Hairui Yin

College Park, MD | yinhr@umd.edu | https://www.linkedin.com/in/hairui-yin/

Education

University of Maryland, College Park 09/2024 - 05/2026
Master of Science, Data Science GPA: 4.00/4.00

University of Wisconsin, Madison 01/2023 - 07/2023
Exchange in Computer Science GPA: 3.80/4.00

Shanghaitech University 09/2020 - 06/2024
Bachelor of Engineering, Computer Science and Technology GPA: 3.53/4.00

Research

Physical-based motion video generation 08/2025 - Present
Intern | Advised by Prof. Chuang Gan at University of Massachusetts, Amherst

- Extending physics-based humanoid control frameworks (InterMimic, PHC) to generate motion videos aligned with motion sequences, constructing a novel dataset for video-motion integration.

Structural Causal model based Diffusion 03/2025 - Present
Advised by Prof. Abdirisak Mohamed at University of Maryland, College Park

- Extended existing counterfactual backtracking methods by removing the restrictive assumption of invertible noise–to–causal node mappings.
- Developing a counterfactual backtracking method for image editing with causal consistency, with the aim of improving controllability in diffusion models.

Tool-Oriented Prompt Injection Attacks on LLM Agents 05/2025 - 06/2025
Advised by Researcher Udari Madhushani Sehwag

- Extended and customized existing multi-agent frameworks (AutoGen, AgentDojo) to support tool injection attack scenarios.

Multi-modal data-driven extraction of genealogy images 09/2023 - 05/2024
Research Assistant | Advised by Prof. Haipeng Zhang at Shanghaitech University

- First to build a large-scale genealogy multimodal dataset (2.8TB) enabling sociological analysis, using multimodal combining OCR, Vision Neural Networks, and LLMs.
- Conducted demographic and sociological analysis, visualizing insights to uncover historical trends and patterns.

Professional Experience

Assistant Data Enginner | Glodon - Shanghai, China 01/2024 - 07/2024

- Enhanced camera-based construction site safety monitoring systems by applying advanced object detection models (YOLO, Faster R-CNN), achieving human-level accuracy (≈94%) and significantly improving real-time detection efficiency, which reduced safety incidents in pilot deployments.
- Developed a robust data processing pipeline (noise reduction, normalization, augmentation with rotation, flipping, cropping), leading to a 10% boost in model accuracy during fine-tuning and enabling the system to reach state-of-the-art performance on internal safety monitoring benchmarks.
- Designed and implemented a scalable synthetic data generation pipeline using Blender, 3D point cloud models, and OpenCV, which cut data collection costs while expanding dataset diversity, thereby accelerating model retraining cycles and improving generalization for few-shot scene.

Security Engineer Intern | NSFOCUS – Shanghai, China 06/2022 - 08/2022

- Implemented robust data validation and preprocessing workflows to ensure data integrity and prevent SQL injection risks in database interactions.
- Developed secure data storage and transmission protocols, including encryption and hashing techniques, to safeguard sensitive information.

Skills

Programming: Python, C++, C, MATLAB, SQL, Markdown
Common Tools: PyTorch, OpenGL, AutoGen, Scikit Learn, OpenCV, Pillow, Hugging Face, Git, Ubuntu Server