Yang Qiao

Atlanta, GA | yang.qiao@emory.edu | Google Scholar | LinkedIn | Homepage

Education

Emory University 2024 - Present
Ph.D. in Computer Science and Informatics, expected 2029 | Advisor: Liang Zhao

Shandong University 2021 - 2024
Master's in Computer Science and Technology | Advisor: Xuemeng Song; co-advisor: Liqiang Nie

Shandong University 2016 - 2020
B.Sc. in Physics

Selected Research and Engineering Projects

LLM Agents and Retrieval

EnterpriseSaga | Causal Dynamics Lab May - Aug 2026
Research Intern | Project & Leaderboard

- Led development of a 1.213-billion-token synthetic enterprise benchmark for reasoning across systems and time. Designed coarse-to-fine company simulation and question-conditioned augmentation to create realistic confounding activities while preserving reference answers.
- Implemented and evaluated seven retrieval pipelines and agents on 966 evidence-grounded questions, using SQLite FTS5 sparse retrieval, Qwen3 embeddings, FAISS indexing, and neural reranking.

AI for Science

XAI-Refine | Collaboration with Emory Radiology and Neurology

- Built an automated brain-age model refinement loop that checks explanations against verified neuroscience literature and converts the evidence into targeted training constraints. Applied minimal corrections to unsupported model behavior, with validation safeguards for predictive accuracy and unintended changes to other explanations.

MolWorld | Collaboration with Merck & Co., Inc.

- Introduced molecule-transfer graphs for sequential molecular optimization. Built a world model that predicts future graph structure by connecting generated molecules through valid transformations, guiding subsequent generation. Evaluated across property optimization and docking tasks.

Profile-Conditioned Latent Dynamics | Collaboration with Emory Psychiatry and Behavioral Sciences

- Proposed a world-model formulation that separates stable psychometric traits from evolving behavioral states. Observed transitions provide training supervision; a profile-conditioned prior predicts transition representations at inference without future-state access.

Multimodal Learning

RCML | Relation-Conditioned Multimodal Learning | Emory

- Developed multimodal representations conditioned on natural-language relations, allowing the same sample to encode different aspects of relevance. Designed a unified contrastive objective that combines image-text alignment with many-to-many relations across samples, extending representation learning beyond isolated image-text pairs.

MILNet | Multimodal Sarcasm Detection | Shandong

- Designed complementary modules for local semantic-guided and global incongruity learning to capture conflicts within and across text and images. Connected them through selective mutual knowledge distillation, combining focused local evidence with broader context to address irrelevant content and incomplete semantic cues.

Technical Skills and Languages

Programming languages: Python, C++
Deep learning: PyTorch, Hugging Face Transformers
Machine learning and data: scikit-learn, NumPy, pandas, SciPy, FAISS
Languages: Mandarin Chinese (native); English

Publications and Preprints

LLM Agents and Retrieval

EnterpriseSaga: A Billion-Token Benchmark for Evidence Retrieval and Reasoning over Evolving Enterprise Knowledge
Yang Qiao, Xuchao Zhang, Liang Zhao, Bowen Zhu, and Hasibul Haque.
Preprint, 2026.

AI for Science

XAI-Refine: An Automated Explanation-Knowledge Loop for Brain-Age Prediction
Yang Qiao, Junjie Wu, Deqiang Qiu, James J. Lah, and Liang Zhao.
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2027.

MolWorld: Molecule World Models for Actionable Molecular Optimization
Yang Qiao, Bo Pan, Hao-Wei Pang, Peter Zhiping Zhang, Liying Zhang, and Liang Zhao.
Preprint, 2026.

Profile-Conditioned Latent Dynamics for Human Behavior Simulation
Ziyuan Qin, **Yang Qiao**, Summer B. Frandsen, Michael T. Treadway, and Liang Zhao.
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2027.

Retrieval-Augmented Language Models for Matched Molecular Pair Transformations to Recapitulate Medicinal Chemistry Intuition
Bo Pan, **Yang Qiao**, Peter Zhiping Zhang, Hao-Wei Pang, Alex Zhu, Xiang Yu, Liying Zhang, and Liang Zhao.
Findings of the Association for Computational Linguistics: EMNLP 2026.

Multimodal Learning

Mutual-Enhanced Incongruity Learning Network for Multi-Modal Sarcasm Detection
Yang Qiao, Liqiang Jing, Xuemeng Song, Xiaolin Chen, Lei Zhu, and Liqiang Nie.
AAAI Conference on Artificial Intelligence (AAAI), 2023.

Multimodal Representation Learning Conditioned on Semantic Relations
Yang Qiao, Yuntong Hu, Bowen Zhu, Hasibul Haque, and Liang Zhao.
Preprint, 2025.

Teaching

Emory University | Teaching Assistant
CS 325: Artificial Intelligence Spring 2026; Fall 2025
CS 253: Data Structures and Algorithms Spring 2025