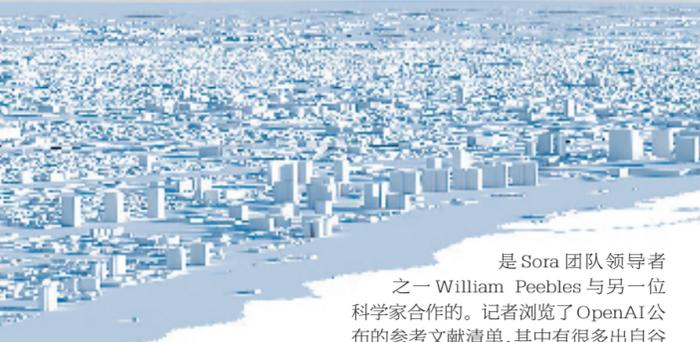


OpenAI 文生视频大模型带来挑战的同时,更多人关注—— Sora 问世,带来哪些新机遇

本报记者 金春华 何冬健 姚颖康 潮新闻记者 朱高祥



分镜头多角度,画面精细流畅,时长是此前市场上产品的2-5倍……2月16日,曾因ChatGPT一炮走红的人工智能研究公司OpenAI推出的文生视频大模型Sora,令全球从事文字、音视频以及科技的工作人士惊叹不已。

就在OpenAI已公布的演示视频中,我们看到一名时尚女性走在充满温暖霓虹灯和动画城市标牌的东京街头;看到两艘海盗船在咖啡杯中缠斗,咖啡如海浪汹涌;看到可爱的水獭穿着黄色救生衣站在滑板上冲浪……惟妙惟肖,与实景拍摄、高清特效无异。网友的评价也出奇一致:炸裂!惊艳!颠覆!

而这,也被视作整个AI产业、视频产业的新挑战、新机遇。连日来,有人为“风口”的到来兴奋不已;有人如音视频领域从业者感受到了危机:“要失业了。”

Sora到底厉害在哪里?它又将掀起哪些新的浪潮?

意料之中与出乎意外

“没想到这一天来得这么快!”“没想到效果这么有颠覆性。”这是记者采访业界人士时听到最多的两句话。

尽管截至记者发稿,Sora尚未公开发布或未公测,有的只是OpenAI公布的数十段演示视频。但前有ChatGPT为证,后有Sora远超其他模型的效果——任谁都会思考,一个新的时代是否来了?

Sora可简单视作一个人工智能视频生成工具:只需我们输入一段文字,它就能创建一个与真实世界接近的虚拟世界,时长可达60秒。此前在文生视频领域,谷歌、Meta、Runway、Pika等产品大都在30秒甚至10秒之内。

更让人啧啧称奇的是,Sora生成的视频,无论是光影色彩的转变,还是多镜头多角度的切换,甚至细微到纹理结构变化,都呈现出“大片感”。

“Sora还能模拟现实世界的物理规律。这尤其有难度。”阿里巴巴大模型算法专家曹宇举例说,其中有一段视频展现的是两艘船在咖啡中战斗,Sora精准呈现了咖啡流动的流体动力学效果。

要知道,在文字、图片、视频等不同体裁中,视频生成是最难的。Sora不仅要理解文本指令的含义,即语义理解能力,不能输入“女孩”出来“老头”,也不能违背逻辑与物理规则,如输入“狗坐在椅子上”出来“狗嵌入了椅子”;生产效果要好,画面流畅程度、稳定性、连贯性等都不能缺,比如太阳在左上角,不能右下角又无故生出阳光;还有就是时长,时间越长,视频表现力越强,AI“露馅”的可能性也越大。

Sora有远超同类的表现,其核心技术于底层逻辑却被普遍认为“还是老一套”,即时空Patch技术和DiT架构。前者的技术论文由谷歌DeepMind团队于2023年7月发表。后者的技术论文,

是Sora团队领导者之一William Peebles与另一位科学家合作的。记者浏览了OpenAI公布的参考文献清单,其中有很多出自谷歌、Meta、微软、斯坦福、MIT、UC伯克利等机构的成果。也就是说,Sora的诞生源于整个AI社区多年来的探索,是“意料之中”的事情。

“事实上,Sora依然遵循OpenAI的规模理论——大量数据、大模型和大量算力。”浙江大学杭州国际科创中心求是科创学者、AI交叉中心专家丁科炎估计,Sora与ChatGPT一样,也是来自大力出奇效的“暴力美学”。

此前,AI视频创业公司Runway和Pika被认为是视频生成赛道的佼佼者,如今却被Sora轻松“吊打”。业内人士分析,Sora制胜之道,一在其使用的图文对、视频等的数量比这两家大一个量级以上;二在算力堆叠的效果更好。

在杭州城市大脑有限公司CEO申永生看来,Sora体现了AI对现实世界的理解已更上一层楼,“它在训练中结合了多种先进的算法,比如masked(遮掩)算法,把一个画面随机遮掩掉30%局部,让AI根据对画面整体理解还原出来,同时利用生成式算法让画面比原作更富有创意。”

曹宇则表示,Sora的基础架构Transformer能实现“作文续写”,相比于传统“完形填空”式的技术拥有更优秀的创作能力。

“打个比方,我们可以将一堆杂乱无章的积木整理好放入一个小盒子中。如此一来,即便面对众多积木,只要找到了这个小盒子就能轻松找到所需积木。”丁科炎也给出自己的解读,由于前期用于训练的海量视频数据被转化成一个“小方块”,当我们向Sora提供一个新任务时,AI就会从视频数据中提取出一些包含时间和空间信息的“小方块”,并将之交给Sora,让其根据这些信息生成新视频。

“某种程度上,Sora已经可以像人一样,对世界有了一定认知。”丁科炎说。

浪漫与现实

Sora带来的产业变革,被视作技术进步产生的“浪漫”。

“相比于ChatGPT,Sora的产业空间更广,影视剧、游戏视频、广告策划、创意和设计可视化、数字孪生、自动驾驶、医疗等都是它的舞台。”申永生表示,视频与文本有着本质的区别,其开放空间很大,“比如我们在做城市大脑新场景应用的推广宣传片时,以前要么是找人拍,要么做动漫,成本很高,创新挑战的压力也很大;有了Sora这类新工具,我们可以用更少的成本、更快的速度、更贴近现实的视频来做推广,效果也会更好。”

这种强大的生成能力,或将打开AIGC新的空间,特别是突破知识产权领域的瓶颈。

浙江震天律师事务所律师钱航告诉记者,是否具有“独特个性”,是传统著作权法理论判断一件创作物是否是作品的关键因素之一。此前不少专家反对人工智能创作物具有独创性,主要认为人工



OpenAI发布的Sora生成视频图。

图源:Sora官网



新年伊始,丁科炎(左)与张强在一起畅聊大模型发展。

本报记者 姚颖康 摄



浙大杭州国际科创中心借助AI工作的“合成生物学自动化科学装置”。

浙大杭州国际科创中心供图

智能的创作物是算法、模板等统一产生的结果,“但现在看来,一方面Sora等新工具的创新力明显更强;另一方面,Sora生成视频的好坏也更受制于提示文本的好坏、文本创作者创意的好坏,其独创性也更明显。”

更重要的是,Sora一旦公开应用,其强大的能力肯定会被大量用于商业用途,成为具备经济价值的交易客体,随之也会产生越来越多的法律问题。钱航认为:“当下应该尽快立法,明确人工智能创作物的法律属性,完善相关权利保障体系,一方面激励人工智能及相关产业发展,一方面维护现有版权秩序。”

很多“浪漫遐想”已经展开。

浙江省工业和信息化学研究院区域经济研究所所长赵立龙认为,随着Sora的到来,其“模拟数字世界”能力将对元宇宙发展产生颠覆性影响。它不仅能够直接根据需求生成虚拟视频,迅速构建元宇宙所需的场景,而且其展现出了直接生成可深入探索的动态虚拟空间的潜力,这意味着元宇宙将进入到全新的发展阶段。

影视行业的变化最受关注。

工信部信息通信经济专家委员会委员、数字经济专家刘兴亮称,AI生成的视频可能减少了对人类演员、导演和其他创意角色的需求,从而影响到这个行业的就业。

“Sora将重塑影视创作等行业的游戏规则,以文生视频模式颠覆电影创作过程,可能今后会有若干人通过AIGC相关技术来生成电影。”浙江大学人工智能研究院院长、计算机与技术学院教授吴飞认为。

确实,去年拿到七项奥斯卡大奖的年度影片《瞬息全宇宙》就使用了AI视频,迪士尼更是完全用AI制作了《秘密入侵》的开头动画。就在2月10日晚,一场AI(人工智能)春晚上演,成为众多AI爱好者的盛宴。

但在杭州时空奇迹互娱科技有限公司CEO胡小凡看来,Sora只是一种工具,“它可以在短时间内把视频产品的产量拉出来,但影视级别的产品目前还是替代不了。”胡小凡判断,未来影视作品的发展方向是多维度多链路的,就像玩游戏一样,按照用户的不同想法推进、分化,最后的结果是不一样的,而Sora未来可以帮助影视公司降低这类游戏化影视作品的生产成本。

浙江省社会治理与传播创新研究院研究员宋哲认为,也许未来的影视创作是由编剧+大模型+算力供应商来完成,不再需要演员、摄像、导演,但实现起来还有很长的路要走。“最现实的限制条件就是算力。目前的算力成本对影视创作来说是极其昂贵的,虽然可能会有第一个吃螃蟹的人尝试用AI做一部电影出来,但是规模化应用于影视产业,我觉得短期内不会实现。”他说。

2月18日,一则“Sora首部AI电影共创计划”的共享文档开始在东阳横店的编剧圈子里流传。文档发布者称,将“共创世界第一部AI电影这一壮举”,现面向全社会征集剧本。网友评价:噱头十足。

现实中,乱象也随之而生。某电商平台,有人在售卖Sora教程,内容据悉90%是网上公开资料;尽管还未公

测,也有商家号称有“内测名额”;而“Sora让人赚到了‘第一桶金’”的消息也是不绝于“屏”……乱象背后,是亟需引起重视的投机与“套路”。

正视差距与奋起直追

“差距很大。”……相比ChatGPT刚出来时国内业界的群情激动,各种相关大模型迅速成“雨后春笋”之势,记者发现,Sora问世,目前几乎没有大厂跟进发声。记者的采访也被纷纷婉拒。

“从ChatGPT到Sora,人工智能发展的地区之间的代差,还是很明显的。”一位从业者说出了大家的心声。

但也有乐观者。业界公认,发展大模型离不开三要素:人才、算力和数据。三者我们与国外有差距,但并不是没有机会。

先看算力。这也是被业界公认OpenAI能超过谷歌这种“老牌大厂”的关键。“OpenAI对算力的投入很坚定。”曹宇观察。

前不久,OpenAI宣布启动“造芯”计划,希望筹得7万亿美元,相当于世界各国GDP加起来的10%。如此天文数字,可窥得OpenAI算力缺口的冰山一角。

在中国,算力等同煤电水等基础资源已日渐成为共识。不少专家分析,尽管目前国内顶级芯片相对缺乏,但总量及增量发展迅速。

2023年11月,工信部公布的一组数据显示,近年来,我国算力产业年增长率近30%,算力总规模位居全球第二。截至去年底,我国算力核心产业规模达

到1.8万亿元。

再看人才。记者查询OpenAI官网发现,Sora团队核心成员包括12人,其中有多位华人。值得注意的是,这支团队十分年轻,成立时间未超过1年。“OpenAI招了很多毕业生,这说明该企业的试错空间很大。我们也要鼓励类似的创业环境。”曹宇说。

相比较,我国差距比较大的是数据。“目前的研发范式,数据一方面看量,另一方面看标准,国际上视频解读标准以及软件基础设施基本都是英语为核心的技术语言底座,这两方面我们相对不足。”申永生说。确实,目前互联网上英语内容约占90%,其他语言才占10%。

不过,随着对新质生产力的进一步重视,我国的竞争力也在不断增强。

2月19日,国务院国资委召开“AI赋能产业焕新”中央企业人工智能专题推进会。会议强调,中央企业要把发展人工智能放在全局工作中统筹谋划,深入推进产业焕新,加快布局和发展智能产业。

前不久召开的杭州市“新春第一会”上,省委副书记、杭州市委书记刘捷两提Sora,提到杭州要将人工智能时代的算力成本降到最低。

目前来看,Sora有自己的局限性,或将成为各方追赶的机遇。

ChatGPT这样评价这位“小弟”：“Sora模型在模拟基本物理交互,如玻璃破碎等方面,不够精确。这可能是因为在训练数据中缺乏足够的这类物理事件的示例,或者模型无法充分学习和理解这些复杂物理过程的底层原理。”

“Sora还无法理解类似中国诗画的深意境。”浙江大学杭州国际科创中心求是科创学者张强说。中国绘画史上有个著名典故叫“深山藏古寺”。有人在山腰间画座古庙,半遮半露;有些只能让古寺露出小角。但最佳答案是,崇山峻岭之中有和尚在挑水。

机遇,还在于比拼的最终方向并不只是Sora,而是通用人工智能(AGI)。一般认为,AGI是指拥有与人类相当甚至超过人类智能的人工智能类型。去年,北京通用人工智能研究院院长朱松纯曾给读者写信道:“如果我们把实现通用人工智能看作‘探月工程’,大模型现在取得的成就如同登上了珠峰,虽然也是了不起的成就,但并不能实现终极目标。”

浙大研究院人工智能中心副主任赵俊博最近在朋友圈发文:“反对很多自媒体把这个技术类比在AGI上面,我们距离AGI还差得远……我觉得一个世界模型需要有能力去输出动作,输出对未来的预测,输出对当前所处状态的判断。Sora大概率是学到了一些世界运转的模式,但是否具备其他上述能力我们不知道。”

对此,业界界的观点也不尽相同。不过,大家都表达了追赶的急迫性。2月16日,360创始人周鸿祎在微博上写道:Sora的诞生意味着AGI的实现可能从10年缩短至一两年。张强在接受记者采访时也表示:“我们距离通用人工智能的目标虽然还有相当的距离,但正在以越来越快的速度来接近它。”



扫一扫,查看更多精彩内容

海南现“祖宗级”棉花

通讯员 周炜

大约20万年前,一粒始发于中美洲西海岸的棉花种子一路向西漂流,在中国的海南岛登陆后生根发芽、繁衍不息。直到2000-3000年前,海南岛的先民发现它的后代,开始了四倍体棉花的驯化栽培。

1月30日,浙江大学团队在Journal of Advanced Research《先进研究期刊》发表研究论文,文中还原了异源四倍体棉花这段数十万年的演化史。

研究团队从基因组水平解析陆地棉海南本土棉花一半野生棉种蓬蓬棉(G. purpurascens)的进化地位,发现中国是四倍体棉花最早驯化栽培的国家之一。

2019年,原新疆闰氏种业公司老总闫有发在海南崖城野外采集到一株棉花,计划用作亲本杂交育种实验。

当其他棉花含芳吐蕊,这株却迟迟没有动情。

“为什么不开花呢?”面对这棵身形略为魁梧的棉花,他们专门请教了浙江

大学张天真教授团队。

“这‘暴露’了它是真正野生的棉花。”课题组成员之一白雪莹教授说,“许多野生的棉花带有光周期敏感基因,花期受光照时间的牵制,一到北方就不开花。”

这个性状会影响棉花的产量与品质,因此当前全球大规模栽培的棉花,都是已经剔除了这个遗传特征的人工选育品种,更适合工业化大规模种植与采摘。

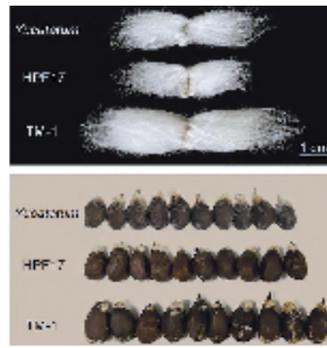
对长期从事作物遗传育种的张天真来说,野生品种是取之不尽的宝藏。它们一方面能帮助人类进一步了解作物在地球上的繁衍与驯化历史;另一方面,随着育种技术的不断发展,野生作物保留下来的各种遗传信息将会被进一步发掘。

为此,棉花研究团队将这株棉花的样本命名为HPP17,对其进行了全基因组测序。

基于现有的陆地棉全基因组重测序数据和11个陆地棉骨干亲本基因组,研究发现,以HPP17为代表的海南本土野生棉花(Hainan Island cotton, HIC)属



海南野生本土棉花及其纤维和种子。



浙大课题组供图

于陆地棉中较原始的蓬蓬棉野生种系。

“异源四倍体陆地棉原产于中美洲。其多年生野生种在美国南部经过人工选择,驯化成栽培的陆地棉或美棉,18世纪中期以后引种传播到世界各地广泛栽培。”张天真简要介绍了世界棉花的栽培

历史,“我国早期栽培的棉种是南亚地区起源的二倍体棉种亚洲棉,1949年以后大规模栽培的四倍体棉花品种也是美棉全球性传播带来的。”

但令人困惑的是,数千年前,我国海南的黎族先民就已经掌握了利用棉纤维

纺织的技术,所使用的是一种独特的本土棉花,这些本土蓬蓬棉是何时来到中国的,以及它在陆地棉中演化过程中的进化地位,一直未得到科学研究的确认。

通过全基因组测序,张天真团队确认了HIC的古老“身份”,它是20万年前原产于中美洲的异源四倍体陆地棉的“直系”后代。“我们的研究表明,早在20万年前,它就与原产于中美洲的异源四倍体陆地棉分家了。”张天真说。

它是如何到达中国海南的?研究团队认为,随着太平洋洋流,它从中美洲西海岸逐渐西进到达中国南海诸岛(海南岛)。

对此他们提供了两个证据。一是洋流分布,海南本土棉花的种子小而坚硬,带绒毛的棉籽能够在水中浮半年以上,目前,人们所知的代表性的野生蓬蓬棉在太平洋岛屿上的分布,与太平洋洋流吻合度高;二是它极强的耐盐性,能在盐水里存活半年以上。

海南岛登陆后,蓬蓬棉是“野生”状态,还是与人类的生产生活有交集?

帮研究人员揭开这一谜团的是手工轧花机——一种给棉花脱籽的机器。

“徒手将纤维从种子上剥离非常费力,这是发展先民棉纺技术遇到的第一个难题。”论文的第一作者程宇介绍。

在研究了不同棉籽的纤维程度后,研究团队进一步确认,只有蓬蓬棉的棉纤维才需要用到轧花机。“木棉的纤维非常短,无法取得用于纺织的纤维;而海岛棉和木棉,能够轻易脱籽。”从某种角度说,正是蓬蓬棉的精籽特征才催生了轧花机的诞生。

研究团队就此提出,在2000-3000年前左右,蓬蓬棉已经在海南岛被初步驯化和小规模栽培,并用于纺织“崖州布”。这个时间点,早在哥伦布时期前,美棉驯化栽培前。“三亚(崖州)可能是陆地棉最早驯化和栽培的地点之一,即中国是四倍体棉花最早驯化栽培的国家之一。”张天真说。