Xingrui WANG

(+1) 2136757298 | xwang378@jhu.edu| 3100 Wyman Park Dr, Baltimore, MD 21211

Education Background

Whiting School of Engineering, Johns Hopkins University; **Baltimore, MD**
Ph.D. in Computer Science; **GPA**: 4.00 / 4.00; *08/2023- Present*
Advisor: Alan L. Yuille.

Viterbi School of Engineering, University of Southern California **Los Angeles, CA**
M.S. in Applied Data Science; **GPA**: 3.92 / 4.00 *08/2021- 05/2023*

School of Statistics, Renmin University of China **Beijing, China**
B.S. in Statistics; Minor in Data Science; **GPA**: 87.04 / 100 *09/2017- 07/2021*

Publications

[1] **PulseCheck457: A Diagnostic Benchmark for Comprehensive Spatial Reasoning of Large Multimodal Models.**
Xingrui Wang, Wufei Ma, Tiezheng Zhang, Celso M de Melo, Jieneng Chen, Alan Yuille.
*Conference on Computer Vision and Pattern Recognition (**CVPR**) 2025*
TL; DR: A benchmark for comprehensive 6D spatial reasoning of large vision language models.

[2] **Compositional 4D Dynamic Scenes Understanding with Physics Priors for Video Question Answering**
Xingrui Wang, Wufei Ma, Angtian Wang, Shuo Chen, Adam Kortylewski, Alan Yuille.
*International Conference on Learning Representations (**ICLR**) 2025.*
TL; DR: A video question answering benchmark and model for 4D physical properties of objects from 3D space.

[3] **3D-Aware Visual Question Answering about Parts, Poses and Occlusions**.
Xingrui Wang, Wufei Ma, Zhuowan Li, Adam Kortylewski, Alan Yuille. 3D-Aware Visual Question Answering about Parts, Poses and Occlusions.
*Advances in Neural Information Processing Systems (**NeurIPS**), 2023*
TL; DR: A benchmark and model for 3D scene understanding in vision question answering, particularly parts, poses, and occlusions.

[4] **Super-CLEVR: A Virtual Benchmark to Diagnose Domain Robustness in Visual Reasoning**
Zhuowan Li, **Xingrui Wang**, Elias Stengel-Eskin, Adam Kortylewski, Wufei Ma, Benjamin Van Durme, Alan Yuille.
*Conference on Computer Vision and Pattern Recognition (**CVPR**) **Highlight**, 2023*
TL; DR: A diagnosis dataset analyzes the factors of domain shift in vision question answering models.

[5] **Contributions of Shape, Texture and Color in Visual Recognition**
Yunhao Ge*, Yao Xiao*, Zhi Xu, **Xingrui Wang**, Laurent Itti.
*European Conference on Computer Vision (**ECCV**)*, 2022
TL; DR: A human-inspired object recognition network which considers the disentangled shape, texture, and color from images.

[6] **Towards Language Hint Attention Reinforcement Learning.**
Haoyu Liu, Yang Liu, **Xingrui Wang**, Hanfang Yang.
IEEE World Congress on Computational Intelligence (***WCCI**), 2022*
TL; DR: A human assisted high-efficient reinforcement learning model.

[7] **Large Scale GPS Trajectory Generation Using Map Based on Two Stage GAN**
Xingrui Wang, Xinyu Liu, Ziteng Lu, Hanfang Yang,
Journal of Data Science. 19(2021), no. 1, 126-141. *DOI 10.6339/21-JDS1004*
TL; DR: A map conditioned GPS trajectory data generation mode with GANs.

Working Experience

Advanced Micro Devices, Inc. | Research Intern *06/2024- present*

- Advisor: Dr. Jiang Liu.
- Research Topic: **Multimodal conditioning video generation.**
- Project Description: Build a video generation diffusion model for dynamical motion conditioned on audio and image input. Evaluate the temporal alignment of given audio and generated video.

Samsung R&D Institute China-Beijing | Research Intern *12/2020- 06/2021*

- Advisor: Dr. Yang Liu
- Research Topic: **Embodied AI; Reinforcement learning.**
- Project Description: (1) Human-Guided Reinforcement Learning**:** Proposed a method that combines language hints with an object template matching module, providing human coarse-grained pre-guided attention to improve the efficiency and performance of the reinforcement learning model. (2) ALFRED benchmark, Embodied AI @ CVPR 2021. Leveraged instance segmentation and depth estimation to ground object positions on the bird's-eye-view obstacle map, generate navigation paths to the grounded objects, and integrate these with language instructions.

Teaching Experiences

University of Southern California

- Course Producer: DSCI 552 - Machine Learning for Data Science

Johns Hopkins University

- Course Producer: DSCI 552 - Machine Learning for Data Science