

Hanyu Wang

+1 (240) 413-4808 | hanyuw66@gmail.com | hywang66.github.io

RESEARCH FOCUS

Model training and data curation for multimodal and embodied AI; visual generation and video representation learning; recursive self-improvement for embodied intelligence.

EXPERIENCE

Research Scientist, **Luma AI**

Feb. 2026 – Present

- **Founding Researcher, Project Opal**, Luma’s open physical AI initiative.
 - * Architected and built Opal’s bimanual YAM robot stack from the ground up, including arm control, camera management, synchronized recording, policy runtime, and a custom remote-serving protocol; integrated and demonstrated Opal’s first diffusion-based VLA policy.
 - * Led the development of fully automated full-episode and subtask annotation pipelines for VLA/WAM models spanning tasks from tabletop object rearrangement to biochemical laboratory procedures, combining rule- and signal-based segmentation with VLM classification; developed a human-in-the-loop auto-research system that uses labeled examples to iteratively improve the pipelines without per-episode correction.
 - * Led the development of a real-robot agent harness that lets coding agents modify policies and prompts within fixed goals, rewards, evaluations, and task definitions, with humans needed only for harness upgrades and rare resets; designed a data-flywheel strategy that uses interpretable, code-generated actions to sample controllable spatiotemporal trajectories for VLA/WAM training.
- **Core Contributor, Uni-1 Series**, Luma’s multimodal reasoning and image generation model.
 - * Built a multi-stage data pipeline for zero-shot subject-driven generation, mining identity-consistent examples from noisy data with VLM/DINO scoring and automatically generating targeted captions using VLM prompts iteratively optimized through a human-in-the-loop auto-research process.
 - * Led continual pre- and post-training of a second-pass text-rendering refinement model; combined SDEdit and on-policy rollouts to generate failure cases and annotated the corrective text edits needed to recover intended renderings, achieving a 90%+ preference rate over Uni-1.0 in paired human evaluations.

Research Intern, **ByteDance US**

2025

- Advanced unified autoregressive MLLMs for joint visual understanding and generation by co-designing model components and text-aligned visual representations.

Research Scientist Intern, **Meta**

2023

- Developed zero-shot text-to-video generation methods using motion-based latent optimization and pretrained image-editing models, improving motion smoothness, semantic control, and visual fidelity.

EDUCATION

University of Maryland, College Park

Ph.D. in Computer Science; Advisor: Abhinav Shrivastava

M.S. in Computer Science

College Park, MD, US

Aug. 2020 – Dec. 2025

Aug. 2018 – May 2020

Xi’an Jiaotong University

B.Eng. in Computer Science and Technology

Xi’an, China

Sept. 2014 – May 2018

SELECTED PUBLICATIONS

Yixuan Ren, **Hanyu Wang**, Hao Chen, Bo He, Abhinav Shrivastava
NeRV-Diffusion: Diffuse Implicit Neural Representation for Video Synthesis.
ICLR 2026. [Project Page](#) / [arXiv](#)

Hanyu Wang*, Jiaming Han*, Ziyang Yang, Qi Zhao, Shanchuan Lin, Xiangyu Yue, Abhinav Shrivastava, Zhenheng Yang, Hao Chen
Growing Visual Generative Capacity for Pre-Trained MLLMs.
ICLR Workshop 2026. [Project Page](#) / [arXiv](#)

Jiaming Han, Hao Chen, Yang Zhao, **Hanyu Wang**, Qi Zhao, Ziyang Yang, Hao He, Xiangyu Yue, Lu Jiang
Vision as a Dialect: Unifying Visual Understanding & Generation via Text-Aligned Representations.
NeurIPS 2025. [Project Page](#) / [arXiv](#) / [Code](#)

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](#)

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](#)

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](#)

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

* Equal contribution.

ACADEMIC SERVICE

Reviewer: ICLR 2025; ICCV 2023, 2025; CVPR 2022–2025; WACV 2024–2025; ECCV 2022, 2024.