Kaiwen TUO

+86 17392763601 | cfintuo@gmail.com | HomePage | GitHub

Education

Tongji University Shanghai, CN
Bachelor of Computer Science and Technology ***(Elite Class)*** *Sep. 2022 – June 2026*
GPA: 4.90 / 5 **Rank**: 1 / 52
Research Interest: Efficient AI, Large Language Model, Multi-modal Large Language Model

Selected Honors

National Scholarship 2023 (**rank 1 / 198**), Tang Zhongying Moral Scholarship (**top 0.5%**)
National Scholarship 2024 (**rank 1 / 73**) , Outstanding Student Model (**top 0.1%**),
National Scholarship 2025 (**rank 1 / 73**) , Inspirational Honorable Mention Award (**top 0.05%**)
9 Competition awards at the provincial or national level, **11 Competition awards** at the school level

Publications **(*Equal contribution)**

SparseSSM: Efficient Selective Structured State Space Models Can Be Pruned in One-Shot
Kaiwen Tuo, Huan Wang.
Submit to ICLR 2026 [arXiv] | [Code] | [Webpage].

RewardMap: Tackling Sparse Rewards in Fine-grained Visual Reasoning via Multi-Stage Reinforcement Learning
Sicheng Feng*, **Kaiwen Tuo***, Song Wang, Lingdong Kong, Jianke Zhu, Huan Wang.
Submit to ICLR 2026 [arXiv] | [Code] | [Webpage] | [Dataset].

Research Experience

Reinforcement Learning on Map-Centric Multimodal Data for Spatial Reasoning Jul. 2025 – Oct. 2025
Advisor: Huan WANG *Encode Lab, Westlake University*

- Expanded ReasonMap into ReasonMap-Plus: 30 cities, 4,018 high-resolution samples; annotated difficulty; 5 question types covering counting and verification.
- Designed a difficulty-aware fine-grained reward and multi-stage RL curriculum to tackle sparse-reward learning.
- Achieved state-of-the-art results on ReasonMap/ReasonMap-Plus, matching or surpassing top open-source and some closed-source MLLMs and delivered consistent gains across 6 external benchmarks. Paper was submitted to **ICLR 2026**

Efficient Selective Structured State Space Models Can Be Pruned in One-Shot Feb. 2025 – Jun. 2025
Advisor: Huan WANG *Encode Lab, Westlake University*

- Developed a training-free OBS-based pruning strategy for Mamba-based LLMs, surpassing current SOTA in performance.
- Existing pruning methods primarily target Transformer architectures, cannot be directly transferred to Mamba.
- Implemented fast, unstructured pruning for Mamba-based LLMs using a single GPU, with no performance degradation at 50% sparsity. Paper was submitted to **ICLR 2026**.

Selected Competitions

China International College Students' Innovation Competition (Top 0.2%) Mar. 2024 – Sep. 2024

- Won Gold Prize of the Shanghai Provincial Competition as the **founder and leader** of the team.
- Led a startup team, developed a comprehensive business model for façade damage detection, designed the hardware layout, and deployed the detection algorithm, Planning to establish a company in the near future.

National Mathematical Modeling Contest, Shanghai Second Prize (Team Leader) Sep. 2024

- Served as **team leader**, responsible for paper writing and algorithm refinement.
- Integrated simulated annealing, SAT-based separating-axis collision detection, binary search, and BFGS optimization to conduct an in-depth study of the bench-dragon movement dynamics.

Second Prize, 15th National College Student Mathematics Competition Nov. 2023
Top Prize, National AI Application Scenario Innovation Competition – Thematic Track Nov. 2024

Leadership

President of Tongji Commercial and Investment Club Oct. 2023 - Present

- Connected and maintained relationships with 300+ outstanding alumni in business and STEM fields.
- Led club operations and handled key responsibilities such as annual event planning, recruitment drives, and alumni-industry engagement. Manage multiple internship information–sharing groups.

"TJeduTeens" Founder & Contact Person Jan. 2023 - Mar. 2023
Tongji University

- Investigated the implementation of double reduction, integrated the strengths of universities and community.
- Liaised with the Guirenmao Community in Yan'an City, Shaanxi, and communicated with several staff members.

Technical Skills

Tech Stack (selected): Python, Pytorch, C/C++, Latex, Linux, Git, Slides Design, Other essential tools for research
Hobbies (selected): Volunteer Service (**106h**), Quantitative trade (Attend J.P. Morgan's AWMC), Chinese characters