综述

《国家科学评论》XX：SEAI，年份
doi:
https://doi.org/10.1093/nsr/XXXX

arXiv:2602.04411v1 [cs.ET] 4 Feb 2026

信息科学

自进化具身人工智能

冯彤彤、王鑫* 和朱文武*

摘要

具身人工智能（AI）是一种由智能体与其环境通过主动感知、具身认知和动作交互共同形成的智能系统。现有具身人工智能仍局限于人工设计环境，在这种环境中，智能体基于给定记忆进行训练，并针对给定任务构建模型，使固定的具身形态能够与相对静态的环境交互。这些方法无法应对真实环境，其特点是具身形态可变且环境动态开放。本文提出了自进化具身人工智能这一新范式。在该范式中，智能体基于自身不断变化的状态和环境运行，并通过记忆自更新、任务自切换、环境自预测、具身形态自适应和模型自进化，旨在实现持续自适应智能与自主进化。

具体而言，我们阐述自进化具身人工智能的定义、框架、组成部分和机制，系统回顾已实现组成部分的前沿研究，探讨实际应用，并指出未来研究方向。我们认为，自进化具身人工智能使智能体能够以类人的方式自主学习并与环境交互，为通用人工智能提供了新的视角。

关键词： 具身人工智能、自我演化、自主

1. 引言

具身人工智能（AI）源于艾伦·图灵于 1950 [1], 提出的具身图灵测试。该测试强调，除了抽象符号计算之外，还应通过智能体在物理世界中的感知和动作能力来评估智能。具体而言，具身人工智能是指由智能体及其环境通过主动感知、具身认知和动作交互共同构成的智能系统 [2]。不同于基于静态数据或抽象符号运行的非具身人工智能，具身人工智能强调感知、认知与动作之间的闭环耦合 [3], ，使智能体能够以扎根于现实的方式与物理环境交互。因此，具身人工智能在众多应用领域展现出广阔前景，包括机器人学 [4], 、自主驾驶 [5], 以及其他现实世界中的交互系统 [6] 等。

尽管进展迅速，现有具身人工智能仍局限于人工构建的设定 [7,8], ，其中智能体利用给定的记忆进行训练，并为给定任务构建模型，从而使固定的具身形态能够与相对静态的环境 [9]进行交互。现有的具身人工智能学习范式 [10] 在预定义目标下针对给定任务进行优化，其性能严重依赖于外部人类指导与经验性规范 [11], ，包括人类给定的任务、预先设计的模型、人工收集的数据集、固定配置的具身形态以及指定的实验环境。

然而，现实世界场景往往处于真实环境中。一方面，开放环境通常是动态且开放的 [12,13], ，其中的物体、动态变化和交互模式不断改变，在训练期间无法完全预先设定或穷尽列举。另一方面，开放环境中的部署通常需要可变具身形态 [14],，不同智能体在形态、感知、驱动、计算配置和物理约束方面各不相同，导致在固定具身形态下学习到的策略和模型难以适应。遗憾的是，现有具身人工智能无法在真实环境中胜任。

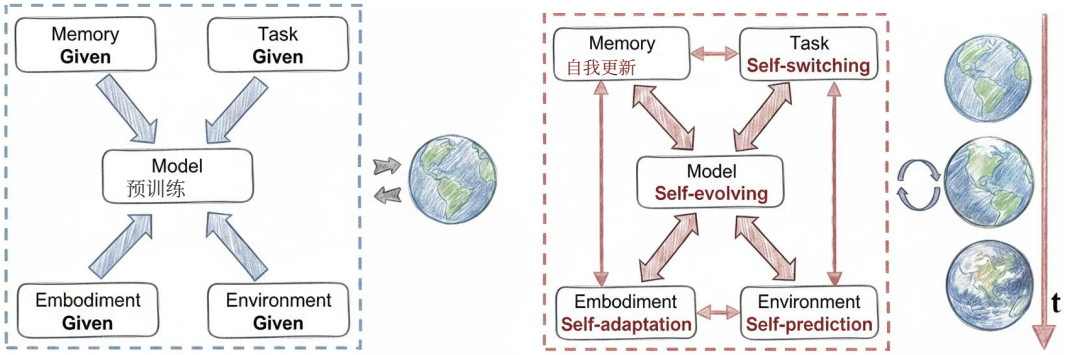
受神经科学 [15] 和认知科学 [16], 的启发，本文提出一种自进化具身人工智能新范式，如

中国北京100084清华大学计算机科学与技术系、北京信息科学与技术国家研究中心

*通讯作者。邮箱：wwzhu@tsinghua.edu.cn；xinwang@tsinghua.edu.cn。

收稿日期： XX XX 年；
修订日期： XX XX 年；
录用日期： XX XX 年

© 作者 Year。由牛津大学出版社代表中国科学出版传媒股份有限公司出版。本文为开放获取文章，依据知识共享署名许可协议（http://creativecommons.org/licenses/by/4.0/）发布。该协议允许在任何媒介中不受限制地再利用、传播和复制，但须适当注明原作出处。



（a）现有具身人工智能　　（b）自进化具身人工智能

图 1. 现有具身人工智能与自进化具身人工智能的比较。**（a）现有具身人工智能**基于给定的任务、给定的记忆、给定的具身形态和给定的环境，对相应模型进行预训练，高度依赖外部人工指导和经验配置。**（b）自进化具身人工智能**基于其不断变化的状态和动态环境运行，并实现记忆自我更新、任务自切换、具身形态自适应、环境自预测以及模型自我进化，旨在随着时间 t 的推移，通过自主进化实现持续自适应智能。

图 1 所示。自进化具身人工智能要求智能体基于自身不断变化的状态和环境运行，并具备记忆自更新、任务自切换、环境自预测、具身形态自适应以及模型自进化能力。在这一范式下，智能体能够通过自主进化实现持续自适应智能，使具身人工智能从人工设计环境走向真实环境。

具体而言，我们首先介绍自进化具身人工智能的定义、框架、组成部分和机制。随后，我们系统性地综述各个自进化组成部分的前沿研究。此外，我们还讨论实际应用，这些应用展示了自进化具身人工智能在可变具身形态和动态开放环境中的优势。最后，我们总结面向构建可控、可信且具备群体自进化能力的具身人工智能的未来研究方向。我们充分相信，自进化具身人工智能能够以类似人类的方式自主学习，并为通用人工智能提供新的视角。

2. 自进化具身人工智能

本节介绍自进化具身人工智能这一新范式，旨在推动具身人工智能走出人工设计环境，迈向真实环境。我们首先给出精确的定义和统一框架，以区分自进化具身人工智能与现有范式。随后，我们详细阐述自进化具身人工智能的核心组成部分，说明每个组成部分演化什么、为何演化、何时演化以及如何演化。最后，我们探讨自进化机制，以及如何通过智能体自身的状态及其与动态开放环境的交互驱动自主进化，从而实现持续自适应智能。

2.1 定义与框架

本小节确立自进化具身人工智能的概念基础。我们首先给出自进化具身人工智能的精确定义，然后介绍一个统一框架，用于描述自我进化的组织方式。

定义 1（自进化具身人工智能）。智能体基于自身不断变化的内部状态和外部环境运行，并具备记忆自更新、任务自切换、具身形态自适应、环境自预测和模型自进化能力，旨在通过自主进化实现持续自适应智能。

框架。图 1(b) 展示了自进化具身人工智能的框架。具体而言，该智能体由五个紧密耦合的模块构成，分别是记忆自更新、任务自切换、环境自预测、具身形态自适应和模型自我演化。这些模块并非彼此独立，而是通过与环境及彼此之间的持续交互共同演化。一个模块的更新可能引发其他模块的调整。这种组织方式体现为各模块随时间进行双向信息交换。例如，具身约束的变化可能改变存储记忆的相关性，进而影响任务切换和模型演化。它们共同构成一个

统一的演化循环，使智能体在具身形态不断变化、开放环境动态变化的情况下，持续具备自适应智能。

作为比较，图1(a)总结了现有具身人工智能的框架。具体而言，智能体在给定记忆、任务和具身配置的情况下，在特定实验环境中学习一个用于实现预定义目标的预训练模型。该框架总体上是单向的，如指向预训练模型的单向箭头所示。

这两个框架之间的对比凸显了具身人工智能在实现方式上的根本差异。现有具身人工智能依赖给定的组成部分和预训练模型，并结合外部人工指导与经验性配置；而自进化具身人工智能则将记忆、任务、环境、具身化和模型视为由智能体自身驱动的演化组成部分。与运行于相对静态环境中的现有具身人工智能相比，自进化具身人工智能通过以自我演化模块之间的双向耦合取代单向优化，使系统能够在动态开放环境中随时间持续适应，并具备自主进化能力。

2.2 组成部分

本小节详细阐述自进化具身人工智能的核心组成部分：记忆自更新、任务自切换、环境自预测、具身化自适应以及模型自进化。对于每个组成部分，我们详细说明什么在进化、为何需要进化、何时触发进化，以及如何具体实现进化。

记忆自我更新。记忆自我更新所演化的是智能体的内部记忆表征。智能体与环境交互时会持续产生经验，而记忆自我更新则决定如何随着时间推移有选择地保留、修订或丢弃这些经验。之所以需要演化，是因为在环境分布偏移以及具身形态或任务不断变化的情况下，固定记忆或静态数据集无法支持长期适应。在真实环境中，存储所有过往经验既不可行，也不可取。当演化被触发时，先前存储的经验相对于智能体当前的具身约束、任务目标或环境动态，可能已经过时、不再相关或具有误导性——演化通过选择性记忆更新机制实现，例如记忆自编辑、记忆自组织和记忆自蒸馏；这些机制会根据相关性、新颖性、不确定性或长期效用来优先处理经验。

任务自切换。任务自切换中进化的是智能体对任务目标的内部表征。它并非优化给定任务的预定义目标，而是提升智能体随时间自主调整其试图达成之事的能力。进化之所以必要，是因为给定任务的预定义目标无法充分刻画动态开放环境中不断变化的目标、约束与机遇。真实环境很少呈现稳定、明确的目标；随着时间推移，给定任务的预定义目标可能限制自主性，并阻碍有效适应。当当前任务表述变得不可行、次优，或与智能体的内部状态、具身化约束或环境动态不一致时，进化即被触发。进化通过自主的任务自切换机制实现，例如任务自主选择和任务自主生成，使智能体能够持续调整其旨在达成的目标，而无需明确的人工重新指定。

环境自预测。环境自预测所演化的是智能体对外部世界的内部表征。它体现了智能体如何持续更新对环境动态的理解，并对外部世界的未来作出预测。之所以有必要进行演化，是因为现实世界的环境具有非平稳性，无法通过固定模型或离线学习的模型进行完整刻画。智能体需要维护并优化预测表征，以预判未来状态、奖励以及交互动作序列。演化的触发条件包括：新观察到的环境状态与先前预测相矛盾、揭示此前未见的动力学，或暴露出过去未能捕捉的长期依赖关系。演化通过基于持续交互不断优化世界模型来实现，例如理解世界模型和生成式世界模型，使智能体能够更新对外部世界的内部表征，并随着时间推移支持自适应规划与决策。

具身自适应。具身自适应所演化的是智能体对自身物理状态的内部表征。自进化具身人工智能需要适应具身形态的异构形态、感知能力、驱动限制、计算配置和物理约束。之所以有必要进行演化，是因为不同平台上的具身形态可能各不相同，也可能因重新配置、磨损、损坏或退化而随时间发生变化，导致在固定具身形态下学习到的策略失效。演化的触发条件是预期的自身物理状态与实际状态之间出现差异，这表明先前的具身假设已不再成立。演化通过具身感知自适应机制来实现，例如具身自重构、具身自校准和具身自恢复，使智能体能够在可变具身形态下保持功能。

模型自我进化。模型自我进化所演进的是智能体的内部模型设计，包括模型架构、优化策略和评估标准，而不仅仅是模型参数。之所以需要进化，是因为固定的架构、既定的训练策略和预先设定的评估指标，无法在记忆、任务、环境和具身形态长期变化的情况下持续发挥作用。随着运行条件不断变化，先前经过良好优化的模型可能会变得低效、失准或无法胜任。当出现持续的性能退化、不确定性增加，或评估结果与现实世界行为之间存在系统性偏差时，便表明当前的模型设计已不再适用，此时就会触发进化。进化通过自适应地修改模型架构、调整优化策略并改进评估标准来实现，例如模型自重构、模型自优化和模型自评估，从而使学习过程本身也能够进化，并与长时程自主运行保持一致。

2.3 自进化机制

自进化机制描述了五个核心组成部分如何在统一的闭环中演化。它们并非各自独立演化，而是通过记忆自更新、任务自切换、环境自预测、具身形态自适应和模型自我演化，在智能体与环境的交互中持续协同进化。一个组成部分的变化会沿闭环传播，并促使其他组成部分产生自适应响应。

在每一次迭代中，智能体首先维持自我状态，包括通过具身形态自适应维持自身物理状态、通过模型自进化维持内部认知状态，以及通过环境自预测维持外部世界状态。基于智能体的自我状态，任务自切换确定要追求的目标，随后记忆自更新有选择地整理与预定目标相关的经验。这些信号进而驱动模型自进化，更新模型架构、优化策略和评估标准。进化后的模型输出交互动作，智能体借此与环境进行交互。随后，模型自进化通过模型评估更新内部认知状态，而具身形态自适应与环境自预测则分别更新自身物理状态和外部世界状态。最后，智能体随时间重复这一循环。

通过这种闭环组织，自进化机制在多个时间尺度和抽象层级上运作。快速自适应发生在单次迭代之内，任务、记忆和动作会根据当前自我状态进行调整；而较慢的进化则在长期交互中重塑内部认知状态、自身物理状态和外部世界状态。重要的是，自我进化由智能体通过动作反馈交互产生的自我状态转移所驱动，而非由外部人类指导和经验配置所驱动。通过在统一循环中协调记忆、任务、模型、具身化与环境的协同进化，自进化机制使持续自适应智能能够随时间在动态开放环境中实现自主进化。

3. 方法

近期研究显示，人们对自进化智能体 [72–74], 的兴趣日益浓厚，这类智能体旨在通过与环境持续交互来提升适应性。具体而言，这些方法通过提示工程 [75,76], 记忆选择 [23,77–79], 和工具优化技术[80], 来探索自我进化，并且通常以预训练大语言模型（大语言模型）作为统一的端到端骨干网络 [32,81]。这些方法展现了无需明确

表 1. 自进化具身人工智能代表性研究的分类。

模块	分类	代表性研究
记忆自我更新	自编辑	SAGE [17], Mem0 [18],记忆-R1 [19],Memento [20]
	自组织	A-MEM [21], MemInsight [22], MemGen [23], ReMe [24],生成式智能体 [25]
	自蒸馏	ExpeL [26],AWM [27],MUSE [28]
Task自切换	自选择	SEC [29], WebRL [30], 智能体0 [31], Mobile-Agent-E [32]
	自生成	ZeroGUI [33], 智能体Evolver [34],WebEvolver [35]
环境自预测	理解工作记忆	DreamerV3 [12], DreamerV4 [36], JEPA [37], V-JEPA [38], EvoAgent [39], NavMorph [40],WorMI [41]
	生成式世界模型	Genie [13], Genie-2 [42], Matrix-Game-2 [43], MineWorld [44],MineDreamer [45],开放获取 [46]
具身化自适应	自重构	GET-Zero [47],BoT [48], PEAC [49]
	自校准	UP-OSI [50],SPI-Active [51],OFCI [52]
	自我恢复	跌倒恢复 [53], 损伤恢复 [54]
模型自我进化	自我重构	MaskTAS [55], PMoE [56], MoE-Adapters4CL [57], D-MoLE [58], L2R [59], EEP [60]
	自我优化	MoE-CL [61], SRT [62], SEAS [63], TTRL [64], 自我进化的具身智能-精炼 [65], Reflexion [66]
	自我评估	以大语言模型为评判者 [67], SER [68], RLME [69], RLIF [70], MAE [71]

人工干预也能实现自我进化的潜力，但环境感知模块和具身配置通常在部署后就被视为固定不变 [82,83], ，这限制了它们在可变具身形态和动态环境下支持长期自主性的能力。

据我们所知，本文首次提出了一个系统性框架，通过整合五个核心组成部分及其自进化机制，对自进化具身人工智能进行形式化定义，使其能够随着时间推移，在动态开放环境中通过自主演化实现持续自适应智能。具体而言，自进化具身人工智能源于记忆自更新、任务自切换、环境自预测、具身形态自适应和模型自进化的闭环协同进化，这一过程由智能体通过动作—反馈交互实现的自身状态转变所驱动。

现有的自进化智能体仅属于自进化具身人工智能的一个子集，而自进化具身人工智能仍处于早期阶段，具有巨大的探索潜力。基于这些观点，本文系统梳理了各个组成部分的现有最先进研究成果。由于自我演化是一个端到端的过程，我们根据现有方法主要针对自我演化循环中的哪个组成部分来进行归类。

3.1 记忆自更新

现有关于记忆自更新的研究大致可分为三类：记忆自编辑、记忆自组织和记忆自蒸馏。这些类别的主要区别在于如何修改、组织和抽象经验，以支持长期适应。

记忆自编辑。这些研究聚焦于直接修改存储记忆内容的显式操作，包括添加、更新、删除、合并和遗忘。SAGE [17]引入了反思机制，利用遗忘曲线调节记忆保留。Mem0 [18] 通过从交互历史中提取显著事实并将其整合到长期记忆中，对记忆编辑进行了形式化。Memory-R1 [19] 通过强化学习学习记忆编辑操作，从而能够自适应地控制何时添加、更新或删除记忆条目。Memento [20] 优化记忆重写和检索策略，以维持紧凑且相关的长期记忆。总之，这些研究将记忆演化视为由操作驱动的编辑过程。

记忆自组织。这些工作强调的是记忆结构的演变，而非

记忆内容。A-MEM [21] 将智能体记忆组织成一个动态索引、相互链接的笔记网络，其灵感源自卡片盒笔记法式的知识管理。MemInsight [22]为原始情景记忆增添语义结构，从而实现更可靠的检索与复用。Mem-Gen [23] 探索生成式潜在记忆，让记忆融入一个随经验不断演变的连续潜在空间。ReMe [24] 和生成式智能体 [25] 等相关系统进一步凸显了结构化记忆在支持长期推理和一致性方面的作用。这些工作将记忆演变视为一种组织过程。

记忆自蒸馏。这些研究旨在将情节性经验转化为可复用的知识、技能或工作流。ExpeL [26]从交互轨迹中提取抽象洞见和规则，以指导未来行为。AWM [27]存储可复用的任务工作流，使智能体能够在不同任务之间迁移程序性知识。MUSE [28]进一步以分层方式组织记忆，在长期过程中积累策略性和程序性能力。总体而言，这些研究将记忆演化视为一种蒸馏过程，把情节性经验转化为持久能力。

3.2 任务自切换

现有关于任务自切换的研究大致可分为两类：任务自选择和任务自生成。这两条研究路线的区别在于，任务是从不断演化的候选集中自适应选择，还是根据智能体的自我状态和环境动力学在线生成。

任务自选择。这些研究将任务切换建模为选择或调度问题，智能体根据学习进度或交互反馈动态决定下一步执行哪个任务。代表性研究包括能够学习自适应任务调度策略的自进化课程（SEC） [29],；用于网络交互、采用自我演化在线课程的 WebRL [30], ；将课程智能体与执行器共同演化的 Agent 0 [31],；以及 Mobile-Agent-E [32], 等分层具身系统，其中高层管理器在执行过程中进行目标选择或子目标选择。

任务自主生成。这些工作着重于在线创建新任务，而不是从预定义池中进行选择，从而实现由智能体的自我状态和探索需求驱动的任务切换。ZeroGUI [33] 根据界面状态自动生成图形用户界面任务，以支持在线学习；AgentEvolver [34] 通过自我提问生成新任务，以指导在新环境中的探索；WebEvolver [35] 则利用协同进化世界模型生成自主指令的交互任务。

3.3 环境自预测

现有关于环境自预测的研究大致可分为两类：理解型世界模型（WMs）和生成式世界模型。前者学习支持预测与规划的预测性潜在表征，但不一定对原始观测进行解码；后者则使用概率生成模型，显式建模观测生成过程（例如下一帧或展开序列）。

理解型世界模型。这些研究专注于学习环境动力学的预测性潜在表征，通常通过循环状态空间模型风格的潜在动力学模型 [12,36] 或 JEPA 风格的预测性表征学习 [37,38]来实现。这些研究能够通过在分布偏移下持续演化预测性潜在表征，在自我演化具身 AI 中实现环境自预测，例如 EvoAgent [39]中针对长时程任务的持续世界模型更新；NavMorph [40]中用于导航的在线自适应世界建模；以及 WorMI [41]中多个特定领域理解的测试时组合。

生成式世界模型。这些工作会明确预测未来的环境观测，通常通过自回归 Transformer 或扩散模型实现 [84]。近期的大规模生成式世界模型，如 Genie [13], Genie-2 [42], 和 Matrix-Game-2 [43], ，从大规模视频数据中学习通用环境动力学，并支持长时程想象。在 Minecraft 及相关开放世界中，Mine World [44]和 Mine-Dreamer [45] 等交互式生成式世界模型可进行动作条件视频生成；OA [46]则在持续强化学习中借助在线世界模型进行规划。这些方法结合持续数据收集和在线优化，为环境自预测奠定了实用基础。在其他应用中，OA [46]也在持续强化学习中借助在线世界模型进行规划。

3.4 具身形态自适应

现有的具身自适应研究大致可分为三类：具身自重构、具身自校准和具身自恢复。具身自适应关注智能体如何在异构形态、感知能力、驱动限制、计算配置和物理约束发生变化时保持适应性。

具身自重构。这些研究聚焦于适应离散配置变化，包括异构形态、传感布局和计算设置。这些研究学习具身条件化或模块化策略，使其能够跨不同机器人结构泛化，从而无需针对每种具身形态重新训练即可复用策略。代表性研究包括 GET-Zero [47], ，它通过图变换器对具身拓扑建模，以实现零样本泛化；Body Transformer（BoT） [48], ，它将具身结构显式编码到策略学习中；以及 PEAC [49],，它预训练可迁移表征，用于跨具身控制。

具身自校准。这些研究针对驱动限制和物理约束的连续变化，例如质量变化、摩擦、接触动力学和传感噪声。这类研究通常依赖在线系统辨识或潜在参数估计，在执行过程中重新校准内部具身模型。一个经典基础是具有在线系统辨识的通用策略（UP-OSI） [50],，近期的扩展则支持在富接触场景中进行主动或持续校准，例如用于足式机器人的 SPI-Active [51]和在线摩擦系数辨识（OFCI） [52] 。

具身自恢复。 这些研究针对感知、驱动或计算资源的突发退化或故障，要求快速恢复可执行行为。这一方向侧重于故障感知自适应与恢复，而非参数优化。代表性工作包括基于质量多样性的损伤恢复 [54], ，通过维护多样化的行为库，实现损伤情况下的快速自适应；以及面向足式机器人的基于学习的跌倒或故障恢复策略 [53]。

3.5 模型自进化

现有的模型自进化研究大致可分为三类：模型自重构、模型自优化和模型自评估。

模型自重构。这些研究通过模块化增长、专家组合或自适应路由来演化模型的内部架构，以适应不同的具身形态、环境和任务。代表性研究包括通过掩码蒸馏进行自监督架构搜索 [55], 用于持续学习的自适应专家混合（MoE）模型 [56,57], 动态适配器或 LoRA 专家组合 [58,59],以及结合专家剪枝的模块化增长 [60]。

模型自优化。这些研究着重通过持续调优、自生成反馈或测试时学习，自主改进训练和参数更新策略。典型研究包括无需人工标签的持续指令微调 [61], 从 SRT 中的语言反馈学习 [62], SEAS 中的对抗式自我改进 [63], 测试时强化学习 [64], 以及迭代式自我精炼和基于反思的优化 [65,66]。

模型自我评估。这些研究演化评估标准和奖励信号，以便在缺乏可靠真实标注监督的情况下引导学习。这一方向的特点包括 LLM-as-a-judge 框架 [67], 自演化奖励学习 [68], 从元评估中进行强化学习 [69], 无外部奖励学习 [70], 以及多评审者协同进化 [71]。

4. 应用

本节讨论三种实际应用，展示自进化具身人工智能如何满足环境分布偏移、有限监督和可变具身形态下的实际需求：自主服务机器人技术、自动驾驶和自主无人机（UAV）。

4.1 自主服务机器人技术

自主服务机器人在开放式的以人为中心的环境中运行，其中物体布局、用户偏好和任务规范会随时间变化 [85]。开放词汇移动操作和通用机器人策略的最新进展，展现出在不同任务和场景中的强大泛化能力 [86–88], ，但

要实现全面部署，仍需持续适应。

在这些应用中，记忆自更新对于优先处理罕见但后果严重的交互故障，而非重复的常规行为，至关重要。任务自切换使机器人能够根据用户意图和情境变化重新解读并调整目标。具身化自适应支持系统部署于感知、驱动和载荷限制各不相同的异构平台。这些组成部分共同展现了自我进化如何实现持续自主，使其超越由人工制定的家庭场景基准测试。

4.2 自动驾驶

自动驾驶是具身人工智能在动态环境中的一种安全关键型应用。现实世界的交通呈现非平稳动力学、罕见的边缘情况，以及与人类之间的复杂交互 [89] ，这些情况无法在训练期间被穷尽地涵盖。近期研究日益依赖预测性世界模型 [90], 和高层推理 [91]来支持规划与仿真 [92] 。

在这些应用中，环境自预测发挥着核心作用，使智能体能够预测未来事件并计算长时域奖励。模型自进化支持系统适应天气、道路结构和交通模式等分布偏移。然而，与其他应用不同，自动驾驶凸显了具身化自适应的必要性，即必须调节智能体的动作，以确保安全性、可解释性和可靠性。这一应用体现了自进化具身人工智能中自主性与控制之间的张力。

4.3 自主无人机

自主无人机 [93,94] 在快速动力学、部分可观测性和严格的物理约束下运行，使其成为自进化具身人工智能的天然测试平台。无人机任务，如巡检 [95], 规划 [96], 和搜救 [97] ，需要持续适应动态环境、 [98], 天气条件[99], 和任务目标 [100]。

在这些应用中，由于能量限制、执行器不确定性和平台异质性，具身化自适应至关重要。环境自预测支持轨迹想象，并能在动力学快速变化时进行在线重新规划。模型自进化则可实现轻量且高效的更新，同时满足机载计算限制，与地面机器人相比，无人机更注重高效的经验自更新，以及具身化、环境和模型之间的紧密协同演化，展现了在严苛物理限制下实现自我演化人工智能的可行性。

5. 未来方向

自进化具身人工智能为推动具身人工智能超越人工设计环境、迈向真实环境提供了新范式，但目前仍处于早期阶段，具有巨大的探索潜力。除了各组成部分的自我进化以及组成部分之间的协同进化之外，未来发展还需要解决三个关键问题：（1）如何控制自我进化，以确保稳定性和效率；（2）如何使自我进化值得信赖，从而实现安全部署；（3）如何将自我进化扩展至群体级系统。

5.1 可控自我演化

自进化具身人工智能面临的一项核心挑战，是确保自我进化过程始终可控。虽然自我进化能让智能体适应动态环境，但失控的演化可能导致不稳定、性能退化或非预期行为。为解决这一问题，未来研究应探索如何调节自我进化的范围、速率和方向。分层控制架构可以将快速控制回路与较慢的演化回路分离，在效率与稳定性之间实现平衡。此外，开发可解释反馈机制对于理解和控制演化过程至关重要。稳定裕度或适应预算等可量化的可控性标准，对于实现可靠且可预测的自我进化必不可少。

5.2 可信自我演化

要部署于现实世界的自进化具身人工智能应用，其自我进化必须值得信赖。这要求确保自我进化过程安全、透明且可问责。一项重大挑战是管理自我进化在环境和智能体自身中引入的非平稳性，这使验证和确认更加复杂。未来的工作应着重于监测、审计和解释自我进化过程，以便人类进行更多监督，并更好地理解发生了哪些变化及其原因。

这包括具备不确定性感知能力的演化，以及实施回滚机制以撤销有害的适应。在缺乏明确监督的情况下，如何让自我进化在长期内与人类价值观保持一致，仍是一个重要的研究方向。

5.3 群体自演化

尽管现有关于自进化具身人工智能的研究主要聚焦于单个智能体，但许多现实世界应用涉及多个智能体作为一个整体协同运行。群体自演化将自我进化的概念扩展到多智能体系统，在此类系统中，自我进化同时发生于个体和群体层面。群体自我进化面临的挑战包括：智能体如何在约束条件下共享经验、协调任务并共同进化其内部模型。未来研究应探索分布式记忆、策略共享和角色分化机制，以支持群体自演化。理解群体自进化智能体所产生的涌现行为，并确保群体层面的进化与集体目标保持一致，也至关重要。群体自演化有望释放超越单个智能体能力的规模化自适应智能，从而实现稳健、灵活且富有韧性的具身系统。

6. 结论

本文提出了自进化具身人工智能这一新范式，使具身人工智能能够从人工设计环境迈向开放环境。自进化具身人工智能是一种智能系统，由智能体及其环境通过感知自我进化、认知自我进化和交互自我进化共同构成。具体而言，智能体依据不断变化的自身状态和环境运行，并实现记忆自更新、任务自切换、环境自预测、具身化自适应以及模型自进化。在这一范式下，智能体能够随着时间推移，通过自主进化实现持续自适应智能。我们认为，自进化具身人工智能能够为通用人工智能提供一种新的视角。

参考文献

1. Turing AM. 计算机器与智能。解析图灵测试：探索思考型计算机过程中的哲学与方法论问题，23–65（2007）。

2. Feng T, Wang X, Jiang YG 等. 具身人工智能：从大语言模型到世界模型 [特征]. IEEE电路与系统杂志 2025年; 25: 14–37. 3. Sun F, Chen R, Ji T 等. 具身智能综合综述：进展、挑战与未来展望. CAAI人工智能研究 2024年; 3: 1. 4. Xu Z, Wu K, Wen J 等. 基础模型赋能机器人学综述：迈向具身人工智能. arXiv预印本 arXiv:2402.02385 2024年; . 5. Ma Y, Song Z, Zhuang Y 等. 具身人工智能视觉-语言-动作模型综述. arXiv预印本 arXiv:2405.14093 2024年; . 6. Hu Y, Xie Q, Jain V 等. 通过基础模型迈向通用机器人：综述与元分析. arXiv预印本 arXiv:2312.08782 2023年; . 7. Duan J, Yu S, Tan HL 等. 具身人工智能综述：从模拟器到研究任务. IEEE计算智能新兴主题汇刊 2022年; 6: 230–244. 8. Srivastava S, Li C, Lingelbach M 等. Behavior：虚拟、交互式和生态化环境中日常家务活动的基准测试. 机器人学习会议 (2022) 477–490. 9. Brohan A, Brown N, Carbajal J 等. Rt-1：用于大规模现实世界控制的机器人变换器. arXiv预印本 arXiv:2212.06817 2022年; . 10. ONeill A, Rehman A, Maddukuri A 等. Open x-embodiment：机器人学习数据集与 rt-x 模型. 2024年 IEEE国际机器人与自动化会议 (ICRA) (2024年) 6892–6903. 11. Zitkovich B, Yu T, Xu S 等. Rt-2：视觉-语言-动作模型将网络知识迁移至机器人控制. 机器人学习会议 (2023) 2165–2183. 12. Hafner D, Pasukonis J, Ba J 等. 通过世界模型掌握多样化控制任务. 《自然》 2025年; 1–7. 13. Bruce J, Dennis MD, Edwards A 等. Genie：生成式交互环境. 第四十一届国际机器学习大会 (2024) . 14. Liu Y, Chen W, Bai Y 等. 使网络空间与物理世界对齐：具身人工智能综合综述. IEEE/ASME机电一体化汇刊 2025年; . 15. Zhu D, Bu Q, Zhu Z 等. 通过终身学习提升自主能力：自主智能系统综述. 《神经机器人学前沿》 2024年; 18: 1385778. 16. Wang L, Zhang X, Su H 等. 持续学习综合综述：理论、方法与应用. IEEE模式分析与机器智能汇刊 2024年; 46: 5362–5383. 17. Liang X, He Y, Xia Y 等. 具备反思与记忆增强能力的自进化智能体. arXiv预印本 arXiv:2409.00872 2024年; . 18. Chhikara P, Khant D, Aryan S 等. Mem0：构建具备可扩展长期记忆、可用于生产环境的人工智能智能体 (2025年)。URL https://arxiv. org/abs/2504.19413 2025年; . 19. Yan S, Yang X, Huang Z 等. Memory-R1：通过强化学习增强大语言模型智能体管理和利用记忆的能力. arXiv预印本 arXiv:2508.19828 2025年; . 20. Zhou H, Chen Y, Guo S 等. Memento：无需微调大语言模型即可微调大语言模型智能体, 2025年。URL https://arxiv. org/abs/2508.16153 ; .

21. Xu W, Liang Z, Mei K 等. A-MEM：面向大语言模型智能体的智能体记忆。*arXiv*预印本 *arXiv:2502.12110* 2025年; . 22. Salama R, Cai J, Yuan M 等.MemInsight：面向大语言模型智能体的自主记忆增强。*arXiv*预印本 *arXiv:2503.21760* 2025年; . 23. Zhang G, Fu M 和 Yan S. MemGen：为自进化智能体编织生成式潜在记忆。*arXiv*预印本 *arXiv:2509.24704* 2025年; . 24. Cao Z, Deng J, Yu L 等. 记住我，完善我：一种由经验驱动的智能体进化动态程序性记忆框架。*arXiv*预印本 *arXiv:2512.10696*2025年; . 25. Park JS, O' Brien J, Cai CJ 等. 生成式智能体：人类行为的交互式模拟。第*36*届*ACM*用户界面软件与技术年度研讨会论文集(2023) 1–22. 26. Zhao A, Huang D, Xu Q等. ExpeL：大语言模型智能体是经验型学习者。*AAAI*人工智能大会论文集，第38卷(2024年) 19632–19642. 27. Wang ZZ, Mao J, Fried D 等. 智能体工作流记忆。*arXiv*预印本 *arXiv:2409.07429*2024年; . 28. Yang C, Yang X, Wen L 等. 在实践中学习：一种由经验驱动、面向长时程任务的自进化智能体。*arXiv*预印本 *arXiv:2510.08002*2025年; . 29. Chen X, Lu J, Kim M 等. 面向大语言模型推理的自进化课程。*arXiv*预印本 *arXiv:2505.14970*2025年; . 30. Qi Z, Liu X, Iong IL 等. WebRL：通过自进化在线课程强化学习训练网页智能体。*arXiv*预印本 *arXiv:2411.02337*2024年; . 31. Xia P, Zeng K, Liu J等.Agent0：通过工具集成推理，在零数据条件下释放自进化智能体的潜力。*arXiv*预印本 *arXiv:2511.16043* 2025年; . 32. Wang Z, Xu H, Wang J 等 . Mobile-Agent-E：面向复杂任务的自进化移动助手。*arXiv*预印本 *arXiv:2501.11733* 2025年; . 33. Yang C, Su S, Liu S 等. ZeroGUI：以零人工成本实现在线 GUI 学习自动化。*arXiv*预印本 *arXiv:2505.23762* 2025年; . 34. Zhai Y, Tao S, Chen C 等. AgentEvolver：迈向高效的自我进化智能体系统。*arXiv*预印本 *arXiv:2511.10395* 2025年; . 35. Fang T, Zhang H, Zhang Z 等.WebEvolver：利用协同进化世界模型增强网页智能体自我改进能力。*arXiv*预印本 *arXiv:2504.21024*2025年; . 36. Hafner D, Yan W 和 Lillicrap T. 在可扩展世界模型内部训练智能体。*arXiv*预印本 *arXiv:2509.24527*2025年; . 37. Assran M, Duval Q, Misra I 等. 利用联合嵌入预测架构从图像中进行自监督学习。*IEEE/CVF*计算机视觉与模式识别会议论文集(2023) 15619–15629. 38. Assran M, Bardes A, Fan D 等. V-JEPA 2：自监督视频模型实现理解、预测与规划。*arXiv*预印本 *arXiv:2506.09985* 2025年; . 39. Feng T, Wang X, Zhou Z 等. EvoAgent：借助持续世界模型实现面向长时程任务的智能体自主进化。*arXiv*预印本 *arXiv:2502.05907*2025年; . 40. Yao X, Gao J 和 Xu C. NavMorph：用于连续环境中视觉语言导航的自我演化世界模型。*arXiv*预印本 *arXiv:2506.23468* 2025年; . 41. Yoo M, Jang J, Yoon S 等用于具身智能体测试时适应的世界模型植入。*arXiv*预印本 *arXiv:2509.03956* 2025年；。42. Parker-Holder J, Ball P, Bruce J 等Genie 2：大型基础世界模型。网址：*https://deepmind.google/discover/blog/genie-2-a-large-scale-foundation-world-model* 2024年；。43. He X, Peng C, Liu Z 等 Matrix-game 2.0：开源实时流式交互式世界模型。*arXiv*预印本 *arXiv:2508.13009*2025年；。44. Guo J, Ye Y, He T 等 Mineworld：Minecraft上的实时开源交互式世界模型。*arXiv*预印本 *arXiv:2504.08388* 2025年；。45. Zhou E, Qin Y, Yin Z 等Minedreamer：通过想象链学习遵循指令，以控制模拟世界。*arXiv*预印本 *arXiv:2403.12037*2024年；。46. Liu Z, Fu G, Du C 等通过在线世界模型规划实现持续强化学习。*arXiv*预印本 *arXiv:2507.09177* 2025年；。47. Patel A and Song S. Get-zero：用于零样本具身泛化的图形具身变换器。*2025*年 *IEEE*机器人与自动化国际会议（*ICRA*）（2025年）14262–14269。48. Sferrazza C, Huang DM, Liu F 等Body transformer：利用机器人具身进行策略学习。*arXiv*预印本 *arXiv:2408.06316* 2024年；。49. Ying C, Zhongkai H, Zhou X 等Peac：用于跨具身强化学习的无监督预训练。神经信息处理系统进展2024年；**37**：54632–54669。50. Yu W, Tan J, Liu CK 等未雨绸缪：通过在线系统辨识学习通用策略。*arXiv*预印本 *arXiv:1702.02453* 2017年；。51. Sobanbabu N, He G, He T等通过主动探索进行基于采样的系统辨识，以实现足式机器人仿真到现实迁移学习。*arXiv*预印本 *arXiv:2505.14266*2025年；。52. Kim H, Kang D, Kim MG 等利用平滑接触梯度在线辨识湿滑地形上足式机器人的摩擦系数。*IEEE*机器人与自动化快报2025年；。53. Yang C, Pu C, Xin G 等学习复杂运动技能，实现足式机器人跌倒恢复。*IEEE*机器人与自动化快报2023年；**8**：4307–4314。54. Allard M, Smith SC, Chatzilygeroudis K 等通过分层质量多样性实现实体机器人的在线损伤恢复。*ACM*进化学习汇刊2023年；**3**：1–23。55. Yan C, Chang X, Li Z 等掩码蒸馏推动自监督变换器架构搜索。第十二届国际学习表征会议（2024年）。56. Jung MJ and Kim J. Pmoe：用于持续学习的渐进式专家混合与非对称变换器。*arXiv*预印本 *arXiv:2407.21571* 2024年；。57. Yu J, Zhuge Y, Zhang L 等通过专家混合适配器提升视觉语言模型的持续学习能力。*IEEE/CVF*计算机视觉与模式识别会议论文集（2024年）23219–23230。58. Ge C, Wang X, Zhang Z 等用于持续多模态指令微调的课程式LoRA专家动态混合。*arXiv*预印本 *arXiv:2506.11672* 2025年；。

59. Araujo V, Moens MF and Tuytelaars T. 学习路由，以在语言模型持续学习中实现动态适配器组合。计算语言学协会论文集：*EMNLP 2024*（2024年）687–696。60. Liu E, Zhu J, Lin Z 等稀疏混合专家语言模型的高效专家剪枝：提升性能并降低推理成本。*arXiv*预印本 *arXiv:2407.00945* 2024年；。61. Kang J, Huang L, Hou C 等通过持续指令微调实现大语言模型自我演化。*arXiv*预印本 *arXiv:2509.18133*2025年；。62. Hu C, Hu Y, Cao H等通过学习语言反馈，教会语言模型自我改进。*arXiv*预印本 *arXiv:2406.07168* 2024年；。63. Diao M, Li R, Liu S 等Seas：面向大语言模型的自我演化对抗性安全优化。*AAAI*人工智能大会论文集，第39卷（2025年）23778–23786。64. Zuo Y, Zhang K, Sheng L 等Ttrl：测试时强化学习。*arXiv*预印本 *arXiv:2504.16084*2025年；。65. Madaan A, Tandon N, Gupta P等Self-refine：通过自反馈进行迭代式优化。神经信息处理系统进展2023年；**36**：46534–46594。66. Shinn N, Cassano F, Gopinath A 等Reflexion：采用语言强化学习的语言智能体。神经信息处理系统进展2023年；**36**：8634–8652。67. Zheng L, Chiang WL, Sheng Y 等使用MT-Bench和Chatbot Arena评判大语言模型评审者。神经信息处理系统进展2023年；**36**：46595–46623。68. Huang C, Fan Z, Wang L 等大语言模型的自演化奖励学习。*arXiv*预印本 *arXiv:2411.00418*2024年；。69. Rentschler M and Roberts J. 基于元评估的强化学习：在没有真实标签的情况下对齐语言模型。*arXiv*预印本 *arXiv:2601.21268*2026年；。70. Zhao X, Kang Z, Feng A 等学习在没有外部奖励的情况下进行推理。*arXiv*预印本 *arXiv:2505.19590* 2025年；。71. Chen Y, Wang Y, Zhu S 等多智能体演化：大语言模型通过协同进化实现自我改进。*arXiv*预印本 *arXiv:2510.23595* 2025年；。72. Fang J, Peng Y, Zhang X 等自演化人工智能智能体综合综述：连接基础模型与终身智能体系统的新范式。*arXiv*预印本 *arXiv:2508.07407* 2025年；。73. Gao Ha, Geng J, Hua W 等自演化智能体综述：迈向人工超级智能。*arXiv*预印本 *arXiv:2507.21046* 2025年；。74. Wang S, Long Z, Fan Z等Benchmark self-evolving：用于动态大语言模型评估的多智能体框架。第*31*届国际计算语言学会议论文集（2025年）3310–3328。75. Ou Y, Zhou W, Ding S 等符号学习助力自进化智能体。*AI Open* 2025年；。76. Yadav S and Roy L. 自适应提示智能：迈向自演化对话智能体。*2025*年 *IEEE*第五届商业、工业与政府信息通信技术国际会议（*ICTBIG*）（2025年）1–6。77. Cai Y, Hao Y, Zhou J 等通过经验驱动的终身学习构建自演化智能体：框架与基准测试。*arXiv*预印本 *arXiv:2508.19005* 2025年；。78. Xu H, Hu J, Zhang K 等Sedm：面向智能体的可扩展自我演化分布式记忆。*arXiv preprintarXiv:2509.09498* 2025年；。79. Huang C, Yu W, Wang X 等R-zero：基于零数据的自我演化推理大语言模型。*arXiv preprintarXiv:2508.05004* 2025年；。80. Liu J, Xiong K, Xia P 等Agent0-vl：探索用于工具集成视觉语言推理的自我演化智能体。*arXiv preprint arXiv:2511.19900* 2025年；。81. Thawakar O, Venkatraman S, Thawkar R等Evolmm：具有连续奖励的自我演化大型多模态模型。*arXiv preprintarXiv:2511.16672* 2025年；。82. Zhang DC, Zhao Y, Wu J 等Evolvesearch：一种迭代式自我演化搜索智能体。*2025*年自然语言处理实证方法会议论文集（2025年）13134–13147。83. Sun Z, Liu Z, Zang Y 等Seagent：通过经验自主学习的自我演化计算机使用智能体。*arXiv preprintarXiv:2508.04700*2025年；。84. Ding J, Zhang Y, Shang Y 等理解世界还是预测未来？世界模型综合综述。*ACM*计算调查2025年；**58**：1–38。85. Zhang W, Li Y, Qiao Y 等通用机器人操控策略的有效调优策略。*2025*年*IEEE*机器人与自动化国际会议（*ICRA*）（2025年）7255–7262。86. Yenamandra S, Ramachandran A, Yadav K 等Homerobot：开放词汇移动操作。*arXivpreprintarXiv:2306.11565* 2023年；。87. Huang S, Chen L, Zhou P 等Enerverse：面向机器人学操作构想具身未来空间。*arXivpreprintarXiv:2501.01895* 2025年；。88. Driess D, Xia F, Sajjadi MS 等Palm-e：一种具身多模态语言模型 2023年；。89. Yurtsever E, Lambert J, Carballo A 等自动驾驶综述：常见实践与新兴技术。*IEEE Access*2020年；**8**：58443–58469。90. Schwarting W, Alonso-Mora J和Rus D。自动驾驶车辆的规划与决策。《控制、机器人与自主系统年度评论》2018年；**1**：187–210。91. Peng Q, Bai C, Zhang G等Navigscene：为超视距自动驾驶桥接局部感知与全局导航。第*33*届*ACM*国际多媒体会议论文集（2025年）4193–4202。92. Sun W, Lin X, Shi Y等Sparsedrive：基于稀疏场景表示的端到端自动驾驶。*2025*年*IEEE*机器人与自动化国际会议（*ICRA*）（2025年）8795–8801。93. Chang Y, Cheng Y, Manzoor U 等无GPS环境下无人机自主导航综述。《机器人与自主系统》2023年；**170**：104533。94. Feng T, Wang X, Han F等U2udata：面向无人机群自主飞行的大规模协同感知数据集。第*32*届*ACM*国际多媒体会议论文集（2024年）7600–7608。95. Li Z, Zhang Y, Wu H 等高压输电线路无人机自主巡检系统的设计与应用。《遥感》2023年；**15**：865。96. Deng W, Feng J和Zhao H。动态环境下基于多策略机制沙猫群优化的无人机自主路径规划。*IEEE*物联网期刊2025年；。

97. Feng T, Wang X, Han F等U2udata+：面向具身长程任务的可扩展无人机群自主飞行数据集。第三十届*AAAI*人工智能会议2025年；。98. Wang C, Wang J, Shen Y等大规模复杂环境中无人机的自主导航：一种深度强化学习方法。*IEEE*车辆技术汇刊2019年；**68**：2124–2136。99. Feng T, Li Q, Wang X 等使用去噪扩散模型的多天气跨视角地理定位。第二届多媒体无人机研讨会论文集：从新视角捕捉世界（2024年）35–39。100. Du Z, Luo C, Min G等无人驾驶航空器（无人机）自主与智能集群综述。*IEEE*智能交通系统汇刊2025年；。