



# 计算机行业研究

买入（维持评级）

行业专题研究报告

证券研究报告

计算机组

分析师：刘高畅（执业 S1130525120005）  
liugaochang@gjzq.com.cn

分析师：郑元昊（执业 S1130525120004）  
zhengyuanhao@gjzq.com.cn

联系人：孙恺祈  
sunkaiqi@gjzq.com.cn

## Optimus V3 前瞻

### 行业观点

■ 怎么看人形机器人的商业化？——四象限的泛化与“干中学”的可能性。

在 2026 年 2 月 25 日的深度报告《具身智能逼近临界点，人形机器人商业化有望揭开序幕》中，我们曾领先市场率先提出了对于人形机器人商业化的思考框架：人形机器人所面向的工作环境必然是“高价值量+低重复性”的场景，上述场景又会在 ToB/C 客群差异、大脑/小脑层面泛化性要求的区分下形成四类买单方画像清晰且现阶段具身智能足以胜任的商业化场景——1) ToB+大脑泛化：导览、导购、导巡场景。2) ToB+小脑泛化：特种行业场景。3) ToC+大脑泛化：个人陪伴场景。4) ToC+小脑泛化：家庭照护场景。

结合前不久我们团队在硅谷调研的见闻，我们认为 Optimus 第一阶段的商业化，不一定是“完全自主替代人”，而更可能是一种“干中学”的过渡，将“劳动力”与“劳动场景”拆开，先做成一种 physical labor cloud 的类型。1) 从官方发心看，Tesla 自己对 Optimus 的官方定义，就是去做“危险、重复、无聊”的工作，而不是先从完全家用万能机器人开始。2) 从行业实践看，Sanctuary AI 公开承认很多公开视频其实是人工遥控操作，并将 teleoperation 视为采集行为数据、训练模型、验证硬件的关键环节。Contoro 将遥控操作模式商业化，它们让远程操作员在客户现场之外控制机器人处理 AI 搞不定的 edge cases，并且“从 day 1 提供可靠运行”，同时把这些人工接管继续喂给模型。3) 从商业逻辑看，以部分美国加州与德州蓝领工作为例，①工业喷涂危险性和重复性高，年均工资差距 38.0%；②商业清洁动作容错高、重复度高、夜班多，年均工资差距 34.8%；③仓储搬运/补货/分拣场景更结构化，且已经有 Contoro 在部署跑通商业模式，年均工资差距 14.8%。

■ 特斯拉 Optimus V3 发布在即，人形机器人有望迎量产元年。

埃隆·马斯克在过去半个月中持续为 Optimus V3 预热，3 月 21 日在 X 发言称把 Optimus 放进“自复制机器人/太空制造”的叙事中，3 月 25 日将 Raptor 3 类比 Optimus V3，暗示这是一次明显的代际升级，同一天发布视频，为 Optimus V3 预热；3 月 31 日在 X 发言称 Optimus V3 已能走动，还仍需最后一些润色才能对外展示。我们认为当前特斯拉 Optimus 已进入大规模量产前的供应链准备阶段，且未来量产指引将从上一个阶段的百万台指引逐渐演进到德州工厂驱动的千万台，行业空间有望十倍打开。

■ 人类视频数据 Scaling Law，顶尖具身大脑商业化试水。

数据侧：召唤感觉运动幽灵，NV EgoScale 验证人类视频数据的 Scaling Law。2026 年 2 月 19 日，NV EgoScale 框架基于 2 万+小时的带动作标签的第一人称(egocentric)人类视频数据集，验证了灵巧手操作泛化性的 Scaling Law。2 月底，NV GROOT 研究员 Danfei Xu 在 X 上发布长文，再次指出“人类数据本质上是伪装的机器人数据”的观点。①理想的人类数据：虽然形态不同、驱动和传感堆栈不同，但通过充分的传感器和状态估计，人类行为可以被捕捉并视 abc 为另一个“机器人”的数据，仍然是由相同物理规律支配的从感知到动作的映射。②相比遥控操作数据：遥控操作会扭曲人类行为，许多自然且细微的感觉运动智能从未进入数据集，而日常人类数据包含丰富的人与人之间的互动。③相比视频数据：在预测的视觉画面和物理执行之间仍然存在最后一英寸（或牛顿）的差距。几毫米的误差、几毫秒的延迟，或单个牛顿的力差都可能决定物体是滑落还是保持稳定。模型侧：顶尖具身大脑公司的商业化进入试水阶段。2 月 24 日，PI 发布 Blog，认为机器人行业也需要一个像大模型 API 那样可复用、可调用的“基础智能层”，从而让应用团队不必每次都从控制器、数据管线和模型训练重新搭起整套系统，同时给出了 Weave 和 Ultra 两家合作伙伴的标杆案例。我们认为，抛出商业化橄榄枝、展示合作伙伴案例秀肌肉本身证明 SOTA 级别的具身大脑渐渐进入可用阶段。

### 相关标的

- 1) 北美链：斯菱智驱、田中精机、科森科技、福赛科技、新泉股份、三花智控、恒立液压、北特科技、金力永磁、均胜电子等。
- 2) 其他机器人链：泛亚微透、唯科科技、领益智造、蓝思科技、禾赛、速腾聚创、海康威视、大华股份等。
- 3) AI 大脑及世界模型：协创数据、索辰科技、群核科技、智微智能、五一视界等。

### 风险提示

■ 北美人形机器人量产节奏不及预期的风险；通用机器人 Day1 L4 路线缺乏商业化基础的风险；仿真合成数据质量不及预期的风险；模型及软件解决方案三方公司长期产业链话语权较低的风险。



## 内容目录

一、怎么看人形机器人的商业化？——四象限的泛化与“干中学”的可能性	3
二、特斯拉 Optimus V3 发布在即，人形机器人有望迎量产元年	5
三、物理 AI：人类视频数据 Scaling Law，顶尖具身大脑商业化试水	8
四、相关标的	15
五、风险提示	15

## 图表目录

图表 1：人形机器人核心商业价值在于处理“高价值/低重复”任务	3
图表 2：对高价值/低重复场景的进一步拆解，ToB/C × 动脑/动手	4
图表 3：Optimus 的官方定义是去做“危险、重复、无聊”的工作	5
图表 4：Sanctuary AI 将遥操作在加拿大零售店落地部署	5
图表 5：Contoro 将遥操作模式商业化	5
图表 6：美国加州与德州部分蓝领工作年均工资差距高达 30+%.	5
图表 7：过去半个月 Elon Musk 持续为 Optimus V3 预热	6
图表 8：Elon Musk 3.31 在 X 发布视频为 Optimus V3 预热	6
图表 9：Optimus 技术逐步成熟，能力持续突破	6
图表 10：Optimus 量产规划明确，预计 2027 年底正式向公众发售	7
图表 11：NV EgoScale 框架验证人类视频数据的 Scaling Law	8
图表 12：NV EgoScale 2 万+小时的人类预训练视频	9
图表 13：PI 展示与 Weave 和 Ultra 两家合作伙伴的标杆案例	9
图表 14：UMI 设备由平行夹爪+软指+扳机构成	10
图表 15：多家具身智能大模型所用训练数据引入仿真合成数据	10
图表 16：银河通用 GraspVLA 大模型工作展示中，在不同光照条件下，模型都能精准执行抓取任务	12
图表 17：NVIDIA 机器人三大计算平台协同解决方案	12
图表 18：三台计算机之间闭环 workflows 与数据协同	12
图表 19：Omniverse 验证机器人步态，支持机器人学习	13
图表 20：机器人在 Isaac Sim 中的不同仿真环境下训练	13
图表 21：Omniverse 免费+订阅模式促进生态建设，形成软件驱动硬件销售模式	13
图表 22：谷歌发布通用世界模型 Genie 3 支持生成多样化交互环境	14
图表 23：LingBot-World 在高动态环境下仍能保持主体一致性	14
图表 24：Gemini 3.1 Pro 实现抽象推理能力翻倍跃升	15
图表 25：Gemini 3.1 Pro 视觉表现力实现显著提升	15



## 一、怎么看人形机器人的商业化？——四象限的泛化与“干中学”的可能性

在 2026 年 2 月 25 日的深度报告《具身智能逼近临界点，人形机器人商业化有望揭开序幕》中，我们曾领先市场率先提出了对于人形机器人商业化的思考框架：

人形机器人的商业化价值几乎必然面向“高价值/低重复”任务。我们将机器人面向的商业场景抽象为四大象限。第一象限（高重复、高价值）是典型的专用设备领域，如汽车组装车间的机械臂或医疗领域的达芬奇手术机器人，客户愿意为特定高价值任务买单，但设备泛化性弱；第二象限（高重复、低价值）多由通用型自动化设备主导，如仓储 AGV 或家用扫地机，以标准化方案解决低附加值劳动；我们认为，人形机器人由于初期研发成本较高，其必须面向具有高额支付意愿的“高价值”领域；同时，其通用形态与泛化智能的溢价，只有在面对非标准化、高度复杂的“低重复性”任务时才能得到真正释放。因此，人形机器人真正的星辰大海在于第四象限（低重复、高价值）。此外，第三象限（低重复、低价值）中长尾且琐碎的日常杂务（如捡拾零散纸团），目前并不具备独立的商业化土壤，我们预计这类场景将在第四象限核心技术成熟后，受益于技术外溢被顺带解决。

图表1：人形机器人核心商业价值在于处理“高价值/低重复”任务



来源：国金证券研究所

在“高价值、低重复”的核心市场中，基于机器人能力侧重点（大脑思考泛化/小脑动作泛化）与客群属性（ToB/ToC）的进一步解耦，人形机器人的商业化落地可划分为四大细分场景。

- 其一，多动脑×ToB端的营销服务场景，主要聚焦商业环境下的逻辑推理与信息交互。该场景对机器人物理运动能力要求较低，仅需完成平地行走、手势引导等基础动作，但对模型“大脑”的动态认知能力要求颇高。小鹏汽车全新人形机器人 IRON 率先落地自有门店“导览、导购、导巡”核心商业场景，在展厅动态复杂环境中，可实时调用 VLM/VLA 架构的底层物理大模型，除完成引路导览外，还能精准解析客户非标准化需求、自主开展产品答疑，并在多轮对话中挖掘客户痛点、输出定制化营销话术，核心商业价值体现在提升门店获客转化率、优化终端运营效率、实现服务人力成本替代上。



- 其二，多动脑 × ToC 端的个人陪伴场景，核心在于提供情绪价值与非结构化对话。该场景对机器人的物理动作复杂度要求有限，基础的肢体回应与表情反馈即可满足需求，但高度依赖大语言模型的语义共情、意图识别与长程多轮交互能力。以春晚小品中亮相的松延动力机器人为例，其面向老年群体的陪伴功能颇具代表性，不仅能流畅完成日常聊天、兴趣互动等对话任务，还可实时感知情绪变化并给予针对性回应。该场景核心是通过高频次的情感交互建立深度用户粘性，将机器人转化为家庭中的智能陪伴终端，从而挖掘长期的服务价值。
- 其三，多动手 × ToB 端的特种行业场景，对机器人的动态平衡、抗冲击力及全地形机动能力提出高要求。以宇树科技在 2026 年春晚表演中的表现为例，其技术底座已具备向特种场景落地的清晰可行性：一是具备较强的高动态运动控制能力，机器人在连续后空翻、武术动作中展现的姿态控制、落地缓冲与动态稳定能力，可为复杂地形通行、应急越障、高危环境作业提供运动基础；二是具备人机协同与精细力控能力，可完成精准器械操作、高速动作协同，并与人员安全无碰撞交互，感知与操控精度已满足协同作业、制式装备操作等场景的硬件条件；三是具备集群协同作业能力，20 余台机器人在直播环境下完成高速编队与队形变换，验证了低延迟通信、路径规划与避障算法的成熟度，形成了多机协同作业的技术闭环。更为关键的是，依托全栈自研，宇树将单机成本压降至 10 万元以内，而海外竞品波士顿动力 Spot 的售价约 53 万元且未实现规模化量产，其成本优势为规模化部署提供了商业基础，其早期产品甚至曾获美国海军陆战队的批量采购用以侦察运输。这种兼具高机动性能、集群协同潜力与低成本量产能力的产业化进展，已引发战略层面关注。2026 年 2 月 14 日，美国国防部将宇树科技列入“1260H”清单，实施投资与采购限制。我们认为，此类管制措施从侧面印证了高机动性小脑机器人在特种行业及关键领域的战略价值与应用潜力。
- 其四，多动手 × ToC 端的家庭照护场景，需要应对高度非结构化且随时变动的家庭物理环境。该场景需要机器人具备稳定的底盘移动能力，以适配居家多样的空间布局，同时要求针对各类柔性物料的精细力控与灵巧操作能力，精准把控施力幅度与动作精度。如银河通用机器人搭载 Sharpa 灵巧手，可灵活完成盘核桃、折叠衣物、平稳端水杯等精细化操作，精准适配居家日常各类操作需求，以此高效替代人类从事兼具精细度与实用性的高价值家政劳动，释放家庭照护中的人力成本。

图表2：对高价值/低重复场景的进一步拆解，ToB/C × 动脑/动手



来源：新京报，NE 时代智能体，国金证券研究所

结合前不久我们团队在硅谷调研的见闻，我们认为 Optimus 第一阶段的商业化，不一定是“完全自主替代人”，而更可能是一种“干中学”的过渡，将“劳动力”与“劳动场景”



拆开，先做成一种 physical labor cloud 的类型。

- 从官方发心看，Tesla 自己对 Optimus 的官方定义，就是去做“危险、重复、无聊”的工作，而不是先从完全家用万能机器人开始。

图表3: Optimus 的官方定义是去做“危险、重复、无聊”的工作



Tesla Optimus

Create a general purpose, bi-pedal, autonomous humanoid robot capable of performing unsafe, repetitive or boring tasks. Achieving that end goal requires building the software stacks that enable balance, navigation, perception and interaction with the physical world. We're hiring deep learning, computer vision, motion planning, controls, mechanical and general software engineers to solve some of our hardest engineering challenges.

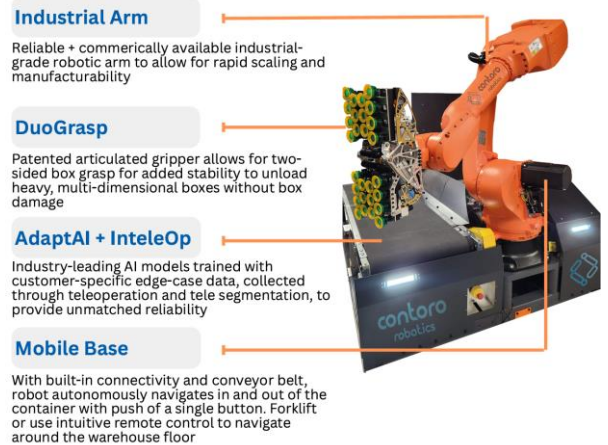
See Opportunities

来源: Tesla, 国金证券研究所

- 从行业实践看，Sanctuary AI 公开承认很多公开视频其实是人工遥操作，并将 teleoperation 视为采集行为数据、训练模型、验证硬件的关键环节，首个此类商业部署发生在加拿大不列颠哥伦比亚省兰利的 Mark's 零售店。CTC 旗下门店为期一周的试点成功在“真实”门店环境中测试了这款通用机器人，完成了 110 项零售相关任务，包括门店前后活动，如拣货和包装、清洁、贴标签、折叠等。Contoro 将遥操作模式商业化，它们让远程操作员在客户现场之外控制机器人处理 AI 搞不定的 edge cases，并且“从 day 1 提供可靠运行”，同时把这些人工接管继续喂给模型。

图表4: Sanctuary AI 将遥操作在加拿大零售店落地部署

图表5: Contoro 将遥操作模式商业化



来源: Sanctuary AI, 国金证券研究所

来源: Contoro, 国金证券研究所

- 从商业逻辑看，以美国加州与德州部分蓝领工作为例，1) 工业喷涂危险性和重复性高，年均工资差距 38.0%；2) 商业清活动作容错高、重复度高、夜班多，年均工资差距 34.8%；3) 仓储搬运/补货/分拣场景更结构化，且已经有 Contoro 在部署跑通商业模式，年均工资差距 14.8%。为“劳动力”与“劳动场景”的分离创造条件。

图表6: 美国加州与德州部分蓝领工作年均工资差距高达 30+%

岗位	California	Texas	年均工资差距
Painters, Construction and Maintenance	\$61,020	\$44,210	38.0%
Janitors and Cleaners	\$41,860	\$31,050	34.8%
Laborers and Freight Movers, Hand	\$43,350	\$36,930	17.4%

来源: U.S. BUREAU OF LABOR STATISTICS, 国金证券研究所

## 二、特斯拉 Optimus V3 发布在即，人形机器人有望迎量产元年

埃隆·马斯克在过去半个月中持续为 Optimus V3 预热，3月21日在 X 发言称把 Optimus 放进“自复制机器/太空制造”的叙事中，3月25日将 Raptor 3 类比 Optimus V3，暗示



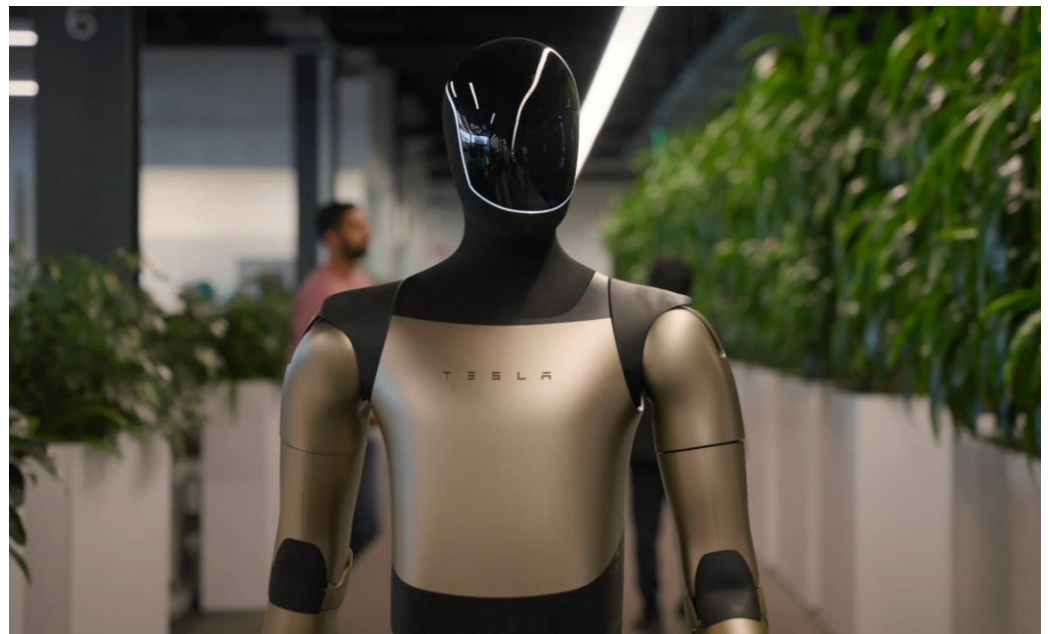
这是一次明显的代际升级，同一天发布视频，为 Optimus V3 预热；3 月 31 日在 X 发言称 Optimus V3 已能走动，还仍需最后一些润色才能对外展示。

图表7: 过去半个月 Elon Musk 持续为 Optimus V3 预热

时间	Elon Musk X 发言
2026-0321	把 Optimus 放进“自复制机器人/太空制造”叙事里，强调长期想象空间
2026-0325	用 Raptor 3 类比 Optimus V3，暗示这是一次明显的代际升级
2026-0325	发布视频，为 Optimus V3 预热
2026-0331	说明 Optimus V3 已能走动，还仍需最后一些润色才能对外展示

来源：X，国金证券研究所

图表8: Elon Musk 3.31 在 X 发布视频为 Optimus V3 预热



来源：X，国金证券研究所

特斯拉 OptimusV3 发布在即，推动人形机器人进入量产阶段。自 2022 年 Bumblebee 原型机开启研发进程以来，特斯拉通过 2023 年 Gen 1、Gen 2 型号的相继推出，在运动控制与环境感知领域实现了质变突破。2025 年 9 月，特斯拉发布 Gen 2.5 版本及其定型设计，进一步升级了灵巧手、运动性能及 AI 架构；同年 Optimus 顺利进入试生产阶段，特斯拉同步启动了弗里蒙特工厂试点产线的建设工作。2026 年 2 月，特斯拉官方正式宣布 Optimus V3 将于 2026 年第一季度亮相，马斯克接受采访称“Optimus 是我投入精力最多的项目，也将是人类历史上最伟大的项目。目前第三版设计已经定型，具备三大核心优势：人类级别的手部灵巧度、AI 大脑以及大规模量产能力”。

图表9: Optimus 技术逐步成熟，能力持续突破

时间	2021 年 8 月	2022 年 9 月	2023 年 3 月	2023 年 12 月	2025 年 9 月
事件	发布特斯拉人形机器人计划，提出人形机器人概念 Tesla Bot，计划将于 2022 年推出该款机器人的原型机	发布 Optimus 原型机 Bumblebee，展示接近量产的第一代 Optimus	发布进展视频，推出 Optimus Gen1	首次线上发布 Optimus Gen2	在 NeurIPS 大会上正式发布 Optimus Gen2.5



<p><b>特征</b></p>	<p>定位通用型机器人，目标代替人类从事危险、重复、无聊的工作</p>	<p>原型机周身直接裸露电线，直接呈现其内部的关节、电路板、电缆等设备，可以直观地向观众展示每个动作细节。</p>	<p>在运动控制、环境感知、动作捕捉等方面取得突破性进展</p>	<p>在核心部件自研化、感知范围、机械机构轻量化、足部仿生设计、运动控制、末端执行器精细化等方面取得突破性进展</p>	<p>硬件平台正式定型，灵巧手升级到22个自由度，能完成精细操作。</p>
<p><b>技术细节</b></p>	<p>全身采用轻量化材料，共有40个电动执行器：手臂12个 + 手部12个 + 腿部12个 + 颈部2个 + 躯干2个。此外，脚部安装有反馈传感器，并采用2轴平衡设计</p>	<p>原型机能够准确定位周围人员或障碍物，主动避让，同时还能执行汽车零部件抓取和放置</p>	<p>端到端操作中，系统仅需视觉输入即可完成复杂任务，全程无需预设代码</p>	<p>新增2自由度颈部关节，全新11自由度的灵巧手结构，每个手指均配备触觉传感器，</p>	<p>头部及躯干主承载结构已由单一金属转向嵌塑件与注塑件方案，在维持强度的同时实现轻量化，首次在量产版集成高性能织物手套兼顾灵敏度与低维护成本</p>
<p><b>图示</b></p>					

来源：人形机器人联盟，慧博资讯，大蒜机器人所国金证券研究所

Optimus 迎量产元年，加州百万台生产线正加速建设，德州千万台产能正加速规划。马斯克在特斯拉 2025 年股东大会上表示，“首先在加州弗里蒙特工厂建立年产 100 万台的生产线，将于 2026 年投产，随后将在德州工厂建设年产 1000 万台的第二条生产线，有望于 2027 年投产”。马斯克还表示，“我认为到今年（26 年）年底，Optimus 将能够完成更复杂的任务，不过它仍将主要被部署在工业环境。明年（27 年）年底，我们将正式向公众（C 端）销售 Optimus”。我们认为当前特斯拉 Optimus 已进入大规模量产前的供应链准备阶段，且未来量产指引将从上一个阶段的百万台指引逐渐演进到德州工厂驱动的千万台，行业空间有望十倍打开。

图10: Optimus 量产规划明确，预计 2027 年底正式向公众发售

时间	事件	量产信息表述
2022 年 10 月	特斯拉 AI DAY	预计 3-5 年后可订购 Optimus
2024 年 1 月	2023Q4 财报电话	Optimus 2025 年可能开始出货
2024 年 6 月	特斯拉股东大会	宣布 Optimus 将于 2025 年有限量产;预计超 1000 个机器人在特斯拉工厂工作
2025 年 1 月	2024 Q4 财报电话会	Optimus2025 年将进入试生产阶段;计划生产 5000 台;已订购 1-1.2 万台组件;目标 2026 年达 5 万台。
2025 年 3 月	社交平台推文	目前设计产线月产能为 1000 台;预计 2026 年下半年可实现向外部客户交付。
2025 年 7 月	2025 Q2 财报电话会	2026 年开启大规模生产;目标五年内实现年产百万台



2025 年 9 月	2025 Q3 财报交流	明确 Optimus V3 展示将于 2026 年 Q1;2026 年底启动大规模产线。
2026 年 1 月	达沃斯论坛	2027 年底正式向公众 C 端发售。

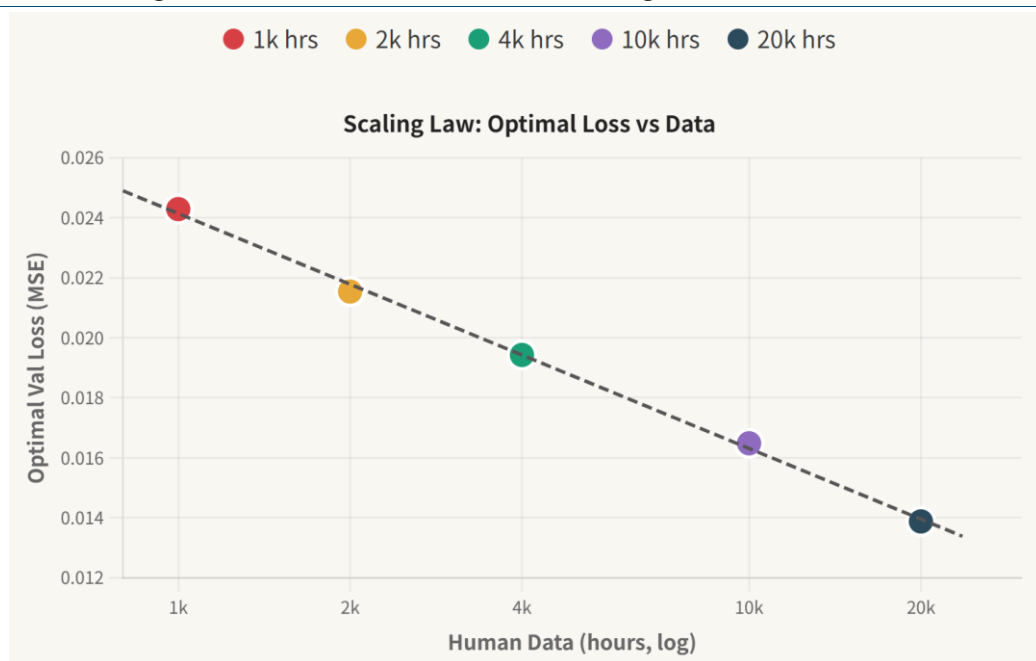
来源：新浪科技，腾讯科技，界面新闻，Auto 有范儿，证券时报，海豚研究，华尔街见闻，国金证券研究所

### 三、物理 AI：人类视频数据 Scaling Law，顶尖具身大脑商业化试水

数据侧：召唤感觉运动幽灵，NV EgoScale 验证人类视频数据的 Scaling Law。

2026 年 2 月 19 日，NV EgoScale 框架基于 2 万+小时的带动作标签的第一人称(egocentric)人类视频数据集，验证了灵巧手操作泛化性的 Scaling Law。研究团队认为，人类行为是学习物理智能最具可扩展性的数据来源之一，EgoScale 框架是一个基于大规模自我中心人类数据构建的人机灵巧操作迁移框架，在超过 20,854 小时的动作标注视频数据（比先前研究规模大 20 倍以上）上训练视觉-语言-动作（VLA）模型，发现人类数据规模与验证损失之间存在对数线性扩展规律。

图表11：NV EgoScale 框架验证人类视频数据的 Scaling Law



来源：NVIDIA，国金证券研究所

人类数据本质上是伪装的机器人数据。2 月底，NV GROOT 研究员 Danfei Xu 在 X 上发布长文，再次指出“人类数据本质上是伪装的机器人数据”的观点。

- 理想的人类数据：虽然形态不同、驱动和传感堆栈不同，但通过充分的传感器和状态估计，人类行为可以被捕捉并视为另一个“机器人”的数据，仍然是由相同物理规律支配的从感知到动作的映射。当前人类数据仍然严重依赖视觉，缺少如声音、力和接触反馈，所以只能获取动作和结果，但很难获取产生它们的控制命令。
- 相比遥操作数据：遥操作会扭曲人类行为，许多自然且细微的感觉运动智能从未进入数据集，而日常人类数据包含丰富的人与人之间的互动。
- 相比视频数据：在预测的视觉画面和物理执行之间仍然存在最后一英寸（或牛顿）的差距。几毫米的误差、几毫秒的延迟，或单个牛顿的力差都可能决定物体是滑落还是保持稳定。



图表12: NV EgoScale 2 万+小时的人类预训练视频



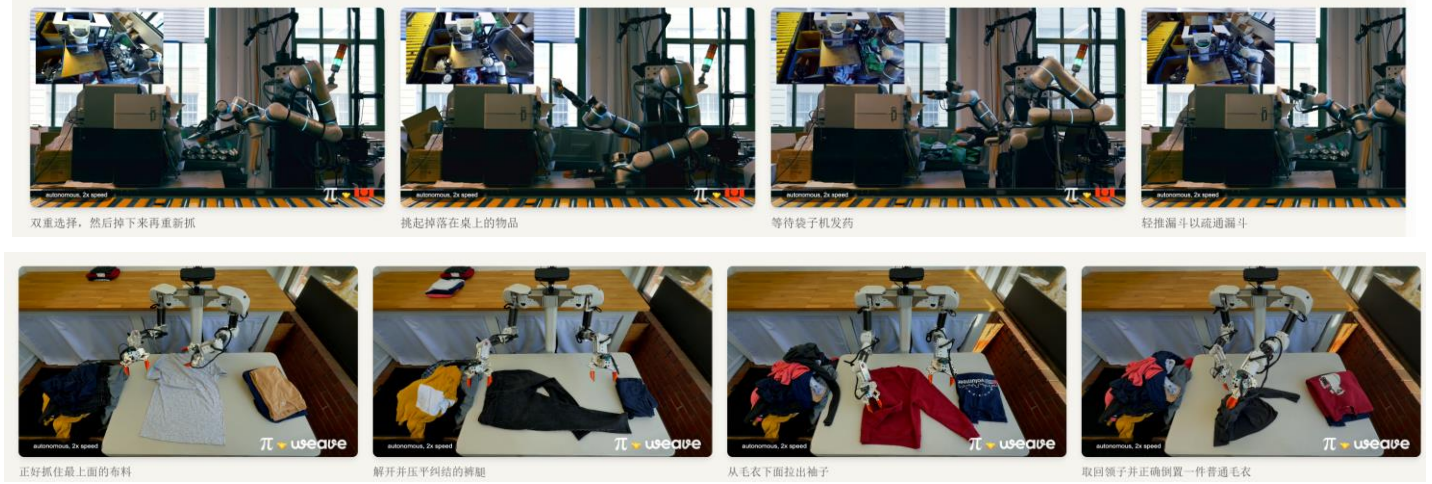
来源: NVIDIA, 国金证券研究所

模型侧: 顶尖具身大脑公司的商业化进入试水阶段

2月24日, PI 发布 Blog, 认为机器人行业也需要一个像大模型 API 那样可复用、可调用的“基础智能层”, 从而让应用团队不必每次都从控制器、数据管线和模型训练重新搭起整套系统, 同时给出了 Weave 和 Ultra 两家合作伙伴的标杆案例。

我们认为, 抛出商业化橄榄枝、展示合作伙伴案例秀肌肉本身证明 SOTA 级别的具身大脑渐渐进入可用阶段。但与 LLM 不同, 具身大脑 API 的商业化在现阶段面临两个追问: 其一是 API 通用性的追问: 具身大脑的通用性面临跨本体与跨场景的挑战, API 是一个非常标准化的产品, 但要在真实世界, 多样的机器人本体上部署还是要经过非常多真实世界的调试, 这可能会限制应用的广泛性。其二是机器人公司灵魂的追问: 具身模型的技术路线可能还未真正收敛, 即便强如 PI 的大脑智能程度也相对初级, 目前主要玩家还不甘心只做应用, 还会持续在基模方向发力。

图表13: PI 展示与 Weave 和 Ultra 两家合作伙伴的标杆案例



来源: PI, 国金证券研究所

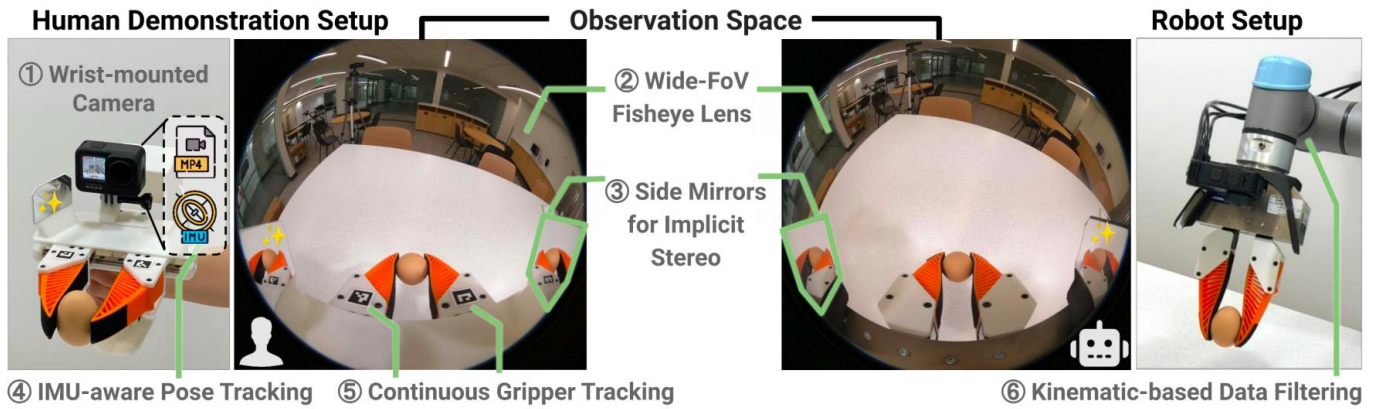
真实机器采集数据是具身智能迭代的核心刚需, 轻量化采集设备的技术突破推动数据获取实现降本提效, UMI 方案则进一步破解传统采集模式的痛点, 成为轻量化真实数据采集的关键路径。

UMI (Universal Manipulation Interface) 作为一种新型轻量化真实数据采集方案, 其核心在于实现数据采集过程与机器人本体的完全解耦。该系统硬件极为精简, 仅由一套手持式 3D 打印平行夹具与腕部视角相机构成, 整套设备成本控制在 400 美元左右。采集过程中, 相机同步



记录相对末端轨迹，而非传统方式中机械臂的绝对关节角度。这种设计将数据采集设备从昂贵的工业机械臂简化为一套“即拿即用”的手持终端，本质上是将机器人操作数据的生产门槛从专业实验室级别降维至消费电子级别。

图表14: UMI 设备由平行夹爪+软指+扳机构成



来源: CSDN, 国金证券研究所

相较于传统采集方案，UMI 在采集场景、数据泛化、成本控制三个核心维度实现突破：

- 采集环境从封闭式向开放式拓展，摆脱对数采工厂的依赖，保障数据多样性与采集高效性。传统遥操作采集依赖昂贵的工业机械臂与动捕设备，仅能在封闭实验室或专业数采工厂开展采集工作。而 UMI 凭借手持终端的轻量化优势，无需固定场地及复杂配套设备，采集人员可便捷进入家庭、办公室、户外等各类非结构化真实场景完成采集，显著丰富了数据的场景多样性与真实度，使采集数据更贴合具身智能的实际应用需求。
- 数据泛化能力显著提升，可实现高质量数据的跨设备复用。传统采集模式高度依赖特定机器人的关节角度数据，导致数据通用性差、难以跨构型复用。而 UMI 以腕部第一视角相机作为唯一观测点，避免采集者体型差异带来的视觉误差，同时重点记录末端执行器的轨迹与姿态，而非具体关节参数。这种“重末端、轻本体”的表征方式，使采集数据具备极强的硬件无关性，可无缝迁移至不同构型、不同尺寸的机器人训练中，有效解决了传统数据复用性差的问题，实现数据价值最大化。
- 商业经济性优势突出，可将真实数据的采集成本击穿至消费级区间。传统真机采集属于重资产投入，高额的场地租金、机械臂折旧及设备运维成本，制约了数据的规模化获取。而 UMI 凭借硬件精简、与机器人本体解耦的设计，有效剥离了高额固定成本。运营成本层面，其部署效率较高，无需搭建固定工装，也无需调试机械臂与动捕系统的联动参数，开机即可启动采集；同时，其采集速度达到传统遥操作模式的 3 倍，可在相同时间内完成更多场景的采集任务，提升单位时间样本产出量。双重优势叠加下，UMI 将采集成本击穿至消费级区间，为具身智能大规模数据采集提供了财务可行性。

一般而言，真机数据的优势在于其高保真度（物理交互的真实性），能够精准反映复杂操作中的力学反馈和不确定性，劣势在于采集成本高昂、采集周期长且现阶段数量较少。仿真合成数据的优势在于低成本、高效率与场景的可扩展性，基于物理引擎可快速生成海量的多样化数据，尤其适合预训练数据积累与极端场景的模拟，劣势在于“仿真鸿沟”——虚拟环境与物理现实的偏差导致模型在真实任务中表现不稳定。目前，英伟达、Skild AI、银河通用三家厂商已明确在训练具身智能大模型时采用了仿真合成数据。

图表15: 多家具身智能大模型所用训练数据引入仿真合成数据

模型名称	厂商	模型类别	模型架构	训练数据类型
RT-2	谷歌 DeepMind	VLA 大模型	单模型架构，以 VLM 模型（PaLI-X 或 PaLM-E）为主干网络	联合微调：大规模互联网数据 + 机器人真机数据



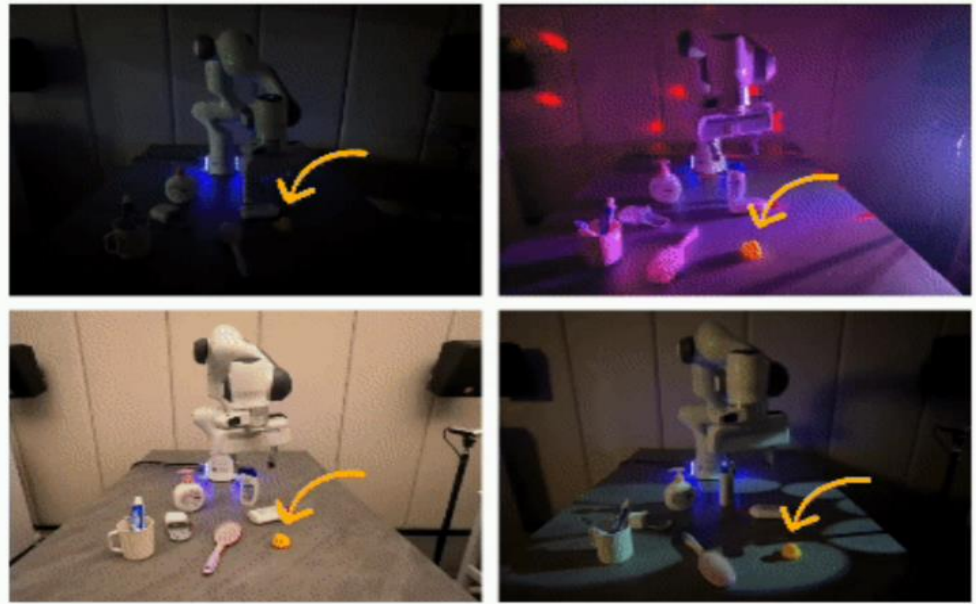
$\pi 0$	Physical Intelligence	VLA 大模型	分层双系统架构: 1) 预训练 VLM: 3B 参数的 PaliGemma; 2) 动作专家模块: 300M 参数规模	1) 预训练: -VLM: 大规模互联网数据; -动作专家模块: 开源真机数据集和基于遥操作采集的真机数据; 2) 后训练: 高质量真机数据
Helix	Figure AI	VLA 大模型	分层双系统架构: 1) 系统 S1: 80M 参数的 Transformer 模型; 2) 系统 S2: 7B 参数的预训练 VLM 模型	1) 系统 S1: 机器人真机数据; 2) 系统 S2: 大规模互联网数据
GR00T N1	英伟达	VLA 大模型	分层双系统架构: 1) 系统 S1: 基于扩散变换器 (DIT) 的动作模块; 2) 系统 S2: 预训练 VLM——Eagle-2	预训练: 真实机器人演示数据、 <b>合成数据</b> 以及互联网上的人类视频数据
Skild Brain	Skild AI	基于视觉的端到端运动模型	分层架构: 用低频率的高层动作策略为高层动作策略提供输入, 而且适用于几乎所有机器人。	预训练: <b>仿真环境与人类操作视频</b> ;后训练: 借助每台联网机器人的真实运行数据进行微调
智元启元大模型 (GO-1)	智元机器人	VLM 基础上发展的 ViLLA 大模型	ViLLA 架构由 VLM(多模态大模型) + MoE(混合专家)组成	AgiBot World 的大规模高质量真机数据+互联网大规模异构视频数据
Seed GR-3	字节跳动	VLA 大模型	采用 Mixture-of-Transformers (MoT) 的网络结构, 把“视觉-语言模块”和“动作生成模块”结合成了一个 40 亿参数的端到端模型	遥操作机器人收集的高质量真机数据+基于 VR 设备的人类轨迹数据+公开、可用的大规模视觉语言数据
ERA-42	星动纪元	VLA 模型	分层系统架构; 1) 高层次规划: 7B 参数的 InstructBLIP 视觉语言模型; 2) 低层次控制: 40M 参数的 Transformer 架构模型	人类操作数据+互联网机器人数据+遥操作方式获取的数据+大规模视频数据 + 机器人数据
GraspVLA	银河通用	VLA 模型	\	1) 预训练: <b>大规模合成数据</b> ; 2) 后训练: 少部分真机数据

来源: Physical Intelligence 工作论文《 $\pi 0$ : A Vision-Language-Action Flow Model for General Robot Control》, 星动纪元官方公众号, 42 号电波公众号, 智元机器人公众号, 字节跳动 Seed 公众号, 银河通用机器人公众号, 国金证券研究所

真机数据采集路线的典型代表是北美具身智能明星企业 Physical Intelligence (简称 PI), 有报道称 PI 租下了 Airbnb 在旧金山的许多公寓用于真实空间和场景的数据采集, 成本极为高昂。仿真合成数据的代表以银河通用为例, 其基于 NVIDIA Isaac 平台构建高效的数据生产管线, 通过升级物理真实性与渲染并行度, 仅一周即生成全球规模最大的十亿级机器人操作数据集。作为全球首个完全基于仿真合成大数据进行预训练的具身大模型, GraspVLA 展现出比 OpenVLA、 $\pi 0$ 、RT-2、RDT 等模型更强大的泛化能力: 光照泛化、干扰物泛化、平面位置泛化、高度泛化、背景泛化、物体类别泛化。



图表16: 银河通用 GraspVLA 大模型工作展示中, 在不同光照条件下, 模型都能精准执行抓取任务

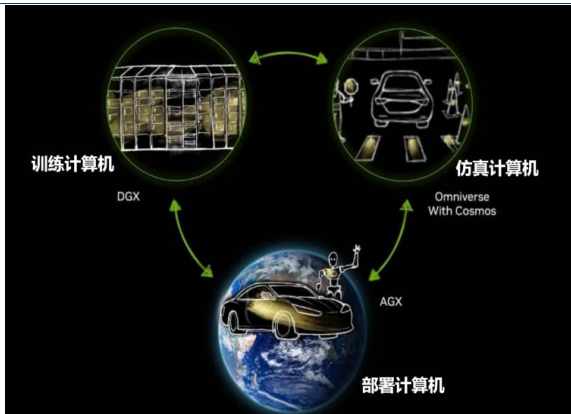


来源: 新智元公众号, 国金证券研究所

NVIDIA CEO 黄仁勋于 2025 年 CES 大会上指出, 每家机器人公司最终都必须构建三台基础计算机协同的解决方案, 形成从训练到优化再到执行的完整体系。对于英伟达而言, “第一台计算机”是 DGX AI 超级计算机, 用于训练。“第二台计算机”是基于 NVIDIA RTX PRO 服务器的 NVIDIA Omniverse 和 Cosmos, 用于生成合成数据、进行强化学习, 同时也是 DGX 和 AGX 的桥梁。“第三台计算机”是 NVIDIA 为边缘计算和自主系统设计的嵌入式平台 NVIDIA Jetson AGX Thor, 用于端侧部署。从训练到仿真再到部署, “三台计算机”打造了英伟达在通用机器人完整框架, 可以缩短人形机器人的开发周期、降低开发成本和风险, 构建开放的生态系统。

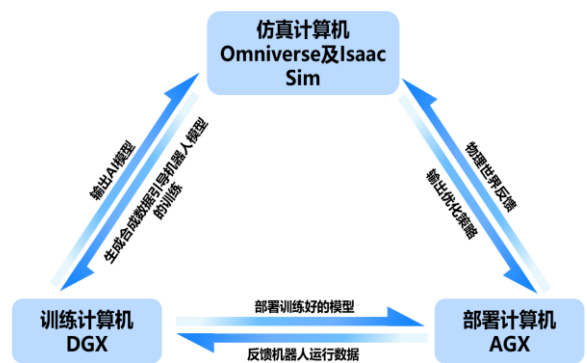
第一步, 在“第一台计算机”上训练模型, 开发者可以在 NVIDIA DGX 平台上使用 NVIDIA NeMo 来训练和微调 AI 模型, 还可以利用 NVIDIA Project GR00T 使人形机器人能够理解自然语言, 并能够通过观察人类行为来模仿动作。第二步, 开发者可利用 Omniverse 和 Cosmos 生成海量符合物理特性的多样化合成数据, 例如 2D 或 3D 图像、分割掩码、深度图或运动轨迹数据, 为模型训练和性能优化奠定基础; 开发者在基于 Omniverse 构建的 NVIDIA Isaac Sim 的仿真环境中零风险验证其机器人策略; 还可以使用“第二台计算机”的 Isaac Lab (一个开源的机器人学习框架), 赋能机器人的强化学习和模仿学习, 加速完善机器人的训练策略。第三步, 训练好的 AI 模型被部署到端侧计算机上。

图表17: NVIDIA 机器人三大计算平台协同解决方案



来源: NVIDIA 官网博客, 国金证券研究所

图表18: 三台计算机之间闭环 workflow 与数据协同



来源: NVIDIA 官网博客, 国金证券研究所

近年来, 英伟达不断强调旗下图形部门中仿真条线的重要性, 即“第二台计算机”的重要性。美国科技媒体《The Information》深入报道了英伟达创始人兼 CEO 黄仁勋子女在该公司任职的情况。报道指出, 黄仁勋现年 34 岁的女儿黄敏珊 (Madison Huang) 于 2020 年加入英伟达,



现任 Omniverse 和机器人技术领域的资深产品营销主管。其 35 岁的儿子黄胜斌 (Spencer Huang) 于 2022 年加入英伟达, 目前担任机器人项目的产品经理。

黄仁勋的管理理念与传统企业不同, 他相信企业架构的扁平化结构与信息的极度透明可以提高企业运转效率, 因此公司高管直接向黄仁勋汇报工作。仿真技术副总裁和英伟达其他重要业务部门副总裁 (如 GPU 工程高级副总裁、DGX Cloud 业务副总裁等) 一并直接向黄仁勋汇报工作。从黄仁勋家族成员在公司的角色配置和岗位安排, 再到仿真技术副总裁的直接汇报, 不难看出黄仁勋本人对于英伟达仿真业务的重视程度之深。

Omniverse 是一个用于大规模构建 3D 应用程序及服务的开发平台, 该平台基于 OpenUSD (通用场景描述) 和 NVIDIA RTX 渲染技术, 支持实时协作、物理精准仿真和生成式 AI 集成, 主要应用于工业数字化、机器人训练、自动驾驶仿真和元宇宙等领域, 已成为工业设计、数字孪生、AI 开发等领域的核心工具。

NVIDIA Isaac Sim 是一款基于 Omniverse 构建的开源应用, 使开发者能够在基于物理的虚拟环境中模拟和测试 AI 机器人解决方案。Isaac Sim 有以下三大功能:

- 生成合成数据: Isaac Sim 支持大规模合成数据生成, 包括感知、移动、基于物理的抓取等, 并提供写实渲染和自动生成真实标签, 用于训练和微调机器人基础模型。
- 执行软件在环测试: Isaac Sim 可通过与真实机器人软件集成, 为完整的机器人堆栈实现软件在环 (software-in-the-loop) 测试, 从而验证机器人控制与感知系统。
- 支持机器人学习: Isaac Lab 是基于 Isaac Sim 平台构建的开源轻量级应用, 专为大规模机器人学习进行优化。通过 Isaac Lab 支持机器人学习流程, 可加速仿真中的训练, 助力模型在现实场景中的快速部署。

图表19: Omniverse 验证机器人步态, 支持机器人学习

图表20: 机器人在 Isaac Sim 中的不同仿真环境下训练



来源: NVIDIA 英伟达企业解决方案公众号, 国金证券研究所

来源: NVIDIA 英伟达企业解决方案公众号, 国金证券研究所

目前, 个人创作者、设计师和开发者可免费使用 NVIDIA Omniverse 的基础版本, 目的是降低设计技术门槛, 吸引用户参与生态建设, 通过培养用户群体, 推动内容创作和技术创新, 进而间接带动企业端需求增长。针对企业用户, NVIDIA Omniverse 专门推出了 Omniverse Enterprise 软件, 企业用户可拥有 90 天试用期, 试用期过后的订阅费用为每 GPU 每年 4,500 美元。但该软件试用版要求拥有搭载 NVIDIA RTX 的工作站或服务器 (官网建议要求具有 16GB VRAM 的支持 RTX 的 GPU)。这种销售模式将在一定程度上带动硬件产品 (GPU 等) 销量增长, 形成“软件驱动硬件”的协同效应。

图表21: Omniverse 免费+订阅模式促进生态建设, 形成软件驱动硬件销售模式

客户类型	使用版本	付费规则	目的
C 端	NVIDIA Omniverse 基础版本	免费使用	吸引用户参与生态建设, 推动内容创作和技术创新, 进而间接带动企业端需求增长
B 端	Omniverse Enterprise 软件	客户拥有 90 天试用期, 试用期过后的订阅费用为每 GPU 每年 4,500 美元	软件试用版要求拥有搭载 NVIDIA RTX 的工作站或服务器, 进而带动硬件产品销量增长, 形成“软件驱动硬件”的协同效应

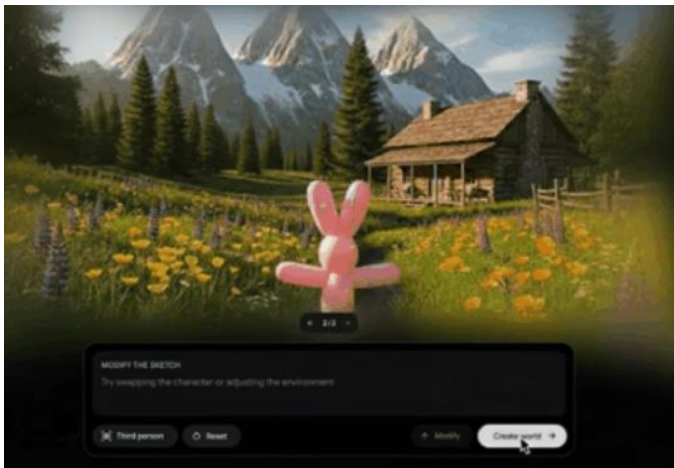


来源: NVIDIA 官网, 国金证券研究所

世界模型的持续迭代, 使机器人可在虚拟环境中精准预判物理规律与动态变化。北美方面, 谷歌 DeepMind 发布的 Genie 3 首次实现了仅凭文本提示即可生成 720P 分辨率、24FPS 的高动态三维交互世界, 该模型不仅具备长达数分钟的物理一致性与空间记忆, 还能对导航、环境干预等外部操作做出符合物理定律的实时响应。国内方面, 蚂蚁灵波科技全面开源了 LingBot-World 及其衍生模型, 实现了极低延迟的键盘/鼠标实时操控与长达 10 分钟的无损连续生成, 克服了长时漂移与细节塌陷问题, 且首创了自回归“视频-动作”一体化框架, 支持机器人同步推演未来物理状态并输出动作指令。这种内化了物理因果律、空间结构与长时交互能力的新一代世界模型, 为具身智能提供了可靠的虚拟仿真与交互推演支撑, 其生成的海量高质量交互合成数据, 可缓解机器人产业的数据供给瓶颈。

图表22: 谷歌发布通用世界模型 Genie 3 支持生成多样化交互环境

图表23: LingBot-World 在高动态环境下仍能保持主体一致性



来源: 新智元, 国金证券研究所

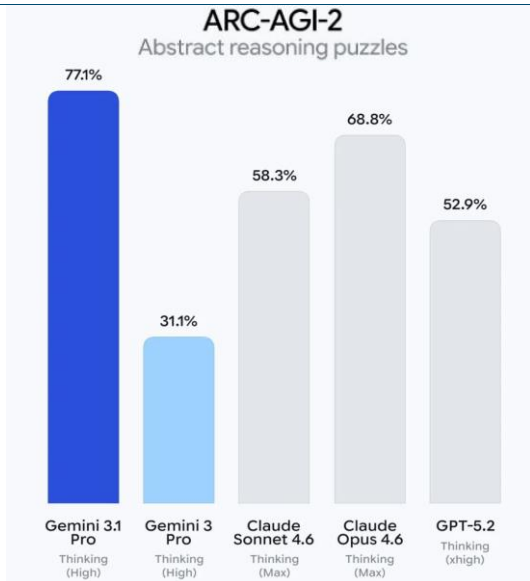
来源: 智东西, 国金证券研究所

AI 大模型多模态理解与生成能力的快速进步, 使得海量人类第一人称及第三人称视频, 正高效转化为机器人训练数据。此前, 受人手与机器夹爪的跨本体形态差异限制, 人类视频难以直接用于机器人训练。随着多模态大模型持续迭代, 该技术瓶颈已被快速突破, 模型可将人类视频中的动作语义精准映射至机器本体。目前, Tesla Optimus、Figure 等北美头部企业已大规模采用人类视频数据开展模型训练。

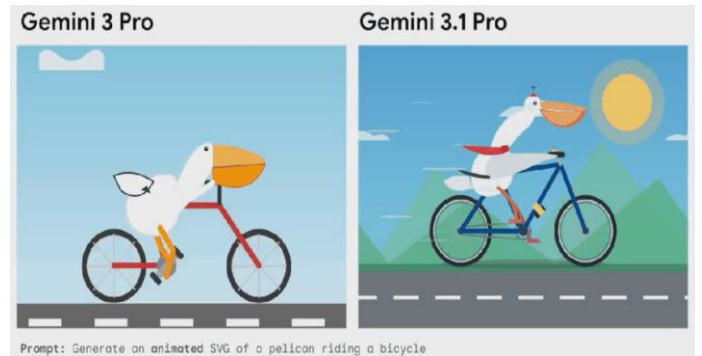
谷歌此前发布的 Gemini 3.0 Pro 已展现出强劲的视觉解析能力: 在多模态屏幕理解任务 (Screenshot-Pro) 中得分达 72.7%, 显著高于 Gemini 2.5 Pro (11.4%) 和 Claude 4.5 (36.2%); 在高阶视觉逻辑推理指标 ARC-AGI-2 上, 得分由 Gemini 2.5 Pro 的 4.9% 提升至 31.1%, 同样领先 GPT-5.1 (17.6%)。在此基础上, 最新迭代的 Gemini 3.1 Pro 实现能力再次跃升, 其 ARC-AGI-2 抽象推理指标得分高达 77.1%, 较 3.0 版本实现翻倍提升, 并大幅领先同期的 Claude Opus 4.6 (68.8%) 与 GPT-5.2 (52.9%)。核心指标的快速提升, 印证了多模态大模型在空间抽象与逻辑推理能力上的高速进化。这一底层能力突破, 将显著提升人类视频向机器人训练数据的转化与映射效率, 充分盘活全网海量存量视频资源, 为机器人理解真实物理世界提供持续、高质量的数据支撑。



图表24: Gemini 3.1 Pro 实现抽象推理能力翻倍跃升



图表25: Gemini 3.1 Pro 视觉表现力实现显著提升



来源: Google DeepMind, 国金证券研究所

来源: Google, 国金证券研究所

## 四、相关标的

- 1) 北美链: 斯菱智驱、田中精机、科森科技、福赛科技、新泉股份、三花智控、恒立液压、北特科技、金力永磁、均胜电子等。
- 2) 其他机器人链: 泛亚微透、唯科科技、领益智造、蓝思科技、禾赛、速腾聚创、海康威视、大华股份等。
- 3) AI 大脑及世界模型: 协创数据、索辰科技、群核科技、智微智能、五一视界等。

## 五、风险提示

- 北美人形机器人量产节奏不及预期的风险。

目前产业聚焦在北美量产指引的兑现节奏上,若新版本机器人发布效果不及预期或量产指引持续 miss,或对产业链放量节奏产生一定扰动。

- 通用机器人 Day1 L4 路线缺乏商业化基础的风险。

目前通用机器人产业发展尚处早期,人形机器人各项软硬件环节成熟度参差不齐,Day1 L4 路线在展业过程中缺乏足够容量的买单客群,或存在商业闭环到数据闭环都无法建立的风险。

- 仿真合成数据质量不及预期的风险。

目前“缺数据”仍是机器人落地的核心卡点之一,假如基于 Sim2Real 模式的仿真合成数据质量不及预期,或对英伟达 Isaac Sim、索辰机器人虚拟平台等主体的发展带来扰动。

- 模型及软件解决方案三方公司长期产业链话语权较低的风险。

参考部分智能驾驶算法厂商与主机厂之间的角力关系,若未来机器人产业渐趋成熟,在国内商业生态普遍对于软件厂商不算友好的环境下,或出现三方数据商、模型算法商产业链话语权较低的风险。



**行业投资评级的说明:**

买入: 预期未来 3—6 个月内该行业上涨幅度超过大盘在 15%以上;

增持: 预期未来 3—6 个月内该行业上涨幅度超过大盘在 5%—15%;

中性: 预期未来 3—6 个月内该行业变动幅度相对大盘在 -5%—5%;

减持: 预期未来 3—6 个月内该行业下跌幅度超过大盘在 5%以上。



**特别声明：**

国金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。

本报告版权归“国金证券股份有限公司”（以下简称“国金证券”）所有，未经事先书面授权，任何机构和个人均不得以任何方式对本报告的任何部分制作任何形式的复制、转发、转载、引用、修改、仿制、刊发，或以任何侵犯本公司版权的其他方式使用。经过书面授权的引用、刊发，需注明出处为“国金证券股份有限公司”，且不得对本报告进行任何有悖原意的删节和修改。

本报告的产生基于国金证券及其研究人员认为可信的公开资料或实地调研资料，但国金证券及其研究人员对这些信息的准确性和完整性不作任何保证。本报告反映撰写研究人员的不同设想、见解及分析方法，故本报告所载观点可能与其他类似研究报告的观点及市场实际情况不一致，国金证券不对使用本报告所包含的材料产生的任何直接或间接损失或与此有关的其他任何损失承担任何责任。且本报告中的资料、意见、预测均反映报告初次公开发布时的判断，在不作事先通知的情况下，可能会随时调整，亦可因使用不同假设和标准、采用不同观点和分析方法而与国金证券其它业务部门、单位或附属机构在制作类似的其他材料时所给出的意见不同或者相反。

本报告仅为参考之用，在任何地区均不应被视为买卖任何证券、金融工具的要约或要约邀请。本报告提及的任何证券或金融工具均可能含有重大的风险，可能不易变卖以及不适合所有投资者。本报告所提及的证券或金融工具的价格、价值及收益可能会受汇率影响而波动。过往的业绩并不能代表未来的表现。

客户应当考虑到国金证券存在可能影响本报告客观性的利益冲突，而不应视本报告为作出投资决策的唯一因素。证券研究报告是用于服务具备专业知识的投资者和投资顾问的专业产品，使用时必须经专业人士进行解读。国金证券建议获取报告人员应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。报告本身、报告中的信息或所表达意见也不构成投资、法律、会计或税务的最终操作建议，国金证券不就报告中的内容对最终操作建议做出任何担保，在任何时候均不构成对任何人的个人推荐。

在法律允许的情况下，国金证券的关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并可能为这些公司正在提供或争取提供多种金融服务。

本报告并非意图发送、发布给在当地法律或监管规则下不允许向其发送、发布该研究报告的人员。国金证券并不因收件人收到本报告而视其为国金证券的客户。本报告对于收件人而言属高度机密，只有符合条件的收件人才能使用。根据《证券期货投资者适当性管理办法》，本报告仅供国金证券股份有限公司客户中风险评级高于C3级（含C3级）的投资者使用；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，不应被视为对特定客户关于特定证券或金融工具的建议或策略。对于本报告中提及的任何证券或金融工具，本报告的收件人须保持自身的独立判断。使用国金证券研究报告进行投资，遭受任何损失，国金证券不承担相关法律责任。

若国金证券以外的任何机构或个人发送本报告，则由该机构或个人为此发送行为承担全部责任。本报告不构成国金证券向发送本报告机构或个人的收件人提供投资建议，国金证券不为此承担任何责任。

此报告仅限于中国境内使用。国金证券版权所有，保留一切权利。

上海	北京	深圳
电话：021-80234211	电话：010-85950438	电话：0755-86695353
邮箱：researchsh@gjzq.com.cn	邮箱：researchbj@gjzq.com.cn	邮箱：researchsz@gjzq.com.cn
邮编：201204	邮编：100005	邮编：518000
地址：上海浦东新区芳甸路1088号 紫竹国际大厦5楼	地址：北京市东城区建国内大街26号 新闻大厦8层南侧	地址：深圳市福田区金田路2028号皇岗商务中心 18楼1806



【小程序】  
国金证券研究服务



【公众号】  
国金证券研究