

特斯拉的FSD时刻

太平洋证券研究院|首席分析师刘虹辰

执业资格证书编号：S1190524010002

2024年4月18日



特斯拉的FSD时刻。V12.3端对端NN重塑FSD。FSD 用户的累计行驶里程从V11扩大推送后加速提升，V12推送后开始指数级提升，到今年4月已超10亿英里，特斯拉FSD迎来奇点时刻。FSD V12.3的推送实现端到端NN，HW4.0大幅提升算力，不再进行识别和标记，摆脱人为规控规则编码，大幅提高自动驾驶感知决策能力。

AI赋能，加速Robotaxi。马斯克称Robotaxi将于今年8月8日落地，近期特斯拉启动了全球范围内的人员调整，自动驾驶预计将获得资源倾斜，大模型加持提升城市NOA感知能力、泛化能力，加速无图城市NOA落地，AI赋能，加速Robotaxi，特斯拉自动驾驶今年累计投资预计超100亿美元。数据鸿沟和算力要求，构成超算的核心壁垒。任何模型的训练都依赖于数量巨大、质量可控的训练数据，而特斯拉FSD自推出以来累计运行超过了10亿英里，为特斯拉训练模型提供了充足的数据。同时，大模型训数据需要强大的算力，特斯拉领先的超算中心，Robotaxi落地路径持续催化值得期待。

无图城市NOA重新定义智能汽车。特斯拉FSD引领“Transformer+BEV”、“Occupancy、Lane算法”、“自动标注”、“超算中心建设”，实现路径收敛。华为ADS 2.0 OTA新增不依赖高精地图的城区智驾领航辅助，问界全系车型实现了全国都能开、有路就能开的City NCA。支持城市道路(主干路/次干路/支路等)、国道、县道、乡道等公开道路,可用路段高达99%，用户体验提升带来渗透率拐点快速提高。

投资建议：智能汽车三条投资主线：①域控制器：德赛西威，华阳集团，科博达。②线控底盘：伯特利，保隆科技。③智能整车：赛力斯，理想汽车，小鹏汽车。

风险提示：智能汽车渗透率提升不及预期，国产替代不及预期，技术路线变更风险。

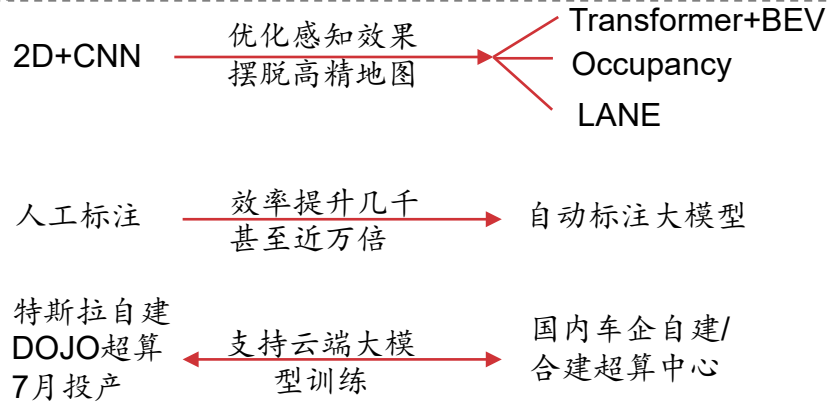


大模型加持提升城市NOA感知能力、泛化能力，加速无图城市NOA落地

接下来一段时间的催化预期？

投资建议

特斯拉引领技术路径收敛



大模型催化FSD时刻

端到端算法有望提升规控准确性

生成式AI提供高质量训练数据

大模型蒸馏训练小模型减少计算资源消耗

多模态模型有望优化感知

BOM成本下降

华为ADS1.0到2.0显著变化为从3颗激光雷达减配至1颗，纯视觉与激光雷达路线之争逐渐走向“折中”方案。高阶辅助驾驶BOM成本持续下降。

华为无图城区NCA全面OTA，特斯拉FSD北美，用户体验提升带来渗透率拐点快速提高。覆盖生活场景从高速NOA的10%到城市NOA的99%。

特斯拉8月8日Robotaxi。

法规：L3智驾政策。

① **域控制器**：德赛西威、科博达、华阳集团。

② **线控底盘**：拓普集团，伯特利、保隆科技。

③ **整车**：赛力斯，理想汽车，小鹏汽车

目录

CONTENTS

01

特斯拉的FSD时刻

02

AI赋能，加速Robotaxi

03

无图城市NOA重新定义智能汽车

04

投资建议

05

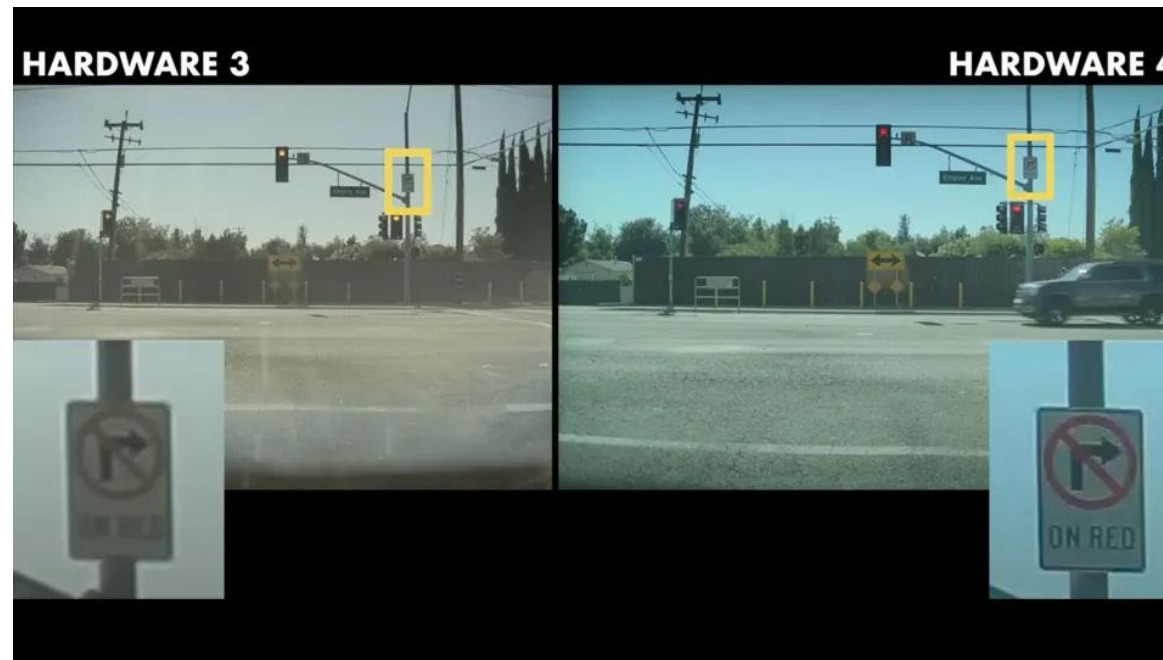
风险提示



特斯拉FSD：V12.3端对端NN重塑FSD

- 2024年3月，特斯拉先后推出FSD（Beta）12.3.2、12.3.2.1以及最新的FSD（Supervised）12.3.3版本。相比较于FSD 11，**FSD v12创新性地采用了端对端的神经网络技术（端对端NN）**，可以更好地理解和处理复杂驾驶环境，减少驾驶人员的干预，提高自动驾驶的精确度和自动化度。
- FSD v12**中采用端到端NN（以前是AI感知+软件规控），大模型仅需输入视频进行学习，而无需单独编码，更灵活更轻代码。感知-决策-执行是FSD的基本流程，在v11中，感知阶段需要通过视觉方式/雷达方式获取周围物体信息并识别和标识分类，决策阶段则依赖于事先认为编写好的控制规则。但是v12采用了端对端NN技术，感知阶段不再需要识别和标识，决策阶段也不需要事先人为编写控制规则，**只需要输入大量视频交给NN学习，就能分辨出在不同情况下需要做什么**，这使得特斯拉在**FSD V12中减少了100倍代码**，使其更轻便、更灵活，同时在没有网络连接的情况下仍能在不熟悉的地形上工作。
- V12同时兼容HW3.0和4.0**，感知力和算力大幅提升。v12在北美的已经适配装备HW4.0的model S、3、X和Y，在中国地区目前只有Model Y可以升级到HW4.0。从3.30号推送以来，截至4.9号，v12.3.3累计安装量达5780。目前，v12同时适配HW3.0和4.0。相比HW3.0,HW4.0在算力、摄像头数量和清晰度等方面具有大幅度的提升。摄像头为8个，算力提升5倍，素有摄像头像素由120万提升至500万。

图 截至4.9号，HW3.0摄像头清晰度（左）vsHW4.0摄像头清晰度（右）





特斯拉FSD：累计里程指数级增长超10亿英里，剑指百亿公里

行业深度报告

图：3月推送FSD v12.3实现端到端NN，大幅提高自动驾驶感知和决策能力

时间	版本进展	版本重大变化
22年12月	FSD Beta 测试版向北美40万人推送	截至2022年12月，特斯拉已经向美国和加拿大的几乎所有买了FSD（约40万）的客户发布了FSD Beta测试版，FSD Beta测试版新增了大量新功能，并可以让特斯拉在城市环境中自动导航。根据马斯克推特，FSD可用于北美90%道路，计划未来在欧洲推送FSD BETA。
23年2月	FSD11.3 版本	合并城市驾驶与高速公路驾驶堆栈，提升辅助驾驶能力。FSD11.3版本于2023年2月向用户推送，单堆栈融合了城市和高速堆栈，单堆栈将使特斯拉能够将其在城市驾驶中的学习内容应用到其自动驾驶高速公路导航功能中。
23年3月	FSD11.3.3 版本	Vision Park Assist使用占用网络代替超声波雷达。自22年10月起，特斯拉逐步移除超声波雷达，此次推送的新功能，代表着特斯拉用视觉占用网络算法替代被移除的超声波雷达，在摄像头的盲区通过记忆数据复现障碍物对车的距离。
23年5月	FSD11.4 版本	通过将自动标记车队数据集进一步增加 80k 个剪辑、提高自动标记算法的准确性以及调整训练监督的分布来实现以下优化：将部分切入的召回率提高了 39%，将由于变道到相邻车道而导致的假阳性切入的准确率提高了 66%，从而使整体变道预测错误减少了 33%。
24年3月	FSD12.3 版本	FSD v12.3采用端到端NN技术，不再进行识别和标记，摆脱人为规控规则编码，并实现了Autopark不依赖USS（超声波传感器）（最早在3.21发布的v11.4.9的更新中实现）

图：截至2024年4月6日，FSD 用户的累计行驶里程已超10亿英里，23年8月后加速上行（系v11扩大推送规模）





新版Autopark不依赖USS, “Tap to Park” 大幅提升用户体验感

- 新版的Autopark基于与FSD v12相同的NN技术, 不再依赖USS: 自2022年特斯拉放弃USS而改用摄像头以来, 未装备USS的新特斯拉的Autopark和Summon功能给用户带来的体验感远不如装备USS的车辆。在今年3月21号推出的FSD v11.4.9的更新中, 特斯拉推出了新版的Autopark, 并在之后的v12.3版本中沿用。在技术层面, 与旧版不同, 新版的Autopark使用了与FSD v12相同的NN技术, 底层技术优于旧版, 相比旧版有较大的提升, 大幅度提升了泊车系统的视觉效果和自动泊车能力。
- 显著改进视觉效果和自动泊车功能, “Tap to Park” 一键泊车降低人为干预需要。之前版本的Autopark具有两大主要缺陷: 1) 很难检测停车位并将其显示在屏幕上, 而且一次只能显示一个停车位。2) 停车速度通常非常慢, 如果周围有其他车辆, 几乎很难完成自动泊车。而新的Autopark基于NN技术, 显著改善了这两个问题, 停车速度更快, 不仅可以在显示屏上勾勒出停车位, 还可以同时显示各种停车位 - 包括平行空间。此外, 这次升级, 特斯拉真正实现了一键泊车 (Tap to Park), 当驾驶员选择好停车位后, 点击“开始按钮”, 特斯拉的车辆能够安全地倒车进入停车位, 驾驶人员只需要保持关注并随时准备在必要时取消自动停车操作即可, 大幅度降低了人为干预需要、提升用户体验。

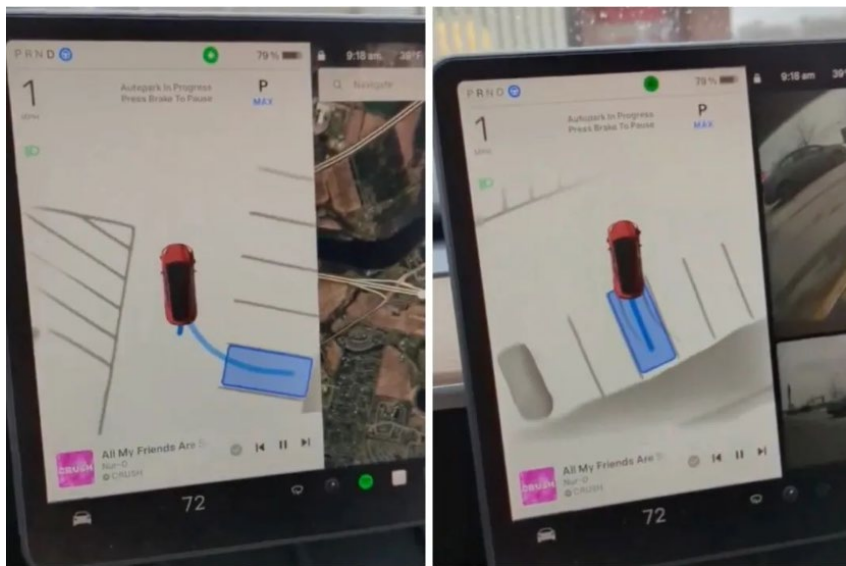


图 新版Autopark视觉体验感显著强于旧版Autopark

图 新版Autopark不再需要驾驶员操作方向盘。自动泊车功能显著改进





HW4.0芯片算力大幅提升

- **芯片改用更先进制程：**HW4.0 FSD芯片采用三星7nm工艺，HW3.0为三星14nm工艺，新的自动驾驶芯片性能将是现款自动驾驶芯片的**5倍**左右，这里的性能或为综合能耗/算力参数，或为单片算力（那么**新平台总算力很可能达到500TOPS**）。
- **摄像头或升级：**摄像头数量或从9个减少到8个，但是摄像头的清晰度从120万像素提升到500像素，这使得HW4.0具有比3.0更强的感知能力，最远探测距离可达424米，扫除盲区和死角。
- **预留雷达接口，但model Y目前未配备雷达：**HW 4.0预留了装备雷达的空间，但是model Y在出厂时并不具备雷达功能，这可能与特斯拉更加偏好视觉方案、降低车辆出厂成本等因素有关，同时也为用户提供了一定的选择空间。
- **CPU内核数量提升66.67%，FSD Computer 2数量增加到3个。**CPU方面，HW4.0由3.0的12核，提升到了20核，内核数量提升66.67%。计算平台方面，虽然HW4.0仍然采用的是FSD Computer 2，但是数量却增加到了3个。
- 目前，海外HW4.0已经支持model 3、X、Y和S。国内，特斯拉为了应对竞争压力，开放了model Y的升级，并且是“加料不加价”。

表：HW3.0和HW 4.0关键参数对比

	HW 3.0	HW 4.0
摄像头	9个，120万像素	8个，500万像素
雷达	1个毫米波雷达，12个超声波雷达	为安装雷达预留空间，但是Model Y目前未配备雷达
计算平台	FSD Computer 2	FSD Computer 2
计算平台数量	1	3
芯片	三星的 Exynos架构	三星的 Exynos架构
CPU	12核	20核
CPU频率	2.2GHz	2.35GHz
神经网络加速器	双核	三核
神经网络频率	2GHz	2.2GHz
最大算力	100TOPS	500TOPS
制程	14nm	7nm



HW5.0：或搭载第二代 3nm N3P 工艺芯片，摄像头或加入加热镜头元件等

行业深度报告

- **HW5.0或搭载第二代 3nm N3P 工艺芯片：**与N3E相比，N3P的晶体管密度更高，从而提高了性能的同时降低了功耗。N3P 性能比N3E 高 5%，同时光学尺寸缩小了 2%，从而晶体管密度达到 1.04 倍。据Digitimes报道，特斯拉已向台积电第二代 3nm N3P 工艺芯片下达了异常大的订单，如果此消息属实，HW5.0的性能将比HW4.0具有显著地提升。
- **摄像头工艺再提升，加入加热镜头元件、水滴排斥涂层和光圈控制技术。**
 - 1) **加热镜头元件：**三星开发的“Weather Proof”摄像头配备了8MP传感器，并采用了一种巧妙的镜头加热系统，能在不到一分钟内融化覆盖在镜头上的冰雪。
 - 2) **水滴排斥涂层：**新型摄像头配备了防水涂层，可以防止水滴停留在镜头上，添加了防水涂层，以防止水滴停留在镜头上，这种涂层耐划痕和损伤，耐用性是现有解决方案的6倍。
 - 3) **光圈控制技术 (IRIS)：**三星开发的摄像头模块引入了光圈控制技术，允许摄像头根据不同的光照条件和自动驾驶需求动态调整进光量，优化图像质量和传感器性能。

图 光圈控制技术 (IRIS) 能够显著优化成像质量



目录

CONTENT

01

特斯拉的FSD时刻

02

AI赋能，加速Robotaxi

03

无图城市NOA重新定义智能汽车

04

投资建议

05

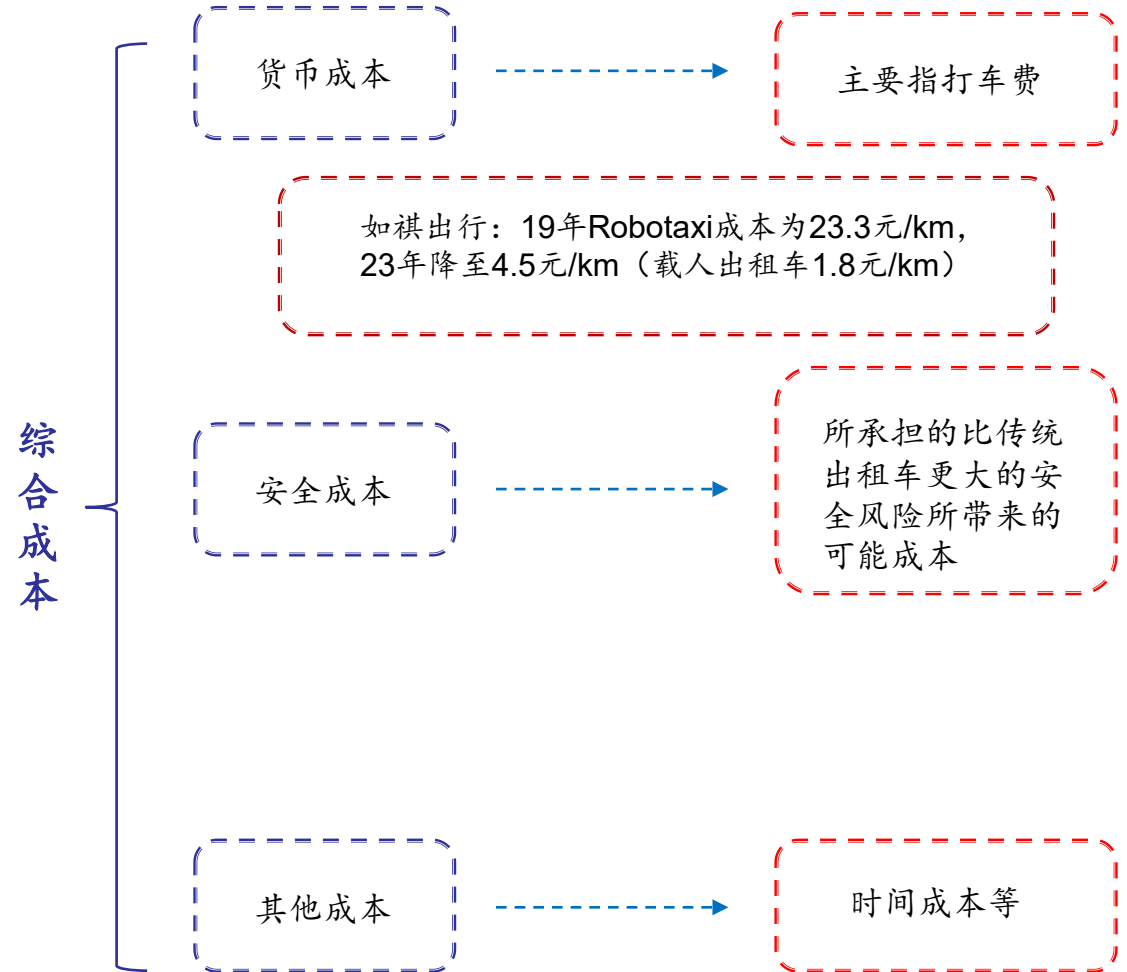
风险提示



Robotaxi: 8月8日落地路径Roadmap持续催化

马斯克称Robotaxi将于今年8月8日落地，近期特斯拉启动了全球范围内的人员调整，自动驾驶预计将获得资源倾斜，落地路径Roadmap持续催化值得期待。

- **乘坐Robotaxi的综合成本：**我们将消费者乘坐Robotaxi的综合成本划分成**货币成本、安全成本、和其他成本**。其中货币成本主要指打车费，安全成本则是消费者乘坐Robotaxi所承担的比传统出租车更大的安全风险，以及由此造成的乘客可能成本（如健康损失、医疗费用），其他成本则主要包括时间成本等（Robotaxi可能因为速度或错误选择匝道进入等导致平均速度相对慢）等。
- **综合成本降低是规模化落地的关键：**Robotaxi面临的竞争对手是传统的出租车行业，其商业模式要想跑得通，就要从传统出租车Uber等公司抢客，客观需要其具有相对低的综合成本吸引顾客。Robotaxi每公里货币成本已经显著下降，**当前更重要的或是降低安全成本**。
- **自驾技术进步是降低综合成本的核心，规模落地依赖L4/L5技术，目前或局限于部分城市的部分区域路段。**目前，国内外都存在Robotaxi的训练，但是仅局限于部分城市部分区域。在这一层级上，L3层级的自动驾驶技术可以满足需求。但是为了规模化推广，意味着自动驾驶技术要能适应更复杂的驾驶环境、降低综合成本，L4则是必须的。基于此，**我们认为马斯克在8.8落地的Robotaxi或将局限于部分城市的部分区域，规模化落地仍有待自动驾驶技术向L4/L5方向进一步发展。**
- **外部条件：监管政策宽松和有关法律完善。**




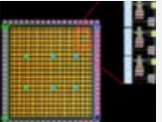




AI赋能，加速Robotaxi，特斯拉自动驾驶今年累计投资预计超100亿美元

行业深度报告

- 超算需巨额资金投入，马斯克预计特斯拉今年在有关领域累计投资超100亿美元。在今年1月27日，马斯克在推特上表示，Dojo目前的算力进相当于1万个英伟达H100，如果想在AI领域具备竞争力，那么目前需要每年投入数十亿美元。随后在4月6号，马斯克再发言，表示特斯拉在训练计算、数据管道和视频存储方面的累计投资将在今年超过100亿美元。AI赋能，加速Robotaxi。
- 数据鸿沟和算力要求，构成超算的核心壁垒。首先，任何模型的训练都依赖于数量巨大、质量可控的训练数据，而特斯拉FSD自推出以来累计运行超过了10亿英里，这位特斯拉训练模型提供了充足的训练数据。其次，算力方面，大模型训数据需要强大的算力，特斯拉目前可能拥有3万-3.5万颗英伟达H100芯片，而下一代的Gork可能需要10万颗H100，如果按照单价4.5万美元/颗，那么10万颗H100总价值高达45亿美元，相当于300多亿人民币。

图 Dojo的数据流近存计算架构

层级	名称	片上SRAM	算力	说明
 内核	Dojo Core	1.25MB	1.024TFLOPS	单个计算核心，64位宽度，具有4个8x8x4的矩阵计算单元，2GHz主频
 芯片级	D1	440MB	362TFLOPS	单芯片，核心数为354，面积645mm ²
 核心级	Dojo Tile	11GB	9050TFLOPS	单个训练模组，每5x5个芯片包括一个训练模组
 整群级	ExaPOD	1320GB	1.1EFLOPS	模拟训练的训练集群，每12个训练模组组成一个机柜，每10个机柜组成ExaPOD，共计3000个D1芯片



AI赋能，加速Robotaxi，无图全城城市NOA加速

- 华为ADS 2.0问世后，城市NOA的竞争愈发向着无图化、全城市可用的方向发展。华为的ADS 2.0首次搭载在华为问界M5上，相比与ADS 1.0, ADS 2.0结合道路拓扑推理网络（RCR）和华为的先进算法,实现了不依赖于高精地图的高速、城区高阶智能驾驶功能。小鹏汽车宣布面向智驾经验用户即刻推送无限 XNGP 智能辅助驾驶功能，号称“不限城市、不限路线，有导航的地方就能用”。

理想、比亚迪、华为、小鹏等智驾头部车企将自动驾驶放在优先发展战略地位，我们认为原因如下：

- 智能驾驶的技术路径逐渐收敛，后发力者技术上具备后发优势。特斯拉的智能驾驶技术路线已得到成功验证，我们看到，当前时点，“Transformer+BEV感知大模型”、“Occupancy、Lane算法”、“自动标注”、“超算建设”等成为共识。从中长期看，纯视觉与激光雷达路线之争逐渐走向“折中”方案、“不依赖高精地图”有望降低智能驾驶软硬件BOM成本。

我们认为在未来大模型将继续赋能自动驾驶。

- 端到端算法有望提升规控准确性、大模型蒸馏训练小模型减少计算资源消耗、生成式AI提供高质量训练数据、多模态模型有望优化感知。

图：不同车企高速NOA及城市NOA Roadmap

品牌	高速NOA	城市NOA	24年计划	指定路线NOA功能
小鹏	√	不限城市	/	AI代驾
阿维塔	√	不限城市 (阿维塔12)	/	AI代驾
问界	√	不限城市	/	AI代驾
理想	√	113城	无图城市NOA预计2H全国开放	通勤NOA
蔚来	√	726城	预计Q2向所有NT2.0用户推送NOP+城区领航辅助	
特斯拉	√	国外开通		
零跑	√	×	预计2H推出	
腾势	√	40+	最快年底覆盖全国	
极氪	√	×		
深蓝	×	×		
宝骏	√			记忆行车

图：华为 ADS2.0和小鹏无限XNGP都支持不限城市、无图NOA

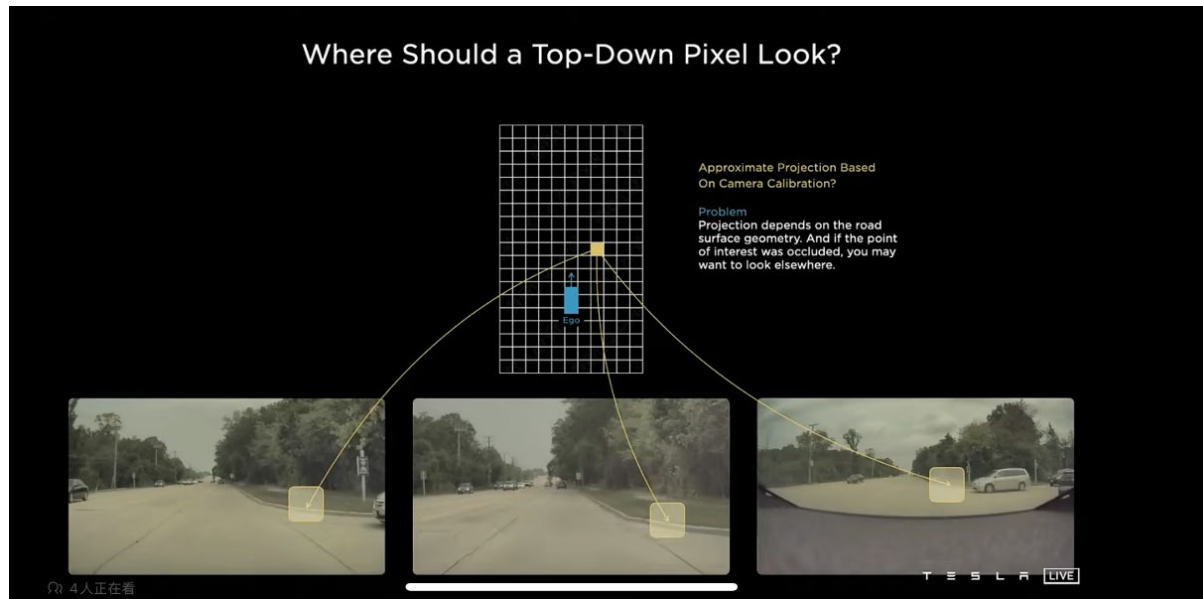
时间	NOA	特点
2023.05	华为 ADS2.0	1、配备1个顶置激光雷达、3个毫米波雷达、11颗高清摄像头以及12个超声波雷达，不受恶劣天气和光线影响，全天候感知信号。 2、BEV+GOD网络（General Objective Detection通用障碍物检测），可以识别通用障碍物白名单外的异形物体。 3、有图无图都能开，覆盖超90%城区
2024.04	小鹏无限 XNGP	不限城市、不限路线，有导航的地方就能用



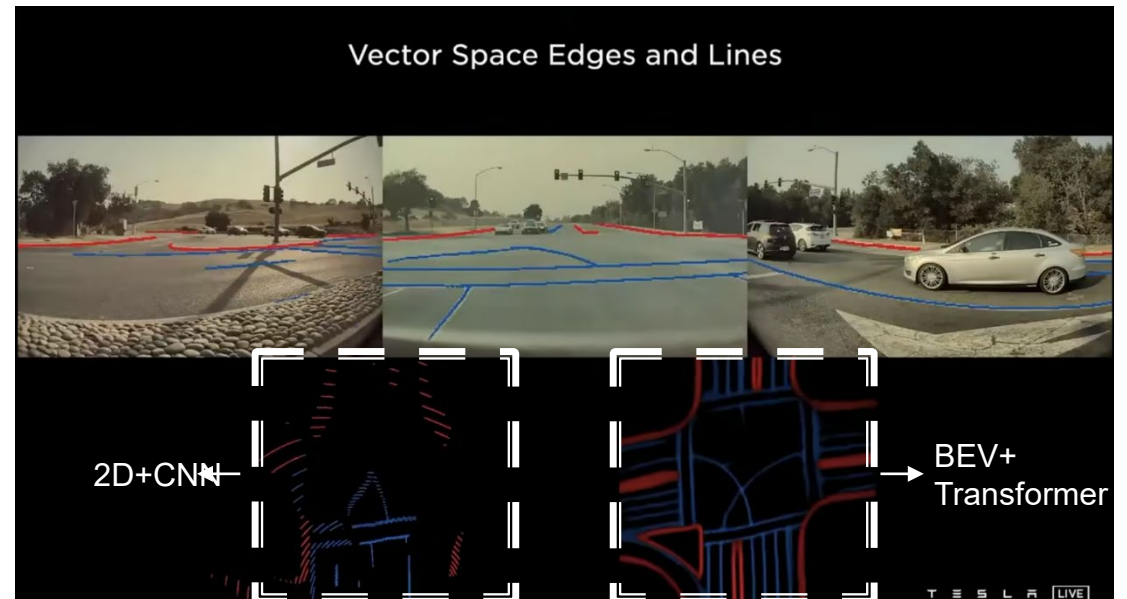
2021年7月AI DAY，特斯拉首次展示基于Transformer的BEV感知方案，是大模型首次用于自动驾驶，实时感知生成向量空间。

- BEV (Bird's-eye-view) 指俯瞰图视角，Transformer利用主干网络对各个摄像头进行特征提取，再利用Transformer将多摄像头数据从图像空间转化为BEV空间（向量空间）。
- Transformer vs CNN: Transformer是一种基于注意力机制（Attention）的神经网络模型，通过注意力机制去挖掘序列中不同元素的联系及相关性，而CNN的局部感受野限制了它们在长距离依赖性建模方面的能力。
- BEV+Transformer vs 2D+CNN: BEV+Transformer可以消除遮挡和重叠，提高检测精度。BEV视角下的物体，不会出现图像视角下的尺度（scale）和遮挡（occlusion）问题。BEV能够实现“局部”端到端优化，感知和预测都在同一个空间进行，输出“并行”结果。

图：自动驾驶跨摄像头和多模态融合背景下的BEV视角



图：BEV+Transformer可以消除遮挡和重叠，提高检测精度

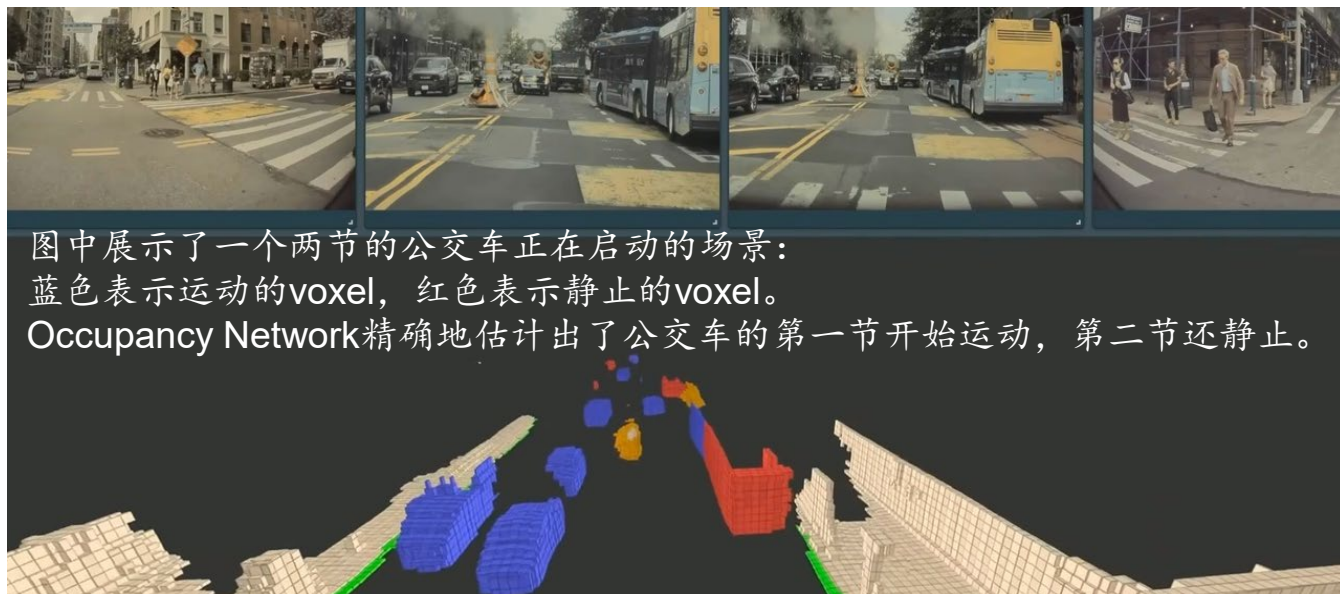




2022年10月AI DAY， 特斯拉Occupancy感知进化， **不识别类别也能判断物体运动状态**， 引领华为GOD、理想Occupancy跟随。

- 2022年10月1日的AI Day中展现了Occupancy Network 感知技术：基本的思想是将三维空间划分成体素voxel，通过0/1赋值对voxel进行二分类，实际中的赋值可以是概率值，表示voxel存在物体的概率。Voxel的属性除了是否被占据，还包含语义信息和速度信息（Occupancy Flow）等。其中速度信息是一个三维向量，用于表述voxel运动的速度和方向，通过occupancy flow可以用于判断物体是否运动。
- 相较BEV+Transformer，Occupancy是又一里程碑式进步：占用网络相较于BEV，有了Z轴上的扩展，是一种通用的2D特征转3D特征的转换，不局限于通用障碍物检测，可以扩展到更多3D空间任务。同时在引入速度信息后，动态描述向4D过渡。
- 2Q23，华为、理想相继发布类似的算法，其中华为创新性地将激光雷达信息融入其GOD网络，理想Occupancy属于其三种重点感知算法之一。

图：通过Occupancy flow可以用于判断物体是否运动



图中展示了一个两节的公交车正在启动的场景：
蓝色表示运动的voxel，红色表示静止的voxel。
Occupancy Network精确地估计出了公交车的第一节开始运动，第二节还静止。

优势

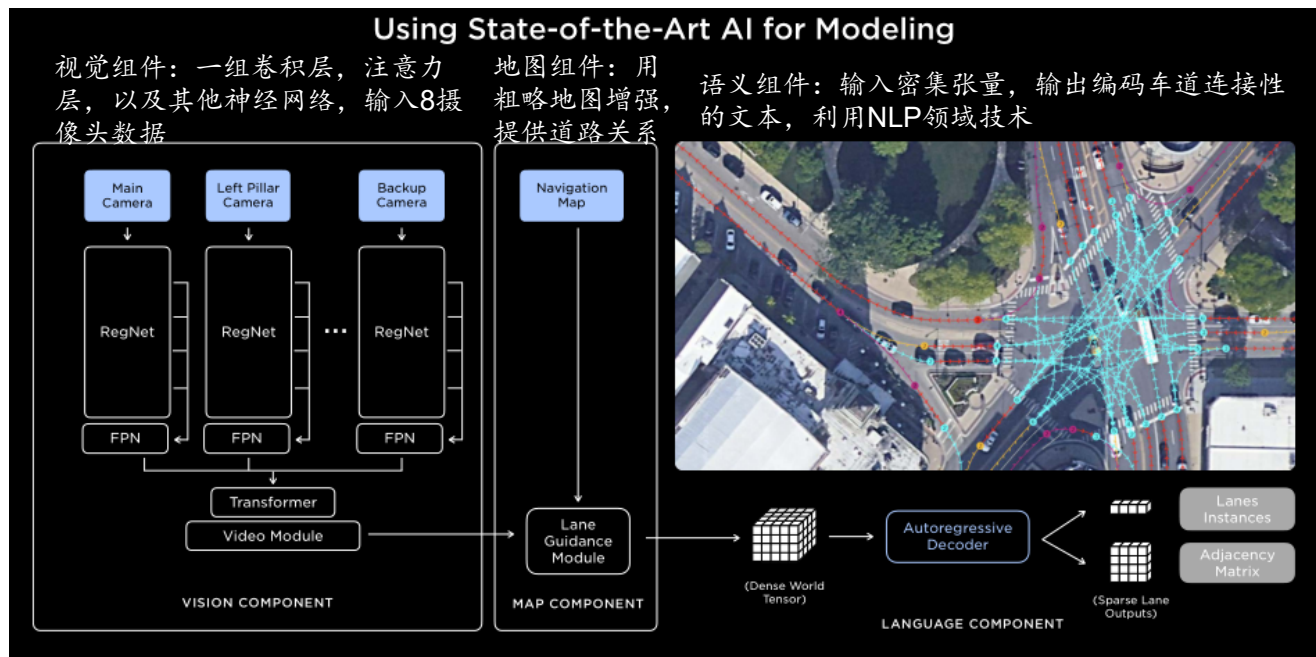
- **可预测道路起伏度**：在BEV空间生成统一体素，可以预测任意一个体素的占用概率，道路边沿的占据栅格和道路表面贴合。
- **运动流，3D方便测算遮挡关系**：能够实时预测被遮挡物体的运动状态，可以为每个体素预测其运动状态，对随机运动进行建模。
- **弱语义，强几何**：可以为每个体素生成对应的语义类别，即使不识别类别也能处理运动物体。
- **Occupancy具有高效的存储和计算优势**：10ms内可以完成计算，各个位置的分辨率可调整，具备BEV空间变焦能力。



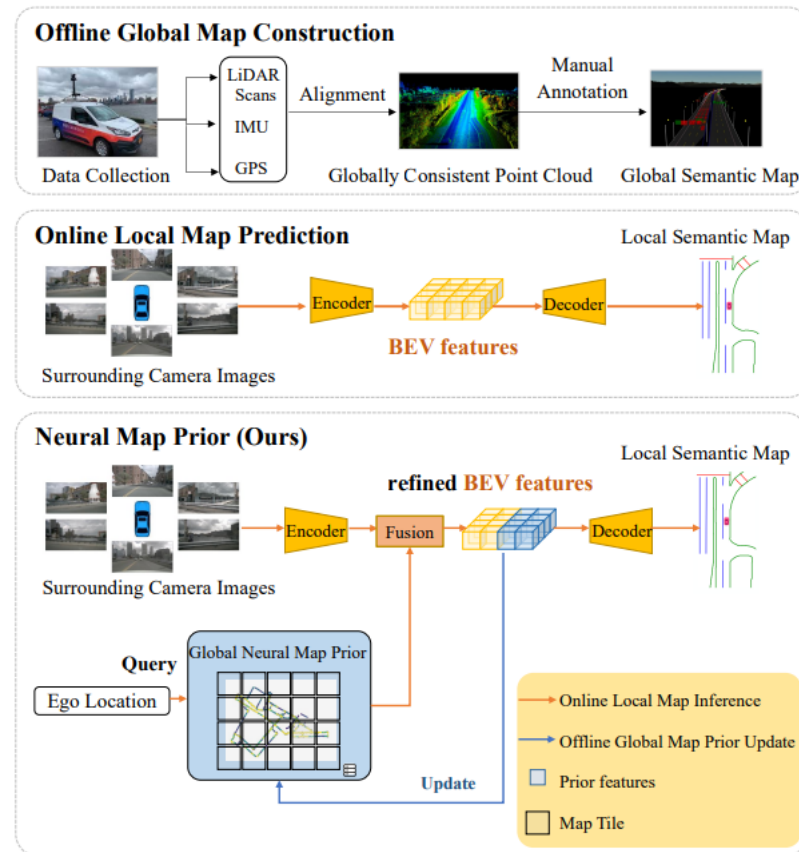
技术路径收敛：Tesla Lane感知车道拓扑，理想提前提取复杂路口特征

行业深度报告

图：Lane神经网络同样基于Transformer模型，是摆脱高精地图的关键算法，可生成道路间的拓扑连接关系



图：针对复杂路口，理想自研NPN网络提前进行路口特征的提取，当车辆再次行驶到该路口时，与车端感知的BEV特征层融合

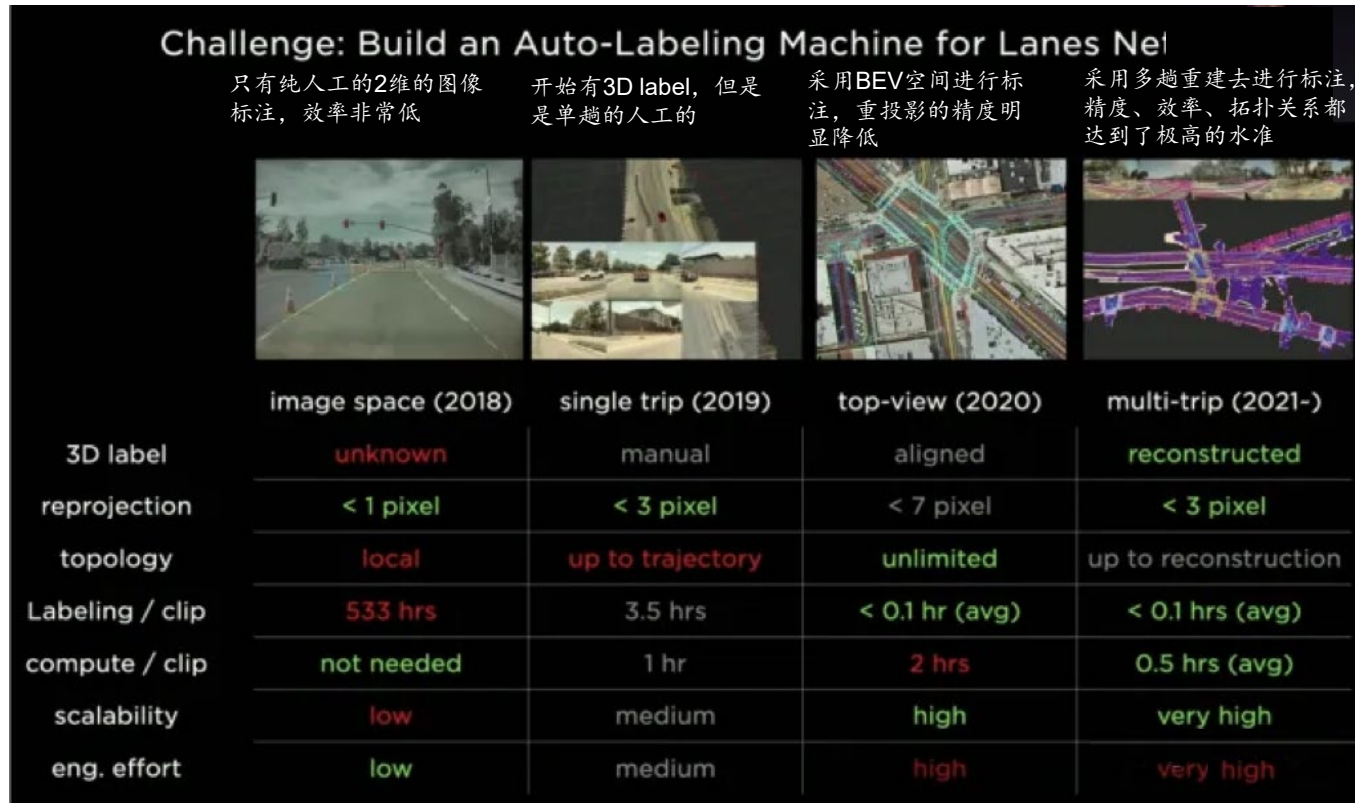




技术路径收敛：大模型自动标注效率是人工的1000-45000倍

行业深度报告

图：特斯拉采用大模型优化数据标注，从2018年只有纯人工的2维标注，2021年后建立自动标注系统可以取代500万小时的人工作业量



图：国产新势力、比亚迪、华为、长城、AI公司纷纷采用自动标注大模型提升效率

企业	自动标注大模型应用及能力
小鹏	引入全自动的标注系统，对比人工标注，全自动标注系统效率高，过去2000个人一年的标注任务，现在用16天左右就能完成，效率可以提升45000倍
理想	使用软件2.0的大模型，通过训练的方式进行自动化标定，过去需要用一年做的事情，基本上3个小时就能完成，效率是人的1000倍
比亚迪	毫末智行在算法模型上，研发了智驾自动标注大模型
华为	华为云研发的标注大模型，基于海量典型数据进行预训练，通过语义分割、对象跟踪技术可以快速完成长时间连续帧的自动标注
毫末智行	DriveGPT的自动化标注能力，可以降本98%数百亿4D Clips (环视连续数据Clips) 的标注成本。
商汤	商汤科技公布了“日日新SenseNova”大模型体系，包含自然语言处理、内容生成、自动化数据标注、自定义模型训练等多种大模型及能力。

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/746204240210010121>