

Qiao SUN

🏠 [qiaosun22.github.io](https://github.com/qiaosun22) ✉ qiaosun22@m.fudan.edu.cn 🌐 github.com/qiaosun22

📍 220 Handan Rd., Yangpu, Shanghai, China ☎ (1)213-583-9491 💬 WeChat

EDUCATION

Massachusetts Institute of Technology

Cambridge, MA, U.S.

🎓 *Visiting Student Researcher*

Jul. 2024 - Dec. 2024

Research Focus: 4D World Model, Embodied AI

Advised by Prof. Chuang Gan & Dr. Yilun Du

Fudan University

Shanghai, China

🎓 *M.Eng.* in Electrical and Computer Engineering | GPA: 3.48/4.0

Sep. 2022 - Jun. 2025

Research Experiences: Natural Language Processing, Computer Vision, Multimodal Learning, Nursing Robots, Memristor-based Analog Computing

Highlighted Courses: Computer Vision (A, 4.0), Robotics (A, 4.0), Frontiers in Intelligent Robotics (A, 4.0), Applied Mathematical Methods (B+, 3.3), Data Science (B+, 3.3)

Advised by Prof. Shijie Guo, Prof. Nanyang Ye & Dr. Qinying Gu

Tianjin University

Tianjin, China

🎓 *B.Eng.* (Major) in Civil Engineering | GPA: 85.10/100

Sep. 2015 - Jun. 2019

Highlighted Courses: Advanced Mathematics (93), Selected Explanation of Mathematical Methods (96), Introduction to College Physics (93), Probability Theory and Mathematic Statistics (92), Basic Techniques of Electrical Engineering (91), Geographic Information System and Engineering (91)

🎓 *B.Mgmt.* (Minor) in Financial Management

Sep. 2016 - Jun. 2019

RESEARCH INTRESTS

- [1] 4D Reconstruction/Generation & 4D Physics-Informed World Model
- [2] Vision-Language-Action Learning & Embodied AI
- [3] 4D Large Vision-Language Models

PUBLICATIONS & PREPRINTS (BY YEAR)

2024

[1] **Learning 4D Embodied World Models**

Under review at *CVPR 2025* [Preprint] [Website]

Haoyu Zhen*, **Qiao Sun***, Pengxiao Han, Yilun Du, Chuang Gan

[2] **VL-Rotate: Vision Model Modulated by Language Model for Few-Shot OoD Rotated Object Detection**

Under review at *CVPR 2025* [Preprint]

Weihan Yin, **Qiao Sun**, Lin Zhu, Liuja Yang, Nanyang Ye.

[3] **MiniConGTS: A Near Ultimate Minimalist Contrastive Grid Tagging Scheme for Aspect Sentiment Triplet Extraction**

Accepted by Main Conference of EMNLP 2024 [Paper] [Code] [PaperWithCode]

Qiao Sun, Liuja Yang, Minghao Ma, Nanyang Ye, Qinying Gu

*Equal Contributions

[4] **Enhancing Nursing and Elderly Care with Large Language Models: A Framework for AI-Driven Patient Monitoring and Interaction**

Accepted by *COLING 2025* [Preprint]

Qiao Sun, Nanyang Ye, Qinying Gu, Jiexin Xie, Shijie Guo

2023

[5] **DV2DM: A Learning-based Visible Difference Predictor for Videos**

Submitted to *TPAMI* [Preprint]

Qi Fan*, Qiao Sun*, Nanyang Ye, Qinying Gu

[6] **Synergistic Development of Perovskite Memristors and Algorithms for Robust Analog Computing**

Submitted to *Nature Communications* [Preprint]

Qinying Gu*, Qiao Sun*, Yifei Wang, Liujia Yang, Nanyang Ye, Huaqiang Wu

RESEARCH EXPERIENCE

Learning 4D Embodied World Models [1]

Research During Visiting at MIT.

June. 2024 - Present

- Dataset Contribution: Harnessing more than 500 GPUs to synchronize and handle the parallel data synthesis procedure.
- Reconstruction Enhancement: Developing novel augmentation techniques to enhance the reconstruction effectiveness and efficiency.
- Contribution: Enabling a quick consistent 4D generation, a quick grasp of dynamic 3D embodied scenarios from a single monocular image, as well as a quick integration with existing foundation language methods.
- Outcome: Submitted a paper to ICLR 2025.

VL-Rotate: Vision-Language Learning for Few-Shot OoD Rotated Object Detection [2]

Participatory Research.

Jan. 2024 - Jun. 2024

- Implemented VL-Rotate to enhance few-shot OoD rotated object detection via vision-language learning, maximizing the semantic potential of encoder models for image-text alignment.
- Integral in the development and incorporation of OOTA and FRTC, bolstering the model's fine-tuning and detection precision.
- Conducted comprehensive evaluations against challenging few-shot OoD datasets, achieving a groundbreaking 75.2% mAP on HRSoD's OoD subset.

MiniConGTS: A Near Ultimate Minimalist Contrastive Grid Tagging Scheme for Aspect Sentiment Triplet Extraction [3]

Independent research.

Oct. 2023 - Feb. 2024

- Proposed the first critical evaluation of the 2D tagging scheme, particularly focusing on the table-filling method. This analysis pioneers in providing a structured framework for the rational design of tagging schemes.
- Introduced a simplified tagging scheme with the few number of label categories to date, integrating a novel token-level contrastive learning approach to enhance PLM representation distribution.
- Demonstrated superior performance over state-of-the-art techniques and Large Language Models like GPT-3.5 and GPT-4 in few-shot learning and Chain-of-Thought prompting scenarios.

Enhancing Nursing and Elderly Care with Large Language Models: A Framework for AI-Driven Patient Monitoring and Interaction [4]

Master's Thesis Project at Fudan University.

Mar. 2024 - Aug. 2024

- Comprehensive Dataset Construction: Curated datasets encompassing high-fidelity text corpora, domain-specific single and multi-turn dialogues, multiple-choice metric dataset, and annotated

*Equal Contributions

image-text pairs, optimizing training for nursing-specific tasks.

- Incremental Pretraining and Mixed Data Finetuning: Implemented a multi-stage finetuning process using a mix of general, domain-specific, and multimodal data to incrementally train large language models (LLMs).
- Chain-of-Thought and Model Architecture: Leveraged a novel model architecture where the LLM functions as the central processing unit, integrating vision encoders and other tools through reasoning.
- Movement and Path Planning: Developing advanced motion planning algorithms to address open-task challenges in nursing robots.

DV2DM: A Learning-based Visible Difference Predictor for Videos [5]

Intern at Shanghai AI Lab.

Sep. 2023 - Mar. 2024

- Developed innovative techniques for fine-tuning large language models (LLMs) and in-context learning to enhance the design of healthcare robots.
- Created tailored metrics and benchmarks to improve model evaluation.
- Collected, processed, cleansed, and validated cutting-edge multimodal datasets.
- Led the project from inception to completion, overseeing all research, development, and procurement activities to ensure seamless integration and execution of project objectives.

Synergistic Development of Perovskite Memristors and Algorithms for Robust Analog Computing Leveraging Bayesian Optimization [6]

Intern at Shanghai AI Lab.

Apr. 2023 - Dec. 2023

- Engaged in the design of a robust architecture for neural network memristors, utilizing Bayesian algorithms.
- Led the experimental efforts, conducting extensive testing of our algorithms across multiple sub-tasks in domains such as computer vision, natural language understanding, 3D perception, and Large Language Models (LLMs).
- Developed the MemSim framework based on PyTorch for simulating memristor computations, enhancing the efficiency and accuracy of hardware emulation.
- Collaborated on the optimization of memristor device fabrication processes to improve performance.

INTERNSHIP

MIT-IBM Watson AI Lab

Jul. 2024 - Present

Supervisor: Prof. Chuang Gan & Dr. Yilun Du

4D World Model, Embodied AI.

JHCCS*, Shanghai Jiao Tong University

Jul. 2023 - Jul. 2024

Supervisor: Prof. Nanyang Ye

Natural Language Processing, Computer Vision, Multimodal Learning.

Shanghai AI Lab

Mar. 2023 - Dec. 2023

Supervisor: Dr. Qinying Gu & Prof. Tianfan Xue

Memristor-Based Analog Computing and Artificial Intelligence.

Western Securities, R&D Center

Sep. 2019 - Jul. 2020

Supervisor: Yumeng Zhang

Data-Driven Financial Engineering, Fund-of-Fund (FoF) Investment Strategies.

SERVICES

Reviewer for ACL ARR October

Nov. 2024

Volunteer at EMNLP, Miami

Nov. 2024

* John Hopcroft Center for Computer Science

HONOR & AWARD

First-class Academic Scholarship at Fudan University	2024
Outstanding Internship Award at Shanghai AI Lab	2023
First-class Academic Scholarship at Fudan University	2023
Asia and Pacific Mathematical Contest in Modeling, Second Prize	2018
Tianjin College Students Innovation and Entrepreneurship Competition, First prize	2017
Henan High School Students' Chemistry Competition, Second Prize	2015
Robot Competition in the National Computer Production Activity, Third Prize	2014

SKILLS

Programmng:

(Proficient) Python (PyTorch, Pillow, OpenCV-python, SciKit-Learn, Open3D, Transformers, etc.), Shell, SLURM, Git, Blender, MuJoCo, L^AT_EX

(Familiar) C/C++, CUDA, MATLAB, HTML/CSS

Math: Matrix Theory, Kolmogorov Probability Theory, Advanced Statistics, Complex Analysis, Differential Geometry

Computer Science: Parallel Computing, Network, Compiling, Architecture