

具身智能技术赋能职业教育： 应用场景、挑战与应对策略

霍自如，孔 婧

浙江工业大学教育学院，浙江 杭州

收稿日期：2025年2月18日；录用日期：2025年3月18日；发布日期：2025年3月25日

摘 要

随着第四次工业革命的深入推进，职业教育领域正面临着技能迭代速度显著加快、产教融合壁垒日益凸显等多重挑战。在此背景下，具身智能技术(Embodied Artificial Intelligence)作为一种能够与环境进行深度交互，并依据环境变化自主做出决策和执行相应行动的前沿技术，为破解职业教育发展困境提供了创新性的解决方案。本研究采用文献研究法，系统探讨了具身智能技术的基本概念体系、关键支撑技术及其在职业教育领域的具体应用场景。研究发现，其应用主要体现在三个维度：在感知能力层面，可支持沉浸式技能训练系统的构建与多模态学习体验的实现；在决策能力层面，能够实现个性化学习路径的动态规划与智能评估反馈机制的建立；在行动能力层面，则可促进机器人辅助教学系统的开发与人机协作学习模式的创新。具身智能技术在职业教育领域应用面临技术成熟度不足、教育模式适应性不足及伦理风险挑战，本研究针对性提出应对策略。展望未来，具身智能技术将在职业教育数字化转型进程中发挥更为重要的作用，推动教育范式从传统的“知识传递”向现代的“能力生成”转变，为培养适应新时代要求的高素质技术技能人才提供强有力的智能支持。

关键词

具身智能技术，职业教育，技能训练，个性化学习

Embodied Intelligence Technology Enabling Vocational Education: Application Scenarios, Challenges, and Response Strategies

Ziru Huo, Jing Kong

College of Education, Zhejiang University of Technology, Hangzhou Zhejiang

Received: Feb. 18th, 2025; accepted: Mar. 18th, 2025; published: Mar. 25th, 2025

Abstract

With the deepening of the Fourth Industrial Revolution, the field of vocational education is facing multiple challenges, including significantly accelerated skill iteration and increasingly prominent barriers to industry-education integration. In this context, Embodied Artificial Intelligence (Embodied AI), as an advanced technology capable of deep interaction with the environment and autonomously making decisions and executing corresponding actions based on environmental changes, provides an innovative solution to address the development dilemmas of vocational education. This study employs a literature review approach to systematically explore the fundamental conceptual framework, key supporting technologies, and specific application scenarios of embodied intelligence technology in the field of vocational education. The research findings indicate that its applications are primarily manifested in three dimensions: at the perceptual level, it supports the construction of immersive skill training systems and the realization of multimodal learning experiences; at the decision-making level, it enables dynamic planning of personalized learning paths and the establishment of intelligent evaluation and feedback mechanisms; and at the action level, it promotes the development of robot-assisted teaching systems and the innovation of human-machine collaborative learning models. The application of embodied intelligence technology in vocational education faces challenges including insufficient technological maturity, inadequate adaptability of educational models, and ethical risks. This study proposes targeted strategies to address these issues. Looking ahead, embodied intelligence technology will play a more significant role in the digital transformation of vocational education, driving the shift in educational paradigms from traditional “knowledge transmission” to modern “competency generation”, and providing robust intelligent support for cultivating high-quality technical and skilled talents that meet the demands of the new era.

Keywords

Embodied Intelligence Technology, Vocational Education, Skill Training, Personalized Learning

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在第四次工业革命与产业智能化转型的叠加效应下, 职业教育正面临技能迭代加速、产教融合壁垒显著、传统实训资源供给不足等结构性矛盾[1][2]。具身智能技术(Embodied AI)通过感知器和执行器与环境进行交互, 并根据环境的变化做出相应的决策和行动, 为破解这些困境提供了新的解决方案。例如, 虚拟现实技术可模拟真实工作环境, 提供沉浸式技能训练[3], 智能导师系统和协作机器人能根据学习者需求提供个性化指导, 提升实践教学效果。这些应用表明, 具身智能技术具有赋能职业教育的巨大潜力。然而, 具身智能技术在职业教育中的落地仍面临诸多挑战。技术层面, 成熟度和成本亟待解决; 教育层面, 教师角色转变、教学模式改革及资源建设需进一步探索; 伦理层面, 数据隐私、算法偏见及人机关系等问题也引发关注。因此, 系统探讨具身智能技术赋能职业教育的应用场景与挑战, 对推动职业教育数字化转型和高质量发展具有重要意义。

本研究基于文献研究方法, 系统梳理了具身智能技术(Embodied Artificial Intelligence)的核心概念体系及其关键技术构成, 重点聚焦于该技术在职业教育领域中的典型应用场景及其实现路径。通过深入分析

具身智能技术在职业教育中的实践应用，本文进一步探讨了其在技术落地过程中所面临的多维度挑战，包括技术层面的局限性、教育场景的适配性问题以及伦理规范的约束等。研究旨在通过理论探讨与实践分析相结合的方式，为职业教育领域的理论体系构建提供新的研究视角，同时为职业教育教学模式的创新实践提供可行性参考。此外，本研究还致力于拓展具身智能技术在教育应用领域的可能性边界，为其在职业教育场景中的深度应用提供创新性的思路与方向。

2. 具身智能技术概述

2.1. 概念与内涵

具身智能的思想萌芽可追溯至人工智能的起源时期。1950 年，图灵在其开创性论文《Computing Machinery and Intelligence》中，不仅为人工智能奠定了理论基础并提出了著名的图灵测试，更在文末前瞻性地展望了人工智能发展的两条可能路径：其一是专注于抽象计算能力(如下棋)的智能发展路线；其二是通过为机器配备先进的传感设备，使其具备与人类交互的能力，并能够像婴儿一样进行自主学习。这一前瞻性的论述，实质上已经勾勒出了非具身智能与具身智能两大研究范式的雏形[4]。中国计算机学会将具身智能定义为是一种基于物理身体进行感知和行动的智能系统，其通过智能体与环境的交互获取信息、理解问题、做出决策并实现行动，从而产生智能行为和适应性。斯坦福大学的李飞飞教授曾经指出，“具身的含义不是身体本身，而是与环境交互以及在环境中做事的整体需求和功能。”英伟达创始人黄仁勋在 ITF World 2023 半导体大会上表示，具身智能是能理解、推理、并与物理世界互动的智能系统，是人工智能的下一个浪潮。基于此提出，具身智能是指一种智能系统或机器(软硬件结合的智能体)能够通过感知和交互环境进行实时互动的能力。具身智能系统通常具备感知、决策和行动的能力，能够通过感知器和执行器与环境进行交互，并根据环境的变化做出相应的决策和行动。

具身智能的核心内涵可归纳为三大要素，即“本体 - 环境 - 迭代”的协同作用，见图 1 所示。首先，其基础是物理实体，例如仿人机器人、多足机器人、自动驾驶车辆等。这些实体作为数字与物理空间的桥梁，不仅具备环境感知、运动控制和任务执行等功能，其性能上限也直接决定了智能体的能力范围。其次，具身智能注重与外部环境的双向互动，既能够接收环境信息，又能通过主动行为改变环境状态，并在这一过程中实现自我优化与适应。最后，智能的持续提升是关键，借助大规模模型的知识处理能力，为各类物理实体注入智能潜能。

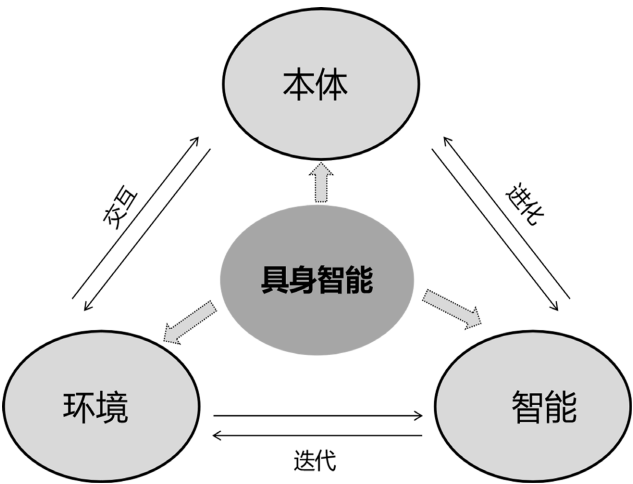


Figure 1. Conceptual connotation of embodied intelligence
图 1. 具身智能概念内涵图

2.2. 关键技术

具身智能的实现依赖于 AI 领域的多项技术, 包括计算机视觉、多模态感知融合技术、移动导航与规划技术、仿真和迁移技术等。计算机视觉技术是指通过计算机和相关设备模拟人类视觉系统的能力, 对图像或视频数据进行处理、分析和理解的一门技术[5]。它结合了图像处理、模式识别、机器学习、人工智能等多个领域的知识, 具身实体旨在从视觉数据中感知信息, 并做出相应的决策或行动。多模态感知融合技术是一种将来自不同传感器或数据源的信息进行整合和处理的技术, 旨在通过综合多种模态的数据, 获得更全面、准确的环境感知和理解能力, 其为具身智能的实现提供了关键的技术支持, 使其能够更高效地感知、理解和交互物理世界[6]。移动导航与规划技术是具身智能的核心组成部分之一, 是指机器人或自动化系统在复杂环境中自主定位、路径规划和导航的能力。通过感知环境和实时计算, 帮助机器人避开障碍物并找到从起点到终点的最优路径[7]。对于具身智能的实现, 移动导航与规划技术提供了空间感知和决策能力, 使得机器人能够在动态环境中自主行动、执行任务, 并与人类或其他设备高效协作。仿真技术是通过建立数学模型并借助计算机模拟真实或设想系统的行为, 以经济、安全且可重复的方式进行分析、优化和决策支持, 而迁移技术则是将数据、应用程序、计算任务或知识在不同硬件、软件环境或任务领域间进行转移, 以实现资源优化、知识复用、长期存取或服务连续性等目标[8]。两者均以计算机技术为核心工具, 前者侧重系统行为的虚拟再现与实验, 后者强调跨环境资源与信息的动态适配与流转。比如英伟达 Isaac Sim 项目训练机器人在虚拟环境中执行复杂任务, 然后将学到的策略迁移到真实机器人, 由此可见, 仿真技术能通过创建虚拟环境进行高效、安全的训练, 减少开发风险; 同时, 迁移学习技术将仿真环境中学到的策略迁移到真实世界, 解决数据稀缺问题[9]。

3. 具身智能技术赋能职业教育的应用场景

具身智能技术通过感知、决策、行动三个核心能力的结合, 为职业教育提供了更真实、沉浸、个性化和智能化的学习体验, 其应用场景主要分为基于感知能力、决策能力和行动能力的应用场景, 见图 2。

3.1. 基于感知能力的应用场景

具身智能系统通过传感器(如摄像头、麦克风、触觉传感器等)感知环境信息, 为职业教育提供更真实、沉浸式的学习体验。基于感知能力的应用场景主要体现在以下两个方面。

3.1.1. 沉浸式技能训练: 虚实融合的职业教育场景重构

具身智能系统通过虚拟现实(VR)、增强现实(AR)等技术构建高仿真工作场景, 使职业教育学生通过穿戴式设备(如触觉手套、动作捕捉装置)与虚拟环境深度交互, 实现“感知-行动-反馈”闭环训练。例如, 在虚拟焊接训练中, 系统可模拟电弧焊、气体保护焊等多种工艺, 实时监测学生的操作参数(电流、电压、焊接速度、焊枪角度), 并通过视觉(熔池动态模拟)、听觉(电弧音效)和触觉(力反馈装置模拟焊枪震动)多通道反馈操作质量。研究表明, 此类训练不仅降低实训材料消耗成本达 60%以上, 还能通过数据驱动的个性化指导(如错误动作矫正提示)显著提升技能掌握效率[10]。

3.1.2. 多模态学习体验: 多感官协同的技能习得机制

具身智能系统整合视觉、听觉、触觉等多模态信息, 突破传统单向知识传递模式, 构建“具身化”学习环境。例如, 在烹饪技能培训中, 触觉反馈设备可模拟刀具切割不同食材的阻力(如胡萝卜的硬度与西红柿的弹性), 热传导模拟器则通过温度反馈装置呈现火候变化, 结合 AR 视觉指导(如翻炒轨迹提示), 使学生形成肌肉记忆与操作直觉。实证表明, 多模态训练可使技能熟练度提升速度加快 40%, 且长期记忆保留率提高 30% [11]。

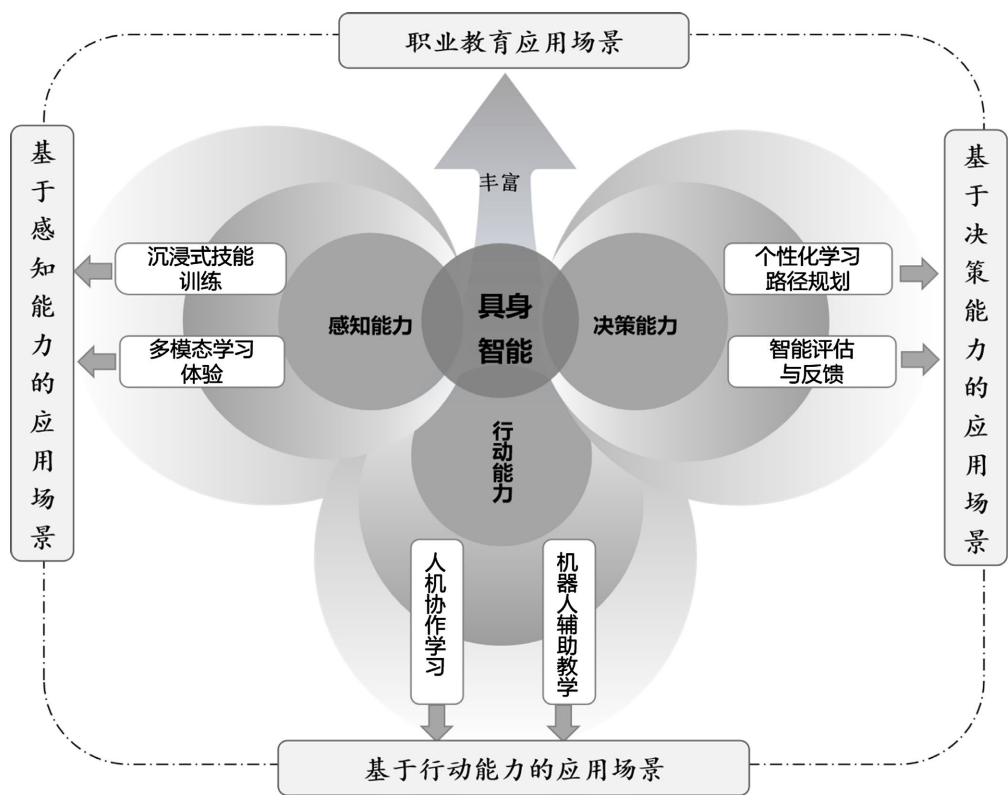


Figure 2. Application scenarios of embodied intelligence empowering vocational education
图 2. 具身智能赋能职业教育应用场景图

3.2. 基于决策能力的运用场景

具身智能系统通过整合多模态感知数据与智能算法，实现环境信息的动态分析与决策优化，为职业教育提供智能化、自适应的学习支持。其核心应用场景体现在以下两方面。

3.2.1. 个性化学习路径规划：动态优化与精准匹配

具身智能系统基于学习者的多维数据(如知识水平、操作熟练度、认知偏好)，通过学习者画像建模与动态知识网络分析，生成个性化的学习路径。例如，在智能制造实训中，系统通过华为昇思 MindSpore 框架的强化学习模块，实时分析学生操作机械臂的抓取精度、任务耗时等数据，动态调整实训项目难度与资源推荐顺序。若学生反复出现焊接角度偏差，系统会自动插入针对性训练模块，并结合多模态反馈强化技能掌握。研究表明，采用项目反应理论构建的动态知识掌握模型，可将学习效率提升 35%，课程完成率提高 25% [12]。

3.2.2. 智能评估与反馈：多模态数据驱动的实时诊断

具身智能系统通过融合动作捕捉、生理传感与机器学习算法，实现技能评估的高精度量化与实时反馈。例如，在编程教学中，AI 系统通过分析代码逻辑结构、调试耗时与错误类型，利用随机森林算法预测知识薄弱点，并推送定制化练习题。

3.3. 基于行动能力的运用场景

具身智能系统能够通过执行器(如机械臂、移动平台等)与环境进行交互，完成特定的任务，为职业教育提供更真实的实践机会。基于行动能力的应用场景主要体现在以下两个方面。

3.3.1. 机器人辅助教学：智能化实训场景构建

机器人辅助教学通过深度融合机器人技术与教学实践,为智能化实训场景构建提供了创新解决方案。首先,机器人能够模拟真实工作环境,如工业生产线、医疗手术等,为学习者打造沉浸式实训体验,提升其操作技能和应对复杂情况的能力。其次,机器人支持个性化学习路径,通过分析学习者的操作数据,识别薄弱环节并提供针对性指导,帮助其高效提升技能水平。此外,机器人还为高风险或高成本的实训项目提供了安全可靠的模拟平台,避免实际操作中的安全隐患和设备损耗,同时记录操作过程,为教学评估和改进提供数据支持。最后,机器人辅助教学还能推动教育资源均衡发展,通过远程控制技术,将优质教育资源输送到偏远地区,促进教育公平。总之,机器人辅助教学通过构建智能化实训场景,提升了教学质量、学习效率和安全性,推动了教育模式的创新与教育资源的均衡发展,未来将在教育领域发挥更加重要的作用。

3.3.2. 人机协作学习：协同任务驱动的能力培养

人机协作学习通过协同任务驱动的模式,为职业教育提供了高度仿真的实践机会。具身智能系统利用机械臂、移动平台等执行器与环境交互,使学习者能够与机器人协作完成特定任务,如工业装配、服务导航等。机器人负责高精度操作,学习者则专注于流程优化和问题解决,同时培养团队协作和沟通能力。系统通过实时反馈和数据分析,为学习者提供个性化指导,帮助其识别不足并优化学习路径。这种模式不仅适用于传统制造业和服务业,还可扩展到智能医疗、农业自动化等新兴领域,为职业教育注入创新活力。基于具身智能系统的人机协作学习,通过协同任务驱动的能力培养,为学习者打造了沉浸式、高互动性的实践环境,助力高素质技能人才的培养和职业教育的创新发展。

4. 具身智能技术赋能职业教育面临的挑战及应对策略

具身智能技术为职业教育带来了巨大的变革潜力,但其在实际应用中仍面临诸多挑战,主要体现在技术、教育和伦理三个层面。

4.1. 技术层面及应对策略

在技术层面主要有技术成熟度和成本问题挑战。尽管具身智能技术取得了显著进展,但其在职业教育中的应用仍处于探索阶段。例如,虚拟现实(VR)和增强现实(AR)技术在沉浸感和交互性方面仍有提升空间,触觉反馈设备的精度和稳定性也需进一步优化。此外,具身智能系统的感知、决策和行动能力尚未完全达到人类水平,难以应对复杂多变的职业教育场景。具身智能技术的硬件设备(如VR头盔、触觉反馈装置、协作机器人等)和软件系统(如仿真平台、智能导师系统等)成本较高,限制了其在职业教育中的大规模普及。此外,技术维护和更新的费用也增加了教育机构的负担。

针对技术成熟度问题应加强产学研协同创新,建立职业教育与技术研发机构的长效合作机制,推动具身智能技术在职业教育场景的定向优化;采用模块化设计思路,将具身智能系统分解为相对独立的功能模块,优先应用成熟度较高的模块,逐步迭代升级其他模块。针对成本问题可以探索“教育+产业”的双向赋能模式,吸引企业投资教育场景的技术应用,形成可持续的商业生态。

4.2. 教育层面及应对策略

在教育层面主要面临着教师角色转变、教学模式改革和教育资源建设等挑战。具身智能技术的引入对教师的角色提出了新的要求。教师需要从传统的知识传授者转变为学习引导者和技术支持者,掌握新技术并灵活运用于教学实践中。然而,部分教师可能缺乏相关技术知识和应用能力,难以适应这一转变;具身智能技术推动了职业教育教学模式的创新,例如沉浸式学习、个性化学习和协作学习等。但是传统

的教学模式(如课堂讲授、实训操作等)仍占主导地位,如何将新技术与传统模式有机结合,实现教学效果的优化,是教育层面面临的重要挑战;具身智能技术的应用需要配套的教育资源支持,例如虚拟实训课程、智能评估系统、多模态学习材料等。然而,目前相关资源的数量和质量仍显不足,难以满足职业教育的多样化需求。

针对教师角色转变问题应建立“技术导师+学科教师”的协同教学团队,通过分工合作降低教师适应门槛,开发教师友好型具身智能教学工具,简化技术操作流程,降低使用难度;建立教师创新社区,促进经验分享与协同创新,加速最佳实践的传播。针对教学模式变革的问题可以开展基于具身智能技术的教学模式实证研究,系统评估不同模式的教学效果,形成循证改进机制;推动教学评价体系创新,将技术支持下的能力表现纳入评价维度,引导教学模式变革。针对教育资源不足的问题可以构建行业主导的资源建设联盟,引入产业资源与标准,确保资源的专业性与实用性;设计模块化、可重组的资源架构,提高资源开发效率与复用率;建立资源质量评估与迭代优化机制,持续提升资源质量。

4.3. 伦理层面及应对策略

在伦理层面主要面临着数据隐私、算法偏见和人机关系三个方面的挑战。具身智能系统在运行过程中会采集大量学习者的个人数据,这些数据可能涉及隐私问题。如何在使用数据的同时保护学习者的隐私权,是伦理层面需要关注的重要问题;具身智能系统的决策和推荐功能依赖于算法模型,而这些模型可能存在偏见,导致不公平的学习体验或评价结果。如何确保算法的公平性和透明度,是伦理层面需要解决的难题;具身智能技术的应用可能改变学习者与教师、学习者与学习者之间的关系。例如,过度依赖智能系统可能导致学习者与教师之间的互动减少,影响教育的人文关怀。如何平衡人机关系,保持教育的温度,是伦理层面需要思考的问题。

针对数据隐私问题建立职业教育数据采集与使用的伦理准则,明确数据权属、使用边界与保护责任;采用数据最小化原则,仅收集教育必需的数据,并实施数据脱敏与匿名化处理;提升学习者数据素养,保障其对个人数据的知情权、同意权与控制权。针对算法偏见问题应增强算法的可解释性设计,使教师与学习者能够理解系统决策的依据;建立人机协同决策框架,保留人类对关键教育决策的最终裁量权;推动多元价值观融入算法设计,避免单一标准导致的评价偏见。针对人机关系的问题可以构建“技术赋能、人文主导”的职业教育理念,强调人文关怀在职业教育中的核心地位。

5. 结语与展望

具身智能技术通过感知、决策和行动能力的深度融合,为职业教育带来了革命性的变革。其在技能训练、个性化学习、协作学习和智能评估等方面的应用,显著提升了职业教育的效率和质量。例如,沉浸式技能训练降低了实训成本并提高了学习效果,个性化学习路径规划优化了学习体验,而机器人辅助教学和人机协作学习则为学习者提供了更真实的实践机会。然而,技术的成熟度、教育模式的适应性以及伦理问题的解决仍是当前面临的主要挑战。

展望未来,具身智能技术将推动职业教育从“知识传递”向“能力生成”跃迁,为职业教育的数字化转型和高质量发展提供了强有力的支持,具有广阔的应用前景。在未来研究方向中,具身智能技术支持下的职业教育评价体系构建研究尤为关键。多模态学习行为分析的能力评价框架将成为重点发展领域,该框架旨在突破传统评价的局限性,通过系统采集学习者在虚实融合环境中的多维行为数据,构建更为精准、立体的能力画像。具体而言,这一评价体系将整合动作捕捉、眼动追踪、操作轨迹、力度控制等微观层面的表现数据,结合认知负荷、注意力分配等生理指标,通过深度学习算法建立专家模型与学习者行为特征的映射关系。未来研究将着力探索多源异构数据的融合机制、基于证据的能力推理模型,以及

能力发展的动态预测系统, 构建具有生态效度的过程性评价框架。同时, 该评价体系还需要解决数据标准化、评价指标权重确定以及模型可解释性等技术挑战, 确保评价结果的科学性与可信度。通过这一评价体系的构建, 将为职业教育中的精准教学与个性化学习提供数据支撑, 推动具身智能技术从辅助工具向教育决策支持系统的转变。

参考文献

- [1] 邱飞岳, 刘雪冰, 张彭丽, 等. 职业教育数字化转型: 内涵、问题与路径[J]. 职业技术教育, 2024, 45(9): 32-38.
- [2] 孟亭含. 职业教育数字化转型的价值、困境及实施路径[J]. 教育理论与实践, 2024, 44(33): 28-32.
- [3] 常馨予. 基于虚拟现实技术的师范生教学技能训练活动设计与效果研究[D]: [硕士学位论文]. 上海: 华东师范大学, 2023.
- [4] Turing, A.M. (2007) Computing Machinery and Intelligence. In: *Parsing the Turing Test*, Springer, 23-65.
https://doi.org/10.1007/978-1-4020-6710-5_3
- [5] 刘恒, 高新燕, 苏新新, 等. 基于计算机视觉技术的人体姿态跟踪方法[J/OL]. 软件导刊, 1-11.
<http://kns.cnki.net/kcms/detail/42.1671.TP.20250217.1550.002.html>, 2025-02-18.
- [6] 李萍, 王丽丽. 国内多模态技术的研究现状与发展趋势: 基于 CiteSpace 的可视化分析[J]. 智能计算机与应用, 2025, 15(1): 194-202.
- [7] 田建林, 陈晓波, 张文昌, 等. 移动机器人自主导航技术研究进展[J/OL]. 自动化技术与应用, 1-6.
<http://kns.cnki.net/kcms/detail/23.1474.TP.20241230.0908.016.html>, 2025-02-18.
- [8] 郭丽媛. 仿真技术在电子电路设计中的应用[J]. 集成电路应用, 2024, 41(8): 244-245.
- [9] Makoviychuk, V., Wawrzyniak, L., Guo, Y., Lu, M., Storey, K., Macklin, M. and State, G. (2021) Isaac Gym: High Performance GPU-Based Physics Simulation for Robot Learning.
- [10] Lavrentieva, O., Arkhypov, I., Kuchma, O. and Uchitel, A. (2020) Use of Simulators Together with Virtual and Augmented Reality in the System of Welders' Vocational Training: Past, Present, and Future.
- [11] Tene, T., Bonilla García, N., Coello-Fiallos, D., Borja, M. and Vacacela Gomez, C. (2024) A Systematic Review of Immersive Educational Technologies in Medical Physics and Radiation Physics. *Frontiers in Medicine*, **11**, Article 1384799. <https://doi.org/10.3389/fmed.2024.1384799>
- [12] Jiang, B., Li, X., Yang, S., Kong, Y., Cheng, W., Hao, C., *et al.* (2022) Data-Driven Personalized Learning Path Planning Based on Cognitive Diagnostic Assessments in MOOCs. *Applied Sciences*, **12**, Article 3982.
<https://doi.org/10.3390/app12083982>