

Education

Simon Fraser University, Canada

Computing Science, M.Sc.

Jun 2021 – Oct 2023

Simon Fraser University, Canada

Computing Science, B.Sc.

Sep 2016 – Jun 2021

Zhejiang University, China

Computer Science and Technology, B.Eng.

Sep 2016 – Jun 2021

Research Interest

Bridging Synthetic & Real-world Data

3D Scene Understanding

Controllable Diffusion Generation

Publications

SpatialLM: Training Large Language Models for Structured Indoor Modeling

Yongsen Mao*, Junhao Zhong*, Chuan Fang, Jia Zheng, Rui Tang, Hao Zhu, Ping Tan, Zihan Zhou
arXiv, 2025

[Project website](#), [Paper](#), [Code](#)

Ego-Exo4D: Understanding Skilled Human Activity from First- and Third-Person Perspectives

Kristen Grauman, Andrew Westbury, Lorenzo Torresani, Kris Kitani, Jitendra Malik, Triantafyllos Afouras, Kumar Ashutosh, Vijay Baiyya, Siddhant Bansal, Bikram Boote, Eugene Byrne, Zach Chavis, Joya Chen, Feng Cheng, Fu-Jen Chu, Sean Crane, Avijit Dasgupta, Jing Dong, Maria Escobar, Cristhian Forigua, Abrahm Gebreselasie, Sanjay Haresh, Jing Huang, Md Mohaiminul Islam, Suyog Jain, Rawal Khirodkar, Devansh Kukreja, Kevin J Liang, Jia-Wei Liu, Sagnik Majumder, **Yongsen Mao**, Miguel Martin, Effrosyni Mavroudi, Tushar Nagarajan, Francesco Ragusa, Santhosh Kumar Ramakrishnan, Luigi Seminara, Arjun Somayazulu, Yale Song, Shan Su, Zihui Xue, Edward Zhang, Jinxu Zhang, Angela Castillo, Changan Chen, Xinzhu Fu, Ryosuke Furuta, Cristina Gonzalez, Prince Gupta, Jiabo Hu, Yifei Huang, Yiming Huang, Weslie Khoo, Anush Kumar, Robert Kuo, Sach Lakhavani, Miao Liu, Mi Luo, Zhengyi Luo, Brighid Meredith, Austin Miller, Oluwatumini Oguntola, Xiaqing Pan, Penny Peng, Shraman Pramanick, Mereray Ramazanova, Fiona Ryan, Wei Shan, Kiran Somasundaram, Chenan Song, Audrey Southerland, Masatoshi Tateno, Huiyu Wang, Yuchen Wang, Takuma Yagi, Mingfei Yan, Xitong Yang, Zecheng Yu, Shengxin Cindy Zha, Chen Zhao, Ziwei Zhao, Zhifan Zhu, Jeff Zhuo, Pablo Arbelaez, Gedas Bertasius, David Crandall, Dima Damen, Jakob Engel, Giovanni Maria Farinella, Antonino Furnari, Bernard Ghanem, Judy Hoffman, C. V. Jawahar, Richard Newcombe, Hyun Soo Park, James M. Rehg, Yoichi Sato, Manolis Savva, Jianbo Shi, Mike Zheng Shou, Michael Wray

Computer Vision and Pattern Recognition (**CVPR**), 2024, **Oral**

[Project website](#), [Paper](#), [Code](#)

Habitat Synthetic Scenes Dataset (HSSD):

An Analysis of 3D Scene Scale and Realism Tradeoffs for ObjectGoal Navigation

Mukul Khanna*, **Yongsen Mao***, Hanxiao Jiang, Sanjay Haresh, Brennan Schacklett, Dhruv Batra, Alexander Clegg, Eric Undersander, Angel Chang, Manolis Savva

Computer Vision and Pattern Recognition (**CVPR**), 2024

[Project website](#), [Paper](#), [Code](#)

MultiScan: Scalable RGBD scanning for 3D environments with articulated objects

Yongsen Mao, Yiming Zhang, Hanxiao Jiang, Angel X. Chang, Manolis Savva

Advances in Neural Information Processing Systems (**NeurIPS**), 2022

[Project website](#), [Paper](#), [Code](#)

OPD: Single-view 3D Openable Part Detection

Hanxiao Jiang, **Yongsen Mao**, Manolis Savva, Angel X. Chang

European Conference on Computer Vision (**ECCV**), 2022, **Oral**

[Project website](#), [Paper](#), [Code](#)

Work Experience

Manycore Tech Inc. [link](#)
Research Engineer (Full Time)

Nov 2023 – Present
Hangzhou, China

- Developed SpatialLM, a large language model for 3D point cloud understanding. Achieving 10k+ downloads / 1k+ likes on Hugging Face and 3.7k+ GitHub stars. Curated a dataset of 12k+ synthetic indoor scenes with ground-truth annotations and experimented training strategies for multimodal adaptation. Delivered state-of-the-art level results in 3D layout estimation and competitive 3D object detection performance, demonstrating LLMs’ potential for robotics, and spatial AI applications.
- Focus on exploring and understanding controllable diffusion models for interior design. Developed controllable diffusion models that preserve room structure and automatically furnish rooms with accurate semantic layouts and realistic visual quality. Open-sourced checkpoint Layout-ControlNet on [HuggingFace Repo](#) and released a demo on [HuggingFace Space](#)

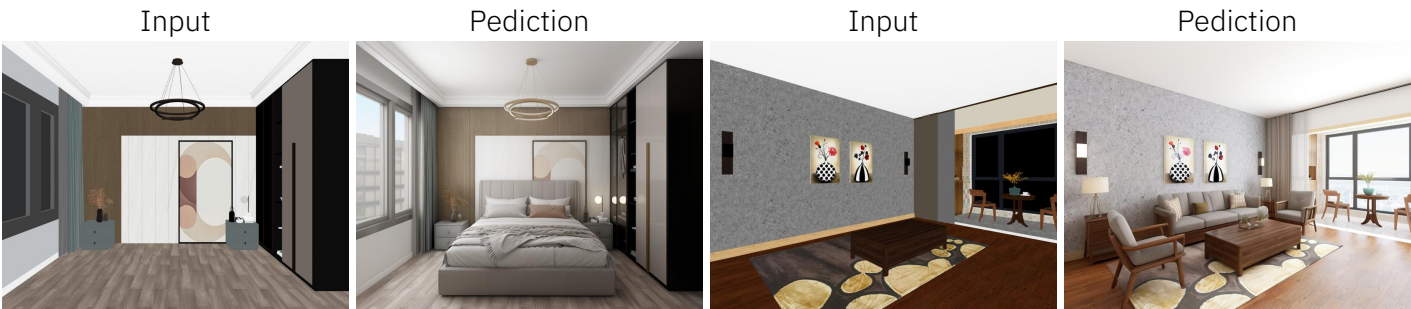


Figure 1: Convert incomplete scenes into fully furnished ones, while preserving texture and structure

DaoAI Robotics Inc. [link](#)
3D Vision Engineer (Undergraduate Coop)

Jan 2020 – Aug 2020
Vancouver, Canada

- Built CAD-based 3D object detection and 6-DoF pose estimation project for Random Bin Picking using point pair features (PPF) based method, the result surpassed the PCL’s PPF Estimation implementation in both quantitative and qualitative terms

Other Projects

SegmentAnyScene: Extends the capabilities of the Segment Anything Model, originally designed for 2D images, to work with 3D scene meshes [Code Video](#)

TA Experience

CMPT 361: Introduction to Computer Graphics

Jan 2022 – Apr 2022

Honors and Awards

Undergraduate Graduation: Joint Honours with Distinction

Fall 2021

Dean’s Honour Roll

Fall 2019

Activities

SFU-UBC AI Research Day

Present project MultiScan in oral and poster session

Dec 2022

CS Research Day

Present project MultiScan in poster session

Dec 2019 – 2022