

# 郭赫天

吉林大学人工智能博士在读 | 研究兴趣: VLA 策略学习、预测式世界模型、视觉触觉操控与全身动作生成

hetiangouml@gmail.com +86 186 8239 8096 个人主页 GitHub

## 教育经历

|                                              |            |
|----------------------------------------------|------------|
| 吉林大学·工学博士，人工智能                               | 2026.08–至今 |
| 研究方向: 具身智能、VLA、世界动作模型与灵巧手控制; 导师: 范子沛教授       |            |
| 佐治亚大学·理学硕士，工程学（机械工程方向）                       | 2024–2025  |
| 研究方向: 智能交通系统、自动驾驶与软硬件联合仿真; 导师: Yunli Shao 教授 |            |
| 南方科技大学·工学学士，计算机科学与技术                         | 2020–2024  |
| 研究方向: 智能交通系统、时空数据挖掘与交通预测; 导师: 宋轩教授           |            |

## 研究与实习经历

|                                                                                                                     |                        |
|---------------------------------------------------------------------------------------------------------------------|------------------------|
| AI 算法实习生·深圳市道和通泰机器人有限公司 (Avant Robotics)                                                                            | 深圳, 中国 2026.01–至今      |
| 在宇树 G1 上部署 Sonic、GMR + AMO 动作执行链路并接入 Inspire 灵巧手; 采集多任务真机遥操作数据, 完成 Psi-0、GR00T N1.7 的微调与部署。                         |                        |
| 开展 SeMoCo 语言条件全身动作生成研究, 成果投稿 AAAI 2027; 将生成动作经 Sonic 映射并在 G1 真机回放, 验证从模型生成到机器人执行的链路。                                |                        |
| 研究实习生·LocationMind Inc.                                                                                             | 东京, 日本 2023.08–2023.10 |
| 研究北海与波罗的海 AIS 航迹中的航路网络提取、交通流趋势/周期分解及静态-动态空间关系建模; 参与构建海事交通流预测模型, 成果发表于 <i>Transportation Research Part E</i> (2025)。 |                        |

## 精选研究与论文

### PLaW-VLA: Predictive Latent World Modeling for Vision-Language-Action Policies

共同第一作者 | CoRL 2026 已接收

- 以预训练视觉表征为预测目标, 在潜在空间建模未来状态, 使 VLA 从任务相关的动力学信息而非像素重建中获得动作上下文。
- 联合预训练状态预测与动作生成, 在 LIBERO / LIBERO-Plus 分别取得 97.4% / 72.7% 平均成功率。

### BeliefVLA: Implicit Belief-State Driven Vision-Language-Action Models for Robust Manipulation

第二作者 | IROS 2026 已接收

- 从历史观测中推断连续隐式 belief state, 将时序记忆与物理动力学先验注入 VLA 策略。
- 利用 V-JEPA2 的预测表征替代单帧反应式决策, 在遮挡和观测扰动场景下提升操控鲁棒性。

### SeMoCo: A Semantic-First Motion Codec for Motion Language Modeling

共同第一作者 | AAAI 2027 投稿中

- 将动作语义与运动学残差分别量化, 构建兼顾语义表达与细节重建的分层动作 token。
- 设计跨时间建模语义、逐 token 细化运动学的双轴生成器, 并构建统一骨架的大规模动作数据集  $\Omega$ -MotionVerse。

### TacDyn-WAM: Learning Implicit Tactile Dynamics in a Heterogeneous Visuo-Tactile World Action Model

ICLR 2027 投稿中

- 学习动态感知的触觉表征, 在多个时间尺度预测潜在接触状态及其变化, 避免直接重建触觉像素。
- 通过视觉与触觉专家联合建模未来状态和动作, 在 UniVTAC 上取得 81.5% 的平均成功率。

## 智能交通系统 (ITS) 相关研究

### HHGNN: Heterogeneous Hypergraph Neural Network for Traffic Agents Trajectory Prediction in Grouping Scenarios

第一作者 | ICRA 2024

- 利用异质超图建模群组交通场景中的高阶交互, 预测交通参与者轨迹。

### A Decomposable Multi-Fusion Spatio-Temporal Marine Traffic Flow Forecasting Algorithm: Taking the North Sea and Baltic Sea Region as an Example

第二作者 | *Transportation Research Part E*, 2025

- 融合航路网络与多尺度时空关系, 对北海和波罗的海海事交通流进行预测。

## 研究与工程能力

研究方法: VLA / 世界动作模型、潜在世界状态预测、视觉触觉表征、语言条件全身动作生成; 策略微调与部署。

真机系统: 双臂/人形操控、多任务遥操作数据采集与动作执行; 使用过 Astribot S1、Realman、ARX R5、AgileX PIPER、宇树 G1。

实现工具: Python、C++、PyTorch、ROS / ROS2、Transformers、Diffusers; Sonic、GMR + AMO。