

曹见闻

瑞士，苏黎世

jianwen.cao@uzh.ch | caojia@student.ethz.ch | +41 766395217 | +86 18094203447

教育背景

苏黎世大学 (UZH)

理学硕士，信息学（主修：人工智能，辅修：数据科学），GPA: 5.57/6

核心课程成绩：数据科学 (6.0/6.0)，自然语言处理机器学习 (6.0/6.0)，移动机器人视觉算法 (5.5/6.0)，ETHZ 3D 视觉 (5.75/6.0)，ETHZ 虚拟现实 (6.0/6.0)。

南京信息工程大学 (NUIST)

工学学士，人工智能专业，GPA: 87.6/100

瑞士，苏黎世

2024 - 至今

中国，南京

2020 - 2024

学术发表

[已录用] 第一作者: “Generative Event Pretraining with Foundation Model Alignment.” IEEE/CVF 计算机视觉与模式识别会议 (CVPR), 2026.

结合图像与事件相机进行多模态预训练, SOTA.

[在投] 第二作者: Memory Over Maps: 3D Object Localization Without Reconstruction IEEE/RISJ 国际智能机器人与系统大会 (IROS), 2026.

本人负责目标定位 (输入 query, 输出目标 3d 位置) 和效率分析部分。最终部署于波士顿动力机器狗。

[已发表] 第一作者: “Fuse Tune: Hierarchical Decoder Towards Efficient Transfer Learning.” 模式识别与计算机视觉大会 (PRCV), 2023.

提出了一种无需主干网络反向传播的高效迁移学习方法。

项目与获奖

机器人与强化学习

- 完整的 vision based 竞速无人机强化学习训练，以及真机部署经验 (GitHub)(Youtube)。
- 在 Isaac Sim 中，通过同步第一人视角视频中的人手的动作，对 Inspire 灵巧手进行遥操作 (GitHub)。
- 正逆运动学、坐标系转换，样条采样，PID 控制，RL, IL, Diffusion Policy。
- 实现强化学习算法 (DQN, PPO) 并在 Gymnasium (Atari 游戏) 上训练智能体。
- 实现模仿学习算法 (BC, Q-chunking, Dagger) 在 Soarm101 上部署有障碍物的方块抓取投放。
- 熟悉 DQN, Policy Gradient, TRPO, PPO, SAC 等强化学习算法原理。

SLAM 与 3D 高斯溅射 (3DGS)

- 基于 PTAM/FAST-LIVO2 的核心理念，在 KITTI 数据集上重新实现了视觉里程计 (VO)。
- 成功复现了 3DGS 原始论文，改进 HI-SLAM2，建立了一套完整的新视角合成 (Novel View Synthesis) 流程。

中国大学生数学建模竞赛

- 负责设计稳健的数学模型与优化算法，全国二等奖。

2022

科研与实习经历

苏黎世联邦理工学院计算机视觉与几何实验室 (ETHZ CVG)

学生研究员 | 视觉-惯性-激光雷达实时高斯溅射 SLAM

复现并改进 GS-LIVO，以实现更好的几何精度和建图效率，达到当时最优 (SOTA) 性能。

苏黎世大学机器人与感知实验室 (UZH RPG)

学生研究员 | 基于事件相机增强机器人运动控制策略

对多模态数据进行自回归预训练，显著提升下游视觉和控制任务的性能与鲁棒性。

合合信息 (IntSig)

视觉算法实习生

基于改进的 DETR 系列架构开发目标检测算法。

瑞士，苏黎世

2025 年 3 月 - 2026 年 2 月

瑞士，苏黎世

2025 年 2 月 - 2026 年 1 月

中国，上海

2023 年 7 月 - 2023 年 10 月

专业技能

机械臂和机器狗部署: 完整的 Mujoco 模仿学习和强化学习训练，ros2 部署 soarm101 经验，包括数据数据采集和状态空间优化
VLA: 熟悉 RT-1, RT-2, Pi0, GR00T, ACT 等 VLA 模型

目标检测与分割: RCNN 系列 (Fast, Faster, Mask, Cascade), DETR, GLIP, GLIP2, SAM

预训练与对比学习: 掩码建模 (BERT, MAE, BEiT, EVA), MOCO, CLIP, DINO

大型语言与多模态模型: 熟悉 GPT, MoE, RoPE, RAG, Engram, 预训练 GPT2 规模的中文语料和 RLHF 后训练的经验

生成式模型: GAN, VAE, VQ-VAE, VQGAN, 隐空间扩散模型

SLAM 与 3D 视觉: PTAM, FAST-LIVO2, 3DGS, HI-SLAM2, GS-LIVO