Xianjin Wu

wuxianjin@hust.edu.cn · wuxianjin633@gmail.com · +86 183-2713-1824
hyperboliccurve.github.io · GitHub · Google Scholar · ORCID

Research Interests

3D scene understanding, spatial intelligence, and embodied AI, with an emphasis on efficient machine learning systems: models that perceive the physical world in 3D from multi-view and video input and act in it in real time, unifying spatial perception, reasoning, and action within scalable foundation models.

Education

Huazhong University of Science and Technology (HUST) Wuhan, China
B.Eng. in Artificial Intelligence Sep 2022 – Jun 2026

- School of Artificial Intelligence and Automation
- GPA: 3.94/5.00; weighted average score: 84.7/100
- Research advisors: Prof. Xiang Bai and Asst. Prof. Dingkang Liang

Publications

* denotes equal contribution.

Conference Papers

[C1] **Xianjin Wu**, Dingkang Liang, Tianrui Feng, Kui Xia, Yumeng Zhang, Xiaofan Li, Xiao Tan, Xiang Bai. **"Generation Models Know Space: Unleashing Implicit 3D Priors for Scene Understanding."** *European Conference on Computer Vision (**ECCV**)*, 2026. [Paper] [Code] — 400+ GitHub stars; #1 on Hugging Face Daily Papers.

Workshop Papers

[W1] **Xianjin Wu***, Heng Fang*, Xuanyang Xi, Yiping Tang, Di Xu, Xiang Bai, Dingkang Liang. **"The Winning Solution for Track 1 of the 2026 AI City Challenge: Gauge Canonicalization for Multi-Camera 3D Perception."** *ECCV 2026 Workshop: 10th AI City Challenge*, 2026. **(Oral)** [Paper] [Code] — 1st place, Track 1.

Under Review

[S1] **Xianjin Wu***, Xian Li*, Jianxiang Tang, Dingkang Liang, Xiang Bai. **"Don't Throw Away the Solution: Reusing Denoising Trajectories for Fast Robot Control."** Under review at *International Conference on Learning Representations (**ICLR**)*, 2027. [Paper] [Code] [Project]

[S2] Hengyi Xie*, Chenfei Yao*, **Xianjin Wu**, Yingying Zhu, Dingkang Liang, Xiang Bai, Han Ding. **"TurboVLA: Real-Time Vision-Language-Action Model at 32 Hz on an RTX 4090 with <1 GB VRAM."** Under review at *International Conference on Learning Representations (**ICLR**)*, 2027. *arXiv:2607.27205*. [Paper] [Code] [Project]

Research Experience

DiDi Voyager (Autonomous Driving), Research Intern Aug 2026 – Present
Mentors: Xuesong Chen and Shaoshuai Shi

VLRLab, HUST, Researcher Oct 2023 – Present
Advisors: Prof. Xiang Bai and Asst. Prof. Dingkang Liang

Research on efficient embodied AI and 3D perception, including 3D spatial understanding in multimodal LLMs, fast inference for robot policies, compact vision-language-action models, and robust multi-camera 3D detection [C1, S1, S2, W1].

Baidu, Multimodal Perception & Understanding Dept., Research Intern Aug 2025 – Apr 2026
Mentors: Kui Xia and Yumeng Zhang

Proposed VEGA-3D [C1], showing that pretrained video generation models carry implicit 3D priors and injecting them into multimodal LLMs by using a video diffusion model as a latent world simulator, which improved 3D scene understanding, spatial reasoning, and embodied manipulation without explicit 3D supervision.

Selected Engineering Projects

LLM Inference Acceleration — SOAR Competition, Grand Final Champion Jun 2026

- Accelerated ultra-long-context inference of MiniCPM-SALA, a hybrid sparse/linear-attention LLM, on a single Blackwell RTX 6000D GPU within SGLang.
- Implemented NVFP4 weight quantization with Marlin/CUTLASS GEMM dispatch, a Triton GLA decode kernel, FlashInfer plan caching, operator fusion, and EAGLE-3 speculative decoding adapted to the hybrid architecture.

LLM Reasoning Fine-Tuning — Tianyi Cloud "Xirang Cup", National Gold Award, Team Leader Jul 2025

- Fine-tuned Qwen2.5-7B for mathematical reasoning on Ascend 910B NPUs with MindSpeed-LLM, leading dataset curation, curriculum learning, distillation with self-correction, and LoRA/vLLM optimization on Ascend.

Large-Scale 3D Reconstruction — Kaggle Image Matching Challenge 2025, Silver Medal Jun 2025

- Built a reconstruction pipeline for unordered image collections: DINOv2 scene clustering, hybrid ALIKED+ LightGlue and LoFTR matching, and RANSAC-based SfM; benchmarked DUSt3R, MASt3R, and VGGT.

Honors and Awards

2026

- 2nd Place, Learn2Design 2026 Challenge (NeurIPS 2026 Competition Track): Automated Design of Gravitational-Wave Detectors — €6,000 prize
- 1st Place, AI City Challenge 2026, Track 1: Multi-Camera 3D Perception (ECCV 2026 Workshop)
- Grand Final Champion, SOAR Sparse Operator Acceleration Competition (OpenBMB, SGLang, NVIDIA) — US$70,000 prize
- Second Prize, ICME 2026 Grand Challenge, Track 2: Fine-Grained Severity Grading for High-Precision Manufacturing
- Bronze Prize, RSIPAC 2026 International Track: Individual Tree Crown Segmentation
- 10th Place, MangoTV AI Challenge: Multi-View Generation

2025

- National Gold Award, Tianyi Cloud "Xirang Cup" University AI Competition, Large Model Application Track — RMB 200,000 prize
- 4th Place, Global Wheat Full Semantic Segmentation (GWFSS) Challenge (ICCV 2025 Workshop)
- Silver Medal (45 / 943 teams), Kaggle Image Matching Challenge 2025 (CVPR 2025 Workshop)
- National Second Prize, 3rd Hikvision "Qizhi Cup", Flexible Few-Shot Defect Detection Track

2024

- National First Prize, 6th Global Campus AI Algorithm Elite Competition (Huawei DIGIX), Steel Surface Defect Detection and Segmentation
- Top 1%, Alibaba Cloud Tianchi Global AI Attack & Defense Challenge: Financial Document Tampering Detection
- Honorable Mention, Mathematical Contest in Modeling (MCM)

Kaggle Competitions Expert (profile)

Academic Service and Open Source

- Reviewer: ECCV, ICLR, *Pattern Recognition* (journal)
- Contributor, vLLM-Omni: co-author of merged PR #7968, which reduces MiniCPM-o 4.5 reference setup and Talker overhead; contributed the Code2Wav prompt-state cache, first proposed in my PR #6653
- Maintainer, Awesome World Action Model: a curated list of vision-language-action and world action model papers

Skills

- **Programming:** Python, C/C++, Triton
- **ML frameworks and systems:** PyTorch, SGLang, vLLM, FlashInfer, CUTLASS, MindSpeed-LLM (Ascend NPU); quantization, speculative decoding, LoRA fine-tuning
- **Robotics:** Teleoperation and policy deployment on the AgileX Piper dual-arm platform
- **Languages:** Mandarin Chinese (native); English (professional proficiency)