

李彦哲. 深度强化学习在游戏开发领域的研究与应用[J]. 智能计算机与应用, 2025, 15(9): 161–167. DOI: 10. 20169/j. issn. 2095–2163. 250925

深度强化学习在游戏开发领域的研究与应用

李彦哲

(中南民族大学 数学与统计学学院, 武汉 430074)

摘要: 深度强化学习结合了深度学习的感知能力和强化学习的决策制定能力, 已成为游戏开发的关键技术。为全面分析深度强化学习在游戏开发领域的研究和应用, 本文分析了深度强化学习的关键技术和算法, 综述了深度强化学习在游戏智能体训练、环境设计和性能优化等游戏开发领域的应用和研究进展, 探讨了深度强化学习在游戏开发领域的技术优势、困难挑战以及研究展望, 为游戏开发带来更多的可能性和机会。

关键词: 深度强化学习; 深度学习; 智能体训练; 游戏开发

中图分类号: TP311

文献标志码: A

文章编号: 2095–2163(2025)09–0161–07

Research and application of deep reinforcement learning in game development

LI Yanzhe

(College of Mathematics and Statistics, South-Central Minzu University, Wuhan 430074, China)

Abstract: Deep reinforcement learning combines the perception ability of deep learning and the decision-making ability of reinforcement learning, and has become the key technology in game development. In order to comprehensively analyze the research and application of deep reinforcement learning in the field of game development, this paper analyzes the key technologies and algorithms of deep reinforcement learning, reviews the application and research progress of deep reinforcement learning in game development fields such as game agent training, environment design and performance optimization. Subsequently, this paper discusses the technical advantages, difficulties and challenges of deep reinforcement learning in the field of game development, and brings more possibilities and opportunities for game development.

Key words: deep reinforcement learning; deep learning; agent training; game development

0 引言

游戏开发是人工智能研究的重要领域之一, 其中深度强化学习的应用已经显示出巨大的潜力^[1–4]。深度强化学习结合了深度学习的强大表征能力和强化学习的决策优化能力^[5], 能够处理复杂的决策问题, 并在不断与环境互动中学习最佳行动策略, 为游戏开发者提供了一种新方法训练智能代理, 也为游戏玩家提供更具挑战性和逼真性的游戏体验。

深度强化学习在游戏开发领域已经取得了一系列令人瞩目的研究成果和应用案例。通过深度强化学习, 智能代理可以学习到高水平的游戏技能, 如在电子竞技游戏中击败人类顶级选手、在角色扮演游戏中自主决策任务等。此外, 深度强化学习还可以

用于游戏内容的生成和优化, 如自动生成游戏关卡、调整游戏平衡性等。然而, 目前相关研究主要围绕应用深度强化学习解决游戏开发的某个阶段, 尚未有关于深度强化学习游戏开发应用研究的全面分析。因此, 本文综述深度强化学习在游戏智能体训练、环境设计和性能优化中的应用和研究现状, 探讨深度强化学习在游戏开发领域的技术优势、困难挑战以及研究前景, 为推动深度强化学习在游戏开发领域的研究与应用奠定基础。

1 深度强化学习游戏开发应用研究现状

深度强化学习在游戏智能体训练、游戏环境设计和游戏性能优化等方面都取得了一定的研究进展。通过自主学习、多智能体协作与竞争、逆强化学习、多任务学习等技术, 深度强化学习为游戏开发者

提供了强大的工具和方法,以实现更智能化、个性化和有趣的游戏体验。同时,研究人员还在游戏环境设计和游戏性能优化方面进行了一些探索,以提高游戏的可玩性和性能。

1.1 游戏智能体的训练

基于深度强化学习的游戏智能体训练是近年来的研究热点,已成功应用于电子竞技游戏、策略游戏、角色扮演游戏等游戏开发领域。研究人员使用深度神经网络来近似值函数或策略函数,通过与环境的交互来训练智能体,在游戏中实现特定目标,设计出更智能、适应性更强的游戏智能体,提供更具挑战性和逼真性的游戏体验。研究人员还探索了多智能体协作与竞争、逆强化学习和多任务学习等技术,以提高智能体的学习效率和决策能力。游戏智能体训练的决策模型主要包括基于价值函数的模型、基于策略梯度的模型、模型预测控制、组合模型等。

(1) 基于价值函数的模型。基于价值函数的模型通常使用深度神经网络来近似值函数,进而评估状态或动作的价值,并选择具有最高价值的动作进行决策。常见的基于价值函数的模型包括 Q-learning、深度 Q 网络 (DQN) 和双重深度 Q 网络 (Double DQN) 等。

2015 年,DeepMind 团队通过结合深度学习和 Q 学习提出 DQN 算法,通过使用经验回放和固定 Q 目标技术解决了训练过程中的不定性和发散问题,在多款雅达利游戏上取得了超越人类的表现^[6]。2017 年,Bellemare 等学者^[7]引入了价值分布的概念对 DQN 算法进行改进,提出 C51 算法,通过学习预期回报的分布而不是单一的期望值来提升游戏智能体训练效果。2018 年,Hessel 等学者^[8]集成了多种提升 DQN 性能的技术,如双重 Q 学习、优先经验回放等,提出了集成智能体 Rainbow 算法,给雅达利 2600 系列的经典视频游戏提供了最先进的性能。上述研究展示了如何在游戏智能体训练中应用基于价值函数的模型来提升决策质量,以及如何通过结合深度强化学习的强大功能来处理高维的状态空间。

(2) 基于策略梯度的模型。基于策略梯度的模型通常使用深度神经网络来参数化策略函数,该函数可以将状态映射到动作的概率分布,通过优化策略函数的参数来最大化累积奖励,并选择具有最高概率的动作进行决策。常见的基于策略梯度的模型包括深度确定性策略梯度 (DDPG)、近似策略优化算法 (PPO)、异步优势行动者评论家算法 (A3C) 及信赖域策略优化算法 (TRPO) 等。

2018 年,OpenAI 团队使用 PPO 算法设计的 OpenAI Five 在多人团战视频游戏 Dota2 中经过大规模训练后,战胜了由职业玩家组成的队伍,且对战在线玩家有 99.4% 的胜率^[9-10]。DeepMind 研发的 AlphaStar 是首个在实时战略游戏《星际争霸 II》中达到人类职业选手水平的 AI 系统,且在对战人类顶级玩家方面有着超过 99.8% 的胜率,其采用了策略梯度方法的变体,并结合了监督学习和强化学习来进行训练^[11]。上述研究展示了深度强化学习在复杂决策环境中的强大能力,尤其是在处理高维动作空间和多智能体协作或竞争时。

(3) 模型预测控制。模型预测控制 (Model Predictive Control, MPC) 是一种用于控制动态系统的方法,其中智能体通过与系统的模型进行交互来学习模型,并使用模型对未来的状态和奖励实现预测。通过优化控制策略,可以选择最优的控制动作序列,以使系统达到期望的性能指标。

在游戏智能体训练领域,MPC 尚未像策略梯度或价值函数那样广泛应用,主要是因为 MPC 通常需要准确的模型预测,而这在复杂的游戏环境中很难获得。DeepMind 研发的 MuZero 算法,在没有访问环境模型的情况下,通过自我学习其内部模型并结合 MPC 的思想进行规划和决策,成功在多个复杂游戏环境(如国际象棋、将棋、围棋和雅达利视频游戏等)上取得超人的表现^[12]。MuZero 通过学习模拟未来可能的行动序列并优化决策,体现了 MPC 在强化学习中的潜力。

(4) 组合模型。组合模型是将多种学习决策和模型结合起来使用的方法,以利用不同模型的优势并弥补单一模型的不足,可以通过联合训练或集成方法来实现。例如,可以将基于价值函数的模型和基于策略梯度的模型结合起来,以充分利用各自的优势。

2016 年,DeepMind 团队继续推出围棋 AI、AlphaGo 及 AlphaZero,其使用了基于价值函数的方法来评估围棋棋局的位置,结合策略网络进行蒙特卡洛树搜索,通过自我对弈学习无需依赖人类知识即可在国际象棋、将棋和围棋等游戏中达到超人的水平,在 2016 年击败了世界围棋冠军李世石。2018 年,OpenAI 团队研发的 Dota2 AI,通过结合策略梯度方法、监督学习和进化策略,能够在复杂、多智能体的环境中与人类顶尖选手竞争^[16]。Bethesda Game Studios 研发的《上古卷轴 V:天际》,可以使用基于价值函数的模型来评估状态的价值,并使用基

于规则的模型来制定任务决策和角色行为^[17]。上述研究中智能体不仅采用多种深度强化学习算法来处理决策和学习,还整合了模型预测和规划来优化长期策略。

1.2 游戏环境的设计

研究人员致力于设计更加复杂和真实的游戏环境,以挑战深度强化学习算法的智能体。游戏环境包括更多的随机性、不确定性和复杂的物理规律,以促进智能体的学习和泛化能力。研究人员通过自动生成游戏关卡、调整游戏平衡性等方式,利用深度强化学习来改善游戏的可玩性和挑战性。此外,还研究探索如何使用深度强化学习来设计更逼真和智能的非玩家角色(NPC),以提供更丰富和沉浸式的游戏体验。深度强化学习在游戏环境设计方面的应用主要包括游戏难度调整、游戏内容生成以及游戏平衡调整等。

(1)游戏难度调整。深度强化学习利用智能体来动态调整游戏的难度,以适应不同玩家的技能水平。通过监控玩家的表现和反馈,使用深度强化学习算法来自动调整游戏的难度,以确保玩家始终面临适当的挑战。自适应的难度调整可以提高游戏的可玩性、娱乐性以及挑战性。

2010年, Pedersen 等学者^[18]提出的自适应游戏机制,利用深度强化学习训练智能体,通过改变游戏中的元素来匹配玩家的技术水平,从而保持游戏的吸引力。2017年, Monterrat 等学者^[19]开发了一个名为“The Ruby Witch”的小游戏,应用深度强化学习智能体来自适应玩家的能力,动态调整游戏难度。智能体通过观察玩家的表现和行为,预测其技能水平,并据此调整游戏难度。上述研究展示了深度强化学习在个性化游戏体验和保持玩家长期参与度方面的潜力。通过自适应难度调整,游戏开发者可以为不同水平的玩家提供更加定制化的体验。

(2)游戏内容生成。深度强化学习可以用于生成游戏中的内容,如地图、关卡、任务、故事情节等。利用深度强化学习,可以创建一个智能体来探索生成内容的高维参数空间,并通过奖励机制来引导其生成具有新颖性、多样性和个性化的游戏内容,可以实现自动生成和无限扩展的游戏体验。

2011年, Togelius 等学者^[20]提出了使用进化算法结合强化学习来自动生成游戏关卡,使得智能体可以根据玩家的技能水平自动设计出挑战性适中的游戏关卡,以提供挑战性和趣味性。2015年, Nielsen 等学者^[21]提出了一个使用深度强化学习智

能体的框架,该智能体通过尝试不同的游戏内容配置来学习如何生成吸引人的游戏关卡。2018年, Justesen 等学者^[22]使用深度强化学习来自动生成游戏内容,使得游戏智能体能够根据游戏玩法的多样性和平衡性来设计新的游戏元素。上述研究表明深度强化学习在游戏内容生成方面的应用潜力,可以帮助游戏开发者自动化和个性化游戏开发过程,创造出与玩家互动并提供个性化体验的游戏智能体。

(3)游戏平衡调整。深度强化学习具有自动化、可持续学习和自适应性,可以用于调整游戏的平衡性。实施游戏平衡调整时,深度强化学习会经历反复的学习过程,其中智能体在游戏环境中进行探索和实验,以寻找最优的游戏设置。在学习过程中,智能体会收集数据,评估游戏平衡,并根据奖励信号调整其策略,可以确保不同策略和角色之间的平衡,提供公平和有趣的游戏体验。

2019年, DeepMind 团队开发了一个多智能体训练框架,用于训练足球游戏智能体。该框架不仅训练智能体执行特定的足球技能,而且还通过自我对弈来调整团队协作策略,从而实现游戏内的战术平衡^[23]。2020年, Moon 等学者^[24]采用了深度强化学习来动态调整多人在线战斗竞技场(MOBA)游戏中的游戏平衡,其开发的系统能够根据玩家的表现和游戏统计数据来调整游戏内的经济系统、角色属性等,以达到更好的平衡。上述研究展示了深度强化学习在游戏平衡调整方面的应用,能够帮助开发者根据玩家行为和游戏数据来微调游戏机制,以维护竞争性和提升游戏的整体体验。

1.3 游戏性能的优化

深度强化学习在游戏性能优化方面也有一些研究。研究人员通过优化深度强化学习模型的结构和参数,以提高智能体的决策速度和效率。此外,还有一些研究探索如何使用分布式深度强化学习来加速训练过程,并实现更高水平的游戏智能,以提升游戏的娱乐性和可玩性。通过利用深度强化学习的能力,可以提高游戏的图形渲染、物理模拟、网络通信和资源管理等方面的性能。

(1)游戏图形优化。深度强化学习可以用于提升游戏图形渲染效果、增强视觉体验和优化性能。深度强化学习能够学习何时以及如何调整分辨率以优化性能和视觉效果,智能体通过学习基于当前的硬件性能、游戏场景复杂度和玩家偏好来做出最佳的渲染决策,用于动态调整游戏的渲染参数(如细节层次、纹理分辨率、光影效果等),以提升游戏性

能同时保持视觉效果。

2020年,Xiao等学者^[25]使用深度强化学习来自动测试和优化游戏引擎中的性能瓶颈。通过这种方式,AI智能体可以学会识别和调整会影响游戏性能的图形设置,从而在不牺牲游戏体验的情况下提升性能。2019年,Facebook的AI研究团队使用深度强化学习自动调整虚拟现实(VR)游戏的图形渲染设置,以在保持视觉质量的同时最大化帧率和减少延迟,其可以动态地根据硬件性能和当前游戏情景调整渲染参数,如纹理质量和野范围^[26]。

(2)游戏物理模拟优化。深度强化学习在游戏物理模拟优化方面的应用通常是为了提升游戏中物理行为的真实性,同时确保计算效率,以便在不牺牲游戏性能的情况下提供沉浸式的体验。使用深度强化学习进行游戏物理模拟优化涉及数据的收集和处理、智能体的训练、以及模型的部署等。实现游戏物理模拟优化,通常需要设计一个可靠的奖励函数,来指导智能体学习如何平衡物理模拟的真实性和计算效率。

2019年,由Ghost Games开发并由Electronic Arts发行的赛车游戏《极品飞车:热度》,使用深度强化学习来优化游戏中的车辆物理模拟,提高驾驶体验和真实感^[27]。《火箭联盟》是一款结合了足球和赛车元素的流行电子游戏,其通过深度强化学习来优化游戏中的球体物理模拟,提高球员和球场的交互效果^[28]。DeepMind团队探索了深度强化学习在改善机器人仿真物理模拟中的应用,其中智能体通过与环境的互动学习最优控制策略,可以减少模拟中不必要的计算,同时保持模拟的准确性^[29]。尽管这项研究的主要焦点是机器人运动,但同样可以适用于游戏物理模拟的场景,以优化性能并降低计算成本。Peng等学者^[30]使用深度强化学习来优化实时物理动画,其训练智能体学习如何在保持视觉真实性的同时简化物理动画中的计算过程,这对于需要大量物理计算的游戏开发来说尤其有价值。上述研究指出了深度强化学习在游戏物理模拟优化中的潜力,尽管在游戏开发领域的应用还不是非常广泛。通过智能地减少物理模拟的计算量而不牺牲其准确度,深度强化学习有望在未来为游戏开发应用带来性能上的重大提升。

(3)游戏网络优化。游戏网络优化对于提供流畅且响应迅速的游戏体验至关重要,尤其是在延迟敏感的游戏如第一人称射击(FPS)或多人在线战斗竞技场(MOBA)游戏中。深度强化学习可以用于优

化游戏中的网络通信和多人游戏体验,为提升多玩家在线游戏的网络性能和稳定性,必须精心设计奖励机制,使得智能体能够在提高网络性能的同时考虑到成本和资源使用效率。

2017年,Bluehole公司开发的《绝地求生》,通过深度强化学习来优化游戏中的网络同步和抗延迟,提高多人游戏的流畅性和公平性^[31]。Mao等学者^[32]提出了一个深度强化学习框架Pensieve,用于自适应视频流质量的选择,减少缓冲并改善用户体验。虽然该研究主要关注视频流,但其背后的原理同样可以应用于游戏网络优化,特别是在云游戏和实时多人在线游戏中。DeepMind团队与谷歌合作,使用深度强化学习来优化数据中心的冷却效率,该技术也可以类比到游戏服务器的资源管理和网络负载均衡上^[33]。尽管直接将深度强化学习应用于游戏网络优化的文献可能较少,但相关技术的潜力和发展趋势表明,深度强化学习可能将随着云游戏和其他网络密集型游戏服务的兴起,在未来的游戏网络优化中扮演重要角色。

(4)游戏资源管理优化。游戏资源管理包括但不限于CPU和GPU的计算资源、内存使用、电源管理以及网络带宽等。深度强化学习可以用于优化游戏中的资源管理和加载。通过训练智能体与环境进行交互,可以学习到一种优化策略,使得游戏在资源管理和加载方面更加高效和智能,提高游戏的加载速度和运行效率,确保在有限的硬件资源下能够提供最佳的玩家体验。

2020年,腾讯人工智能实验室研发了觉悟AI,其能够学习如何收集资源、扩展基地、建造单位,并与敌人作战,在游戏王者荣耀上击败顶级的职业电竞玩家^[34]。Tytarenko^[35]利用深度强化学习动态地管理游戏中的LOD,以在不同距离和视角下调整模型的细节级别,可以根据玩家的位置和相机视角来决定何时加载或卸载不同细节级别的模型,从而优化游戏的性能和资源利用。Gamal等学者^[36]利用深度强化学习动态地预测玩家的行为和游戏场景,以提前加载可能需要的资源,从而减少游戏中的延迟和卡顿,可以根据当前的游戏状态和玩家的行为来预测未来可能需要的资源,并在合适的时机进行预加载。

2 深度强化学习游戏开发应用优势与挑战

2.1 深度强化学习游戏开发应用优势分析

深度强化学习在游戏开发领域具有自主学习能

力和自主智能代理、个性化游戏体验、游戏内容自动生成以及游戏平衡性优化等优势,可以提升游戏的智能性、可玩性和娱乐性,为玩家提供更好的游戏体验。

(1)自主学习能力和自主智能代理。深度强化学习可以通过与环境的交互来自主学习游戏策略,无需人工设计复杂的规则,通过试错和反馈机制不断优化策略,逐渐提高游戏性能;也可以训练出具备自主决策能力的智能代理,使得游戏中的 NPC 角色或虚拟对手能够更加智能化地与玩家进行互动,可以提供更具挑战性和逼真性的游戏体验。

(2)个性化游戏体验。深度强化学习可以根据玩家的行为和偏好,实时调整游戏内容和难度,以满足不同玩家的需求和偏好,从而提供个性化的游戏体验,可以增加玩家的参与度和满意度。

(3)游戏内容自动生成。深度强化学习可以用于生成游戏中的关卡、道具、敌人等内容。通过训练模型来学习游戏规则和设计原则,从而自动生成丰富多彩的游戏内容,可以提高游戏的可玩性和重复性,为游戏玩家提供更多的游戏选择和体验。

(4)游戏平衡性优化。深度强化学习通过对游戏进行模拟和优化,可以调整游戏参数和规则,确保游戏中的各种元素和机制之间的平衡和公平性,使得游戏更加公平和有趣。

2.2 深度强化学习游戏开发应用困难挑战

虽然利用深度强化学习解决游戏开发问题具有诸多优势和广阔前景,但与实际应用仍有差距,在训练效率和样本获取、模型泛化能力、人机交互和用户体验以及解释性和可解释性等方面仍面临着挑战。

(1)训练效率和样本复杂性。深度强化学习需要大量的训练数据和计算资源,而在游戏开发中,获取大规模的训练数据可能会面临困难。此外,游戏中的状态空间和动作空间通常非常庞大,需要更多的数据来进行有效的训练,也增加了训练的复杂性。

(2)游戏多样性和泛化能力。游戏中存在各种各样的场景、规则和任务,要求深度强化学习模型具备良好的泛化能力,能够适应不同的游戏环境和任务。然而,游戏的多样性可能导致模型在新游戏中表现不佳,需要进一步研究如何提高模型的泛化能力。

(3)人机交互和用户体验。深度强化学习在游戏中的应用需要考虑与玩家的交互和用户体验。模型的决策和行为应该与玩家的期望和意图相符合,以提供良好的游戏体验。如何设计合适的交互方式

和反馈机制,使得玩家能够与模型进行有效的互动,并提供令人满意的游戏体验,是目前需要解决的问题。

(4)解释性和可解释性。深度强化学习模型通常是黑盒模型,难以解释其决策过程和内部机制。在游戏开发中,解释性和可解释性对于开发者和玩家来说都非常重要。因此,需要研究如何提高深度强化学习模型的解释性和可解释性,以便开发者和玩家能够理解和信任模型的决策。

3 深度强化学习游戏开发应用研究展望

深度强化学习在游戏开发领域具有广阔的应用前景。未来的研究可以在多任务学习、逆强化学习、多智能体协作、可解释性和可视化等方面进行深入探索,以进一步推动游戏智能化的发展,提升游戏的智能性、可玩性和娱乐性。同时,与其他技术的融合也将为游戏开发带来更多创新和可能性。

(1)多任务学习。目前的深度强化学习研究主要集中在单一任务上,但实际游戏中存在多个任务和目标。未来的研究可以探索如何将深度强化学习应用于多个游戏任务的学习中。通过共享模型参数和知识传递,可以提高模型的泛化能力和学习效率,将使得智能代理能够在不同类型的游戏中更好地适应和学习,实现更复杂的游戏策略。

(2)逆强化学习。逆强化学习是一种从观察到的行为中推断出潜在奖励函数的方法。在游戏开发中,逆强化学习可以用于从专家玩家的行为中学习更好的奖励函数,从而提高智能代理的性能。未来的研究可以探索如何将逆强化学习与深度强化学习相结合,以生成更具挑战性和逼真性的智能对手,实现更高水平的游戏智能。

(3)多智能体协作。游戏中通常存在多个智能代理之间的协作和竞争关系。未来的研究可以探索如何使用深度强化学习来训练多个智能代理进行协作和竞争,以实现更复杂和逼真的游戏体验。

(4)可解释性和可视化。深度强化学习模型通常是黑盒模型,难以解释其决策过程和内部机制。未来的研究可以探索如何提高模型的解释性,并开发可视化工具来帮助开发者和玩家理解智能代理的行为和决策。

(5)深度强化学习与其他技术的融合。深度强化学习可以与其他技术如图像处理、自然语言处理等进行融合,以进一步提高游戏开发的效果。例如,结合计算机视觉技术可以实现基于图像的游戏智能

代理,结合自然语言处理技术可以实现与玩家的自然对话交互。

4 结束语

深度强化学习结合深度学习的感知能力和强化学习的决策制定能力,在游戏开发领域具有巨大的潜力和广阔的应用前景。由于深度强化学习自主学习能力和自主智能代理、个性化游戏体验、游戏内容自动生成以及游戏平衡性优化等优势,在游戏智能体训练、游戏环境设计以及游戏性能优化等方面得到广泛应用,其能够提高游戏智能体的决策质量,设计更具挑战性和逼真性的游戏环境,以及提升游戏整体的性能和用户体验。然而,深度强化学习在训练效率和样本获取、模型泛化能力、人机交互和用户体验以及解释性和可解释性等方面仍面临着挑战。未来需进一步加强深度强化学习在多任务学习、逆强化学习、多智能体协作、可解释性和可视化等方面的研究,为玩家提供更加丰富、沉浸和个性化的游戏体验。

参考文献

- [1] BUI V H, HUSSAIN A, SU W. A dynamic internal trading price strategy for networked microgrids: a deep reinforcement learning-based game-theoretic approach[J]. IEEE Transactions on Smart Grid, 2022, 13(5): 3408–3421.
- [2] CHEN Miaojiang, LIU Wei, ZHANG Ning, et al. GPDS: A multi-agent deep reinforcement learning game for anti-jamming secure computing in MEC network[J]. Expert Systems with Applications, 2022, 210: 118394.
- [3] 沈宇, 韩金朋, 李灵犀, 等. 游戏智能中的 AI: 从多角色博弈到平行博弈[J]. 智能科学与技术学报, 2020, 2(3): 205–213.
- [4] 熊水彬. 基于深度强化学习的对抗攻击和防御在动态视频中的应用[J]. 通信技术, 2023, 56(9): 1115–1120.
- [5] 杨思明, 单征, 丁煜, 等. 深度强化学习研究综述[J]. 计算机工程, 2021, 47(12): 19–29.
- [6] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015, 518(7540): 529–533.
- [7] BELLEMARE M G, DABNEY W, MUNOS R. A distributional perspective on reinforcement learning[C]//Proceedings of the International Conference on Machine Learning. New York: ACM, 2017: 449–458.
- [8] HESSEL M, MODAYIL J, HASSELT V H, et al. Rainbow: Combining improvements in deep reinforcement learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence. New Orleans, USA: AAAI, 2018: 2668–3603.
- [9] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal policy optimization algorithms[J]. arXiv preprint arXiv, 1707.06347, 2017.
- [10] BERNER C, BROCKMAN G, CHAN B, et al. Dota 2 with large scale deep reinforcement learning[J]. arXiv preprint arXiv, 1912.06680, 2019.
- [11] VINYALS O, BABUSCHKIN I, CZARNECKI W M, et al. Grandmaster level in StarCraft II using multi-agent reinforcement learning[J]. Nature, 2019, 575(7782): 350–354.
- [12] SCHRITTWIESER J, ANTONOGLOU I, HUBERT T, et al. Mastering atari, go, chess and shogi by planning with a learned model[J]. Nature, 2020, 588(7839): 604–609.
- [13] SILVER D, HUANG A, MADDISON C J, et al. Mastering the game of Go with deep neural networks and tree search[J]. Nature, 2016, 529(7587): 484–489.
- [14] SILVER D, SCHRITTWIESER J, SIMONYAN K, et al. Mastering the game of go without human knowledge[J]. Nature, 2017, 550(7676): 354–359.
- [15] SILVER D, HUBERT T, SCHRITTWIESER J, et al. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play[J]. Science, 2018, 362(6419): 1140–1144.
- [16] FANGASADHA E F, SOEROREDJO S, GUNAWAN A A S. Literature Review of OpenAI Five's Mechanisms in Dota 2's Bot Player [C]//Proceedings of 2022 International Seminar on Application for Technology of Information and Communication (iSemantic). Piscataway, NJ: IEEE, 2022: 183–190.
- [17] BJØRKELO K A. Elves are jews with pointy ears and gay magic: White nationalist readings of the elder scrolls V: Skyrim [J]. Game Studies, 2020, 20: 232051778.
- [18] PEDERSEN C, TOGELIUS J, YANNAKAKIS G N. Modeling player experience for content creation[J]. IEEE Transactions on Computational Intelligence and AI in Games, 2010, 2(1): 54–67.
- [19] MONTEERRAT B, LAVOUÉ É, GEORGE S. Adaptation of gaming features for motivating learners[J]. Simulation & Gaming, 2017, 48(5): 625–656.
- [20] TOGELIUS J, YANNAKAKIS G N, STANLEY K O, et al. Search-based procedural content generation: A taxonomy and survey[J]. IEEE Transactions on Computational Intelligence and AI in Games, 2011, 3(3): 172–186.
- [21] NIELSEN T S, BARROS G A, TOGELIUS J, et al. Towards generating arcade game rules with VGDL [C]//Proceedings of 2015 IEEE Conference on Computational Intelligence and Games (CIG). Piscataway, NJ: IEEE, 2015: 185–192.
- [22] JUSTESEN N, TORRADO R R, BONTRAGER P, et al. Illuminating generalization in deep reinforcement learning through procedural level generation [J]. arXiv preprint arXiv, 1806.10729, 2018.
- [23] JADERBERG M, CZARNECKI W M, DUNNING I, et al. Human-level performance in 3D multiplayer games with population-based reinforcement learning[J]. Science, 2019, 364(6443): 859–865.
- [24] MOON H S, SEO J. Dynamic difficulty adjustment via fast user adaptation[C]//Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology. New York: ACM, 2020: 13–15.
- [25] XIAO Qingyu, YE Jin, PANG Chengjie, et al. Adaptive video streaming via deep reinforcement learning from user trajectory preferences[C]//Proceedings of 2020 IEEE 39th International Performance Computing and Communications Conference (IPCCC). Piscataway, NJ: IEEE, 2020: 1–8.
- [26] NACHUM O, GU Shixiang, LEE H, et al. Data-efficient hierarchical reinforcement learning[C]//Proceedings of the 32nd International Conference on Neural Information Processing Systems

(NIPS'18). Montreal, Canada; NIPS Foundation, 2018; 3307–3317.

[27] KIEFFER J C. The need for speed [J]. Facets, 2021, 6(1): 1390–1408.

[28] MATHONAT R, BOULICAUT J F, KAYTOUE M. A behavioral pattern mining approach to model player skills in rocket league [C]//Proceedings of 2020 IEEE Conference on Games (CoG). Piscataway, NJ: IEEE, 2020; 267–274.

[29] TASSA Y, DORON Y, MULDAL A, et al. Deepmind control suite [J]. arXiv preprint arXiv, 1801.00690, 2018.

[30] PENG Xuebin, ABBEEL P, LEVINE S, et al. Deepmimic: Example-guided deep reinforcement learning of physics-based character skills [J]. ACM Transactions On Graphics (TOG), 2018, 37(4): 1–14.

[31] AL-QAHTANI A A, ALENZI A A S, ALI A S. Player unknown's battlegrounds: Yet another internet gaming addiction [J]. Journal of Ayub Medical College Abbottabad, 2020, 32(1): 145–146.

[32] MAO Hongzi, NETRAVALI R, ALIZADEH M. Neural adaptive video streaming with pensieve [C]//Proceedings of the Conference of the ACM Special Interest Group on Data Communication. New York: ACM, 2017; 197–210.

[33] CHEN J, WANG Y, OU J, et al. ALBRL: Automatic load-balancing architecture based on reinforcement learning in software-defined networking [J]. Wireless Communications and Mobile Computing, 2022, 2022: 1–17.

[34] YE Deheng, CHEN Guibin, ZHANG Wen, et al. Towards playing full moba games with deep reinforcement learning [J]. Advances in Neural Information Processing Systems, 2020, 33: 621–632.

[35] TYTARENKO M. Optimizing immersion: Analyzing graphics and performance considerations in Unity3D VR development [J]. Asian Journal of Research in Computer Science, 2023, 16(4): 104–114.

[36] GAMAL G, AL-SHAikh M, SAEED M A, et al. Evaluating the performance of machine learning models for dynamic resource allocation in NFV [C]//Proceedings of the 3rd International Conference on Emerging Smart Technologies and Applications (eSmarTA). Piscataway, NJ: IEEE, 2023; 1–9.