

Xiaoxing Hu

✉ xiaoxinghhh@gmail.com 📞 15111480307

Education

Shanghai Jiao Tong University	Ph.D. Student, Control Science and Engineering	Sep. 2026 – Present
Shanghai Innovation Institute	Ph.D. Student (Joint Program)	Sep. 2026 – Present
Beijing Institute of Technology	M.S., Information and Communication Engineering	Sep. 2023 – Jun. 2026
Central South University	B.Eng., Automation	Sep. 2019 – Jun. 2023

Research Interests

- Multimodal foundation models: vision–language representation alignment, multimodal retrieval, and cross-modal understanding.
- Audio-visual understanding and evaluation: evidence-grounded reasoning benchmarks and automated evaluation pipelines.
- Agent systems: data pipeline design, framework engineering, evaluation methodology, and model optimization.
- Transfer and domain generalization of vision foundation models for remote sensing imagery.

Publications

Bold denotes the author. [†] Equal contribution. * Student first author.

EvaOmni: An Evidence-Grounded Benchmark for Audio-Visual Reasoning

Xiaoxing Hu, Wenxi Yao, Chaoyuan Shi, Ziqi Ye, Xiaorong Guo, Peixuan Li, Zhuoran Sun, Ziyang Gong, Peiyuan Zhang, Mingxin Liu, Shuran Ma, Hongbo Liu, Xiaomin Yu, Yu Tian, Wentao Wang, Zhenjie Yang, Jian Hu, Yunhang Shen, Haoyu Cao, Xue Yang
Under Review

Project Code Paper

ProCLIP: Progressive Vision-Language Alignment via LLM-based Embedder

Xiaoxing Hu, Kaicheng Yang, Ziyang Gong, Qi Ming, Zonghao Guo, Yu Tian, Xiang An, Ziyong Feng, Xue Yang
NeurIPS 2026 (Spotlight) · CCF-A

Code Paper

Earth-Adapter: Bridge the Geospatial Domain Gaps with Mixture of Frequency Adaptation

Xiaoxing Hu[†], Ziyang Gong, Yupei Wang, Yuru Jia, Fei Lin, Dexiang Gao, Ke An, Jianhong Han, Zhuoran Sun, Gen Luo, Xue Yang

AAAI 2026 · CCF-A

Code Paper

Decoupled Global-Local Alignment for Improving Compositional Understanding

Xiaoxing Hu[†], Kaicheng Yang, Jun Wang, Haoran Xu, Ziyong Feng, Yupei Wang

ACM MM 2025 · CCF-A

Project Code Paper

Boosting Domain Generalization in Remote Sensing Image Segmentation via Style Mapping and General Prototypical Contrast

Yupei Wang, **Xiaoxing Hu**^{*}, Yongkang Hu, Zhuoran Sun, Shanghang Zhang, Liang Chen

International Journal of Computer Vision (IJCV) 2025 · CCF-A

Paper

Internships

Tencent Youtu Lab · ShenNong Research Center *Research Intern* Dec. 2025 – Sep. 2026

- Contributed to omni-modal model development, owning the data, training, and evaluation loop.
- Designed and built an automated evaluation system for capability comparison and regression testing across multimodal models.
- Helped establish dataset evaluation and cleaning workflows; wrote and submitted one paper (under review).

DeepGlint · Foundation Model Group *Research Intern* Jun. 2025 – Oct. 2025

- Conducted research on multimodal retrieval foundation models, focusing on vision–language representation alignment.
- The resulting first-authored work was accepted to NeurIPS 2026 as a Spotlight (ProCLIP).

Honors & Awards

- **Xu Teli Scholarship Candidate** — 6 graduate students selected university-wide Jun. 2026
- **National Scholarship** Oct. 2025

Skills

- **Languages:** Python (proficient), C/C++, Shell
- **Tools:** PyTorch, Hugging Face Transformers, DeepSpeed, Git, $\LaTeX$
- **Domain:** multimodal pretraining and alignment, instruction tuning, retrieval, benchmark construction