Design and Testing of Vision Algorithms for an Interactive Soft Robotic Hand

Matthew Du [1] Zhanfeng Zhou [1] Xinyu Liu [1 2]

[1]Department of Mechanical and Industrial Engineering

[2]Department of Biomedical Engineering

Problem Definition

Soft robotic grippers are safer to use for interacting with humans than hard robotic grippers because of their deformable structure and dexterity. Despite being safe [1], there is a lack of research on the advantages of using soft robotic grippers in human-robot interaction applications. This project looks for ways to locate objects grasped in a human hand and plan an optimal grasp using an in-hand RGB-D camera & segmentation / detection algorithms.

Objective goals:

Goal	Metric
Identify unique objects from a complex background	# objects detected
Grab objects securely without contact with human hand	# successful grasps

Table 1. Objective Goals

Related Work

Mechanical Search: looking for objects in various environments.

Figure 1. Mechanical Search

In-Hand Manipulation: Manipulation of objects using only hand movements.

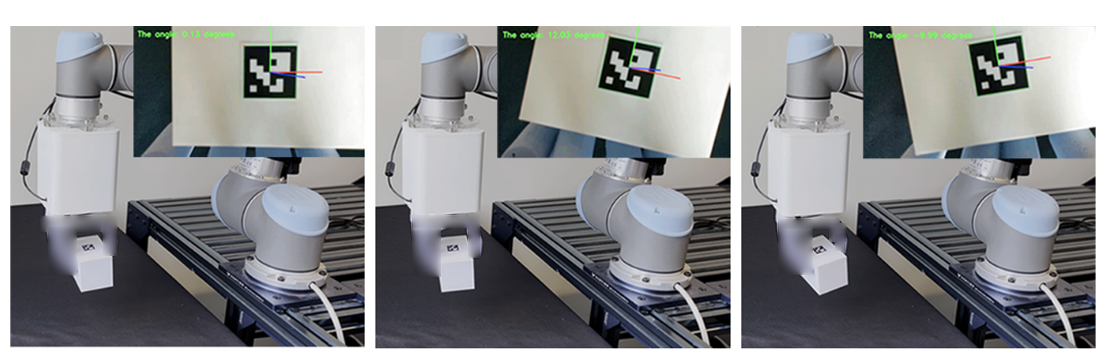


Figure 2. In-Hand Manipulation

Grasping Planning: Find optimal grasps based on geometric features [2].

Experimental Setup

Arm: Universal Robots UR5e robot; **Camera**: Intel Realsense LiDAR Camera L515; **Hand**: Soft 5-Fingered Robotic Hand

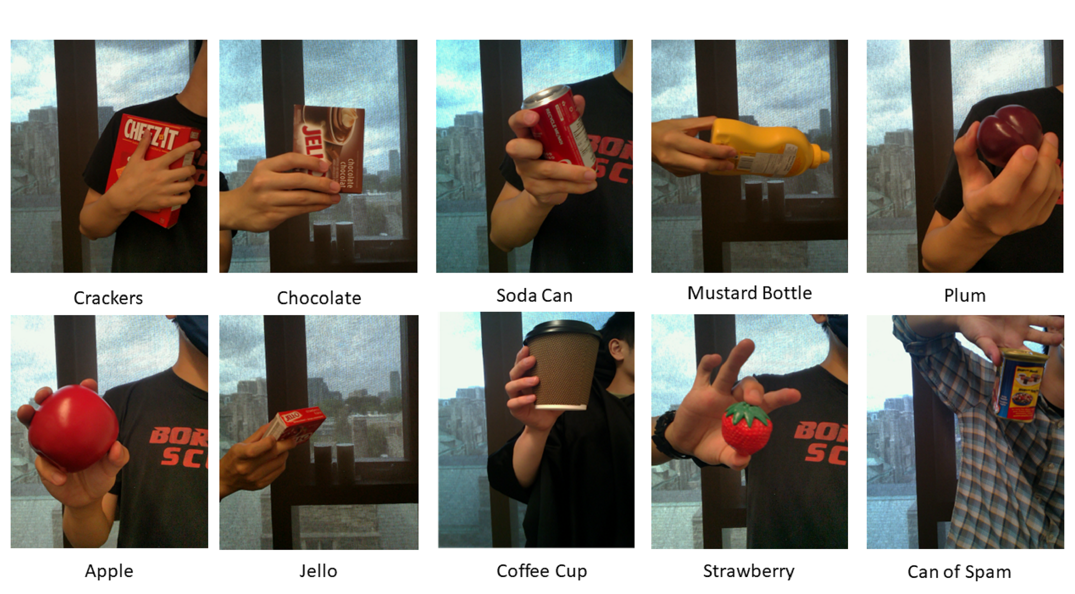


Figure 3. Sample of Training Set

Algorithm Description

The RGB image is inputted into a YOLO algorithm to identify a bounding box around the object. A mask of the region on the RGB and Depth image is created and used to form the point cloud, which is filtered and sent to the grasp planning algorithm.

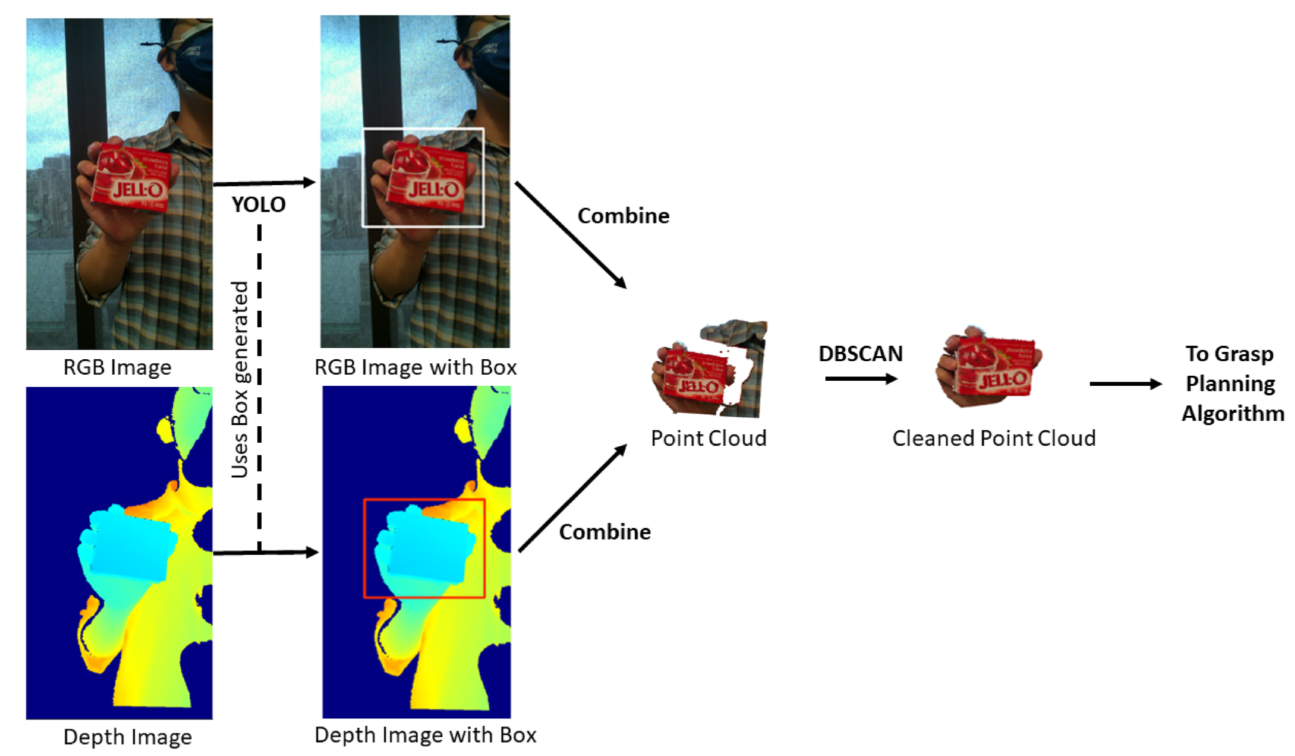
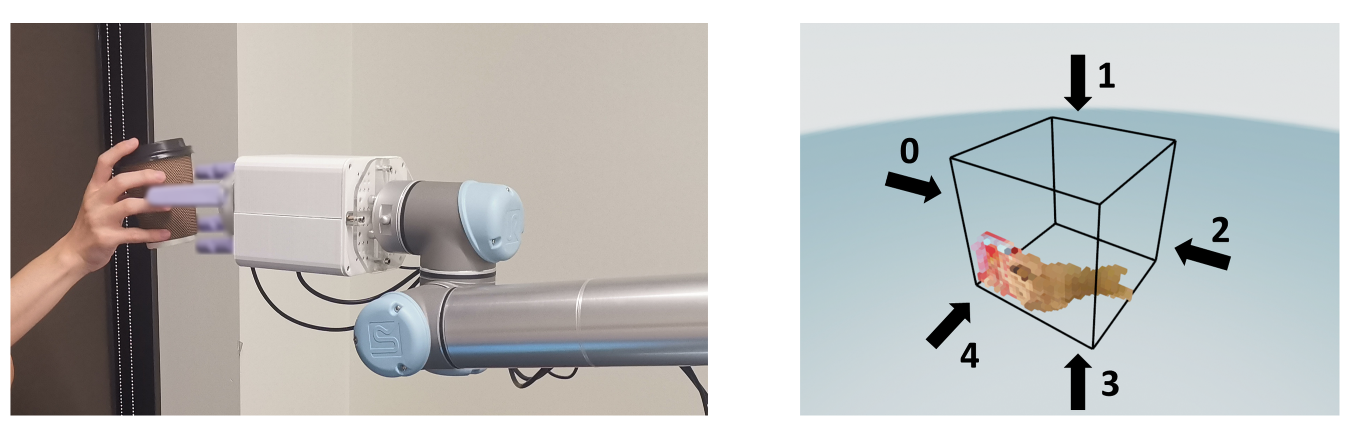


Figure 4. Diagram of Algorithm

Deep Learning Grasp Planning Model

Input: $32 \times 32 \times 32$ voxel grid generated using the scaled Point Cloud.

Output: A value from $\{0, 1, 2, 3, 4\}$, representing a different grasp position which can be used by a lookup table to obtain the grasp position.



(a) Sample Grasp (b) Trained Grasp Positions

Figure 5. Deep Learning Model

Figure 6 shows the 3D-CNN Architecture.

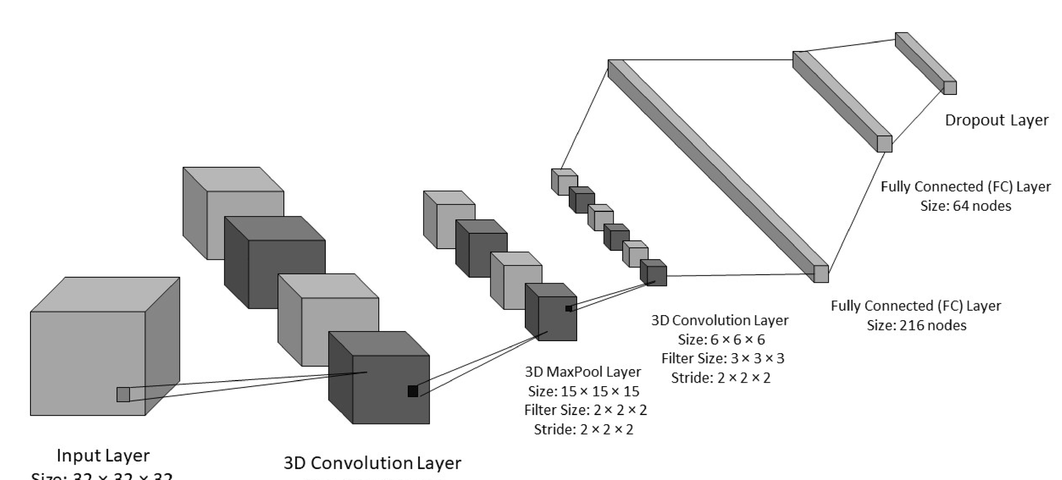


Figure 6. CNN Architecture

Applications

The onset of the COVID-19 Pandemic has increased demand for methods to limit human-to-human contact. This soft robot system could be used to hand objects to people safely in the manufacturing, hospitality, and food and beverage industries to reduce the amount of direct human contact. A safe soft robotic arm could also see applications in Health Care and Senior care for its ability to safely assist in grabbing objects.

State Machine Model

Inputs: Point Cloud **Outputs**: $\mathbb{R}^3 \times \mathbb{R}^3$ position and angle to grasp

For all $i \in \mathbb{N}$, define thumb position as $\vec{t}_i$ and finger position as $\vec{f}_i$. As shown in Figure 7, the state machine model treats hand positions as states, and searches the state machine for a state that satisfies a constraint function.

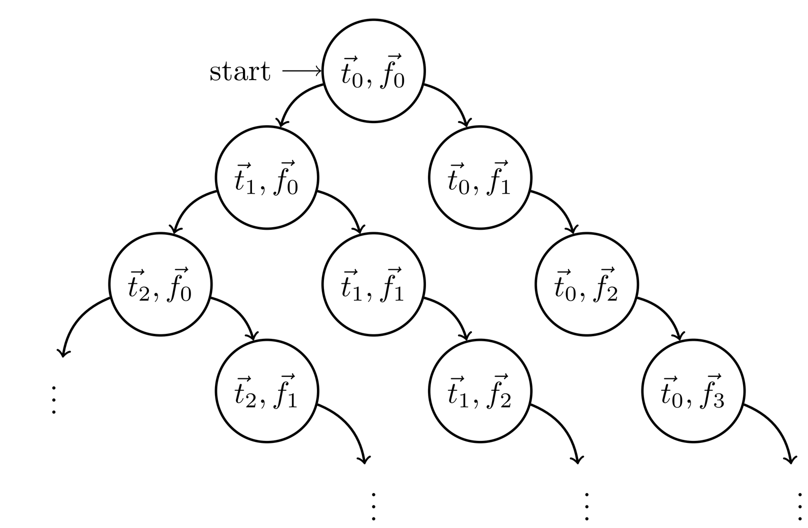
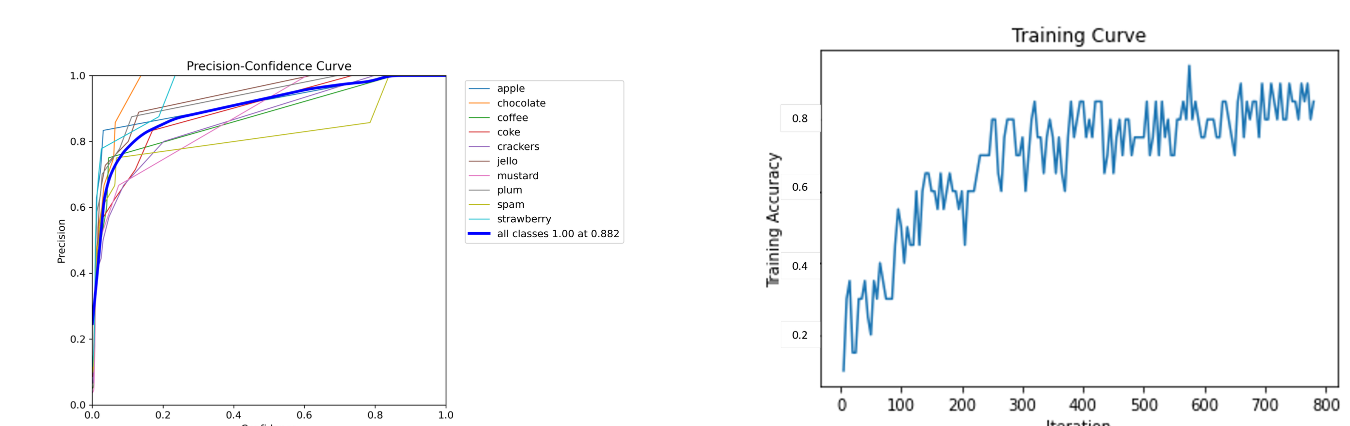


Figure 7. State Machine Model

Results

Machine Learning Models

The precision of the YOLO and 3DCNN are shown below in Figures 8b and 8a. The precision in getting the correct classification from the YOLO 95.8% with Batch size 64 and Epoch 60. The precision in getting the *correct output* from the 3DCNN is 85%.

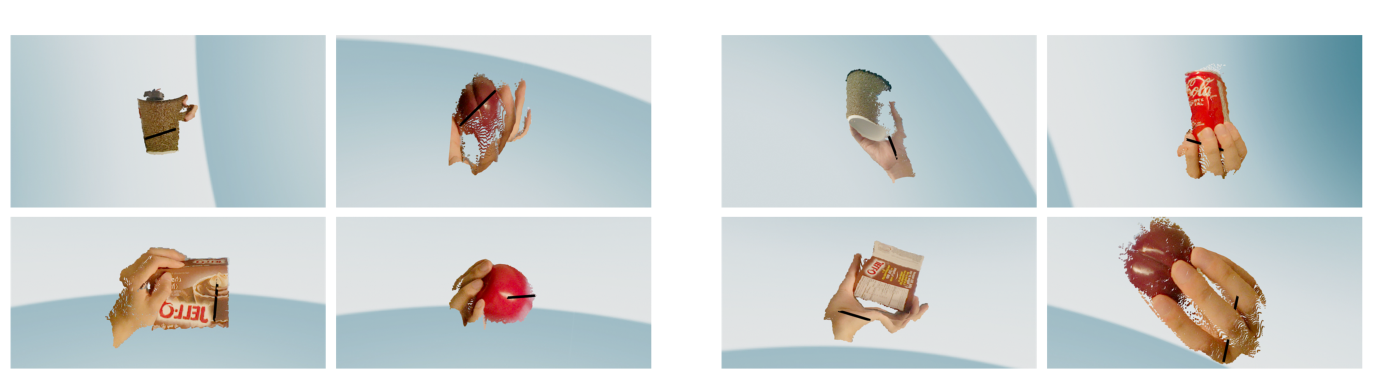


(a) YOLO Precision vs Confidence (b) 3DCNN Training Curve

Figure 8. Deep Learning Model Details

State Machine Model

The constraint function we used attempts to identify the hand by color. The precision was 60%, but a better constraint function could improve results. Sample of grasps are shown in Figure 9. Each end of the black line is the predicted position of the hand.



(a) Good Grasps (b) Poor Grasps

Figure 9. Sample of Grasps Generated

Future Work

We plan on trying other image segmentation algorithms, such as Mask R-CNN [3], to help improve image segmentation and classification abilities.

Additional work in collecting better dataset for machine learning, and better constraint functions would allow for more accurate and diverse set of grasping positions.

References

[1] D. Rus and M. T. Tolley, "Design, fabrication and control of soft robots," Nature, vol. 521, no. 7553, pp. 467–475, 2015. DOI: 10.1038/nature14543.

[2] P. Song, Z. Fu, and L. Liu, "Grasp planning via hand-object geometric fitting," The Visual Computer, vol. 34, no. 2, pp. 257–270, 2018.

[3] M. Danielczuk, M. Matl, S. Gupta, et al., "Segmenting unknown 3d objects from real depth images using mask r-cnn trained on synthetic data," in Proc. IEEE Int. Conf. Robotics and Automation (ICRA), 2019.