Qiao Sun

⌂ Address: 220 Handan Rd., Yangpu, Shanghai, China
✉ E-mail: qiaosun22@m.fudan.edu.cn
 GitHub: github.com/qiaosun22
 Homepage: josphersun.ink

EDUCATION EXPERIENCE

Fudan University (GPA: 3.42/4.0, Rank: 17/69)   Shanghai, China
M.Eng. in Electrical and Computer Engineering   *Sep. 2022 - Jun. 2025*
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)

Tianjin University (GPA: 85.10/100, Rank: 27/108)   Tianjin, China
B.Eng. (Major) in Civil Engineering   *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*

PUBLICATIONS

[1] Weihan Yin, **Qiao Sun**, *et al.* VL-Rotate: Vision-Language Learning for Few-Shot OoD
Rotated Object Detection. To be submitted to ***AAAI***, 2024. [Paper]
[2] **Qiao Sun**, *et al.* MiniConGTS: A Near Ultimate Minimalist Contrastive Grid Tagging
Scheme for Aspect Sentiment Triplet Extraction. Under Review in ***EMNLP***, 2024. [Preprint]
[3] Nanyang Ye[*], **Qiao Sun**[*], *et al.* Synergistic Development of Perovskite Memristors and
Algorithms for Robust Analog Computing. Under Review in ***Nature Communications***,
2024. [Paper]
[4] Qi Fan, **Qiao Sun**, *et al.* DV2DM: A Learning-based Visible Difference Predictor for Videos.
Submitted to ***TPAMI***, 2024. [Paper]

RESEARCH EXPERIENCE

**Exploration of Physical Dynamics Through Motion Trajectory Prediction Using
Large Language Models and Blender-generated Animations**
Ongoing Independent Research Project.   *June. 2024 - Present*

- Performance Evaluation: Conducting comprehensive evaluation of current large language
  models (LLMs) on predicting motion trajectories from Blender-generated animations, us-
  ing metrics such as Mean Squared Error (MSE) and R-squared (R^2).
- Data Collection: Creating a diverse dataset of Blender animations, including physics-
  based simulations and keyframe animations, to support model training and evaluation.
- Model Enhancement: Fine-tuning LLMs with physics-informed data and integrating phys-
  ical simulation modules for real-time feedback to enhance the model's ability to generate
  realistic motion trajectories.
- Trajectory Optimization: Applying chain-of-thought reasoning and reinforcement learning
  to improve motion trajectory predictions, focusing on ensuring physical plausibility and
  accuracy.
- Evaluation and Metrics: Developing custom evaluation metrics to assess model perfor-
  mance, with a focus on trajectory smoothness, physical plausibility, and prediction accu-
  racy.

Design and Implementation of an Embodied Intelligent Agent for Nursing Robots Using Multimodal Large Models to Enhance Patient Care and Interaction

Master's Thesis Project at Fudan University.   *Mar. 2024 - Present*

- **Comprehensive Dataset Construction**: Curated datasets encompassing high-fidelity text corpora, domain-specific single and multi-turn dialogues, multiple-choice metric dataset, and annotated 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). This approach balances general knowledge retention with specialized nursing expertise, enhancing the model's contextual understanding and task-specific performance.
- **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. This architecture allows for open vocabulary detection, using outputs as vision prompts for subsequent steps, thereby enhancing the robot's reasoning and task execution capabilities.
- **Movement and Path Planning**: Developing advanced motion planning algorithms to address open-task challenges in nursing robots. The approach includes three key steps: (1) extracting spatial constraints from scenes using video language models (VLMs) and converting them into symbolic constraint sets; (2) translating these constraints into scoring functions via program synthesis; and (3) employing kinematic algorithms for trajectory optimization, ensuring task completion while adhering to all constraints.

VL-Rotate: Vision-Language Learning for Few-Shot OoD Rotated Object Detection [4]

Participatory Research in Multimodal Models.   *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 HRSOoD's OoD subset.

MiniConGTS: A Near Ultimate Minimalist Contrastive Grid Tagging Scheme for Aspect Sentiment Triplet Extraction [1]

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.

DV2DM: A Learning-based Visible Difference Predictor for Videos [3]

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 [2]

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 subtasks 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.

AWARD

Outstanding Internship Award (Shanghai AI lab) *2023*
First-class Academic Scholarship (Fudan University) *2023*

INTERN EXPERIENCE

Shanghai Artificial Intelligence Laboratory *Mar. 2023 - Present*
Research Intern on Memristor-based Analog Computing and Artificial Intelligence.

Western Securities R&D Center, Financial Engineering Group *Dec. 2019 - Jul. 2020*
Research Assistant on Data-driven Fund-of-fund Investment Strategies.

SKILLS

- **Programmng**: Python, Shell, C++, PyTorch, scikit-learn, R, Stata
- **Math**: Matrix Theory, Kolmogorov Probability Theory, Probabilistic Graphical Models, Advanced Statistics
- **Hardware**: Memristor-based Neuromorphic Analog Computing, Circuits design
- **Computer Science**: Data Structure, Computer Network, Fundamentals of Compiling, Computer Architecture, Operation System
- **Tools**: PyCharm, VS Code, Visual Studio, Linux, GitHub
- **Language**: Mandarin (Native), English (Fluent)