



2024 AI+研发数字峰会

AI+ Development Digital summit

AI驱动研发迈进数智化时代

中国·上海 05/17-18

世界模型增强下的自动驾驶

严骏驰 上海交通大学

科技生态圈峰会 + 深度研习



—1000+ 技术团队的选择



K+全球软件研发行业创新峰会

时间: 2024.06.21-22



K+思考周®研习社

时间: 2024.10.17-19



K+思考周®研习社

时间: 2024.11.10-12



K+峰会详情



Ai+研发数字峰会

时间: 2024.05.17-18



Ai+研发数字峰会

时间: 2024.08.16-17



Ai+研发数字峰会

时间: 2024.11.08-09



AiDD峰会详情

目录

CONTENTS

1. 世界模型介绍
2. 世界模型增强下的自动驾驶 – 视频生成
3. 世界模型增强下的自动驾驶 – 闭环决策



严骏驰

上海交通大学计算机系教授/国家优青

严骏驰，上海交通大学计算机系教授，支部书记，CCF杰出会员/杰出讲者。国家优青、IET Fellow、科技部2030新一代人工智能重大项目、基金委人工智能重大研究计划项目负责人，教育部深度学习资源建设首席专家。曾任IBM研究院首席研究员。主要研究兴趣为机器学习及交叉应用。发表第一/通讯作者CCF-A类论文超百篇，谷歌引用过万次，成果先后获PaperDigest评选的最具影响力AAAI21、IJCAI23论文榜首，及陕西省自然科学一等奖（排三）。长期任机器学习三大会议ICML/NeurIPS/ICLR领域主席，模式识别旗舰期刊PRJ编委。指导学生获得挑战杯特等奖、CCF优博/CV新锐奖、交大学术之星等荣誉和本科生自然科学基金。

PART 01

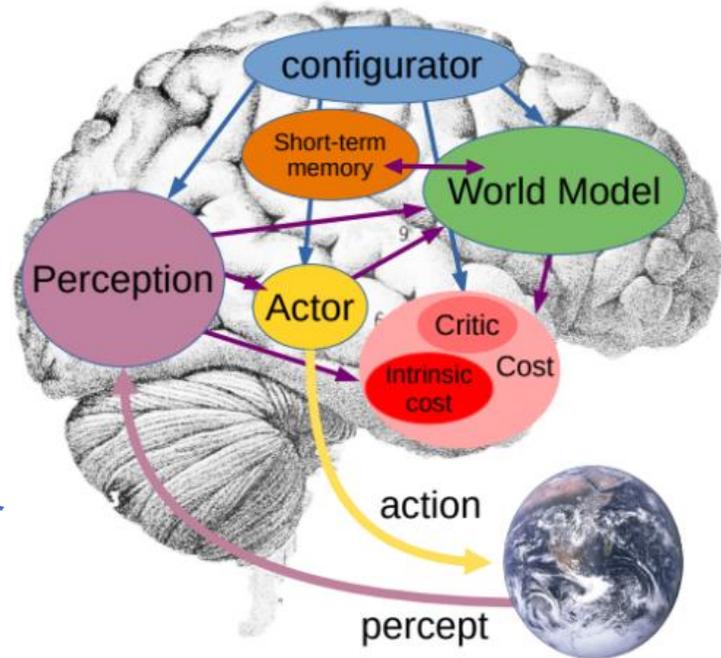
世界模型介绍

▶ 世界模型介绍

Yann LeCun构想的人脑模块[I]

- 总控 (Configurator) : 协调所有模块
- 感知 (Perception) : 获得环境当前状态
- 评价 (Cost) : 评估在当前以及可能的未来下, 个体的“舒适度”
- 短期记忆 (Short Memory) : 存储当前状态以及预测的未来以及对应的“舒适度”
- 行动 (Actor) : 给出在当前状态下最大化个体“舒适度”的行为
- **世界模型 (World Model) : 给定当前状态, 基于打算做的动作, 预测可能的未来状态**
- “人类和动物能够通过观察和以独立于任务、**无监督**的方式学习大量有关世界如何运作的背景知识, 也就是我们所谓的**常识 - 世界模型**...世界模型模块构成了体系结构中最复杂的部分。其角色是双重的: (1) **估计**感知模块未提供的有关世界状态的缺失信息 (2) **预测**世界的合理未来状态。世界模型可以预测世界的自然演变, 也可以预测由执行器模块提出的一系列动作导致的未来世界状态”

[I] A Path Towards Autonomous Machine Intelligence. 2022. Yann Lecun



Yann Lecun的人脑模型[I]

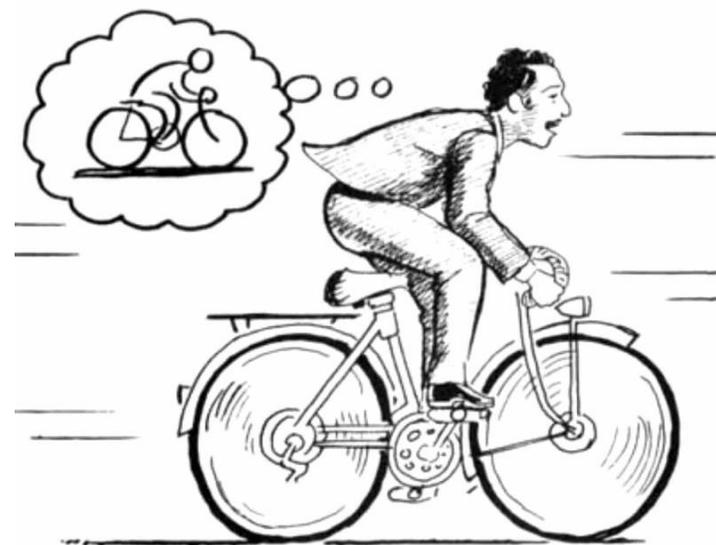


图灵奖得主 Yann Lecun

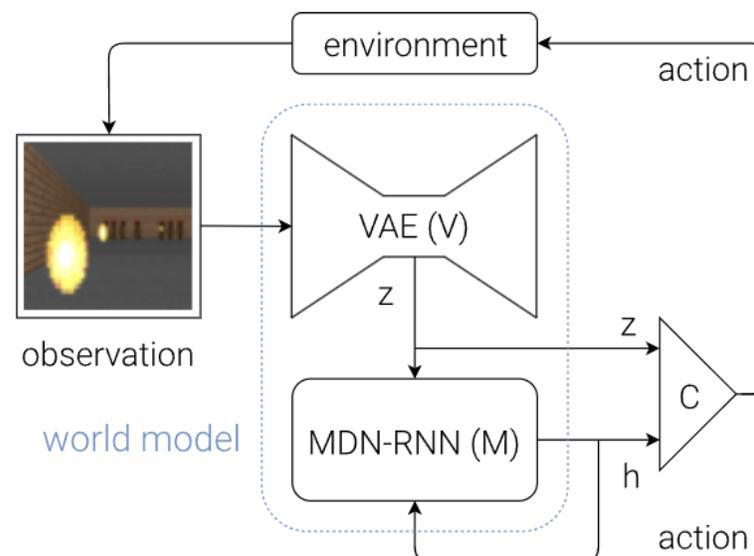
▶ 世界模型介绍

世界模型的神经网络实现初步探索[II]

- LSTM之父Jürgen Schmidhuber 在其2018年 NeurIPS的文章中，探索了神经网络能否在自己想象中的环境中学习？
- 环境状态编码器 – VAE
- 未来状态预测器 – MD-RNN
- 控制器（决策）- C
- 可以完成简单的游戏



世界模型示例。源自[I]，在[II]中被使用，而成为经典示例



Jürgen Schmidhuber



一个基于神经网络的世界模型[II]

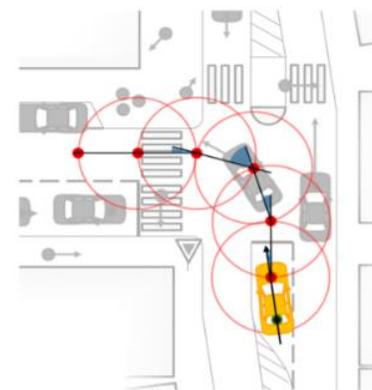
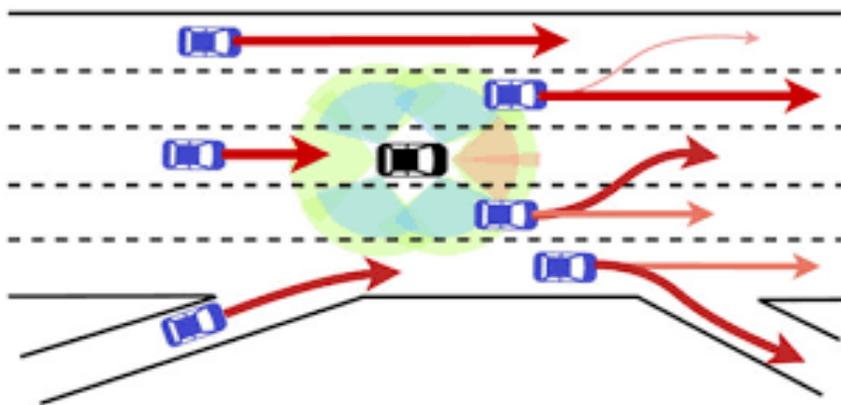
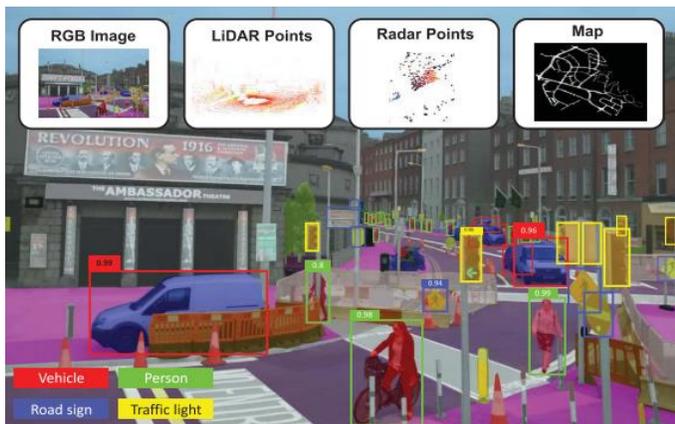
[I] Understanding comics: the invisible art. S. McCloud. Tundra Publishing. 1993

[II] Recurrent World Models Facilitate Policy Evolution. Ha, David, and Jürgen Schmidhuber. NeurIPS 2018

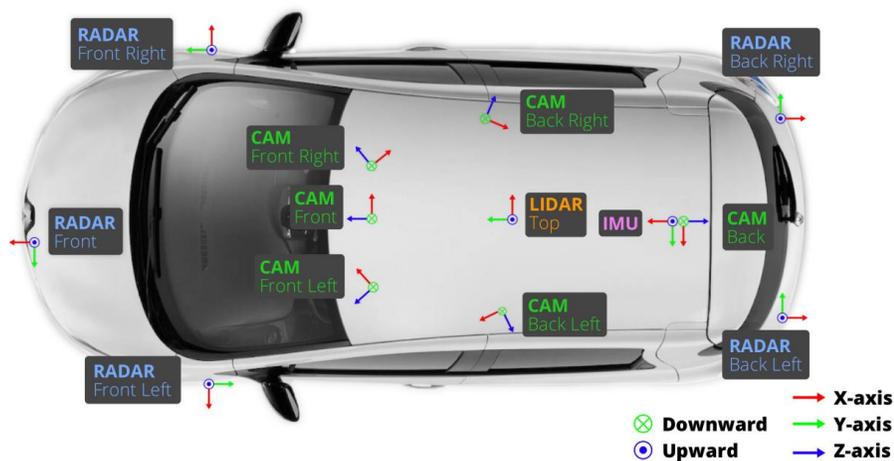
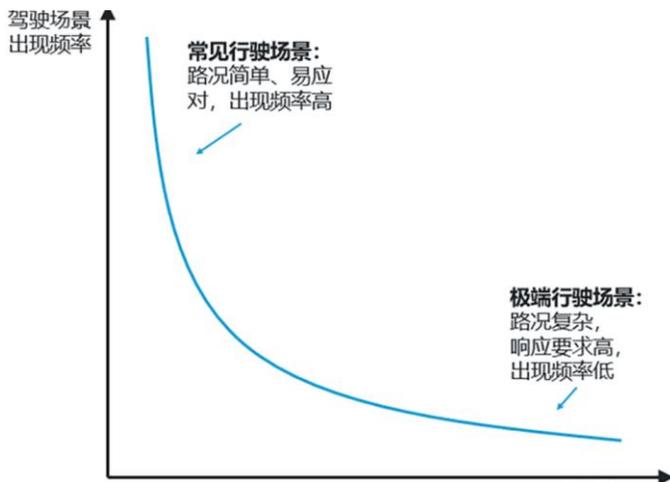
▶ 世界模型与自动驾驶

自动驾驶

- 子任务多且繁杂（感知、预测、决策） -> 标注成本高



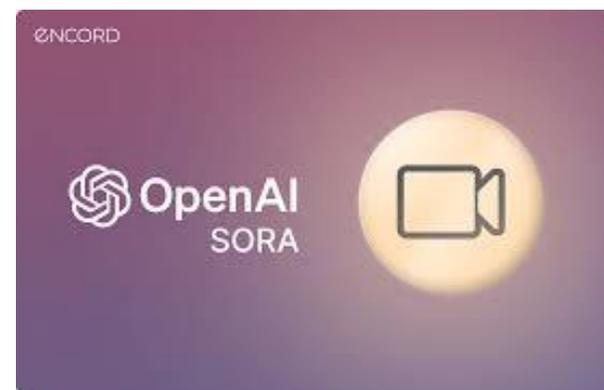
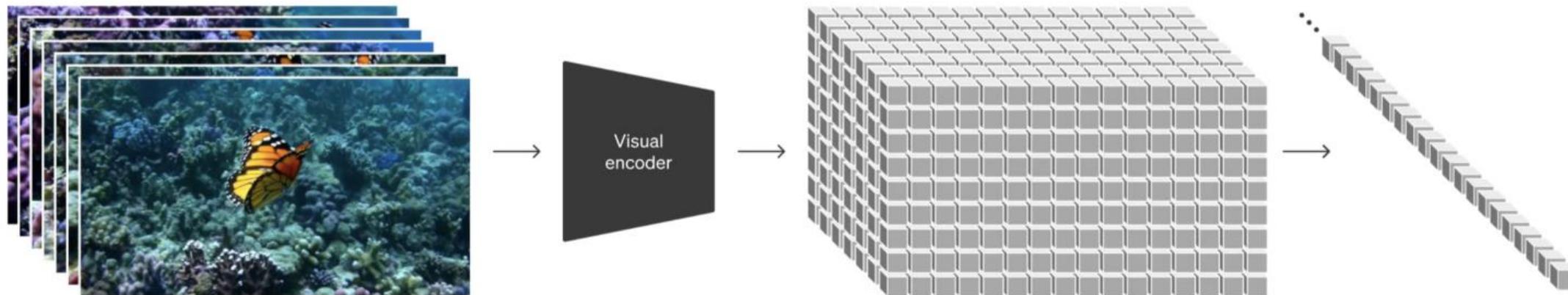
- 数据成长尾分布，多样性要求高 -> 要求丰富的时间、地点、行为 -> 高采集成本



▶ 世界模型与自动驾驶

视频生成模型

- OpenAI – Sora展示了**无监督Scale Up**视频生成模型的强大潜力



Video by Sora: "SUV in the dust"

▶ 世界模型与自动驾驶

世界模型增强下的自动驾驶

- 显式生成逼真视频 -> 数据增强, 感知训练

GAIA-1

DriveDreamer/DriveDreamer2

GenAD

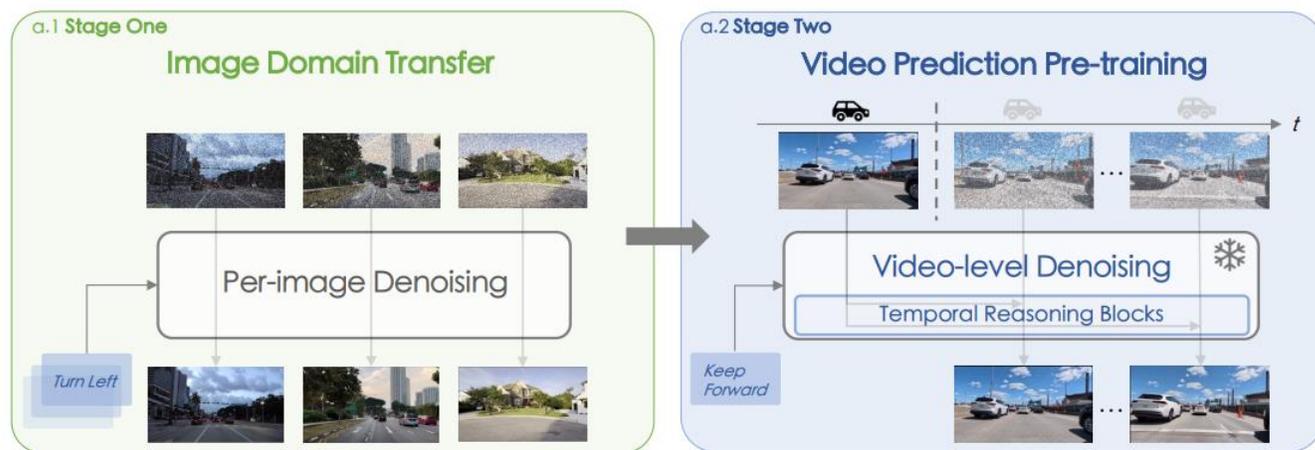
Drive-WM

ADriver-I

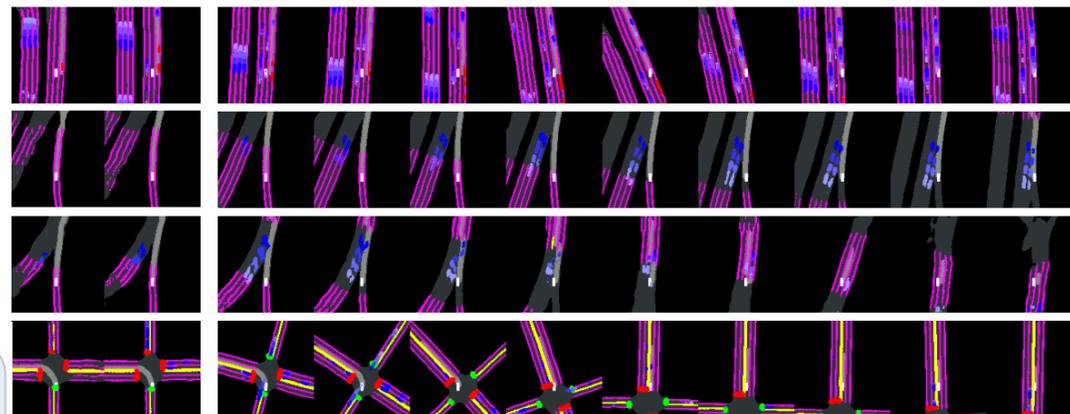
etc

- 隐空间预测未来世界 -> 神经仿真器, 闭环决策训练

Think2Drive (上海交通大学)



视频生成模型, 以GenAD为例



隐空间预测模型, 以Think2Drive为例

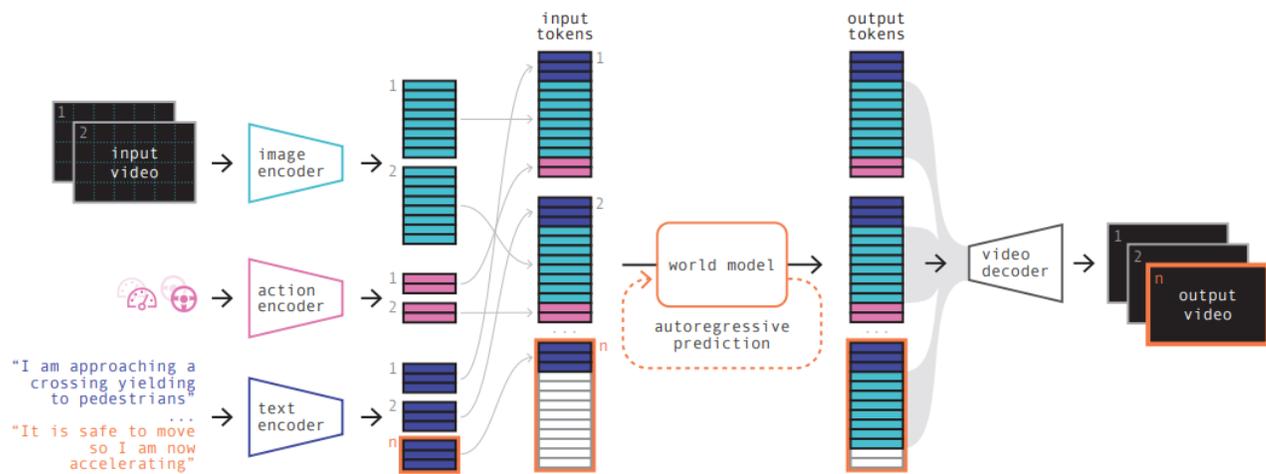
PART 02

世界模型增强下的自动驾驶 – 视频生成

▶ 世界模型增强下的自动驾驶 – 视频生成

GAIA-1 (Generative Artificial Intelligence for Autonomy, Wayve, 2023)

- 世界模型: Autoregressive Transformer, 预测下一个图片Token (仅前视相机)
- 解码器: 将预测出的Token转换成视频, 可以于世界模型分别训练
- 4700小时训练集, 世界模型6.5B参数 (128 A100 * 15天), 解码器2.6B参数 (32 A100 * 15天)



GAIA-1模型图

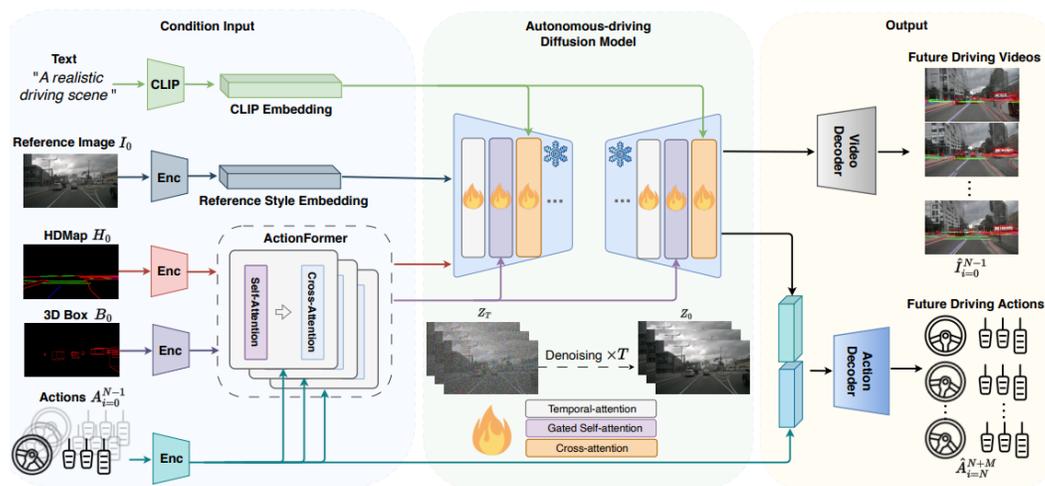


GAIA-1: A Generative World Model for Autonomous Driving. arXiv:2309.17080

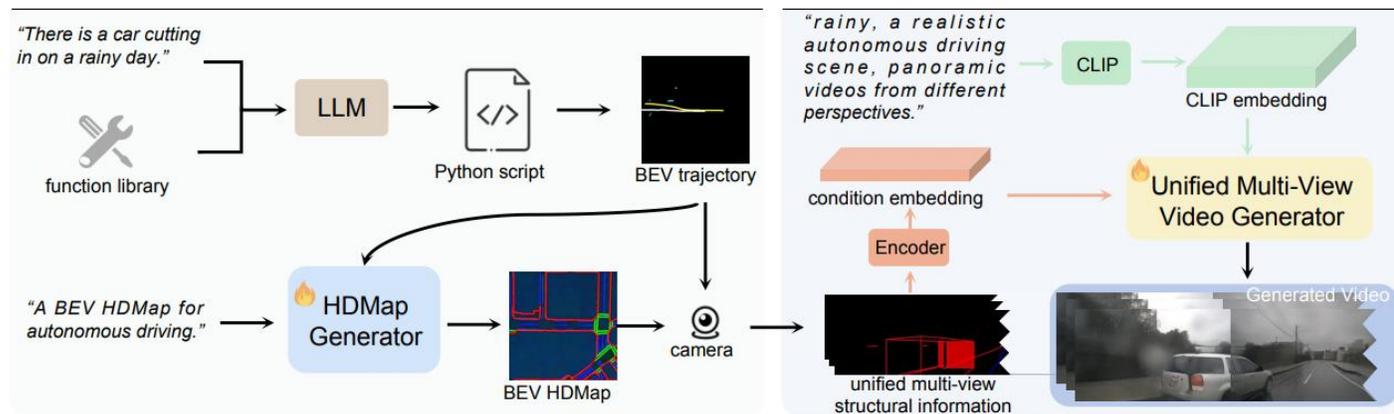
▶ 世界模型增强下的自动驾驶 – 视频生成

DriveDreamer (GigaAI+清华大学, 2023)/DriveDreamer-2 (GigaAI+自动化所, 2024)

- 专注Conditioned Generation, 给定3D框、高精地图、自车动作、文本描述等约束, 同时生成未来视频与动作
- 分阶段训练生成图片、视频, 可利用图片预训练
- V2版本和LLM结合自动化生成车辆轨迹, 根据轨迹训练一个高精地图生成器, 扩充数据
- nuScenes: 3.89小时训练数据



DriveDreamer模型图



DriveDreamer2模型图

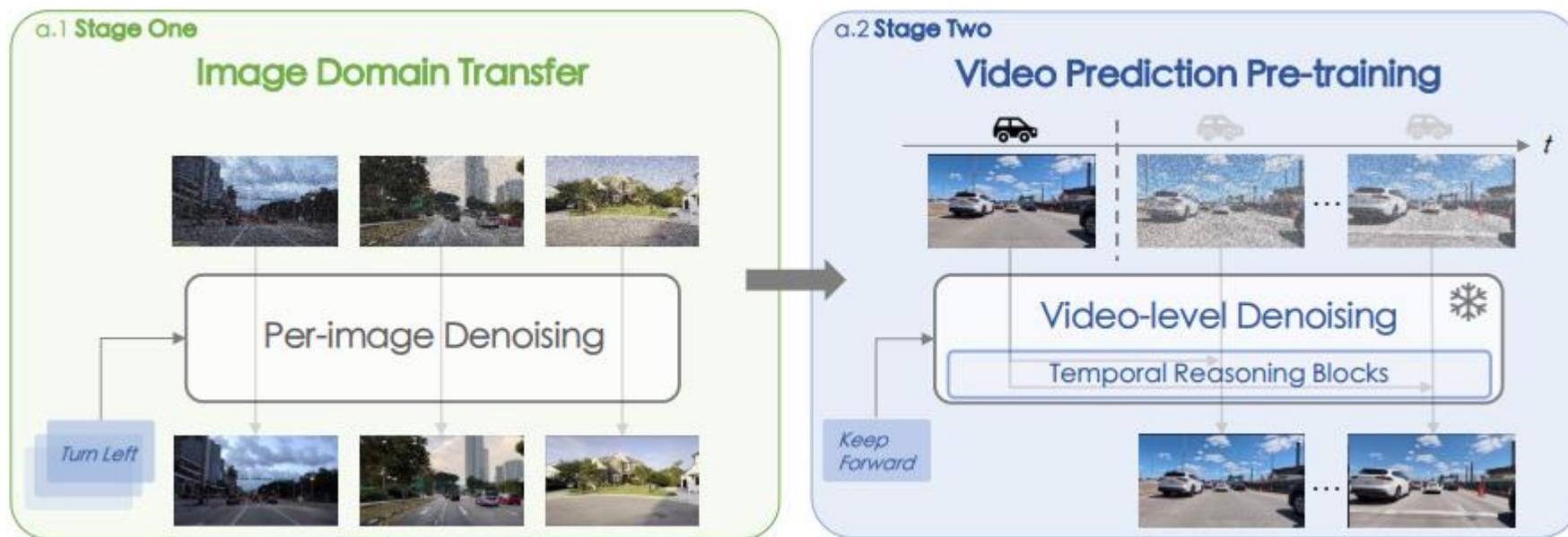
DriveDreamer: Towards Real-world-driven World Models for Autonomous Driving. arXiv:2309.09777

DriveDreamer-2: LLM-Enhanced World Models for Diverse Driving Video Generation. arXiv 2403.06845

▶ 世界模型增强下的自动驾驶 – 视频生成

GenAD (CVPR 2024, 上海人工智能实验室OpenDriveLab)

- Language Condition + 仅前视
- 自采公开数据: 2000+小时Youtube视频, 64 A100
- Zero-shot 泛化数据集能力



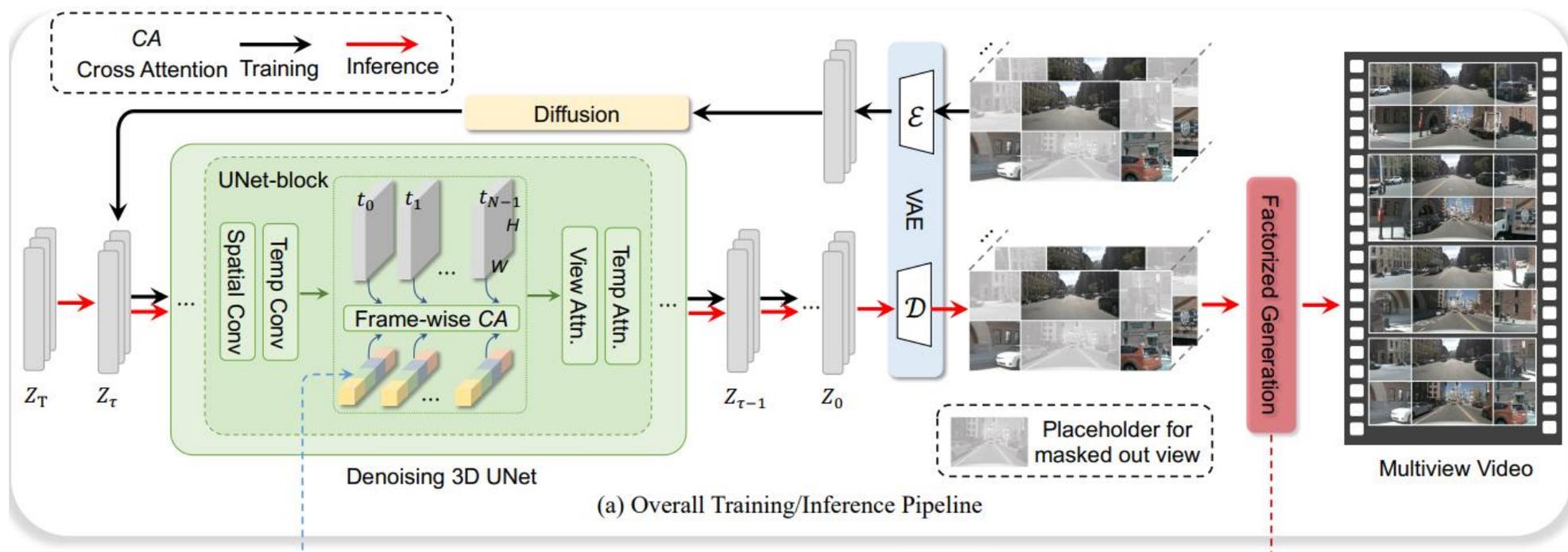
GenAD模型图

Generalized Predictive Model for Autonomous Driving. CVPR 2024

▶ 世界模型增强下的自动驾驶 – 视频生成

Drive-WM (CVPR24, 自动化所)

- Conditioned Generation + 环视一致性
- 结合下游目标检测与高精地图检测模型，计算未来帧的合法性，进而通过Tree-Search，完成决策
- nuScenes数据集：3.89小时训练数据



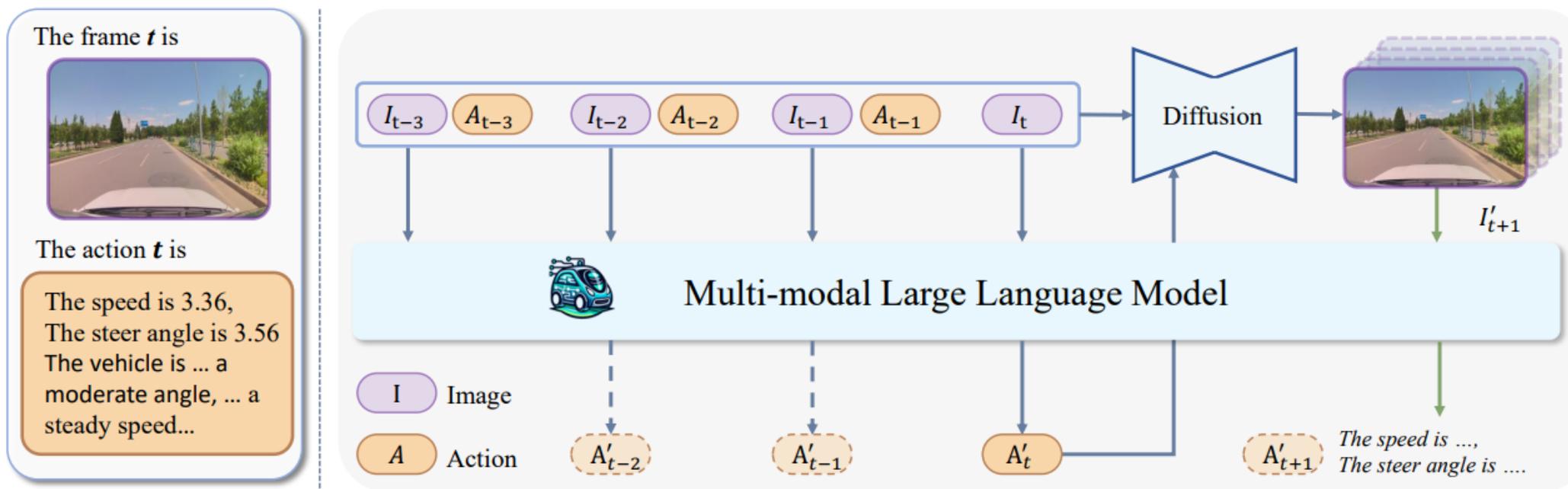
Drive-WM模型图

Driving into the Future: Multiview Visual Forecasting and Planning with World Model for Autonomous Driving. CVPR 2024

▶ 世界模型增强下的自动驾驶 – 视频生成

ADriver-I (旷视, 2023)

- 图文大模型：根据当前图片与自车状态，产生动作
- 扩散视频生成模型：根据当前图片、自车状态、打算执行的动作，产生未来图片
- 二者可联合进行rollout



ADriver-I模型图

ADriver-I: A General World Model for Autonomous Driving. arXiv 2311.13549

▶ 世界模型增强下的自动驾驶 – 视频生成

- 总结

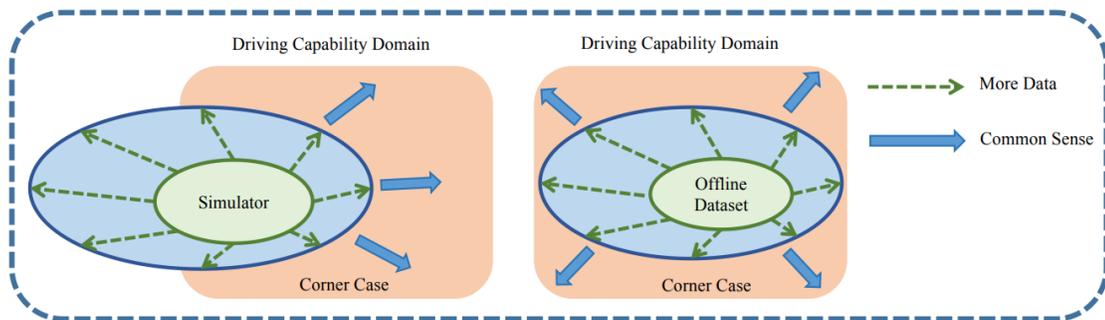
方法名	输入		输出		数据规模
	Condition 3D框	Condition 高精地图	环视输出	预测动作	
GAIA-1	x	x	✓	✓	4700小时
DriveDreamer/ DriveDreamer2	✓	✓	x/✓	✓	3.89小时
GenAD	x	x	x	✓	2000小时
Drive-WM	✓	✓	✓	✓	3.89小时
ADriver-I	x	x	✓	✓	3.89小时

ADriver-I: A General World Model for Autonomous Driving. arXiv 2311.13549

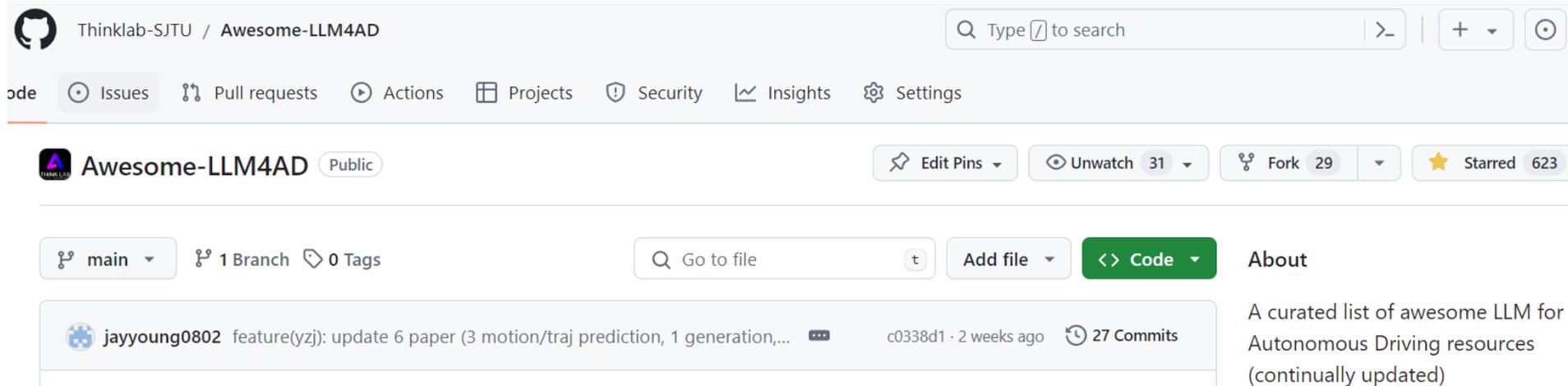
▶ 世界模型增强下的自动驾驶 – 视频生成

- 总结

通过结合Foundation Model中的常识，视频生成有望为长尾问题的解决提供助力



更多讨论见 [LLM4Drive: A Survey of Large Language Models for Autonomous Driving. \(arxiv.org/abs/2311.01043\)](https://arxiv.org/abs/2311.01043)

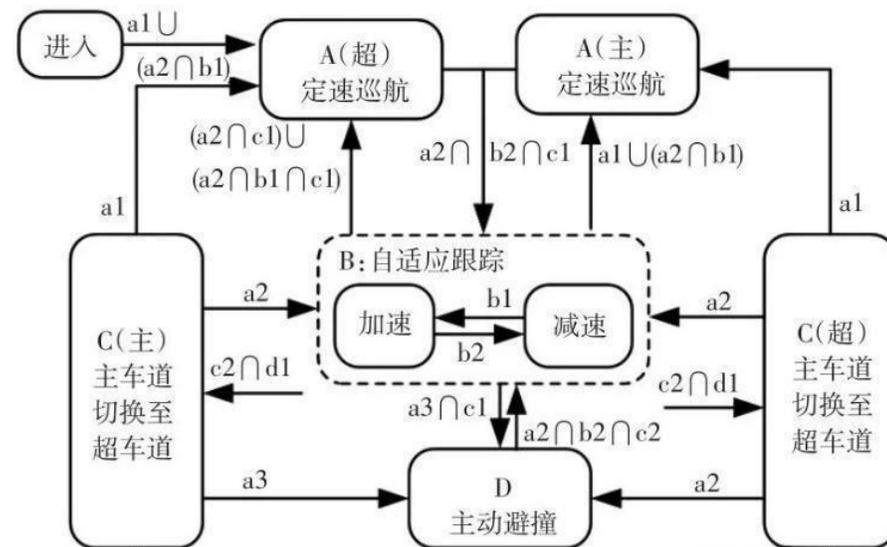
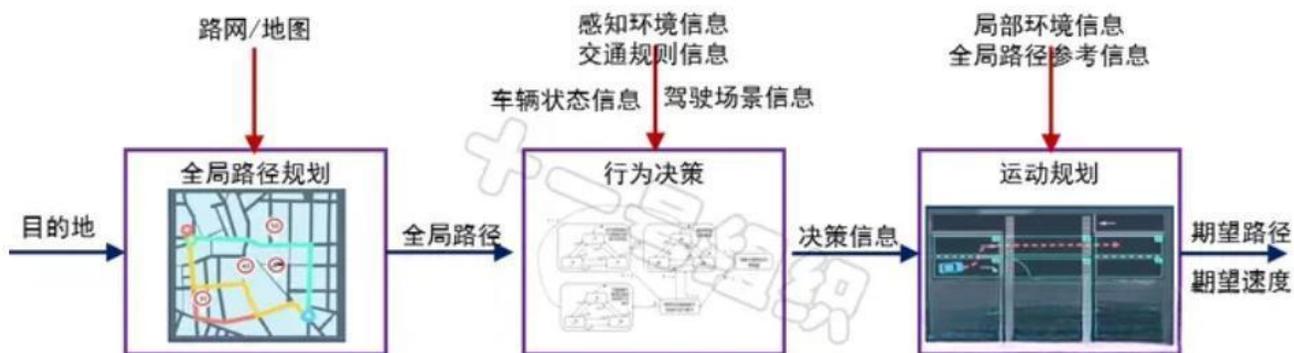


PART 03

世界模型增强下的自动驾驶 – 闭环决策

世界模型增强下的自动驾驶 - 闭环决策

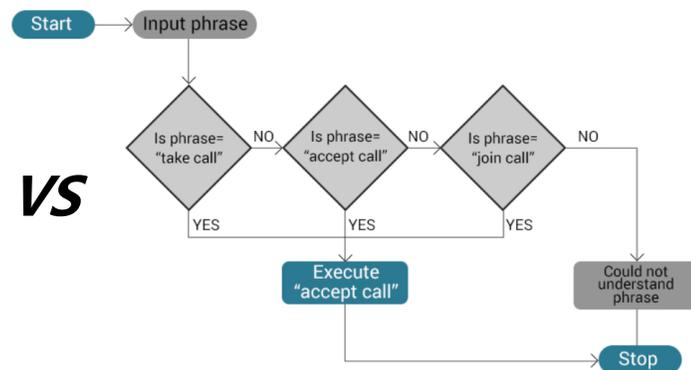
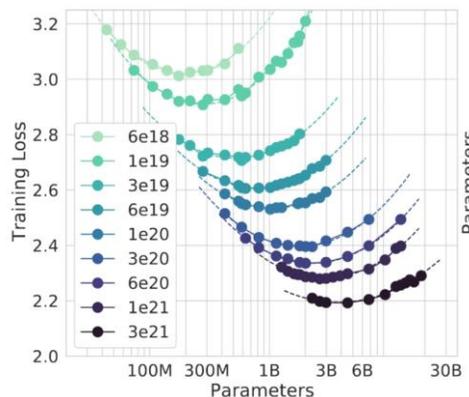
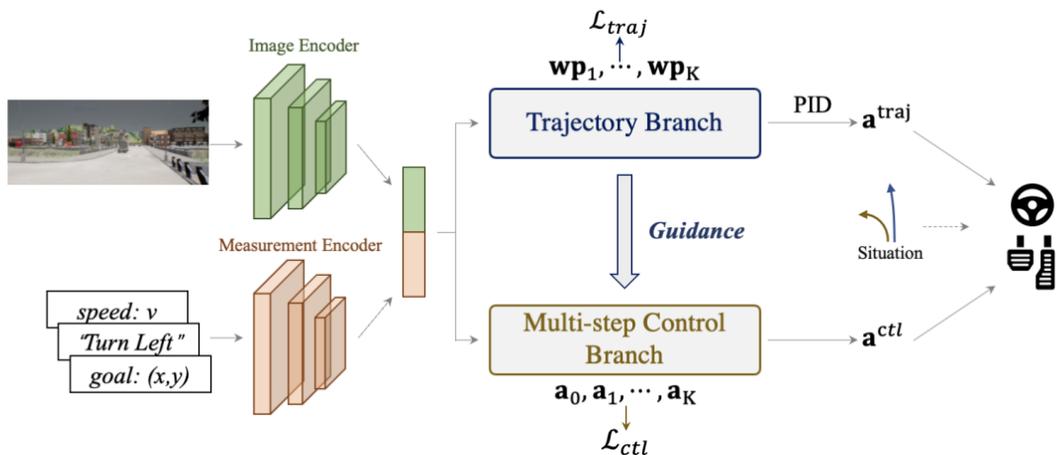
传统的基于规则的车端自动驾驶系统决策方案：



- 状态机 + 路径搜索 + 运动规划等多个模块结合的**人工**定义决策方案
- “小鹏官方称，与高速公路NGP（Navigation Guided Pilot 导航辅助驾驶）相比，城市NGP的代码量是高速NGP的6倍、感知模型数量多出4倍、**预测和规控的代码量多出了88倍**”
- 复杂场景下的手工设计决策的难度指数上升，长尾场景越多，新增代码难度越大（新旧规则兼容） -> 反向 Scaling Law

▶ 世界模型增强下的自动驾驶 - 闭环决策

基于学习的数据驱动的车端自动驾驶系统决策方案:



VS

- ChatGPT击败基于规则的聊天机器人 -> **数据驱动**范式
- 从人类专家数据与自主探索中学习: 数据 + 算力 + 训练时间越多, 性能越好 -> Scaling Law
- 自动驾驶公司特斯拉CEO马斯克: 最新的FSDV12中, 基于神经网络的决策模块代码量相比基于规则的**减少30万行**



Replying to @Scobleizer

Vehicle control is the final piece of the Tesla FSD AI puzzle. That will drop >300k lines of C++ control code by ~2 orders of magnitude.

It is training as I write this. Our progress is currently training compute constrained, not engineer constrained.

7:04 AM · Aug 2, 2023

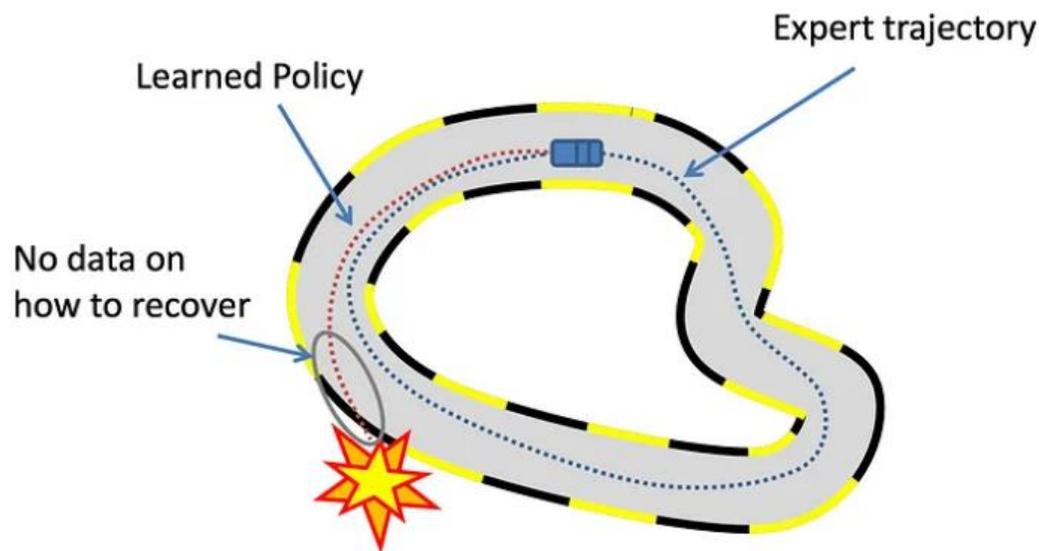
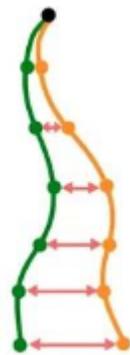
5.4K Reply Copy link

Read 371 replies

▶ 世界模型增强下的自动驾驶 – 闭环决策

• 模仿学习

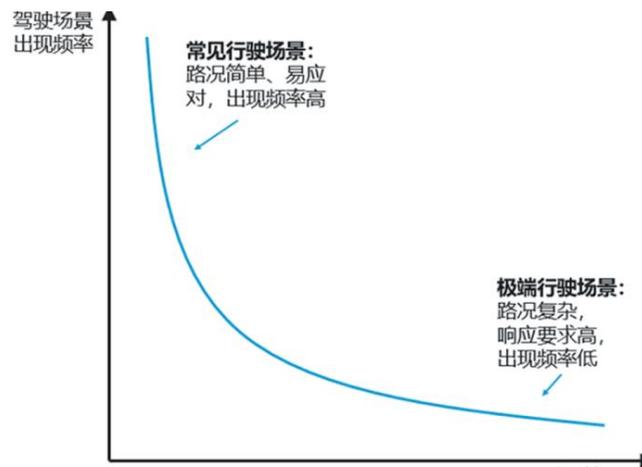
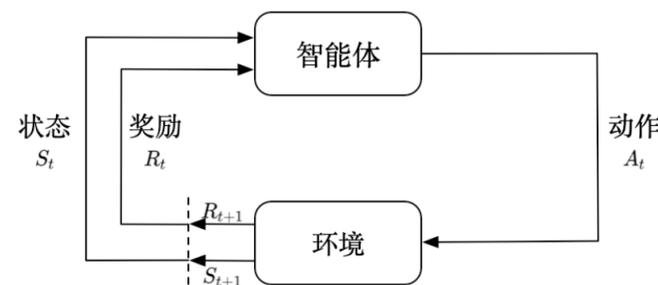
- 监督学习人类驾驶数据
- 只知道怎么开是对的，不知道怎么开是错的，不知道怎么从错误中回复（知其然，不知其所以然）
- **因果倒置、鲁棒性差、泛化能力差、闭环实车评测效果不佳[1]**



[1] Parting with Misconceptions about Learning-based Vehicle Motion Planning, CoRL 2023

• 强化学习

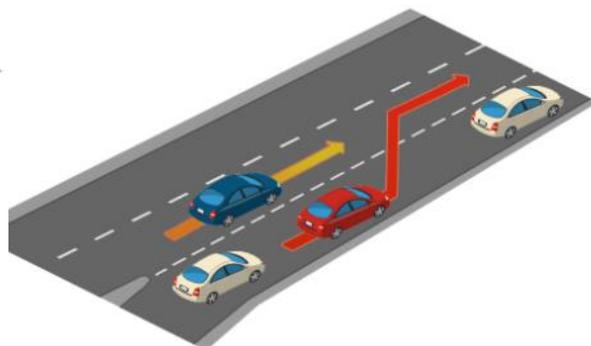
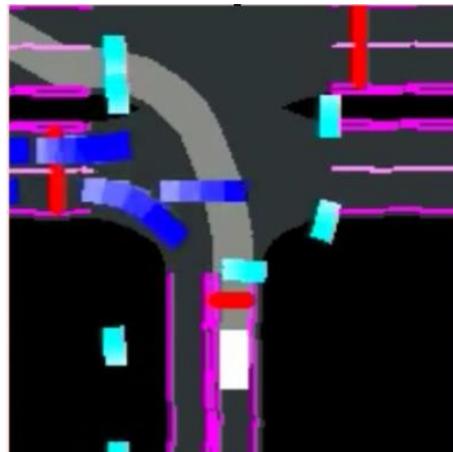
- 通过探索学习驾驶策略
- 奖励与惩罚 -> **知其所以然**
- 随机探索解决自动驾驶中的长尾分布**效率**存在挑战



▶ 世界模型增强下的自动驾驶 – 闭环决策

• 已有的基于强化学习的方案 – Roach

- 基于经典Model-Free RL模型PPO，纯决策模块
- 使用鸟瞰图下的（BEV）表征，解耦感知与决策
- 可学会基本驾驶技能：跟车、变道、急刹、红灯停绿灯行
- 探索效率相比Model-based RL较低，**不具备**超车、避障、汇流、窄道会车等复杂交互能力



End-to-End Urban Driving by Imitating a Reinforcement Learning Coach. ICCV 21 (Oral)

▶ 世界模型增强下的自动驾驶 – 闭环决策

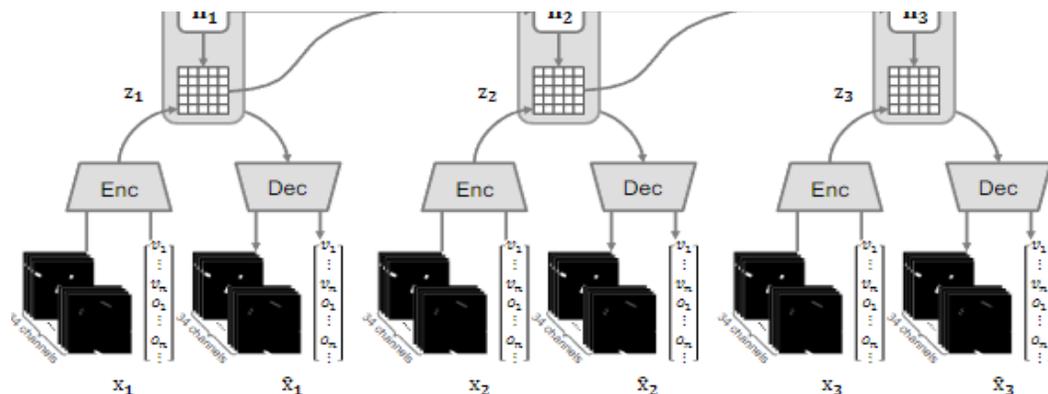
• Think2Drive

- 基于RSSM world model (Google Deepmind最新工作)
- 通过在训练出的世界模型中学习, 可提高数十倍采样效率, 模型训练仅需单卡A6000-48G GPU 3天
- 第一个, 也是唯一一个解决其中CARLA v2中所有39种长尾场景的方案

测试路径:

长度: 7千米

复杂场景: 共52个



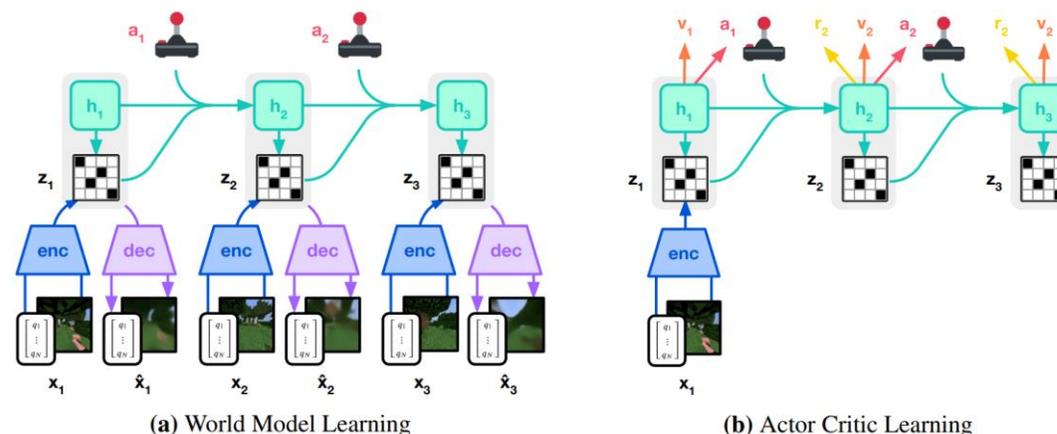
Think2Drive: Efficient Reinforcement Learning by Thinking in Latent World Model for Quasi-Realistic Autonomous Driving (in CARLA-v2). arXiv 2024

▶ 世界模型增强下的自动驾驶 – 闭环决策

• Think2Drive

- World Model实现 (RSSM) :
 - 编码器 (Encoder) : 将输入图片映射为状态 (隐变量) - **编码信息**
 - 隐空间预测模型 (Sequence Model + Dynamics Predictor + Reward Predictor + Termination Predictor) : 根据当前状态和执行动作, 预测下一时刻的状态、得到的奖励 - **预测未来**
 - 解码器 (Decoder) : 将状态转换回图片 - **提供监督信号**
- 结合World Model的强化学习
 - 1. 用当前驾驶策略网络 + 随机扰动, 在物理仿真中采集数据, 训练world model
 - 2. 在world model中, 用RL算法, 训练驾驶策略网络
 - 1、2交替进行

$$\text{RSSM} \begin{cases} \text{Sequence model:} & h_t = f_\theta(h_{t-1}, z_{t-1}, a_{t-1}) \\ \text{Encoder:} & z_t \sim q_\theta(z_t | h_t, x_t) \\ \text{Dynamics predictor:} & \hat{z}_t \sim p_\theta(\hat{z}_t | h_t) \\ \text{Reward predictor:} & \hat{r}_t \sim p_\theta(\hat{r}_t | h_t, z_t) \\ \text{Termination predictor:} & \hat{c}_t \sim p_\theta(\hat{c}_t | h_t, z_t) \\ \text{Decoder:} & \hat{x}_t \sim p_\theta(\hat{x}_t | h_t, z_t) \end{cases} \quad (2)$$



Think2Drive: Efficient Reinforcement Learning by Thinking in Latent World Model for Quasi-Realistic Autonomous Driving (in CARLA-v2). arXiv 2024

▶ 世界模型增强下的自动驾驶 – 闭环决策

• Think2Drive

- 基于世界模型的强化学习在自动驾驶中的优势：
 - 物理仿真较慢，难以大量并行，是强化学习的**探索效率瓶颈**。在世界模型的隐空间中训练，相当于Neural仿真，可利用GPU的Tensor层面并行，**单卡就可同时进行上千个环境仿真**
 - 自动驾驶BEV下的世界预测难度相对较低，易得到**好的世界模型**
- 额外结合最前沿的强化学习技术，进一步提高性能
 - **课程学习** – 由易到难
 - **优先级采样** – 在有挑战的场景中学习
 - **退火世界模型训练** – 世界模型后期已经非常精准，专注训练策略网络
 - **动作稳定性** – 减少RL算法内在的动作波动

Think2Drive: Efficient Reinforcement Learning by Thinking in Latent World Model for Quasi-Realistic Autonomous Driving (in CARLA-v2). arXiv 2024

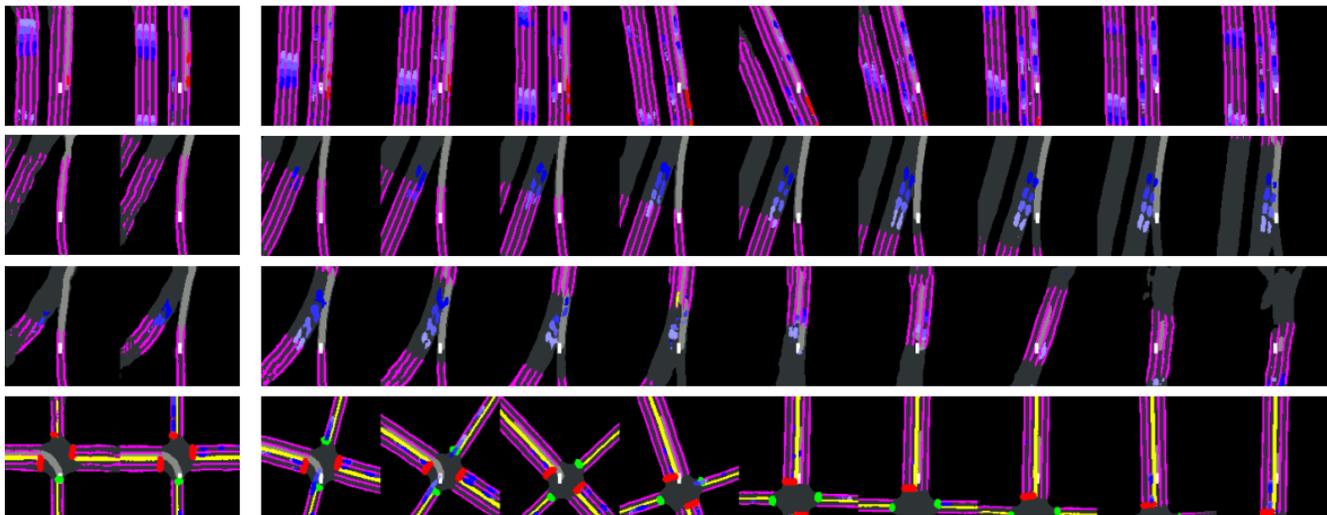
▶ 世界模型增强下的自动驾驶 – 闭环决策

• Think2Drive

- 性能断崖式领先 (91.7 vs 0.6)

Method	benchmark	Driving Scores	Weighted Driving Score	Route Complete %
Roach (Expert)	CARLA v1	85.0	–	72.0
Think2Drive (Ours)		90.2	90.2	99.7
PPO (Expert)	CARLA v2	0.7	0.6	1.0
Think2Drive (Ours)		56.8	91.7	98.6

- 世界模型 (左一输入, 后续为世界模型预测)



Think2Drive: Efficient Reinforcement Learning by Thinking in Latent World Model for Quasi-Realistic Autonomous Driving (in CARLA-v2). arXiv 2024

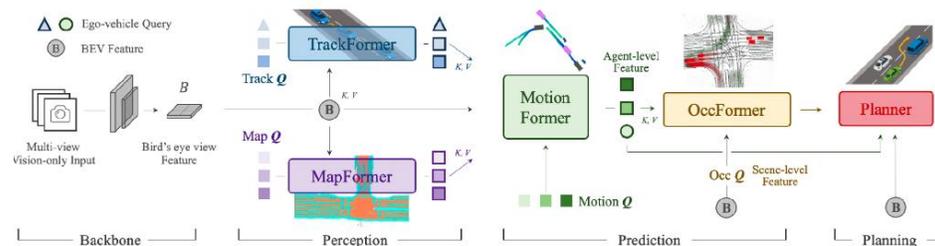
▶ 总结与展望

• 总结

- 世界模型作为一种Neural Simulator，使得自动驾驶中的策略学习可以计算驱动，**更多的GPU = 更多的探索**
- 高几个数量级级别的探索，对于解决自动驾驶策略学习中的**长尾问题**至关重要

• 展望

- 海量无监督自动驾驶数据中学习World Model
- 基于强化学习的策略网络嵌入端到端自动驾驶框架
- 强化学习 + 人类驾驶数据对齐 (SFT/RLHF)



科技生态圈峰会 + 深度研习



—1000+ 技术团队的选择



K+全球软件研发行业创新峰会

时间: 2024.06.21-22



K+思考周®研习社

时间: 2024.10.17-19



K+思考周®研习社

时间: 2024.11.10-12



K+峰会详情



Ai+研发数字峰会

时间: 2024.05.17-18



Ai+研发数字峰会

时间: 2024.08.16-17



Ai+研发数字峰会

时间: 2024.11.08-09



AiDD峰会详情

THANKS

