Yi Pan — Curriculum Vitae

CONTACT INFORMATION

School of Computing,
The University of Georgia
Athens, Georgia, U.S.

Email: ypan24@uga.edu
Personal Website
Mobile: +1(706)614-8093

SUMMARY

Second-year Ph.D. candidate at the University of Georgia focusing on **multi-modal large language model, multi-agent system, quantum AI, and their applications on medical AI**. My research aims to advance the frontiers of **multi-agent systems** and **quantum AI** with a dedicated focus on transforming **healthcare and medical intelligence**. I specialize in developing **intelligent agentic workflows, quantum-enhanced machine learning, and scalable AI systems** to solve complex clinical and healthcare challenges. I seek to contribute to **multi-agent workflows**, **quantum AI**, and their **applications on medical AI**. I have co-authored over 30 papers (including IEEE TBME, TNNLS, QAI, AAAI QIML, EMNLP, and ICLR) with 1700+ citations.

RESEARCH INTERESTS

- **Core Areas:** Multi-agent Systems, Quantum AI, Medical AI, Healthcare Intelligence, Efficient and Trustworthy AI
- **Methods:** Agentic Workflow Orchestration, Quantum Machine Learning, Foundation Model Adaptation (SFT/RLHF), Multimodal Representation, Scalable AI Infrastructure

EDUCATION

School of Computing, 2024 - Now
The University of Georgia, Athens
Georgia, U.S.
Ph.D.
Computer Science

James Watt School of Engineering, 2020 - 2024
University of Glasgow
Glasgow, UK
*Bachelor of Engineering (**First Class Honours**)*
Electronics and Electrical Engineering
GPA: 3.87/4.00 **(TOP 10%)**

Glasgow College, 2020 - 2024
University of Electronic Science and Technology of China
Chengdu, China
Bachelor of Engineering
Electronic Information Engineering
GPA: 3.87/4.00 **(TOP 10%)**

SELECTED PROJECT EXPERIENCES

• **Biomedical Multi-agent Systems via Foundation Model Adaptation** May 2025 – Present
Project Lead
Massachusetts General Hospital (MGH) and Harvard Medical School (HMS)
My Role: Led the development of an **agentic AI system** that leverages **BiomedGPT** as a foundational model for biomedical knowledge acquisition, domain-specific adaptation, and autonomous fine-tuning.

- Designed an **automated fine-tuning pipeline** that enables the foundation model to self-adapt across multiple biomedical subdomains (e.g., radiology, genomics, pharmacology) using public benchmark datasets.
- Implemented a modular **agent orchestration framework** for dynamically spawning domain-specific agents capable of continual learning, evaluation, and integration

into a larger biomedical reasoning system.

- Developed **cross-domain coordination protocols** among agents, allowing for inter-agent communication and knowledge transfer via shared latent representations.
- Established large-scale experimental infrastructure, integrating **distributed fine-tuning, multi-agent communication, and automatic evaluation workflows**.
- Collaborating with clinicians and computational scientists at MGH/HMS to align model objectives with biomedical reasoning and hypothesis testing needs.

Outcome: Building a scalable **biomedical multi-agent system** capable of autonomous adaptation and cross-domain generalization, grounded in physics- and biology-informed foundation models. The system aims to advance the next generation of **AI for Science** in biomedical discovery.

• **LLM-Driven Discovery of Energy-Efficient and Scalable Neuromorphic Designs** Mar. 2025 – Present

Project Lead

My Role: Led a project that integrates large language models with **neuromorphic hardware design** and scientific computing to accelerate the discovery of energy-efficient architectures.

- Designed a structured **evaluation framework** assessing neuromorphic devices on energy efficiency and scalability metrics to construct a comprehensive performance dataset.
- Developed an **LLM-based reasoning agent** that autonomously proposes design combinations by synthesizing strengths across existing device architectures.
- Implemented iterative feedback loops allowing the agent to refine hypotheses based on performance trade-offs, aligning with physics-informed optimization principles.
- Collaborating with cross-disciplinary teams to integrate domain knowledge from materials science, microelectronics, and AI modeling for interpretable and reproducible results.

Outcome: Established an AI-guided scientific workflow for **neuromorphic device design discovery**, advancing energy-efficient computing through foundation-model-assisted reasoning.

• **Diffusion Brain: Multi-Agent System with Integrated Memory for Scientific Reasoning** Mar. 2025 – Present

Major Contributor

My Role: Developed a biologically inspired **multi-agent system** with integrated long-term memory and auto-regressive sequencing to simulate intelligent reasoning processes.

- Designed the system architecture that coordinates multiple agents through shared memory and dynamic context updating, inspired by cognitive diffusion in human brains.
- Implemented a modular **agent coordination layer** enabling each agent to query, reason, and contribute insights collaboratively to solve domain-specific problems.
- Proposed and prototyped a "**diffusion brain**" mechanism capable of generalizing to broader scientific and engineering tasks through autonomous knowledge retrieval.
- Leading ongoing experiments to apply this architecture to structured scientific design queries, demonstrating cross-domain adaptability and reasoning coherence.

Outcome: Formulated a scalable cognitive-inspired agentic framework that bridges **LLM-driven reasoning and scientific computing**, providing a prototype for AI-augmented scientific exploration.

• **AQCF: Adaptive Quantum–Classical Fusion Transformer for Next-Generation Language Models** Mar. 2025 – May 2025

Project Lead

School of Computing, The University of Georgia

My Role: Led the development of **AQCF**, a hybrid Transformer architecture integrating classical and quantum computation for adaptive, entropy-aware language modeling.

- Proposed an **entropy-driven adaptive circuit** to dynamically adjust quantum depth and gate configuration, mitigating barren plateaus.
- Designed **quantum memory banks** for unified quantum–classical attention and efficient information retrieval.
- Built a **fusion controller** enabling stable training and 35–40% quantum utilization under NISQ constraints.

Outcome: Introduced a scalable framework toward **Quantum-Enhanced LLMs**, accepted at *AAAI Quantum Intelligence in Machine Learning (QIML 2025)*.

• **MolQAE: Quantum–Classical Hybrid Autoencoder for Molecular Representation Learning** Mar. 2025 – May 2025

Project Lead

School of Computing, University of Georgia

My Role: Led the development of **MolQAE**, the first quantum autoencoder to process complete molecular structures for representation learning, integrating principles from quantum mechanics, chemistry, and machine learning.

- Designed a **quantum encoding framework** mapping SMILES molecular sequences directly into parameterized quantum states, preserving structural and chemical semantics.
- Constructed a **hybrid quantum–classical architecture** with encoder–decoder circuits and trainable latent–ancillary qubits, enabling scalable molecular compression and reconstruction under NISQ-era constraints.
- Developed a dual-objective **hybrid optimization strategy** combining fidelity maximization with quantum coherence preservation to achieve stable convergence and high reconstruction fidelity (up to 96.8%).
- Conducted extensive experiments on the QM9 molecular dataset, analyzing compression–fidelity trade-offs and establishing the first **sublinear quantum scaling law** for molecular representation depth.
- Implemented the end-to-end system with TorchQuantum, Qiskit, and PyTorch, ensuring full reproducibility across hybrid simulation environments and GPU–QPU co-training pipelines.

Outcome: Demonstrated the first quantum autoencoder capable of high-fidelity molecular reconstruction and quantum-efficient compression, published in *IEEE QAI 2025*. This work establishes a foundation for **physics- and chemistry-informed AI for scientific discovery** in molecular and biomedical modeling.

• **ChatRadio-Valuer: A Chat Large Language Model for Generalizable Radiology Report Generation** Feb. 2023 – Aug. 2025

Data & System Engineering Lead

Collaborating Institutions: UESTC, NWPU, UGA, and multi-center hospitals

My Role: Led the design and implementation of a scalable **supervised fine-tuning (SFT)** framework for large language models in biomedical domains.

- Conducted extensive **data curation, cleaning, and normalization** across multi-institution and multi-system radiology report datasets to ensure domain and device generalizability.
- Designed the overall **experimental framework and evaluation strategy**, specifying training schedules, ablation protocols, and model selection criteria for multi-institutional data scenarios.

- Implemented a unified **distributed training and evaluation pipeline** in PyTorch, allowing collaborators to reproduce and extend large-scale SFT experiments under heterogeneous environments.
- Executed the majority of large-scale **training and inference experiments**, managing compute scheduling, logging, and model performance tracking across multiple GPUs.
- Collaborated with radiologists to align annotation and diagnostic objectives with model training targets, ensuring clinical relevance and interpretability.
- Contributed to the TBME paper's experimental methodology and result analysis sections, emphasizing **cross-domain robustness, reproducibility, and model scalability**.

Outcome: Developed a multi-institution, multi-system **foundation model** for radiology text generation demonstrating superior generalization and transferability, published in *IEEE Transactions on Biomedical Engineering (TBME), 2025.*

SELECTED PUBLICATIONS

• **Pan, Y.**, Jiang, H., Chen, J., Li, Y., Zhao, H., Zhao, L., Abate, Y., Wang, Y. and Liu, T., Bridging Classical and Quantum Computing for Next-Generation Language Models.
First AAAI Symposium on QIML. 2025

• Jahin, A., **Pan, Y.**, Wang, Y., Liu, T., and Zhang W., Quantum-Classical Hybrid Molecular Autoencoder for Advancing Classical Decoding.
First AAAI Symposium on QIML. 2025

• **Pan, Y.**, Jiang, H., Ruan, W., Zhu, D., Li, X., Abate, Y., Wang, Y. and Liu, T., *MolQAE: Quantum Autoencoder for Molecular Representation Learning.*
IEEE QAI. 2025

• Zhao, H., Li, J., **Pan, Y.**, Liang, S., Yang, X., Dou, F., Liu, T., and Lu, J., HELENE: Hessian Layer-wise Clipping and Gradient Annealing for Accelerating Fine-Tuning LLM with Zeroth-Order Optimization.
EMNLP Main Conference. 2025

• Zhong, T., Zhao, W., Zhang, Y., **Pan, Y.**, Dong, P., Jiang, Z., Jiang, H., Zhou, Y., Kui, X., Shang, Y., et al., *ChatRadio-Valuer: A Chat Large Language Model for Generalizable Radiology Report Generation Based on Multi-institution and Multi-system Data.*
IEEE TBME. 2025

• Liu, Z., Li, Y., Shu, P., Zhong, A., Jiang, H., **Pan, Y.**, Yang, L., Ju, C., Wu, Z., Ma, C., et al., *Radiology-GPT: a large language model for radiology.*
Meta-Radiology. 2025

• Zhong, T., **Pan Y.**, Zhang, Y., Wei, Y., Yang, L., Wu, Z., Liu, Z., Wei, X., Li, W., Yao, J., Ma, C., Han, Y., Li, X., Zhu, D., Jiang, X., Shen, D., Han, J., and Zhang, T., *ChatABL: Abductive Learning via Natural Language Interaction with ChatGPT.*
IEEE TNNLS. 2025

• Ruan, W., Lyu, Y., Zhang, J., Cai, J., Shu, P., Ge, Y., Lu, Y., Gao, S., Wang, Y., Wang, P., Zhao, L., Wang, T., Liu, Y., Fang, L., Liu, Z., Liu, Z., Li, Y., Wu, Z., Chen, J., Jiang, H., **Pan, Y.**, Yang, Z., Chen, J., et al., *Large Language Models for Bioinformatics.*
Quantitative Biology. 2025

• **Pan, Y.**, Jiang, H., Chen, J., Li, Y., Zhao, H., Zhou, Y., Shu, P., Wu, Z., Liu, Z., Zhu, D., Li, X., Abate Y., and Liu T., *EG-SpikeFormer: Eye-Gaze Guided Transformer*

on Spiking Neural Networks for Medical Image Analysis.
IEEE ISBI (Oral Presentation). 2025

• Li, Y., Kim, S., Wu, Z., Jiang, H., **Pan, Y.**, Jin, P., Song, S., Shi, Y., Liu, T., Li, Q. and Li, X., *ECHOPulse: ECG Controlled Echocardio-gram Video Generation.*
ICLR. 2025

• Zhong, T., Liu, Z., **Pan, Y**., Zhang, Y., Zhou, Y., Liang, S., Wu, Z., Lyu, Y., Shu, P., Yu, X., et al., *Evaluation of OpenAI o1: Opportunities and Challenges of AGI.*
Arxiv. **Co-first Author** 2024

• Zhang, Y., **Pan, Y**., Zhong, T., Dong, P., Xie, K., Liu, Y., Jiang, H., Liu, Z., Zhao, S., Zhang, T., Jiang, X., Shen D., Liu T., and Zhang X., *Potential of Multimodal Large Language Models for Data Mining of Medical Images and Free-text Reports.*
Meta-Radiology. **Co-first Author** 2024

• Chen, Y., Xiao, Z., **Pan, Y**., Zhao, L., Dai, H., Wu, Z., Li, C., Zhang, T., Li, C., Zhu, D. and Liu, T., Mask-Guided Vision Transformer for Few-Shot Learning.
IEEE TNNLS. 2024

• Xiao, Z., Chen, Y. , Yao, J., Zhang, L., Liu, Z., Wu, Z., Yu, X., **Pan, Y.**, Zhao, L., Ma, C., Liu, X., Liu, W., Li, X., Yuan, Y., Shen, D., Zhu, D., Yao, D., Liu, T., and Jiang, X., Instruction-ViT: Multi-modal prompts for instruction learning in vision transformer.
Information Fusion. 2024

• Liu Y., He H., Han T., Zhang X., Liu M., Tian J., Zhang Y., Wang J., Gao X., Zhong T., **Pan Y.**, Xu S., Wu Z., Liu Z., Zhang X., Zhang S., Hu X., Zhang T., Qiang N., Liu T., and Ge B., Understanding LLMs: A Comprehensive Overview from Training to Inference.
Neurocomputing. 2024

• Wang, J., Liu, Z., Zhao, L., Wu, Z., Ma, C., Yu, S., Dai, H., Yang, Q., Liu, Y., Zhang, S., Shi, E., **Pan, Y.**, Zhang, T., Zhu, D., Li, X., Jiang, X., Ge, B., Yuan, Y., Shen, D., Liu, T., and Zhang, S., Review of large vision models and visual prompt engineering.
Meta-Radiology. 2023

• Zhao, H., Ling, Q., **Pan, Y.**, Zhong, T., Hu, J.Y., Yao, J., Xiao, F., Xiao, Z., Zhang, Y., Xu, S.H., Wu, S.N., Kang, M., Wu, Z., Liu, Z., Jiang, X., Liu, T., and Shao Y., Ophtha-LLaMA2: A Large Language Model for Ophthalmology.
Arxiv. **Co-first Author** 2023

• Wang, J., Shi, E., Yu, S., Wu, Z., Ma, C., Dai, H., Yang, Q., Kang, Y., Wu, J., Hu, H., Yue, C., Zhang, H., Liu, Y., **Pan, Y.**, Li, X., Ge, B., Zhu, D., Yuan, Y., Shen, D., Liu, T., Zhang, S., Prompt engineering for healthcare: Methodologies and applications.
Arxiv. 2023

HONORS & AWARDS

• NSF Student Travel Award, AAAI FSS25 (QIML) *Nation-level*
• Outstanding Graduate **(Ratio: 10%)**, UESTC *University-level*
• **First Prize** Scholarship for Academic Excellence in Academic Year 2021-2022 **(Ratio: 8%)**, UESTC *University-level*
• Scholarship for English Proficiency in Academic Year 2021-2022 **(Ratio: 6.25%)**, Glasgow College, UESTC *College-level*
• **First Prize** Scholarship for Academic Excellence in Academic Year 2020-2021 **(Ratio: 8%)**, UESTC *University-level*
• Academic Scholarship in Academic Year 2020-2021 (Ratio: 5%, 30,000RMB), Glasgow College, UESTC *College-level*

• **Second Prize** in "NECCS" (National English Competition for College Students) in Academic Year 2020-2021 *Nation-level*
• **Second Prize** in "FLTRP (Foreign Language Teaching and Research Press)—National Talent Cup"—English Writing Contest, Sichuan Division (ranked 32^{nd} in Sichuan Province & the sole Second Prize from UESTC) *Province-level*
• **First Prize** in "FLTRP—National Talent Cup"—Preliminary Contest at School Level, National English Writing Contest (one of the two selected for participating in following contests as the representative of UESTC) *University-level*

ACADEMIC SERVICE

Professional Memberships:
- IEEE Student Member
- AAAI Student Member

Journal and Conference Reviewer:
- ***Journals***
  - IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
  - IEEE Transactions on Artificial Intelligence (TAI)
  - Frontiers in Oncology
  - European Journal of Radiology Artificial Intelligence
- ***Conference***
  - International Conference on Learning Representations (ICLR) 2026
  - International Conference on Machine Learning (ICML) 2025

SKILLS

Languages: Python, C/C++, MATLAB, Bash
DL/LLM: PyTorch (*DDP*), DeepSpeed, Transformers
Scientific Computing: NumPy, SciPy, RDKit, TorchQuantum
Experimentation: multi-GPU training, experiment tracking, reproducible pipelines
Quantum: CUDA-Q, PennyLane, Qiskit
Focus: Physics-/chemistry-informed modeling, multimodal LLMs, agentic workflows, efficient fine-tuning/inference

INTERNSHIP

Graduate Research Intern — Massachusetts General Hospital and Harvard Medical School
May 2025. - Aug. 2025 — Boston, U.S.

TEACHING EXPERIENCE

• **Teaching Assistant** — School of Computing, UGA
Aug. 2025. - Now — Athens, U.S.

• **Teaching Assistant** — Glasgow College, UESTC
Sep 2023. - Jun. 2024 — Chengdu, China

RELEVANT PROGRAMME

• **Artificial Intelligence Internship Programme** — *Business AI Lab NTU*
Distinction Grade

• **Artificial Intelligence and Public Health** — *UCLA*
Project-based Learning

• **Introduction to Data Analytics** *IBM*
Coursera Online Certificate

• **Introduction to Programming with MATLAB** *Vanderbilt University*
Coursera Online Certificate