

具身智能赋能教育的内涵、框架与研究路向

王建虎, 王静, 万亚玲

(新疆师范大学教育科学学院, 新疆乌鲁木齐 830017)

[摘要] 当前人工智能驱动的学习存在显著的具身性缺失与情境脱嵌,主要表现为认知过程与身体经验的二元割裂以及知识建构与真实情境的疏离,这种技术异化现象正不断加深教育领域中认知与实践的鸿沟。具身智能作为人工智能演进的重要方向,为重塑教育的具身化特征、实现个性化发展提供了新范式,已成为培育教育新质生产力的关键动能。鉴于此,研究在阐述具身智能理论内涵及教育领域的应用的的基础上,尝试构建了具身智能赋能教育领域的实现框架,提出了具身智能赋能教育领域的可能研究路向,为具身智能在教育领域的应用提供理论参考和实践借鉴。

[关键词] 具身认知; 具身智能; 人形机器人; 大模型; 人工智能

[中图分类号] G434 **[文献标志码]** A

[作者简介] 王建虎(1980—),男,新疆乌鲁木齐人。教授,博士研究生,主要从事具身智能、知识管理、教师专业发展、数字化学习研究。E-mail:bsyx1102@qq.com。万亚玲为通信作者,E-mail:wylxjnu@163.com。

一、引言

2024年2月,《北京具身智能科技创新与产业培育行动计划(2025—2027年)》发布,提出要推动具身智能技术的科技创新和产业创新深度融合,加快活跃产业生态,培育人工智能发展新赛道^[1]。2025年3月,第十四届全国人民代表大会第三次会议《政府工作报告》中提出“建设未来产业投入增长机制,培育生物制造、量子科技、具身智能、6G等未来产业”^[2]。以ChatGPT4、DeepSeek为代表的大语言模型在逻辑推理、知识生成、数据处理方面表现卓越,这些数据驱动的智能,不依赖物理实体与环境的互动,是典型的“离身智能”,与此对应,具身智能则是通过机器人等实体与外界环境进行交互,实现自主学习和对外界环境的不断适应。具身智能作为人工智能从抽象计算模型到物理实体世界转变的进一步延伸,其从“离身”到“具身”的演进是人工智能发展的必然趋势。因此,具身智能的发展也被视为一条实现通用人工智能的基本途

径。这一技术的发展将推动教育领域从“离身认知”向“具身共生”的范式转型,为破解教育领域长期存在的“知行分离”的困境提供了新思路。

本研究聚焦于在具身智能技术快速发展和教育数字化转型的双重背景下,具身智能如何赋能教育领域新生态?将从具身智能的理论基础、技术应用和教育实践三个维度出发,首先阐释其理论内涵,追溯起源脉络并解析核心要素;其次以“功能递进”为轴线,划分具身智能的教育应用层级,构建具身智能赋能教育的实现路径框架;最后提炼未来研究的关键路向。

二、具身智能的理论内涵

(一)具身智能的历史源起

具身智能的理论萌芽可追溯至人工智能的初期。1950年,图灵在《Computing Machinery Intelligence》中提出“机器是否能思考”的问题^[3],引发了人们对机器与智能关系的遐想。1986年,机器人学家布鲁克斯提出

基金项目:2025年度新疆维吾尔自治区高校本科教育教学研究与改革项目综合教改项目“深入推进慕课西部行计划,东西部高校优质教育资源共享研究”(项目编号:XJGXJGZH-2025015);新疆维吾尔自治区2024高校基本科研业务费项目“数智技术赋能南疆县域教师精准教研模式研究”(项目编号:XJEDU2024J066)

“行为机器人”的概念,强调智能是具身化和情境化的,引导研究者从算力研究转向身体和环境交互的研究,并认为具身智能是走向通用人工智能的必经之路。21世纪初,人形机器人和仿生机器人的出现,进一步推动了具身智能的发展,也为具身智能的实际应用提供了条件。

(二)具身智能的概念解析

具身智能从字面意思上理解为“具身化的人工智能”。“具身”指具有物理实体,并能通过感知、交互、行动等能力执行任务。“智能”则体现为对具身感知到的多模态信息(文本、视觉、语音等)的理解与转换能力。具身智能强调智能受到大脑、身体和环境紧密耦合的影响,通过感知信息和与环境的物理交互实现自主学习和进化^[4]。首先,具身智能不是大语言模型(Large Language Models, LLMs)等技术与机器人的简单结合,大语言模型虽然具有语言理解与生成等能力,但缺乏主观感知能力,而具身智能则通过“感知—行动”闭环实现对环境的自主决策和自适应行动。其次,具身智能的目标之一就是让机器人像人一样,灵活、高效、稳健地行动。人形机器人虽然被视为具身智能理想的应用形态,但具身智能也不能等同于人形机器人。最后,具身智能不等同于智能体(Agent),智能体通过感知周围环境采取相应的行动,其可以是虚拟智能体(如聊天机器人、智能助手等),也可以是真实存在的智能实体(如智能机器人、工业机械臂等)。因此,具身智能下的智能体特指拥有物理实体的智能体,其可以通过与环境进行情景化的交互,主动获取物理世界的真实反馈^[5]。

以上观点都承认了“智能性”和“具身性”的重要作用。根据2024年发布的《具身智能发展报告》的描述:具身智能是指通过机器人等物理实体与环境交互,能进行环境感知、信息认知、自主决策和采取行动,并能够从经验反馈中实现智能增长和行动自适应的智能系统^[6]。本研究后续论述均以此定义为基础。

(三)具身智能的要素构成

具身智能一般由三大要素构成:本体、智能和环境^[6](如图1所示)。本体是具身智能的物理基础,通常指具有多种形态的物理实体机器人。智能指嵌入具身本体的智能大脑,如大语言模型(Large Language Models, LLMs)、视觉语言模型(Vision Language Model, VLM)、视觉语言动作模型(Vision Language Action, VLA)等。环境要素方面,具身智能本体与环境的交互不仅能够感知环境,还能通过行动影响环境,在与环境的交互中不断学习,不断适应。就三者关系

而言,环境是支撑学习过程与策略优化的“试验场”;本体是连接物理世界和虚拟世界的“感官和四肢”,拥有“感官和四肢”的本体感知周围环境并与环境交互获取信息、理解问题;智能大脑是处理信息、知识学习和思维推理的“神经中枢”,通过不断增强本体的感知、认知、决策及行动能力来指挥本体实施行动。拥有智能大脑的具身本体和环境进行双向交互,形成“感知—行动”的闭环,实现智能进化。

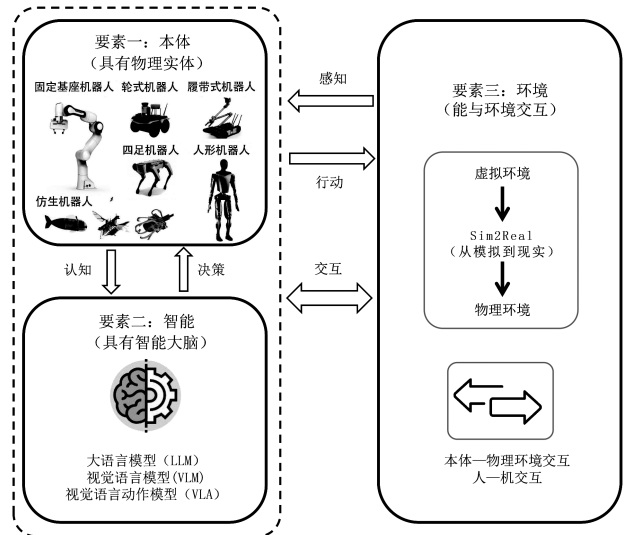


图1 具身智能的核心要素

三、具身智能赋能教育的应用层级

具身智能是具身认知理论在人工智能领域的实践延伸。具身认知强调认知活动的形成是大脑、身体和环境之间相互作用的过程^[7]。在具体学习情境中身体参与所获得的认识,不仅是思维的演绎,更是情感、态度、直觉、体验等非理性思维的形成^[8],且具有涉身性、体验性与环境性等特点^[9]。人们的基本概念来自感知觉体验,主要来自人们对身体、情境等的认识^[10]。在知识学习的过程中,环境、身体和认知是不可缺少的,缺失具身的情境或是脱离了身体的实践,阻碍了“身体参与认知”^[11]。本研究从具身认知的理论视角下将教育领域的具身智能应用划分为三个层级:初级具身、中级具身和高级具身(如图2所示)。

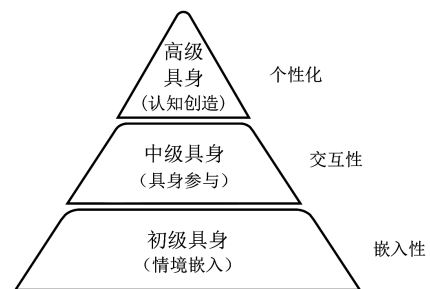


图2 教育场景中具身智能的应用层级

首先,学习是一种“嵌入”身体和环境的活动^[7]。只有在具体情境中,学生才能获得切身体验,身体的意志和归属^[2],感知与知识接触的快乐。在初级具身智能阶段通过情境嵌入的方式突破单一学习场景限制,借助虚实融合的环境,例如,在具身型混合现实学习环境中^[13],学生能获得沉浸式的具身体验,激发其学习兴趣。其次,学习是全身心参与的过程。教育心理学家 Francesconi 等人指出,“大脑仅仅是身体的一个特殊器官,思维源于整体的人,源于有机体”^[14]。中级具身智能阶段强调具身的交互,学生能借助具身智能在虚实融合的环境中通过操作、实践、观察、反馈来内化知识,比被动听讲或阅读更高效。例如,张金莹等人研究发现,不同具身程度的虚拟实验对大学生的学习效果都具有较高的促进作用,且在不同学科和知识类型也有不同程度的影响^[15]。高级具身则强调创造性认知和认知的个性化,通过构建虚实融合的学习环境,并借助具身智能在与环境交互的过程中(如手势塑造虚拟几何体、虚拟实验操作、全身演奏可视化音乐等),将“抽象的概念”转化为“具身的体验”,促进学习者知识结构的深度重构与创新。通过“情境嵌入—具身参与—认知创造”三个层级,学习者经历从初级具身、中级具身到高级具身的认知深化过程,三个层级并非相互独立,而是构成了一个动态整合的有机整体。其中,情境构成了认知发生的具身化场景,实体则作为“感知—行动”闭环的实现载体,支持知识在具身实践中的创造性重构和跨情景迁移,层级的划分为构建具身智能赋能教育领域的应用框架提供了结构化范式。

四、具身智能赋能教育的实现框架

基于教育场景的具身智能层级,本研究从具身认知理论、人工智能技术和教育教学实践角度出发,提出具身智能赋能教育的实现框架(如图3所示)。该框架由虚实环境、具身交互和智能大脑三大部分组成,并对应三个层级,分别是情境嵌入、具身参与和认知创造。该框架通过具身智能技术赋能,强调虚实融合的学习环境、具身参与的学习体验和智能大脑的认知深化对教育的作用,可为具身化、情境化和个性化的未来教育提供理论和技术支撑。

(一)情境嵌入:虚实融合的学习场景构建

具身智能需要通过与物理世界(物理或虚拟)的交互来学习知识、适应环境和优化行动策略。在物理环境下,具身智能通过物理载体借助多模态传感器接收外部环境的视觉、听觉、触觉、嗅觉等信息,为学习者提供多感官刺激和丰富的学习体验。而在虚拟环境中,通过虚拟仿真平台构建一个逼真的模拟环境,进行演示转换和重放,自动生成用于模仿学习的训练数据,将学习到的能力或行为转移到现实世界。总体上,情境嵌入中的环境包括真实物理环境、虚拟仿真环境和混合现实环境等。在真实物理环境中,学习活动基于真实物理空间与操作场景,实现沉浸式学习体验。在虚拟仿真环境中,学习者借助VR眼镜、VR动捕数据手套等进入沉浸式学习环境^[16-17]。在混合现实环境中,虚拟元素通过投影物理空间实现虚拟世界与现实世界的叠加^[18],弥补虚拟空间与物理环境的隔阂,引领

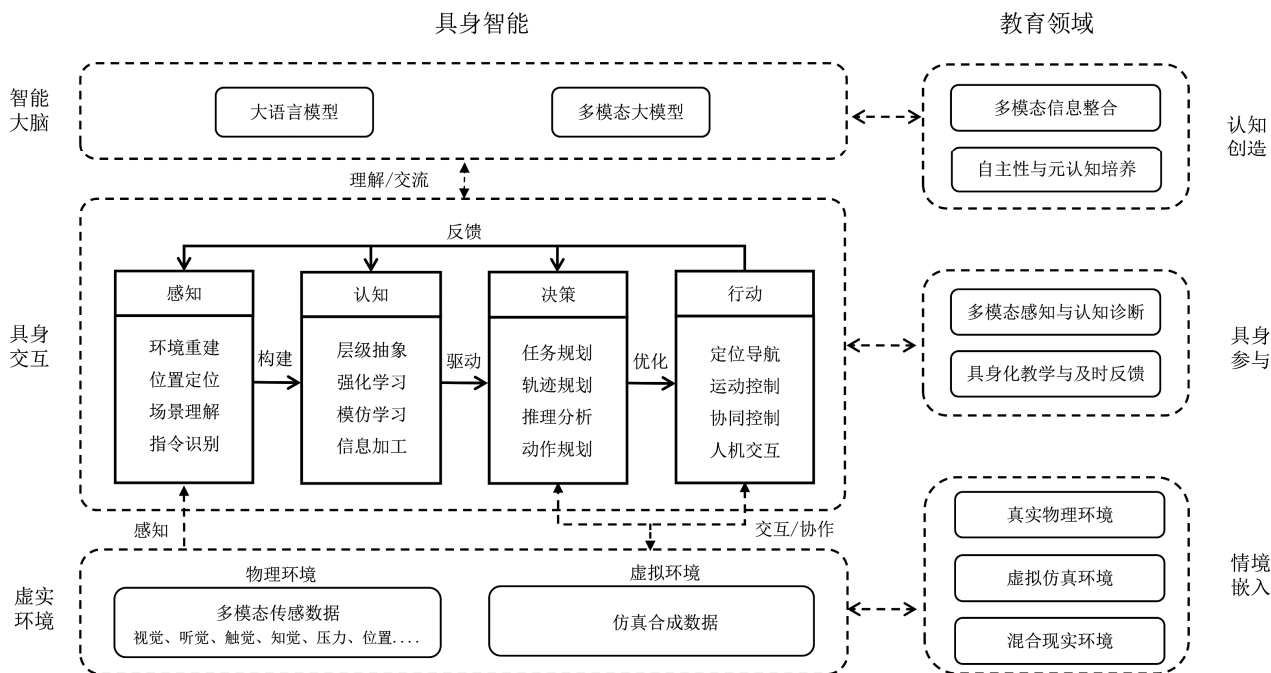


图3 具身智能赋能教育的实现框架

学生从“离身”向“具身”的转化,进而为学习者提供实时交互的混合现实环境^[19-20]。

(二)具身参与:多模态感知的具身交互

具身智能通过其物理化躯干的架构(如机械臂、仿生关节,机器人等)构建起与物理世界实时交互的基础,并以第一人称视角主动感知、规划、决策,驱动具身智能完成自主行动,实现“感知—认知—决策—行动”一体化^[21]。首先,具身智能通过感知能力全面感知环境信息,借助传感器、摄像头、麦克风等设备采集视觉、听觉和触觉等多模态数据,并依据多模态信息理解三维场景(推断物体几何与物理属性,有效识别目标对象)、识别有效指令、确认自身位姿。在此基础上通过即时定位与地图构建(Simultaneous Localization and Mapping, SLAM)等技术在移动中持续扫描环境,构建空间地图并完成三维环境重建。其次,具身智能通过与环境交互,不断总结过去经验、反思视觉感知结果以形成认知地图^[22]。在认知的过程中,模仿学习提供行为先验,强化学习(与环境的试错交互)驱动自主进化,具身智能将感知到的信息进行层级化抽象并逐级提炼,通过信息加工实现层级间的信息流动与行为输出,为决策和行动提供支持。最后,具身智能需要理解外界指令(如自然语言与场景语义),将复杂任务进行分解,规划子任务,通过轨迹规划生成最优可行路径,通过动作规划生成精确运动指令,依据推理分析来适应复杂环境,完成智能决策,并根据智能决策实现目标定位与位置导航,从而控制和优化行动。总之,如果说感知是具身智能的“五官”,那么认知就是它的“大脑”,决策是它的“神经中枢”,行动便是它最终的“归宿”。具身智能在教育中的应用体现了“具身认知”理论的核心主张,即认知过程本质上依赖于身体与环境的互动。借助具身智能能够实时感知学习者的状态,包括面部微表情、语音、肢体动作等生理行为特征,同时获取学习过程、交互记录等学习行为数据,通过对这些数据的深度分析,精准识别学习者的知识掌握程度、认知特点及情感状态。物理具身(如教育人形机器人)的介入,能够实现教学示范、即时反馈和情境化指导等功能,同时感知学习者的情感状态(参与度、挫折感),形成一个完整的教学闭环,以克服传统教育评估滞后性与主观性的矛盾。借助具身智能,可以更加有效地突破传统人机交互局限,实现更加自然、沉浸式的学习体验,进一步根据每个学习者的认知特点和情感需求,提供适切的个性化教育服务。

(三)认知创造:智能驱动的知识生成与创新

具身智能的“智能大脑”依托大语言模型和多模

态大模型(Multimodal Large Language Models, MLLMs)等技术,实现了对自然语言的理解、生成,以及对多模态数据的整合和分析。多模态大模型能够同时理解和整合不同感官的信息,赋予了具身智能一种近乎人类的理解能力。例如,谷歌的PaLM-E具身多模态大语言模型^[23]和OpenAI开发的CLIP多模态(文本和图像)预训练模型^[24]。其中,CLIP模型通过将文本和图像嵌入到共同的语义空间中,从而实现跨模态的理解。在具身智能技术支持下,通过对学生行为数据的理解分析,构建精准的学生画像,并融合情绪识别与课堂行为分析等多模态感知技术,动态调整教学策略和交互方式,从而激发学生的学习积极性、提高其自主学习能力。Memarian等人构建了涵盖“环境—身体—大脑”三元维度的分类体系,其中,每个维度均可能有人类、人工智能技术、人类+人工智能技术的参与^[25],为理解具身智能赋能教育提供了理论依据,对应本框架中的虚实环境、具身交互和智能大脑三部分,在赋能教育场景中,分别对应情境嵌入,具身参与和认知创造。虚实融合的环境提供了教育教学中情境嵌入的基础,具身交互创造了身体参与和交互的可能性,智能大脑则驱动认知创造、知识生成的质变效应。这一框架不仅突破了传统教育中的“离身教育”的局限性,也通过具身智能赋能学习者获得自主学习能力和主动建构和创新知识的能力。

五、可能的研究路向

作为新一代人工智能技术体系的核心构成,具身智能已在工业制造、自动驾驶、智慧医疗等多个领域展现出变革性潜能和技术穿透力,通过物理实体与智能系统的深度融合,持续重塑产业格局、驱动社会进步并深度介入人类生活场景,如工业场景中多模态机械臂实现柔性生产线的自主调优,医疗领域通过触觉反馈机器人辅助外科手术精准化,服务行业借助人形机器人完成复杂场景的递送与交互等。教育作为知识传承与认知塑造的核心场域,具身智能的技术特性与教育本质具有深层耦合关系,也为教育领域突破“离身认知”的范式困境创造了条件,为知识传承与认知塑造提供了新范式,这必将推动教育教学方法、模式的深度变革。具身智能在教育领域的落地应用仍需进一步探索可行的研究路向与应用场景。基于此,本研究从学习者、教师、人机协同及学习环境四大核心视角切入,尝试提出具身智能赋能教育的具体研究路向。

(一)具身个性化学习

从学习者的视角来看,具身智能通过整合多模态

感知、物理交互与自适应决策技术,为个性化学习提供新的技术路径,其核心在于将学习者的身体动作、环境交互与认知过程深度融合,突破传统数字化学习“离身性”局限。例如:在 STEAM 教育中,学生可通过编程控制仿生机器人,在物理操作中直观理解力学原理与算法逻辑,通过机器人操作系统(Robot Operating System, ROS)集成环境感知模块,实时反馈力学实验数据,使抽象概念通过具身实践内化为可迁移的技能^[26]。在特殊教育领域,人形机器人作为交互伙伴和治疗工具,在环境中以受控且可预测的方式复现社交场景,进而优化自闭症儿童注意力分配并降低社交焦虑,为其制定更全面、个性化的治疗计划^[27]。在体育训练中,融合动作捕捉系统精确捕捉人体运动数据,通过触觉反馈、视觉叠加等具身交互方式提供实时纠正指导,加速技能内化并预防错误动作固化^[28]。具身智能的教育价值不仅体现于个性化适配的精准性,更在于重构学习者的主体性体验。当前,具身智能正推动教育从“离身认知”向“身心合一”跃迁,但其最终目标并非取代教师,而是通过人机协同释放教育者的人文关怀潜能^[29],让技术真正服务于“全人发展”的教育本质。

(二)具身教学代理

具身教学代理以物理形态存在,该形态下的具身代理机器人可伴随教师实时监控教学动态,通过多模态感知设备捕捉学生学习行为,并借助手势、表情及语音在物理环境中与学生实现自然交互,辅助教师提升教学质量,消解学习者的“人机隔阂”。具身教学代理作为学习伙伴和教学导师正随着技术的迭代不断浮现。2022年,瑞典初创公司开发的机器人 Furhat 不仅可以通过自然语言,还可以通过面部表情等非自然语言提示与人类互动。在学习语言时,通过模拟真人口语交谈场景,帮助学生练习外语,增强其语言运用能力。2024年,松灵机器人推出了面向科研教育的具身智能移动协作机器人 Cobot S Kit,深度整合了移动机器人平台、精确传感、AI 大模型技术,以及灵活的机械臂设计,为科研教育提供了一个优质的模拟实验环境。2025世界数字教育大会正式发布了中小学科学教育智能导师原型,中小学科学教育智能导师关键技术也被列为国家重点研发计划。此外,具身智能代理也为特殊需求的师生打开了一个全新的学习领域,通过图像、声音、视频的交互和多模态信息的感知,能够极大程度上赋予特殊群体儿童之前无法感受的学习体验,还可以应用于远程支教或人智协同教研,营造一种身临其境的教育感受。

(三)多智能体支持的协同工作

黄荣怀教授提出:“具象化的智能体作为大模型的‘执行层’,承担从认知到行动的关键环节,打通大模型与实际教学场景的‘最后一公里’”^[30]。单一智能体往往难以有效应对现实世界中的复杂问题,需要具身多智能体协同工作。随着大语言模型和多模态大模型的快速发展,具身智能在推理和泛化能力上不断取得突破,加之其感知和自主决策能力的提升,将催生教育教学方式的改变,推动教育从“知识灌输式”向“具身认知建构式”转型,形成人机共生的新型教学模式,并促使教师角色由知识学习的“灌输者”转为知识学习的设计者和引导者。例如,可使用多个智能体协同支持项目式学习(每个智能体扮演不同的角色,模拟真实社会分工),或通过脑机接口实现知识的内化,突破传统的认知边界,构建具身的教育元宇宙空间。仇星月等人设计了6个智能体协同工作的教育研究场景应用^[31]。吴永和等人构建了基于大模型的智能体“眼—脑—手”三维能力框架,并提出了多智能体系统智能性提升的内外双循环框架^[32]。Kannan 等人提出了一个为具身多机器人任务规划而设计的创新框架(Smart-LLM 框架),该框架为各类具身机器人赋予了不同的能力,让多个机器人协同完成任务^[33]。此外,如何实现多智能体之间的协调控制、动态环境适应和信息共享也是亟须不断去探究的问题。

(四)具身参与的沉浸式学习

基于具身认知的沉浸式教学将成为智能时代背景下兴起的新型教学形态,强调学习时的“身体在场”与学习的具身化体验,借助虚拟现实(VR)、混合现实(MR)、增强现实(AR)及元宇宙等沉浸式技术,创建可感知、可交互的具身化体验情境。在此情境中,学习者能够通过虚拟化身进行自然互动,增强学习的具身感和沉浸感。例如:Liu 等人发现,身体确实参与了认知过程,基于 VR 的具身学习不仅可以有效提高学习成绩,还对学生的学习参与度和学习兴趣产生积极影响^[34]。Cicarelli 等人则通过沉浸式教学案例表明,基于具身认知理论,使用游戏设计技术,将身体、感知和情感整合到人工智能增强的学习环境中,具有重要的教育意义和实践价值^[35]。

六、结束语

具身智能作为实现通用人工智能的重要途径,是人工智能链接现实产业的重要依托,属于关键性的技术创新领域,也是推动新质生产力建设的重要引擎,对于推进“人工智能+”和教育数字化转型的战略布局

具有重要意义。具体而言,具身智能为教育领域破解“离身认知”范式困境,实现“具身共生”的转型提供了新的思路,并有效应对了教育领域长期存在的“知行分离”难题。本研究系统阐述了具身智能的理论内涵,具身智能赋能教育的应用层级、实现框架及可能的研究路向,以期为具身智能在教育领域的应用提供理论参考与实践探索。教育是一个复杂的、涉及多维度交互以及动态发展的系统工程,未来研究需要重点突破多模态数据的采集、多智能体协作、多场景

自适应迁移、多任务并行学习与泛化,以及多模态交互等关键技术挑战。同时,具身智能与环境深度交互、持续感知的技术特性,也不可避免地会引发一系列的隐私威胁和伦理问题,如何有效应对技术所带来的伦理风险,是具身智能教育应用过程中需要关注和应对的问题。另外,具身智能必须建立可解释性强的理论和方法,发展安全、可控、易扩展的具身智能技术,并推动具身智能在教育领域中的创新性应用和健康发展。

[参考文献]

- [1] 北京市人民政府. 北京具身智能科技创新与产业培育行动计划(2025—2027年)[EB/OL]. (2025-02-28) [2025-05-10]. https://www.beijing.gov.cn/zhengce/zhengcefagui/202503/t20250304_4024579.html.
- [2] 中华人民共和国国务院办公厅. 政府工作报告——2025年3月5日在第十四届全国人民代表大会第三次会议 [EB/OL]. (2025-03-05) [2025-05-10]. https://www.gov.cn/gongbao/2025/issue_11946/202503/content_7015861.html.
- [3] TURING A M. Computing machinery and intelligence[J]. *Mind*, 1950, 59(236): 433-460.
- [4] LIU H, GUO D, CANGELOSI A. Embodied intelligence: a synergy of morphology, action, perception and learning[J]. *ACM computing surveys*, 2025, 57(7): 1-36.
- [5] 刘华平,郭迪,孙富春,等. 基于形态的具身智能研究:历史回顾与前沿进展[J]. *自动化学报*, 2023, 49(6): 1131-1154.
- [6] 中国信息通信研究院,北京人形机器人创新中心有限公司. 具身智能发展报告(2024年)[EB/OL]. (2024-08-27) [2025-05-10]. http://www.caict.ac.cn/kxyj/qwfb/bps/202408/t20240827_491546.htm.
- [7] 叶浩生. 身体与学习:具身认知及其对传统教育观的挑战[J]. *教育研究*, 2015, 36(4): 104-114.
- [8] 钟柏昌,刘晓凡. 论具身学习环境:本质、构成与交互设计[J]. *开放教育研究*, 2022, 28(5): 56-67.
- [9] 殷明,刘电芝. 身心融合学习:具身认知及其教育意蕴[J]. *课程·教材·教法*, 2015, 35(7): 57-65.
- [10] GIBBS JR R W. Metaphor interpretation as embodied simulation[J]. *Mind & language*, 2006, 21(3): 434-458.
- [11] 赵雪梅,钟绍春. 具身认知视域下促进高阶思维发展的多模态交互机制研究[J]. *电化教育研究*, 2021, 42(8): 65-71, 87.
- [12] 邱关军. 从离身到具身:当代教学思维方式的转型[J]. *教育理论与实践*, 2013, 33(1): 61-64.
- [13] SHUGUANG L, LIN B. Construction and implementation of embodied mixed-reality learning environments [C]//2021 International Conference on Big Data Engineering and Education (BDEE). Guiyang: IEEE, 2021: 126-131.
- [14] FRANCESCONI D, TAROZZI M. Embodied education: a convergence of phenomenological pedagogy and embodiment[J]. *Studia phaenomenologica*, 2012, 12: 263-288.
- [15] 张金莹,杨现民,胡永斌. 具身程度对大学生虚拟实验学习的影响研究——基于77项实验和准实验数据的元分析[J]. *电化教育研究*, 2024(4): 1-7.
- [16] 顾君忠. VR、AR和MR——挑战与机遇[J]. *计算机应用与软件*, 2018, 35(3): 1-7, 14.
- [17] HULUSIC V, GUSIA L, LUCI N, et al. Tangible user interfaces for enhancing user experience of virtual reality cultural heritage applications for utilization in educational environment[J]. *ACM journal on computing and cultural heritage*, 2023, 16(2): 1-24.
- [18] 黄红涛,孟红娟,左明章,等. 混合现实环境中具身交互如何促进科学概念理解[J]. *现代远程教育研究*, 2018(6): 28-36.
- [19] ROKHSARITALEMI S, SADEGHI-NIARAKI A, CHOI S M. A review on mixed reality: current trends, challenges and prospects [J]. *Applied sciences*, 2020, 10(2): 636.
- [20] XU S, WEI Y, ZHENG P, et al. LLM enabled generative collaborative design in a mixed reality environment [J]. *Journal of manufacturing systems*, 2024, 74: 703-715.
- [21] 沈甜雨,陶子锐,王亚东,等. 具身智能研究的关键问题:自主感知、行动与进化[J]. *自动化学报*, 2025, 51(1): 43-71.
- [22] CHEN S, WU Z, ZHANG K, et al. Exploring embodied multimodal large models: development, datasets, and future directions[J]. *Information fusion*, 2025, 122: 103198.

- [23] DRIESS D, XIA F, SAJJADI M S M, et al. Palm-e: an embodied multimodal language model [EB/OL]. (2023-03-06)[2025-05-10]. <https://arxiv.org/pdf/2303.03378>.
- [24] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision [C]// International Conference on Machine Learning. Virtual: PmLR, 2021,139:8748-8763.
- [25] MEMARIAN B, DOLECK T. Embodied AI in education: a review on the body, environment, and mind [J]. Education and information technologies, 2024,29(1):895-916.
- [26] GAN Y, ZHANG B, SHAO J, et al. Embodied intelligence: bionic robot controller integrating environment perception, autonomous planning, and motion control[J]. IEEE robotics and automation letters, 2024,9(5):4559-4566.
- [27] GHIGLINO D, FLORIS F, DE TOMMASO D, et al. Enhancing theory of mind in autism through humanoid robot interaction in a randomized controlled trial[J]. Scientific reports, 2025,15(1):27650.
- [28] SUO X, TANG W, MAO L, et al. Digital human and embodied intelligence for sports science: advancements, opportunities and prospects[J]. The visual computer, 2025,41(4):2477-2493.
- [29] SIMS M. Self-concern across scales: a biologically inspired direction for embodied artificial intelligence [J]. Frontiers in neurorobotics, 2022,16:857614.
- [30] 黄荣怀, 达婷. 智能体打开教育拥抱大模型的入口[N]. 中国教育报, 2025-02-15(4).
- [31] 仇星月, 陈向东, 陈鹏, 等. 大语言模型支持的元综合研究: 基于智能体的方法[J]. 现代教育技术, 2025,35(1):63-72.
- [32] 吴永和, 姜元昊, 陈圆圆, 等. 大语言模型支持的多智能体: 技术路径、教育应用与未来展望[J]. 开放教育研究, 2024,30(5):63-75.
- [33] KANNAN S S, VENKATESH V L N, MIN B C. Smart-llm: Smart multi-agent robot task planning using large language models[C]// 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Abu Dhabi:IEEE, 2024:12140-12147.
- [34] LIU Z, WANG L. Study on the style construction of embodied cognitive learning supported by VR technology [J]. Education and information technologies, 2025,30(1):403-419.
- [35] CICCARELLI S, TAFURI M G. Embodied digital learning: new educational scenarios between artificial intelligence and the rediscovery of corporeality[J]. Journal of inclusive methodology and technology in learning and teaching, 2024,4(1):1-10.

The Connotation, Framework, and Research Direction of Embodied Intelligence Empowering Education

WANG Jianhu, WANG Jing, WAN Yaling

(School of Educational Sciences, Xinjiang Normal University, Urumqi Xinjiang 830017)

[Abstract] Current AI-driven learning suffers from significant deficiencies in embodiment and contextual disembedding, primarily manifested as the binary separation between cognitive processes and bodily experiences, as well as the disconnect between knowledge construction and real-world situations. This technological alienation is continuously widening the gap between cognition and practice in the field of education. As a crucial direction in the evolution of AI, embodied intelligence provides a new paradigm for reshaping the embodied characteristics of education and achieving personalized development, and has become a key driving force in cultivating new productive forces in education. In view of this, based on elucidating the theoretical connotation of embodied intelligence and its application in the field of education, the study attempts to construct an implementation framework for empowering education through embodied intelligence, and proposes possible research directions for applying embodied intelligence to the field of education, providing theoretical references and practical insights for the application of embodied intelligence in educational settings.

[Keywords] Embodied Cognition; Embodied Intelligence; Humanoid Robot; Large Model; Artificial Intelligence