

Ziqi CHEN

ziqichen1@link.cuhk.edu.cn | chenziqiadam.github.io | [LinkedIn Profile](#)

Research Interests: Efficient LLM Architectures, Multi-Agent Systems, and evolving AI systems

EDUCATION

The Chinese University of Hong Kong, Shenzhen (CUHKSZ)

Shenzhen, China

B.S. in Computer Science and Engineering

Sep. 2023 – May 2027 (Expected)

CGPA: 3.68/4.00 | GRE: 329 + 3.5 | TOEFL: 111

PUBLICATIONS & PREPRINTS

DocMaster: A Hierarchical Structure-Aware System for Document Analysis

Ziqi Chen, Yingli Zhou, Fangyuan Zhang, Quanqing Xu, Chuanhui Yang, Yixiang Fang.

Under Review. 2026. | [Paper Website](#)

RESEARCH EXPERIENCE

FreedomAI Lab, CUHKSZ

Advisor: Prof. Benyou Wang and Dr. Yan Hu

- **AdaLLM: An Adaptive LLM Architecture for Inference Adaptation** Feb. 2026 – Present
 - Inspired by *Nested Learning* HOPE architecture and proposed a parameter-efficient fine-tuning architecture that enables task-specific specialization during inference without modifying model parameters.
 - Designed a selective unfreezing mechanism for context-aware memory module layers at inference time, eliminating the need for costly full fine-tuning while preserving generalization.
- **GRouter: A Learned Model Routing System for Multi-Agent Systems** Dec. 2025 – Present
 - Developed a learned model routing system for multi-agent systems that assigns cost-effective base LLM models based on the specific tasks while balancing cost and accuracy.
 - Curated multi-LLM graph data, applied GRPO-based topology pruning, and trained a lightweight MLP router for real-time model selection.
- **MAS-Designer: Optimizing Multi-Agent Systems via GRPO** Oct. 2025 – Dec. 2025
 - Refined *G-Designer* by inducing GRPO and incorporating a utility-based reward function to stabilize the training process. Formulated multi-agent communication topology optimization as a graph pruning problem and solved it via GRPO.
 - Achieved a **9.7%** token cost reduction while preserving **89.1%** accuracy across GSM8K and HumanEval benchmarks.

Prof. Fang's Lab, CUHKSZ

Advisor: Prof. Yixiang Fang and Dr. Yingli Zhou

- **DocMaster: A Hierarchical Structure-Aware Document Analysis System** May. 2025 – Sep. 2025
 - Designed an end-to-end system for document retrieval and filtering, enabling comparison of indexing methods on accuracy and LLM token cost.
 - Implemented three indexing paradigms: structure-aware document tree, constraint-based semantic clustering (PC-Kmeans + LLM-juegded links), and hypergraph retrieval for cross-section semantics.

INDUSTRIAL EXPERIENCE

Founding Engineer, Mortar

Dec. 2025 – Feb. 2026

Built a data-driven AI platform for real-world camera feed analysis at scale on Vercel and Supabase. Secured \$30,000 in startup funding from Anthropic and NVIDIA.

NLP Engineer, Lalamove

May 2025 – Aug. 2025

- Optimized an AI appointment scheduling agent via context engineering and fallback strategies, reducing appointment error rates by **67.7%**.
- Fine-tuned a multilingual BERT model on QA dialogues, achieving **95.7% consistency** with existing services and saving approximately \$50,000 annually.

SELECTED PROJECTS

Narrative Navigator (*Capstone Project at Tencent*)

Developed a narrative planner module for interview-based memoir generation, increasing life-event graph coverage by **13.4%** through structured interview progression.

RVQ Tokenizer

Designed a Reparameterized Vector Quantization Tokenizer that resolves codebook collapse, achieves **100% codebook utilization**, and supports scalable integration with autoregressive models.

Obsidian Plugin: Daily News Briefing

Built a highly customizable AI-powered plugin for generating daily news briefings within Obsidian; 2,100+ downloads.

PROFESSIONAL SERVICE

Reviewer, ICLR 2026 Workshop on AI with Recursive Self-Improvement

Mar. 2026

Presenter, The 1st CUHK(SZ) Undergraduate Research Conference

Feb. 2025

AWARDS & HONORS

- Undergraduate Research Award Mar. 2026
- Dean's List 2023 – 2025
- Undergraduate Student Teaching Fellows 2024 – 2026
- GBA International Programming Contest: First Prize Apr. 2025
- ShanHaiWoo 2025 Scholarship (\$1,000) Aug. 2025
- CUMCM (Contemporary Undergraduate Mathematical Contest in Modeling): Third Prize Sep. 2025
- ICM/MCM: Honorable Mention May 2025

SKILLS

- **Relevant Courses:** Natural Language Processing, Reinforcement Learning, Deep Learning
- **Communities:** CAMEL Contributor, Sentient Ambassador
- **Interests:** Reading, Table Tennis, Swimming