

2026年中国AI视频生成模型行业概览： 人工智能，使好莱坞为之轰动 (精华版)

2026 China GenAI Video Model Industry
2026 年中国の映像生成AIモデル産業

概览标签：生成式AI模型、影视娱乐、知识产权

2026/03

报告提供的任何内容（包括但不限于数据、文字、图表、图像等）均系头豹研究院独有的高度机密性文件（在报告中另行标明出处者除外）。未经头豹研究院事先书面许可，任何人不得以任何方式擅自复制、再造、传播、出版、引用、改编、汇编本报告内容，若有违反上述约定的行为发生，头豹研究院保留采取法律措施，追究相关人员责任的权利。头豹研究院开展的所有商业活动均使用“头豹研究院”或“头豹”的商号、商标，头豹研究院无任何前述名称之外的其他分支机构，也未授权或聘用其他任何第三方代表头豹研究院开展商业活动。

研究目的&摘要

研究目的

本报告为2026年中国AI视频生成模型行业概览报告，将梳理中国AI视频生成模型行业的相关生产及发展情况，对该行业的产业链，竞争格局做出具体分析。

本报告关键问题：

1. 中国AI视频生成模型产业链上、中、下游由哪几部分构成？
2. 中国AI视频生成模型行业市场规模如何，未来行业将如何发展？
3. 中国AI视频生成模型行业的竞争情况如何？

摘要

- **行业产业链：**中国AI视频生成模型行业已形成上游算力基建、中游模型研发与工具、下游场景应用的完整闭环产业链：上游以AI芯片、GPU服务器、智算中心、数据服务为核心，为模型训练与推理提供硬件与数据支撑，并且上游的自主可控将直接决定下游模型工具定价、API收费和商业化空间；中游环节是产业核心，字节跳动Seedance系列、快手可灵AI系列与阿里通义万相等国产模型快速迭代，聚焦文生视频、图生视频、长视频生成与音画同步能力，同时配套视频编辑、转码优化、MaaS服务等工具层，形成技术与商业化双轮驱动；下游环节则广泛渗透影视短剧、电商广告、营销传媒、文旅教育等B端主流场景与C端创作需求，依托“降本增效、批量生产、快速迭代”的核心价值，推动AI视频工业化落地，全链条呈现巨头引领、创业公司跟进、国产替代加速、商业化快速兑现的发展格局。
- **市场规模：**2020年以来，中国AI视频生成模型市场保持快速增长，规模从约1.5万元攀升至12.1亿元，期间年均复合增长率高达965.0%，核心驱动力主要来自两大维度：技术层面，国产算力优化、模型量化与MaaS服务普及大幅降低推理成本，叠加工程化与产品化能力持续提升，推动行业从实验室科研阶段走向可商用、可规模化落地的成熟阶段，为市场从极低基数快速增长提供了核心技术支撑；需求层面，企业与专业创作者对降本增效、批量生产的需求持续释放，叠加内容平台生态赋能与C端大众创作需求觉醒，形成了多层次需求格局；未来，随着AI视频生成模型性能升级与行业渗透扩张，中国AI视频生成模型市场规模将由2026年的36.6亿元跃升至2030年的1,497.0亿元，年均复合增长率为152.9%。
- **竞争格局：**全球AI视频生成模型行业呈现出较高的市场集中度，头部厂商凭借领先的技术实力与完善的生态体系占据行业主导地位，并持续强化自身竞争优势；当前全球AI视频生成行业由中美共同主导，且随着字节跳动、快手等企业的持续发力与OpenAI相关业务的退出，中国企业有望凭借突出优势抢占全球行业领先地位。



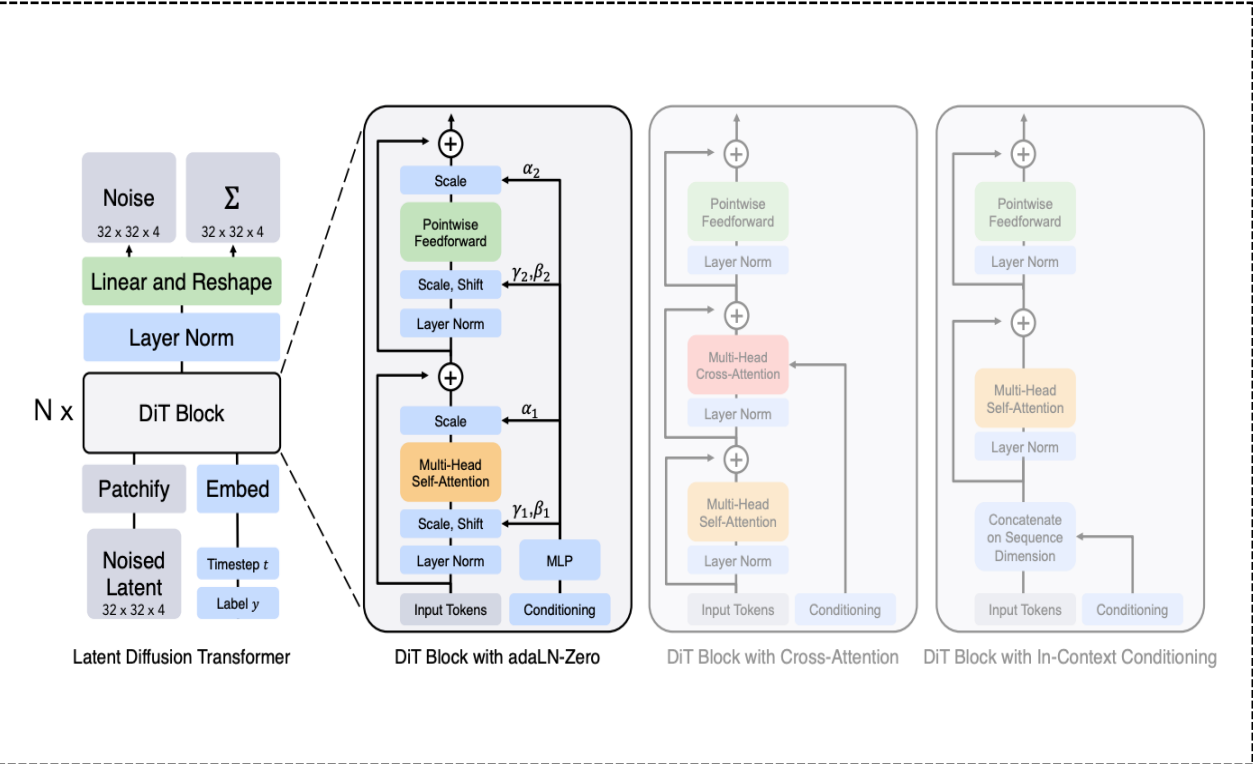
中国AI视频生成模型行业综述——Diffusion Transformer架构定义

Diffusion Transformer (DiT) 是Seedance、OpenAI Sora等前沿视频生成模型的核心架构，同时也是当前行业的主流架构，具备更强的长程依赖捕捉能力与扩展性

Diffusion Transformer (DiT) 架构介绍

➤ **Diffusion Transformers (DiT) 架构**属于基于隐空间扩散模型的生成方法，其核心价值在于证明了扩散模型的性能上限不依赖于U-Net的卷积归纳偏置，而是取决于通用架构的建模能力与计算资源投入。它将Transformer的长程依赖捕捉优势与扩散模型的生成稳定性完美结合，不仅推动了图像/视频生成质量的飞跃，更确立了“Transformer+扩散”的主流范式，为后续大规模生成模型的发展提供了关键架构参考。

Diffusion Transformer (DiT) 架构模型示意图



- ❑ 当前，Diffusion Transformers (DiT) 架构已成为Sora等前沿视频生成模型的核心技术架构，具有更强的长程依赖捕捉能力和扩展性。其核心设计思路在于将Transformer的全局注意力机制与扩散模型的加噪-去噪生成逻辑进行深度融合，以Transformer网络全面替代传统扩散模型所采用的U-Net骨干结构。
- ❑ DiT的前向加噪过程与传统扩散模型保持一致，通过确定性高斯噪声注入将清晰的图像或视频帧潜变量特征逐步转化为纯噪声，构建马尔可夫链无序状态；在反向去噪前，先借助Patchify模块将潜变量特征图分割为固定大小图像块并经线性变换转为一维Token序列，同时将时间步、类别标签等条件信息嵌入为额外Token并拼接至输入序列，再由多个DiT块构成的Transformer网络利用全局自注意力机制捕捉图像块间关联、动态预测噪声分布，最后通过线性解码器对处理后的Token序列进行反块化操作以重构噪声预测结果，完成单步去噪，重复该流程即可从纯噪声中生成清晰样本。
- ❑ 以 OpenAI的Sora模型为例，该模型首先将视频压缩至低维潜在空间，再将其表征分解为时空块以实现视频的分块处理，并采用DiT架构，在多个应用领域体现出优异的扩展性能；相较于传统视频生成模型，Sora不仅显著延长了生成视频的时长，还具备更强的自然语言理解能力与物理世界规律建模能力。

AI视频生成模型行业产业链分析

中国AI视频生成模型行业形成了以上游算力与数据为基础壁垒、中游模型研发与训练为核心枢纽、下游场景为商业变现主体的产业链，三者协同驱动行业从技术研发走向规模化、合规化与商业化落地

AI视频生成模型产业链



来源：企业官网、头豹研究院

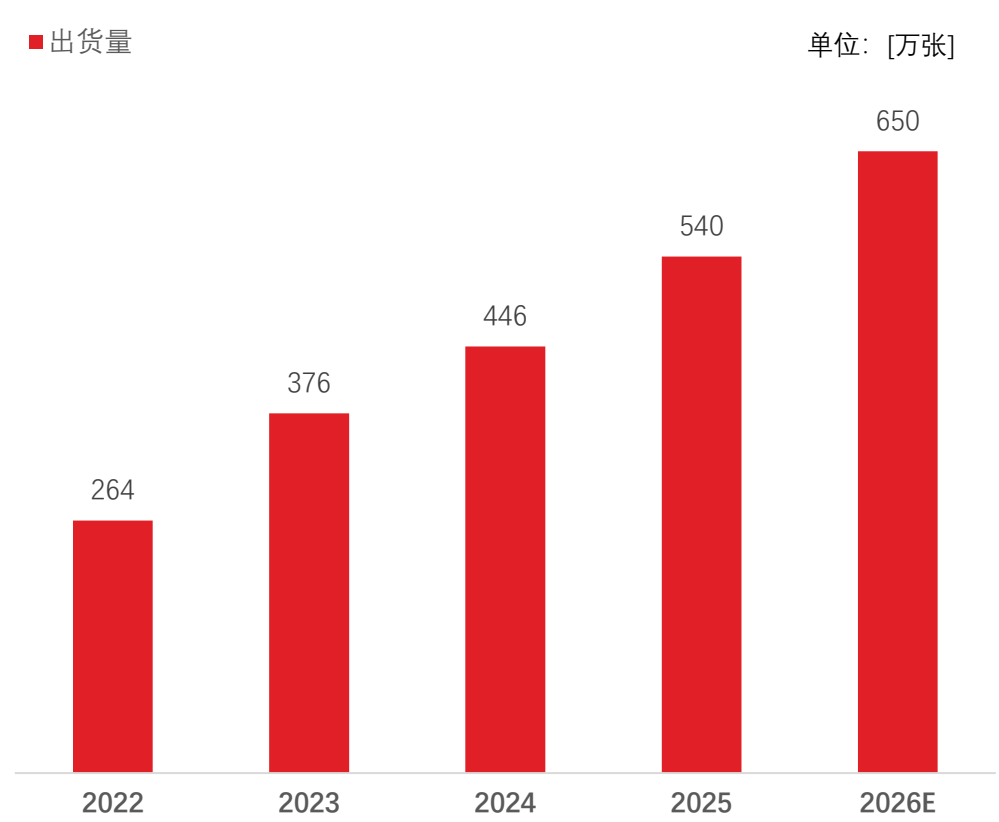
AI视频生成模型行业产业链——上游：算力芯片

英伟达市占率超过95%，在全球高性能GPU（图形处理器）市场中占据了主导地位，同时其芯片架构与系统的技术正在实现进一步的迭代升级

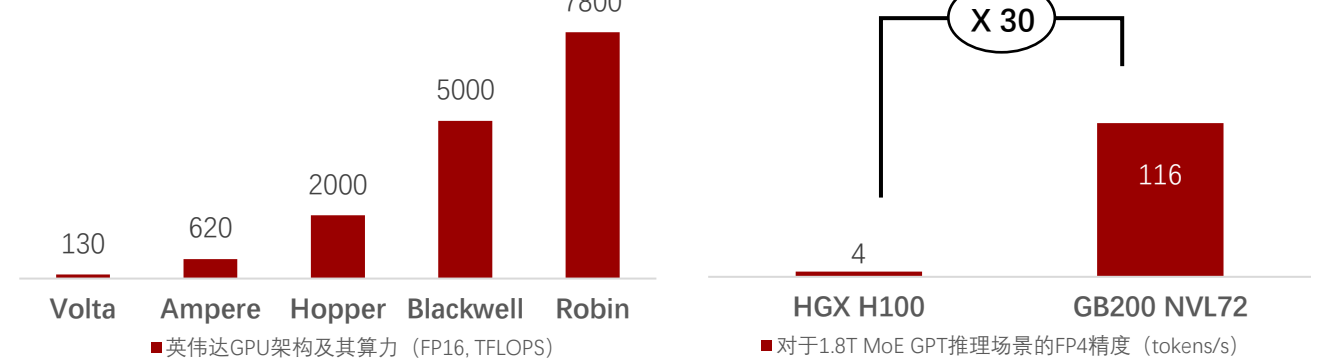
AI视频生成模型上游——算力芯片

算力芯片是专门为人工智能、高性能计算等场景设计，以高效处理大规模并行计算为核心目标的专用集成电路，也被称为AI芯片或计算加速芯片。它更侧重矩阵运算、浮点计算等AI与科学计算任务，通过专用架构（如GPU、NPU、TPU）实现对深度学习模型训练、推理及复杂科学模拟的算力支撑，是支撑大模型、生成式AI等现代算力需求的核心硬件底座。

英伟达人工智能GPU（图形处理器）出货量，2022-2026E



英伟达芯片层与系统层性能演化



- 当前，全球高性能GPU（图形处理器）市场呈现出英伟达公司一家独大的格局，其市占率已经超过了95%。2022-2025 年，英伟达人工智能GPU出货量已经从264万张快速增长至约540万张，期间内年均复合增长率高达26.9%。市场预计，2026年英伟达公司的人工智能GPU出货量将以约20.4%的增速进一步提升至650万张左右，同时伴随芯片架构迭代，持续增强芯片与系统的算力水平。另外，多家媒体透露，中国有关部门已批准多家中国企业从英伟达公司采购H200人工智能GPU产品，这是第一款能以每秒4.8TB的速度提供141GB HBM3e内存的GPU，容量几乎是英伟达H100 GPU的两倍，内存带宽则提升1.4倍。H200更大、更快的内存能加速生成式人工智能和LLM，同时提升高性能计算工作负载的科学运算。
- 芯片层方面，从Volta到下一代Robin架构，英伟达单GPU的FP16算力已经从130 TFLOPS跃升至约7,800 TFLOPS，实现了超60倍的指数级增长，奠定了AI大模型发展的硬件基础；系统层方面，对比表明在1.8T MoE GPT推理场景下，英伟达新一代GB200 NVL72系统相比HGX H100实现了约30倍的性能提升（从4 tokens/秒提升至116 tokens/秒），印证了芯片架构升级与系统级整合优化能将底层算力转化为实际业务效率的巨大突破。

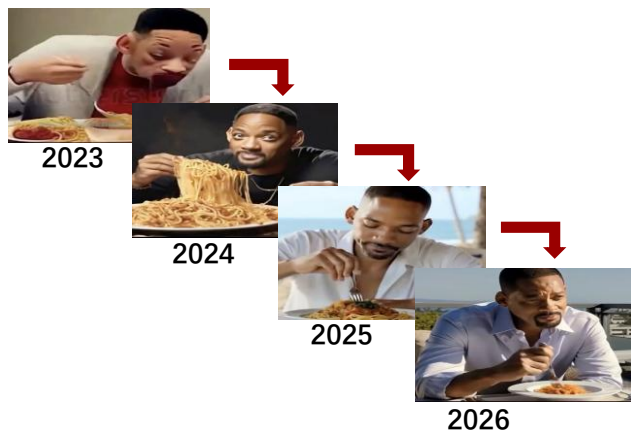
来源：英伟达、路透社、头豹研究院

AI视频生成模型行业产业链——中游：模型研发与训练

从相关测试来看，国产模型已实现全球领跑并重塑行业研发格局；同时，国产模型具备显著性价比优势，未来将朝着实时生成的方向全面迈向成熟商业化阶段

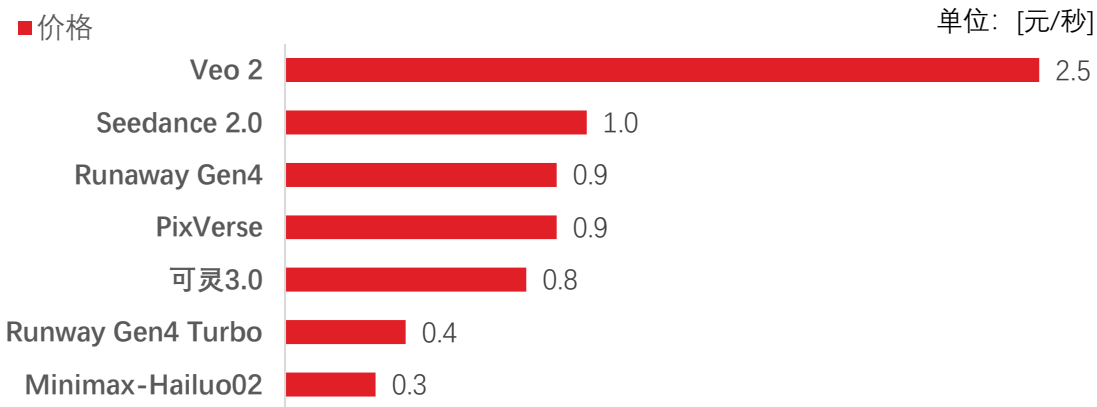
AI视频生成模型中游——模型研发与训练

AI视频生成模型“Will Smith Eating Spaghetti Test”非正式基准测试进程，2023-2026



- “Will Smith Eating Spaghetti Test”作为人工智能行业的非官方评测标准，主要用于衡量生成式视频模型还原真实人体动作与面部表情的水平，其源自2023年网络上流传的一段威尔·史密斯吃面条的AI视频，因画面动作僵硬失真而成为评判AI视频能力与短板的重要参照；2025年，谷歌推出的视频生成工具Veo 3测试版在面部精度、动作流畅度及音画同步上实现明显优化，但仍存在面条音效异常等细节瑕疵；到2026年，技术得到显著突破，快手可灵与字节跳动Seedance 2.0成功成为全球仅有的两款基本通过该测试的模型，可精准复刻参考素材的镜头语言与运镜方式，使用户生成高度逼真、足以乱真的影视级场景内容。
- 可灵与Seedance 2.0的成功，意味着全球AI视频生成模型行业的中游研发与训练环节迎来了技术成熟度的关键分水岭，推动全球厂商从解决基础动作僵硬问题转向优化音画同步、细节质感等高阶能力；而2026年中国两款模型成为全球仅有的通过者，标志着全球AI视频研发格局发生重大转变，中游技术主导权由海外转向中国，同时也确立了影视级逼真生成成为行业下一代核心技术竞争方向，倒逼全球厂商在动作建模、多模态对齐、镜头语言理解等训练环节加速技术追赶与创新。

AI视频生成模型API单秒价格，2026年



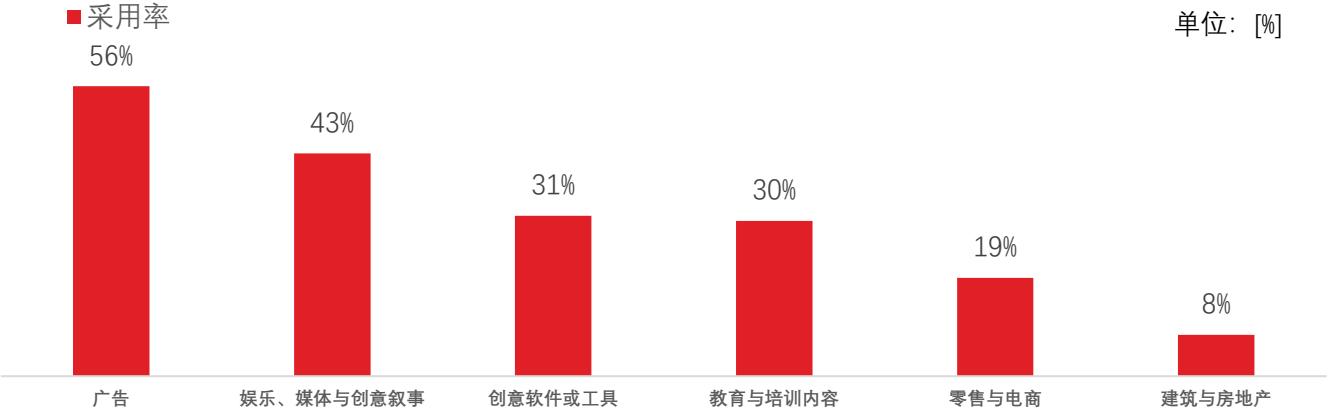
- ❑ 目前，全球主流AI视频生成模型普遍采用积分计费模式，且大多支持API接口调用，从定价水平来看，国内外主流模型的API单秒计费区间集中在0.2-1元/秒，其中定价较高的Google Veo 2单秒成本约0.4美元（折合人民币约2.5元）；相较之下，国产AI视频模型具备显著的价格优势，例如技术实力处于第一梯队的可灵3.0，单秒定价约0.8元，而火山引擎旗下技术领先的Seedance 2.0已正式公布定价标准，带视频输入的生成服务为28元/百万tokens，纯视频生成服务为46元/百万tokens，折算后单秒成本约1.0元，价格水平显著低于Google Veo 2等同期海外竞品。
- ❑ 市场预测，2027至-2029年AI视频生成模型将实现关键跃升，依托自然语言等交互方式可完成对内容的精准控制，深度理解物理世界运行规律，其生成效果的稳定性与丰富度将在各应用场景全面满足商业化要求，1分钟视频的生产时长可压缩至数秒级别，同时平均成本有望降至0.5元/秒以下。

AI视频生成模型行业产业链——下游：模型应用

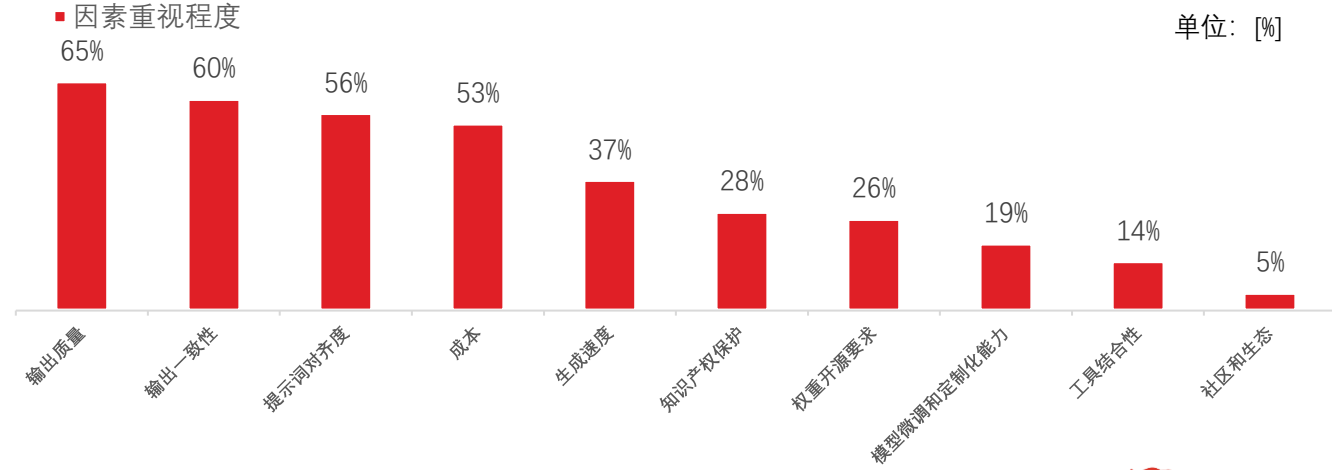
AI视频生成技术当前已在广告、媒体娱乐等多行业开展应用，呈现出“效率优先”的渗透格局；另外，下游用户更加聚焦输出质量、一致性、指令遵循度等核心指标，整体行业已进入务实落地阶段

AI视频生成模型下游——模型应用

不同垂直行业对AI图像和AI视频生成的采用率，2025



用户AI视频生成模型选择标准，2025



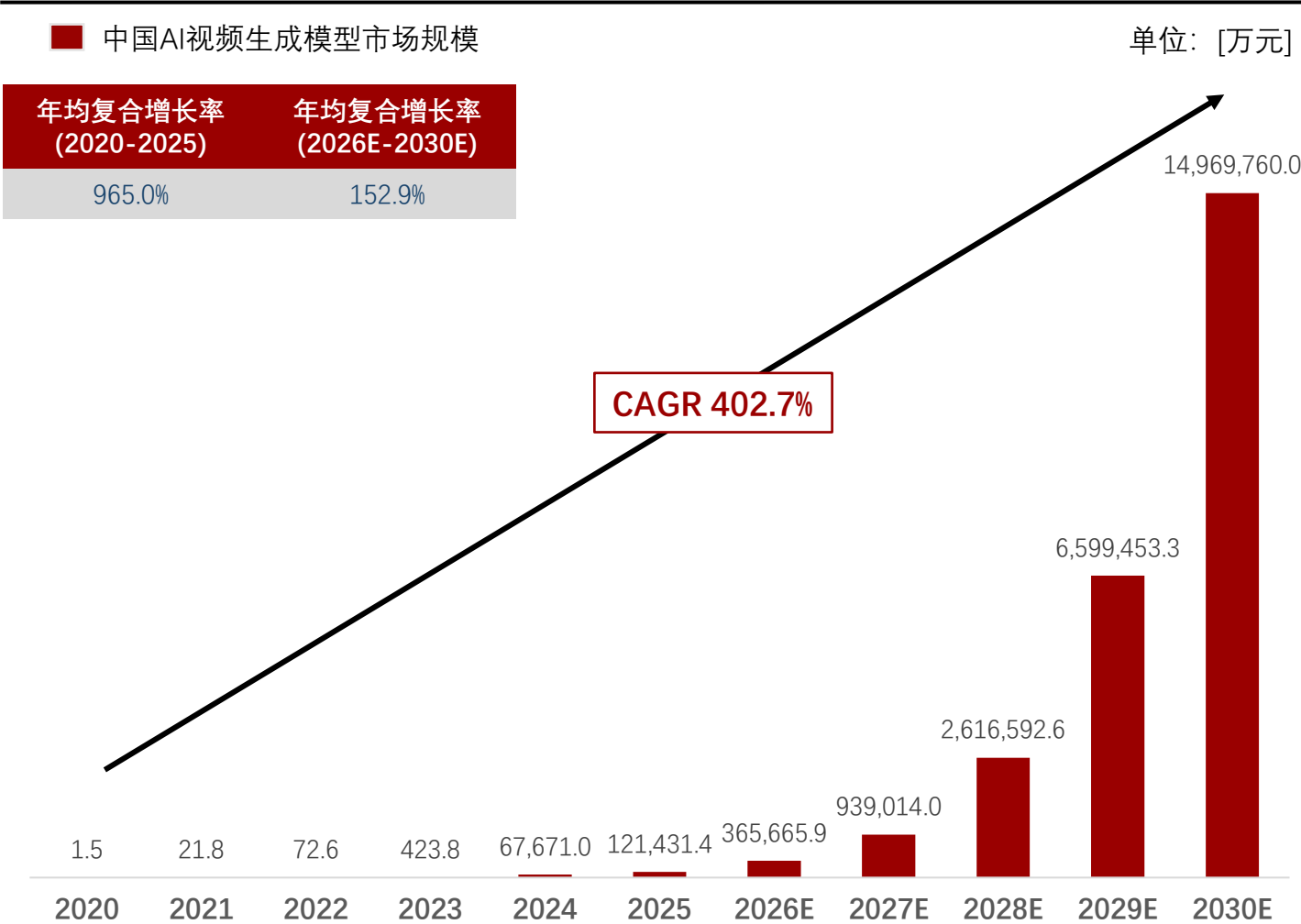
- 目前，AI视频生成模型主要应用于广告、娱乐、媒体、创意、教育、零售与建筑等相关行业或领域当中。具体来看，广告行业以56%的采用率位居首位，主要用于批量快速制作营销视觉素材、横幅广告与社交平台图文；娱乐、媒体与创意叙事领域以43%紧随其后，聚焦分镜制作、视觉预演、特效制作及短视频推广片段；创意软件或工具、教育与培训内容的采用率分别为31%和30%，对应设计平台、视频编辑工具及互动学习视频等方向；零售与电商、建筑与房地产的采用率相对较低，分别为19%和8%，应用场景集中在自动化商品摄影、虚拟试穿样机及3D渲染、项目概念图等专业领域。
- 从数据来看，AI视频生成技术的行业渗透呈现明显的“效率优先”特征。广告与内容创作类场景因对规模化、高效率生产的需求迫切，成为当前商业化最成熟的核心阵地，而零售电商、建筑地产等垂直场景因对精度、真实感要求更高，仍处于早期渗透阶段。这一分布既反映了技术落地的规律，即先解决高频、标准化的效率提升需求，再逐步向高定制化、高精度的专业场景延伸，也预示着未来随着模型物理仿真能力与成本优化，教育、电商等潜力场景将快速增长，成为行业新的增长极。
- 对于下游用户而言，输出质量、输出一致性、提示词对齐度和成本是用户最核心的四大决策依据，分别占用户总量的65%、60%、56%与53%，构成模型商业化的核心竞争力；生成速度、知识产权保护、权重开源要求等效率与可控性指标紧随其后，重视程度分别达到了37%、28%与25%；而模型微调定制化能力、工具结合性及社区生态目前影响相对有限，重视程度均在20%以下。这一分布直观反映出用户已从技术尝鲜转向务实的商业化落地需求，将稳定交付高质量、符合预期且成本可控的内容作为首要考量。

来源：《State of Generative Media Survey Report 2025》、头豹研究院

中国AI视频生成模型市场规模

2025年中国AI视频生成模型行业市场规模为12.1亿元，同比增长77.9%；预计2030年中国AI视频生成模型行业市场规模将达到1,497.0亿元，复合年增长率为402.7%

中国AI视频生成模型行业市场规模，2020-2030E



- 2020年至2025年，中国AI视频生成模型行业市场规模呈现爆发式增长态势，市场规模从约1.5万元攀升至12.1亿元，期间年均复合增长率高达965.0%。这一跨越式增长的核心驱动力主要来自两大维度：技术层面，AI视频生成模型架构从GAN、VAE逐步迭代至DiT扩散Transformer架构，实现了生成时长、分辨率与内容一致性的显著突破，同时国产算力优化（国产AI芯片通过低精度量化、算力深度优化与高算力密度设计，显著提升了视频模型推理效率；云厂商依托自研分布式架构与液冷智算集群，实现了多卡并行调度与存算网协同，大幅提升视频生成任务的并发处理能力）、模型量化与MaaS服务普及大幅降低推理成本，叠加工程化与产品化能力持续提升，推动行业从实验室科研阶段走向可商用、可规模化落地的成熟阶段，为市场从极低基数快速增长提供了核心技术支撑；需求层面，国内广告营销、电商短视频、影视创作及教育等领域呈现出强烈的视频化生产刚需，企业与专业创作者对降本增效、批量生产的需求持续释放，叠加内容平台生态赋能与C端大众创作需求觉醒，形成了覆盖B端、专业用户与消费端的多层次需求格局，为行业市场规模快速攀升提供了坚实的应用落地基础。
- 未来，预计中国AI视频生成模型行业规模将由2026年的36.6亿元继续跃升至2030年的1,497.0亿元，期间年复合增长率为152.9%。供给侧层面，国产AI芯片与智算云依托低精度量化、分布式并行计算及液冷集群技术，持续降低模型推理成本，实现单卡承载中大型模型运行，同时云厂商提供的MaaS服务与弹性算力供给进一步降低了企业的应用门槛；需求侧方面，AI视频生成技术不断向电影制作、短剧生产及自媒体短视频创作等场景渗透，在内容供给快速扩张与行业监管日趋规范的双重驱动下，凭借剧本可视化、多镜头叙事、角色一致性生成及高效成片等核心能力，有力支撑行业市场规模实现高速增长。

来源：中国互联网络信息中心、JP Morgan、头豹研究院

全球智能科技创新奖评选 AIX Award

AI + X · 无限乘数 · 让智能科技赋能真实场景

我们旨在筛选最具竞争力与用户价值的终端创新，让真正的好产品被看见、被信赖。

聚焦 终端制造领域（消费智能科技核心载体）



个人移动智能终端

涵盖智能手机、AIPC、平板电脑、便携式计算设备等核心移动载体



智能穿戴终端

聚焦智能手表、手环、AR/VR眼镜、智能服饰等可穿戴式设备



智能家居家电终端智能

覆盖智能音箱、门锁、安防监控、白/黑电、环境控制等智能联动硬件



智能文娱与教育陪伴终端

包含游戏主机、AI学习机、家庭影院、智能投影等沉浸式内容交互设备



智能办公与外设配件终端

包含智能办公本、键盘、鼠标、显示器、远程会议终端等生产力工具



智能出行与外场终端

涵盖消费级无人机、运动相机、智能出行等出行相关硬件



消费级/商用场景智能机器人

包含扫地机器人、商用服务机器人及工业协作机器人等

核心价值



权威公信力背书

依托头豹全球产业研究积淀，为企业第三方验证；入选全球创新案例库，助力企业拓展国际视野，加速全球化布局



聚焦真实用户价值

评审侧重应用场景适配度与用户真实体验，解决技术强但难感知的市场痛点



长效荣誉资产沉淀

专属奖杯及全球标识授权，构建品牌护城河，强化市场认知

关键节点 · 申报流程

01. 申报节点

2026年6月10日 全面开放申请通道

02. 评审期

7月11日-25日 多维度交叉评估

🏆 全球颁奖典礼

2026年8月4日 · 上海年度盛典

联系方式

陈夏琳Sharlin.chen@leadleo.com



方法论

- ◆ 头豹研究院布局中国市场，深入研究19大行业，532个垂直行业的市场变化，已经积累了近100万行业研究样本，完成近10,000多个独立的研究咨询项目。
- ◆ 研究院依托中国活跃的经济环境，研究内容覆盖整个行业的发展周期，伴随着行业中企业的创立，发展，扩张，到企业走向上市及上市后的成熟期，研究院的各行业研究员探索和评估行业中多变的产业模式，企业的商业模式和运营模式，以专业的视野解读行业的沿革。
- ◆ 研究院融合传统与新型的研究方法，采用自主研发的算法，结合行业交叉的大数据，以多元化的调研方法，挖掘定量数据背后的逻辑，分析定性内容背后的观点，客观和真实地阐述行业的现状，前瞻性地预测行业未来的发展趋势，在研究院的每一份研究报告中，完整地呈现行业的过去，现在和未来。
- ◆ 研究院密切关注行业发展最新动向，报告内容及数据会随着行业发展、技术革新、竞争格局变化、政策法规颁布、市场调研深入，保持不断更新与优化。
- ◆ 研究院秉承匠心研究，砥砺前行的宗旨，从战略的角度分析行业，从执行的层面阅读行业，为每一个行业的报告阅读者提供值得品鉴的研究报告。

法律声明

- ◆ 本报告著作权归头豹所有，未经书面许可，任何机构或个人不得以任何形式翻版、复刻、发表或引用。若征得头豹同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“头豹研究院”，且不得对本报告进行任何有悖原意的引用、删节或修改。
- ◆ 本报告分析师具有专业研究能力，保证报告数据均来自合法合规渠道，观点产出及数据分析基于分析师对行业的客观理解，本报告不受任何第三方授意或影响。
- ◆ 本报告所涉及的观点或信息仅供参考，不构成任何证券或基金投资建议。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告或证券研究报告。在法律许可的情况下，头豹可能会为报告中提及的企业提供或争取提供投融资或咨询等相关服务。
- ◆ 本报告的部分信息来源于公开资料，头豹对该等信息的准确性、完整性或可靠性不做任何保证。本报告所载的资料、意见及推测仅反映头豹于发布本报告当日的判断，过往报告中的描述不应作为日后的表现依据。在不同时期，头豹可发出与本报告所载资料、意见及推测不一致的报告或文章。头豹均不保证本报告所含信息保持在最新状态。同时，头豹对本报告所含信息可在不发出通知的情形下做出修改，读者应当自行关注相应的更新或修改。任何机构或个人应对其利用本报告的数据、分析、研究、部分或者全部内容所进行的一切活动负责并承担该等活动所导致的任何损失或伤害。

头豹业务合作

全球视野 · 本土洞察 · 研究数据 · 可信知识网络



行业数据API

开放原创报告与研究数据接口，
支持企业知识库、系统平台及AI
应用高效接入和调用



KNIT解决方案

构建企业可信内容体系，提升品
牌在AI搜索与问答中的可见度、
准确性与转化效果



报告会员账号

可阅读全部原创报告和百万数据，
提供PC及移动端，方便触达平台
内容



定制报告/白皮书

对产业及细分行业进行现状
梳理和趋势洞察，输出全局
观深度研究报告



商业尽调

面向投资并购和商业决策，
评估标的公司的商业前景、
价值及风险



招股书引用

研究覆盖国民经济19+核心产
业，内容可授权引用至上市
文件、年报



报告作者

- 陈夏琳 | 首席分析师
- 梁霄同 | 行业分析师



service@leadleo.com



业务咨询

- 客服电话：400-072-5588
- 官方网站：www.leadleo.com



深圳办公室

广东省深圳市南山区粤海街道华润
置地大厦E座4105室
邮编：518057



上海办公室

上海市静安区南京西路1717号会
德丰国际广场2701室
邮编：200040



南京办公室

江苏省南京市栖霞区经济开发区兴
智科技园B栋401
邮编：210046