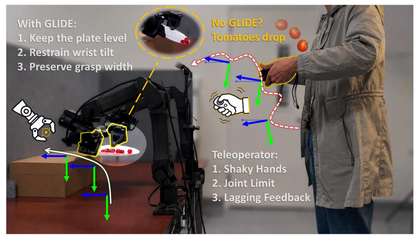
GLIDE: Learning Beyond What Humans Can Demonstrate

Yuchen Song, Aditya Mittal, Unnat Jain
University of California, Irvine

When Teleop Fails

Can robots learn tasks their operators cannot demonstrate?



- Carry a plate with tomatoes, no spills
- Hand over a marker, stand it up
- Pour from a bottle into a glass

Raw VR teleop success: **0%** · **10%** · **0%**

Impossible to demonstrate, but **failures are easy to describe**: the plate tilts, the marker topples…

GLIDE: Generate, Learn, Iterate, Deploy

1 Generate task-specific guardrails

The agent infers failure modes from task description ℓ.

$$G^{(0)}_{\mathrm{GLIDE}} = \mathrm{Write}(\ell, T)$$

2 Learn and Iterate on rollouts

Each round, the agent diagnoses successes and failures, learns from them, and revises the guardrail code.

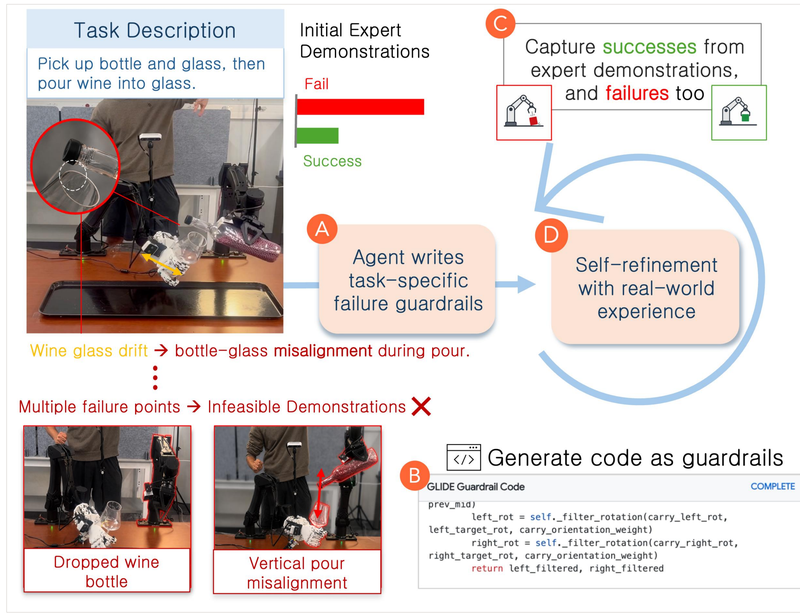
$$\tilde{a}_t = G^{(k)}_{\mathrm{GLIDE}}(p_t, a_t), \quad G^{(k+1)}_{\mathrm{GLIDE}} = \mathrm{Revise}\big(\ell, G^{(k)}_{\mathrm{GLIDE}}, \mathcal{D}^{(k)}\big)$$

3 Deploy the policy with guardrails on

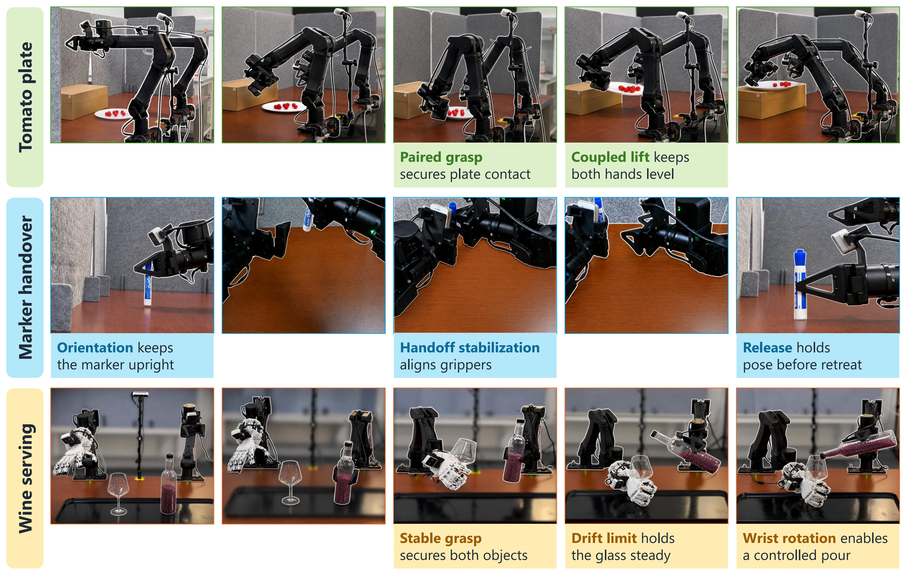
The same guardrails G^* filter the policy's predicted actions.

$$\tilde{a}_t = G^{*}_{\mathrm{GLIDE}}\big(p_t, \hat{\pi}_{\mathrm{mix}}(o_t, \ell)\big)$$

Raw a_t (human or policy) → **Guardrail G** (agent-generated) → **Feasible $\tilde{a}_t$** (robot command)



Phase-Aware Guardrails



Generated guardrails in depth

✨ **emergent**: beyond what an expert would hard-code

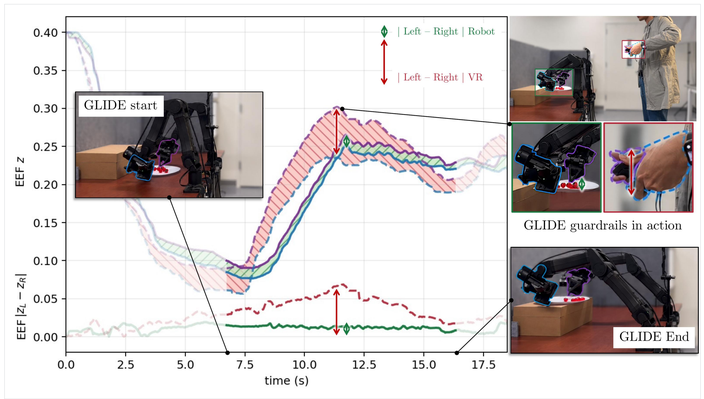
Tomato plate
- Close grippers **in sync** ✨
- **Slow, coupled** carry ✨
- Keep both sides **level**
- Preserve **grasp width**
- Restrain **wrist tilt**

Marker handover
- **Smooth**, capped motion ✨
- Keep grippers **apart** ✨
- **Damp** the handoff ✨
- **Pause 0.55s** at release ✨
- Keep marker **upright**

Wine serving
- **Boost** pour tilt **1.85×** ✨
- Aim mouth **16 cm** over rim ✨
- Keep the glass **upright** ✨
- Hands **≥ 10.5 cm** apart ✨
- **Gentle** grip on the glass ✨

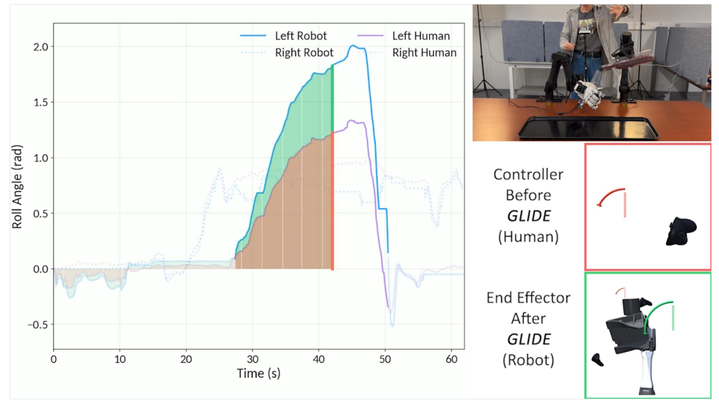
Raw Commands vs. Guarded Motion

Coupled carry keeps the plate level



VR hands drift apart (**red**); GLIDE keeps both grippers level (**green**).

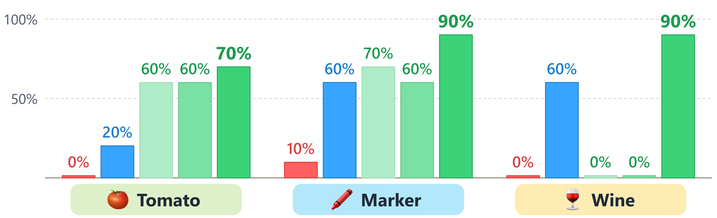
Wrist-roll boost makes pouring easier



GLIDE amplifies wrist roll (**purple** → **blue**) so the bottle reaches a pouring angle.

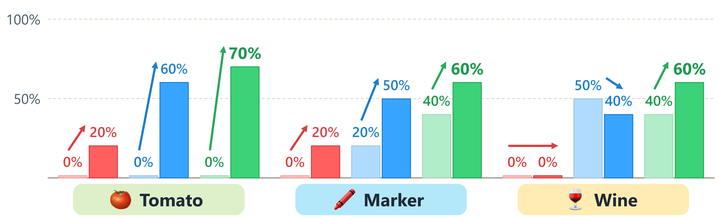
Guarded Demos and Policies

Guarded teleop collects better demos



+70 / +80 / +90 pts over vanilla teleop, **+50 / +30 / +30** over manual guardrails

Guarded policies make use of imperfect demos



Same policy, guard off → on: **0/40/40%** → **70/60/60%**