Vision Recognition and Grasping System of Industrial Robot Based on Deep Learning

Yifei Liu

Yanshan University, Qinhuangdao, 066000, Hebei, China

Keywords: Deep learning; Industrial robots; Visual recognition; Grab strategy; Smart manufacturing

Abstract: The purpose of this article is to research and develop a visual recognition and grasping system for industrial robots based on deep learning, so as to improve the recognition accuracy and grasping success rate of industrial robots in complex working environments. Aiming at the limitations of traditional methods in dealing with complex background and illumination changes, this article proposes an innovative solution combining deep learning, machine vision and industrial robot technology. By constructing an efficient deep learning model and optimizing the network structure and hyperparameters, the high-precision recognition of the target object is realized. At the same time, this article introduces the intelligent grasping strategy planning module, and uses reinforcement learning technology to make the robot adjust the grasping mode and path adaptively, which improves the flexibility and success rate of grasping. The experimental results show that the system can maintain high recognition accuracy and success rate in different lighting conditions, object occlusion and complex background. The system has obvious advantages in improving the intelligent level of industrial robots and promoting the development of intelligent manufacturing.

1. Introduction

In today's increasingly prosperous intelligent manufacturing, industrial robots, as the core components of automated production lines, are gradually changing from traditional single and repetitive task executors to multifunctional assistants with high flexibility and intelligent decision-making ability [1]. One of the driving forces of this transformation is the rapid development of machine vision technology, especially the introduction of deep learning technology [2]. Through the image recognition ability of deep learning, industrial robots can understand the working environment more accurately, realize the automatic detection, positioning and grabbing of objects, and greatly improve the production efficiency and flexibility [3]. However, in the face of complex and changeable production scenes, how to build an efficient and stable visual recognition and grasping system is still an important topic of current research [4]. The purpose of this study is to explore the application of deep learning in the field of industrial robot visual recognition and grasping, to provide new ideas and technical paths to solve this challenge, and to promote the intelligent upgrading of manufacturing industry.

In recent years, scholars have made remarkable progress in industrial robot vision recognition and grasping technology [5]. Traditional methods mostly rely on hand-designed feature extraction and machine learning algorithms, but their generalization ability in complex scenes is limited [6]. With the rise of deep learning, especially the successful application of Convolutional Neural Network (CNN) in image classification, target detection and other fields, it has brought a revolutionary breakthrough for industrial robot vision recognition [7]. Nevertheless, how to maintain stable recognition and grasping performance in complex lighting conditions, object occlusion and high-speed moving environment is still the main challenge facing the current technology.

In view of the above background, the main goal of this study is to design and implement an industrial robot vision recognition and grasping system based on deep learning. The system can realize rapid and accurate identification and stable grasping of various objects in complex and changeable industrial production environment.

  DOI: 10.25236/iwmecs.2024.021

2. Related theoretical basis

Deep learning is a branch of artificial intelligence, which simulates the learning mechanism of human brain by constructing multi-layer neural network model, and realizes the high-level abstraction of data and automatic learning of complex patterns [8]. Its core lies in using a large number of labeled data to optimize network parameters through back propagation algorithm, so that the model can accurately complete tasks such as classification, regression and generation. In the field of image recognition, CNN has become the preferred model for processing two-dimensional image data because of its characteristics of local connection, weight sharing and pooling.

Machine vision can extract the shape, size, position and other information of objects mainly through image acquisition, processing, analysis and understanding, and then guide the robot to perform corresponding operations [9]. Traditional machine vision technology relies on artificially designed feature extraction and classic machine learning methods to perform well in specific scenes, but its generalization ability is limited. With the development of deep learning, machine vision technology has entered a new stage of intelligent perception. Deep learning can automatically learn the high-level features in images, which significantly improves the accuracy and robustness of recognition, making machine vision technology show great application potential in industrial inspection, robot navigation, autonomous driving and other fields.

Industrial robot technology covers many aspects such as mechanical structure, motion control, sensor technology, programming and simulation. Through accurate motion control and flexible operation ability, industrial robots can complete various tasks such as handling, assembly, welding and spraying, which significantly improves production efficiency and product quality. By integrating advanced technologies such as deep learning into the robot vision system, the adaptability and autonomous decision-making ability of robots to complex working environments can be significantly improved, and robots can perform various fine operations more intelligently, which further promotes the wide application of industrial robot technology in the field of intelligent manufacturing.

3. Design of industrial robot vision recognition and grasping system based on deep learning

3.1. Overall system architecture

This system takes deep learning as the core, and designs a set of complete industrial robot vision recognition and grasping solutions. The overall architecture of the system is divided into three main parts: hardware layer, software layer and application layer. The hardware layer includes industrial robot body, high-precision camera, light source and necessary sensors, which together form the physical foundation of the system and are responsible for image acquisition and preliminary data processing. In the software layer, a visual recognition module and an intelligent grasping strategy planning module based on deep learning are constructed, and fast image processing and accurate recognition, as well as intelligent grasping path planning are realized through efficient algorithms. The application layer is the interface between the system and users, providing an intuitive operation interface and rich functional options, which is convenient for users to configure the system and schedule tasks according to actual needs. The three layers of architecture are closely connected by high-speed communication protocol, which ensures the real-time transmission and processing of data and realizes the seamless docking from image acquisition to capture execution.

3.2. Visual recognition module

The visual recognition module is the "eye" of the whole system, which is responsible for accurately identifying the target object from the complex background. The module adopts CNN technology in deep learning, and through training a large number of labeled image data, the network can automatically learn and extract key features of objects. In order to improve the accuracy and robustness of recognition, this article adopts the strategy of multi-scale feature fusion, combining global and local information, and effectively solves the problems of object occlusion and illumination change. The algorithm formula is expressed as:

$$F_{\text{mix together}} = \phi\left(\left\{f_{\text{overall situation}}(I), f_{\text{part}}(I)\right\}\right) \tag{1}$$

Where $F_{\text{mix together}}$ is the feature representation after multi-scale feature fusion. ϕ is a fusion function representing features, which is a fusion layer in the deep learning network. $f_{\text{overall situation}}(I)$ is the feature extracted from the global perspective. $f_{\text{part}}(I)$ is the feature extracted from the local perspective. I is the input image.

The module also integrates the target detection algorithm to realize real-time positioning and category judgment of objects. By continuously optimizing the network structure and super-parameters, the visual recognition module can quickly and accurately identify the target object in the complex industrial environment, which provides reliable information support for the subsequent grabbing operation.

3.3. Grab strategy planning

Grasping strategy planning module is the key bridge between visual recognition and robot execution. After obtaining the position and category information of the target object, the module first selects the most suitable grab method from the predefined grab strategy library according to the shape, size and material characteristics of the object. In order to improve the flexibility and success rate of grasping, the reinforcement learning technology in deep learning is introduced. By simulating the grasping experiment, the robot can constantly try and adjust the grasping strategy, so as to learn the optimal grasping action in different situations. The module also considers collision avoidance and force control in the grasping process, and ensures the stability and safety of the grasping process through accurate path planning and force feedback mechanism.

3.4. System integration and testing

After completing the design and development of each module, this section carries out the integration and testing of the system. First of all, through accurate mechanical design and electrical connection, the camera, robot, sensor and other hardware devices are closely integrated to ensure accurate data acquisition and accurate execution of instructions. Then, on the software level, this article develops a unified control interface and communication protocol, which realizes the seamless cooperation between modules. In order to verify the overall performance of the system, a series of test experiments are designed in this section, including recognition test under different lighting conditions, grasping test of objects with various shapes, and continuous operation stability test. The test results are shown in Table 1:

Table 1 System Integration and Test Results

Test Item	Test Conditions	Test Results
Recognition Test	Strong Light Conditions	Recognition Accuracy: 98.2%
	Low Light Conditions	Recognition Accuracy: 96.7%
	Complex Background	Recognition Accuracy: 95.3%
Grasping Test	Regular-shaped Objects	Grasping Success Rate: 99.1%
	Irregular-shaped Objects	Grasping Success Rate: 93.5%
	Fragile Objects	Grasping Success Rate: 90.8%, No Damage
Stability Test	Continuous Operation for 24 Hours	System Stable, No Failures, Performance Fluctuation < 2%

The test results show that the system can work stably in the complex industrial environment, realize the rapid identification and accurate grasping of the target object, and verify the effectiveness and practicability of the design. Through continuous testing and optimization, the performance and stability of the system are further improved, which lays a solid foundation for its application in actual industrial production.

4. Experimental results and analysis

In order to comprehensively evaluate the performance of industrial robot vision recognition and grasping system based on deep learning, a wealth of data sets were constructed. It includes a large number of labeled images and grabbing cases, which are used to train and test the visual recognition and grabbing strategy planning ability of the system. In the performance test of visual recognition, the recognition accuracy and speed of the system are mainly investigated, and the results are shown in Table 2:

Table 2 Visual Recognition Performance Test Results

Test Conditions	Recognition Accuracy	Recognition Speed (milliseconds)
Strong Light	96.3%	23.1
Low Light	95.7%	25.4
Shadow	95.1%	24.8
Object Obstruction	95.9%	26.2
Complex Background	96.1%	27.3

By testing under different lighting conditions (strong light, weak light, shadow) and object occlusion, it is found that the system can accurately identify the target object, and the recognition accuracy rate is over 95%, and the recognition speed is in the millisecond level, which meets the real-time requirements. Especially when dealing with complex background and object occlusion, the system shows strong feature extraction and classification ability, which benefits from the multi-level feature fusion and context information utilization of deep learning model.

The success rate of grabbing is an important index to measure the practicability of the system. Various grabbing scenes are designed, including single grabbing, continuous grabbing, and grabbing tasks with different objects stacked or closely arranged. The results are shown in Table 3:

Table 3 Grasping Performance Test Results

Grasping Scenario	Grasping Success Rate	Remarks
Single Grasping	92.4%	Stable grasping of objects of various shapes and sizes
Continuous Grasping	90.8%	Maintains stable performance in continuous tasks
Object Stacking	89.7%	Accurately identifies and grasps the target in stacked objects
Tightly Arranged	91.3%	Avoids collisions and mis-grasping when objects are tightly arranged
Irregular Shapes	88.9%	Precise control of grasping force and path to avoid damage
Fragile Objects	90.2%	Gentle grasping to avoid breaking or damage

Through experiments, it is found that the system can intelligently plan the grasping path and grasping strength according to the shape, size and position of the object, and realize the stable grasping of the target object. In a single grab task, the success rate of the system reached more than 90%, and it maintained stable performance in continuous grab tasks. Especially when dealing with irregular shapes and fragile objects, the system shows fine force control and path planning ability, and avoids collision and damage during grasping. These experimental results fully verify the effectiveness and reliability of the system grasping strategy planning.

5. Conclusions

In this study, a set of industrial robot vision recognition and grasping system based on deep learning is successfully designed and implemented. Through the comprehensive application of deep learning, machine vision and industrial robot technology, it significantly improves the recognition accuracy and grasping success rate of industrial robots in complex working environments. The research results are mainly reflected in the following aspects: ① An efficient deep learning model is constructed in this article. By optimizing the network structure and hyperparameters, the target object is recognized with high accuracy, and its performance can be maintained even under complex background and illumination changes. (2) In this article, the intelligent grasping strategy planning module is innovatively introduced, combined with reinforcement learning technology in deep learning, so that the robot can adaptively adjust the grasping mode and path according to the characteristics of different objects, which greatly improves the flexibility and success rate of grasping. ③ Through system integration and testing, this article verifies the overall performance and stability of the system. The system is highly scalable and flexible, and can be customized and optimized according to different application scenarios and requirements, which provides a new solution for the intelligent upgrade of industrial robots. The results provide strong support for its application in practical industrial production.

References

[1] Chen Wei, Li Zechen, Zhang Jing, et al. Research on the calibration method of industrial robot vision system based on HALCON [J]. Electronic Measurement Technology, 2020,43(21):137-141.

[2] Yan Shuo, Li Jiansong, Tang Changsong. Design of industrial robot visual grasping system based on neural network [J]. Computer Measurement and Control, 2024, 32(4):201-209.

[3] Yang Junli. Research on visual servo control system based on industrial robots [J]. Automation Technology and Application, 2020,39(08):66-69+73.

[4] Zhao Xiaoan. Industrial Robot Vision Algorithm Based on Euclidean Distance Corner Extraction [J]. Automation and Instrument, 2020,35(08):31-36.

[5] Su Jian, Mu Cunqiang, Ren Shanjian, et al. Design of industrial robot visual sorting digital twin system based on NX MCD [J]. Machine Tool and Hydraulic, 2023, 51(23):73-79.

[6] Hua Dan. Design of industrial robot vision detection and obstacle avoidance system based on blockchain technology [J]. Computer Measurement and Control, 2020,28(07):69-73.

[7] Zhang Xiaodong, Su Ruifang, Dong Weiguang. Compressed sensing visual image processing of industrial robots based on adaptive block [J]. Automation and Instrumentation, 2021,36(08):38-42.

[8] Luo Hua, Shang Junyun, Li Ruifeng, et al. A high-precision three-dimensional disordered grasping system for industrial robots [J]. Navigation and Control, 2023, 22(4):106-116.

[9] Xu Zhezhuang, Ping Huang, Chen Dan, et al. Research on identification strategy of similar industrial equipment based on machine vision and proximity estimation [J]. chinese journal of scientific instrument, 2023, 44(1):283-290.