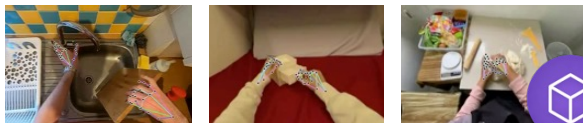


1 Scaling Robotic Manipulation Data

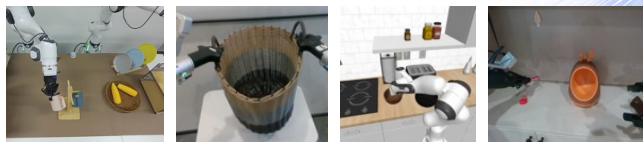
Human-to-Robot Synthesis (15 platforms)



Diverse Robot Embodiments



Multi-source Data Curation



> 38,100 Hours

Heterogeneous Manipulation Data

Robot Data Human Videos Synthetic Data

+ Vision-Language Co-training Data

2 Unified Cross-Embodiment Alignment



Representation Alignment

Shared canonical state vector



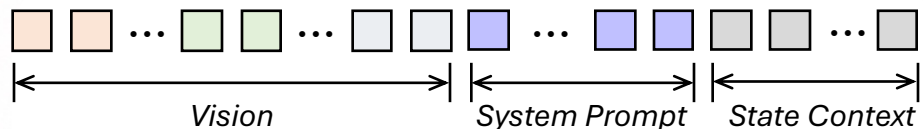
Motion Alignment

Camera-centric consistency



Qwen-RobotManip

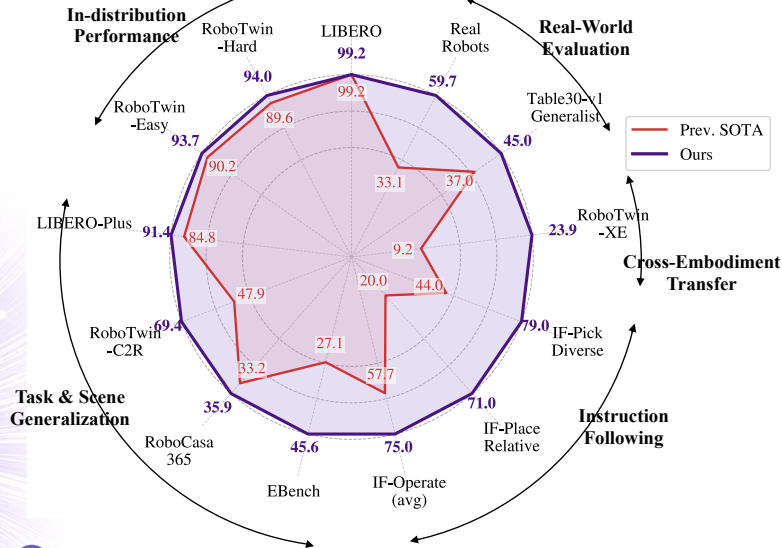
Vision-Language-Action Model



Behavior Alignment

System prompt and In-context adaptation

3 Performance



4 Scaling Law (Emergent Generalization)

