Andy Chen

• **London, UK** • **(+44) 7464253195** • **likeyandy1025@gmail.com** • **LinkedIn:** www.linkedin.com/in/likeyandy1025

EDUCATION

M.S. **Tsinghua University (QS World Ranking #17, Top University in China)** Aug 2022 - Jun 2025
Data Science and Information Technology
Thesis: "Topological and Graph Theoretical Data Analysis for Brain Dynamics Studies"
Courses: Machine Learning (A-), Learning from Data (A-), Bayesian Learning (A-), Computational Photography (A)

B.Eng. **The University of Glasgow** Aug 2018 - Jun 2022
Electronics and Electrical Engineering
1:1 First-Class Honours Degree
Courses: Computing System (A1), Programming (A3), Embedded Processors (A1), Digital Communication (A1)

B.Eng. **University of Electronic Science and Technology of China (UESTC)** Aug 2018 - Jun 2022
Electronic Information Engineering
GPA: 3.99/4.0, ranked 2/231
Excellent Graduate Award, Excellent Bachelor's Degree Thesis, National Scholarship
Courses: Python Data Analysis (92), Signals and Systems (99), Probability (93), Calculus (97), Digital Logic (91)

WORK EXPERIENCE

Algorithm Developer | AI, (M)LLM Training & Eval, 10+ Projects Apr - Sep 2025
Tencent (China's leading internet company that developed WeChat)

- **LLM Chatbot Song Information Detection | LLM Fine-Tuning, SFT/RLHF**
  Design training datasets in **OpenAI format** from user logs for detecting song info (title, artist, etc.) from chat histories.
  Fine-tune **Qwen** and **GPT-OSS** base models using **LLaMA Factory** and **EasyR1** frameworks.
  Song information detection reached **95%+ accuracy** with **Qwen1.7B/4B** models.
- **VLM Chatbot Meme Understanding and Collection | MCP, RAG, API, VLM Eval**
  Automate collection of **2000+** memes from the internet using **RAG** search with **MCP** tools and dataset **APIs**.
  Meme search quality control using computer vision models for face detection and MCP settings for relevance scores.
  Interpret memes using **VLM (Gemma) inference**, annotating the suitable context, meme purpose, etc.
- **VLM Food Datasets and Calorie Detection | VLM Traning & Eval, CV**
  Fine-tuned and benchmarked 10+ VLM base models (Qwen-VL, LLaVA, intern-VL, Kimi-VL, Yi-VL, etc.) on 20+ datasets with LLaMA Factory, VLLM, EvalScope, and VLMEvalKit, exploring 10+ parameters/config settings (lora rank, gradient accumulation step, epoch, etc.).
  Implement CV methods (YOLO-World etc.) to compare with VLM detection accuracy
- **VLM & Diffusion Model Visual Effect Generation | GenAI, Prompt Eng., Eval**
  Evaluate the image and video generation of MLLM and diffusion models for visual special effects.
  Tested object detection, math datasets, and video question answering performances of Vision-Language Models (VLMs).
- **MLLM Prediction of Movie Watching Brain Response**
  Extract video features using **MLLMs** and **vision-transformer (ViT)** models using **PyTorch**.
  Predict brain response (fMRI) using regression models.

RESEARCH PROJECTS

- **Dynamic Functional Connectivity Analysis using Topological and Graph Theoretical Methods**
  [Conference Paper] [Oral Presentation] | **Comp. Neuro., Nilearn, Python, Sklearn, Pandas, MATLAB**
  Applied graph theory and topological data analysis to extract brain networks in MATLAB.
  Extracted graph/simplicial features of brain time-point networks in Python/Jupyter.
  Processed fMRI data of schizophrenia, bipolar disorder, autism spectrum disorder, etc. using Nilearn
  Examined group differences of metric distributions using statistical testing.
  Discovered **statistically significant ($p<0.05$)** biomarkers, backed by previous study findings.
- **Multi-View 3D Reconstruction** [Conference Paper] | **CV, PyTorch, C++, 3D Modeling**

Generated and processed 3D point cloud data for multiview reconstruction datasets using **Blender** and **CloudCompare**.
Enhanced sparse view 3D reconstruction using shape prior representations with **RANSAC** algorithm and **CGAL** package.
Improved the PMVS multiview stereopsis algorithm in **C++** & Reduced the overall mean distance by **22%**.

- **Drone Image Stitching** [Journal Article (Chinese)] | **MATLAB, CV, Img. Processing**
  Proposed a distance-based stitching method to solve ghosting and irregular overlap problems.
  Studied Euclidean Mapping, Optimal Seamline, and Multi-Resolution Image Fusion algorithms.
  Investigated feature detection methods, including SURF and SIFT, and implemented them in MATLAB.

AI EVENTS & PROJECTS

Projects:

- **Encode x Arc DeFi Hackathon | TypeScript, Python, PostgreSQL, Gamma**
  On-Chain Lending Smart Contract App for Freelancer Economy.
- **Granola & V0 Hackathon | React/TypeScript, REST API/Python, V0/Vercel**
  Online meeting software with participant emotion analysis.
- **Website Accessibility Test with Computer-Use Agents | Agents, LangChain, Eval**
  Automate testing of WCAG 2.2 website accessibility using **OpenAI CUA** with web viewing tools.
- **Metaverse and Digital Asset Creator Program - University of Hong Kong | Roblox, Unreal Engine, Unity**
  Interactive 3D digital architectural assets and virtual classrooms.
- **Sentiment Analysis of Restaurant Reviews Using BERT Models | HuggingFace, BERT, RNN, Transformers**
  Sentiment analysis of Meituan dataset. Visualized attention and word importance.
- **Obstacle Course Navigation Car Robot | Embedded C, OpenMV, microprocessors**
  Built a robot car for obstacle courses using processors, sensors, cameras, motors, and antenna communication modules.

Events:

NH(A)I Global Summit - London; AI for Science: Scientists to Builders; Weights & Biases Fully Connected London 2025; GenZ Immersive Technologies Summit; GenAI London 2025; International Conference of Brain Informatics Bangkok 2024

SKILLS

AI Frameworks:	PyTorch	TensorFlow	Transformers	HuggingFace	(M)LLMs
AI Techniques:	Prompt Eng.	NLP	Neural Networks	API Integration	Software Dev.
Agentic Systems:	MCP	A2A	RAG	Vector DBs	LangChain
LLM Ecosystem:	Qwen	GPT-OSS	Gemma	InternVL	LLaMA
LLM Training Tools:	LLaMA-Factory	EasyR1	VLLM	SFT/RLHF	EvalScope
Programming:	Python	C/C++	Java	Linux	Git
Data Science:	NumPy	Pandas	Scikit-learn	MATLAB	
Data Analytics:	SQL	Tableau	Power BI	R	Excel
Frontend:	React	TypeScript	HTML/CSS	Swift	Bootstrap
Hardware & Embedded:	LTSpice	Multisim	Keil	Verilog	Assembly
Neuroscience:	Nilearn	Conn. Workbench	SPM		
3D & Design:	Unity	Roblox	Unreal Engine	CloudCompare	Meshlab