Community datasets

SmolVLA

Affordable robots

LeRobot

a_t a_{t+1} … a_{t+h}

Vision-Language Model

Self-Attention

Self-Attention

Self-Attention

KV

QKV

KV

Action Expert

Cross-Attention

Self-Attention

Cross-Attention

Task: Grasp the object and put it in the bin

State

Noised Action