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. candidate in Computer Science, advised by Prof. Mohit Bansal. — *Aug. 2022 - Exp. May. 2027*

University of Electronic Science and Technology of China — *Chengdu, China*

Bachelor of Engineer, Major: Software Engineering. Overall GPA: 3.92. — *Sept. 2018 - Jun. 2022*

Research Interest

My research interest is video-language understanding and multimodal AI. Particularly, I am interested in the challenge of reasoning over long and complex videos.

Papers

2025	**SiLVR: A Simple Language-based Video Reasoning Framework** Ce Zhang*, Yan-Bo Lin*, **Ziyang Wang**, Mohit Bansal, Gedas Bertasius	*Preprint*
2025	**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	*Preprint*
2025	**Video-RTS: Rethinking Reinforcement Learning and Test-Time Scaling for Efficient and Enhanced Video Reasoning** **Ziyang Wang***, Jaehong Yoon*, Shoubin Yu, Md Mohaiminul Islam, Gedas Bertasius, Mohit Bansal, ***EMNLP 2025 (Main)***	*Suzhou, China*
2025	**MEXA: Towards General Multimodal Reasoning with Dynamic Multi-Expert Aggregation** Shoubin Yu*, Yue Zhang*, **Ziyang Wang**, Jaehong Yoon, Mohit Bansal, ***EMNLP 2025 (Findings)***	*Suzhou, China*
2025	**VideoTree: Adaptive Tree-based Video Representation for LLM Reasoning on Long Videos** **Ziyang Wang***, Shoubin Yu*, Elias Stengel-Eskin*, Jaehong Yoon, Feng Cheng, Gedas Bertasius, Mohit Bansal, ***CVPR 2025***	*Nashville, USA*
2025	**DAM: Dynamic Adapter Merging for Continual Video QA Learning** Feng Cheng*, **Ziyang Wang***, Yi-Lin Sung, Yan-Bo Lin, Mohit Bansal, Gedas Bertasius, ***WACV 2025***	*Tucson, USA*
2024	**Unified Embeddings for Multimodal Retrieval via Frozen LLMs** **Ziyang Wang**, Heba Elfardy, Markus Dreyer, Kevin Small, Mohit Bansal, ***EACL 2024 (Findings)***	*Malta*
2024	**A Simple LLM Framework for Long-Range Video Question-Answering** Ce Zhang, Taixi Lu, Md Mohaiminul Islam, **Ziyang Wang**, Shoubin Yu, Mohit Bansal, Gedas Bertasius, ***EMNLP 2024 (Main)***	*Maimi, USA*
2023	**Unified Coarse-to-Fine Alignment for Video-Text Retrieval** **Ziyang Wang**, Yi-Lin Sung, Feng Cheng, Gedas Bertasius, Mohit Bansal, ***ICCV 2023***	*Paris, France*
2022	**Language-Augmented Pixel Embedding for Generalized Zero-Shot Learning** **Ziyang Wang**, Yunhao Gou, Jingjing Li, Lei Zhu, Heng Tao Shen, ***IEEE Transactions on Circuits and Systems for Video Technology (TCSVT 2022)***	*Journal*
2021	**Region Semantically Aligned Network for Zero-Shot Learning** **Ziyang Wang***, Yunhao Gou*, Jingjing Li, Yu Zhang, Yang Yang, ***CIKM 2021, long oral***	*Virtual*

Work Experience

UNC-NLP — *Chapel Hill, NC*

Graduate Research Assistant, advised by Prof. Mohit Bansal, also work closely with Prof. Gedas Bertasius — *Aug. 2022 - PRESENT*

- Built models to tackle challenging video understanding tasks like retrieval and long video understanding.
- Published several academic papers in CV/NLP conferences including ICCV, EACL, EMNLP.

Salesforce AI Research

Palo Alto, CA

Research Intern

May. 2025 - Dec. 2025

- Advised by Dr. Juan Carlos Niebles (manager), Dr. Michael S. Ryoo, Dr. Junnan Li and Dr. Honglu Zhou.
- Work on the agentic video understanding via tool calling and efficient video inference.

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.
- Work on the next generation video foundation model.
- Work 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.

Service

Top CV/NLP conferences

Reviewer

- Engaged in the peer-review process in ECCV, ICLR, CVPR, COLM, WACV, ACL Rolling Reviews, IJCV.

Transformers for Vision (T4V) workshop @ CVPR 2023 & 2024 & 2025

Organizer and Program Commitee

- Co-organize the T4V workshop @ CVPR 2025.
- Worked as a program committee member to help the reviewing process and other aspect of the workshop (2023 and 2024).