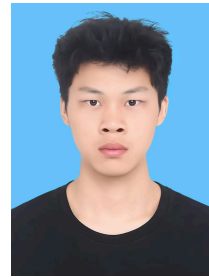


# Sinan Wang

✉ sinanwang@hust.edu.cn    ☎ 150-0270-9274  
🌐 github.com/wsn18381187    🌐 WeChat: Phone Number  
🏠 Personal Homepage    📍 Wuhan



## Education Background

---

### Huazhong University of Science and Technology

Sep. 2023 – Jun. 2027

Undergraduate    Computer Science and Technology    Weighted Ave. 83.2 / 100

### National University of Singapore, School of Computing

Jul. 2024

Summer Workshop    Structure and Interpretation of Computer Programs    Game Developing with Unity Package in Java

## Additional Information

---

**Professional Skills** Proficient in C/C++, Python, and data structures/algorithms, with a solid grasp of Linux/Git workflows and CS fundamentals (Networking/Systems), plus experience in  $\LaTeX$ /Markdown documentation.

**Research Interests** LLM-as-a-Judge, Multimodal, LLM Reasoning, Agent System

**English Proficiency** IELTS Score:6.5, proficient in reading English technical documentation and academic literature.

## Research Projects

---

### Label-Free LLM-as-a-Judge Evaluation Framework and Applications

Aug. 2025 – Mar. 2026

Defined and validated the "Situational Preference" phenomenon in LLM evaluation and proposed a "Label-Free" metric to assess LLM robustness.

- **Co-first Author:** Implemented all project code and contributed to experimental design, paper drafting, and visualization.
- Developed an LLM-as-a-Judge framework based on "Situational Preference" to automatically quantify stability variances across LLMs; demonstrated high correlation with mainstream benchmarks, validating the framework's feasibility for performance optimization.
- Developed automated evaluation scripts in Python using OpenAI SDK and JSON-related libraries.[Project Repository]

### LearnAgent —AI Agent Framework

Apr. 2026 – Present

Self-designed and developed a versatile AI Agent framework supporting seamless integration of various LLMs and external tools.

- Designed and implemented **CodeAct mode**, replacing JSON function calls with Python code generation to eliminate parameter escaping issues and significantly reduce formatting error rates.
- Developed **Condition Flow** (Plan-Execute-Judge-Conclude architecture) to automate task extraction, decomposition, execution, and review, enabling iterative execution of complex tasks.
- Integrated a dual-memory mechanism based on RAG retrieval and dynamic user profiling for long-term/short-term memory; designed multi-tool scheduling modules to build a complete toolchain.
- Developed in Python using OpenAI SDK for model orchestration and ChromaDB for RAG functionality; project open-sourced on GitHub. [Project Repository]

### MultiRef: Controllable Image Generation with Multiple Visual References

Mar. 2025 – Aug. 2025

Developed "RefBlend," a data engine for multi-dimensional decomposition of real and synthetic image datasets, and established the high-quality "MultiRef-dataset" and "MultiRef-benchmark."

- Built a pipeline to construct the "MultiRef-dataset" and "MultiRef-benchmark" based on various signals and instructions from generated target images.
- Curated and open-sourced the dataset and benchmark on Hugging Face, and responsible for subsequent management and maintenance.
- Participated in local model deployment using vLLM and data annotation tasks. [Project Repository]

### **Reinforced Visual Perception with Tools**

Jan. 2025 – May. 2025

A GRPO-based training framework for MLLMs using external tools for visual Chain-of-Thought (CoT) reasoning

- Reproduced benchmark data for visual Chain-of-Thought reasoning from relevant research papers.
- Built a pipeline to construct visual Chain-of-Thought reasoning datasets for training.

## **Papers&Awards**

---

### **MultiRef: Controllable Image Generation with Multiple Visual References**

Aug. 2025

Accepted by ACM MM 2025 [Paper Link]

### **Are We on the Right Way to Assess LLM-as-a-Judge?**

Dec. 2025

Under Review at EMNLP 2026 [Paper Link]

### **Reinforced Visual Perception with Tools**

Sep. 2025

Arxiv Preprint [Paper Link]

### **Interdisciplinary Contest in Modeling**

May. 2025

Honorable Mention