Haolin He

Email: harlandzzc@gmail.com
Address: Jockey Club Postgraduate Hall 1, The Chinese University of Hong Kong
https://harlandzzc.github.io/

EDUCATION

The Chinese University of Hong Kong, Department of Electronic Engineering

Degree:	*Doctor of Philosophy* (Major: Electronic Engineering)	08/2024-07/2028
Supervisor:	Qiuqiang Kong	
Research topic:	Audio Processing, Computational Neuroscience, Artificial Intelligence	

University of Electronic Science and Technology of China, School of Information and Communication Engineering

Degree:	*Bachelor of Engineering* (Honors, Major: Communication Engineering)	09/2020-07/2024
GPA:	3.95/4.0	
Core Courses:	Principle of Communication (99/100), Digital Signal Processing (96/100), Digital Logic Circuit and System (95/100), Stochastic Signal Analysis (94/100), Signals and Systems (93/100), Computer Communication Network (90/100), High-Level Language Programming (89/100), etc.	
Scholarships:	National Scholarship (2022-2023 Academic Year; Rank 2/217)	
	Scholarship for Outstanding Students (2022-2023 Academic Year)	
	Scholarship for Outstanding Students (2021-2022 Academic Year)	
	Scholarship for Outstanding Students (2020-2021 Academic Year)	

INTERNSHIP

Antgroup, Shanghai   04/2025-present

Research topic:	Large Audio Language Models, Post-Training
Achievement:	Antgroup's first championship at DCASE (track: Audio Question Answering), Completed the paper "Measuring Audio's Impact on Correctness: Audio-Contribution-Aware Post-Training of Large Audio Language Models"

REASEARCH PROJECTS

Measuring Audio's Impact on Correctness: Audio-Contribution-Aware Post-Training of Large Audio Language Models
Research Project | Supervisor: Qiuqiang KONG   04/2025-present

- Developed AudioMCQ, a high quality large-scale audio multiple-choice question dataset comprising 571k samples with structured and unstructured chain-of-thought annotations for Large Audio Language Models (LALMs) post-training, and investigated the zero audio-contribution phenomenon where models derive correct answers solely from textual information without processing audio content
- Proposed Audio-Contribution Filtering methodology to partition data into weak and strong audio-contribution subsets, and developed two effective post-training paradigms (Weak-to-Strong and Mixed-to-Strong) achieving state-of-the-art performance on MMAU, MMAR, and MMSU benchmarks
- Achieved first place in the DCASE 2025 Audio-Question-Answering challenge by using AudioMCQ
- Completed the paper (1st author) "Measuring Audio's Impact on Correctness: Audio-Contribution-Aware Post-Training of Large Audio Language Models" which has been submitted to *ICLR 2026*

DeepNuParc: A Novel Deep Clustering Framework for Fine-scale Parcellation of Brain Nuclei Using Diffusion MRI Tractography
Research Project | Supervisor: Fan ZHANG   02/2025-08/2025

- Developed DeepNuParc, a novel deep learning pipeline for automated fine-scale parcellation of brain nuclei (amygdala and thalamus) using diffusion MRI tractography
- Designed innovative streamline clustering-based connectivity features with cluster dilation and Gaussian smoothing, and created an adaptive k-means-friendly autoencoder with dense CNN architecture achieving superior parcellation performance with high spatial continuity and good correspondence to established atlases
- Completed the paper (1st author) "DeepNuParc: A novel deep clustering framework for fine-scale parcellation of brain nuclei

using diffusion MRI tractography" which has been accepted by *NeuroImage* (CAS Q1 Top, JCR Q1 in Neuroimaging)

Musimple: A Simplified Music Generation System with Diffusion Transformer   09/2024-11/2024
Research Project | Supervisor: Qiuqiang KONG
- Proposed a streamlined framework for text-to-music generation using mel spectrograms and a 2D Diffusion Transformer (DiT), avoiding the complexity of Variational Autoencoders (VAEs)
- Designed an efficient text-conditioning mechanism leveraging Sentence Transformer, reducing computational overhead while maintaining semantic relevance
- Developed a fine-tuning strategy with reinitialized positional embeddings to improve model performance for generating coherent long-duration music
- Co-authored (2nd author) the paper "Musimple: A Simplified Music Generation System with Diffusion Transformer," accepted by *ICASSP 2025 Workshop SALMA (Speech and Audio Language Models)*

The Application of Deep Learning in Amygdala Parcellation Based on Diffusion Weighted Imaging (DWI)   06/2023-07/2024
Research Project | Supervisor: Fan ZHANG
- Purely used DWI to avoid T1-assisted segmentation and manual segmentation
- Re-clustered fiber of interest and used fiber dilation convolution and Gaussian smoothing filters to improve data quality
- Designed a K-means-friendly dense autoencoder with a tailored training framework to divide the amygdala into nine subregions
- Completed the paper (1st author) "A Novel Deep Clustering Framework for Fine-scale Parcellation of Amygdala Using dMRI Tractography" which has been accpeted by the *2024 International Symposium on Biomedical Imaging (ISBI2024)*
- Selected as one of the 50 Student Travel Grant Winners worldwide in *ISBI2024*

COMPETITIONS

DCASE 2025 Audio-Question-Answering Challenge   04/2025-07/2025
- Collaborated with a multi-institutional team to develop comprehensive audio question answering systems for the DCASE 2025 Challenge, creating DCASE-AQA-Boost (AudioMCQ) dataset with 571k samples featuring diverse question types (sound, music, speech, temporal)
- Innovatively integrated Supervised Fine-Tuning (SFT) with Group Relative Policy Optimization (GRPO) in a three-stage training
- Achieved first place globally in the DCASE 2025 Audio-Question-Answering Challenge with Kimi-Audio-SFT-12B (78.18% accuracy) and Qwen2-Audio-R1-8B (77.66% accuracy), and completed the technical report (1st author)

The Mathematical Contest in Modeling / The Interdisciplinary Contest in Modeling (MCM/ICM)   02/2023
- Collaborated with a group to develop a novel index for assessing and preventing the risks of light pollution
- Innovatively integrated the EM algorithm from machine learning with the Gaussian mixture model (GMM) and optical flow estimation algorithm from computer vision, achieving impressive results in light pollution assessment during the competition
- Achieved the finalist designation (F award, top 2% approximately) in the competition with over 25,000 teams worldwide

ADDITIONAL INFORMATION

Software and Tools:	MATLAB, Python, Verilog, C, Latex
Language Proficiency:	Chinese (Native), English (IELTS: Overall 7.5)
Awards:	First Prize of DCASE 2025 Audio-Question-Answering Challenge (International-Level)
	Outstanding Graduate of Sichuan Province (Provincial-Level)
	Outstanding Graduate of the University of Electronic Science and Technology of China (University-Level)
	Honors Bachelor, Honors Courses, Honors Research (University-level)
	F Award in the MCM/ICM (Top 2%; International-Level)