COLA: Learning Human-Humanoid Coordination for Collaborative Object Carrying

Yushi Du [*1,2] Yixuan Li [*2,3] Baoxiong Jia [*2†] Yutang Lin [2,4] Pei Zhou [1]
Wei Liang [3†] Yanchao Yang [1†] Siyuan Huang [2†]

[1]The University of Hong Kong [2]State Key Laboratory of General Artificial Intelligence, BIGAI
[3]Beijing Institute of Technology [4]Peking University
[*]Equal contribution †Corresponding authors

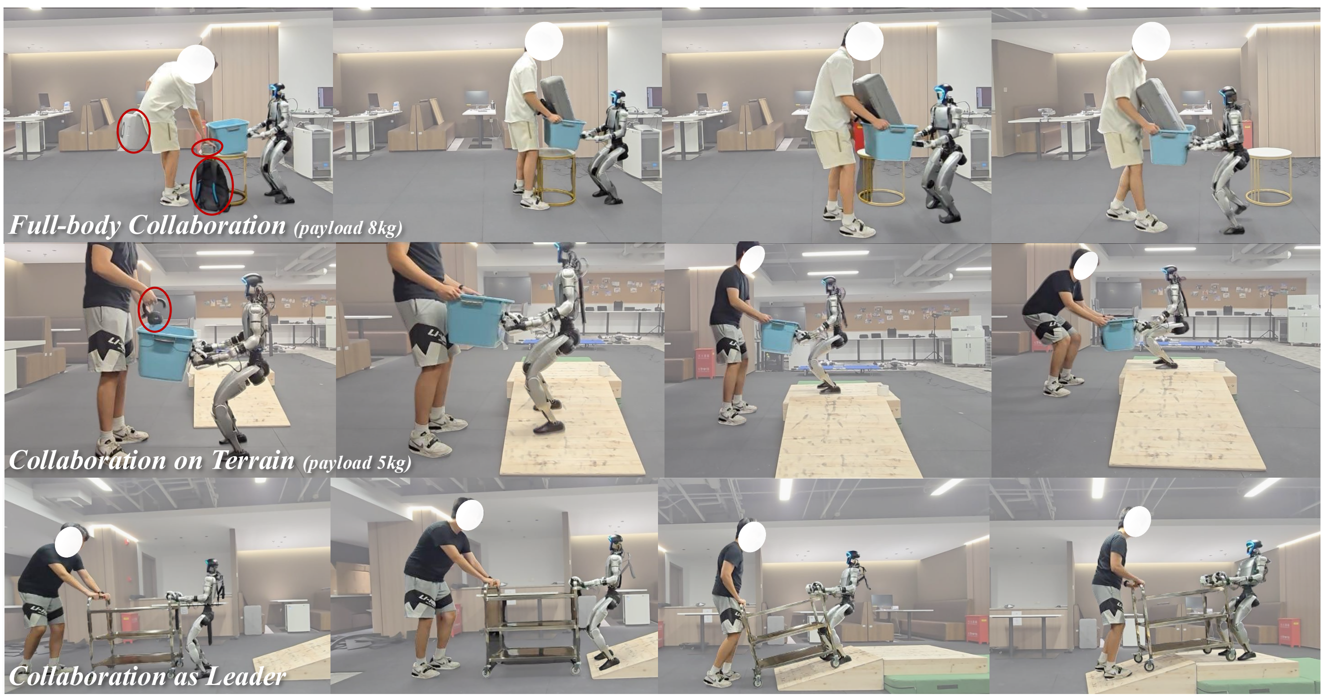


Figure 1. COLA provides a proprioception-only policy that enables compliant human-humanoid collaboration for carrying diverse objects across various movement patterns. (a) Horizontal trajectory coordination: the humanoid adapts to human motion patterns. (b) Whole-body coordination during collaborative object-lowering. (c) COLA acts as the leader and collaborates with a human to drag a cart up a slope.

Abstract

Human-humanoid collaboration shows significant promise for applications in healthcare, domestic assistance, and manufacturing. While compliant robot-human collaboration has been extensively developed for robotic arms, enabling compliant human-humanoid collaboration remains largely unexplored due to humanoids' complex whole-body dynamics. We propose COLA, a proprioception-only reinforcement learning approach that combines leader and follower behaviors within a single policy. COLA is trained in a closed-loop environment with dynamic object interactions, enabling it to predict object motion patterns and human intentions implicitly and to maintain load balance through coordinated trajectory planning. Simulation experiments show that COLA reduces human effort by 24.7% *over baselines while maintaining object stability. Real-world experiments validate robust collaborative carrying across diverse objects (boxes, desks, stretchers,* etc.*) and movement patterns (straight-line, turning, slope climbing). A user study with* 23 *participants* *confirms an average improvement of* 27.4% *over baselines. COLA thus enables compliant collaborative carrying without external sensors or complex interaction models. Project website:* `https://yushi-du.github.io/COLA/`.

1. Introduction

Recent years have witnessed significant progress in humanoid robot development, including agile locomotion [8, 14, 28, 33], teleoperation [11, 26], and dexterous manipulation [19, 27]. While these advances highlight growing versatility and robustness of humanoid control, progress in enabling humanoid robots to collaborate effectively with humans remains limited. Human-robot collaboration has been a long-standing challenge [5, 17, 18, 25, 31, 32], requiring the modeling of diverse human behaviors, adaptive responses to dynamic interactions, and coordinated planning for shared tasks. As humanoid robots develop increasingly reliable motor and control abilities, addressing collaboration becomes timely and essential to realize their central role in supporting human life.

Object carrying [2, 3, 7] is a representative task for advancing such collaboration. Its core challenges arise from adapting to diverse terrains, responding compliantly to human motion under limited or no direct force sensing, and dynamically allocating leader or follower roles. These interdependent requirements make the task particularly difficult for humanoids: true collaboration requires integrating all aspects to ease the human partner, rather than addressing a single constraint as in prior work on environment-conditioned locomotion [14, 29, 33], compliance behavior learning [23–25, 30], or high-level intention prediction in open-loop object-finding or serving tasks [6, 13, 20, 31].

To address these challenges, we propose COLA, a learning-based policy for human-humanoid collaborative carrying that uses reinforcement learning to model dynamic and versatile interactions. COLA lets humanoids share loads with humans compliantly and flexibly switch between leader and follower roles. Our design rests on two insights: (i) offsets between joint states and their targets provide a proxy for estimating interaction forces, and (ii) the carried object's state encodes implicit collaboration constraints such as stability and coordination. We adopt a teacher-student framework where the teacher, trained with proprioceptive and privileged object-state information, is supervised by rewards on humanoid motion (*e.g.* robust locomotion across terrains) and object status (*e.g.* keeping a stretcher level). The student is distilled from the teacher and relies solely on proprioceptive inputs for real-world deployment. Role allocation is controlled by a velocity command (zero velocity = follower).

Our contributions are:

- A unified residual policy that relies solely on proprioception for whole-body collaborative carrying, supporting compliant, coordinated, and generalizable collaboration across diverse movement patterns.
- A three-step training framework and closed-loop training that explicitly model humanoid-object interactions, enabling the robot to implicitly learn object dynamics and assist humans through compliant collaboration.
- Simulation and real-world experiments demonstrating superior effort reduction and trajectory coordination over baselines; a user study with 23 participants confirms that COLA achieves more compliant collaboration.

2. Related Work

Robot-human Collaboration. Robot-human collaboration spans from robotic arms to legged robots [10, 13, 20]. While arm-based systems assist in confined workspaces, humanoid robots are increasingly used for open-environment assistance. Existing human-humanoid methods [5, 17, 32] mostly rely on model-based control with heuristic subtasks [1, 2] or multi-modal intention prediction [15, 16]. H^2-COMPACT [3] learns horizontal velocity commands from haptic cues but operates in a limited scope; none of these methods exploit whole-body coordination [22] for tasks such as ground pickup or slope climbing. We instead propose a residual learning framework that enables humanoids to collaborate using whole-body coordination, broadening the range of feasible collaborative scenarios.

Compliant Whole-body Control. Position-only control lacks compliance for human-humanoid interaction [9, 25], since it ignores force. Force regulation is essential for contact-rich collaboration [12]. Recent work shows that explicit contact-force estimation [4, 30] or implicit force-adaptive behavior [24] can improve compliance on robotic arms and legged robots. How force-aware control benefits human-humanoid collaboration, however, remains underexplored. We incorporate force considerations implicitly into our framework, enabling natural cooperative interactions without external force sensing.

3. Methodology

Task. We define human-humanoid collaborative carrying as a humanoid assisting a human partner to transport an object that is challenging for a single person due to its size or weight. With the partner already engaged in carrying, the robot must (i) coordinate its motion to align with the human's velocity, (ii) support the object's weight to reduce the human's physical burden, and (iii) stabilize the object's orientation throughout transport.

Three-step pipeline. As shown in Fig. 2, our pipeline consists of: (a) base whole-body controller training, (b) residual teacher policy training for collaboration in a closed-loop environment, and (c) student policy distillation for deployment.

Whole-body controller (WBC). The whole-body control

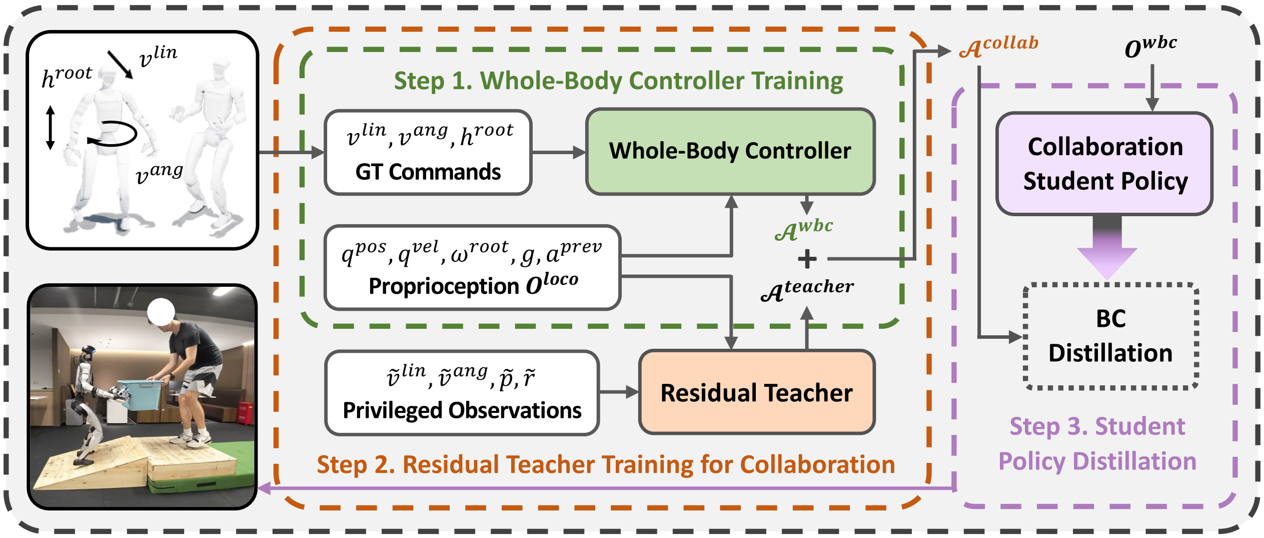


Figure 2. **Overview of COLA.** (i) Train a base whole-body control policy to provide a robust whole-body controller. (ii) In the closed-loop training environment, train a residual teacher policy on top of the controller using privileged information about the carried object. (iii) Distill the teacher into a student policy for real-world deployment via behavioral cloning.

(WBC) policy receives goal commands $\mathcal{G} = [\mathcal{G}^{\text{lower}}, \mathcal{G}^{\text{upper}}]$ — lower-body locomotion goals $[v^{\text{lin}}, v^{\text{ang}}, h^{\text{root}}]$ and upper-body end-effector goals $[p^{\text{ee}}, r^{\text{ee}}]$ — together with proprioception $\mathcal{O}_t^{\text{wbc}} = [q^{\text{pos}}_{t-l:t}, q^{\text{vel}}_{t-l:t}, \omega^{\text{root}}_{t-l:t}, g_{t-l:t}, a^{\text{prev}}_{t-(l+1):t-1}]$. Actions $\mathcal{A}^{\text{wbc}} \in \mathbb{R}^N$ are target joint positions for the N=29 joints of the G1 robot (excluding fingers), executed via PD control. We train with Proximal Policy Optimization (PPO) following [21, 29] and apply external forces at the end-effectors to improve payload robustness.

Residual teacher policy. We then introduce a closed-loop environment that explicitly models the humanoid–object–human dynamics: a simulated support body, connected to the object through a 6-DoF joint, plays the human partner. Random velocities, target angular velocities, and target heights are applied to the support body via PD control, so the support body's dynamics are transmitted to the object through joint friction, damping, and limits. The teacher observes both proprioception and privileged object state $\mathcal{O}_t^{\text{priv}} = [\widetilde{v}^{\text{lin}}_{t-l:t}, \widetilde{v}^{\text{ang}}_{t-l:t}, \widetilde{p}_{t-l:t}, \widetilde{r}_{t-l:t}]$, and outputs a residual action that adjusts the WBC output: $\mathcal{A}^{\text{collab}} = \mathcal{A}^{\text{wbc}} + \mathcal{A}^{\text{teacher}}$. We hypothesize this implicit, residual paradigm is effective because: (i) the WBC's exploration may not cover all collaboration motion constraints; (ii) interaction nuances are hard to encode with explicit commands. Training rewards combine WBC rewards with task-specific terms (Tab. 1).

Student distillation. We distill the teacher ($\mathcal{F}^{\text{wbc}} + \mathcal{F}^{\text{teacher}}$) into a student $\mathcal{F}^{\text{student}} : \mathcal{O}^{\text{wbc}} \to \mathcal{A}^{\text{student}}$ that uses only proprioception, via behavioral cloning with loss $\mathcal{L}_{\text{distill}} = \mathbb{E}[\, \|\mathcal{A}^{\text{student}} - \mathcal{A}^{\text{collab}}\|^2\,]$. Role allocation is realized through the goal command: in COLA-F (*Follower*) all networks receive zero goal commands; in COLA-L (*Leader*) a sampled goal command within the WBC's range is provided.

Table 1. **Reward functions** for collaboration-policy training. $\phi(x) = e^{-\|x\|}$; $v_{\text{lin}}^{\text{applied}}$, $v_{\text{ang}}^{\text{goal}}$ are the applied/target linear and angular velocities; $h_{1,2}^{\text{obj}}$ are predefined heights of the object's far ends; $\mathcal{F}^{\text{support-obj}}$ is the horizontal force between support body and object.

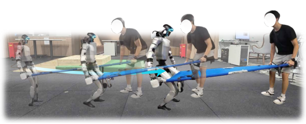
Term	Expression	Weight
Lin. Vel. Tracking	$\phi(v_{\text{lin}}^{\text{CoM}} - v_{\text{lin}}^{\text{applied}})$	1.0
Yaw Vel. Tracking	$\phi(v_{\text{ang}}^{\text{CoM}} - v_{\text{ang}}^{\text{goal}})$	1.0
Z-axis Vel. Penalty	$-\|v_z^{\text{obj}}\|$	0.05
Height Diff. Penalty	$-\|h_1^{\text{obj}} - h_2^{\text{obj}}\|$	10.0
Force Penalty	$-\lvert\mathcal{F}^{\text{support-obj}}\rvert$	0.002

Training setup. We train in Isaac Lab on a single RTX 4090D with PPO. The WBC actor/critic and the residual teacher/student each use a three-layer multi-layer perceptrons (MLPs) of size $(512, 256, 128)$, and observations use a history of $l = 25$ frames at 50 Hz. The WBC, teacher, and distillation stages take 350k, 250k, and 250k environment steps respectively, totaling about 48 hours.

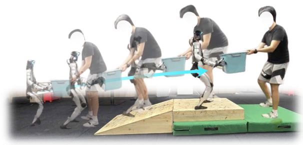
4. Experiments

We address: (Q1) Do the residual teacher and distillation enable effective compliant collaboration? (Q2) Is the COLA student architecture concise and effective? (Q3) Do the results transfer to real-world deployment?

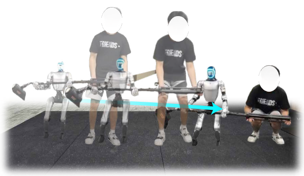
Baselines and metrics. We compare COLA against *Vanilla MLP* (single MLP trained end-to-end with PPO), *Explicit Goal Estimation* (teacher predicts WBC goal command in-



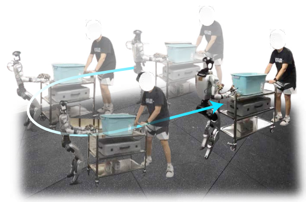
(a) Stretcher



(b) Box on slope



(c) Rod height tracking



(d) Cart pushing

Figure 3. **Qualitative human-humanoid collaboration with COLA.** COLA carries diverse objects (3 kg rod, 8 kg box, 11 kg stretcher, 20 kg cart) under various grasps on flat ground and slopes.

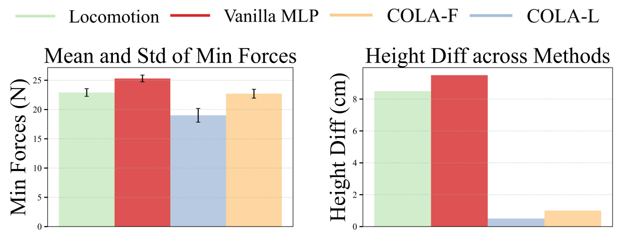
stead of residual), and *Transformer* (student MLP replaced with a Transformer), and ablate history length (10 vs. 50). Metrics: mean linear/angular velocity tracking error w.r.t. the human (*Lin. Vel.*, *Ang. Vel.*), object height-tracking error (*Height Err.*), and average horizontal interaction force (*Avg. E.F.*, reflecting human effort).

Table 2. **Quantitative evaluation in simulation.** Velocity/height tracking and average external force on the carried object reflect collaboration accuracy and human effort. Best in **bold**.

Methods	Lin. V.↓ (m/s)	Ang. V.↓ (rad/s)	H. Err.↓ (m)	E.F.↓ (N)
Explicit Goal Est.	0.235	0.335	0.102	19.07
Vanilla MLP	0.173	0.261	0.080	17.95
Transformer	0.178	0.310	0.077	19.38
COLA-F-Hist. 10	0.121	0.131	0.037	15.44
COLA-F-Hist. 50	0.116	0.132	0.036	14.57
COLA-F	0.109	0.118	**0.031**	14.58
COLA-L-Hist. 10	0.118	0.106	0.039	13.92
COLA-L-Hist. 50	0.112	0.103	0.036	13.50
COLA-L	**0.102**	**0.098**	0.038	**12.30**

(Q1) Residual teacher and distillation are effective. Tab. 2: COLA beats all baselines on every metric; the lowest *Avg. E.F.* indicates strongest compliance. *Vanilla MLP* tracks linear velocity but fails on angular/height, showing the teacher–student distillation is critical. *Explicit Goal Estimation* performs worst, indicating that implicit learning of object dynamics in a closed-loop environment is more effective than predicting WBC goal commands.

(Q2) The student architecture is concise and effective. COLA outperforms the *Transformer*, which needs $>2\times$ more training to converge: long-range temporal mixing introduces unnecessary complexity, while the compact MLP adapts more promptly. Histories shorter than 25 frames miss context; longer than 25 give negligible gains, so we choose 25. COLA-L consistently outperforms COLA-F, since the goal command provides cues under interaction noise.



(a) Real force (b) Real height

Figure 4. **Real-world quantitative results.** (Left) Minimum force to move the robot; COLA follows external forces beyond ∼15 N. (Right) Height difference between human- and humanoid-held ends; COLA reduces error by ∼3/4.

Table 3. **User study** with 23 participants rating *Height Tracking* and *Smoothness* on a 1–5 scale.

Method	Height Tracking ↑	Smoothness ↑
Locomotion	2.96	2.61
Vanilla MLP	3.09	3.09
COLA	**3.96**	**3.96**

(Q3) Real-world validation. Fig. 3: COLA carries diverse objects (boxes, carts, stretchers) under various grasps and on slopes; it implicitly interprets human intent through force interaction, so a directional push by the human is followed autonomously. Fig. 4 shows that COLA reduces height-tracking error by ∼3/4 and exhibits much stronger force compliance than the baseline. Tab. 3 reports a user study with 23 participants (videos shuffled to reduce bias); COLA achieves the highest scores on both height tracking and smoothness. Finally, the humanoid's behavior is primarily sensitive to forces at the hands/arms, while torso/leg forces leave it stable — evidence that COLA learns the humanoid–object–human interaction dynamics through joint state-target offsets alone.

5. Conclusions and Limitations

We presented COLA, a unified approach for human-humanoid collaboration based on a three-step residual learning framework that supports both leader and follower modes. COLA distills privileged object-state information into a student that operates solely on proprioceptive feedback, enabling compliant, generalizable whole-body coordination without external sensors. The closed-loop training environment explicitly models humanoid–object-human interactions, letting the robot infer human motion and adapt implicitly. While COLA already achieves effective collaboration through proprioception alone, multi-modal perception (vision, tactile) and autonomous planning for assistive behaviors are promising directions for future work.

References

[1] Don Joven Agravante, Andrea Cherubini, Antoine Bussy, Pierre Gergondet, and Abderrahmane Kheddar. Collaborative human-humanoid carrying using vision and haptic sensing. In *International Conference on Robotics and Automation (ICRA)*, 2014. 2

[2] Don Joven Agravante, Andrea Cherubini, Alexander Sherikov, Pierre-Brice Wieber, and Abderrahmane Kheddar. Human-humanoid collaborative carrying. *Transactions on Robotics (T-RO)*, 35(4):833–846, 2019. 2

[3] Geeta Chandra Raju Bethala, Hao Huang, Niraj Pudasaini, Abdullah Mohamed Ali, Shuaihang Yuan, Congcong Wen, Anthony Tzes, and Yi. Fang. H2-compact: Human-humanoid co-manipulation via adaptive contact trajectory policies. In *International Conference on Humanoid Robots*, 2025. 2

[4] Yushi Du, Ruihai Wu, Yan Shen, and Hao Dong. Learning part motion of articulated objects using spatially continuous neural implicit representations. In *British Machine Vision Conference (BMVC)*, 2023. 2

[5] Matthew L Elwin, Billie Strong, Randy A Freeman, and Kevin M Lynch. Human-multirobot collaborative mobile manipulation: The omnid mocobots. *IEEE Robotics and Automation Letters (RA-L)*, 8(1):376–383, 2022. 2

[6] Xiaoshan Gao, Liang Yan, Gang Wang, and Chris Gerada. Hybrid recurrent neural network architecture-based intention recognition for human–robot collaboration. *IEEE Transactions on Cybernetics*, 53(3):1578–1586, 2021. 2

[7] Antonio Gonzalez-Morgado, J Soueidan, G Heredia, A Ollero, Philippe Fraisse, Marco Tognon, and Marco Cognetti. A nonlinear mpc for physical human-aerial robot interaction in collaborative transportation tasks. *IEEE Robotics and Automation Letters (RA-L)*, 2025. 2

[8] Zhaoyuan Gu, Junheng Li, Wenlan Shen, Wenhao Yu, Zhaoming Xie, Stephen McCrory, Xianyi Cheng, Abdulaziz Shamsah, Robert Griffin, C Karen Liu, et al. Humanoid locomotion and manipulation: Current progress and challenges in control, planning, and learning, 2025. 2

[9] Adrian Hartmann, Dongho Kang, Fatemeh Zargarbashi, Miguel Zamora, and Stelian Coros. Deep compliant control for legged robots. In *International Conference on Robotics and Automation (ICRA)*, 2024. 2

[10] Chenzui Li, Xi Wu, Yiming Chen, Tao Teng, Xuefeng Zhang, Sylvain Calinon, Darwin Caldwell, and Fei Chen. Human-like robot impedance regulation skill learning from human-human demonstrations. *IEEE Transactions on Cognitive and Developmental Systems*, 2025. 2

[11] Yixuan Li, Yutang Lin, Jieming Cui, Tengyu Liu, Wei Liang, Yixin Zhu, and Siyuan Huang. Clone: Closed-loop whole-body humanoid teleoperation for long-horizon tasks. In *Conference on Robot Learning (CoRL)*, 2025. 2

[12] Yanghong Li, Li Zheng, Yahao Wang, Erbao Dong, and Shiwu Zhang. Impedance learning-based adaptive force tracking for robot on unknown terrains. *Transactions on Robotics (T-RO)*, 41(3):1404–1420, 2025. 2

[13] Haotian Liu, Yuchuang Tong, Guanchen Liu, Zhaojie Ju, and Zhengtao Zhang. Idagc: Adaptive generalized human-robot collaboration via human intent estimation and multimodal policy learning. In *International Conference on Intelligent Robots and Systems (IROS)*, 2025. 2

[14] Le Ma, Ziyu Meng, Tengyu Liu, Yuhan Li, Ran Song, Wei Zhang, and Siyuan Huang. Styleloco: Generative adversarial distillation for natural humanoid robot locomotion, 2025. 2

[15] Davide Nicolis, Andrea Maria Zanchettin, and Paolo Rocco. Human intention estimation based on neural networks for enhanced collaboration with robots. In *International Conference on Intelligent Robots and Systems (IROS)*, 2018. 2

[16] Oskar Palinko, Francesco Rea, Giulio Sandini, and Alessandra Sciutti. Robot reading human gaze: Why eye tracking is better than head tracking for human-robot collaboration. In *International Conference on Intelligent Robots and Systems (IROS)*, 2016. 2

[17] Rémy Rahem, Christopher Yee Wong, and Wael Suleiman. Human-humanoid robot cooperative load transportation: model-based control approach. In *International Conference on Intelligent Robots and Systems (IROS)*, 2022. 2

[18] Leonel Rozo, Sylvain Calinon, Darwin G Caldwell, Pablo Jimenez, and Carme Torras. Learning physical collaborative robot behaviors from human demonstrations. *Transactions on Robotics (T-RO)*, 32(3):513–527, 2016. 2

[19] Rutav Shah, Shuijing Liu, Qi Wang, Zhenyu Jiang, Sateesh Kumar, Mingyo Seo, Roberto Martín-Martín, and Yuke Zhu. Mimicdroid: In-context learning for humanoid robot manipulation from human play videos. In *International Conference on Robotics and Automation (ICRA)*, 2026. 2

[20] Yubo Sheng, Yiwei Wang, Haoyuan Cheng, Huan Zhao, and Han Ding. Human-like robot action policy through game-theoretic intent inference for human-robot collaboration. *Transactions on Robotics (T-RO)*, 2025. 2

[21] Sun Wandong. Legged lab: Direct isaaclab workflow for legged robots, 2025. 3

[22] Weitian Wang, Rui Li, Yi Chen, Z Max Diekel, and Yunyi Jia. Facilitating human–robot collaborative tasks by teaching-learning-collaboration from human demonstrations. *IEEE Transactions on Automation Science and Engineering*, 16(2): 640–653, 2018. 2

[23] Christopher Yee Wong, Lucas Vergez, and Wael Suleiman. Vision-and tactile-based continuous multimodal intention and attention recognition for safer physical human–robot interaction. *IEEE Transactions on Automation Science and Engineering*, 21(3):3205–3215, 2023. 2

[24] Botian Xu, Haoyang Weng, Qingzhou Lu, Yang Gao, and Huazhe Xu. Facet: Force-adaptive control via impedance reference tracking for legged robots. In *Conference on Robot Learning (CoRL)*, 2025. 2

[25] Xinbo Yu, Bin Li, Wei He, Yanghe Feng, Long Cheng, and Carlos Silvestre. Adaptive-constrained impedance control for human–robot co-transportation. *IEEE Transactions on Cybernetics*, 52(12):13237–13249, 2021. 2

[26] Yanjie Ze, Zixuan Chen, JoÃĢo Pedro AraÃšjo, Zi-ang Cao, Xue Bin Peng, Jiajun Wu, and C Karen Liu. Twist: Teleoperated whole-body imitation system. In *Conference on Robot Learning (CoRL)*, 2025. 2

[27] Yanjie Ze, Zixuan Chen, Wenhao Wang, Tianyi Chen, Xialin He, Ying Yuan, Xue Bin Peng, and Jiajun Wu. Generalizable

humanoid manipulation with 3d diffusion policies. In *International Conference on Intelligent Robots and Systems (IROS)*, 2025. 2

[28] Tong Zhang, Boyuan Zheng, Ruiqian Nai, Yingdong Hu, Yen-Jen Wang, Geng Chen, Fanqi Lin, Jiongye Li, Chuye Hong, Koushil Sreenath, et al. Hub: Learning extreme humanoid balance. *Conference on Robot Learning (CoRL)*, 2025. 2

[29] Yuanhang Zhang, Yifu Yuan, Prajwal Gurunath, Tairan He, Shayegan Omidshafiei, Ali akbar Agha-mohammadi, Marcell Vazquez-Chanlatte, Liam Pedersen, and Guanya Shi. Falcon: Learning force-adaptive humanoid loco-manipulation. In *Annual Learning for Dynamics & Control Conference*, 2026. 2, 3

[30] Peiyuan Zhi, Peiyang Li, Jianqin Yin, Baoxiong Jia, and Siyuan Huang. Learning unified force and position control for legged loco-manipulation. In *Conference on Robot Learning (CoRL)*, 2025. 2

[31] Peiyuan Zhi, Zhiyuan Zhang, Yu Zhao, Muzhi Han, Zeyu Zhang, Zhitian Li, Ziyuan Jiao, Baoxiong Jia, and Siyuan Huang. Closed-loop open-vocabulary mobile manipulation with gpt-4v. In *International Conference on Robotics and Automation (ICRA)*, 2025. 2

[32] Jihong Zhu, Michael Gienger, Giovanni Franzese, and Jens Kober. Do you need a hand?–a bimanual robotic dressing assistance scheme. *Transactions on Robotics (T-RO)*, 40: 1906–1919, 2024. 2

[33] Ziwen Zhuang, ShenZhe Yao, and Hang Zhao. Humanoid parkour learning. In *Conference on Robot Learning (CoRL)*, 2024. 2