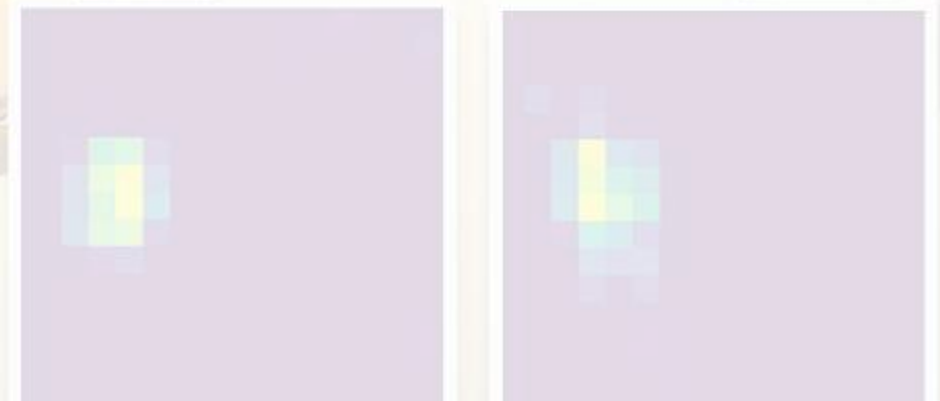




Touch: The Missing Sense for Physical Intelligence

Binghao Huang @ Columbia



I build robots that work in the physical world



Binghao Huang

Final-year Ph.D. @ Columbia CS · advised by Yunzhu Li

M.S. @ UC San Diego MechE · advised by Xiaolong Wang

NVIDIA Robotics

Amazon FAR

THE SENSE

**Flexible
Tactile Skin**



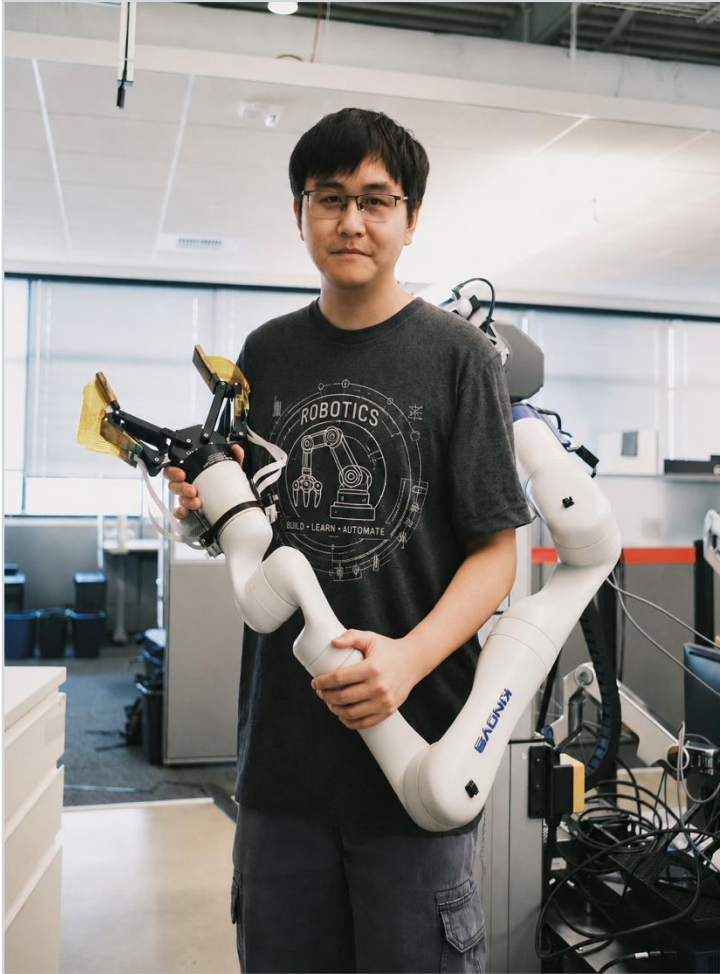
THE BRAIN

**Visuo-Tactile
Perception**



THE SCALE

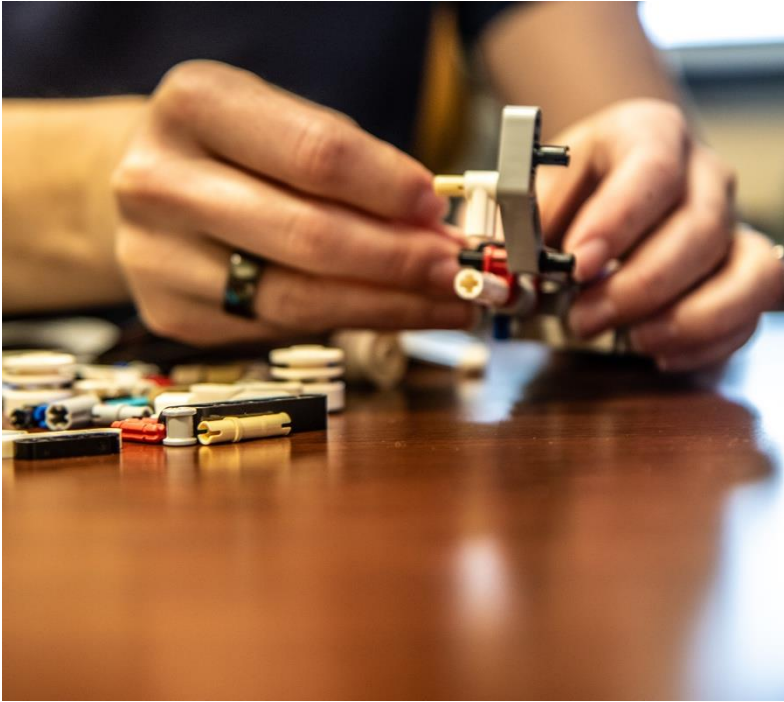
Data: Real + Sim



It starts with vision and touch — so I build the full stack around it

Touch is helpful in our human's daily life

Assemble Lego



Fine alignment

Pick Strawberries

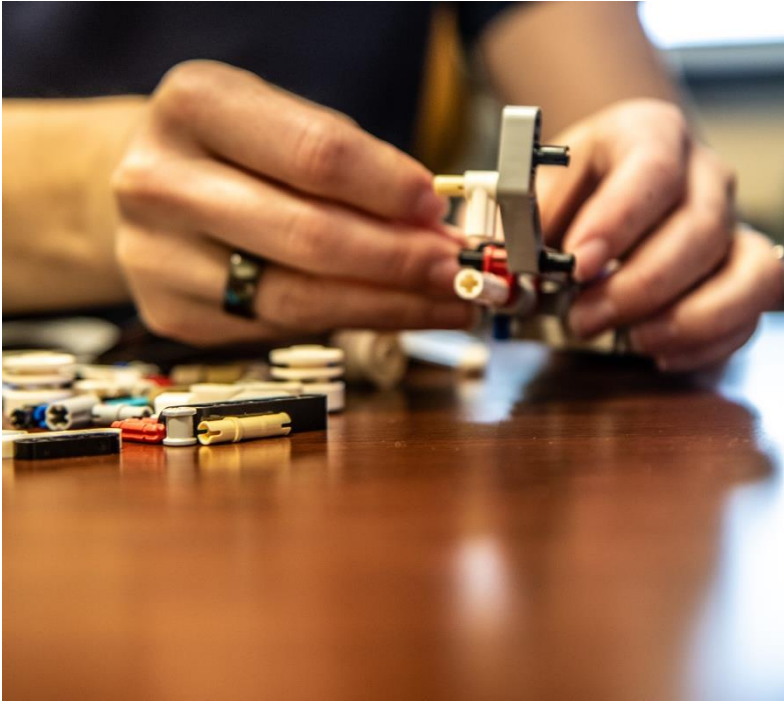


Use Lab Dropper



Touch is helpful in our human's daily life

Assemble Lego



Fine alignment

Pick Strawberries



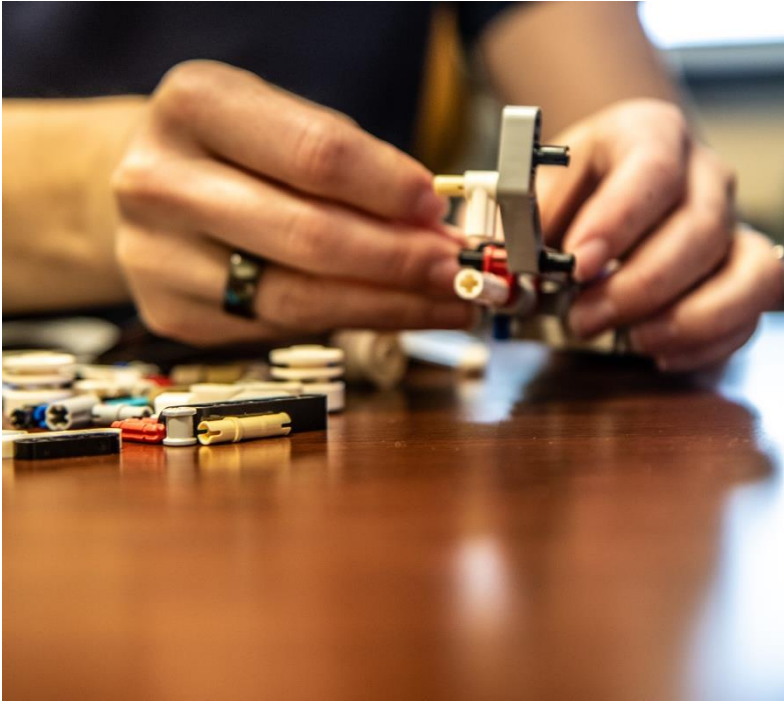
Handling fragile objects

Use Lab Dropper



Touch is helpful in our human's daily life

Assemble Lego



Fine alignment

Pick Strawberries



Handling fragile objects

Use Lab Dropper



Precise force control

Overview

One full stack for scalable tactile manipulation
— build it, perceive with it, learn at scale.



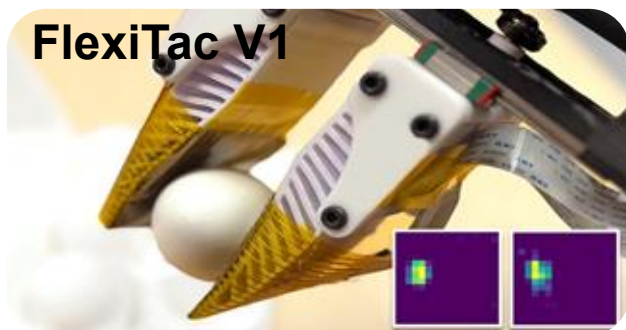
Hardware



Perception



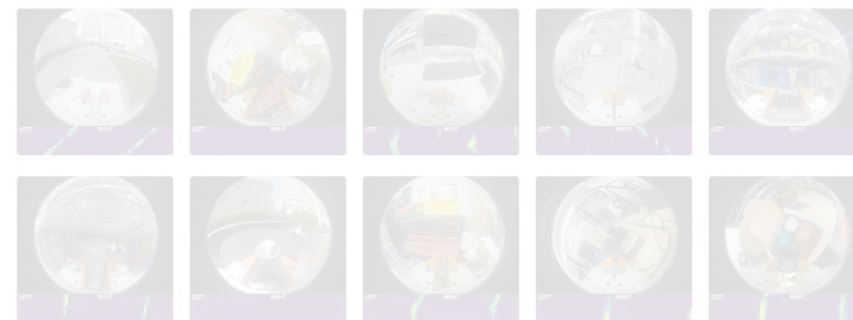
Data



FlexiTac V1



FlexiTac V2



Large-Scale Tactile in Real



Large-Scale Tactile in Simulation

Overview

One full stack for scalable tactile manipulation
— build it, perceive with it, learn at scale.



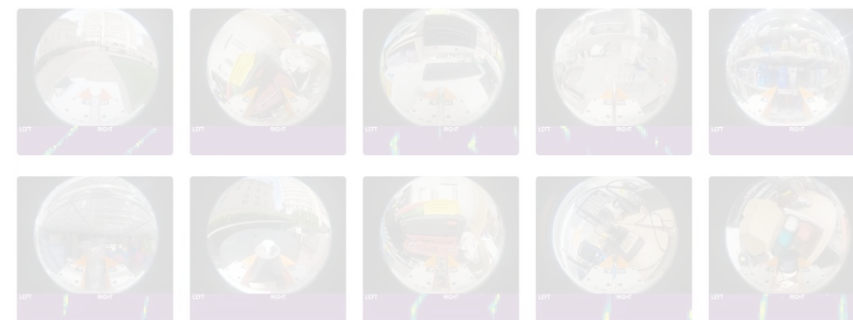
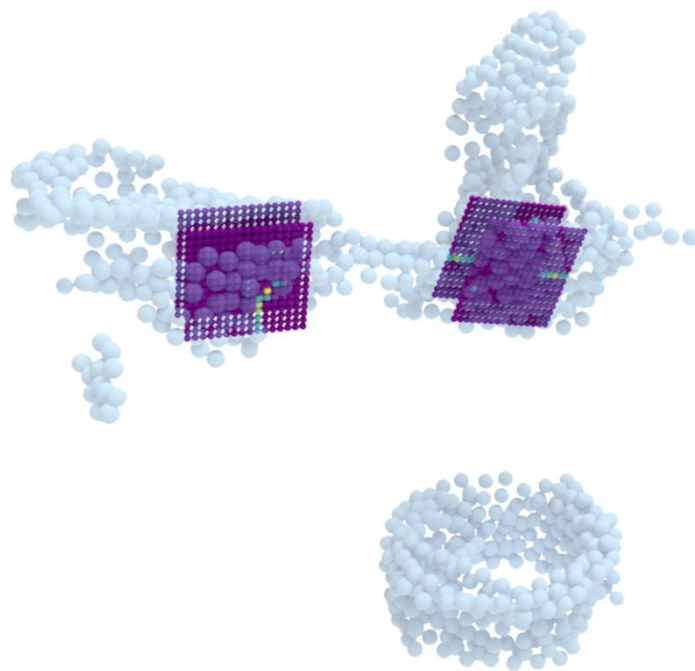
Hardware



Perception



Data



Large-Scale Tactile in Real



Large-Scale Tactile in Simulation

Overview

One full stack for scalable tactile manipulation
— build it, perceive with it, learn at scale.



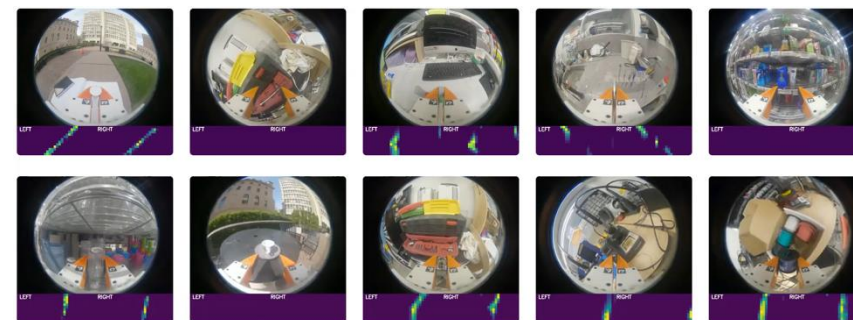
Hardware



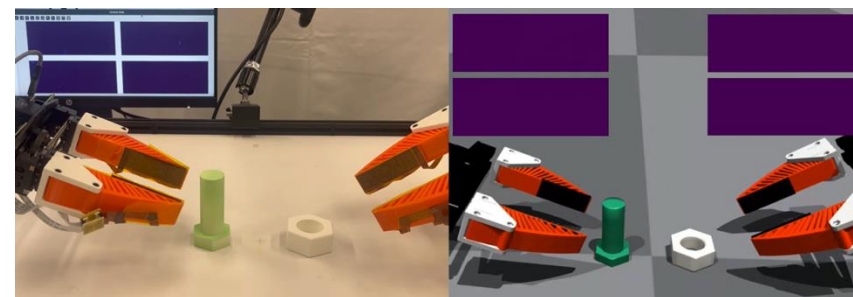
Perception



Data



Large-Scale Tactile in Real



Large-Scale Tactile in Simulation

Overview

One full stack for scalable tactile manipulation
— build it, perceive with it, learn at scale.



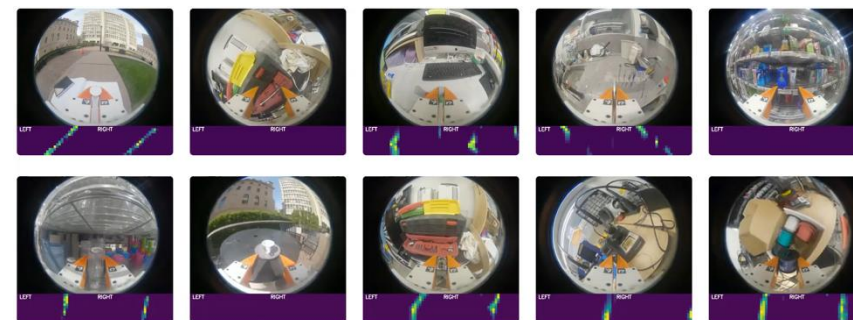
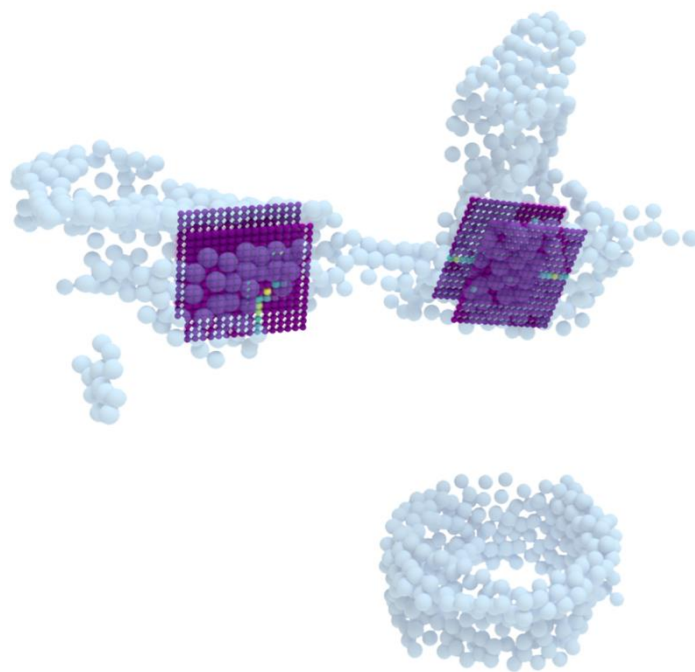
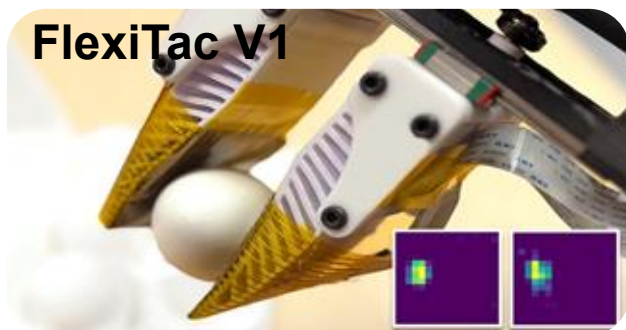
Hardware



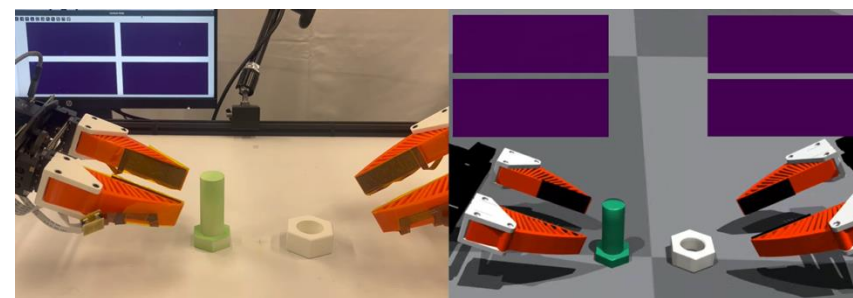
Perception



Data



Large-Scale Tactile in Real

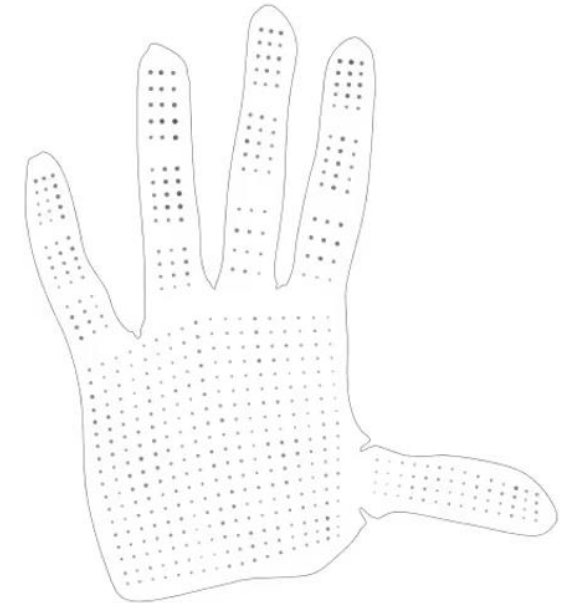


Large-Scale Tactile in Simulation

Start with Hardware



*Tactile glove, MIT Museum Collection
(Sundaram et al., Nature 2019)*



Tactile Glove Demo

Multisensory Machines

Hands-on Robots



Can robots identify and characterize an object simply through touch?

Designing a robot that can feel, grasp, and infer information about objects the way people can is a challenge, because we don't fully understand how tactile signals work in the human grasp. Using an array of sensors and neural networks to assist in object recognition and classification, this tactile glove models how different regions of the hand, whether individually or in concert with each other, can identify an object, make inferences about its weight, and understand the force needed to grasp it. This insight into the nature of the human grasp is vital to informing the design of robot-grasping tools as well as personal prosthetic devices.

Scalable Tactile Glove

To help understand the signatures of the human grasp, this limited glove uses an array of 540 sensors consisting of a piezoelectric film and a network of conductive thread electrodes. The dataset recorded with this glove includes interactions with 20 different objects over 100,000 frames.

Scalable tactile glove
Sensory: Sanyal, Poon, Tachibana, Yoshida, Li, Jan-Hui Zhu, Antonio Torralba, and Wallyash Mahesh
2018
Courtesy of MIT Computer Science and Artificial Intelligence Laboratory

How people used it

Hand poses

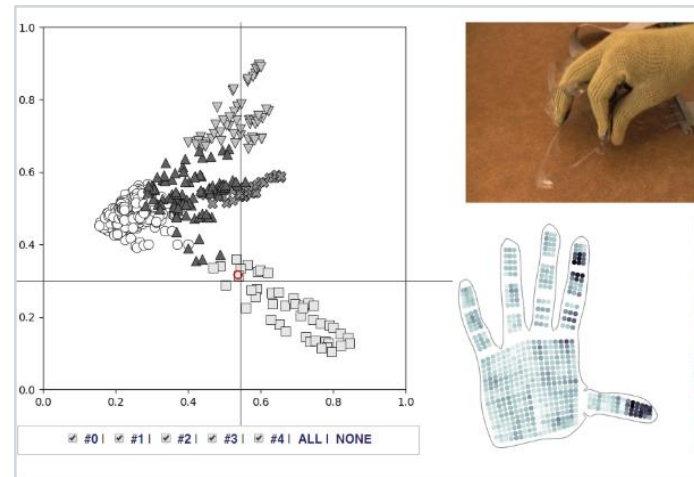
Start with Hardware



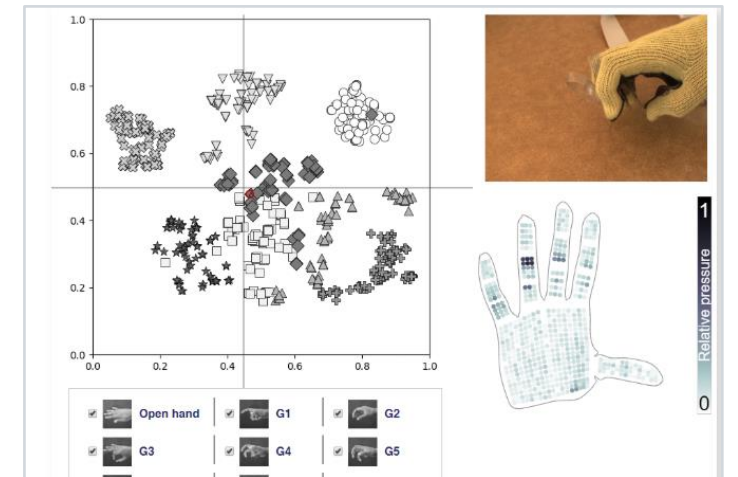
*Tactile glove, MIT Museum Collection
(Sundaram et al., Nature 2019)*

How people used it

Offline analysis of recorded grasps — not robot control



Frame clustering



Hand poses

Deployable on a real robot?

What Touch Needs to Reach Real Robots



*Tactile glove, MIT Museum Collection
(Sundaram et al., Nature 2019)*



GATE 1

Deployable on a real robot?

The best research sensor failed in all following points

- ✗ **~6 hours** — per sensor to fabricate (square form factor)
- ✗ **Messy cabling** — not suitable on an end-effector
- ✗ **Hard to reproduce** — open-source, but no one else use

What Touch Needs to Reach Real Robots



*Tactile glove, MIT Museum Collection
(Sundaram et al., Nature 2019)*



GATE 1

Deployable on a real robot?



GATE 2

Useful for the task/policy?

The best research sensor failed in all following points



~6 hours — per sensor to fabricate (square form factor)



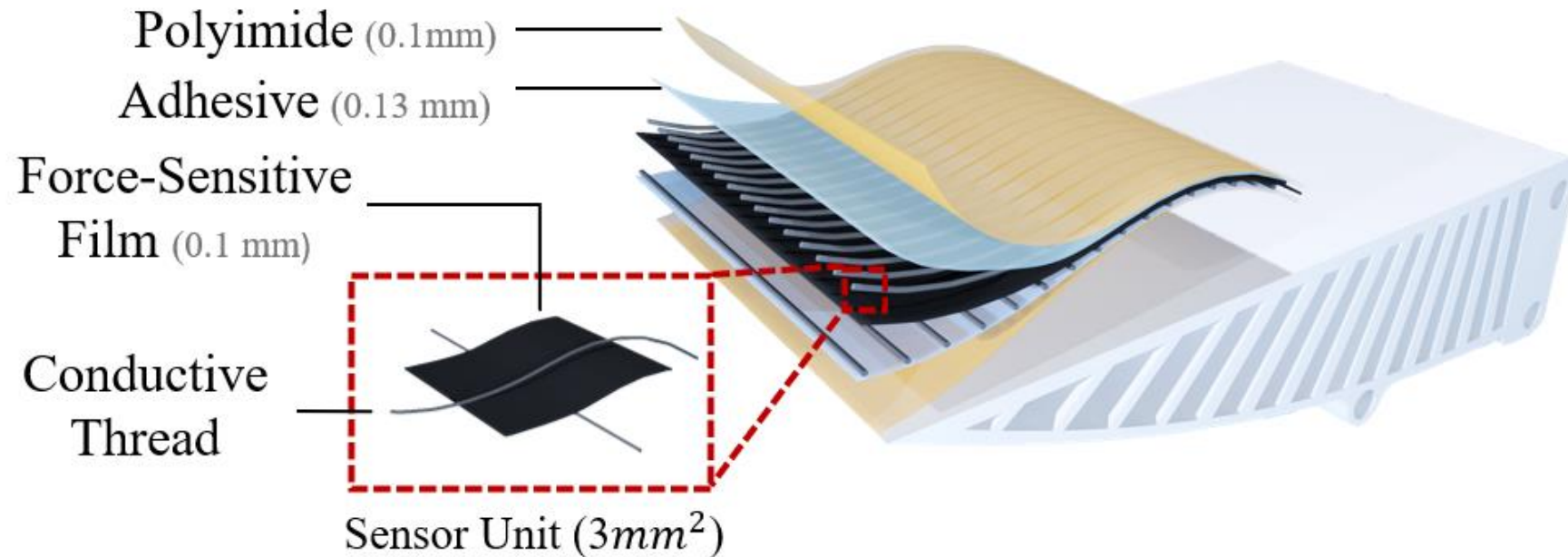
Messy cabling — not suitable on an end-effector



Hard to reproduce — open-source, but no one else use

Quick Background: How the Pressure Sensor Works

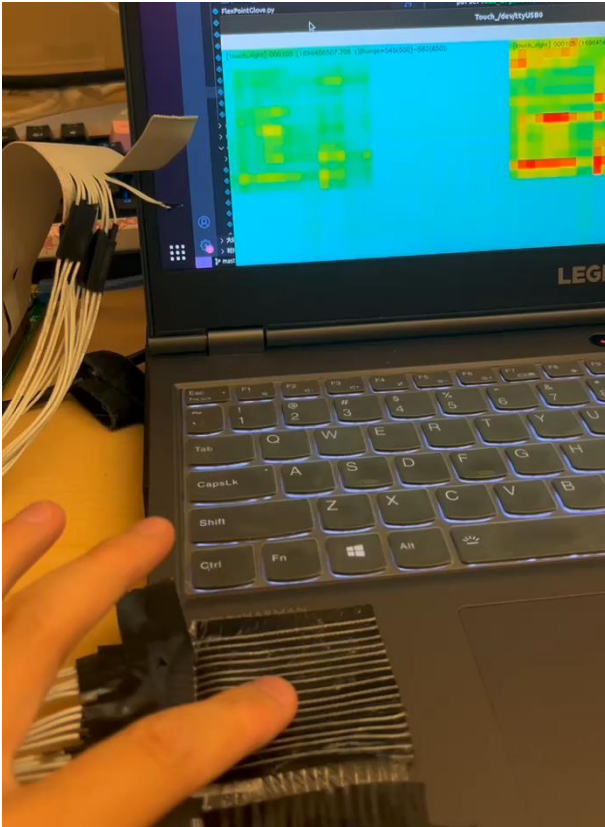
SENSING PRINCIPLE



A simple resistive "sandwich"

My first prototype — built by hand

BACK TO 2023



Hand-built prototype, 2023

Where it started — four real problems:



6 hours

lead time per sensor



Inconsistent

piece-to-piece
variation



Noisy

unstable raw signals

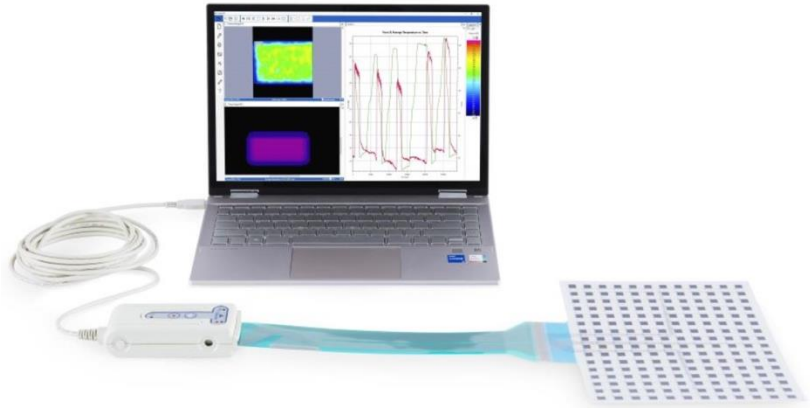


Unknown

robustness / lifetime

Fail to reproduce it well, I tried to buy it.

BACK TO 2023



Commercial Pressure Sensor
(Tekscan)

COMMERCIAL

Tekscan and similar

✗ ~\$1,000 for system, and no customization.

FACTORY

OEM

✗ Cheap off-the-shelf, but ~\$6K to customize.

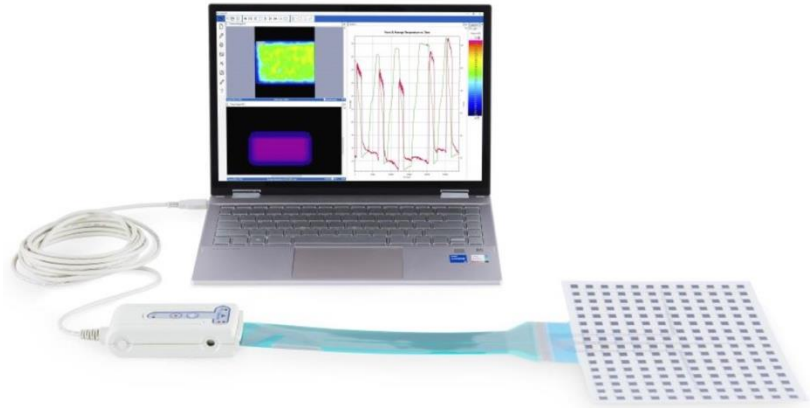
ACADEMIC

Lab's prototype

✗ No response, and a steep licensing fee.

Fail to reproduce it well, I tried to buy it.

BACK TO 2023



Commercial Pressure Sensor
(Tekscan)

COMMERCIAL

Tekscan and similar

✗ ~\$1,000 for system, and no customization.

FACTORY

OEM

✗ Cheap off-the-shelf, but ~\$6K to customize.

ACADEMIC

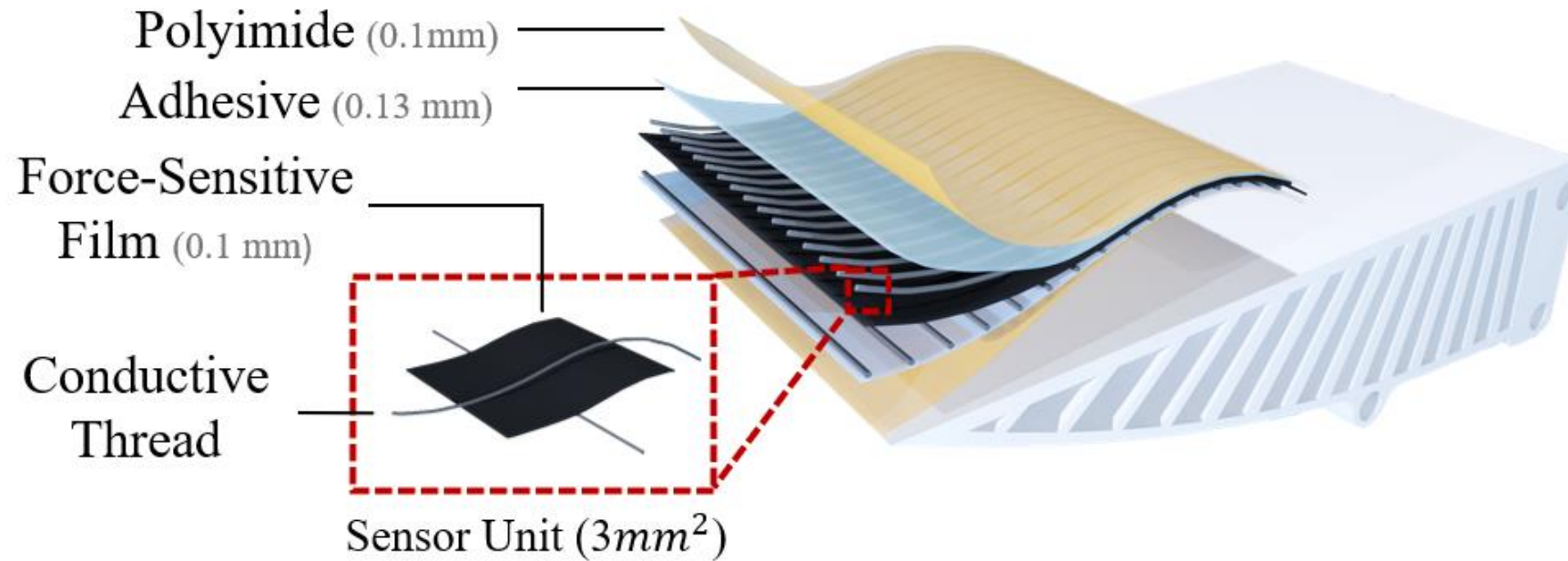
Lab's prototype

✗ No response, and a steep licensing fee.

Every option was a dead end — I almost gave up on this sensor.

Why not build from scratch, from principle?

SENSING PRINCIPLE



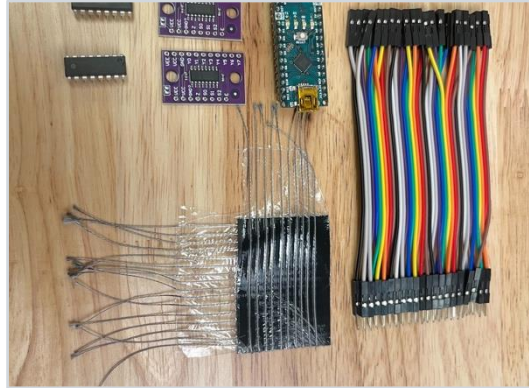
A simple resistive "sandwich"

Why not build from scratch, from principle?

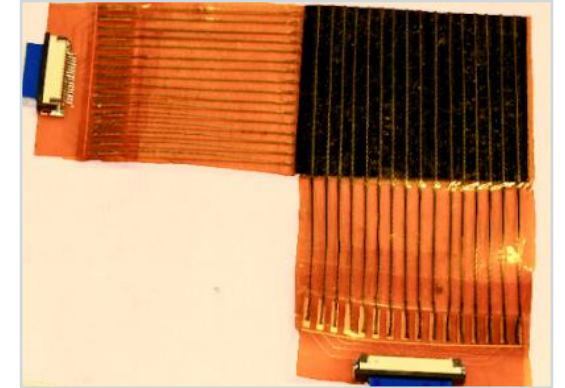
THE SENSOR · faster to fabricate each version



6 Hours

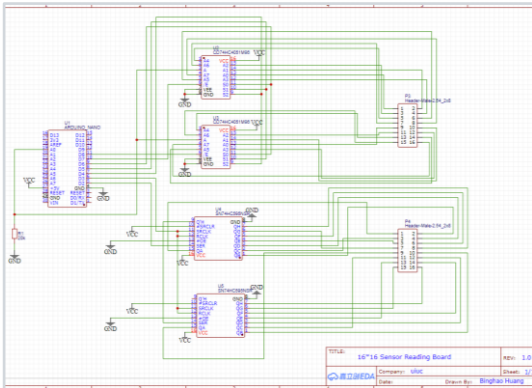


2 Hours

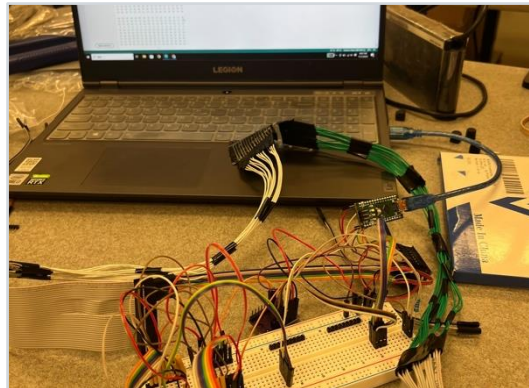


30 Minutes

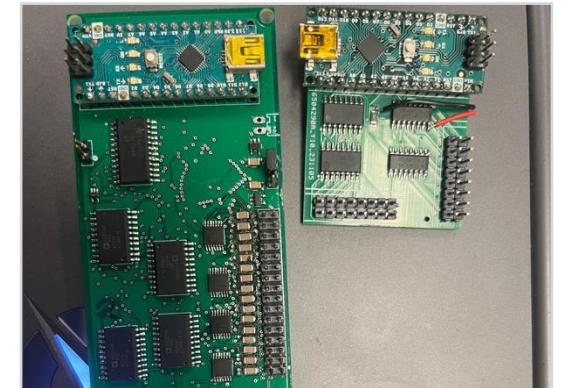
THE READOUT ELECTRONICS · from design to production



Design

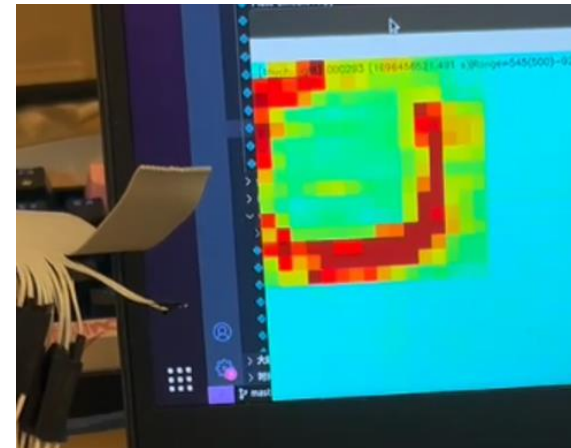
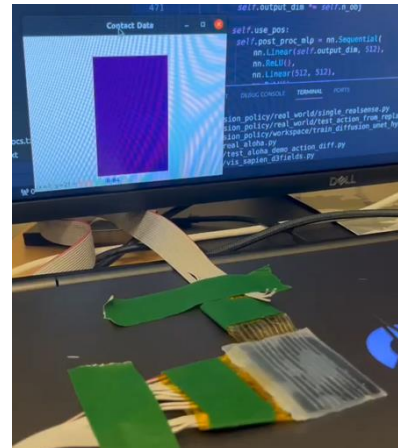
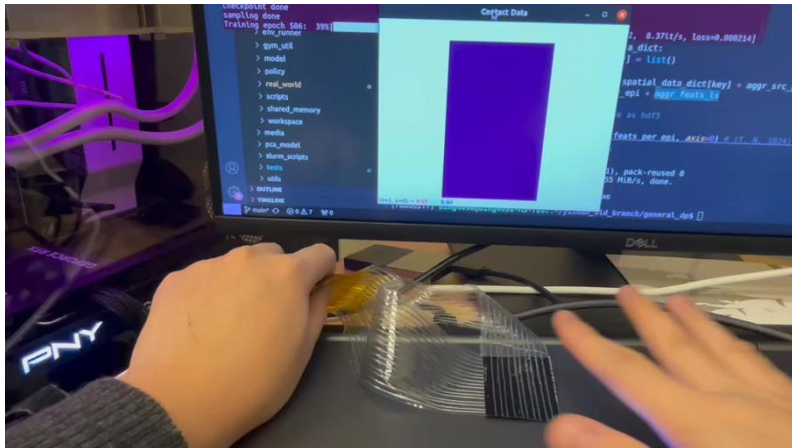
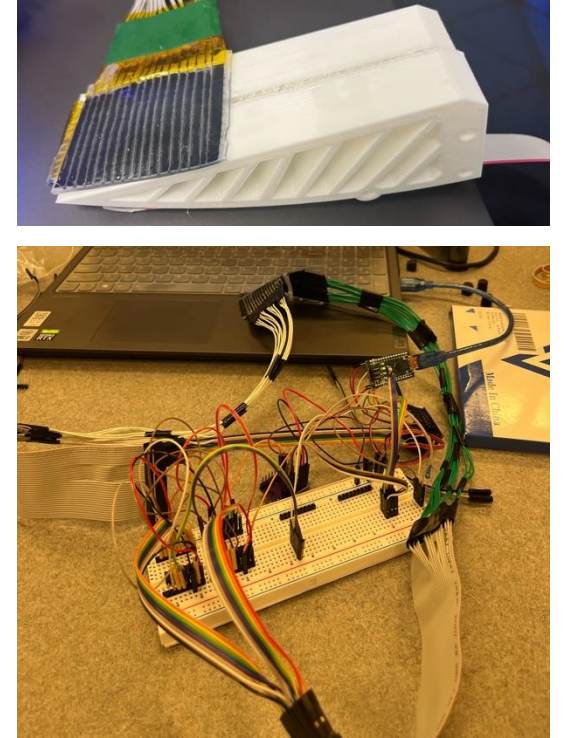
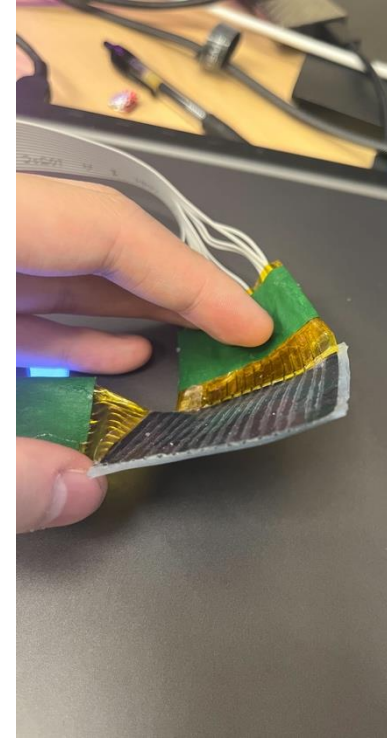
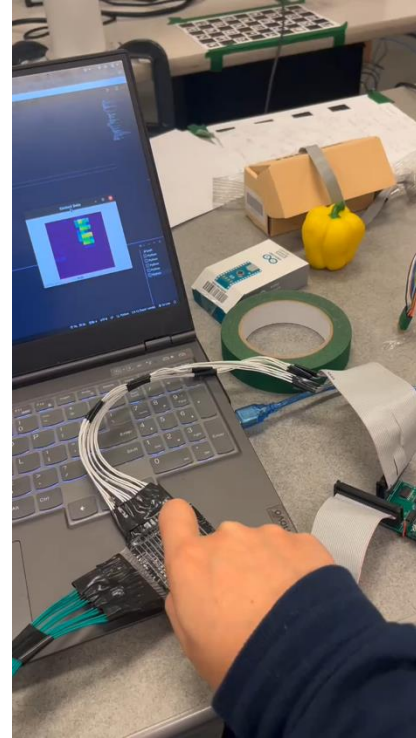
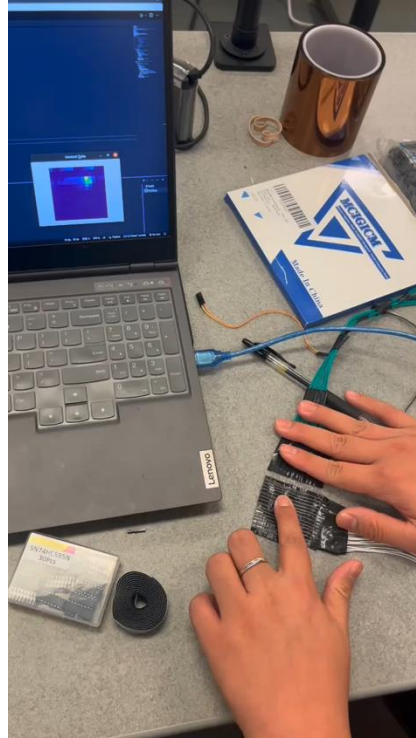


Testing

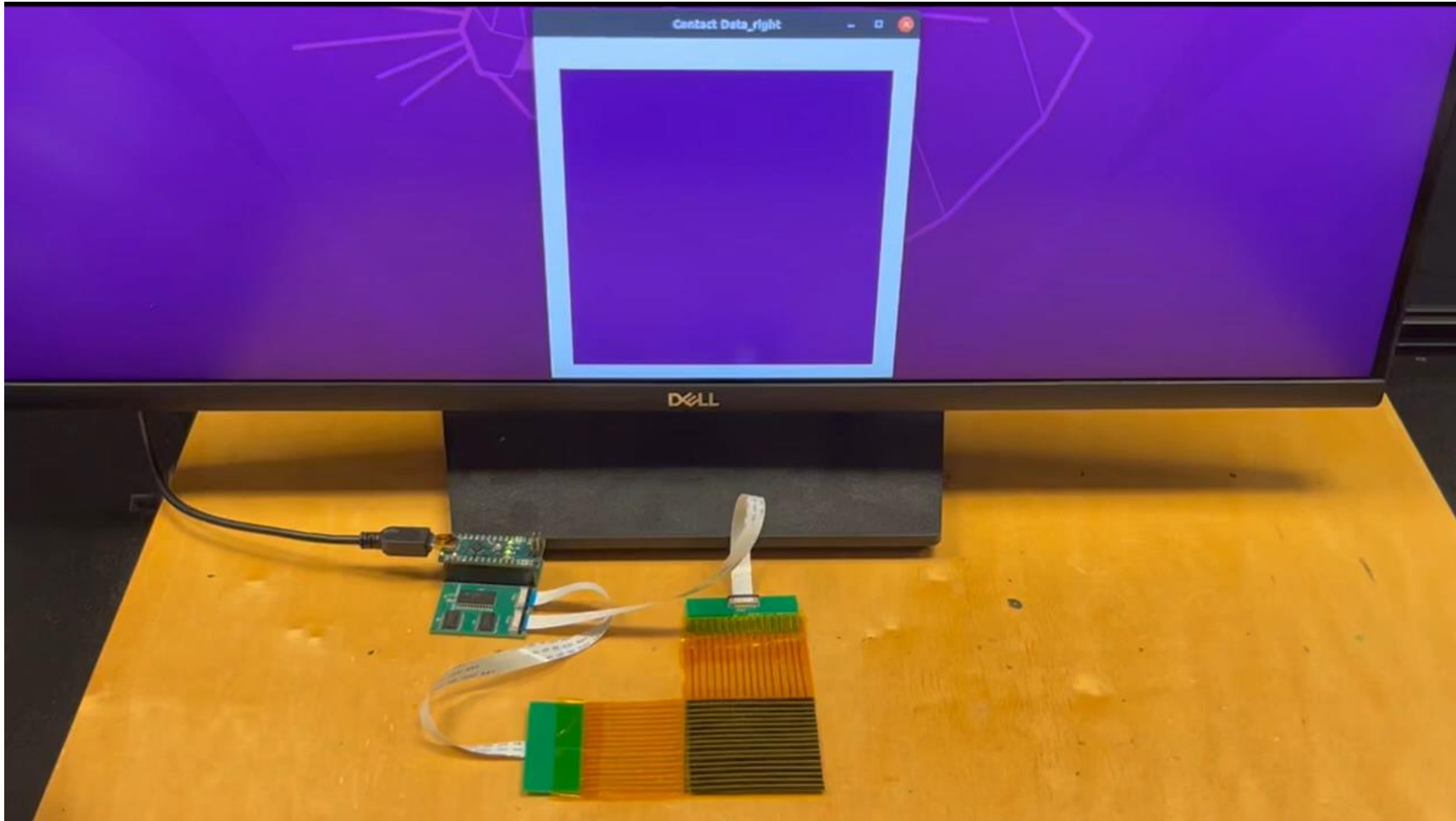


Production

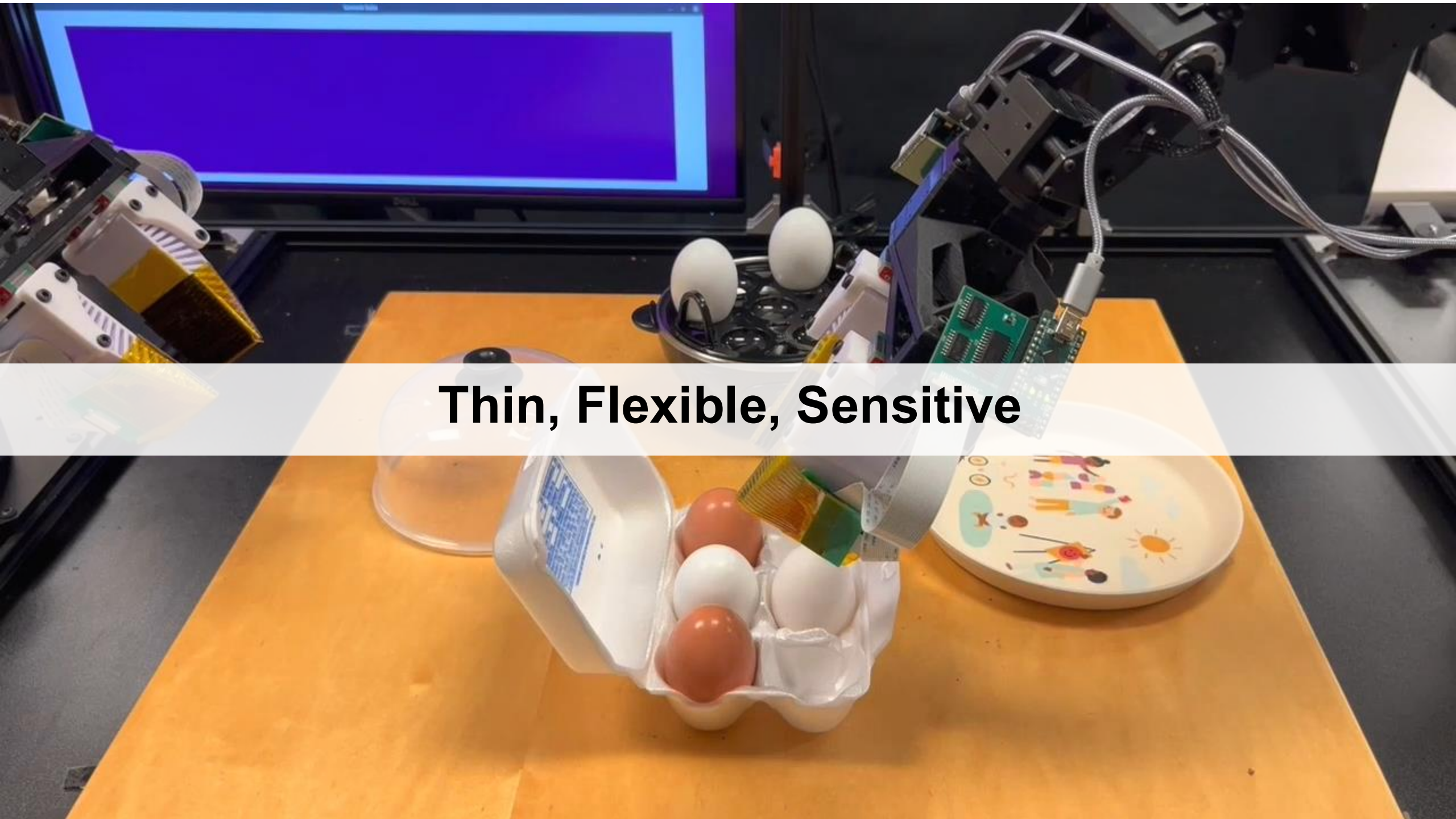
After months of testing and iteration



Finally, we got a stable version



Accessible, Low-cost, Robust!

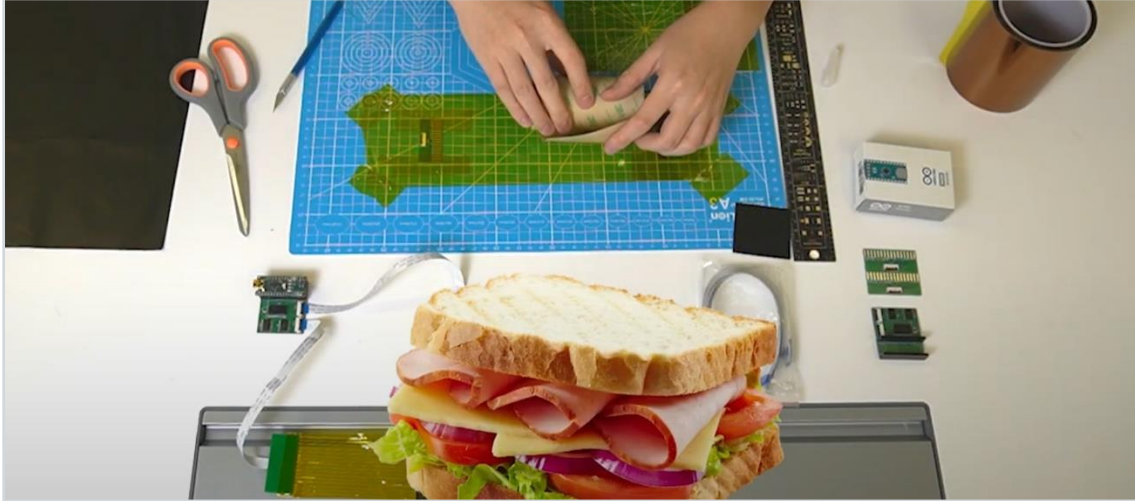


Thin, Flexible, Sensitive

A laboratory setting featuring a computer monitor in the background displaying a blue interface with four dark rectangular panels. In the foreground, a glass beaker containing a clear liquid sits on a white surface next to a glass thermometer with a blue bulb. The scene is framed by orange and white structural elements on the left and right sides.

Our Platform Today

Make Tactile Sensor Accessible to Everyone



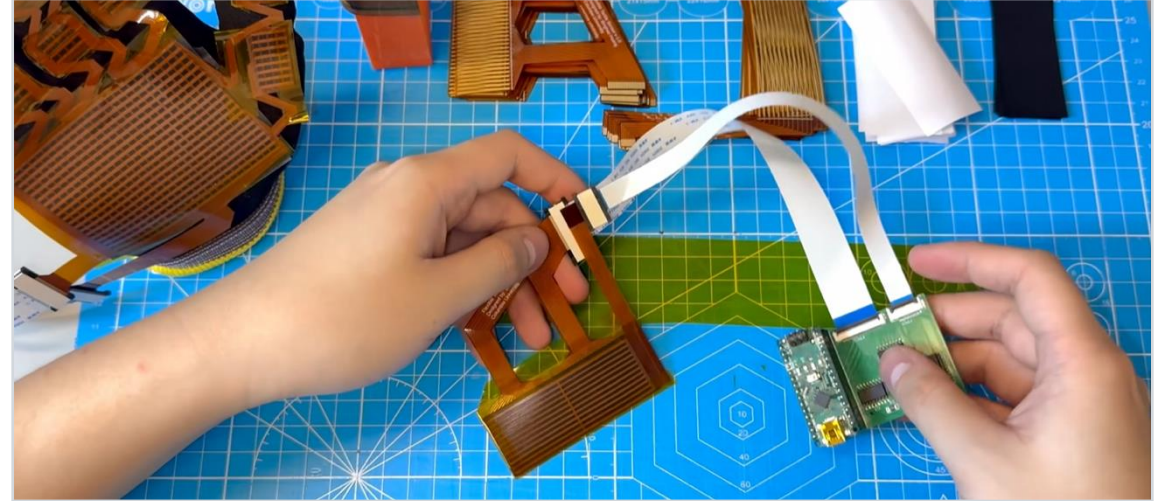
FlexiTac-V1

Fabrication ~30 min

Price ~\$5/piece

Customize hand-made

Lead time in-lab ~30 min



FlexiTac-V2

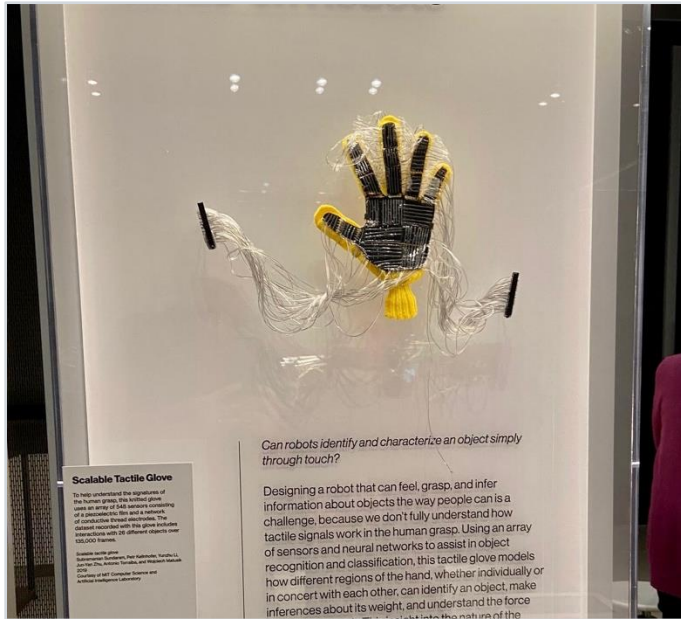
Fabrication ~3 min

Price \$4.19/piece (20 unit)
\$1.36/piece (1000 unit)

Customize from PCB factory

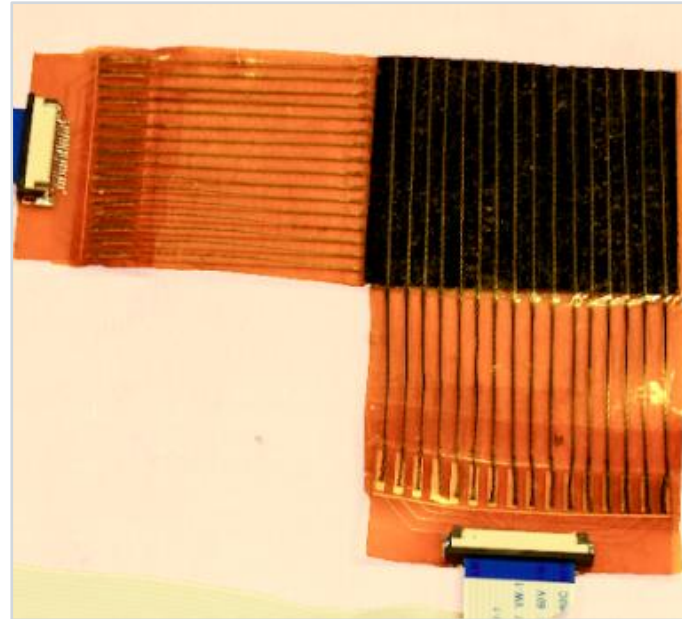
Lead time manufacture + ship < 1 week

Why Adoption Matters in Robotic Touch



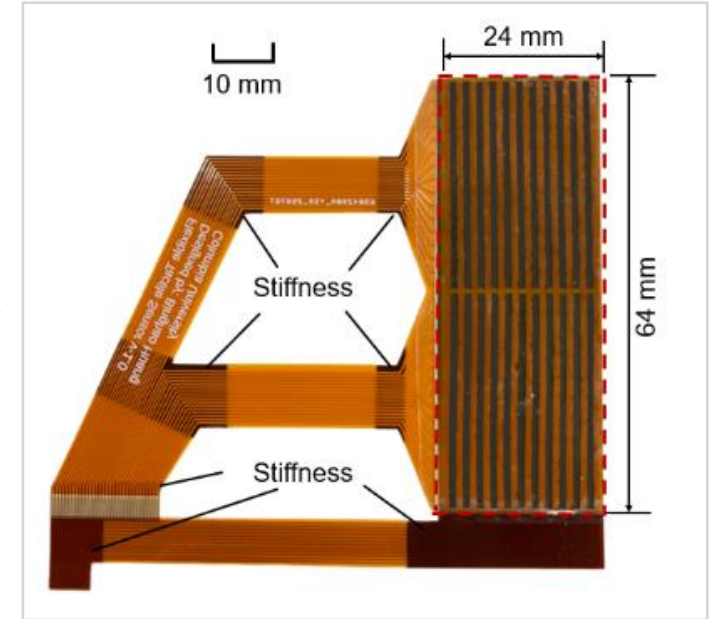
2019

Only MIT used it



2024

More research labs adopt it



2026



Broad academia + industry use

Our claim: make sensors accessible, and the field builds on them.

We Open-Source FlexiTac!

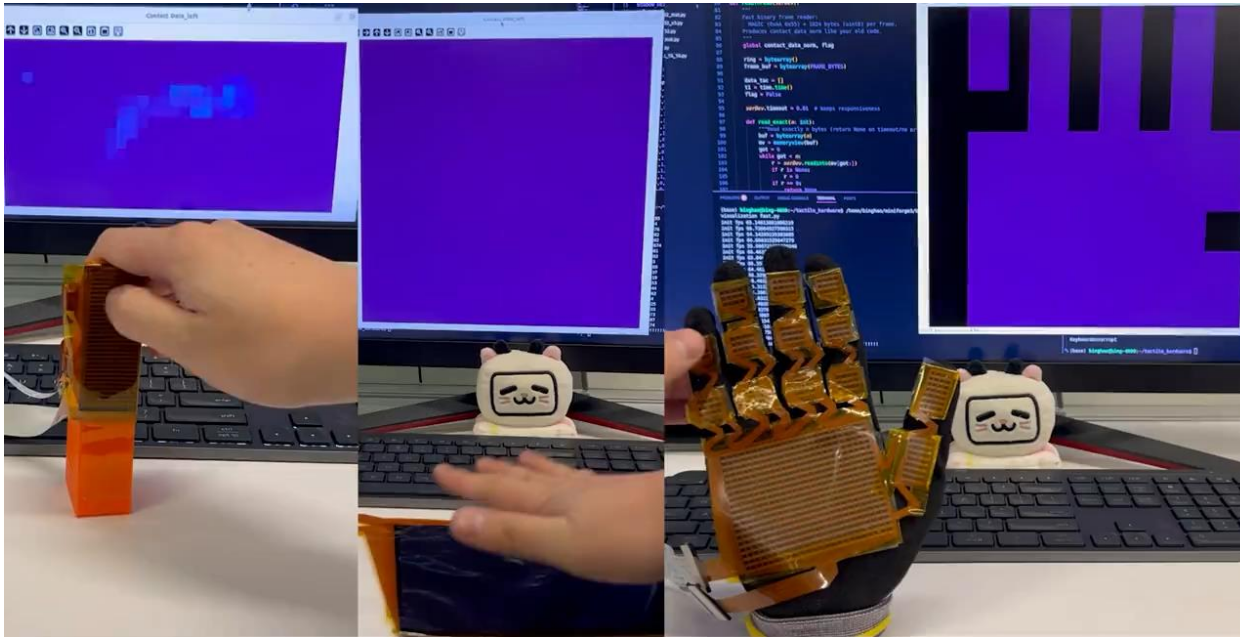
FlexiTac Hardware Fabrication Guide

Binghao Huang¹, Yunzhu Li¹

¹Columbia University

March 2026

Simple but Works!



Get Your Tactile Sensor in 3 mins!

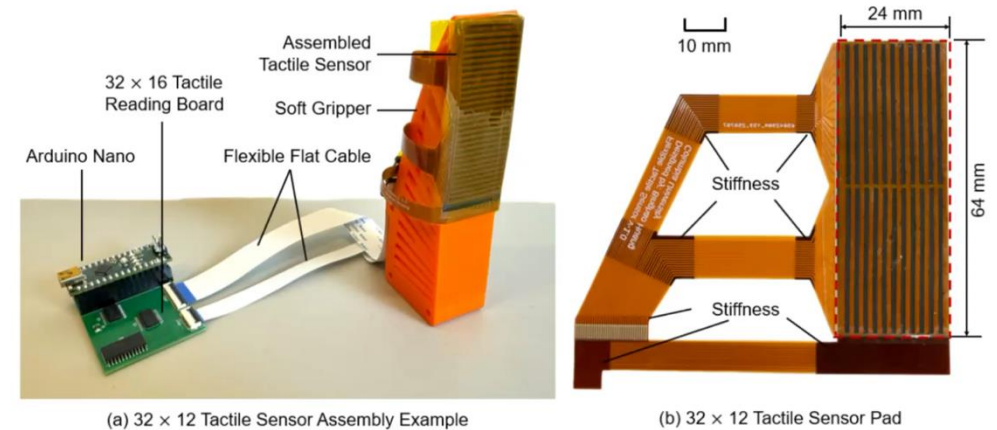
FlexiTac Hardware Fabrication Guide

Binghao Huang¹, Yunzhu Li¹

¹Columbia University

March 2026

Simple but Works!



(a) 32 × 12 Tactile Sensor Assembly Example

(b) 32 × 12 Tactile Sensor Pad

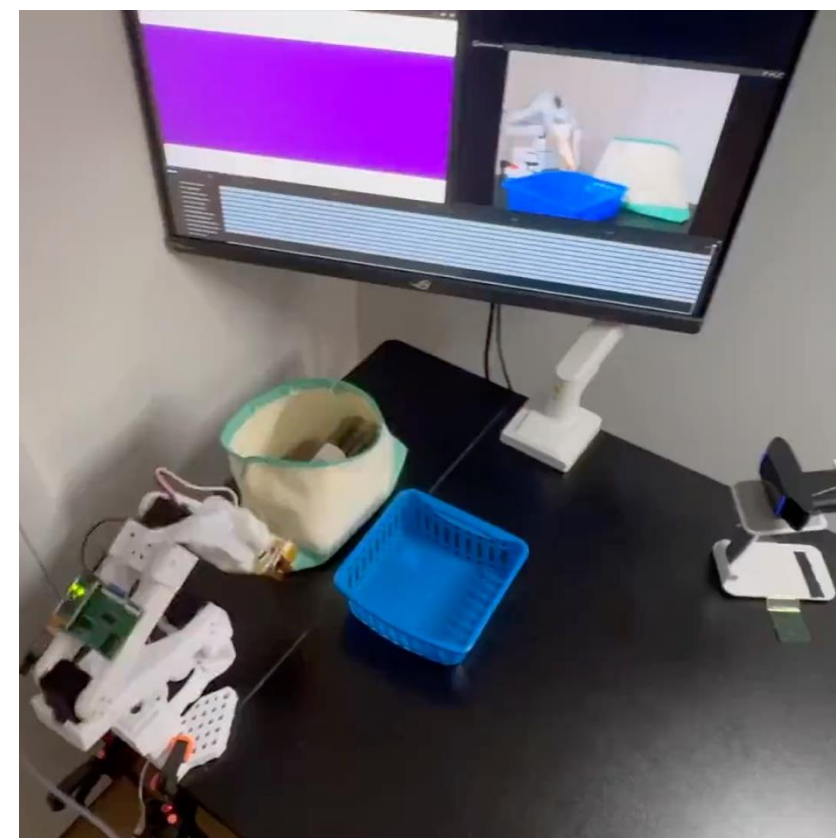
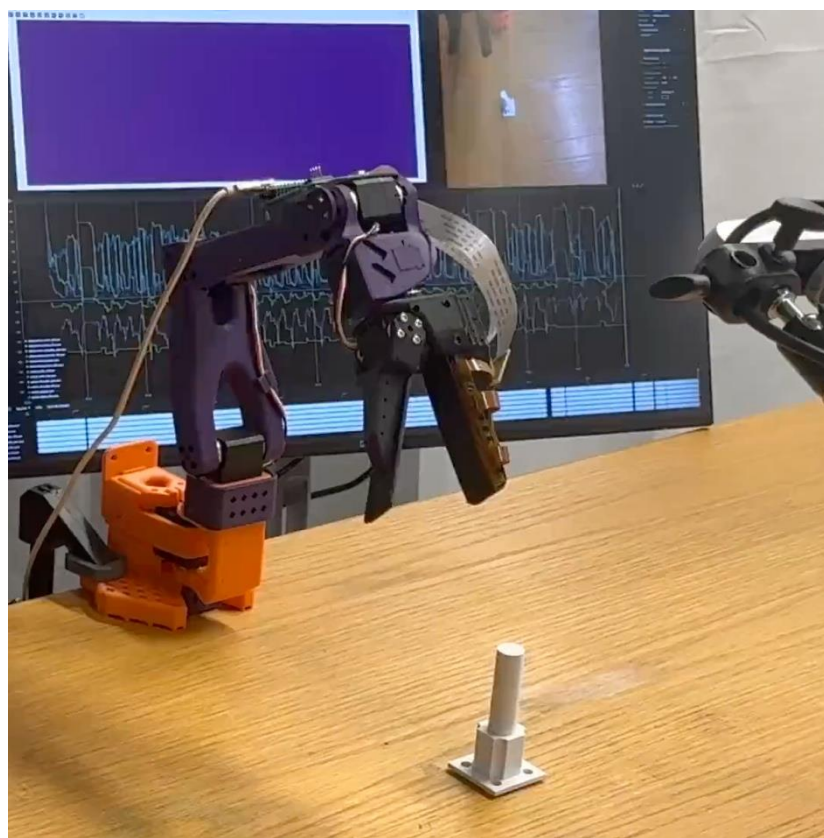
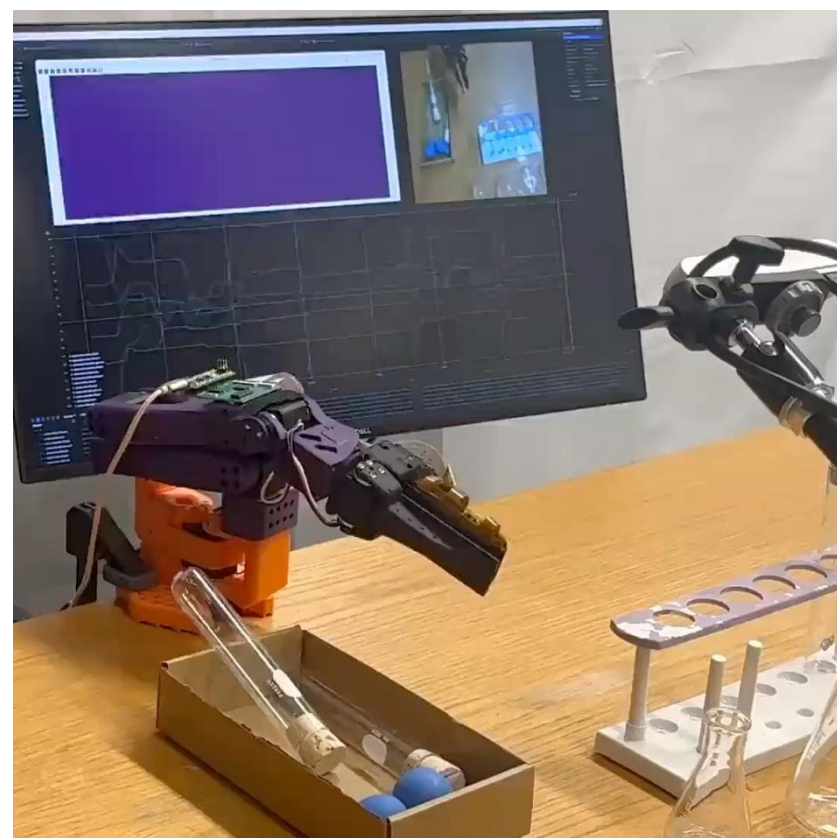
Get Your Tactile Sensor in 3 mins!

Resources

✓ Contact

Please email binghao.huang@columbia.edu for any questions.

... and Integrate into the LeRobot Community



Naian Tao, Yifan He, Wesley Maa, Binghao Huang, and Yunzhu Li, "LeFlexiTac: Giving Robots a Sense of Touch,"
Columbia University RoboPIL Blog, 2026. [\[Link\]](#)

One Year Later: A Growing FlexiTac Ecosystem

Awesome FlexiTac

FlexiTac Homepage

Join Discord

Categories

Filter Research

All Works (25)

Hardware Design (12)

Glove/Hand (4)

Multi-Modal Learning (11)

Simulation (2)

Locomanipulation (3)

Industry (2)

Dataset (2)

AWESOME LIST

Awesome FlexiTac

The growing ecosystem of research building on FlexiTac, together with the prior works.

Maintained by [FlexiTac Team](#)

Have a relevant work? [Open a PR or issue](#) to add it.

All Works

Hardware Design

Glove/Hand

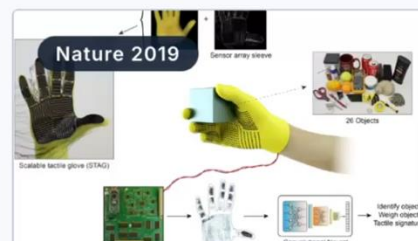
Simulation

Multi-Modal Learning

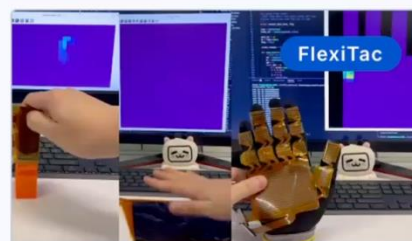
Locomanipulation

Industry

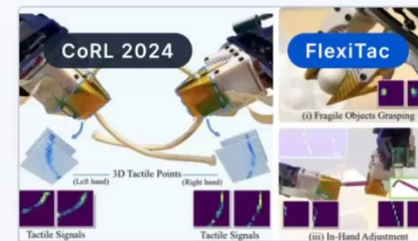
Dataset



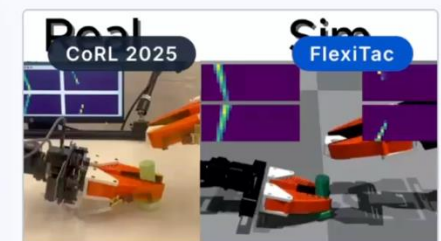
Learning the Signatures of the Human Grasp Using a Scalable Tactile Glove



FlexiTac: An Open-Source, Scalable Tactile Solution for Robotic Systems



3D-ViTac: Learning Fine-Grained Manipulation with Visuo-Tactile Sensing



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



One Year Later: A Growing FlexiTac Ecosystem

npj Robotics 2026



Zhou et al., “Tactile-Reactive Gripper with an Active Palm for Dexterous Manipulation,” **Nature Partner Journals Robotics 2026.**

ICRA 2026



Chen et al., “Multi-Modal Manipulation via Multi-Modal Policy Consensus,” **ICRA 2026.**

ECCV 2026



Li et al., “Deform360: A Massive Multi-view Visuotactile Dataset for Deformable World Models,” **ECCV 2026.**

IROS 2026



Zhou et al., “Learning Tactile-Aware Quadrupedal Loco-Manipulation Policies,” **IROS 2026.**

Under Review



Ma et al., “TacCoRL: Integrating Tactile Feedback into VLA via Simulation,” 2026.

Under Review



Lin & Zhao, “ART-Glove: Articulated Tactile Glove for Contact-Grounded Dexterous Interaction Capture,” 2026.

Acknowledgement



Massachusetts
Institute of
Technology



Carnegie
Mellon
University



UNIVERSITY OF
ILLINOIS
URBANA-CHAMPAIGN

Georgia
Tech.



USC

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK



NVIDIA®

Hardware

Perception

Data

Market Tailwind: The Industry Is Moving to Touch

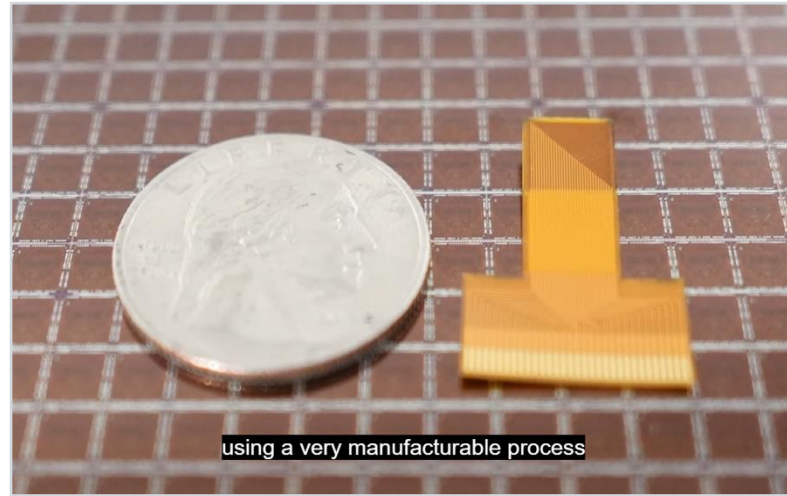
NEW PRODUCT

Tactile sensing at 5× the resolution of a human finger



ADVANCED PRODUCTION

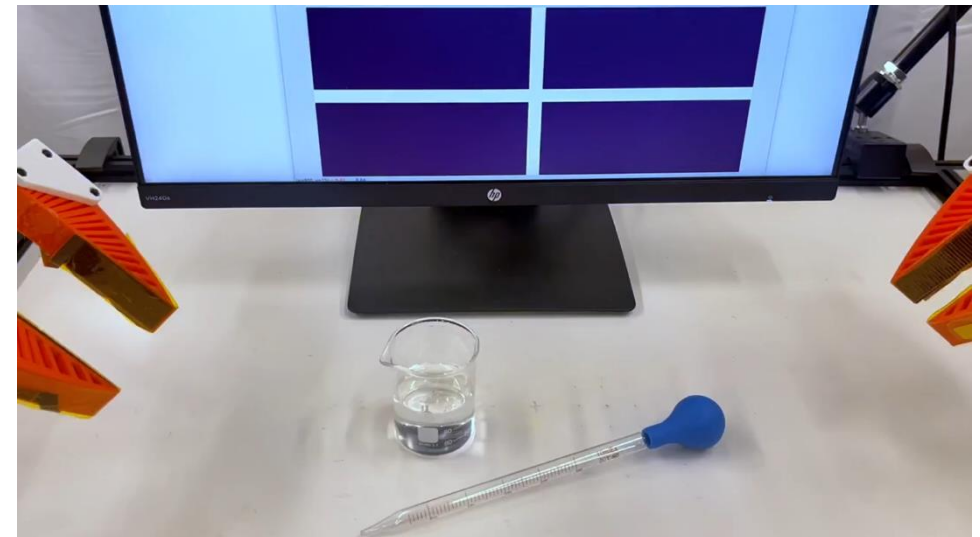
Built on semiconductor fabrication technology



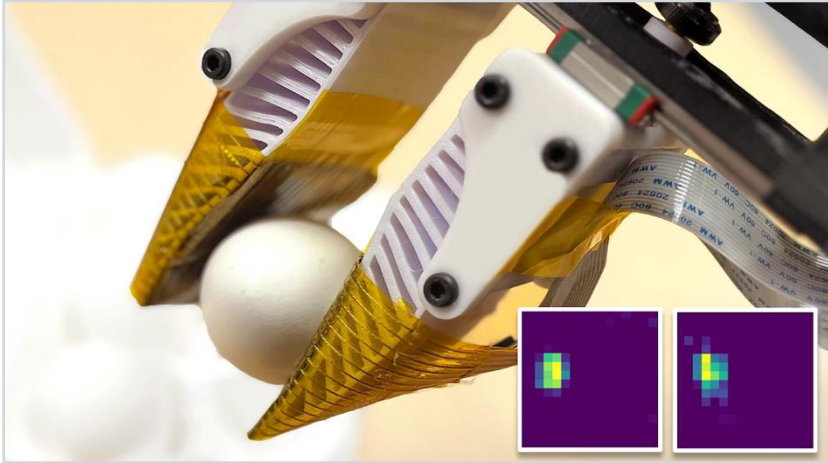
[1] Redefining Tactile Sensing. <https://www.analog.com/en/resources/media-center/videos/6379738445112.html>

[2] THE TOUCH THAT WILL TRANSFORM ROBOTICS. <https://www.analog.com/en/signals/articles/the-touch-that-will-transform-robotics.html>

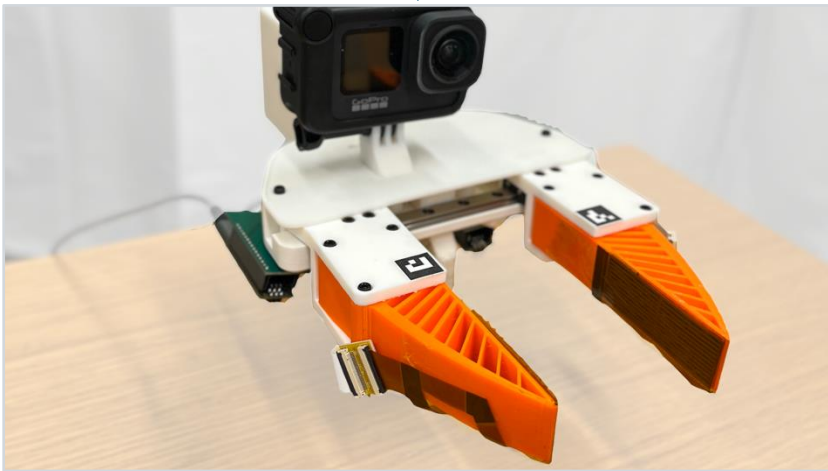
- **Scalable**
- **Flexible**
- **Customizable**
- **Robustness/Consistency**
- **Easy to Simulate**



What Touch Needs to Reach Real Robots



FlexiTac V1



FlexiTac V2



GATE 1

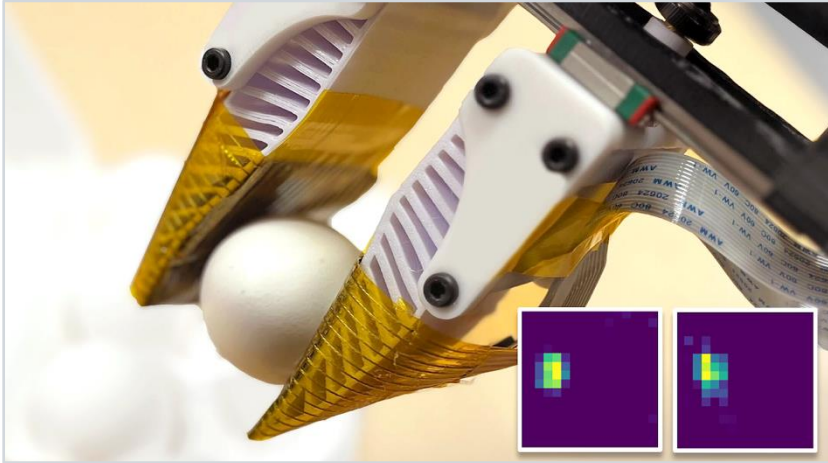
Deployable on a real robot?



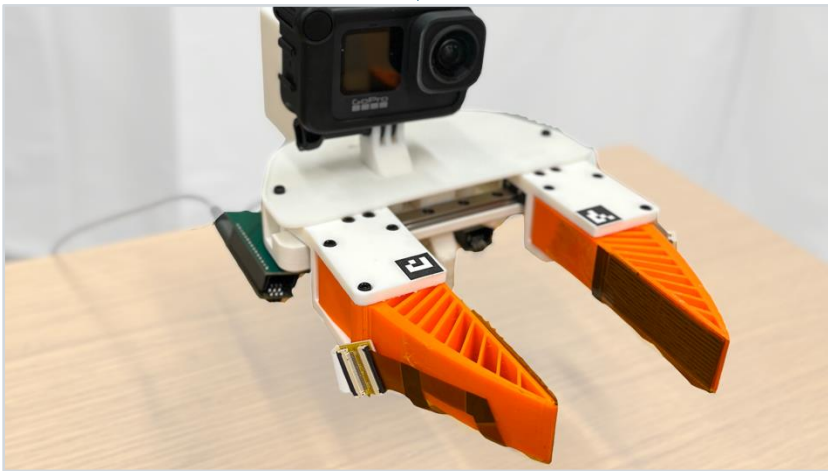
GATE 2

Useful for the task/policy?

What Touch Needs to Reach Real Robots



FlexiTac V1



FlexiTac V2



GATE 1

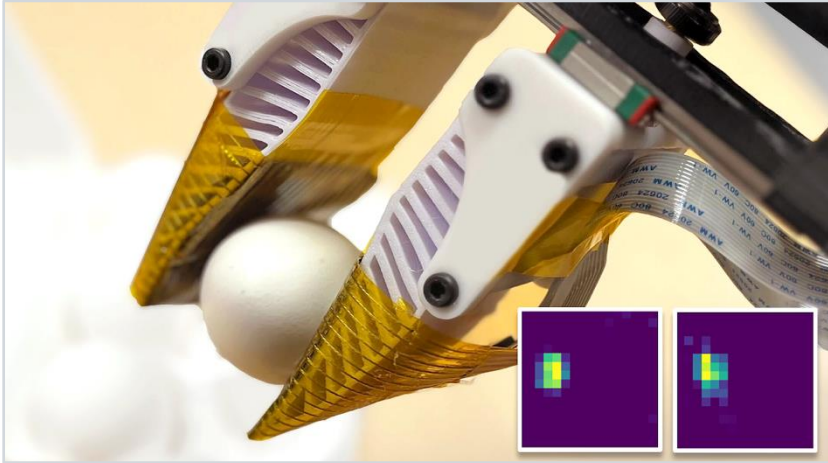
Deployable on a real robot?



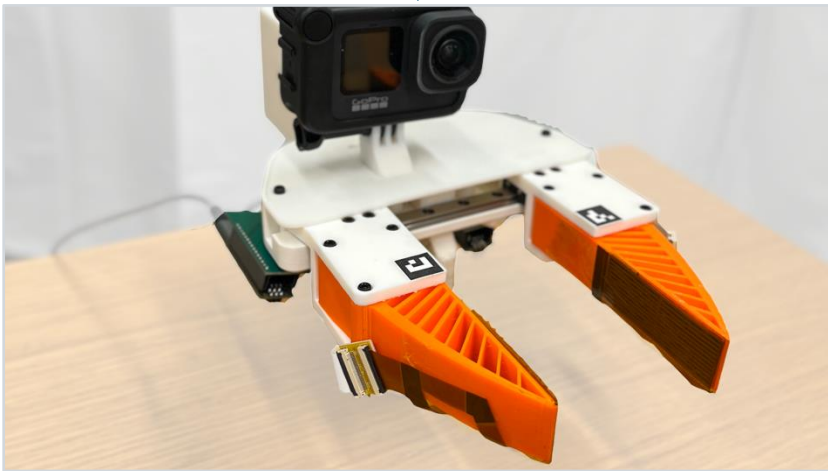
GATE 2

Useful for the task/policy?

What Touch Needs to Reach Real Robots



FlexiTac V1



FlexiTac V2



GATE 1

Deployable on a real robot?



GATE 2

Useful for the task/policy?

Hardware alone isn't enough.

What makes the difference?

Overview

One full stack for scalable tactile manipulation
— build it, perceive with it, learn at scale.



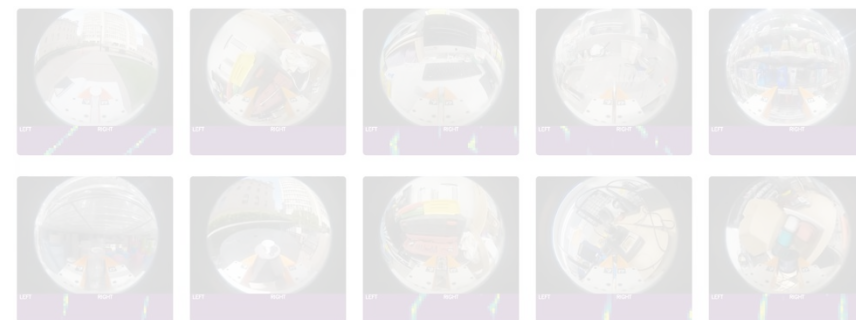
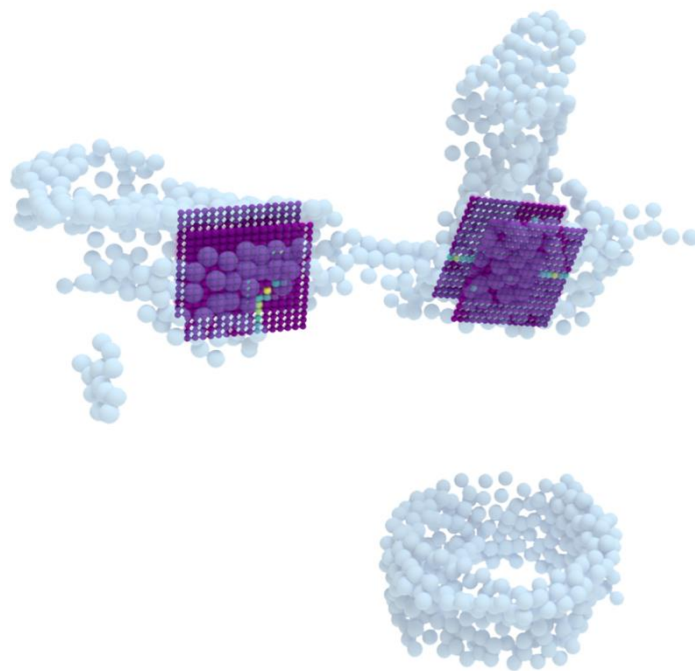
Hardware



Perception



Data



Large-Scale Tactile in Real

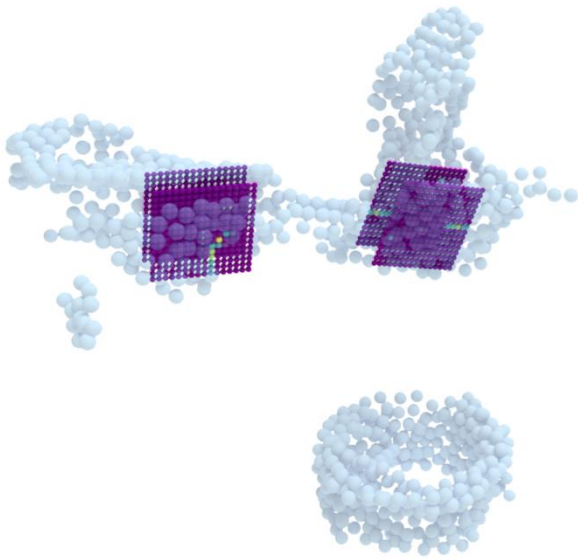


Large-Scale Tactile in Simulation

Overview



Perception



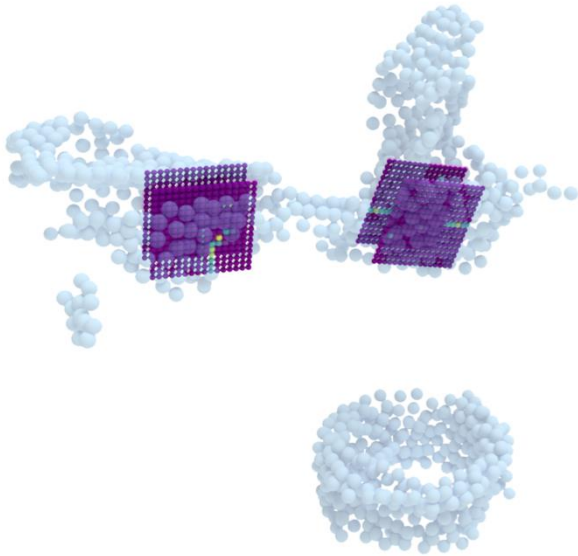
Adding touch should help.

So **why** does performance **drop**?

Overview



Perception



Adding touch should help.

So **why** does performance **drop**?

THE REASON: the two signals are distinct

VISION

High-dimensional
Global scene

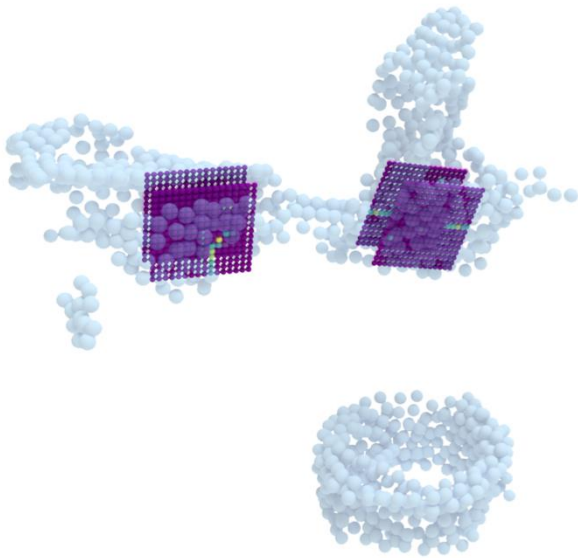
TOUCH

Low-dimensional
Local contact

Overview



Perception



Adding touch should help.

So **why** does performance **drop**?

THE REASON: the two signals are distinct

VISION

High-dimensional
Global scene

TOUCH

Low-dimensional
Local contact

The key is: How we fuse it?

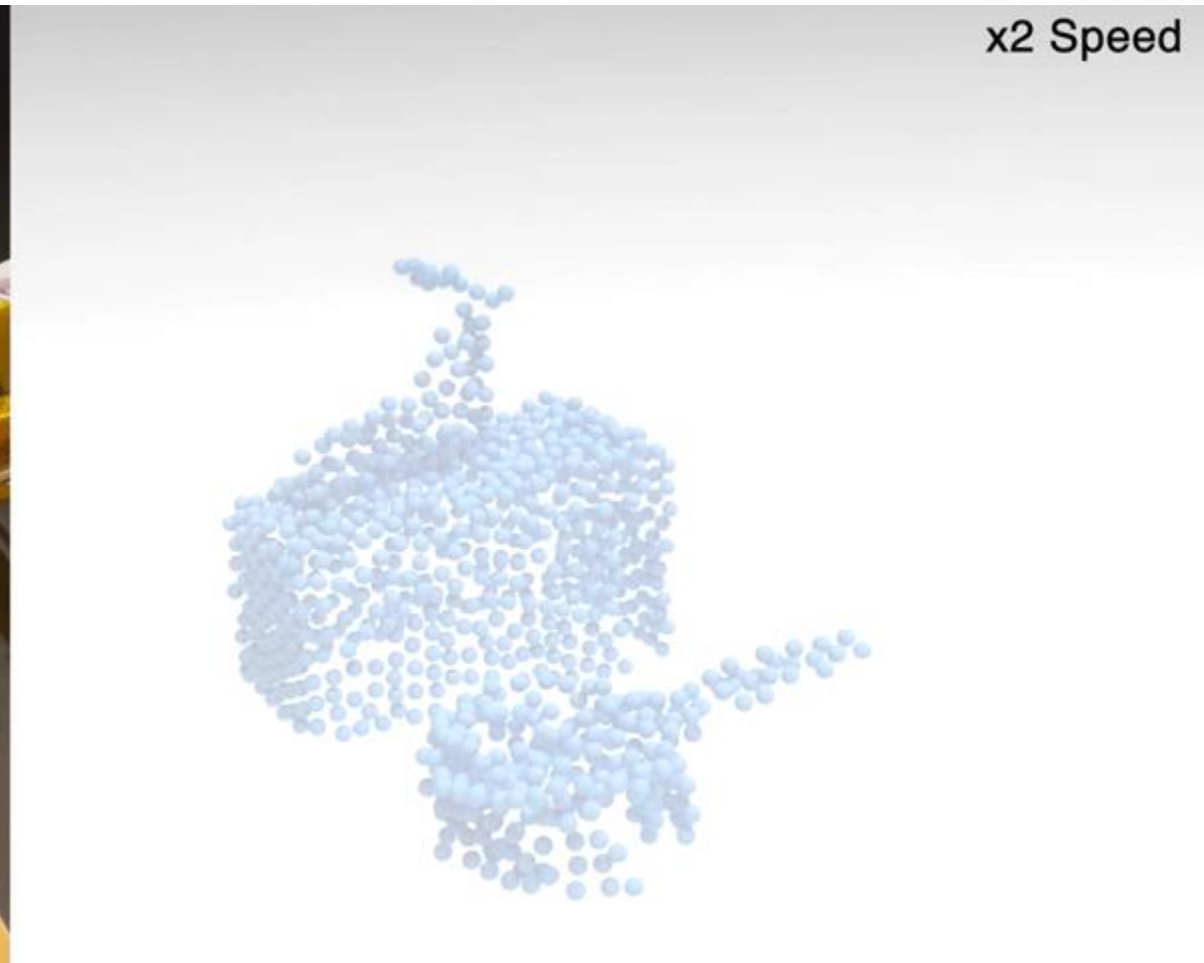
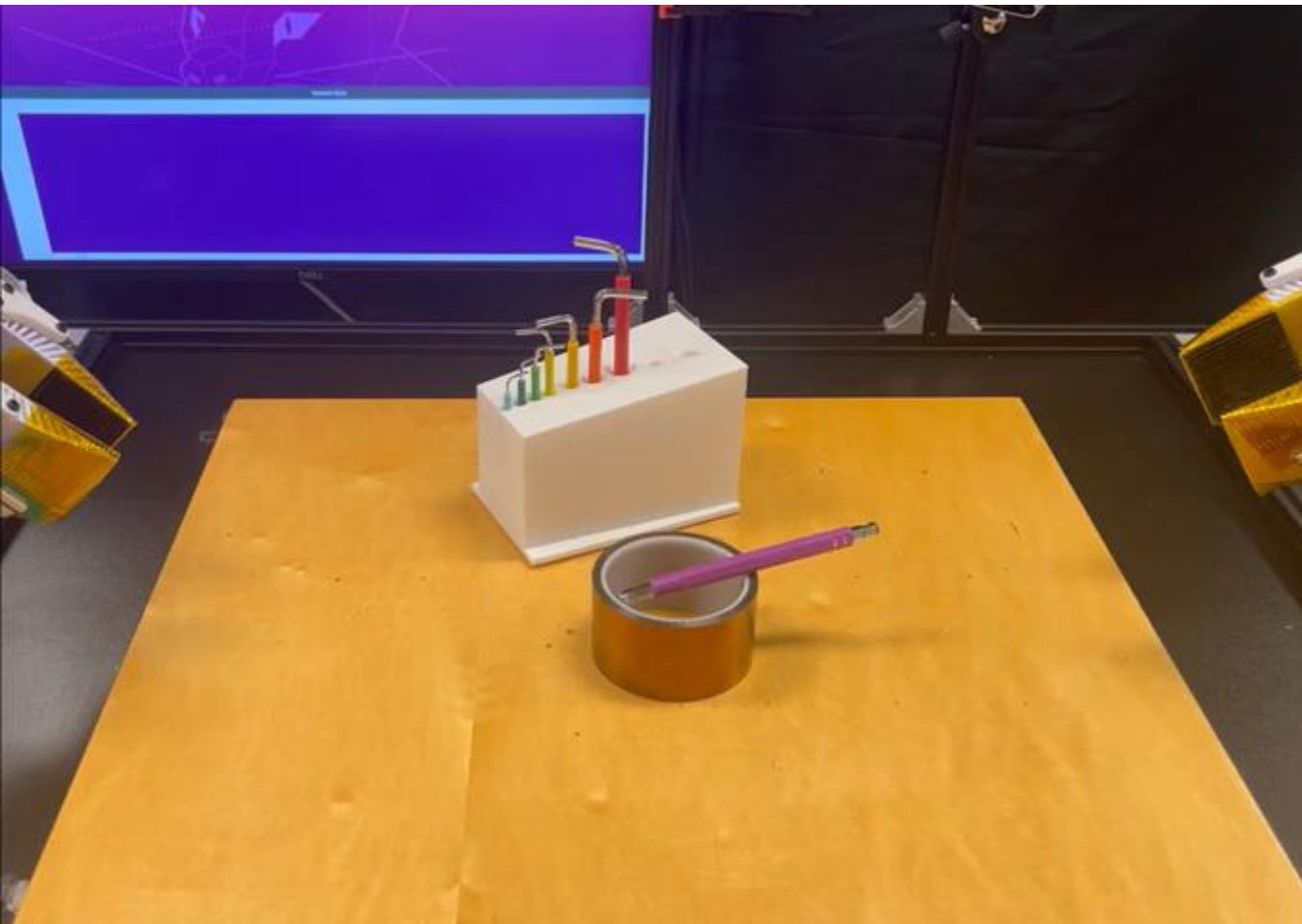
3D-ViTac:

Learning Fine-Grained Manipulation with Visuo-Tactile Sensing

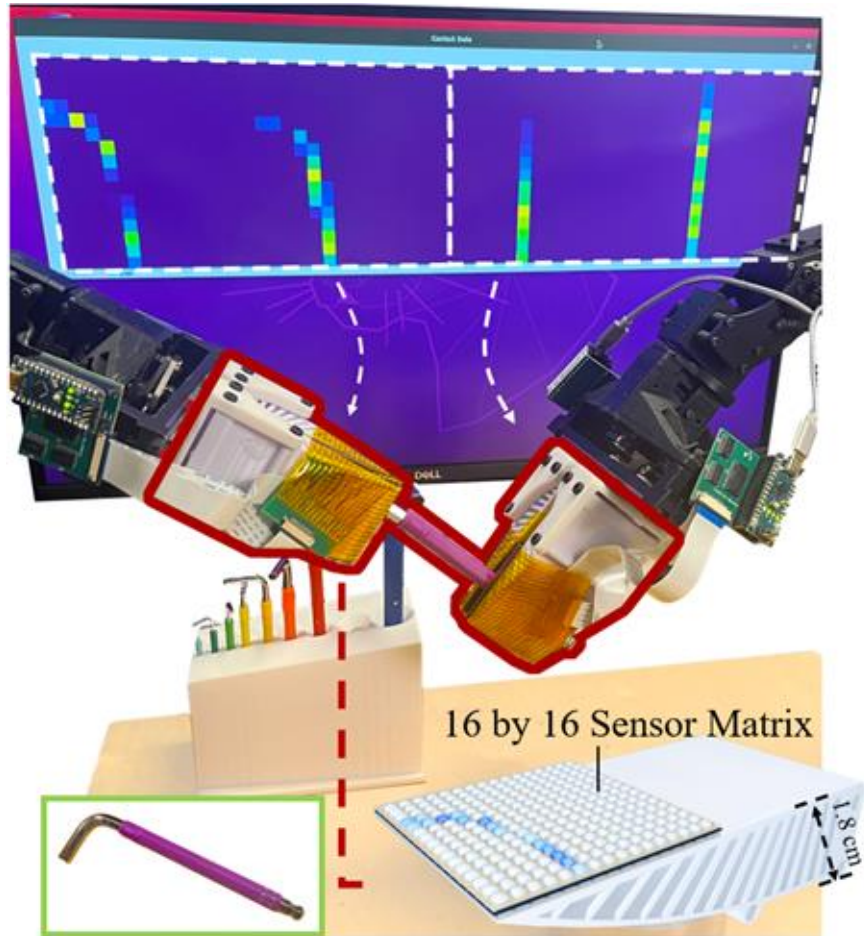
Binghao Huang¹, Yixuan Wang¹, Xinyi Yang², Yiyue Luo³, Yunzhu Li¹

¹Columbia University, ²University of Illinois Urbana-Champaign, ³University of Washington

Conference on Robot Learning (CoRL) 2024



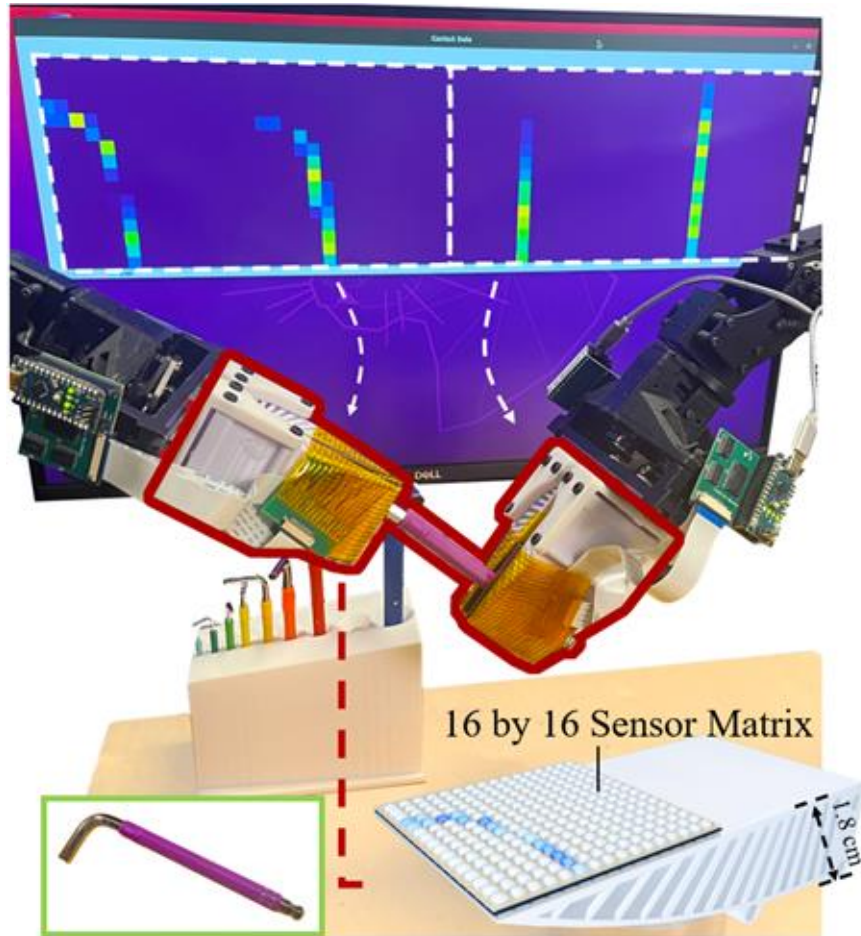
How We Build Multi-Sensory System?



High Resolution Tactile with Multi-View RGBD Camera

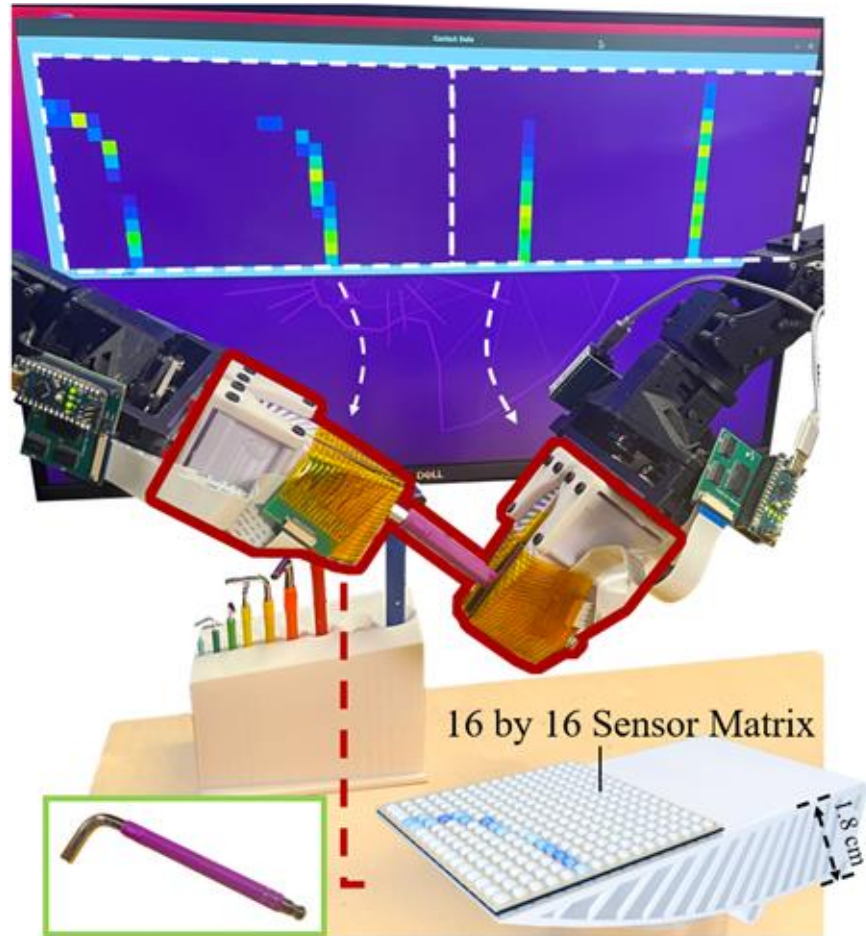
How We Build Multi-Sensory System?

Visual Haptic Feedback



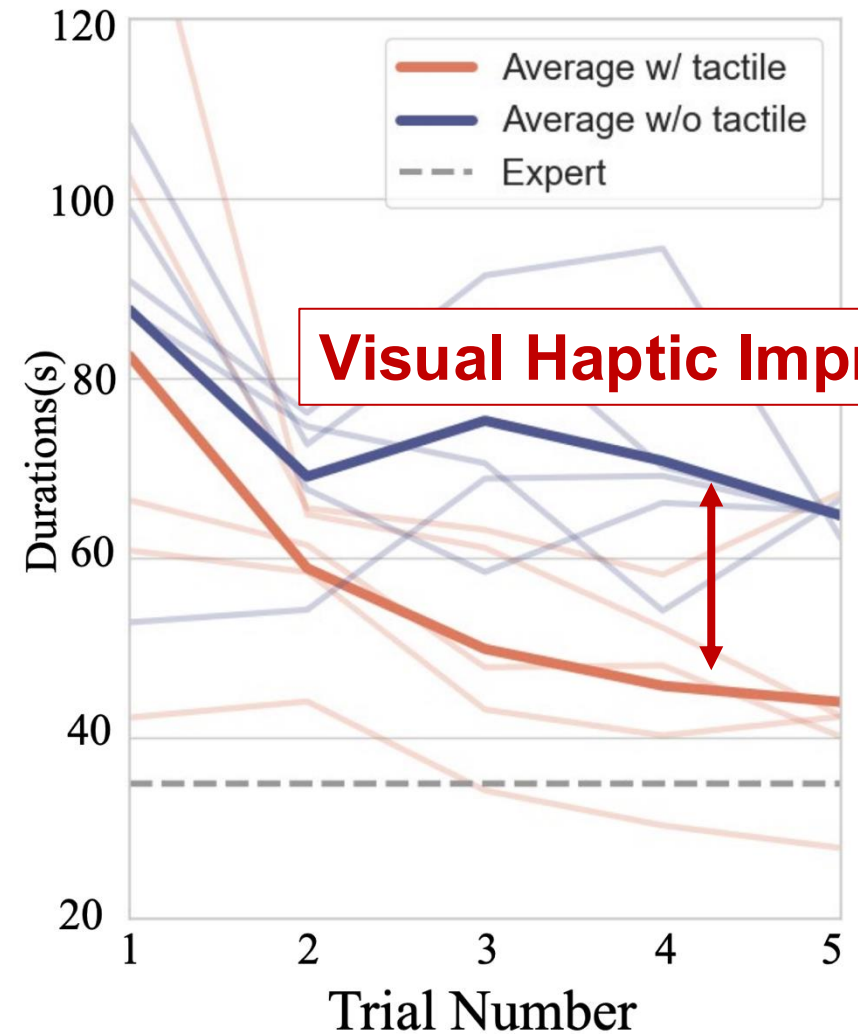
High Resolution Tactile with Multi-View RGBD Camera

How We Build Multi-Sensory System?

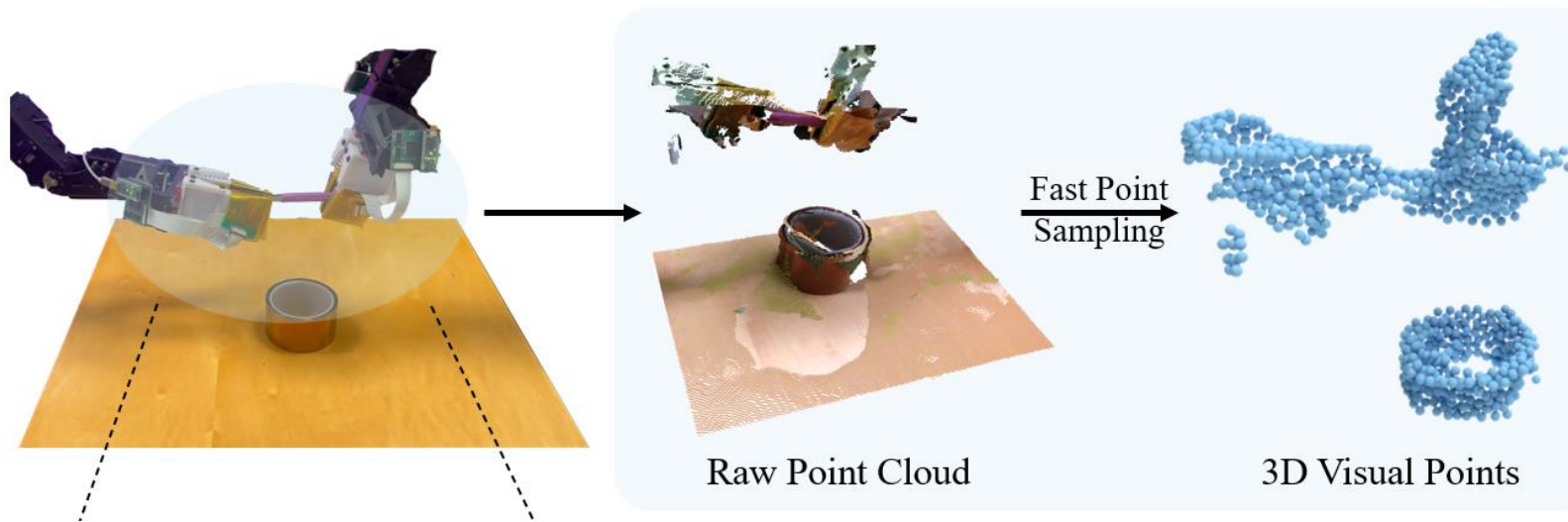


High Resolution Tactile with Multi-View RGBD Camera

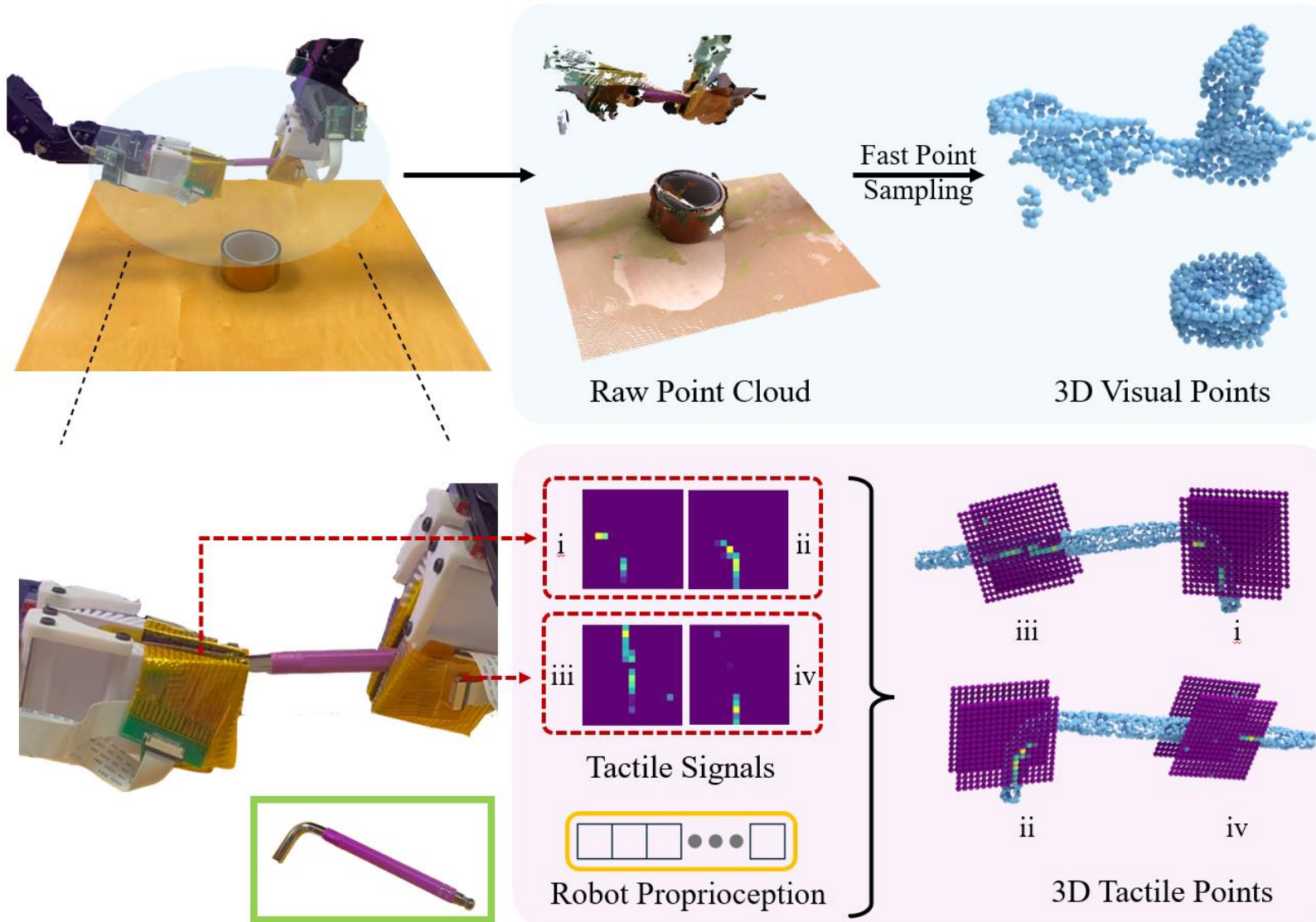
Visual Haptic Feedback



Learning with Visuo-Tactile Policy



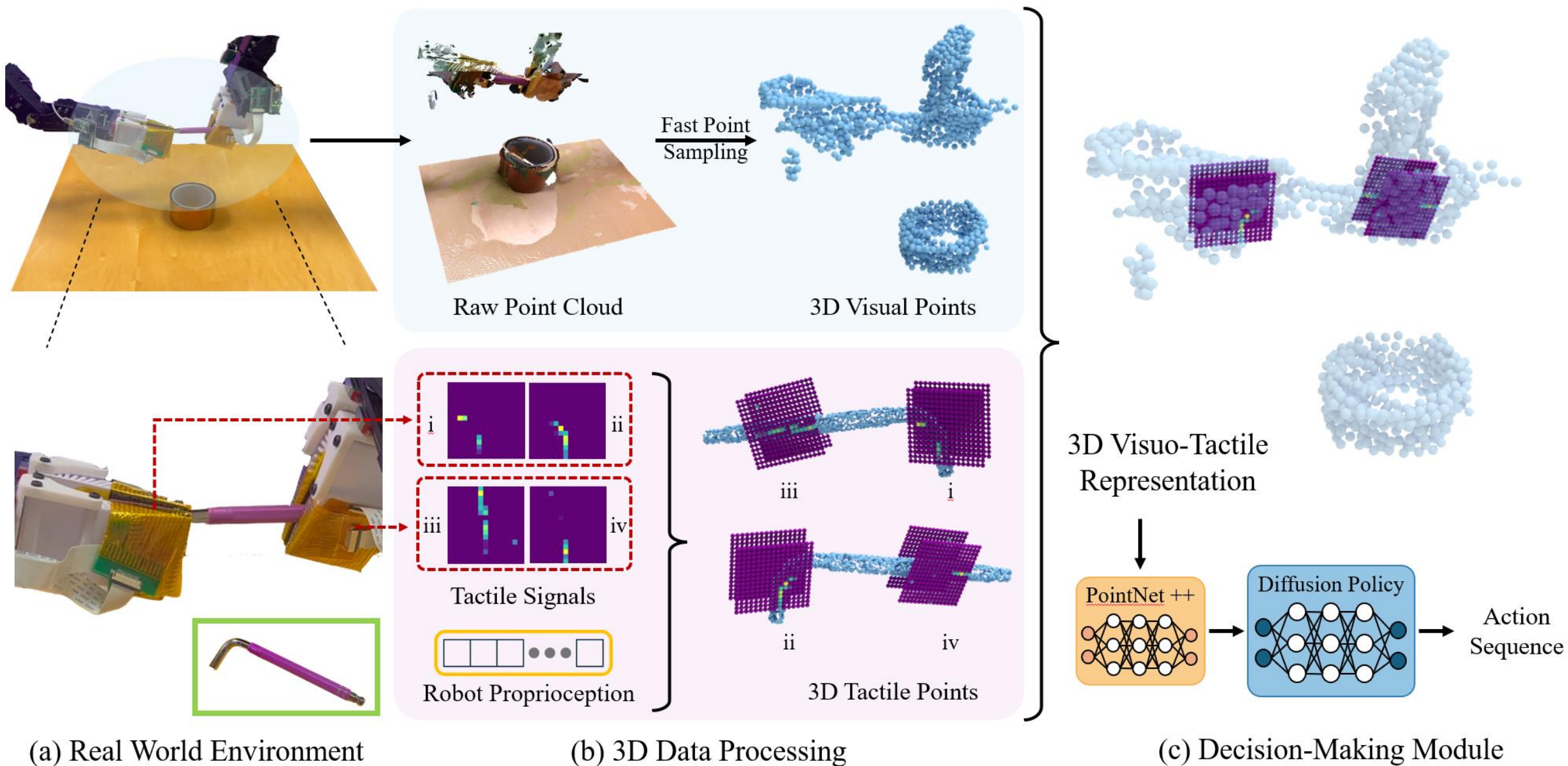
Learning with Visuo-Tactile Policy



(a) Real World Environment

(b) 3D Data Processing

Learning with Visuo-Tactile Policy



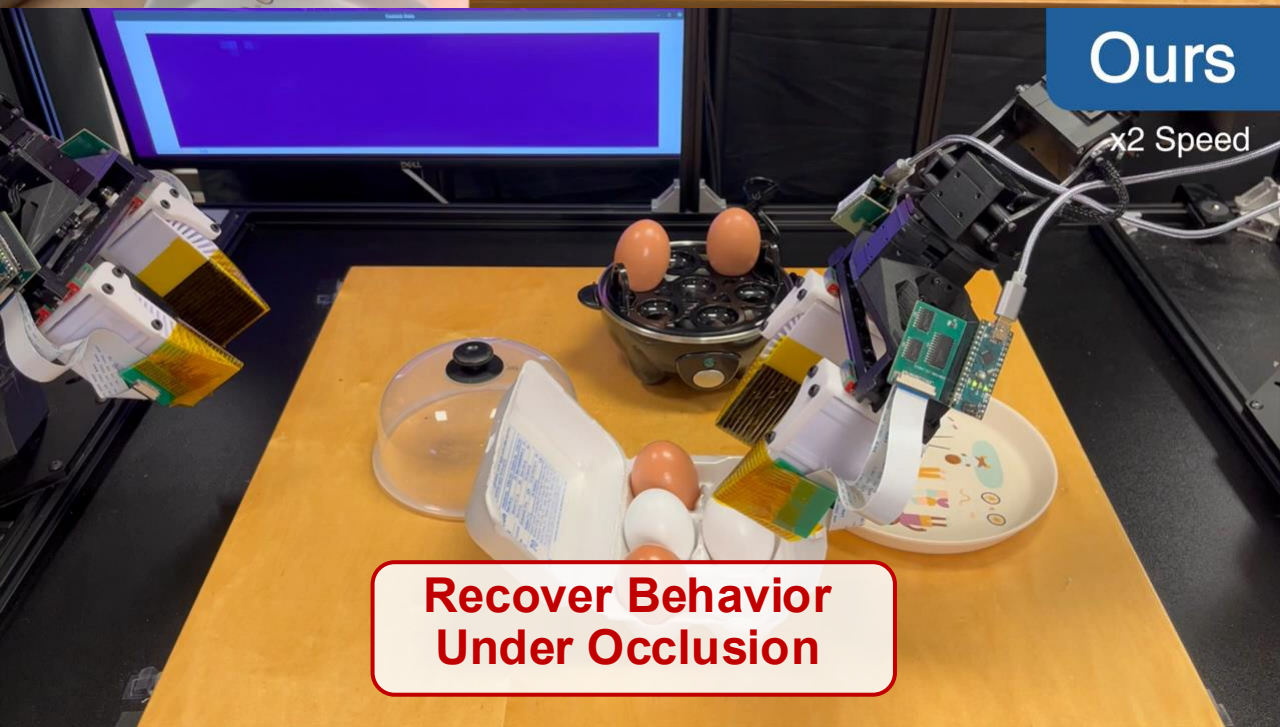


**Occluded Environments
with Fragile Objects**



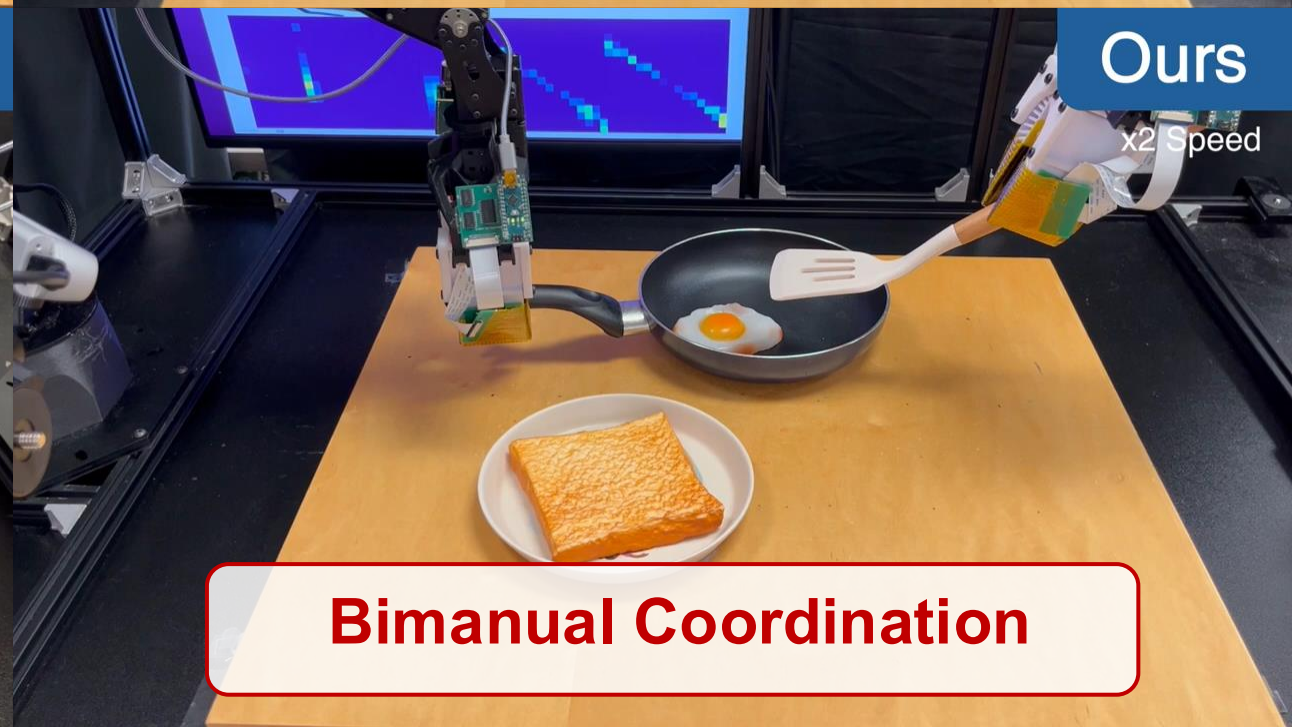
Ours
x2 Speed

In-Hand State Information



Ours
x2 Speed

**Recover Behavior
Under Occlusion**

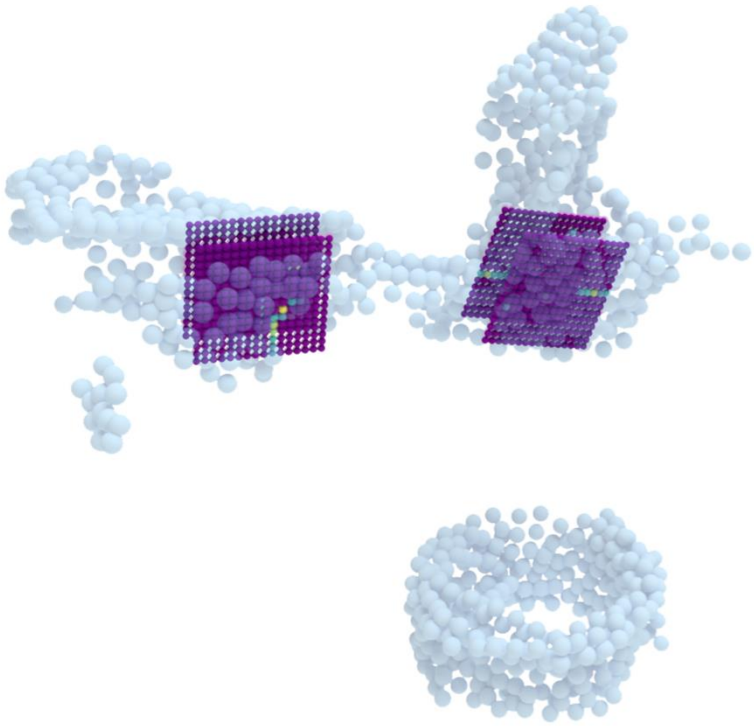


Ours
x2 Speed

Bimanual Coordination

Visuo-Tactile Fusion Is Crucial to the Policy

Two Different Approach to Fuse



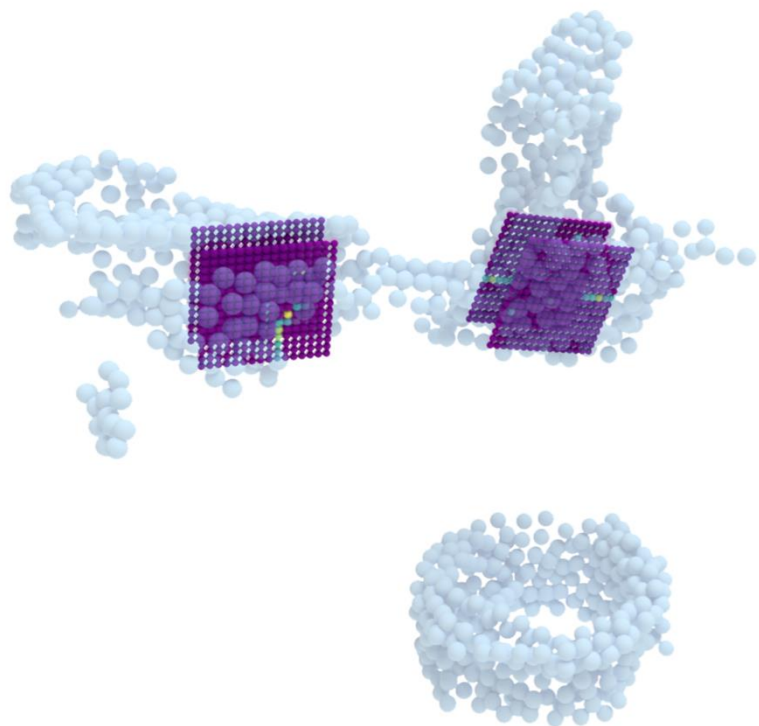
3D Tactile Fusion^[1] (before policy)

[1] Huang et al., “3D-ViTac: Learning Fine-Grained Manipulation with Visuo-Tactile Sensing,” CoRL 2024.

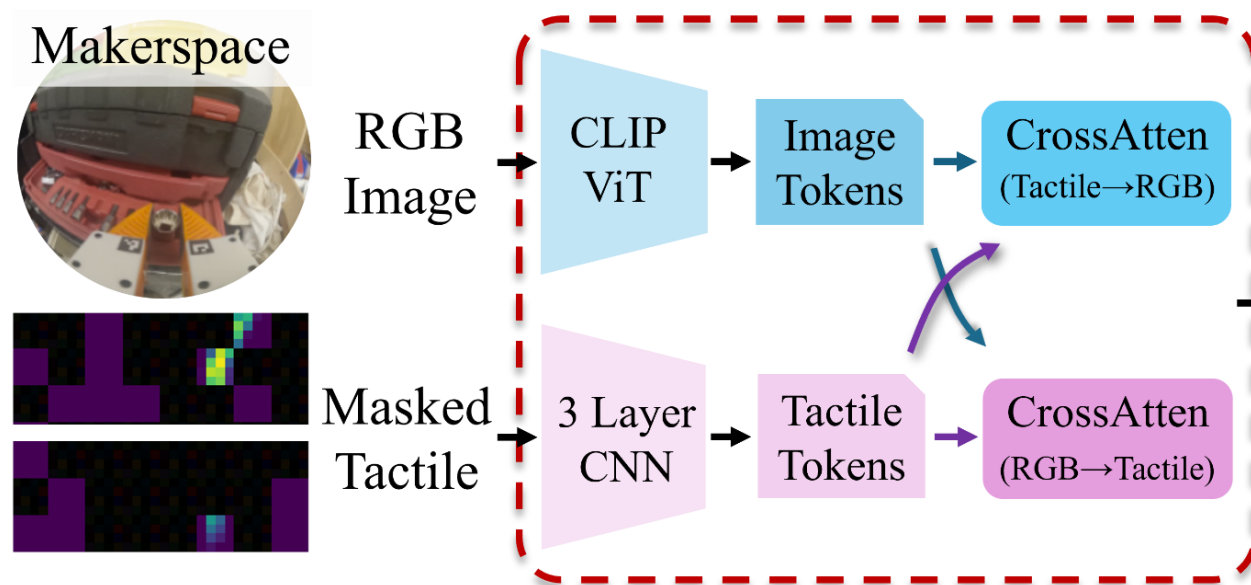
[2] Zhu et al., “Touch in the Wild: Learning Fine-Grained Manipulation with a Portable Visuo-Tactile Gripper,” NeurIPS 2025.

Visuo-Tactile Fusion Is Crucial to the Policy

Two Different Approach to Fuse



3D Tactile Fusion^[1] (before policy)



Tactile Fusion Encoder^[2] (cross-attention)

[1] Huang et al., "3D-ViTac: Learning Fine-Grained Manipulation with Visuo-Tactile Sensing," CoRL 2024.

[2] Zhu et al., "Touch in the Wild: Learning Fine-Grained Manipulation with a Portable Visuo-Tactile Gripper," NeurIPS 2025.

Touch in the Wild

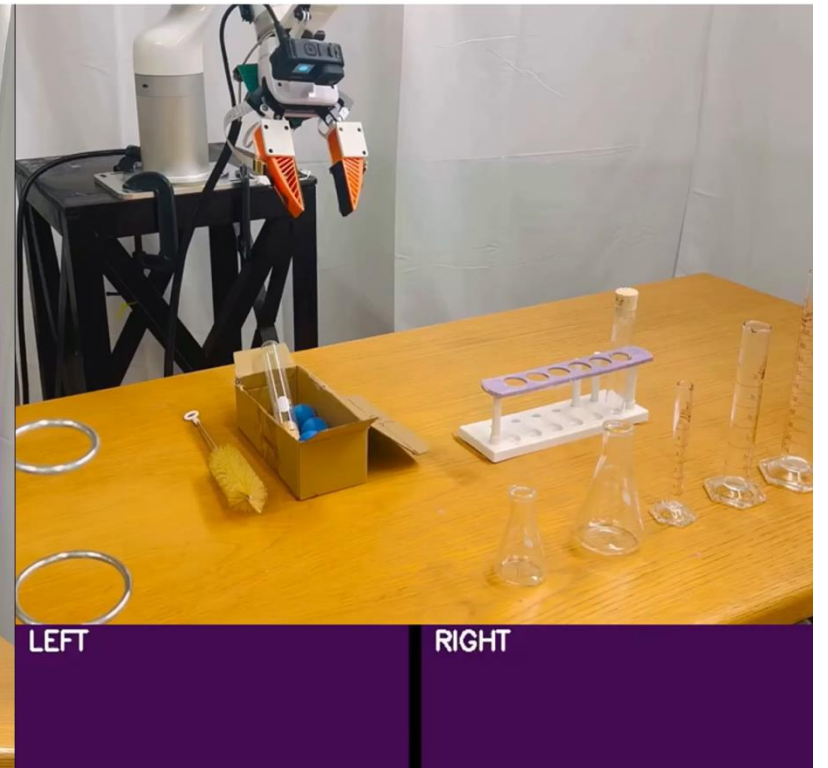
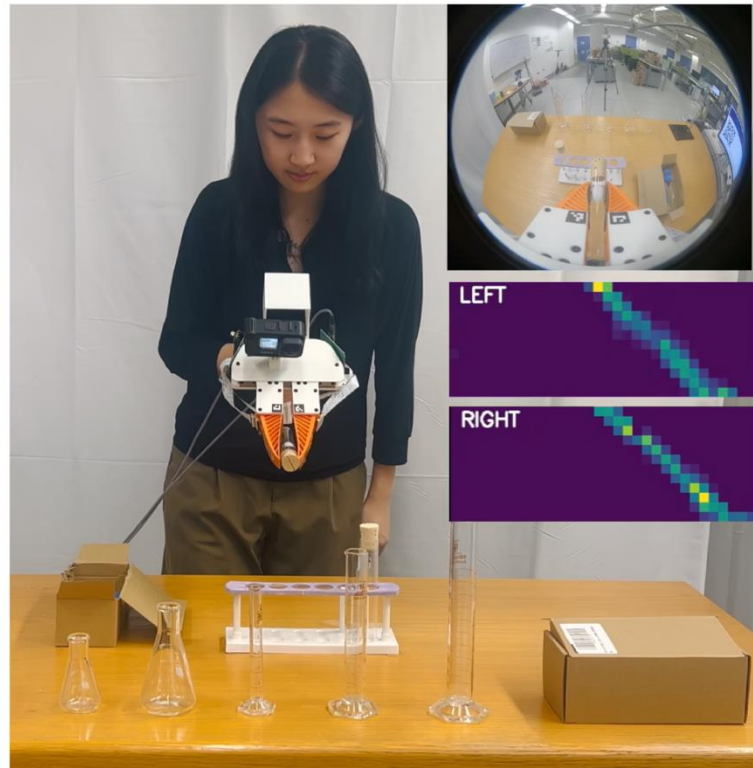
Learning Fine-Grained Manipulation with a Portable Visuo-Tactile Gripper

Xinyue Zhu*, Binghao Huang*, Yunzhu Li

NeurIPS 2025

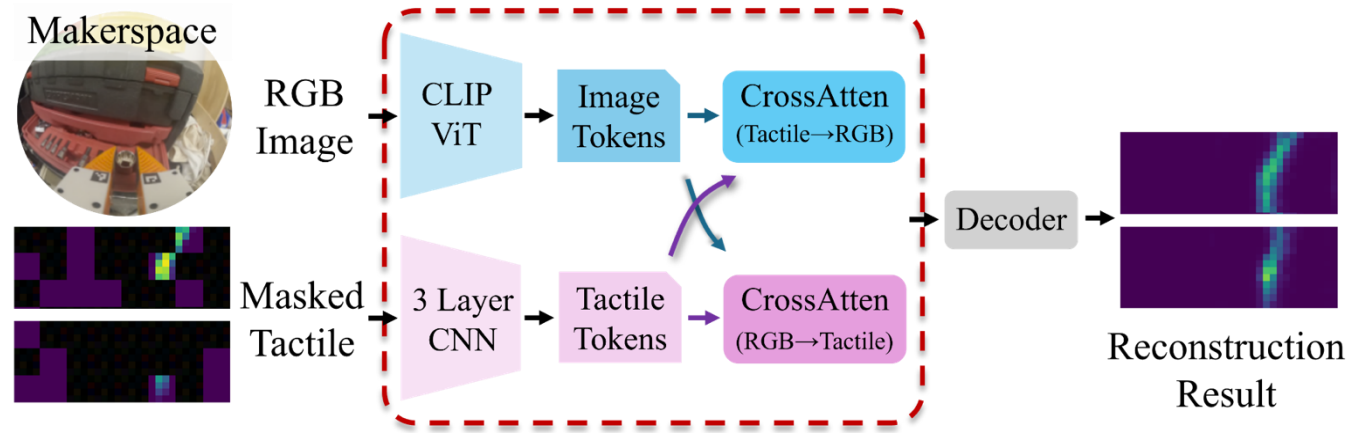
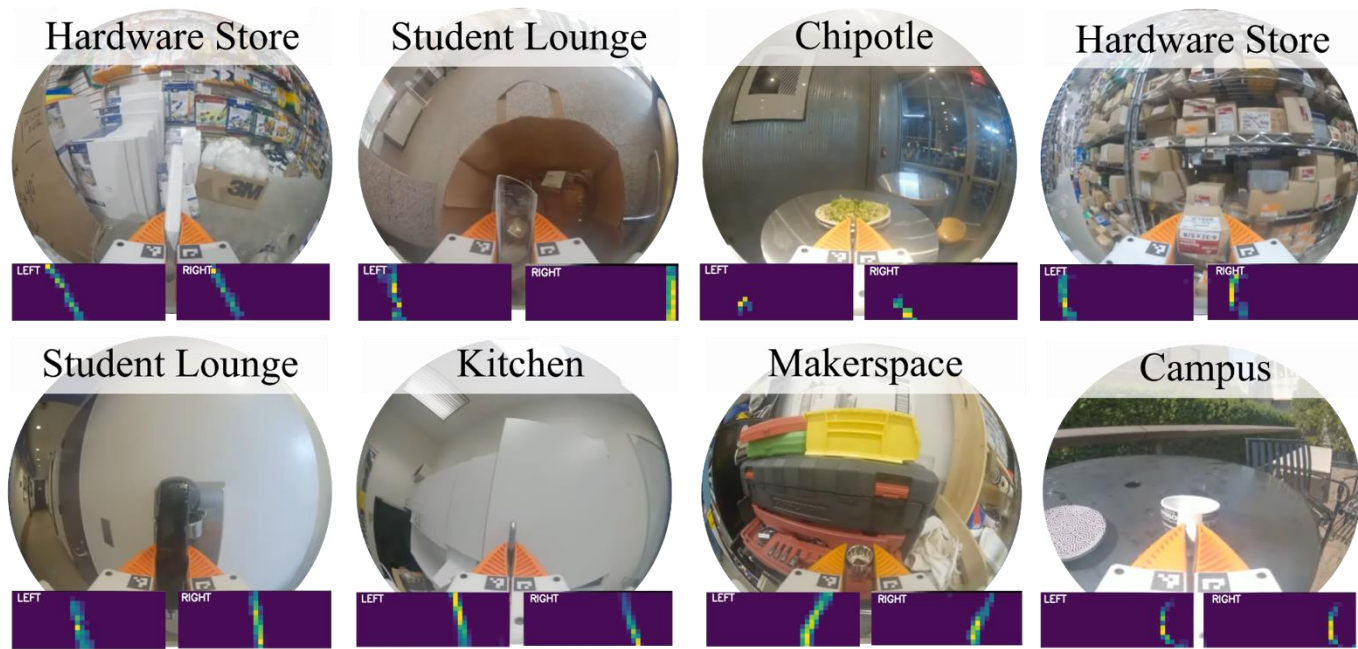
Best Demo Award at RSS 2025 Workshop on Robot Hardware-Aware Intelligence

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



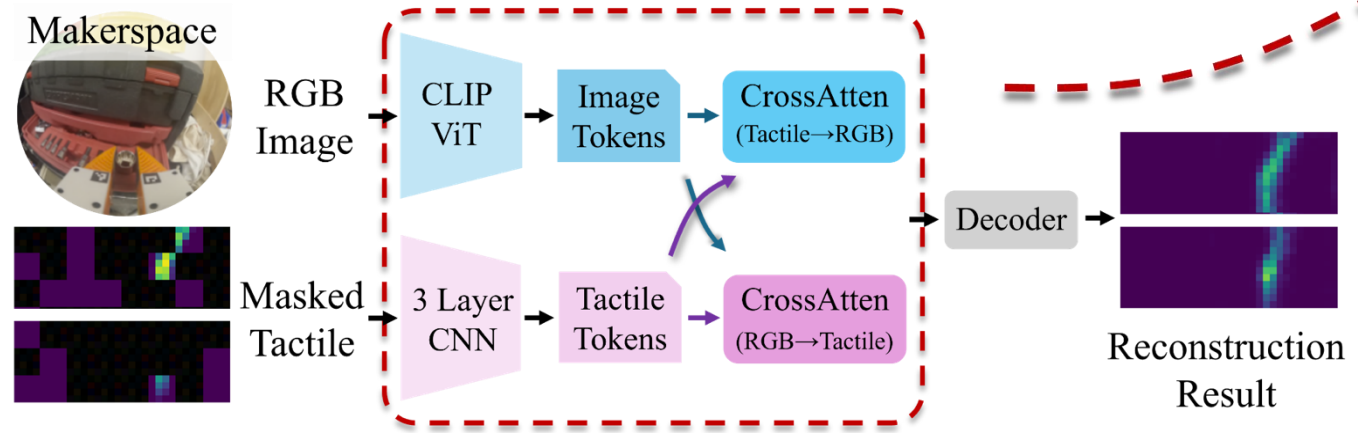
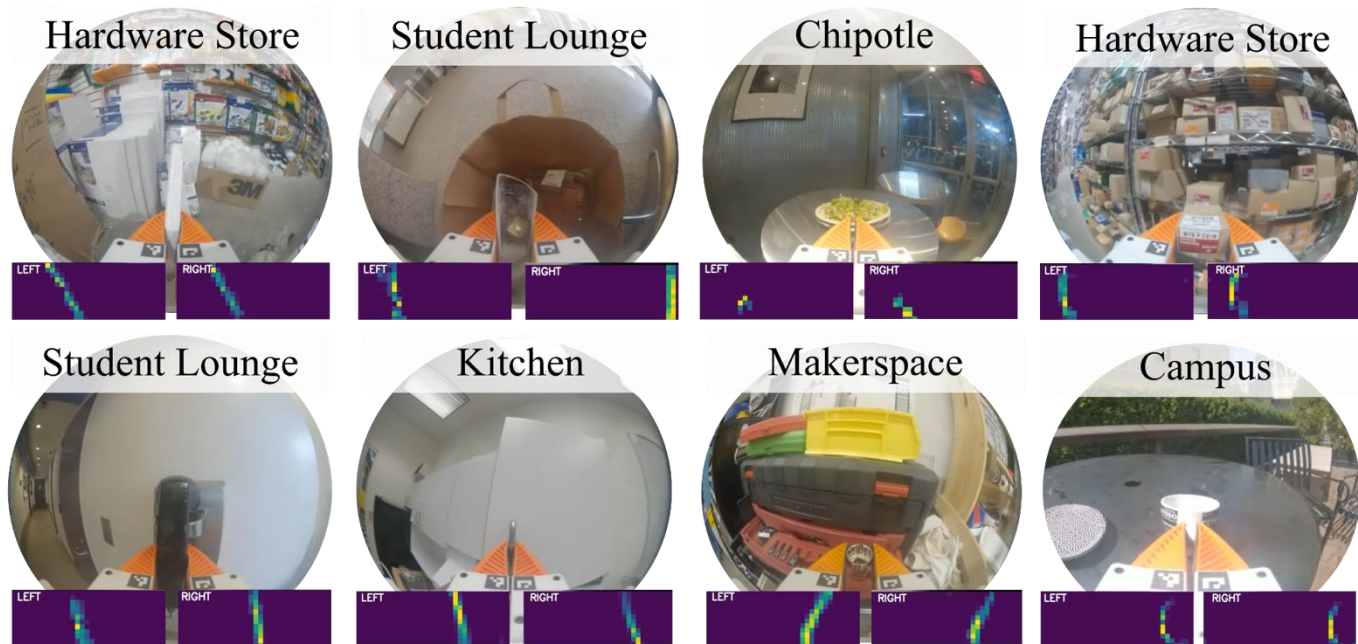
Visuo-Tactile Pre-Training & Downstream Learning

Pre-Train



Visuo-Tactile Pre-Training & Downstream Learning

Pre-Train



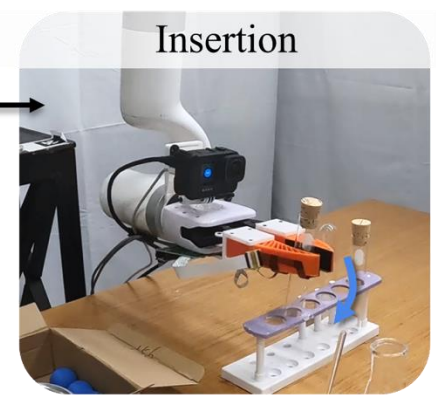
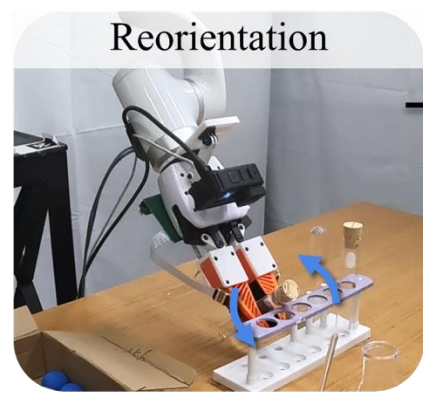
Downstream Task



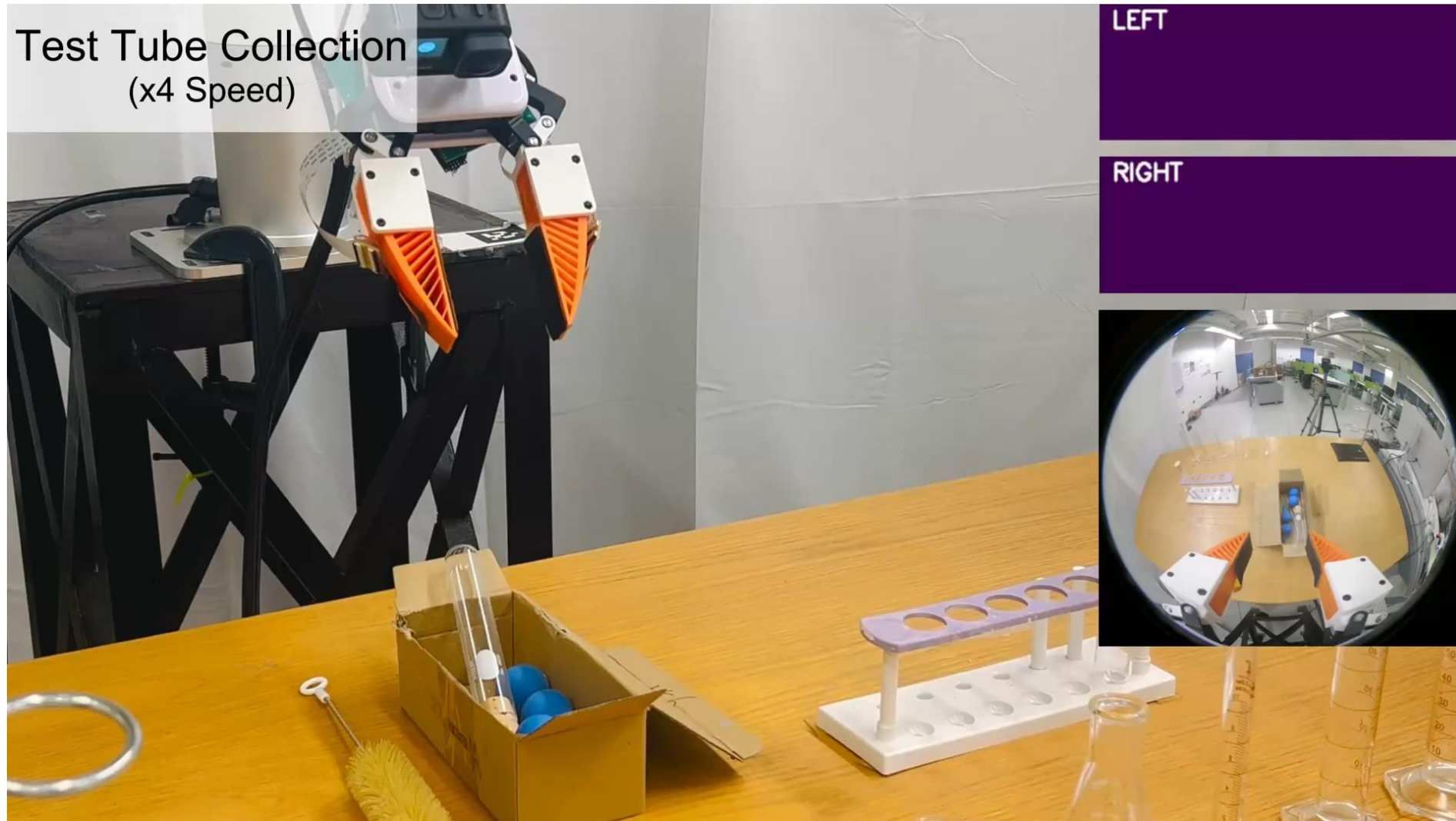
Visuo-Tactile Encoder

Robot Proprioception
Relative EE Pose,
Gripper Width

Diffusion Policy

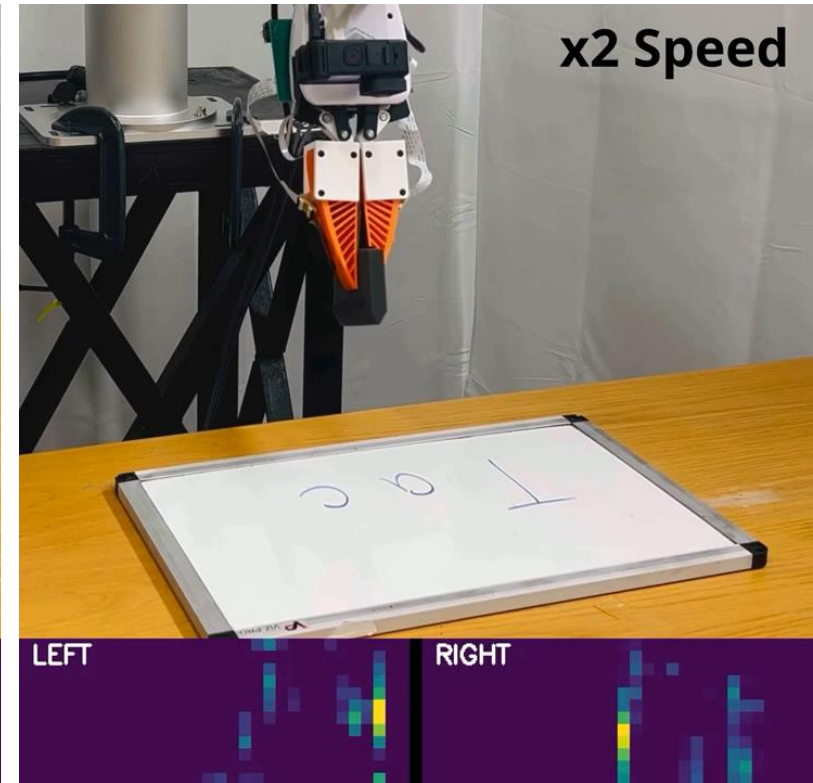
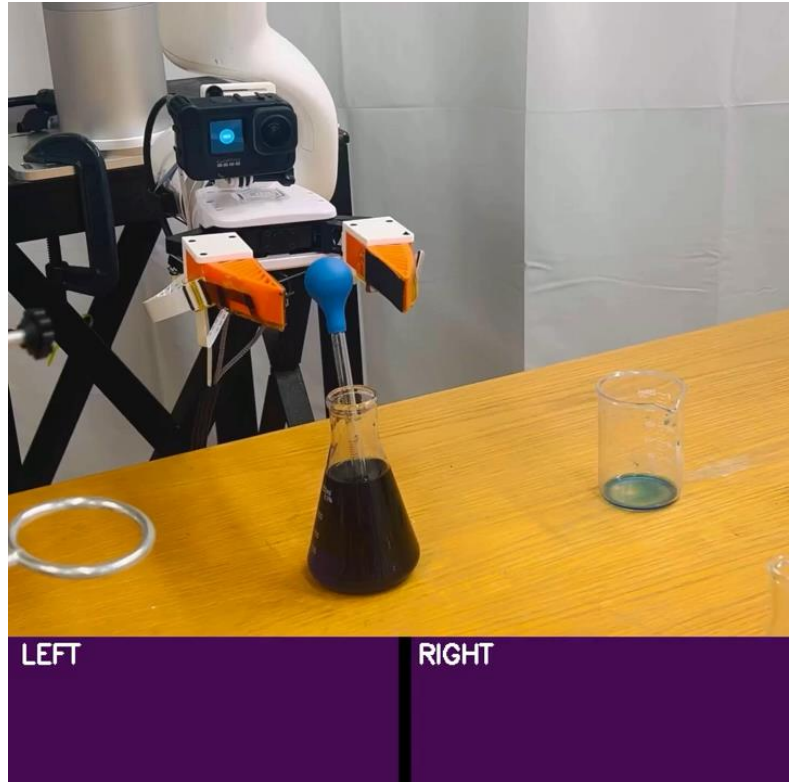
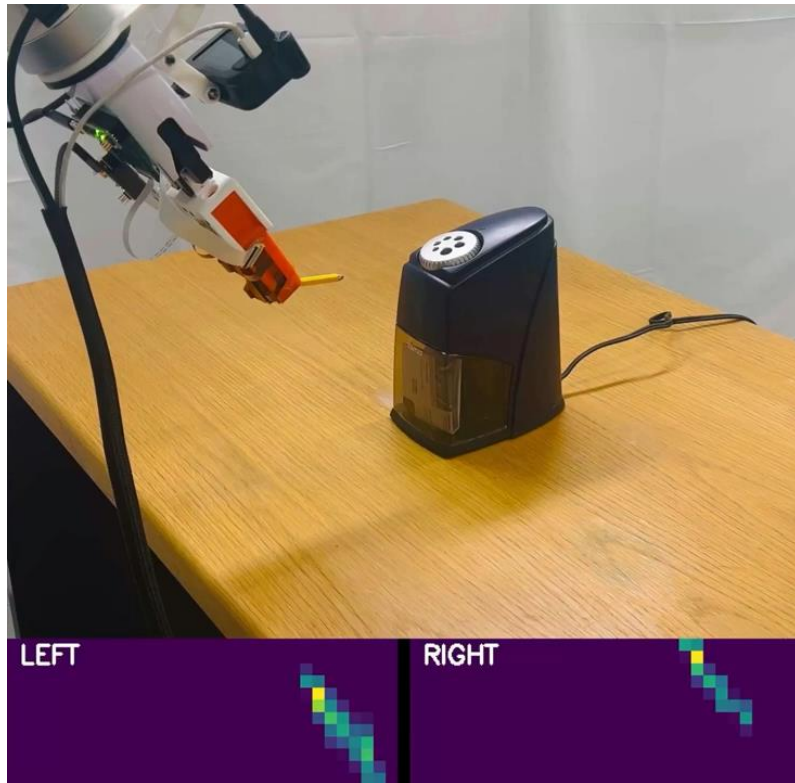


How well does the downstream policy work?



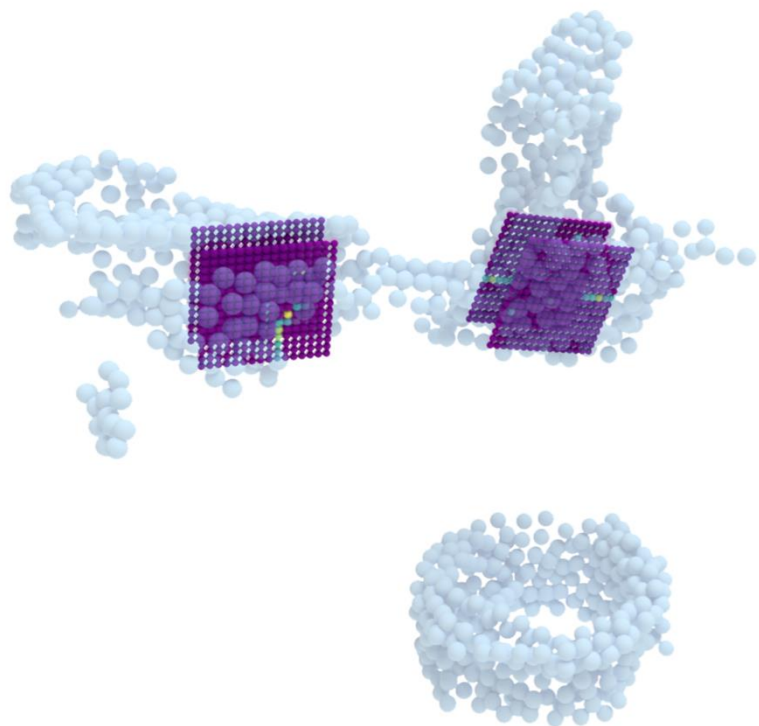
Xinyue Zhu, Binghao Huang, and Yunzhu Li, "Touch in the Wild: Learning Fine-Grained Manipulation with a Portable Visuo-Tactile Gripper," NeurIPS 2025. **Best Demo Award**, RSS 2025 Workshop on Robot Hardware-Aware Intelligence

How well does the downstream policy work?

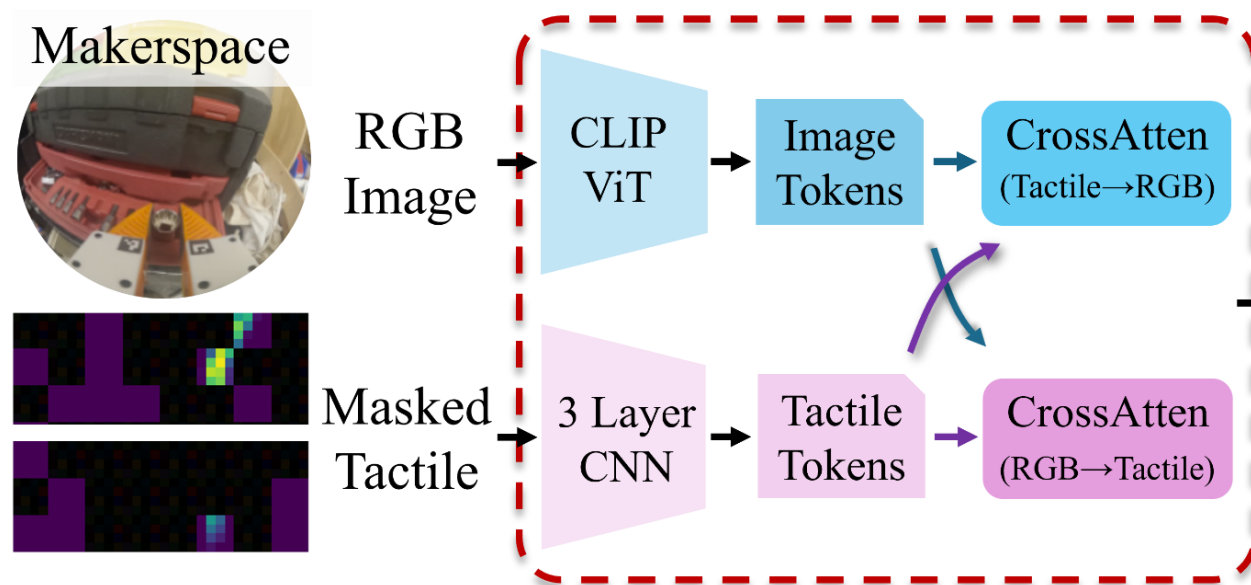


Visuo-Tactile Fusion Is Crucial to the Policy

Two Different Approach to Fuse



3D Tactile Fusion^[1] (before policy)



Tactile Fusion Encoder^[2] (cross-attention)

[1] Huang et al., “3D-ViTac: Learning Fine-Grained Manipulation with Visuo-Tactile Sensing,” CoRL 2024.

[2] Zhu et al., “Touch in the Wild: Learning Fine-Grained Manipulation with a Portable Visuo-Tactile Gripper,” NeurIPS 2025.

Learning Visuo-Tactile Models

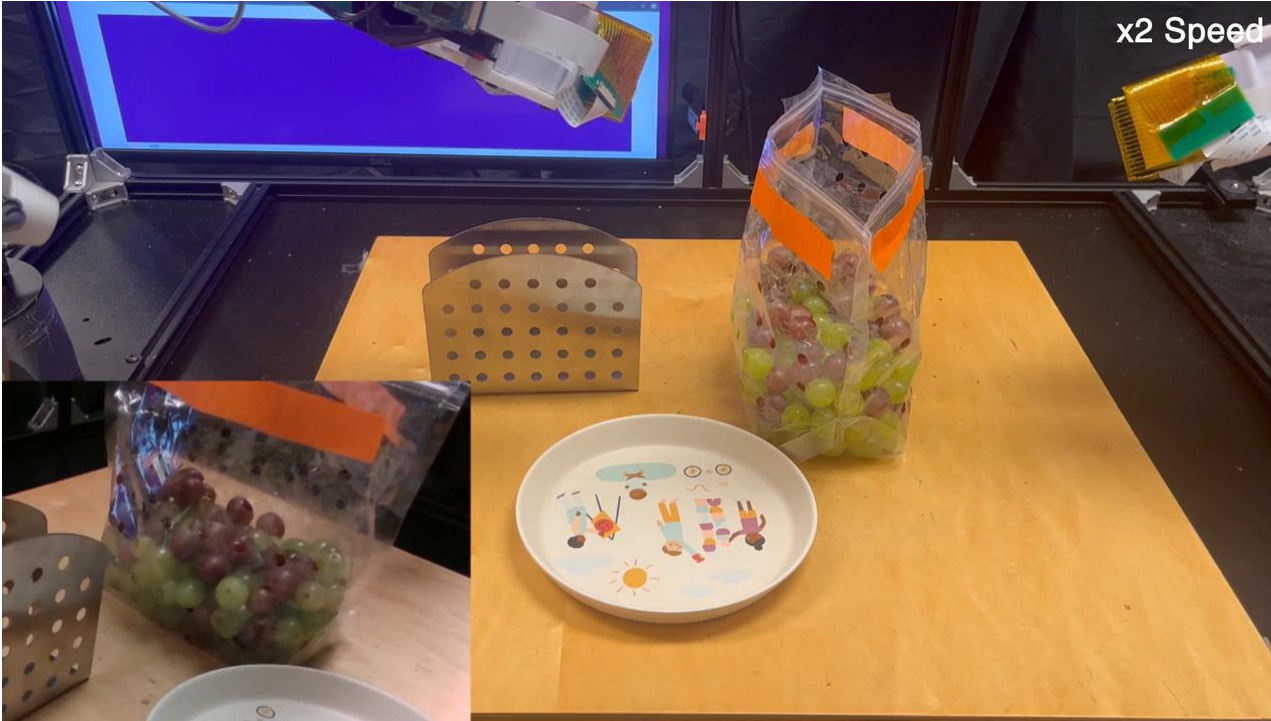


GATE 2

Useful for the task/policy?

Yes — in most cases.

Learning Visuo-Tactile Models



GATE 2

Useful for the task/policy?

Yes — in most cases.

“Most cases” isn’t every case.

Next: scale up the data.

Overview



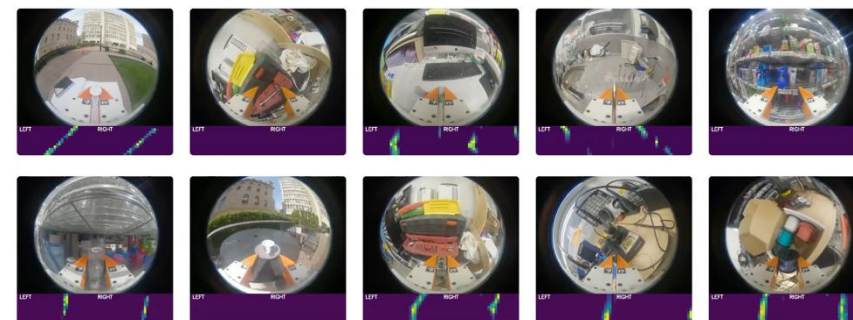
Hardware



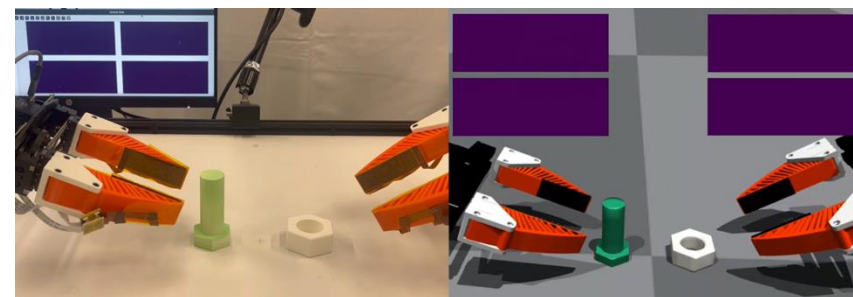
Perception



Data



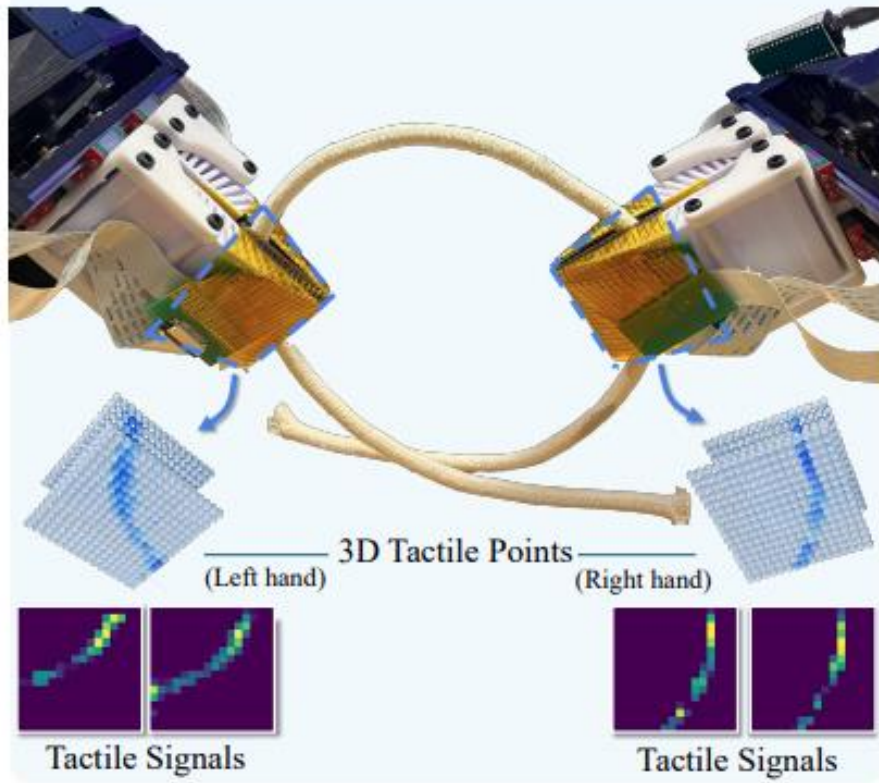
Large-Scale Tactile in Real



Large-Scale Tactile in Simulation

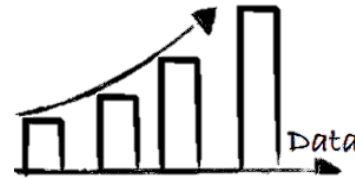
One full stack for scalable tactile manipulation
— build it, perceive with it, learn at scale.

Scaling Up Tactile Data

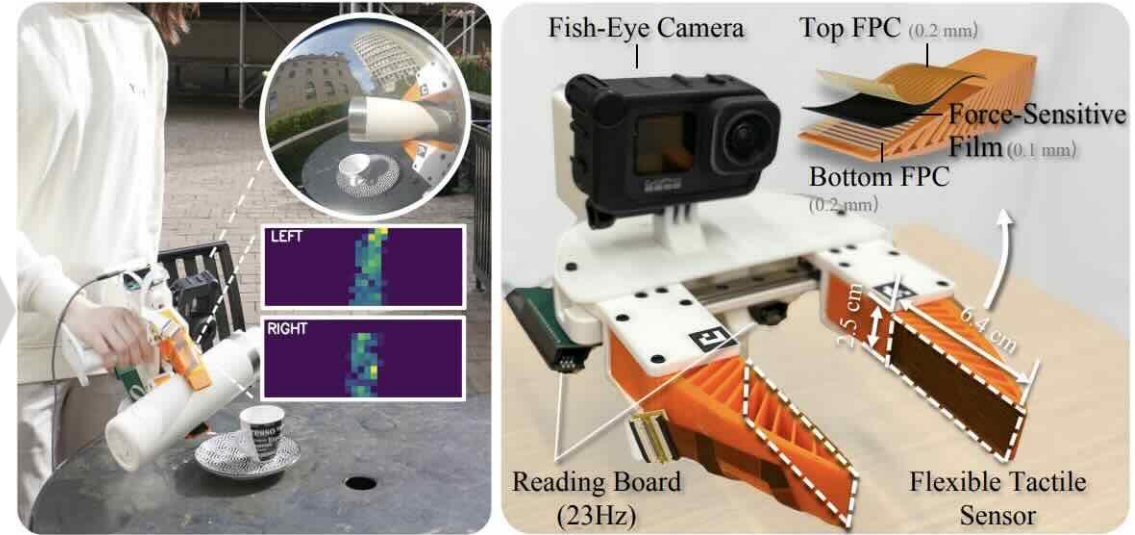


[1] Huang et al., “3D-ViTac: Learning Fine-Grained Manipulation with Visuo-Tactile Sensing,” CoRL 2024.

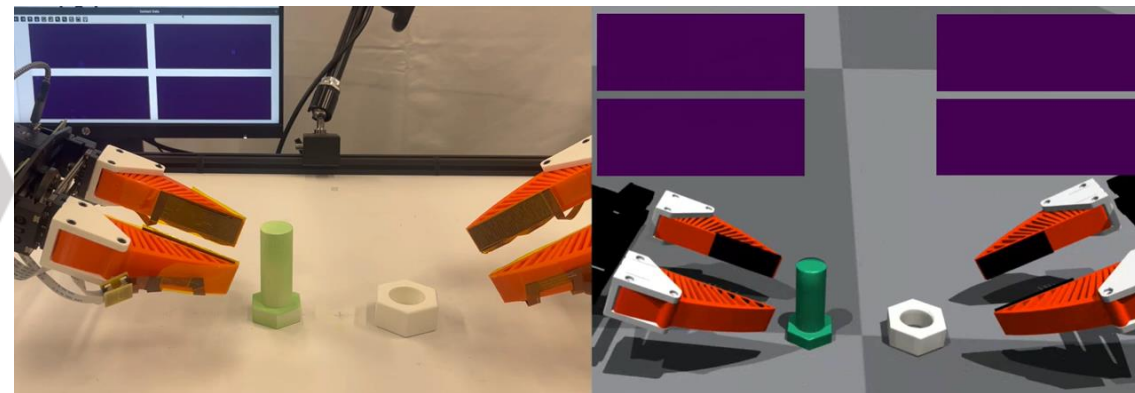
Scale in Real



Scale in Sim

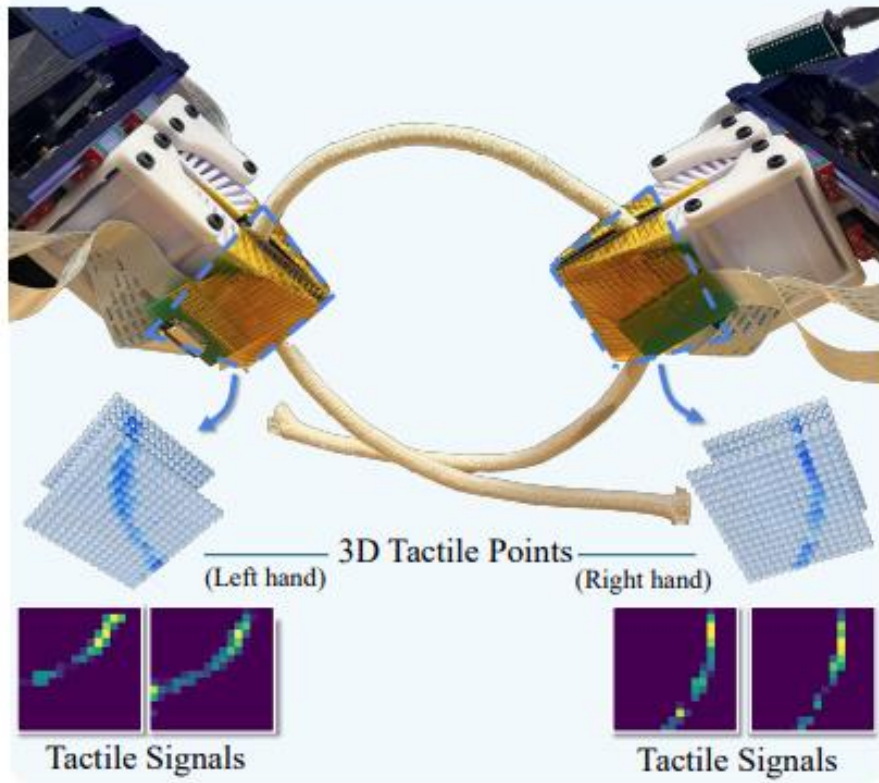


[2] Zhu et al., “Touch in the Wild: Learning Fine-Grained Manipulation with a Portable Visuo-Tactile Gripper,” NeurIPS 2025.



[3] Huang et al., “VT-Refine: Learning Bimanual Assembly with Visuo-Tactile Feedback via Simulation Fine-Tuning,” CoRL 2025.

Scaling Up Tactile Data



[1] Huang et al., "3D-ViTac: Learning Fine-Grained Manipulation with Visuo-Tactile Sensing," CoRL 2024.

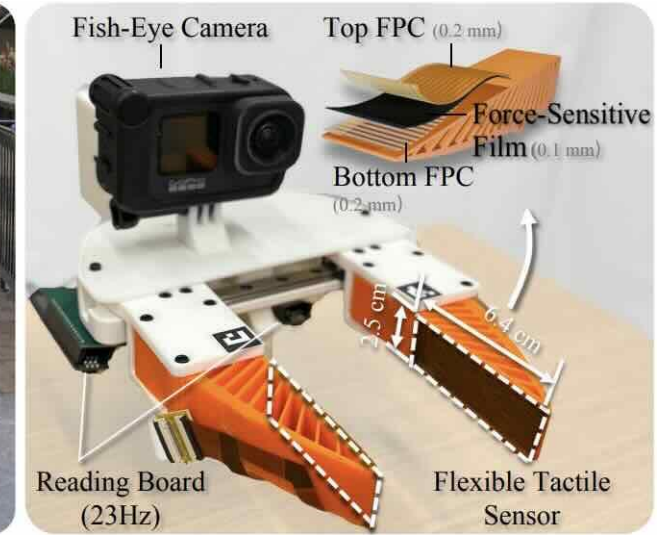
Scale in Real



Scale in Sim



[2] Zhu et al., "Touch in the Wild: Learning Fine-Grained Manipulation with a Portable Visuo-Tactile Gripper," NeurIPS 2025.



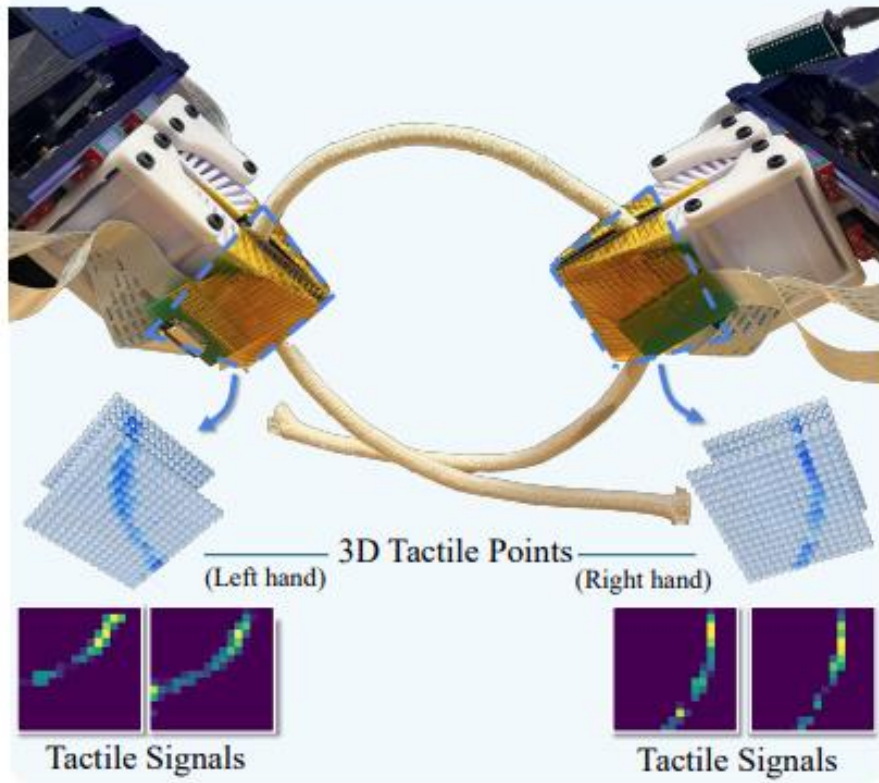
[3] Huang et al., "VT-Refine: Learning Bimanual Assembly with Visuo-Tactile Feedback via Simulation Fine-Tuning," CoRL 2025.

Step 1: Prepare Gripper





Scaling Up Tactile Data



[1] Huang et al., "3D-ViTac: Learning Fine-Grained Manipulation with Visuo-Tactile Sensing," CoRL 2024.

Scale in Real



Scale in Sim

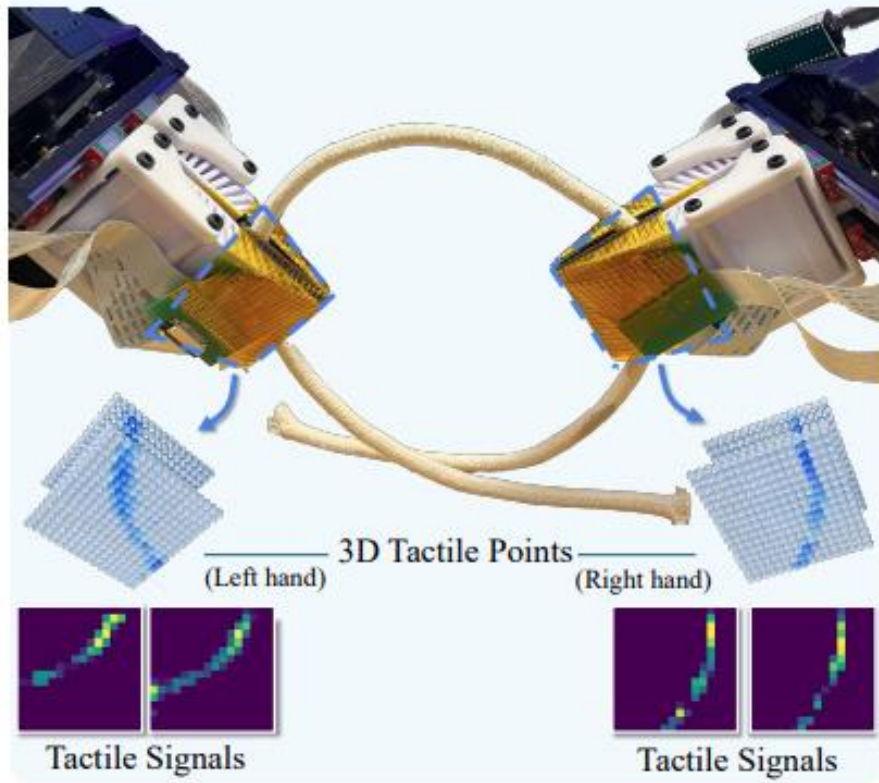


[2] Zhu et al., "Touch in the Wild: Learning Fine-Grained Manipulation with a Portable Visuo-Tactile Gripper," NeurIPS 2025.



[3] Huang et al., "VT-Refine: Learning Bimanual Assembly with Visuo-Tactile Feedback via Simulation Fine-Tuning," CoRL 2025.

Scaling Up Tactile Data



[1] Huang et al., "3D-ViTac: Learning Fine-Grained Manipulation with Visuo-Tactile Sensing," CoRL 2024.

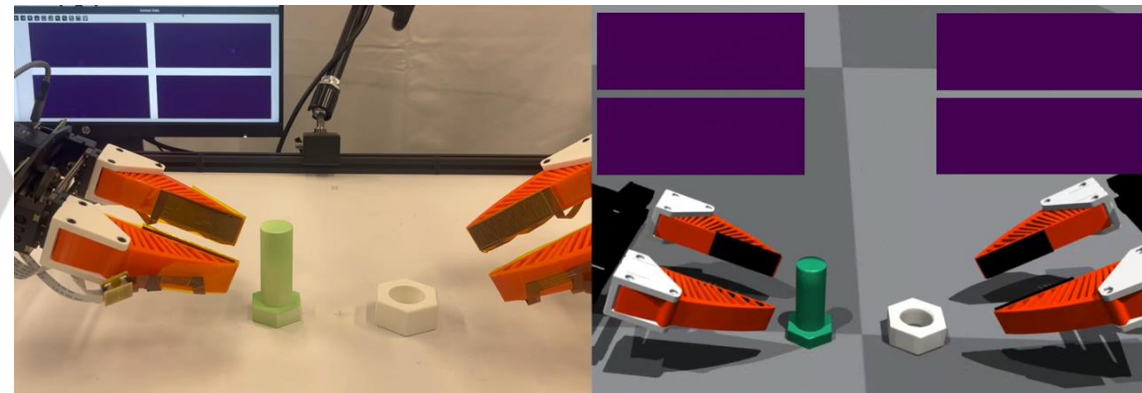
Scale in Real



Scale in Sim



[2] Zhu et al., "Touch in the Wild: Learning Fine-Grained Manipulation with a Portable Visuo-Tactile Gripper," NeurIPS 2025.



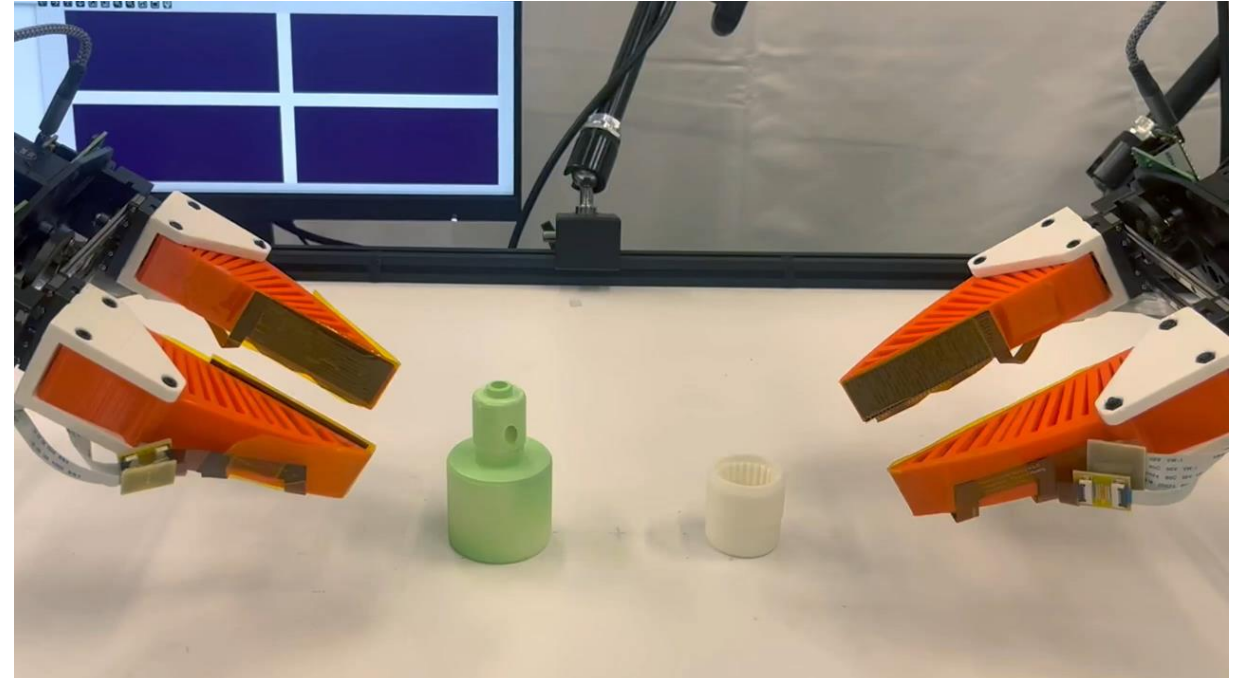
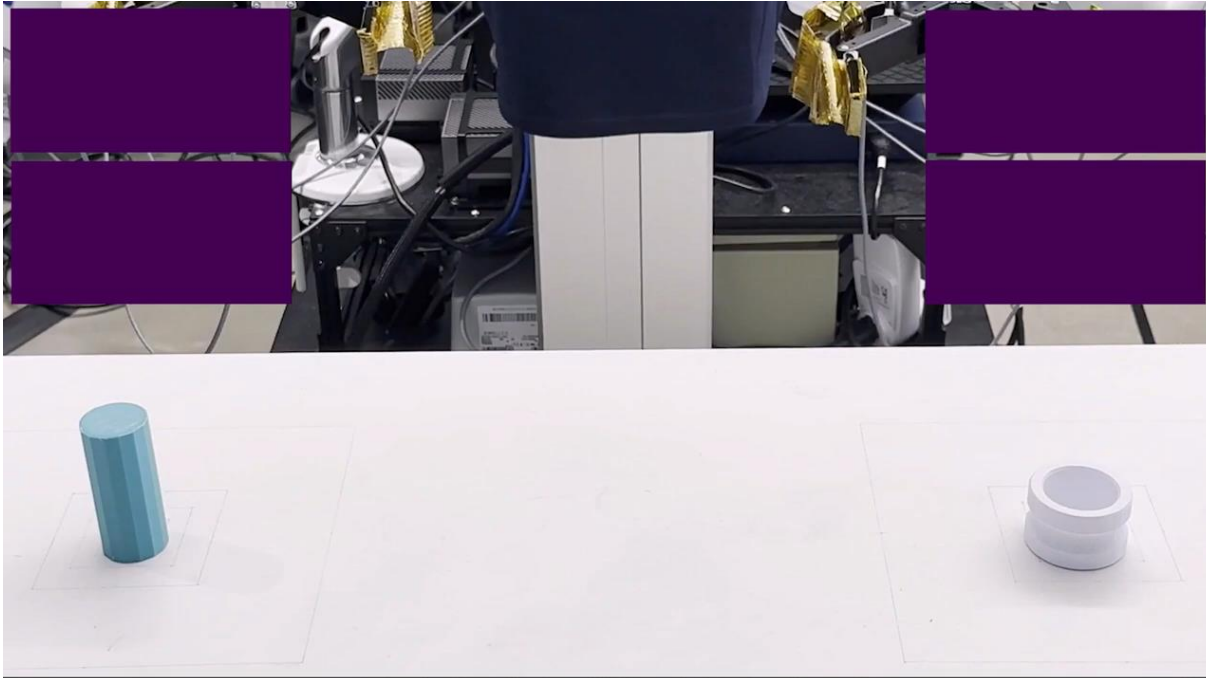
[3] Huang et al., "VT-Refine: Learning Bimanual Assembly with Visuo-Tactile Feedback via Simulation Fine-Tuning," CoRL 2025.

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

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

CoRL 2025, Best Paper Award at IROS 2025 AHFHR Workshop

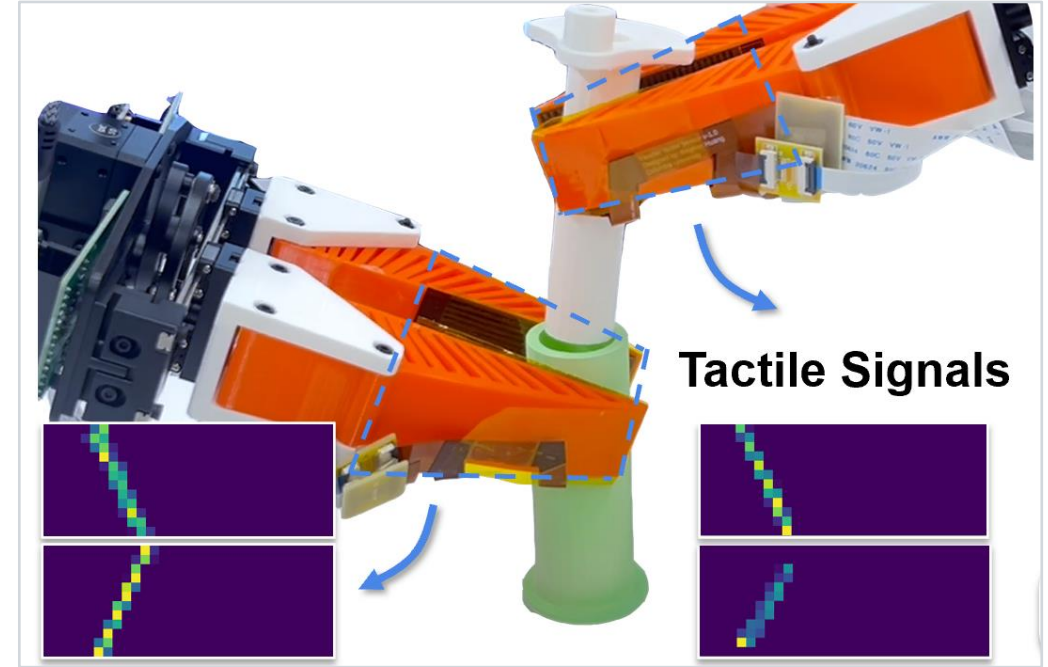
In collaboration with NVIDIA Robotics



Humans Rely on Touch for Compact Insertions



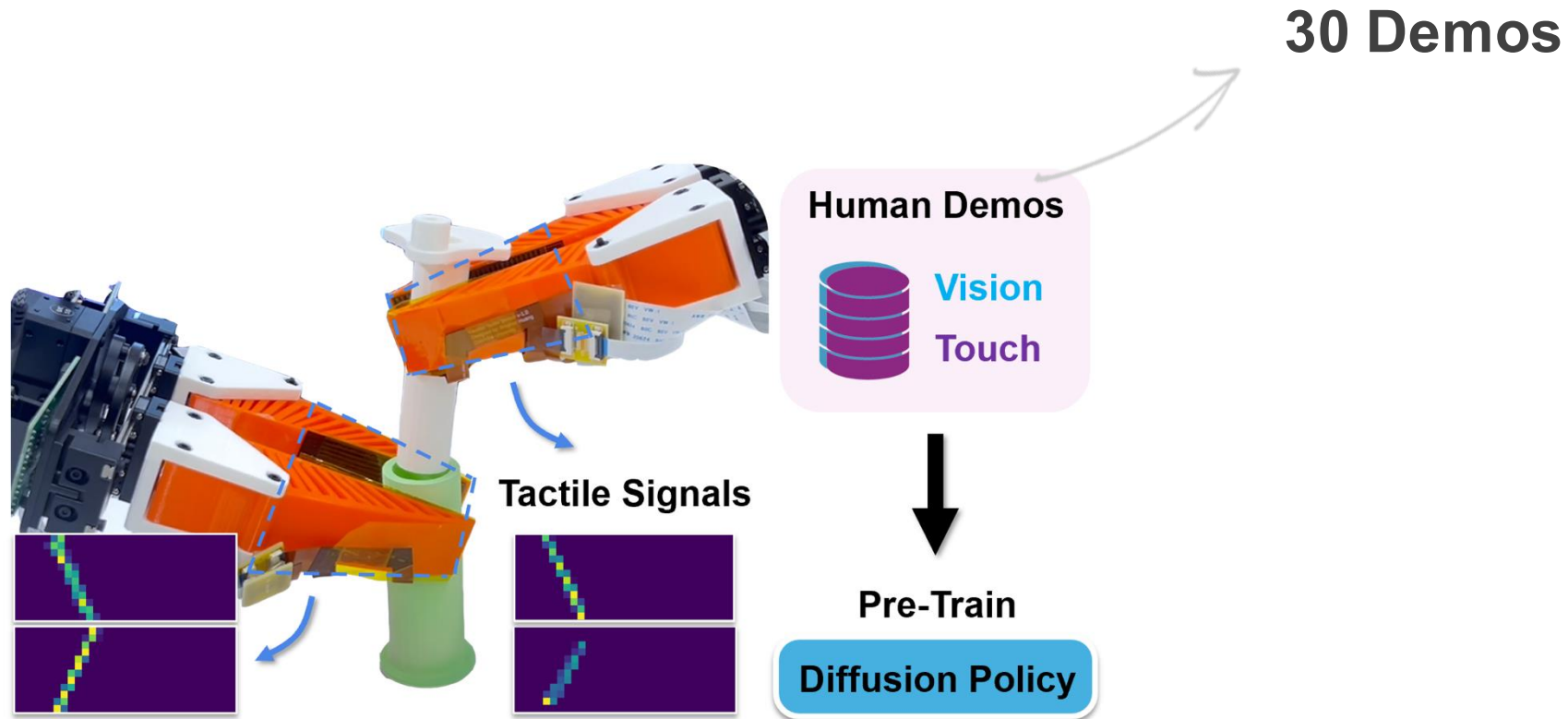
Humans: insert by feel



Robot: vision + tactile feedback

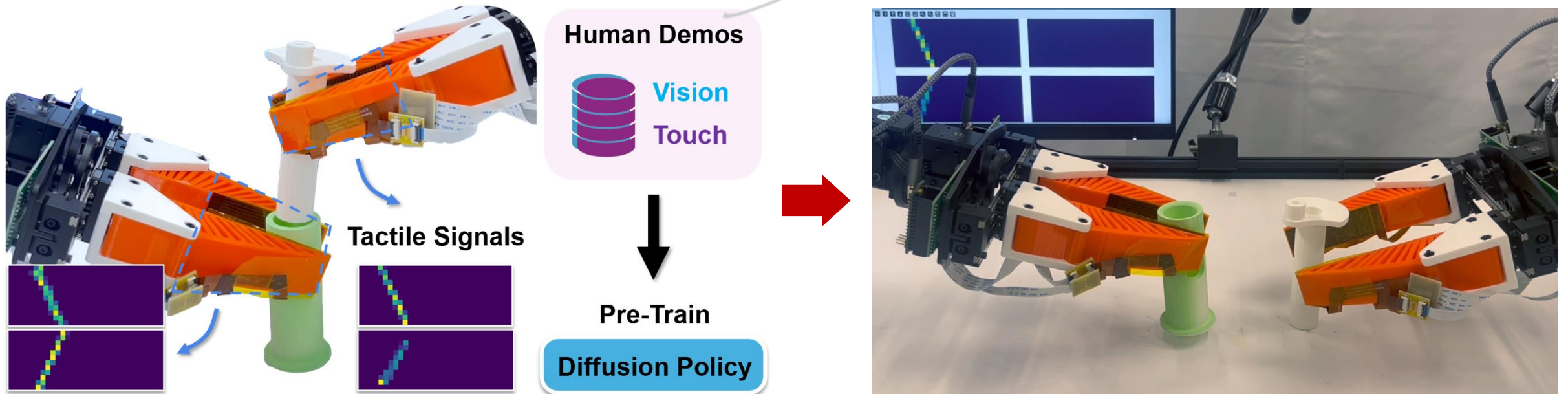
Tight insertions are contact-rich — you can't do them by looking alone.

Even with Touch, imitation learning suffers....

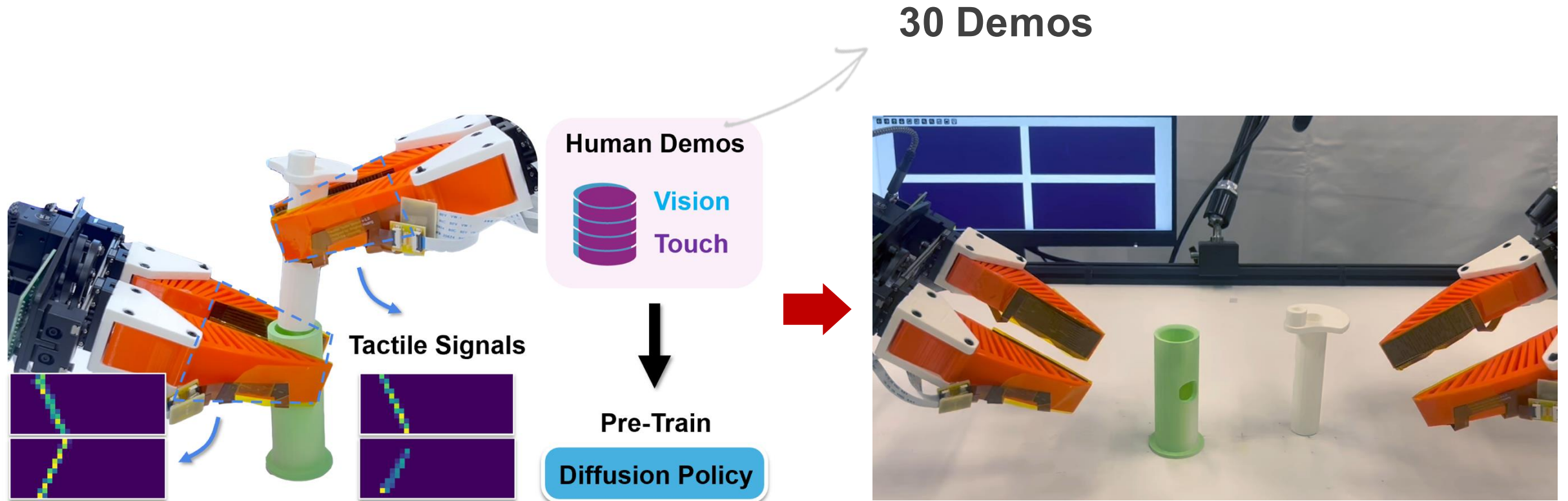


Even with Touch, imitation learning suffers....

30 Demos

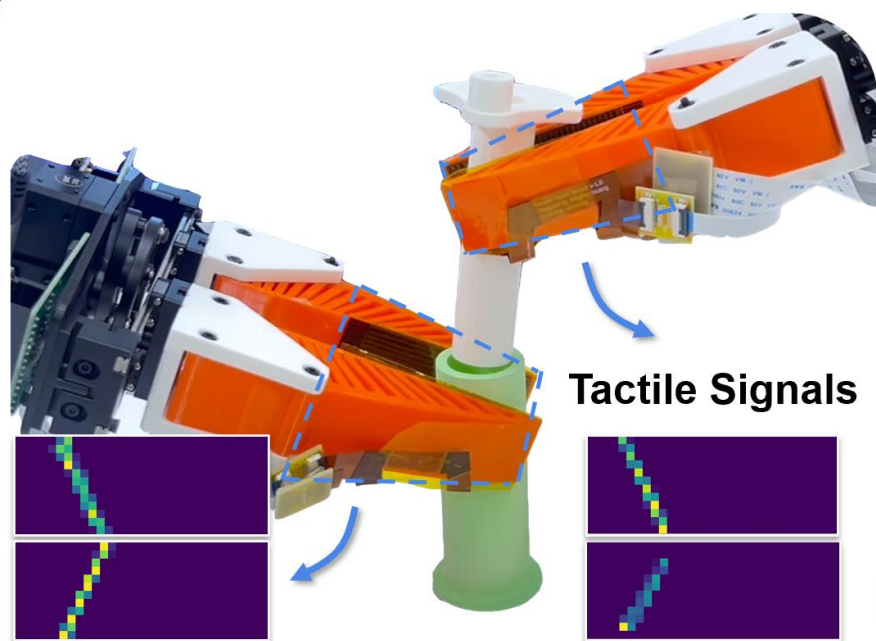


Even with Touch, imitation learning suffers....



Train on a small amount of real data? **Not enough.**

Real World



Human Demos



Vision

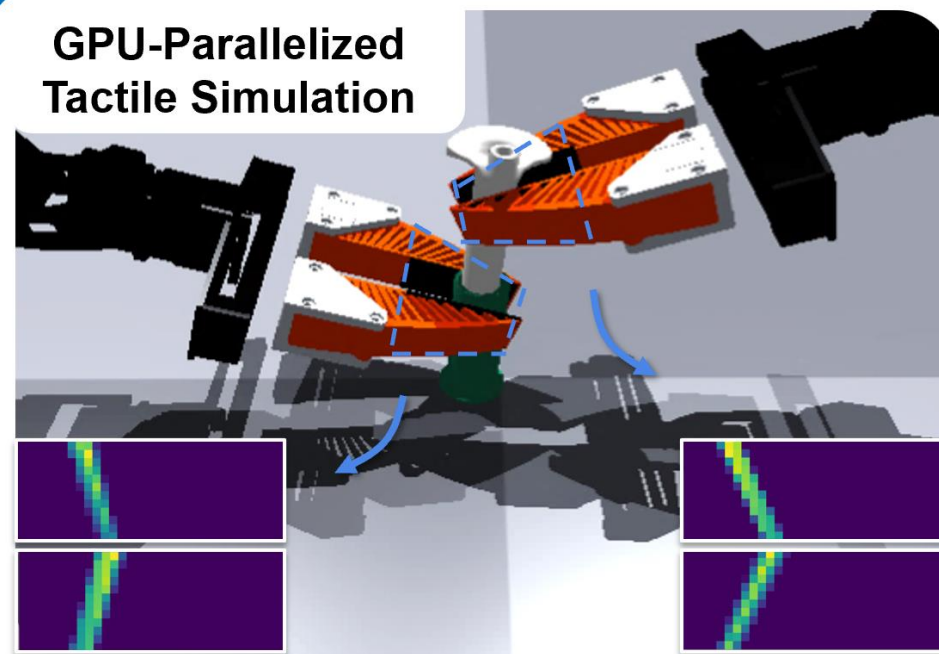
Touch

Pre-Train

Diffusion Policy

Simulation

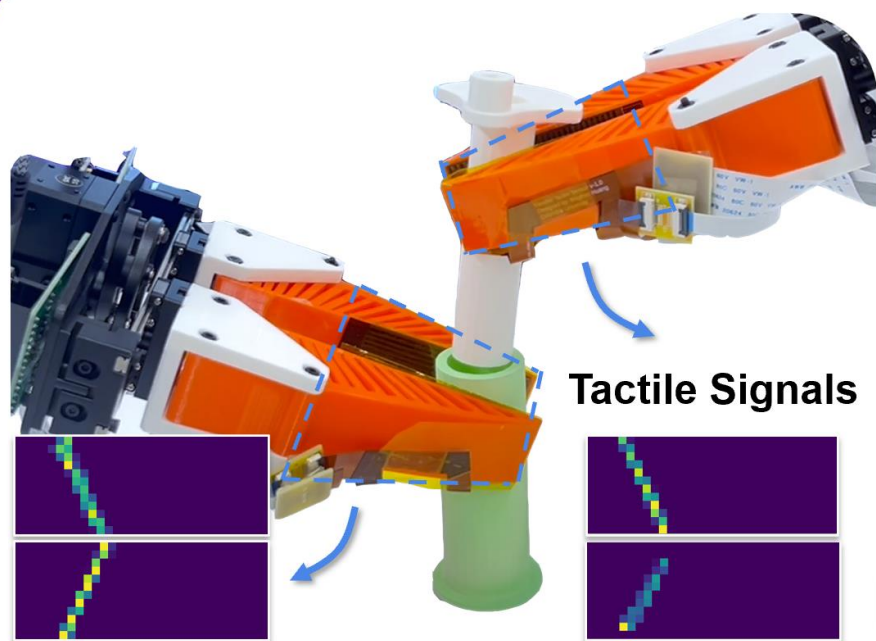
GPU-Parallelized
Tactile Simulation



Large-Scale RL Fine-Tuning



Real World



Human Demos

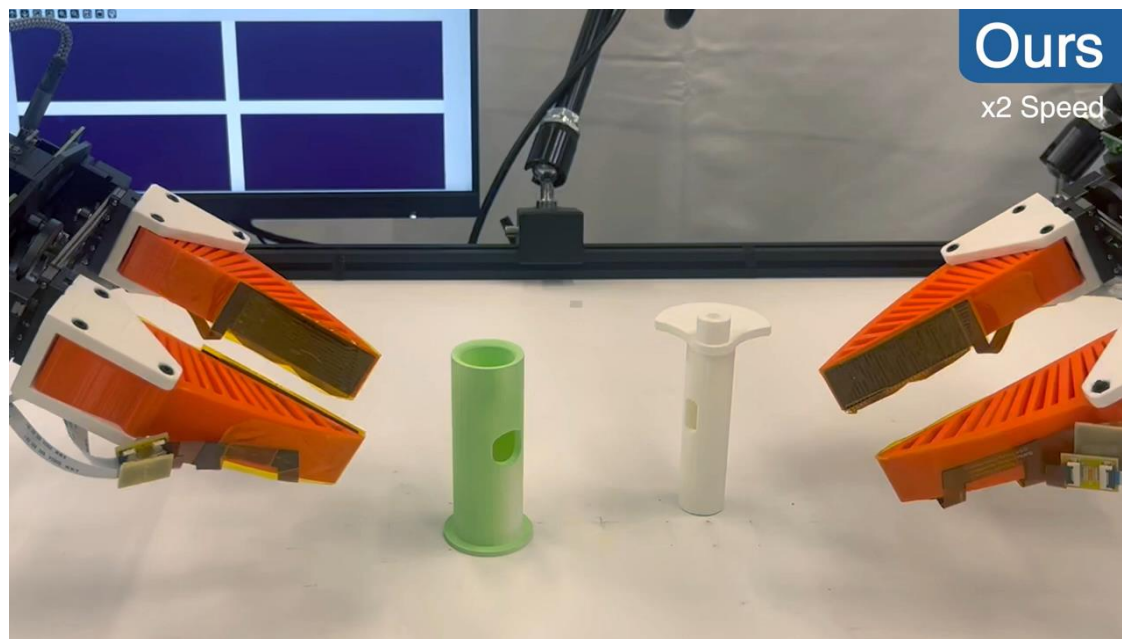


Vision

Touch

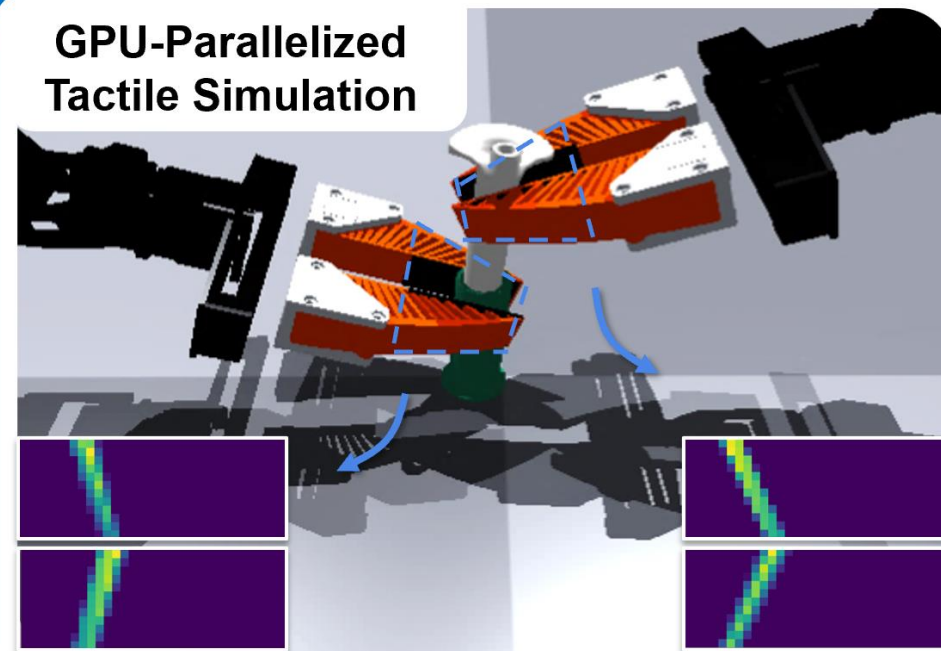
Pre-Train

Diffusion Policy

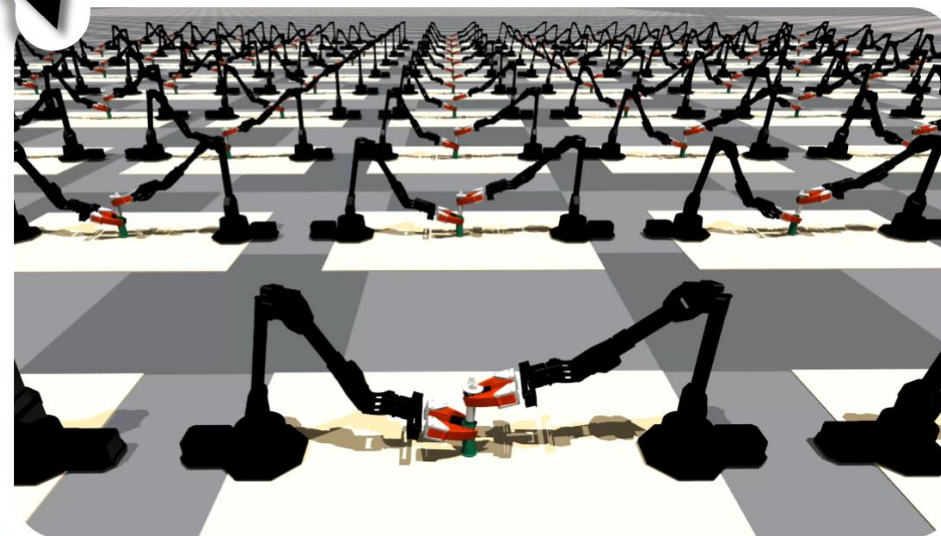


Simulation

GPU-Parallelized
Tactile Simulation



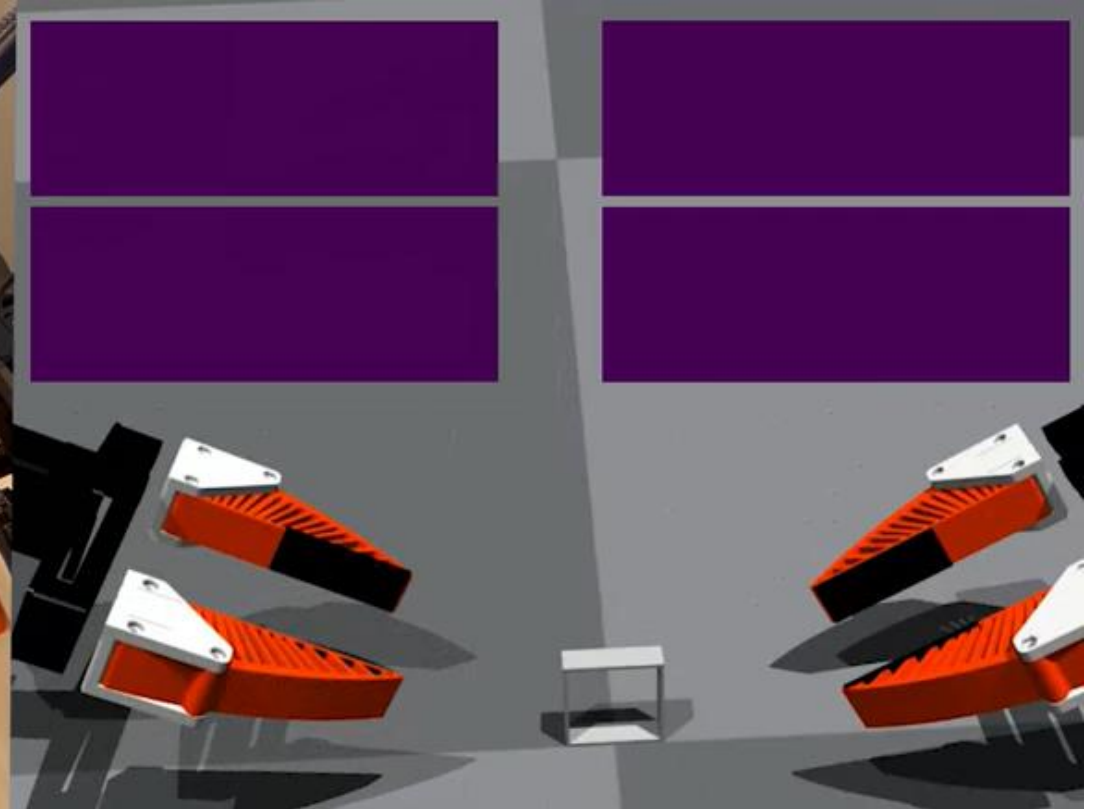
Large-Scale RL Fine-Tuning



Our Tactile Simulation



Real

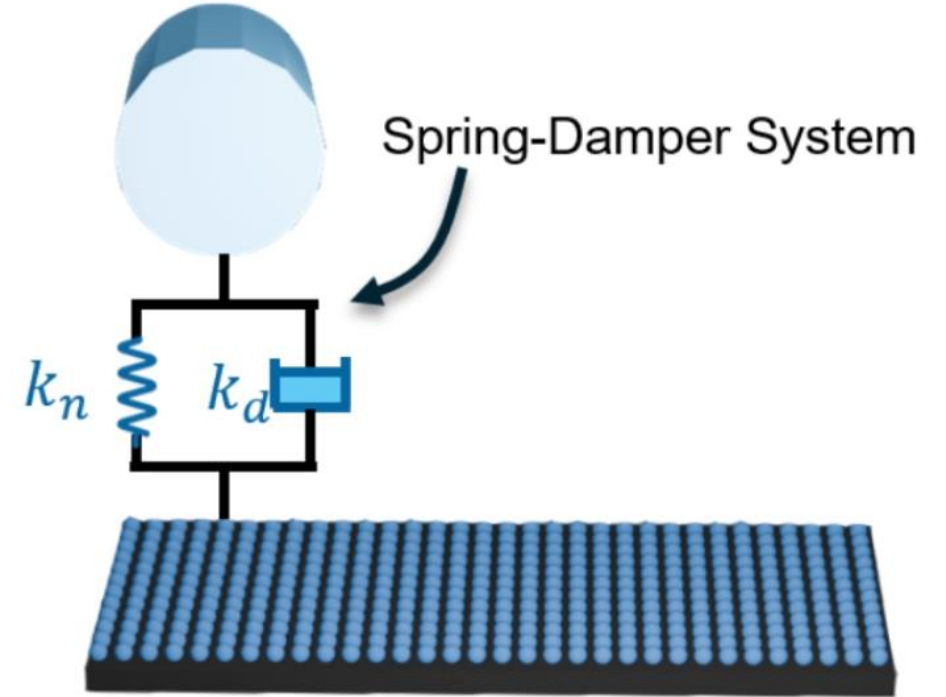
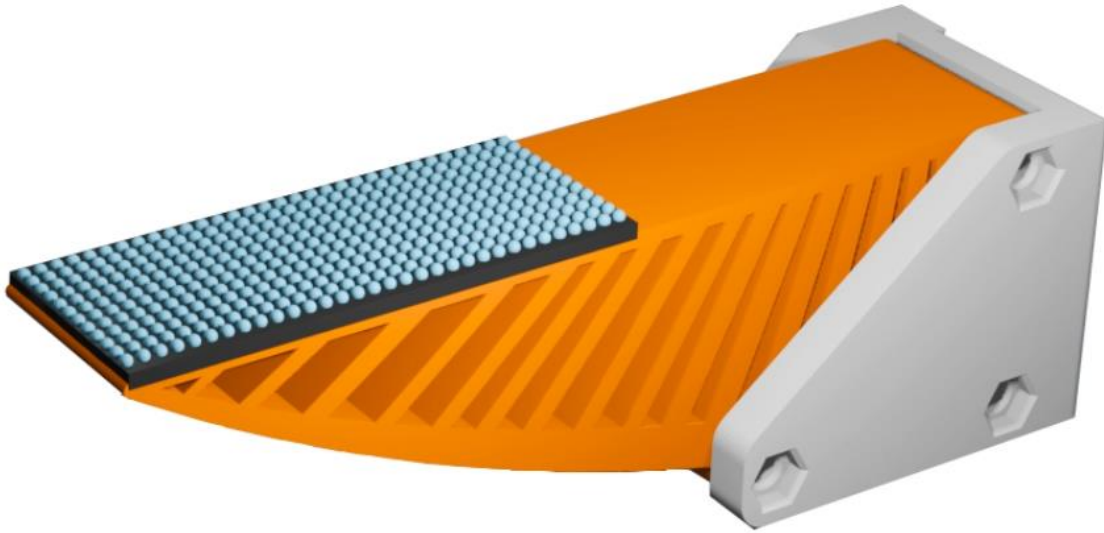


Sim

Binghao Huang, Jie Xu, Iretiayo Akinola, Wei Yang, Balakumar Sundaralingam, Rowland O'Flaherty, Dieter Fox, Xiaolong Wang, Arsalan Mousavian, Yu-Wei Chao, and Yunzhu Li, "VT-Refine: Learning Bimanual Assembly with Visuo-Tactile Feedback via Simulation Fine-Tuning," CoRL 2025.

Best Paper Award, IROS 2025 AHFR Workshop.

Our Tactile Simulation



Kelvin–Voigt model:

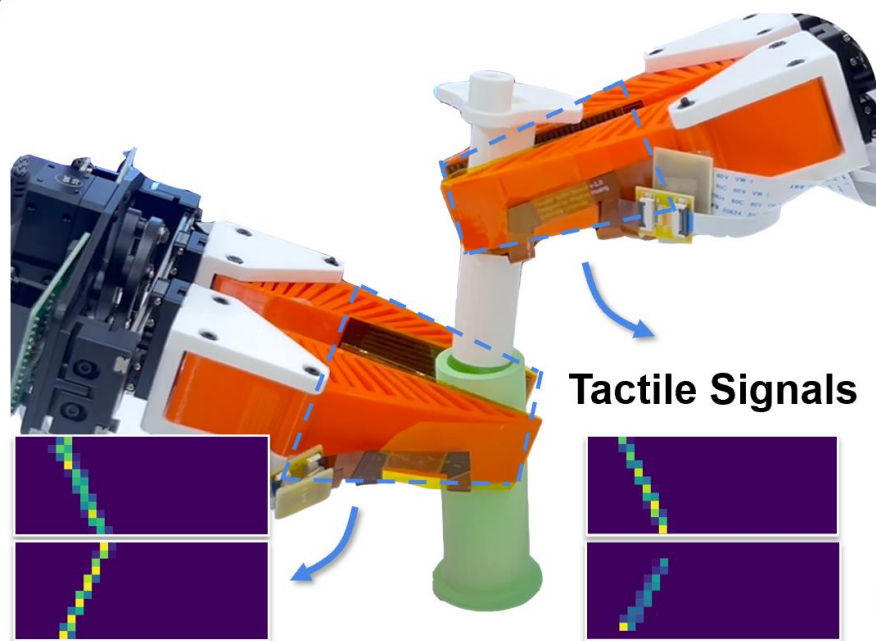
$$f_n = -(k_n d + k_d \dot{d}) \mathbf{n}$$

- k_n, k_d = stiffness and damping coefficients
- $d, \dot{d}$ = penetration depth and velocity

Binghao Huang, Jie Xu, Iretiayo Akinola, Wei Yang, Balakumar Sundaralingam, Rowland O’Flaherty, Dieter Fox, Xiaolong Wang, Arsalan Mousavian, Yu-Wei Chao, and Yunzhu Li, “VT-Refine: Learning Bimanual Assembly with Visuo-Tactile Feedback via Simulation Fine-Tuning,” CoRL 2025.

Best Paper Award, IROS 2025 AHFR Workshop.

Real World



Human Demos

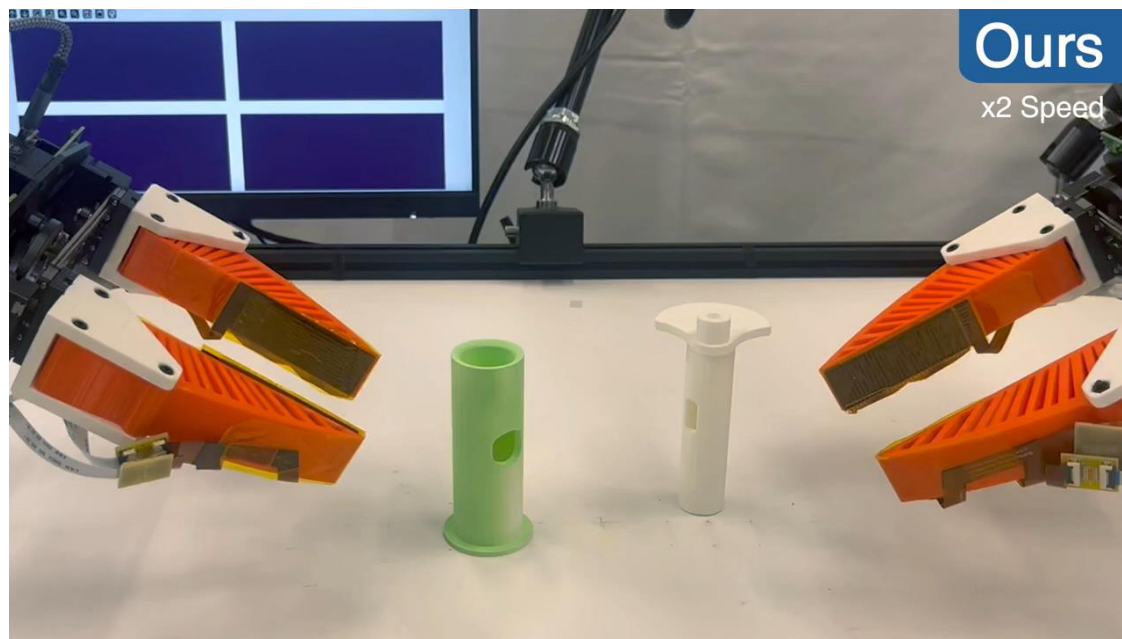


Vision

Touch

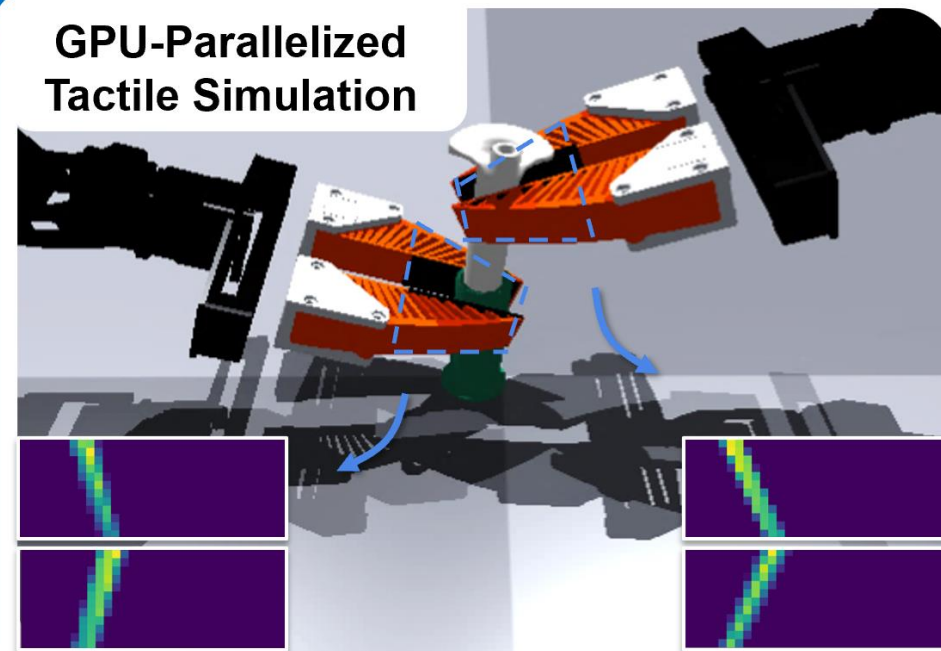
Pre-Train

Diffusion Policy

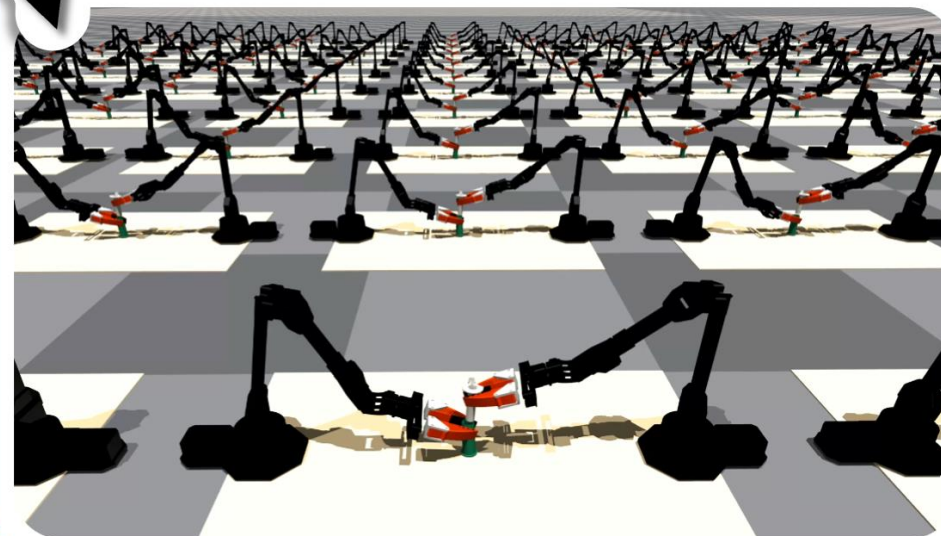


Simulation

GPU-Parallelized
Tactile Simulation



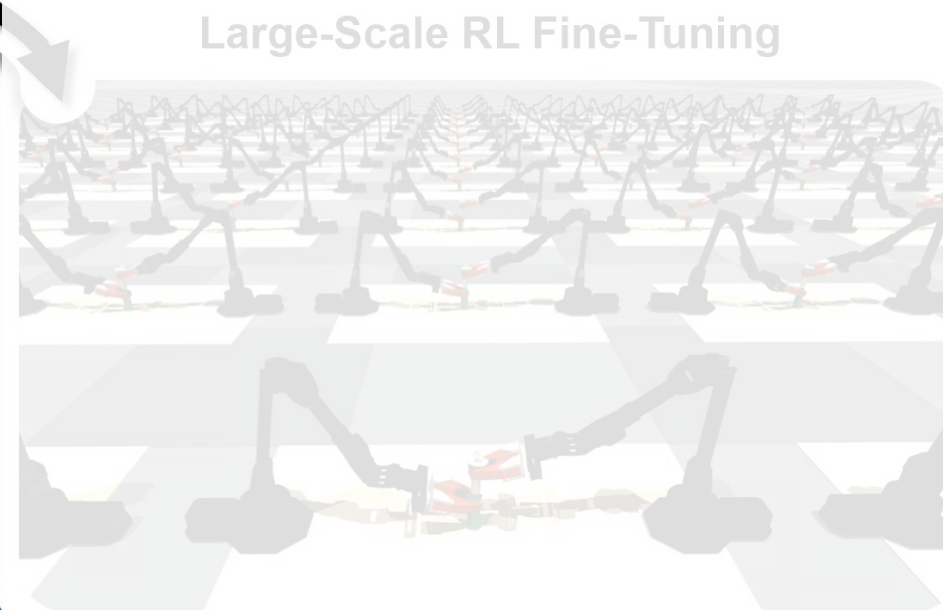
Large-Scale RL Fine-Tuning



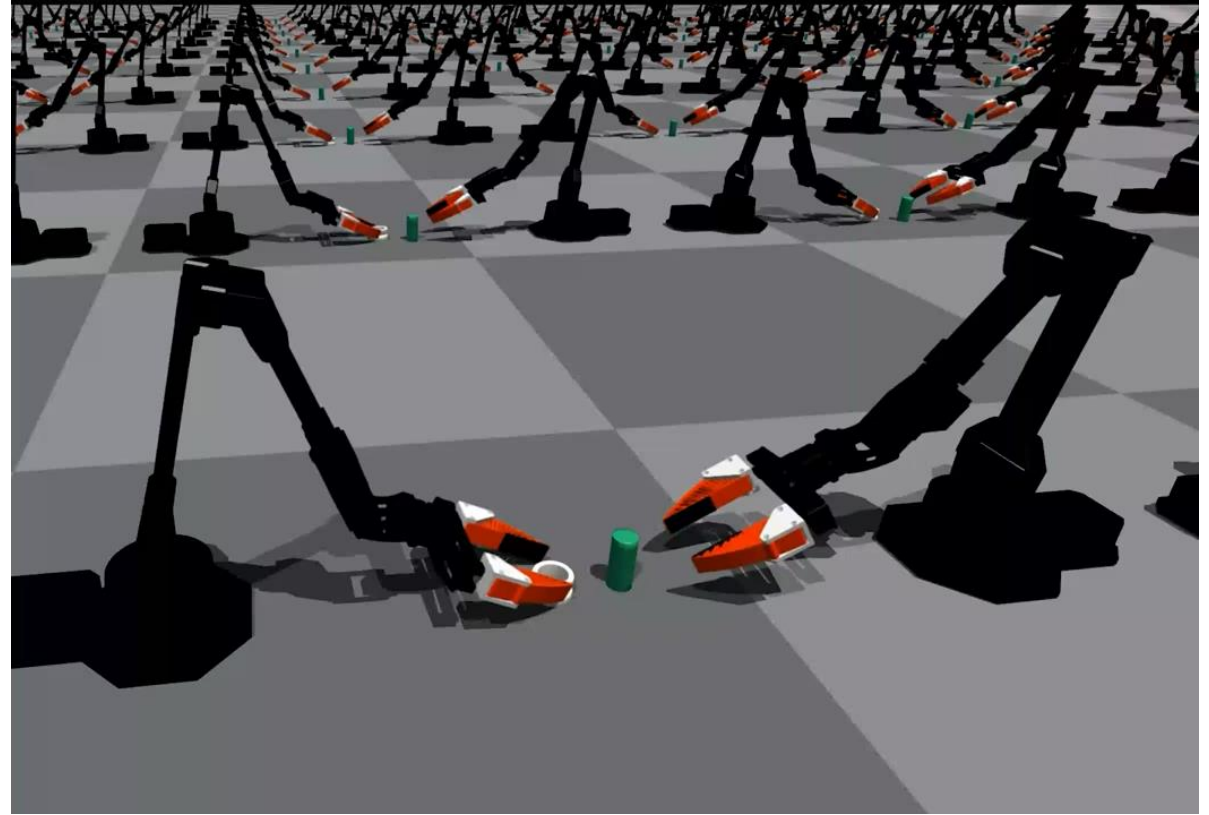
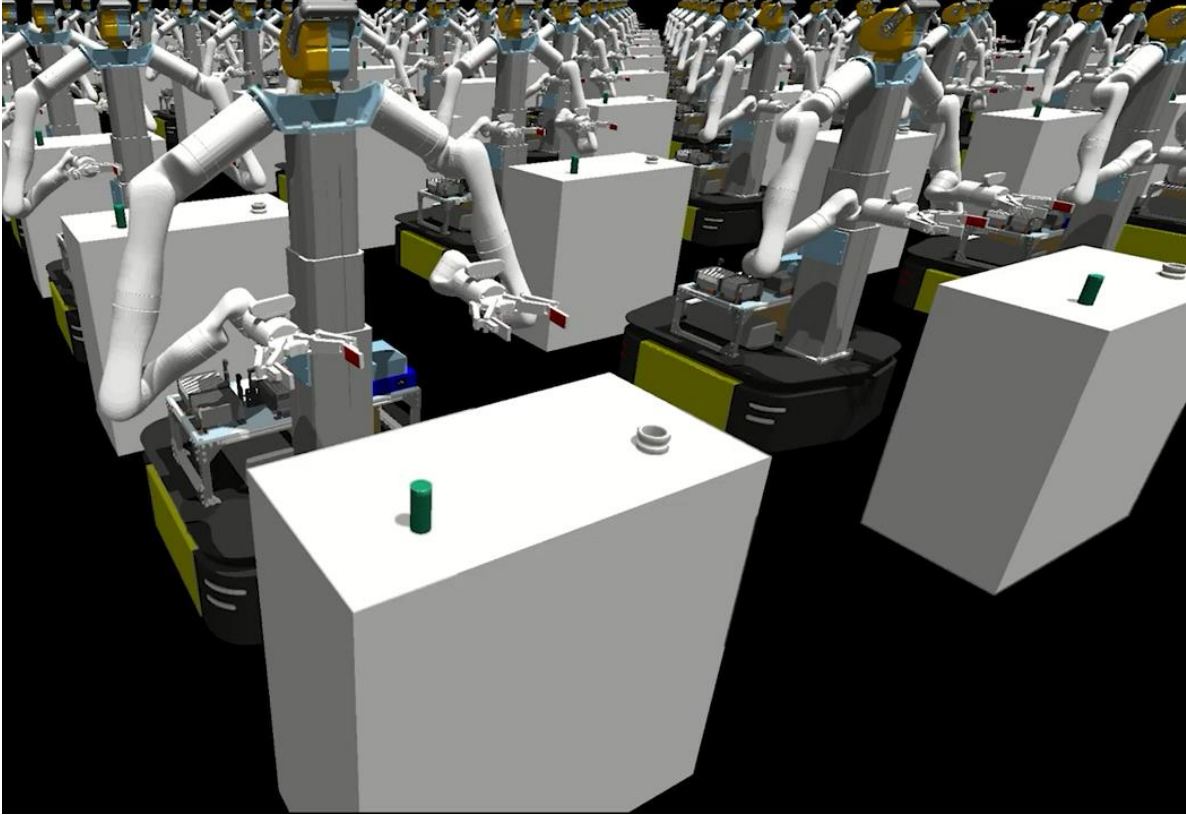
Real World



Simulation



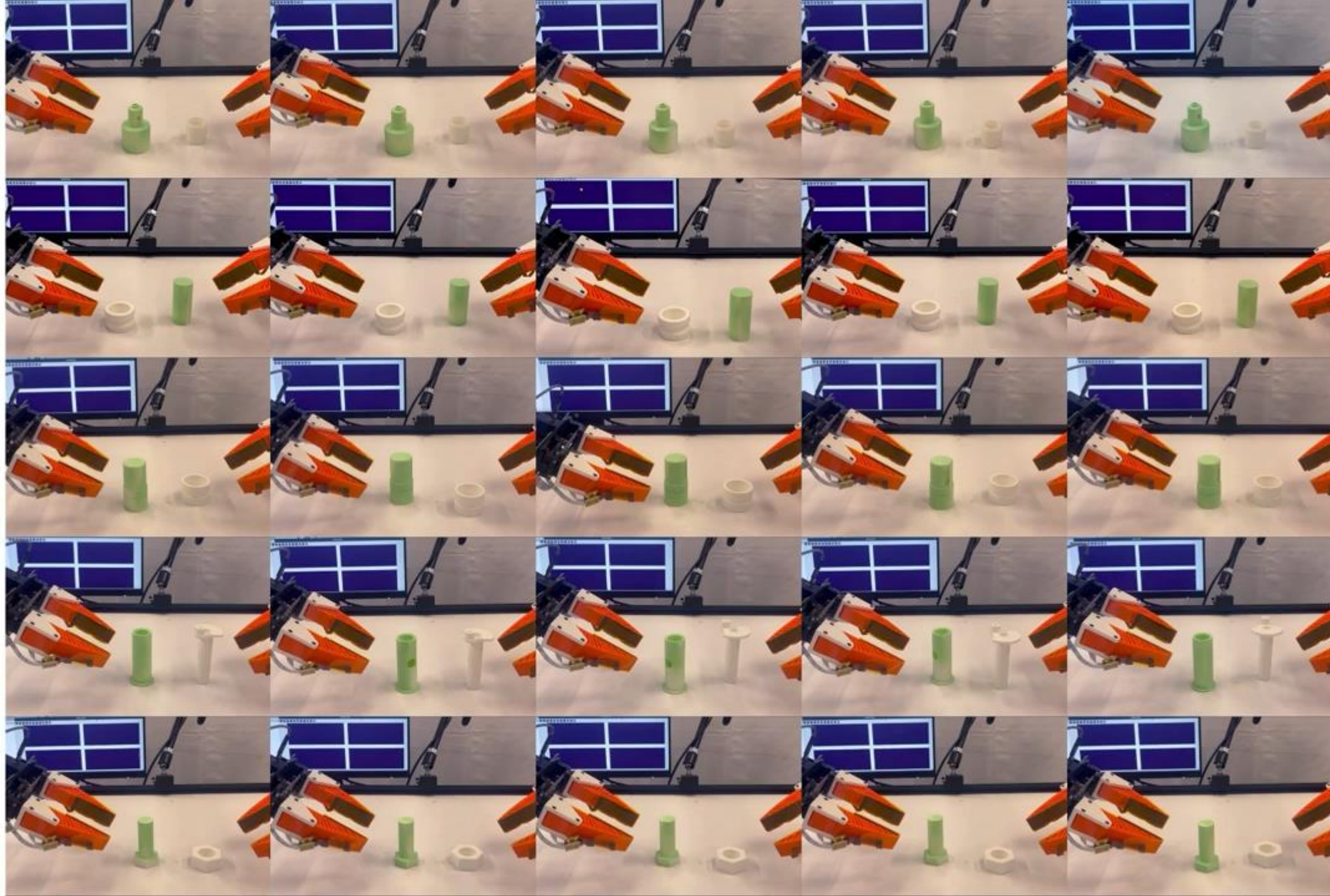
Real → Sim → Real: Scalable RL Training



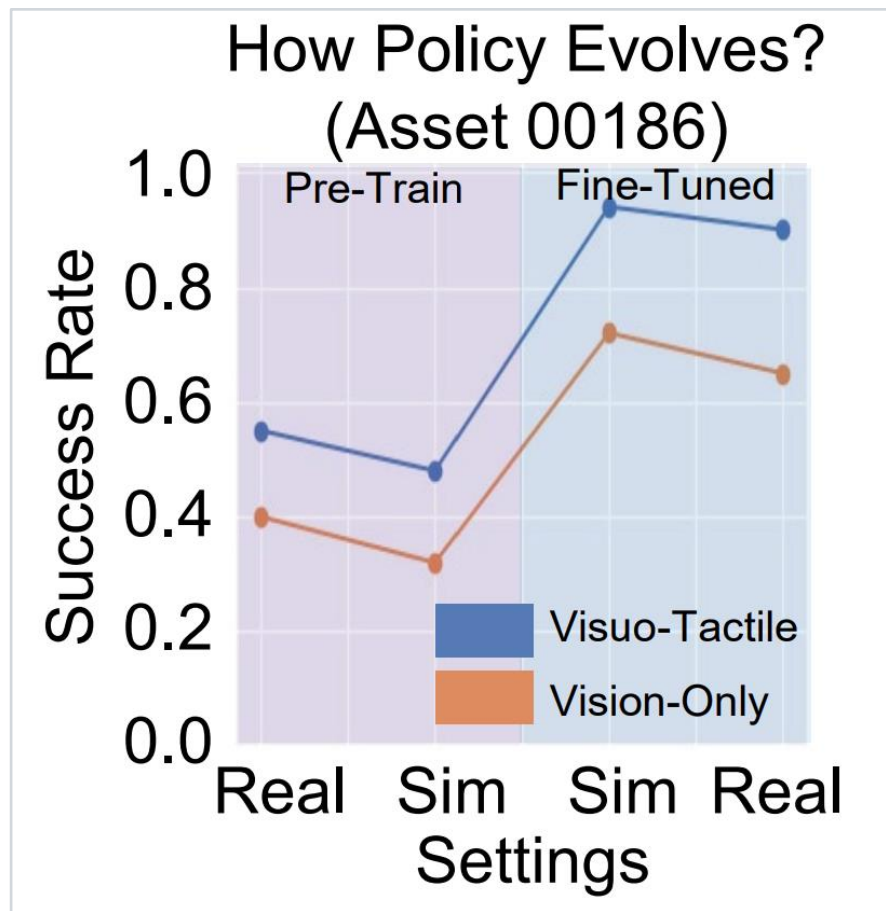
Binghao Huang, Jie Xu, Iretiayo Akinola, Wei Yang, Balakumar Sundaralingam, Rowland O’Flaherty, Dieter Fox, Xiaolong Wang, Arsalan Mousavian, Yu-Wei Chao, and Yunzhu Li, “VT-Refine: Learning Bimanual Assembly with Visuo-Tactile Feedback via Simulation Fine-Tuning,” CoRL 2025.

Best Paper Award, IROS 2025 AHFR Workshop.

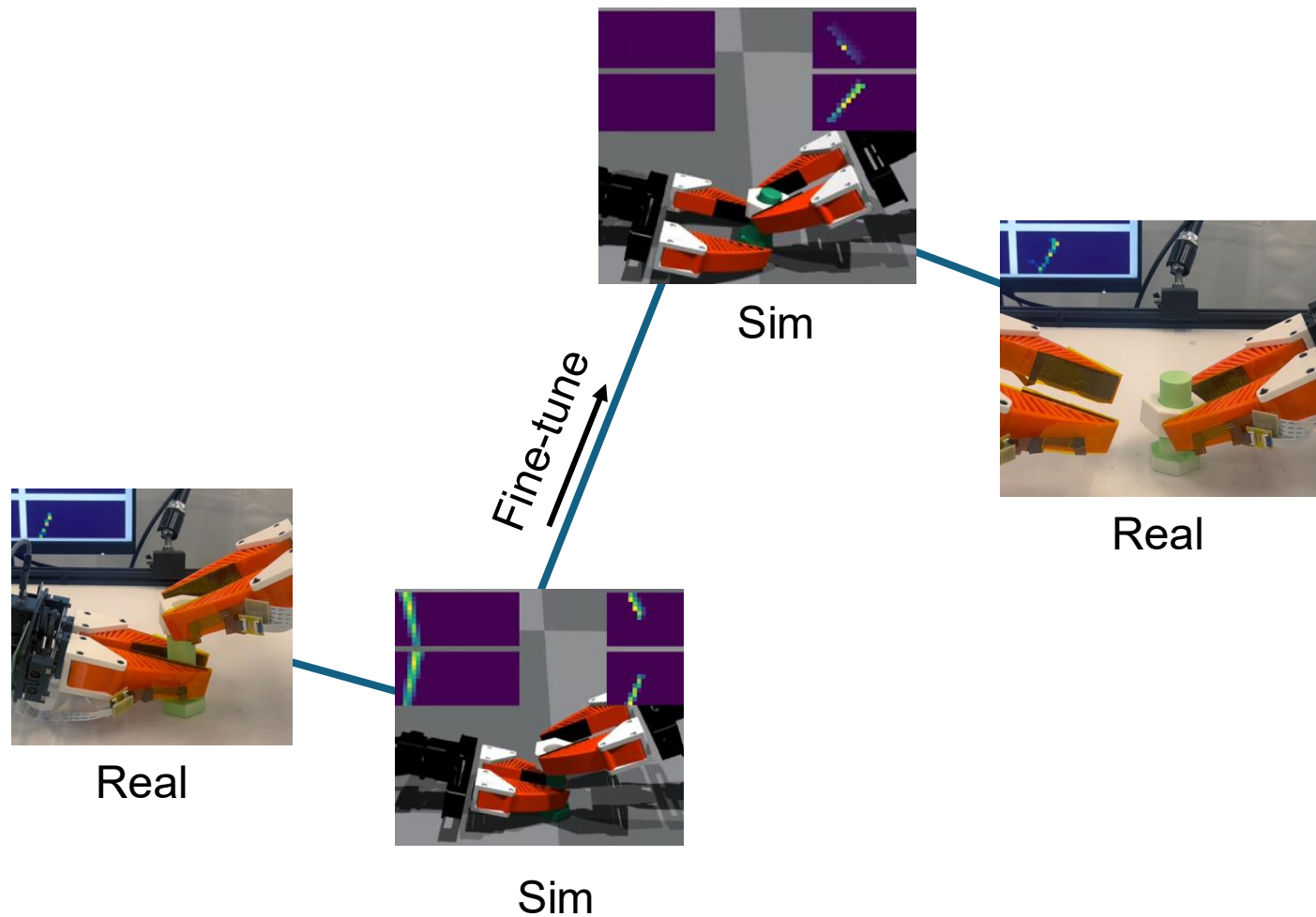




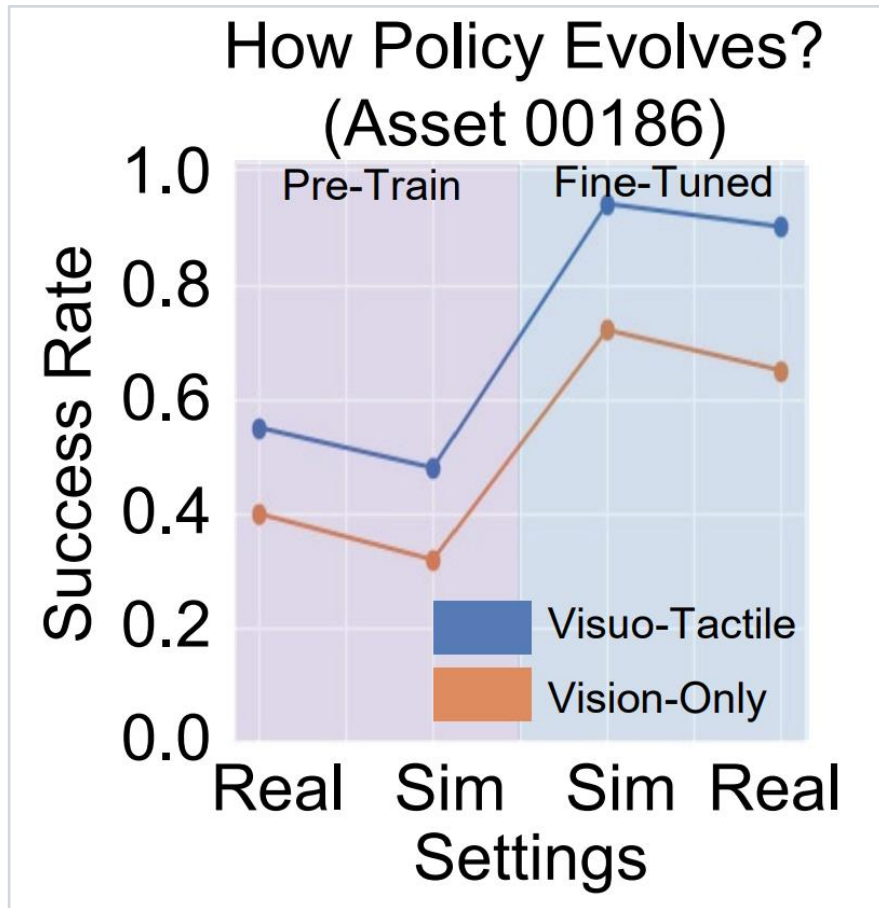
Real → Sim → Real: How the Policy Evolves



Success rate across Real / Sim,
pre-train → fine-tuned



Real → Sim → Real: How the Policy Evolves



Success rate across Real / Sim,
pre-train → fine-tuned

KEY TAKEAWAYS

1 Sim-to-real gap remains

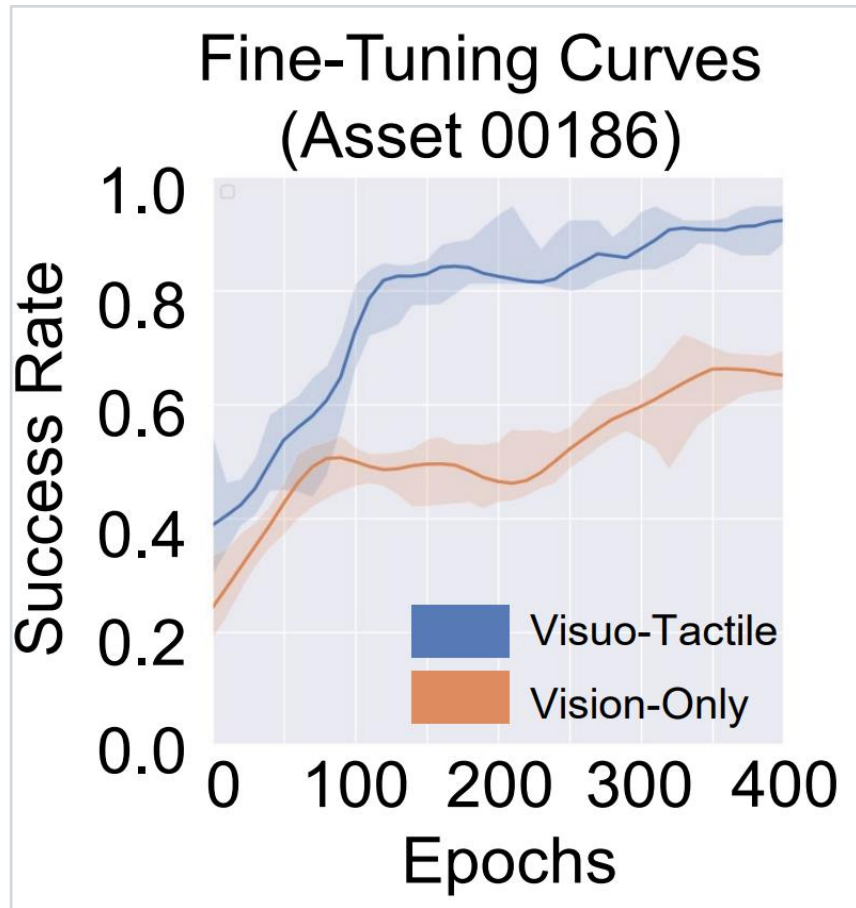
Sim scores below Real in each phase.

2 Fine-tuning lifts both policies

And visuo-tactile gains the most.

Sim fine-tuning helps **even with the gap.**

Results: Visuo-Tactile vs. Vision-Only



Fine-tuning success rate vs. epochs

KEY TAKEAWAY

40% → 95%

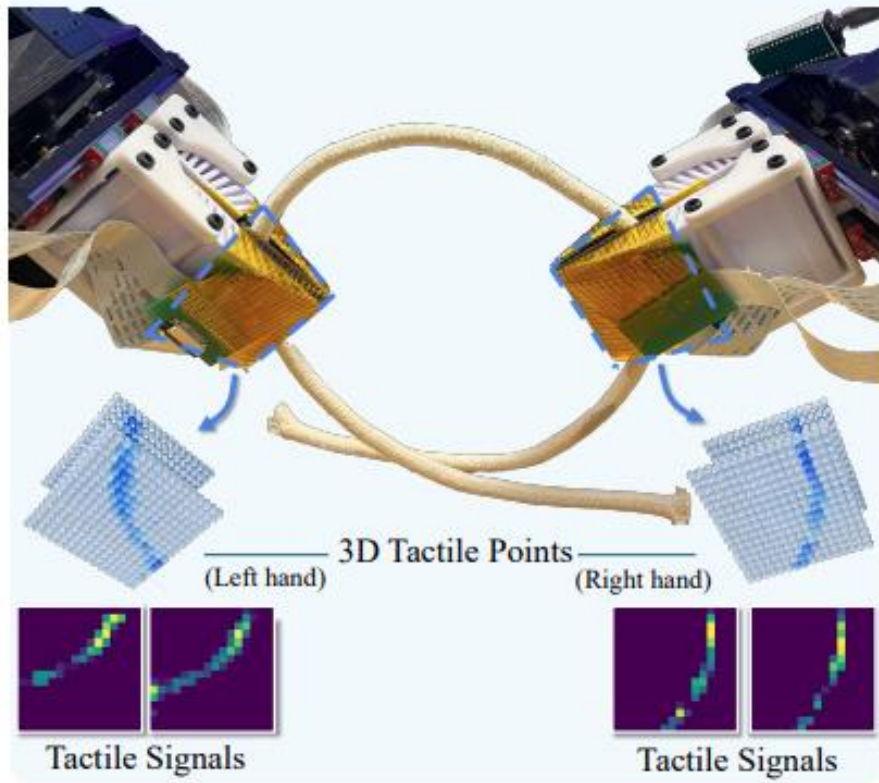
Visuo-Tactile

22% → 65%

Vision-Only

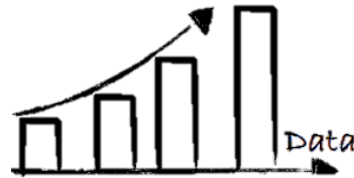
More reliable, more robust manipulation.

Scaling Up Tactile Data



[1] Huang et al., “3D-ViTac: Learning Fine-Grained Manipulation with Visuo-Tactile Sensing,” CoRL 2024.

Scale in Real



Scale in Sim



Portable devices to collect real data fast

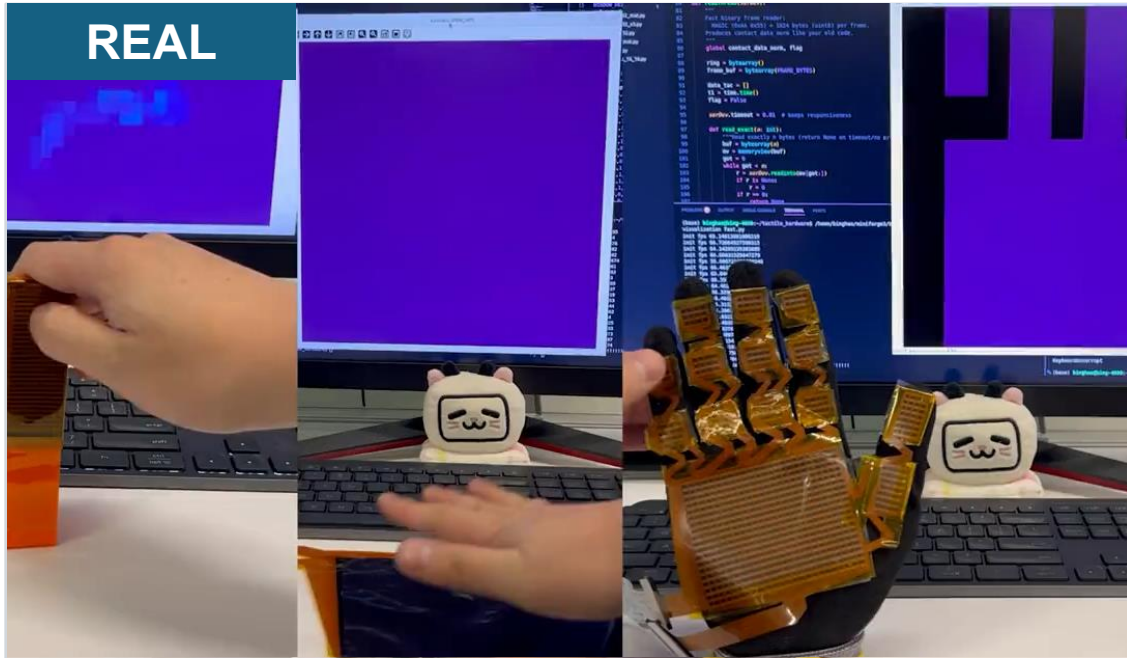
[2] Zhu et al., “Touch in the Wild: Learning Fine-Grained Manipulation with a Portable Visuo-Tactile Gripper,” NeurIPS 2025.



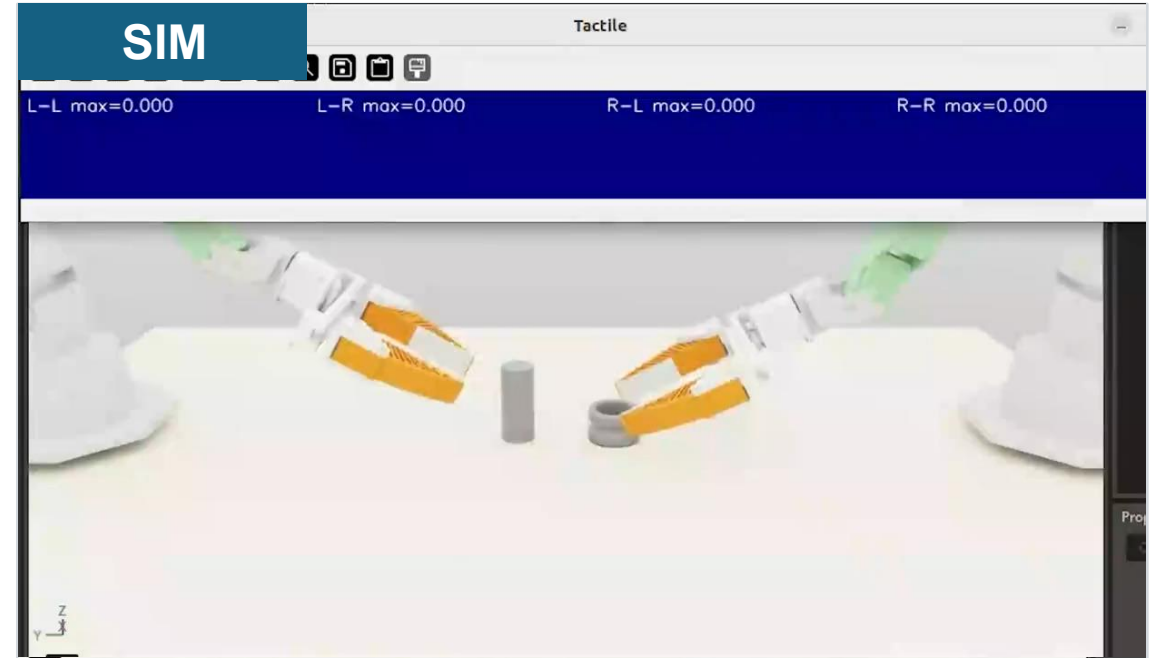
Simulation to increase local exploration

[3] Huang et al., “VT-Refine: Learning Bimanual Assembly with Visuo-Tactile Feedback via Simulation Fine-Tuning,” CoRL 2025.

We Open-Source FlexiTac — **Real + Simulation**

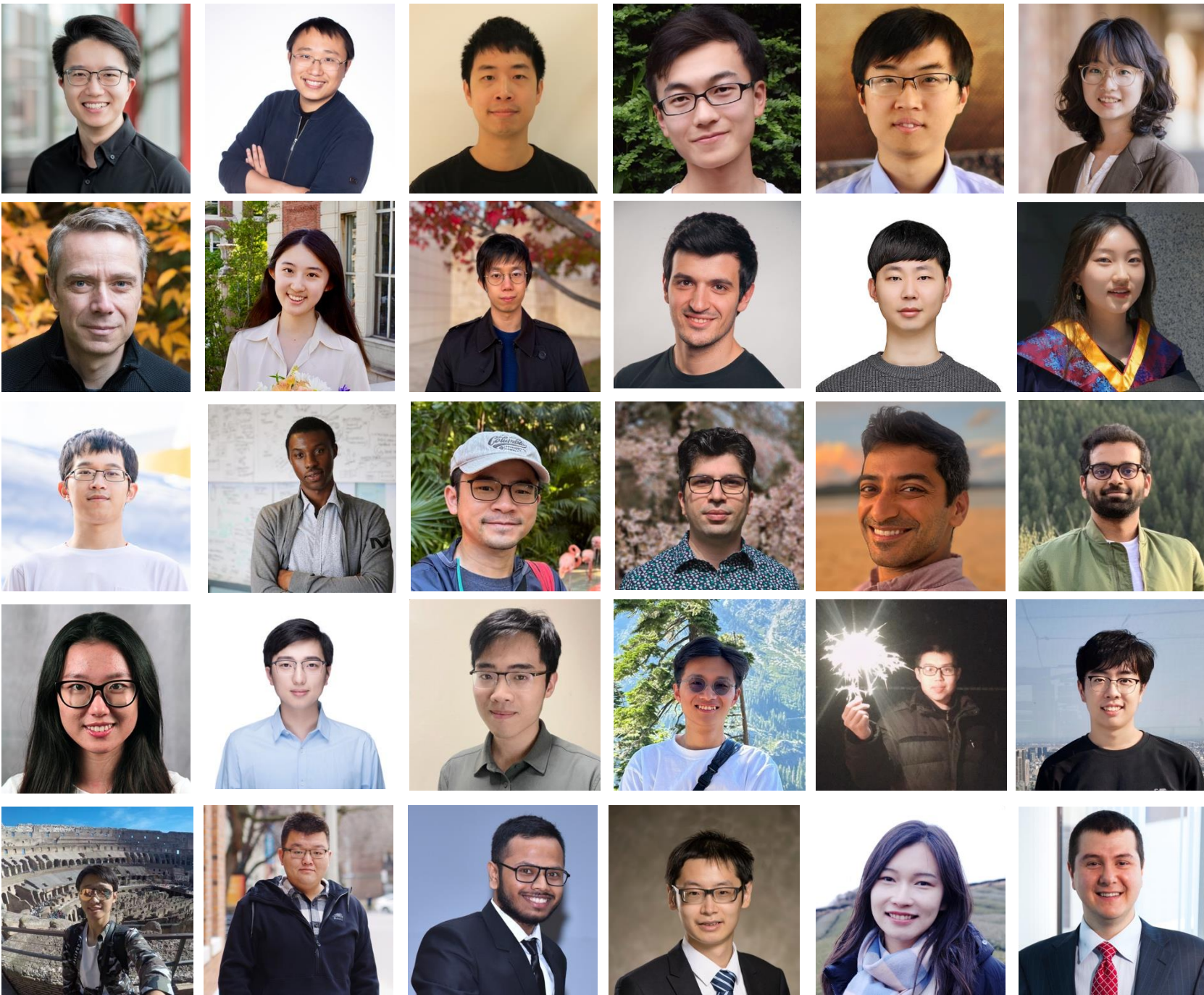


Open Hardware: Build & Deploy



GPU-Parallel Tactile Simulation

Consider both hardware and software — build better robots.



Acknowledgement



Massachusetts
Institute of
Technology



Carnegie
Mellon
University



Georgia
Tech.



UNIVERSITY OF
ILLINOIS
URBANA-CHAMPAIGN



USC



COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK



NVIDIA®

Next: 3D Multi-Modal World Model

