KAIST

RIRO LAB
Robust Intelligence & Robotics Lab

A Reinforcement Learning Testbed for Deformable Object Manipulation using Visuotactile Sensing

Presenter: Chanyoung Ahn

Master's Thesis

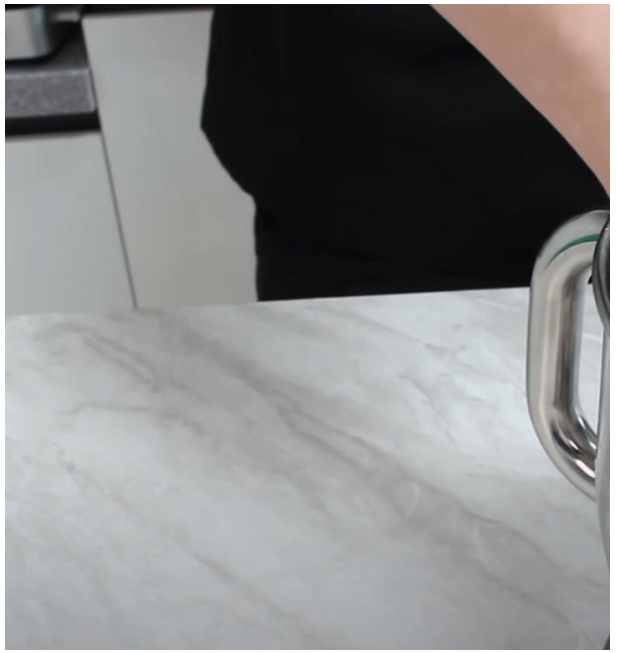
Robotics Program, KAIST

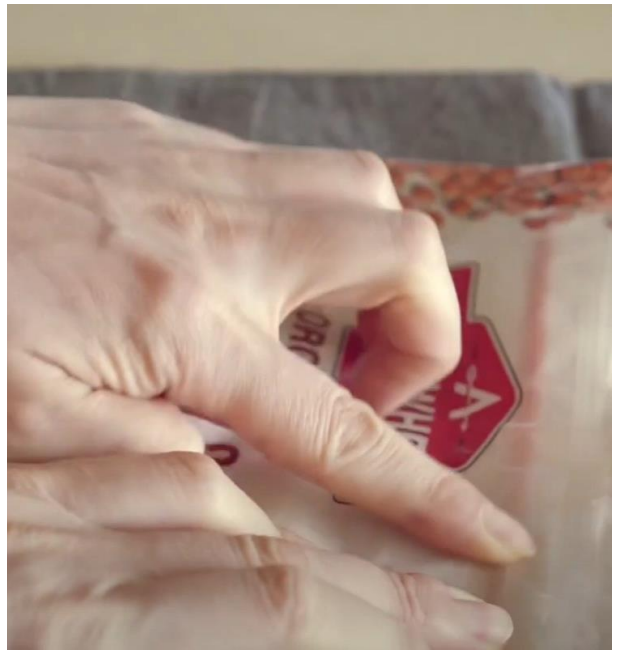
Advisor: Daehyung Park

❖ Introduction

❖ Motivation

- **Deformable Object Manipulation (DOM)**
  - Manipulation of deformable objects can have various applications in daily life.





https://youtu.be/OXzW3QB0anc?si=DAEriBhcRaMJr1qa
https://youtube.com/shorts/cLJWYsiDf0Q?si=sJCZmZVb6iNSuCUk
https://youtu.be/1VXX6vUCwV4?si=1Me_JrcRGPDU0O-z

❖ Introduction

❖ Motivation

- **Deformable Object Manipulation (DOM)**

Vision

Vision
+
Contact

Vision
+
Tactile

❖ Introduction

❖ Motivation (Wrap up)

Heterogeneous
Deformable Object Manipulation

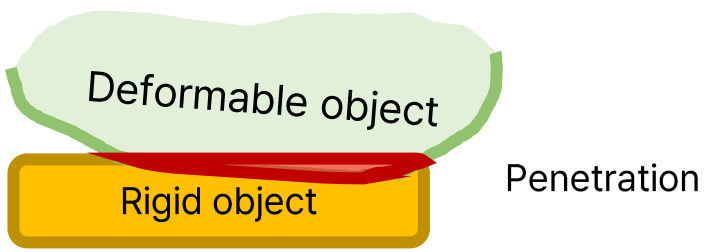
Vision + Tactile sensing

Simulated Testbed

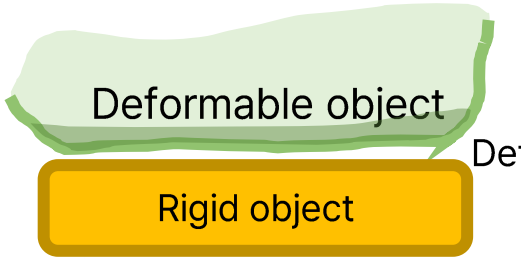
❖ Problem statement

We aim to build a testbed for learning DOM skills with visuotactile sensing

Conventional testbeds do not provide a direct way to obtain tactile information from deformable objects



Penetration



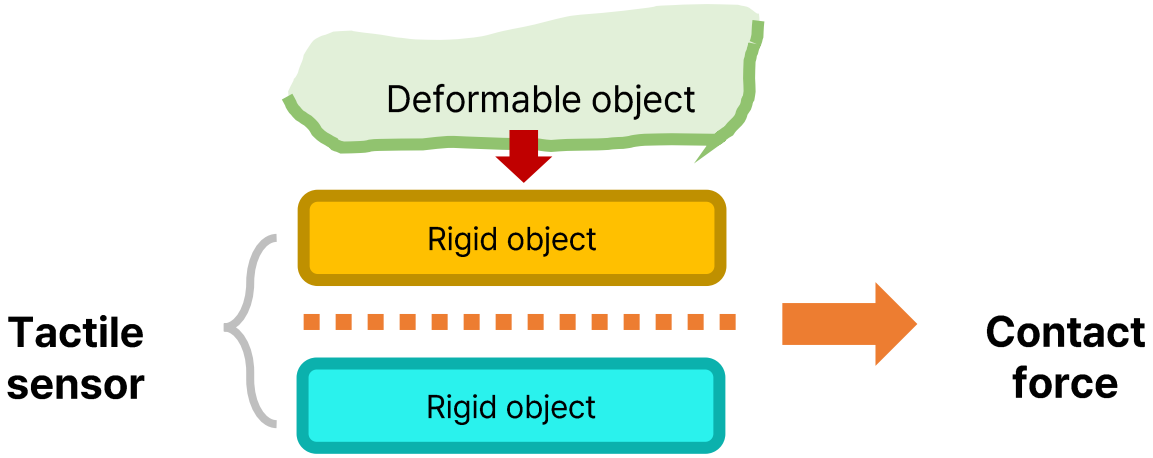
Deformation

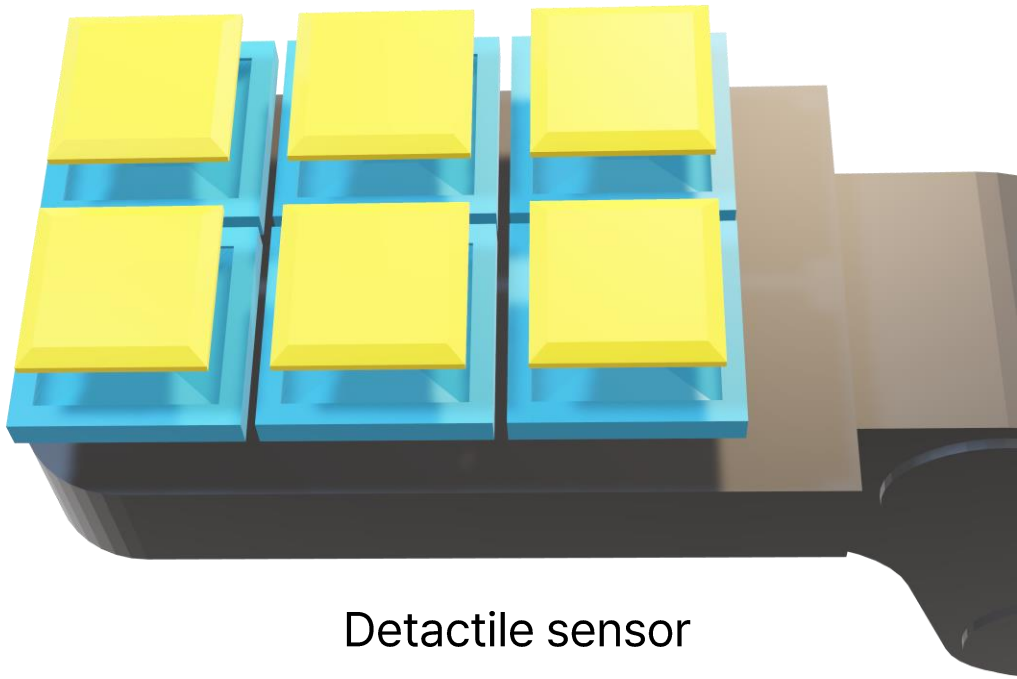
Penetration
+
Deformation
+
Relative velocity

Contact force

❖ Proposed method

We introduce a visuotactile testbed for deformable object manipulation, integrating a novel architecture of tactile sensor leveraging collision cascades.





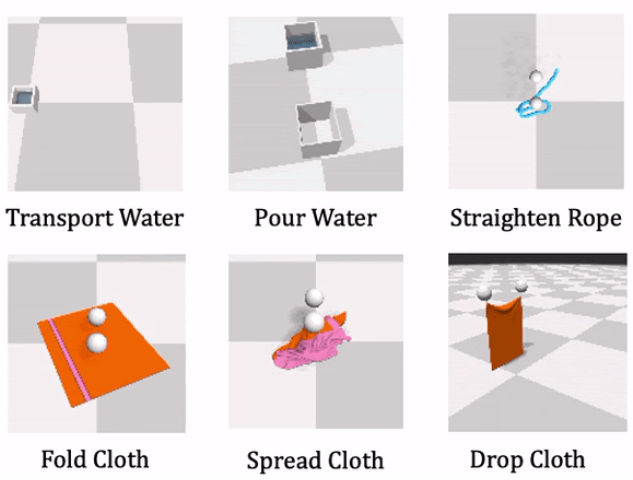
Detactile sensor

❖ Related Work

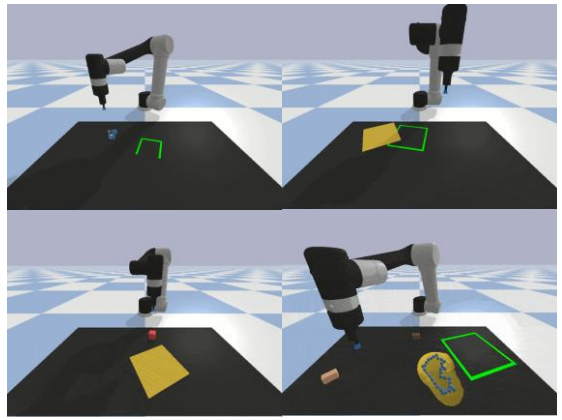
❖ Testbed for Deformable Object Manipulation (DOM)

- Previous most testbed developments focus on **only vision-based DOM**

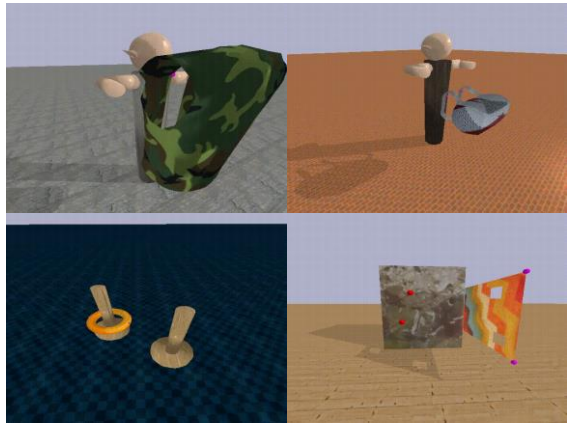
SoftGym [2]



DeformableRavens [3]



DEDO [4]



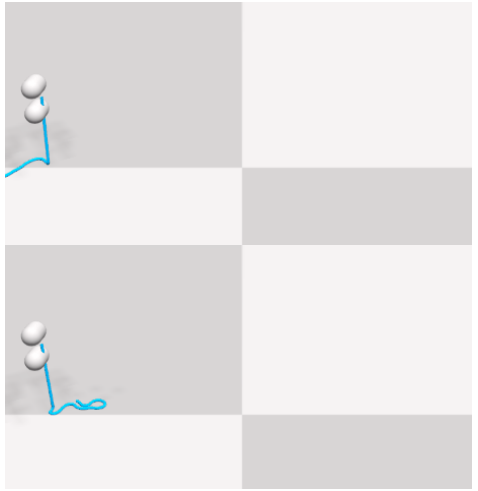
[2] Lin, Xingyu, et al. "Softgym: Benchmarking deep reinforcement learning for deformable object manipulation." Conference on Robot Learning. PMLR, 2021.
[3] Seita, Daniel, et al. "Learning to rearrange deformable cables, fabrics, and bags with goal-conditioned transporter networks." 2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2021.
[4] Antonova, Rika, et al. "Dynamic environments with deformable objects." Thirty-fifth conference on neural information processing systems datasets and benchmarks track (Round 2). 2021.

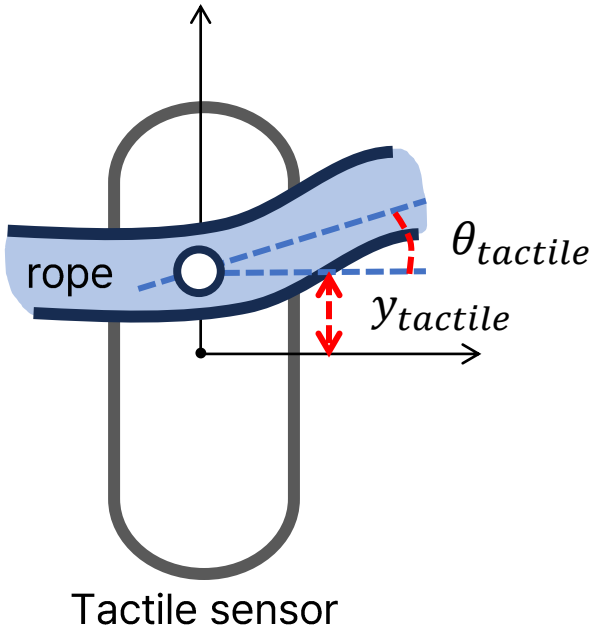
❖ Related Work

❖ Testbed for Deformable Object Manipulation with Tactile Sensor

- Previous **tactile-based gym** provides **binary tactile sensor** for DOM

SoftSlidingGym [5]





Tactile sensor

Restrictive tactile data

[y, theta]
Binary sensing
(contact/ no contact)

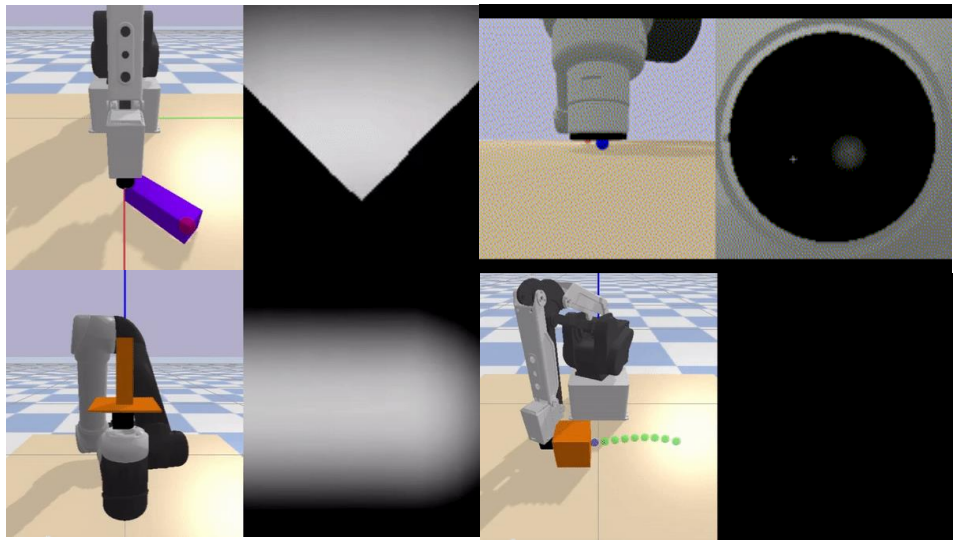
[5] Pecyna, Leszek, Siyuan Dong, and Shan Luo. "Visual-tactile multimodality for following deformable linear objects using reinforcement learning." 2022 IEEE IROS

❖ Related Work

❖ Testbed for Tactile-based manipulation

- Previous **tactile-based gym** provides tactile sensor **for rigid body manipulation**

Tactile Gym 2.0 [6]



Optical tactile sensor

Only rigid body
No deformable

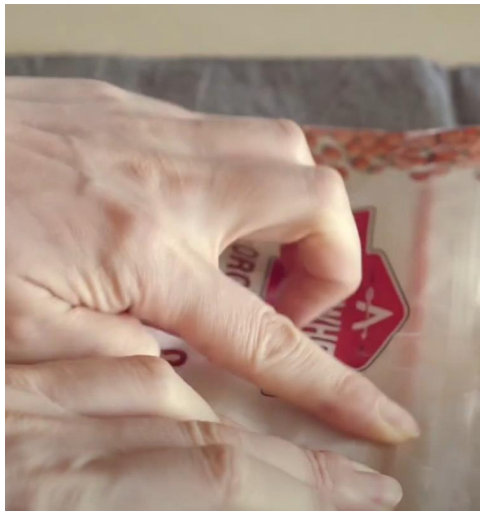
[6] Lin, Yijiong, et al. "Tactile gym 2.0: Sim-to-real deep reinforcement learning for comparing low-cost high-resolution robot touch." IEEE Robotics and Automation Letters 7.4 (2022): 10754-10761.

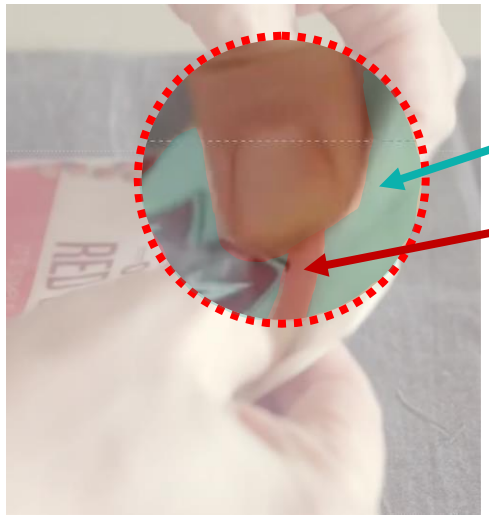
❖ Related Work

Gym	deformable	heterogenous deformable	vision	tactile	DOM RL	Physics engine	
SoftGym [CoRL 20]	O	X	O	X	O	FleX	Vision
DeformableRavens [ICRA 21]	O	X	O	X	O	Pybullet	
DEDO [NeurIPS 21]	O	X	O	X	O	Pybullet	
SoftSlidingGym [IROS 22]	O	X	O	O, binary	O	FleX	Vision + Tactile
Tactile Gym 2.0 [RA-L 22]	X	X	O	O, Image	X	Pybullet	
DetactGym(ours)	O	O	O	O, 3 axis force	O	PhysX 5	

❖ Method

❖ DetactGym



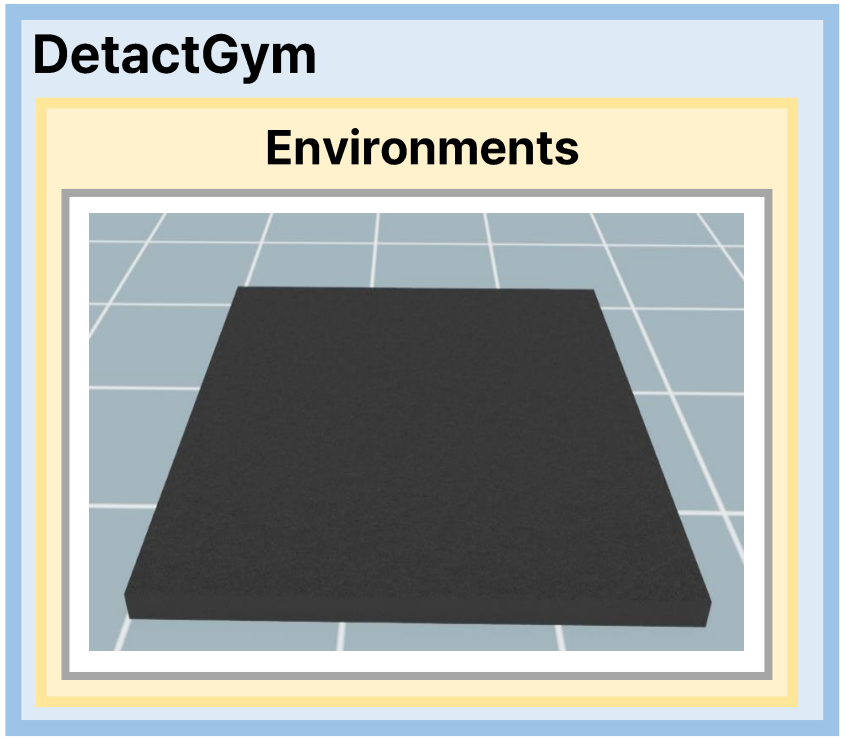


Heterogeneous
deformable

Vision
+
Tactile

❖ Method

❖ DetactGym

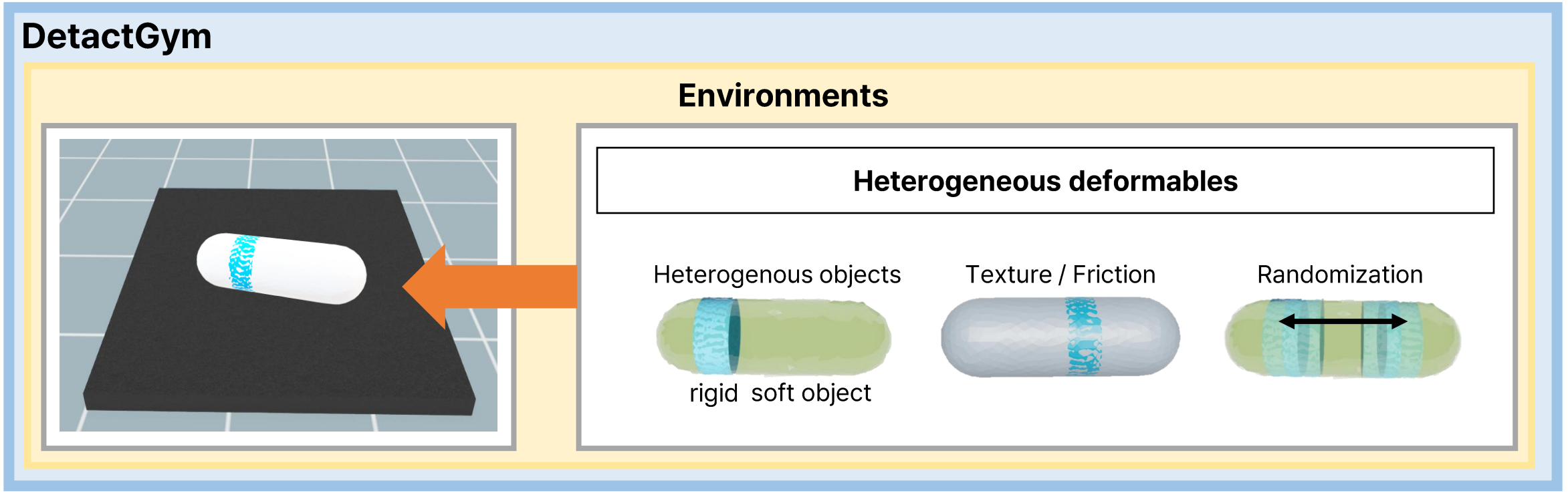


PhysX by NVIDIA + NVIDIA Isaac Sim

- CPU/GPU parallel computing performance for large simulations
- Realistic deformable(FEM method)/fluid/cloths

❖ Method

❖ DetactGym: Deformable objects

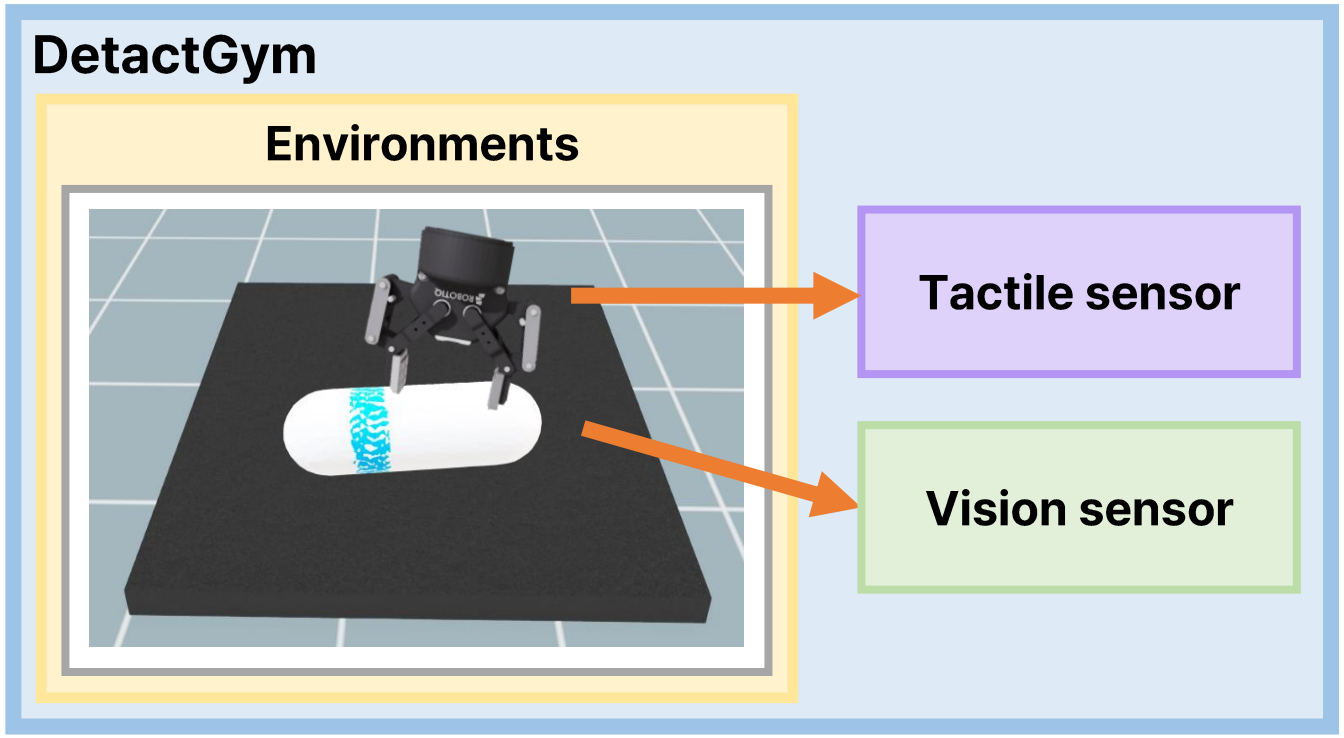


Heterogeneous DOM

- **Heterogeneous deformables**, including soft objects with rigid parts

❖ Method

❖ DetactGym: Vision and Tactile sensor

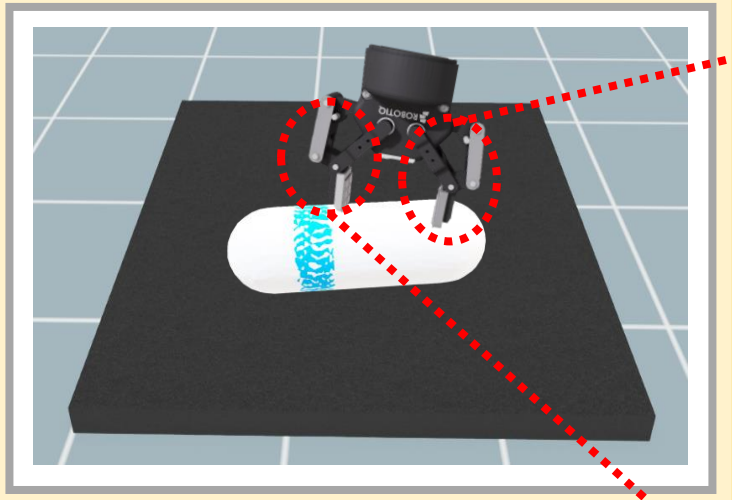


❖ Method

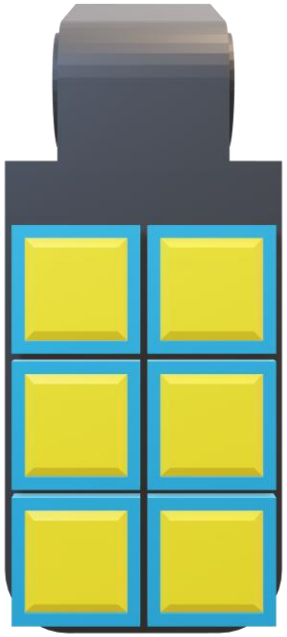
❖ DetactGym: Tactile sensor

DetactGym

Environments

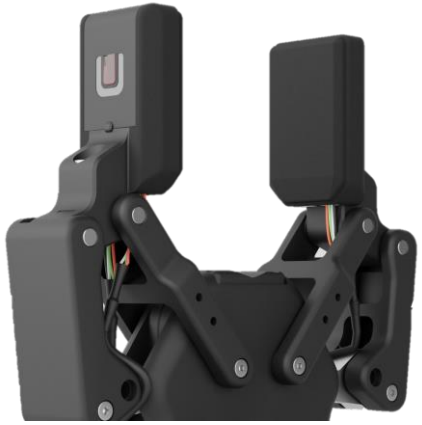


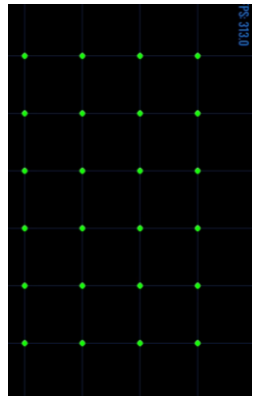
Tactile sensor



Detactile

Simulated tactile sensor





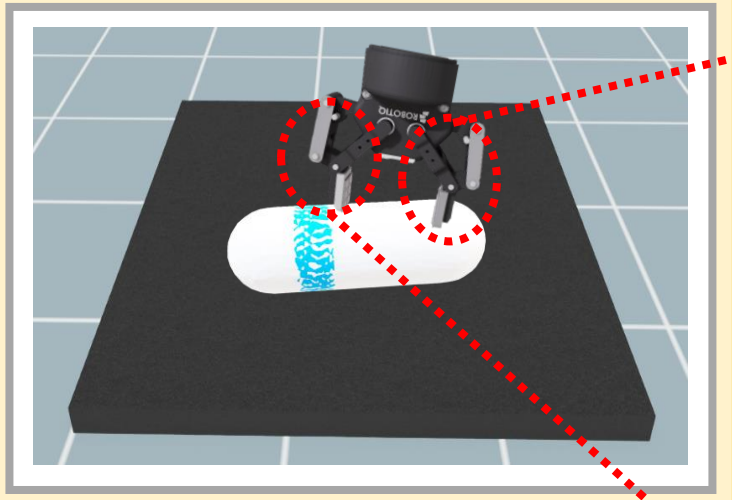
XR1946
xela robotics

❖ Method

❖ DetactGym: Tactile sensor

DetactGym

Environments



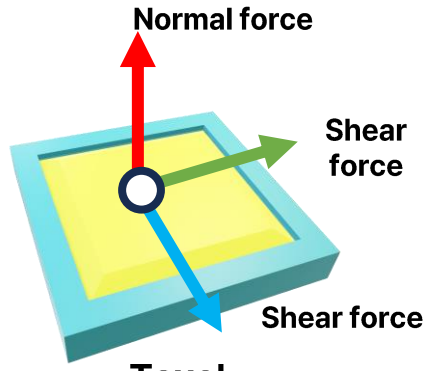
Tactile sensor

2x3 Taxels

Detactile

Simulated tactile sensor

2X3 tactile array
36 tactile data

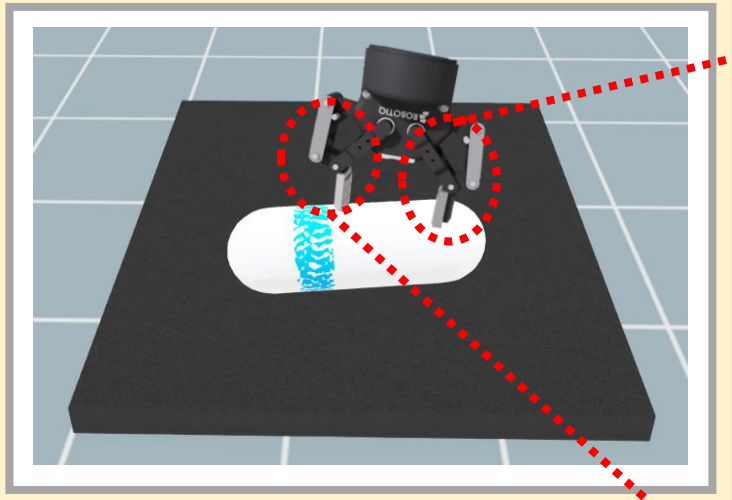


❖ Method

❖ DetactGym: Tactile sensor

DetactGym

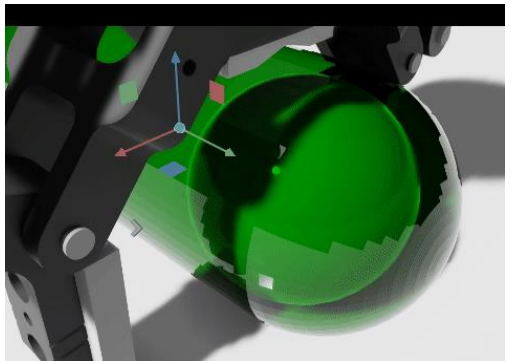
Environments

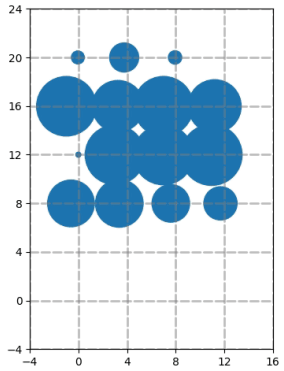


Tactile sensor

Detactile

Simulated tactile sensor



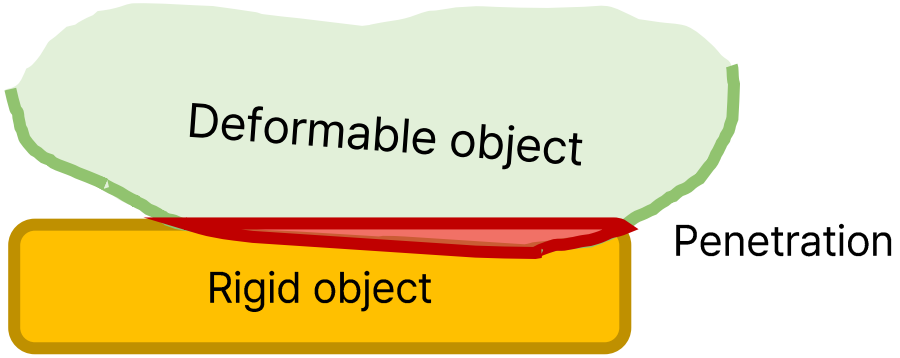


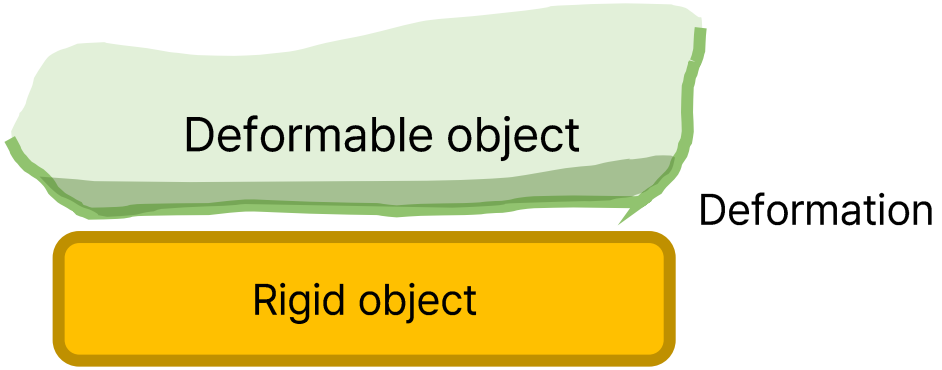
Sensing data visualization

❖ Method

❖ DetactGym: Tactile sensor

✓ Why is it difficult to provide tactile sensing in DOM testbed?





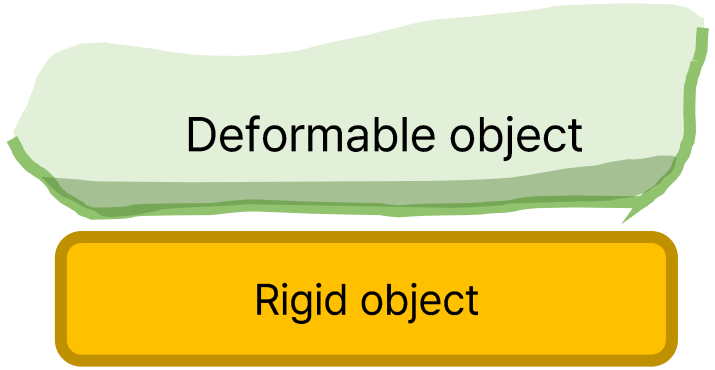
Penetration
+
Deformation
+
Relative velocity

Contact force

❖ Method

❖ DetactGym: Tactile sensor

✓ Why is it difficult to provide tactile sensing in DOM testbed?



Surface impulse, contact mesh (Bullet 3)

Stress, surface mesh normal (FleX)

.

.

Bullet 3, https://github.com/bulletphysics/bullet3/pull/4413 2023.02.25
FleX, https://developer.nvidia.com/isaac-gym, Isaacgym release 4

❖ Method

❖ DetactGym: Tactile sensor

✓ Why is it difficult to provide tactile sensing in DOM testbed?

Surface impulse, contact mesh (Bullet 3)

Stress, surface mesh normal (FleX)

.

.

X (PhysX)

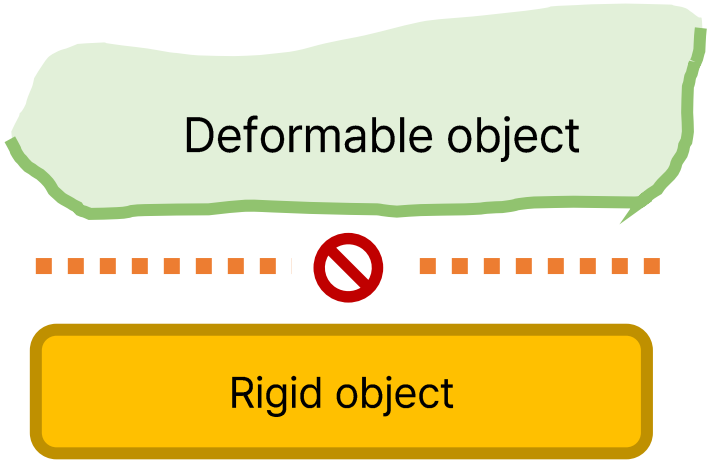
What if we are unable to get deformable information from simulators?

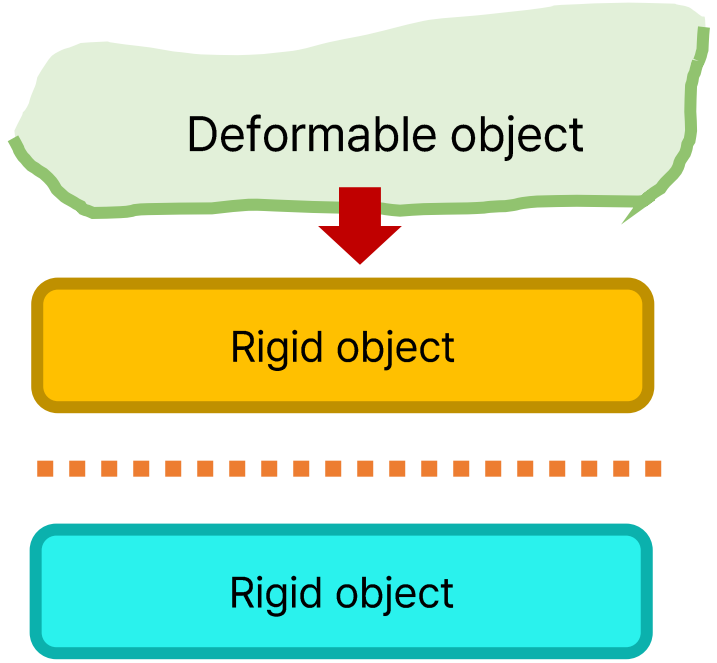
Bullet 3, https://github.com/bulletphysics/bullet3/pull/4413 2023.02.25
FleX, https://developer.nvidia.com/isaac-gym, Isaacgym release 4

❖ Method

❖ DetactGym: Tactile sensor

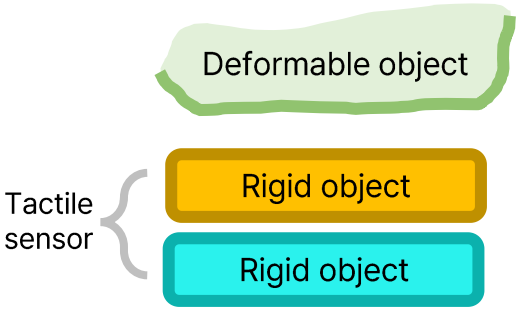
What if we are unable to get deformable information from simulators?

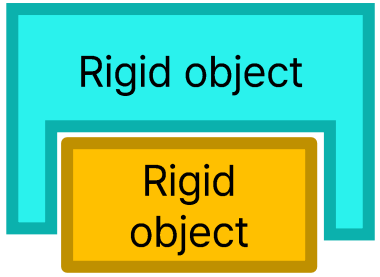


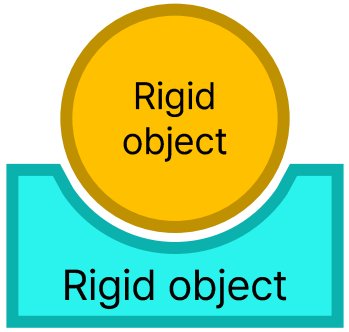


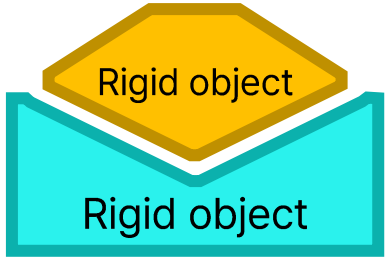
❖ Method

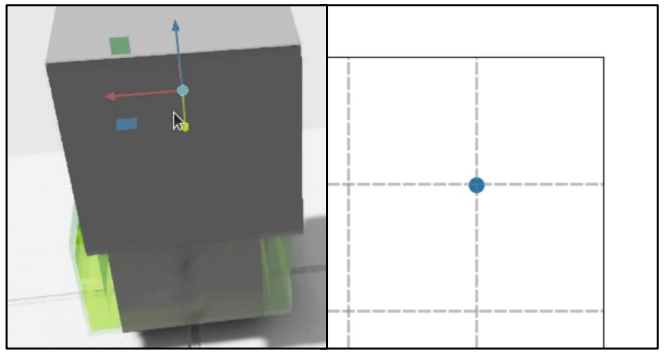
❖ DetactGym: Tactile sensor

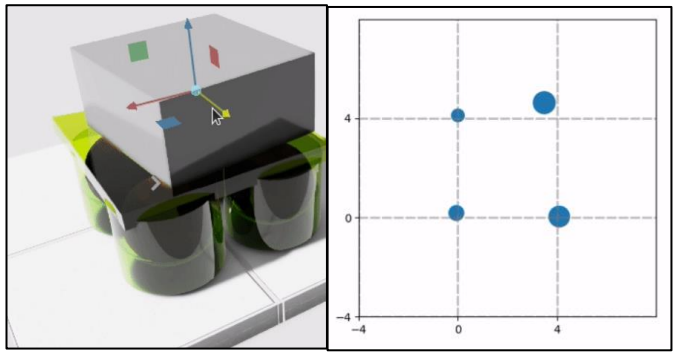


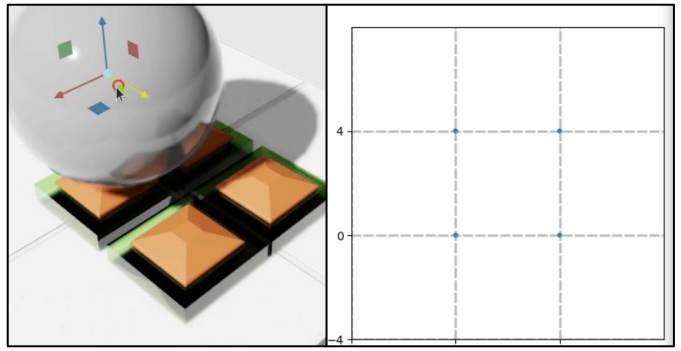






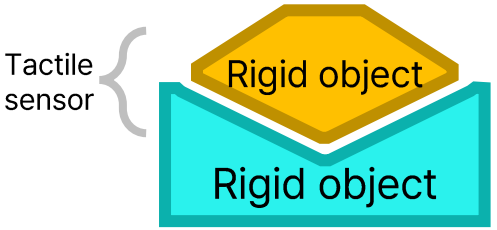


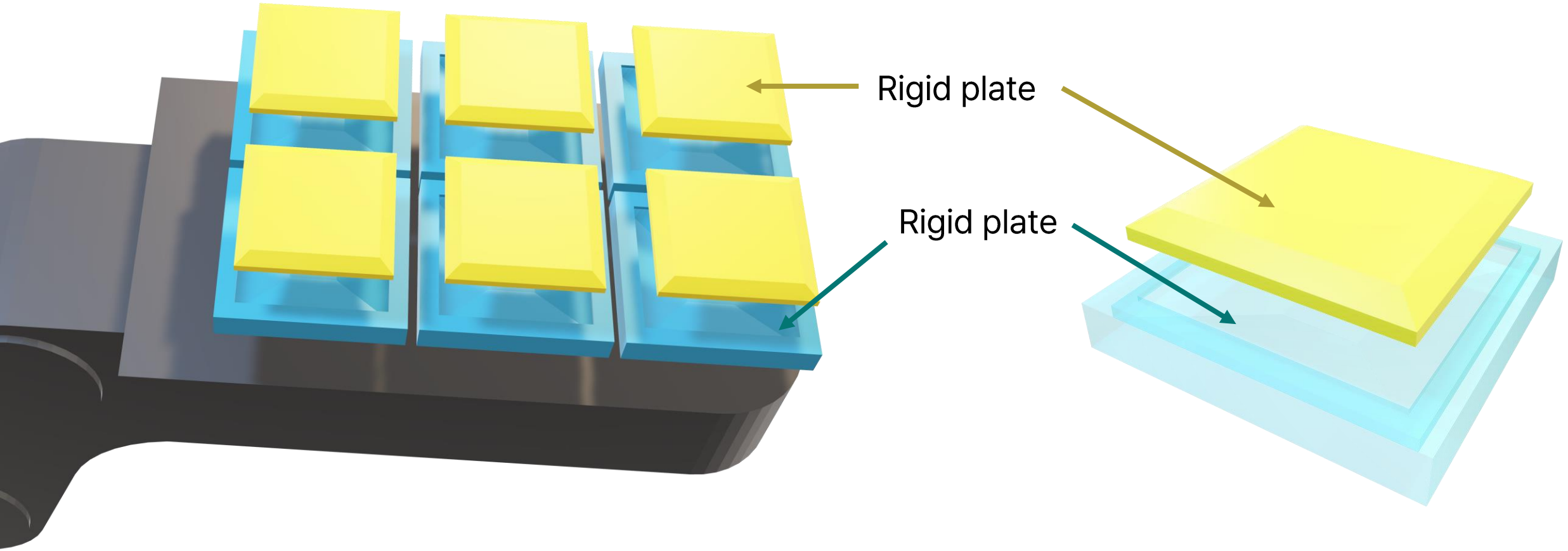




❖ Method

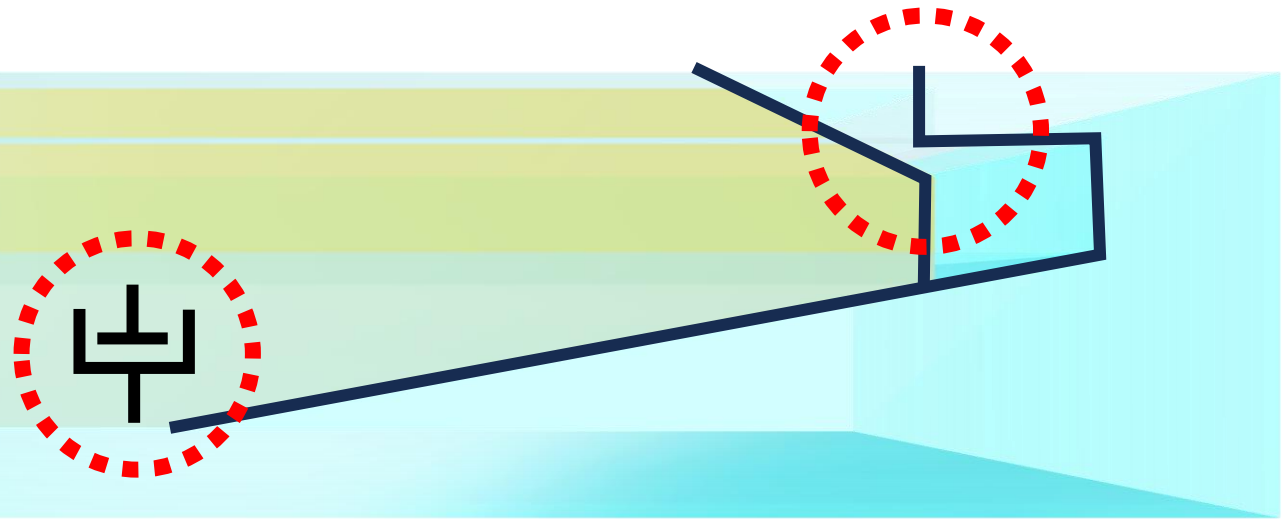
❖ DetactGym: Tactile sensor

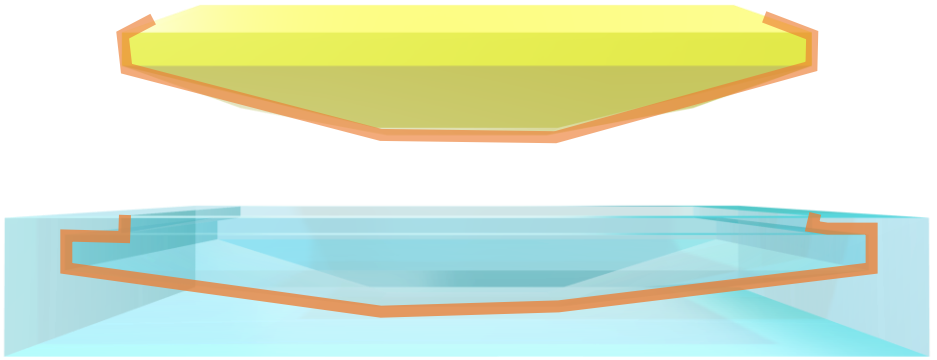




❖ Method

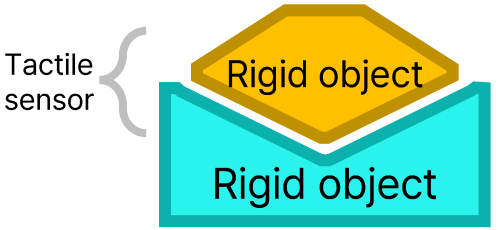
❖ DetactGym: Tactile sensor

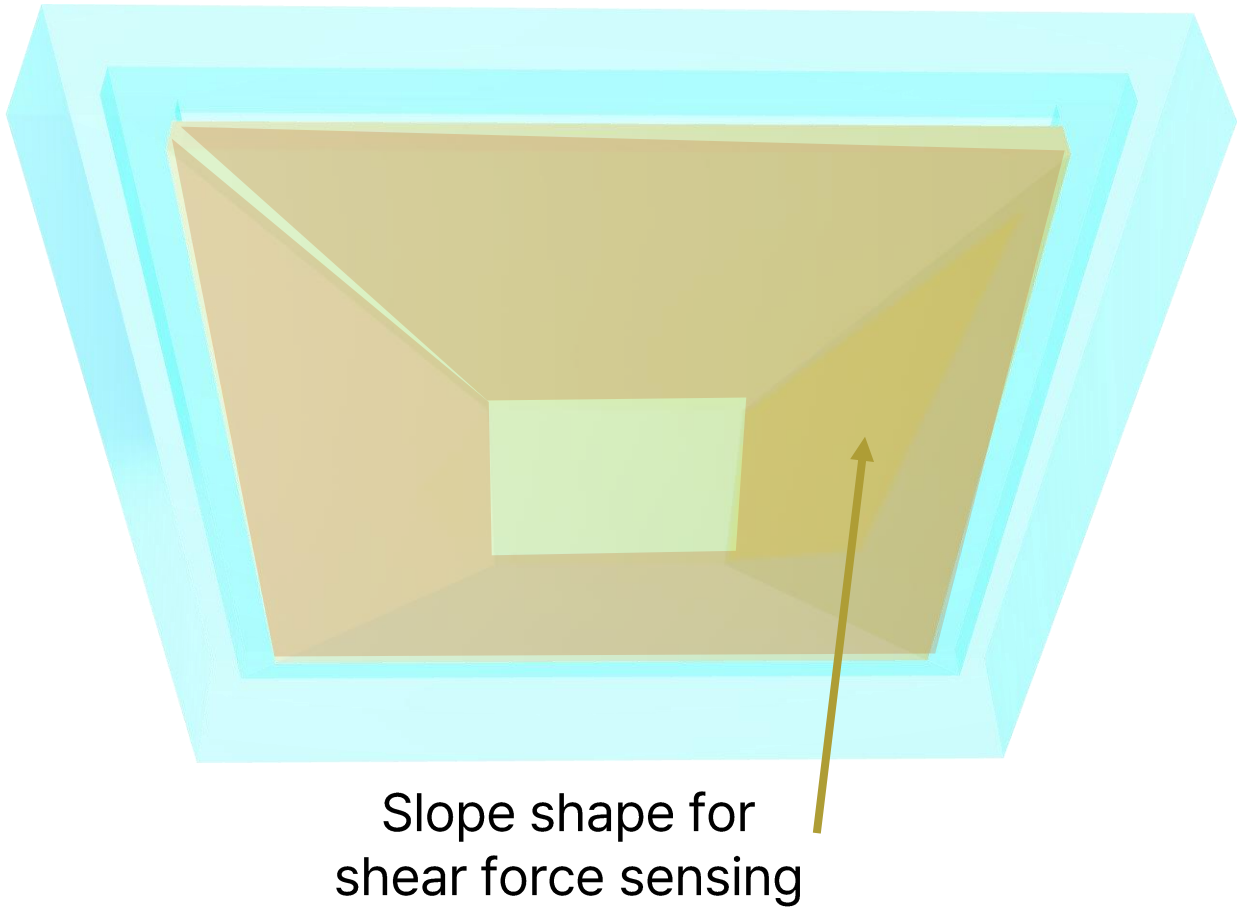


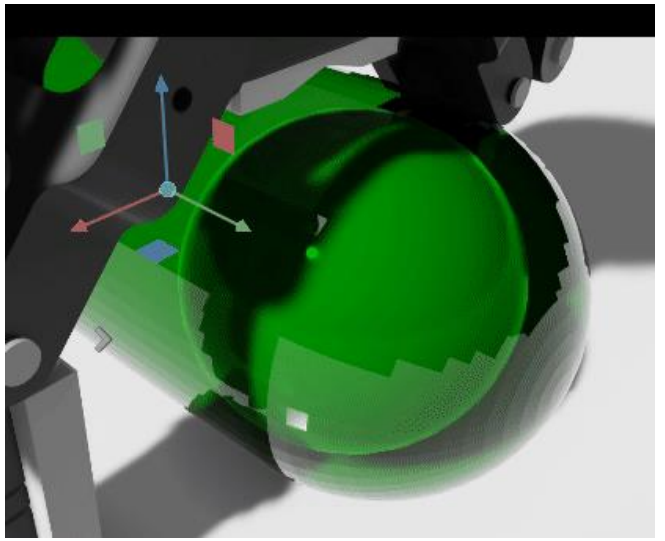


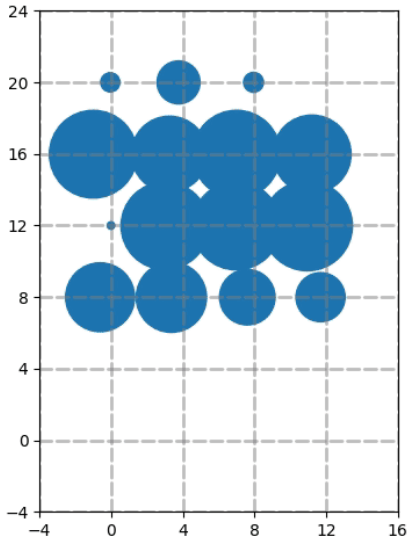
❖ Method

❖ DetactGym: Tactile sensor



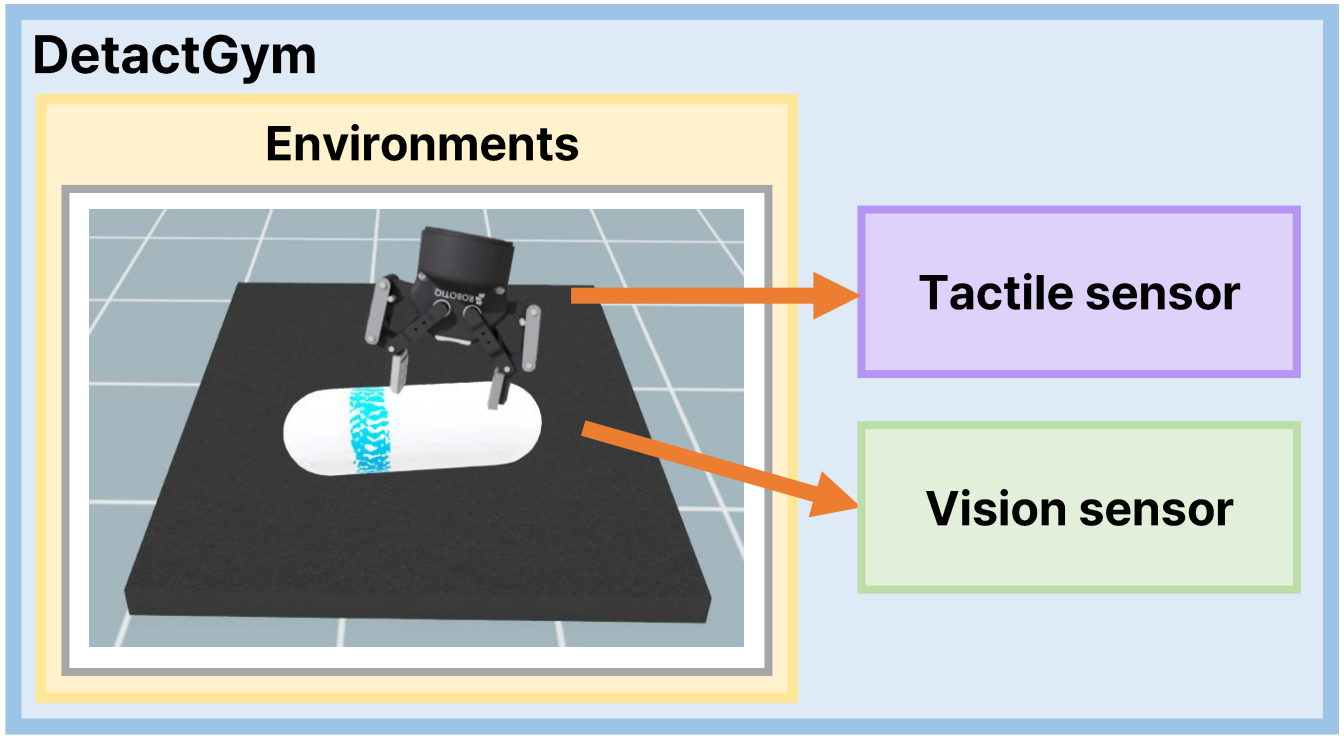






❖ Method

❖ DetactGym: Overall Framework

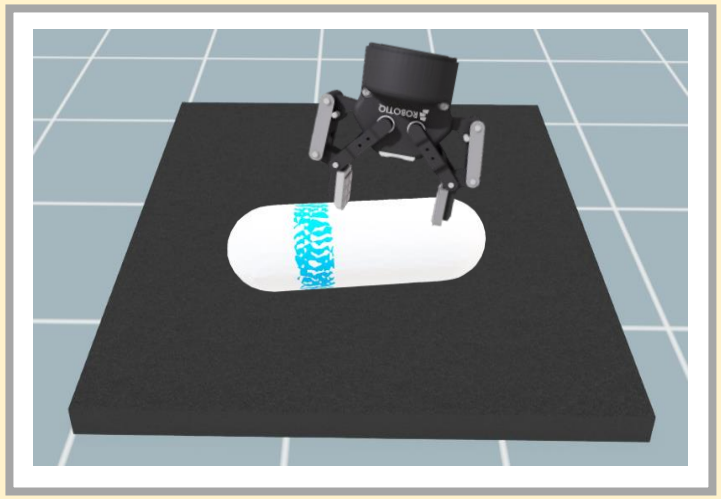


❖ Method

❖ DetactGym: Vision sensor

DetactGym

Environments

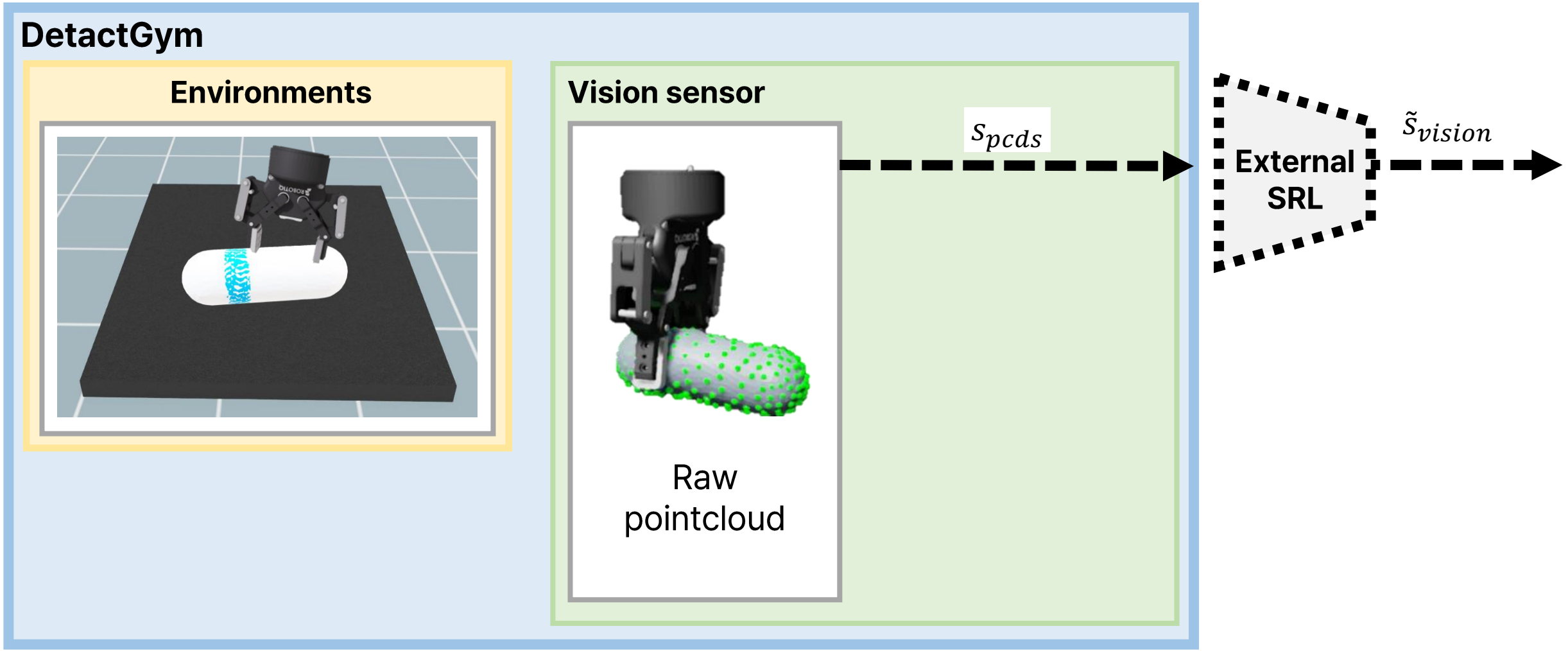


Vision sensor

Raw pointcloud

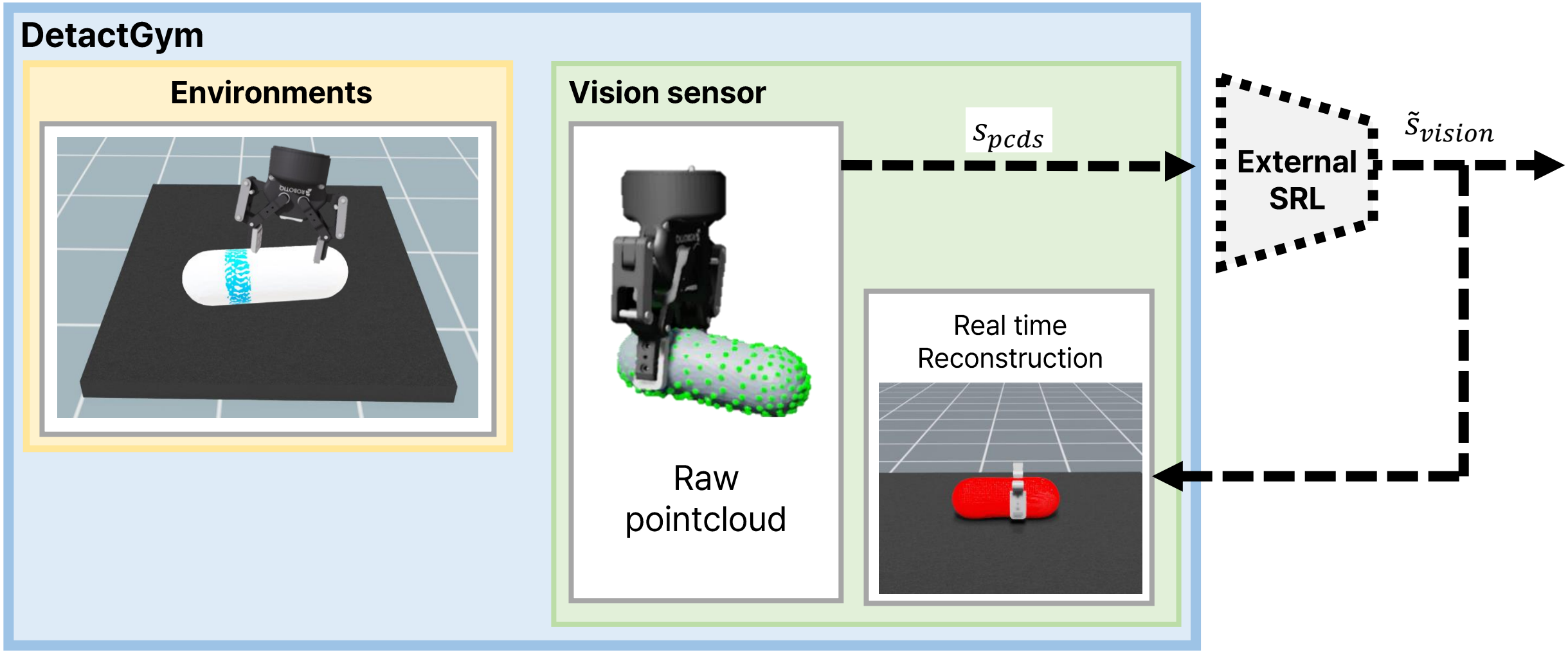
❖ Method

❖ DetactGym: Vision sensor



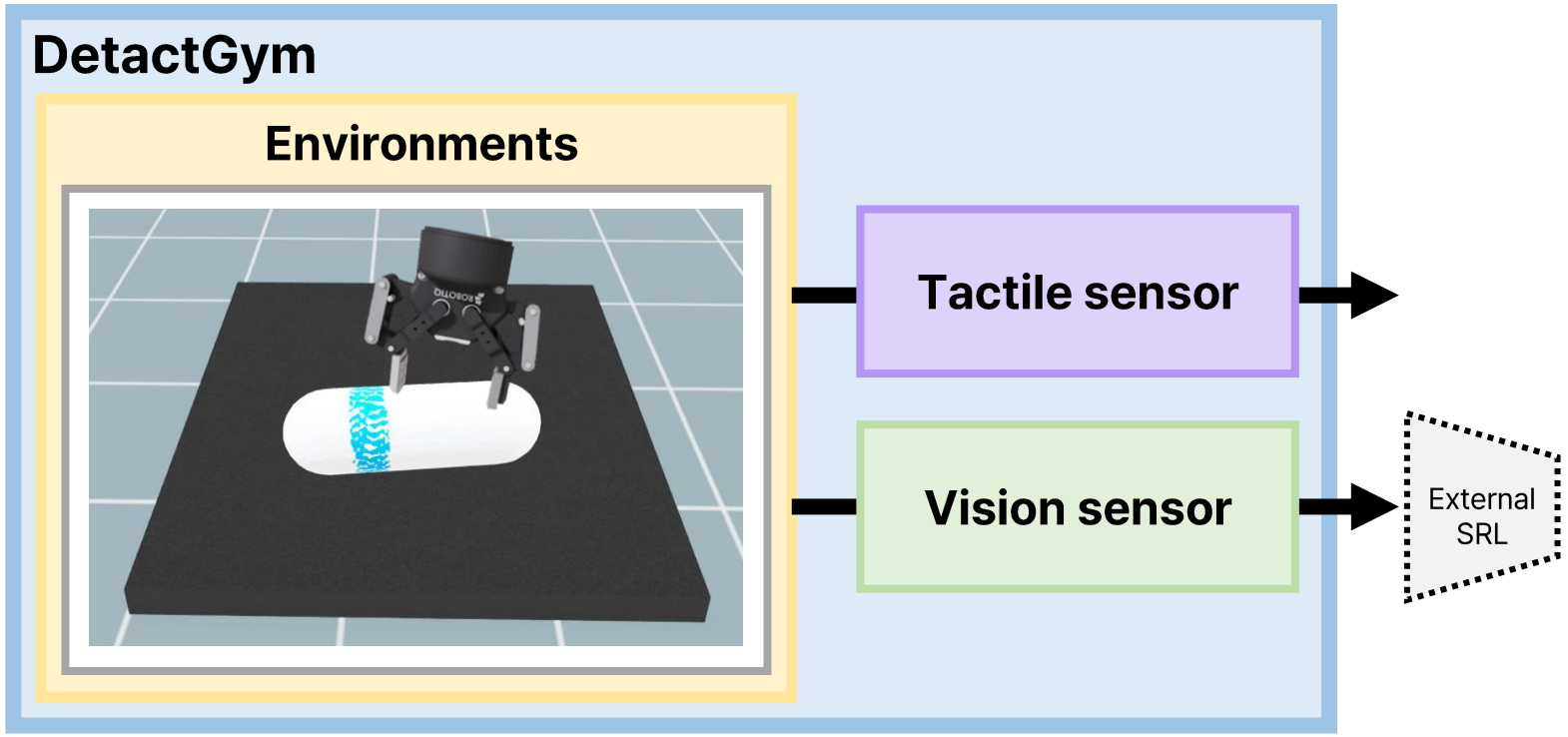
❖ Method

❖ DetactGym: Vision sensor



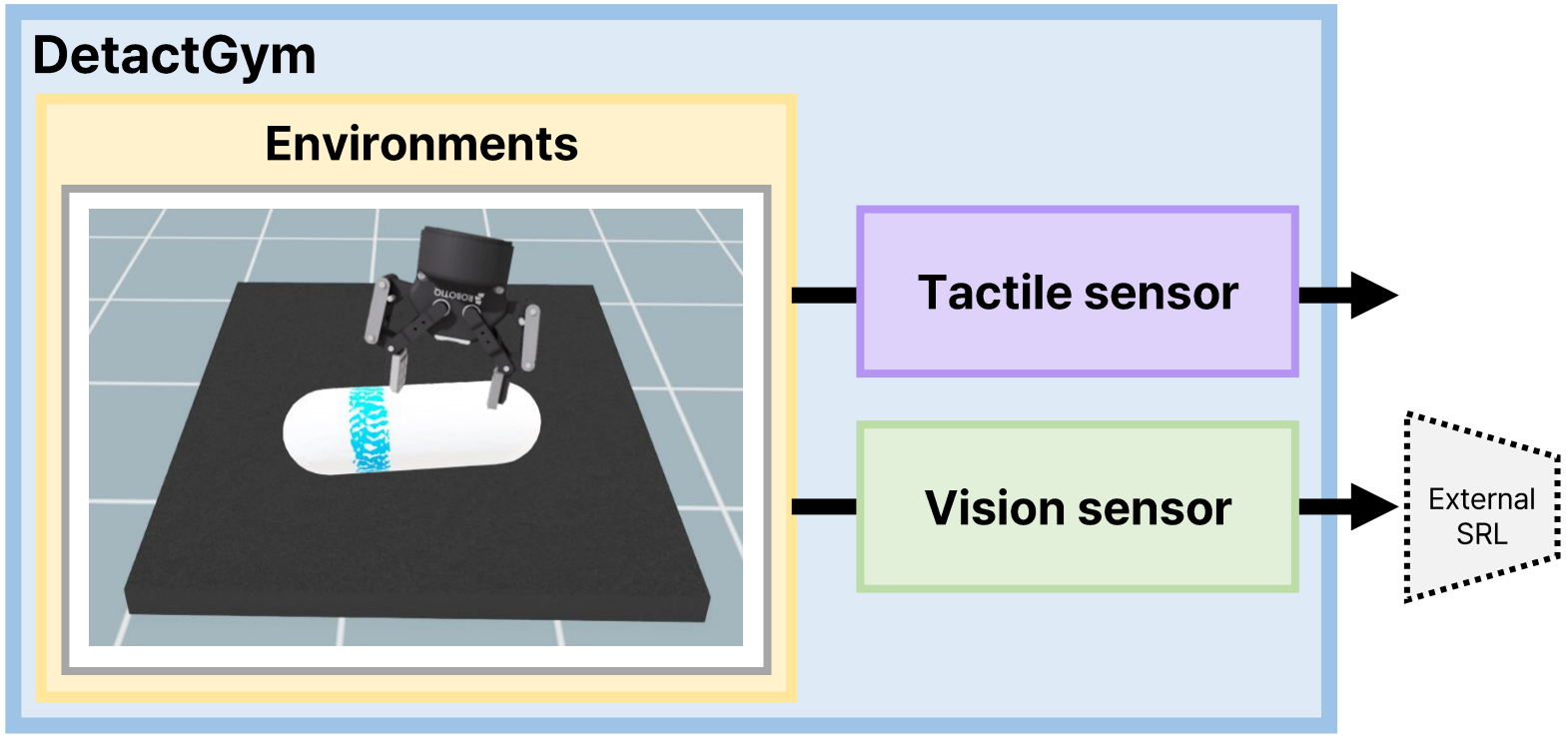
❖ Method

❖ DetactGym: Overall Framework

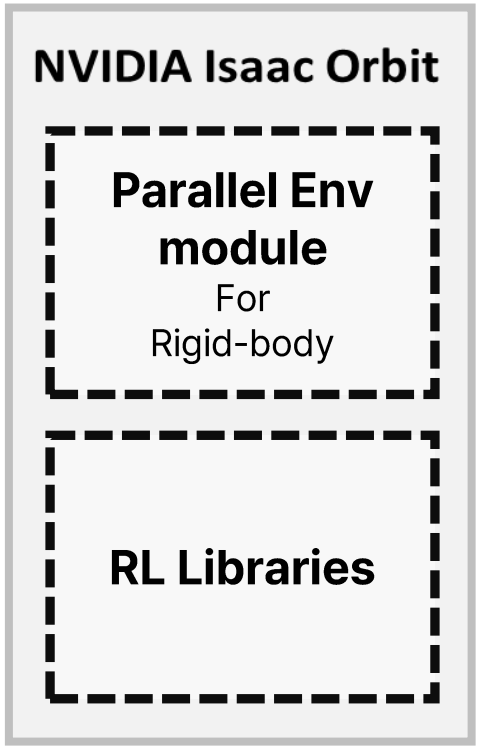


❖ Method

❖ DetactGym: Overall Framework



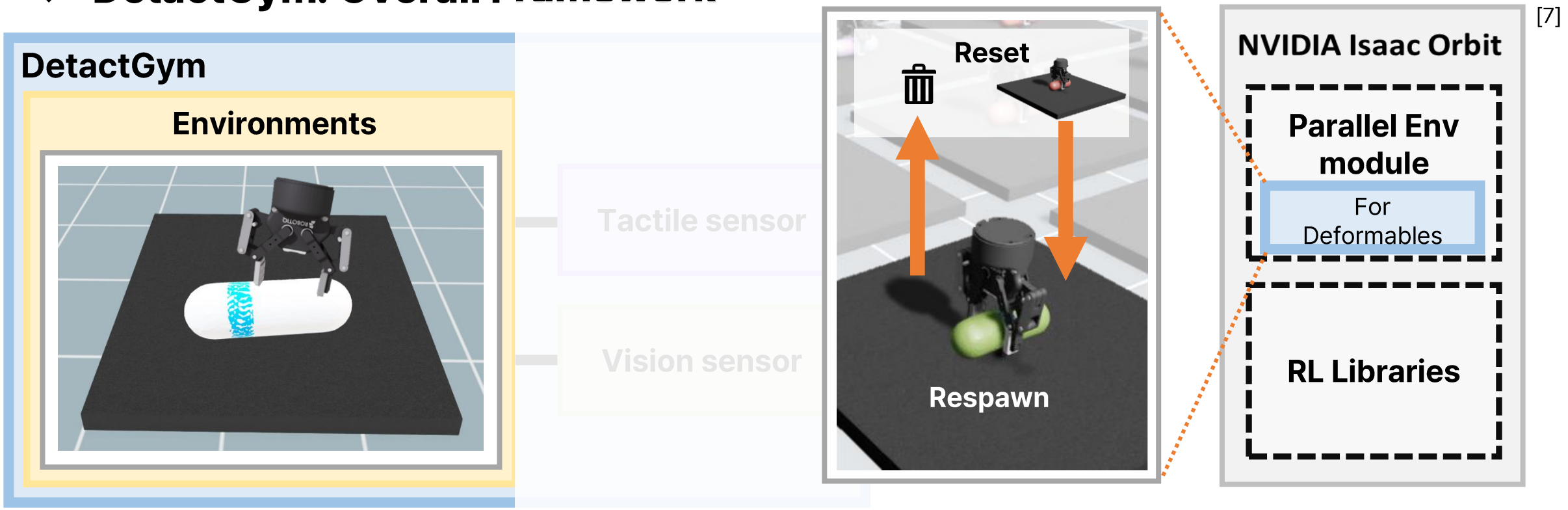
[7]



[7] Mittal, Mayank, et al. "Orbit: A unified simulation framework for interactive robot learning environments." IEEE Robotics and Automation Letters (2023).

❖ Method

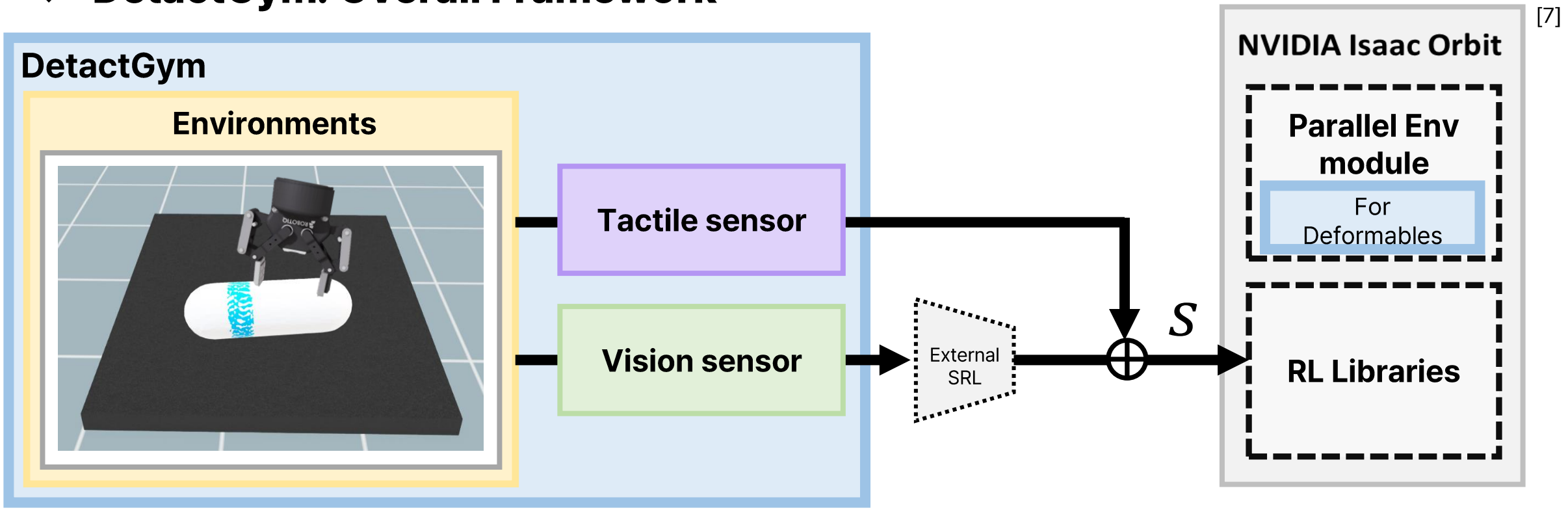
❖ DetactGym: Overall Framework



[7] Mittal, Mayank, et al. "Orbit: A unified simulation framework for interactive robot learning environments." IEEE Robotics and Automation Letters (2023).

❖ Method

❖ DetactGym: Overall Framework

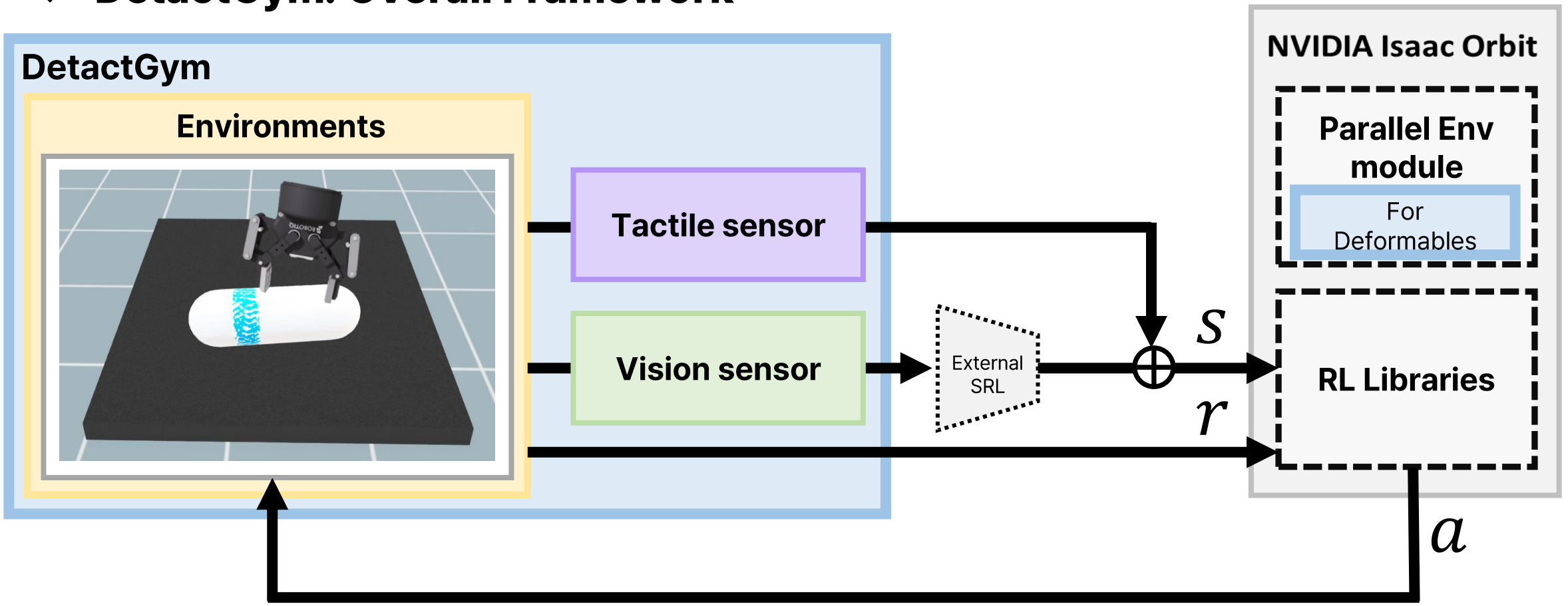


[7]

[7] Mittal, Mayank, et al. "Orbit: A unified simulation framework for interactive robot learning environments." IEEE Robotics and Automation Letters (2023).

❖ Method

❖ DetactGym: Overall Framework



[7]

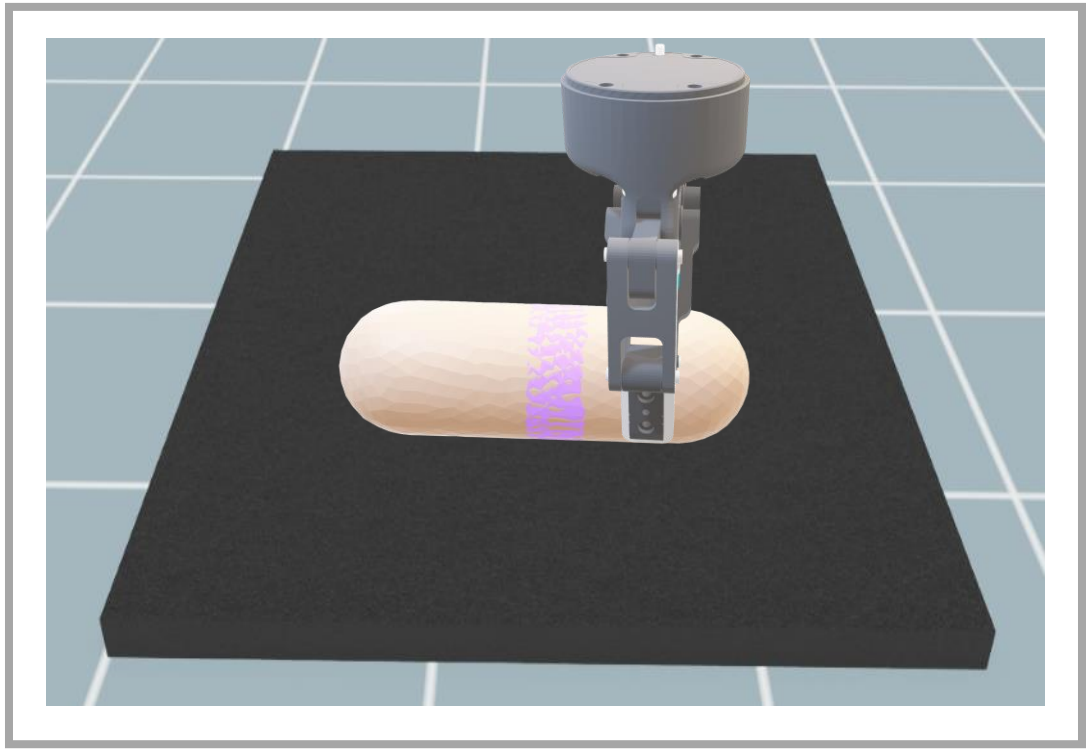
[7] Mittal, Mayank, et al. "Orbit: A unified simulation framework for interactive robot learning environments." IEEE Robotics and Automation Letters (2023).

❖ Experiments

❖ Soft-lift-v0

Goal

- **Minimize deformation** during lift deformable objects



Minimize deformation

Fail Success

❖ Experiments

❖ Soft-lift-v0

Goal

- **Minimize deformation** during lift deformable objects

Problem: How to evaluate deformation?

➢ Provide '**Deform Chamfer Distance**' metric

* Appendix: Chamfer Distance

❖ Experiments

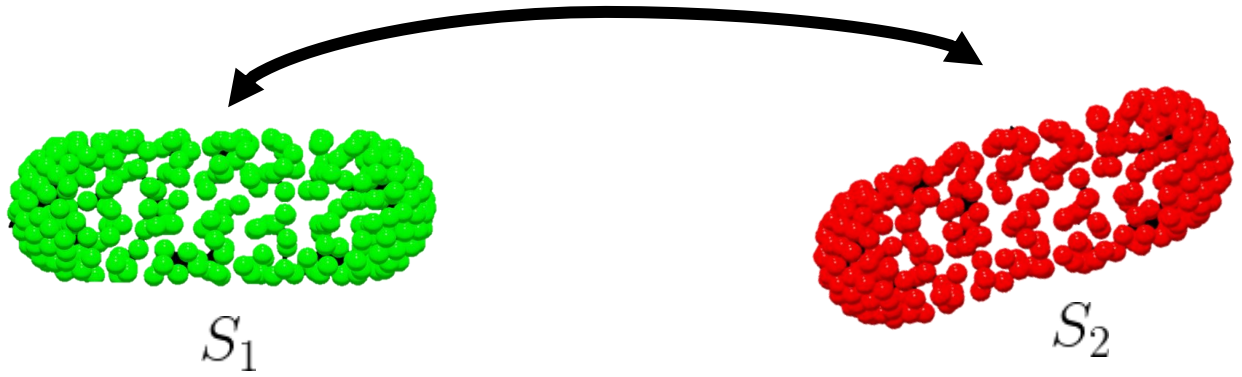
❖ Soft-lift-v0

Chamfer Distance of Deformable Object (CDDO)

- **Chamfer Distance (CD)**

  Measure the shape **dissimilarity between point clouds** [8]

$$d_{CD} = (S_1, S_2) = \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|_2^2 + \sum_{y \in S_2} \min_{x \in S_1} \|x - y\|_2^2$$



[8]Lin, Fangzhou, et al. "Hyperbolic chamfer distance for point cloud completion." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.

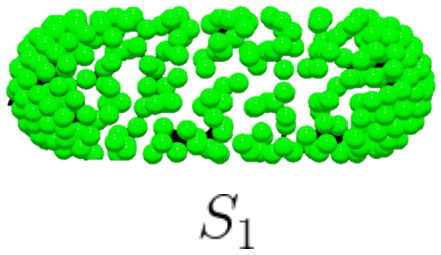
❖ Experiments

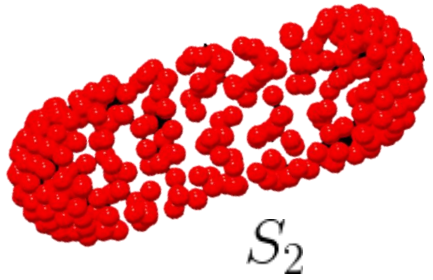
❖ Soft-lift-v0

Chamfer Distance of Deformable Object (CDDO)

- **Chamfer Distance (CD)**

  Measure the shape **dissimilarity between point clouds** [8]



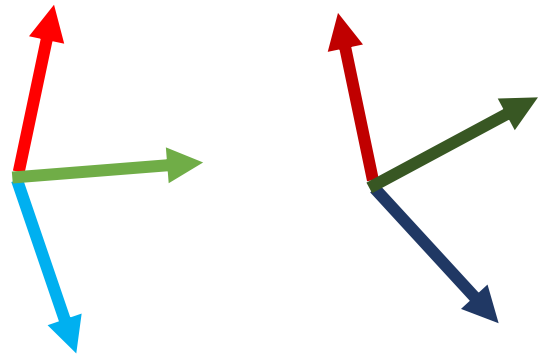


- CD consider **translation, rotation, deformation ...**

[8]Lin, Fangzhou, et al. "Hyperbolic chamfer distance for point cloud completion." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.

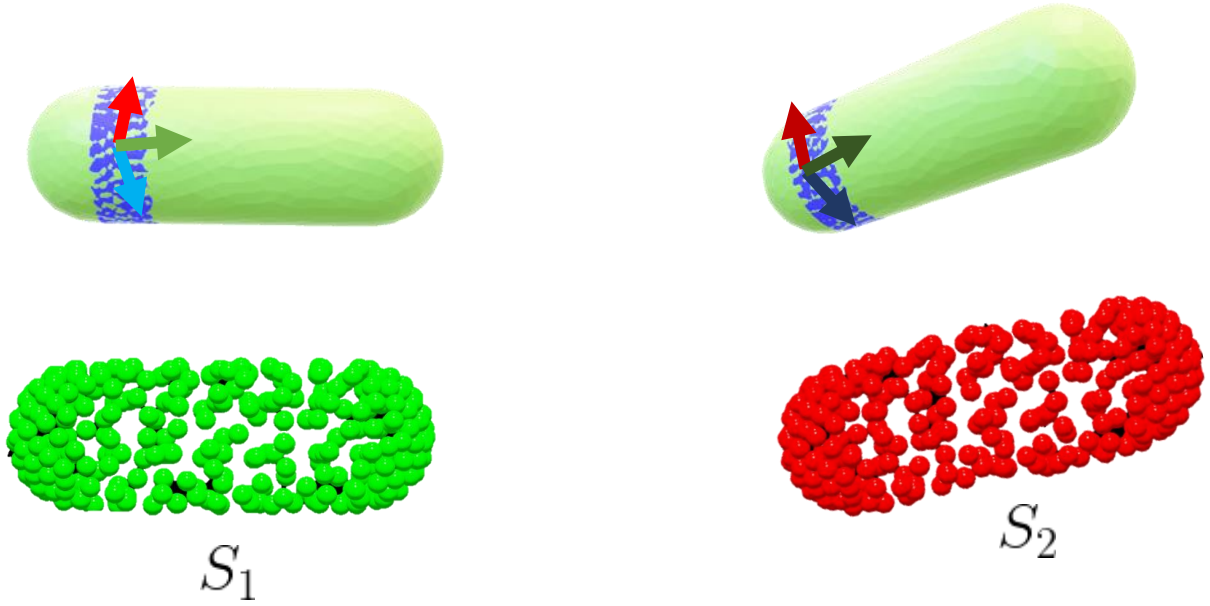
❖ Experiments

❖ Soft-lift-v0



Chamfer Distance of Deformable Object (CDDO)

Measure the shape **deformation between point clouds**

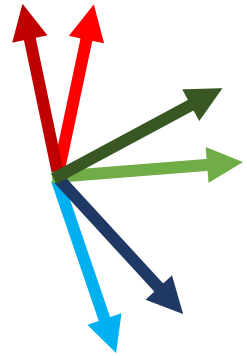


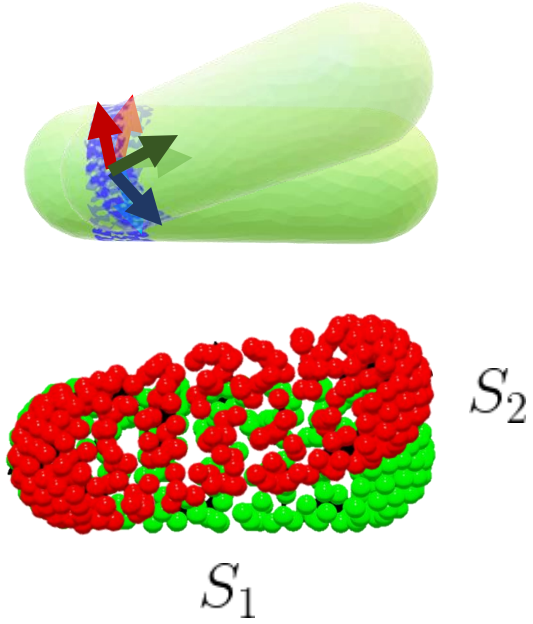
❖ Experiments

❖ Soft-lift-v0

Chamfer Distance of Deformable Object (CDDO)

Measure the shape **deformation between point clouds**





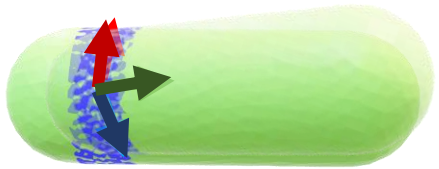
(1) Remove difference of point cloud by **translation**

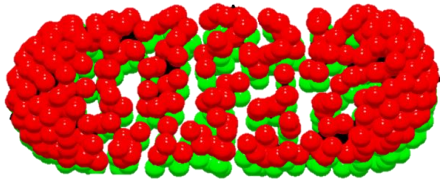
❖ Experiments

❖ Soft-lift-v0

Chamfer Distance of Deformable Object (CDDO)

Measure the shape **deformation between point clouds**





S_2

S_1

(2) Remove difference of point cloud by **rotation**

$$S_2' = \begin{bmatrix} R & T_{xyz} \\ 0 & 1 \end{bmatrix} y_{\in S_2}$$

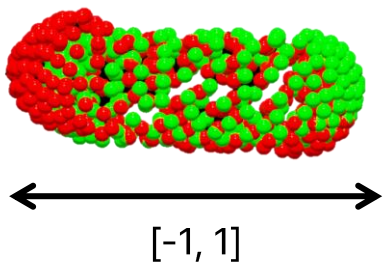
[8]Lin, Fangzhou, et al. "Hyperbolic chamfer distance for point cloud completion." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.

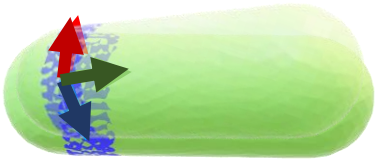
❖ Experiments

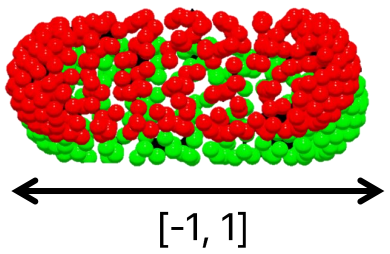
❖ Soft-lift-v0

Chamfer Distance of Deformable Object (CDDO)

Measure the shape **deformation between point clouds**







(3) Remove difference of point cloud by
scale through normalization

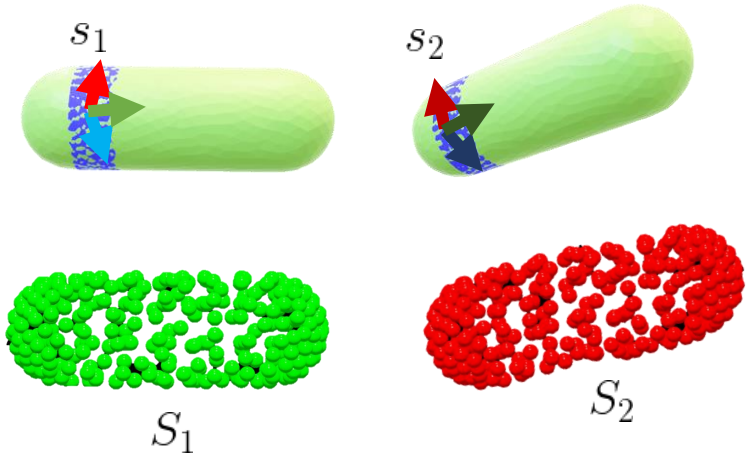
$$y'_n = \frac{y' - x_{min}}{x_{max} - x_{min}}, y' \in S'_2 \quad x_n = \frac{x - x_{min}}{x_{max} - x_{min}}, x \in S_1$$

❖ Experiments

❖ Soft-lift-v0

Chamfer Distance of Deformable Object (CDDO)

Measure the shape **deformation between point clouds**



$$d_{CDC} = \sum_{y'_n} \min_{x_n} \|x_n - y'_n\|_2^2 + \sum_{x_n} \min_{y'_n} \|x_n - y'_n\|_2^2$$

❖ Experiments

❖ Soft-lift-v0

Goal

- **Minimize deformation** (CDDO < thereshold) during lift deformables

Problem: How evaluate deformation?

➢ Provide '**Deform Chamfer Distance**' metric

$$d_{CDC} = \sum_{y'_n} \min_{x_n} \left\| x_n - y'_n \right\|_2^2 + \sum_{x_n} \min_{y'_n} \left\| x_n - y'_n \right\|_2^2$$

* Appendix: Chamfer Distance, Deform Chamfer Distance

❖ Experiments

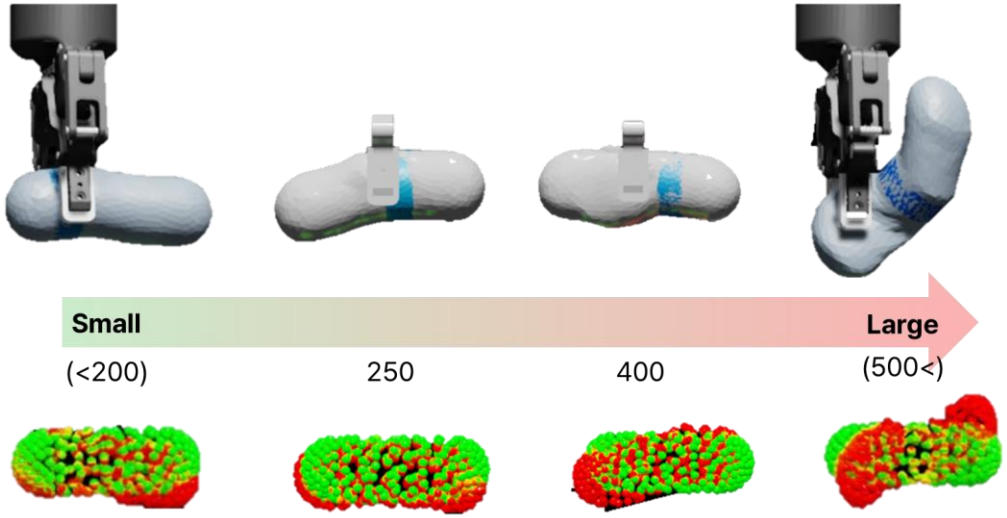
❖ Soft-lift-v0

Goal

- **Minimize deformation** (CDDO < thereshold) during lift deformables

Problem: How evaluate deformation?

➢ Provide ‘Chamfer Distance of Deformable Object’ metric



* Appendix: Chamfer Distance, Deform Chamfer Distance

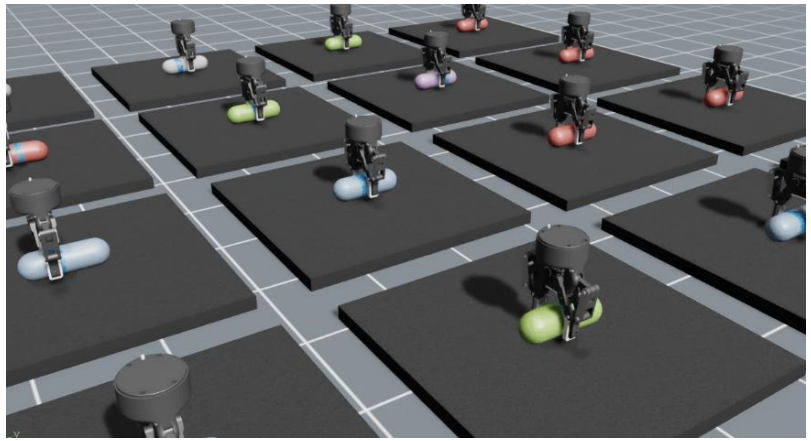
❖ Experiments

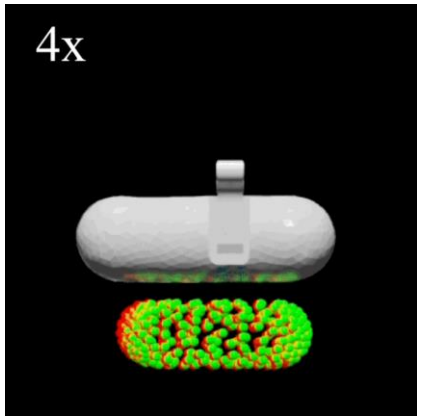
❖ Soft-lift-v0

- Evaluating success rate of *Soft-lift-v0* with 200 episodes in visuotactile, tactile, force, and vision

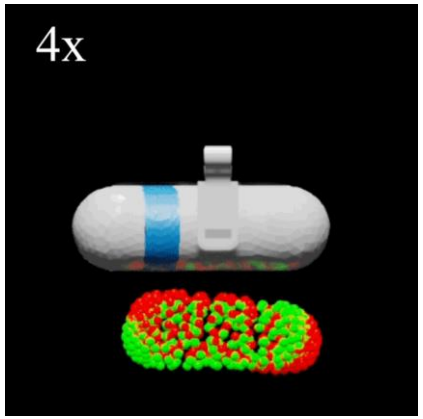
Rate of Success	Visuotactile	Tactile 2*3 force array	Force	Vision
PPO	**80%**	70%	55%	54%

* The training spans 24,000 steps across four parallel environments





Success case



Fail case

❖ Experiments

❖ Soft-lift-v0

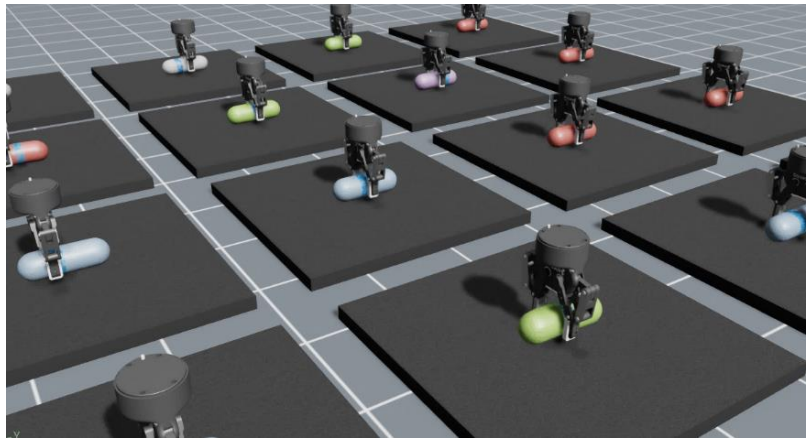
- Evaluating success rate of *Soft-lift-v0* with 200 episodes in visuotactile, tactile, force, and vision

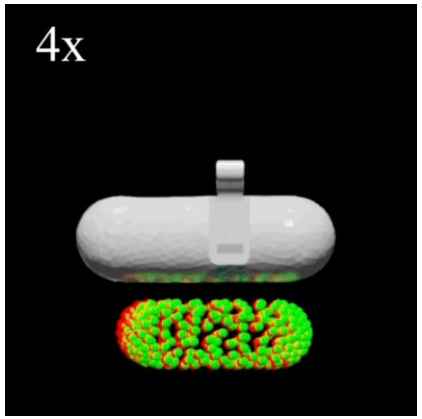
Rate of Success	Visuotactile	Tactile 2*3 force array	Force	Vision
PPO	**80%**	70%	55%	54%

* The training spans 24,000 steps across four parallel environments

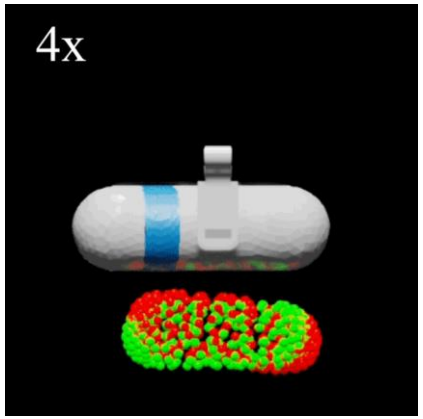
Tactile
2*3 force array

Force





Success case



Fail case

❖ Conclusion

❖ Contributions

- **Provide a simulated testbed for training and assessing DOM skill using visuotactile sensing**
- DetactGym enables learning skills in soft-lift-v0 with visual or tactile sensing through RL.
- We highlight the effectiveness of tactile information over visual cues in manipulating deformable while minimizing deformation.

❖ Next research plan

- Provide tactile sensing from deformation energy..
- Develop a complex environment that assess tactile sensors' benefits on DOM
- Provide realistic vision sensor(partially point cloud) for parallel RL environment

Thank you for your attention

Any question or feedback will be welcome