

科技速览

2024年第3期

(总第10期)

科普中国智库粤港澳大湾区研究中心

广州市科普工作联席会议办公室

2024年3月20日

【编者按】2月18日，广东省召开全省高质量发展大会，提出“推进产业科技创新、发展新质生产力是广东的战略之举、长远之策，也必将是一场艰苦的竞速赛、耐力赛、接力赛”。2月19日，广州市接着召开高质量发展大会，提出广州要继续在高质量发展方面发挥领头羊和火车头作用，必须在发展壮大新质生产力上走前列、作示范，把握“创新”这个特点勇攀科技高峰，催生更多新产业、新模式、新动能；把握“质优”这个关键优化产业结构，强化产业科技互促双强，驱动生产力向新的质量形态跃升；把握“先进生产力”这个本质深化改革扩大开放，促进各类生产要素向新质生产力顺畅流动，实现从科技强到企业强、产业强、经济强。最近，文生视频引发广泛关注，可能带来颠覆性变革。本期《科技速览》收集整理了sora（由美国人工智能研究机构OpenAI开发的一个文生视频大模型，输入文本即可得到视频）专题信息，供参阅。

一、Sora 的出现引发关注

1.OpenAI Sora 可一次生成多个视频拼接展示。OpenAI 研究科学家比尔·皮布尔斯(Bill Peebles)在 X(“推特”转型后改名“X”)上发文称，Sora 可以同时生成多个并排在一起的视频。根据其推文展示的动图，Sora 生成了 5 个人在雪天漫步、玩雪的视频。目前，OpenAI 尚未向公众开放 Sora，正在接受测试，只分享给一批精选的研究人员和学者研究，以寻找该模型被滥用的风险。(摘编自 2 月 18 日凤凰网科技)

2.Sora 可以做什么？OpenAI 发布首个视频生成模型 Sora 可以呈现“具有多个角色、特定类型的动作以及主题和背景的准确细节的复杂场景”。如 AI 想象中的龙年春节，红旗招展，人山人海，有紧跟舞龙队伍抬头好奇观望的儿童，还有不少人掏出手机边跟边拍；雨后东京街头，潮湿地面反射霓虹灯，光影效果堪比 RTX ON (RTX 是指采用了英伟达 Turing GPU 架构的高性能显卡，提供实时光线追踪和 AI；RTX ON 形容各种画质突然提升的场景)；行驶中的列车窗外偶遇遮挡，车内人物倒影短暂出现，非常惊艳。目前 Sora 已经能生成具有多个角色、包含特定运动的复杂场景，不仅能理解用户在提示中提出的要求，还了解这些物体在物理世界中的存在方式。比如一大群纸飞机在树林中飞过，Sora 知道碰撞后会发生什么，并表现其中的光影变化。Sora 还可以在单个视频中创建多个镜头，并依靠对语言的深入理解准确地解释提示词，保留角色和视觉风格。对于 Sora 存在的弱点，

OpenAI 也指出它可能难以准确模拟复杂场景的物理原理，并且可能无法理解因果关系。例如“五只灰狼幼崽在一条偏僻的碎石路上互相嬉戏、追逐”，狼的数量会变化，一些狼凭空出现或消失。该模型还可能混淆提示的空间细节，例如混淆左右，并且可能难以精确描述随着时间推移发生的事件。目前有一些视觉艺术家、设计师和电影制作人以及 OpenAI 员工获得了 Sora 访问权限。（摘编自 2 月 16 日财联社）

3.Sora 的突破性。360 集团创始人周鸿祎说，最重要的是 Sora 的技术思路完全不一样，在这之前我们做视频做图用的都是 Diffusion（一种生成模型，用于在深度学习和人工智能领域生成高质量、逼真的数据，尤其是在图像生成方面表现出色），你可以把视频看成是多个真实图片的组合，它并没有真正掌握这个世界的知识。现在所有的文生图、文生视频都是在 2D 平面上对图形元素进行操作，并没有适用物理定律。但 Sora 产生的视频里，它能像人一样理解坦克是有巨大冲击力的，坦克能撞毁汽车，而不会出现汽车撞毁坦克这样的情况。这次 OpenAI 利用大语言模型优势，把 LLM（大型语言模型）和 Diffusion 结合起来训练，实现了对现实世界的理解和对世界的模拟两层能力，这样产生的视频才是真实的，并跳出 2D 模拟真实的物理世界。（摘编自 2 月 19 日中环狙击手）

4.Sora 刷屏视频仍有一些瑕疵。Sora 视频的质量令人惊叹，但目前，除了作为展示 AI 潜力的范例作用外，实用落地价值仍

待观察。例如针对 OpenAI 定位“世界模拟器”的观点，图灵奖得主、Meta 首席 AI 科学家杨立昆表示，根据提示词生成的大部分逼真视频并不表明这样的 AI 系统理解物理世界，生成视频的过程与基于世界模型的因果预测完全不同。如老奶奶吹蜡烛火苗却不动、东京街头步伐错误的女郎、玻璃杯碎后错误的液体流向、在跑步机上反向跑步的男子等。导演牛萌琛表示，Sora 视频中的画质，对广告片拍摄来说，也可以用这种样板来与客户更好地进行概念上的沟通和确认。但如果涉及镜头运动角度，与更精细内容调控，比如色彩、光影、道具、人物动作等，目前还不能达到令客户满意的程度。摄影专业人士孟凡表示，Sora 视频帧率较高，说明计算能力比较强，但是 Sora 视频的逻辑性会差一点。（摘编自 2 月 18 日第一财经）

二、Sora 带来的深刻变革

1.Sora 将引领新一轮行业变革。2 月 16 日，OpenAI 公司发布的 Sora 有可能促使人工智能引领新一轮行业变革。在部分样片中，Sora 还展现了对“物理规律”超强的学习能力，无论是飘逸的毛发还是水体波纹，都能“合乎常理”地呈现出来。360 集团创始人周鸿祎连续发声称，Sora 的诞生意味着 AGI（通用人工智能）实现时间可能从 10 年缩短至一两年。昆仑万维 CEO 方汉认为，从技术上来看，Sora 领先国内同行大概半年。Sora 的问世对影视、视频、广告等行业或将带来不小的冲击。（摘编自 2 月 19 日上海证券报）

2.Sora 会冲击谁？ OpenAI 一次性发布了 48 个由 Sora 生成且未经任何修改的视频，这些视频的时长各不相同，最短的仅有 9 秒，而最长的则达到了 60 秒。一个具备理解、感知及创造内容能力的 AI 视频模型，正将开启无尽的可能性。首当其冲的便是影视行业，以及广告、新闻、教育等领域，一些简单的、重复性的视频制作工作可能会被 Sora 等技术所取代。更为关键的是，如果说之前的 AI “文生视频” 工具仅仅停留在 “模拟现实” 的层面，那么 Sora 则已经跃升到了 “构建现实” 的新高度。在现实中，我们每咬下一口食物，食物上都会留下咬痕，这是遵循物理规则的自然现象。而 Sora 生成的视频，同样能够精准地再现这一细节，做到 “咬下去有痕迹”，从而在虚拟世界中完美再现了现实世界的真实感。而这是其他文生视频产品做不到的。（摘编自 2 月 19 日风暴眼）

3.Sora 意味着 AGI 实现将从 10 年缩短到 1 年。 360 集团创始人周鸿祎认为，Open AI 训练这个模型应该会阅读大量视频，大模型加上 Diffusion 技术需要对这个世界进行进一步了解，学习样本就会以视频和摄像头捕捉到的画面为主。一旦人工智能接上摄像头，把所有的电影都看一遍，把 YouTube 上和 TikTok 的视频都看一遍，对世界的理解将远远超过文字学习，一幅图胜过千言万语，而视频传递的信息量又远远超过一幅图，这就离 AGI 真的就不远了。（摘编自 2 月 17 日证券时报）

4.AGI 只给人类留下 5 年时间？ 英国卫报采访了多位担忧 AI

技术将给人类带来极大负面影响的人士。在加州伯克利的非营利机构“机器智能研究所”担任首席研究员 Yudkowsky 看来，人类所剩的时间可能 5 年，反正肯定不超过 10 年。他所指的“剩余时间”其实就是直到人类遭遇由机器人引发的末日灾难的时间。他说，未来人类要生存下来，可能需要对 AI 的数据中心进行空袭，甚至使用核武器，才能保存自己。纽约作家 Edward Ongweso Jr 和艺术家 Molly Crabapple 也将自己定义为“愚人”。作为典型的 AI 保守主义者，他们不是要放弃创新，而是要求审视每项创新的优点、社会公平性及其潜在的隐蔽恶意。（摘编自 2 月 21 日新智元）

5.Sora 让好莱坞紧张了？ Sora 令 IT 界、艺术界和普通爱好者兴奋不已，同时也令娱乐产业感到担忧。《综艺》评论称，Sora 的“现实主义效果”将有望在高端娱乐中发挥作用，只是因为某些限制，目前还无法胜任好莱坞电影这样的制作规模。人工智能公司 Metaphysics 的 CEO 汤姆·格拉汉姆认为：“如今创意工作者要求对表演、场景进行全面把控，Sora 这样的模型要想生成好莱坞电影，还有很长的路要走”。在 Sora 横空出世之前，业内已经存在 Runway（一种 AI 文生视频、图生视频软件）和 Pika（另一种 AI 文生视频、图生视频软件）等功能相似的 AI 视频生成、编辑模型。这些模型可以基于文本提示、2D 图片生成短视频，由静态图像衍生出动画，修复、替换或插入新的视觉元素，以及将镜头扩展到原始帧之外，填充上下文相关内容的“画外

画”，这些都丰富了人们所需的视频内容。此外，版权等法律问题也是好莱坞目前无法直接使用 Sora 的一大原因。（摘编自 2 月 20 日环球时报）

三、国内 Sora 相关产业动态

1.我国多家公司积极布局视频生成模型赛道。萨摩耶云科技集团首席经济学家郑磊表示，在视频生成领域，已有海康威视、佳都科技、昆仑万维、大华股份、萤石网络、万兴科技、因赛集团、拓尔思、当虹科技等多家 A 股上市公司积极布局。商务部研究院电子商务研究所副研究员洪勇表示，中国企业在某些细分领域已具备国际竞争力，但仍有短板，当下需要清醒意识到差距，加强基础科研、培养高层次人才、建立合规体系，并积极参与国际合作与竞争，进一步提升在全球人工智能领域的地位。（摘编自 2 月 19 日证券日报）

2.国内首部 AIGC（即 AI Generated Content，是指利用人工智能技术来生成内容）动画开播。2 月 26 日，首部中国原创文生视频 AI 系列动画《千秋诗颂》于 CCTV-1 综合频道播出。业内人士分析，随着国内首部 AI 动画开播，文生视频应用落地有望提速，将重塑影视行业制作流程，进一步推动影视动画制作、游戏开发、广告设计等降本增效。据介绍，在同等预算的条件下，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，在 AIGC 赋能下，制作周期缩短至 4 个月。（摘编自 2 月 26 日智通财经）

3.集中国内所有 AI 企业的算力或许是方法之一。360 集团创始人周鸿祎认为，Sora 的技术路线如果被开源，国内将能很快赶上，但在追赶 Sora 时，算力有可能成为门槛，对此，集中国内所有 AI 企业的算力或许是方法之一。(摘编自 2 月 24 日新京报)

4.AI 时代将开启资本市场的伟大征程。深圳东方港湾投资管理股份有限公司董事长但斌表示，继电子硬件时代、互联网时代、移动互联网之后，第四个人工智能时代将开启资本市场的伟大征程。电子硬件时代，美国纳指大概 16 年涨了 6.5%，互联网时代大概 10 年涨了 13 倍，移动互联网时代大概 13 年涨了 10 倍，人工智能是比前三个时代更没有上限的技术革命。(摘编自 3 月 3 日环球市场播报)

报：永航书记、志洋市长、衍诗主任、贻伟主席

市委常委、人大副主任、副市长、政协副主席

送：市科普工作联席会议成员单位；各区科普工作联席会议办公室。

总 编：徐 柳

编 审：张 勇

编 辑：陈晓萍、冯振行

编辑单位：科普中国智库粤港澳大湾区研究中心

市科普工作联席会议办公室 电 话：61106390

市博士科技创新研究会 电 话：33975064
