

Deep Learning for Object Manipulation

Deep learning is a topic that is attracting increasing attention in machine learning, computer vision, and robotics. There exist various CNN-based deep architectures for object segmentation and recognition. However, there are not many deep architectures addressing the task of object manipulation in robotics. We suggest creating a deep architecture which takes the raw camera images as input and informs the robot how to manipulate the object. In the literature, one can find an already trained deep neural network for vision-based scene representation. We utilize one of them as the first part of our deep architecture and construct the second part which receives the vision result and outputs the optimal robot's pose to push or manipulate the object.

The student working on this project should learn how to adjust or use a deep neural network for the first part of our deep architecture, and also construct a reasonable deep structure and learning strategy for the second part by themselves. They also need to collect and annotate data for training the second part of the deep architecture.

The result of the project should be a well-trained deep architecture which can suggest where and in which direction to push the object, so that the robot can better grasp or manipulate the object. Applicants must be interested in deep learning and should possess good programming skills either in Python or C/C++.

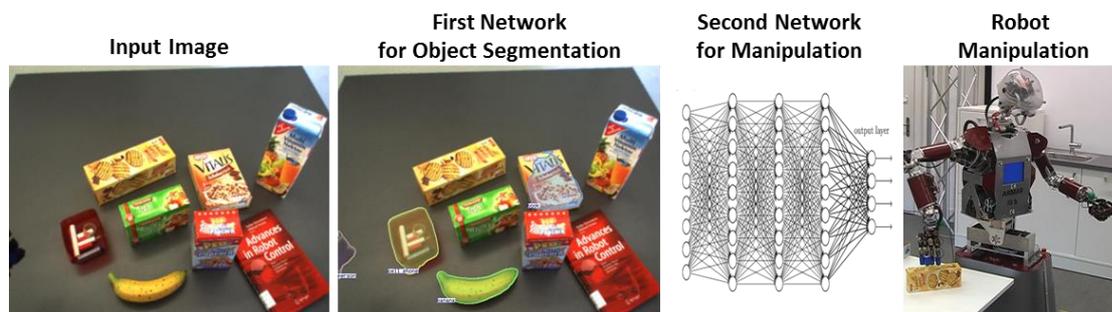


Figure 1 The input image is processed by the first deep network to segment and detect objects. The second network will output how and where to push one of the detected objects by the robot.