Using affordances for assembly: Towards a complete Craft Assembly System

Vitor H. Isume[1], Kensuke Harada[1,2], Weiwei Wan[1,2], Yukiyasu Domae[2]

[1]Graduate School of Engineering Science, Osaka University
[2]National Institute of Advanced Industrial Science and Technology (AIST)

ICCAS2021

Overview

- Introduction
- Related works
- Craft Assembly Task
- Method
- Results
- Experiment
- Conclusions

OSAKA UNIVERSITY

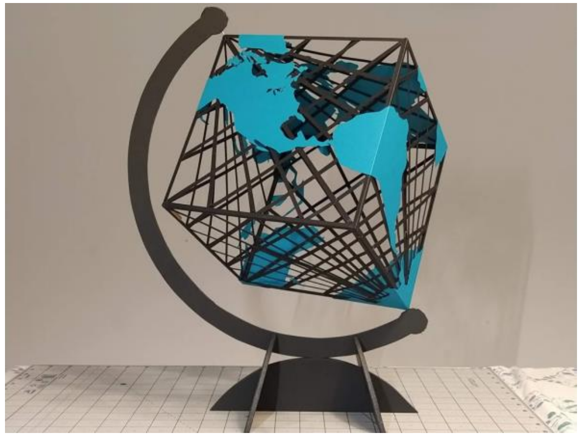
HARADA LAB
Robotic Manipulation

Introduction

- Do-it-yourself (DIY) products – custom, homemade objects that are similar to commercially available objects

- How to perform this kind of task in a robotic system?





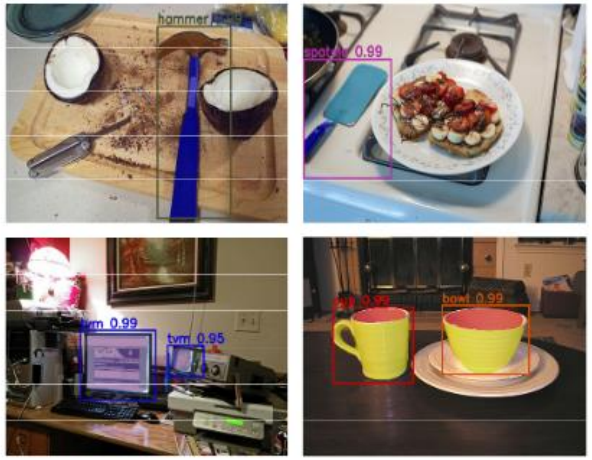

DIY projects taken from instructables.com

Related works

"The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill."

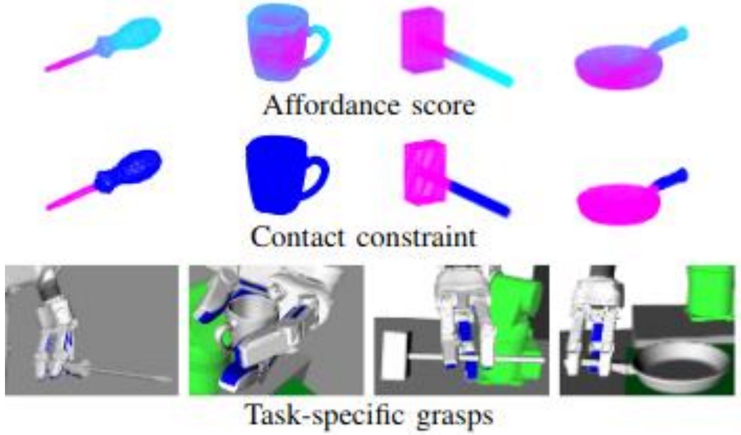
- James J. Gibson, 1979

- Detecting affordances



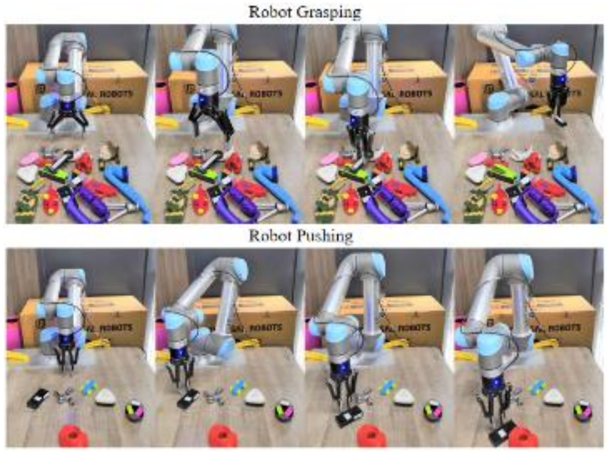
Do, T.T., Nguyen, A. and Reid, I. – "AffordanceNet: An End-to-End Deep Learning Approach for Object Affordance Detection" (2018)

- Task-specific Grasping



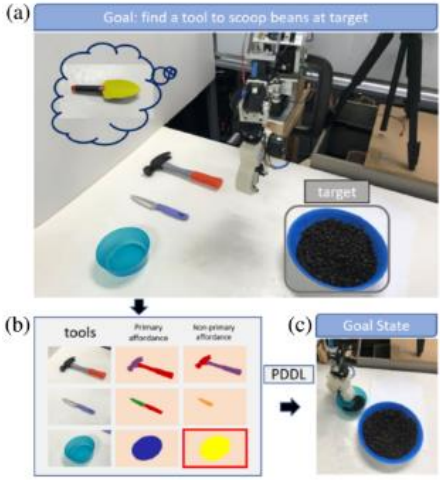
M. Kokic, J. A. Stork, J. A. Haustein, and D. Kragic – "Affordance detection for task-specific grasping using deep learning" (2017)

- Robotic manipulation

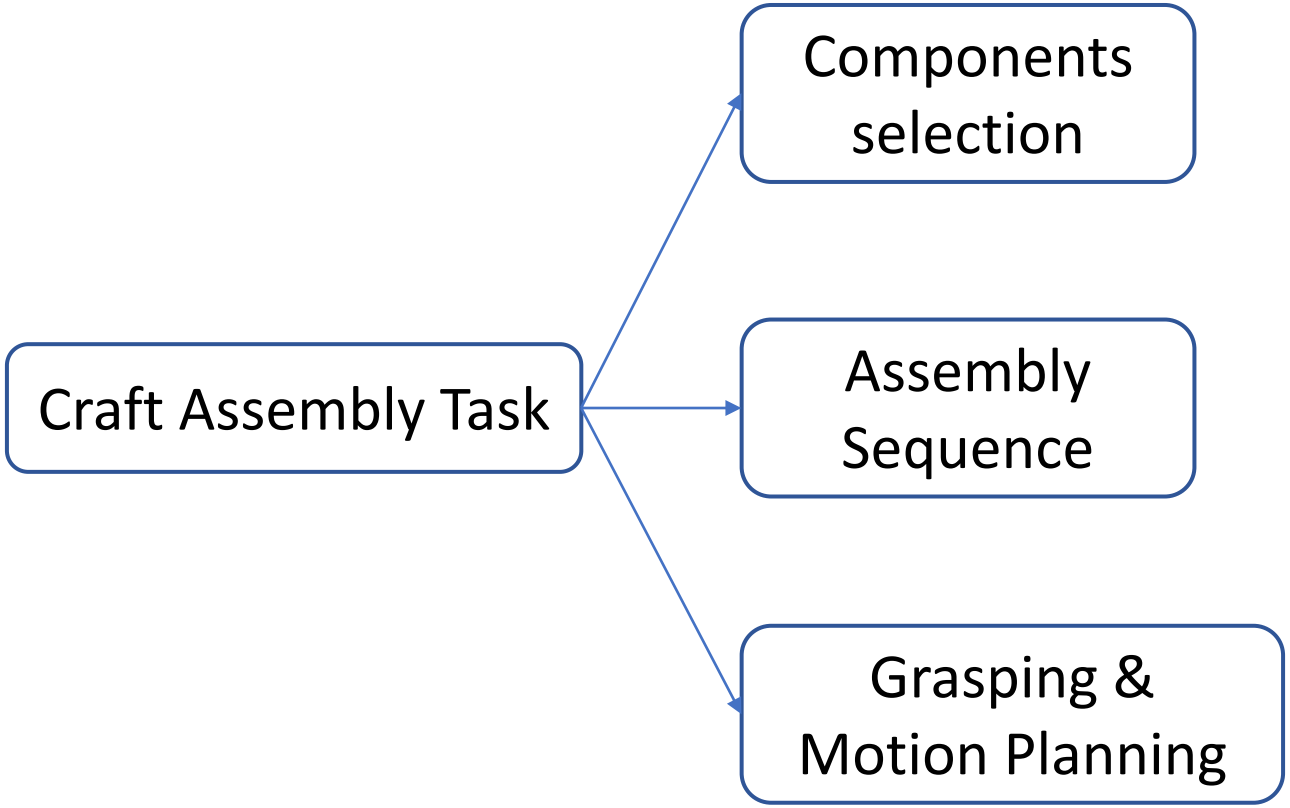


Wu, Z. Zhang, H. Cheng, K. Yang, J. Liu, and Z. Guo – "Learning affordance space in physical world for vision-based robotic object manipulation" (2020)

- Alternative solutions



F- J. Chu, R. Xu, L. Seguin, and P. A. Vela – "Toward affordance detection and ranking on novel objects for real-world robotic manipulation" (2019)



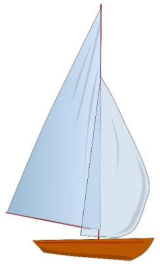
Affordance usage:

-Classify functionality of parts

-Determine connectivity between the components

-Determine grasp regions

Goal product:



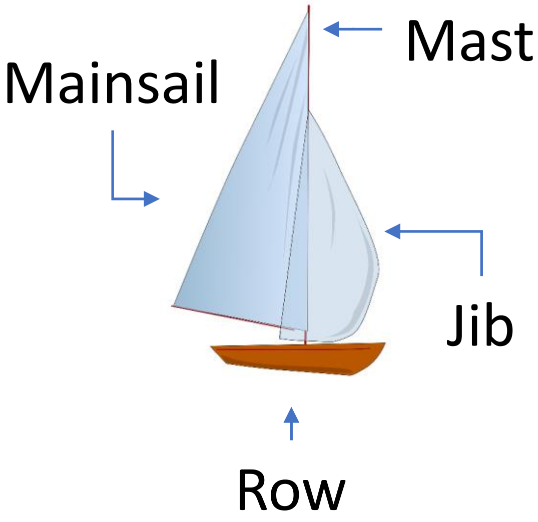
Craft Assembly Task

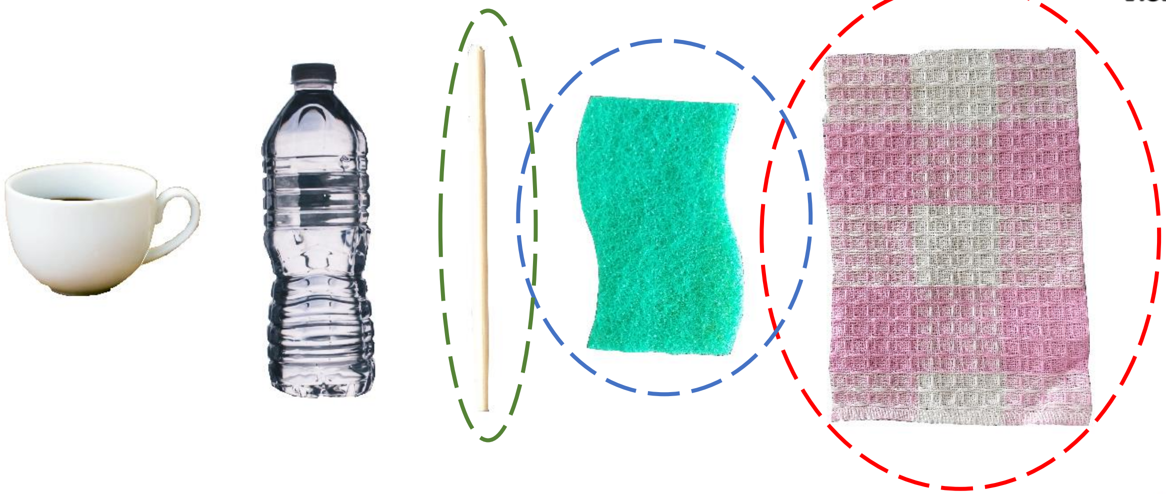
Components selection

Assembly Sequence

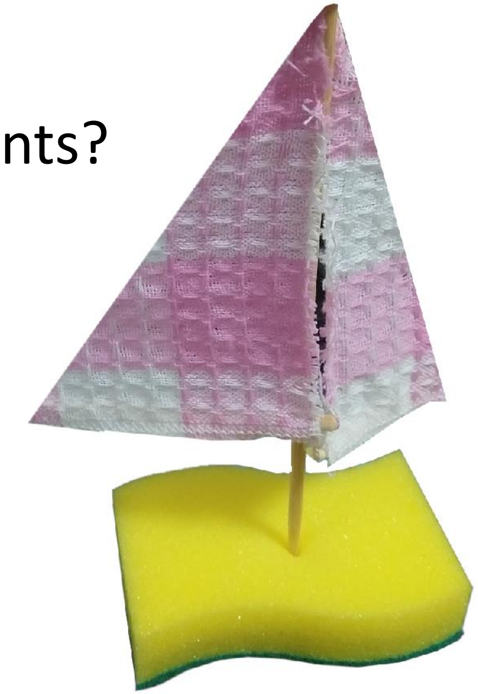
Grasping & Motion Planning

Components selection



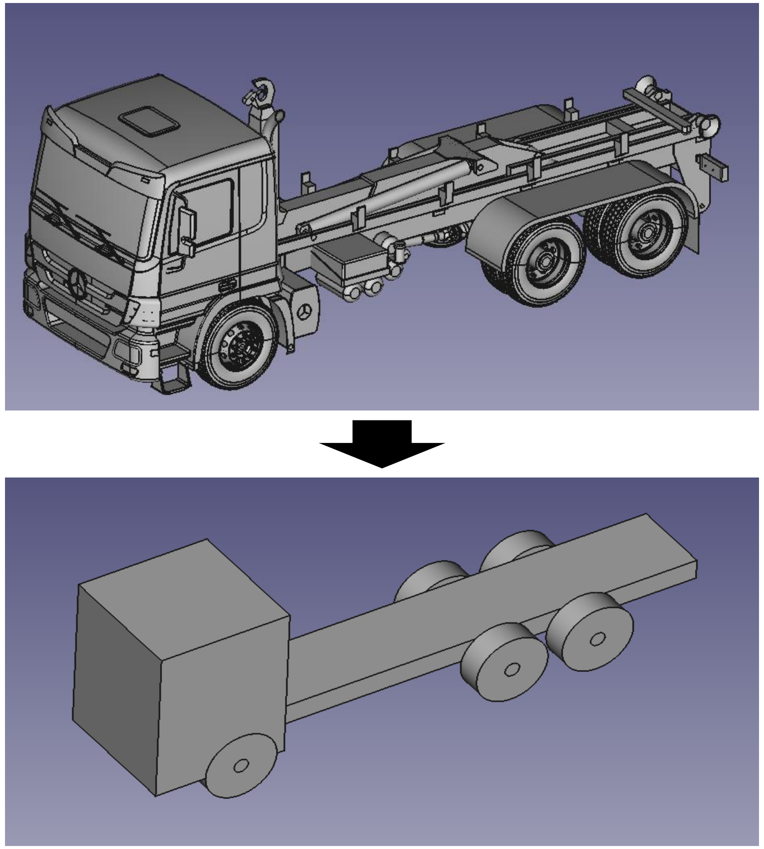


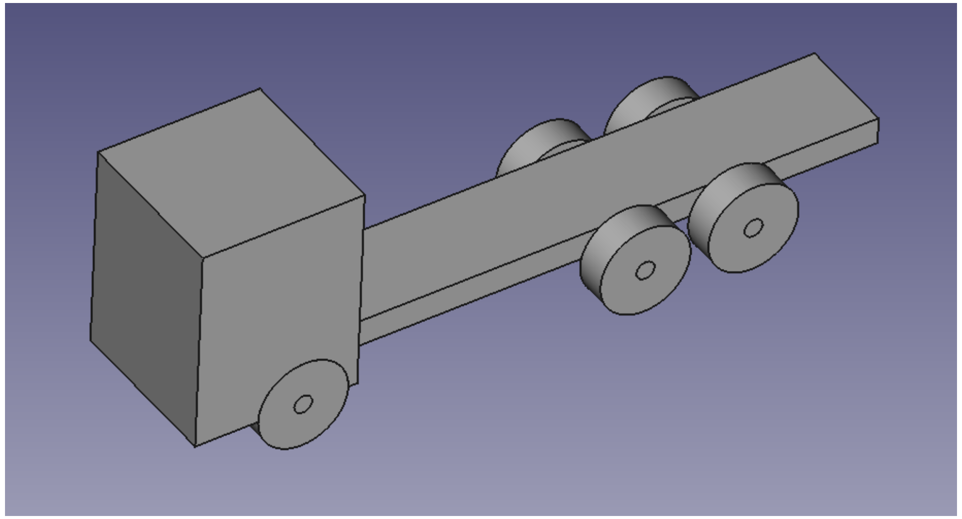
- How to select the components?
  - Appearance
  - Functionality



Problem Definition

- Goal: Build a toy truck
  - Simplified 3D CAD model of a truck, where each component is a primitive shape with affordance labels
- System:
  - For each component in the goal product, find candidate that matches the required **affordances**, **shape** and has the most similar **size**





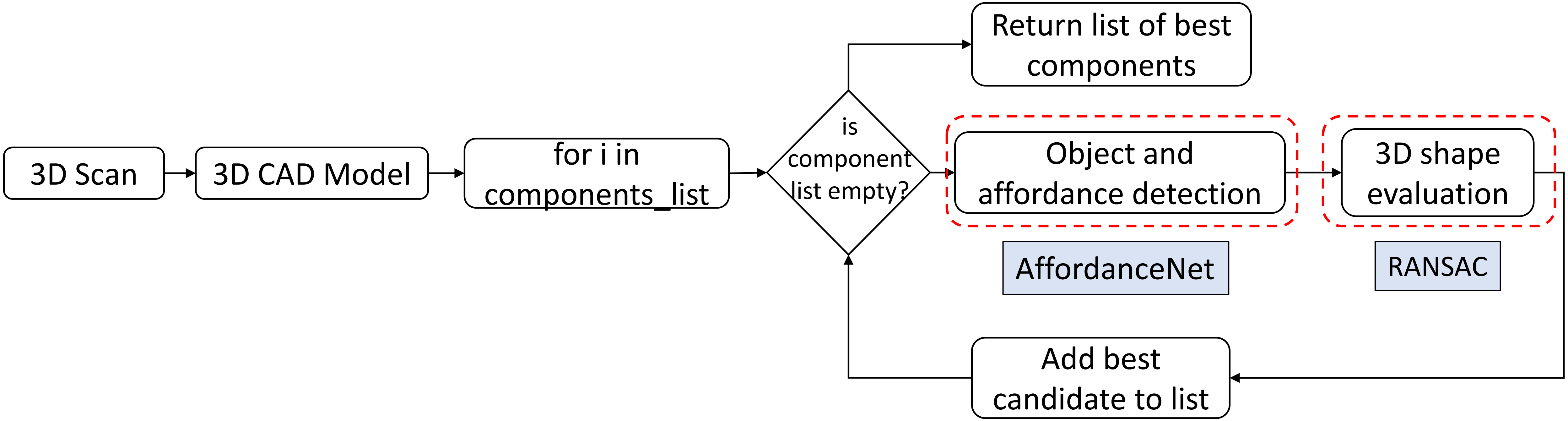
Cuboids	Length x Width x Height (mm)	Affordance label
Cabin	180 x 180 x 230	contain; support
Chassis	500 x 110 x 30	support

Cylinders	Diameter x Height (mm)	Affordance label
Axle	20 x 180	rollable
Wheel	100 x 35	rollable

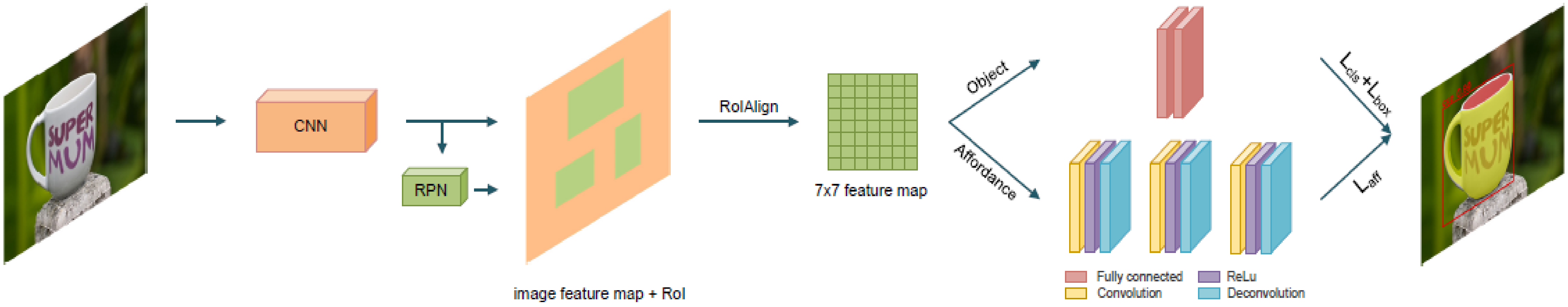
大阪大学 OSAKA UNIVERSITY

HARADA LAB Robotic Manipulation

Method



AffordanceNet



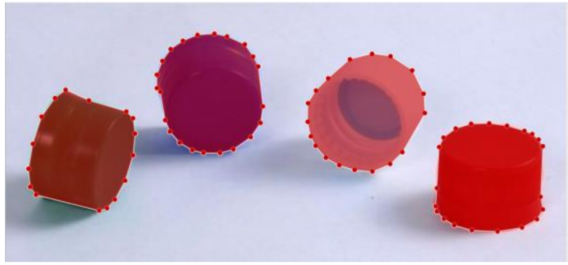
Do, Thanh-Toan, Anh Nguyen, and Ian Reid. "Affordancenet: An end-to-end deep learning approach for object affordance detection." (2018)

- Framework of CNNs
  - One branch for Object detection
  - One branch for Affordance detection

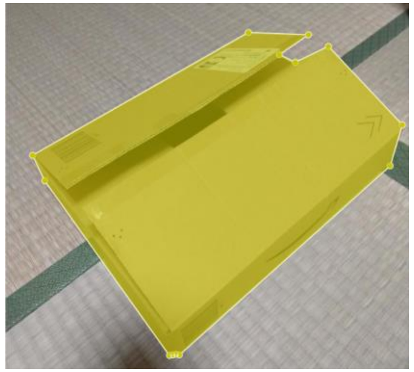
Database

- 5 Object classes and 3 affordance classes
- 1190 images of mostly isolated objects
  - Augmented 4 times – total of 5950 images
- 80% for training, 20% for testing.
- Training for 170k iterations.

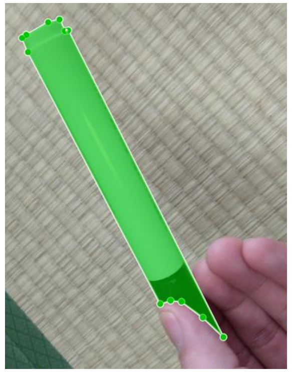
Object Class	Affordance label
PET bottle	Contain + rollable
Bottle cap	Contain + rollable
Cardboard box	Contain + support
Cup	Contain + rollable
Marker	Rollable



contain+rollable



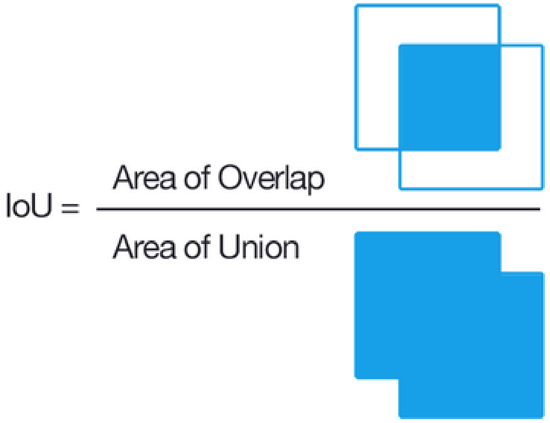
contain+support



rollable

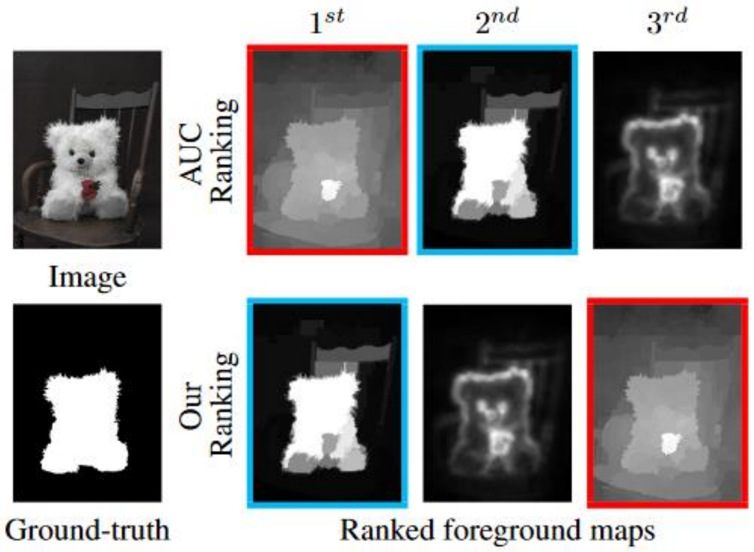
Evaluation

- Object detection:
  - Mean Average Precision (mAP)
    - IoU @0.5
- Affordance detection:
  - F_β^ω



$$Precision^\omega = \frac{TP^\omega}{TP^\omega + FP^\omega} \qquad Recall^\omega = \frac{TP^\omega}{TP^\omega + FN^\omega}$$

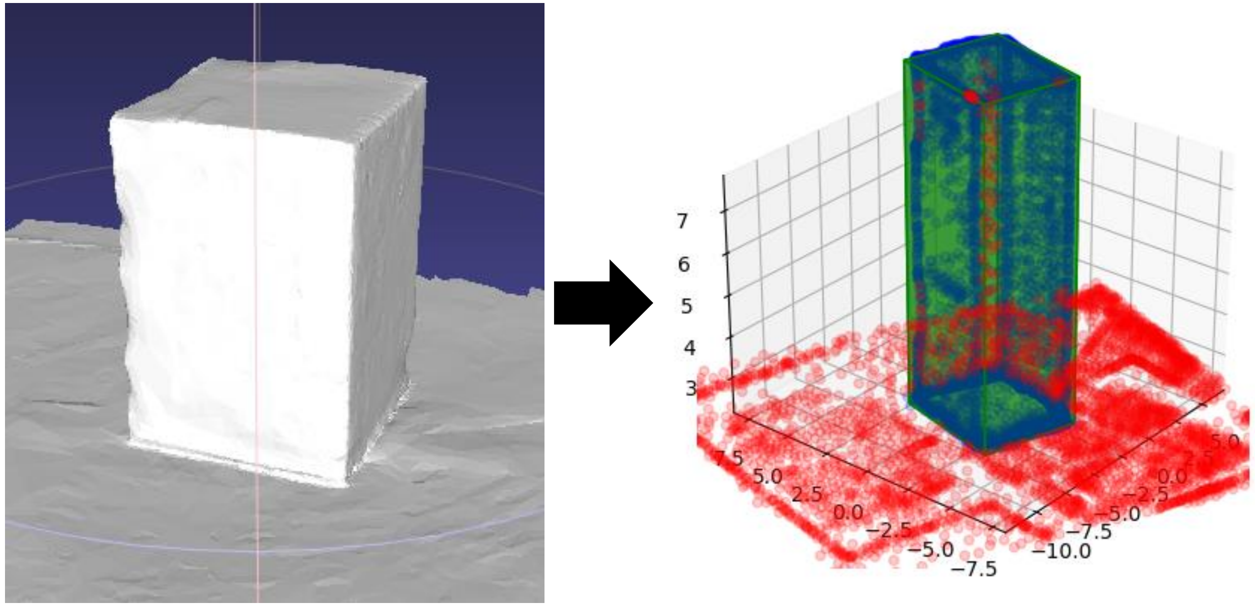
$$F_\beta^\omega = (1+\beta^2)\frac{Precision^\omega \cdot Recall^\omega}{Precision^\omega + Recall^\omega}$$

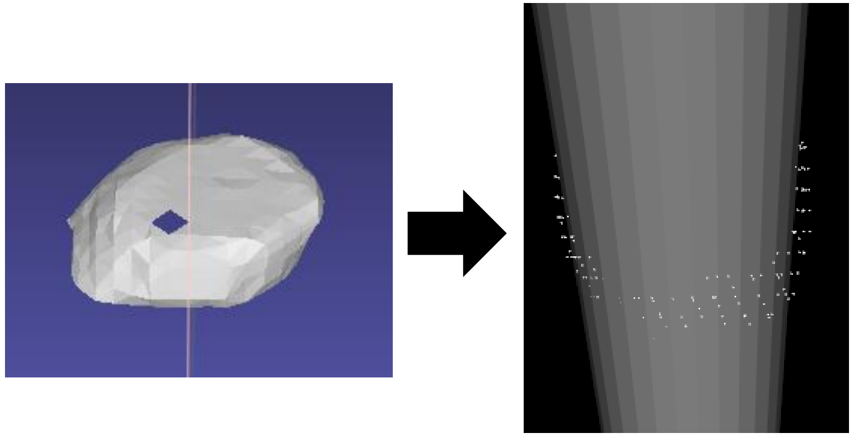


R. Margolin, L. Zelnik-Manor, and A. Tal, “How to evaluate foreground maps?”. (2014)

3D Primitive shape estimation

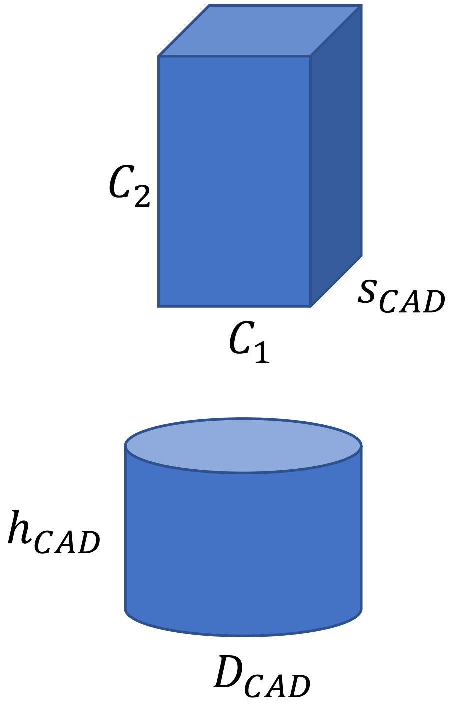
- SCHNABEL, R., WAHL, R., KLEIN, R. “***Efficient RANSAC for Point-Cloud Shape Detection***”. Computer Graphics Forum, Vol. 26, p. 214-226. 2007.
    - Improved computational time



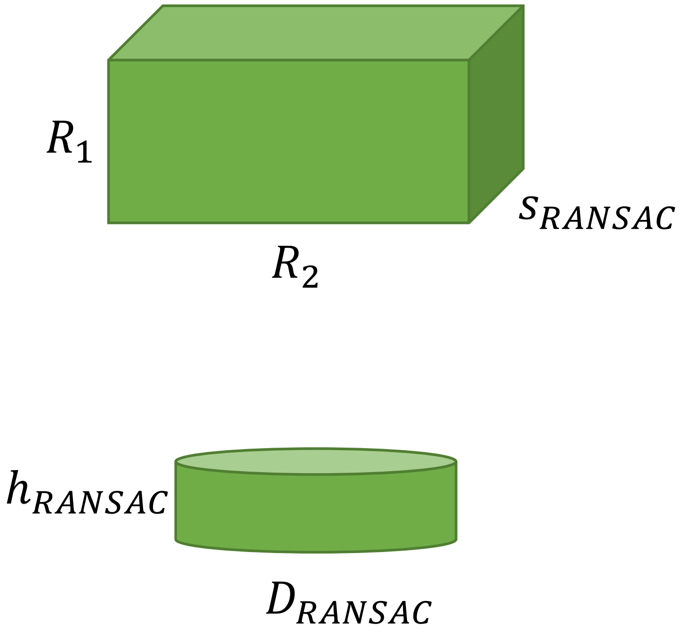


Dimension evaluation

CAD Component (groundtruth)



Object shape (RANSAC estimated)



Smallest dimension error

$$e_s = \frac{|s_{RANSAC} - s_{CAD}|}{s_{CAD}}$$

$$e_s = \frac{|D_{RANSAC} - D_{CAD}|}{D_{CAD}}$$

Proportion error

$$e_{r,1} = \left| \frac{R_1}{s_{RANSAC}} - \frac{C_1}{s_{CAD}} \right|$$

$$e_{r,2} = \left| \frac{R_2}{s_{RANSAC}} - \frac{C_2}{s_{CAD}} \right|$$

$$e_{r,1} = \left| \frac{h_{RANSAC}}{s_{RANSAC}} - \frac{h_{CAD}}{s_{CAD}} \right|$$

$$e_T = e_s + \sum e_{r,i}$$

Results

- AffordanceNet

Object Class	Average Precision
PET bottle	0.9377
bottle cap	0.7966
cardboard box	0.9406
cup	0.9726
marker	0.8357
mAP	**0.8967**

Affordance class	F_β^ω
contain + rollable	0.9357
contain + support	0.9425
rollable	0.7123

3D shape estimation

- Cylinder

	Measured (cm)		Estimated (RANSAC) (cm)		Absolute Difference		Score
	Radius	Height	Radius	Height	Radius	Height	
PET bottle	4.5	23	4.7	23.8	+0.2	+0.8	0.068
Bottle cap	3	3	3.6	2.9	+0.6	-0.1	0.516
Marker	0.8	14	1.2	13.2	+0.4	-0.8	3.75

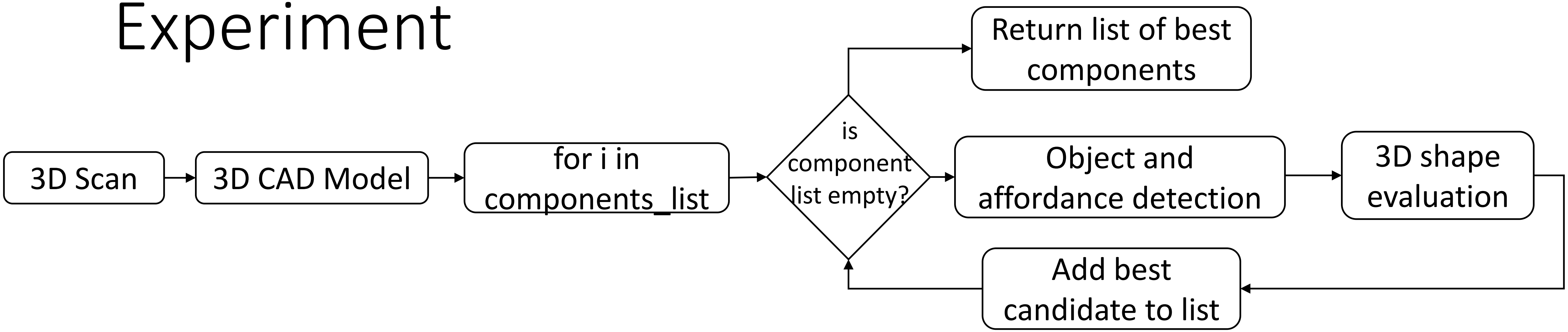
- Cuboid

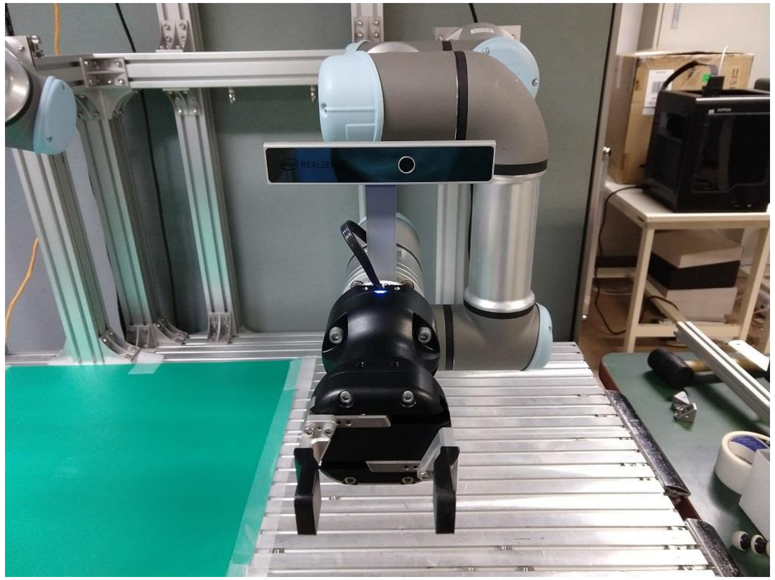
	Measured (cm)			Estimated (RANSAC) (cm)			Absolute Difference			Score
	Length	Width	Height	Length	Width	Height	Length	Width	Height	
Cardboard #1	31	8	4	24.4	7.0	3.4	-6.6	-1.0	-0.6	0.782
Cardboard #2	22	22	16	18.2	17.3	12.1	-3.8	-4.7	-3.9	0.428

OSAKA UNIVERSITY 大阪大学

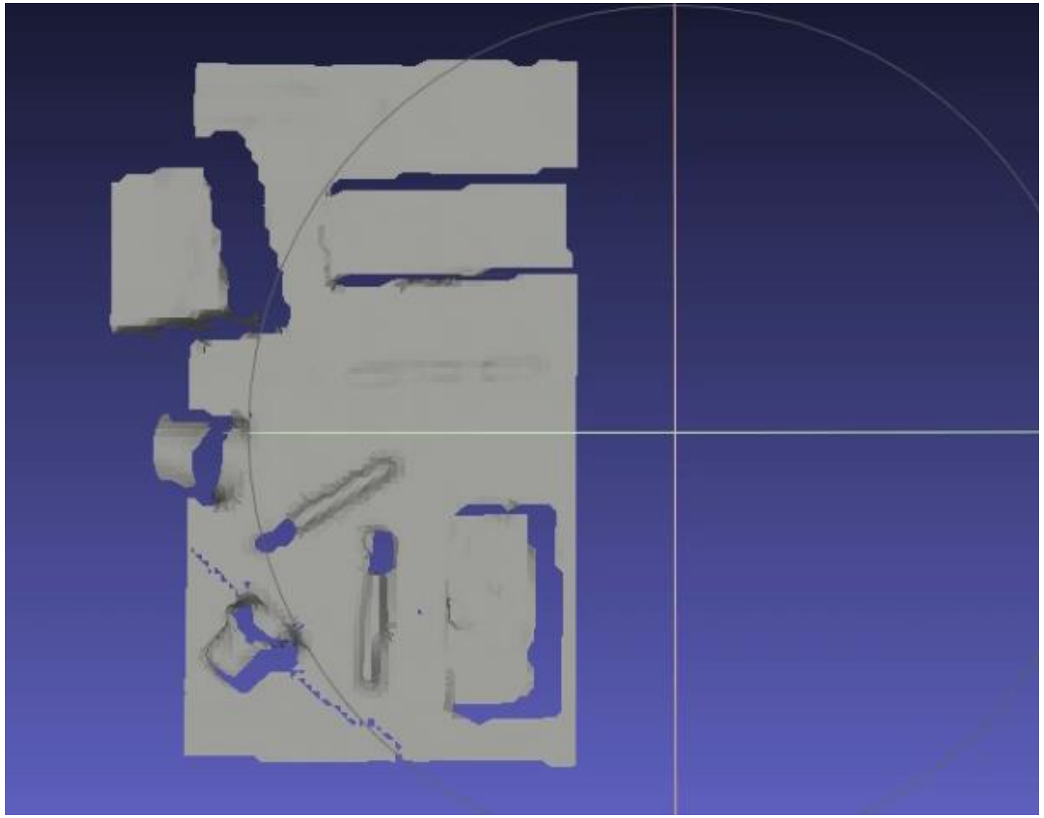
HARADA LAB Robotic Manipulation

Experiment





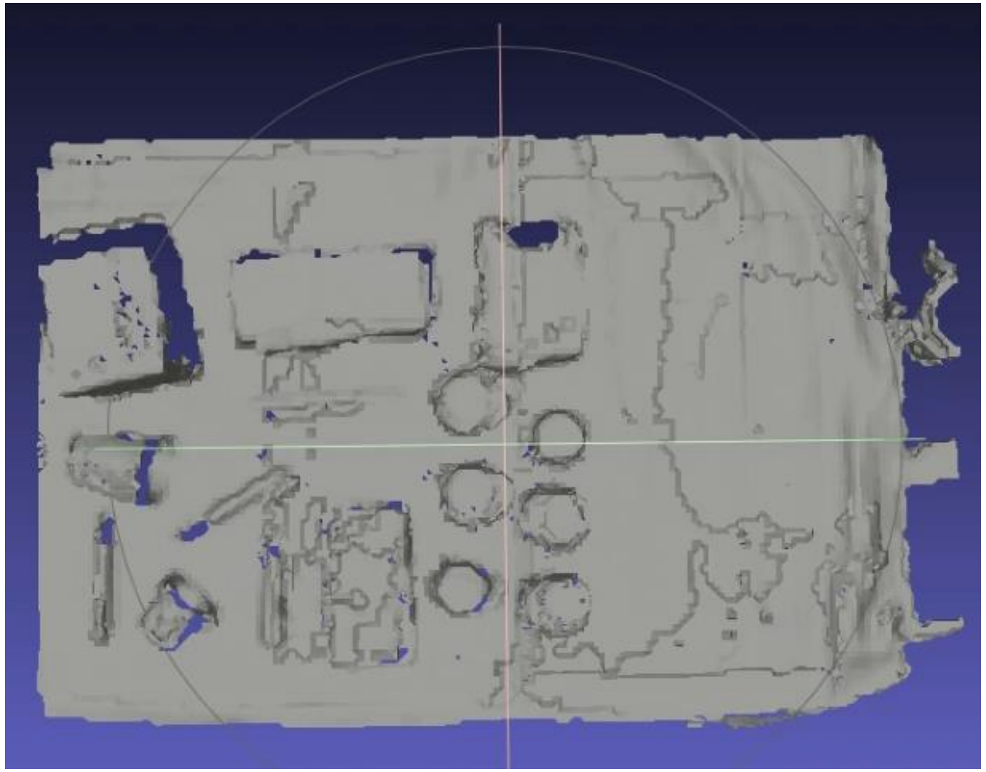
3D Scan



大阪大学 OSAKA UNIVERSITY

HARADA LAB Robotic Manipulation

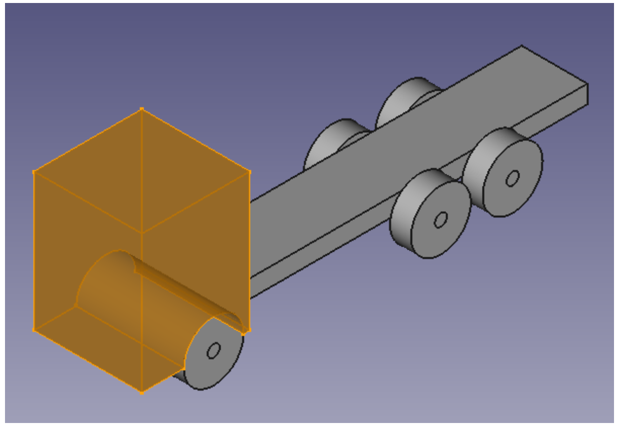
3D Scan



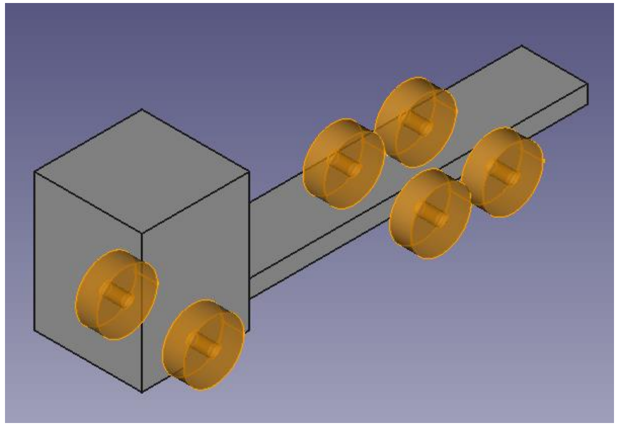
大阪大学 OSAKA UNIVERSITY

HARADA LAB Robotic Manipulation

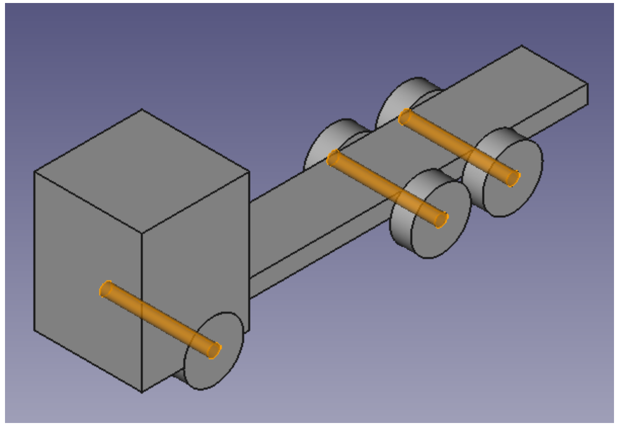
3D Scan → 3D CAD Model



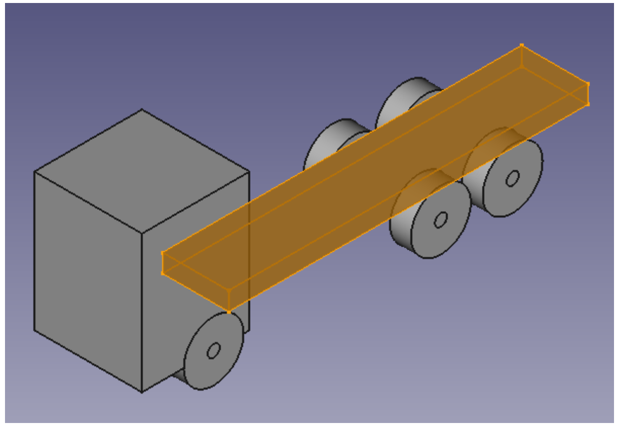
Part: Cabin
Shape: Cuboid
Dimensions: 18 x 18 x 23 cm
Req. affordances: contain; support



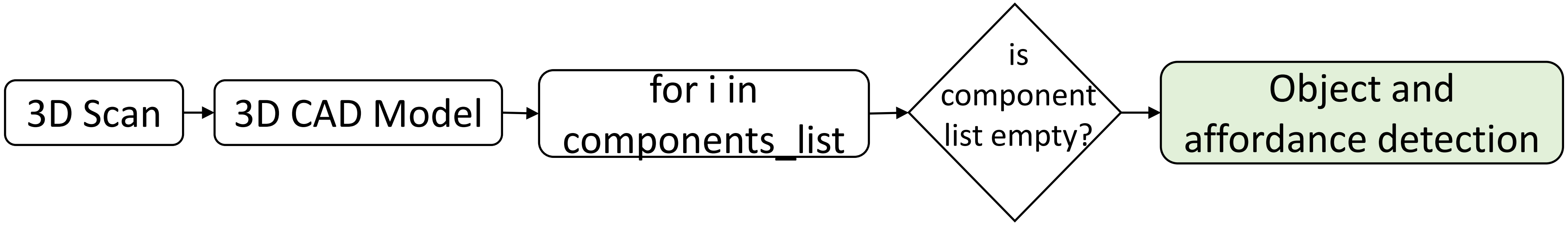
Part: Wheel
Shape: Cylinder
Dimensions (RxH): 5 x 3.5 cm
Req. affordances: rollable

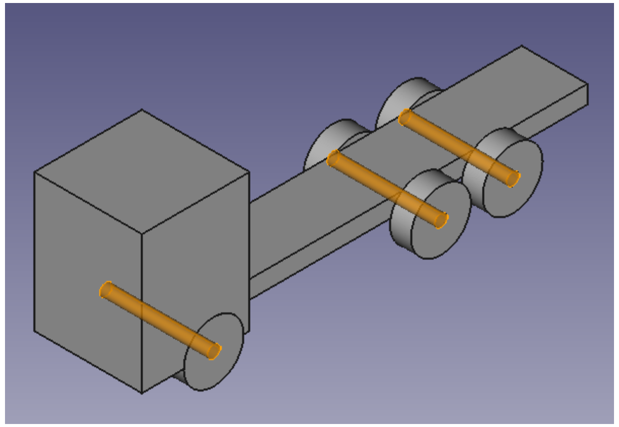


Part: Axle
Shape: Cylinder
Dimensions (RxH): 1 x 18 cm
Req. affordances: rollable

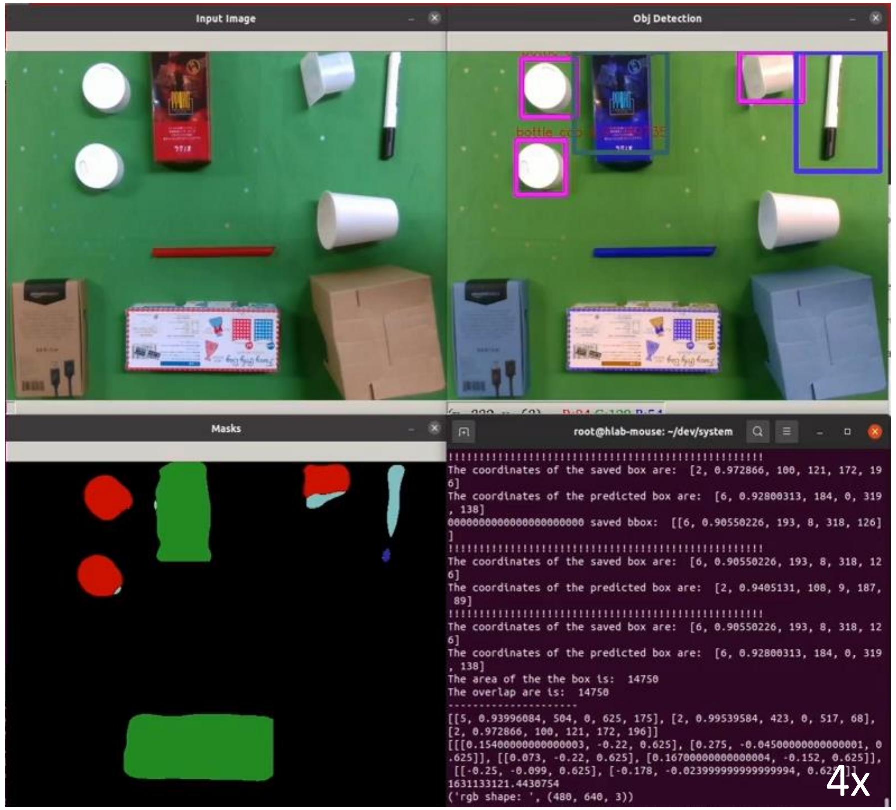


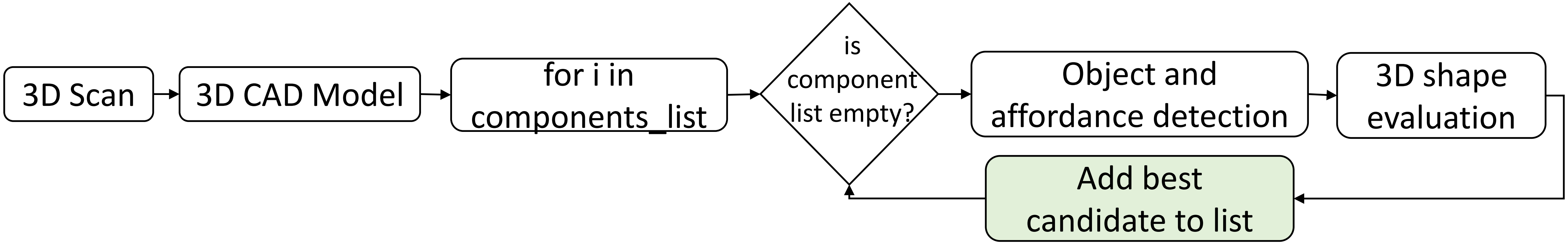
Part: Chassis
Shape: Cuboid
Dimensions: 50 x 11 x 3 cm
Req. affordances: support

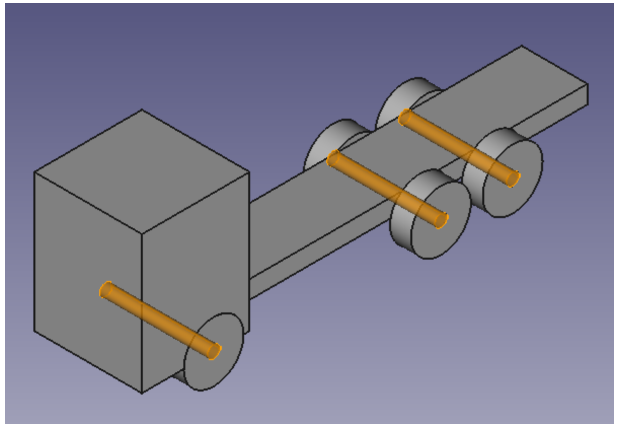




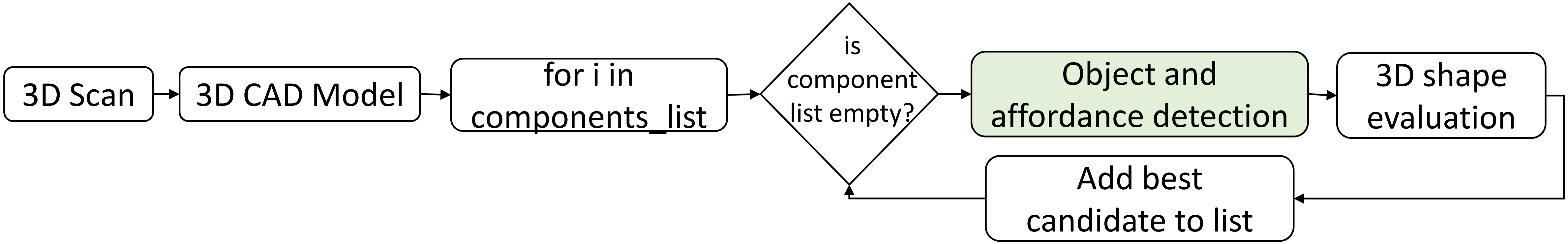
Part: Axle
Shape: Cylinder
Dimensions (RxH): 1 x 18 cm
Req. affordances: rollable

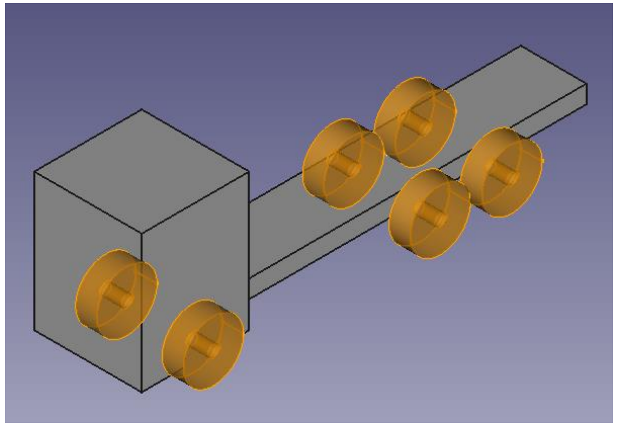


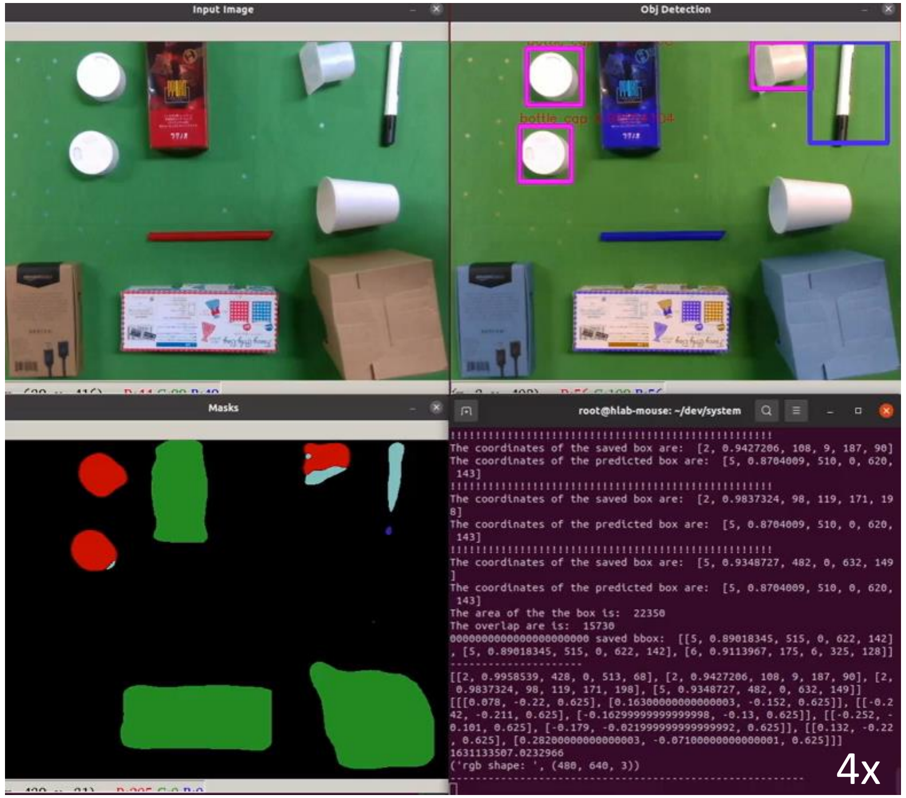




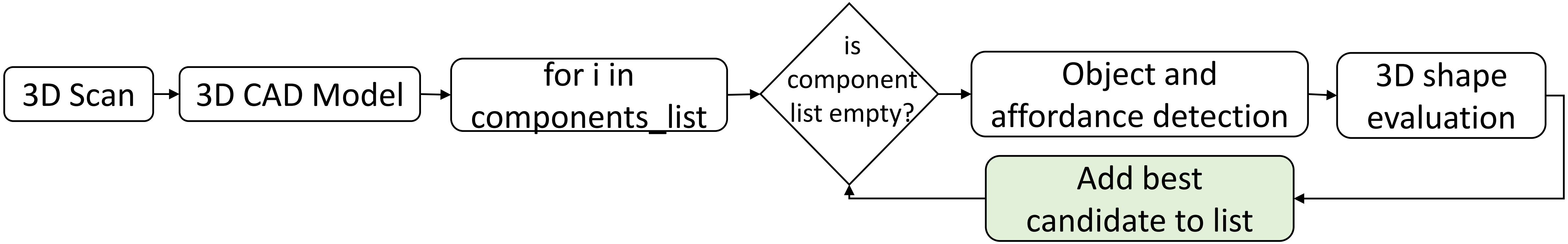
Part: Axle
Shape: Cylinder
Dimensions (RxH): 1 x 18 cm
Req. affordances: rollable

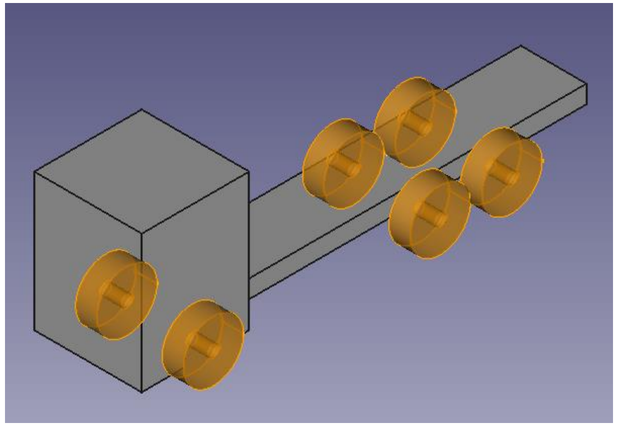




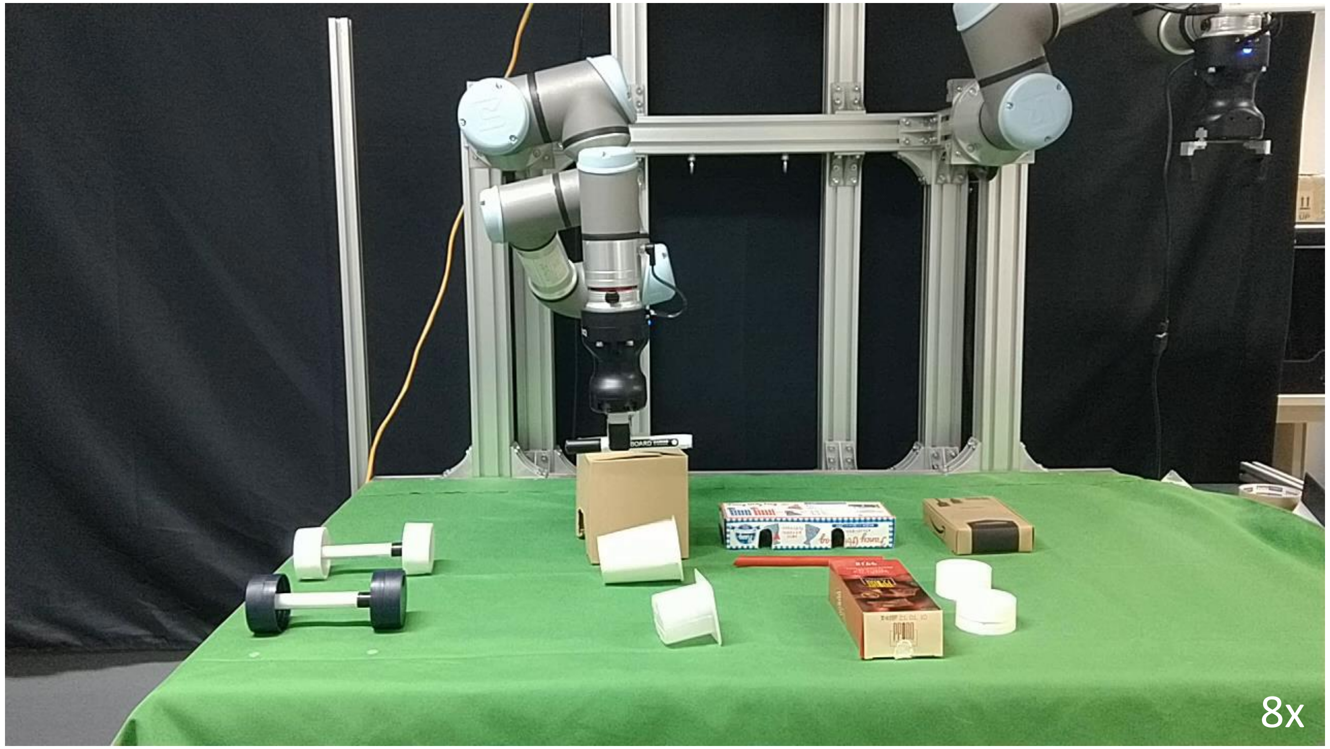


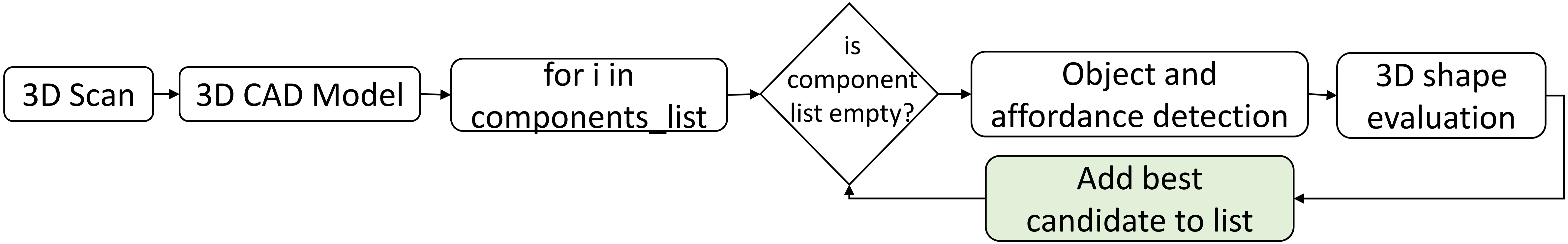
Part: Wheel
Shape: Cylinder
Dimensions (RxH): 5 x 3.5 cm
Req. affordances: rollable

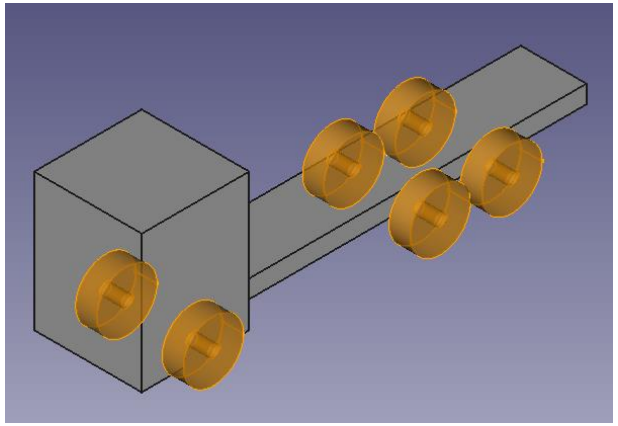




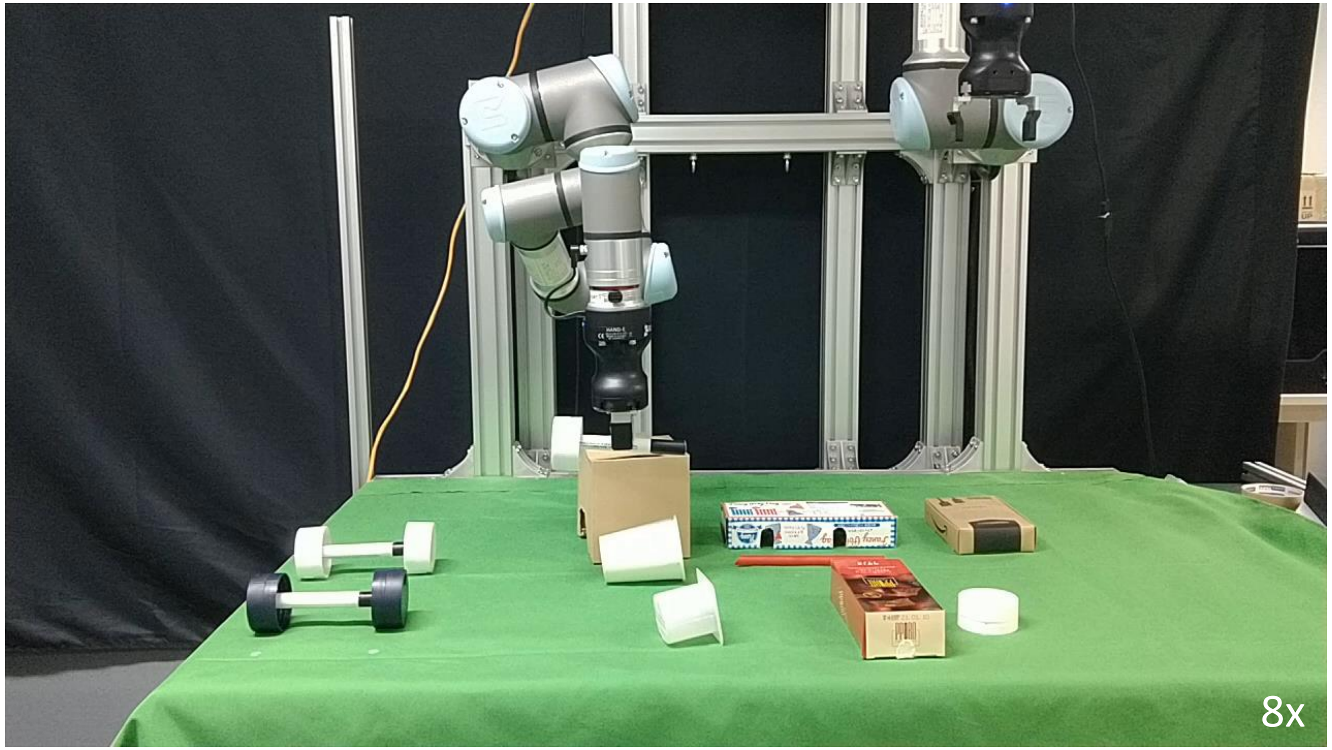
Part: Wheel
Shape: Cylinder
Dimensions (RxH): 5 x 3.5 cm
Req. affordances: rollable





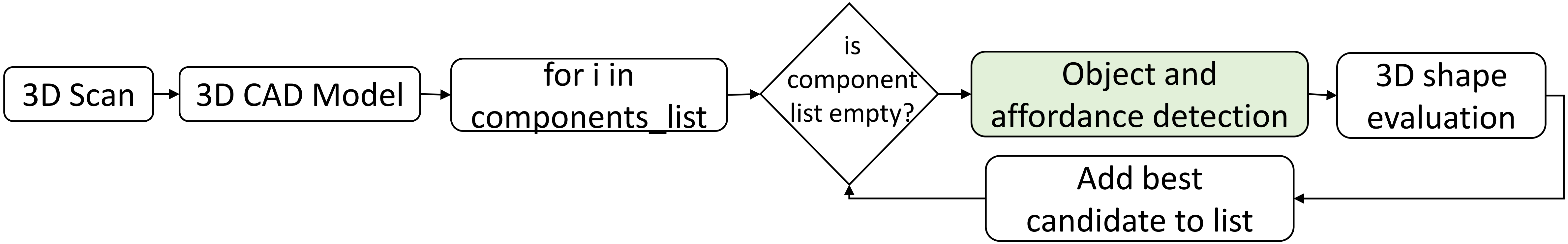


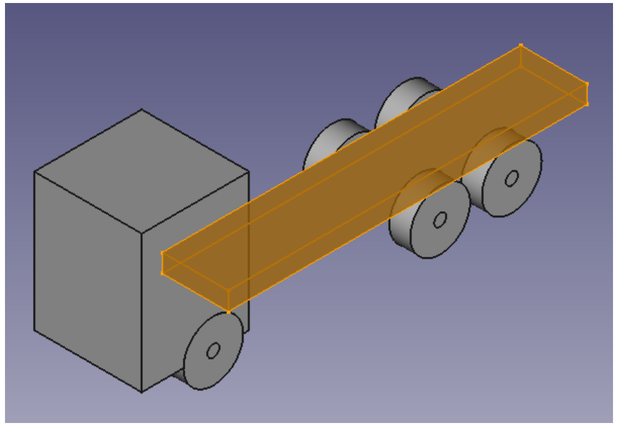
Part: Wheel
Shape: Cylinder
Dimensions (RxH): 5 x 3.5 cm
Req. affordances: rollable



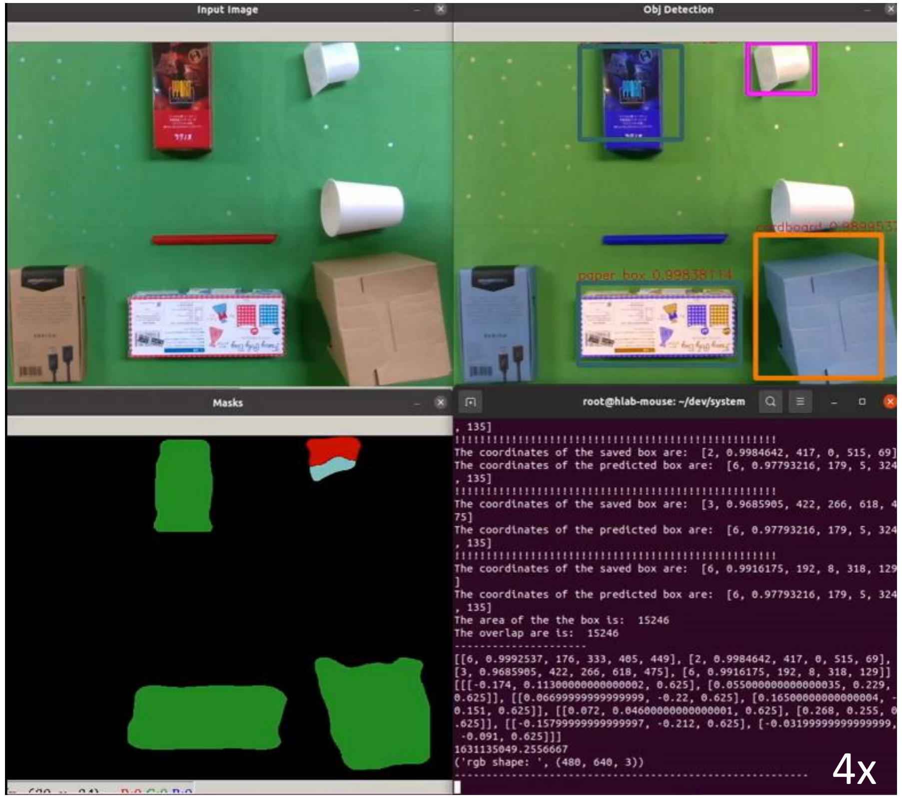
大阪大学 OSAKA UNIVERSITY

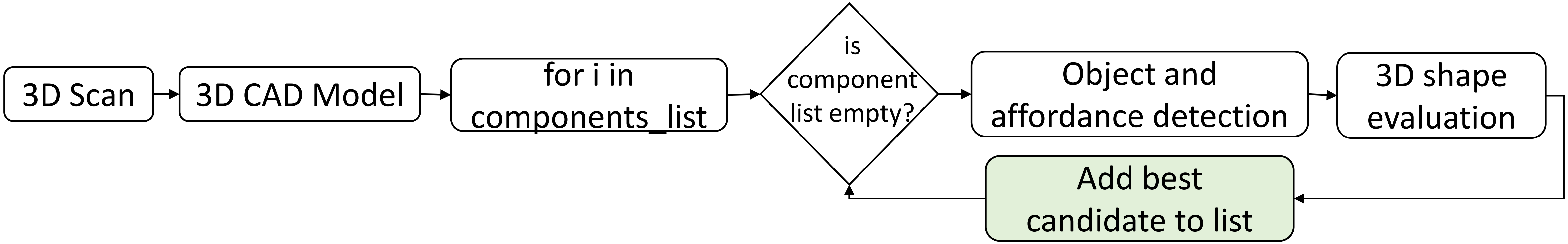
HARADA LAB Robotic Manipulation

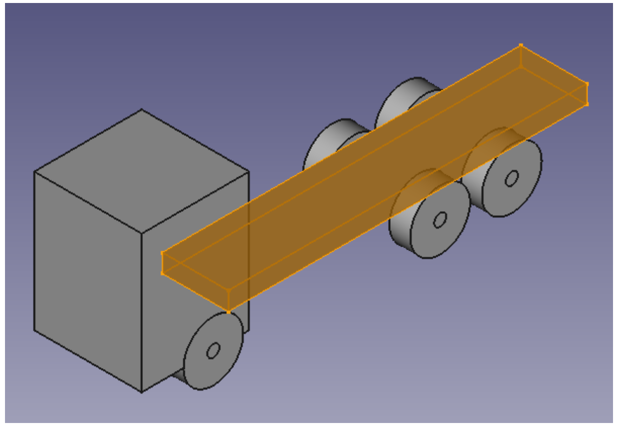




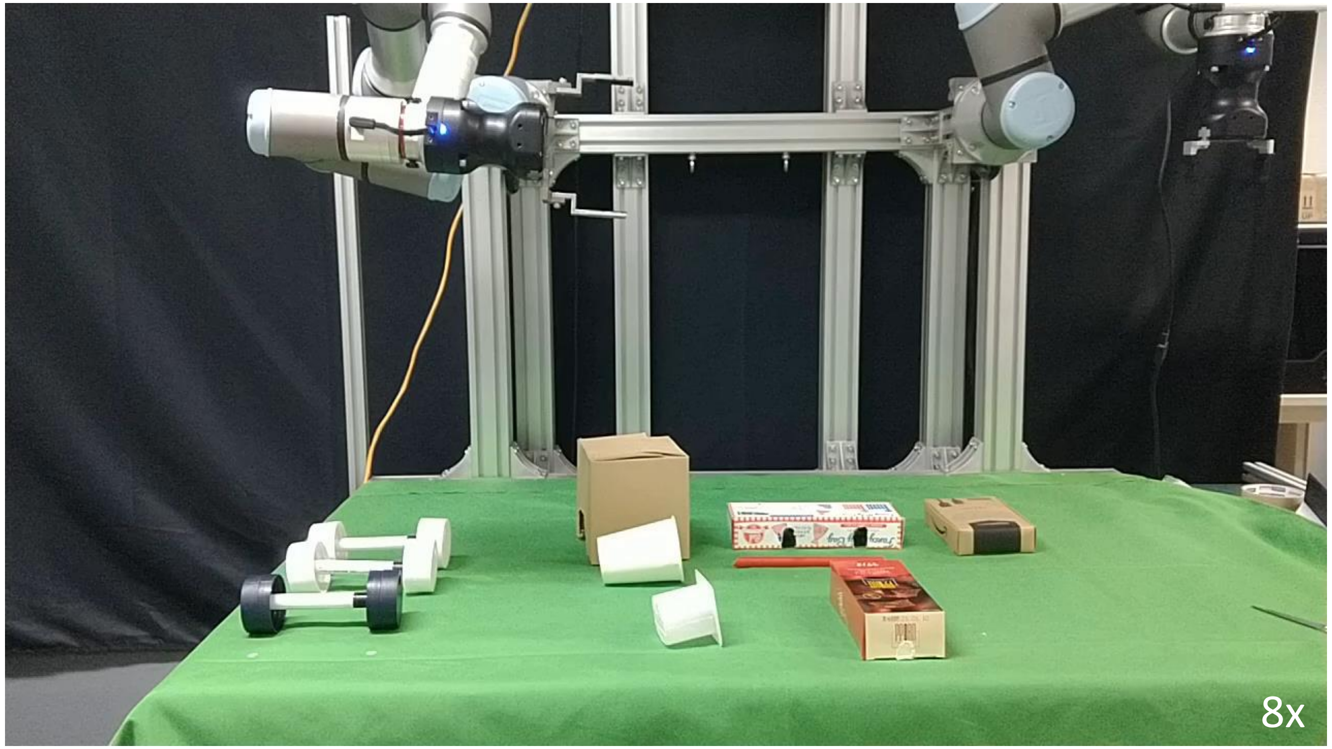
Part: Chassis
Shape: Cuboid
Dimensions: 50 x 11 x 3 cm
Req. affordances: support

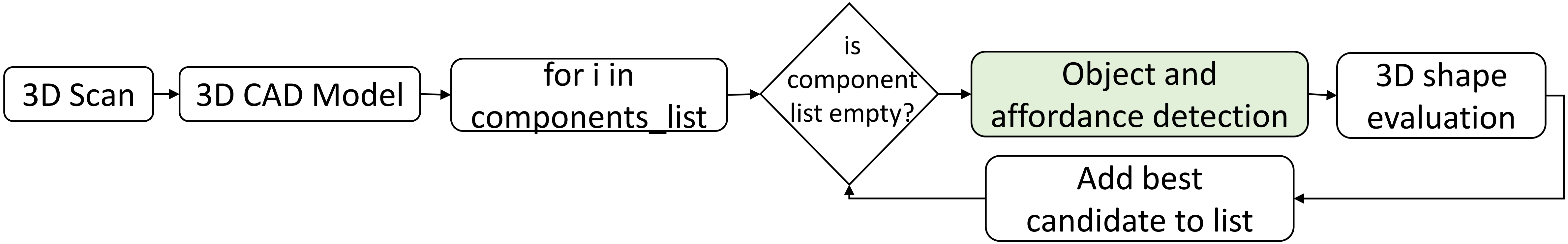


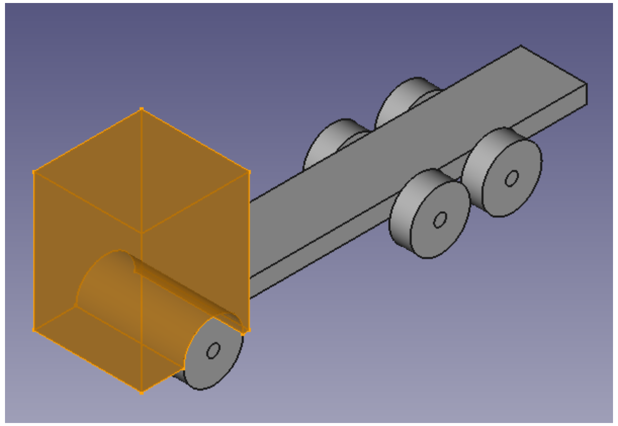




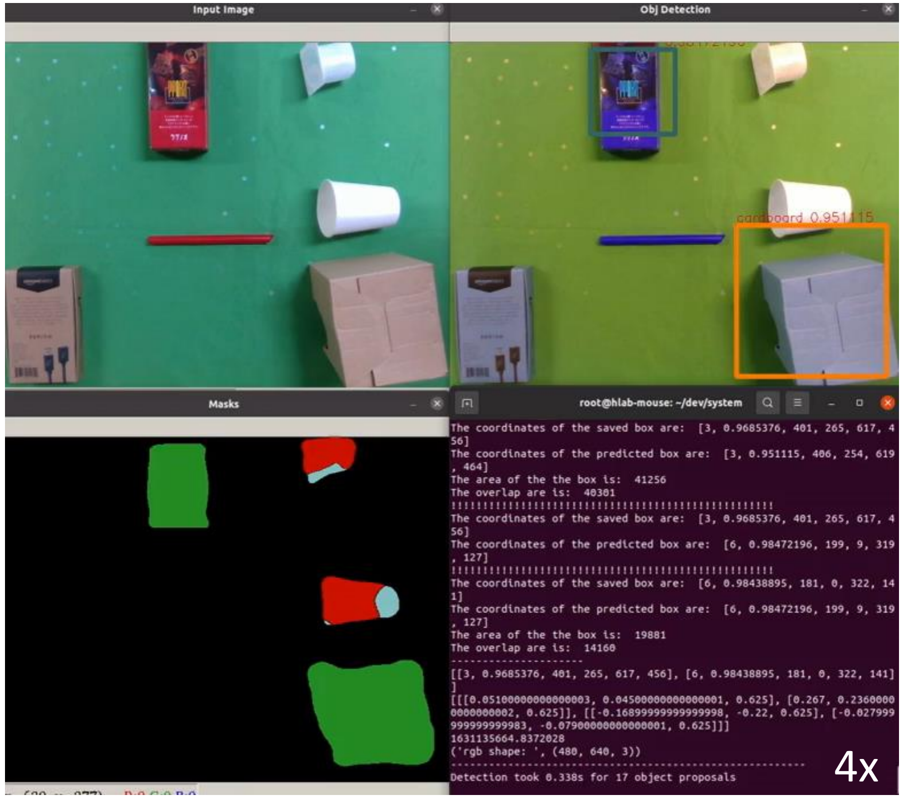
Part: Chassis
Shape: Cuboid
Dimensions: 50 x 11 x 3 cm
Req. affordances: support

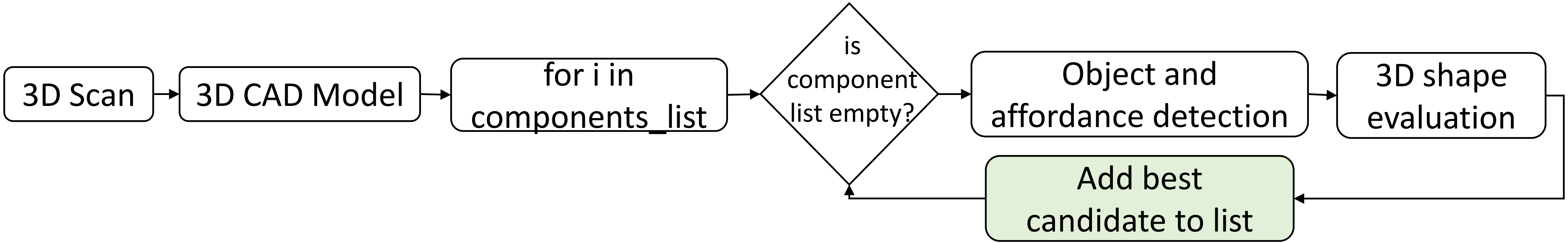


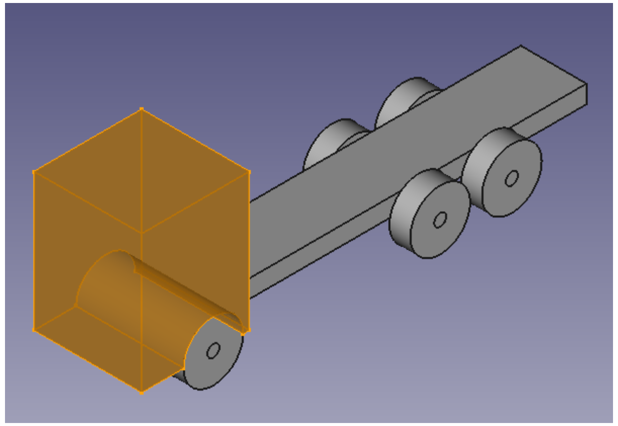




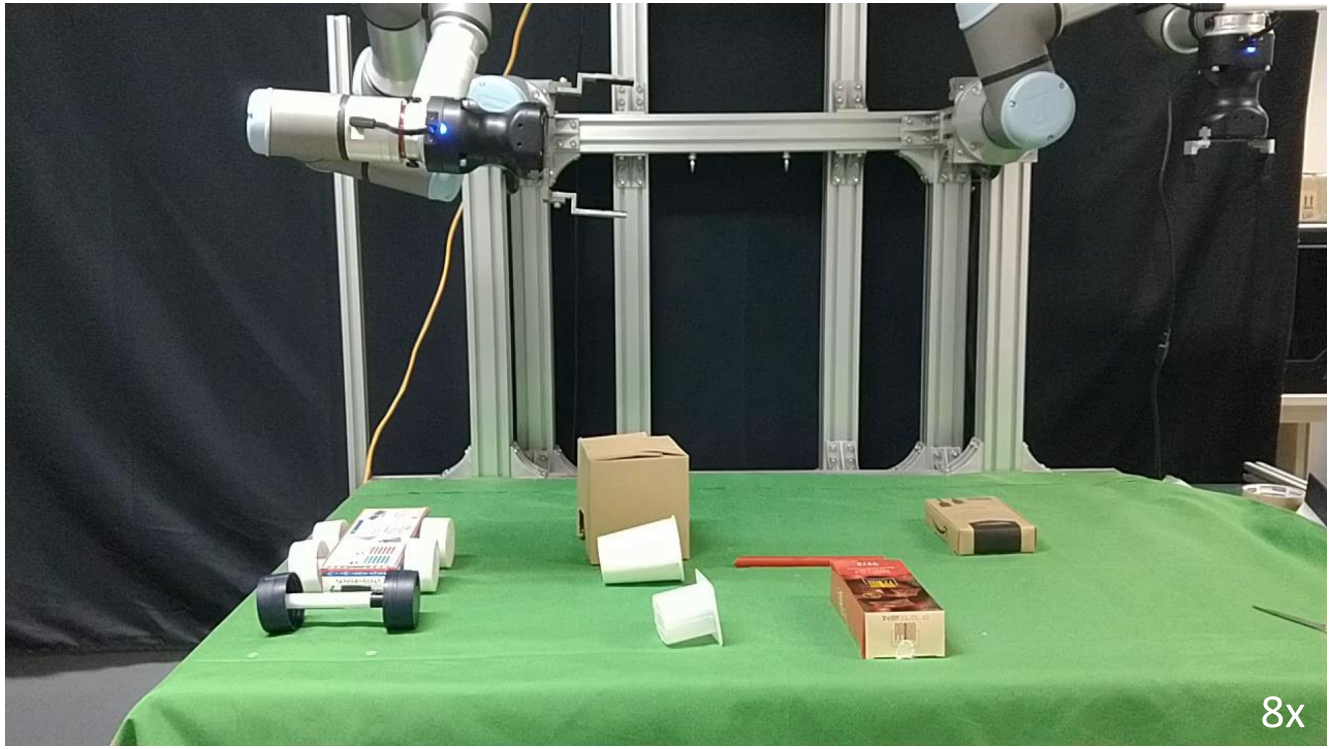
Part: Cabin
Shape: Cuboid
Dimensions: 18 x 18 x 23 cm
Req. affordances: contain; support





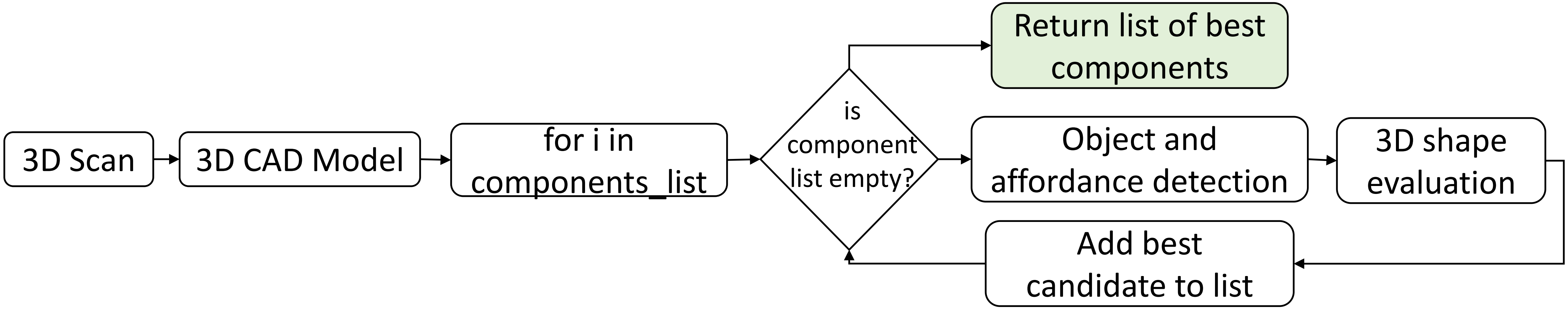


Part: Cabin
Shape: Cuboid
Dimensions: 18 x 18 x 23 cm
Req. affordances: contain; support



大阪大学
OSAKA UNIVERSITY

HARADA LAB
Robotic Manipulation



Conclusions

- We introduced the Craft Assembly Task and proposed a system to solve the first step: selecting the materials for the craft
- A database with affordance masks for common, everyday materials was built
- RANSAC shows some errors in dimension estimation, which might be propagated in future steps
- <u>Future Work</u>
- Add remaining steps of the Craft Assembly Task
- Perform the experiment for building the entire toy truck

大阪大学 OSAKA UNIVERSITY

HARADA LAB Robotic Manipulation

Thank you for listening!