

Ziyang Wang

Homepage: ziyangw2000.github.io | [Google Scholar](#) | Email: ziyangw@cs.unc.edu

EDUCATION

The University of North Carolina at Chapel Hill

Chapel Hill, NC

Ph.D. student in Computer Science department

Sept 2022 - May 2027 (Expected)

- Advisor: [Prof. Mohit Bansal](#)
- Research Interest: multimodal learning, video-language understanding

University of Electronic Science and Technology of China

Chengdu, China

Bachelor of Software Engineering

Sept 2018 - July 2022

- Graduate from Elite Program
- GPA: 3.92/4

RESEARCH INTEREST

Multimodal learning, vision-language understanding

- In general, I am interested in the fundamental challenges in multimodal machine learning. Particularly, I am enthusiastic about video-language understanding. Recently, I have been working on multiple topics on video-language alignment, video LLM, and neuro-symbolic methods to accomplish complex and explainable video understanding.

RESEARCH EXPERIENCE

Research Assistant

September 2022 – present

UNC Chapel Hill

Chapel Hill, NC

- Advised by [Prof. Mohit Bansal](#), also work closely with [Prof. Gedas Bertasius](#)
- Have a broad interest in multi-modal learning, video-language understanding

Applied Scientist Intern

May 2023 – Oct 2023

Amazon Alexa AI

Seattle, WA

- Work with [Heba Elfardy](#), [Kevin Small](#), [Markus Dreyer](#) on multimodal retrieval

Research Intern

October 2021 – May 2022

Tsinghua University

Beijing, China

- Work with [Prof. Jingjing Liu](#) on multimodality

Research Assistant

January 2021 – September 2021

UESTC

Chengdu, China

- Worked with [Prof. Jingjing Li](#) on transfer learning

PUBLICATION

1. **Ziyang Wang**, Heba Elfardy, Markus Dreyer, Kevin Small, Mohit Bansal. Unified Embeddings for Multimodal Retrieval via Frozen LLMs. **EACL2024** Findings.
2. Ce Zhang*, Taixi Lu*, Md Mohaiminul Islam, **Ziyang Wang**, Shoubin Yu, Mohit Bansal, Gedas Bertasius. A Simple LLM Framework for Long-Range Video Question-Answering. Arxiv preprint. [Link](#)
3. **Ziyang Wang**, Yi-Lin Sung, Feng Cheng, Gedas Bertasius, Mohit Bansal. Unified Coarse-to-Fine Alignment for Video-Text Retrieval. In Proceedings of the IEEE/CVF International Conference on Computer Vision (**ICCV 2023**) [Link](#)
4. **Ziyang Wang**, Yunhao Gou, Jingjing Li, Lei Zhu, Heng Tao Shen. Language-Augmented Pixel Embedding for Generalized Zero-shot Learning. IEEE Transactions on Circuits and Systems for Video Technology (**TCSVT 2022**) [Link](#)
5. **Ziyang Wang***, Yunhao Gou*, Jingjing Li, Yu Zhang, and Yang Yang. Region Semantically Aligned Network for Zero-Shot Learning. In Proceedings of the 30th ACM International Conference on Information and Knowledge Management (**CIKM 2021 oral**) [Link](#)

SERVICE

1. Reviewer: ACL Rolling Review (ARR), CIKM, ACM Multimedia
2. Program Committee: [T4V @ CVPR 2023](#)