Hao Xu

 \blacksquare haoxu.live@outlook.com \blacksquare +44 7536203893 xiao001010.github.io

in linkedin.com/in/hao-xu-xiao

Q github.com/Xiao001010

Education

University of Edinburgh, Edinburgh, UK

MSc in High Performance Computing with Data Science

• Main courses: Thread Programming (OpenMP), Advanced Message Passing Programming (MPI), Accelerated Systems (CUDA), HPC Architecture, Performance Programming, Machine Learning at Scale, Data Management, Practical Software Development

University of Sussex, Brighton, UK

BSc in Computer Science & Artificial Intelligence

Soochow University, Soochow, China

International Bachelor-Sussex Excellence AI Programme at Soochow University

• Main courses: Database, Machine Learning, Neural Networks, Computer Vision, Natural Language Engineering, Program Analysis

EXPERIENCE

 Natural Language Engineering Teaching Assistant University of Sussex Taught second-year undergraduate students concepts and practice in Natural Language Guided students through coding assignments and lab exercises. Provided academic support for students with difficulties. 	Oct, 2023 – Dec, 2023 re Engineering.
Junior Research Associate University of Sussex (Predictive Analytics Lab) Cell cycle stage classification from partial information, Supervised by Dr Ivor Simpson	Jul, $2023 - Sept$, 2023
 Fine-tuned Cellpose for cell segmentation in micrographs images. Used VAE to reconstruct missing channels in cell images. Classified cell cycle stages based on image data through CNN. 	
 AI Research & Development Intern Focus Technology Co., Ltd. Used CNN for prohibited product image recognition. Detected sensitive logos with YOLO. Reproduced Image Matting to enhance item edge sharpness for improved extraction. Trained CLIP model for Ali Tianchi's CVPR 2022 AliProducts Challenge for text-image 	Sept, 2021 – Jun, 2022 e retrieval.
 Data Analysis Intern Nanjing Maiyuchuang IT Co., Ltd Participated in the development of a public opinion analysis system. Scraped and preprocessed news data with Scrapy. Trained BERT models for sentiment analysis and text classification. 	Jun, 2021 – Aug, 2021

Projects

Vision Transformer Training Acceleration | Python, Profiling, Distributed Training Mar, 2025 – Apr 2025

- Performed end-to-end **profiling** and **optimisation** for **ViT** training on Cirrus HPC platform.
- Used **DLProf** and **Nsight Systems** to identify CPU/GPU and memory bottlenecks.
- Enabled AMP, cudnn.benchmark; migrated loss accumulation to GPU.
- Compared **DP**, **DDP**, and **Pipeline** parallelism strategies; achieved 4x throughput boost with no loss in precision.

Parallel Optimization of Brain Simulation System | C, MPI, Distributed Training

- Parallelised brain simulation code using C/MPI and Event-Based Coordination pattern.
- $\bullet\,$ Designed reusable event-based coordination framework, separated mechanism and strategy.
- Used MPI_Allgatherv for inter-process communication; tested scalability on Cirrus cluster.
- Supported single/multi handler mode, dynamically scales FIFO queues to optimise memory and ordering.

Mar, 2025

First-Class Honors

Sept, 2024 - Sept 2025

Sept, 2022 - June 2024

Grade: 84% Sept. 2019 – Jun. 2021

Percolation Model CUDA Parallelization | C++, CUDA

- Ported 2D percolation model to **CUDA**; optimized grid mapping with 16x16 thread blocks.
- Used atomic operations and host-device sync for termination detection.
- Achieved up to 52.8x speedup on NVIDIA V100.

2D Decomposition for Cellular Automaton | C, MPI

- Designed 2D decomposition for cellular automaton using MPI.
- Managed data distribution and non-blocking communication for grid boundaries.
- Evaluated scalability for different problem sizes and process counts.

Paper Retrieval and Analysis System | Flask, SQLite, NLTK, ML, Docker, CI/CD Oct, 2024 – Apr 2024

- Developed academic literature search and analysis platform with Python/Flask and SQLite.
- Designed backend APIs for search and metadata statistics.
- Integrated arXiv API with local caching; used **NLTK VADER** for sentiment scoring.
- Applied **TF-IDF**, **KMeans**, and PCA for topic clustering and visualisation.
- Automated deployment and test with Docker and GitLab CI/CD.

LLM based Review Assistant | Python, LLM, RAG, Langchain, Streamlit Oct, 2023 – June, 2024

- Developed a web application using LLMs (Large Language Models) to generate quizzes to assist users in consolidating knowledge from uploaded books, notes, and lecture audio/videos.
- Processed various file formats via LangChain, including transcribing audio and video with the Whisper model.
- Explored MapReduce and Clustering techniques to manage long text inputs for quiz generation.
- Integrated RAG (Retrieval-Augmented Generation) for feedback on incorrect answers.
- Designed and deployed an interactive interface with **Streamlit**.

Face Alignment and Landmark Detection System | *Python, CV*

- Built the system for detecting 44 facial landmarks with a full dataset and a subset containing only 5 landmarks.
- Implemented Cascade Model: predicted 5 landmarks, then all 44 after rotation correction.
- Performed extensive hyperparameter, data augmentation, model and loss experiments.
- Achieved NME (Normalised Mean Error) as low as 5% after detailed quantitative and qualitative analysis.

SKILLS

Programming: Python, C/C++ Technologies: Linux, LaTeX, Git, MySQL, MongoDB Parallel Computing: OpenMP, MPI, CUDA, SYCL Big Data & ML: Hadoop, Spark, PyTorch Languages: Mandarin, English

Feb, 2025

Nov, 2024 – Dec 2024

Feb, 2023 – May, 2023