2023 3rd International Conference on Image Processing and Intelligent , Kuala Lumpur, Malaysia

Research on Object Detection of Robotic based on Convolutional Neural Network

Xi Lin; Yuge Chen; Donglai Liu Dalian University of Technology, School of Mechanical Engineering

1. Abstract

With the continuous dev. 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: CV, Robotics, Object detection, CNN, Computation cost.

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

2. Introduction

Robot

Robot is a crucial mechanical equipment in modern industry. With the continuous development of artificial intelligence technology, intelligent robots emerged. They play a significant role in production, social life, and labor shortage solutions. Thus, enhancing robot controllability for identifying multi-target objects in complex environments holds significant value.

Object Detection

Obj. detection is critical in robotics and locating objects in images/videos. It's core to apps like autonomous navigation, obj. manipulation, and recognition. Alg. can detect indoor/outdoor obj. in static/dynamic scenes, in 2D/3D and various sizes/colors. It can ID obj. in real-time, enabling robots to interact accurately. [3].

Convolution

CNNs (Convolutional Neural Networks) are used for image recognition and classification. They consist of multiple layers of neurons that analyze images. Each layer is connected to the next, and the output of each layer generates the next. CNNs learn complex patterns in images and can classify/detect objects or generate images. They have various applications, including computer vision, natural language processing, and robotics. CNNs are also utilized in medical imaging for detecting and classifying tumors and abnormalities.

Application

As shown in Figure 1 (a) is an automatic screw tightening robot, 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. 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 .

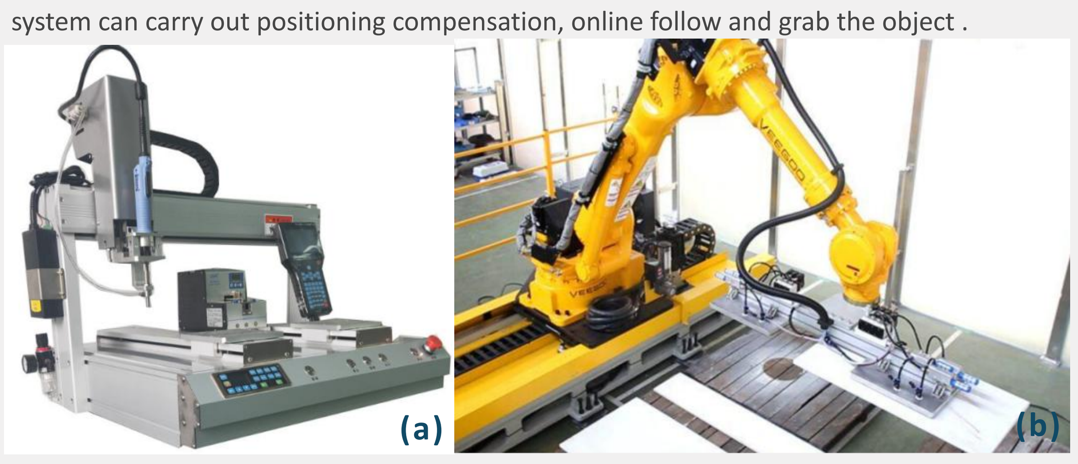

Figure 1. Existing robotic machine with object detection components.

3. Preliminaries

Related works

Initially, vision-based robot target recognition requires a combination of multi-faceted algorithms and actual physical hardware facilities. Researcher Zhang used DL 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 auto-identify target objects and capture the targets . With the deeper investigation of DL algorithm development, researcher Yun proposed a multi-modal network model based on DL, which generates a new feature sharing layer through RGB and deep image feature extraction and fusion, and then obtains the objects identification results through the iterations of learning training.

Primary parameters description

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

4. Model Framework

Primary procedures illustration

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

1. 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.

2. 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.

3. 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.

4. 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.

5. 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.

6. 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.

7. 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.

Convolution neural network framework

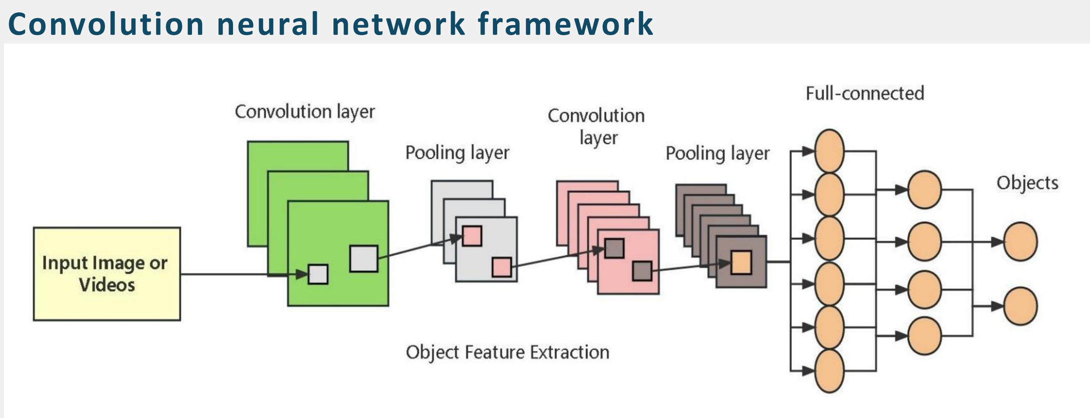

Figure 2. Convolutional neural network general structure demonstration.

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:

$$c = \mathrm{conv}(p_1 + p_2)$$

5. Experimental Result and Analysis

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.

Experimental results and analysis

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.

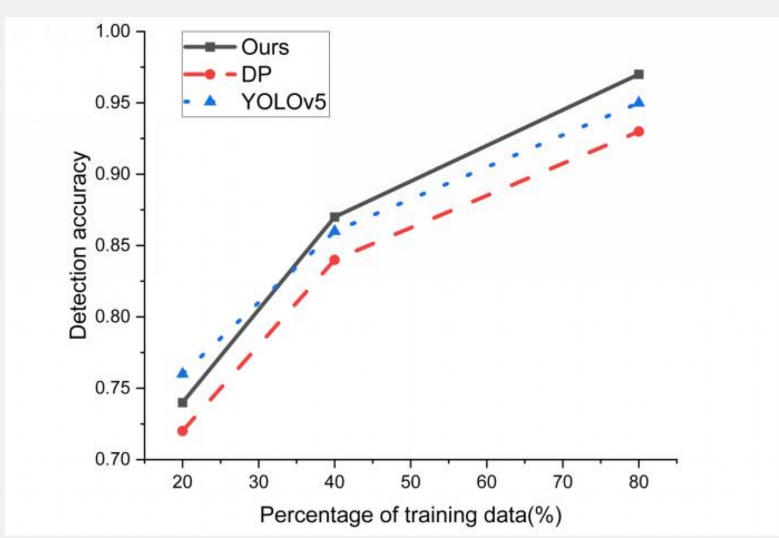

Figure 3. Simulation detection accuracy comparison results..

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.

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

Table 2. Average epoch response computation costs comparison results.

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.

6. Conclusions

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.

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.