

Coding Agents | Agent Evaluation | Context Engineering

lux-liang.top / lux.liang@outlook.com / github.com/lux-liang

I study how coding agents explore repositories and select useful context. My future research interests include user intent and memory in long-running agents.

## EDUCATION

### Xinjiang University

Sep. 2024 - Jun. 2028 (expected)

B.Eng. in Information Security | GPA: **4.16 / 5.00** | Rank: **5 / 78**

Relevant coursework: Algorithms, Operating Systems, Machine Learning, Computer Networks

## SELECTED PAPERS & PREPRINTS

### SWE-Explore: Benchmarking How Coding Agents Explore Repositories

Co-first author | NeurIPS 2026, under review | [Paper](#) / [Code](#)

- **Study:** A repository-exploration benchmark with **848 instances, 203 repositories and 10 languages**, formulated as line-level ranked context selection. Trajectory-grounded supervision and Coverage, Ranking and Context Efficiency metrics are validated through downstream repair.
- **My contribution:** Co-led benchmark construction, evaluation protocol design, large-scale baseline experiments and manuscript preparation.

Featured on Hugging Face Daily Papers

### SWE-Pruner Pro: The Coder LLM Already Knows What to Prune

Co-author | ACL Rolling Review, Aug. 2026 submission | [Paper](#) / [Code](#)

- **Study:** Lightweight in-agent context pruning from relevance signals in coding-LLM hidden representations, without a separate pruning LLM.
- **Selected result:** With Qwen3-Coder-Next on SWE-QA-Pro, total tokens (prompt + completion) decreased **39.4%** versus NoPruning (607K to 368K); mean judge score (1-10) increased from **7.60 to 7.84**.
- **My contribution:** Contributed to experimental evaluation, result analysis and manuscript preparation.

### Temporal-Physiological Gaze Tracking under Edge Resource Constraints

Co-author | AAAI 2027, under review

- **Study:** A temporal-physiological gaze tracker achieved **0.712°** mean error at 128 × 128 / 15 FPS, with **78.38%** fewer full-model executions through adaptive inference.
- **My contribution:** Conducted experimental evaluation and ablation analysis, including Raspberry Pi 5 deployment demonstrating a **38.63%** latency reduction.

RESEARCH EXPERIENCE

|                                                                                                                                                                       |                       |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| <b>AI for Multimodality and Science Lab (AIMS)</b><br>Hong Kong University of Science and Technology (Guangzhou)<br>Research student   Advised by Prof. Jun Xia   Lab | Aug. 2026 - Present   |
| <b>Visual Perception and Security Group (VPSG)</b><br>Xinjiang University<br>Research student   Advised by Prof. Zhiqing Guo   Group                                  | Dec. 2025 - Present   |
| <b>LLM for Software Engineering Lab (LLMSE)</b><br>Shanghai Jiao Tong University<br>Research student   Mentored by Prof. Xiaodong Gu   Lab                            | Jan. 2026 - Jun. 2026 |

SELECTED AWARDS

|                                                                                                |      |
|------------------------------------------------------------------------------------------------|------|
| <b>National College Student Information Security Competition</b><br>National Third Prize, team | 2026 |
| <b>Huawei ICT Competition, China Final - Ascend AI</b><br>Third Prize                          | 2026 |
| <b>Chinese Collegiate Computing Competition</b><br>National Third Prize, team                  | 2026 |
| <b>Global Algorithm Elite Competition</b><br>National Third Prize                              | 2025 |
| <b>ICPC Xinjiang Provincial Contest</b><br>19th place                                          | 2025 |
| <b>National Encouragement Scholarship</b>                                                      | 2025 |

SKILLS

**Research:** Coding agents, agent evaluation, context engineering, efficient agent systems

**Programming:** Python, C++

**Research stack:** PyTorch, Docker, SWE-bench, mini-SWE-agent, Linux, Git