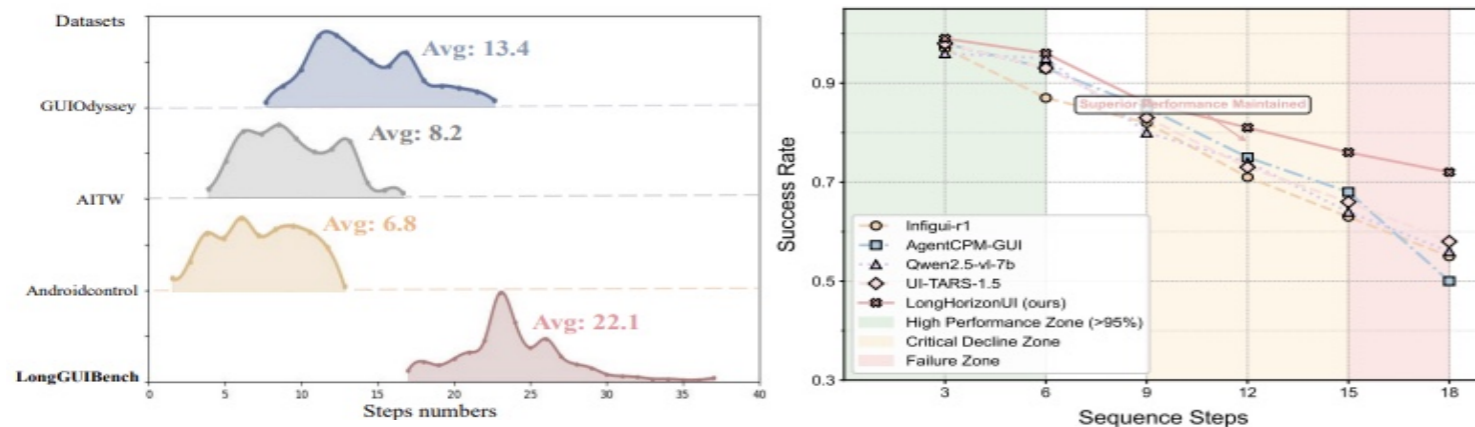
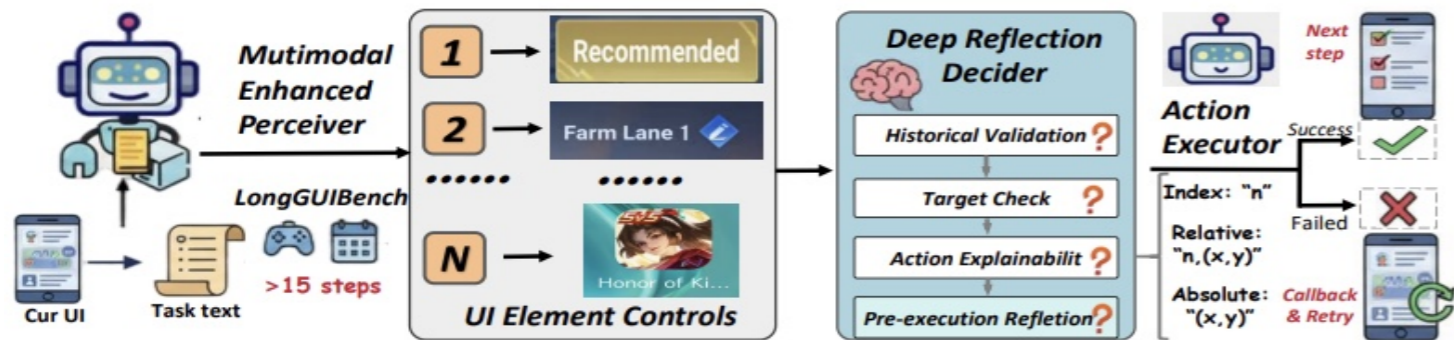


Motivation: Why Long-Horizon GUI Fails?

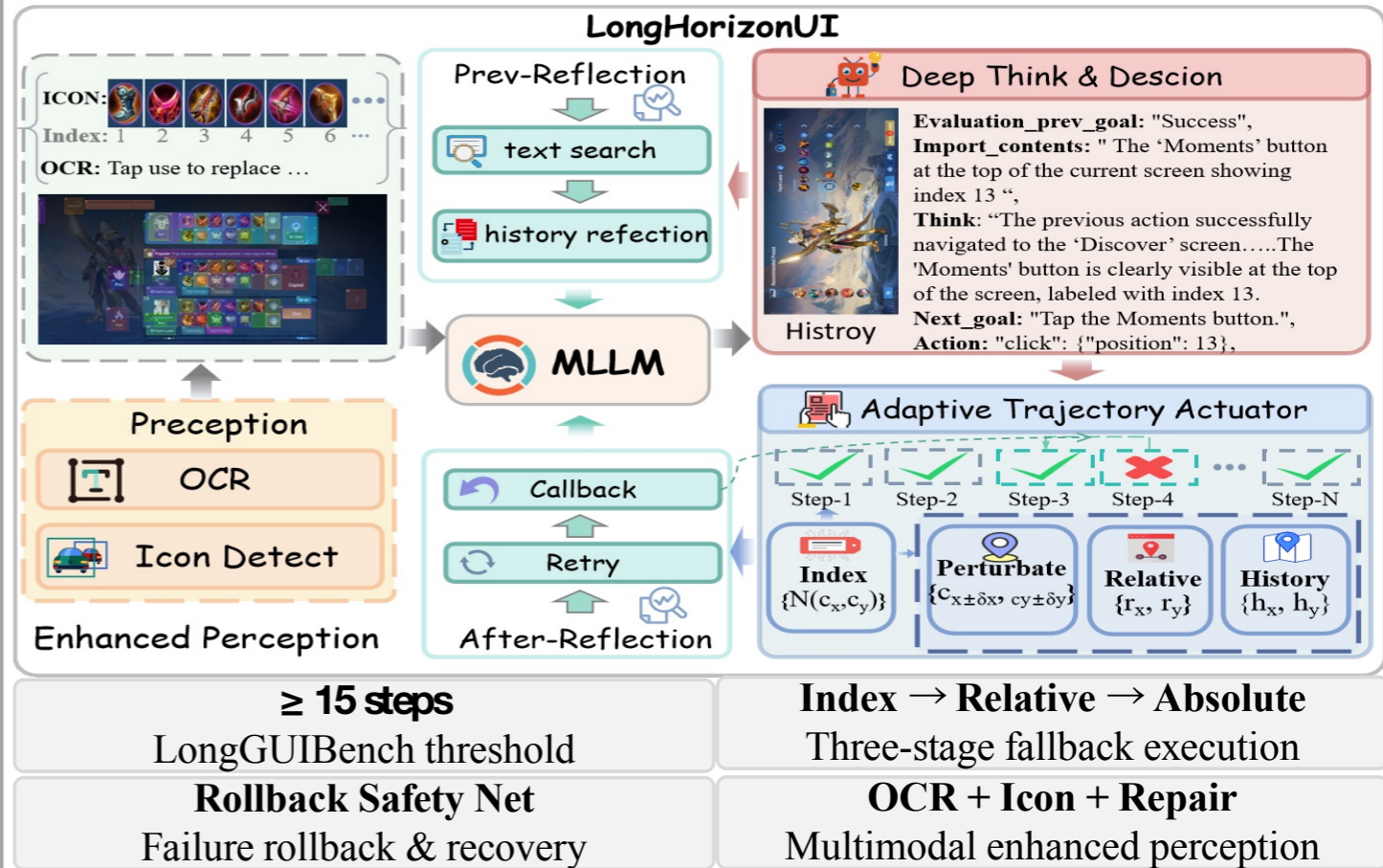


Step insufficient

Errors accumulate



LongHorizonUI: Reflect, Ground, Recover



LongGUIBench: Beyond Short GUI Episodes

Scene: General
Step nums: 18~31

Scene: Game
Step nums: 22~34

371 scenarios across apps and games/Average 22.1 steps

Code: <https://github.com/kane2kang/LongHorizonUI>

Page: <https://kane2kang.github.io/LongHorizonUI/>

Experiments Results

GUI Models	90.0	83.0	79.3	35.6	91.8	61.0	70.4	20.1	66.4
OmniParser (Lu et al., 2024b)	90.0	83.0	79.3	35.6	91.8	61.0	70.4	20.1	66.4
AgentCPM-GUI (Zhang et al., 2025b)	92.1	81.2	82.4	37.1	89.7	66.5	74.1	25.8	68.6
InfiGUI-R1 (Liu et al., 2025)	93.2	79.7	56.7	23.8	92.9	67.2	53.9	19.4	61.8
UI-TARS-1.5 (Qin et al., 2025)	93.6	79.2	75.4	21.8	88.2	69.5	77.8	18.9	65.8
LongHorizonUI	93.5	85.3	78.0	52.3	93.8	83.9	79.7	52.1	77.3

The user opens the "Honor of Kings" application, selects the "Animist" hero in Solo Practice mode, changes its skin to "Duke Lepus", and switches the skill to "Sprint", then closes the background apps.

Task Execution Steps:

- Step1: "click": {"position": 19}
- Step2: "click": {"position": 14}
- Step3: "click": {"position": 40}
- Step4: "click": {"position": 21}
- Step5: "click": {"position": 26}
- Step6: "click": {"position": 6}
- Step7: "click": {"position": 15}
- Step8: "click": {"position": 19}
- Step9: "click": {"position": 36}
- Step23: "press key": {"recent"}
- Step24: "click": {"position": 58}

Final actions: Oppo app → Solo Practice mode → Changes hero's skin → Switches the skill → Closing process