

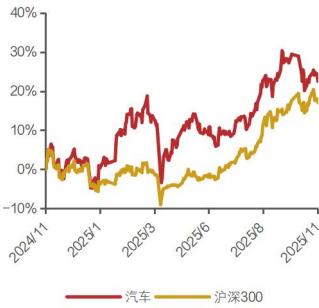
投资评级：看好（维持）

证券分析师

李泽
SAC: S1350525030001
lize@huayuanstock.com
陈嵩
SAC: S1350525070005
chensong@huayuanstock.com

联系人

板块表现：



ICCV 2025 Foundational Model for FSD 专题演讲解读

——汽车行业双周报（20251020-20251102）

投资要点：

- 特斯拉选择“端到端框架”的核心原因：模块化框架难以编码人类价值观、感知与预测规划之间的接口定义会带来信息传递损失、端到端能应对现实世界的长尾问题/同构计算带来的确定性延迟，端到端也能更好地契合AI领域的Scaling-law。
- 端到端面临的主要挑战以及特斯拉如何解决：维度灾难（特斯拉采用复杂的触发机制来回传长尾场景数据）、怎么保证模型的可解释性与安全性（允许模型输出多样的中间结果（例如占用、其它目标、交通信号灯、交通标志、道路边界等），这些中间结果可通过车机渲染展示给驾驶员与乘客）；通过类似理想、小鹏的VLA/VLM框架，更好解释为什么做出某个决策&应对长尾场景、如何评估模型（多元化的评价体系、重视闭环测试、评估模型决策可能造成的结果、复用历史问题库测评新模式；构建世界模拟器这一强大的测评仿真工具）。
- 模型具备强大的可迁移性：Robotaxi&Optimus实质上复用了同一套FSD神经网络，只要在训练数据中增加了相应数据样本，模型便能泛化到其他机器人形态。
- 投资分析意见：建议关注数据积累丰富+技术路线选择近似特斯拉的头部智驾整车公司（理想、小鹏）、核心环节Tier1（线控转向（耐世特、浙江世宝））、线控制动（伯特利）、大算力域控制器（德赛西威、均胜电子、经纬恒润、中科创达）、增量零部件（激光雷达（禾赛科技、速腾聚创、图达通）、大算力芯片（地平线机器人、黑芝麻智能））、Robotaxi运营（文远知行、小马智行、曹操出行、千里科技）、RoboVan（九识智能、新石器）等。
- 风险提示：1) 法规进展不及预期；2) 竞争格局恶化；3) 技术路线迭代风险。

内容目录

1. 前言	4
2. 特斯拉为什么放弃显式模块化的算法框架	5
2.1. 模块化方法虽更易调试模型，但要精确编码人类价值观非常困难	5
2.2. 感知与预测规划之间的接口定义很困难	5
2.3. 端到端架构的其他优势	6
3. 端到端的挑战	6
3.1. 如何解决维度灾难（Curse of dimensionality）	7
3.2. 端到端系统怎么调试&开发？怎么保证可解释性？（Interpretability and safety guarantees）	9
3.3. 3D 高斯溅射（3D Gaussian Splatting）	10
3.4. 模型评测	12
4. 未来规划	15
5. 总结	16
6. 投资建议	16
7. 风险提示	17

图表目录

图表 1: 特斯拉从 V12 版本开始转向基于海外数据训练、运行在 36Hz、直接输出控制指令的端到端神经网络	4
图表 2: 特斯拉算法迭代历程总结	4
图表 3: 为什么需要从模块化走向端到端	5
图表 4: 编码人类价值观是非常困难的	5
图表 5: 实拍鸡、鹅过马路场景, FSD 依托场景理解做了不同决策	6
图表 6: 从输入像素到输出控制指令的端到端系统面临三大主要挑战	6
图表 7: 维度灾难指输入的数据非常庞大, 却要映射到非常简单的输出	7
图表 8: 特斯拉是如何收集长尾场景数据	8
图表 9: 前车撞到护栏后弹回, FSD 精确判断事故车可能回弹, 提前减速	8
图表 10: 端到端不意味着不能输出中间过程	9
图表 11: 所有的中间结果都是模型的一部分, 也都可以单独输出中间结果信息	9
图表 12: 特斯拉也采用了类似理想的“快慢”双系统框架, 通过 LLM 辅助决策	10
图表 13: 基于 2D 图像做 3D 场景重建的算法迭代史	10
图表 14: NeRF 训练过程	11
图表 15: 3DGS 原理	11
图表 16: 3DGS VS NeRF	11
图表 17: 特斯拉的生成式 3DGS VS 传统 3DGS	12
图表 18: 特斯拉是如何测评模型的	12
图表 19: 特斯拉基于“状态-动作”对构建世界模拟器	13
图表 20: 世界模拟器可以与车辆的策略网络闭环运行	13
图表 21: 基于单一网络生成 8 路摄像头不同视角的长视频	13
图表 22: 可以用历史问题库中的场景来检验新模式表现	14
图表 23: 基于相同场景自由变换场景中的某一元素	14
图表 24: 降低计算精度就可以获得接近实时渲染的仿真效果	15
图表 25: FSD 的神经网络模型迁移到 Optimus 上	15

1. 前言

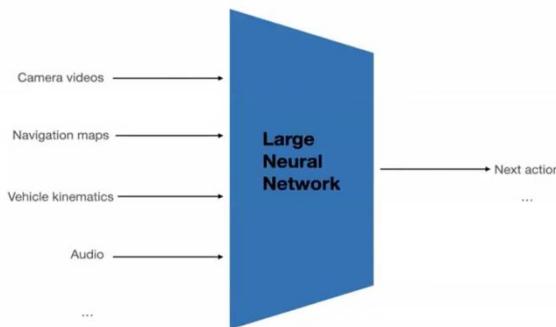
本文主要内容援引自 Tesla AI Software VP 及目前 Optimus 人形机器人项目 leader Ashok Elluswamy (阿肖克·埃鲁斯瓦米) 于 2025 年 10 月在计算机视觉顶会 ICCV 上的分享。

共识信息——V12 版本开始特斯拉 FSD 全面转向端到端：改用单一大型神经网络，把像素以及其他传感器数据作为输入，直接输出下一步要执行的动作。也就是说，在车辆行驶过程中，它不再显式地去做诸如车辆识别、道路边界之类的感知。这些信息可以作为隐式表征存在，或作为辅助任务来训练。

图表 1：特斯拉从 V12 版本开始转向基于海外数据训练、运行在 36Hz、直接输出控制指令的端到端神经网络

End-to-end foundation model for all of driving

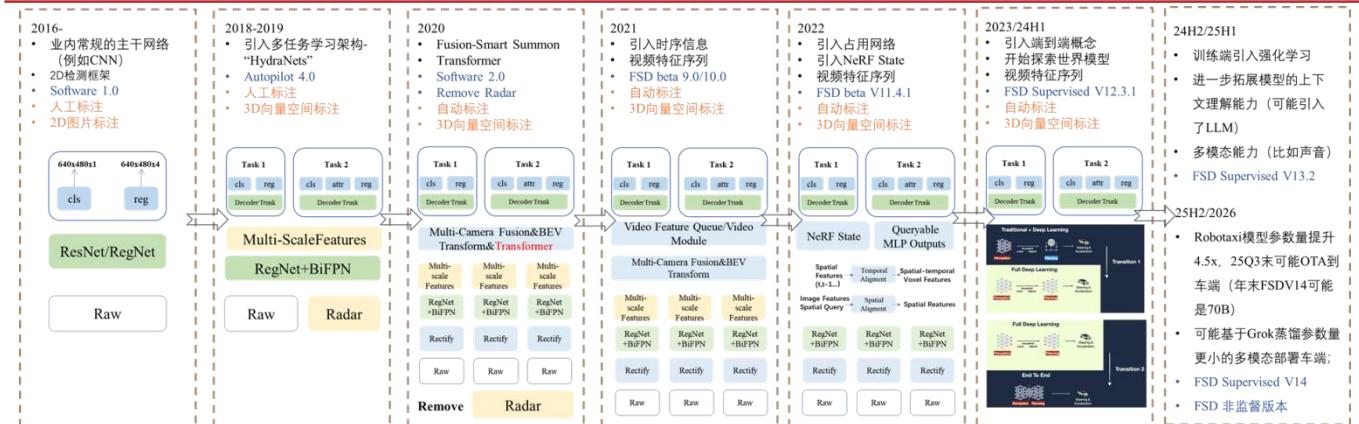
A large model, trained with tons of data, has large context, runs at 36 Hz, directly produces control actions



资料来源：Tesla ICCV 2025 Foundational Model for FSD (作者：Ashok Elluswamy)，华源证券研究所

分享的价值——Ashok 系统展示了特斯拉基于端到端基础模型驱动物理 AI (乘用车 L3/L4 & 人形机器人) 的完整叙事链条：从数据→模型→可解释&安全保障→评测→仿真→跨形态迁移，讨论了端到端方法的优势&难点以及模型评估的重要性，为我们了解特斯拉 FSD 为什么从模块化演化到端到端提供了最新洞见。

图表 2：特斯拉算法迭代历程总结



资料来源：青源 talk，《BEVFormer：新一代自动驾驶环视感知算法》代季峰，汽车之家公众号等，华源证券研究所

2. 特斯拉为什么放弃显式模块化的算法框架

2.1. 模块化方法虽更易调试模型，但要精确编码人类价值观非常困难

例子 1：遇到障碍物时，是选择刹车还是绕过去；是为了尽快通过晚点刹车，还是为了驾驶平缓提早刹车。这些取舍本质上是对人偏好的对齐，而不是靠诸如“刹车力度多大、何时施加”之类的某种客观指标界定。且这种权衡可能也跟场景相关，比如低速与高速最佳策略差别很大。

例子 2：车行驶在双向单车道上。自车道前方有个水坑，是否应该为了避开水坑借用对向车道？如果尝试用显式函数编码这个策略会很困难，压过水坑的代价 VS 借用对向车道的风险如何权衡？但如果是人开车，直觉上就很容易判断借道的风险大小（例如能见度高+道路够宽，可能很自然地就借道了）。

图表 3：为什么需要从模块化走向端到端

The case for end-to-end driving

- Codifying human values is incredibly difficult
- Interface between perception, prediction and planning is ill-defined
- Easily scalable to handle the fat and long tail of real-world robotics
- Homogenous compute with deterministic latency
- Overall, on the correct side of scaling w.r.t. the bitter lesson

资料来源：AI 工业公众号，华源证券研究所

图表 4：编码人类价值观是非常困难的

Codifying human values is difficult

Real-world is filled with tiny trolley problems

What's better:

1. Going over the small puddle
2. Going into the oncoming lane



资料来源：AI 工业公众号，华源证券研究所

2.2. 感知与预测规划之间的接口定义很困难

传统模块化自动驾驶系统中，各模块间通过预设接口通信。例如使用边界框（Bounding Box）定义车辆行人，用多段线描述车道线。一些特殊场景表面接口间的信息传递极易造成信息损失，比如：图表 5 左图，FSD 等待所有鸡都过完马路才继续前进；图表 5 右图，路上有鹅，FSD 等待了一会儿，发现路中的鹅实际是静止的，于是果断左转通过。鸡、鹅过马路的实拍视频表明，传统感知接口难以定义这些动物的类型、过马路意图等属性。但如果像人类一样端到端（E2E）地看整个场景，答案其实一目了然（最大程度减少信息传递损失，确保决策模块获取更全面的环境信息）。

图表 5：实拍鸡、鹅过马路场景，FSD 依托场景理解做了不同决策

Interface between perception & planning is ill-defined

Understanding that the last chicken is waiting to cross



Understanding that the geese aren't crossing and should go around



资料来源：AI 工业公众号，华源证券研究所

2.3. 端到端架构的其他优势

应对现实世界的长尾问题：减少信息传递损失，确保在罕见场景下，决策系统仍能基于丰富的输入信息做出合理判断，并通过学习人类驾驶行为获得处理长尾场景的能力。

同构计算带来的确定性延迟：自动驾驶系统对时延极其敏感。传统基于规则和优化的规控方案，其求解时间受环境复杂性、初始解质量等多种因素影响，难以保证稳定。而端到端神经网络具有固定的模型结构和参数量，其单帧计算延迟是确定性的，有利于控制系统的时问波动。

更好地契合 AI 领域的数据规模效应：端到端自动驾驶是彻底的数据驱动范式，它摒弃了大量人为设计的规则和评价指标，从而能更充分地利用计算力和数据的增长来提升系统性能。

3. 端到端的挑战

三大主要挑战为：维度灾难、可解释性&安全保证、如何评估模型。

图表 6：从输入像素到输出控制指令的端到端系统面临三大主要挑战

Main Challenges of learning pixels -> control

1. Curse of dimensionality
2. Interpretability and safety guarantees
3. Evaluation

资料来源：AI 工业公众号，华源证券研究所

3.1. 如何解决维度灾难 (Curse of dimensionality)

几十亿 token 的输入要映射到 2 个 token 输出：考虑车端 7–8 个摄像头+其他必要传感器，以 36Hz 的帧率采集过去 30s 的 7 路摄像头视频数据、长达数英里的导航地图、100Hz 的车速与 IMU 信息、48KHz 的音频信息（FSD V14 已可以处理音频信息），数据总量可以达到 20 亿 tokens，而输出仅为方向盘和加减速信号，约等于 2 个 token。

图表 7：维度灾难指输入的数据非常庞大，却要映射到非常简单的输出

1. Curse of dimensionality

Extremely large context length is a minimum requirement for driving

- Input context length of 2 billion tokens:
 - 7 cameras x 36 FPS x 5 Mega pixels x 30s history / (5x5 pixel patch)
 - Navigation maps and route for next few miles
 - 100 Hz kinematic data such as speed, IMU, odometry, etc
 - 48 KHz audio data
- Output tokens:
 - Next steering and acceleration
- Need to learn the correct causal mapping of 2 billion tokens -> 2 tokens

资料来源：AI 工业公众号，华源证券研究所

特斯拉通过强大的数据引擎 (Data Engine) 来应对此挑战，致力于采集大量高质量数据。Tesla 车队每日可产生相当于 500 年驾驶时长的数据，但其中多数为常规场景；为此特斯拉采用了复杂的触发机制来回传长尾场景数据，包括：

- 使用显式触发器或面向特定场景的小型神经网络来抓取样本；
- 基于预测偏差回传 bad cases；
- 收集所有用户接管数据；
- 收集感知状态突变的场景。

图表 8：特斯拉是如何收集长尾场景数据

1. Curse of dimensionality

**Tesla fleet can provide 500 years of driving data every single day
But most of driving data is boring**

- Sophisticated trigger based data collection
- Tiny NNs catch specific corner cases, such as emergency vehicles
- Can evaluate prediction versus what happened in hindsight
- Any user intervention during autonomous operation
- Any large change in state space

资料来源：AI 工业公众号，华源证券研究所

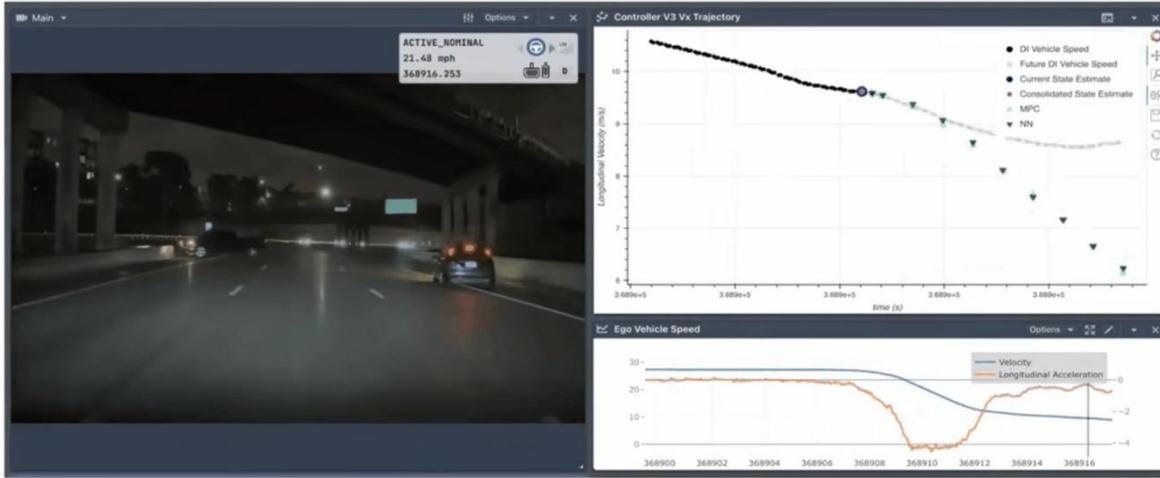
高效的数据筛选与回传机制，使得 Tesla 能收集海量的极端场景和主动避险数据，确保 FSD 模型具备极强的泛化能力，面对极端场景不至于到最后一秒才被动反应。以下是个体现模型主动安全能力的例子（Can evaluate prediction versus what happened in hindsight）：

前车打滑失控，随后撞向护栏；FSD 在该车出现打滑情况时开始刹车，FSD 提前判断前车这次碰撞不是一次碰撞，而是可能撞到护栏后发生回弹的二次碰撞。同时，也没把这种迹象误判为对方的正常变道，系统预判到了风险，并请求了 4 m/s^2 的减速度。只有当拥有海量数据并覆盖了这些角落场景时，才做得到这种提前量。智能化程度更低的系统往往要等到方向真的改变了，或者坏事已经发生了，才会反应——而拥有大量数据支撑的“直觉式”端到端系统，可以更早做出正确动作。

图表 9：前车撞到护栏后弹回，FSD 精确判断事故车可能回弹，提前减速

1. Curse of dimensionality

Large data gives extreme generalization and proactive safety



资料来源：AI 工业公众号，华源证券研究所

3.2. 端到端系统怎么调试&开发？怎么保证可解释性？ (Interpretability and safety guarantees)

端到端并不意味着不能输出其他中间结果：同一个模型可以通过提示词（prompt）来预测任意辅助信息：例如占用、其它目标、交通信号灯、交通标志、道路边界等；甚至可以用自然语言问它“为什么这样决策”、“是否理解这个场景”。这些能力有助于解释模型做出的决策，并在一定程度上为系统安全性提供佐证。另外，端到端也不意味着不能有面向特定传感器的专用编码器，统一用一个超大的Transformer来做所有的运算就可能需要输入海量token，那可能会面临算力瓶颈+达不到实时延迟的要求。**需要贯彻始终的原则只有梯度必须端到端贯通。**

图表 10：端到端不意味着不能输出中间过程

图表 11：所有的中间结果都是模型的一部分，也都可以单独输出中间结果信息

2. Interpretability and safety guarantees

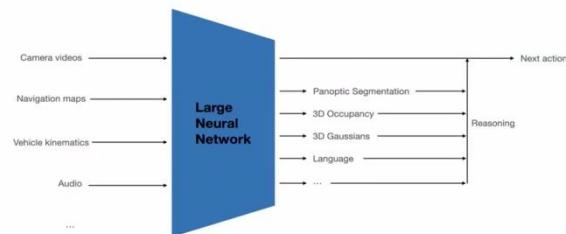
Chain-of-thought and process verification to the rescue

- Foundation model also predicts many interpretable outputs such as
- 3D occupancy and flow
- Objects such as vehicles, pedestrians, bicyclists, etc.,
- Traffic controls
- Road boundaries, lanes and semantics
- Speed limits and other road attributes
- Decisions expressed as plain language

资料来源：AI 工业公众号，华源证券研究所

2. Interpretability and safety guarantees

Looks more like this



资料来源：AI 工业公众号，华源证券研究所

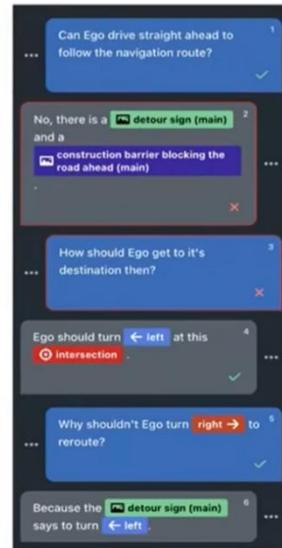
这些中间结果信息不仅仅用于车机可视化渲染，也可以通过条件概率（LLM 中广泛应用的思维链（COT，Chain-of-Thought）形式以及过程校验手段），确保最终输出的驾驶控制信号正确。这说明特斯拉的技术路线实际跟国内理想、小鹏等车企的VLA/VLM方案有异曲同工之妙。

具体如下：模型可以通过提示词来解释场景中的要素/解释为什么做出某个决策。实时驾驶时可能不需要把所有的解释功能都开启。但需要的时候，模型可以思考更久、产生更多推理 token，再输出与推理一致的控制动作。也就是车上运行同一个模型，该模型可以根据情景选择是否启用“长思考”，它可以启用这些能力，也可以直接产出控制。如果对所有大小事件都展开深度推理，那延迟会太大。所以需要时，再延长推理以给出正确答案。

图表 12：特斯拉也采用了类似理想的“快慢”双系统框架，通过 LLM 辅助决策

2. Interpretability and safety guarantees

Natural language can help in “system 2 thinking” cases

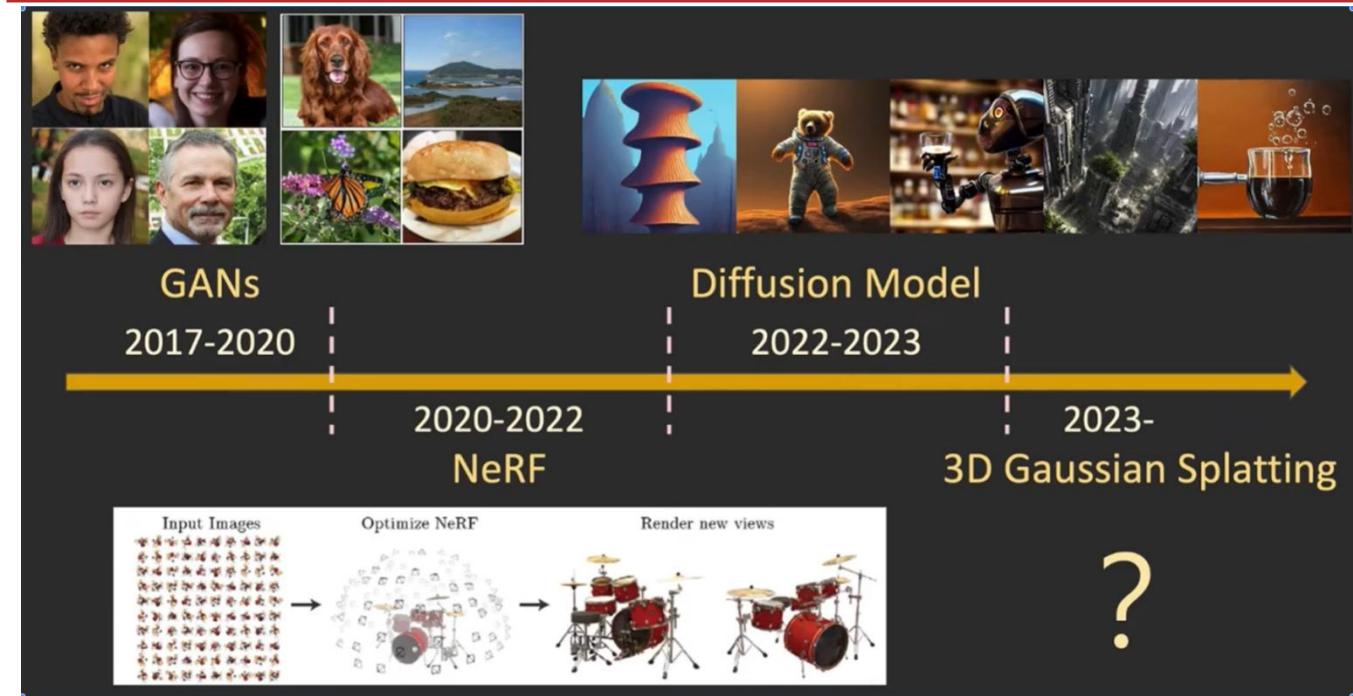


资料来源：AI 工业公众号，华源证券研究所

3. 3. 3D 高斯溅射（3D Gaussian Splatting）

3DGS VS NeRF: 二者是自动驾驶方向热门的两种场景重建方法（场景重建——根据三维场景中拍摄的多张 2D 图片还原该场景的三维模型）。

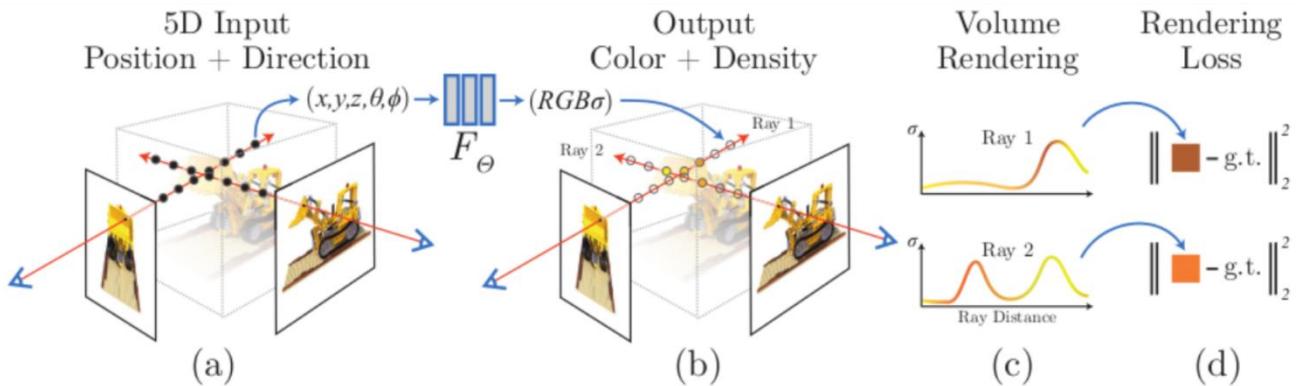
图表 13：基于 2D 图像做 3D 场景重建的算法迭代史



资料来源：CSDN，华源证券研究所

NeRF (Neural Radiance Fields)：三维空间中的任意一个点 (x, y, z)，当我们从任意角度 (θ, ϕ) 观察它时，它的颜色 (RGB) 和体密度 (σ ，跟透明度相关) 是已知的，那么这个三维空间在任意视角下的可视化结果就可以被计算出来了。NeRF 就是通过训练一个神经网络建模来表示这个辐射场 (x, y, z, θ, ϕ) \rightarrow (RGB, σ)；NeRF 并没有直接建模三维世界物体的形状，而是隐式地表示了这个三维世界“看起来是什么样”。

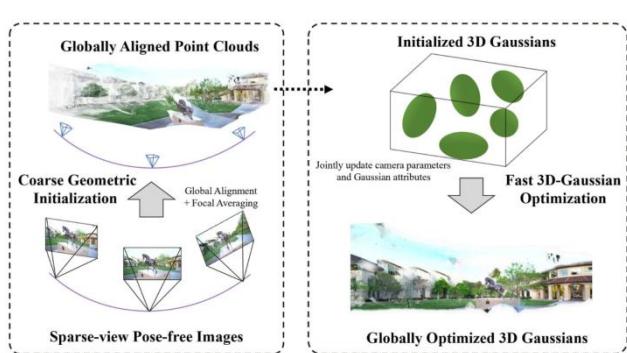
图表 14: NeRF 训练过程



资料来源：腾讯云开发者社区，华源证券研究所

3DGS (3D Gaussian Splatting)：3D 高斯溅射是一种通过显式地存储和优化数十万个可自定义位置、形状和颜色的 3D 高斯球，并利用高效的 Alpha 混合技术进行实时、逼真渲染的 3D 场景重建方法。与 NeRF 的核心区别在于“用光栅化代替光线追踪+把隐式表达变成显式表达”，正好对应它的字面拆分——3D Gaussian (显式表示法) 、Splatting (光栅化，把物体“砸”向相机屏幕看像素有没有被覆盖)。

图表 15: 3DGS 原理



资料来源：Radiance Fields，华源证券研究所

图表 16: 3DGS VS NeRF



资料来源：CSDN，华源证券研究所

较之传统的 3DGS，特斯拉的 3DGS 在相同的受限相机视角下，效果整体好得多（渲染时间大幅缩短+不需要 3D 关键点位的初始化+动态对象渲染+高质量新视角生成+可以附加语义信息）。这样的能力一方面解决了 Tesla 端到端模型训练时候，单纯监督驾驶动作带来的监督信号过于稀疏单一问题，也确保了 Tesla 对周围环境的良好空间理解，类似重建技术也支撑特斯拉实现闭环仿真系统。

图表 17：特斯拉的生成式 3DGS VS 传统 3DGS

2. Interpretability and safety guarantees

Foundation model feedforward generates 3D gaussian splats



	Traditional GS	Tesla's generative GS
Optimization duration	~30 mins	220 ms
Initialization	3D key points	Not required
Dynamic Objects	Difficult	Yes
Novel View Quality	Low	Much higher

资料来源：AI 工业公众号，华源证券研究所

3. 4. 模型评测

特斯拉认为这个环节对于端到端系统来说是最重要的，至少包括以下几个点：

- 无论数据集质量多高，训练 Loss 都不能代表模型性能（需要有更多元的评价体系）
- 开环指标好坏不能保证闭环性能（闭环测试是刚需）
- 自动驾驶存在多种方式来避免驾驶任务失败，评价体系需要捕捉驾驶方式的多元性
- 评估模型性能的方式之一是对驾驶行为造成的结果进行预测

图表 18：特斯拉是如何测评模型的

3. Evaluation

The hardest of the three problems

- Even with a high quality dataset, loss isn't a sufficient indicator of performance
- Good open-loop performance does not guarantee great closed-loop results
- Many different ways to avoid a failure. Metric needs to capture this multimodality
- One way is to evaluate consequence-of-action predictions
- Needs balanced and thorough evaluation sets
- Tedious work, but secretly v.important

资料来源：AI 工业公众号，华源证券研究所

如果只是从车队里随机抽样，大部分都会是枯燥的高速巡航，这不足以评估模型。特斯拉依托强大的数据引擎构建覆盖面均衡的评测集，这非常重要，但也是繁琐的体力活。

一种有效的方法是利用车队数据收集“状态 - 动作”对，据此构建一个世界模拟器——给定过去的状态和采取的动作，模拟器就能条件生成后续的新状态。

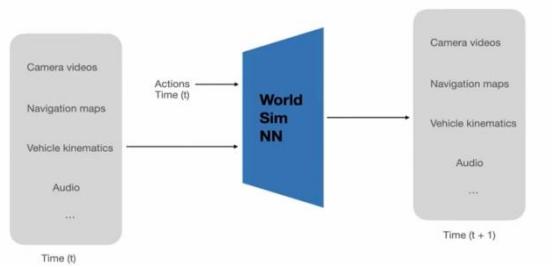
这样的数据很容易收集，而且不需要“最优驾驶”；甚至出错/碰撞的驾驶数据也有价值，因为模拟器需要覆盖边界场景。有了它，就可以把这个模拟器与驱动车辆的策略网络（Policy Engine）连接起来，以闭环方式串联运行，从而模拟整个世界。

图表 19：特斯拉基于“状态-动作”对构建世界模拟器

3. Evaluation

Neural Network closed-loop simulator

Can be trained with cheap to collect state-action data

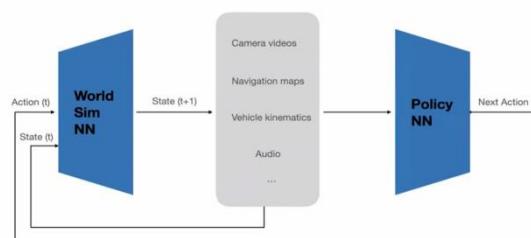


资料来源：AI 工业公众号，华源证券研究所

图表 20：世界模拟器可以与车辆的策略网络闭环运行

3. Evaluation

Can then be used as a closed loop simulator



资料来源：AI 工业公众号，华源证券研究所

世界模拟器的一个例子：基于 8 路 500 万像素视频流由单一神经网络同时生成 1 分 30s 不同角度摄像头视角的视频；且这个网络可以把任何新的指令作为输入，操纵整个网络生成新的画面，跨相机的一致性非常好，甚至车轮轮辋、交通灯等细节都保持一致。

图表 21：基于单一网络生成 8 路摄像头不同视角的长视频



资料来源：AI 工业公众号，华源证券研究所

此外，还能把历史问题拿来重放——比如取一年前的一个案例，用最新的神经网络再跑一遍，看它会如何表现。图表 22 左图是早期模型任务失败的案例，图表 22 右图则是新模型在同一场景下的表现：策略会更早开始横向避让——一旦行人从前车后方露出身位，它就提前变道。这种评估方式非常有用：与其只在新增里程上复测，更好的做法是把历史问题库逐一回放，确认新模型在旧问题上也做得更好。

图表 22：可以用历史问题库中的场景来检验新模式表现



资料来源：AI 工业公众号，华源证券研究所

我们还可以基于相同场景合成新问题：例如图表 23 左图原始视频里，白车一直沿本车道直行；而图表 23 右图，可让白车变道，注入这种对抗性事件来测试系统的反应。可以看到，其余场景保持一致、其他车辆按原样运行，只有这辆车横切——借此就能批量构造合成数据集，验证自动驾驶系统在极端边缘条件下的性能。

图表 23：基于相同场景自由变换场景中的某一元素

3. Evaluation

Evaluate policy models on synthetic adversarial issues



资料来源：AI 工业公众号，华源证券研究所

如果愿意稍微降低测试时的计算量，就能获得接近实时的渲染。还是同一个模型渲染 8 路相机，只是在测试端降低了算力占用；会发现它可以实时驾驶：即便在生成 8 路 500 万像素的视频流，帧率仍然足够高。显然它会响应用户的驾驶指令：转向、制动、在虚拟世界里像现实驾驶那样导航——但整个画面都是神经网络生成的视频流。仿真时还可以特地把车开上路沿，来检验极端场景的模拟效果，体现了模拟器出众的泛化能力。整段视频长 6 分钟，可以在很长时间里稳定生成并对控制指令做出响应。由此可见，这个工具既强于评测，也适合做闭环强化学习：让车在其中长时间驾驶，并验证它是否长时间无碰撞。

图表 24：降低计算精度就可以获得接近实时渲染的仿真效果



资料来源：AI 工业公众号、华源证券研究所

4. 未来规划

特斯拉 FSD 的神经网络模型可迁移到 Robotaxi&其他机器人上：例如 Optimus，模拟其在特斯拉工厂内自主导航，所有画面都是生成视频，跨视角保持了高度一致性，且我们可以类似 FSD 对车的云端模拟一样，对机器人/场景内的其他元素做动作变更，模型就能条件生成与该动作相一致的像素序列。用的还是同一套神经网络，只是在训练中加入了更多来自 Optimus 的数据，就能泛化到其他机器人形态。

图表 25：FSD 的神经网络模型迁移到 Optimus 上

Neural Simulation models from FSD scale to Optimus



资料来源：AI 工业公众号、华源证券研究所

5. 总结

Ashok Elluswamy 关于 Tesla Foundational Model for FSD 的分享基本阐明了以下内容：

1) 特斯拉选择“端到端框架”的核心原因：模块化框架难以编码人类价值观、感知与预测规划之间的接口定义会带来信息传递损失、端到端能应对现实世界的长尾问题/同构计算带来的确定性延迟，端到端也能更好地契合 AI 领域的 Scaling-law。

2) 端到端面临的主要挑战以及特斯拉如何解决：维度灾难（特斯拉采用复杂的触发机制来回传长尾场景数据）、怎么保证模型的可解释性与安全性（允许模型输出多样的中间结果（例如占用、其它目标、交通信号灯、交通标志、道路边界等），这些中间结果可通过车机渲染展示给驾驶员与乘客）；通过类似理想、小鹏的 VLA/VLM 框架，更好解释为什么做出某个决策&应对长尾场景、如何评估模型（多元化的评价体系、重视闭环测试、评估模型决策可能造成的结果、复用历史问题库测评新模式；特斯拉构建了世界模拟器这一强大的测评仿真工具）。

3) 模型具备强大的可迁移性：Robotaxi&Optimus 实质上复用了同一套 FSD 神经网络，只是在训练数据中增加了相应数据，模型便能泛化到其他机器人形态。

基于以上我们还可以做一些推断：

- 1) 乘用车、Robotaxi、人形机器人以及其他类型机器人可以复用相同的感知决策网络；
- 2) 理想、小鹏选择的 VLA/VLM 路线与特斯拉目前的技术路线较为吻合；
- 3) 端到端可以允许存在多种传感器编码器、中间结果、显式表达，重要的是保证梯度的端到端贯穿；
- 4) 模型测评非常重要+数据&筛选长尾数据的能力非常重要+世界模拟器/世界模型是强大的仿真测试工具。

6. 投资建议

建议关注：数据积累丰富+技术路线选择近似特斯拉的头部智驾整车公司（理想、小鹏）、核心环节 Tier1（线控转向（耐世特、浙江世宝））、线控制动（伯特利）、大算力域控制器（德赛西威、均胜电子、经纬恒润、中科创达）、增量零部件（激光雷达（禾赛科技、速腾聚创、图达通）、大算力芯片（地平线机器人、黑芝麻智能））、Robotaxi 运营（文远知行、小马智行、曹操出行、千里科技）、RoboVan（九识智能、新石器）等。

7. 风险提示

- (1) **法规进展不及预期风险：**如 Robotaxi、自动驾驶重卡等相关法规/政策放开速度低于预期，可能导致市场规模增长不及预期。
- (2) **竞争格局恶化风险：**如行业主要供应商为扩大市场份额采取更激进的产品报价策略，可能引发行业市场规模与盈利水平可能不及预期。
- (3) **技术路线迭代风险：**目前智驾算法尚处于快速迭代阶段，技术路线不统一，仍可能出现颠覆性的技术变化导致行业大洗牌。

证券分析师声明

本报告署名分析师在此声明，本人具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，本报告表述的所有观点均准确反映了本人对标的证券和发行人的个人看法。本人以勤勉的职业态度，专业审慎的研究方法，使用合法合规的信息，独立、客观的出具此报告，本人所得报酬的任何部分不曾与、不与，也不将会与本报告中的具体投资意见或观点有直接或间接联系。

一般声明

华源证券股份有限公司（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。

本报告是机密文件，仅供本公司的客户使用。本公司不会因接收人收到本报告而视其为本公司客户。本报告是基于本公司认为可靠的已公开信息撰写，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、工具、意见及推测等只提供给客户作参考之用，并非作为或被视为出售或购买证券或其他投资标的的邀请或向人作出邀请。该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。客户应对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特殊需求，必要时就法律、商业、财务、税收等方面咨询专家的意见。对依据或使用本报告所造成的一切后果，本公司及/或其关联人员均不承担任何法律责任。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

本报告所载的意见、评估及推测仅反映本公司于发布本报告当日的观点和判断，在不同时期，本公司可发出与本报告所载意见、评估及推测不一致的报告。本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。除非另行说明，本报告中所引用的关于业绩的数据代表过往表现，过往的业绩表现不应作为日后回报的预示。本公司不承诺也不保证任何预示的回报会得以实现，分析中所做的预测可能是基于相应的假设，任何假设的变化可能会显著影响所预测的回报。本公司不保证本报告所含信息保持在最新状态。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。

本报告的版权归本公司所有，属于非公开资料。本公司对本报告保留一切权利。未经本公司事先书面授权，本报告的任何部分均不得以任何方式修改、复制或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。如征得本公司许可进行引用、刊发的，需在允许的范围内使用，并注明出处为“华源证券研究所”，且不得对本报告进行任何有悖原意的引用、删节和修改。本公司保留追究相关责任的权利。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

本公司销售人员、交易人员以及其他专业人员可能会依据不同的假设和标准，采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论或交易观点，本公司没有就此意见及建议向报告所有接收者进行更新的义务。本公司的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

信息披露声明

在法律许可的情况下，本公司可能会持有本报告中提及公司所发行的证券并进行交易，也可能为这些公司提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。本公司将会在知晓范围内依法合规的履行信息披露义务。因此，投资者应当考虑到本公司及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突，投资者请勿将本报告视为投资或其他决定的唯一参考依据。

投资评级说明

证券的投资评级：以报告日后的 6 个月内，证券相对于同期市场基准指数的涨跌幅为标准，定义如下：

买入：相对同期市场基准指数涨跌幅在 20%以上；

增持：相对同期市场基准指数涨跌幅在 5% ~ 20%之间；

中性：相对同期市场基准指数涨跌幅在 -5% ~ +5%之间；

减持：相对同期市场基准指数涨跌幅低于-5%及以下。

无：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。

行业的投资评级：以报告日后的 6 个月内，行业股票指数相对于同期市场基准指数的涨跌幅为标准，定义如下：

看好：行业股票指数超越同期市场基准指数；

中性：行业股票指数与同期市场基准指数基本持平；

看淡：行业股票指数弱于同期市场基准指数。

我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议；

投资者买入或者卖出证券的决定取决于个人的实际情况，比如当前的持仓结构以及其他需要考虑的因素。投资者应阅读整篇报告，以获取比较完整的观点与信息，不应仅仅依靠投资评级来推断结论。

本报告采用的基准指数：A 股市场基准为沪深 300 指数，香港市场基准为恒生中国企业指数（HSCEI），美国市场基准为标普 500 指数或者纳斯达克指数。