

# Yinghao Tang

 [Homepage](#) |  [GitHub](#) |  [Email](#) |  [Google Scholar](#)

## ABOUT

---

Yinghao Tang is a second-year Ph.D. student in Computer Science at Zhejiang University, advised by Prof. Wei Chen in the VAI group. His research lies at the intersection of AI systems and intelligent computation, focusing on building efficient infrastructures for training, serving, and optimizing large-scale models. He is particularly interested in AI system design, scheduling, and resource optimization for deep learning frameworks.

## EDUCATION

---

- 2023 - present Ph.D. in Computer Science at Zhejiang University  
Advisor: Prof. Wei Chen
- 2019 - 2023 B.Eng. in Computer Science at Sichuan University  
Advisor: Prof. Mingjie Tang

## EXPERIENCE

---

- Dec 2023 - Jul 2024 Research Intern at Ant Group, Remote  
Computation Resource Scheduling of Deep Learning Training Framework

## PUBLICATIONS

---

- **Y. Tang**, T. Lan, X. Huang, H. Lu, and W. Chen. “SCORPIO: Serving the Right Requests at the Right Time for Heterogeneous SLOs in LLM Inference.” *Preprint*, 2025. [arXiv:2505.23022](#)
- Z. Wen, L. Weng, **Y. Tang**, R. Zhang, Y. Liu, B. Pan, M. Zhu, and W. Chen. “Exploring Multimodal Prompt for Visualization Authoring with Large Language Models.” *Preprint*, 2025. [Link](#)
- L. Weng, **Y. Tang**, Y. Feng, Z. Chang, R. Chen, H. Feng, C. Hou, D. Huang, Y. Li, H. Rao, H. Wang, C. Wei, X. Yang, Y. Zhang, Y. Zheng, X. Huang, M. Zhu, Y. Ma, B. Cui, P. Chen, and W. Chen. “DataLab: A Unified Platform for LLM-Powered Business Intelligence.” In *41st IEEE International Conference on Data Engineering (ICDE)*, 2025. [arXiv:2412.02205](#)
- Q. Wang, T. Lan, **Y. Tang**, B. Sang, Z. Huang, Y. Du, H. Zhang, J. Sha, H. Lu, and Y. Zhou. “DLRover-RM: Resource Optimization for Deep Recommendation Models Training in the Cloud.” In *50th International Conference on Very Large Data Bases (VLDB)*, 2024. [Link](#)
- D. Li, Y. Ma, N. Wang, Z. Ye, Z. Cheng, **Y. Tang**, Y. Zhang, L. Duan, J. Zuo, C. Yang, and M. Tang. “MixLoRA: Enhancing Large Language Models Fine-Tuning with LoRA-based Mixture of Experts.” *Preprint*, 2024. [arXiv:2404.15159](#)

## HONORS AND AWARDS

---

- 2020 Outstanding Student Award at Sichuan University  
2022 National Encouragement Scholarship