

(a) Object Pose Variation

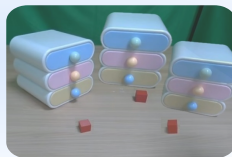
Real-world initial configurations



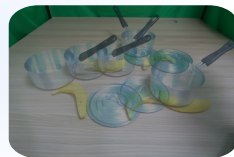
Letter-Aligning



House-Building



Block-Storing



Banana-in-Pot

+18 pp Real Robot Success Rate

(b) EquiVLA



EQUI-PERCEPTOR

Token-level Frame Average
Equivariant Feature Fusion

Pretrained VLM

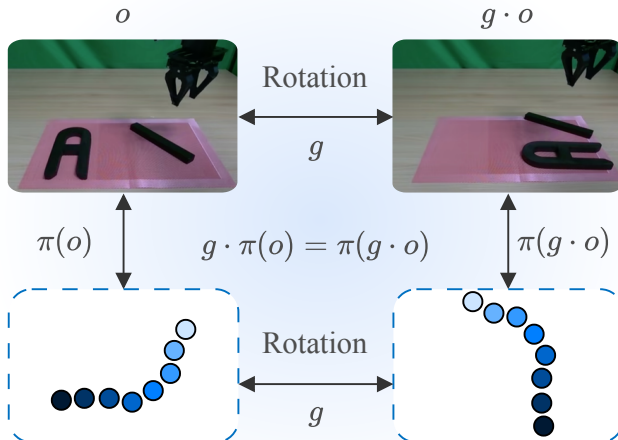


EQUI-ACTOR

Equivariant Flow-matching DiT

Preserve Pretrained VLM Weights

(c) Consistent Task Alignment



27.3×Lower Equivariance Error