ZHIHUI CHEN

Phone: (+65) 80389703 ◇ Email: zhihui.chen@u.nus.edu
Objective: *LLM Research Intern*
HomePage - https://richardchenzhihui.github.io/ ◇ LinkedIn

EDUCATION

National University of Singapore *Jan 2025 – Dec 2028 (Expected)*
Ph.D. Student in Artificial Intelligence
Research: Multimodal LLM, Medical AI, Trustworthy & Responsible AI

The University of Hong Kong *Sep 2022 – Jul 2024*
M.Sc. in Artificial Intelligence

The Chinese University of Hong Kong, Shenzhen *Sep 2018 – May 2022*
B.Sc. in Statistics, Data Science Stream

RESEARCH INTERESTS

Multimodal Large Language Models • Medical AI & Healthcare Fundation Model • Trustworthy / Alignment Techniques • Reinforcement Learning for Preference Modelling

PUBLICATIONS

- **Zhihui Chen**, Kai He, Yucheng Huang, Yunxiao Zhu, Mengling Feng. "DivScore: Zero-Shot Detection of LLM-Generated Text in Specialized Domains." Proceedings of EMNLP 2025, Main Conference (Top Conference in NLP).
- **Zhihui Chen**, Kai He, Qingyuan Lei, Mengling Feng. "MedForge: Interpretable Medical Deepfake Detection via Forgery-aware Reasoning". Submitted to ACL 2026, (Top Conference in NLP).
- **Zhihui Chen**, Mengling Feng, "Med-Banana-50K: A Cross-modality Large-Scale Dataset for Text-guided Medical Image Editing" Arxiv 2025.

RESEARCH EXPERIENCE

DivScore: Zero-Shot LLM Text Detection Jan 2024 – Present
Lead Researcher, NUS Saw Swee Hock School of Public Health

- Proposed **DivScore**, a zero-shot detection framework leveraging normalized entropy-based scoring and unsupervised domain knowledge distillation to identify LLM-generated text in specialized domains.
- Achieved state-of-the-art performance with a **14.4% improvement in AUROC** and **64.0% higher recall** (at 0.1% FPR) compared to best baseline on medical and legal benchmarks.
- Conducted theoretical analysis on the divergence between human and LLM text distributions, proving the effectiveness of entropy normalization under domain shifts.

MedForge: Interpretable Medical Deepfake Detection Jan 2025 – Present
Lead Researcher, Health Innovation lab at Emory University

- Developed **MedForge-Reasoner**, an MLLM-based detector utilizing a "localize-then-analyze" reasoning mechanism to identify tamperings in medical scans, achieving **99.23% accuracy**.
- Introduced **Forgery-aware Group Sequence Policy Optimization (GSPO)** to enforce grounding on visual evidence, reducing hallucinations by 16.2% and ensuring clinically verifiable explanations.
- Constructed **MedForge-90K**, a large-scale benchmark with 90k images across 19 pathologies (CXR, MRI, Fundus), annotated with hierarchical expert reasoning guidelines.

Med-Banana-50K: Medical Image Editing Foundation Model May 2025 – Present
Lead Researcher, NUS Medicine × National University Hospital

- Constructed **Med-Banana-50K**, the first large-scale dataset for instruction-based medical image editing, containing 50K high-quality editing pairs across Chest X-ray, Brain MRI, and Fundus photography.
- Designed a rigorous quality control pipeline using **LLM-as-Judge** with medical-specific rubrics (fidelity, structural plausibility) and history-aware iterative refinement to ensure anatomical correctness.
- Released 37k negative samples with full conversation logs to support preference learning (DPO) and robustness research in medical multimodal generation.

SELECTED PROJECTS

Smart Word Agent: Intelligent Document Assistant Feb 2024 – Apr 2024
Creator, Principle Developer GitHub Open-source

- Built an agentic pipeline that parses, edits and re-formats MS Word documents through natural-language instructions based on ReAct Agent Principle.
- Supports multi-document context, attachment reasoning, bulk formatting, table manipulation, low latency with Kimi-K2 as core LLM.
- Released single-file portable exe version with 1,000 downloads in 1 week and 100+ stars in GitHub.

Quant Trading Agent: Automated HK Stock Trading May 2024 – Present
Lead Development Scientist AQUMON (Hong Kong)

- Developed an autonomous quantitative trading agent based on **LangChain** architecture to orchestrate market analysis and decision-making workflows.
- Integrated **Futu OpenAPI** to implement real-time market data streaming and automated order execution for Hong Kong stocks, enabling programmatic trading strategies.

Legal Audio Transcription System with Whisper Large-v2 Jan 2025 – Feb 2025
Freelance AI Consulting HAIWEN & PARTNERS LLP

- Delivered GPU-accelerated legal-domain transcription service; deployed with Docker + FastAPI, average WER 7.8% for Haiwen LLP (HK).

TEACHING EXPERIENCE

Teaching Assistant — "Generative AI and LLM" Feb 2024 – Apr 2024
NUS Business School, Prof. Pang Yan

- Designed the course "Generative AI and LLM" curriculum & knowledge framework
- Developed lecture notes and LLM-oriented course projects, created coding tutorials (PyTorch, Huggingface, LLM-pretraining, SFT, RLHF) and coding assignments

AWARDS

- Full Ph.D. Scholarship, National University of Singapore (2025–2029)
- Outstanding Graduate, CUHK-Shenzhen (Top 5 %)
- Undergraduate Research Excellence Award, CUHK-Shenzhen

SKILLS

Coding	Python, R, C++, MATLAB, MySQL
DL Framework	PyTorch, TensorFlow, scikit-learn, HuggingFace
AI Tools	Linux, Docker, Git, Llama.cpp
Languages	English (TOEFL: 106, GRE: 320), Mandarin (Native), Cantonese (Native)
Interests	Photography, Cycling, Violin, Running