

让人形机器人
走进“360行”具身智能成风口
竞争关键在“小脑”

本报记者 马亚宁 叶薇

焦点

背上时空智算背包,机器人能给人指路,还能在多类场景稳定工作、完成任务;国内首个由 VLA 模型 ERA-42 驱动的星动 L7,既能“整活”,表演街舞、调酒、360°旋转;还能干活,在生产线上智能分拣、搬运扫码,到科研实验室里采集数据,精准打螺丝……本届世界人工智能大会上,以人形机器人、机器狗等为代表的新一代具身机器人,正在从封闭式专用系统向通用化智能平台、从单一限制场景向复杂全域空间加速转型。

从实验场到应用场,由“舞台秀”转战“360行”,人形机器人真正实现具身智能,还有多长的路要走?

目前大模型能模仿的动作仅200个,远不及人类手指40个关节能完成的操作

今年,人形机器人似乎一下子就火了。上半年,一场妙趣横生的人形机器人半马赛,让人形机器人的跑动能力“捧”出圈;这两天,2025世界人工智能大会现场,人形机器人阵容强大,让人类有种被机器人包围的感觉;而且借力深度学习、生成式大模型,人形机器人的“大脑”也越来越好用。

“我们已经站上了具身智能爆发的临界点。”在人形机器人与具身智能论坛的圆桌环节,世界数字科学院国际首席人工智能官杜兰说,人形机器人被视为具身智能的关键载体之一。从人形机器人跑马拉松,到机器人格斗赛,人形机器人的边界已经突破了实验室。它通过“大脑、小脑、身体”三部分协同,展现了国产机器人在机械结构、运动能力、量产成本控制等方面的出色表现。

不过,AI能写诗,却很难稳稳地拿起一只玻璃杯。要让人形机器人一直保持“身子稳、大脑灵”这个人类都难以一直在线的状态,还需要翻越技术创新的“珠穆朗玛峰”。特别是物理世界的数据采集成本极高,且多样性严重不足,让其数据饥渴严重。正如行业专家们所观察到的,训练人形机器人完成简单的洗衣动作,特斯拉就需投入20亿美元。即便如此,大模型能模仿的动作仅200个,远不及人类手指40个关节所能完成的复杂操作。

“机器人在不同场景中的稳定性控制,需要海量数据支撑,但现实中室内空间数据的复杂性和多样性获取难度极大”,这导致机器人在陌生环境中往往“水土不服”。更关键的是,不同实验室的数据缺乏兼容性,难以形成合力,即便有开源模型,也很少有人能清晰说明其风险与验证方法。而且,人形机器人能耗极高,完成相同任务的能耗是人类的5倍,且平衡能力薄弱。

智澄AI的联合创始人及首席战略合作官徐明强指出:“若机器人在亚毫米精度控制上失误摔坏,机械臂损坏将导致前期数据训练白费。”这使得家庭服务等场景的落地难度远超预期,例如,训练机器人做家务时,机器人打碎8个碟子后,竟围着碎片“跳科目三”。因此,在业内人士看来,训练人形机器人做家务以期替代家政阿姨,要比AI深度学习写医生、律师的文章,难得多。

在这个不断迸发科技火焰的蓝色星球上,当AI与机器人在山巅相遇,“具身智能”开始向人类亲切招手。“大家好,我是青龙pro,面向全场景应用的人形通用智能机器人。”在上周末两天举行的2025世界人工智能大会上,身高185厘米,具备自主行走、语音教育与灵巧操作能力的人形机器人青龙3.0公版机系列,集体亮相人形机器人与具身智能论坛。它骄傲地说:“未来的智能世界,我准备好了!”



▲ 行走的机器人成为观众追逐的焦点

◀ 从事装配工作的机器人

本报记者 陶磊 摄

从走跑跳到能干活,具身智能需要打通大脑、小脑、肢体的技术路径

面对技术难题的层层包裹,以人形机器人为代表的具身智能,在我国却成为科技竞争和科技创业的火热赛道。近一年来,中国人形机器人产品不断取得创新突破,从站得住、走得稳、跑得快,到进入街区、工厂、商务楼等实验性应用场景。在资本和政策的加持下,人形机器人技术创新者们纷纷走上台前,人形机器人迎难而上,迎来从实验室迈入规模化量产和商业应用化的井喷期。

走进2025世界人工智能大会的“WAIC·技能大舞台”,具身智能在民生服务领域的落地地图徐徐展开。偌大的展品“应用场”里,大模型技术已全面进入机器人的“大脑”(决策)和“小脑”(控制)。例如,在智造车间区域,青龙机器人正进行自动化物料分拣。其核心是基于600万真机数据和VLA预训练,形成的2.8B操作大模型。

更令人瞩目的是,国家地方共建人形机器人创新中心和浦江实验室团队预计年底前将完成2500万整机数据积累,规模达谷歌同类数据集的10倍。通过扩大数据规模,显著提升了机器人的泛化能力和涌现能力,使其能够基本满足核电、汽车、3C等复杂工业场景的作业要求。另一个是上核管孔清理机器人,它可以在管板直

径达5米、近两万多个管孔中完成清洁工作,高效解决了劳动工人长时间弯腰重复作业的问题。

在邻里文艺站才艺表演区,智元新创远征系列机器人正展示其“听得懂人话,敲得响乐鼓,写得了书法”的多才多艺。这得益于“具身智能+核心零部件”的创新设计思想。中国的人形机器人成功实现了用不到美国Atlas、日本Asimo等顶级机器人1/100的成本,达到相当的性能水平。这条既要低成本、又要高智能的道路,充分结合了中国制造业优势和雄厚的具身智能技术基础,展现出独特的“中国路径”。

此外,云百生剥鹌鹑蛋机器人、希夕做饭机器人、擎朗智能服务机器人、复旦抓豆腐机器人、新松修理机器人、银河通用销售机器人等,体现了“一脑多形”的通用具身大脑正从想象变为现实。这些形态各异的机器人,通过共享先进的VLA具身智能模型,同样能实现高效、精准的作业,让普通观众更直观地理解具身智能的内涵,预示着具身智能赋能的“智造”时代即将全面来临。

国家地方共建人形机器人创新中心首席科学家江磊认为,机器人能走、跑、跳,并不是炫技,是对人形机器人控制水平硬件平台的一个综合考验。下一步,是从走跑跳到能干活,竞争的关键在于“小脑”。要让机器人能干活,要解决它的感知局限、决策断层、泛化瓶颈。具身智能不光是加在大脑,不光加在小脑,还要加

在肢体上,需要打通大脑、小脑、肢体的具身智能技术路径。

从整个产业链来看,具身智能赋能的机器人,已经发生了明显的改变,从过去的老四大件——感知、控制、驱动、传动,发展到今天新四大件——高端芯片、动力电池、高性能材料以及连接器和标准器组件,还有新四大组件——感知头、灵巧手、机械臂、电子皮肤,还有很多挑战。“我们做人形机器人的,特别希望直接能用到感知头,而不是用视觉传感器;用到灵巧手,而不是直接用应变片的传感器,我们能不能为人形机器人打造出新的组件?具身智能是有身体的AI,我们不当只关注AI而忽略身体,要软硬一体化打造。我们将围绕标准、创新、开源跟生态这四方面来打造。”

具身智能,未来像人类一样理解物理规律、优化动作策略、积累操作常识

当人工智能技术革命进入深海时,以人形机器人为代表的具身智能,自然浮出水面。在清华大学计算机科学与技术系教授孙富春看来,具身智能与传统的

离身智能不同,离身智能体不具备物理形态,只能被动接受人类采集制作好的数据,仅仅停留在“纸上谈兵”的阶段,就如同ChatGPT等大模型,虽能处理文字、与用户聊天输出信息,却无法直接作用于物理世界。而具身智能则是赋予机器物理身体或形式,使其能够与物理世界交互,例如家用服务机器人、无人车等。

“人形机器人一定是未来具身智能本体的最主要部分,而它将来也必然会进入我们的工作和生活

中。例如,制造业的生产车间都是按人体结构设计的,人形机器人要比机械手、机械臂,有更大作为。”孙富春教授指出,具身智能的核心任务是打造通用智能体,让人形机器人不再局限于执行单一任务,而是具备完成多种任务的能力,真正成为人类的得力帮手。

他坦言,目前引人注目的人形机器人还只是第一代,能跑会跳,但没有智能。未来,人形机器人的人体计算机会考虑肌肉力的分布,神经功能的尝试性增强,还能像人类一样去学习,通过模仿训练、知识迁移、理解感知、联想互动等,不断自我提升、自我完善。例如,在认知阶段,机器人可借助视觉、触觉捕捉人类动作细节,结合能力反馈实时纠偏,实现带反馈的模仿学习;在自主阶段,也就是人类学习的高阶阶段,机器人通过在不同场景中反复学习,借助世界模型模拟物理环境,让感知与行为持续优化,最终可以像人类一样灵活应对复杂任务,实现具身智能质的跨越。

无论是科学界还是产业界,创新者们相信,人形机器人的终极目标,是融入人类设计的物理世界——工厂车间的操作台、家庭的厨房、办公室的桌椅,这些按人类尺度设计的环境,注定是人形机器人的主场。当它们能像人类一样理解物理规律、优化动作策略、积累操作常识时,不仅会革新制造业(替代重复劳动),更会在服务业、医疗、救灾等领域释放巨大潜力。

从人类百万年的进化到机器人30年的追赶,本质是一场“向自然学习”的革命。人形机器人的每一步突破,都是对具身智能的重新定义和突破——它不仅是机械与算法的结合,更是对人类行为、认知与智慧的理解与尊重。

未来已来,而这场“让机器像人一样思考与行动”的创新探索,才刚刚开始。