Scalable **C**ollaborative **H**uman–R**o**b**o**t **L**earning
(SCHooL)

Sept 2017 - Aug 2021

Scalable **C**ollaborative **H**uman–R**o**b**o**t **L**earning
(SCHooL)

Sept 2017 - Aug 2021

Ken Goldberg

Pieter Abbeel

Anca Dragan

Stuart Russell

Berkeley
UNIVERSITY OF CALIFORNIA

BAIR
BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

AUTOLAB

schoolproject.berkeley.edu

Learning from Demonstrations

- **Inverse Reinforcement Learning (IRL):** learn reward structure
  - → **more robust, generalizable, and explainable**
- Pioneered by co-PI Russel (1998)
- Apprenticeship Learning (Abbeel & Ng, 2004)
- **Cloud Robotics** (Kuffner 2015, Kehoe, Abbeel, Goldberg, 2017)
- Inference-based (Ramachandran & Amir, 2007; Dimitrakakis & Rothkopf, 2011; Levine et al., 2011)
- Entropy-constrained (Ziebart et al., 2008; Boularias et al., 2011)
- Active Learning (Lopes et al., 2009)

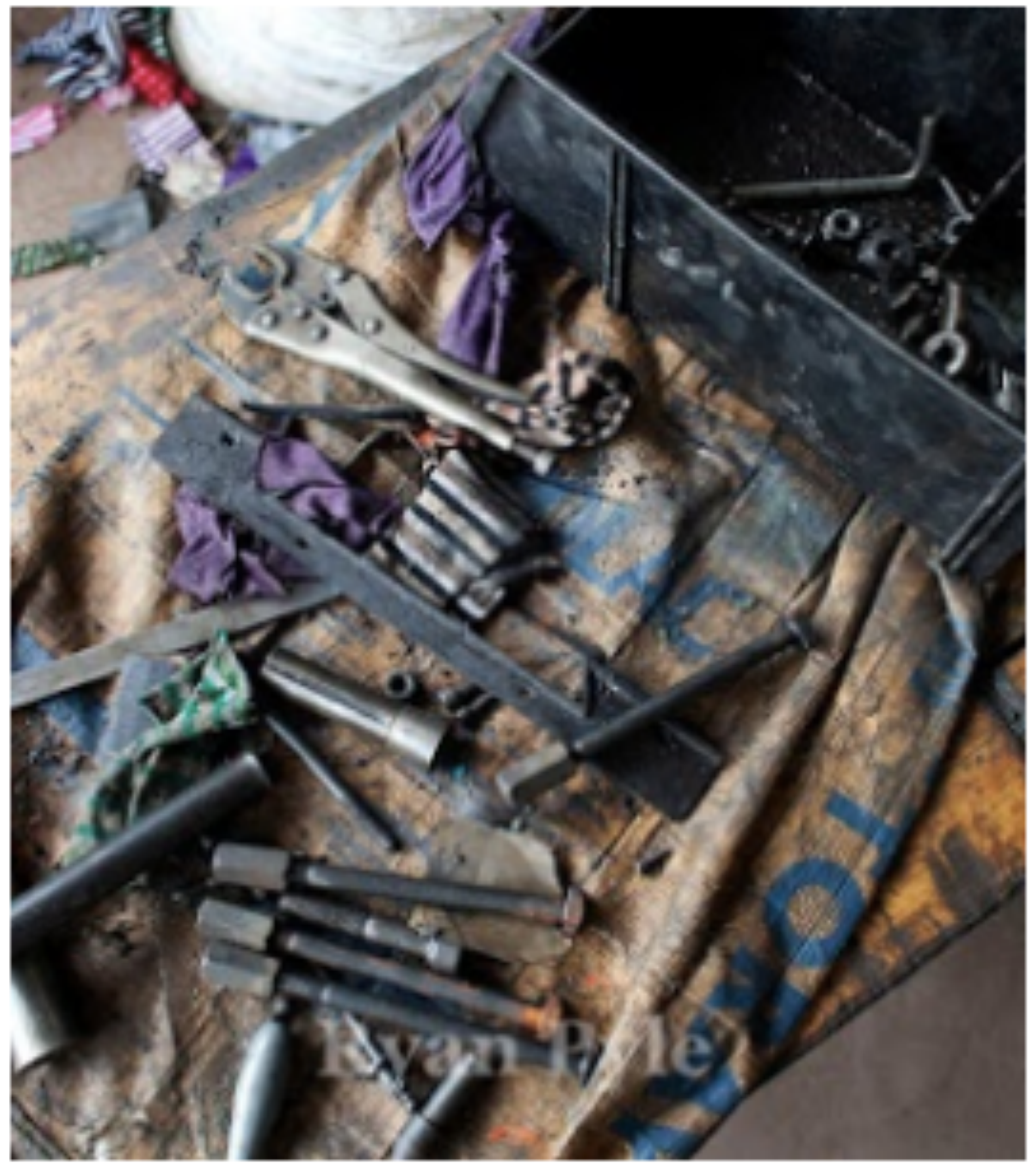
Formal Framework: Cooperative IRL

- A CIRL is a 2-player cooperative Markov game $\langle \mathcal{S}, (\mathcal{A}^H, \mathcal{A}^R), P, R_\theta \rangle$
- Human and robot take simultaneous actions, get same reward parametrized by θ
- Human preference R_θ initially unknown to the robot
- This incentivizes the human to teach and the robot to learn this preference
- Both agents can reason about the robot's belief state, making it a sufficient state representation (together with the environment state)
- Example: human signals which objects should not be decluttered by replacing them in the environment when the robot removes them

4 Research Objectives

$$\tau^{\mathbf{H}} \leftarrow \underset{\tau}{\operatorname{argmax}}\ \phi(\tau)^{\top}\theta - \eta||\phi_\theta - \phi(\tau)||^2$$

1. Extend CIRL Formal Framework:
2. Distributed Sensing, Reward Models using Deep Learning
3. Learning Hierarchical Task and Reward Structure
4. Bidirectional / Active Human–Robot Communication

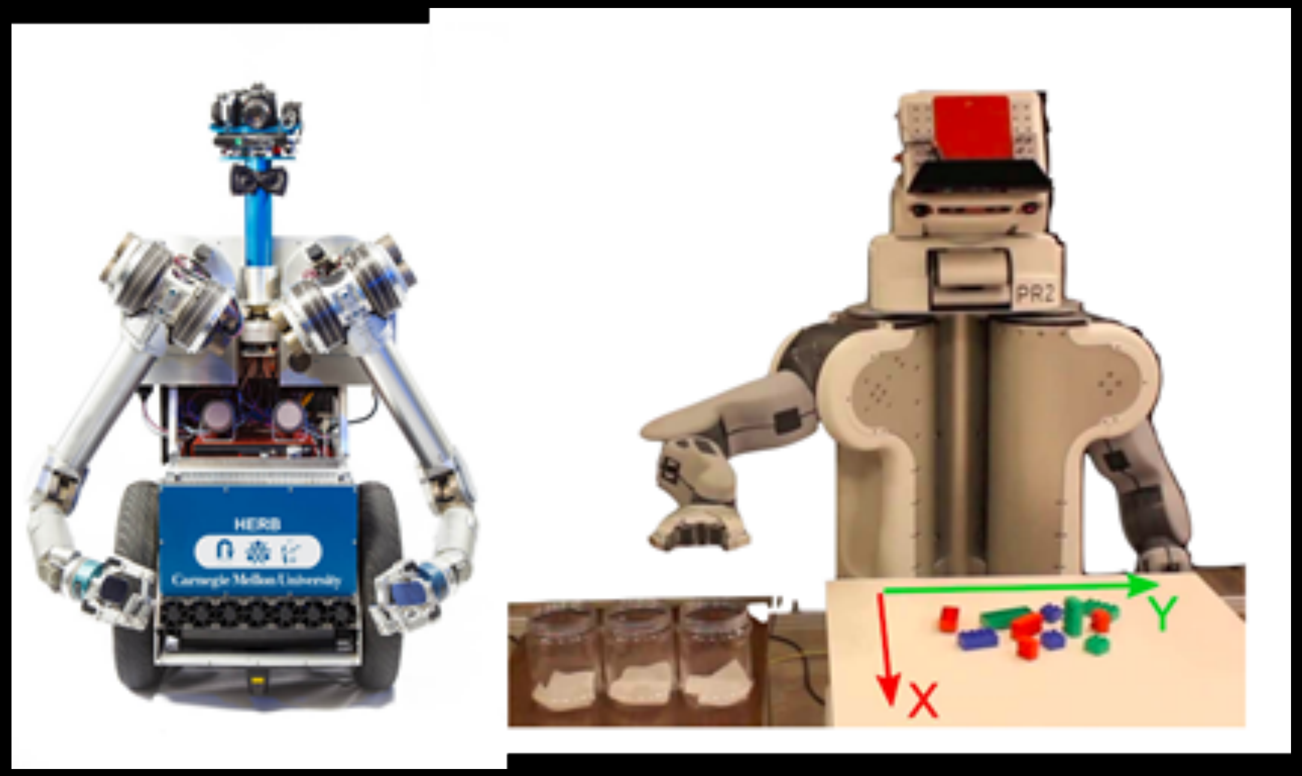

Integrative Application:

Surface Decluttering

To increase productivity and safety in homes, machine shops, warehouses, offices, and retail stores.

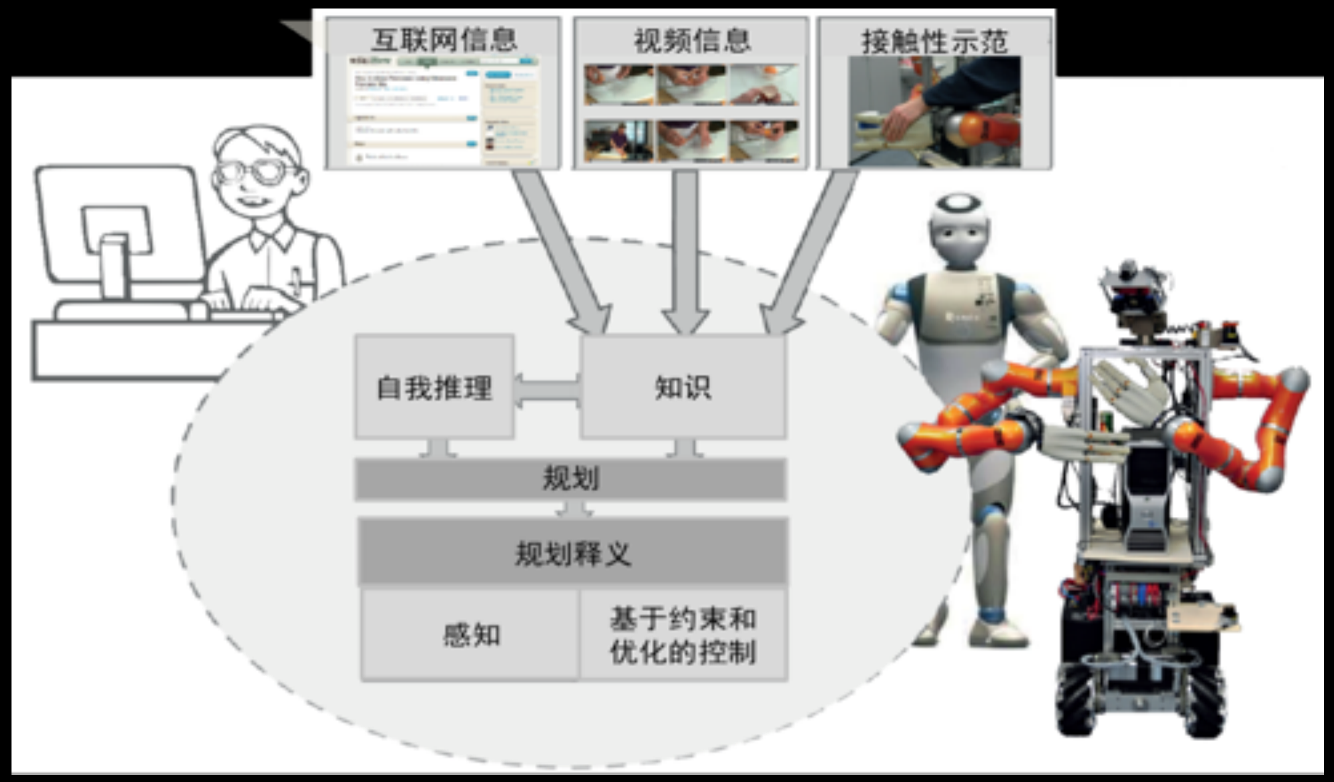

fruit
screwdriver
wrench
cup
tube
scissors
bottle
hammer
toy
utility
assembly
tape
20x

Related Work



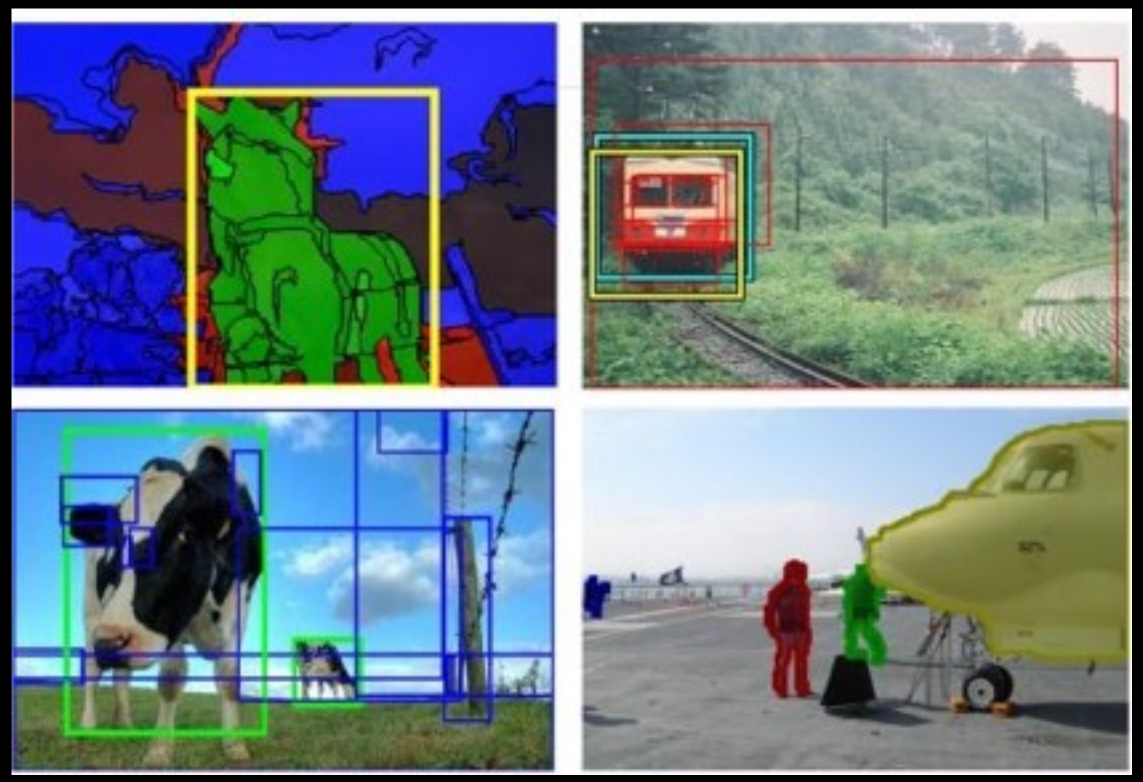
Surface Decluttering

[Tanwani et al., 2018]
[Preferred Networks, 2018]
[Srinivasa et al, 2010]
[Gupta et al, 2015]



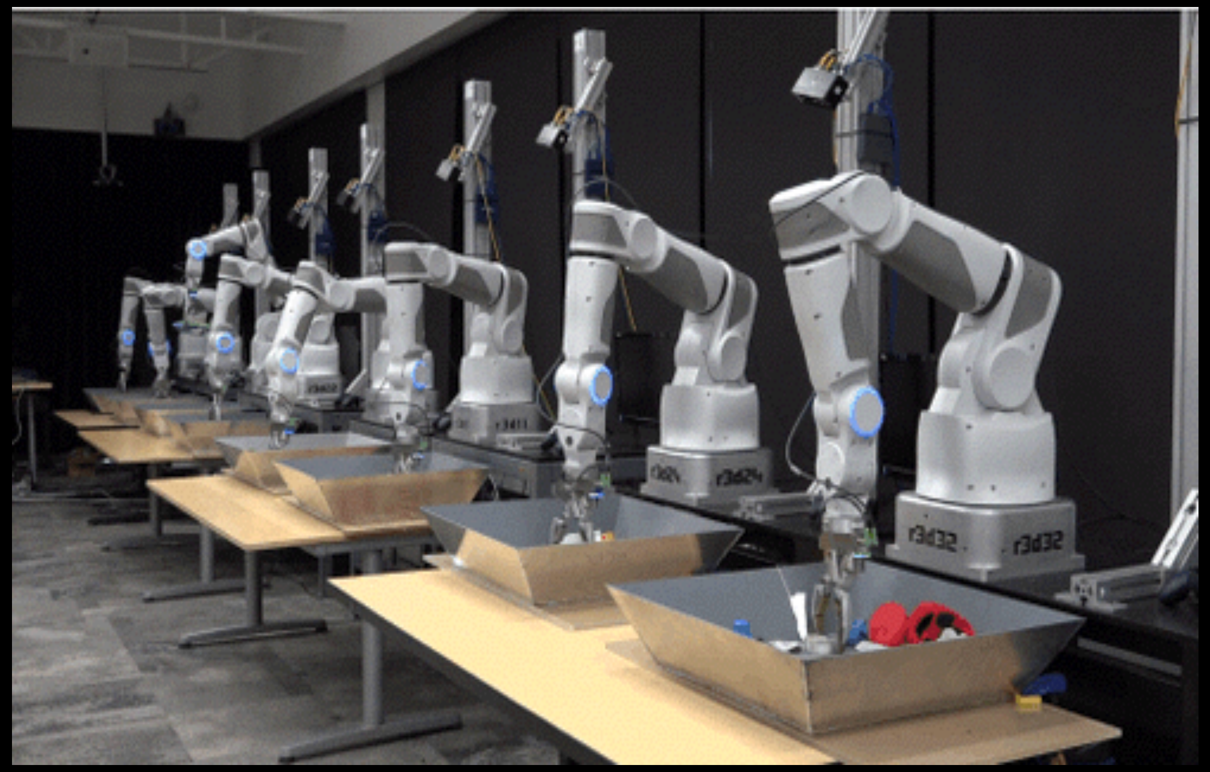
Cloud and Fog Robotics

[Tanwani et al, 2018]
[Gianni et al, 2016]
[Kehoe et al, 2015]



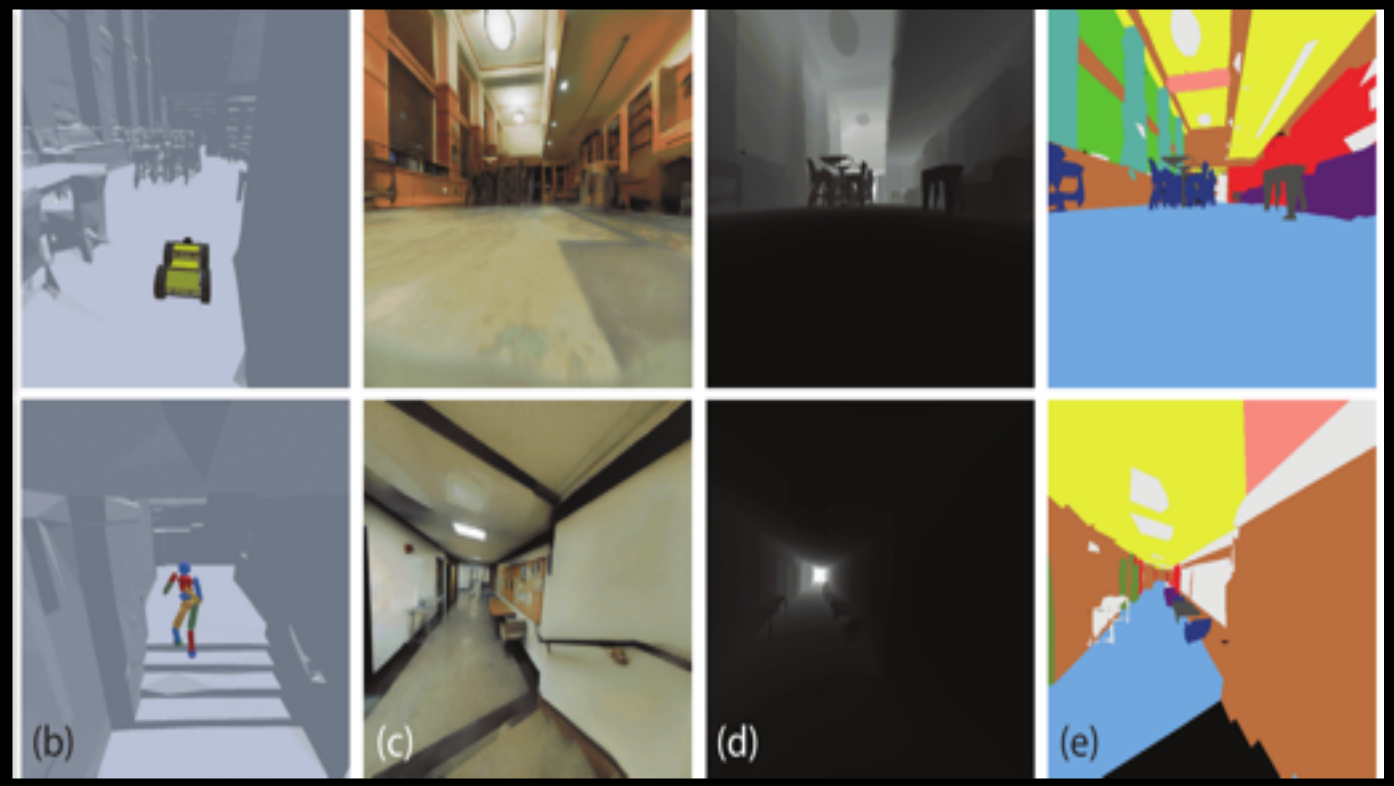
Object Recognition

[Wang et al, 2017]
[Max et al, 2017]



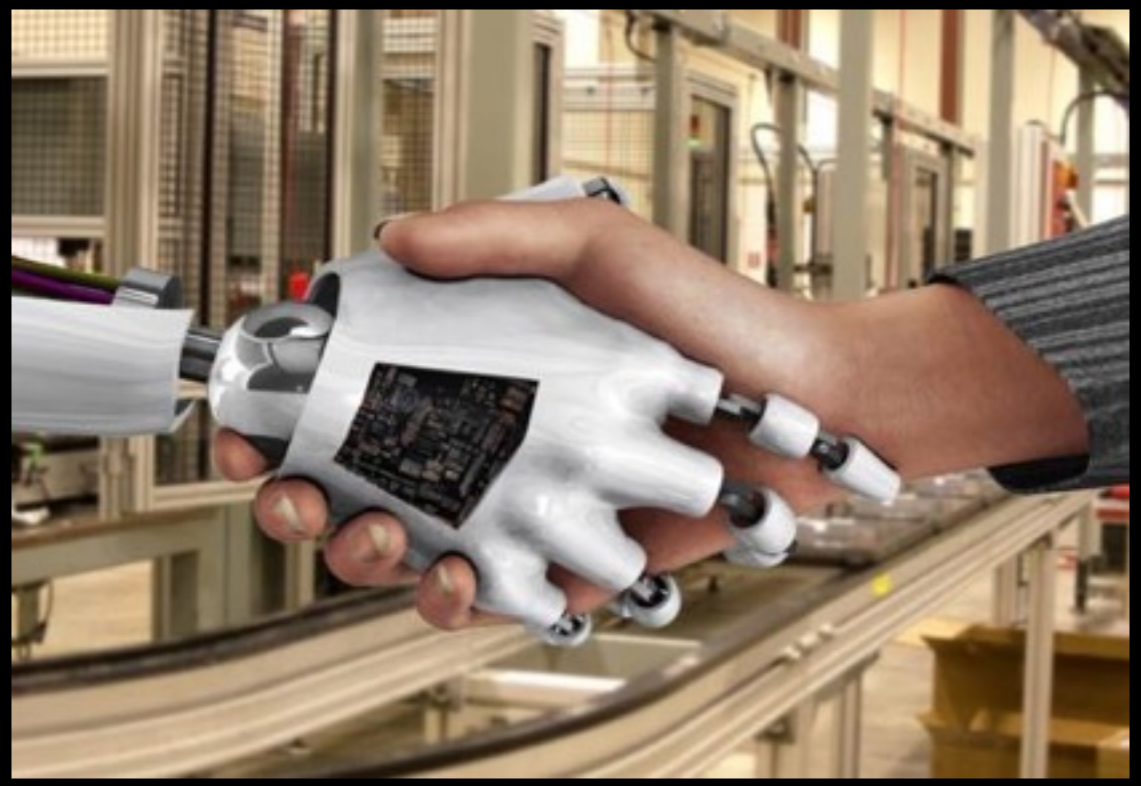
Robust grasping

[Lerrel et al, 2015]
[Sergey et al, 2016]
[Mahler et al., 2017]



Sim2real Transfer

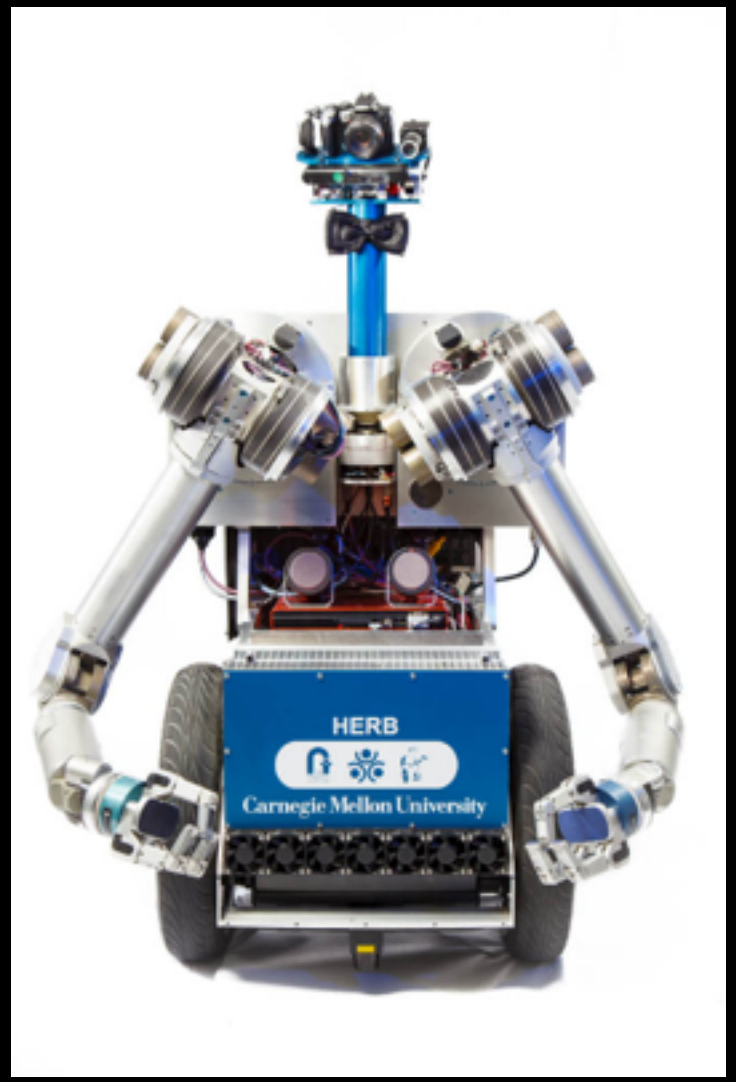
[Zamir et al, 2018]
[Peng et al. 2017]

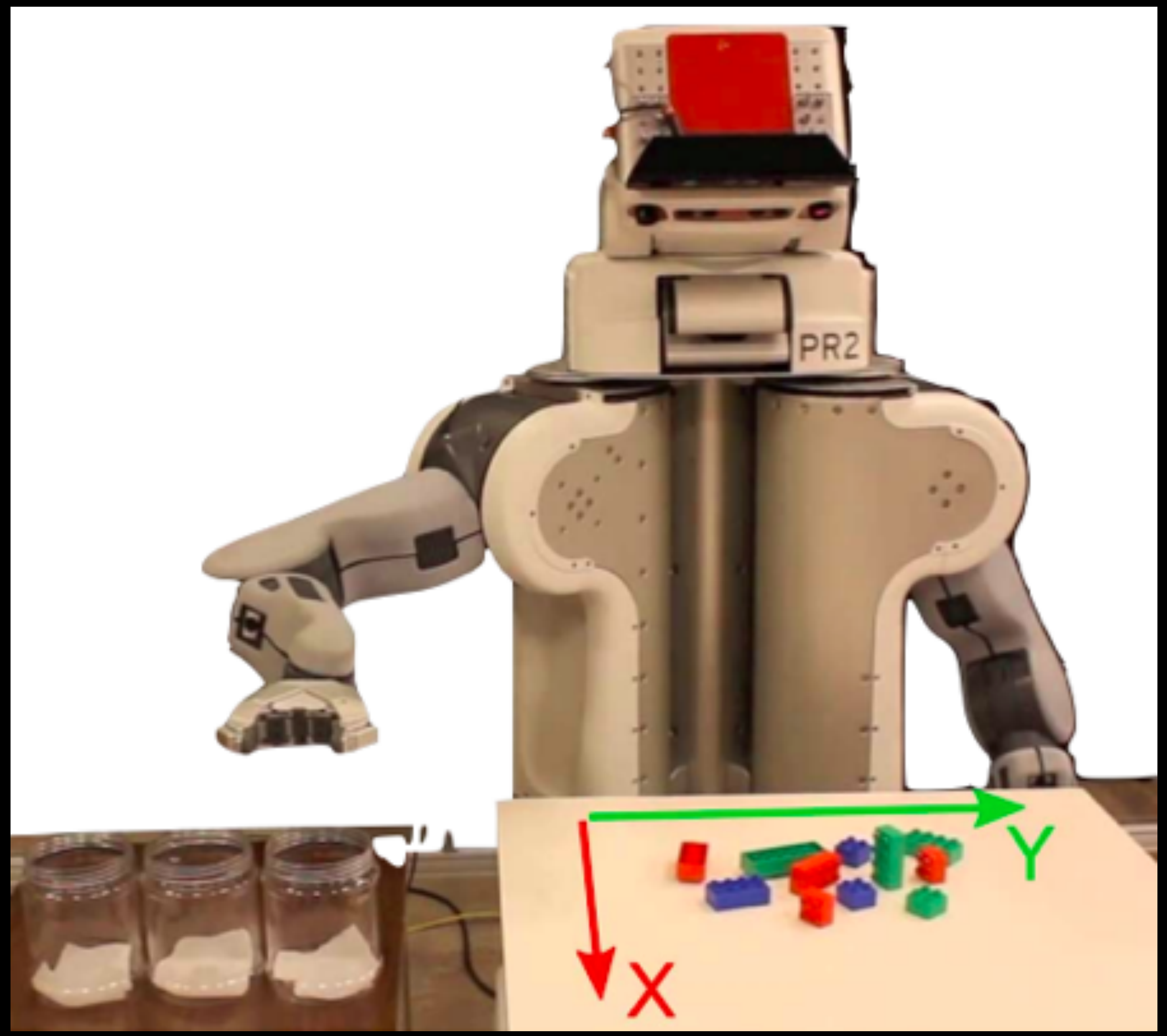


Human-Robot Interaction

[Thomas et al, 2016]
[Leonel et al, 2016]

Related Work: Mobile Manipulator Decluttering





HERB CMU 2009
- Unstructured grasping

[Srinivasa et al., 2010]

PR2 USC 2005
- Grasping and Pushing

[Gupta et al., 2012, 2015]

Low cost mobile robots
- Unstructured grasping in homes

[Gupta et al., 2018]

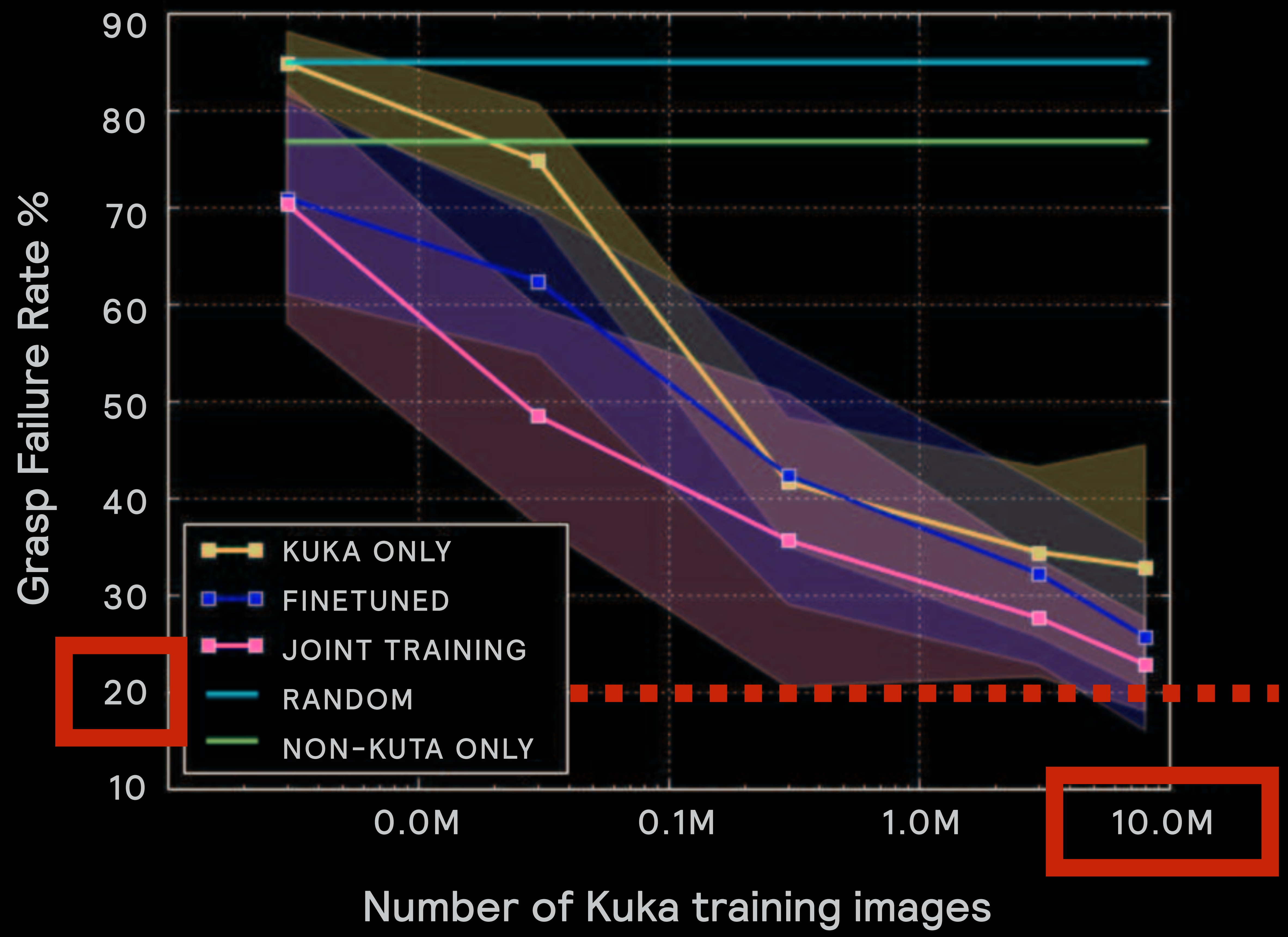
Surface Decluttering
- Grasping and Object Recognition
- Fog Computing
- Sim-to-Real Transfer

[Tanwani et al., 2019]

Universal Picking:

grasp diversely shaped and sized novel objects

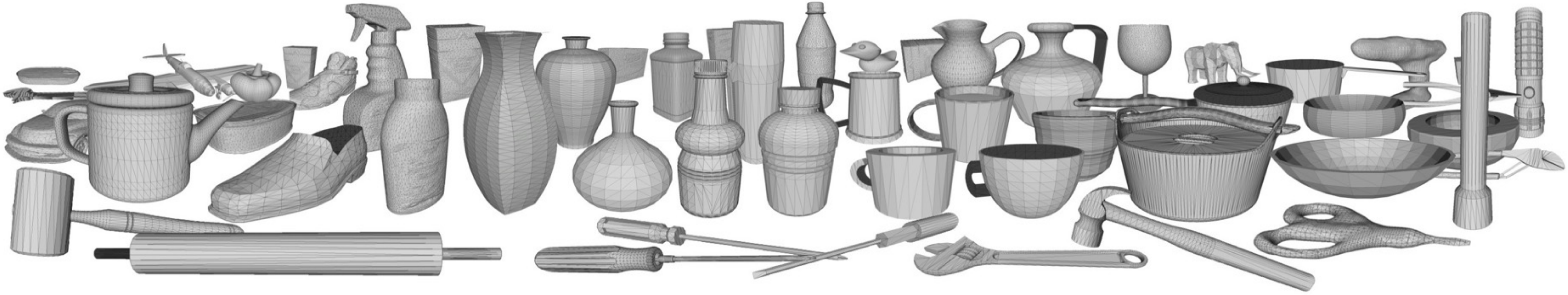
“Arm farm”



S. LEVINE ET AL. “LEARNING HAND-EYE COORDINATION FOR ROBOTIC GRASPING WITH DEEP LEARNING AND LARGE-SCALE DATA COLLECTION.” IJRR 2017.

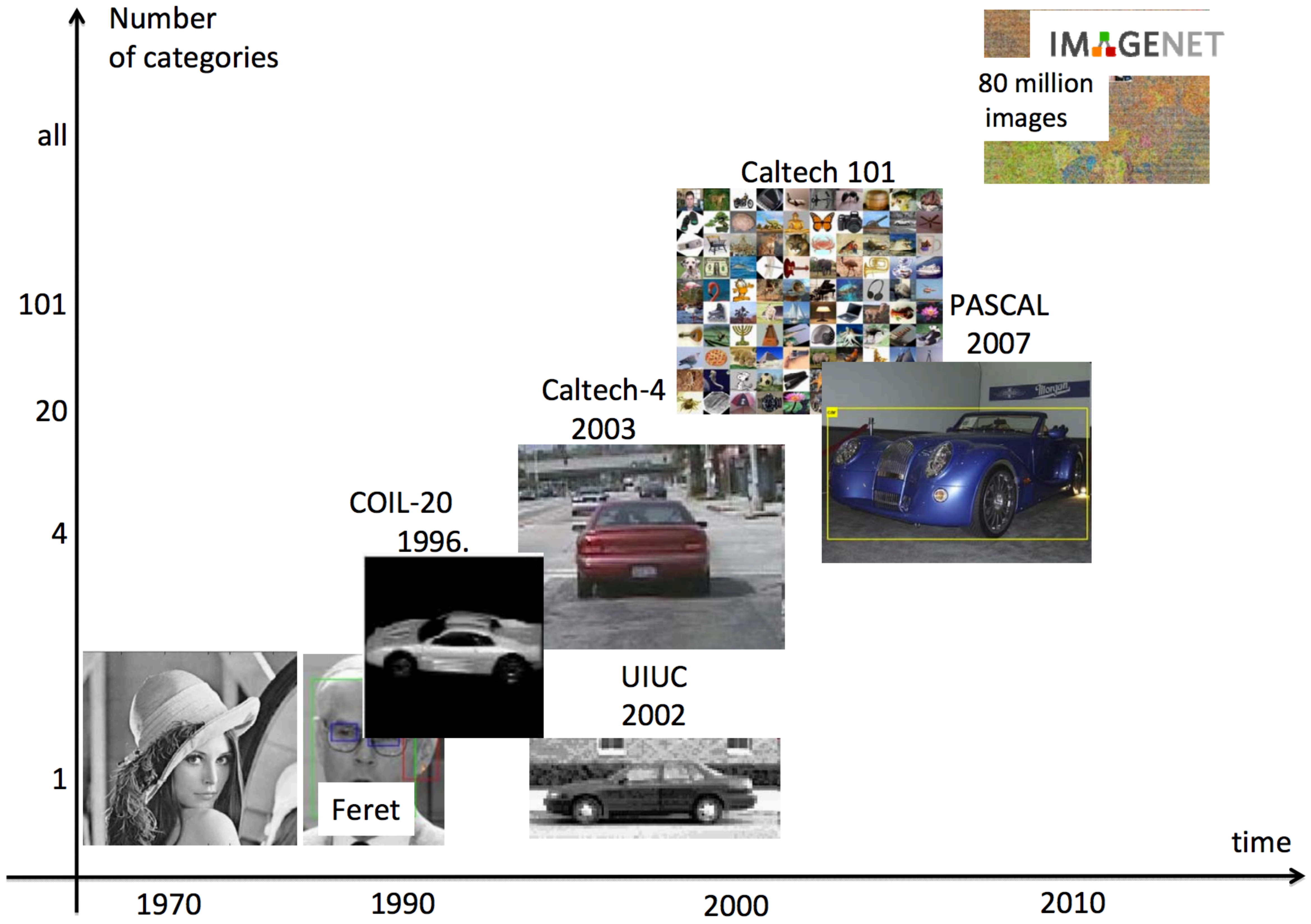
Dex-Net
(Dexterity Network)

Jeff Mahler

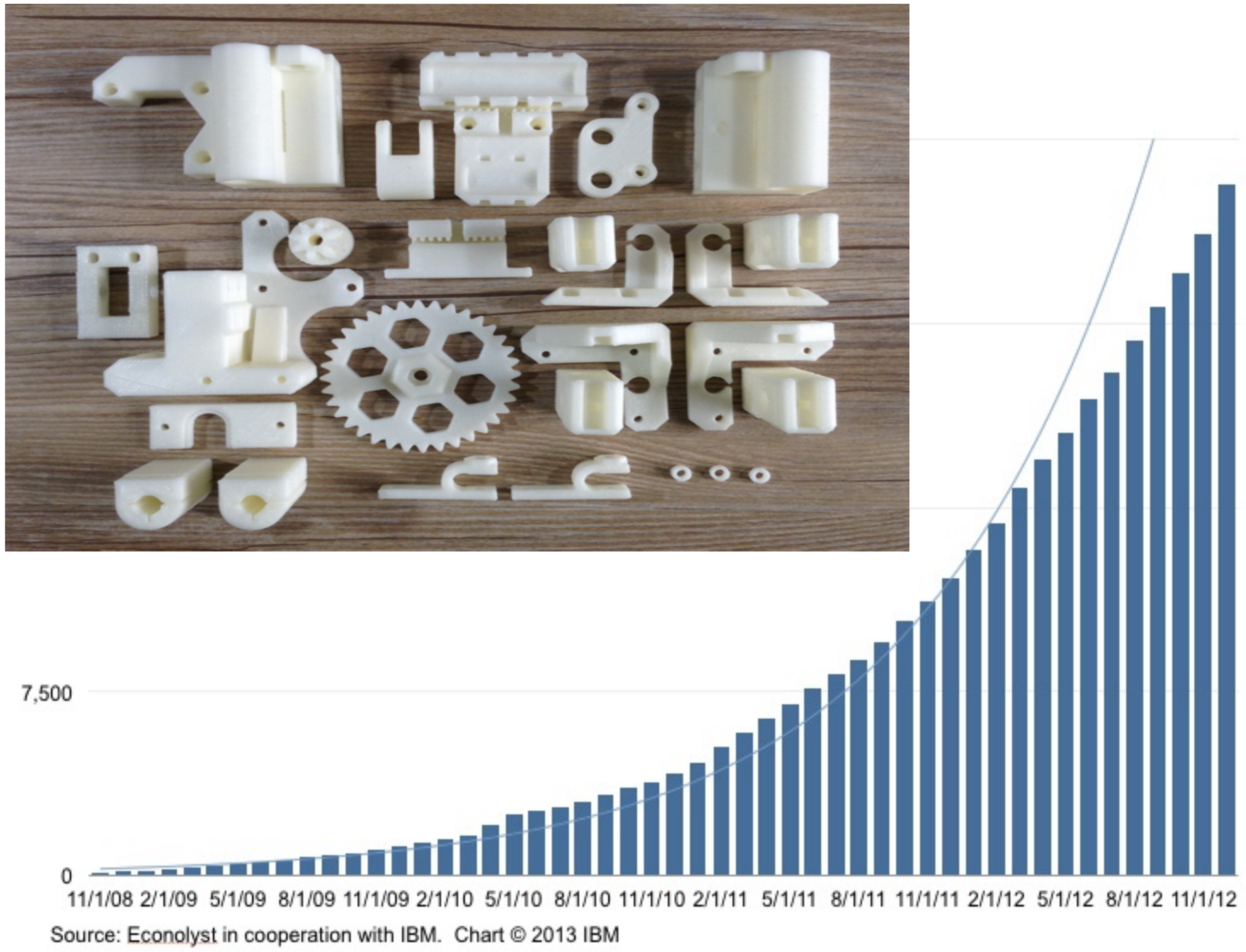


Jeff Mahler, Florian T. Pokorny, Brian Hou, Melrose Roderick, Michael Laskey, Mathieu Aubry, Kai Kohlhoff, Torsten Kroeger, James Kuffner, Ken Goldberg. “Dex-Net 1.0: A Cloud-Based Network of 3D Objects for Robust Grasp Planning Using a Multi-Armed Bandit Model with Correlated Rewards.” IEEE International Conference on Automation Science and Engineering (ICRA) 2016

Large Datasets of 3D Models



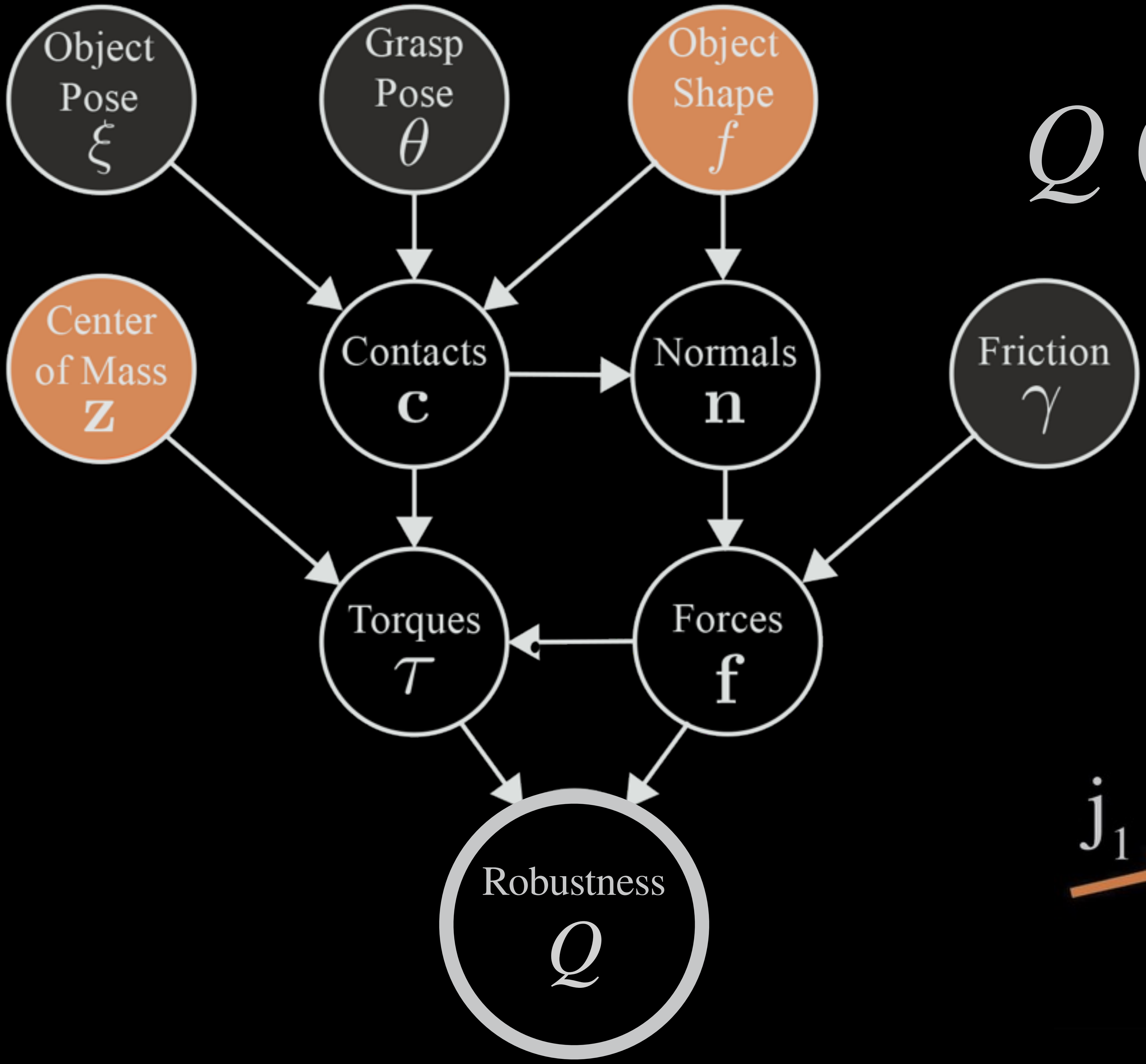
Computer Vision

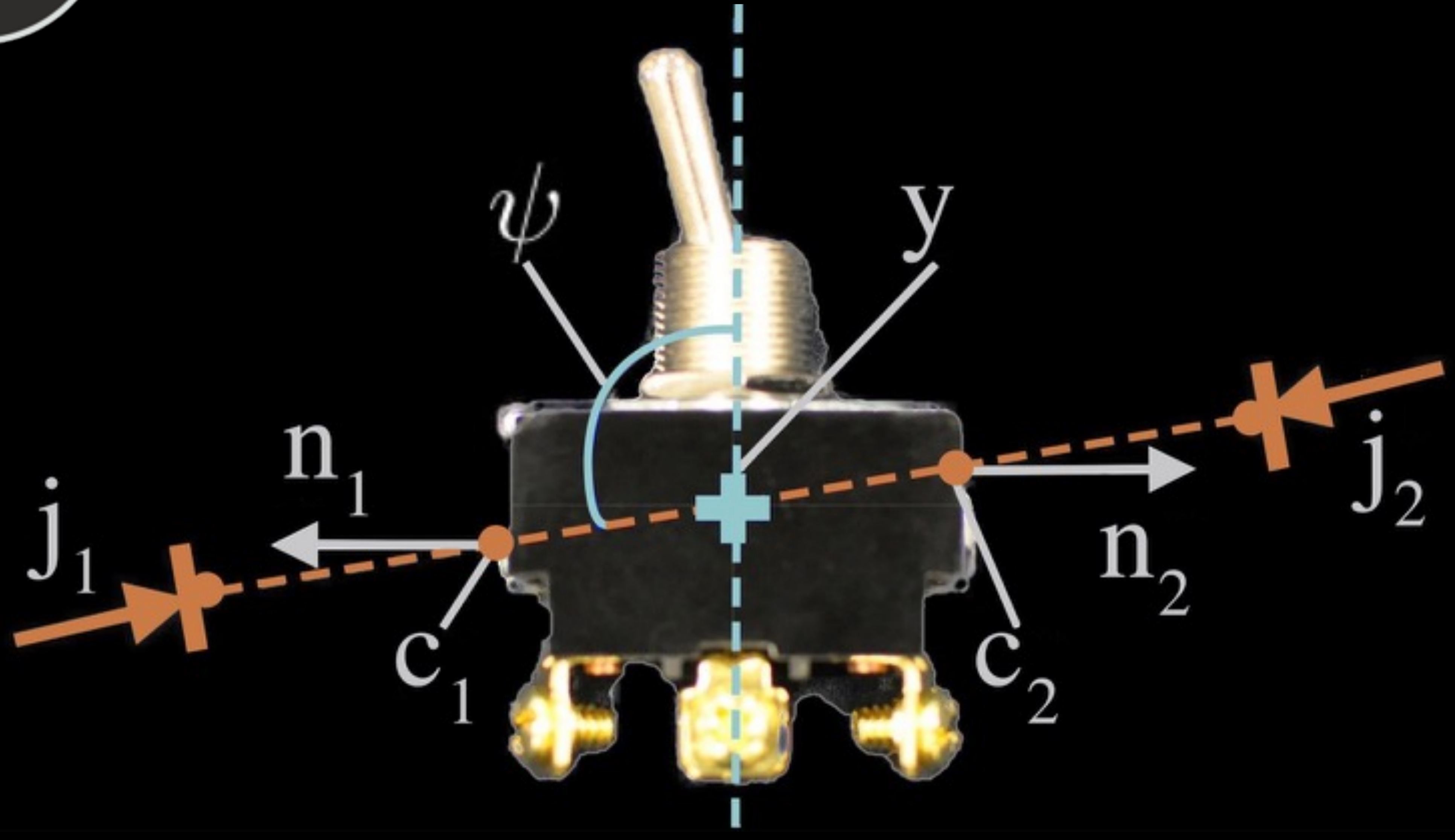


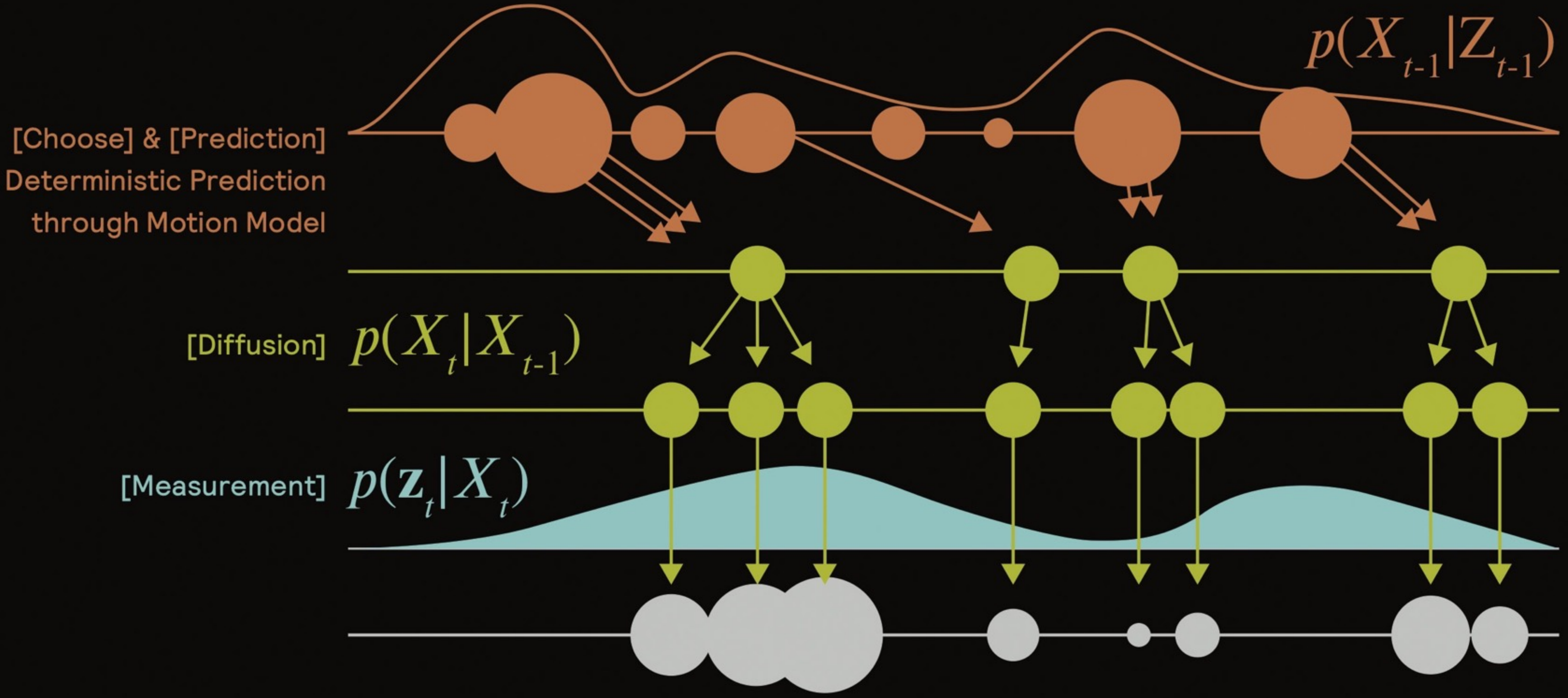
Robot Grasping

First Wave:
Stochastic
Analytic
Methods

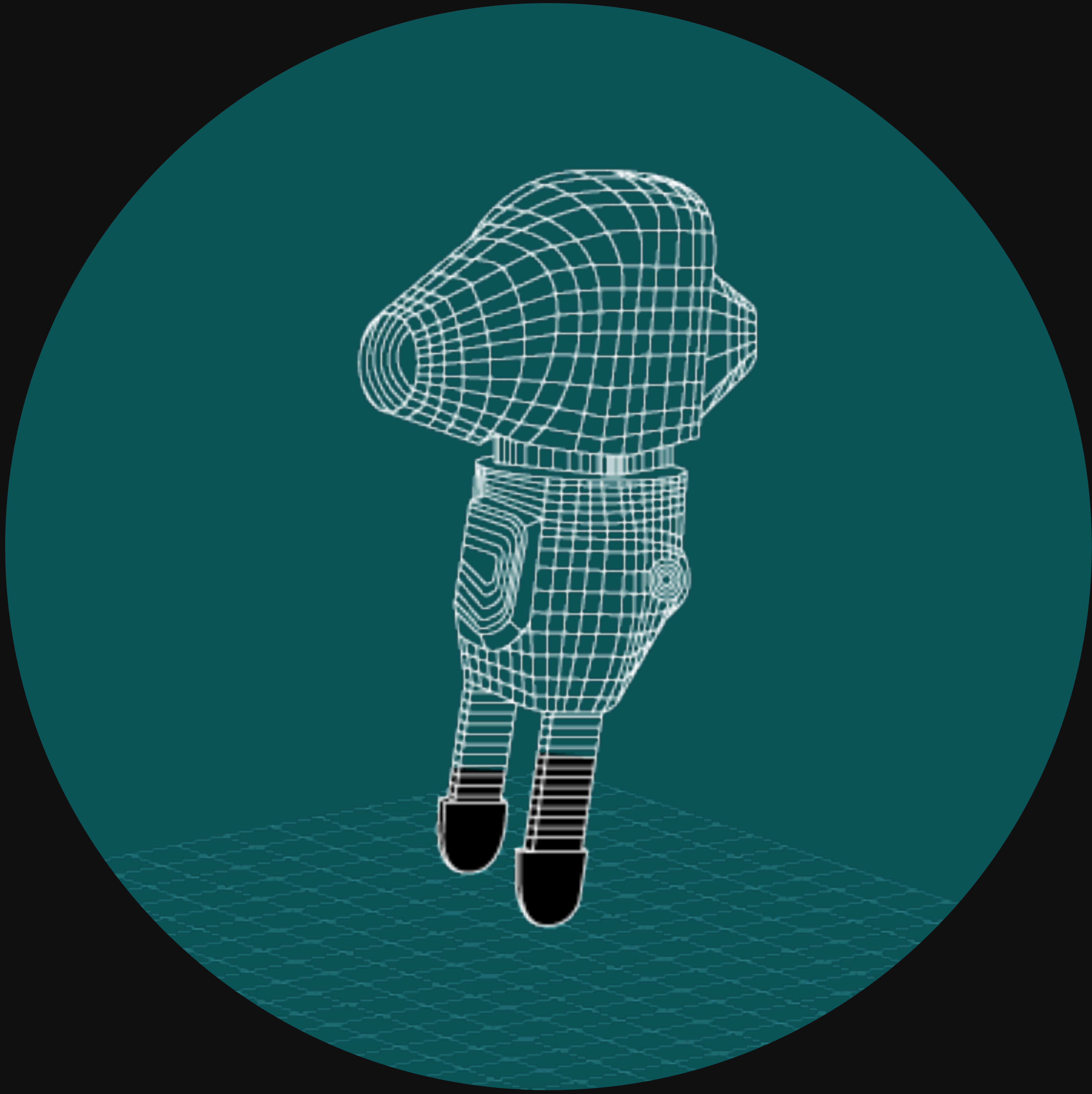
$$\tilde{\mathbf{x}} = \hat{\mathbf{x}} + \epsilon$$
$$\tilde{\mathbf{u}} = \mathbf{u} + \delta$$
$$Q\left(\mathbf{x}, \mathbf{u}\right) = \mathbb{E}\left[R(\tilde{\mathbf{x}}, \tilde{\mathbf{u}})\right]$$





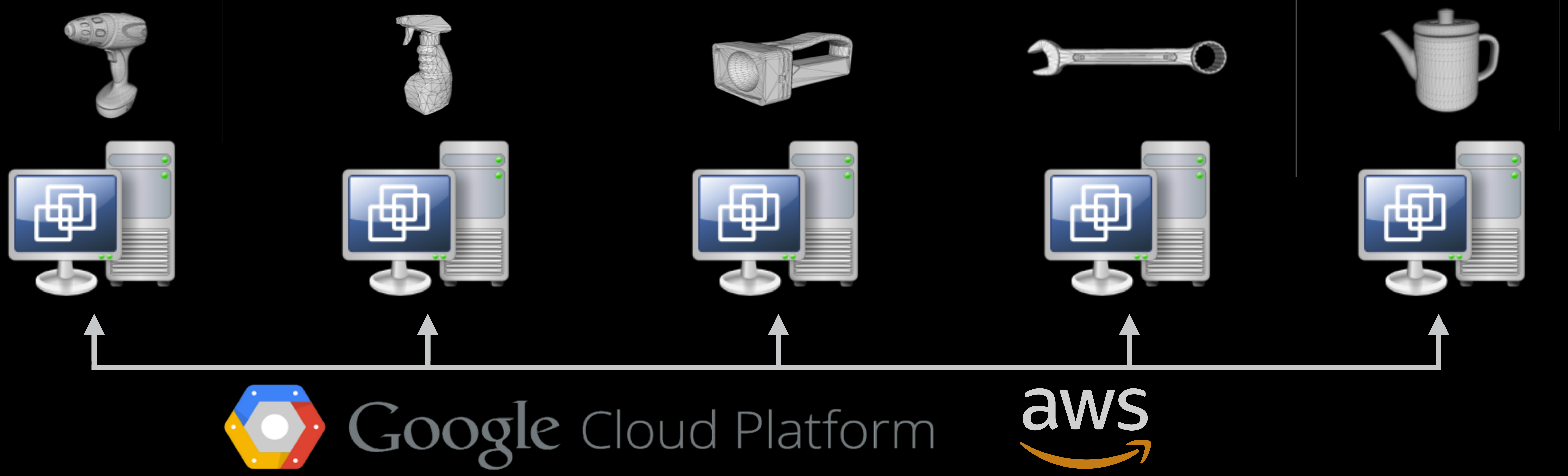


$$p(X_t|Z_t)=p(z_t|\boldsymbol{X}_t)\int p(\boldsymbol{X}_t|\boldsymbol{X}_{t-1})p(\boldsymbol{X}_{t-1}|ZX)dX_{t-1}$$



Stochastic Grasp Analysis

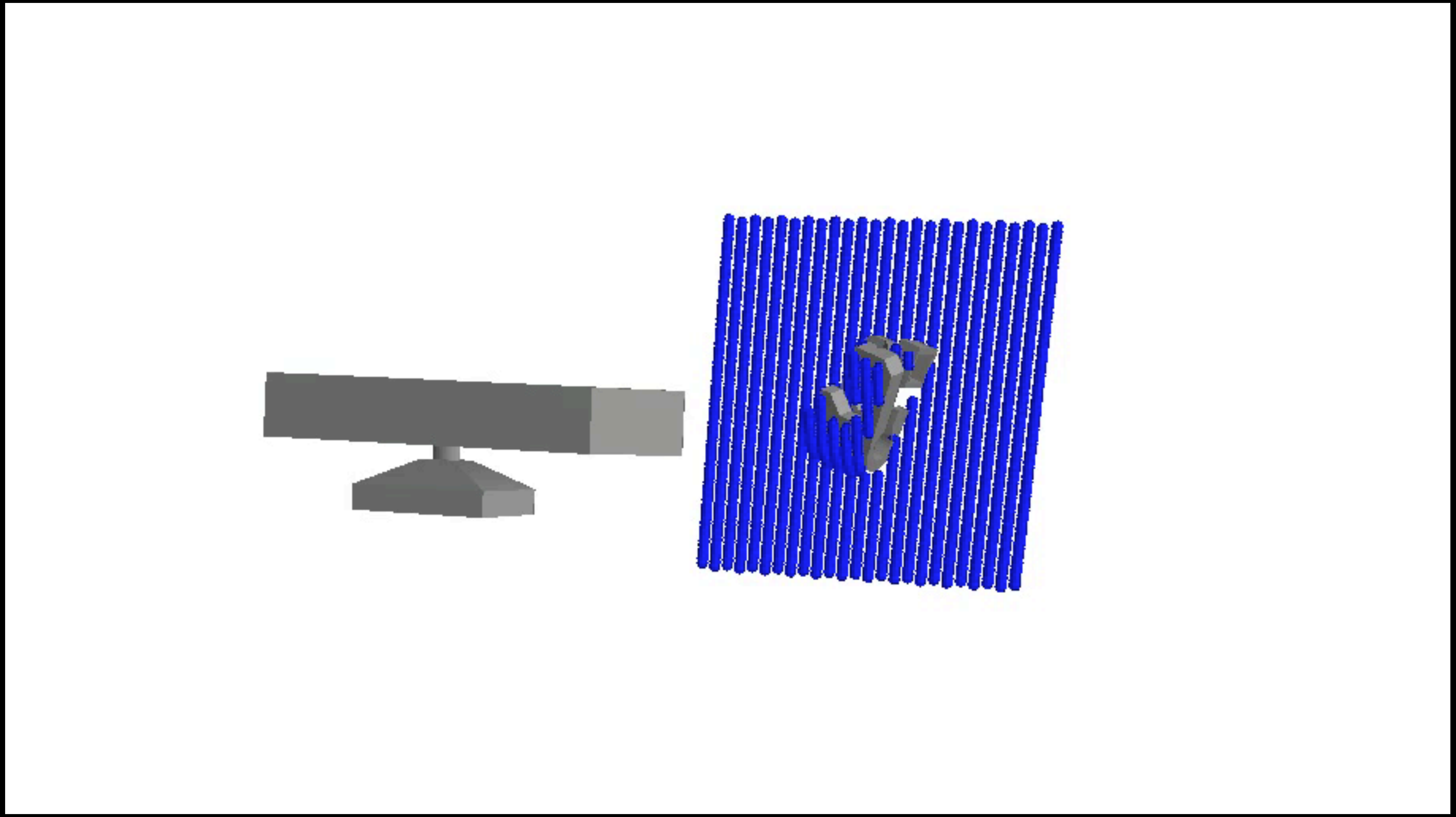
- 1,000+ facets per 3D object
- 1,000,000+ candidate grasps per object
- 1,000+ perturbations per grasp
- 1 billion grasp evaluations per object
- 1000 3D objects
- 1 trillion grasp evaluations



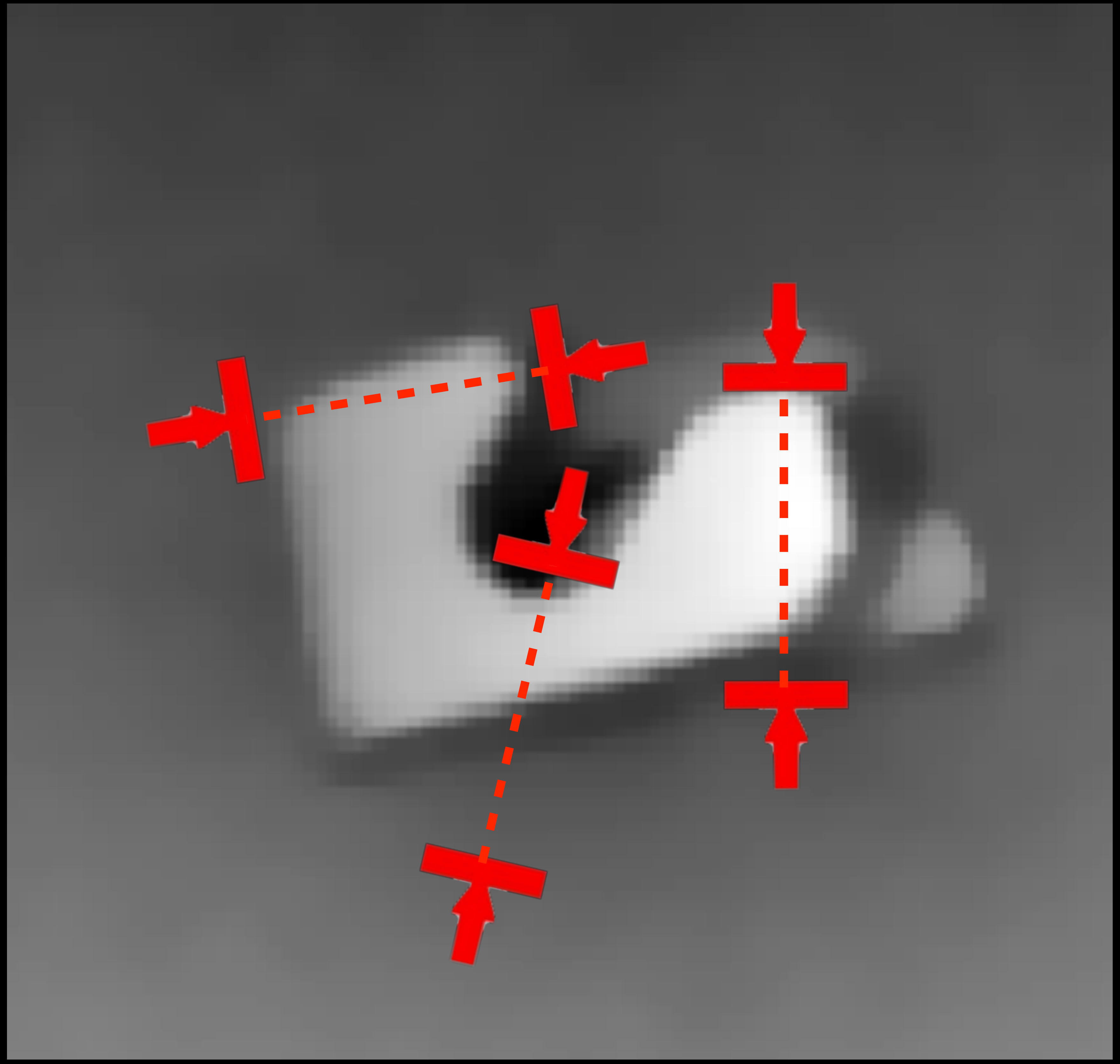
Synthetic LIDAR Images

$$\mathcal{D} = \{(\mathbf{y}_i, \mathbf{u}_i, R_i)\}_{i=1}^{I}$$

Rendering



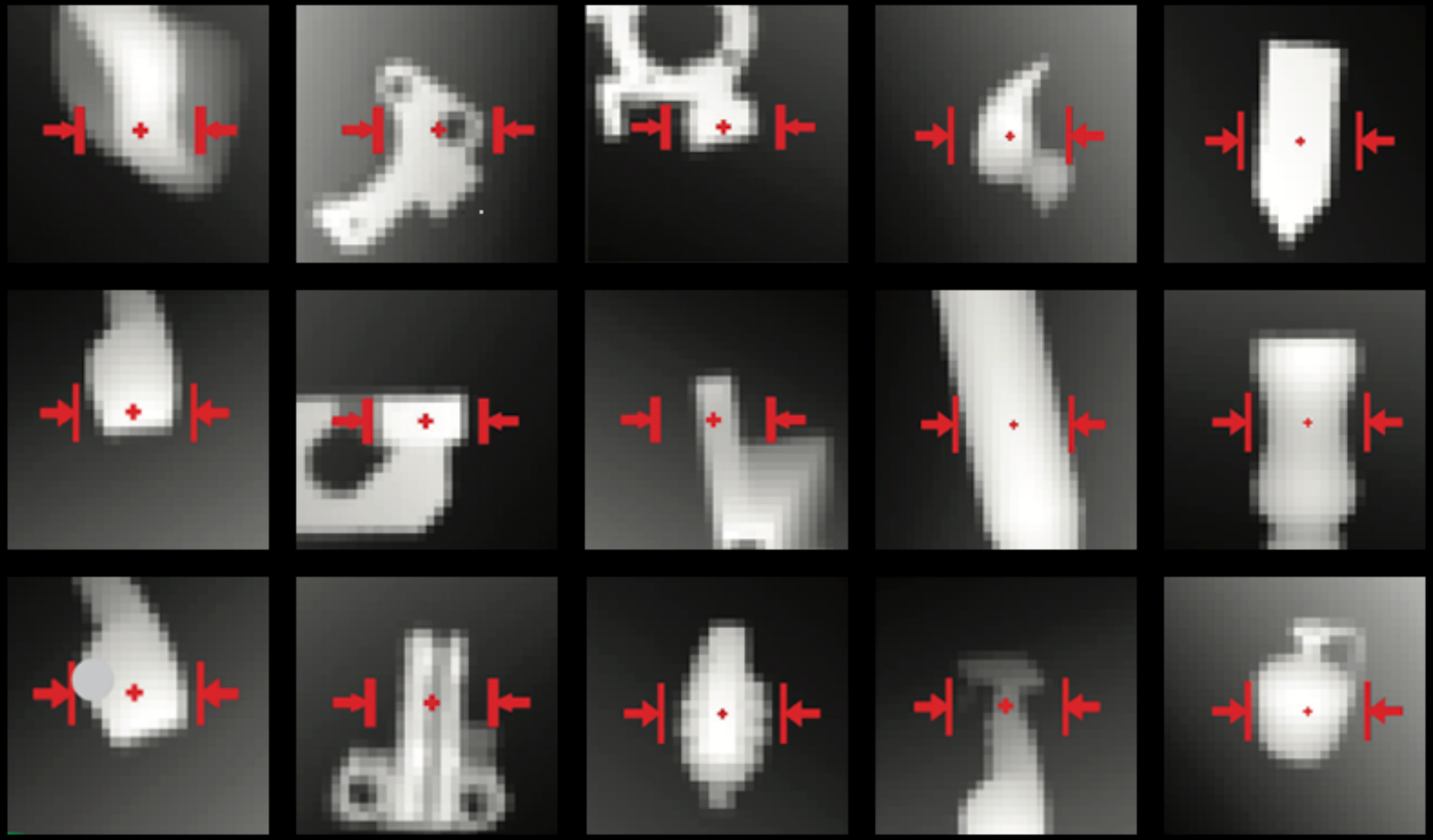
Sample Grasps



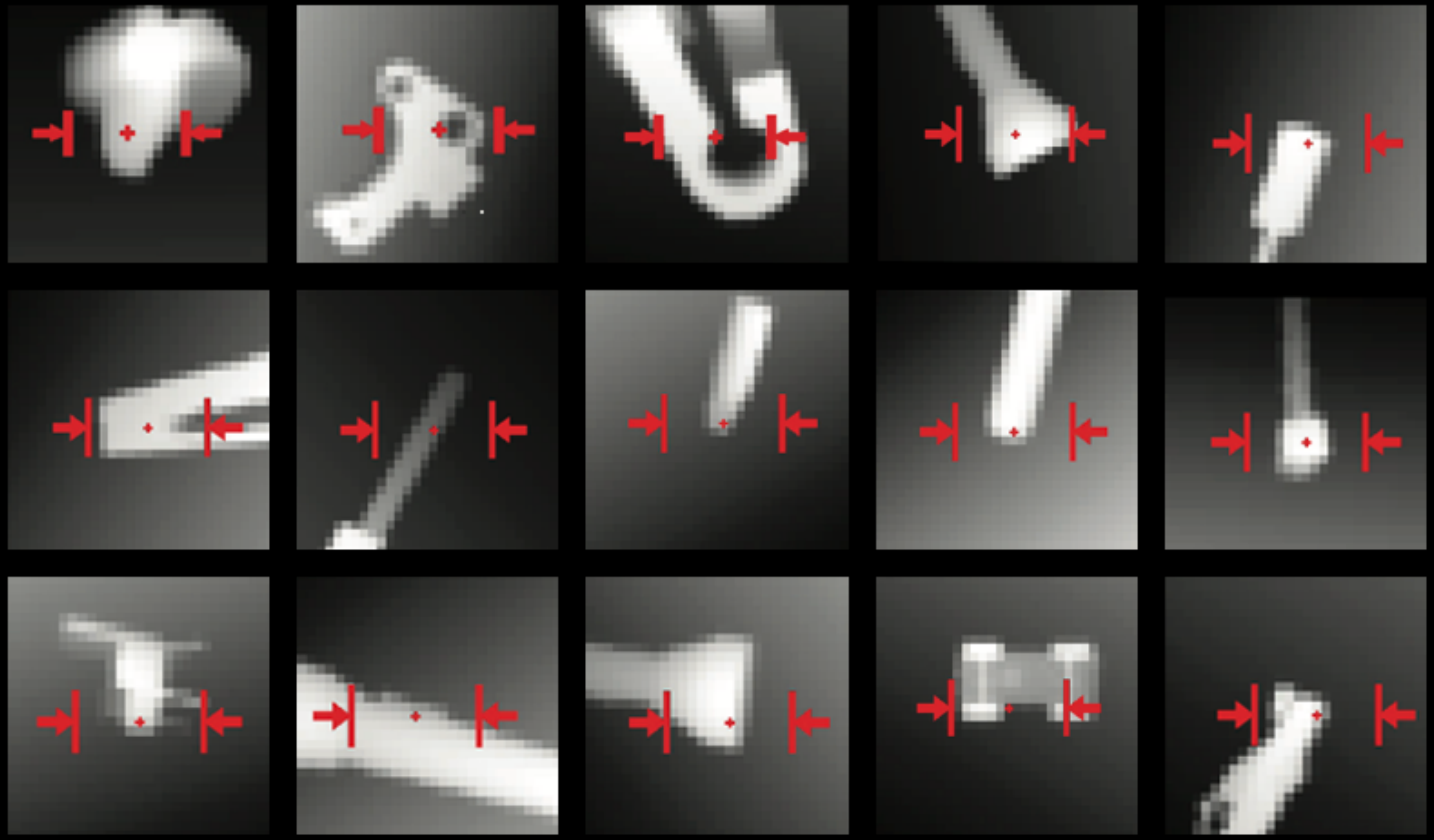
Dex-Net
6.7 million examples

$$\mathcal{D} = \{(\mathbf{y}_i, \mathbf{u}_i, R_i)\}_{i=1}^{I}$$

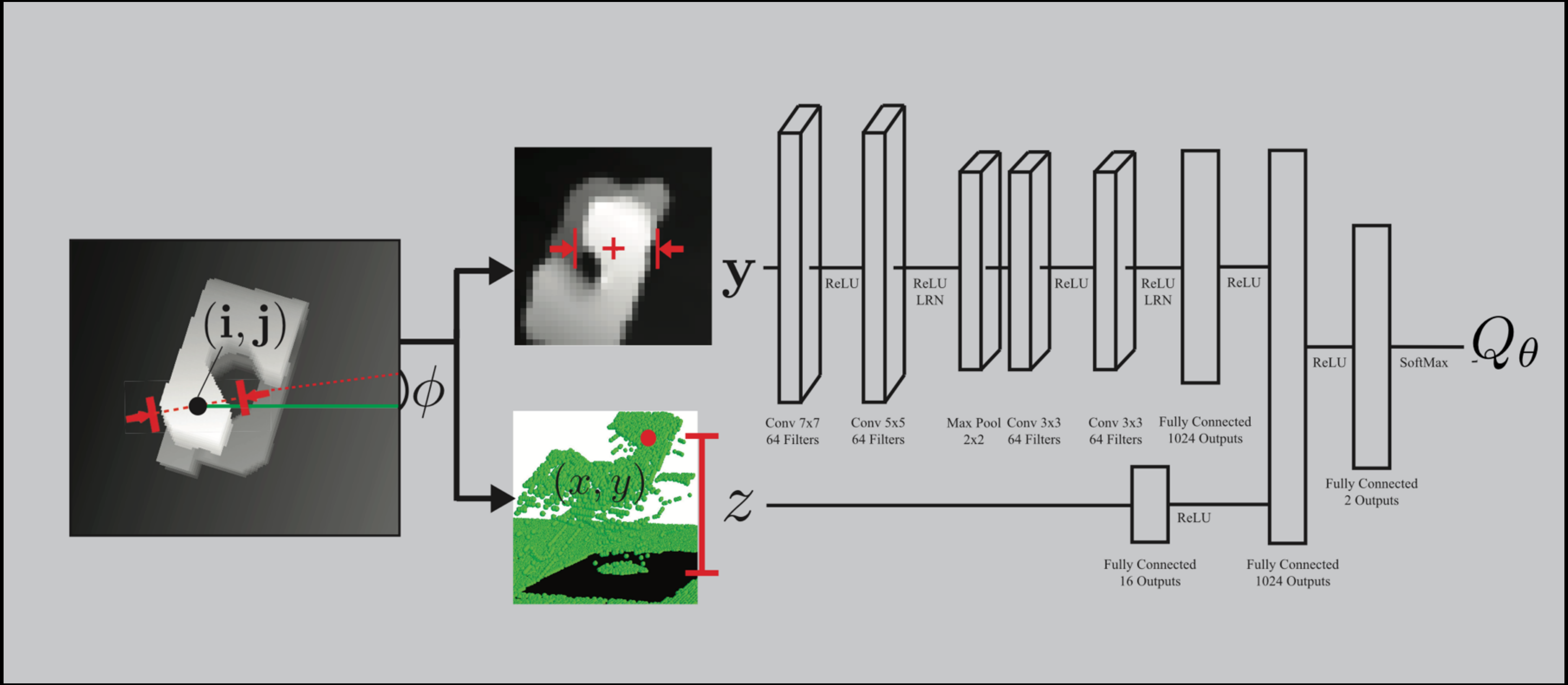
Positive



Negative

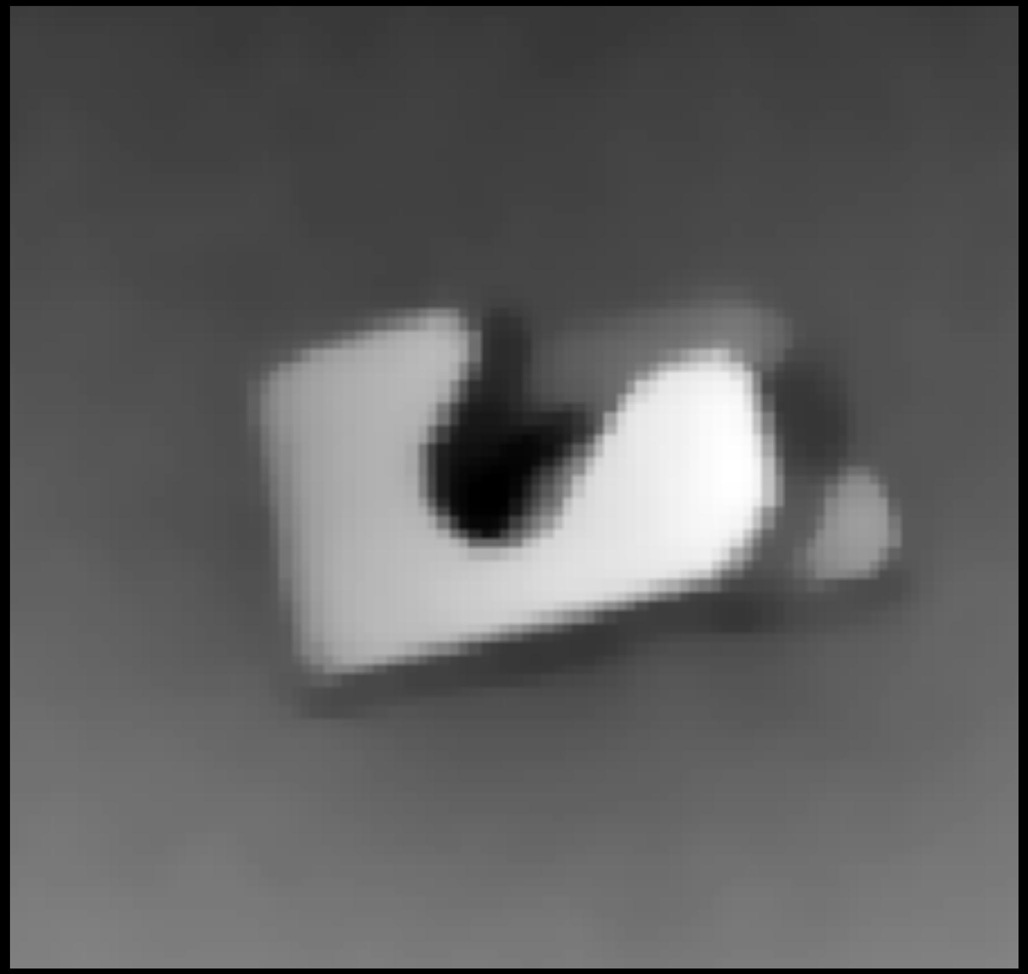


Grasp Quality CNN

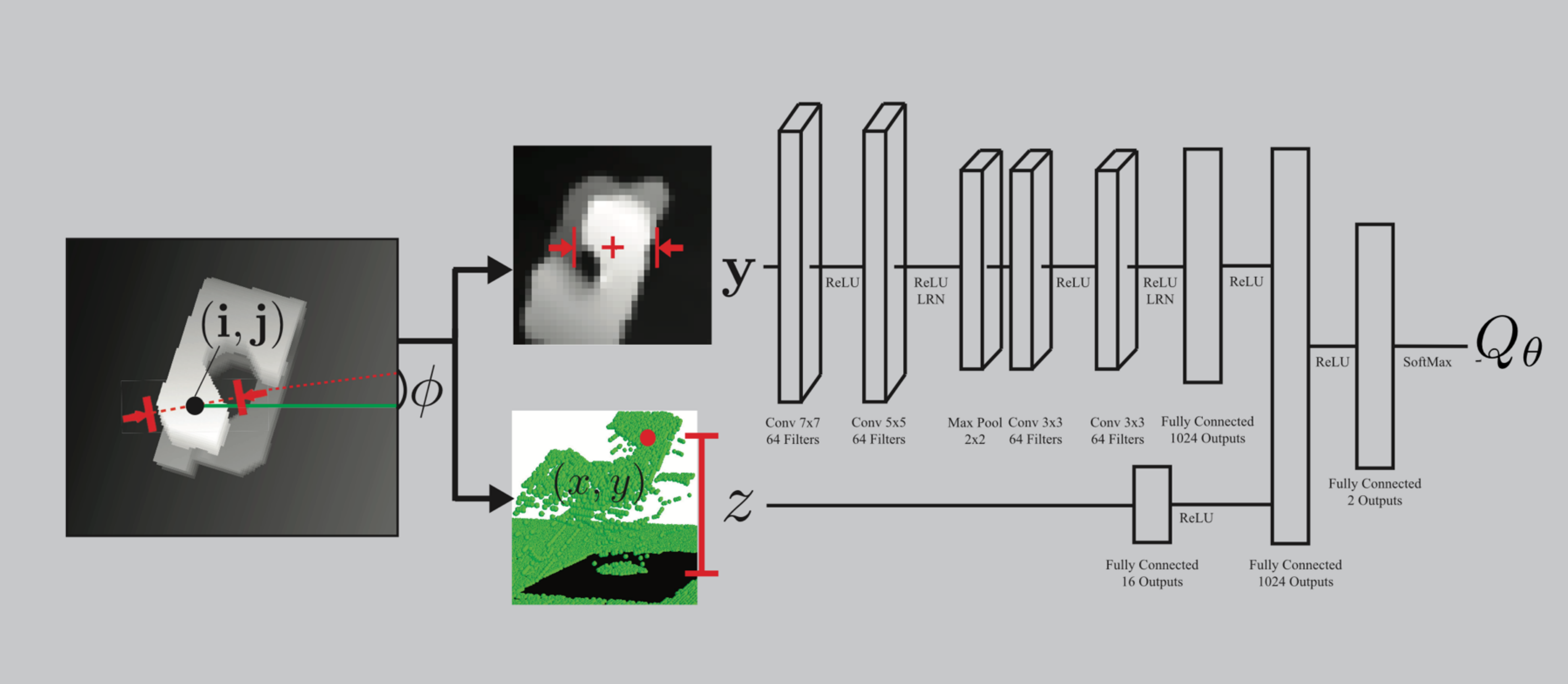


Dex-Net

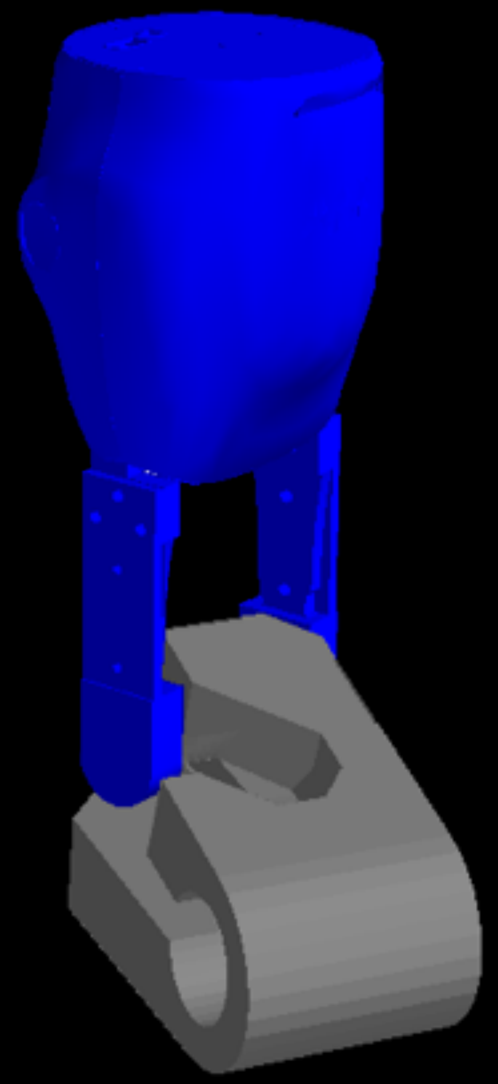
Lidar image



Grasp Quality CNN



Execute best grasp

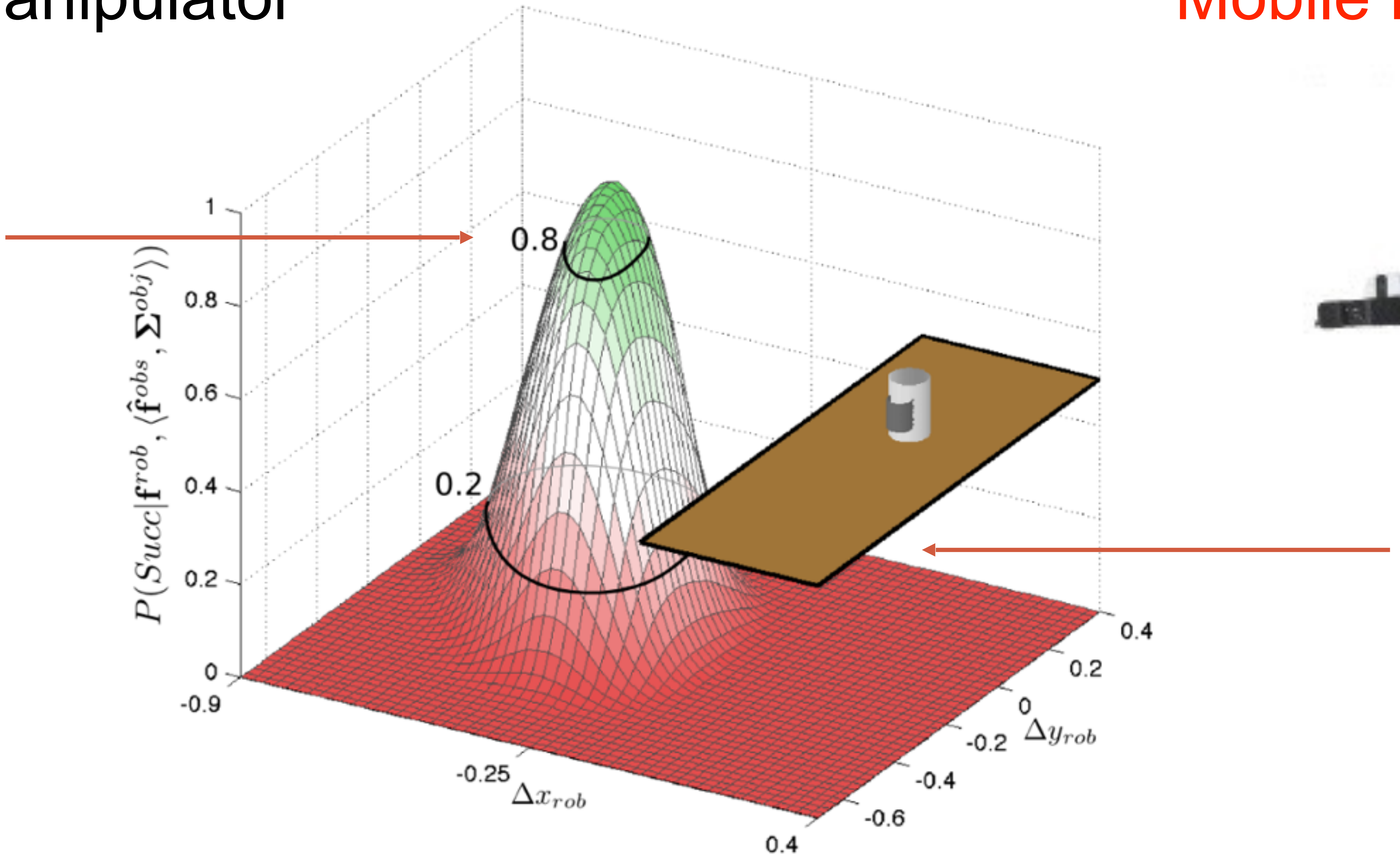



5x
AUTOLAB
Q-tips
Q-tips
750


fruit
screwdriver
wrench
cup
tube
scissors
bottle
hammer
toy
utility
assembly
tape
20x

From Fixed Base to Mobile Manipulators

Fixed Base Industrial Manipulator



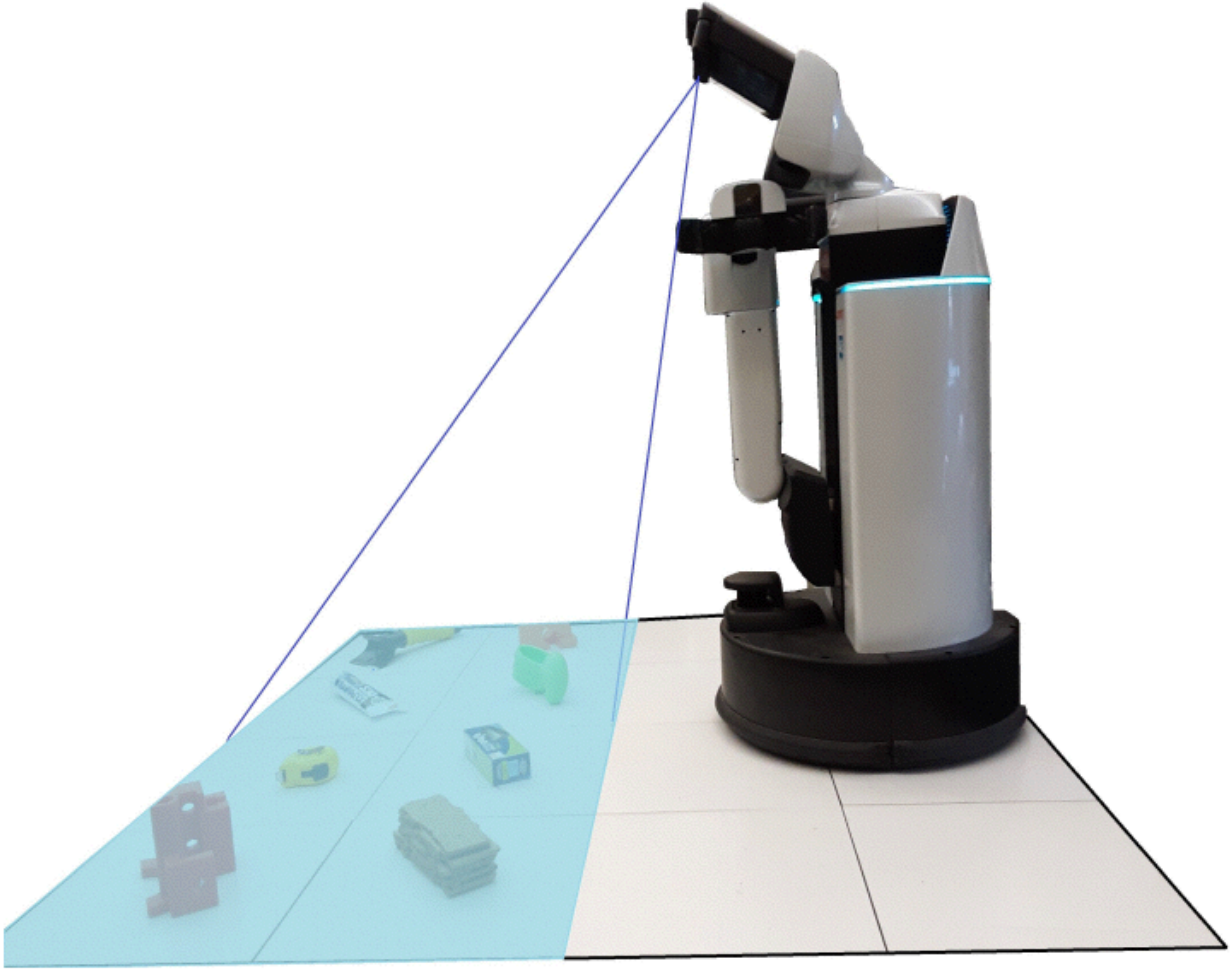
Mobile Manipulator

Lower Resolution, Non-Overhead Depth Sensor
Wider, Less Precise Gripper

Toyota
Human
Support
Robot
(HSR)

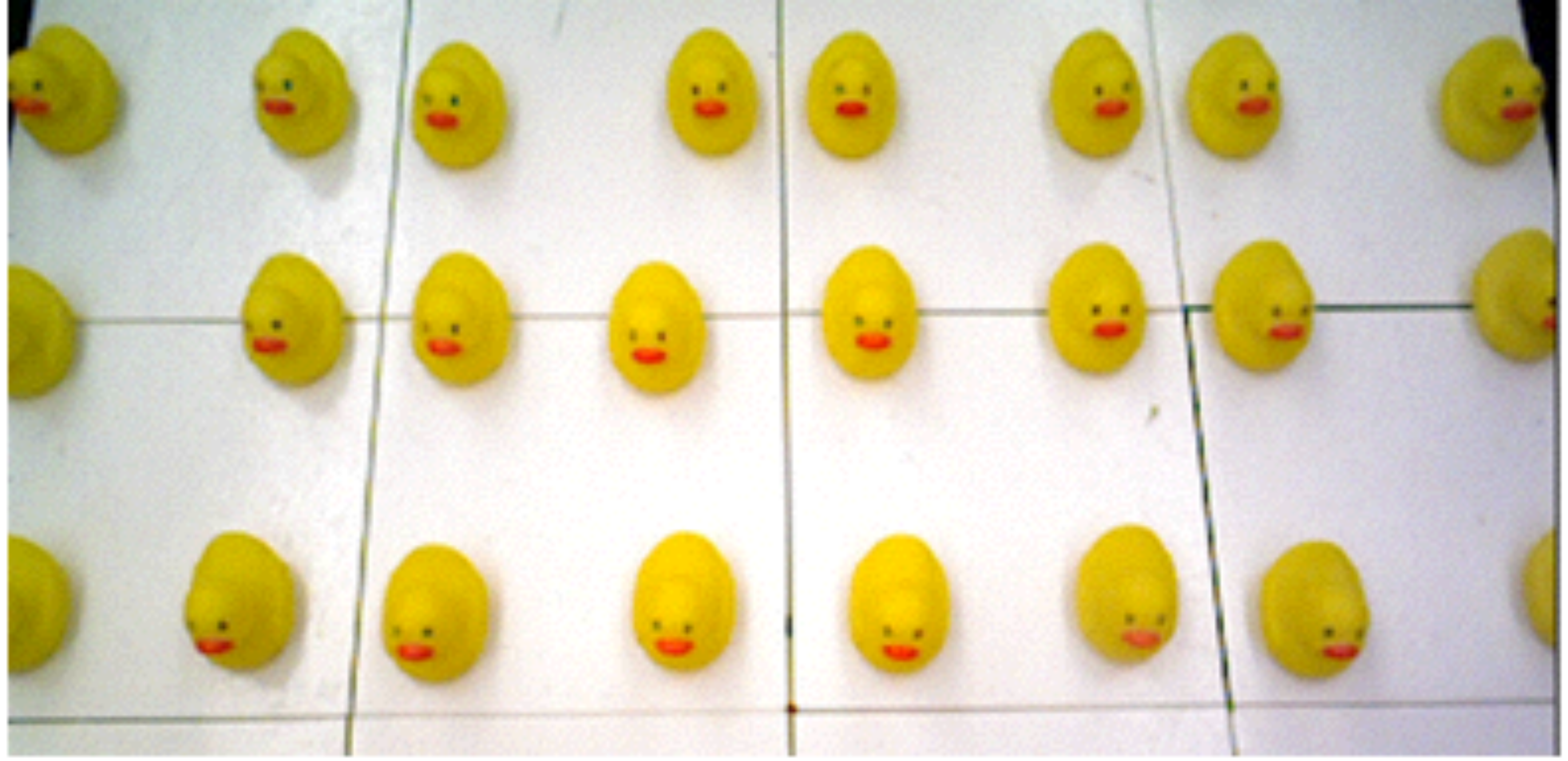
Sensor Configuration not Overhead

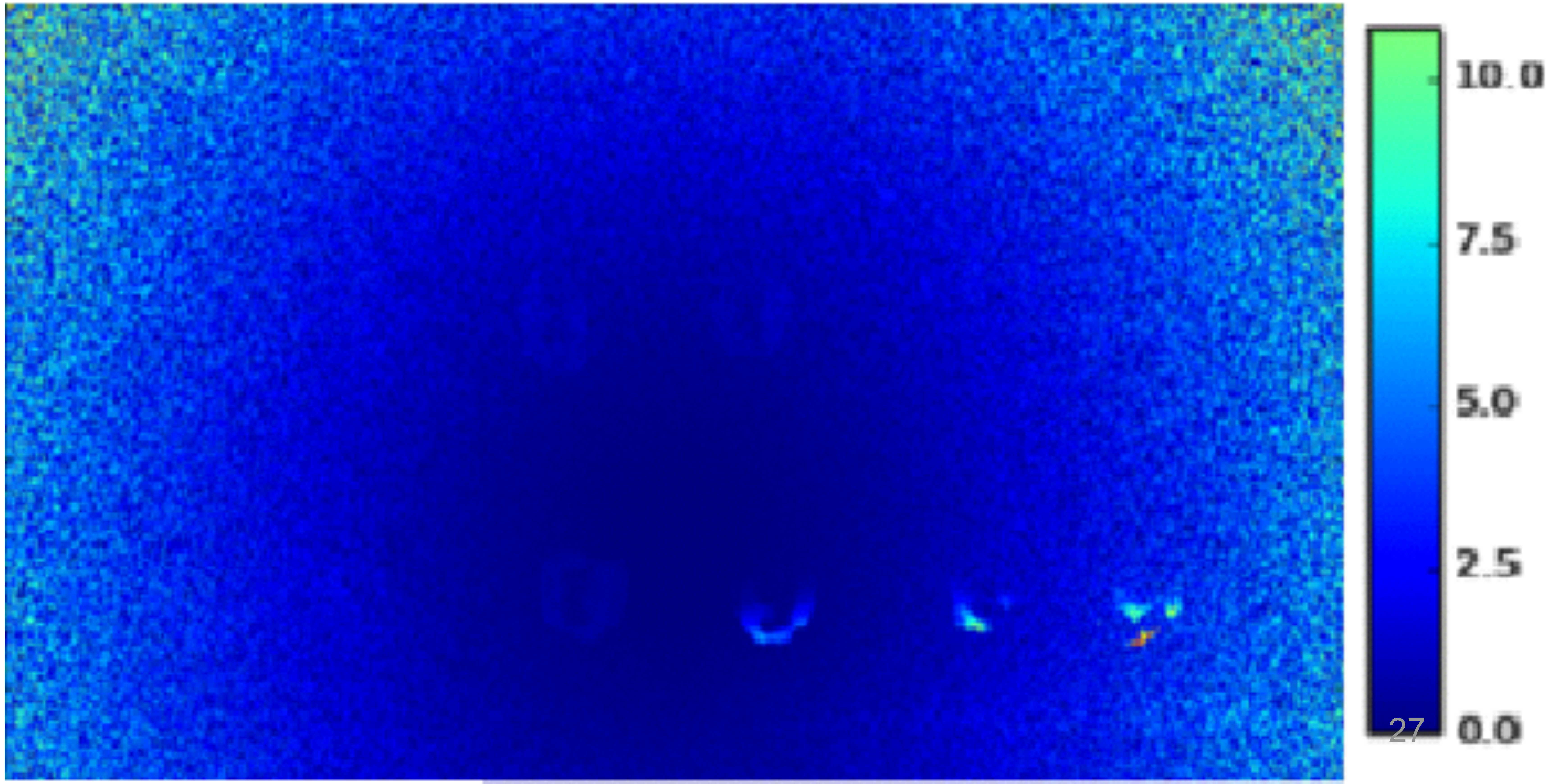
- Camera elevation angle:
- 14 degrees from vertical



RealSense Depth Sensor

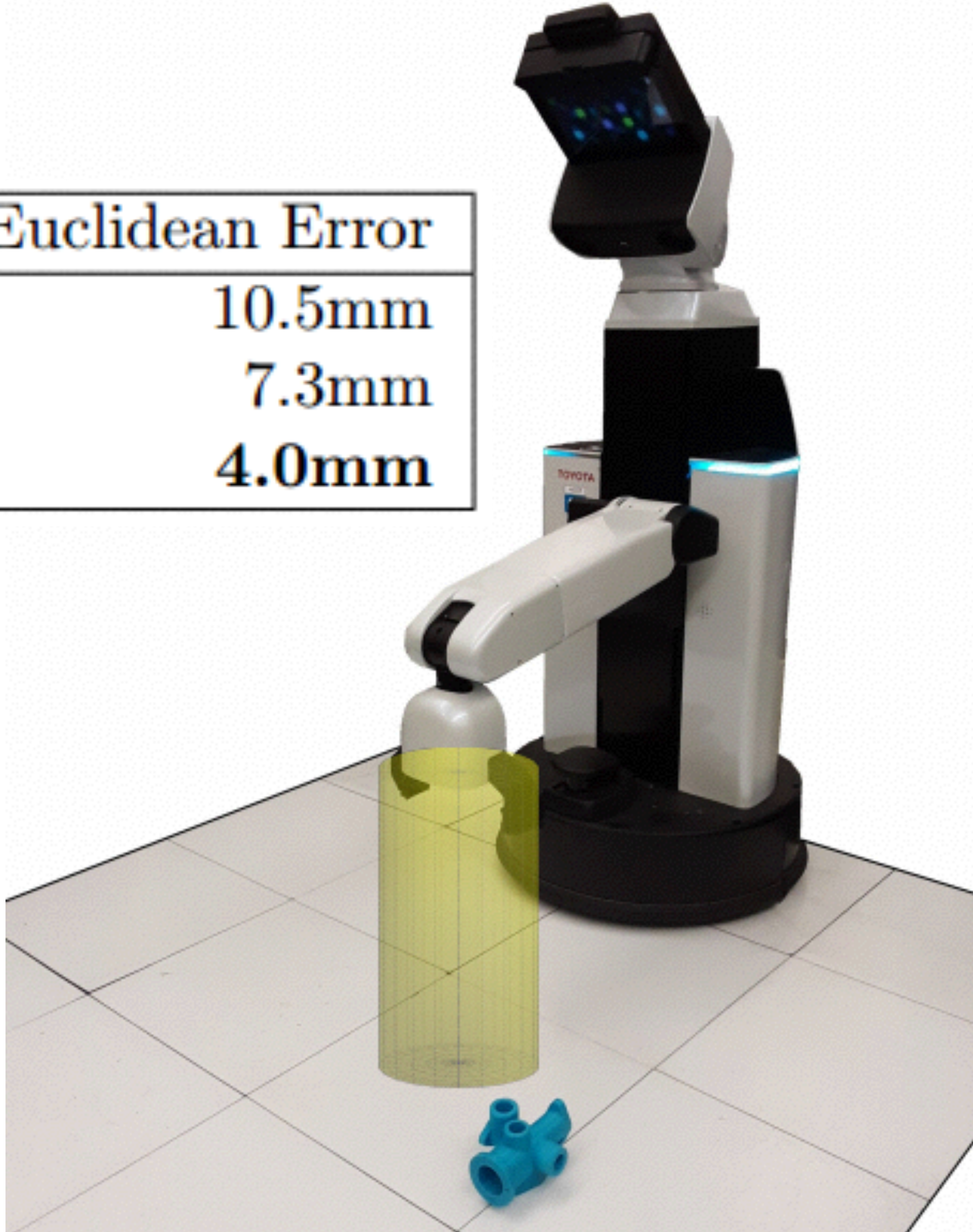
- Depth map noise radially increases outwards up to 10 mm from 2 mm in the center of the image





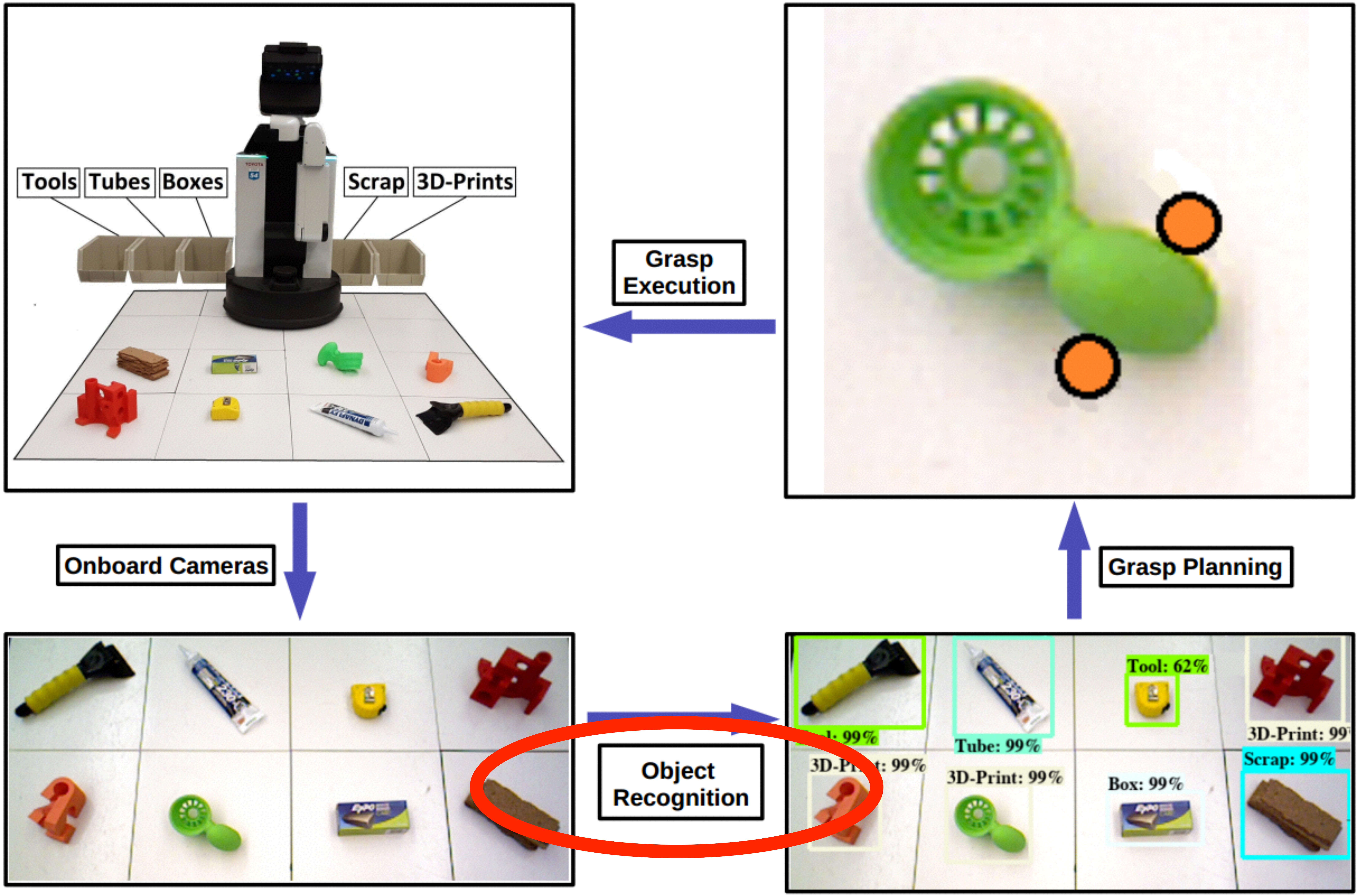
Control Imprecision

Controller	Error in X	Error in Y	Euclidean Error
τ_{IK}	6.9mm	7.9mm	10.5mm
$\tau_{IK,tuned}$	5.9mm	4.3mm	7.3mm
τ_{cyl}	**2.6mm**	**3.0mm**	**4.0mm**



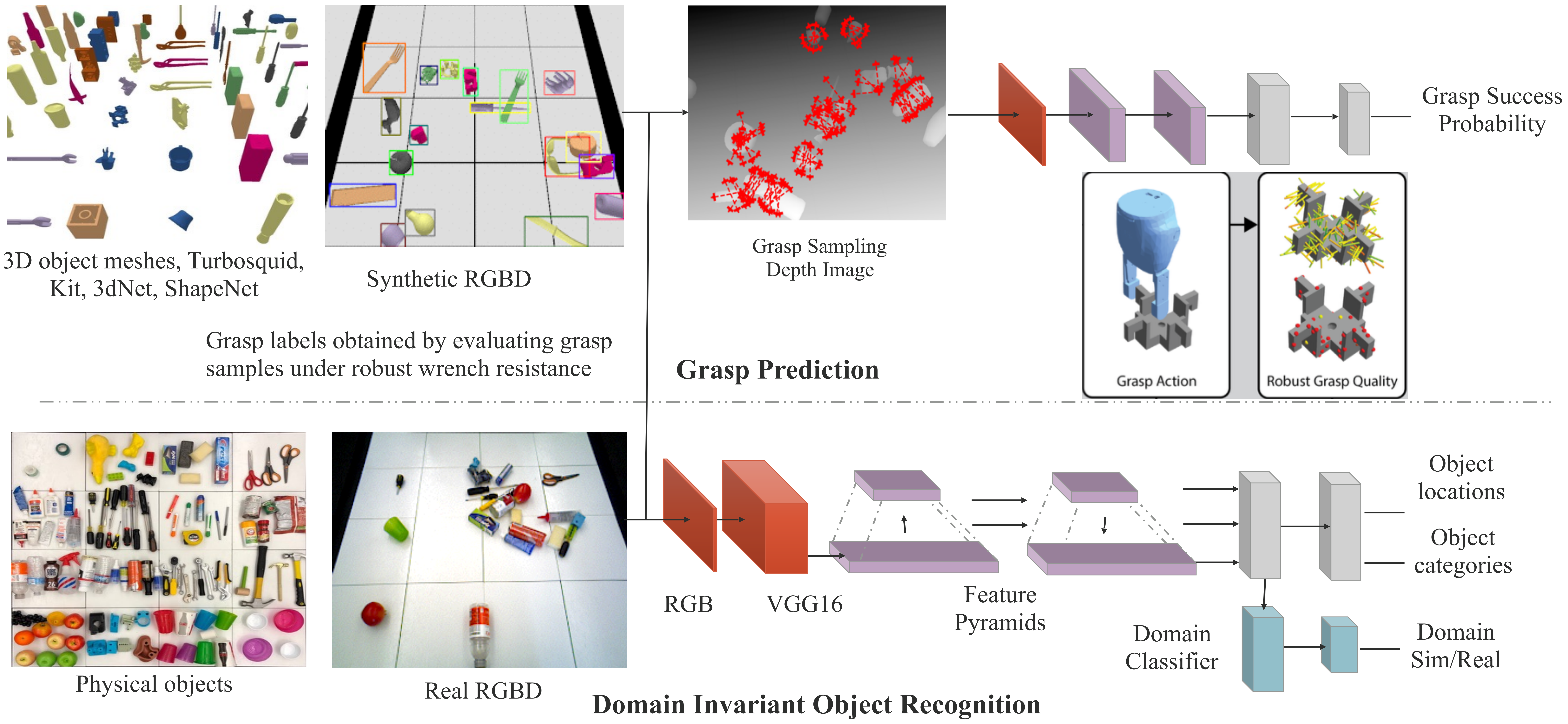
Conjecture:

- Retrain Dex-Net with HSR Sensing, Gripper, and Precision parameters.
- Combine with Deep Learning Image Classifier for Multiple Bins: a composite policy for surface decluttering.
- Adapt Dex-Net 4.0 grasping policy to the HSR low precision Mobile Manipulator: **Dex-Net MM**

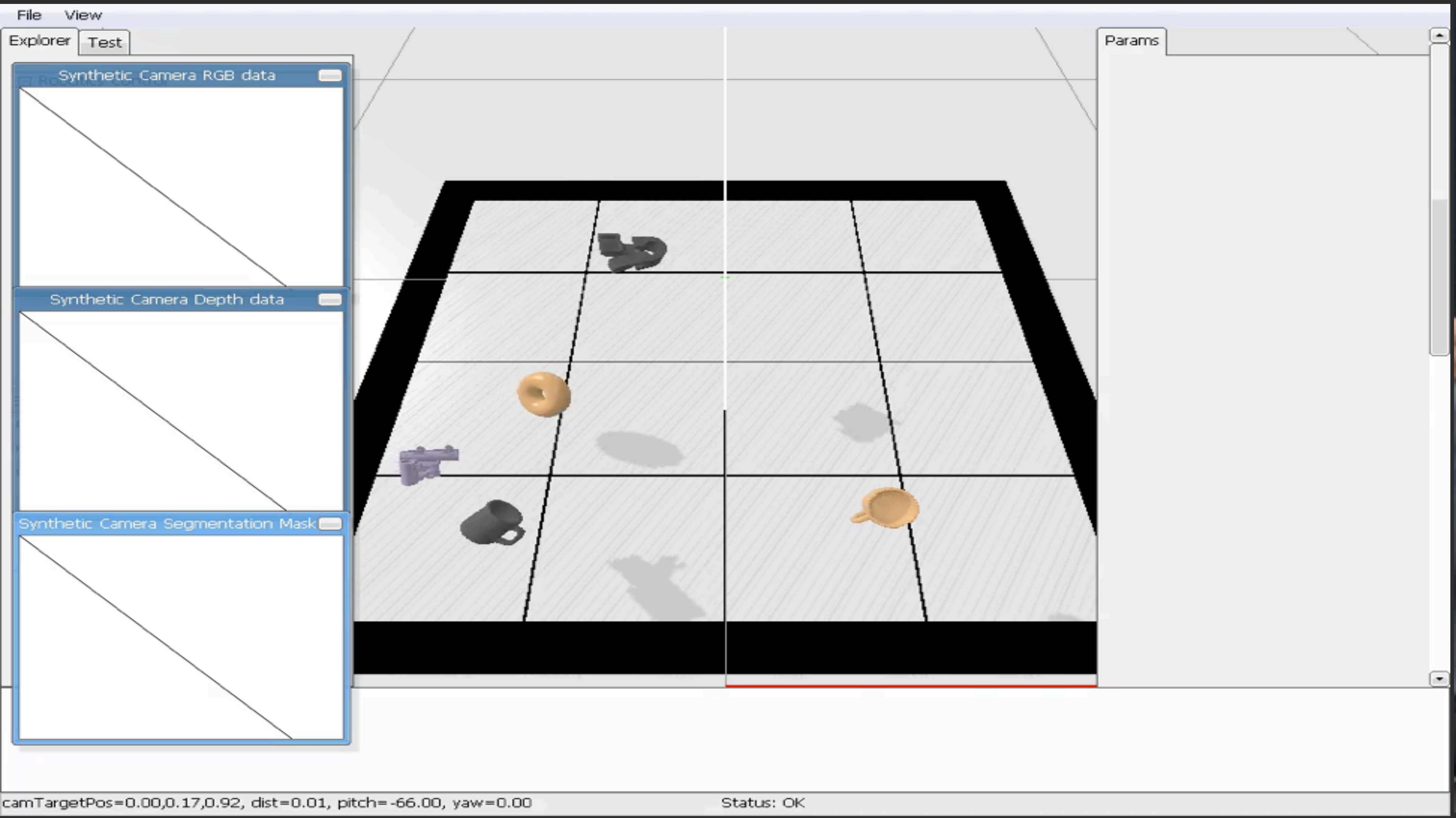


Tools
Tubes
Boxes
Scrap
3D-Prints
Grasp Execution
Onboard Cameras
Grasp Planning
Object Recognition
Tool: 62%
Tube: 99%
3D-Print: 99%
Box: 99%
Scrap: 99%

Deep Object Recognition



A.K. Tanwani, N. Mor, J. Kubiatowicz, J. Gonzalez, K. Goldberg. "A Fog Robotics Approach to Deep Robot Learning: Application to Object Recognition and Grasp Planning in Surface Decluttering", IEEE International Conference on Robotics and Automation (ICRA), 2019.



Sim Learning

- Modeling Inaccuracies
- Adaptation to real
- Synthetic (public) data

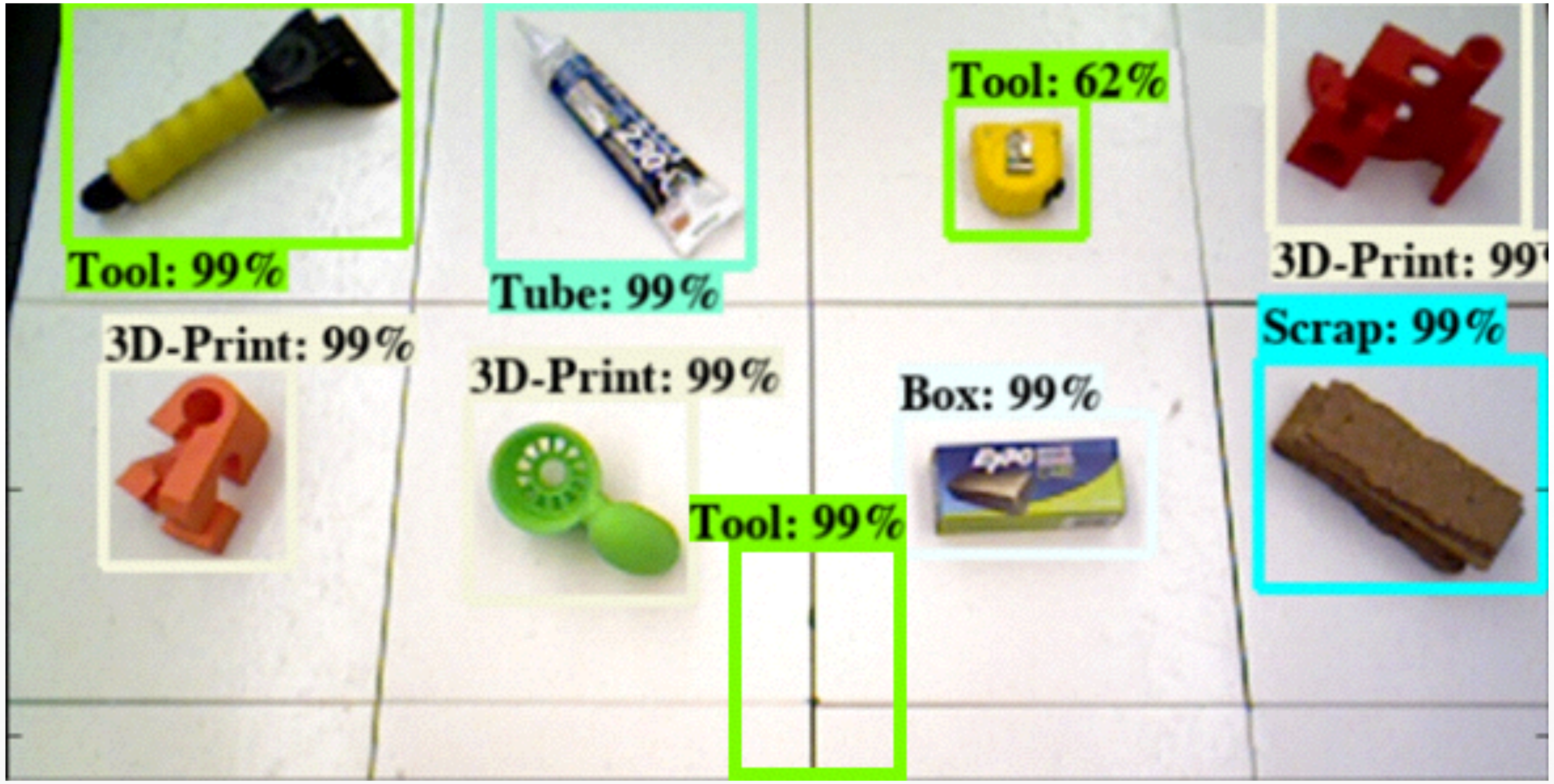
Real Learning

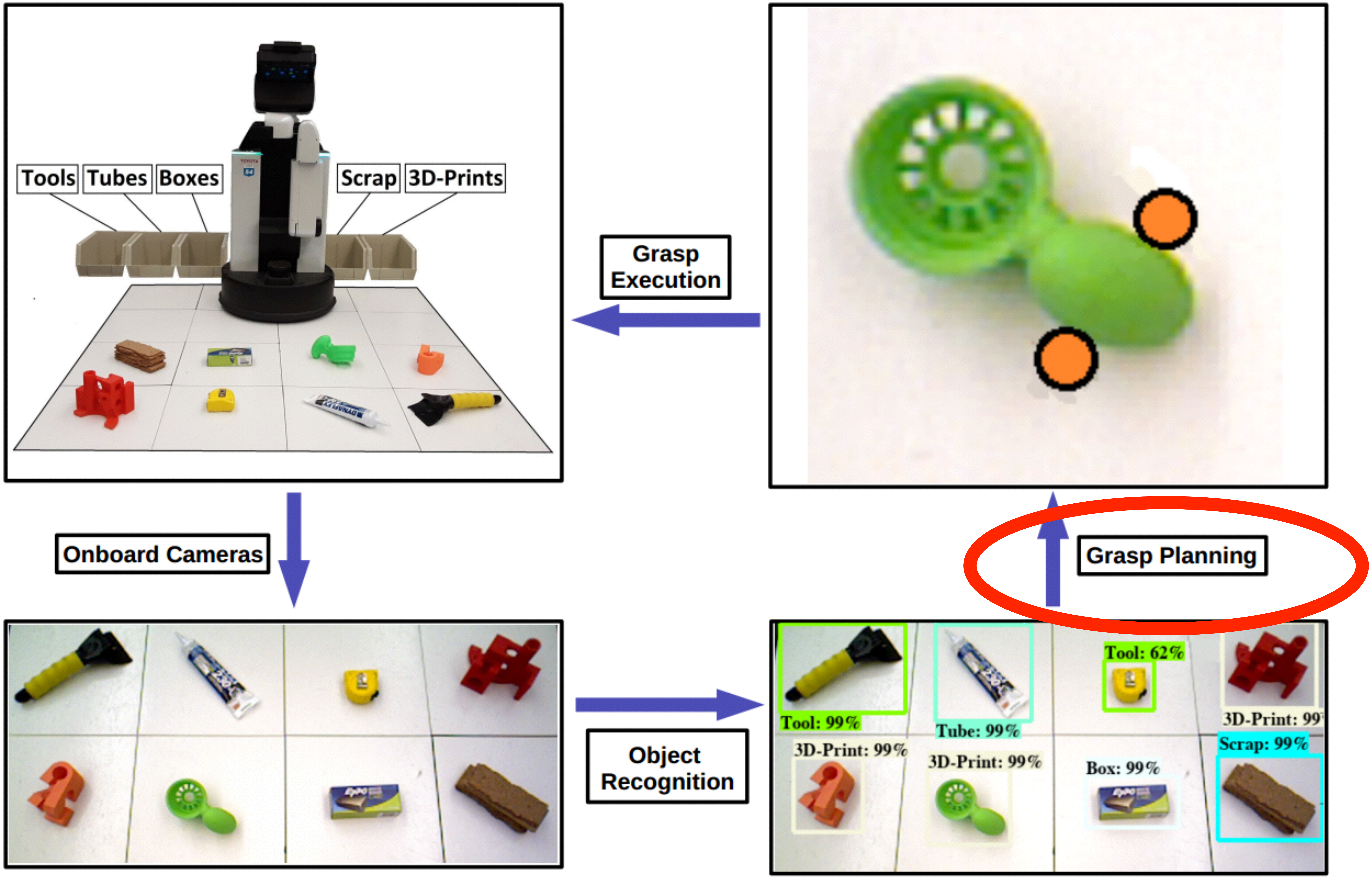
- On-premise
- High sample and time complexity
- Privacy and security

fruit: 89%
toy: 75%
fruit: 99%
toy: 60%
assemblyPart: 99%
assemblyPart: 99%
cup: 96%
fruit: 9%
wrench: 47%
hammer: 98%
screwdriver: 98%
assemblyPart: 99%
tube: 98%
bottle: 66%
cup: 99%
cup: 99%
fruit: 99%

Object Recognition

- Domain Invariant Object Recognition (DIOR) model
  - Sim-to-real transfer
  - 20k synthetic images
  - 200 real images

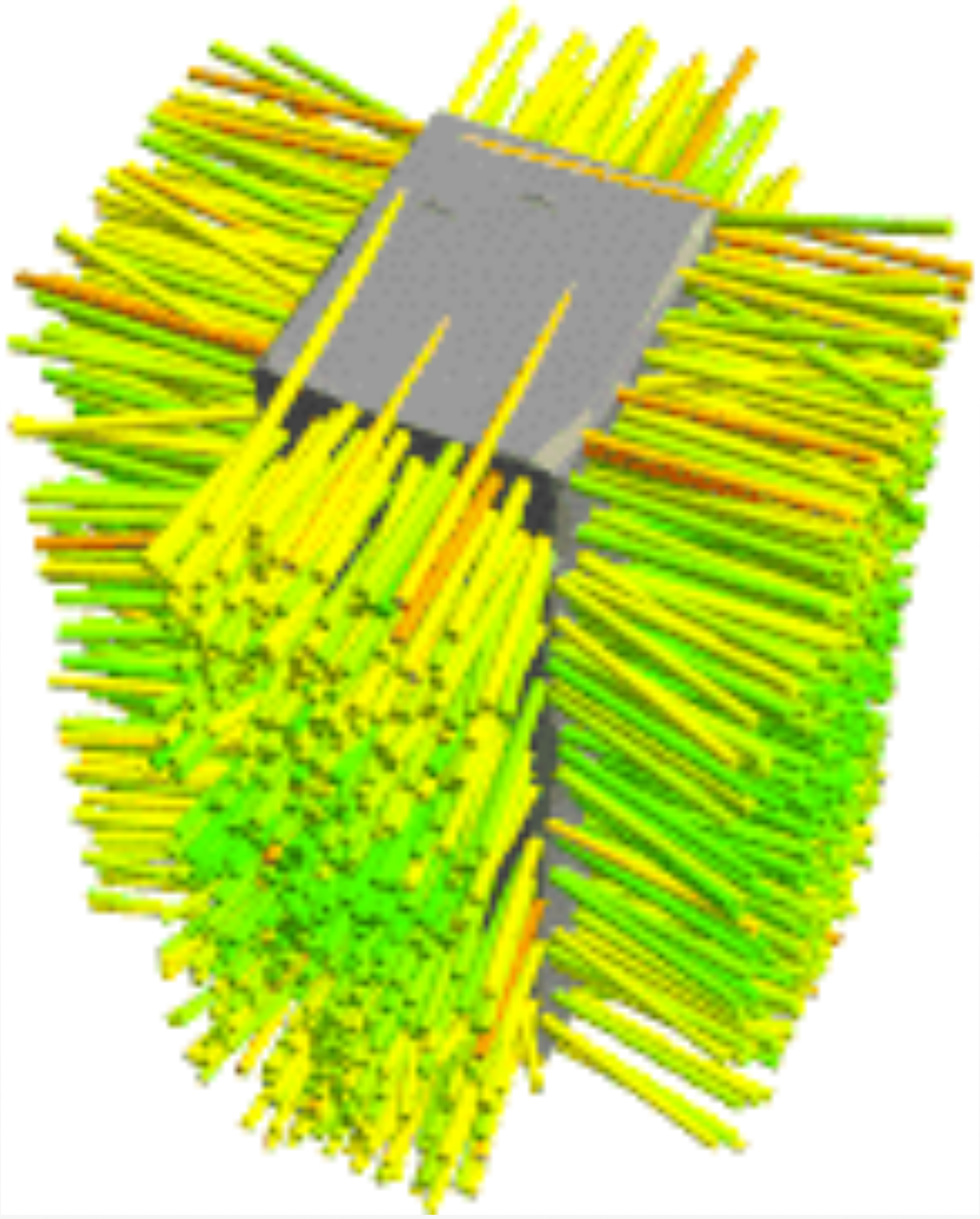


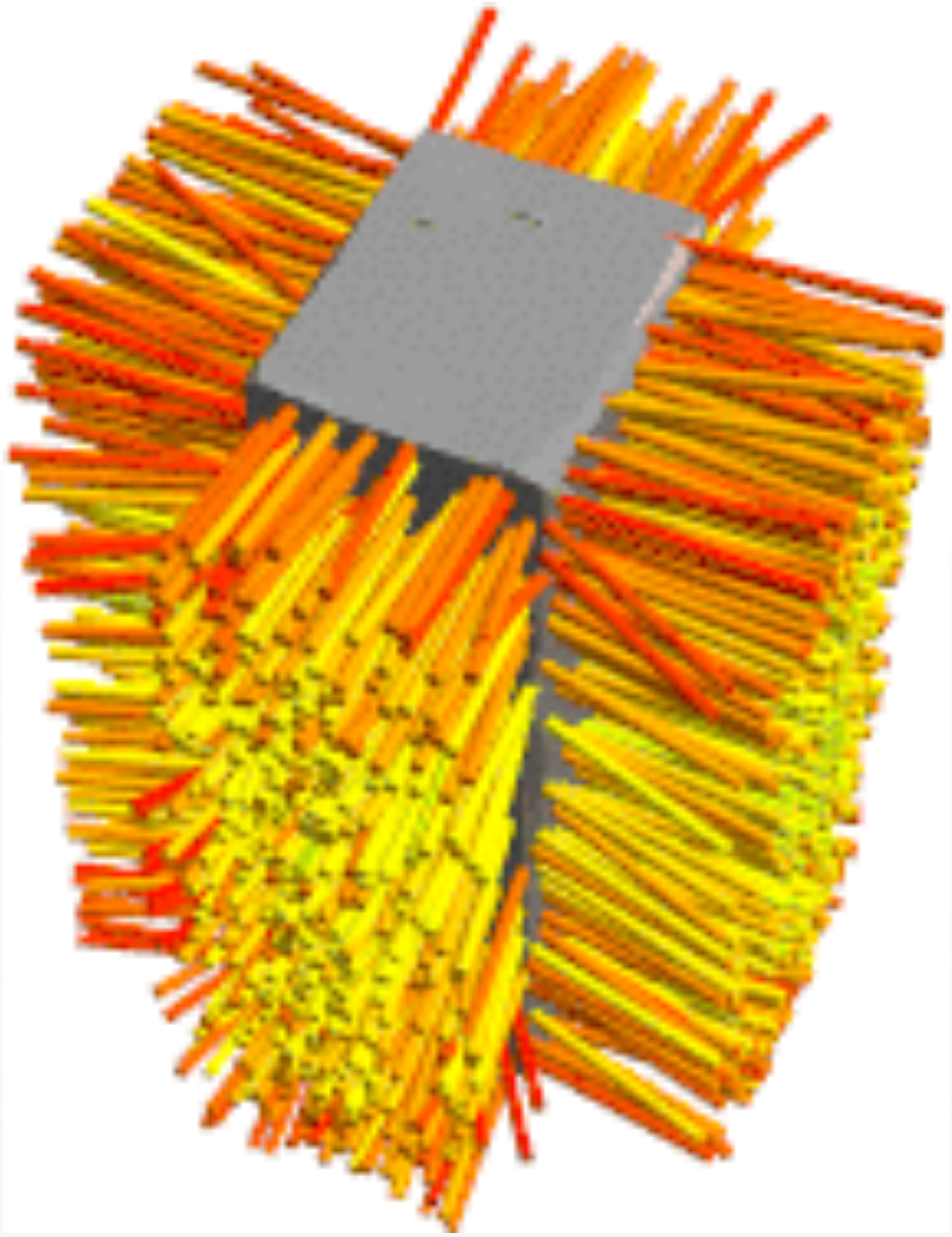

Tools
Tubes
Boxes
Scrap
3D-Prints
Grasp Execution
Onboard Cameras
Object Recognition
Grasp Planning
Tool: 99%
Tube: 99%
Tool: 62%
3D-Print: 99
3D-Print: 99%
3D-Print: 99%
Box: 99%
Scrap: 99%

Dex-Net MM Dataset

- Camera elevation angle [12-17]°
- HSR camera
- 3-10 objects per scene
- 1000 scenes
- 2.6M datapoints

- Effect of increasing motion imprecision on grasp labels in the training dataset on reducing grasp quality

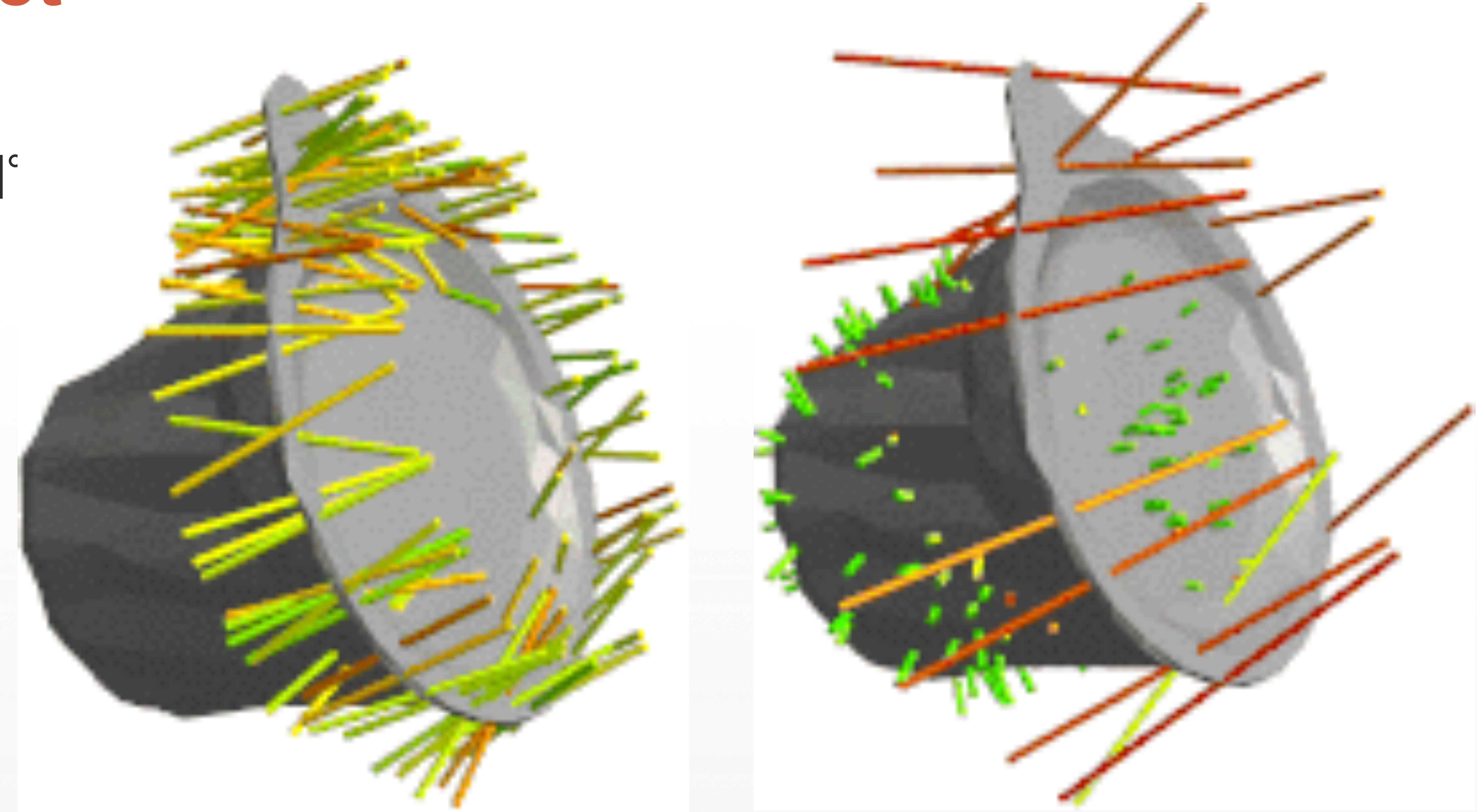




Dex-Net MM Dataset

- Camera elevation angle [12-17]°
- HSR camera
- 3-10 objects per scene
- 1000 scenes
- 2.6M datapoints

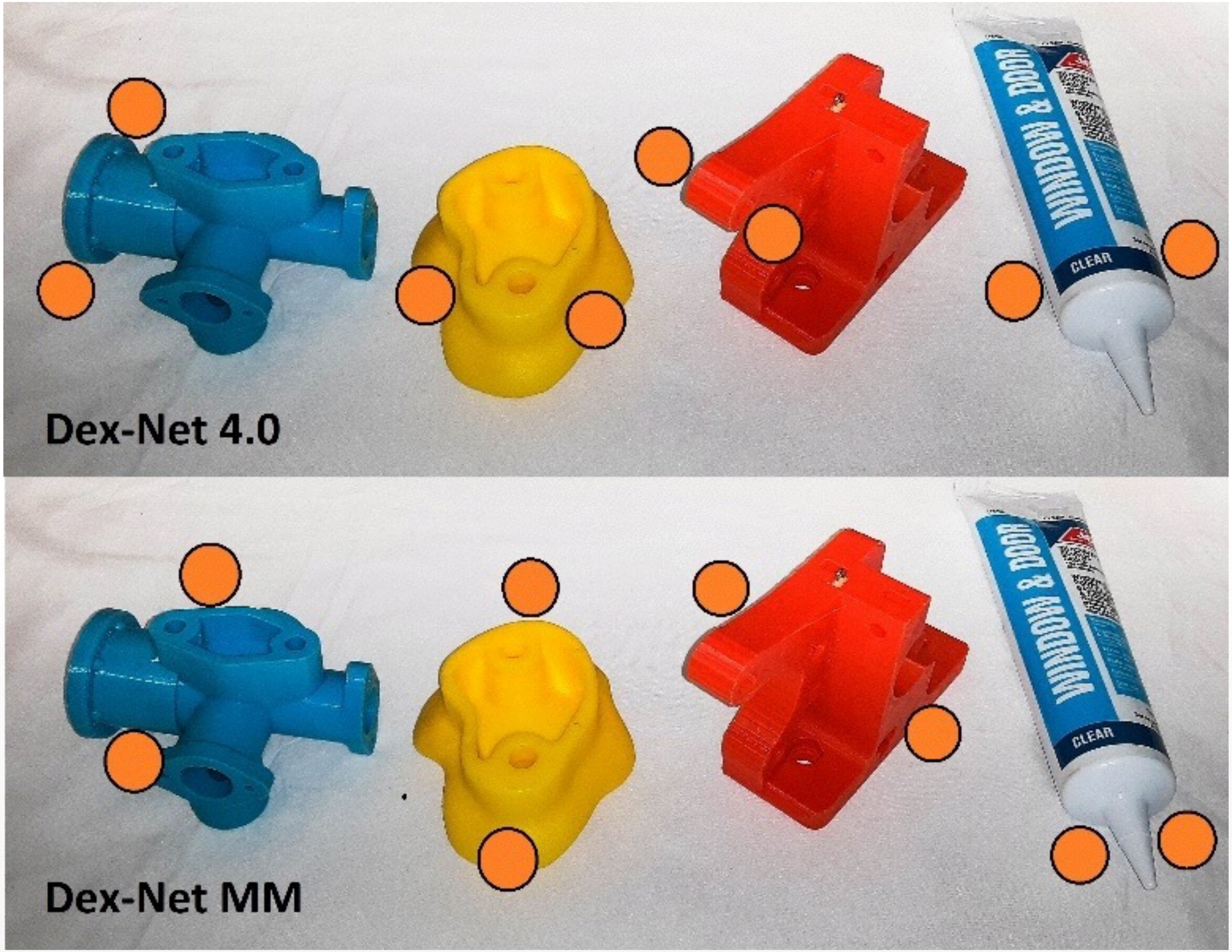
- Effect of increasing gripper throw



Dex-Net 4.0

vs.

Dex-Net MM



Surface Decluttering Setup

Experiments: Surface Decluttering

Object Categories

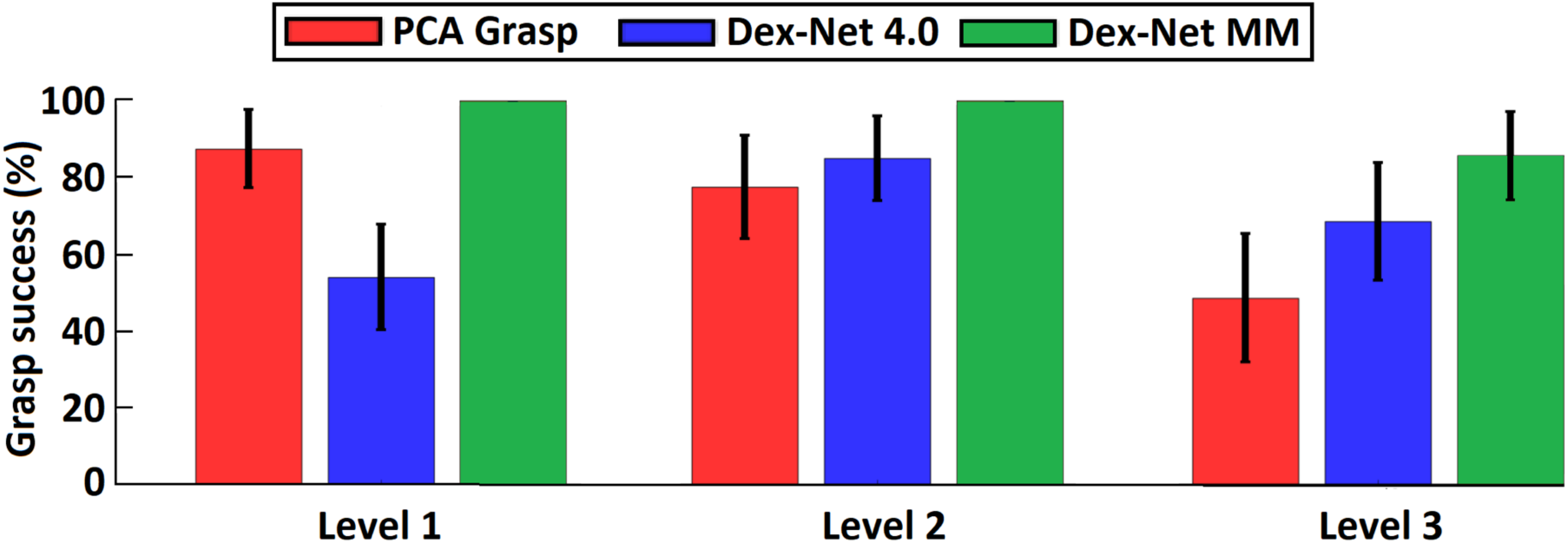
1
easy

2:
typical

3:
adversarial

4:
pathological

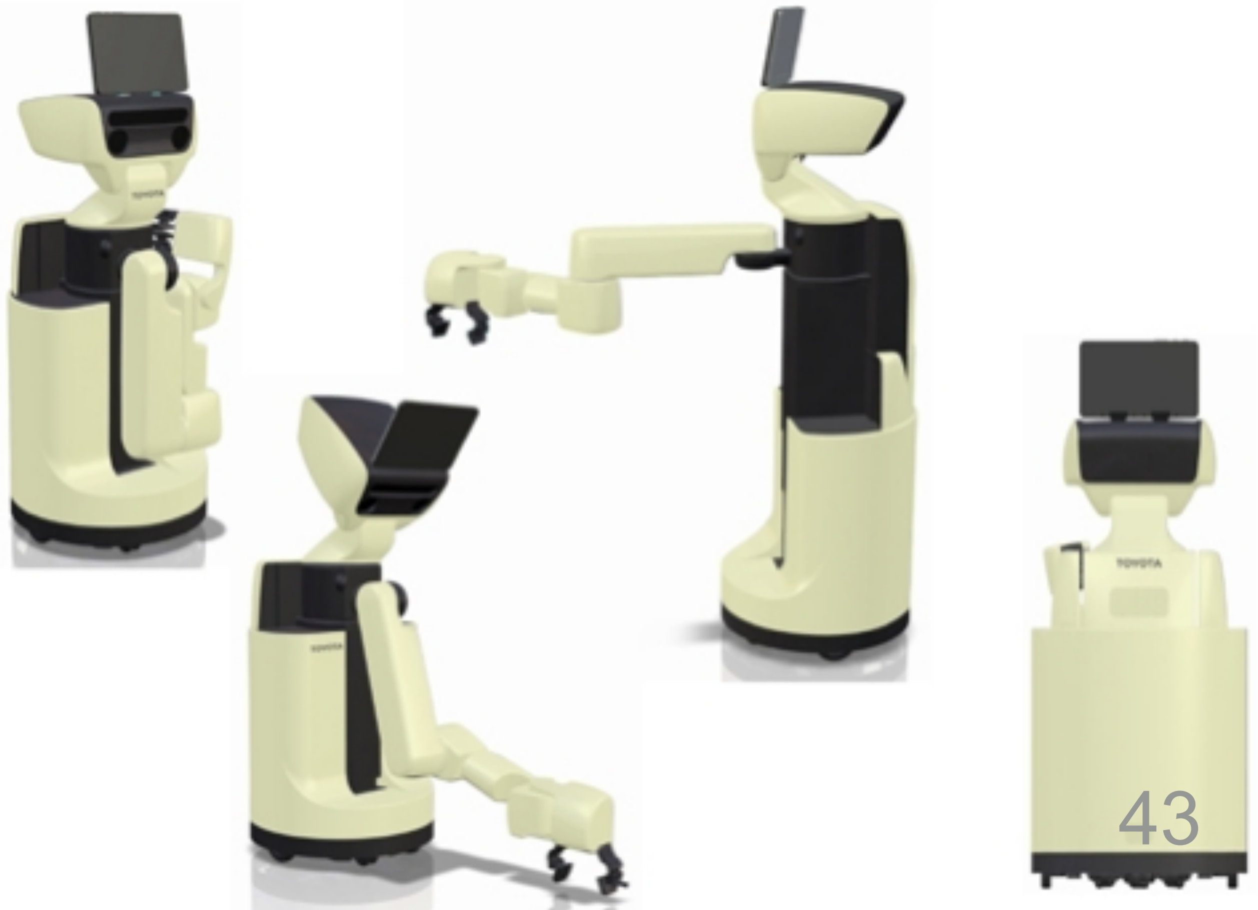
1 2: 3: 4:


PCA Grasp
Dex-Net 4.0
Dex-Net MM
Grasp success (%)
100
80
60
40
20
0
Level 1
Level 2
Level 3

Cloud

Edge:

Robots:

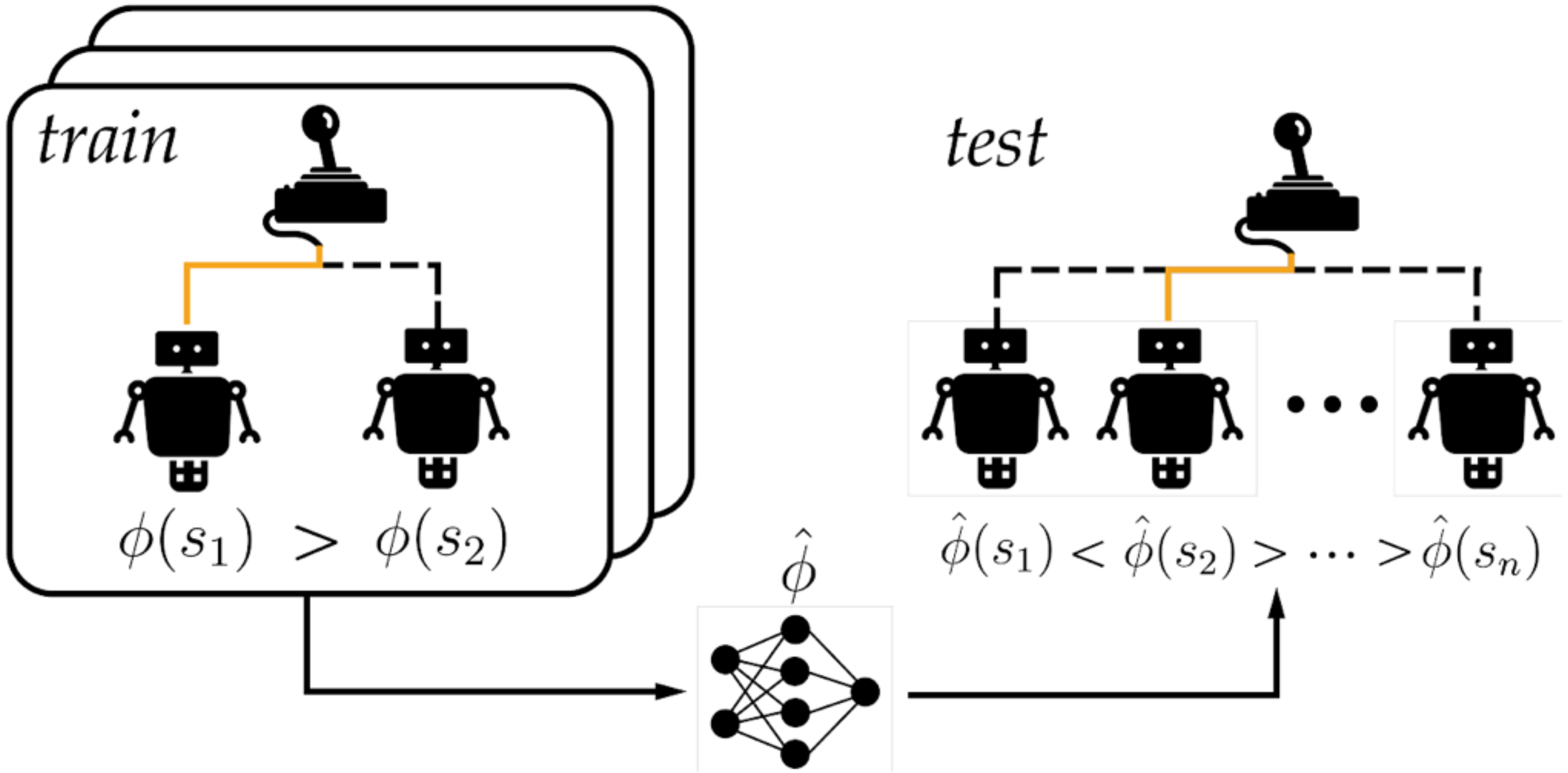


Scaled Autonomy: Enabling Human Operators to Control Robot Fleets

Gokul Swamy, Siddharth Reddy, Sergey Levine, Anca D. Dragan, ICRA 2020.

(Optional) Step 0: *Train baseline policy or learn from demos.*

Step 1: *Let users freely choose which of a few robots to teleop.*



Step 2: *To learn a function that generalizes to an arbitrary number of robots, we model the user as making rational choices to maximize utility across all robots:*

$$\phi = \arg\max_{\phi\in\Phi} \mathbb{E}\left[\sum_{i=1}^{n}\sum_{t=0}^{T-1} R(s_i^t, a_i^t) \mid \pi_H, \pi_R\right] \qquad \mathbb{P}[i_H^t = i] = e^{\phi(s_i^t)} \Big/ \sum_{j=1}^{n} e^{\phi(s_j^t)}$$

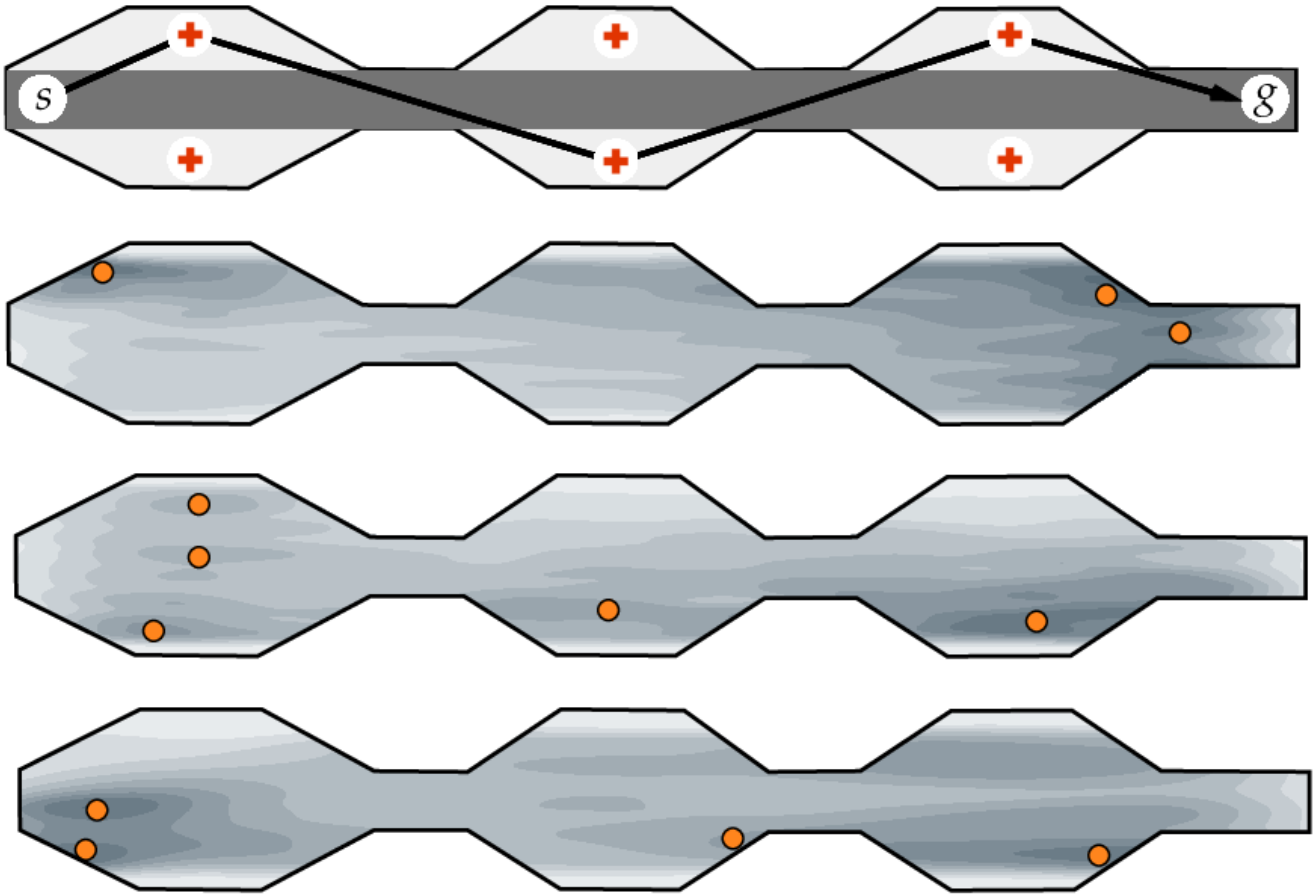
Then, we train a network to mimic these choices by maximizing the log likelihood:

$$\ell(\boldsymbol{\theta};\mathscr{D}) = \sum_{(s_1^t,\dots,s_n^t,i_H^t)\in\mathscr{D}} -\phi_{\boldsymbol{\theta}}\left(s_{i_H^t}\right) + \log\left(\sum_{j=1}^{n} e^{\phi_{\boldsymbol{\theta}}(s_j^t)}\right)$$

Step 3: *Take the argmax over the learned function to automatically choose a robot for the user.*

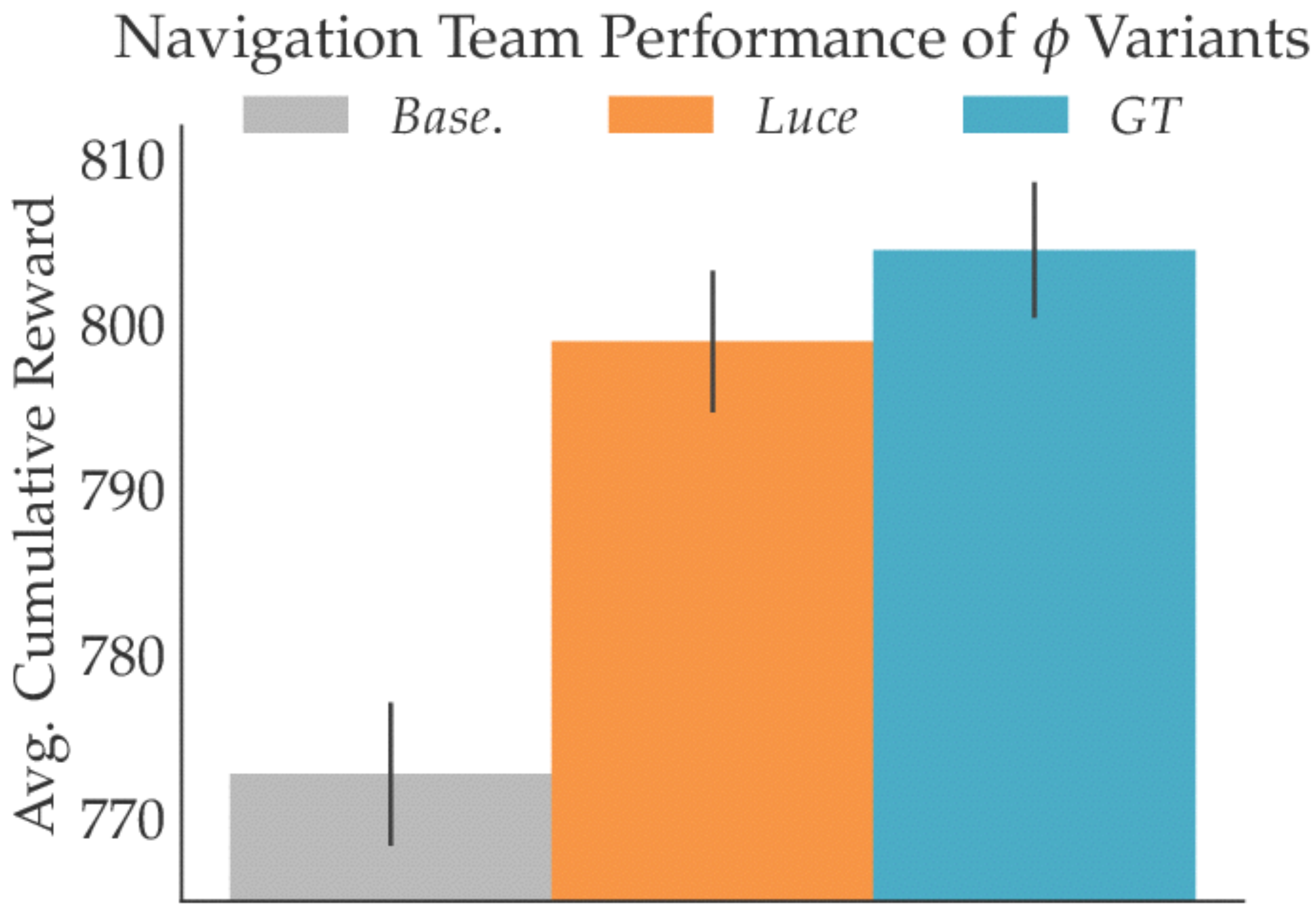
Step 4: *Use collected data to improve robot policy.*

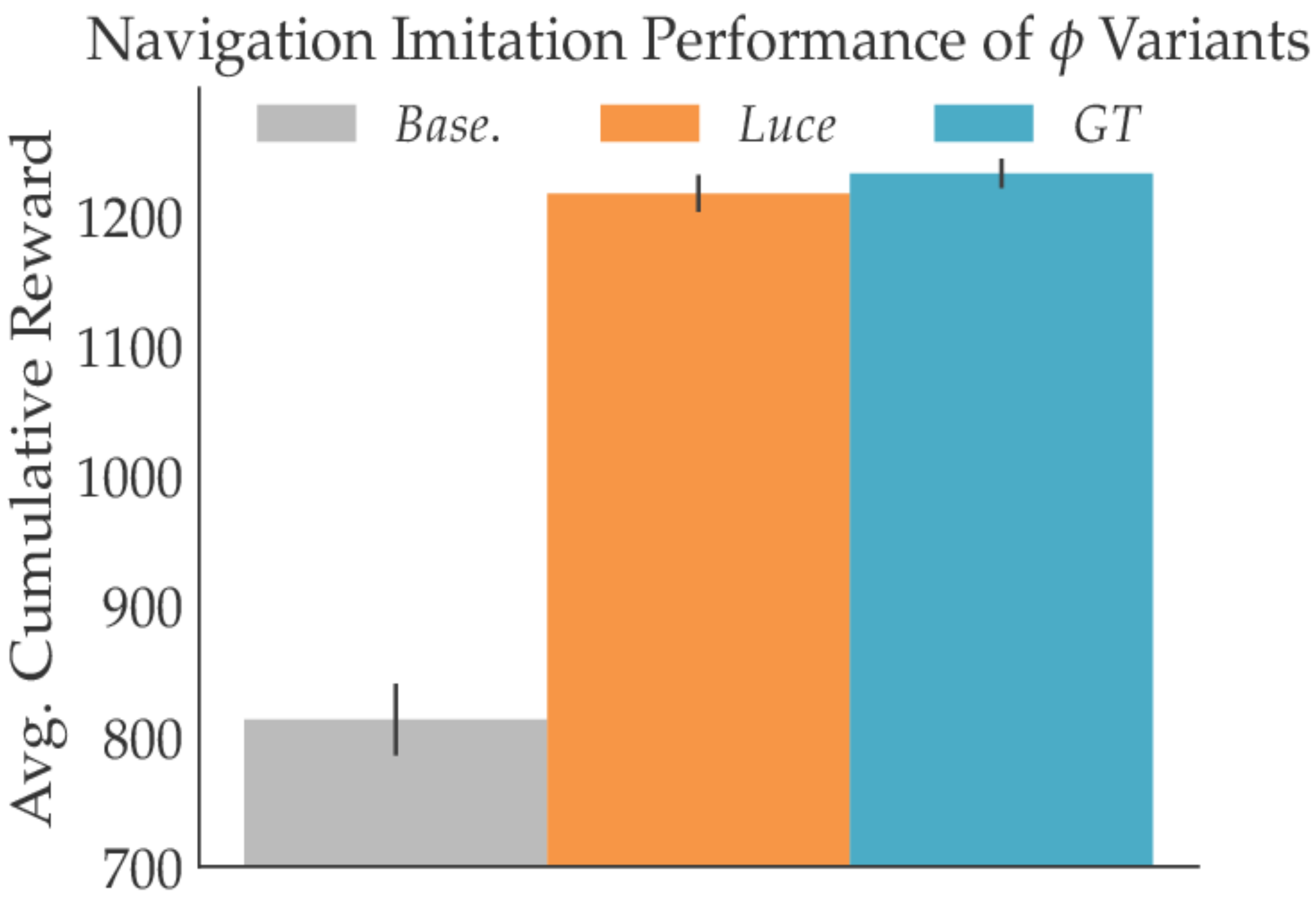
Go back to Step 1


s
g

Scaled Autonomy: Enabling Human Operators to Control Robot Fleets

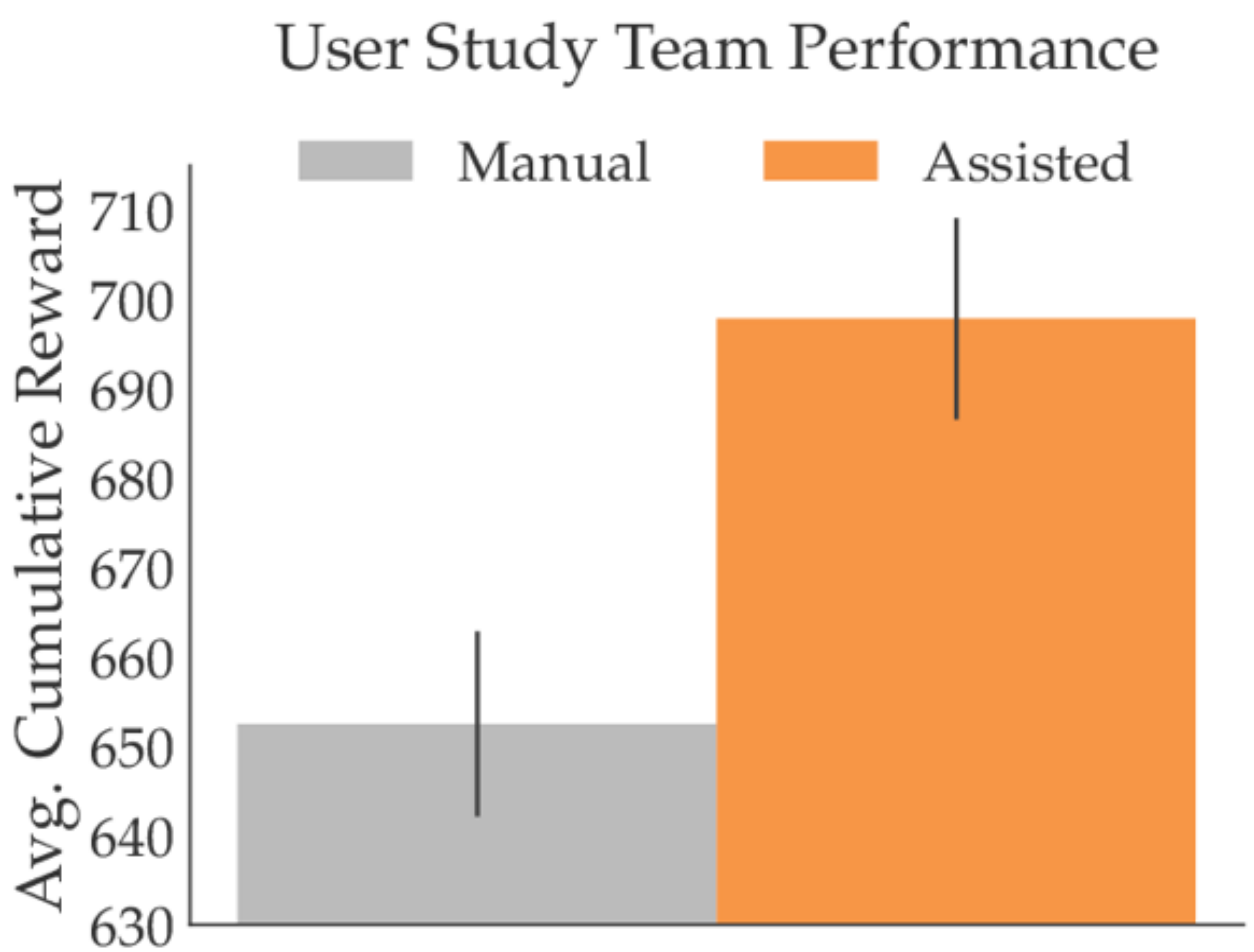
Gokul Swamy, Siddharth Reddy, Sergey Levine, Anca D. Dragan, ICRA 2020.



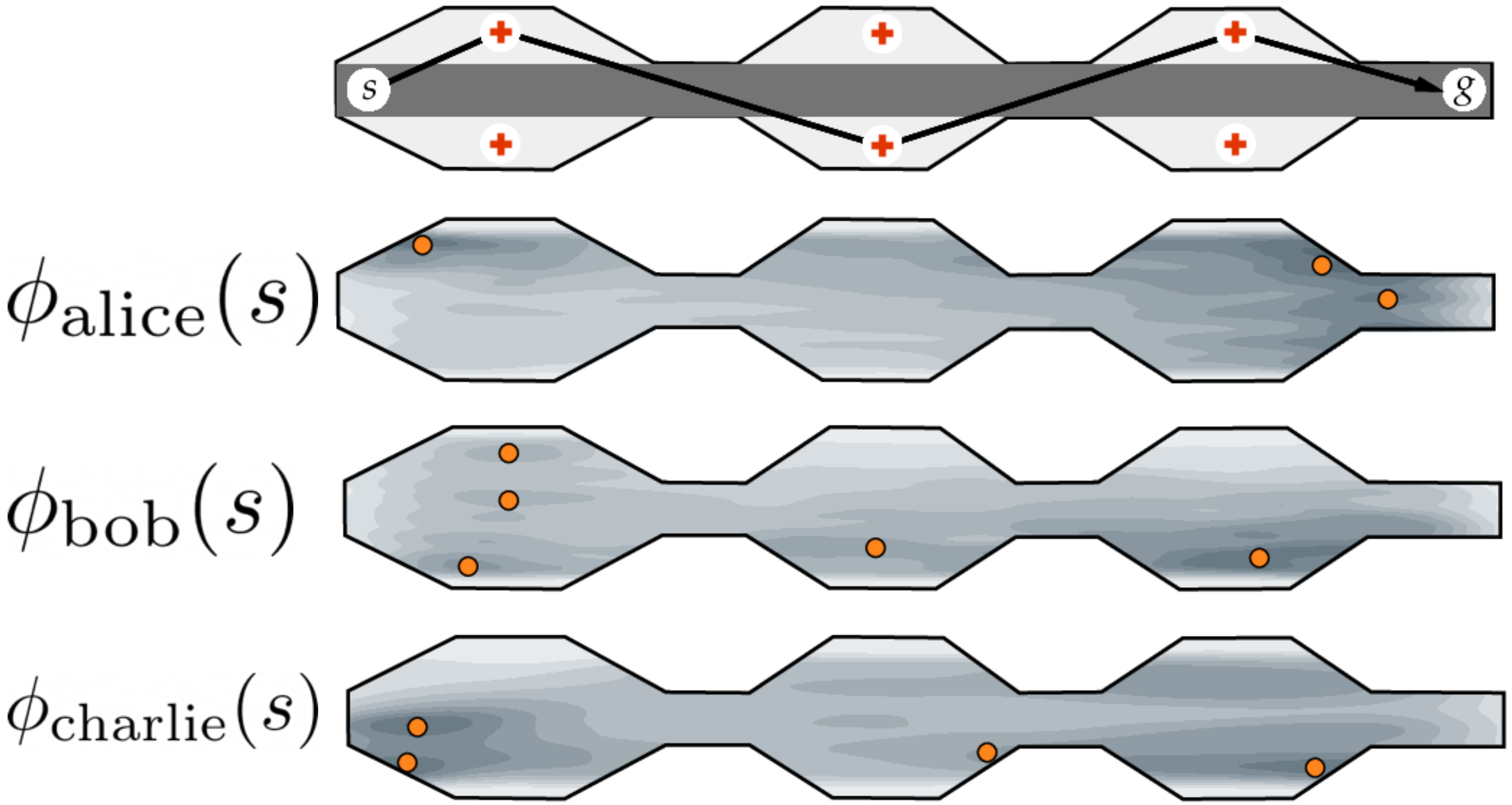


Scaled Autonomy: Enabling Human Operators to Control Robot Fleets

Gokul Swamy, Siddharth Reddy, Sergey Levine, Anca D. Dragan, ICRA 2020.



Real User + Assisted Choice
(n=12 human subjects)

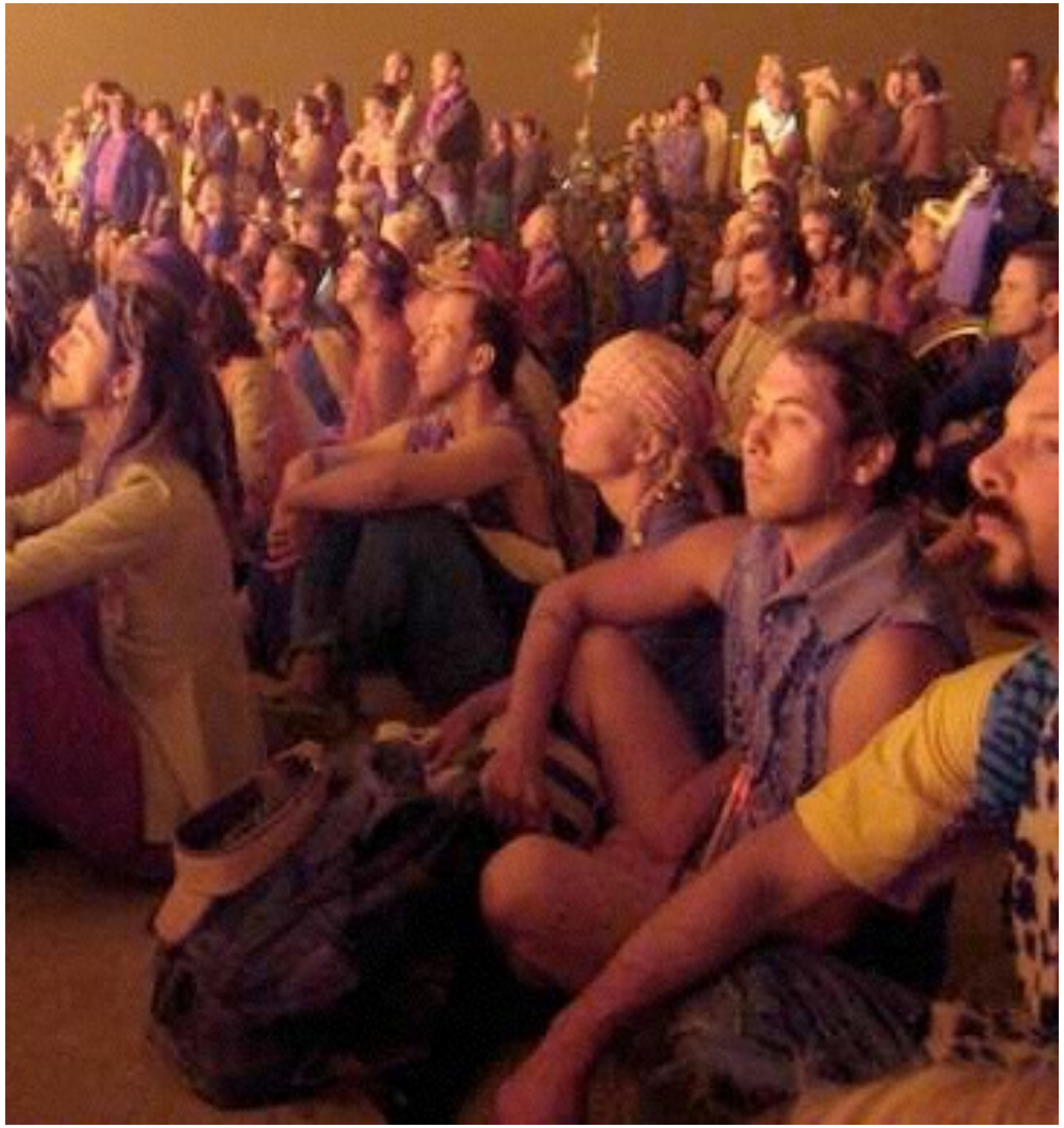


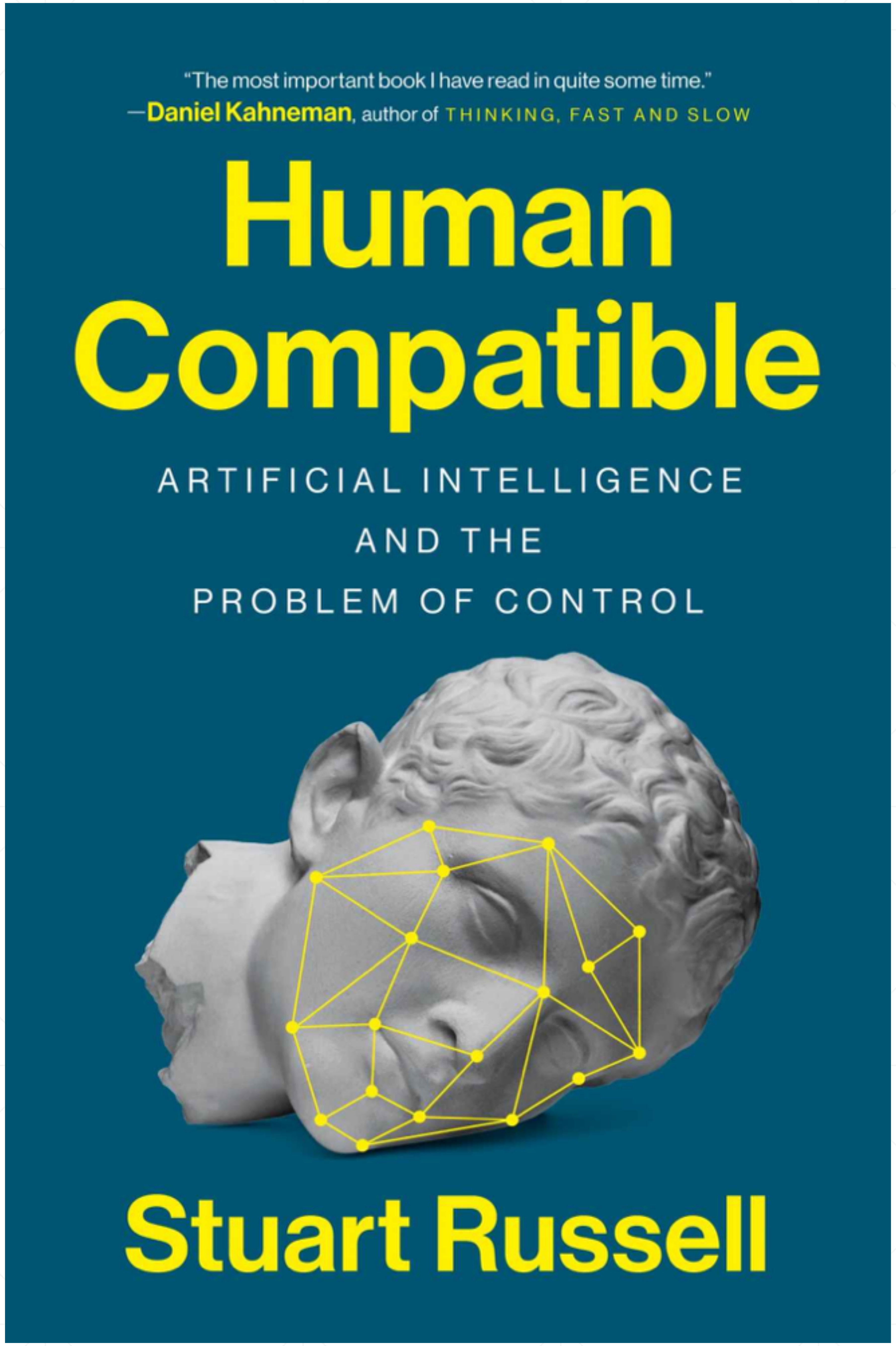
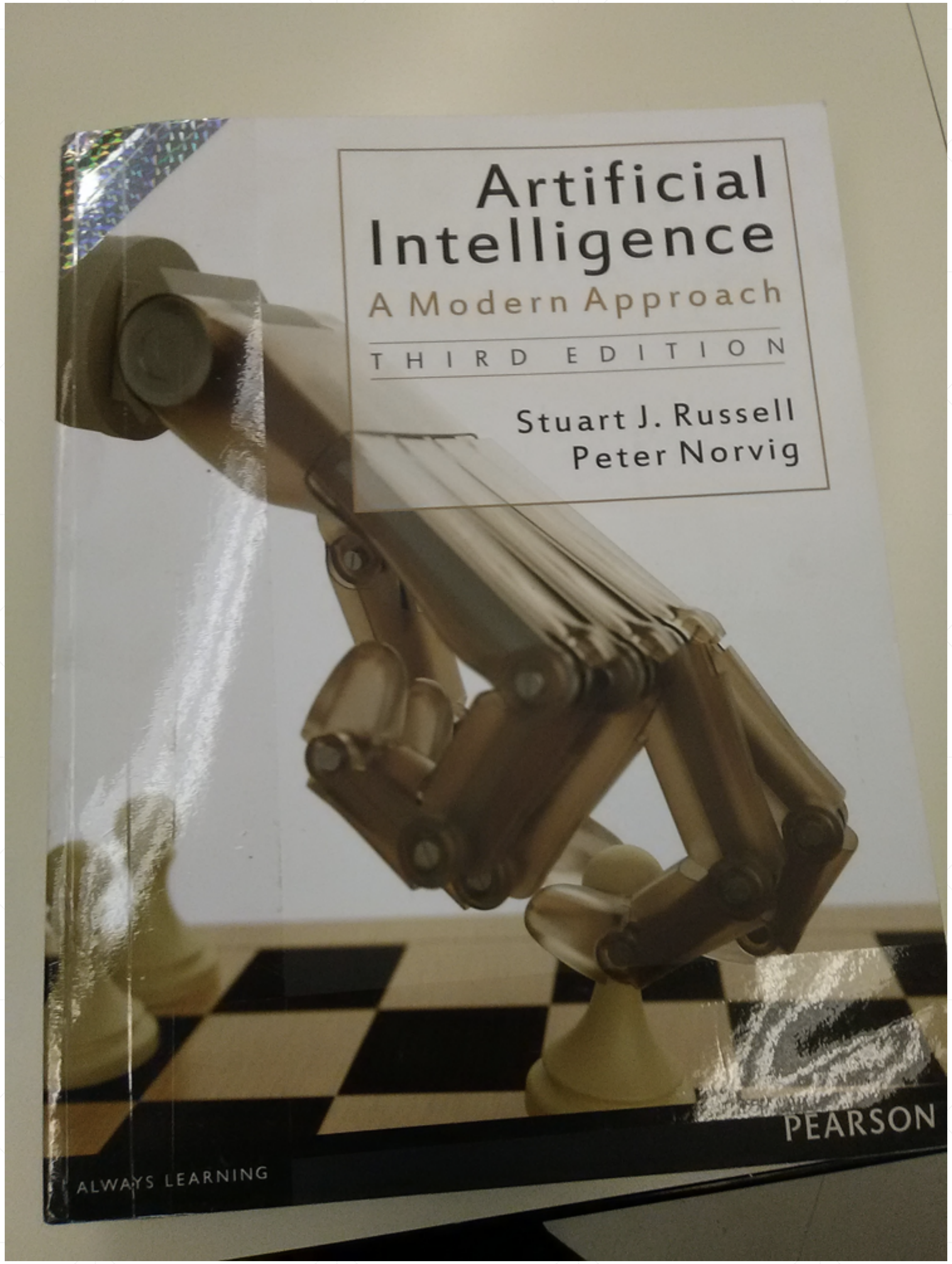
	Unassisted	**Assisted**	F(1,11)	*p*-value
Q1: On average, it was easy to guide the robots to their goals.	2.92	4.50	17.49	$<$ **0.01**
Q2: I was successful at guiding the robots.	2.25	3.92	13.75	$<$ **0.01**
Q1, after objective measures revealed	2.67	4.17	17.47	$<$ **0.01**
Q2, after objective measures revealed	2.92	3.25	0.88	0.37

Fig. 5. Survey responses on a 7-point Likert scale, where 1 = Strongly Disagree, 4 = Neither Disagree nor Agree, and 7 = Strongly Agree.

Broader Impacts

- Integrate results into AI Textbook
- Develop children's book on human–robot learning
- Integrated into curricular materials




"The most important book I have read in quite some time."
—Daniel Kahneman, author of THINKING, FAST AND SLOW
Human Compatible
ARTIFICIAL INTELLIGENCE
AND THE
PROBLEM OF CONTROL
Stuart Russell

Artificial Intelligence
A Modern Approach
THIRD EDITION
Stuart J. Russell
Peter Norvig
PEARSON
ALWAYS LEARNING

THE LAWRENCE
HALL OF SCIENCE
UNIVERSITY OF CALIFORNIA, BERKELEY

The Lawrence Hall of Science

- **UC Berkeley's Public Science Center**

We provide a window into Berkeley's cutting-edge research, exciting discoveries and transformative innovation

- **A Learning Lab**

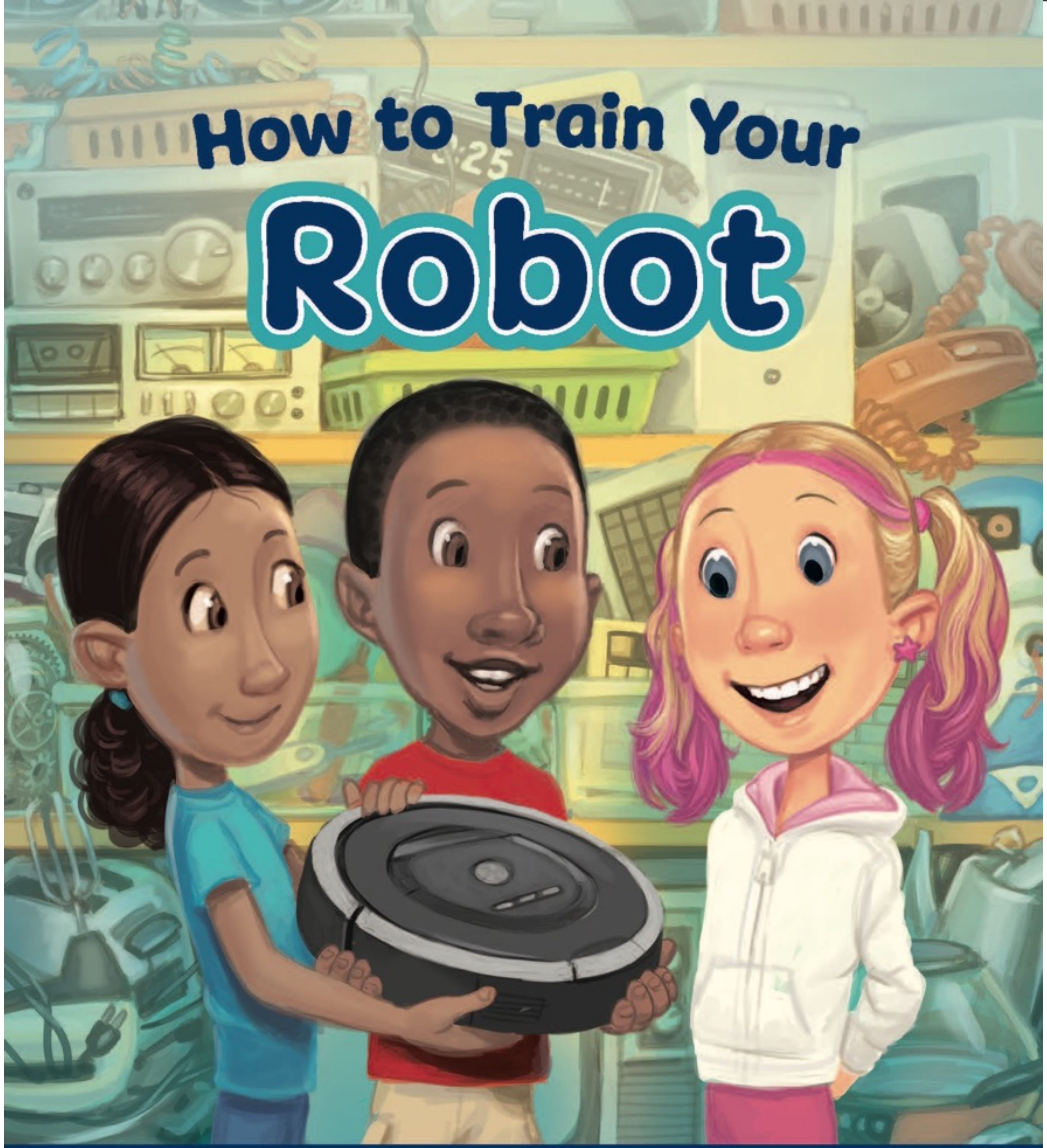
We design, develop, test, and study model STEM learning programs and products for youth, families, and adults from diverse backgrounds in our specialized facilities

- **Global Impact**

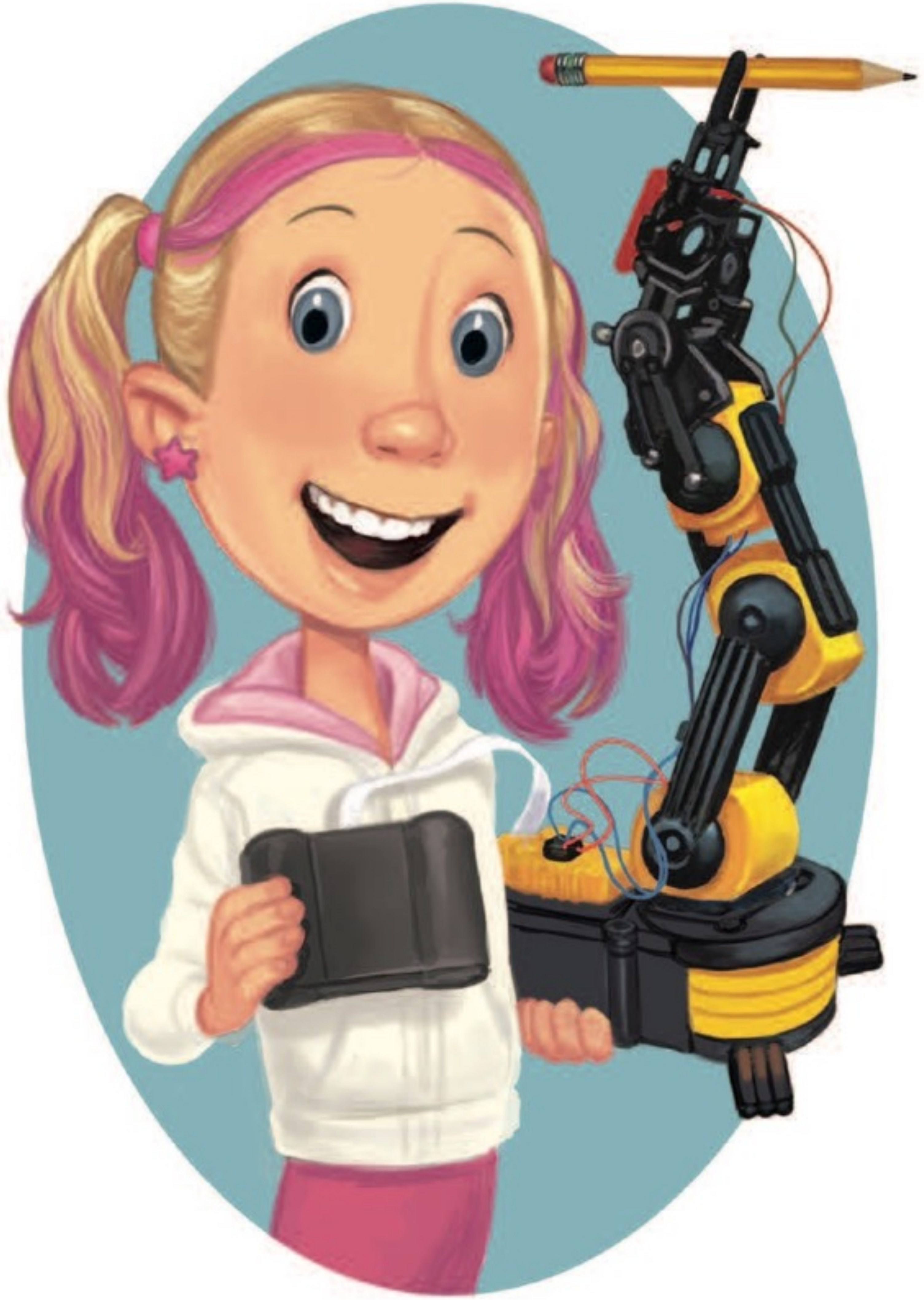
We disseminate & scale-up our effective and innovative learning programs and materials in ways that broaden participation in STEM and have local, national, and global impact.

to inspire and foster science, technology, engineering, and mathematics learning for all, especially those with limited access to science

How to Train Your Robot



by Blooma Goldberg, Ken Goldberg, and Ashley Chase
illustrated by Dave Clegg



I'm Blooma. I love art, science, math, dancing, basketball, and roller skating, but my favorite activity is inventing new things… especially robots! And I'm super excited to tell you how I figured out the best way to train a robot.

I'm the president of the Razzle-Dazzle Robot Club. We meet after school in a room we call the Lightning Lab, where we **design** and build robots. Our robots do hip-hop dance moves, escape from mazes, launch water balloons, and other interesting stuff. Once we hid a robot under a cardboard box and took it out on the schoolyard during recess. We freaked out some fifth graders by making the box move just slightly when they walked by it. People said it was the best prank ever.

One day, we were in the middle of building bug robots when our club advisor, Mariana, spoiled the fun. "Five o'clock: cleanup time," she announced. Cleaning up was the only bad part of robot club. We all wanted to keep working on our robots. Suddenly, I had an idea: maybe we could do both at once. "Let's make a robot that can clean for us. We can call it Clark—the Cleaning Robot!" The entire club wanted to help and everybody had ideas for building our new robot.

The next afternoon, we started work on Clark. Step one was to visit my favorite spot: we call it the Twisted Treasure Closet. It's filled with old machines like computers, phones, TVs, toasters, radios, fans, cameras, popcorn poppers, electric toothbrushes, and lots of motorized toys. Most of the machines don't work anymore, so we just take them apart and use the motors, gears, and other parts for building new inventions.

My friends Tyler and Rada picked up something that looked like a giant hockey puck with wheels. "It's an old robot vacuum cleaner," said Tyler. "It drives around and sucks up dirt."

"I saw one in a video," said Rada: "Somebody's cat was riding it." We turned it on, and it still worked! We decided that old robot vacuum could become part of Clark. It would be the perfect way for Clark to move around.

We needed a way to control Clark, and Rada had an idea. She remembered that robot vacuums have computers and Wi-Fi connections. We went online and found a way to get the club's laptop computer to send commands to the little computer inside Clark. After a few days of work, we got the robot vacuum to follow directions from the laptop computer. We could use the laptop to steer Clark around the workshop. The robot did a good job sucking up dust, but couldn't pick up things like tools and trash. Clark needed a way to clear bigger stuff off the floor.

"Okay, brainiacs: time for a club meeting!" I called. "Anybody have ideas for parts to help Clark pick things up?" People came up with all kinds of stuff: shovels, claws, scoops, suction cups, magnets, and even a laser zapper to blast trash into dust! I thought the laser zapper sounded really fun, but Tyler said Clark needed to put things away, not destroy them. Annoying, but I had to admit he was right. After comparing all the design ideas, we decided to try out the scoop idea first.

In the Twisted Treasure Closet, we found a dustpan that would work as a scoop. We also found an old toy truck you could drive with a **remote control**. We pulled out the motors so we could use them to power Clark's arm.

We built a robot arm with two motors. We bolted the scoop to the arm and then connected the whole thing to Clark. We used the remote control to make the scoop go up and down. The arm wasn't easy to control, but after a while Tyler got the hang of it and taught us. He's an ace at video games!

In our first experiment, we scattered some foam packing peanuts over the floor. We started up Clark and lowered the scoop. The scoop slid under the packing peanuts and picked them up easily. We put a box on one side of the room for Clark to dump the peanuts into. When Clark got near the box, I used the remote to slowly raise the scoop. The problem was that the peanuts didn't fall out into the box, they just stayed in the scoop. I lowered the scoop and raised it up again, but it still didn't work.

Then I tried raising the scoop really fast to get the peanuts to pop out of the scoop. They popped out all right. The packing peanuts went flying into the air!

It was just my luck that this was right when Mariana came over to see how we were doing with the testing. She got showered with packing peanuts! Awkward. It was pretty funny, though. Mariana didn't get too mad.

Mariana reminded us that workshop cleanup doesn't mean dumping everything into a box: each object needs to be put away in the right place. Tools go in the tool storage area, trash goes into the trash can, extra parts go into the closet. "Don't worry," I said. "We'll work on a new design! And sorry about the packing peanuts. You've still got one more stuck in your hair, by the way."

So… the scoop wasn't working out. Clark needed to be able to pick things up one by one and put each one in the right place. We started brainstorming different kinds of robot grippers. We drew lots of **diagrams** to design a motorized gripper, and then worked together to build it.

To test the new gripper, we scattered some tools on the floor. Then we took turns driving Clark over to a tool and using the remote controls to make the gripper arm pick up the tool and put it away. Some tools were really easy to pick up, but Clark had a LOT of trouble picking up other tools.

We must have tried about 20 times to get Clark to pick up the hammer. It kept dropping before Clark could put it away. Then Tyler guided Clark to grab the head of the hammer in just the perfect spot so it could pick it up and put it on the tool bench. Finally!

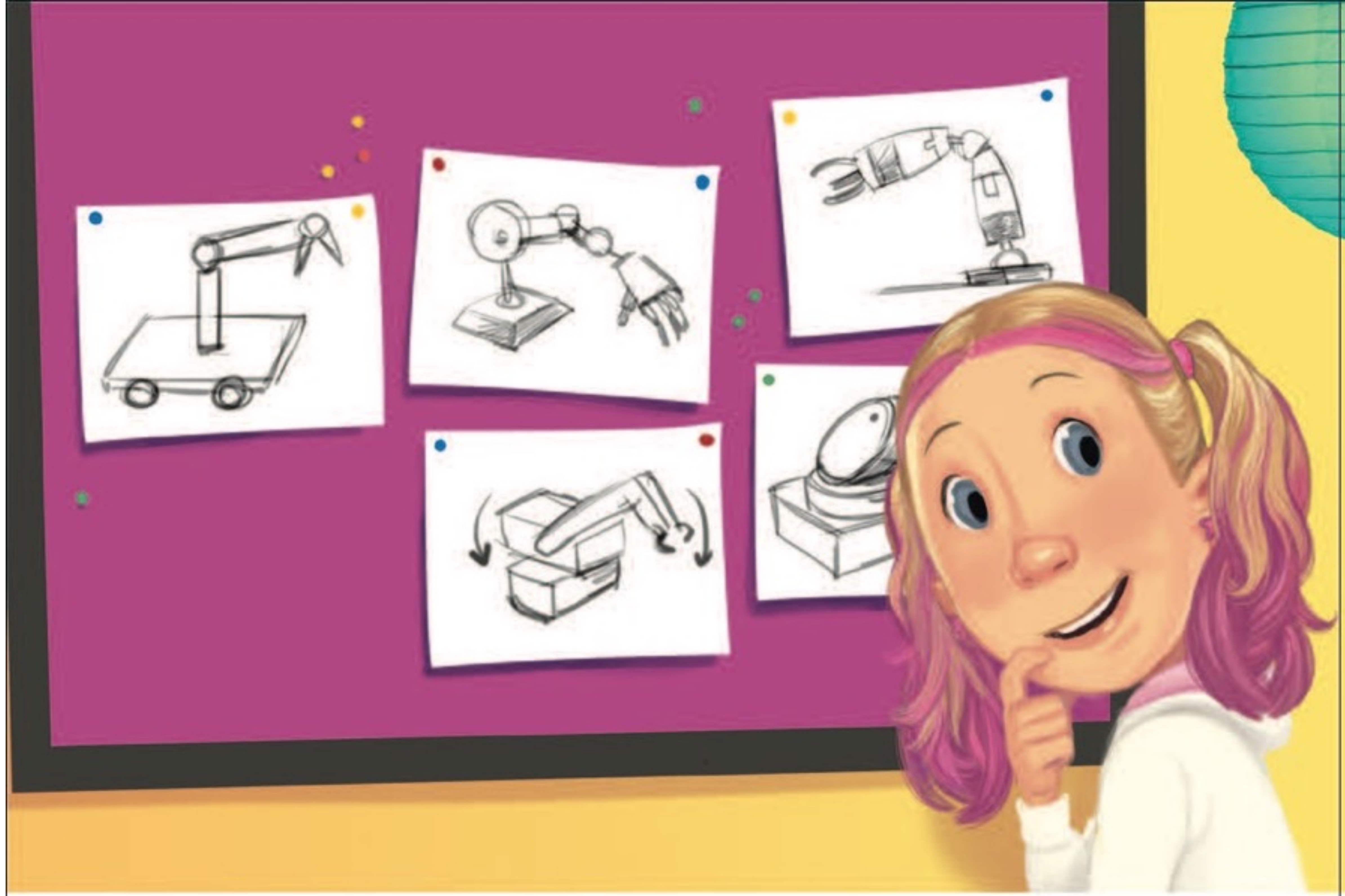
Controlling Clark's gripper was fun… for a while. After everybody had taken a turn at the controls, we still had tons of tools left to pick up. It was taking way too long, and we were working even harder than we did before when we cleaned up the Lightning Lab ourselves. Clark needed to step up and do more of the work. (Okay, Clark had wheels instead of legs, so Clark couldn't step up exactly, but you know what I mean.)

We did another brainstorming session thinking about what Clark would need to clean up without us working the controls. We realized the computer could operate Clark's motors on its own. But before Clark could pick up an object and put it away, the robot would need to find the object in the first place. Clark needed a way to see.

Rada was rummaging around the Twisted Treasure Closet when she yelled "Eureka!" She had found an old webcam that could send video to a computer. We attached the webcam to Clark and saw a video pop up on the club computer screen. The video showed the workshop floor. Rada bent down and wiggled her fingers in front of Clark, and we saw the wiggling fingers on our computer screen. The screen showed the world from Clark's point of view. Clark could see!

Now it was time to do some **coding**. In the Razzle-Dazzle Robot Club, we've all gotten pretty good at coding, so we wrote some **code** giving Clark instructions on how to clean up. This is what we instructed Clark to do: Look at the floor with the webcam and find the objects on the floor. Drive over to an object and pick it up. Go to the tool storage area and put the object down in the correct spot. Find another object and do the same thing. Repeat until there's nothing on the floor.

Sounds easy, right? Well, it wasn't easy for Clark. Clark could find objects just fine, and would try to pick them up... but Clark kept dropping things. Clark was still a klutz!



We thought maybe the problem was Clark's gripper, so we experimented with different designs. We tested a suction cup gripper and grippers with two jaws, three jaws, and four jaws. We even tried a gripper shaped like a human hand. Each design picked up some stuff, but dropped other stuff. Some grippers were better at picking up round things, and others were better at picking up squishy things. No gripper was perfect at picking up everything.

No matter how much training we did, Clark didn't seem to be getting better at cleaning up. For example, you know that hammer that had to be picked up just the right way or it would drop? Clark usually picked it up eventually, when the gripper finally happened to close the right way.

Still, the next day, Clark would go right back to trying to pick up the hammer the wrong way again. "Silly robot! When will you ever learn?" said Tyler.

"AHA!" That gave me an idea. Maybe the problem wasn't the gripper, but the coding in Clark's robot brain. Clark needed to be able to learn how to pick things up. Tyler went online and searched ROBOT LEARNING. A bunch of stuff popped up on the screen, but one thing on the list caught my eye—a robot lab at the university in our town. In their lab, Professor Mason and his students were investigating how to get robots to pick things up, just like we were! I could hardly believe it. We told Mariana about it, and she set up a field trip for us to go and visit Professor Mason's lab.

Walking into Professor Mason's lab was amazing. It was the Lightning Lab in overdrive! I counted eight different robots around the room. Everywhere, robot arms were reaching for things. While he was showing us around, Professor Mason told us about robot learning.

"Robots usually do only what people tell them to do," Professor Mason said. "They just follow instructions. But in the last few years researchers are starting to build robots that can learn from experience, like Dexter over here. We call it **deep learning**."

Professor Mason brought us over to a robot with a pair of robot arms. "Want to try it out?" he asked me, pointing to my special plastic keychain shaped like a shark. I dropped it in front of Dexter's robot arms, and one of the robot arms reached right down and picked it up. Wow!

"Dexter has never picked up a shark-shaped object like this one before," said Professor Mason, "and we didn't code Dexter for this object. Instead, we wrote code that gave Dexter the ability to learn. Dexter can pick up new things because it has practiced the best ways to pick up millions of objects with different shapes."

"It must have taken forever to get Dexter to do all that practicing," I said, thinking of Clark.

Professor Mason replied, "Great point! Actually, the practicing went really fast because all of it was **simulated** on a computer."

I was starting to understand. "It's like Dexter dreamed about picking up lots of different things, and that's how it learned!"

"Exactly!" said Professor Mason. "That's how this robot learned to pick up objects of almost any shape."

"Even a hammer?" asked Tyler. We explained all about Clark the Cleaning Robot, and how much trouble it was having picking up our hammer. Professor Mason and his students dug out some tools in different shapes, and we did some experiments with them. Dexter picked up some tools but dropped others. Professor Mason and his students got really interested in this challenge and said they would help us.

For days, the Razzle-Dazzle Robot Club was trying to guess how Professor Gold and his students were going to help us. "Maybe they'll send Dexter over to teach Clark a few tricks," Rada joked. A week later, we got an email from Professor Mason with a file attached. In the email, he explained that the file had code in it to help Clark start learning from practice. We couldn't wait to try it!

We installed the new code on Clark. Then we put some tools on the floor and started Clark up. Clark moved toward a hammer and started reaching for it, but Clark was reaching for the middle of the hammer again. I carefully guided Clark's gripper up high near the head of the hammer, so that it wouldn't drop. We took turns showing Clark the best way to pick up the tools. Then Clark started experimenting on its own!

"Five o'clock, clean up time!" said Mariana. There were still some tools on the floor. We begged Mariana to let us keep working with Clark. We watched as Clark taught itself how to pick the tools up. When Clark finally picked up the hammer without dropping it, we cheered and gave each other high-fives! Soon the floor was clean.

From now on, Clark would help us keep the Lightning Lab clean. We were learning about Clark, and Clark was learning about us. So what's the best way to train a robot? Teach it to train itself!

Glossary

code: instructions that tell a computer how to do something

coding: writing instructions for a computer

deep learning: a new approach to artificial intelligence in which millions of numbers in a network are tuned to match the examples that are given

design: (noun) something new made to solve a problem
(verb) to try to make something new that solves a problem

diagram: an illustration that shows how something works or what its parts are

remote control: a controller that can send signals to a machine and tell it what to do

simulated: recreated inside a computer

Can robots learn?

Blooma and her friends in the Razzle-Dazzle Robot Club hope so. They build a robot and try to train it to clean up their workshop, but that turns out to be harder than it sounds. Will Clark the Cleaning Robot ever learn to clean up?

Blooma Goldberg is 9 years old. **Ken Goldberg** is the Principal Investigator of the Scalable Collaborative Human-Robot Learning (SCHooL) project supported by the National Science Foundation and is the William S. Floyd Jr. Distinguished Chair in Engineering, UC Berkeley. Prof. Goldberg continues to learn from his students (and daughters). **Ashley Chase** is a science writer and editor in the Learning Design Group at the Lawrence Hall of Science. She has authored dozens of educational books for children of all ages.

THE LAWRENCE HALL OF SCIENCE

ISBN 978-1-94715-919-8
50799
9 781947 159198

Publications

Over 30 papers, see: schoolproject.berkeley.edu

- G. Swamy, S. Reddy, S. Levine, A. Dragan, “Scaled Autonomy: Enabling Human Operators to Control Robot Fleets”, ICRA 2020
- A. Tanwani, P. Sermanet, A. Yan, R. Anand, M. Phielipp, K. Goldberg, “Motion2Vec representation learning from surgical videos”, ICRA 2020
- Y. Du, S. Tiomkin, E. Kiciman, A. Dragan, P. Abbeel, “Goal Agnostic Assistance through Human Empowerment”, ICML, 2020 (submitted)
- Y. Wu, W. Wu, A. Tamar, S. Russell, G. Gkioxari, Y. Tian, “Bayesian Relational Memory for Semantic Visual Navigation”, ICCV, 2019
- I. Huang, S. Huang, R. Pandya, A. Dragan, “Nonverbal Feedback for Human Teachers”, CoRL, 2019
- S. Russell. Human Compatible: AI and the Problem of Control. Penguin Books Ltd., 2019
- B. Goldberg, K. Goldberg, A. Chase. How to Train Your Robot. Lawrence Hall of Science, 2019.
- R. Fox, R. Berenstein, I. Stoica, K. Goldberg, “Multi-Task Hierarchical Imitation Learning for Home Automation”, CASE 2019

NSF NRI 2.0 — Research Themes

- **Collaboration**
  - **Collaborate and coordinate effectively with multiple people and robots;**
  - Learn efficiently from direct experience, people, other robots, and digital media;
  - Inform and instruct multiple people and robots.
- **Interaction**
  - Reliably recognize and predict the activities of others;
- **Scalability**
  - Perform a variety of tasks in a variety of situations;

Scalable **C**ollaborative **H**uman–R**o**b**o**t **L**earning
(SCHooL)

Sept 2017 - Aug 2021

Scalable **C**ollaborative **H**uman–R**o**b**o**t **L**earning
(SCHooL)

Sept 2017 - Aug 2021

Ken Goldberg

Pieter Abbeel

Anca Dragan

Stuart Russell

Berkeley
UNIVERSITY OF CALIFORNIA

BAIR
BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

AUTOLAB

schoolproject.berkeley.edu

SCHooL: Scalable Collaborative Human–Robot Learning

Ken Golderg, Pieter Abbeel, Anca Dragan, Stuart Russell

contact: <goldberg@berkeley.edu>

University of California, Berkeley

schoolproject.berkeley.edu

1. Overview

- Learning from Demonstrations (LfD) paradigms lack a theoretical framework for scalable human–robot cooperative learning and hierarchical planning.
- The SCHooL project aims to fill this gap by investigating scalable robot manipulation, where multiple robots collaboratively learn from multiple humans with a unified game-theoretic inverse reinforcement learning framework.
- Integrative Application: "Surface Decluttering": robots that keep specified surfaces clear by identifying, grasping, and appropriately relocating objects with applications in homes, schools, warehouses, offices, manufacturing and machine shops, retail stores using an emerging class of mobile manipulators such as the Fetch robot.

2. Research Objectives

1. Formalize framework for scalable collaborative inverse reinforcement learning (SCIRL) using theory of multiagent games and collaborative learning in multiple distributed domains.
2. Develop deep learning representations of visuospatial features and reward functions to extract and share deep learning representations for scalable human-robot learning.
3. Develop hierarchical task and reward structure to increase planning horizon and decrease sample complexity by partitioning complex tasks into sub tasks.
4. Develop new models to represent and share awareness of robot capabilities and robot models of human intent to support distributed learning.

These objectives support the NRI 2.0 themes: Collaboration, Interaction, and Scalability and for Broader Outreach we are partnering with Lawrence Hall of Science and Penguin Books.

3. Problem Formulation

- Cooperative inverse reinforcement learning (CIRL) is posed as an n-player cooperative Markov game with asymmetric information: only humans observe the reward parameters:

$$G = \langle \mathcal{S}, \{\mathcal{A}^{\mathrm{H}}, \mathcal{A}^{\mathrm{R}}\}, P(\cdot|\cdot,\cdot,\cdot), P_0(\cdot), R_\theta(\cdot,\cdot,\cdot), \gamma \rangle$$

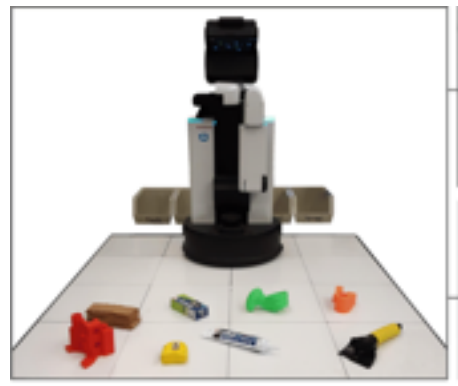
- Robots try to learn the reward parameters by querying the humans and by observations of other robot actions.
- Robots learn parameters incrementally from human actions and humans incrementally learn to convey intentions to the robots.

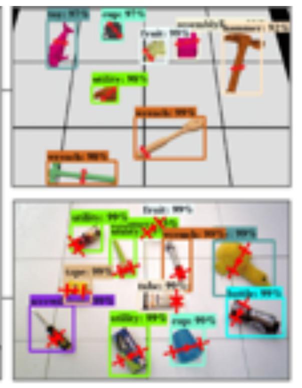
4. Primary Results

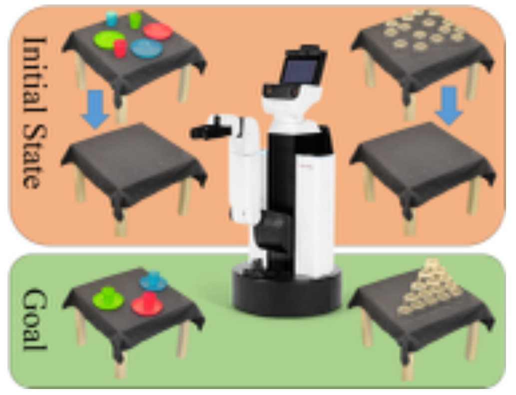
- A reformulation of AI replacing the standard model (optimizing a fixed, known objective) with optimizing human objectives that are not fully observed.
- New Formal models exploring irrationality in reward inference; learning to control a fleet of robots by humans; learning efficient representation for intrinsic motivation.
- Combining depth sensing and sim-to-real transfer for extracting hierarchical task and reward structure.

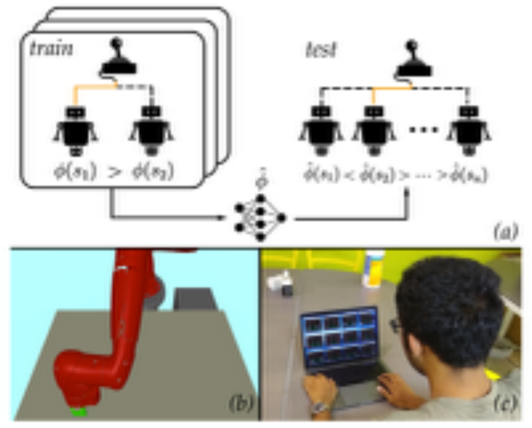
Broader Impacts:

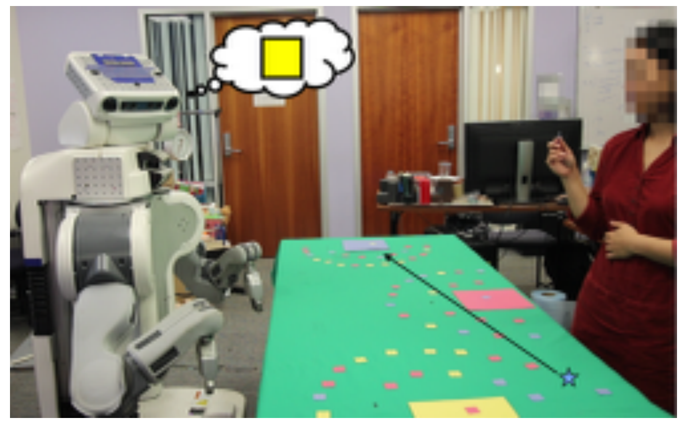
- Project incorporated into 4th Edition of Russel and Norvig: **AI: A Modern Approach**. 2020 textbook.
- **How To Train Your Robot.** Elementary school book based on the project to inspire young and under-represented minority readers to explore AI and robot learning. Freely distributing 1800 copies to schools and student clubs, featured in IEEE Robot Gift Guide.

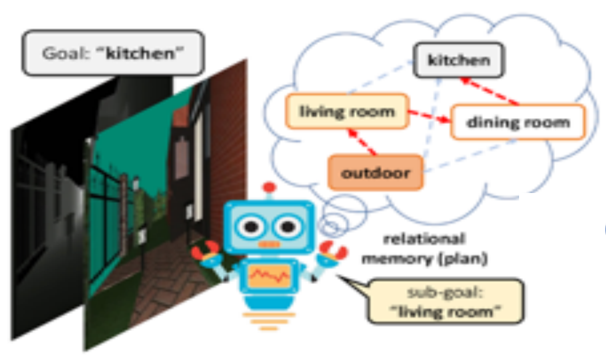


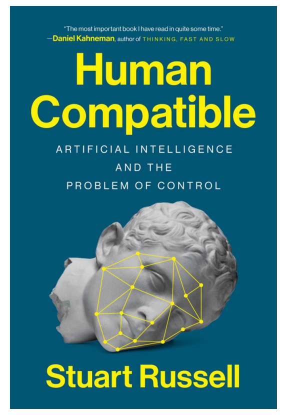












5. Selected Publications

See website above for comprehensive list of over 30 papers.

- G. Swamy, S. Reddy, S. Levine, A. Dragan, "Scaled Autonomy: Enabling Human Operators to Control Robot Fleets", ICRA 2020
- A. Tanwani, P. Sermanet, A. Yan, R. Anand, M. Phielipp, K. Goldberg, "Motion2Vec representation learning from surgical videos", ICRA 2020
- Y. Du, S. Tiomkin, E. Kiciman, A. Dragan, P. Abbeel, "Goal Agnostic Assistance through Human Empowerment", ICML, 2020 (submitted)
- Y. Wu, W. Wu, A. Tamar, S. Russell, G. Gkioxari, Y. Tian, "Bayesian Relational Memory for Semantic Visual Navigation", ICCV, 2019
- I. Huang, S. Huang, R. Pandya, A. Dragan, "Nonverbal Feedback for Human Teachers", CoRL, 2019
- S. Russell. Human Compatible: AI and the Problem of Control. Penguin Books Ltd., 2019
- B. Goldberg, K. Goldberg, A. Chase. How to Train Your Robot. Lawrence Hall of Science, 2019.
- R. Fox, R. Berenstein, I. Stoica, K. Goldberg, "Multi-Task Hierarchical Imitation Learning for Home Automation", CASE 2019