



Yang Jiao

☎ (+86)13258238851 ✉ yjiao23@m.fudan.edu.cn
💻 <https://scholar.google.com/citations?user=5gA7Wv0AAAAJ&hl=zh-CN>

ACADEMIC EXPERIENCE

Fudan University

Doctor Computer Science and Technology

GPA: 3.57/4.0

Scholarship : National Scholarship; Dong's Enterprise Scholarship

Sep 2021 - Jun 2026

Shanghai

INDUSTRIAL EXPERIENCE

Meituan

Large Multimodal Model Intern

- Develop a large multimodal model equipped with versatile vision-centric capabilities
- Develop a native large multimodal model for unified image generation and visual comprehension

Sep 2023 - Oct 2024

Beijing

Meituan

Algorithm Intern

- Image super-resolution algorithm development and deployment
- Video frame interpolation algorithm development and deployment

Oct 2020 - May 2021

Beijing

SELECTED RESEARCH

Lumen: Unleashing Versatile Vision-Centric Capabilities of Large Multimodal Models

First Author, Accepted to NeurIPS 2024

- Introduce **Lumen**, a LMM with versatile vision-centric functions without significantly compromising general-purpose VQA capabilities.

Fusing LiDAR and Camera at Multiple Scales with Multi-Depth Seeds for 3D Object Detection

First Author, Accepted to CVPR 2023

- Enhance the fine-grained LiDAR-camera fusion with a proposed framework, abbreviated as **MSMDFusion**.
- Rank 1st and 2nd on nuScenes detection and tracking leaderboard among all single-model approaches by the time of submission.
- Github (<https://github.com/SxJyJay/MSMDFusion>) 100+ stars.

MORE: Multi-Order Relation Mining for Dense Captioning in 3D Scenes

First Author, Accepted to ECCV 2022

- Mitigate the modality gap between 3D and language with a proposed **Multi-Order RElation (MORE)** mining network.

PUBLICATION

Two-stage Visual Cues Enhancement Network for Referring Image Segmentation (ACM MM 2021)

Yang Jiao, Zequn Jie, Weixin Luo, Jingjing Chen, Yu-Gang Jiang, Xiaolin Wei and Lin Ma

MORE: Multi-order Relation Mining for Dense Captioning in 3D Scenes (ECCV 2022)

Yang Jiao, Shaoxiang Chen, Zequn Jie, Jingjing Chen, Lin Ma and Yu-Gang Jiang

MSMDFusion: Fusing LiDAR and Camera at Multiple Scales with Multi-depth Seeds for 3D Object Detection (CVPR 2023)

Yang Jiao, Zequn Jie, Shaoxiang Chen, Jingjing Chen, Lin Ma and Yu-Gang Jiang

Suspected Object Matters: Rethinking Model's Prediction for One-stage Visual Grounding (ACM MM 2023)

Yang Jiao, Zequn Jie, Jingjing Chen, Lin Ma and Yu-Gang Jiang

Instance-aware Multi-Camera 3D Object Detection with Structural Priors Mining and Self-Boosting Learning (AAAI 2024)

Yang Jiao, Zequn Jie, Shaoxiang Chen, Jingjing Chen, Lin Ma and Yu-Gang Jiang

Lumen: Unleashing Versatile Vision-Centric Capabilities of Large Multimodal Models (NeurIPS 2024)

Yang Jiao, Shaoxiang Chen, Zequn Jie, Jingjing Chen, Lin Ma and Yu-Gang Jiang

NuScenes-QA: A Multi-modal Visual Question Answering Benchmark for Autonomous Driving Scenario (AAAI 2024)

Tianwen Qian, Jingjing Chen, Linhai Zhuo, Yang Jiao and Yu-Gang Jiang

From Canteen Food to Daily Meals: Generalizing Food Recognition to More Practical Scenarios (TMM 2024)

Guoshan Liu, Yang Jiao, Jingjing Chen, Bin Zhu and Yu-Gang Jiang

WildRIS: Towards Comprehensive Referring Image Segmentation in the Wild (UnderReview)

Haoxiang Chen, Yang Jiao, Jingjing Chen, and Yu-Gang Jiang

EAGLE: Towards Efficient Arbitrary Referring Visual Prompts Comprehension for Multimodal Large Language Models (UnderReview)

Jiacheng Zhang, Yang Jiao, Shaoxiang Chen, Jingjing Chen, Yu-Gang Jiang

Eyes Can Deceive: Benchmarking Counterfactual Reasoning Abilities of Multimodal Large Language Models (UnderReview)

Yan Li, Wentao Tian*, Yang Jiao, Jingjing Chen, Yu-Gang Jiang*