# Runsen Xu

runsxu@gmail.com · (+86)13388603499 · runsenxu.com · google scholar

## Education

**The Chinese University of Hong Kong (CUHK)** — Hong Kong, China
Ph.D. Student at MMLab, Information Engineering — Aug. 2022 – June 2026 (Expected)
- Advisor: Prof. Dahua Lin

**Zhejiang University (ZJU)** — Hangzhou, China
B.Eng. in Computer Science and Technology — Sep. 2018 – June 2022
- Pursuit Science Class (Qiushi Honor's Program, 20 students only), Chu Kochen Honors College

## Publications

(Some papers are under review. * equal contribution, # project lead, ✉ corresponding author)

[1] **ChangingGrounding: 3D Visual Grounding in Changing Scenes**
Miao Hu, Zhiwei Huang, Tai Wang, Jiangmiao Pang, Dahua Lin, Nanning Zheng, **Runsen Xu**✉
In submission, 2025. [Paper]

[2] **MMSI-Bench: A Benchmark for Multi-Image Spatial Intelligence**
Sihan Yang*, **Runsen Xu***#, Yiman Xie, ..., Dahua Lin, Tai Wang✉, Jiangmiao Pang✉
arXiv Preprint, 2025. [Paper]

[3] **Multi-SpatialMLLM: Multi-Frame Spatial Understanding with Multi-Modal Large Language Models**
**Runsen Xu**, Weiyao Wang, Hao Tang, ..., Fu-Jen Chu, Dahua Lin, Matt Feiszli, Kevin J. Liang
arXiv Preprint, 2025. [Paper]

[4] **PointLLM-V2: Empowering Large Language Models to Better Understand Point Clouds**
**Runsen Xu***, Shuai Yang*, Xiaolong Wang, Tai Wang✉, Yilun Chen, Jiangmiao Pang✉, Dahua Lin
Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2025. [Paper]

[5] **VLM-Grounder: A VLM Agent for Zero-Shot 3D Visual Grounding**
**Runsen Xu**, Zhiwei Huang, Tai Wang, Yilun Chen, Jiangmiao Pang✉, Dahua Lin
Conference on Robot Learning (CoRL), 2024. [Paper]

[6] **PointLLM: Empowering Large Language Models to Understand Point Clouds**
**Runsen Xu**, Xiaolong Wang, Tai Wang✉, Yilun Chen, Jiangmiao Pang✉, Dahua Lin
European Conference on Computer Vision (ECCV), 2024, Best Paper Candidate. [Paper]

[7] **MV-JAR: Masked Voxel Jigsaw and Reconstruction for LiDAR-Based Self-Supervised Pre-Training**
**Runsen Xu**, Tai Wang, Wenwei Zhang, Runjian Chen, Jinkun Cao, Jiangmiao Pang✉, Dahua Lin
Computer Vision and Pattern Recognition (CVPR), 2023. [Paper]

[8] **VFlowOpt: A Token Pruning Framework for LMMs with Visual Information Flow-Guided Optimization**
Sihan Yang, **Runsen Xu**, Chenhang Cui, Dahua Lin, Tai Wang, Jiangmiao Pang
International Conference on Computer Vision (ICCV), 2025. [Paper]

[9] **OST-Bench: Evaluating the Capabilities of MLLMs in Online Spatio-temporal Scene Understanding**
Jingli Lin*, Chenming Zhu*, **Runsen Xu**, Xiaohan Mao, Xihui Liu, Tai Wang, Jiangmiao Pang
arXiv Preprint, 2025. [Paper]

[10] **Learning Video Generation for Robotic Manipulation with Collaborative Trajectory Control**
Xiao Fu, Xintao Wang, Xian Liu, Jianhong Bai, **Runsen Xu**, Pengfei Wan, Di Zhang, Dahua Lin
arXiv Preprint, 2025. [Paper]

[11] **Grounded 3D-LLM with Referent Tokens**
Yilun Chen*, Shuai Yang*, Haifeng Huang*, Tai Wang, **Runsen Xu**, Ruiyuan Lyu, et al.
arXiv Preprint, 2024. [Paper]

[12] **MMScan: A Multi-Modal 3D Scene Dataset with Hierarchical Grounded Language Annotations**
Ruiyuan Lyu*, Tai Wang*, Jingli Lin*, Shuai Yang*, Xiaohan Mao, Yilun Chen, **Runsen Xu**, et al.

Neural Information Processing Systems, Datasets and Benchmarks Track (NeurIPS), 2024. [Paper]

[13] **Chat-Scene: Bridging 3D Scene and Large Language Models with Object Identifiers**
Haifeng Huang, Yilun Chen, Zehan Wang, Rongjie Huang, **Runsen Xu**, et al.
Neural Information Processing Systems (NeurIPS), 2024. [Paper]

[14] **EmbodiedScan: A Holistic Multi-Modal 3D Perception Suite Towards Embodied AI**
Tai Wang*, Xiaohan Mao*, Chenming Zhu*, **Runsen Xu**, et al.
Computer Vision and Pattern Recognition (CVPR), 2024. [Paper]

[15] **Fine-Grained Cross-View Geo-Localization Using a Correlation-Aware Homography Estimator**
Xiaolong Wang, **Runsen Xu**, Zuofan Cui, Zeyu Wan, Yu Zhang
Neural Information Processing Systems (NeurIPS), 2023. [Paper]

[16] **CO$^3$: Cooperative Unsupervised 3D Representation Learning for Autonomous Driving**
Runjian Chen, Yao Mu, **Runsen Xu**, Wenqi Shao, Chenhan Jiang, Hang Xu, Zhenguo Li, Ping Luo
International Conference on Learning Representations (ICLR), 2023. [Paper]

[17] **RNIN-VIO: Robust Neural Inertial Navigation Aided Visual-Inertial Odometry in Challenging Scenes**
Danpeng Chen, Nan Wang, **Runsen Xu**, Weijian Xie, Hujun Bao, Guofeng Zhang
International Symposium on Mixed and Augmented Reality (ISMAR), 2021. [Paper]

[18] **LIFE: Lighting Invariant Flow Estimation**
Zhaoyang Huang*, Xiaokun Pan*, **Runsen Xu**, Yan Xu, Kachun Cheung, Guofeng Zhang, Hongsheng Li
arXiv Preprint, 2021. [Paper]

## Research Experiences

**OpenRobotLab, Shanghai AI Laboratory** Jan. 2025 - June 2025
Research intern Shanghai, China
- Worked on multi-modal large language models (MLLMs) for spatial understanding.

**Perception Pillar, FAIR, Meta** June 2024 - Dec. 2024
Research scientist intern Menlo Park, CA, USA
- Worked on multi-modal large language models (MLLMs) for spatial understanding.

**OpenRobotLab, Shanghai AI Laboratory** Mar. 2023 - June 2024
Research intern Shanghai, China
- Worked on 3D vision and language, multi-modal large language models (MLLMs).

**Shanghai AI Laboratory** Nov. 2021 - July 2022
Research intern Shanghai, China
- Worked on self-supervised 3D representation learning.

**3D Vision Group, CAD&CG State Key Lab, Zhejiang University** Mar. 2020 - Oct. 2021
Research assistant Hangzhou, Zhejiang, China
- Worked on indoor AR navigation, indoor navigation solutions for individuals with visual impairments, and optical flow estimation.

**Mobile SLAM, SenseTime Group Ltd.** Mar. 2021 - June 2021
Research intern Hangzhou, Zhejiang, China
- Worked on neural inertial navigation with IMU data.

## Selected Awards

- ECCV Best Paper Candidate 2024
- Hong Kong PhD Fellowship (most prestigious scholarship for Ph.D. studies in Hong Kong) 2022
- CUHK Vice-Chancellor's PhD Scholarship 2022
- Outstanding Graduates of Zhejiang University 2022

- Outstanding Undergraduate Thesis 2022
- National Scholarship (highest honor nationwide for Chinese undergraduates) 2019

## Teaching

- Teaching Assistant: IERG4998 - Final Year Project Spring 2023
- Teaching Assistant: IERG4998 - Final Year Project Fall 2022

## Computer and Language Skills

- Programming Proficiency: Python, C/C++, PyTorch, Shell Script, Java, MATLAB, JavaScript/HTML/CSS
- Language Proficiency: Mandarin (native speaker), English (proficient)