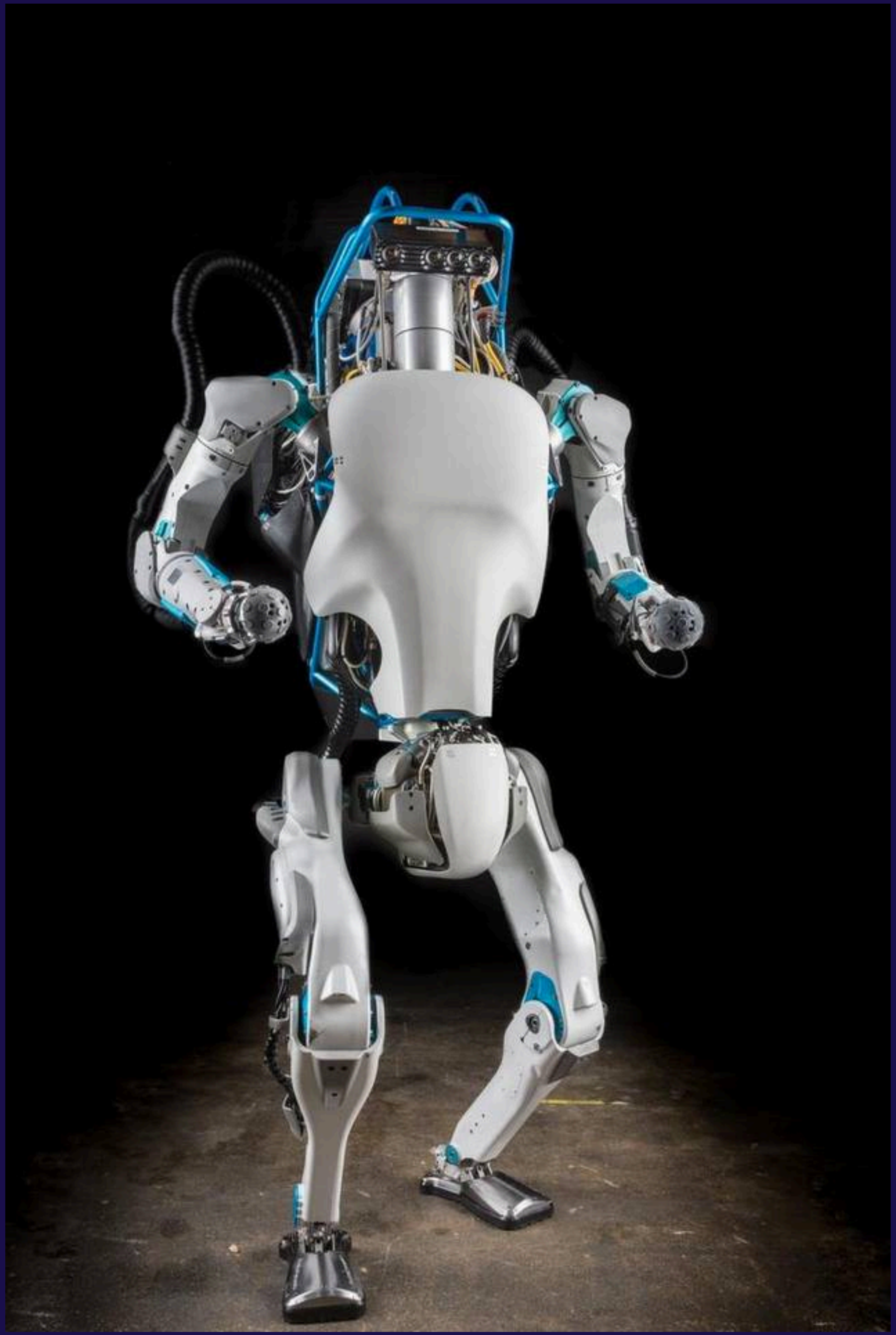
OpenAI

Learning Dexterity

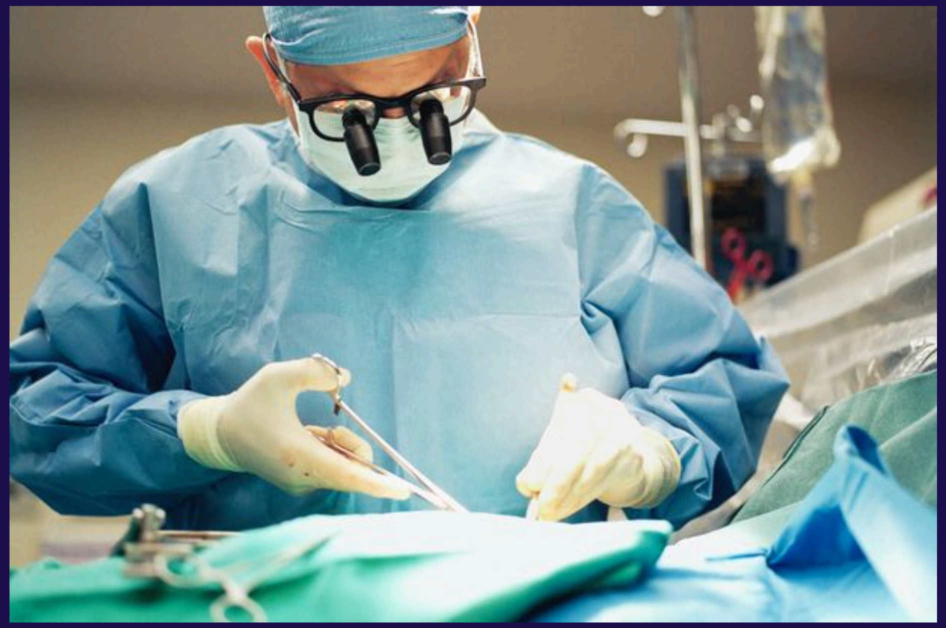
OpenAI Robotics

Alex Paino, Arthur Petron, Ilge Akkaya, Jerry Tworek, Jonas Schneider,
Josh Tobin, Lilian Weng, Maciek Chociej, Mateusz Litwin, Matthias Plappert,
Nikolas Tezak, Peter Welinder, Qiming Yuan, Wojciech Zaremba

AUGUST, 2019

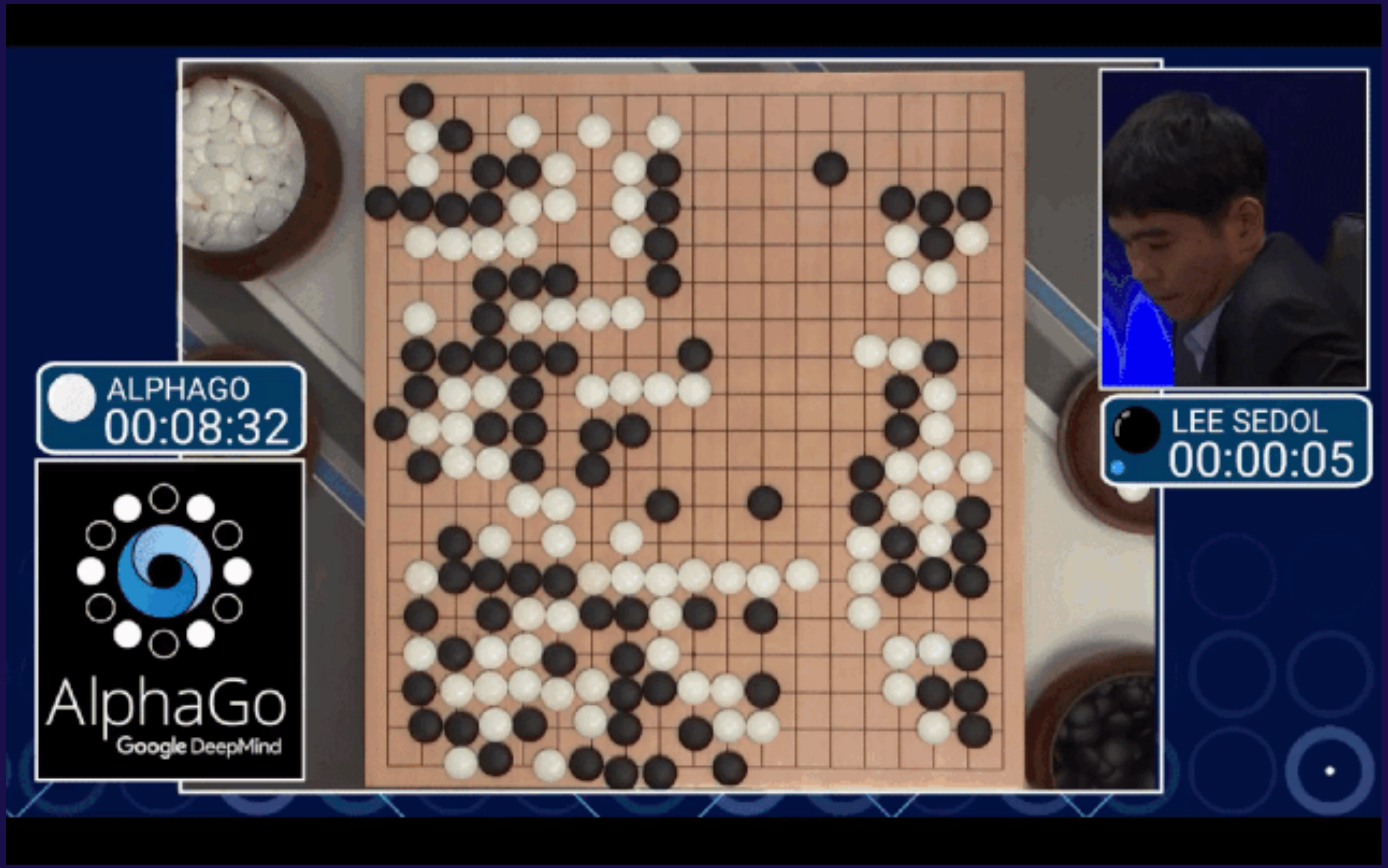



FRAGILE
FRAGILE




WAYMO

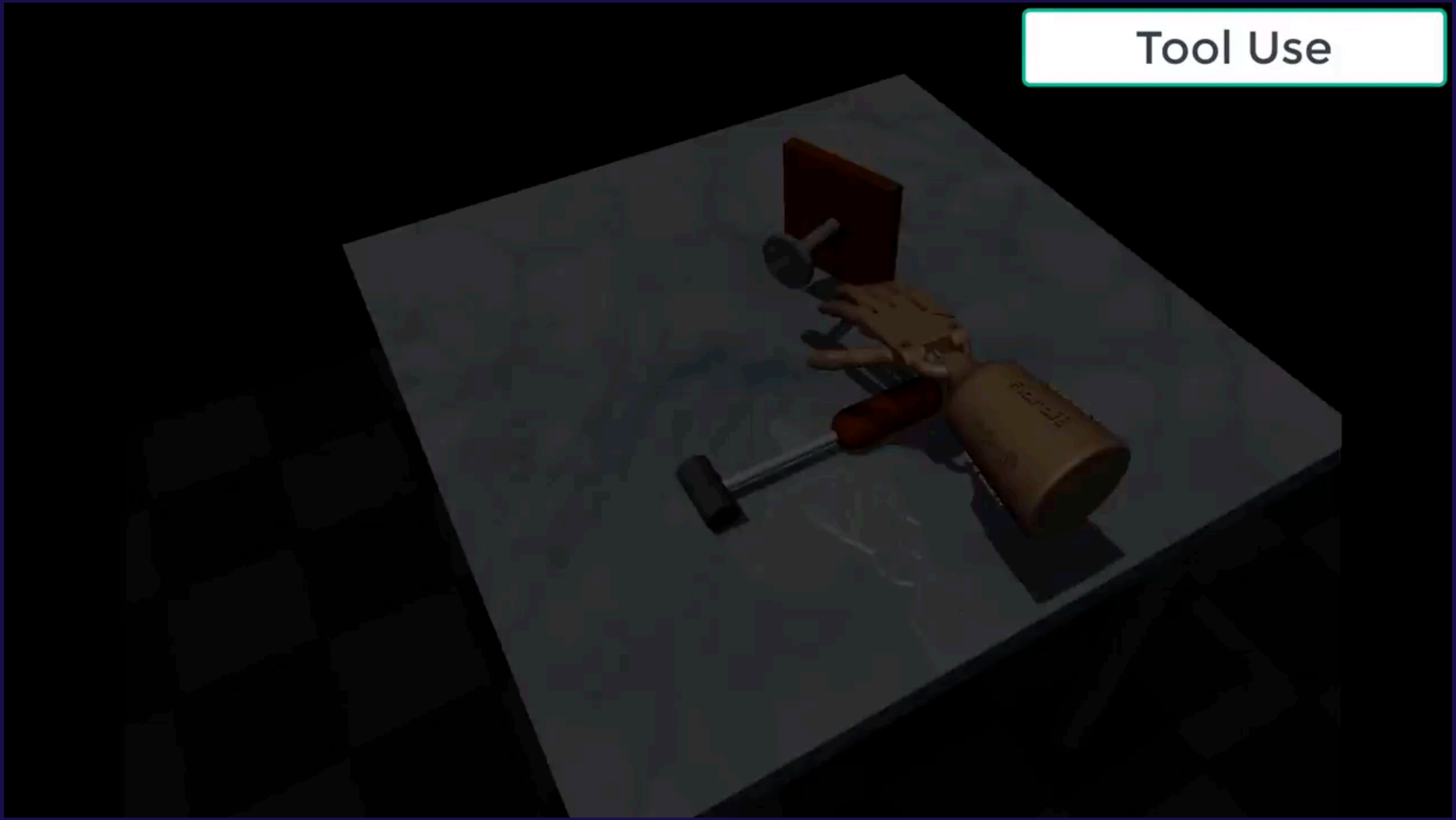
Reinforcement Learning (RL)



GO (ALPHAGO ZERO)

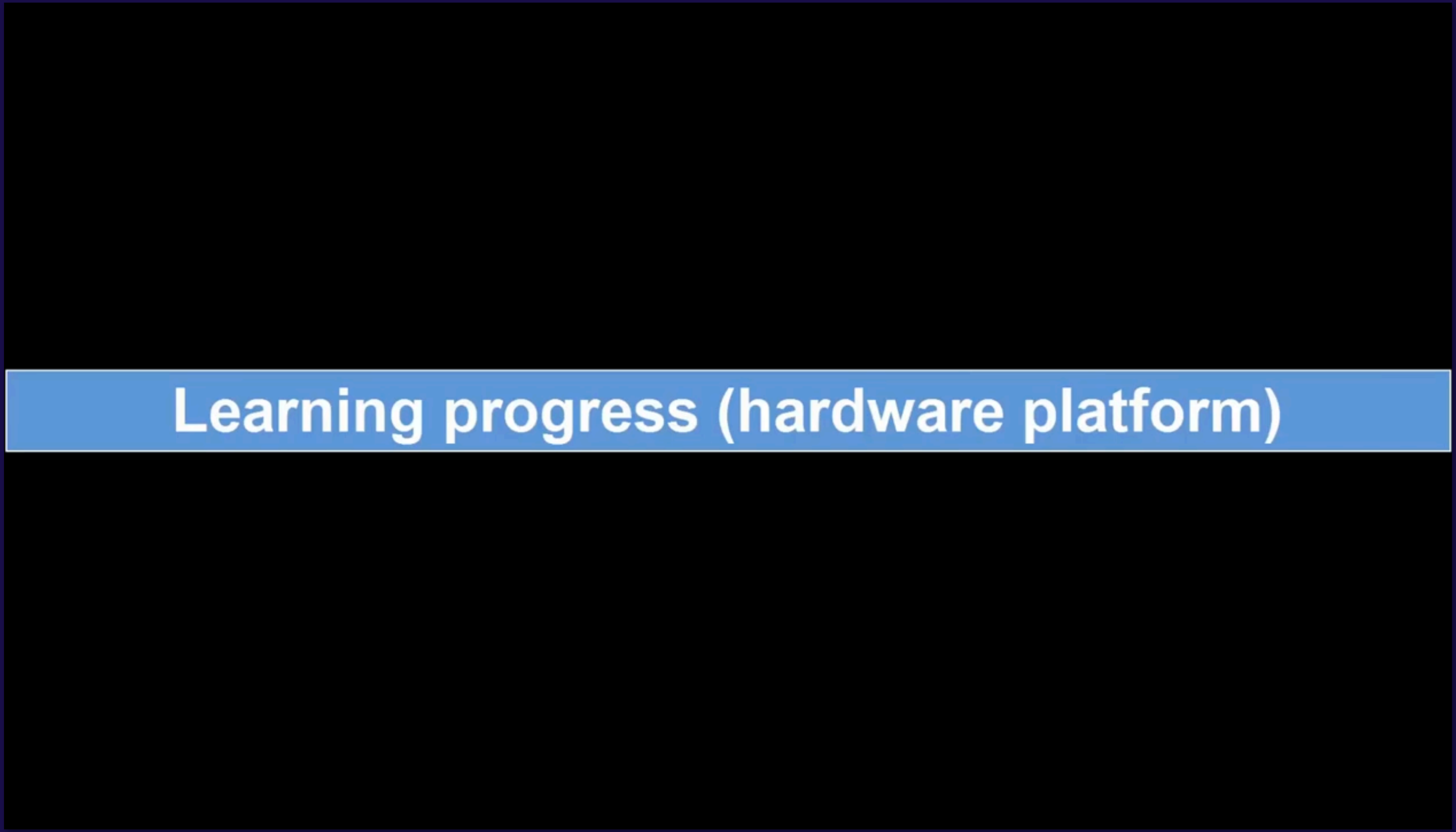
DOTA 2 (OPENAI FIVE)

RL for Robotics (1)



Rajeswaran et al. (2017)

RL for Robotics (2)



Kumar et al. (2016)

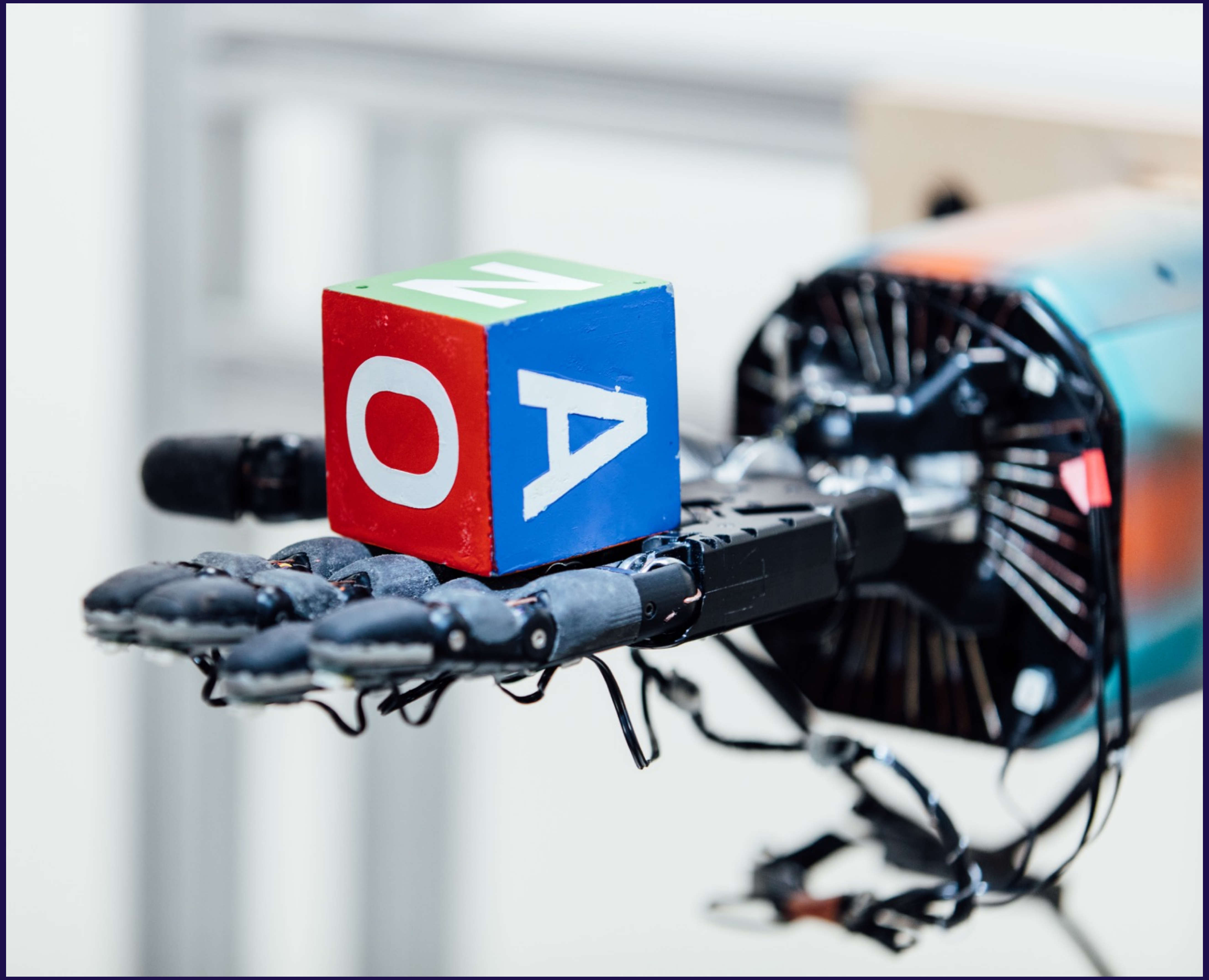
RL for Robotics (3)

Levine et al. (2018)

Can we train complex policies
only in simulation
but still run on the real robot?

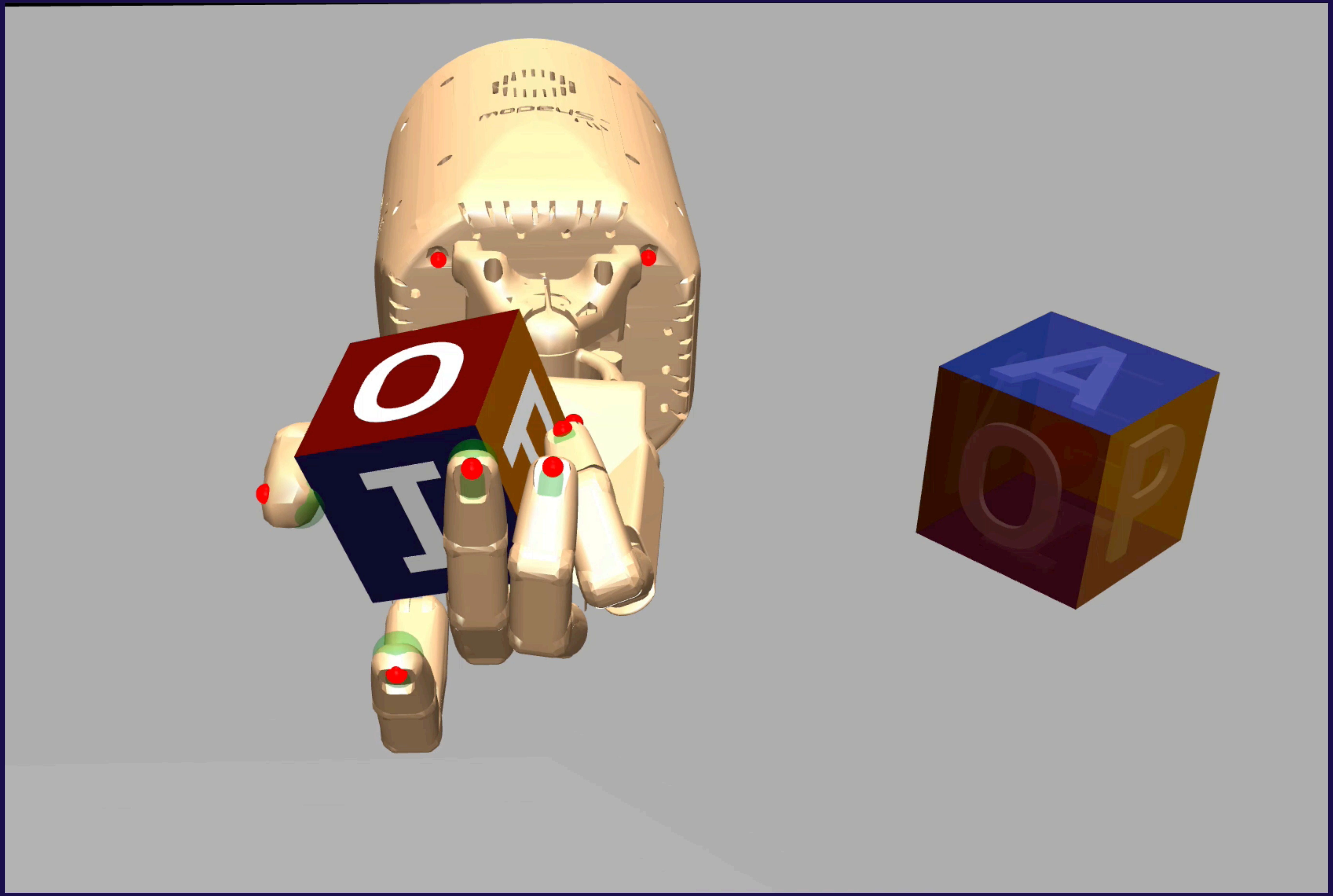
Learning Dexterity

Dexterous In-Hand Manipulation



- A humanoid 5-fingered hand
- A human hand is a universal end-effector
- Long standing unachieved goal for classical robotics

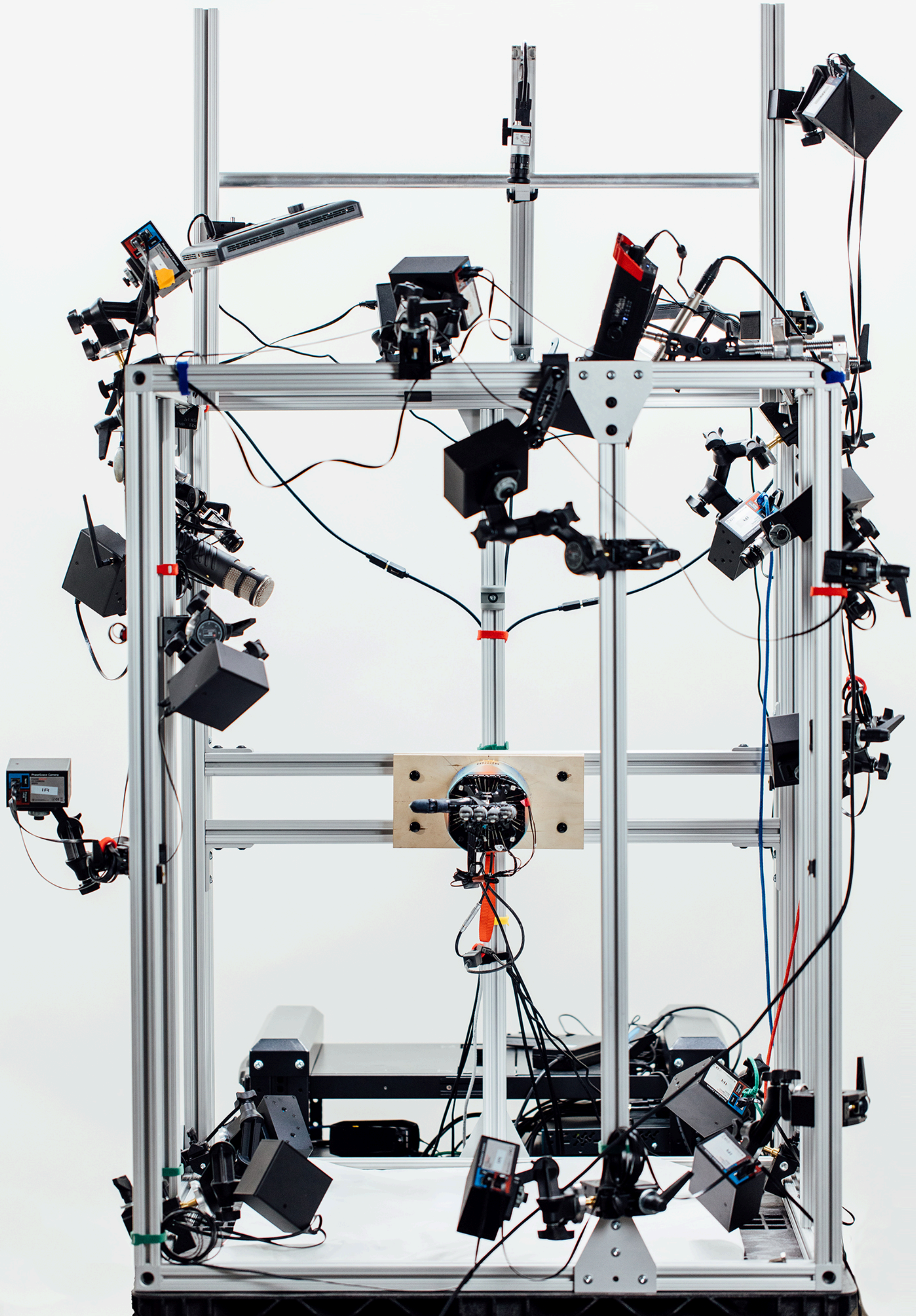
Simulation



Task: reorient the object in-hand

Sim2Real

Transfer



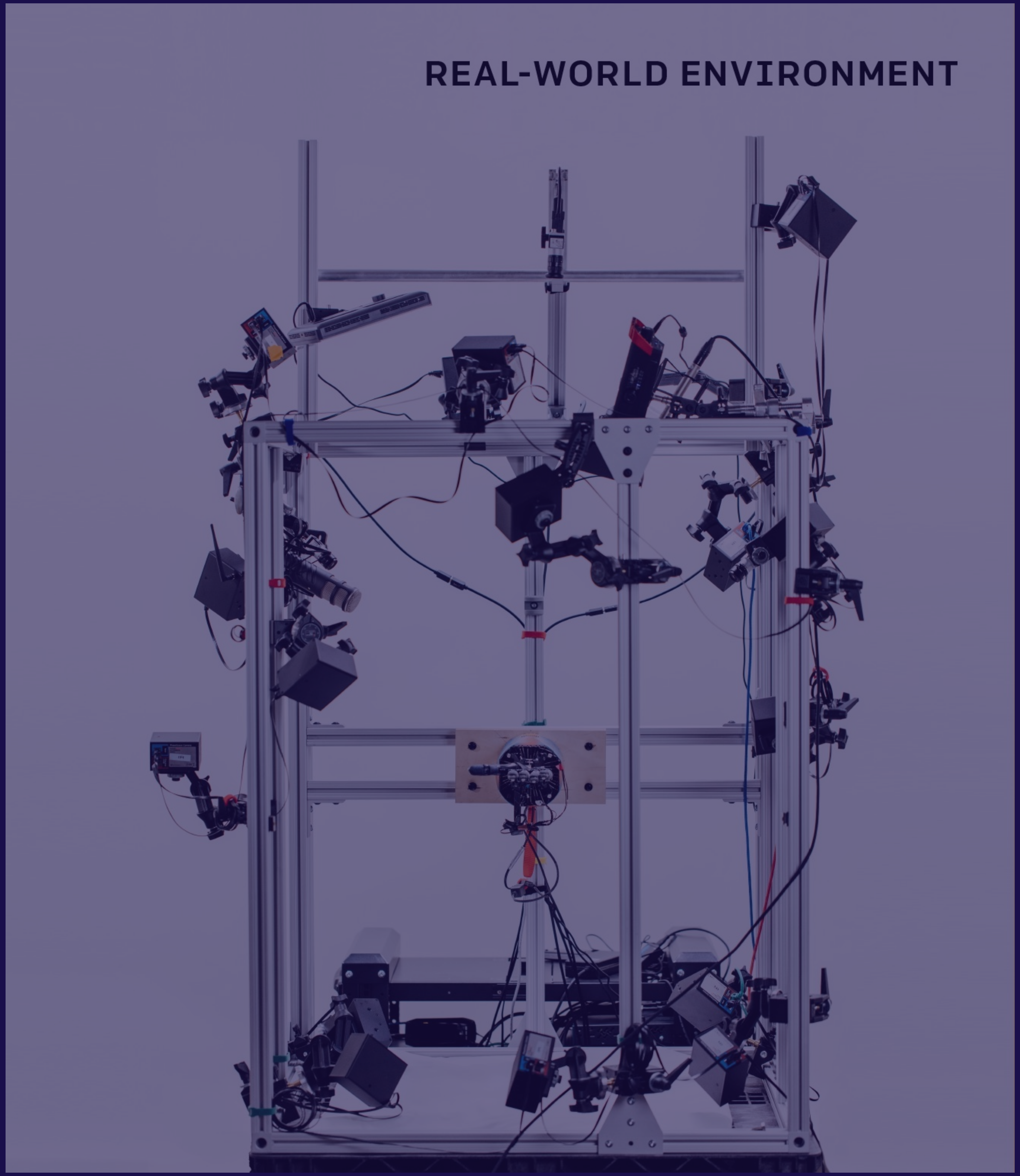

Shadow Dexterous Hand


PhaseSpace tracking


Top RGB camera
Right RGB camera
Left RGB camera

Sim2Real

Transfer



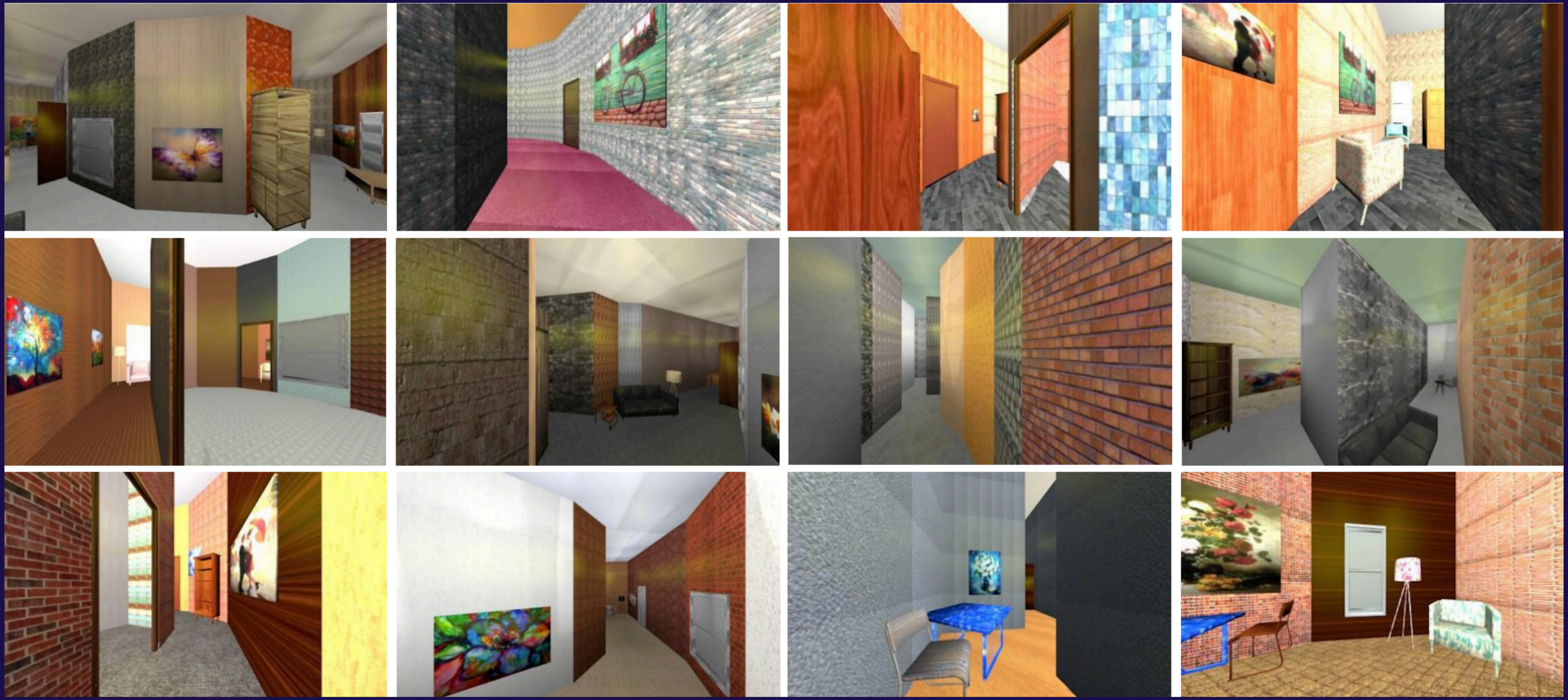
Reinforcement Learning

+

Domain Randomization

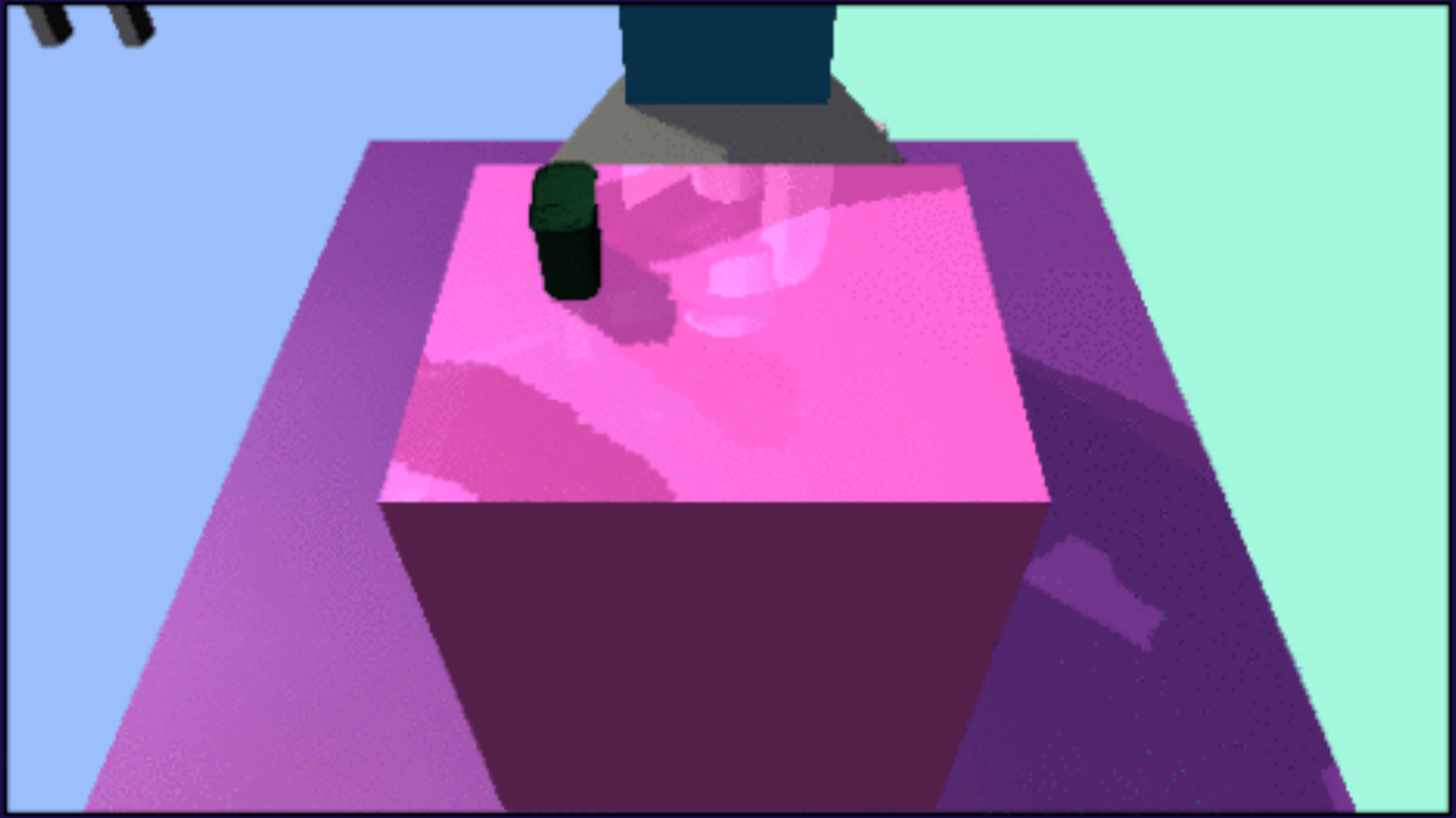
Domain Randomization

Domain Randomization



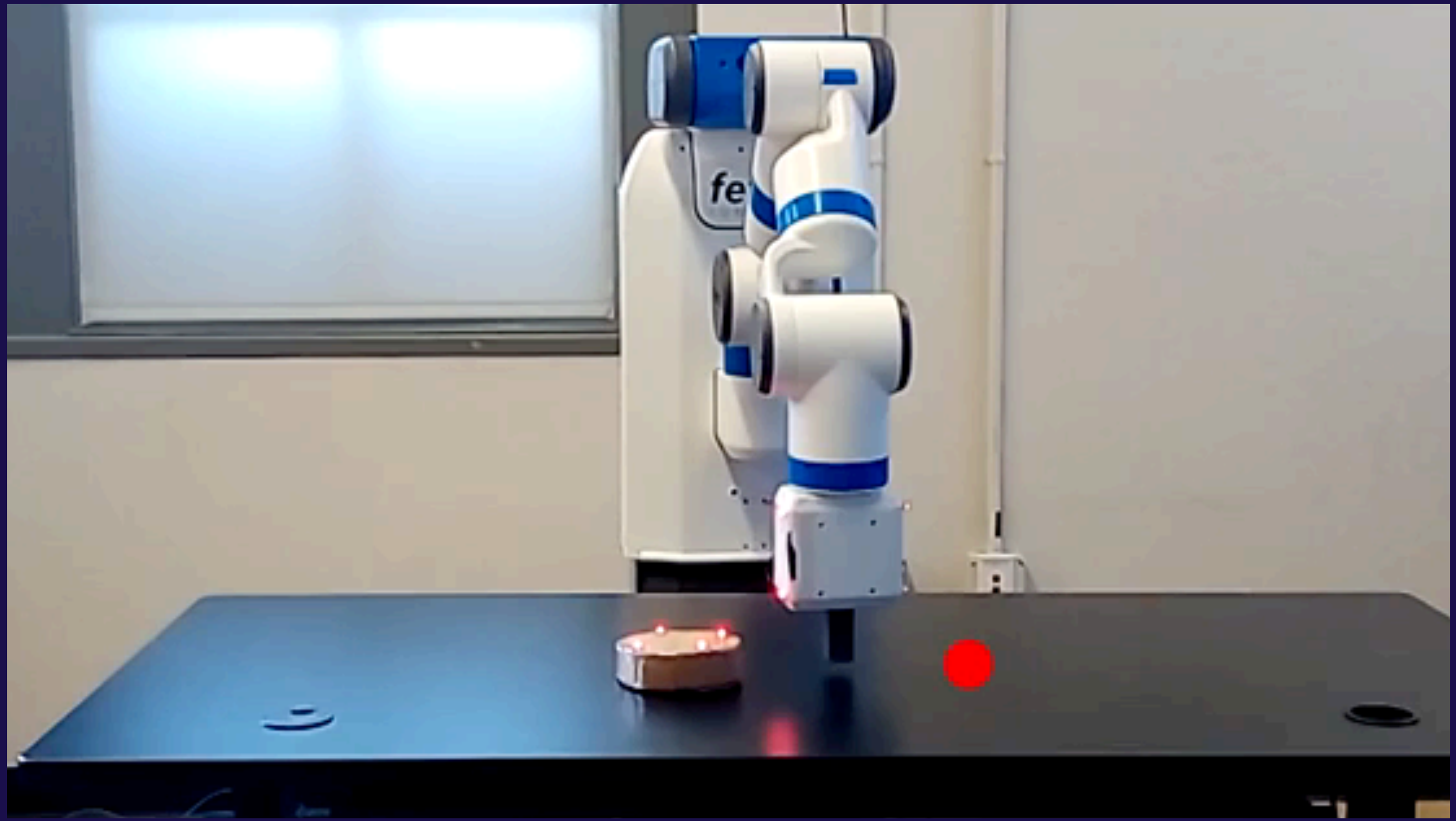
Sadeghi & Levine (2016)

Domain Randomization

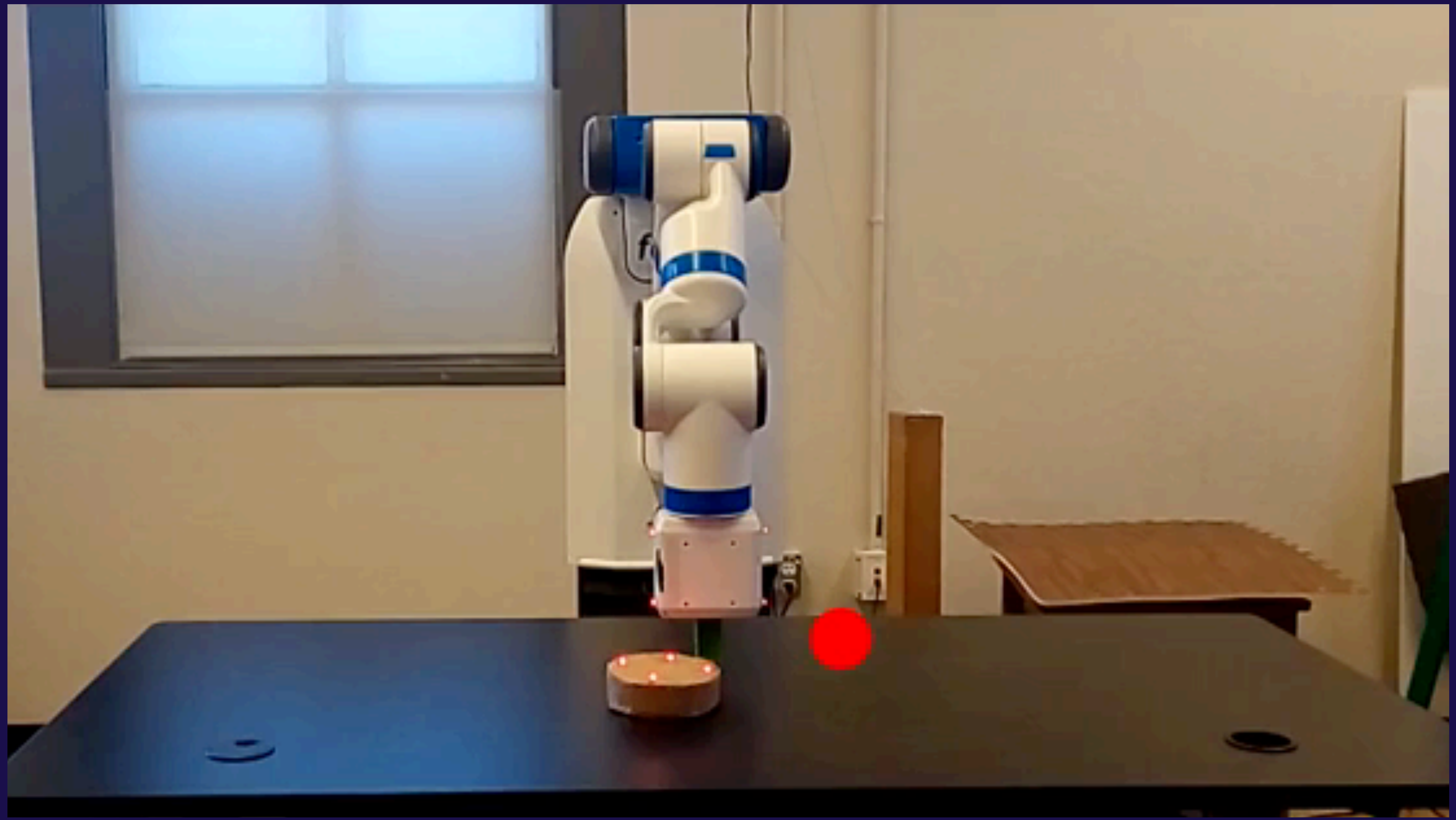


Tobin et al. (2017)

Physics Randomization



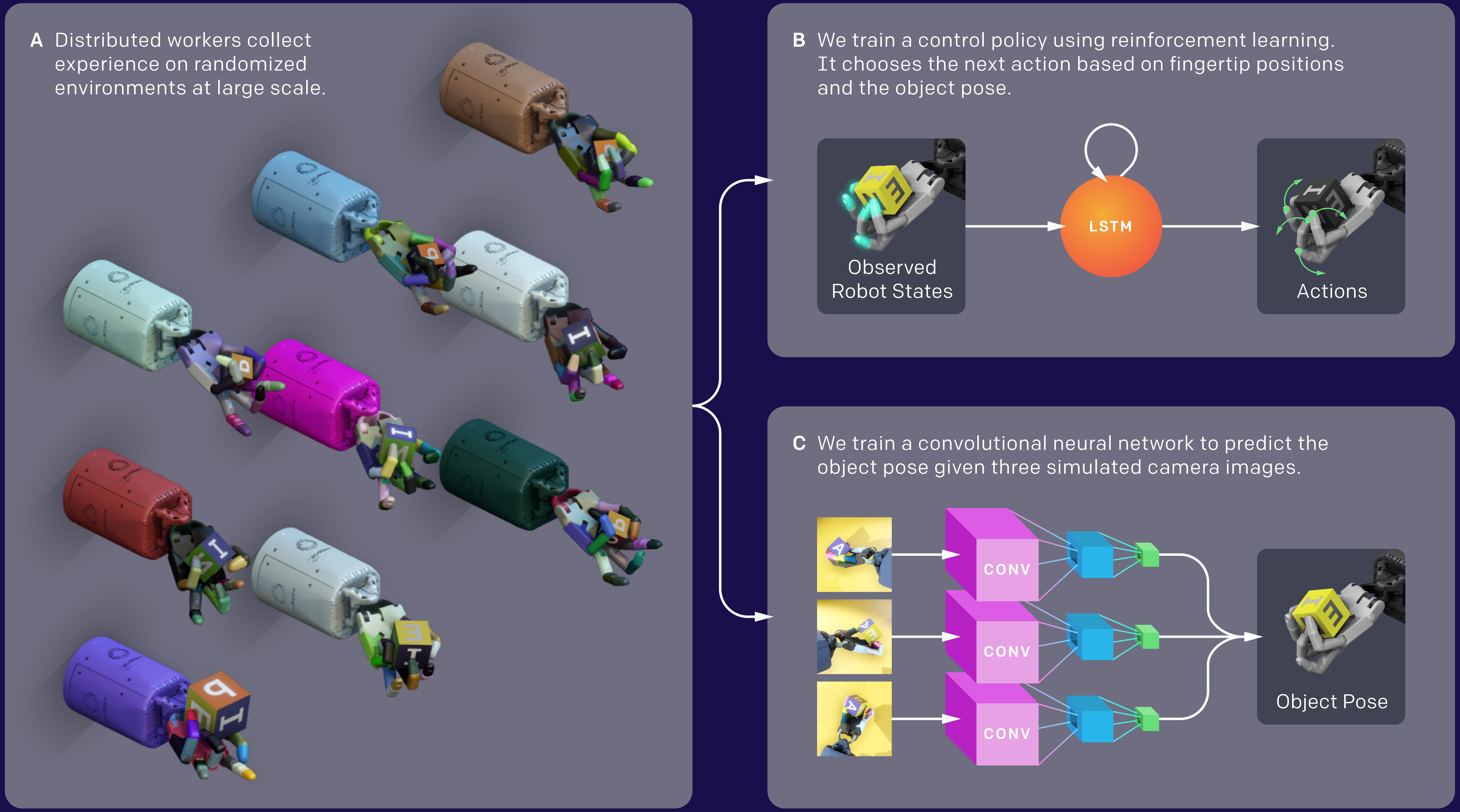
Physics randomization



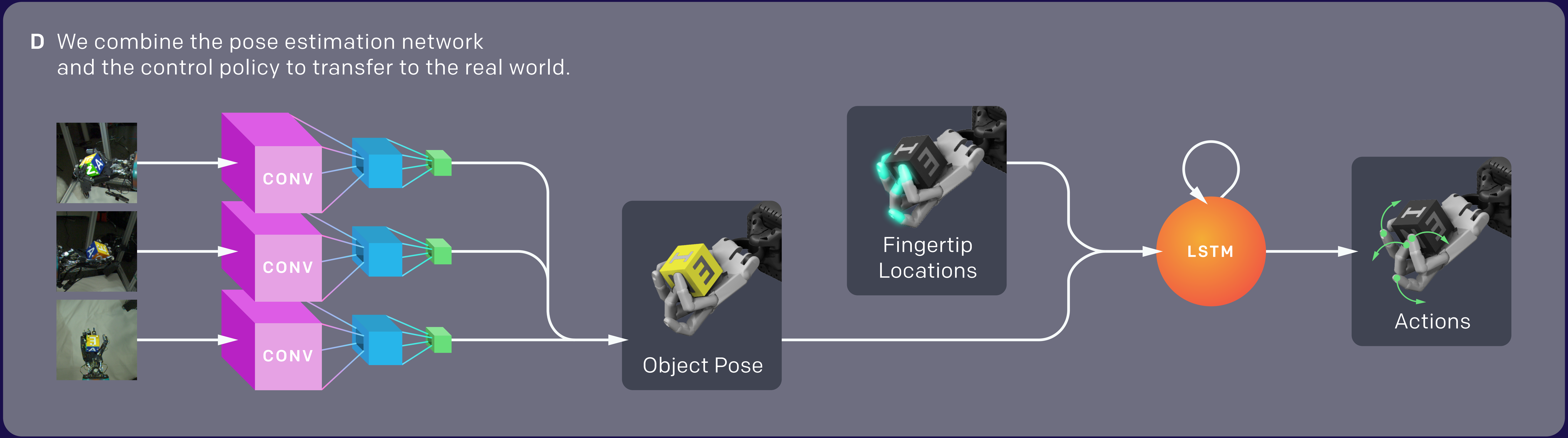
No randomizations

Peng et al. (2017)

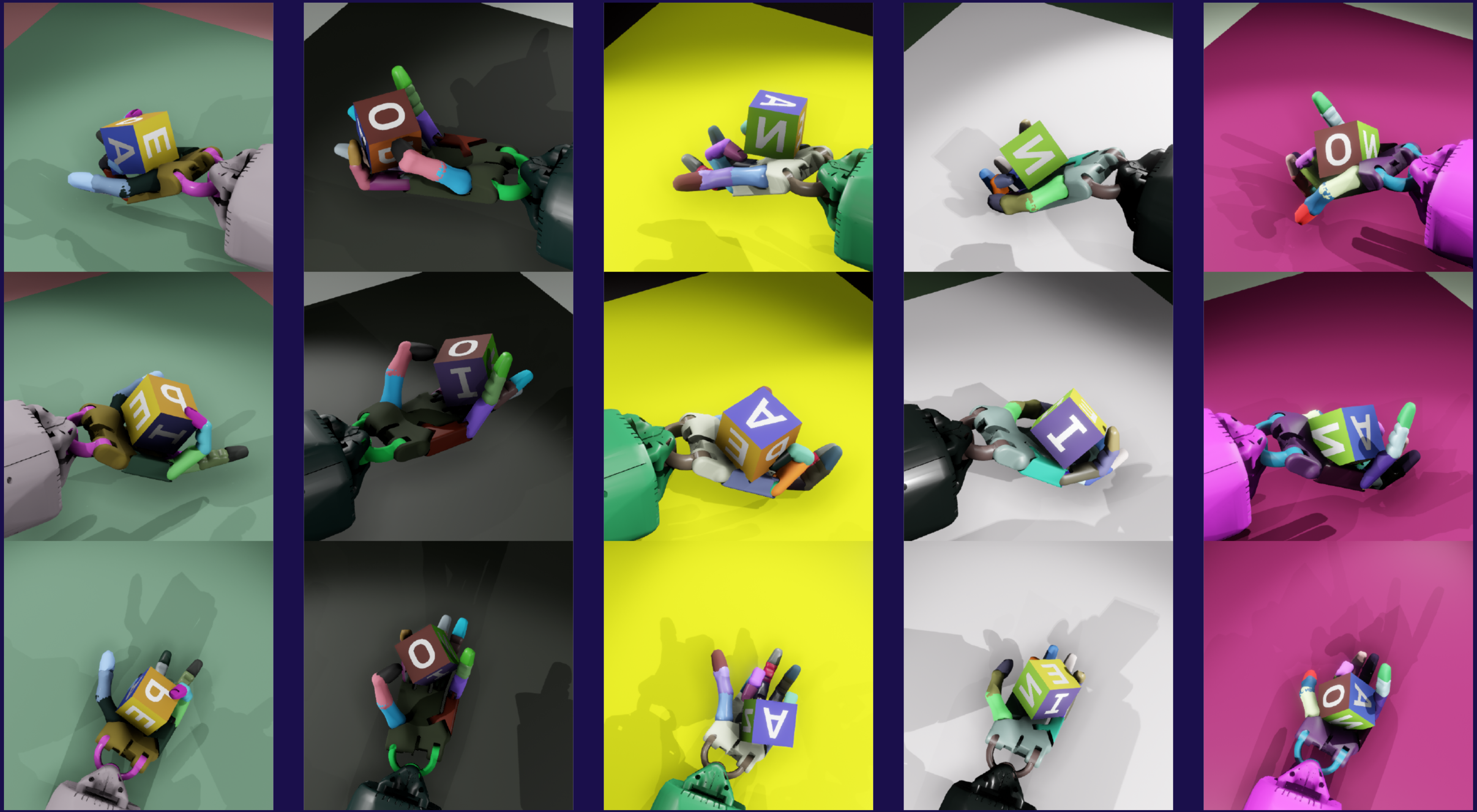
Our Approach


A Distributed workers collect experience on randomized environments at large scale.
B We train a control policy using reinforcement learning. It chooses the next action based on fingertip positions and the object pose.
Observed Robot States
LSTM
Actions
C We train a convolutional neural network to predict the object pose given three simulated camera images.
CONV
CONV
CONV
Object Pose

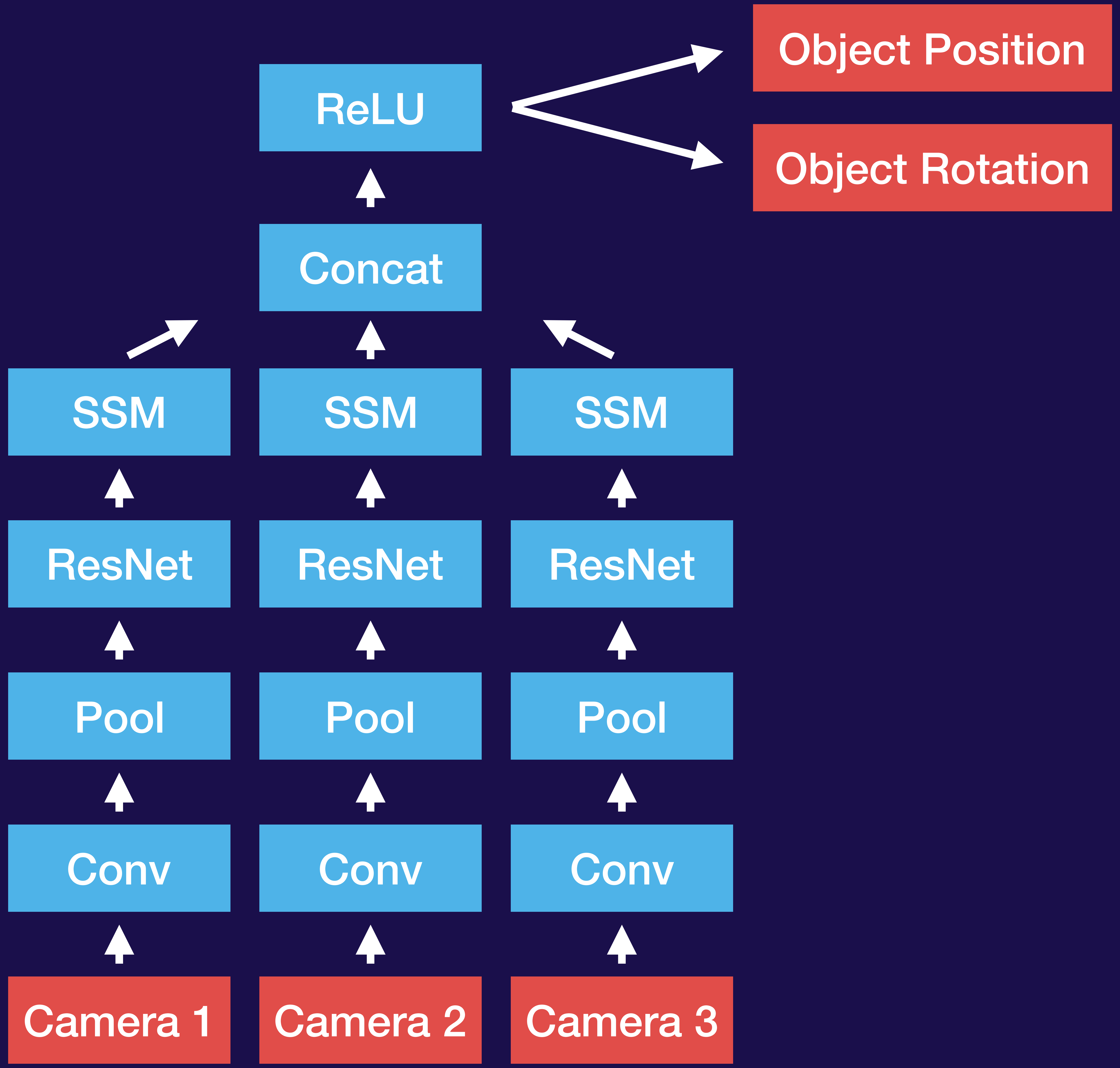
Transfer to the Real World



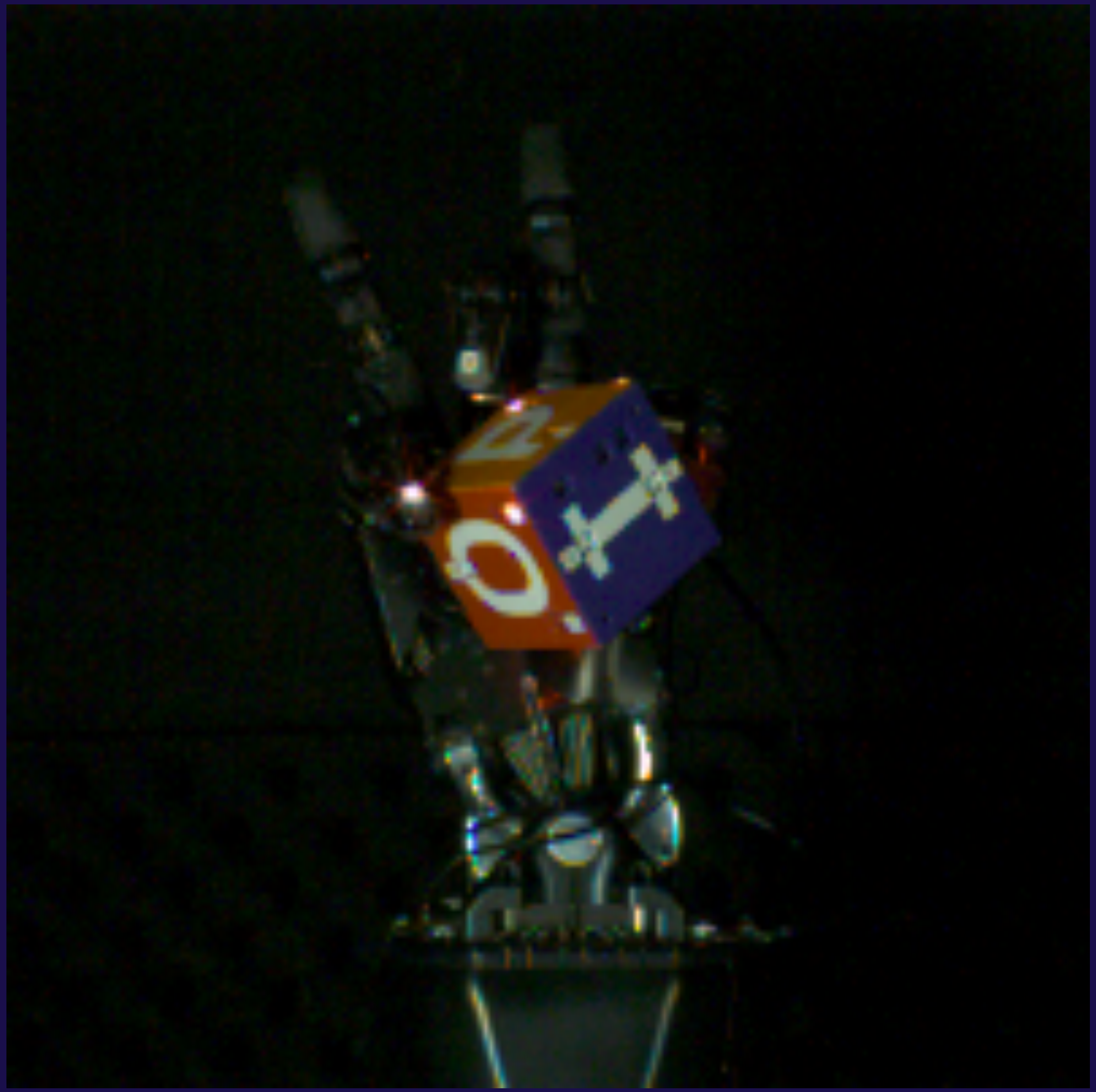
Appearance Randomizations

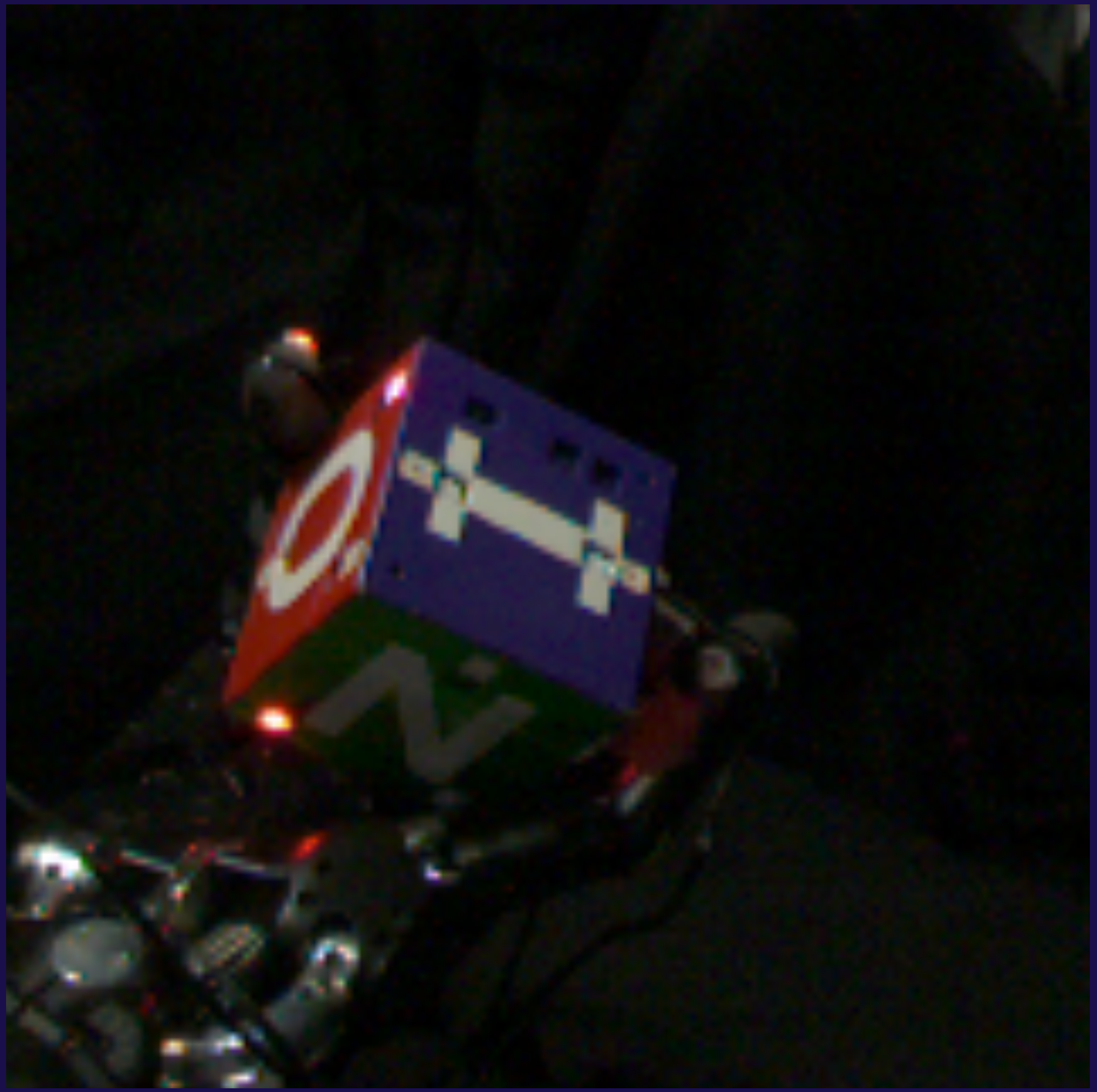


Vision Architecture



What the Model Sees





Physics Randomizations

object dimensions

object and robot link masses

surface friction coefficients

robot joint damping coefficients

actuator force gains

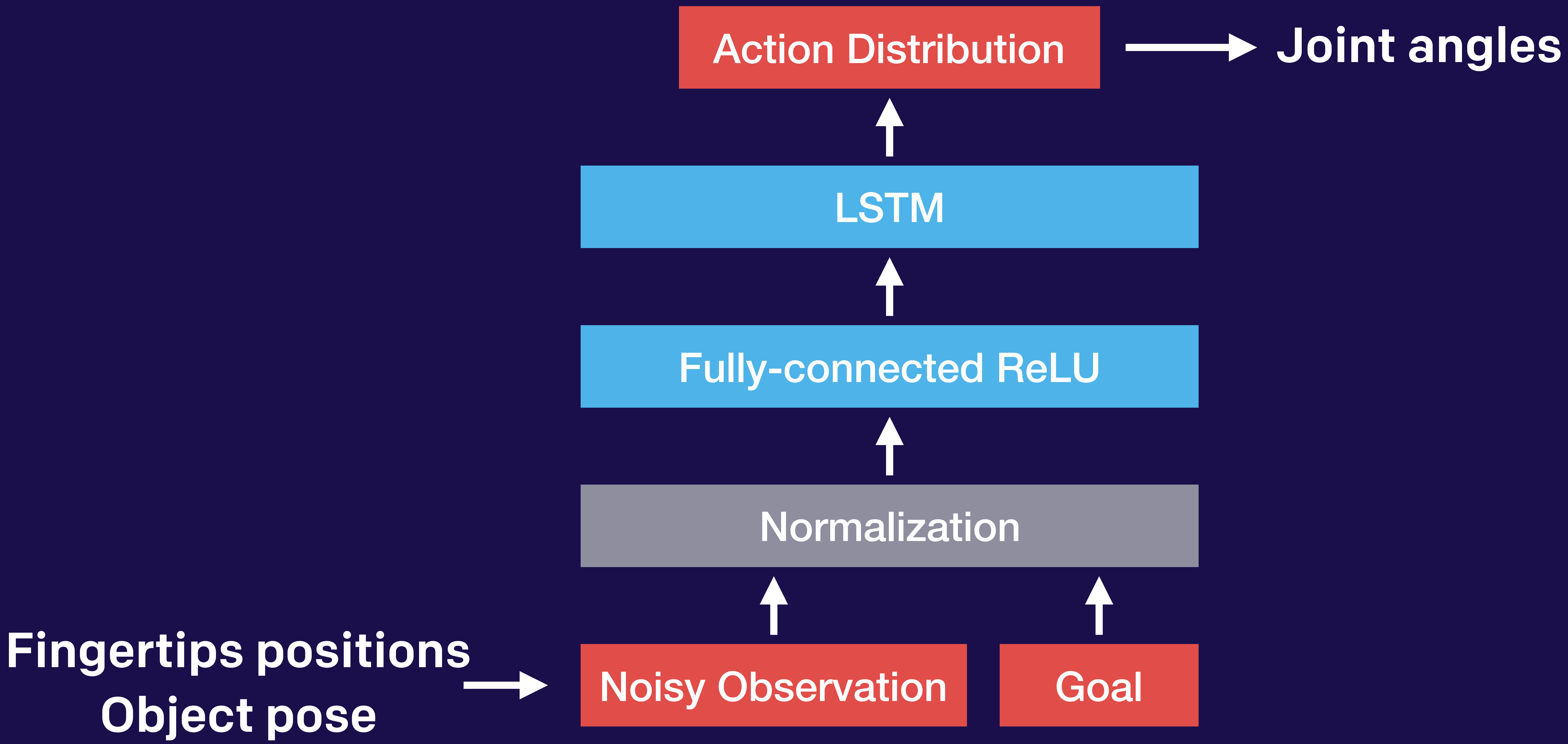
joint limits

gravity vector

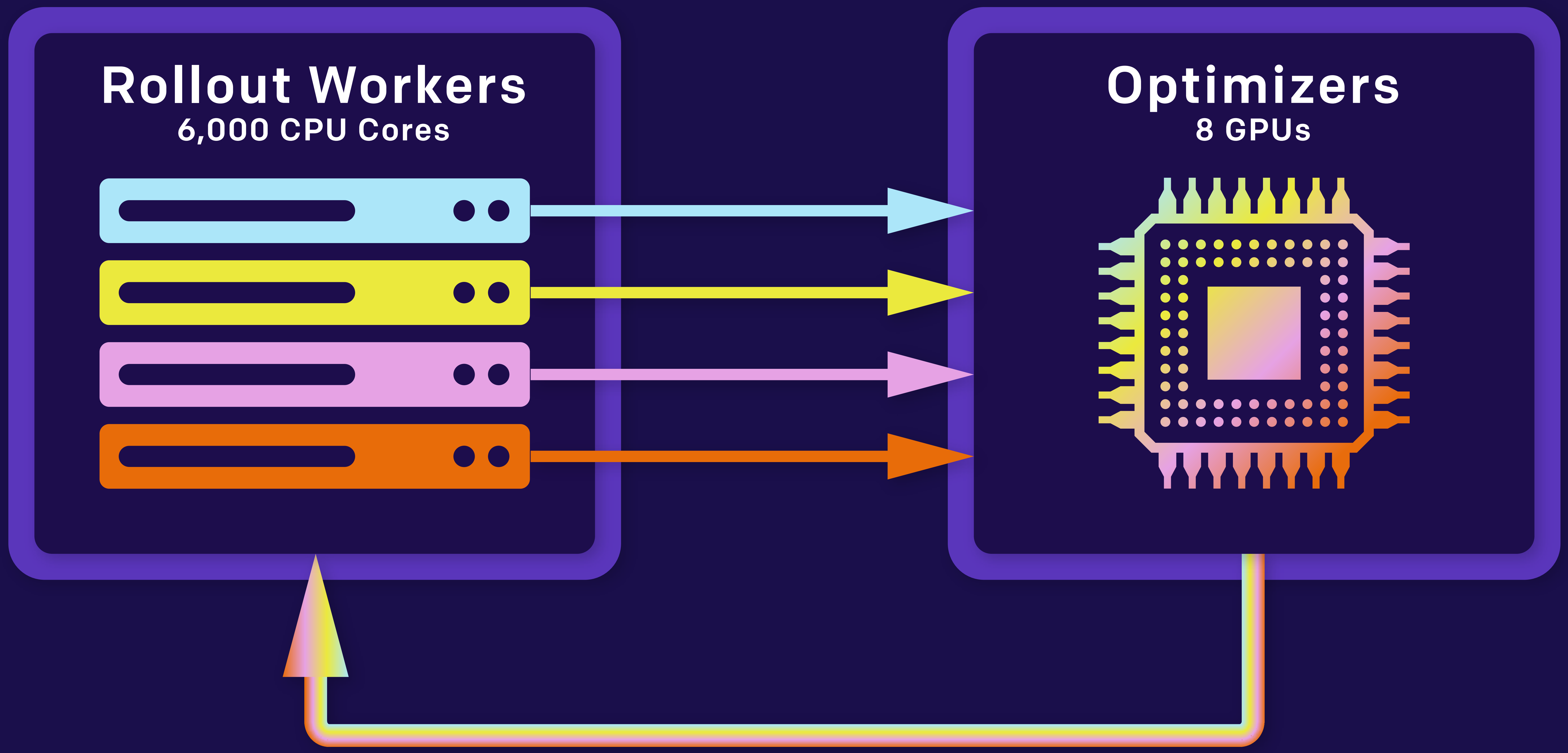
noisy observations

noisy actions

Policy Architecture



Distributed Training with PPO



Policy Parameters

Value Function

policy
inputs
(noisy)

fingertip positions

object pose

target orientation

hand joints angles

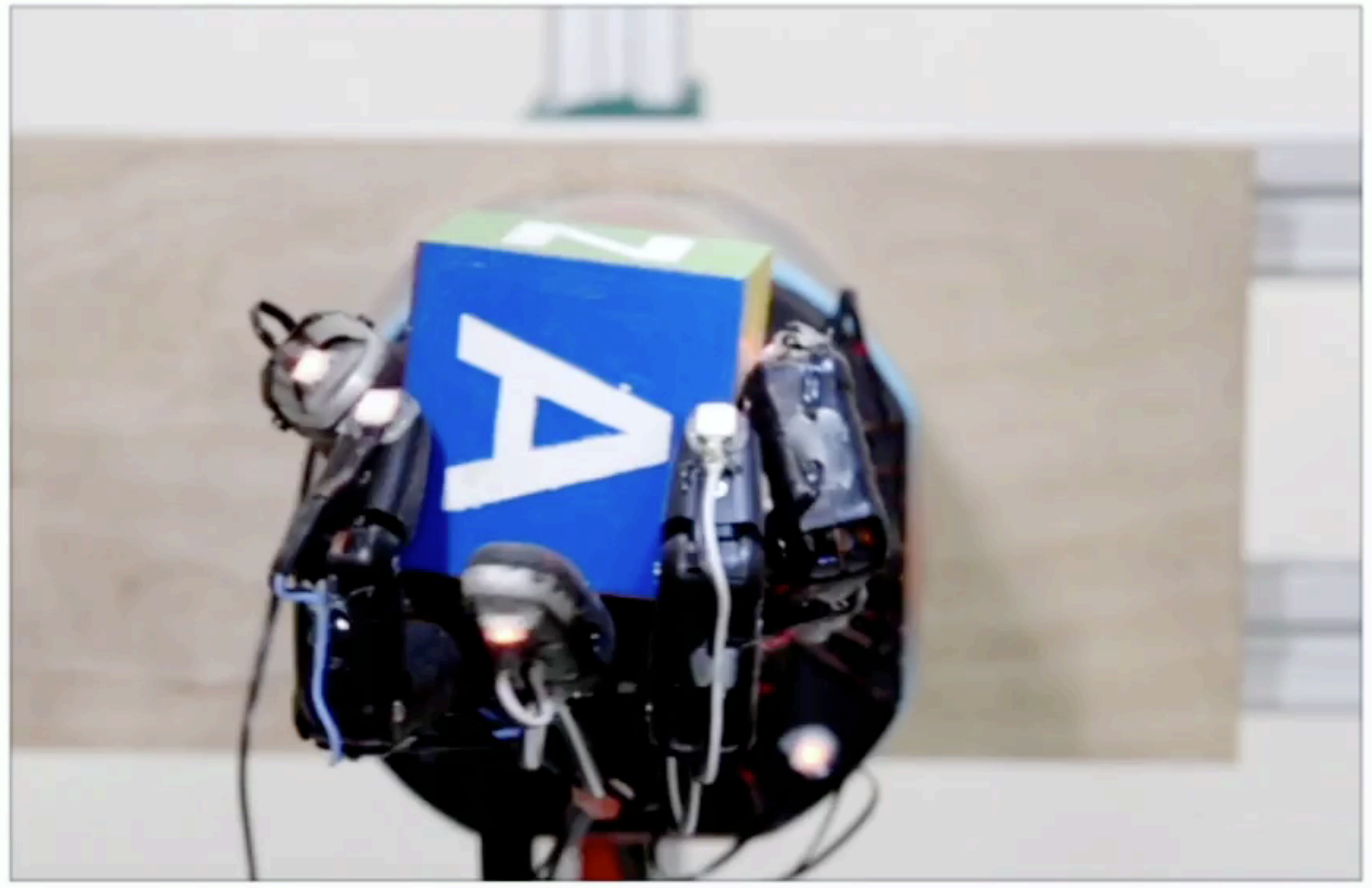
hand joints velocities

object velocity

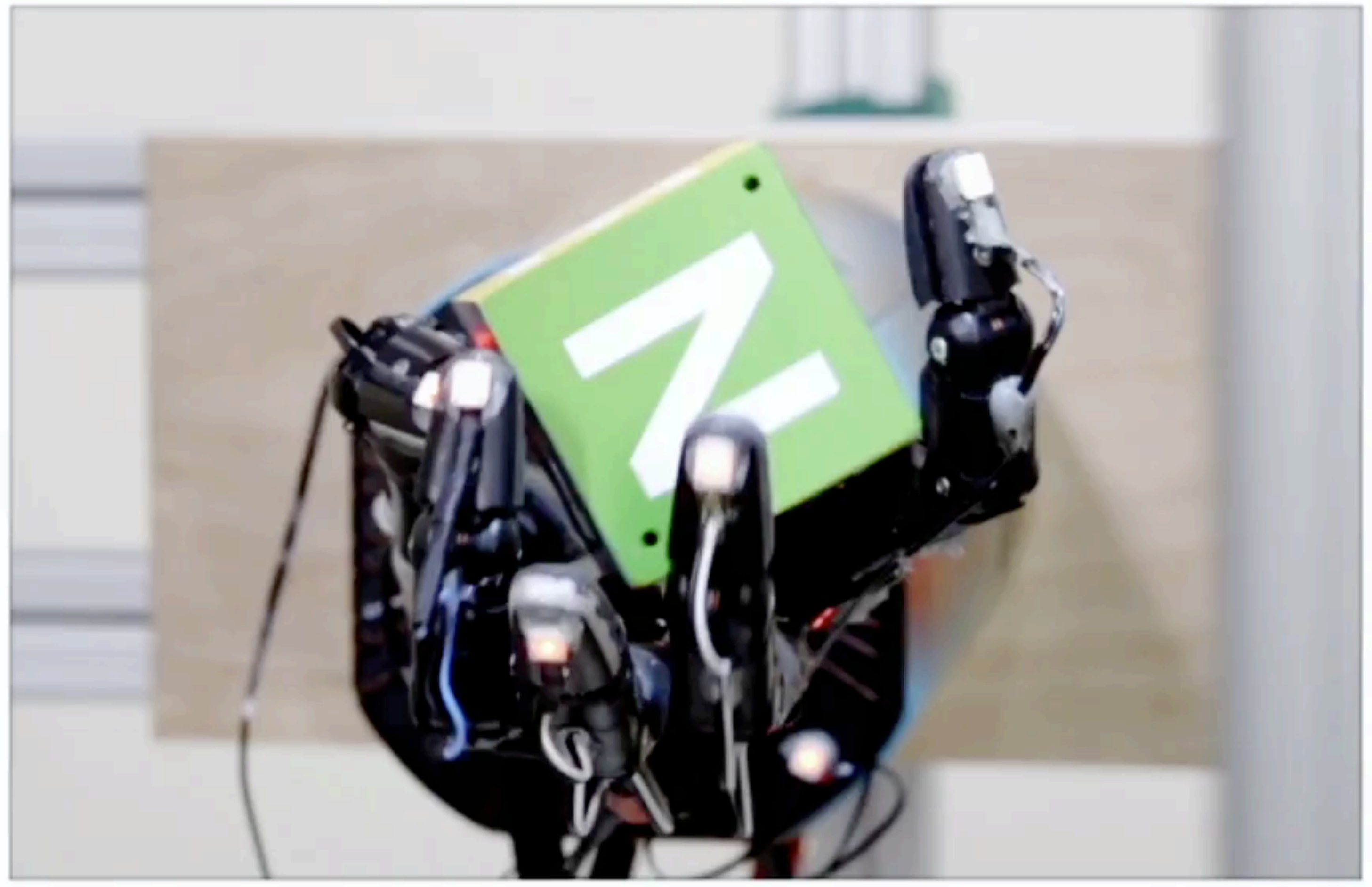
value function
inputs
(noise-free)

Results

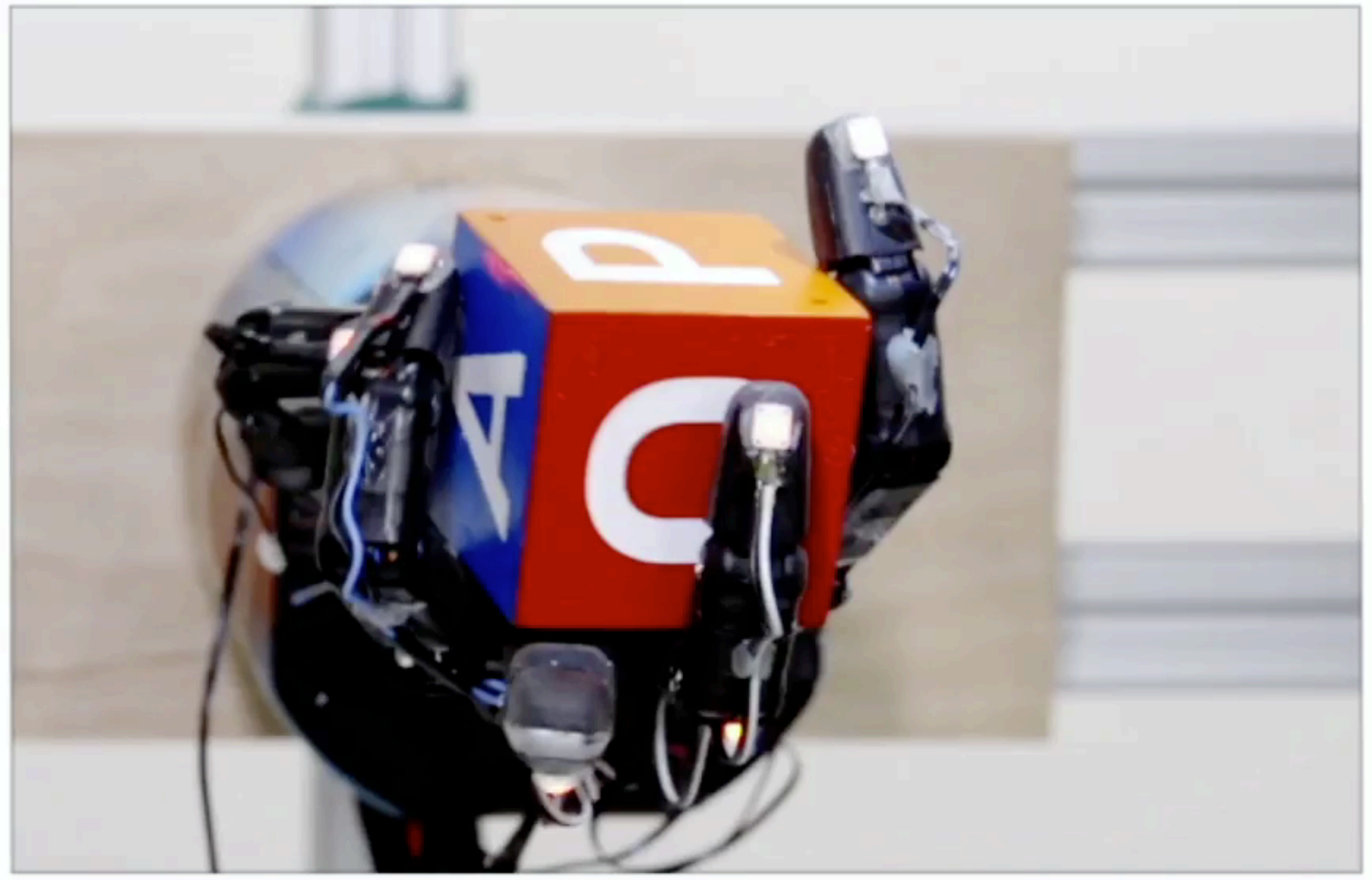
Emergent Behaviors



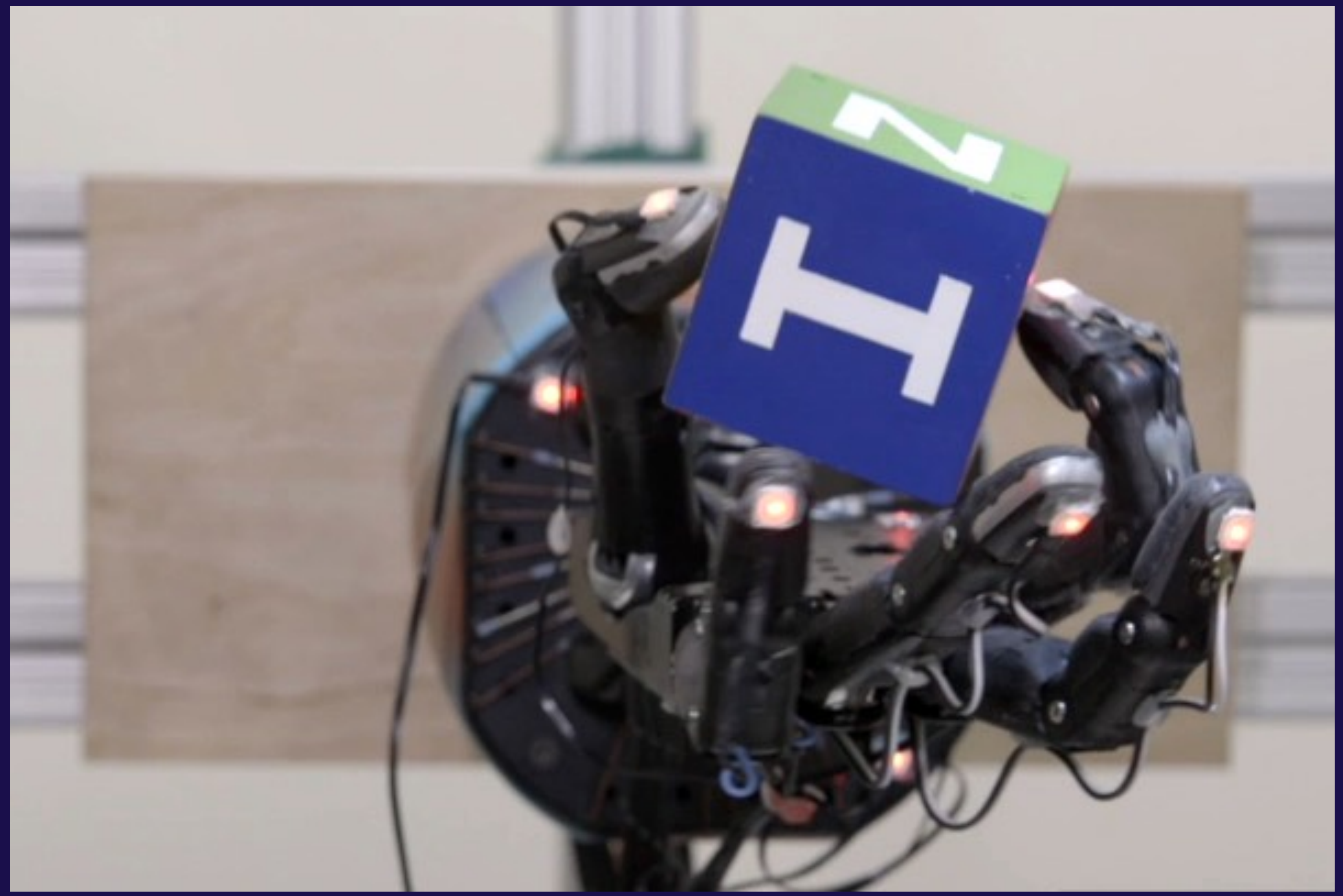
FINGER PIVOTING



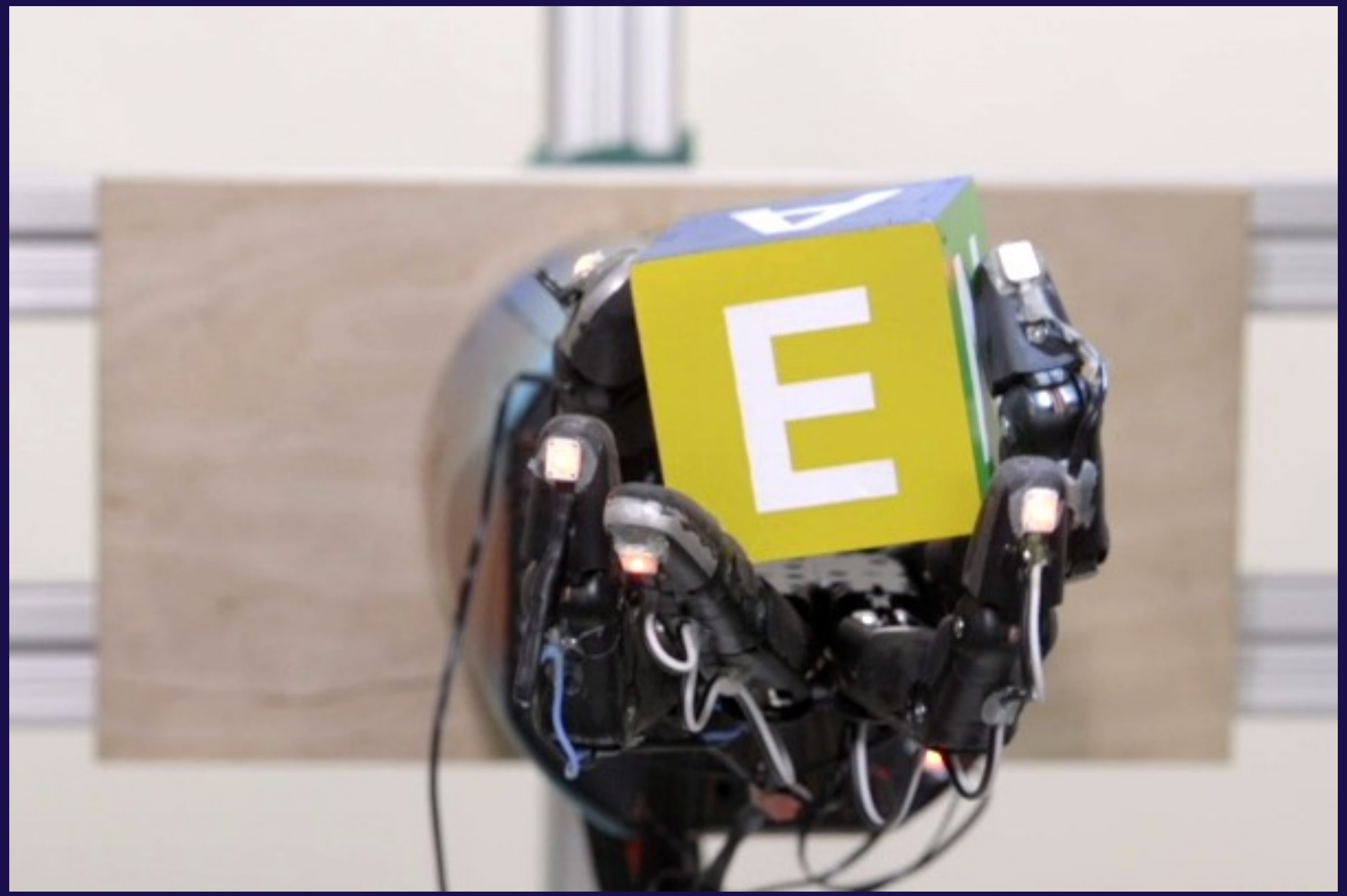
SLIDING



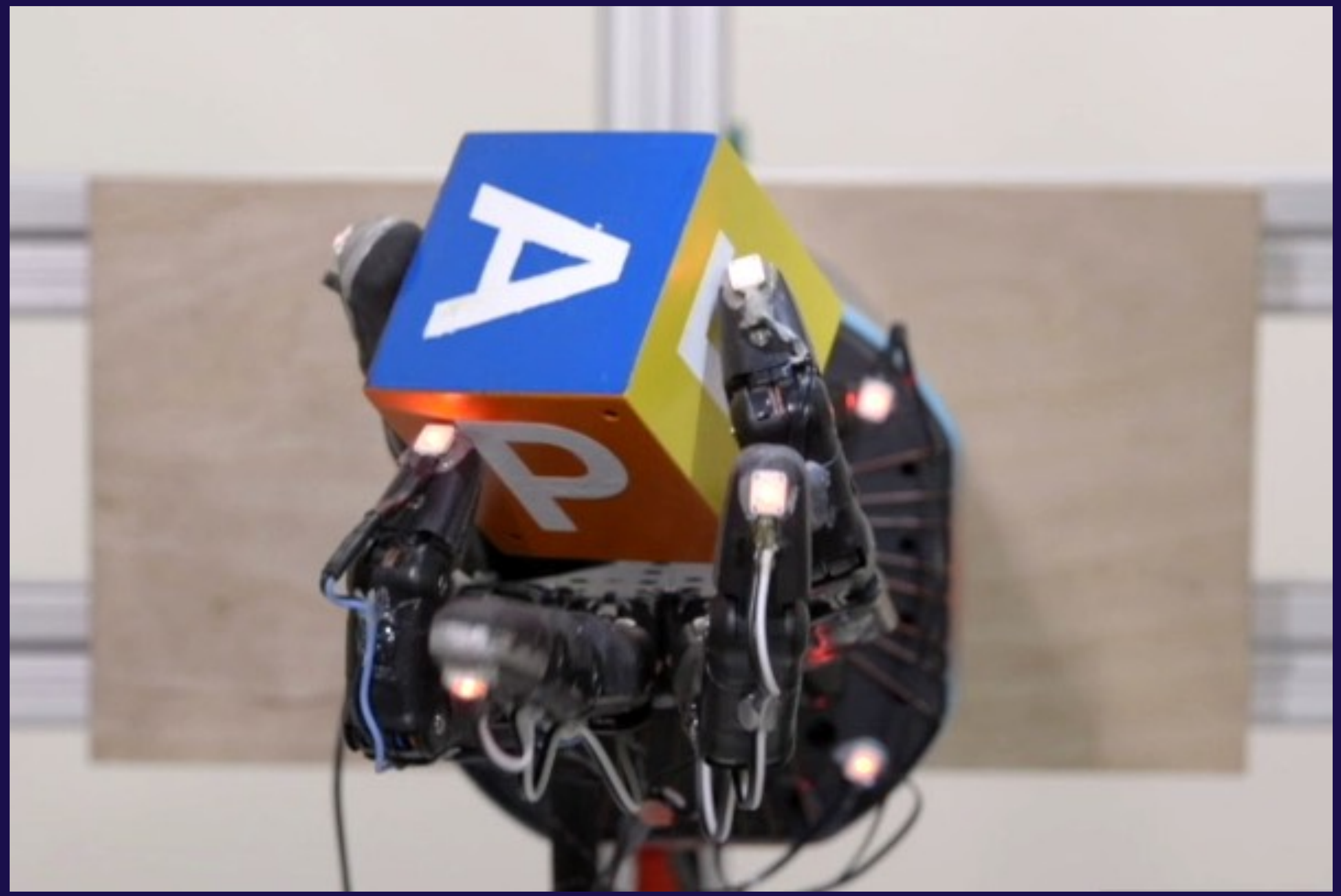
FINGER GAITING



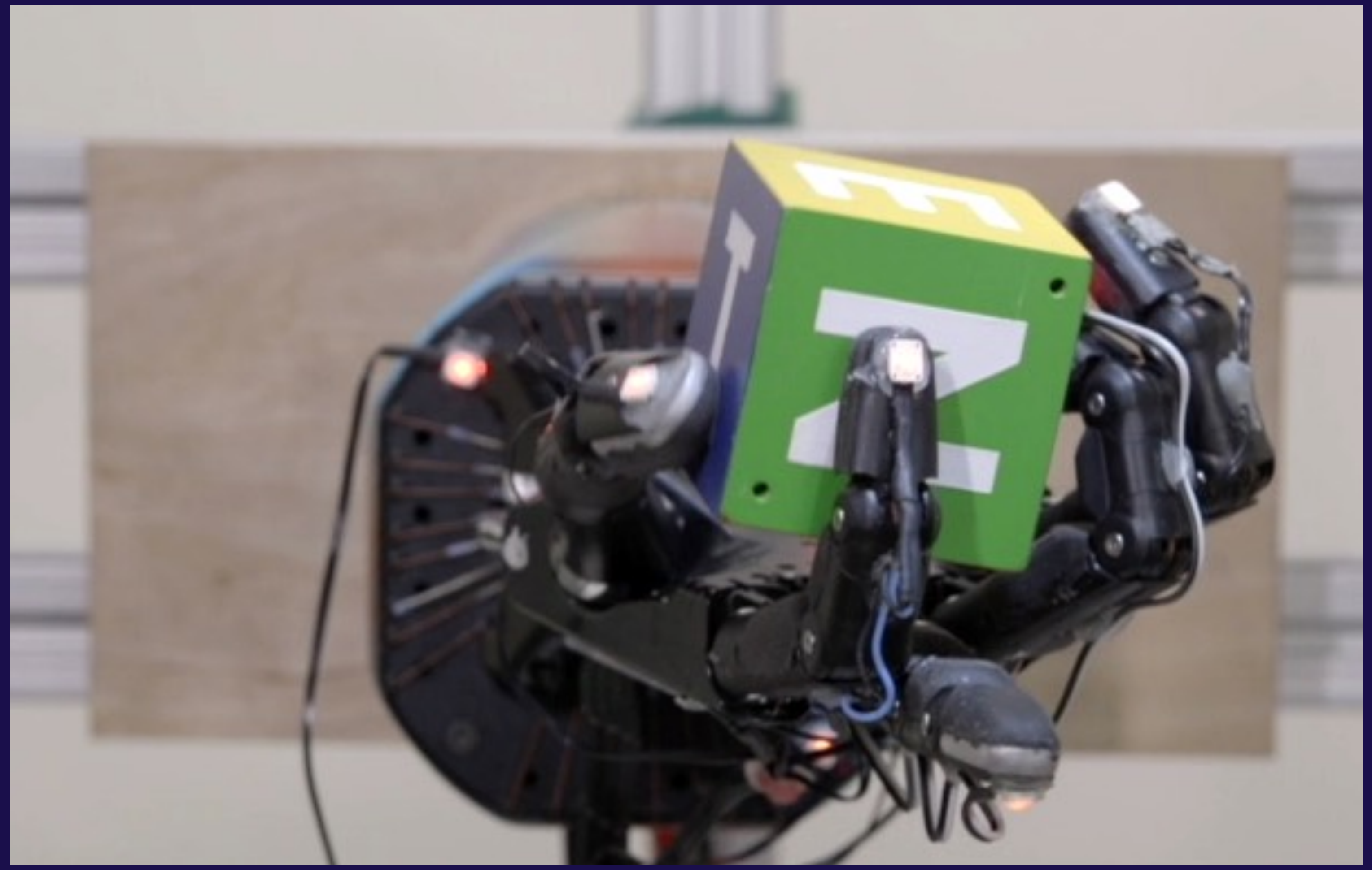
Tip Pinch Grasp



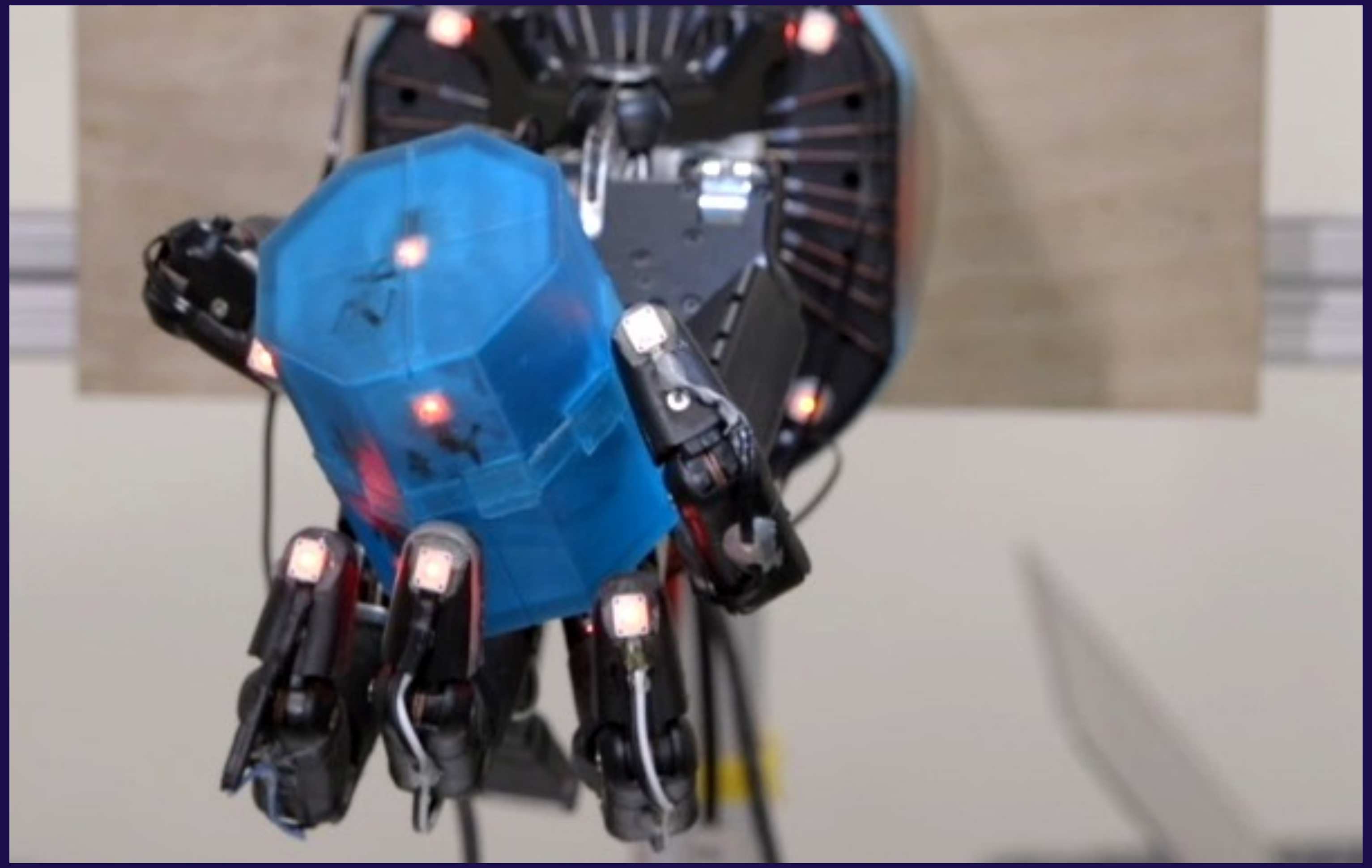
Palmar Pinch Grasp



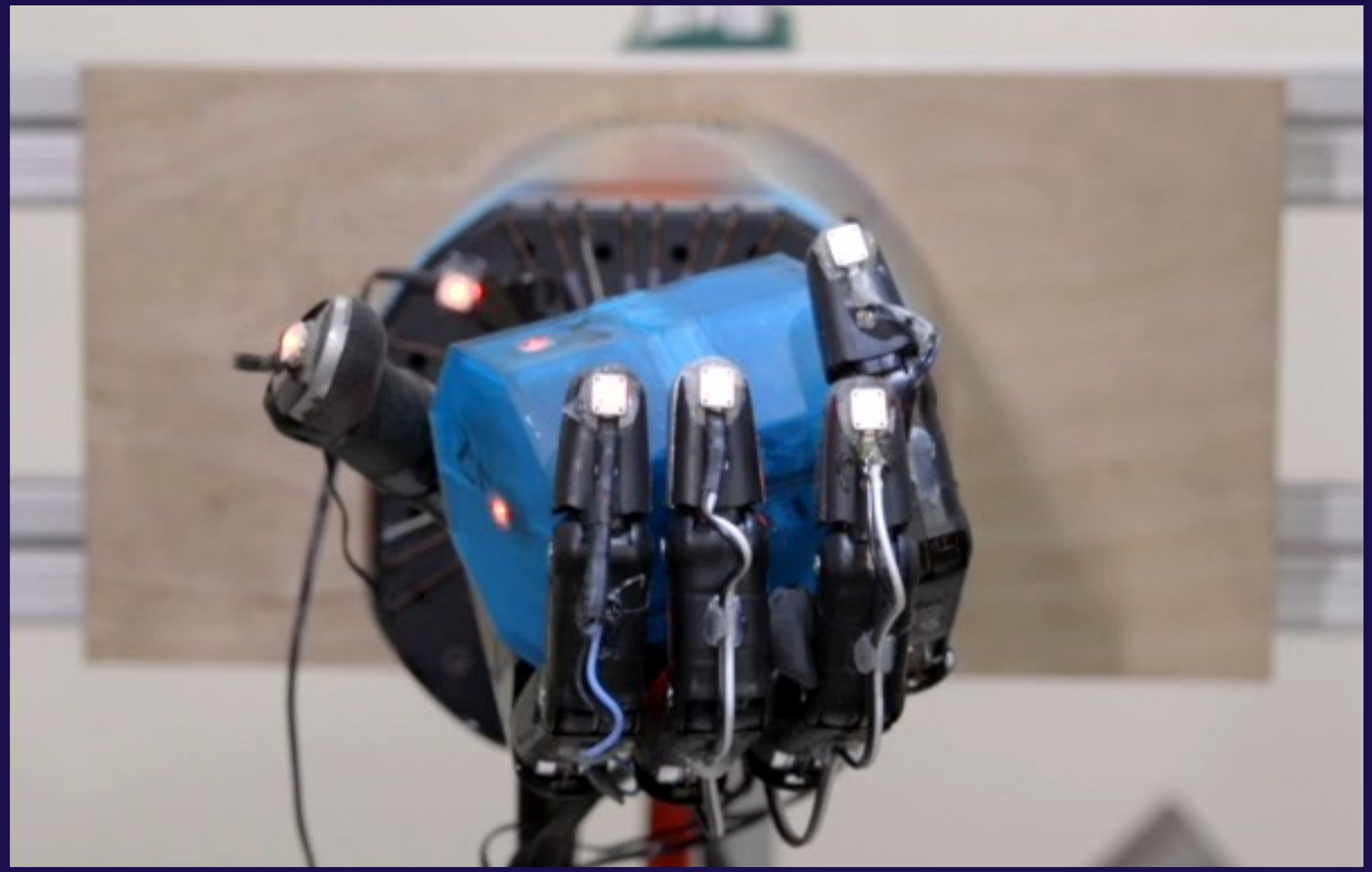
Tripod Grasp



Quadpod Grasp



5-Finger Precision Grasp



Power Grasp

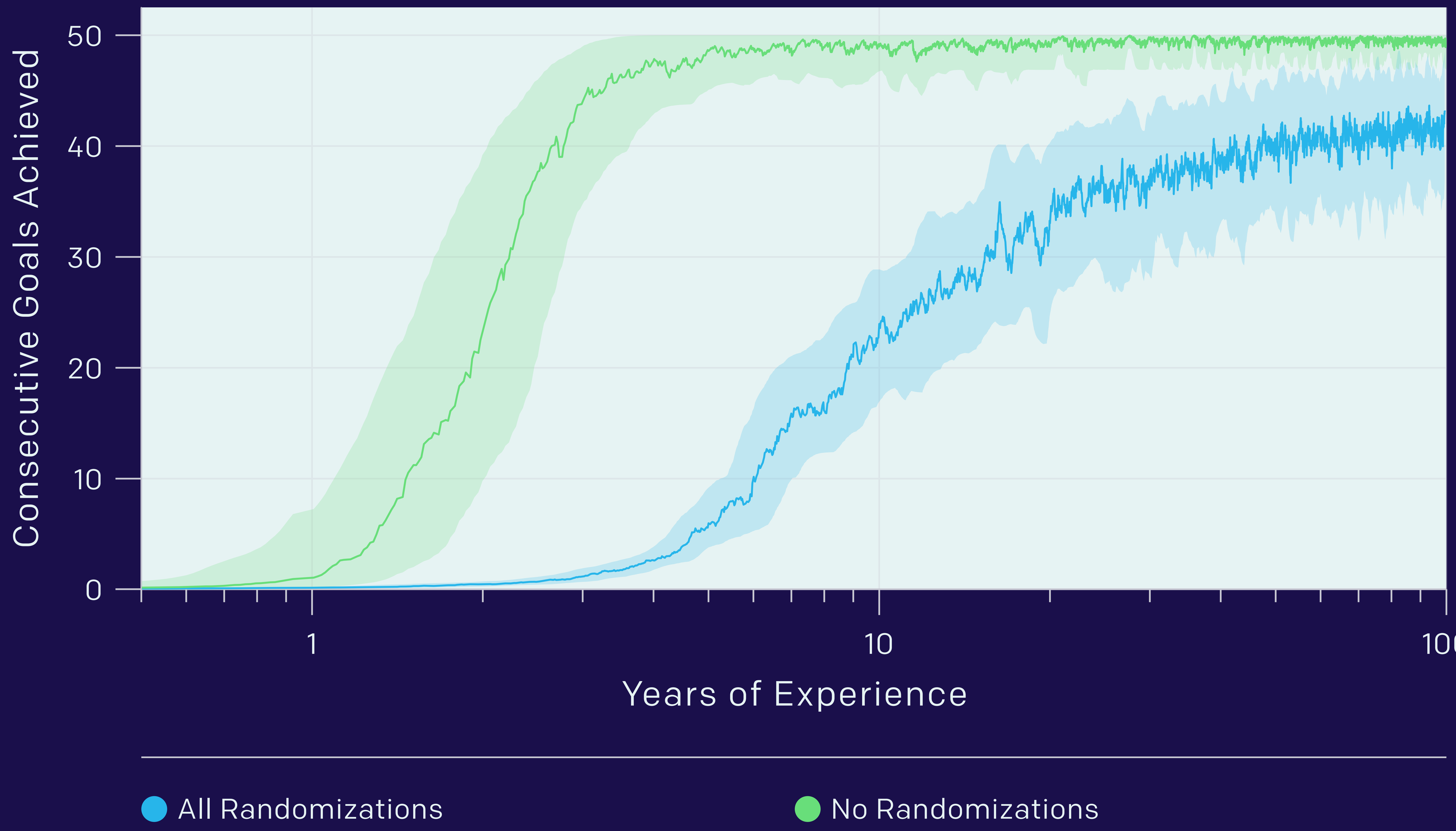
Quantitative results

RANDOMIZATONS	OBJECT TRACKING	MEDIAN NUMBER OF SUCCESSES
None	Motion tracking	0
All	Motion tracking	13
All	Vision	11.5

Effect of Memory

POLICY	VALUE FUNCTION	MEDIAN NUMBER OF SUCCESSES
LSTM	LSTM	13
Feedforward	LSTM	3.5
Feedforward	Feedforward	3

Training time



Distribution of environments

+

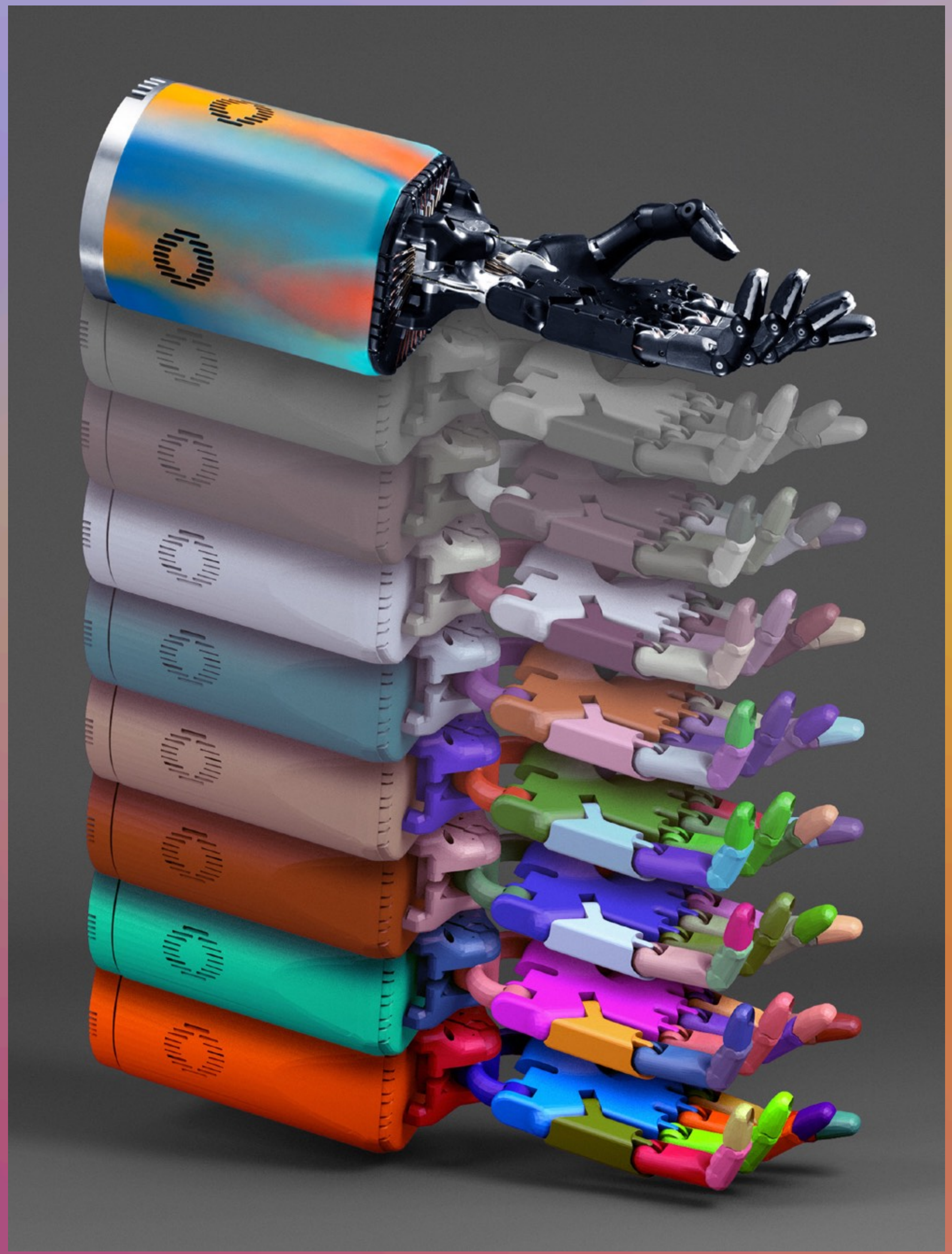
Memory

=

Meta-Learning

Thank You!

Blog Post



blog.openai.com/learning-dexterity

Paper

arXiv:1808.00177v2 [cs.LG] 28 Aug 2018

Learning Dexterous In-Hand Manipulation

OpenAI*

Initial configuration

Goal configuration

Figure 1: A five-fingered humanoid hand trained with reinforcement learning manipulating a block from an initial configuration to a goal configuration using vision for sensing.

Abstract

We use reinforcement learning (RL) to learn dexterous in-hand manipulation policies which can perform vision-based object reorientation on a physical Shadow Dexterous Hand. The training is performed in a simulated environment in which we randomize many of the physical properties of the system like friction coefficients and an object's appearance. Our policies transfer to the physical robot despite being trained entirely in simulation. Our method does not rely on any human demonstrations, but many behaviors found in human manipulation emerge naturally, including finger gaiting, multi-finger coordination, and the controlled use of gravity. Our results were obtained using the same distributed RL system that was used to train OpenAI Five [43]. We also include a video of our results: `https://youtu.be/jwSbzNHGflM`.

1 Introduction

While dexterous manipulation of objects is a fundamental everyday task for humans, it is still challenging for autonomous robots. Modern-day robots are typically designed for specific tasks in constrained settings and are largely unable to utilize complex end-effectors. In contrast, people are able to perform a wide range of dexterous manipulation tasks in a diverse set of environments, making the human hand a grounded source of inspiration for research into robotic manipulation.

The Shadow Dexterous Hand [58] is an example of a robotic hand designed for human-level dexterity; it has five fingers with a total of 24 degrees of freedom. The hand has been commercially available

*Built by a team of researchers and engineers at OpenAI (in alphabetical order).

Marcin Andrychowicz Bowen Baker Maciek Chociej Rafał Józefowicz Bob McGrew Jakub Pachocki Arthur Petron Matthias Plappert Glenn Powell Alex Ray Jonas Schneider Szymon Sidor Josh Tobin Peter Welinder Lilian Weng Wojciech Zaremba

arxiv.org/abs/1808.00177

FOLLOW @OPENAI ON TWITTER

RANDOMIZATONS	OBJECT TRACKING	POLICY	NUMBER OF SUCCESSES	
			MEDIAN	MAX
None	Motion tracking	LSTM	0	6
All	Motion tracking	LSTM	13	50
All	Vision	LSTM	11.5	46
All	Motion tracking	FF	3.5	15

RANDOMIZATONS	OBJECT TRACKING	POLICY	NUMBER OF SUCCESSES	
			MEDIAN	MAX
None	Motion tracking	LSTM	0	6
All	Motion tracking	LSTM	13	50
All	Vision	LSTM	11.5	46
All	Motion tracking	FF	3.5	15

RANDOMIZATONS	OBJECT TRACKING	POLICY	NUMBER OF SUCCESSES	
			MEDIAN	MAX
None	Motion tracking	LSTM	0	6
All	Motion tracking	LSTM	13	50
All	Vision	LSTM	11.5	46
All	Motion tracking	FF	3.5	15

RANDOMIZATONS	OBJECT TRACKING	POLICY	NUMBER OF SUCCESSES	
			MEDIAN	MAX
None	Motion tracking	LSTM	0	6
All	Motion tracking	LSTM	13	50
All	Vision	LSTM	11.5	46
All	Motion tracking	FF	3.5	15

RANDOMIZATONS	OBJECT TRACKING	POLICY	NUMBER OF SUCCESSES	
			MEDIAN	MAX
None	Motion tracking	LSTM	0	6
All	Motion tracking	LSTM	13	50
All	Vision	LSTM	11.5	46
All	Motion tracking	FF	3.5	15