

分析师：乔琪
登记编码：S0730520090001
qiaoqi@ccnew.com 021-50586985

OpenAI 发布 Sora，AI 视频技术巨大突破

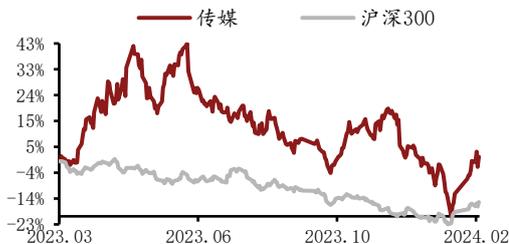
——传媒行业分析报告

证券研究报告-行业分析报告

强于大市(维持)

传媒相对沪深 300 指数表现

发布日期：2024 年 03 月 01 日



资料来源：中原证券，聚源

相关报告

《传媒行业专题研究：春节档数据创同期最高，内容端呈高集中度特征》 2024-02-21

《传媒行业月报：游戏版号再破百，关注春节档表现》 2024-02-07

《传媒行业月报：电影市场 550 亿元票房收官，游戏市场规模首超 3000 亿元》
2024-01-15

联系人：马崑琦

电话：021-50586973

地址：上海浦东新区世纪大道 1788 号 16 楼

邮编：200122

投资要点：

- **重磅 AI 视频模型发布。**2024 年 2 月 15 日，OpenAI 发布重磅产品 AI 视频模型 Sora，引发市场强烈关注。根据 OpenAI 官网介绍，Sora 能够根据用户输入的指令生成最高长达 1 分钟、清晰度最高可至 1080p 的高质量视频，生成的视频内容能够包含多个角色、特定类型的运动以及精确的主题和背景细节等。Sora 不仅能够充分理解用户的提示词 (Prompt)，还能够理解提示词中的物体在物理世界中的存在方式；同时 Sora 还能够在生成的单个视频中实现运镜、镜头切换等更加复杂的内容呈现并较为准确地保留角色和视觉风格，实现较高的一致性。
- **Sora 能够实现多种视频功能。**根据 OpenAI 官网的演示视频，目前 Sora 能够实现由 AI 完成的文生视频、图生视频、视频拓展、视频编辑、视频拼接、图像生成等功能。
- **技术进步有望大幅拓宽 AI 视频应用场景。**相比此前 AI 生成的视频，Sora 生成的视频不论是长度、运镜、镜头切换能力还是生成画面的细节、一致性、稳定性以及物理特性等方面都有了明显的突破，AI 生成视频即将从探索性阶段向实用性阶段逐渐过渡，随着这一过程的演变其应用场景也将得到大幅拓宽。
- **传媒产业有望受益显著。**从内容生产的角度来看，Sora 已经展现出了比较强的多模态内容生产能力，未来以 Sora 为代表的 AI 视频工具在游戏、广告营销、影视等多个内容产品领域都具有广泛的应用空间。一方面通过 AI 工具提升内容生产的效率，有效减少视频拍摄和制作环节所花费的时间和成本，缩短内容创作周期，未来的内容产业中，部分非核心内容的创作需求可能被 AI 模型替代；另一方面借助 AI 工具的力量能够降低内容创作的门槛，使更多的创作者参与到内容创作之中，提升内容产品的丰富性和创意性，同时非专业团队或内容创作者通过使用 AI 模型也能够创作出媲美专业团队的高质量内容产品，提升内容产品的质量。

建议关注：恺英网络、三七互娱、完美世界、吉比特、芒果超媒、风语筑、光线传媒、中国电影、万达电影

风险提示：AI 生成内容存在伦理、道德和法律风险；AI 工具发展进展不及预期

内容目录

1. 重磅 AI 视频模型发布，取得突破性进展	3
1.1. AI 视频模型 Sora 重磅发布	3
1.2. Sora 拥有多种视频生成功能	3
1.3. Sora 实现突破性进步	6
1.4. Sora 以 patch 作为基本训练单元，采用 diffusion + transformer 架构	7
1.5. 部分技术仍有瑕疵，静待技术再次进步	8
2. AI 技术进步扩大应用场景，传媒产业受益显著	8
3. 风险提示	9

图表目录

图 1: Sora 能够基于文字指令生成视频	4
图 2: Sora 能够结合左侧的图片和文字指令生成右侧的视频	4
图 3: Sora 视频拓展功能（三个不同的视频开头引导向同一个视频结尾）	5
图 4: Sora 能够通过添加指定元素或改变视频风格实现视频编辑	5
图 5: Sora 视频拼接功能（将左侧视频与右侧视频拼接并生成中间的视频）	5
图 6: Sora 图片生成功能	6
图 7: Sora 通过视觉编码器将视觉数据维度降低	7
图 8: Sora 通过去除噪声的方式生成清晰的视频	7
图 9: 男子跑步方向和跑步机相反	8
图 10: 篮球没有触碰篮筐而是直接穿过	8
表 1: AI 视频模型对比	6

1. 重磅 AI 视频模型发布，取得突破性进展

1.1. AI 视频模型 Sora 重磅发布

2024 年 2 月 15 日，OpenAI 发布重磅产品 AI 视频模型 Sora，引发市场强烈关注。根据 OpenAI 官网介绍，Sora 能够根据用户输入的指令生成最高长达 1 分钟、清晰度最高可至 1080p 的高质量视频，生成的视频内容能够包含多个角色、特定类型的运动以及精确的主题和背景细节等。Sora 不仅能够充分理解用户的提示词（Prompt），还能够理解提示词中的物体在物理世界中的存在方式；同时 Sora 还能够在生成的单个视频中实现运镜、镜头切换等更加复杂的内容呈现并较为准确地保留角色和视觉风格，实现较高的一致性。

目前 Sora 尚未对大众用户开放使用，但从 OpenAI 官网展示的视频来看，Sora 生成的视频已经接近实拍的效果，具有丰富的光影色彩、细腻的画面以及逼真的人物和物体造型，视频展示的物体和人物的运动轨迹也基本符合现实世界中的物理规律

1.2. Sora 拥有多种视频生成功能

根据 OpenAI 官网的演示视频，目前 Sora 能够实现由 AI 完成的文生视频、图生视频、视频拓展、视频编辑、视频拼接、图像生成等功能。

文生视频

在 OpenAI 的官网中展示了多个由 Sora 基于提示词生成视频的范例，例如根据提示词“一位时尚女性走在充满温暖霓虹灯和生动城市标志的东京街头。她穿着黑色的皮夹克、红色长裙、黑色靴子并拿着黑色皮包；戴着太阳镜、涂着口红；走路时自信又随意。街道潮湿且有反光，并与彩灯形成了镜面效应。周围还有一些行人。”生成了一段约 60s 的视频。从最终效果来看，视频基本能够包含提示词中所要求的全部要素。在画面移动的过程中人物与镜头的距离、背景中街道的场景以及地面积水处的反光也会随之不断变化，人物的动作姿态、背景元素以及画面风格都能够保持比较好的连贯性和一致性；同时镜头画面切换至脸部特写时，墨镜的反光、人物的表情等细节方面也得到了比较好的处理。

图 1: Sora 能够基于文字指令生成视频



资料来源: OpenAI 官网, 中原证券

图生视频

Sora 能够根据给定的图片和输入的提示词, 将静态图片转化为动态视频。

图 2: Sora 能够结合左侧的图片和文字指令生成右侧的视频



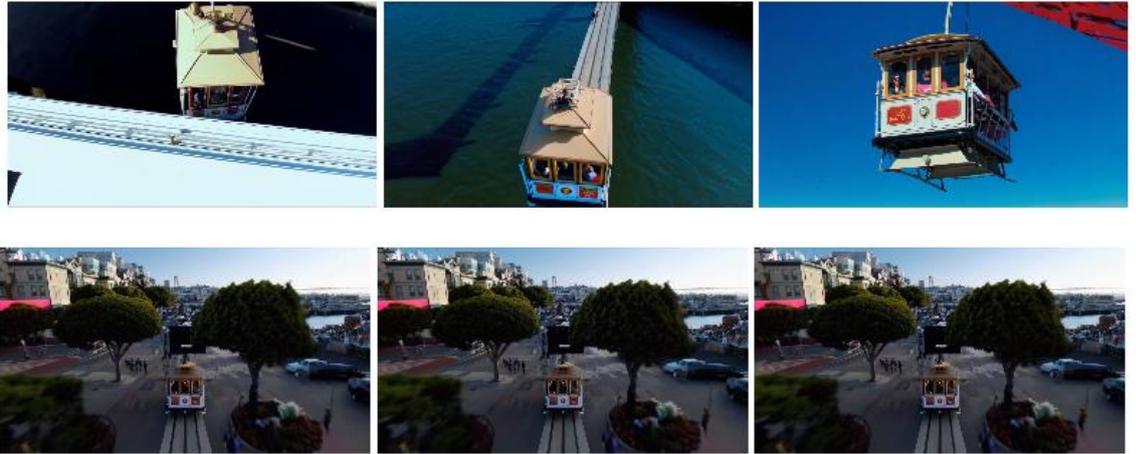
In an ornate, historical hall, a massive tidal wave peaks and begins to crash. Two surfers, seizing the moment, skillfully navigate the face of the wave.

资料来源: OpenAI 官网, 中原证券

视频拓展

根据给定的视频, 由 AI 模型根据时间线对该视频向前或向后的内容进行拓展补充, 此外也可以通过同时向前和向后拓展生成一个无限循环视频。OpenAI 在演示中展示了 3 个结尾完全相同的视频, 但由 Sora 补充的向前拓展的视频内容却完全不同。

图 3: Sora 视频拓展功能 (三个不同的视频开头引导向同一个视频结尾)



资料来源: OpenAI 官网, 中原证券

视频编辑

通过零拍摄的方式将输入视频的风格和环境按照需求进行要素的添加或风格的转换并输出新视频。

图 4: Sora 能够通过添加指定元素或改变视频风格实现视频编辑



资料来源: OpenAI 官网, 中原证券

视频拼接

将两段主体和场景完全不同的视频进行拼接或无缝转场, 生成一个融合两个不同视频要素的新视频。

图 5: Sora 视频拼接功能 (将左侧视频与右侧视频拼接并生成中间的视频)



资料来源: OpenAI 官网, 中原证券

图像生成

通过在一帧的空间网格中加入一块块高斯噪声来实现图像生成功能, 生成的图像可以拥有

不同的尺寸，最高可达到 2048*2048 的分辨率。

图 6: Sora 图片生成功能



Close-up portrait shot of a woman in autumn, extreme detail, shallow depth of field

Vibrant coral reef teeming with colorful fish and sea creatures

资料来源: OpenAI 官网, 中原证券

1.3. Sora 实现突破性进步

相比于此前市场上的 AI 视频模型, Sora 的进步十分明显。目前市场上主要的 AI 视频模型包括 Pika、Gen-2、Lumiere、Stable Video Diffusion 等。根据 OpenAI 官网的技术文档以及展示的视频效果, 相比其他 AI 视频模型, Sora 在生成视频的长度、视频尺寸、拓展性、多镜头能力、运镜以及视频的一致性、稳定性等方面都展现出了比较显著的优势。

具体来看: (1) Sora 生成的视频长度最高可达 60s, 而市场上的主流竞品模型最高支持到十几秒的视频生成; (2) Sora 在拓展视频时可以实现向前拓展、向后拓展和双向拓展的多种方式, 竞品模型通常仅支持向后拓展; (3) Sora 能够生成横屏 1920*1080 以及竖屏 1080*1920 之间任意尺寸视频, 还能够直接根据不同设备的原始屏幕高宽比生成相应视频内容, 而竞品模型通常仅支持固定比例尺寸视频, 同时 Sora 能够在生成完整分辨率视频之前快速生成同一模型的小尺寸内容原型, 进一步提高视频生成的效率; (4) 经过大规模训练后 Sora 涌现出了新的模拟能力, 仅通过规模效应而非归纳偏好的情况下使 Sora 也能够模拟来自于物理世界的人物、动物和环境的某些方面; (5) Sora 展现出了比较高的 3D 一致性, 在镜头运动过程中, 人物和场景元素也会在三维空间一致移动; (6) Sora 具有较好的长时间的连贯性和物体的持久性, 当人物或物体被遮挡或在同一个视频中即使生成同一个角色或物体的多个镜头也能够保持其外观的一致。

依靠这些优势, 相比于市面上的其他 AI 视频模型, Sora 能够生成更加复杂的视频内容, AI 视频模型的实用性也得到较大程度的提升。

表 1: AI 视频模型对比

模型	公司	最高视频时长	视频拓展性	视频尺寸
Sora	OpenAI	60s	向前/向后或双向同时拓展	1920*1080 至 1080*1920 之间任意尺寸

Pika 1.0	Pika Labs	3s,可拓展至 7s	可向后拓展	16:9、9:16、1:1、5:2、4:5、4:3
Lumiere	Google	5s	-	1024*1024
Stable Video Diffusion	Stability.AI	2-5s	-	576*1024
Gen-2	Runway	4s,升级会员后可拓展至 16s	可向后拓展	16:9、9:16、1:1、4:3、3:4、21:9

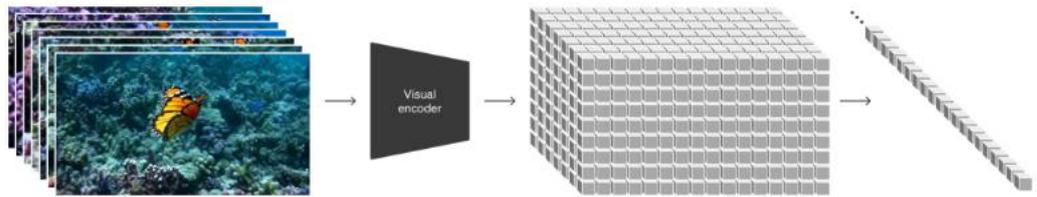
资料来源：Pikalabs 官网，OpenAI 官网，Lumiere 官网，Runway 官网，Stability.AI 官网，中原证券

1.4. Sora 以 patch 作为基本训练单元，采用 diffusion + transformer 架构

大语言模型通过大规模数据的训练获得了通用能力，其成功部分得益于通过 token 统一了代码、数学以及不同的自然语言等文本。在视频内容的生成方面，OpenAI 从大语言模型的训练中受到启发并沿用这一思路，将不同的视频和图像数据经过压缩后转化为 patch，patch 类似于大语言模型中的 token，能够作为视觉数据的基本单元。

OpenAI 训练了一个降低视觉数据维度的网络，将未经处理的原始视频输入，并输出经过时间和空间两个维度上压缩后的隐性表示，Sora 在压缩后的隐性空间中进行训练并生成视频，同时 OpenAI 还训练了一个解码器模型并将生成的隐性表示映射回像素空间。

图 7: Sora 通过视觉编码器将视觉数据维度降低



资料来源：OpenAI 官网，中原证券

根据 Sora 的技术报告显示，Sora 是一种扩散模型（diffusion model），从生成类似静态噪声的视频开始并通过逐步去噪的方式来最终形成清晰的视频。同时 Sora 还采用了 Transformer 架构，Transformer 架构已经在语言模型、计算机视觉和图片生成等不同领域展现出了显著的扩展性；在视频模型领域，diffusion transformer 也同样能够实现有效扩展。

图 8: Sora 通过去除噪声的方式生成清晰的视频



资料来源：OpenAI 官网，中原证券

为了提高 Sora 在文生视频方面的语言理解能力，OpenAI 基于过去对 DALL·E 和 GPT 模型的研究基础上，将 DALL·E 3 的重描述技术应用到视频领域。首先训练一个高度描述性的文字说明模型并用来为训练集中的所有视频生成文字描述，通过对高度描述性的视频文字的训练能够显著改善文本的保真度和视频的整体质量。

与 DALL·E 3 类似，OpenAI 还利用 GPT 将简短的用户提示词转化为更长的细节文字描述并发送给视频模型，使 Sora 能准确地根据用户的提示词生成高质量的视频。

1.5. 部分技术仍有瑕疵，静待技术再次进步

虽然相比以往的 AI 视频，目前 Sora 生成的视频已经实现了突破性的进步，但 OpenAI 同时也表示，目前 Sora 生成的部分视频中仍存在一些问题，在复杂场景下无法很好地模拟物体的物理规律以及部分场景下模型可能无法理解事物的因果关系，例如当生成一个人吃饼干的视频时饼干可能没有留下咬痕。在 OpenAI 官网展示的演示视频中，也展示了人物跑步方向和跑步机相反、篮球触碰篮筐没有反弹而是直接穿过篮筐等问题。

不可否认，此类问题的出现将会影响视频生成的准确性和真实程度，导致生成视频的实用性下降。但相信随着技术的推进，这些问题也有望被逐渐克服，AI 视频生成技术的应用潜力也将得到进一步释放。

图 9：男子跑步方向和跑步机相反



资料来源：OpenAI 官网，中原证券

图 10：篮球没有触碰篮筐而是直接穿过



资料来源：OpenAI 官网，中原证券

由于目前与生成式 AI 技术相关的法律还不够健全，AI 生成的内容仍存在版权、伦理、争议等多方面的潜在未知风险，因此 Sora 暂未向公众开放，仅有少部分视觉艺术家、设计师、影视制作人以及网络安全专家能够进行测试使用，OpenAI 也正在收集相关人员的反馈并优化模型，未来有望向机构及大众用户开放，带动 AI 内容生产的新机会。

2. AI 技术进步扩大应用场景，传媒产业受益显著

相比此前 AI 生成的视频，Sora 生成的视频不论是长度、运镜、镜头切换能力还是生成画面的细节、一致性、稳定性以及物理特性等方面都有了明显的突破，AI 生成视频即将从探索性阶段向实用性阶段逐渐过渡，随着这一过程的演变其应用场景也将得到大幅拓宽。

从内容生产的角度来看，Sora 已经展现出了比较强的多模态内容生产能力，未来以 Sora 为代表的 AI 视频工具在游戏、广告营销、影视等多个内容产品领域都具有广泛的应用空间。一方面通过 AI 工具提升内容生产的效率，有效减少视频拍摄和制作环节所花费的时间和成本，缩

短内容创作周期，未来的内容产业中，部分非核心内容的创作需求可能被 AI 模型替代；另一方面借助 AI 工具的力量能够降低内容创作的门槛，使更多的创作者参与到内容创作之中，提升内容产品的丰富性和创意性，同时非专业团队或内容创作者通过使用 AI 模型也能够创作出媲美专业团队的高质量内容产品，提升内容产品的质量。

但从短期来看，考虑到目前 Sora 展现出的能力仍有一定缺陷，同时相比 AI 文字模型和 AI 图像模型，AI 视频模型对于算力的需求和使用成本也会更高，因此 Sora 等 AI 视频模型的短期定位仍将是面向企业或团队的生产辅助工具。长期的 AI 模型进步以及边际成本的下降有望带动更多 UGC 或 PUGC 内容的增量空间。

在游戏产业，游戏产品的研发和发行过程涉及大量内容要素的创作，利用 Sora 等 AI 模型的多样化功能能够生成概念图、人物动态立绘、宣传发行素材、游戏动画等内容要素，降低原本由人工创作所需要的创作周期和成本。同时 OpenAI 的演示视频中还显示了 Sora 在高保真地渲染游戏世界和动态的同时还能够利用基本策略操控游戏角色。

在广告产业，AI 视频模型能够将抽象的广告文案快速地转化为具象的广告视频 DEMO，同时相比人类创作内容，AI 模型创作的内容可能更加不受逻辑的限制，也能为广告创作人员提供全新的创意参考。总体而言，利用 AI 视频模型能够提升广告创作过程中的产出效率和创意效果，缩短广告内容的创作时间，降低成本。

在影视产业，AI 视频模型可用于动画、短视频、影视剧等内容的辅助创作，前期的剧本概念、分镜等内容的创作、中期的影视拍摄以及后期的特效制作等环节能够通过使用 AI 模型工具的方式降低成本和风险，同时减少对于实景拍摄的需求，进一步压缩制作成本。这其中我们认为动画类影视作品的制作将会受益最大，动画影视主要依赖数字和技术工具进行制作，AI 模型具有天然优势，能够有效提升制作效率。2 月 26 日，由中央广播电视总台制作的中国首部文生视频 AI 动画片《千秋诗颂》在 CCTV 开播。根据财联社的相关报道，《千秋诗颂》由央视听媒体大模型制作，其中美术设计、动效生成和后期成片等环节均有 AI 辅助制作。在同等预算条件下，如果按照传统动画制作流程计算，《千秋诗颂》至少需要 8 个月时间，依托大模型后，制作周期缩短至 4 个月。

建议关注：恺英网络、三七互娱、完美世界、吉比特、芒果超媒、风语筑、光线传媒、中国电影、万达电影

3. 风险提示

AI 生成内容存在伦理、道德和法律风险；AI 工具发展进展不及预期

行业投资评级

强于大市：未来 6 个月内行业指数相对沪深 300 涨幅 10% 以上；

同步大市：未来 6 个月内行业指数相对沪深 300 涨幅-10% 至 10% 之间；

弱于大市：未来 6 个月内行业指数相对沪深 300 跌幅 10% 以上。

公司投资评级

买入：未来 6 个月内公司相对沪深 300 涨幅 15% 以上；

增持：未来 6 个月内公司相对沪深 300 涨幅 5% 至 15%；

谨慎增持：未来 6 个月内公司相对沪深 300 涨幅-10% 至 5%；

减持：未来 6 个月内公司相对沪深 300 涨幅-15% 至 -10%；

卖出：未来 6 个月内公司相对沪深 300 跌幅 15% 以上。

证券分析师承诺

本报告署名分析师具有中国证券业协会授予的证券分析师执业资格，本人任职符合监管机构相关合规要求。本人基于认真审慎的职业态度、专业严谨的研究方法与分析逻辑，独立、客观的制作本报告。本报告准确的反映了本人的研究观点，本人对报告内容和观点负责，保证报告信息来源合法合规。

重要声明

中原证券股份有限公司具备证券投资咨询业务资格。本报告由中原证券股份有限公司（以下简称“本公司”）制作并仅向本公司客户发布，本公司不会因任何机构或个人接收到本报告而视其为本公司的当然客户。

本报告中的信息均来源于已公开的资料，本公司对这些信息的准确性及完整性不作任何保证，也不保证所含的信息不会发生任何变更。本报告中的推测、预测、评估、建议均为报告发布日的判断，本报告中的证券或投资标的价格、价值及投资带来的收益可能会波动，过往的业绩表现也不应当作为未来证券或投资标的表现的依据和担保。报告中的信息或所表达的意见并不构成所述证券买卖的出价或征价。本报告所含观点和建议并未考虑投资者的具体投资目标、财务状况以及特殊需求，任何时候不应视为对特定投资者关于特定证券或投资标的的推荐。

本报告具有专业性，仅供专业投资者和合格投资者参考。根据《证券期货投资者适当性管理办法》相关规定，本报告作为资讯类服务属于低风险（R1）等级，普通投资者应在投资顾问指导下谨慎使用。

本报告版权归本公司所有，未经本公司书面授权，任何机构、个人不得刊载、转发本报告或本报告任何部分，不得以任何侵犯本公司版权的其他方式使用。未经授权的刊载、转发，本公司不承担任何刊载、转发责任。获得本公司书面授权的刊载、转发、引用，须在本公司允许的范围内使用，并注明报告出处、发布人、发布日期，提示使用本报告的风险。

若本公司客户（以下简称“该客户”）向第三方发送本报告，则由该客户独自为其发送行为负责，提醒通过该种途径获得本报告的投资者注意，本公司不对通过该种途径获得本报告所引起的任何损失承担任何责任。

特别声明

在合法合规的前提下，本公司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头寸并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问等各种服务。本公司资产管理部、自营部门以及其他投资业务部门可能独立做出与本报告意见或者建议不一致的投资决策。投资者应当考虑到潜在的利益冲突，勿将本报告作为投资或者其他决定的唯一信赖依据。