Jie Zhu

Last Update: 09/20/2025

zhujie4@msu.edu — +1 (202) 758-8919 — Google Scholar — LinkedIn — Github — Personal Website

Background I am a third-year CS Ph.D. student advised by Dr. Xiaoming Liu (Fellow of IEEE and IAPR) at Michigan State University, and collaborate closely with Dr. Anil Jain (NAE Member, Fellow of ACM and IEEE). Before that, I received my Master's degree in Computer Science at George Washington University in 2023, and Bachelor's degree in Computer Science at Northeastern University in 2020.

Research Interests

1. Multimodal: MLLMs (Under-review), Agents (Submitting to CVPR26), VQA (MM23), Human Recognition (ICCV25).
2. Biometrics (TPAMI (under review)).

PUBLICATIONS

Conference Papers

- **Jie Zhu**, Yiyang Su, Minchul Kim, Anil Jain, and Xiaoming Liu. A Quality-Guided Mixture of Score-fusion Experts Framework for Human Recognition. **ICCV, 2025**.
  Keywords: Multi-modal, Biometrics, MoE
- Junwen Chen, **Jie Zhu**, and Yu Kong. 2023. ATM: Action Temporality Modeling for Video Question Answering. **ACM MM, 2023**.
  Keywords: VQA, Action Understanding
- **Jie Zhu**, Mengsha Hu, Amy Zhang, and Rui Liu. Fairness-Sensitive Policy-Gradient Reinforcement Learning for Reducing Bias in Robotic Assistance. **IEEE ROMAN, 2024.**
  Keywords: Reinforcement Learning, Fairness

Under Review

- **Jie Zhu**, Xiao Guo, and Xiaoming Liu. SapiensAgent: A Multimodal Agent with Dynamic Model Selection for Human Recognition (**Submitting to CVPR26**).
  Keywords: Agents, MLLMs, Refinforcement Learning, Biometrics
- **Jie Zhu**, and Xiaoming Liu. ReFine-RFT: Improving Reasoning Capability of Fine-grained Recognition for Multi-modal LLMs (**Under review**).
  Keywords: MLLMs, Refinforcement Learning, Fine-grained Understanding
- Liu Feng, ..., **Jie Zhu**, et al. Person Recognition at Altitude and Range: Fusion of Face, Body Shape and Gait (**TPAMI (under review)**).
- **Jie Zhu**, Minchul Kim, Zhizhong Huang, and Xiaoming Liu. Subtoken Image Transformer (SiT) for Generalizable Fine-grained Recognition (**Under review**).
  Keywords: Fine-grained Recognition, Image Tokenization

EDUCATION

Michigan State University, United States — Aug 2023 – Apr 2028
Doctor of Philosophy in Computer Science — GPA: 4.0/4.0
Research Areas: Representation Learning, Multi-modal, and Biometric Recognition

George Washington University, Washington, DC, United States — Sep 2021 – May 2023
Master of Science in Computer Science — GPA: 3.9/4.0

Northeastern University, Shenyang, China — Aug 2016 – Jun 2020
Bachelor of Science in Computer Science — GPA: 3.2/4.0

PROJECT EXPERIENCE

BRIAR, MSU & IARPA — United States
Lead Student Researcher - (Human Recognition, Biometrics) — Aug 2023 – Now

- Served as the lead student investigator on the BRIAR program, spearheading the development of the **FarSight biometric recognition system**.
- Independently designed and implemented a novel multimodal fusion framework. Achieved a breakthrough **34.3 percentage point improvement in True Accept Rate**, directly enhancing the operational capabilities of U.S. intelligence and security agencies.
- Published research at the top-tier conference **ICCV**, with ongoing work **under review at TPAMI** and in preparation for **CVPR26**.

ACADEMIC EXPERIENCE

ACTION Lab, Michigan State University United States
Research Intern - (VQA, Action Understanding) Feb 2022 – Nov 2022

- We propose the **ATM** to address VideoQA featuring temporal dynamic reasoning by faithful action modeling. Our action-centric contrastive learning learns action-aware representations from both vision and text modalities.
- We present an **Action-centric Contrastive Learning (AcCL)** for action-plentiful cross-modal representation.
- We fine-tune the model with a newly developed **temporal sensitivity-aware confusion loss (TSC)** that mitigates static bias in temporality reasoning.
- Comprehensive experimental results demonstrate the effectiveness of ATM, especially for temporal reasoning and action understanding with **+2.1%** improvement on *NExT-QA* and **+5.8%** on *TGIF-QA*. The work is accepted by **ACM MM**.

Cognitive Robotics and AI Lab, Kent State University United States
Research Intern - (Reinforcement Learning, Fairness) Mar 2022 – Dec 2022

- We identify four types of **fairness issues** that appear in Human-Robot Interaction in restaurant scenarios to evaluate robots fairness performance.
- We propose a method called Fairness-Sensitive Policy-Gradient Reinforcement Learning for Reducing Bias in Robotic Assistance (FSPGRL) to mitigate robot bias. We demonstrate the effectiveness of our method using **PPO** and **REINFORCE** RL algorithms.
- We developed a logistic regression model for timely **robot bias detection** during service. We set up a questionnaire to survey attitudes toward robot behavior to collect data for model training. The work is accepted by **IEEE ROMAN**.

WORK EXPERIENCE

Inter-American Development Bank United States
AI Analytics Consultant - (LLM, Web Design) Jun 2023 – Aug 2023

- Engineered web scraping pipelines using BeautifulSoup and Scrapy to process multilingual news content from 50+ media sources.
- Developed **ChatGPT-powered dashboard** for automated summarization and trend analysis of text/video news.
- Developed a framework for multimedia content extraction using Automated Speech Recognition (ASR) and ChatGPT.

Research of Institute of Tsinghua, Pearl River Delta Guangzhou, China
AI Engineer (Text-to-Speech) Sep 2020 – Aug 2021

- Developed phoneme-based text normalization pipeline for Text-to-Speech (TTS) systems using **Tacotron 2**.
- Implemented Speech Quality Assessment system with Automatic Speech Recognition (ASR) and feature similarity.
- Built proprietary Mandarin speech dataset containing 100,000+ clean/noisy audio samples with text transcriptions.
- Filed **14 CN patents** with 2 as first inventor. 10 patents granted.

Seeking AI Co. Ltd. Guangzhou, China
R&D Intern Dec 2019 – Apr 2020

- Developed automated dimensional analysis tool using OpenCV contour detection.
- Contributed to CI/CD pipelines using GitLab for model deployment on edge devices.

HONORS & AWARDS

Graduate Tuition Fellowship Aug 2022
Faculty Awards of Computer Animation Dec 2021
Third Prize Scholarship Sep 2018 – Jul 2020

ACADEMIC SERVICES

- **Reviewer:** TPAMI 2025; FG 2024-2025; IJCBLRR 2024
- **Teaching Experience:** Computer Animation (Fall 2022), Computer Graphics II (Spring 2023)

PATENTS

CN113194348B, "Virtual human lecture video generation method, system, device and storage medium", granted: July 2022.
CN112562720B, "Lip-sync video generation method, device, equipment and storage medium", granted: July 2024.
CN113192161B, "Virtual human image video generation method, system, device and storage medium", granted: October 2022.
CN113192162B, "Method, system, device and storage medium for driving image by voice", granted: December 2022.

CN112487978B, "A Method for Speaker Localization in Video", granted: April 2024.
CN112562721B, "Method and device for positioning speaker in video and computer storage medium", granted: April 2024.
CN113179449B, "Method, system, device and storage medium for driving image by voice and motion", granted: April 2022.
CN112562719B, "Method, system, device and storage medium for matching synthesized voice with original video", granted: March 2024.
CN112530401B, "Speech synthesis method, system and device", granted: May 2024.
CN112565885B, "Video segmentation method, system, device and storage medium", granted: January 2023.
CN112530400A, "Method, system, device and medium for generating voice based on text of deep learning", filed: November 2020 (pending).
CN113259778A, "Method, system and storage medium for using virtual character for automatic video production", filed: April 2021 (pending).

SKILLS

- **Programming:** Python, PyTorch, Hugging Face, MuJoCo, Unity (AR/VR development), HTML, Golang, C++.
- **Languages:** Chinese (Native), Cantonese (Native), TOEFL 102 (Speaking: 24).