

# XINPENG LI

Dallas, TX 75080, United States

◊ xinpeng.li@utdallas.edu ◊ <https://sampson-lee.github.io/>

## SUMMARY

Ph.D. student in Computer Science at UT Dallas (advised by Prof. [Yapeng Tian](#)). My research includes multimodal large language models, audio-visual social signal processing, and dataset/benchmark development for human-centric interaction. I have published papers in ACM MM, IEEE TMM, NeurIPS and released multiple datasets, benchmarks, and open-source systems.

## EDUCATION

**University of Texas at Dallas**, Dallas, TX, U.S.A.

*Aug. 2024 - Present*

**Doctor of Philosophy** in Computer Science

Advisor: Prof. [Yapeng Tian](#)

**South China University of Technology**, Guangzhou, P.R.China

*Sept. 2018 - June 2021*

**Master of Engineering** in Signal and Information Processing

Advisor: Prof. [Dacheng Tao](#)

**South China University of Technology**, Guangzhou, P.R.China

*Sept. 2014 - June 2018*

**Bachelor of Engineering** in Information Engineering

## RESEARCH EXPERIENCE

**Computer Vision and Multimodal Computing Lab, UT Dallas**

*Aug. 2024 - Present*

Research Assistant

Supervisor: Prof. [Yapeng Tian](#)

- Research on **audio-visual pipelines** for multimodal social interaction datasets, covering speech diarization, active speaker detection, and representation learning

- Developed **LLM-based systems**, integrating conversation forecasting, chain-of-thought reasoning, and agentic tool usage, that operate on raw text, audio and video signals.

**MIPS Lab, SZTU & VIP Lab, SUSTech**

*Jun. 2021 – Jul. 2024*

Research Assistant

Mentors: Profs. [Xiaojiang Peng](#), [Feng Zheng](#), [Jinbao Wang](#)

- Led projects on railway detection, emotion recognition and context-aware affective computing (MIPS) and contributed to dataset and benchmark of 3D anomaly detection (VIP, Jan. 23 to Jul. 23)

- Released datasets, benchmarks, and baseline models for railway detection, emotion recognition and anomaly detection, enabling reproducible multimodal research

## PUBLICATIONS (SELECTED)

### 1. Omni-MMSI: Towards Identity-aware Social Interaction Understanding.

**Xinpeng Li**, Bolin Lai, Hardy Chen, Shijian Deng, Cihang Xie, Yuyin Zhou, James M. Rehg, Yapeng Tian.

*Submitted to CVPR, 2026.*

Introduced a LLM-based framework that reasons with leveraging references and using tools.

### 2. Towards Online Multi-Modal Social Interaction Understanding. [PDF] [Code]

**Xinpeng Li**, Shijian Deng, Bolin Lai, Weiguo Pian, James M. Rehg, Yapeng Tian.

*Submitted to TMLR, 2025.*

Proposed a VLM pipeline with forecasting and prompting for online social interaction understanding.

3. **Two in One Go: Single-stage Emotion Recognition with Decoupled Subject-context Transformer.** [\[PDF\]](#) [\[Code\]](#)  
**Xinpeng Li**, Teng Wang, Jian Zhao, Shuyi Mao, Jinbao Wang, Feng Zheng, Xiaojiang Peng, Xuelong Li.  
*ACM Multimedia, 2024.*  
Designed a decoupled subject–context transformer for effective visual emotion recognition.
4. **Facial Action Units as A Bridge of Joint Dataset Training for Facial Expression Recognition.** [\[PDF\]](#) [\[Code\]](#)  
Shuyi Mao, **Xinpeng Li**, Fan Zhang, Xiaojiang Peng, Yang Yang.  
*IEEE Transactions on Multimedia, 2024.*  
Leveraged facial action units as intermediate representations to improve multi-dataset learning.
5. **Real3D-AD: A Dataset of Point Cloud Anomaly Detection.** [\[PDF\]](#) [\[Code\]](#)  
Jiaqi Liu, Guoyang Xie, Ruitao Chen, **Xinpeng Li**, Jinbao Wang, Yong Liu, Chengjie Wang, Feng Zheng.  
*NeurIPS Datasets and Benchmarks Track, 2023.*  
Released the first real-world 3D point cloud anomaly detection dataset with standardized benchmarks.
6. **Rail Detection: An Efficient Row-based Network and A New Benchmark.** [\[PDF\]](#) [\[Code\]](#)  
**Xinpeng Li**, Xiaojiang Peng.  
*ACM Multimedia, 2022.*  
Introduced a new row-based rail detection dataset and proposed an efficient row-based model.

## HONORS & AWARDS

---

Top 11% in Competition on Affective Behavior Analysis in-the-wild (ECCV)	2022
Top 6% in Multimodal Sentiment Analysis Challenge (ACM MM)	2021
Top 9% in Short Video Face Parsing Challenge (CVPR)	2021
Multiple scholarships including National and University-level	2015-2020

## ACADEMIC SERVICES

---

### Conference Reviewer:

ACM Multimedia 2024, ACM Multimedia 2023, ICCV 2025, CVPR 2026

### Teaching Assistant:

Computer Science II Summer'25, Operating System Fall'25, Operating System Spring'26