Predictive Temporal Alignment in Motor Skill Acquisition: A Bio-inspired Computational Model of Quadrupedal Locomotion

Ligan Jia (jialg@stu.xju.edu.cn)[1], Minchi Kuang* (kuangmc@mail.tsinghua.edu.cn)[2], Yongchang Kong[3] & Rui Xing[3]

[1]School of Automation, Northwestern Polytechnical University
[2]Department of Precision Instrument, Tsinghua University
[3]School of Computer Science and Technology, Xinjiang University

Abstract

This study explores how agents achieve stable locomotion in complex terrain. They do this relying only on proprioception, without visual guidance. We propose the Asymmetric Reinforcement Learning Contrastive Optimization (ARLC) framework. This framework simulates predictive coding in biological neural systems through a contrastive learning mechanism. The core of this model is to establish temporal consistency between historical action sequences and future expected actions. Experimental results show that this model has excellent generalization ability in the physical world. This includes zero-shot transfer from simulation to reality. More importantly, its learned latent representation space shows a self-organizing mapping of terrain categories. This provides a new computational perspective for understanding the evolution of internal representations in biological motor control.

Keywords: Cognitive science; Quadrupedal Locomotion; intelligent behavior; Reinforcement Learning

Introduction

Cognitive Science Background

Motor control is very important in cognitive science. It is the most direct way for an intelligent agent to interact with the physical world. It is also key to understanding how embodied cognition develops high-level cognitive functions from simple sensorimotor feedback. In the mid-20th century, Nikolai Bernstein proposed the famous Degrees of Freedom Problem. He pointed out that organisms must use cognitive strategies to simplify control dimensions. This is needed to achieve coordinated movement when facing redundant combinations of joints and muscles. This coordination requires compliance with physical laws. It also requires the nervous system to establish extremely robust control representations in dynamic and unpredictable environments.

In contemporary cognitive neuroscience, this representation is often modeled as the evolution of Internal Models (Ehlers et al., 2026). According to the classic theory by Wolpert et al., the brain simulates the sensory consequences of actions through Forward Models (Wolpert et al., 1998). It uses Inverse Models to calculate motor commands for achieving goals. This simulation mechanism allows organisms to predict changes in proprioception caused by movement. This enables feedforward adjustments before the feedback loop is complete. This predictive cognitive strategy greatly reduces risks from neural transmission delays. It also provides necessary physiological assurance for the temporal coherence of complex behaviors.

Furthermore, Predictive Coding and the Free Energy Principle offer a unified computational framework for understanding this mechanism. This theory suggests that the core task of a cognitive system is to minimize prediction error. This is the discrepancy between actual sensory input and internal expectations. During motor learning, an agent continuously compares the expected trajectory of its motor commands with real proprioceptive feedback. This error-based adaptive process prompts brain regions like the Cerebellum to establish precise alignment mechanisms for time, force, and spatial position (Yu et al., 2023). Therefore, motor skill acquisition is essentially a process where a predictive model constantly refines itself. It ultimately achieves high temporal closure between historical experience and future expectations.

However, proprioception is considered the sixth sense for biological motor stability. Its role in computational models is often simplified to instantaneous state feedback. This simplification overlooks proprioception's potential in building long-term body schemas and environmental classification representations (Wei et al., 2022). How agents rely solely on high-dimensional, noisy proprioceptive flow is a challenge. They use temporal contrast and alignment to induce semantically valuable terrain classification representations. This is not only a robotics challenge but also a key way to reveal how natural intelligence distills spatial and physical cognitive laws from time. By building a computational model of this process, we can provide a concrete validation platform for the cognitive hypothesis of action-driven perception (Jang et al., 2020; Jones et al., 2021; Milne et al., 2021).

Research Questions

Deep Reinforcement Learning (DRL) has made significant progress in robot motion control (Agarwal et al., 2023; J. Lee et al., 2020; Rudin, Hoeller, et al., 2022). However, existing computational models still have a clear gap. They struggle to simulate the fluidity and adaptability of biological motion. The most prominent contradiction is the huge difference between high-frequency action sequence jitters and smooth, coherent biological movements. Traditional DRL algorithms often overemphasize maximizing instantaneous rewards when optimizing strategies. They lack constraints on action consistency over time. This reliance solely on the immediate feedback mechanism of Markov Decision Processes (MDP) leads to highly random and oscillatory command streams. This

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

happens when agents face complex or sudden terrain changes. From a cognitive computing perspective, this reflects a lack of a regularization mechanism in existing models that can maintain temporal expectations. This prevents agents from maintaining stability and physical plausibility in their actions, unlike biological organisms (Miki et al., 2022).

Biological control systems are not just passive reactors to current stimuli. They are active systems that constantly generate future expectations (Li et al., 2022; Sohrabi et al., 2022). However, in standard Actor-Critic architectures, policy networks often make isolated decisions. They lack long-range contextual understanding. This mechanism ignores the inherent correlation of movement trajectories over time. Past action experiences should resonate with future expected situations within the same internal representation space. Without this past-future alignment, agents cannot form deep models of their own movement inertia and environmental dynamics at a subconscious level. This leads to inefficient skill learning. It also results in poor generalization and adaptability when facing unseen real-world environments (Smith et al., 2022).

Finally, implicit recognition of environmental information becomes a huge challenge when the system relies only on proprioception (rather than external vision) (Ji et al., 2022; J. Lee et al., 2020). Organisms can infer ground friction, hardness, and geometry from subtle changes in limb touch (Son et al., 2024; H. Zhang et al., 2024). This process is essentially complex cognitive inference (Kirby et al., 2022). However, existing reinforcement learning frameworks often treat proprioceptive signals as meaningless numerical stacks. They fail to extract semantic environmental representations from them. How to induce agents to spontaneously evolve a structured understanding of the physical world (e.g., implicit classification of terrain like stairs and ramps) using only time-series data of limb-environment interaction, without exteroceptive organs, remains an unsolved problem(S. Lee et al., 2022; Oh et al., 2022). The lack of this representation ability directly limits the agent's autonomous navigation and survival in unstructured real environments.

Research Hypotheses

Based on predictive coding theory and the embodied cognition hypothesis, we propose three progressive research hypotheses on temporal alignment-driven intelligent behavior:

Hypothesis on Behavioral Coherence: We believe that biological-level motor fluidity does not come from precise preset muscle strength. Instead, it comes from the internal model's pursuit of temporal consistency in the action flow. We hypothesize that introducing contrastive constraints in the latent representation space can effectively suppress policy drift and high-frequency control noise common in reinforcement learning. This is done by forcing alignment between historical behavior features and future expected trends (Sohrabi et al., 2022; Z. Zhang et al., 2024). This mechanism will induce the agent to shift from passive state feedback to active temporal prediction.

Hypothesis on Representation Self-organization: We assume that structured representations of the environment can emerge without explicit labels. This happens as a byproduct of maintaining predictive consistency. This idea is based on the action-driven perception view (Jones et al., 2021; Milne et al., 2021). Specifically, the latent space self-organizes when an agent compresses proprioceptive time series into consistent latent vectors. It clusters based on different constraints that various terrains (like stairs, ramps, and flat ground) impose on limb dynamics. This means a deep understanding of physical laws doesn't need external visual maps. Instead, it can be achieved by predictively aligning proprioceptive sensory data streams (Ji et al., 2022; Miki et al., 2022).

Hypothesis on Cognitive Generalization: We hypothesize that internal models built on temporal alignment have stronger cross-domain adaptability. This is compared to models that solely fit environmental parameters. This alignment mechanism gives the agent a robustness similar to a biological body schema. It allows the agent to ignore non-essential physical differences between simulated and real-world environments. Examples include minor noise fluctuations or unmodeled resistance. It retains only core dynamic interaction rules. Therefore, this predictive computational framework should have superior Zero-shot Transfer capabilities. It should be able to handle complex, unknown, and unstructured natural environments (J. Lee et al., 2020; Rudin, Kolvenbach, et al., 2022; Smith et al., 2022).

Major Contributions

This paper provides the following four main research contributions:

- **Proposed ARLC Cognitive Computing Framework:** This framework integrates contrastive learning with an asymmetric Actor-Critic architecture. It simulates the complete skill acquisition loop in organisms, from cerebellar prediction to motor internalization.
- **Introduced Predictive Temporal Alignment Mechanism:** By minimizing the representational distance between historical and future action sequences, this mechanism reveals that temporal coherence is key to suppressing motor noise and achieving fluid behavior.
- **Discovered Self-Organizing Terrain Representation:** Experiments confirmed that the model's internal latent space can spontaneously generate terrain classifications of the physical environment. This provides computational evidence for action-driven perceptual classification.
- **Validated Bio-Level Generalization Capability:** Using quadrupedal robots, the model achieved zero-shot transfer in complex real-world environments. This demonstrates the robustness of this model as a general motor intelligence model.

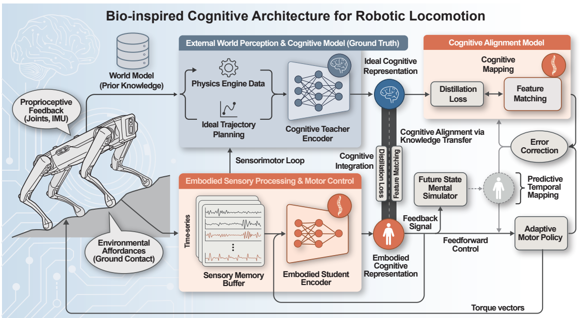

Figure 1: The ARLC Framework: An asymmetric actor-critic architecture featuring a contrastive temporal alignment module for skill internalization.

Methodology

ARLC Architecture

The design of the ARLC (Asymmetric Reinforcement Learning Contrastive) architecture is deeply informed by the teacher-student paradigm of skill internalization and consolidation found in biological motor control systems. It constructs a highly asymmetric neural network feedback loop, the specific structure of which is illustrated in Figure 1.

At the neurobiological level, this architecture simulates how higher motor cortices use rich multimodal information (e.g., visual evaluation, fine touch) as teachers. They train lower motor centers that rely only on proprioception. In our computational model, this asymmetry appears as an information gap between the Critic network and the Actor network. The Critic network acts as a privileged evaluator. During training, it directly accesses physical ground truth in the simulated environment. This includes precise friction coefficients, slope gradients, real-time contact forces at each foot, and center-of-mass linear velocity. This privileged information simulates how organisms acquire experience under supervised learning or visual guidance (Ji et al., 2022; Miki et al., 2022).

In contrast, the Actor network simulates motor skills that are eventually internalized and solidified in the spinal cord and primary motor cortex. It is strictly limited to non-privileged sensory dimensions, receiving only signal streams from the robot's proprioceptors (joint positions, angular velocities, gravity vectors, and action history). To achieve this privilege to proprioception knowledge transfer, we introduced a History Encoder. Its task is to implicitly infer hidden environmental physical features by minimizing the representation distance in the latent space between the privileged encoder and the history encoder:

$$L_{priv} = \|z_{student} - z_{teacher}\|^2 \quad (1)$$

The Actor network is thus forced to learn to reconstruct environmental schemas relying only on the sense of touch and movement inertia (Smith et al., 2022).

The ARLC architecture's most critical cognitive component is its embedded Contrastive Optimizer. This component resembles the biological cerebellum, acting as a predictive hub responsible for coordinating temporal information. We abstract this as temporal alignment in the latent representation space. The hybrid encoder captures historical actions $H_{hist} = \{a_{t-6}, ..., a_t\}$ and future predicted actions $H_{future} = \{a_{t-5}, ..., a_{t+1}\}$. We introduce a contrastive loss function $L_{contrastive}$ based on covariance matrix regularization:

$$L_{contrastive} = L_{diag} + \lambda L_{off\_diag} + D(z_{hist}, z_{future}) \quad (2)$$

This strong temporal coupling constraint forces the Actor network to consider the predictability of future states, effectively simulating the feedforward inhibition function in biological nervous systems (Yu et al., 2023).

Domain Randomization: Adaptive Learning

To give the agent robust adaptability across heterogeneous terrains, we implemented a large-scale Domain Randomization strategy. This can be seen as a computational abstraction of ecological diversity in nature. During biological evolution, species must identify movement invariances despite constant fluctuations in physical environment parameters. Specifically, we extensively perturbed the physical properties of the ground, such as friction coefficients (0.2 to 3.1) and restitution coefficients (0.1 to 1.0) (Ji et al., 2022; J. Lee et al., 2020).

We randomized the agent's Body Schema in addition to environmental features, simulating physiological instability during development or injury. This includes adjusting the robot's base mass (-1.0kg to +5.0kg) and perturbing the Center of Mass (Rudin, Kolvenbach, et al., 2022). Most importantly, we simulated time delays in biological neural conduction, forcing the model to compensate through its internal temporal alignment mechanism. Additionally, we introduced periodic random external impulses (Push Disturbance) of 0 to 10 Newtons to test the efficiency of Perturbation Recovery (Agarwal et al., 2023; Chen et al., 2021). Detailed randomization parameters are provided in Table 1.

Table 1: Domain Randomization Environment Design.

Randomization Name	Range	Unit
Friction	[0.2, 3.1]	-
Restitution	[0.1, 1.0]	-
Base Mass	[-1.0, 5.0]	kg
Center of Mass (CoM)	[-0.1, 0.2]	m
Push Disturbance	10	N
Motor Strength Factor	[0.5, 1.2]	-
k_p (Proportional)	[0.5, 1.1]	-
k_d (Derivative)	[0.5, 1.1]	-

Results

Latent Space Emergence

To explore the evolution of internal cognitive structures as an agent acquires motor skills, we conducted a large-scale sampling and visualization study of the 128-dimensional latent vectors generated by the History Encoder in the ARLC architecture. During this process, we did not provide the agent with any prior labels regarding terrain categories. Instead, we extracted latent representations solely from its movement trajectories across different terrains (flat ground, rough terrain, ramps, stairs).

By applying the t-SNE algorithm to reduce these high-dimensional vectors to a two-dimensional plane, we observed a highly insightful cognitive phenomenon: highly structured cluster centers spontaneously emerged in the latent space. These clusters were not only spatially independent but also showed astonishing consistency with the physical characteristics of the external environment (Oh et al., 2022). This self-organization phenomenon strongly demonstrates that complex motor interactions are sufficient to induce agents to form implicit representations similar to biological terrain concepts (Daunizeau et al., 2025).

In the visualized terrain map, feature vectors for flat ground and slightly rough terrain show large overlaps or adjacent distributions. From a cognitive dynamics perspective, this proximity reflects the agent's essential extraction of their physical commonalities. That is, the gait frequency and the oscillation pattern of the body's center of mass (COM) required to maintain balance are highly similar on both terrains. Conversely, stairs and slopes evolve into independent and clearly bounded clusters. The latent cluster for stair terrain shows a unique step-like periodicity, corresponding to the explosive torque compensation required for discontinuous support surfaces. The cluster for slope terrain shows linear stretching along the direction of the gravity vector's offset, revealing that the model has successfully modeled the continuous functional relationship between posture inclination and gravity compensation (Li et al., 2022).

This self-organizing process deeply reveals the essence of Action-driven Perception. In the ARLC framework, the accuracy of prediction largely depends on whether the model can accurately identify the dynamic constraints imposed by the environment. To minimize prediction error, the encoder must learn to extract the most distinctive terrain features from disorganized proprioceptive flow without supervision. This leap from low-level signals to high-level categorical representations is a key path for embodied cognitive systems to build a world model.

Zero-Shot Transfer: Cognitive Generalization

The model's performance in the physical world provides key experimental evidence for understanding cognitive generalization. In cognitive computing, zero-shot transfer is often considered the gold standard (J. Lee et al., 2020). It measures whether an agent truly grasps underlying physical logic, rather than simply fitting environmental noise. In our experiments, the ARLC-driven agent showed strong adaptability in new outdoor environments, including irregular stone steps, uneven grass, and natural slopes (see Figure 2).

This generalization ability stems from the temporal consistency representation induced by contrastive learning, which is naturally immune to real-world physical disturbances such as minor ground collapse or motor torque fluctuations. From a cognitive modeling perspective, the agent successfully extracted motor schemas that bridge the simulation-to-reality gap (Rudin, Hoeller, et al., 2022).

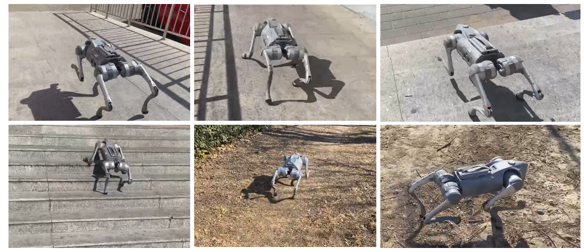

Figure 2: Control Model Deployment in Real Environments.

When switching from an idealized simulator to a real environment full of uncertainty, traditional reinforcement learning strategies often fail due to perceptual bias. However, ARLC agents develop a movement anticipation ability that allows them to offset sudden changes in environmental parameters. In field deployment tests, the robot maintained stable gait frequency even on soft grass and uneven terrain (Agarwal et al., 2023). This success verifies that the Body Schema constructed by this model has cross-platform universality (Handler et al., 2023; Wei et al., 2022). By focusing on conserved features over time, such as momentum transfer laws, the agent exhibits biological-like elastic cognition when facing complex mechanical feedback.

Ablation Study

To analyze each component's contribution to emergent motor intelligence in the ARLC framework, this study designed control experiments in a dynamic physics simulator. We focused on observing cognitive degradation when the agent lost temporal prediction or privileged guidance. Performance evaluation criteria included average episode length and reward value curves (see Figures 3 and 4).

- **Baseline (Standard Markov RL):** This approach uses only the traditional PPO algorithm without temporal alignment or asymmetric architecture. The agent exhibited typical sense-react myopic behavior, generating severe high-frequency control tremors due to a lack of internal modeling for action sequence continuity (Peng et al., 2017). As shown in Figures 3 and 4, the performance only began to rise after 1000 training rounds with significant fluctuations, leading to high mechanical wear.
- **No-Contrastive (Stripping Predictive Alignment):** By removing the contrastive learning loss while keeping the asymmetric architecture, the model lost its predictive coding ability. Although the agent could complete flat-ground tasks, the gait fluidity was significantly reduced. The agent exhibited ataxia-like characteristics, unable to smooth

transitions between actions through feedforward prediction. This strongly supports the hypothesis that temporal alignment is the core cognitive bridge connecting perception and action.

- **No-Asymmetric (Stripping Privileged Information):** We limited the Critic's access to the same dimensions as the Actor. Without the guidance of a privileged teacher, the algorithm stabilized at a lower performance ceiling. The lack of deep understanding of physical ground truth forced the agent to explore blindly in noisy perceptual streams, resulting in low-quality internal models (Ji et al., 2022; Miki et al., 2022).

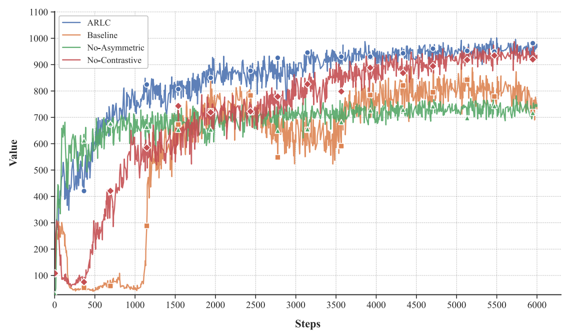

Figure 3: Comparison of average episode length across different ablation models during the training process.

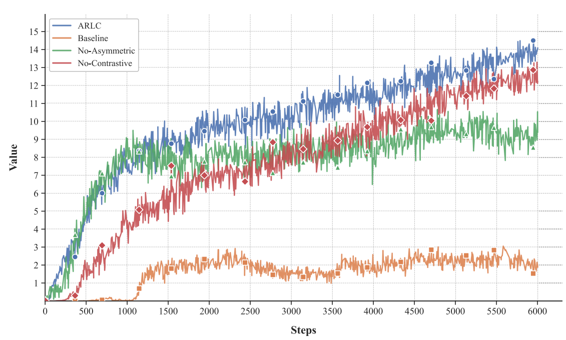

Figure 4: Learning curves of cumulative reward for ARLC and baseline architectures in dynamic environments.

ARLC achieved a peak average reward of 14–15, a nearly five-fold performance increase over the Baseline (around 2–3). In terms of motion stability, its average training steps remained consistently at 900–1000 steps, significantly better than the No-Asymmetric scheme's approximately 750 steps.

More importantly, the torque-velocity distribution in Figure 5 shows that ARLC's motion trajectories exhibit the most compact and geometrically regular clustering features in phase space. This compactness is an effective computational manifestation of solving the degrees of freedom problem, proving that the agent achieves complex coordinated control with extremely low cognitive load and high energy efficiency.

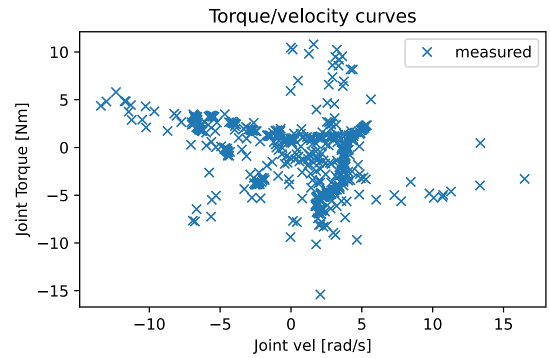

Figure 5: Phase space distribution of joint torque versus angular velocity for the ARLC-driven quadruped.

Discussion

Predictive Coding and Behavioral Fluidization

The results of this study provide strong computational evidence for applying Predictive Coding theory in complex embodied interaction. According to the free energy principle, an intelligent system's core task is to minimize prediction error—the gap between internally generated expectation models and external real-time sensory feedback (Sohrabi et al., 2022).

In the ARLC framework, the Contrastive Loss function acts as this error regulator. It forces the historical representation z_{hist} and future expectation z_{future} in the latent space to align, effectively building an internal probabilistic model of momentum evolution. When the model's trajectory deviates from the optimal path determined by physical inertia, the contrastive loss creates cognitive tension, forcing the Actor network to return to a temporally coherent trajectory. This mechanism matches the neural strategy biological organisms use to maintain movement smoothness, utilizing feedforward predictions to counteract the inherent delay of sensory feedback (Wei et al., 2022).

ARLC reveals that motor smoothness emerges spontaneously as a side effect of minimizing predictive uncertainty. In traditional RL, which lacks temporal ties, agents respond to each state disturbance independently, resembling biological ataxia—jerky movements lacking foresight. ARLC's bidirectional temporal alignment coordination simulates the asynchronous firing of motor units and continuous feedforward regulation found in biological nervous systems (Yu et al., 2023). This proves that temporal predictive consistency is the computational essence of building highly coordinated, fluid behaviors.

Implications for Internal Models

Experimental findings from the ARLC framework provide new insights into the nature of internal models. Traditional symbolism views internal models as static mirrors or explicit maps of the physical environment. However, our results support a more dynamic perspective: internal models are essentially action-oriented implicit expectations. In the ARLC latent space, terrain classification does not exist as abstract geological knowledge; instead, it emerges as a byproduct of system dynamics to maintain predictive coherence (Daunizeau et al., 2025; Oh et al., 2022).

This shift demonstrates the minimal representation principle in embodied cognition. The ARLC agent does not require precise coordinates of obstacles; instead, it infers the evolution of the next action using proprioceptive time series. The observed self-organization in latent clusters represents a functional clustering of environmental challenges based on how different terrains create unique perturbation patterns for limb momentum (J. Lee et al., 2020).

Furthermore, this action-oriented model explains how organisms achieve cognitive coupling between the self and the environment. By establishing long-range temporal alignment, the agent learns a dynamic probability distribution representing the potential for self-motion. The model no longer focuses on what the environment is, but rather on how the system can and should change given the current proprioceptive flow (Kirby et al., 2022). This implicit mechanism incorporates physical causality and may serve as an efficient survival algorithm from natural evolution, allowing agents to adapt to complex ecological niches through deep prediction of limb dynamics, even without fine visual guidance.

The ARLC framework simulates predictive coding at a computational level. However, a fundamental difference exists between our fixed-step discrete prediction and real synaptic delays. Biological synaptic delays are highly random and nonlinear. Our current simplifcation improves training efficiency. But it may limit the agent's "temporal elasticity" in extreme dynamic environments. Therefore, future research will introduce temporal jitter or asynchronous processing into the prediction alignment mechanism. We will also conduct targeted robustness experiments. These tests will verify the model's adaptability to the real temporal dynamics of biological neural systems.

Conclusion

This study proposes the ARLC framework, which achieves a significant performance leap in quadrupedal motion control and provides a computational paradigm for studying the emergence of motor intelligence in cognitive science. By integrating a contrastive learning mechanism that simulates cerebellar predictive functions into an Actor-Critic architecture, our work demonstrates that temporal consistency is the core link connecting perception, prediction, and coordinated movement.

We deployed the motion control model on a physical robot and conducted experiments on real-world terrains, including stairs, slopes, and grass (see experimental results at: https://www.youtube.com/playlist?list=PLFJjnb32o0n9yDMpd1Xt_kVLXfChyBZHD). The results show how the agent maintains fluid motion by relying on the predictive alignment of proprioception rather than exteroceptive signals. The spontaneous emergence of a self-organized cognitive map during this process successfully distinguishes between different terrains (Agarwal et al., 2023; J. Lee et al., 2020). This self-organized cognitive map is a byproduct of efficient skill acquisition and also serves as the cognitive foundation for zero-shot generalization across environments.

From a broader perspective, the success of the ARLC framework reveals an underlying logic that may guide biological intelligence: the construction of action-oriented internal expectation models to offset the limitations of sensory feedback. This implicit representation, built on temporal associations, possesses stronger ecological adaptability than traditional static maps and is more biophysically plausible (Handler et al., 2023). This conclusion challenges the traditional static view of internal models. Furthermore, it provides empirical support for the predictive brain within the field of embodied cognition. Our research demonstrates that complex cognitive classification and environmental inference do not require explicit symbolic guidance; instead, they emerge from the dynamic evolution required to maintain behavioral consistency (Milne et al., 2021).

Future research will further explore the performance of this model in more challenging embodied tasks, such as extending ARLC to complex robot systems with robotic arms for loco-manipulation tasks (H. Zhang et al., 2024). Additionally, we aim to use experimental data from neuroscience to compare the orthogonalization constraints of our contrastive loss function with cerebellar neuron firing patterns. This comparison will help verify the biological authenticity of our framework. In summary, this research provides a direction for developing artificial systems with biological-level flexibility and offers new theoretical references and simulation tools to understand how natural intelligence achieves generalization and skill acquisition in uncertain environments.

References

Agarwal, A., Kumar, A., Malik, J., & Pathak, D. (2023). Legged locomotion in challenging terrains using egocentric vision. *Proceedings of The 6th Conference on Robot Learning*, *205*, 403–415. https://proceedings.mlr.press/v205/agarwal23a.html

Chen, G., Sun, D., Dong, W., Sheng, X., Zhu, X., & Ding, H. (2021). Computationally efficient trajectory planning for high speed obstacle avoidance of a quadrotor with active sensing. *IEEE Robotics and Automation Letters*, *6*(2), 3365–3372. https://doi.org/10.1109/LRA.2021.3062332

Daunizeau, T., Haliyo, S., Gueorguiev, D., & Hayward, V. (2025). Dynamic shaping of multi-touch stimuli by programmable acoustic metamaterial. *Nat Commun*, *16*(1), 8562. https://doi.org/10.1038/s41467-025-63560-6

Ehlers, V. L., Sriram, A., Stuart, B. A. R., Mecca, C. M., & Stucky, C. L. (2026). Sensory neuron piezo1 deletion inhibits dynamic light touch sensitivity in uninjured mice, prevents neuropathic light touch hypersensitivity, and drives compensatory changes in dorsal root ganglia.

Pain, *167*(1), 192–203. https://doi.org/10.1097/j.pain.0000000000003781

Handler, A., Zhang, Q., Pang, S., et al. (2023). Three-dimensional reconstructions of mechanosensory end organs suggest a unifying mechanism underlying dynamic, light touch. *Neuron*, *111*(20), 3211–3229.e9. https://doi.org/10.1016/j.neuron.2023.08.023

Jang, D., Yoo, J., Son, C. Y., Kim, D., & Kim, H. J. (2020). Multi-robot active sensing and environmental model learning with distributed gaussian process. *IEEE Robotics and Automation Letters*, *5*(4), 5905–5912. https://doi.org/10.1109/LRA.2020.3010456

Ji, G., Mun, J., Kim, H., & Hwangbo, J. (2022). Concurrent training of a control policy and a state estimator for dynamic and robust legged locomotion. *IEEE Robotics and Automation Letters*, *7*(2), 4630–4637. https://doi.org/10.1109/LRA.2022.3151396

Jones, T. K., Allen, K. M., & Moss, C. F. (2021). Communication with self, friends and foes in active-sensing animals. *J Exp Biol*, *224*(22), jeb242637. https://doi.org/10.1242/jeb.242637

Kirby, E., Zenha, R., & Jamone, L. (2022). Comparing single touch to dynamic exploratory procedures for robotic tactile object recognition. *IEEE Robotics and Automation Letters*, *7*(2), 4252–4258. https://doi.org/10.1109/LRA.2022.3151261

Lee, J., Hwangbo, J., Wellhausen, L., Koltun, V., & Hutter, M. (2020). Learning quadrupedal locomotion over challenging terrain. *Science Robotics*, *5*(47), eabc5986. https://doi.org/10.1126/scirobotics.abc5986

Lee, S., Lim, H., & Myung, H. (2022). Patchwork++: Fast and robust ground segmentation solving partial under-segmentation using 3d point cloud. *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 13276–13283. https://doi.org/10.1109/IROS47612.2022.9981561

Li, T., Wang, C., Meng, M. Q.-H., & de Silva, C. W. (2022). Attention-driven active sensing with hybrid neural network for environmental field mapping. *IEEE Transactions on Automation Science and Engineering*, *19*(3), 2135–2152. https://doi.org/10.1109/TASE.2021.3077689

Miki, T., Lee, J., Hwangbo, J., Wellhausen, L., Koltun, V., & Hutter, M. (2022). Learning robust perceptive locomotion for quadrupedal robots in the wild. *Science Robotics*, *7*(62), eabk2822. https://doi.org/10.1126/scirobotics.abk2822

Milne, A. O., Orton, L., Black, C. H., Jones, G. C., Sullivan, M., & Grant, R. A. (2021). California sea lions employ task-specific strategies for active touch sensing. *J Exp Biol*, *224*(21), jeb243085. https://doi.org/10.1242/jeb.243085

Oh, M., Jung, E., Lim, H., Song, W., Hu, S., Lee, E. M., Park, J., Kim, J., Lee, J., & Myung, H. (2022). Travel: Traversable ground and above-ground object segmentation using graph representation of 3d lidar scans. *IEEE Robotics and Automation Letters*, *7*(3), 7255–7262. https://doi.org/10.1109/LRA.2022.3182096

Peng, X. B., Berseth, G., Yin, K., & Van De Panne, M. (2017). Deeploco: Dynamic locomotion skills using hierarchical deep reinforcement learning. *ACM Trans. Graph.*, *36*(4), 41. https://doi.org/10.1145/3072959.3073602

Rudin, N., Hoeller, D., Reist, P., & Hutter, M. (2022). Learning to walk in minutes using massively parallel deep reinforcement learning. *Proceedings of the 5th Conference on Robot Learning*, *164*, 91–100. https://proceedings.mlr.press/v164/rudin22a.html

Rudin, N., Kolvenbach, H., Tsounis, V., & Hutter, M. (2022). Cat-like jumping and landing of legged robots in low gravity using deep reinforcement learning. *IEEE Transactions on Robotics*, *38*(1), 317–328. https://doi.org/10.1109/TRO.2021.3084374

Smith, L., Kew, J. C., Bin Peng, X., Ha, S., Tan, J., & Levine, S. (2022). Legged robots that keep on learning: Fine-tuning locomotion policies in the real world. *2022 International Conference on Robotics and Automation (ICRA)*, 1593–1599. https://doi.org/10.1109/ICRA46639.2022.9812166

Sohrabi, F., Jiang, T., Cui, W., & Yu, W. (2022). Active sensing for communications by learning. *IEEE Journal on Selected Areas in Communications*, *40*(6), 1780–1794. https://doi.org/10.1109/JSAC.2022.3155496

Son, C., Kim, J., Kang, D., et al. (2024). Behavioral biometric optical tactile sensor for instantaneous decoupling of dynamic touch signals in real time. *Nat Commun*, *15*(1), 8003. https://doi.org/10.1038/s41467-024-52331-4

Wei, Y., McGlone, F. P., Marshall, A. G., et al. (2022). From skin mechanics to tactile neural coding: Predicting afferent neural dynamics during active touch and perception. *IEEE Trans Biomed Eng*, *69*(12), 3748–3759. https://doi.org/10.1109/TBME.2022.3177006

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in cognitive sciences*, *2*(9), 338–347.

Yu, J., Guo, X., Zheng, S., & Zhang, W. (2023). A dedicate sensorimotor circuit enables fine texture discrimination by active touch. *PLoS Genet*, *19*(1), e1010562. https://doi.org/10.1371/journal.pgen.1010562

Zhang, H., Li, H., & Li, Y. (2024). Biomimetic electronic skin for robots aiming at superior dynamic-static perception and material cognition based on triboelectric-piezoresistive effects. *Nano Lett*, *24*(13), 4002–4011. https://doi.org/10.1021/acs.nanolett.4c00623

Zhang, Z., Jiang, T., & Yu, W. (2024). Localization with reconfigurable intelligent surface: An active sensing approach. *IEEE Transactions on Wireless Communications*, *23*(7), 7698–7711. https://doi.org/10.1109/TWC.2023.3344021