Learning Visually Guided Latent Actions for Assistive Teleoperation

Siddharth Karamcheti

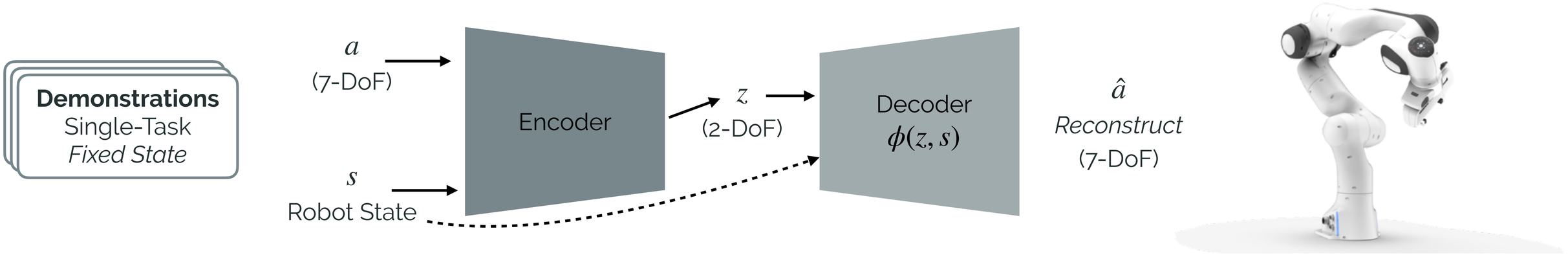
Albert Zhai

Dylan P. Losey

Dorsa Sadigh

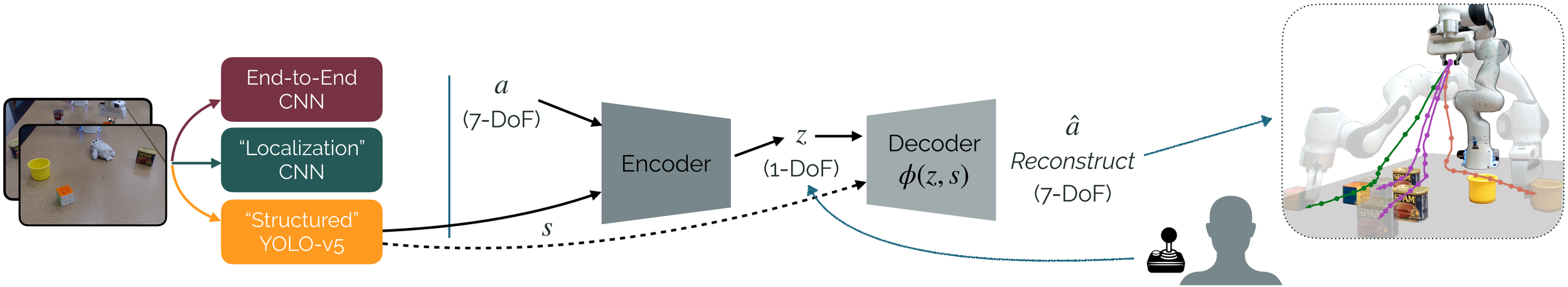
Learned Latent Actions

[Prior Work] Latent Actions — Low-dimensional, task aware controllers for high-dimensional robots.



[Problem] Generalization — dynamic, changing states; few-shot learning "similar" tasks!

[This Work] Perception — how to encode visual information and enable generalization?



Visually Guided Latent Actions — Results

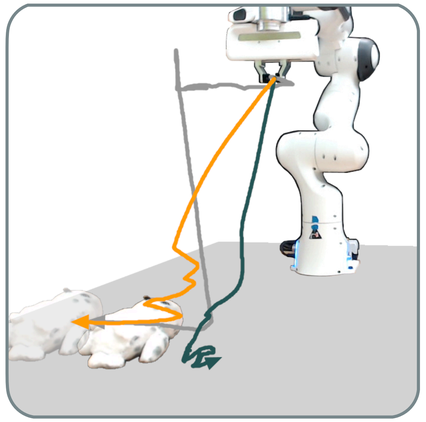
User Study — Which controllers are effective?

Few-shot generalization performance from *just 3 demonstrations?*

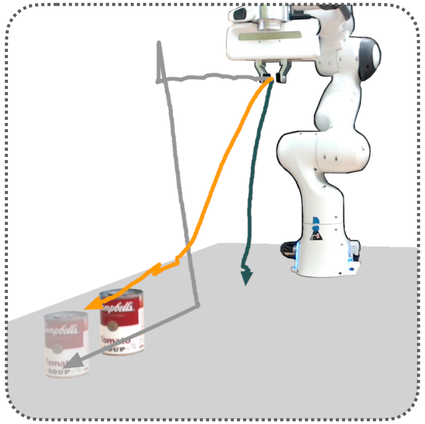
End-Effector Control
(6 x 1 DoF)

Localization-Only
(Oracle Classification)

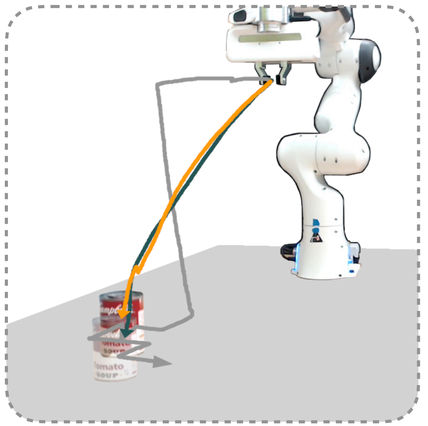
YOLO-v5
(Structured & Pretrained)



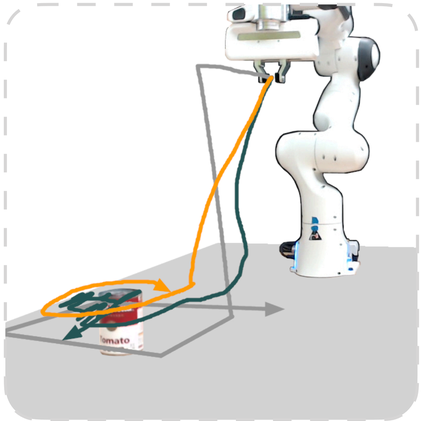
Train Task
Push the animal west.



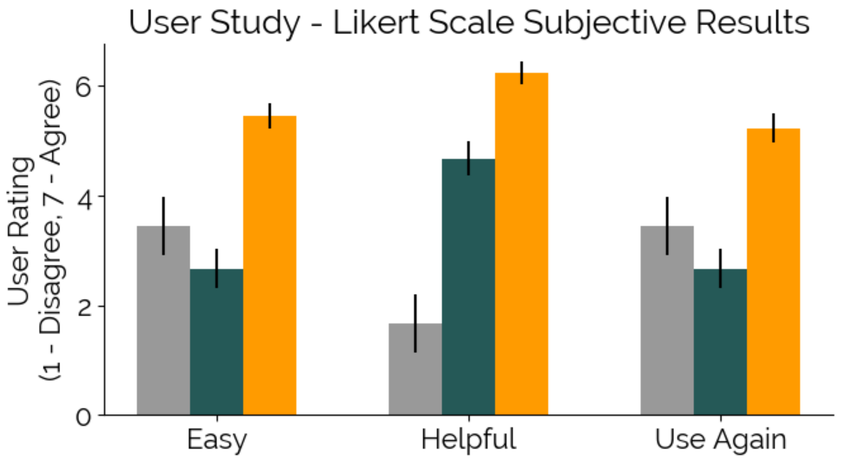
Seen Task (Few-Shot)
Push the soup can south.



Near Task (Few-Shot)
Push soup can southeast.



Far Task (Few-Shot)
Rotate around the soup.



Check out our paper for more details on methods, simulations, and qualitative results!

Thanks so much!

If you have questions or want to chat, feel free to email me — skaramcheti@cs.stanford.edu