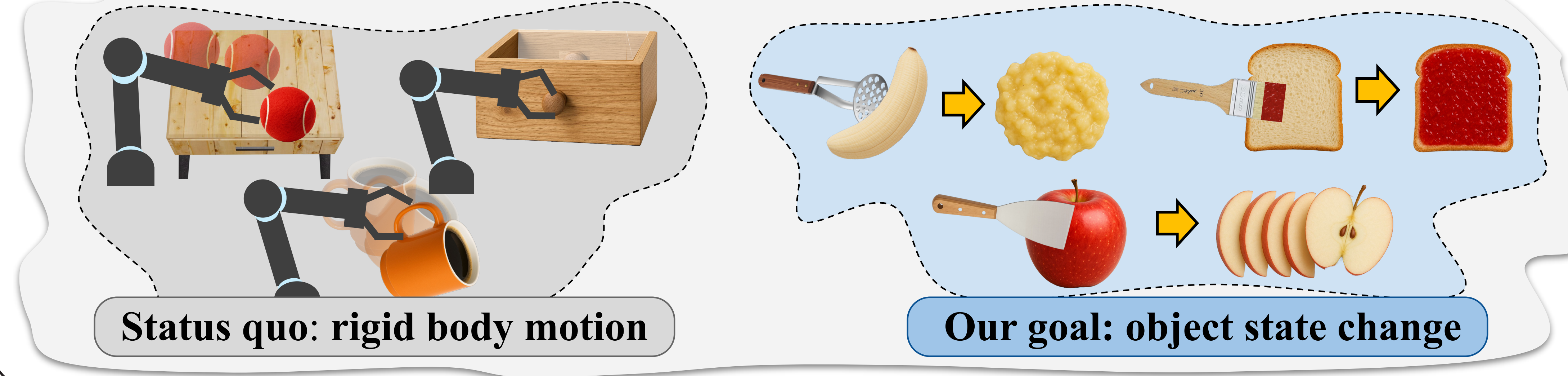




MOTIVATION

- Dominant paradigm in robot manipulation: rigid body motion
- Many real-world manipulations involve: object state changes

Space of object manipulation tasks

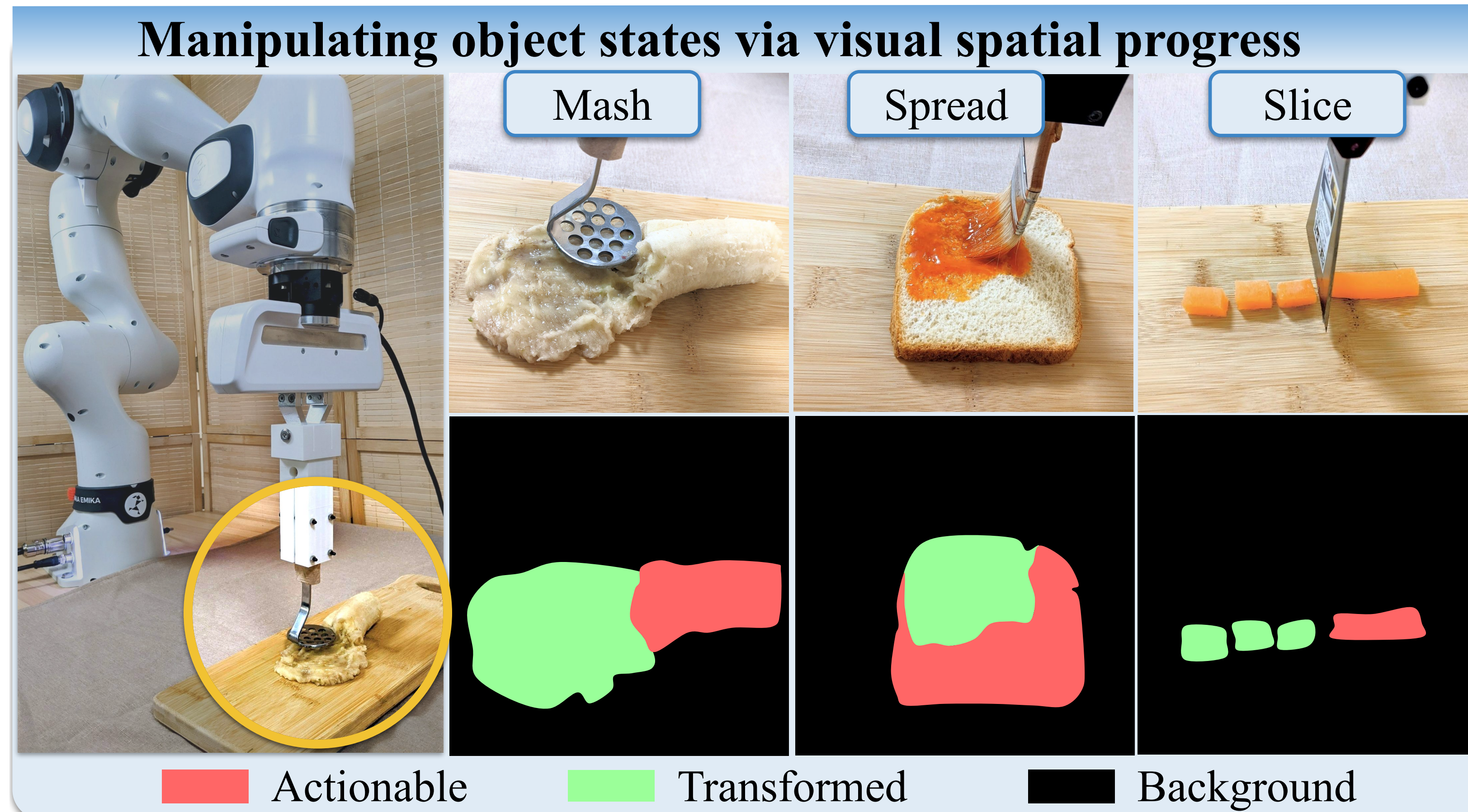


PRIOR WORK

- Simulation-heavy; require general-purpose vision feedback
- Address each state change task *in isolation*: Slicing (DiSECT, RoboNinja, RoboCook, Slice-It!), Spreading (DiffusionPolicy)

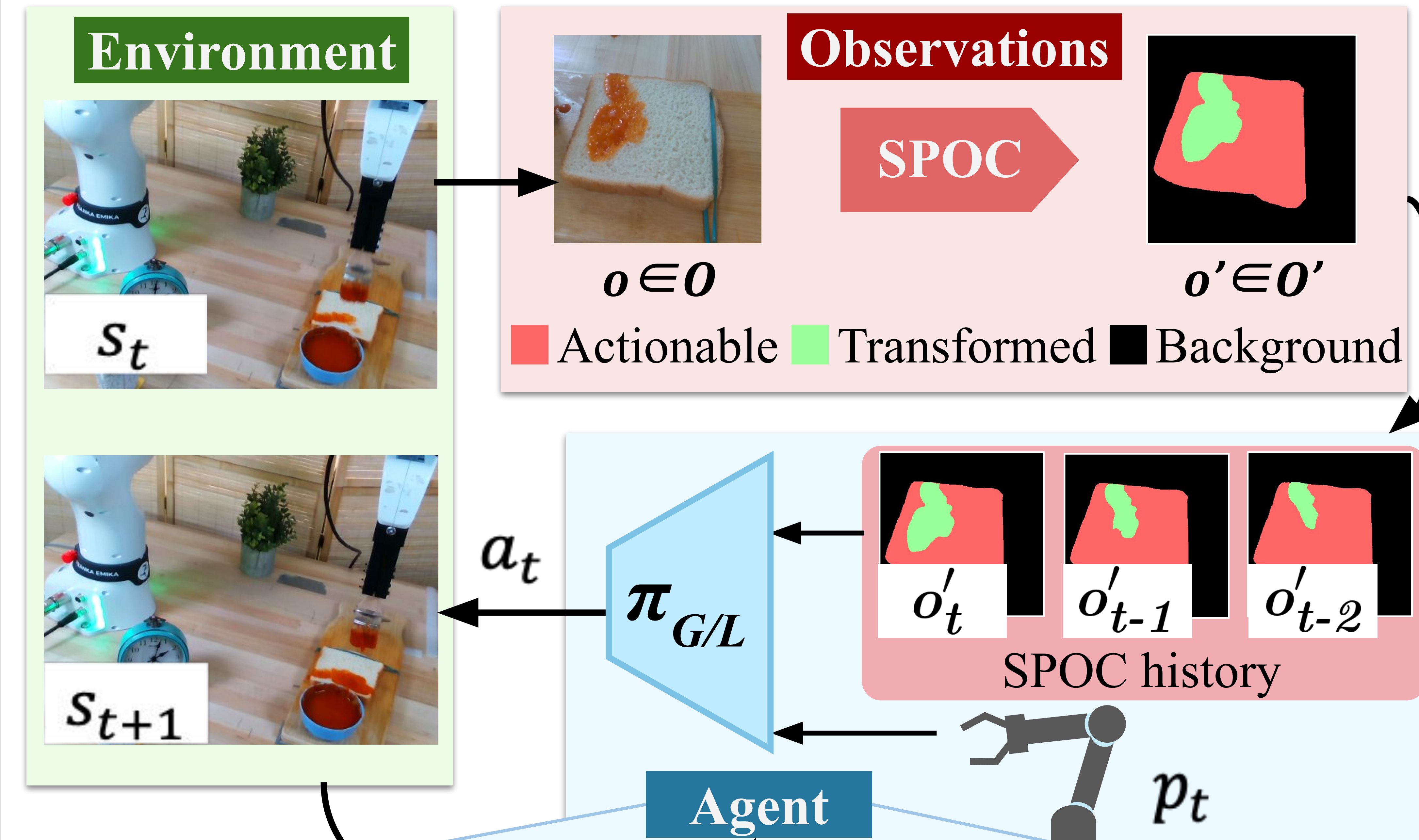
OUR IDEA

- Develop a *unified framework* across the broad family of object state change manipulation tasks
- Different spatial regions progressively evolve: *actionable* \rightarrow *transformed*
- Shared visual representation across objects and tasks



OVERVIEW

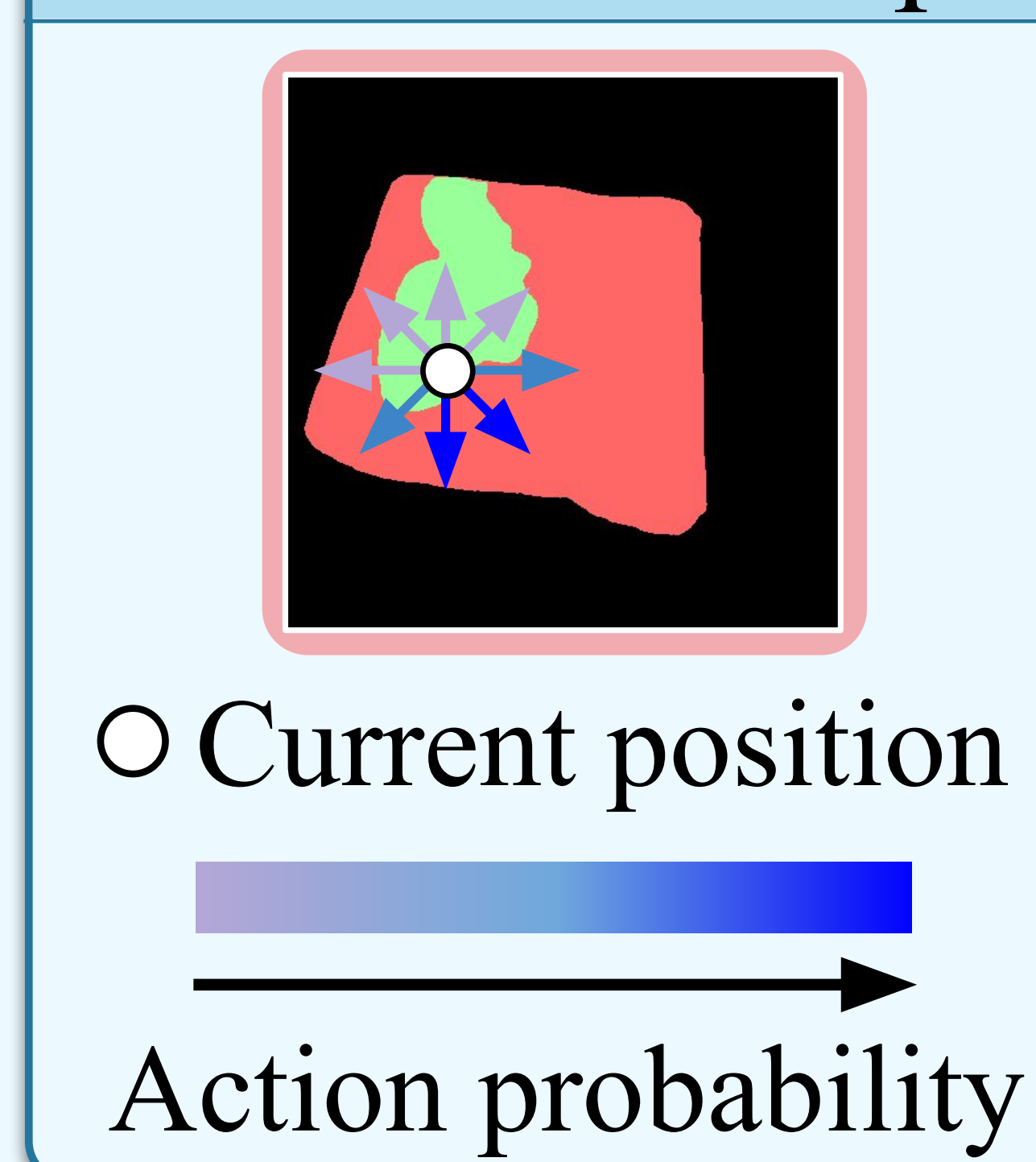
SPARTA enables two policy variants:
RL training \rightarrow complex high-precision tasks
Greedy policy \rightarrow fast deployment with no training overhead



a) π_L : Reinforcement Learning with SPOC rewards

$$R_t^{spoc} = \frac{o_{t+1}^{trf} - o_t^{trf}}{o_t^{act}}$$

b) π_G : Greedy with SPOC maps



RESULTS

Model	Spread					Slice			Mash			
	Seen	Unseen				Seen	Unseen		Seen	Unseen		
RANDOM												
SPARSE												
LIV [10]	0.24	0.42	0.27	0.29	0.23	0.13	0.15	0.14	0.18	0.14	0.23	0.20
SPARTA-G	0.14	0.10	0.07	0.11	0.13	0.09	0.08	0.09	0.13	0.08	0.09	0.18
SPARTA-L	0.17	0.14	0.12	0.16	0.12	0.10	0.09	0.11	0.13	0.09	0.10	0.08
SPARTA-G	0.44	0.49	0.55	0.66	0.39	0.52	0.48	0.51	0.75	0.69	0.71	0.75
SPARTA-L	0.61	0.55	0.58	0.63	0.42	0.78	0.69	0.72	0.77	0.72	0.62	0.68

