JO-KU CHENG

Homepage: https://chengruogu0915.github.io/
Email: chengruogu@stu.pku.edu.cn

EDUCATION

Peking University, School of Mathematical Sciences Sep 2022 - Present
M.S. in Applied Mathematics, Specialization in Intelligent Information Processing
Supervisor: Prof. Jinwen Ma GPA: 3.66/4.0
Related courses: Artificial Intelligence, Machine Learning and Optimization for Deep Learning

Beijing Normal University, School of Mathematical Sciences Sep 2018- July 2022
B.S. in Mathematics and Applied Mathematics

RESEARCH INTERESTS

My research interests include **multi-modal reasoning**, **natural language processing**, and **computer vision**, with a particular focus on using synthetic data to enhance the performance of LLMs.

RESEARCH EXPERIENCE

Unified Model for Geometry Problem Solving and Diagram Generation Oct 2024 - Present

Master's Thesis Project (Ongoing)

- Developing a unified multi-modal model that integrates geometric diagram generation, geometric problem solving, and diagram understanding.
- The research employs the VQ-GAN image tokenizer, combining OCR-aware loss and geometry-specific loss functions to optimize the reconstruction of geometric figures and captions. An auto-regressive method is applied to predict the next token for both text and images. Additionally, a set of evaluation metrics is being developed, focusing on isomorphism, structural accuracy, and adherence to mathematical rules in the generated diagrams.

Diagram Formalization Enhanced Geometry Problem Solver [1] Jan 2024 - Sep 2024

- Developed and open-sourced the multimodal reasoning framework, Diagram Formalization Enhanced Geometry Problem Solver (DFE-GPS), which integrates visual features, geometric formal language, and natural language to enhance MLLM performance in solving geometry problems.
- Achieved an accuracy of 82.38% on the publicly available FormalGeo7k dataset, significantly outperforming existing multi-modal and language models, including GPT-4.
- Proposed a synthetic data generation pipeline and created a large-scale dataset SynthGeo228K, improving the vision encoder's ability to extract features from geometric diagrams and enabling LLMs to acquire geometric knowledge.
- Proposed an evaluation method that leverages LLMs to assess the models generated solution process with calculation accuracy, logical coherence, and conciseness.

State-owned Assets and Enterprises LLM[2] Nov 2023 - Sep 2024

- Participated in a collaborative project that proposed a novel training strategy for LLMs tailored for state-owned assets and enterprises (SOAEs), integrating domain-adaptive pre-training with a replay mechanism to prevent catastrophic forgetting and enhancing preference alignment with innovative data scheduling.

Document Layout Analysis [3] Sep 2023 - Apr 2024

- Developed a fusion framework combining Whitespace Smear Cutting (WSC) and Swin Transformer for enhanced Document Layout Analysis (DLA), achieving superior performance in processing Chinese documents.
- Introduced a novel unsupervised segmentation method, WSC, to accurately segment connected content blocks by eliminating whitespace smears, preserving document structure details.
- Leveraged Swin Transformer for semantic segmentation, implementing a continuous training paradigm to adapt the model to specific data distributions, significantly improving semantic classification accuracy.
- Demonstrated the effectiveness of the proposed method on a custom Chinese dataset and the open POD dataset, showcasing its practicality for downstream applications.

AI Smart Monitoring Algorithms for Ships Jul 2023 - Nov 2023

- Took responsibility for developing two key functions: long-term sitting detection for personnel and cargo hold cover monitoring. Applied object detection algorithms to improve model performance on real datasets and deployed the model to edge devices on ships.

Wild Animal Object Detection and Video Classification [4] Dec 2021 - Jul 2022

- Achieved automatic object detection and video classification for 17 species of wild animals based on recordings collected from the Northeast China Tiger & Leopard National Park, using deep learning object detection models. This significantly reduced manual data processing time for the research team at Beijing Normal University's Tiger and Leopard Research Project.
- Collected and organized wild animal datasets, including image and video data, and performed data preprocessing and annotation. Built and trained the YOLOv5 model based on the PyTorch framework to implement wild animal target detection and video classification.

PUBLICATIONS

Accepted and Published * Co-first authors.

[1] Zeren Zhang*, **Jo-Ku Cheng***, Jingyang Deng, Lu Tian, *et al.*, "Diagram formalization enhanced multi-modal geometry problem solver," in *International Conference on Acoustics, Speech and Signal Processing*, 2025.

[2] Jingyang Deng, Zeren Zhang, **Jo-Ku Cheng**, and Jinwen Ma, "Enhancing large language models on domain-specific tasks: A novel training strategy via domain adaptation and preference alignment," in *International Conference on Acoustics, Speech and Signal Processing*, 2025.

[3] Ran Chen*, **Jo-Ku Cheng***, and Jinwen Ma, "A fusion framework of whitespace smear cutting and swin transformer for document layout analysis," in *International Conference on Intelligent Computing*, Springer, 2024, pp. 338–353.

[4] Mengyu Tan*, Wentao Chao*, **Jo-Ku Cheng**, Mo Zhou, *et al.*, "Animal detection and classification from camera trap images using different mainstream object detection architectures," *Animals*, vol. 12, no. 15, p. 1976, 2022.

[5] Zeren Zhang*, Haibo Qin*, Jiayu Huang, **Jo-Ku Cheng**, *et al.*, "Swaptalk: Audio-driven talking face generation with one-shot customization in latent space," in *International Conference on Acoustics, Speech and Signal Processing*, 2025.

INTERNSHIPS

01 AI Jan 2024 - Jul 2024

- Mentor: Lu Tian
- Contributed to the MLLM-based Diagram Formalization Enhanced Geometry Problem Solver project.
- Worked on the LLM SOAE project for domain-specific large language model training.

Institute of Information Engineering, Chinese Academy of Sciences July 2023 - Aug 2023

- Mentor: Associate Prof. Hua Zhang
- Participated in the Chat Privacy project, aimed at protecting user privacy. Applied deep learning techniques such as semantic segmentation, mask repair, and text-to-image models to automate tasks like batch segmentation, image erasure, and redrawing based on textual prompts.

Sexual Health Education Research Group, Beijing Normal University July 2021 - July 2022

- Mentor: Prof. Wenli Liu
- Participated in collaborative projects within the Sexual Health Education Research Group, aiming to promote public awareness on the importance of sexual health and sex education in China. These projects included producing content for groups social media accounts, such as street interviews, informational videos and articles on WeChat account.

ACHIEVEMENTS & AWARDS

Ministry of Education Taiwan Student Scholarship (Second Class Award)	2024
2022 MDPI Best Paper Award [4]	2024
Ministry of Education Taiwan Student Scholarship (Third Class Award)	2022, 2023
Outstanding Undergraduate Thesis, School of Mathematical Sciences, BNU	2022
Ministry of Education Taiwan Student Scholarship (Third Class Award)	2020, 2021
Second Prize in Beijing Universities Mathematical Modeling Campus Competition	2020

ADDITIONAL SKILLS

Programming Languages	Python, MATLAB
Machine Learning Tools	Pytorch, Sklearn, Pandas, Numpy
Language	English (Fluent, TOEFL IBT score 94); Mandarin Chinese (Native) Turkish (Beginner)
Other	boxing, CrossFit, baking