
游戏画面理解与结构化标注系统需求文档

(基于 DiMo-GUI方向论文<https://arxiv.org/pdf/2507.00008>的一点想法)

一、项目背景与痛点

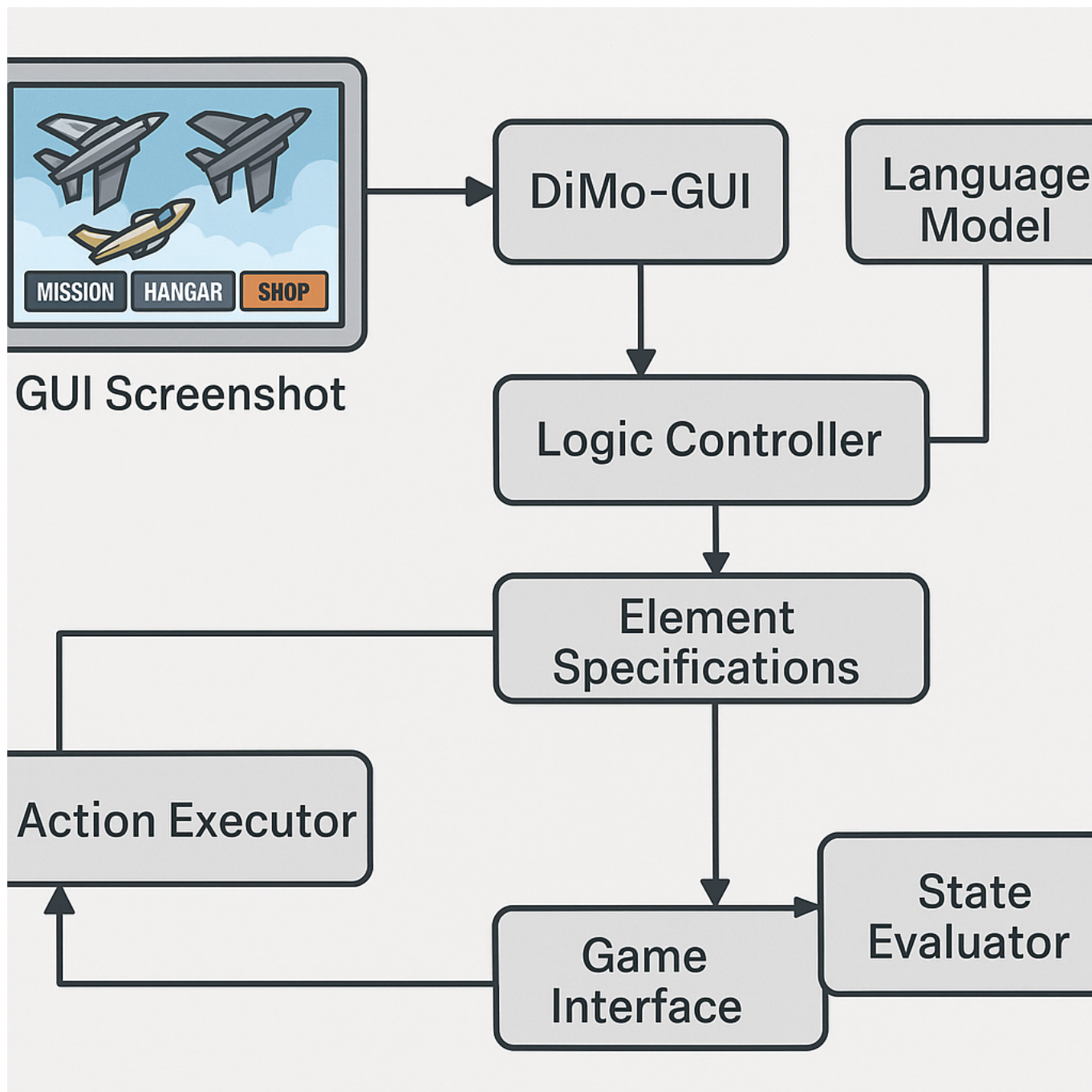
当前针对游戏进行分析、理解与标注时，面临如下核心问题：

- 1. **游戏内部信息缺失严重：**
 - 无法直接获取游戏内具体付费内容信息（比如花了18块钱到底干嘛了？），难以还原游戏消费设计。对用户-游戏的付费理解受限。
 - 缺乏对游戏“节奏节点”的刻画，如关卡起承转合、奖励频率、付费诱导位置。
- 2. **无法从画面进程中构建完整理解链路：**
 - 游戏进展存在非线性、多场景切换，难以构建统一理解。
 - 依赖人工标注或试玩，效率低、主观性强。
- 3. **小游戏数量大但标注价值低下：**
 - 小游戏画面单一，重复性强，但人工标注成本高。
 - 标注缺乏统一性与结构化标准，难以形成知识库。

二、系统目标与价值定位

能力点	目标描述
游戏关键付费内容点提取	自动识别游戏内的付费点、礼包、货币、商城入口及其界面表现（位置、文案、诱导方式）
游戏内容节奏理解构建	基于连续截图流，识别游戏节奏结构，如剧情过渡、任务流、奖励频次、重复周期等
游戏程序化打标（适配小游戏）	针对小游戏自动化识别玩法、UI结构、交互机制，并输出结构化标签（图标/文案/进程）

三、系统功能设计



1. 游戏截图识别与切片流分析

- **输入：**游戏连续截图序列（含点击链路）
- **输出：**按逻辑/节奏节点分段的画面流结构

2. DiMo-GUI视觉识别与目标聚焦模块

- 局部聚焦关键模块（如“商店”、“购买”、“礼包”）
- 输出每个页面中的关键要素bounding box + 类型

3. 游戏语义理解与进程建模模块

- 抽取按钮文本、奖励提示、操作指引
- 构建“游戏进展链”：从登录到任务完成的路径流

4. 付费点提取与诱导分析模块

```
{
  "page": "商城-首充礼包",
  "pay_type": "首充礼包",
  "position": [430, 230, 620, 300],
  "cta_text": "立即购买",
  "诱导标签": ["限时", "超值", "一次性"]
}
```

5. 小游戏识别与打标模块

- 自动提取 UI 模板（左右控制、弹射、计分器等）
- 快速匹配玩法结构并输出标签

四、数据输出格式标准（随便举个栗子）

```
{
  "game_name": "XXX",
  "screenshots": [
    {
      "page_type": "任务完成页",
      "element_tags": [
        {"type": "按钮", "text": "领取奖励", "bbox": [200, 800, 500, 900]},
        {"type": "道具图标", "label": "钻石", "bbox": [300, 700, 350, 750]}
      ],
      "progress_phase": "奖励引导阶段",
      "pay_incentive_detected": true
    }
  ]
}
```

五、预期产出价值

应用维度	收益描述
游戏付费相关算法模型特征	快速还原游戏内付费结构、节奏结构，构建模型特征 应用于ltv，礼券等场景
游戏打标	为模型训练/推荐系统提供结构化UI语义、进程、玩法标签

六、技术支撑（待讨论 我就是瞎写）

- 图像识别模型：DiMo-GUI结构 + OCR + 图标分类器
- 语言模型能力：GPT-4, Claude 3.5
- 前端工具：基于Canvas或Streamlit展示标注与流程图

功能拓展与延伸方向畅想

一、集成方向

1. AI辅助评测系统

- 生成游戏评测摘要图文报告，用于平台或发行商评审（可能只能服务于微小cp）

2. 自动生成游戏路径/流程图

- 可视化展示游戏从首页到关键交互/消费节点的完整路径

3. 通用化“游戏知识图谱”

- 构建统一的玩法与付费路径图谱，服务于聚类/推荐/检索

二、自动化生成方向

4. 自动生成“新手引导页面”（在当前尝试版本基础上融合）

- 自动识别新手期关键步骤并生成引导语与结构化内容

5. 自动生成游戏知识问答题库

- 结合结构与上下文，生成游戏流程的知识性测试问答

6. 设计生产辅助工具

- 针对大量游戏内设计，支持任务型标签生产：如“首充入口”、“红色按钮”、“付费诱导文案”等