

CEBench

Single-arm manipulation (Sim)

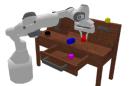
Static Environment



Env A



Env B



Env C



Env D

Instruction



Pick up the blue block and rotate it.

Proprioception



$\Delta T = [-0.9, 0.3, 0.1]$

$\Delta R = [9^\circ, 12^\circ, 3^\circ]$

$\Delta \text{Gripper}$

Bimanual-arm manipulation (Sim)

Seen



lift pot



place bottles

Domain Randomization



place dual shoes



click bell

Bimanual mobile manipulation (Real)

Seen



seen object & background & positions

Domain Randomization



unseen objects & distractors



unseen background
unseen spatial position

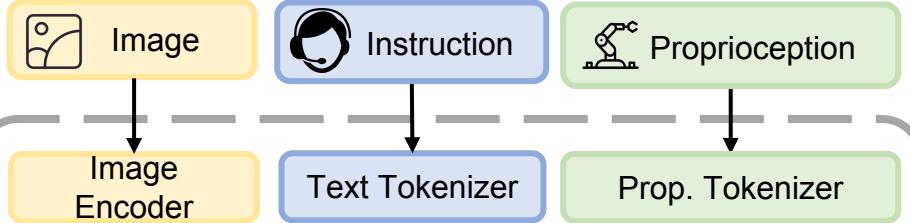
Research Focus

Q1: To what extent does model performance depend on parameter scale, and which techniques enable small models to achieve comparable performance to their larger counterparts?

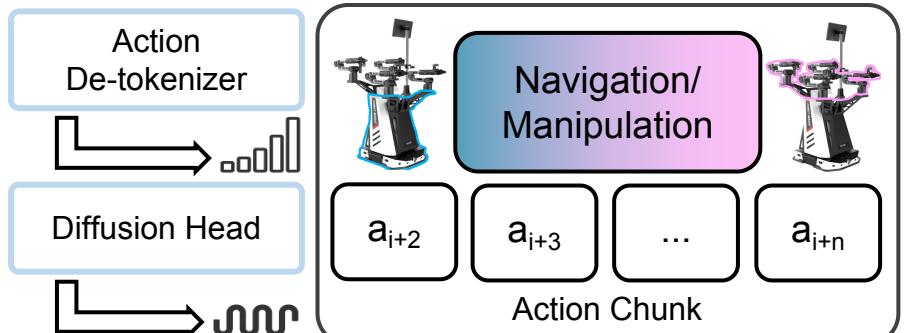
Q2: Is pre-training necessary for small models to accomplish tasks in specific scenarios?

Q3: How to define a hybrid action space for cross-embodiment manipulation, including fixed-base and mobile ones?

Framework of LightVLA



Small Vision-Language Model (SVLM)



Training Pipeline

