Hanyu Wang

+1 (240) 413-4808 | hywang66@umd.edu | hywang66.github.io

Education

University of Maryland – College Park — College Park, MD, US
Ph.D. in Computer Science; Advisor: Abhinav Shrivastava — *Aug. 2020 – Present*

University of Maryland – College Park — College Park, MD, US
M.S. in Computer Science; GPA: 3.94/4.0 — *Aug. 2018 – May 2020*

Xi'an Jiaotong University — Xi'an, China
B.Eng. in Computer Science and Technology; GPA: 3.9/4.0 — *Sept. 2014 – May 2018*

Research Interests

My research focuses on **Computer Vision** and **Generative AI**, with an emphasis on **visual content generation** under various conditions such as class labels, reference images, and text descriptions, as well as in unconditional settings.
My long-term goal is to build multimodal **foundation models that unify the understanding and generation** of different data types, enabling ***cross-modal learning*** and ***mutual enhancement across modalities***.

Internship Experience

Research Intern, **ByteDance US** — **2025**

- Mentors: Hao Chen and Lu Jiang
- Exploring unified multimodal LLMs for both understanding and generation.

Research Scientist Intern, **Meta** — **2023**

- Mentor: Chenyang Zhang
- Explored zero-shot text-to-video generation.

Research Intern, **Snap Inc.** — **2021 - 2022**

- Mentors: Pengxiang Wu, Chen Wang, Kevin Dela Rosa
- Explored the problem of multimodality-guided image style transfer.

Research Intern, **Institute of Automation, CAS** — **2017 - 2018**

- Mentors: Jianwei Guo, Dong-Ming Yan, Xiaopeng Zhang
- Explored learning-based 3D keypoint descriptor.

Publications

Jiaming Han, Hao Chen, Yang Zhao, **Hanyu Wang**, Qi Zhao, Ziyan Yang, Hao He, Xiangyu Yue, Lu Jiang
Vision as a Dialect: Unifying Visual Understanding & Generation via Text-Aligned Representations.
arXiv preprint 2025. Project Page / arXiv / Code / Demo

Hanyu Wang, Saksham Suri, Yixuan Ren, Hao Chen, Abhinav Shrivastava
LARP: Tokenizing Videos with a Learned Autoregressive Generative Prior.
ICLR 2025 (Oral). Project Page / arXiv / Code

Hanyu Wang, Pengxiang Wu, Kevin S Dela Rosa, Chen Wang, Abhinav Shrivastava
Multimodality-guided Image Style Transfer using Cross-modal GAN Inversion.
WACV 2024. Project Page / arXiv / Code

Haoyue Tang, Tian Xie, Aosong Feng, **Hanyu Wang**, Chenyang Zhang, Yang Bai
Solving General Noisy Inverse Problem via Posterior Sampling: A Policy Gradient Viewpoint.
AISTATS 2024. arXiv / Code

Nirat Saini*, **Hanyu Wang***, Archana Swaminathan, Vinoj Jayasundara, Bo He, Kamal Gupta, Abhinav Shrivastava
Chop & Learn: Recognizing and Generating Object-State Compositions.
ICCV 2023. Project Page / arXiv / Data

Bo He, Xitong Yang, **Hanyu Wang**, Zuxuan Wu, Hao Chen, Shuaiyi Huang, Yixuan Ren, Ser-Nam Lim, Abhinav Shrivastava
Towards Scalable Neural Representation for Diverse Videos.
CVPR 2023. Project Page / arXiv / Code

Shishira R Maiya*, Sharath Girish*, Max Ehrlich, **Hanyu Wang**, Kwot Sin Lee, Patrick Poirson, Pengxiang Wu, Chen Wang, Abhinav Shrivastava
NIRVANA: Neural Implicit Representations of Videos with Adaptive Networks and Autoregressive Patch-wise Modeling.
CVPR 2023. Project Page / arXiv

Hanyu Wang, Kamal Gupta, Larry Davis, Abhinav Shrivastava
Neural Space-filling Curves.
ECCV 2022. Project Page / arXiv / Code

Hao Chen, Bo He, **Hanyu Wang**, Yixuan Ren, Ser Nam Lim, Abhinav Shrivastava
NeRV: Neural Representations for Videos.
NeurIPS 2021. Project Page / arXiv / Code

Jianwei Guo, **Hanyu Wang**, Zhanglin Cheng, Xiaopeng Zhang, Dong-Ming Yan
Learning Local Shape Descriptors for Computing Non-rigid Dense Correspondence.
Computational Visual Media 2020. Paper

Hanyu Wang*, Jianwei Guo*, Dong-Ming Yan, Weize Quan, Xiaopeng Zhang
Learning 3D Keypoint Descriptors for Non-Rigid Shape Matching.
ECCV 2018. Paper

Academic Services

ICCV 2025, CVPR 2025, WACV 2025, ECCV 2024, CVPR 2024, WAVC 2024, ICCV 2023, CVPR 2023, CVPR 2022, ECCV 2022 reviewer.

Teaching Experiences

Head Teaching Assistant, CMSC472, Introduction to Deep Learning, Fall 2024

Teaching Assistant, CMSC828I, Advanced Techniques in Visual Learning & Recognition, Fall 2021

Teaching Assistant, CMSC351, Algorithms, Spring 2021

Teaching Assistant, CMSC320, Introduction to Data Science, Fall 2020

Technical Skills

Programming Languages: Python, C/C++, MATLAB, Java, etc.
Frameworks & Libraries: Transformers, Pytorch, Tensorflow, Numpy, Jupyter, etc.