Co-Design and Control of a Biomimetic Snake Robot and its Contact Surfaces with Reinforcement Learning

Liza Darwesh
Vrije Universiteit Amsterdam
Amsterdam, The Netherlands
l.darwesh@student.vu.nl

Sylvia Cressman
Aalto University
Helsinki, Finland
sgrcressman@gmail.com

Riccardo Pretto
University of Padua
Padua, Italy
riccardo.pretto@student.unipd.it

Tiia Tikkala
Technische Universität München
Munich, Germany
tiia.tikkala@tum.de

Shivam Chaubey
Aalto University
Helsinki, Finland
shivam.chaubey@aalto.fi

Kevin Sebastian Luck
Vrije Universiteit Amsterdam
Amsterdam, The Netherlands
k.s.luck@vu.nl

Abstract

Snake robots offer promising capabilities for locomotion in complex environments and terrain due to their modular structure and distributed actuation. However, achieving efficient movement in the real world requires not only effective control policies but also morphological designs adapted to the terrain. This paper investigates real-world co-adaptation in snake robots by coupling Soft Actor-Critic (SAC) reinforcement learning. Specifically, we investigate the effect on locomotion nd control of bio-inspired scale-designs on the contact-surfaces of a snake robot. We evaluate four morphological designs across three physical terrains in the real world. Results demonstrate that morphology significantly influences learning speed and final performance, and that certain designs generalize better across environments. This is a first step towards bio-inspired and -mimetic snake designs utilizig optimized scales combining reinforcement learning with parameterized scale designs.

1 Introduction

The development of robotic snakes capable of navigating confined or cluttered terrains, due to their flexibility and distributed motors, has garnered significant interest in recent years [1, 2]. Their effective deployment in real-world settings requires coordination between body design and control policy. However, the impact of individual design parameters in snake designs in combination with reinforcement learning is not well studied, due to a combination of experimental cost and the simulation-to-reality-gap: Optimizing and studying design parameter of contact surfaces in classic robot simulators, such as Mujoco, is hard, as the interaction between contact-surfaces and ground is hard to model accurately. This leads to a performance gap when applied in physical environments with variable friction and external factors, due to the high number of contacts between snake body and environment. This can be overcome by real-world experiments which are, however, costly and time-intensive. This work presents a first study into bio-mimetic snake skin/scale designs and their evaluation with reinforcement learning in the real world. We deploy Soft Actor-Critic

18th European Workshop on Reinforcement Learning (EWRL 2025).

(SAC), a state-of-the-art reinforcement learning method, to learn effective locomotion strategies directly on hardware, without relying on prior simulation. To create directional friction and improve propulsion, 3D-printed friction scales are attached to the underside of each segment of the snake robot, mimicking the anisotropic surface properties found in biological snakes. Unlike prior work that only adjusts control policies, this approach also explores how changes to physical design can influence performance.

2 Related Work

2.1 Snake Locomotion

The study of snake-inspired locomotion provides the baseline for understanding movement efficiency across diverse terrains. Snakes inhabit a wide range of environments, but across all species, four main forms of locomotion are commonly observed. The first is lateral undulation, a sinusoidal wave-like motion that propels the snake forward [3]. The second is concertina movement, where snakes move through confided spaces by folding and unfolding their bodies like an accordion [4]. The third is rectilinear movement, in which the skin along the underside and lower sides of the snake contracts and extends, allowing it to crawl forward in a straight line [5]. The final form is side-winding, typically used on loose or slippery surfaces, where the snake uses the mid-section of its body to glide sideways across the ground [6]. See Figure 1.

The most common form is the lateral undulation, as it is effective on both smooth and heterogeneous terrains. This is due to their anisotropic scales on their skin. Frictional anisotropy is the property of many surfaces that usually facilitate the generation of motion in a preferred direction [8]. For snakes there is a low coefficient of friction in cranial and lateral direction, and a high friction on the caudal direction.

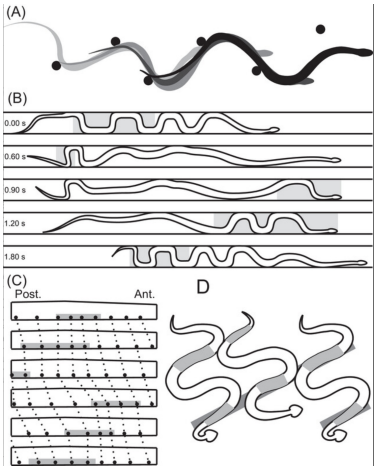


Figure 1: Different locomotion modes in snakes as presented in [7]: (A) Lateral undulation, (B) Concertina movement, (C) Rectilinear movement. (D) Sidewinding.

2.2 Frictional Anisotropic Skins in Robotics

Snake robots mimic this biological principle using artificial frictional surfaces. Two major categories exist: wheeled snake robots and legless designs. The former use passive wheels or ratchets to emulate directional resistance, as seen in Fu and Li's [9] design for navigating large obstacles. In contrast, legless snake robots, such as the design by Moattari and Bagherzadeh [10], utilize interlocking scale structures and compliant joints to generate propulsion through surface friction. Studies show that features like scale angle, density, geometry, and embedding material significantly influence locomotion efficacy.

2.3 Co-adaptation in Robotics

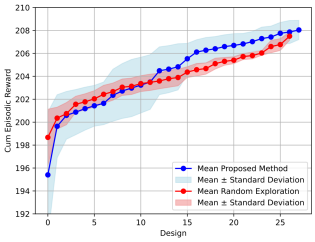
Co-adaptation refers to the simultaneous optimization of morphology and control. While extensively explored in simulation, [11] [12], real-world applications remain rare due to the reality gaps. Luck et al. [13] proposed a deep RL-based framework where a critic guides the search for optimal morphology-control pairs in a bi-level optimization loop to improve efficiency and minimizing the need for physical prototyping. Recent approaches also include real-time adaptive control using Central Pattern Generators (CPGs), such as the work by Ryu et al. [14], which adjusts oscillator frequency in response to changing surface friction.

Despite these advances, real-world co-adaptation in snake robots, especially involving material and shape variability in scale design, has not been deeply studied. Our work builds directly upon these foundations and implements a bi-level co-adaptive loop in a physical snake robot.

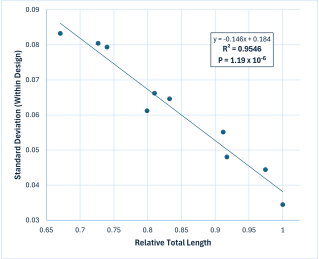
3 Problem Statement

Achieving adaptive and efficient locomotion in real-world snake robots remains a significant challenge in reinforcement learning and bio-inspired robotics. Existing methods often rely on fixed morphologies and conduct training exclusively in simulation, resulting in policies that fail to generalize across diverse real-world terrains. In addition, these approaches often overlook the importance of morphological properties such as anisotropic friction and material compliance, key features that contribute to effective locomotion in biological snakes.

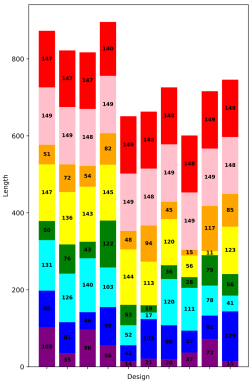
Simulating anisotropic scale interactions accurately remains an open challenge due to the complex frictional and compliant behaviors of soft and structured materials under real-world conditions. As a result, this work addresses the challenge of real-world co-adaptation by using the Soft Actor-Critic (SAC) for policy optimization with empirical evaluation of physical morphology variants. Specifically, we investigate whether a snake robot can autonomously improve its forward locomotion performance by jointly adapting its control policy and scale configuration through repeated physical trials in varied environments. Our central research question is: *Can a snake robot learn to locomote efficiently by co-adapting its control policy and physical morphology directly in the real world?*



(a) Performance per design of randomly selected and optimized designs.



(b) Among top performing designs, within-design variation in segment length correlates negatively with total length,



(c) Best 10 design through all experiments - Higher reward design to the left - First segment (head) is purple, last segment is the red. Lenght values are in mm.

Figure 2: Experimental results of co-adapting the morphology and behaviour of a snake robot. Link segment lengths are adapted over 55 designs, experiment was repeated five times.

4 Co-Design of Module Lengths and Behavior of Snake Robots in Simulation

In our first study we co-optimize the morphology parameters a snake robot in simulation, here the link-lengths of the individual modules, as well as a closed-loop neural network controller producing motor commands. This leads to a bi-level optimization problem with $\max_\xi \max_\pi \mathbb{E}_{\pi,\xi}\left[\sum_t \gamma^t R(s_t, a_t)\right]$ in which we optimize the length parameters ξ in the outer loop and the policy parameters π in the inner loop. In our experiments we use the recently introduced *Fast Evolution through Actor-Critic Reinforcement Learning (FEAR)* [15] algorithm as co-adaptation algorithm, and Soft Actor-Critic as underlying reinforcement learning algorithm. For specific details about FEAR we refer the interested reader to [15]. Initially, 5 randomly selected robot designs are selected and policies learned, which then serve as the starting point for subsequent training phases. Following this, 55 designs are selected and optimized with FEAR. During this phase, the design selection process alternates between selecting a design proposed by the neural network surrogate and selecting a design randomly from a uniform distribution.

The training is conducted with a simulation of the real snake robot using the Mujoco Physics Engine [16]. The learning objective of the robot is to reach a specific goal in $[x, y]$ from its starting position. The reward function defined for the task with $R(s,a) = R_g(s,a) * w_g - R_c(s,a) * w_c$ The reward components are given with $R_g(s,a) = -w * d(s,s_g)^2 - v * log(d(s,s_g)^2 + 10^{-\alpha})$ and $R_c(s,a) = ctrl_{cost} = \frac{\sum_{i=0}^{n-1} |a_i|}{n}$, where $d(s,s_g)$ represents the distance to the goal s_g. w, v, and α are adaptable parameters. The parameter's values are $w = v = 1$ and $\alpha = 3$. The action a_t is the torque applied at the i^{th} servomotor. The weight w_g and w_c are respectively 0.8 and 0.2.

4.1 Simulation Experiments

The morphology, here the individual link-lengths, and behaviour of the snake robot are adapted in a first study in the physics simulator Mujoco. Rewards and goal position $x, y = [2, 2]m$ were chosen to encourage the robot to learn and evolve to traverse the terrain as fast as possible. The design

parameters of the robot were adapted every 300 episodes. The design parameters, i.e. link lengths, can vary from a minimum of $0mm$ and a maximum of $149mm$.

Five experiments were conducted, producing 28 optimized designs and 27 randomly selected designs each. To evaluate the performance of a specific snake robot design we measure the maximal episodic return achieved during the training time of the robot (Fig.2a). Figure 2a compares the designs found using an optimization strategy and using random design sampling. We can find that optimization shows increases performance and data-efficiency versus random sampling, however the results also indicate that a larger size of design adaptations may be needed. Overall, the results underscores the advantage of adapting robot designs and the potential of optimizing the shape of snake robots to increase performance and energy-efficiency.

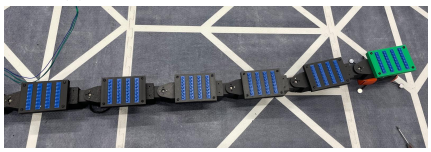
To understand the characteristics of the highest-performing snake designs, we examine the top 10 designs across all experiments. In Fig.(2c), a discernible pattern emerges: the last two segments are consistently close to the longest possible size (red and pink), followed by a shorter third-to-last segment (orange). The fourth segment is medium to long size(yellow), and the fifth segment is short again (green). The first three segments vary across different designs (light blue, blue, and purple). Among top-performers, there is also a strong, significant negative correlation ($\rho = -0.977, \mathrm{P} = 1.19 \times 10^{-6}$) between the total length of a given design and the amount of variation in its segment lengths. Longer designs tend to have more evenly-sized segments, mirroring what is observed in the skeletal structure of snakes [17].

5 Control and Design of Snake Scales in the Real World

5.1 Robot Design and Setup

The robot consists of eight PLA-printed segments connected via three Dynamixel XL430-W250-T servo motors and four XL430-W240-T servo motors , see Figure 3a. Unlike uniform modular designs, each segment in our robot has a distinct length. These lengths were chosen based on the prior simulation experiments that showed improved locomotion performance when varying segment lengths were used (see previous section) . Specifically, we adopted the segment dimensions identified in that work with the highest overall locomotion efficiency. Each segment supports a baseplate designed to accept interchangeable scales, see Figure 3b. The motor actuation range is restricted to $[-55°, 55°]$ to prevent collision with the other segments. OptiTrack cameras (Flex13) are used to track the global position of the snake by using reflective markers mounted on the head of the robot.

(a) Snake Robot in Environment



(b) Bottom view of the snake robot with baseplate and interchangeable scales.

Figure 3: The real world snake robot used in the scale-adaptation experiments.

5.2 Control Policy: Soft Actor-Critic (SAC)

We use the Soft Actor-Critic (SAC) algorithm, a state-of-the-art off-policy reinforcement learning method that used the standard actor-critic framework with entropy regularization. This encourages the policy to promote exploration by introducing random actions during training.

The SAC framework uses two soft Q-functions $Q_1(s, a)$ and $Q_2(s, a)$, a stochastic policy $\pi(a|s)$, and corresponding target networks for both critics. The policy aims to maximize both the expected return and the entropy of the action distribution. The actor objective is given by:

$$J(\pi) = \mathbb{E}_{(s,a)\sim D}\left[Q(s, a) - \alpha \log \pi(a|s)\right], \qquad (1)$$

where α is a fixed temperature coefficient controlling the balance between exploration and exploitation. In our implementation, $\alpha = 0.01$.

The critic networks are trained to minimize the soft Bellman residual using the target Q-values:

$$L_Q = \mathbb{E}_{(s,a,r,s')} \left[\left(Q(s,a) - \left(r + \gamma \min_{i=1,2} Q_i'(s',a') - \alpha \log \pi(a'|s') \right) \right)^2 \right], \tag{2}$$

where $\gamma = 0.99$ is the discount factor and Q_i' denotes the target Q-network. The target networks are updated through an exponential moving average:

$$\theta_{\text{target}} \leftarrow \tau\theta + (1-\tau)\theta_{\text{target}}, \tag{3}$$

with $\tau = 0.01$ as the soft update coefficient.

Network Architecture Each Q-function and policy network is implemented as a fully connected multilayer perceptron (MLP) with three hidden layers of 256 units each, using ReLU activations. The policy outputs the mean and standard deviation of a Gaussian distribution, which is transformed by a Tanh squashing function to ensure bounded actions. We clip the gradient across all layers with a maximum absolute value of 1.0 to prevent exploding gradients.

Optimization Details The actor and both critic networks are optimized using the Adam optimizer with a learning rate of 1×10^{-3}. We use automatic entropy tuning, where the entropy coefficient α is adapted during training to match the target entropy, which is set to the number of motors used in the robot.

Training Protocol Training is conducted by sampling from a global replay buffer. During each episode, the networks undergo 1000 gradient update steps. The replay buffer stores the observation from real-world interaction, with one update batch comprising 32 samples. The observation includes global position changes ($\Delta x, \Delta y, \Delta z$), the robot's orientation in 3D space ($\theta_x, \theta_y, \theta_z$), and the current motor angles (m_1 through m_6). The full observation vector is shown in Equation 4.

$$S_t = [\Delta x, \Delta y, \Delta z, \theta_x, \theta_y, \theta_z, m_1, m_2, m_3, m_4, m_5, m_6] \tag{4}$$

To mitigate latency between sensing and action selection, we employ multi-threaded execution to decouple the control loop from motion tracking and learning updates. the OptiTrack thread continuously reads global position data and posts it to a shared variable, the motor thread monitors motor encoder values and writes them to a separate shared state variable, and the training loop thread synchronizes these inputs to compute and execute actions using the policy network.

5.3 Reward Function

The reward function used for training encourages forward movement and penalizes unstable body orientation. It is defined in Equation 5, where x_{target} is the target x-position, x_{curr} is the current x-position, d_{max} is the maximum expected distance (set to 40), and θ_y is the pitch angle.

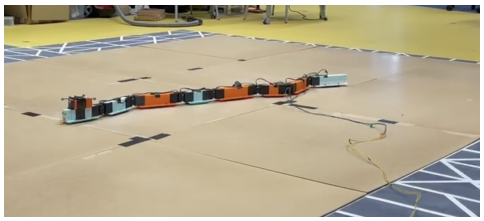
$$r = \exp\left(1 - \frac{|x_{\text{target}} - x_{\text{curr}}|}{d_{\text{max}}}\right) + (0.3 - |\theta_y|) \tag{5}$$

6 Experimental Setup

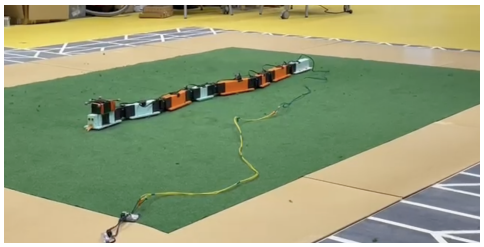
6.1 Environments

We conduct experiments in three environments: a foam mat, a wooden cardboard sheet, and a carpet. These were selected to represent a range of physical properties relevant to locomotion, such as texture, and frictional resistance.

(a) Foam mat



(b) Wooden cardboard



(c) Carpet

Figure 4: The three different terrains used for training.

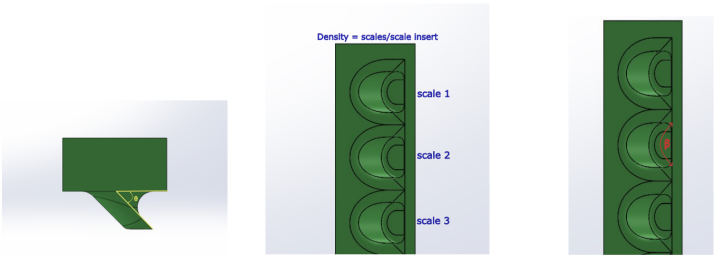


Figure 5: Snake scale geometric parameters that are able to be varied between designs.

The foam mat provides a compliant and smooth surface that introduces damping during contact, see Figure 4a. It simulates terrain similar to soft outdoor flooring. The wooden cardboard sheet, made of stiff compressed fiberboard, offers a flat and smooth surface with consistent low friction, see Figure 4b. The carpet, with woven synthetic fibers adds high surface roughness and resistance, challenging the robot's ability to generate forward thrust and maintain lateral stability, see Figure 6b.

6.2 Morphological Designs

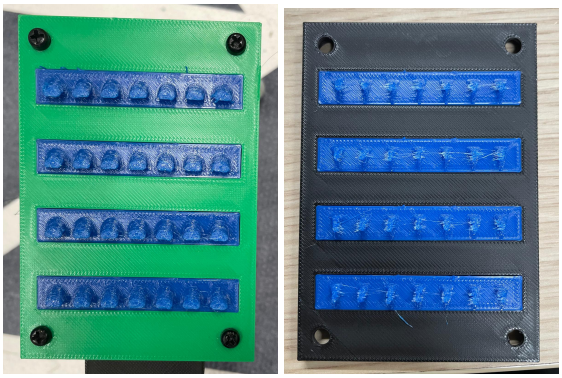
Each scale design provides capability to adjust the angle of attack of the scales, the scale density or the number of scales per insert, and the geometry of the scales through a beta β angle value. These parameters are outlined in Figure 5.

We tested four configurations. For each design trial every robot link was installed with the same scale design thus the level of frictional anisotropy in this experimentation did not vary along the snake body. See Table 1 for designs used during experiments.

Table 1: Design Specifications

	Filament Type	Beta Angle	Design Bottom View
No Scales	-	-	baseline
Design 1	PLA	180	
Design 2	TPU	180	
Design 3	TPU	30	

Each insert includes seven spikes and is mounted in arrays of four per segment, see Figure 6.



(a) Design 2 printed with TPU.

(b) Design 3 printed with TPU.

Figure 6: Example of the printed scale designs used on the robot snake.

All scales are mounted using a modular plug-in system allowing rapid configuration changes. The TPU-based scales provide greater compliance and a bit of friction, whereas PLA scales are rigid and produce higher contact forces. The combination of material stiffness, and orientation governs the directional friction generated during the learned gait.

6.3 Training Protocol

Each design-environment pair is trained using 30 episodes of 175 steps. Each episode spans approximately 1 minute and 20 seconds, and a 30-second cool-down period to prevent performance drift due to temperature fluctuations, resulting in one hour of real-time interaction per trial. The torque velocity of each servo motor is limited to 120 (out of 1023) to prevent overheating and mechanical stress during prolonged runs.

Each episode the robot is placed at the same starting point. The robot's position and orientation are recorded via OptiTrack and used to compute displacement-based rewards.

7 Experimental Evaluation

7.1 Performance Across Designs

Figure 8,7 and 9 show the x-axis displacement over time across all terrains. On the foam mat, Design 2 shows highest performance and clearest upward learning trend, especially after episode 20. This suggests that that flexible filament interacts well with the compliant surface, and allows for effective propulsion. Design 3 also performs well on the foam mats, followed by design 1, while no scales achieves minimal movement.

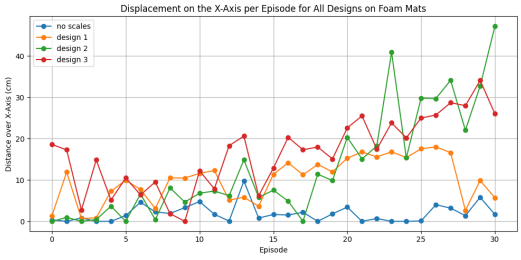


Figure 7: X-axis displacement of all designs on the foam mat.

On wooden cardboard, Design 3 consistently outperforms the other designs. The sharp drops in Design 3 is usually due to overheating of the motors. The use of TPU material and its sharp geometric design is the reason for its high performance as it allows for traction over the hard and smooth surface.

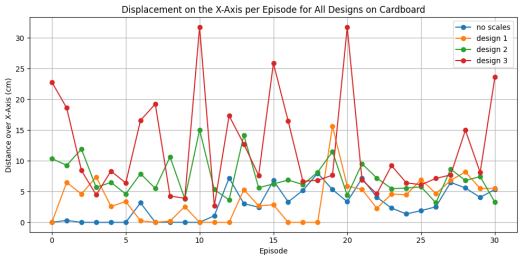


Figure 8: X-axis displacement of all designs on the cardboard.

In the carpet environment, Design 1 achieves the best performance, with consistent displacement

over the carpet. This confirms that the design is suitable for high-friction environments, due to the smooth and rigid nature of the filament. Design 2 shows moderate performance, while the robot was unable to train with Design 3, as the scales tended to catch and latch into the carpet fibers, preventing forward movement. The use of no scales performed the worst in this environment as it did not manage to go forward at any point.

7.2 Performance Evaluation

To access the locomotion potential of each design, we analyze the top 20% of episodes in term of highest forward displacement across all environments. Figure 10 presents the mean of the 6 top performing episodes for each design and terrain.

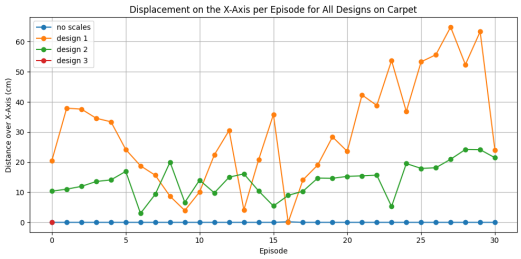


Figure 9: X-axis displacement of all designs on the carpet.

On foam mats and cardboard, Design 3 achieves the highest average displacement, closely followed by Design 2. Both designs are printed using TPU filament, which likely facilitates propulsion on these smooth low-friction surfaces. In contrast, the benchmark shows minimal displacement across all environments. The geometry of Design 3 proves ineffective on the carpet, whereas Design 2 performs slightly better on the carpet than it does on the foam mats. On the carpet, Design 1 outperforms all other designs, demonstrating its suitability for high-friction terrains, unlike Design 3, which performs best on low-friction surfaces.

7.3 Cross-Terrain Analysis

We further analyze terrain generalization by comparing how each design performs across all three environments, see Figure 11. The no scales baseline shows minimal displacement across all surfaces, reaffirming the importance of anisotropic friction. Design 1 shows the best performance on the carpet, but performance moderately on the cardboard and foam mats, indicating that the design work effectively on smooth surfaces. Design 2 shows its best performance on the foam mats with a strong upward trajectory over the episodes, but does show limiting performance on the cardboard. Design 3 stands out on the foam mats with high displacement and a learning stability, but shows instable learning on the cardboard and no performance on the carpet.

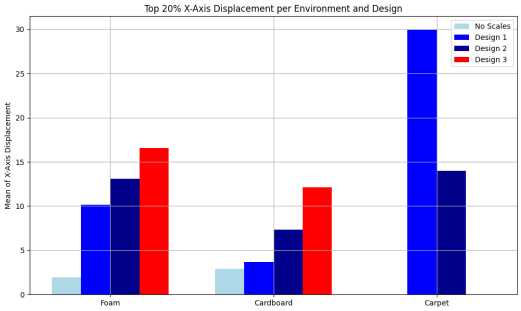
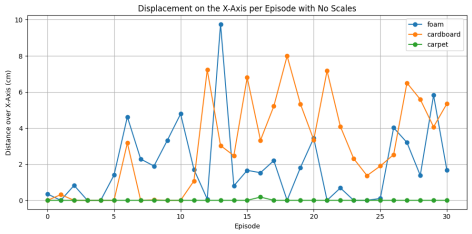


Figure 10: Mean top 20% x-axis displacement per environment and design.

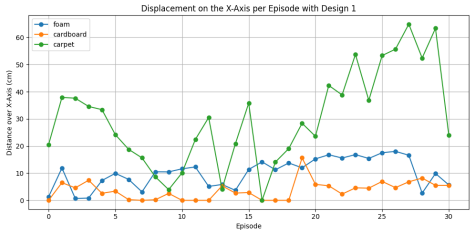
8 Discussion and Limitations

Our findings demonstrate that morphological features such as scale geometry, material compliance, and density significantly impact the performance of real-world snake robot locomotion. Specifically, we observe that certain scale designs result in distinct advantages depending on the terrain, which underscores the importance of alignment between morphology and environment. For instance, TPU-based designs outperform rigid PLA structures on soft foam surfaces due to their compliance, whereas PLA scales excel on dense carpets by providing more consistent contact forces.

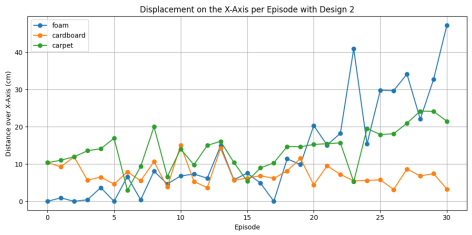
While SAC proved effective for control learning, training was limited to 30 episodes per configuration due to real-world time constraints. This may prevent full policy convergence for certain morphology-environment pairs. Another limitation is that the morphological variations were limited to three designs. Broader exploration of design space, including different scale length or asymmetrical configurations, may uncover richer trends. A last thing to add is the integration of an automated morphology search method, such as Particle Swarm Optimization (PSO), to streamline the design



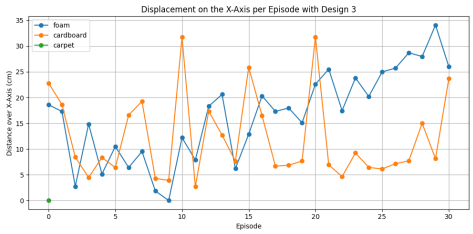
(a) No-scale benchmark x-axis displacement across all terrains.



(b) Design 1 x-axis displacement across all terrains.



(c) Design 2 x-axis displacement across all terrains.



(d) Design 3 x-axis displacement across all terrains.

Figure 11: Cross-terrain evaluation of each morphological design and the no-scale baseline.

selection process by identifying high-performing configurations, thereby reducing the reliance on manual prototyping.

9 Conclusion and Future Work

This paper presents a real-world evaluation of morphology-aware control in snake robots, using Soft Actor-Critic for locomotion policy learning and modular physical scales to study the role of directional friction. Our results show that scale configuration substantially affects terrain-specific performance and that even without formal design optimization, evaluation can guide effective morphology-policy combinations. The proposed method bridges the simulation-to-reality gap by demonstrating co-adaptation principles directly in physical systems.

In future work, we plan to extend this study by incorporating more diverse morphological parameters, such as asymmetrical scale distribution and varying segment lengths. We aim to implement a bi-level optimization loop with PSO as the outer loop to optimize the morphology, allowing the robot to autonomously adapt its structure across terrain transitions, while also improving its control policy. Finally, integrating lightweight simulation pretraining with real-world fine-tuning may further reduce sample complexity and improve generalization.

10 Acknowledgement

Kevin Sebastian Luck is supported by project number NGF.1609.241.015 of the research programme AiNed XS Europe which is financed by the Dutch Research Council (NWO). We are also grateful to the Computational Intelligence group for their assistance with the hardware experiments, and in particular to Fuda van Diggelen and Jed Muff for their valuable discussions and feedback.

References

[1] J. Shi, T. Dear, and S. D. Kelly, "Deep reinforcement learning for snake robot locomotion," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 9688–9695, 2020, 21st IFAC World Congress. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405896320333772

[2] C. Wright, A. Johnson, A. Peck, Z. McCord, A. Naaktgeboren, P. Gianfortoni, M. Gonzalez-Rivero, R. Hatton, and H. Choset, "Design of a modular snake robot," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2007, pp. 2609–2614.

[3] A. J. Ijspeert, "Amphibious and sprawling locomotion: From biology to robotics and back," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, pp. 173–193, 2020. [Online]. Available: https://www.annualreviews.org/doi/10.1146/annurev-control-091919-095731

[4] H. Marvi and D. L. Hu, "Friction enhancement in concertina locomotion of snakes," *Journal of The Royal Society Interface*, vol. 9, no. 76, pp. 3067–3080, 2012. [Online]. Available: https://royalsocietypublishing.org/doi/10.1098/rsif.2012.0132

[5] B. C. Jayne, "What defines different modes of snake locomotion?" *Integrative and Comparative Biology*, vol. 60, no. 1, pp. 156–170, 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7391877/

[6] C. Gong, M. Tesch, D. Rollinson, and H. Choset, "Snakes on an inclined plane: Learning an adaptive sidewinding motion for changing slopes," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, September 2014, pp. 1114–1119. [Online]. Available: https://biorobotics.ri.cmu.edu/papers/paperUploads/slopesidewinding.pdf

[7] J. L. Tingle, K. L. Garner, and H. C. Astley, "Functional diversity of snake locomotor behaviors: A review of the biological literature for bioinspiration," *Annals of the New York Academy of Sciences*, vol. 1533, no. 1, pp. 16–37, February 2024. [Online]. Available: https://nyaspubs.onlinelibrary.wiley.com/doi/10.1111/nyas.15109

[8] H. T. Tramsen, S. N. Gorb, H. Zhang, P. Manoonpong, Z. Dai, and L. Heepe, "Inversion of friction anisotropy in a bio-inspired asymmetrically structured surface," *Journal of The Royal Society Interface*, vol. 15, no. 138, p. 20170629, 2018. [Online]. Available: https://doi.org/10.1098/rsif.2017.0629

[9] Q. Fu and C. Li, "Robotic modelling of snake traversing large, smooth obstacles reveals stability benefits of body compliance," *Royal Society Open Science*, vol. 7, no. 2, p. 191192, 2020.

[10] M. Moattari and M. Bagherzadeh, "Flexible snake robot: Design and implementation," 04 2013, pp. 1–5.

[11] A. Gupta, S. Savarese, S. Ganguli, and L. Fei-Fei, "Embodied intelligence via learning and evolution," *Nature Communications*, vol. 12, no. 1, p. 5721, 2021.

[12] R. Riccardo Pretto, S. Cressmann, S. Chaubey, M. A. Jansen, V. Kyrki, and K. S. Luck, "Investigation into bio-inspired snake robot designs with co-adaptation of morphology and behaviour," in *40th Anniversary of the IEEE International Conference on Robotics and Automation (ICRA@40)*. Abstract, 2024.

[13] K. S. Luck, H. B. Amor, and R. Calandra, "Data-efficient co-adaptation of morphology and behaviour with deep reinforcement learning," in *Conference on Robot Learning*. PMLR, 2020, pp. 854–869.

[14] J.-K. Ryu, N. Chong, B.-J. You, and H. Christensen, "Locomotion of snake-like robots using adaptive neural oscillators," *Intelligent Service Robotics*, vol. 3, pp. 1–10, 01 2009.

[15] K. S. Luck, H. B. Amor, and R. Calandra, "Data-efficient co-adaptation of morphology and behaviour with deep reinforcement learning," in *Proceedings of the Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, L. P. Kaelbling, D. Kragic, and K. Sugiura, Eds., vol. 100. PMLR, 30 Oct–01 Nov 2020, pp. 854–869. [Online]. Available: https://proceedings.mlr.press/v100/luck20a.html

[16] E. Todorov, T. Erez, and Y. Tassa, "Mujoco: A physics engine for model-based control," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012, pp. 5026–5033.

[17] B. C. Jayne, T. J. Ward, and H. K. Voris, "Morphology, reproduction, and diet of the marine homalopsine snake bitia hydroides in peninsular malaysia," *Copeia*, pp. 800–808, 1995.