

HOTICE: Whole-Body Humanoid Object Transportation in Cluttered Environments

Anonymous Authors



Fig. 1: We introduce **HOTICE**, a humanoid loco-manipulation method that enables humanoid robots to transport objects through cluttered environments populated by diverse types of obstacles, including side obstacles – (1) and (2), overhead obstacles – (3), and ground obstacles – (4). More results can be found at our anonymous website <https://hotice2027.github.io>.

Abstract—Object transportation is a fundamental capability for humanoid robots operating in real-world, human-centric environments, yet existing methods struggle when clutter constrains free space around both the robot and its carried payload. We present **HOTICE**, a whole-body humanoid learning framework for transporting objects through such cluttered environments. First, we introduce Humanoid-Object Decoupled Potential Fields, which jointly encode collision-avoidance guidance for the robot and the carried object, enabling coordinated, obstacle-aware motion for both. Second, to address the large action space inherent to whole-body loco-manipulation, we design a dual-agent reinforcement learning architecture that decouples upper- and lower-body control while preserving whole-body coordination via shared state observations and rewards. To train a policy that generalizes across diverse cluttered scenes, we further employ a specialist-to-generalist distillation strategy, in which privileged teacher policies are distilled into a single deployable student policy. We evaluate **HOTICE** in MuJoCo simulation and on a real Unitree G1 humanoid, demonstrating effective and robust object transportation across cluttered scenarios for objects of varying shapes. Our results show that **HOTICE** reliably coordinates whole-body motion and object-aware collision avoidance, generalizing effectively to previously unseen cluttered environments while achieving strong performance in sim2real deployment.

I. INTRODUCTION

Object transportation is one of the most important problems in robotics, underpinning a wide range of applications including last-mile delivery, warehouse logistics, and household assistance. Traditionally, this problem has been mainly addressed with robotic manipulators mounted on fixed bases [1], wheeled mobile platforms [2], or aerial drones [3], which achieve high reliability in structured, engineered settings. However, these platforms rely on assumptions that break down in open-world, cluttered environments: fixed-base arms are constrained by a limited workspace and cannot reposition to navigate around obstacles, wheeled bases require traversable, relatively flat ground and struggle with thresholds and narrow or irregular passages common

in human-centric spaces, and drones are limited by small payload capacities and the need for open airspace.

In contrast, the humanoid form factor is well suited to these challenges [4]. Whole-body loco-manipulation potentially allows humanoids to step over ground-level obstacles and adjust their posture to duck under overhead obstructions while maintaining a stable hold on the carried payload, and their human-scale footprint enables passage through narrow side clearances and dense clutter built for people. Moreover, a single humanoid embodiment can generalize across diverse environments, mitigating the need for specialized, environment-specific robot platforms.

Despite these advantages, humanoid object transportation in cluttered environments remains a challenging and under-investigated problem. Humanoid loco-manipulation requires coordinated control over both the upper and lower body, resulting in a large, high-dimensional action space that is further complicated in cluttered scenarios, where the robot must often adopt extreme, non-default poses to simultaneously maintain balance and navigate through obstacles [4]. This challenge is compounded when the robot must also maintain a secure hold on the transported object, particularly when it is heavy or large, as the added dynamics and reduced clearance make safe traversal substantially more difficult.

In this work, we propose **Humanoid Object Transportation In Cluttered Environments (HOTICE)**, a whole-body humanoid learning framework that enables humanoid robots to transport diverse objects through complex, cluttered environments. First, we introduce the Humanoid-Object Decoupled Potential Fields (**HOD-PF**), which encode the geometric relations between the humanoid, the transported object, and surrounding obstacles. HOD-PF enables collision-free transportation by providing guidance signals for jointly avoiding humanoid-obstacle and object-obstacle collisions, while maintaining tight synchronization between the robot and the object throughout transport. Moreover, to address the large

action space inherent to whole-body control, we design a dual-agent learning architecture that decouples the upper- and lower-body action spaces, substantially improving the robot’s loco-manipulation capability in complex environments. Finally, we apply a scalable specialist-to-generalist distillation scheme that consolidates many scene-specific privileged teacher policies into a single generalist student policy capable of handling a wide range of obstacle configurations. Together, our system enables a Unitree G1 humanoid robot to effectively transport objects of different shapes in diverse cluttered scenarios both in MuJoCo simulation [5] and real-world deployment.

Our contributions can be summarized as follows:

- We propose HOD-PF, a framework that effectively encodes the relationships between the humanoid robot, object, and obstacles for collision-free object transportation in cluttered environments.
- We present a dual-agent learning system to deal with the large action space difficulty of whole-body loco-manipulation under challenging conditions.
- By applying specialist-to-generalist distillation from various obstacle settings, our policy enables successful object transportation across various scenarios in both simulation and real-world deployment.

II. RELATED WORK

A. Humanoid Traversal in Cluttered Environments

Traversing spatially constrained environments has been studied extensively in legged robotics, from model-based whole-body planning [6], [7], [8] to learned quadrupedal policies that crouch under overhangs, squeeze through gaps, and negotiate irregular voids [9], [10], [11], [12]. Extending these capabilities to humanoids is considered harder given their higher dimensionality and narrower stability margins, yet recent work has demonstrated vision-based parkour [13], [14], contact-driven whole-body control [15], [16], balance on narrow terrain [17], [18], and terrain-conditioned skill composition [19], [20]. More recently, [21] proposed HumanoidPF that encodes humanoid-obstacle relationships as collision-free motion directions, enabling object-free traversal across ground, lateral, and overhead obstacles. TANGO [22] targeted similar scenarios but for long-horizon text-driven navigation. These methods, however, assume an unencumbered robot whose collision geometry follows entirely from its own kinematics, an assumption that breaks under object transportation, where the payload extends the collision volume and perturbs balance precisely when extreme postures are required. Our HOD-PF targets this gap by decoupling the humanoid’s and object’s collision avoidance into separate guidance vector fields that jointly shape a collision-free carrying motion in cluttered environments.

B. Humanoid Object Transportation

Humanoid object transportation was first studied as a planning problem, in which collision-free trajectories are computed jointly for the robot and the carried object [23],

[24]. Such methods can be highly reliable, but typically rely on simplified dynamics assumptions. Learning-based approaches relax these assumptions and generalize more readily across objects and scene configurations, from sim2real box transport under varying weight and geometry [25], [26] to contact-based whole-body strategies that embrace bulky objects [27], [28] and load-aware controllers that stabilize locomotion against payload-induced disturbances [29], [30]. More recent systems incorporate exteroceptive perception, enabling humanoids to carry objects over uneven terrain [31], [32], execute long-horizon transport [33], [26], and perform visual loco-manipulation [34], [35], [36]. Across both paradigms, however, existing work targets open or lightly obstructed scenes, where clearance can be recovered by walking around sparse, low-lying obstacles. Real-world clutter instead constrains the robot from the ground, sides, and above simultaneously, so both the robot and its carried object must fit through what free space remains. Our **HOTICE** targets this regime, enabling effective humanoid object transportation through complex, highly constrained scenes.

III. PROBLEM STATEMENT

We investigate a humanoid loco-manipulation object transportation setting. Given a cluttered environment populated by obstacles of different types (side, ground, and overhead obstacles), the robot starts at a starting position $\mathbf{s} \in \mathbb{R}^3$ holding an object $\mathcal{O}$ and is required to transport $\mathcal{O}$ to a target position $\mathbf{g} \in \mathbb{R}^3$ without colliding with the obstacles $\mathcal{C}$. The object $\mathcal{O}$ can be of different shapes. In this paper, we mainly refer to $\mathcal{O}$ as boxes of cuboid shape, but our method, as demonstrated later in Section V, is highly generalizable to objects of other shapes. Note that since we mainly care about the transportation of the object and collision avoidance during the traversal, we assume the object $\mathcal{O}$ is already held by the robot at the beginning of the task in our method description and experiments. However, we do train a policy to pick up the object from a support surface to facilitate the complete pipeline of object picking and transporting to the goal position. The task is considered successful if the robot reaches the goal $\mathbf{g}$ while still holding $\mathcal{O}$, without either the robot or $\mathcal{O}$ colliding with obstacles along the way; any robot fall, drop of the object, or collision constitutes a failure.

IV. METHOD

We present **Humanoid Object Transportation In Cluttered Environments (HOTICE)**, an RL whole-body humanoid learning framework for transporting objects while avoiding collisions with diverse obstacles. **HOTICE** consists of three main components: Humanoid-Object Decoupled Potential Fields (HOD-PF) that provide obstacle-aware guidance for both the robot and the carried object (Section IV-A), a dual-agent learning architecture that decouples upper- and lower-body control to manage the large whole-body action space (Section IV-B), and a specialist-to-generalist distillation scheme that trains scene-specific teachers and distills them into a single generalizable policy (Section IV-C). See Figure 2 for an overview of our full proposed system.

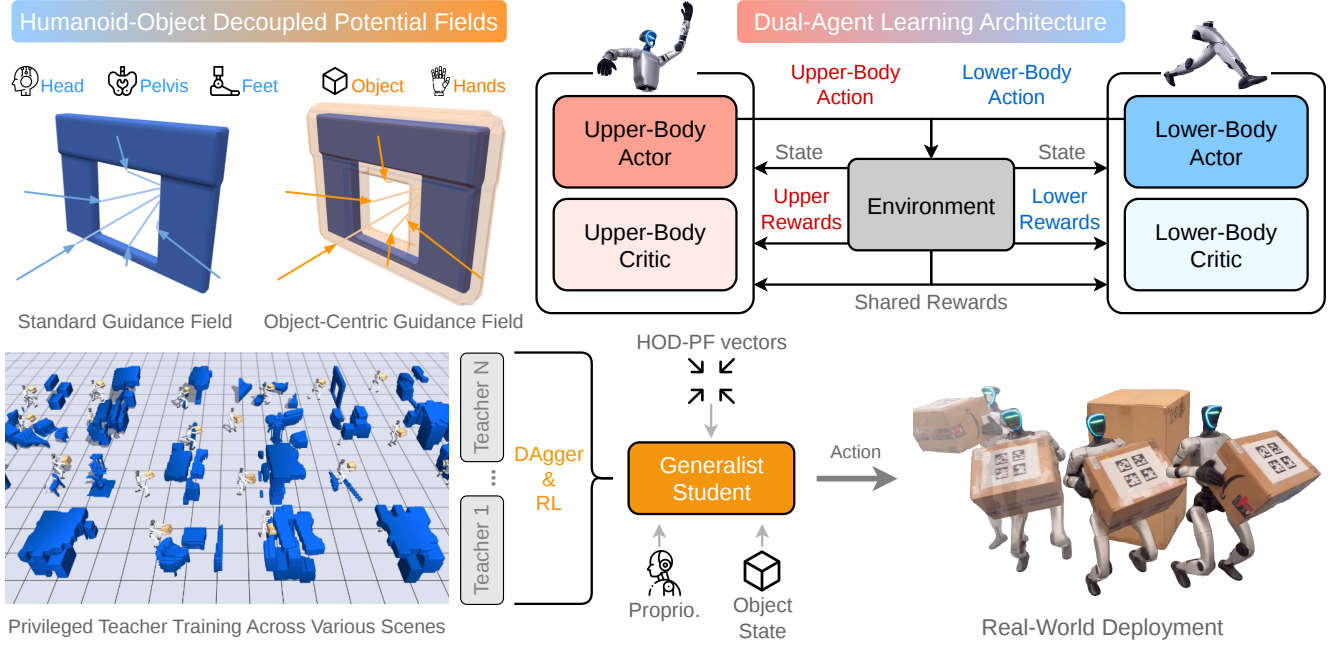


Fig. 2: **Method Overview.** **HOTICE** is composed of three main components. **Humanoid-Object Decoupled Potential Fields** (top left): a Standard Guidance Field guides the robot’s head, pelvis, and feet, while an Object-Centric Guidance Field guides the carried object and the robot’s hands. **Dual-Agent Learning Architecture** (top right): the whole-body control is split between an upper-body actor-critic and a lower-body actor-critic, each with its own body-specific reward set, but remain coordinated via unified state observations and shared whole-body rewards. **Specialist-to-Generalist Training** (bottom): privileged teacher policies are trained across diverse procedurally generated and realistic scenes, then distilled via DAgger and RL fine-tuning into a single generalist student policy that is deployed on the real Unitree G1 humanoid robot.

A. Humanoid-Object Decoupled Potential Fields

Our HOD-PF builds on, and significantly extends, the Humanoid Potential Field (HumanoidPF) introduced in [21]. Fundamentally, HumanoidPF instantiates the Artificial Potential Field [37] by querying it at multiple body keypoints. Despite achieving strong results in collision-free humanoid traversal, HumanoidPF is designed only to control the robot’s legs and does not account for a carried object. In contrast, our proposed HOD-PF is designed for whole-body control and object transportation, where the object adds considerable payload mass and volume that make the task substantially more difficult. Specifically, HOD-PF is composed of two decoupled guidance fields: a Standard Guidance Field that drives the robot’s head, pelvis, and feet to enable traversal and collision avoidance, and an Object-Centric Guidance Field that guides the object’s motion while maintaining robot-object synchronization during traversal.

Standard Guidance Field. We follow HumanoidPF [21] to construct our Standard Guidance Field (SGF), which is derived from an attractive field $\mathbf{U}_{\text{att}} : \mathbb{R}^3 \rightarrow \mathbb{R}$ and a repulsive field $\mathbf{U}_{\text{rep}} : \mathbb{R}^3 \rightarrow \mathbb{R}$. In particular, at any position $\mathbf{x}$ in the 3D space, the values of $\mathbf{U}_{\text{att}}$ and $\mathbf{U}_{\text{rep}}$ at $\mathbf{x}$ are computed as:

$$\mathbf{U}_{\text{att}}(\mathbf{x}) = \eta D_{\text{geo}}(\mathbf{x}, \mathbf{g}), \quad (1)$$

$$\mathbf{U}_{\text{rep}}(\mathbf{x}) = \begin{cases} \frac{1}{2} \xi \left(\frac{1}{\text{sdf}(\mathbf{x})} - \frac{1}{d_0} \right)^2, & \text{sdf}(\mathbf{x}) \leq d_0, \\ 0, & \text{sdf}(\mathbf{x}) > d_0. \end{cases} \quad (2)$$

Here, $D_{\text{geo}}(\mathbf{x}, \mathbf{g})$ is the geodesic distance between $\mathbf{x}$ and the goal $\mathbf{g}$, i.e., the length of the shortest path from $\mathbf{x}$ to $\mathbf{g}$ without intersecting obstacles $\mathcal{C}$, $\text{sdf}(\mathbf{x})$ is the signed distance function of $\mathbf{x}$ with respect to $\mathcal{C}$, d_0 determines the influence range of obstacles, and η and ξ are scaling factors. We set $d_0 = 0.2\text{m}$, $\eta = 0.6$, and $\xi = 1 \times 10^{-3}$, inherited from [21]. Intuitively, $\mathbf{U}_{\text{att}}$ decreases toward $\mathbf{g}$, while $\mathbf{U}_{\text{rep}}$ grows sharply as the robot approaches $\mathcal{C}$. The SGF $\mathbf{F}_{\text{standard}} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is then obtained as the negative gradient of their sum:

$$\mathbf{F}_{\text{standard}}(\mathbf{x}) = -\nabla_{\mathbf{x}} \left(\mathbf{U}_{\text{att}}(\mathbf{x}) + \mathbf{U}_{\text{rep}}(\mathbf{x}) \right). \quad (3)$$

At each position $\mathbf{x}$, $\mathbf{F}_{\text{standard}}(\mathbf{x})$ therefore points in the direction of steepest decrease of the combined potential, pulling the robot toward $\mathbf{g}$ while pushing it away from obstacles $\mathcal{C}$. An example visualization of $\mathbf{F}_{\text{standard}}$ can be seen in Figure 3. Our RL policy’s input then contains the $\mathbf{F}_{\text{standard}}$ guidance vectors queried at the robot’s head, shoulders, torso, root (pelvis), knees, and feet. In addition, the policy is optimized with rewards that encourage each body part to move along the direction indicated by its corresponding guidance vector. Concretely, let $\mathbf{x}$ be the 3D position of a robot body part among head, pelvis, and feet, and let $\mathbf{v}$ be that part’s Cartesian velocity, the policy is trained to maximize the cosine similarity between $\mathbf{F}_{\text{standard}}(\mathbf{x})$ and $\mathbf{v}$.

Object-Centric Guidance Field. In addition to SGF, we implement a second vector field, the Object-Centric Guidance Field (OGF), which provides collision-avoidance signals specifically for the carried object $\mathcal{O}$ and the robot’s

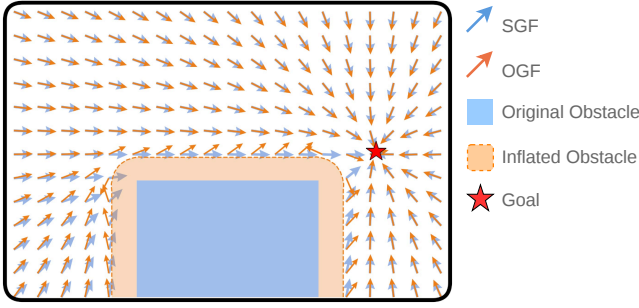


Fig. 3: **Humanoid-Object Decoupled Potential Fields.** The Object-Centric Guidance Field (OGF) provides earlier collision-avoidance signals and affords the robot more head-room than the Standard Guidance Field (SGF).

hands. To construct OGF, we first inflate the obstacles $\mathcal{C}$ by a margin m to obtain $\mathcal{C}_{\text{inflated}}$. The attractive field, repulsive field, and the resulting $\mathbf{F}_{\text{object}}$ guidance field are then computed in the same manner as in SGF, but with respect to $\mathcal{C}_{\text{inflated}}$. Figure 3 shows an example comparison between the resulting SGF and OGF. Computing OGF with respect to $\mathcal{C}_{\text{inflated}}$ offers two main benefits. First, because the object is not directly actuated by the robot’s joints, the margin adds robustness to real-world conditions such as sensor noise and control delay. Second, it makes the object’s avoidance behavior more anticipatory: unlike other robot parts, which can react just before contact, e.g., feet stepping over a ground bar only as it is reached, the robot must begin repositioning the object earlier, since a heavy or large object is considerably harder to maneuver once close to an obstacle. We set $m = 8\text{cm}$ as it balances between collision avoidance performance and traversal feasibility.

On the carried object $\mathcal{O}$, we query $\mathbf{F}_{\text{object}}$ guidance vectors at its keypoints, e.g., the eight corners of box objects. Note that our state-based RL policy is lightweight (MLP-based), thus supporting a wide range of keypoint counts without hurting latency and enabling controllable granularity. In addition to object keypoints, $\mathbf{F}_{\text{object}}$ vectors are also queried for the robot’s hands that control the object to maintain robust robot-object interaction and synchronization. Those guidance vectors are included in the policy’s input along with $\mathbf{F}_{\text{standard}}$ vectors of other body parts. OGF-based steering reward functions are also implemented. Concretely, let $\mathbf{x}$ and $\mathbf{v}$ be the position and velocity of a robot hand or an object keypoint, the policy is trained to maximize the cosine similarity between $[\mathbf{F}_{\text{object}}(\mathbf{x}) - \mathbf{c}]$ and $[\mathbf{v} - \mathbf{c}]$, where $\mathbf{c}$ is the overall navigation command aggregated from $\mathbf{F}_{\text{standard}}$ vectors of the robot’s head, pelvis, and feet. Subtracting $\mathbf{c}$ removes the motion that locomotion already produces, so the reward credits only the residual, obstacle-driven adjustments the hands are actually responsible for.

B. Dual-Agent Learning Architecture

Although the proposed HOD-PF provides helpful and specialized collision-avoidance signals, simultaneously traversing complex scene geometries and maintaining control over an unwieldy object remains a significant challenge for whole-

body loco-manipulation learning. To tackle this, inspired by [38], instead of treating the whole RL policy as a single learning agent, **HOTICE** employs a dual-agent learning architecture, splitting the high-dimensional action space into an upper-body part that controls 8 arm joints and a lower-body part that controls 12 leg joints. Note that we exclude the waist and wrist joints, which are found to be unnecessary for our setting. In our dual-agent design, each agent maintains its own actor and privileged critic, but both read the same observation and share termination conditions, so the two policies stay coupled through the environment rather than through a monolithic network. The two actors emit $\mathbf{a}_{\text{low}} \in \mathbb{R}^{12}$ and $\mathbf{a}_{\text{up}} \in \mathbb{R}^8$, concatenated into the full 20D action $\mathbf{a} = [\mathbf{a}_{\text{low}}, \mathbf{a}_{\text{up}}]$, which is then fed into a PD controller to produce torque commands. We partition the reward terms into a lower set $\mathcal{L}$ (locomotion, balance), an upper set $\mathcal{U}$ (object reaching, grasp maintenance), and a shared set $\mathcal{S}$ (object lifting and HOD-PF rewards, whose values depend jointly on arm articulation and torso pose), yielding two separate reward streams:

$$r_{\text{low}} = \sum_{i \in \mathcal{L} \cup \mathcal{S}} w_i r_i, \quad r_{\text{up}} = \sum_{i \in \mathcal{U} \cup \mathcal{S}} w_i r_i. \quad (4)$$

Each agent then forms its own advantage estimate (denoted as $\hat{A}_{\text{low}}$ and $\hat{A}_{\text{up}}$) from its own reward stream and value function, and the two PPO [39] clipped surrogate objectives are optimized jointly:

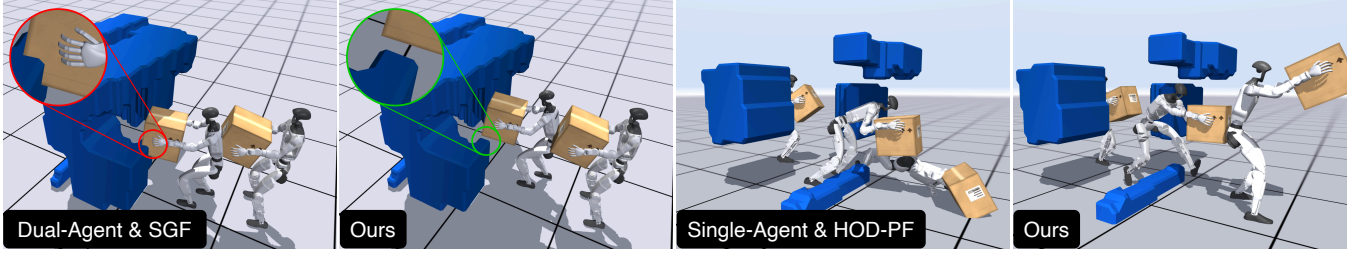
$$\mathcal{J} = \mathcal{J}_{\text{low}}^{\text{PPO}}(\hat{A}_{\text{low}}) + \mathcal{J}_{\text{up}}^{\text{PPO}}(\hat{A}_{\text{up}}). \quad (5)$$

Because the shared terms enter both reward streams, the two agents remain jointly accountable for the carry, while the locomotion and grasping signals no longer compete inside a single advantage estimate.

C. Specialist-to-Generalist Training

To enable a policy that can effectively handle a wide range of cluttered settings, we apply a scalable specialist-to-generalist distillation similar to [21]. More specifically, we first train in parallel multiple scene-specific teacher policies that have access to noiseless, non-delayed privileged states. We employ the scene construction pipeline in [21], which supports procedurally generated scenes and realistic scenes cropped from the 3D-FRONT dataset [40]. While the former enables the creation of highly complex scenes that push the limits of the learned policy, the latter improves generalization to real-world scenarios. Note that we do not perform any scene-specific reward tuning in training teacher policies, as our decoupled potential fields already serve as a generalizable representation and reward structure. The trained teacher policies are then distilled into a single deployable generalist student policy via DAgger distillation [41]. On top of DAgger distillation, we also do PPO training to further refine the distilled student policy. The PPO training of the student policy follows the same dual-agent framework used to train its teachers, described in Section IV-B. To mitigate the sim2real gap for the student policy, we simulate sensor noises, control delays, etc., in its training. We employ MuJoCo Playground [42] for our policy training.

(A) Qualitative Comparisons



(B) Our Method's Performance in Various Scenes

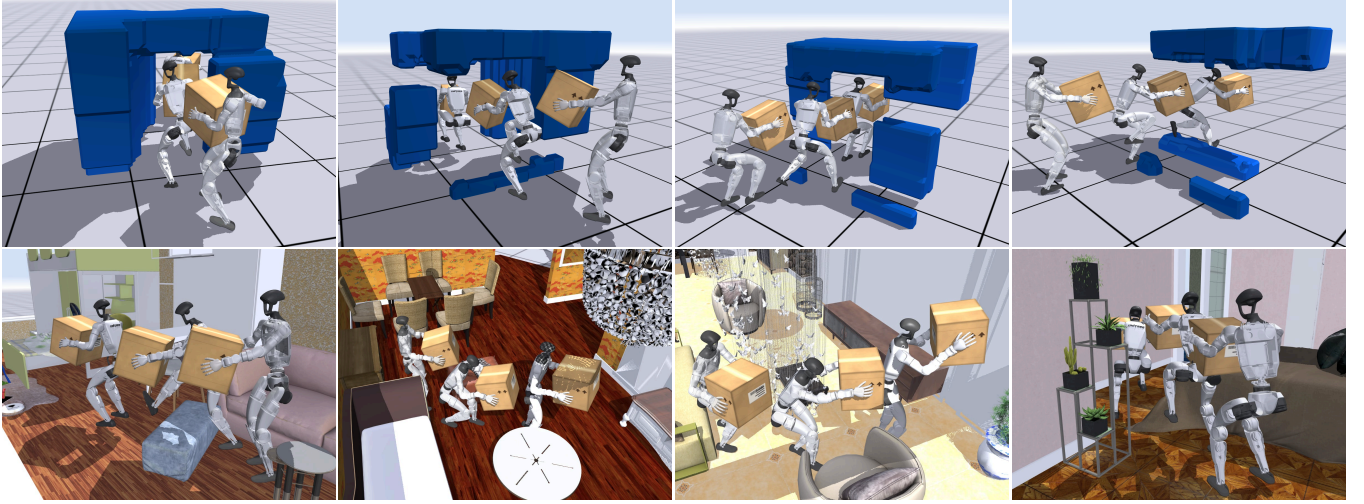


Fig. 4: **Qualitative Results in Simulation.** (A)-*left comparison* – The proposed decoupled potential fields enable **HOTICE** to more reliably avoid collisions during object transportation. (A)-*right comparison* – Our method is capable of maintaining the robot’s balance and stability under challenging obstacle conditions thanks to the dual-agent architecture. (B) – Our method effectively handles a wide variety of challenging procedurally generated scenes (*upper row*) and realistic indoor scenes (*lower row*), without any scene-specific reward tuning.

V. EXPERIMENTS

We conduct extensive experiments in both simulation (Section V-A) and the real world (Section V-B) to validate the performance and generalizability of our proposed **HOTICE**.

A. Simulation Experiments

Main Experiments. We first investigate the effectiveness of our method in enabling humanoid robots to transport objects in cluttered scenes while avoiding collisions with complex obstacles. Since no prior method directly addresses our problem setting, we assess the significance of our key contributions through ablations of **HOTICE**. In particular, we design three baselines to compare with our method: (1) a variant that employs only the Standard Guidance Field, effectively a naive application of HumanoidPF [21], to all keypoints of the robot and the carried object, without a dedicated Object-Centric Guidance Field; (2) a single-agent variant that uses HOD-PF but replaces our dual-agent architecture with a unified policy; and (3) a variant using neither the OGF nor the dual-agent architecture.

For each of the four methods, we train teacher policies for 400M timesteps on 20 scenes of diverse types of obstacles and geometries and then distill them into a student policy using our distillation process described in Section IV-C.

Refer to Figure 4 for some example scenes we use in this experiment. We use cuboid-shaped boxes as our primary objects unless stated otherwise. Box mass and size are randomized, capped at a maximum mass of 3.5 kg and a maximum size of $40 \times 40 \times 40 \text{ cm}^3$, chosen to match the payload and scale of the G1 robot and to ensure traversal feasibility under extreme conditions in our simulation scenes. In evaluation, we apply input noises, delays, and perturbations to reflect the real-world conditions. We benchmark all methods across 20 scenes, each with 1,000 rollouts, using two metrics: (1) success rate (SR), the percentage of trials in which the robot reaches within 0.1 m of the target location within 5 s without falling, dropping the box, or colliding with any obstacle (either the robot or the box); and (2) mean distance to goal (mDist, m), the average distance between the robot’s root and the target at the end of every trial.

Table I shows that our proposed **HOTICE** outperforms other baselines on both metrics by a large margin, demonstrating the importance of both the dual-agent architecture and the Humanoid-Object Decoupled Potential Fields. More qualitative results of **HOTICE** and comparisons to its ablations can be seen in Figure 4, where the effectiveness of our method in avoiding collisions and handling diverse challenging scenarios is further illustrated.

Method	SR (%) $\uparrow$	mDist (m) $\downarrow$
Single-Agent & SGF	58.7	0.296
Single-Agent & HOD-PF	65.7	0.267
Dual-Agent & SGF	77.1	0.195
Dual-Agent & HOD-PF (Ours)	88.5	0.130

TABLE I: **Quantitative Simulation Results.** Success rates (SR) and mean distances to goal (mDist) are reported over 20,000 trials in 20 scenes with diverse obstacle conditions.

Generalization to Unseen Scenes. We distill a generalist policy from 75 specialist teacher policies trained on 75 scenes covering diverse obstacle configurations and evaluate it on 50 unseen environments. To better interpret the results, we measure the similarity between each unseen scene and the training scenes in terms of clutter configuration. Specifically, for a pair of an unseen scene and a training scene, we discretize their SGFs and OGFs at a resolution of $(0.04\text{ m})^3$ and compute the average cosine similarity between guidance vectors at corresponding positions, across both fields and all positions. The similarity between an unseen scene and the training set is then defined as the maximum similarity across all training scenes. We present SR and mDist against this similarity measure for every unseen scene.

The results in Figure 5 indicate that success rate tends to decrease and the mean distance to goal tends to increase as scenes become less similar to the training scenes, as expected. Nonetheless, our method maintains strong performance across unseen scenes, achieving an average success rate of 80.1% (with a floor of approximately 70%) and an average mDist of 0.27 m. Overall, these results highlight both the generalizability of HOD-PF and the effectiveness of teacher-to-student policy distillation.

Generalization to Different Object Shapes. We next investigate whether **HOTICE** generalizes to object shapes beyond the cuboid box, namely cylinders and spheres. The main difference between the policies for these shapes and the box policy lies in the keypoints at which the Object-Centric Guidance Field vectors are queried on the object. Specifically, for cylinder objects, we query the field at four evenly spaced points on the top rim and four on the bottom rim; for sphere objects, we use eight fixed world-frame octant directions at the sphere’s radius. Since every shape uses exactly eight keypoints, the observation layout, network architecture, and action space remain unchanged across all three shapes. Beyond this, we apply only minimal shape-specific reward adaptations, e.g., for hand placement and object uprightness, while all other reward terms, weights, and hyperparameters are inherited unchanged from the box policy. We train and evaluate shape-specific policies on the same 20 scenes as in the main experiment.

As shown in Table II, performance on both metrics remains consistent across all three object types, demonstrating the effectiveness of our method across widely-used shape primitives. More qualitative results of the cylinder and sphere policies performing in different procedurally generated scenes and realistic scenes are illustrated in Figure 6.

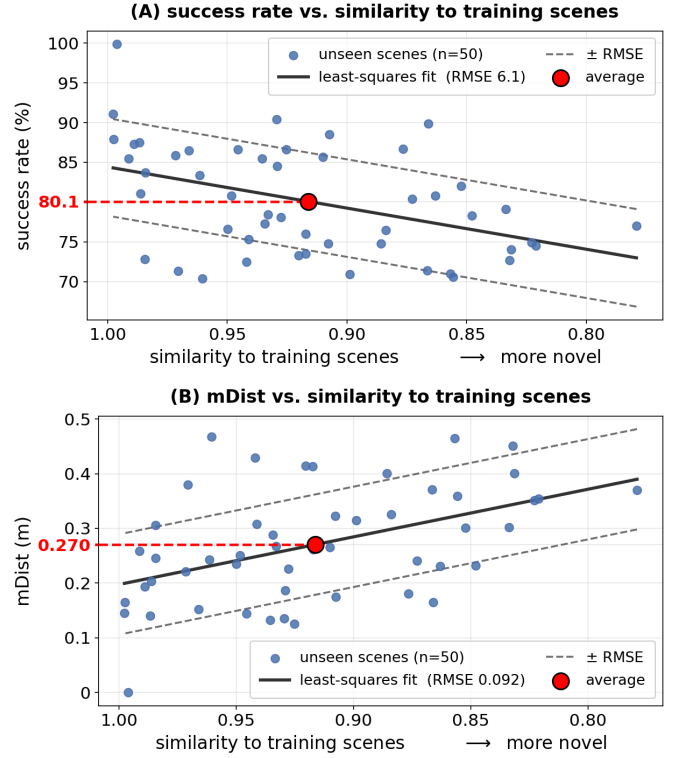


Fig. 5: **Results on Unseen Scenes.** We evaluate our generalist policy on 50 unseen scenes in simulation and plot (A) success rate and (B) mean distance to goal against similarity to the training set. Least-squares fits (shaded by $\pm\text{RMSE}$) indicate a graceful degradation in both metrics with increasing novelty, while our method still maintains strong average performance.

	Cuboid 	Cylinder 	Sphere 
SR (%) $\uparrow$	88.5	86.7	87.2
mDist (m) $\downarrow$	0.130	0.127	0.131

TABLE II: **Results on Different Object Shapes.** Results are reported for three shape-specific policies distilled from corresponding teachers trained on the same 20 scenes. The cuboid results match those of our method in Table I since we use the same policy here.

B. Real-World Results

Deployment Details. We perform sim2real deployment for our method on a Unitree G1 robot. Obstacle scenes are first reconstructed via SLAM [43] by walking the robot around the environment without the box. During inference, HOD-PF guidance vectors are inferred from the robot’s real-time localization and joint angles, and the box pose, which we estimate via the AprilTags [44] attached to the box and the robot’s head camera. Note that our method could in principle support live LiDAR-based scene reconstruction, similar to [21]. However, the G1 is equipped with only a single LiDAR sensor, whose field of view is significantly occluded by the carried box. We attribute this to a hardware limitation rather than a limitation of the method itself.

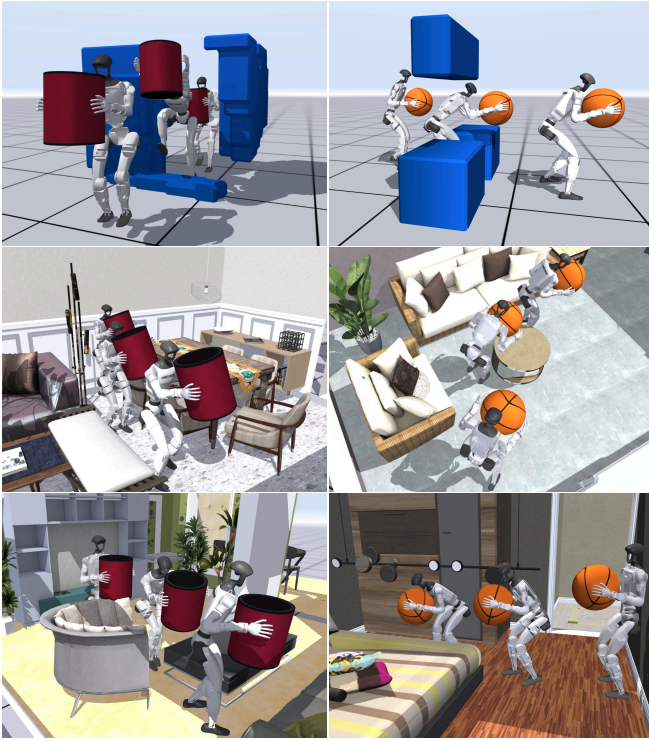


Fig. 6: **Results on Other Objects.** **HOTICE** is also generalizable to cylinder (left) and sphere object shapes (right) with only minimal shape-specific reward tuning.

Real-World Experiments. We benchmark our method and its ablations on three representative obstacle types: side, overhead, and ground obstacles. For each obstacle type, we perform 6 trials per method. At the start of each trial, a shared box-pickup policy first picks up the box from a support surface, followed by the transportation policy; both run at 50 Hz. The success rate results are reported in Table III. In particular, our method obtains the best real-world results, succeeding in all side obstacle trials and achieving 5/6 success rates on overhead and ground obstacle scenes. Examples of real-world rollouts are shown in Figure 7, where **HOTICE** demonstrates high performance across diverse settings, while the SGF-only policy collides more frequently and the single-agent policy struggles to remain stable during challenging motions such as stepping over the ground pipe obstacle. Overall, these results demonstrate the effectiveness of our proposed **HOTICE** in real-world environments.

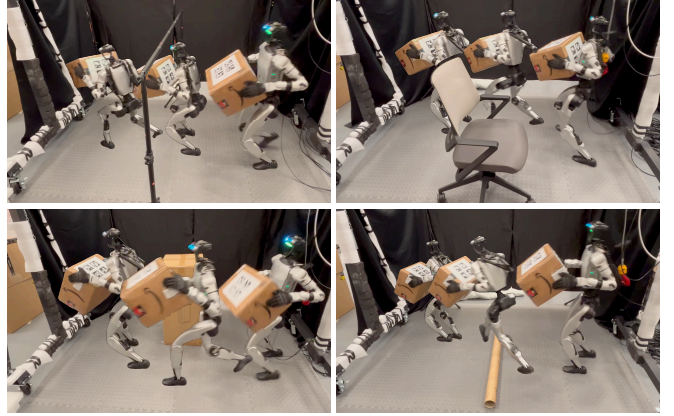
VI. DISCUSSION AND LIMITATIONS

Despite achieving strong performance across several settings, **HOTICE** is not without limitations. First, while our method exhibits strong collision-avoidance capability, we do not model post-collision scenarios. As a result, if a collision applies a disturbance large enough to exceed the robot’s recovery capacity, the robot falls. Moreover, although our loco-manipulation policy is trained to manipulate the carried object to avoid obstacles, more complex manipulation strategies remain out of reach for our method, such as carrying the object on the robot’s shoulder or tucking it against the hip. We also do not tackle dynamic scenes with

Method	Side	Overhead	Ground
Single-Agent & SGF	3/6	3/6	1/6
Single-Agent & HOD-PF	6/6	3/6	2/6
Dual-Agent & SGF	4/6	4/6	4/6
Dual-Agent & HOD-PF (Ours)	6/6	5/6	5/6

TABLE III: **Real-World Experimental Results.** For each method, we conduct 6 trials per obstacle type.

Ours



Dual-Agent & SGF - collide with the side obstacle



Single-Agent & HOD-PF - strike the ground obstacle and fall

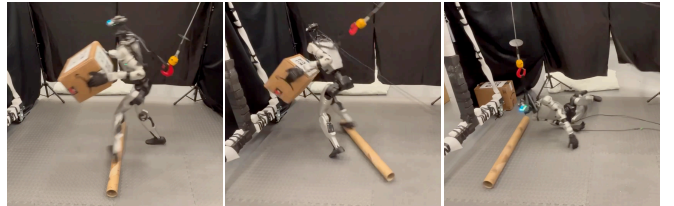


Fig. 7: **Real-World Qualitative Results.** **HOTICE** performs effectively across different types of obstacles, whereas the SGF-only policy is more prone to collisions, and the single-agent policy underperforms on challenging motions such as hurdling over the ground obstacle.

moving obstacles in this work. Enabling object transportation in such dynamic settings could be significantly beneficial for human-centric, open-world environments. Finally, one could investigate scenarios involving “acceptable” collisions alongside harder ones, for instance, the robot leaning against a wall to navigate a tight space, or brushing against a soft obstacle such as a bean bag chair without penalty. More broadly, the robot could also interact with obstacles to actively make way, e.g., using its feet to push small obstacles aside or using its back to push a door. We leave these interesting research ideas for future work.

VII. CONCLUSION

We present **HOTICE**, a whole-body humanoid learning framework for transporting objects through cluttered environments that constrain the robot and its payload from the ground, sides, and above. By encoding humanoid-obstacle and object-obstacle geometry into decoupled guidance fields, splitting whole-body control across a dual-agent architecture, and consolidating scene-specific teachers into a single generalist policy, **HOTICE** achieves reliable, collision-aware transport across diverse clutter and object shapes. Extensive experiments in MuJoCo and on a Unitree G1 demonstrate strong performance on previously unseen scenes and robust sim2real transfer. We hope **HOTICE** inspires follow-up work on humanoid loco-manipulation in clutter.

REFERENCES

- [1] A. S. Brown, “Amazon’s robot arms break ground in safety and technology,” *Amazon Science*, 2022.
- [2] D. K. Mehta, A. Joglekar, and V. Krovı, “Deep reinforcement learning for coordinated payload transport in biped-wheeled robots,” in *ICRA*, 2025.
- [3] H. Lee, H. Kim, and H. J. Kim, “Path planning and control of multiple aerial manipulators for a cooperative transportation,” in *IROS*, 2015.
- [4] Z. Gu, J. Li, W. Shen, W. Yu, Z. Xie, S. McCrory, X. Cheng, A. Shamsah, R. Griffin, C. K. Liu *et al.*, “Humanoid locomotion and manipulation: Current progress and challenges in control, planning, and learning,” *IEEE/ASME Transactions on Mechatronics*, 2026.
- [5] E. Todorov, T. Erez, and Y. Tassa, “Mujoco: A physics engine for model-based control,” in *IROS*, 2012.
- [6] R. Buchanan, L. Wellhausen, M. Bjelonic, T. Bandyopadhyay, N. Kottege, and M. Hutter, “Perceptive whole-body planning for multilegged robots in confined spaces,” *Journal of Field Robotics*, 2021.
- [7] R. Buchanan, T. Bandyopadhyay, M. Bjelonic, L. Wellhausen, M. Hutter, and N. Kottege, “Walking posture adaptation for legged robot navigation in confined spaces,” *RA-L*, 2019.
- [8] J.-R. Chiu, J.-P. Sleiman, M. Mittal, F. Farshidian, and M. Hutter, “A collision-free mpc for whole-body dynamic locomotion and manipulation,” in *ICRA*, 2022.
- [9] D. Hoeller, N. Rudin, D. Sako, and M. Hutter, “Anymal parkour: Learning agile navigation for quadrupedal robots,” *Science Robotics*, 2024.
- [10] T. Miki, J. Lee, L. Wellhausen, and M. Hutter, “Learning to walk in confined spaces using 3d representation,” in *ICRA*, 2024.
- [11] Z. Xu, A. H. Raj, X. Xiao, and P. Stone, “Dexterous legged locomotion in confined 3d spaces with reinforcement learning,” in *ICRA*, 2024.
- [12] T. He, C. Zhang, W. Xiao, G. He, C. Liu, and G. Shi, “Agile but safe: Learning collision-free high-speed legged locomotion,” in *RSS*, 2024.
- [13] Z. Zhuang, S. Yao, and H. Zhao, “Humanoid parkour learning,” in *CoRL*, 2024.
- [14] Z. Wu, X. Huang, L. Yang, Y. Zhang, X. Chen, P. Abbeel, R. Duan, A. Kanazawa, C. Sferrazza, G. Shi *et al.*, “Perceptive humanoid parkour: Chaining dynamic human skills via motion matching,” in *RSS*, 2026.
- [15] C. Zhang, W. Xiao, T. He, and G. Shi, “Wococo: Learning whole-body humanoid control with sequential contacts,” in *CoRL*, 2024.
- [16] L. Yang, X. Huang, Z. Wu, A. Kanazawa, P. Abbeel, C. Sferrazza, C. K. Liu, R. Duan, and G. Shi, “Omni-target: Interaction-preserving data generation for humanoid whole-body loco-manipulation and scene interaction,” in *ICRA*, 2026.
- [17] H. Wang, Z. Wang, J. Ren, Q. Ben, T. Huang, W. Zhang, and J. Pang, “Beamdojo: Learning agile humanoid locomotion on sparse footholds,” in *RSS*, 2025.
- [18] S. Zhao, Y. Zhang, Z. Lu, P. Abbeel, R. Duan, K. Sreenath, Y. Wang, C. K. Liu, and G. Shi, “Ladderman: Learning humanoid perceptive ladder climbing,” in *CoRL*, 2026.
- [19] Z. Zhang, K. Wen, M. Xu, J. He, C. Li, T. Miki, C. Schwarke, C. Zhang, X. B. Peng, and M. Hutter, “Learning whole-body humanoid locomotion via motion generation and motion tracking,” *RA-L*, 2026.
- [20] Y. Wang, T. Leng, C. Lin, S. Liu, S. Simon, B. Chen, J. Francis, and D. Zhao, “Apex: Learning adaptive high-platform traversal for humanoid robots,” in *IROS*, 2026.
- [21] H. Xue, S. Liang, Z. Zhang, Z. Zeng, Y. Liu, Y. Lian, J. Wang, Q. Liu, X. Shi, and L. Yi, “Collision-free humanoid traversal in cluttered indoor scenes,” *RA-L*, 2026.
- [22] A. Li, Y. Chen, Z. Li, Z. Cao, J. Ren, M. Tomizuka, and D. Shah, “Tango: Humanoid navigation in cluttered environments with a whole-body vision-language-action model,” in *CoRL*, 2026.
- [23] E. Yoshida, C. Esteves, I. Belousov, J.-P. Laumond, T. Sakaguchi, and K. Yokoi, “Planning 3-d collision-free dynamic robotic motion through iterative reshaping,” *T-RO*, 2008.
- [24] K. Harada, S. Kajita, H. Saito, M. Morisawa, F. Kanehiro, K. Fujiwara, K. Kaneko, and H. Hirukawa, “A humanoid robot carrying a heavy object,” in *ICRA*, 2005.
- [25] J. Dao, H. Duan, and A. Fern, “Sim-to-real learning for humanoid box loco-manipulation,” in *ICRA*, 2024.
- [26] M. Kim, K.-C. Chen, A. Shrestha, L. Fuxin, S. Lee, and A. Fern, “Humanoid hanoi: Investigating shared whole-body control for skill-based box rearrangement,” in *CoRL*, 2026.
- [27] C. Zheng, K. Chen, Z. Bi, Y. Li, L. Pan, J. Zhou, H. Li, and J. Ma, “Embracing bulky objects with humanoid robots: Whole-body manipulation with reinforcement learning,” in *ICRA*, 2026.
- [28] Z. Bi, Q. Zhang, G. Zhao, J. Cao, X. Luo, Y. Zhang, J. Xu, R. Geng, Y. Li, A. F. Luo *et al.*, “Genhoi: Contact-aware humanoid-object interaction by imitating generated videos without task-specific training,” *arXiv preprint arXiv:2606.12995*, 2026.
- [29] L. Fu, Y. Zhong, X. Li, Y. Liu, Z. Xu, J. Tang, and S. Li, “Load-aware locomotion control for humanoid robots in industrial transportation tasks,” *IEEE Transactions on Industrial Electronics*, 2026.
- [30] J. Kang, H. Cho, S. Kang, and D. Koo, “Splitadapter: Load-aware humanoid loco-manipulation via factorized adaptation,” *arXiv preprint arXiv:2606.03297*, 2026.
- [31] X. Cui, L. Feng, Y. Zhou, H. Han, Z. Liu, and H. Wang, “Pilot: A perceptive integrated low-level controller for loco-manipulation over unstructured scenes,” *arXiv preprint arXiv:2601.17440*, 2026.
- [32] Y. Zhang, Y. Seo, J. Chen, Y. Yuan, K. Sreenath, P. Abbeel, C. Sferrazza, K. Liu, R. Duan, and G. Shi, “Rpl: Learning robust humanoid perceptive locomotion on challenging terrains,” in *CoRL*, 2026.
- [33] Z. Gu, Y. Chen, Z. Chai, A. Cueva, T. Nguyen, Y. Wu, H. Xue, M. Kim, I. Legene, F. Liu *et al.*, “Refine-dp: Diffusion policy fine-tuning for humanoid loco-manipulation via reinforcement learning,” *RA-L*, 2026.
- [34] S. Yin, Y. Ze, H.-X. Yu, C. K. Liu, and J. Wu, “Visualmimic: Visual humanoid loco-manipulation via motion tracking and generation,” *arXiv preprint arXiv:2509.20322*, 2025.
- [35] Y.-J. Wang, J. Li, S. Chen, T. E. Truong, P. Xu, P. Abbeel, R. Duan, K. Sreenath, A. Kanazawa, C. Sferrazza *et al.*, “Vlk: Learning humanoid loco-manipulation from synthetic interactions in reconstructed scenes,” in *CoRL*, 2026.
- [36] T. Xie, H. Zhang, J. Park, Z. Wang, B. Wen, J. Li, X. Li, Q. Ben, H. Weng, Y. Ye *et al.*, “Grail: Generating humanoid loco-manipulation from 3d assets and video priors,” in *CoRL*, 2026.
- [37] O. Khatib, “Real-time obstacle avoidance for manipulators and mobile robots,” *IJRR*, 1986.
- [38] Y. Zhang, Y. Yuan, P. Gurunath, I. Gupta, S. Omidshafiei, A.-a. Aghamohammadi, M. Vazquez-Chanlatte, L. Pedersen, T. He, and G. Shi, “Falcon: Learning force-adaptive humanoid loco-manipulation,” in *LADC*, 2026.
- [39] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” *arXiv preprint arXiv:1707.06347*, 2017.
- [40] H. Fu, B. Cai, L. Gao, L.-X. Zhang, J. Wang, C. Li, Q. Zeng, C. Sun, R. Jia, B. Zhao *et al.*, “3d-front: 3d furnished rooms with layouts and semantics,” in *ICCV*, 2021.
- [41] S. Ross, G. Gordon, and D. Bagnell, “A reduction of imitation learning and structured prediction to no-regret online learning,” in *AISTATS*, 2011.
- [42] K. Zakka, B. Tabanpour, Q. Liao, M. Haiderbhai, S. Holt, J. Y. Luo, A. Allshire, E. Frey, K. Sreenath, L. A. Kahrs *et al.*, “Mujoco playground,” in *RSS*, 2025.
- [43] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang, “Fast-lid2: Fast direct lidar-inertial odometry,” *T-RO*, 2022.
- [44] E. Olson, “Apriltag: A robust and flexible visual fiducial system,” in *ICRA*, 2011.