

Xiao Wang

(315) 278-8449
Buffalo, NY
xwang277@buffalo.edu

Generative AI/ Efficient AI / LLMs

WebPage:WangXiaoShawn
LinkedIn: Xiao Wang

I am pursuing a Ph.D. in Computer Science and Engineering at the University at Buffalo, where I focus on tightly integrating computer vision, multimodal large language models, and robotics within a real-world perception–decision–control loop to build a new generation of efficient, scalable, and interpretable intelligent agents.

EDUCATION

University at Buffalo, SUNY | Department of Computer Science and Engineering

Ph.D. Student in Computer Science GPA:4.00 Advisor: Prof. Venu Govindaraju

Syracuse University | College of Engineering & Computer Science

MS of Computer Science GPA:3.861

PUBLICATIONS

- **Xiao Wang***, Lu Dong*, Jingchen Sun, Ifeoma Nwogu, Srirangaraj Setlur, Venu Govindaraju. "MistyPilot: An Agentic Fast–Slow Thinking LLM Framework for Misty Social Robots." *IEEE International Conference on Robotics and Automation (ICRA)*, under review.
- **Xiao Wang***, Lu Dong*, Sahana Rangasrinivasan, Ifeoma Nwogu, Srirangaraj Setlur, Venu Govindaraju. "AutoMisty: A Multi-Agent LLM Framework for Automated Code Generation in the Misty Social Robot." *International Conference on Intelligent Robots and Systems (IROS 2025)*.
- **Xiao Wang*** Lu Dong*, Srirangaraj Setlur, Venu Govindaraju, Ifeoma Nwogu."Ig3D: Integrating 3D Face Representations in Facial Expression Inference" *The 18th European Conference on Computer Vision (ECCV 2024 Workshop)*.
- Lu Dong, **Xiao Wang**, Ifeoma Nwogu. "Word-Conditioned 3D American Sign Language Motion Generation" *The 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP 2024)*.
- Bhavin Jawade, Alexander Stone, Deen Dayal Mohan, **Xiao Wang**, Srirangaraj Setlur, Venu Govindaraju. "ProxyFusion: Face Feature Aggregation Through Sparse Experts." *The 2024 Conference and Workshop on Neural Information Processing Systems (NeurIPS 2024)*.
- Lu Dong, Lipisha Nitin Chaudhary, Fei Xu, **Xiao Wang**, Mason Lary, Ifeoma Nwogu. "SignAvatar: Sign Language 3D Motion Reconstruction and Generation." *The 18th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2024)*

PROFESSIONAL EXPERIENCE

PhD Research Intern

May 2025 — Aug 2025

MSAI Reasoning and Language Understanding Group, Microsoft (*Researcher: a reasoning agent for Microsoft 365 Copilot*)

USA

- Collaborated on the design and implementation of a four-stage, multi-agent meeting simulation pipeline (*Scene Planning → Character Modeling → Directed Interaction → Post-production*), replacing one-shot script generation with iterative, collaborative agent workflows.
- Coordinated agents with distinct personas and specialized knowledge bases to produce realistic dialogues averaging 40–60 minutes and over 150 turns, overcoming the “short, flat, fake” limitations of conventional scripted meetings.
- Built an automated fact-evaluation framework to assess the accuracy of key discussion points and detect significant factual deviations.
- Integrated fact-evaluation results as feedback signals into the generation pipeline, enabling automatic quality optimization.
- Delivered hundreds of high-fidelity meeting transcripts, now serving as synthetic training data for multiple internal initiatives.

Data Systems Engineer

2015 Sep—2021 May

SHAANXI HAINA ELECTRONIC TECHNOLOGY CO., LTD

China

- Developed a C++ data transfer and database management system, increasing network speed, improving data security, and boosting database update efficiency.
- Created a MySQL-based C++ API for flexible, efficient data management tailored to diverse business requirements.
- Engineered a process management system with automated modules, ensuring system stability and optimized task scheduling.
- Built a C++ API-driven file transfer system, improving data aggregation and warehousing for enhanced data utilization.
- Optimized database performance through table design, indexing, and stored procedures, resulting in faster data retrieval and increased reliability.
- Applied data analysis and modeling for business insights, supporting customer feature extraction, capacity forecasting, and assembly line strategy simulation.

RESEARCH EXPERIENCE

National AI Institute for Exceptional Education

08/2024-now

Position: Research Assistant, Advisor: Dr. Venu Govindaraju @UB, USA

- **Multimodal Intelligence & Real-Time Interaction:** Integrated large language and vision-language models to enhance contextual understanding. Implemented real-time pose estimation and emotion detection, enabling Misty to extract human posture and behavioral cues from video streams. Dynamically adjusted interaction strategies based on real-time analysis for a more responsive and personalized user experience.
- **AutoMisty: LLM-Powered Multi-Agent Code Generation:** AutoMisty, a multi-agent framework powered by large language models (LLMs), overcomes the inaccessibility of the Misty robot's open API for non-programmers by enabling seamless natural language-to-code generation through specialized agents for task decomposition, assignment, problem-solving, and result synthesis, incorporating self-reflection and human-in-the-loop optimization to ensure transparent reasoning, iterative refinement, and precise execution, significantly outperforming ChatGPT-4o and ChatGPT-o1 in benchmark evaluations.

3D Motion Reconstruction and Generative Modeling for Sign Language Avatars

07/2023-08/2024

Position: Research Assistant, Advisor: Dr. Ifeoma Nwogu @UB, USA

- Developed **SignDiffusion**, a diffusion-based model that yields diverse, realistic, syntax-aligned 3D sign-language avatars from multimodal inputs (audio, text, image, video).
- Proposed **SignAvatar**, a generative framework combining a Transformer-based VAE, CLIP semantic alignment, and curriculum training, setting a new state-of-the-art baseline for 3D signing.
- Built the **ASLGloss103** dataset using 3D SMPLX models to enable quantitative evaluation of sign-language generation and analysis.
- Achieved superior avatar naturalness and accuracy, verified through comprehensive objective metrics and human evaluations.

TECHNICAL EXPERIENCE

MARL-INTERSECT: A Multi-Agent Reinforcement Learning Algorithm for Autonomous Cars

The source code and demo are available on GitHub at: [MARL-INTERSECT](#).

- **Multi-Agent Framework:** Developed independent Actor-Critic models for each vehicle agent, leveraging parallel computing and multi-threaded data sampling to accelerate training and improve stability and scalability.
- **Enhanced Algorithm:** Integrated A3C with PPO and incorporated entropy-based exploration to constrain policy updates and incorporate advantage estimation, ensuring smoother convergence and reliable policy improvements.
- **Simulation Testing:** Validated in a simulated unsignalized intersection scenario, outperforming traditional independent A3C in terms of convergence speed, policy performance, and interaction safety.
- **Scalability Verification:** Demonstrated adaptability and scalability in both small-scale (4 agents) and large-scale (10 agents) environments.

XiaoStyle - Customizable & Secure eCommerce Platform

The source code and demo are available on GitHub at: [XiaoStyle](#).

- Developed a scalable eCommerce platform using Python Django, featuring a custom user model, product/category management, cart functionality, and secure payment integration.
- Implemented comprehensive post-order processing, including stock reduction, invoice generation, and real-time email notifications.
- Integrated a review and rating system with interactive features, secure user authentication, and session management for a seamless user experience.

Ultimate Data Navigator: Efficient Data Management System

The source code is available on GitHub at: [Ultimate Data Navigator](#).

- Built a high-performance C++ system automating data scanning, gathering, and uploads to a centralized database, handling diverse data types and large volumes efficiently.
- Implemented modular utilities for scheduling, resource management, and file transfers, allowing flexible and scalable data operations.
- Enabled cost-effective processing of millions of data entries weekly using open-source components.