

Curriculum Vitae

Personal details

Surname / Family Name: Luo

First name(s): Yaxin

Mobile Phone: +971 585699266 (UAE); +86 17882057622 (China)

Email address: Yaxin.Luo@mbzuai.ac.ae & luoyaxin999@gmail.com

Personal Website: <https://yaxin9luo.github.io/>

Google Scholar: <https://scholar.google.com/citations?user=tEaSCzYAAAAJ&hl=en>

Research Interests

My research vision centers on building Autonomous Agents that can perceive, reason, and act in both digital and physical environments. To achieve this, I focus on developing Native Multimodal Foundation Models with unified understanding, reasoning, generation and agentic capabilities—these models serve as the core intelligence that powers agentic systems to effectively execute complex real-world tasks (computer/devices use, autonomous research like doing AI research to updating model's own parameters.)

Education

MBZUAI

PhD of Machine Learning

Expected graduation date: June,2029

Start date: 15/08/2025

Advisor: Prof. Zhiqiang Shen & Prof. Mohsen Guizani

Technical University of Denmark

Bachelor of General Engineering (specialized in Machine Learning)

2800 Kongens Lyngby

Graduation date: 21/03/2025

Start date: 01/09/2021

Bachelor thesis advisor: Prof. Dimitrios Papadopoulos

University of Edinburgh

Bachelor of Mathematics (Overall grade of taken courses: UK first-class)

Start date: 21/09/2020

Date of withdrawing: 19/03/2021

Reason for withdrawing: change major and country

Working Experience

MBZUAI

Research Assistant

Jan 2025 – Aug 2025

Conducting a project of analyzing LLM's generalization ability on pure vision tasks using only image data. Also, explore reasoning in MLLMs.

Conference/Journal Publications

- NeurIPS 2025** Luo Y, Li Z, Liu J, Cui J, Zhao X, Shen Z. Open CaptchaWorld: A Comprehensive Web-based Platform for Testing and Benchmarking Multimodal LLM Agents. arXiv preprint arXiv:2505.24878. 2025 May 30.
- NeurIPS 2025** Cui, J., Bi, X., Luo, Y., Zhao, X., Liu, J. and Shen, Z., 2025. FADRM: Fast and Accurate Data Residual Matching for Dataset Distillation. arXiv preprint arXiv:2506.24125.
- ACL 2025 Main** Chen, J., Myrzakhan, A., Luo, Y., Khan, H.M., Bsharat, S.M. and Shen, Z., 2025. DRAG: Distilling RAG for SLMs from LLMs to Transfer Knowledge and Mitigate Hallucination via Evidence and Graph-based Distillation. arXiv preprint arXiv:2506.01954.
- CVPR 2025** Chen, X., Luo, Y., Luo, G., Ji, J., Ding, H. and Zhou, Y., 2025. DVIN: Dynamic Visual Routing Network for Weakly Supervised Referring Expression Comprehension. In Proceedings of the Computer Vision and Pattern Recognition Conference (pp. 14347-14357).
- ICLR 2025** Luo, Y., Luo, G., Ji, J., Zhou, Y., Sun, X., Shen, Z. and Ji, R., 2024. γ -MoD: Exploring Mixture-of-Depth Adaptation for Multimodal Large Language Models. arXiv preprint arXiv:2410.13859.
- ECCV 2024** Luo, Y., Ji, J., Chen, X., Zhang, Y., Ren, T. and Luo, G., 2024, September. APL: Anchor-Based Prompt Learning for One-Stage Weakly Supervised Referring Expression Comprehension. In European Conference on Computer Vision (pp. 198-215). Cham: Springer Nature Switzerland.

Other Experience

IEEE Cybermatics Congress 2024

Aug 2024

Conference Local Team Member

Acting as a conference helper and the session chair of the Smart Data workshop

SciSec2024

Aug 2024

Conference helper

Other Skills

- Good practical skills in HPC training (Slurm), DeepSpeed for LLMs, and some theoretical knowledge in Distributed Computing
- Experience in 3 Course Projects in Embedded System Programming and Circuits
- Signal Processing and Knowledge in Acoustics
- Some Basic Knowledge in Bioinformatics from DTU bachelor's biology courses