

Da Chang

✉ changda24@mailsucas.ac.cn
🌐 Personal website: maechd.github.io
📖 Google Scholar: Da Chang
🐙 GitHub: github.com/MaeChd
🌐 LinkedIn: linkedin.com/in/da-chang-4a13262b8

Professional Experience

07/2025 – Present  **Jointly Educated Ph.D. Candidate**, Pengcheng Laboratory
Shenzhen, China
Conduct research on deep learning optimization, LLM training infrastructure, parameter-efficient fine-tuning, and optimization theory. Work on large-scale training systems and optimizer design for modern deep learning models.

Education

09/2024 – 06/2029  **Ph.D.**, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences
Pattern Recognition and Intelligent Systems
Jointly educated Ph.D. candidate with Pengcheng Laboratory
Advisors: [Ganzhao Yuan](#) and [Huihui Zhou](#); Mentor: [Yongxiang Liu](#)

09/2020 – 06/2024  **B.Eng.**, Central South University, School of Automation
Intelligent Science and Technology
Outstanding Undergraduate Thesis, Central South University

Research Areas: an Overview

My research focuses on LLM pre-training, agent policy learning, and generative AI. Specific topics include deep learning optimization and generalization; optimization for large-scale machine learning and LLM training; efficient preconditioning and Muon-type optimizers; parameter-efficient fine-tuning; and deep learning applications.

Projects

- 1 Distributed Muon-based MoE Training on 64 NPUs**
Systems for LLM Training · MoE / Muon / ZeRO / Megatron / MindSpeed / 64 NPUs
 - Built Muon and Distributed Muon in a Megatron- and MindSpeed-based 64-NPU MoE training framework.
 - Supported TP, CP, EP, and PP parallelism.
 - Observed that Muon consistently outperformed AdamW during training.
 - Reduced memory usage with Distributed Muon, with additional communication overhead.
 - Identified a loss-spike issue caused by the interaction between MoE experts and Muon redundancy.

Preprints & Under Review

- 1 [Muse: Representation Geometry of Muon Beyond Normalized Momentum](#)
Da Chang*, Qiankun Shi*, Lvqiang Zhang*, Di He, Yaoshuai Ma, Ganzhao Yuan, Yongxiang Liu
arXiv preprint, Jul. 2026
- 2 [A Note on Stability for Orthogonalized Matrix Momentum with Client Sampling](#)
Da Chang, Qiankun Shi, Lvqiang Zhang, Yu Li, Ruijie Zhang
arXiv preprint, Jun. 2026
- 3 [When Does Value-Aware KV Eviction Help? A Fixed-Contract Diagnostic for Non-Monotone Cache Compression](#)
Ruijie Zhang, Haozhe Liang, **Da Chang**, Li Hu, Fanqi Kong, Huaxiao Yin, Yu Li
arXiv preprint, May 2026
- 4 [MuonEq: Balancing Before Orthogonalization with Lightweight Equilibration](#)
Da Chang, Qiankun Shi, Lvqiang Zhang, Yu Li, Ruijie Zhang, Yao Lu, Yongxiang Liu, Ganzhao Yuan
arXiv preprint, Mar. 2026 · [Code](#)
- 5 [On the Convergence of Muon and Beyond](#)
Da Chang, Yongxiang Liu, Ganzhao Yuan
arXiv preprint, Sept. 2025 · [Code](#)

Publications

- 1 [KG-SAM: Injecting Anatomical Knowledge into Segment Anything Models via Conditional Random Fields](#)
Yu Li, **Da Chang**, Xi Xiao
IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2026 Oral), 2026
- 2 [Calibrating and Rotating: A Unified Framework for Weight Conditioning in PEFT](#)
Da Chang, Peng Xue, Yu Li, Yongxiang Liu, Pengxiang Xu, Shixun Zhang
AAAI Conference on Artificial Intelligence (AAAI 2026), 2026 · [Code](#)
- 3 [MGUP: A Momentum-Gradient Alignment Update Policy for Stochastic Optimization](#)
Da Chang, Ganzhao Yuan
Annual Conference on Neural Information Processing Systems (NeurIPS 2025 Spotlight), 2025 · [Code](#)
- 4 [SfMDiffusion: self-supervised monocular depth estimation in endoscopy based on diffusion models](#)
Yu Li, **Da Chang**, Die Luo, Jin Huang, Lan Dong, Du Wang, Liye Mei, Cheng Lei
International Journal of Computer Assisted Radiology and Surgery, JCR-Q2, 2025 · [Code](#)
- 5 [Mixed Text Recognition with Efficient Parameter Fine-Tuning and Transformer](#)
Da Chang*, Yu Li*
International Conference on Neural Information Processing (ICONIP 2024), 2024 · [Code](#)

Selected Awards & Honors

- 2026  Silver Reviewer Award, International Conference on Machine Learning (ICML)
- 2023  Second prize, 8th National Biomedical Engineering Innovation Design Competition for College Students, China
- 2023  Third prize, 9th National Statistical Modeling Competition for College Students, China
- 2023  Second Class Scholarship, Central South University (Top 15%)
- 2022  “ShanHe Excellent Student” Second Class Scholarship, Central South University (Top 5%)
- 2022  First-Class Scholarship, Central South University (Top 5%)

Academic Services

- Reviewer for TPAMI 2025, ICML 2026, and NeurIPS 2026.

Skills

- **Languages:** Python, C/C++, MATLAB, LaTeX.
- **Frameworks:** PyTorch, scikit-learn, Megatron, MindSpeed.
- **Distributed LLM Training:** TP, PP, EP, CP, MoE training, and NPU-based large-scale training.
- **AI-assisted Development:** Claude Code and Codex.