Ziteng Wang

Personal Webpage

Email : u13zw21@abdn.ac.uk

Mobile : +86-13711191538

EDUCATION

University of Aberdeen Aberdeen, Scotland
BSc. Computing Science; First Class Honor; GPA: 19.03/23.00 *Sep. 2021 – Jun. 2025*

South China Normal University Guangzhou, China
BEng. Software Engineering (Dual degree); GPA: 85.7/100 *Sep. 2021 – Jun. 2025*

PUBLICATIONS

[C1] VITRIX-CLIPIN: Enhancing Fine-Grained Visual Understanding in CLIP via Instruction-Editing Data and Long Captions [Project] NeurIPS 2025
Ziteng Wang, Siqi Yang, Limeng Qiao, Lin Ma
[C2] From What to Where: Reward-Driven Post-Training for Spatially-Aware MLLMs Submitted to CVPR 2026
Ziteng Wang, Jiancheng Huang, Yu Gao, Liang Xiang
[R1] LongCat-Flash-Omni Technical Report [Report]
Ziteng Wang and Meituan LongCat Team

RESEARCH EXPERIENCE

LongCat Research, Meituan Beijing, China
Advisor: Dr. Lin Ma, Dr. Siqi Yang *Sep. 2024 - May. 2025*

- Proposed to leverage symmetric hard negative loss and instruction editing data to enhance CLIP's performance on fine-grained visual understanding and compositional visual understanding, as well as zero-shot classification and retrieval.
- Simultaneously improving MLLM's performance on general and fine-grained visual understanding tasks.
- Related publication: [C1].

AML, ByteDance Seed Seattle, WA
Advisor: Dr. Jiancheng Huang, Dr. Liang Xiang *May. 2025 - Nov. 2025*

- Developed and implemented a post-training framework using SFT and RLHF algorithms, incorporating spatially-aware reward signals to enhance the ability of MLLMs to understand object location, distance, and spatial relationships within images.
- Related publication: [C2].

NeuroAI LAB, University of Washington Seattle, WA
Advisor: Prof. Eli Shlizerman *Nov. 2025 - Present*

- Working on a project regarding audio-visual spatial joint understanding. The goal is to build a spatially-aware language and audio vision encoder.
- Be responsible for the data curation pipeline on FOV binaural video.

The Chinese University of HongKong, Shenzhen Shenzhen, China
Advisor: Prof. Haizhou Li *Jul. 2024 - Sep. 2024*

- Implemented an end-to-end Arabic Automatic Speech Recognition (ASR) system

INDUSTRIAL EXPERIENCE

Meituan LongCat Beijing, China
Research & MLE Intern *Sep. 2024 - May. 2025*

- Implemented a data pipeline for large scale language models training, including data ratio and accuracy verification.
- Related Project: [R1][P1][P2][P3].

- **Google** — Shanghai, China
  SDE Intern — *Jul. 2023 - Sep. 2023*
  - Building Efficient AI infra structures for large scales language models training and inference.

Research Interest

- Multimodal Visual Representation. I am interested in how to make MLLMs see more clearly and correctly (i.e. Better visual features extraction/fusion).

- Post-training pipelines for MLLMs. I am also interested in enhancing MLLMs' performance on specific tasks (i.e. maths reasoning) by reinforcement learning-related algorithms.

Projects

[P1] LongCat-Flash: A 560-billion-parameter Mixture-of-Experts (MoE) language model designed for both computational efficiency and advanced agentic capabilities.[Demo][Model][Code][Report]

[P2] LongCat-Flash-Thinking: An efficient 560-billion-parameter open-source Mixture-ofExperts (MoE) reasoning model.[Demo][Model][Code][Report]

[P3] LongCat-Vedio: A foundational video generation model with 13.6B parameters, delivering strong performance across multiple video generation tasks.[Model][Code][Report]

Honors & Award

NeurIPS Student Scholar Award
University of Aberdeen Academic Achievement Award (3.75 cGPA or higher)

Programming Skills

- **Languages**: Python, Javascript, SQL, Java   **Training frameworks**: megatron, Openclip, OpenRLHF, verl