Zheng Liu

15251464261 | 2501213330@stu.pku.edu.cn | zhengliu03.github.io | starriver030515

EDUCATION

- **Nanjing University, School of Computer Science** *Sep. 2021 – Jul. 2025*
  Bachelor of Science in Computer Science GPA: 4.62/5.0, Ranking: 1/210
- **Peking University, School of Computer Science** *Sep. 2025 – Jul. 2028*
  Master of Science in Computer Science, supervised by Prof. Wentao Zhang and Prof. Bin Cui

SELECTED PUBLICATIONS *(* denotes equal contribution)*

SynthVLM: Towards High-Quality and Efficient Synthesis of Image-Caption Datasets for Vision-Language Models

Zheng Liu, Hao Liang*, Bozhou Li, Wentao Xiong, Chong Chen, Conghui He, Wentao Zhang, Bin Cui* *ACM MM 2025(CCF-A)*

- Proposed a caption-to-image synthesis pipeline using diffusion models with CLIPScore-based filtering, reversing the traditional paradigm to create precisely aligned image-text pairs.
- Curated SynthVLM-100K; with only 18% of the pretraining data, SynthVLM-7B/13B surpass LLaVA-1.5 on most VQA benchmarks while preserving language capabilities.

FLARE: Fully Integration of Vision-Language Representations for Deep Cross-Modal Understanding

***Zheng Liu**, Mengjie Liu, Jingzhou Chen, Jingwei Xu, Bin Cui, Conghui He, Wentao Zhang* *ICLR 2026(CCF-A)*

- Introduced a full-pipeline vision-language integration framework with Text-Guided Vision Encoding for pixel-level alignment, Context-Aware Recursive Alignment Decoding for query-level semantic fusion, Dual-Supervised Semantic Mapping Loss for modality-level bridging, and a language-driven QA synthesis strategy for data-level scaling.
- Our model achieves comparable performance against Qwen2.5-VL with significantly fewer vision tokens.

From Uniform to Heterogeneous: Tailoring Policy Optimization to Every Token's Nature

Zheng Liu, Mengjie Liu*, Siwei Wen, Mengzhang Cai, Bin Cui, Conghui He, Wentao Zhang* *ACL 2026 (Under Review)*

- Proposed HAPO, a token-aware RL framework that adapts sampling temperature, advantage redistribution, and clipping boundaries based on per-token entropy, embedding heterogeneous treatment into every optimization stage.
- Consistently outperforms DAPO on math reasoning and code benchmarks with negligible computational overhead.

ChartVerse: Scaling Chart Reasoning via Reliable Programmatic Synthesis from Scratch

Zheng Liu, Honglin Lin*, Chonghan Qin, . . . , Bin Cui, Conghui He, Lijun Wu, Wentao Zhang* *ACL 2026 (Under Review)*

- Designed a complexity-aware chart synthesis engine and an answer-first QA generation pipeline with programmatic verification, eliminating reliance on external proprietary models.
- Curated 600K verifiable chart reasoning data, enabling significant gains on chart reasoning benchmarks for VLMs.

Scientific Image Synthesis: Benchmarking, Methodologies, and Downstream Utility

Honglin Lin, **Zheng Liu***, Chonghan Qin*, Qizhi Pei, Yu Li, Zhanping Zhong, . . . , Conghui He, Lijun Wu* *ACL 2026 (Under Review)*

- Systematically benchmarked scientific image synthesis paradigms with metrics emphasizing scientific correctness; analyzed downstream utility for boosting multimodal reasoning.

MMFineReason: Closing the Multimodal Reasoning Gap via Open Data-Centric Methods

Honglin Lin, **Zheng Liu***, Yun Zhu*, Chonghan Qin, Xiaoran Shang, Conghui He, Wentao Zhang, Lijun Wu* *ECCV 2026 (Under Review)*

- Contributed to an open pipeline for long-form multimodal reasoning traces (collection, cleaning, verification), supporting scalable instruction tuning.
- Open-sourced MMFineReason-1.8M. An 8B model trained on our dataset surpasses Gemini-2.5-Flash and approaches GPT-5-mini-High and Qwen3-VL-32B-Thinking on multimodal reasoning benchmarks.

PROJECTS

MinerU2.5: A Decoupled Vision-Language Model for Efficient High-Resolution Document Parsing

Junbo Niu, **Zheng Liu***, Zhuangcheng Gu*, Bin Wang*, . . . , Bowen Zhou, Dahua Lin, Wentao Zhang, Conghui He*

- Developed a decoupled coarse-to-fine VLM pipeline for high-resolution PDF/document parsing, improving robustness on dense text, tables, and formulas.
- Contributed to system optimization and evaluation workflows, improving end-to-end stability and efficiency on real-world documents.

- Garnered 54K+ GitHub stars, reflecting broad community adoption.

OpenDataArena: A Fair and Open Arena for Benchmarking Post-Training Dataset Value

Mengzhang Cai, Xin Gao*, Yu Li*, Honglin Lin*, **Zheng Liu***, . . . , Dahua Lin, Conghui He, Lijun Wu*

- Built an open evaluation pipeline to quantify post-training dataset value with multi-dimensional scoring and fair cross-dataset comparisons.
- Supported large-scale experimentation, enabling reproducible dataset contribution measurement for post-training.

EXPERIENCE

- **OpenDataLab, Shanghai AI Laboratory** *2024.09 – 2025.08*
  Research Intern | *Multimodal LLM Understanding, OCR-Free Visual Understanding* Supervised by Dr. Conghui He
- **OpenDataLab, Shanghai AI Laboratory** *2025.09 – 2026.02*
  Research Intern | *Multimodal LLM Reasoning, Scientific Image Understanding, Reasoning Data Synthesis* Supervised by Dr. Lijun Wu

HONOURS AND AWARDS

- GuoXieBirong Scholarship, Nanjing University (Top 2) *2023*
- National Scholarship, Nanjing University (Top 1) *2024*
- Gold Award, 20th Programming Competition, Nanjing University *2024*

SKILLS

Programming: C, C++, Python.
Frameworks & Tools: PyTorch, Transformers, LLaVA, Verl, LLaMA-Factory
Research Interests: Vision-Language Models, Multimodal Reasoning, Document Parsing, Synthetic Data.