ZHENYI SHEN

(+44) 07566802085 | zhenyi.shen@kcl.ac.uk | https://www.zhenyishen.com/

EDUCATION

King's College London

Ph.D. of Computer Science Research, Supervisor: Prof. Yulan He

Research Interests: Natural Language processing, LLM reasoning, LLM efficiency, Multimodal Learning

Imperial College London

Master of Engineering in Electrical and Electronic

Achieved First Class Honour, ranked the top 20% overall (Dean's List 2016).

PUBLICATIONS

- Zhenyi Shen, Hanqi Yan, Linhai Zhang, Zhanghao Hu, Yali Du, Yulan He. "CODI: Compressing Chain-of-Thought into Continuous Space via Self-Distillation." Submitted to ACL 2025.
- Zhanghao Hu, Hangi Yan, Qingling Zhu, Zhenyi Shen, Yulan He, Lin Gui. "Beyond Prompting: An Efficient Embedding Framework for Open-Domain Question Answering." Submitted to ACL 2025.

WORK EXPERIENCES

iFLYTEK Co. Ltd.

Speech Synthesis Researcher

Developed high-quality speech synthesis models for low-resource Chinese dialects (e.g., Taiwanese Mandarin, Suzhou, Shanghai, and Southern Min), achieving MOS >4.0, and conducted in-depth research on a text-analysis module to enhance rhythm, intonation, and overall controllability in dialectal TTS systems.

Zhuofan Information Technology Co. Ltd.

Computer Vision Engineer

Developed and deployed computer vision solutions, including facial recognition, object detection, and classification, to enhance E-Government IT systems. Automated key administrative processes such as surveillance, customer registration, and license issuance, improving efficiency and reducing manual workload.

MediaTek Inc.

Verification Engineer Intern, Supervisor: Dr. Dimitris Nalbantis

Completed the functional verification of MediaTek's next-gen commercial 5G cellular RFIC via unit testing and integration testing of all simulated modules.

RESEARCH EXPERIENCES

Temporally Coded Spiking Neural Networks

Imperial College London, Supervisors: Professor Pier Luigi Dragotti and Dr. Vincent C.H. Leung London, UK Conducted research on Spiking Neural Networks (SNNs). Designed a hybrid SNN architecture that seamlessly integrates temporally-coded and rate-coded modules via a converter, harnessing the advantages of both paradigms to enhance accuracy and efficiency on real-world neuromorphic datasets (N-MNIST, DVS-128).

Undergraduate Research Opportunity

Imperial College London, Supervisors: Professor George A. Constantinides and Dr. James J. Davis London, UK Engineered an FPGA-based device checker, capable of determining the failure frequency of on-chip device, testing up to an upper limit of 800MHz with negligible error margins.

SKILLS

- Programming: Proficient in Python (PyTorch, NumPy, HuggingFace Transformers), Bash, LATEX, Git
- Languages: Fluent in English (IELTS: 7.5), and native in Mandarin.

12.2021-8.2024

Shanghai, China

8.2020-12.2021

Shanghai, China

4.2019-10.2019

Kent, UK

6.2018-9.2018

10.2019-6.2020

London, UK

London, UK 10.2016-7.2020

2024.10-Present