Ziyang Wang

ziyangw@cs.unc.edu | ziyangw2000.github.io/ | Ziyang412 | ziyang-wang-882188203

Education

University of North Carolina at Chapel Hill — Chapel Hill, NC
Ph.D. in Computer Science, advised by Prof. Mohit Bansal — Aug. 2022 - Exp. May 2027

University of North Carolina at Chapel Hill — Chapel Hill, NC
M.S. in Computer Science (earned en route to Ph.D.) — Dec. 2024

University of Electronic Science and Technology of China — Chengdu, China
B.Eng. in Software Engineering, GPA: 3.92/4.0 — Sept. 2018 - Jun. 2022

Research Interest

I build multimodal systems that perceive and reason over long and complex videos. My recent work follows three directions: (1) **agentic video understanding**, where agents actively decide what to watch and when to stop (VideoTree, Active Video Perception); (2) **efficient video reasoning with LLMs**, through data-efficient training and test-time scaling (Video-RTS, LLoVi); (3) **long-horizon memory and verifiable reasoning**, with answers grounded in evidence (EgoMem-Reason, Multimodal Fact-Level Attribution).

Honor

Bloomberg Data Science Ph.D. Fellowship Award — New York City, NY
Bloomberg L.P. — March. 2026

- Recognized for my research in long context multimodal understanding and multimodal agents.

Work Experience

UNC-NLP — Chapel Hill, NC
Graduate Research Assistant — Aug. 2022 - Present

- Advised by Prof. Mohit Bansal.
- Research on agentic and efficient long video understanding with LLMs, and on long-horizon memory and verifiable multimodal reasoning.
- Published at CVPR, ICCV, ECCV, NeurIPS, ICML, EMNLP and EACL.

Bloomberg L.P. — New York City, NY
Data Science Fellow and Research Intern — May. 2026 - Aug. 2026

- Advised by Dr. Debanjan Mahata and Dr. Daniel Preoţiuc-Pietro.
- Worked on large-scale multimodal document understanding.

Salesforce AI Research — Palo Alto, CA
Research Intern — May. 2025 - Feb. 2026

- Advised by Dr. Juan Carlos Niebles (manager), Dr. Michael S. Ryoo, Dr. Junnan Li and Dr. Honglu Zhou.
- Worked on the agentic long video understanding.
- Published “Active Video Perception: Iterative Evidence Seeking for Agentic Long Video Understanding” in CVPR 2026.

Meta FAIR Perception — Menlo Park, CA
Research Intern — May. 2024 - Dec. 2024

- Advised by Dr. Ronghang Hu (manager), Dr. Christoph Feichtenhofer, Dr. Po-Yao (Bernie) Huang and Dr. Daniel Bolya.
- Worked on the next generation video foundation encoders.
- Worked on long video encoder training and zero-shot application for short video encoder on long video tasks.

Amazon Alexa AI — Seattle, WA
Applied Scientist Intern — May. 2023 - Oct. 2023

- Advised by Dr. Heba Elfardy (manager), Dr. Kevin Small, Dr. Markus Dreyer
- Improved the multimodal large language model’s ability to retrieve both visual and textual outputs.
- Published the paper named ”Unified Embeddings for Multimodal Retrieval via Frozen LLMs” in EACL2024 (findings).

Tsinghua University Beijing, China
Research Intern, advised by Prof. Jingjing Liu Oct. 2021 - May. 2022

- Worked on a project to improve the visual question answering ability of the MLLMs.
- Also work on the project that focuses on Vehicle-Infrastructure Cooperative 3D Object Detection.

Papers

2026 **Searching Videos as Trees: Self-Correcting Agents for Grounded Long Video QA** Code
Ce Zhang, **Ziyang Wang**, Yulu Pan, Oluwatumininu Oguntola, Pranav Wagh, Qiyu Wu, Hiromi Wakaki, Mohit Bansal, Gedas Bertasius
NeurIPS 2026

2026 **EgoMemReason: A Memory-Driven Reasoning Benchmark for Long-Horizon Egocentric Video Understanding** Code
Ziyang Wang*, Yue Zhang*, Shoubin Yu, Ce Zhang, Zengqi Zhao, Jaehong Yoon, Hyunji Lee, Gedas Bertasius, Mohit Bansal
COLM 2026

2026 **Evidence-Backed Video Question Answering** Code
Shijie Wang, Honglu Zhou, **Ziyang Wang**, Ran Xu, Caiming Xiong, Silvio Savarese, Chen Sun, Juan Carlos Niebles
ECCV 2026

2026 **Multimodal Fact-Level Attribution for Verifiable Reasoning**
David Wan, Han Wang, **Ziyang Wang**, Elias Stengel-Eskin, Hyunji Lee, Mohit Bansal
ICML 2026

2026 **Active Video Perception: Iterative Evidence Seeking for Agentic Long Video Understanding** Project
Ziyang Wang, Honglu Zhou, Shijie Wang, Junnan Li, Caiming Xiong, Silvio Savarese, Mohit Bansal, Michael S. Ryoo, Juan Carlos Niebles
CVPR 2026 Findings

2026 **SiLVR: A Simple Language-based Video Reasoning Framework** Code
Ce Zhang*, Yan-Bo Lin*, **Ziyang Wang**, Mohit Bansal, Gedas Bertasius
TMLR 2026

2026 **TimeRefine: Temporal Grounding with Time Refining Video LLM**
Xizi Wang, Feng Cheng, **Ziyang Wang**, Huiyu Wang, Md Mohaiminul Islam, Lorenzo Torresani, Mohit Bansal, Gedas Bertasius, David Crandall
WACV 2026

2025 **Video-RTS: Rethinking Reinforcement Learning and Test-Time Scaling for Efficient and Enhanced Video Reasoning** Code
Ziyang Wang*, Jaehong Yoon*, Shoubin Yu, Md Mohaiminul Islam, Gedas Bertasius, Mohit Bansal
EMNLP 2025 Main

2025 **MEXA: Towards General Multimodal Reasoning with Dynamic Multi-Expert Aggregation** Code
Shoubin Yu*, Yue Zhang*, **Ziyang Wang**, Jaehong Yoon, Mohit Bansal
EMNLP 2025 Findings

2025 **VideoTree: Adaptive Tree-based Video Representation for LLM Reasoning on Long Videos** Code
Ziyang Wang*, Shoubin Yu*, Elias Stengel-Eskin*, Jaehong Yoon, Feng Cheng, Gedas Bertasius, Mohit Bansal
CVPR 2025

2025 **DAM: Dynamic Adapter Merging for Continual Video QA Learning** Code
Feng Cheng*, **Ziyang Wang***, Yi-Lin Sung, Yan-Bo Lin, Mohit Bansal, Gedas Bertasius
WACV 2025

2024 **Unified Embeddings for Multimodal Retrieval via Frozen LLMs**
Ziyang Wang, Heba Elfardy, Markus Dreyer, Kevin Small, Mohit Bansal
EACL 2024 Findings

2024 **A Simple LLM Framework for Long-Range Video Question-Answering** Code

Ce Zhang, Taixi Lu, Md Mohaiminul Islam, **Ziyang Wang**, Shoubin Yu, Mohit Bansal, Gedas Bertasius
EMNLP 2024 Main

2023 **Unified Coarse-to-Fine Alignment for Video-Text Retrieval** Code
Ziyang Wang, Yi-Lin Sung, Feng Cheng, Gedas Bertasius, Mohit Bansal
ICCV 2023

2022 **Language-Augmented Pixel Embedding for Generalized Zero-Shot Learning**
Ziyang Wang, Yunhao Gou, Jingjing Li, Lei Zhu, Heng Tao Shen
TCSVT 2022

2021 **Region Semantically Aligned Network for Zero-Shot Learning**
Ziyang Wang*, Yunhao Gou*, Jingjing Li, Yu Zhang, Yang Yang
CIKM 2021 (Long Oral)

Service

Transformers for Vision (T4V) Workshop @ CVPR

Co-organizer (2025, 2026) and Program Committee Member (2023, 2024) 2023 - 2026

- Co-organized the workshop at CVPR 2025 and 2026.
- Served on the program committee in 2023 and 2024, reviewing submissions and supporting workshop logistics.

Conferences and Journals

Reviewer

- CVPR, ECCV, WACV, ICLR, COLM, ACL Rolling Review, IJCV.