

AI+汽车智能化系列之五—— 智驾算法步入深水区，头部玩家有望持续领跑

汽车行业证券分析师：黄细里
执业编号：S0600520010001
联系邮箱：huanxl@dwzq.com.cn
联系电话：021-60199790

汽车行业证券分析师：杨惠冰
执业编号：S0600523070004
联系邮箱：yanghb@dwzq.com.cn

2024年5月31日

- **何谓“大模型”？大数据喂养神经网络算法，设定规则机制使其自成长。** 1) 多模态数据是基础，包括文本/图像/音频/视频等在内的多类型数据喂养，驱动算法更好完成理解/生成等任务。2) **Transformer是核心**，Self-Attention机制强化算法抽象特征提取能力，并支持并行计算，高能且高效，衍生ViT/DiT支持多模态数据理解/生成。3) **ChatGPT及Sora为代表应用**，在大参数加持下，开发多模态自然语言处理以及文生视频等功能。4) **世界模型为未来方向**，算法能力由数据驱动演变为认知驱动，模型具备反事实推理和理解物理客观规律的能力，提升通用泛化特性。
- **大模型重塑车端算法架构，加速云端算法迭代；世界模型或为完全自动驾驶最优解。** 智驾一阶段（L2~L3）**脱胎换骨**：由场景驱动转向数据驱动，大模型带来底层架构质变；智驾二阶段（L3~L4）**厚积薄发**：由数据驱动转向认知驱动，数据和算力逐步累计驱动能力提升，量变引起质变。1) **车端**：上层感知/规控应用层算法随功能需求提升持续进化，“场景理解处理能力泛化”是核心诉求。L2~L3，感知端为升级核心，Transformer加持BEV+ 占用网络算法落地感知端到端，解决长尾场景识别难题；L3~L4以规控算法升级为核心，精准识别并快速处理，Learning-base 逐步取代Rule-base，端到端拉高场景处理能力的天花板。2) **云端**：数据闭环为前提，加速大数据有效利用，采集/标注/仿真/训练/部署一体化。Transformer赋能自动标注，数据驱动场景仿真泛化，降低对有限的实际路测数据的依赖。3) **世界模型【通用具身智能】或为自动驾驶最优解。** 车端场景生成泛化，将自动驾驶问题转化为预测视频的下一帧，类人模式处理，实现泛化至数据场景以外的能力；并可快速生成标准化仿真数据，结合大算力加速云端训练。
- **算法实现高壁垒+数据闭环硬要求，未来智驾算法产业格局趋于集中化。** 当前下游L3+高阶算法方案以OEM自研为主，华为以“算法赋能，数据回传”的形式深度绑定OEM；L3以下算法呈现OEM+独立算法商+硬件商三足鼎立格局。我们认为，考虑高阶智驾功能的提升对于算法能力/组织架构/超算中心/完整数据链等的要求，未来“**掌握硬件的基础上去发展软件**”或为**主流**，即**掌握壁垒最高的硬件——芯片；提供性价比最高的硬件——传感器；掌握粘性最强的硬件——整车。**

■ 投资建议：汽车AI智能化转型大势所趋，算法为主干，看好头部算法玩家持续领先铸就高壁垒。

- **全行业加速智能化转型，产业趋势明确。** 下游OEM玩家+中游Tier供应商均加大对汽车智能化投入，大势所趋；智驾核心环节【软件+硬件+数据】均围绕下游OEM展开，数据催化算法提效进而驱动硬件迭代。以特斯拉为代表，应用算法向全栈端到端-世界模型持续迭代，功能落地兑现。
 - OEM整车厂商&核心芯片硬件厂商&智驾传感器厂商&独立算法商加速布局端到端算法开发，场景驱动-数据驱动-认知驱动持续进化；智驾算法产业发展进入深水区，高投入赋能【大算力+大数据】，**方能走通L3有条件自动驾驶至L4完全自动驾驶之路。**
 - **看好智驾头部车企以及智能化增量零部件：** 1) 华为系玩家【长安汽车+赛力斯+江淮汽车】，关注【北汽蓝谷】； 2) 头部新势力【小鹏汽车+理想汽车】； 3) 加速转型【吉利汽车+上汽集团+长城汽车+广汽集团】； 4) **智能化核心增量零部件：** 域控制器（德赛西威+经纬恒润+华阳集团+均胜电子等）+线控底盘（伯特利+耐世特+拓普集团等）。
- **风险提示：** 智能驾驶相关技术迭代/产业政策出台低于预期；华为/小鹏等车企新车销量低于预期。



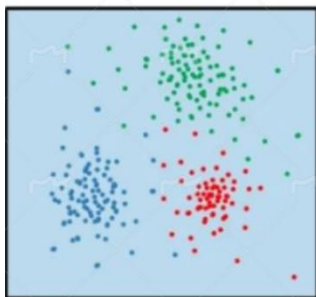
- 一、何谓“大模型”？
- 二、车端：大模型重塑智驾算法架构
- 三、云端：大模型加速智驾算法迭代
- 四、当前产业玩家的智驾算法能力如何？
- 五、投资建议与风险提示

一、何谓“大模型”？

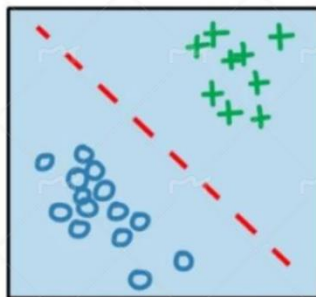
图：机器学习

machine learning

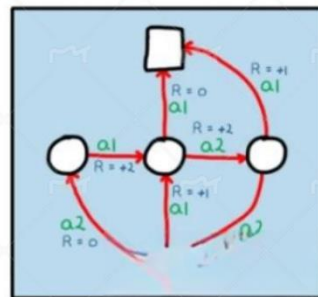
unsupervised learning



supervised learning



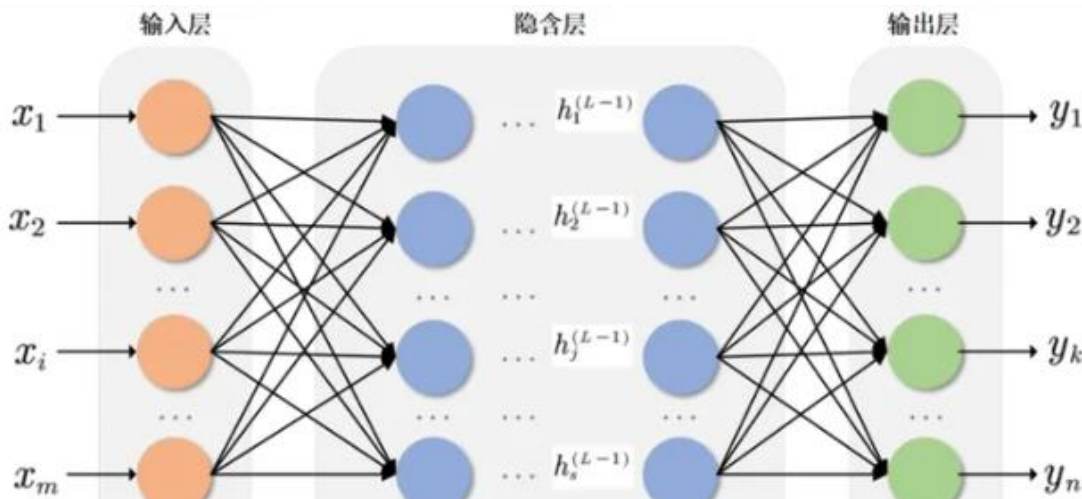
reinforcement learning



机器学习 > 神经网络 > 深度学习 ≈ 深度神经网络。机器学习用于解决由人工基于 if-else 等规则开发算法而导致成本过高的问题，想要通过帮助机器“发现”它们“自己”解决问题的算法来解决；机器学习可以分为有监督学习、无监督学习和强化学习等三类。

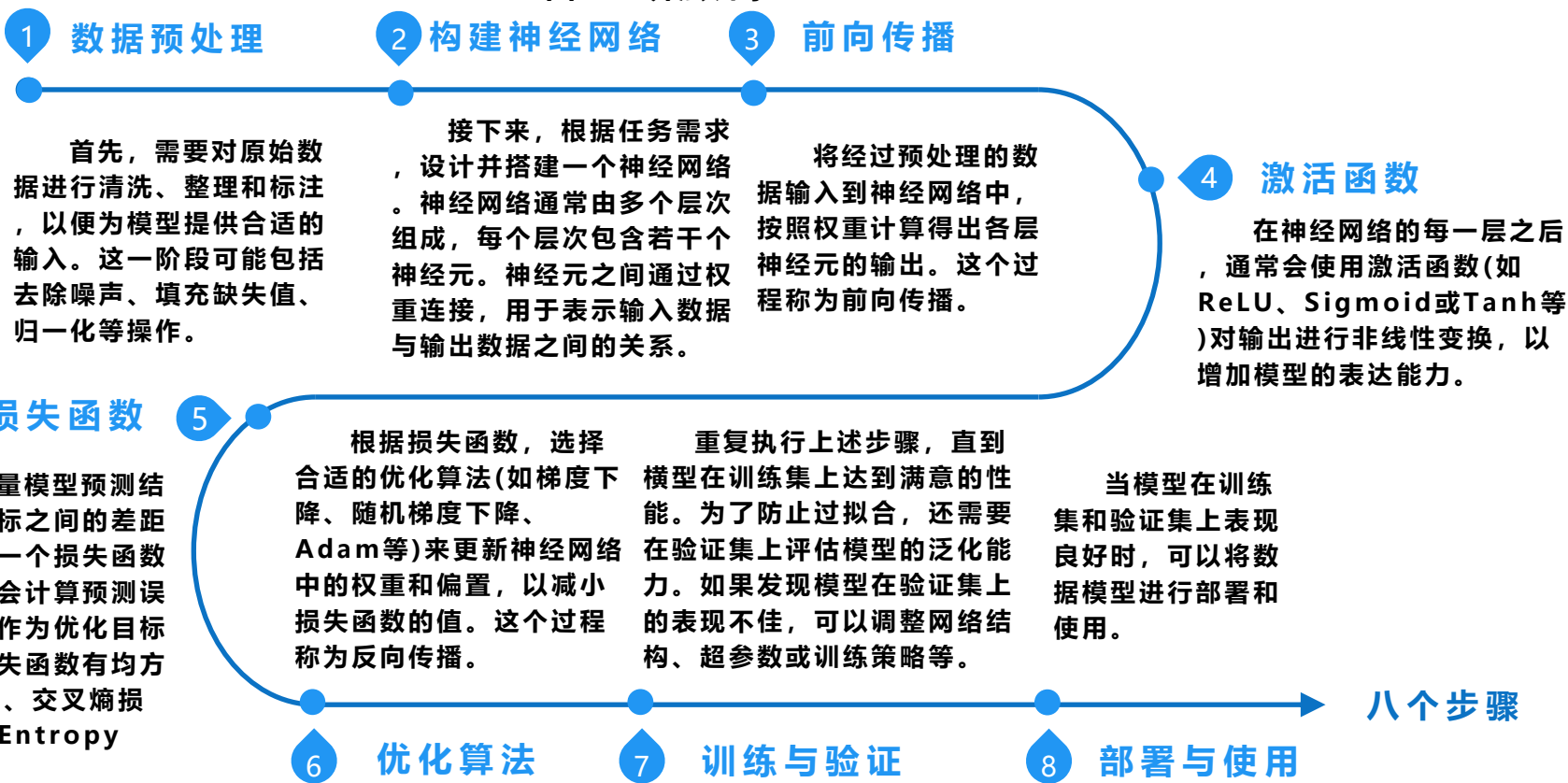
深度学习是基于深度神经网络的，而神经网络算法是机器学习模型的一个分支，包括卷积神经网络 CNN/循环神经网络 RNN 等等，自注意力机制 (Transformer) 则是基于全连接神经网络和循环神经网络的衍生。

深度学习使用多层神经网络，从原始输入中逐步提取更高层次更抽象的特征用于后续算法识别，处理大规模数据是其核心优势。当前，深度学习已经应用到包括图像识别、自然语言处理、语音识别等各领域。



AI大模型是指具有超大规模参数（通常在十亿个以上）、超强计算资源的机器学习模型，其目标是通过增加模型的参数数量来提高模型的表现能力，它们能够处理海量数据，完成各种复杂任务。AI大模型的原理是基于神经网络和大量数据的训练，模型通过模拟人脑的神经元结构，对输入数据进行多层抽象和处理，从而实现对复杂任务的学习和预测。AI大模型的训练主要分为：**数据预处理、模型构建、模型训练、模型评估**等几大步骤，如下：

图：AI 算法训练



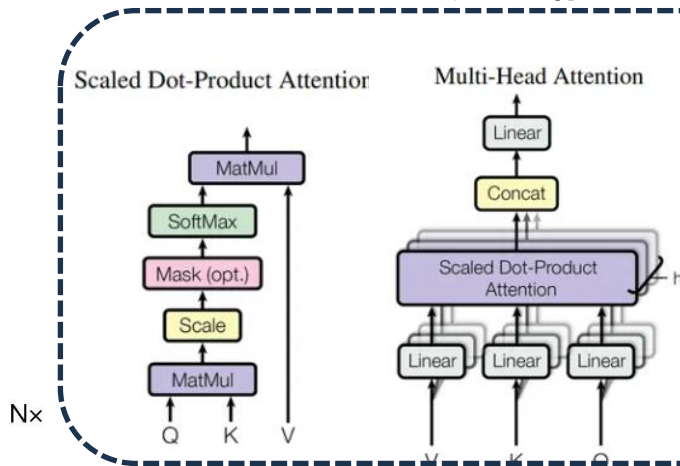
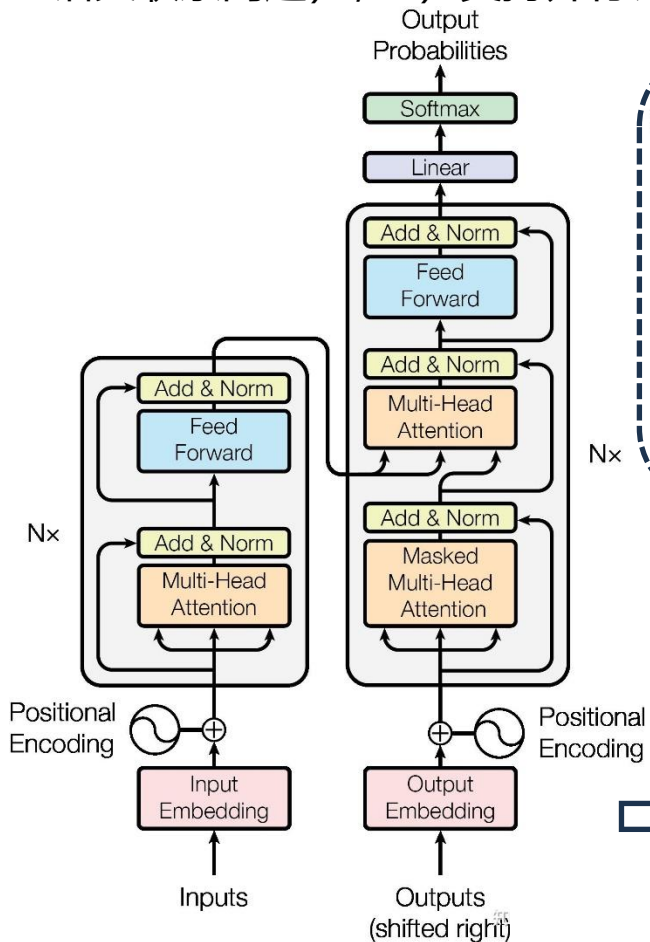
AI大模型能够处理以下几类核心问题：**1) 自然语言处理**：以GPT-3和BERT为例，AI大模型通过学习海量的语料库和上下文，让计算机更加准确地理解和处理自然语言，如翻译、问答、分词、文本生成等领域。**2) 计算机视觉**：以ResNet和EficientNet为例，AI大模型通过学习大量的图像数据和构建更深更复杂的神经网络，使计算机能够对图像进行更加准确的识别和分析，包括目标检测、图像分类、语义分割等领域。**3) 语音识别和生成**。通过以上几类问题的解决，AI大模型可以进一步通过自动化和智能化的方式提高生产效率，在部分工业领域可以实现人机合作或自动化，减少人力成本。

图：AI大模型的优劣势分析



- **Transformer**是本轮大模型颠覆全行业算法架构的核心，也是多模态模型相比之前单一文本理解模型提升的关键，Transformer赋予算法更精准的特征提取能力，强化理解和识别功能，其独特的**自注意力机制**是灵魂，即**Attention is all you need**。
- **Transformer的优势在于**：1) 自注意力机制赋予的长依赖语义问题（捕捉间隔较远的词之间的语义联系问题）；2) 支持并行计算，可极大的提升大模型数据处理效率。

图：Transformer的原理机制



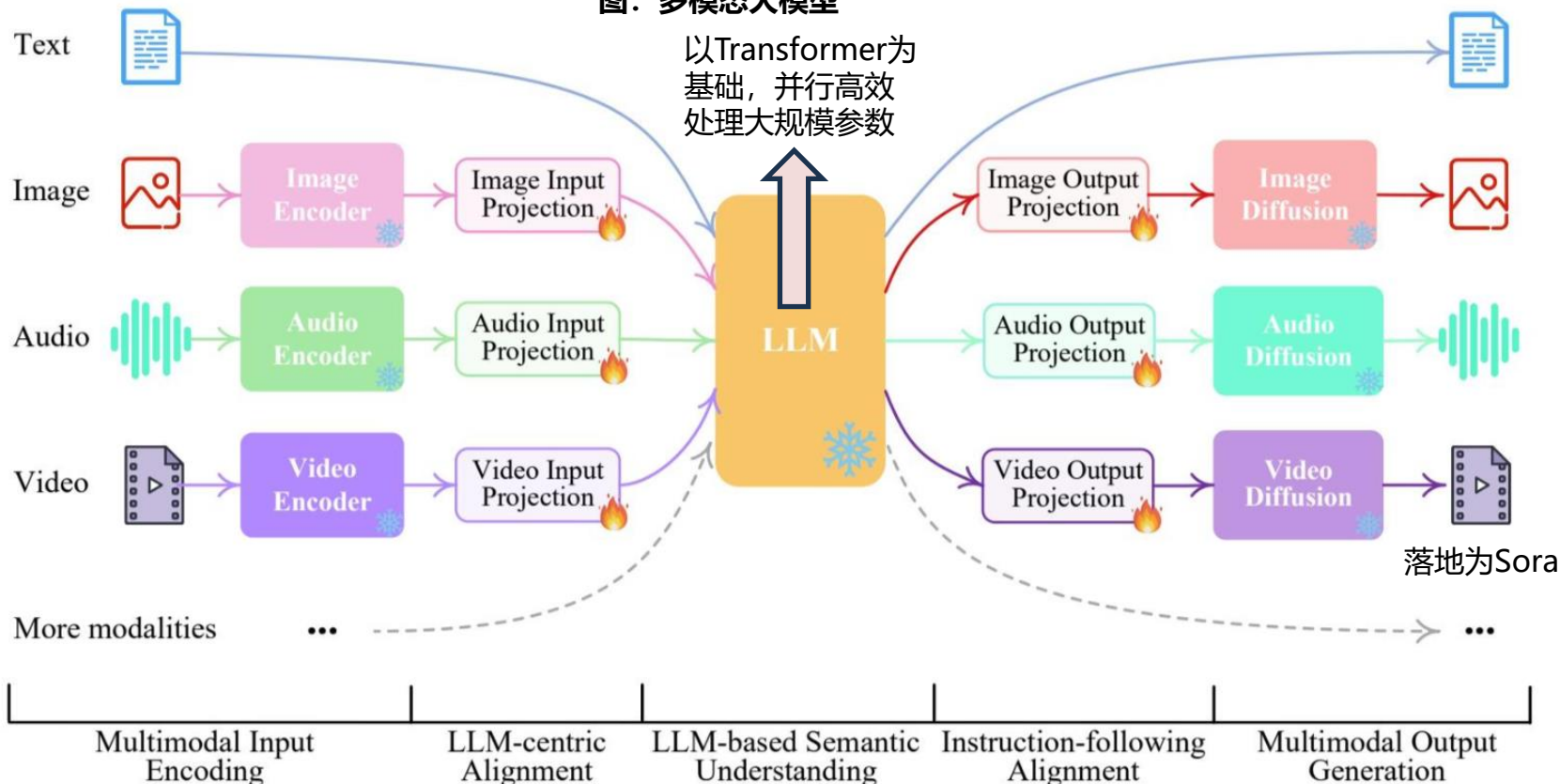
自注意力机制：查询/定位，赋予权重（即注意力），多头即多个维度计算相似度特征，保证充分性。

优势：自注意力机制不同于RNN，无主次先后，可支持并行计算，同时利用位置编码和权重赋予进行定位和筛选。

- input embedding是编码器的输入，
- output embedding是解码器的输入
- Positional Encoding代表位置编码，记住顺序
- Encoder和Decoder分别有6层 (blocks)，Encoder负责特征提取，Decoder则利用特征来完成识别、分类、回归等任务
- 解码后输出结果

- **由理解内容至生成内容，多模态大语言模型持续进化。**多模态模型是一种能够处理多种类型数据（如文本、图像、音频和视频）的人工智能模型。这种模型的目标是通过结合不同类型的数据来提供更全面、更准确的信息。在自然语言处理（NLP）领域，多模态模型可以用于机器翻译、情感分析、文本摘要等任务。在计算机视觉领域，多模态模型可以用于图像分类、目标检测、人脸识别等任务。**多模态大语言（MM-LLMs）**即是将多模态模型与具备强大推理和生成能力的大语言模型结合的产物，其难点在于**如何对齐本不兼容的图像/视频/文本等的编码器**。

图：多模态大模型

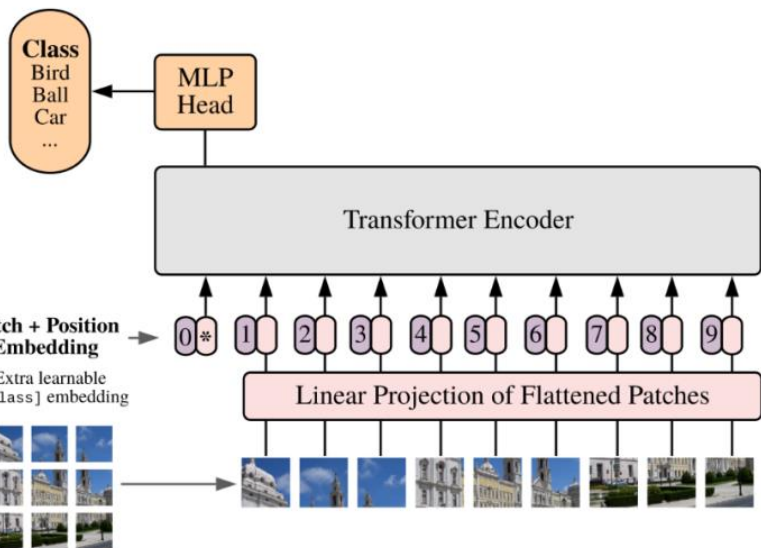


■ **视觉表征主框架由CNN切换Transformer，即ViT**，其**多头自注意力机制**赋予模型**并行高效计算**以及**把握前后长时间依赖关系**的能力，能够同时接收来自文本/图像/音频/视频的特征Token，并接入全链接层服务于下游的分类任务。**ViT成功的秘诀在于大量的数据做预训练**，如果没有这个过程，在开源任务上直接训练，其效果仍会逊色于具有更强归纳偏置的CNN网络。

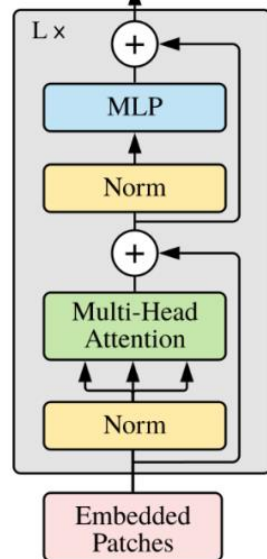
■ ViT步骤分为三大步：

- 图形切块 Patch Embedding；位置编码 Position Embedding；
- **特征提取** Class Token；**注意力权重**赋予 Transformer Encoder；
- 多头输出 MLP Head。

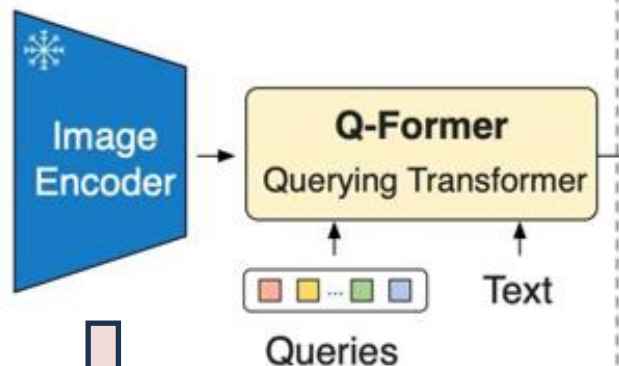
Vision Transformer (ViT)



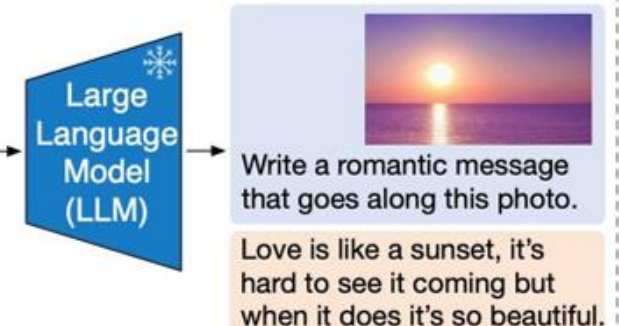
Transformer Encoder



Vision-and-Language Representation Learning



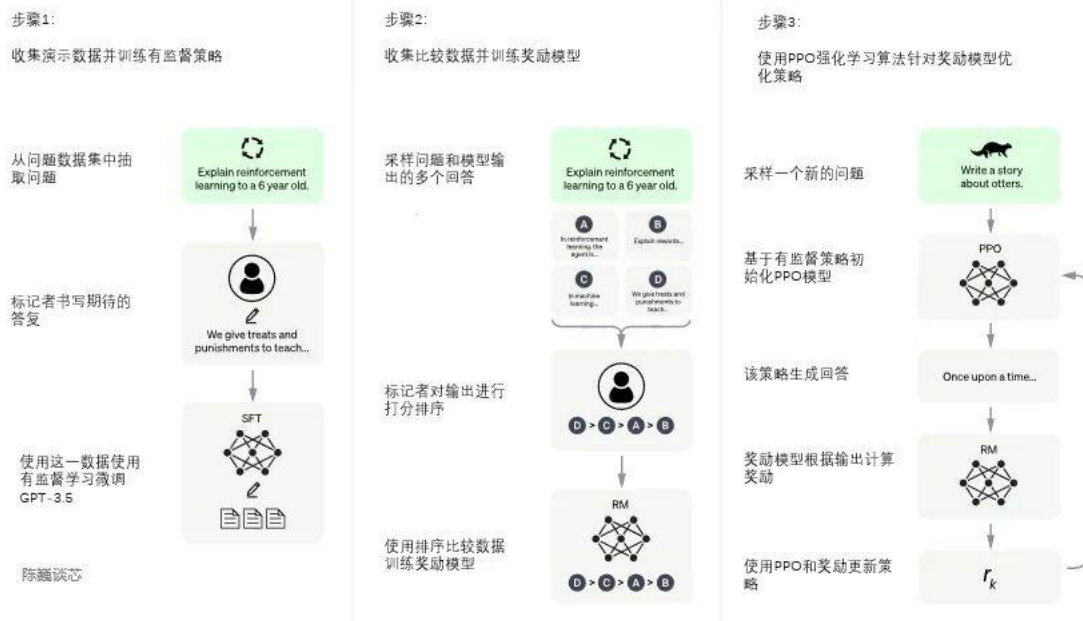
Vision-to-Language Generative Learning



- **GPT1**: 用Transformer的解码器和大量的无标签样本去预训练一个语言模型，然后在子任务上提供少量的标注样本做微调，就可以很大的提高模型的性能。
- **GPT2**: **Zero-shot**，在子任务上不去提供任何相关的训练样本，而是直接用足够大的预训练模型去理解自然语言表达的要求，并基于此做预测。但GPT2性能差，有效性低。
- **GPT3**: **few-shot learning**，兼顾少样本和有效性。用有限的样本，模型可以迅速学会任务。
- **GPT4**: GPT1~3本质还是通过海量的参数学习海量的数据，然后依赖transformer强大的拟合能力使得模型能够收敛，因此不具备文本生成能力。ChatGPT则在此基础上，依赖**指令学习 (Instruction Learning)** 和**人工反馈强化学习 (RLHF)** 进行训练，对原有模型进行有监督的微调（人工优化数据集）+强化学习对模型生成结果进行打分，提高泛化能力，在此基础上loss持续迭代，生成正确结果。相比GPT3规模大幅提升，从1750亿提升至1.8万亿，算力需求大幅提升。

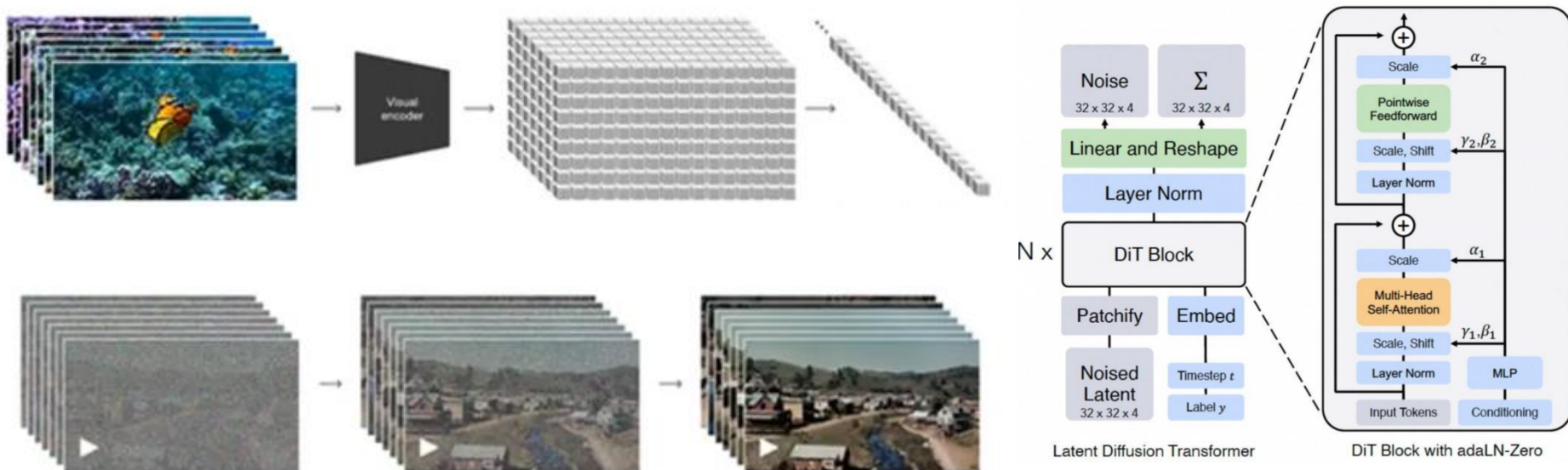
■ **GPT4o**: **完全统一多模态**。o即 omni，意为“全体”、“所有”或“全面的”，打通多模态间输入交互，延迟降低并完全开放免费，并进一步增强推理能力。但其依然是基于Transformer架构去实现对于模态信息的理解和生成，因此并未有底层架构的创新。

图：ChatGPT模型的训练过程



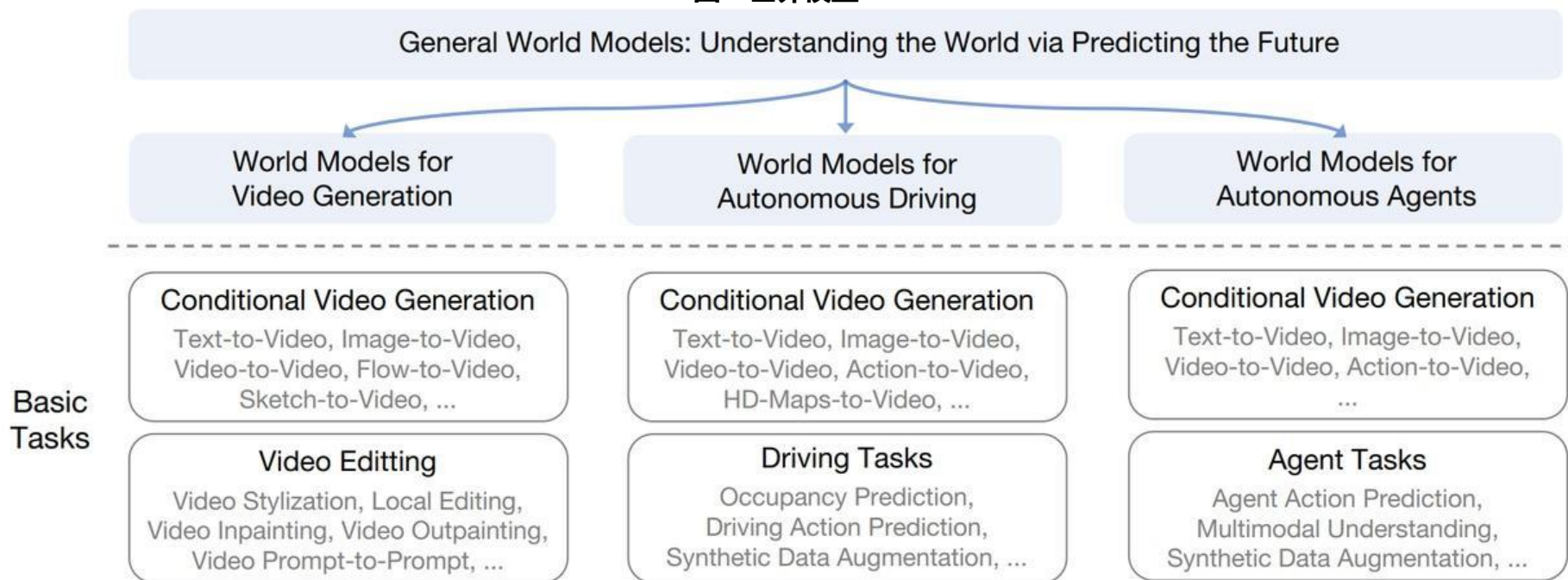
- Sora模型充分利用**扩散模型 (Diffusion Model)** 精细化生成能力以及**Transformer**的前后文全局关系捕捉能力，实现视频每一帧的图像精确生成以及前后的时空一致性。
- Sora可以理解为是**Open AI大模型技术的完全集成**，其原理可以分为三步：1) 首先，模型将视频压缩到**低维潜在空间**中，然后将其分解为patch（类似于GPT中的Token），从而将视频完全压缩。2) 其次，视频patch在低维空间中训练，**扩散模型**通过在训练数据上逐步添加高斯噪声并学习如何逆向去除噪声的过程来生成新数据，并整合了Transformer的多头注意力机制和自适应层归一化，提升扩散模型在处理连续大规模视频数据时的**稳定性和一致性**。3) 最后，模型通过对应解码器，将生成的元素**映射回像素空间**，完成视频生成任务。
- Sora的核心**DiT模型**：**Transformer架构取代原有UNet架构**（本质为**CNN卷积神经网络**），提升长时间性能的同时具备可扩展性+灵活性，并支持跨领域扩展，或可应用至自动驾驶领域。

图：Sora文生视频，DiT模型为基础原理



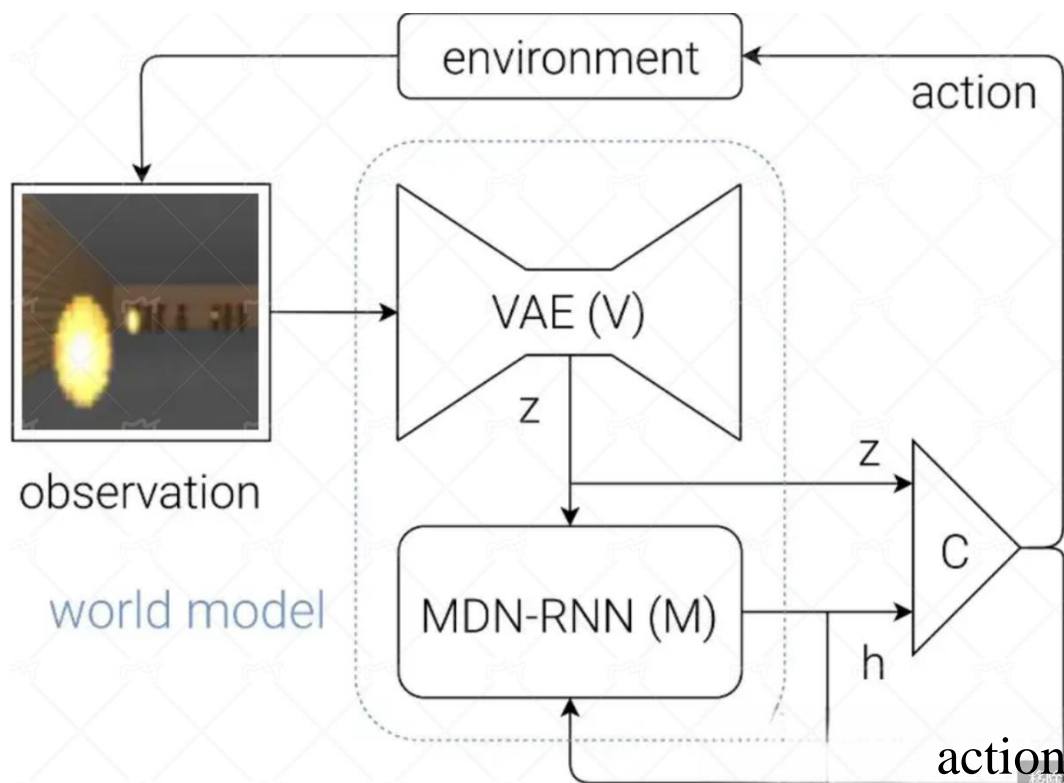
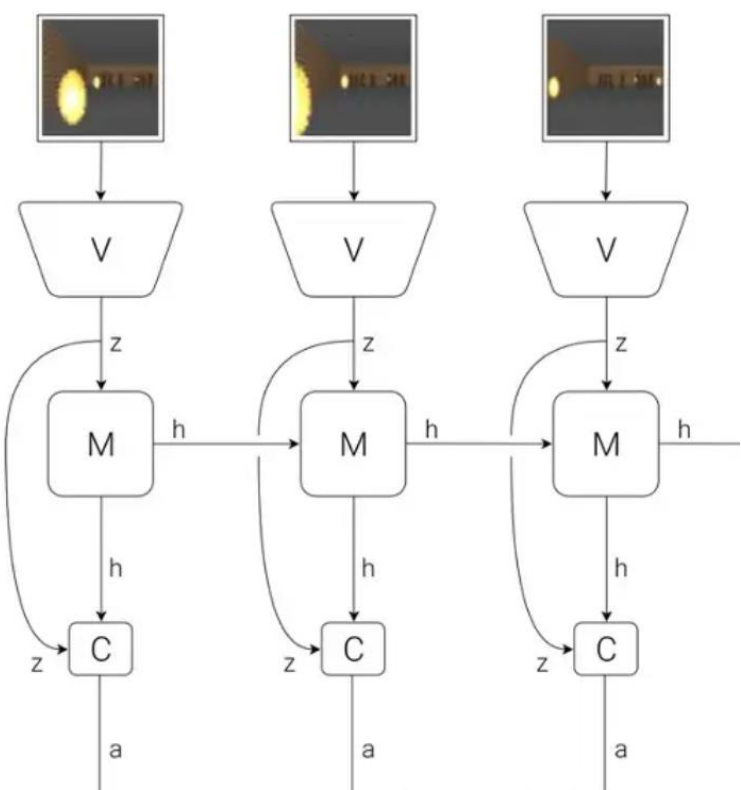
- **视频生成Video Generation是世界模型World Model的基础，Sora的成功使得DiT扩散模型收敛为视频生成的主导方式。** 基于对历史信息的理解生成对于未来的预测进而合理表征，并结合可能的行为action进一步辅助预测可能的结果，可应用于包括直接的视频生成以及自动驾驶和机器人等通用人工智能多个领域。
- **基于数据，超越数据。** 传统大模型要适应新的精细任务时，必须基于数据对模型参数进行全面微调，依赖【预训练】环节，不同任务与不同数据一一对应，模型专业化。相比过往大模型强调的模态信息理解的能力，世界模型更加注重【因果和反事实推理、模拟客观物理定律】等规划和预测的能力，并具备**强泛化和高效率**等性能表现。

图：世界模型



- 传统的多子函数模块化算法演变为世界模型数据闭环训练，持续优化提升算法认知。世界模型在进行端到端训练的过程为：
- 观测值经过V（基于Transformer的Vision Model，图中为早先采用的VAE架构）提取feature，然后经过M（基于LSTM的Memory）得到h（预测值），最后预测值和历史信息合并至C（Controller）得到动作，基于动作和环境交互的结果产生新的观测值，保障车辆做出合理行为。

图：传统模块化算法架构【V+M+C】转变为端到端【VM输出C】



二、车端：大模型重塑智驾算法架构

- **底层调动支持：**1) **异构化底层软件配合SoC异构芯片支持不同类型的功能实现落地。**用于智驾域控的SoC异构芯片是高度集成化的，除多核CPU外，还会集成DSP/NPU或其它专用计算加速单元。在此硬件上进行匹配的异构软件，多核CPU经过虚拟化之后，可以在不同的虚拟机上运行不同的操作系统，例如运行Linux来充分利用Linux社区丰富的软件生态，运行QNX/VxWorks 来达到更好的实时性。2) **SOA中间件提供标准化接口服务。**SOA中间件能屏蔽不同服务间软硬件异构平台的差异性，一方面让异构平台都能够以SOA的方式交互，另一方面让通用的功能逻辑能在不同硬件平台之间移植。比如将DSP或NPU的使用封装在特定的库中，对外呈现为标准SOA服务。3) **软件框架：**即可复用的设计构件，它规定了应用的体系结构，代表已经完成项目的底层开发基础搭建，可以在其基础上差异化开发的半成品。可以让使用者减少很多重复的代码、让代码的结构更加清晰，耦合度更低，后期维护方便。
- **上层应用赋能：算法：**即解决某一类问题的特定策略机制，不同算法解决同一问题时的质量优劣不同，这将影响程序的效率。一个算法的评价主要从时间复杂度和空间复杂度来考虑。



图：汽车智能驾驶软件层级架构

多自研，匹配上层应用算法

多采用第三方成熟系统

- **What:** 传统汽车时代，狭义的操作系统指操作系统内核；当前智电汽车域控时代，系统更加复杂，因此需要基于内核进行大量工程优化以后方可直接开发上层应用软件，中间件和功能软件即由工程优化算法标准化后所得。当前广义操作系统包括内核以及中间件和功能软件等几部分。
- **How:** 高性能、高安全、跨平台、高效开发应用。操作系统要求较强实时性，系统任务调度时钟周期在毫秒级；且提供高可靠性和较强功能安全。未来，行业舱驾一体化等集中式EE架构还要求操作系统保证合理共享算力资源，支持系统快速OTA迭代。
- **底层支持软件相对标准化，通用性较强。** 1) **操作系统内核 (Kernel) :** 即为狭义操作系统，如OSEK OS、VxWorks、RT-Linux等。内核提供操作系统最基本的功能，负责管理系统的进程、内存、设备驱动程序、文件和网络系统，决定着系统的性能和稳定性。2) **中间件:** 处于应用和操作系统之间的软件，实现异构网络环境下软件互联和互操作等共性问题，提供标准接口、协议，具有较高的移植性，如POSIX/ARA（自适应AutoSAR运行时环境即中间件API接口）和DDS（分布式实时通信中间件）。3) **功能软件:** 主要指自动驾驶的核心共性功能模块。

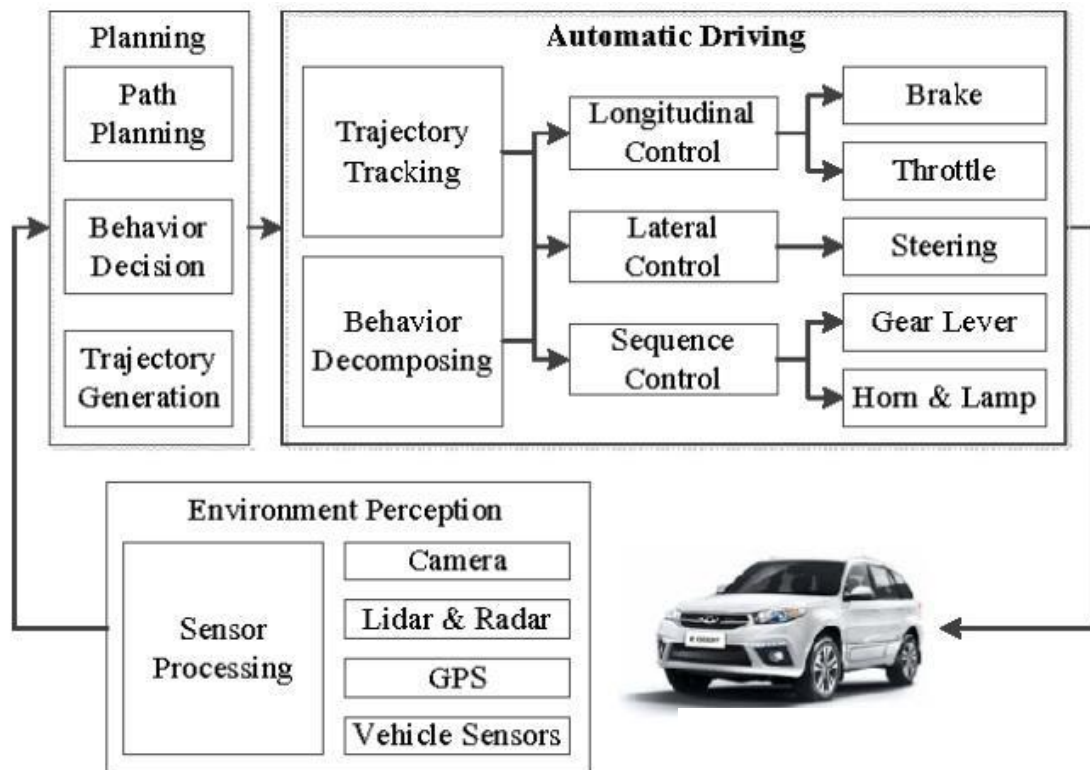


图：汽车操作系统

主流操作系统	特点
QNX	可靠性很好，协议栈、各种外设驱动稳定，只是运行所需资源有些多
Linux	实时性较差，但软件资源内容丰富
VxWorks	开发工具功能强大，使用方便，但是价格昂贵

- 车端上层应用算法分为感知 (Perception) /规控 (Planning and Decision) /执行 (Motion and control) 三大模块。感知模块的输入可以是各种传感器：摄像头、激光雷达、毫米波雷达等，同时感知也会接受车身传感器的信息来对感知结果进行修正。决策规划模型又分为行为决策、轨迹规划、轨迹生成等模块。控制模块主要分为横向控制与纵向控制，横向控制主要是控制转向系统，纵向控制涉及油门与刹车的控制。
- 模块化的软件算法早期具备可解释性强、安全性强、灵活性高的优点，但过于依赖先验数据，无法实时更新，且穷举法终究没办法覆盖所有长尾场景，因此随智驾等级提升，正逐渐被大模型端到端所取代。

图：智能驾驶应用层算法分类

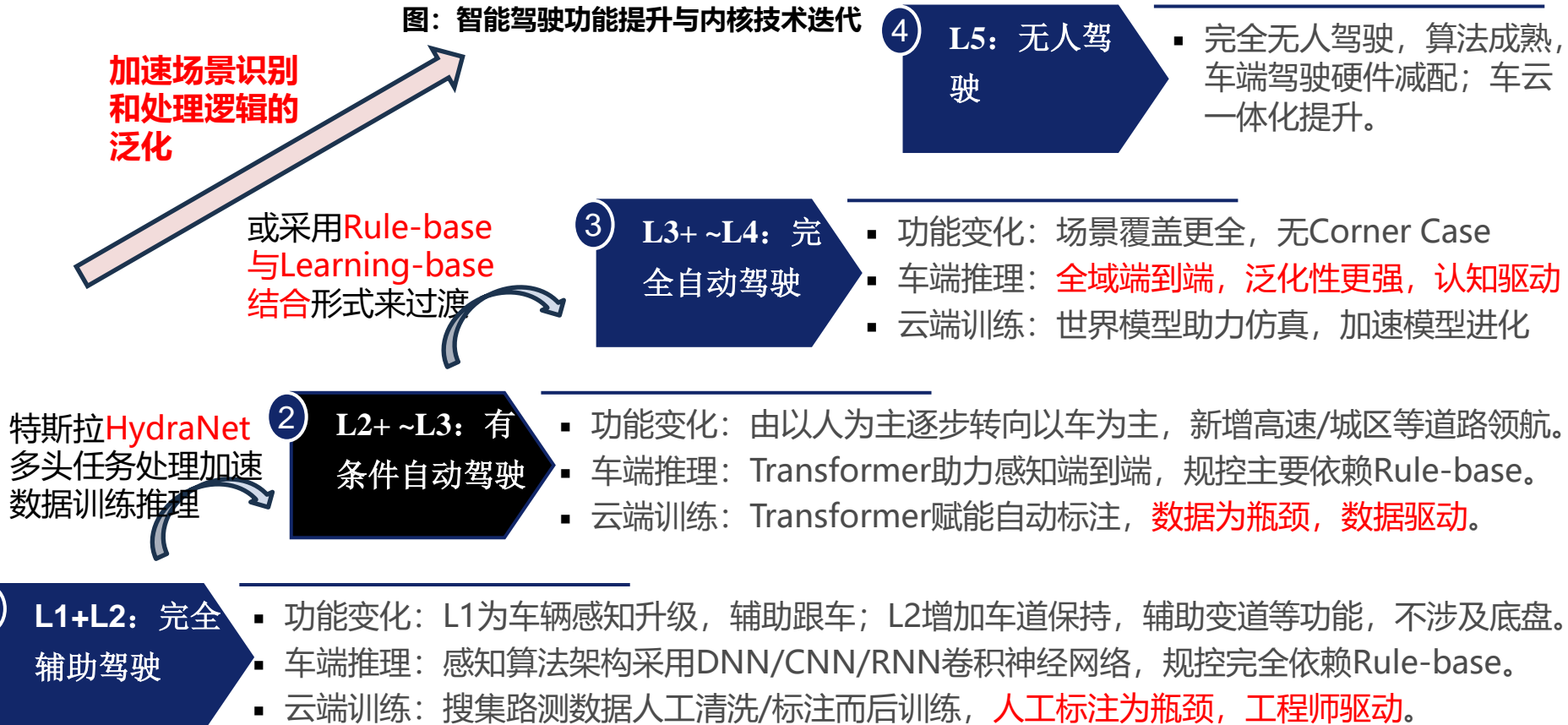


- **多模态数据喂养软件算法，驱动功能进化。** 硬件端升级提供多样化数据，智驾传感器由单目视觉向双目视觉以及视觉+雷达的方向演变，多传感器支持对周围环境的更精确识别，图像/点云等数据呈现方式多样化，数据赋能算法升级，支持智驾功能由L3级别以下向L3及以上迭代升级。
- **场景泛化是智驾能力提升的重要方向，数据需求量激增驱动算法由小模型切换为大模型。** 智驾软件上层应用算法中，感知能力提升核心系**场景识别能力的泛化**，规控能力提升核心系**场景处理逻辑的泛化**，均以**有效数据量为核心驱动**，数据处理/利用等是瓶颈。DNN/CNN/RNN小模型切换为Transformer大模型赋能**多模态数据处理**，重塑车端算法架构，以支持更高级别智驾功能。
- **世界模型（自学习，强泛化，高效率）或为通向L4级完全自动驾驶的必由之路。** 数据赋能催化模型自学习，理解物理世界基础运行规律，模型“认知能力”提升，或可实现完全自动驾驶。

类型	算法架构	特征
小模型	DNN	通过对像素级别进行逐个对比，通过图像特征进行识别，精确，存储简单，但计算量较大
	CNN	以“卷积-池化”的方式，将整幅图按兴趣区域 ROI 进行特征分割后，分别对 ROI 进行局部像素对比识别，相当于在DNN基础上进行降维，特征提取降低计算量需求
	RNN	在CNN基础上引入循环核，便于处理时间序列图像特征
大模型	Transformer	基于特有的 自注意力机制 ，能够有效捕捉序列信息中 长距离依赖关系并支持并行计算 ，适合处理大量且多类型的数据
	ResNet	深度 卷积神经网络，为解决随网络层数增加带来的网络退化（梯度消失和梯度爆炸）现象，残差网络使用跳跃连接实现信号跨层传播
	VQ-VAE	以自编码器将图像压缩为有限数量的向量集合，从而实现了高效的图像表示和重构
	生成对抗网络GAN	GAN由生成器和判别器组成，生成器经由输入的噪声向量输出为假设图像，判别器区分假设图像和真实图像并反馈训练，驱动生成更加逼真的图像
	世界模型（基于Transformer的复合）	人工智能系统构建的对于现实世界的模拟与表达，模型充分理解物理世界的客观运行规律

- 产业对于智驾功能持续迭代的诉求驱动车端/云端算法持续升级。1) 阶段一：L3以下。OEM搜集数据清洗标注后训练，算法基于特定规则，场景泛化能力较差，数据以及标注工程师均为瓶颈。2) 阶段二：L3级别。Transformer赋能，大模型端到端算法落地，自动化标注多模态数据，并快速提升数据利用效率，数据与算法能力形成正循环，**场景识别的泛化能力提升**。3) 阶段三：L3以上。车端算法完全端到端，Learning-base主导算法训练，强化算法感知以及规控端的泛化能力，**算法认知规划能力提升**，逐步支持完全自动驾驶。

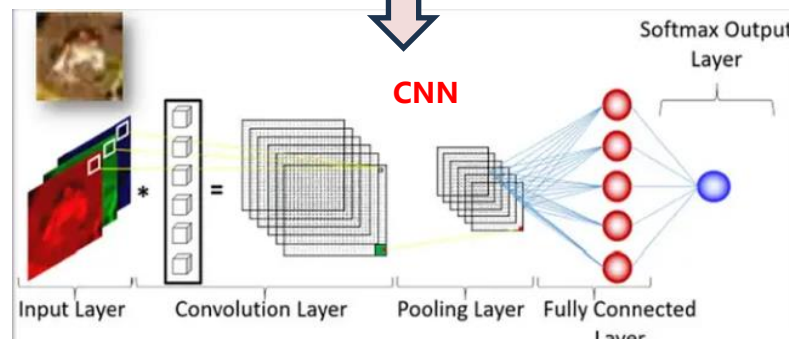
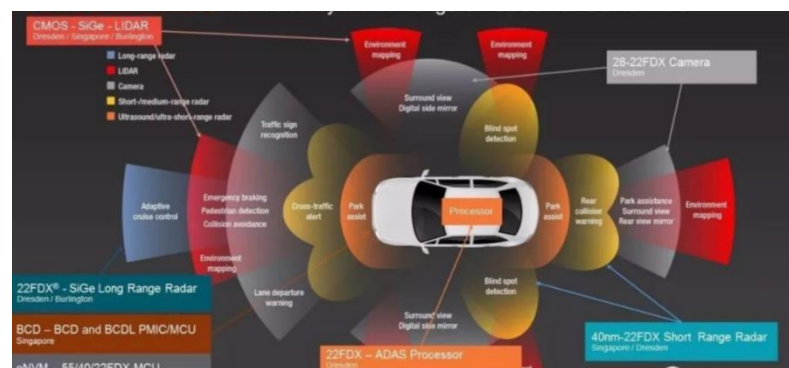
图：智能驾驶功能提升与内核技术迭代



■ **精确感知为跃升L3的瓶颈, BEV+Transformer架构上车支持产业化解决方案落地。** L1-L2: 智驾迭代主要集中在感知维度的升级, 不包含规控以及执行端底盘的变化。由L2-L3的过程, 则是由以人为主体逐步切换为人车共驾, 【外部环境的精确感知问题】为瓶颈。当前, 低成本的软件算法升级 (BEV+Transformer) 基本取代高成本解决方案 (激光雷达+高精地图), 推动产业化。

➤ **传统CV小模型架构下, 感知能力升级更依赖硬件。** 依赖摄像头+激光雷达/毫米波雷达等能够感知三维深度的硬件实时描绘周围环境, 并配合高精度地图提供的先验信息辅助判断; 算法领域, 以CNN卷积神经网络进行图像的特征提取和目标识别, RNN大规模记忆处理时间序列数据。

图: 智能驾驶感知



■ 暴露问题:

- ✓ 能力提升依赖庞大数据量, 但数据量过大带来: 1) 算法难以捕捉长距离依赖关系; 2) 数据处理依赖人工, 利用效率较低。
- ✓ 雷达硬件成本过高且高精度地图覆盖面较窄。

精确感知为跃升L3的瓶颈, BEV+Transformer架构上车支持产业化解决方案落地。

- **BEV+Transformer大模型架构赋能精准感知【处理长距离依赖关系+并行计算+统一融合】。**
- ✓ **Transformer以自注意力机制完成时间/空间关联, 处理长序列数据信息同时支持并行计算, 提高效率。** 1) 高效对多模态数据进行特征提取, 利用空间元素间的关联关系进行全局建模, 实现特征增强; 2) 利用元素与前后体之间的关联关系, 高效处理长序列信息, 避免重复递归/卷积。
- ✓ **BEVFusion统一框架, 实现多模态信息的精确端到端融合。** 用独立流程分别处理雷达和摄像头数据, 然后在鸟瞰视图(BEV)层面进行融合, 并扩展到多任务、多传感器框架, 兼顾信息完整度和效率, 标准化框架也便于数据清洗标注。

图: Self-Attention (自注意力机制) 原理

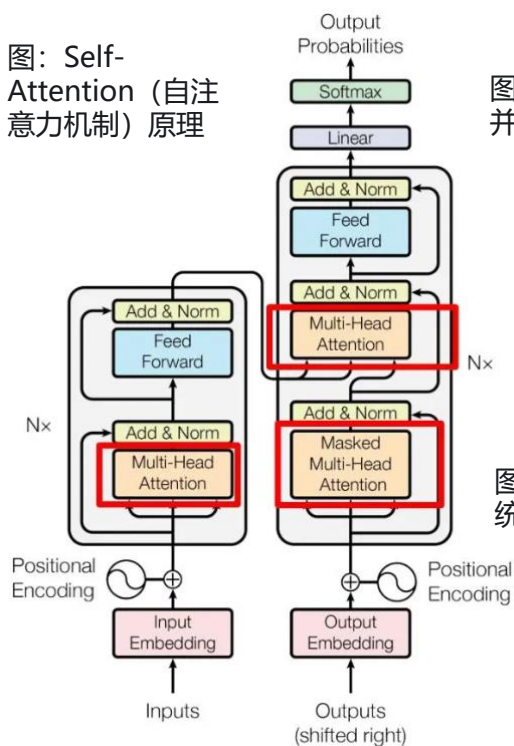


图: Transformer 并行特征提取

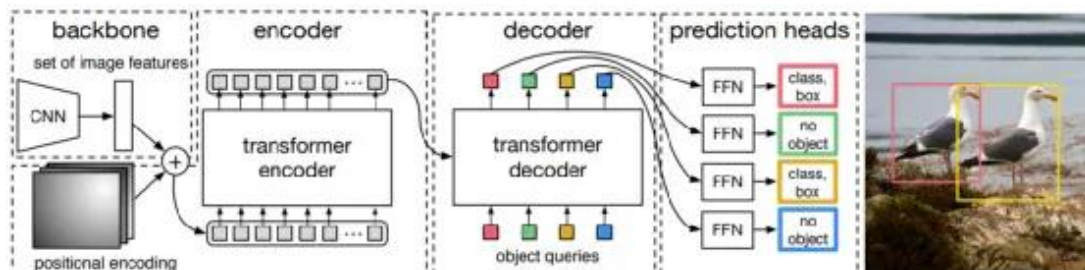
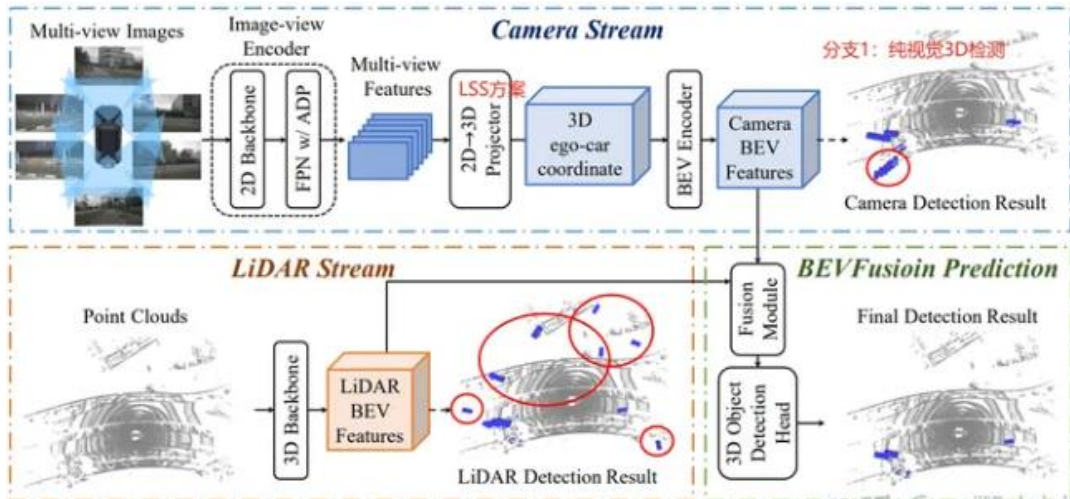


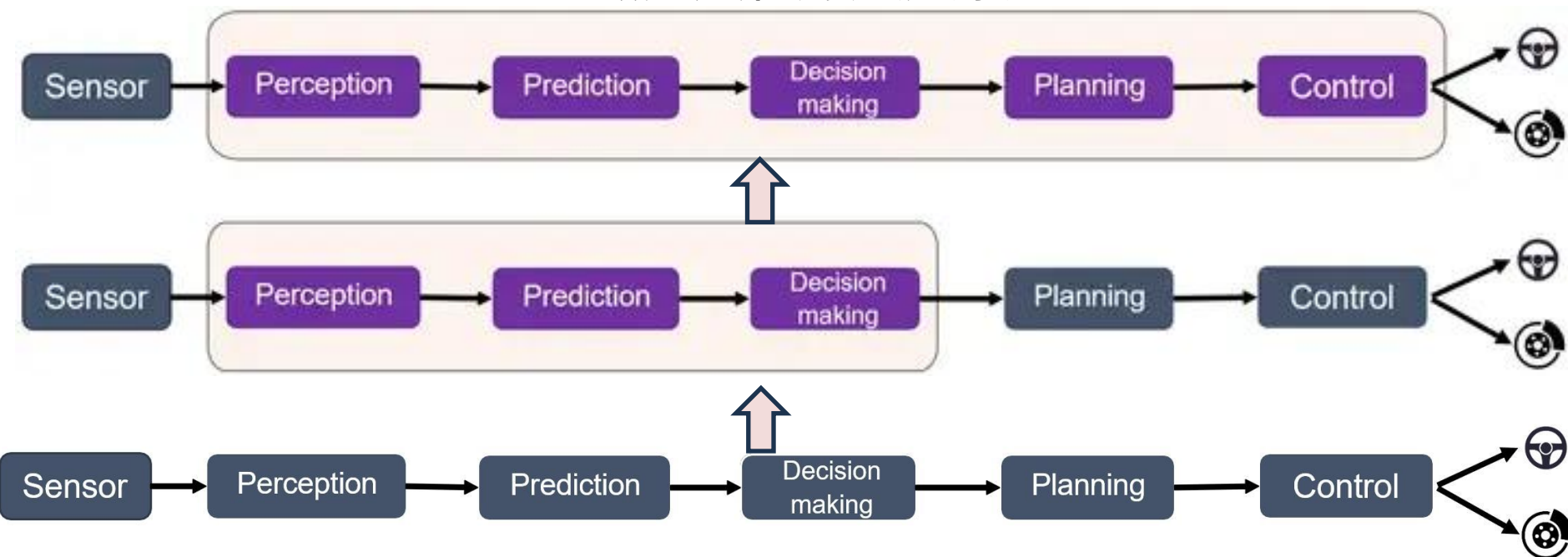
图: BEV视图进行统一特征级融合



■ **全面泛化**为跃升L4的瓶颈, World Model世界模型 (本质为模型基于对客观世界物理规律的理解进行精准预测) 或为终局解决方案。L4智驾要求完全自动驾驶, 相比L3智驾, 其对车辆运行过程中的场景识别以及预测规划提出更高要求, 强调【100%全面泛化】。因此, 车辆自主控制阶段的算法瓶颈转变为【如何正确预测周围环境变化并合理规划行驶路径】。

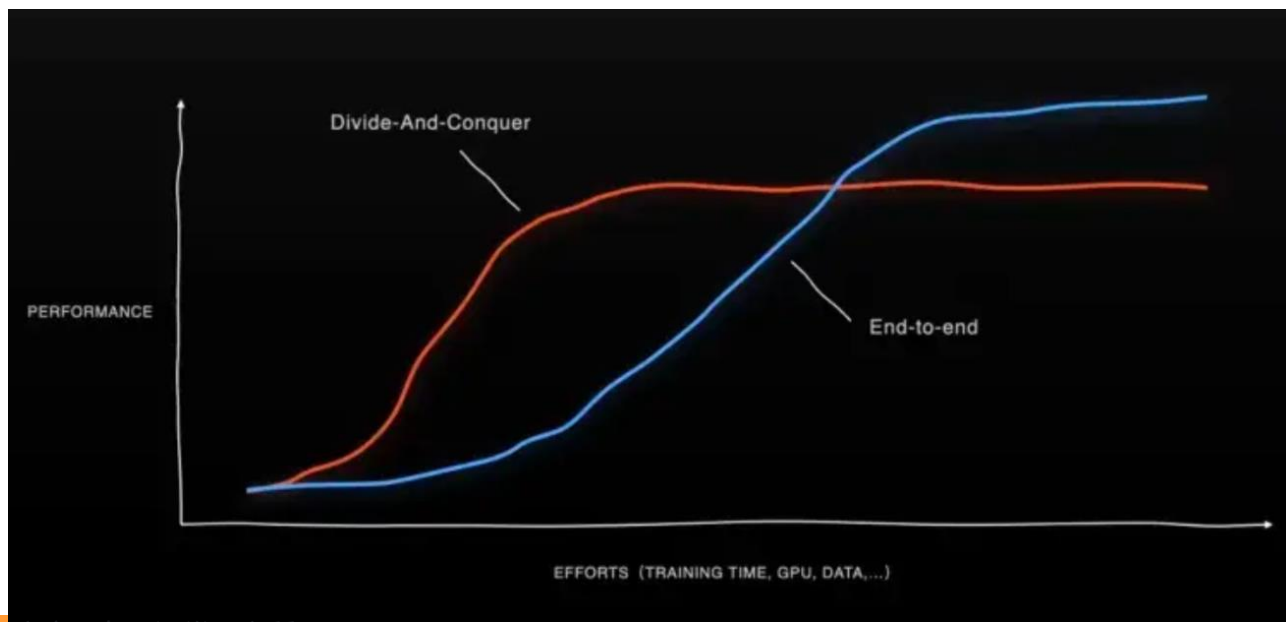
➤ **世界模型或为L4完全自动驾驶终局解决方案**, 通过End-to-End算法+大数据闭环训练, 培养算法认知泛化能力, 实现【通用具身智能】。World Model的核心增量: 反事实推理 (Counterfactual reasoning), 也即对于数据中没有见过的决策, 在world model中都能推理出决策的结果。

图: 模块化算法架构逐步迭代为端到端



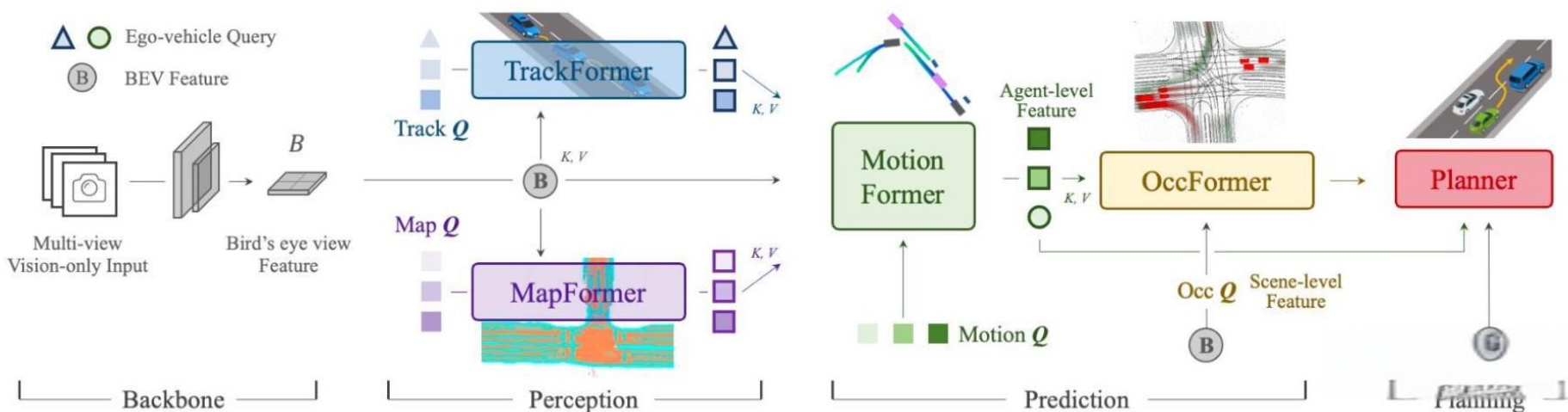
■ 端到端架构将感知-规控-执行模块串联统一训练。

- **优点：**1) **结构简单直接，算法工作量较低，减少维护成本；**传统策略基于完整规则进行控制，行为机械且代码量大，据小鹏2022年，城市NGP代码量是高速的88倍，端到端有明显优势，架构和代码方面一劳永逸，高度整合，省去独立模块开发以及数据人工标注和规则设计，降低成本。2) **更利于追求“全局最优解”，能力上限高，泛化能力更强，完美应对长尾场景；**基于规则的方案遇到未学习过的场景时无法做出良好的处理策略，且基于已有数据难以保证100%的场景数据覆盖，传统模式的最优解理论上无法达成，端到端是通过对场景的理解进行判断，天花板更高。3) **减少数据时延，便于车端迅速判断，提升安全系数。**
- **缺点：**数据的针对性更弱、模型训练的信号更弱，意味着提升性能**所需的数据量和算力规模更大。**端到端的优势在数据量达到一定程度后性能显著提高，但数据量较小时候性能上升缓慢，远低于解耦的传统基于专家模型的策略（特斯拉当前高速依然保留FSDv11的代码）。

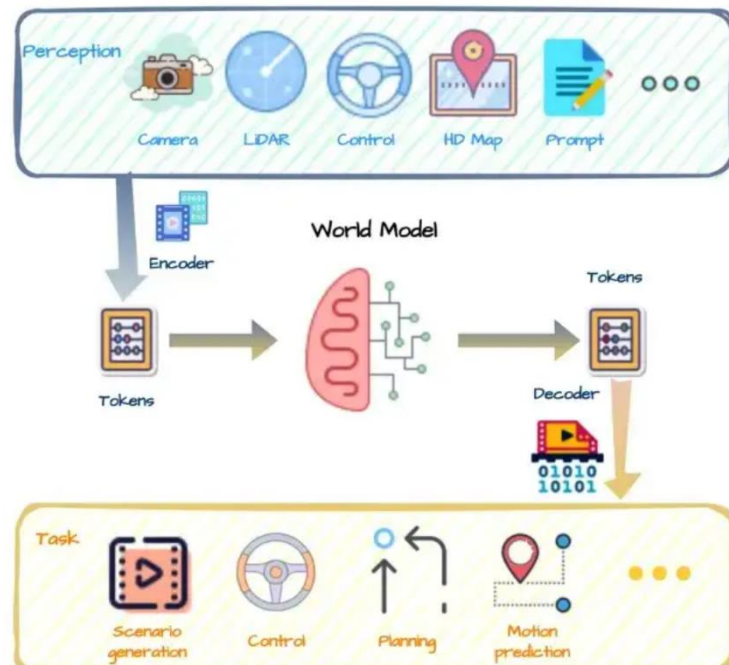


- UniAD (感知Perception+规划Planning)**：多个共享BEV 特征的Transformer网络将跟踪+建图+轨迹预测+占据栅格预测统一到一起，并且使用不依赖高精地图的Planner 作为一个最终的目标输出，同时使用Plan结果作为整体训练的loss 来源。相比于特斯拉FSDv12的黑盒端到端，其在感知/规控等模块间虽然一体化整合，但并非一个整体黑盒网络，仍可以对各个模块进行分别的监测和优化，各个模块间有了相当的可解释性，也有利于训练和Debug，**【解耦白盒】是核心**。
- UniSim (仿真)**：使用多个NeRF 网络，重建静态背景，并且将每个动态物体也构建出来，同时保存动态物体的轨迹。之后在仿真环境中，静态物体作为整体世界背景，所有的动态物体根据轨迹转换到世界坐标系中，进行世界环境的模拟。利用动静态解耦的方式，为之后仿真物体的删除和插入提供便利，同时对危险场景的模拟提供充足的自由度，保证场景泛化。

图：UniAD算法框架



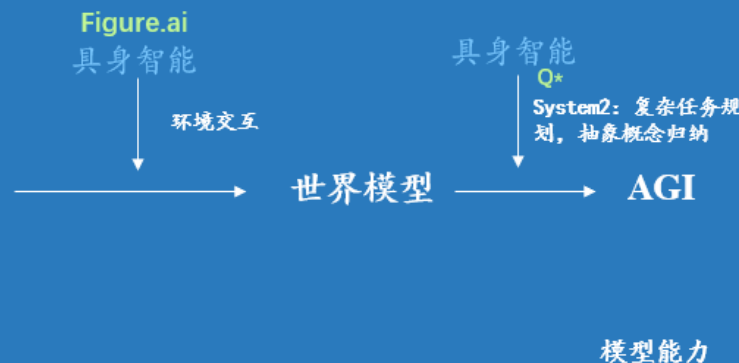
- **World Model ≠ 端到端**，但或为实现端到端的理想方式。
 - 1) 数据需求：端到端算法训练需要大量低成本+高互动的标准化数据，世界模型可制造大量仿真数据，无需标注直接应用于算法训练；
 - 2) 利用数据大规模进行强化学习，培养模型的规律认知能力，类似于人类学车/开车的行为，完成感知/记忆/动作的连贯行为。
- 特斯拉早于2023年提出世界模型World Model，**即通过神经网络结构化表示事物，并预测未来，将自动驾驶决策问题转为预测视频的下一帧**。世界模型用于理解世界，不仅能用于自动驾驶也能用于机器人，是特斯拉自身端到端智驾解决方案的基础原理，自动驾驶是世界模型的具身智能。



图：世界模型算法



通往AGI的路径方向



早期阶段：
各个模态独立发展

当前主流：
多模融合：模态融合&任务融合

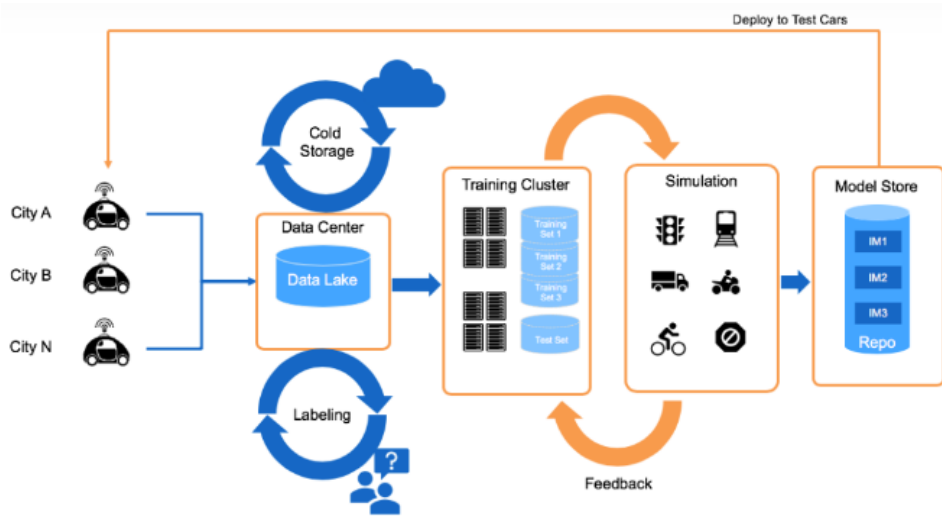
未来阶段：端到端

三、云端：大模型加速智驾算法迭代

数据闭环是产业玩家高效利用数据的核心。 智驾数据主要指车辆在行驶过程中由摄像头/雷达/地图等传感器感知到的周遭环境以及车辆自身状态响应等信息，其为智驾算法模型的学习反馈和迭代提供了必备的“良师”。

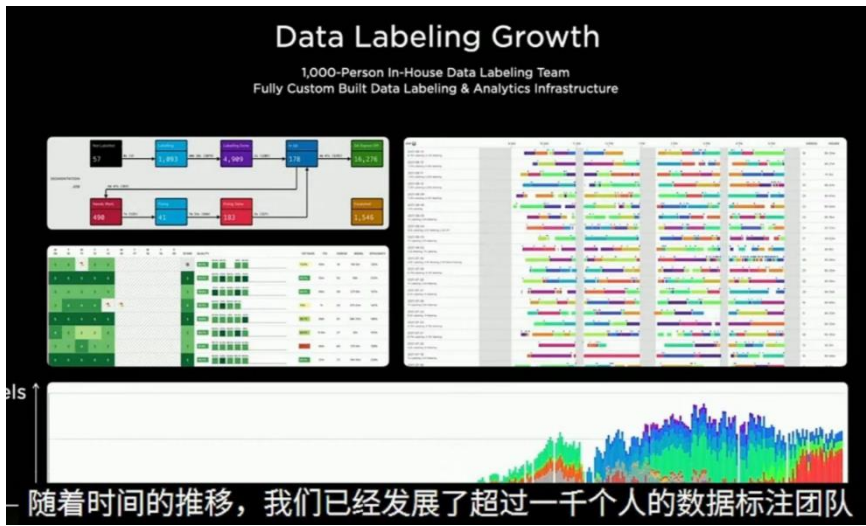
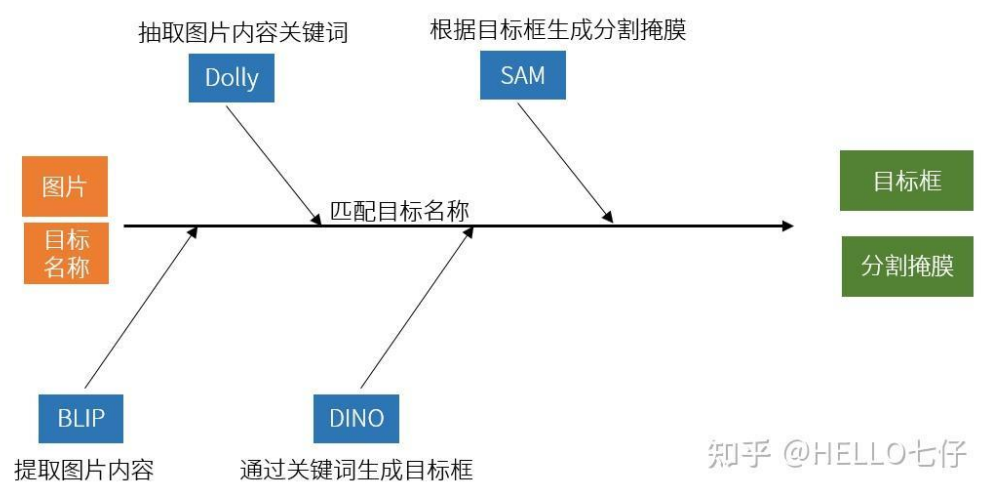
- **数据记录：**将车身传感器感知到的信息转化为数字信号，通常上传至云端，便于访问管理。
- **数据处理：**将基于不同传感器感知的数据进行预处理、滤波和融合，生成高质量统一格式的数据。
- **数据标注：**即数据信息归类，包括目标检测（特征提取）+轨迹生成/优化等，将智驾面临的场景一一归类，是算法预训练前最为繁琐的环节之一。
- **数据仿真：**模拟真实的交通场景和车辆行驶状态，方便对车辆的算法和系统进行测试和验证。
- **模型训练和算法部署：**基于已有数据对智驾算法进行云端训练，并裁剪部署至车端实时推理应用。

图：智能驾驶数据利用全流程



- **“场景驱动”** 是小模型时代智驾算法能力的重要特征，体现为数据的人工标注。传统CV小模型时代，智驾覆盖广度的提升依赖于长尾场景Corner Case的持续完善，工程师需要手动区分场景信息中的新增部分并人工分类标注，进一步修改“if-else”规则下的规控算法，“补丁”式小碎步提升。21年之前特斯拉数据标注团队规模在千人级别，工程师数量是限制算法迭代的瓶颈之一。
- **基于Transformer的SAM等自动标注算法打开智驾算法数据利用的天花板。** 其原理在于：首先基于少量的人工处理构建标注真值系统，进而将未标注的数据应用至真值系统得到标注结果，以SAM为例的标注算法即是基于Transformer机制完成批量化特征提取和标注。特斯拉2021年AI day宣布裁撤旗下千人左右的数据标注团队，正式大规模应用基于大模型的自动标注，并可于一周内实现人工标注几个月时间才能完成的一万个片段的标注任务。

图：自动标注



知乎 @HELLO七仔

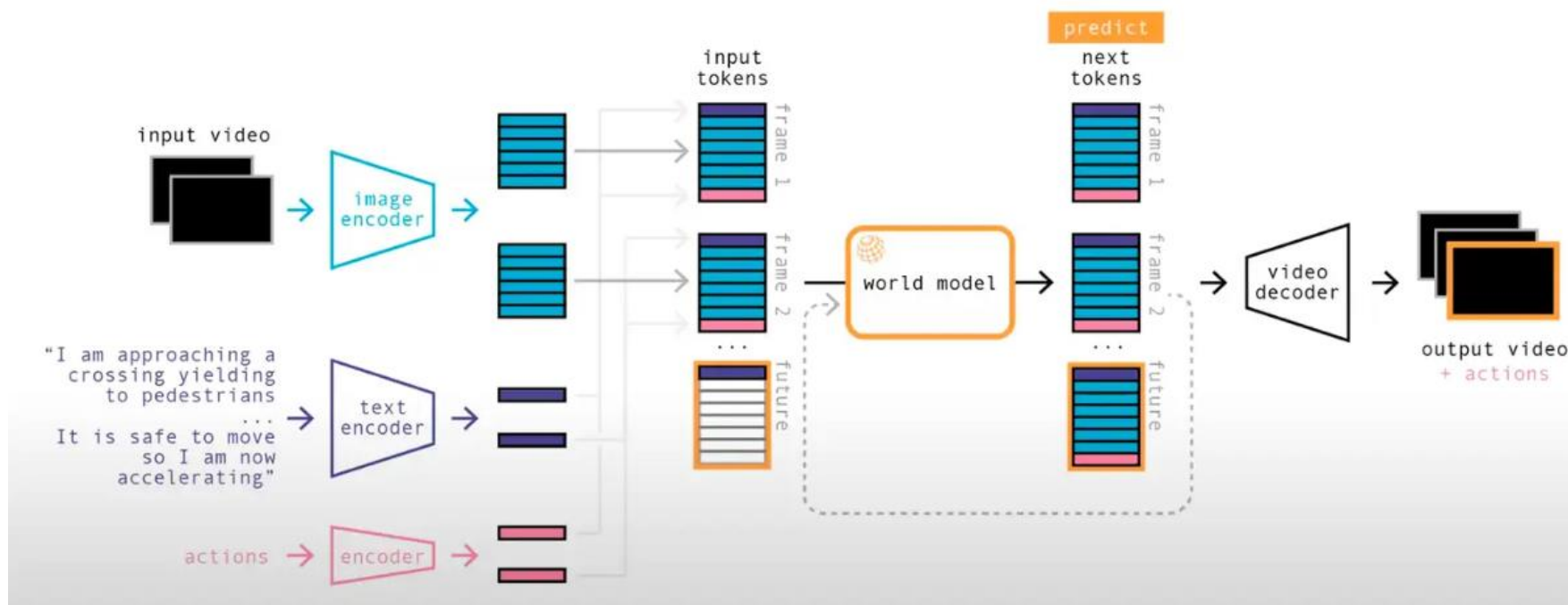
- **数据仿真是云端算法训练的重要组成部分。** 智驾仿真测试，即以建立车辆模型并将其应用场景进行数字化还原，建立尽可能接近真实世界的系统模型，如此通过软件仿真即可对自动驾驶系统和算法进行测试。其包含了虚拟的驾驶场景、车辆的动力学系统、感知系统、并预留对接ADAS/自动驾驶系统的通信接口。现实世界路测数据搜集【长尾场景信息不全+耗时且成本高昂】的短板极大促进了场景仿真技术的发展。
- **从基于模型设计到数据驱动迭代，以场景真实性作为方向。** 仿真系统包括静态场景搭建+动态场景搭建+车辆动力学仿真+传感器仿真几大部分。传统仿真多依据V-Shape流程，以动力学仿真为主，基于车控进行建模，随L2发展，“弱动力，重场景”的建模方式逐步被引入。但L3级别智驾所需的场景数目以及代码量飞速增长，**以数据驱动的“场景重建”仿真技术成为主流**：基于游戏引擎开发，以实现对静态场景的高保真渲染和物理模拟。例如英伟达的DRIVE Sim即是一种基于Omniverse引擎构建的仿真工具，应用 Omniverse Replicator 以缩小仿真到真实的域差距。



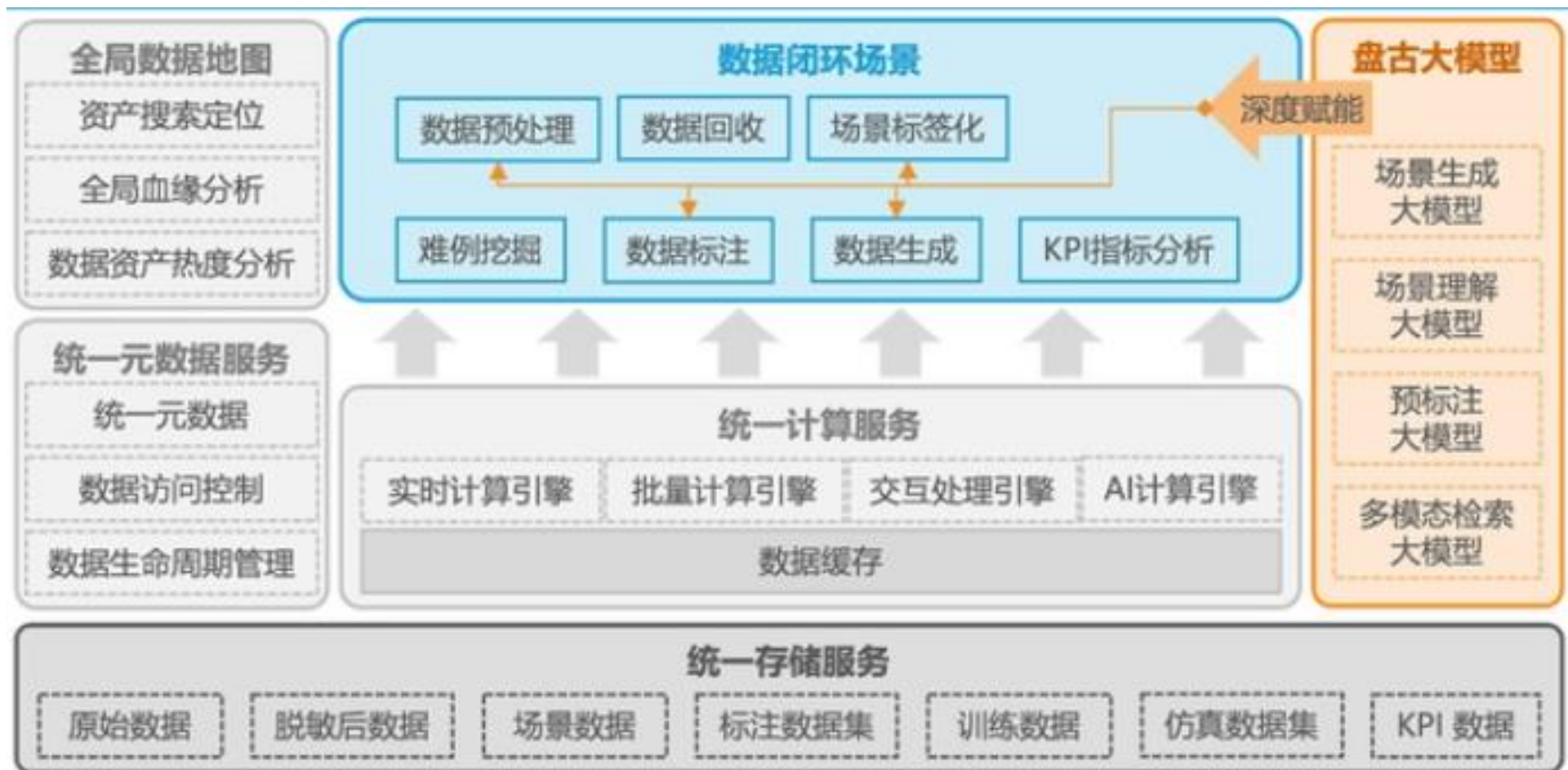
知乎 @ 佐思汽车研究

- 世界模型有望提供真实且多样化的仿真数据，即加速泛化，加速云端算法迭代。
- 1) 当前智能驾驶仿真的主流技术方案为【NeRF+素材库+游戏引擎，基于真实数据进行道路重建，并保证和真实场景的相似程度，再依靠交通元素的不同排列组合进行有条件泛化，最后通过游戏引擎输出高质量的仿真画面】。
- 2) 世界模型与之相比，具备能够凭借对于物理世界运行规律的理解，自身泛化到样本以外的能力，覆盖更多Corner Case，有效降低厂商对于真实数据的依赖程度，提升效率；但大模型的基础前提决定其比较依赖大算力+大规模数据喂养持续训练，方能避免出现违背客观规律的情况。Tesla可以做到多个摄像头同步联合预测；并支持prompt嵌入，可以提出要求生成车辆直行或者对向车道视角的未来数据。Sora面世即是世界模型所代表的场景生成的新范式代表。

图：Wayve的GAIA-1 high-level architecture



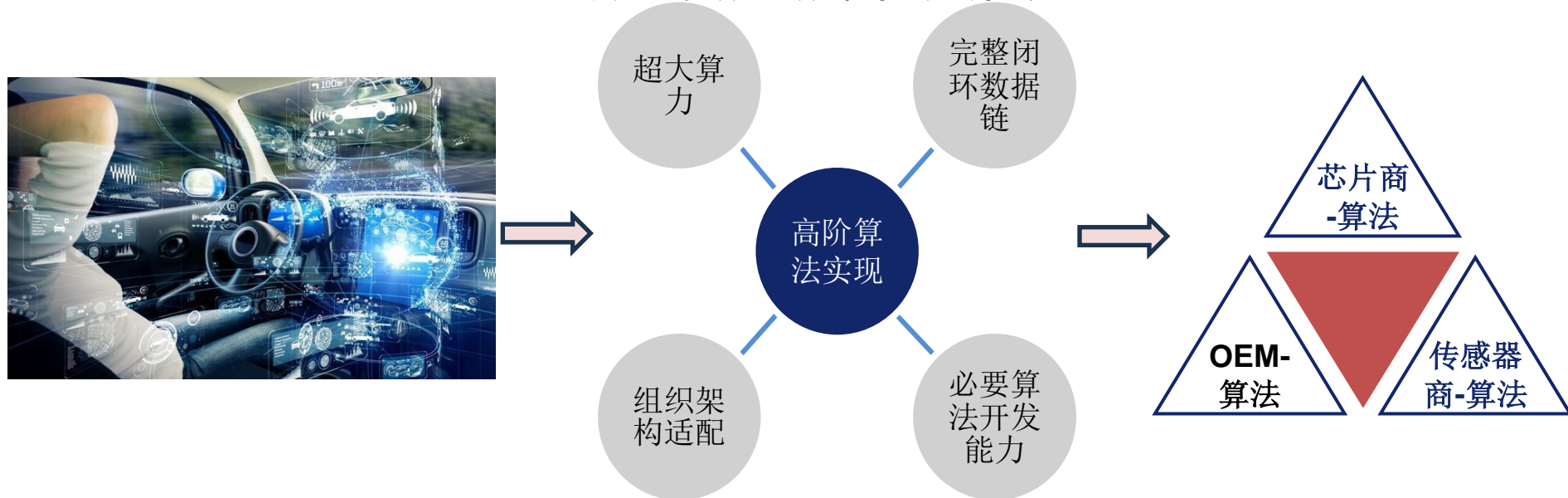
- **云服务为基，搭建数据闭环坚实底座。** 华为云ModelArts平台提供DataTurbo、TrainTurbo、InferTurbo 3层加速，分别提供数据加载、模型训练、模型推理三方面服务，能够助力数据读取时间缩短50%、训练效率提升40%+。统一的数据湖让数据存储成本下降20%、冷数据下沉率最高96%。华为云服务能够完整覆盖数据闭环各个环节，提供数据管理平台、**AI模型训练平台、仿真评测平台等**，形成全栈式研发工具链。串联互通的工具，增强了业务连续性，缩短了模型训练时间，大大减少了调试成本和人力成本，提升研发效率。



四、当前产业玩家的智驾算法能力如何？

- **高阶智驾算法持续升级，产业趋势明确；数据/算力为壁垒。** 北美特斯拉以及国内市场华为/小鹏等玩家引领，Transformer赋能多模态感知融合（FSD v10）及更进一步的端到端（FSD v11/12）驱动**产品功能体验持续升级，低时延+强泛化**，同时也明确产业升级方向。端到端方案依赖大算力+极多高质量数据训练，同时要求玩家数据链闭环，保证数据采集/存储/清洗/标注更加高效。
- **算法实现高壁垒+数据闭环硬要求，未来智驾算法产业格局趋于集中化。** 当前整车下游L3高阶智驾算法方案以OEM自研为主，华为以“算法赋能，数据回传”的形式深度绑定OEM；L3以下智驾算法呈现OEM+硬件商+独立算法商三足鼎立格局，OEM玩家/硬件商占据行业主导位置。
- 我们认为，考虑高阶智驾功能的实现对于算法能力、组织架构、超算中心、完整数据链等的要求，未来“掌握硬件的基础上去发展软件”或为主流，即**掌握壁垒最高的硬件——芯片；提供性价比最高的硬件——传感器；掌握粘性最强的硬件——整车。**

图：自动驾驶驱动多模态大模型，带来算法架构变化



4.1 OEM玩家及相关企业

- 2020年为特斯拉“硬件为先，软件随后”策略的重要窗口，【大模型】是特斯拉引领本次变革的核心抓手。【HydraNet/FSD芯片+Transformer/Dojo】等6次硬件变化升级+3次软件架构的重大创新变革，均体现特斯拉第一性原理的经营思想。

图：特斯拉边缘端/云端软硬件历史迭代

		2014.10	2016.10	2017.07	2019.04	2021.07	2022.04	2023Q2	2024Q1	
边缘端	FSD历史重要阶段 (软件)	Autopilot1.0	Autopilot2.0		Autopilot3.0	FSD Beta v9	FSD Beta v10.11	FSD Beta v11.3	FSD v12	
	硬件解决方案	名称	HW1.0	HW2.0	HW2.5	HW3.0			HW4.0	
		芯片	Mobileye Q3	英伟达 drive PX2	英伟达 drive PX2+	自研FSD1.0			自研FSD2.0, 算力五倍	
		传感器	1摄像头+1毫米波雷达+12超声波雷达	8摄像头+1毫米波雷达 (2.5升级)+12超声波雷达		8摄像头+1毫米波雷达+12超声波雷达			12摄像头+1毫米波雷达	
	软件解决方案	标注	人工标注				BEV+Transformer架构, 数据驱动深度学习, 使得自动标注落地			
		感知	外采Mobileye黑盒	自研图像识别+多传感器后融合策略		自研芯片硬件, Hydra Net驱动多头任务实现, 进一步提升	BEV+Transformer上车, 特征级融合落地, 实现感知大模型, 无图城市领航功能上车	占用网络上车, 泛化能力增强, 功能维度统一高速与城市领航辅助		感知+规控全域端到端落地, 持续算法训练, 追求具身智能
		规控	自研rule-base				rule为主, 开发learning-base			
	备注	2015.04开启自研软件	2017.03自研算法, 硬件升级但功能降级	2018年开发更好实现多传感器融合的策略	硬件储备完成	感知能力升级驱动无图化		泛化能力增强驱动传感器简化	硬件再次迭代, 支持软件进一步升级	北美开放, 即将入华
	云端						2021年8月官宣Dojo, 利用海量的数据, 做无监督的标注和仿真训练, 更好驱动大模型; 世界模型加速大模型仿真			

- **HydraNet赋能多传感器融合，特斯拉基于全自研软硬件实现有限范围智能领航辅助驾驶，追平行业。** 2016年起特斯拉自研软件算法上车，2D网络的CNN架构+人工数据标注的迭代模式下，特斯拉智驾功能表现逊于长期深耕的Mobileye；2019年4月配合自研FSD芯片，**HydraNet九头蛇架构**实现更好的特征提取以及多头任务并行，催化特斯拉追平Mobileye L2+级别智驾功能。
- **BEV+Transformer开启大模型时代，软件能力跨时代跃迁，引领行业变革。** 1) Transformer赋能2D图像升维实现BEV转换，占用网络（加入时序的BEV）进一步泛化异形场景，**实现无高精度地图化及传感器精简化**；2) 特征级融合取代后融合，**减少信息损耗，还原更真实的场景信息**；3) 人工标准转向自动标注，庞大数据预训练大模型，进行云端预标注，车端蒸馏落地。
- **FSD升级，全栈端到端落地。** 车端感知规控全融合，大模型覆盖范围深化，数据赋能加速迭代。

图：特斯拉自动驾驶算法迭代历程



智能驾驶软件算法架构历经两次框架变化：

- 1) 2D+CNN小模型向Transformer大模型进化，智驾迭代由**工程师驱动转为数据驱动**：2020年特斯拉BEV上车，**数据后融合变为特征级融合**，提升数据利用效率和结果准确性，同时**数据标注由人工转为模型自动**，迭代提效；2022年占用网络落地，**泛化能力进一步增强**，实现感知维度端到端；2023年数据驱动算法持续迭代，规控环节端到端落地，**rule-base比例降低，能力优化**。
- 2) 模块化端到端进化为全栈端到端，**World Model逐步完善，数据驱动变为认知驱动，向L4迭代**。特斯拉FSD v12起转向智驾环节全栈端到端，数据闭环后World Model逐步演化完善，基于历史数据去理解未来的范式进步为基于常识和规律自我学习进化的范式，实现无接管的L4智驾。

图：智驾软件算法路径变化以及各玩家阶段对应

底层感知逻辑	算法架构	对应智驾功能	对应FSD版本	行业玩家掌握程度
传统CV	2D+CNN	L2	Autopilot阶段	2020年以前行业的主流方案
ViT	BEV	L3	FSD v9	当前国内主流OEM+Tier玩家主要竞争环节
	占用网络		FSD v10~v11	
DiT	World Model	L4	FSD v12	2024年特斯拉FSD v12首次采用

备注：1) FSDv11相比v10核心迭代在于learning-base神经网络基本主导规控环节，因此体验更流畅；2) 国内车企对比以各家公开的产品功能落地和研究宣发为准，不代表车企自身的内部非公开研发规划；3) 参考特斯拉，各玩家智驾等级迭代较快，需高频跟踪，当前水平划分不代表长期。

特斯拉FSD版本	算法架构	对应智驾功能	对应落地时间	国内车企当前水平对应阶段
FSD Beta v9~v10	BEV+Transformer	城市领航落地	2021.07之后	小米
FSD Beta v10.11	占用网络	城市领航快速铺开	2022.04	理想/蔚来/百度
FSD Beta v11.3	感知端到端，规控环节rule-base比例降低	城市+高速+泊车等全场景贯通	2023Q2	华为/小鹏
FSD v12	世界模型下全栈端到端	全场景贯通且逐步迭代为零接管	2024Q1	24H2起华为/小鹏陆续迭代方向（规划阶段，尚未落地）

华为：ADS3.0落地端到端，数据引擎加速迭代

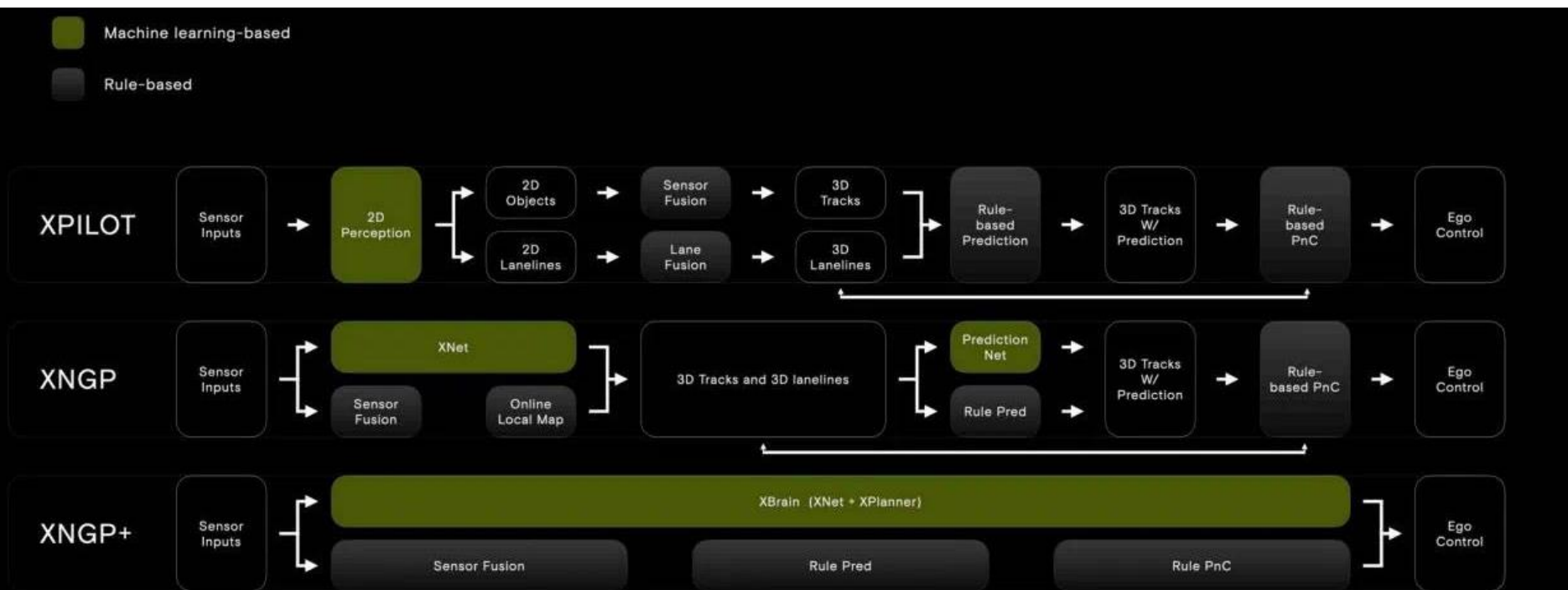
- **华为ADS智驾系统历经三次迭代**：2021~2023年，BEV+GOD（类似于占用网络）支持感知端大模型化，实现感知维度端到端，在感知硬件逐步简化的同时支持无图模式城市NCA加速开放；2024年华为重磅迭代模块化端到端，覆盖感知及规控环节，全场景贯通。

图：华为ADS持续进化历史迭代

版本		ADS 1.0	ADS 2.0	ADS 3.0
发布时间		2021.4	2023.4	2024.4
软件	架构	模块化	模块化，感知端到端	感知GOD大网，规控决策PDP端到端
		BEV网络	BEV网络+GOD网络	GOD大网，输入PDP端到端落地
	感知方式	白名单目标+道路结构，需要高精地图	识别异形障碍物，无图化，泛化能力提升	全面的物理世界理解，感知场景语义
	增强功能	L2级别LCC	城区道路NAC、LAEB、GAEB、ELKA、城区LCC PLUS、哨兵模式	全场景贯通NCA，CAS 3.0、ESA、车位到车位NCA、窄空间泊车
硬件	视觉传感器	13颗	11颗，前挡风减少2颗	摄像头+雷达全融合
	毫米波雷达	6颗，3D毫米波雷达	2颗，3D毫米波雷达	升级为4D毫米波雷达，性能提升35%
	激光雷达	3颗，华为等效96线半固态前保1颗+前保侧面2颗	1颗，速腾聚创（车顶）	192线，增强全天候、小目标检测能力
	云端算力		2.8EFLOPS（截至2023年11月）	3.5EFLOPS
智驾功能定位		L2	L2+	L3-L5
解决方案		全系标配	全系标配	中高端：含1/3颗激光雷达+4D毫米波雷达 入门级：视觉ADS，支持高速NCA
应用车型		北汽极狐α、阿维塔	问界/智界/享界全系	-

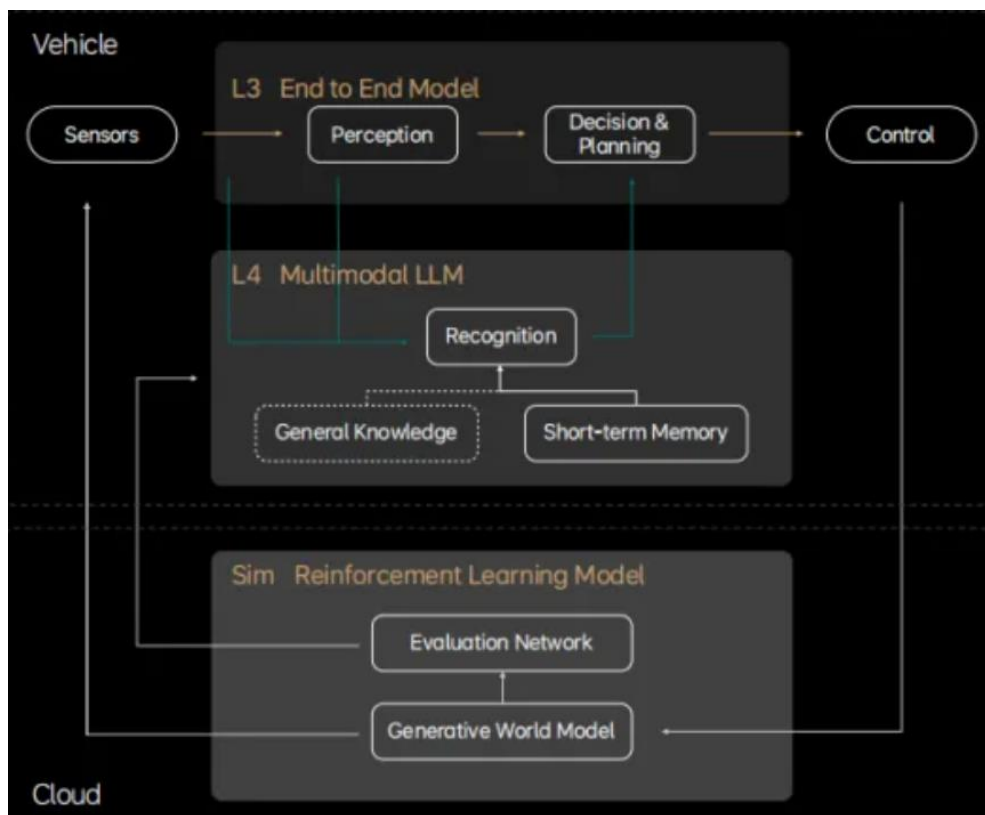
- **软件维度：**小鹏Xpilot/XNGP/XNGP+迭代围绕【增加learning-base使用率，端到端全覆盖】的目标架构，XNGP落地BEV+Transformer架构实现感知维度端到端，规控环节逐步引入learning-base；2024年XNGP+有望在rule-base基础上实现感知-规控模块化端到端XBrain。
- **Xbrain：**XNet 2.0融合了行业最高精度的纯视觉占据网络，可实现动/静态BEV、占据网络三网合一；基于神经网络的XPlanner可结合分钟级以上的时序连续动机，并依据周边环境信息及时变通，生成最佳运动轨迹。5月OTA上车的XNGP+将实现上述感知大模型升级和规控大模型上车。

图：小鹏汽车Xbrain终极架构



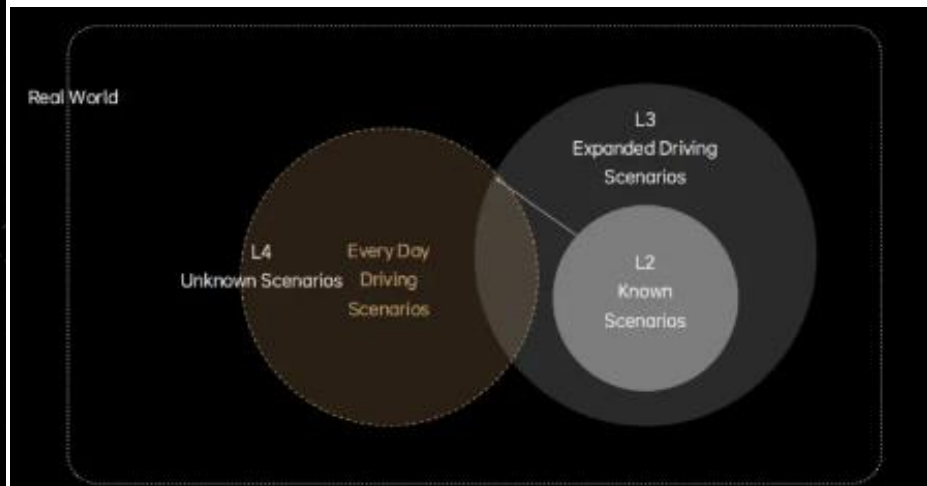
- **软件维度：**理想AD Max实现感知大模型落地，AD Pro预计24年中迭代与Max技术路线趋同；2024年理想预计完成规控环节端到端落地，实现全技术栈可训练的“端到端模型化”。
- 展望未来，理想基于【认知模型】（即特斯拉世界模型）做预研，目标是开发L4场景，基于1.4EFLOPS云端算力，更充分的利用多模态AIGC，短期进行场景重建和衍生，加速仿真数据生成和算法训练；中长期做到知识驱动形式的“场景理解”，真正实现L4覆盖100%的Corner Case。

图：理想未来智驾算法迭代的思维框架



图：理想关于不同级别智驾的看法

- 规则驱动——L2：2D/Mona 3D
- 数据驱动——L3：BEV/端到端
- 认知驱动——L4：VLM/世界模型



■ 蔚来自NT2.0平台起加速自研算法迭代，24年4月底迭代全域领航NOP+，对标FSD v10+。

- **硬件维度**：蔚来自NT2.0平台起全系切换英伟达，标配4*OrinX芯片（2主控/1冗余/1训练）以及1激光雷达/11摄像头+5毫米波雷达+12超声波雷达，标配硬件支持**3.8EFLOPS端云一体化算力**。
- **功能维度**：蔚来于23年初上车自研高速NOP领航智驾，23年中/底分别将BEV/占用网络架构迭代上车，23年底/24Q1城市领航分别开城6/20万km，**24年4月底实现全域领航辅助NOP+推送**。
- **软件维度**：**落地感知维度端到端，全域推送NOP**。蔚来打造NAD Arch智能驾驶架构，该架构包含Lane 2.0感知网络、NADCloudM云、NADHVVH等算法模型应用。其中，Lane 2.0可支持城市场景路口通行的实时感知，NADCloud M可借助云端大模型，提升感知能力；而NADHVVN可通过数据驱动的规划网络，使全域领航辅助NOP+拥有更细腻的交互能力。

表：蔚来NT2.0平台车型智驾功能迭代历史

日期	版本	新增智驾内容
2022/8/19	Banyan 1.1.0	视觉融合泊车SAPA、前向碰撞预警FCW、自动紧急制动AEB
2022/12/20	Banyan 1.2.0	ET7车辆近距召唤
2023/3/17	Banyan 1.3.0	23年元旦试运行，3月底向全量用户推送高速领航 ：NOP+Beta增强领航辅助功能优化、全车型近距召唤、视觉融合泊车增强、紧急车道保持
2023/6/30	Banyan 2.0.0	动态环境模拟现实2.0（ESD）、辅助遥控泊车（RPA）、 全新BEV升级NOP+连续性
2023/11/15	Banyan 2.2.0	高速领航辅助驾驶”零接管“成为可能、泊车高频一把泊入
2023/11/22	Banyan 2.3.0	更名为“全域领航辅助 NOP+”，开城6万km
2024/1/27	Banyan 2.4.0	4D路况舒适领航、GOA通用障碍物预警 、辅助Beta通用障碍物识别、全向AEB
2024/4/30	Banyan 2.6.0	全域领航辅助 NOP + 全量推送 ：增强车道居中辅助、全场景误加速抑制辅助、GOA通用障碍物预警及辅助

极越以吉利+百度合作赋能，利用百度自研LD车道级地图，坚持纯视觉智驾方案，迅速追赶。

- **硬件维度**：极越外采英伟达双Orin芯片，采用11摄像头/5毫米波雷达/12超声波雷达感知硬件，通过无激光雷达的视觉方案实现L3高阶智驾功能；同时云端百度2.2EFLOPS算力加速算法训练。
- **功能维度**：极越于2024年初OTA V1.3实现北上深杭广五城高精地图加持下的城市PPA，预计24H1开放300城，24年全年实现全国覆盖（有百度地图的地方均可使用）。
- **软件维度**：**1) 大模型通用纯视觉方案上车，类比特斯拉FSDv10+阶段**：百度VTA视觉大模型实现感知端到端，OCC感知模型实现占用网络+Transformer加持后对通用异形障碍物的检测。**2) LD智驾车道地图形成差异化**：由视觉大模型端到端生成，保留必要精度基础上增加经验地图+安全图层+实时图层，目前已覆盖全国360城，支持PPA全国都能开。

图：百度OCC占据网络-通用障碍物检测能力覆盖行泊全场景

表：百度极越智驾关键迭代历程

视觉OCC 占据网络「检测」能力升级 – 提供行泊全场景模型

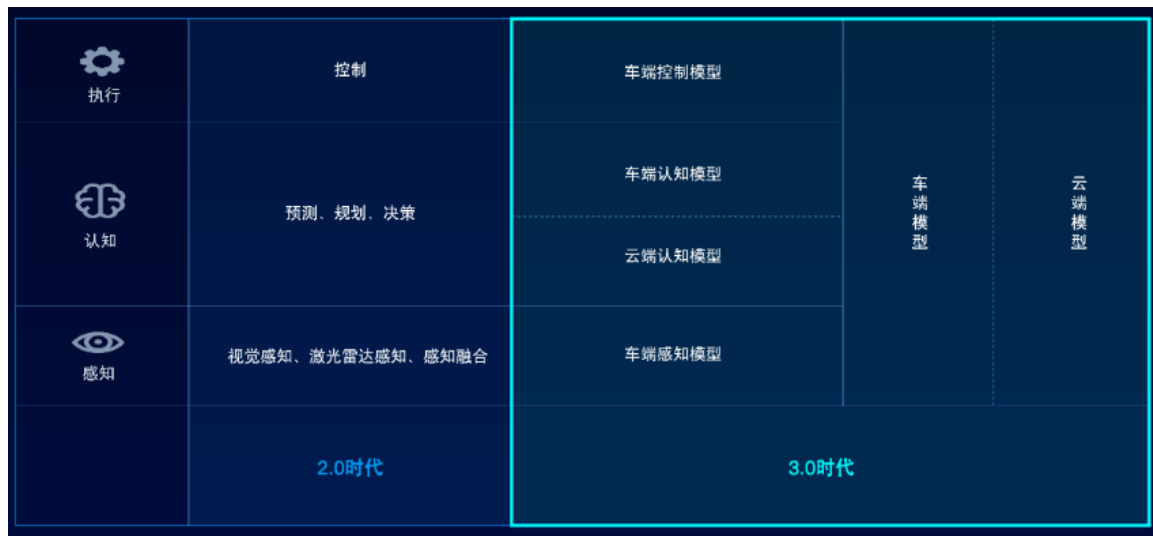
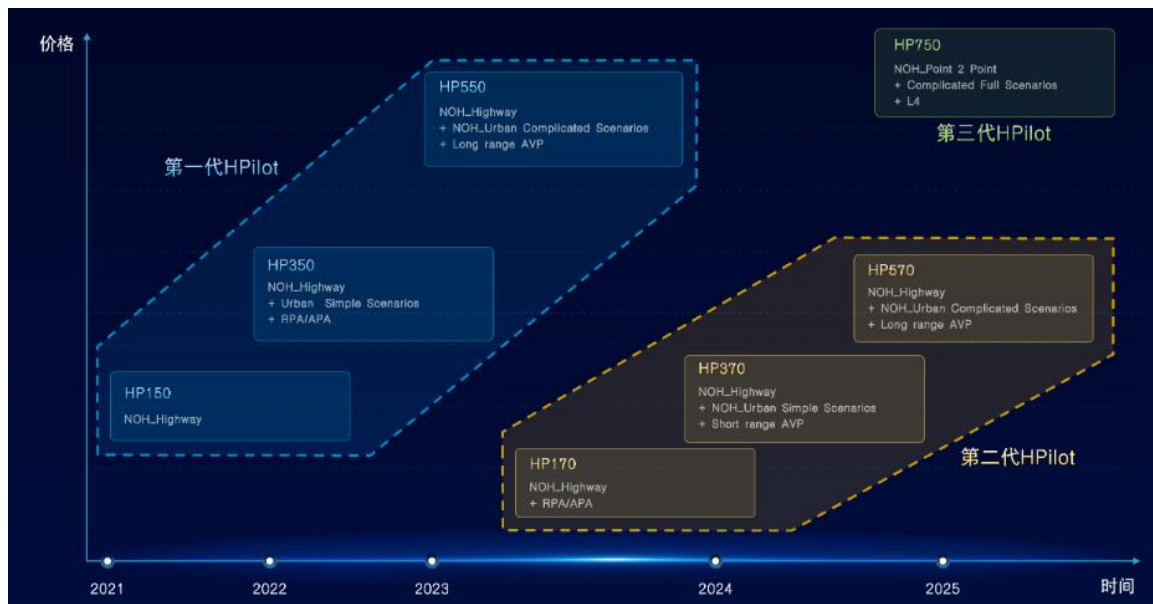


迭代历程	时间	功能
OTA V1.3	2024.01.14	OCC模型上车，PPA开城北上深杭，泊车效率优化，行车逻辑优化
OTA V1.4.0	2024.03.25	1、视觉大模型发布上车，升级静态检测、时序跟踪、实时建图、场景理解等能力，新增开城广州； 2、推出百度LD车道智驾地图
OTA V 2.0	即将上线	全国都能开的PPA

➤ **长城毫末持续深耕智驾大模型上车，自研生成式大模型 DriveGPT 雪湖·海若充分展现公司在数据的筛选挖掘、自动标注、生成仿真以及认知可解释性等领域的技术积累，算法领域聚焦感知/认知（即规控）大模型车云一体端到端，感知端引入多模态，认知端借助外部大语言模型 LLM 的海量知识来辅助给出驾驶决策。**

➤ **产品维度，毫末已推出三代七款智能驾驶产品，可满足高、中、低价位不同车型的量产需求。其中，HP170、HP370、HP570 三款千元级辅助驾驶产品，已陆续进入交付状态。**

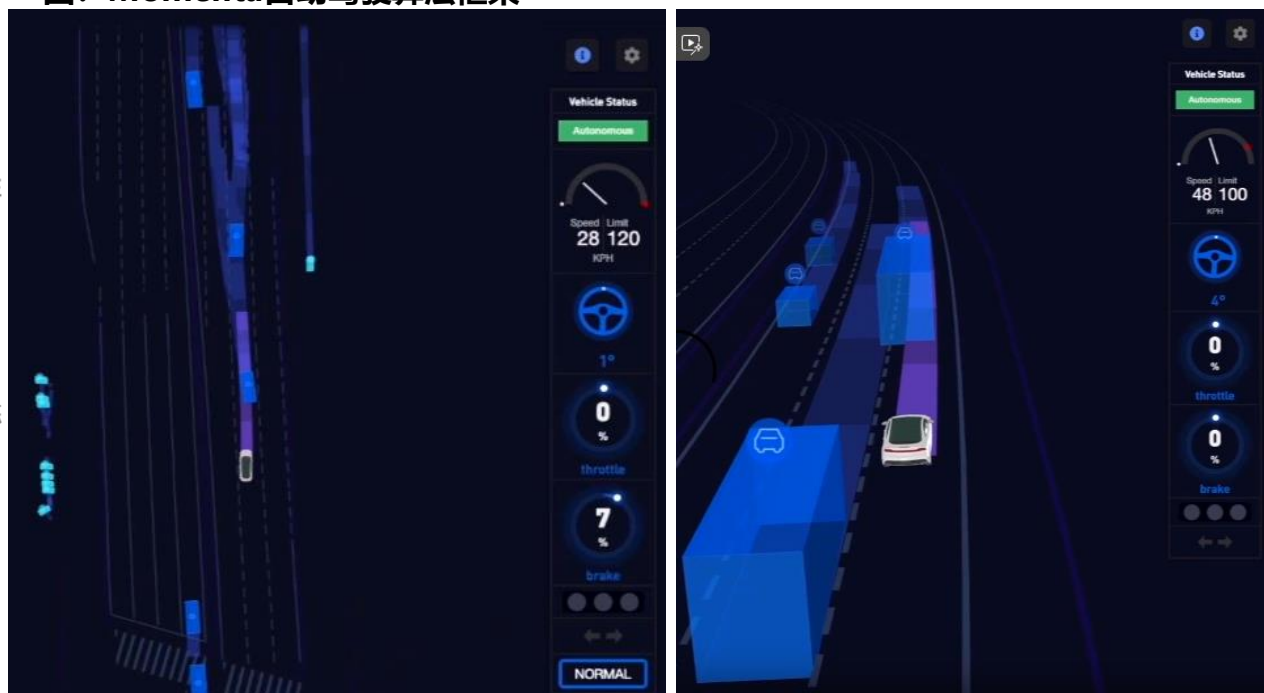
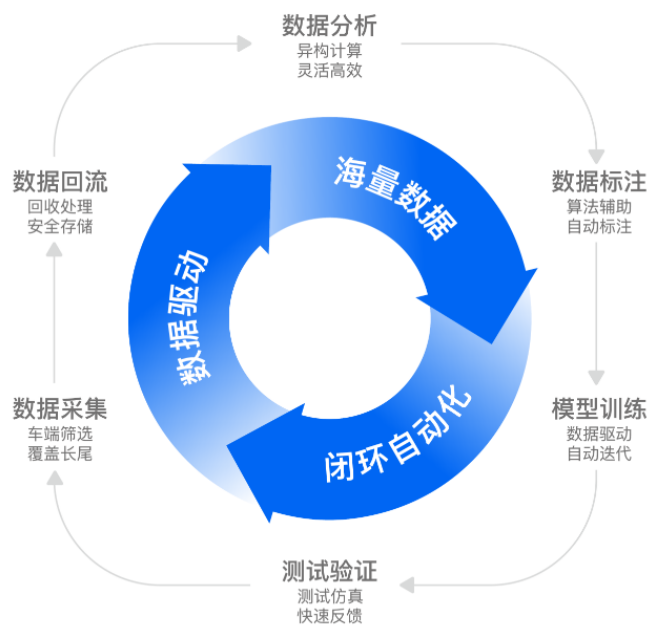
图：长城汽车毫末技术栈以及产品栈



4.2 第三方智驾算法公司

- **Momenta**，成立于2016年，以“Better AL, Better Life”为使命，历史融资12+亿美元，先后完成大数据&大模型平台建设、感知规控全流程算法构建、不同级别智驾软件解决方案三阶段。
- **核心优势**：1) **全流程数据驱动算法**，包括感知/融合/预测和规控等算法模块，都可通过数据驱动的方式低成本高效的迭代更新；2) **闭环自动化**，自动筛选海量黄金数据，驱动算法的自动迭代。
- **产品解决方案**：1) **MSD (Momenta Self-Driving)**，L4，是Momenta 完全无人驾驶解决方案，广泛应用于出租车和私家车等场景；2) **Mpilot**是针对私家车前装可量产的高度自动驾驶全栈式解决方案，主要的核心产品包括Mpilot X 等端到端的全场景，连续的高度自动驾驶解决方案，并可通过Open Solution适配英伟达、高通等主流芯片硬件和传感器平台，逐步落地端到端。

图：Momenta自动驾驶算法框架



- 商汤科技于2016年开始布局车辆业务，聚焦舱驾两面，2022年底推出基于5R11V方案的高速NOA，而后持续发展布局城市等全场景道路。
- 2024年商汤绝影UniAD方案首秀，将感知、决策、规划等模块都整合到一个全栈Transformer端到端模型，实现感知决策一体化，不需要对感知数据进行抽象和逐级传递，“所见即所得”，将原始信息直接输入到端到端模型中，然后以自车轨迹规划为准进行指令输出。算力层面，商汤当前布局1.2EFLOPS的云端算力，支持更好的数据利用和更快的算法迭代。
- Drive AGI：数据驱动转向认知驱动的起点。类似于世界模型定义，Drive AGI具备对世界规律更进一步的深入理解，具备更强的推理能力、决策能力以及交互能力，是目前自动驾驶中最贴近人类思维模式、最能理解人类意图并有最强解决驾驶困难场景能力的技术方案；同时，商汤利用Drive AGI进一步打破座舱和智驾的界限，推动舱驾一体的架构变革。

图：商汤自动驾驶算法框架



- **量产产品维度：**商汤绝影针对32 TOPS到500 TOPS+的不同算力平台，提供适配的智能解决方案，满足全场景行泊一体智驾需求，最高平台可实现L3级别解放驾驶员双手双眼的高速领航功能。具备更高能力上限，更快迭代效率，更低系统成本等核心优势。
- 综合商汤自身在AI基础大模型软件和超算中心硬件的前期技术积淀，公司具备行业领先的人工智能基础设施【完善的自研底层工具链+超算中心】+算法强泛化性+主流平台适配+量产经验丰富+研发工具链丰富等优势，与本田、戴姆勒等全球车企以及国内主流新势力开展深度合作。

图：商汤自动驾驶算法解决方案



智能驾驶软件解决方案-初创企业汇总

智驾初创公司	发展类别	创立时间	智驾相关主营产品及解决方案	下游客户	创始团队	历史融资金额累计	研发团队规模
小马知行	L4与L2并进	2016	自动驾驶出行服务PonyPilot+、自动驾驶卡车、乘用车智能驾驶业务 ，基于 英伟达芯片 三款驾乘体验一流的智能驾驶方案：提供差异化的配置和智驾功能，量产智驾域控，提供闭环软件工具链	与丰田合作在国内推广无人驾驶出租车；与三一联合打造L4级智驾重卡，由一骥智卡生产，已向青骊物流交付；与一汽、广汽等合作乘用车智驾	百度系 前百度北美研发主任架构师彭军	11亿美元+	-
文远知行	L4与L2并进	2017	自动驾驶Robotaxi；Robobus无人巴士；WeRide L2++解决方案	广汽；奇瑞星途星纪元系列搭载WeRide L2++	百度系 韩旭 (Tony Han)，2007 UIUC CS PhD，曾任百度自动驾驶事业部首席科学家	10亿美元+	700+ (截至23年底)
驭势科技	L4	2016	自动驾驶Robotaxi；Robobus无人巴士；无人运货物流	已与上汽大众、中国一汽、上汽通用五菱、宇通客车达成商业合作	Tech firm 系 CEO. 吴甘沙，前英特尔中国研究院首席工程师	约13亿元	约400 (未知时间)
元戎启行	L4	2019	自动驾驶Robotaxi；L4级自动驾驶解决方案提供	东风风神、曹操出行合作自动驾驶Robotaxi	百度系 周光，2017 德州大学 AI PhD，曾在百度硅谷无人车团队负责标定、感知等方面的工作	4亿美元+	300+ (截至21年9月)
智加科技	L4	2018	重卡自动驾驶解决方案	一汽解放、江湖汽车	刘万千斯坦福电子工程系博士毕业，曾有过成功创业项目 InformLink；CTO郑皓曾是雅虎北京研究院院长	7亿元+	约400 (截至22年底)
轻舟智航	L4	2019	自动驾驶巴士：龙舟系列无人驾驶小巴士；轻舟矩阵智驾软件工具链	为理想提供算法开发辅助	谷歌Waymo系 4位创始人 于骞、侯聪、大方、汪堃	-	-
智行者	L4	2015	自动驾驶Robotaxi；无人运货物流	无人配送物流车 (和京东合作)、蜗小白无人驾驶环卫车 (和百度合作)	清华系 CEO张德兆：清华大学博士、博士后；CTO 王肖，清华 PhD	-	-
东软睿驰	L2+	2015	硬件 ：X-Cube系列L2级别智能前视摄像头、X-Box系列L2+及以上级别行泊一体域控制器、X-Center系列跨域式中央计算单元和驾驶员疲劳监测系统DMS产品等主流智能驾驶产品； 软件 ：提供自研的智驾生态软件工具包	软件赋能广汽集团星灵架构，与岚图签订战略合作协议，并与极氪、零跑、吉利等合作	依托东软集团强大软件能力赋能	6亿元+	-

五、投资建议与风险提示

■ 投资建议：汽车AI智能化转型大势所趋，算法为主干，看好头部算法玩家持续领先铸就高壁垒。

- **全行业加速智能化转型，产业趋势明确。** 下游OEM玩家+中游Tier供应商均加大对汽车智能化投入，大势所趋；智驾核心环节【软件+硬件+数据】均围绕下游OEM展开，数据催化算法提效进而驱动硬件迭代。以特斯拉为代表，应用算法向全栈端到端-世界模型持续迭代，功能落地兑现。
- OEM整车厂商&核心芯片硬件厂商&智驾传感器厂商&独立算法商加速布局端到端算法开发，场景驱动-数据驱动-认知驱动持续进化；智驾算法产业发展进入深水区，高投入赋能【大算力+大数据】，**方能走通L3有条件自动驾驶至L4完全自动驾驶之路。**
- **看好智驾头部车企以及智能化增量零部件：** 1) 华为系玩家【长安汽车+赛力斯+江淮汽车】，关注【北汽蓝谷】； 2) 头部新势力【小鹏汽车+理想汽车】； 3) 加速转型【吉利汽车+上汽集团+长城汽车+广汽集团】； 4) **智能化核心增量零部件：** 域控制器（德赛西威+经纬恒润+华阳集团+均胜电子等）+线控底盘（伯特利+耐世特+拓普集团等）。

■ 风险提示：

- **智能驾驶相关技术迭代/产业政策出台低于预期。** 若智能驾驶相关技术迭代节奏低于预期，可能会对消费者对智驾的认知和接受度产生影响，政策出台节奏低于预期也可能影响节奏。
- **华为/小鹏等头部车企新车销量低于预期。** 头部车企智驾新车销量表现低于预期，可能拖累智驾渗透率提升，对板块产生负面影响。

免责声明

东吴证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本研究报告仅供东吴证券股份有限公司（以下简称“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客户。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议，本公司及作者不对任何人因使用本报告中的内容所导致的任何后果负任何责任。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。

在法律许可的情况下，东吴证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

市场有风险，投资需谨慎。本报告是基于本公司分析师认为可靠且已公开的信息，本公司力求但不保证这些信息的准确性和完整性，也不保证文中观点或陈述不会发生任何变更，在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。

本报告的版权归本公司所有，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。经授权刊载、转发本报告或者摘要的，应当注明出处为东吴证券研究所，并注明本报告发布人和发布日期，提示使用本报告的风险，且不得对本报告进行有悖原意的引用、删节和修改。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

东吴证券投资评级标准

投资评级基于分析师对报告发布日后6至12个月内行业或公司回报潜力相对基准表现的预期（A股市场基准为沪深300指数，香港市场基准为恒生指数，美国市场基准为标普500指数，新三板基准指数为三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的），北交所基准指数为北证50指数），具体如下：

公司投资评级：

买入：预期未来6个月个股涨跌幅相对基准在15%以上；

增持：预期未来6个月个股涨跌幅相对基准介于5%与15%之间；

中性：预期未来6个月个股涨跌幅相对基准介于-5%与5%之间；

减持：预期未来6个月个股涨跌幅相对基准介于-15%与-5%之间；

卖出：预期未来6个月个股涨跌幅相对基准在-15%以下。

行业投资评级：

增持：预期未来6个月内，行业指数相对强于基准5%以上；

中性：预期未来6个月内，行业指数相对基准-5%与5%；

减持：预期未来6个月内，行业指数相对弱于基准5%以上。

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应当充分考虑自身特定状况，如具体投资目的、财务状况以及特定需求等，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。

东吴证券研究所
苏州工业园区星阳街5号
邮政编码：215021

传真：（0512）62938527

公司网址：<http://www.dwzq.com.cn>

东吴证券 财富家园