

KAIYUAN HOU

New York, NY • (303) 808-9867 • kh3119@columbia.edu
hou-kaiyuan.github.io • [Google Scholar](#)

SUMMARY

Ph.D. candidate at Columbia University working on agentic AI, multimodal sensing systems, and embodied intelligence. Built LLM agents, VLM training and inference systems, and deployable physical-world AI systems, with publications at MobiCom, SenSys, ICRA, IPSN, and IMWUT. Industry experience in VLM post-training and efficient inference, and LLM-agent optimization for task automation. Available full-time from Jan 2027.

EDUCATION

Columbia University Ph.D. in Electrical Engineering, GPA 4.03/4.00 Advisor: Prof. Xiaofan (Fred) Jiang, Intelligent and Connected Systems Lab (ICSL)	09/2021 – 12/2026 (expected)
University of Colorado Boulder B.S. in Electrical and Computer Engineering, GPA 3.98/4.00, summa cum laude with highest honors	09/2017 – 05/2021

EXPERIENCE

Amazon – Applied AI (now QuickSight) <i>Applied Scientist Intern</i> Agent optimization: Built a framework for distilling LLM-orchestrated agents into deterministic workflow agents, reducing inference cost by 37% at comparable task success; validated on both an internal benchmark and public datasets.	Summer 2026
---	--------------------

Ambarella <i>Software Engineer Intern</i> Reconfigurable VLM training: Developed a pipeline for training vision-language models (VLMs) with interchangeable vision encoders and language backbones, along with model-serving and evaluation pipelines. Inference optimization: Investigated vision-token pruning and quantization for efficient VLM inference.	Summer 2025
--	--------------------

Intelligent and Connected Systems Lab, Columbia University <i>Graduate Research Assistant</i> Agentic & Embodied AI: Developed intelligent systems that connect foundation-model agents with the physical world through sensing and actuation. Built reconfigurable embodied agents for natural-language task execution and adaptive decision-making, including autonomous drone and smart-home systems capable of dynamically composing sensing and actuation capabilities. (SenSys'25, MobiCom'24 Best Demo Runner-Up, MobiCom'23, IoTDI'23) Multimodal Foundation Models: Investigated how foundation models understand and reason over physical-world and non-traditional multimodal inputs, including aerial imagery, egocentric perception, IMU signals, and IoT sensor data. Developed benchmarks and evaluation methods for model robustness and reliability, and explored techniques for improving sensory representations for foundation models. (IMWUT'26, MobiCom'24, FMSys'24/25, arXiv) Human-Centered Intelligent Sensing: Developed deployable sensing and interactive systems for health and physical environments, spanning thermal fever screening, acoustic breathing monitoring, AR-assisted cardiac auscultation, and spatial visualization of IoT sensor data. A deployed fever-screening system has operated continuously for over three years, screening more than 40,000 individuals. (BuildSys'26, IPSN'22/23, SenSys'22)	09/2021 – Present
--	--------------------------

AIoT Lab, The Chinese University of Hong Kong <i>Research Assistant</i> Generative AR for sensor data: Developed an embedding-blending method that generates coherent scenes that reflect the changes in sensor readings.	01/2024 – 06/2024
---	--------------------------

SKILLS

Programming	Python, C/C++, SQL, MATLAB, Verilog, RISC-V
AI/ML Frameworks	PyTorch, TensorFlow, LangChain/LangGraph
Systems & Tools	Docker/Podman, ROS, Linux, Android, EAGLE, SolidWorks

AWARDS & SERVICE

Best Paper Award, ACM HumanSys, 2025

Best Demo Runner-Up, ACM MobiCom, 2024

Technical Program Committee: ACM IASA, 2025

Reviewer: ACL; ACM TIOT/TOSN/Health; IEEE TMC/Pervasive Computing

TEACHING & MENTORING

Research mentor (2022 – 2025): Mentored undergraduate researchers (NSF REU, Columbia SURE); one mentee received a Best Presentation Award; one mentee had a peer-reviewed publication.

Teaching Assistant, Columbia (2021 – 2025): IoT; Reinforcement Learning; Blockchain; Embedded AI.

PUBLICATIONS

** denotes co-first authors*

L. Xu, B. Yang, S. Jiang, K. Liu, **K. Hou**, Y. Fan, H. Chen, Z. Yan, X. Jiang. Pro²Assist: Continuous Step-Aware Proactive Assistance with Multimodal Egocentric Perception for Long-Horizon Procedural Tasks. *IMWUT*, 2026.

K. Hou, X. Di, X. Jiang. T-Gated: A Scene-Aware Framework for Occlusion-Resilient Passenger Analytics in Bus Interiors. *BuildSys*, 2026.

Y. Liu, M. Zhao, **K. Hou**, J. Xia, C. Carver, S. Xia, X. Zhou, X. Jiang. Moth: A Low-cost IR-based Approach Towards Autonomous Precision Drone Landing. *ICRA*, 2026.

M. Zhao*, J. Xia*, **K. Hou***, Y. Liu, S. Xia, X. Jiang. FlexiFly: Interfacing the Physical World with Foundation Models Empowered by Reconfigurable Drone Systems. *SenSys*, 2025.

L. Xu*, **K. Hou***, X. Jiang. Exploring the Capabilities of LLMs for IMU-based Fine-grained Human Activity Understanding. *FMSys*, 2025.

Y. Sui, Y. Zhang, Y. Liu, M. Zhao, **K. Hou**, J. Nie, X. Jiang, S. Xia. DomAIn: Towards Programless Smart Homes. *HumanSys*, 2025. **Best Paper Award**.

K. Hou*, M. Zhao*, L. Xu, Y. Fan, X. Jiang. TDBench: Benchmarking Vision-Language Models in Understanding Top-Down Images. *arXiv:2504.03748*, 2025.

K. Hou, Y. Guo, H. Fu, H. Chen, Z. Yan, G. Xing, X. Jiang. Improving On-Device LLMs' Sensory Understanding with Embedding Interpolations. *MobiCom*, 2024.

M. Zhao*, **K. Hou***, J. Xia, S. Xia, X. Jiang. EmbodiedRDA: Connecting Foundation Models with the Physical World using Reconfigurable Drone Agents. *MobiCom*, 2024. **Best Demo Runner-Up**.

S. Xia, M. Zhao, C. Adhivarahan, **K. Hou**, Y. Chen, J. Nie, E. Wu, K. Dantu, X. Jiang. Anemol: A Low-cost Sensorless Indoor Drone System for Automatic Mapping of 3D Airflow Fields. *MobiCom*, 2023.

K. Hou, S. Xia, E. Bejerano, J. Wu, X. Jiang. ARSteth: Enabling Home Self-Screening with AR-Assisted Intelligent Stethoscopes. *IPSN*, 2023.

M. Zhao, S. Xia, J. Nie, **K. Hou**, A. Dhupar, X. Jiang. LegoSENSE: An Open and Modular Sensing Platform for Rapidly-Deployable IoT Applications. *IoTDI*, 2023.

K. Hou, S. Xia, J. Wu, M. Zhao, E. Bejerano, X. Jiang. AI Stethoscope for Home Self-Diagnosis with AR Guidance. *SenSys*, 2022.

M. Zhao, Y. Liu, A. Dhupar, **K. Hou**, S. Xia, X. Jiang. A Modular and Reconfigurable Sensing and Actuation Platform for Smarter Environments and Drones. *MobiSys*, 2022.

K. Hou, Y. Liu, P. Wei, C. Yang, H. Kang, S. Xia, T. Spada, A. Rundle, X. Jiang. A Low-Cost In-Situ System for Continuous Multi-Person Fever Screening. *IPSN*, 2022.