



Yang Jiao

ACADEMIC EXPERIENCE

Doctor Computer Science and Technology Shang GPA: 3.57/4.0 Scholarship: National Scholarship; Dong's Enterprise Scholarship INDUSTRIAL EXPERIENCE INDUSTRIAL EXPERIENCE Meituan Sep 2023 - Oct 20 Large Multimodal Model Intern Beij • Develop a large multimodal model equipped with versatile vision-centric capabilities Oct 2020 - May 20 Meituan Oct 2020 - May 20		
Scholarship: National Scholarship: Dong's Enterprise Scholarship INDUSTRIAL EXPERIENCE Methan Sep 2023 - Oct 20 Large Multimodal Model Intern Beij Develop a targe multimodal model equipped with versatile vision-centric capabilities Develop a targe multimodal model equipped with versatile vision-centric capabilities Algorithm Intern Oct 2020 - May 20 Scholarship: National Sporthm development and deployment Beij Image super-resolution algorithm development and deployment Beij Scholarship: National Sporthm development and deployment Scholarship Scholarship: National Sporthm development and deployment Scholarship Scholarship: Scholarship Scholarship Scholarship: Scholarship Scholarship Scholarship Viao frame interpolation algorithm development and deployment Scholarship Scholarship Schoarship Scholarship <t< th=""><th>Doctor Computer Science and Technology</th><th>Sep 2021 – Jun 2026 Shangha</th></t<>	Doctor Computer Science and Technology	Sep 2021 – Jun 2026 Shangha
Metiuan Sep 2023 - Oct 20 Large Multimodal Model Intern Beij Develop a targe multimodal model equipped with versatile vision-centric capabilities Beij Develop a native large multimodal model for unified image generation and visual comprehension Metiuan Agorithm Intern Oct 2020 - May 20 Beij Unage super-resolution algorithm development and deployment English and the provide of the provid	Scholarship : National Scholarship; Dong's Enterprise Scholarship	
Large Multimodal Model Intern Develop a large multimodal model equipped with versatile vision-centric capabilities Develop a large multimodal model for unified image generation and visual comprehension Metuan Oct 2020 - May 22 Agorithm Intern Develop a large multimodal model for unified image generation and visual comprehension Metuan Oct 2020 - May 22 Agorithm Intern Develop a large multimodal model for unified image generation and visual comprehension Oct 2020 - May 22 Agorithm Intern Undes frame interpolation algorithm development and deployment 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 LDAR-camera fusion with a proposed framework, abbreviated as MSMDFusion. Github (https://github.com/Sxk/jag/MSMDFusion) 100 - stars. MORE: Multi-Order Relation Mining for Dense Captioning in 3D Seenes 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 Relering Image Segmentation (ACM MM 2021) Yang Jiao, Zequn Jie, Weizin Luo, Jinging Chen, Jin Ma and Yu-Gang Jiang Subocted Object Matters: Retinking Othen, Sur Gang Jiang, Xiaolin Wei and Lin Ma MORE: Multi-order Relation Mining for Dense Captioning in 3D Seenes First Author, Accepted Lo Cuer 2023 Yang Jiao, Zequn Jie, Shaoxiang Chen, Jinging Chen, Lin Ma and Yu-Gang Jiang Suspected Object Matters: Retinking Chen, Lin Ma and Yu-Gang Jiang Suspected Object Matters: Retinkinging Chen, Lin Ma and Yu-Gang Jiang	INDUSTRIAL EXPERIENCE	
 Develop a large multimodal model equipped with versatile vision-centric capabilities Develop a native large multimodal model for unified image generation and visual comprehension Metuan Oct 2020 - May 20 May 20 <		Sep 2023 - Oct 2024
 Develop a native large multimodal model for unified image generation and visual comprehension Metiuan Oct 2020 - May 20 Majorithm Intern Develop an native large multimodal model for unified deployment Oct 2020 - May 20 Majorithm Intern Develop an interpolation algorithm development and deployment 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 NeurIPS 2024 Introduce Lumen, a LMM with versatile vision-centric functions without significantly compromising general-purpose VQA capabilities. Fusing LiDAR and 2nd on nuScenes detection and tracking leaderboard among all single-model approaches by the time of submission. Github (https://github.com/Sx.lylay/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 Jako, Zegun Jie, Mining for Dense Captioning in 3D Scenes (ECCV 2022) Yang Jako, Zegun Jie, Maionin Qing Dense Captioning in 3D Scenes (ECCV 2022) Yang Jako, Shaoxiang Chen, Jingjing Chen, Lin Ma and Yu-Gang Jang, MSMDFusion: Fusing LIDAR and Camera at Multiple Scales with Multi-depth Seeds for 3D Object Detection (CVPR 2023) Yang Jako, Zegun Jie, Shaoxiang Chen, Jingjing Chen, Lin Ma and Yu-Gang Jang Suspected Object Matters: Rethinking Model's Prediction for One-stage Visual Grounding (ACM MM 2021) Yang Jako, Zegun Jie, Shaoxiang Chen, Jingjing Chen, Lin Ma and Yu-Gang Jiang Lumen: Unleasting Versatile Vision-Centric Capabilities of Large Multimodal Models (NeurIPS 2024) Yang Jako,	•	Deijing
Algorithm Intern Image super-resolution algorithm development and deployment Video frame interpolation algorithm development and deployment SELECTED RESEARCH Image super-resolution algorithm development and deployment SELECTED RESEARCH Image: 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. First Author, Accepted to CVPR 2023 Enhance the fine-grained LIDAR-camera fusion with a proposed framework, abbreviated as MSMDFusion. Earls tat 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 More: Multi-Order Relation Mining for Dense Captioning in 3D Scenes First Author, Accepted to ECCV 2022 More: Multi-Order Relation Mining for Dense Captioning in 3D Scenes First Author, Accepted to ECCV 2022 More: Multi-Order Relation Mining for Dense Captioning in 3D Scenes (ECCV 2022) Mary Jao, Saquu Jue, Wakin Luo, Junging Chen, Jur Agang Jiang, Xiaolin Wei and Lin Ma MORE: Multi-Order Relation Mining for Dense Captioning in 3D Scenes (ECCV 2022) Mary Jao, Saquu Jue, Maxima Chen, Junging 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) Marg Jao, Zequu Jue, Shaoxiang Chen, Junging Chen, Lin Ma and Yu-Gang Jiang Mang Jian, Zequu Jue, Shaoxiang Chen, Jinging 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) Marg Jao, Zequu Jue, Shaoxiang Chen, Jinging Chen, Lin Ma and Yu-Gang Jiang MSMDFusion: Fusing LIDAR and Camera at Multiple Scales with Multi-depth Seeds (Sud CM MM 2023) Marg Jao, Zequu Jue, Shaoxiang Chen, Jinging		
 Image super-resultion algorithm development and deployment Video frame interpolation algorithm development and deployment SELECTED RESEARCH Lumen: Unleashing Versatile Vision-Centric Capabilities of Large Multimodal Models First Author, Accepted to NeurIPS 2024 Introduce Lumen, a LIMW 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/SkyJay/MSMDFusion) 100- stars. MORE: Multi-Order Relation Mining for Dense Captioning in 3D Scenes First Author, Accepted to ECCV 2022 Viang Jako, Zequn Jie, Weiwin Luo, Jingling Chen, Yu-Gang Jiang, Xiaolin Wei and Lin Ma MORE: Multi-Order Relation Mining for Dense Captioning in 3D Scenes (ECCV 2022) Yang Jako, Sequn Jie, Weiwin Luo, Jingling 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 Jako, Zequn Jie, Jingling Chen, Lin Ma and Yu-Gang Jiang MSMDFusion: Fusing LIDAR and Camera at Multiple Scales Visual Grounding (ACM MM 2023) Yang Jako, Zequn Jie, Jingling Chen, Lin Ma and Yu-Gang Jiang MSMDFusion: Fusing LIDAR and Camera at Multiple Scales Visual Grounding (ACM MM 2023) Yang Jako, Zequn Jie, Shaoxiang Chen, Jingling Chen, Lin Ma and Yu-Gang Jiang Instance-aware Multi-Camera 3D Object Detection with Structural Priors Mining and Self-Boosting Learning (AAI		Oct 2020 - May 2021
Video frame interpolation algorithm development and deployment 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. Early that a 2nd on nuScenes detection and tracking leaderboard among all single-model approaches by the time of submission. Github (https://github.com/Sx/yay/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 Jaio, Zequn Jie, Wei/in Luo, Jinging Chen, Yu-Gang Jiang, Xiaolin Wei and Lin Ma MORE: Multi-Order Relation Mining for Dense Captioning in 3D Scenes (ECCV 2022) Yang Jaio, Zequn Jie, Maxim, Jainging Chen, Jin Ma and Yu-Gang Jiang MSMDFusion: Fusing LIDAR and Camera at Multiple Scales with Multi-depth Seeds for 3D Object Detection (CVPR 2023) Yang Jaio, Zequn Jie, Jingjing Chen, Jin Ma and Yu-Gang Jiang Instance-aware Multi-Camera 3D Object Detection with Structural Priors Mining and Self-Boosting (AAM 2024) Yang Jaio, Shaoxiang Chen, Jingjing Chen, Lin Ma and Yu-Gang Jiang Instance-aware Multi-Camera 3D Object Detection in Ma and Yu-Gang Jiang Instance-aware Multi-Camera 3D Object Detection with Structural Priors Mining and Self-Boosting Learning (AAAI 2024) Yang Jaio, Shaoxiang Chen, Jingjing Chen, Lin Ma and Yu-Gang Jiang Instance - Strue, Shaoxiang Chen, Jingjing Chen, Lin Ma	•	Beijing
SELECTED RESEARCH Lumen: Unleashing Versatile Vision-Centric Capabilities of Large Multimodal Models First Author, Accepted to NeurIPS 2024 Introduce Lumen, a LIMM with versatile vision-centric functions without significantly compromising general-purpose VQA capabilities. 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 (Intrips/)github.com/SxJJa/MIMB/Fusion) 100+ stars. WORE: Kulti-Order Relation Mining for Dense Captioning in 3D Scenes First Author, Accepted to ECV 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 Jaio, Shaoxiang Chen, Zequu 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 Jaio, Zequu 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, Zequu Jie, Jingjing Chen, Lin Ma and Yu-Gang Jiang Issuece Avare Multi-Camerera 3D Object Detection with Structural		
First Author, Accepted to NeurIPS 2024 Introduce Lumen, a LMM with versatile vision-centric functions without significantly compromising general-purpose VOA capabilities. Fusing LiDAR and Camera at Multiple Scales with Multi-Depth Seeds for 3D Object Detection First Author, Accpeted to CVPR 2023 Enhance the fine-grained LIDAR-camera fusion with a proposed framework, abbreviated as MSMDFusion. Arank 1st and 2nd on nuScenes detection and tracking leaderboard among all single-model approaches by the time of submission. Glinbu (https://glinbu.com/SxJ/Jay/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, Zaqun Jie, Webin 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, Yu-Gang Jiang, Xiaolin Wei and Lin Ma MORE: Sulti-order Relation Mining for Dense Captioning in 3D Scenes (ECCV 2021) Yang Jiao, Shaoxiang Chen, Zequn Jie, 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, Shaoxiang Chen, 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 Instance-aware Multi-Camera 3D Object Detection neith Structural Priors Mining and Self-Boosting Learning (AAAI 2024) Yang Jiao, Shaoxiang Chen, Jangjing Chen, Lin Ma and Yu-Gang Jiang Instance-aware Multi-Camera 3D Object Detection neith Structural Priors Mining and Self-Boosting Learning (AAAI 2024) Yang Jiao, Shaoxiang Chen, Jangjing Chen, Lin Ma and Yu-Ga		
 Introduce Lumen, a LMM with versatile vision-centric functions without significantly compromising general-purpose VQA capabilities. Fusi 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/SXJ/Jay/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. PUELCATION 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 Suspected Object Matters: Rethinking Model's Prediction for One-stage Visual Grounding (ACM MM 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 Lumen: Unleashing Versatile Vision-Centric Capabilities of Large Multimodal Models (NeurIPS 2024) Yang Jiao, 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, Jingjing Chen, Lin Ma and Yu-Gang Jiang Lumen: Unleashing Versatile Vision-Centric Capabilities of Large Multimodal Models (NeurIPS 2024) Yang Jiao, Shaoxiang Chen, Jingjing Chen, Lin Ma and Yu-Gan		
 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/SxJ/Jay/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 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 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 NuScenes-OA: A Multi-modal Visual Question Answering Benchmark for Autonomous Driving Scenario (AAAI 2024) Yang Jiao, Shaoxiang Chen, Jingjing Chen, Lin Ma and Yu-Gang Jiang NuScenes-OA: A Multi-modal Visual Question Answering Benchmark for Autonomous Driving Scenario (AAAI 2024) Yang Jiao, Jingjing Chen, Jingjing Chen, Lin Ma and Yu-Gang Jiang From Canteen Food to Daily Meals: Generalizing Fo	• Introduce Lumen, a LMM with versatile vision-centric functions without significantly compromising general-purpos Fusing LiDAR and Camera at Multiple Scales with Multi-Depth Seeds for 3D Object Detection	e VQA capabilities.
 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 Iumen: 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) Tiamwen 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, And Yu-Gang Jiang WildRIS: Towards Comprehensive Referring Ima		
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 2021) 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-OA: A Multi-modal Visual Question Answering Benchmark for Autonomous Driving Scenario (AAAI 2024) Tianwen Olan, Jingjing Chen, Lin Ajia and Yu-Gang Jiang From Canteen Food to Daily Meats: Generalizing Food Recognition to More Practical Scenarios (TIMI 2024) Gueshan 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 EVGLE: Towards Efficient Arbitrary Referring Image Segmentation in the Wild (UnderReview) Haoxiang Chen, Yang Jiao, Jingjing Chen, Jingjing Chen, Yu-Gang Jiang EvgLE: Towards Efficient Arbitrary Referring Issual Prompts Comprehension for Multimodal Large Lang		ne of submission.
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) <i>Yang Jiao, Zequn Jie, Weixin Luo, Jingjing Chen, Yu-Gang Jiang, Xiaolin Wei and Lin Ma</i> MORE: Multi-order Relation Mining for Dense Captioning in 3D Scenes (ECCV 2022) <i>Yang Jiao, Shaoxiang Chen, Zequn Jie, Jingjing Chen, Lin Ma and Yu-Gang Jiang</i> MSMDFusion: Fusing LiDAR and Camera at Multiple Scales with Multi-depth Seeds for 3D Object Detection (CVPR 2023) <i>Yang Jiao, Zequn Jie, Shaoxiang Chen, Jingjing Chen, Lin Ma and Yu-Gang Jiang</i> Suspected Object Matters: Rethinking Model's Prediction for One-stage Visual Grounding (ACM MM 2023) <i>Yang Jiao, Zequn Jie, Jingjing Chen, Lin Ma and Yu-Gang Jiang</i> Instance-aware Multi-Camera 3D Object Detection with Structural Priors Mining and Self-Boosting Learning (AAAI 2024) <i>Yang Jiao, Shaoxiang Chen, Jingjing Chen, Lin Ma and Yu-Gang Jiang</i> Lumen: Unleashing Versatile Vision-Centric Capabilities of Large Multimodal Models (NeurIPS 2024) <i>Yang Jiao, Shaoxiang Chen, Zequn Jie, Jingjing Chen, Lin Ma and Yu-Gang Jiang</i> NuScenes-OA: A Multi-modal Visual Question Answering Benchmark for Autonomous Driving Scenario (AAAI 2024) <i>Tianwen Qian, Jingjing Chen, Linhai Zhuo, Yang Jiao and Yu-Gang Jiang</i> From Canteen Food to Daily Meals: Generalizing Food Recognition to More Practical Scenarios (TMM 2024) <i>Guoshan Liu, Yang Jiao, Jingjing Chen, Bin Zhu and Yu-Gang Jiang</i> WildRIS: Towards Comprehensive Referring Image Segmentation in the Wild (UnderReview) <i>Haoxiang Chen, Yang Jiao, Jingjing Chen, And Yu-Gang Jiang</i> EAGLE: Towards Efficient Arbitrary Referring Image Segmentation in the Wild (UnderReview) <i>Jachen Yang Jiao, Jingjing Chen, And Yu-Gang Jiang</i> Eyes Can Deceive: Benchmarking Counterfactual Reasoning Abilities of Multimodal Large Language Models (UnderReview) <i>Jiachen Zhang, Yang Jiao, Shaoxiang Chen, Jingjing Chen, Y</i>		
 Mitigate the modality gap between 3D and language with a proposed Multi-Order RElation (MORE) mining network. PUELCATION Two-stage Visual Cues Enhancement Network for Referring Image Segmentation (ACM MM 2021) Yang Jiao, Zequn Jie, Weikin 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 Autonmous Driving Scenario (AAAI 2024) Tianwen Qian, Jingjing Chen, Linha Jino, 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 Vusual Prompts Comprehension for Multimodal Large Language Models (UnderReview) Jacheng Zhang, Yang Jiao, Shaoxiang Counterfactual Reasoning Abilities of Multimodal Large Langu		
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) Yang Jiao, 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 WildRIIS: Towards Compr		'k.
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 Olan, 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, Bin Zhu and Yu-Gang Jiang <th></th> <th></th>		
 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, 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) 		
 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) 		
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, Jingjing Chen, Yu-Gang Jiang Eyes Can Deceive: Benchmarking Counterfactual Reasoning Abilities of Multimodal Large Language Models (UnderReview)		2023)
 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) 		
 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) 		24)
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)		
 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) 	NuScenes-QA: A Multi-modal Visual Question Answering Benchmark for Autonomous Driving Scenario (AAAI 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) 		
 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) 		
 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) 		
 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) 		
Eyes Can Deceive: Benchmarking Counterfactual Reasoning Abilities of Multimodal Large Language Models (UnderReview)	EAGLE: Towards Efficient Arbitrary Referring VIsual Prompts Comprehension for Multimodal Large Language Models	(UnderReview)
Yian Li", wentao Tian^, Yang Jiao , Jingjing Chen, Yu-Gang Jiang	Eyes Can Deceive: Benchmarking Counterfactual Reasoning Abilities of Multimodal Large Language Models (UnderF	Review)
	Yian Li", wentao Tian^, Yang Jiao , Jingjing Chen, Yu-Gang Jiang	