Unified Perceptive Fall Safety for Humanoids

Osher Azulay, Zhengjie Xu, Andrew Scheffer, and Stella X. Yu

***Abstract*—Reliable fall recovery is critical for humanoids operating in cluttered environments. Unlike quadrupeds or wheeled robots, humanoids undergo high-energy impacts, complex whole-body contact, and large viewpoint changes during falls, making recovery essential. Prior work treats fall safety as separate problems (avoidance, impact mitigation, stand-up) or learns end-to-end policies without vision, typically on flat terrain. More fundamentally, it couples pose, dynamics, and terrain into a single data space, requiring exhaustive coverage and limiting generalization. We present a unified visually grounded fall-safety approach spanning all phases of recovery. It builds on two insights: (1) human fall-recovery poses are structured and transferable across terrains via alignment, and (2) effective recovery depends on interpreting the recovery goal in relation to the surrounding terrain. We train a privileged teacher using sparse human demonstrations and simulated terrains, and distill it into a deployable student that relies on egocentric depth and proprioception. The student predicts a goal-in-context latent representing the recovery target in relation to its surroundings, guiding coordinated whole-body reactions. Experiments in simulation and on a Unitree G1 demonstrate robust zero-shot fall safety across diverse terrains.**

***Index Terms*—Humanoid Robot Systems, Whole-Body Motion Planning and Control, Visual Learning.**

I. Introduction

HUMANOIDS are intended to operate in the same cluttered and uneven environments as humans, yet they remain vulnerable when failures occur. Minor disturbances can quickly cascade into coupled rotations, impacts, and contact transitions, unfolding faster than conventional stabilization can correct. Humanoid fall safety demands whole-body coordination under severe time pressure, as early actions constrain later contacts and recovery.

Fall safety cannot be reduced to a post-impact reflex: How a robot falls determines which supports and contacts remain feasible, and whether standing up is possible at all. These choices depend on situational and terrain awareness, since contact feasibility and momentum redirection are shaped by local geometry invisible to proprioception alone. Fall safety thus requires the integration of terrain perception and coordinated whole-body control across the fall-recovery process. Despite the intrinsic coupling between fall and recovery, existing approaches fragment fall safety along two axes and typically rely on blind, proprioceptive-only sensing.

First, many methods decompose fall-related behaviour into isolated subproblems: fall avoidance (preventing loss of balance), impact mitigation (controlling descent and contact to reduce impact), or stand-up recovery (restoring stable standing). Balance controllers focus on preventing loss of stability [1], [2], while stand-up controllers are typically invoked only after the robot reaches rest in a small set of predefined poses [3]. These components are designed largely in isolation and rely primarily on proprioceptive sensing, implicitly assuming flat ground and benign contact geometry. Recovery therefore cannot be separated from *how the robot falls*, nor from *where and how contact occurs*.

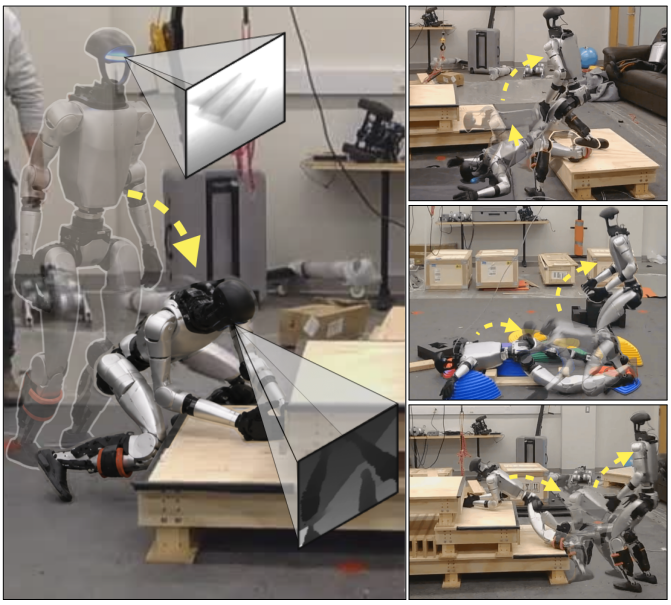


Fig. 1. **Vision-enabled unified fall safety for humanoids**. Left: fall-mitigation recovery after a push towards stairs demands vision-guided anticipatory arm contact to arrest the fall. Right: recovery behaviors across diverse terrain and initial conditions highlight the generality of vision-guided contact selection and whole-body rebalancing.

Second, learning-based approaches that aim to address fall-related behaviors end-to-end often treat the problem as one of monolithic data complexity, and are typically trained without access to visual terrain perception. Reinforcement learning (RL) methods usually require extensive reward engineering, frequently leading to unnatural motions [4]. Imitation learning methods rely on dense trajectory demonstrations, often sourced from internet-scale human motion data, which transfer poorly across terrain geometry and contact conditions [5]–[7]. Both families of methods often treat environmental complexity as a monolithic data problem, entangling kinematics, dynamics, and terrain and requiring exhaustive coverage of their combinations.

In contrast, we adopt a factorized view of data complexity (Fig. 2): Our first insight is that natural human fall and recovery poses are far more constrained than they appear: Poses required on complex terrain often admit spatially aligned

Manuscript received May 5, 2026; revised October 2, 2026; accepted October 13, 2026.

This paper was recommended for publication by Editor Olivier Stasse upon evaluation of the Associate Editor and Reviewers' comments.

All authors are with the University of Michigan, Ann Arbor, MI, USA. Corresponding Author: osher@umich.edu.

Project page: vigor2026.github.io.

Digital Object Identifier (DOI): see top of this page.

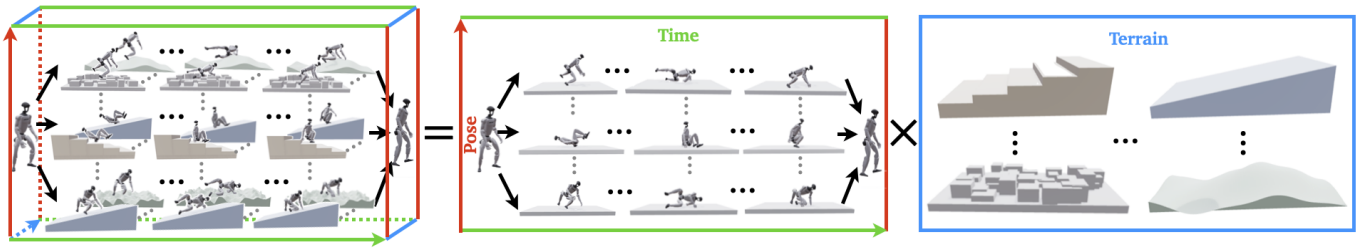


Fig. 2. **Factorized data generation yields sample-efficient imitation and scalable adaptation for humanoid fall safety learning.** Rather than treating pose, time, and terrain as a single monolithic data space requiring exhaustive coverage, we factorize it into a small set of human pose trajectories from real-world demonstrations on flat terrain and independently varying terrain geometry in simulation.

and physically compatible counterparts on flat terrain. This observation allows pose and terrain variation to be treated as largely independent factors. As a result, a few human demonstrations collected on flat ground can serve as priors on kinematics and dynamics, while variation in terrain geometry, contact timing, and momentum redirection is handled through RL via interaction with the environment.

Our second insight is that recovery goals must be understood in relation to the terrain on which they are realized. A motion prior provides whole-body coordination, while the current body state and local geometry determine which contacts can support that recovery.

We therefore build the policy directly on a compact visual goal-in-context latent representation, which integrates the next target pose with the local terrain and body context in a single perceptual-motor space.

Based on these insights, we introduce VIGOR (*Visual Goal-In-Context Inference*), a perceptive control framework that unifies fall avoidance, impact mitigation, and stand-up recovery. A privileged teacher policy is trained via RL using sparse human demonstrations on flat terrain together with access to local terrain geometry, yielding terrain-aware reactive behaviours. This knowledge is then distilled into a deployable student policy that operates using only egocentric depth observations and short-term proprioceptive history. Latent supervision guides the student's interpretation of its recovery target, while action imitation guides the resulting control.

Our contributions are threefold:

- We formulate humanoid fall safety as a unified, *visually grounded* control problem spanning fall mitigation, impact handling, and stand-up recovery within a single policy. To our knowledge, this is the first learning-based framework for unified perceptive humanoid recovery.
- We introduce a factorized data formulation that decouples human fall-recovery motion from terrain variation, enabling sample-efficient learning from a small set of flat-ground human demonstrations combined with large-scale terrain randomization.
- We introduce a visual recovery policy trained through intermediate latent supervision alongside action imitation. A privileged teacher guides the student to infer recovery targets suited to the surrounding terrain, using only egocentric depth and proprioceptive history.

We evaluate VIGOR extensively in simulation and demonstrate zero-shot transfer to a real humanoid across diverse fall and recovery conditions, validating a unified vision-enabled approach to humanoid fall safety.

II. Related Work

A. Humanoid Fall Mitigation and Standing-Up Control

Humanoid fall safety requires managing impacts and restoring stable standing. Prior work typically decomposes recovery into two disjoint stages: fall mitigation during descent and standing up after impact. Classical fall mitigation methods focus on maintaining balance or redirecting momentum to reduce impact forces [1], [2], while learning-based approaches similarly emphasize safe landing or damage reduction without explicitly addressing subsequent recovery [8]. Stand-up controllers typically act after the robot has settled into predefined configurations [9], although recent learning-based methods extend recovery to more diverse initial poses [4], [10]. Only a few works attempt to unify fall mitigation and standing-up within a single policy [11], [12], but these approaches typically rely on limited sensing or blind operation. In contrast, we learn a *unified, visually grounded* policy that explicitly couples fall mitigation and recovery under diverse terrain conditions.

B. Visual Whole-Body Control

Methods in vision-based control have recently enabled humanoids to perform complex whole-body behaviors under partial observability [13]. Visual input has been used to support loco-manipulation and contact-rich interaction [14]–[16], as well as to guide navigation and locomotion across challenging environments [17]–[19]. Visual reinforcement learning has also produced active perception during whole-body manipulation in simulated humanoids [20]. Despite this progress, fall recovery remains perceptually challenging: rapid body rotations, self-occlusion, and intermittent ground contact produce unstable visual observations. Visually grounded fall recovery is still largely unexplored, especially for policies spanning fall mitigation, impact, and stand-up recovery.

C. Motion Priors and Style-Constrained RL

Learning whole-body control requires coordinating many joints through changing contacts. Prior work therefore relies on carefully shaped multi-term reward functions [13], [21] or constrains the control problem through reduced body representations or limited link actuation [18]. Human motion

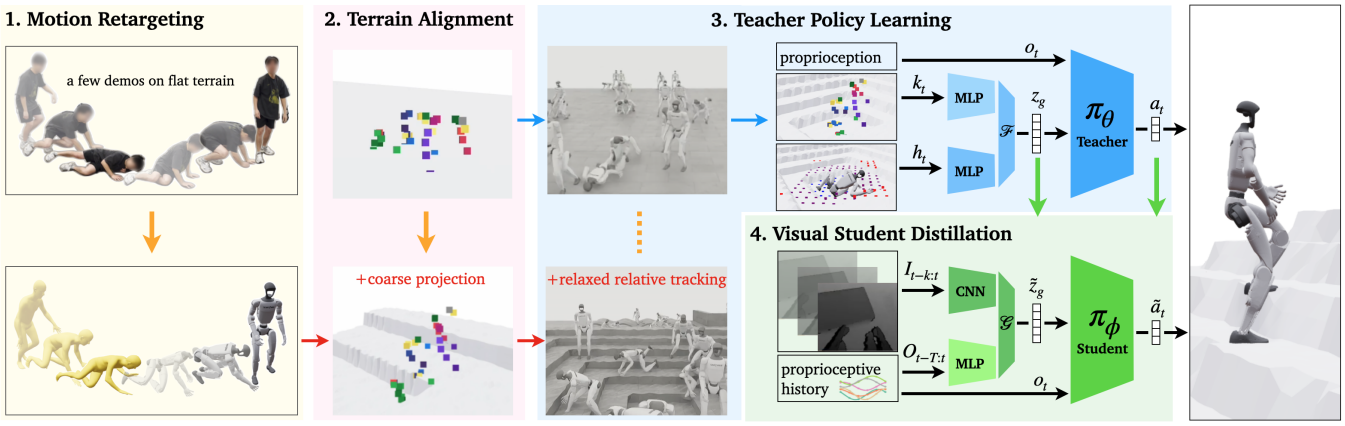


Fig. 3. **Overview of VIGOR. 1)** Human fall–recovery demonstrations collected on flat ground are kinematically retargeted to the robot. **2)** Reference poses are used directly on flat terrain and coarsely projected onto uneven terrain to provide sparse tracking targets. **3)** A privileged RL teacher learns a goal-in-context latent representing the immediate recovery target in relation to local terrain. **4)** A student learns through latent supervision and action imitation, using only egocentric depth and proprioceptive history at deployment.

priors provide an alternative form of regularization, inducing behavior with minimal reward engineering.

Reference-guided RL methods such as DeepMimic [22] and AMP [23] demonstrate strong whole-body control on flat terrain, but often rely on dense trajectory tracking or periodic motion assumptions, which become limiting in highly non-periodic, contact-rich behaviors such as fall recovery. Recent work relaxes dense tracking through higher-level structural guidance, including keyframe-based objectives [5], style-constrained learning [6], or hybrid formulations that combine flat-terrain imitation with residual adaptation on uneven terrain [7]. These methods allow deviation from demonstrations when needed, but still depend on sufficient motion–terrain coverage. We therefore treat human fall-recovery demonstrations as *sparse structural priors* rather than dense trajectory targets.

III. Method

We propose VIGOR, a unified framework for humanoid stand-up, fall mitigation, and recovery in unstructured environments, as illustrated in Fig. 3. Unlike pipelines that treat falling, impact mitigation, and standing as separate modules, VIGOR learns the full process with a single control policy conditioned on onboard egocentric sensing.

Following a destabilizing disturbance, the robot may enter arbitrary configurations involving high-energy impacts, whole-body contact, and large changes in body orientation. At each timestep t, the robot receives proprioception $\mathbf{o}_t = [\boldsymbol{\omega}_t, \mathbf{g}_t, \mathbf{q}_t, \dot{\mathbf{q}}_t, \mathbf{a}_{t-1}]$, where $\boldsymbol{\omega}_t$ is the base angular velocity, $\mathbf{g}_t$ is projected gravity, $\mathbf{q}_t$ and $\dot{\mathbf{q}}_t$ are joint positions and velocities, and $\mathbf{a}_{t-1}$ is the previous action. The policy outputs joint-space PD targets $\mathbf{a}_t$ for the 23-DoF humanoid.

To learn recovery under these conditions, we adopt a privileged teacher–student design. During training, the teacher π_θ observes the current proprioception $\mathbf{o}_t$, sparse keyframe targets $\mathbf{k}_t$, and local terrain samples $\mathbf{h}_t$, which provide structured recovery intent and terrain context. The deployable student π_ϕ does not access these privileged inputs. Instead, it receives proprioceptive history $\mathbf{o}_{t-T:t}$ and stacked egocentric depth $\mathbf{I}_{t-k:t}$ from a head-mounted camera, and learns to reconstruct the teacher's *visual goal-in-context* latent. This enables unified vision-based fall mitigation and recovery from onboard sensing.

A. Motion–Terrain Factorization

Motion–terrain factorization removes the need for terrain-matched demonstrations: flat-ground motions supply recovery structure, while RL learns terrain-specific contacts and dynamics. We obtain fall–recovery demonstrations from monocular human videos recorded on flat ground, covering prone, forward, sideways, and backward recoveries. We process the videos with VideoMimic [24] and retarget the reconstructed motions to the Unitree G1 using kinematics-aware mapping with conservative pelvis and hip constraints to reduce artifacts. We retain nine high-quality recovery sequences from three body sizes. Rather than tracking full trajectories, we uniformly subsample sparse keyframes at $\sim$5 Hz. These keyframes provide coarse structural priors for the privileged teacher while allowing RL to adapt contact timing, momentum redirection, and terrain-dependent recovery. During training, keyframes are coarsely aligned to terrain by a vertical shift:

$$\Delta z = \max_{i=1,\dots,N_{\text{links}}} \left(h(\mathbf{p}_i^{\text{ref}}) - z_i^{\text{ref}} \right),$$

where $h(\cdot)$ is the terrain height at link position $\mathbf{p}_i^{\text{ref}}$. All reference poses are shifted by Δz to maintain terrain clearance and are initialized near the terrain center with small x–y perturbations. To increase coverage, we randomize the reference trajectory, temporal phase, global transformations, and occasionally add a short free-fall interval by suppressing a random subset of joint torques before control begins.

B. Privileged Goal-in-Context Teacher

We train a privileged fall-mitigation and recovery teacher policy π_θ using PPO [25]. Training proceeds in two stages: first on flat terrain to acquire basic fall mitigation and stand-up behaviors, and then on randomized non-flat terrains to handle diverse contact geometry and perturbations. The teacher receives proprioception $\mathbf{o}_t$, sparse keyframe targets $\mathbf{k}_t$, and privileged terrain samples $\mathbf{h}_t$. The reference and terrain signals

are fused into a single *goal-in-context* latent $\mathbf{z}_g = \mathcal{F}(\mathbf{k}_t, \mathbf{h}_t)$, which encodes the immediate recovery target conditioned on local terrain context. We normalize the projected latent as $\mathbf{z}_g = \sqrt{d}\,\bar{\mathbf{z}}_g/\|\bar{\mathbf{z}}_g\|_2$ to maintain a consistent scale for policy conditioning and distillation. The teacher actor conditions on this latent and proprioception to produce joint targets: $\mathbf{a}_t = \pi_\theta(\mathbf{z}_g, \mathbf{o}_t)$. While the teacher is trained in a tracking-based manner similar to prior motion imitation policies [22], the reference signals are intentionally sparse and enforced only at a coarse, relative level. Together with privileged terrain observations, this under-specification allows RL to resolve contact timing, body placement, and terrain-dependent execution details beyond what is prescribed by the demonstrations.

C. Reward Design

We design a composite reward function to support fall mitigation, stabilization, and recovery within a single policy:

$$r_t = r_t^{\text{imit}} + r_t^{\text{reg}} + r_t^{\text{post}}$$

where r_t^{imit} is a DeepMimic-style tracking term [22], r_t^{reg} aggregates standard motion-regularization penalties, and r_t^{post} provides post-recovery stabilization by rewarding upright hold and suppressing residual motion once standing. All reward terms are summarized in Table I.

Motion imitation. We project reference poses, recorded on flat ground, onto the simulated terrain before computing tracking errors, as described in Sec. III-A. This preserves the demonstrated motion structure but is not physically exact on uneven terrain, leaving small residual offsets. Tracking full world-frame positions $\|\mathbf{p}_{i,t}^{\text{ref}} - \mathbf{p}_{i,t}\|$ would therefore penalize projection residuals and terrain-dependent deviations needed for recovery. Instead, we define imitation targets in a root-relative canonical frame by comparing $(\mathbf{p}_{i,t}^{\text{ref}} - \mathbf{p}_{0,t}^{\text{ref}}) - (\mathbf{p}_{i,t} - \mathbf{p}_{0,t})$, preserving relative body configuration while remaining insensitive to projection artifacts, allowing demonstrations to act as structural priors rather than exact pose targets.

Regularization and safety. The term r_t^{reg} aggregates standard regularization penalties, including joint limit violations, joint velocities, accelerations, momentum change, action smoothness, undesired contacts, and contacts near terrain edges. These terms stabilize learning and reduce unsafe behaviors without prescribing a specific recovery strategy.

Post-recovery stabilization. Once the robot reaches an upright configuration, r_t^{post} encourages stable upright posture by suppressing residual base motion and joint velocities.

D. Egocentric Student Policy

At deployment time, the robot relies only on onboard sensing. The student therefore learns to infer the teacher's goal-in-context latent from egocentric depth and short-term proprioceptive history. The same sparse demonstrations used to train the teacher define the space of recovery behaviors available to the student, which infers an appropriate recovery target without explicit reference signals.

The student receives egocentric depth $\mathbf{I}_{t-k:t}$ and proprioceptive history $\mathbf{o}_{t-T:t}$. These inputs are fused into a predicted *visual* goal-in-context latent $\tilde{\mathbf{z}}_g = \mathcal{G}(\mathbf{I}_{t-k:t}, \mathbf{o}_{t-T:t})$, which is designed to capture the terrain-conditioned recovery objective. To infer this latent, the student encodes depth with a CNN and proprioceptive history with a temporal encoder. The resulting features are concatenated and projected through an MLP, then normalized as in the teacher. The student actor conditions on this latent and current proprioception to produce joint targets: $\tilde{\mathbf{a}}_t = \pi_\phi(\tilde{\mathbf{z}}_g, \mathbf{o}_t)$.

TABLE I
REWARD TERMS USED FOR FALL-RECOVERY LEARNING.

Reward	Definition	Weight
Motion imitation		
Pos. track.	$\exp(-\|\Delta x\|^2/\sigma)$	1.25
Rot. track.	$\exp(-\|\Delta\theta\|^2/\sigma)$	0.5
Lin. vel. track.	$\exp(-\|\Delta v\|^2/\sigma)$	0.125
Ang. vel. track.	$\exp(-\|\Delta\omega\|^2/\sigma)$	0.125
Joint pos. track.	$\exp(-\|\Delta q\|^2/\sigma)$	0.5
Joint vel. track.	$\exp(-\|\Delta\dot{q}\|^2/\sigma)$	0.125
Regularization & safety		
Torque penalty	$\|\tau\|^2$	-10^{-6}
Torque limit	$\max(0, \|\tau\| - \tau_{\max})$	-0.1
Joint pos. limit	$\max(0, q - q_{\max}) + \max(0, q_{\min} - q)$	-10
Joint vel. limit	$\max(0, \|\dot{q}\| - \dot{q}_{\max})$	-5
Joint acc.	$\|\ddot{q}\|^2$	$-2.5{\cdot}10^{-7}$
Momentum change	$\|\Delta p\|$	$-5{\cdot}10^{-3}$
Body yank	$\|\Delta F\|^2$	$-2{\cdot}10^{-6}$
Joint vel. penalty	$\|\dot{q}\|^2$	-10^{-4}
Action rate	$\|a_t - a_{t-1}\|^2$	-0.1
Undesired contacts	r_{contacts} [26]	-0.1
Foothold	r_{foothold} [21]	-1
Post-recovery stabilization		
Head height	$\exp(-\|\Delta h\|^2/\sigma)$	0.25
Base lin. vel.	$\|\mathbf{v}_{\text{base}}\|^2$	-0.01
Base ang. vel.	$\|\boldsymbol{\omega}_{\text{base}}\|^2$	-0.05

We train the student by distilling from the privileged teacher using a hybrid of behavioral cloning and DAgger-style [27], [28] rollouts. Supervision is applied over a mixture of teacher- and student-induced state distributions,

$$\rho = \alpha\rho_{\pi_\theta} + (1-\alpha)\rho_{\pi_\phi},$$

where ρ_{π_θ} and ρ_{π_ϕ} denote the distributions induced by teacher and student rollouts, respectively. The mixing coefficient α is annealed from 1 toward 0 during training, gradually shifting rollouts from teacher-driven supervision to student-induced states. The student is trained using weighted behavior-cloning and goal-in-context reconstruction losses:

$$\mathcal{L}_{\text{distill}} = \mathbb{E}_{t\sim\rho}\left[\lambda_z \left\|\tilde{\mathbf{z}}_g - \mathbf{z}_g\right\|^2 + \lambda_a \left\|\tilde{\mathbf{a}}_t - \mathbf{a}_t\right\|^2\right].$$

Teacher rollouts provide stable demonstrations that rapidly imprint recovery behavior, while student rollouts expose the policy to off-distribution states and improve adaptability.

Although the teacher is trained with motion-tracking signals that include intentional falls, the student is distilled from unstable initial states, which biases it toward stabilization rather than deliberate falling. Therefore, without disturbance the policy maintains a stable posture; under disturbance, it mitigates fall severity and recovers.

E. Domain Randomization

To improve robustness and sim-to-real transfer, following prior works [16], [28], we randomize both dynamics and perception at the start of each episode and throughout rollout. On the dynamics side, we randomize friction and restitution, and further vary the episode initialization described in Sec. III-A, including the reference motion clip and phase, planar po-

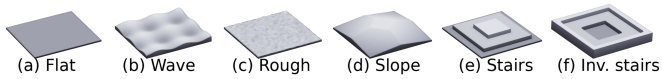


Fig. 4. Training terrains, from left to right: flat, wave, rough, slope, stairs, and inverted stairs. One representative difficulty is shown.

sition, yaw, initial pose, and joint states. We additionally apply stochastic external pushes and occasional joint torque dropouts to mimic disturbances and partial actuator failures. On the perception side, we perturb depth observations using depth clipping and non-linear remapping, multiplicative noise, spatial and temporal dropout, synthetic occlusions, and small camera pose jitter. These perturbations encourage the policy to rely on stable geometric structure.

IV. Experiments

A. Implementation Details

We summarize the main components of our simulation, training, and deployment setup below. For more visual results, see the supplementary multimedia.

Simulation Setup. We use a 23-DoF Unitree G1 with a head-mounted depth camera in IsaacGym and IsaacSim, controlled at 50 Hz with a 7.5 s recovery horizon. Training proceeds in two phases. The policy is first trained on flat terrain to acquire core stand-up and impact-mitigation behaviors, and is then continued on randomized non-flat terrains, as shown in Fig. 4. We train the teacher with PPO using 4,096 parallel environments, then distill it into the visual student using 512 environments.

Real-World Setup: Hardware experiments are performed on a Unitree G1 equipped with an Intel RealSense head mounted depth camera. Proprioception streams at 500 Hz and depth at 30 Hz. The student policy runs at 50 Hz and outputs joint-space position targets to the low-level PD controller. Depth preprocessing mirrors simulation, and deployment is *fully zero-shot* without any real-world fine-tuning.

B. Simulated Experiments

We evaluate policies under two initialization regimes on terrains sampled from the families shown in Fig. 4. In both cases, initial planar position and yaw are randomized.

In *Stand-Up*, episodes start from low-body configurations sampled around the lowest keyframe of each demonstration, with pose noise. In *Fall-Recovery*, episodes start earlier in the motion, near the onset of falling, with random perturbations and nonzero base velocity, requiring the policy to handle impact mitigation and recovery on varying terrain geometry. Unless otherwise stated, results are reported in IsaacGym.

Metrics. Simulation results in this section use 500 trials per seed across three different evaluation seeds. Table II defines the evaluation metrics.

C. Simulated Results

Teacher Ablations. We evaluate teacher-side ablations to isolate the contribution of privileged structure and supervision.

TABLE II
Evaluation metrics.

Metric	Definition
SR ↑ (%)	Success with upright stabilization within 7.5 s.
SSR ↑ (%)	Stable Success without re-fall.
Time ↓ (s)	Time to recovery.
PIF ↓	Peak internal-force magnitude.
F_c ↓ (kN)	Peak contact force across all links.
FR ↓ (%)	Subsequent re-fall rate.
Drift ↓ (m)	Horizontal pelvis displacement.
MPJPE ↓ (cm)	Pelvis-relative joint tracking error.

TABLE III
Teacher ablations under Stand-Up and Fall-Recovery.

Method	Stand-Up			Fall-Recovery		
	SR↑	SSR↑	Time↓	SR↑	SSR↑	Time↓
NoKeypoints	57.4 (± 4.1)	53.2 (± 3.0)	6.9 (± 1.3)	76.1 (± 3.1)	73.2 (± 3.4)	6.3 (± 1.5)
DofKeypoints	86.6 (± 3.1)	80.5 (± 3.4)	5.8 (± 1.0)	84.1 (± 2.9)	80.1 (± 2.4)	5.5 (± 0.9)
AbsTrack	73.1 (± 3.0)	70.0 (± 3.5)	5.7 (± 2.0)	70.4 (± 3.5)	67.3 (± 3.8)	5.6 (± 2.1)
NoScandots	84.4 (± 3.8)	70.0 (± 3.4)	5.6 (± 1.7)	88.3 (± 3.7)	77.6 (± 4.6)	6.4 (± 1.6)
Teacher	**98.0** (± 0.3)	**97.6** (± 0.2)	**4.2** (± 1.0)	**96.3** (± 0.8)	**95.7** (± 1.2)	**5.3** (± 0.9)

NoKeypoints removes keypoint-based observations. DofKeypoints replaces spatial keypoint positions with joint angle targets. AbsTrack trains the teacher using absolute pose tracking objectives instead of relative tracking. NoScandots removes access to privileged terrain information, while Teacher denotes the full privileged teacher.

Table III highlights the role of structural priors in the teacher. Removing keypoint guidance causes the largest degradation in Stand-Up, while replacing relative tracking with absolute tracking causes the largest degradation in Fall-Recovery. Similarly, replacing spatial keypoints with joint-angle targets (DofKeypoints) partially recovers performance but remains below the full model, suggesting that geometric body configuration provides stronger guidance than joints alone. Removing privileged terrain samples (NoScandots) limits the teacher's scene awareness. The full privileged Teacher achieves the best overall performance.

Baselines. We compare VIGOR with pretrained HOST [4] and FIRM [12] policies, using HOST's official checkpoints and a FIRM checkpoint provided by its authors. HOST learns separate RL stand-up policies for different initial configurations, while FIRM retrieves keyframe goal features to condition a diffusion policy. As no prior visual baseline exists for humanoid fall recovery, we analyze perception through student ablations.

Table IV reports average recovery performance across all simulated terrains for both Stand-Up and Fall-Recovery.

VIGOR achieves substantially higher recovery and stable success than the pretrained baselines in the evaluated settings.

Student Ablations. We ablate student-side components of VIGOR while keeping training conditions fixed. w.o Shared removes intermediate latent supervision, retaining action imitation. w.o Vision removes egocentric depth input, yielding a proprioception-only student. w.o History removes temporal observation history, restricting the student to single-step observations. VIGOR includes egocentric depth, temporal history,

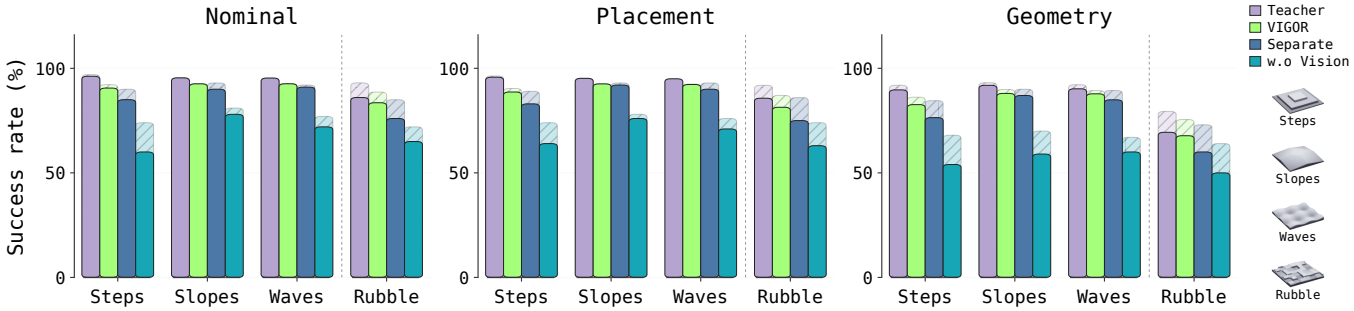


Fig. 5. Recovery across terrain conditions. Solid bars show SSR; ▨ bars show SR (%). Dividers mark unseen rubble, excluded from training.

TABLE IV
AVERAGE SIMULATED RECOVERY PERFORMANCE.

Method	Stand-Up			Fall-Recovery		
	SR↑	SSR↑	Time↓	SR↑	SSR↑	Time↓
HOST [4]	17.0 (±3.1)	16.1 (±3.1)	6.9 (±1.6)	–	–	–
FIRM [12]	35.3 (±4.6)	30.4 (±4.1)	6.2 (±1.5)	22.1 (±6.5)	19.3 (±3.6)	6.9 (±2.1)
VIGOR	**90.8** (±1.2)	**88.7** (±1.7)	**5.1** (±2.2)	**91.3** (±0.9)	**90.5** (±0.7)	**5.5** (±1.0)



Fig. 6. Demonstration scaling. Recovery success (SR) and stable success (SSR) for visual students supervised by teachers trained with one, four, or nine human demonstrations, under Stand-Up and Fall-Recovery.

and intermediate latent supervision. Separate retains latent supervision but replaces the shared goal–terrain projection with separate encodings, whose concatenation the student predicts.

Table V shows that vision and intermediate latent supervision both contribute substantially to recovery. Removing vision causes the largest reduction in stable success in both tasks, while removing intermediate latent supervision also substantially degrades recovery, supporting recovery-level guidance beyond action imitation. The blind student inherits strong recovery skills from the teacher and can find support and regain a stable posture even after terrain collisions. Separate achieves comparable recovery success but lower stable success. An additional evaluation also finds higher peak contact forces for Separate than VIGOR (2.4 versus 2.0 kN). Temporal history provides a clearer benefit in Stand-Up, where successive contacts must be coordinated. Its smaller benefit in Fall-Recovery is consistent with greater reliance on immediate feedback during rapid descent.

Recovery Across Terrain Conditions. Factorization lets flat-ground demonstrations guide recovery across varied terrain contacts. We evaluate the teacher and students on steps, slopes, waves, and unseen rubble (Fig. 5), without new demonstrations or fine-tuning. *Nominal* uses random starts on standard terrain; *Placement* varies starts and headings around terrain features; *Geometry* increases terrain severity by doubling heights and gradients.

TABLE V
STUDENT ABLATIONS UNDER STAND-UP AND FALL-RECOVERY.

Method	Stand-Up			Fall-Recovery		
	SR↑	SSR↑	Time↓	SR↑	SSR↑	Time↓
w.o Vision	74.3 (±3.2)	60.9 (±1.5)	5.8 (±2.0)	80.6 (±3.0)	71.2 (±1.9)	5.5 (±1.8)
w.o Shared	75.8 (±3.7)	70.9 (±2.1)	6.3 (±2.1)	76.1 (±3.2)	72.5 (±2.4)	6.8 (±2.0)
w.o History	86.4 (±3.0)	84.2 (±1.8)	5.1 (±1.9)	89.4 (±1.4)	88.2 (±1.5)	**5.4** (±1.4)
Separate	88.9 (±0.8)	84.3 (±1.7)	**5.0** (±2.1)	89.3 (±0.7)	85.3 (±0.8)	**5.4** (±1.1)
VIGOR	**90.8** (±1.2)	**88.7** (±1.7)	5.1 (±2.2)	**91.3** (±0.9)	**90.5** (±0.7)	5.5 (±1.0)

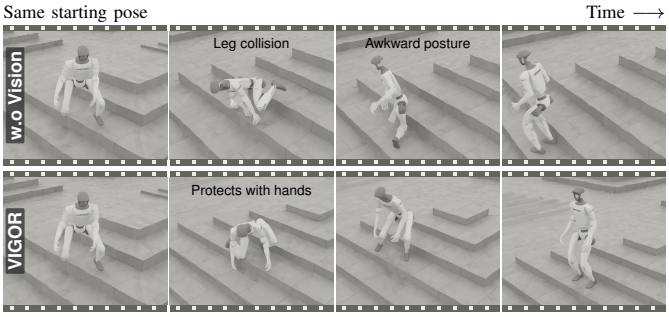


Fig. 7. Stair recovery illustration: w.o Vision (top) and VIGOR (bottom), progressing from left to right.

Placement changes have relatively little effect, whereas increased geometry reduces recovery, especially on rubble. The teacher's performance on unseen rubble supports transfer of flat-ground motion priors. On unseen rubble, VIGOR and Separate reach standing at similar rates, but VIGOR more often maintains stable recovery; the blind policy performs worse on both metrics. We hypothesize that avoiding reconstruction of a separate terrain latent helps VIGOR maintain recovery on unfamiliar terrain.

Demonstration Scaling. We compare visual students supervised by teachers trained with one, four, or the full demonstration set. Evaluation uses identical held-out starts across four fall directions and the tested terrains. The gain from one to four demonstrations (Fig. 6) suggests that complementary recovery patterns broaden the student's coverage.

Safety and Recovery Quality. We compare the teacher, VIGOR, and the blind student from matched initial states under both tasks. Table VI suggests that vision supports more controlled recovery: improved reliability is accompanied by lower contact forces, less drift, and closer motion tracking. The reduction in contact peaks is more pronounced than the change in PIF, highlighting the benefit of vision for controlling robot–

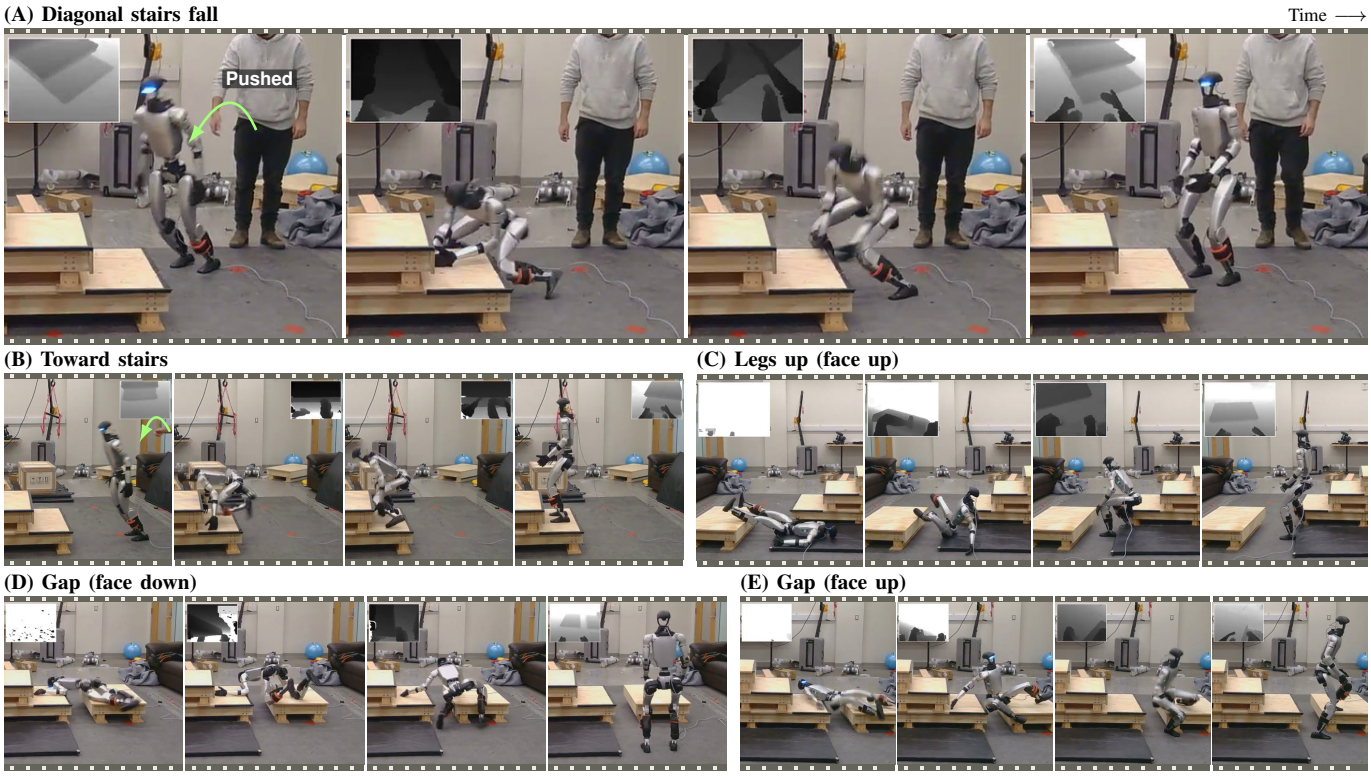


Fig. 8. Real-world recovery sequences from the project teaser: (A) diagonal stairs fall, (B) fall toward stairs, (C) stand-up with raised legs, and stand-up across a gap from (D) face-down and (E) face-up configurations. Four selected frames per example show the starting pose or fall, contact transitions, and recovery from left to right; insets show synchronized depth.

TABLE VI
SAFETY AND RECOVERY QUALITY.

Method	PIF↓	F_c↓ (kN)	FR↓ (%)	Drift↓ (m)	MPJPE↓ (cm)
w.o Vision	44.4 (±7.7)	3.3 (±1.1)	22.1 (±3.0)	0.9 (±0.6)	21.5 (±6.6)
VIGOR	40.6 (±7.6)	2.0 (±1.0)	7.3 (±1.6)	0.5 (±0.5)	17.3 (±5.9)
Teacher	39.3 (±6.0)	1.8 (±1.1)	3.1 (±2.5)	0.5 (±0.4)	11.5 (±4.8)

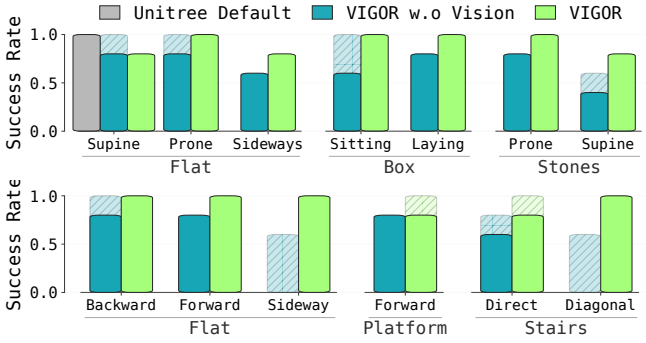


Fig. 9. Real-world recovery performance across surfaces and initial configurations. Top: stand-up. Bottom: fall recovery. Bars show success over 5 trials; ▨ denotes successful recoveries with unsafe upper-body contact.

terrain loading. The teacher's stronger reliability and tracking indicate further room for improvement in distillation. Fig. 7 shows recorded recovery sequences for the visual and blind students.

D. Real-World Experiments

We evaluate VIGOR on a Unitree G1 without task-specific tuning or parameter changes. Each setting is tested for 5 trials under the *Stand-Up* and *Fall-Recovery* regimes, with representative sequences shown in Fig. 8. See the supplementary video for the setups.

Stand-Up. We test stand-up on flat ground, stones, and a box obstacle. Flat-ground trials include prone, supine, and sideways initializations; stones include prone and supine initializations; and box trials include seated and torso-on-box configurations. In all cases, joint configurations are *randomized*, and the policy must recover to an upright stance.

Fall-Recovery. We test fall recovery on flat ground, a raised platform, and stairs. Flat-ground trials use backward, forward, and sideways pushes; platform trials push the robot forward off the edge; and stair trials use direct and diagonal pushes toward the stairs. To ensure a genuine dynamic fall, the policy is initially disabled and activated only after the projected gravity vector indicates that the robot has reached a predefined fall angle (20°). This delayed activation evaluates mitigation and recovery after destabilization.

Results. Fig. 9 reports real-world performance on a Unitree G1 under *Stand-Up* and *Fall-Recovery*, grouped by terrain and initialization or disturbance type. Fig. 8 shows corresponding recovery behaviors. *Unitree* controller only succeeds for face-up flat-ground stand-up, so we omit it from further evaluation. For stand-up, both policies achieve high recovery rates across initial poses on flat terrain. As terrain complexity increases, especially on stones and box obstacles,

the benefit of visual grounding becomes clearer: VIGOR maintains higher success under uneven contacts, indicating more reliable terrain-aware stabilization. During fall recovery, both policies recover reliably from forward and backward pushes on flat terrain. Under more challenging disturbances, including sideways pushes and stair interactions, VIGOR achieves more recoveries without unsafe upper-body contact, with the strongest gains in settings that require terrain-aware reactions such as diagonal stair pushes. These results show why recovery success alone is insufficient to characterize fall safety. In sideways falls and diagonal stair pushes, the blind policy sometimes reaches standing despite unsafe upper-body contacts, whereas VIGOR achieves more recoveries without these contacts. The recovery sequences in Fig. 8 illustrate how depth helps the policy identify nearby support and obstacles before impact and coordinate contacts during recovery.

V. Conclusion

We present VIGOR, a unified framework for humanoid fall safety that integrates fall mitigation and recovery in a single vision-conditioned policy. By factorizing data into sparse human motion priors and independently varying terrain, and by representing recovery goals through a compact goal-in-context latent, VIGOR enables RL to resolve contact timing and terrain-dependent execution without dense demonstrations. A privileged teacher learns terrain-aware recovery from sparse keyframes and terrain access, and is distilled into a deployable student operating from egocentric perception. Simulation and real-world results demonstrate robust recovery across diverse terrains and initial conditions, validating structured teacher-student learning as a practical path to zero-shot sim-to-real humanoid fall safety. Future work will explore integration with ongoing tasks, such as locomotion, to support task continuation after recovery.

References

[1] J. Englsberger, T. Koolen, S. Bertrand, J. Pratt, C. Ott, and A. Albu-Schäffer, "Trajectory generation for continuous leg forces during double support and heel-to-toe shift based on divergent component of motion," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2014, pp. 4022–4029.

[2] D. Ferigo, R. Camoriano, P. M. Viceconte, D. Calandriello, S. Traversaro, L. Rosasco, and D. Pucci, "On the emergence of whole-body strategies from humanoid robot push-recovery learning," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 8561–8568, 2021.

[3] D. Luo, Y. Ding, Z. Cao, and X. Wu, "A multi-stage approach for efficiently learning humanoid robot stand-up behavior," in *2014 IEEE international conference on mechatronics and automation*. IEEE, 2014, pp. 884–889.

[4] T. Huang, J. Ren, H. Wang, Z. Wang, Q. Ben, M. Wen, X. Chen, J. Li, and J. Pang, "Learning Humanoid Standing-up Control across Diverse Postures," in *Proceedings of Robotics: Science and Systems*, LosAngeles, CA, USA, June 2025.

[5] F. Zargarbashi, J. Cheng, D. Kang, R. Sumner, and S. Coros, "RobotKeyframing: Learning Locomotion with High-Level Objectives via Mixture of Dense and Sparse Rewards," in *8th Annual Conference on Robot Learning*, 2024.

[6] K. Wen, C. Li, J. He, and M. Hutter, "Constrained Style Learning from Imperfect Demonstrations under Task Optimality," in *9th Annual Conference on Robot Learning*, 2025.

[7] Z. Zhang, C. Li, T. Miki, and M. Hutter, "Motion Priors Reimagined: Adapting Flat-Terrain Skills for Complex Quadruped Mobility," in *9th Annual Conference on Robot Learning*, 2025.

[8] V. C. Kumar, S. Ha, and C. K. Liu, "Learning a unified control policy for safe falling," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 3940–3947.

[9] J. Stückler, J. Schwenk, and S. Behnke, "Getting back on two feet: Reliable standing-up routines for a humanoid robot." in *IAS*, 2006, pp. 676–685.

[10] X. He, R. Dong, Z. Chen, and S. Gupta, "Learning Getting-Up Policies for Real-World Humanoid Robots," in *Proceedings of Robotics: Science and Systems*, LosAngeles, CA, USA, June 2025.

[11] C. Gaspard, M. Duclusaud, G. Passault, M. Daniel, and O. Ly, "Frasa: An end-to-end reinforcement learning agent for fall recovery and stand up of humanoid robots," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2025, pp. 15 994–16 000.

[12] Z. Xu, Y. Li, K.-y. Lin, and S. X. Yu, "Unified humanoid fall-safety policy from a few demonstrations," *arXiv preprint arXiv:2511.07407*, 2025.

[13] Z. Zhuang, S. Yao, and H. Zhao, "Humanoid Parkour Learning," in *8th Annual Conference on Robot Learning*, 2024.

[14] Z. Fu, X. Cheng, and D. Pathak, "Deep Whole-Body Control: Learning a Unified Policy for Manipulation and Locomotion," in *6th Annual Conference on Robot Learning*, 2022.

[15] 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 *8th Annual Conference on Robot Learning*, 2024.

[16] 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.

[17] A. Loquercio, A. Kumar, and J. Malik, "Learning visual locomotion with cross-modal supervision," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 7295–7302.

[18] K.-Y. Lin and S. X. Yu, "Let humanoids hike! integrative skill development on complex trails," in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 22 498–22 507.

[19] H. Duan, B. Pandit, M. S. Gadde, B. J. van Marum, J. Dao, C. Kim, and A. Fern, "Learning Vision-Based Bipedal Locomotion for Challenging Terrain," *arXiv preprint arXiv:2309.14594*, 2023.

[20] Z. Luo, C. Tessler, T. Lin, Y. Yuan, T. He, W. Xiao, Y. Guo, G. Chechik, K. Kitani, L. Fan *et al.*, "Emergent active perception and dexterity of simulated humanoids from visual reinforcement learning," *arXiv preprint arXiv:2505.12278*, 2025.

[21] H. Wang, Z. Wang, J. Ren, Q. Ben, T. Huang, W. Zhang, and J. Pang, "Beamdojo: Learning agile humanoid locomotion on sparse footholds," *arXiv preprint arXiv:2502.10363*, 2025.

[22] X. B. Peng, P. Abbeel, S. Levine, and M. Van de Panne, "Deepmimic: Example-guided deep reinforcement learning of physics-based character skills," *ACM Transactions On Graphics (TOG)*, vol. 37, no. 4, pp. 1–14, 2018.

[23] 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)*, vol. 40, no. 4, pp. 1–20, 2021.

[24] A. Allshire, H. Choi, J. Zhang, D. McAllister, A. Zhang, C. M. Kim, T. Darrell, P. Abbeel, J. Malik, and A. Kanazawa, "Visual imitation enables contextual humanoid control," in *Proceedings of the Conference on Robot Learning (CoRL)*, 2025.

[25] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[26] Q. Liao, T. E. Truong, X. Huang, Y. Gao, G. Tevet, K. Sreenath, and C. K. Liu, "Beyondmimic: From motion tracking to versatile humanoid control via guided diffusion," *arXiv preprint arXiv:2508.08241*, 2025.

[27] S. Ross, G. Gordon, and D. Bagnell, "A reduction of imitation learning and structured prediction to no-regret online learning," in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 2011, pp. 627–635.

[28] O. Azulay, D. M. Ramesh, N. Curtis, and A. Sintov, "Visuotactile-based learning for insertion with compliant hands," *IEEE Robotics and Automation Letters*, vol. 10, no. 4, pp. 4053–4060, 2025.