

## INFORMATION SCIENCE

## Self-evolving Embodied AI

Tongtong Feng, Xin Wang\* and Wenwu Zhu\*

### ABSTRACT

Embodied Artificial Intelligence (AI) is an intelligent system formed by agents and their environment through active perception, embodied cognition, and action interaction. Existing embodied AI remains confined to human-crafted setting, in which agents are trained on given memory and construct models for given tasks, enabling fixed embodiments to interact with relatively static environments. This paradigm of existing embodied AI heavily relies on external human guidance and empirical configurations. Such methods fail in in-the-wild setting characterized by variable embodiments and dynamic open environments. This paper introduces *self-evolving embodied AI*, a new paradigm in which agents operate based on their changing state and environment with memory self-updating, task self-switching, environment self-prediction, embodiment self-adaptation, and model self-evolution, aiming to achieve continually adaptive intelligence with autonomous evolution. Specifically, we present the definition, framework, components, and mechanisms of self-evolving embodied AI, systematically review state-of-the-art works for realized components, discuss practical applications, and point out future research directions. We believe that self-evolving embodied AI enables agents to autonomously learn and interact with environments in a human-like manner and provide a new perspective toward general artificial intelligence.

**Keywords:** Embodied AI, Self-evolving, Autonomous

### 1. Introduction

Embodied Artificial Intelligence (AI) originates from the Embodied Turing Test by Alan Turing in 1950 [1], which emphasizes that intelligence should be evaluated through an agent's ability to perceive and act in the physical world in addition to abstract symbol computation. Specifically, embodied AI refers to intelligent systems [2] formed by agents and their environment through active perception, embodied cognition, and action interaction. Unlike disembodied AI that operates on static data or abstract symbols, embodied AI emphasizes the closed-loop coupling between perception, cognition, and action [3], enabling agents to interact with physical environments in a grounded manner. As such, embodied AI benefits in promising potential in a wide range of applications, including robotics [4], autonomous driving [5], and other real-world interactive systems [6] etc.

Despite its rapid progress, existing embodied AI remains confined to human-crafted setting [7,8], in which agents are trained with given memory and construct models for given tasks,

thus enabling fixed embodiments to interact with relatively static environments [9]. The current embodied AI learning paradigm [10] is optimized for given tasks under predefined objectives, and the performance relies heavily on external human guidance and empirical specification [11], including human-given tasks, pre-designed models, manually-collected datasets, fixed-configuration embodiments, and specified experimental environments.

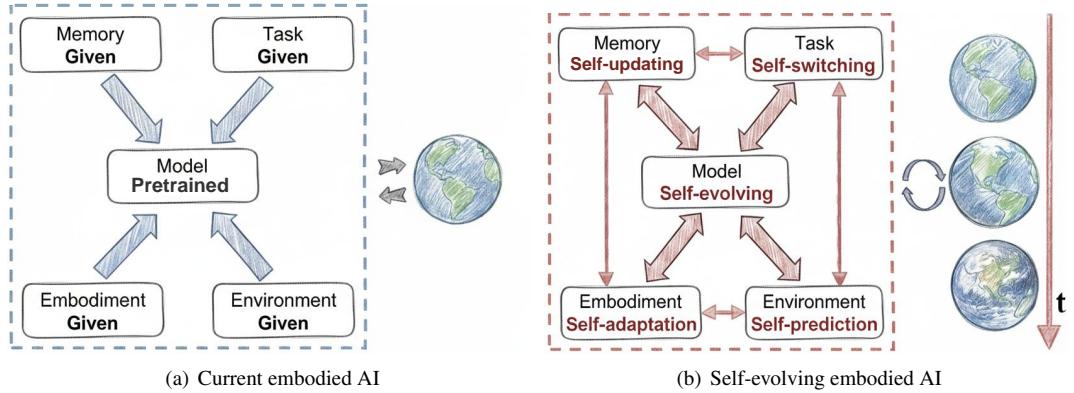
However, real-world scenarios often involve in-the-wild setting. On the one hand, in-the-wild environments are often dynamic and open [12, 13], with continually changing objects, dynamics, and interaction patterns that cannot be fully predefined or exhaustively enumerated during training. On the other hand, in-the-wild deployment often requires variable embodiments [14], where agents differ in morphology, sensing, actuation, computational configurations, and physical constraints, resulting in policies and models learned under fixed embodiments being difficult to adapt. Unfortunately, existing embodied AI fails in in-the-wild setting.

Motivated by neuroscience [15] and cogni-

Department of  
Computer Science and  
Technology, Beijing  
National Research  
Center for Information  
Science and  
Technology, Tsinghua  
University, Beijing  
100084, China

\*Corresponding authors.  
Email:  
wwzhu@tsinghua.edu.cn;  
xin.wang@tsinghua.edu.cn.

Received: XX XX Year;  
Revised: XX XX Year;  
Accepted: XX XX Year



**Figure 1.** Comparison between current embodied AI and self-evolving embodied AI. **(a) Current embodied AI** operates based on *given* task, *given* memory, *given* embodiment, and *given* environment to *pretrain* the corresponding model, which relies heavily on external human guidance and empirical configurations. **(b) Self-evolving embodied AI** operates based on its changing state and dynamic environment with memory *self-updating*, task *self-switching*, embodiment *self-adaptation*, environment *self-prediction*, and model *self-evolution*, aiming to achieve continually adaptive intelligence over time  $t$  with autonomous evolution.

tive science [16], we introduce a new paradigm of *self-evolving embodied AI* in this paper, as illustrated in Figure 1. Self-evolving embodied AI requires agents to operate based on their changing state and environment with memory self-updating, task self-switching, environment self-prediction, embodiment self-adaptation, as well as model self-evolution. In this paradigm, agents can achieve continually adaptive intelligence with autonomous evolution, enabling embodied AI to move from human-crafted setting to in-the-wild setting.

Specifically, we first present the definition, framework, components, and mechanisms of self-evolving embodied AI. We then systematically review state-of-the-art works for each self-evolving component. Furthermore, we discuss practical applications that demonstrate the advantages of self-evolving embodied AI in variable embodiments and dynamic open environments. Finally, we summarize future research directions toward building controllable, trustworthy, and swarm self-evolving embodied AI. We anticipate with full confidence that self-evolving embodied AI is able to learn autonomously in a human-like manner and provide a new perspective toward general artificial intelligence.

## 2. Self-evolving Embodied AI

This section introduces *self-evolving embodied AI*, a new paradigm that aims to move embodied AI beyond human-crafted setting toward in-the-wild setting. We first provide a precise definition and unified framework that distinguishes self-evolving embodied AI from existing paradigms. We then elaborate on the core components of self-evolving embodied AI, ex-

plaining what evolves, why it evolves, when it evolves, and how it evolves for each component. Finally, we discuss self-evolving mechanisms, how to achieve continually adaptive intelligence with autonomous evolution driven by the agent's own state and interaction with dynamic open environments.

### 2.1 The Definition and Framework

This subsection establishes the conceptual foundation of self-evolving embodied AI. We provide a precise definition of self-evolving embodied AI and then introduce a unified framework to describe how self-evolution is organized.

**Definition 1** (Self-evolving embodied AI). *The agent operates based on its own changing internal state and external environment with memory self-updating, task self-switching, embodiment self-adaptation, environment self-prediction, and model self-evolution, aiming to achieve continually adaptive intelligence with autonomous evolution.*

**Framework.** Figure 1(b) illustrates the framework of self-evolving embodied AI. Specifically, the agent is organized around five tightly coupled modules, namely memory self-updating, task self-switching, environment self-prediction, embodiment self-adaptation, and model self-evolution. These modules are not independent, but evolve jointly through continuous interaction with the environment and with each other. Updates in one module may induce adjustments in others. This organization is reflected by bidirectional information exchange among modules across time. For example, changes in embodiment constraints may alter the relevance of

stored memory, which in turn affects task switching and model evolving. Together, they form a unified evolutionary loop that supports continually adaptive intelligence under variable embodiments and dynamic open environments.

For comparison, Figure 1(a) summarizes the framework of existing embodied AI. Specifically, the agent, with given memory, task, and embodiment configuration, learns a pretrained model for predefined objectives in a specific experimental environment. This framework is largely unidirectional, shown as one-way arrows pointing toward the pretrained model.

The contrast between the two frameworks highlights a fundamental difference in how embodied AI is realized. Existing embodied AI relies on given components and pretrained models with external human guidance and empirical configurations, whereas self-evolving embodied AI treats memory, task, environment, embodiment, and model as evolving components driven by the agent itself. Compared to existing embodied AI operating under relatively static environments, by replacing unidirectional optimization with bidirectional coupling among self-evolving modules, self-evolving embodied AI enables continuous adaptation with autonomous evolutionary capability in dynamic open environments over time.

## 2.2 The Components

This subsection elaborates on the core components of self-evolving embodied AI: memory self-updating, task self-switching, environment self-prediction, embodiment self-adaptation, and model self-evolution. For each component, we explain what evolves, why evolution is necessary, when evolution is triggered, and how evolution is realized in details.

*Memory self-updating.* What evolves in memory self-updating is the agent's internal memory representation. Experience is continuously generated through the agent's interaction with the environment, while memory self-updating determines how experience is selectively retained, revised, or discarded over time. Why evolution is necessary lies in the fact that fixed memories or static datasets cannot support long-term adaptation under environmental distribution shift and changing embodiments or tasks. In in-the-wild environments, storing all past experiences is neither feasible nor desirable. When evolution is triggered, previously stored experiences become outdated, irrelevant, or misleading with

respect to the agent's current embodiment constraints, task objectives, or environmental dynamics. How evolution is realized is through selective memory updating mechanisms, such as memory self-editing, memory self-organization, and memory self-distillation, which prioritize experiences based on relevance, novelty, uncertainty, or long-term utility.

*Task self-switching.* What evolves in task self-switching is the agent's internal representation of task objectives. Rather than optimizing predefined objectives of given tasks, task self-switching improves the agent's ability to autonomously adjust what it is trying to achieve over time. Why evolution is necessary lies in the fact that predefined objectives of given tasks cannot adequately capture the changing goals, constraints, and opportunities encountered in dynamic open environments. In-the-wild setting rarely present stable, well-defined goals; over time, predefined objectives of given tasks may limit autonomy and prevent effective adaptation. When evolution is triggered, current task formulations become infeasible, suboptimal, or misaligned with the agent's internal state, embodiment constraints, or environmental dynamics. How evolution is realized is through autonomous task self-switching mechanisms, such as task self-selection and task self-generation, enabling agents to continuously adjust what they aim to achieve without explicit human re-specification.

*Environment self-prediction.* What evolves in environment self-prediction is the agent's internal representation of the external world. It captures how agents continuously update their understanding of environmental dynamics and make future predictions of the external world. Why evolution is necessary lies in the fact that real-world environments are non-stationary and cannot be fully characterized by fixed or offline-learned models. Agents need to maintain and refine predictive representations that anticipate future states, rewards, and interaction action sequences. When evolution is triggered, newly observed environment states contradict prior predictions, reveal previously unseen dynamics, or expose long-term dependencies that were not captured before. How evolution is realized is through continual refinement of world models based on ongoing interaction, such as understanding world models and generative world models, allowing agents to update their internal representation of the external world and supporting adaptive planning and decision-making over

time.

*Embodiment self-adaptation.* What evolves in embodiment self-adaptation is the agent's internal representation of its own physical state. Self-evolving embodied AI needs to adapt the embodiment's heterogeneous morphology, sensing capabilities, actuation limits, computational configurations, and physical constraints. Why evolution is necessary lies in the fact that embodiments may vary across platforms or change over time due to reconfiguration, wear, damage, or degradation, causing policies learned under fixed embodiments to fail. When evolution is triggered, discrepancies arise between expected and actual own physical states, indicating that previous embodiment assumptions are no longer valid. How evolution is realized is through embodiment-aware adaptation mechanisms, such as embodiment self-reconfiguration, embodiment self-calibration, and embodiment self-recovery, enabling agents to maintain functionality across variable embodiments.

*Model self-evolution.* What evolves in model self-evolution is the agent's internal model design, including model architectures, optimization strategies, and evaluation criteria, rather than only model parameters. Why evolution is necessary lies in the fact that fixed architectures, given training strategies, and predefined evaluation metrics cannot remain effective under long-term changes in memory, tasks, environments, and embodiments. As operating conditions evolve, models that were previously well-optimized may become inefficient, misaligned, or inadequate. When evolution is triggered, persistent performance degradation, increasing uncertainty, or systematic mismatch between evaluation outcomes and real-world behavior indicates that existing model designs are no longer appropriate. How evolution is realized is through adaptive modification of model architectures, adjustment of optimization strategies, and refinement of evaluation criteria, such as model self-restructuring, model self-optimization, and model self-evaluating, enabling the learning process itself to evolve and remain aligned with long-horizon autonomous operation.

### 2.3 Self-evolving Mechanisms

Self-evolving mechanisms describe how the five core components evolve in a unified closed-loop. Rather than evolving independently, there is continuous co-evolution of memory self-updating,

task self-switching, environment self-prediction, embodiment self-adaptation, and model self-evolution through the agent's interaction with the environment. Changes in one component propagate through the loop and induce adaptive responses in others.

At each iteration, the agent first maintains a self-state, including its own physical state through embodiment self-adaptation, its internal cognition state through model self-evolution, and its external world state through environment self-prediction. Based on the agent's self-state, task self-switching determines the objectives to pursue, followed by memory self-updating that selectively curates experience relevant to the current objectives. These signals then drive model self-evolution, which updates model architectures, optimization strategies, and evaluation criteria. The evolved model outputs interaction actions, through which the agent engages with the environment. Subsequently, model self-evolution updates the internal cognition state through model evaluation, while embodiment self-adaptation and environment self-prediction update the own physical state and external world state, respectively. Finally, the agent repeats this loop over time.

Through this closed-loop organization, self-evolving mechanisms operate across multiple time scales and levels of abstraction. Fast adaptation occurs within individual iterations as tasks, memory, and actions are adjusted to the current self-state, while slower evolution reshapes internal cognition state, own physical state, and external world state over extended interaction. Importantly, self-evolution is driven by the agent's self-state transitions by action-feedback interactions rather than by external human guidance and empirical configurations. By coordinating the co-evolution of memory, task, model, embodiment, and environment within a unified loop, self-evolving mechanisms enable continually adaptive intelligence with autonomous evolution in dynamic open environments over time.

### 3. Methodologies

Recent studies have shown a growing interest in self-evolving agents [72–74], which aim to improve adaptability through continual interaction with the environment. Specifically, those methods explore self-evolution by using prompt engineering [75,76], memory selection [23,77–79], and tool refinement technologies [80], and are often built upon pretrained Large Language Models (LLMs) as unified end-to-end

**Table 1.** Taxonomy of representative works for self-evolving embodied AI.

Module	Taxonomy	Representative Works
Memory self-updating	Self-editing	SAGE [17], Mem0 [18], Memory-R1 [19], Memento [20]
	Self-organization	A-MEM [21], MemInsight [22], MemGen [23], ReMe [24], Generative Agents [25]
	Self-distillation	ExpeL [26], AWM [27], MUSE [28]
Task Self-switching	Self-selection	SEC [29], WebRL [30], Agent0 [31], Mobile-Agent-E [32]
	Self-generation	ZeroGUI [33], AgentEvolver [34], WebEvolver [35]
Environment Self-prediction	Understanding WM	DreamerV3 [12], DreamerV4 [36], JEPA [37], V-JEPA [38], EvoAgent [39], NavMorph [40], WormI [41]
	Generative WM	Genie [13], Genie-2 [42], Matrix-Game-2 [43], MineWorld [44], MineDreamer [45], OA [46]
Embodiment Self-adaptation	Self-reconfiguration	GET-Zero [47], BoT [48], PEAC [49]
	Self-calibration	UP-OSI [50], SPI-Active [51], OFCI [52]
	Self-recovery	Fall Recovery [53], Damage Recovery [54]
Model Self-evolution	Self-restructuring	MaskTAS [55], PMoE [56], MoE-Adapters4CL [57], D-MoLE [58], L2R [59], EEP [60]
	Self-optimization	MoE-CL [61], SRT [62], SEAS [63], TTRL [64], Self-Refine [65], Reflexion [66]
	Self-evaluating	LLM-as-a-Judge [67], SER [68], RLME [69], RLIF [70], MAE [71]

backbones [32,81]. Those methods demonstrate the potential of self-evolution without explicit human intervention, but environment perception blocks and embodiment configurations are typically treated as fixed once deployed [82,83], limiting their ability to support long-term autonomy under variable embodiments and dynamic environments.

To the best of our knowledge, this paper introduces the first systematic framework that formalizes self-evolving embodied AI by integrating five core components and their self-evolving mechanisms, which can achieve continually adaptive intelligence with autonomous evolution in dynamic open environments over time. Specifically, self-evolving embodied AI arises from the closed-loop co-evolution of memory self-updating, task self-switching, environment self-prediction, embodiment self-adaptation, and model self-evolution, driven by the agent’s self-state transitions by action-feedback interactions.

Existing self-evolving agents belong only to a subset of self-evolving embodied AI, which is still in its early stages and has enormous potential for exploration. Based on these perspectives, this section systematically reviews existing state-of-the-art works for each component. Since self-evolution is an end-to-end process, we organize existing methods according to which component

of the self-evolving loop they primarily address.

### 3.1 Memory self-updating

Existing works on memory self-updating can be broadly categorized into three classes: memory self-editing, memory self-organization, and memory self-distillation. These categories primarily differ in how experience is modified, structured, and abstracted to support long-horizon adaptation over time.

**Memory self-editing.** Those works focus on explicit operations that directly modify stored memory content, including addition, update, deletion, merging, and forgetting. SAGE [17] introduces reflective mechanisms that regulate memory retention using forgetting curves. Mem0 [18] formalizes memory editing by extracting salient facts from interaction histories and consolidating them into long-term memory. Memory-R1 [19] learns memory editing operations via reinforcement learning, enabling adaptive control over when to add, update, or delete memory entries. Memento [20] optimizes memory rewriting and retrieval policies to maintain compact and relevant long-term memory. Together, these works treat memory evolution as an operation-driven editing process.

*Memory self-organization.* Those works emphasize evolving the structure of memory rather than the content. A-MEM [21] organizes agent memory as a dynamically indexed and linked note network inspired by Zettelkasten-style knowledge management. MemInsight [22] augments raw episodic memory with semantic structure, enabling more robust retrieval and reuse. MemGen [23] explores generative latent memory, where memory is woven into a continuous latent space that evolves with experience. Related systems such as ReMe [24] and Generative Agents [25] further highlight the role of structured memory in supporting long-term reasoning and consistency. These works view memory evolution as an organizational process.

*Memory self-distillation.* Those works aim to transform episodic experience into reusable knowledge, skills, or workflows. ExpeL [26] extracts abstract insights and rules from interaction trajectories to guide future behavior. AWM [27] stores reusable task workflows, enabling agents to transfer procedural knowledge across tasks. MUSE [28] further organizes memory hierarchically to accumulate strategic and procedural competence over long horizons. Collectively, these works treat memory evolution as a distillation process that converts episodic experiences into a persistent capability.

### 3.2 Task Self-switching

Existing works on task self-switching can be broadly categorized into two classes: task self-selection and task self-generation. These two lines differ in whether tasks are adaptively selected from an evolving candidate set or generated online conditioned on the agent's self-state and environment dynamics.

*Task self-selection.* Those works formulate task switching as a selection or scheduling problem, where an agent dynamically chooses which task to pursue next based on learning progress or interaction feedback. Representative works include Self-Evolving Curriculum (SEC) [29], which learns adaptive task scheduling policies; WebRL [30], which employs a self-evolving online curriculum for web interaction; Agent0 [31], which co-evolves a curriculum agent with an executor; as well as hierarchical embodied systems such as Mobile-Agent-E [32], where high-level managers perform goal or subgoal selection during execution.

*Task self-generation.* Those works focus on creating new tasks online rather than selecting from a predefined pool, enabling task switching driven by the agent's self-state and exploration demands. ZeroGUI [33] automatically generates GUI tasks from interface states to support online learning; AgentEvolver [34] generates new tasks via self-questioning to guide exploration in novel environments; and WebEvolver [35] leverages a coevolving world model to generate self-instructed interaction tasks.

### 3.3 Environment Self-prediction

Existing works on environment self-prediction can be broadly categorized into two classes: understanding World Models (WMs) and generative world models. The former learns predictive latent representations that support predicting and planning without necessarily decoding raw observations, while the latter explicitly models the observation-generation process (e.g., next-frame or rollouts) with probabilistic generative models.

*Understanding world models.* Those works focus on learning predictive latent representations of environment dynamics, typically instantiated via RSSM-style latent dynamics models [12,36] or JEPAs-style predictive representation learning [37,38]. These works can achieve environment self-prediction in self-evolving embodied AI by continually evolving predictive latent representations under distribution shifts, e.g., continual world-model updates for long-horizon tasks in EvoAgent [39]; online-adaptive world modeling for navigation in NavMorph [40]; test-time composition of multiple domain-specific understanding in WorMI [41].

*Generative world models.* Those works explicitly predict future environment observations, commonly implemented via autoregressive Transformers or diffusion models [84]. Recent large-scale generative world models, such as Genie [13], Genie-2 [42], and Matrix-Game-2 [43], learn general-purpose environment dynamics from large-scale video data and enable long-horizon imagination. In Minecraft and related open-world, interactive generative world models such as MineWorld [44] and Mine-Dreamer [45] provide action-conditioned video generation, and OA [46] plans with online world models in continual RL, forming a practical foundation for environment self-prediction when combined with continual data collection and online refinement. In other applications, OA [46]

plans with online world models in continual RL.

### 3.4 Embodiment Self-adaptation

Existing works on embodiment self-adaptation can be broadly categorized into three classes: embodiment self-reconfiguration, embodiment self-calibration and embodiment self-recovery. Embodiment self-adaptation addresses how the agent maintains adaptability under changes in heterogeneous morphology, sensing capabilities, actuation limits, computational configurations, and physical constraints.

*Embodiment self-reconfiguration.* Those works focus on adapting to discrete configuration changes, including heterogeneous morphology, sensing layouts, and computational setups. Those works learn embodiment-conditioned or modular policies that generalize across robot structures, enabling policy reuse without retraining for each embodiment. Representative works include GET-Zero [47], which models embodiment topology via graph transformers for zero-shot generalization; Body Transformer (BoT) [48], which explicitly encodes embodiment structure into policy learning; PEAC [49], which pre-trains transferable representations for cross-embodiment control.

*Embodiment self-calibration.* Those works target continuous changes in actuation limits and physical constraints, such as mass variation, friction, contact dynamics, and sensing noise. This line typically relies on online system identification or latent parameter estimation to recalibrate internal embodiment models during execution. A classical foundation is the universal policy with online system identification (UP-OSI) [50], with recent extensions enabling active or continual calibration in contact-rich scenarios, such as SPI-Active [51] and Online Friction Coefficient Identification (OFCI) [52] for legged robots.

*Embodiment self-recovery.* Those works address sudden degradation or failure of sensing, actuation, or computational resources, requiring rapid restoration of executable behavior. This line emphasizes fault-aware adaptation and recovery rather than parameter refinement. Representative works include quality-diversity based damage recovery [54], which maintains diverse behavioral repertoires for fast adaptation under damage, and learning-based fall or failure recovery policies for legged robots [53].

### 3.5 Model Self-evolution

Existing works on model self-evolution can be broadly categorized into three classes: model self-restructuring, model self-optimization and model self-evaluating.

*Model self-restructuring.* Those works evolve the internal architecture of models through modular growth, expert composition, or adaptive routing to meet across embodiments, environments, and tasks. Representative works include self-supervised architecture search via masked distillation [55], adaptive Mixture of Expert (MoE) models for continual learning [56,57], dynamic adapter or LoRA expert composition [58,59], and modular growth with expert pruning [60].

*Model self-optimization.* Those works focus on autonomously improving training and parameter update strategies through continual tuning, self-generated feedback, or test-time learning. Typical works include continual instruction tuning without human labels [61], learning from language feedback in SRT [62], adversarial self-improvement in SEAS [63], test-time reinforcement learning [64], and iterative self-refinement and reflection-based optimization [65,66].

*Model self-evaluating.* Those works evolve evaluation criteria and reward signals to guide learning in the absence of reliable ground-truth supervision. This line is characterized by LLM-as-a-judge frameworks [67], self-evolved reward learning [68], reinforcement learning from meta-evaluation [69], learning without external rewards [70], and multi-judge co-evolution [71].

## 4. Applications

This section discuss three practical applications, illustrating how self-evolving embodied AI aligns with practical requirements under environment distribution shifts, limited supervision, and variable embodiments: autonomous service robotics, autonomous driving, and autonomous Unmanned Aerial Vehicles (UAVs).

### 4.1 Autonomous Service Robotics

Autonomous service robots operate in open-ended human-centered environments where object layouts, user preferences, and task specifications change over time [85]. Recent advances in open-vocabulary mobile manipulation and generalist robot policies demonstrate strong generalization across tasks and scenes [86–88], but

comprehensive deployment still requires continual adaptation.

In these applications, memory self-updating is essential for prioritizing rare but consequential interaction failures over redundant routine behavior. Task self-switching enables robots to reinterpret and adjust objectives based on user intent and contextual changes. Embodiment self-adaptation supports deployment across heterogeneous platforms with varying sensing, actuation, and payload constraints. Together, these components illustrate how self-evolution enables sustained autonomy beyond human-crafted household benchmarks.

#### 4.2 Autonomous Driving

Autonomous driving represents a safety-critical instantiation of embodied AI [89] in dynamic environments. Real-world traffic exhibits non-stationary dynamics, rare corner cases, and complex interactions with humans [90], which cannot be exhaustively captured during training. Recent work increasingly relies on predictive world models [91] and high-level reasoning [92] to support planning and simulation.

In these applications, environment self-prediction plays a central role by enabling agents to predict future events and compute long-horizon rewards. Model self-evolution supports adaptation to distribution shifts such as weather, road structure, and traffic patterns. However, unlike other applications, autonomous driving highlights the necessity of embodiment self-adaptation, where the agent's action must be regulated to ensure safety, interpretability, and reliability. This application exemplifies the tension between autonomy and control in self-evolving embodied AI.

#### 4.3 Autonomous UAVs

Autonomous UAVs [93,94] operate under fast dynamics, partial observability, and strict physical constraints, making them a natural testbed for self-evolving embodied AI. UAV missions such as inspection [95], planning [96], and search-and-rescue [97] require continual adaptation to dynamic environments [98], weather conditions [99], and task objectives [100].

In these applications, embodiment self-adaptation is critical due to energy limits, actuator uncertainty, and platform heterogeneity. Environment self-prediction supports trajectory imagination and online replanning under fast-changing dynamics. Model self-evolution enables lightweight and efficient updates com-

patible with onboard computational constraints. Compared with ground robots, UAVs emphasize efficient experience self-updating and tightly coupled evolution across embodiment, environment, and model, demonstrating the feasibility of self-evolving AI under stringent physical constraints.

### 5. Future Directions

Self-evolving embodied AI presents a new paradigm for moving embodied AI beyond human-crafted setting to in-the-wild setting, which is still in its early stages and has enormous potential for exploration. In addition to the self-evolution of each component and the co-evolution between components, future progress has three critical questions to address: (1) how to control self-evolution to ensure stability and efficiency, (2) how to make self-evolution trustworthy for safe deployment, and (3) how to scale self-evolution to swarm-level systems.

#### 5.1 Controllable Self-evolution

A core challenge for self-evolving embodied AI is ensuring that the self-evolution process remains controllable. While self-evolution allows agents to adapt to dynamic environments, uncontrolled evolution may result in instability, performance degradation, or unintended behaviors. To address this, future research should explore methods for regulating the scope, rate, and direction of self-evolution. Hierarchical control architectures could separate fast control loops from slower evolution loops, providing a balance between efficiency and stability. Moreover, developing interpretable feedback mechanisms is crucial for understanding and controlling the evolution process. Quantifiable controllability criteria, such as stability margins or adaptation budgets, will be essential for reliable and predictable self-evolution.

#### 5.2 Trustworthy Self-evolution

The applications of self-evolving embodied AI to be deployed in real-world, their self-evolution must be trustworthy. This requires ensuring that self-evolution processes are safe, transparent, and accountable. One significant challenge is managing the non-stationarity introduced by self-evolution, both in the environment and within the agent itself, complicating verification and validation. Future work should focus on monitoring, auditing, and explaining self-evolution processes to allow for greater hu-

man oversight and understanding of what has changed and why. This includes uncertainty-aware evolution and the implementation of rollback mechanisms to undo harmful adaptations. Aligning self-evolution with human values over long periods remains a critical avenue for research when explicit supervision is limited.

### 5.3 Swarm Self-evolution

While existing studies on self-evolving embodied AI primarily focus on individual agents, many real-world applications involve multiple agents operating as a collective. Swarm self-evolution extends the concept of self-evolution to multi-agent systems, where self-evolution occurs simultaneously at both the individual and swarm levels. Challenges in swarm self-evolution include how agents share experiences, coordinate tasks, and co-evolve their internal models under constraints. Future research should explore mechanisms for distributed memory, policy sharing, and role differentiation to support swarm self-evolution. Understanding emergent behaviors from swarm self-evolving agents and ensuring that swarm-level evolution aligns with collective objectives is also crucial. Swarm self-evolution has the potential to unlock adaptive intelligence on a scale that exceeds individual agent capabilities, enabling robust, flexible, and resilient embodied systems.

## 6. Conclusion

This paper presents a new paradigm of self-evolving embodied AI, enabling embodied AI to move from human-crafted setting to in-the-wild setting. Self-evolving embodied AI is an intelligent system formed by agents and their environment through perception self-evolution, cognition self-evolution, and interaction self-evolution. Specifically, agents operate based on their changing state and environment with memory self-updating, task self-switching, environment self-prediction, embodiment self-adaptation, as well as model self-evolution. In this paradigm, agents can achieve continually adaptive intelligence with autonomous evolution over time. We believe that self-evolving embodied AI can provide a new perspective toward general artificial intelligence.

## REFERENCES

1. Turing AM. Computing machinery and intelligence. *Parsing the Turing test: Philosophical and methodological issues in the quest for the thinking computer*, 23–65 (2007).
2. Feng T, Wang X, Jiang YG et al. Embodied ai: From llms to world models [feature]. *IEEE Circuits and Systems Magazine* 2025; **25**: 14–37.
3. Sun F, Chen R, Ji T et al. A comprehensive survey on embodied intelligence: Advancements, challenges, and future perspectives. *CAAI Artificial Intelligence Research* 2024; **3**: 1.
4. Xu Z, Wu K, Wen J et al. A survey on robotics with foundation models: toward embodied ai. *arXiv preprint arXiv:2402.02385* 2024; .
5. Ma Y, Song Z, Zhuang Y et al. A survey on vision-language-action models for embodied ai. *arXiv preprint arXiv:2405.14093* 2024; .
6. Hu Y, Xie Q, Jain V et al. Toward general-purpose robots via foundation models: A survey and meta-analysis. *arXiv preprint arXiv:2312.08782* 2023; .
7. Duan J, Yu S, Tan HL et al. A survey of embodied ai: From simulators to research tasks. *IEEE Transactions on Emerging Topics in Computational Intelligence* 2022; **6**: 230–244.
8. Srivastava S, Li C, Lingelbach M et al. Behavior: Benchmark for everyday household activities in virtual, interactive, and ecological environments. *Conference on robot learning* (2022) 477–490.
9. Brohan A, Brown N, Carbajal J et al. Rt-1: Robotics transformer for real-world control at scale. *arXiv preprint arXiv:2212.06817* 2022; .
10. ONeill A, Rehman A, Maddukuri A et al. Open x-embodiment: Robotic learning datasets and rt-x models. *2024 IEEE International Conference on Robotics and Automation (ICRA)* (2024) 6892–6903.
11. Zitkovich B, Yu T, Xu S et al. Rt-2: Vision-language-action models transfer web knowledge to robotic control. *Conference on Robot Learning* (2023) 2165–2183.
12. Hafner D, Pasukonis J, Ba J et al. Mastering diverse control tasks through world models. *Nature* 2025; 1–7.
13. Bruce J, Dennis MD, Edwards A et al. Genie: Generative interactive environments. *Forty-first International Conference on Machine Learning* (2024) .
14. Liu Y, Chen W, Bai Y et al. Aligning cyber space with physical world: A comprehensive survey on embodied ai. *IEEE/ASME Transactions on Mechatronics* 2025; .
15. Zhu D, Bu Q, Zhu Z et al. Advancing autonomy through lifelong learning: a survey of autonomous intelligent systems. *Frontiers in neurorobotics* 2024; **18**: 1385778.
16. Wang L, Zhang X, Su H et al. A comprehensive survey of continual learning: Theory, method and application. *IEEE transactions on pattern analysis and machine intelligence* 2024; **46**: 5362–5383.
17. Liang X, He Y, Xia Y et al. Self-evolving agents with reflective and memory-augmented abilities. *arXiv preprint arXiv:2409.00872* 2024; .
18. Chhikara P, Khant D, Aryan S et al. Mem0: building production-ready ai agents with scalable long-term memory (2025). URL <https://arxiv.org/abs/2504.19413> 2025; .
19. Yan S, Yang X, Huang Z et al. Memory-r1: Enhancing large language model agents to manage and utilize memories via reinforcement learning. *arXiv preprint arXiv:2508.19828* 2025; .
20. Zhou H, Chen Y, Guo S et al. Memento: Fine-tuning llm agents without fine-tuning llms, 2025. URL <https://arxiv.org/abs/2508.16153>; .

21. Xu W, Liang Z, Mei K *et al.* A-mem: Agentic memory for llm agents. *arXiv preprint arXiv:2502.12110* 2025; .
22. Salama R, Cai J, Yuan M *et al.* Meminsight: Autonomous memory augmentation for llm agents. *arXiv preprint arXiv:2503.21760* 2025; .
23. Zhang G, Fu M and Yan S. Memgen: Weaving generative latent memory for self-evolving agents. *arXiv preprint arXiv:2509.24704* 2025; .
24. Cao Z, Deng J, Yu L *et al.* Remember me, refine me: A dynamic procedural memory framework for experience-driven agent evolution. *arXiv preprint arXiv:2512.10696* 2025; .
25. Park JS, O'Brien J, Cai CJ *et al.* Generative agents: Interactive simulacra of human behavior. *Proceedings of the 36th annual ACM symposium on user interface software and technology* (2023) 1–22.
26. Zhao A, Huang D, Xu Q *et al.* Expel: Llm agents are experiential learners. *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38 (2024) 19632–19642.
27. Wang ZZ, Mao J, Fried D *et al.* Agent workflow memory. *arXiv preprint arXiv:2409.07429* 2024; .
28. Yang C, Yang X, Wen L *et al.* Learning on the job: An experience-driven self-evolving agent for long-horizon tasks. *arXiv preprint arXiv:2510.08002* 2025; .
29. Chen X, Lu J, Kim M *et al.* Self-evolving curriculum for llm reasoning. *arXiv preprint arXiv:2505.14970* 2025; .
30. Qi Z, Liu X, long IL *et al.* Webrl: Training llm web agents via self-evolving online curriculum reinforcement learning. *arXiv preprint arXiv:2411.02337* 2024; .
31. Xia P, Zeng K, Liu J *et al.* Agent0: Unleashing self-evolving agents from zero data via tool-integrated reasoning. *arXiv preprint arXiv:2511.16043* 2025; .
32. Wang Z, Xu H, Wang J *et al.* Mobile-agent-e: Self-evolving mobile assistant for complex tasks. *arXiv preprint arXiv:2501.11733* 2025; .
33. Yang C, Su S, Liu S *et al.* Zerogui: Automating online gui learning at zero human cost. *arXiv preprint arXiv:2505.23762* 2025; .
34. Zhai Y, Tao S, Chen C *et al.* Agentevolver: Towards efficient self-evolving agent system. *arXiv preprint arXiv:2511.10395* 2025; .
35. Fang T, Zhang H, Zhang Z *et al.* Webevolver: Enhancing web agent self-improvement with coevolving world model. *arXiv preprint arXiv:2504.21024* 2025; .
36. Hafner D, Yan W and Lillicrap T. Training agents inside of scalable world models. *arXiv preprint arXiv:2509.24527* 2025; .
37. Assran M, Duval Q, Misra I *et al.* Self-supervised learning from images with a joint-embedding predictive architecture. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023) 15619–15629.
38. Assran M, Bardes A, Fan D *et al.* V-jepa 2: Self-supervised video models enable understanding, prediction and planning. *arXiv preprint arXiv:2506.09985* 2025; .
39. Feng T, Wang X, Zhou Z *et al.* Evoagent: Agent autonomous evolution with continual world model for long-horizon tasks. *arXiv preprint arXiv:2502.05907* 2025; .
40. Yao X, Gao J and Xu C. Navmorph: A self-evolving world model for vision-and-language navigation in continuous environments. *arXiv preprint arXiv:2506.23468* 2025; .
41. Yoo M, Jang J, Yoon S *et al.* World model implanting for test-time adaptation of embodied agents. *arXiv preprint arXiv:2509.03956* 2025; .
42. Parker-Holder J, Ball P, Bruce J *et al.* Genie 2: A large-scale foundation world model. *URL: https://deepmind.google/discover/blog/genie-2-a-large-scale-foundation-world-model* 2024; .
43. He X, Peng C, Liu Z *et al.* Matrix-game 2.0: An open-source real-time and streaming interactive world model. *arXiv preprint arXiv:2508.13009* 2025; .
44. Guo J, Ye Y, He T *et al.* Mineworld: a real-time and open-source interactive world model on minecraft. *arXiv preprint arXiv:2504.08388* 2025; .
45. Zhou E, Qin Y, Yin Z *et al.* Minedreamer: Learning to follow instructions via chain-of-imagination for simulated-world control. *arXiv preprint arXiv:2403.12037* 2024; .
46. Liu Z, Fu G, Du C *et al.* Continual reinforcement learning by planning with online world models. *arXiv preprint arXiv:2507.09177* 2025; .
47. Patel A and Song S. Get-zero: Graph embodiment transformer for zero-shot embodiment generalization. *2025 IEEE International Conference on Robotics and Automation (ICRA)* (2025) 14262–14269.
48. Sferrazza C, Huang DM, Liu F *et al.* Body transformer: Leveraging robot embodiment for policy learning. *arXiv preprint arXiv:2408.06316* 2024; .
49. Ying C, Zhongkai H, Zhou X *et al.* Peac: Unsupervised pre-training for cross-embodiment reinforcement learning. *Advances in Neural Information Processing Systems* 2024; **37**: 54632–54669.
50. Yu W, Tan J, Liu CK *et al.* Preparing for the unknown: Learning a universal policy with online system identification. *arXiv preprint arXiv:1702.02453* 2017; .
51. Sobanbabu N, He G, He T *et al.* Sampling-based system identification with active exploration for legged robot sim2real learning. *arXiv preprint arXiv:2505.14266* 2025; .
52. Kim H, Kang D, Kim MG *et al.* Online friction coefficient identification for legged robots on slippery terrain using smoothed contact gradients. *IEEE Robotics and Automation Letters* 2025; .
53. Yang C, Pu C, Xin G *et al.* Learning complex motor skills for legged robot fall recovery. *IEEE Robotics and Automation Letters* 2023; **8**: 4307–4314.
54. Allard M, Smith SC, Chatzilygeroudis K *et al.* Online damage recovery for physical robots with hierarchical quality-diversity. *ACM Transactions on Evolutionary Learning* 2023; **3**: 1–23.
55. Yan C, Chang X, Li Z *et al.* Masked distillation advances self-supervised transformer architecture search. *The Twelfth International Conference on Learning Representations* (2024) .
56. Jung MJ and Kim J. Pmoe: Progressive mixture of experts with asymmetric transformer for continual learning. *arXiv preprint arXiv:2407.21571* 2024; .
57. Yu J, Zhuge Y, Zhang L *et al.* Boosting continual learning of vision-language models via mixture-of-experts adapters. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2024) 23219–23230.
58. Ge C, Wang X, Zhang Z *et al.* Dynamic mixture of curriculum lora experts for continual multimodal instruction tuning. *arXiv preprint arXiv:2506.11672* 2025; .

59. Araujo V, Moens MF and Tuytelaars T. Learning to route for dynamic adapter composition in continual learning with language models. *Findings of the Association for Computational Linguistics: EMNLP 2024* (2024) 687–696.

60. Liu E, Zhu J, Lin Z *et al.* Efficient expert pruning for sparse mixture-of-experts language models: Enhancing performance and reducing inference costs. *arXiv preprint arXiv:2407.00945* 2024; .

61. Kang J, Huang L, Hou C *et al.* Self-evolving llms via continual instruction tuning. *arXiv preprint arXiv:2509.18133* 2025; .

62. Hu C, Hu Y, Cao H *et al.* Teaching language models to self-improve by learning from language feedback. *arXiv preprint arXiv:2406.07168* 2024; .

63. Diao M, Li R, Liu S *et al.* Seas: Self-evolving adversarial safety optimization for large language models. *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39 (2025) 23778–23786.

64. Zuo Y, Zhang K, Sheng L *et al.* Ttrl: Test-time reinforcement learning. *arXiv preprint arXiv:2504.16084* 2025; .

65. Madaan A, Tandon N, Gupta P *et al.* Self-refine: Iterative refinement with self-feedback. *Advances in Neural Information Processing Systems* 2023; **36**: 46534–46594.

66. Shinn N, Cassano F, Gopinath A *et al.* Reflexion: Language agents with verbal reinforcement learning. *Advances in Neural Information Processing Systems* 2023; **36**: 8634–8652.

67. Zheng L, Chiang WL, Sheng Y *et al.* Judging llm-as-a-judge with mt-bench and chatbot arena. *Advances in neural information processing systems* 2023; **36**: 46595–46623.

68. Huang C, Fan Z, Wang L *et al.* Self-evolved reward learning for llms. *arXiv preprint arXiv:2411.00418* 2024; .

69. Rentschler M and Roberts J. Reinforcement learning from meta-evaluation: Aligning language models without ground-truth labels. *arXiv preprint arXiv:2601.21268* 2026; .

70. Zhao X, Kang Z, Feng A *et al.* Learning to reason without external rewards. *arXiv preprint arXiv:2505.19590* 2025; .

71. Chen Y, Wang Y, Zhu S *et al.* Multi-agent evolve: Llm self-improve through co-evolution. *arXiv preprint arXiv:2510.23595* 2025; .

72. Fang J, Peng Y, Zhang X *et al.* A comprehensive survey of self-evolving ai agents: A new paradigm bridging foundation models and lifelong agentic systems. *arXiv preprint arXiv:2508.07407* 2025; .

73. Gao Ha, Geng J, Hua W *et al.* A survey of self-evolving agents: On path to artificial super intelligence. *arXiv preprint arXiv:2507.21046* 2025; .

74. Wang S, Long Z, Fan Z *et al.* Benchmark self-evolving: A multi-agent framework for dynamic llm evaluation. *Proceedings of the 31st international conference on computational linguistics* (2025) 3310–3328.

75. Ou Y, Zhou W, Ding S *et al.* Symbolic learning enables self-evolving agents. *AI Open* 2025; .

76. Yadav S and Roy L. Adaptive prompt intelligence: Towards self-evolving conversational agents. *2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG)* (2025) 1–6.

77. Cai Y, Hao Y, Zhou J *et al.* Building self-evolving agents via experience-driven lifelong learning: A framework and benchmark. *arXiv preprint arXiv:2508.19005* 2025; .

78. Xu H, Hu J, Zhang K *et al.* Sedm: Scalable self-evolving distributed memory for agents. *arXiv preprint arXiv:2509.09498* 2025; .

79. Huang C, Yu W, Wang X *et al.* R-zero: Self-evolving reasoning llm from zero data. *arXiv preprint arXiv:2508.05004* 2025; .

80. Liu J, Xiong K, Xia P *et al.* Agent0-vl: Exploring self-evolving agent for tool-integrated vision-language reasoning. *arXiv preprint arXiv:2511.19900* 2025; .

81. Thawakar O, Venkatraman S, Thawakar R *et al.* Evolmm: Self-evolving large multimodal models with continuous rewards. *arXiv preprint arXiv:2511.16672* 2025; .

82. Zhang DC, Zhao Y, Wu J *et al.* Evolvesearch: An iterative self-evolving search agent. *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing* (2025) 13134–13147.

83. Sun Z, Liu Z, Zang Y *et al.* Seagent: Self-evolving computer use agent with autonomous learning from experience. *arXiv preprint arXiv:2508.04700* 2025; .

84. Ding J, Zhang Y, Shang Y *et al.* Understanding world or predicting future? a comprehensive survey of world models. *ACM Computing Surveys* 2025; **58**: 1–38.

85. Zhang W, Li Y, Qiao Y *et al.* Effective tuning strategies for generalist robot manipulation policies. *2025 IEEE International Conference on Robotics and Automation (ICRA)* (2025) 7255–7262.

86. Yenamandra S, Ramachandran A, Yadav K *et al.* Homerobot: Open-vocabulary mobile manipulation. *arXiv preprint arXiv:2306.11565* 2023; .

87. Huang S, Chen L, Zhou P *et al.* Enerverse: Envisioning embodied future space for robotics manipulation. *arXiv preprint arXiv:2501.01895* 2025; .

88. Driess D, Xia F, Sajjadi MS *et al.* Palm-e: An embodied multimodal language model 2023; .

89. Yurtsever E, Lambert J, Carballo A *et al.* A survey of autonomous driving: Common practices and emerging technologies. *IEEE access* 2020; **8**: 58443–58469.

90. Schwarting W, Alonso-Mora J and Rus D. Planning and decision-making for autonomous vehicles. *Annual Review of Control, Robotics, and Autonomous Systems* 2018; **1**: 187–210.

91. Peng Q, Bai C, Zhang G *et al.* Navigscene: Bridging local perception and global navigation for beyond-visual-range autonomous driving. *Proceedings of the 33rd ACM International Conference on Multimedia* (2025) 4193–4202.

92. Sun W, Lin X, Shi Y *et al.* Sparsedrive: End-to-end autonomous driving via sparse scene representation. *2025 IEEE International Conference on Robotics and Automation (ICRA)* (2025) 8795–8801.

93. Chang Y, Cheng Y, Manzoor U *et al.* A review of uav autonomous navigation in gps-denied environments. *Robotics and Autonomous Systems* 2023; **170**: 104533.

94. Feng T, Wang X, Han F *et al.* U2udata: A large-scale cooperative perception dataset for swarm uavs autonomous flight. *Proceedings of the 32nd ACM International Conference on Multimedia* (2024) 7600–7608.

95. Li Z, Zhang Y, Wu H *et al.* Design and application of a uav autonomous inspection system for high-voltage power transmission lines. *Remote Sensing* 2023; **15**: 865.

96. Deng W, Feng J and Zhao H. Autonomous path planning via sand cat swarm optimization with multi-strategy mechanism for unmanned aerial vehicles in dynamic environment. *IEEE Internet of Things Journal* 2025; .

97. Feng T, Wang X, Han F *et al.* U2udata+: A scalable swarm uavs autonomous flight dataset for embodied long-horizon tasks. *Thirtieth AAAI Conference on Artificial Intelligence* 2025; .
98. Wang C, Wang J, Shen Y *et al.* Autonomous navigation of uavs in large-scale complex environments: A deep reinforcement learning approach. *IEEE Transactions on Vehicular Technology* 2019; **68**: 2124–2136.
99. Feng T, Li Q, Wang X *et al.* Multi-weather cross-view geo-localization using denoising diffusion models. *Proceedings of the 2nd Workshop on UAVs in Multimedia: Capturing the World from a New Perspective* (2024) 35–39.
100. Du Z, Luo C, Min G *et al.* A survey on autonomous and intelligent swarms of uncrewed aerial vehicles (uavs). *IEEE Transactions on Intelligent Transportation Systems* 2025; .