

证券研究报告

OpenAI Sora 模型发布，视频生成技术迎来突破性升级

行业研究

2024 年 2 月 19 日

点评报告

计算机

投资评级：看好

上次评级：看好

庞倩倩 计算机行业首席分析师

执业编号：S1500522110006

邮箱：pangqianqian@cindasc.com

- **事件：**2月15日，OpenAI 发布 Sora AI 视频生成模型，文本生成视频模型迎来重大突破。此次发布的 Sora 模型能够根据用户的文本描述生成长达 60 秒、1080P 高质量视频，其中包含精细复杂的场景、生动的角色表情以及复杂的镜头运动。
- **模型亮点：**1) 视频生成时间长达 1 分钟，大幅领先其他 AI 视频生成模型。2) 单视频多角度镜头。3) 3D 一致性：Sora 可以生成摄像机动态运动的视频。随着摄像机的移动和旋转，人物和场景元素在 3D 空间中的移动会保持一致。4) Sora 可以向前向后拓展视频，连接视频。5) 模拟数字世界：Sora 还能够模拟人工过程，例如视频游戏。
- **Sora 模型的技术实现特点：**Sora 基于扩散模型，并且在其中使用了 Transformer 架构，将视频和图像分解为 Patch 小数据单元，这些 Patch 类似于 GPT 中的 Token，用于在 Transformer 模型中进行训练和生成；同时 OpenAI 将 DALL·E3 引入到 Sora 里，使得 Sora 能够精准地还原用户的文本提示，生成高质量的长视频。
- **应用展望：**随着文生图、图片对话技术的成熟，文生视频可能成为多模态大模型下一步发展的方向。该技术有望极大地降低短剧制作的成本，并且会进一步向影视、自媒体、游戏等领域渗透，提高视频创作效率等。
- **风险提示：**行业竞争加剧、金融科技发展不及预期。

信达证券股份有限公司

CINDA SECURITIES CO., LTD

北京市西城区闹市口大街9号院1号楼

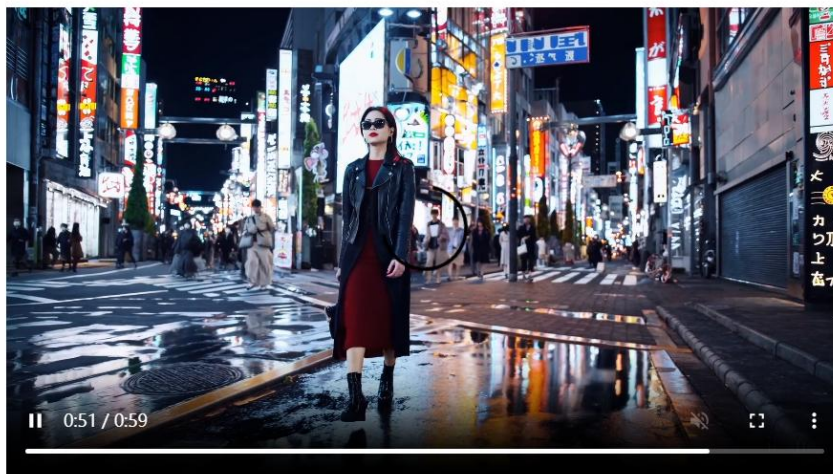
邮编：100031

## 1. OpenAI 发布 Sora 模型，视频生成亮点突出

2月15日，OpenAI 发布 Sora AI 视频生成模型，文本生成视频模型迎来重大突破。此次发布的 Sora 模型能够根据用户的文本描述生成长达 60 秒、1080P 高质量视频，其中包含精细复杂的场景、生动的角色表情以及复杂的镜头运动。

1) 视频生成时间长达 1 分钟，大幅领先其他 AI 视频生成模型。目前文本视频生成热门模型中，pika 1.0 默认生成 3 秒视频，可拓展到 7 秒；而 runaway gen2 模型最高可生成 18 秒视频。

图 1: Sora 生成 1 分钟视频示例

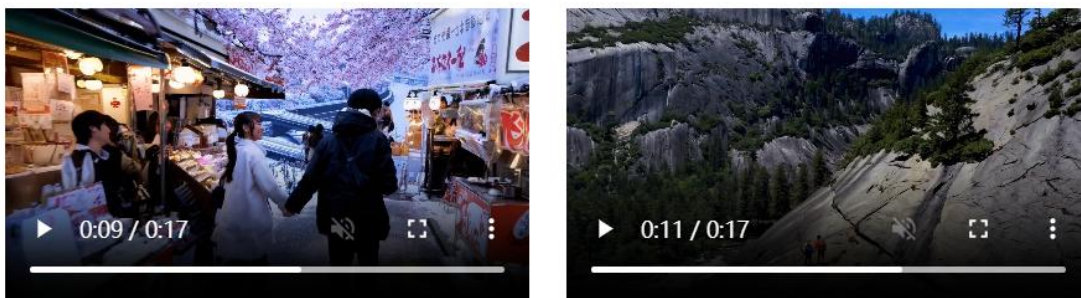


资料来源: OpenAI 官网, 信达证券研发中心

2) 单视频多角度镜头: 在单个样本中生成同一角色的多个镜头, 实现各种中、近、远景的切换, 并在整个视频中稳定保持其外观。

3) 3D 一致性: Sora 可以生成摄像机动态运动的视频。随着摄像机的移动和旋转, 人物和场景元素在 3D 空间中的移动会保持一致。长距离一致性和对象持久性, 是生成高质量长视频的一个重大挑战。Sora 经常能够有效地模拟短距离和长距离依赖关系。例如, 即使人、动物和物体被遮挡或离开画面, 也能保持它们的存在。

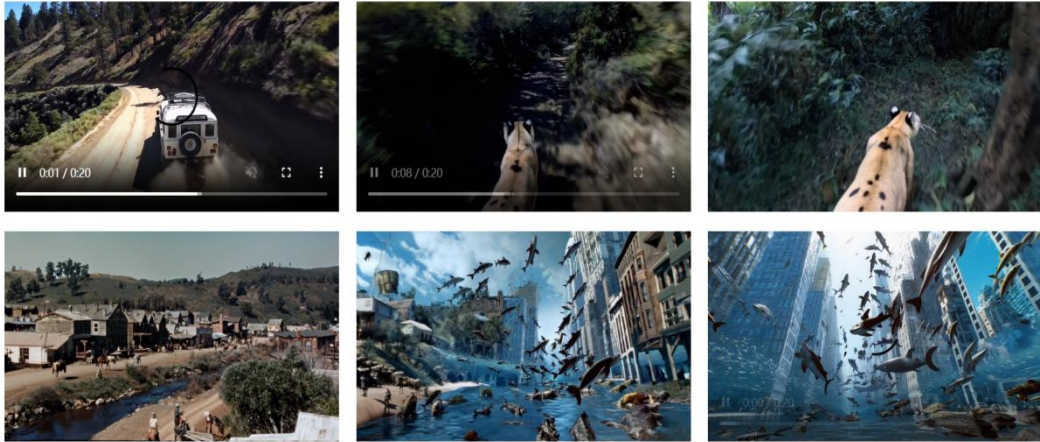
图 2: Sora 生成动态运动视频示例



资料来源: OpenAI 官网, 信达证券研发中心

4) Sora 可以向前向后拓展视频，连接视频。可以使用 Sora 在两个输入视频之间逐渐插值，在具有完全不同主题和场景构图的视频之间创建无缝过渡。

图 3: Sora 连接视频示例，中间的视频在左侧和右侧的相应视频之间进行插值。



资料来源: OpenAI 官网, 信达证券研发中心

5) 模拟数字世界: Sora 还能够模拟人工过程, 例如视频游戏。Sora 可以同时通过基本策略控制 Minecraft 中的玩家, 同时还可以高保真地渲染世界及其动态。

**视频生成质量、能力上领先其余模型。**Sora 模型不仅在视频时长上远超其他文生视频模型, 其在视频长宽比、视频拓展等功能上也领先于其他模型。OpenAI Sora 支持 1920x1080 与 1080x1920 之间任意尺寸变换, 而其他工具 (Runway Gen2、Pika 等) 只有固定尺寸如 16: 9、1: 1 等; 另外相比于其余模型只能向后拓展视频, Sora 支持前后双向拓展。另外在视频连接、运动相机模拟等能力上也有一定程度的优势。

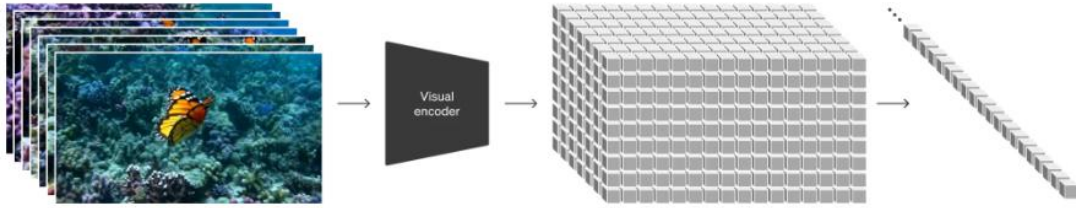
## 2.Sora 模型的技术实现特点分析

Sora 基于扩散模型, 其机理是从静态噪声开始, 通过多个步骤逐渐去除噪声, 最终生成清晰视频。这种模型能够处理视频和图片中时空片段的潜代码。并且在其中使用了 Transformer 架构, 这种架构在处理语言模型、计算机视觉和图像生成等领域表现出了卓越的扩展能力。

并且 Sora 使用“视频补丁”Patch, OpenAI 发现, 视频补丁是一种高度可扩展且有效的表示形式, 可用于在不同类型的视频和图像上训练生成模型。在较高维度上, OpenAI 首先将视频压缩到低维潜在空间中, 然后将其分解为时空 Patch, 这些 Patch 类似于 GPT 中的 Token, 用于在 Transformer 模型中进行训练和生成。Sora 的扩散 Transformer 模型能够有效地扩展, 处理不同分辨率、持续时间和纵横比的视频和图片。



图 4：Sora 将视频分解为 Patch



资料来源：OpenAI 官网，信达证券研发中心

**精确文本语义理解：**同时 OpenAI 将 DALL·E3 引入到 Sora 里，首先训练一个高度描述性的字幕模型，然后用它为训练集中的所有视频制作文本字幕。OpenAI 表示，在高度描述性的视频字幕上进行训练，可提高文本的保真度以及视频的整体质量。与 DALL·E3 一样，OpenAI 也通过 GPT 将简短的用户提示转化为较长的详细字幕，并发送给视频模型。这使得 Sora 能够精准地还原用户的文本提示，生成高质量的长视频。

**模型局限性：**Sora 目前还存在一些局限性。例如，它不能准确模拟许多基本交互的物理现象，如玻璃碎裂。其他交互，如吃食物，并不总能产生正确的物体状态变化。官方主页列举了该模型的其他常见失效模式，例如长时间样本中出现的不一致性或物体的自发出现等。

### 3.Sora 模型的应用展望

随着文生图、图片对话技术的成熟，文生视频可能成为多模态大模型下一步发展的方向。该技术有望极大降低短剧制作的成本，并且会进一步向影视、自媒体、游戏等领域渗透，提高视频创作效率等。Sora 展示了 AI 在理解和模拟物理世界方面的能力，这被认为是实现通用人工智能（AGI）的关键一步（AGI 是指能够在多种任务和环境中灵活应用知识的 AI）。

## 研究团队简介

**庞倩倩**，计算机行业首席分析师，华南理工大学管理学硕士。曾就职于华创证券、广发证券，2022 年加入信达证券研究开发中心。在广发证券期间，所在团队 21 年取得：新财富第四名、金牛奖最佳行业分析师第二名、水晶球第二名、新浪金麒麟最佳分析师第一名、上证报最佳分析师第一名、21 世纪金牌分析师第一名。

**姜恺非**，计算机行业研究员，悉尼大学商学硕士，2023 年加入信达证券研究所，主要覆盖金融 it、网络安全等领域。

## 分析师声明

负责本报告全部或部分内容的每一位分析师在此申明，本人具有证券投资咨询执业资格，并在中国证券业协会注册登记为证券分析师，以勤勉的职业态度，独立、客观地出具本报告；本报告所表述的所有观点准确反映了分析师本人的研究观点；本人薪酬的任何组成部分不曾与，不与，也将不会与本报告中的具体分析意见或观点直接或间接相关。

## 免责声明

信达证券股份有限公司(以下简称“信达证券”)具有中国证监会批复的证券投资咨询业务资格。本报告由信达证券制作并发布。

本报告是针对与信达证券签署服务协议的签约客户的专属研究产品，为该类客户进行投资决策时提供辅助和参考，双方对权利与义务均有严格约定。本报告仅提供给上述特定客户，并不面向公众发布。信达证券不会因接收人收到本报告而视其为本公司的当然客户。客户应当认识到有关本报告的电话、短信、邮件提示仅为研究观点的简要沟通，对本报告的参考使用须以本报告的完整版本为准。

本报告是基于信达证券认为可靠的已公开信息编制，但信达证券不保证所载信息的准确性和完整性。本报告所载的意见、评估及预测仅为本报告最初出具日的观点和判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会出现不同程度的波动，涉及证券或投资标的的历史表现不应作为日后表现的保证。在不同时期，或因使用不同假设和标准，采用不同观点和分析方法，致使信达证券发出与本报告所载意见、评估及预测不一致的研究报告，对此信达证券可不发出特别通知。

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，也没有考虑到客户特殊的投资目标、财务状况或需求。客户应考虑本报告中的任何意见或建议是否符合其特定状况，若有必要应寻求专家意见。本报告所载的资料、工具、意见及推测仅供参考，并非作为或被视为出售或购买证券或其他投资标的的邀请或向人做出邀请。

在法律允许的情况下，信达证券或其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，并可能会为这些公司正在提供或争取提供投资银行业务服务。

本报告版权仅为信达证券所有。未经信达证券书面同意，任何机构和个人不得以任何形式翻版、复制、发布、转发或引用本报告的任何部分。若信达证券以外的机构向其客户发放本报告，则由该机构独自为此发送行为负责，信达证券对此等行为不承担任何责任。本报告同时不构成信达证券向发送本报告的机构之客户提供的投资建议。

如未经信达证券授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。信达证券将保留随时追究其法律责任的权利。

## 评级说明

投资建议的比较标准	股票投资评级	行业投资评级
本报告采用的基准指数：沪深 300 指数（以下简称基准）； 时间段：报告发布之日起 6 个月内。	<b>买入</b> ：股价相对强于基准 20% 以上；	<b>看好</b> ：行业指数超越基准；
	<b>增持</b> ：股价相对强于基准 5%~20%；	<b>中性</b> ：行业指数与基准基本持平；
	<b>持有</b> ：股价相对基准波动在±5% 之间；	<b>看淡</b> ：行业指数弱于基准。
	<b>卖出</b> ：股价相对弱于基准 5% 以下。	

## 风险提示

证券市场是一个风险无时不在的市场。投资者在进行证券交易时存在赢利的可能，也存在亏损的风险。建议投资者应当充分深入地了解证券市场蕴含的各项风险并谨慎行事。

本报告中所述证券不一定能在所有的国家和地区向所有类型的投资者销售，投资者应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专业顾问的意见。在任何情况下，信达证券不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任，投资者需自行承担风险。