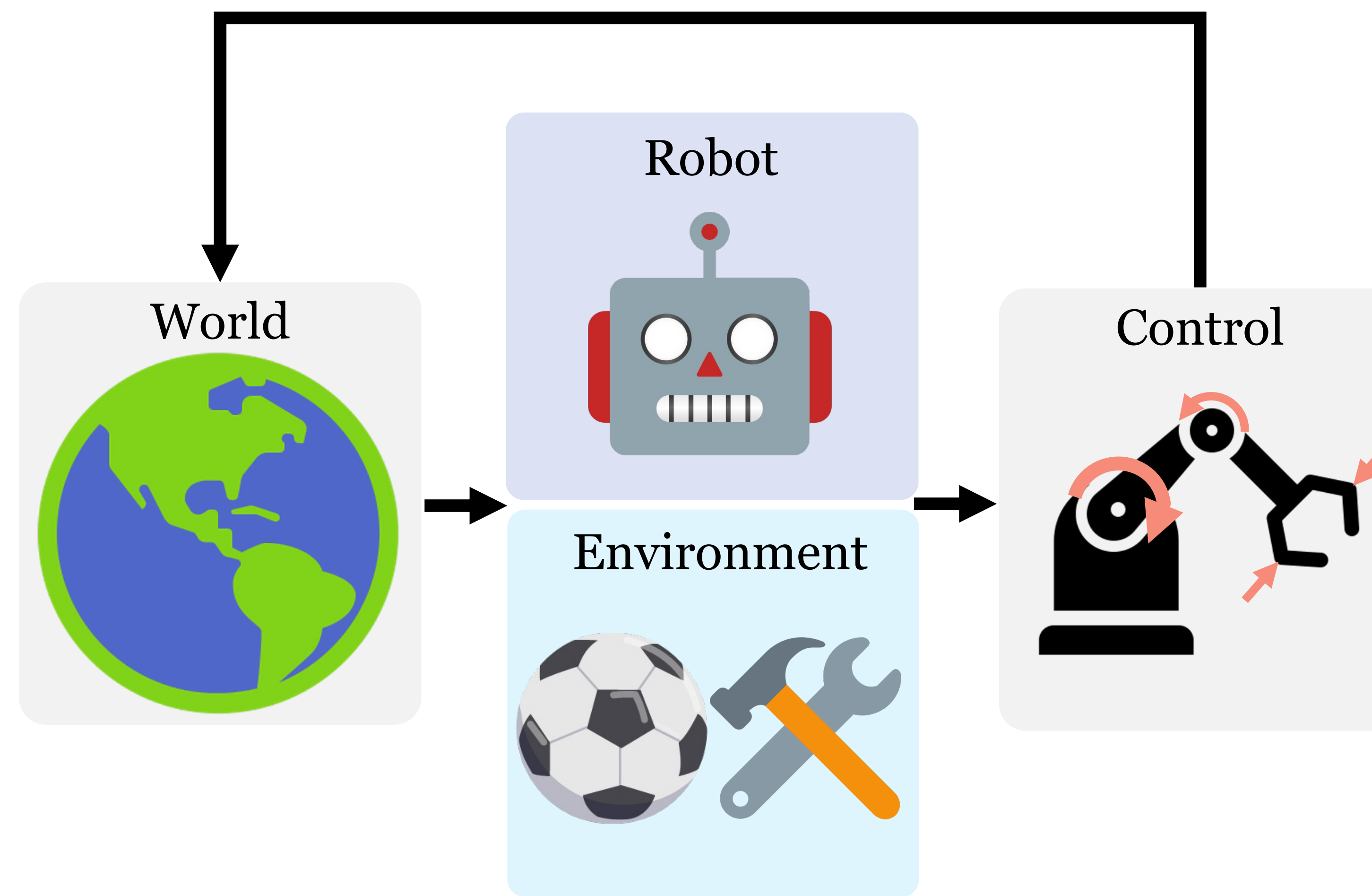
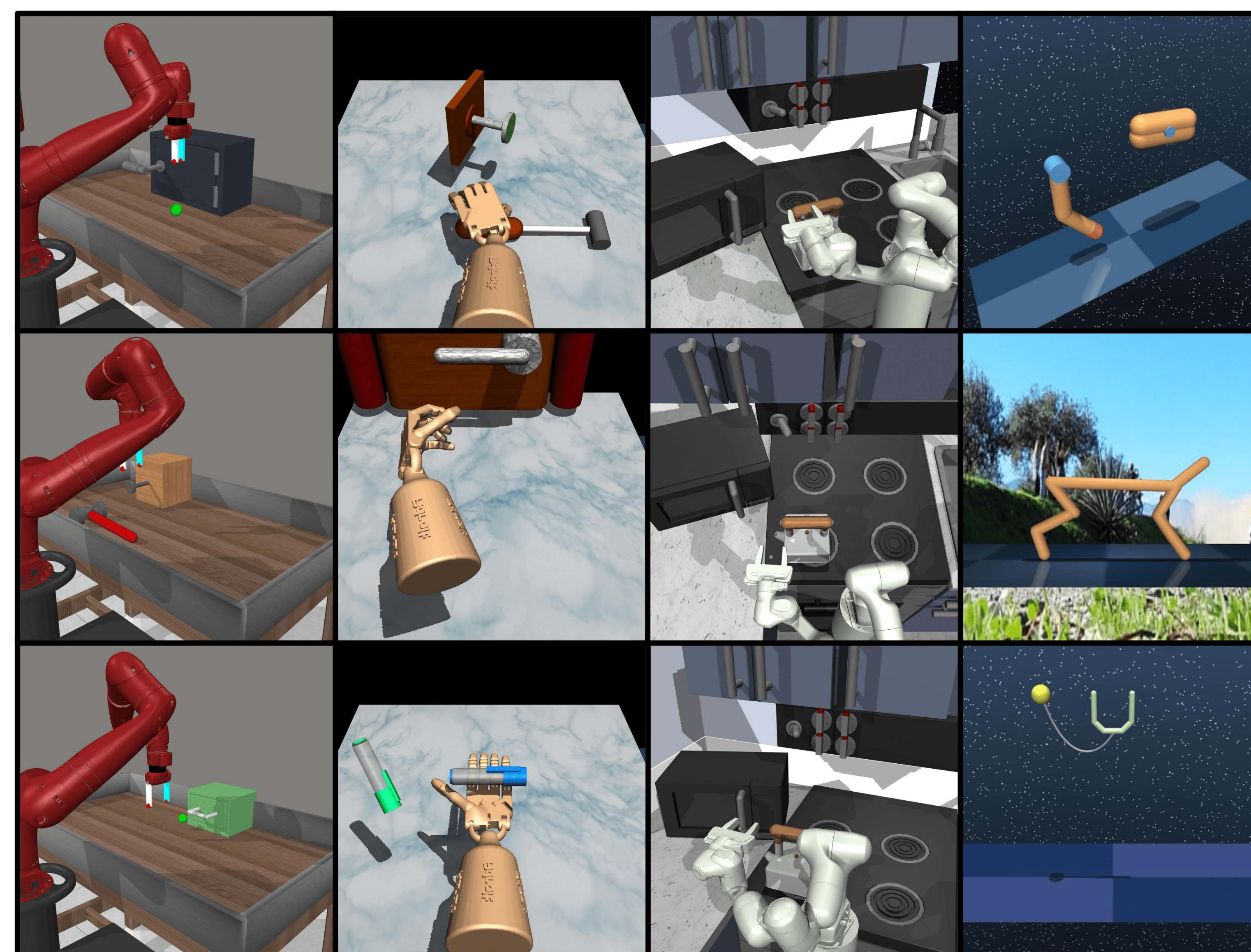




## Structured Environment Agent Representations (SEAR)



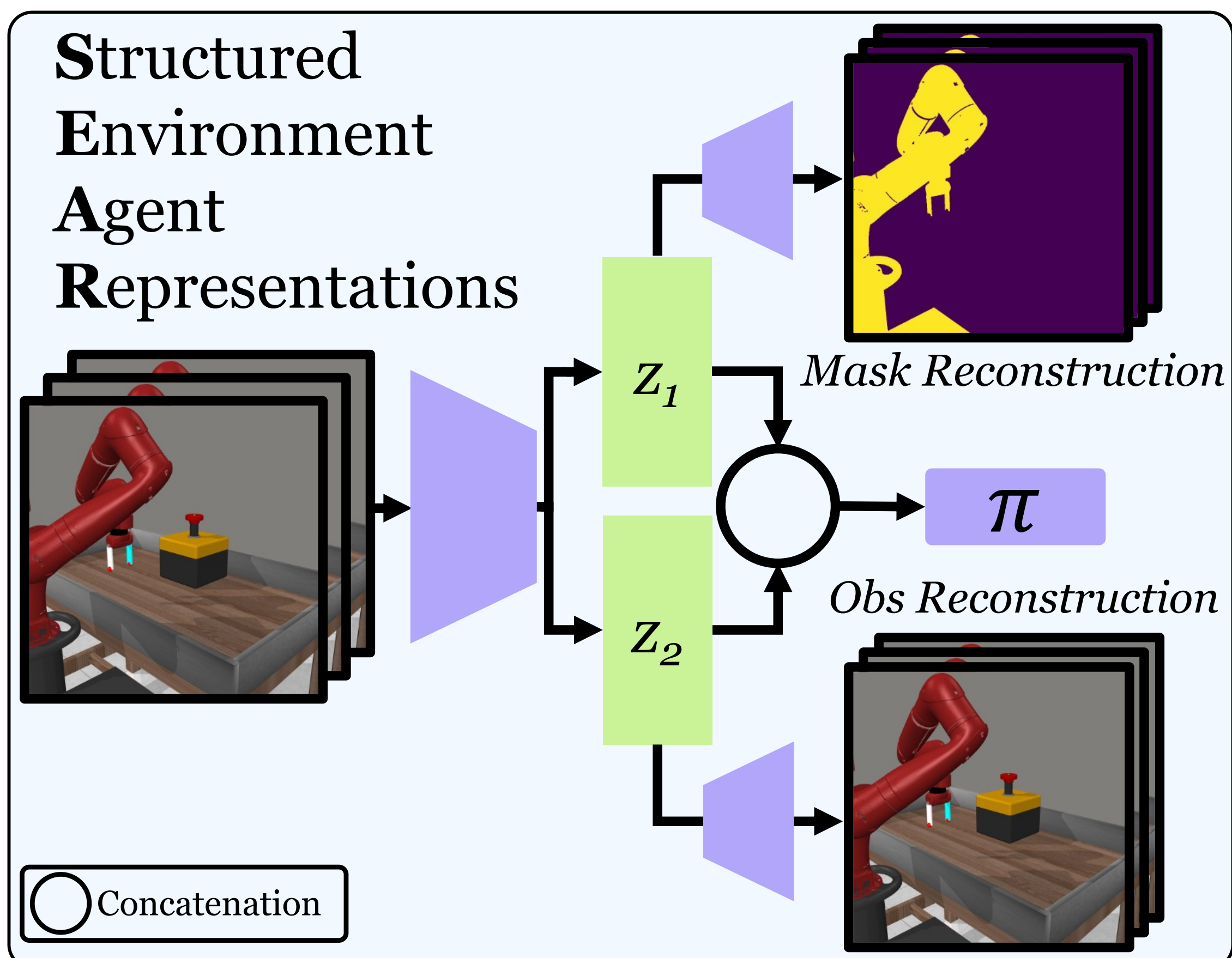
Use **visual knowledge of the agent** to learn the agent-environment split, by **reconstructing the agent mask**



**18 tasks spanning 5 robots across 4 simulation suites**

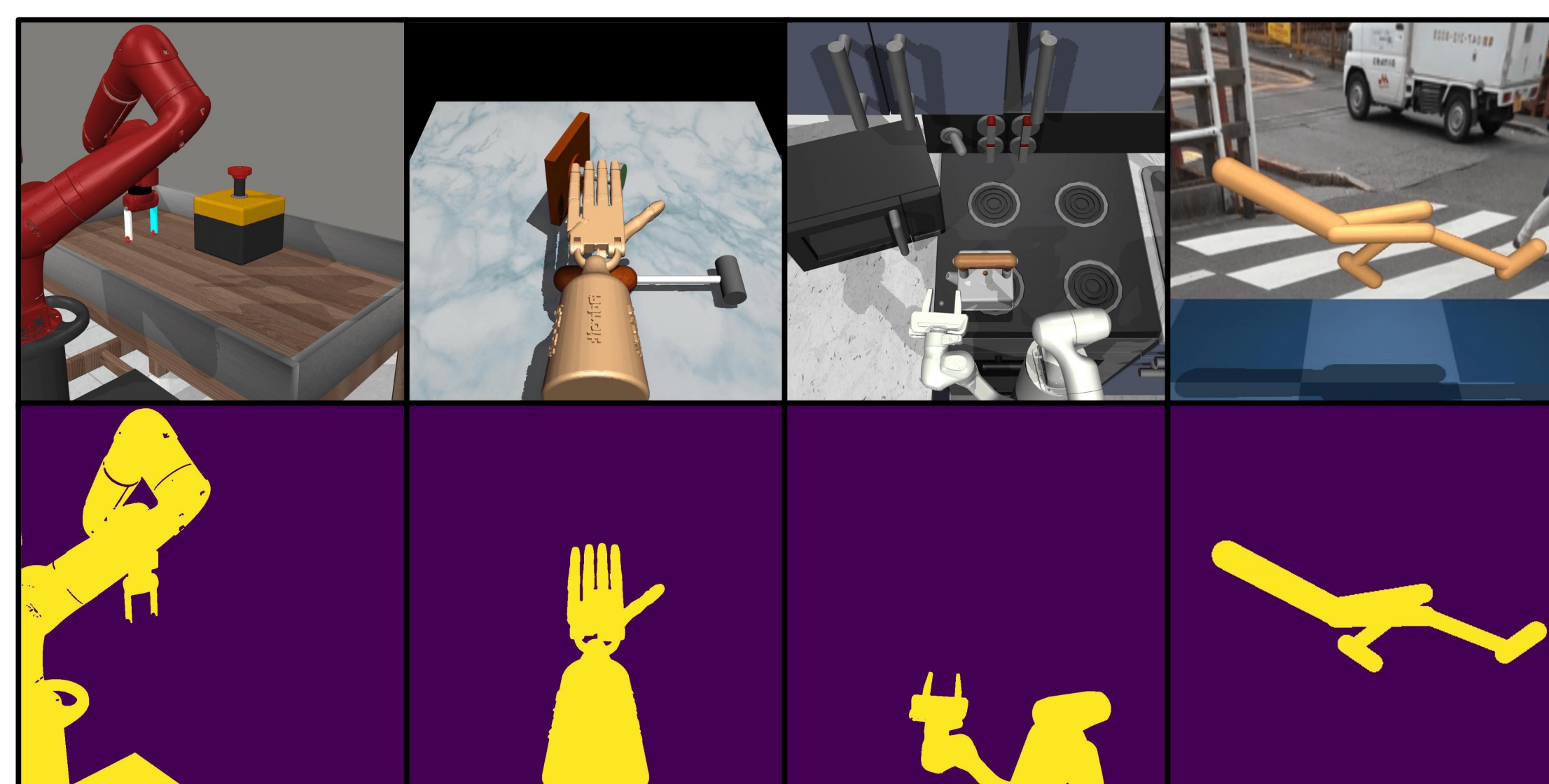
### How can we disentangle agent from environment in a visual RL setup?

We augment the RL loss with agent-centric and environment-centric losses

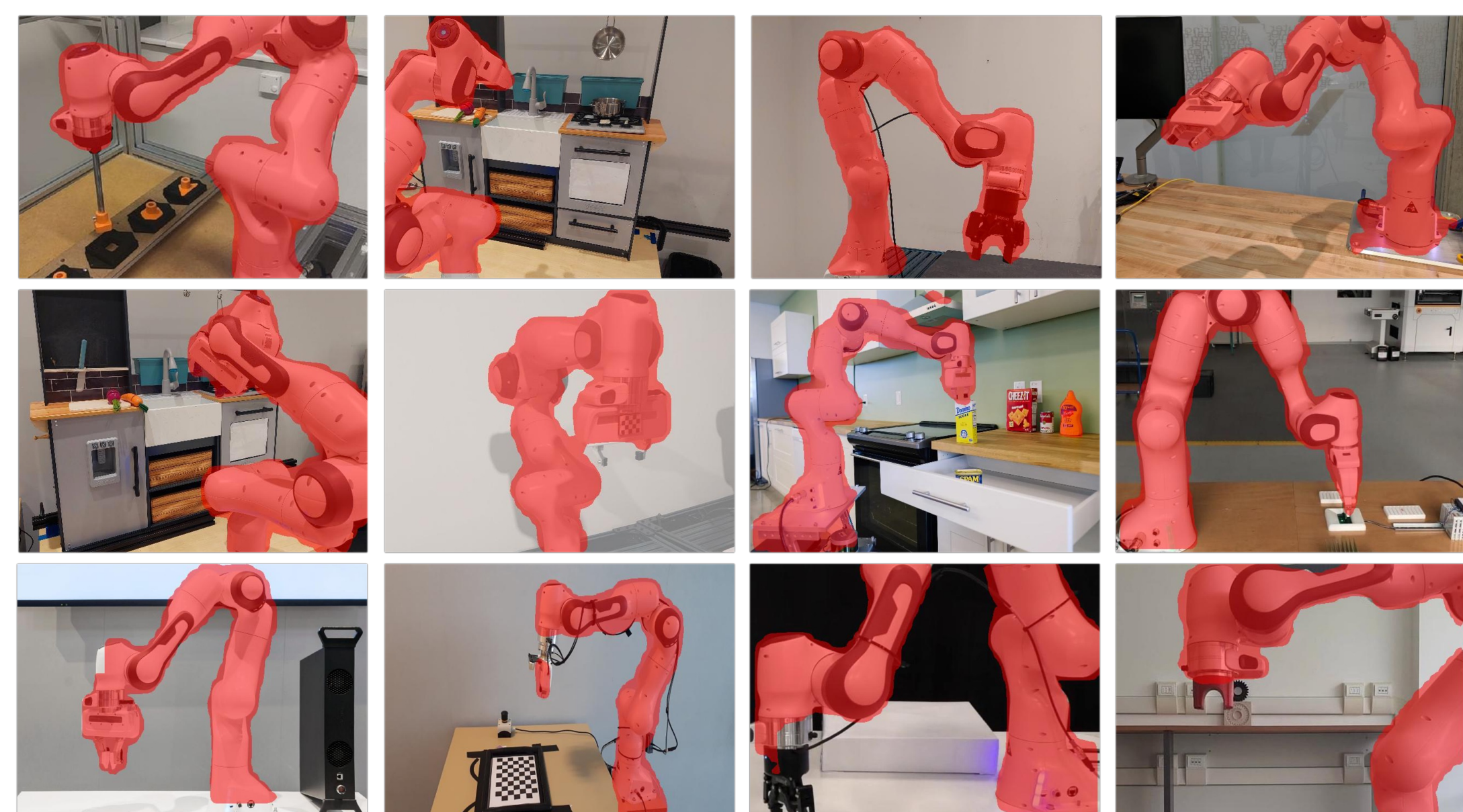


### How do we obtain robot masks?

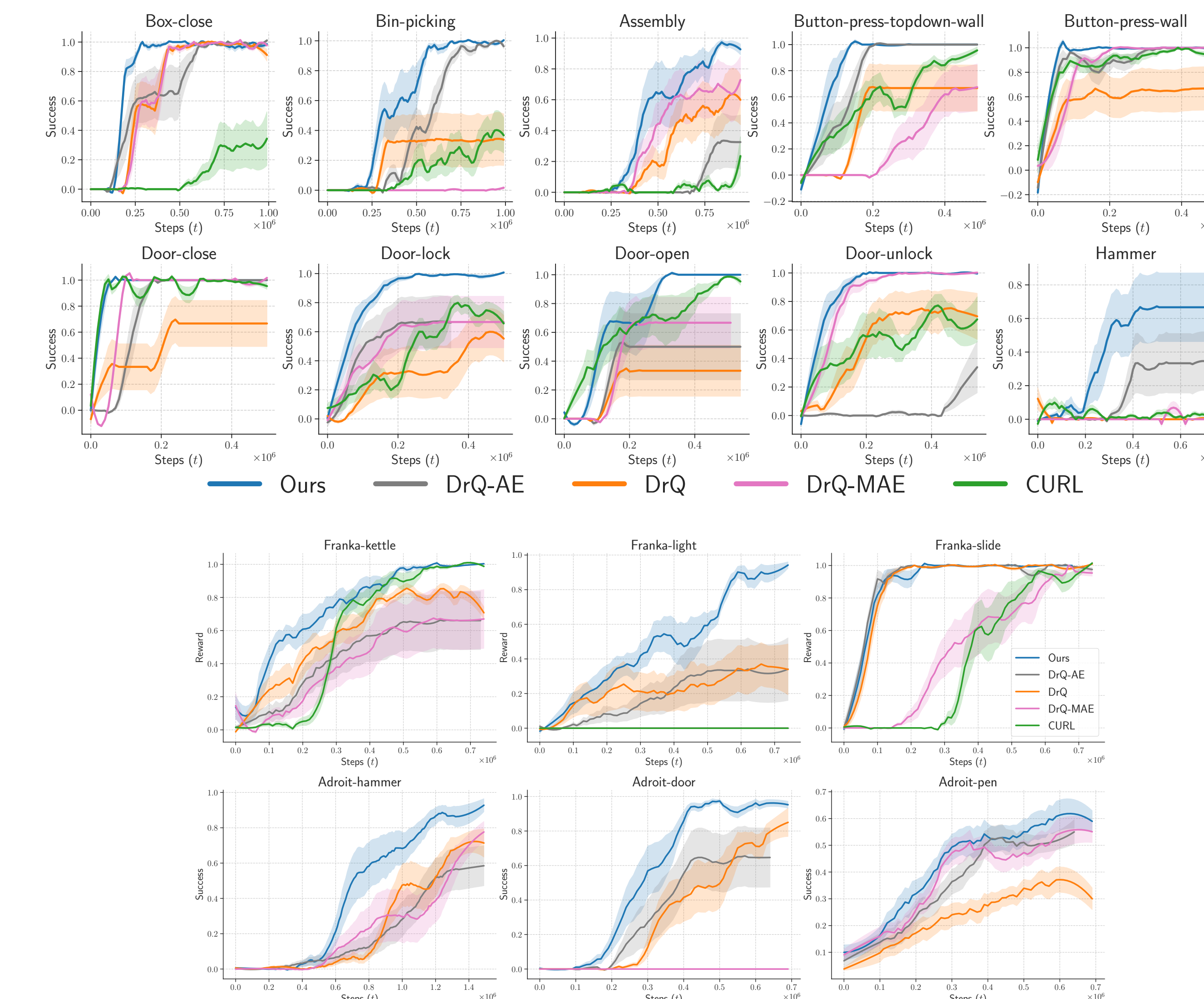
We can directly get masks from a simulator



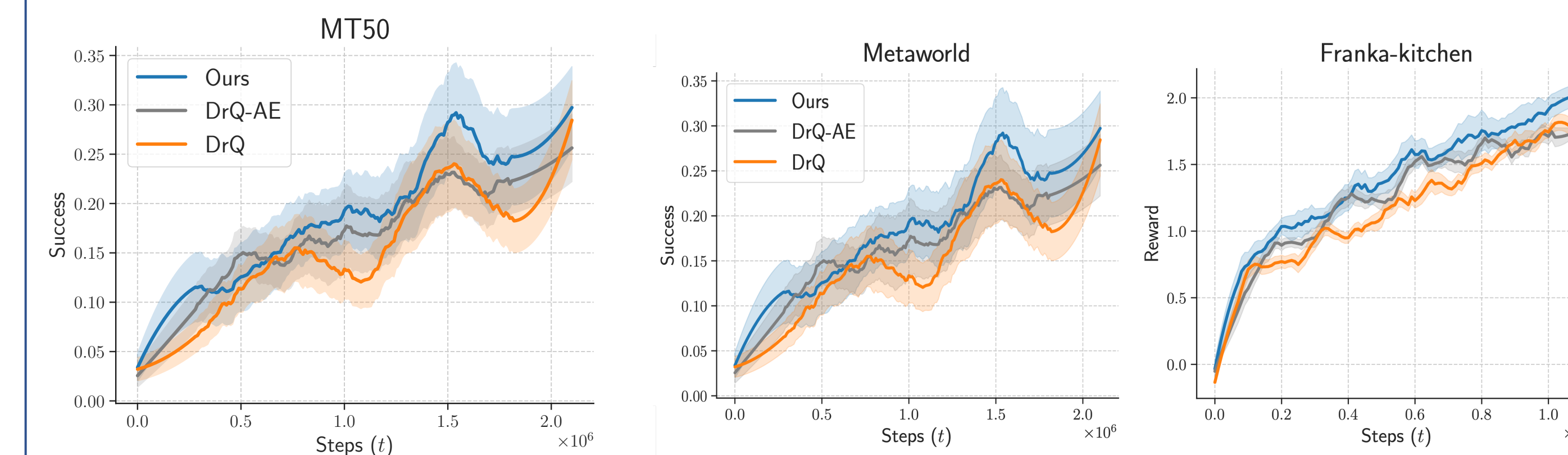
Or fine-tune a segmentation model



### How does it perform in single-task environments?



### What about multi-task environments?



### What about noisy masks?

