Haojun Yan

CURRICULUM VITAE

hy431@cam.ac.uk
(+44) 7865 136545
github.com/Hao7un
Google Scholar

Research Interests

World Models, Robot Learning, and 3D/4D Vision.

Education

University of Cambridge

Oct 2025 – Aug 2026

MPhil in Machine Learning and Machine Intelligence

Cambridge, UK

- **Thesis:** Efficient 4D Reconstruction from Temporally Sparse Keyframes.
- **Selected Coursework:** Introduction to Machine Learning, Probabilistic Machine Learning, Advanced Computer Vision, Reinforcement Learning, Introduction to Robotics, Advanced Human-Computer Interaction.

Beihang University

Sep 2021 – Jun 2025

B.Eng. in Computer Science and Technology

Beijing, China

- **GPA**: 3.86/4.00; **Rank** 1/204 through Year 3 (GPA 3.90)
- **Selected Coursework:** Linear Algebra, Pattern Recognition, Algorithm Design and Analysis, Mathematical Modeling, Discrete Mathematics, Optimization in Computer Engineering.
- **Undergraduate Research:** Institute of Advanced Computing Technology, Apr 2023 – Sep 2024; co-author of *Ionformer* (IEEE TGRS, 2025).

Publications

* equal contribution

[1] **OMeGa: Joint Optimization of Explicit Meshes and Gaussian Splats for Robust Scene-Level Surface Reconstruction**
Yuhang Cao*, **Haojun Yan***, Danya Yao
IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2026 ORAL

[2] **Ionformer: A Data-Driven Deep Learning Baseline for Global Ionospheric TEC Forecasting**
Lanhao Li, Yang Liu, Haoyi Zhou, Kunlin Yang, **Haojun Yan**, Jianxin Li
IEEE Transactions on Geoscience and Remote Sensing (TGRS), 2025

Research Experience

AGIBOT

Aug 2026 – Present

World Model Research Intern • Genie Team

Shanghai, China

- Developing a GE-Sim 2.0 benchmark for evaluating robot policies on contact-rich tasks involving fluids and deformable materials that are challenging for conventional rigid-body simulators. Defining the task suite and evaluation protocols, and implementing policy evaluation using generated video rollouts.
- Implementing and training a DiT-based Value Model to predict frame-level task-success scores from world-model video latents, using binary cross-entropy on success annotations from teleoperated demonstrations. Evaluating its ability to score generated trajectories automatically without manual frame-by-frame evaluation.

Visual Structure Learning Lab, University of Cambridge

Mar 2026 – Present

Visual Reconstruction and Video Generation Researcher • advised by Prof. Ayush Tewari

Cambridge, UK

- Adapted NeoVerse into a feedforward 4D Gaussian Splatting framework for reconstructing dynamic scenes from temporally sparse monocular keyframes. Introduced two zero-initialized residual waypoints per Gaussian to model nonlinear motion with endpoint-preserving cubic trajectories, while keeping the backbone and camera head frozen.
- Constructed dense 3D track targets by lifting long-range 2D tracks with LiDAR-aligned monocular depth, and trained with visibility-aware motion losses and depth supervision. Added learned gauge correction and cross-window rendering consistency to align overlapping windows and stabilize geometry throughout long video sequences.
- Evaluated on 202 held-out Waymo validation clips at 0.3–0.7 s keyframe gaps, achieving 1.03–2.81 dB higher PSNR and 19–44% lower LPIPS than released NeoVerse, with feedforward inference requiring no per-scene optimization.

LatenCore AI May 2026 – Aug 2026

Physical AI Intern • advised by Dr. Chuang Yu (UCL Interaction Center) Cambridge, UK

- Designed and implemented an HG-DAgger pipeline for real-robot $\pi_{0.5}$ rollouts with operator takeover and control handback, aggregating corrective demonstrations with the original dataset for iterative LoRA fine-tuning.
- Designed a demonstration-weighting scheme for policy training, scoring bimanual trajectories by the fraction of waypoints successfully solved by inverse kinematics to down-weight demonstrations that are difficult to retarget.
- Built a UMI-based bimanual data collection framework using Vive Tracker 3.0 to track the 6-DoF poses of both grippers, enabling human demonstrations to be collected independently of robot hardware and arm availability.
- Developed a teleoperation system for dual-arm AgileX Piper / Nero robots, integrating Meta Quest 3 controllers and Pika handles to support low-latency bimanual control, trajectory replay, and demonstration quality control.

Galaxea AI Jun 2024 – Feb 2025

3D Computer Vision Research Intern • advised by Prof. Hang Zhao (Tsinghua University) Beijing, China

- Proposed OMeGa, a framework that jointly optimizes explicit triangle meshes and 2D Gaussian splats through flexible binding and iterative mesh refinement to improve reconstruction in textureless indoor scenes. Reduced Chamfer-L_1 by 47–60% over the 2DGS baseline on MuSHRoom, ScanNet, and ScanNet++ (WACV 2026, Oral).
- Developed a 2DGS reconstruction pipeline for 10,000 m^2 outdoor scenes, achieving centimeter-level accuracy from consumer-grade cameras and extending the trainer to support multi-node distributed scene optimization.
- Integrated SAM-based instance segmentation for object-level scene decomposition and layout randomization, generating diverse simulation environments from a single capture to support manipulation policy training.
- Integrated depth and surface-normal priors from Depth Anything and StableNormal to improve geometry in textureless and reflective regions, and applied sky and dynamic-object masks to suppress reconstruction artifacts.

Honors & Awards

- 1st Place (Vienna Regional Site), What Bimanuals Can Do Competition, ICRA 2026
- National Scholarship, Ministry of Education of China 2022
- Outstanding Student Scholarship, Beihang University 2022

Skills

PROGRAMMING	Python, C / C++, CUDA, MATLAB
ML STACK	PyTorch, Hugging Face, Accelerate, Hydra, Weights & Biases
ROBOTICS	MuJoCo, Isaac Sim, ROS, Pinocchio, LeRobot, WebXR, URDF
LANGUAGES	Mandarin (Native), Cantonese (Native), English
TEST SCORES	IELTS: 8.0 GRE: 329