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; Pattern Recognition; Optimization for Deep Learning; Intelligent Computing; Meta-Learning Seminar

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

RESEARCH INTERESTS

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

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 (ICASSP 2025)*, 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 (ICASSP 2025)*, 2025.

[3] 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 (ICASSP 2025)*, 2025.

[4] 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 (ICIC 2024)*, Springer, 2024, pp. 338–353.

[5] 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.

RESEARCH EXPERIENCE

Unified Model for Geometry Problem Solving and Diagram Generation Oct 2024 - Present
Master's Thesis Project at Peking University, Supervised by Prof. Jinwen Ma (Ongoing)

· Develop a unified multi-modal model that integrates geometric diagram generation, geometric problem solving, and diagram understanding.

Diagram Formalization Enhanced Geometry Problem Solver [1] Jan 2024 - Sep 2024
Supervisor: Prof. Jinwen Ma and Lu Tian Peking University and 01AI

· Developed the multi-modal 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 accuracy 82.38% on the public FormalGeo7k dataset, significantly outperforming existing MLLMs and LLMs, 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.
- The model and the SynthGeo228k dataset can be found on Hugging Face as open-source resource for public use.

State-owned Assets and Enterprises LLM[2] Nov 2023 - Sep 2024
Supervisor: Prof. Jinwen Ma and Lu Tian Peking University and 01AI

- Participated in a collaborative project that proposed a novel training strategy for LLMs tailored for state-owned assets and enterprises (SOAEs).
- Proposed an improved domain-adaptive strategy in pre-training with a replay mechanism to prevent catastrophic forgetting, ensuring that the model not only acquires domain-specific knowledge but also retains its general language abilities.
- Utilized a portion of domain-specific data for supervised fine-tuning (SFT) and integrated low-quality data with the remaining SFT data to create tailored preference datasets using the Kahneman-Tversky Optimization technique.

Document Layout Analysis [4] Sep 2023 - Apr 2024
Supervisor: Prof. Jinwen Ma Peking University

- Developed a fusion framework combining Whitespace Smear Cutting (WSC) and Swin Transformer for Document Layout Analysis, achieving superior performance in processing Chinese documents, showcasing its practicality for downstream applications.
- Introduced a novel unsupervised segmentation method, WSC, to accurately segment connected content blocks by eliminating whitespace smears, preserving document structure details.
- Utilized Swin Transformer for semantic segmentation and applied transfer learning to adapt the model to specific data distributions, enhancing semantic segmentation performance.

Wild Animal Object Detection and Video Classification [5] Dec 2021 - Jul 2022
Supervisor: Senior Engineer Lian Yu and Prof. Limin Feng Beijing Normal University

- Participated in a interdisciplinary project that completed automatic object detection and video classification for 17 species of wild animals, using deep learning object detection models based on recordings collected from the Northeast China Tiger & Leopard National Park. This reduced manual data processing time for the research team at BNU's Tiger and Leopard Research Project.
- Collected and organized wild animal datasets, including images and videos, and performed data preprocessing and annotation. Built and trained the YOLOv5 model for wild animal detection and video classification, achieving 97.2%-97.9% precision and 97.1%-97.8% recall rate, demonstrating the most robust performance among all models.

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 State-owned Assets and Enterprise LLMs training.
- Engaged in technical exchanges with the company's engineering team, sharing insights and collaborating on project improvements.

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

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

Sexual Health Education Research Group, Beijing Normal University Jul 2021 - Jul 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 the group's social media accounts, such as street interviews, informational videos, and articles on WeChat.

ACHIEVEMENTS & AWARDS

Ministry of Education Taiwan Student Scholarship (Second Class Award)	2024
2022 MDPI Best Paper Award [5]	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), Mandarin Chinese (Native), Turkish (Beginner)
Other	Boxing, CrossFit, Baking