DOI: 10.53469/jtpes.2023.03(12).04

# Grasp and Inspection of Mechanical Parts based on Visual Image Recognition Technology

**Yuxiang Liu[1,*], Shiheng Duan[2], Zepeng Shen[3], Zheng He[4],Linxiao Li[5]**

[1]Computer Engineering, university of California, Davis, California, USA
[2]Computer Engineering, Northwestern University, Atlanta, Georgia, USA
[3]Network Engineering, Shaanxi University of Technology, Shaanxi 723001, China
[4]Applied Analytics, Columbia University, NY, USA
[5]Communication Engineering,Peking University, Beijing, China

**Abstract:** ***In the process of industrial production, whether the machine can work stably is an important issue related to production efficiency and production safety. The cleaning of mechanical parts is an important factor affecting the stable operation of mechanical equipment, so the cleaning of mechanical parts is one of the important manufacturing links before assembly and manufacturing. The important value of the application of image recognition technology in the quality inspection of mechanical parts is briefly analyzed, and then the common problems in the quality inspection of mechanical parts are studied, the relevant measures of the application of image recognition technology in the quality inspection of mechanical parts are discussed, and finally the case analysis of the quality inspection of mechanical parts is carried out. The main purpose is to rationally use image recognition technology to complete the quality inspection of mechanical parts, so as to improve the overall level of production automation and give full play to the maximum role of this technology. At present, a variety of mechanical parts classification methods based on deep learning are studied. According to the demand of industrial machine parts, the corresponding machine parts data set is created. According to the production requirements, the SPP structure in the traditional YOLOV5 network is improved, and a new SPC structure is proposed: The YOLOv5 network with SPC structure is characterized by no pooling layer, which improves the accuracy and recall rate compared with the traditional YOLOv5 network. Secondly, the control method of robot grasping mechanical parts is studied. Using the camera to capture the image and the upper computer to calculate the mechanical parts under the camera coordinate system; Through camera calibration and hand-eye calibration, the transformation relationship between the grasping coordinates of mechanical parts in the camera coordinate system and the grasping coordinates of mechanical parts in the robot base coordinate system is obtained. The grasping coordinates and part type information of the mechanical parts in the robot base coordinate system are transmitted from the host computer to the robot through Modbus/TCP communication protocol.***

**Keywords:** Mechanical parts; Image capture; Quality inspection; Visual technique

## 1. INTRODUCTION

With the rapid development of computers and the Internet, emerging technologies such as deep learning and image analysis have emerged, bringing a new round of industrial revolution. In order to better grasp the new round of industrial revolution, countries in the world have put forward their own "fourth industrial Revolution" strategic deployment. The United States put forward the National Advanced Manufacturing Strategic Plan in 2012; In 2013, Germany and France put forward the "German Industry 4.0 Strategy" and the "New Industrial France Plan". In 2014, South Korea proposed an implementation plan for Future Growth drivers. After many countries put forward their own "strategic deployment of the fourth industrial revolution, China also followed suit and put forward the" Made in China 2025 "in 2015. The development of computer vision has led to a new round of industrial change. It is now possible to install cameras for industrial robots, making industrial robots more intelligent, and getting rid of the problem that traditional robots can only perform fixed procedures and move in too fixed a position. Compared with the traditional industrial robot system, the industrial robot system with vision system can identify the types of mechanical parts. By identifying the types of mechanical parts, the classification task of various mechanical parts can be better completed, and the position of these mechanical parts can be located[1]. The coordinate information of the mechanical parts is transmitted to the robot through a certain communication protocol, so that the robot can automatically grasp these mechanical parts.

At present, image recognition technology has been well applied in different industries, but the production of mechanical parts is prone to errors and surface damage related situations, mechanical parts quality inspection work needs to invest a lot of human resources, material resources, time and energy, money, etc[2]. Even so, the probability of leakage and false detection is also high. Therefore, it is recommended to apply image recognition technology in the quality inspection of mechanical parts to ensure that the quality of mechanical parts meets the actual needs.

## 2. RELATED WORK

There are many ways to grab mechanical parts. Buchholz et al. 21 divided the grasping of mechanical parts into two parts, namely, the grasping pose solution and the grasping pose. This obviously increases the calculation time. LeBI uses the deep learning method based on instance segmentation to identify the pose of objects. Drost et al[3]. use four-dimensional vectors as the features of point pairs in the point cloud, and build a hash table according to the feature vectors of each group of point pairs. Finally, according to the hash table, a rotation matrix R and a translation matrix T are obtained. Che et al. transformed raw sensor data (accelerometers, gyroscopes, magnetometers) into precise position predictions through meticulous preprocessing, extracting crucial parameters like absolute velocity and accelerations along different axes. Employing a 2D Convolutional Neural Network (2D-CNN), the study utilizes deep learning to model predictions, aided by an Asymmetric Gaussian loss function tailored for real-world sensor data. This interdisciplinary study merges mechanical engineering and computer science to enhance logistics automation[4]. Cui et al. 's network classification of mechanical parts for Drost et al.' s scheme does not have the function of target detection itself, and is not suitable for mechanical parts with complex shapes and small volumes[5]. Iriondo et al. proposed a scheme for picking random mechanical parts using deep learning. By obtaining high-precision two-dimensional and three-dimensional images, deep learning is used to analyze the material, shape, color and texture of mechanical parts, and finally the mechanical parts are classified, but the calculation speed is slow and not suitable for industrial production. Che et al. proposed a deep learning model that can identify point clouds directly, which splits a whole point cloud into up to four point clouds, and trains the point clouds of all the whole point clouds in the training model, which holds significant implications for our research[6]. By analyzing which point clouds are included in the overall point cloud, we can determine what object the shape of the point cloud is, but usually the objects to be detected are for larger objects, and are not suitable for the identification of smaller mechanical parts.

One of the application directions of object detection problem industry is to obtain the type information of mechanical parts. In the initial development stage of object detection, the method of object detection used by people is sliding window detection. The sliding window detection method needs to know the size and color of the object to be detected in the image in advance. Then according to the set detection step, the pixels in the detection image are compared and matched pixel by pixel. Sliding window detection method is the most original detection method. Its implementation means is simple, but this method not only consumes a lot of computing resources, the recognition rate is also very low[7-8]. As convolutional neural networks come into people's view. People begin to combine convolutional neural networks with digital image processing and apply these two technologies to object detection. During this time AlexNetl was proposed, AlexNet first applied an unsaturated activation function, the ReLU activation function, in convolutional neural networks, and used the Dropout mechanism to prevent overfitting, while AlexNet first applied Gpus to neural network calculations. The algorithm applied by deep learning to target detection is Region-CNNI9 (R-CNN for short) proposed in 2014. Unlike traditional object detection, R-CNN uses deep convolutional networks to acquire image features. R-CNN uses the deep convolutional network to select some candidate regions in advance in the image to be detected as candidate regions for target detection, instead of traversing all pixels in the image for target detection like the sliding window detection method, which greatly improves detection efficiency. On the basis of the R-CNN also developed in addition to better performance of the Fast R-CNN! And Faster RCNN[9-11].

## 3. THE IMPORTANCE OF IMAGE RECOGNITION TECHNOLOGY

### 3.1 Improve detection accuracy and efficiency

In the traditional assembly process, it is necessary to use measuring tools and manual measurements to check whether the accuracy of each part of the spindle meets the requirements. This method is time-consuming and laborious, and there are some errors in the accuracy of manual measurement. After the introduction of computer vision detection technology, the image of the spindle component can be obtained through the camera and transmitted to the computer for processing. Therefore, the use of image recognition technology in the process of system detection has a good effect, mainly reflected in the role of technology to obtain image features and timely processing of images, not only that, image recognition also has input function to effectively avoid detection errors, which is conducive to improving the accuracy of information processing. Visual inspection technology refers to the use of computer and digital image processing technology, the simulation and expansion of human vision system, to achieve the automatic understanding and analysis of image and video data, image recognition technology can be based on information technology for mechanical parts quality inspection, compared with traditional inspection technology[12].

The most commonly used algorithm of this kind of technology is the Haar-like feature extraction form, which is a classic image feature extraction algorithm. It was first used for face description by Papage origiou. The application of this technology has outstanding advantages in saving time, controlling cost and simplifying identification process, and can obtain accurate detection results during actual identification, which is helpful to improve the overall quality of mechanical parts quality inspection.

Adaboost algorithm based on Haar-like features can effectively identify objects in the image and determine whether there is a vehicle. First, the feature is extracted by training samples, then the eigenvalues are calculated and the weak classifier is built, and finally it is constructed into a strong classifier by weighted voting.

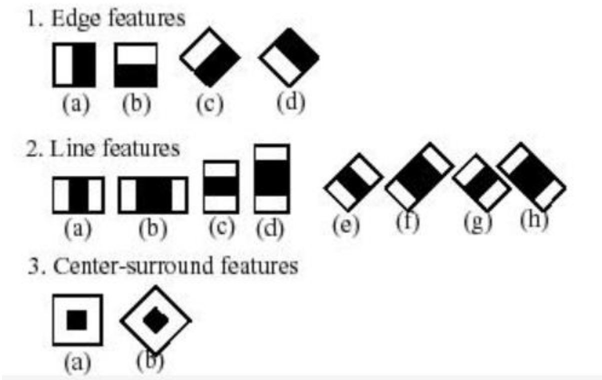


**Figure 1:** Haar-like features

The size of each feature box is not fixed, but the proportion of the black and white matrix is always fixed. The difference of the sum of the gray values of all pixels in the white area and the black area is taken as the feature value, and the sum of pixels in the area can be quickly obtained through the integral graph. To improve the detection rate,

**3.2 Reduce quality inspection costs**

Computer vision inspection technology can be applied to the detection of surface defects of mechanical parts. For example, in the manufacturing process, the surface of the part may appear cracks, wear, scratches and other defects, and these defects will have a great impact on the service life and performance of the part. Using computer vision technology, the surface of parts can be scanned and analyzed with high precision, the location and size of surface defects can be accurately detected, and the corresponding report can be generated. This can help manufacturers find and repair surface defects of parts in a timely manner, improve product quality and user satisfaction. Therefore, under the premise of achieving high-quality detection, detection costs can be reduced through computer recognition technology, in fact, to improve the accuracy of detection, production enterprises need to hire professional and technical personnel to carry out related work, and reasonable use of image recognition technology processing, the purpose is to reduce labor costs and detection costs, effectively alleviate financial pressure.

**3.3 Accelerate image analysis and capture**

Visual inspection technology can be used to classify and rank defects of mechanical parts. During the production process, a variety of defects may occur in mechanical parts, which may have different effects on the service life and performance of the parts. Through computer vision technology, different defects can be classified and sorted and dealt with accordingly. For example, defects can be marked and numbered to facilitate unified management and treatment by staff. Therefore, in the process of quality detection and image capture, binary image processing and computer processing are implemented in image analysis, and the main features of images should be deeply analyzed[13]. The binary image can be processed effectively by analyzing the image carefully after separating the image. There are many methods for image separation analysis, such as indirect method, direct method and multi-threshold method, among which the application of multi-threshold method can be combined with gray level and target region to segment the image. The features of image edge detection are gray scale, texture, line, amplitude, transformation coefficient, etc.

The commonly used algorithm here is HOG feature algorithm. HOG feature and SVM algorithm are a feature based on image gradient change and a machine learning algorithm based on binary classification. HOG feature can effectively describe the shape and edge information of the object to be detected, and SVM algorithm can classify and detect the object. The algorithm has good robustness and accuracy, but its performance is not good enough when dealing with large-scale changes and occlusion.

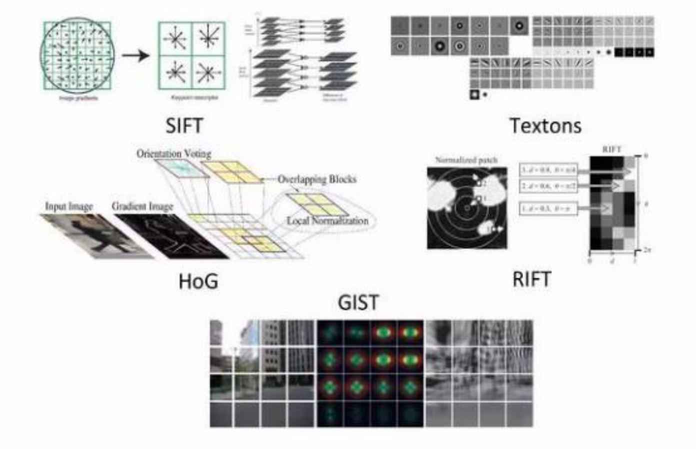


**Figure 2:** Computer Vision features--HOG

Select the characteristic parameters reasonably when identifying the image of the mechanical part, and then identify the common quality problems of the mechanical part, such as irregular defects, breaks, surface/internal cracks and other problems. When there are quality problems in the actual quality inspection of mechanical parts, the defects or quality problems can be found in time through the image recognition of the equipment. If cracks occur in the process of parts production, it is recommended to use image recognition technology to process and analyze images to obtain relevant data information and integrate information

### 3.4 YOLO features

Another feature extraction method for quality detection algorithm is (You Only Look Once) YOLO algorithm, YOLO algorithm is a real-time object detection algorithm based on deep learning. The algorithm uses convolutional neural network to detect and classify the detected images from end to end. YOLO algorithm has the characteristics of fast detection speed, high accuracy and good robustness, and has excellent performance in practical application. However, the algorithm is not effective for the detection of small objects. Therefore, this kind of algorithm is rarely used in the grasping and quality inspection of most mechanical parts.

Table 1 below is the analysis of the test results of the test set of the three algorithms in the mechanical parts testing experiment.

**Table 1:** Test results of the test set in the mechanical part inspection experiment

| *Algorithm* | *Detection rate* | *False drop rate* |
|---|---|---|
| Haar-like | 90.5% | 2.5% |
| HOG | 93.2% | 0.75 |

| | | |
|---|---|---|
| YOLO | 97.3% | 1.8% |

In addition, the running time of the three algorithms is compared. The experimental results show that the running time of Haar-like feature +AdaBoost classifier and HOG feature +SVM classifier is shorter. The YOLO algorithm has a longer running time, but it also has high real-time performance in practical applications. To sum up, although YOLO algorithm performs best in mechanical parts inspection, it is also necessary to consider the needs and scenarios of practical applications and select the most appropriate algorithm for application.

## 4. PROBLEMS AND SOLUTION

### 4.1 Problems

In the inspection of mechanical parts, there are some problems to be solved. First of all, there are many kinds of mechanical parts, and the shape, material, color and other characteristics of each part are not the same, so they need to be targeted at different types different detection algorithms are used for parts. Secondly, the accuracy and robustness of the algorithm will be affected by the possible noise and occlusion in the image. Finally, traditional algorithms often require manual characterization

The amount of work and the effect of classifier training is not stable.

(1) The fuzzy outline image of mechanical parts can not guarantee the accuracy of the detection result, which is closely related to the low resolution of the camera and insufficient illumination, and the illumination environment is easily affected by the radiating error boundary factor of the part.

(2) Because the deviation of visual image capture detection position is easy to cause curve error and system deviation, the two curve center points should be coincident together to ensure the same Angle when positioning the contour curve.

(3) The resolution of the shooting picture is very high, so the image should be clear and the detection position is accurate. In addition to ensuring that the related hardware meets the image resolution requirements, image interpolation should also be used to improve the image resolution, so as to meet the actual requirements.

(4) The recognition stress of the contour curve of the part should be accurate, and the curve interference should be eliminated in time to ensure the accurate position of the curve, and the algorithm can be improved by boundary detection method. The contour is the representative closed or unclosed geometric shape of the generated entity, the feature modeling plane graphics, which is convenient for users to draw a certain reference plan, and the contour curve is closely related to the current feature creation.

### 4.2 Solution

In order to solve the above problems, the following improvement schemes are proposed

(1) Deep learning algorithm is introduced. Deep learning algorithms can automatically extract feature and classification information by learning a large amount of data, avoiding the tedious process of traditional algorithms requiring manual feature design and classifier training. For example, deep learning algorithms such as convolutional neural network (CNN) and recurrent neural network (RNN) can be used to detect mechanical parts.

(2) Engaging data enhancement technology. Data enhancement technology can generate more diverse data sets by rotating, sub shifting and scaling the original data, so as to improve the robustness and generalization ability of the algorithm. For example, image enhancement technology can be used to process images of mechanical parts to reduce noise and distortion due to factors such as light, Angle, etc.

(3) Introduce multi-modal information. Multi-modal information can detect mechanical parts from different angles by collecting image, sound, vibration and other data at the same time, improving detection accuracy and robustness. For example, sound signal analysis technology and vibration signal analysis technology can be used

to carry out multi-directional detection of mechanical parts.

## 5. CONCLUSION

This paper introduces the commonly used mechanical parts inspection algorithms from two aspects of image processing and deep learning, and carries out experimental verification. The experimental results show that compared with the traditional algorithm, the deep learning algorithm is better in mechanical parts inspection.

In the test, it has higher accuracy and robustness, especially better performance in complex cases. The comparison between the deep learning algorithm in the improved scheme and the data enhancement technology is also analyzed. It can be seen that compared with the traditional algorithm, the deep learning algorithm has higher accuracy, especially better performance in complex cases such as occlusion, uneven illumination and image noise. Therefore, in the daily quality inspection and realization of mechanical parts, advanced technologies such as deep learning and AI can be combined. Based on features such as Haar-like, HOG, YOLO, etc., a new feature set and its feature calculation formula are proposed, and a better classifier can be trained by using new features to improve the detection accuracy[14-15]. The Adaboost algorithm can also be used to train different vehicle attitude corresponding classifiers, combine the detection results of each classifier, and realize the multi-channel cascade classifier. It can improve the accuracy of the weak classifier in the training process of Adaboost, especially in the detection of vehicles with various states in the same image[16-18]. Compared with the original Haar-like feature classifier, the detection rate of the trained strong classifier is obviously improved, and it can also meet the requirements of real-time detection.

## ACKNOWLEDGMENT

Image recognition technology holds crucial application value in the field of mechanical industry, such as in the robotic arms on automobile production lines. We appreciate the specialized knowledge provided by Bo Liu and Chang Che in mechanical and computer vision-related areas. Their research in the detection of defects in industrial castings and vehicles' components using artificial intelligence provides us with significant guidance in this field.

## REFERENCES

[1] JIN L S, JIA M, SUN Y O, et al. Detection of backside vehicle on freeway in daytime [i] . Journal of Southwest Jiao tong University ,2010 ,45(2): 231 – 237.
[2] FLETCHER L, PETERSSON L, ZELINSKY A. Driver assistance systems based on vision in and out of vehicles [C] / /Proceedings of Intelligent Vehicles Symposium. IEEE ,2003.
[3] VIOLA P, JONES M. Robust real-time face detection m. International Journal of Computer Vision ,2004 ,57(2): 137 - 154.
[4] Chang Che, Bo Liu, Shulin Li, Jiaxin Huang, and Hao Hu. Deep learning for precise robot position prediction in logistics. Journal of Theory and Practice of Engineering Science, 3(10):36–41, 2023.DOI: 10.1021/acs.jctc.3c00031.
[5] Hao Hu, Shulin Li, Jiaxin Huang, Bo Liu, and Change Che. Casting product image data for quality inspection with xception and data augmentation. Journal of Theory and Practice of Engineering Science, 3(10):42–46, 2023. https://doi.org/10.53469/jtpes.2023.03(10).06
[6] Chang Che, Qunwei Lin, Xinyu Zhao, Jiaxin Huang, and Liqiang Yu. 2023. Enhancing Multimodal Understanding with CLIP-Based Image-to-Text Transformation. In Proceedings of the 2023 6th International Conference on Big Data Technologies (ICBDT '23). Association for Computing Machinery, New York, NY, USA, 414–418. https://doi.org/10.1145/3627377.3627442
[7] Lin, Q., Che, C., Hu, H., Zhao, X., & Li, S. (2023). A Comprehensive Study on Early Alzheimer's Disease Detection through Advanced Machine Learning Techniques on MRI Data. Academic Journal of Science and Technology, 8(1), 281–285.DOI: 10.1111/jgs.18617
[8] Che, C., Hu, H., Zhao, X., Li, S., & Lin, Q. (2023). Advancing Cancer Document Classification with R andom Forest. Academic Journal of Science and Technology, 8(1), 278–280. https://doi.org/10.54097/ajst.v8i1.14333
[9] S. Tianbo, H. Weijun, C. Jiangfeng, L. Weijia, Y. Quan and H. Kun, "Bio-inspired Swarm Intelligence: a Flocking Project With Group Object Recognition," *2023 3rd International Conference on Consumer*

*Electronics and Computer Engineering (ICCECE)*, Guangzhou, China, 2023, pp. 834-837, DOI: 10.1109/ICCECE58074.2023.10135464.

[10] Y. Wang, K. Yang, W. Wan, Y. Zhang and Q. Liu, "Energy-Efficient Data and Energy Integrated Management Strategy for IoT Devices Based on RF Energy Harvesting," in IEEE Internet of Things Journal, vol. 8, no. 17, pp. 13640-13651, 1 Sept.1, 2021, DOI: 10.1109/JIOT.2021.3068040.

[11] Wang, Y., Yang, K., Wan, W., & Mei, H. (2020). Adaptive energy saving algorithms for Internet of Things devices integrating end and edge strategies. *Transactions on Emerging Telecommunications Technologies, 32*.DOI: https://doi.org/10.1002/ett.4122

[12] Xu, J., Pan, L., Zeng, Q., Sun, W., & Wan, W. Based on TPUGRAPHS Predicting Model Runtimes Using Graph Neural Networks. https://api.semanticscholar.org/Corpus

[13] IZADINIA H, RAMAKRISHNA V, KITANI KM, et al. Multi-pose multi-target tracking for activity understanding [C] / /Proceedings of Applications of Computer Vision ( WACV) . Tampa: IEEE 2013

[14] VIOLA P, JONES M . Rapid object detection using a boosted cascade of simple features [C] / /Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and PatternRecognition.CVPR,2001.

[15] Che, C., Yang, C., Liu, T., & Danlin, L. (2023). THE APPLICATION OF COMPUTER VISION IN THE ANALYSIS OF VEHICLE COMPONENT. In *СТУДЕНТ И НАУКА: АКТУАЛЬНЫЕ ВОПРОСЫ СОВРЕМЕННЫХ ИССЛЕДОВАНИЙ* (pp. 12-16).

[16] Yao, J., Zou, Y., Du, S., Wu, H., & Yuan, B. (2023). Progress in the Application of Artificial Intelligence in Ultrasound Diagnosis of Breast Cancer. Frontiers in Computing and Intelligent Systems, 6(1), 56-59.

[17] Zhou, H., Lou, Y., Xiong, J., Wang, Y., & Liu, Y. (2023). Improvement of Deep Learning Model for Gastrointestinal Tract Segmentation Surgery. Frontiers in Computing and Intelligent Systems, 6(1), 103-106.

[18] Liu, B., Yu, L., Che, C., Lin, Q., Hu, H., & Zhao, X. (2023). Integration and Performance Analysis of Artificial Intelligence and Computer Vision Based on Deep Learning Algorithms. arXiv preprint arXiv:2312.12872.