

Mingyu “Myles” Ouyang

yyyangwhu@gmail.com | mingyuouyang.com | (+65) 88851437 | [Google Scholar](#) | [LinkedIn](#)

Research Focus Areas

My research focuses on computer-use and game agents, including agent-environment interaction and the productivity gains these agents enable. I currently focus on game environments with rich visual observations, long-horizon goals, and dynamic mechanics. I aim to improve agents’ perception, reasoning, and control to support their future use in embodied settings.

Education

National University of Singapore

Ph.D. in Computer Science

Singapore

Jan. 2025 – Present

Supervisors: Prof. Hwee Tou Ng and Prof. Mike Zheng Shou

National University of Singapore

M.Sc. in Data Science and Machine Learning

Singapore

Aug. 2023 – Dec. 2024

Wuhan University

B.Sc. in Geospatial Informatics and Digitalized Technology

Wuhan, China

Sep. 2019 – Jul. 2023

Publications (* equal contribution; † corresponding author)

[Under review, 2026.10] Player-1: Learning to Play from Synthetic Experience

Mingyu Ouyang, Yiran Wang, Siyuan Hu, Xingyilang Yin, Huiqiang Sun, Guangzhi Wang, Junfu Pu, Cheng Cheng, Honghao Chen, Ying Shan, Mike Zheng Shou[†], Hwee Tou Ng[†]

- Synthetic environments improve transferable agent capabilities: An agentic post-training framework that learns general game control beyond supervised demonstrations.

[ECCV ’26] GameWorld: Towards Standardized and Verifiable Evaluation of Multimodal Game Agents

Mingyu Ouyang*, Siyuan Hu*, Kevin Qinghong Lin, Hwee Tou Ng[†], Mike Zheng Shou[†]

- Evaluates computer-use and general-purpose game agents with 34 games and 170 tasks.
- Received the **Best Paper Runner-Up Award** at the Multimodal Digital Agents Workshop, ECCV 2026.
- Adopted for multimodal agent evaluation in the [Seed-2.1 model card](#) from ByteDance Seed and [VISTA](#) from MIT.

[CVPR ’26] FocusUI: Efficient UI Grounding via Position-Preserving Visual Token Selection

Mingyu Ouyang, Kevin Qinghong Lin, Mike Zheng Shou[†], Hwee Tou Ng[†]

- Selects instruction-relevant visual tokens for low-cost, efficient UI grounding.
- Received the **Compute Transparency Champion Award** at CVPR 2026.

[AAAI ’26] SlideTailor: Personalized Presentation Slide Generation for Scientific Papers

Wenzheng Zeng*, Mingyu Ouyang*, Langyuan Cui*, Hwee Tou Ng[†]

- Generates preference-aligned presentation slides for scientific papers.

[arXiv ’24] The Dawn of GUI Agent: A Preliminary Case Study with Claude 3.5 Computer Use

Siyuan Hu, Mingyu Ouyang, Difei Gao, Mike Zheng Shou[†]

- An out-of-the-box framework for desktop GUI agents, supporting multiple operating systems and models.

[CVPR ’24] AssistGUI: Task-Oriented PC Graphical User Interface Automation

Difei Gao, Lei Ji, Zechen Bai, Mingyu Ouyang, Peiran Li, Dongxing Mao, Qinchun Wu, Weichen Zhang, Peiyi Wang, Xiangwu Guo, Hengxu Wang, Luowei Zhou, Mike Zheng Shou[†]

- A desktop GUI agent framework for task decomposition, GUI parsing, action generation, and reflection.

[TIP ’24] JPEG Quantized Coefficient Recovery via DCT Domain Spatial-Frequency Transformer

Mingyu Ouyang, Zhenzhong Chen[†]

- Recovers JPEG-quantized DCT coefficients with spatial-frequency modeling across compression quality factors.

Professional Experience

ARC Lab, Tencent IEG

Research Intern (Project Up)

2026 – Present

- Research on multimodal digital agents for game environments and agent-environment interaction.

Uselt AI

Singapore

Co-founder

Mar. 2025 – Oct. 2025

- Led GUI agent development and workflow orchestration, focusing on creating, editing, and deploying personal agents.

SenseTime Research Singapore

Singapore

Research Intern

Dec. 2023 – Oct. 2024

- Developed efficient low-level deep learning models for RAW image processing, including noise synthesis and learned image signal processing for built-in camera and smartphone modules.

Tencent Map, Tencent

Beijing, China

Algorithm Development Intern

Mar. 2022 – Oct. 2022

- Implemented GPS data processing using Hadoop and Spark. Developed and maintained an internal service that automatically executes daily tasks, handling over 10 billion callbacks per hour.

Springturn Education Technology

Zhuhai, China

Co-founder, Software Developer

Nov. 2020 – Sep. 2021

- Co-founded the company and secured angel investment exceeding ¥1 million in the second year of business.

Research Experience

Show Lab, NUS

Aug. 2023 – Present

- Research on multimodal digital agents and agent-environment interaction, with a current focus on game environments.

IIP Lab, Wuhan University

Oct. 2022 – Jul. 2023

- Research focus: Low-level vision tasks, including image super-resolution, denoising, and JPEG artifact removal.

State Key Lab of Info. Engineering in Surveying, Mapping and Remote Sensing

Sep. 2021 – Jun. 2022

- Research focus: 3D scene perception, fusing point clouds and BEV images for road-scene understanding.

Open-Source Projects

[GitHub 2k stars] Computer-Use OOTB: [showlab/computer_use_ootb](#)

- Core developer and maintainer. A desktop GUI agent framework for Windows/macOS that supports API-based and self-hosted models, with local execution.

[GitHub 1.9k stars] ShowUI: [showlab/showui](#)

- Maintainer. A training and inference codebase for a lightweight (2B) VLA model for GUI agents.

Miscellaneous

Wuhan University Football Team

- Served as captain for three years and led the team to win the championship at the 16th Hubei Province Games in 2022.

Technical Skills

- **Agentic Coding:** Proficient with Codex and Claude Code.
- **AI/ML Frameworks:** PyTorch, Transformers, vLLM, VeRL.
- **Computer-Use Agent Infrastructure:** On-device control (Windows, macOS), cloud services (AWS, Azure), agent framework orchestration, and browser-based roll-out infrastructure.