

Zun Wang

☎ (+1) 984-369-3714 • ✉ zunwang@cs.unc.edu • 🌐 zunwang1.github.io

🎓 [Zun's Google Scholar](#) • 🌐 [wz0919](#)

Education

University of North Carolina, Chapel Hill, Chapel Hill, North Carolina **Jul 2024 – Present**

Ph.D. in Computer Science

Advisor: [Prof. Mohit Bansal](#)

The Australian National University, Canberra, Australia **Jul 2020 – Jul 2023**

Master of Machine Learning and Computer Vision

Advisor: [Prof. Stephen Gould](#)

Graduated with [Postgraduate Medal for Academic Excellence](#) (**Top 1** out of the CECC college)

University of Science and Technology of China, Hefei, China **Sep 2015 – Jun 2019**

Bachelor of Mathematics and Applied Mathematics

Specialization in Probability and Statistics

Research Interest

My ultimate research goal focuses on building **multimodal**, **generative**, and **embodied** agents via both model-centric and data-centric approaches, with my current research interests lying in

- Multimodal Understanding and Generation
- Scalable Learning for Embodied Agents
- Multimodal Data Generation and Curation

Papers

* denotes equal contribution

1. **DreamRunner: Fine-Grained Storytelling Video Generation with Retrieval-Augmented Motion Adaptation** [**Video Generation**]
Preprint, Submitted to CVPR 2025
Zun Wang, Jialu Li, Han Lin, Jaehong Yong, Mohit Bansal
2. **Bootstrapping Language-Guided Navigation Learning with Self-Refining Data Flywheel** [**Embodied AI**][**Synthetic Data**][**Multimodal AI**]
ICLR 2025
Zun Wang, Jialu Li, Yicong Hong, Songze Li, Kunchang Li, Shoubin Yu, Yi Wang, Yu Qiao, Yali Wang, Mohit Bansal, Limin Wang
3. **Scaling Data Generation in Vision-and-Language Navigation** [**Embodied AI**][**Synthetic Data**]
ICCV 2023 (Oral)
Zun Wang*, Jialu Li*, Yicong Hong*, Yi Wang, Qi Wu, Mohit Bansal, Stephen Gould, Hao Tan, Yu Qiao
4. **Bridging the Gap Between Learning in Discrete and Continuous Environments for Vision-and-Language Navigation** [**Embodied AI**]
CVPR 2022
Yicong Hong*, Zun Wang*, Qi Wu, Stephen Gould
5. **Vision-and-Language Navigation Today and Tomorrow: A Survey in the Era of Foundation Models** [**Survey**][**Embodied AI**][**Multimodal AI**]
TMLR 2024
Yue Zhang*, Ziqiao Ma*, Jialu Li*, Yanyuan Qiao*, Zun Wang*, Joyce Chai, Qi Wu, Mohit Bansal, Parisa Kordjamshidi

6. **Navgpt-2: Unleashing Navigational Reasoning Capability for Large Vision-Language Models** [Embodied AI][Multimodal AI]
ECCV 2024
Gengze Zhou, Yicong Hong, **Zun Wang**, Xin Eric Wang, Qi Wu
7. **ETPNav: Evolving Topological Planning for Vision-Language Navigation in Continuous Environments** [Embodied AI]
TPAMI 2024
Dong An, Hanqing Wang, Wenguan Wang, **Zun Wang**, Yan Huang, Keji He, Liang Wang
8. **SAME: Learning Generic Language-Guided Visual Navigation with State-Adaptive Mixture of Experts** [Embodied AI]
Preprint, Under Review
Gengze Zhou, Yicong Hong, **Zun Wang**, Chongyang Zhao, Mohit Bansal, Qi Wu
9. **MVBench: A Comprehensive Multi-modal Video Understanding Benchmark** [Multimodal AI]
CVPR 2024 (*Highlight*)
Kunchang Li, Yali Wang, Yinan He, Yizhuo Li, Yi Wang, Yi Liu, **Zun Wang**, Jilan Xu, Guo Chen, Ping Luo, Limin Wang, Yu Qiao
10. **InternVideo: General Video Foundation Models via Generative and Discriminative Learning** [Multimodal AI]
Preprint, Under Review
Yi Wang, ..., **Zun Wang**, ..., Yali Wang, Limin Wang, Yu Qiao
11. **InternVideo2: Scaling Video Foundation Models for Multimodal Video Understanding** [Multimodal AI]
ECCV 2024
Yi Wang, ... , **Zun Wang**, ..., Yu Qiao, Yali Wang, Limin Wang

Research Experience

University of North Carolina, Chapel Hill, Chapel Hill, North Carolina **Aug 2024 - Present**

Ph.D. Student – Advised by Prof. Mohit Bansal

- Topics: Video Generation. Multimodal Learning.
- Building a powerful storytelling video generation model via LLM planning, RAG and attention/lora decomposition, could be generalized to compositional text-to-video generation.
- Paper under review at CVPR 2025.

Shanghai AI Laboratory, GeneralVideo Group, Shanghai, China **Sept 2023 - Aug 2024**

Research Intern – Led by Prof. Limin Wang and Prof. Yu Qiao

- Topics: Language-guided Navigation.
- Building a strong language-guided navigation agent that **surpasses human performance** on the R2R dataset + SoTAs on eight VLN tasks with the proposed navigator-generator-data self-refining flywheel.
- paper accepted to ICLR 2025.

Shanghai AI Laboratory, GeneralVideo Group, Shanghai, China **Mar 2022 – Mar 2023**

Research Intern – Led by Dr. Yi Wang and Prof. Yu Qiao

- Topics: Video Understanding, Vision-and-Language Navigation.
- Improving VLN SoTAs by a large margin via large-scale data generation (accepted as Oral at ICCV 2023).
- Enhancing navigation agents with video foundation models.
- 2022 REVERIE VLN Challenge winner among 50+ international teams.

Sensetime & Shanghai AI Laboratory, AGI Group, Shanghai, China **Dec 2021 – Mar 2022**

Research Intern – Led by Dr. Jing Shao

- Won the RxR-Habitat Competition in Embodied AI Workshop at CVPR 2022.

The Australian National University, Canberra, Australia **Mar 2021 – Dec 2021**

Research Student – Advised by Dr. Yicong Hong and Prof. Stephen Gould

- Projects on discrete-to-continuous Vision-and-Language Navigation (accepted by CVPR 2022).

Awards

Postgraduate Medal for Academic Excellence, ANU **Jul 2023**

Top 1 postgraduate student out from the College of Engineering, Computing & Cybernetics, university-wide prize, highest honor for master students

Chancellor's Letter of Commendation, ANU **Nov 2022**

Awarded to full-grade students over the past year

CSIG 2022, REVERIE Challenge channel 1 & channel 2, 1st places, 14800 USD Bonus **Aug 2022**

Zun Wang, Yi Wang, Yinan He, Yu Qiao

CVPR 2022 Embodied AI Workshop, RxR-Habitat Competition, 1st place **Jun 2022**

Dong An*, **Zun Wang***, Yangguang Li, Yi Wang, Yicong Hong, Yan Huang, Liang Wang, Jing Shao

Academic Activities

- Conference Reviewer: CVPR2023, EMNLP2023, CVPR2024, EMNLP2024, CVPR2025, ACL2025
- Winner invited talk at Competition Forum @ CSIG2022
- Winner invited talk at Embodied AI Workshop @ CVPR2022

References

Mohit Bansal

John R. & Louise S. Parker Distinguished Professor

University of North Carolina, Chapel Hill

mbansal@cs.unc.edu

Stephen Gould

Professor

The Australian National University

stephen.gould@anu.edu.au

Yu Qiao

Lead Scientist

Shanghai AI Laboratory

qiaoyu@pjlab.org.cn

Limin Wang

Professor

Nanjing University

lmwang.nju@gmail.com

Yi Wang

Research Scientist

Shanghai AI Laboratory

wangyi@pjlab.org.cn

Yicong Hong

Research Scientist

Adobe Research

mr.yiconghong@gmail.com