

XIANGYU LI

Ph.D. Student (3rd Year)

Institute for AI Industry Research (AIR), Tsinghua University



CONTACT

Address: Qidi Technology Building, Block C, Floor 11
Haidian, Beijing, 100084, China

Mobile: +86-159-1091-2361

Email: lixiangy22@mails.tsinghua.edu.cn

EDUCATION

Ph.D. Electronic Science and Technology *Sep. 2022 - Now*
Department of Electronic Engineering, Tsinghua University, Beijing, China

B.Eng. Electronic Information Science and Technology *Sep. 2018 – Jun. 2022*
Department of Electronic Engineering, Tsinghua University, Beijing, China

RESEARCH INTEREST

I am advised by Professor **Yunxin Liu** at **Institute for AI Industry Research (AIR), Tsinghua University**. My research interest lies in the intersection of **Mobile Computing** and **Efficient Deep Learning**. I am currently focused on the efficient deployment of LLMs (Large Language Models) on edge/mobile devices.

PUBLICATIONS

[MobiCom 2024] X. Li, Y. Li, Y. Li, T. Cao, and Y. Liu, “FlexNN: Efficient and Adaptive DNN Inference on Memory-Constrained Edge Devices”, in *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*, 2024.

Y. Li, H. Wen, W. Wang, X. Li, Y. Yuan, G. Liu, J. Liu, W. Xu, X. Wang, Y. Sun, et al., “Personal LLM Agents: Insights and Survey about the Capability, Efficiency and Security”, arXiv preprint, 2024.

[ISCA 2022] G. Dai, Z. Zhu, T. Fu, C. Wei, B. Wang, X. Li, Y. Xie, H. Yang, and Y. Wang, “DIMMining: Pruning-Efficient and Parallel Graph Mining on Near-Memory-Computing”, in *Proceedings of the 49th Annual International Symposium on Computer Architecture*, 2022.