Xinmiao Yu

xmyu@ir.hit.edu.cn · Google Scholar · Linkedin

Research Interest

My research interest focuses on developing advanced multimodal models and LLM agents capable of enhancing their intelligence through interaction with the real world.

Education

Politecnico di Milano, Milano, Italy Sep. 2024 – Jun. 2025(expected)

Double-degree Master candidate School of Computer Science(CS)

Harin Institute of Technology *C9 First-class* Harbin, China Sep. 2023 – March. 2026(expected)

Master candidate School of Computer Science(CS) Supervisors: Prof.Bing Qin ***Outstanding Scholarship***

Harin Institute of Technology *C9 First-class* Harbin, China Sep. 2019 – Jun. 2023

B.S. School of Computer Science(CS) *GPA: 3.78 (5/83)* ***National Scholarship, Outstanding Graduate***

UCLA Visiting Student Los Angeles, United States Mar. 2023-Jun. 2023

Major Course Machine Learning A Behavioral Finance A+ Principles of Accounting A

Publications

1. **Xinmiao Yu**, Xiaocheng Feng, Yun Li, Minghui Liao, Ya-Qi Yu, Xiachong Feng, Weihong Zhong, Ruihan Chen, Mengkang Hu, Jihao Wu, Duyu Tang, Dandan Tu, Bing Qin. Cross-Lingual Text-Rich Visual Comprehension: An Information Theory Perspective. submitted to AAAI 2024
2. Mengkang Hu, Yao Mu, **Xinmiao Yu**, Mingyu Ding, Shiguang Wu, Wenqi Shao, Qiguang Chen, Bin Wang, Yu Qiao, Ping Luo. Tree-Planner: Efficient Close-loop Task Planning with Large Language Models [arxiv][ICLR 2024]
3. **Xinmiao Yu**, Meng Qu, Xiaocheng Feng, Bing Qin. GraphAgent: Exploiting Large Language Models for Interpretable Learning on Text-attributed Graphs. [arxiv]
4. Zhangyin Feng, Yuchen Ren, **Xinmiao Yu**, Xiaocheng Feng, Duyu Tang, Shuming Shi, Bing Qin. Improved Visual Story Generation with Adaptive Context Modeling: ACL 2023
5. Xuehui Yu, Jingchi Jiang, **Xinmiao Yu**, Yi Guan*,Xue Li. Causal Coupled Mechanisms: A Control Method with Cooperation and Competition for Complex System. BIBM 2022

Experience

HUAWEI Shanghai, China Oct. 2023 – Aug. 2024

Research Intern Supervisors: Duyu Tang and Prof. Xiaocheng Feng

- Developed data construction and alignment training methods for visual language models, incorporating reinforcement learning techniques to enhance OCR and complex reasoning capabilities on text-rich images.
- Constructed XT-VQA, a cross-lingual text-rich visual QA benchmark, revealing performance gaps in LVLMs due to insufficient visual information activation across languages.
- Developed MVCL-MI, a policy-inspired approach maximizing vision-language mutual information, reducing cross-lingual disparities through knowledge distillation from monolingual to multilingual contexts.

Mila-Quebec AI institute Quebec, Canada June. 2023 – Oct. 2023

Research Intern Supervisor: PhD.Meng Qu

- Developed GraphAgent, a novel approach reframing text-attributed graph learning as an agent planning problem, leveraging Large Language Models to explore both structural and textual features in graphs.
- Implemented a policy-driven framework where the LLM-parameterized agent takes actions tailored for text-attributed graphs, achieving improved performance and interpretability through a process analogous to state-action planning in reinforcement learning.

University of Southern California GLAMOR lab Los Angeles, U.S. May. 2023 – Present

Research assistant Supervisors: Prof. Jesse Thomason and PhD. Ishika Singh

- Scene Graph Construction: Developed an automated scene graph pipeline that enabled robots to autonomously construct and refine 3D maps from interactions, enhancing scene accuracy and robot autonomy.
- Task Planning Innovation: Integrated a Chain-of-Thought methodology in task planning, significantly improving planning precision and efficiency.

HIT SCIR Lab Harbin, China Jun. 2022 – Jan. 2023

Master candidate Supervisors: Prof. Xiaocheng Feng and PhD. Zhangyin Feng.

- Developed an improved visual story generation model with adaptive context modeling, addressing the limitation of treating historical images equally by implementing a more nuanced approach to historical context.
- Implemented a novel guidance mechanism in the sampling stage, enhancing global consistency and achieving SOTA FID scores on story visualization and continuation tasks on PororoSV and FlintstonesSV datasets.

Languages and Skills

Python, Java, C++, Pytorch, Tensorflow, NLP, Large Language Models

Honors and Awards

First-Class Graduate Scholarship(Top 20%)	2024
Deloitte Digital Elite Challenge Runner-up(Top 0.2%)	2024
Graduate Entrance Outstanding Scholarship(Top 20%)	2023
Outstanding graduate award (Top 1%)	2023
1st Prize, HUAWEI Cup(The only), on National Undergraduate Internet of Things Contest	2022
National Scholarship (Top 1%)	2022
People's Scholarship	2020,2021,2022
Excellent Student Cadre (Top 5%)	2020,2021,2022