

Yucheng Mao

San Diego, US — yucheng.mao.cs@gmail.com — +(1) 619-629-3136 — [Personal Webpage](#)

EDUCATION

University of California, San Diego , San Diego, US Master of Computer Science, Advisor: Zhuowen Tu	Enrolled: 09 2025 — Now
University of Science and Technology Beijing , Beijing, China Bachelor of Science	Enrolled: 09 2020 — 06 2024
University of California, Berkeley , Berkeley, US Visiting Student	Enrolled: 09 2023 — 12 2023

PUBLICATIONS

AnySplat: Feed-forward 3D Gaussian Splatting from Unconstrained Views. <i>Y Mao*, L Jiang*, L Xu, T Lu, K Ren, X Xu, M Yu, J Pang, F Zhao, D Lin, B Dai</i>	SIG Asia 2025, ACM TOG
Diffusion-Based Generative Models for 3D Occupancy Prediction in Autonomous Driving. <i>Y Wang*, Y Liu*, T Yuan*, Y Mao, Y Liang, X Yang, H Zhang, H Zhao</i>	ICRA 2025
Sparse Image Sets Restoration with Multi-View Diffusion Model. <i>Y Mao*, B Wang*, N Kulkarni, J Park</i>	CVPR 2025
BEVScope: Enhancing Self-Supervised Depth Estimation Leveraging Bird's-Eye-View in Dynamic Scenarios. <i>Y Mao, R Zhao, T Zhang, H Zhao</i>	Arxiv
Occ3D: A large-scale 3d occupancy prediction benchmark for autonomous driving. <i>X Tian, T Jiang, L Yun, Y Mao, Y Wang, Y Wang, H Zhao</i>	NeurIPS 2023
PreSight: Enhancing Autonomous Vehicle Perception with City-Scale NeRF Priors. <i>T Yuan, Y Mao, J Yang, Y Liu, Y Wang, H Zhao</i>	ECCV 2024

RESEARCH EXPERIENCES

University of California, San Diego Research Assistant	Supervised by Prof. Zhuowen Tu June 2025 - Present
<ul style="list-style-type: none">Project: Dynamic Image Tokenization with Quadtree Representation<ul style="list-style-type: none">Ongoing research project. We are trying to build an autoencoder and decoder which encode image into a variant number of tokens. We also aim to train a image tokenization policy which can use more tokens to represent high-frequency information in the image and use less tokens for low-frequency region of the image.	
Shanghai Artificial Intelligence Laboratory Research Assistant	January 2025 - Present
<ul style="list-style-type: none">Project: Scene-Level 3D Tokenizer<ul style="list-style-type: none">We build a scene level 3D tokenizer with 1B parameters, this tokenizer could directly take multi-view pose-free images as input, and squeeze the information of 3D scene into 1d tokens. Then we also train a decoder which can decode point cloud and images with given pose. Based on this tokenization strategy, we can easily adopt this kind of neural implicit representation for novel view synthesis, spatial understanding and 3D scene generation.	
Project: Large Feed-Forward 3D Gaussian Splatting Model <ul style="list-style-type: none">We build a large view synthesis model with 1.2B parameters, it can directly predict 3D gaussian primitives based on given images without any camera pose as extra input. Through the predicted gaussian primitives, we can render novel view images and geometry information(depth, normal) from it. Compared to previous optimization-based/feed-forward gaussian splatting method, we can predict 3D gaussian primitives from any number of unconstrained views within 3 seconds.	
<ul style="list-style-type: none">Project: Video Generation For Robotics Learning<ul style="list-style-type: none">Although current Vision-Language-Action model is pretrained on a large scale dataset, it still relies on fine-tuning with real world collected data in novel scene to handle novel task, generally, this kind of data collection requires around 500 episodes to make model working well on new task which is expensive and time-consuming. We aim to solve this problem by training a in-context learning video generation model to augmenting the observation of the robot system with novel background and novel manipulated objects.	
University of Michigan Research Assistant	Supervised by Prof. Jeong Joon Park April 2024 - January 2025

- **Project: Sparse View 3D reconstruction**

- Finetune a Stable Diffusion model for consistent depth estimation and sparse view reconstruction. Estimate the depth maps of multiple images jointly so that they are consistent with each other. The output system will be useful for computer graphics in rendering the target scene from new viewpoints not captured in the original input images.

- **Project: 3D reconstruction with degraded observation.**

- This work addresses image restoration by leveraging information from multiple degraded photographs of the same scene, a novel approach compared to traditional single-view methods. By hypothesizing that multi-view images contain complementary information, a powerful multi-view diffusion model is implemented to jointly denoise and enhance the views. The model achieves superior performance in tasks like image deblurring and super-resolution while ensuring 3D consistency, making it particularly valuable for applications such as 3D reconstruction and image correspondence.

MARS Lab in Tsinghua University
Research Assistant

Supervised by Prof. Hang Zhao
October 2022 - March 2024

- **Project: LiDAR Point Cloud Generation**

- Based on unconditional LiDAR point cloud generation, we propose a conditional LiDAR point cloud generation paradigm based on BEV layout.
- Our goal is to use generated LiDAR point cloud to replace sensor data collected on real-world data. This work is still in progress.

- **Project: BEVScope: Enhancing Self-Supervised Depth Estimation Leveraging Bird's-Eye-View in Dynamic Scenarios**

- Based on the high cost of existing depth sensing functions of the sensors of self-driving vehicles, determined the research direction to explore the use of cheaper sensors to obtain depth information; took the self-supervised paradigm as the solution.
- Improved depth estimation of dynamic objects by redesigning an adaptive photometric loss function; analyzed the issue of pose estimation and mutual consistency in multi-view depth maps using constraint conditions such as the adaptive photometric loss function and the camera pose consistency loss function.
- Accomplished multi-view information utilization by the proposal of the BEVScope, demonstrating competitive performance on datasets for multi-camera depth estimation.

- **Project: Occ3D: A large-scale 3d occupancy prediction benchmark for autonomous driving**

- Developed a collection of pipelines with automated labeling capabilities to represent detailed 3D geometric information; addressed the limitations of existing approaches in robotics and autonomous driving perception systems, such as overlooking significant geometric details and the lack perception of out-of-vocabulary objects.
- Assessed the validity and dependability of the auto-label system by a comparison between the automatically labeled data and the existing manually data; specifically, projecting the labeled occupancy data with semantic tags back to each perspective view and computing iou with its already labeled semantic segmentation tags.

- **Project: PreSight: Enhancing Autonomous Vehicle Perception with City-Scale NeRF Priors**

- Leverages historical traversal data to create static prior memories that enhance online perception during subsequent navigations.
- The experimental results on the nuScenes dataset demonstrate PreSight's high compatibility with various online perception models, showing significant improvements in tasks such as high-definition map construction and occupancy prediction. Especially for static components in self-driving scene.

PROJECTS

Unofficial implementation of **GS-LRM**

<https://github.com/InternRobotics/gs-lrm-unofficial>

- Reproduce the ECCV 2024 paper GS-LRM on Re10K dataset.

Unofficial implementation of **UltraLiDAR**

https://github.com/myc634/UltraLiDAR_nusc_waymo

- Reproduce the CVPR 2023 paper UltraLiDAR on LiDAR point cloud generation, and add the experiment on nuScenes and Waymo dataset

Merging **DETR3D** into **MMDetection3D Framework**

<https://github.com/open-mmlab/mmdetection3d/tree/main/projects/DETR3D>

- Refactored and trained the deep learning model to work with the new framework version, and integrated the 3D target detection model(DETR3D) into the open source MMdetection3D framework

Honors and Awards

- **University of Science and Technology Beijing People's Third Class Scholarship (Top 25%)** 09/2021
- **National Undergraduate Statistical Modeling Competition (Top 10%)** 05/2022
- **University of Science and Technology Beijing People's Third Class Scholarship (Top 25%)** 09/2023