



行业研究 | 深度报告 | 软件与服务

# 从特斯拉智能驾驶看人形机器人发展路径

## 报告要点

当前人形机器人产品落地的主要瓶颈在于模型侧。模型侧对机器人多模态对齐数据有较高的需求。数据驱动的大模型需要大量机器人多模态对齐数据投喂，对机器人数据的数量和质量都有较高的需求。机器人厂商解决机器人数据短缺主要靠两种途径：1) 通过为数众多的机器人在物理世界中收集数据；2) 通过 AIGC、数字孪生等方式合成仿真数据。因此建议关注模拟仿真相关标的。

## 分析师及联系人



宗建树

SAC: S0490520030004

SFC: BUX668

## 从特斯拉智能驾驶看人形机器人发展路径

### 机器人软件：越过 L2，直接从具身智能开始

AI 模型的发展已经接近关键节点，多模态模型具备了一定的成熟度，为人形机器人系统的发展铺平了道路。复盘特斯拉智能驾驶的发展路径可以发现，人形机器人模型可以采用自上而下的形式，直接从 L3 级开始构建模型系统。在机器人时代，主流软件厂商都选择了数据驱动的开发范式。数据驱动的机器人模型具备较高的泛化能力，不仅可用于高复杂性场景，还能在环境沟通中自主学习。受限于算力不足以及数据集短缺，模型控制层暂时只能以规则驱动。特斯拉正在以数据驱动快速迭代机器人系统。目前特斯拉已将为数众多的 Optimus 放进汽车生产工厂，安排其协助完成工厂内的电池分拣工作，形成了数据飞轮。

### 机器人硬件：性能优化与云边端协同

机器人从汽车感知系统上继承了摄像头、超声波雷达、激光雷达等传感器，并且提高了传感器数量和性能。此外，为了实现更精准的操控，机器人新增加了一维力和一维力矩传感器、关节电机内部的位置传感器、IMU、用于腕关节和脚部的六维力传感器、触觉传感器(指尖、手掌、高碰撞风险区)等部件。

当前条件下，如何平衡云端的计算延迟与端侧的算力不足是机器人 AI 大脑的核心矛盾之一。放眼未来，机器人大脑在一段时间内或采用云端大脑与端侧模型并行的技术路径。在云端，特斯拉自研 DOJO 超算平台支撑人形机器人的底层算力。DOJO 芯片架构平台由 D1 芯片组成，D1 芯片采用了台积电的 7 纳米制程工艺，计算能力在 BF16/CFP8 上达到 362TFLOPS，FP32 峰值算力为 22.6 TFLOPS，TDP(热设计功耗)为 400W。预计到 2024 年 10 月，Dojo 超算平台的有望匹配 30 万块英伟达 A100 芯片的算力。特斯拉自研的 Dojo 超算平台有望为公司节省超过 65 亿美元的算力投入。

### 商业模式：从制造到运营，机器人有望复刻路径

具身智能将带来产业链的重构，其中软件将在产业链中具备最高价值量。未来随着大模型的渗透率提升，机器人智能化程度有望大幅度提高。参考智能汽车领域的“软件定义汽车”逻辑，软件算法或成为人形机器人厂商的技术护城河，是机器人主机厂实现产品差异化的关键因素。

特斯拉 Robotaxi 将于今年 10 月 10 日发布，而人形机器人有望复刻 RoboTaxi 商业路径。从长期来看，人形机器人有望在近两年快速降本，仿照 RoboTaxi 的形式，在某些高价值场景中的成本低于人力成本，快速达到生产厂商的投资回报曲线门槛。

### 投资建议：关注模拟仿真相关标的

当前人形机器人产品落地的主要瓶颈在于模型侧。模型侧对机器人多模态对齐数据有较高的需求。数据驱动的大模型需要大量机器人多模态对齐数据投喂，对机器人数据的数量和质量都有较高的需求。机器人厂商解决机器人数据短缺主要靠两种途径：1) 通过为数众多的机器人在物理世界中收集数据；2) 通过 AIGC、数字孪生等方式合成仿真数据。因此建议关注模拟仿真相关标的。

### 风险提示

- 1、人形机器人技术发展不及预期；
- 2、人形机器人产品需求不及预期。

### 市场表现对比图(近 12 个月)



资料来源：Wind

### 相关研究

- 《Neuralink 公布脑机接口新进展，产品完成度迅速提升》2024-08-28
- 《Neuralink 公布脑机接口新进展，产品完成度迅速提升》2024-08-28
- 《关键词：泛信创崛起 2024 年第 34 周计算机行业周报》2024-08-26

## 目录

软件：从规则驱动到数据驱动，大模型带来产业链重构 .....	6
智能驾驶软件：先做加法再做减法 .....	7
机器人软件：越过 L2，直接从具身智能开始 .....	12
硬件：性能触及边界之后的再优化 .....	15
智能驾驶硬件：性能与成本之间的平衡 .....	15
机器人硬件：性能优化与云边端协同 .....	19
商业模式：从制造到运营，机器人有望复刻路径 .....	22
软件在产业链中具备最高价值量 .....	23
人形机器人或复刻 RoboTaxi 路径 .....	24
投资建议：关注模拟仿真相关标的 .....	26
风险提示 .....	28

## 图表目录

图 1：智能驾驶技术通常被分为 L0 到 L5 共 6 个级别 .....	6
图 2：L3 级的人形机器人算法转向数据驱动 .....	7
图 3：智能驾驶感知、决策与控制之间的关系 .....	7
图 4：规则驱动下的智能驾驶算法 .....	8
图 5：规则驱动下的算法难以覆盖长尾场景 .....	9
图 6：理想几千人的团队解决 Corner Case .....	9
图 7：高通新一代舱驾融合智驾芯片 .....	9
图 8：特斯拉 FSD V12 实现了端到端的自动驾驶 .....	10
图 9：特斯拉 HydraNet 方案 .....	10
图 10：2021 年特斯拉 HydraNet 智驾系统结构 .....	10
图 11：Occupancy Network 可以识别 3D 空间的体素 .....	11
图 12：2022 年特斯拉 Occupancy Network 智驾系统结构 .....	11
图 13：截至 2024 年 6 月，FSD 里程数据已突破 16 亿英里 .....	12
图 14：AI 模型的演进方向 .....	13
图 15：谷歌 RT-2 模型采用大模型作为底层框架 .....	13
图 16：VLA 是端到端的机器人系统，具备较高的运行效率 .....	14
图 17：机器人模型的两条技术路径 .....	14
图 18：拖动示教快速丰富机器人动作数据 .....	15
图 19：Optimus 机器人在工厂 .....	15
图 20：智能驾驶感知硬件功能俯视图 .....	16
图 21：视觉派和激光雷达派各自代表厂商 .....	17
图 22：特斯拉纯视觉传感器方案 .....	17
图 23：特斯拉智驾系统算力迅速提升 .....	18
图 24：特斯拉 HW3.0 智驾芯片内部构造 .....	18
图 25：Optimus 手指采用触觉传感器 .....	20

图 26: Optimus 脚部采用的力和力矩传感器或为六维力传感器 .....	20
图 27: 云端大脑与端侧模型特点对比 .....	20
图 28: Dojo D1 芯片参数 .....	21
图 29: DOJO D1 性能超越英伟达 A100 .....	21
图 30: Dojo 超算平台为特斯拉带来了什么 .....	21
图 31: 到 2024 年 10 月 Dojo 超算平台算力有望达到 100EFlops .....	21
图 32: 英伟达 NVlink 架构 .....	22
图 33: 特斯拉 Dojo 超算中心结构 .....	22
图 34: 人形机器人产业链图谱 .....	23
图 35: 全球 FSD 搭载比率 .....	23
图 36: FSD 在全球各地的订阅收入 (百万美元) .....	23
图 37: 国内不同城市无人驾驶落地节奏 .....	24
图 38: 特斯拉设想的手机 App 的打车服务界面 .....	25
图 39: 人形机器人有望仿照 RoboTaxi 模式, 快速实现投资回报 .....	26
图 40: 超过 20 个机器人使用 AutoRT 自主收集现实世界数据集 .....	27
图 41: ISSAC 数字孪生系统提升了机器人可用训练数据规模 .....	27
表 1: 不同场景下的部分智驾功能模块 .....	8
表 2: FSD 迭代历程 .....	11
表 3: 特斯拉机器人发展的关键节点 .....	15
表 4: 特斯拉智能驾驶硬件 .....	16
表 5: 不同芯片种类的优缺点 .....	19
表 6: AutoPilot 和 FSD 价格变化 (美国, 美元) .....	24

## 软件：从规则驱动到数据驱动，大模型带来产业链重构

智能驾驶行业与人形机器人行业具备多处共性。智能驾驶的技术进步本质上是通过人工智能高技术供给，驱动交通行业从数字化向智能化转型的过程。在此过程中，车辆与环境互动能力提升，交互门槛降低；与此相对应的，人形机器人也是通过人工智能的高技术供给，驱动制造业/服务业从数字化向智能化转型的过程，在此过程中机器人与环境互动能力提升，交互门槛降低。因此，复盘智能驾驶的发展历程，对研究人形机器人的产业变革具备较强的参考意义。

智能驾驶技术按照能力划分通常被分为 L0 到 L5 共 6 个级别。L0-L2 级智能驾驶又被称为辅助驾驶（ADAS），主要由人类驾驶者负责车辆的操控与决策，并且为最终行驶结果负责，智驾系统只起到辅助作用；而 L3-L5 级的智能驾驶由智驾系统负责车辆的操控与决策，其中 L3 级在特定条件下需要人类接管，L4 级和 L5 级无需人类监督。目前，L0 至 L2 级辅助驾驶系统已广泛普及，而 L3 至 L5 级的智驾也逐渐落地。

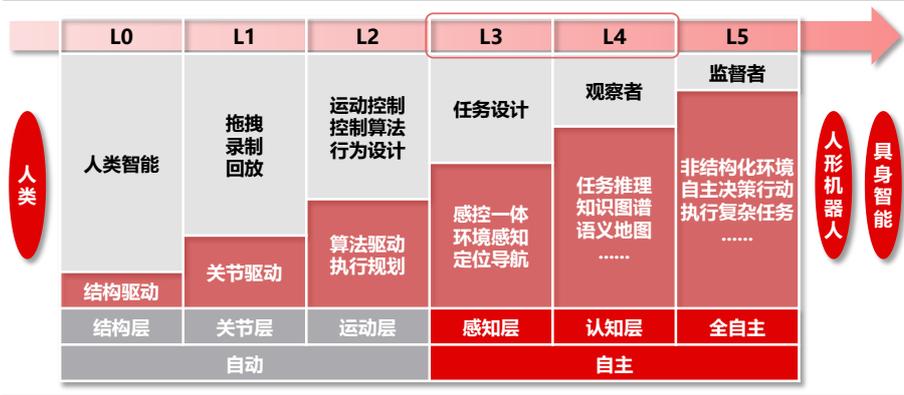
图 1：智能驾驶技术通常被分为 L0 到 L5 共 6 个级别

等级	L0	L1	L2	L3	L4	L5
驾驶员需做到	无论这些驾驶员支持功能在驾驶车辆，即使您的脚			当这些自动驾驶功能启动上，也无须驾驶		
	您必须时刻监督这些辅助功能制动或加速以保证安全			当功能请求时您必须接管	这些自动驾驶功能不会接管车辆	
	半自动驾驶			全自动驾驶		
功能	仅限于提供警告和瞬时协助	为驾驶员提供转向或制动/加速支持	为驾驶员提供转向或制动/加速支持	可以在有限的条件下运行，除非满足所有条件，否则		可在所有条件下驾驶车辆
举例	<ul style="list-style-type: none"> <li>AEB 自动紧急制动</li> <li>盲区警告</li> <li>车道偏离警告</li> </ul>	<ul style="list-style-type: none"> <li>车道保持</li> <li>自适应巡航控制</li> </ul>	<ul style="list-style-type: none"> <li>车道保持</li> <li>自适应巡航控制</li> </ul>	<ul style="list-style-type: none"> <li>交通阻塞自动驾驶模式</li> </ul>	<ul style="list-style-type: none"> <li>局域无人出租车</li> <li>踏板/方向盘可能不被安装</li> </ul>	<ul style="list-style-type: none"> <li>与 L4 级相同，但无条件限制</li> </ul>

资料来源：SAE 官网，长江证券研究所

人形机器人系统也可以分成 L0-L5 六个等级。在软件算法层面，特斯拉机器人同样复刻了智能驾驶算法的开发模式。与智能驾驶类似，端到端的人形机器人系统同样由规则驱动转向数据驱动。传统机器人算法本质上由规则驱动，开发者基于机器人作业规则和需求，参考人类动作的经验，制定预定义的规则和逻辑来进行决策，算法中包含着若干规则算法库；场景越复杂，规则设定也越复杂；但由于现实场景存在较多 Corner Case，所以仅靠规则算法库难以覆盖所有场景，泛化性较低。相比之下，具身智能机器人由数据驱动，开发者先收集大量机器人对齐数据，再使用 AI 算法对数据进行训练和学习，以得出最优的活动策略。

图 2: L3 级的人形机器人算法转向数据驱动



资料来源：甲子光年，长江证券研究所

## 智能驾驶软件：先做加法再做减法

### 从 L0 到 L2 级：规则驱动下系统高度复杂

**L0-L2 智能驾驶软件系统通常分为感知、决策、控制三个模块。**感知模块主要负责处理雷达、激光雷达、摄像头等传感器收集起来的周围环境的数据。收集到传感器信号后，感知算法会将信息融合，创建一个实时的三维环境模型，传输给控制模块；决策模块基于感知模块提供的信息，分析路况环境，识别潜在危险，规划行驶路径，并做出如变道、加速、刹车等驾驶决策。最后，控制模块将决策模块的指令转化为实际的车辆动作，包括调整油门、刹车和转向。

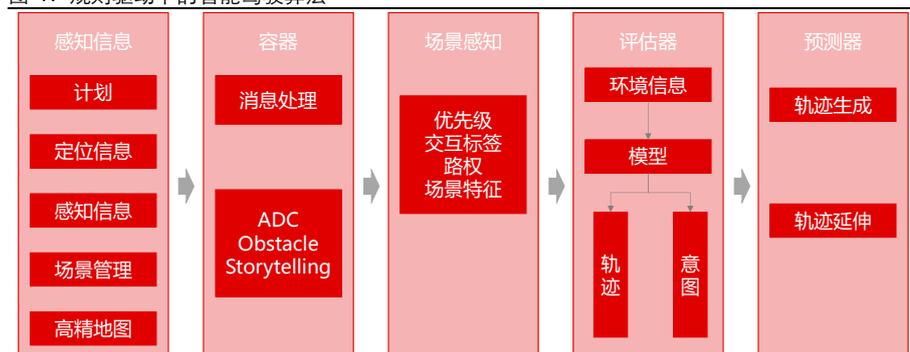
图 3: 智能驾驶感知、决策与控制之间的关系



资料来源：艾瑞咨询，长江证券研究所

从软件层面来看，L0-L2 级的智驾系统主要由规则驱动。系统由多个不同的模块组合而成，不同模块的算法内置了不同场景的应对规则；在进入新场景时，智驾系统需要算法模块之间的衔接。这些模块通过预设的规则集进行交互，确保在不同情境下能够适当地响应，但它们并不具备自主学习或适应新情况的能力，因为决策主要由人类驾驶员监控和干预。

图 4：规则驱动下的智能驾驶算法



资料来源：Apollo 开发者社区，长江证券研究所

从 L0 到 L2 级的软件迭代是功能模块丰富的过程，算法呈现极度复杂化的趋势。由于算法能力不足导致的规则不完善，早期的 L0 级智驾仅能对驾驶员进行路况的提醒；而中后期的 L2 级系统通过补全算法规则，已经能实现自动泊车、交通拥堵辅助等特定场景的接管。在此期间，智驾算法为了尽可能应对所有场景，呈现极度复杂化的趋势。

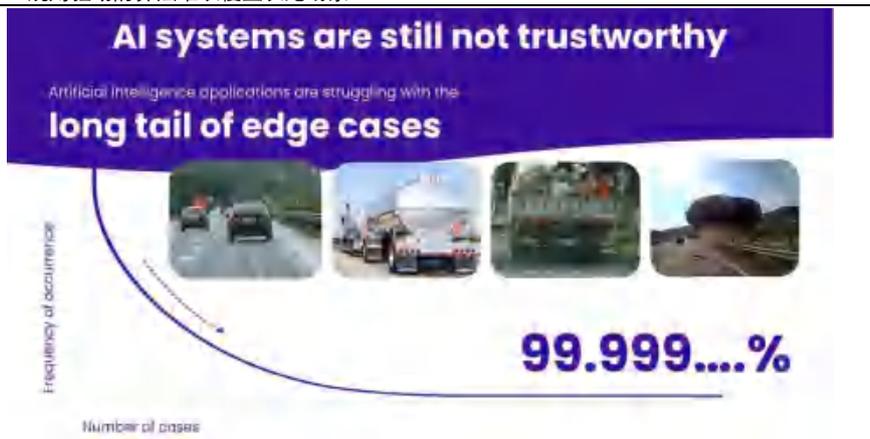
表 1：不同场景下的部分智驾功能模块

场景	功能	功能缩写	功能介绍	智驾等级
智能行车功能	自适应巡航	ACC	纵向自动控制，实现加减速	L1
	车道居中控制	LCC	横向自动控制，保持居中行驶	L1
	自动变道辅助	ALC	横纵向自动控制，实现自动变道	L1
	交通拥堵辅助	TJA	横纵向自动控制，交通拥堵时辅助驾驶	L2
智能泊车功能	自动泊车	APA	自动泊入、泊出选定的车位	L2
	遥控泊车	RPA	通过遥控装置，控制车辆自动泊车	L2
主动安全功能	前向碰撞预警	FCW	前方有碰撞风险时，发出预警	L0
	自动紧急制动	AEB	前方有碰撞风险时，自动制动	L1
	前方交叉区域辅助	FCTB	前方交叉区域有碰撞风险，自动制动	L1
	车道偏离预警	LDW	车辆偏离车道时，发出预警	L0
	车道保持辅助	LKA	车辆偏离车道时，自动横向控制保持	L1

资料来源：芝能智驾，长江证券研究所

规则驱动的辅助驾驶算法存在较多 **Corner Case**。规则驱动的智能驾驶算法依赖于预定义的规则和情景，这种方法的优点是相对直观和易于理解，因为规则是基于人类的直观判断和现有的交通规则。然而，它也有局限性，因为现实世界的道路状况复杂多变，很难为所有可能的情况预先编写规则。所以，当遇到罕见或复杂的 **corner cases**，如极端天气、突然的路面障碍或未曾预见的驾驶行为，算法可能无法做出恰当的决策，这也是辅助驾驶需要人类接管的原因。

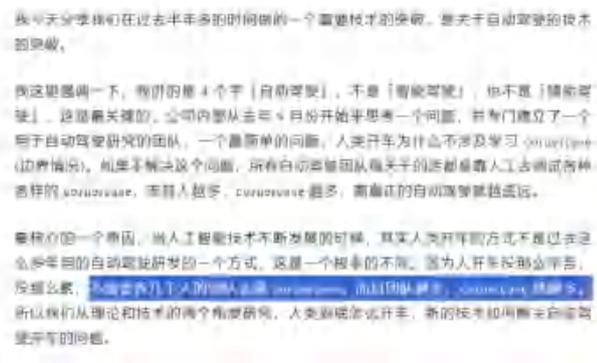
图 5：规则驱动算法难以覆盖长尾场景



资料来源：medium，长江证券研究所

**规则驱动的开发范式带来边际成本递增。**在 L2 级的后期，规则驱动的算法开发范式只能通过查漏补缺的方式解决 Corner Case；而高度复杂化的算法规则也带来了成本端的快速上涨，为了提升 L2 级智驾在中低端车型的渗透率，智驾系统出现了舱驾融合等新技术路径。

图 6：理想几千人的团队解决 Corner Case



资料来源：极客公园，长江证券研究所

图 7：高通新一代舱驾融合智驾芯片

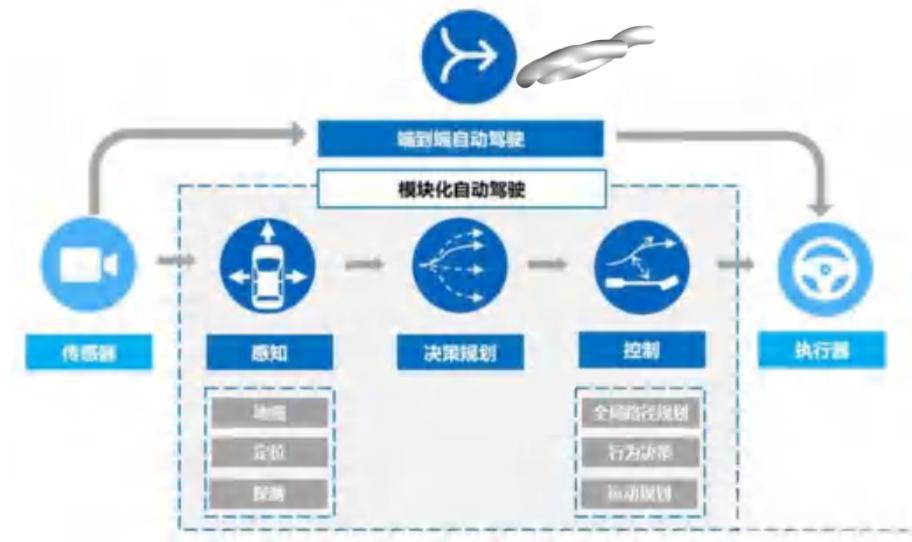


资料来源：高通官网，长江证券研究所

## 从 L3 到 L5 级：数据驱动的系统化繁为简

**数据驱动实现了系统化繁为简。**自 L3 级智驾开始，特斯拉 FSD V12 实现了从规则驱动逻辑到端到端神经网络架构的重大转变。在 L3 级智能驾驶出现之前，特斯拉依赖 30 多万行 C 语言代码来为 FSD 划定“围栏”，根据环境情况翻阅规则库给出驾驶指示；而 V12 版则给系统投喂海量人类司机在海量的场景中的驾驶行为数据来训练自动驾驶神经网络大模型，系统本身精简到只有 3000 行代码；车辆在行驶时，FSD V12 会让神经网络 AI 根据实时环境自主决定驾驶方式，彻底重塑了智能驾驶系统的逻辑。

图 8：特斯拉 FSD V12 实现了端到端的自动驾驶

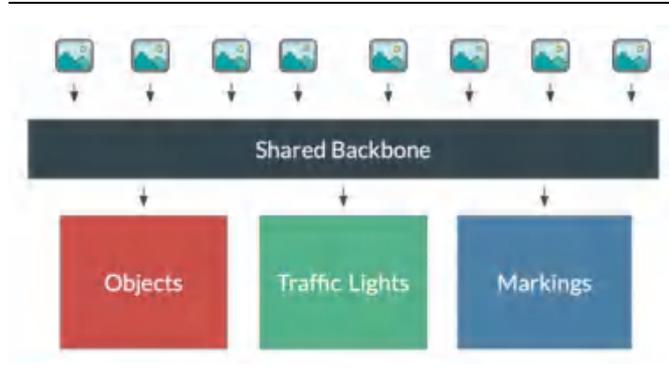


资料来源：智驾最前沿，长江证券研究所

**数据驱动的系统泛化能力与执行效率更高。**基于雷达控制的智能驾驶系统是代码驱动，可能在绝对安全性上的保障要更好，但海量的场景越积越多，算法就会变得越来越庞大和臃肿，导致自动驾驶决策越来越慢，决策效率下降；而基于神经网络大模型是数据驱动，随着算力和训练数据的飞速提升，大模型的能力和覆盖场景也能飞速进化，且执行端代码非常简洁，执行效率越来越高，执行结果比肩人类司机。

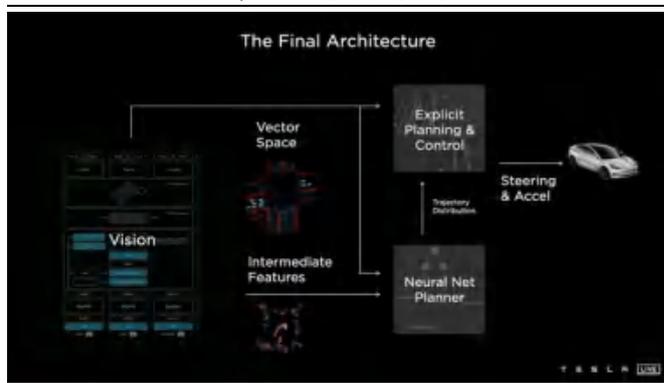
从 2021 年开始，特斯拉智能驾驶系统经历了 HydraNets、Occupancy Network 和端到端系统三次迭代。在早期阶段，特斯拉曾移植 MobileEye 的算法来构建智驾系统；后来随着特斯拉搭建自身的智能驾驶系统，在感知层采用了 Backbone（骨干网）架构；在 2021 年 AI Day，特斯拉发布了 HydraNet 神经网络；HydraNets 神经网络实现了 20 多个感知神经网络的统一，能够完成多头任务，减少重复的卷积计算，减少主干网络计算数量，可以同时处理多个摄像头的的数据。HydraNet 设计用于同时处理多种感知任务，能够在统一的框架内学习多个任务，优化共享特征的提取，提高了处理效率和性能。

图 9：特斯拉 HydraNet 方案



资料来源：thinkautonomous，长江证券研究所

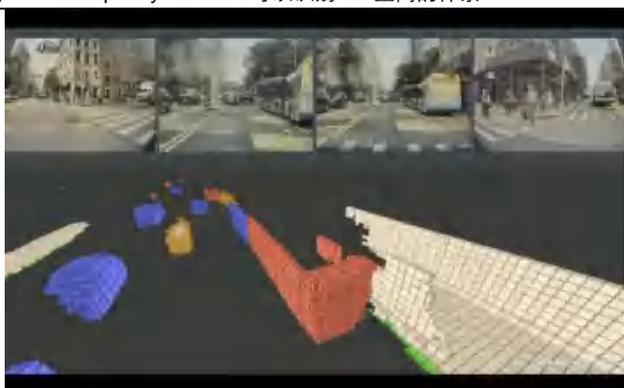
图 10：2021 年特斯拉 HydraNet 智驾系统结构



资料来源：特斯拉 2021 AI Day，长江证券研究所

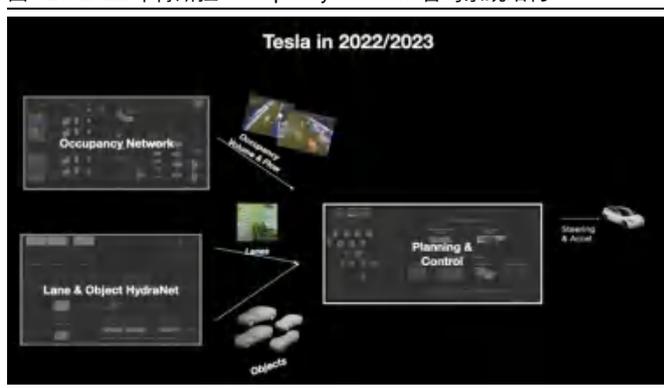
**2022 年，特斯拉发布了 Occupancy Network。**Occupancy Network 是在 HydraNet 系统上的一次升级，把感知层从一个系统切割成 HydraNet 和 Occupancy Network 两部分；HydraNet 负责识别不同类型的物体、标志等信号，而 Occupancy Networks 可以直接感知 3D 空间中的体素 (voxel)，将世界划分为多个大小一致的立方体，快速识别每个体素是否被占用，继而判断车辆是否要躲避。

图 11: Occupancy Network 可以识别 3D 空间的体素



资料来源：特斯拉 2022 AI Day，长江证券研究所

图 12: 2022 年特斯拉 Occupancy Network 智驾系统结构



资料来源：thinkautonomous，长江证券研究所

**2023 年 11 月，特斯拉推出了 FSD V12，实现了端到端的自动驾驶。**端到端的系统将感知、决策、控制整合为一个统一的神经网络模型。在之前的模块化算法中，车辆控制决策和路线规划算法是通过逻辑控制程序来实现的，而端到端方案则把这部分逻辑控制代码负责的功能也交由神经网络模型来实现，取消了传统的规则或者条件判断代码，也无需地图数据。

表 2: FSD 迭代历程

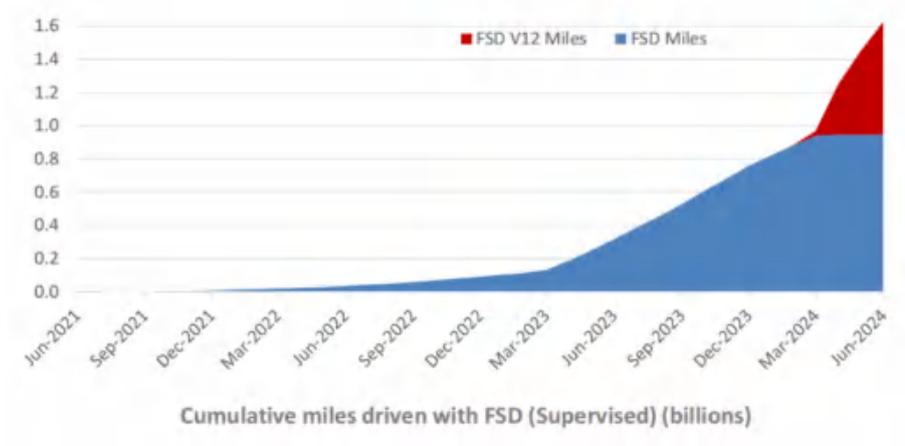
发布时间	FSD 版本	主要效果和影响
2021-07-10	FSD Beta V9	仅使用纯视觉感知实现城市自动辅助驾驶；约 2000 名客户内测，强调用户专注于驾驶
2021-09-11	FSD Beta V10	通过 7 天安全考核的车主中筛选
2021-10-11	FSD Beta V10.2	首个向公众推送的 FSD 版本，安全分满分可订阅
2022-08-21	FSD Beta V10.69	改进原先的视觉系统，增加了占用网络，能够对低速物体进行测量
2022-11-12	FSD Beta V11	高速公路启用 FSD Beta，统一高速公路内外的视觉和规划系统
2022-11-18	FSD Beta V10.69.3.1	取消 FSD Beta 安全分限制，推送向北美所有订阅用户
2023-02-14	FSD Beta V11.3	强化感知规控能力，利用基于神经网络的自车轨迹估计，代替近似的运动学模型，合并高速和城区 FSD
2023-04-17	FSD Beta V11.4	增强了行人识别和响应能力，改进车道引导，将减少 64% 的驾驶员干预
2023-05-10	FSD Beta V11.4.1	在占用网络中添加车道引导输入，减少 43% 由其他车辆切入车道引起的干预
2023-06-19	FSD Beta V11.4.4	提高间隙选择的自信，更准确地预测弱势道路使用者的未来意图
2023-11-24	FSD Beta V12	实现全新的“端到端自动驾驶”
2024-03-12	FSD Beta V12.3	大规模向用户推送
2024-03-30	FSD Supervised V12.3.3	更名为 Supervised，开启所有用户为期 1 个月免费试用
2024-06-06	FSD Supervised V12.4.1	放弃方向盘扭力监测系统，改用车内摄像头来监测驾驶员状态（基于神经网络）

2024-07-23	FSD Supervised V12.5	可处理更复杂的场景，包括在狭窄的单向道路上行驶，以及处理意外的道路封闭
------------	----------------------	-------------------------------------

资料来源：notateslaapp, 长江证券研究所

**数据飞轮使得系统迭代速度大幅度提升。**2024 年 1 月起，特斯拉向特定用户开放了 FSD V12 的使用权限，并随着版本迭代陆续扩大用户触及面；与此同时，FSD 收集的驾驶里程数据也在指数级增长，截至 2024 年 7 月已超过 13 亿英里。特斯拉 FSD 系统已形成正向的数据飞轮，而数据飞轮加速了 FSD 系统的迭代速度。

图 13：截至 2024 年 6 月，FSD 里程数据已突破 16 亿英里



资料来源：特斯拉财报 2024Q2, 长江证券研究所

## 机器人软件：越过 L2，直接从具身智能开始

**多模态模型为具身智能铺平道路。**在智能驾驶技术刚起步的 2010 年代，人工智能的技术发展还不完善，只能凭借堆叠深度学习模型（CNN、RNN）的形式，将成百上千算法拼接起来，自下而上地塑造智能驾驶模型；而如今 AI 模型的发展已经接近关键节点，多模态模型具备了一定的成熟度，为人形机器人系统的发展铺平了道路。因此，人形机器人模型可以采用自上而下的形式，直接从 L3 级开始构建模型系统。

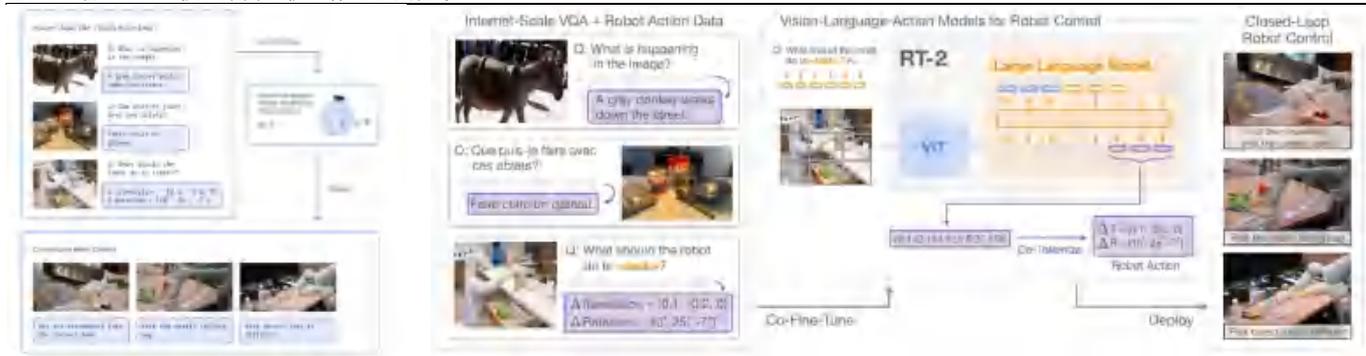
图 14: AI 模型的演进方向



资料来源: 甲子光年, 长江证券研究所

**数据驱动是机器人系统的最佳开发范式。**由智能驾驶的发展历程可以得知, 如果以规则驱动的模式来开发算法, 在功能丰富的过程中, 算法必定极度复杂化, 后期会存在大量 corner case 无法解决; 此外, 由于机器人的能力门槛比智能驾驶更高, 算法的开发难度与 corner case 也会更高。因此在机器人时代, 主流软件厂商都选择了数据驱动的开发范式。

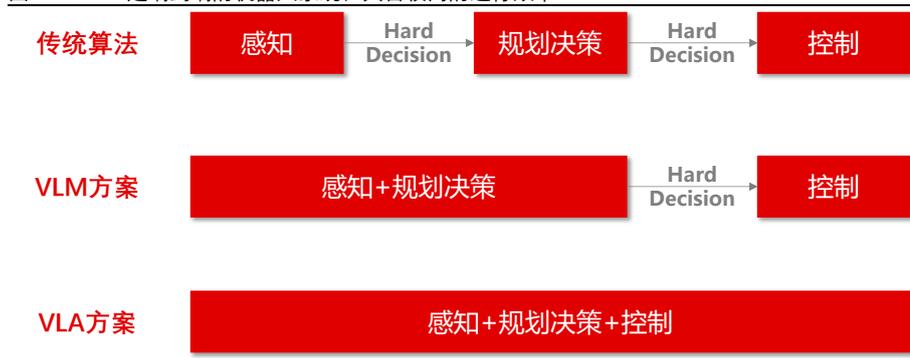
图 15: 谷歌 RT-2 模型采用大模型作为底层框架



资料来源: 谷歌 RT-2 官网, 《RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control》, Brohan et al., 长江证券研究所

**数据驱动的机器人模型具备较高的泛化能力。**目前规则驱动的只能根据算法规则做出相应的判断, 无法面对复杂性较高的场景, 泛化能力较低。相比之下, 数据驱动的机器人借力具身智能大脑, 不仅可用于高复杂性场景, 还能在环境沟通中自主学习。目前大语言模型 (LLM) 已有较高的成熟度, 图像-语言多模态模型 (VLM) 也在快速迭代的过程中; 而随着图像-语言-动作多模态模型 (VLA) 的推出, 大模型实现了数据与处理任务的跃升。从 LLM 到 VLM 再到 VLA, 大模型的数据模态逐渐丰富, 数据规模的数量级迅速增长, 大模型的应用场景和价值量也成比例扩张。

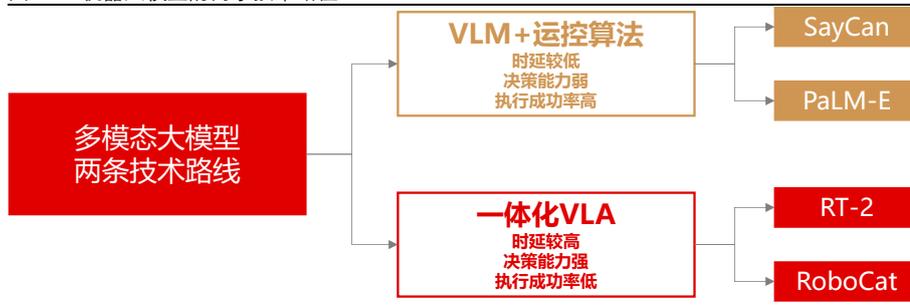
图 16: VLA 是端到端的机器人系统，具备较高的运行效率



资料来源：长江证券研究所

受限于算力不足以及数据集短缺，模型控制层暂时只能以规则驱动。目前主流机器人大多模型偏向于任务理解和拆分，对于机器人运动控制的涉及较少，只是用端到端的训练方式生成了简单且离散分布的机械臂末端位置和底盘移动指令，未考虑连续路径和轨迹规划等更偏机器人领域的内容。以 RT-2 为例，RT-2-PaLI-X-55B 只能做到 1-3Hz 的指令生成速度，在移动马克笔的任务中，因为其运控速度远低于马克笔的滚动速度而失败；并且其生成的运动指令只是机械臂的末端位置和姿势，未涉及传统机器人运控范畴。因此，目前机器人大多模型更擅长任务级和技能级的控制，传统机器人算法擅长底层运控。以 Figure 01 为例，Figure 01 机器人采用了 VLM+控制小模型方案，VLM 部分由 OpenAI 赋能，频率达到了 200Hz，下游连接的控制小模型负责机器人的运控，指令生成速度达到 1kHz，可以较好应用于现实场景。

图 17: 机器人模型的两条技术路径



资料来源：高工移动机器人，长江证券研究所

**Optimus 机器人发展包含 6 个重要时间节点：** 1) 2021 年 8 月马斯克首次透露了要造人形机器人 (Tesla Bot) 的设想；2) 2022 年 10 月 Optimus 人形机器人首次亮相。在 Demo 演示中，Optimus 展示了其搬箱子、浇花、在特斯拉工厂工作的画面；3) 2023 年 5 月，Optimus 具备了流畅行走和抓取物体的能力；4) 2023 年 9 月，Optimus 大脑快速进化，实现神经网络完全端到端训练，可以自主对物体进行分类；5) 2023 年 12 月，特斯拉发布 Optimus Gen-2，在稳定性、灵活性、机动性方面有较大提升。6) 2024 年 1 月 16 日，Optimus 可以实现叠衣服功能，动作控制更加精准。

表 3: 特斯拉机器人发展的关键节点

时间	事件	内涵
2021.08	Tesla 宣布要造人形机器人	特斯拉敏锐抓住人工智能的关键节点，尝试把智能驾驶汽车的成功经验复制到机器人领域
2022.10	Tesla 在 2022 AI Day 上发布人形机器人 BumbleBee	确定了机器人雏形与大致能力范围，公布了人形机器人的身高、体重、行走速度等关键参数细节
2023.05	Tesla 公布了机器人在工厂中行走，以及人类动作数据示教的视频	Optimus Gen-1 机器人的运控能力与感知能力得到大幅度提升
2023.09	机器人实现神经网络完全端到端训练，可以自主对物体进行分类	机器人泛化能力得到提升
2023.12	Tesla 发布 Optimus Gen-2	机器人稳定性、灵活性、机动性方面有较大提升，并在手指搭载触觉传感器
2024.01	机器人可以实现叠衣服功能，动作控制更加精准	思维链能力提升，可以将复杂任务拆解为简单步骤；控制能力同样得到了提升

资料来源：特斯拉机器人 X 账号，长江证券研究所

**特斯拉以数据驱动快速迭代机器人系统。**2023 年 9 月起，Optimus 机器人即装载了端到端的神经网络。但早期的系统成熟度不高，在特斯拉公布的视频中可以看到机器人有机动力不足、控制不够精准、感知反馈迟缓等问题；但随着拖动示教等项目推进，Optimus 得到了大量动作数据的投喂，性能得到了快速提升。目前特斯拉已将为数众多的 Optimus 放进汽车生产工厂，安排其协助完成工厂内的电池分拣工作，形成了数据飞轮。

图 18: 拖动示教快速丰富机器人动作数据



资料来源：特斯拉 2022 AI Day，长江证券研究所

图 19: Optimus 机器人在工厂



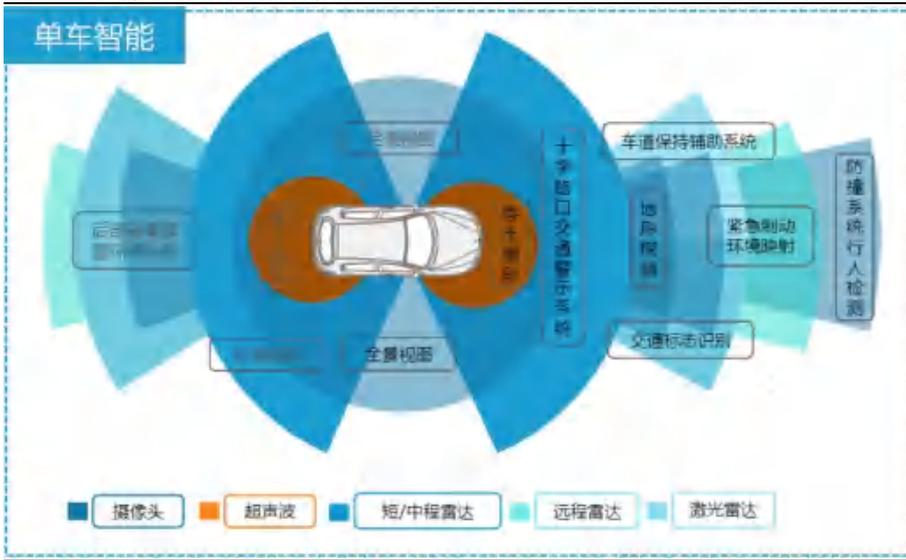
资料来源：Tesla X 账号，长江证券研究所

## 硬件：性能触及边界之后的再优化

### 智能驾驶硬件：性能与成本之间的平衡

**硬件性能的提升给智能驾驶整体发展提供了有力的基石。**智能驾驶硬件主要可以分为摄像头、超声波雷达、毫米波雷达等传感器以及计算芯片。传感器负责收集周围环境的数据，而计算芯片则处理这些数据以实现决策和控制。近年来，传感器的数量、种类和性能都得到了明显增强，提升了车辆对周边环境的感知能力；与此同时，计算芯片对更大规模的环境数据进行处理，确保了更高的决策准确性。

图 20：智能驾驶感知硬件功能俯视图



资料来源：亿欧智库，长江证券研究所

### 传感器：多传感器融合与纯视觉的技术路线之争

从 L0 到 L2 级，智能驾驶硬件同样逐渐复杂化。以特斯拉为例，特斯拉智驾硬件经历了从 HW1.0 到 HW4.0 的多轮迭代。HW1.0 于 2014 年发布，最早被搭载于部分 Model S 车型上，并逐步应用到 Model X 车型中。HW1.0 时期的传感器部分仅搭载了 1 颗前置摄像头，1 个毫米波雷达与 12 个超声波雷达；此外，由于当时的特斯拉缺少芯片自研能力，智驾芯片采用的是 MobileEye 的 Q3 芯片，算力仅为 0.256Tops@int8。相比之下，特斯拉在 2019 年实现了 L2.9 级的智能驾驶，此时的硬件无论是数量还是性能都得到提升。总计 8 个摄像头覆盖了长焦、中焦与广角三个层面，毫米波雷达与超声波雷达的性能也得到了提升，搭载的智驾芯片由两个大 FSD 芯片和两个小 NPU 组成，总算力达到了 144TOPS@int8。

表 4：特斯拉智能驾驶硬件

项目	HW1.0	HW2.0	HW2.5	HW3.0	HW4.0
发布时间	2014 年 9 月	2016 年 10 月	2017 年 8 月	2019 年 3 月	2023 年 3 月
毫米波雷达		160m	170m	170m	300m
超声波雷达	Lidar*12(5m)		Lidar*12(8m)	Lidar*12(8m)	-
前置摄像头	1 颗		3 颗： 长焦 (35°): 250 m 中焦 (50°): 150 m 广角(120°): 60 m		升级为 12 摄像头(1 个冗余，可用 11 个)；
侧前摄像头	-		两颗摄像头 (90°): 80 m		摄像头像素从原来的 120 万提升到了 500 万，探测距离至少 424 米
侧后摄像头	-		两颗摄像头(60°): 100 m		
后摄像头	-		1 颗摄像头(150°)		
芯片/核心处理器	Mobileye EyeQ3*1	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	FSD 1 芯片 (12 核心) *2	FSD 2 芯片 (20 核心) *2
内存	256 MB	6 GB	8GB	8GB *2	8GB *2
FPS	36	110	110	2300	-

算力 (TOPS)	0.256	12	12	144	-
-----------	-------	----	----	-----	---

资料来源: teslatap, 评智, 智能车情报局, 长江证券研究所

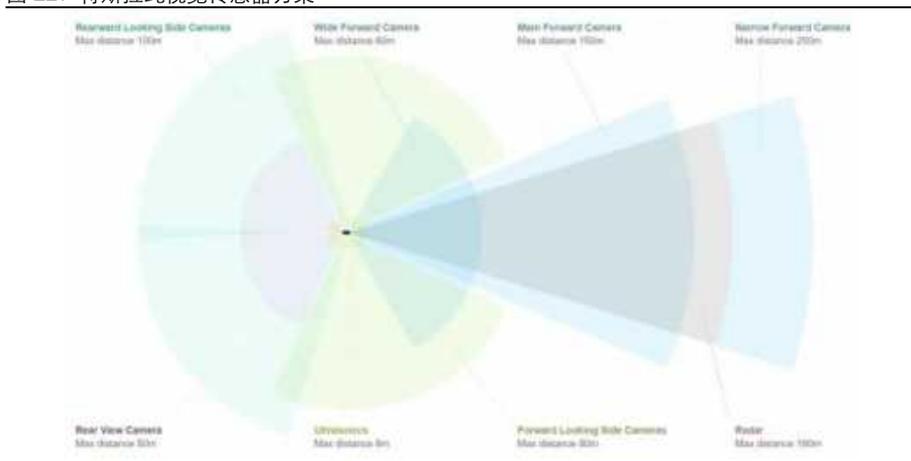
数据驱动的系统降低了对雷达等传感器的需求。在配备了激光雷达、毫米波雷达的驾驶系统中，雷达传感器的优先级必须高于摄像头，才能实现雷达对安全性的兜底作用。这种系统本身就是一种规则设定，而不是经过神经网络进行决策。相比之下，由数据驱动的智能驾驶系统模拟人类的驾驶逻辑，舍弃了针对传感器和图像数据的控制算法和规则判决，对雷达等传感器的需求较低。因此，**特斯拉在 HW3.0 后期舍弃了全部车载雷达，转向纯技术方案。**

图 21: 视觉派和激光雷达派各自代表厂商



资料来源: 量子位, 长江证券研究所

图 22: 特斯拉纯视觉传感器方案

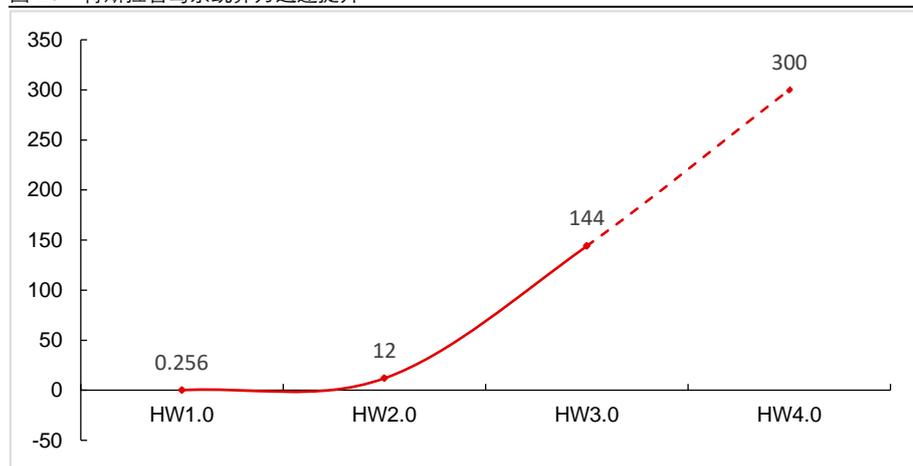


资料来源: 高工智能汽车, 长江证券研究所

## 智驾芯片：性能提升与异构化

智驾系统性能提升也带动了芯片的算力增长。早期特斯拉智驾系统移植了是 MobileEye 的 Q3 芯片，算力仅为 0.256TOPS@int8；随着算法模块数量增长，芯片算力也随之提升。截至 2024 年，新推出的 HW4.0 芯片算力有望在 HW3.0 的基础上更进一步，突破 300TOPS@int8。

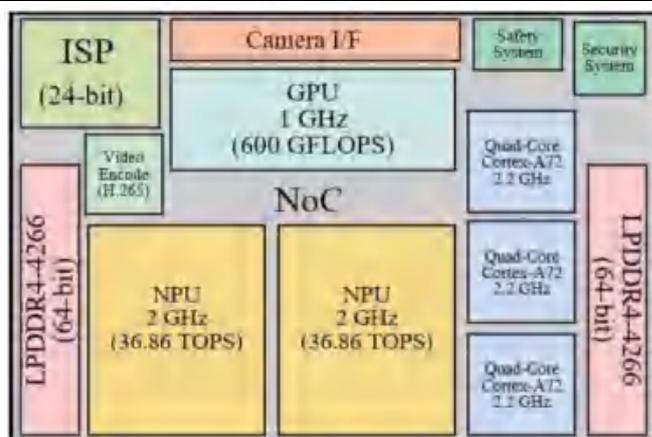
图 23：特斯拉智驾系统算力迅速提升



资料来源：智能车情报局，长江证券研究所

智驾软硬件性能的提升除了推动了智驾芯片的算力提升，也促使智驾芯片正朝着专用性强，特定领域效率高的方向发展。智驾芯片逐渐复杂化，从早期的 CPU、GPU 发展成 CPU、GPU、FPGA、ASIC 相结合的模式。其中 CPU、GPU、FPGA 均是前期较为成熟的芯片架构，属于通用型芯片，而 ASIC 属于为 AI 特定场景定制芯片。目前通用的 CPU、GPU 等都能执行 AI 算法，但执行效率差异较大，而 FPGA 与 ASIC 每个逻辑单元的功能在重编程时就已经确定，不需要指令，因此计算效率较高。在智能驾驶领域，环境感知、物体识别等深度学习需要芯片保持高响应速度，且在保证高性能的同时低功耗，不能对智能汽车的续航里程造成较大影响，所以 CPU、GPU 的组合不适合智能驾驶场景，开发 FPGA、ASIC 成了必然选择。

图 24：特斯拉 HW3.0 智驾芯片内部构造



资料来源：wikichip，长江证券研究所

表 5: 不同芯片种类的优缺点

芯片类型	基本情况
CPU	<p><b>中央处理器, Central Processing Unit。</b>是运算核心 (Core) 和控制核心 (Control Unit)。它的功能主要是解释计算机指令以及处理计算机软件中的数据</p> <p><b>优点:</b> 通用芯片, 有强大的调度、管理、协调能力, 应用范围广灵活性最高, 擅长逻辑控制、串行的运算</p> <p><b>缺点:</b> 运算量低, 性能和功耗效率相对较低</p>
GPU	<p><b>图形处理器, Graphics Processing Unit。</b>又称显示核心、视觉处理器、显示芯片</p> <p><b>优点:</b> 最初专用于图形处理制作, 后逐渐应用于计算, 适合执行复杂的数学和几何计算, 尤其是并行计算。通用性较高, 相比 CPU, 性能好。</p> <p><b>缺点:</b> 功耗效率依然不高</p>
ASIC	<p><b>专用集成电路, Application Specific Integrated Circuit,</b> 为某种单一用途专门制作的电路</p> <p><b>优点:</b> 计算能力和效率都需要根据特定算法进行定制, 具备极高的性能和功耗效率。定制芯片成本最低, 适合量产</p> <p><b>缺点:</b> 研发成本高, 制作周期长、可复制性差、灵活性不足</p>
FPGA	<p><b>现场可编程门阵列, Field Programmable Gate Array,</b> 半定制芯片。</p> <p><b>优点:</b> 可针对具体应用对算法结构进行深度定制, 达到更高的性能和功耗效率, 可以无限次编程; FPGA 相比 GPU 具有稳定和延迟低的优势, 适用于流式的计算密集型任务和通信密集型任务</p> <p><b>缺点:</b> 基本单元算力有限, 开发难度大、只适合定点运算、价格比较昂贵</p>

资料来源: CSDN、艾瑞咨询, 长江证券研究所

## 机器人硬件: 性能优化与云边端协同

### 传感器: 场景拓宽与性能优化

机器人继承了智能驾驶的传感器系统, 并在此基础上根据场景增加了新传感器种类。

机器人从汽车感知系统上继承了摄像头、超声波雷达、激光雷达等传感器, 并且提高了传感器数量和性能。此外, 为了实现更精准的操控, 机器人新增加了一维力和一维力矩传感器、关节电机内部的位置传感器、IMU、用于腕关节和脚部的六维力传感器、触觉传感器(指尖、手掌、高碰撞风险区)等部件。

- **力矩传感器:** 关节的一维力和一维力矩传感器主要针对关节的运用状态、输出力或力矩进行监测和反馈。以旋转执行器为例, 若采用谐波减速器会产生瞬时传动比不稳定等问题, 因此需要一维力矩传感器对关节的力矩输出状态做实时监测。
- **六维力传感器:** 一般用于人形机器人的脚部和手腕处, 能随时精确测量机器人关节在空间中所受力及力矩, 为优化手部运动控制及行走控制的重要部件。
- **触觉传感器:** 协助人形机器人获取如接触力大小和方向、温度、湿度、形状纹理等信息, 对非结构化环境中的稳定抓取、路径规划和避障等至关重要。根据敏感材料和工作原理的不同可分为压阻式、压电式、电容式、磁场式、光电式、超声波式等。
- **IMU:** 运用于人形机器人实现协助身体平衡、定位导航等功能。IMU 具备加速度计、陀螺仪等结构能够负责感知和控制人形机器人姿态和平衡, 同时, 还可以与摄像头、力传感器等多传感器数据融合, 达到预测人形机器人速度和轨迹并进行定位导航等功能。

- **视觉模块：**让人形机器人能“看得见”，包括环境感知、物体的识别和跟踪等，导入 AI、深度学习等技术后，人形机器人视觉可以实现理解并智能规划行进路线，准确判断物体类别、跟踪物体运动等功能。

图 25: Optimus 手指采用触觉传感器



资料来源：Tesla X 账户，长江证券研究所

图 26: Optimus 脚部采用的力和力矩传感器或为六维力传感器



资料来源：Tesla X 账户，长江证券研究所

## 算力芯片：云边端协同与异构化

**机器人模型部署包含云端大脑和端侧模型两条路径。**当前条件下，如何平衡云端的计算延迟与端侧的算力不足是机器人 AI 大脑的核心矛盾之一。一方面，云端大脑受数据传输速率影响，计算频率往往只有 1-3Hz，导致机器人动作高度时延，不利于日常工作，因此机器人部分计算必须由端侧模型完成。另一方面，受限于端侧芯片的算力不足，目前超过百亿参数的模型难以下沉至端侧与边缘侧，因此端侧模型的参数量需要压缩至百亿级以下才具备实用价值。放眼未来，机器人大脑在一段时间内或采用云端大脑与端侧模型并行的技术路径：端侧 VLA 模型有望蒸馏至百亿以下，主要负责日常低时延任务场景，同时端侧芯片算力也得到大幅度提升；云端大脑通过算力集群搭配千亿级别 VLA 模型，辅助负责需要高泛化能力的场景。

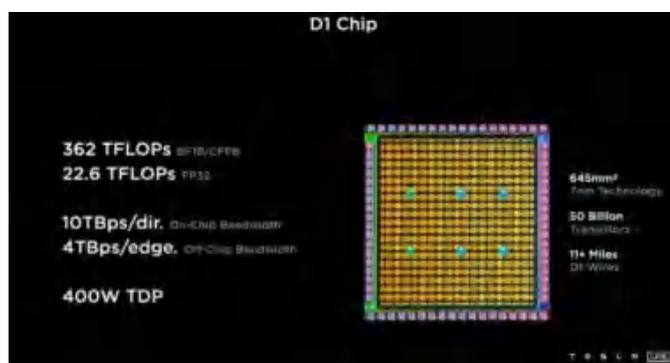
图 27: 云端大脑与端侧模型特点对比



资料来源：《RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control》，Brohan, et al., Figure AI X 账号，长江证券研究所

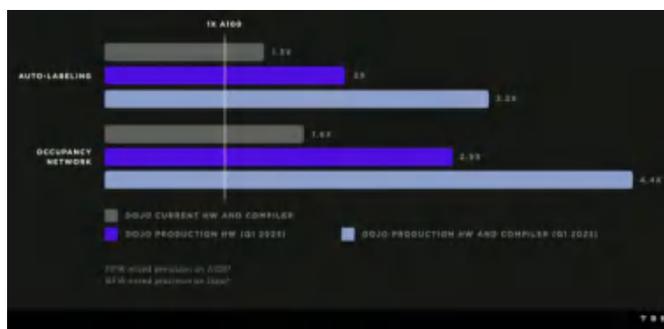
在云端，特斯拉自研 Dojo 超算平台支撑人形机器人的底层算力。特斯拉机器人 Optimus 采用 Dojo 超算平台作为底层 AI 算力支撑。Dojo 芯片架构平台由 Dojo D1 芯片组成，拥有一个大规模计算平面，极高宽带和低延迟，训练模块最大程度上实现了带宽的保留。D1 芯片采用了台积电的 7 纳米制程工艺，芯片面积为 645 平方毫米，拥有 500 亿个晶体管。每个 D1 芯片包括 360 个内核，出于良率和处理器核心的稳定性考虑，其中只有 354 个内核是可用的。D1 芯片的计算能力在 BF16/CFP8 上达到 362TFLOPS，FP32 峰值算力为 22.6 TFLOPS，TDP(热设计功耗)为 400W。D1 每个芯片节点都具备一个内核，拥有专用的 CPU 内存和 I/O 接口，每个内核还拥有一个 1.25MB SRAM 作为主存储器 D1 架构通过矩阵计算单元来增强计算能力，SRAM 能以 400GB/s 的速度加载数据，并以 270GB/s 的速度进行数据存储。

图 28: Dojo D1 芯片参数



资料来源：特斯拉 2021 AI Day，长江证券研究所

图 29: DOJO D1 性能超越英伟达 A100



资料来源：特斯拉 2022 AI Day，长江证券研究所

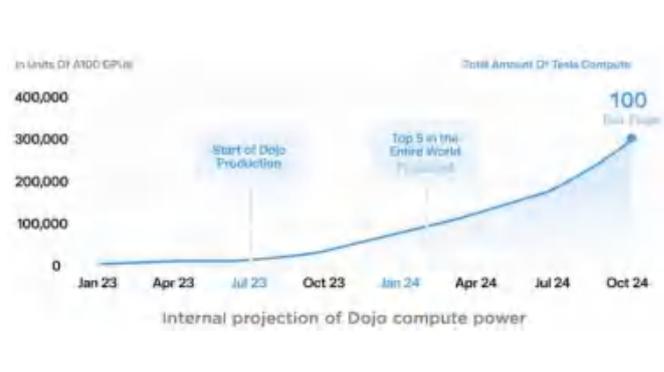
DOJO 算力集群构建了特斯拉核心竞争力。D1 芯片采用定制化的 ASIC 架构，聚焦于智能驾驶视觉数据，比传统 GPU 架构拥有更高的计算效率。据 2022 AI Day 数据显示，与英伟达的 A100 相比，一颗 D1 芯片在配合特斯拉自研编译器的情况下，在自动标注任务中最高能够实现 3.2 倍的计算性能，在占用网络任务中最高能够实现 4.4 倍的计算性能。预计到 2024 年 10 月，Dojo 超算平台的有望匹配 30 万块英伟达 A100 芯片的算力。据摩根斯坦利数据显示，特斯拉自研的 Dojo 超算平台有望为公司节省超过 65 亿美元的算力投入（对比购入同等算力下的 A100 芯片算力集群）。

图 30: Dojo 超算平台为特斯拉带来了什么



资料来源：摩根斯坦利，长江证券研究所

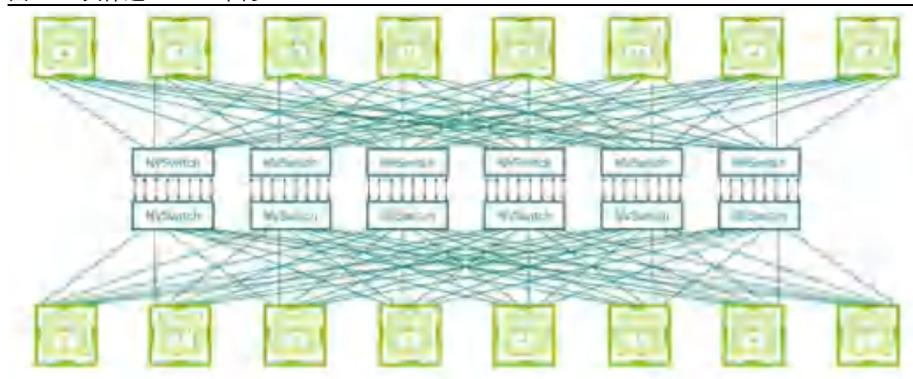
图 31: 到 2024 年 10 月 Dojo 超算平台算力有望达到 100EFlops



资料来源：Verge，长江证券研究所

**Dojo D1 超算芯片优化了 GPU 通信协议，相比 NVLink 架构效率更高。**在大模型训练时，单独一块超算芯片的显存难以储存大模型所有的参数量，而是需要上千块超算芯片的显存分别承载其中一部分；为了保证超算显存之间的高效联通，就需要英伟达 NVLink 这样的 GPU 通信协议。Dojo 超算平台优化了超算芯片结构，将 D1 芯片以 5x5 的方式封装形成一个瓦片 (training tile)，然后将 6 个瓦片组合成一个服务器。瓦片之间可以进行快速互联，不需要经过中间步骤。数据可以通过接口处理器进行连接，通过充电瓦片之间的二维扩展，获得比 GPU 更高的互联带宽。此外，特斯拉在软件算法方面通过变更自己的配套式框架支持和编译的 LVM 价值取代驱动，从应用层到服务器层形成了架构的完整融合。Dojo 超算平台的架构使其在功耗、算力和成本方面都优于英伟达的 GPU。

图 32：英伟达 NVlink 架构



资料来源：nextplatform，长江证券研究所

图 33：特斯拉 Dojo 超算中心结构



资料来源：特斯拉 2022 AI Day，长江证券研究所

## 商业模式：从制造到运营，机器人有望复刻路径

**具身智能将带来产业链的重构。**传统的机器人产业链主要分为上游的零部件供应商、软件系统提供商、中游的设备制造商和下游的应用场景。而在具身智能时代，机器人产业链或将发生重大变化。从上游角度来看，具身智能机器人多为人形机器人，从成本、续航能力、零部件性能等方面对机器人提出了新的需求。在传感器方面，由于机器人多用于复杂场景，所需传感器数量和性能需求都得到大幅度提升；与此同时，芯片的算力需求也大幅度提升。中游的 OEM 厂商部分会选择自研软件系统，部分会采购

上游的软硬一体化解决方案。最后，下游的应用场景将不断拓展。具身智能机器人可以应用于制造业、服务业、医疗保健等多个领域，应用场景和能力边界都得到了拓展。

图 34：人形机器人产业链图谱



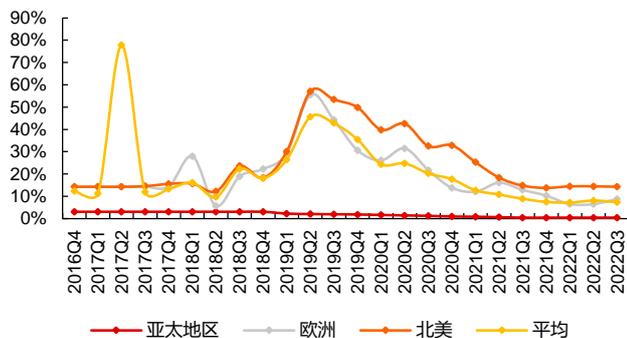
资料来源：优必选招股说明书，行行查，长江证券研究所

## 软件在产业链中具备最高价值量

**产品价值量有望向下游转移。**回顾智能手机产业与新能源汽车产业链可以发现，在产业落地的早期阶段，由于产品成熟度不高，OEM 主要扮演零部件整合的角色。在此阶段，上游的 Tier1 厂商议价权较高，OEM 的核心竞争力在于产业链整合能力；而随着产品大规模渗透，产业链逐渐成熟，产品价值量也在向下游转移，此时的 OEM 更重视软件生态、芯片等关键零部件自研能力。

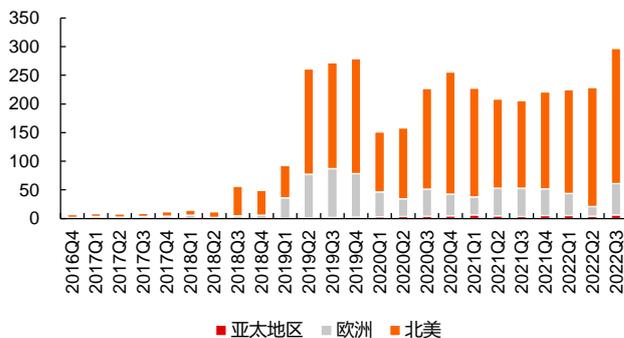
**软件在人形机器人产业链中具备最高价值量。**当前人形机器人产业链的硬件部分已有商业化落地的方案，产业驱动逻辑在于量产规模效应实现降本；而软件部分尚无成熟的技术方案，因此成为限制人形机器人发展的主要瓶颈。未来随着大模型的渗透率提升，机器人智能化程度有望大幅度提高。参考智能汽车领域的“软件定义汽车”逻辑，软件算法或成为人形机器人厂商的技术护城河，是机器人主机厂实现产品差异化的关键因素。

图 35：全球 FSD 搭载比率



资料来源：patreon，长江证券研究所

图 36：FSD 在全球各地的订阅收入（百万美元）



资料来源：patreon，长江证券研究所

表 6: AutoPilot 和 FSD 价格变化 (美国, 美元)

开始时间	基础 AutoPilot	增强 AutoPilot 价格	FSD 价格
2016.10-2019.02	-	5000	+3000
2019.02-2019.04	-	3000	+6000
2019.04-2019.08	免费	-	6000
2019.08-2020.06	免费	-	7000
2020.07-2020.10	免费	-	8000
2020.10-2022.01	免费	-	10000
2022.01-2022.06	免费	-	12000
2022.06-2022.09	免费	6000	12000
2022.09-2023.09	免费	6000	15000
2023.09-2024.04	免费	-	12000
2024.04-至今	免费	-	8000

资料来源: patreon, notateslaapp, 长江证券研究所

## 人形机器人或复刻 RoboTaxi 路径

2024 年 5 月起, 萝卜快跑在武汉推出了全天候 24 小时运营的 RoboTaxi 服务。萝卜快跑的 RoboTaxi 与以往的不同点在于, 车内无安全员监控, 所有自动驾驶行为监控都发生在云端。萝卜快跑基于百度 Apollo 智驾平台, 最新的第六代无人车是与江铃新能源合作, 采用磷酸铁锂电池, 改造前的整车的单价为 20.47 万元, 总体投放数量超过 400 辆。在具体的人员配置上, 萝卜快跑在武汉配备了 200 名左右的测试人员, 200 左右的地勤, 还有 60 到 70 个云端驾驶人员, 和其他的后台人员。

图 37: 国内不同城市无人驾驶落地节奏

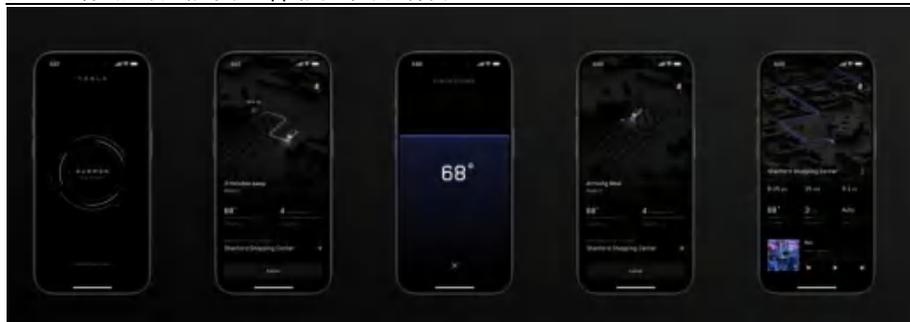


资料来源: 远川汽车评论, 长江证券研究所

RoboTaxi 有望颠覆智能驾驶商业模式。在 2024 年二季度财报电话会上, 特斯拉首席执行官马斯克表示, Robotaxi 将于今年 10 月 10 日发布, 预计今年年底前进入中国市场, 目标是今年内全球上线。马斯克将特斯拉 RoboTaxi 服务表述成 Airbnb 和 Uber 的

结合体；拥有搭载 HW3.0 或 HW4.0 的特斯拉电动车车主可以在空闲时间，将车辆“出租”成为 RoboTaxi，从而参与收入分成。

图 38：特斯拉设想的手机 App 的打车服务界面

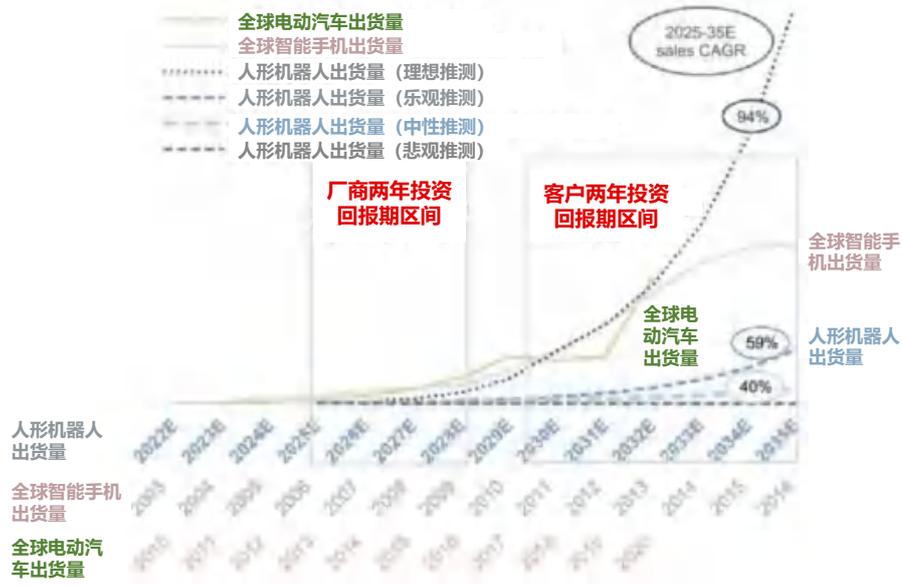


资料来源：2024 年 Q1 财报会，长江证券研究所

**数据驱动的 RoboTaxi 落地成本更低，渗透速度更快。**规则驱动智能驾驶厂商将车辆改装成 RoboTaxi，需要加装多个激光雷达、工控机、供电系统，雷达自清洁系统等部件，整体落地成本往往超过百万量级；而特斯拉仅需摄像头即可实现可靠的自动驾驶能力。因此，特斯拉 RoboTaxi 相比 Waymo、萝卜快跑等厂商具备较大的成本优势。另一方面，规则驱动的智驾公司仍需依赖高精地图，每新开一座城市的服务均需 Waymo 团队在该城市进行绘图、测试和验证，无法像特斯拉 FSD 一样快速泛化至多个城市。因此，Waymo 为代表的科技公司，难以在 Robotaxi 的商业化竞争中击败特斯拉。

**人形机器人有望复刻 RoboTaxi 商业路径。**4 月在特斯拉的一季度财报电话会议上，马斯克表示，“如果你拥有一个有感知能力的人形机器人，能够驾驭现实并根据要求执行任务，那么经济规模就没有任何有意义的限制。”从长期来看，人形机器人有望在近两年快速将本，仿照 RoboTaxi 的形式，在某些高价值场景中的成本低于人力成本，快速达到生产厂商的投资回报曲线门槛。

图 39：人形机器人有望仿照 RoboTaxi 模式，快速实现投资回报



资料来源：高盛，长江证券研究所

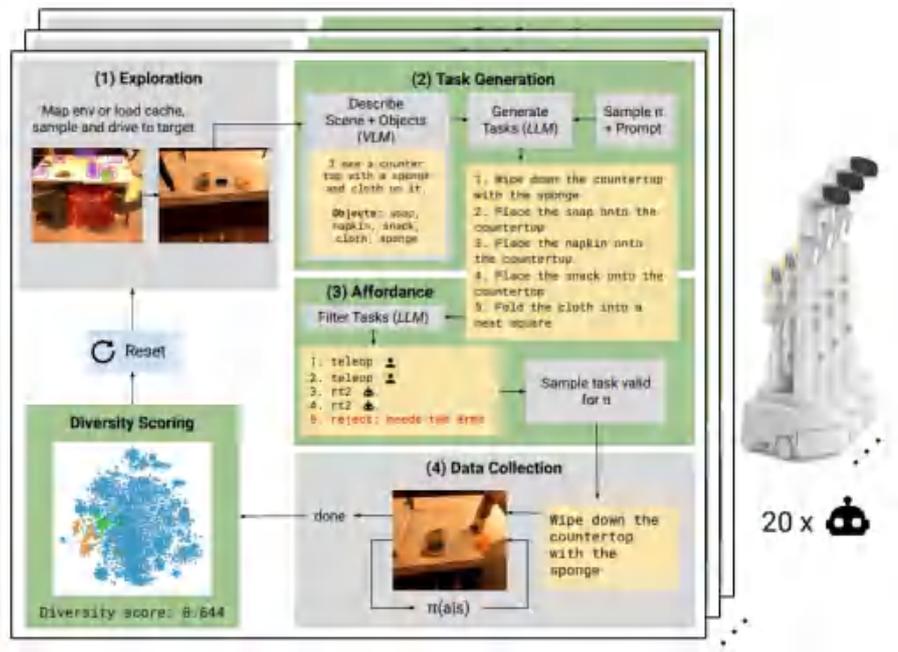
## 投资建议：关注模拟仿真相关标的

当前人形机器人产品落地的主要瓶颈在于模型侧。当前 Optimus 等人形机器人落地节奏推迟的主要原因在于模型能力不足。受限于人形机器人的模型发展不成熟，当前机器人尚未实现较高的泛化能力，导致机器人在物理世界难以取代人类。

模型侧对机器人多模态对齐数据有较高的需求。数据驱动的大模型需要大量机器人多模态对齐数据投喂，对机器人数据的数量和质量都有较高的需求。机器人厂商解决机器人数据短缺主要靠两种途径：

- 1) 通过为数众多的机器人在物理世界中收集数据。回顾人形机器人商业化落地，特斯拉 Optimus、优必选 Walker S、Figure 02 等人形机器人都选择汽车工厂作为首选落地方式。汽车工厂创造了较好的环境，可供机器人同时采集视觉、听觉、运控等多模态对齐数据。具体案例如谷歌在近期发布了 AutoRT 平台，可以同时指导多个配备视觉系统和末端执行器的机器人，在各种各样环境中执行多样化的任务，以此来主动收集机器人感知、决策与控制数据。

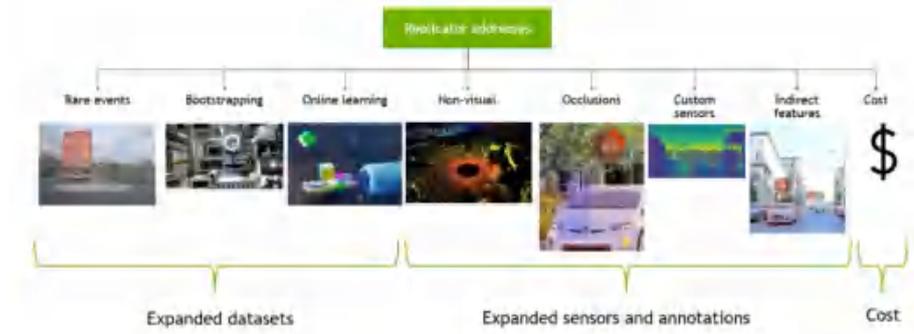
图 40：超过 20 个机器人使用 AutoRT 自主收集现实世界数据集



资料来源：《AUTORT: EMBODIED FOUNDATION MODELS FOR LARGE SCALE ORCHESTRATION OF ROBOTIC AGENTS》，Ahn et al.，长江证券研究所

2) 通过 AIGC、数字孪生等方式合成仿真数据。仅凭机器人在物理世界收集运控数据往往不够，模拟仿真通过另一种方式给机器人提供了充沛的数据源。具体案例如英伟达发布了 Isaac Lab、GROOT 等多款机器人开发中间件，可以运行数千个用于机器人学习的并行仿真。Isaac 软件解决方案包含 Isaac 机器人操作系统和 Issac 软件开发平台，可以快速实现机器人的软件开发。

图 41：ISSAC 数字孪生系统提升了机器人可用训练数据规模



资料来源：英伟达官网，长江证券研究所

## 风险提示

- 1、人形机器人仍处于技术突破的关键时期，若未来技术突破不及预期，人形机器人难以实现较强的智能化水平，产业发展或受阻。
- 2、人形机器人产业处于发展初期阶段，产品需求具备较高弹性。若未来人形机器人产品需求不及预期，产业发展或受阻。

## 投资评级说明

**行业评级** 报告发布日后的 12 个月内行业股票指数的涨跌幅相对同期相关证券市场代表性指数的涨跌幅为基准，投资建议的评级标准为：

看 好： 相对表现优于同期相关证券市场代表性指数

中 性： 相对表现与同期相关证券市场代表性指数持平

看 淡： 相对表现弱于同期相关证券市场代表性指数

**公司评级** 报告发布日后的 12 个月内公司的涨跌幅相对同期相关证券市场代表性指数的涨跌幅为基准，投资建议的评级标准为：

买 入： 相对同期相关证券市场代表性指数涨幅大于 10%

增 持： 相对同期相关证券市场代表性指数涨幅在 5%~10%之间

中 性： 相对同期相关证券市场代表性指数涨幅在-5%~5%之间

减 持： 相对同期相关证券市场代表性指数涨幅小于-5%

无投资评级： 由于我们无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使我们无法给出明确的投资评级。

**相关证券市场代表性指数说明：** A 股市场以沪深 300 指数为基准；新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准；香港市场以恒生指数为基准。

## 办公地址

### 上海

Add /虹口区新建路 200 号国华金融中心 B 栋 22、23 层

P.C / (200080)

### 武汉

Add /武汉市江汉区淮海路 88 号长江证券大厦 37 楼

P.C / (430015)

### 北京

Add /西城区金融街 33 号通泰大厦 15 层

P.C / (100032)

### 深圳

Add /深圳市福田区中心四路 1 号嘉里建设广场 3 期 36 楼

P.C / (518048)

## 分析师声明

本报告署名分析师以勤勉的职业态度，独立、客观地出具本报告。分析逻辑基于作者的职业理解，本报告清晰准确地反映了作者的研究观点。作者所得报酬的任何部分不曾与，不与，也不将与本报告中的具体推荐意见或观点而有直接或间接联系，特此声明。

## 法律主体声明

本报告由长江证券股份有限公司及/或其附属机构（以下简称「长江证券」或「本公司」）制作，由长江证券股份有限公司在中华人民共和国大陆地区发行。长江证券股份有限公司具有中国证监会许可的投资咨询业务资格，经营证券业务许可证编号为：10060000。本报告署名分析师所持中国证券业协会授予的证券投资咨询执业资格证书编号已披露在报告首页的作者姓名旁。

在遵守适用的法律法规情况下，本报告亦可能由长江证券经纪（香港）有限公司在香港地区发行。长江证券经纪（香港）有限公司具有香港证券及期货事务监察委员会核准的“就证券提供意见”业务资格（第四类牌照的受监管活动），中央编号为：AXY608。本报告作者所持香港证监会牌照的中央编号已披露在报告首页的作者姓名旁。

## 其他声明

本报告并非针对或意图发送、发布给在当地法律或监管规则下不允许该报告发送、发布的人员。本公司不会因接收人收到本报告而视其为客户。本报告的信息均来源于公开资料，本公司对这些信息的准确性和完整性不作任何保证，也不保证所包含信息和建议不发生任何变更。本报告内容的全部或部分均不构成投资建议。本报告所包含的观点、建议并未考虑报告接收人在财务状况、投资目的、风险偏好等方面的具体情况，报告接收者应当独立评估本报告所含信息，基于自身投资目标、需求、市场机会、风险及其他因素自主做出决策并自行承担投资风险。本公司已力求报告内容的客观、公正，但文中的观点、结论和建议仅供参考，不包含作者对证券价格涨跌或市场走势的确定性判断。报告中的信息或意见并不构成所述证券的买卖出价或征价，投资者据此做出的任何投资决策与本公司和作者无关。本研究报告并不构成本公司对购入、购买或认购证券的邀请或要约。本公司有可能会与本报告涉及的公司进行投资银行业务或投资服务等其他业务(例如:配售代理、牵头经办人、保荐人、承销商或自营投资)。

本报告所包含的观点及建议不适用于所有投资者，且并未考虑个别客户的特殊情况、目标或需要，不应被视为对特定客户关于特定证券或金融工具的建议或策略。投资者不应以本报告取代其独立判断或仅依据本报告做出决策，并在需要时咨询专业意见。

本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可升可跌，过往表现不应作为日后的表现依据；在不同时期，本公司可以发出其他与本报告所载信息不一致及有不同结论的报告；本报告所反映研究人员的不同观点、见解及分析方法，并不代表本公司或其他附属机构的立场；本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。本公司及作者在自身所知范围内，与本报告中所评价或推荐的证券不存在法律法规要求披露或采取限制、静默措施的利益冲突。

本报告版权仅为本公司所有，。未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布给其他机构及/或人士（无论整份和部分）。如引用须注明出处为本公司研究所，且不得对本报告进行有悖原意的引用、删节和修改。刊载或者转发本证券研究报告或者摘要的，应当注明本报告的发布人和发布日期，提示使用证券研究报告的风险。本公司不为转发人及/或其客户因使用本报告或报告载明的内容产生的直接或间接损失承担任何责任。未经授权刊载或者转发本报告的，本公司将保留向其追究法律责任的权利。

本公司保留一切权利。