JIAHAO ZHAN

(+86) 177-1853-4889 ◇ 22307140116@m.fudan.edu.cn ◇ Website

EDUCATION

Fudan University, Shanghai Sept. 2022 - Present
Bachelor of Artificial Intelligence (Honor Class, Computer Science Track)
GPA: 3.91/4.0, Major Rank: 1/30 (Artificial Intelligence), College Rank: 2/208, National Scholarship (2022 - 2023, Top 1 of the school), National Scholarship (2023 - 2024, Top 1 of the school)
Selected Courses: Linear Algebra (A), Data Structure (A), Probability and Mathematical Statistics (A), Computer Graphics A (A), Human-Computer Interaction (A), Artificial Intelligence (A), Method of Optimization (A), Operating Systems(H) (A), Natural Language Processing and Knowledge Representation (A)

Hong Kong University of Science and Technology, Hong Kong Jan. 2025 - Jun. 2025
Exchange Student in Computer Science

RESEARCH INTERESTS

Hybrid Generative Simulator, World Model, AIGC

PUBLICATION

- **PerpetualWonder: Long-Horizon Action-Conditioned 4D Scene Generation**
  Jiahao Zhan*, Zizhang Li*, Hong-Xing Yu, Jiajun Wu. (Submitted to CVPR 2026)
- **Fake it till You Make it: Reward Modeling as Discriminative Prediction**
  Runtao Liu*, **Jiahao Zhan***, Yingqing He, Chen Wei, Alan Yuille, Qifeng Chen. (Submitted to ICLR 2026)
- **MAC: A Live Benchmark for Multimodal Large Language Models in Scientific Understanding**
  Mohan Jiang*, Jin Gao*, **Jiahao Zhan**, and Dequan Wang. (COLM 2025)
- **Generalizing Motion Planners with Mixture of Experts for Autonomous Driving**
  Qiao Sun*, Huimin Wang*, **Jiahao Zhan**, Fan Nie, Xin Wen, Leimeng Xu, Kun Zhan, Peng Jia, Xianpeng Lang, Hang Zhao (ICRA 2025)

RESEARCH EXPERIENCES

Bytedance MMlab Nov. 2025 - Present
Algorithm Engineer Intern, mentored by Qunliang Xing and Shijie Zhao

Stanford Vision and Learning Lab Jun. 2025 - Nov. 2025
Research Intern (UGVI), advised by Prof. Jiajun Wu

- **PerpetualWonder: Long-Horizon Action-Conditioned 4D Scene Generation:**
  - We tackle the task of long-horizon action-conditioned 4D scene generation, enabling sequential action interactions.
  - We propose PerpetualWonder, a novel hybrid generative simulator that features a unified representation for both physical state and visual appearance, and a multi-view optimization mechanism for consistent scene updates.
  - We demonstrate that PerpetualWonder consistently outperforms prior work in action-conditioned 4D scene generation, including both long-horizon interaction abilities and scene consistency.

HKUST Visual Intelligence Lab Jan. 2025 - Jun. 2025
Research Intern, advised by Prof. Qifeng Chen

- **Fake it till You Make it: Reward Modeling as Discriminative Prediction:**
  - Existing preference alignment methods rely heavily on costly human annotations or require combining multiple metrics to construct rewards, which is both expensive and suboptimal.
  - Instead, we propose to infer preferences directly from the target data distribution, eliminating the need for extensive external supervision.

– Our method enables learning a robust reward model from minimal data, facilitating efficient sampling and RLHF optimization, also enabling multi-round GAN-like optimization.

Shanghai Qi Zhi Institute May. 2024 - Jan. 2025
Research Intern, advised by Prof. Hang Zhao

- **Generalizing the Motion Planners in Autonomous Driving:**
  - We propose a scalable MoE-based autoregressive model to learn different explicit rewards for motion planning and our method outperforms previous state-of-the-art methods by scaling with self-supervisions.
  - We comprehensively benchmark and analyze the generality of previous methods across multiple test sets with multiple Closed-Loop metrics by speeding up the model inference during simulations.
  - We unprecedentedly present scaling experiments with up to billions of diverse real-world urban driving scenarios.

Shanghai AI Lab Mar. 2023 - May. 2024
Research Intern, advised by Prof. Dequan Wang

- **MAC: A Live Benchmark for Multimodal Large Language Models in Scientific Understanding:**
  - We introduce MAC, a continuously updating benchmark for evaluating multimodal scientific understanding in MLLMs. Based on live scientific journals, our benchmark is built using a live mechanism that adapts to model progress.
  - We study the latest yearly snapshot of MAC, MAC-2025, drawn from over 2,000 curated journal issues, and provide a thorough evaluation of seven advanced MLLMs on both Image2Text and Text2Image tasks.
  - We propose an inference-time approach, Description and Deduction (DAD). It significantly enhances MLLMs' scientific concept reasoning by bridging cross-modal information between MLLMs and a reasoning language model.
  - We investigate the live attribute of our MAC through temporal data analysis and adaptive benchmark construction, showing the necessity of growing scientific journals and evolving construction using embedding models.

PROJECT EXPERIENCES

Shape Completion and Reconstruction of Sweet Peppers Challenge (ECCV workshop) Aug. 2024
Team Member

- Won the third prize.
- Constructed a point completion algorithm based on current researches.

Intel LLM-based Application Innovation Contest Nov. 2023
Project Leader

- Developed OptiPrompt, a user-friendly framework that can automatically optimize prompts to explicit users' questions and promote LLMs' performance based on Intel BigDL platform.
- Released the prototype in the final of the competition, winning the second prize.

SKILLS

Programming Languages: C/C++, Python, Bash, Git, LaTeX
Frameworks: Pytorch, Huggingface Transformers