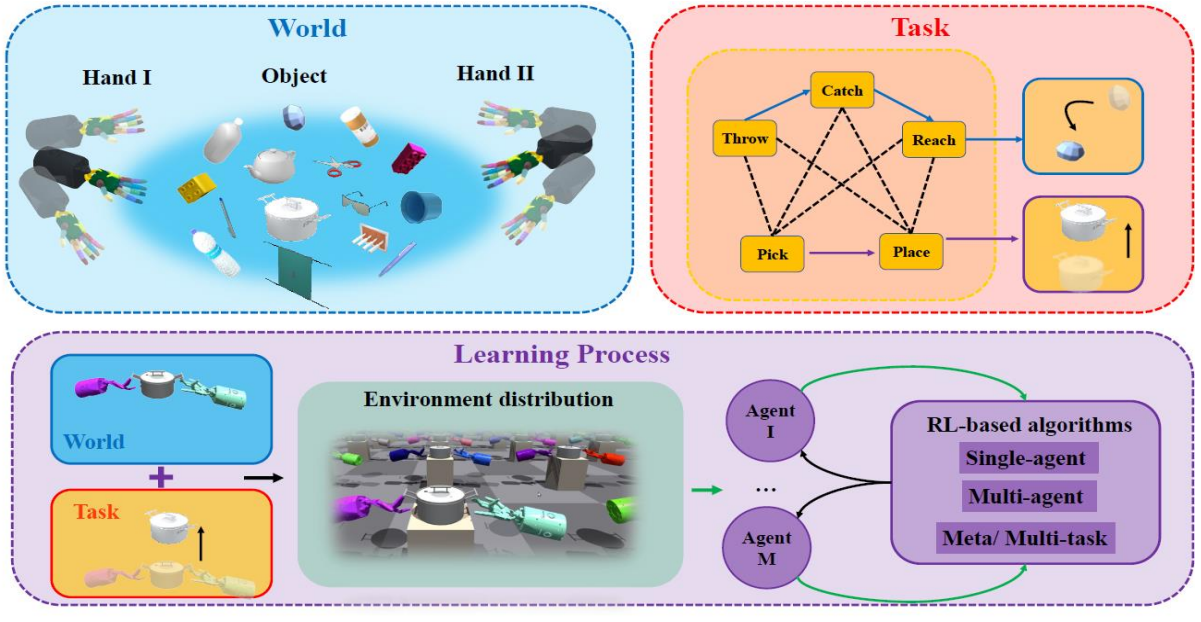
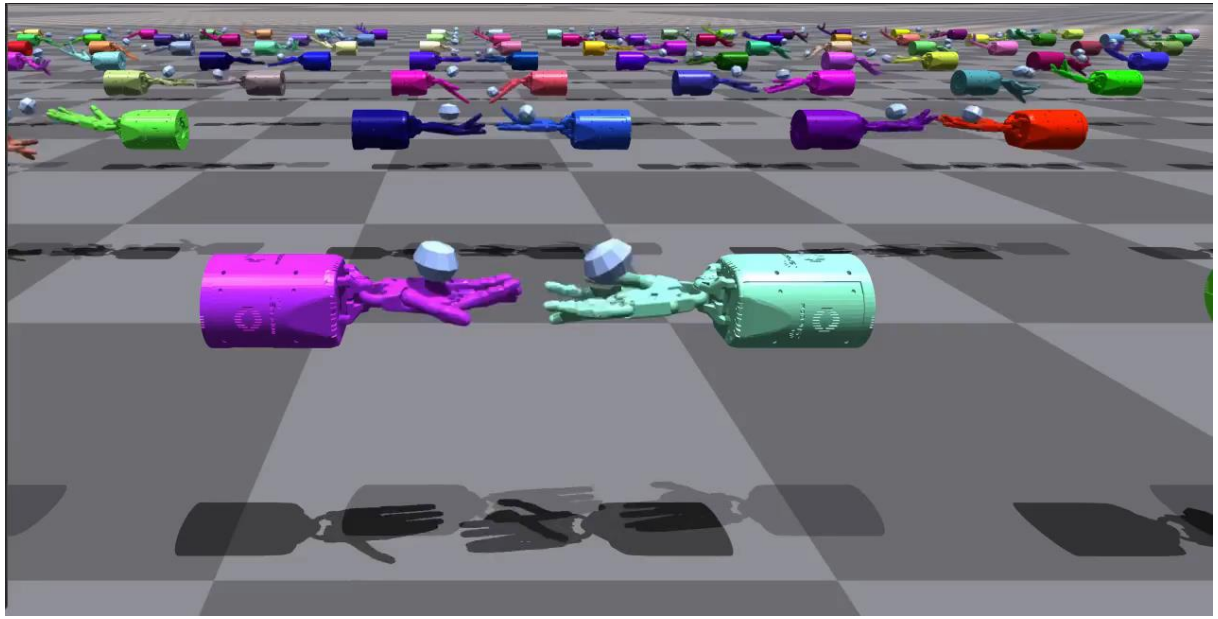
Research Proposal

Yuanpei Chen

Bi-DexHands: Bimanual Dexterous Manipulation with RL

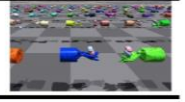




- ✓ **Isaac Efficiency**： Using IsaacGym for parallel environment simulation, greatly improves the sampling efficiency of reinforcement learning algorithms.
- ✓ **RL/MARL Benchmark**： We provide the first bimanual manipulation task environment for common RL, MARL, offline RL, multi-task RL, and Meta RL practitioners, along with a comprehensive benchmark for SOTA continuous control model-free RL methods.
- ✓ **Task Generalization**： We introduced a variety of dexterous manipulation tasks (e.g., hand over, lift up, throw, place, put...) as well as enormous target objects from the YCB and SAPIEN dataset, thus allowing meta-RL and multi-task RL algorithms to be tested on the task generalization front.
- ✓ …

For more information, please refer to the github repository： https://github.com/PKU-MARL/DexterousHands

Bi-DexHands: Bimanual Dexterous Manipulation with RL

Task Name	Human's Skill Description	Age (months)	Demo
Push Block Easy	Child extends one or both arms forward and touches the block with any part of either hand	5-6 [18, Chapter 3]	
Open Scissor & Open Pen Cap Easy	They use one hand to hold a toy and the other hand manipulate it	7 [51, Chapter 4]	
Turn Button ON/OFF Easy	They can push and squish soft stuff or push hard things, like a button on a toy phone or popup toy	11 [52, 11 months]	
Swing Cup Easy	They can turn a ball on their toy mobile, a steering wheel on a toy car, or the faucet in the tub	11 [52, 11 months]	
Lift Pot & Lift Cup Easy	They can put a sippy cup to their mouth to drink	12 [52, 12 months]	
Door Open & Close Easy	Toddlers can open and close cupboards and oven doors	13 [52, 13 months]	
Re-Orientation Medium	Infant further refines this ability to differentiate individual finger movement and manipulate objects	18 [51, Chapter 4]	
Stack Block(2,6,8) Medium	Child stacks at least 2/6/8 blocks in any trial.	2:22-28 6,8:33-42 [18, Chapter 3]	
Pull a Ball into Bucket Medium	Child place s 10 pellets in the bottle in 60 seconds or less, one pellet at a time.	22-28 [18, Chapter 3]	
Open Bottle Cap Medium	Uses hands to twist things, like turning doorknobs or unscrewing lids	30 [53, Table 6]	
Catch Underarm Hard	Catches a large ball most of the time	48 [53, Table 6]	
Pour Water Hard	Serves himself food or pours water, with adult supervision	48 [53, Table 6]	
Two Catch Underarm Hard	Some adults can throw objects between two hands like magic	adult	

Bi-DexHands: Bimanual Dexterous Manipulation with RL

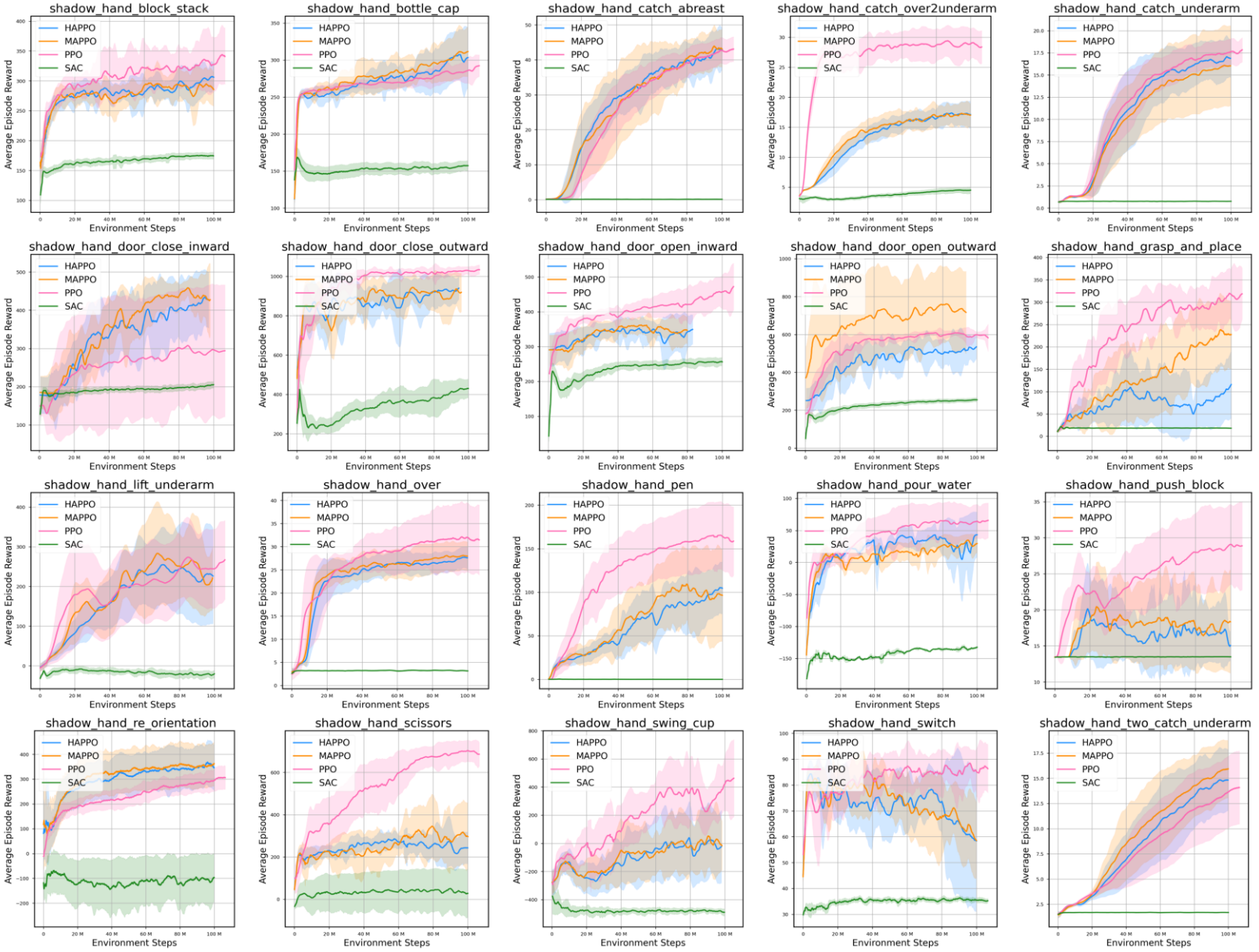


Table 2: Mean and standard deviation of FPS (frame per second) of the environments in Bi-DexHands.

Algorithms	CatchUnderarm	CatchOver2Underarm	CatchAbreast	TwoCatchUnderarm
PPO	35554 ± 613	35607 ± 344	35164 ± 450	32285 ± 898
HAPPO	23929 ± 98	23827 ± 135	23456 ± 255	23205 ± 168

Table 3: Normalized score in offline tasks.

Tasks	Datasets	Online PPO	BC	BCQ	TD3+BC	IQL
Hand Over	random	100.0	0.7 ± 0.2	1.0 ± 0.1	0.9 ± 0.2	0.7 ± 0.4
	replay	100.0	17.5 ± 3.5	61.6 ± 4.9	$\mathbf{70.1} \pm 2.1$	43.1 ± 2.3
	medium	100.0	61.6 ± 1.0	$\mathbf{66.1} \pm 1.9$	$\mathbf{65.8} \pm 2.2$	57.4 ± 1.5
	medium-expert	100.0	63.3 ± 1.4	$\mathbf{81.7} \pm 4.9$	$\mathbf{84.9} \pm 5.3$	67.2 ± 3.6
Door Open Outward	random	100.0	2.1 ± 0.6	23.8 ± 2.9	$\mathbf{34.9} \pm 4.3$	3.8 ± 1.0
	replay	100.0	36.9 ± 4.3	48.8 ± 4.5	$\mathbf{60.5} \pm 2.6$	31.7 ± 2.0
	medium	100.0	63.9 ± 0.7	60.1 ± 2.3	$\mathbf{66.3} \pm 0.7$	56.6 ± 1.2
	medium-expert	100.0	69.0 ± 6.4	$\mathbf{73.7} \pm 4.5$	$\mathbf{71.9} \pm 3.5$	53.8 ± 1.8

Table 4: The average reward of all tasks for MT1, ML1, MT4, ML4, MT20, and ML20 on 10 seeds.

Method	MT1	MT4	MT20	Method	ML1		ML4		ML20	
					train	test	train	test	train	test
Ground Truth	15.2	24,3	32.5	Ground Truth	15.0	15.8	28.0	13.1	33.7	26.1
Multi-task PPO	9.4	5.4	8.9	ProMP	0.95	1.2	2.5	0.5	0.02	0.36
Random	0.61	1.1	-2.5	Random	0.59	0.68	1.5	0.24	-2.9	0.27

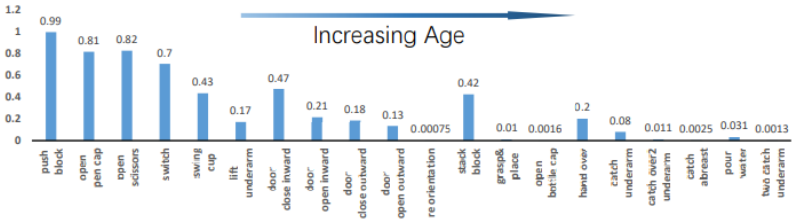


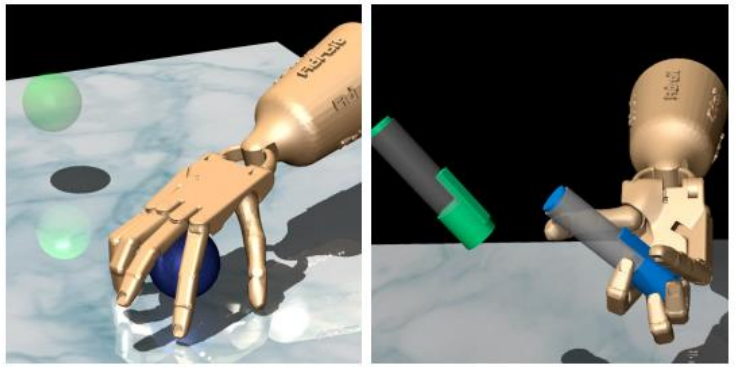
Figure 4: The normalized reward run by the MTPPO algorithm under the MT20 setting. The tasks from left to right according to the increase of corresponding age. The normalized score is computed by score $= \frac{\text{reward-random reward}}{\text{ground truth reward-random reward}}$.

Trends and challenges in robot manipulation

---- Learning for dexterous manipulation

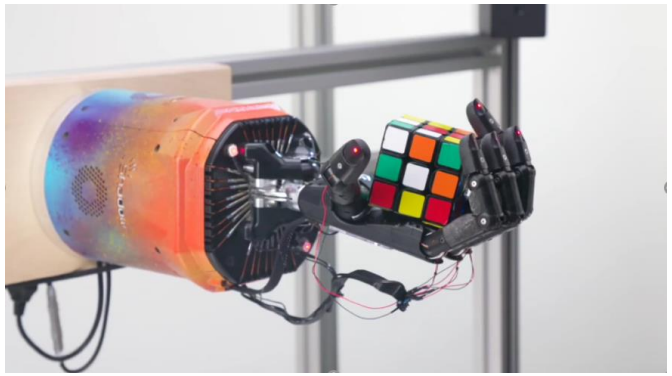
From Bayley Scales of Infant Development: Third Edition | SpringerLink:

Basic skills

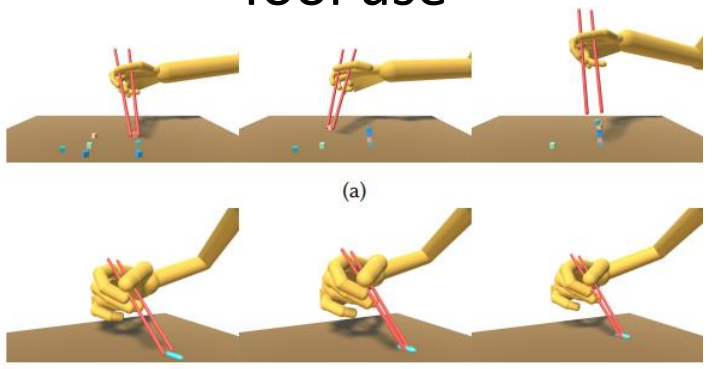


Precise control

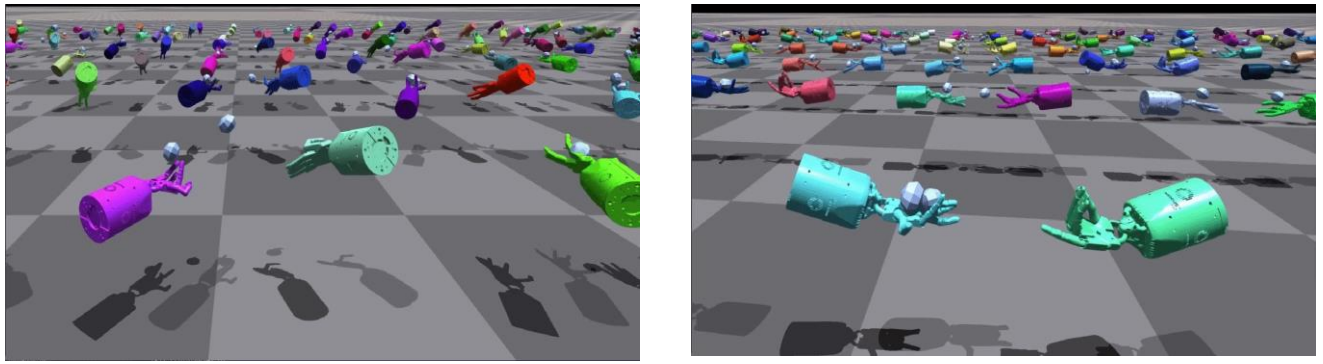
Finger dexterity



Tool use



Dynamic interaction (ours)

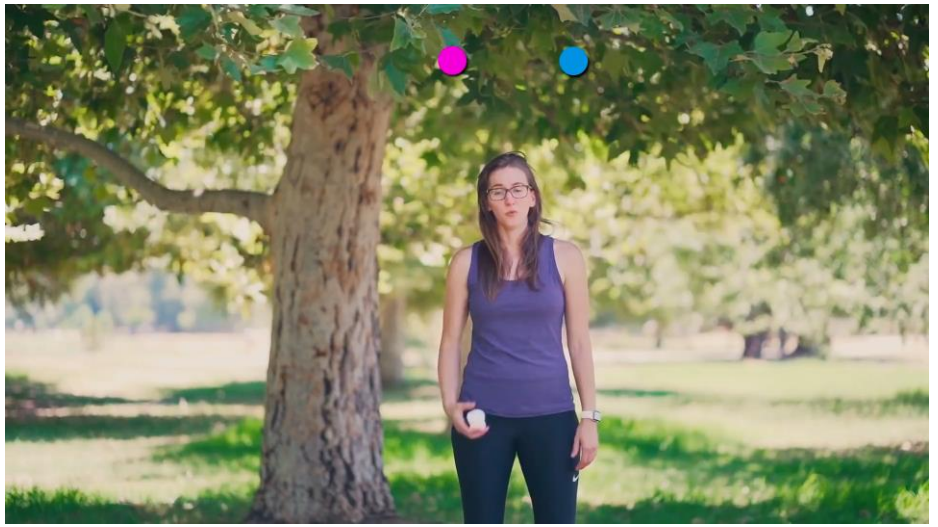


Learn to Juggle 3 Balls

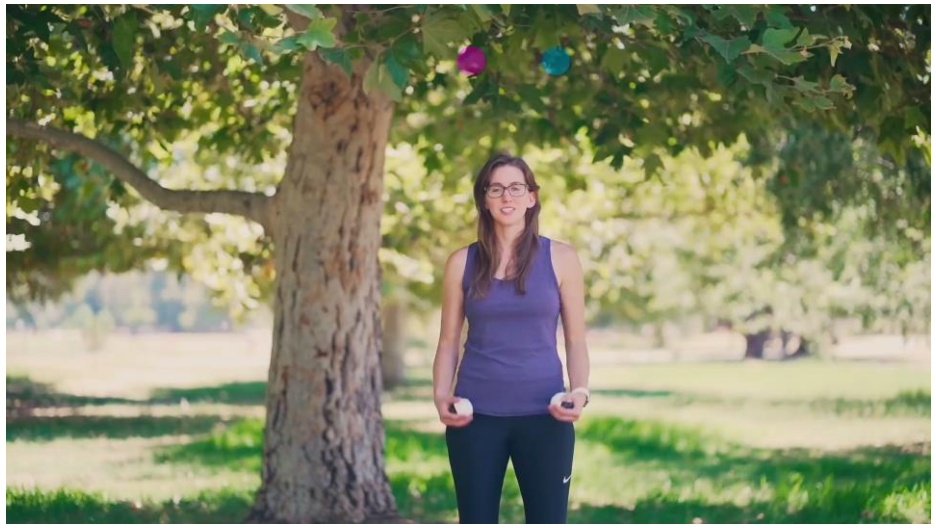
Learn how to juggle 3 balls with professional juggler Taylor Glenn! In-depth explanations, tips, and troubleshooting to help you learn to juggle like an awesome person.

[Learn to JUGGLE 3 BALLS - Beginner Tutorial - YouTube]

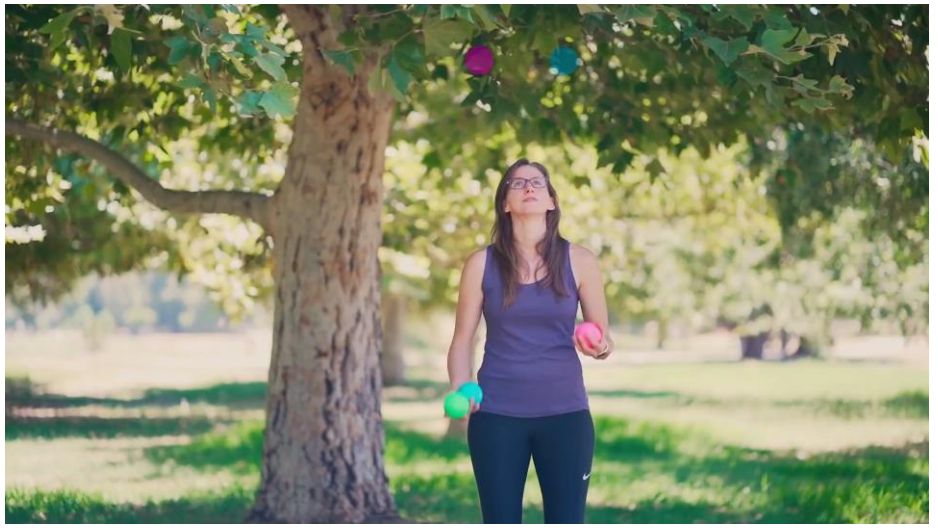
1 ball



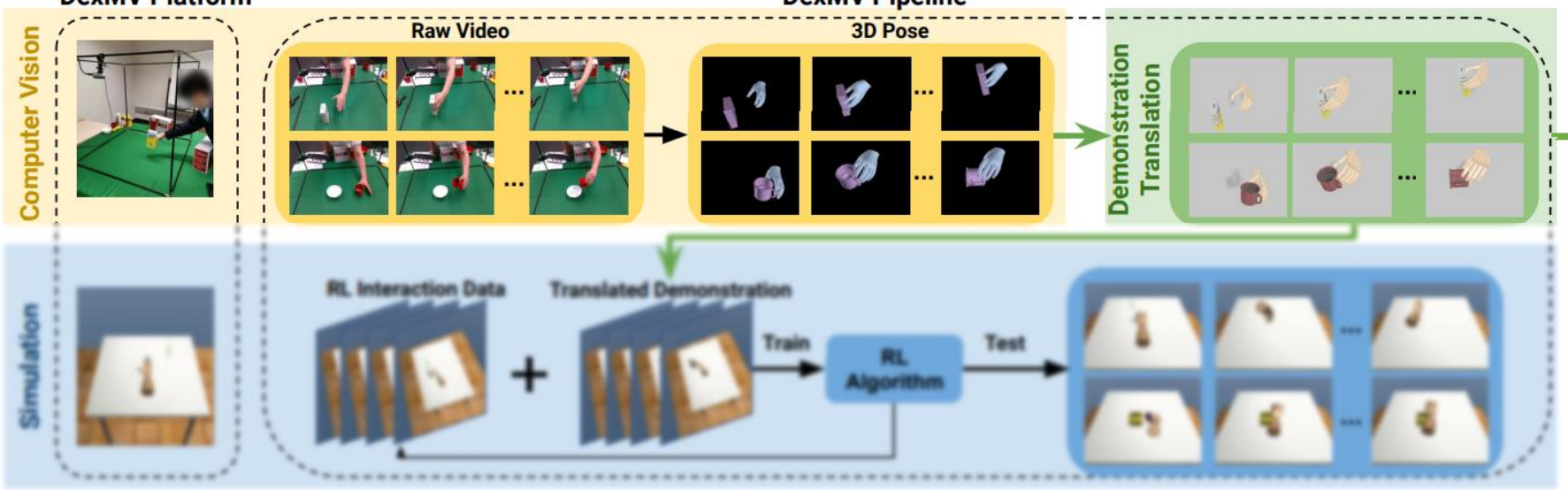
2 balls



3 balls



Learning Pipeline



DAPG

Other technologies that may be useful:

Affordance

Hierarchical RL

Multi-Agent RL

Mixture of Expert

Purturbation

Further more

More skills

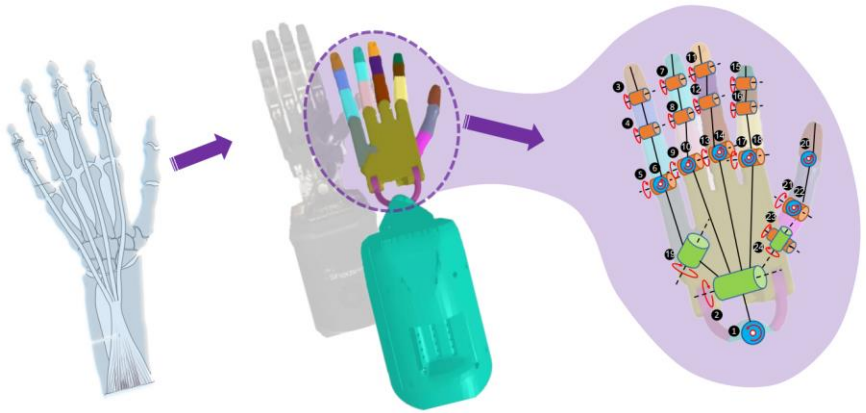
More balls

More objects

More agents

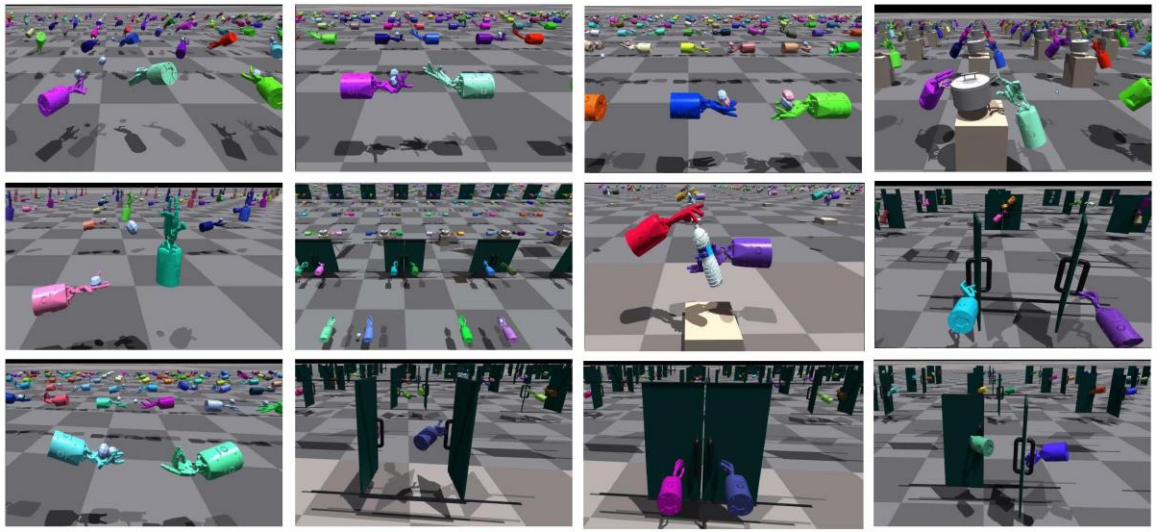
Sim2Real Transfer

Robot Dynamics Model



Training Bi-Dexterous Hand Manipulation Skills in Simulation

Training in Simulation



Real Robot Deployment

Sim2Real Method：

- System Identification
- Domain Randomization
- Domain Adaptation
- Learning with Disturbances