Email: myz20@tsinghua.org.cn https://yz-mao.github.io GitHub: https://github.com/yz-mao

EDUCATION

• Tsinghua University

Beijing, China

M.S. in Data Science and Information Technology; GPA: 3.85 / 4.00

Sep. 2020 - Jun. 2023

• Wuhan University

Wuhan, China

B.E. in Information Security; GPA: 3.87 / 4.00 (Graduate with Distinction)

Sep. 2016 - Jun. 2020

Work Experience

• Tsinghua-UC Berkeley Shenzhen Institution (TBSI)

Shenzhen, China

Research Assistant, Advisor: Prof. Wenbo Ding

Sep. 2023 - Present

- Efficient and Generalized LLMs: Explored quantization and low-rank adaptation techniques to reduce the computational complexity and memory footprint of LLMs for deployment on resource-constrained platforms.
- Code Contributions: https://github.com/yz-mao/Quantized-LoRA-Finetuning-of-GPT-2
- Meituan Technology

Shenzhen, China

Algorithm Engineer Intern, UAV Group, Manager: Tianjian Chen

Jun. 2022 - Sep. 2022

- o Texture Generation: Empowered a 2D to 3D platform for large-scale UAV simulations by creating realistic textures from 2D images.
- Tencent Technology

Shenzhen, China

Algorithm Engineer Intern, Robotics-X Lab, Manager: Cheng Zhou

Jun. 2021 - Sep. 2021

- Optimization Acceleration with GPU: Enhanced the performance of an open-source JAX-based rigid body dynamics algorithm library by leveraging GPU acceleration for faster computations.
- Code Contributions: https://github.com/Tencent-RoboticsX/jbdl
- Institute for AI Industry Research (AIR), Tsinghua University

Beijing, China

Research Assistant, Advisor: Prof. Yang Liu

Aug. 2020 - Jun. 2021

o Efficient and Reliable Federated Learning: Designed and experimented with adaptive quantization methods to reduce communication overhead in distributed learning systems.

Additional Research Projects

• Deep Learning-based Text-to-Speech Synthesis System

Wuhan, China

Distinct Undergraduate Thesis of Wuhan University in 2020

Jan. 2020 - Jun. 2020

- Speech Synthesis: Studied end-to-end speech synthesis models and designed a deep learning algorithm that uses inter-frame audio features for synthesized speech detection, achieving over 90% accuracy.
- Deep Learning-based Deepfake Video Detection System

Wuhan, China

Second Prize, National College Student Information Security Contest

Jan. 2019 - Aug. 2019

- GAN-based Video Generation and Detection: Utilized both temporal and spatial features from video streams for deepfake video detection, achieving over 96% accuracy.
- Autonomous Vehicle Robotics Engineering

Singapore

Summer Workshop, School of Computing, National University of Singapore

Jun. 2018 - Sep. 2018

• Object Detection: Developed a CNN-based object detection algorithm, achieving first place in the road test.

Key Courses

• Graduate Studies

Sep. 2020 - Jun. 2023

- \circ Learning from Data: 4.0/4.0
- Optimization Theory and Machine Learning: 4.0/4.0
- Advanced Signal Processing: 4.0/4.0

• Undergraduate Studies

Sep. 2016 - Jun. 2020

- Data Structures: 4.0/4.0
- Probability Theory and Statistics: 4.0/4.0
- Operating Systems: 4.0/4.0
- Database Principles and Security: 4.0/4.0
- Pattern Recognition: 4.0/4.0

(* denotes equal contribution)

REFEREED JOURNAL ARTICLES

- [1] SAFARI: Sparsity-enabled Federated Learning with Limited and Unreliable Communications Yuzhu Mao*, Zihao Zhao*, Meilin Yang, Le Liang, Yang Liu, Wenbo Ding, Tian Lan, Xiao-Ping Zhang IEEE Transactions on Mobile Computing (TMC), 2023.
- [2] AQUILA: Communication-efficient Federated Learning with Adaptive Quantization in Device Selection Strategy Zihao Zhao*, Yuzhu Mao*, Zhenpeng Shi, Yang Liu, Tian Lan, Wenbo Ding, Xiao-Ping Zhang IEEE Transactions on Mobile Computing (TMC), 2023.
- [3] Towards Efficient Communications in Federated Learning: A Contemporary Survey Zihao Zhao, **Yuzhu Mao**, Yang Liu, Linqi Song, Ye Ouyang, Xinlei Chen, Wenbo Ding *Journal of the Franklin Institute*, 2023.
- [4] Communication-efficient Federated Learning with Adaptive Quantization Yuzhu Mao, Zihao Zhao, Guangfeng Yan, Yang Liu, Tian Lan, Linqi Song, Wenbo Ding ACM Transactions on Intelligent Systems and Technology (TIST), 2022.

CONFERENCE PROCEEDINGS

[6] FL-TAC: Enhanced Fine-tuning in Federated Learning via Low-rank, Task-specific Adapter Clustering Siqi Ping*, Yuzhu Mao*, Yang Liu, Xiao-Ping Zhang, Wenbo Ding International Conference on Learning Representations (ICLR) Workshop on Large Language Model (LLM) Agents, 2024.

AWARDS AND HONORS

- Tsinghua Graduate Scholarship for Excellent Academic Performance (2020-2021 and 2021-2022, First-class, Top 3%)
- Wuhan University Scholarship for Outstanding Undergraduates (2020, **Top 3%**)
- National Cyber Security Scholarship (2019, Top 1%)
- National Scholarship (2018, Top 1%)
- Wuhan University Scholarship for Outstanding Students (2016-2017, 2017-2018, and 2018-2019, First-class, Top 3%)

Programming Skills

• Tools: PyTorch, TensorFlow, Git, Linux Languages: Python, C, C++, Java, MATLAB, LaTex

Language Skills

• English: IELTS overall Band 8.0