

OpenHEART: Opening Heterogeneous Articulated Objects with a Legged Manipulator

Seonghyeon Lim¹, Hyeonwoo Lee¹, Seunghyun Lee¹,
 I Made Aswin Nahrendra^{2†}, and Hyun Myung^{1*}, *Senior Member, IEEE*

Abstract—Legged manipulators offer high mobility and versatile manipulation. However, robust interaction with heterogeneous articulated objects, such as doors, drawers, and cabinets, remains challenging because of the diverse articulation types of the objects and the complex dynamics of the legged robot. Existing reinforcement learning (RL)-based approaches often rely on high-dimensional sensory inputs, leading to sample inefficiency. In this paper, we propose a robust and sample-efficient framework for opening heterogeneous articulated objects with a legged manipulator. In particular, we propose Sampling-based Abstracted Feature Extraction (SAFE), which encodes handle and panel geometry into a compact low-dimensional representation, improving cross-domain generalization. Additionally, Articulation Information Estimator (ArtIEst) is introduced to adaptively mix proprioception with exteroception to estimate opening direction and range of motion for each object. The proposed framework was deployed to manipulate various heterogeneous articulated objects in simulation and real-world robot systems. Videos can be found on the project website: <https://openheart-icra.github.io/OpenHEART/>.

I. INTRODUCTION

A legged manipulator, combining the locomotion capability of a quadruped robot with a manipulating arm, offers the versatility needed for a wide range of tasks in everyday environments [1]–[4]. The everyday tasks usually involve opening heterogeneous articulated objects, such as doors, cabinets, and drawers. The heterogeneous articulated objects can vary in appearance and joint directions. However, opening heterogeneous articulated objects with a legged manipulator remains challenging. Not only the different handle shapes and object articulation types require different manipulation strategies, but also the robot itself has complex dynamics due to its high degree of freedom (DoF) and floating base configuration.

Recently, reinforcement learning (RL)-based methods have demonstrated remarkable success in contact-rich tasks [5]–[9], including articulated object manipulation [10]–[14]. However, legged manipulators are usually limited to homogeneous door types [10], [11]. Because these methods

*Corresponding author: Hyun Myung

†Work done during the time at 1.

¹Seonghyeon Lim, Hyeonwoo Lee, Seunghyun Lee, and Hyun Myung are with the School of Electrical Engineering, KAIST (Korea Advanced Institute of Science and Technology), Daejeon, 34141, Republic of Korea. {shlim, hyeonwoolee, kevin9709, hmyung}@kaist.ac.kr

²I Made Aswin Nahrendra is with KRAFTON, Seoul, 06142, Republic of Korea. anahrendra@krafton.com

This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT). (No. RS-2024-00348461) and in part by the National Research Council of Science & Technology(NST) grant by the Korea government (MSIT) (No. GTL25041-000). The students are supported by the BK21 FOUR (Republic of Korea).

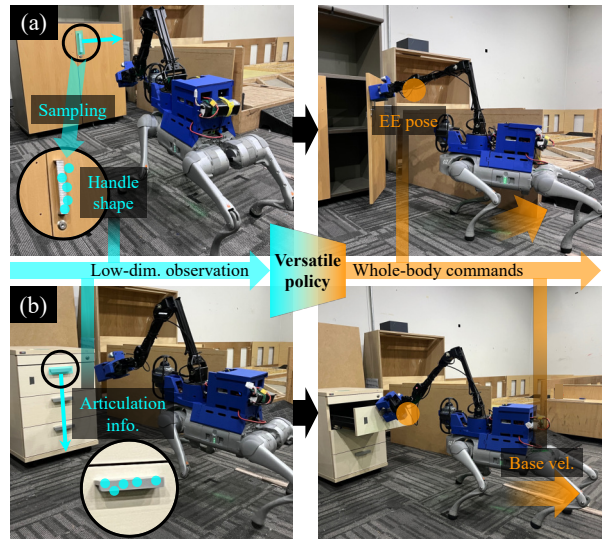


Fig. 1. We present a framework for opening heterogeneous articulated objects with a legged manipulator. The robot can open diverse articulated objects, such as (a) a revolute cabinet with a vertical handle and (b) a prismatic drawer with a horizontal handle without any object-specific models.

mainly focused on a single type of door with a similar shape, the position of the handle and doorway is sufficient to represent objects [10]. Unlike homogeneous objects, heterogeneous articulated objects can be represented with other geometric features, such as the shape of the handle and panel. These representations are essential for the robot to generate appropriate motions to manipulate the objects. To this end, high-dimensional exteroception such as point clouds [12], [13] and images [15] has often been used to represent the object. Although such high-dimensional inputs provide abundant information about the object, they are inefficient for learning contact-rich tasks with legged manipulators. The complexity of the robot’s dynamics and task could increase the number of required training samples, making sample efficiency more important. However, training an RL policy directly from high-dimensional observations is typically sample inefficient [16], [17].

In this study, we propose a framework for opening heterogeneous articulated objects using a legged manipulator. We introduce a low-dimensional representation of objects that contains the geometric features, such as handle shape and articulation information, to open heterogeneous articulated objects with a single versatile policy (Fig. 1). The articulation information includes the joint direction of the object and the distance of the handle from the joint axis. Such features are

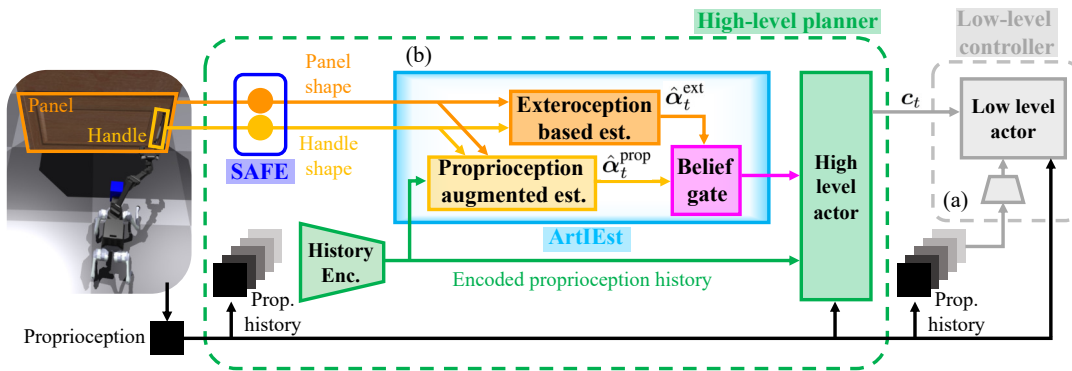


Fig. 2. The framework is hierarchically structured, with a high-level planner and (a) a low-level controller. SAFE is introduced to efficiently represent the object shape, while mitigating overfitting. (b) ArtIEst estimates the articulation information by adaptively mixing exteroception and proprioception. The history encoder extracts features from proprioception history.

required to estimate the grasping pose as well as the direction and range of opening motion. We validate that the proposed low-dimensional representation enables learning a versatile policy to open various heterogeneous articulated objects with a legged manipulator.

Our contributions can be summarized as follows:

- A hierarchical framework is proposed to open heterogeneous articulated objects with a legged manipulator. To the best of our knowledge, this is the first approach that achieves autonomous manipulation of heterogeneous articulated objects with a legged manipulator without a precise object model.
- Articulation Information Estimator (ArtIEst) is proposed to attain object features required to infer the opening motion. ArtIEst adaptively fuses proprioception with exteroception, achieving a lower estimation error than using exteroception only and an estimator that monolithically fuses both sensor modalities.
- Sampling-based Feature Extraction (SAFE) is used to abstract object shapes into low-dimensional representations and reduce overfitting through a simple post-processing step.

II. RELATED WORK

RL-based Articulated Object Manipulation. RL has been widely applied to articulated object manipulation across various robotic platforms [12]–[15], [18]–[23]. However, most of these works focus on fixed-base manipulators [12], [13], [18], [19], [23], [24] or wheeled mobile manipulators [15], which are not subject to floating-base configurations and generally have lower DoF than legged manipulators. In contrast, legged manipulators involve floating-base configurations and multi-contact dynamics, which makes learning opening tasks more challenging. These complex dynamics of robots can lead to a larger number of required training samples; therefore, sample efficiency can be crucial. Zhang *et al.* [10] and Sleiman *et al.* [11] demonstrate that legged manipulators can open a single type of door with various dimensions. Zhang *et al.* [10] represent the door using the position of the handle and doorway. Although this representation is sample efficient and sufficient for opening a homogeneous door, it is insufficient for heterogeneous

articulated objects with various shapes and articulation types. In contrast, we propose a low-dimensional representation that can efficiently encode relevant features for opening heterogeneous articulated objects.

Object Articulation Estimation. Exteroception-based methods [25]–[29] can estimate articulation information without physical interaction. However, such methods can suffer from visual ambiguity in estimation when the appearance of an object implies multiple possible articulation models, especially with symmetric objects. In contrast, proprioception-based methods [10], [30] are proposed to estimate the articulation information after grasping. However, such approaches are applicable only when the robot is in contact with the object. Additionally, they do not leverage observable geometric cues. Buchanan *et al.* [31] address these limitations by fusing both sensor modalities using a factor graph framework. However, applying such a model-based approach to a legged manipulator is challenging. The high degrees of freedom, floating-base dynamics, and contact-rich interactions complicate accurate model-based estimation. Zhang *et al.* [10] present a learning-based approach combining both inputs, enabling a legged manipulator to estimate door types, such as push/pull and left/right. However, their approach is limited to a single articulation type and monolithically fuses both sensor modalities. In comparison, our proposed method, ArtIEst, can estimate various articulation types by adaptively fusing exteroception and proprioception based on the contact state.

III. METHODOLOGY

Our framework is hierarchically structured with a low-level controller and a high-level planner. The low-level controller is responsible for generating joint target positions of the legged manipulator to track the given command, \mathbf{c}_t . The command consists of the end-effector (EE) pose, the velocity of the quadruped base, and a boolean for gripper closing. The high-level planner generates an appropriate \mathbf{c}_t to open the target articulated object. We pre-trained the low-level controller (Fig. 2 (a)) with RL, using a proprioception history encoder to estimate environmental states [7], [32]. Subsequently, the high-level planner was trained with heterogeneous articulated objects.

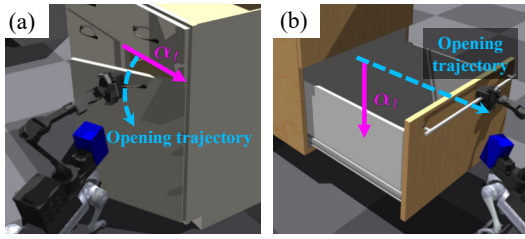


Fig. 3. Defined articulation information (pink solid arrow) for objects with (a) a revolute joint and (b) a prismatic joint. The expected opening trajectory is shown as a blue dashed arrow.

A. SAFE: Sampling-based Abstracted Feature Extraction

SAFE is designed to represent the object shape with a low-dimensional feature, while reducing the risk of overfitting to the training object sets. For the opening task, one important feature is the relative length of each side of the handle and panel. In particular, the relative length of the handle determines the grasping strategy, and that of the panel implies the opening direction. Thus, we abstracted the detailed appearance of the handle and panel into an enveloping cuboid, while preserving relative length features. During the training, such cuboids are provided from part annotations in the dataset [13], [33]. Additionally, such an enveloping cuboid can be readily obtained during deployment, using off-the-shelf object detection algorithms [34], [35].

One major challenge in transferring an RL agent to the real world is the distributional difference between training and test objects. Although numerous articulated object datasets [36], [37] can be simulated during training, the diversity of assets is substantially lower than that of real-world objects. Therefore, if an RL agent overfits to the training assets, it may not generalize well to diverse real-world objects.

The risk of overfitting can be mitigated when the distributional difference between training and test sets is reduced. We selected Kullback-Leibler (KL) divergence as the metric to measure this discrepancy.

Remark 3.1: Let P_C and Q_C be the distributions of cuboid corner points $C \in \mathbb{R}^{8 \times 3}$ corresponding to the handle shapes in P and Q , respectively. The data processing inequality [38] states that no processing can increase the distinguishability of random variables. As a result, any post-processing either decreases or preserves the KL divergence between distributions. Let $f_C : \mathbb{R}^{8 \times 3} \rightarrow \mathbb{R}^{N \times 3}$ be any post-processing function, then the following holds:

$$D_{KL}(P_C \| Q_C) \geq D_{KL}(f_C(P_C) \| f_C(Q_C)). \quad (1)$$

Equality holds if f_C preserves all the information distinguishing P_C and Q_C .

In the post-processing step, we replaced the corner points of each cuboid with points randomly sampled from its interior according to a uniform distribution. This post-processing does not preserve all the information about the distinction between P_C and Q_C . Therefore, the KL-divergence between the transformed distributions is expected to be smaller than that between the original distributions. In this way, the

distributional difference between the training and test sets can be reduced.

Additionally, the ordering of points was randomly shuffled during the sampling process. This randomness makes it difficult for the agent to learn a consistent mapping from the observation to the action. Therefore, considering that three axes can always be defined to align with the cuboid's sides, we sorted the sampled points along the lateral axis of the cuboid. In this way, the sampled points can preserve the order consistency.

To handle the motion blur in the image and the occlusion during manipulation, we captured the enveloping cuboids and sampled the points at the beginning of each episode, and used them throughout the episode. All sampled points were transformed to the robot's base frame.

B. ArtIEst: Articulation Information Estimator

We defined articulation information, α_t , which can provide appropriate motion instructions after grasping. In the case of revolute joints, α_t is defined as the perpendicular vector from the center of the handle to the joint axis of the object. Therefore, the direction of the vector implies the opening direction, and the scale indicates the range of motion. For prismatic joints, we use a downward unit vector, which has the same direction as α_t for the revolute-down joint, since their motion during the opening is similar. Fig. 3 illustrates the defined α_t for revolute and prismatic joints.

ArtIEst is proposed to estimate α_t without a precise model of the object. ArtIEst consists of an exteroception-based estimator, a proprioception-augmented estimator, and a belief gate that adaptively mixes the two estimations. The overall structure of ArtIEst is illustrated in Fig. 2 (b). All modules in ArtIEst were concurrently trained with other components of the high-level planner, and the mean squared error (MSE) between the estimated and ground truth (GT) α_t was used as the training loss.

1) *Exteroception-based Estimator:* The exteroception-based estimator was trained to estimate $\hat{\alpha}_t^{\text{ext}}$ from geometric characteristics of the object. The relative position of the handle on the panel can be an important cue for $\hat{\alpha}_t^{\text{ext}}$. For example, when a handle is placed on the left side of the panel, α_t is typically determined as a vector pointing to the right side of the panel. In contrast, when the same handle is placed on the right side of the panel, α_t should be defined as a vector pointing to the left side. Thus, we segment the panel and handle, and apply SAFE to each part independently. It differs from previous approaches [25], [26] that estimate object joint information from the holistic point cloud of the object. Subsequently, SAFE features from both parts are concatenated and fed into the estimator. Additionally, we discovered that including the robot's orientation and EE pose in the input improves the estimation accuracy, thus they are also included in the input.

2) *Proprioception-augmented Estimator:* Although the exteroception-based estimator can infer α_t before manipulation, estimation can be ambiguous when visual features suggest multiple candidates. For instance, a cabinet with

TABLE I. Reward function elements. \mathbf{p}_t^{EE} and \mathbf{p}_t^{H} denote the position of the end-effector and handle, respectively. θ_t is the angle between the dorsal-palmar axis of EE and the longer side of the handle. $\mathbb{I}\{\cdot\}$ is the indicator function, which returns one if the condition is satisfied, and zero otherwise.

Reward	Equation	Weight
Opening	ξ_t^{norm}	6.5
EE approach	$\begin{cases} \exp(-4\ \mathbf{p}_t^{\text{EE}} - \mathbf{p}_t^{\text{H}}\ ^2), & \text{if } \xi_t^{\text{norm}} < 0.2 \\ 1, & \text{otherwise} \end{cases}$	1.9
EE alignment	$\cos^2 \theta_t$	0.75
Grasping	$\mathbb{I}\{\text{grasp}\}$	1.0
Command rate	$60\lambda_t \cdot \ \mathbf{c}_t - \mathbf{c}_{t-1}\ ^2$	-0.1
EE command rate	$30\lambda_t \cdot \ \mathbf{c}_t^{\text{EE}} - \mathbf{c}_{t-1}\ ^2$	-0.1
Command smoothness	$30\lambda_t \cdot \ \mathbf{c}_t - 2\mathbf{c}_{t-1} + \mathbf{c}_{t-2}\ ^2$	-0.1
Collision	$\mathbb{I}\{\text{non-gripper collision}\}$	-1.0

a horizontally elongated handle at the upper center may appear to open left, right, or downward. To resolve such visual ambiguities, the proprioception-augmented estimator incorporates proprioceptive information with exteroceptive information during manipulation. The estimator takes the current proprioception, encoded proprioception history, and SAFE features of the handle and panel as input.

However, proprioception can provide meaningful information about α_t only when there is contact between the robot and the object. Additionally, we find that using the errors of the proprioception-augmented estimation, $\hat{\alpha}_t^{\text{prop}}$, as training signals without contact can lead to incorrect predictions. Thus, we trained the proprioception-augmented estimator only while the door was opening. Specifically, when the door was closed, the MSE loss of $\hat{\alpha}_t^{\text{prop}}$ was not used for training.

3) *Belief Gating Mechanism*: Exteroception-based estimation is appropriate when there is no interaction with the object, whereas proprioception-augmented estimation could be more accurate while the robot is opening the object. Thus, we implemented a belief gating mechanism to adaptively mix the two estimations. The belief gate predicts the linear interpolation ratio, γ_t , between the two estimations. Subsequently, the mixed estimation is calculated as $\hat{\alpha}_t^{\text{mix}} = (1 - \gamma_t)\hat{\alpha}_t^{\text{ext}} + \gamma_t\hat{\alpha}_t^{\text{prop}}$, where $0 \leq \gamma_t \leq 1$. The belief gate takes the encoded proprioception history and both estimations as input.

C. High-level Actor

The high-level actor was trained to infer appropriate commands for opening the target articulated object. A history encoder encodes the proprioception history for the actor. The encoder was trained with the β -variational autoencoder (VAE) loss [39], [40], while the corresponding decoder reconstructed the last proprioception.

1) *High-level Actor Observation*: The observation of the high-level actor includes the proprioception, encoded proprioception history, SAFE features of the handle, and $\hat{\alpha}_t^{\text{mix}}$. The proprioception consists of joint position, joint position target, previous command, base linear and angular velocity, EE pose, projected gravity vector, and finger distance of the gripper. The base linear velocity is estimated from the low-level controller using proprioception history [7]. Addition-

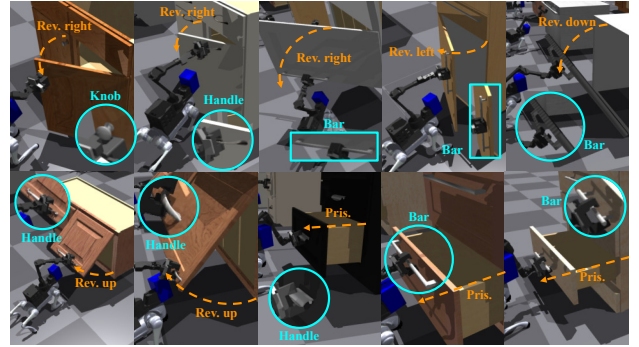


Fig. 4. Demonstration of a heterogeneous articulated object opening in the simulation. The objects differ in their handle shape, dimension, and opening direction.

ally, a vector from the EE to the handle and the heading angle of the robot are included in the observation.

2) *Reward Function*: The reward function consists of three components: opening, auxiliary, and shaping rewards.

The opening reward is designed to encourage the robot to open the target articulated object. Therefore, it is proportional to the movement of the object, ξ_t , which is defined as the opening angle for revolute joints and the displacement for prismatic ones. The angle and displacement are normalized, ξ_t^{norm} , using the corresponding joint limits.

The auxiliary reward guides the robot to grasp the handle in an appropriate pose. It consists of three components: EE approach, EE alignment, and grasping reward. The EE approach reward encourages the robot to approach its EE close to the center of the handle. To avoid conflicts with the opening reward, it is set to the maximum value after the opening begins. The EE alignment reward encourages the robot to align the dorsal-palmar axis of EE with the longer side of the handle, which is the proper grasping pose. The grasping reward is given when the gripper grasps the handle.

Finally, the shaping reward is designed to prevent inappropriate motions, such as aggressively changing commands. The command change rate and its derivative are penalized. Note that the relative amount of shaping reward becomes smaller as the training progresses. Thus, we increased the scaling factor of the shaping reward as the average success rate, λ_t , rose. The detailed equation and scales of reward functions are summarized in Table I.

IV. EXPERIMENTS

A. Experiment Setup

We trained all high-level modules in the Isaac Gym [41]. The training environment consisted of 41 heterogeneous articulated objects selected from the PartManip dataset [13]. The objects were selected to cover diverse combinations of articulation and handle types. The articulation types include revolute joints with axes opened up, down, left, or right, and prismatic joints. The handle types were heuristically classified into bar, handle, and knob depending on their shape. Examples of each articulation and handle type are illustrated in Fig. 4. The relative pose between the robot and the object was randomly sampled at the beginning of each

episode, while assuming that the robot is facing the object. The initial pose of the robot was sampled within a 2 m box, centered 2 m away from the object, with a random heading angle in $[-\pi/4, \pi/4]$. Additionally, the friction coefficient of the manipulator was set to zero, preventing the robot from scrubbing the object, as referred to Zhang *et al.* [10]. Instead, to compensate for the limited grasping force and friction of the finger, a small hook is attached to the tip of the gripper.

Training was performed in 2,048 parallel environments, with different object assets distributed across them. The high-level actor was trained using Proximal Policy Optimization (PPO) [42]. Similar to [43], an adaptive learning rate based on the KL-divergence was used for the high-level actor and its history encoder. Additionally, the learning rate was set to 0.0005 for ArtIEst modules, since the diversity of openable objects increases during training, and the proprioception-based estimation should update accordingly. All training was performed on a desktop PC with an Intel Core i9-13900 CPU (up to 5.6 GHz), 48 GB RAM, and an NVIDIA RTX 4090 GPU.

The robot platform comprises a 12-DoF Unitree Go2 [44] legged robot and a 6-DoF ViperX 300 arm [45], which is equipped with a parallel two-finger gripper. Additionally, two Livox Mid-360 LiDARs and an XSens MTi-300 IMU are mounted on the robot for perception. The high-level planner, odometry, and perception algorithm were run on an ASUS NUC 13 Pro while the low-level controller was run on a separate NVIDIA Orin NX. The high-level planner and low-level controller were run at 50 Hz. We accumulated 25 steps of proprioception at 0.02s intervals for the history encoder.

The corner positions of the handle and panel cuboids were provided in the fixed world frame at the beginning of each experiment and not updated during manipulation. Instead, their relative positions in the robot base frame were updated at 10 Hz through LiDAR odometry and provided to the high-level planner.

B. Compared Methods

We compare our method with the following baselines:

- **Center-based teacher [10]:** This baseline was adapted from the method proposed by Zhang *et al.* [10]. The door was represented using only the center position of the handle, whereas our method incorporates the handle and panel shape. The position of the doorway is excluded, since our task does not require the robot to traverse it. Privileged information, such as GT α_t and joint states, was included in the observations. Furthermore, due to differences in task settings and robot hardware, this baseline was trained with our reward function and low-level controller.
- **Point cloud-based policy:** This baseline represents the object using a high-dimensional point cloud. We sampled 1,024 points on the frontal surface of the object. A custom lightweight version of PointNet [46] is used as a feature extractor. We reduced the number of hidden layers of the PointNet to 2, and set the dimensions to

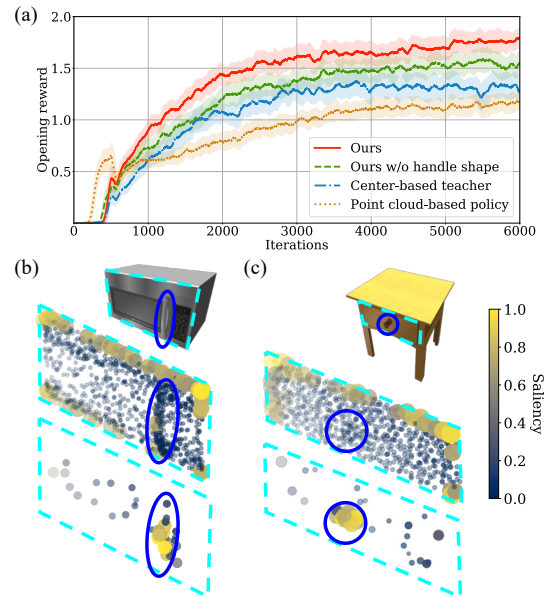


Fig. 5. (a) Learning curve of opening reward. **Ours** shows the highest performance compared with the baselines. Saliency maps are shown for the object with (b) revolute joint and (c) prismatic joint. Object appearances are depicted at the top line, the saliency maps from **Point cloud-based policy**, and **Ours** are shown in the middle and below each subfigure.

64 and 128 due to the large memory usage, but it still uses about three times larger GPU memory than **Ours**.

Additionally, to investigate the effectiveness of each component in our method, we conduct an ablation study with the following variants:

- **Ours w/o handle shape:** This variant excludes the handle shape feature, while still providing the center of the handle to the high-level policy.
- **Ours w/o sampling:** Sampling process is ablated in this variant, thus the corner points of the enveloping cuboid are used to represent the handle and panel shape.
- **Ours w/ mono. est.:** A monolithic estimator is employed for calculating α_t instead of separating modules as in ArtIEst. This monolithic estimator is similar to that of Zhang *et al.* [10], which jointly takes both proprioception and exteroception as inputs and estimates α_t .
- **Ours w/o prop. est.:** Without the proprioception-augmented estimator, only the exteroception-based estimator is used, similar to the existing vision-based approaches [25]–[27].

All the methods above were trained with the same reward functions and training environment described in Sec. III.C.2 and Sec. IV.A, respectively. For a fair comparison, exponential linear units (ELUs) [47] were used as the activation function in all networks.

C. Efficiency of the Proposed Representation

Fig. 5 (a) shows that **Ours** achieves a higher opening reward compared with the baselines, highlighting the efficiency of the proposed representation. Although privileged information, including α_t , is provided for the **Center-based teacher**, representing an object with handle centers appears

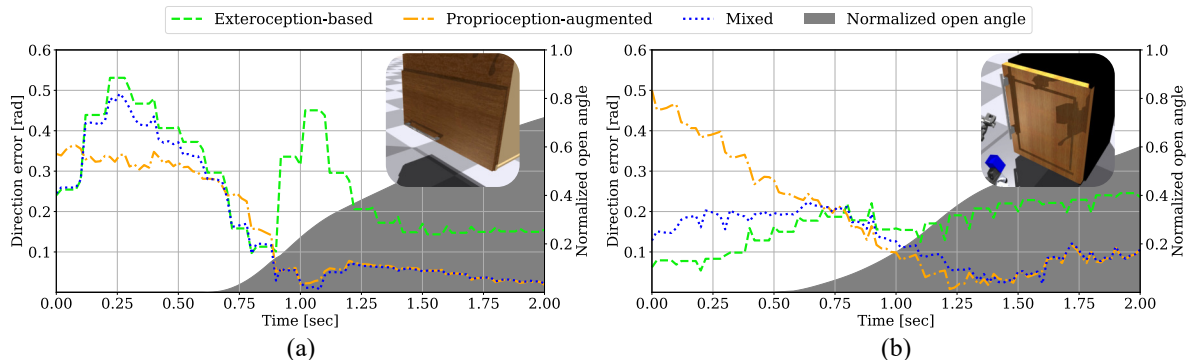


Fig. 6. Estimation error transition of exteroception-based, proprioception-augmented, and mixed estimations during the manipulation. The progress of the opening motion is indicated by the object’s open angle normalized using the joint limits. (a) The appearance of the object can be inferred to be opened in either a rightward or upward direction, while the correct direction is upward, leading to visual ambiguity in exteroception-based estimation. (b) In contrast, no visual ambiguity exists in this case.

insufficient for heterogeneous articulated objects, resulting in lower rewards. Another difference between **Ours** and **Center-based teacher** is the inclusion of encoded proprioception history, which may help the policy to capture contact-rich dynamics between the robot and the object. The comparison with **Ours w/o handle shape** further clarifies the role of the encoded proprioception history. Since **Ours w/o handle shape** retains proprioception history but excludes handle-shape features, its efficiency falls between **Ours** and **Center-based teacher**. This suggests that both handle-shape representation and proprioception history play important roles in effective learning.

Figs. 5 (b) and (c) show the saliency map of the actions of the **Point cloud-based policy** and **Ours** from the objects with (b) revolute joint and (c) prismatic joint at the top of each subfigure. The saliency is obtained by computing the gradient of the action with respect to the input point cloud for **Point cloud-based policy** (middle) and sampled points for **Ours** (below). Higher saliency indicates higher attention of the policy to a specific region in the input when making decisions. Subsequently, the saliency map is obtained by projecting the accumulated saliency over 100 steps (2 s) onto the corresponding points. Higher saliency values are depicted with larger points and brighter colors. **Ours** consistently focuses on the handle shape for both objects. However, the **Point cloud-based policy** typically focuses on the edges of the objects rather than the handle shape. This results in slower convergence and lower performance of the **Point cloud-based policy** than that of **Ours** as shown in Fig. 5 (a). These results indicate that directly using high-dimensional observations is less efficient for learning contact-rich tasks with legged manipulators.

D. Heterogeneous Articulated Object Manipulation

We demonstrated heterogeneous articulated object manipulation with a legged manipulator in the simulation environment. As shown in Fig. 4, our single versatile policy successfully opens a variety of articulated objects. The objects differ in handle shape, dimension, and articulation type. Therefore, the robot needs to adapt its manipulation strategy to each object. In particular, the robot has to align its gripper to grasp the handle firmly and determine the opening direction

TABLE II. ArtIEst evaluation. The estimation errors are separately measured before and during the contact.

Method	Estimation direction error (rad) ↓		
	Before contact	During contact	Entire episode
Ours	0.2293	0.0687	0.1701
Ours w/o prop. est.	0.2623	0.2228	0.2482
Ours w/ mono. est.	0.2574	0.0606	0.1827

and range of motion for opening the object. The proposed representation, encoding handle shape and opening direction, enables heterogeneous articulated object manipulation.

E. Belief Gating Mechanism of ArtIEst

Fig. 6 shows the error transition of $\hat{\alpha}_t^{\text{ext}}$, $\hat{\alpha}_t^{\text{prop}}$, and $\hat{\alpha}_t^{\text{mix}}$ during manipulation, (a) with visual ambiguity in the articulation information and (b) without ambiguity. The errors are measured as the angle difference between the estimated and GT articulation information. When there is visual ambiguity, multiple candidates for α_t are inferred, depicted as the oscillating error of $\hat{\alpha}_t^{\text{ext}}$ in Fig. 6 (a). However, errors of $\hat{\alpha}_t^{\text{prop}}$ decrease as the manipulation proceeds, since proprioception during the contact can help resolve the visual ambiguity. Therefore, the belief gate shifts the belief from $\hat{\alpha}_t^{\text{ext}}$ to $\hat{\alpha}_t^{\text{prop}}$, leading to a decrease in the error of $\hat{\alpha}_t^{\text{mix}}$. In contrast, when there is no ambiguity in the exteroception-based estimation, the error of $\hat{\alpha}_t^{\text{ext}}$ is low and stable, as shown in Fig. 6 (b). However, $\hat{\alpha}_t^{\text{prop}}$ shows lower error from the beginning of the contact. Therefore, the belief transitions from $\hat{\alpha}_t^{\text{ext}}$ to $\hat{\alpha}_t^{\text{prop}}$.

F. Modular Architecture of ArtIEst

We evaluated the contribution of each component within ArtIEst. The estimation errors are measured across test object sets and averaged over five episodes. The test set was randomly selected from the dataset [13], excluding the objects used for training. To investigate the role of excluding proprioception before contact, the estimation error was separately measured before and during the contact.

As shown in Table II, **Ours** achieves the lowest estimation error for the entire episode among the methods. Compared with **Ours w/o prop. est.**, the consistently reduced error before and during contact demonstrates the benefit of incorporating proprioceptive information. Additionally, **Ours**

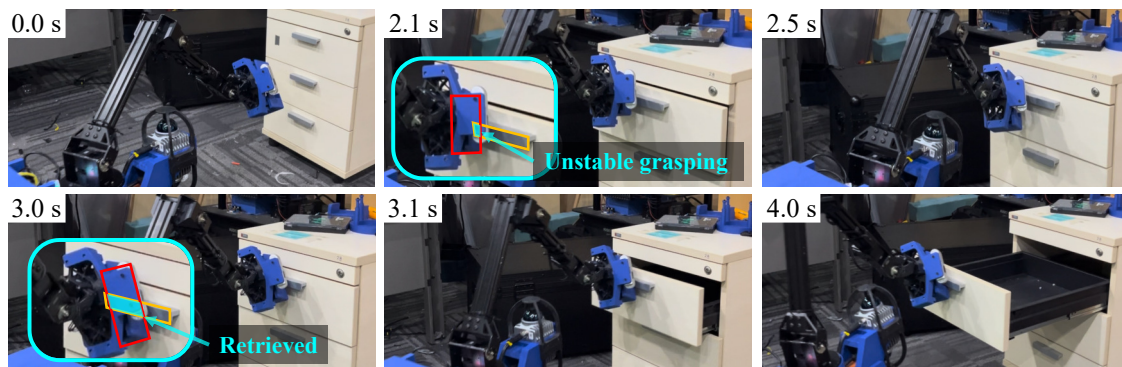


Fig. 7. Real-world demonstration of opening a drawer with a prismatic joint and a horizontally elongated handle. Although the first grasp pose (2.1 s) was not stable, the robot autonomously retrieved the handle (3.0 s) and successfully opened the drawer.

TABLE III. Success rate across training and test sets. The proposed method shows the highest performance retention.

Method	Success rate (%) \uparrow		Test/Train (%) \uparrow
	Train set	Test set	
Ours	79.35	79.02	99.35
Ours w/o sampling	77.56	70.23	92.92
Center-based teacher [10]	62.43	50.60	81.05
Point cloud-based policy	56.58	41.39	73.15

w/ **mono. est.** shows higher error than **Ours** when there is no contact. This indicates that excluding proprioceptive inputs before contact helps improve the estimation accuracy. Although **Ours w/ mono. est.** shows slightly lower error during contact, the difference is smaller than that of before contact, and the overall accuracy is still lower than **Ours**.

G. Cross-domain Generalization of SAFE

To evaluate the enhanced cross-domain generalization resulting from the sampling process of SAFE method, we compared the success rate over the training and test sets. An episode was considered successful if the robot opened the object at least 33% of its maximum range of motion, which is typically a stricter criterion than that used in dataset [13] (20% for drawers and 30° for doors, while doors usually can be opened more than 90°). The success rate was measured over ten episodes for each object, with a maximum length of 15 s per episode.

Table III shows that **Ours** achieves the highest success rate on both the training and test sets. Additionally, the proposed method demonstrates robust cross-domain generalization, as evidenced by a 99.35% test/train ratio. Compared with **Ours w/o sampling**, the sampling process in SAFE improves the test/train ratio by 6.43%, confirming that the sampling process contributes to enhanced generalization. **Point cloud-based policy** shows the lowest test/train ratio, indicating its limited ability to generalize. This is because directly using high-dimensional visual input without sufficient abstraction can make the policy overfit to the specific appearance details of the objects, as depicted in Figs. 5 (b) and (c).

H. Real-world Demonstration

We deployed the proposed framework on a real-world legged manipulator. The robot successfully opens a cabinet with a revolute joint and a vertical handle (Fig. 1 (a)) as well

as a prismatic drawer with a horizontal handle (Fig. 1 (b)). Although these real-world objects were not included in the training dataset, the robot successfully opened them, demonstrating the generalization capability of the proposed framework.

The sequence of opening motions for the drawer is demonstrated in Fig. 7. Although the first grasping (2.1 s) was not stable due to the misaligned gripper, the robot autonomously regrasped the handle (3.0 s) and successfully opened the drawer. Notably, such auto-retrying behaviors are important for the robot to adapt to unexpected changes in the environment and successfully complete the task. However, such behaviors can be difficult to achieve with model-based approaches, as they require precise modeling of contact uncertainties and adaptive replanning, which is often impractical in real-world settings.

V. CONCLUSION

In this study, we presented a framework for opening heterogeneous articulated objects with a legged manipulator. We proposed a low-dimensional representation to extract features from the objects, enabling sample-efficient training. The proposed framework demonstrated the ability to open diverse objects with a single policy. The estimation error for articulation information is decreased with ArtIEst architecture, and SAFE shows improved cross-domain generalization. Currently, the framework was demonstrated with a fixed object pose measurement given at the beginning of the episode. For future work, we seek to integrate the framework with onboard object pose estimation.

REFERENCES

- [1] X. Liu, B. Ma, C. Qi, Y. Ding, G. Zhang, P. Chen, K. Liu, Z. Jia, C. Guan, Y. Mo, *et al.*, “MLM: Learning multi-task loco-manipulation whole-body control for quadruped robot with arm,” *arXiv preprint arXiv:2508.10538*, 2025.
- [2] H. Ha, Y. Gao, Z. Fu, J. Tan, and S. Song, “UMI on legs: Making manipulation policies mobile with manipulation-centric whole-body controllers,” in *Proc. PMLR Conf. Robot Learn. (CoRL)*, 2025, pp. 5254–5270.
- [3] J. Bruedigam, A. A. Abbas, M. Sorokin, K. Fang, B. Hung, M. Guru, S. G. Sosnowski, J. Wang, S. Hirche, and S. Le Cleac’h, “Jacta: A versatile planner for learning dexterous and whole-body manipulation,” in *Proc. PMLR Conf. Robot Learn. (CoRL)*, 2025, pp. 994–1020.
- [4] M. Liu, Z. Chen, X. Cheng, Y. Ji, R.-Z. Qiu, R. Yang, and X. Wang, “Visual whole-body control for legged loco-manipulation,” in *Proc. PMLR Conf. Robot Learn. (CoRL)*, 2025, pp. 234–257.

- [5] X. Cheng, K. Shi, A. Agarwal, and D. Pathak, "Extreme parkour with legged robots," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, 2024, pp. 11 443–11 450.
- [6] J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, "Learning quadrupedal locomotion over challenging terrain," *Sci. Robot.*, vol. 5, no. 47, p. eabc5986, 2020.
- [7] I. M. A. Nahrendra, B. Yu, and H. Myung, "DreamWaQ: Learning robust quadrupedal locomotion with implicit terrain imagination via deep reinforcement learning," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, 2023, pp. 5078–5084.
- [8] Y. Ma, F. Farshidian, and M. Hutter, "Learning arm-assisted fall damage reduction and recovery for legged mobile manipulators," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, 2023, pp. 12 149–12 155.
- [9] S. Lim, H. Lee, S. Lee, and H. Myung, "Learning robust whole-body control for legged manipulation via learned terrain estimation," in *Proc. Int. Conf. Control Automat. Syst. (ICCAS)*, 2024, pp. 1510–1515.
- [10] M. Zhang, Y. Ma, T. Miki, and M. Hutter, "Learning to open and traverse doors with a legged manipulator," in *Proc. PMLR Conf. Robot Learn. (CoRL)*, 2025, pp. 2913–2927.
- [11] J. P. Sleiman, M. Mittal, and M. Hutter, "Guided reinforcement learning for robust multi-contact loco-manipulation," in *Proc. PMLR Conf. Robot Learn. (CoRL)*, 2025, pp. 531–546.
- [12] Y. Li, X. Zhang, R. Wu, Z. Zhang, Y. Geng, H. Dong, and Z. He, "UniDoorManip: Learning universal door manipulation policy over large-scale and diverse door manipulation environments," *arXiv preprint arXiv:2403.02604*, 2024.
- [13] H. Geng, Z. Li, Y. Geng, J. Chen, H. Dong, and H. Wang, "PartManip: Learning cross-category generalizable part manipulation policy from point cloud observations," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2023, pp. 2978–2988.
- [14] Y. Urakami, A. Hodgkinson, C. Carlin, R. Leu, L. Rigazio, and P. Abbeel, "DoorGym: A scalable door opening environment and baseline agent," *arXiv preprint arXiv:1908.01887*, 2019.
- [15] H. Xiong, R. Mendonca, K. Shaw, and D. Pathak, "Adaptive mobile manipulation for articulated objects in the open world," *arXiv preprint arXiv:2401.14403*, 2024.
- [16] A. Banino, A. P. Badia, J. Walker, T. Scholtes, J. Mitrovic, and C. Blundell, "Coberl: Contrastive bert for reinforcement learning," in *Proc. Int. Conf. on Mach. Learn. (ICML) Workshop on Unsupervised Reinforcement Learn.*, 2021.
- [17] B. M. Lake, T. D. Ullman, J. B. Tenenbaum, and S. J. Gershman, "Building machines that learn and think like people," *Behavioral and Brain Sci.*, vol. 40, p. e253, 2017.
- [18] S. Gu, E. Holly, T. Lillicrap, and S. Levine, "Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, 2017, pp. 3389–3396.
- [19] M. Sharma, J. Liang, J. Zhao, A. Lagrassa, and O. Kroemer, "Learning to compose hierarchical object-centric controllers for robotic manipulation," in *Proc. PMLR Conf. Robot Learn. (CoRL)*, 2021, pp. 822–844.
- [20] A. Rajeswaran, V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, and S. Levine, "Learning complex dexterous manipulation with deep reinforcement learning and demonstrations," *Robot. Sci. Syst. (RSS)*, 2018.
- [21] C. Schwarke, V. Klemm, M. Van der Boon, M. Bjelonic, and M. Hutter, "Curiosity-driven learning of joint locomotion and manipulation tasks," in *Proc. PMLR Conf. Robot Learn. (CoRL)*, vol. 229, 2023, pp. 2594–2610.
- [22] S. Luo, Q. Peng, J. Lv, K. Hong, K. R. Driggs-Campbell, C. Lu, and Y.-L. Li, "Human-agent joint learning for efficient robot manipulation skill acquisition," in *Proc. PMLR Conf. Robot Learn. (CoRL) Workshop on Learn. Robot Fine and Dexterous Manipulation: Perception and Control*, 2024.
- [23] X. Ma, S. Patidar, I. Haughton, and S. James, "Hierarchical diffusion policy for kinematics-aware multi-task robotic manipulation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2024, pp. 18 081–18 090.
- [24] M. Dalal, M. Liu, W. Talbott, C. Chen, D. Pathak, J. Zhang, and R. Salakhutdinov, "Local policies enable zero-shot long-horizon manipulation," in *Proc. PMLR Conf. Robot Learn. (CoRL) Workshop on Learn. Effective Abstractions for Planning*, 2024.
- [25] J. Wang, W. Liu, Q. Yu, Y. You, L. Liu, W. Wang, and C. Lu, "RPMart: Towards robust perception and manipulation for articulated objects," in *Proc. IEEE/RJSJ Int. Conf. Intell. Robot. Syst. (IROS)*, 2024, pp. 7270–7277.
- [26] Q. Yu, J. Wang, W. Liu, C. Hao, L. Liu, L. Shao, W. Wang, and C. Lu, "GAMMA: Generalizable articulation modeling and manipulation for articulated objects," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, 2024, pp. 5419–5426.
- [27] V. Zeng, T. E. Lee, J. Liang, and O. Kroemer, "Visual identification of articulated object parts," in *Proc. IEEE/RJSJ Int. Conf. Intell. Robot. Syst. (IROS)*, 2021, pp. 2443–2450.
- [28] B. Eisner, H. Zhang, and D. Held, "FlowBot3D: Learning 3D articulation flow to manipulate articulated objects," *Robot. Sci. Syst. (RSS)*, 2022.
- [29] X. Li, H. Wang, L. Yi, L. J. Guibas, A. L. Abbott, and S. Song, "Category-level articulated object pose estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2020, pp. 3706–3715.
- [30] T. Lips and F. Wyffels, "Revisiting proprioceptive sensing for articulated object manipulation," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA) Embracing Contacts Workshop*, 2023.
- [31] R. Buchanan, A. Röfer, J. Moura, A. Valada, and S. Vijayakumar, "Online estimation of articulated objects with factor graphs using vision and proprioceptive sensing," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, 2024, pp. 16 111–16 117.
- [32] Z. Fu, X. Cheng, and D. Pathak, "Deep whole-body control: learning a unified policy for manipulation and locomotion," in *Proc. PMLR Conf. Robot Learn. (CoRL)*, 2023, pp. 138–149.
- [33] H. Geng, H. Xu, C. Zhao, C. Xu, L. Yi, S. Huang, and H. Wang, "GaPartNet: Cross-category domain-generalizable object perception and manipulation via generalizable and actionable parts," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2023, pp. 7081–7091.
- [34] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, real-time object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 779–788.
- [35] A. Wang, L. Liu, H. Chen, Z. Lin, J. Han, and G. Ding, "YOLOE: Real-time seeing anything," *arXiv preprint arXiv:2503.07465*, 2025.
- [36] T. Mu, Z. Ling, F. Xiang, D. Yang, X. Li, S. Tao, Z. Huang, Z. Jia, and H. Su, "ManiSkill: Generalizable manipulation skill benchmark with large-scale demonstrations," in *Adv. Neural Inf. Process. Syst. (NeurIPS), Track on Datasets and Benchmarks*, 2021.
- [37] L. Liu, W. Xu, H. Fu, S. Qian, Q. Yu, Y. Han, and C. Lu, "AKB-48: A real-world articulated object knowledge base," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2022, pp. 14 809–14 818.
- [38] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ: Wiley, 1991.
- [39] D. P. Kingma, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.
- [40] I. Higgins, L. Matthey, A. Pal, C. Burgess, X. Glorot, M. Botvinick, S. Mohamed, and A. Lerchner, " β -VAE: Learning basic visual concepts with a constrained variational framework," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2017.
- [41] V. Makovychuk, L. Wawrzyniak, Y. Guo, M. Lu, K. Storey, M. Macklin, D. Hoeller, N. Rudin, A. Allshire, A. Handa, et al., "Isaac Gym: High performance GPU-based physics simulation for robot learning," in *Adv. Neural Inf. Process. Syst. (NeurIPS), Track on Datasets and Benchmarks*, 2021.
- [42] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.
- [43] N. Rudin, D. Hoeller, P. Reist, and M. Hutter, "Learning to walk in minutes using massively parallel deep reinforcement learning," in *Proc. PMLR Conf. Robot Learn. (CoRL)*, 2022, pp. 91–100.
- [44] Unitree Robotics, "Unitree Go2," Accessed: Sep. 11, 2025. [Online]. Available: <https://shop.unitree.com/products/unitree-go2>
- [45] Trossen Robotics, "ViperX 300S," Accessed: Sep. 11, 2025. [Online]. Available: <https://www.trossenrobotics.com/viperx-300>
- [46] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 652–660.
- [47] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units ELUs," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015, pp. 2921–2929.