Zhongze LUO

(+86)136-0797-9355 | luozhongze0928@foxmail.com | https://luozhongze.github.io

Education

The Chinese University of Hong Kong, Shenzhen *(Expected) 2025.09—2027.06*
M.Phil. of Computer and Information Engineering *GPA: 3.82/4.00*

Northeast Forestry University (**Project 211**) *2021.09—2025.06*
B.Eng of Communication Engineering *GPA: 88.08/100 Ranking: 14/73 (Top 20%)*

Research interests

Multimodal Understanding, LLM Reasoning and Evaluation, Knowledge Graph.

Project

VRGR - Verifiable and Reflective Grounded Reasoning for Honest MLLMs

Methodology: Introduced a structured reasoning framework for MLLMs under visual and logical uncertainty, enforcing explicit visual grounding with bounding boxes and answer-or-refusal decisions. Built the dataset by converting existing VQA benchmarks into answerable and unanswerable cases spanning visual contradiction, perceptual limitation, and reasoning inconsistency. Used VLM for generation of grounded reasoning chains, with data validation via format checks and visual consistency verification using Grounding DINO. Trained models using LoRA-based SFT and GRPO-based reinforcement learning with joint rewards over outcome correctness and format compliance.

Results: Consistently improved trustworthiness, refusal recall, and refusal rationality across multimodal benchmarks. Demonstrated that lightweight format constraints combined with outcome-based rewards effectively mitigate shortcut learning and induce grounded, honest multimodal reasoning.

Multi-Physics - Benchmark for Multimodal LLMs

Methodology: Engineered a specialized benchmark to evaluate Multimodal Large Language Models on 1,412 high-school physics problems across 11 subjects. Leveraged the Mathpix API for high-fidelity OCR extraction of text and formulas from complex pedagogical documents. Developed a difficulty-scoring system based on computational steps and visual complexity. Implemented a dual-scoring framework using a LLM-as-a-judge approach to verify both final answer accuracy and the logical consistency of generated Chain-of-Thought reasoning steps.

Results: Benchmarked 20 state-of-the-art models to identify critical failures in spatial reasoning and multi-step physical deduction. Established a new standard for evaluating how visual context influences scientific problem-solving in Chinese-centric multimodal environments.

KG2QA - Knowledge Graph-Enhanced RAG Framework

Methodology: Developed a Retrieval-Augmented Generation pipeline integrating a domain-specific Knowledge Graph with fine-tuned LLMs for communication standards. Constructed a graph database in Neo4j comprising 13,906 entities and 13,524 relations through LLM-assisted triple extraction and a custom ontology. Executed supervised fine-tuning on Qwen2.5-7B-Instruct using LoRA adapters on a specialized dataset of 6,587 high-quality instruction-following pairs derived from ITU-T recommendations.

Results: Achieved a dramatic performance increase by raising the BLEU-4 score from 18.86 to 66.90. Demonstrated superior factual grounding compared to Llama-3-8B and significantly reduced inference latency, effectively eliminating hallucinations in technical consultations for telecommunication standards.

Internship

BYD Company Limited [link] *2024.01—2024.06*

As a hardware intern, I learned to weld small capacitors, resistors and other components, including 0201MLCC, and MCU. I also learned hardware circuit knowledge, including DC-DC power supply, EMC protection, ADI power supply design, mastered the real vehicle measurement method, participated in the oscilloscope test of vehicle rear domain control circuit board, recorded steady state, shock, stall current, input waveform no-action voltage and action voltage.

Skills

Familiar with python programming, and the training process and model deployment of deep learning models.

Familiar with the use of Altium Designer and other PCB design software, also familiar with Linux operating systems.

IELTS Academic: 6.5(6) Listening 7 Reading 7.5 Writing 6 Speaking 6.

Honors

Outstanding Undergraduate Graduation Thesis of Northeast Forestry University (3%) *2025.06*

China College Student Computer Design Competition - Third Prize *2024.08*

National College Students Mathematical Contest in Modeling - Provincial First Prize *2022.11*

University scholarship - First Prize *2022.03*

Papers

- **Luo, Z.**, Yin, Z., Guo, Y., Wang, Z., Zhu, J., Tang, X. (2026). Multi-Physics: A Comprehensive Benchmark for Multimodel LLMs Reasoning on Chinese Multi-subject Physics Problems. ICASSP 2026. [link]
- **Luo, Z.**, Wan, W., Zhang, T., Wang, D., Tang, X. (2026). KG2QA: Knowledge Graph-enhanced Retrieval-augmented Generation for Communication Standards Question Answering. ICASSP 2026. [link]
- Sun, J., & **Luo, Z.** (2025). ForPKG: A Framework for Constructing Forestry Policy Knowledge Graph and Application Analysis. In 2025 International Joint Conference on Neural Networks (IJCNN). [link]
- Sun, J., **Luo, Z.**, & Li, Y. (2025). A compliance checking framework based on retrieval augmented generation. In Proceedings of the 31st International Conference on Computational Linguistics (COLING). [link]
- Zheng, Q., **Luo, Z.**, Guo, M., Wang, X., Wu, R., Meng, Q., & Dong, G. (2025). HGO-YOLO: advancing anomaly behavior detection with hierarchical features and lightweight optimized detection. Journal of Real-Time Image Processing, 22(4), 1-15. [link]
- **Luo, Z.**, Jia, S., Niu, H., Zhao, Y., Zeng, X., & Dong, G. (2024). Elderly fall detection algorithm based on improved yolov5s. Information Technology and Control, 53(2), 601-618. [link]
- Sun, J., Dai, C., **Luo, Z.**, Chang, Y., & Li, Y. (2024). LawLuo: A Multi-Agent Collaborative Framework for Multi-Round Chinese Legal Consultation. arXiv preprint arXiv:2407.16252. [link]