Adaptive Multi-layered Perception for Picking in Clutter with Parameter Reinforce Learning via Experience in Task

The University of Tokyo
wada@jsk.imi.i.u-tokyo.ac.jp
◯ Kentaro Wada, Kei Okada, Masayuki Inaba

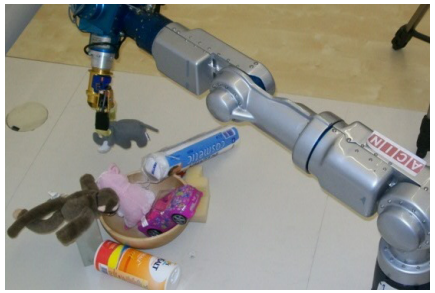
Introduction

- Picking Task
  1. Various Objects
  2. Clutter Environment
- Issue

  Various object/obstacle state

Necessity of adaptive perception system for these various environmental state.

1. Various Objects



David F. 2012, Shape based Learning for Grasping Novel Objects in Cluttered Scenes

Amazon Picking Challenge 2015

2. Clutter Environment

Kei Okada. 2007, Multi-cue 3D Object Recognition in Knowledge-based Vision-guided Humanoid Robot System

Ashutosh Saxena, 2007, A Vision-based System for Grasping Novel Objects in Cluttered Environments

Picking System with Adaptive Perception

Picking system

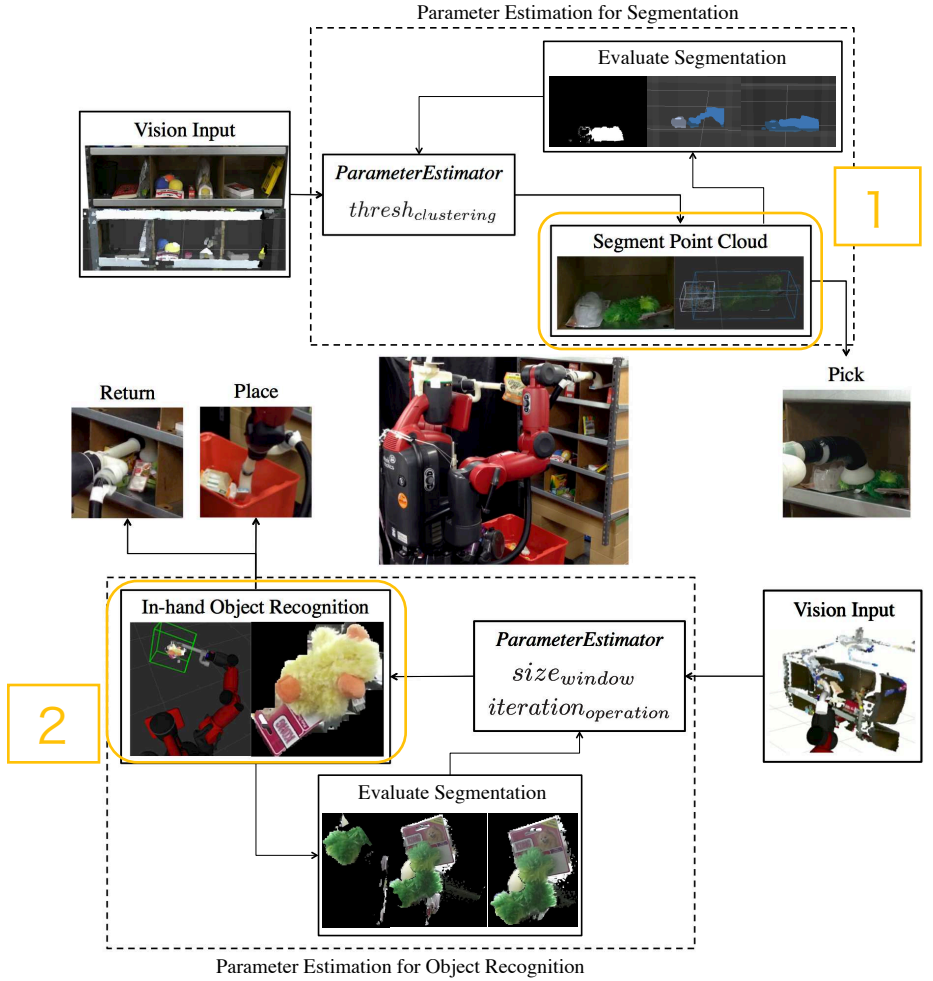
1. Segment based on point cloud
2. Pick
3. In-hand object recognition
4. Place or Return

Picking System with Adaptive Perception

Picking system

1. Segment based on point cloud
2. Pick
3. In-hand object recognition
4. Place or Return

Adaptive perception system

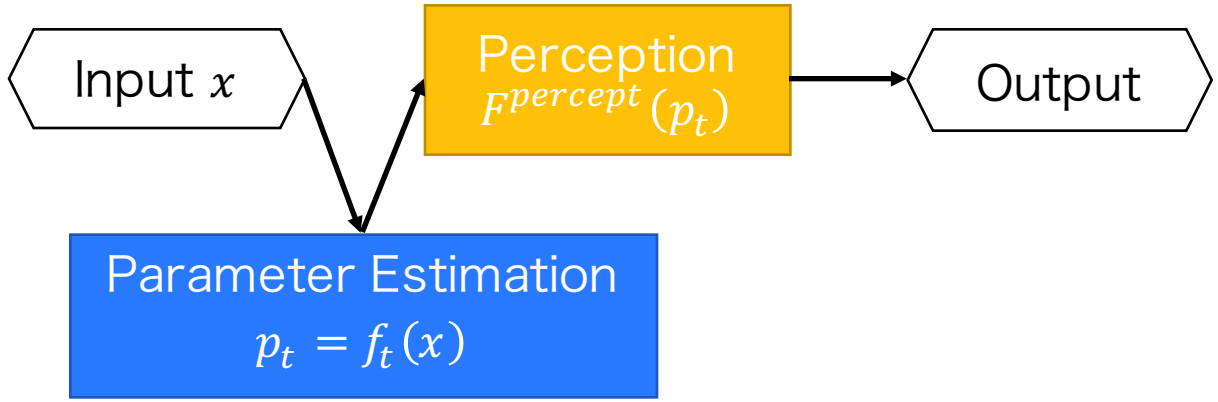


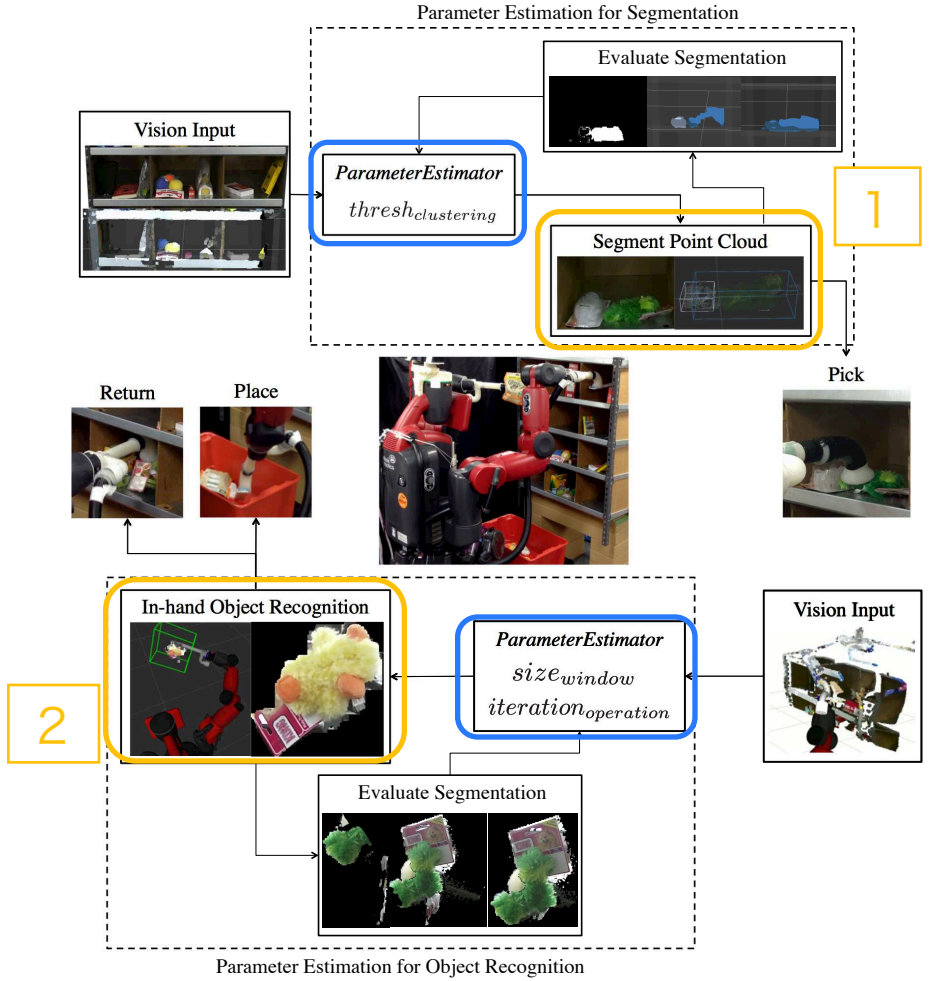
Picking System with Adaptive Perception

Picking system

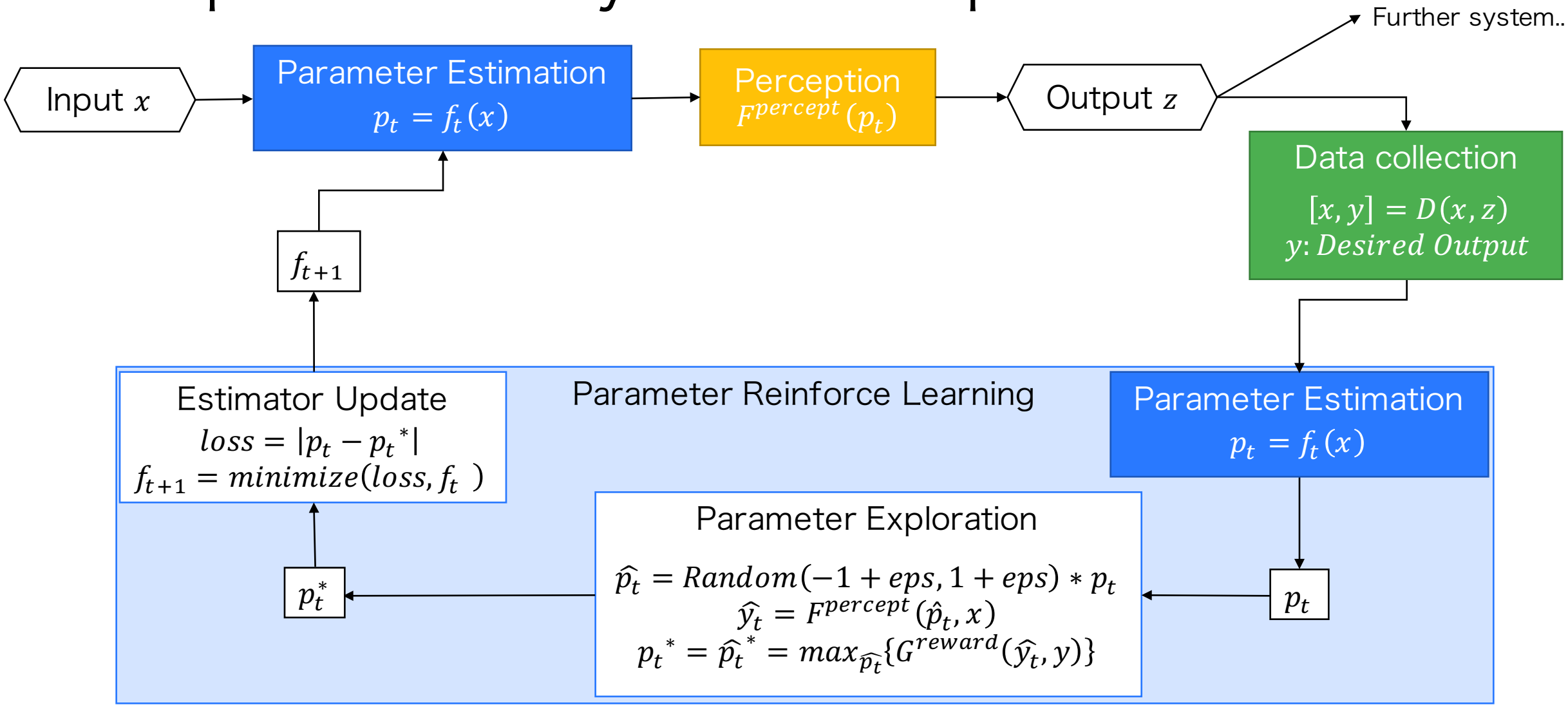
1. Segment based on point cloud
2. Pick
3. In-hand object recognition
4. Place or Return

Adaptive perception system



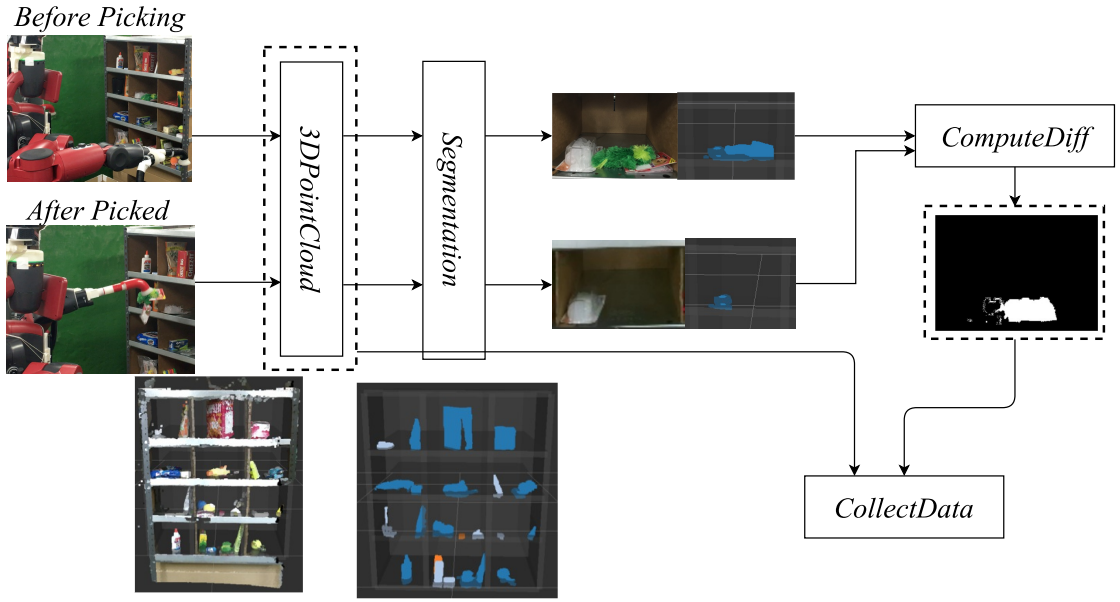


Parameter Reinforce Learning for Adaptive Multi-layered Perception

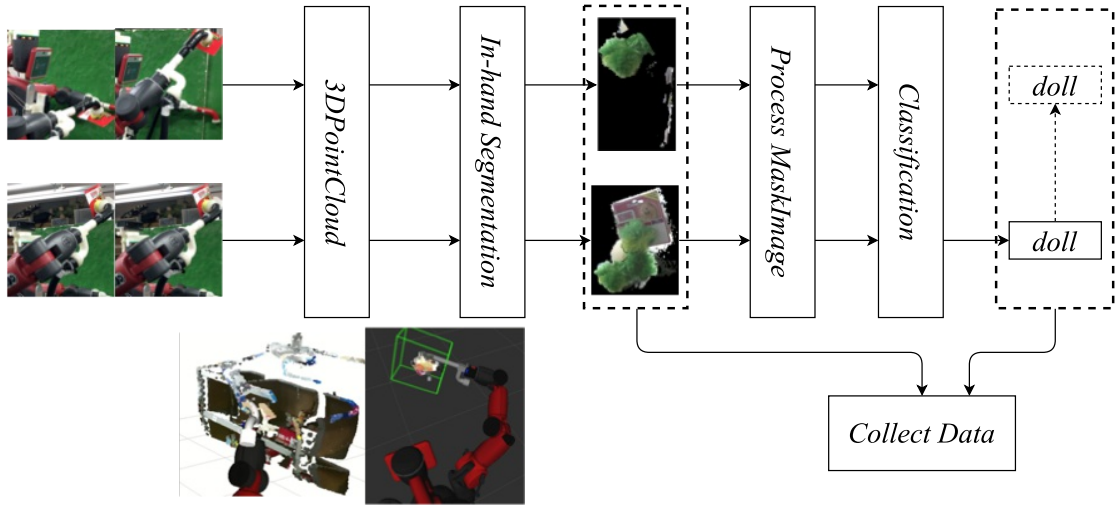


Data Collection

1. Object segmentation
   - Object can be segmented by picking.
   - Desired segmentation output is collectable by comparing before/after picking.
2. Target object verification
   - Recognition result while hand is stopping is reliable.
   - Desired recognition result is collectable by stopping the hand.



1. Data collection for object segmentation



2. Data collection for target object verification

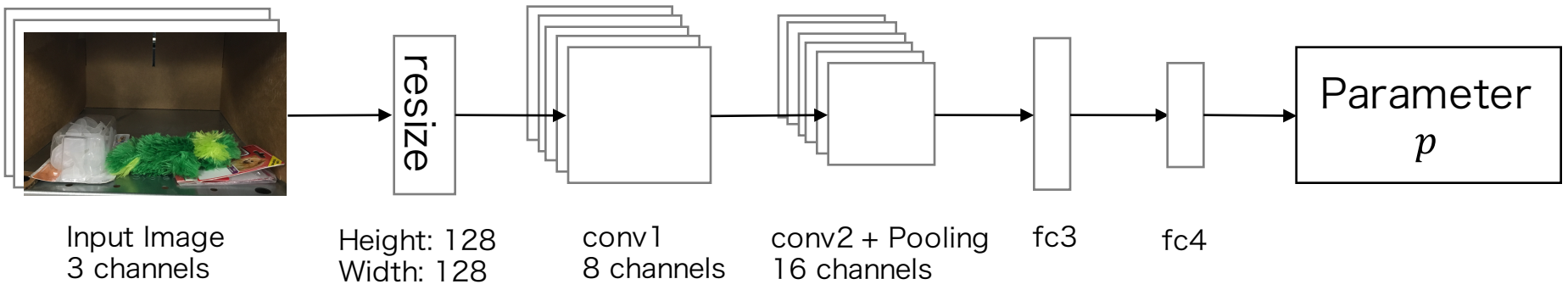
Parameter Estimation Network: *PNet*

Network structure

- 2 Convolutional Network Layers + 2 Fully Connected Layers.
- Input image is resized before the first layer.
- Initial parameter should be assigned by human, so it learns the scale to it.

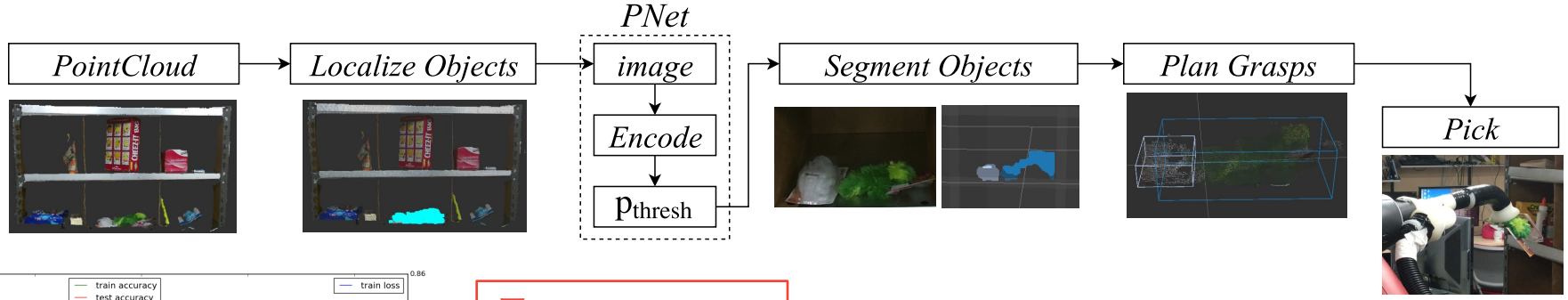
Pretraining with convolutional auto-encoder

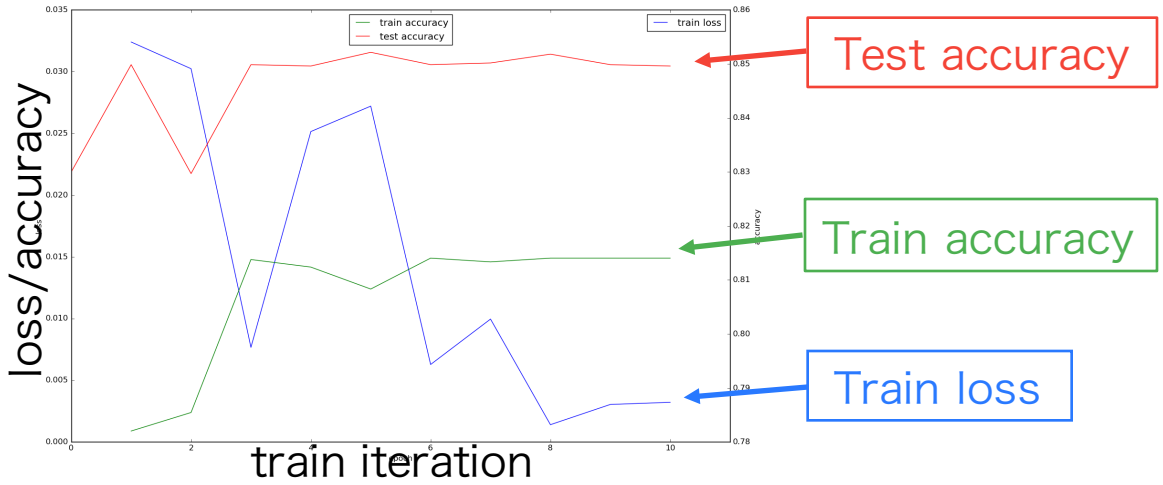
- Convolutional auto-encoder with deconvolution + unpooling.
- Encoding layers is used for estimation.
- The encoded units are initialized as ones vector.



Experiment

1. Object segmentation
   - Localize object point cloud by location
   - Parameter estimation from depth (x) : $p^{thresh} = f(x)$
   - Euclidean clustering segmentation : $clusters = EuclidClustering\left(PointCloud, p^{thresh}\right)$
   - Plan Grasp Coordinates



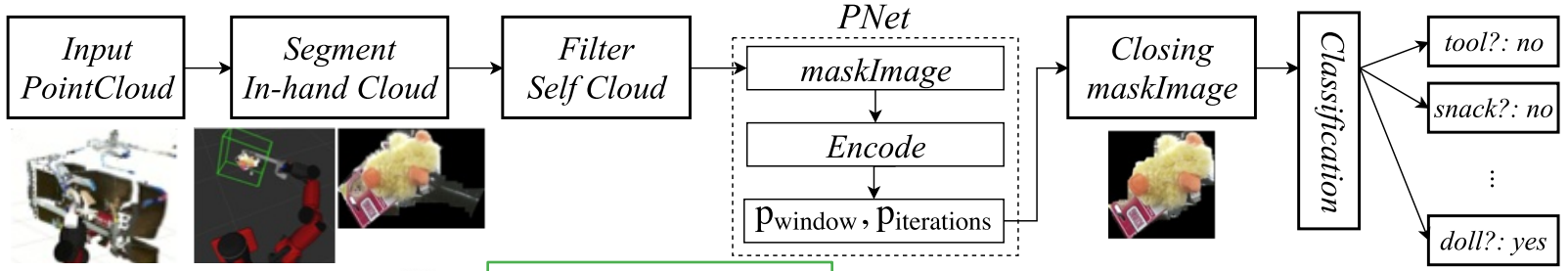


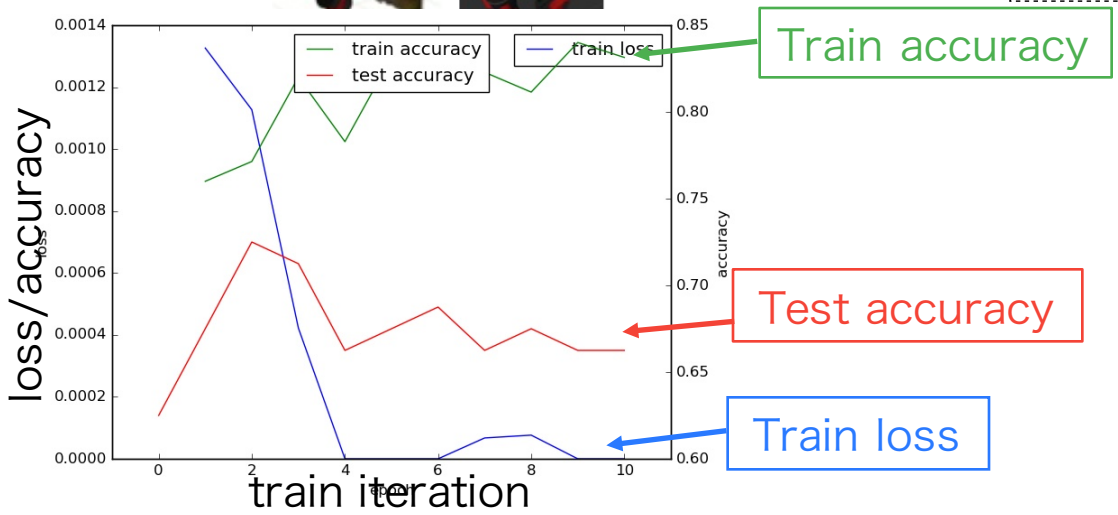
	Accuracy iter=0	Accuracy iter=10
train	78.5%	*81.5%*
test	82%	*85%*

Training result with 5 collected data

Experiment

2. Target object verification
   - In-hand point cloud segmentation with self filtering
   - Parameter estimation from mask image (x) : $p^{thresh} = f(x)$
   - Closing mask image : $mask^{dst} = Closing\left(mask^{src}, p^{iteration}, p^{window}\right)$
   - Classify the mask applied image





	Accuracy iter=0	Accuracy iter=10
train	76%	*83%*
test	62.5%	*66.3%*

Training result with 130 collected data

Experiment

Conclusion

- We propose a learning-based method to introduce adaptive perception into robotic system.
- The parameter estimation network with data-collection system can be used for adaptive vision-based perception in general.
- We applied the proposed method to 2 perceptions in a picking system and got **3 - 4 %** accuracy rises.