

热点聚焦

Sora诞生半年后,它的“挑战者”纷至沓来,连“等不及”又“追不上”的英伟达也亲自下场。

迄今为止,Sora依旧只释放小样、未开放使用,而快手可灵、智谱清影、Vidu已率先打开体验大门,走向大众。

尽管“一键生成”的初体验称不上完美,却搅动了内容产业的一池春水。身边不少微短剧、广告、动画已经开始用上AI这个“效率搭子”。人工智能生成技术,从前不久的文生图,到如今的文生视频、图生视频、视频生成视频,“AIGC宇宙”不断扩张。

AI,是华夏神话里的“神笔马良”吗?它能让多少想象力、创造力活起来、动起来?

■“文生视频”,如何“生”

“文生视频是一颗重磅炸弹。”半年来,从大厂到独角兽的Sora复现潮无一不在说明产业界对“生成”的重视。

视频生成,简言之是通过生成式人工智能技术,将文本、图片等多模态输入,转化为视频信号。

当前,视频生成的技术路线主要有两种。一种是扩散模型,其中又分为两类,一类是基于卷积神经网络的扩散模型,如Meta的EmuVideo、腾讯等推出的Video-Crafter;一类是基于Transformer架构的扩散模型,如OpenAI的Sora、快手的可灵AI、生数科技的Vidu等。另一种是自回归路线,如谷歌的VideoPoet、Phenaki等。

2024年7月26日,中国科技企业智谱AI面向全球用户发布其自主研发的人工智能生成视频模型清影(Ying)。

目前,基于Transformer架构的扩散模型是视频生成模型的主流选择,也称“DiT”(Di为Diffusion缩写,T为Transformer缩写)。

文本“扩散”为视频?“扩散在此指一种建模方式。”北京大学信息工程学院助理教授、博士生导师袁粒举了一个生动例子——

米开朗琪罗在凿刻著名的大卫雕像时,说过这样一句话:雕塑本来就在石头里,我只是把不要的部分去掉。“这句话很形象地形容了‘扩散’这一建模过程。原始的纯噪声视频好比未经雕琢的石块。如何敲打这个大石块,敲除多余的部分,直到把它敲成轮廓清晰的‘大卫’,这样的方式就是‘扩散’。”袁粒说。

袁粒进一步解释:“Transformer就是一个神经网络,遵从‘规模规则’,执行敲石头的过程。它能处理输入的时空信息,通过理解其内部复杂关系来理解现实世界,使模型具备推理能力,既能捕捉视频帧之间的细微联系,也能确保视觉上的连贯、时间上的流畅。”

■“效率搭子”,有多快

一只憨态可掬的北极熊被闹钟叫醒,背起行囊,乘坐直升机,转乘高铁、换乘出租车、登上轮船,跨越山河湖海、历经艰难

人工智能时代来了。过去、现在、未来,AI快速改变我们的生活与工作。快!有多快?AI以涌现之力掀起浪潮,迭代、突破、颠覆,一浪高过一浪!

从“图灵测试”到“深蓝”系统胜利,再到GPT系列模型从1、2、3、3.5直抵ChatGPT系列……AI通过模仿、对齐、学习像人一样说话、思考,甚至有分析称,最接近“人”的AI大模型,智商将达到145,碾压多数人。

这样的变革,带来的不仅是技术突破,还有对未来的焦虑——算法歧视、隐私外泄、虚假信息泛滥以及失业忧虑等。而这样的涌现,远没有达到终点。越来越多的人开始担心:当人工智能无限接近乃至超过人,人会不会被替代?人类未来的出路在哪儿?

剑桥大学物理化学博士刘煜炎,近期在社交媒体上发布了一条视频:“我可能要面临失业了。从教40年,今天感觉到我的工作、我的未来,被人工智能严重挑战了。”

刘煜炎说,作为老师,他一直用科学家的方法教授并培养学生思考,但没有想到ChatGPT的出现,让他忧心忡忡。人工智能分分钟就掌握了他经过40多年才练就的“功力”。这意味着,家长将无需再找一对一老师或是培训机构,人工智能完全可以胜任,而且能教各个科目。那么,真实世界里的老师要去哪里找工作?他担心会面临失业。

同样,近期无人驾驶网约车服务“萝卜快跑”试运行,也引发广泛关注。有舆论称,伴随无人驾驶的大面积普及,未来网约车、出租车可能多数是无人驾驶,相关人员面临就业挑战。

可以看到,随着“百模大战”(各类“大規模深度学习模型”在应用领域竞相发展的态势)如火如荼,各类AI应用将在真实商业场景中加速落地。



险阻,终于到达南极,与企鹅相会……

这部时长1分半、名为《一路向南》的动画短片,由视频生成模型Vidu完成。原本1个月的工作量,有了AI这个“效率搭子”的加入,仅用1周时间就制作出精良作品——效率是过去的4倍。

这让北京电影节AIGC短片单元最佳影片得主、Ainimate Lab AI负责人陈刘芳心生感慨:视频生成技术,让高水平动画不再是大厂才敢玩的“烧钱游戏”。

AI动画《一路向南》的创作团队仅由3人构成:一名导演、一名故事版艺术家、一名AIGC技术应用专家。而以传统流程制作的话,需要20人。算下来,仅制作成本就降低90%以上。

正如快手视觉生成和互动中心负责人

万鹏飞所言,视频生成的本质是从目标分布中采样计算得到像素。这种方式能以更低的成本,达到更高的内容自由度。

进入Vidu的视频生成页面,笔者也体验了一把“一键生成”的自由。上传一张照片设置为“起始帧”或作为“参考的人物角色”,在对话框里输入想要生成的场景的文字描绘,点击“生成”,一条灵动精彩的短视频就自动生成了。从进入页面到下载完毕,不足1分钟。

将一张图片发给国产视频大模型Vidu,一段动画视频随即自动生成。

“‘人人成为设计师’‘人人成为导演’的时代将会到来,就像当年‘人人拥有麦克风’一般。”智谱AI首席执行官张鹏说。

■“世界模拟器”,有戏吗

视频生成,颠覆的仅仅是内容产业吗?这显然不是OpenAI的初衷。“生成视频”只是一道“开胃菜”。

Sora诞生之前,OpenAI并未将其定位为AIGC的实现工具,而是复刻物理世界的“容器”——世界模拟器。这一容器里,运行着真实世界的物理规律、环境行为、交互逻辑,恰似《黑客帝国》描绘的虚拟世界,冲击着我们的想象与感官。

然而,物理世界是三维的,目前的Sora等模型还只是基于二维运作,并非真实物理引擎,也就谈不到深层次的物理世界模拟。

“多年来,我一直表示,‘看到’世界即为‘理解’世界。但是现在我愿意将这个概念推进一步,‘看到’不仅仅是为了‘理解’,而是为了‘做到’。”斯坦福大学讲席教授李飞飞公开表示,空间智能的底线是将“看到”和“做到”联系在一起,有一天,AI将会做到这一点。

当“看到”还不等于“做到”时,人工智能的创造就不能停。最近,又有新的技术路线出现了。不同路线之间你追我赶,共同向前,推进这个由向量与模型构造的智能世界。

未来的“世界观”,依旧是一道尚未揭晓的谜题。正如美国物理学家费曼所说:“我不能创造一个我不理解的世界。”但这并不意味着,理解了一个世界,就一定能够创造出一个世界。

此刻,依旧是颠覆到来的前夜。这就是为什么当我们向技术探索者抛出关于未来的问题时,会得到截然不同的答案。也许“不确定”,正是这个时代的幸事。

(新华社9月10日电)



卷技术、卷资金、卷市场、卷人才……不难想象,在这样的趋势下,科技突飞猛进,AI将在竞争格局下“卷”出新高度。在2024年世界人工智能大会上,智能感知表现突出:用“脑电”技术感知人的意图,辅助人类做康复训练;国产机器人“青龙”增添了手指“触觉”,可以像真人一样抓取食物;利用AI语音声库,让世界首位中文虚拟歌手洛天依拥有堪比歌手的音色,让AI听起来不再是AI……人工智能在应用领域的各个赛道,想象空间愈发广阔。

世界知识产权组织发布的报告显示,2014年至2023年,中国生成式人工智能专利申请量超3.8万件,居世界第一。

从国内发展看,科技革命之光已经照进现实。AI应用不再局限于聊天、画画、写作,而是立足更复杂的生活与工作场景:自动驾驶出租车已经上路,自动送货机器人整装待发,家庭服务机器人不再是梦……似乎所有工具性的劳动都有可能被AI取代。

当人工智能被“卷”得越来越像人,越来越多的人自我调侃被“卷”成“工具人”。一旦被“卷”成“工具人”,意味着996、007,意

成为我们的知识库、帮手、助理。利用AI的“超级能力”提高学习与工作效率,我们可以从工具性、机械性事务中解脱出来,腾出精力去做更有价值的事。

AI终归是为人类福祉服务的,而不是背道而驰。当人工智能越来越像人,人自己是不是应该更像人,而不是更像工具或AI?比如,机器人固然能作出像模像样的一般性诗作,在词汇、格律、技法等浅层次似乎可以无限接近人,但基于人类本质力量的想象力、情感力、思悟力以及创作个性、风格等深层次东西,则是AI不那么容易企及的,而这些恰恰最能体现出古往今来一个真正诗人孜孜开掘并创造性表达的精神维度。

从另一个角度看,如果每个人都活得像AI一样,投喂、训练AI的素材也会越来越少,AI会停止进步、迭代,直至消亡。

AI再发达,也是要达成人和人、人和物的高效协同。人工智能的核心还是“人”。人工智能的蓬勃发展,不应取代人而应解放人,让我们活得更像人本身,更能激发出人之为人的本质力量。

(新华社9月12日电)

12秒下线一块组件
AI“担纲”光伏制造大脑

新华社北京9月19日电 智能制造给光伏行业带来了什么?作为目前全球光伏行业中唯一的“灯塔工厂”,隆基绿能嘉兴基地给出了具象化的答案。

这座自动化率超过90%的工厂,有着51条智能化生产线,每隔12秒就会有一块完整的光伏组件下线。隆基绿能嘉兴基地灯塔工厂通过实施超过30项数字化用例,让AI成为生产管理的大脑,实现产品品质可靠性提升43%,产品生产交付周期缩短84%,基地整体单位能耗降低20%。

穿戴好无尘车间专用鞋套和网帽,记者走进隆基绿能嘉兴基地一期工厂。在位于生产车间一侧的数字化运营管理中心内,数字大屏实时展示着工厂运营管理过程中的订单交付、生产进度、产品质量、设备运行、人力状况、能耗、安全等数据。

“隆基绿能嘉兴基地主要生产隆基BC(背接触电池技术)系列高效太阳能电池组件,整体组件产能超35GW。每隔12秒,就会有一块完整的光伏组件下线,其中包含4道检测程序。”隆基绿能嘉兴基地智能化项目负责人杜国祥说。

电池片的隐裂用肉眼极难察觉,传统条码追溯不适用、虚拟码追溯准确率低一直是困扰光伏组件生产的难题。如今,嘉兴基地创新应用AI精准追溯技术,在生产过程中每18秒就可以判断出12串组件是否有缺陷。

“缺陷是哪条流水线哪个机台生产出来的,都可以进行识别和追溯。这一技术不仅能保证整个生产过程的高效率,还提高了对客户响应速度。”杜国祥指着工厂中央智能复判室里的电脑屏幕说,发现异常后,AI会快速生成解决方案发送给屏幕前负责监控的质检人员,如果质检人员能够给出不同于AI的新解决方案,也会被AI存储学习。

这仅仅是隆基绿能嘉兴基地数字化、智能化的一个缩影。在这里,通过大规模采用工业互联网、大数据、人工智能、数字孪生等新技术,30余项数字化用例已成功实施。其中包括五大核心“黑科技”,即机器视觉赋能的柔性自动化、AI赋能的全流程检测及追溯、订单生产交付周期智慧管理、AI算法赋能的电池资源匹配及动态纠偏、智能人力管理。

近年来,光伏行业作为我国经济发展的新动能不断壮大,产业链各环节产量和装机规模连续多年保持全球第一,并成为“新三样”产品出口的代表。在阶段性供需错配、同质化竞争加剧的当下,推动技术升级和降本增效至关重要,数智化深度转型成为打造高质量发展的重要方向。

“据我们统计,通过智能化技术的应用,2016年至2023年我国多晶硅、电池和组件的人均产出率出现了3至4倍的增长。”中国光伏行业协会副秘书长江华表示,智能化技术的应用不仅带来了生产效率的提升,还推进了光伏与半导体产业、数字化技术发展的耦合。

(《经济参考报》记者王璐)

上万AI“质检员”
把关汽车“智造”全链条

新华社北京9月20日电 全自动无人化的导向车(AGV)将数千种零部件准确送往生产的各个环节;AI加持的智能制造系统如同上万个“在线”质检员实时检测关键工艺;大模型持续积累生产数据以优化各个环节……智能化正在帮助理想汽车实现更高精度和更高效率的生产。

在理想汽车北京绿色智能制造基地,走进高度自动化的焊装车间,记者看到,不少关键性的焊装环节旁,坐着一些着装稍有不同的员工,他们正实时关注着电脑上显示的数据。

“这些员工主要负责的是基于生产过程中不断积累的数据,持续完善自研的大模型生产系统。”理想汽车北京绿色智能制造基地负责人王垒告诉记者:“这正是我们走向智能化的关键,也是理想北京工厂走在汽车智能制造最前沿的底气。”

“例如在合车环节,需要将数十颗螺栓拧紧,把底盘和车身连接起来。以往可能只对螺栓最终是否足够紧密这一结果进行检测,出现问题排除起来十分复杂。我们的连山数据监控平台系统可以通过对螺栓拧紧的全过程进行检测,并利用AI技术与正常的扭力变化曲线进行对比,更精准更高效地定位问题环节。”王垒说:“有了‘连山平台’的加持,我们的产线又多出来了上万个专业的AI质检员,不知疲倦地去对每台车的每个工艺细节进行实时的‘在线CT’,这在国内乃至世界的汽车制造领域,都是最前沿的技术。”

“未来,我们不仅可以实现让计算机回望任意时刻、任意场景的微观产品和整体生产数据,甚至可以让计算机去预测用户的潜在需求和产品的潜在风险,让制造业真正拥抱AI。”王垒表示。

据透露,基于连山系统等智能制造手段,北京基地新工厂实施周期缩短3个月,生产效率提高20%以上,运营成本降低25%,能源利用提高11.3%。

智能制造正在深度参与汽车产业的发展。中国汽车工业协会副总工程师许海东表示,汽车行业的自动化水平一直走在各个产业的前列,正大量应用“数字孪生技术”来实现智能制造。

国办印发的《新能源汽车产业发展规划(2021—2035年)》提出,加快新能源汽车智能制造仿真、管理、控制等核心工业软件开发和集成,开展智能工厂、数字化车间应用示范。

对于如何将智能制造进一步加以推广落地,理想汽车相关负责人表示,工厂的作用不仅仅是制造产品,在智能化加持下,工厂成为不断成长迭代的有机体。“未来的生产制造中,我们将围绕核心系统全栈自研、高度自动化、数字孪生、AI赋能及智慧物流等方面,高效应用,持续迭代。”

“随着数字化技术、智能化技术的发展,汽车行业的智能制造将持续提升,无人工厂和‘黑灯工厂’这样的标杆项目会越来越多。”许海东说。

(《经济参考报》记者袁小康)

新知