

Yun-Ta Hsieh

Philadelphia, PA | yunta@seas.upenn.edu | github.com/Kiky-88 | [Google Scholar](#) | [Homepage](#)

Education

University of Pennsylvania, M.A.S. in Computer Science Expected graduated in May.2026
University of Michigan, M.S. in Digital and Material Technologies Expected graduated in Dec.2026
Huazhong University of Science and Technology, Bachelor of Architecture Graduated in Jun.2024

Research Interest

Fields: Multimodal Agents, Vision-Language-Action Systems, Efficient MLSys

Publications

Selected Publications

QuantVLA: Scale-Calibrated Post-Training Quantization for Vision-Language-Action Models

Jingxuan Zhang*, *Yunta Hsieh**, Zhongwei Wan, Haokun Lin, Xin Wang, Ziqi Wang, Yingtie Lei, Mi Zhang
CVPR 2026

OS-Omni: A Cross-Platform Benchmark for Generalist Computer-Using Agents

Hui Shen*, *Yunta Hsieh**, Jianing Ma*, Ziyuan Liu*, Qi Han*, Xiuqi Xu*, Yanheng Shang*, Chenqi Yin, Zesen Zhao, Boyuan Zheng, Jisen Li, Xiaoxia Wu, Ruoyan Zhang, et al.
Submitted to NeurIPS 2026

Preprints

MathGen: Revealing the Illusion of Mathematical Competence through Text-to-Image Generation

R Liu, Hui Shen, Peng Zhang, *Yunta Hsieh*, Yuyue Zhang, J Xu, S Chen, et al.
arXiv preprint, 2026

MMSpec: Benchmarking Speculative Decoding for Vision-Language Models

Hui Shen, Xin Wang, Peng Zhang, *Yunta Hsieh*, Qi Han, Zhongwei Wan, Zhe Zhang, Jingxuan Zhang, et al.
arXiv preprint, 2026

SkillEvolBench: Benchmarking the Evolution from Episodic Experience to Procedural Skills

Yingtie Lei, Zhongwei Wan, Jiankun Zhang, Samiul Alam, Zixuan Zhong, Peizhou Huang, Xin Wang, Jingxuan Zhang, Donghao Zhou, *Yunta Hsieh*, Zhihao Dou, Hui Shen, et al.
Submitted to NeurIPS 2026

Speculative Decoding for Multimodal Models: A Survey

Yuyue Zhang, Y Wang, *Yunta Hsieh*, Xin Wang, Peng Zhang, Z Yang, J Ma, Z Zhao, et al.
TMLR 2026

MMFormalizer: Multimodal Autoformalization in the Wild

Jing Xiong, Qi Han, *Yunta Hsieh*, Hui Shen, Hongsheng Xin, Chaofan Tao, Chenyu Zhao, et al.
arXiv preprint, 2026

PhyX: Does Your Model Have the “Wits” for Physical Reasoning?

Hui Shen, Taiqiang Wu, Qi Han, *Yunta Hsieh*, Jizhou Wang, Yuyue Zhang, Yuxin Cheng, et al.
arXiv preprint, 2025

Research Experience

Graduate Research Assistant, Prof. Mi Zhang

Aug, 2025 – Present

AIoT and Machine Learning Systems Lab, Department of Computer Science and Engineering, The Ohio State University

- Conduct research on multimodal agents, vision-language-action systems, and GUI-oriented evaluation benchmarks.
- Work on efficient multimodal model deployment and benchmark construction for computer-using agents.

Skills

Programming Languages: Python, C/C++, Java, SQL, HTML/CSS

Tools and Frameworks: Git, PyTorch, LaTeX, Docker, Markdown