

YUXUAN WANG

Email: flagwyx@gmail.com

Beijing, China

[Google Scholar](#) \diamond [Github](#) \diamond [Homepage](#)

EDUCATION

Peking University

June 2023

M.S. in Data Science (Computer Science and Technology)

Northeastern University

June 2019

B.E. in Internet of Things Engineering

WORKING EXPERIENCE

Qwen Team, Alibaba Inc.

April 2025 - Present

Leader: Dr. Jin Xu and Junyang Lin

Research Engineer

- Qwen-omni series

Beijing Institute for General Artificial Intelligence

July 2023 - Mar. 2025

Leader: Dr. Zilong Zheng and Prof. Songchun Zhu

Research Engineer

- Large omni language model (multiplexing omni model [1])
- Large video language model (efficient temporal grounding [2], long video modeling [3])
- Multimodal dialogue: (scene-topic modeling [4], visual narrative reasoning [5], speaker identification [6])
- Logographic language modeling (Chinese pretraining [7])

INTERN EXPERIENCE

UC. Santa Cruz

Mar. 2024 - Sep. 2024

Supervisors: Prof. Cihang Xie

Part-time Research Intern

- Hallucination and long-context modeling in video-language modeling [3], [8]

Johns Hopkins University

May 2023 - Sep. 2023

Supervisors: Dr. Zhuowan Li and Prof. Alan Yuille

Summer Intern

- Compositional visual reasoning/multimodal agent with verifiers [9]

Peking University

Jan. 2021 - June 2023

Supervisors: Prof. Dongyan Zhao

Research Assistant

- Interpretable multimodal dialogue (multimodal semantic graph [10], auditable intermedia NMN [11])

PUBLICATIONS

- [1] **Yuxuan Wang**, Y. Wang, B. Chen, T. Wu, D. Zhao, and Z. Zheng, "OmniMMI: A Comprehensive Multi-modal Interaction Benchmark in Streaming Video Contexts," *CVPR*, 2025.
- [2] **Yuxuan Wang**, Y. Wang, P. Wu, J. Liang, D. Zhao, and Z. Zheng, "Efficient Temporal Extrapolation of Multimodal Large Language Models with Temporal Grounding Bridge," *EMNLP*, 2024.
- [4] **Yuxuan Wang**, Z. Zheng, X. Zhao, J. Li, Y. Wang, and D. Zhao, "VSTAR: A Video-grounded Dialogue Dataset for Situated Semantic Understanding with Scene and Topic Transitions," *ACL*, 2023.

- [6] Y. Wang, X. Meng, **Yuxuan Wang**, J. Liang, Q. Liu, and D. Zhao, “Friends-MMC: A Dataset for Multi-modal Multi-party Conversation Understanding,” *AAAI*, 2025.
- [7] **Yuxuan Wang**, J. Wang, D. Zhao, and Z. Zheng, “Shuo Wen Jie Zi: Rethinking Dictionaries and Glyphs for Chinese Language Pre-training,” *Findings of ACL*, 2023.
- [9] **Yuxuan Wang**, A. Yuille, Z. Li, and Z. Zheng, “ExoViP: Step-by-step Verification and Exploration with Exoskeleton Modules for Compositional Visual Reasoning,” *COLM*, 2024.
- [10] X. Zhao*, **Yuxuan Wang***, C. Tao, C. Wang, and D. Zhao, “Collaborative Reasoning on Multi-Modal Semantic Graphs for Video-Grounded Dialogue Generation,” in *Findings of EMNLP*, 2022.
- [11] Y. Wang, **Yuxuan Wang**, K. Chen, and D. Zhao, “STAIR: Spatial-Temporal Reasoning with Auditable Intermediate Results for Video Question Answering,” *AAAI*, 2024.

PREPRINTS

- [3] **Yuxuan Wang**, C. Xie, Y. Liu, and Z. Zheng, “VideoLLaMB: Long Video Understanding with Recurrent Memory Bridges,” *Preprint*, 2024.
- [5] J. Wang*, **Yuxuan Wang***, D. Zhao, and Z. Zheng, “MoviePuzzle: Visual Narrative Reasoning through Multimodal Order Learning,” *Preprint*, 2023.
- [8] **Yuxuan Wang**, Y. Wang, D. Zhao, C. Xie, and Z. Zheng, “VideoHalluciner: Evaluating Intrinsic and Extrinsic Hallucinations in Large Video-Language Models,” *Preprint*, 2024.

AWARDS

Competition

- The Chinese Mathematics Competition (CMC), **First Prize**, 2018
- COMAP’S Mathematical Contest in Modeling (MCM), **Honorable Mention**, 2017

Scholarship

- Academic Scholarship, awarded by Peking University, 2023
- The First/Second/Third Prize Scholarship, awarded by Northeastern University, 2016-2018

SERVICES

Organizer

- NLPCC 2023 Shared Task 10
- NLPCC 2022 Shared Task 4

Reviewer

- ACL ARR 2024 December - now, Area Chair
- ACL ARR 2023 - now (Great Reviews Mention) & CVPR 2025

Teaching Assistant

- Course “Introduction to Computing (2021Fall)” at Peking University

SKILLS/HOBBIES

Programming Languages

Python, C/C++

Machine Learning Tools

Pytorch, Tensorflow