Rao Ziyang

MPhil. Candidate in Artificial Intelligence
Hong Kong University of Science and Technology
zrao538@connect.hkust-gz.edu.cn — +86 15113849177

SUMMARY

I am an MPhil. candidate in the AI Thrust of HKUST, supervised by Prof. Hui Xiong and Dr. Xuming Hu. My research interests focus on **mechanistic interpretability and multi-modal LLMs**, with additional knowledge in data mining and graph analysis. I also have experience in AI startups, industrialization, and knowledge transfer.

EDUCATION

Hong Kong University of Science and Technology — Hong Kong
Mphil. Candidate in Artificial Intelligence — Sep. 2023 — Jul. 2025 **(expected)**

- **CGPA:** 4.0 / 4.3
- **Supervisor:** Prof. Hui Xiong, Associate Vice President, AAAI, AAAS, and IEEE Fellow
- **Co-supervisor:** Dr. Xuming Hu, Assistant Professor
- Postgraduate scholarship award

Renmin University of China — Beijing, China
BSc. in Economic Statistics — Sep. 2019 — Jul. 2023

- **CGPA:** 3.4 / 4.0
- Outstanding Graduation Thesis of Renmin University of China **(top 5 of the cohort)**

NUS School of Computing — Singapore, Singapore
Summer Workshop — May. 2022 — Jul. 2022

- Specialised in data mining and visual computing

RESEARCH

Modelling and Interpreting Information Flow in Multi-modal LLM — Nov. 2024 — present
Lead Researcher

- Preparing for submission to **ACM MM 2025**.
- Employed **sparse autoencoders (SAEs)** in the residual stream of multiple VLMs to model the internal information flow, extracted **human-understandable** visual and text concepts and revealed the concept hierarchy inside VLMs.
- Identified **modality-specific features** with entropy metrics and interpreted the dynamics of **modality gap** between visual and text with feature behavior.
- Located the fine-grained origin of **visual hallucinations** with activation patching on extracted features.

RGB-Event ISP: The Dataset and Benchmark — May. 2024 — Sep. 2024
Contributing Researcher

- Accepted by **ICLR 2025**.
- Proposed a new task of event camera guided image signal processor (event-ISP), collected a large scale pixel-level aligned RGB-event signal paired dataset and benchmarked SOTA baselines on our dataset.
- **Contributions:** collection and organization of the entire dataset, development of benchmarking pipeline and reproduction of baseline models, writing of the experiment and analysis section.

WORK EXPERIENCES

Guangzhou QiWu Co.,Ltd. an AI startup — Guangzhou, China
Co-founder — Aug. 2024 — present

- Responsible for **business plan pitching** and product-market fit.
- Developed demos of **algorithm** training, visualization and deployment to meet customer needs.
- Providing 3D-related AI solutions for our business partners: Bonsai design, interior design, large scale scene reconstruction for construction sites.

HKUST Fok Ying Tung Research Institute Guangzhou, China
Research Assistant Oct. 2022 — Aug. 2023

- **Advisor:** Prof. Hui Xiong, Associate Vice President, AAAI, AAAS, and IEEE Fellow.
- Participated in a **big data mining** project to develop a large-scale academic talent network sponsored by Tencent.
- Designed an **internet crawler** to scrape related information of academic talents.
- Developed an information parser with **NLP techniques** (e. g. word2vec, BERT-based models) to retrieve structured data from the raw crawl results. **before ChatGPT api launch*

PROJECTS

Multi-modal LLM for Brain to Text Generation Apr. 2024 — present
MPhil. Graduation Theis

- Explored the **multimodal capabilities** of LLMs to enhance electroencephalogram (EEG) to text generation.
- Discovered that existing SOTA methods suffer from high signal-to-noise ratio and modality gap, generating barely unreadable outputs.
- Modeled the **modality gap** and information bottleneck with information theory and dimensionality reduction methods.
- The project inspired my further research on *modelling and interpreting information flow in multi-modal LLM*.

Large Language Model for Spatio-Temporal Graph Oct. 2023 — Mar. 2024
Individual Academic Project

- Explored the potential of LLMs to handle **spatio-temporal graph (STG)** with no pretrained encoder available.
- Trained a new STG tokenizer to encode STGs into the token space of GPT-2, then **finetuned GPT-2** with soft prompt and adapter modules.
- Research not published due to the inferior performance on zero-shot tasks. Yet key insights about the multi-modality potential of LLMs and practical skills in LLM research were gained for further research.

Knowledge Graph Analysis on Bloomberg Tech News Nov. 2023
Data Mining Project

- Crawled more than 20,000 news articles for Bloomberg news tech sector.
- Extracted the entities and relations from raw crawl text with NER and built a **knowledge graph**.
- Performed sentiment analysis on news articles with **BERT-based model**, and integrated into the knowledge graph.
- Performed **data mining** on the graph (e. g. communities detection, shortest path search) and analysed with finance domain expert for further interpretation.
- Explored the potential capabilities of **LLMs** to reason and perform analysis on knowledge graph.

Intelligent Production Line Control System Aug. 2023 — Aug. 2024
Industrial Project

- Cooperated with GJSS Co, Ltd. (a Japanese steel sheet manufacturer) to improve the efficiency of the steel sheet production line.
- Performed on-site survey and discussed with production line engineers to formulate the exact problems to solve with AI.
- Developed multiple models for different stages of production including time series prediction, parameters selection, production line status visualization.
- Deployed the models on the factory intranet server and informed the engineers of their functions.

A Simple Framework for Contrastive Learning on Spatio-Temporal Graph Prediction Sep. 2022 — Mar. 2023
BSc. Graduation Thesis

- Identified the difficulties of designing positive pairs for **contrastive learning** on spatio-temporal graph (STG) models.
- Designed a novel contrastive framework for STG models which utilize a simple Dropout layer instead of carefully designed augmentation methods to automatically generate positive pairs for contrastive task.
- The framework could be integrated to existing STG models in a plug-and-play manner, and significantly boosted the performance of most baseline models on multiple datasets.

ACADEMIC SERVICE

- **Conference Reviewer:** ICLR 2025

SKILLS

- **Programming Languages:** Python, R, Matlab, SQL, C++
- **Python Libraries:** PyTorch, transformers, HuggingFace, Accelerate, TransformerLens, Selenium, Playwright...
- **Language:** Chinese, Cantonese, English (IELTS 7.5), German (B1)