



具身智能狂飙：技能“破圈”、资本“输血”和自我“造血”

文/曲忠芳

2025年，具身智能几乎贯穿全年“走红”。上半年，在春晚舞台上，宇树机器人扭秧歌转手绢，凭借灵动的动作技能成功“破圈”；半马跑道上，天工机器人完赛夺冠；之后的运动会与大小型活动现场，机器人成为不可或缺的参与者。下半年，资本竞速，国资、大厂、专业机构争相下场，生怕错过具身智能的“造富”洪流，热度持续到年底仍未消减；与此同时，冲刺上市、并购借壳、增发扩产，具身智能明星企业的野心也日益凸显。

这一年，被业界称为具身智能的“量产元年”，繁花盛开之下亦有暗流涌动，市场的终极版图如迷雾中的群峰，远未显现全貌，尚有待在2026年持续角逐与厮杀。在政策利好、资本竞速的趋势中，谁能最终握住那通往未来的“上桌门票”，不仅取决于眼下的技术身手，更取决于在残酷的商业大逃杀中，谁能率先完成从依赖资本“输血”到自我“造血”的跨越，显然这将成为2026年的精彩看点。

资本游戏：一级市场热捧，二级市场蓄力

截至2025年12月26日，IT桔子平台共收录了国内“具身智能”赛道发生的共317起股权融资事件。《中国经营报》记者基于对平台数据的分析对比观察到，2025年上半年具身智能赛道平均每月约15—20起融资事件，自7月出现第一个波峰（37起）后，到12月以45起达到全年最高峰，年底冲刺效应显著。从融资金额来看，银河通用机器人于12月19日完成了3亿美元的新一轮融资，斩获全年最大规模的单笔融资。

需要指出的是，多家具身智能企业在2025年获得了资本的不断加持，至少有14家企业发生了4次及以上的融资事件，例如松延动力共发生6起融资事件，累计融资金额约11.5亿元；众擎机器人、智平方、鹿明机器人各完成5起融资事件。与此同时，有11家具身智能企业在2025年的融资规模总额达到10亿元以上，银河通用、自变

量、它石智航、星动纪元、乐聚机器人的年内融资总额更是超过15亿元。这些数据表明，具身智能领域的融资密度极高，行业竞争节奏非常快。

从投资方来看，在具身智能赛道，国资基金及其关联企业出手频繁，其中以北京国管（北京国有资本运营管理有限公司）为典型代表，其在2025年共参与了14起投资事件，如星动纪元、银河通用、加速进化、灵巧手、星海图等，其投资高度集中在“清华系”及北京的头部人形机器人标的。在专业投资机构中，英诺天使基金是早期投资中最活跃的“搏手”，参与14起具身智能领域的种子轮与天使轮融资事件。除此之外，百度、阿里、联想、美团、小米、京东、腾讯、华为等科技大厂也成为投资队列中的重要力量，尤其是百度风投出手14次，参与智元机器人、星动纪元、加速进化、灵巧手、千寻智能等企业的股权融资。

值得注意的是，许多具身智能企业在作为被投标的谋求股权融资的同时，也开始以“产业投资

者”或“生态构建者”的身份，通过财务或战略投资，向下游核心零部件、算法或垂直应用企业输出资金与资源。如智元机器人在2025年参与了9起股权融资事件，银河通用、乐聚机器人分别出手了4次和3次。

一级市场火热异常，二级市场同样暗潮涌动。宇树科技、乐聚机器人、云深处等已陆续开始IPO进程。其中走得最快的宇树科技已在监管部门备案，于2025年11月完成了上市辅导阶段，预计不日将提交IPO申请。2025年12月23日，云深处宣布启动上市辅导，辅导阶段预计持续到2026年6月。随后，云深处完成了数亿元Pre-IPO轮融资。

除了走从股权融资到IPO上市的募路路径之外，在A股市场通过“协议转让+要约收购/增资控股母公司”的模式实现战略及财务目标，也成为许多具身智能企业的选择。最典型的是2025年7月，智元机器人通过资本运作成为上市公司上纬新材的控股股东；到2025年12月24日，港股上

市公司优必选拟斥资16.65亿元收购锋龙股份。

圈地“造血”：争抢订单，工业场景占主流

“具身智能的竞争会不断加剧，你追我赶，在这个过程中市场的挤出效应会不断增强，那些没有足够强融资造血能力的企业很容易就会掉队，甚至被淘汰出局。”谦询智库创始合伙人龚斌指出，人形机器人要跨越周期，必须在这个时间节点融到足够的钱，所谓“高筑墙、广积粮”，才能扛住未来五至十年的持久战。尤其值得警惕的是，在近几个月来，AI包括具身智能行业存在“泡沫”的声音甚嚣尘上，业界争论莫衷一是，而一旦2026年或者未来两三年内行业“泡沫”破灭，那么许多企业会面临生存危机。基于此，当前具身智能领域的资本游戏热闹异常，自然也在意料之中。

无论是一级市场的股权融资，还是二级市场的资本运作，对于具身智能企业来说，本质上都属于外部“输血”，而后续随着市场理性的

回归，如何实现自我“造血”、跑通商业化路径才是对具身智能企业的核心考验。

据记者对2025年13家公开披露订单的具身智能机器人企业的初步汇总整理，订单合同的金额总和超过80亿元。除了常见的科研教育、文旅商业场景外，大额订单往往流向工业场景，涵盖工厂制造、仓储物流等场景。其中，优必选在2025年12月24日给出的官方数据称，其工业人形机器人累计获得近14亿元订单，Walker S2在2025年11月开启量产交付，当月产能超过300台，预期全年交付500余台，到2026年预期全年产能将提升至10000台。智元机器人创始人、董事长、CEO邓泰华于2025年12月22日对外透露，该公司在2025年真实出货量超过5000台，销售额超过10亿元，预计从2026年起，机器人出货量与销售收入有望保持数倍增长。从此前披露的采购订单中，智元机器人的采购方包括中国移动、龙旗科技、富临精工、均普智能等。

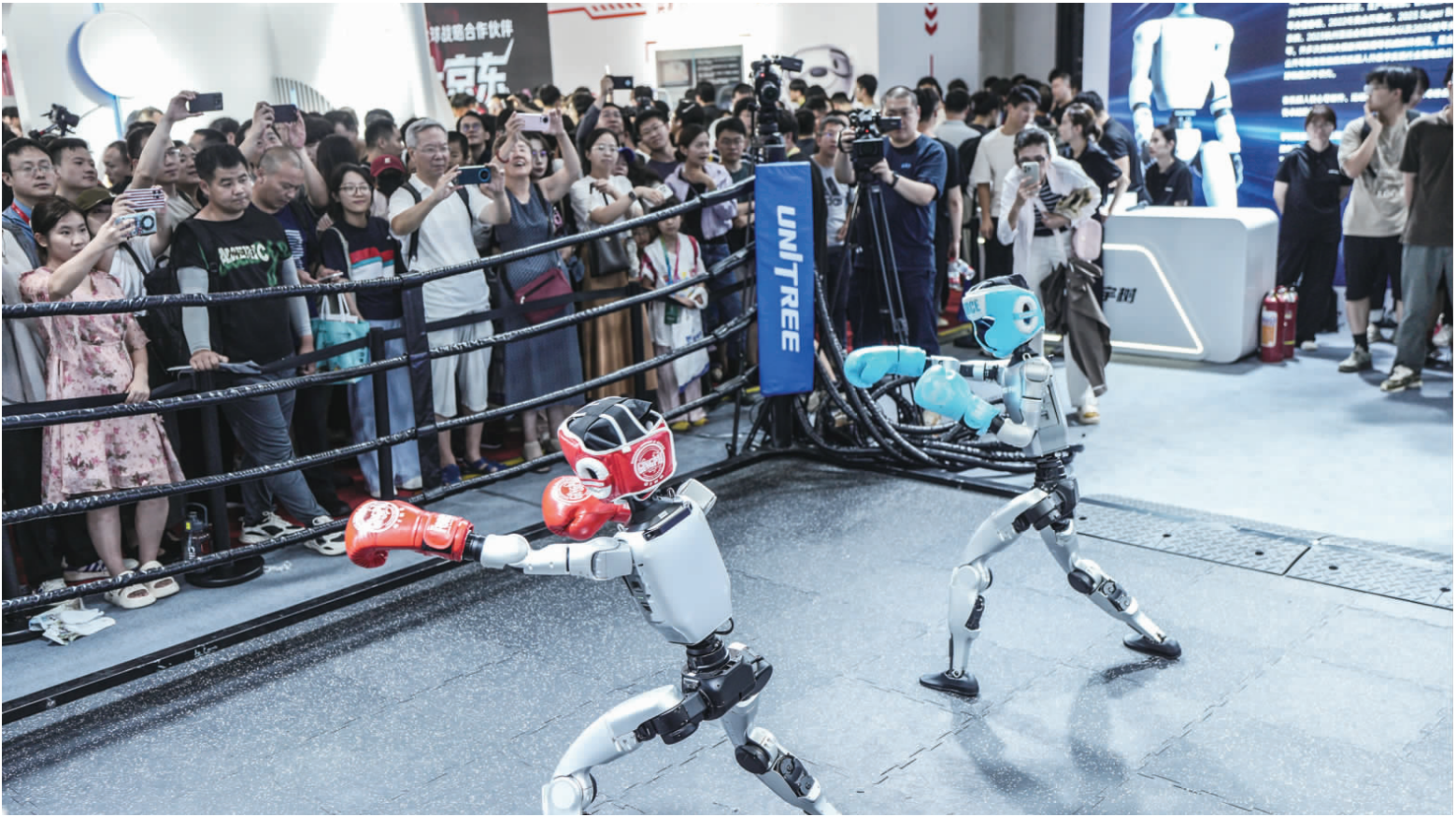
从这些商业订单中可以看到，具身智能企业都在推动技术落地

与规模化量产，争抢行业订单，一方面锚定教育、工业等场景切入行业应用，另一方面也在试图绑定行业龙头公司构建供应链与生态壁垒。企业的竞争逻辑从单一的产品技术比拼，演变为“技术研发+产能建设+生态合作”的综合比拼。而市场竞争升温带来的必然结果之一，便是供应链成本及产品服务价格的下降，这为具身智能市场的真正爆发奠定了基础条件，这从2025年人形机器人整机的单价从数十万元降至十万元以内甚至万元以内也可见一斑。

对于具身智能领域的火热订单潮，智源研究院院长王仲远在接受记者采访时指出，机器人出货量增长是好现象，因为硬件要靠打磨，只有当出货量达到一定规模，才能知道硬件存在什么样的问题，从而不断快速迭代，使机器人真正从“可用”进入“好用”阶段。与此同时，王仲远提醒道，希望行业的量产是由真实需求形成的量产，这样才是可持续的，而并不只是在政策和投资推动、市场尝鲜等因素汇聚下促成的量产。假设是非真实需求形成的量产，那它在经历一定热度后，很容易因为技术不够成熟、应用价值达不到需求、客户不再继续购买，出现市场冷却，进入所谓的“泡沫破裂期”或“低谷期”。

摩根士丹利在2025年12月发布的《机器人年鉴》报告显示，预测到2040年全球将有超过10亿台人形机器人投入运营，中国在制造人形机器人的竞赛中遥遥领先，过去五年里以7705项专利领先全球，在全球供应链中的占比超过六成；在全球100家企业名单中，56%的公司来自中国。从全球范围来看，当前整体市场仍处在“培养期”，实际规模尚有限，预计在2030年至2035年逐渐进入快速增长阶段。

毫无疑问，具身智能走过2025年，已从技术“秀场”真正踏入了产业“战场”。在这场刚刚起跑不久的“马拉松”耐力跑中，谁最终能稳住产线质量、跑通真实的商业闭环，最终拥有持续留在赛场的资格？2026年的具身智能赛道更值得期待与关注。



具身智能走过2025年，已从技术“秀场”真正踏入了产业“战场”。

视觉中国/图

视频大模型从“玩具”走向生产力工具

文/李静

过亿条的视频正在被AI大模型生成。2025年，视频生成大模型在技术飞速进步与商业化探索中，正经历着从“玩具”向“生产力工具”的关键转型。

“如果成本下不来，用户基本不会用。我们希望让更多人能够用起来。”百度副总裁陈一凡在谈及蒸汽机大模型的免费策略时曾如此表示。2025年下半年，百度通过推出限时免费的Turbo版本正式进军视频生成领域，犹如一条鲇鱼搅动了原本平静的市场。

2025年，中美视频生成大模型双轨竞争格局基本成型：OpenAI发布Sora2并巧妙融入社交功能，引发全球用户创作热潮；与此同时，国内快手的可灵、百度的蒸汽机、字节跳动的Seedance等产品在2025年密集迭代，行业在技术可控性、多模态融合与商业落地方面取得显著进展。

站在2025年年末回望，这一年既是视频生成大模型的技术突破年，也是市场竞争白热化的一年。展望2026年，行业或将迎来从“能生成”向“好生成、快生成、可交互生成”的深度演进。

技术跃进

2025年，视频生成大模型厂家相继推出多个版本，技术焦点从“能否生成”转向“如何更好地生成”。在国内厂商中，快手可灵AI

自2024年6月发布以来，在一年半时间内已完成超30次迭代升级，平均每半个月就有一次优化。阿里巴巴的通义万相视频大模型在2025年完成5次重要迭代，其12月推出的2.6版本新增角色扮演与分镜控制功能。字节跳动在2025年年末发布的Seedance1.5 pro，则重点优化了音视频生成效率与泛化能力。

华创证券在研报中指出，从模型层看，2025年多款主流模型视频生成效果朝着画面更高清、时长更长、物理更拟真等多方位进步；主体跨帧一致性、复杂语义理解等长期难题得到缓解。

Omdia高级首席分析师詹墨磊对记者指出，2025年视频生成大模型行业实现了两大关键技术突破：一是对物理规律的理解增强，如国外的Sora2和Runway Gen-4.5能模拟物体重量、碰撞反弹、液体流动与布料飘动；二是音画同步能力成熟，如阿里WAN、火山引擎Seedance1.5 Pro、百度蒸汽机模型，均支持端到端生成带口型匹配的声音视频，无须后期配音。

百度蒸汽机团队在2025年10月展示了实时交互与长视频生成能力。百度商业研发首席架构师李双龙透露道：“团队将窗口期的扩散生成转变为回归流式的扩展生成，解决了长视频生成的技术瓶颈。”

火山引擎总裁谭待在接受《中国经营报》记者采访时特别强调，

音画同步不仅是技术指标，更是“语义同步”的体现。“要做好这一点，不仅需要强大的视频模型，还需要顶尖的大语言模型和端到端语音模型支撑。”他透露，Seedance在中文及方言口型同步上已达到行业领先水平。

与此同时，生成时长与交互能力成为新战场。百度蒸汽机在10月宣布支持“逻辑上无限生成”的长视频，并实现“实时交互编辑”——用户可在生成过程中随时打断、修改、续写。李双龙解释，这是通过将扩散模型与自回归架构融合，兼顾质量与线性成本的结果。

快手可灵2.0则推出Multi-modal Visual Language (MVL)交互框架，允许用户结合图像、视频片段、文字甚至声音输入，实现多维度创意表达。

如果说2025年年初，视频大模型更多被视为一种“玩具”，那么随着技术持续迭代，越来越多的行业开始将其深度应用于实际场景。例如在电商领域，值得买科技CTO王云峰对记者指出，电商商家已经普遍在使用AI生成的短视频投流内容，视频模型支持从脚本、分镜到成片的全流程自动化，“而且部分AI生成的视频在投流效果方面已经超越人工制作水平”。

尽管如此，技术挑战依然存在。在复杂场景（如多人互动、精细物理交互）中，模型仍易出现人物形象漂移、动作畸变或流体不自

然等问题。詹墨磊直言：“目前的AI视频生成技术好比一个聪明但是不细心或者经验不够丰富的助手，最典型的例子是流体动力学问题。虽然Sora2和Runway Gen-4.5都声称改进了液体模拟，但实际生成的水花、油溅、烟雾还是带有‘人工合成’的痕迹，不够自然。”

技术的成熟直接推动了用户增长。截至2025年7月27日，可灵AI已在全球拥有超过4500万创作者，累计生成超2亿个视频和4亿张图片，服务超过2万家企业客户。

随着技术的进步，AI生成的视频内容正从“奢侈品”逐渐转化为“日用品”。“随着基础设施价格的定价趋于成熟，ComfyUI (AI绘图工具)、Diffusers（专注于提供预训练的扩散模型）等开源工具链逐渐完善，创作者可将AI无缝集成到After Effects（影视后期特效软件）中，使得工具链和视频生成技术融合达到一定的发展阶段，商业化落地也得到了进一步发展。”

快手财报显示，可灵AI在2025年前三季度累计营收超7亿元，全年预计达10亿元，成为国内首个披露明确收入规模的视频生成产品。其中，Prosumer（专业创作者）订阅贡献近70%收入，客户涵盖自媒体、广告营销从业者等兼具传播力与付费能力的群体。

然而，可灵AI的收入增速正在放缓。数据显示，可灵AI在2025年第一季度和第二季度的收入分

别为1.5亿元、2.5亿元，第三季度在3亿元以上，但按全年目标推算，第四季度预计仅完成约3亿元，环比持平，同比增速明显回落。这一变化与下半年的行业竞争加剧密切相关。一个“卷”字可以直观反映出国内视频生成赛道的激烈竞争态势。

2025年7月，百度正式入局，推出蒸汽机大模型，并采取“基础版限时免费+分层付费”策略。陈一凡坦言：“如果成本过高，用户很难接受。我们希望让更多人用起来。”此举显然直接冲击现有付费模式。此前，多数模型仅提供少量免费积分，深度使用需订阅。而百度将免费入口嵌入搜索与广告后台，形成“需求—生成—分发—反馈”闭环，迅速拉动调用量。

面对竞争，谭待则显得较为从容：“竞争本身并不关键。假设对手用户增长了三倍或五倍，但如果2026年整个市场能增长十倍呢？”

2026年趋势展望

展望2026年，视频生成大模型行业将从高速增长转向高质量发展，技术创新、场景深耕与商业模式升级将呈现新的趋势特征。

行业的共识是，通用模型难以满足高要求场景，专业化微调将成为主流。詹墨磊指出，电商、教育、医疗仿真等领域存在明确需求，但需通过RAG（检索增强生成）注入行业知识库，打造“开箱即用”的专业模型。百度已在短

剧、动漫广告等场景快速迭代运镜语言模型；快手则通过API向电商、广告主开放虚拟试穿、批量渲染等模块。

此外，“实时生成”将成为技术竞争新高地。百度蒸汽机团队已规划将音视频一体化能力扩展至长视频，并支持多人对话；火山引擎也在研发Draft功能，以降低样片制作成本。谭待强调：“to B商业应用必须‘真的有用’，而非‘玩一玩’。”这意味着低延迟、高一致性、可中断编辑的实时工作流将成为企业采购的关键考量。

从技术突破到商业探索，2025年的视频生成大模型行业呈现出高速发展与理性回归并存的特征。厂商们既在技术上不断突破边界，又在商业上寻求可持续模式。

陈一凡将视频生成的发展类比为通讯技术演进：“我觉得视频大模型有点类似于通讯行业的发展，现阶段基本是从2G往3G时代走，就是基本实现了可生成。再往下就是实时互动，一定是从‘可生成’走向‘可实时互动’。”

2026年，随着视频生成大模型在实时交互技术的成熟和行业解决方案的落地，视频生成大模型有望在更多领域展现其价值，推动整个行业向更加成熟、理性的方向发展。

招银国际在研报中预计，到2027年，全球创意应用市场规模有望达到546亿美元，其中AI创意应用渗透率有望达到11%。