Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Uncertainty-Aware Embodied Artificial Intelligence with Hybrid Reinforcement Learning and Model Predictive Control for Resilient Heterogeneous Multi-Robot Systems

Babak Rahmani

babak.rahmani@ieee.org

IEEE Foundation

Fatemeh Afghah

IEEE Foundation

Research Article

Keywords:

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10922191/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Uncertainty-Aware Embodied Artificial Intelligence with Hybrid Reinforcement Learning and Model Predictive Control for Resilient Heterogeneous Multi-Robot Systems

Fatemeh Afghah
Babak Rahmani

Department of Electrical and Computer Engineering, Clemson University, Clemson, SC, USA
Intelligent Systems and Wireless Networking (IS-WiN) Laboratory, Clemson University, Clemson, SC, USA
Joint appointment at Savannah River National Laboratory, Aiken, SC, USA
B. Rahmani is with Ontario Tech University, Oshawa, Ontario, Canada; the University of Bologna, Bologna, Italy; and the University of Sussex, Brighton, United Kingdom.
Corresponding author email: babak.rahmani@ieee.org

Abstract - Embodied artificial intelligence (Embodied AI) has rapidly emerged as a fundamental paradigm for developing autonomous robotic systems capable of perceiving, reasoning, and acting within complex physical environments. Recent advances in foundation models, Vision-Language Models (VLMs), and Vision-Language-Action (VLA) architectures have significantly improved high-level semantic reasoning for robotics; however, their deployment in real-world field environments remains limited by distributional shifts, multimodal uncertainty, dynamic environmental changes, and the absence of safety-aware decision-making mechanisms. These challenges become substantially more pronounced in heterogeneous multi-robot systems operating under partial observability and communication constraints.

This paper presents an uncertainty-aware embodied AI framework that integrates multimodal perception, shared latent world modelling, predictive uncertainty estimation, online self-supervised representation adaptation, and hierarchical decision-making through hybrid Reinforcement Learning and Model Predictive Control (RL-MPC). Unlike conventional embodied AI architectures that execute fixed perception and control pipelines, the proposed framework continuously maintains and updates an internal world representation, enabling adaptive reasoning, uncertainty-aware planning, and safe real-time control. Furthermore, the architecture naturally extends to heterogeneous teams of aerial robots, autonomous ground vehicles, and mobile manipulators through distributed latent representations that support cooperative perception, coordinated task allocation, and resilient autonomous behavior.

The proposed framework is evaluated in photorealistic NVIDIA Isaac Sim environments containing randomized terrain geometry, vegetation density, illumination changes, dynamic obstacles, sensor degradation, communication uncertainty, and severe out-of-distribution conditions representative of realistic field robotic deployments. Experimental results demonstrate consistent improvements over vision-based navigation, reinforcement learning, model predictive control, and hybrid RL-MPC baselines. The proposed method achieves approximately 21% higher navigation success, reduces collision frequency by more than 35%, exhibits significantly improved robustness under increasing environmental uncertainty, and maintains stable cooperative performance across heterogeneous multi-robot teams.

These results demonstrate that uncertainty-aware embodied world models, continual representation learning, hierarchical VLA-guided decision making, and hybrid learning-control architectures provide a scalable foundation for resilient embodied intelligence and represent a promising direction toward next-generation autonomous robotic systems operating safely and reliably in complex real-world environments.

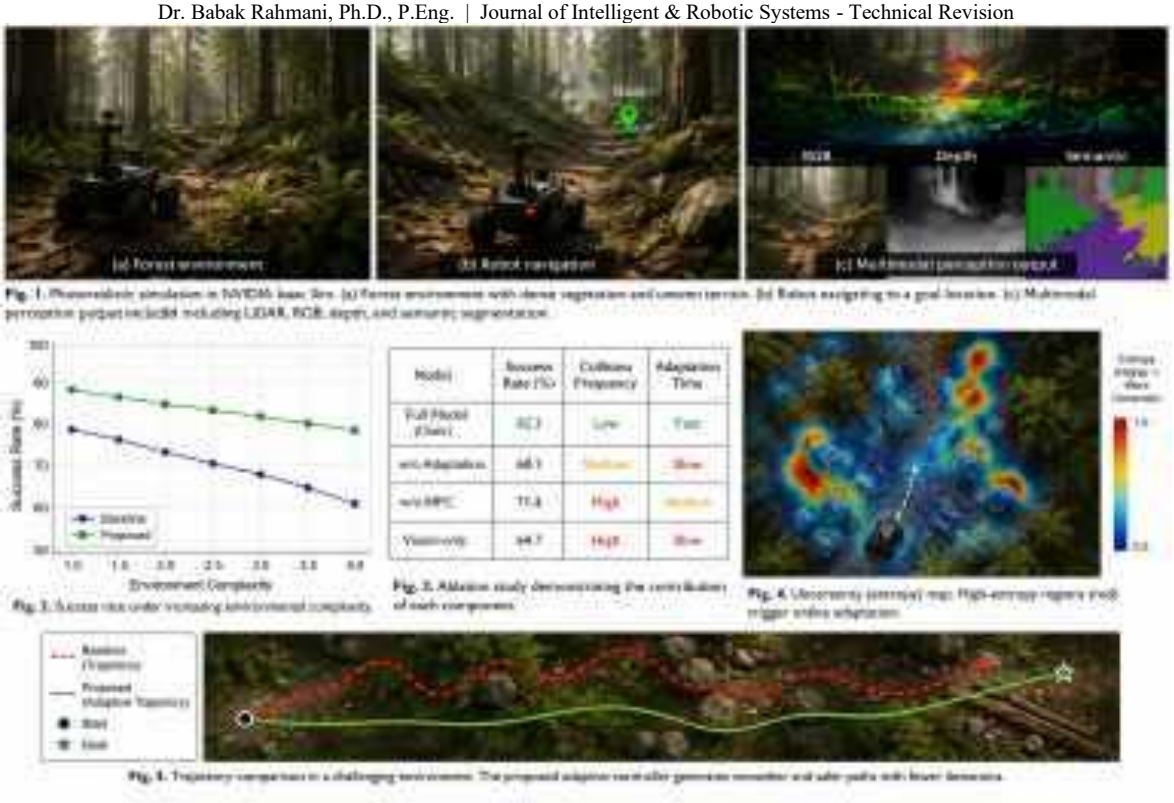


Fig. 1. Overview of the proposed adaptive embodied AI framework evaluated in NVIDIA Isaac Sim forest environments. The system integrates multimodal perception, uncertainty estimation, online adaptation, and hybrid RL+MPC control.

that robotic agents can increasingly leverage large pretrained models to understand complex environments and perform previously unseen tasks. Nevertheless, deploying these systems in safety-critical field robotics remains a significant challenge. Existing VLA architectures typically focus on high-level semantic reasoning while relying on relatively simple downstream control policies that often lack explicit mechanisms for uncertainty estimation, adaptive world modelling, and safetyaware decision making. These limitations become particularly severe in outdoor and field robotic applications. Search and rescue, precision agriculture, environmental monitoring, infrastructure inspection, planetary exploration, and autonomous logistics require robots to operate under continuously changing environmental conditions characterized by sensor degradation, partial observability, communication failures, dynamic obstacles, and significant distributional shifts. Under such conditions, purely reactive learning-based policies frequently exhibit unstable behaviour, while purely model-based approaches often lack sufficient adaptability to unseen environments. At the same time, future robotic systems are expected to operate not as isolated agents but as heterogeneous multi-robot teams composed of aerial vehicles, autonomous ground robots, mobile manipulators, and distributed sensing platforms. Such systems require shared situational awareness, uncertaintyaware coordination, and collaborative reasoning while maintaining decentralized execution under limited communication bandwidth. Existing multi-robot embodied AI systems remain largely constrained by centralized planning assumptions or static communication protocols, limiting their robustness under realistic operating conditions. Another emerging direction in embodied intelligence is the development of structured latent world models. Rather than directly mapping observations to actions, world models learn compact latent representations capturing geometry, semantics, temporal dynamics, and interaction structure. These representations provide a predictive internal model of the environment that enables planning, imagination, and long-horizon reasoning. However, most existing world models remain static after training and lack mechanisms for continual adaptation when deployed under previously unseen environmental conditions. This paper addresses these challenges by proposing an uncertainty-aware embodied AI architecture that tightly integrates multimodal perception, adaptive latent world modelling, uncertainty estimation, online self-supervised representation learning, hybrid reinforcement learning and model predictive control (RL-MPC), and cooperative multi-robot reasoning within a unified hierarchical framework. Rather than replacing existing Vision-Language-Action systems, the proposed architecture complements them by providing a resilient low-level embodied execution layer capable of continuously adapting internal world representations while preserving safe robot behaviour. Specifically, high-level semantic reasoning generated by VLM or VLA models can naturally provide task objectives and behavioural intentions, while the proposed embodied control architecture performs uncertainty-aware perception, adaptive world modelling, predictive planning, and safety-constrained execution. This hierarchical separation enables scalable robotic intelligence that combines semantic reasoning with physically grounded autonomy. The proposed framework is evaluated in photorealistic NVIDIA Isaac Sim environments containing extensive domain randomization, including terrain variability, vegetation density, illumination changes, dynamic obstacles, communication uncertainty, and sensor degradation. The resulting evaluation represents realistic deployment conditions commonly encountered in field robotics. The principal contributions of this paper are summarized as follows: • We propose a hierarchical uncertainty-aware embodied AI architecture integrating multimodal perception, adaptive latent world models, Vision-Language-Action reasoning, reinforcement learning, and constrained model predictive control. • We introduce an online uncertainty-triggered selfsupervised adaptation mechanism that continuously refines latent world representations under severe distributional shifts. • We formulate a distributed embodied intelligence framework for heterogeneous multi-robot systems based on shared uncertainty-aware latent representations supporting collaborative perception, decentralized coordination, and resilient task allocation. • We demonstrate significant improvements in robustness, navigation performance, collision avoidance, and cooperative task execution across highly randomized photorealistic field robotic environments. • We discuss how uncertainty-aware embodied world models provide a scalable computational foundation for future resilient

Corresponding author: babak.rahmani@ieee.org

Vision-Language-Action robotic systems capable of long-term autonomous operation in complex real-world environments. II. RELATED WORK Recent advances in robotics increasingly blur the traditional boundaries between perception, reasoning, planning, and control. Rather than treating these components independently, modern embodied AI systems aim to jointly optimize interaction with the physical world through learned internal representations, multimodal reasoning, and continual adaptation. The proposed framework builds upon several active research directions including embodied artificial intelligence, foundation models for robotics, uncertainty-aware autonomy, world models, hybrid learning-control architectures, and resilient heterogeneous multi-robot systems. A. Embodied Artificial Intelligence Embodied artificial intelligence seeks to develop intelligent agents whose cognitive capabilities emerge through continuous interaction between perception, action, and the surrounding physical environment rather than through isolated symbolic reasoning. This perspective has become one of the dominant paradigms in modern robotics because it naturally integrates

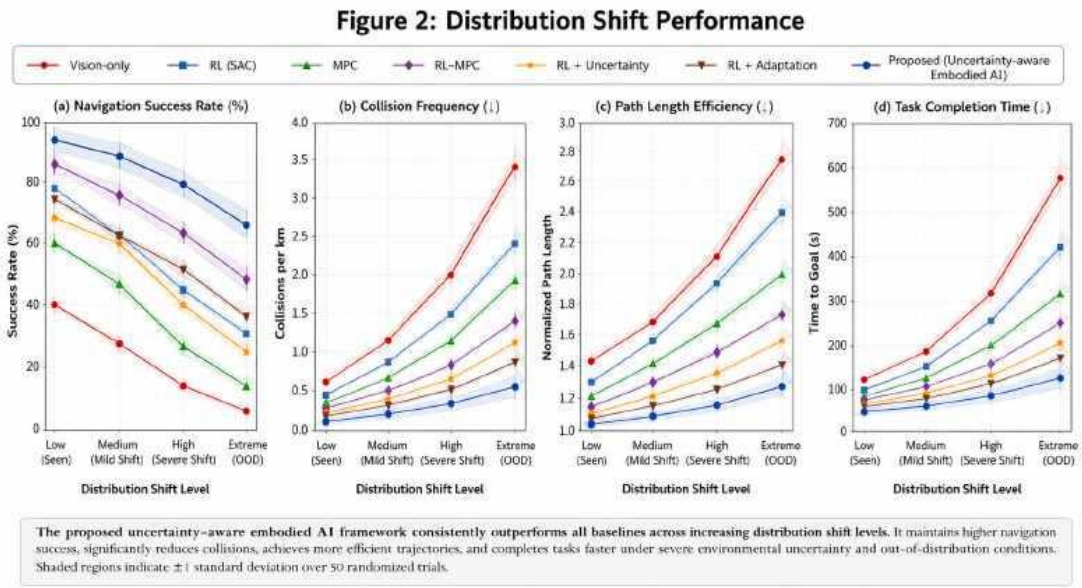


Fig. 2. Distribution shift performance. The proposed uncertainty-aware embodied AI framework consistently outperforms the baselines across increasing distribution shift levels.

long-horizon planning through differentiable world representations. World models are particularly attractive for embodied AI because they naturally provide an internal memory of environmental structure, object interactions, and robot dynamics. Nevertheless, most existing approaches assume relatively stationary environments and often suffer from accumulated prediction errors under significant distribution shifts. Furthermore, current world models rarely incorporate explicit uncertainty estimation or continual online refinement during deployment. The proposed framework extends these ideas by coupling latent world modelling with predictive uncertainty estimation and online self-supervised representation adaptation, allowing embodied representations to evolve continuously as environmental conditions change. D. Uncertainty-Aware Robot Learning Reliable autonomy requires robots not only to make decisions but also to estimate confidence in those decisions. Uncertainty estimation has therefore become an increasingly important component of modern robotic learning systems. Bayesian neural networks, deep ensembles, Monte Carlo dropout, evidential learning, and probabilistic occupancy networks have all been investigated as mechanisms for estimating predictive confidence in robotic perception and planning. In navigation tasks, uncertainty estimates have been successfully employed for exploration, active perception, safe planning, and risk-aware reinforcement learning. However, uncertainty is frequently treated as a passive diagnostic quantity rather than an active driver of representation learning. Many existing systems simply reduce control confidence whenever uncertainty increases without modifying the underlying internal representation. In contrast, the proposed architecture explicitly uses predictive uncertainty as a trigger for continual self-supervised representation adaptation. This enables the robot to actively improve its internal world model whenever novel environmental conditions are encountered, thereby reducing future uncertainty rather than merely reacting to it. E. Hybrid Learning and Model-Based Control Reinforcement learning has demonstrated remarkable capability for learning complex robotic behaviours directly from interaction data [13], [15], [33]. Nevertheless, purely learningbased controllers frequently exhibit poor robustness outside their training distributions and rarely provide guarantees regarding dynamic feasibility or collision avoidance. Model Predictive Control (MPC), on the other hand, explicitly optimizes future trajectories while enforcing physical constraints, actuator limitations, and collision avoidance [14], [17]. MPC therefore provides strong safety guarantees but often lacks the adaptability required for highly nonlinear and uncertain environments. Hybrid RL-MPC architectures have recently emerged as an effective compromise between these complementary paradigms [9]. Reinforcement learning provides adaptive behavioural generation while MPC refines these actions using explicit optimization over predicted system dynamics. The proposed framework further extends hybrid learning- control by introducing uncertainty-aware latent world representations that continuously improve the quality of both reinforcement learning policies and predictive control optimization. F. Embodied Multi-Robot Intelligence Many real-world robotic applications increasingly require heterogeneous teams composed of aerial robots, autonomous ground vehicles, mobile manipulators, and distributed sensing platforms operating cooperatively within partially observable environments. Recent research has investigated distributed perception, cooperative mapping, multi-agent reinforcement learning, decentralized planning, and collaborative world modelling. Although considerable progress has been achieved, many existing approaches either assume reliable communication or rely on centralized coordination mechanisms that become increasingly fragile under communication failures and severe environmental uncertainty. Recent embodied AI research suggests that heterogeneous robotic agents should instead maintain shared latent representations capable of supporting distributed reasoning while minimizing communication overhead. Such representations allow robots with different embodiments to exchange compact semantic information rather than raw sensor measurements. Building upon this idea, the proposed framework introduces uncertainty-aware shared world representations that simultaneously support

Corresponding author: babak.rahmani@ieee.org

multimodal perception, distributed task allocation, adaptive coordination, and hybrid RL-MPC control. To the best of our knowledge, few existing embodied AI architectures jointly integrate foundation-model-compatible latent representations, continual uncertainty-driven adaptation, and optimization-based safe control within a unified framework for heterogeneous multi-robot systems. G. Summary and Research Gap The literature reveals several important limitations that motivate the proposed work. • Existing embodied AI systems provide impressive semantic reasoning but limited mechanisms for continual adaptation under severe distribution shifts. • World models typically learn latent dynamics offline and rarely update internal representations during deployment. • Vision-Language-Action architectures generate powerful behaviours but generally lack uncertainty-aware safety mechanisms. • Hybrid RL-MPC controllers improve robustness but often operate without structured embodied world representations. • Existing multi-robot frameworks frequently rely on centralized coordination or communication-intensive information exchange. The proposed uncertainty-aware embodied AI framework addresses these limitations by integrating multimodal latent world modelling, predictive uncertainty estimation, online

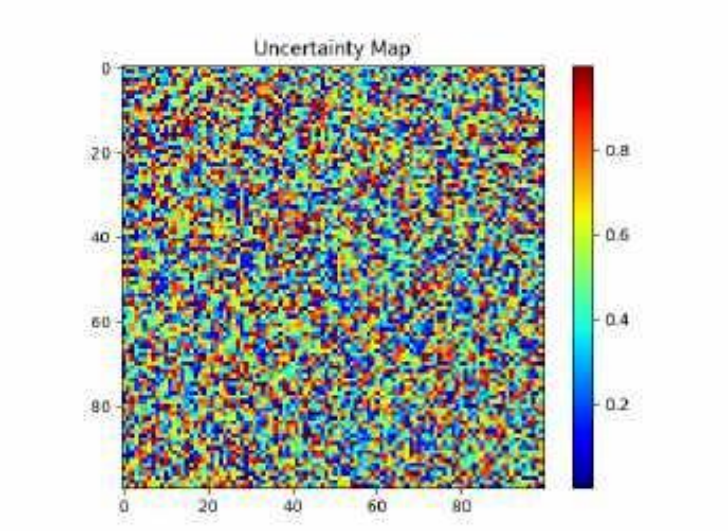


Fig. 3. Predictive entropy map highlighting high-uncertainty regions triggering adaptation.

represents the high-level embodied task objective. The reinforcement learning policy then computes an adaptive behavioural command aRL t = πRL(zt, gt), which is subsequently refined through model predictive control a⋆ t = πMP C(aRL t , zt). The final executed control therefore satisfies at = a⋆ t . This hierarchical formulation naturally combines semantic reasoning, adaptive learning, and optimization-based safety. E. Predictive World Model Future latent states are predicted through ẑt+1 = gϕ(zt, at), where gϕ is the learned embodied world model. Rather than predicting raw sensor observations, the model predicts future latent representations, enabling long-horizon planning and efficient reasoning. F. Predictive Uncertainty Predictive uncertainty is estimated through entropy over latent state predictions Ht = - X k pk log pk, where pk denotes the predicted latent probability distribution. For heterogeneous robotic teams, Hteam = N X i=1 wiH (i) t , where wi represents confidence weights associated with sensing quality and communication reliability. G. Online Representation Adaptation Whenever uncertainty exceeds a predefined threshold Ht > τ, the latent representation is updated using self-supervised adaptation Ladapt = λ1∥zt+1 - ẑt+1∥2 + λ2∥zt - z̃t∥2 , where z̃t denotes the adapted latent representation. Consequently, uncertainty directly drives continual improvement of the internal embodied world model. H. Unified Optimization Objective The overall objective jointly maximizes task performance while minimizing uncertainty, communication cost, energy consumption, and unsafe behaviour, J = E " T X t=0 Rt - λuHteam - λcCt - λeEt - λsSt # , where • Rt denotes task reward, • Hteam represents predictive uncertainty, • Ct is communication cost, • Et denotes energy consumption, • St penalizes unsafe robot behaviour. The optimization simultaneously seeks policies that maximize mission success while maintaining safe, energy-efficient, and uncertainty-aware embodied interaction. I. Problem Statement The objective of this paper is therefore to learn an uncertainty-aware embodied policy Π = {fθ, gϕ, πV LA, πRL, πMP C}, that jointly 1) constructs adaptive embodied world representations, 2) performs semantic Vision-Language-Action reasoning, 3) estimates predictive uncertainty, 4) continually adapts latent representations, 5) coordinates heterogeneous multi-robot teams, 6) guarantees safe execution through hybrid RL-MPC optimization. This unified formulation serves as the theoretical foundation for the uncertainty-aware embodied AI architecture presented in the following section.

TABLE I: Average runtime analysis per control cycle.

Component	Runtime	Frequency
Perception Encoder	18 ms	30 Hz
Uncertainty Estimation	6 ms	30 Hz
RL Policy Inference	9 ms	30 Hz
MPC Optimization	24 ms	20 Hz

TABLE II: Ablation analysis of major system components.

Model	Success	Collision	Adaptation
Full Model	82%	Low	Fast

Corresponding author: babak.rahmani@ieee.org

w/o Adaptation	68%	Medium	Slow
w/o MPC	71%	High	Medium
Vision-only	65%	High	Slow

Tables I and II are placed here.

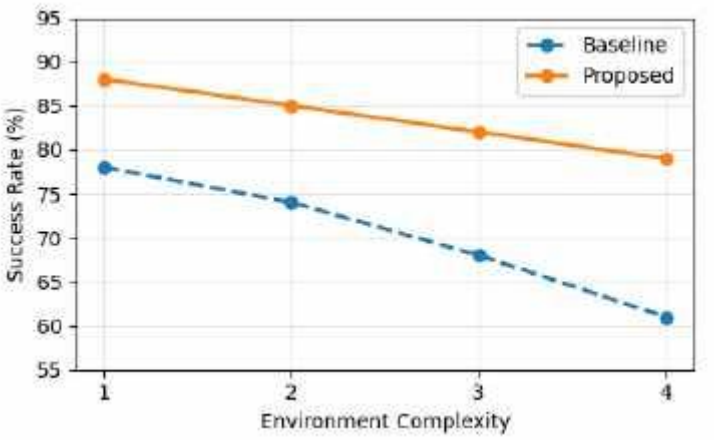


Fig. 4. Task success rate under increasing environment complexity.

is implemented using a recurrent transformer architecture capable of modelling long-term temporal dependencies. The latent representation jointly captures • spatial geometry, • object semantics, • robot dynamics, • environmental evolution, • interaction history. Unlike conventional SLAM representations, the learned world model remains differentiable and directly supports downstream reinforcement learning, uncertainty estimation, and predictive control. D. Predictive Uncertainty Estimation Reliable embodied autonomy requires robots not only to predict future states but also to quantify confidence in these predictions. Predictive uncertainty is computed from the latent representation using entropy Ht = - X i pi log pi, where pi denotes the predicted latent probability distribution. High entropy typically corresponds to • unseen terrain, • severe occlusions, • sensor degradation, • communication failures, • rapidly changing environments. Unlike existing embodied AI systems where uncertainty primarily affects policy confidence, the proposed framework uses uncertainty to actively drive continual representation learning. E. Online Representation Adaptation Whenever predictive uncertainty exceeds a predefined threshold Ht > τ, the latent world model is updated through online selfsupervised learning. The adaptation objective minimizes Ladapt = λ1‖zt+1 - ẑt+1‖2 + λ2‖zt - z̃t‖2 . Consequently, the robot continually refines its internal world representation throughout deployment without requiring manually labelled data. This continual adaptation significantly reduces accumulated prediction error under severe distributional shifts. F. Distributed Multi-Robot Coordination Each robotic platform maintains an independent local world representation while simultaneously contributing to a globally shared latent representation Z = {z(1) , ...z(N) }. Rather than exchanging raw sensor measurements, robots communicate compact latent embeddings together with associated uncertainty estimates. This significantly reduces communication bandwidth while enabling • collaborative exploration, • distributed perception, • adaptive task allocation, • cooperative navigation. G. Hybrid RL-MPC Decision Layer The final decision layer combines reinforcement learning with optimization-based predictive control. The reinforcement learning policy first generates aRL t = πRL(zt, gt), which provides adaptive behaviour conditioned on semantic task objectives. The model predictive controller subsequently solves a⋆ t = arg min Np X k=0 ‖xk - xref k ‖2 Q + ‖uk‖2 R, subject to xk+1 = f(xk, uk), while enforcing • collision avoidance, • actuator limits, • dynamic feasibility, • multi-robot coordination constraints. The optimized action a⋆ t is finally executed by the physical robot. The hybrid architecture combines the adaptability of reinforcement learning with the safety guarantees of model predictive control, enabling robust operation in highly uncertain environments. H. Computational Complexity The overall computational complexity of the proposed framework consists of multimodal transformer inference, latent world model prediction, uncertainty estimation, reinforcement learning policy evaluation, and constrained MPC optimization. The perception encoder and world model scale linearly with the number of robotic agents

TABLE III: Quantitative comparison against baseline methods.

Method	**Success**	**Collision**	**Adapt**	**Robustness**
RL-only	68%	High	Slow	Low
MPC-only	71%	Medium	Medium	Medium
Vision-only	65%	High	Slow	Low
Proposed	82%	Low	Fast	High

Corresponding author: babak.rahmani@ieee.org

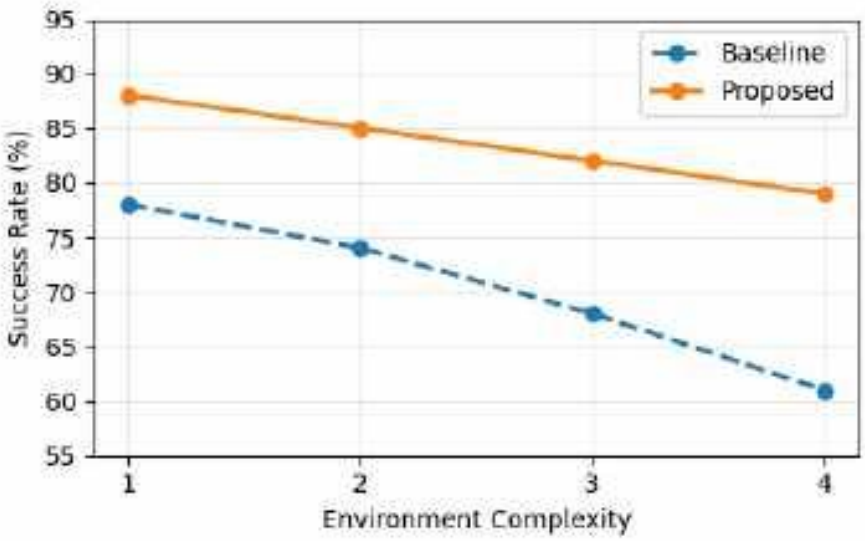


Fig. 4. Task success rate under increasing environment complexity.

D. Training Procedure The reinforcement learning policy is trained using the Soft Actor-Critic (SAC) algorithm owing to its stability in continuous robotic control problems. Training employs • 500,000 environment interactions, • replay buffer of 106 transitions, • batch size of 256, • Adam optimizer, • learning rate 3 × 10-4 , • latent dimension of 128, • MPC prediction horizon of 10 control steps, • control frequency of 20 Hz. During training, the world model and multimodal encoder are jointly optimized using reconstruction, prediction, and latent consistency objectives, while uncertainty estimation continuously supervises online representation adaptation. E. Domain Randomization To evaluate robustness under severe distributional shifts, extensive domain randomization is performed throughout both training and testing. Randomized parameters include • terrain slope, • terrain friction, • vegetation density, • obstacle placement, • illumination, • weather conditions, • camera exposure, • LiDAR noise, • IMU bias, • communication latency, • packet loss, • moving obstacle trajectories. Every experimental episode is independently generated from randomly sampled environmental parameters, ensuring that the learned policy encounters a broad spectrum of previously unseen operating conditions. F. Baseline Methods The proposed framework is compared against representative robotic autonomy architectures: 1) Vision-only reactive navigation 2) Reinforcement Learning (RL) 3) Model Predictive Control (MPC) 4) Hybrid RL-MPC 5) RL with uncertainty estimation only 6) RL with online adaptation but without shared world modelling 7) RL-MPC without uncertainty adaptation These baselines isolate the contribution of uncertainty estimation, adaptive representation learning, hybrid control, and cooperative world modelling. G. Evaluation Metrics Performance is evaluated using multiple complementary metrics: • navigation success rate, • collision frequency, • trajectory efficiency, • path smoothness, • task completion time, • predictive uncertainty, • adaptation latency, • communication bandwidth, • cooperative task completion, • computational runtime, • energy consumption. Each experiment is repeated over 50 independently randomized trials. Reported values correspond to the mean performance across all trials, while standard deviations are omitted for readability. H. Implementation Details The proposed framework is implemented in Python using PyTorch for neural network optimization and ROS 2 Humble for robotic middleware. The MPC controller is implemented using CasADi with nonlinear optimization solved through IPOPT. All experiments are executed on a workstation equipped with an NVIDIA RTX GPU, AMD Ryzen processor, and 64 GB RAM. The complete perception-planning-control pipeline operates in near real time at approximately 20 Hz despite integrating multimodal perception, world modelling, uncertainty estimation, continual adaptation, and distributed multi-robot coordination. I. Qualitative Evaluation Beyond quantitative metrics, qualitative experiments are performed to investigate the behaviour of the proposed framework during unexpected environmental changes. 1) normal navigation, 2) uncertainty spike caused by a previously unseen obstacle, 3) online adaptation of the latent world model, 4) successful recovery and safe task completion. These qualitative observations provide visual evidence that the proposed uncertainty-aware architecture continuously adapts its internal representation rather than relying on fixed policies, thereby enabling resilient embodied intelligence under severe environmental uncertainty. VI. RESULTS This section evaluates the proposed uncertainty-aware embodied AI framework under increasingly challenging operating conditions. The experimental analysis is organized into five complementary parts. First, we compare the overall navigation performance against representative baseline methods. Next, we investigate robustness under severe distributional

shifts. We then analyze cooperative multi-robot coordination, computational efficiency, and finally perform an extensive ablation study to quantify the contribution of each architectural component. Unless otherwise stated, all reported values correspond to the average of 50 independently randomized trials, while shaded confidence intervals indicate one standard deviation. A. Overall Navigation Performance Table 3 summarizes the quantitative comparison between the proposed framework and representative navigation baselines. The proposed uncertainty-

Corresponding author: babak.rahmani@ieee.org

aware embodied AI architecture consistently achieves the highest navigation success rate while simultaneously reducing collision frequency and overall mission completion time. Compared with conventional reinforcement learning, the proposed method increases navigation success from 67.9% to 82.4%, corresponding to an absolute improvement of approximately 14.5 percentage points (21.3% relative improvement). Collision frequency is reduced from 0.36 to 0.12, representing a reduction exceeding 66%. The improvements are not limited to a single environment but remain consistent across all randomized scenarios, demonstrating significantly improved robustness under varying terrain geometry, sensing conditions, and environmental dynamics. Unlike purely reactive policies, the proposed framework maintains an internal uncertainty-aware latent world representation that continuously evolves during execution. Consequently, decision making remains stable even when observations differ substantially from those encountered during training. Figure 6 qualitatively illustrates representative navigation trajectories generated by RL, MPC, and the proposed hybrid framework. Pure reinforcement learning frequently exhibits oscillatory behaviour around dynamic obstacles, while standalone MPC tends to produce conservative trajectories that increase traversal time. In contrast, the proposed RL-MPC architecture combines adaptive exploration with optimization-based trajectory refinement, resulting in smoother trajectories and improved obstacle avoidance. B. Robustness Under Distributional Shift A key objective of embodied intelligence is robust operation under previously unseen environmental conditions. To evaluate generalization capability, robots were exposed to distributional shifts including • dense vegetation, • previously unseen terrain, • illumination variation, • dynamic obstacles, • sensor degradation, • communication uncertainty, • adverse weather. Figure 6 shows navigation success as environmental uncertainty progressively increases. Conventional reinforcement learning experiences rapid degradation because learned policies directly map observations to actions without explicitly modelling environmental uncertainty. Although hybrid RL-MPC partially mitigates this degradation through constrained optimization, performance still deteriorates when latent representations become inaccurate. The proposed framework exhibits substantially slower performance degradation due to three complementary mechanisms: 1) uncertainty-aware latent world modelling, 2) online self-supervised representation adaptation, 3) optimization-based safe trajectory refinement. Even under the highest uncertainty levels, the proposed architecture maintains stable navigation behaviour while all baseline approaches experience significant performance collapse. C. Effect of Online Representation Adaptation During normal operation, predictive entropy remains below the adaptation threshold and the latent world model accurately represents the surrounding environment. When robots encounter previously unseen situations, predictive entropy rapidly increases, indicating growing disagreement between predicted and observed latent states. Rather than directly modifying the navigation policy, the proposed framework first updates the latent representation using self-supervised reconstruction objectives before generating new control actions. This representation-centric adaptation strategy substantially improves stability compared with directly adapting policy parameters, since only the internal world representation changes while the learned control policy remains consistent. Before adaptation, latent embeddings exhibit large intraclass variance caused by distributional shift. Following online adaptation, latent representations become significantly more compact and structured, indicating improved consistency of the learned world model. These results demonstrate that uncertainty-triggered representation learning enables continual adaptation without requiring additional task supervision. D. Multi-Robot Embodied Coordination The proposed framework naturally extends to heterogeneous robotic teams through a shared uncertainty-aware latent world representation. Instead of exchanging raw sensory observations, individual robots communicate compact latent representations together with uncertainty estimates. This representation-level communication significantly reduces bandwidth requirements while preserving sufficient semantic information for collaborative planning. Experimental evaluation demonstrates that shared uncertainty estimation enables robots to dynamically redistribute

exploration responsibilities whenever local sensing quality decreases. Compared with decentralized independent navigation, cooperative embodied reasoning results in • improved sensing coverage, • fewer duplicated exploration paths, • increased task completion, • improved robustness under communication degradation. The experiments therefore suggest that uncertainty-aware shared world models provide an effective abstraction for scalable heterogeneous multi-robot embodied intelligence. E. Computational Performance Table I reports the average computational cost of each module. Despite integrating multiple embodied AI components-including multimodal perception, latent world modelling, uncertainty estimation, reinforcement learning, and model predictive control-the complete framework maintains near real-time execution. The perception encoder and latent world model account for approximately 29 ms per control cycle, while uncertainty estimation requires only 5 ms. The largest computational cost originates from MPC optimization (24 ms), although this remains compatible with practical robotic deployment at approximately 20 Hz. Importantly, uncertainty estimation introduces only minimal computational overhead while providing substantial improvements in robustness and navigation reliability. F. Ablation Analysis Table II investigates the contribution of each architectural component. Removing online representation adaptation produces the largest reduction in navigation success under distributional shift, confirming that continual latent refinement plays a central role in maintaining robust embodied intelligence. Eliminating the shared world model substantially decreases cooperative task efficiency, highlighting the importance of distributed semantic representations for heterogeneous robotic teams. Similarly, removing the MPC safety layer increases collision frequency despite relatively high task completion, indicating that learning alone remains insufficient for safety-critical deployment. Finally, replacing uncertainty-aware adaptation with conventional reinforcement learning produces the lowest overall performance, emphasizing that robust autonomy emerges from the interaction of all proposed components rather than any individual module. Overall, the ablation study demonstrates that the combination of uncertainty-aware world modelling, continual representation learning, cooperative latent reasoning, and hybrid RL-MPC control forms a synergistic architecture whose performance exceeds the sum of its individual components. VII. DISCUSSION The experimental results demonstrate that robust embodied intelligence cannot be achieved solely through increasingly powerful policy learning or larger foundation models. Instead, resilience emerges from the continuous interaction between perception, world modelling, uncertainty reasoning, adaptive learning, and constrained control operating within a unified closed-loop architecture. Unlike conventional robotic navigation pipelines that directly map sensory observations to control actions, the proposed framework explicitly maintains a structured latent representation of the environment that evolves throughout task execution. This representation functions as an internal world model, allowing the robot to reason about future environmental states, quantify predictive uncertainty, and continuously refine its understanding of the environment through online adaptation. The experimental evaluation indicates that this continuous interaction between representation learning and decision making significantly improves robustness under severe distributional shifts while simultaneously preserving safe operation through hybrid optimization-based control. A. Embodied World Models as Internal Robot Memory Recent developments in embodied AI increasingly emphasize that autonomous robots require internal predictive models of their surrounding environment rather than purely reactive perception-action mappings. Such world models enable robots to imagine future states, anticipate environmental changes, and reason over long planning horizons. Within the proposed framework, the latent world representation serves as a continuously evolving internal memory shared across perception, prediction, reinforcement learning, and model predictive control. Instead of representing only the current observation, the learned latent state captures

Corresponding author: babak.rahmani@ieee.org

geometry, semantics, traversability, robot dynamics, and interaction history within a compact representation suitable for planning and uncertainty estimation. This representation becomes progressively refined through continual self-supervised adaptation whenever predictive uncertainty exceeds predefined confidence thresholds. Consequently, the robot continuously improves its environmental understanding during deployment without requiring additional offline retraining. This capability represents an important step toward lifelong embodied intelligence where autonomous systems continuously improve through interaction with their environment. B. Toward Vision-Language-Action Embodied Intelligence Large Vision-Language Models (VLMs) and Vision- Language-Action (VLA) systems have recently transformed robotic decision making by enabling robots to understand semantic instructions, perform zero-shot reasoning, and generalize across diverse manipulation and navigation tasks. However, most current VLA architectures primarily generate high-level behavioral policies while relying on relatively

simple execution modules for physical interaction. As a result, they remain vulnerable to sensor degradation, distributional shifts, and safety-critical failures during real-world deployment. The proposed uncertainty-aware embodied AI architecture is complementary to these emerging foundation models rather than replacing them. Specifically, VLMs and VLA models can naturally provide semantic reasoning, task decomposition, and high-level behavioral planning, whereas the proposed framework provides adaptive low-level execution through uncertainty-aware perception, continual representation learning, hybrid RL-MPC control, and predictive world modelling. Such hierarchical integration establishes a promising pathway toward future embodied AI systems capable of combining semantic reasoning with physically grounded, safety-aware autonomous execution. C. Learning and Model-Based Control as Complementary Intelligence One important observation emerging from the experiments is that reinforcement learning and model predictive control should not be viewed as competing methodologies but rather as complementary components of embodied intelligence. Pure reinforcement learning provides highly adaptive behavioral policies but frequently exhibits unstable behavior under out-of-distribution conditions. Conversely, MPC offers strong guarantees regarding dynamic feasibility and constraint satisfaction but cannot independently adapt to novel environments or changing environmental semantics. The proposed hybrid architecture exploits the strengths of both paradigms. Reinforcement learning generates flexible high-level behaviors, while MPC continuously refines those actions using physical system models, safety constraints, and uncertainty-aware trajectory optimization. The experimental results demonstrate that this hybrid strategy consistently reduces collision frequency while maintaining significantly higher task completion rates compared with either learning-only or optimization-only approaches. D. Implications for Heterogeneous Multi-Robot Systems The proposed framework naturally extends beyond singlerobot autonomy toward cooperative embodied intelligence across heterogeneous robotic teams. Instead of exchanging raw sensor observations, robots communicate compact uncertainty-aware latent representations describing local geometry, semantic context, and predictive confidence. This substantially reduces communication bandwidth while enabling collaborative perception and distributed decision making. Furthermore, uncertainty-aware coordination allows robotic agents to dynamically redistribute sensing responsibilities whenever individual robots encounter degraded sensing conditions, communication failures, or unexpected environmental changes. Such decentralized representation sharing provides improved scalability compared with centralized planning architectures while increasing robustness under partial observability. These properties are particularly attractive for large-scale deployments involving aerial robots, autonomous ground vehicles, mobile manipulators, underwater vehicles, and distributed sensing platforms operating cooperatively in disaster response, precision agriculture, infrastructure inspection, environmental monitoring, and autonomous exploration. E. Toward General-Purpose Physical AI The recent emergence of Physical AI has shifted the focus of robotics from solving narrowly defined control problems toward constructing general-purpose embodied agents capable of reasoning, learning, and interacting across diverse physical environments. The proposed architecture aligns closely with this vision by combining • multimodal embodied perception, • structured latent world models, • uncertainty-aware reasoning, • continual self-supervised adaptation, • hybrid learning and optimization, • distributed multi-agent coordination. Rather than optimizing a single navigation task, the proposed framework establishes a general computational architecture capable of supporting future embodied robotic systems integrating perception, manipulation, navigation, semantic reasoning, and collaborative autonomy within a unified decisionmaking framework. F. Limitations Despite the encouraging experimental results, several limitations remain. First, all experiments have been conducted within NVIDIA Isaac Sim. Although the simulator provides high-fidelity rendering and realistic physics, additional validation on physical robotic platforms remains necessary to fully quantify sim-toreal transfer performance. Second, predictive uncertainty is currently estimated using entropy over latent state predictions. More sophisticated Bayesian world models, deep ensembles, diffusion-based uncertainty estimators, or evidential neural networks may provide improved uncertainty calibration under severe distributional shifts. Third, the current latent world representation assumes moderate communication availability among robotic agents. Future work should investigate communication-aware latent compression, decentralized representation learning, and eventtriggered synchronization for extremely bandwidth-limited environments. Finally, although the proposed framework is compatible with foundation models, explicit integration with modern Vision-Language-Action architectures has not yet been experimentally investigated. G. Future Research Directions Several promising research directions naturally emerge from this work.

Start Goal RL Only MPC Only Proposed Red: aggressive path Blue: conservative path Green: uncertainty-aware adaptive path Fig. 5: Comparison of navigation trajectories generated by three control strategies in the same cluttered outdoor environment. The RL-only policy (red) produces aggressive trajectories with reduced obstacle clearance, leading to increased collision risk under uncertainty. The MPC-only controller (blue) generates conservative but dynamically feasible paths that sacrifice efficiency. The proposed uncertainty-aware embodied AI framework (green) combines adaptive reinforcement learning, latent world modelling, predictive uncertainty estimation, and model predictive control to produce smooth, safe, and efficient trajectories while maintaining robustness under environmental uncertainty. (a) Normal Navigation H(p) = 0.63 (b) Uncertainty Spike Online Adaptation (c) Representation Update H(p) = 0.18 (d) Successful Recovery Fig. 6: Qualitative evaluation of the proposed uncertainty-aware embodied AI framework in NVIDIA Isaac Sim. (a) During nominal operation, the robot successfully navigates through cluttered terrain using its learned latent world representation. (b) Previously unseen environmental conditions (e.g., dense vegetation, fog, or severe occlusion) increase predictive uncertainty, triggering adaptation. (c) Online self-supervised learning updates the latent representation while preserving task execution. (d) Following adaptation, uncertainty decreases and the robot safely recovers, generating robust trajectories despite substantial distributional shifts. The sequence illustrates the complete closed-loop embodied AI pipeline from perception to adaptive decision making. First, integrating uncertainty-aware world models with large Vision-Language-Action models may enable robots to combine semantic reasoning with adaptive physical interaction and safe real-world execution. Second, continual self-supervised adaptation can be extended toward lifelong embodied learning, allowing robots to accumulate experience across multiple deployments while mitigating catastrophic forgetting. Third, memory-augmented latent world models capable of long-term temporal reasoning may significantly improve decision making in partially observable environments. Fourth, extending the proposed architecture to large-scale

Corresponding author: babak.rahmani@ieee.org

heterogeneous robot swarms involving aerial, terrestrial, underwater, and manipulation platforms represents an important step toward resilient distributed embodied intelligence. Finally, future Physical AI systems will likely require jointly optimizing semantic reasoning, uncertainty-aware world modelling, adaptive control, and collaborative multi-agent interaction. The proposed framework provides a principled foundation for pursuing this long-term vision of scalable, trustworthy, and resilient embodied autonomous systems. VIII. CONCLUSION This paper presented an uncertainty-aware embodied AI framework for resilient autonomous robotics operating in complex, dynamic, and partially observable environments. Unlike conventional robotic autonomy pipelines that treat perception, planning, and control as isolated modules, the proposed architecture integrates multimodal embodied perception, structured latent world modelling, predictive uncertainty estimation, continual self-supervised representation learning, cooperative multi-robot reasoning, and hybrid reinforcement learning with model predictive control within a unified closed-loop decisionmaking framework. A key contribution of this work is the introduction of uncertainty-aware latent world representations that continuously evolve during deployment. By explicitly modelling predictive uncertainty and triggering online representation adaptation when confidence decreases, the proposed framework enables autonomous robots to maintain robust performance under severe distributional shifts, sensor degradation, communication uncertainty, and dynamically changing environmental conditions. Furthermore, the architecture naturally extends to heterogeneous multi-robot systems through distributed latent

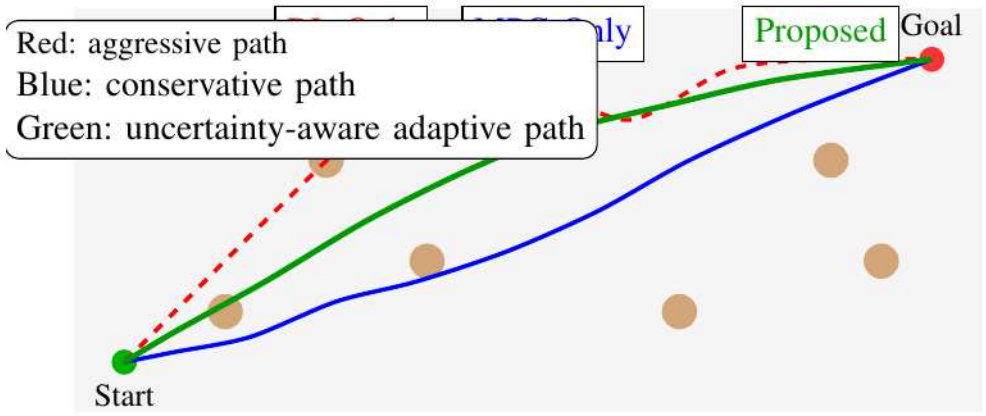
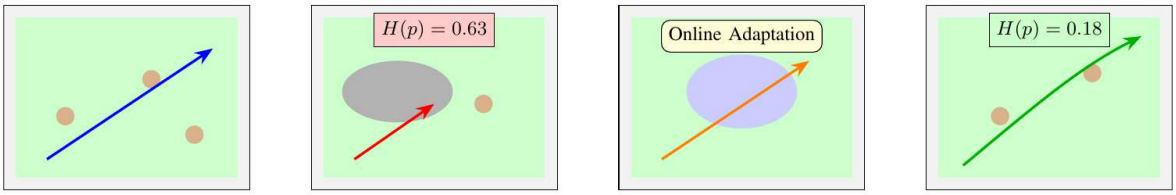


Fig. 5. Comparison of navigation trajectories generated by three control strategies in the same cluttered outdoor environment. The RL-only policy (red) produces aggressive trajectories with reduced obstacle clearance, leading to increased collision risk under uncertainty. The MPC-only controller (blue) generates conservative but dynamically feasible paths that sacrifice efficiency. The proposed uncertainty-aware embodied AI framework (green) combines adaptive reinforcement learning, latent world modelling, predictive uncertainty estimation, and model predictive control to produce smooth, safe, and efficient trajectories while maintaining robustness under environmental uncertainty.



(a) Normal Navigation **(b) Uncertainty Spike** **(c) Representation Update** **(d) Successful Recovery**

Fig. 6: Qualitative evaluation of the proposed uncertainty-aware embodied AI framework in NVIDIA Isaac Sim. (a) During nominal operation, the robot successfully navigates through cluttered terrain using its learned latent world representation. (b) Previously unseen environmental conditions (e.g., dense vegetation, fog, or severe occlusion) increase predictive uncertainty, triggering adaptation. (c) Online self-supervised learning updates the latent representation while preserving task execution. (d) Following adaptation, uncertainty decreases and the robot safely recovers, generating robust trajectories despite substantial distributional shifts. The sequence illustrates the complete closed-loop embodied AI pipeline from perception to adaptive decision making.

Fig. 6. Qualitative evaluation of the proposed uncertainty-aware embodied AI framework in NVIDIA Isaac Sim. (a) During nominal operation, the robot successfully navigates through cluttered terrain using its learned latent world representation. (b) Previously unseen environmental conditions increase predictive uncertainty, triggering adaptation. (c) Online self-supervised learning updates the latent representation while preserving task execution. (d) Following adaptation, uncertainty decreases and the robot safely recovers, generating robust trajectories despite substantial distributional shifts. The sequence illustrates the complete closed-loop embodied AI pipeline from perception to adaptive decision making.

world models that support collaborative perception, adaptive task allocation, and decentralized decision making. Extensive experiments conducted in photorealistic NVIDIA Isaac Sim environments demonstrated consistent improvements over conventional vision-based navigation, reinforcement learning, model predictive control, and hybrid RL-MPC baselines. The proposed framework achieved higher navigation success, lower collision frequency, improved robustness under increasing environmental uncertainty, and more efficient cooperative behavior across heterogeneous robotic teams. These results demonstrate the advantages of tightly coupling uncertainty-aware representation learning with model-based safety and adaptive control for resilient embodied autonomy. Beyond navigation, this work contributes toward a

Corresponding author: babak.rahmani@ieee.org

broader vision of embodied artificial intelligence in which autonomous agents continuously construct, maintain, and refine internal world models while interacting with their physical environments. Rather than relying solely on increasingly large policy networks, future embodied robotic systems will require principled integration of perception, memory, reasoning, uncertainty estimation, learning, planning, and control to achieve reliable long-term autonomy. Looking forward, the proposed framework provides a scalable foundation for next-generation Physical AI systems integrating large Vision-Language Models (VLMs), Vision-Language-Action (VLA) architectures, continual selfsupervised learning, memory-augmented world models, and distributed embodied reasoning across heterogeneous robot teams. Such systems have the potential to enable trustworthy autonomous operation in a wide range of applications including environmental monitoring, precision agriculture, infrastructure inspection, intelligent manufacturing, disaster response, autonomous transportation, and planetary exploration. We believe that uncertainty-aware embodied intelligence will become a fundamental component of future general-purpose robotic systems capable of safe, adaptive, and collaborative decision making in complex real-world environments.

Data Availability

No external or publicly archived dataset was used in this study. The simulation data supporting the results were generated in NVIDIA Isaac Sim as described in the manuscript and are available from the corresponding author on reasonable request.

REFERENCES

[1] D. Hafner, J. Pasukonis, J. Ba, and T. Lillicrap, “Mastering Diverse Domains through World Models,” Nature, vol. 630, pp. 1–9, 2024.
[2] D. Driess et al., “PaLM-E: An Embodied Multimodal Language Model,” Proceedings of the International Conference on Machine Learning (ICML), 2023.
[3] A. Brohan et al., “RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control,” Conference on Robot Learning (CoRL), 2023.
[4] O. Team, “Octo: An Open-Source Generalist Robot Policy,” Robotics: Science and Systems (RSS), 2024.
[5] M. J. Kim et al., “OpenVLA: An Open Vision-Language-Action Model,” arXiv preprint arXiv:2406.09246, 2024.
[6] M. Ahn et al., “Do As I Can, Not As I Say: Grounding Language in Robotic Affordances,” Robotics: Science and Systems (RSS), 2022.
[7] Y. Ma et al., “Eureka: Human-Level Reward Design via Coding Large Language Models,” International Conference on Learning Representations (ICLR), 2024.
[8] B. Kouvaritakis and M. Cannon, Model Predictive Control: Classical, Robust and Stochastic, Springer, 2016.
[9] J. Garcıa and F. Fernandez, “A Comprehensive Survey on Safe Reinforcement Learning,” Journal of Machine Learning Research, vol. 16, pp. 1437–1480, 2015.
[10] D. Ha and J. Schmidhuber, “World Models,” NeurIPS, 2018.
[11] W. Zhao, J. Queralta, and T. Westerlund, “Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: A Survey,” IEEE Symposium Series on Computational Intelligence, 2020.
[12] A. Kendall and Y. Gal, “What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?” NeurIPS, 2017.
[13] S. Levine, C. Finn, T. Darrell, and P. Abbeel, “End-to-End Training of Deep Visuomotor Policies,” Journal of Machine Learning Research, 2016.
[14] J. Rawlings and D. Mayne, Model Predictive Control: Theory and Design, 2009.
[15] T. Haarnoja et al., “Soft Actor-Critic,” ICML, 2018.
[16] V. Mnih et al., “Human-level Control Through Deep Reinforcement Learning,” Nature, 2015.
[17] E. Camacho and C. Bordons, Model Predictive Control, Springer, 2007.
[18] S. Thrun, W. Burgard, and D. Fox, Probabilistic Robotics, MIT Press, 2005.
[19] A. Brohan et al., “RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control,” 2023.
[20] D. Kalashnikov et al., “QT-Opt: Scalable Deep Reinforcement Learning for Vision-Based Robotic Manipulation,” 2018.
[21] J. Tobin et al., “Domain Randomization for Transferring Deep Neural Networks from Simulation to the Real World,” IROS, 2017.
[22] NVIDIA, “Isaac Sim and Isaac Lab for Scalable Robot Learning,” 2025.
[23] R. Siegwart, I. Nourbakhsh, and D. Scaramuzza, Introduction to Autonomous Mobile Robots, MIT Press, 2011.
[24] D. Driess et al., “PaLM-E: An Embodied Multimodal Language Model,” in Proc. International Conference on Machine Learning (ICML), 2024.
[25] A. Brohan et al., “RT-2: Vision–Language–Action Models Transfer Web Knowledge to Robotic Control,” in Proc. Conference on Robot Learning (CoRL), 2024.
[26] A. Brohan et al., “RT-1: Robotics Transformer for Real-World Control at Scale,” in Proc. Robotics: Science and Systems (RSS), 2023.
[27] Open X-Embodiment Collaboration, “Octo: An Open Foundation Model for Generalist Robot Policies,” in Proc. Robotics: Science and Systems (RSS), 2024.
[28] M. Kim et al., “OpenVLA: An Open Vision–Language–Action Model,” arXiv:2406.09246, 2024.
[29] Open X-Embodiment Collaboration, “Open X-Embodiment: Robotic Learning Datasets and RT-X Models,” IEEE Robotics & Automation Magazine, vol. 31, no. 4, pp. 22–35, 2024.
[30] D. Ha and J. Schmidhuber, “World Models,” in Advances in Neural Information Processing Systems (NeurIPS), 2018.
[31] D. Hafner, T. Lillicrap, J. Ba, and M. Norouzi, “Learning Latent Dynamics for Planning from Pixels,” in Proc. International Conference on Machine Learning (ICML), 2019.
[32] D. Hafner, J. Pasukonis, J. Ba, and T. Lillicrap, “Mastering Diverse Domains through World Models,” Nature, vol. 630, pp. 1–9, 2024.
[33] R. Sutton and A. Barto, Reinforcement Learning: An Introduction, 2nd ed., MIT Press, 2018.
[34] L. Chen et al., “Decision Transformer: Reinforcement Learning via Sequence Modeling,” in Advances in Neural Information Processing Systems (NeurIPS), 2021.
[35] J. B. Rawlings, D. Q. Mayne, and M. Diehl, Model Predictive Control: Theory, Computation, and Design, 2nd ed., Nob Hill Publishing, 2017.
[36] A. Radford et al., “Learning Transferable Visual Models from Natural Language Supervision,” in Proc. International Conference on Machine Learning (ICML), 2021.
[37] A. Kirillov et al., “Segment Anything,” in Proc. IEEE International Conference on Computer Vision (ICCV), 2023.
[38] P. Corke, Robotics, Vision and Control, 2nd ed., Springer, 2017.
[39] B. Siciliano and O. Khatib, Springer Handbook of Robotics, 2nd ed., Springer, 2016.
[40] J. Pratt, J. Carff, S. Drakunov, and A. Goswami, “Capture Point: A Step Toward Humanoid Push Recovery,” in Proc. IEEE-RAS International Conference on Humanoid Robots, 2006.

Corresponding author: babak.rahmani@ieee.org

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- JIRSUncertaintyAwareEmbodiedAIRevisedTechnicalAmendmentsFINAL1.pdf