Learning Visually Guided Latent Actions for Assistive Teleoperation

Siddharth Karamcheti, Albert J. Zhai, Dylan P. Losey, Dorsa Sadigh
skaramcheti@cs.stanford.edu | https://youtu.be/6nTH5ALsJQ

Latent Actions: Low-dimensional, task aware controllers for high-dimensional robots.

Problem: Generalizing to new tasks & objects; *what’s the right inductive bias for perception?*

User Study:
Few-Shot Generalization from *just 3 demos?*

- **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.*