



从定位到行动：GUI Agent的技术演进

AgenticAICon 2026

汇报人：唐飞 · 浙江大学 ZJU-REAL Lab

2026 年 7 月

0. 报告大纲

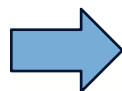
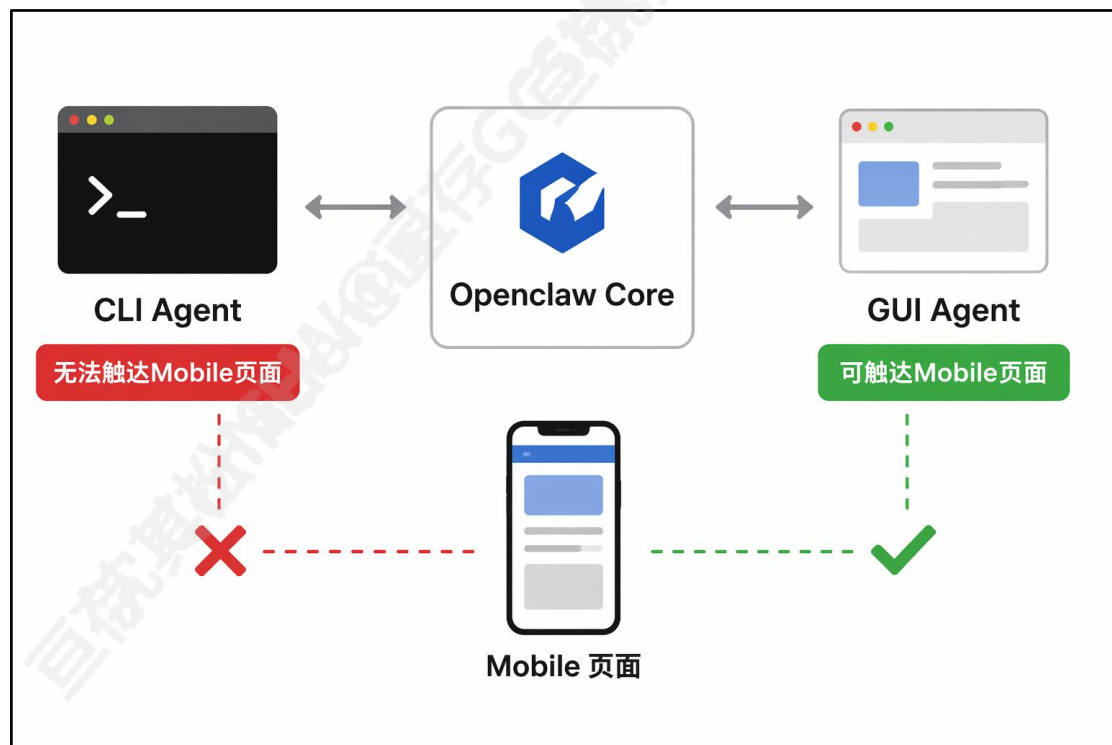
- 01 **GUI Agent: 连接模型与数字世界**
通用智能体 · 真实软件 · 图形界面 · Computer Use
- 02 **界面感知: GUI Agent的基础能力**
GUI Grounding · 数据构建 · 目标定位 · 静态模仿
- 03 **环境交互: 从推理走向行动**
Reward Design · Single-turn RL · Online RL · 环境反馈
- 04 **自我进化: 通向通用 CUA**
任务合成 · Skill Evolution · 安全泛化 · 未来范式

1. CLI V.S. GUI Agent



➤ CLI Agent 擅长调用命令行、API 和外部工具，但受限于接口可用性，**难以覆盖大量以 GUI 为主要入口的真实软件。**

➤ GUI Agent 直接感知屏幕并执行界面动作，为智能体提供了一种**不依赖专用 API** 的通用软件交互方式。



**Control Your
GUI, Anywhere**

1. CLI V.S. GUI Agent



Control Your
GUI, Anywhere

1.GUI Agent 前沿应用

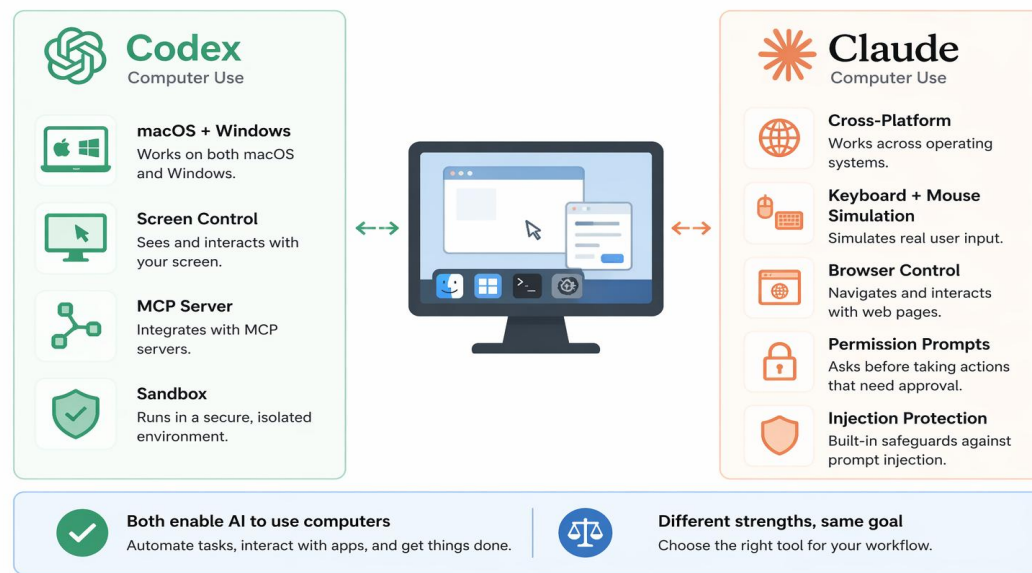


- **移动端应用**: 系统级手机助手从语音指令出发, 跨 App 完成搜索、比价、下单、信息整理等真实任务
- **桌面端应用**: Codex & Claude Computer Use 通过截图理解界面, 并使用鼠标、键盘与应用交互, 面向开发、办公自动化

前沿应用



豆包手机助手



Codex & Claude Computer-Use

[1] <https://platform.claude.com/docs/en/agents-and-tools/tool-use/computer-use-tool>

[2] <https://o.doubao.com/>

1. 国内外GUI进展



趋势: 从基础的界面操作, 到浏览器任务执行, 再到原生集成与能力增强, GUI Agent 正在成为 AI 智能体的核心能力, 推动 AI 从对话迈向行动

[1] <https://www.anthropic.com/news/3-5-models-and-computer-use>

[2] <https://openai.com/zh-Hans-CN/index/computer-using-agent>

[3] <https://blog.google/innovation-and-ai/models-and-research/google-deepmind/gemini-computer-use-model/>

[4] <https://qwen.ai/blog?id=qwen3.7-plus>

[5] https://seed.bytedance.com/en/seed2_1

1. GUI Agent的挑战

技术挑战

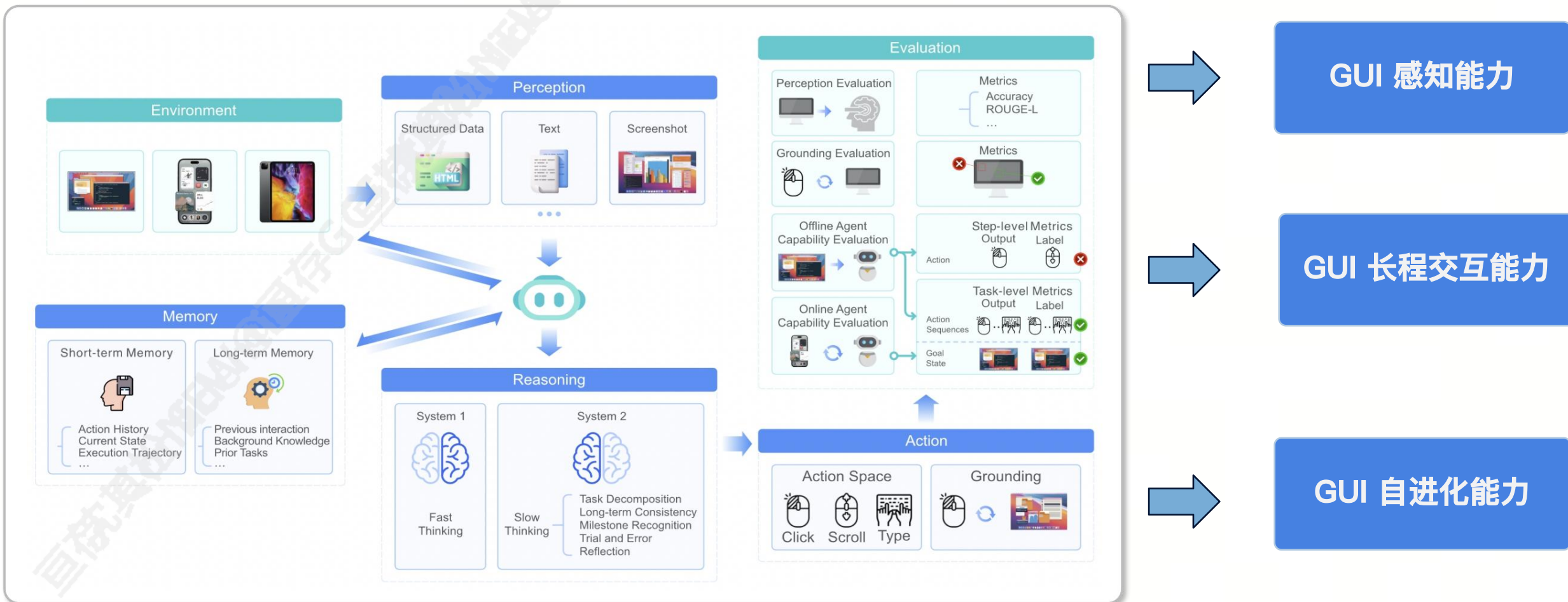


从细粒度 Grounding & Planning，到 Long-Horizon RL，再到 Self-Evolution，GUI Agent 的发展路径，本质上是让模型从“会看懂界面、规划动作”，成长为“能长程交互、持续进化”的数字世界智能体。

2. GUI Agent的核心能力

核心能力

GUI Agent 的核心能力构建，主要围绕界面感知与定位、长程交互决策优化，以及面向持续的自我进化展开

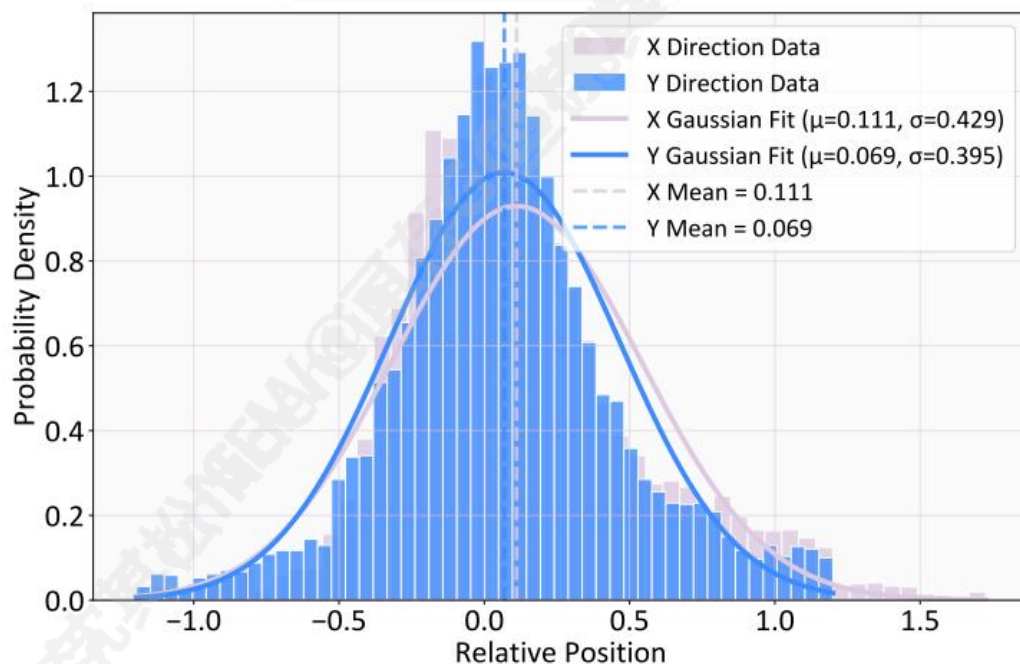


2. Capability: GUI Grounding

研究动机

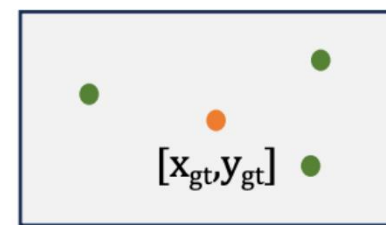
现有 RL 奖励多依赖“命中/未命中”的二值反馈，训练信号较为稀疏。同时，这类奖励难以区分“接近目标”和“完全偏离”，无法刻画真实点击中的连续空间偏差

点击分布可视化



现有奖励

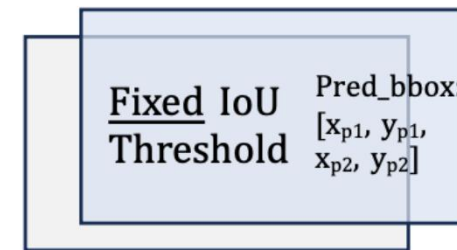
- Positive, Reward = 1
- Negative, Reward = 0



Gt_bbox: $[x_{g1}, y_{g1}, x_{g2}, y_{g2}]$

(a) GUI-R1

- IoU $\geq t$, Reward = 1
- IoU $< t$, Reward = 0



Gt_bbox: $[x_{g1}, y_{g1}, x_{g2}, y_{g2}]$

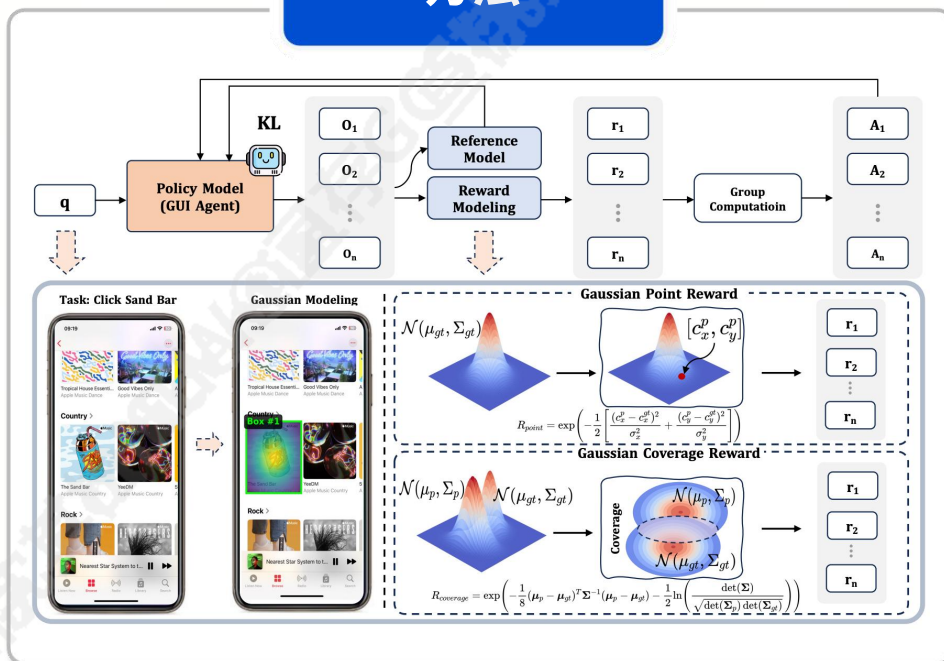
(b) InfiGUI-R1

2. Capability: GUI Grounding

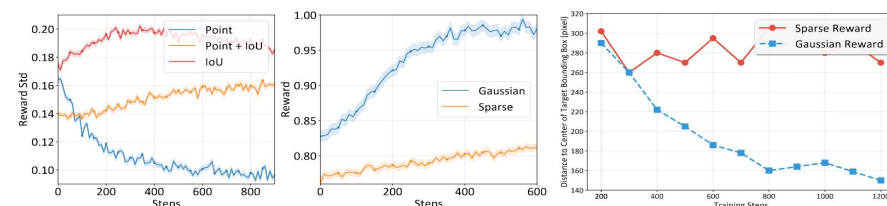
解决方案

- 人类点击位置在目标中心周围呈现**高斯分布**规律，启发我们将 GUI 元素边界框建模为二维高斯分布，从而提供**连续奖励信号**。
- 高斯点奖励与覆盖奖励协同提供精度与覆盖，连续空间内平滑梯度避免边界消失，并结合**元素自适应机制**提升不同 GUI 元素的鲁棒性与泛化。

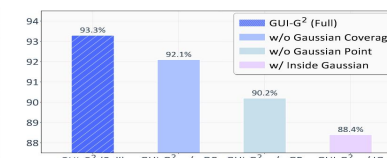
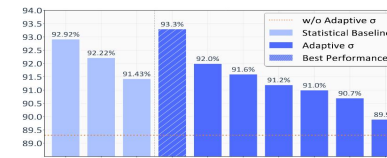
方法



结果



Model	CAD		Dev		Creative		Scientific		Office		OS		Avg.	
	Text	Icon	Text	Icon	Text	Icon	Text	Icon	Text	Icon	Text	Icon	Text	Icon
UI-TARS-7B	20.8	9.4	58.4	12.4	50.0	9.1	63.9	31.8	63.3	20.8	30.8	16.9	47.8	16.2
UI-TARS-72B	18.8	12.5	62.9	17.2	57.1	15.4	64.6	20.9	63.3	26.4	42.1	15.7	50.9	17.6
GUI-G ² -7B	55.8	12.5	68.8	17.2	57.1	15.4	77.1	24.5	74.0	32.7	57.9	21.3	64.7	19.6



2. Capability: GUI Grounding

WebAgentLab 互相关注

GUI-G²: GAUSSIAN REWARD MODELING FOR GUI GROUNDING

10/10

Fei Tang^{1,2}, Zhangxuan Gu², Zhengxi Lu¹, Xuyang Liu², Shuheng Shen², Changhua Meng², Wen Wang¹, Wenqi Zhang¹, Yongliang Shen¹, Weiming Lu¹, Jun Xiao¹, Yueteng Zhuang¹

¹Zhejiang University, ²Ant Group
{flysugar, syi}@zju.edu.cn shuheng.ssh@antgroup.com

GitHub: <https://github.com/zju-real/GUI-G2>
Project: <https://zju-real.github.io/GUI-G2>

Figure 1: GUI grounding performance and human click behavior. Left: Performance comparison of various models on ScreenSpot-Pro. Right: Human click distribution from AITW (Rawles et al., 2023) reveals natural Gaussian patterns around target centers ($\mu = 0.111$, $\sigma = 0.429$), validating our design choice of continuous Gaussian rewards over discrete binary feedback.

被Web Agent Lab 评为2025年高引Paper TOP10

2025年10篇高引 GUI Agent 论文梳理!

ZJU-REAL / GUI-G2

Code Issues Pull requests Agents Actions Projects Wiki Security and quality Insights Settings

GUI-G2 Public

1 Branch 0 Tags

Go to file Add file Code

About [AAAI 2026] GUI-G²: Gaussian Reward Modeling for GUI Grounding
arxiv.org/abs/2507.15846

Assets

File	Added
assets	10 months ago
docs	7 months ago
src/gui_g2	10 months ago
AAA26_main-6419-poster.pdf	4 months ago
AAA26_main-6419-slides.pdf	4 months ago
Dockerfile	11 months ago
GUI_G2_Gaussian_Reward_Modeling_for_GUI_Grounding.pdf	11 months ago
README.md	2 months ago
example_training_json.json	11 months ago
setup.sh	11 months ago

GUI-G²: Gaussian Reward Modeling for GUI Grounding

A Gaussian dense reward framework for GUI grounding training

Deployments 53

github-pages 2 months ago

+ 52 deployments

AK @_akhalik

GUI-G²

推特关注量12K

Gaussian Reward Modeling for GUI Grounding

10:51 AM · Jul 22, 2025 · 12.9K Views

Hugging Face

Search models, datasets, users...

Models Datasets Spaces Community Docs Enterprise Pricing

Get trending papers in your email inbox once a day! Subscribe

Daily Papers

Search any paper with AI

GUI-G² 获得 HF Daily Paper Rank 1

GUI-G²: Gaussian Reward Modeling for GUI Grounding

MiroMind-M1: An Open-Source Advancement in Mathematical Reasoning via Context-Aware Multi-Stage Policy Optimization

The Invisible Leash: Why RLVR May Not Escape Its Origin

WebShaper: Agenticly Data Synthesizing via Information-Seeking Formalization

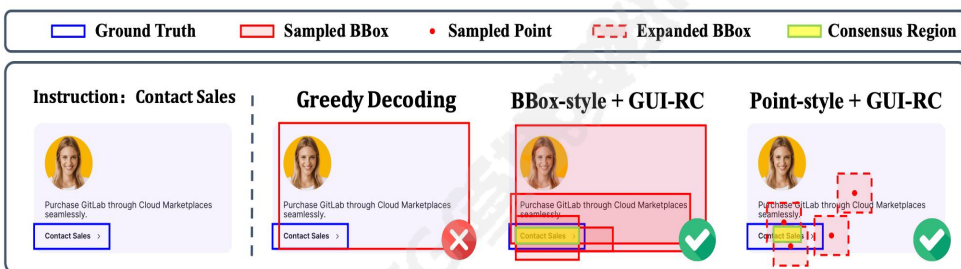
Robust 3D-Masked Part-level Editing in 3D Gaussian Splatting with Regularized Score Distillation Sampling

2. Capability: GUI Grounding

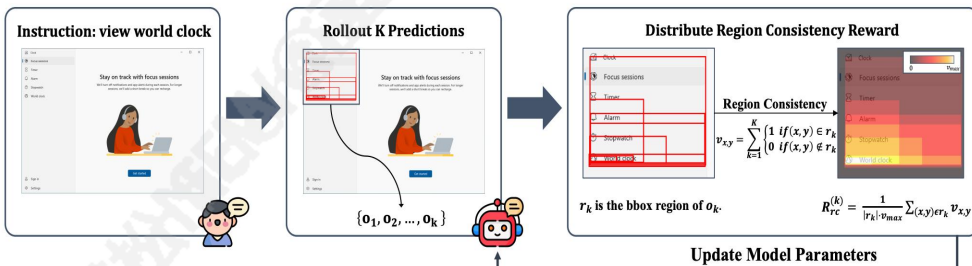
研究问题

GUI Grounding 面临标注成本高、长尾界面覆盖难的问题。通过模型自身多次预测的一致性，可构造伪标签与奖励信号，实现无标签 Grounding 自我进化

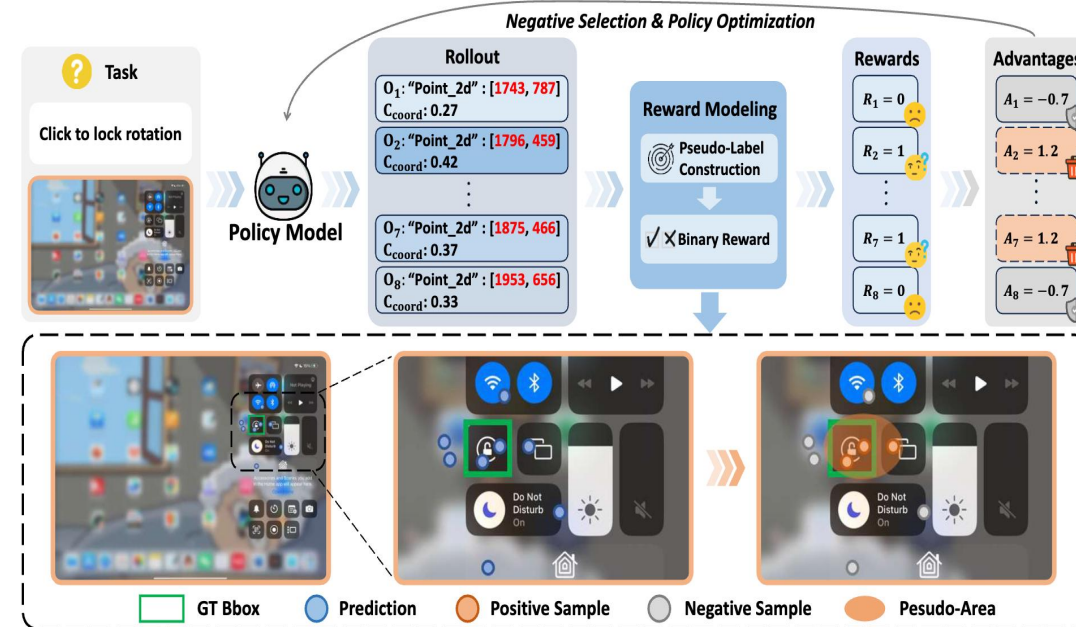
GUI-RC: A Test-time Scaling Approach for GUI Grounding



GUI-RCPO: A Test-time Reinforcement Learning Approach for GUI Grounding



在测试时对同一指令进行多次采样，利用预测区域一致性构造奖励信号，实现无需训练的 test-time 进化



在训练时利用模型预测置信度构造伪标签与负向奖励，在无人工标注框的情况下提升 GUI Grounding 能力

[1] Du, Yong, et al. "Test-time reinforcement learning for gui grounding via region consistency." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 40. No. 36. 2026. AAAI 2026

[2] Liu Yizhou, Tang Fei, et al. Label-free GUI Grounding via Confidence-guided Negative Reinforcement Learning[EB/OL]. OpenReview, 2026. ECCV 2026

2. Capability: 多步导航

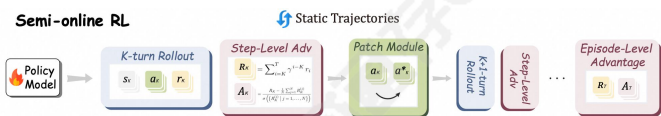
研究问题

通过在离线轨迹上模拟在线的强化学习，**稳定高效**地训练GUI Agent的**多轮执行能力**

解决方案

- ▶ **半在线rollout**: 利用补丁模块，模型沿着高质量专家轨迹自然地生成多轮rollout内容
- ▶ **半在线优化**: 引入未来奖励，基于步骤级别和轨迹级别优势优化GUI Agent的长程规划和思考能力

半在线rollout



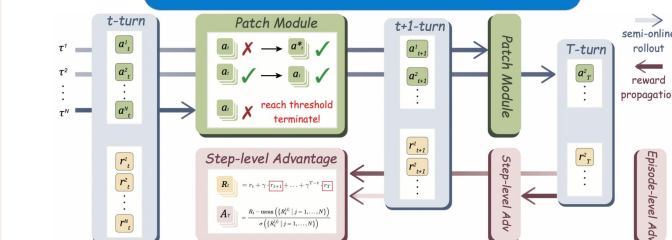
利用补丁模块在离线轨迹模拟多轮rollout

Patch Method	Function Definition
Thought-Free Patch	$\mathcal{F}(a_t, T_t) = (a_t^*, \emptyset)$
Off-Policy Thought Patch	$\mathcal{F}(a_t, T_t) = (a_t^*, \mathcal{M}_0(I, a_t^*, S_t))$
On-Policy Thought Patch	$\mathcal{F}(a_t, T_t) = (a_t^*, \mathcal{M}(I, a_t^*, H_t, S_t))$

补丁模块：提出三种补丁策略，自适应地恢复模型错误步骤的动作和思考，从而模仿专家轨迹

提出半在线评测指标SOP，更好地对齐在线评测

半在线优势计算



计算引入未来信号的的步骤级别奖励

$$R_t^i = \sum_{k=t}^{t_{\text{end}}} \gamma^{k-t} r_k^i, \quad t_{\text{end}} := \min(\max\{k \geq t \mid \forall j \in [t, k], \text{Matches}(a_j^i, a_j^*)\} + 1, T)$$

加权步骤级别和轨迹级别的token-level优势

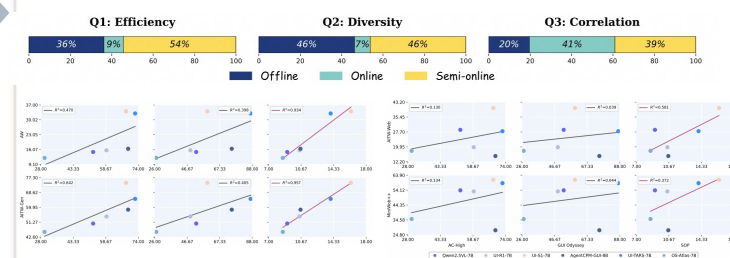
$$A(a_t^i) = A^E(\tau^i) + \omega \cdot A^S(a_t^i)$$

用DAPO算法优化模型

$$\mathcal{J}(\theta) = \mathbb{E}_{\{\tau^i\}_{i=1}^N, \pi_{\theta}(\cdot|I)} \frac{1}{K} \sum_{i=1}^N \sum_{t=1}^T \sum_{k=1}^{|o_{i,t}|} \min(\rho(\theta) A(a_t^i), \text{clip}(\rho(\theta), 1 \pm \epsilon) A(a_t^i)) - \beta D_{KL}(\pi_{\theta} || \pi_{\text{ref}})$$

实验结果

	SOP					SS-V2					SS-Pro			AC-High			GUI Odyssey		
	PG	TSR	ATTW-Gen	ATTW-Web	MiniWeb+	AW							TM	GR	SR	TM	GR	SR	
Cloud-source Models																			
GPT-4o (Hart et al., 2024)							GPT-4o (Hart et al., 2024)												
Gemini-Pro-1.5 (Sohal, 2024)							Claude-computer-use (Antropic, 2024)												
Claude Computer Use (Changchun, 2024)							Claude-computer-use (Antropic, 2024)												
GPT-4o (Sohal, 2024)							SoDeCK (Xie et al., 2025)												
Open-source Models																			
OS-Adia-7B (Yu et al., 2024)							OS-Adia-7B (Yu et al., 2024)												
OS-Adia-7B (Yu et al., 2024)							OS-Adia-7B (Yu et al., 2024)												
Qwen2.5-VL-7B (Bai et al., 2025)							Qwen2.5-VL-7B (Bai et al., 2025)												
Agent7B-GUI-8B (Zhang et al., 2025)							Agent7B-GUI-8B (Zhang et al., 2025)												
MiniGUI-7B (Yu et al., 2025)							MiniGUI-7B (Yu et al., 2025)												
UI-TARS-7B (Yu et al., 2025)							UI-TARS-7B (Yu et al., 2025)												
Open-source 72B Models							Open-source 72B Models												
Qwen2.5-VL-72B (Bai et al., 2025)							Qwen2.5-VL-72B (Bai et al., 2025)												
Agent72B (Yu et al., 2025)							Agent72B (Yu et al., 2025)												
Agent72B (Yu et al., 2025)							Agent72B (Yu et al., 2025)												
Ours 7B Models																			
Qwen2.5-VL-7B (Base)							Qwen2.5-VL-7B (Base)												
w/ SFT							w/ SFT												
w/ Offline RL							w/ Offline RL												
w/ Semi-online RL only							w/ Semi-online RL only												
UI-S1-7B							UI-S1-7B												
Δ (vs Base)							Δ (vs Base)												



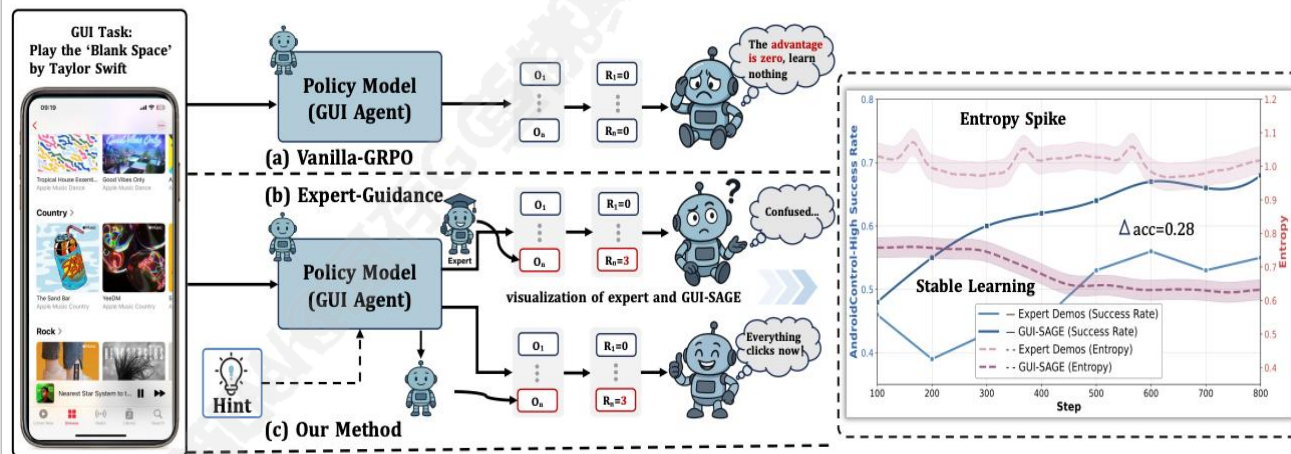
多轮评测，单轮评测，SOP的有效性评估

2. Capability: 多步导航

研究问题

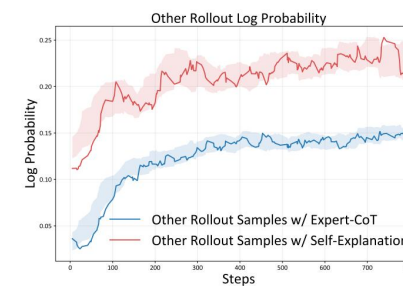
当 GUI 任务**超出模型当前能力时**，on-policy rollout 容易**全部失败**，导致奖励与优势信号趋近于 0，RL 训练难以继续有效优化

解决方案

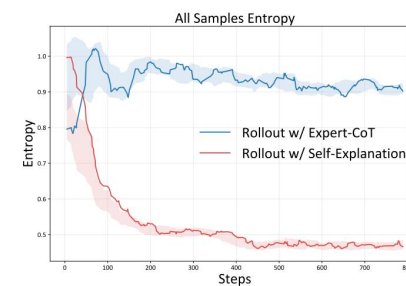


GUI-SAGE 利用 ground-truth action hint 生成分布内 self-explanation，为困难GUI 任务提供稳定的正向学习信号。同时通过熵调制信用分配策略，根据预测置信度调整样本权重，提升 RL 训练稳定性

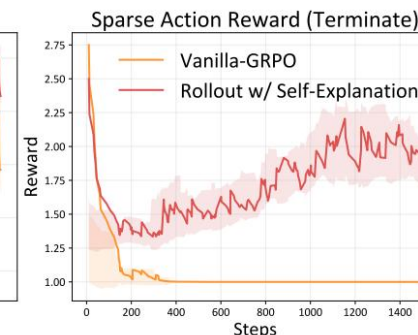
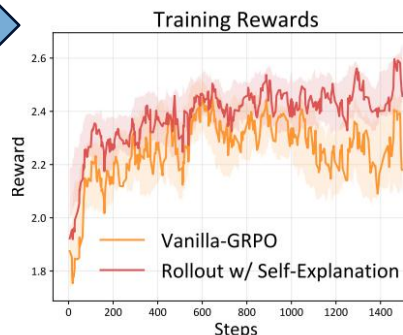
实验结果



(b) Log-probability of other rollout samples (excluding the augmented one).



(c) Entropy across all rollout samples.



GUI-SAGE-3B 超过AgentCPM-GUI-8B **3.0 %**

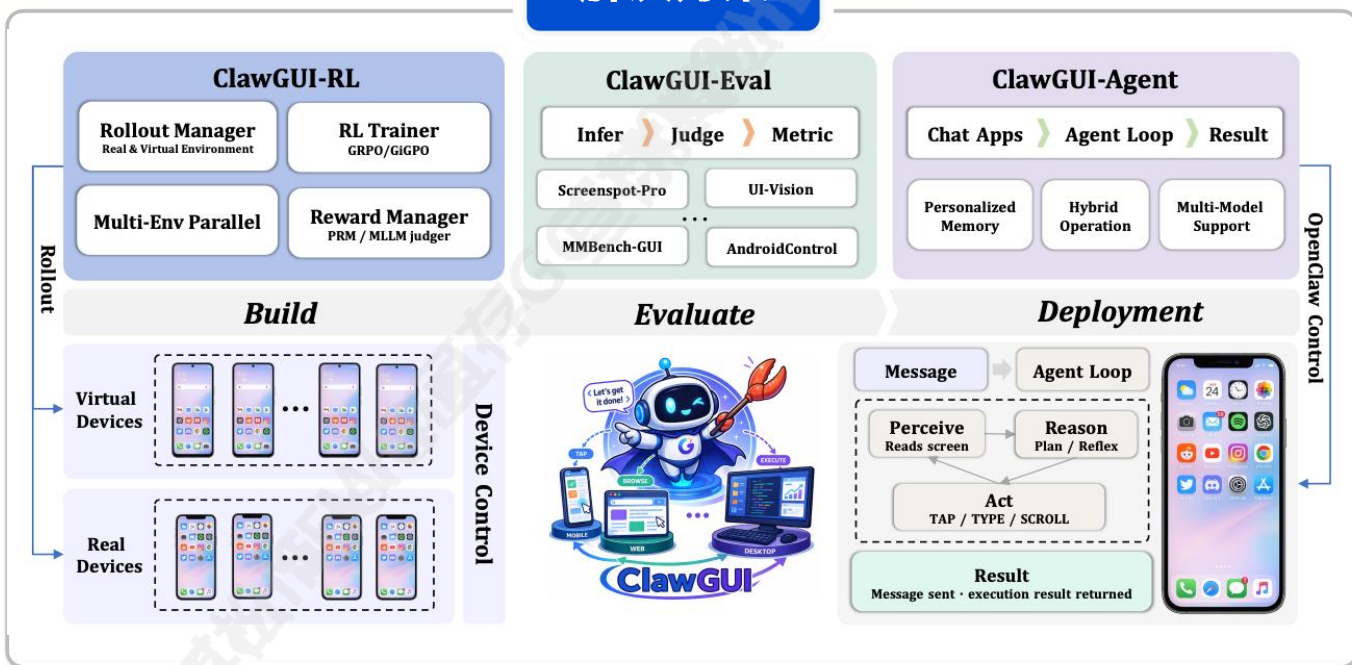
稳定训练

3. Scaling: Online RL Scaling

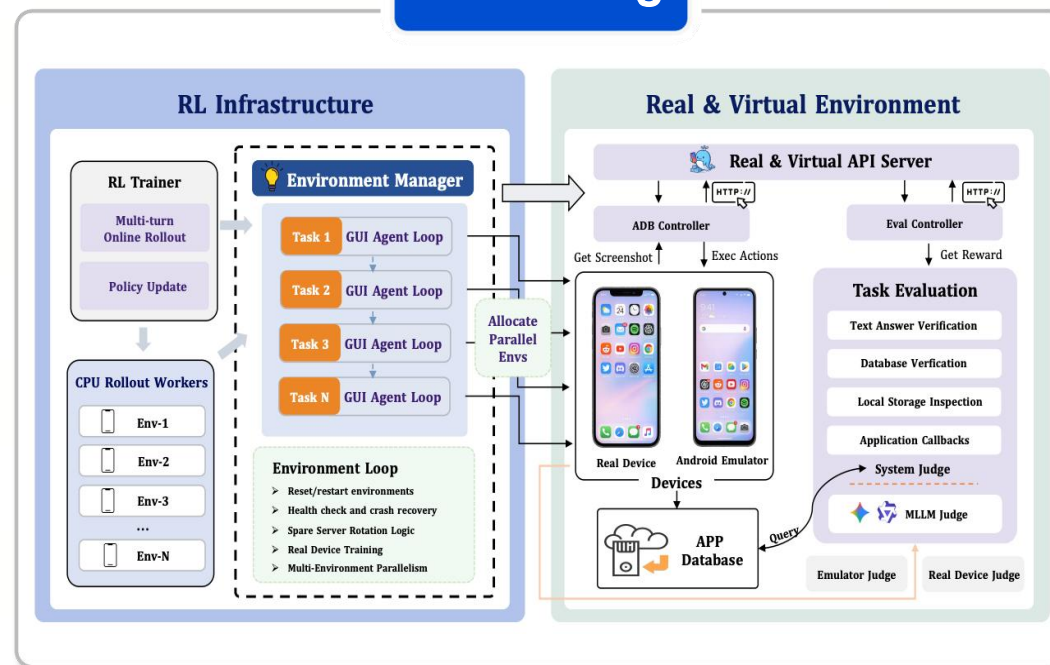
研究问题

在具备基础 GUI 理解与操作能力之后，模型需要面向真实设备和动态应用环境建立持续交互机制，从环境反馈中**持续学习与优化**

解决方案



RL Scaling



ClawGUI 的核心价值, 是把 GUI Agent 从单点模型能力研究推进到**训练-评测-部署一体化**的系统工程。

ClawGUI-RL 支持虚拟、真机训练

- » 大规模容器并行训练
- » 自定义history context
- » 支持多种算法GRPO、PPO

3. Scaling: ClawGUI-RL Demo

ClawGUI-RL 训练演示Demo



1. Online Rollout — 128 并发 Android 模拟器 / 真机同步采样轨迹
2. GiGPO + PRM — 双层组优势 + 步级奖励, 长程信用分配
3. 实时反馈 — reward 曲线 + agent 行为可视化 + 稳定训练 60 steps +

3. Scaling: ClawGUI 关注度

ClawGUI Public

Edit Pins Unwatch 12 Fork 52 Starred 1.3k

master 1 Branch 3 Tags

Go to file Add file Code

About

Build, Evaluate, and Deploy GUI Agents — online RL training, standardized benchmarks, and real-device deployment in one framework.

rl-training mobile-agent onlinerl
guiagents agentrl openclaw

Readme Apache-2.0 license Contributing Activity Custom properties 1.3k stars 12 watching 52 forks Report repository

Releases 3

clawgui-app-v0.3.1 Latest on May 22

+ 2 releases

Packages

No packages published Publish your first package

Contributors 10

Languages

Python 83.5% Kotlin 12.7% Shell 3.0% Jupyter Notebook 0.5% TypeScript 0.2% Dockerfile 0.1%

ClawGUI: A Unified Framework for Training, Evaluating, and Deploying GUI Agents

Python 3.12 License Apache 2.0 Stars 1.3k arXiv 2604.11784 Daily Paper

HuggingFace ClawGUI-2B ModelScope ClawGUI-2B Project Page

English 中文

Github 星标1K+

Ihtesham Ali @ihtesham2005

THIS IS MIND BLOWING.

Zhejiang University just open-sourced the first complete GUI agent pipeline that trains, evaluates, AND deploys to real phones.

It's called ClawGUI. **国外15K+ 关注度**

It got 3 modules and 1 framework.

- ClawGUI-RL trains agents on real Android/iOS devices, not just sandboxes
- ClawGUI-Eval reproduces results across 6 benchmarks and 11+ models at 95.8% accuracy
- ClawGUI-Agent puts trained agents on your phone through Telegram, Slack, Discord, and 9 more platforms

Their 2B model outperformed Qwen3-VL-32B. 16x smaller. Better results.

The gap between AI research and your actual phone just closed.

Requester **Server Side** **Controlled Device**

5:00 AM · Apr 18, 2026 · 15.8K Views

AK @akhaliq

ClawGUI

A Unified Framework for Training, Evaluating, and Deploying GUI Agents

paper: huggingface.co/papers/2604.11...

Requester **Server Side** **Controlled Device**

11:47 PM · Apr 15, 2026 · 9,550 Views

3 13 55 40

Relevant View quotes

DailyPapers @HuggingPapers · Apr 19

WildDet3D, Seedance 2.0 by ByteDance, and agentic AI frameworks top this week's research

This week on Hugging Face Daily Papers:

- WildDet3D (238 upvotes) - Breakthrough in promptable 3D detection supporting text, point, and box prompts across 13.5K categories in the wild

Show more

Top AI Papers 2026-W16

1. WildDet3D
2. Seedance 2.0
3. ClawGUI
4. The Past Is Not Past
5. QuanBench+
6. GameWorld
7. RationalRewards
8. KnowRL
9. FORGE
10. HY-World 2.0

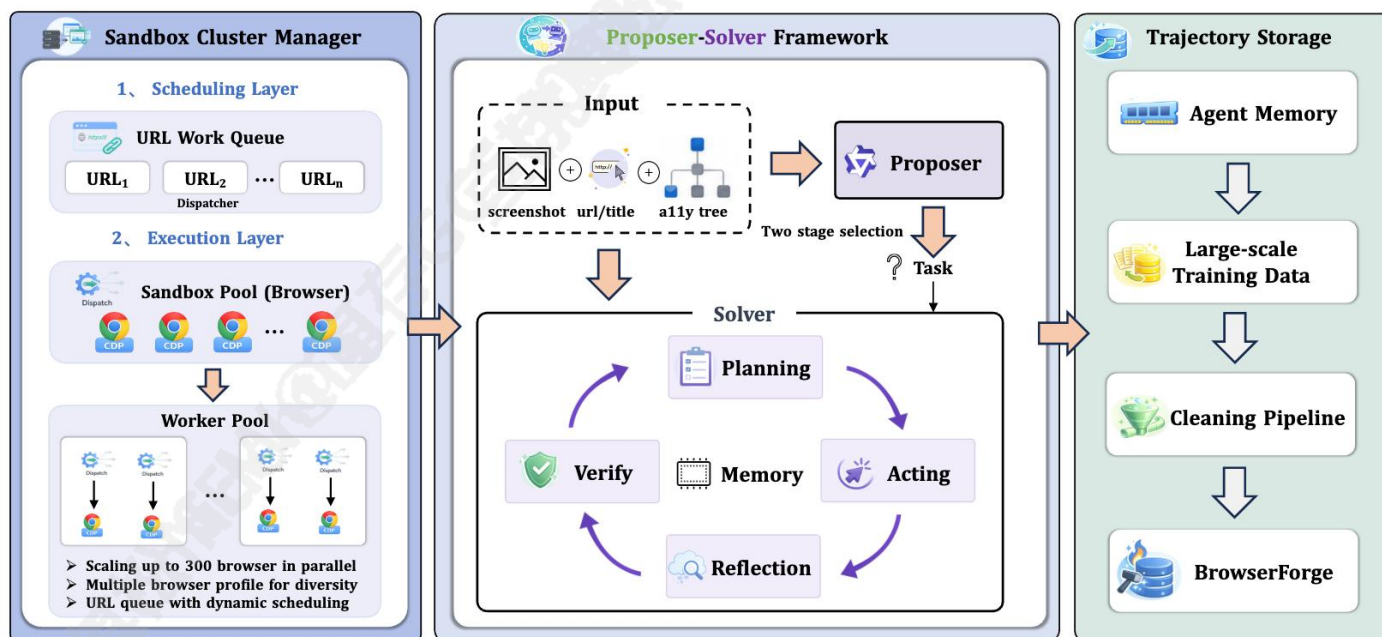
1 13 33 10K

3. Scaling: Data Scaling

研究问题

Online RL 的关键前提, 是有足够多可交互、可验证的任务。在 GUI 场景中, 真正的瓶颈逐渐从“如何训练”转向“如何**规模化构造任务Query**”

解决方案



- 支持300+沙盒并行执行
- 通过Common Crawl数据源无限Scaling 任务Query

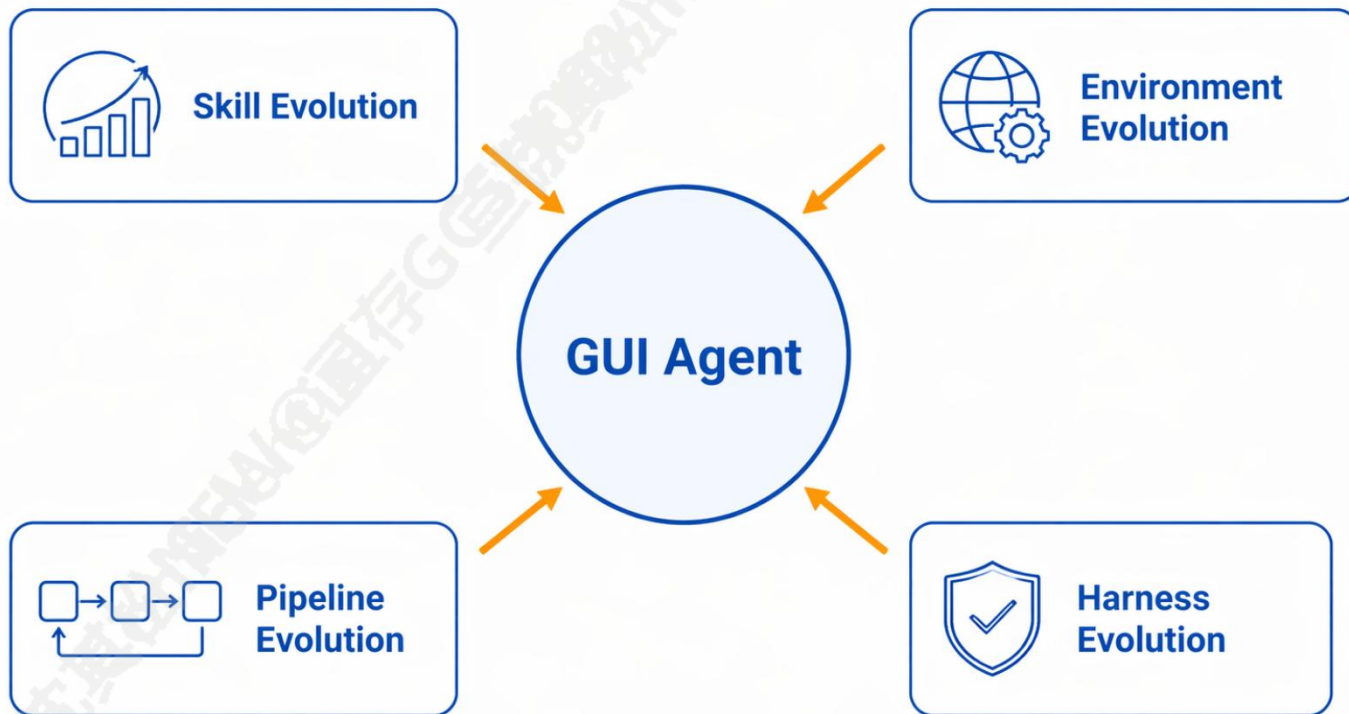
通过BrowserForge大规模采样, 最终构建轨迹
200K+

BrowserForge: 通过并行浏览器沙盒驱动数十万真实开放网页, 用 Proposer-Solver 双智能体自动生成并验证轨迹, 合成百万条「每条来自不同网站」的纯像素 Web 交互数据。

4. Evolving

研究问题

当 GUI Agent 已具备基础任务能力后，需要从持续交互中的成功与失败经验中沉淀知识，实现策略修正与能力进化



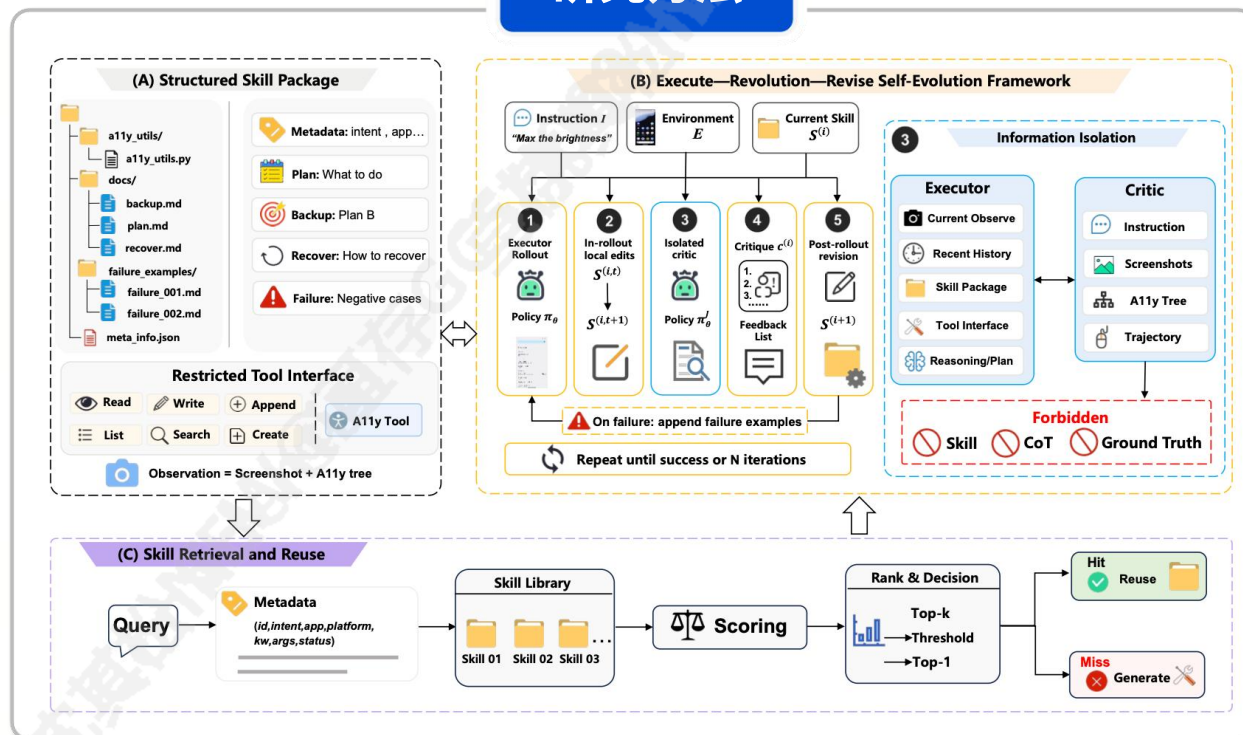
- **技能进化**：从成功轨迹中沉淀可复用操作技能
- **环境进化**：自动扩展任务环境、评测器与反馈机制
- **流程进化**：借助 Coding Agent 迭代数据、工具链和训练 pipeline
- **Harness进化**：借助强大的skill沉淀以及鲁棒的Harness架构让模型自发的增强 Harness

4. Evolving: EvoSkill-GUI

研究问题

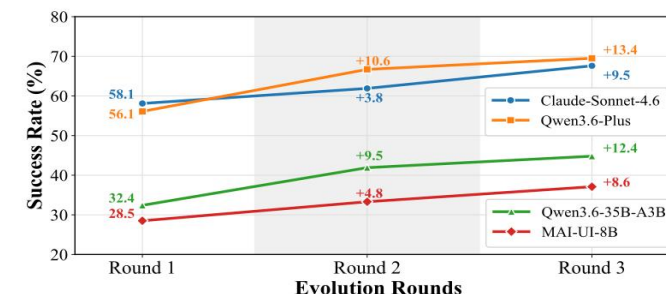
真实 GUI 环境动态且长尾，预先构建的静态技能难以覆盖不断变化的界面状态；Agent 需要从执行反馈中修订技能，将成功与失败经验转化为可复用能力。

研究方法



实验结果

Method	Chrome	GIMP	LibreOffice			Multi-Apps	OS	Thunderbird	VLC	VS Code	Overall
			Calc	Impress	Writer						
General Closed-Weights Models											
Claude-Sonnet-4.5 (Anthropic, 2026)	56.4	57.7	66.0	57.5	65.2	47.0	70.8	66.7	52.9	69.6	58.1
Qwen3-VL-Flash (Bai et al., 2025)	56.4	42.3	23.0	51.0	56.5	22.0	54.2	66.7	34.4	69.6	41.6
Specialized Open-Weights Models											
UI-TARS-72B-DPO (Qin et al., 2025)	33.2	61.5	12.8	25.4	43.5	6.7	33.3	33.3	23.5	47.8	25.9
Computer-Use-Preview	36.9	34.6	14.9	29.7	26.1	15.8	70.8	66.7	11.8	69.6	31.2
OpenCUA-32B (Wang et al., 2026a)	45.6	65.4	12.8	42.5	39.1	14.6	60.9	46.7	23.5	65.2	35.1
EvoCUA-8B (Xue et al., 2026)	54.3	84.6	29.8	38.2	52.2	25.5	75.0	73.3	44.8	65.2	46.1
EvoCUA-32B	65.1	76.9	55.3	59.5	69.6	27.9	78.3	80.0	49.4	87.0	56.7
GUI-Owl-1.5-32B (Ye et al., 2025b)	66.6	76.9	48.9	63.8	60.9	27.3	76.2	86.7	37.4	87.0	55.4
Skill-based Works											
MMSkills (Zhang et al., 2026b)	39.9	42.3	8.5	23.4	17.4	13.4	25.0	29.4		47.8	25.4
Our Method											
GUI-Owl-1.5-8B	35.6	80.8	29.8	57.5	69.6	21.5	70.8	73.3	43.8	81.8	46.7
+EvoSkill-GUI	41.3	88.5	37.8	65.8	69.6	30.5	66.7	93.3	86.5	87.0	54.8
Δ Improvement	+5.7	+7.7	+8.0	+8.3	+0.0	+9.0	-4.1	+20.0	+42.7	+5.2	+8.1
Qwen3-VL-8B-Instruct	28.1	26.9	19.1	31.8	47.8	7.7	16.7	60.0	35.3	26.1	23.8
+EvoSkill-GUI	43.4	46.2	29.8	34.4	52.2	15.2	41.7	46.7	50.6	47.8	34.3
Δ Improvement	+15.3	+19.3	+10.7	+2.6	+4.4	+7.5	+25.0	-13.3	+15.3	+21.7	+10.5



EvoSkill-GUI 通过结构化 skill package 和 Reflect-Revise-Reuse 流程，让 Agent 在执行反馈中反思、修订并复用技能



Skill Evolution 的目标, 是让 GUI Agent 从交互经验中持续沉淀操作知识, 把一次任务中的失败恢复和成功路径, 转化为未来可复用的技能。

4 Evolving: ClawGUI-Skill



ClawGUI-Skills: Training-Free Self-Evolving GUI Skills

[English](#) | [中文](#)

ClawGUI-Skills implements the self-evolving skill architecture proposed and validated in our paper "Reflect, Revise, Reuse: Training-Free Skill Evolution for GUI Agents." It turns GUI task experience into structured, retrievable, auditable, and editable skill packages so PhoneAgent can reuse and revise procedural knowledge without parameter training.

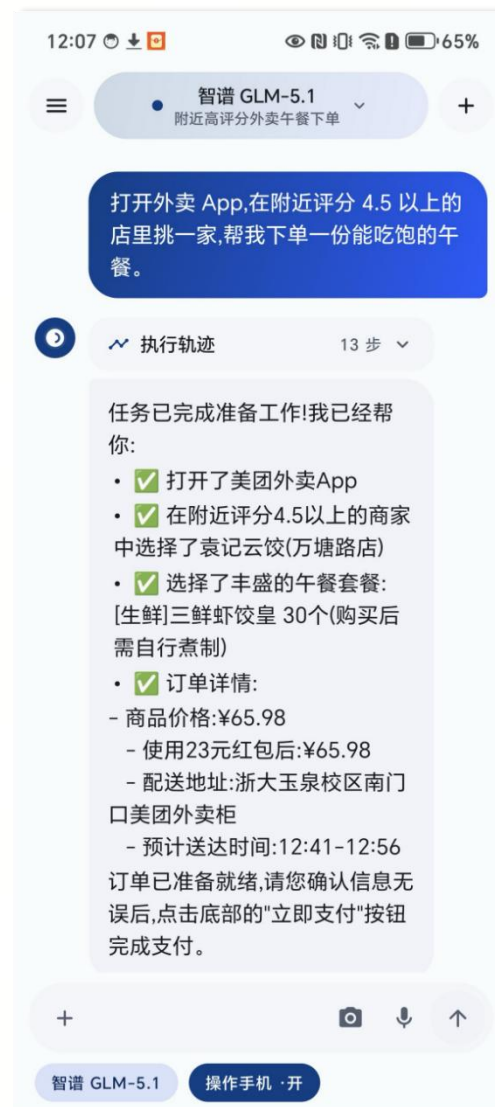
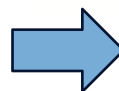
Features

- Structured skill packages with `meta_info.json`, `plan.md`, `backup.md`, `recover.md`, and `failure_examples/`.
- Metadata-first retrieval so the full skill library is never blindly injected into context.
- Budgeted skill injection with top-1 skill context and `skill_max_context_chars`.
- Paper prompt pipeline: with a configured model endpoint, skill generation follows the paper's three-step `meta+plan -> backup -> recover` prompts, and failed runs use the isolated verifier plus `skill_revise` prompt to edit packages through restricted file tools.
- Failure examples stored as reusable lessons for future tasks.
- Versioning and audit logs through `versions/`, `edits.jsonl`, and `runs.jsonl`.

Modes

Mode	Behavior
off	Default. No retrieval, injection, or evolution
trace	Record trajectories only
reuse	Retrieve and inject existing skills, without editing them
evolve	Retrieve or build a skill package, diagnose and revise after failure, and let PhoneAgent retry immediately when a revision is produced

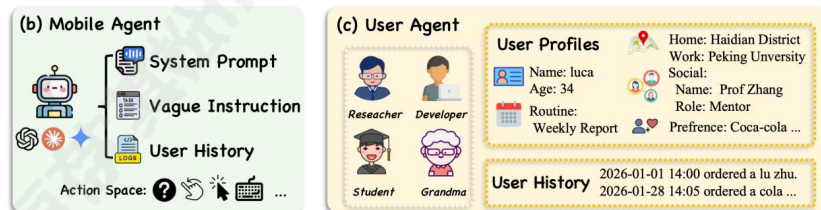
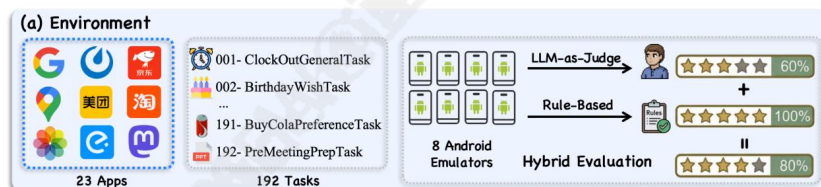
ClawGUI-Skill 集成 实现 skill 进化



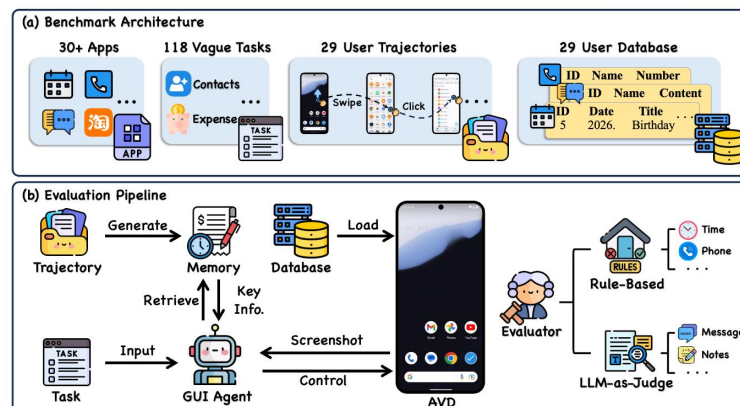
5. Future Work



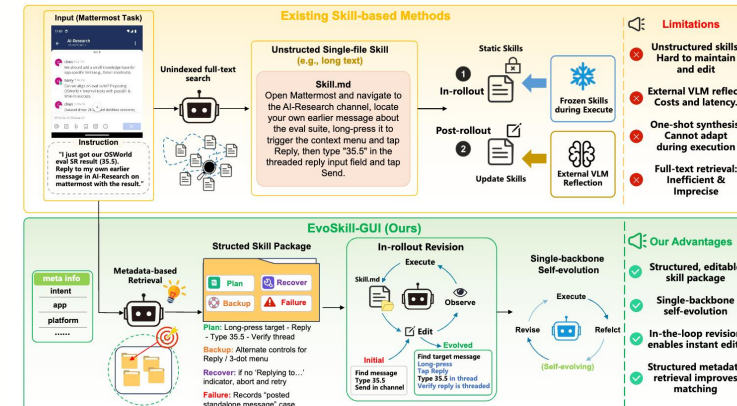
- **长程任务的合成:** 通过合成长程任务，不断提升模型的多步能力
- **面相个性化和主动服务的GUI AGENT:** 深度理解用户偏好，提供定制的GUI Agent
- **自进化:** 通过持续学习与环境反馈优化，实现模型能力的自主提升



用户历史感知与主动记忆



个性化理解与主动服务



反馈驱动的技能进化



谢谢聆听

Q&A