Yitao XU

111 Renai Road, Suzhou, China 215000
+86-189-1525-1158 yxu879@gatech.edu www.linkedin.com/in/yitaoxu20

Education

Georgia Institute of Technology **Atlanta, GA, USA**
Master of Science in Electrical and Computer Engineering *2024.8 – 2026.6 (Expected)*

Duke University (Kunshan Campus) **Suzhou, China**
Computer Vision, GSI Programme, Advisor: Prof. Kaizhu Huang *2023.7 – 2023.8*

University of Liverpool (UoL) **Liverpool, UK**
Xi'an Jiaotong-Liverpool University (XJTLU) **Suzhou, China**
Bachelor of Science in Information and Computing Science *2020.9 – 2024.7*

- Major GPA: 3.85 / 4.0, Graduate with a first class honours degree.
- Awarded two degrees from XJTLU and the University of Liverpool.

Publications

Context-Enhanced Multimodal Sentiment Analysis: Multi-Granular Alternative Telescopic Displacement **2024**
In submission to Information Fusion (Journal)
- Author: Jiahao Qin, **Yitao Xu**, Zong*, and Lu Feng Liu**.

Challenging Data Normalization in Multimodal Sentiment Analysis: An Empirical Study on the MOSI and MOSEI Datasets **2024**
In submission to AAAI 2025
- Author: Jiahao Qin, **Yitao Xu**, and Lu Zong*.

MTSA-SNN: A Multi-modal Time Series Analysis Model Based on Spiking Neural Network **2024**
Received by International Conference on Pattern Recognition 2024
- Author: Chengzhi Liu, Chenghao Liu, Zheng Tao, Zihong Luo and **Yitao Xu***.

MC-DBN: A Deep Belief Network-Based Model for Modality Completion **2024**
In submission to International Conference on Neural Information Processing 2024
- Author: Zihong Luo*, Chengzhi Liu, Zheng Tao, Kexin He, and **Yitao Xu**.

Multimodal Deep Learning for Enhanced Arrhythmia Detection Using ECG Time Series and Image Data **2023**
Received by IEEE International Conference on Cyber-enabled distributed computing and knowledge discovery 2023
- Author: **Yitao Xu**, Jiahao Qin*, Zihong Luo, Zhengxu Jing, Bihao You.

Research Experience

LLM-Driven Dialogue State Tracking Project | *Collaborated with PhD candidate Minyu Jin* **2024.6 – Present**
- Developed a dialogue state tracking system leveraging large language models (LLMs) to enhance conversational AI capabilities, focusing on improving context retention and user interaction.
- Implemented innovative tracking techniques that capture and maintain dialogue context more effectively by studying the reasoning and planning abilities of LLMs, determining whether they genuinely exhibit reasoning or primarily rely on knowledge retrieval from pre-training data.
- Investigated efficient and effective LLM prompting and fine-tuning techniques for information extraction tasks, including named entity recognition and relation extraction, to ensure robust dialogue state tracking.
- Contributed to the development and evaluation of new methodologies for state tracking, demonstrating significant improvements in tracking accuracy and performance across various conversational contexts.

Alignment Optimization for Multimodal Sentiment Analysis | *Supervised by Asst. Prof. Fangyu Wu* **2023 – 2024**
- Study the difficult problems in multimodal sentiment analysis, such as semantic contradictions, asynchronous and heterogeneous problems of multimodal data.

- Develop algorithms to address challenges in aligning text, visual, and auditory data, with a special focus on aligning over high-dimensional representation spaces.
- Investigate the potential value of the classical optimal transmission algorithm in multimodal alignment.
- Apply the designed algorithm and structure to the MSA-single and MSA-multiple data sets and achieved SOTA.

Anomaly Detection System Based on Computer Vision | *Supervised by Prof. Kaizhu Huang* **2023.8**

- Preprocessed datasets using techniques such as normalization, data augmentation, and noise reduction to improve model performance.
- Implemented advanced image processing techniques and utilized ResNet-based models, incorporating a custom-designed contrastive learning method for semantic segmentation to accurately identify anomalies.
- Validate the system's performance and robustness using real-world data and scenarios.

Multimodal Alignment with Deep Learning for Financial Analysis | *Supervised by Assoc. Prof. Zong Lu* **2023.4**

- Scraped ten years of historical data from NASDAQ 300 and company news data, preprocessed using BERT, and packaged into a comprehensive dataset.
- Investigated existing multimodal financial analysis models, with a focus on long-term semantic polarity analysis of news text.
- Designed a multimodal financial analysis model incorporating data, candlestick charts, and news data, emphasizing alignment and fusion issues.
- Developed a bidirectional gated unit to prevent potential contamination from the text modality, thereby maintaining high performance.

Technical Skills

Languages: Python, Java, C, HTML/CSS, JavaScript, SQL

Technologies/Frameworks: Pytorch, GitHub, JUnit, LaTeX

Other Interests: Cycling, Photography, Reading