Humanoid Goalkeeper: Learning from Position Conditioned Task-Motion Constraints

Junli Ren [*,1,2]   Junfeng Long[*,2]   Tao Huang[2]   Huayi Wang[2]   Zirui Wang[2]   Feiyu Jia[2]
Wentao Zhang[2]   Jingbo Wang[2,†]   Ping Luo[1,2,†]   Jiangmiao Pang[2,†]

[1]The University of Hong Kong   [2]Shanghai AI Laboratory   *Equal Contribution   †Equal Advising
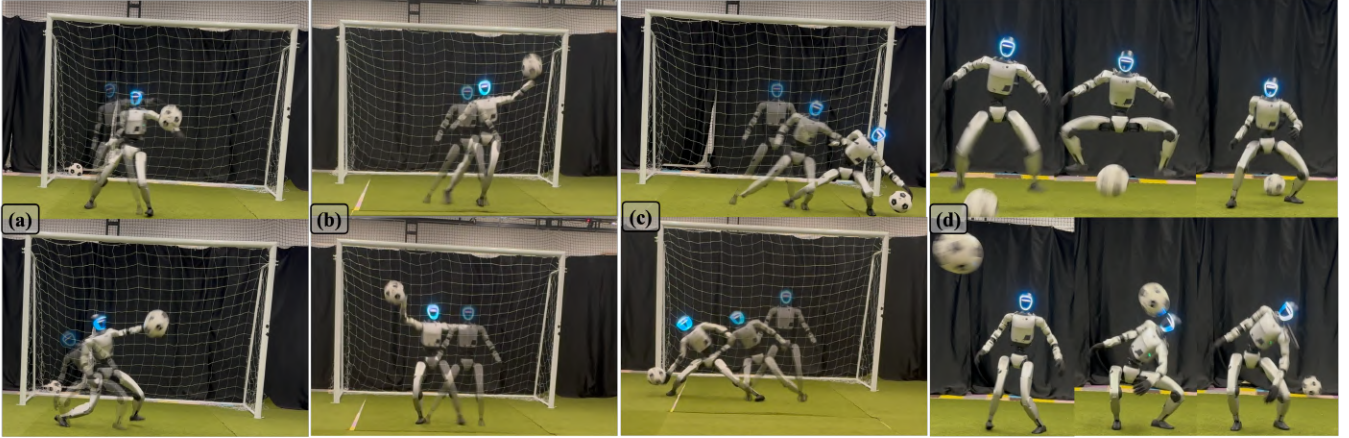
Fig. 1: We present **Humanoid Goalkeeper**, capable of performing goalkeeping tasks across various regions with a wide operational range (b, c), and initiating from arbitrary postures (a). Our policy enables the robot to escape an incoming ball with jump and squat motions (d).

Abstract— **We present a reinforcement learning framework for autonomous goalkeeping with humanoid robots in real-world scenarios. While prior work has demonstrated similar capabilities on quadrupedal platforms, humanoid goalkeeping introduces two critical challenges: (1) generating natural, human-like whole-body motions, and (2) covering a wider guarding range with an equivalent response time. Unlike existing approaches that rely on separate teleoperation or fixed motion tracking for whole-body control, our method learns a single end-to-end RL policy, enabling fully autonomous, highly dynamic, and human-like robot-object interactions. To achieve this, we integrate multiple human motion priors conditioned on perceptual inputs into the RL training via an adversarial scheme. We demonstrate the effectiveness of our method through real-world experiments, where the humanoid robot successfully performs agile, autonomous, and naturalistic interceptions of fast-moving balls. In addition to goalkeeping, we demonstrate the generalization of our approach through tasks such as ball escaping and grabbing. Our work presents a practical and scalable solution for enabling highly dynamic interactions between robots and moving objects, advancing the field toward more adaptive and lifelike robotic behaviors. More details, including videos and implementation, are available on our project website and GitHub repository.**

I. INTRODUCTION

Developing robotic soccer capabilities inherently requires tight integration of perception, decision-making, and agile motor control, making it an appealing benchmark [1–5] for evaluating motion intelligence. Within this domain, goalkeeping stands out as a distinct subtask with many underexplored challenges. Unlike dribbling [6, 7] or shooting [8, 9], which primarily involve the lower limbs and interact with slow or static objects, goalkeeping requires rapid, full-body responses to highly dynamic stimuli. Advancing robotic goalkeeping is therefore a crucial step toward fully autonomous, real-time physical intelligence.

Despite the growing interest in robotic goalkeeping, including recent advances in dynamic skills on quadrupeds, existing systems remain limited in several critical aspects: being confined to kid-sized platforms [10], relying on fixed motion primitives [11] and operating within narrow interception ranges. Meanwhile, emerging developments in reinforcement learning (RL) [12, 13], imitation-driven motion priors [14, 15], and human motion recovery pipelines [16–18] open new avenues for addressing these limitations. Coupled with the advent of high-performance humanoid hardware and scalable learning frameworks, it is now feasible to explore goalkeeper behaviors that: (1) exhibit human-like, whole-body interactive motions; (2) respond effectively under extreme time constraints; and (3) perform highly dynamic actions to cover broad and diverse interception regions.

In this paper, we present the Humanoid Goalkeeper, a reinforcement learning–based framework for agile and autonomous goalkeeping on humanoid robots. Our method learns a single end-to-end control policy that integrates task-specific rewards with motion priors through an adversarial training scheme, enabling policies that optimize both task performance and motion realism. To address the challenge of wide operational coverage, we condition motion learning on ball landing positions, effectively dividing the constraint space. To ensure real-world feasibility, we incorporate perception noise, trajectory estimation, and multi-modal sensing into the training loop to close the sim-to-real gap. Through extensive evaluations in both simulation and hardware, we demonstrate that our approach enables fast, reactive, and naturalistic

interception of high-speed flying balls. Furthermore, we show that the learned policy generalizes to related dynamic interaction tasks such as escaping and grabbing, highlighting the framework's potential for broader lifelike and adaptive humanoid behaviors.

II. RELATED WORK

Dynamic Object Interactions with floating base. Recent advances have yielded impressive results in interactions between dynamic object and mobile robots, with notable applications in tasks such as badminton [19, 20], table tennis [21–25], mobile-manipulator-based catching [26, 27], throwing [28], and soccer [1, 4, 6, 7, 29, 30]. These studies have successfully demonstrated millisecond-level interactions between robots and dynamic objects, with a strong integration of perception, planning, and control systems. These methods excel in specific domains such as performing consecutive actions [21], multi-agent coordination [4, 31] or professional level skill [22], but they are constrained by narrow movement ranges or lack the dynamic whole-body agility required for humanoid systems. In this work, we address these gaps by developing a humanoid goalkeeper capable of covering wide movement ranges, while executing agile and human-like whole-body motions.

Whole body control with motion priors Integrating imitation learning with the RL framework has become a popular approach for transferring human motion skills to humanoid robots. A majority of these methods strictly mimic reference trajectories following the fixed time phase [32–36]. These methods are typically limited to non-interactive tasks, such as dancing or locomotion [37, 38], or are restricted to predefined interaction moments within the reference [21, 39–43]. In contrast, adversarial motion priors (AMP) [14] offer an alternative approach by integrating expert data without enforcing strict temporal alignment. However, this potential has yet to be fully explored in real-world humanoid tasks, with current frameworks limited to simpler locomotion tasks [44, 45]. Meanwhile, multi-motion prior synthesis has been explored through task-condition tokens [46] or sub-task policy selection [47] in the animation domain. Similarly, previous quadruped goalkeepers have learned to select from a set of pre-learned low-level skills [29] for different ball target regions. Building upon these experiences, our approach trains the robot with task rewards and motion discriminators specific to different ball target regions, activating only the corresponding constraints during each trial. The proposed pipeline results in a one-stage, end-to-end, hardware-deployable policy that achieves dynamic interactions while maintaining reference motion resemblance.

III. METHOD

We create the humanoid goalkeeper that can autonomously perform successfully keeping skills while adaptively conforming to human motion priors. Our approach (Fig. 2) centers on an end-to-end reinforcement learning (RL) policy that generates highly dynamic and fully autonomous keeping behaviours. The policy operates directly on real-time ball

position observations, which also serve as conditioning variables for integrated human motion priors. These priors are incorporated into the RL process via an adversarial training scheme, promoting robust and naturalistic motion generation without reliance on predefined trajectories or strict motion tracking.

Our method advances prior work by addressing two critical challenges: (1) We propose a unified framework that jointly enforces task success and motion realism through observation-conditioned adversarial training, eliminating the need for separate optimization stages; and (2) we demonstrate hardware-feasible, high-dynamic performance on real-world humanoid robots.
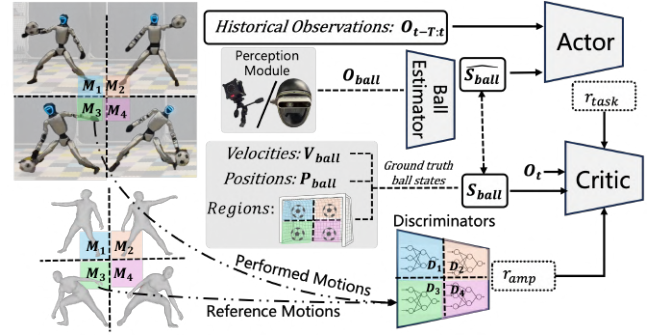

Fig. 2: Method framework: We train our policy using an end-to-end reinforcement learning (RL) framework, incorporating position-conditioned task and motion constraints. Our method is capable of operating with either an onboard camera or the motion-capture (MoCap) system.

A. Goalkeeper Training through RL

Our learning system is designed to optimize two primary objectives: (1) task success and (2) motion resemblance to human demonstrations. This section details our reinforcement learning–based optimization process for enabling successful goalkeeping skills.

1) Training Environment: We train policies using the standard PPO algorithm [12], implemented in the IsaacGym simulator [13]. The policy trains the robot to stop a flying ball before hitting into the goal line.

The observation space includes the observed ball's position O_{ball}, combined with proprioceptive observations O_p. These are used as actor inputs O_t, with a history of length T to capture temporal information necessary for predicting the ball's motion trajectory.

Each episode begins with a ball launched toward the robot. The training environment is divided into k regions, denoted as $\mathcal{R} = 0, 1, 2, \ldots, k$, where each region corresponds to a specific area on the goal line where the ball may arrive. The region is determined at the start of each episode, then we sample a random landing position $\mathbf{p}_{land}$ within that region and assign the ball's velocity accordingly to reach the target. Table I outlines the observation space used during training. All ball-related observations are computed in the robot's local frame, facilitating direct deployment with onboard perception.

TABLE I: Observation spaces for the actor and critic network.

Observation	Actor	Critic
Ball position (in local frame) ($\mathbf{O}_{\text{ball}} \in \mathbb{R}^3$)	✓	✓
Base angular velocity ($\boldsymbol{\omega}_{\text{base}} \in \mathbb{R}^3$)	✓	✓
Projected gravity vector ($\mathbf{g}_{\text{base}} \in \mathbb{R}^3$)	✓	✓
Joint positions ($\mathbf{q} \in \mathbb{R}^{29}$)	✓	✓
Joint velocities ($\dot{\mathbf{q}} \in \mathbb{R}^{29}$)	✓	✓
Previous action ($\mathbf{a}_{\text{last}} \in \mathbb{R}^{29}$)	✓	✓
Base linear velocity ($\mathbf{v}_{\text{base}} \in \mathbb{R}^3$)	–	✓
End target region ($\mathcal{R} = 0, 1, 2, \ldots, k$)	–	✓
End effector target ($\mathbf{p}_{\text{target}} \in \mathbb{R}^3$)	–	✓
Ball velocity (in local frame) ($\mathbf{v}_{\text{ball}} \in \mathbb{R}^3$)	–	✓
Left hand position ($\mathbf{p}_{\text{hand}}^{\text{left}} \in \mathbb{R}^3$)	–	✓
Right hand position ($\mathbf{p}_{\text{hand}}^{\text{right}} \in \mathbb{R}^3$)	–	✓
Reach Distance ($\|\mathbf{p}_{\text{target}} - \mathbf{p}_{\text{hand}}^{(\mathcal{R})}\|$)	–	✓

2) Position-Conditioned Task Rewards: We introduce **position-conditioned task rewards** r_t that guide the robot to keep the ball away from the goal line by tracking a dynamic target. The end target $\mathbf{p}_{\text{target}}$ is designed to prompt the robot to predict the landing point of the approaching ball, touch and punch it away when it is close, and then hold a static position after the ball is stopped (Fig. 3). When the ball is flying towards the rbobot, we have:

$$\mathbf{p}_{\text{target}} = \mathbf{p}_{\text{land}} \cdot \mathbf{1}_{d_x > d_{th}} + \mathbf{p}_{\text{ball}} \cdot \mathbf{1}_{d_x \leq d_{th}}, \quad (1)$$

where $\mathbf{1}_{\text{condition}}$ selects the corresponding target based on the distance between the ball and the robot, d_x denotes the distance between the ball and the goal line, and d_{th} stands for the approaching threshold. Then we define the following smooth, distance-based reward using a sigmoid function to encourage the end-effector (hand) to reach the target position:

$$r_t = \left\{ 1 - \frac{1}{1 + \exp[-\sigma_{\text{keep}}(\|\mathbf{p}_{\text{target}} - \mathbf{p}_{\text{hand}}^{(\mathcal{R})}\| - d_{\text{keep}})]} \right\} \times (1 + \nu^{(\mathcal{R})}), \quad (2)$$

where σ_{keep} is the scaling factor and d_{keep} is a catch distance threshold indicating when the end-effector is sufficiently close to the ball. The reward r_t is conditioned on the landing region $\mathcal{R}$ through computing the distance between the selected hand and the end target $\|\mathbf{p}_{\text{target}} - \mathbf{p}_{\text{hand}}^{(\mathcal{R})}\|$. For example, if the ball is flying toward the left region, the robot should use its left hand to intercept it, and similarly for the right region.
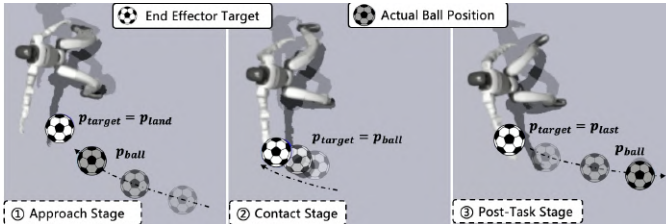

Fig. 3: This design ensures that when the ball is far from the robot, the interaction is based on the predicted landing point, whereas for closer balls, the target position is the actual ball location, allowing for more precise handling near the robot.

We introduce a **position-conditioned dynamics modulation term** $\nu^{(\mathcal{R})}$ to encourage whole-body motions such as lateral shifts or jumps, depending on the region $\mathcal{R}$:

$$\nu^{(\mathcal{R})} = \nu_y \cdot \mathbf{1}_{\mathcal{R} \in \text{right}} + (-\nu_y) \cdot \mathbf{1}_{\mathcal{R} \in \text{left}} + \nu_z \cdot \mathbf{1}_{\mathcal{R} \in \text{up}}, \quad (3)$$

$\nu^{(\mathcal{R})}$ promotes consistent whole-body motions that support coordinated end-effector movements during interception.

In addition to the position-conditioned reward r_t, we incorporate additional reward terms to enforce motion constraints and promote hardware-feasible behavior. The full reward formulation is provided in the appendix.

3) Post-Task Stability: We set the episode length to 3 seconds, which is considerably longer than the ball's flight time (typically 0.4– seconds), in order to allow for post-task stabilization. To encourage the robot to maintain a normalized posture after the ball has passed, we introduce post-task reward terms that promote balance and recovery from the high-dynamic keeping motion. These rewards are generally defined as:

$$r_{\text{post}} = \exp\left(-\sigma_{\text{stable}} \cdot \|\mathbf{Err}_{\text{stable}}\|\right) \cdot \mathbf{1}_{\text{ball stopped}}, \quad (4)$$

where $\mathbf{Err}_{\text{stable}}$ denotes the deviation of key stability-related quantities—such as base height, angular velocity, and orientation—from their nominal values. The indicator function $\mathbf{1}_{\text{ball stopped}}$ activates the reward only after the ball has been kept or flying behind the robot. In addition, we reset a subset of environments not to the default joint configuration, but instead to the joint positions sampled from other ongoing environments. This forces the robot to begin a new trial in the middle of a previous keeping motion, rather than from a nominal reset state. In addition to promoting stability, we found this strategy effective for learning continuous keeping behaviors, as the robot can initiate new episodes without returning to a default pose.

B. Optimization with Motion Constraints

We include a position-conditioned adversarial motion prior (AMP) reward to impose motion constraints during training, encouraging the robot to exhibit appropriate movement styles based on the ball's landing region (see Fig. 2). Inspired by [29], which addresses region-specific goalkeeping in quadrupeds, we incorporate imitation learning into reinforcement learning to promote region-consistent behaviors. These motion constraints are derived from a task-specific motion library constructed from human demonstrations.

1) Reference Motion Curation: We first leverage existing pipelines (GVHMR [16]) that extract humanoid motions from self-recorded RGB videos (Fig. 4). Next, we retarget the extracted motions to the Unitree G1 robot and store the resulting joint position sequences q_t in a region-specific motion reference buffer.

2) Optimize with Motion Constraints: The motion constraints are implemented using adversarial motion prior (AMP) rewards [14], which encourage the robot to exhibit region-consistent motion styles during interception. To this end, we define position-conditioned AMP rewards by associating each region $\mathcal{R}$ with a dedicated reference motion slot and a corresponding discriminator $D^{(\mathcal{R})}$. Each discriminator is trained to distinguish between state transitions from the
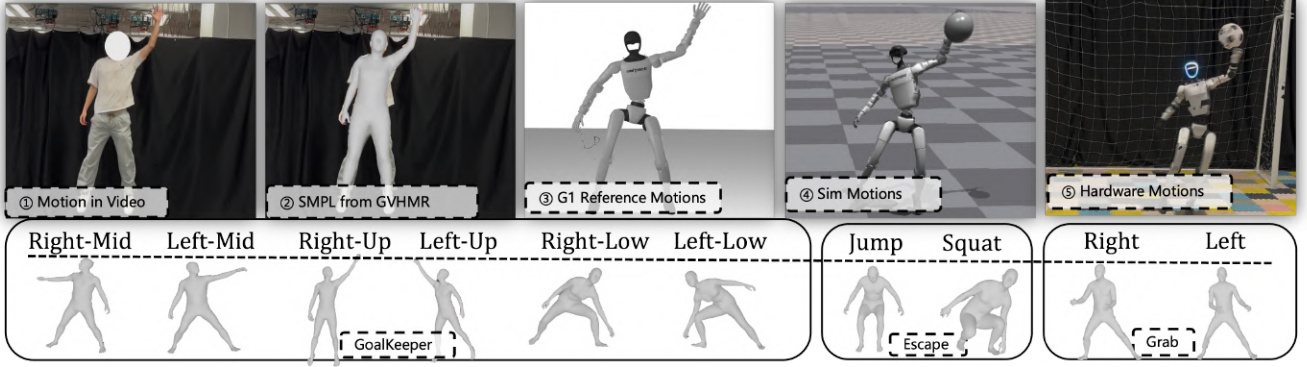
Fig. 4: The motion curation and formulation pipeline, demonstrating the process from human video to robot-executed motion.

reference motions and those generated by the policy. The training objective for $D^{(\mathcal{R})}$ is:

$$\arg\min_{D^{(\mathcal{R})}} \quad \mathbb{E}_{(q_t,q_{t+1})\sim d_M^{(\mathcal{R})}} \left[(D^{(\mathcal{R})}(q_t,q_{t+1}) - 1)^2 \right]$$
$$+ \mathbb{E}_{(q_t,q_{t+1})\sim d_\pi^{(\mathcal{R})}} \left[(D^{(\mathcal{R})}(q_t,q_{t+1}) + 1)^2 \right], \quad (5)$$

where $d_M^{(\mathcal{R})}$ and $d_\pi^{(\mathcal{R})}$ denote the distributions of joint position transitions (q_t, q_{t+1}) from expert reference motions and from the policy, respectively, conditioned on region $\mathcal{R}$.

Given the high accuracy requirements of our task, we apply motion constraints more softly to avoid conflicts with task-specific objectives. Unlike the original AMP formulation—which directly scores executed motions and rewards those that closely match the reference distribution—we generate Gaussian samples around each executed motion and reward only the one with the highest discriminator score. This encourages smooth alignment with the reference motion distribution, allowing the policy to receive similar rewards for nearby motions while letting the task objectives determine the precise execution. The AMP reward is computed as:

$$r_{\text{amp}} = \max_{j\in\{1,...,N\}} \left\{ \max\left[0, 1 - 0.25 \left(D(\tilde{q}_t^{(j)}, \tilde{q}_{t+1}^{(j)}) - 1 \right)^2 \right] \right\},$$
$$(6)$$

where each $(\tilde{q}_t^{(j)}, \tilde{q}_{t+1}^{(j)})$ is sampled from a Gaussian centered at the executed transition (q_t, q_{t+1}).

C. Aligning Sim2Real Perception

Acquiring hardware-deployable ball perception and bridging the sim-to-real gap are critical challenges in dynamic object interaction tasks. In this work, the robot receives the ball position relative to its torso frame, represented by the observation vector $O_{\text{ball}} \in \mathbb{R}^3$. We implement two hardware-based perception systems to acquire the ball position O_{ball}: 1) a motion capture (MoCap) system using markers placed on the robot's head and the ball, where O_{ball} is computed from the relative position between the two as reported by the MoCap system; 2) a depth camera (Intel RealSense D435i) mounted on the robot's head with an infrared (IR) filter [4], which detects a high-reflectivity ball and directly estimates O_{ball} from the camera's perception pipeline. While both systems provide reasonably accurate ball

position estimates, performance gaps remain due to calibration errors and perception limitations—such as occlusions in the MoCap system and field-of-view constraints in the camera-based system. To mitigate these issues, we incorporate a ball position estimator into the policy and introduce training-time noise to improve robustness across both perception modalities.

1) Ball Estimator: We integrate a ball estimator into the training loop, which is effective in identifying regions (for task and motion conditions) and accurately predicting ball positions and moving trajectories (Fig. 2), fully utilizing historical observations. The position estimator is trained with Mean Squared Error (MSE) loss to minimize the prediction error of the ball's position, while the region estimator is trained using Cross-Entropy Loss to classify the correct region where the ball is likely to land.

2) Training Noises: To simulate real-world perception uncertainties, we introduce noise into the ball observation O_{ball} during training. Specifically, we apply: 1) position noise—random perturbations of up to $5cm$, which is half the ball's $10cm$ radius—to mimic typical sensing errors; and 2) random dropouts—O_{ball} is occasionally set to zero after $0.4s$ of flight, simulating occlusions or out-of-FOV cases encountered in hardware perception when the ball approaches. Additionally, once the ball has stopped flying (e.g. after being intercepted or landing), we set $O_{\text{ball}} = 0$ to reflect the absence of further tracking. The robot is then encouraged to hold its final pose, as described in Section III-A.3.

IV. EVALUATIONS

We conduct a series of experiments to comprehensively evaluate the proposed method, focusing on the following objectives: 1) to validate the effectiveness of our method in mastering humanoid goalkeeping skills (in terms of success rate and coverage), 2) to assess the resemblance between robot-executed motions and expert demonstrations, and 3) to evaluate the benefit of the proposed position-conditioned learning scheme compared to non-division baselines. To this end, we design evaluation metrics that reflect task performance, motion fidelity, and action smoothness:

TABLE II: Simulation Results.

| Method	Task		Motion			Smoothness	
	$E_{\text{succ}}\uparrow$	$E_{\text{ee task}}\downarrow$	$E_{\text{match}(\mathcal{R})}\uparrow$	$E_{\text{pos}(min)}\downarrow$	$E_{\text{pos}(\mathcal{R})}\downarrow$	$E_{\text{dof acc}}\downarrow$	$E_{\text{smth}}\downarrow$
Ablation Motion / Task Constraints							
w.o. Task Constraints	9.44 ± 4.82	0.55 ± 0.018	64.33 ± 5.37	$\mathbf{0.76}\pm0.012$	$\mathbf{0.98}\pm0.060$	$\mathbf{14.26}\pm0.275$	$\mathbf{0.030}\pm0.001$
w.o. Motion Constraints	31.69 ± 1.18	0.98 ± 0.172	8.83 ± 4.50	1.73 ± 0.018	2.03 ± 0.025	62.23 ± 4.368	0.151 ± 0.015
Humanoid Goalkeeper (ours)	$\mathbf{80.92}\pm1.72$	$\mathbf{0.12}\pm0.004$	67.84 ± 1.49	1.24 ± 0.017	1.31 ± 0.021	22.80 ± 0.404	0.042 ± 0.001
Ablation Position Division							
w.o. AMP Division	76.11 ± 2.19	0.14 ± 0.010	25.73 ± 5.03	1.39 ± 0.007	1.51 ± 0.015	23.18 ± 0.326	0.043 ± 0.000
w.o. Task Division	62.08 ± 4.30	0.14 ± 0.005	39.18 ± 3.60	1.62 ± 0.007	1.84 ± 0.030	28.69 ± 0.329	0.054 ± 0.000
w.o. Division	64.83 ± 5.07	0.14 ± 0.004	19.09 ± 5.39	1.62 ± 0.023	1.87 ± 0.057	28.56 ± 0.450	0.052 ± 0.001
Humanoid Goalkeeper (ours)	$\mathbf{80.92}\pm1.72$	$\mathbf{0.12}\pm0.004$	67.84 ± 1.49	$\mathbf{1.24}\pm0.017$	$\mathbf{1.31}\pm0.021$	$\mathbf{22.80}\pm0.404$	$\mathbf{0.042}\pm0.001$
Evaluation Settings							
Speed-Easy $(0.7s \sim 1.0s)$	80.95 ± 4.21	0.12 ± 0.012	50.88 ± 3.28	1.26 ± 0.008	1.37 ± 0.021	23.20 ± 0.432	0.042 ± 0.001
Range-Easy $(\pm1.0m)$	$\mathbf{84.64}\pm2.75$	0.13 ± 0.011	63.16 ± 4.47	1.26 ± 0.007	1.32 ± 0.012	22.23 ± 0.244	0.041 ± 0.000
Speed-Hard $(0.4s \sim 0.7s)$	64.25 ± 7.23	0.13 ± 0.009	$\mathbf{70.47}\pm3.53$	1.26 ± 0.013	$\mathbf{1.29}\pm0.015$	$\mathbf{21.85}\pm0.324$	$\mathbf{0.041}\pm0.000$
Range-Hard $(\pm2.0m)$	63.91 ± 2.29	0.15 ± 0.007	63.45 ± 7.92	1.26 ± 0.009	1.33 ± 0.012	24.42 ± 0.449	0.044 ± 0.001
Default $(0.5s \sim 1.0s, \quad \pm1.5m)$	80.92 ± 1.72	$\mathbf{0.12}\pm0.004$	67.84 ± 1.49	$\mathbf{1.24}\pm0.017$	1.31 ± 0.021	22.80 ± 0.404	0.042 ± 0.001
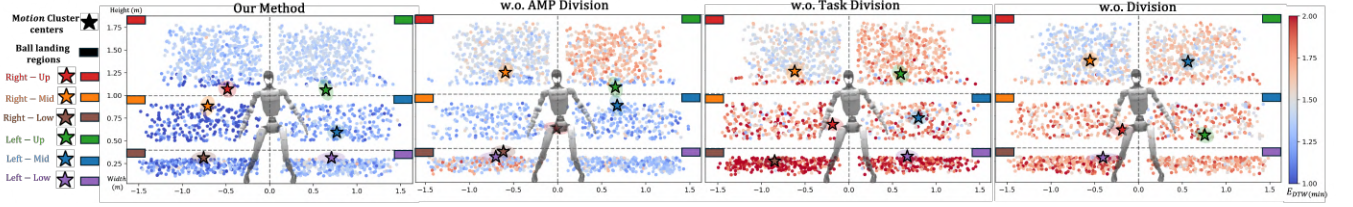
Fig. 5: Each scatter point represents a trial, with the position corresponding to the ball target of the trial. The color gradient of the points indicates $E_{\text{pos}}(\text{min})$ for each trial. The pentagram represents the cluster center of trials that has the same matching region (regions with minimal E_{pos}). A well-learned position-conditioned policy should exhibit a clear correspondence between the cluster centers and their respective ball target regions

$E_{\text{succ}}\uparrow$	Success rate of blocking the ball.
$E_{\text{ee task}}\downarrow$	Closest distance between the end effector (hand) and ball during a trial.
$E_{\text{pos}(\mathcal{R})}\downarrow$	pose error to expert motion (processed through DTW [48]) from the ball target region $\mathcal{R}$.
$E_{\text{pos}(min)}\downarrow$	Lowest pose error between executed motion and expert motion among all the regions.
$E_{\text{match}(\mathcal{R})}\uparrow$	Whether the closest-matching motion aligns with the ball target region $\mathcal{R}$.
$E_{\text{dof acc}}\downarrow$	Joint acceleration: $\frac{1}{\Delta t}|\dot{q}_t - \dot{q}_{t-1}|$.
$E_{\text{smth}}\downarrow$	Action smoothness: $|a_t - 2a_{t-1} + a_{t-2}|$.

During training and evaluation in simulation, the goal area is partitioned into six predefined ball landing regions, with corresponding expert motions visualized in Fig. 4 and the region boundaries illustrated in Fig. 5. We conduct 500 simulation trials and 5 hardware trials for each region. The ball flight duration ranges from $0.5\,\text{s}$ to $1.0\,\text{s}$, with flying distances varying from $3.0\,\text{m}$ to $5.0\,\text{m}$. The goal used in hardware evaluations corresponds to the standard dimensions of a 5-a-side soccer match, with a width of $3.0\,\text{m}$ (between the goalposts) and a height of $2.0\,\text{m}$.

A. Simulation Results

Table II summarizes the primary numerical results from our simulation experiments. The proposed method achieves the highest task success rate while also demonstrating superior motion resemblance compared to all baselines. These outcomes support our core hypothesis: aligning task constraints with appropriate motion constraints yields better overall performance. Notably, the *w.o. AMP Division* variant—which lacks explicit motion segmentation but retains region-based task constraints—achieves moderately competitive success and motion metrics. This result suggests that correct task conditioning can guide action selection to some extent. However, its performance still lags behind the full method, highlighting the importance of explicit motion correspondence in achieving high-quality, task-aligned behaviors.

Given the default keeping ranges of $(\pm1.5\,\text{m})$ in width and $(1.8\,\text{m})$ in height, the proposed method achieves a notably high success rate—despite operating over one of the widest goalkeeping areas reported in the robotic goalkeeping literature. To further assess robustness, we evaluate the method under varying ball speeds and expanded guarding regions. As expected, the results indicate a gradual decline in performance as task difficulty increases, validating the method's scalability while revealing its current limitations under more challenging conditions.

The smoothness metric serves as a useful indicator for assessing how well the alignment between motion and task constraints translates into natural robot behavior. Our results support this connection: our policy that guided by accurate expert motions achieve better smoothness. Interestingly, the baseline without explicit task constraints—relying solely on AMP and regularization rewards—achieves the best

TABLE III: Hardware Results.

Method	Right-Low			Right-Mid			Right-Up		
	$E_{\text{succ}}\uparrow$	$E_{\text{pos}(min)}$	$E_{\text{dof acc}}\downarrow$	E_{succ}	$E_{\text{pos}(min)}$	$E_{\text{dof acc}}$	E_{succ}	$E_{\text{pos}(min)}$	$E_{\text{dof acc}}$
Humanoid Goalkeeper (camera)	1/5	1.67	**21.79**	3/5	1.46	21.30	**3/5**	1.56	32.99
w.o. AMP Division (mocap)	0/5	**1.43**	31.02	3/5	1.36	**20.62**	0/5	1.47	31.33
w.o. Division (mocap)	1/5	1.66	36.13	0/5	1.46	31.09	2/5	1.57	34.47
Humanoid Goalkeeper (mocap)	**5/5**	1.49	33.34	**4/5**	**1.28**	24.56	1/5	**1.43**	**29.43**

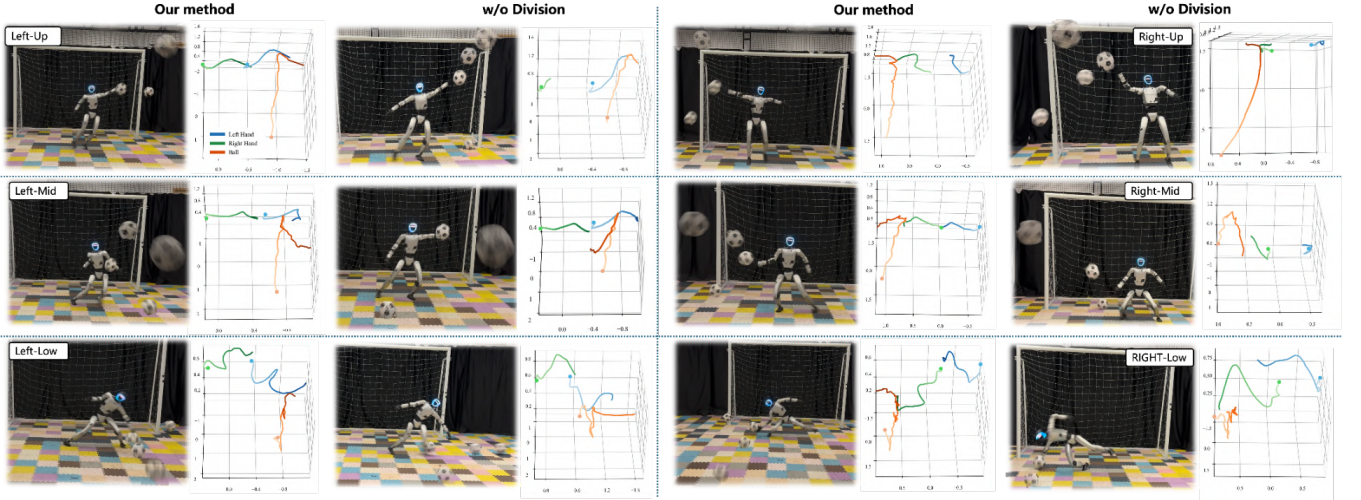| Method	Left-Low			Left-Mid			Left-Up		
	$E_{\text{succ}}\uparrow$	$E_{\text{pos}(min)}$	$E_{\text{dof acc}}\downarrow$	E_{succ}	$E_{\text{pos}(min)}$	$E_{\text{dof acc}}$	E_{succ}	$E_{\text{pos}(min)}$	$E_{\text{dof acc}}$
Humanoid Goalkeeper (camera)	1/5	1.70	**24.04**	3/5	1.53	21.71	3/5	1.58	28.12
w.o. AMP Division (mocap)	5/5	**1.48**	35.25	2/5	1.54	**18.60**	3/5	1.47	27.68
w.o. Division (mocap)	3/5	1.48	29.88	3/5	**1.43**	23.77	3/5	1.68	38.87
Humanoid Goalkeeper (mocap)	5/5	1.68	**28.76**	**3/5**	1.44	22.08	3/5	**1.40**	**27.16**

Fig. 6: We visualize the 3D trajectory of the ball, as well as the left and right hand trajectories for one trial in each region, and compare the performance of the proposed method with that of the non-division baselines.

smoothness performance. This suggests that, in the absence of region-based goals, the robot may prioritize generic motion regularity over task effectiveness, resulting in smoother but less purposeful behaviors.

Fig. 5 provides a region-based visualization of motion behavior to validate our approach. In each subplot, the pentagram represents the cluster center of robot-performed motions that best match the specific expert motion. The results show that our method learns region-conditioned motion effectively, with distinct and consistent clusters emerging for each region. In contrast, baseline methods tend to mix motions across regions, resulting in less coherent behaviors and higher joint position errors. This highlights the effectiveness of our region-conditioned policy learning in preserving motion structure aligned with task semantics.

B. Region-Specific Hardware Evaluation

Table III presents region-specific hardware evaluation results on the goalkeeping task. The proposed method achieves a total success rate of 21/30, with most failure cases concentrated in the upper regions. These upper regions demand stricter alignment between the end-effector and the ball trajectory, as well as wide-range, whole-body motions—posing greater challenges compared to the lower and side regions.

Despite these challenges, our method exhibits the smallest sim-to-real performance drop among all compared methods. In contrast, non-division baselines tend to overfit to one side of the goal and underperform in other regions. Specifically, their performance on the left side consistently exceeds that on the right, indicating a lack of generalization across the full goal area. We further illustrate this phenomenon in Fig. 6, which visualizes the ball trajectory and end-effector motion (recorded via MoCap). The visualization shows that, for balls targeting the right-side regions, the robot controlled by a non-division baseline fails to initiate an early lateral step, resulting in missed saves due to the short response window. Conversely, our method accurately guides the end-effector to the appropriate region from the onset, resulting in successful interception. This highlights the effectiveness of region-aware motion learning in achieving balanced, robust performance across the entire goal.
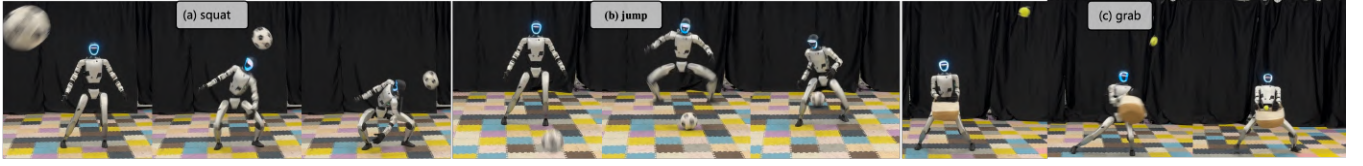
Fig. 7: We demonstrate the generalization ability of the extended tasks, including ball grabbing and ball escaping.
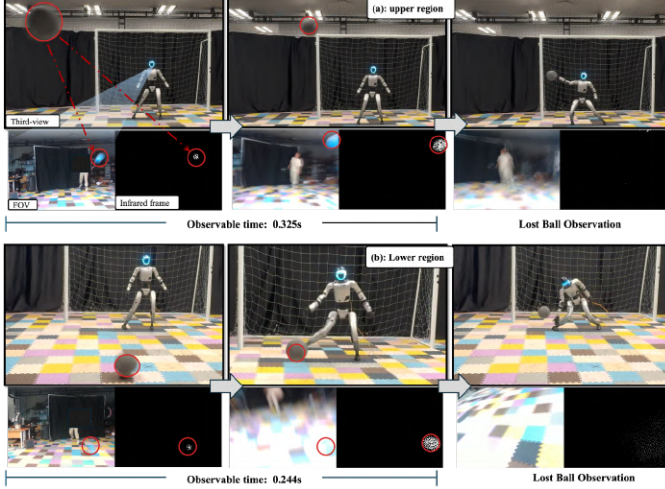

Fig. 8: the first-view illustration of on-board camera perception.

Fig. 9: Snapshots of the robot performing continuous goalkeeping without resetting to the default posture.

entering a recovery phase. This capability highlights the potential of our method for real-world scenarios that demand uninterrupted, match-like behavior.

E. Generalization Tasks

Centered on the goalkeeper task, we introduce an effective framework that enables humanoid robots to perform whole-body interactions with dynamic objects by leveraging learned motion priors. To demonstrate the generalization capability of our framework beyond goalkeeping, we extend it to two additional tasks that also demand highly dynamic responses: (1) escaping from an incoming flying ball, and (2) grabbing a moving ball using a soft bag.

These motions are developed using the same position-conditioned approach employed in goalkeeping. Specifically, we partition the input space based on the ball's spatial characteristics: high vs. low positions correspond to jump and squat escapes, respectively; left vs. right positions correspond to left- and right-stepped ball grabs. We show this region-based division ensures appropriate motion selection under varying conditions. Table IV presents the quantitative results of our method on these tasks.

TABLE IV: Escape Task Evaluation.

Task	$E_{\text{succ}}\uparrow$	$E_{\text{pos}(min)}$	$E_{\text{dof acc}}\downarrow$
Jump Escape	3/5	1.49	46.79
Squat Escape	5/5	1.12	16.19
Grab Tennis	1/5	0.99	18.25

In the escape task, the robot achieves a high success rate and closely replicates expert motion Fig. 7, validating the effectiveness of our generalzaition ability. The robot is capable of reacting to an incoming ball and executing a successful escape maneuver within 0.5 seconds. Meanwhile, jumping from a stationary stance constitutes a highly dynamic maneuver, leading to substantial joint accelerations. The grabbing task exhibits a lower success rate, due to the absence

C. Performance with Camera perception

The performance under camera-based perception drops to 14/30 successful trials, compared to the MoCap setup. The primary limitation stems from the narrow camera field of view (FoV), which reduces the observable time window of the ball. As shown in Fig. 8, the high-reflectivity ball is easily detectable through the camera's infrared channel at a distance. However, as the ball approaches the robot, it quickly disappears from view, leaving less than 0.3 seconds of observable time for lower regions. This limited visibility contributes to the decline in success rate due to the lack of consistent ball tracking.

Despite the FoV constraints, the camera-based system still achieves high success rates in the upper regions, where the ball remains visible for longer periods. This result demonstrates the robustness and adaptability of the proposed method when relying solely on onboard perception. Furthermore, the perceptual constraint naturally suppresses large, abrupt motions (such as lateral jumps), resulting in smoother behaviors, as reflected by improved $E_{\text{dof acc}}$ scores compared to the MoCap-based evaluations.

D. Perform Continuous Keeping

Fig. 9 demonstrates that the proposed pipeline enables the robot to continuously perform goalkeeping skills without requiring a manual reset to default states between trials. This is evidenced by two consecutive successful interceptions from both sides of the goal. As shown by the timestamps on the snapshots, the robot transitions directly from the terminal posture of the first trial to the initiation of the second, without
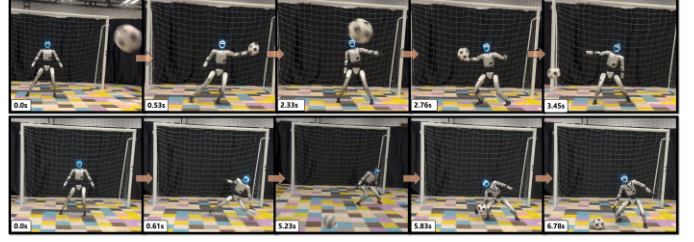
of bag during training. In simulation, a successful grab is defined as the ball falling between the robot's two hands, implicitly assuming a consistently open and rigid bag structure. However, in real-world trials, the deformability and varying orientation of the soft bag's mouth frequently violate this assumption, as illustrated in Fig. 10. Despite these difficulties, we observe that the robot reliably moves toward the ball's predicted landing location.

Fig. 10: The uncontrollable nature of the bag contributes to the low grabbing success rate.

V. DISCUSSION

We have developed a humanoid robotic goalkeeper capable of executing agile, human-like motions to intercept flying balls, as well as performing tasks such as escaping a ball using jump and squat movements. Our results show that explicitly aligning task and motion constraints based on task observations significantly improves both motion quality and task completion. Extensive experimental evaluations, along with alternative perception modules, provide strong evidence supporting the effectiveness of our method.

However, as the first humanoid-level autonomous goalkeeper, there is still room for improvement before achieving professional-level goalkeeping performance. Specifically:

(1) the lack of global observations and a high-level planning module limits our framework's ability to perform consecutive goalkeeping actions at the match level.

(2) the perception module requires further refinement to react appropriately to complex situations in real-world scenarios. Currently, the robot responds to any ball detected within its field of view, even when the ball is not targeting the goal, which is not valid in real matches.

Future work could focus on adding a high-level planner to enable continuous goalkeeping actions, such as repositioning after each trial. Additionally, improving the perception system by integrating active sensing capabilities to overcome camera field-of-view (FOV) limitations and eliminating the reliance on fixed motion-capture systems will be crucial for enhancing robustness and flexibility in real-world settings. Furthermore, to advance the robot's current motions, we believe training it to grab the ball with its hand, rather than simply punching it away, would present an appealing challenge and a valuable demonstration for humanoid-object interaction skills.

VI. ACKNOWLEDGMENT

REFERENCES

[1] T. Haarnoja et al., "Learning agile soccer skills for a bipedal robot with deep reinforcement learning," *Science Robotics*, 2023.

[2] M. Beukman, B. Ingram, G. N. Tasse, B. Rosman, and P. Ranchod, "Robocupgym: A challenging continuous control benchmark in robocup," *ArXiv*, 2024.

[3] D. Tirumala et al., "Learning robot soccer from egocentric vision with deep reinforcement learning," in *Conference on Robot Learning*, 2024.

[4] "Toward real-world cooperative and competitive soccer with quadrupedal robot teams," *Arxiv*, 2025.

[5] G. I. Fernandez, R. Hou, A. Xu, C. Togashi, and D. W. Hong, "Clap: Clustering to localize across n possibilities, a simple, robust geometric approach in the presence of symmetries," *ArXiv*, 2025.

[6] Y. Ji, G. Margolis, and P. Agrawal, "Dribblebot: Dynamic legged manipulation in the wild," in *International Conference on Robotics and Automation (ICRA)*, 2023.

[7] Z. Wang, J. Zhou, and Q. Wu, "Dribble master: Learning agile humanoid dribbling through legged locomotion," *Arxiv*, 2025.

[8] Y. Ji et al., "Hierarchical reinforcement learning for precise soccer shooting skills using a quadrupedal robot," 2022.

[9] D. Marew, N. Perera, S. Yu, S. Roelker, and D. Kim, "A biomechanics-inspired approach to soccer kicking for humanoid robots," in *IEEE-RAS 23rd International Conference on Humanoid Robots (Humanoids)*, 2024.

[10] B. Robotics, *Booster t1: Humanoid robot*, Accessed: 2025-09-15, 2025.

[11] Y. Ze et al., "Twist: Teleoperated whole-body imitation system," *ArXiv*, 2025.

[12] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *Arxiv*, 2017.

[13] N. Rudin, D. Hoeller, P. Reist, and M. Hutter, "Learning to walk in minutes using massively parallel deep reinforcement learning," in *Conference on Robot Learning*, 2022.

[14] X. B. Peng, Z. Ma, P. Abbeel, S. Levine, and A. Kanazawa, "Amp: Adversarial motion priors for stylized physics-based character control," *ACM Transactions on Graphics (ToG)*, 2021.

[15] A. Tang et al., "Humanmimic: Learning natural locomotion and transitions for humanoid robot via wasserstein adversarial imitation," in *International Conference on Robotics and Automation (ICRA)*, 2024.

[16] Z. Shen et al., "World-grounded human motion recovery via gravity-view coordinates," in *SIGGRAPH Asia 2024 Conference Papers*, 2024.

[17] Y. Zhang et al., "Humanmm: Global human motion recovery from multi-shot videos," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025.

[18] Z. Shen et al., "World-grounded human motion recovery via gravity-view coordinates," 2024.

[19] Y. Ma, A. Cramariuc, F. Farshidian, and M. Hutter, "Learning coordinated badminton skills for legged manipulators," *Science Robotics*, 2025.

[20] H. Wang et al., "Integrating learning-based manipulation and physics-based locomotion for whole-body badminton robot control," *International Conference on Robotics and Automation (ICRA)*, 2025.

[21] Z. Su et al., "Hitter: A humanoid table tennis robot via hierarchical planning and learning," *Arxiv*, 2025.

[22] D. B. D'Ambrosio et al., "Achieving human level competitive robot table tennis," in *International Conference on Robotics and Automation (ICRA)*, 2025.

[23] K. M. Lee et al., "Learning diverse robot striking motions with diffusion models and kinematically constrained gradient guidance," *International Conference on Robotics and Automation (ICRA)*, 2025.

[24] M. Hu et al., "Towards versatile humanoid table tennis: Unified reinforcement learning with prediction augmentation," *ArXiv*, 2025.

[25] D. Nguyen, Z. Zaidi, K. Karol, J. Hodgins, and Z. Xie, "Whole body model predictive control for spin-aware quadrupedal table tennis," *ArXiv*, 2025.

[26] Y. Zhang, T. Liang, Z. Chen, Y. Ze, and H. Xu, "Catch it! learning to catch in flight with mobile dexterous hands," in *International Conference on Robotics and Automation (ICRA)*, 2025.

[27] S. Abeyruwan et al., "Agile catching with whole-body mpc and blackbox policy learning," in *Conference on Learning for Dynamics & Control*, 2023.

[28] Y. Ma, Y. Liu, K. Qu, and M. Hutter, "Learning accurate whole-body throwing with high-frequency residual policy and pullback tube acceleration," *ArXiv*, 2025.

[29] X. Huang et al., "Creating a dynamic quadrupedal robotic goalkeeper with reinforcement learning," in *International Conference on Intelligent Robots and Systems (IROS)*, 2023.

[30] R. Hou, M. Zhu, H. Nam, G. I. Fernandez, and D. W. Hong, "Fast and robust localization for humanoid soccer robot via iterative landmark matching," *Arxiv*, 2025.

[31] Z. Xu et al., "Volleybots: A testbed for multi-drone volleyball game combining motion control and strategic play," *ArXiv*, 2025.

[32] Z. Chen, M. Ji, X. Cheng, X. Peng, X. B. Peng, and X. Wang, "Gmt: General motion tracking for humanoid whole-body control," *ArXiv*, 2025.

[33] T. He et al., "Asap: Aligning simulation and real-world physics for learning agile humanoid whole-body skills," in *Robotics: Science and Systems*, 2025.

[34] Q. Liao, T. E. Truong, X. Huang, G. Tevet, K. Sreenath, and C. K. Liu, "Beyondmimic: From motion tracking to versatile humanoid control via guided diffusion," *ArXiv*, 2025.

[35] A. Allshire et al., "Visual imitation enables contextual humanoid control," *ArXiv*, 2025.

[36] T. He et al., "Omnih2o: Universal and dexterous human-to-humanoid whole-body teleoperation and learning," in *Conference on Robot Learning*, 2024.

[37] W. Xie et al., "Kungfubot: Physics-based humanoid whole-body control for learning highly-dynamic skills," *ArXiv*, 2025.

[38] M. Ji et al., "Exbody2: Advanced expressive humanoid whole-body control," *ArXiv*, 2024.

[39] Y. Liu, B. Yang, L. Zhong, H. Wang, and L. Yi, "Mimicking-bench: A benchmark for generalizable humanoid-scene interaction learning via human mimicking," *ArXiv*, 2024.

[40] H. Xue et al., "Leverb: Humanoid whole-body control with latent vision-language instruction," *ArXiv*, 2025.

[41] L. Yang et al., "Omniretarget: Interaction-preserving data generation for humanoid whole-body loco-manipulation and scene interaction," *ArXiv*, 2025.

[42] S. Yin, Y. Ze, H.-X. Yu, C. K. Liu, and J. Wu, "Visualmimic: Visual humanoid loco-manipulation via motion tracking and generation," *ArXiv*, 2025.

[43] H. Weng et al., "Hdmi: Learning interactive humanoid whole-body control from human videos," *ArXiv*, 2025.

[44] S. Lin, G. Qiao, Y. Tai, A. Li, K. Jia, and G. Liu, "Hwc-loco: A hierarchical whole-body control approach to robust humanoid locomotion," *ArXiv*, 2025.

[45] L. Ma et al., "Styleloco: Generative adversarial distillation for natural humanoid robot locomotion," *ArXiv*, 2025.

[46] L. Pan et al., "Tokenhsi: Unified synthesis of physical human-scene interactions through task tokenization," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025.

[47] Y. Wang et al., "Skillmimic: Learning basketball interaction skills from demonstrations," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025.

[48] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE transactions on acoustics, speech, and signal processing*, 2003.