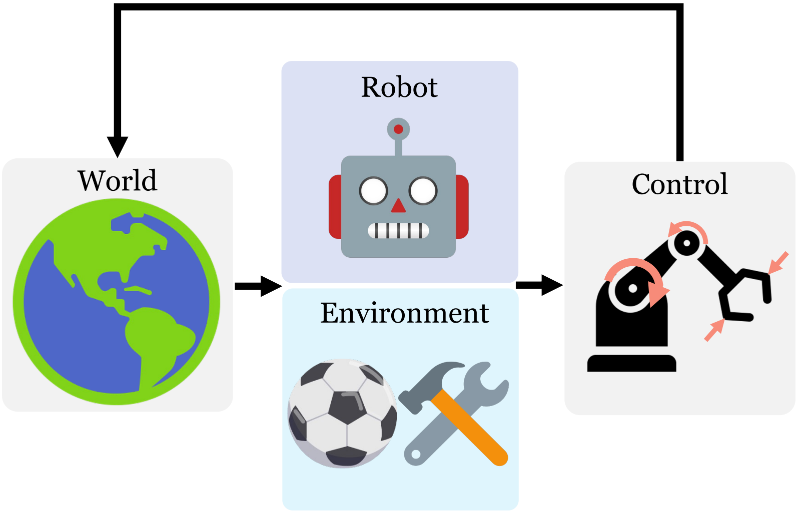
Efficient RL via Disentangled Environment and Agent Representations



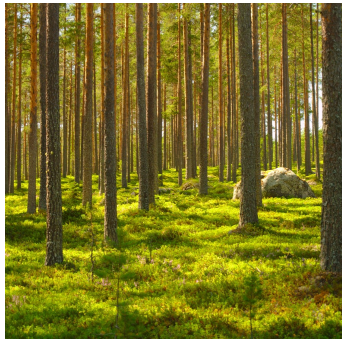

Kevin Gmelin* Shikhar Bahl* Russell Mendonca Deepak Pathak



Carnegie Mellon University

THE ROBOTICS INSTITUTE

Goal: General-Purpose Robots



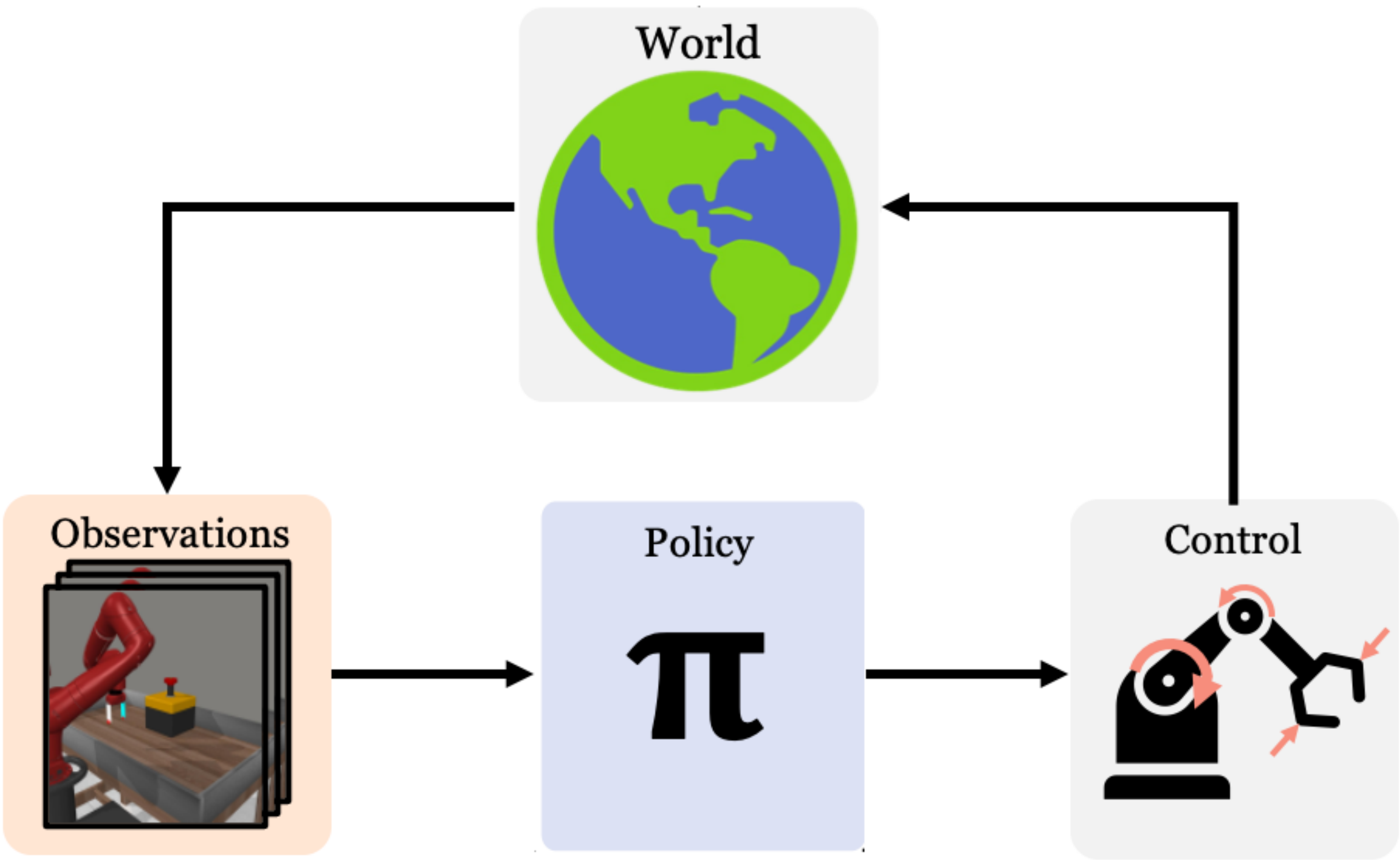


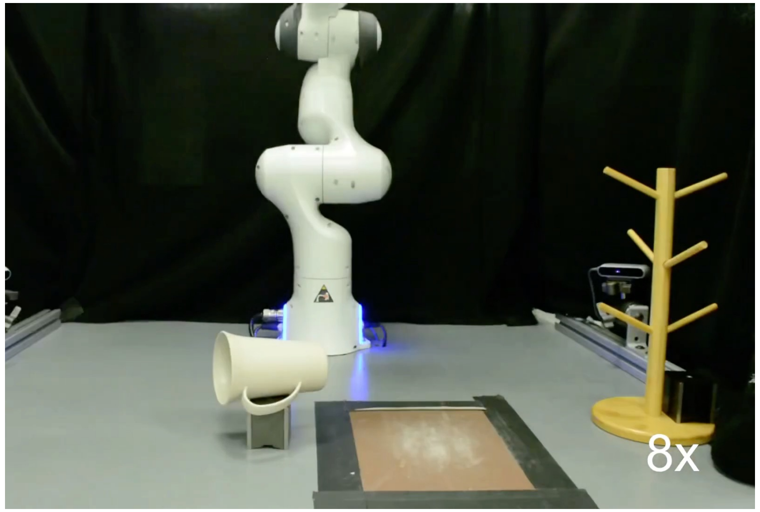
Robots that can do thousands of tasks in thousands of environments

Visual Reinforcement Learning

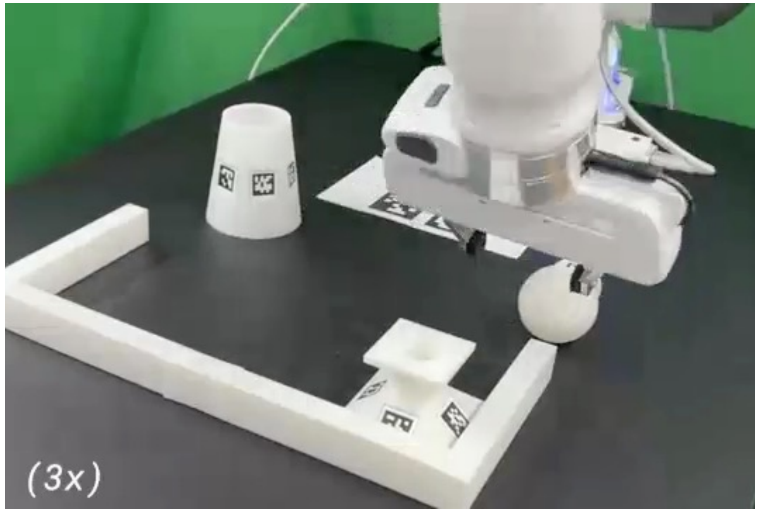
Goal: Learn mapping from images to actions

Want **sample-efficient** algorithms

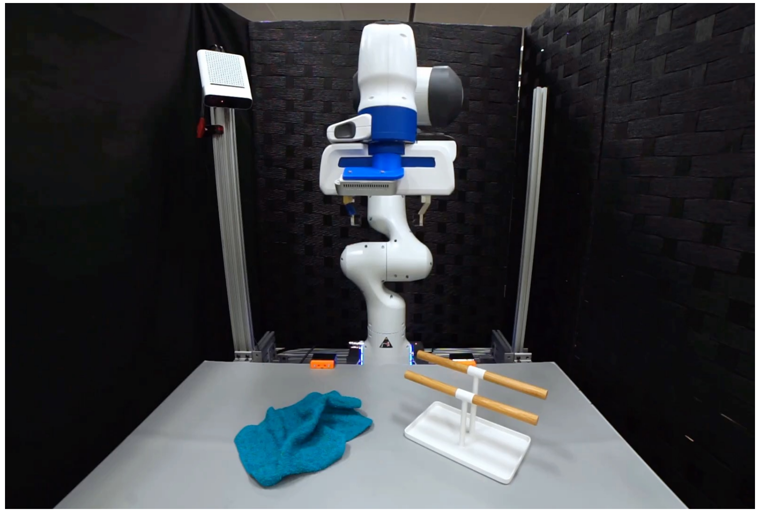




Simeonov et al., 2022

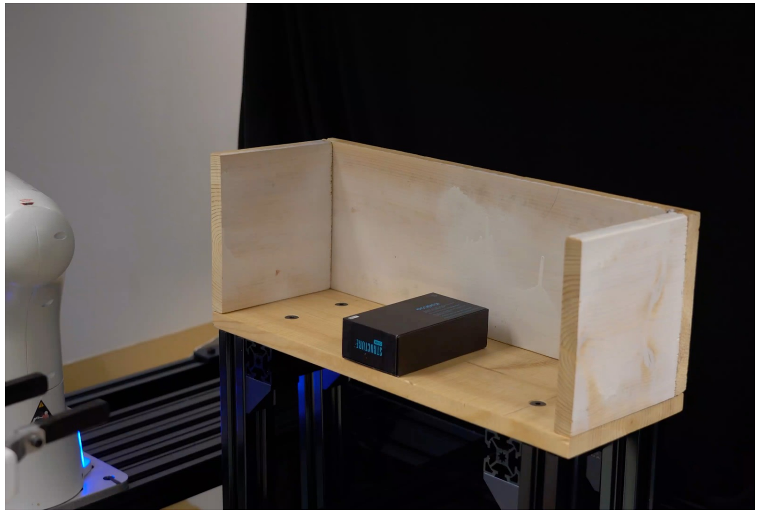


Heo et al., 2023



Huang et al., 2023

Mendonca et al., 2023



Liang et al., 2022

Zhou et al., 2023

Agent is common across many tasks

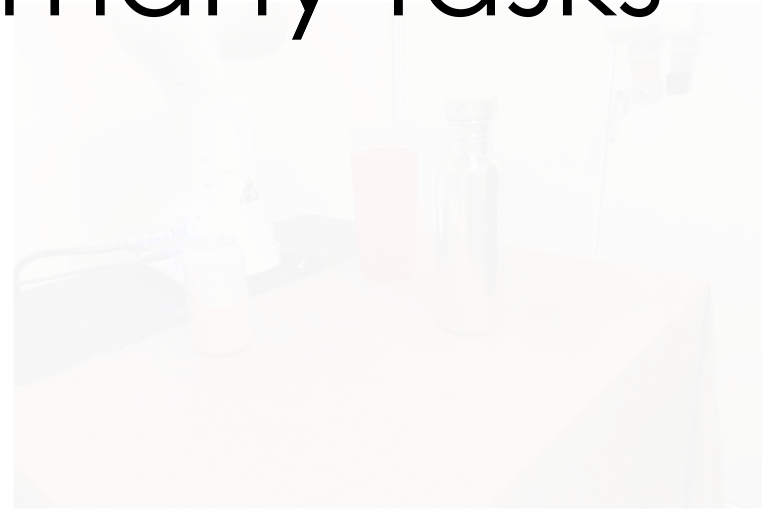
Simeonov et al., 2022

Heo et al., 2023

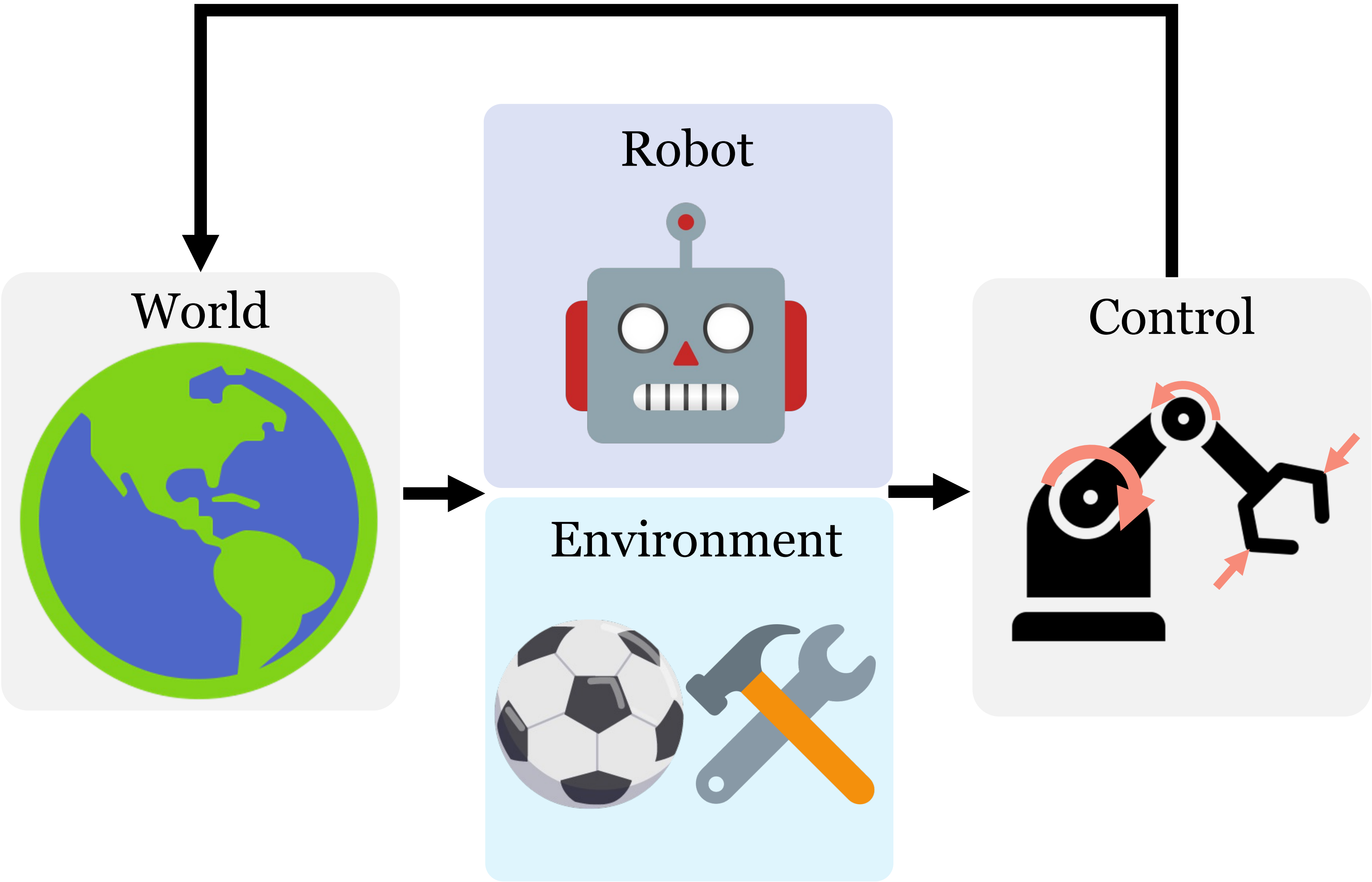
Huang et al., 2023


Mendonca et al., 2023


Liang et al., 2022


Zhou et al., 2023

Agent vs Environment Rep. for Control



Agent vs Environment Rep. for Control

How can we learn disentangled **agent** & **environment** representations?

How to obtain supervision?

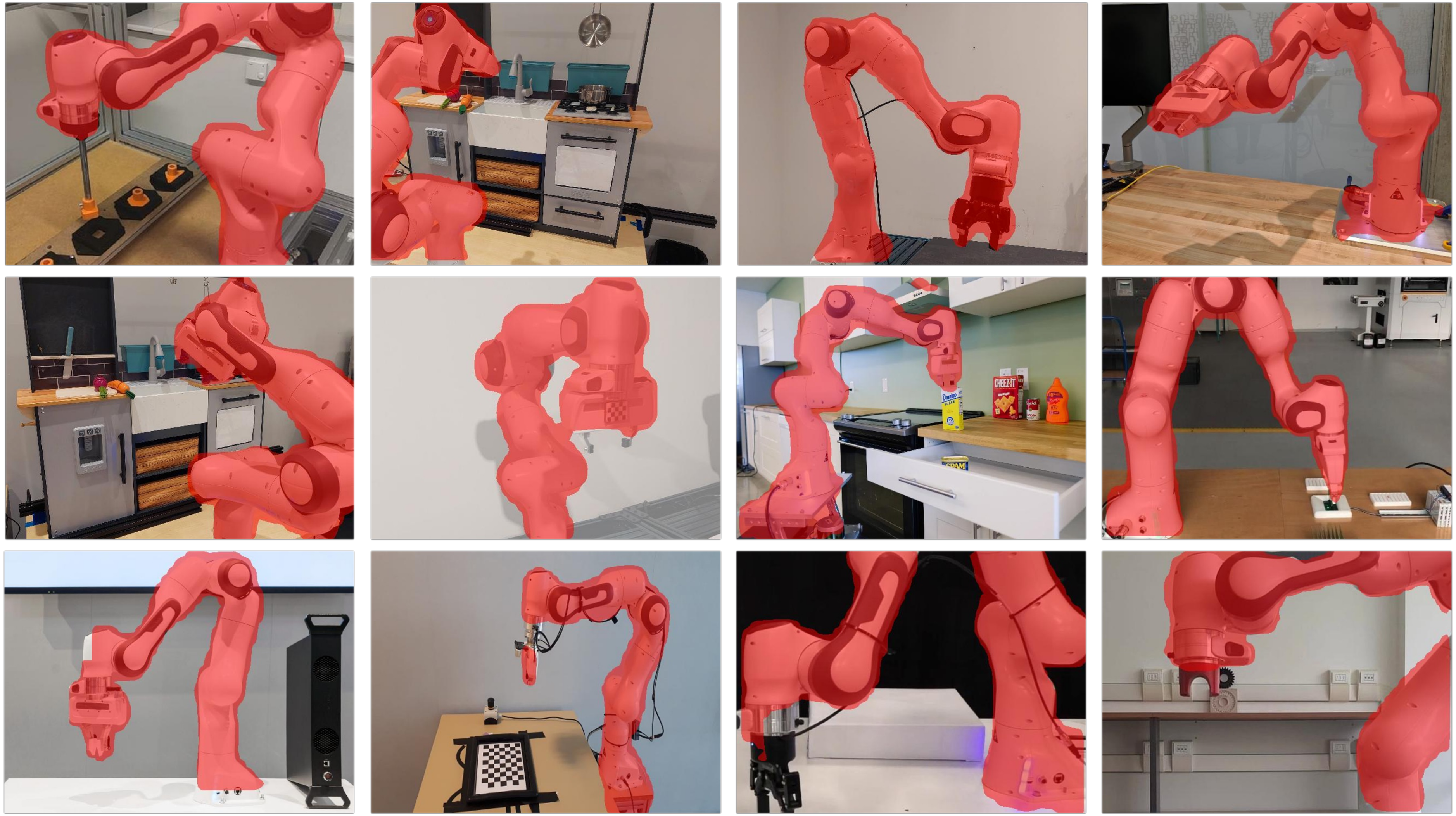
Environment?

- Can directly use scene observation

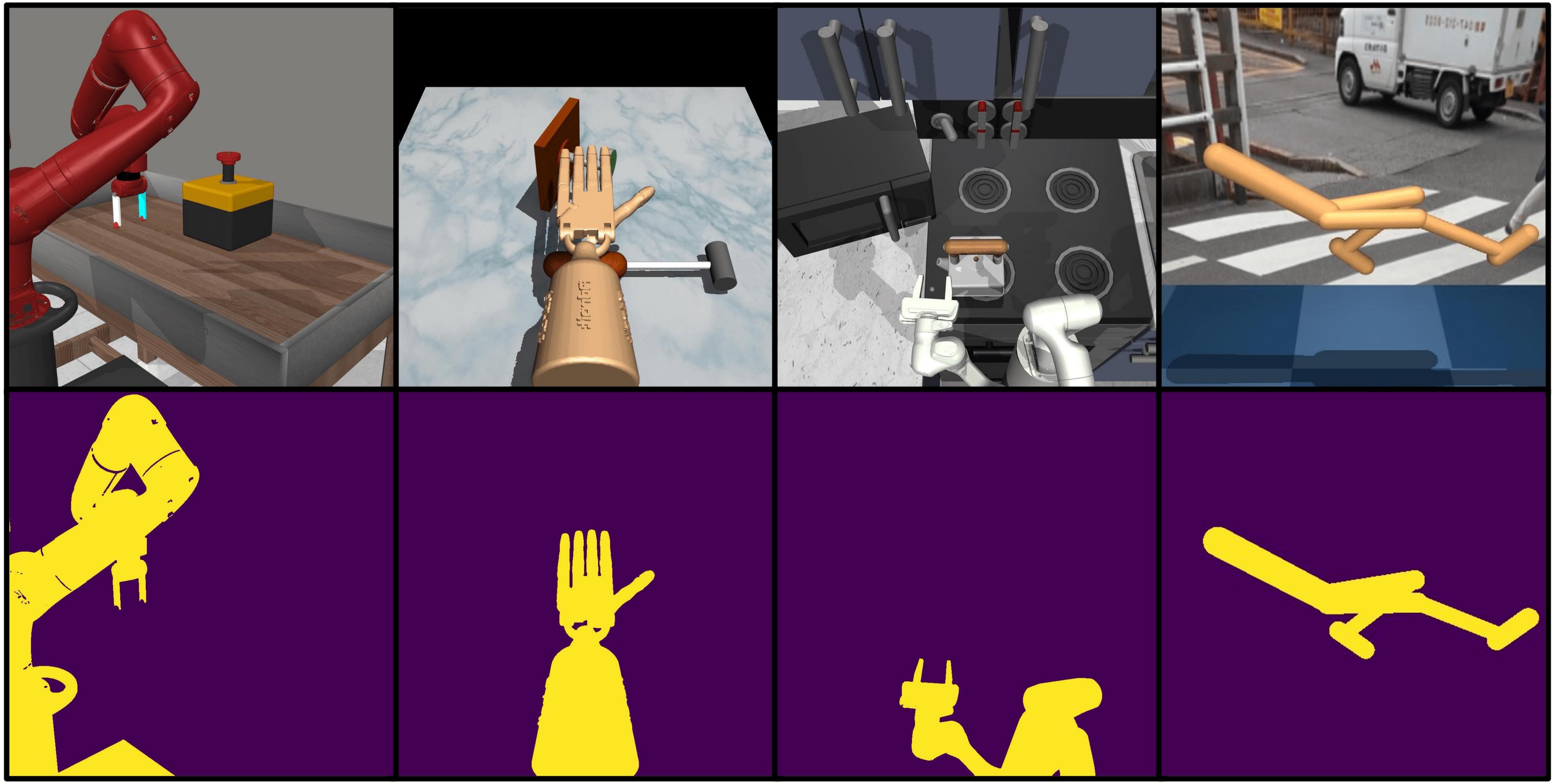
What about for the agent?

- Common to know what robot looks like
- **Masks** are a good proxy for agent info
- Use **off-the-shelf models**

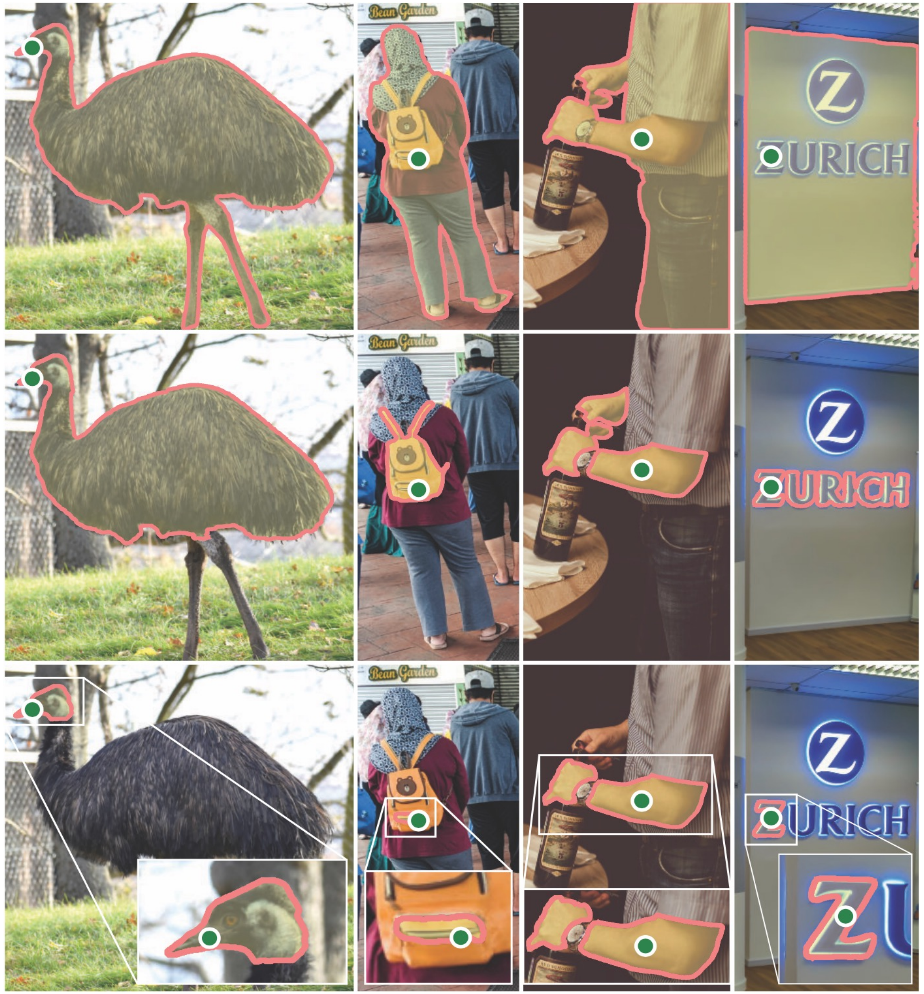
Real World Masks



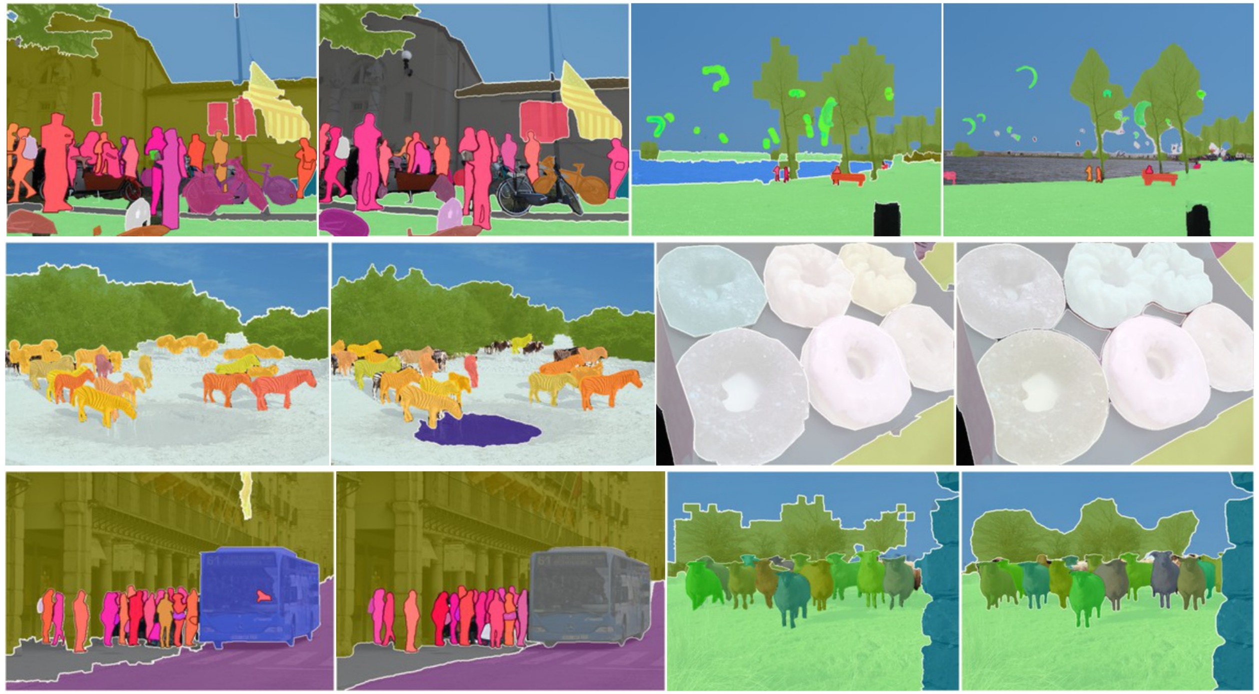
Simulation



Shelf-Supervised Agent Masks



Segment Anything
Kirillov et al., 2023



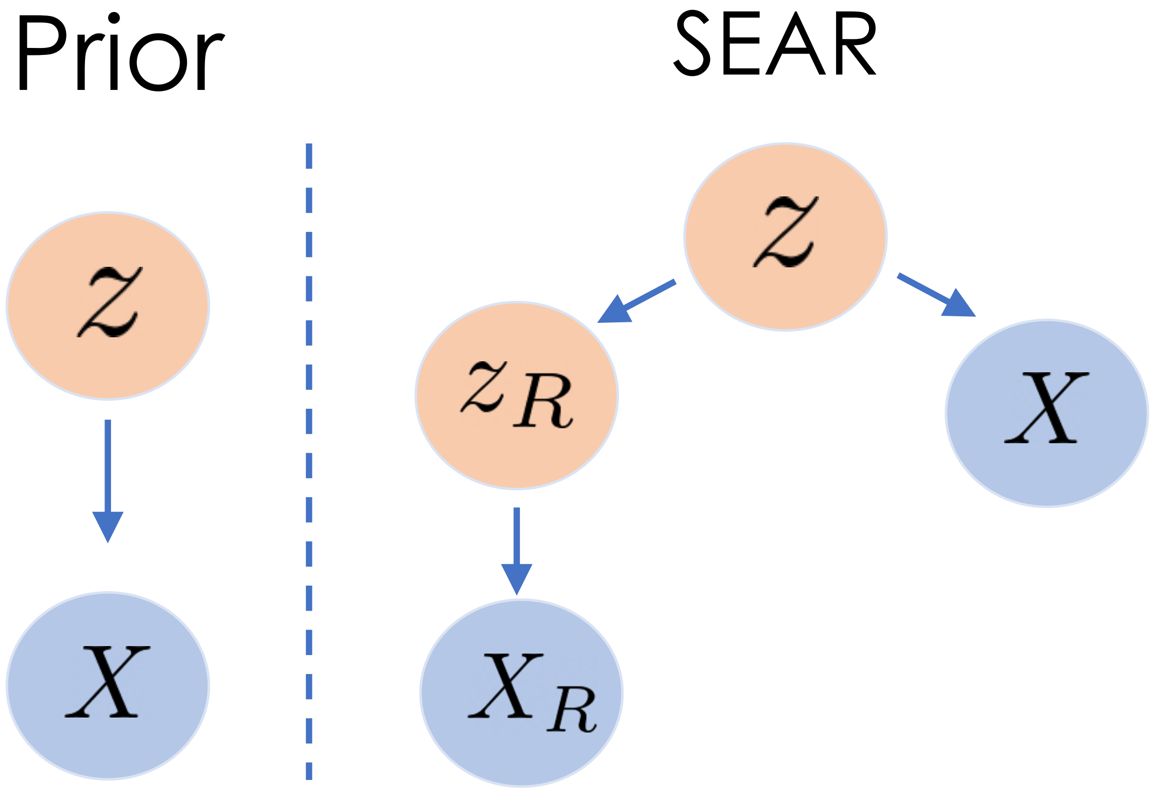
Mask2Former
Cheng et al., 2022

How do we incorporate this into RL?

Agent/Environment Latent Structure

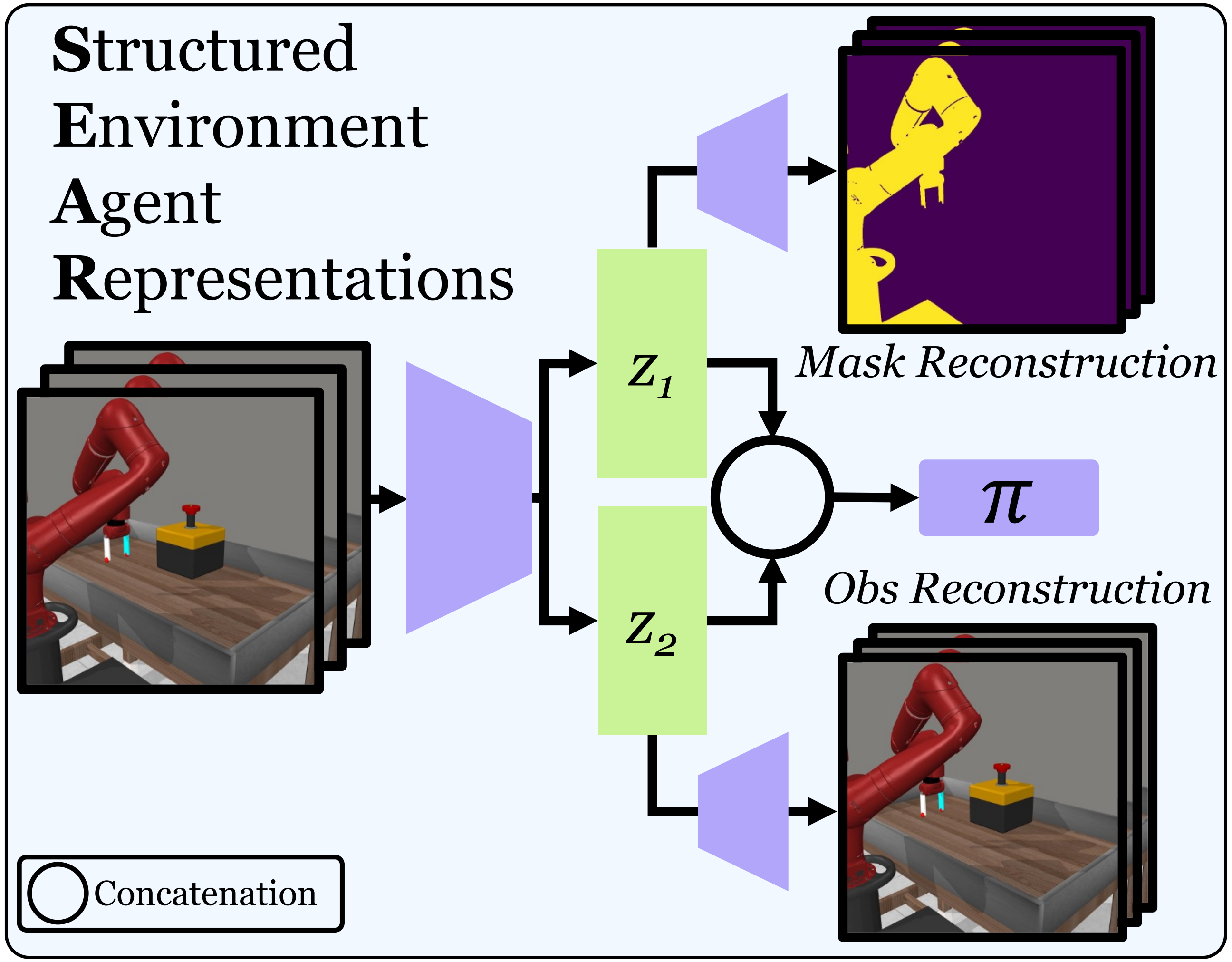
$$p(X, X_R, z, z_R) = p(z).p(z_R|z).p(X|z).p(X_R|z_R)$$

$$\mathcal{L} = \mathop{\mathbb{E}}_{z,z_R\sim q}[\log p(X|z)] + \mathop{\mathbb{E}}_{z,z_R\sim q}[\log p(X_R|z_R)]$$
$$- D_{KL}(q(z, z_R|X))||(p(z, z_R))$$



SEAR: Structured Environment-Agent Representations

Structure latent into
agent (z_1), **env** (z_2)

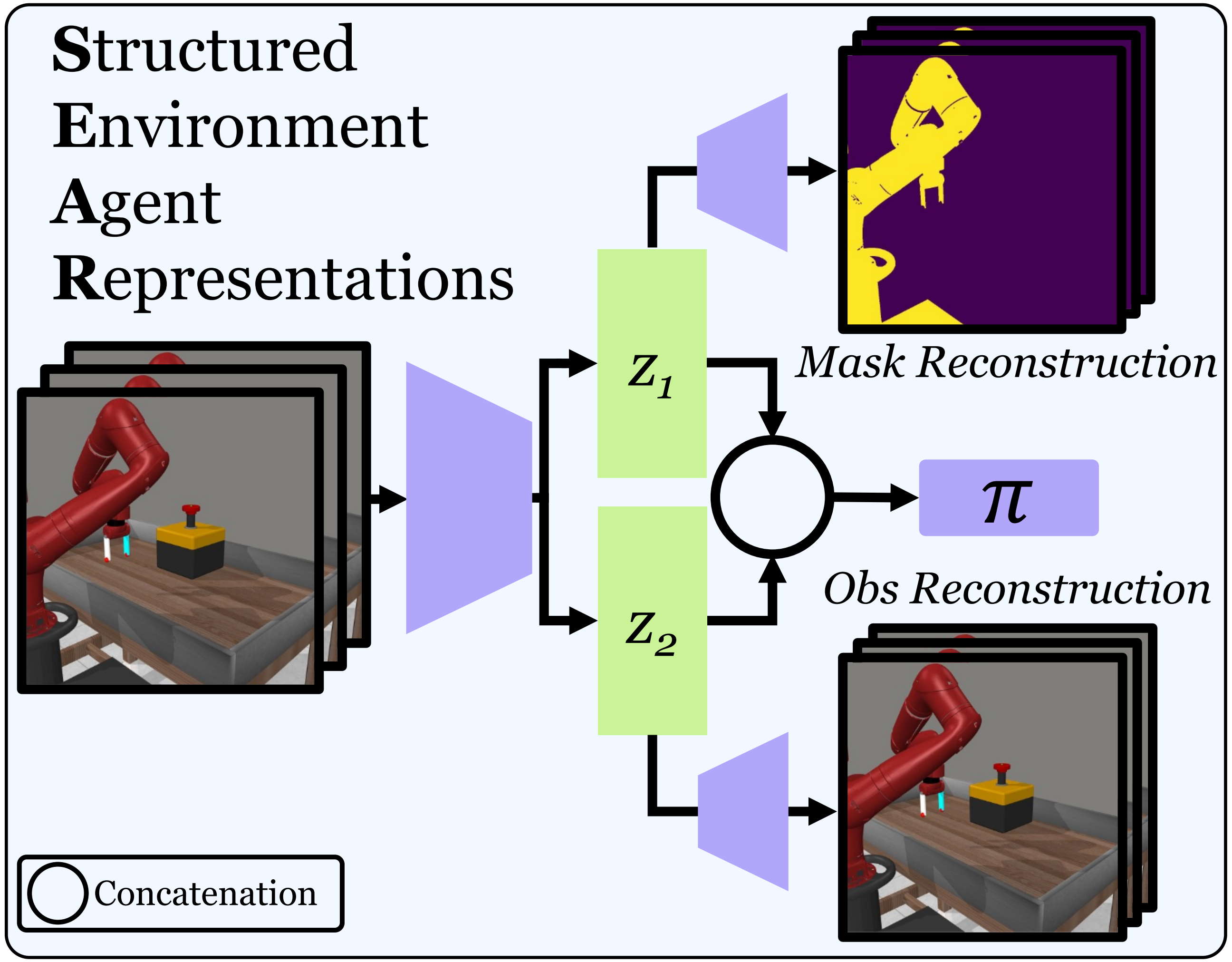


SEAR: Structured Environment-Agent Representations

Leverage agent **mask** as supervision

$$\mathcal{L}_{mask} = M \log P_\phi(M|z_R) + (1-M)\log(1 - P_\phi(M|z_R))$$

$$\mathcal{L} = \mathcal{L}_{critic} + c_1\mathcal{L}_{recon} + c_2\mathcal{L}_{mask}$$



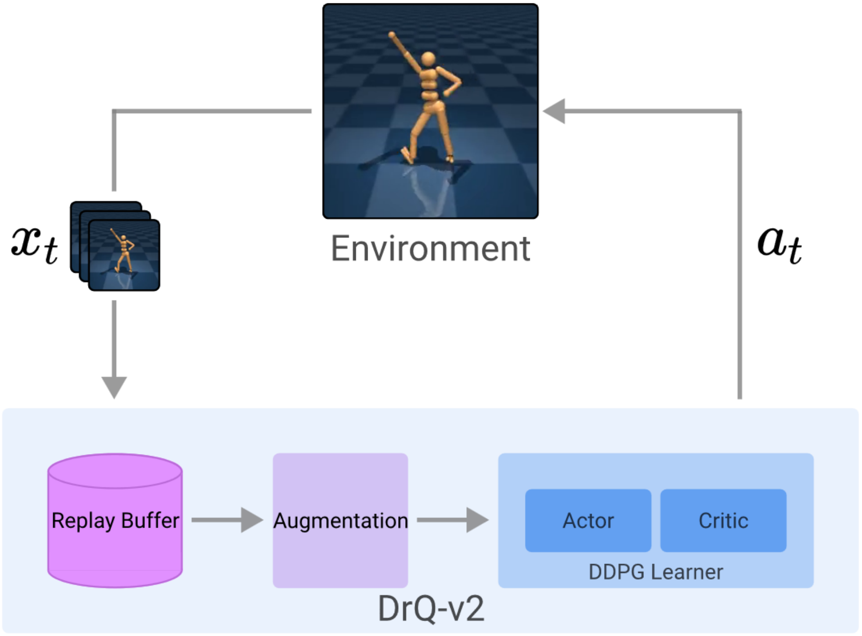
Existing Approaches

Data-Augmentations
- RAD
- DrQ-v2

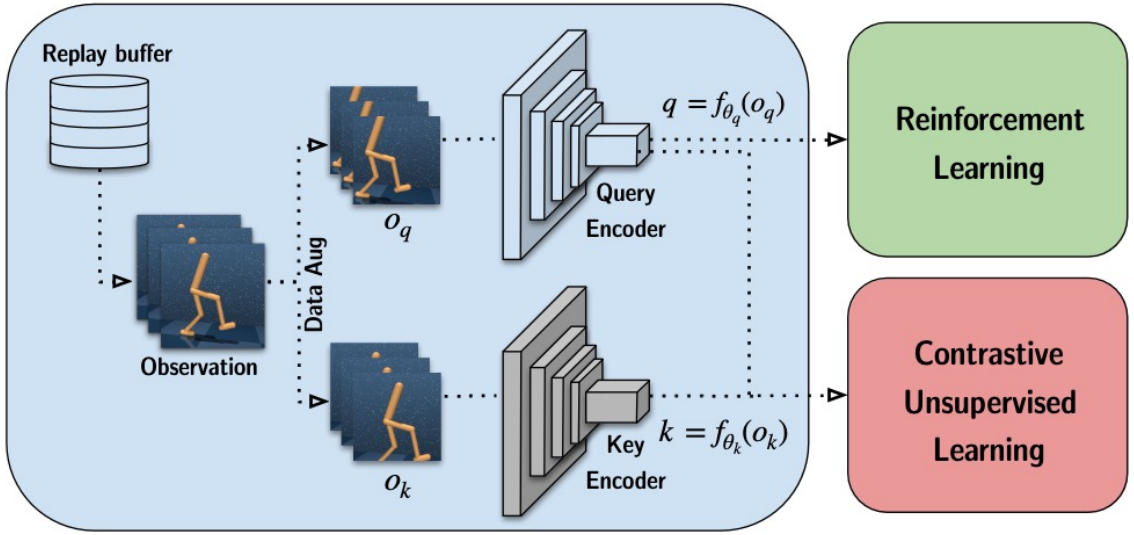
Auxillary Losses
- CURL
- SAC-AE

Model-Based
- Dreamer

And many more…



Yarats et al., 2021a)



Laskin et al., 2020b

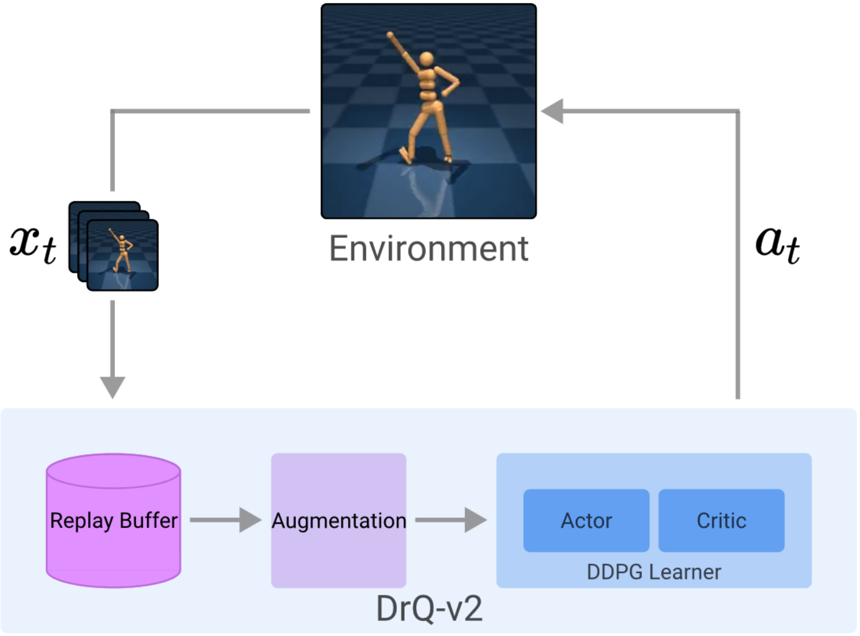
Existing Approaches

Data-Augmentations
- RAD
- DrQ-v2

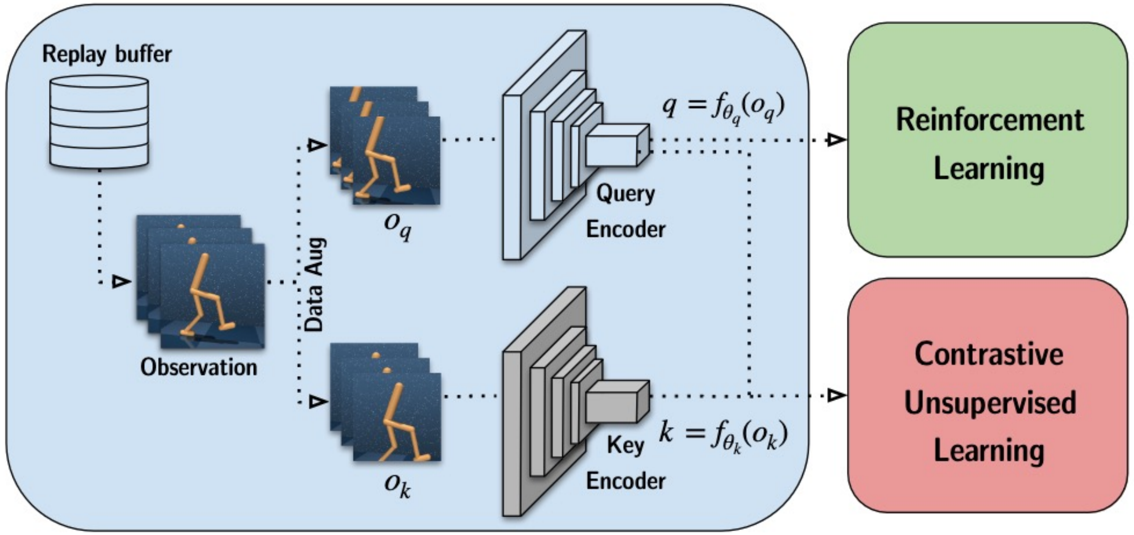
Auxillary Losses
- CURL
- SAC-AE

Model-Based
- Dreamer

And many more…



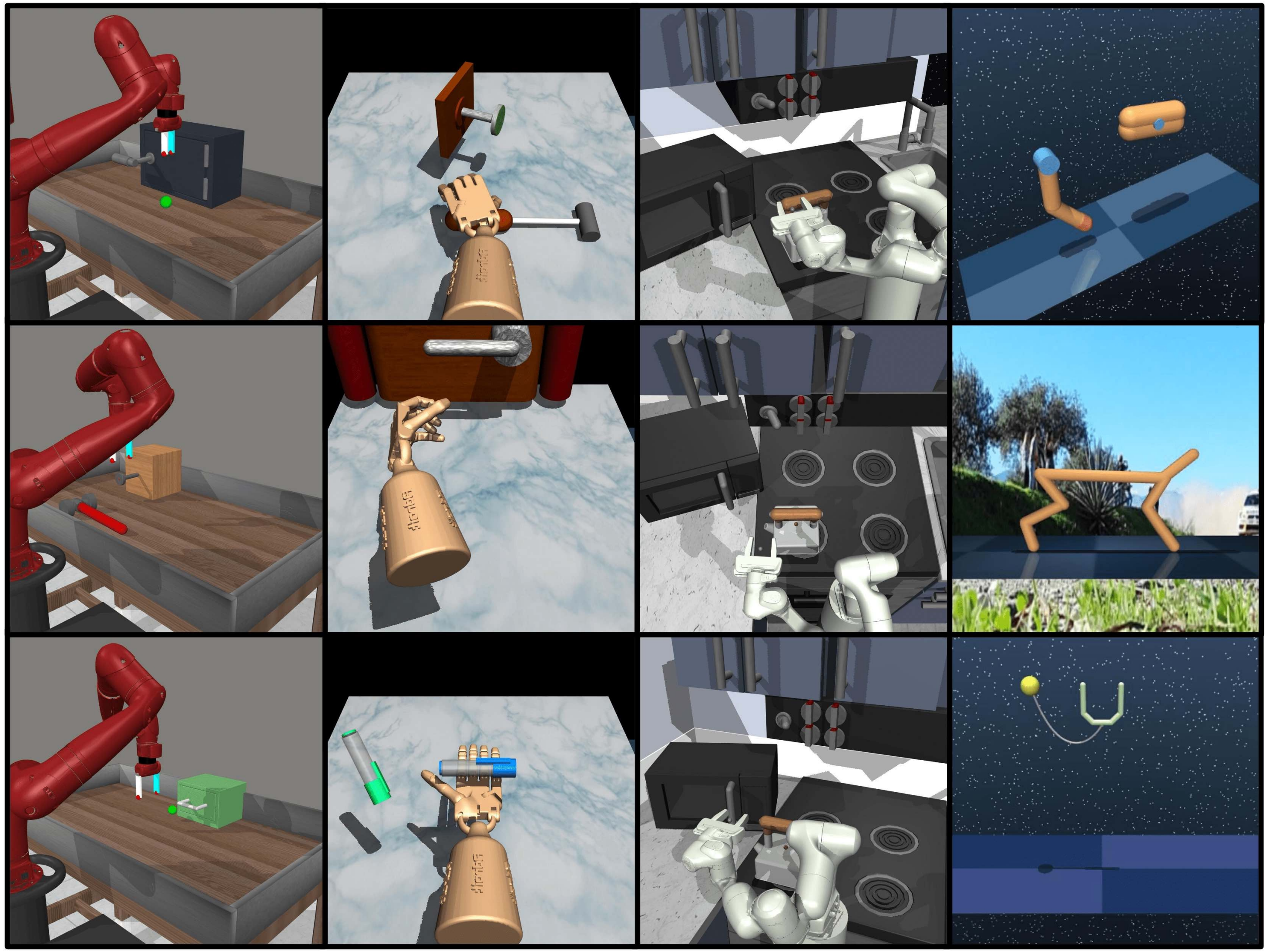
Yarats et al., 2021a)



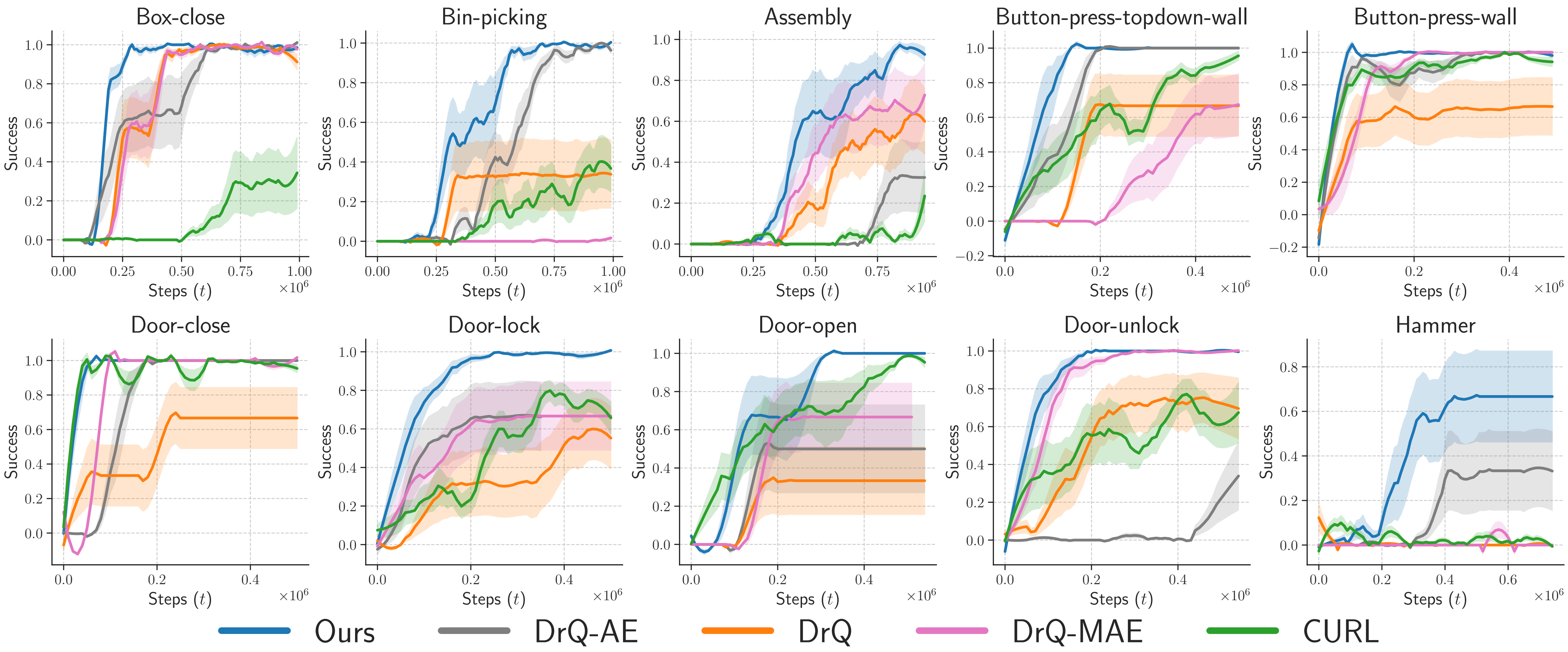
Laskin et al., 2020b

Environments

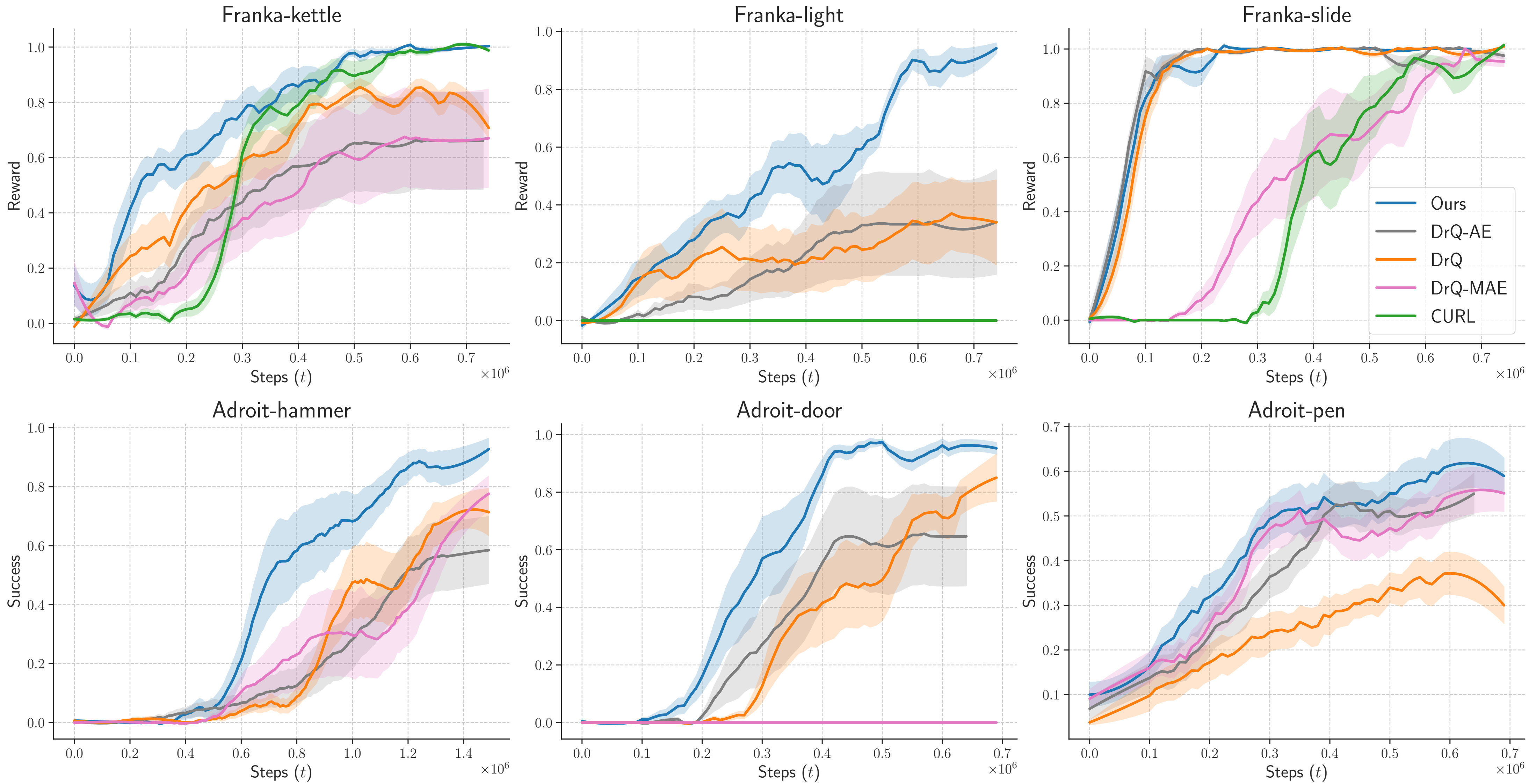
- Meta-World (Yu et al., 2020)
- Hand Manipulation Suite (Rajeswaran et al., 2017)
- Franka Kitchen (Gupta et al., 2019)
- Distracting Control Suite (Stone et al., 2021)



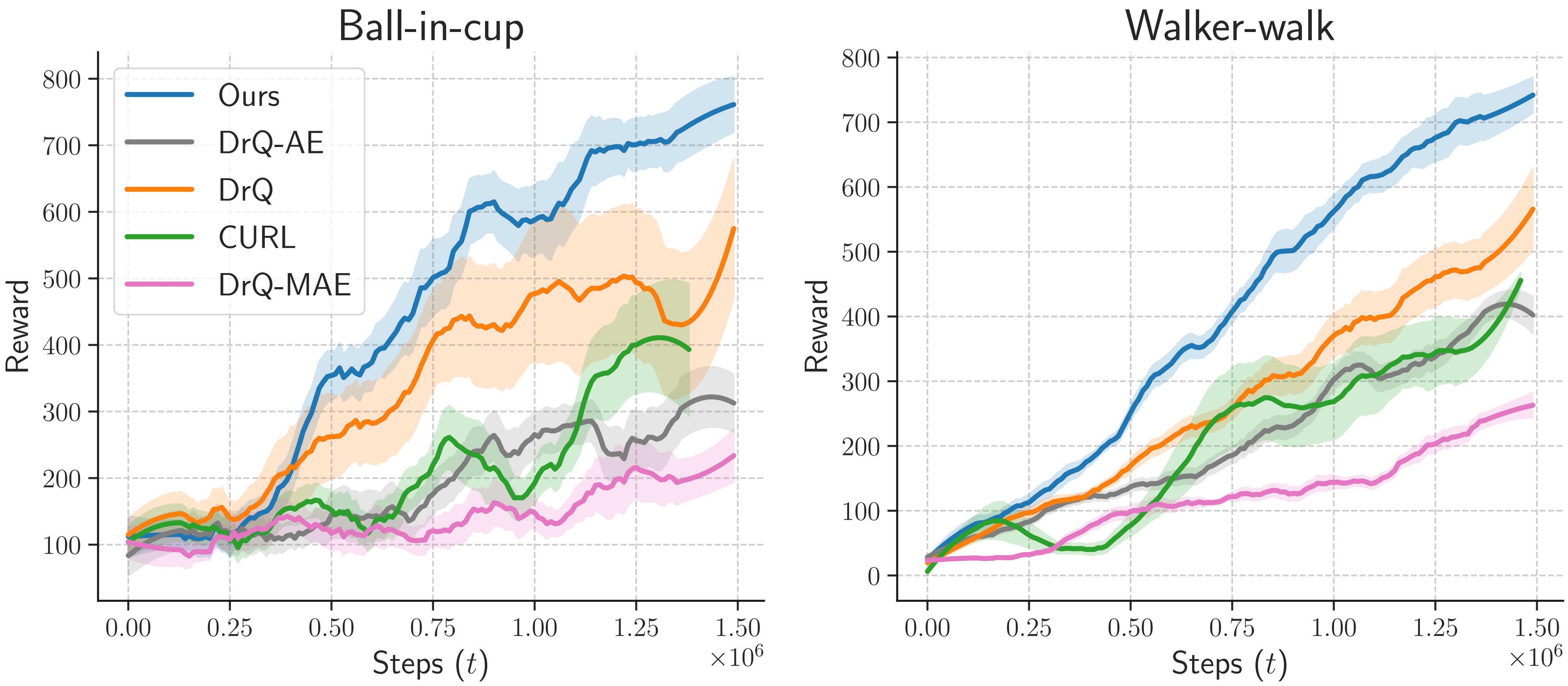
Meta-World


Box-close
Bin-picking
Assembly
Button-press-topdown-wall
Button-press-wall
Door-close
Door-lock
Door-open
Door-unlock
Hammer
Success
Steps (t)
$\times 10^6$
Ours
DrQ-AE
DrQ
DrQ-MAE
CURL

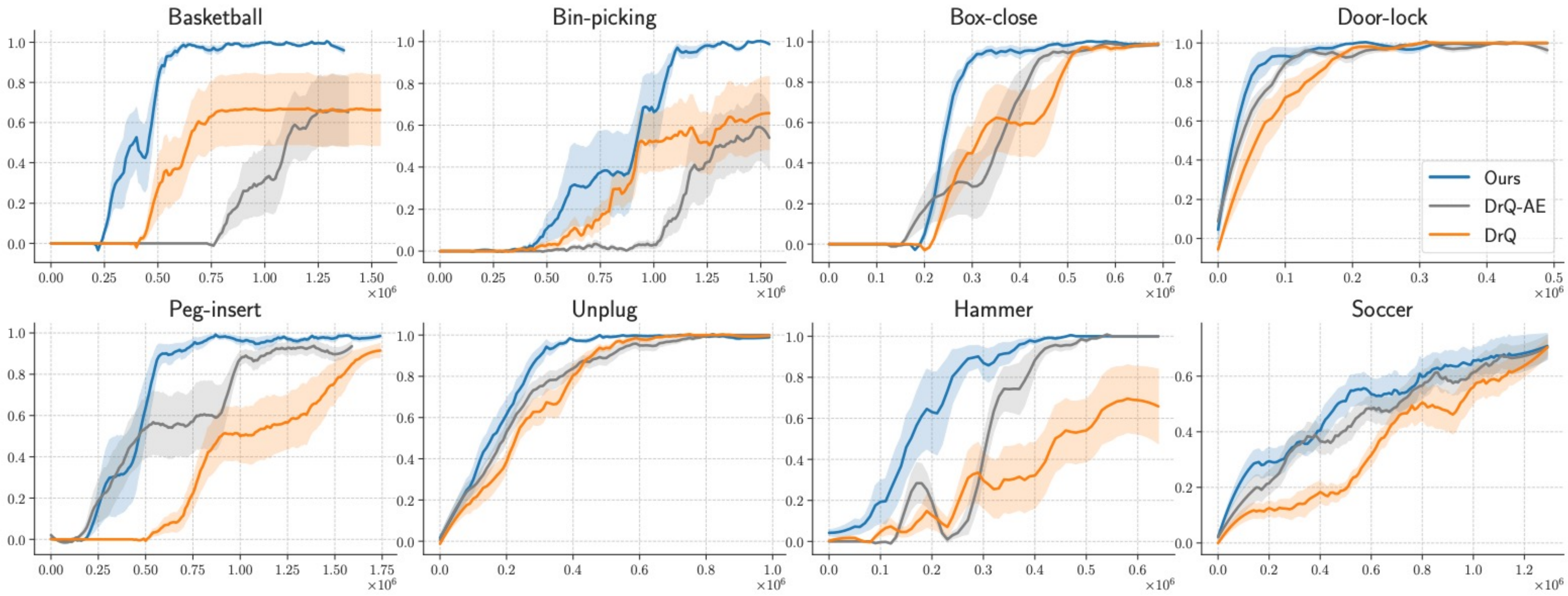
Franka Kitchen & Adroit



Distract Gym

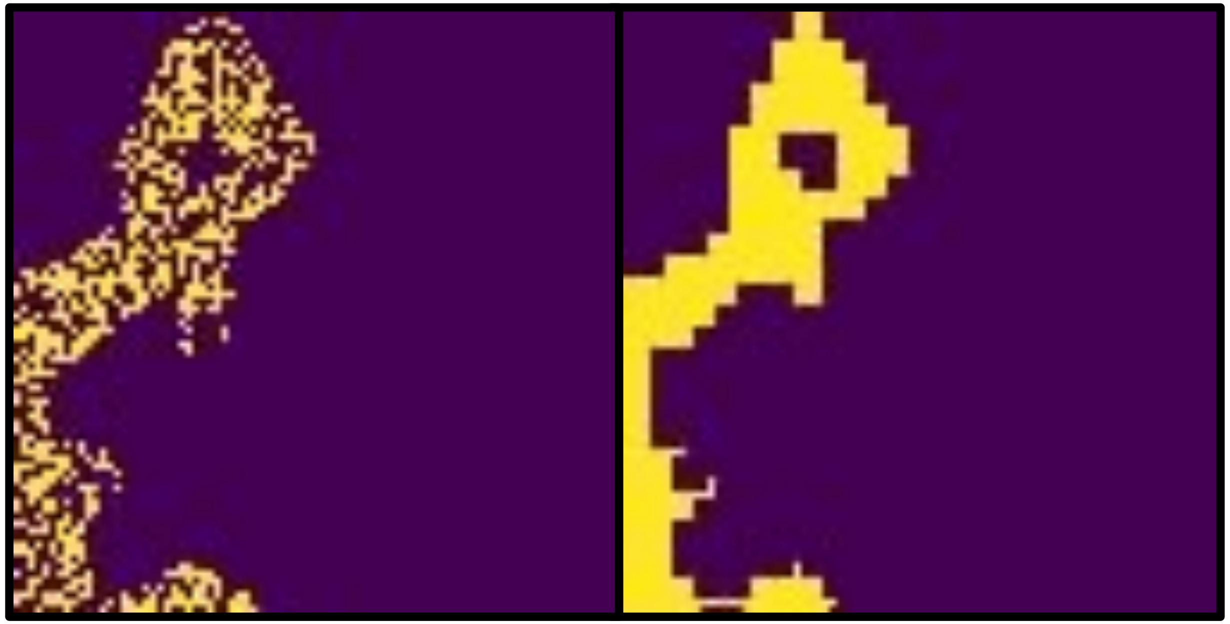


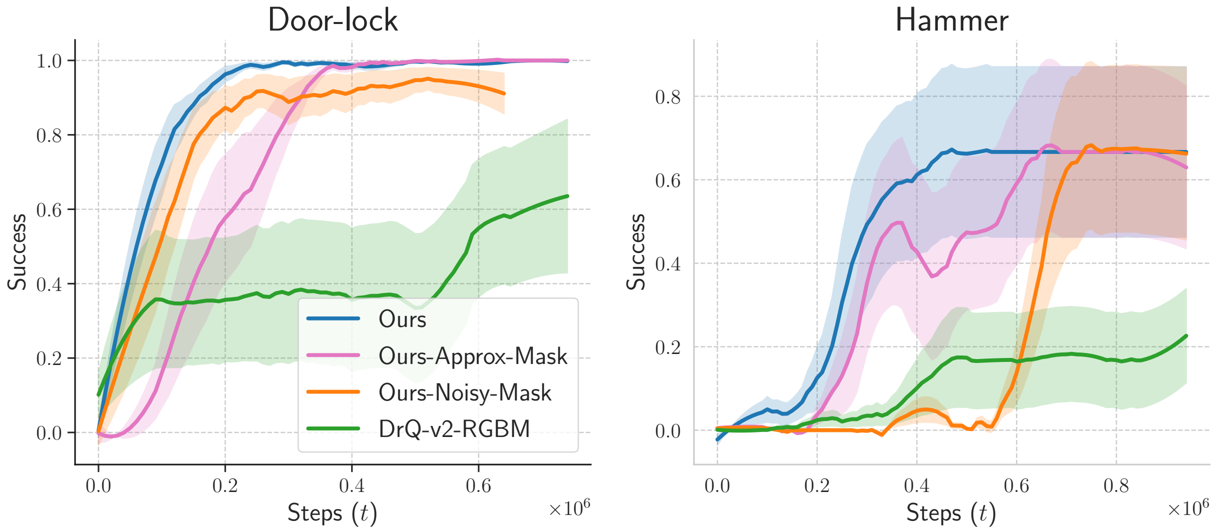
Transfer to new environments



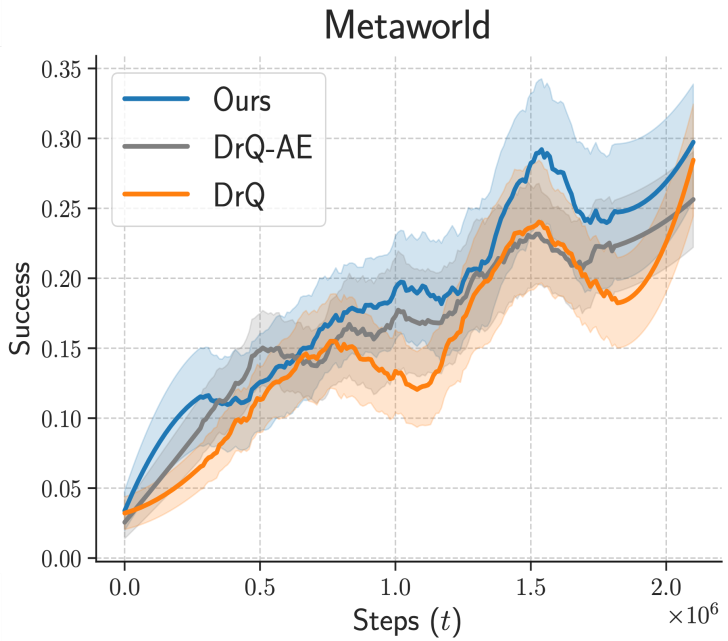
Robustness to Mask Quality

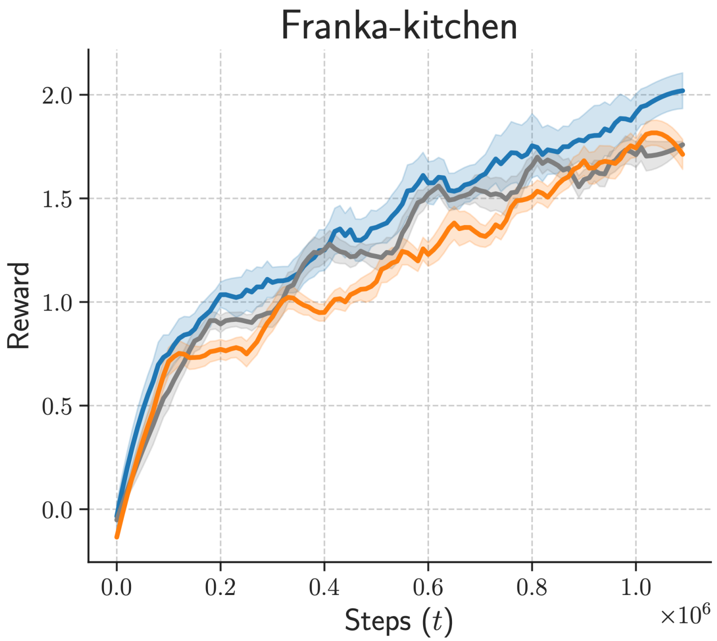
We train with noisy mask (left) and lower resolution, approximate mask (right)



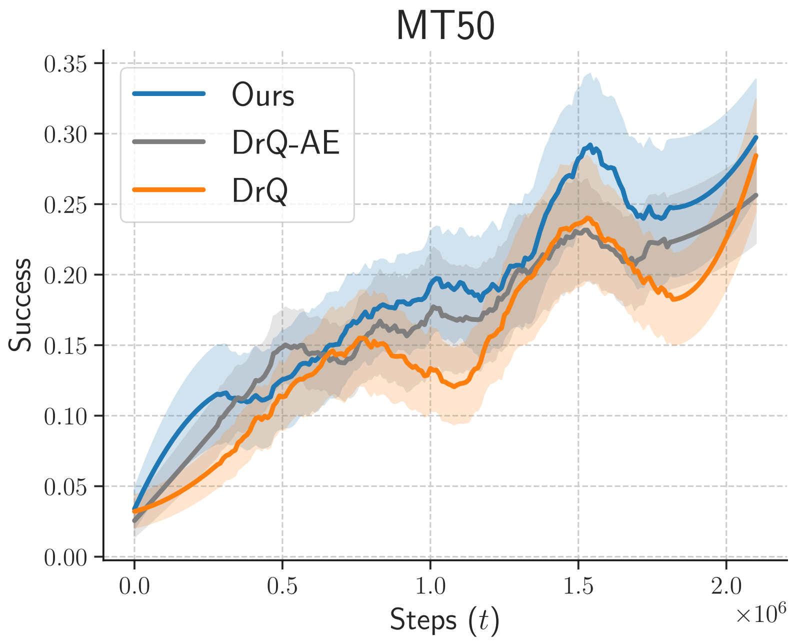


What about multi-task settings?





Small-Scale Setups



Larger-Scale Setups

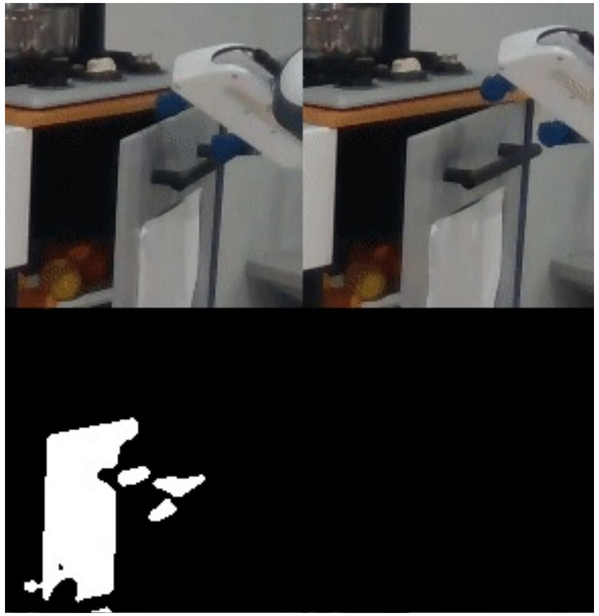
Applications in real robots?

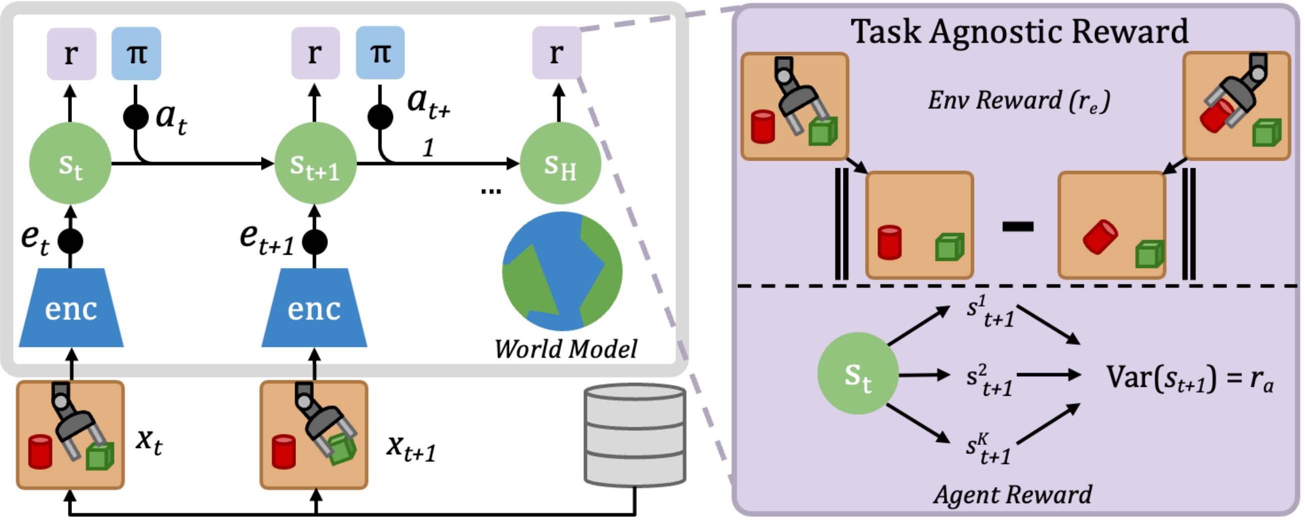
Environment-centric Exploration

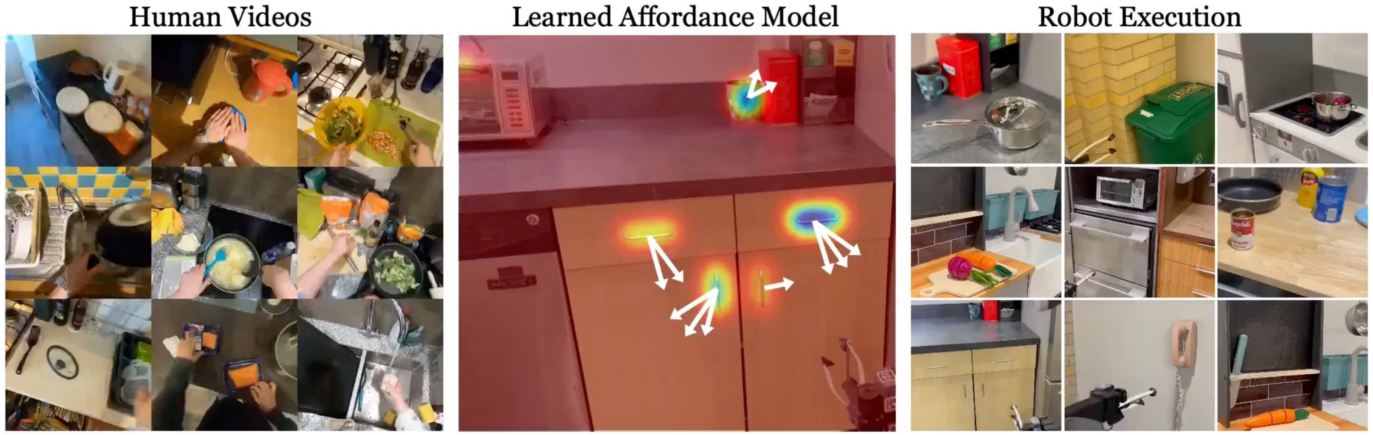
Env. Chang Exploration Objective

$$c_k = \max_{i,j} \left\|\Phi_f\left(R_{k,i}\right) - \Phi_f\left(R_{k,j}\right)\right\|$$

Human-to-Robot Imitation in the Wild.
Bahl, Gupta and Pathak. *RSS, 2022*







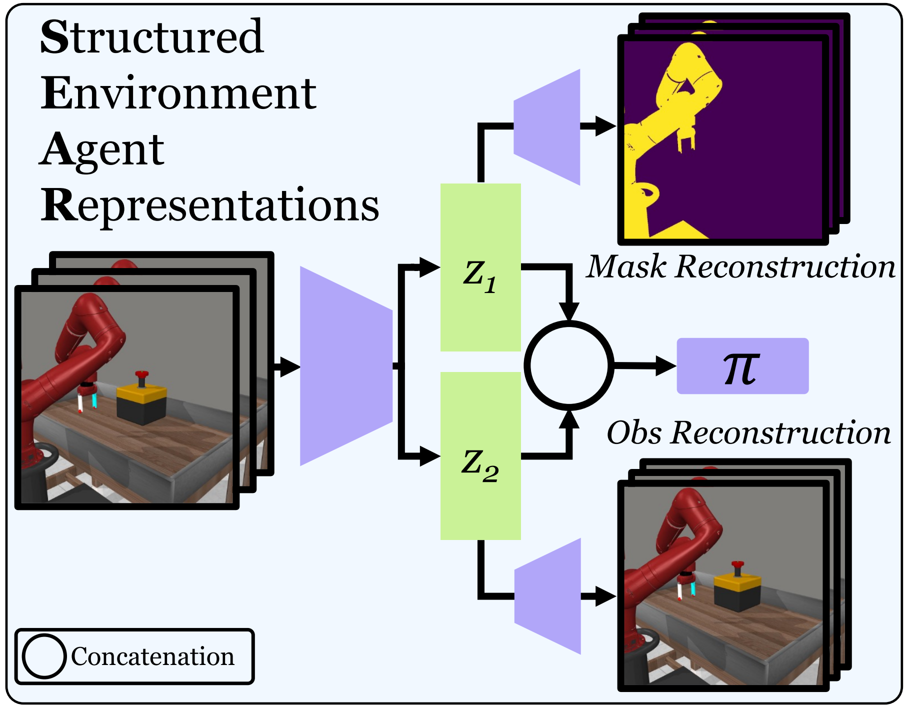
Agent-Agnostic, change-based exploration can be useful

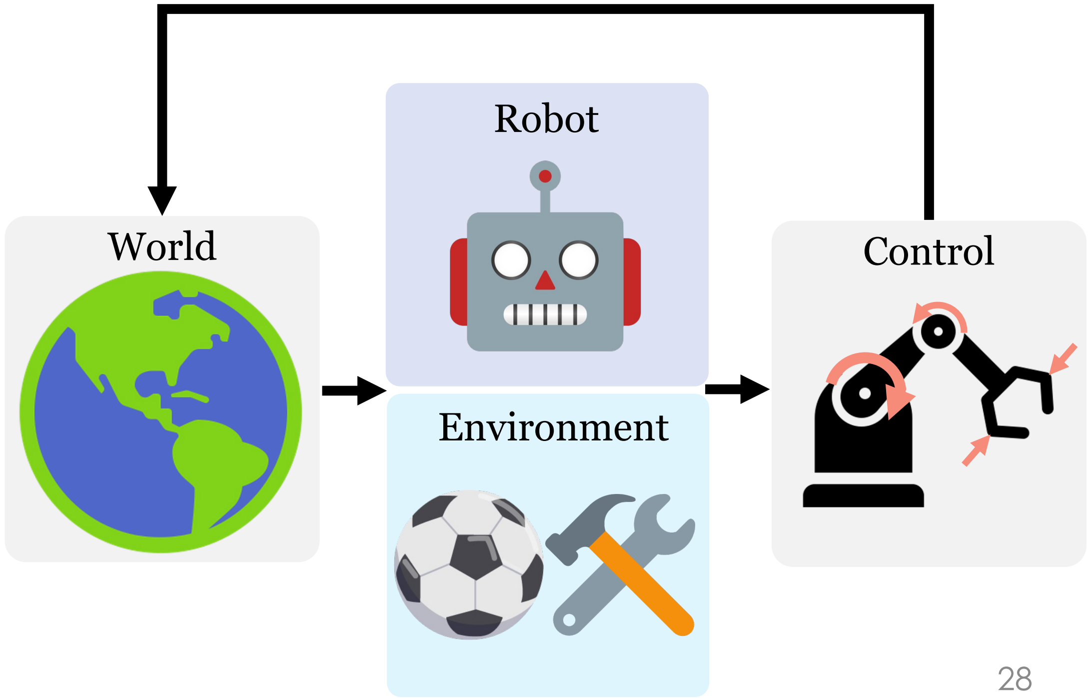
Autonomously Exploring Robotic Agents in the Real World.
Mendonca, Bahl and Pathak. *ICRA, 2023*

Visual Affordances as a Versatile Representation For Robotics .
Bahl*, Mendonca*, Chen, Jain and Pathak. *ICVPR, 2023*

Limitations

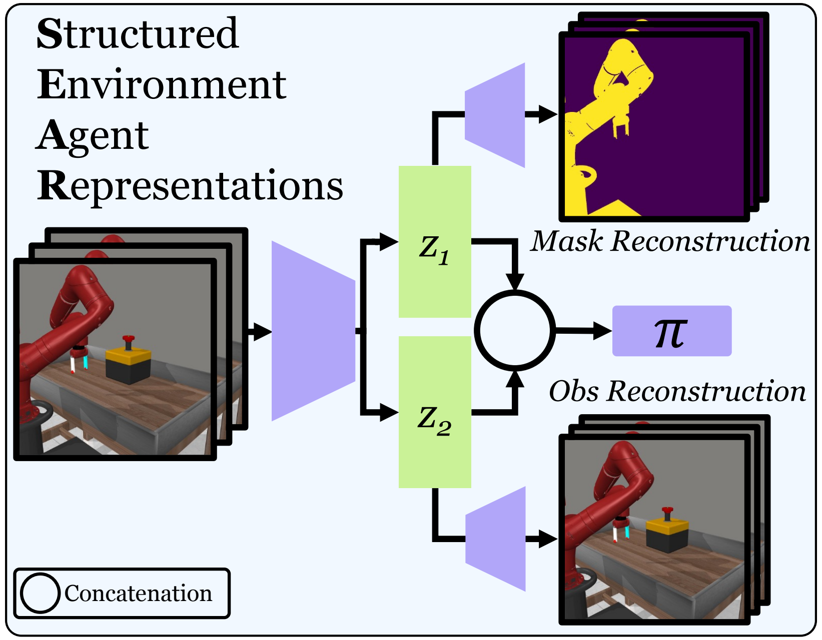
- Further investigation into **multi-task** setting needed
- Robot must be **visible** in image
- Only examined training from **scratch**
- Only added SEAR onto **DrQ-v2**





Final Thoughts

- Decoupled representation **boosts performance**
- SEAR can help with **transfer**
- Masks are **readily available** from shelf-supervised models
- Can be added to **any visual RL** approach



sear-rl.github.io