



教育经历

复旦大学	2021年09月 - 至今
博士四年级 计算机科学技术学院	上海
获奖情况：国家奖学金（2021年）；董氏企业奖学金（2022年）；博士研究生一等奖奖学金（2023年）	
电子科技大学	2017年09月 - 2021年06月
学士 计算机科学与工程学院（“成电英才计划”实验班）	成都
绩点：3.99/4.0（前5%）	

研究方向及代表性工作

自动驾驶

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

- 提出MSMDFusion，一个增进细粒度激光雷达与相机信息融合的框架；
- 在自动驾驶国际知名榜单nuScenes中分别获得检测赛道和跟踪赛道单模型第1名和第2名的成绩；
- 开源仓库(<https://github.com/SxJyJay/MSMDFusion>)170+ stars；

多模态大模型

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

- 提出Lumen，一个支持包括目标检测、实例分割、姿态估计、语言指代的视觉物体定位和分割等多种视觉能力的多模态大模型；
- 所提出的Lumen不仅在多项视觉任务上取得了领域内最优的结果，且保持了与领域内先进模型可媲美的对话能力；

图像理解生成统一的多模态大模型（预研项目）

- 开发类Chameleon的支持图像理解生成统一的多模态大模型，保证高质量文生图的同时，兼顾对标业界多模态大模型的图文理解性能

论文发表

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 (Under Review)

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

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

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

Look Before You Decide: Prompting Active Deduction of MLLMs for Assumptive Reasoning

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

EventHallusion: Diagnosing Event Hallucinations in Video LLMs (Under Review)

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

科研项目

国家自然科学基金青年学生基础研究项目（博士研究生）：“基于指令驱动的视觉内容理解与生成研究”

主持

基于指令驱动的视觉内容理解与生成旨在根据输入指令，理解视觉场景中所重点指代的对象，进而对其进行定位或编辑。目前相关工作已经取得长足发展，但其大多仅局限于理解简单且有限的文本指令，无法应对现实场景中用户多样、复杂的指令，从而无法解读用户的真正意图。针对上述问题，本项目拟开展以下几个方面的研究：（1）大规模、层次化的“指令—图像”数据集构建；（2）多样化指令驱动的多层级视觉内容理解通用模型构建；（3）渐进式指令驱动的视觉内容生成研究。其中，数据集的构建是算法研究的基础；通用模型的构建可精确解读多样的用户指令、理解用户意图，并定位至相应的视觉对象；在大规模数据集与通用模型的加持下，渐进式的内容生成方法进一步以视觉形式将结果反馈给用户。本课题拟推动视觉内容理解与生成领域的智能化发展，为复杂场景下的用户需求提供关键技术支持。

项目经历

美团	2023年09月 - 2024年12月
大模型实习生	北京
• 提升多模态大模型的多样化视觉能力，例如检测、分割、关键点等	
• 研究具有图文理解与文生图统一能力的原生多模态大模型	
美团	2020年10月 - 2021年05月
算法实习生	北京
• 图像超分辨率与视频插帧算法的开发与部署	