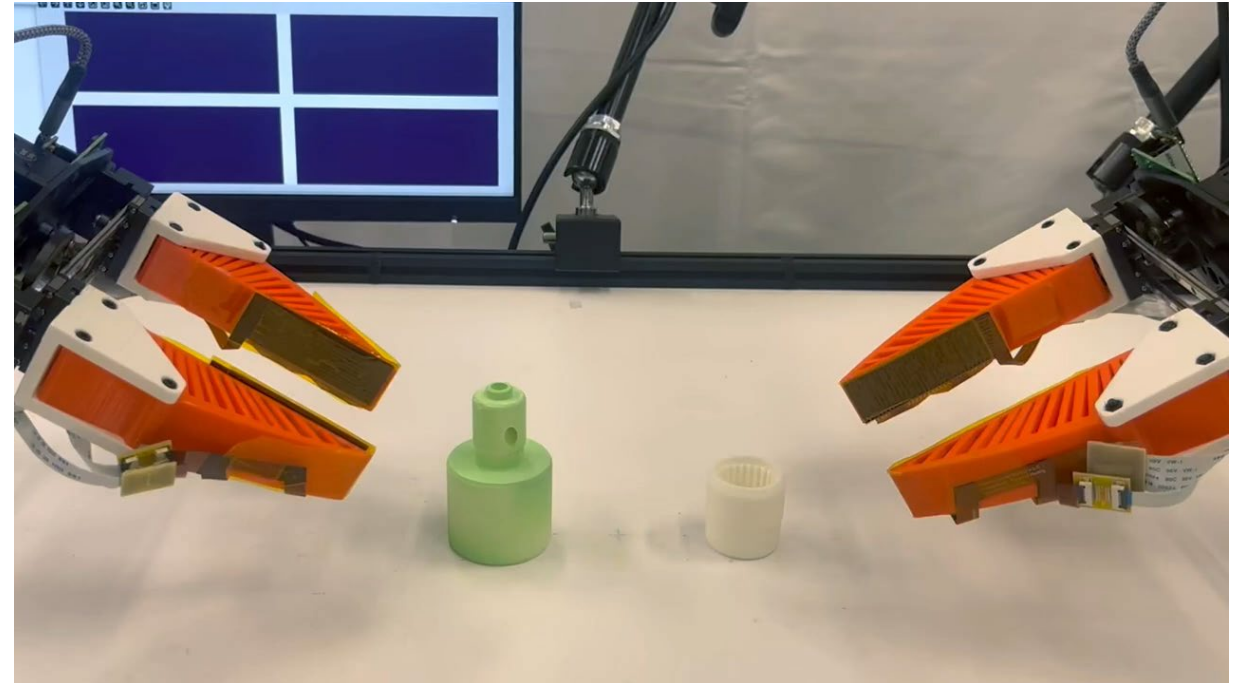
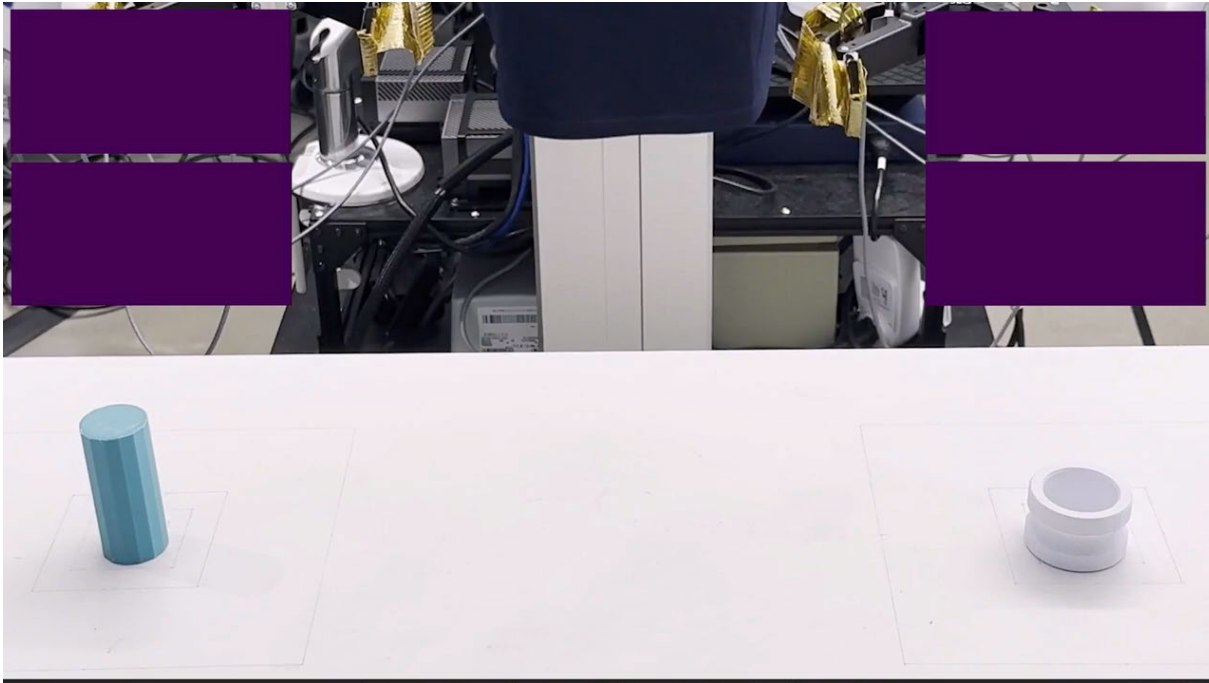


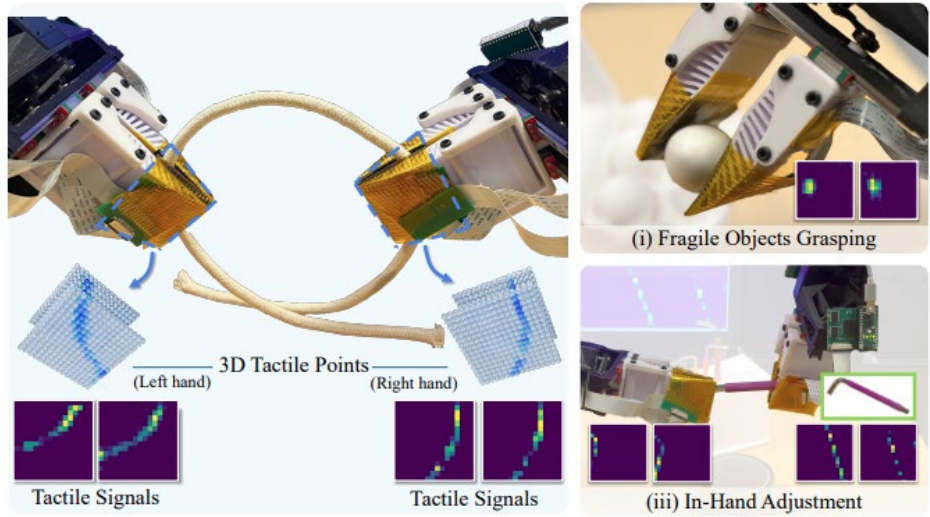
VT-Refine: Learning Bimanual Assembly with Visuo-Tactile Feedback via Simulation Fine-Tuning

Presenter: Binghao Huang



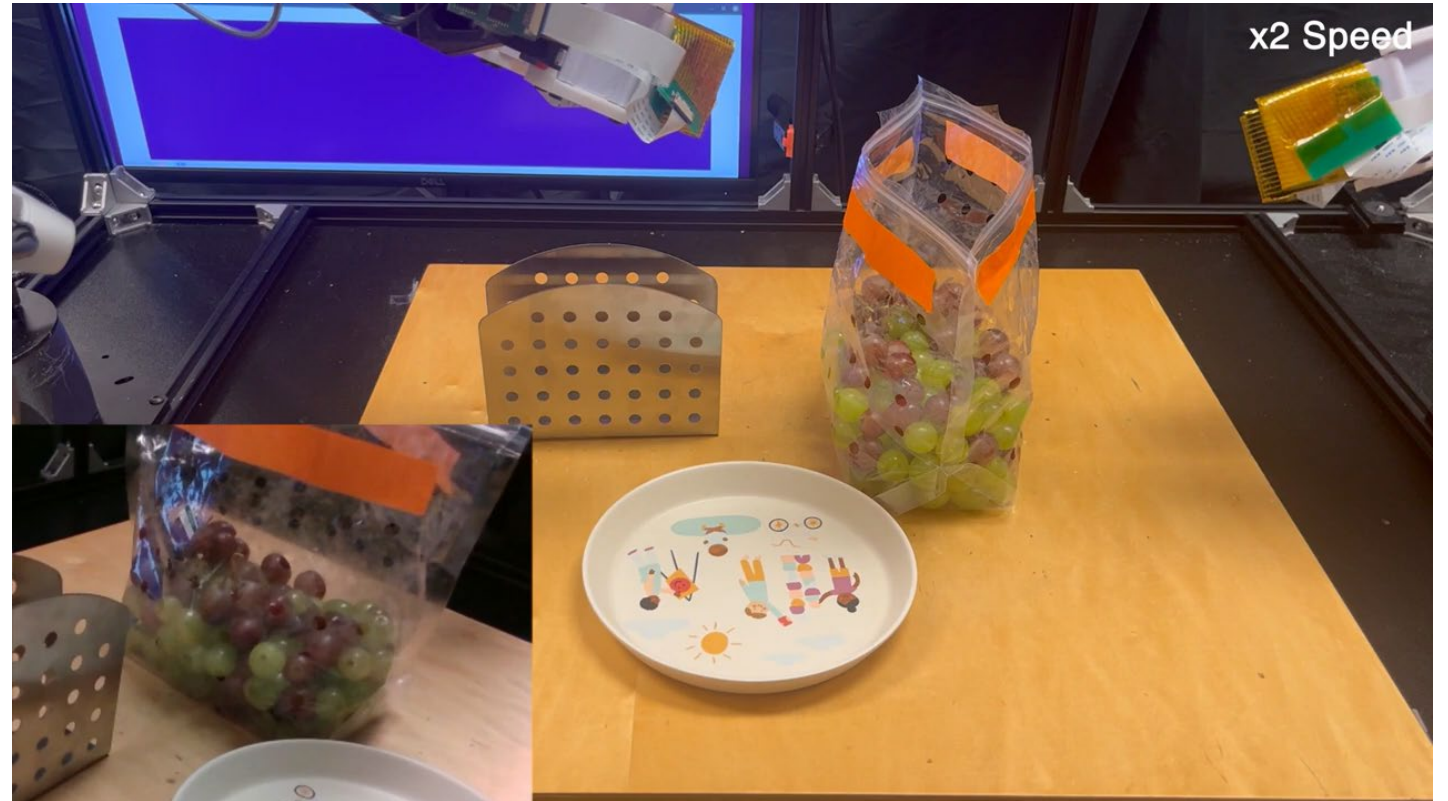
RESEARCH OVERVIEW

Project 1



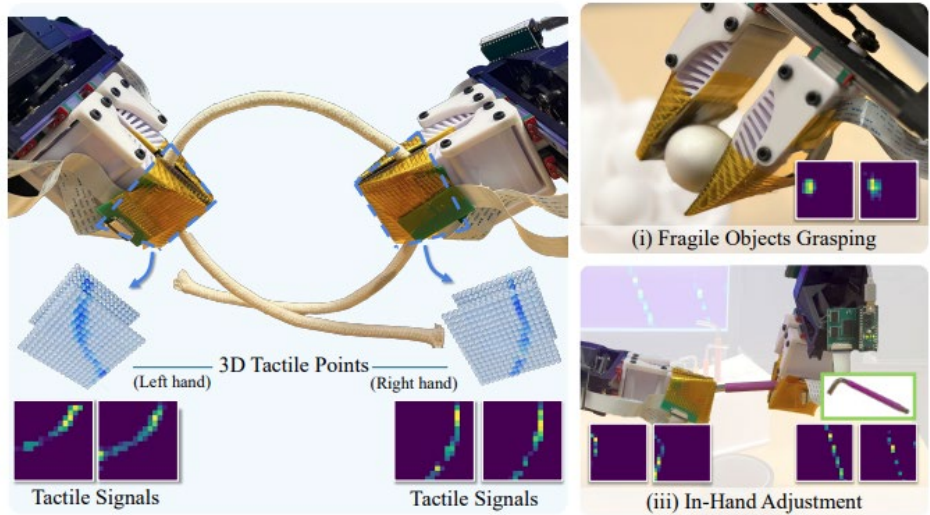
3D-VITAC

Learning with Real Visuo-Tactile Data



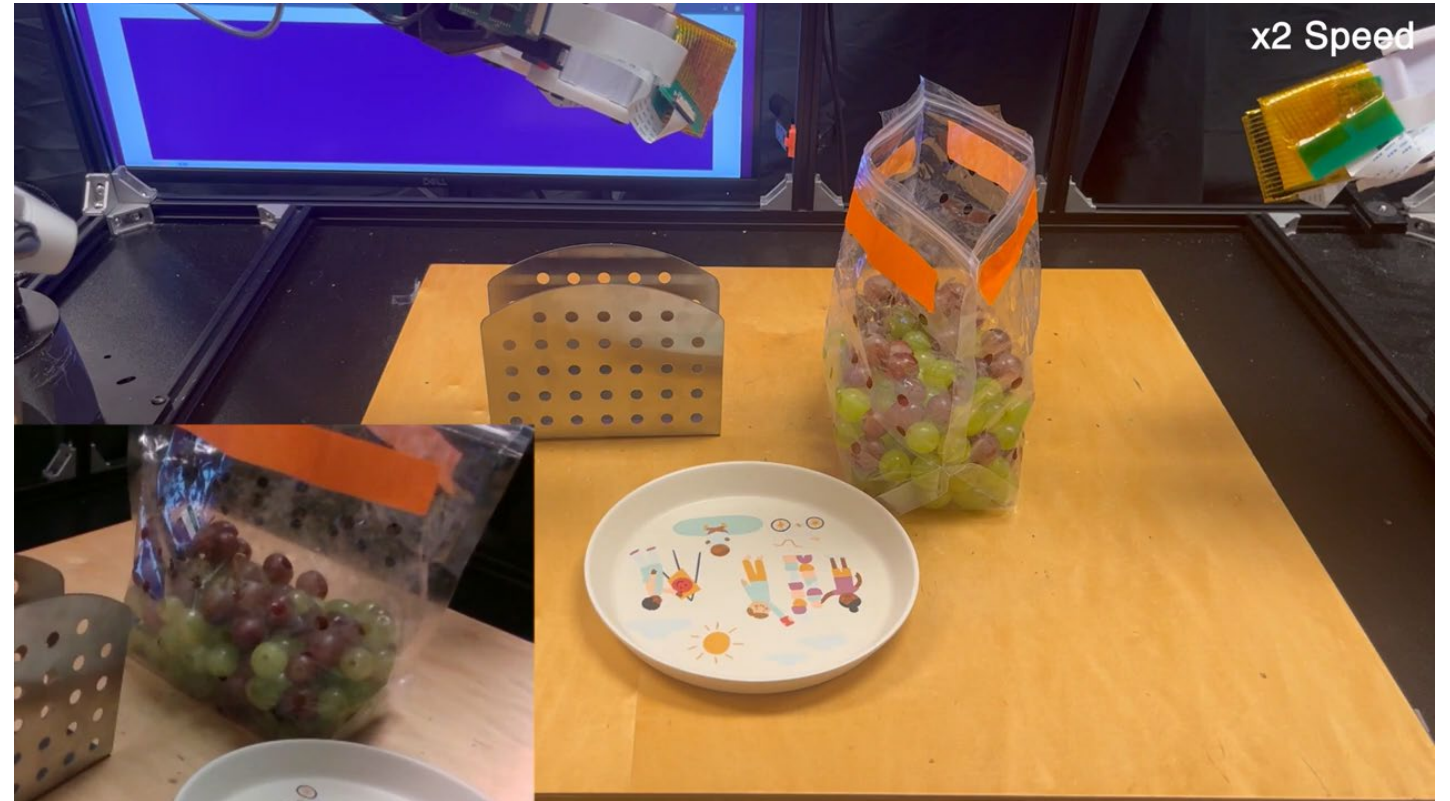
RESEARCH OVERVIEW

Project 1



3D-VITAC

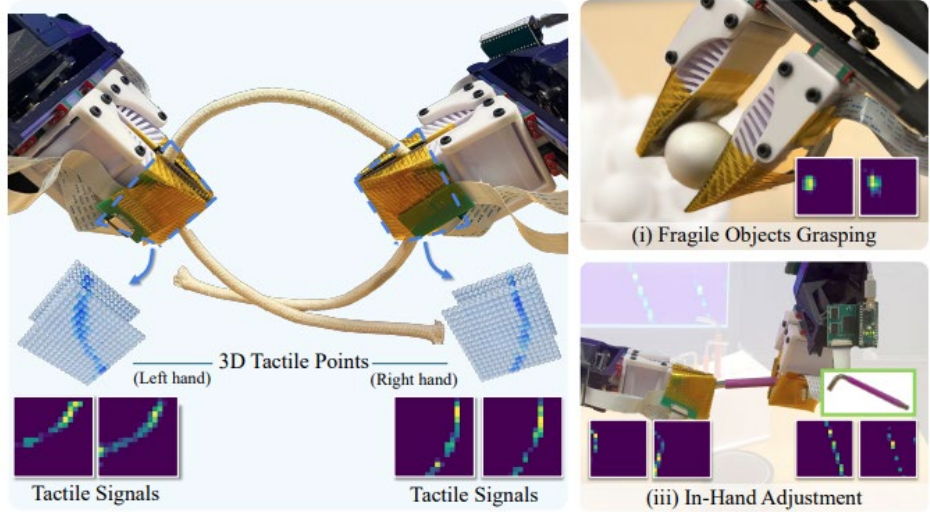
Learning with Real Visuo-Tactile Data



Still not Satisfied with Performance?

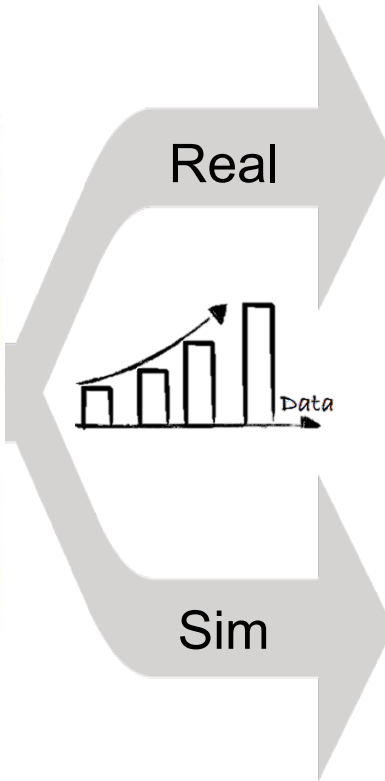
RESEARCH OVERVIEW

Project 1



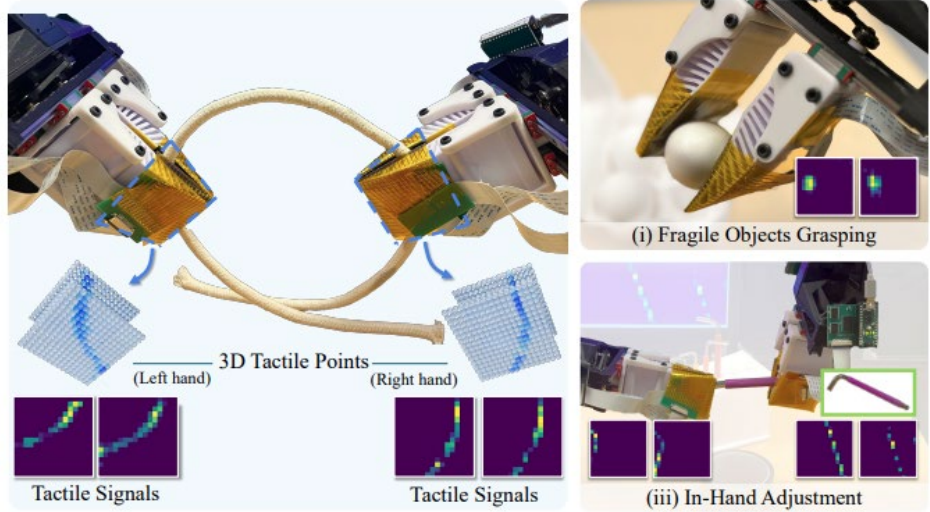
3D-ViTac

Learning with Real Visuo-Tactile Data



RESEARCH OVERVIEW

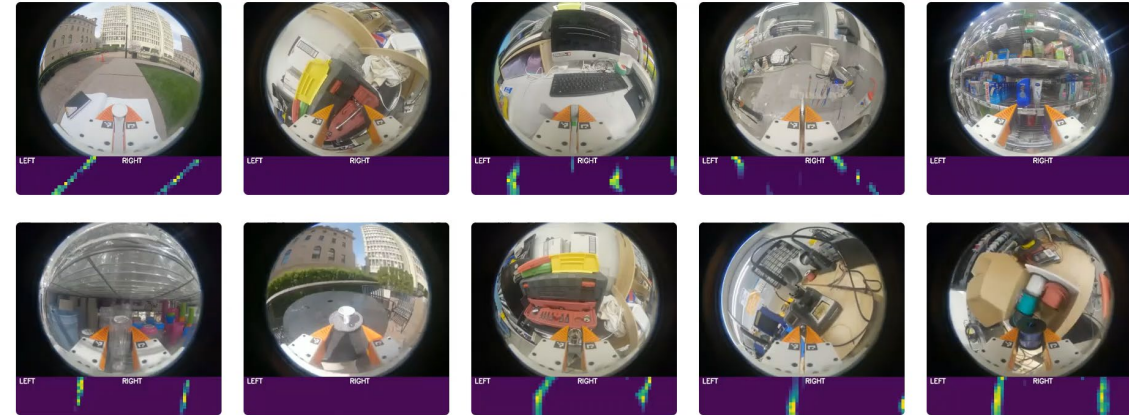
Project 1



3D-VITAC

Learning with Real Visuo-Tactile Data

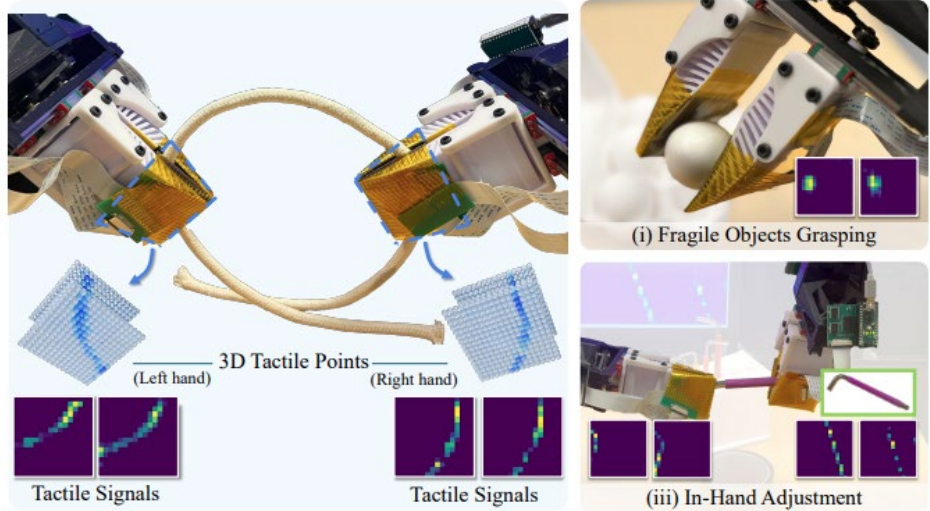
Project 2



Scaling **Large-Scale Real** Visuo-Tactile Data

RESEARCH OVERVIEW

Project 1



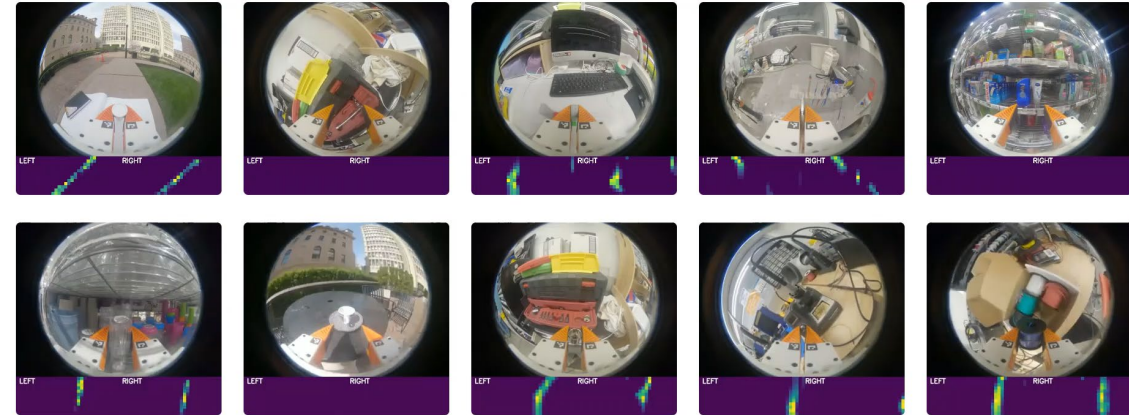
3D-VITAC

Learning with Real Visuo-Tactile Data

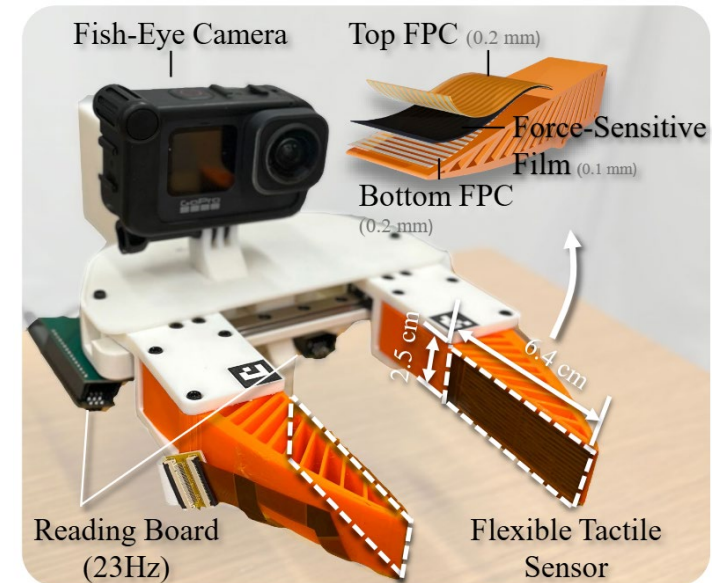
Real



Project 2

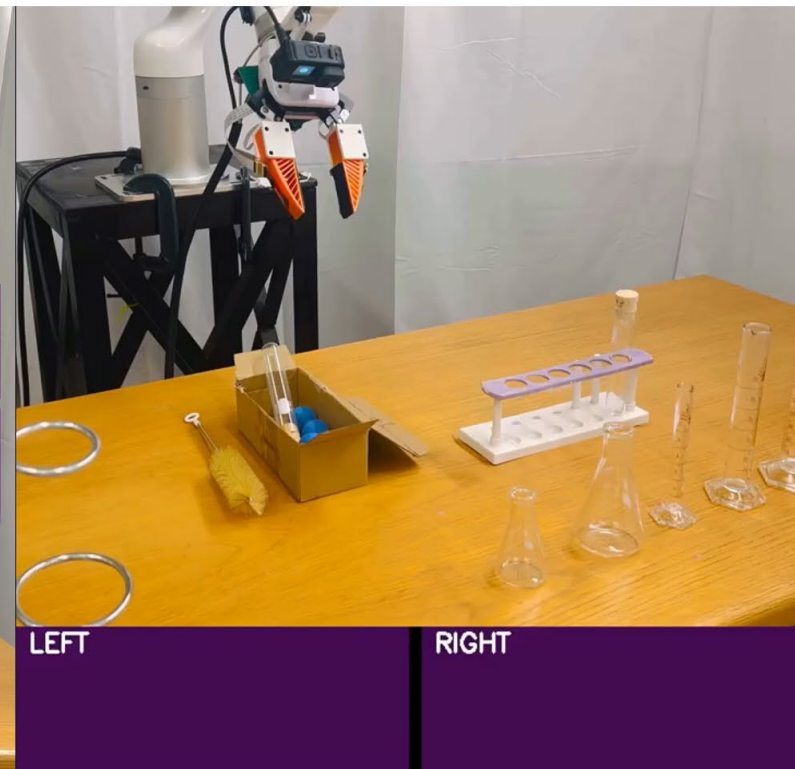
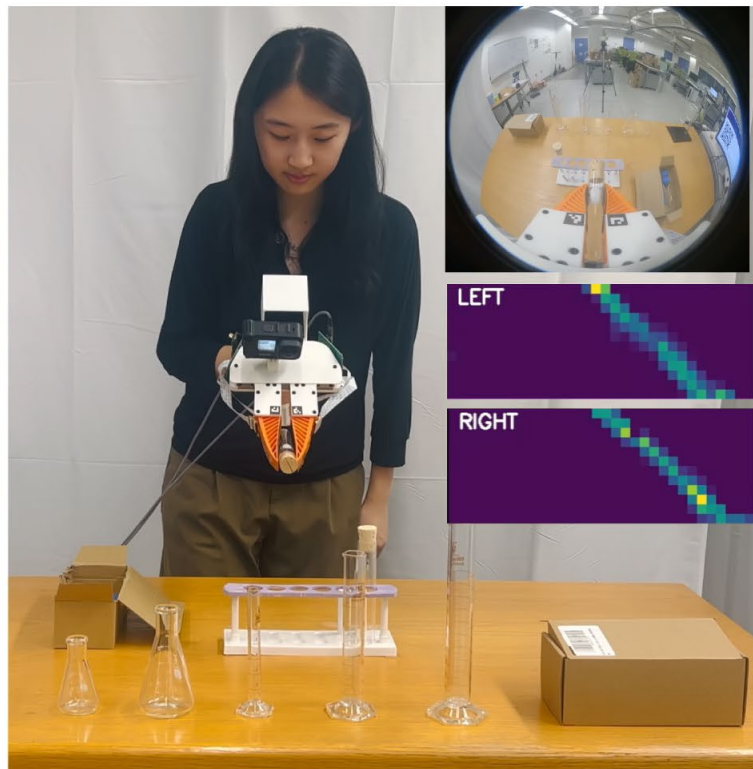


Scaling **Large-Scale Real** Visuo-Tactile Data



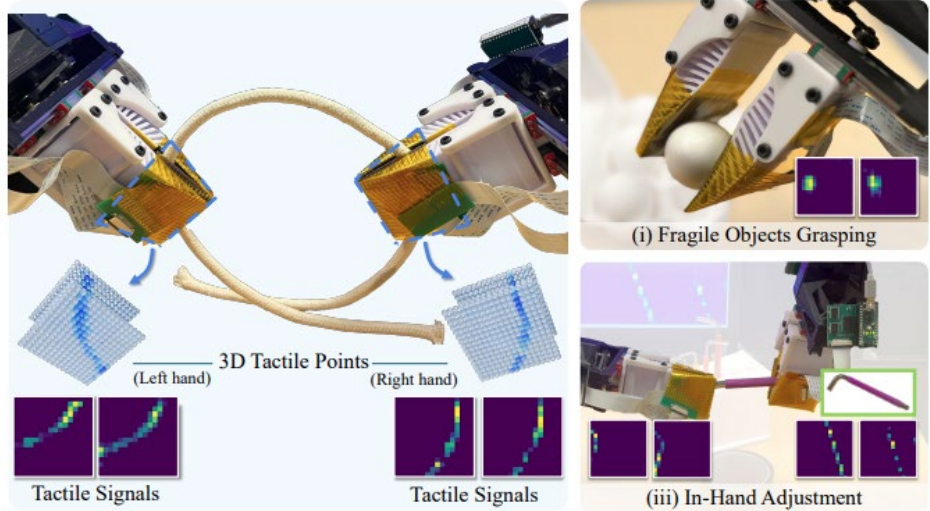
In-the-wild Tactile Data can help with fine-grained manipulation!

In-the-Wild Data Pre-Train → Downstream Data Fine-Tuning → Autonomous Policy Execution



RESEARCH OVERVIEW

Project 1



3D-ViTac

Learning with Real Visuo-Tactile Data

Real



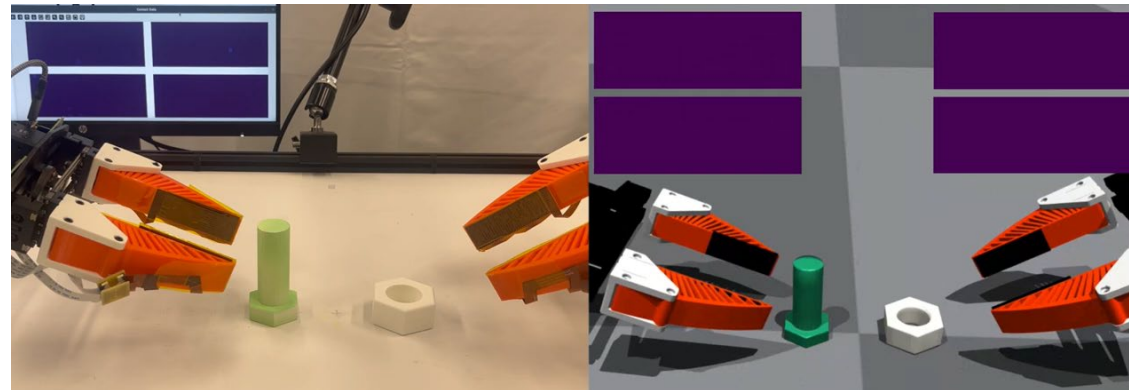
Sim

Project 2



Scaling Large-Scale Real Visuo-Tactile Data

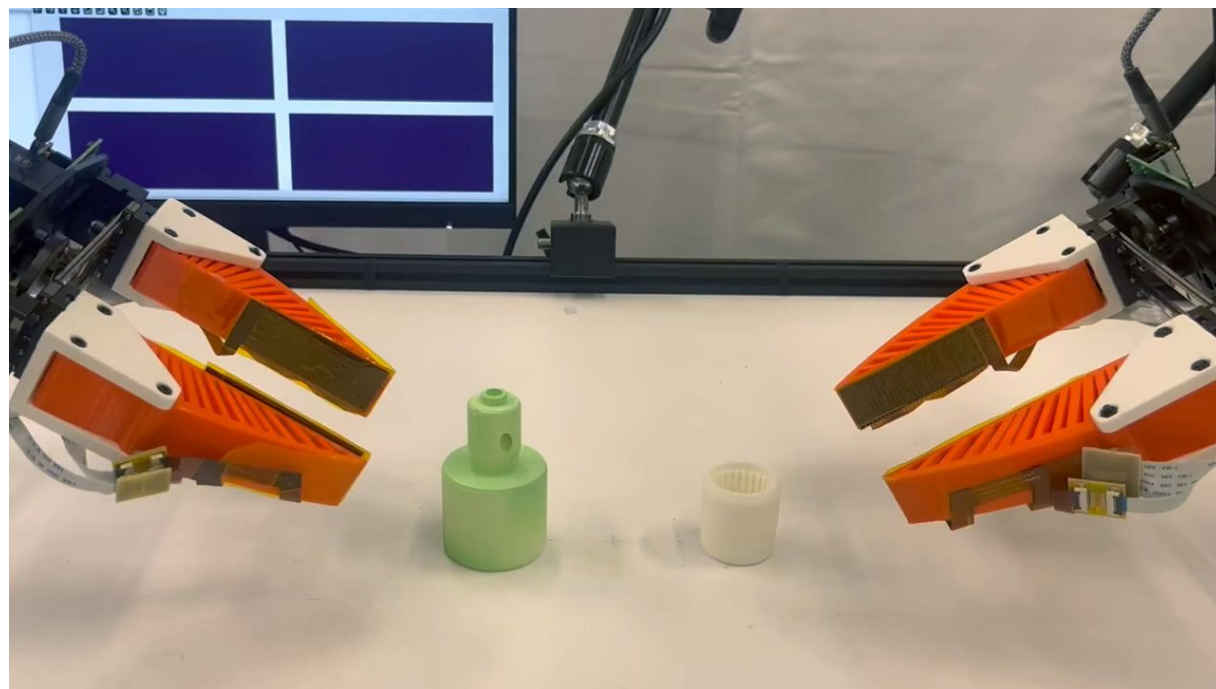
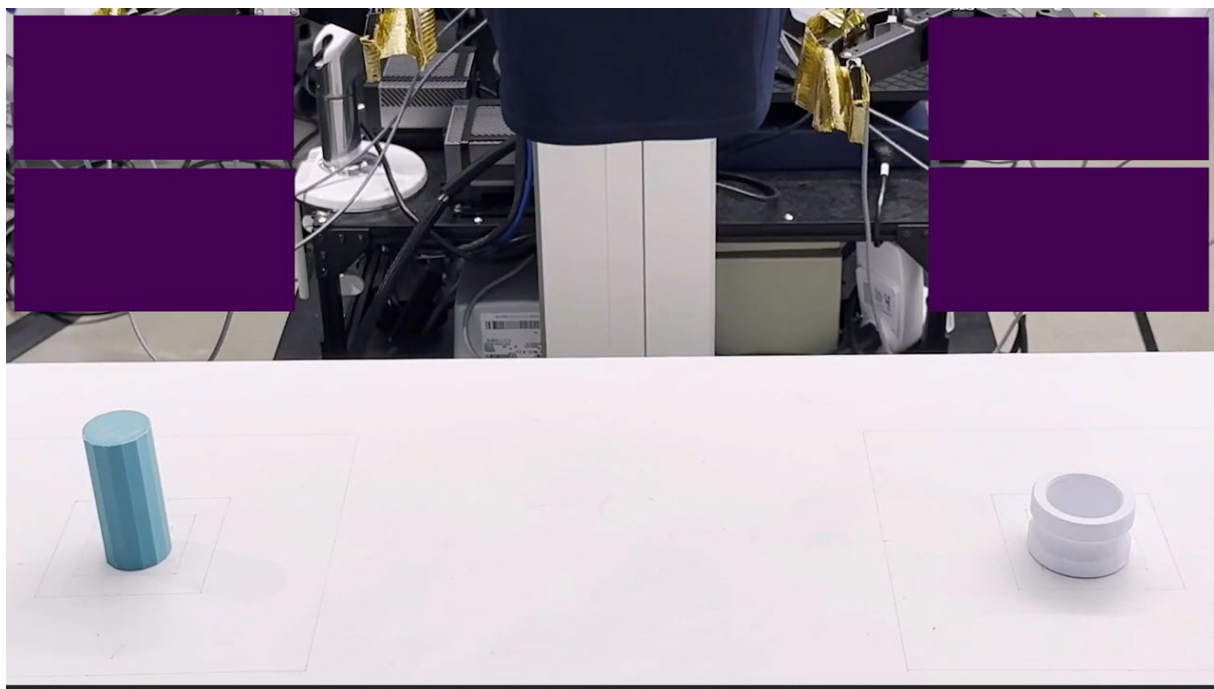
Project 3



Scaling Visuo-Tactile Data in **Simulation**

VT-Refine: Learning Bimanual Assembly with Visuo-Tactile Feedback via Simulation Fine-Tuning

Binghao Huang, Jie Xu, Iretiayo Akinola, Wei Yang, Balakumar Sundaralingam, Rowland O'Flaherty,
Dieter Fox, Xiaolong Wang, Arsalan Mousavian, Yu-Wei Chao*, Yunzhu Li*



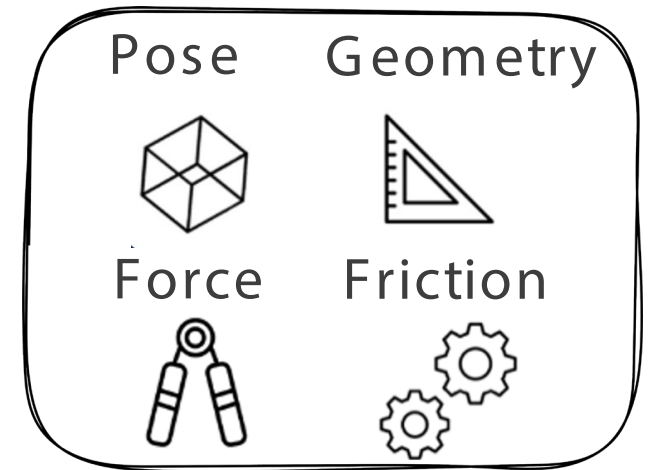
CHALLENGES



CHALLENGES



Action



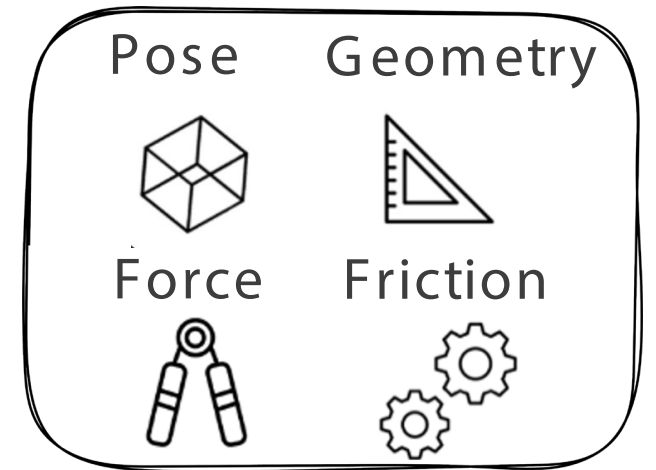
Task-relevant information



CHALLENGES



Action



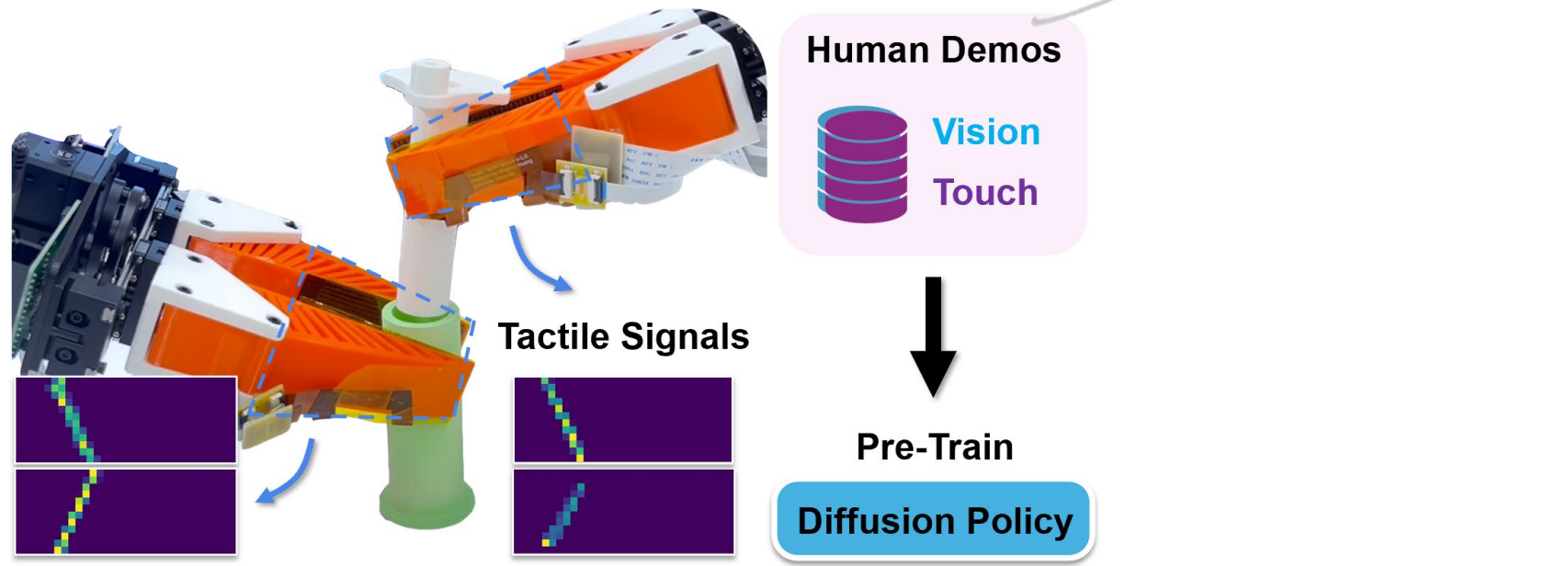
Task-relevant information



Can we enable robots
with those abilities?

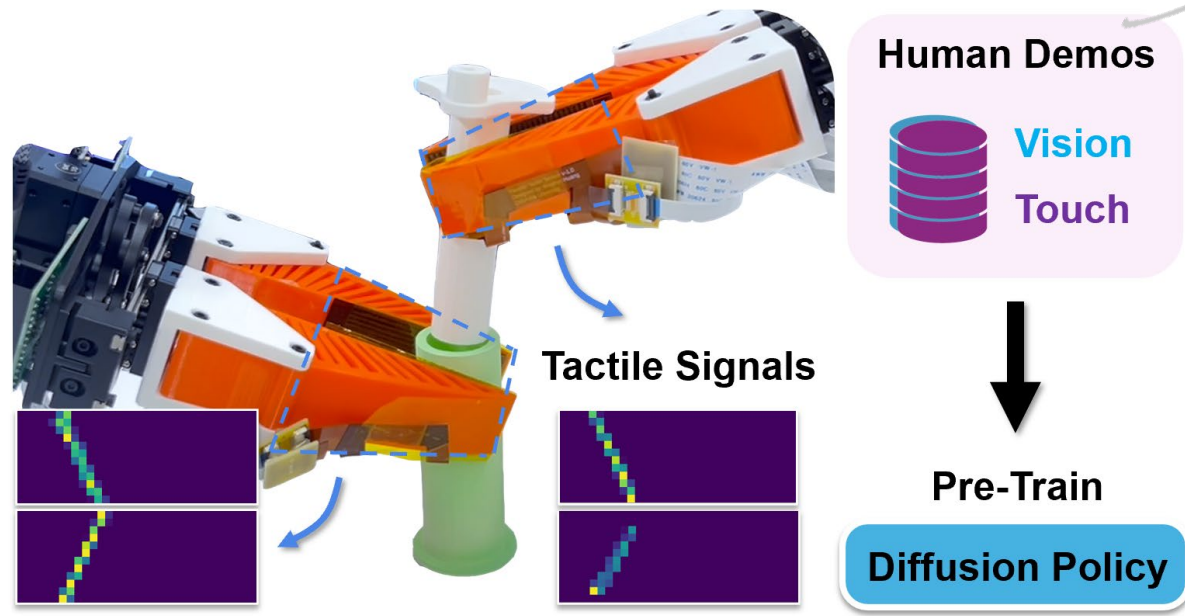


Train a Diffusion Policy

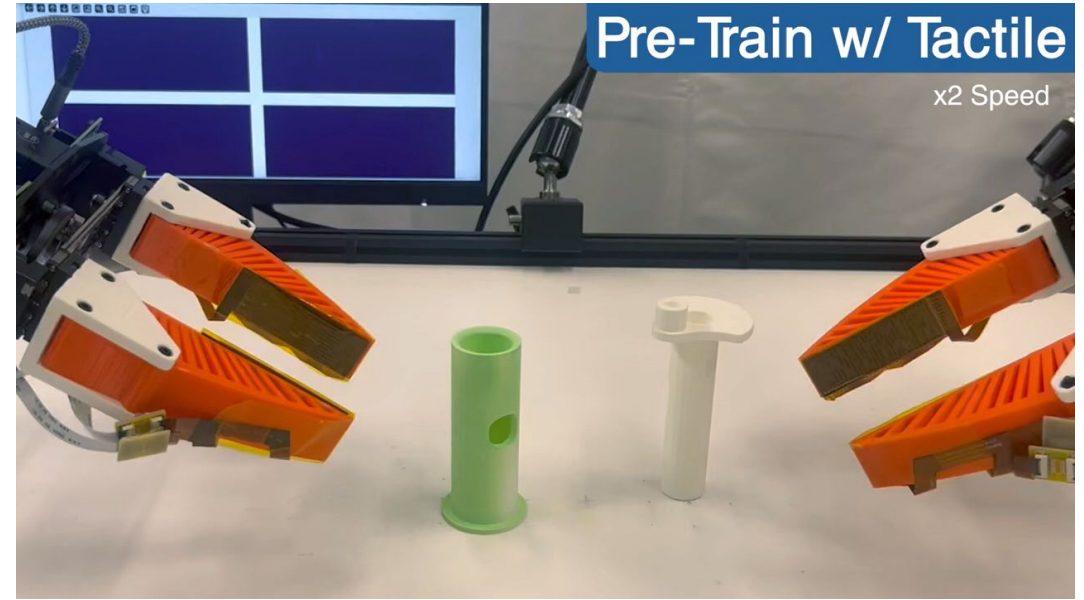


Train on 30 real data?

Train a Diffusion Policy

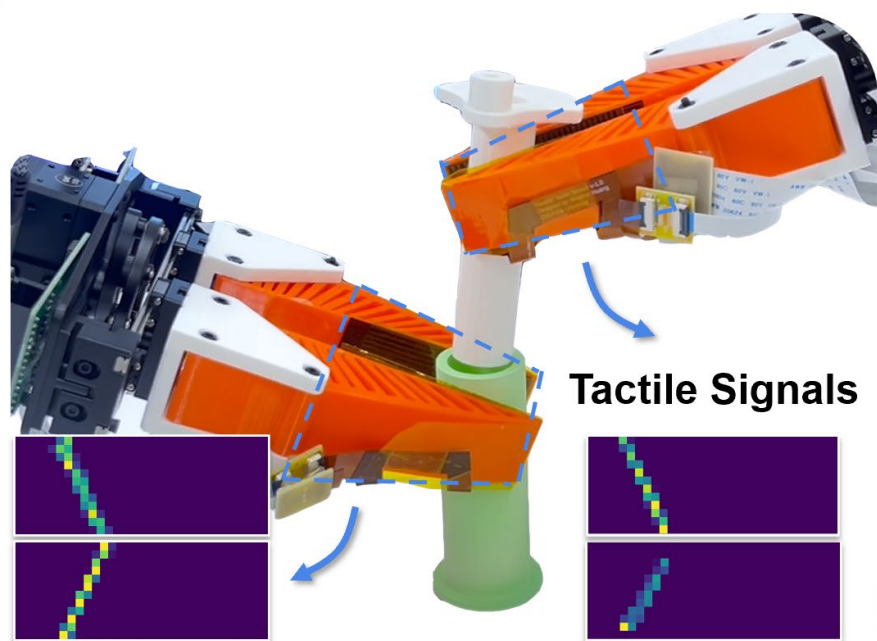


30 Demos



Train on 30 real data? **Not Satisfied!**

Real World



Human Demos



Vision

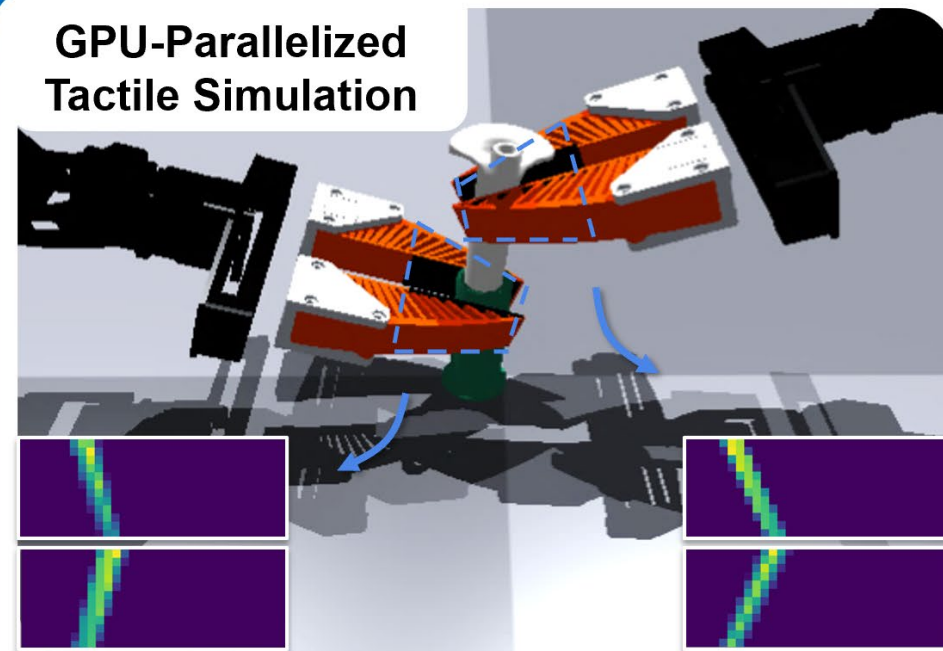
Touch

Pre-Train

Diffusion Policy

Simulation

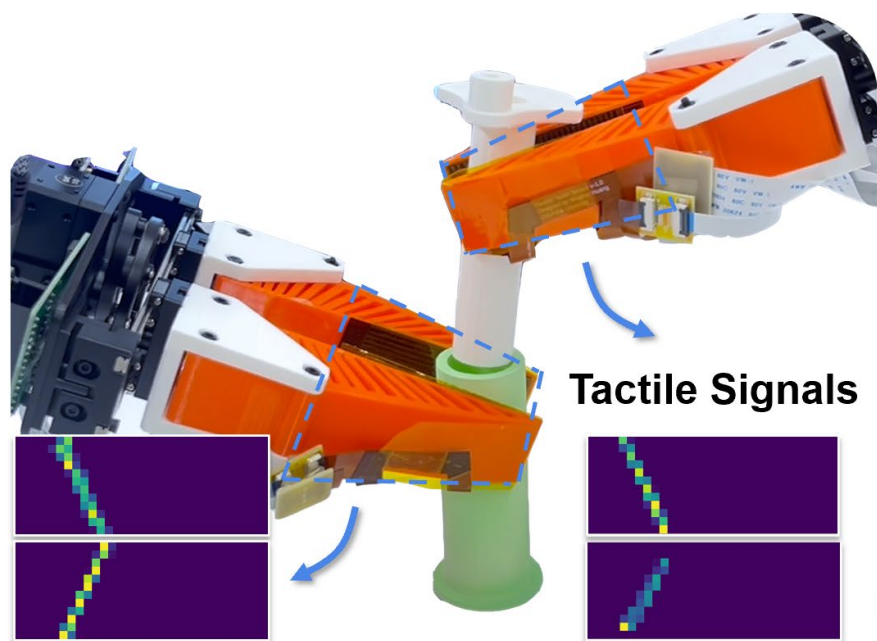
GPU-Parallelized
Tactile Simulation



Large-Scale RL Fine-Tuning



Real World



Human Demos

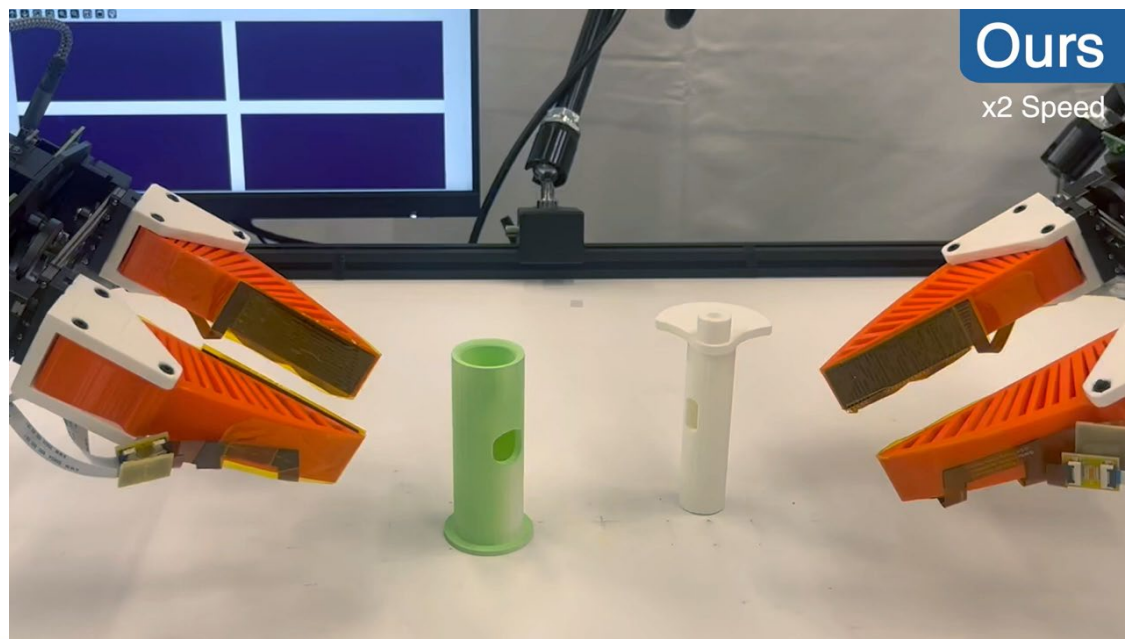


Vision

Touch

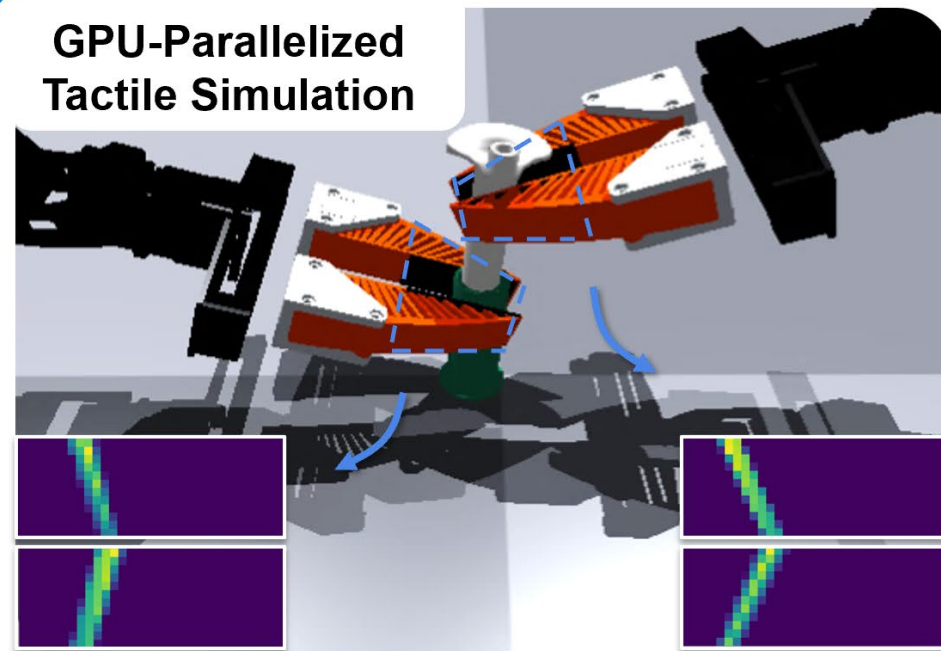
Pre-Train

Diffusion Policy

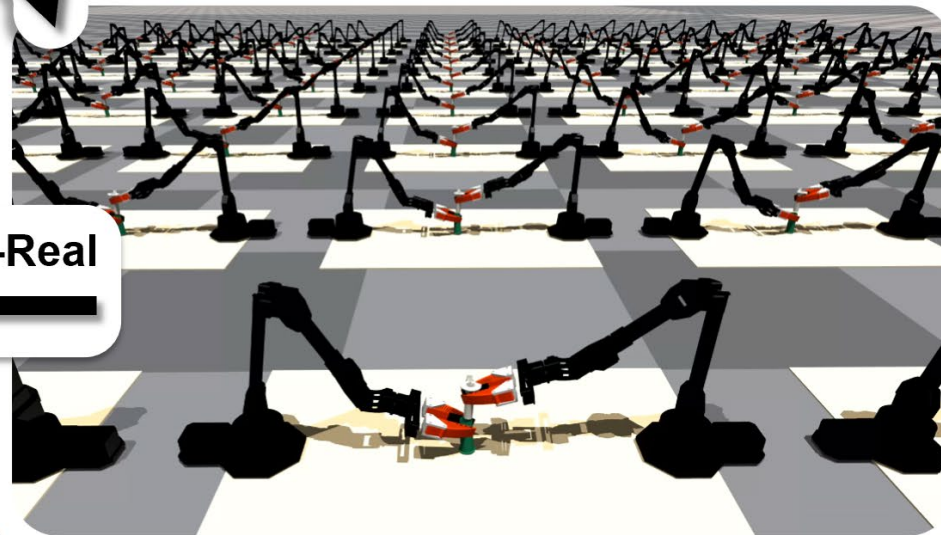


Simulation

GPU-Parallelized
Tactile Simulation

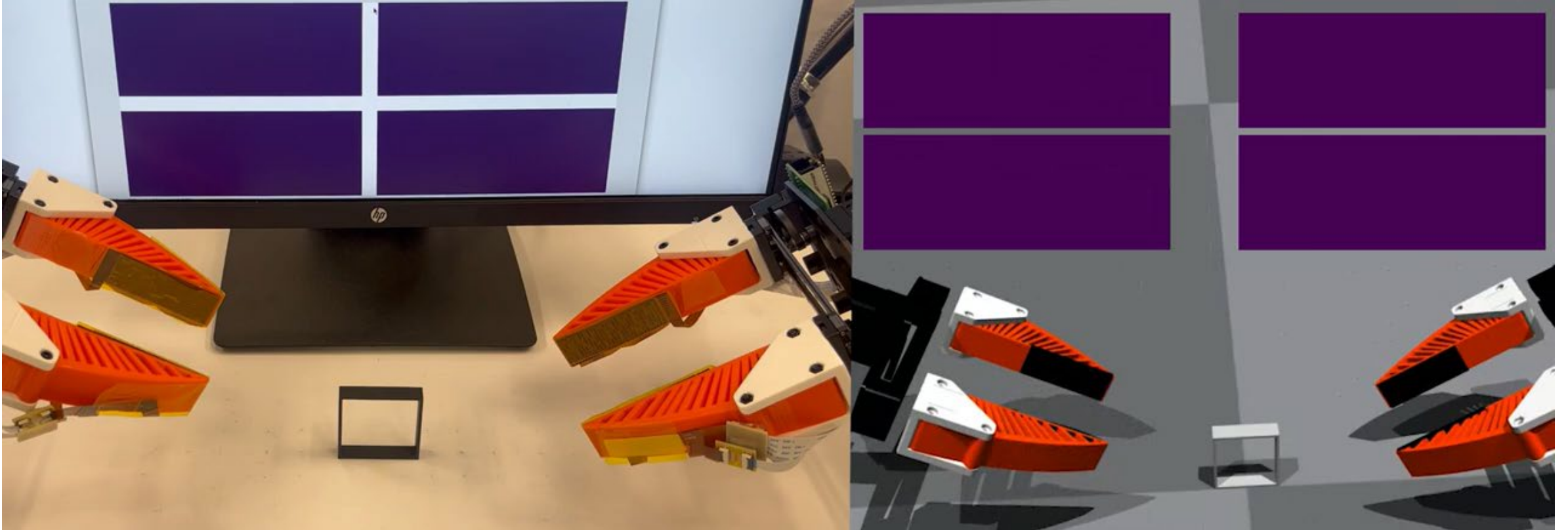


Large-Scale RL Fine-Tuning

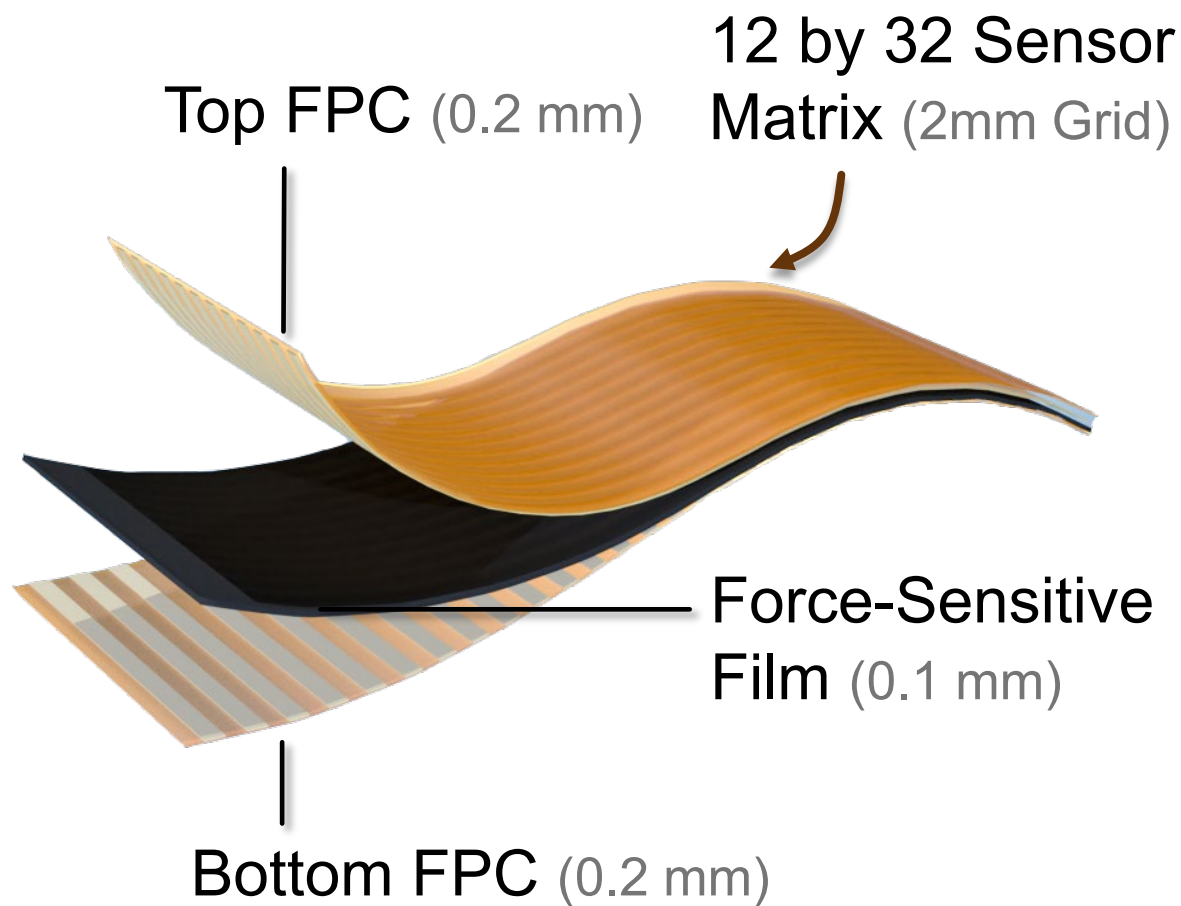


Sim-to-Real

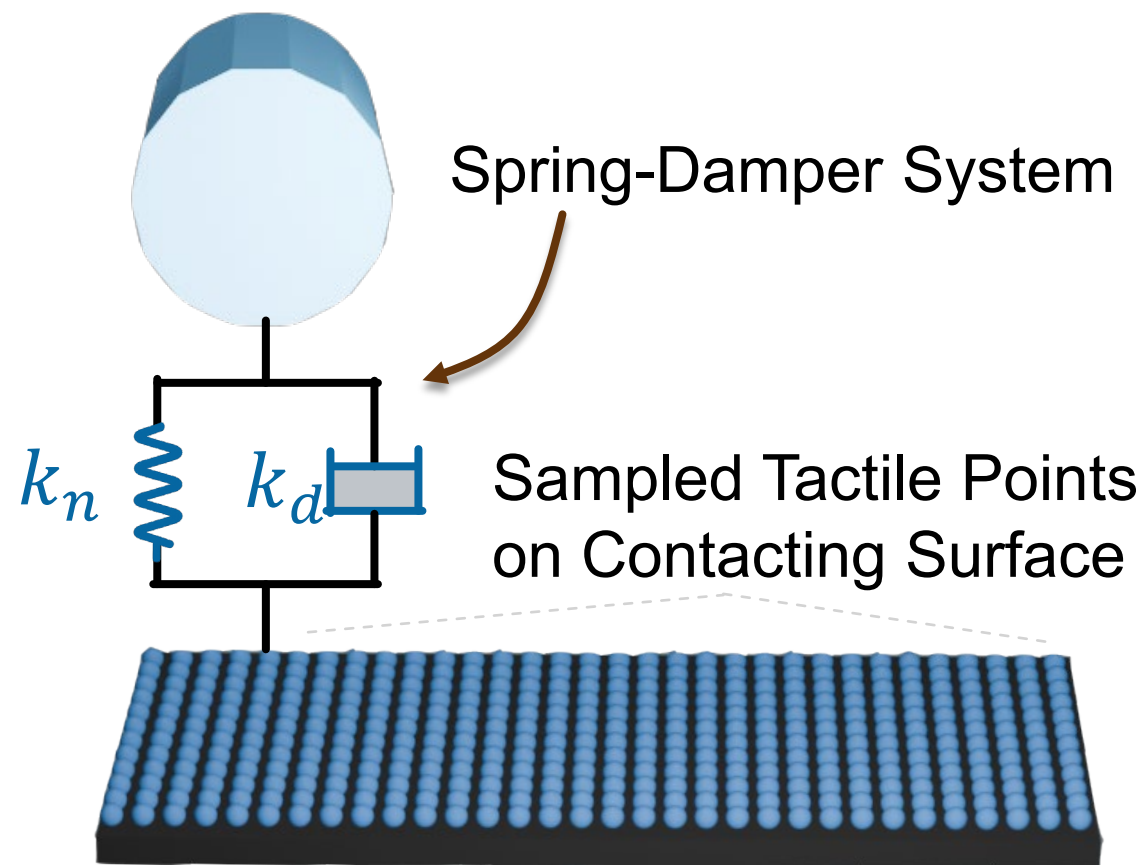
OUR TACTILE SIMULATION



OUR TACTILE SIMULATION

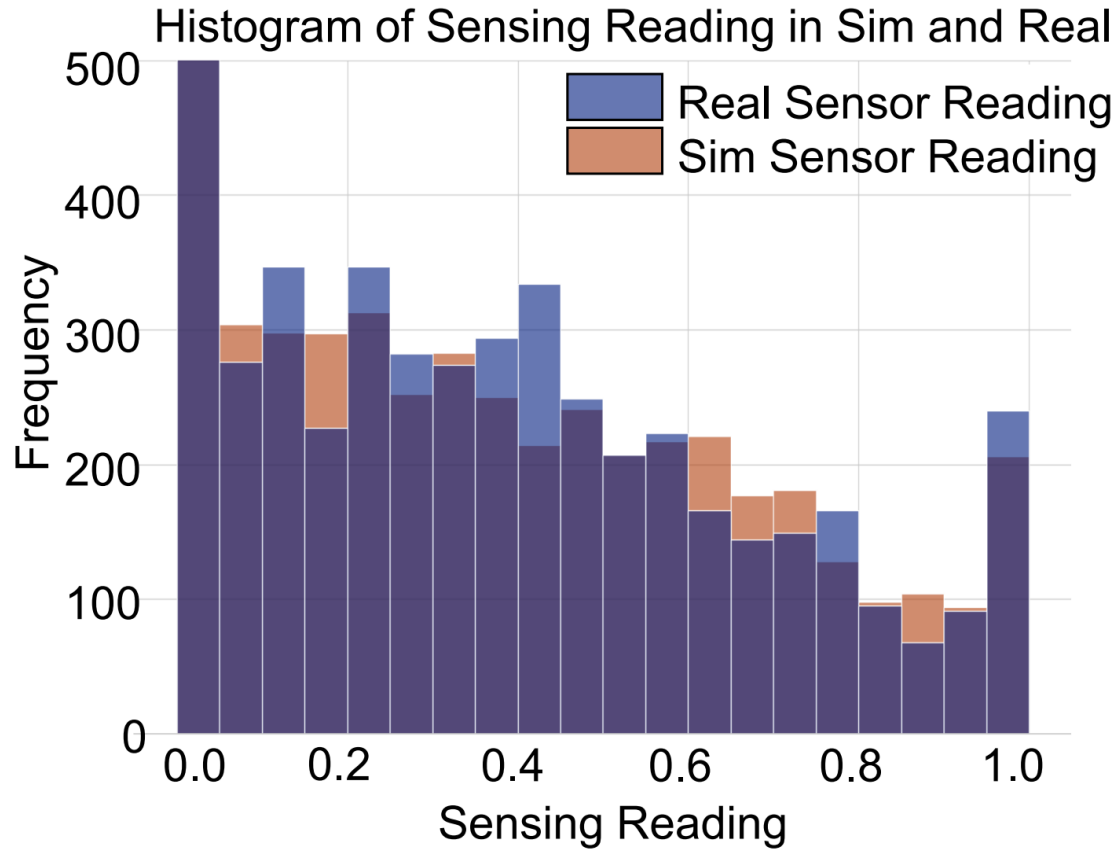


Real Tactile Sensor



Sim Tactile Sensor

OUR TACTILE SIMULATION

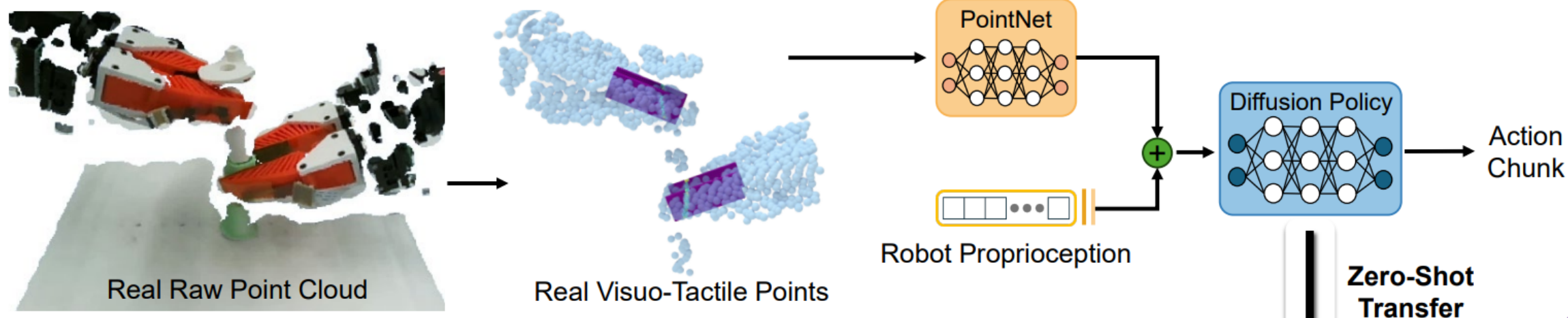


Key Takeaways:

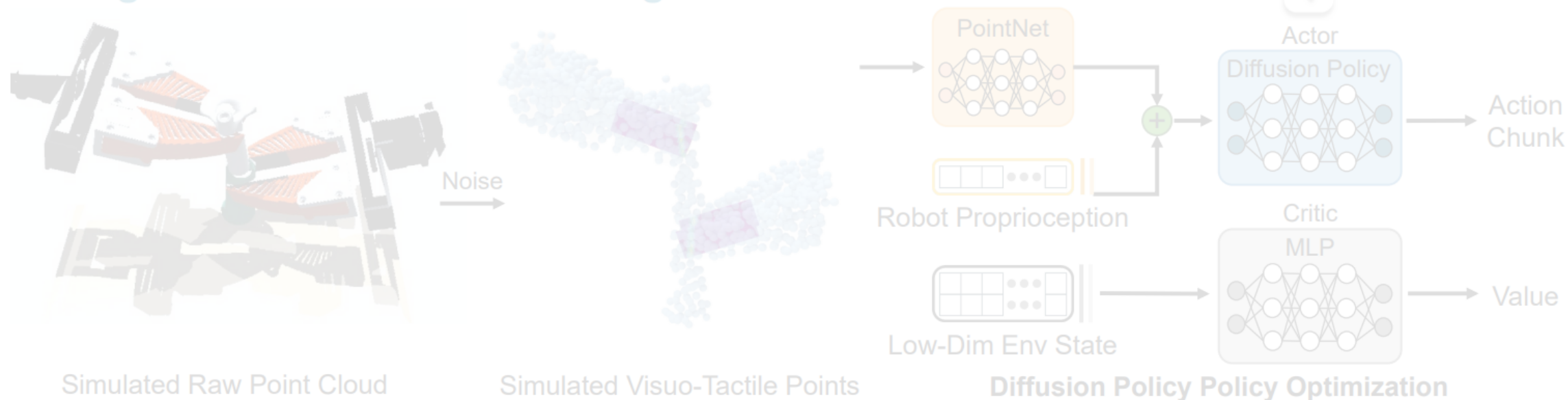
Normal force is easier to simulate!

Two Stage Policy Learning

Stage 1: Real World Pre-Training

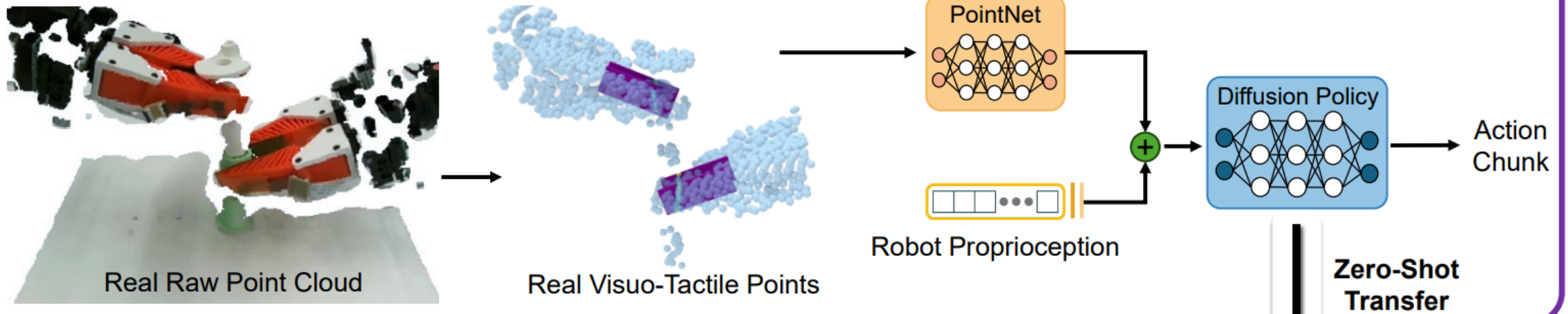


Stage 2: Simulation Fine-Tuning

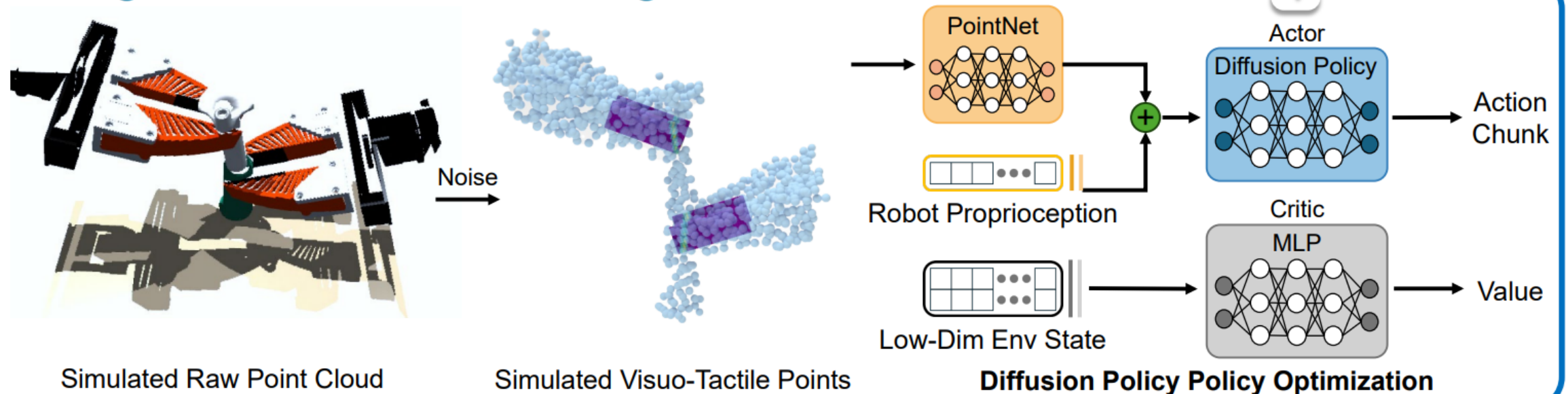


Two Stage Policy Learning

Stage 1: Real World Pre-Training



Stage 2: Simulation Fine-Tuning



Two Stage Policy Learning

Stage 1: Real World Pre-Training

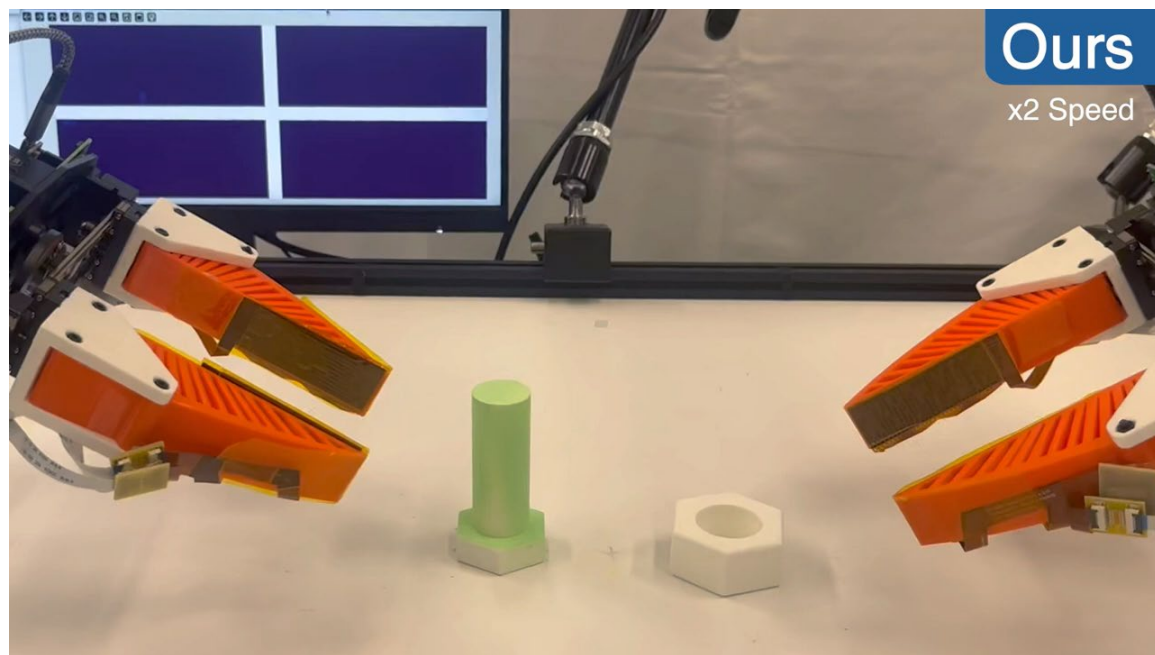


Stage 2: Simulation Fine-Tuning



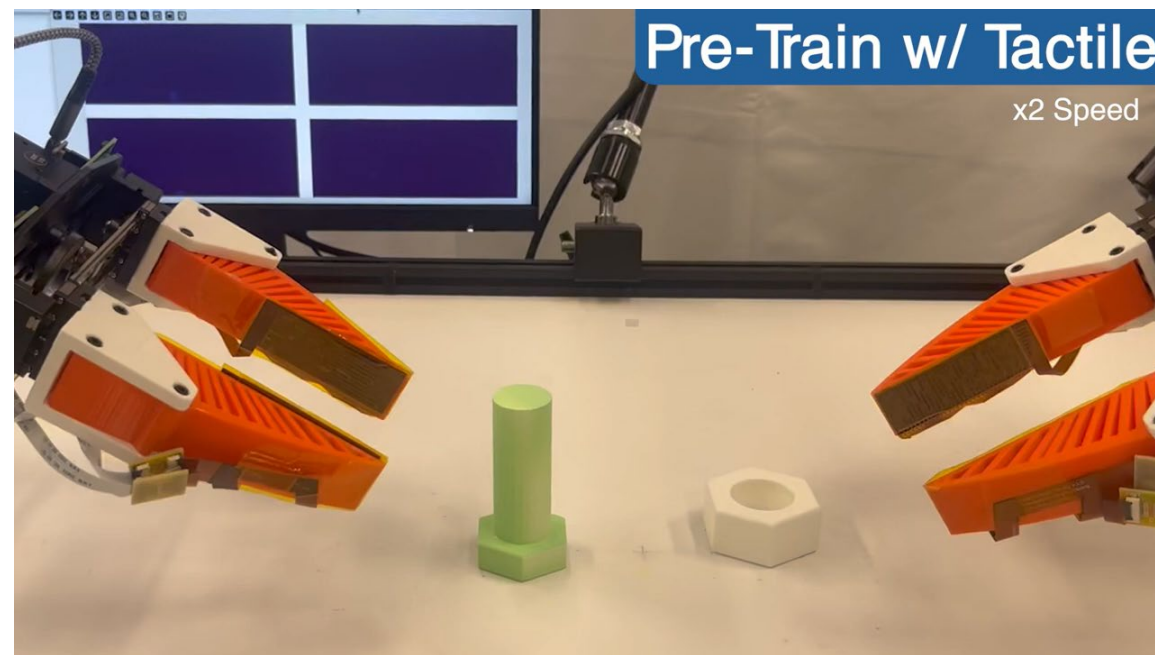
COMPARISONS: w/ and w/o Fine-Tune

Fine-Tune w/ Tactile (Ours)



Successful In-hand Adjustment

Pre-Train w/ Tactile

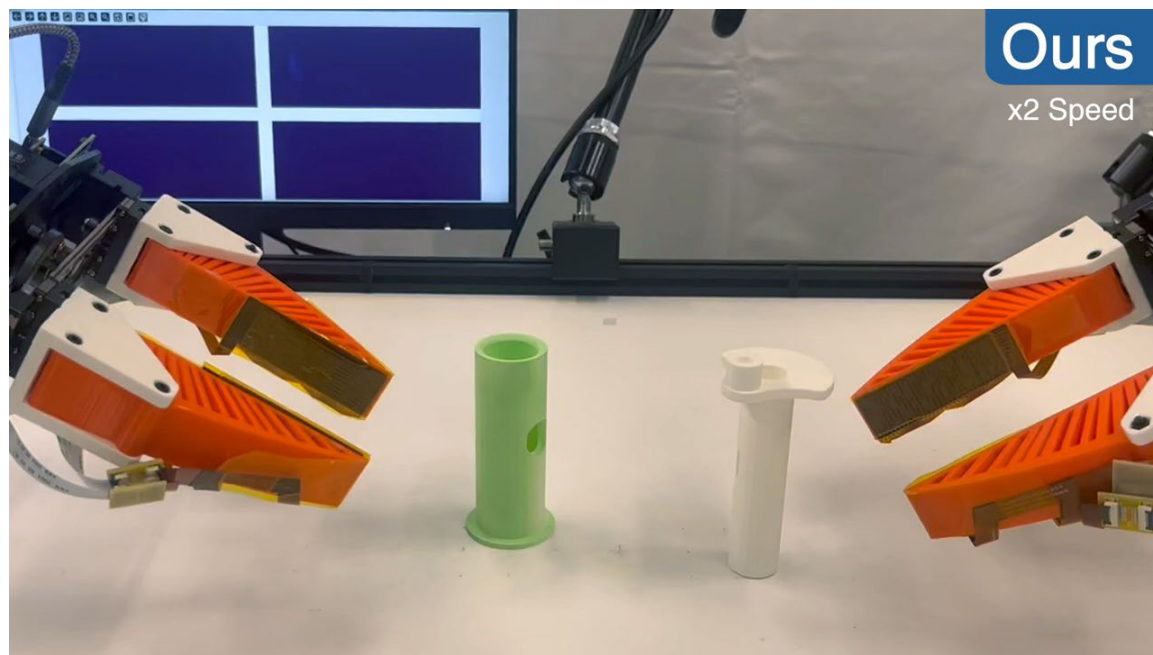


Failed to Re-Align the Socket

COMPARISONS

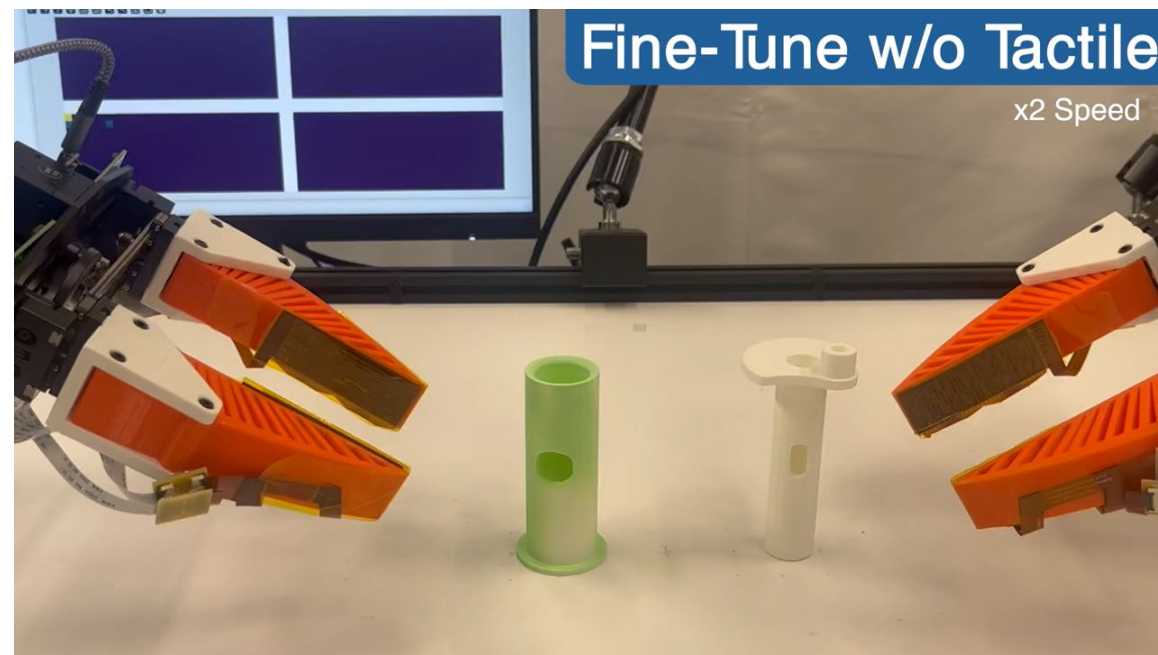
w/ and w/o Tactile

Fine-Tune w/ Tactile (Ours)



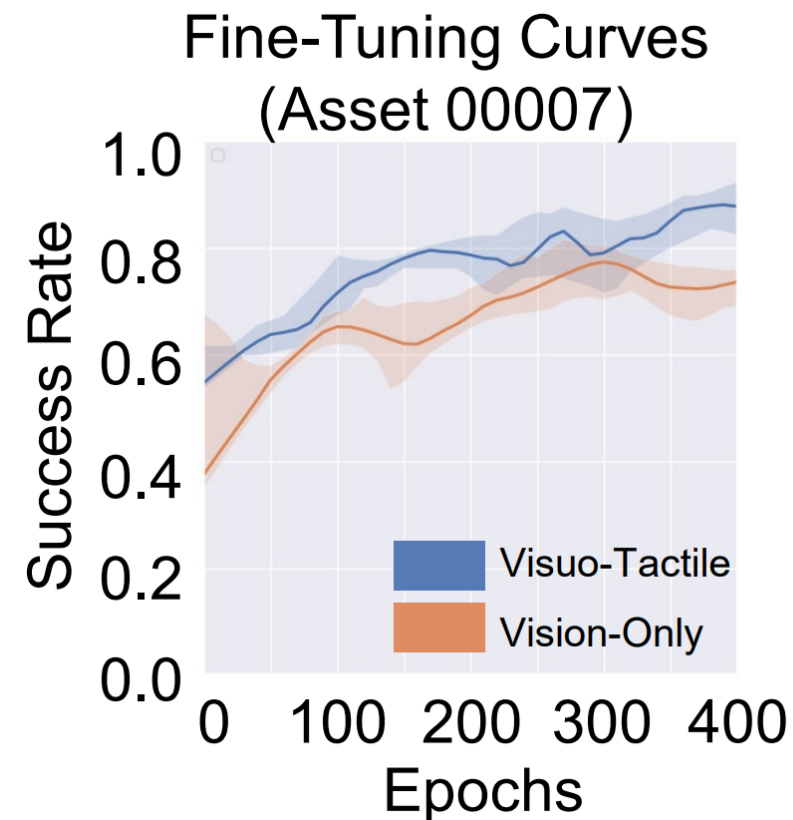
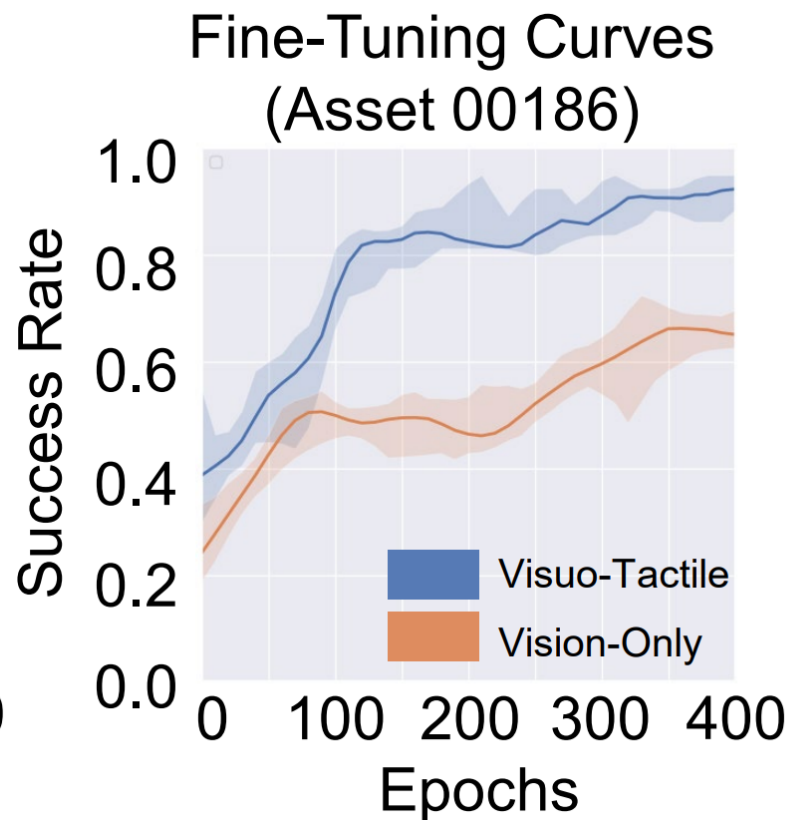
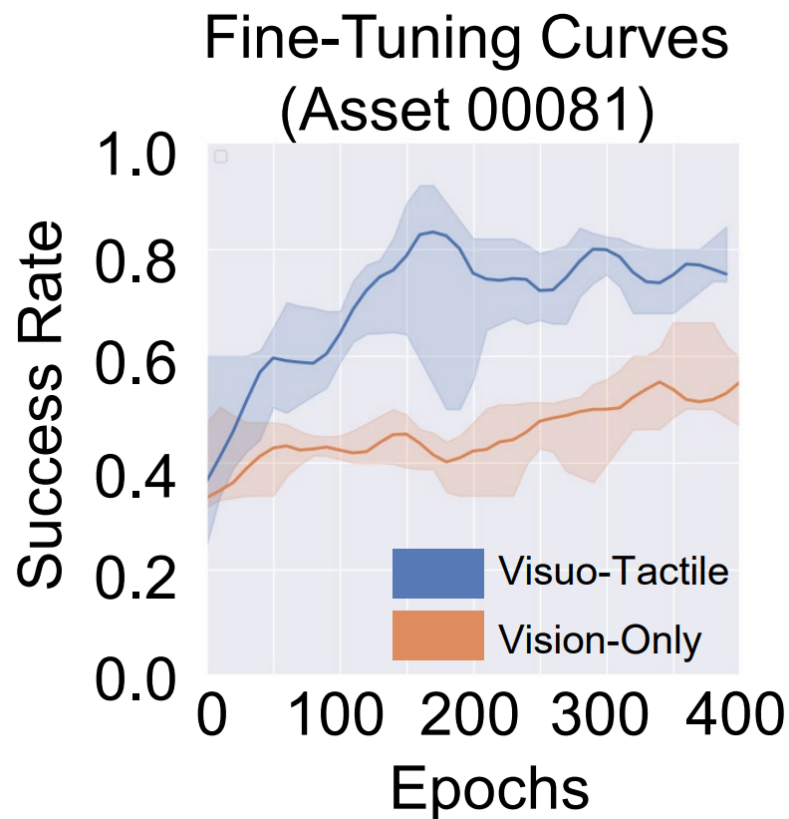
Accurate Insertion

Fine-Tune w/o Tactile

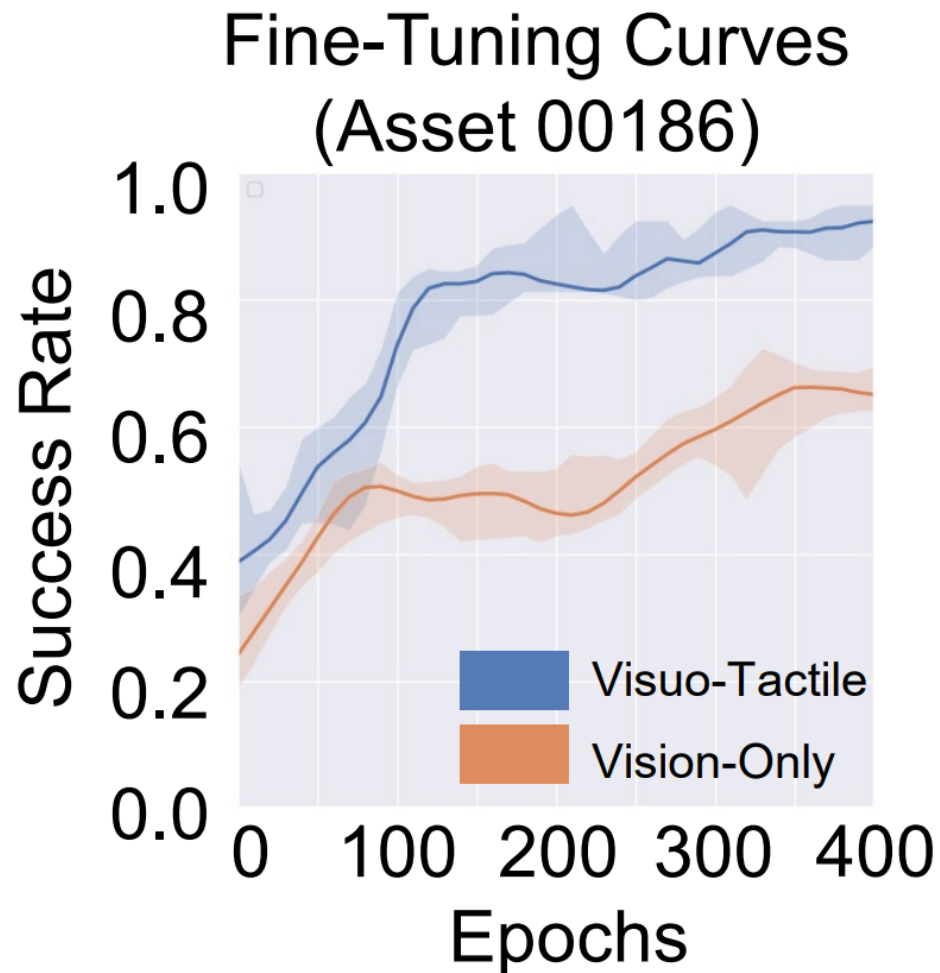


Failed Re-Adjustment Attempt

RESULTS: TRAINING CURVE



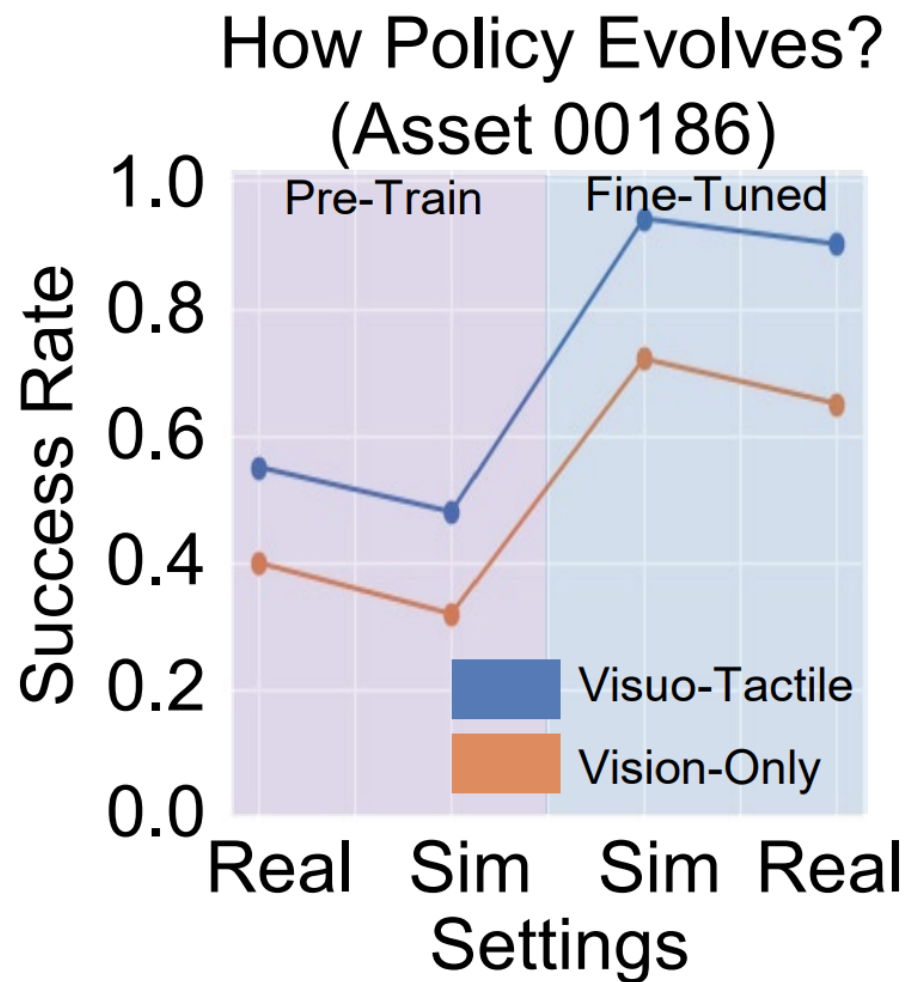
RESULTS



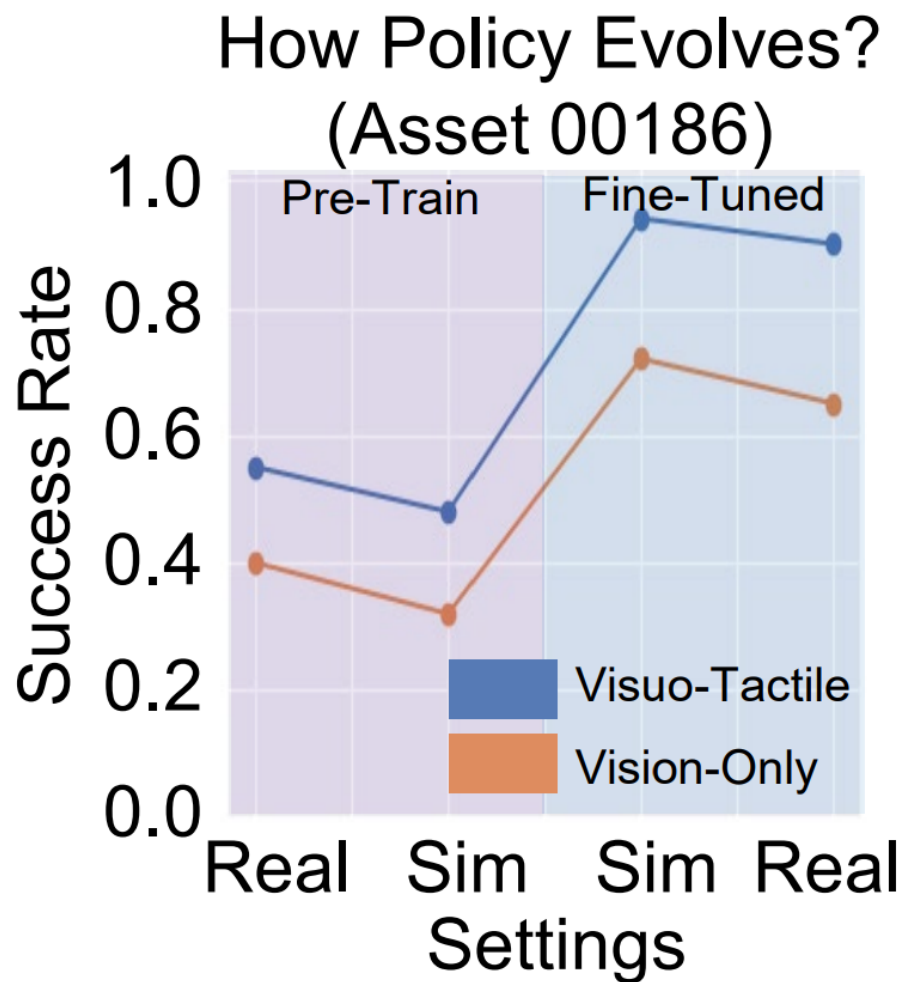
Key Takeaways:

Visuo-tactile input leads to more reliable and robust manipulation!

RESULTS



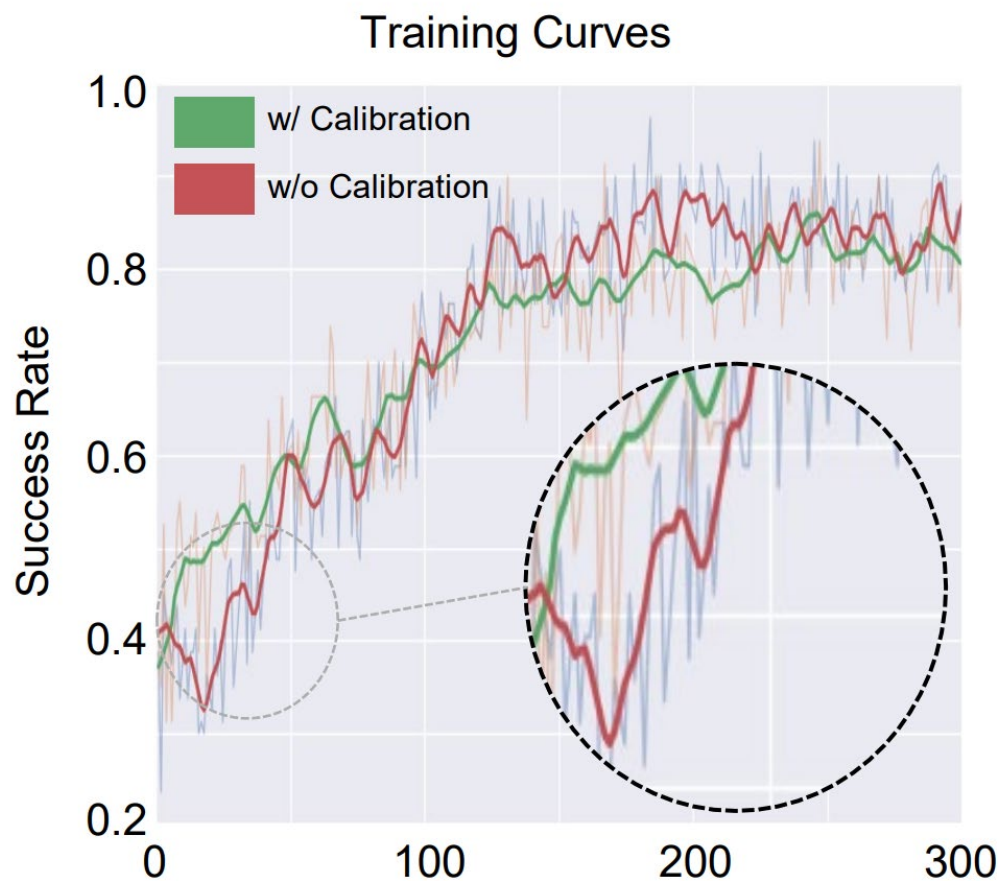
RESULTS



Key Takeaways:

The Sim-to-Real gap still exists —
but fine-tuning helps *even more!*

RESULTS



Key Takeaways:

Vision, tactile, and control each play a role in the Sim2Real gap!

w/ and w/o Tactile Calibration

We hope to democratize our sensor to robot learning community!

Reproduction Resources



Hardware
Codebase

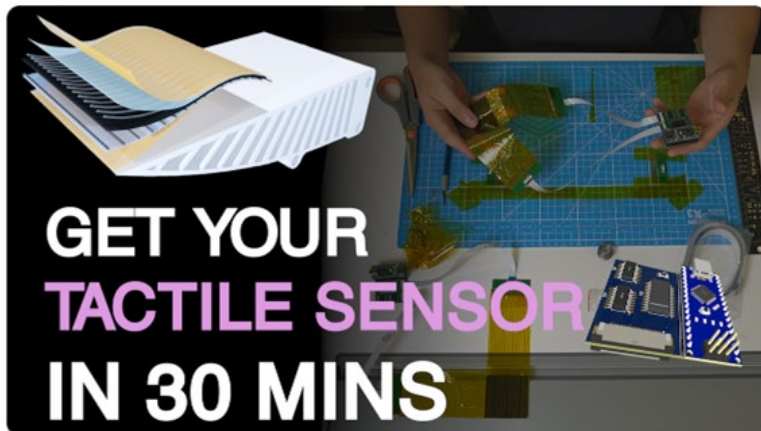


Hardware
Guide

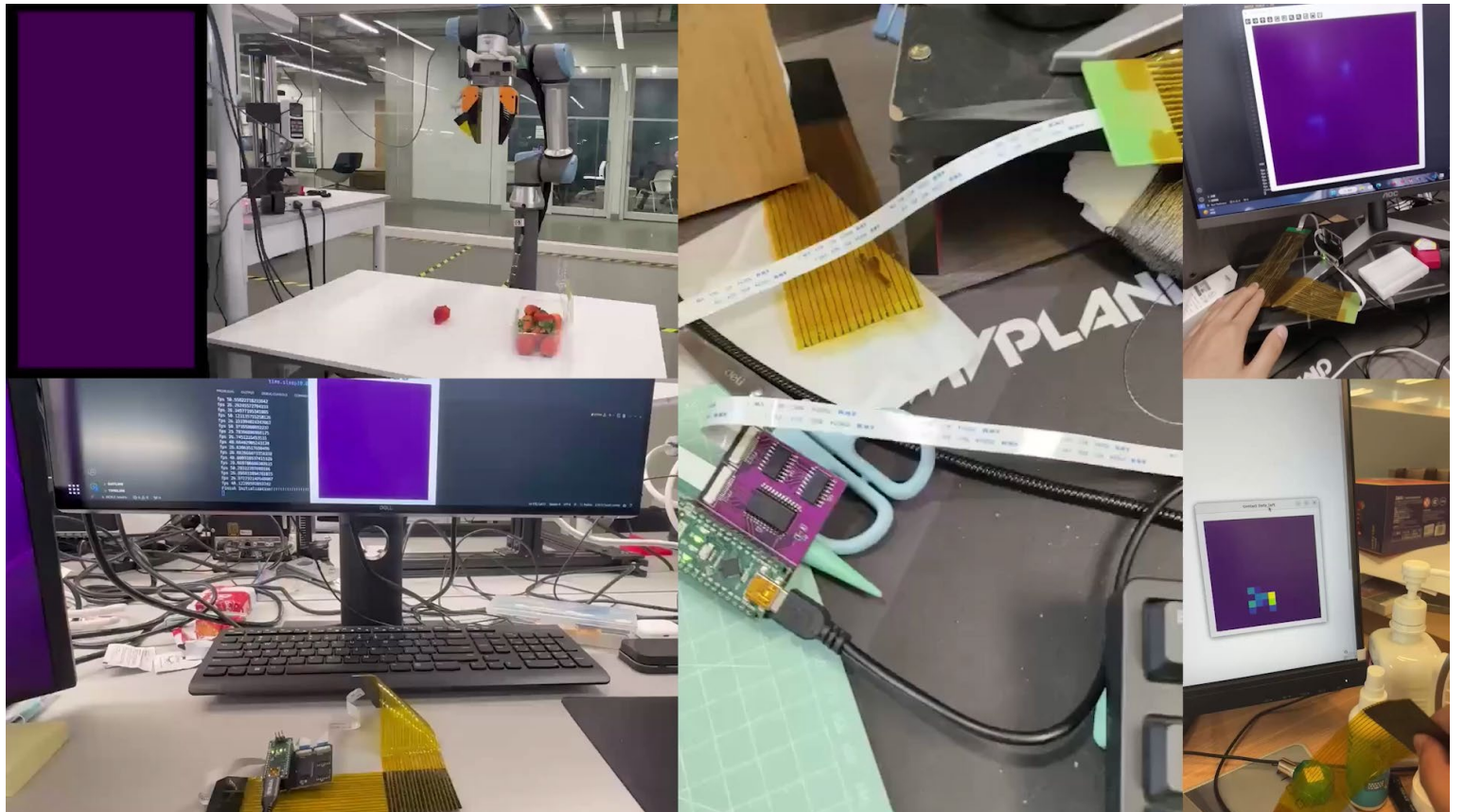


Bill of
Materials

Sensor Fabrication and PCB Tutorial



Videos back from both academia and industry



We hope to democratize our sensor to robot learning community!

Reproduction Resources

Videos back from both academia and industry



Hardware
Codebase



Hardware
Guide



Bill of
Materials

Sensor Fabrication and PCB Tutorial



GET YOUR
TACTILE SENSOR
IN 30 MINS

**We are continuously
improving our sensor!**



FlexiTac-V2 (Lead time: ~2 mins)

(FPCB-Based, 2 mm Grid, 32×12 Resolution, \$2.5 each sensor piece)

Request a Sample

