2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

COLA: Learning Human-Humanoid Coordination for Collaborative Object Carrying

Yushi Du [*,1,2], Yixuan Li [*,3,2], Baoxiong Jia [*,†,2], Yutang Lin [4,2], Pei Zhou [1]
Wei Liang [†,3,5], Yanchao Yang [†,1], Siyuan Huang [†,2]
https://yushi-du.github.io/COLA/

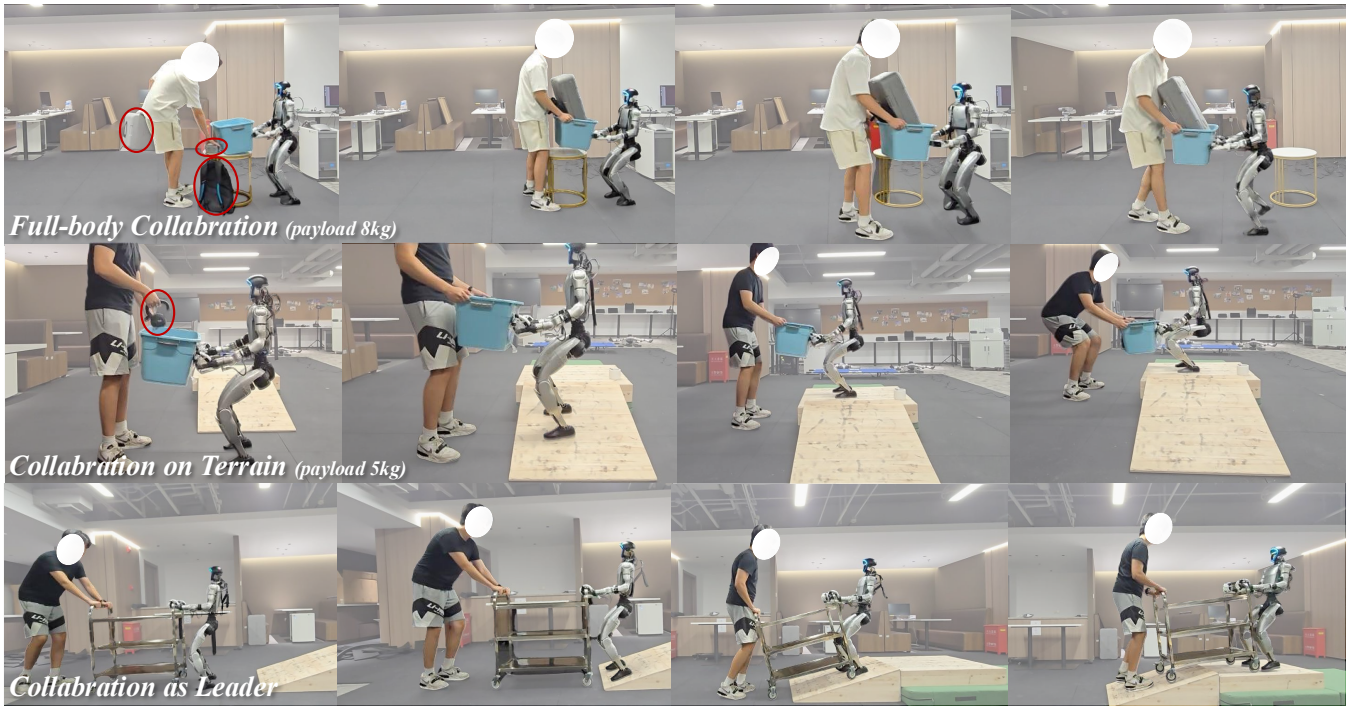


Fig. 1: **COLA** provides a proprioception-only policy that enables compliant human-humanoid collaboration for carrying diverse objects across various movement patterns. (a) demonstrates horizontal trajectory coordination where the humanoid adapts to human motion patterns. (b) illustrates whole-body coordination during collaborative object-lowering tasks. (c) shows **COLA** acting as the leader in collaboration 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. In this paper, we propose a proprioception-only reinforcement learning approach, COLA, that combines leader and follower behaviors within a single policy. The model is trained in a closed-loop environment with dynamic object interactions to predict object motion patterns and human intentions implicitly, enabling compliant collaboration to maintain load balance through coordinated trajectory planning. We evaluate our approach through comprehensive simulator and real-world experiments on collaborative carrying tasks, demonstrating the effectiveness, generalization, and robustness of our model across various terrains and objects. Simulation experiments demonstrate that our model reduces human effort by 24.7% compared to baseline approaches while maintaining object stability. Real-world experiments validate robust collaborative carrying across different object types (boxes, desks, stretchers, *etc*.) and movement patterns (straight-line, turning, slope climbing). Human user studies with 23 participants confirm an average improvement of 27.4% compared to baselines. Our method enables compliant human-humanoid collaborative carrying without requiring external sensors or complex interaction models, offering a practical solution for real-world deployment.**

I. Introduction

Recent years have witnessed significant progress in humanoid robot development, including agile locomotion [8,

∗ Equal contributions.
† Corresponding authors. Emails: jiabaoxiong@bigai.ai, liangwei@bit.edu.cn, yanchaoy@hku.hk, and syhuang@bigai.ai
[1] The University of Hong Kong [2] State Key Laboratory of General Artificial Intelligence, Beijing Institute for General Artificial Intelligence (BIGAI) [3] Beijing Institute of Technology [4] Yuanpei College, Peking University [5] Yangtze Delta Region Academy of Beijing Institute of Technology
This work is supported by the Early Career Scheme of the Research Grants Council (grant # 27207224), the National Natural Science Foundation of China (NSFC) under Grant No.62172043, and in part by the JC STEM Lab of Autonomous Intelligent Systems funded by The Hong Kong Jockey Club Charities Trust.

15, 29, 34], teleoperation [12, 27], and dexterous manipulation [13, 20, 28]. While these advances highlight the growing versatility and robustness of humanoid control, progress in enabling humanoid robots to collaborate effectively with humans remains limited. Human-robot collaboration remains challenging [5, 18, 19, 26, 32, 33], as it requires modeling diverse human behaviors, adaptive responses to physical interactions, and coordinated planning for shared tasks.

Object carrying [2, 3, 7] has become a representative task for advancing human-robot collaboration. Its core challenges arise from adapting to diverse environments (*e.g.*, maintaining stable support of objects across varying terrains), responding compliantly to human motions (*e.g.*, standing up together) often with limited or no direct force sensing, and dynamically allocating roles such as leading or following. These interdependent requirements make the task particularly difficult for humanoids, as collaboration requires integrating all aspects to ease the human partner, rather than addressing a single constraint as in prior work on environment-conditioned locomotion [15, 30, 34], compliance behavior learning [24–26, 31], or high-level intention prediction in open-loop object-finding or serving tasks [6, 14, 21, 32]. Thus, we propose a policy that unifies interactions, implicit constraints, and dynamic coordination into a coherent framework for humanoid collaborative carrying.

To address these challenges, we propose a learning-based policy for human-humanoid collaborative carrying that leverages reinforcement learning to model dynamic and versatile interactions. Our policy enables humanoid robots to share physical loads with human partners compliantly while facilitating seamless transitions between leader and follower roles. This role allocation is modulated via a velocity command, where a zero-velocity input implicitly designates a following behavior. Our design is built on two key 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. To incorporate these, we adopt a teacher-student framework in which the teacher policy, trained with both proprioceptive and privileged object-state information, is learned with rewards on the humanoid motion (*e.g.*, robust locomotion on terrains) and the object status (*e.g.*, maintaining a stretcher). The student policy, distilled from the teacher, relies solely on proprioceptive inputs for real-world inference and deployment.

We conduct extensive simulation experiments to demonstrate that **COLA** reduces human effort by 31.47% in collaborative carrying tasks compared to baseline approaches. Trajectory analysis shows that our method achieves 10.2 cm/s mean linear velocity tracking error and 0.1 rad/s mean angular tracking error relative to human motion, indicating precise coordination. Real-world experiments validate that **COLA** successfully tracks human movement patterns while assisting with object lifting, lowering, and transport along both straight and curved trajectories. Human user studies with 23 participants **confirm compliant collaborative carrying across diverse scenarios, demonstrating the practical effectiveness of our approach for human-robot collaborative object transport**.

Overall, our contributions can be summarized as follows:

- We propose a unified residual model that relies solely on proprioception for whole-body collaborative carrying, enabling compliant, coordinated, and generalizable collaboration across diverse movement patterns.
- We develop a three-step training framework and closed-loop training environment that explicitly models humanoid-object interactions, enabling the robot to implicitly learn object movements and assist humans through compliant collaboration.
- We demonstrate our policy in simulation and real-world settings, where our method achieves superior effort reduction and trajectory coordination compared to baseline approaches. The human user study confirms that our model achieves more compliant collaboration.

II. Related Works

A. Robot-human Collaboration

Robot-human collaboration has been a long-standing research topic spanning from robotic arms to legged robots [10, 14, 21]. While robotic arm-based collaboration systems typically assist humans in confined workspaces, researchers increasingly employ humanoid robots to provide assistance in open environments, leveraging their superior mobility and human-like morphology. However, current human-humanoid collaboration methods [5, 18, 33] rely primarily on model-based approaches. Existing works [1, 2] use heuristic rules that predefine a set of subtasks, including basic walking patterns, and identify primitive behaviors necessary for collaborative carrying. Some works [16, 17] focus on predicting human intention from multi-modal data and performing subtasks accordingly. While H2-COMPACT [3] proposes a learning-based model that uses haptic cues to predict horizontal velocity commands, it still operates with limited scope. All these approaches neglect whole-body coordination capabilities in human-humanoid collaborative carrying [23], thus lacking the ability to perform complex collaborative tasks such as picking up objects from the ground or carrying objects while climbing slopes. In this work, we propose a residual learning framework that enables humanoids to collaborate with humans using whole-body coordination, significantly broadening the range of collaborative carrying scenarios that humanoid robots can handle.

B. Compliant Whole-body Control

Position-only control lacks the compliance required for human-humanoid interaction [9, 26], as it operates without force awareness. Force regulation is crucial for collaborative tasks [11], particularly those involving human contact. Recent research [31] has demonstrated the effectiveness of force and compliance control in contact-rich manipulation tasks [4]. These methods explicitly estimate contact forces and integrate them into control policies, achieving improved performance on tasks such as force tracking and compliant responses to varied force and position inputs. Moreover,

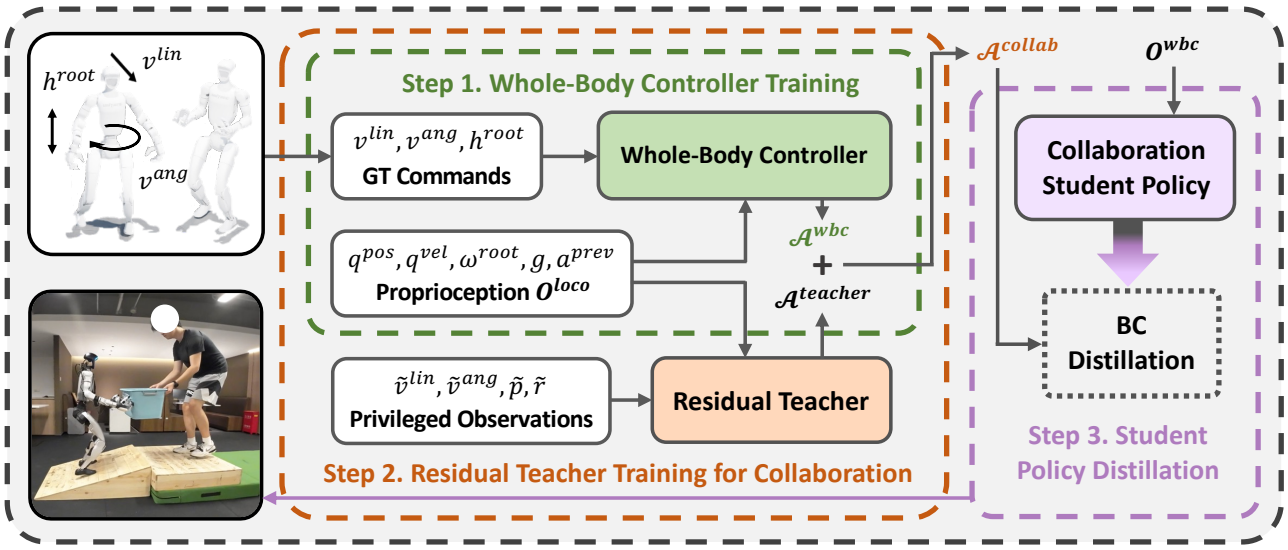


Fig. 2: **Overview of COLA.** Our Policy mainly consists of three steps: (i) We train a base whole-body control policy to provide a robust whole-body controller. (ii) In the closed-loop training environment, we train a residual teacher policy on top of the whole-body control policy with privileged information for human-humanoid collaboration. (iii) We distill the knowledge from the teacher policy into a student policy for real-world deployment using behavioral cloning.

other approaches [25] learn force characteristics implicitly to enable compliant and force-adaptive behaviors on legged robots. While these advances demonstrate that incorporating force feedback provides significant advantages for robotic interaction tasks, how force-aware control benefits human-humanoid collaboration remains underexplored. Building on these insights, we implicitly incorporate force considerations into our human-humanoid collaborative framework, enabling more natural and intuitive cooperative interactions.

III. Methodology

A. Overview

We define the task of 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. We assume the human partner is engaged in carrying the object, and the robot's objectives are three-fold: (i) to coordinate its movement by aligning with the human's velocity, (ii) to support the object's weight, thereby reducing the human's physical burden, and (iii) to stabilize the object's orientation throughout the transportation.

Our training pipeline is composed of three distinct learning steps: **a. Whole-body controller training**, **b. Residual teacher policy training for collaboration**, and **c. Student policy distillation**. We elaborate on this three-step design in the subsequent section.

B. Whole-body Control Policy

In the first step, we train a robust Whole-Body Control (WBC) policy within the simulator, establishing a foundational motion repertoire without additional task-specific constraints. The goal command is defined as $\mathcal{G} = [\mathcal{G}^{\text{lower}}, \mathcal{G}^{\text{upper}}]$, where $\mathcal{G}_t^{\text{lower}} \triangleq [v_t^{\text{lin}}, v_t^{\text{ang}}, h_t^{\text{root}}]$ denotes the lower-body locomotion targets, comprising linear velocity, angular velocity, and root height, and $\mathcal{G}_t^{\text{upper}} = [p^{\text{ee}}, r^{\text{ee}}]$ specifies the upper-body end-effector objectives for target position and rotation The observation space $\mathcal{O}_t^{\text{wbc}}$ is designed to capture temporal dynamics through a history window of length l:

$$\mathcal{O}_t^{\text{wbc}} \triangleq [q_{t-l:t}^{\text{pos}}, q_{t-l:t}^{\text{vel}}, \omega_{t-l:t}^{\text{root}}, g_{t-l:t}, a_{t-(l+1):t-1}^{\text{prev}}], \quad (1)$$

where $q_t^{\text{pos}} \in \mathbb{R}^N$ and $q_t^{\text{vel}} \in \mathbb{R}^N$ represent joint positions and velocities for $N = 29$ joints, respectively; $\omega_t^{\text{root}} \in \mathbb{R}^4$ denotes the root orientation; $g_t \in \mathbb{R}^3$ signifies the gravity vector in the robot's root frame; and a^{prev} represents the history of prior actions. The action space, $\mathcal{A}^{\text{wbc}}$, corresponds to the target joint positions of the humanoid, excluding the fingers.

We employ the Proximal Policy Optimization (PPO) algorithm to train our collaboration-ready WBC policy. Specifically, the WBC policy is trained to map a goal command $\mathcal{G}$, comprising lower-body locomotion targets $[v^{\text{lin}}, v^{\text{ang}}, h^{\text{root}}]$ and upper-body end-effector pose objectives $[p^{\text{ee}}, r^{\text{ee}}]$, into joint-level actions. The end-effector pose commands are sampled from a task-specific distribution detailed in Sec. IV. The policy is formally defined as:

$$\mathcal{F}^{\text{wbc}} : \mathcal{G} \times \mathcal{O}^{\text{wbc}} \to \mathcal{A}^{\text{wbc}}, \quad \mathcal{A}^{\text{wbc}} \in \mathbb{R}^N. \quad (2)$$

We train the WBC policy with rewards following prior works [22, 30]. To improve the robustness under payloads, we apply external forces to the humanoid's end-effectors during training, improving its force-adaptive capabilities.

C. Residual Teacher Policy

Effective human-humanoid collaboration requires real-time adaptation to the shifting dynamics of a shared object. We formulate the collaboration policy as a residual layer on the top of the base WBC policy, enabling the model to implicitly infer collaborative intent from physical

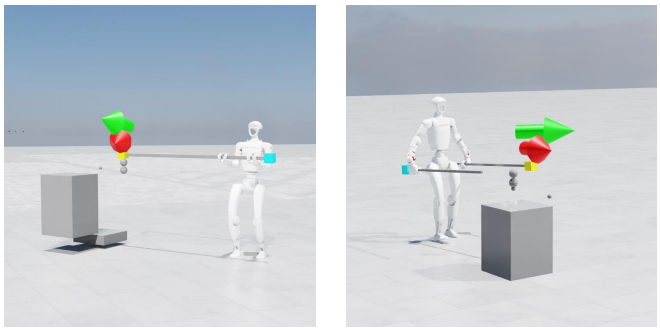


Fig. 3: **Closed-loop Training Environment**. This figure illustrates our closed-loop training environment in simulation. The green arrow represents the goal velocity of the carried object, while the red arrow indicates its current velocity.

interactions. This implicit learning paradigm is chosen for two reasons: (i) the exploration space of the base WBC may not fully encompass the complex motion constraints of real-world collaboration, and (ii) the nuances of physical interaction are difficult to encode via explicit, manually designed commands.

In the second training phase, we introduce a closed-loop environment to explicitly model the dynamic interactions between the human, the object, and the humanoid. By attaching an object to the robot's hands, the end-effectors are physically constrained by the object's geometry and the human partner's motions. Under these constraints, we train a residual teacher policy, $\mathcal{F}^{\text{teacher}}$, which receives an augmented observation space $\mathcal{O}_t^{\text{teacher}} \triangleq [\mathcal{O}_t^{\text{wbc}}, \mathcal{O}_t^{\text{priv}}]$. The privileged component, $\mathcal{O}_t^{\text{priv}}$, provides ground-truth object states over a history window l:

$$\mathcal{O}^{\text{priv}}t \triangleq [\tilde{v}^{\text{lin}}t - l : t, \tilde{v}^{\text{ang}}t - l : t, \tilde{p}t - l : t, \tilde{r}_{t-l:t}], \quad (3)$$

which comprises the linear and angular velocities, position, and orientation of the object.

The teacher policy leverages both privileged information and robot proprioception ($\mathcal{O}^{\text{wbc}}$) to output a residual action, $\mathcal{A}^{\text{teacher}}$, representing a corrective adjustment to the WBC output. This ensures compliant load sharing and precise trajectory coordination. The policy and resulting collaborative action, $\mathcal{A}^{\text{collab}}$, are defined as:

$$\begin{aligned} &\mathcal{F}^{\text{teacher}} : [\mathcal{O}^{\text{wbc}}, \mathcal{O}^{\text{priv}}] \rightarrow \mathcal{A}^{\text{teacher}}, \quad \mathcal{A}^{\text{teacher}} \in \mathbb{R}^N, \\ &\mathcal{A}^{\text{collab}} = \mathcal{A}^{\text{wbc}} + \mathcal{A}^{\text{teacher}}. \end{aligned} \quad (4)$$

During this step, the goal command of the model is modified based on the settings described in Sec. IV. The teacher's learning is guided by a composite reward function that combines base whole-body control rewards with task-specific rewards, as detailed in Tab. I.

D. Knowledge Distillation

In the distillation step, we transfer the expertise of the teacher policy ($\mathcal{F}^{\text{wbc}} + \mathcal{F}^{\text{teacher}}$) into a student policy, $\mathcal{F}^{\text{student}}$, designed for real-world deployment. This student policy operates without access to privileged information, relying solely on the proprioceptive observations, $\mathcal{O}^{\text{wbc}}$:

$$\mathcal{F}^{\text{student}} : \mathcal{O}^{\text{wbc}} \rightarrow \mathcal{A}^{\text{student}}, \text{ where } \mathcal{A}^{\text{student}} \in \mathbb{R}^N. \quad (5)$$

We use behavioral cloning to distill the teacher policy into a student policy, training the student to mimic the teacher by minimizing the mean squared error between their outputs during interactions with the environment. The optimization is defined by the following loss function:

$$\mathcal{L}_{\text{distill}} = \mathbb{E}\left[\left\|\mathcal{A}^{\text{student}} - \mathcal{A}^{\text{collab}}\right\|^2\right]. \quad (6)$$

Furthermore, we define two distinct experimental settings based on whether the model observes the goal command during collaboration: **COLA-F (Follower)** and **COLA-L (Leader)**. In the **COLA-F** setting, all networks receive a goal command input of zero, whereas in the **COLA-L** setting, the policy is provided with a sampled goal command within the same range used for the whole-body controller.

TABLE I: **Reward Functions** for collaboration policy training.

Term	Expression	Weight
Linear 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	$-\|\mathcal{F}^{\text{support-obj}}\|$	0.002

*Note: $v_{\text{lin}}^{\text{applied}}$ is the applied end-point linear-velocity(Sec. IV); $v_{\text{ang}}^{\text{goal}}$ is the target angular velocity of the object; h_1^{obj}, h_2^{obj} are the predefined heights of the object's far ends; $\mathcal{F}^{\text{support-object}}$ is the horizontal force between the support body and object; $\phi(x) = e^{-\|x\|}$.

IV. Implementation Details

A. Training Setup

We train our policy in Isaac Lab on a single RTX 4090D GPU using PPO with 4096 parallel environments. The actor and critic networks for the base WBC policy are three-layer Multi-Layer Perceptrons (MLPs) of size (512, 256, 128), while the residual teacher and studnet policy network employs two additional MLPs with the same dimensions. The training of the WBC, residual teacher, and distillation policies takes 350k, 250k, and 250k environment steps, respectively, which correspond to abot 15k, 10k, and 10k PPO update steps. The full training time is 48 hours.

B. Observation Space Details

We sample WBC commands from a pre-defined range. The end-effector goal command, representing the 6−DoF target pose of the robot's wrist, is generated using Spherical Linear Interpolation (SLERP). Since our task focuses on human-humanoid collaborative carrying rather than complex upper-body manipulation, the robot primarily needs to execute fine-scale arm adjustments to modify the object's pose and velocity. Therefore, we do not sample large-range upper-body motions. Instead, we sample end-effector goal commands in the vicinity of the default grasping pose, with positions randomly sampled within a small cubic region and

orientations within a conical region around the nominal grasp orientation. Our whole-body controller achieves a tracking error of $5.6cm$ for end-effector goal position and 7° for end-effector goal orientation.

C. Closed-loop Training Environment

To precisely simulate the dynamic interactions in real-world collaborative carrying tasks, we set up a closed-loop training environment that explicitly models the collaboration setting as illustrated in Fig. 3. The environment consists of the humanoid, a supporting body that simulates the human carrier, and the carried object to be transported, which is connected to the support body via a 6-DoF joint. Once the environment is initialized, the object is placed in the robot's hand while the hands are fixed in a predefined grasp pose.

During training, we randomly sample a goal command $\mathcal{G}$ including the lower and upper-body motion of the robot and the target dynamics of the object. We also sample a velocity v^{applied} and apply it to the supporting base body at the end of the object opposite the robot-held end. The held end is predetermined for each object type (*e.g.*, the far side of a box or the handle of a stretcher). We sample the magnitude of v^{applied} from the range $(0, \mathcal{G})$ with added random noise and update it at twice the frequency of the goal command $\mathcal{G}$. For angular velocity control, we set a target angular velocity and use a PD controller to apply torque to the support body. For height control, we randomly sample a target height for the support body within a predefined range and apply a PD-controlled force to adjust its height accordingly, without requiring the robot to maintain a fixed height.

It is worth noting that although the carried object and the support body are connected through a 6-DoF joint, the inherent friction, damping, and joint limits ensure that any movement of the supporting base body directly affects the object. In this way, the dynamics of the support body are faithfully transmitted to the carried object.

The optimization objectives are summarized in Tab. I.

V. Experiments

A. Overview

In this section, we conduct experiments to evaluate the effectiveness of **COLA**. We aim to address the following key research questions via empirical analysis and discussion:

- Does the residual teacher policy and the distillation training enable effective and compliant collaboration?
- Is the architecture of **COLA** designed in a concise and effective manner?
- Do the results demonstrate practical value in real-world scenarios, such as assisting humans in object transportation and reducing physical effort?

B. Baselines

1) **Vanilla MLP**: We implement the policy as an MLP, initialize it with the weights of the whole-body controller, and train it end-to-end with PPO.
2) **Explicit Goal Estimation**: We replace the whole-body control command with the predicted one, remove the residual component from the teacher policy, and distill the resulting policy into the student.
3) **Transformer**: We replace the student policy's original architecture with a Transformer.

C. Metrics

We use the following metrics to evaluate the performance of the proposed method in terms of trajectory following, height tracking, and coordination with the human: *Linear velocity tracking error (Lin. Vel.)*: Mean linear velocity tracking error relative to the human over the entire episode. *Angular velocity tracking error (Ang. Vel.)*: Mean angular velocity tracking error relative to the human over the entire episode. *Height Error (Height Err.)*: Height tracking error between the object ends held by the human and humanoid, measuring stability of height coordination. *Average external force (Avg. E.F.)*: Average horizontal interaction force between the human and the object, reflecting human effort required to move the carried object along the intended direction.

D. Do the residual teacher policy and the distillation training contribute to effective and compliant collaboration?

We conduct simulation experiments to evaluate the performance of our model against baseline methods and analyze the choice of model architecture. As shown in the upper part of Tab. II, our method outperforms all baselines across the evaluated metrics. The superior *Lin. Vel.*, *Ang. Vel.*, and *Height Err.* indicate that our model achieves better collaboration with humans, while the highest *Avg. E.F.* reflects the strongest compliance exhibited by the model.

Although the *Vanilla MLP* achieves relatively high performance among the baseline models, it struggles to accurately track *Ang. Vel.* and *Height Err.*. This result indicates that linear movements are easier to infer compared to angular and vertical movements. The **teacher-student distillation framework offers a promising approach to learn these complex interaction patterns using privileged information** that is difficult to acquire directly.

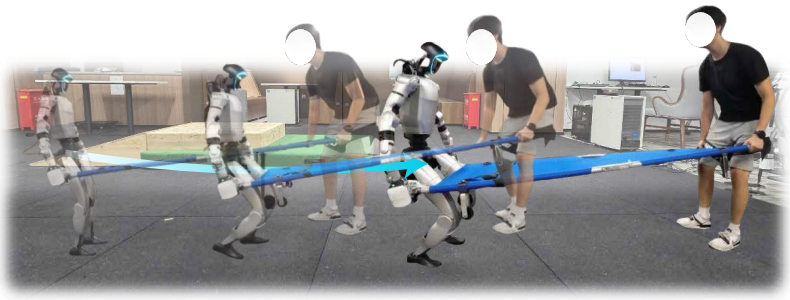
While the *Explicit Goal Estimation* baseline performs the poorest, this result highlights that collaborative carrying is not merely a task of predicting whole-body control commands. The dynamic interactions among the humanoid, object, and human introduce additional challenges for human-humanoid collaboration. Consequently, **implicitly learning object dynamics and human movements within a closed-loop environment proves more effective for maintaining object stability and achieving coordinated collaboration**.

E. Is the architecture of **COLA** designed in a concise and effective manner?

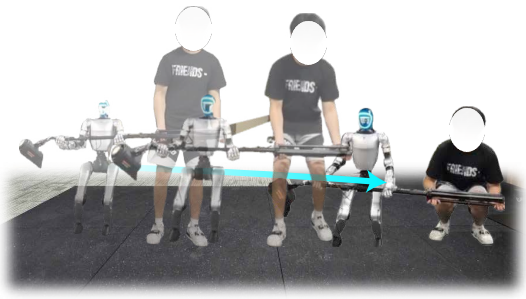
The lower part of Tab. II presents an ablation study on different choices of our model architecture of the student policy. **COLA** outperforms the *Transformer*, which requires twice the number of training steps to converge, demonstrating that **a compact model can achieve superior performance**. This is likely because the Transformer's long-term temporal processing introduces unnecessary complexity, whereas the

TABLE II: **Quantitative evaluation in simulation.** We report results on velocity and height tracking, as well as the average external force between the robot and the carried object, to evaluate the effort required for joint carrying and movement.

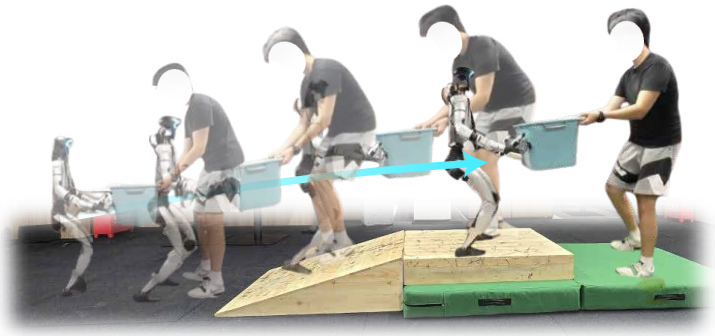
Methods	Lin. Vel. $(m/s)\downarrow$	Ang. Vel. $(rad/s)\downarrow$	Height Err. $(m)\downarrow$	Avg. E.F. $(N)\downarrow$
Explicit Goal Estimation	0.235	0.335	0.102	19.067
Transformer	0.178	0.310	0.077	19.382
COLA-F-History10	0.121	0.131	0.037	15.435
COLA-F-History50	0.116	0.132	0.036	14.574
COLA-F	0.109	0.118	**0.031**	14.576
COLA-L-History10	0.118	0.106	0.039	13.924
COLA-L-History50	0.112	0.103	0.036	13.495
COLA-L	**0.102**	**0.098**	0.038	**12.298**



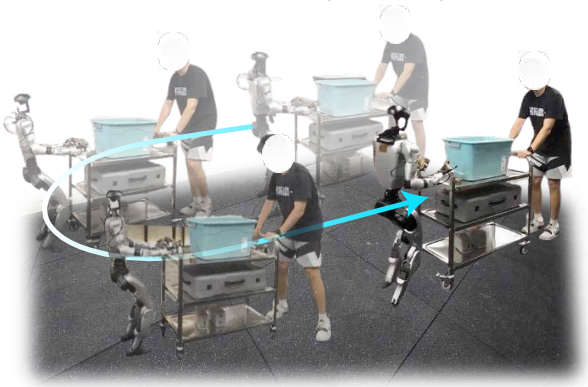
(a) Stretcher Carrying



(b) Rod Height Tracking



(c) Box Carrying on Slope



(d) Cart Pushing

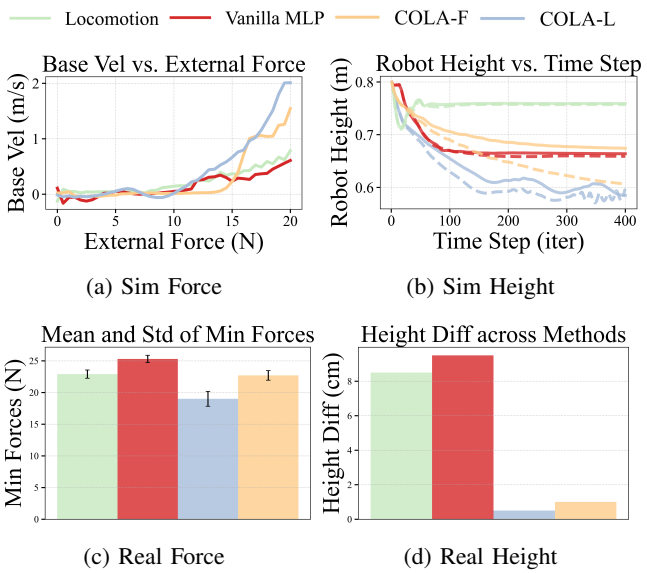
Fig. 4: **Qualitative visualizations of human-humanoid collaboration using COLA.** This figure showcases our model's ability to carry diverse objects and perform collaborative skills such as horizontal-velocity and height tracking, including under challenging conditions such as sloped terrains. The interacted objects include a 3kg rod, an 8kg box, an 11kg stretcher, and a 20kg cart, demonstrating both versatility and generalizability of the proposed method.

MLP-based model adapts more promptly to diverse human movements, which is crucial for collaborative tasks. For example, when the robot collaboratively carries an object with a human and they begin moving from a stationary state, the robot should focus on the current object motion rather than earlier frames that encouraged it to stay still. Relying on outdated information can cause the robot to hesitate between continuing its movement or stopping, thereby degrading the smoothness of human–humanoid cooperation.

We also ablate the history length of **COLA**. We find that a shorter history provides insufficient information for the policy to implicitly learn collaboration with human movements from state observations. Increasing the history length to 50 yields little improvement. We therefore select 25 as a balance between performance and learning efficiency. This suggests that the task is not highly sensitive to long-term joint position changes, consistent with our earlier findings on the Transformer-based student policy baseline.

F. How compliant is the model to external forces in both simulation and real-world experiments?

The results in Fig. 5a and Fig. 5b illustrate how the velocity and height respond to external forces applied along the humanoid's x-axis. In Fig. 5a, the baseline model remains nearly stationary, whereas **COLA** begins to follow the external force once it exceeds 15N. Forces below 15N are interpreted as cues for adjusting motion to stabilize the humanoid rather than for initiating movement. Fig. 5b illustrates how the humanoid responds to externally applied



(a) Sim Force (b) Sim Height

(c) Real Force (d) Real Height

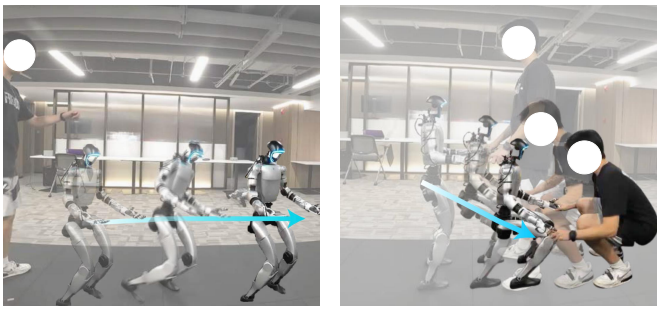
Fig. 5: **Quantitative Results on the Effectiveness of Collaborative Carrying.** (a) illustrates the robot's velocity with force applied to the hands of the robot, where the force is linearly increased throughout a time sequence of 10.0 seconds. (b) illustrates the height of the robot's pelvis over time when applying vertical external forces. Solid lines show the robot's height with 10N applied to each hand, whereas dashed lines show the height with 20N in the simulator. (c) illustrates the minimal force required to move the robot in the real world. (d) illustrates the height difference between the human-held end and the humanoid-held end of the object in real-world experiments.

vertical forces. The *Locomotion* policy maintains a nearly constant height under the applied force, while the *Vanilla MLP* squats to a fixed height regardless of the magnitude of the ascending force. This indicates that the *Vanilla MLP* only supports the external force without actively complying in the vertical direction. In contrast, both **COLA** settings effectively respond to the applied force, demonstrating agile full-body motions that comply with vertical disturbances.

It is observed that **COLA-L** consistently outperforms **COLA-F**. We attribute this to the **goal command, which helps the policy learn to collaborate with humans more actively and precisely**. When noise and disturbances are consistently present in dynamic interactions, the goal command provides additional informative cues that enhance human-humanoid collaboration.

G. Do the results demonstrate practical value in real-world scenarios?

We also conduct real-world experiments to validate the effectiveness of our approach in practical scenarios. Qualitative results are shown in Fig. 4, demonstrating that our method achieves promising performance in carrying diverse objects (*e.g.*, boxes, carts, and stretchers) across various grasping poses and terrains. Additionally, we demonstrate that our model implicitly learns to interpret human intentions through force-based interaction. When a human applies directional forces to guide the robot, the humanoid infers the desired movement command and continues executing that motion autonomously. This force-aware adaptation capability broadens the range of intuitive human-robot collaboration scenarios which highlight performance that surpasses existing methods.



(a) Kicking torso (b) Dragging hand

Fig. 6: **Movement Analysis.** When a continuous external force is applied to the robot's torso, it resists to maintain a stable stance. In contrast, when a smaller force is applied to the robot's end-effector, it tends to follow the force.

We quantitatively evaluate our model using two metrics: *Min. Force*: The minimum force required to move the robot, reflecting the robot's compliance during collaborative transportation. *Height Diff.* The difference in height between the two ends of the carried object, held by the human and the humanoid, reflecting the height-tracking performance.

As shown in Fig. 5c and Fig. 5d, our model **reduces height-tracking error by approximately** 75% and demonstrates stronger compliance in following external forces compared to the baseline. These findings highlight that our model provides practical assistance for collaborative carrying tasks by maintaining stable object pose tracking while reducing the physical effort of the human operator.

We also recruited 23 participants to rate the compliance and height-tracking ability of **COLA** during the carrying process on a scale from 1 to 5. To reduce personal bias and ensure fair evaluation, we recorded the collaboration videos, shuffled their order, and published them online for participants to provide ratings. The results shown in Tab. III indicate that **COLA** gets the best results in both metrics. These findings further demonstrate that our model provides effective assistance to humans in real-world scenarios.

H. Implicitly estimating interaction forces from joint states.

We observe that the humanoid's behavior during transportation is primarily sensitive to forces applied at specific joints. As shown in Fig. 6, when humans apply forces to the hand or arm during carrying, the humanoid tends to follow the human's lead. In contrast, when forces are applied to other joints, such as the torso or legs, the humanoid remains stable. These results demonstrate that our model **effectively learns the interaction dynamics among the humanoid, object, and human through the offsets between joint states and their targets**.

VI. Conclusions and Limitations

We present a unified three-step residual learning framework for human-humanoid collaboration, enabling the robot to act as either a leader or a follower. By distilling privileged

TABLE III: **User study results** evaluated by 23 participants on the performance of *Height Tracking* and *Smoothness*.

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

object data into a student policy, our method achieves compliant, whole-body coordination using only proprioceptive feedback, eliminating the need for external sensors. A closed-loop training environment explicitly models robot-object interactions, allowing the humanoid to implicitly infer human movement and adapt.

While effective, future work could enhance coordination by integrating multi-modal perception, such as visual and tactile sensing. Additionally, developing autonomous planning for proactive human assistance remains a key research direction. We hope this framework advances human-humanoid interaction and inspires new collaborative applications.

REFERENCES

[1] D. J. Agravante, A. Cherubini, A. Bussy, P. Gergondet, and A. Kheddar, "Collaborative human-humanoid carrying using vision and haptic sensing," in *International Conference on Robotics and Automation (ICRA)*, 2014. 2

[2] D. J. Agravante, A. Cherubini, A. Sherikov, P.-B. Wieber, and A. Kheddar, "Human-humanoid collaborative carrying," *Transactions on Robotics (T-RO)*, vol. 35, no. 4, pp. 833–846, 2019. 2

[3] G. C. R. Bethala, H. Huang, N. Pudasaini, A. M. Ali, S. Yuan, C. Wen, A. Tzes, and Y. Fang, "H2-compact: Human-humanoid co-manipulation via adaptive contact trajectory policies," in *International Conference on Humanoid Robots*, 2025. 2

[4] Y. Du, R. Wu, Y. Shen, and H. Dong, "Learning part motion of articulated objects using spatially continuous neural implicit representations," in *British Machine Vision Conference (BMVC)*, 2023. 2

[5] M. L. Elwin, B. Strong, R. A. Freeman, and K. M. Lynch, "Human-multirobot collaborative mobile manipulation: The omnid mocobots," *IEEE Robotics and Automation Letters (RA-L)*, vol. 8, no. 1, pp. 376–383, 2022. 2

[6] X. Gao, L. Yan, G. Wang, and C. Gerada, "Hybrid recurrent neural network architecture-based intention recognition for human–robot collaboration," *IEEE Transactions on Cybernetics*, vol. 53, no. 3, pp. 1578–1586, 2021. 2

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

[8] Z. Gu, J. Li, W. Shen, W. Yu, Z. Xie, S. McCrory, X. Cheng, A. Shamsah, R. Griffin, C. K. Liu *et al.*, "Humanoid locomotion and manipulation: Current progress and challenges in control, planning, and learning," 2025. 1

[9] A. Hartmann, D. Kang, F. Zargarbashi, M. Zamora, and S. Coros, "Deep compliant control for legged robots," in *International Conference on Robotics and Automation (ICRA)*, 2024. 2

[10] C. Li, X. Wu, Y. Chen, T. Teng, X. Zhang, S. Calinon, D. Caldwell, and F. Chen, "Human-like robot impedance regulation skill learning from human-human demonstrations," *IEEE Transactions on Cognitive and Developmental Systems*, 2025. 2

[11] Y. Li, L. Zheng, Y. Wang, E. Dong, and S. Zhang, "Impedance learning-based adaptive force tracking for robot on unknown terrains," *Transactions on Robotics (T-RO)*, vol. 41, no. 3, pp. 1404–1420, 2025. 2

[12] Y. Li, Y. Lin, J. Cui, T. Liu, W. Liang, Y. Zhu, and S. Huang, "Clone: Closed-loop whole-body humanoid teleoperation for long-horizon tasks," in *Conference on Robot Learning (CoRL)*, 2025. 2

[13] Y. Lin, J. Cui, Y. Li, B. Jia, Y. Zhu, and S. Huang, "Lessmimic: Long-horizon humanoid interaction with unified distance field representations," 2026. 2

[14] H. Liu, Y. Tong, G. Liu, Z. Ju, and Z. 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

[15] L. Ma, Z. Meng, T. Liu, Y. Li, R. Song, W. Zhang, and S. Huang, "Styleloco: Generative adversarial distillation for natural humanoid robot locomotion," 2025. 1, 2

[16] D. Nicolis, A. M. Zanchettin, and P. Rocco, "Human intention estimation based on neural networks for enhanced collaboration with robots," in *International Conference on Intelligent Robots and Systems (IROS)*, 2018. 2

[17] O. Palinko, F. Rea, G. Sandini, and A. 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

[18] R. Rahem, C. Y. Wong, and W. Suleiman, "Human-humanoid robot cooperative load transportation: model-based control approach," in *International Conference on Intelligent Robots and Systems (IROS)*, 2022. 2

[19] L. Rozo, S. Calinon, D. G. Caldwell, P. Jimenez, and C. Torras, "Learning physical collaborative robot behaviors from human demonstrations," *Transactions on Robotics (T-RO)*, vol. 32, no. 3, pp. 513–527, 2016. 2

[20] R. Shah, S. Liu, Q. Wang, Z. Jiang, S. Kumar, M. Seo, R. Martín-Martín, and Y. Zhu, "Mimicdroid: In-context learning for humanoid robot manipulation from human play videos," in *International Conference on Robotics and Automation (ICRA)*, 2026. 2

[21] Y. Sheng, Y. Wang, H. Cheng, H. Zhao, and H. Ding, "Human-like robot action policy through game-theoretic intent inference for human-robot collaboration," *Transactions on Robotics (T-RO)*, 2025. 2

[22] S. Wandong, "Legged lab: Direct isaaclab workflow for legged robots," 2025. [Online]. Available: https://github.com/Hellod035/LeggedLab 3

[23] W. Wang, R. Li, Y. Chen, Z. M. Diekel, and Y. Jia, "Facilitating human–robot collaborative tasks by teaching-learning-collaboration from human demonstrations," *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 2, pp. 640–653, 2018. 2

[24] C. Y. Wong, L. Vergez, and W. Suleiman, "Vision-and tactile-based continuous multimodal intention and attention recognition for safer physical human–robot interaction," *IEEE Transactions on Automation Science and Engineering*, vol. 21, no. 3, pp. 3205–3215, 2023. 2

[25] B. Xu, H. Weng, Q. Lu, Y. Gao, and H. Xu, "Facet: Force-adaptive control via impedance reference tracking for legged robots," in *Conference on Robot Learning (CoRL)*, 2025. 2, 3

[26] X. Yu, B. Li, W. He, Y. Feng, L. Cheng, and C. Silvestre, "Adaptive-constrained impedance control for human–robot co-transportation," *IEEE Transactions on Cybernetics*, vol. 52, no. 12, pp. 13 237–13 249, 2021. 2

[27] Y. Ze, Z. Chen, J. P. AraÃšjo, Z.-a. Cao, X. B. Peng, J. Wu, and C. K. Liu, "Twist: Teleoperated whole-body imitation system," in *Conference on Robot Learning (CoRL)*, 2025. 2

[28] Y. Ze, Z. Chen, W. Wang, T. Chen, X. He, Y. Yuan, X. B. Peng, and J. Wu, "Generalizable humanoid manipulation with 3d diffusion policies," in *International Conference on Intelligent Robots and Systems (IROS)*, 2025. 2

[29] T. Zhang, B. Zheng, R. Nai, Y. Hu, Y.-J. Wang, G. Chen, F. Lin, J. Li, C. Hong, K. Sreenath *et al.*, "Hub: Learning extreme humanoid balance," *Conference on Robot Learning (CoRL)*, 2025. 1

[30] Y. Zhang, Y. Yuan, P. Gurunath, T. He, S. Omidshafiei, A. akbar Agha-mohammadi, M. Vazquez-Chanlatte, L. Pedersen, and G. Shi, "Falcon: Learning force-adaptive humanoid loco-manipulation," in *Annual Learning for Dynamics & Control Conference*, 2026. 2, 3

[31] P. Zhi, P. Li, J. Yin, B. Jia, and S. Huang, "Learning unified force and position control for legged loco-manipulation," in *Conference on Robot Learning (CoRL)*, 2025. 2

[32] P. Zhi, Z. Zhang, Y. Zhao, M. Han, Z. Zhang, Z. Li, Z. Jiao, B. Jia, and S. Huang, "Closed-loop open-vocabulary mobile manipulation with gpt-4v," in *International Conference on Robotics and Automation (ICRA)*, 2025. 2

[33] J. Zhu, M. Gienger, G. Franzese, and J. Kober, "Do you need a hand?–a bimanual robotic dressing assistance scheme," *Transactions on Robotics (T-RO)*, vol. 40, pp. 1906–1919, 2024. 2

[34] Z. Zhuang, S. Yao, and H. Zhao, "Humanoid parkour learning," in *Conference on Robot Learning (CoRL)*, 2024. 1, 2