Object Detection and Pose Estimation based on Deep Learning

Present Name : 12010611相昊阳

2024-3-26

AncoraSIR.com

SUSTech

Project Summary

- Our project investigates automated parcel sorting with a robotic arm in a simulated environment. With the rapid development of the logistics industry, traditional manual sorting can no longer cope with heavy sorting loads. Therefore, vision-based intelligent logistics sorting systems have significant research value. Detecting objects' positions, categories, and attitude information accurately and quickly has become a key issue in the implementation of intelligent sorting systems.
- We'll delve into literature on object detection, pose estimation, and robotics control for background. And the detailed content of the references will be listed in the following slides.
- We plan to utilize well-established datasets such as YCB, Linemod, COCO, and the Cornell Grasping Dataset, known for their diversity in object types and complexity in scenes, providing a robust foundation for training our models. If necessary, we might collect new data from simulated environments tailored to our project's specific needs, then we will label the object's position and orientation data ourselves.

AncoraSIR.com

SUSTech

Project Summary

- We propose to use the YOLO model for object detection due to its efficiency and accuracy. For pose estimation, we will explore deep learning-based pose estimation methods, such as DenseFusion. While there are existing implementations available, we plan to fine-tune these models with our collected data to enhance their performance.
- To evaluate our results, we decided to adopt both score assessment and ratio assessment methods. Score assessment refers to using a specific score range to represent the quality of experimental results, where a high score indicates good results, and conversely, a low score suggests the results are worse than expected. Ratio assessment means quantifying experimental results through objective ratios, such as accuracy, precision, and recall, etc.

AncoraSIR.com

What is the problem that you will be investigating?

Why is it interesting?

- Develop a logistics sorting system based on deep learning for parcel classification and visual sorting. The system should be able to recognize the dimensions of parcels and sort them in a disorderly arrangement.
- This research problem arises from the rapid development of e-commerce, which has led to a significant increase in parcel logistics volume. Traditional manual classification and sorting methods are no longer sufficient to meet the demands of efficient operations in modern logistics centers. Therefore, the development of robotic systems capable of automatically completing sorting tasks is particularly important. Solving this problem not only improves the efficiency and accuracy of logistics processing but also has the potential to significantly reduce logistics costs and enhance customer satisfaction.

AncoraSIR.com

SUSTech

What reading will you examine?

To provide context and background

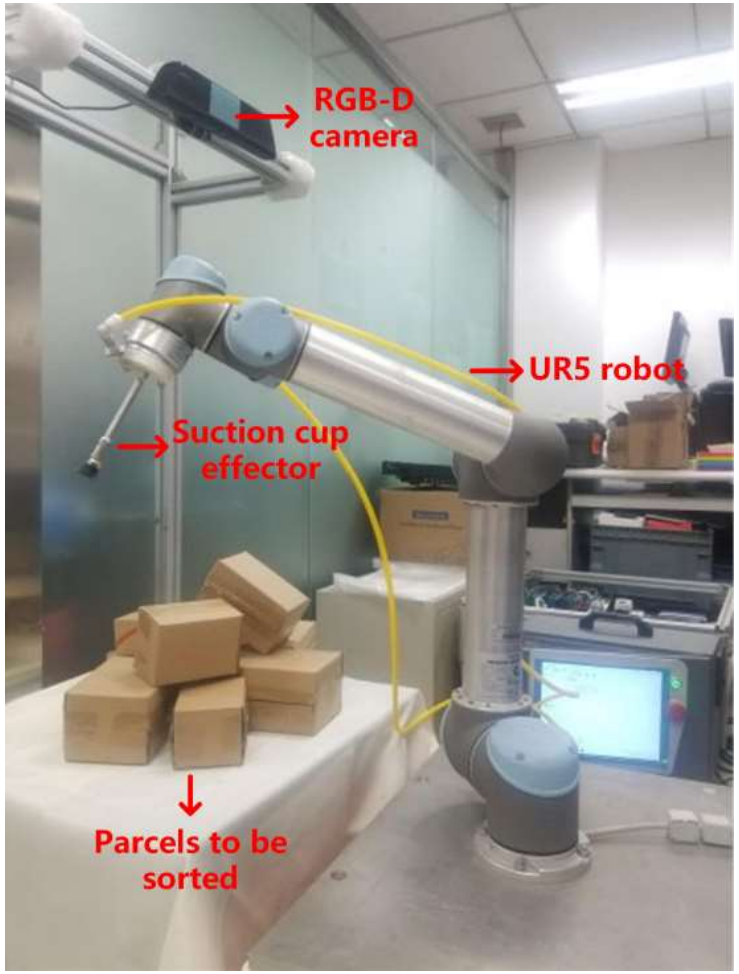
1. Han, Song, Xiaoping Liu, Xing Han, Gang Wang, and Shaobo Wu. 2020. "Visual Sorting of Express Parcels Based on Multi-Task Deep Learning" Sensors 20, no. 23: 6785. https://-doi.org/10.3390/s20236785

The article "Visual Sorting of Express Parcels Based on Multi-Task Deep Learning" focuses on enhancing the efficiency and accuracy of sorting express parcels in complex scenarios using a robot sorting method powered by multi-task deep learning.

In addition, we will read more papers on object detection and pose estimation including:

2. Xing, H.; Xiao-Ping, L. Robotic sorting method in complex scene based on deep neural network. J. BeijingUniv. Posts Telecommun. 2019, 42, 22–28.

3. Andy, Z.; Shuran, S.; Kuan-Ting, Y.; Elliott, D.; Alberto, R. Robotic pick-and-place of novel objects in clutter with multi-affordance grasping and cross-domain image matching. In Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, Australia, 21–26 May 2018.



AncoraSIR.com

SUSTech

What data will you use?

If you are collecting new data, how will you do it?

- Dataset from web
  - We will use well-established datasets such as YCB, Linemod and COCO to train our model.
- New data
  - Get object photos and pose information from RGB-D camera.
  - Do the image processing using OpenCV.
  - Label the object's position and orientation data ourselves.

AncoraSIR.com

SUSTech

What method or algorithm are you proposing?

Computer Vision and Deep learning

- Computer Vision
  - **Data Acquisition**: imagery and depth information from **RGB-D cameras**
  - **Image Preprocessing**: reading, cropping, resizing, enhancement using **OpenCV**
- Deep Learning

  Utilizing the **TensorFlow** deep learning framework
  - **Object Detection**: pre-trained **YOLO** model
    - Finetune the YOLO model with annotated data generated from the environment
  - **Pose Estimation**: **pose estimation** method
    - Explore available pose estimation (e.g. **DenseFusion**)

AncoraSIR.com

SUSTech

What method or algorithm are you proposing?

Potential Future Improvements

- Algorithm Optimization
  - Adjust **network structure**, **loss functions**, and **training strategies** to enhance accuracy and real-time performance.
- System Integration
  - Integrate **object detection**, **pose estimation**, and **robot arm gripping** into a unified system.

AncoraSIR.com

How will you evaluate your results?

Two kinds of evaluation methods

Score	
60>	Unable to complete the task at all, serious errors occurred during the gripping process.
60-70	The box can be gripped to the specified area, but there are issues with uneven placement or unstable gripping.
70-80	The box can be gripped to the correct position and roughly placed neatly.
80-90	The box can be accurately gripped and placed.
90-100	The box can be gripped and placed quickly and accurately.

To evaluate our result, we will apply some objective evaluation criteria, depending on the specific circumstances.

1) Accuracy: the proportion of objects that are correctly grasped and placed among all objects.

2) Precision: the proportion of samples correctly predicted as positive out of all samples predicted as positive by the model.

3) Recall, F1 Score …

AncoraSIR.com

Thank you for listening

AncoraSIR.com

SUSTech