Manipulation

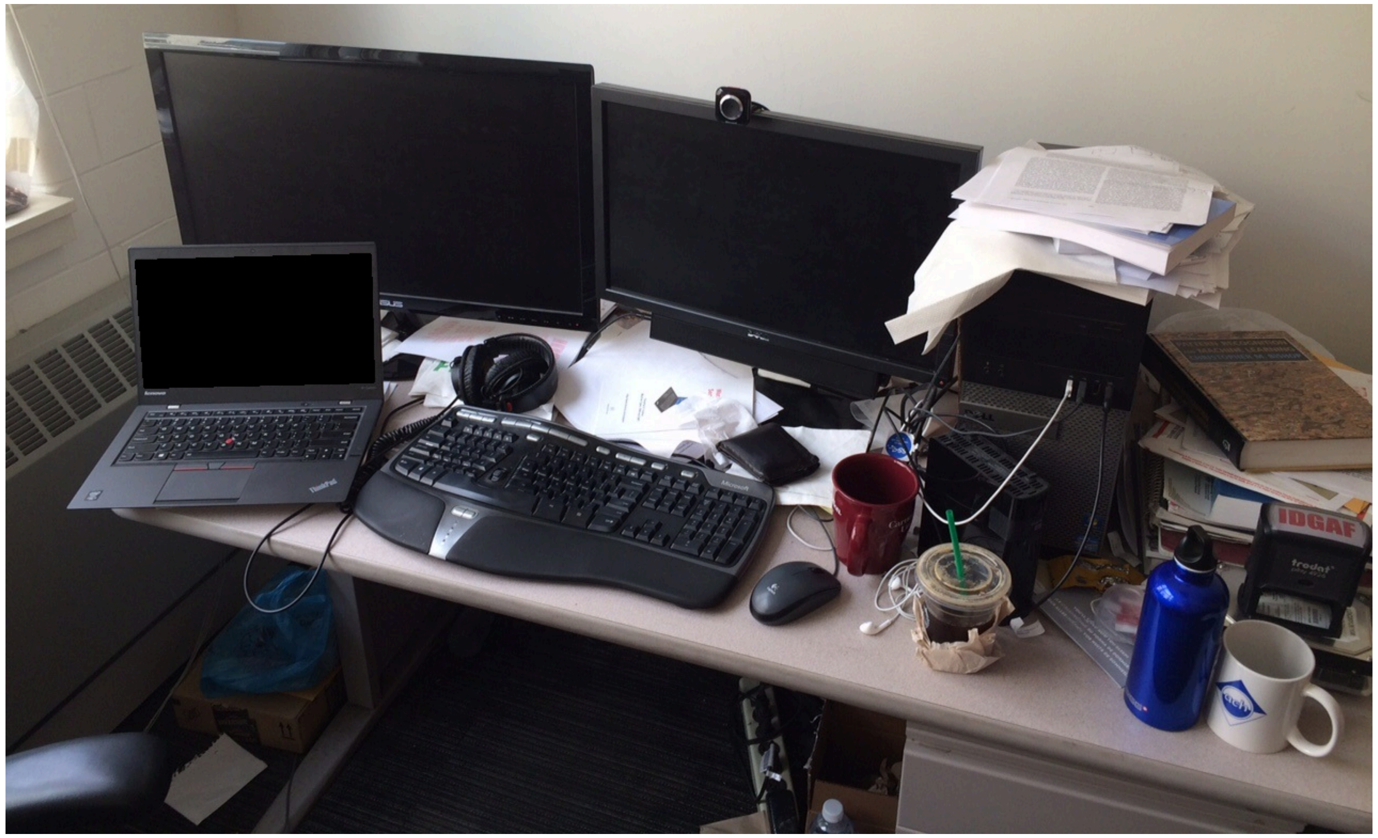
Saurabh Gupta

Today's Class

- Task, and challenges in manipulation
- Discussion about the role of learning for manipulation tasks
- DexNet 2.0 Paper

Example Tasks

- Tidying up a table
- Folding laundry
- Taking out keys from pocket
- Inserting key into lock
- Cutting a potato
- Scrubbing a dish



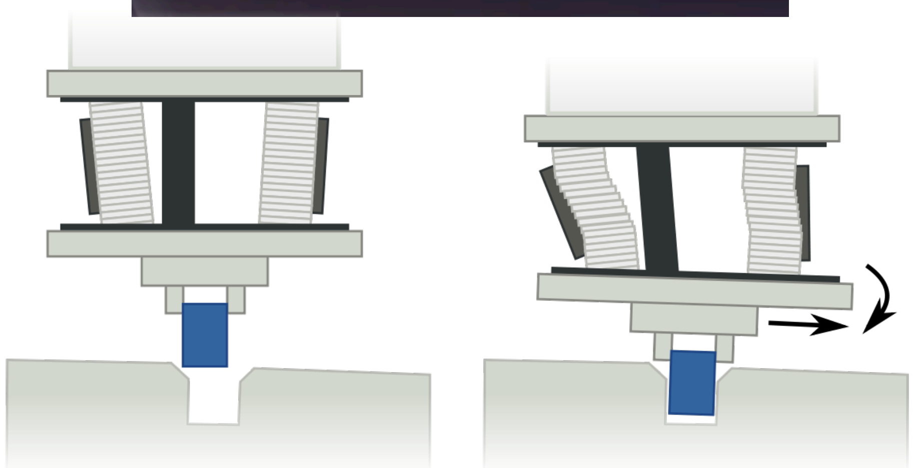
Tasks

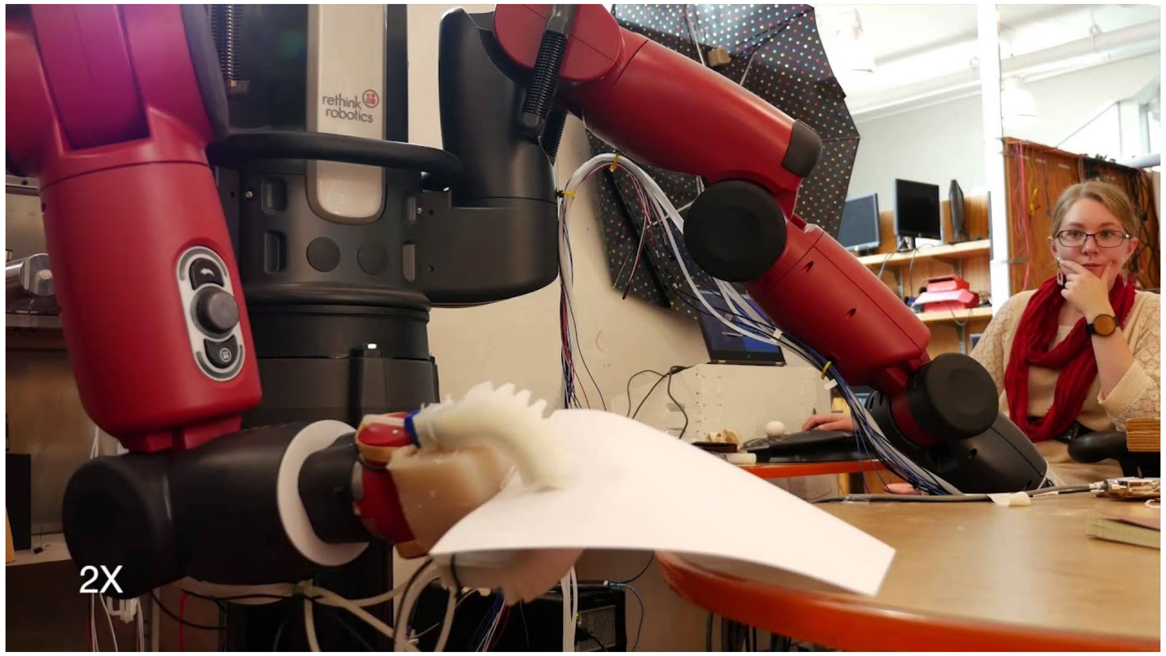
- Programmed motion
- Compliant motion
- Structured pick-and-place
- Unstructured pick-and-place
- Mechanical assembly and task mechanics
- In-hand manipulation
- Non-prehensile manipulation
- Whole body manipulation
- Task-oriented grasping
- Manipulation of deformable objects
    - cloth, granular media

Towards Robotic Manipulation. Matthew Mason 2018.

Actuation / End effector design

- Parallel jaw grippers
- Task specific end-effectors
  - Eg: Suction cups, remote Center Compliance for peg insertion
- Multi-finger hands
- Soft robots





Sensing

- Over head camera / hand-in-eye camera, etc.

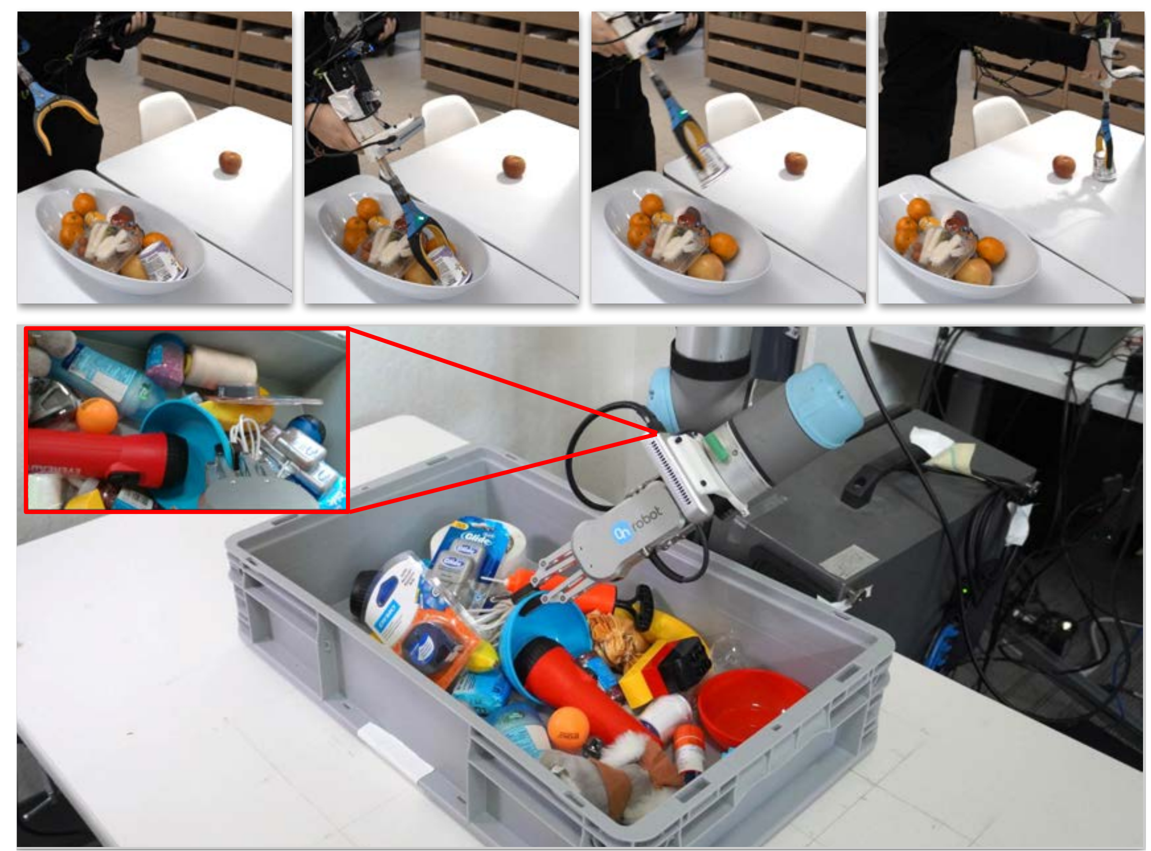


Image Source: L. Pinto, S. Song

Sensing

- Over head camera / hand-in-eye camera, etc.
- Just visual sensing may not be enough, eg: consider inserting a key in cold

Video Source: Dr. Roland Johansson

Sensing

- Over head camera / hand-in-eye camera, etc.
- Just visual sensing may not be enough, eg: consider inserting a key in cold

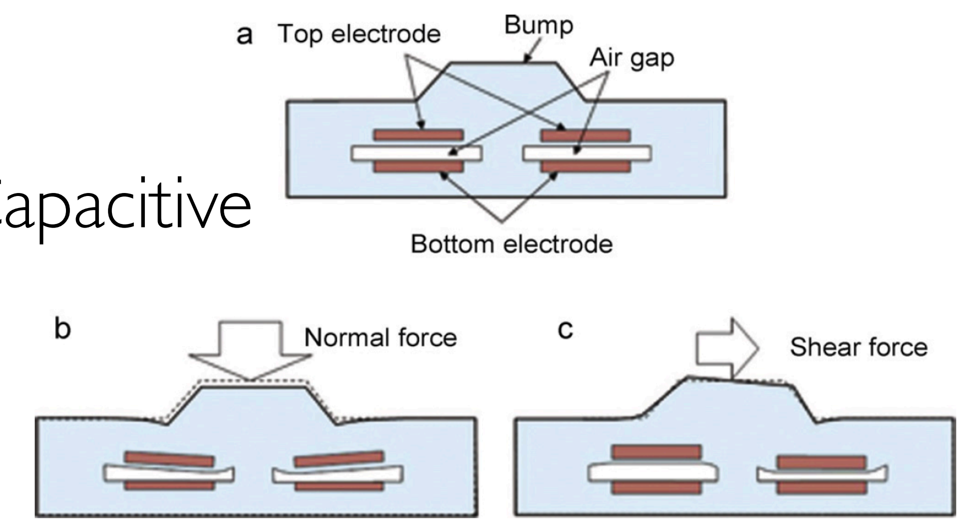
Video Source: Dr. Roland Johansson

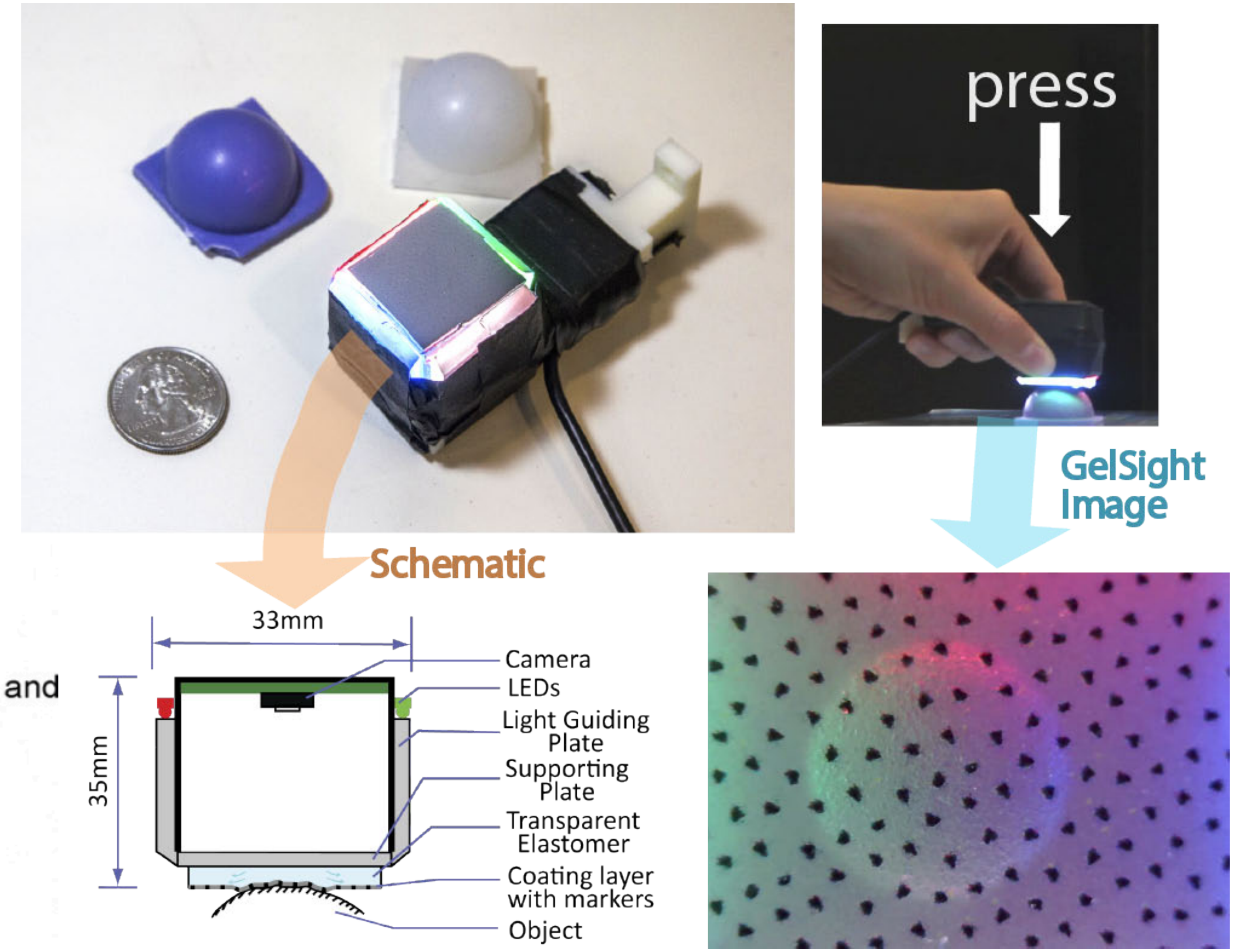
Sensing

- Over head camera / hand-in-eye camera, etc.
- Just visual sensing may not be enough, eg: consider inserting a key in cold
- Tactile sensing / haptic feedback may be crucial

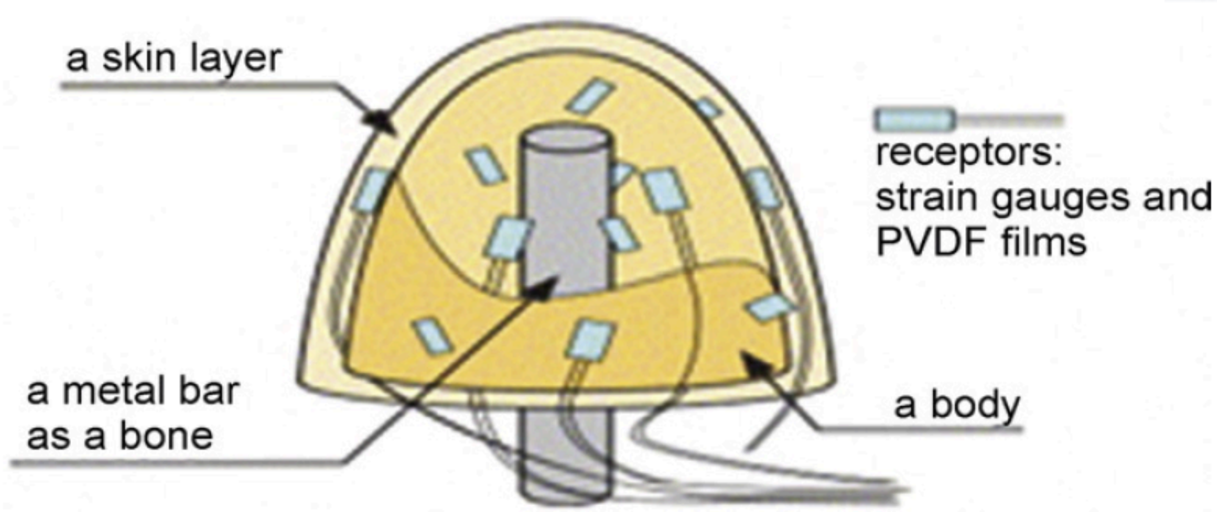
Optical tactile sensors

Capacitive





Piezoelectric

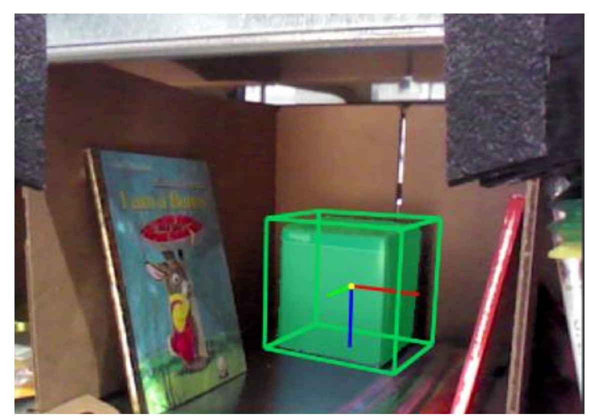


Typical Robotics Pipeline

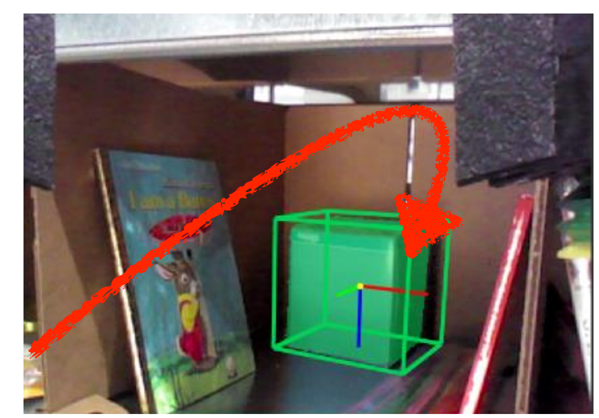
Observations → State Estimation → Planning → Control

Manipulation

Observed Images



6DOF Pose



Grasp Motion Planning

Discussion

- What, if any, are some ways in which classical techniques may fall short for manipulation tasks?
- Would it be possible to fix any of these via machine learning?

Manipulation is hard

- Actuators and sensors are far from mature
- Contact is hard to model
- High-dimensional systems can get hard to control
- Tasks are very varied

Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics

Jeffrey Mahler*, Jacky Liang*, Sherdil Niyaz*, Michael Laskey*, Richard Doan*, Xinyu Liu*,
Juan Aparicio Ojea†, and Ken Goldberg*
*Dept. of EECS, University of California, Berkeley
Email: {jmahler, jackyliang, sniyaz, laskeymd, rdoan, xinyuliu, goldberg}@berkeley.edu
† Siemens Corporation, Corporate Technology
Email: juan.aparicio@siemens.com

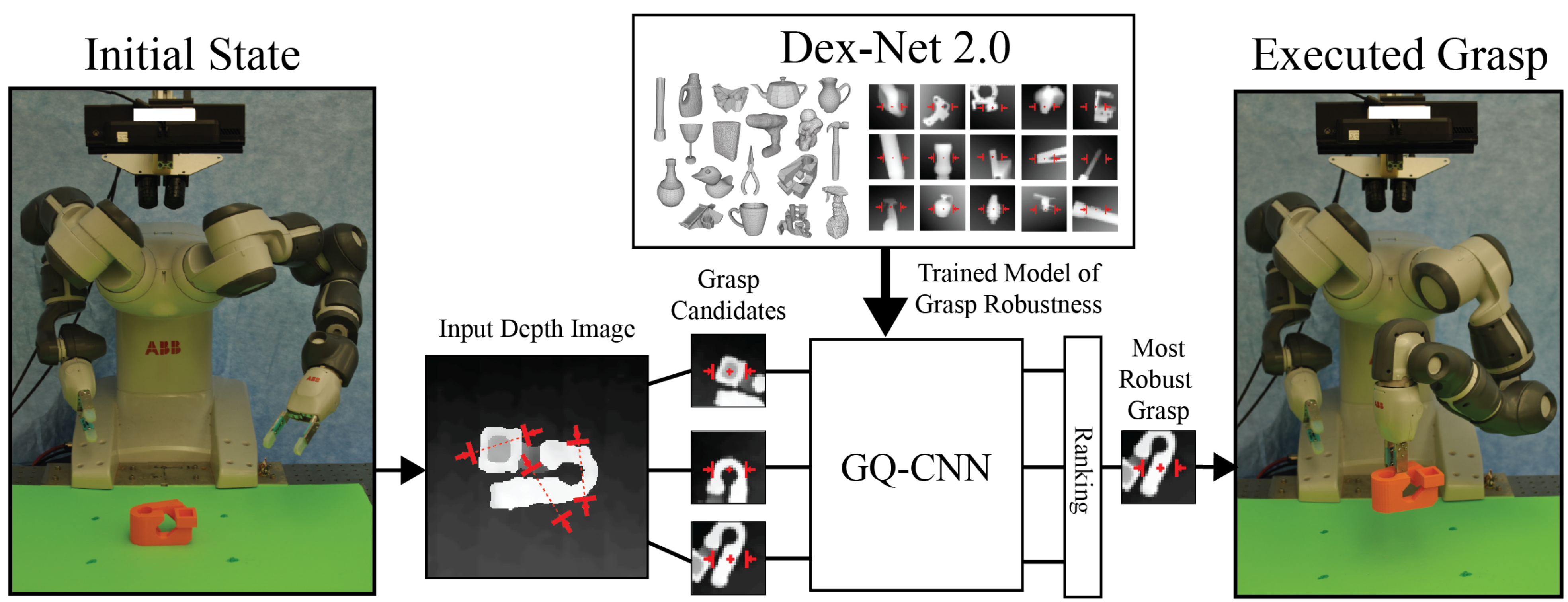
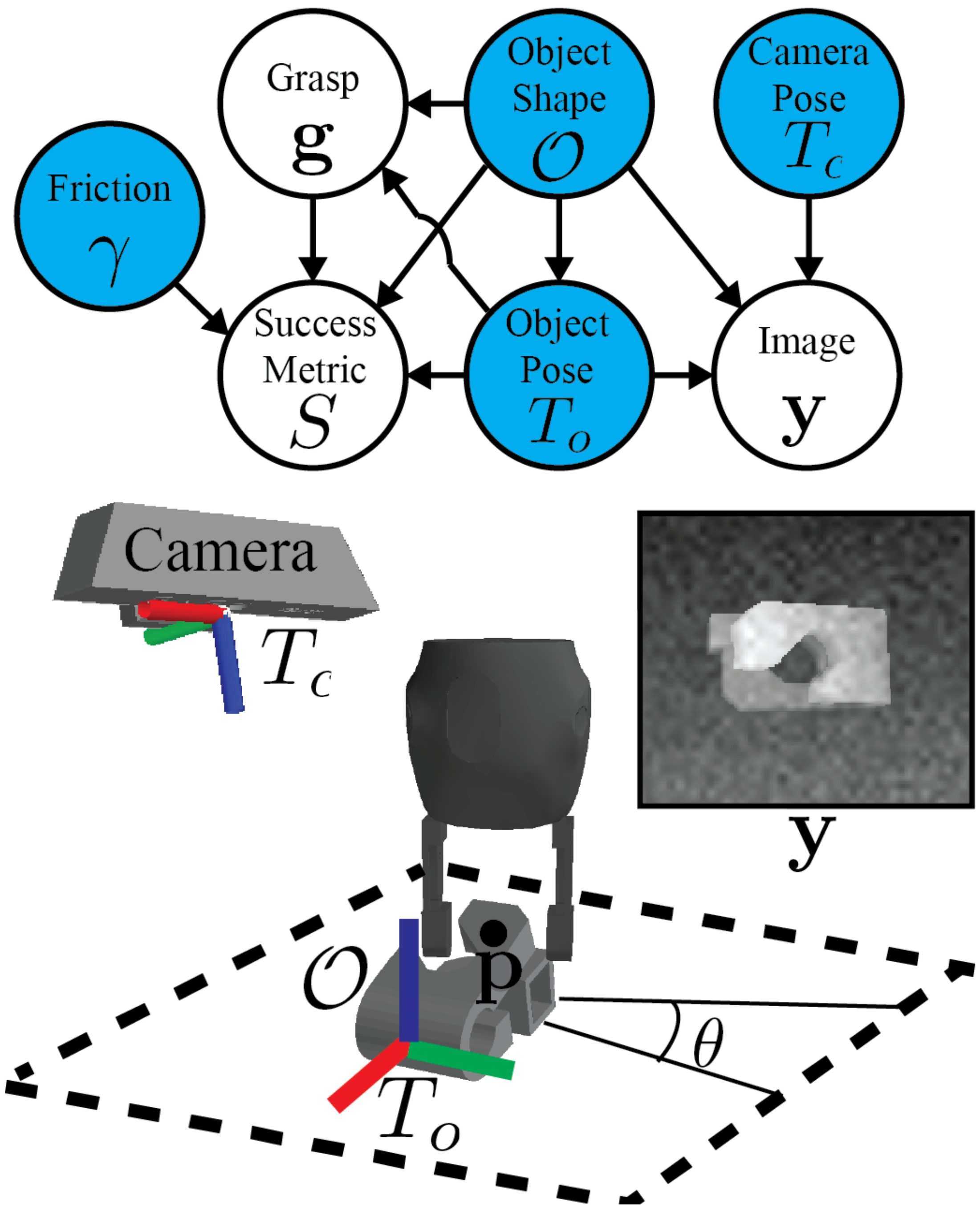


ABB
4x

Data Generation

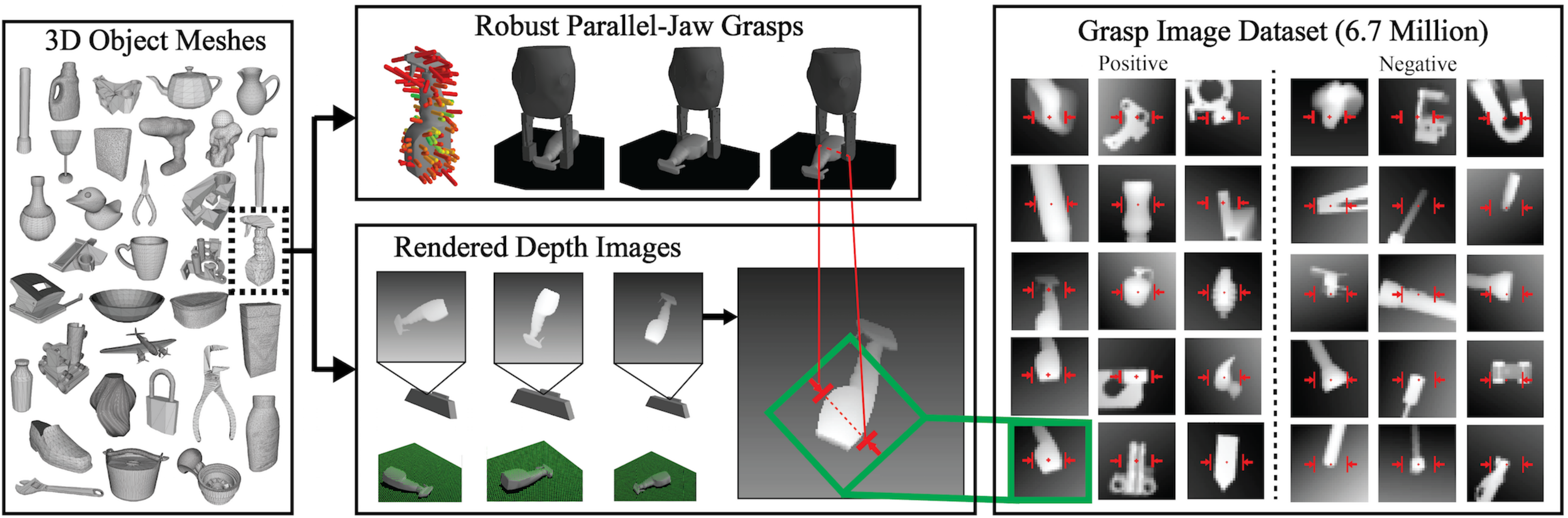


Distribution	Description
$p(\gamma)$	truncated Gaussian distribution over friction coefficients
$p(\mathcal{O})$	discrete uniform distribution over 3D object models
$p(T_o \mid \mathcal{O})$	continuous uniform distribution over the discrete set of object stable poses and planar poses on the table surface
$p(T_c)$	continuous uniform distribution over spherical coordinates for radial bounds $[r_\ell, r_u]$ and polar angle in $[0, \delta]$

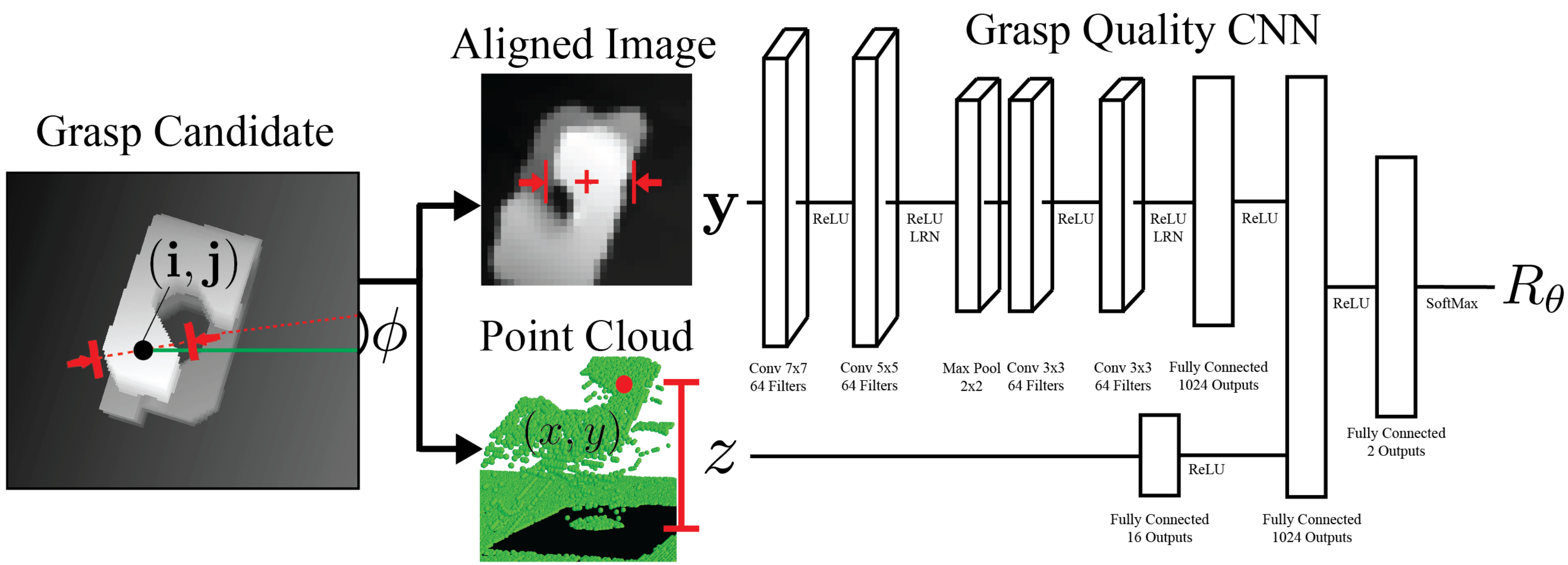
$$\mathbf{y} = \alpha \hat{\mathbf{y}} + \epsilon$$

$$S(\mathbf{u}, \mathbf{x}) = \begin{cases} 1 & E_Q > \delta \text{ and } collfree(\mathbf{u}, \mathbf{x}) \\ 0 & otherwise \end{cases}$$

Data Generation



GQ-CNN

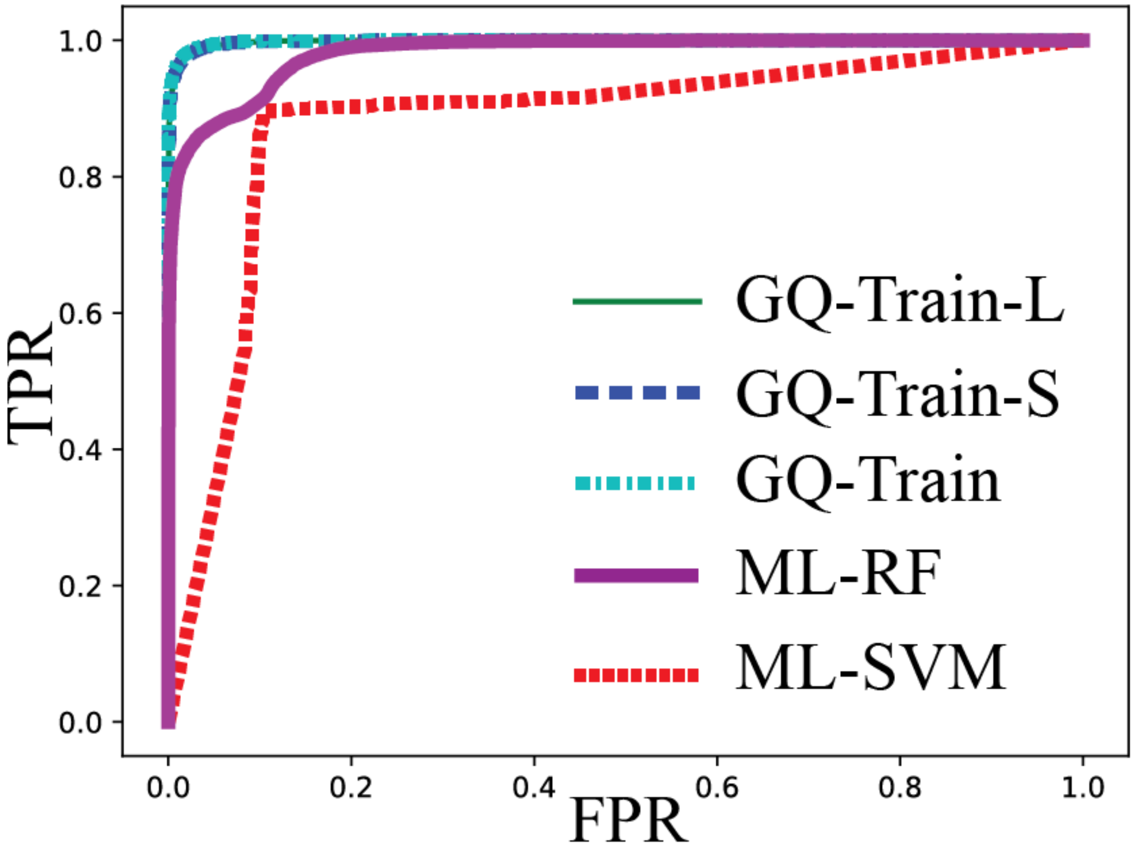


- Learn a Q-function, $Q_\theta(u, y)$
- CNN on pose aligned depth images
- Trained via supervised learning
- Test time:
  - $\pi_\theta(\mathbf{y}) = \text{argmax}_{\mathbf{u}\in\mathcal{C}} Q_\theta(\mathbf{u}, \mathbf{y})$
  - Optimized via CEM

Input: Num rounds m, Num inital samples n, Num CEM samples c, Num GMM mixture k, Friction coef μ, Elite percentage γ, Robustness function Q_θ
Result: $\mathbf{u}$, most robust grasp

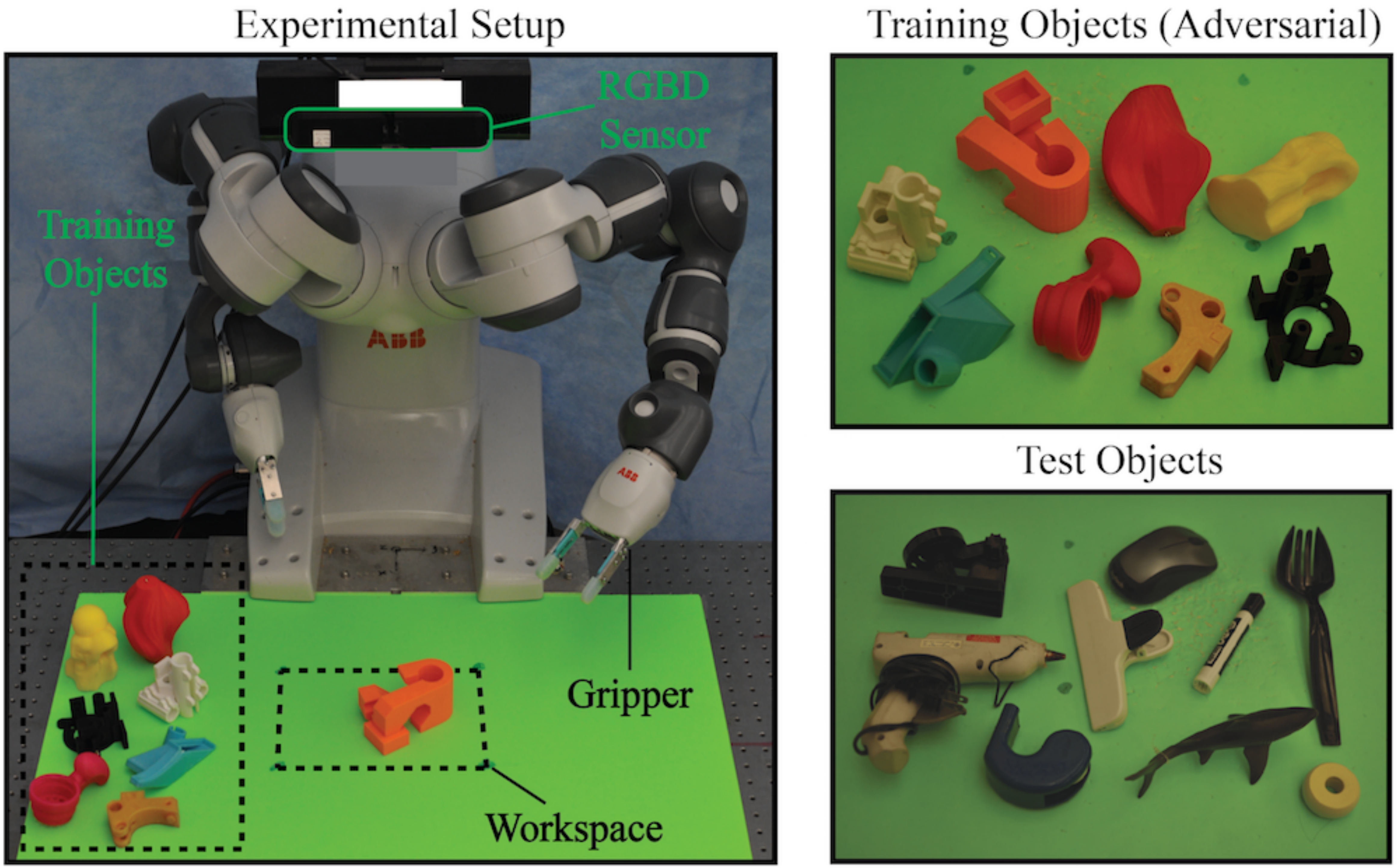
```
U ←uniform set of n antipodal grasps;
for i = 1, ..., m do
    E ←top γ−percentile of grasps ranked by Q_θ;
    M ←GMM fit to E with k mixtures;
    G ← c iid samples from M;
end
return argmax_{u∈U} Q_θ(u, y);
```

Experiments (Offline)



Model	**Accuracy (%)**
ML-SVM	89.7
ML-RF	90.5
GQ-S-Adv	97.8
GQ-L-Adv	97.8
GQ-Adv	98.1

Experiments



All methods other than point cloud registration used the antipodal grasp sampling method described in Section V with the same set of parameters to generate candidate grasps, and each planner executes the highest-ranked grasp according to the method.

Experiments

- Metrics
  - Success rate
  - Precision: success rate for confident grasps
  - Robust grasp rate: how many grasps were confident
  - Planning time
- Baselines:
  - Image-based grasp quality metrics
  - Point cloud registration
  - Alternative ML Models
- Results (Test Objects)

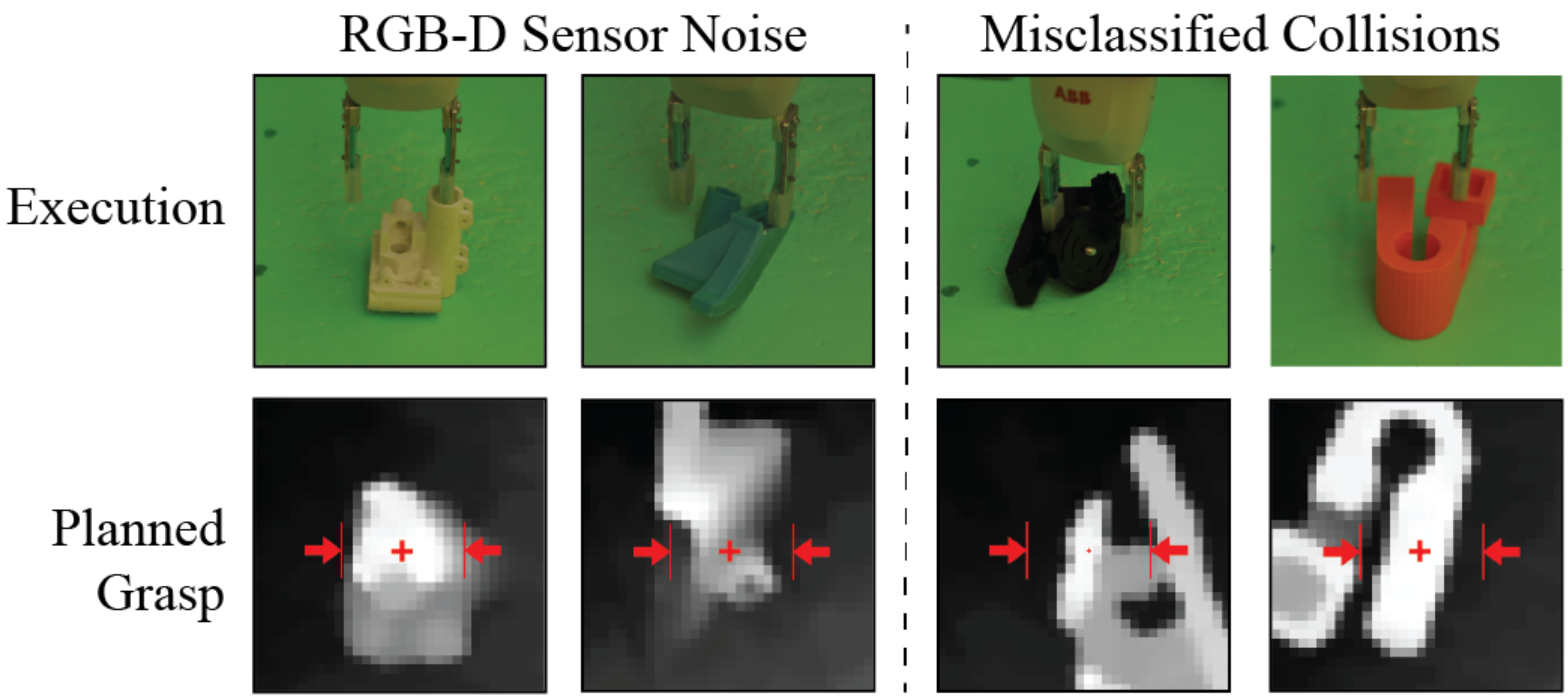
	IGQ	REG	GQ-Adv-Phys	GQ-Adv	GQ-S	GQ
Success Rate (%)	60±13	52±14	68±13	74±12	72±12	**80±11**
Precision (%)	N/A	N/A	68	87	92	**100**
Robust Grasp Rate (%)	N/A	N/A	100	30	48	**58**
Planning Time (sec)	1.8	3.4	0.7	0.7	0.8	0.8

Grasping System Evaluation

Generalization Objects

- The Dex-Net 2.0 grasp planner achieves 94% success and 99% precision on a dataset of 40 novel household objects, some of which are articulated or deformable. And it took an average of 2.5s to plan grasps.

Failure Modes



Thank you