

YIZHE XIONG (熊翊哲)

✉ xiongyizhe2001@163.com · 📄 Google Scholar · in linkedin.com/in/bostoncake · 🏠 xiongyizhe.xyz

🎓 EDUCATION

Tsinghua University, Beijing, China 2022 – Present (expected June 2027)

Ph.D. in Software Engineering, GPA 3.9/4.0

Research focus: Transfer Learning, Pretraining and Fine-Tuning Foundation Models

Tsinghua University, Beijing, China 2018 – 2022

B.S. in Computer Science and Technology, GPA 3.7/4.0

🔧 SKILLS

- Programming Language: C, C++, Python
- Deep learning development: PyTorch, torchvision, Huggingface transformers, timm, MMCV, etc.

👤 EXPERIENCE

Kuaishou Technology October 2023 – Present

Machine Learning Intern @ Department of Foundation Models and Multimedia

Researching in Large Language Models (LLMs) and Multimodal Large Language Models (MLLMs), focusing on pretraining, fine-tuning, and enhancing model capabilities.

- Investigated the learning dynamics of LLMs and submitted a first-author conference paper: Temporal Scaling Law for Large Language Models, currently under review.
- Contributed to the development of the tokenizer and Mixture-of-Experts (MoE) feature in the KuaiYii model, leading to the submission of Scaffold-BPE and MaskMoE.
- Developed the hyperparameter selection method for the KwaiYii model using Temporal Scaling Law for Large Language Models.
- Developed the fine-tuning process to enable long-context support for the KwaiYii model.

Microsoft (China) Co. Ltd May 2021 – August 2021

Software Engineering Intern @ Cloud and AI (C+AI), Microsoft Azure

Enhanced Synapse and developer tooling in Azure Powershell.

- Developed and maintained key features for managing Synapse workspaces.
- Streamlined the PowerShell development process by optimizing the integration pipeline for the AutoRest tool.

📖 PUBLICATIONS

PYRA: Parallel Yielding Re-Activation for Training-Inference Efficient Task Adaptation

European Conference on Computer Vision (ECCV) 2024

Authors: **Yizhe Xiong**, Hui Chen, Tianxiang Hao, Zijia Lin, Jungong Han, Yuesong Zhang, Guoxin Wang, Yongjun Bao, Guiguang Ding

Confidence-based Visual Dispersal for Few-shot Unsupervised Domain Adaptation

International Conference on Computer Vision (ICCV) 2023

Authors: **Yizhe Xiong**, Hui Chen, Zijia Lin, Sicheng Zhao, Guiguang Ding

Breaking the Stage Barrier: A Novel Single-Stage Approach to Long Context Extension for Large Language Models

International Conference on Computational Linguistics (COLING) 2025

Authors: Haoran Lian*, Junmin Chen*, Wei Huang*, **Yizhe Xiong***, Wenping Hu, Guiguang Ding, Hui Chen, Jianwei Niu, Zijia Lin, Fuzheng Zhang, Di Zhang (* denotes equal contribution)

Scaffold-BPE: Enhancing Byte Pair Encoding for Large Language Models with Simple and Effective Scaffold Token Removal

AAAI Conference on Artificial Intelligence (AAAI) 2025

Authors: Haoran Lian, **Yizhe Xiong**, Jianwei Niu, Shasha Mo, Zhenpeng Su, Zijia Lin, Peng Liu, Hui Chen, Guiguang Ding

★ AWARDS

<i>Academic Scholarship</i> , School of Software, Tsinghua University	December 2024
<i>First Place and Gold Prize</i> , VISION'24 Data Challenge, ECCV 2024	September 2024
<i>Academic Scholarship</i> , School of Software, Tsinghua University	October 2023
<i>Outstanding Graduate</i> , Department of Computer Science and Technology, Tsinghua University	June 2022

📄 ADDITIONAL INFORMATION

- GitHub: <https://github.com/bostoncake>
- Fluent in English (TOEFL 113)