

Hongrui Wu

Homepage [◇](#) Google Scholar [◇](#) LinkedIn [◇](#) Email how036@ucsd.edu

EDUCATION

Tongji University

Undergraduate in Computer Science and Technology (**Score:** 91.4/100) *Apr 2024 - Jul 2026*
AI in Construction, Civil Engineering (switched) *Sep 2021 - Apr 2024*

RESEARCH INTERESTS

My research interests span open-vocabulary understanding of 3D scenes, 3D scene reconstruction from both multi-view images and a single image, across explicit and implicit 3D representations .

PUBLICATIONS

- [1] **Hongrui Wu**, Z. Gao, J. Cao, K. Yao, W. Shen, and Z. Wei, “Folk: Fast open-vocabulary 3d instance segmentation via label-guided knowledge distillation,” 2025. arXiv: [2510.08849 \[cs.CV\]](#).
- [2] J. Cao*, **Hongrui Wu***, Z. Feng, H. Bao, X. Zhou, and S. Peng, “Universe: Unleashing the scene prior of video diffusion models for robust radiance field reconstruction,” *ICCV*, 2025.

EXPERIENCE

MLPC lab in UC San Diego Advisor: [Prof. Zhuowen Tu](#) July 2025 - Present

- We propose a scene-level 3D reconstruction framework that outputs artist-created triangle meshes, addressing the current gap where most mesh-based methods are restricted to object-level generation.
- Our framework aims to reconstruct detailed instance-level geometry while aligning scene-level layout with a single input image. (**Preparing for CVPR 2026**)

Tongji University [1] Advisor: [Prof. Zhihua Wei](#) and [Prof. Wen Shen](#) Feb 2025 - Aug 2025

- FOLK is a fast open-vocabulary 3D instance segmentation framework that eliminates expensive 3D-to-2D projection by distilling open-vocabulary knowledge from 2D CLIP into a 3D student model.
- We introduce a novel teacher-student architecture where the teacher leverages multi-view CLIP embeddings with high visibility and viewpoint diversity, and the student learns a direct 3D embedding space via a label-guided distillation mechanism, ensuring semantic consistency and fast inference.
- FOLK achieves **state-of-the-art AP50 of 35.7 on the ScanNet200 dataset**, while being up to **6.0×-152× faster** than prior methods during inference. (**Under review [Paper]**)

Zhejiang University [2] Advisor: [Prof. Sida Peng](#) and [Prof. Xiaowei Zhou](#) Dec 2024 - Mar 2025

- UniVerse is a video generative model for robust 3D reconstruction from inconsistent multi-view images. Given unstructured multi-view inputs, we first estimate a camera trajectory, insert blank frames to form a pseudo video, and then restore the sequence and extract consistent frames for reconstruction.
- This work highlights the potential of decoupling robust reconstruction into **restoration and reconstruction**, instead of directly performing reconstruction on inconsistent inputs. (**Accepted by ICCV 2025.** [\[Website\]](#) [\[Paper\]](#) [\[code\]](#))

AWARDS

National Scholarship (Top 0.2% nationwide, highest scholarship in China).
TianXiang Scholarship (2/600)
Interdisciplinary Contest In Modeling hosted by COMAP Finalist Prize (Top 1.8% worldwide)
National Undergraduate Mathematics/Physics Competition: Third Prize/Second Prize
ASCE Concrete Canoe Competition (2nd Place in California Section)