Chongjie Si

Email | Google Scholar | 知乎 Zhihu | Homepage

Shanghai, China

Summary

Chongjie Si is now a Ph.D Candidate at the Artificial Intelligence Institute, Shanghai Jiao Tong University. His research interests lie in efficient LLM training and machine learning (especially math) beyond.

Education

- **Shanghai Jiao Tong University** *Sept. 2022 - Present*
  Ph.D, Artificial Intelligence, supervised by Prof. Wei Shen Shanghai, China
- **Southeast University** *Sept. 2018 - Jun. 2022*
  BSc, Chien-Shiung Wu College, Artificial Intelligence, rank 1/22 Nanjing, China

Selected Projects

- **Efficient LLM Training** *Mar. 2024 - Now*
  Optimizer and Training Strategies.
  - **Optimizer-Muon and AdaMuon [GitHub]**
    - Muon scaling law & productionization: Led the validation of Muon's scaling law; applied Muon to Xiaohongshu's LLM training, improving large-batch stability and reducing compute waste during long runs.
    - AdaMuon: Designed AdaMuon that combines Muon's Newton–Schulz orthogonalization with adaptive second-momentum scheduling; delivered greater training efficiency.
  - **Training strategies — PEFT at scale [GitHub], addressing several challenges as follows:**
    - Explanation / Unification for the research area
    - 2D methods for high-dimensional parameter space
    - Initialization
    - Task-related alignment
    - Direction & Magnitude decomposition
    - Dynamic parameter budget allocation
    - High-rank updates
- **Machine Learning Research** *Jan. 2022 - Now*
  Convex Optimization, Matrix Theory, Linear Algebra, Probability Theory
  - **Partial Label Learning**, investigate three essential problems:
    - What information can be used for disambiguation
    - How to supervise the process of disambiguation
    - Why accurate model can learn from partial labels
  - **Multi Label Learning**, investigate two problems:
    - High-order MLL method
    - Ground-truth Labels with partial labels

Work Experience

- **Xiaohongshu [Web]** *Jun 2025 - Now*
  REDstar Intern, mentored by Dr. Debing Zhang Shanghai, China
  - LLM pre-training.
  - Explore efficient pre-training strategies through optimizer design, architecture choices, and model interpretability.
- **Alibaba Group [Web]** *Mar. 2025 - Jun. 2025*
  Research Intern, mentored by Yadao Wang and Dr. Wenbo Su Hangzhou, China
  - LLM post-training.
  - Investigated efficient training strategies from the perspective of fully fine-tuning.
  - Explored model merging techniques to enhance the capabilities of foundation models without additional training.

PUBLICATIONS

Book

- 2025, Wei Shen, **Chongjie Si**, Chen Yang, Yong Yu. ***Hands on Computer Vision***. Posts & Telecoms Press.

Journal or Conference

- 2025. **Chongjie Si**, Debing Zhang, Wei Shen. **AdaMuon: Adaptive Muon Optimizer.**
- 2025. **Chongjie Si**, Xuankun Yang, Muqing Liu, Yadao Wang, Xiaokang Yang, Wenbo Su, Bo Zheng, Wei Shen. **Weight Spectra Induced Efficient Model Adaptation.**
- 2025, NeurIPS, Spotlight. Jingjing Jiang, **Chongjie Si**, Jun Luo, Hanwang Zhang, Chao Ma. **Co-Reinforcement Learning for Unified Multimodal Understanding and Generation.**
- 2025, NeurIPS. Xuehui Wang*, **Chongjie Si***, Xue Yang, Yuzhi Zhao, Wenhai Wang, Xiaokang Yang, Wei Shen. **OPMapper: Enhancing Open-Vocabulary Semantic Segmentation with Multi-Guidance Information.**
- 2025. **Chongjie Si***, Yidan Cui*, Fuchao Yang, Xiaokang Yang, Wei Shen. **Revisiting Sparsity Constraint Under High-Rank Property in Partial Multi-Label Learning.**
- 2025, COLM. Zhiyi Shi, Binjie Wang, **Chongjie Si**, Yichen Wu, Junsik Kim, Hanspeter Pfister. **DualEdit: Dual Editing for Knowledge Updating in Vision-Language Models.**
- 2025. **Chongjie Si**, Jingjing jiang, Xiaokang Yang, Wei Shen. **Unveiling the Mystery of Weight in Large Foundation Models: Gaussian Distribution Never Fades**.
- 2025. **Chongjie Si**, Kangtao Lv, Jingjing Jiang, Yadao Wang, Yongwei Wang, Xiaokang Yang, Wenbo Su, Bo Zheng, Wei Shen. **NAN: A Training-Free Solution to Coefficient Estimation in Model Merging.**
- 2025. **Chongjie Si**, Zhiyi Shi, Yadao Wang, Xiaokang Yang, Susanto Rahardja, Wei Shen. **MAP: Revisiting Weight Decomposition for Low-Rank Adaptation.**
- 2025, ICCV. **Chongjie Si**, Zhiyi Shi, Xuehui Wang, Yichen Xiao, Xiaokang Yang, Wei Shen. **Generalized Tensor-based Parameter-Efficient Fine-Tuning via Lie Group Transformations.**
- 2025. **Chongjie Si***, Yidan Cui*, Fuchao Yang, Xiaokang Yang, Wei Shen. **Why Can Accurate Models Be Learned from Inaccurate Annotations?**
- 2025, ICLR. **Chongjie Si***, Zhiyi Shi*, Shifan Zhang, Xiaokang Yang, Hanspeter Pfister, Wei Shen. **Unleashing the Power of Task-Specific Directions in Parameter Efficient Fine-tuning**.
- 2025, ICLR. **Chongjie Si***, Xuehui Wang*, Xue Yang, Zhengqin Xu, Qingyun Li, Jifeng Dai, Yu Qiao, Xiaokang Yang, Wei Shen. **Maintaining Structural Integrity in Parameter Spaces for Parameter Efficient Fine-tuning**.
- 2024. **Chongjie Si**, Xiaokang Yang, Wei Shen. **See Further for Parameter Efficient Fine-tuning by Standing on the Shoulders of Decomposition**.
- 2024. **Chongjie Si**, Xuehui Wang, Yan Wang, Xiaokang Yang, Wei Shen. **Appeal: Allow Mislabeled Samples the Chance to be Rectified in Partial Label Learning**.
- 2024, ECCV. **Chongjie Si**, Xuehui Wang, Xiaokang Yang, Wei Shen. **Tendency-driven Mutual Exclusivity for Weakly Supervised Incremental Semantic Segmentation**
- 2024, AAAI, Oral. **Chongjie Si**, Zekun Jiang, Xuehui Wang, Yan Wang, Xiaokang Yang, Wei Shen. **Partial Label Learning with a Partner**.
- 2023, TKDE. **Chongjie Si**, Yuheng Jia, Ran Wang, Min-Ling Zhang, Yanghe Feng, Chongxiao Qu. **Multi-label Classification with High-rank and High-order Label Correlations**.
- 2023, KDD, Oral. **Chongjie Si***, Yuheng Jia*, Min-ling Zhang. **Complementary Classifier Induced Partial Label Learning**.

HONORS AND AWARDS

- **Doctoral National Scholarship (Top 1%)** *Oct. 2024*
  Chinese Ministry of Education
- **Outstanding Graduate Student** *Jun. 2022*
  Southeast University
- **National Scholarship (Top 1%)** *Oct. 2019*
  Chinese Ministry of Education
- **Model of Merit Student (Top 1%)** *Oct. 2019 & 2020*
  Southeast University

SKILLS

- **Programming Languages:** Python, MATLAB, C++
- **Software & Tools:** PyTorch, Megatron-LM, DeepSpeed, OpenCV, LaTeX, PyQt5