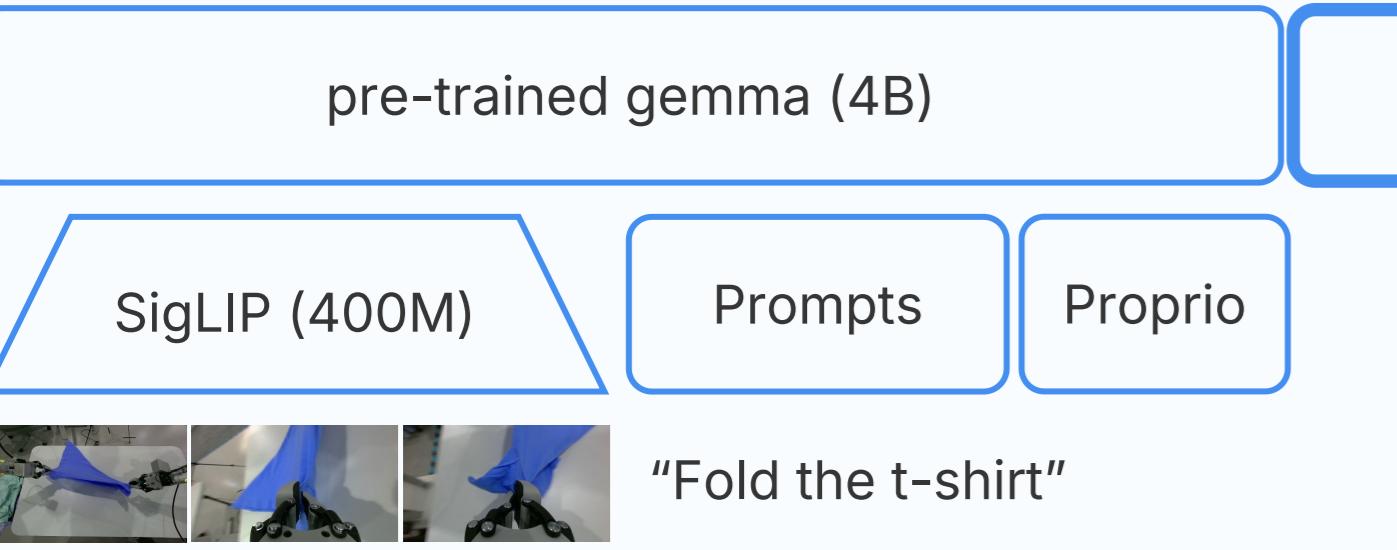
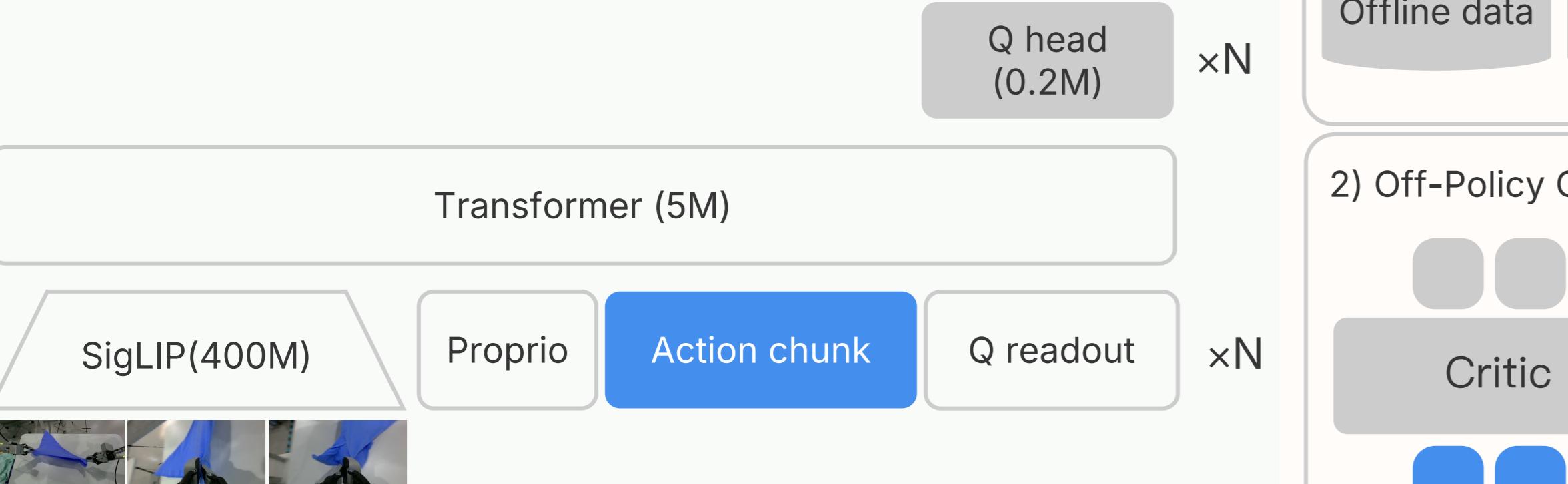


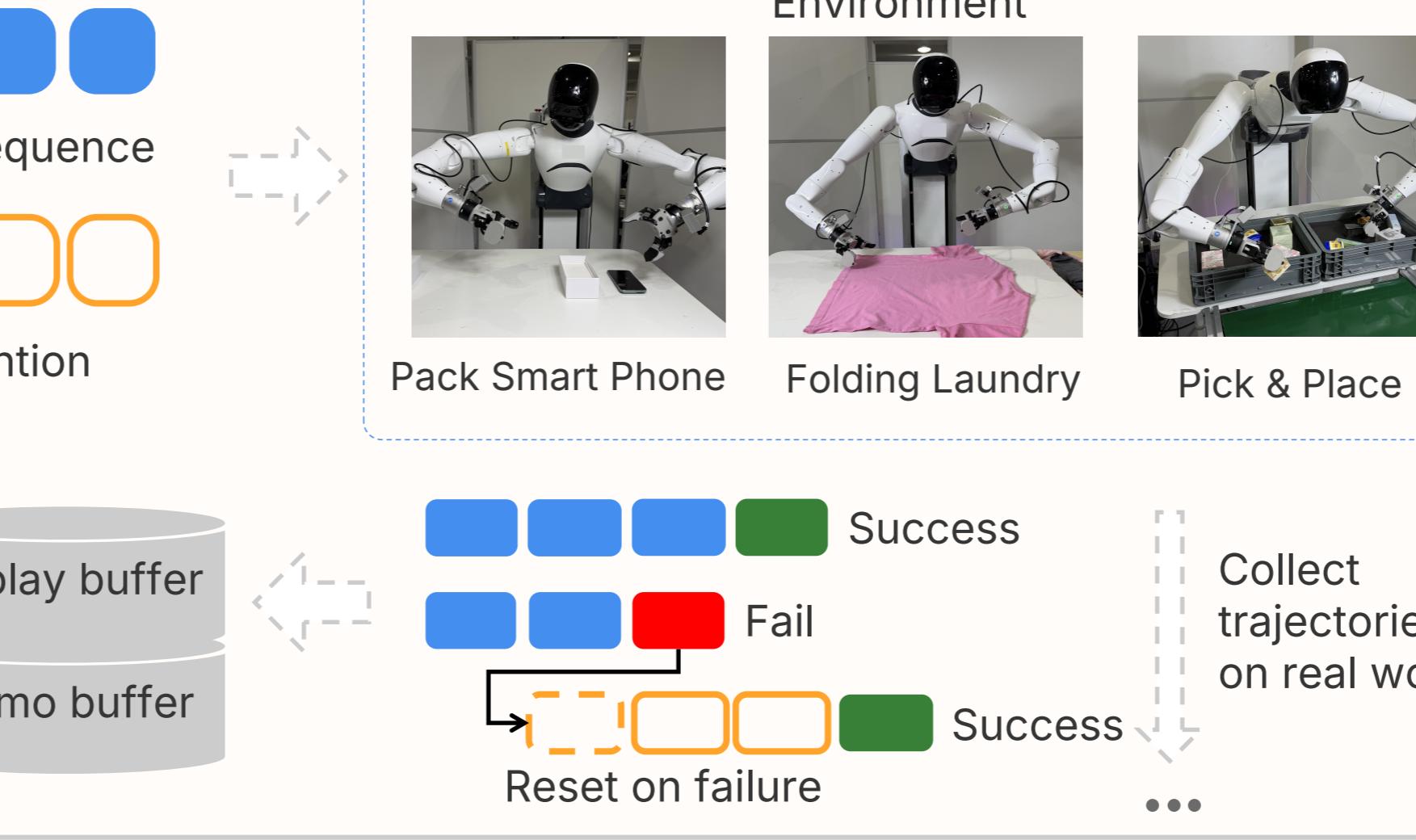
Actor: Flow-Based VLA Model



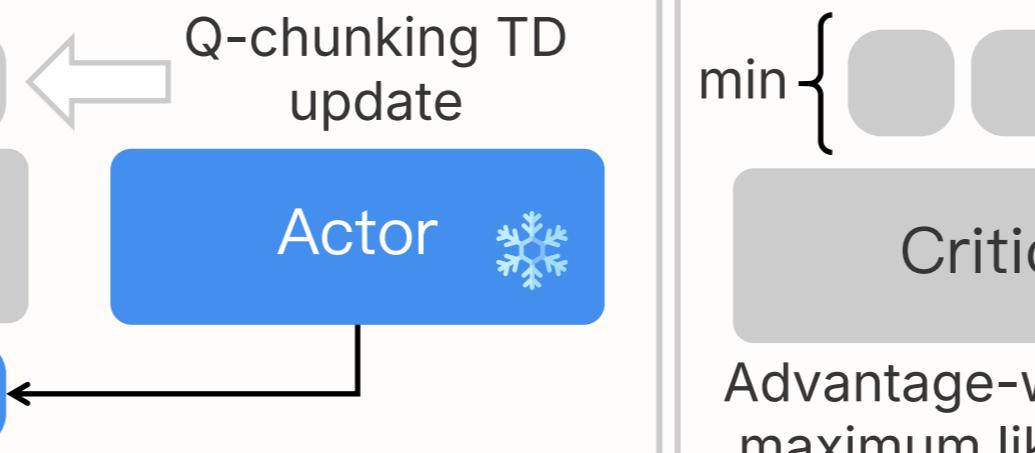
Critic: Chunk-Based Ensemble Q Network



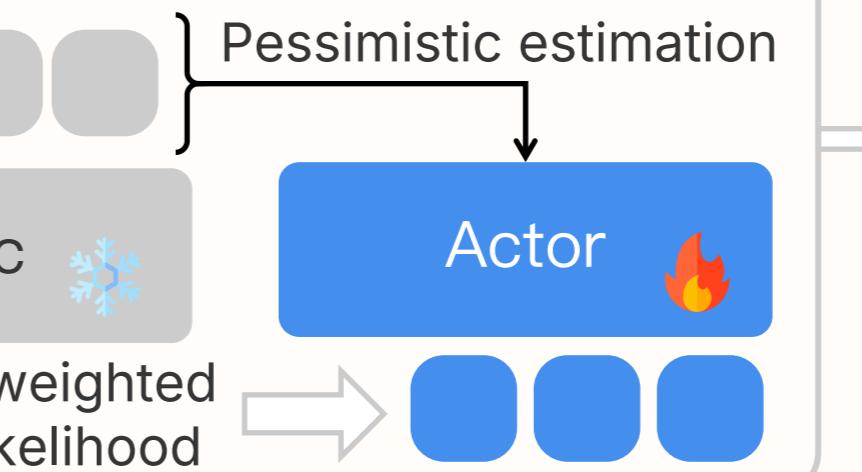
1) Data Collection Under Human Intervention



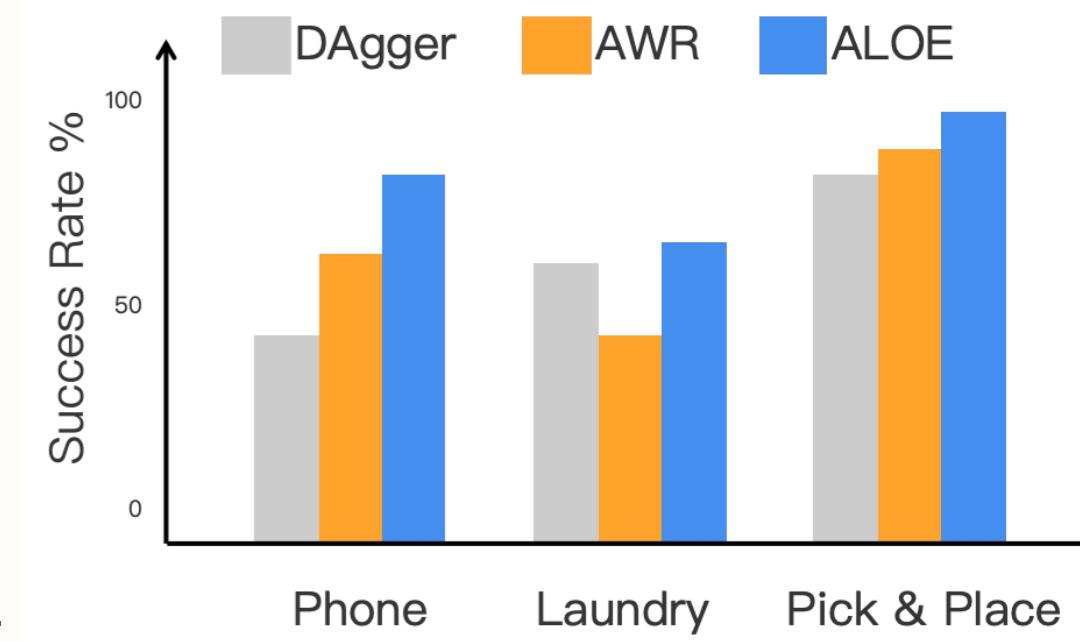
2) Off-Policy Critic Estimation



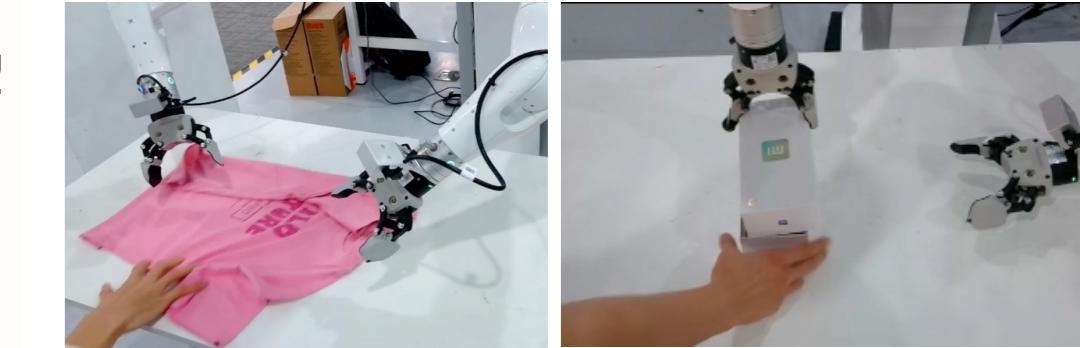
3) Policy Improvement



a) Tasks Performance



b) Robustness against Disturbances



c) OOD Generalization

