

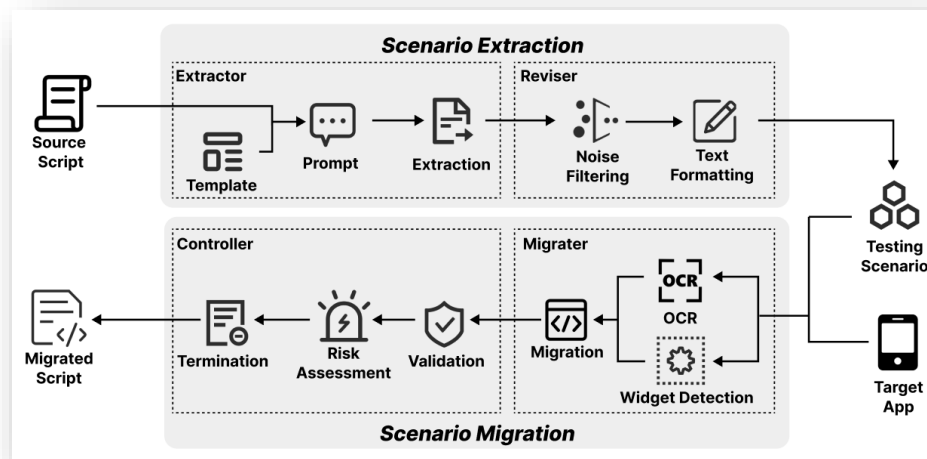
GUI Test Migration via LLM with Scenario-Granularity Understanding

Quan ZHOU, Shengcheng YU, Chunrong FANG,
Junyang XING, Hongyi YANG, Jia LIU, Wenbin
GUO, Zhenyu CHEN

Frontiers of Computer Science, DOI: [10.1007/s11704-025-50270-x](https://doi.org/10.1007/s11704-025-50270-x)

Problems & Ideas

- Problems of existing operation-matching or widget-matching approaches:
 - Those matching approaches often fail to handle divergent operations or widgets between apps.
 - They primarily focus on local matching results, and ignore significant testing scenarios, facing challenges in practical migration scenarios.
- Ideas: A novel GUI test migration approach which extracts testing scenarios from scripts and GUI, and organizes scenario-driven migration via LLM.



The framework of the proposed approach, which organizes scenario-driven migration based on the multi-agent mechanism.

Main Contributions

- Contributions:
 - We introduce ScriptRR, a novel framework that performs cross-app test migration at the granularity of testing scenarios for the first time.
 - We propose a method that combines the computer vision technology, vision language models, and a multi-agent mechanism to enable efficient and dynamic exploration of mobile apps.
 - We conduct empirical experiments on popular apps and representative baselines to evaluate our approach.

Experiment results on SR, GSR and OSR. DLLM and Craft are abbreviations of DirectLLMRR and Craftdroid, which are representative baselines. S-GPT and S-Qwen refer to the ScriptRR powered by GPT-4o and Qwen-VL-MAX respectively.

| Category | SR | | | | GSR | | | | OSR | | | |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | DLLM | Craft | S-GPT | S-Qwen | DLLM | Craft | S-GPT | S-Qwen | DLLM | Craft | S-GPT | S-Qwen |
| to-do | 0.400 | 0.140 | 0.500 | 0.560 | 0.460 | 0.140 | 0.520 | 0.480 | 0.400 | 0.140 | 0.420 | 0.440 |
| browser | 0.104 | 0.167 | 0.792 | 0.771 | 0.083 | 0.125 | 0.667 | 0.708 | 0.042 | 0.125 | 0.646 | 0.708 |
| shopping | 0.375 | 0.000 | 0.531 | 0.500 | 0.375 | 0.000 | 0.563 | 0.500 | 0.375 | 0.000 | 0.531 | 0.500 |
| mail | 0.500 | 0.188 | 0.479 | 0.521 | 0.500 | 0.188 | 0.479 | 0.500 | 0.500 | 0.229 | 0.479 | 0.500 |
| news | 0.104 | 0.188 | 0.771 | 0.750 | 0.188 | 0.188 | 0.833 | 0.729 | 0.083 | 0.188 | 0.750 | 0.708 |
| music | 0.208 | 0.167 | 0.521 | 0.563 | 0.104 | 0.063 | 0.542 | 0.542 | 0.021 | 0.063 | 0.458 | 0.542 |
| comm. | 0.375 | 0.125 | 0.750 | 0.750 | 0.125 | 0.000 | 0.375 | 1.000 | 0.125 | 0.000 | 0.375 | 0.750 |
| Total | 0.287 | 0.148 | 0.610 | 0.609 | 0.277 | 0.120 | 0.599 | 0.578 | 0.227 | 0.127 | 0.546 | 0.561 |