Jie Zhu

(EMAIL) zhujie4@msu.edu — (TEL) +1 (202) 758-8919 — LinkedIn — Github —Personal page

Summary

I am a second-year CS PhD Student advised by Prof. Xiaoming Liu at Michigan State University. 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. I have 2 publications on ACM MM, ROMAN, and 2 under-reviewed ICCV papers.
Research Interests: Representation Learning, Multi-modal, VLMs, Image Understanding, Biometric Recognition

PUBLICATIONS

Conference Papers

- 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**, Minchul Kim, Zhizhong Huang, and Xiaoming Liu. Subtoken Image Transformer (SiT) for Generalized Category Discovery (**ICCV2025 Under review**).
  Keywords: Fine-grained Classification, Image Tokenization
- **Jie Zhu**, Yiyang Su, Minchul Kim, Anil Kumar Jain, and Xiaoming Liu. A Quality-Guided Mixture of Score-fusion Experts Framework for Human Recognition (**ICCV2025 Under review**).
  Keywords: Biometrics, Multi-modal, Score-fusion

EDUCATION

Michigan State University, United States — Aug 2023 – Apr 2028
Computer Science Doctoral Program — Cumulative 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 — Cumulative GPA: 3.9/4.0

Northeastern University, Shenyang, China — Aug 2016 – Jun 2020
Bachelor of Science in Computer Science — Rank: Top 30

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*.

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.

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 TTS systems using Tacotron 2.
- Implemented Speech Quality Assessment system with ASR and feature similarity.
- Built proprietary Mandarin speech dataset containing 100,000+ clean/noisy audio samples with text transcriptions.

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.

PROJECTS

Intellectual Property Management System Based on Blockchain Shenyang, China
Graduation Project Mar 2020 – Jun 2020

- Design the system structure by analyzing the demand of the market.
- Use Flask to build the website, which includes IP register, IP trade, and electrical evidence generation functions.

Fake News Information Analysis and Visualization Platform Shenyang, China
Developer Jun 2018 – Jan 2019

- Checked and cleansed data, calculated similarities, and extracted keywords for 10k+ copies of texts using ELMO model.
- Detected duplicates between new data and the database, with an accuracy rate of 92% in the screening test.

HONORS & AWARDS

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

Academic Services & Activities

- **Reviewer**
  FG 2024-2025
  IJCBLRR 2024
- **Teaching Experience**
  Computer Animation Fall 2022
  Computer Graphic II Spring 2023

SKILLS

- **Programming:** Python, HTML, CSS, JavaScript, Golang, Java, C++.
- **Deep Learning Frameworks:** PyTorch, Hugging Face, TensorFlow.
- **Languages:** Chinese (Native), Cantonese (Native), TOEFL 102 (Speaking: 24).
- **Others:** Linux, LaTeX, GitHub, MuJoCo, Unity (AR/VR development).