

➤ **概述：FSD 软硬件全链路闭环 技术端快速进化。**特斯拉 FSD 是一套包含感知、规控、执行在内的全链路自动驾驶软硬件架构，在算法、算力、数据、芯片等层面实现了高度集成。FSD 技术端快速进化，V12 为首个端到端自动驾驶系统，能够模拟人类驾驶行为，实现感知决策一体化。特斯拉 Robotaxi 即将正式推出，有望成为特斯拉发展历程重要里程碑。

➤ **算法端：创新算法优化 实现高效神经网络推理。**1) **感知模块**，采用 HydraNets 架构，整合多个视觉识别任务到单一网络，通过 BEV+Transformer 技术，实现对车辆周围环境的感知；2) **规划模块**，引入基于神经网络的规划模块和蒙特卡洛树搜索，提高自动驾驶决策质量。通过算法端全栈自研，特斯拉以低成本感知硬件实现高阶智驾能力，快速优化迭代自动驾驶算法。

➤ **算力端：布局 Dojo 超级算力 支撑远期算力需求。**特斯拉从算力芯片开始，完整构建 Dojo 超级计算机系统，以处理自动驾驶所需海量数据，旨在摆脱对英伟达 GPU 的依赖，为远期算力瓶颈进行前瞻布局。其中，自研 D1 芯片作为 Dojo 核心，采用存算一体架构，高效优化数据传输和处理效率，为 AI 训练提供强大支持。

➤ **芯片端：自研 FSD 芯片 集成高效计算能力。**特斯拉自动驾驶硬件平台 HW1.0、HW2.0 分别与 Mobileye、英伟达合作推出，2019 年推出 HW3.0 系统，正式转向硬件平台全面自研，下一代全自动驾驶硬件 AI 5 预计将于 2025 年下半年投产。特斯拉自研 FSD 芯片采用异构设计，计算性能和高效的数据处理能力卓越。通过芯片自研，特斯拉能够实现硬件方案的持续快速迭代，与软件算法进行更好的整合，从而实现更优的系统性能。

➤ **数据端：自动化数据闭环 优化数据驱动训练。**1) **自动标注**，特斯拉开发自动标注系统生成 4D 标签，实现人工与机器相结合的数据标注，大幅提高标注效率；2) **仿真模拟**：构建虚拟仿真空间，安全且高效地重现和构建极端及复杂场景以进行算法训练，赋能模型迭代；3) **数据引擎**，特斯拉将采集的数据通过影子模式筛选出具有语义信息的有效数据，回传至云端，实现数据和模型持续优化。特斯拉形成综合训练数据集，用于训练车端的在线模型，形成闭环的数据流，实现自动驾驶系统的持续优化。

➤ **FSD 技术端快速进化 实现感知决策一体化。**2023 年 12 月，特斯拉正式推出 FSD Beta V12，是首个实现端到端 AI 自动驾驶的系统，采用单一深度学习模型处理从原始输入到驾驶决策的全过程，提升自动驾驶的效率和适应性。特斯拉 FSD 快速进化，V13 即将 10 月推出，有望于 2025Q1 进入中国和欧洲，智能驾驶拐点已至。

➤ **投资建议：**FSD 技术端快速进化，“We, Robot” 发布会即将开幕，Robotaxi 即将推出，有望成为特斯拉发展历程重要里程碑，并与 FSD V13 发布形成共振，共同成为板块强劲催化，智能驾驶奇点已至：推荐 1) 智能驾驶-【**伯特利、德赛西威、经纬恒润-W、科博达**】+智能座舱-【**上声电子（智能声学）、继峰股份（智能座椅）、光峰科技（车载光学）**】；2) T 链-【**拓普集团、新泉股份、爱柯迪、旭升集团、双环传动、岱美股份、银轮股份**】。

➤ **风险提示：**智能化渗透率提升不及预期，智能驾驶技术发展不及预期，Robotaxi 商业化进展不及预期。

推荐

维持评级



分析师 崔琰

执业证书：S0100523110002

邮箱：cuiyan@mszq.com

研究助理 完颜尚文

执业证书：S0100124040021

邮箱：wanyanshangwen@mszq.com

相关研究

1. 汽车和汽车零部件行业周报 20241007：假期港股汽车走强 看多强势车企+智驾-2024/10/07

2. 新势力系列点评十一：新势力销量新高 供给驱动加速-2024/10/03

3. 汽车和汽车零部件行业周报 20240929：汽车板块投资思考：看多强势车企+智驾-2024/09/29

4. 汽车和汽车零部件行业周报 20240922：2024Q3 前瞻：自主崛起提速 规模效应加强-2024/09/22

5. 摩托车行业系列点评八：中大排增势强劲 出口加速恢复-2024/09/20

目录

1 引言	3
2 概览：软硬件全链路闭环 全栈自研构筑强大技术基石	5
2.1 发展历程：软硬件持续迭代 由外部合作到全栈自研	5
2.2 系统架构：软硬件全链路闭环 实现高度集成	6
3 算法端：创新算法优化 实现高效神经网络推理	8
3.1 感知：先进感知技术，确保精准环境理解和物体识别	8
3.2 规划：采用交互搜索框架 快速选择最优方案	18
4 算力端：Dojo 超级算力支撑 保持能效比优势	21
4.1 Dojo 系统：特斯拉自研超算平台 研发进展迅猛	21
4.2 D1 芯片：基于存算一体架构 实现高效数据传输	24
4.3 瓦片集群：由基本单元出发 构成大规模算力集群	26
4.4 软件系统：采用高效同步机制 性能较 GPU 大幅提升	27
5 芯片端：自研 FSD 芯片 集成高效计算能力	30
5.1 硬件架构：由合作迈向自研 运算性能强大	30
5.2 编译推理：高效优化 最大化计算资源利用	34
6 数据端：自动化数据闭环 优化数据驱动训练	36
6.1 自动标注：自动化 4D 标签生成 提升标注效率	36
6.2 仿真模拟：构建虚拟场景 优化算法表现	39
6.3 数据引擎：集成多源数据 加速模型训练	42
7 端到端架构：FSD V12 引领 实现感知决策一体化	46
7.1 特斯拉的 AI 关键时刻：FSD V12 首次实现端到端 V13 即将推出	46
7.2 大模型成就端到端自动驾驶 推动感知决策一体化	48
7.3 端到端架构演进：感知端到端到 One Model 端到端	49
8 投资建议	51
9 风险提示	52
插图目录	53
表格目录	54

1 引言

本报告为特斯拉 FSD 专题报告，从算法端、算力端、芯片端、数据端四个层面出发，对 FSD 系统底层技术原理进行全面深度拆解，并结合端到端架构演进趋势，对 FSD 系统的最新发展变化进行梳理，对板块后续核心催化进行展望。

FSD 是一套包含感知、规控、执行在内的全链路自动驾驶软硬件架构，在算法、算力、数据、芯片等层面实现了高度集成：

1) 算法端：感知规划算法全栈自研，实现从纯视觉信息输入到规划方案输出。

1) 感知。特斯拉采用 BEV+Transformer 架构，将 2D 图像转化为对周围环境的准确 3D 感知。而后，特斯拉将该架构升级为 Occupancy Network，能够直接在向量空间产生体积占用，精准识别物体运动状态差异；2) 规划。特斯拉采用交互搜索框架，以任务分解的方式对一系列可能的行驶轨迹进行研究，实现对规划方案的实时评估。通过算法端全栈自研，特斯拉以低成本感知硬件进行高阶智驾能力输出，快速实现自动驾驶算法优化迭代。

2) 算力端：从 0 到 1 构建超级计算机系统，为远期算力提供强大支撑。特斯拉从算力芯片开始，完整构建 Dojo 超级计算机系统，以处理自动驾驶所需海量数据。2021 年 8 月，Dojo 在特斯拉首届 AI Day 上正式亮相，定位为超高速训练计算机，采用分布式计算架构设计，算力分为内核级、芯片级、格点级、集群级等四个层级，实现从训练节点到训练集群的完整构建。特斯拉从 0 到 1 构建超级计算机系统，旨在摆脱对英伟达 GPU 的依赖，为远期算力瓶颈进行前瞻布局。

3) 芯片端：由合作迈向自研，实现高性能算力集成。特斯拉自动驾驶硬件平台初期与 Mobileye、英伟达等合作，2019 年正式发布基于自研 FSD 芯片的 HW 3.0 系统，开始转向硬件平台全面自研，下一代全自动驾驶 (FSD) 硬件——AI 5，预计将于 2025 年下半年投产。FSD 硬件计算平台采用两颗 SoC 芯片，以双系统设计提升自动驾驶功能安全冗余。特斯拉构建了神经网络编译器与链接器，以最大化计算资源利用率、吞吐量，并最小化延迟。通过芯片自研，特斯拉能够实现硬件方案的持续快速迭代，与软件算法进行更好的整合，从而实现更优的系统性能。

4) 数据端：高效自动标注+构建仿真场景，实现数据驱动训练。2020 年，特斯拉开始研发并使用数据自动标注系统，能够在 12 小时内自动标注一万个驾驶旅程，可抵充 500 万个小时的人工标注工作，极大提高了标注效率。仿真模拟则可以提供现实世界中难以获得或是难以标记的数据，从而加速 FSD 能力的训练，赋能模型迭代。结合真实数据和标签，以及仿真和手动校准的数据，特斯拉形成综合训练数据集，用于训练车端的在线模型，涉及网络占用、车道线和障碍物检测以及规划算法，形成闭环的数据流，实现自动驾驶系统的持续优化。

FSD 技术端快速进化，V12 为首个端到端自动驾驶系统，能够模拟人类驾驶行为，实现感知决策一体化。特斯拉 FSD v12 于 2023 年底推出，采用端到端大

模型，消除了自动驾驶系统的感知和定位、决策和规划、控制和执行之间的断面，将三大模块合在一起，形成了一个大的神经网络，直接从原始传感器数据到车辆操控指令，简化了信息传递过程，因而减少了延迟和误差，提高了系统的敏捷性和准确性。特斯拉 FSD 快速进化，V13 即将 10 月推出，有望于 2025 年 Q1 进入中国和欧洲，智能驾驶拐点已至。

“We, Robot” 发布会即将开幕，有望成为智驾板块强劲催化。 特斯拉将于北京时间 10 月 11 日在美国洛杉矶发布新品，活动主题口号为 “We, Robot”。特斯拉 Robotaxi 即将正式推出，有望成为特斯拉发展历程重要里程碑，并与 FSD V13 发布形成共振，共同成为板块强劲催化。

本篇报告与市场不同之处：

- 1) 从算法端、算力端、芯片端、数据端四个层面出发，对 FSD 系统底层技术原理进行全面深度拆解；
- 2) 结合 FSD V12 系统发展路径，对感知端到端到 One Model 端到端的技术架构演进趋势展开研究，探析自动驾驶领域全新发展路径；
- 3) 对特斯拉 FSD 的未来发展路径进行分析，判断特斯拉自动驾驶的重要技术发展节点，并对未来智驾板块的潜在事件催化进行展望。

2 概览：软硬件全链路闭环 全栈自研构筑强大技术基石

2.1 发展历程：软硬件持续迭代 由外部合作到全栈自研

特斯拉自动驾驶系统发展始于 2013 年，初期采用外部合作方式。2013 年 9 月，马斯克在推特上首次提到 AP（Autopilot 系统），表示特斯拉正在进行自动驾驶领域的探索。2014 年 10 月，特斯拉与视觉处理芯片独角兽公司 Mobileye 进行合作，正式推出第一代 Autopilot 硬件（HW1.0），率先搭载于 Model S。在未来一年多的时间里，特斯拉通过 OTA 不断更新固件，使车辆获得更完善的驾驶辅助或自动驾驶功能。2016 年 10 月，特斯拉推出第二代 Autopilot 硬件（HW2.0），采用了英伟达的 DRIVE PX2 平台，硬件平台进一步升级。

图1：2013 年 9 月，马斯克在 Twitter 中首次提到 AP



Elon Musk
@elonmusk

Intense effort underway at Tesla to develop a practical autopilot system for Model S

6:00 AM · Sep 19, 2013 · Twitter Web Client

资料来源：界面新闻，民生证券研究院

2019 年 4 月，特斯拉推出 HW3.0，正式开启全栈自研。HW3.0 放弃了英伟达的 DRIVE PX2 平台，转而采用特斯拉全栈自研的 FSD 芯片。2020 年 10 月，特斯拉小范围推送 FSD Beta，对 Autopilot 基础架构进行了重大重写。2021 年 7 月，特斯拉开始推送 FSD Beta V9，该版本采用纯视觉自动驾驶方案，摒弃了传统的毫米波雷达和超声波雷达，是特斯拉在自动驾驶技术的重要发展节点。2024 年 1 月，特斯拉 FSD V12 正式向用户推送，将城市街道驾驶堆栈升级为端到端神经网络，该神经网络由数百万个视频片段训练而成，取代了超过 30 万行的 C++ 代码。2024 年 2 月，特斯拉 Model Y 迎来 HW4.0 自动辅助驾驶硬件升级，与 HW3.0 相比，HW4.0 算力提升 5 倍，在硬件设计上实现并行处理能力增强、内存管理优化和专用加速器集成等多项创新。从最初的辅助驾驶系统，到全栈自研自动驾驶技术，特斯拉持续引领智能驾驶技术发展浪潮。

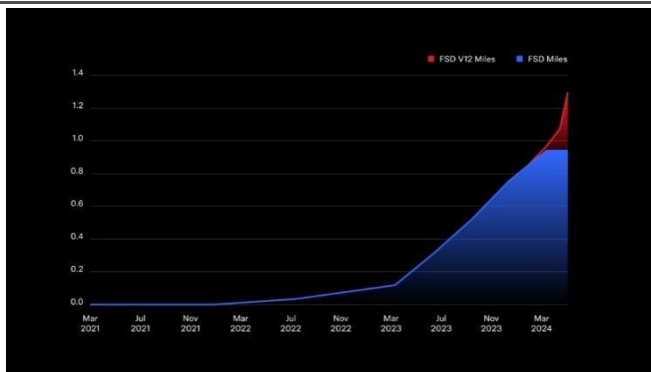
图2：特斯拉自动驾驶系统发展历程



资料来源：特斯拉官网，Teslarati，搜狐网等，民生证券研究院

FSD 累计行驶里程快速增长，商业化拐点已至。2024年4月12日，为降低FSD 体验门槛，吸引更多人订阅，特斯拉 FSD 推出单月付费优惠，价格从 199 美元/月调降 50%至 99 美元/月，FSD 买断价格维持 1.2 万美元不变。FSD V12 版本更新、订阅价格下降，共同驱动 FSD 累计行驶里程数快速增长，截至 2024 年 8 月 7 日，FSD 累计行驶里程达 1.6 亿公里，商业化拐点已至。

图3：FSD 累计行驶里程



资料来源：Teslarati，民生证券研究院

图4：FSD 订阅价格降至 99 美元/月



资料来源：阿尔法工厂研究院，民生证券研究院

2.2 系统架构：软硬件全链路闭环 实现高度集成

FSD 是一套包含感知、规控、执行在内的全链路自动驾驶软硬件架构，在算法、算力、数据、芯片等层面实现了高度集成。FSD 架构在 Tesla AI Day 2022 完整提出，核心组件包括规划 (Planning)、神经网络 (Neural Networks)、训练数据 (Training Data)、训练基础设施 (Training Infra)、AI 编译与推理 (AI

Compiler & Inference) 等。**FSD 核心组件在算法端、算力端、数据端和硬件端展现出强大的整合能力：**

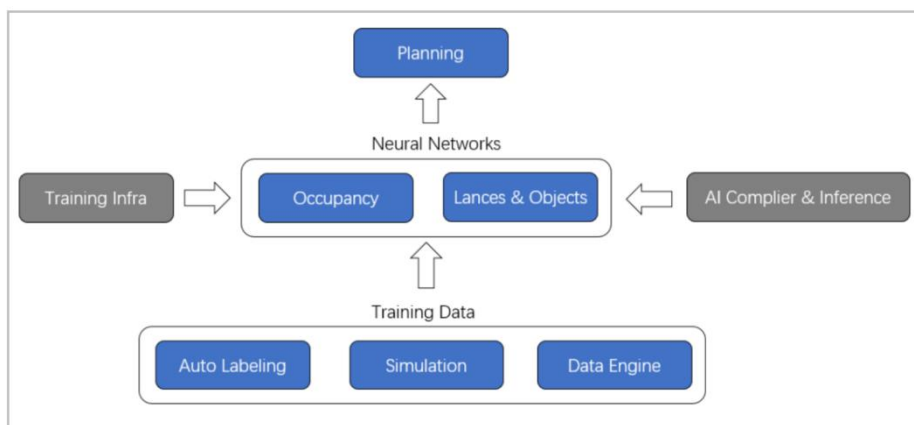
1) 算法端，规划组件专注于解决复杂的多物体关联路径规划问题，通过精确处理自我车辆与周围所有对象的行进轨迹，为汽车提供执行动作的指导。同时，**神经网络**则利用视频流等信息，输出车辆的运动学状态，如位置、速度、加速度等，以实现精确控制。

2) 算力端，训练基础设施提供了强大的计算支持，包括 CPU、GPU 和神经网络加速器单元 (Neural Network Accelerator)。这些硬件资源通过 AI 编译器的优化，能够高效地支持神经网络所需的新操作，并映射到最合适的硬件上，从而提升整体的计算效率。

3) 芯片端，AI 编译与推理组件确保了神经网络能够在计算机上高效运行。通过将神经网络的执行分配到两个独立的芯片系统上，FSD 实现了高性能的并行计算，进一步提升了自动驾驶系统的响应速度和处理能力。

4) 数据端，训练数据通过 **4D 自动标注技术、模拟仿真和数据引擎**，实现了数据的自动化和精准化处理，形成了一个闭环的数据系统，为算法的训练和优化提供了高质量的数据基础。

图5：特斯拉 FSD 系统架构



资料来源：Tesla AI Day 2022，民生证券研究院

3 算法端：创新算法优化 实现高效神经网络推理

感知规划算法全栈自研，实现从纯视觉信息输入到规划方案输出。1) 感知。特斯拉采用 BEV+Transformer 架构，将 2D 图像转化为对周围环境的准确 3D 感知。而后，特斯拉将该架构升级为 Occupancy Network，能够直接在向量空间产生体积占用，精准识别物体运动状态差异；2) 规划。特斯拉采用交互搜索框架，以任务分解的方式对一系列可能的行驶轨迹进行研究，实现对规划方案的实时评估。通过算法端全栈自研，特斯拉以低成本感知硬件进行高阶智驾能力输出，快速实现自动驾驶算法优化迭代。

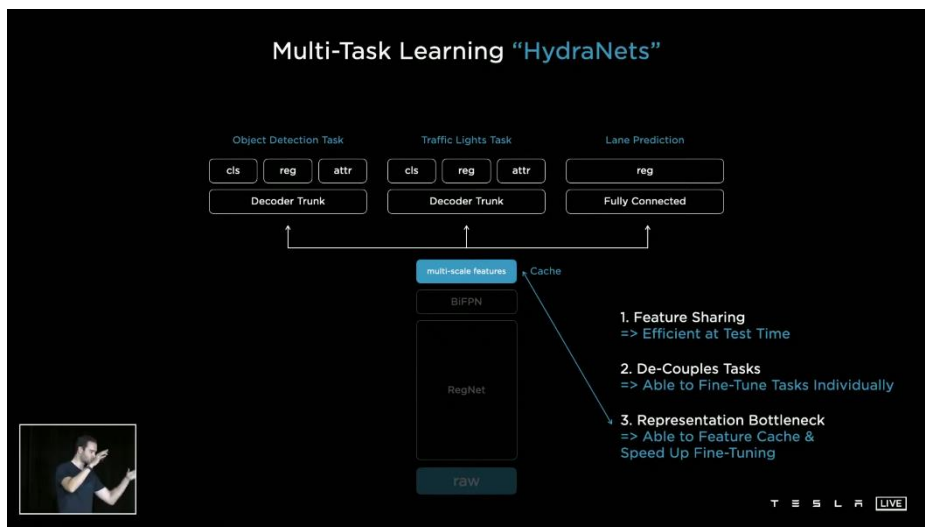
3.1 感知：先进感知技术，确保精准环境理解和物体识别

3.1.1 HydraNets：感知算法整体框架 以标准化流程处理视觉信息

特斯拉视觉感知系统采用 HydraNets 架构，以标准化流程进行视觉信息处理。1) **Input**。首先，系统接收来自摄像头的原始视觉数据，每个摄像头采集分辨率为 1280×960、36hz、12 bit 的视频图像；2) **Backbone**。用于提取图像特征，特斯拉主要采用由 RegNets (Residual Neural Networks, 残差神经网络) 组成的特征网络结构，能够通过不同层次的特征提取，捕捉图像的细节以及整体上下文信息；3) **Neck**。用于提取更复杂的特征，特斯拉采用 BiFPNs (Bi-directional Feature Pyramid Networks, 双向特征金字塔网络)，通过引入双向信息流，实现多个尺度之间信息交流共享，增强了对多尺度目标的检测性能；4) **Head**。由多个 Task Specific Heads 组成，负责最终的检测任务，如物体检测、交通信号和车道识别等。

HydraNets 架构能够实现特征共享、任务解耦与特征缓存。1) **特征共享**。HydraNets 通过共享 Backbone 和 BiFPNs 特征金字塔网络，减少了重复计算工作；2) **任务解耦**。每个子任务在主干网络上独立工作和微调，而不影响其他子任务，从而可以在不影响其他任务的情况下，单独对某个任务的数据集或头部架构进行更改和优化；3) **特征缓存**。HydraNets 可以缓存多尺度级别特征，在进行微调工作流程时，可以只使用这些缓存的特征来微调模型的头部，而无需重复计算整个网络。

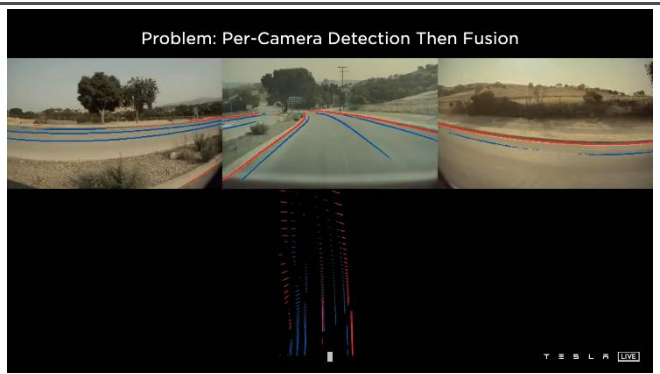
图6：特斯拉视觉感知系统算法采用 HydraNets 架构



资料来源：Tesla AI Day 2021，民生证券研究院

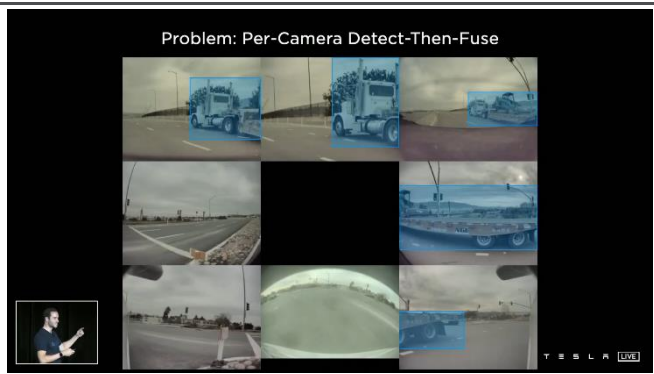
特斯拉早期方案为先在二维图像空间实现感知，再投射至三维向量空间。摄像头采集到的数据为 2D 图像级，与现实世界不在一个维度，因此要实现完全自动驾驶能力，则需要将二维数据变换至三维空间。特斯拉早期采取的方案是先在二维图像空间 (Image Space) 实现感知，将其投射至三维向量空间 (Vector Space)，再将所有摄像头的结果进行融合，但该方法需要对每个像素的信息进行精准深度预测，难度极大。此外，该方法无法对被遮挡的区域进行预测，因此如果物体横跨多个摄像头，且没有任何摄像头能够检测到物体全貌，则难以对多个摄像头采集到的信息进行准确融合，从而无法对物体进行准确预测。

图7：图像空间预测投射到向量空间后出现较大偏差



资料来源：Tesla AI Day 2021，民生证券研究院

图8：单相机检测无法解决物体横跨多相机的问题

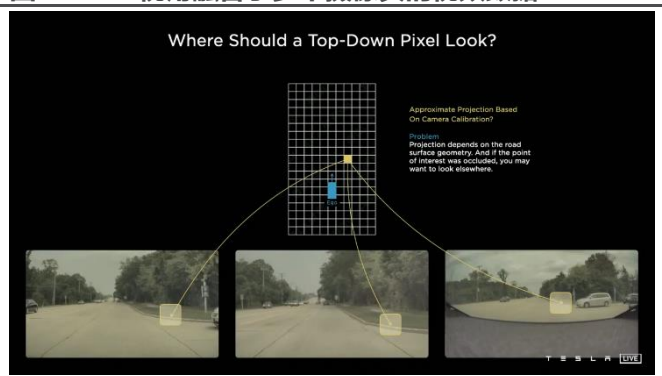


资料来源：Tesla AI Day 2021，民生证券研究院

特斯拉采用 BEV+Transformer 架构，将 2D 图像转化为 3D 感知。为了构建三维向量空间，网络需要进行物体深度信息输出。大部分自动驾驶公司的方案是

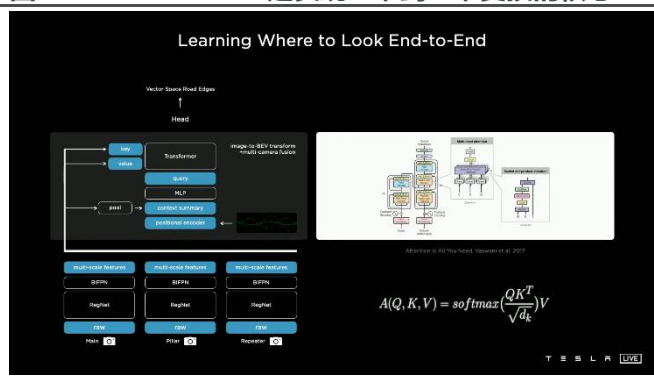
采用激光雷达、毫米波雷达等传感器来获取深度信息，与视觉感知结果进行融合，而特斯拉坚持使用纯视觉视频数据来计算深度信息，在网络结构中引入一层 BEV (Bird Eye's View, 鸟瞰图) 空间转换层，用以构建网络的空间理解能力。特斯拉采用“前融合”方案，将车身多个摄像头获得的视频数据直接进行融合，并采用同一套神经网络进行训练，实现特征从二维图像空间到三维向量空间的变换。Transformer 神经网络是实现二维到三维变换的核心，通过自注意力机制 (Self-Attention) 和多头注意力 (Multi-Head Attention) 模块，将每个相机对应的图像特征转换为 Key (键) 和 Value (值)，然后训练模型以查表的方式自行检索需要的特征用于预测，实现对车辆周围环境的准确感知。

图9: BEV 视角融合了多个摄像头的视频数据



资料来源: Tesla AI Day 2021, 民生证券研究院

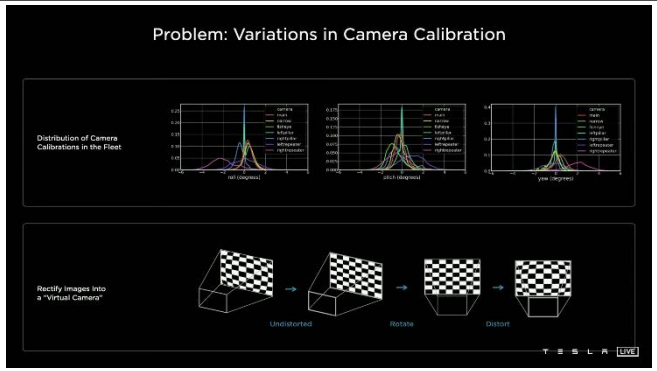
图10: Transformer 是实现二维到三维变换的核心



资料来源: Tesla AI Day 2021, 民生证券研究院

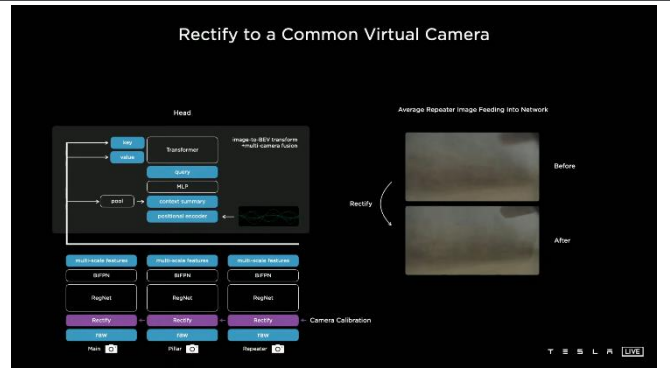
通过虚拟标准摄像头实现图像校准，消除外参误差。不同车辆由于摄像头安装外参的差异，可能导致采集的数据存在微小偏差，为此特斯拉在感知框架中加入了一层虚拟标准摄像头 (synthetic virtual camera)，引入摄像头标定外参将每辆车采集到的图像数据通过去畸变、旋转等方式处理后，统一映射到同一套虚拟标准摄像头坐标中，从而实现各摄像头原始数据的校准 (Rectify)，消除外参误差，确保数据一致性。

图11：通过图像校准解决摄像头采集数据偏差问题



资料来源：Tesla AI Day 2021，民生证券研究院

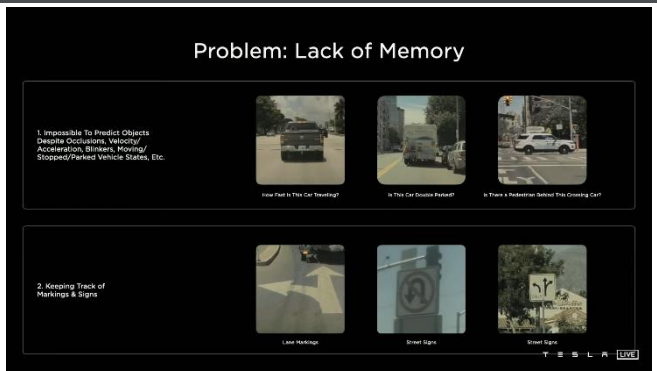
图12：加入虚拟标准摄像头以校准图像数据偏差



资料来源：Tesla AI Day 2021，民生证券研究院

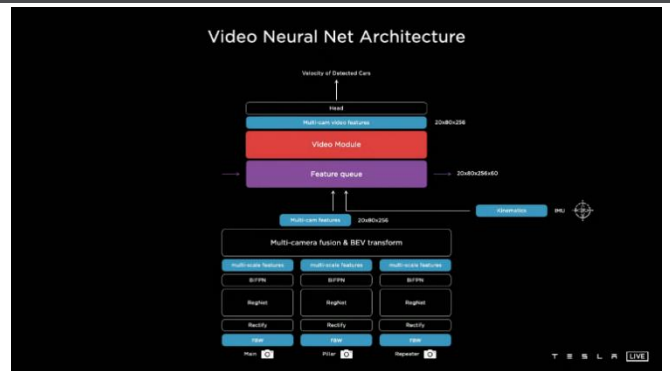
特斯拉引入时空序列特征层，以进一步提升环境感知准确性。在引入 BEV+Transformer 后，感知网络已经具备三维向量空间的感知能力，但仍是对瞬时的图像片段进行感知，只能根据当前时刻感知到的信息进行判断，感知不到世界空间内部分特征。特斯拉通过引入时空序列特征层，使得感知网络拥有类似于司机的短时记忆，可以对当前时刻的场景做出判断，并根据一段时间内的数据特征推演出目前场景下的可能结果。

图13：感知网络仍是对瞬时图像片段进行感知



资料来源：Tesla AI Day 2021，民生证券研究院

图14：特斯拉引入时空序列特征层



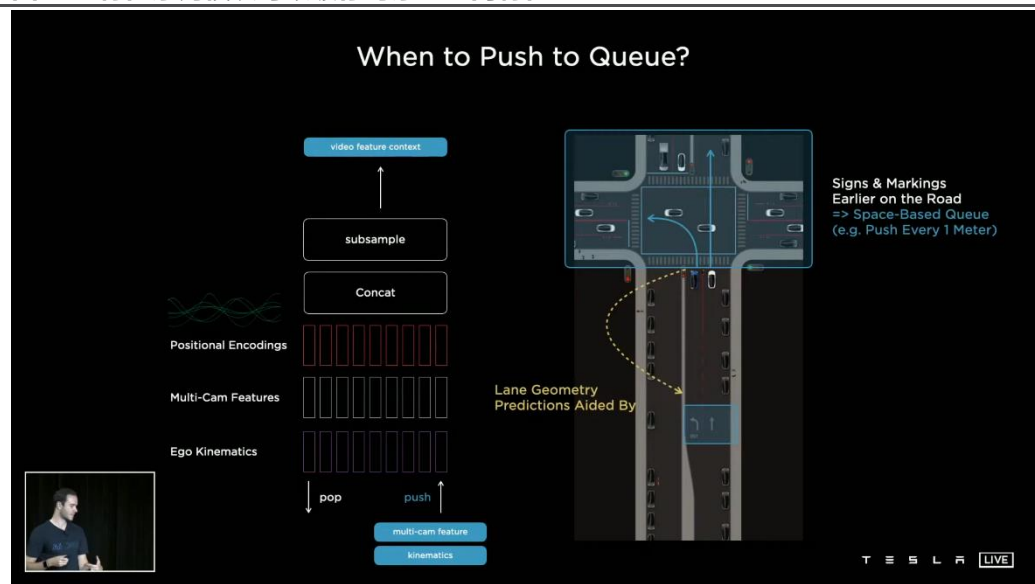
资料来源：Tesla AI Day 2021，民生证券研究院

时空序列特征层主要包括两部分：

1) 特征队列模块 (Feature Queue)，用来缓存时序与空间特征。其中，时序特征队列每过 27ms 将一个特征加入队列，可以稳定感知结果的输出，比如运动过程中发生的目标遮挡，模型可以找到目标被遮挡前的特征来预测感知结果；空间特征队列每行驶一定固定距离，将一个特征加入队列，用于等红绿灯一类需要长时间静止等待的状态，在该状态下一段时间之前的时序特征队列中的特征会出队而丢失，因此需要用空间特征队列来记住一段距离之前路面的箭头或是路边的标

牌等交通标志信息;

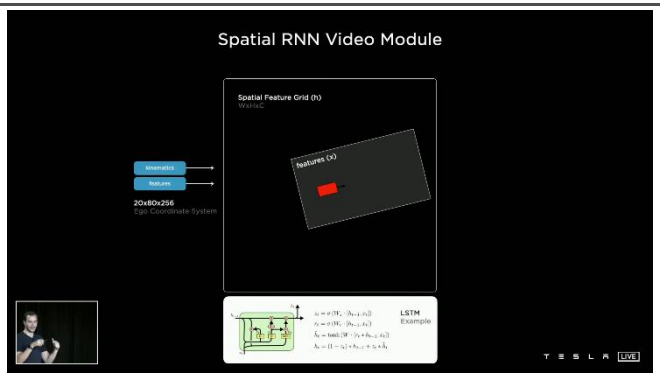
图15: 特征序列模块可以缓存时序与空间特征



资料来源: Tesla AI Day 2021, 民生证券研究院

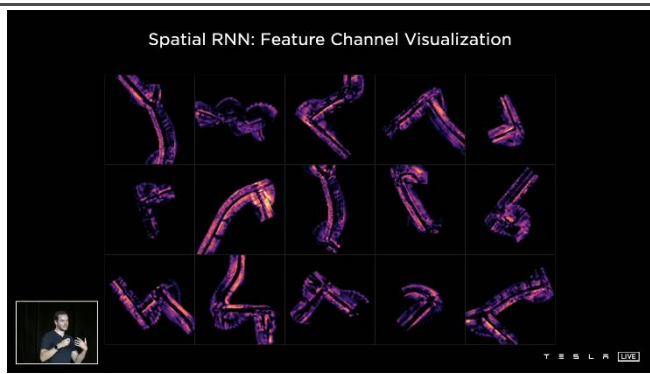
2) 视频模块 (Video Queue) , 用来整合时序上的信息。 特斯拉使用 RNN 结构作为视频模块, 命名为空间 RNN 模块 (Spatial RNN Module) 。车辆在二维平面上前进, 可以将隐状态 (Hidden State) 组织成一个二维网格。当车辆前进时, 只更新网格上与车辆当前视野相关的部分, 同时使用车辆的运动学状态以及隐特征 (Hidden Features) 更新车辆位置。空间 RNN 的隐状态可包含多个通道, 每个通道可以跟踪道路的不同方面, 如道路中心、边缘、标线等, 网格可以同时处理多种类型的环境信息。网络可以根据当前的能见度选择性地更新隐藏状态, 如果某个区域被其他车辆遮挡, 网络可以选择不更新那个区域的状态, 直到能见度恢复。视频模块能够提升感知系统对于时序遮挡的鲁棒性、对于距离和目标移动速度估计的准确性。

图16: 隐状态可组织成二维网格



资料来源: Tesla AI Day 2021, 民生证券研究院

图17: 空间 RNN 的隐状态可包含多个通道



资料来源: Tesla AI Day 2021, 民生证券研究院

3.1.2 Occupancy Network: 对 HydraNets 进行重要改进 直接在向量空间产生体积占用

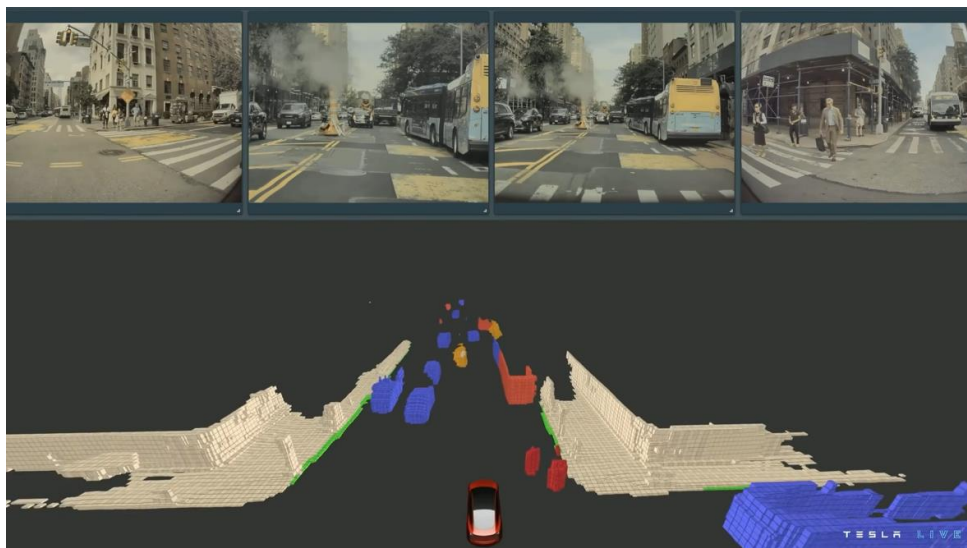
Occupancy Network 能够对长尾障碍物进行更好表达。在自动驾驶过程中,对常见障碍物如车辆、行人,可以通过 3D 物体检测的方式来估计其位置和大小,但还有更多长尾障碍物也会对行驶产生重要影响,例如: 1) 可变形的障碍物,如两节的挂车,不适合用 3D bounding box 来准确表示; 2) 异形障碍物,如翻倒的车辆,难以用传统的 3D 姿态来表示; 3) 不在已知类别中的障碍物,如路上的石子、垃圾等,无法进行分类。因此,我们希望能找到一种更好的表达来描述这些长尾障碍物,完整估计 3D 空间中每一个位置的占据情况 (Occupancy), 甚至是语义 (Semantics) 和运动情况 (Flow)。因此,特斯拉引入了 Occupancy Network (占用网络)。

Occupancy Network 是对 HydraNets 的重要改进,能够直接在向量空间产生体积占用。Occupancy Network 算法受到机器人领域中 occupancy grid mapping 启发,将 3D 空间分为大小一致的体素网格 (Grid cell), 然后判断每个 cell 是否被占用。Occupancy Network 以车辆摄像头产生的视频流作为输入,直接在向量空间产生单一统一的体积占用,对车辆周围 3D 位置被占用的概率进行预测,并可以通过视频信息对被遮挡物体情况进行即时预测。对于每个位置,Occupancy Network 能够产生一组语义,如路缘、汽车、行人和路上的碎片。Occupancy Network 通过高效的计算能力,在 10 毫秒内快速更新对周围环境的感知,同时提供物体尺寸的近似估计,支持动态及静态场景的全面预测,具有低延迟和低内存占用的特点。

Occupancy Network 能够精准识别物体运动状态差异,与传统目标检测网络相比优势明显。中一辆两节的公交车正在启动,其中蓝色表示运动的体素,红色表示静止的体素,Occupancy Network 精确捕捉到公交车第一节已经启动,而第二节还处于静止状态的细微差别,并可对公交车的精确曲率进行预测,而这一过程

对传统的目标检测网络来说非常复杂。

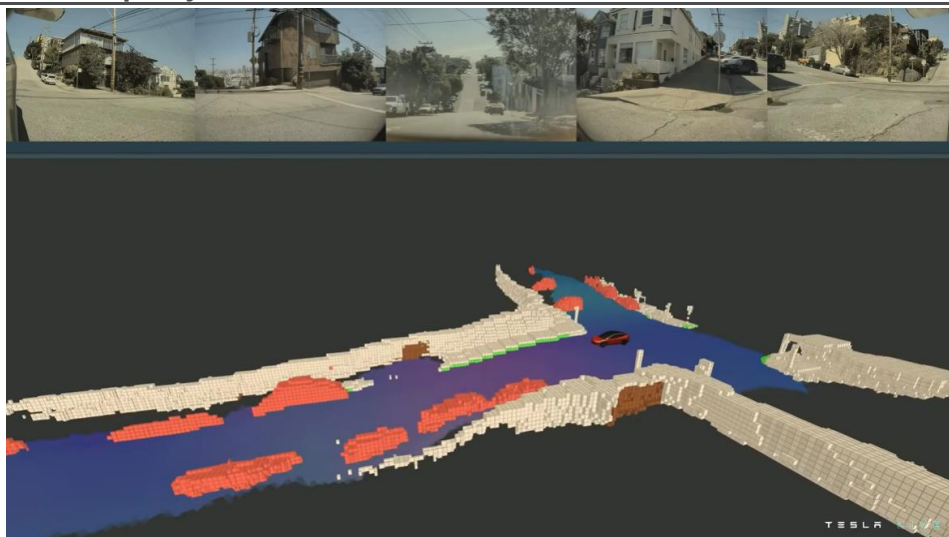
图18: Occupancy Network 对正在启动的两节公交车运动状态进行精准捕捉



资料来源: Tesla AI Day 2022, 民生证券研究院

Occupancy Network 能够生成可行驶表面, 增强复杂地形自动驾驶车辆控制能力。除体素网格外, Occupancy Network 还能够生成可行驶表面 (drivable surface), 可行驶表面具有 3D 几何形状与语义信息, 能够增强在多山、弯曲道路等复杂地形上, 自动驾驶车辆的控制能力。决策层可利用可行驶表面信息, 更好的进行加速、减速等运动决策。

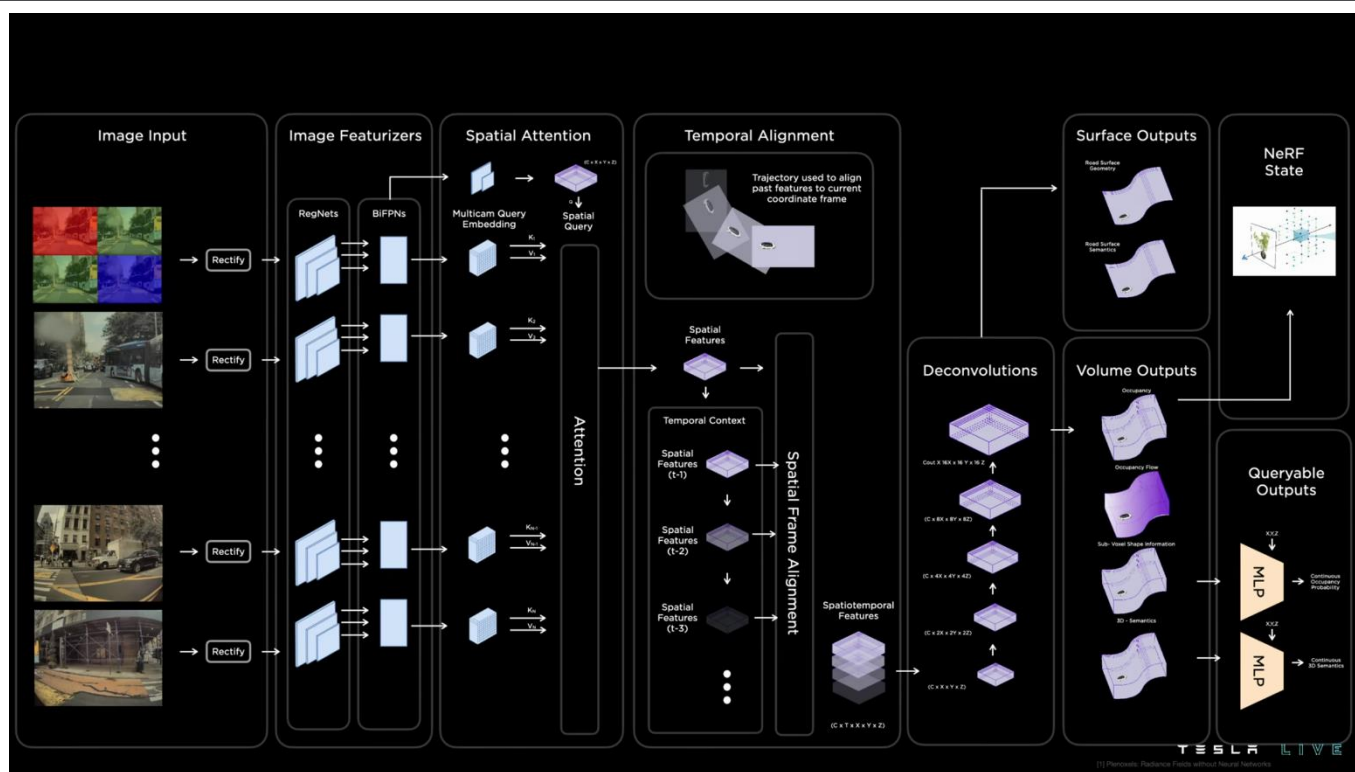
图19: Occupancy Network 能够生成可行驶表面



资料来源: Tesla AI Day 2022, 民生证券研究院

Occupancy Network 引入 Spatial Attention 机制。Occupancy Network 首先利用 RegNet 和 BiFPN 从多相机获取特征，然后采用带有 3D 空间位置信息的 Spatial Query，基于 Spatial Attention 注意力机制，实现对多个相机的 3D 空间位置信息和 2D 图像的信息融合，模型从中学习对应的特征关系，最终输出高维的空间特征。在进行特征融合后，基于反卷积 (Deconvolution) 的解码器会解码出每个 3D 空间位置的占用情况和占用流输出，形成固定大小的体素网络。模型的最后额外设计了一个隐式 Queryable MLP Decoder，输入任意坐标值(x,y,z)，可解码出该空间位置的信息，即 Occupancy, Semantics, Flow，打破了模型分辨率的限制。

图20: 基于 Attention 机制的 Occupancy Network 占用网络

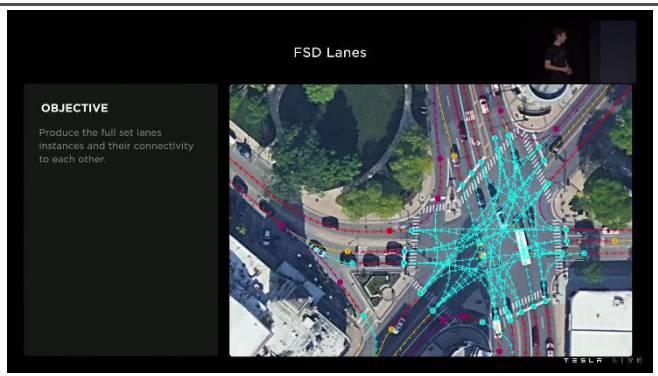


资料来源: Tesla AI Day 2022, 民生证券研究院

3.1.3 Lanes Network: 提供关键车道拓扑信息 优化轨迹与变道决策

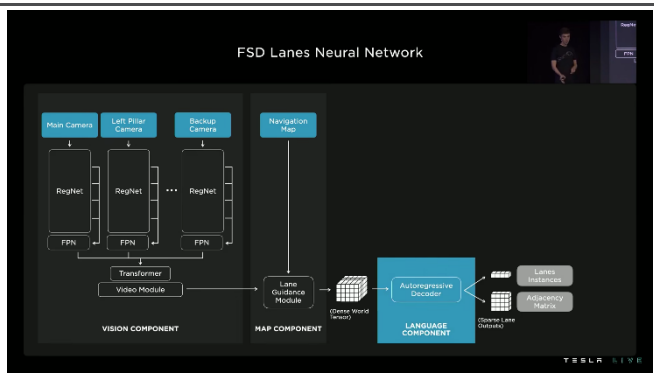
Lanes Network 旨在为自动驾驶车辆提供关键的车道拓扑信息，以优化轨迹规划和车道变换决策。初期，特斯拉将车道检测问题建模为图像空间即时分割任务，只能从几种不同类型的几何形状中进行车道预测，适用于高速公路等高度结构化的道路，但无法应对交叉路口等复杂、多样的道路拓扑场景。为此，特斯拉运用神经网络来预测车道与车道之间的连接性。车道检测神经网络由三部分组成，其中第一部分包括卷积层、注意力层与其他神经网络层，对车辆摄像头采集的视频信息进行编码，产生丰富的视觉表示。之后，特斯拉采用涵盖有关交叉口内车道拓扑、各条道路上的车道数等信息的低精度地图，对车道检测神经网络生成的丰富视觉表示进行增强，输出密集张量信息，并最终转化为车道及其连接性的信息。

图21: Lanes Network 旨在生成车道与连接信息



资料来源: Tesla AI Day 2022, 民生证券研究院

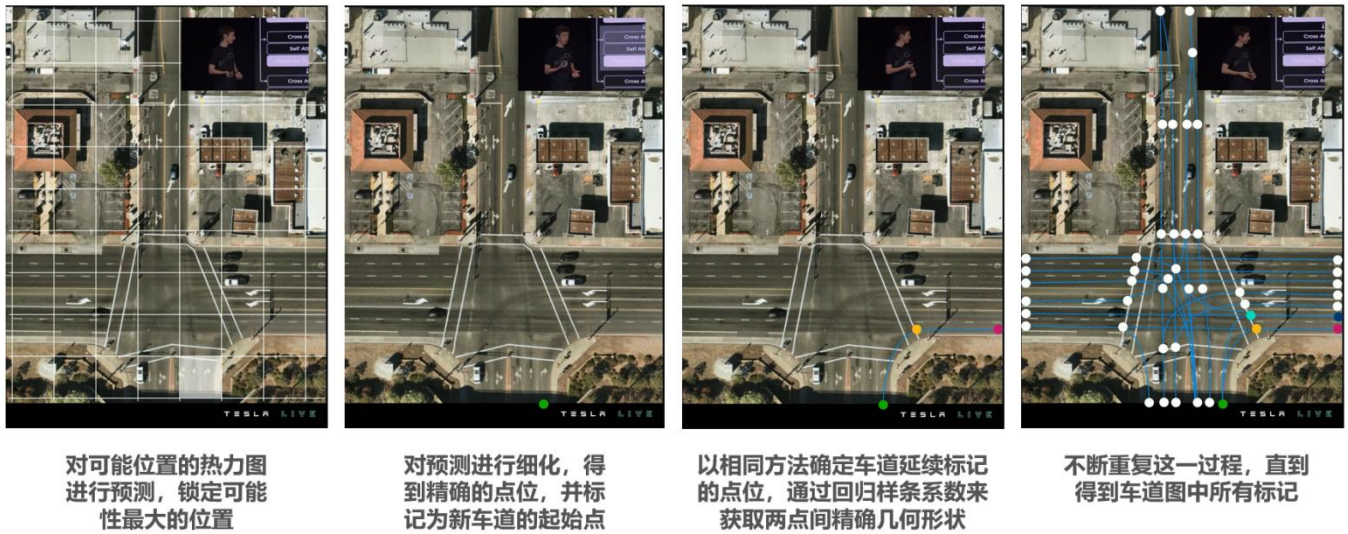
图22: 特斯拉采用低精度地图对视觉表示进行增强



资料来源: Tesla AI Day 2022, 民生证券研究院

特斯拉采用离散化处理+样条系数回归的方法，进行车道线预测。特斯拉首先将现实世界进行离散化处理，引入粗略划分的网格，然后对可能位置的热力图进行预测，锁定可能性最大的位置，并在此基础上对预测进行细化，以得到精确的点位。不同点位的标记类型不同，新车道的起点为起始标记，其他点为延续标记，通过回归样条系数来获取两点间的精确几何形状。之后，不断重复这一过程，直到得到车道图中所有标记。

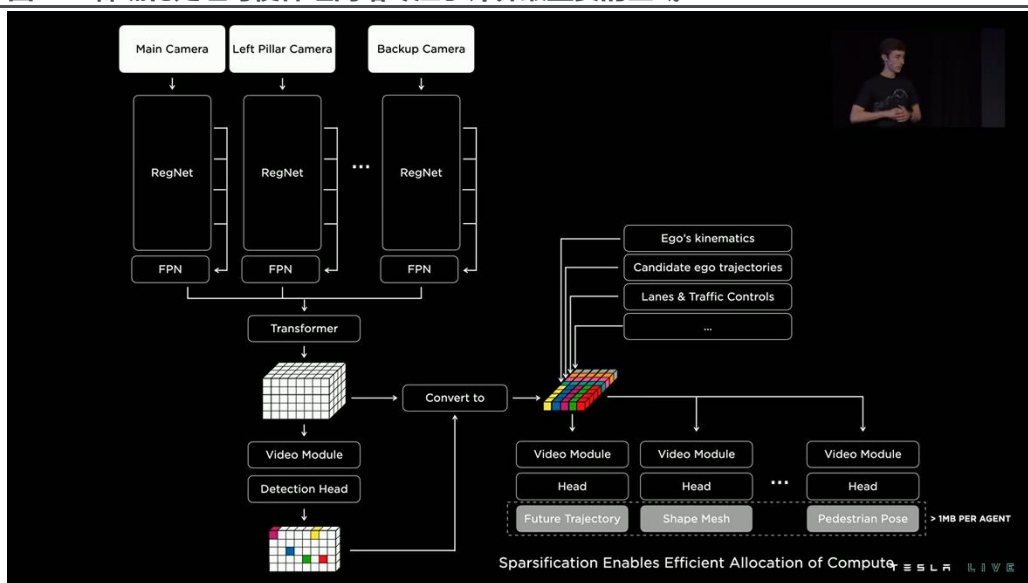
图23: Lanes Network 工作原理



资料来源: Tesla AI Day 2022, 民生证券研究院

神经网络可专注于计算最重要的区域，以较低延迟对道路上其他物体行为信息进行预测。自动驾驶系统神经网络的运行分为两步：1) 神经网络快速识别出 3D 空间中代理（即车辆或物体）的位置；2) 神经网络从这些位置提取张量，结合车辆速度、方向等其他数据，进行后续处理。通过稀疏化处理方式，神经网络可以专注于计算最重要的区域，以较低延迟实现卓越性能，对道路上其他物体行为信息进行预测。自动驾驶系统不仅可以理解环境中的位置和运动状态，还能够对其丰富的语义信息进行预测，从而在保证驾驶安全性的同时，实现类人的驾驶操作。

图24: 稀疏化处理可使神经网络专注于计算最重要的区域

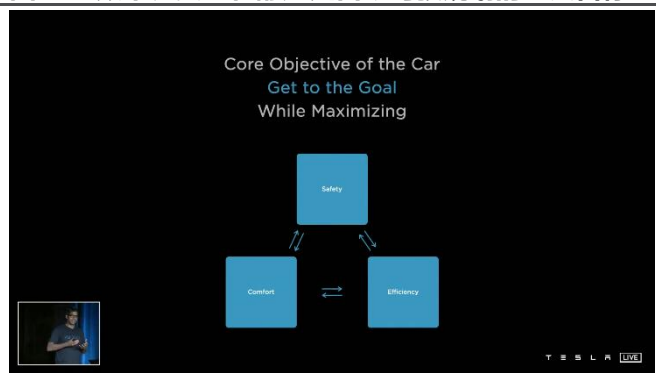


资料来源: Tesla AI Day 2022, 民生证券研究院

3.2 规划：采用交互搜索框架 快速选择最优方案

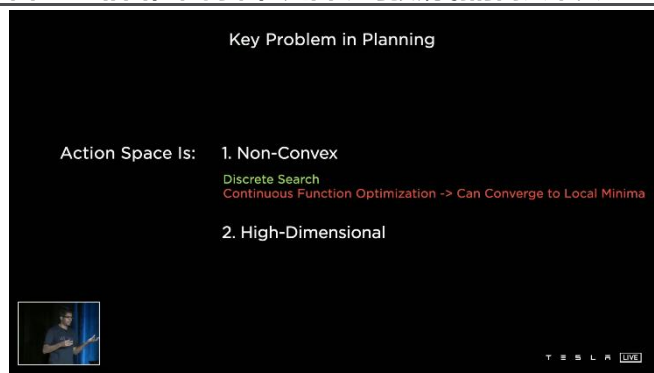
效率、安全和舒适是自动驾驶规划的三大目标，非凸性和高维性是自动驾驶规划的两大难点。与高速路况相比，城市路况更为复杂，需要应对临时施工路段、穿行车辆与行人。自动驾驶规划的难点主要体现在两个方面：1) 非凸性，行动空间是非凸的，意味着存在多个可能的解，但难以找到全局一致的解决方案，规划可能会陷入局部最小值；2) 高维性，车辆需要对未来 10-15 秒的行为做出规划，涉及位置、速度、加速度等大量参数，导致规划问题呈现出高维的特征。

图25：效率、安全和舒适是自动驾驶规划的三大目标



资料来源：Tesla AI Day 2021，民生证券研究院

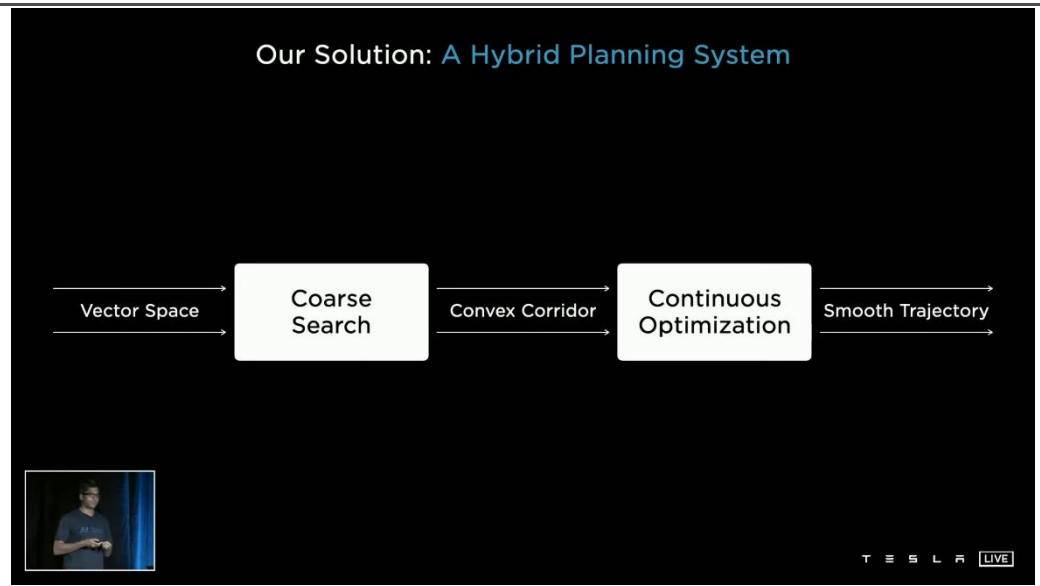
图26：非凸性和高维性是自动驾驶规划的两大难点



资料来源：Tesla AI Day 2021，民生证券研究院

特斯拉将规划问题进行分层分解，先采用离散搜索方法降低非凸性，缩小选择范围，再采用连续优化方法进行优化，得出最终解决方案。对于非凸问题，特斯拉采用离散搜索方法解决，因为离散搜索不会陷入局部最小值，而连续函数优化容易陷入局部最小值；对于高维问题，特斯拉采用连续优化方法解决，因为离散搜索方法不使用任何梯度信息，需要逐点评估每个点的好坏，效率较低，而连续优化方法利用基于梯度的方法，可以快速找到好的解决方案。

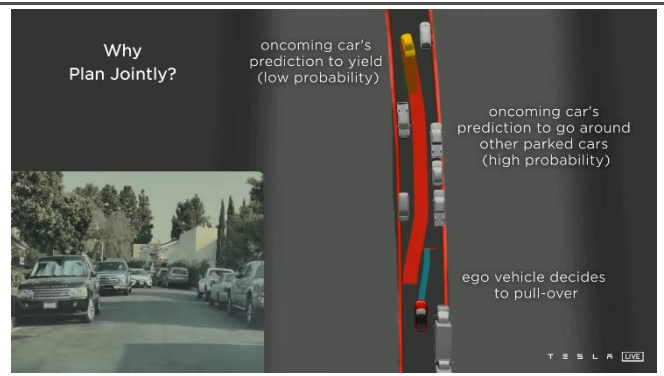
图27：特斯拉将规划问题进行分层分解



资料来源：Tesla AI Day 2021，民生证券研究院

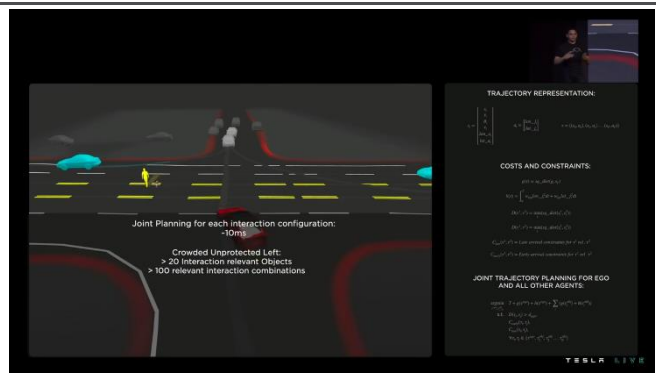
自动驾驶规划需要解决多代理联合轨迹规划问题，快速选出最优行驶方案。自动驾驶系统需要解决多代理联合轨迹规划的问题，考虑自己和所有其他车辆、行人的运动轨迹，对所有可能的行驶方案进行评估，快速选出最优行驶方案。评估过程至少需要 10 毫秒左右，而在面对繁忙路口等复杂场景时需要 50 毫秒。

图28：自动驾驶系统需要实现多代理联合轨迹规划



资料来源：Tesla AI Day 2021，民生证券研究院

图29：自动驾驶行驶方案评估至少需要 10 毫秒



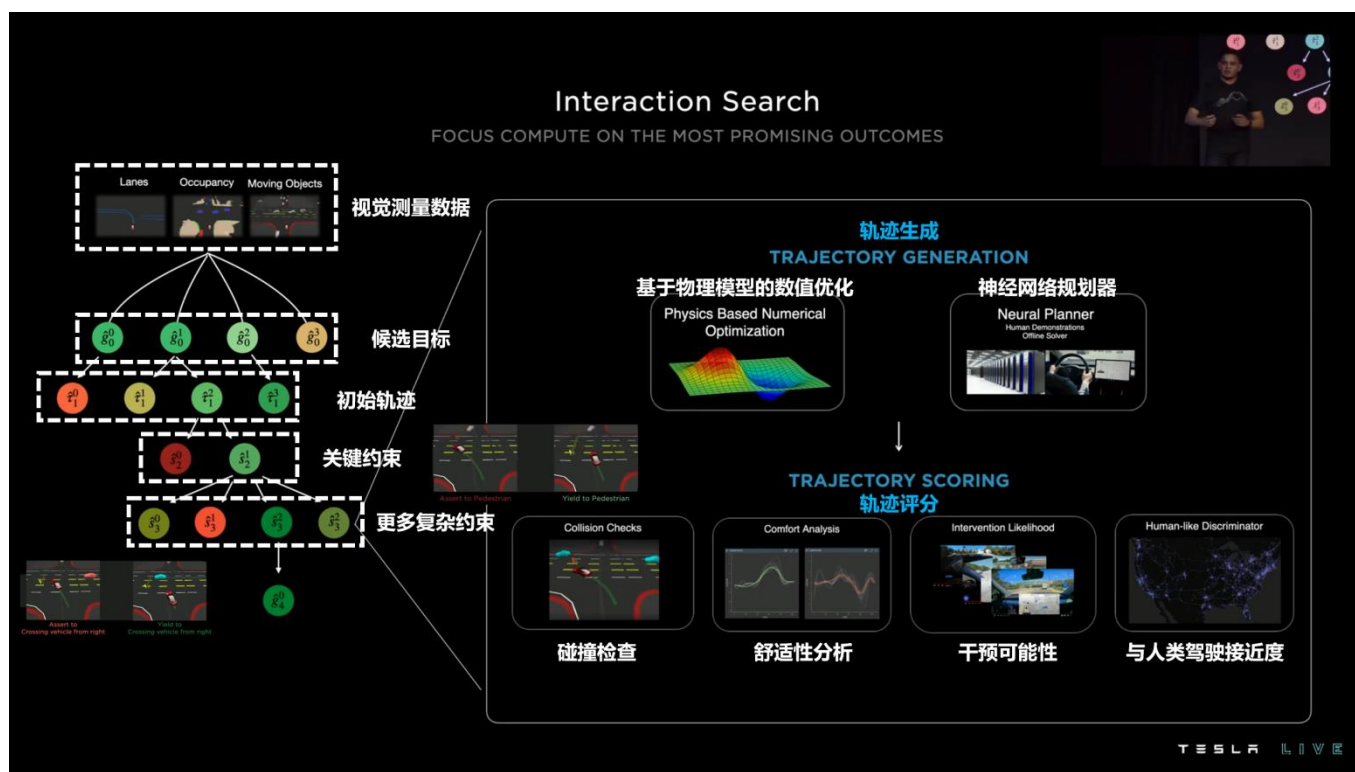
资料来源：Tesla AI Day 2022，民生证券研究院

为实现实时方案评估，特斯拉采用交互搜索框架，以任务分解的方式对一系列可能的行驶轨迹进行研究。1) 自动驾驶系统首先收集车道、障碍物和周围移动物体的视觉测量数据，这些数据被表示为稀疏抽象 (Sparse Abstraction) 和潜在特征 (Latent Features)。自动驾驶系统利用这些信息生成一组候选目标，使用经典优化方法与神经网络规划器来创建初始轨迹；2) 得到初始轨迹之后，特斯拉采

用递增式的方法，在关键约束的基础上，不断加入新的约束条件，利用较少约束下的最优解作为初值，逐步求解更复杂的优化问题；3) 构建轻量级可查询网络，该网络由人类驾驶数据与宽松时间限制下的计算数据进行训练，能够在 100 微秒内对规划轨迹进行评分，显著提升规划效率。

特斯拉主要从四个方面对规划轨迹进行评分，进行决策树剪枝：1) 碰撞检查，以确保路径安全；2) 舒适性分析，以评估乘客体验；3) 干预可能性，以评估候选路径导致人工接管的概率；4) 与人类驾驶接近度，基于人类驾驶数据，评估候选路径与人类驾驶员行为的接近程度。通过综合评估，特斯拉规划系统能够高效地筛选出最优路径，优化自动驾驶的决策过程。

图30：特斯拉采用交互搜索框架，实现实时方案评估



资料来源：Tesla AI Day 2022，民生证券研究院

4 算力端：Dojo 超级算力支撑 保持能效比优势

从 0 到 1 构建超级计算机系统，为远期算力提供强大支撑。特斯拉从算力芯片开始，完整构建 Dojo 超级计算机系统，以处理自动驾驶所需海量数据。2021 年 8 月，Dojo 在特斯拉首届 AI Day 上正式亮相，定位为超高速训练计算机，采用分布式计算架构设计，算力分为内核级、芯片级、格点级、集群级等四个层级，实现从训练节点到训练集群的完整构建。特斯拉从 0 到 1 构建超级计算机系统，旨在摆脱对英伟达 GPU 的依赖，为远期算力瓶颈进行前瞻布局。

4.1 Dojo 系统：特斯拉自研超算平台 研发进展迅猛

Dojo 于 2021 年 8 月正式亮相，研发进展迅猛。2019 年 4 月，马斯克在特斯拉自动驾驶日（Autonomous Day）上提到了 Dojo，称 Dojo 是“能够利用海量视频数据，做无人监管标注和训练的超级计算机”；2021 年 8 月，Dojo 超级计算机在特斯拉首届 AI Day 上首次正式亮相，特斯拉重点展示了 Dojo 的关键组成单元 D1 芯片，由特斯拉自主研发，用于神经网络训练；2022 年 9 月，特斯拉在第二届 AI Day 上展示了 Dojo 的最新进展，包括散热方案、Dojo POD 机柜集成架构等；2023 年 7 月，马斯克表示，特斯拉计划在 2024 年底前对 Dojo 项目投资超过 10 亿美元，自主研发超级计算机，以处理发展自动驾驶软件所需的海量数据，摆脱对英伟达 GPU 的依赖；2024 年 1 月，特斯拉将投资超过 5 亿美元，在纽约超级工厂建造一个巨型 Dojo 超级计算机集群。同时，马斯克表示特斯拉 2024 年在英伟达硬件上的投资将超过 5 亿美元，并将开始购买 AMD 的硬件。

图31：特斯拉 Dojo 发展历程

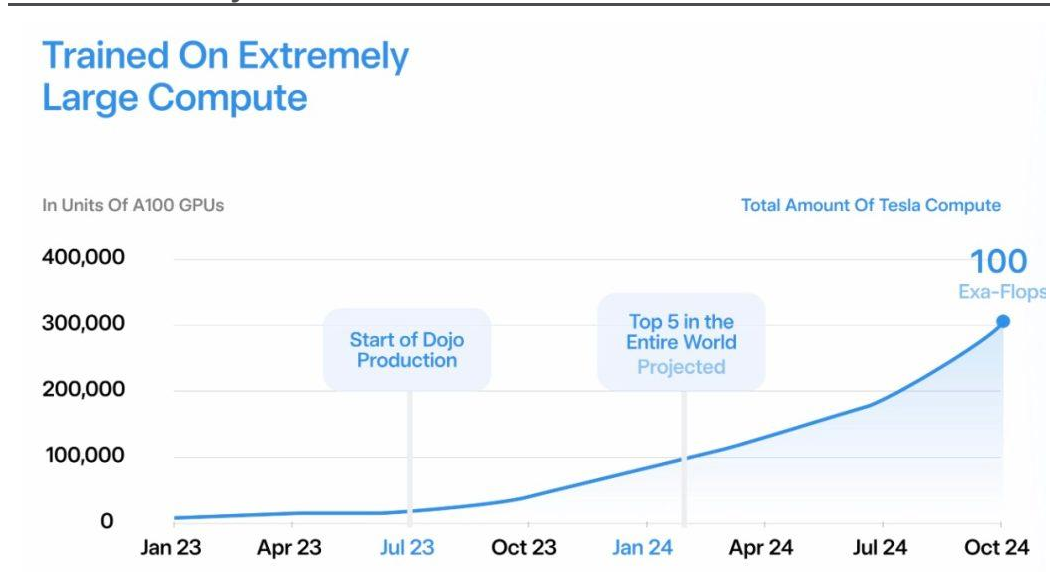


资料来源：腾讯科技，电动星球，界面新闻，electrek，民生证券研究院

Dojo 算力规划明确，计划于 2024 年 10 月达到 100 Exa-Flops。根据特斯

拉的规划：1) 2023 年 7 月，Dojo 进入投产阶段，拉开特斯拉算力集群快速建设阶段的帷幕；2) 2024 年 2 月，Dojo 将成为全球最强大的五台超级计算机之一；3) 2024 年 10 月，Dojo 的算力总规模达到 100Exa-Flops，相当于 30 万块英伟达 A100GPU 的算力总和。

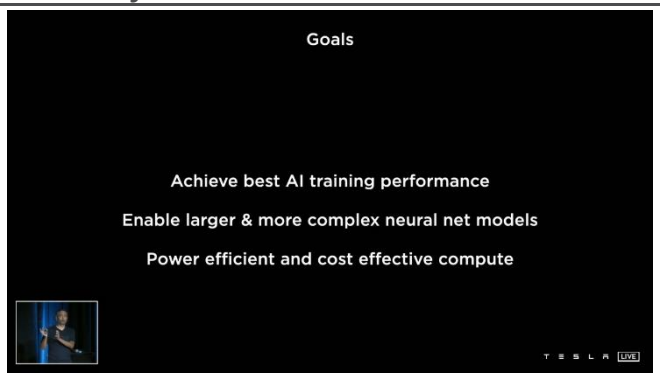
图32：特斯拉 Dojo 算力规划



资料来源：electrek，民生证券研究院

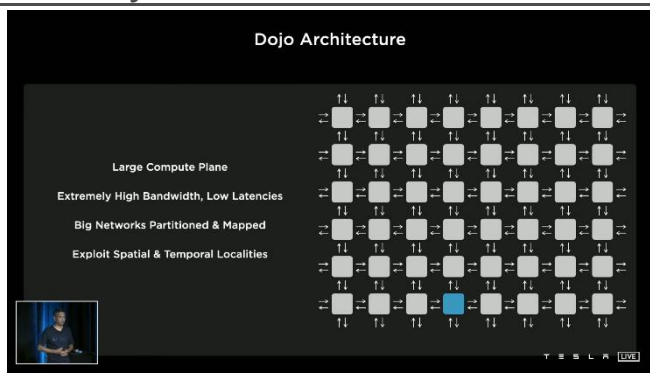
Dojo 定位为超高速训练计算机，采用分布式计算架构设计。2021 年 AI Day 上，特斯拉表示 Dojo 项目的三大目标为实现最佳 AI 训练性能、助力更大更复杂的神经网络模型训练、实现节能与成本效益。为此，特斯拉采用分布式计算架构设计，以强大的计算单元组成巨大计算平面，计算单元之间通过高带宽、低延迟的网络结构进行连接，形成二维网格结构。在进行大型神经网络训练时，Dojo 可以将神经网络分成若干小块，每个计算单元可以同时处理网络的一部分。神经网络编译器会考虑数据在时间和空间上的接近性，优化信息在计算单元之间的传递，从而提高工作效率。

图33: Dojo 超级计算机三大目标



资料来源: Tesla AI Day 2021, 民生证券研究院

图34: Dojo 系统二维网格结构



资料来源: Tesla AI Day 2021, 民生证券研究院

Dojo 采用分布式 2D 架构，算力分为内核级、芯片级、格点级、集群级等四个层级。按照层次划分，每 354 个 Dojo 核心组成一块 D1 芯片，而每 25 颗芯片组成一个训练模组，最后 120 个训练模组组成一组 ExaPOD 计算集群，共计 3000 颗 D1 芯片。**1) 内核级:** Dojo Core (训练节点)，为单个计算核心，64 位位宽，具有 4 个 8x8x4 的矩阵计算核心，2GHz 主频；**2) 芯片级:** D1，为单个芯片，核心数为 354，面积 645mm²；**3) 格点级:** Dojo Tile，为单个训练模组，每 5x5 个芯片组成一个训练模组；**4) 集群级:** ExaPOD，为特斯拉训练集群，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。

表1: Dojo 算力分为内核级、芯片级、格点级、集群级等四个层级

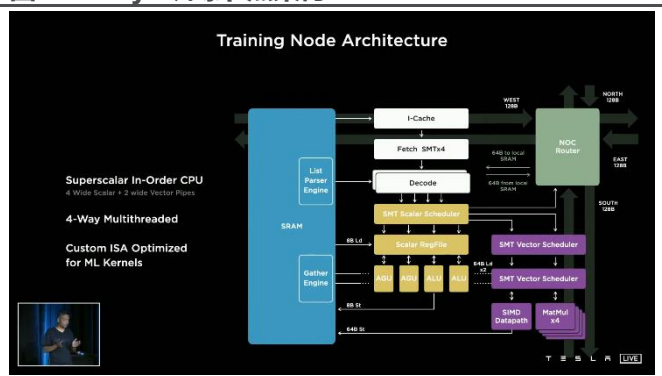
图示	层级	名称	片上 SRAM	算力	说明
	内核级	Dojo Core (训练节点)	1.25MB	1.024TFLOPS	单个计算核心，64 位位宽，具有 4 个 8x8x4 的矩阵计算核心，2GHz 主频
	芯片级	D1	440MB	362TFLOPS	单芯片，核心数为 354，面积 645mm ²
	格点级	Dojo Tile (训练瓦片)	11GB	9050TFLOPS	单个训练瓦片，每 5x5 个芯片组成一个训练瓦片
	集群级	ExaPOD	1320GB	1.1EFLOPS	特斯拉的训练集群，每 12 个训练模组组成一个机柜，每 10 个机柜组成 ExaPOD，共计 3000 个 D1 芯片

资料来源: 焉知汽车, 民生证券研究院

4.2 D1 芯片：基于存算一体架构 实现高效数据传输

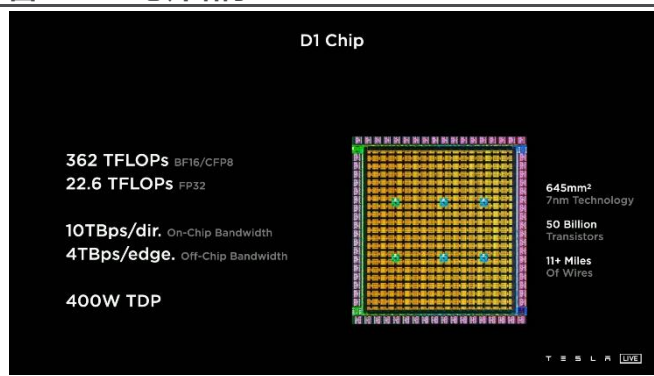
训练节点是 Dojo 最小的内部计算单元，基于存算一体架构设计。训练节点 (Training Node) 是 Dojo 最小的内部计算单元，大小关系到同步速度与硬件复杂度。降低延迟和提高带宽是训练节点的主要优化方向，每个训练节点基于存算一体架构设计，不仅具备向量计算和矩阵计算能力，还包含完整的取指、译码、执行部件，并以 2GHz 的频率运行。每个训练节点还配备了 1.25MB 的 SRAM 作为主存、而非缓存使用，能够以 400GB/s 的速度进行数据加载和 270GB/s 的速度进行数据存储。训练节点包含了一个 64 位超标量 CPU，针对矩阵乘法和向量 SIMD 进行了优化，支持 FP32、BFP16、CFP8 等多种浮点数格式运算。超标量 CPU 具备 4 路多线程能力，可以同时处理多个指令，以提高运行效率。同时，CPU 的指令集针对机器学习工作负载进行了优化，能够实现转置、聚集、链接遍历、广播等多种功能。

图35: Dojo 训练节点架构



资料来源: Tesla AI Day 2021, 民生证券研究院

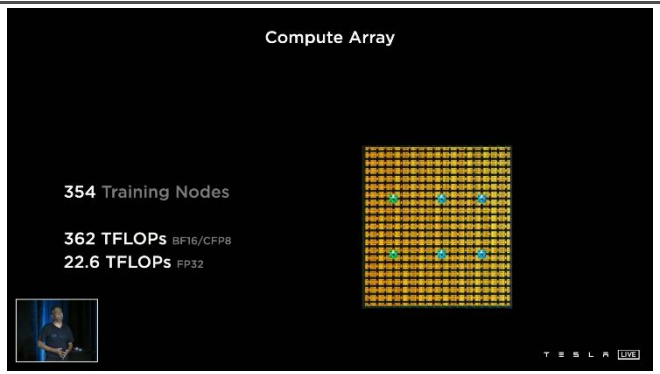
图36: D1 芯片结构



资料来源: Tesla AI Day 2021, 民生证券研究院

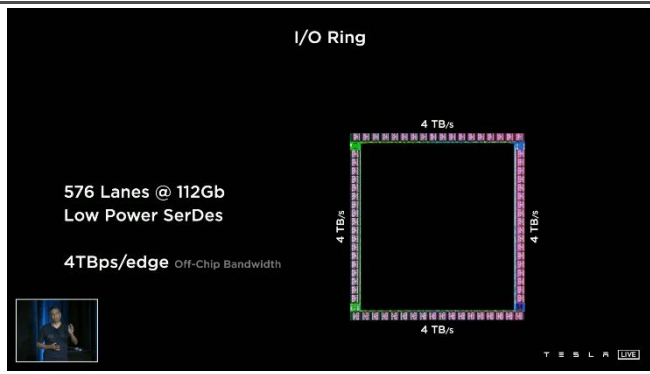
D1 芯片计算阵列由 54 个训练节点组成，能够实现高效数据传输。训练节点采用模块化设计，可以灵活组合扩展，形成更大的计算平面。D1 芯片采用 18x20 的网格布局，计算阵列由 354 个训练节点组成，采用台积电 7 纳米制造工艺，算力达 362TFLOPs，100%面积用于机器学习训练和带宽。在计算阵列周围，D1 芯片布置了 576 个高速低功耗串行器/解串器 (Serializer/Deserializer, 简称 SerDes)，I/O 带宽高达 10TB/s，大约是最先进的网络交换芯片的两倍，在进行机器学习训练时能够实现高效的数据传输，从而提升整体的计算性能。

图37: D1 芯片计算阵列



资料来源: Tesla AI Day 2021, 民生证券研究院

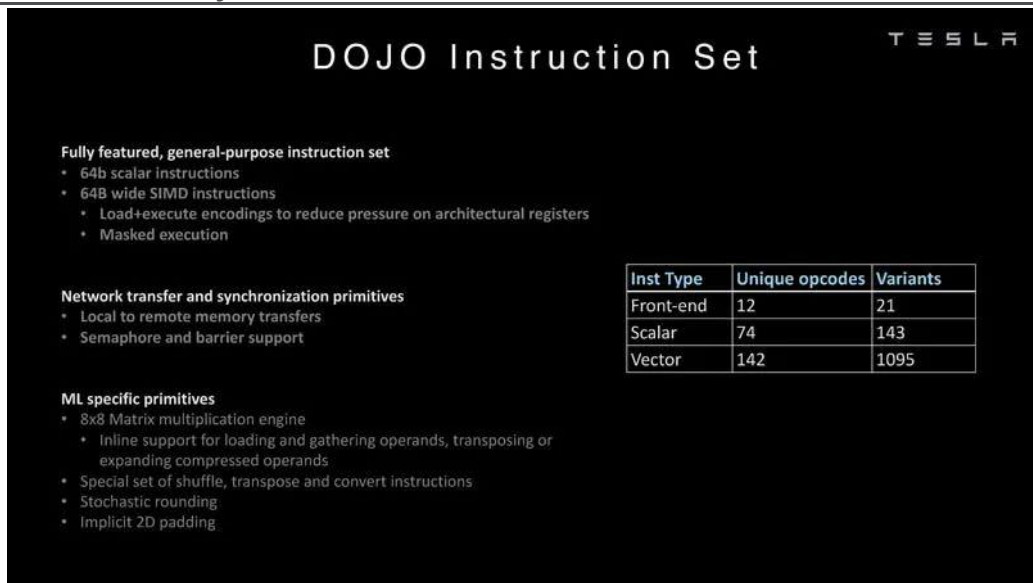
图38: D1 芯片串行器/解串器分布



资料来源: Tesla AI Day 2021, 民生证券研究院

D1 处理器的指令集设计通过扩展 RISC-V 架构, 提供多样化的计算格式和编译器支持。D1 处理器基于 RISC-V 架构的 ISA 进行了扩展, 以增强其计算能力。D1 核心支持 FP32 和 FP16 这两种标准的浮点计算格式, 并特别引入了 BFP16 格式, 以优化推理 (Inference) 过程。为了进一步提升性能, D1 处理器还集成了 8 位 CFP8 格式, 这种格式在降低精度的同时提高了计算吞吐量。Dojo 编译器的设计允许在尾数精度上进行动态调整, 从而覆盖更广的精度范围, 以适应不同的计算需求。此外, D1 处理器能够同时支持最多 16 种不同的矢量格式, 灵活性高, 能够显著提升处理器算力, 使其能够更高效的处理各种计算任务。

图39: 特斯拉 Dojo 指令集

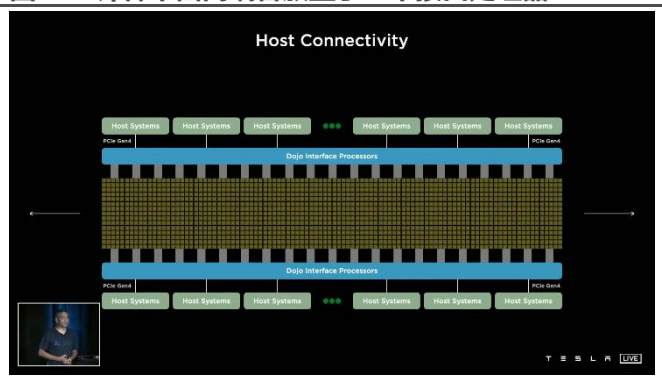


资料来源: 焉知汽车, 民生证券研究院

4.3 瓦片集群：由基本单元出发 构成大规模算力集群

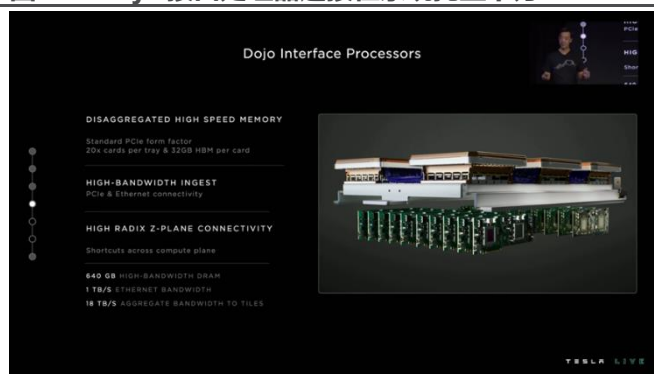
特斯拉将 50 万个训练节点集成成巨大的计算平面，以接口处理器提升训练效率。特斯拉把 50 万个训练节点进行组合，形成巨大的计算平面，该计算平面由 1500 个 D1 芯片密集连接而成。在计算平面的两端，特斯拉各放置了一个 Dojo 接口处理器，通过第四代 PCIe 接口与计算平面连接，能够实现数据中心主机与计算平台的高速通信。接口处理器为计算平面配备了高带宽 DRAM 共享内存，可以快速存储训练瓦片所需的大量数据；使用特斯拉定制协议 TTP 在整个加速器中进行通信，确保训练瓦片能够获得全部内存带宽，从而提升训练效率。

图40：计算平面两端各放置了一个接口处理器



资料来源：Tesla AI Day 2021，民生证券研究院

图41：Dojo 接口处理器连接在系统托盘下方



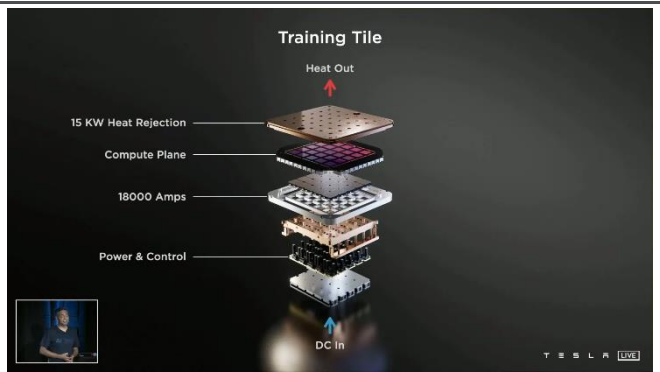
资料来源：Tesla AI Day 2021，民生证券研究院

训练瓦片是计算平面的基本单元，最终构成大规模训练集群：

1) **训练瓦片**。特斯拉创造性的设计了训练瓦片结构，作为计算平面的基本单元。一个训练瓦片上集成了 25 个 D1 芯片，采用高带宽连接器进行数据传输，可以提供 9 PFLOPS 的算力。训练瓦片具有巨大的输入/输出 (I/O) 带宽，可以快速传输和处理大量数据；

2) **系统托盘**。系统托盘上放置了 6 个训练瓦片，呈 2x3 矩阵式排布。一个系统托盘拥有 20 个接口处理器,以及 640GB 的高带宽 DRAM,能够实现机柜内部、不同机柜之间计算瓦片的无缝连接；

图42: 训练瓦片上集成了 25 个 D1 芯片



资料来源: Tesla AI Day 2021, 民生证券研究院

图43: 系统托盘上训练瓦片呈 2x3 矩阵式排布

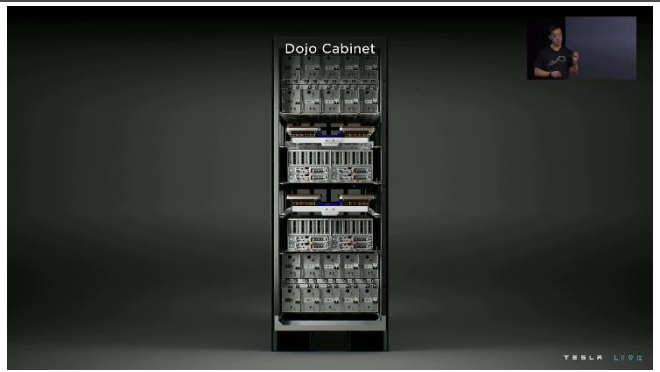


资料来源: Tesla AI Day 2021, 民生证券研究院

3) **训练机柜**。训练机柜中集成了两个系统托盘，总算力达 1000PFLOPS。机柜配置了冗余电源供应系统，以保证计算系统稳定运行；

4) **训练集群**。训练集群（ExaPOD）由 10 个训练机柜组成，总算力达 1 ExaFLOPS。

图44: Dojo 训练机柜中集成了两个系统托盘



资料来源: Tesla AI Day 2022, 民生证券研究院

图45: 系统托盘上训练瓦片呈 2x3 矩阵式排布

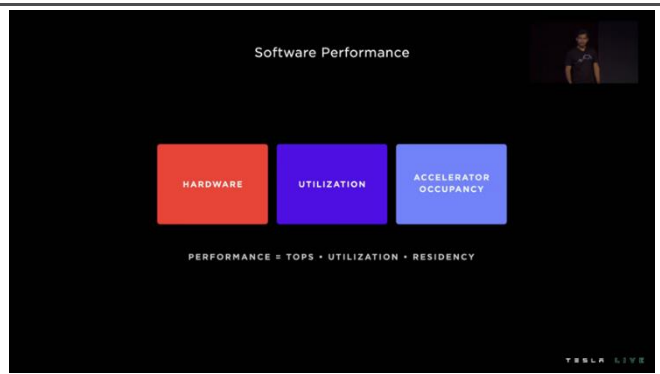


资料来源: Tesla AI Day 2021, 民生证券研究院

4.4 软件系统：采用高效同步机制 性能较 GPU 大幅提升

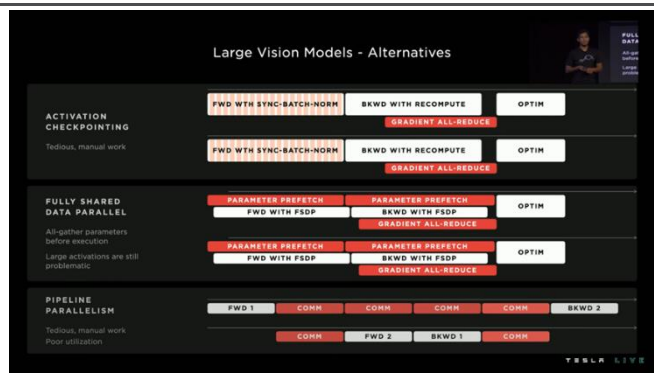
软件系统整体性能由硬件性能、利用率和加速器占用率共同决定。编译器的任务是从硬件中提取性能，而数据流的任务是确保数据能够以足够的吞吐量供给硬件，避免硬件空闲等待数据。当模型非常大时，单个加速器无法一次性处理所有数据，因为每个加速器能处理的批量大小通常比批量归一化所需的批量大小要小。为解决这一问题，需要在多个加速器上同步运行批量归一化，但这会带来前向传播通信延迟的问题。

图46：系统性能由硬件、利用率和加速器占用率决定



资料来源：Tesla AI Day 2022，民生证券研究院

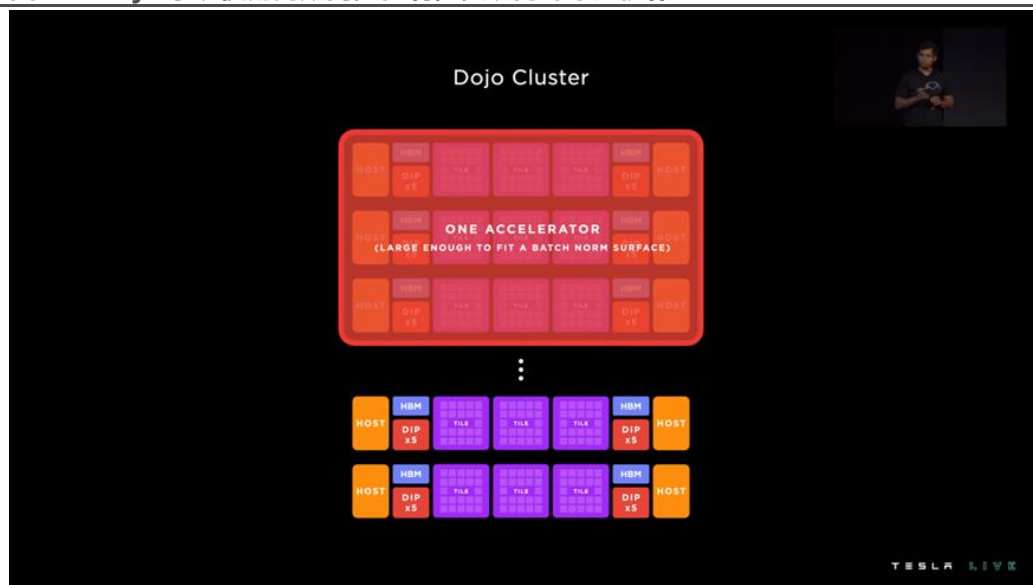
图47：多加速器运行批量归一化会导致前向传播延迟



资料来源：Tesla AI Day 2022，民生证券研究院

Dojo 使用高效同步机制，实现不同单元协同工作。Dojo 不仅可以加速模型中那些需要大量计算的部分，还可以提升那些受延迟或带宽限制的部分的性能，比如批量归一化或梯度合并、参数收集等操作。Dojo 可以进行灵活切割，以适配模型批量归一化所需大小，切割部分可作为独立加速器运行。Dojo 内部使用了高效的同步机制，实现不同计算单元之间的协调工作，形成单一可扩展的计算平面。数据被分散存储在快速的存储器中，并且在需要的时候才复制到相应的计算单元，高带宽可以助力数据快速复制。大多数模型可以在 Dojo 系统上直接运行，无需进行额外配置。

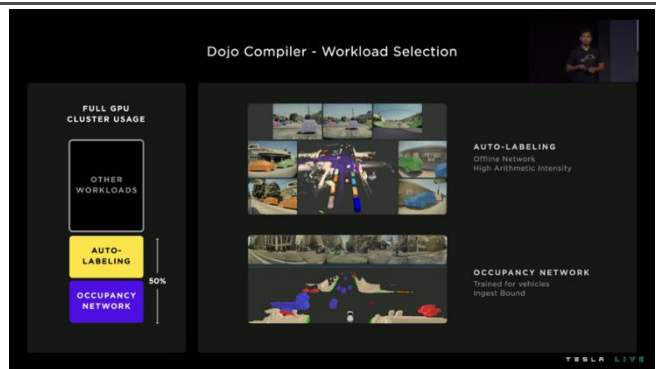
图48：Dojo 系统使用高效同步机制，实现内部单元协作



资料来源：Tesla AI Day 2022，民生证券研究院

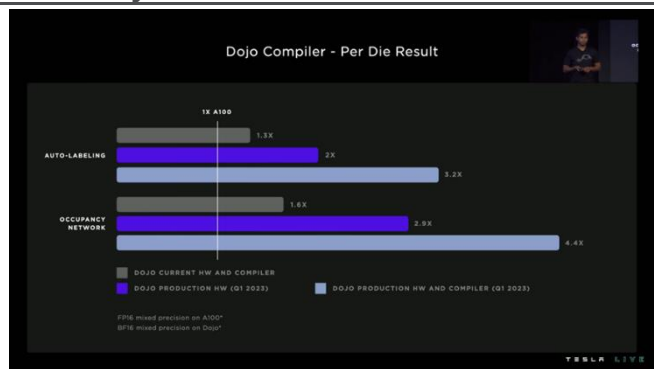
与 GPU 相比, Dojo 性能提升明显: 1) 批量归一化运行速度。在 Dojo 上, 一个批量均一化操作运行仅需 5 微秒, 而在 24 个 GPU 上运行需要 150 微秒。Dojo 运行速度具备数量级优势; 2) ResNet 50 模型运行速度。Dojo 的性能与 100 个 GPU 相当; 3) 自动标注网络运行速度。Dojo 性能已超越 A100 GPU, 采用新硬件后性能可以达到 A100 性能的两倍, 使用关键编译器优化后, 性能可达 A100 的三倍。

图49: 自动标注与占用网络占特斯拉 GPU 使用量一半



资料来源: Tesla AI Day 2022, 民生证券研究院

图50: Dojo 处理自动标注与占用网络速度大幅提升



资料来源: Tesla AI Day 2022, 民生证券研究院

5 芯片端：自研 FSD 芯片 集成高效计算能力

由合作迈向自研，实现高性能算力集成。 特斯拉自动驾驶硬件平台初期与 Mobileye、英伟达等合作，2019 年正式发布基于自研 FSD 芯片的 HW 3.0 系统，开始转向硬件平台全面自研，下一代全自动驾驶（FSD）硬件——AI 5，预计将于 2025 年下半年投产。FSD 硬件计算平台采用两颗 SoC 芯片，以双系统设计提升自动驾驶功能安全冗余。特斯拉构建了神经网络编译器与链接器，以最大化计算资源利用率、吞吐量，并最小化延迟。通过芯片自研，特斯拉能够实现硬件方案的持续快速迭代，与软件算法进行更好的整合，从而实现更优的系统性能。

5.1 硬件架构：由合作迈向自研 运算性能强大

5.1.1 发展历程：由合作迈向自研 AI 5 预计 2025 年下半年推出

HW1.0 向 HW4.0 快速迭代，硬件性能持续升级。 特斯拉的硬件平台（Hardware Platform，简称 HW）是指在其电动汽车中用于自动驾驶功能的计算和传感系统。特斯拉的硬件平台经历了几代的演变，每一版都代表着技术上的进步和对自动驾驶能力的提升：

1) HW1.0。 2014 年 10 月，特斯拉基于 Mobileye 芯片 Mobileye EyeQ3 发布第一代硬件 Hardware1.0；

2) HW2.0/HW2.5。 2016 年 10 月，特斯拉推出 HW2.0，采用了 NVIDIA 的 Drive PX2 平台，并配置 8 个摄像头+12 个远程超声波雷达+1 个前置毫米波雷达，在功能上实现辅助驾驶。而于 2017 年 8 月推出的 HW2.5，在 HW2.0 的基础上增加了额外的 NVIDIA Tegra Parker 芯片于增强计算能力；

3) HW3.0。 2019 年 4 月，特斯拉发布 HW 3.0 系统，特斯拉自研的 FSD 芯片首次亮相，整体算力达 144TOPS，标志着特斯拉开始全面掌握从芯片设计到软件开发的全栈技术；

4) HW4.0。 2024 年 2 月，推出的 HW4.0 搭载 FSD2.0 芯片，相较于 FSD1.0 芯片，采用更先进的制程技术，算力大幅提升 5 倍；同时，NNA 的数量从 2 个增加到 3 个，工作频率也从 2.0GHz 提升至 2.2GHz，这有助于更高效地处理深度学习任务，尤其是针对视觉数据的分析。

5) AI 5。 在 2024 年 6 月的特斯拉股东大会上，马斯克首次向公众介绍了下一代全自动驾驶（FSD）硬件——AI 5，预计将于 2025 年下半年投产。与现有的 HW4.0 版本相比，AI 5 在性能上将实现显著飞跃，预计整体性能提升约 10 倍，特别是在推理能力上，这一提升可能高达 50 倍，同时在整体能耗方面也比 HW4.0 提升了 4-5 倍。

表2：特斯拉自动驾驶硬件平台不同版本对比

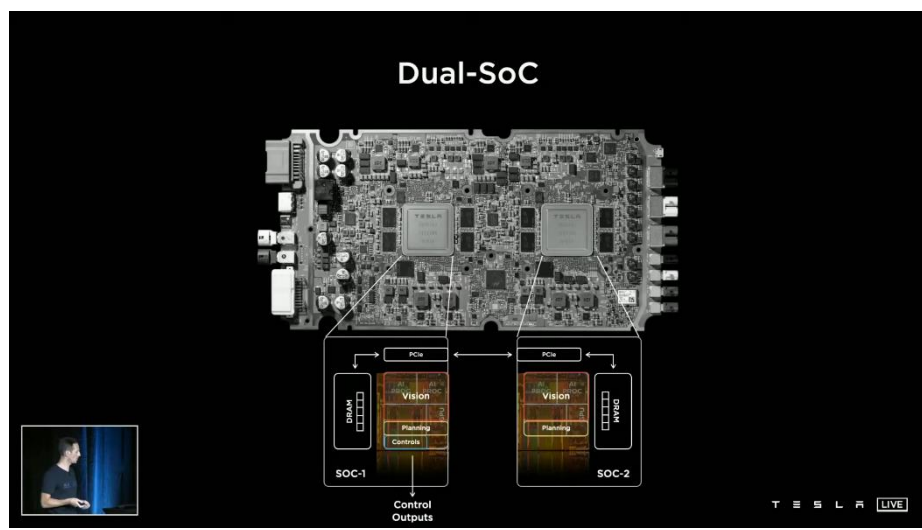
硬件版本	SOP	处理平台/主芯片	冗余控制	算力 (TOPS)	功耗 (W)	图像处理能力 (fps)
HW1.0	2014.10	Mobileye EyeQ3	无	0.256	25	36
HW2.0	2016.10	Nvidia Drive PX2	部分	20	250	110
HW2.5	2017.08	Nvidia Drive PX2+	完全	20	300	110
HW3.0	2019.04	Tesla FSD	完全	144	220	2,300
HW4.0	2024.02	Tesla FSD	完全	720	80	—
AI 5	2025 下半年	Tesla FSD	完全	7200	320-400	—

资料来源：Tesla 官网，汽车之家，Twitter，智能车参考，民生证券研究院

5.1.2 系统架构：双冗余 SoC 设计 集成高性能核心组件

FSD 硬件计算平台采用两颗 SoC 芯片，以双系统设计提升自动驾驶功能安全冗余。其中，SOC-1 作为主控单元，输出最终的控制指令，而 SOC-2 则提供扩展计算支持。作为两套完全独立的芯片系统，每套系统都配备了独立的 CPU、GPU、NNA（神经网络加速器）以及内存，两套系统中的一个作为主系统运行，另一个作为热备份，在主系统出现故障时立即接管，以实现冗余，提高安全性。双系统架构的另一个优势在于相互验证能力。在面对相同的驾驶情境时，两套独立系统会分别经过感知和规划算法（或一个完整的端到端算法）处理，得出的驾驶决策应当是一致的。这种设计提供了决策冗余保障，有效提升了自动驾驶系统的功能安全性。

图51：特斯拉 FSD 双芯片系统设计



资料来源：Tesla AI Day 2021，民生证券研究院

特斯拉的 FSD 芯片自 2016 年启动设计规划以来，经历了数年的发展和迭代。2017 年 12 月，特斯拉进行了 FSD 芯片的首次试产，在随后进行了必要的设计调整。2018 年 4 月，B0 样片投产，并在同年 7 月获得认证。2018 年 12 月，特斯

拉开始用新的硬件和软件对员工用车进行改装。2019年3月，特斯拉开始在 Model S、Model X 和 Model 3 车型上部署 FSD 芯片，用于 HW 3.0 系统，标志着 FSD 芯片开始得到批量化应用。

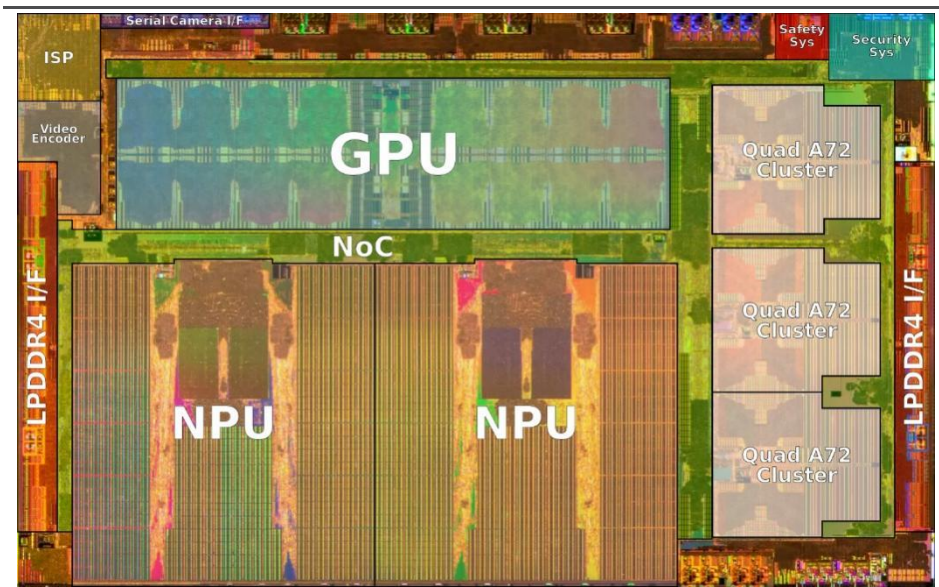
图52：特斯拉 FSD 芯片发展历程



资料来源：Tesla 官网，汽车 ECU 开发，民生证券研究院

FSD 芯片采用异构设计，集成了 CPU、GPU、NNA，和硬件加速器等多个处理单元。FSD 芯片采用三星 14nmFinFet 技术制造，支持单精度和双精度浮点运算，并配备了运行于 2133MHz 的 128bit LPDDR4 内存，展现出卓越的计算性能和高效的数据处理能力：1) **CPUs**。3 个四核 Cortex-A72 集群，共计 12 个 CPU 核心，运行频率为 2.2GHz；2) **GPU**。1 个 Mali G71 MP12 GPU，工作频率为 1GHz，支持 FP16 和 FP32 浮点运算；3) **2 个神经网络加速器 (NN Accelerator, 简称 NNA)**。用于深度学习推理，这是 FSD 芯片的核心部分，用于处理自动驾驶所需的大量视觉和传感器数据，运行频率为 2GHz。每个 NNA 配备了 32MB 的 SRAM 缓存，用于存储模型权重和加速数据访问。

图53：特斯拉 FSD 芯片架构



资料来源：Wikichip，民生证券研究院

FSD 2.0 通过扩大 CPU 核心数和增强 NPU 核心，提供更强大的数据处理和神经网络运算性能。2023 年 2 月，FSD2.0 芯片开始在汽车上部署，用于 HW 4.0。FSD2.0 芯片的设计与第一代非常相似，采用更密集的晶体管布局和优化电路设计，性能预计提升三倍以上。在 CPU 中，Cortex-A72 内核从 12 个增至 20 个，分布在 5 个四核集群中。此外，第二代芯片采用 3 个 NPU 核心，每个核心配备了 32MB SRAM 用于存储模型权重和激活，通过 96x96 的 MAC 网格实现每个周期 9216 个 MAC 和 18432 个操作的高效率运算。NPU 核心以 2.2 GHz 频率运行，使得单颗芯片算力达到 121.651 TOPS，显著提升了自动驾驶的数据处理和神经网络运算性能。

表3：FSD 芯片 1.0 和 2.0 性能对比

Model	FSD1	FSD2
CPU Cores	12(3*4) A72	20 (5*4)
CPU Frequency	2.2GHz	2.35GHz
GPU	Mali G71 MP12	Mali G71 MP12
GPU Frequency	1.0 GHz	—
NPUs	2	3
NPU Frequency	2.0 GHz	2.2 GHz
NPU Performance (Individual)	36.86 TOPS	40.55 TOPS
NPU Performance (Total in SoC)	73.7 TOPS	121.65TOPS
Memory	128-bit LPDDR4	128-bit GDDR6
Memory Speed	4266 MT/S	14000 MT/S
Memory Capacity	8 GB	16 GB
Memory Bandwidth	68.3 GB/s	224 GB/s

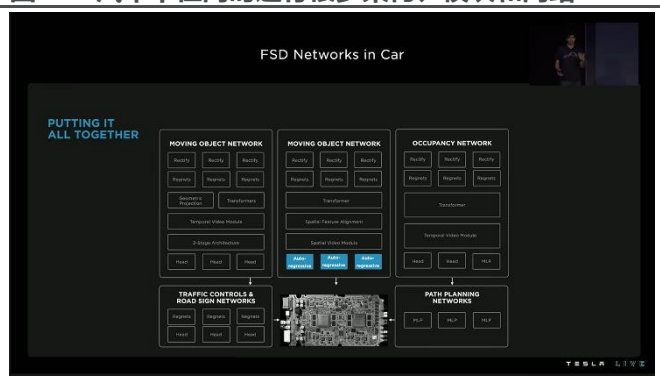
uArch	Cortex-A72	Cortex-A72
Manufacturing node	Samsung 14nm	Samsung —
TDP	36 Watts	—

资料来源：ADS 智库，民生证券研究院

5.2 编译推理：高效优化 最大化计算资源利用

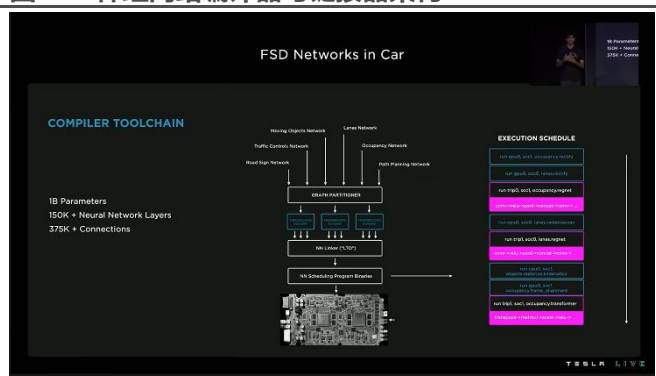
特斯拉构建了神经网络编译器与链接器，以最大化计算资源利用率、吞吐量，并最小化延迟。汽车中在同时运行很多架构、模块和网络，共有多达 10 亿个参数、1000 个神经网络信号。为了提升计算效率，特斯拉建立了：1) 神经网络编译器，用于接收复杂的神经网络图，然后将其分割为独立的子图，并分别进行编译；2) 神经网络链接器，用于链接各个编译好的部分，形成完整系统。计算过程需要进行离线优化，从而能够在有限的内存和带宽条件下完成计算。

图54：汽车中在同时运行很多架构、模块和网络



资料来源：Tesla AI Day 2022，民生证券研究院

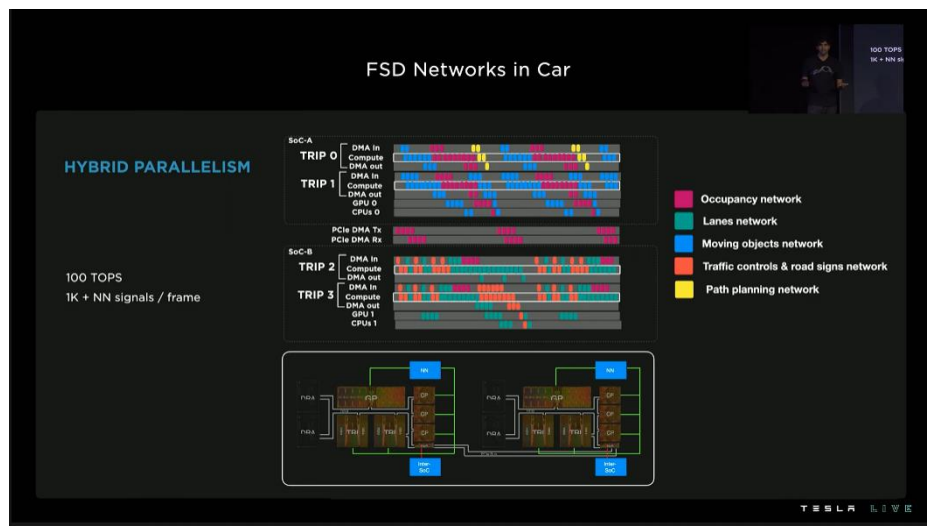
图55：神经网络编译器与链接器架构



资料来源：Tesla AI Day 2022，民生证券研究院

特斯拉设计了混合调度系统，以实现算力的充分利用。实现特斯拉设计了混合调度系统，能够在在一个 SOC 上实现异构调度，并在两个 SOC 之间进行分布式调度，以模型并行方式运行网络。特斯拉从软件所有层级进行优化，采用调整网络结构、改进编译器、在两个 SOC 之间采用低延迟高带宽的 RDMA 链接等多种方式，以实现 100 TOPS 算力的充分利用。

图56：特斯拉混合调度系统架构



资料来源：Tesla AI Day 2022，民生证券研究院

6 数据端：自动化数据闭环 优化数据驱动训练

高效自动标注+构建仿真场景，实现数据驱动训练。2020年，特斯拉开始研发并使用数据自动标注系统，能够在12小时内自动标注一万个驾驶旅程，可抵充500万个小时的人工标注工作，极大提高了标注效率。仿真模拟则可以提供现实世界中难以获得或是难以标记的数据，从而加速FSD能力的训练，赋能模型迭代。结合真实数据和标签，以及仿真和手动校准的数据，特斯拉形成综合训练数据集，用于训练车端的在线模型，涉及网络占用、车道线和障碍物检测以及规划算法，形成闭环的数据流，实现自动驾驶系统的持续优化。

6.1 自动标注：自动化4D标签生成 提升标注效率

从最初的外包第三方人工标注，到开发自动标注（Auto Labeling）系统，特斯拉实现了数据标注的高效化和规模化。在特斯拉的自动驾驶方案中，无论是在感知还是规控层面，核心算法基本都是由数据驱动的，数据的数量和质量决定了算法的性能，因此构建一套高效获取、标注及仿真训练数据的闭环至关重要。特斯拉每年售出近百万辆汽车，通过这些汽车日常运行，可以采集到超大规模的原始数据集，对这些数据集的标注方面经历了显著的发展：**1) 外包第三方进行人工数据标注。**2018年，特斯拉与第三方公司合作，采用人工标注，该方式标注效率低且沟通成本高。**2) 自建超千人团队进行手工标注。**而后为提升标注效率和质量，特斯拉自建标注团队，人员规模近千人。**3) 开发自动标注系统，实现人工与机器相结合的数据标注模式。**随着自动驾驶数据持续增长，所需标注人员的规模进一步扩大，使得人力成本快速增长，使得2020年特斯拉开始研发并使用数据自动标注系统，通过大量数据训练大模型，再用大模型训练车端小模型。新算法能够在12小时内自动标注一万个驾驶旅程，可抵充500万个小时的人工标注工作，极大地提高了标注效率。

图57：特斯拉数据标注发展历程

	image space (2018)	single trip (2019)	top-view (2020)	multi-trip (2021-)
3D label	unknown	manual	aligned	reconstructed
reprojection	< 1 pixel	< 3 pixel	< 7 pixel	< 3 pixel
topology	local	up to trajectory	unlimited	up to reconstruction
Labeling / clip	533 hrs	3.5 hrs	< 0.1 hr (avg)	< 0.1 hrs (avg)
compute / clip	not needed	1 hr	2 hrs	0.5 hrs (avg)
scalability	low	medium	high	very high
eng. effort	low	medium	high	very high

资料来源：Tesla AI Day 2022，民生证券研究院

从 2D 图像标注，到 4D 的向量空间标注，特斯拉 FSD 系统实现更精细化的路径规划和决策制定。通过引入 BEV（鸟瞰图）视角，特斯拉实现了从 2D 图像到 3D 车身自坐标系的转变，但最初这仅是对瞬时图像片段的感知，缺乏时空连续性。为了增强感知网络的短时记忆能力，特斯拉在感知网络架构中引入了时空序列特征层，使用视频片段代替单张图像来训练神经网络，从而提升了对场景的理解能力。到了 2022 年，特斯拉进一步升级了 BEV 感知，引入了 Occupancy Network，推动数据标注向 4D 升级，这不仅包括 3D 空间信息，还融入了时间维度，实现了对动态物体运动轨迹和参数的精确标注。

图58：基于 2D 图像进行标注



资料来源：Tesla AI Day 2021，民生证券研究院

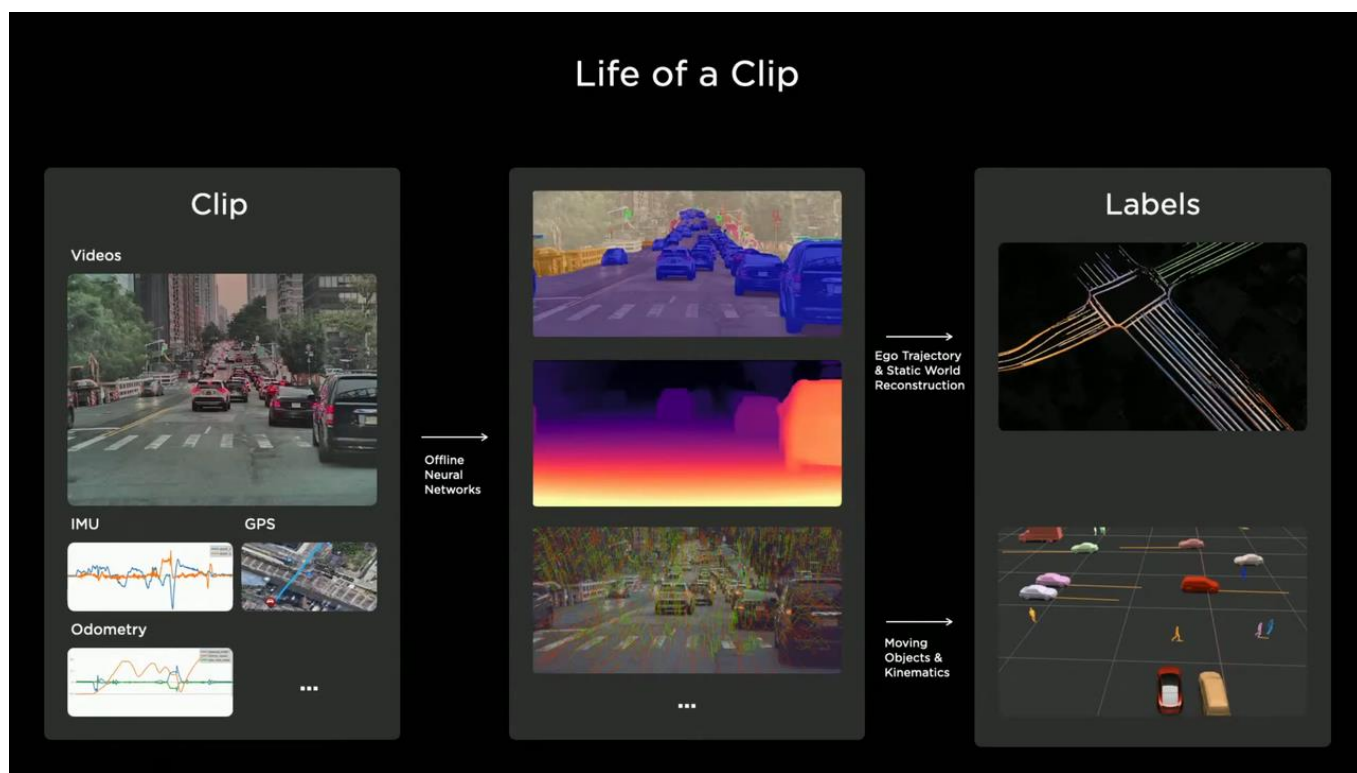
图59：在 BEV 空间下进行 4D 自动标注



资料来源：Tesla AI Day 2021，民生证券研究院

特斯拉的自动标注方案通过结合车辆采集的多源数据和先进的神经网络技术，实现了对自动驾驶所需数据的快速和高效处理。具体来说，这一方案首先利用车辆在一段时间内采集到的视频、IMU、GPS、里程表等数据构成最小标注单元 Clip，然后通过离线神经网络进行预测，得到深度图、坐标、目标物等中间层结果。最终，通过特征重建，获得 3D 标签集 (Labels) 以及所有目标的运动趋势，包括行车轨迹、静态环境重建、动态物体和运动学参数等，为自动驾驶算法的训练提供了大量高质量的标注数据。这一流程不仅提高了数据标注的效率，还保证了数据的准确性和一致性。

图60：特斯拉 4D 自动标注流程



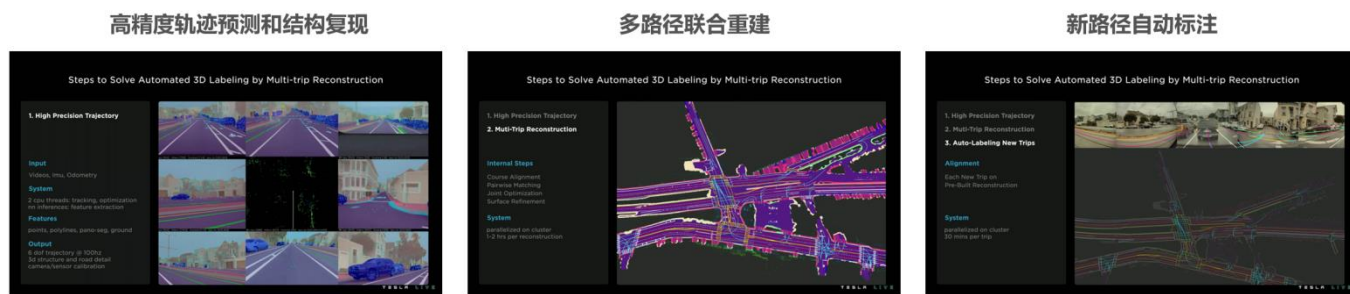
资料来源：Tesla AI Day 2021，民生证券研究院

特斯拉的 4D 自动标注技术通过三个关键步骤实现高精度的轨迹和结构复现：

- 1) **高精度轨迹预测和结构复现。**利用多摄像头和惯性里程计，通过粗对齐、关联、联合非线性优化及最终曲面优化，实现多重轨迹重建。这一步骤确保了轨迹数据的准确性和可靠性；
- 2) **多路径联合重建。**自动标记新轨迹的过程使用多轨迹对齐引擎，结合现有重建结果和新行驶轨迹，进一步细化路面细节。所有特征都是通过神经网络从视频中自动推断出来，并在向量空间进行跟踪和重构。这一自动化过程显著提高了标注效率，每个新轨迹的自动标记仅需 30 分钟，大大减少了人工标记的时间和劳动强度；
- 3) **新路径自动标注。**整个 4D 自动标注过程在集群上实现并行

化，整个过程从轨迹重建到最终标记的确定，仅需 1-2 小时。这种高效的自动化流程不仅提高了可扩展性，而且复建的路面和道路细节也为后续的人工验证提供了有力的指导，确保了标注的准确性和一致性。

图61：4D 自动标注的三个关键步骤



资料来源：Tesla AI Day 2022, YouTube, 民生证券研究院

6.2 仿真模拟：构建虚拟场景 优化算法表现

仿真模拟 (Simulation) 可以提供现实世界中难以获得或是难以标记的数据，从而加速 FSD 能力的训练，赋能模型迭代。由于路测条件的限制，导致积累数据和训练算法的效率偏低且成本高昂。为了更高效的实现数据训练，特斯拉构建了一个真实世界的虚拟仿真空间，来加速 FSD 能力的训练。自动驾驶的仿真是在模拟环境中，通过调整各类交通参与者及环境的模型参数以构建各种虚拟场景，以训练算法应对不同场景的性能。仿真模拟在自动驾驶领域中的价值体现在能够安全且高效地重现和构建极端及复杂场景以进行算法训练，实现快速且准确的数据标注，提供无风险的测试环境，优化规控算法，并通过持续的闭环场景训练提升特定功能如泊车技能，同时能够重现真实世界中的失败案例进行针对性优化，从而显著增强自动驾驶系统的安全性和可靠性。

特斯拉仿真模拟流程包含五大关键步骤：1) 传感器准确模拟。对真实摄像头属性进行模拟，包括传感器噪声、运动模糊、光学畸变等，以生成车辆上的摄像头和其他传感器实际会检测到的内容；2) 真实视觉渲染。特斯拉采用神经网络渲染技术与光线追踪方法，以使渲染效果能够逼近真实世界；3) 多元化素材库。特斯拉的素材库中，车辆、人物、动物等真实素材数量多达数千个，能够避免素材数量过少导致的神经网络过拟合问题。特斯拉还对大量地点进行了建模，以生成模拟环境；4) 可扩展场景生成。特斯拉基于算法，程序化地创建具有各种参数（如曲率、不同树木、锥体、杆子、不同速度的汽车等）的道。此外，特斯拉采用机器学习技术，使神经网络能够对故障点进行检测，并在故障点周围创造更多数据，形成闭环，

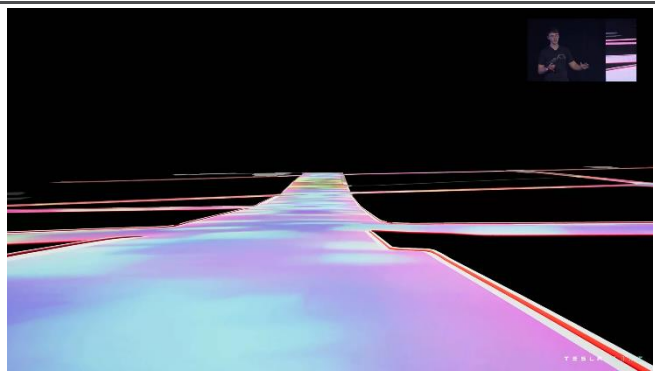
以优化网络性能；5) **场景重建**。特斯拉利用真实行驶视频片段，构建出与现实相同的虚拟场景，然后在虚拟场景中运行自动驾驶系统，观察系统运行情况。如果系统在模拟中失败，可以分析失败原因，从而对系统进行调整和优化，不断提升系统性能。

图62：特斯拉仿真模拟流程包含五大关键步骤



资料来源：Tesla AI Day 2021，民生证券研究院

特斯拉采用程序化方法进行场景生成，大幅提升运行效率。 特斯拉采用自动化工具进行场景生成，能够快速生成复杂模拟环境，过程包括：1) 通过自动化标签生成道路网格和车道；2) 使用线条数据创建车道标记；3) 利用中线边缘生成路中心分道区，并用随机植被填充；4) 以随机化启发式规则生成外部世界，包括建筑物、消防栓、树木等；5) 引入地图数据确定交通信号灯和停车标志的位置，并收集车道数量等信息。特斯拉可以在模拟器中更改道路真实标记，创造全新模拟数据，从而进行更有针对性的训练，提升预测准确性。

图63：通过自动化标签生成道路网格和车道


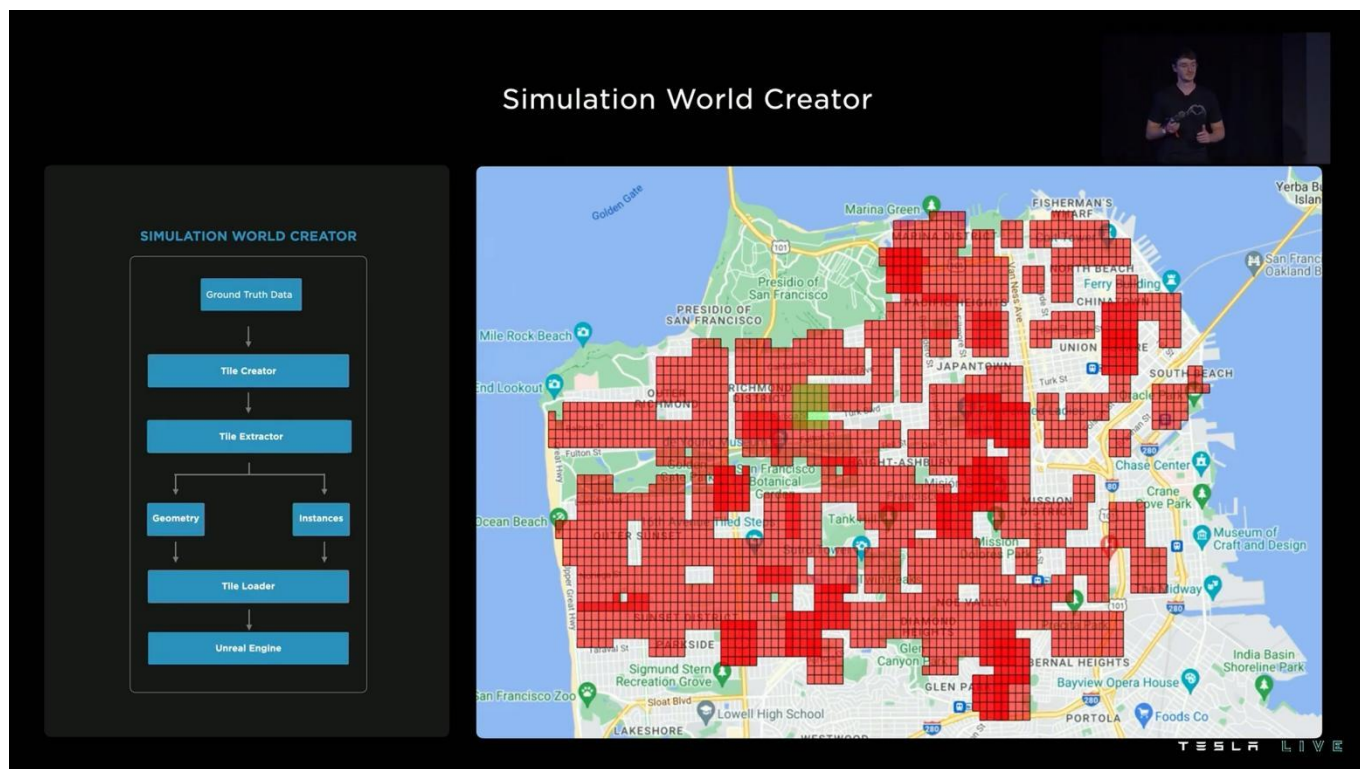
资料来源：Tesla AI Day 2022，民生证券研究院

图64：以随机化启发式规则生成外部世界


资料来源：Tesla AI Day 2022，民生证券研究院

特斯拉仿真模拟的主要创新和亮点在于高效的虚拟验证架构，能够将复杂的道路交通信息和场景元素系统化并高效运行。 特斯拉通过瓦片生成器（Tile Creator），将真实标签数据转化为具体的仿真元素，如车道线、路缘石和建筑物等。这些元素随后被瓦片提取器（Tile Extractor）分割成 150 平米的 Geohash 单元，每个单元都拥有一个独特的 ID 以便于快速加载和调用。这种处理方式使得建模信息更加简洁，提高了加载和渲染的效率。利用瓦片加载器（Tile Loader），特斯拉可以根据 Geohash ID 编码快速加载所需的仿真场景切片，专注于加载用户感兴趣的地点及其周边环境。最终，通过虚拟引擎生成完整的场景。通过这种方式，一个工程师在短短两周内即可生成旧金山街道的虚拟世界，显著减少了创建虚拟世界所需的时间。同样特斯拉可以利用此项 PDG 技术快速拓展到其他城市与国家，或者更新原有的虚拟世界，确保数据依据现实动态发展。

图65：特斯拉仿真世界创建流程

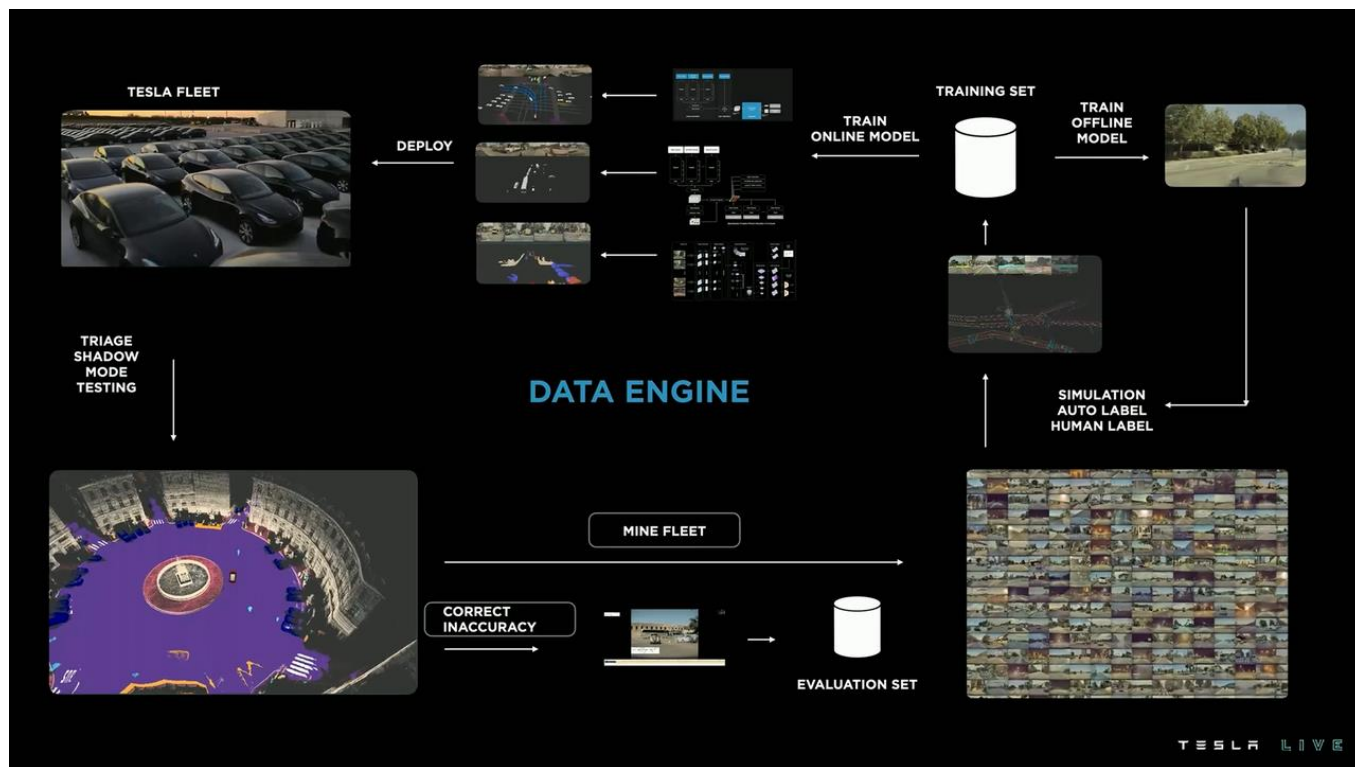


资料来源：Tesla AI Day 2022，民生证券研究院

6.3 数据引擎：集成多源数据 加速模型训练

特斯拉通过其自成闭环的数据引擎（Data Engine），有效地优化了神经网络。特斯拉通过标配自动驾驶硬件的车队进行数据采集，并通过规则和影子模式筛选出具有语义信息的有效数据，并回传至云端。在云端，利用工具对 AI 的错误输出进行纠正，并将这些数据整合入数据集。这些数据进一步用于训练车端的在线模型和云端的离线模型。最终，通过影子模式在车端进行新模型的测试和不同版本指标的比较，确保经过验证的新模型得以部署，实现数据和模型的持续优化。

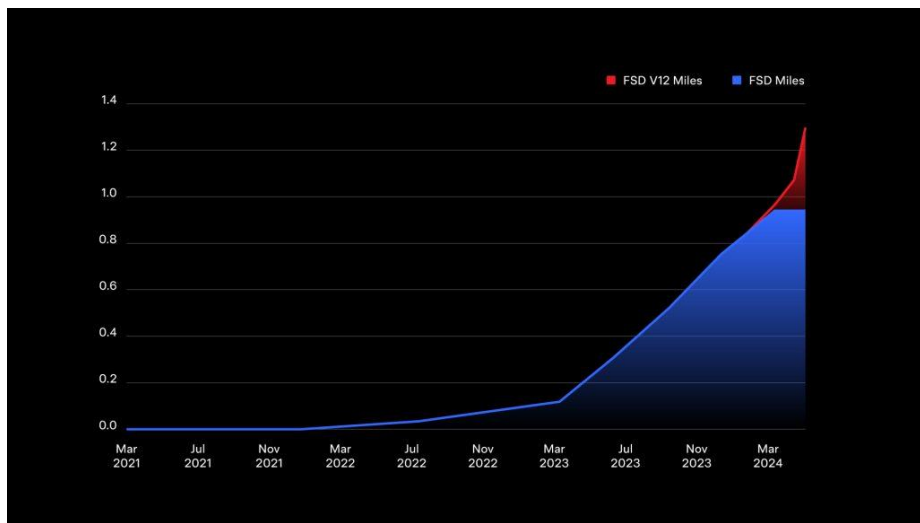
图66：特斯拉数据引擎示意图



资料来源：Tesla AI Day 2022，民生证券研究院

影子模式加速数据采集，训练质量有望大幅提升。特斯拉通过影子模式实现了量产车上的自动驾驶系统和传感器的持续运行与数据采集。在有人驾驶状态下，系统进行模拟决策并与驾驶员行为对比，不一致时触发数据回传，从而积累大量“极端工况”数据。这种模式不仅利用了量产车的广泛覆盖和低成本优势，还显著加速了数据的收集和训练模型的质量提升。截至 2024 年 4 月底，特斯拉全自动驾驶（FSD）用户的累计行驶里程已超过 13 亿英里，随着更多驾驶员试用并可能订购 FSD，这一数字有望出现大幅增长。

图67：特斯拉 FSD 用户的累计行驶里程已超过 13 亿英里

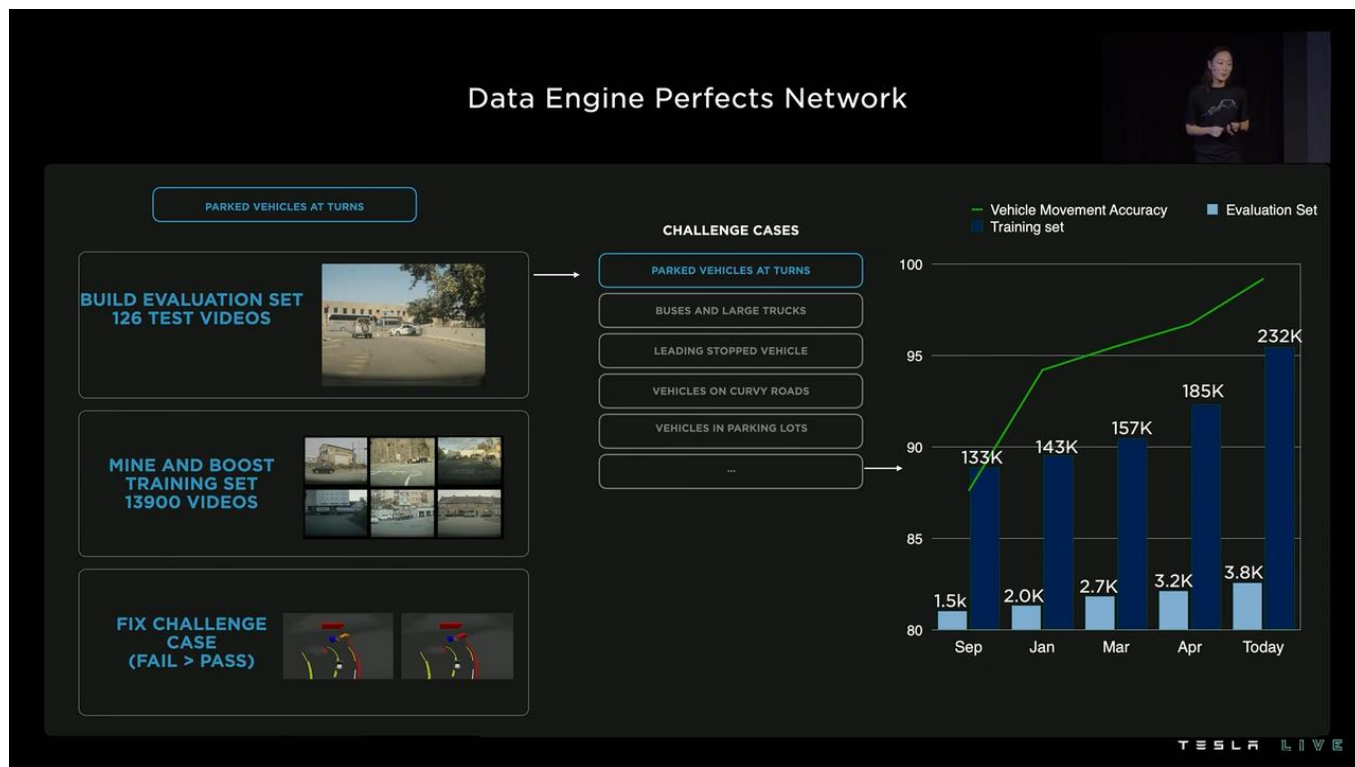


资料来源：Teslarati，民生证券研究院

影子模式是特斯拉自动驾驶系统中的关键技术，通过内置的触发器记录异常情况。在 2021 年 CVPR WAD 会议上，该模式已集成了 221 个触发器。当异常被触发时，影子模式会捕获异常发生前后的原始数据，这些数据经过清洗后，一部分形成了验证集，而其余数据则通过离线自动标注算法生成标签。

结合真实数据和标签，以及仿真和手动校准的数据，形成了综合训练数据集。这个数据集被用于训练车端的**在线模型**，涉及网络占用、车道线和障碍物检测以及规划算法。同时，该数据集也用于训练云端的**离线模型**，包括重建模型、感知模型的自动标注以及基于优化的规划模型。随着在线算法的更新和新数据的采集，影子模式再次捕获关键案例，并通过更新的离线模型进行自动标注，形成一个闭环的数据流和持续优化的自动驾驶系统。

图68：自成闭环的数据引擎能够更好地优化神经网络



资料来源：Tesla AI Day 2022，民生证券研究院

7 端到端架构：FSD V12 引领 实现感知决策一体化

FSD V12 为首个端到端自动驾驶系统，实现感知决策一体化。特斯拉 FSD v12 采用端到端大模型，消除了自动驾驶系统的感知和定位、决策和规划、控制和执行之间的断面，将三大模块合在一起，形成了一个大的神经网络，直接从原始传感器数据到车辆操控指令，简化了信息传递过程，因而减少了延迟和误差，提高了系统的敏捷性和准确性。FSD V12 能够模拟人类驾驶决策，成为自动驾驶领域全新发展路径。

7.1 特斯拉的 AI 关键时刻：FSD V12 首次实现端到端 V13 即将推出

特斯拉 FSD Beta V12 是首个实现端到端 AI 自动驾驶的系统，标志着自动驾驶技术的重大突破。2023 年 12 月 21 日，特斯拉正式向公众发布了最新版本的完全自动驾驶系统 FSD Beta V12。这一版本是首个实现端到端 AI 自动驾驶的系统，它通过一个单一的神经网络处理从摄像头获取的图像数据到输出车辆控制指令的全部过程。这种转变意味着，与之前的版本相比，V12 版本大幅减少了对传统编程的依赖，从超过 30 万行 C++ 代码减少至仅有 2000 行，更多地依赖于神经网络进行车辆的转向、加速和制动控制。

FSD Beta V12 的神经网络经过了数百万个视频剪辑的训练，能够模拟人类的驾驶决策。这一系统的设计哲学是尽可能地模仿人类的驾驶方式，让车辆能够自主地处理复杂的交通环境和不同的驾驶场景。此前，马斯克在 2023 年 8 月试驾直播中强调了 FSD Beta V12 的这一创新之处，将其描述为“端到端人工智能”，并称之为“Baby AGI（婴儿版通用人工智能）”，这表明了其在感知和理解现实世界复杂性方面的高级能力。特斯拉 FSD Beta V12 的发布，展示了通过端到端神经网络实现自动驾驶的新可能性，为自动驾驶领域提供新的发展路径。

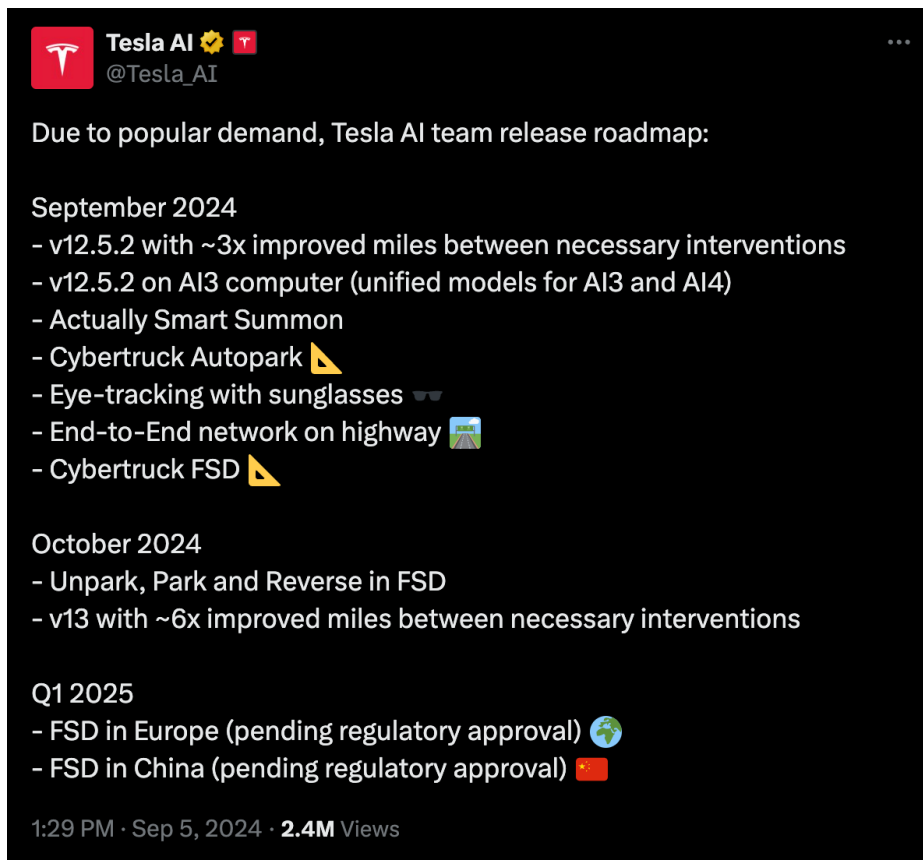
图69：马斯克强调：FSDv12 运行速度快了 10 倍，可以替代 30 万行代码



资料来源：42 号车库，民生证券研究院

FSD V13 即将 10 月推出，有望于 2025 年 Q1 进入中国和欧洲。2024 年 9 月 5 日，特斯拉 AI 官方推特发布 FSD 发展路线图：1) 2024 年 9 月，FSD V12.5.2 版本，必要干预的间隔里程提高 3 倍；V12.5.2 版本在 HW 3.0 平台上进行推送，和 HW4.0 统一模型；推送智能召唤功能；在 Cybertruck 上推送自动泊车和 FSD；上线基于太阳镜的注意力检测能力；推送高速端到端。2) 2024 年 10 月，将泊车和倒车功能整合进 FSD（端到端实现泊车、倒车功能）；FSD V13 版本的必要干预之间的里程数提高约 6 倍。3) 2025 年 Q1，在中国和欧洲推出 FSD（有待监管部门批准）。

图70: 特斯拉 FSD 发展路线图



资料来源: 赛博汽车, 民生证券研究院

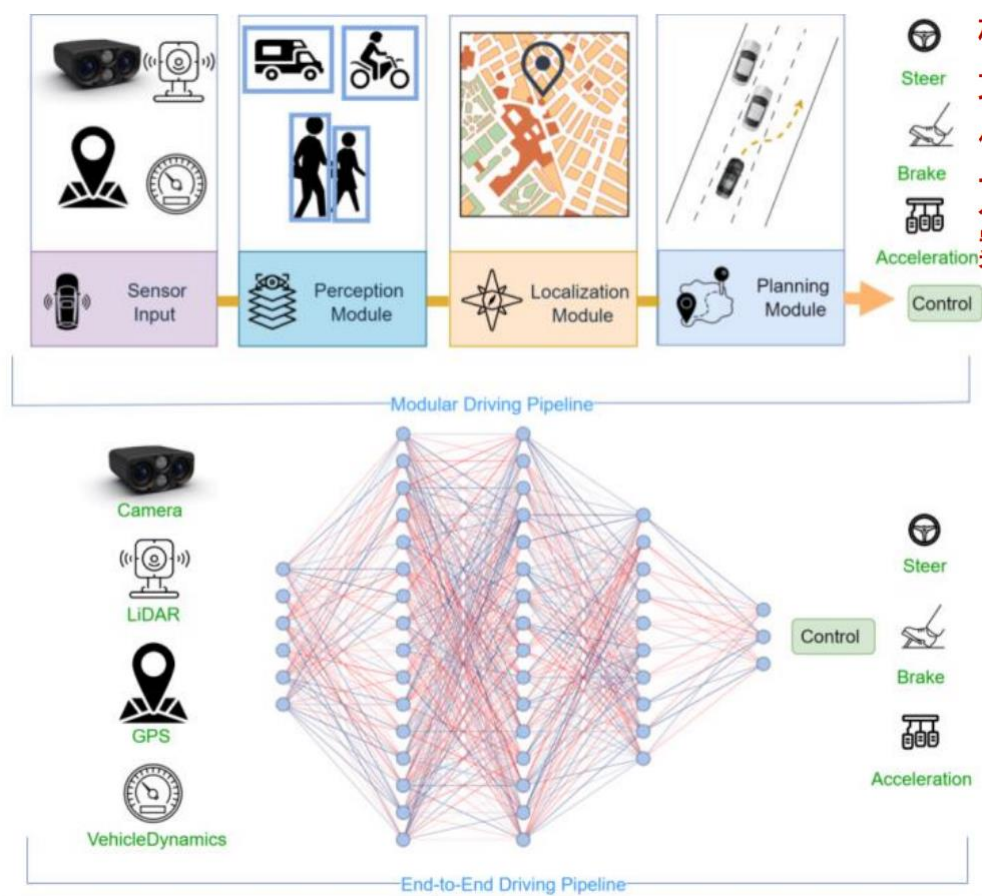
7.2 大模型成就端到端自动驾驶 推动感知决策一体化

目前, 自动驾驶系统的设计主要分为两大技术路径: 模块化方案和端到端方案。

1) 模块化路径。 将自动驾驶系统拆分为众多模块, 例如感知、规划和控制, 每个一级模块下又分为众多子模块, 每个模块可基于不同的规则或算法。由于每个独立模块负责单独的子任务, 因此出现问题时可及时回溯, 并易于调试, 具有较强的解释性; **2) 端到端路径。** 将自动驾驶系统视为一个整体, 在自动驾驶的应用中, 端到端模型可以将感知、规划和控制环节一体化, 通过将车载传感器采集到的信息直接输入神经网络, 经过处理后直接输出自动驾驶的驾驶命令, 潜在性能更佳、优化效率更高。

特斯拉 FSD v12 采用端到端大模型, 消除了自动驾驶系统的感知和定位、决策和规划、控制和执行之间的断面, 将三大模块合在一起, 形成了一个大的神经网络。通过这一神经网络架构, 直接从原始传感器数据到车辆操控指令, 简化了信息传递过程, 因而减少了延迟和误差, 提高了系统的敏捷性和准确性。但同时, 端到端的学习模式也带来了系统解释性相对较弱、问题定位相对困难等问题, 这是技术发展中需要持续关注并解决的挑战。

图71：多模块化方案 VS 端到端方案



资料来源：《Recent Advancements in End-to-End Autonomous Driving using Deep Learning: A Survey》Pranav Singh Chib 等，民生证券研究院

7.3 端到端架构演进：感知端到端到 One Model 端到端

截至目前，自动驾驶架构的演进可以分为四个主要阶段：

第一阶段：感知“端到端”。这一阶段，整个自动驾驶架构被拆分成了感知和预测决策规划两个主要模块，其中，感知模块已经通过基于多传感器融合的 BEV (Bird Eye View, 鸟瞰图视角融合) 技术实现了模块级别的“端到端”。通过引入 transformer 以及跨传感器的 cross attention 方案，感知输出检测结果的精度及稳定性相对之前的感知方案都有比较大的提升，不过，规划决策模块仍然以 Rule-based 为主。

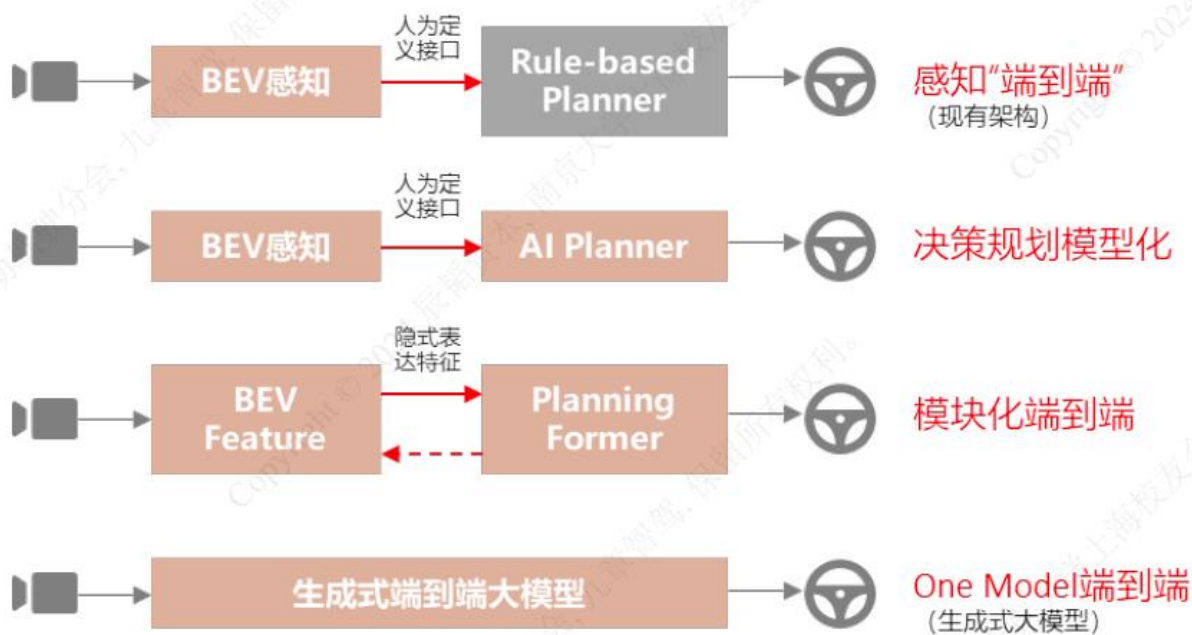
第二阶段：决策规划模型化。这个阶段，整个自动驾驶架构被仍然分为感知和预测决策规划两个主要模块，其中，感知端仍保持上一代的解决方案，但预测决策规划模块的变动比较大——从预测到决策到规划的功能模块已经被集成到同一个神经网络当中。值得注意的是，虽然感知和预测规划决策都是通过深度学习实现，但是这两个主要模块之间的接口仍然基于人类的理解定义（如障碍物位置，道路边

界等)；另外，在这一阶段，各模块仍然会进行独立训练。

第三阶段：模块化端到端。从结构上来讲，这一阶段的结构和上一阶段比较类似，但是在网络结构的细节及训练方案上有很大不同。首先，感知模块不再输出基于人类理解定义的结果，而更多给出的是特征向量。相应地，预测决策规划模块的综合模型基于特征向量输出运动规划的结果。除了两个模块之间的输出从基于人类可理解的抽象输出变为特征向量，在训练方式上，这个阶段的模型必须支持跨模块的梯度传导——两个模块均无法独立进行训练，训练必须通过梯度传导的方式同时进行。

第四阶段：One Model /单一模型端到端。在这一阶段，就不再有感知、决策规划等功能的明确划分。从原始信号输入到最终规划轨迹的输出直接采用同一个深度学习模型。基于实现方案的不同，这一阶段的 One Model 可以是基于强化学习 (Reinforcement Learning, RL) 或模仿学习 (Imitation Learning, IL) 的端到端模型，也可以通过世界模型这类生成式模型衍生而来。

图72：自动驾驶架构演进示意图



资料来源：辰韬资本，南京大学上海校友会自动驾驶分会，九章智驾，民生证券研究院

8 投资建议

智能驾驶奇点已至，海外特斯拉引领产业化进程，特斯拉高阶智驾技术端不断突破，FSD V12 已实现“端到端自动驾驶”，Robotaxi 将于 10 月发布。**国内三重拐点向上**，**1) 政策端**，L3 路测政策出台，国家、地方层面自动驾驶汽车全无人商业化运营相关政策不断落地；**2) 供给端**，萝卜快跑 Robotaxi 厂商商业化进程迅猛，已在包括北京、武汉、重庆、深圳和上海在内的 11 个城市启动了载人测试运营，计划到 2024 年底在武汉部署 1000 辆无人驾驶出租车，并力争在 2025 年实现盈利。华为、小鹏、理想、小米无图城市 NOA 2024 年将加速开通，特斯拉 FSD V12 端到端版本将进一步开放测试及使用；**3) 需求端**，社会公众对于全无人自动驾驶汽车接受度不断提升，智驾逐步成为购车重要影响因素。2024 年智能驾驶将迎来为产业化加速的元年，看好高阶智能驾驶功能 (NOA) 渗透率稳步提升。

投资建议：FSD 技术端快速进化，“We, Robot” 发布会即将开幕，Robotaxi 即将推出，有望成为特斯拉发展历程重要里程碑，并与 FSD V13 发布形成共振，共同成为板块强劲催化，智能驾驶奇点已至：推荐 1) 智能驾驶-【伯特利、德赛西威、经纬恒润-W、科博达】+智能座舱-【上声电子 (智能声学)、继峰股份 (智能座椅)、光峰科技 (车载光学)】；2) T 链-【拓普集团、新泉股份、爱柯迪、旭升集团、双环传动、岱美股份、银轮股份】。

表4：重点推荐标的

证券代码	证券简称	股价 (元)	EPS			PE			评级
			2023A	2024E	2025E	2023A	2024E	2025E	
603596.SH	伯特利	52.15	2.15	1.93	2.58	24	27	20	推荐
002920.SZ	德赛西威	131.77	2.81	3.76	4.78	47	35	28	推荐
688326.SH	经纬恒润	102.79	-1.82	-0.12	2.45	-	-	42	推荐
603786.SH	科博达	63.6	1.52	2.09	2.70	42	30	24	推荐
688533.SH	上声电子	36.99	0.99	1.75	2.38	37	21	16	推荐
603997.SH	继峰股份	13.57	0.18	0.05	0.67	-	-	20	推荐
601689.SH	拓普集团	49.6	1.95	1.76	2.34	25	28	21	推荐
603179.SH	新泉股份	51.87	1.65	2.22	2.91	31	23	18	推荐
600933.SH	爱柯迪	16.1	1.03	1.19	1.54	16	14	10	推荐
603305.SH	旭升集团	11.86	0.77	0.60	0.76	15	20	16	推荐
002472.SZ	双环传动	30.28	0.97	1.24	1.58	-	-	19	推荐
603730.SH	岱美股份	11.17	0.51	0.53	0.63	22	21	18	推荐
002126.SZ	银轮股份	20.54	0.77	1.03	1.26	27	20	16	推荐

资料来源：iFind，民生证券研究院（注：股价为 2024 年 10 月 8 日收盘价）

9 风险提示

智能化渗透率提升不及预期：智能化正处于发展初期，若受制于成本、技术等因素，后续渗透率提升可能不及预期；

智能驾驶技术发展不及预期：智能驾驶技术突破存在延迟可能性，导致自动驾驶汽车的安全性和可靠性无法满足商业化的要求；

Robotaxi 商业化进展不及预期：技术发展缓慢、政策监管的不确定性与消费者接受度的不确定性可能导致市场需求低于预期，影响 Robotaxi 服务的推广和应用。

插图目录

图 1: 2013 年 9 月, 马斯克在 Twitter 中首次提到 AP	5
图 2: 特斯拉自动驾驶系统发展历程.....	6
图 3: FSD 累计行驶里程.....	6
图 4: FSD 订阅价格降至 99 美元/月.....	6
图 5: 特斯拉 FSD 系统架构.....	7
图 6: 特斯拉视觉感知系统算法采用 HydraNets 架构.....	9
图 7: 图像空间预测投射到向量空间后出现较大偏差.....	9
图 8: 单相机检测无法解决物体横跨多相机的问题.....	9
图 9: BEV 视角融合了多个摄像头的视频数据.....	10
图 10: Transformer 是实现二维到三维变换的核心.....	10
图 11: 通过图像校准解决摄像头采集数据偏差问题.....	11
图 12: 加入虚拟标准摄像头以校准图像数据偏差.....	11
图 13: 感知网络仍是对瞬时图像片段进行感知.....	11
图 14: 特斯拉引入时空序列特征层.....	11
图 15: 特征序列模块可以缓存时序与空间特征.....	12
图 16: 隐状态可组织成二维网格.....	13
图 17: 空间 RNN 的隐状态可包含多个通道.....	13
图 18: Occupancy Network 对正在启动的两节公交车运动状态进行精准捕捉.....	14
图 19: Occupancy Network 能够生成可行使表面.....	14
图 20: 基于 Attention 机制的 Occupancy Network 占用网络.....	15
图 21: Lanes Network 旨在生成车道与连接信息.....	16
图 22: 特斯拉采用低精度地图对视觉表示进行增强.....	16
图 23: Lanes Network 工作原理.....	17
图 24: 稀疏化处理可使神经网络专注于计算最重要的区域.....	17
图 25: 效率、安全和舒适是自动驾驶规划的三大目标.....	18
图 26: 非凸性和高维性是自动驾驶规划的两大难点.....	18
图 27: 特斯拉将规划问题进行分层分解.....	19
图 28: 自动驾驶系统需要实现多代理联合轨迹规划.....	19
图 29: 自动驾驶行驶方案评估至少需要 10 毫秒.....	19
图 30: 特斯拉采用交互搜索框架, 实现实时方案评估.....	20
图 31: 特斯拉 Dojo 发展历程.....	21
图 32: 特斯拉 Dojo 算力规划.....	22
图 33: Dojo 超级计算机三大目标.....	23
图 34: Dojo 系统二维网格结构.....	23
图 35: Dojo 训练节点架构.....	24
图 36: D1 芯片结构.....	24
图 37: D1 芯片计算阵列.....	25
图 38: D1 芯片串行器/解串行器分布.....	25
图 39: 特斯拉 Dojo 指令集.....	25
图 40: 计算平面两端各放置了一个接口处理器.....	26
图 41: Dojo 接口处理器连接在系统托盘下方.....	26
图 42: 训练瓦片上集成了 25 个 D1 芯片.....	27
图 43: 系统托盘上训练瓦片呈 2x3 矩阵式排布.....	27
图 44: Dojo 训练机柜中集成了两个系统托盘.....	27
图 45: 系统托盘上训练瓦片呈 2x3 矩阵式排布.....	27
图 46: 系统性能由硬件、利用率和加速器占用率决定.....	28
图 47: 多加速器运行批量归一化会导致前向传播延迟.....	28
图 48: Dojo 系统使用高效同步机制, 实现内部单元协作.....	28
图 49: 自动标注与占用网络占特斯拉 GPU 使用量一半.....	29
图 50: Dojo 处理自动标注与占用网络速度大幅提升.....	29
图 51: 特斯拉 FSD 双芯片系统设计.....	31
图 52: 特斯拉 FSD 芯片发展历程.....	32
图 53: 特斯拉 FSD 芯片架构.....	33
图 54: 汽车中在同时运行很多架构、模块和网络.....	34

图 55: 神经网络编译器与链接器架构	34
图 56: 特斯拉混合调度系统架构	35
图 57: 特斯拉数据标注发展历程	37
图 58: 基于 2D 图像进行标注	37
图 59: 在 BEV 空间下进行 4D 自动标注	37
图 60: 特斯拉 4D 自动标注流程	38
图 61: 4D 自动标注的三个关键步骤	39
图 62: 特斯拉仿真模拟流程包含五大关键步骤	40
图 63: 通过自动化标签生成道路网格和车道	41
图 64: 以随机化启发式规则生成外部世界	41
图 65: 特斯拉仿真世界创建流程	42
图 66: 特斯拉数据引擎示意图	43
图 67: 特斯拉 FSD 用户的累计行驶里程已超过 13 亿英里	44
图 68: 自成闭环的数据引擎能够更好地优化神经网络	45
图 69: 马斯克强调: FSDv12 运行速度快了 10 倍, 可以替代 30 万行代码	47
图 70: 特斯拉 FSD 发展路线图	48
图 71: 多模块化方案 VS 端到端方案	49
图 72: 自动驾驶架构演进示意图	50

表格目录

表 1: Dojo 算力分为内核级、芯片级、格点级、集群级等四个层级	23
表 2: 特斯拉自动驾驶硬件平台不同版本对比	31
表 3: FSD 芯片 1.0 和 2.0 性能对比	33
表 4: 重点推荐标的	51

分析师承诺

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格并登记为注册分析师，基于认真审慎的工作态度、专业严谨的研究方法与分析逻辑得出研究结论，独立、客观地出具本报告，并对本报告的内容和观点负责。本报告清晰地反映了研究人员的研究观点，结论不受任何第三方的授意、影响，研究人员不曾因、不因、也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

评级说明

投资建议评级标准	评级	说明
以报告发布日后的 12 个月内公司股价（或行业指数）相对同期基准指数的涨跌幅为基准。其中：A 股以沪深 300 指数为基准；新三板以三板成指或三板做市指数为基准；港股以恒生指数为基准；美股以纳斯达克综合指数或标普 500 指数为基准。	推荐	相对基准指数涨幅 15%以上
	谨慎推荐	相对基准指数涨幅 5% ~ 15%之间
	中性	相对基准指数涨幅-5% ~ 5%之间
	回避	相对基准指数跌幅 5%以上
行业评级	推荐	相对基准指数涨幅 5%以上
	中性	相对基准指数涨幅-5% ~ 5%之间
	回避	相对基准指数跌幅 5%以上

免责声明

民生证券股份有限公司（以下简称“本公司”）具有中国证监会许可的证券投资咨询业务资格。

本报告仅供本公司境内客户使用。本公司不会因接收人收到本报告而视其为客户。本报告仅为参考之用，并不构成对客户的投资建议，不应被视为买卖任何证券、金融工具的要约或要约邀请。本报告所包含的观点及建议并未考虑个别客户的特殊状况、目标或需要，客户应当充分考虑自身特定状况，不应单纯依靠本报告所载的内容而取代个人的独立判断。在任何情况下，本公司不对任何人因使用本报告中的任何内容而导致的任何可能的损失负任何责任。

本报告是基于已公开信息撰写，但本公司不保证该等信息的准确性或完整性。本报告所载的资料、意见及预测仅反映本公司于发布本报告当日的判断，且预测方法及结果存在一定程度局限性。在不同时期，本公司可发出与本报告所刊载的意见、预测不一致的报告，但本公司没有义务和责任及时更新本报告所涉及的内容并通知客户。

在法律允许的情况下，本公司及其附属机构可能持有报告中提及的公司所发行证券的头寸并进行交易，也可能为这些公司提供或正在争取提供投资银行、财务顾问、咨询服务等相关服务，本公司的员工可能担任本报告所提及的公司的董事。客户应充分考虑可能存在的利益冲突，勿将本报告作为投资决策的唯一参考依据。

若本公司以外的金融机构发送本报告，则由该金融机构独自为此发送行为负责。该机构的客户应联系该机构以交易本报告提及的证券或要求获悉更详细的信息。本报告不构成本公司向发送本报告金融机构之客户提供的投资建议。本公司不会因任何机构或个人从其他机构获得本报告而将其视为本公司客户。

本报告的版权仅归本公司所有，未经书面许可，任何机构或个人不得以任何形式、任何目的进行翻版、转载、发表、篡改或引用。所有在本报告中使用的商标、服务标识及标记，除非另有说明，均为本公司的商标、服务标识及标记。本公司版权所有并保留一切权利。

民生证券研究院：

上海：上海市浦东新区浦明路 8 号财富金融广场 1 幢 5F； 200120

北京：北京市东城区建国门内大街 28 号民生金融中心 A 座 18 层； 100005

深圳：广东省深圳市福田区益田路 6001 号太平金融大厦 32 层 05 单元； 518026