Coordinating Humanoid Loco-Manipulation through Action-Conditioned Latent Foresight

Supplementary Material

Unseen object configurations

Table S1: Unseen object configurations.

Task	Configurations
Dispose of Trash	Tissue packet; tea bottle.
Collect Toy with Cart	Two arrangements of the blue target toy and two neighboring distractors.
Water Plant	Watering can shifted left; shifted right; spout reversed.
Pack and Stow Power Strip	Power strip rotated; repositioned on the tabletop.
Load Washer	Purple or pink T-shirt, placed on either the right or left side of the washer top.
Retrieve Shirt from Drawer	Target T-shirt repositioned; gray T-shirt nearby; pink T-shirt nearby.
Place Basket on Table	Blue-rimmed basket replacing the red-rimmed basket.

Visual subgoals

Dispose of Trash



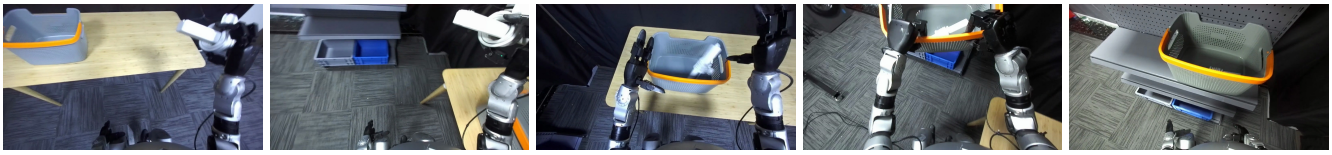
Collect Toy with Cart



Water Plant



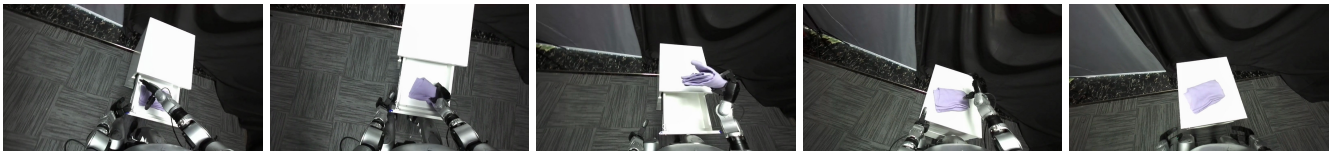
Pack and Stow Power Strip



Load Washer



Retrieve Shirt from Drawer



Place Basket on Table



Figure S1: Visual subgoals extracted with the HIW-500-fine-tuned V-JEPA 2.1 encoder. Each row shows five temporally ordered subgoals from one task demonstration.