

目录

FSD v12 全面升级，“端到端”大模型来临..... 6

 算法与算力共振，正式步入“端到端”大模型时代..... 6

 FSD v12 版本推送，“端到端”自驾算法落地..... 8

 拥抱“端到端”的自动驾驶，打开智驾算法天花板..... 9

 数据端储备优势显著，投产超算 Dojo，训练平台算力持续领先..... 14

特斯拉 Robotaxi 官宣，无人车商业化加速..... 16

 酝酿多时，特斯拉 Robotaxi 预计 8 月发布..... 16

 Robotaxi 商业化进程：机遇与挑战并行..... 17

 展望：Robotaxi 空间广阔，远期成本大幅低于网约车..... 22

国内：智驾渗透率快速提升，算力增配仍在继续..... 24

 智驾级别：技术不断成熟，高阶智驾渗透率快速提升..... 24

 感知：摄像头数量持续加配，激光雷达在高阶智驾中仍不可或缺..... 26

 决策：算法驱动，算力增配仍在继续..... 28

 执行：线控升级，渗透率快速提升..... 30

投资建议：“端到端”技术飞跃+平价时代，引领智驾加速渗透..... 32

风险提示..... 33

图表目录

图 1：特斯拉自动驾驶发展历程..... 6

图 2：特斯拉智能驾驶硬件持续升级，HW4.0 目前已经实现全部车型上车..... 7

图 3：传统自动驾驶算法与“端到端”算法差异对比..... 9

图 4：“端到端”自动驾驶关键成就发展图..... 11

图 5：不同思路的“端到端”算法结构设计思路..... 12

图 6：UnitAD 以规划为导向的端到端自动驾驶模型..... 13

图 7：特斯拉“端到端”自动驾驶算法框架..... 13

图 8：特斯拉 FSD Beta 累计行驶超过 10 亿英里（约 16 亿公里）..... 14

图 9：特斯拉 FSD 链路..... 14

图 10：2023 年 7 月投产 Dojo，特斯拉 2024 年算力预期 100EFLOPS..... 15

图 11：特斯拉 Dojo 架构..... 15

图 12：特斯拉 D1 芯片与英伟达 A100 芯片性能对比..... 15

图 13：Dojo 系统软件栈对接 PyTorch..... 15

图 14：马斯克推特官宣下一代车型 Robotaxi 发布时间..... 16

图 15：2019 年特斯拉对于 Robotaxi 的展望..... 16

图 16：Robotaxi 商业化发展的三个阶段及关键特征..... 17

图 17：各国 Robotaxi 商业化进展程度..... 18

图 18：2023 年 DMV 自动驾驶公司车辆登记数量（辆）..... 19

图 19：2023 年 DMV 自动驾驶公司里程累积数（万公里）..... 20

图 20: 2023 年 DMV 自动驾驶公司接管情况 (单位(万公里/次))	20
图 21: Robotaxi 全无人自动驾驶的商业化运营挑战	20
图 22: Cruise 商业化进程中遇到的问题	21
图 23: 2023 年 DMV 自动驾驶路试安全员介入原因统计	21
图 24: 美国 Robotaxi“技术+整车+平台”的“金三角”合作模式	22
图 25: Robotaxi 预计在 2026 年左右成本达到拐点	23
图 26: 成熟形态下国内 Robotaxi 单车成本	23
图 27: 2021-2023 各智能驾驶级别月度销量 (万辆)	25
图 28: 2021-2023 各智能驾驶级别渗透率	25
图 29: 新能源乘用车各智能驾驶级别渗透率	25
图 30: 2023 年智能驾驶渗透率: 新能源 VS 油车	25
图 31: 2023 年 L0 及 L1 级别智能驾驶分价格带装配率	26
图 32: 2023 年 L2 及以上方案智能驾驶分价格带装配率	26
图 33: 前视摄像头渗透率持续提升	26
图 34: 前视摄像头仍以 200 万以下像素为主, 800 万方案开始渗透	26
图 35: 2023 年车端摄像头配置数量持续增加	27
图 36: 分价格带, 2023 年配置激光雷达车型销量	27
图 37: 2023 年配置激光雷达车型销量占乘用车总销量的比例	27
图 38: 2023 年配置激光雷达企业市占率	28
图 39: 2023 年中高算力方案及特斯拉 FSD 销量 (辆)	28
图 40: 2023 年中高算力方案及特斯拉 FSD 销量占比	28
图 41: 2023 年不同算力方案销量 (辆) 及占比	29
图 42: 2023 年智驾域控制器供应商市场份额	29
图 43: 2023 年智驾域控制器供应商销量 (万辆)	29
图 44: 2023 年智驾域控制器供应商市场份额变化图	30
图 45: 新能源车线控制动渗透率	30
图 46: 2023 年新能源车 One-Box 线控制动供应商市占率	31
图 47: 2023 年新能源车 Two-Box 线控制动供应商市占率	31
图 48: 小鹏汽车 MONA 定位 10-15 万级市场	32
表 1: 特斯拉 FSD 芯片与英伟达智驾芯片对比	8
表 2: 特斯拉 FSD Beta12 版本性能升级主要内容	8
表 3: 自动驾驶解决方案分类及优缺点	10
表 4: 美国主要 Robotaxi 公司情况梳理	18
表 5: 全球市场 Robotaxi 空间测算 (亿美元)	24
表 6: 线控制动能回收策略	31

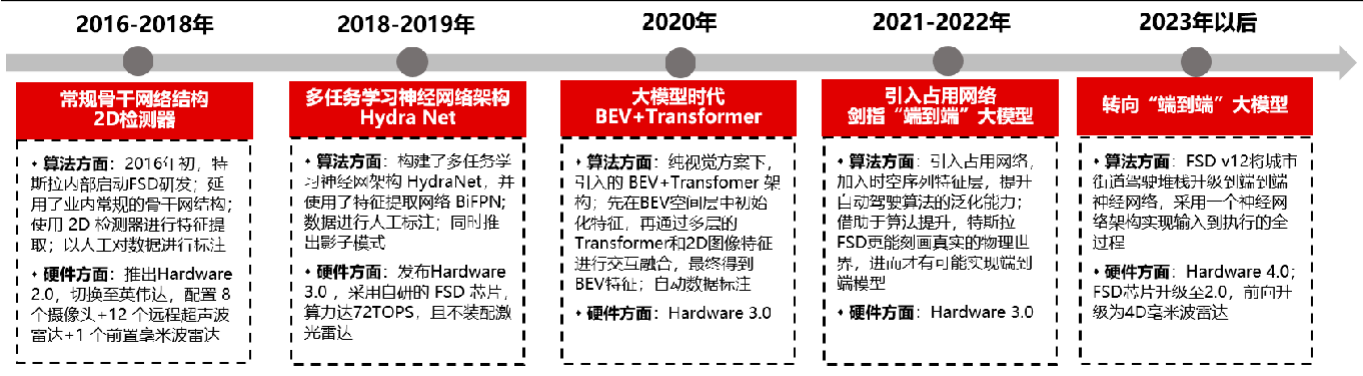
FSD v12 全面升级，“端到端”大模型来临

算法与算力共振，正式步入“端到端”大模型时代

特斯拉自动驾驶主要采用纯视觉方案，自研之路主要经历了以下五大阶段：

- 1) **2016-2018 年 · 自研早期 · 常规骨干网络 2D 检测**：2016 年 7 月，Mobileye 宣布和特斯拉终止合作，特斯拉正式开始开启自研算法之路。其实早在 2016 年初，特斯拉内部已启动 FSD 研发，主要沿用业内常规的骨干网络结构，使用 2D 检测器进行特征提取，并通过人工进行数据标记，这一套自动驾驶算法相对比较原始。硬件方面，特斯拉推出 Hardware 2.0，域控芯片切换至英伟达方案，配置 8 个摄像头、12 个远程超声波雷达、1 个前置毫米波雷达。
- 2) **2018-2019 年 · 引入多任务学习神经网络架构 Hydra Net · 创新性推出“影子模式”**。特斯拉构建了多任务学习神经网络架构 Hydra Net，并使用特征提取网络 BiFPN。相较此前算法，Hydra Net 能够减少重复的卷积计算，减少主干网络计算数量，还能够将特定任务从主干中解耦出来，进行单独微调。在融合方式上，特斯拉采用的仍是后融合策略，数据进行人工标注，且自动驾驶算法仍旧是小模型；但创新性推出“影子模式”。硬件方面，2019 年 4 月发布 Hardware 3.0，采用自研 FSD 芯片，算力达到 72TOPS，不装配激光雷达。
- 3) **2020 年 · BEV + Transformer 自动驾驶算法引领行业新路径**。纯视觉方案下，特斯拉引入 BEV+Transformer 架构，自动驾驶进入大模型时代；先在 BEV 空间层中初始化特征，再通过多层 Transformer 和 2D 图像特征进行交互融合，最终得到 BEV 特征，加强各模块信息传递与融合，同时改为自动数据标注。
- 4) **2021-2022 年 · 引入占用网络 · 提升自驾算法泛化能力**。随着自驾面临场景逐渐多元化，corner case 增多，特斯拉在感知网络架构引入了时空序列特征层，对 BEV 进行了升级，即引入占用网络，以提升自动驾驶算法的泛化能力。而借助于算法提升，特斯拉 FSD 更能刻画真实的物理世界，进而才有可能实现端到端模型。
- 5) **2023 年以后 · 转向“端到端”大模型**。FSD v12 将城市街道驾驶堆栈升级到端到端神经网络，采用一个神经网络架构实现输入到执行的全过程。硬件方面，升级至 Hardware 4.0，FSD 芯片升级至 2.0，算力持续升级，前向升级为 4D 毫米波雷达。

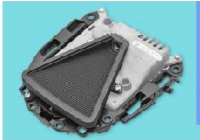
图 1：特斯拉自动驾驶发展历程



资料来源：汽车之心，

特斯拉智驾硬件全线升级至 HW4.0，感知层摄像头像素提升，增配 4D 毫米波雷达，智驾升级硬件增配，算力提升仍是智驾主旋律。特斯拉经过智驾领域 10 年投入，目前已经成为智能驾驶领先的整车厂。历史来看，特斯拉智驾经历过 HW1.0 和 HW2.0 的核心芯片外采后，2019 年官宣的 HW3.0 智驾域平台上使用 FSD1.0 芯片，加强数据自主可控能力，持续丰富“影子模式”数据集，迭代自动驾驶算法能力。2023 年 4 月，特斯拉发布 HW4.0 平台，硬件接口数量从 HW3.0 的 9 个摄像头接口减少至 HW4.0 的 8 个，但摄像头的清晰度从 120 万像素升级到 500 像素，视觉感知能力大幅提升，最远探测距离可达 424 米。同时，核心芯片升级为 FSD2.0 芯片，预计算力提升至 300-500TOPS 左右，是 HW3.0 算力的 2-3 倍，处理能力预计是 HW3.0 的三倍，同时功耗显著下滑。

图 2：特斯拉智能驾驶硬件持续升级，HW4.0 目前已经实现全部车型上车

硬件	HW1.0	HW2.0	HW2.5	HW3.0	HW4.0
域控制器					
域控制器					
上市时间	2014年9月	2016年10月	2017年8月	2019年4月	2023年4月
对应车型	Model S、Model X	Model S、Model X	全部车型	全部车型	Model 3、X、Y和S
芯片方案	Mobileye EyeQ3	Nvidia Parker Soc *1 Nvidia Pascal GPU *1 英飞凌 TriCore MCU *1	Nvidia Parker Soc *2 Nvidia Pascal GPU *1 英飞凌 TriCore MCU *1	FSD1.0 *2	FSD2.0 *2
核心芯片方案算力 (tops)	0.256	1	2	144	300-500
芯片方案分类	芯片在摄像头上	多种不同功能的芯片同时存在			SoC芯片阶段
RAM	256MB	6GB	8GB	8GB*2	8GB*2
闪存				4GB*2	4GB*2
存储方案	芯片在外购过程中数据重视程度较低			增加闪存，保证数据传输的稳定性	
处理能力	1x	40x	40x，带有冗余	420x，带有冗余	1260x，带有冗余
每秒处理帧数	36	110	110	2300	6900
功率 (w)	25	250	300	220	300
电源类型	单独电源	单独电源	冗余电源	冗余电源	冗余电源
功耗 (W/tops)	0.010	0.004	0.007	0.655	
感知层硬件					
前置摄像头	1 (EQ3前视摄像头)		35°窄幅摄像头*1 50°主摄像头*1 120°广角摄像头*1		2
车顶摄像头			0		2
侧摄像头	0			2	
侧后摄像头	0			2	
座舱摄像头			1	1	1
毫米波雷达	1 (博世)	1 (人陆)		1 (人陆，去除较多功能)	150° *1，4D毫米波雷达
超声波雷达			12		
感知层硬件方案	成熟感知层产品	感知层开发能力增强，芯片端能力短期较弱			感知层开发能力增强，补全芯片能力
阶段方案特斯拉智驾能力	智驾能力较弱	感知层开发能力增强，芯片端能力短期较弱			感知层开发能力增强，补全芯片能力

资料来源：TeslaTap，ZAKER，车东西，

智能驾驶芯片是决定智驾能力的重要因素，特斯拉自研芯片算力提升，并保持智驾域芯片的领先。智能驾驶芯片图谱来看，可将现有芯片分为低算力芯片（30TOPS 以下）、中算力芯片（30-200TOPS）和高算力芯片（200TOPS 以上）。特斯拉 FSD1.0 芯片采用 14nm 工艺，单颗算力为 72TOPS。FSD2.0 芯片单颗芯片算力显著提升，单颗芯片算力为 216TOPS。与英伟达芯片对比来看，制成上 FSD1.0 和英伟达 Xavier 采用接近的制成，FSD2.0 预测采用 7nm 工艺打造，而英伟达 Thor 的制成则预测是 5nm，算力上 FSD 落后于英伟达。

表 1：特斯拉 FSD 芯片与英伟达智驾芯片对比

项目	特斯拉 FSD1.0	特斯拉 FSD2.0 (预测)	英伟达 Xavier	英伟达 Orin	英伟达 Thor (预测)
芯片算力 (TOPS)	72	200 以上	32	128-254	2000
功率 (W)	100	-	30-40	40-60	
制成 (nm)	14	7	12	7	5

资料来源：特斯拉官网，英伟达官网，芯智讯，

FSD v12 版本推送，“端到端”自驾算法落地

在智驾领域进行多维度迭代，特斯拉自动驾驶能力快速提升，FSD v12 推送使智驾算法步入“端到端”的新纪元。2023 年 8 月 28 日，埃隆·马斯克进行了直播，试驾了特斯拉 FSD v12 版本，虽然仍为测试版本，但却是有史以来特斯拉推出的第一个端到端 AI 自动驾驶系统，标志特斯拉在自动驾驶技术上迈出了一大步。2023 年 12 月 24 日，特斯拉在北美，向一部分内部员推送 FSD Beta v12.1，新算法将城市街道驾驶堆栈升级为单一的端到端神经网络，该神经网络经过数百万个视频片段的训练，替代了超过 30 万行 C++代码。

2024 年 3 月，特斯拉在北美开始向更多 HW4 车主推送 FSD Beta v12.3 更新，同时新增“自动设定速度偏移功能”，能够综合道路类型、交通流量、环境条件、速度限制和 FSD 驾驶模式因素之后，帮助汽车更自然地与道路上的其他车辆一起行驶。3 月 26 日，马斯克宣布全美所有适用 FSD 的车辆可以免费试用 FSD 一个月。

表 2：特斯拉 FSD Beta12 版本性能升级主要内容

版本	升级项目	升级内容	适用车型
FSD Beta V12.1 (2023/12/24)	FSD Beta v12.1	将城市街道驾驶的软件栈升级为单一的端到端神经网络，该网络经过数百万个视频片段的训练，替代了超过 30 万行 C++代码	S/3/X/Y
	Over-the-Air (OTA) Recall	1、通过增加文本大小和将通知移动到更突出的位置，提高了触摸屏上驾驶员监控警告警报的可见性	3/Y
		2、添加了通过单次触动操纵杆，而不是两次来激活 Autopilot 功能的选项，以帮助简化 Autopilot 的激活和接管过程	S/3/X/Y
		3、提高了使用自动转向系统以及接近高速公路外交通灯和停车标志时驾驶员注意力的严格要求	S/3/X/Y
		4、引入了一项停用政策，如果检测到驾驶员的不当使用行为，将限制用户使用 Autosteer 功能一周。不当使用行为是指您或您车辆的其他驾驶员收到五次“强制自动驾驶接管”	S/3/X/Y
FSD Beta V12.3 (2024/03/12)	U-Turns	在某些特定的路段可以实现 U 型掉头	S/3/X/Y
	Automatic Speed Offset	能够综合道路类型、交通流量、环境条件、速度限制和FSD 驾驶模式因素之后，允许自动驾驶以最自然的速度行驶	S/3/X/Y
	Removal of TACC with FSD Beta	取消了原先用于激活交通感知巡航控制（TACC）的两次拉动方式，在启用FSD Beta 的同时，无法直接使用 TACC 功能	S/3/X/Y

资料来源：Not a Tesla App，

拥抱“端到端”的自动驾驶，打开智驾算法天花板

“端到端”架构在神经网络化的大模型中整体实现自动驾驶的所有操作，完成从感知结果到控制命令的直接映射，能够提高算法响应速度、减少累计误差，拥有更高的“智驾天花板”。从当前设计思想来看，自动驾驶系统主要分为传统分治法和端到端两大类方式，当前属于从分任务的分治法向“端到端”演进的过程中。

图 3：传统自动驾驶算法与“端到端”算法差异对比

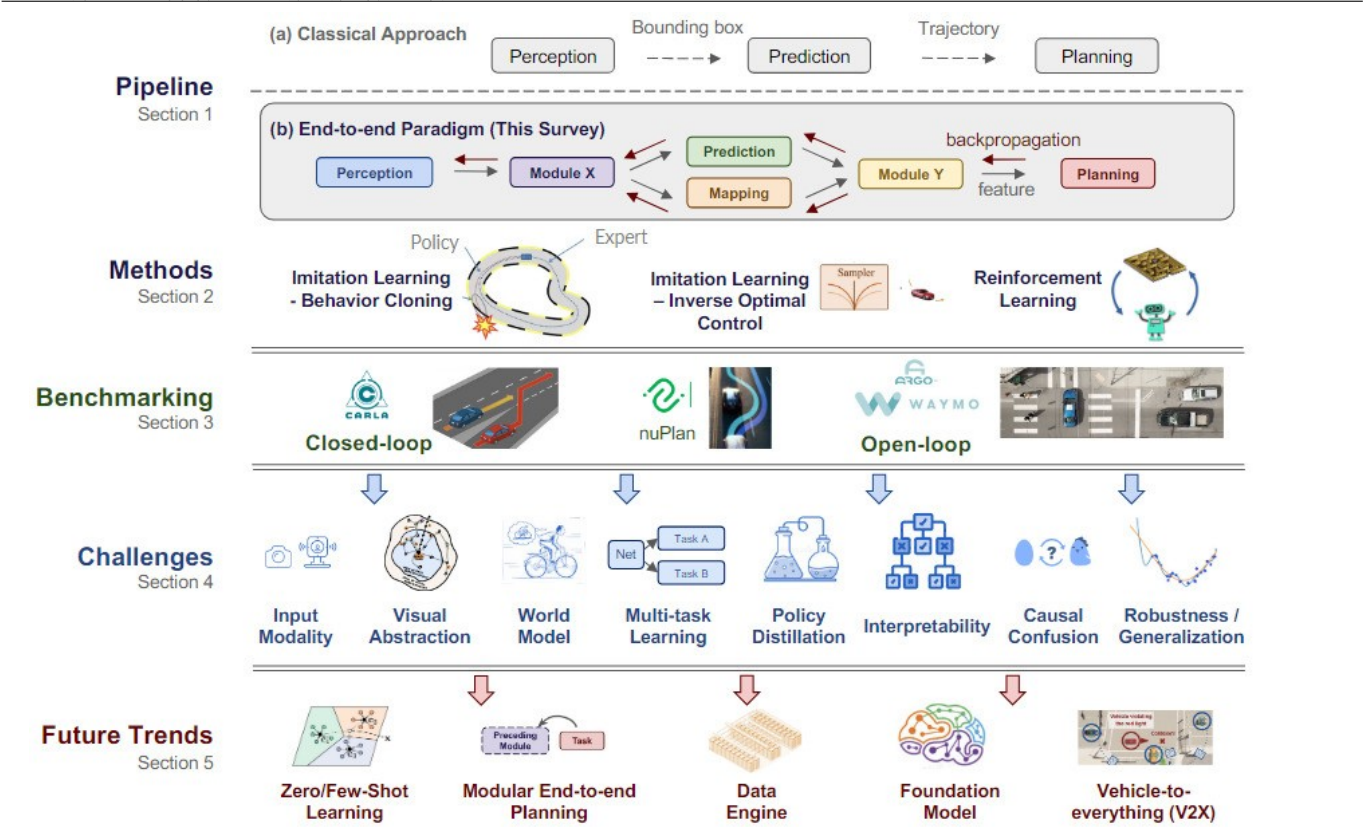


Fig. 1: Our Survey at a glance. I. Pipeline and methods. We define end-to-end autonomous driving as a learning-based algorithm framework with raw sensor input and planning/control output. We classify the 250+ papers surveyed into imitation learning (IL) and reinforcement learning (RL). II. Benchmarking. We group popular benchmarks into closed-loop and open-loop evaluation, respectively. We cover various aspects of closed-loop simulation and the limitations of open-loop evaluation for this problem. III. Challenges. This is the main section of our work. We list key challenges from a wide range of topics and extensively analyze why these concerns are crucial. We also comment on promising resolutions to these challenges. IV. Future trends. We discuss how end-to-end autonomous driving could benefit from the aid of the rapid development of foundation models, visual pre-training, etc. Some photos by courtesy of online resources.

资料来源: Arxiv, ；引用自论文: End-to-end Autonomous Driving: Challenges and Frontiers, L. Chen, P. Wu, K. Chitta, B. Jaeger, A. Geiger and H. Li, arXiv:2306.16927

总体来看，传统自动驾驶算法是多任务的，将每个任务模块的输出直接输入到后续单元中（上图虚线箭头）。而“端到端”算法是单任务的，特征可在不同组件之间传递（上图灰色实线箭头），优化函数被定义为过程全局最优以及通过反向传播方法将错误最小化（上图红色箭头）。

传统分治法是将自动驾驶任务进行切分，定义多个子任务，每个子系统负责解决驾驶过程中的某些特定问题，单独进行开发和训练，再进行系统集成来完成整个自动驾驶任务。

例如，传统智驾方案将自动驾驶分模块为感知、规划、控制，分别负责不同功能，例如**感知模块**负责通过各类传感器（如摄像头、雷达等）收集环境信息，识别道路、车辆、行人、交通标志等元素；**规划模块**基于感知模块提供的信息进行路径规划和决策制定；**控制模块**则依据规划结果执行具体动作，如调整车速、转向等。从目前的行业应用看，**分任务模块化设计更适合团队分工合作，易于引入专家规则，大幅度降低了系统开发难度，更具有工程落地能力，这也是目前多数企业采用的落地方案**。同时可以针对每个模块的输入和输出进行白盒化的分析，系统可解释性更强，对车企而言，一旦发生事故与故障，更易找到原因。但这种方式存在的弊端同样明显，比如模块太多、集成困难、错误累加、响应延迟等，同时由于系统设计中引入较多人为先验信息，导致智驾系统泛化能力和驾驶能力上限较低。

“端到端”大模型打破这种分任务模块化的界限，构建统一的神经网络结构，同时实现多个任务，直接从原始传感数据输入开始，经过深度学习算法处理后，一步到位输出对车辆的操控指令，更加接近最真实的人类驾驶。“端到端”算法优势显著，能够有效减少各模块间的信息传递延迟和误差累计，使得自动驾驶系统的反应更加敏捷、准确，同时也简化了系统结构，更接近于人类真实驾驶过程中的操作方式。但同时，“端到端”的弊端也相对明显：1）**模型可解释性较差**，工作过程是“黑盒”，问题定位相对困难；2）**以数据驱动，对海量高质量数据和算力高度依赖**，大型神经网络训练需要依靠极其庞大的高质量数据支持，同时对于超算中心的算力提出了更高要求。

表 3：自动驾驶解决方案分类及优缺点

模型类型	模块化情况	算法规则	拟人化程度	优点	缺点
传统规则算法驱动的模块化小模型	【感知】、【定位】、【决策】、【规划】、【控制】	算法规则驱动	机械化	数据较易获取,以严谨的算法规则驱动,系统可解释性强	系统复杂,数据计算量大,对硬件和算力要求高,规则构建复杂且难以覆盖全场景
半拟人化部分算法规则驱动	【感知-定位-决策】、【规划】、【控制】	部分算法规则驱动	半拟人	系统复杂度较低,对硬件和算力要求降低,具备部分可解释性	数据获取难度较大,部分规则算法,对应的规则驱动部分仍存在难以覆盖的场景
拟人化无规则端到端AI大模型	【感知-定位-决策-规划-控制】	数据驱动	完全拟人	系统复杂度低且高效,数据获取成本低,不需人为制定规则,模型以数据驱动	系统需根据不同车辆和传感器进行校准,可解释性差

资料来源：虎嗅智库，

“端到端”自动驾驶发展历史可以追溯到 1988 年的 ALVINN，通过摄像头和激光测距仪输入视觉信息，并通过一个简单的神经网络生成转向指令输出。“端到端”自动驾驶算法在 2021 年出现重要转折点，随着车端传感器增加以及算力提升，算法开始引入更高级的架构（比如 Transformer），更好地捕捉全局且更具代表的特征。结合模拟环境，新的架构使得在闭环 CARLA 基准（Closed-loop CARLA Benchmark）下性能大大提升。为了提升自动驾驶系统的可解释性和安全性，将 NEAT、NMP 和 BDD-X 等方法纳入各辅助模块，以更好地监督学习过程。当前，研究方向主要关注数据生成、基础大模型预训练、模块化的端到端规划理念，同时引入新的、功能更强的 CARLA v2 和 nuPlan 基准。

图 4：“端到端”自动驾驶关键成就发展图

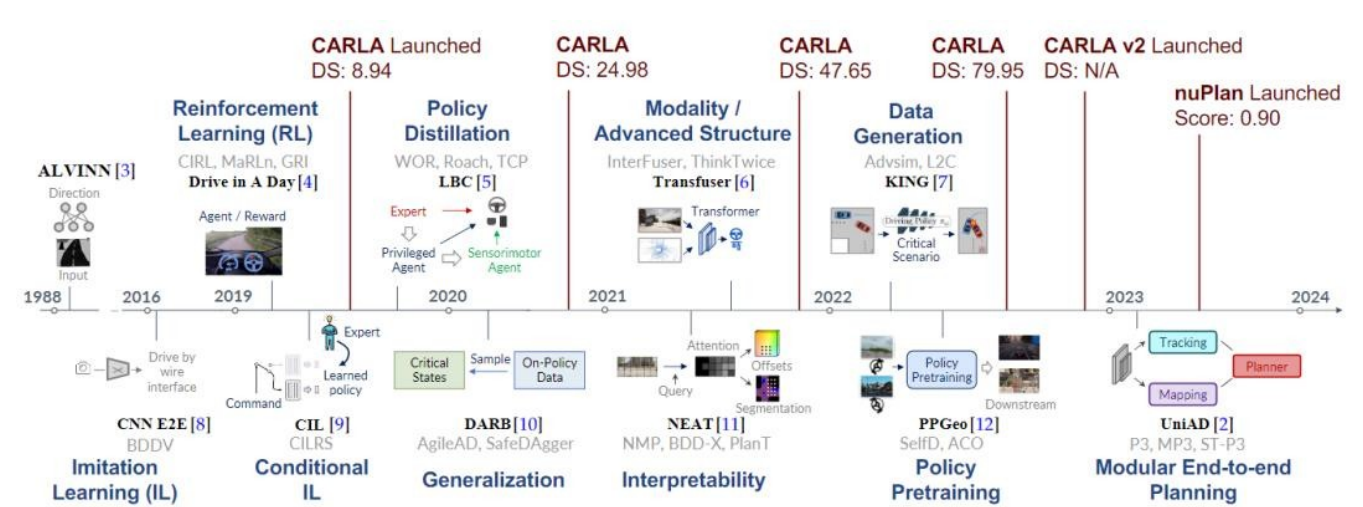


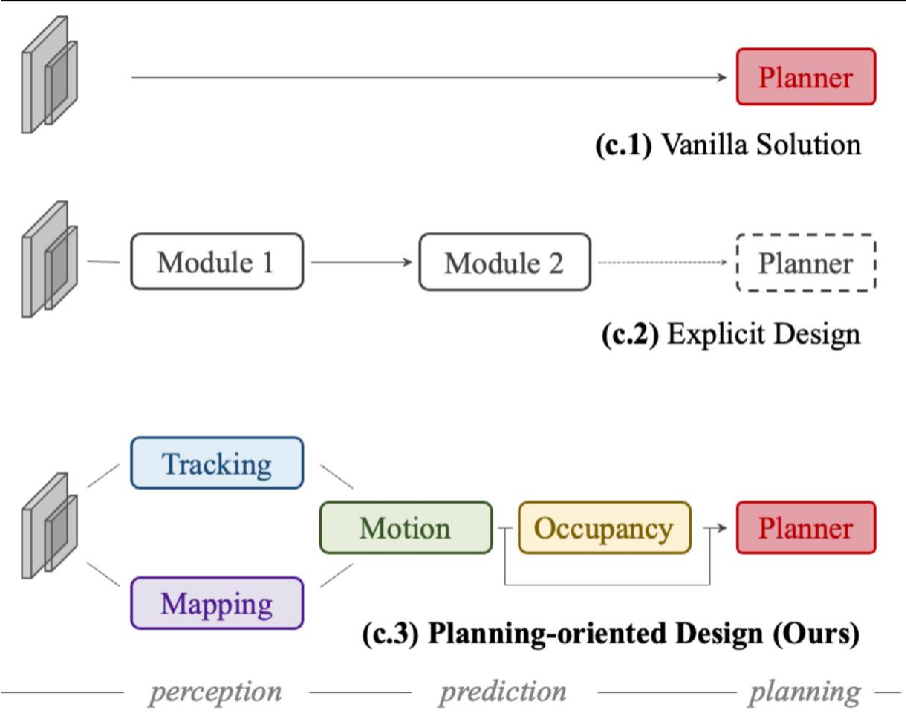
Fig. 2: Roadmap of End-to-end Autonomous Driving. We present the key milestones chronologically, grouping similar works under the same theme. The representative or first work is shown in bold with an illustration, while the date of the rest of the literature in the same theme may vary. We also display the score for each year's top entry in the CARLA leaderboard [13] (DS, ranging from 0 to 100) and the recent nuPlan challenge [14] (Score ranging from 0 to 1).

资料来源: Arxiv, ; 引用自论文: End-to-end Autonomous Driving: Challenges and Frontiers, L. Chen, P. Wu, K. Chitta, B. Jaeger, A. Geiger and H. Li, arXiv:2306.16927

值得注意的是，经常会有一个误区，即认为“传统自动驾驶是多模块的，而端到端自动驾驶一定是单模块的”，这其实把分任务和分模块概念混淆了。传统自动驾驶算法是分任务的，因此必然是多个模块。而“端到端”自动驾驶可以通过单模块来实现，也可以通过多模块来实现，但模型整体优化目标是单任务的，即实现全局最优的表现。其最主要的区别之一体现在模型训练，分任务系统是每个任务独立训练、独立优化、独立测评的，而“端到端”系统是把所有模块看成一个整体进行端到端训练、端到端测评。

从结构来看，“端到端”算法有多种实现路径：1) 简单方式 (Vanilla Solution)，使用单一模块，直接获得规划的最优化；2) 多模块设计 (Explicit Design)，将算法内部划分成不同部分，以实现统一的规划最优化；3) 规划导向的设计 (Planning-oriented Design)，设计一个理想的系统，以路径规划为导向，并正确组织先前的模块以便于进行规划，实现感知决策规划一体化，该方式是在多模块设计 (Explicit Design) 基础上的进一步延伸和深化。

图 5：不同思路的“端到端”算法结构设计思路



资料来源：Arxiv，；引用自论文：Planning-oriented Autonomous Driving, Yihan Hu, Jiazhi Yang, Li Chen, Keyu Li, Chonghao Sima, Xizhou Zhu, Siqi Chai, Senyao Du, Tianwei Lin, Wenhai Wang, Lewei Lu, Xiaosong Jia, Qiang Liu, Jifeng Dai, Yu Qiao, Hongyang Li, arXiv:2212.10156

最新来看，“端到端”模型学术界的代表是由上海人工智能实验室、武汉大学和商汤科技合作研发的自动驾驶通用大模型 **UniAD**，以路径规划为导向，实现感知决策规划一体化。UniAD 是一种分模块端到端训练方式，这种方式通过端到端训练避免了多任务的融合难题，实现全局最优解，又保留了分模块系统的优势，可以抛出中间模块的结果进行白盒化分析，相比单一模块模型，反而更具灵活性，更易部署。

UniAD 模型将检测、跟踪、建图、轨迹预测、占据格栅预测整合到基于 **Transformer** 的“端到端”网络框架下，通过通用 **token** 融合环视图片映射得到 **BEV** 特征，实现目标跟踪、在线建图、目标轨迹预测和障碍物预测。联合优化通过连续时序视频流输入，映射到 **BEV** 空间，通过 **Transformer**、**Mapformer**、**Motionformer** 和 **Occformer** 实现特征融合和多步未来占用预测，最终通过 **Planner** 进行端到端的多模块联合优化和可微分训练。UniAD 通过先训练感知能力，再训练整体模型，实现了自动驾驶领域的创新突破。

图 6: UnitAD 以规划为导向的端到端自动驾驶模型

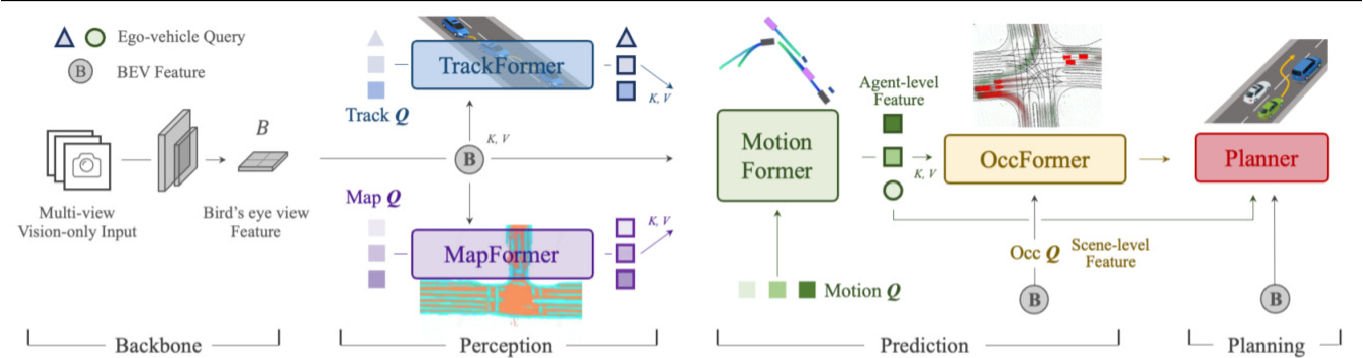


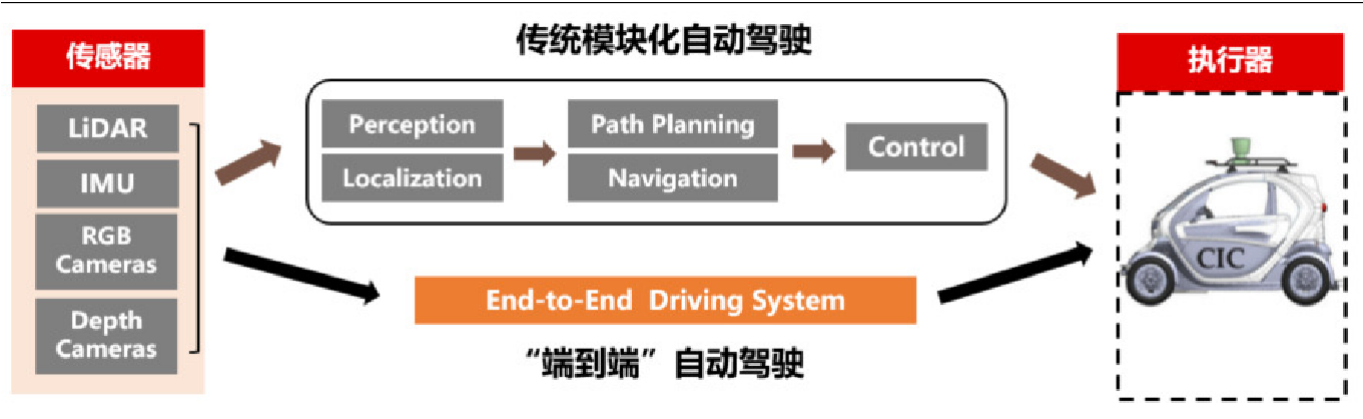
Figure 2. **Pipeline** of Unified Autonomous Driving (UniAD). It is exquisitely devised following planning-oriented philosophy. Instead of a simple stack of tasks, we investigate the effect of each module in perception and prediction, leveraging the benefits of joint optimization from preceding nodes to final planning in the driving scene. All perception and prediction modules are designed in a transformer decoder structure, with task queries as interfaces connecting each node. A simple attention-based planner is in the end to predict future waypoints of the ego-vehicle considering the knowledge extracted from preceding nodes. The map over occupancy is for visual purpose only.

资料来源: Arxiv, ; 引用自论文: Planning-oriented Autonomous Driving, Yihan Hu, Jiazhi Yang, Li Chen, Keyu Li, Chonghao Sima, Xizhou Zhu, Siqi Chai, Senyao Du, Tianwei Lin, Wenhai Wang, Lewei Lu, Xiaosong Jia, Qiang Liu, Jifeng Dai, Yu Qiao, Hongyang Li, arXiv:2212.10156

在工业界，以特斯拉 FSD Beta v12 为代表，率先实现了“端到端”自动驾驶算法的量产落地。特斯拉 FSD Beta v12 的自动驾驶解决方案采用了感知决策一体化模型，将“感知”和“决策”两个模块融合到一个模型中，通过纯视觉感知方案，利用影子模式实时收集数据并训练模型，实现仅依靠车载传感器和神经网络识别道路和交通情况，并作出相应决策。

FSD v12 使用海量数据进行训练，超过 1000 万段分布多样、高质量的视频数据，并采用端到端方式训练。相比 FSD Beta v11 运用超过 30 万行代码来处理各类场景，FSD v12 则仅有 3000 多行 C++ 代码，不再针对任何场景定制策略，完全依靠数据训练神经网络大模型，拥有 50 帧/秒的极高运行效率。

图 7: 特斯拉“端到端”自动驾驶算法框架



资料来源: Tesla Live,

数据端储备优势显著，投产超算 Dojo，训练平台算力持续领先

在传统基于规则算法向数据驱动神经网络大模型转变过程中，对于高质量数据的获取和大模型训练算力的需求呈指数式增长，“智驾基建”成为各车企重点发力的一方面，特斯拉在数据积累和算力建设方面优势显著。

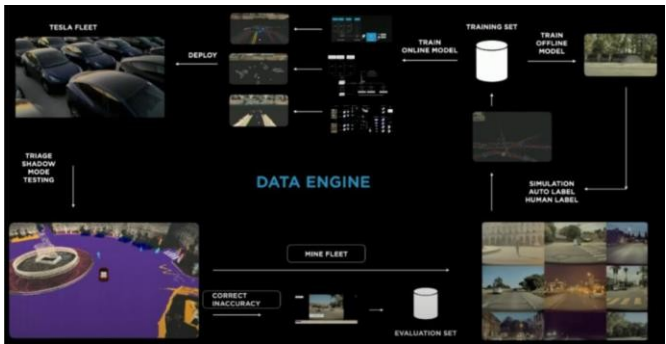
特斯拉影子模式提升数据量，数据端储备优势显著，打造高效数据训练闭环。“端到端”感知模型训练依赖海量数据，特斯拉数据方面优势显著，截至 2023 年底，特斯拉 FSD Beta 累计行驶超过 10 亿英里，积累海量数据。从 FSD 智驾算法数据链来看，主要包括视频数据采集，并将数据通过离线大模型进行自动标注，在云端通过大模型参数训练得到车端精炼的小模型，最后通过 OTA 同步至车端。数据获取主要依赖影子模式和专门数据采集车，采集的 Video 数据会在云端被重建为 4D 空间，并通过大模型进行自动标注，形成带标签的 pair 数据，用于神经网络模型训练。

影子模式提升数据量，影子模式支持量产车上运行并采集数据，等同于将已售车辆作为数据采集车，同时特斯拉销售车辆多，2022 年交付超过 130 万辆，2023 年交付超 180 万辆，充分发挥量产车覆盖场景广、成本低的优势。

图 8：特斯拉 FSD Beta 累计行驶超过 10 亿英里（约 16 亿公里） 图 9：特斯拉 FSD 链路



资料来源：Tesla AI，

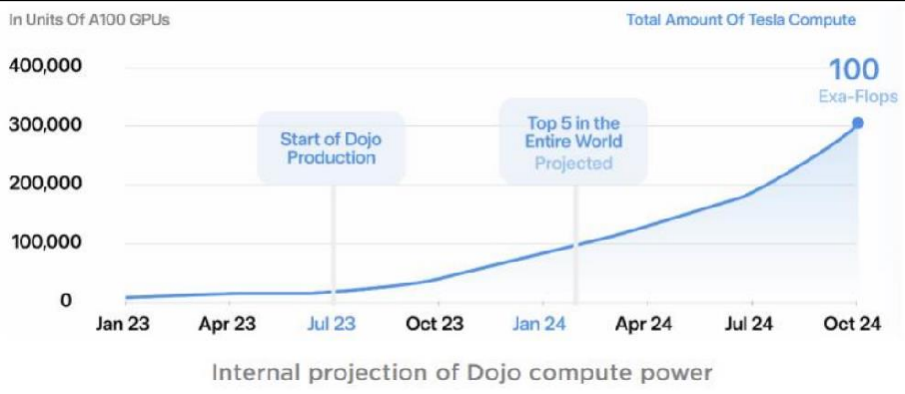


资料来源：Tesla AI Day(2022)，

特斯拉自建 Dojo 超算中心，满足海量视频数据带来的大模型训练算力需求。2021 年特斯拉首次发布 Dojo，2023 年 7 月投产，2024 年 10 月特斯拉的算力总规模将达到 100 Exa-Flops，相当于 30 万块英伟达 A100 显卡的算力总和。

特斯拉拥有海量数据库，数据价值亟待挖掘。Dojo 能够处理海量的视频数据，有望突破瓶颈，一方面加速特斯拉 FSD 系统的迭代，另一方面还能为特斯拉的人形机器人 Optimus 提供算力支持。2024 年 3 月 22 日，马斯克表示特斯拉的人工智能训练将不再受算力限制，表明近期特斯拉计算能力或出现大幅提升。

图 10: 2023 年 7 月投产 Dojo, 特斯拉 2024 年算力预期 100EFLOPS



资料来源：特斯拉官网，

Dojo 采用存算一体架构，单个可扩展计算平面、全局寻址快速存储器和统一的高带宽和低延迟。ExaPOD 架构整体包含 120 个训练模块（Tile），即 3000 个 D1，拥有超过 100 万个训练节点，算力达到 1.1EFLOP；ExaPOD 包含 10 个 Cabinets，1 个 Cabinets 包含 2 个 SystemTray，1 个 SystemTray 包含 6 个 Tile，1 个 Tile 包含 25 个 D1。

图 11: 特斯拉 Dojo 架构

	名称	片上SRAM	算力	说明
	Dojo Core	1.25MB	1.024 TFLOPS	单个计算核心，64bit，4个8x8x4矩阵计算核心，2GHz±频
	Dojo D1	440MB	362 TFLOPS	单芯片，354核心数，654mm^2
	Dojo Tile	11GB	9.05 PFLOPS	单个训练模组，每5x5个芯片组成一个训练模组
	ExaPOD	1320GB	1.1 EFLOPS	训练集群，每12个训练模组组成一个机柜，每10个机柜组成一个ExaPOD，一共3000个D1芯片

资料来源：ADS 智库，

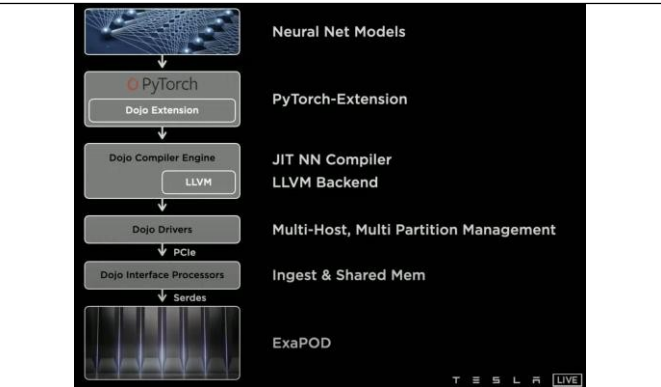
在配套软件方面，特斯拉通过自研与开源相结合的方式为 Dojo 项目打造了专属全栈软件系统，涵盖底层驱动软件、编译器引擎、PyTorch 插件和上层神经网络模型等。

图 12: 特斯拉 D1 芯片与英伟达 A100 芯片性能对比

	特斯拉 D1	英伟达 A100
系统架构	500亿个晶体管，裸晶面积645平方毫米，热设计功耗仅为400W	540亿个晶体管，裸晶面积826平方毫米，采用GPU安培架构，最大功率达到400W，PCIe版250W
AI算力	FP32单精度浮点计算性能达22.6TFlops(每秒22.6万亿次)，BF16/CFP8计算性能则可达362TFlops(每则可达362TFlops(每秒362万亿次);	A100支持TF32运算，浮点性能156TFLOPS，搭载HBM2显存，存储带宽1.6TB/S
核心技术	特斯拉的垂直运算系统设计非常新颖，但芯片制造的核心技术还是台积电掌握	英伟达的张量核心技对AI产生了显著的加速效果

资料来源：佐思汽研，

图 13: Dojo 系统软件栈对接 PyTorch



资料来源：特斯拉官网，

特斯拉 Robotaxi 官宣，无人车商业化加速

酝酿多时，特斯拉Robotaxi 预计 8 月发布

2024 年，4 月 6 日，马斯克在 X 平台宣布，特斯拉将在今年 8 月 8 日发布无人驾驶出租车（Robotaxi）产品，再次引发市场对于 Robotaxi 商业化和 L4 智驾功能的热议，也为自去年10月 Cruise被吊销自动驾驶运营牌照后的Robotaxi行业带来一剂“催化剂”。

特斯拉的 Robotaxi 计划已经酝酿多时。早在 2006 年马斯克刚摸索量产 Roadster 时，就有了 Robotaxi 的想法。2016 年，特斯拉公布“宏图”第二章时提到通过大量汽车行驶数据，开发比人类手动驾驶安全 10 倍的自动驾驶技术（FSD），以及让车辆在闲置的时候，通过分享来为你赚钱（ Robotaxi ）。2019 年特斯拉自动驾驶日发布会上，马斯克表示 Robotaxi 第一阶段将类似网约车、共享汽车的业务模式，特斯拉 Robotaxi 的成本可以低到不到 0.18 美元/英里，而有人驾驶出租车的成本是 2~3 美元/英里。2022 年 Q1 财报会上，马斯克又重申有望在 2024 年量产，并将采用无方向盘或踏板的设计。在去年的高盛科技峰会上，特斯拉投资者关系主管马丁·维查表示，Robotaxi 业务将成为特斯拉第三代平台，除了销售汽车外，特斯拉还会以类似优步模式，从每个订单中抽取相应分成。

图 14：马斯克推特官宣下一代车型 Robotaxi 发布时间

图 15：2019 年特斯拉对于 Robotaxi 的展望



资料来源：推特，

资料来源：Tesla AI Day(2019)，

特斯拉自动驾驶技术全球领先 · FSD 累计超 10 亿英里行驶里程助力其落地 L4 级 Robotaxi 自动驾驶。由于路测里程不足、长尾效应导致的使用和安全性问题、政策因素等原因，一直制约着 Robotaxi 的落地及商业化进程。L4 级自动驾驶的最低路测里程数为 10 亿公里，相当于 100 万辆车以每天 10 小时的频率不间断行驶一年。最新来看，由于车队规模优势赋能 FSD 技术，特斯拉相比其他厂商更接近完全自动驾驶，FSD 累计行驶里程超 10 亿英里（约 16 亿公里）。此外，特斯拉 FSD 升级至 v12 版本，“端到端”的自动驾驶算法上车，智驾水平进一步提升，有望解决更多长尾 corner case 问题，拥有更好地智驾体验。同时美国已在 2023 年开始发放全无人自动驾驶的商业化运营资质，我们认为对于特斯拉发布 Robotaxi 的时机也逐渐成熟。

另外特斯拉入局 Robotaxi，一方面特斯拉在全球范围内拥有庞大的量产车队，Robotaxi 专用车和旗下乘用车运营 Robotaxi 能进一步提升车辆无人驾驶功能使用率，加速规模优势下的自动驾驶数据闭环运行，提升数据采集量，反哺模型训练，加快模型优化。另一方面，入局 Robotaxi 庞大市场，通过创新 Robotaxi 运营商业模式，进一步实现靠软

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/548120070032006075>