Research on Object Detection of Robotic based on Convolutional Neural Network

Lin Xi*[a], Yuge Chen[b] and Donglai Liu[b]
[a] School of Mechanical Engineering, Dalian University of Technology, Dalian, China;
[b] School of Mechanical Engineering, Dalian University of Technology, Dalian, China.
[*] Corresponding author: 20211044031@mail.dlut.edu.cn

ABSTRACT

With the continuous development of computer vision, objects detection methods are widely devised to achieve the target objects identification for robotics. Indeed, object detection of robotics has become an essential component to recognize the avoid conflict among these objects and robotics. However, existing detection are primarily used deep learning or select the target area to recognize the objects, which ignores the convolutional neural network and leaks investigation for detecting multiple objects for robotics. In this work, we proposed a novel objects detection model of robotics through utilizing the convolutional neural network, which can also utilize to dispose the classification issue. Initially, we capture the robotic input images through robotic vision and subsequently the trained convolutional neural network is utilized to identify the input image objects from split input data. From our extensive experimental results, we can conclude that the proposed model can achieve the objects detection with acceptable identification accuracy and reasonable computation cost.

Keywords: Computer vision, Robotics, Object detection, Convolutional neural network, Computation cost.

1. INTRODUCTION

Robot is an important mechanical equipment of modern industry, with the continuous development and iteration of artificial intelligence technology, intelligent robots came into being, intelligent robots not only play an important role in production or social life, but also an effective method to solve labor shortages [1]. Therefore, improving the controllability of robots so that they can identify multi-target objects in complex object environments is of great significance and promotion value.

Object detection is a critical component of robotics, as it allows robots to identify and interact with objects in their environment. Object detection is the process of identifying and locating objects in the format of a series image or videos. The objects detection is a core component of numerous robotic-based downstream applications including autonomous navigation, object manipulation and object recognition [2]. Above all, the object detection algorithms can be utilized to detect objects in a variety of scenarios including indoor and outdoor environments, and in both static and dynamic scenes. Object detection algorithms can be utilized to detect objects of various sizes, shapes, and colors, and can be used to detect objects in both 2-dimension and 3-dimension space [3]. Furthermore, the object detection algorithms can also be utilized to detect objects in real-time, allowing robots to effectively and accurately identify and interact with objects in their environment.

Object detection is an important research object in computer-aided vision, and the robot target detection task includes the recognition of target location distance and target category attributes, which is also an important component principle of machine vision [4]. With the popularization of computer technology, its technology has been applied in many fields such as remote sensing target detection, pedestrian detection, face recognition, product defect detection and industrial safety detection [5]. Object detection involves knowledge in many fields, including computer science, artificial intelligence and machine learning, etc., is a relatively complex science and technology, using deep learning methods to detect objects in a large number of images, which is also a more common application [6].

Convolutional neural networks are a type of deep learning algorithm that are used for image recognition and classification. Convolutional neural networks are composed of multiple layers of neurons that process and analyze images [7]. Each layer of neurons is connected to the next layer, and the output of each layer is used to generate the output of the next layer. Convolutional neural networks are able to learn complex patterns in images, and can be used to classify objects in images, detect objects in images and generate images from scratch. Indeed, the convolutional neural networks are used in a variety of applications, including computer vision, natural language processing and robotics [8]. Additionally, the convolutional neural networks are also utilized in medical imaging, where the trained models are utilized to detect and classify tumors and other abnormalities.

In the robot system, the object recognition element and the algorithm are important execution devices, and many object recognition figures can be seen in service and collaborative work scenarios. As shown in Figure 1 (a) is an automatic screw tightening robot, in the actual industrial product production process, tightening the screw is a monotonous and repetitive task, there are still many factories are workers holding a tightening gun to tighten the screw, the robot through the identification system to identify the specific position of the screw and through the equipped tightening module can tighten the screw in the specified position [9]. Figure 1 (b) is the express sorting robot, under the guidance of computer vision, the robot follows the express on the conveyor belt to pick, and then places it in the target area, its advantage is that the vision system can carry out positioning compensation, online follow and grab the object [10]. Therefore, the development of robot technology from simple to complex, has higher demand and research significance for the accuracy and speed of object recognition.

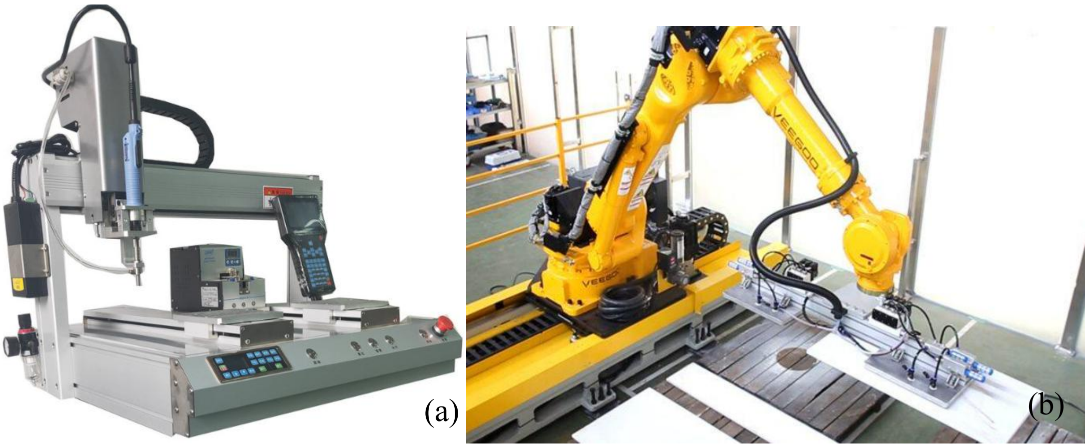

Figure 1. Existing robotic machine with object detection components.

Indeed, the research on robot target recognition system based on convolutional neural network has important research significance and value. This paper aims at the detection and recognition task of robot input vision objects, so that robot vision can accurately identify target objects, and can also apply this design to many other assembly line production systems to help solve monotonic and repetitive identification work and improve the detection accuracy of industrial products. In addition, the emergence of such new technologies and applications can further promote the further improvement of robot vision, promote the development of robot technology and the expansion of social and economic benefits, and have a wide range of prospects for large-scale robot practice in the future.

The reminder of this paper will be arranged as an introduction about related objects detection of robotics and primary parameters of proposed model is shown in Section 2. Subsequently, Section 3 will provide the general procedures and demonstrate the specific framework of neural cells and convolution layers. Indeed, the experimental results and comparison analysis demonstrates in the Section 4. Finally, we conclude the whole paper and illustrate several possible methods for future improvements.

2. PRELIMINARIES

2.1 Related works

Initially, vision-based robot target recognition requires a combination of multi-faceted algorithms and actual physical hardware facilities. Researcher Zhang used deep learning technology to obtain a grasping and positioning model combining the robot end fixture and the core point mapping of the target object in the captured image, and applied it to the robotic arm grasping system to achieve automatically identify target objects and capture the targets [11].

Indeed, with the deeper investigation of deep learning algorithm development, researcher Yun proposed a multi-modal network model based on deep learning, which generates a new feature sharing layer through red-green-blue (RGB) and deep image feature extraction and fusion, and then obtains the objects identification results through the iterations of learning training [12].

2.2 Primary parameters description

Subsequently, the following Table 1 describes the specific parameter symbols and its corresponding explanations for the proposed model.

Table 1. Proposed model primary parameter explanations.

Parameters	Explanations
M	Number of input images
P	Pixel matrix of input
I	Number of row pixels
C	Convolution operation results
conv	Convolution product

3. MODEL FRAMEWORK

In this section, we initially illustrate the main procedures of proposed model and describe the specific functions of these procedures. Subsequently, we illustrate the general model framework with detail execution sequences and components of the model.

3.1 Primary procedures illustration

Initially, following items show the detail procedures of proposed model and demonstrate the execution explanation for the proposed method.

- Collect the data-set of images that contain the object to be detected. The images should be of high quality and contain the object in various orientations and sizes.
- Pro-process the images aims to ensure that the input images or videos are suitable for training a convolutional neural network with certain sizes and shapes. The module includes re-sizing the images, normalizing the pixel values and converting the images to a suitable format.
- Design and establish a convolutional neural network architecture that is aimed to accomplish the muti-objects detection. This includes selecting the number of layers, the type of layers and the hyper-parameters inner the convolution operation.
- Train the convolutional neural network on the data-set of images. This includes selecting an appropriate optimizer, setting the learning rate and training for a suitable number of epochs.
- Test the convolutional neural network on a separate data-set of images to evaluate its performance. This includes measuring the accuracy, precision and recall of the model.
- Deploy the trained convolutional neural network on the robotic platform. This includes converting the model to a suitable format and loading it onto the robotic platform.
- Through utilizing the convolutional neural network to detect the object in real-time on the robotic platform. The module includes using the model to detect the object in each frame of the video stream and providing feedback to the robotic platform.

3.2 Convolution neural network framework

Subsequently, we demonstrate the general framework of proposed model in following Figure 2 with the robotic input components and the object feature extraction and identification process components. Finally, after the full-connected layer, the detected objects are labeled.

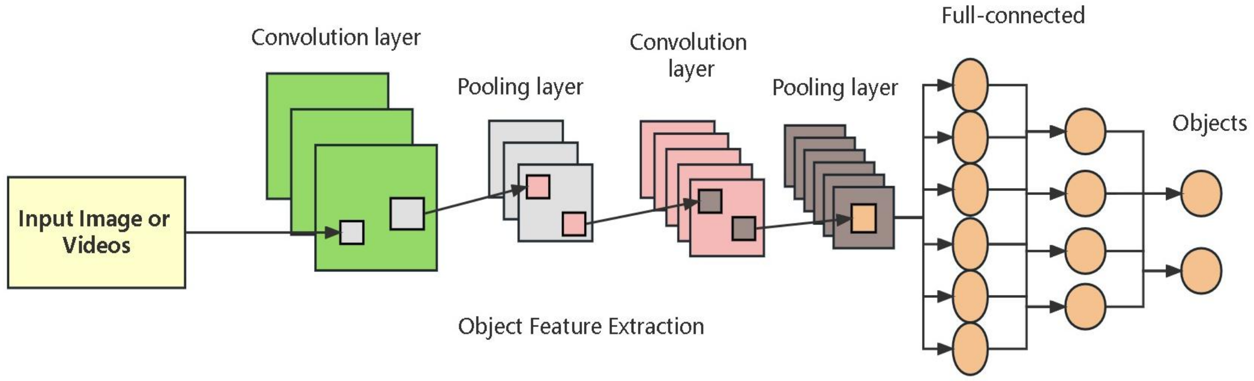

Figure 2. Convolutional neural network general structure demonstration.

Following Equation 1 describes the convolution operation with the calculation process, where the symbol p represents the pixel value of input images and the conv is the convolution product of these two values

$$c = \mathrm{conv}(p_1, p_2) \tag{1}$$

4. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we initially introduce the used data-set and the related algorithms to accomplish the objects detection tasks of robotics. Subsequently, we demonstrate the simulation results and analyze the comparison results through simulating the identical environment with related detection algorithms.

4.1 Data-set and related algorithm description

In experimental simulation procedure, we utilize the existing data-set VOC2012, which is a large-scale image data-set for object detection and segmentation. It contains 20,000 images from the PASCAL VOC 2012 challenge, which is a popular benchmark for object detection and segmentation.

The data-set includes images from various categories such as people, animals, vehicles, and indoor scenes. The images are annotated with bounding boxes and segmentation masks for each object in the image. The data-set also includes a set of annotations for each image, such as object class, object size, and object location. The used data-set is widely used for training and evaluating object detection and segmentation algorithms.

Additionally, we utilize existing identification method deep learning (DP) mentioned in previous related works and YOLOv5 [13] methods to compare with our proposed model. YOLOv5 is a state-of-the-art object detection system developed by the YOLO team.

It is based on the YOLOv4 architecture and is designed to be fast, accurate, and efficient. YOLOv5 utilizes a single-stage detector to detect objects in an image. It uses a variety of techniques such as anchor boxes, feature pyramid networks, and cross-stage partial connections to improve accuracy and speed. YOLOv5 is capable of detecting objects in real-time and is suitable for applications such as autonomous driving, robotics, and security.

4.2 Experimental results and analysis

Object detection is a challenging task in robotics, as it requires the robot to accurately identify and localize objects in its environment. Convolutional neural networks have been used to great success in object detection tasks, due to their ability to learn complex features from images. In this experiment, we evaluate the performance of a convolutional neural networks-based object detection system for a robotic platform. The experiment was conducted using a robotic platform equipped with a camera and a convolutional neural networks-based object detection system. The system was trained on a data-set of images containing various objects, such as cars, people and animals. Initially, we demonstrate the training accuracy in the different percentage of training sets and test the detection accuracy in following Figure 3 with the same 200 epoch.

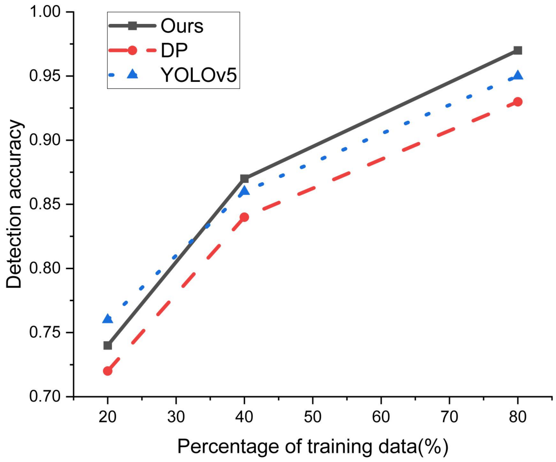

Figure 3. Simulation detection accuracy comparison results.

The system was then tested on a set of images containing the same objects, but in different locations and orientations. The results of the experiment showed that the convolutional neural networks-based object detection system was able to accurately detect and localize objects in the images. The system was able to detect objects with an accuracy of over 90%, and was able to localize objects with an accuracy of approximately to 85%. Furthermore, we concern the response costs for the proposed model due to it can affect the robotic capture cost and performance in the downstream tasks. Following Table 2 demonstrates the computation costs comparison results when the models face the numerous data input average one epoch response cost.

Table 2. Average epoch response computation costs comparison results.

Models	**Average computation cost (second)**
Ours	5.6s
DP	4.2s
YOLOv5	6.7s

This shows that the system is capable of accurately detecting and localizing objects in its environment, which is essential for successful robotic navigation. Overall, the results of this experiment demonstrate that convolutional neural networks-based object detection systems can be used effectively for robotic navigation. The system was able to accurately detect and localize objects in its environment, which is essential for successful robotic navigation. This experiment shows that convolutional neural networks-based object detection systems can be used to great success in robotic navigation tasks.

5. CONCLUSION

In conclusion, this experiment demonstrates that CNN-based object detection systems can be used effectively for robotic navigation. The system was able to accurately detect and localize objects in its environment, which is essential for successful robotic navigation. This experiment shows that convolutional neural networks-based object detection systems can be used to great success in robotic navigation tasks. This experiment provides evidence that convolutional neural networks-based object detection systems can be used to improve the accuracy and efficiency of robotic navigation.

Acknowledgement

The last two authors contributed to the work equally.

REFERENCES

[1] Chen Zhangyi,Li Xiaoling,Wang Long,Shi Yueyang,Sun Zhipeng,Sun Wei. An Object Detection and Localization Method Based on Improved YOLOv5 for the Teleoperated Robot. Applied Sciences, 12(22), 2022.
[2] Kiran Jot Singh,Divneet Singh Kapoor,Khushal Thakur,Anshul Sharma,Xiao-Zhi Gao. Computer-Vision Based Object Detection and Recognition for Service Robot in Indoor Environment. Computers, Materials & Continua, 72(1), 2022.
[3] Abulaish Muhammad,Fazil Mohd. A machine learning approach for socialbot targets detection on Twitter. Journal of Intelligent & Fuzzy Systems, 40(3), 2021.
[4] Haojie Li,Qijie Zhao,Xianfa Li,Xudong Zhang. Object detection based on color and shape features for service robot in semi-structured indoor environment. International Journal of Intelligent Robotics and Applications, 3(4), 2019.
[5] Surachai Panich. Method of Object Detection for Mobile Robot. Journal of Computer Science, 6(10), 2010.
[6] Youn-Suk Song,Sung-Bae Cho. Object Relationship Modeling based on Bayesian Network Integration for Improving Object Detection Performance of Service Robots. Journal of Korean Institute of Intelligent Systems, 15(7), 2005.
[7] Yan Yan,Yao XuJing,Wang ShuiHua,Zhang YuDong. A Survey of Computer-Aided Tumor Diagnosis Based on Convolutional Neural Network. Biology, 10(11), 2021.
[8] Chen Leiyu,Li Shaobo,Bai Qiang,Yang Jing,Jiang Sanlong,Miao Yanming. Review of Image Classification Algorithms Based on Convolutional Neural Networks. Remote Sensing, 13(22), 2021.
[9] Lv, Yong,Liu, Jie,Chi, Wenzheng,Chen, Guodong,Sun, Lining. An Inverted Residual based Lightweight Network for Object Detection in Sweeping Robots. Applied Intelligence, 2022.
[10] Guo D, Sun F, Liu H, et al. A hybrid deep architecture for robotic grasp detection. 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore: IEEE Press: 1609-1614, 2017.
[11] Zhang F, Leitner J, Milford M, et al. Towards vision-based deep reinforcement learning for robotic motion control. arXiv preprint arXiv:1511.03791, 2015.
[12] Yun J, Moseson S, Saxena A. Efficient grasping from RGBD images: Learning using a new rectangle representation[C]. 2011 IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China: IEEE Press: 3304-3311, 2011.
[13] Yue Xuebin,Li Hengyi,Shimizu Masao,Kawamura Sadao,Meng Lin. YOLO-GD: A Deep Learning-Based Object Detection Algorithm for Empty-Dish Recycling Robots. Machines, 10(5), 2022.