

具身智能技术演进、工业应用 Embodied Intelligence Technology Evolution, Industrial Application and Future Outlook

实践与未来展望

王晓思,林家宁,白琳(中国联合网络通信集团有限公司,北京 100033)

Wang Xiaosi, Lin Jianing, Bai Lin (China United Network Communications Group Co., Ltd., Beijing 100033, China)

摘要:

具身智能作为人工智能领域中的前沿方向,正从技术验证走向实际应用,展现出产业革新潜力。阐述了具身智能的核心概念与技术架构,探讨了其在多领域的应用案例和边界,特别是工业应用实践、挑战和突破点;同时关注其发展所面临的关键问题与挑战,展望了未来可能的演进路径。此研究表明,具身智能不仅是技术发展的必然趋势,更是推动社会智能化转型的重要力量,将在全球范围内掀起新一轮科技革命,为全球高质量发展注入强劲动能。

关键词:

具身智能;技术架构;应用案例;发展挑战;未来趋势

doi: 10.12045/j.issn.1007-3043.2025.07.006

文章编号: 1007-3043(2025)07-0035-06

中图分类号: TP18

文献标识码: A

开放科学(资源服务)标识码(OSID):



Abstract:

As a cutting-edge forefront in the field of Artificial Intelligence, Embodied Intelligence is transitioning from technological validation to practical applications, demonstrating strong potential for industrial transformation. It states the core concepts of Embodied Intelligence and outlines its technical architecture. It introduces application cases and boundaries across multiple domains, with a particular focus on industrial practices, related challenges, and potential breakthroughs. Meanwhile, it focuses on the key issues and challenges in its development and forecasts possible future development trends. This study shows that Embodied Intelligence is not only an inevitable trend in technological development but also a driver of more intelligent society. It will trigger a new round of scientific and technological revolution globally, injecting strong momentum into global high-quality development.

Keywords:

Embodied intelligence; Technical architecture; Application cases; Development challenges; Future trends

引用格式: 王晓思,林家宁,白琳. 具身智能技术演进、工业应用实践与未来展望[J]. 邮电设计技术, 2025(7): 35-40.

0 引言

在科技飞速发展的时代,我们有理由相信,人工智能(Artificial Intelligence)会加速推动各行业特别是传统行业的颠覆性变革和提升。具身智能作为人工智能领域的前沿方向,正从理论研究阶段迈向实际应用,开启人机深度协同的新纪元^[1]。

2023年以来,具身智能在技术突破和应用拓展方面取得了显著进展。计算机视觉中激光雷达、深度相

机等感知技术的进步,使具身物理本体能够实现毫米级环境感知^[2];强化学习、模仿学习、自适应控制等人工智能算法的发展,让具身智能在虚拟仿真环境中通过大量试错优化策略,并将经验无缝迁移至真实场景^[3]; Vision-Language-Action 大模型的出现,进一步增强了具身智能的泛化能力,使其能够适应不同的应用场景^[4];分布式协同技术则构建了多智能体协作网络,提升了群体智能的规模效应^[5]。

随着技术的不断成熟,具身智能的不同初级形态已广泛应用于工业制造、医疗健康、自动驾驶、家庭服务等多个领域,为各行业的智能化升级提供了强大动

收稿日期: 2025-05-28

力^[6]。尽管具身智能得到了一定程度的应用,但现阶段并没有实现技术的全面突破,也没有形成规模化市场,其发展仍面临诸多挑战。数据获取与仿真鸿沟、硬件限制、软件生态标准化缺失以及伦理与安全等问题,制约着具身智能的进一步发展和广泛应用^[6]。因此,深入研究具身智能的技术架构、应用案例、发展挑战及未来趋势,具有重要的理论和现实意义^[7]。

1 具身智能的核心概念

1.1 具身智能的定义与内涵

具身智能是指具有物理本体的人工智能 agent(如机器人、无人机、智能汽车等)通过物理实体与环境实时交互,实现感知、认知、决策和行动一体化的智能系统^[8]。其核心理念在于强调智能的本质必须通过物理实体与环境的动态互动来塑造和体现,突破了传统人工智能仅依赖符号推理和虚拟计算的“离身性”局限。具身智能的智能体能够利用自身的传感器感知周围环境的信息,通过对这些信息的理解和分析做出决策,并通过执行器执行相应的动作,从而实现与环境的有效交互并完成任务^[7]。

1.2 具身智能与传统人工智能的区别

传统人工智能主要基于符号主义和联结主义,侧重于通过算法和模型对数据进行处理和分析,以实现对问题的求解和决策;这些处理和分析都是在计算机虚拟空间进行的,解决的问题也集中在抽象和虚拟问题^[9]。然而,这种方式往往缺乏与真实物理世界的直接交互,难以应对复杂多变的现实环境。相比之下,具身智能的优势在于其能够让人工智能通过物理实体与环境进行实时交互,在交互过程中不断学习和适应环境变化,从而具备更强的环境感知能力、决策能力和行为能力^[10]。具身智能能够将感知、认知和行动紧密结合,形成一个闭环的智能系统,使智能体能够更加灵活、智能地应对各种实际问题^[7]。

1.3 具身智能与传统机器人、工业机器人的区别

具身智能的起源可追溯至20世纪中叶,计算机科学之父 Alan Turing 在《计算机与智能》中首次提出“机器能否思考?”的哲学命题,预示了具有物理本体的人工智能 agent 通过物理交互实现认知的可能性^[11]。然而,受限于当时的技术条件,这一构想在随后数十年里未能取得实质性突破。MIT 前计算机科学和人工智能系主任,前 iRobot 公司创始人和首席执行官,机器人学家 Rodney Brooks 推动了机器人行为主义方法论

的发展,主张智能应由身体与环境的实时交互自然涌现,确立了“感知—行动”闭环系统的可行性^[12],成为具身智能的奠基性理念。自2023年美国斯坦福大学 Mobile-ALOHA 家务机器人横空出世之后,具身智能的前沿概念开始被大众所知晓^[13]。本次具身智能的浪潮,是以人形机器人的技术突破为核心,使其逐渐成为人工智能领域的研究热点。近年来,随着计算机视觉、深度学习、强化学习等技术的快速发展,具身智能在理论研究和实际应用方面都取得了显著进展,迎来了新的发展机遇。

工业机器人在自动化产线作业已经几十年了,这难道不是具身智能?确实,从二十世纪八十年代,大量重复性、固定程式固定步骤的作业工序逐渐由工业机器人承接并完成。工业机器人被广泛应用于产线作业是自动化时代重要的落地应用之一。但是工业机器人是不具有自主感知、认知、决策和执行、接入自主修正等能力的,依赖编程逻辑;固定用于某一道工序的工业机器人,如果更换成另一道工序,则需要重新编程,可以认为经典工业机器人依赖规则设定和专家经验。近年来,工业机器人也在逐渐增加人工智能的组件和模块,将 AI 视觉目标检测组件融合至经典工业机器人可以实现目标智能检测和传统编程操作的结合,但是这种方式是局部智能、局部规则,而不是 end-to-end 全局具身智能。同样的,送餐机器人也是类似的局部智能、局部规则。从广义层面,以上应用都可以认为是具身智能的初级形态,但是具身智能的大众化以及当下所热议的具身智能主要是指狭义层面的具身智能,也就是以人形机器人为代表,以前沿人工智能技术全局驱动物理本体为核心研究方向和应用落地方向的具身智能。

现在所探讨的具身智能概念,更强调其人工智能的属性,更强调人工智能通过物理实体与环境实时交互,实现感知、认知、决策和行动一体化。送餐机器人、智能辅助驾驶车辆都是具身智能的初级形态,而具身智能的终极目标是实现全自主学习、自主执行、自主调整,稳定可控的、end-to-end 具身智能。

2 具身智能的技术架构

2.1 多模态感知技术

多模态感知技术是具身智能的基础,它通过融合视觉、触觉、力觉、听觉等多种传感器的信息,实现对环境的全面感知和理解。视觉传感器如摄像头、主动

式距离/深度传感器以及激光雷达等能够获取环境的图像和三维信息,用于目标识别、定位和导航;触觉和力觉传感器可以感知物体的表面特征、形状和力的大小,使具身智能本体能够实现精细操作和安全交互;听觉传感器则用于语音识别和声音定位,实现人机语音交互。通过多模态感知技术,具身智能本体能够构建更加准确、全面的环境模型,为后续的决策和行动提供可靠依据。例如,人形机器人结合视觉反馈系统、激光雷达和触觉反馈系统,能够在复杂的生产环境中精确地抓取和操作物体。在实验室实现中,训练好的具身智能双臂夹爪能够在虚拟环境进行任意点位木块感知,决策,抓取,最后拔插。无论红色和蓝色物体出现在桌面的任何点位,具身智能上肢和双手都能实现感知、决定操作、抓取(见图1)和拔插(见图2)。

2.2 强化学习与自适应控制技术

强化学习是一种通过具身智能物理本体与环境进行交互,根据环境反馈的奖励信号来学习最优策略的机器学习方法。在具身智能中,强化学习被广泛应用于决策和控制,使机器人能够在不断试错的过程中

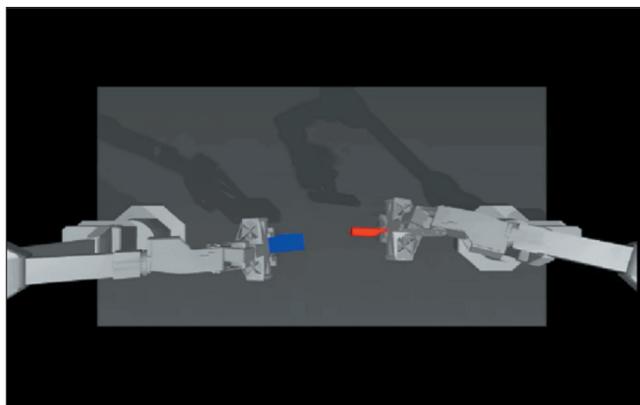


图1 仿真环境中桌面任意点位木块抓取示意

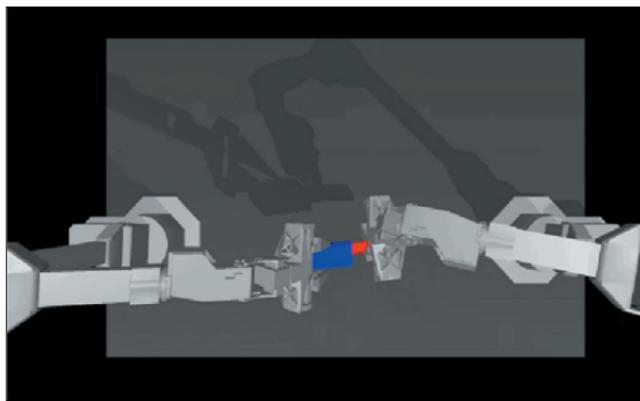


图2 将红色木块插入蓝色木块中心凹槽处示意

学习如何在不同环境下做出最优决策,以实现目标任务。自适应控制技术则使机器人能够根据环境的变化实时调整自身的控制策略,以保持良好的性能和稳定性。通过强化学习与自适应控制技术的结合,具身智能机器人能够在复杂多变的环境中进行自主学习和适应,实现更加灵活、智能的行为。

2.3 具身大模型

具身大模型是具身智能的核心技术之一,它通过对大规模数据的学习和训练,使具身智能 agent 具备强大的语言理解、视觉感知和决策能力。具身大模型能够将语言、视觉、行动等多种模态的信息进行融合和处理,实现对复杂任务的理解和执行。同时,具身大模型还具有良好的泛化能力,能够在不同的应用场景中进行快速适应和学习,为具身智能的发展提供了有力支持。例如,PaLM-E 多模态大模型将语言理解与机器人控制相结合,使机器人能够根据语言指令执行各种任务,推动具身智能进入了语义交互的新纪元。

2.4 虚拟仿真环境中的大规模强化训练

通过不断扩大训练数据的规模并增加其多样性,具身智能 agent 能够在虚拟环境里持续地进行学习与试错(见图3)。在这个过程中,具身智能 agent 可以掌握更丰富的行为模式,识别更多的环境特征,进而提升自身的智能和行为能力,最终达到提高模型性能和智能水平的目的。如此一来,模型在面对复杂现实场景时,应对能力也会显著增强,智能体也能更充分地探索环境,发现新的行为策略。可以说,虚拟仿真环境中的大规模强化训练,是具身智能迈向广泛应用的坚实基础,为其未来的蓬勃发展奠定了牢固基础。

2.5 分布式协同技术

分布式协同技术用于构建多具身智能 agent 协作网络,使多个具身智能 agent 能够在分布式环境下协同工作,实现共同目标。通过分布式协同技术,具身智能 agent 之间可以进行信息共享、任务分配和冲突消解,提高群体智能的效率和性能。在智慧物流场景中,上百台自动导引车(AGV)通过动态任务分配与冲突消解算法,能够高效地完成仓储分拣搬运任务,展现了群体智能的规模效应。分布式协同技术的发展,为具身智能在复杂场景下的应用提供了更加可靠的解决方案。

3 具身智能的应用案例和前景

3.1 工业制造领域



图3 仿真环境中大规模强化学习示意

在工业制造领域,具身智能的应用能够显著提高生产效率、降低成本、提升产品质量。例如,开普勒人形机器人在通用汽车厂代替人工冲压上料^[14],优必选人形机器人在吉利工厂代替人工从AGV搬运到冲压环节^[15]。具身机器人还能够在危险、恶劣的工作环境中代替人类工作,保障工人的安全和健康。

3.2 医疗健康领域

具身智能在医疗健康领域的应用前景广阔,能够为医疗服务带来革命性的变化。加州大学圣地亚哥分校的研究团队为宇树人形机器人开发了一种双手远程操作系统,这套远程操作系统能让机器人上岗手术台,学会体检、精密针刺等技术^[16]。此外,具身智能还可用于康复辅助、智能护理等领域,为患者提供更加个性化、智能化的医疗服务。

3.3 自动驾驶领域

自动驾驶是具身智能的重要应用场景之一,通过融合激光雷达、摄像头、毫米波雷达等多传感器数据,以及深度学习、强化学习等算法,自动驾驶汽车能够实现道路环境的实时感知、决策和控制,实现高度自动化驾驶。例如,智能辅助驾驶技术通过融合激光雷达与视觉感知,能够实现车辆在复杂城市路况下的实时避障和安全行驶;百度Apollo结合大模型优化路径规划,已在Robotaxi中部署,为人们提供更加便捷、

高效的出行服务。自动驾驶技术的发展,将有望减少交通事故、缓解交通拥堵,改变人们的出行方式和城市交通格局。

3.4 家庭服务领域

随着人们生活水平的提高,对家庭服务的智能化需求也日益增长。具身智能在家用机器人中的应用,使机器人能够为家庭提供清洁、陪伴、教育等多种服务。腾讯Robotics X实验室的小五机器人,能够实现照顾老人等任务,为家庭生活带来便利^[17]。伏羲理疗机器人通过AI视觉定位穴位,模拟名家按摩手法,为用户提供个性化的理疗服务^[18]。来自深圳的星尘智能机器人在收拾杂物、拖地扫地、擦拭桌面等家庭任务中大显身手^[19]。未来,随着具身智能技术的不断发展,家用机器人将更加智能化、人性化,成为家庭生活的好帮手。

4 具身智能工业应用实践与挑战

如前文所述,目前具身智能的在工业应用的初级形态要聚焦于单一工序的智能化改造,以“感知+执行”的轻量化组合为特征。以基于AI视觉检测的机械臂抓放系统为例。在3C产品组装线、汽车零部件分拣场景中,AI视觉系统通过深度学习模型识别工件位置与姿态,引导机械臂完成抓取—搬运—放置动作。此

类系统的机械臂的运动轨迹、抓取力度等参数需通过编程或离线仿真预设。当工件尺寸公差超过预设范围、来料姿态随机变化(如散乱堆放的铸件),或工序调整(如新增部件)时,需工程师重新编写运动控制代码。此外,在粉尘车间(如铸造厂)、强光直射工位(如焊接区),此类系统的视觉传感器易受干扰导致可用性下降;对于柔性工件(如布料),触觉反馈缺失可能导致抓取变形率超标。因此,当前行业正从“感知—执行”向“全流程自主作业”跃迁,核心突破点在于在单一工序中实现类人决策能力,实现从“局部优化”到“全局智能”跃迁。

5 具身智能的未来趋势

5.1 自主学习与泛化能力提升

未来具身智能将通过大模型与强化学习等技术的深度融合,进一步提升自主学习和泛化能力。智能体将能够在更少的人工干预下,通过与环境的交互自主学习知识和技能,并且能够将在一个场景中学习到的能力快速应用到其他不同场景中,实现更加通用的智能。这将使具身智能系统能够更好地适应复杂多变的现实环境,完成更加多样化的任务。

5.2 多模态感知深化

随着传感器技术的不断发展,具身智能将实现触觉、嗅觉、味觉等更多感官的集成,进一步深化多模态感知。通过融合多种感官信息,智能体能够更加全面、准确地理解环境,增强对环境的感知维度和认知能力。例如,在医疗领域,结合触觉和嗅觉感知的具身智能设备可以更准确地检测疾病;在家庭服务领域,具备嗅觉感知的家用机器人可以更好地完成清洁任务,检测异味等。

5.3 稳定性和鲁棒性提升

具身智能是物理实体与智能系统深度融合的前沿领域,其外在不再局限于执行单一任务,而是朝着适应复杂环境的方向大步迈进。提升具身智能稳定性与鲁棒性,已然成为突破产业化瓶颈的核心任务。需打破传统控制理论与深度学习之间的协同壁垒,将动态环境下的模型预测控制(MPC)与实时强化学习(RL)紧密耦合,构建起时空连续的决策优化框架。借助这一框架,有效解决感知噪声累积、动作执行延迟等动态失配问题。同时,搭建“环境扰动—硬件反馈—算法容错”的闭环强化机制,强化系统应对复杂情况的能力。在硬件上采用冗余驱动设计与自适应柔

顺控制技术,提高系统的可靠性和鲁棒性。推动具身智能从当前“有限场景可用”向“复杂世界可信”的全新范式转变,为其实现规模化应用筑牢根基。

5.4 工作效率和质量提升

在具身智能的未来发展中,提升工作效率和输出质量是关键课题。优化感知与决策算法起着至关重要的作用,利用更先进的多模态融合技术,快速精准地理解环境信息,减少决策时间,引入强化学习和迁移学习,让智能体快速适应新任务,避免重复学习,提高执行效率。同时,升级硬件设施不可或缺,配备高性能处理器、采用新型材料和设计优化机械结构,提升数据处理速度,降低磨损、提高精度。此外,通过构建完善的任务规划与调度系统,依据任务优先级和资源状况合理分配任务,进一步提升工作效率和输出质量。

5.5 低功耗硬件突破

为了解决具身智能硬件的能耗问题,固态电池、神经拟态芯片等低功耗硬件技术的研发将成为重点。固态电池具有高能量密度、长寿命等优点,有望为具身智能设备提供更持久的续航能力;神经拟态芯片则模仿人类大脑的神经元结构和工作方式,能够实现高效的计算和低功耗运行。这些低功耗硬件技术的突破将为具身智能的发展提供更强大的硬件支持,推动具身智能设备的小型化、便携化和长时间运行。

5.6 跨行业融合

具身智能将与元宇宙、数字孪生等新兴技术进行深度融合,拓展虚拟训练场景和应用领域。在元宇宙中,具身智能体可以进行虚拟训练和模拟操作,获取丰富的经验和技能,然后将这些经验迁移到真实世界中。数字孪生技术则可以为具身智能提供精确的物理模型和实时仿真环境,实现对真实系统的实时监测、优化和控制。具身智能与其他技术的跨行业融合,将创造出更多的创新应用和商业模式,推动各行业的数字化转型和智能化升级。

6 结论

具身智能作为人工智能领域的前沿方向,通过具身智能机器人与环境的实时交互实现了感知、认知、决策和行动的一体化,为各行业的智能化升级带来了新的机遇。英伟达创始人兼CEO黄仁勋讲,在广义层面,人工智能本质上就是智能机器人技术。然而现实情况是,更接近甚至超越人类的全自主感知、决策并

行动的机器人,能够模拟人类学习复杂环境、连续决策、连续行动的机器人技术之前并不存在。传统机器人是无法只通过反复学习就能实现感知、决策和行动的,而基于人工智能的机器人,也就是具身智能,则可以实现。传统机器人在容错率小的场景下严重依赖精准的控制程序和指令,主做一项任务的机器人如果换到另一项任务中就需要重新修改程序和指令。事实上,大部分机器人只能做固定类型的工作,换个环境或决策复杂的场景,要么机器人无法达到人类水平,要么需要重新人为修改程序和指令。具身智能技术的进步慢慢会带来更拟人、更类人、某些方面更超人的机器人。

本文通过对具身智能的核心概念、技术架构、应用案例、发展挑战及未来趋势的全面分析,揭示了具身智能在推动社会智能化转型中的重要作用。尽管具身智能在发展过程中仍面临诸多挑战,如数据获取与仿真鸿沟、硬件限制、软件生态标准化缺失以及伦理与安全等问题,但随着技术的不断进步和创新,这些问题有望逐步得到解决。未来,具身智能将在自主学习与泛化能力提升、多模态感知深化、稳定性和鲁棒性提升、工作效率和质量提升、低功耗硬件突破和跨行业融合等方面取得显著进展,实现更加广泛的应用和发展。具身智能不仅将改变人们的生产生活方式,还将在全球范围内掀起新一轮科技革命,为经济的高质量发展注入强劲动能。

因此,各国应高度重视具身智能的发展,加大政策支持、资本投入和产学研合作力度,加快具身智能的核心技术攻关和商业化应用,构建全球竞争力,共同推动具身智能技术的发展和应用,为人类创造更加美好的未来。

参考文献:

[1] 人民网理论频道. 何为“具身智能”. [EB/OL]. [2025-05-17]. <https://theory.people.com.cn/n1/2025/0319/c40531-40441929.html>.

[2] 学习时报. 何为“具身智能”. [EB/OL]. [2025-05-17]. <https://www.infoobs.com/article/20250320/68768.html>.

[3] ZHAO W S, PEÑA QUERALTA J, WESTERLUND T. Sim-to-real transfer in deep reinforcement learning for robotics: a survey [EB/OL]. [2025-05-17]. <https://doi.org/10.48550/arXiv.2009.13303>.

[4] SAPKOTA R, CAO Y, ROUMELIOTIS K I, et al. Vision-language-action models: concepts, progress, applications and challenges [EB/OL]. [2025-05-17]. <https://arxiv.org/abs/2505.04769>.

[5] TRAN K T, DAO D, NGUYEN M D, et al. Multi-agent collaboration mechanisms: a survey of LLMs [EB/OL]. [2025-05-17]. <https://arxiv.org/abs/2501.06322>.

[6] SUN F C, CHEN R F, JI T Y, et al. A Comprehensive survey on embodied intelligence: advancements, challenges, and future perspectives[J]. CAAI Artificial Intelligence Research, 2024, 3: 9150042.

[7] JIANG J, CHEN C, FENG S, et al. Embodied Intelligence: The Key to Unlocking Generalized Artificial Intelligence. [EB/OL]. [2025-05-17]. <https://doi.org/10.48550/arXiv.2505.06897>

[8] ZHAO Z K, WU Q X, WANG J, et al. Exploring embodied intelligence in soft robotics: a review[J]. Biomimetics, 2024, 9(4): 248.

[9] ZHANG L, LI H. Symbolic AI vs. Embodied AI: Symbolic AI vs. Embodied AI: a comparative study. [J]. AI & Machine Learning Journal, 2023, 22(4): 84-95.

[10] LIU Y, CHEN W X, BAI Y J, et al. Aligning cyber space with physical world: a comprehensive survey on embodied AI [EB/OL]. [2025-05-17]. <https://arxiv.org/abs/2407.06886>.

[11] TURING A M. Computing machinery and intelligence [J]. Mind, 1950, 59(236): 433-460.

[12] BROOKS R A. Intelligence without representation [J]. Artificial Intelligence, 1991, 47(1/3): 139-159.

[13] FU Z P, ZHAO T Z, FINN C. Mobile ALOHA: learning bimanual mobile manipulation with low-cost whole-body teleoperation [EB/OL]. [2025-05-17]. <https://arxiv.org/abs/2401.02117>.

[14] 中国机器人网. Figure CEO点赞的开普勒“大黄蜂”,正式落地上汽“打工” [EB/OL]. [2025-05-17]. <https://baijiahao.baidu.com/s?id=1830535665728707561&wfr=spider&for=pc>.

[15] 南方都市报. 优必选人形机器人“入职”极氪,从事搬运任务 [EB/OL]. [2025-05-17]. <https://baijiahao.baidu.com/s?id=1806518250691435959&wfr=spider&for=pc>.

[16] 中国科学报. 中国机器人在美国当医生了? 宇树机器人拿起手术刀 [EB/OL]. [2025-05-17]. <https://weibo.com/2427364747/PkD-kVngHi>.

[17] IT之家. 腾讯机器人实验室推出第5代机器人,可给老人取快递、抱老人起床 [EB/OL]. [2025-05-17]. <https://baijiahao.baidu.com/s?id=1811053255349866556&wfr=spider&for=pc>.

[18] 网易伏羲. 理疗按摩机器人: 现代健康管理的新兴解决方案 网易伏羲 [EB/OL]. [2025-05-17]. <https://fuxi.163.com/database/1908>.

[19] 经济日报新闻客户端. 星尘智能发布新一代 AI 机器人助理 [EB/OL]. [2025-05-17]. http://www.ce.cn/xwzx/gnsz/gdxw/202408/21/t20240821_39112402.shtml.

作者简介:

王晓思,博士,中国联通人工智能首席专家,《泰晤士报》全英科技女性新星,长期从事数据科学与人工智能领域科技攻关、产业化和全球化工作,近期主要从事工业具身智能大小脑的技术突破和产业化工作;林家宁,硕士,主要从事具身智能在实际场景的可行路径、软硬件协同机制与应用推广策略等研究工作;白琳,硕士,主要从事自然科学基金项目的统筹管理工作。