

大模型企业闯进“纯视觉时代”

▶ 本报记者 李洋

视频生成有望成为通用知识学习方法,在现实世界充当“人工大脑”。近日,豆包大模型团队宣布其携手北京交通大学与中国科学技术大学共同研发的视觉生成实验模型“VideoWorld”正式开源,首次实现无需依赖语言模型即可认知世界。

民生证券发布研报称,豆包开源视频大模型“VideoWorld”的发布,让视频生成成为通用知识学习方法,在现实世界充当“人工大脑”,仅靠“视觉”即可学习知识,并“理解”因果关系,“预测”未来,有望激活“视觉市场”为其打开增长空间。

仅靠“视觉”即可预测未来

目前,现有模型大多依赖语言或标签数据学习知识,很少涉及纯视觉信号的学习。然而,语言作为一种表达工具,其局限性

在于无法全面捕捉真实世界中的所有复杂信息。

据了解,VideoWorld摒弃语言模型,通过纯视觉信号进行统一的理解、执行和推理。同时,VideoWorld基于一种潜在动态模型,可高效压缩视频连续帧间的变化信息,显著提升知识学习效率和效果。

“视觉模型在处理连续帧变化时能够捕捉动态信息。这在理解动作序列、预测未来状态等方面有独特优势。”资深人工智能专家郭涛认为,这类模型能够直接处理和理解图像或视频数据,不需要先转换成文本描述再进行理解,因此在图像识别、场景理解方面更为直观和准确。此外,视觉信息本身含有丰富的细节和上下文,可为自动驾驶过程中的实时路况判断、医疗影像的精确分析等特殊场景提供更

精细的判断依据。

“这种视频生成模型不依赖语言模型,进一步减少了对数据标注和预训练的依赖,降低了开发成本和难度。”北京市社会科学院副研究员王鹏说。

“这无疑大大提升了模型的学习效率;此外,这种大模型不需要中间的语言转换,更符合人类的直观认知方式。”科方得智库研究负责人张新原认为。

为AI应用带来新的可能性

业界专家表示,VideoWorld的发布,也为人工智能领域带来新的可能性。比如,在医疗影像分析中,模型能够通过学习大量医学影像数据,辅助医生进行诊断和治疗;在娱乐产业,可用于电影特效制作、虚拟偶像生成等;在教育领域,能够创建互动式学习材料;在零售行业则可应用于商品

展示和虚拟现实购物体验等。

民生证券发布的研究报告认为,视觉认知世界的能力有望激活“视觉市场”。

“这些应用场景只是‘冰山一角’,随着技术进步,更多创意应用将会涌现。”郭涛说。

依赖文本大模型有危机

“随着多模态技术的发展,结合文本和图像的综合模型将成为趋势,单一依赖文本的大模型可能会逐渐失去竞争力。”在郭涛看来,这并不意味着语言大模型会被淘汰,而是促使其向更高级、更综合的方向发展,比如增强对图像的理解能力或与视觉模型更好地协同工作。

此次豆包大模型团队选择开源VideoWorld,或许希望通过此举让更多的研究者和开发者可以接触并使用VideoWorld,以

推动纯视觉认知技术的快速发展和应用。从长远看,作为首个无需依赖语言模型的视觉认知模型,VideoWorld的开源有助于其成为行业标准,影响未来相关技术的发展方向。

VideoWorld的开源,为研究者提供了一个强大的基准模型,有助于探索更多基于视觉的学习方法和算法。“随着更多开发者的加入,有望围绕VideoWorld形成一个新的技术社区和生态系统,促进不同领域间交叉合作和技术融合,进一步拓展人工智能的应用边界。”郭涛说。

关于开源面临的一些挑战,业界人士认为,开源使得竞争对手可以轻松获取技术细节,可能加剧市场竞争。当然还有社区管理方面的挑战,维护一个活跃且健康的开源社区需要投入大量资源,包括技术支持、文档维护等。

我国AI大模型市场 预计明年突破700亿元

▶ 王查娜

近期,艾媒咨询发布《2024-2025年中国AI大模型市场现状及发展趋势研究报告》(以下简称“《报告》”)。《报告》预计,2026年中国AI大模型市场规模将突破700亿元。中国AI大模型行业正处于爆发式发展阶段。

当前,在人工智能的浪潮中,AI大模型技术以其强大的数据处理能力和广泛的应用前景,成为推动社会进步和产业升级的重要力量。

AI大模型正重塑行业生态

《报告》显示,2024年,中国AI大模型市场规模约为294.16亿元。AI大模型以其泛化性、通用性和涌现性三大特征,展现出强大的数据处理和任务执行能力。这些模型能够从海量数据中学习规律,适应多种任务,并在训练后展现出新的能力和行为。

“人工智能的底层技术和重要底座是通用大模型,只有优先发展本土通用大模型,才能培育具有本土特色的人工智能大模型体系。”麒麟合盛集团创始人李涛近日接受本报采访时强调了AI大模型的重要性。

2024年,众多行业加速探索并集成AI大模型能力,使其在研发生产、个性化学习、辅助诊断、行程规划、智能办公等方面发挥关键作用。同时,在金融、教育、

医疗健康、工业制造、传媒娱乐、公共交通等场景,大模型正逐步重塑行业生态与价值链。目前,华为、百度、月之暗面等企业在AI大模型领域的表现尤为突出。华为的盘古大模型、百度的文心一言以及月之暗面的Kimi等产品,在技术研发和商业应用上展现了各自的特色和优势。

据了解,按部署方式,AI大模型可分为云侧和端侧两大类。云侧模型部署在云端,拥有更大的参数和算力资源;端侧模型则运行在终端设备上,如手机和PC,具有隐私保护和本地化运行的优势。目前,AI大模型已在金融、互联网办公、电商等多个行业得到应用,通过数据训练和智能分析,为行业提供实时监测、风险预警和决策支持等服务。

《报告》指出,AI大模型企业发展的驱动力包括3个方面,一是政策支持。国家和地方政府的政策支持是推动AI大模型产业发展的重要因素。例如,“十四五”规划中提出要加强数字基础设施建设,为AI大模型的发展提供政策保障。二是算力提升。云计算市场的快速发展带动了算力的迅猛增长,为AI大模型的训练和应用提供强大的技术支持。三是数字经济的推动。中国数字经济规模的不断扩大,为企业数字化转型提供了



近日,2025年苏州市“人工智能+”创新发展推进大会暨人工智能赋能新型工业化深度行(苏州站)活动举行。本次活动以“共启AI+新质赋未来”为主题,活动现场,人形机器人、智能控制无人系统、AI大模型互动虚拟产品等百余件展品亮相,集中展示苏州人工智能领域发展情况,吸引众多观众前来参观。

图为活动现场展示的脑机接口智能装备。

新华社记者 李博/摄

广阔的市场空间,也为AI大模型的发展带来了新的机遇。

仍有诸多问题待解决

《报告》也指出了我国AI大模型的不足之处:与国外先进企业相比,国产大模型在训练数据集、算力支持和应用场景等方面仍有差距。技术缺陷仍是AI大模型面临的主要挑战,特别是在复杂逻辑任务的处理上。

在技术层面,随着模型规模

及复杂度的递增,算力受限、数据短缺、模型性能待提升等问题凸显。同时,高端算力设备供应不足,数据供给遭遇困境,端侧大模型融合进展缓慢。此外,高端算力芯片如GPU(图形处理器)、高速存储芯片供应匮乏,极大束缚了行业的自主发展。

在人才层面,顶尖人才和复合型人才短缺问题制约研发创新。大模型领域的快速发展导致专业人才供不应求,我国人工

智能顶尖算法人才、芯片人才严重不足,没有形成顶尖人才聚集的“马太效应”。同时,跨领域复合型人才短缺,与发达国家存在明显差距,限制了大模型的研发和创新速度。

在监管层面,政策滞后与内容风险引发担忧,AI大模型生成的虚假误导内容引发伦理与合规问题的讨论,行业亟待通过完善法规、明晰标准等举措突破困境。