Xiang Zhang

(607) 239 - 8929 | zxiang4@binghamton.edu | www.xianng.com | xianng | Google Scholar

Summary

Innovative Ph.D. candidate in Computer Science with 5+ years of experience in computer vision, multi-modal learning, facial analysis, and generative AI.

Expertise in deep learning, large-scale data processing, and **generative models** (GANs, Diffusion, NeRF), with a focus on text-driven AI, **large language models (LLMs)**, and **language-vision** integration. Proven track record of publishing in top-tier conferences and developing state-of-the-art AI technologies.

Hands-on experience in deploying machine learning solutions and collaborating across disciplines. Industrial experience includes autonomous driving planning and AI research at XPENG (xmotors.ai). Seeking an internship or full-time role in AI research and development. Expected graduation: **Winter 2025**.

Education

Binghamton University

Binghamton, NY

Ph.D. Candidate in Computer Science; *Aug. 2019 - Dec. 2025(expected)*

Binghamton University

Binghamton, NY

Master of Science in Computer Science, GPA: 3.81/4.0 *Aug. 2017 - May 2019*

Experience

XPENG (xmotors.ai)

San Diego, CA

Machine Learning Engineer Intern *Aug. 2024 - Jan. 2025*

- Work with cross-teams to deploy machine learning solutions to continuously improve SOTA performance for multiple perception tasks, e.g., object detection, segmentation, etc.
- Deliver pre-trained and fine-tuned models to achieve the class's most intelligent autonomous driving system.
- Contributed to autonomous driving planning, and optimizing decision-making algorithms for self-driving vehicles.

Research Experiences

Advanced Research Projects Agency (ARPA) and Kitware Inc

Biometric Recognition And Identification At Altitude And Range(BRIAR) *05/31/2023 - 06/30/2024*

- Researched long-distance face recognition in extreme conditions, using advanced Generative AI models, including **stable diffusion, GANs, and neural radiation fields (NeRF)**.
- Developed a multi-stage **face restoration** and **3D multi-view generation** pipeline to enhance recognition under adverse conditions (e.g., atmospheric turbulence, low resolution).
- Collaborate with **Kitware Inc.** to transfer research advances into their framework, contributing to the overall face recognition system.
- Published findings in **Enhancing Face Recognition in Low-Quality Images Based on Restoration and 3D Multiview Generation. [IJCB 2024]**.

Binghamton University GAIC Lab

MULTI-MODAL LEARNING RESEARCH *Aug. 2019 - Present*

- Study in multi-modal learning using texture, thermal, 3D data, text, and EEG data for facial behavior analysis, such as **AU detection** and **facial expression recognition**.
- Built large-scale **human emotion databases** (BP4D++, BU-EEG, ReactioNet) with multi-modal signals for **affective computing** research.
- Published papers in **ICCV, WACV, FG,** and **TAC**, advancing state-of-the-art methods in multi-modal AI and human-centered computing.

TEXT-DRIVEN AND LLM-RELATED WORK

- This work explores **text-driven contrastive learning (CLIP-based)** to analyze facial behaviors, combining textual descriptions with visual data.
- It also ties into advancements in **Large Language Model (LLM)** and **language-vision model**, where textual descriptions enhance the model's ability to recognize and understand facial expressions.
- This work is reported in the research paper: **Weakly-Supervised Text-driven Contrastive Learning for Facial Behavior Understanding. [ICCV 2023]**

Selected Publications

** denotes equal contribution*

- **Zhang, Xiang**, Xiaotian Li, Taoyue Wang, and Lijun Yin. *Enhancing Face Recognition in Low-Quality Images Based on Restoration and 3D Multiview Generation.* [IJCB 2024]
- **Zhang, Xiang**, Huiyuan Yang, Taoyue Wang, Xiaotian Li and Lijun Yin.*Multimodal Channel-Mixing: Channel and Spatial Masked AutoEncoder on Facial Action Unit Detection*. [WACV 2024]
- **Zhang, Xiang**, Taoyue Wang, Xiaotian Li, Huiyuan Yang and Lijun Yin.*Weakly-Supervised Text-driven Contrastive Learning for Facial Behavior Understanding*. [ICCV 2023]
- Li, Xiaotian, **Xiang Zhang**, Taoyue Wang and Lijun Yin. *Knowledge-Spreader: Learning Facial Action Dynamics from Single Label Clips via Progressive Knowledge Distillation*. [ICCV 2023]
- Li, Xiaotian, Taoyue Wang, Geran Zhao, **Xiang Zhang**, Xi Kang and Lijun Yin.*ReactioNet: Learning High-order Facial Behavior from Universal Stimulus-Reaction by Dyadic Relation Reasoning*. [ICCV 2023]
- Xiaotian Li, Zheng Zhang, **Xiang Zhang**, Taoyue Wang, Zhihua Li, Huiyuan Yang, Umur Ciftci, Qiang Ji, Jeffrey Cohn, and Lijun Yin.*Disagreement Matters: Exploring Internal Diversification for Redundant Attention in Generic Facial Action Analysis.*. [TAC 2023]
- **Zhang, Xiang** and Lijun Yin.*Multi-Modal Learning for AU Detection Based on Multi-Head Fused Transformers*. [FG 2021]
- Li, Xiaotian*, **Xiang Zhang***, Huiyuan Yang, Wenna Duan, Weiying Dai and Lijun Yin.*An EEG-Based Multi-Modal Emotion Database with Both Posed and Authentic Facial Actions for Emotion Analysis*. [FG 2020]

Skills

Programming Languages	Python, C/C++, Java, JavaScript, Objective-C, Matlab
Frameworks	PyTorch, PyTorch-Lighting, Tensorflow, Scikit-Learn, TensorRT, ONNX
Generative AI	GANs, Diffusion Models, NeRF, Transformers
Large-Scale Model Training	Distributed Training, Model Optimization (LoRA)
Platforms & Tools	Linux, Docker, Git, Azure, CI/CD pipeline