

《自动化测试 2024》GUI 测试方向大作业

方向助教：虞圣呈 yusc@smail.nju.edu.cn

凌毓辰 yuchenling@smail.nju.edu.cn

选题一：大模型引导的场景感知 GUI 探索测试

传统 GUI 自动化探索测试生成大量测试结果，并从中发现待测应用中的缺陷。然而，已有的探索策略缺乏对于真实业务逻辑的理解，从而难以生成复杂交互操作序列以全面对待测应用的功能点进行覆盖。在自动化探索中引入对测试场景的感知能够从用户视角出发，做到更有效的测试覆盖。因此，本项目拟设计一种基于大语言模型引导的场景感知 GUI 自动化探索测试技术，通过大模型理解测试场景、生成并执行测试操作，将基于场景领域知识融入自动化 GUI 测试，实现更加可靠且可解释的 GUI 测试生成。

a) 测试场景的感知与理解

- i. 通过大语言模型理解测试场景的目标
- ii. 通过大语言模型实时分析理解待测应用的状态
- iii. 通过大语言模型理解应用状态与目标场景的联系

b) 基于场景的测试操作生成

- i. 通过大语言模型生成符合场景要求的 GUI 测试操作（基于对测试场景的感知和理解）
- ii. 结合已有的自动化测试工具/接口，如 ADB、Appium 等，执行生成的测试操作

c) 基于场景的应用状态验证

- i. 通过大语言模型检查生成的测试操作是否被正确执行
- ii. 通过大语言模型检查待测应用状态是否符合预期（符合场景要求）

d) 完整的自动化 GUI 测试探索 workflow

- i. 通过大语言模型对偏离目标场景要求的应用状态进行纠正
- ii. 通过大语言模型对未能正确执行的操作进行修改
- iii. 通过循环将场景感知、操作生成与执行、状态验证与修正串联，形成完整的自动化 GUI 测试探索 workflow

e) 交互界面设计与实现

- i. 本工具要求基于 B/S 架构实现，实现测试配置、测试启动与终止、以及基本的自动化 GUI 测试过程的可视化即可。

f) 领域知识引导的 GUI 场景感知测试生成（额外附加分）

- i. 有效将人工开发测试脚本、人工编写测试脚本等测试产物中领域知识进行提取，通过 RAG 方式增强大模型的场景感知理解能力

工具实现评分细则

工具功能：60%（每项关键需求占 15%，每项重要需求占 10%，每项普通需求占 5%）			
测试场景感知与理解	关键需求		
基于场景的测试操作生成	关键需求		
基于场景的应用状态验证	重要需求		
完整的自动化 GUI 测试探索 workflow	关键需求		
交互界面设计与实现	普通需求		
实现程度	关键需求	重要需求	普通需求
采用合理、较为先进算法，或集成较新研究成果，运行效果好	13 – 15	8 – 10	—
采用简单算法完成该需求，运行效果良好	10 – 12	5 – 7	3 – 5
该需求可完整运行，但无法达到预期效果	6 – 9	3 – 4	—
无法连续运行该需求，仅简单实现部分该需求对应算法片段	1 – 5	1 – 2	1 – 2
未完成该需求	0	0	0
项目文档：20%			
详细描述功能模块及交互、运行要求，可根据文档运行工具		16 – 20	
基本描述功能描述功能模块及交互、运行要求		10 – 16	
描述部分功能及交互		4 – 9	
文档内容匮乏		0 – 3	
项目视频：10%			
视频配解说、字幕，能够较好呈现运行过程、关键技术挑战		7 – 10	
视频包含讲解，完整呈现工具运行过程		4 – 6	
基本呈现工具运行过程		1 – 3	
无视频		0	
代码规范：10%			
注释完整，项目结构清晰		7 – 10	
关键部分注释完整，项目结构较为清晰		4 – 6	
包含若干注释		1 – 3	
无注释，结构混乱		0	

选题二：基于机械臂视觉识别的非侵入式 GUI 界面自动化探索（不超过 3 组）

传统 GUI 自动化探索测试往往使用侵入式测试框架完成，例如 Appium。这些框架通过侵入移动应用安装文件来捕获 GUI 元素的标识符，从而完成测试任务。然而，当前的方法存在诸多问题。首先，由于这些框架需要侵入移动应用安装文件，可能会带来用户隐私问题；其次，这些框架依赖于移动操作系统平台底层特性，因此无法在不同系统之间通用（如安卓、iOS、鸿蒙等），导致测试脚本重复开发的工作量大；第三，由于这些方法直接从应用安装文件捕获 GUI 元素，无法真正从用户视角完成测试，有些问题难以被发现，如异形屏 GUI 元素遮挡问题等。因此，本项目希望利用机械臂模拟真正用户对 app 的使用，利用计算机视觉技术，完成待测应用及 GUI 元素的识别与分析，进一步结合已有的 GUI 界面自动化探索策略，从而实现基于机械臂视觉识别的非侵入式 GUI 界面自动化探索。（选择本题的同学请提前发邮件至助教确认名额）

a) 目标识别

- i. 通过计算机视觉技术，识别机械臂摄像头拍摄照片中设备 GUI 界面；
- ii. 通过计算机视觉技术，识别 GUI 界面中的控件元素；

b) 机械臂操作

- i. 实现上下左右前后六个方向的基本移动操作；
- ii. 实现点击、长按、滑动、双击、等单点触控操作；
- iii. 实现机械臂摄像头的拍摄与照片传输；

c) GUI 界面探索策略

- i. 集成已有 GUI 界面探索策略，如 Monkey，Stoat 等
- ii. 可在现有算法基础上根据本场景进行优化；
- iii. 实现机械臂命令转化，以及与服务器的命令传输；

d) 机械臂调度

- i. 基于已有 GUI 界面探索策略，设计新的机械臂移动调度策略
- ii. 将机械臂的移动距离作为影响测试效率的因素对 GUI 界面探索策略进行优化（即移动总路径较低的情况下，实现更高的覆盖率）

e) 交互界面设计与实现

- i. 本工具要求基于 B/S 架构实现，要求实现 web 端机械臂状态监控及命令传输等最基本功能即可。

工具实现评分细则

工具功能：60%（每项关键需求占 15%，每项重要需求占 10%，每项普通需求占 5%）			
目标识别	关键需求		
机械臂操作	关键需求		
GUI 界面探索策略	重要需求		
机械臂调度	关键需求		
交互界面设计与实现	普通需求		
实现程度	关键需求	重要需求	普通需求
采用合理、较为先进算法，或集成较新研究成果，运行效果好	13 – 15	8 – 10	—
采用简单算法完成该需求，运行效果良好	10 – 12	5 – 7	3 – 5
该需求可完整运行，但无法达到预期效果	6 – 9	3 – 4	—
无法连续运行该需求，仅简单实现部分该需求对应算法片段	1 – 5	1 – 2	1 – 2
未完成该需求	0	0	0
项目文档：20%			
详细描述功能模块及交互、运行要求，可根据文档运行工具		16 – 20	
基本描述功能描述功能模块及交互、运行要求		10 – 16	
描述部分功能及交互		4 – 9	
文档内容匮乏		0 – 3	
项目视频：10%			
视频配解说、字幕，能够较好呈现运行过程、关键技术挑战		7 – 10	
视频包含讲解，完整呈现工具运行过程		4 – 6	
基本呈现工具运行过程		1 – 3	
无视频		0	
代码规范：10%			
注释完整，项目结构清晰		7 – 10	
关键部分注释完整，项目结构较为清晰		4 – 6	
包含若干注释		1 – 3	
无注释，结构混乱		0	