

教育背景

赫瑞瓦特大学

博士研究生

导师: [Ahmed H. Elsheikh](#)、[Nanzhe Wang](#)

研究方向: (i) 借助更具表达能力的策略类 (*diffusion policy* 与 *flow policy*) 提升基于世界模型的强化学习 (*model-based RL*) 的数据效率; (ii) 将基于世界模型的强化学习算法应用于流动控制等工程问题

广州大学

机器人工程 工学学士

英国 · 爱丁堡

10/2023–01/2028 (预计)

中国 · 广州

09/2019–06/2023

代表性论文

- [1] **M. Ye**, A. H. Elsheikh. “DPMPC: Diffusion Policies as Expressive Priors for Model Predictive Control.” *ICLR 2027*, 审稿中。
- [2] **M. Ye**, A. H. Elsheikh. “Model-based reinforcement learning for active flow control.” *Physics of Fluids*, 2025.
- [3] **M. Ye**, J. Zhang. “Mobip: A lightweight model for driving perception using MobileNet.” *Frontiers in Neurorobotics*, 2023.

研究经历

博士研究生, 赫瑞瓦特大学, 英国 · 爱丁堡

10/2023–至今

- **DPMPC** ——以 *Diffusion Policy* 作为模型预测控制 (*MPC*) 的策略先验:
 - 发现采用 Gaussian 策略先验的 MPC 方法持续存在 29–82% 的价值估计误差, 并将其根源归结为该先验无法表示规划器的多峰动作分布
 - 提出一套高效训练 *diffusion policy* 先验的方法, 使其匹配规划器的动作分布, 并将其集成到 MPC 框架 (TDMPC) 中
 - 在三个高难度基准 (DMControl、MyoSuite 和 HumanoidBench, 共 14 个任务) 上显著超越强基线, 在 61 维动作空间的 Unitree H1 人形机器人控制任务上提升尤为明显
- 用于主动流动控制的世界模型方法:
 - 从零设计两个控制问题 (状态空间、奖励函数等) ——圆柱绕流的合成射流减阻, 以及地下油藏产量最大化
 - 将两类基于世界模型的算法——规划方法 (PETS) 与 Dyna 式强化学习 (MBPO) ——适配到流动控制问题上; 引入多步预测损失改进世界模型, 缓解自回归 rollout 中的误差累积问题
 - 两种 MBRL 算法均优于 PPO 基线: 样本效率提升 2–9 倍, 实际运行时间下收敛速度快 2–7 倍; 代码已在 Gym-preCICE 生态中开源

访问研究员, 同济大学, 中国 · 上海

09/2025–11/2025

- 面向地下流动控制的降阶世界模型:
 - 利用本征正交分解 (POD) 将高维流场状态 (压力场与速度场) 压缩为紧凑的隐空间表示, 并直接在隐空间中学习动力学模型; 隐空间模型的 rollout 速度快于全阶模型, 从而支持实时规划与控制

访问研究员, NORCE Research, 挪威 · 卑尔根

07/2025–08/2025

- 有限观测条件下的地下钻井轨迹控制:
 - 将实时地下钻井 (geosteering) 建模为部分可观测控制问题: 钻井轨迹需始终保持在储层内, 但钻头前方的地质情况未知, 只能观测到间接且含噪声的伽马测井数据
 - 设计世界模型以预测周围地层的伽马响应, 得到地下地层的概率预测模型; 将该预测与考虑不确定性的基于模型的规划相结合, 降低钻出储层的风险

研究助理，广州大学，中国·广州

03/2022–09/2023

• *MobiP* ——轻量级多任务驾驶感知：

- 构建轻量级自动驾驶多任务感知网络，可在嵌入式硬件而非专用 GPU 上运行，拓宽了此类算法的部署场景
- 设计共享 MobileNetV2 编码器 + 双解码头的架构——一个检测头负责交通目标检测，一个多类别分割头覆盖可行驶区域与车道线
- MobiP 仅有 7.5M 参数，性能与 YOLOP (SOTA 方法) 相当，速度快 45%：在 NVIDIA Tesla V100 上达到 58 FPS，在 Raspberry Pi 4B 上达到 4.8 FPS

荣誉奖项

- James Watt 奖学金（全额资助博士），赫瑞瓦特大学 2023
- 广州大学二等奖学金（前 5%），两次获得 2020, 2023
- 第十二届全国大学生数学竞赛三等奖 2020

受邀报告

- “Model-Based Reinforcement Learning for Efficient Subsurface Flow Control”, NORCE Research, 挪威·卑尔根 08/2025

专业技能

编程语言： Python, C, MATLAB, Simulink, $\text{\LaTeX}$, HTML/CSS

机器学习： PyTorch, JAX, OpenCV

机器人： ABB IRB 1600, Arduino, Raspberry Pi, STM32, MuJoCo, Gazebo

软件工具： SLURM, Docker

语言： 中文、英语