



CLONE: Closed-Loop Whole-Body Humanoid Teleoperation for Long-Horizon Tasks

Yixuan Li^{*,1,2}

lyx@bit.edu.cn

Tengyu Liu^{2,7}

liutengyu@bigai.ai

Yutang Lin^{*,3,4,5,6}

yutang.lin@stu.pku.edu.cn

Jieming Cui^{2,3,4,6}

cuijieming@stu.pku.edu.cn

Wei Liang^{✉,1}

liangwei@bit.edu.cn

Yixin Zhu^{✉,3,4,6,8}

yixin.zhu@pku.edu.cn

Siyuan Huang^{✉,2,7}

syhuang@bigai.ai

* Equal contributors. ¹ School of Computer Science and Technology, Beijing Institute of Technology

² State Key Laboratory of General Artificial Intelligence, BIGAI

³ School of Psychological and Cognitive Sciences, Peking University

⁴ Institute for Artificial Intelligence, Peking University ⁵ Yuanpei College, Peking University

⁶ Beijing Key Laboratory of Behavior and Mental Health, Peking University

⁷ Joint Laboratory of Embodied AI and Humanoid Robots, BIGAI & UniTree Robotics

⁸ Embodied Intelligence Lab, PKU-Wuhan Institute for Artificial Intelligence

<https://humanoid-clone.github.io/>



Figure 1: **CLONE** employs an MoE-based policy to teleoperate a humanoid, enabling **long-horizon task execution** and **holistic humanoid-environment interactions**.

Abstract: Humanoid robot teleoperation plays a vital role in demonstrating and collecting data for complex interactions. Current methods suffer from two key limitations: (i) restricted controllability due to decoupled upper- and lower-body control, and (ii) severe drift caused by open-loop execution. These issues prevent humanoid robots from performing coordinated whole-body motions required for long-horizon loco-manipulation tasks. We introduce **CLONE**, a whole-body teleoperation system that overcomes these challenges through three key contributions: (i) a Mixture-of-Experts (MoE) whole-body control policy that enables complex coordinated movements, such as “picking up an object from the ground” and “placing it in a distant bin”; (ii) a closed-loop error correction mechanism using LiDAR odometry, reducing translational drift to 12cm over 8.9-meter trajectories; and (iii) a systematic data augmentation strategy that ensures robust performance under diverse, previously unseen operator poses. In extensive experiments, **CLONE** demonstrates robust performance across diverse scenarios while maintaining stable whole-body control. These capabilities significantly advance humanoid robotics by enabling the collection of long-horizon interaction data and establishing a foundation for more sophisticated humanoid-environment interaction in both research and practical applications.

1 Introduction

Humans can effortlessly perform tasks like navigating to distant locations or retrieving objects from the ground, owing to their precise and coordinated whole-body control. Humanoid robots, designed with a similar morphology, have demonstrated the ability to replicate various human-like movements through whole-body controllers. However, achieving seamless humanoid-environment interactions in long-horizon tasks remains challenging, particularly in maintaining accurate global positioning and coordinating complex whole-body movements.

Recent advances in humanoid whole-body control, including loco-manipulation and teleoperation [1–8] have made notable progress. Nevertheless, existing methods still struggle with precise teleoperation over extended durations and fall short of enabling the holistic control necessary for humanoid-environment interaction. Two key challenges persist in bridging the capability gap between humans and humanoids.

First, existing methods often neglect the critical coordination between upper- and lower-body movements and undermine the role of hand orientation in executing comprehensive and humanlike interactions. Some works proposed decoupled approaches that separately control the upper and lower body to ensure stability [6, 8]. This separation often leads to suboptimal coordination, limiting the humanoid’s ability to perform holistic locomotion tasks—capabilities that humans execute naturally. Other approaches focus on controlling humanoid robots by mimicking human motions captured through motion capture (MoCap) systems [1, 3, 7, 9–11]. These methods typically prioritize stability over full action-space utilization, with constraints implicitly imposed by the training data, thereby restricting the humanoid’s potential to execute dexterous full-body movements. Moreover, these approaches unanimously overlook hand orientation—a critical component for manipulation.

Second, all existing whole-body teleoperation systems operate in an open-loop manner without accurate global position feedback. This limitation stems from the inherent difficulty in accurately estimating a humanoid robot’s global position and orientation during locomotion. Unlike fixed-base systems or wheeled platforms with relatively simple odometry, humanoid robots experience complex, non-holonomic motion patterns with significant foot-ground interaction variations, making reliable state estimation particularly challenging. As the teleoperator guides the humanoid through the environment, small errors in global position and orientation accumulate with each step, resulting in significant drift over long-horizon tasks. This accumulated error eventually compromises the operator’s ability to accurately control the robot, particularly for manipulation tasks requiring precise positioning relative to environmental objects. Furthermore, existing systems often require complex motion capture setups or multiple sensors distributed across the operator’s body, creating barriers to intuitive and accessible teleoperation.

In this paper, we propose **CLONE**, a novel Mixture-of-Experts (MoE)-based closed-loop teleoperation system that bridges the capability gap between humans and humanoids through three key technical innovations:

- **Model Architecture:** We adopt an MoE framework that enables a unified policy to learn diverse motion skills while synthesizing lower-body motions coordinated with upper-body actions.
- **System Integration:** We incorporate LiDAR odometry [12] and Apple Vision Pro tracking to provide closed-loop global pose feedback, enabling real-time drift correction during teleoperation. Our system captures the operator’s head and hand poses as intuitive control signals.
- **Data Curation:** We curate a large-scale dataset **CLONED** by enhancing a subset of the AMASS dataset [13] with sampled hand orientations and additional motion-captured dataset, ensuring robust generalization to dexterous and dynamic whole-body motions.

Our extensive evaluations in both simulation and real-world settings demonstrate **CLONE**’s exceptional performance in long-horizon teleoperation and loco-manipulation tasks. Using only the opera-

tor’s head and hand positions captured by an Apple Vision Pro headset as control signals, the system exhibits positional drift of 12cm over a 8.9-meter teleoperation session—a substantial improvement over open-loop approaches. **CLONE** enables complex whole-body actions such as picking up objects from the ground while squatting, and produces smooth, stable walking regardless of stance and direction, highlighting its ability to significantly enhance motion quality and precision in long-horizon tasks while maintaining an intuitive control interface.

The contributions of this work include: (i) an **MoE-based teleoperation framework** that enables full action-space control with coordinated whole-body movements; (ii) a **closed-loop error correction system** that mitigates positional drift for long-horizon tasks; (iii) a **comprehensive training dataset** ensuring coverage of diverse full-body motions, particularly for dexterous loco-manipulation; and (iv) extensive empirical validation demonstrating **substantial improvements in real-world humanoid-environment interaction** capabilities.

2 Related Work

Humanoid Whole-Body Teleoperation Humanoid teleoperation enables robots to replicate human movements for complex task execution, using MoCap systems [11, 14, 15], haptic feedback devices [16–18], or virtual reality (VR) interface [3, 7, 19–22], with control policies balancing robot stability and motion tracking. However, current methods rely on typical control policies and struggle to reproduce the full diversity and fluidity of human motions [23], often due to monolithic Multi-Layer Perceptrons (MLP)-based architectures that inadequately handle conflicting objectives across diverse motion types (*e.g.*, walking vs. crouching) [24, 25]. While mixture-based models have shown promise in other domains [26–29], their potential for humanoid teleoperation remains underexplored.

Long-Horizon Tasks for Humanoid Long-horizon task execution has been extensively explored in robotics research, particularly for fixed-base robotic arm systems in structured environments [30–32]. Mobile manipulation has extended these capabilities, with recent work by [33] using a bimanual wheeled robot trained through teleoperation for household tasks. While humanoid teleoperation remains largely constrained to short-horizon motion replication [7, 10, 11], often in open-loop due to difficulties in real-time global state estimation for bipedal systems. Unlike wheeled systems, humanoids exhibit complex, non-holonomic locomotion and foot-ground dynamics that hinder odometry. Although multi-sensory odometry has shown promise in legged robots [34, 35], its integration into humanoid teleoperation for long-horizon, demonstration-rich tasks remains underexplored.

Humanoid Learning Dataset Large-scale motion capture datasets such as AMASS [13] and LAFAN1 [36] have been instrumental in training low-level humanoid control policies [3, 7, 10, 37]. To improve motion diversity, ExBody2 [11] filtered AMASS and applied a Conditional Variational Autoencoder (CVAE)-based generative model to synthesize novel trajectories. However, these datasets were primarily designed for animation and graphics [38], rather than robotics-specific needs. This creates an inherent domain gap in motion representation. Although they contain semantically distinct actions (*e.g.*, waving, hugging, drinking), they underrepresent the breadth of kinematic configurations and transitions required for robust and generalizable controller training. This disparity between available human motion data and the needed for effective humanoid control remains a significant challenge in the field.

3 Method

In this section, we provide an overview of our teleoperation framework, as illustrated in Fig. 2. Our system captures a minimal set of control signals from the teleoperator, consisting solely of the 6D poses (position and orientation) of the head and both wrists, tracked using an Apple Vision Pro headset. These three points serve as the complete control interface, providing an intuitive yet powerful means of directing the humanoid’s full-body motion while maintaining a simple setup.

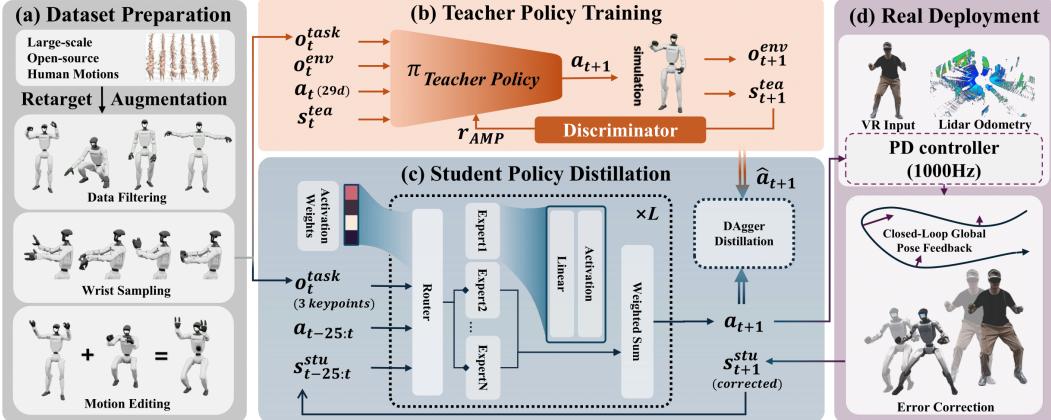


Figure 2: Framework and structure of **CLONE**. **CLONE** curates and augments the retargeted AMASS [13] dataset through motion editing to introduce diverse humanoid motions and detailed hand movements. We employ an MoE network as the student policy, distilling it from a teacher policy trained with privileged information. For the real-world deployment, we integrate LiDAR odometry into the system to obtain real-time humanoid states, enabling closed-loop error correction.

Sec. 3.1 details the policy learning approach that transforms these sparse control signals into coordinated whole-body movements, while Sec. 3.2 explains the closed-loop error correction mechanism that ensures positional accuracy during extended operation. Subsequently, Sec. 3.3 outlines the design of the reward system and the randomization techniques employed in our approach. Finally, Sec. 3.4 describes the dataset curation process. Implementation details can be found in Appendix C.

3.1 Policy Learning

We employ a teacher-student training strategy for the teleoperation policy, following the overall framework of OmniH2O [7] (see Appendix A.1 for the problem formulation). A teacher policy is first trained using privileged information, and a student policy is then distilled from it to operate using only observations available in the real world.

Teacher Policy Training The teacher policy, π_{tea} , is implemented as an MLP. At each timestep t , the teacher policy takes the observation $\mathbf{o}_t^{\text{tea}} = [\mathbf{s}_t^{\text{tea}}, \mathbf{o}_t^{\text{task}}, \mathbf{a}_t, \mathbf{o}_t^{\text{env}}]$ as input and outputs the action $\mathbf{a}_{t+1} \in \mathbb{R}^{29}$. Here, \mathbf{a}_{t+1} represents the target joint positions, which are used to compute motor torques through a PD controller.

The privileged states, $\mathbf{s}_t^{\text{tea}}$, consist of $[\mathbf{p}_t, \theta_t, \mathbf{v}_t, \omega_t]$, where \mathbf{p}_t is the angular positions of each joint, θ_t , \mathbf{v}_t , and ω_t are the 6D poses, linear velocities, and angular velocities of all robot links. Task-related observations, $\mathbf{o}_t^{\text{task}} = [\hat{\mathbf{p}}_{t+1} - \mathbf{p}_t, \hat{\theta}_{t+1} - \theta_t, \hat{\mathbf{v}}_{t+1} - \mathbf{v}_t, \hat{\omega}_{t+1} - \omega_t, \hat{\mathbf{p}}_{t+1}, \hat{\theta}_{t+1}]$, capture the reference motion (denoted by $\hat{\cdot}$) and the differences between the reference motion and the current robot state. Environmental observations, $\mathbf{o}_t^{\text{env}}$, include environment and humanoid-specific intrinsic properties, such as the friction coefficient of the ground and the mass distribution of the robot.

Student Policy Distillation The student policy, deployable on the real robot, replaces privileged states with a sequence of observations directly obtainable on a real robot: $a_{t+1} = \pi_{\text{stu}}(s_{t-25:t}^{\text{stu}}, a_{t-25:t}, o_t^{\text{task}})$, where the robot state sequence $s_{t-25:t}^{\text{stu}}$ includes the joint positions q , joint velocities \dot{q} , root angular velocity ω^{root} , and root gravity vector g obtained from on-device IMU over the past 25 frames. The task-related observations, $\mathbf{o}_t^{\text{task}}$ consist of $\hat{\mathbf{p}}_{t+1} - \mathbf{p}_t, \hat{\mathbf{p}}_{t+1}, \hat{\mathbf{p}}_{t+1}, h_t$ and \hat{h}_{t+1} , where p_t is the 3d positions of the head and two wrists obtained from LiDAR odometry and forward kinematics, \hat{p}_{t+1} and \hat{h}_{t+1} are the target 3d positions and velocities of those links from the reference motion. h_t and \hat{h}_{t+1} represent the current and target wrists orientations.

The student policy is implemented using an MoE architecture, as shown in Fig. 2. It consists of L MoE layers, each comprising N experts. Each expert functions as an independent feed-forward sub-layer with its own parameters. A router dynamically selects which experts are activated to process the hidden state at each layer. The outputs are propagated through the network, enabling the model to specialize different experts for distinct motion patterns and mitigate optimization interference.

For each MoE layer, the router generates a weight distribution over all experts based on the input. The layer output is a weighted sum of the outputs from the top- k experts with the highest routing weights. Specifically, the output of the MoE layer is: $f = \sum_i^k w_i \cdot E_i(\cdot)$, where w_i is the routing weight for the i -th selected expert, and $E_i(\cdot)$ is the output of the i -th expert given the input h . To prevent the model from collapsing into using only a few experts, we introduce a balancing loss to regularize the router and encourage a uniform selection of experts during training and inference. The balancing loss is defined as:

$$\mathcal{L}_{balance} = \sum_{l=1}^L \sum_{e=1}^N [\max(p_e - \frac{1+\epsilon}{N}, 0) + \min(\frac{1-\epsilon}{N} - p_e, 0)], \quad (1)$$

where $p_e = \mathbb{E}[w_e]$ represents the expected activation probability of expert e , and ϵ is a slack constant that allows slight deviations from perfect uniformity.

3.2 Closed-Loop Error Correction

Traditional humanoid teleoperation systems operate in an open-loop configuration, where small errors in position tracking accumulate over time, leading to significant drift during extended operations. To address this fundamental limitation, we implement a closed-loop error correction mechanism that continuously monitors and compensates for positional discrepancies between the teleoperator and the humanoid robot. Our system utilizes LiDAR odometry to maintain accurate global position estimates for both the humanoid robot and the teleoperator. Specifically, we employ FAST-LIO2 [12], a state-of-the-art algorithm that tightly couples IMU and LiDAR data through an iterated Kalman filter to provide robust real-time state estimation even during dynamic movements. The humanoid’s global position $p \in \mathbb{R}^3$ is computed from onboard sensors, while the teleoperator’s position $\hat{p} \in \mathbb{R}^3$ is similarly tracked through VR hardware equipped with a comparable odometry pipeline. The student teleoperation policy consumes the difference between p and \hat{p} , generating actions that systematically reduce positional drift.

3.3 Reward Design and Domain Randomization

We build upon the reward terms and domain randomizations from OmniH2O [7] as the foundation of our approach, with specific enhancements to address the challenges of real-world teleoperation. Detailed reward functions and domain randomization settings are provided in [Appendix B](#).

To enhance robustness against LiDAR odometry errors, we introduce a velocity-dependent Stochastic Differential Equation (SDE) noise model during training. For the head position \vec{p}_{head} , we define the randomized position $\tilde{\vec{p}}_{\text{head}}$ as:

$$d\tilde{\vec{p}}_{\text{head}} = \dot{\vec{p}}_{\text{head}} dt + \left(\frac{\|\dot{\vec{p}}_{\text{head}}\|}{c_{\text{vel}}} + c_{\text{min}} \right) d\vec{W}, \quad (2)$$

where \vec{W} is a standard Wiener process, and c_{vel} and c_{min} are constants that scale the noise proportionally to movement speed and establish a minimum randomization level. This formulation mirrors real-world dynamics, where faster movements tend to produce greater odometry errors. We use forward kinematics to compute other body positions based on the randomized head position, while periodically resetting and constraining the maximum deviation to avoid unrealistic drift.

To tackle the absence of explicit lower-body reference motions in our teleoperation system, we employ an Adversarial Motion Priors (AMP) reward [39] to regularize lower-body movements and encourage natural, stable behavior. Through this combination of specialized domain randomization and reward design, our system learns to generate robust lower-body behaviors while maintaining precise upper-body control aligned with operator commands.

3.4 Dataset Curation

The training dataset **CLONED** comprises three complementary components to support robust whole-body teleoperation: (i) an augmented AMASS [13] subset of 149 curated sequences featuring di-

verse pairings of upper- and lower-body movements, enhanced via targeted motion editing to increase compositional diversity and policy generalization; (ii) 14 custom sequences captured with an IMU-based Xsens MoCap system to fill coverage gaps, emphasizing continuous transitions and diverse upper-body poses critical for manipulation; and (iii) systematic hand orientation augmentation through procedurally generated 6D wrist targets, smoothed via Spherical Linear Interpolation (SLERP) to ensure coherent and natural hand motions for teleoperation.

4 Experiments

In this section, we present a comprehensive evaluation of **CLONE** through both extensive simulated experiments and real-world deployment. Our evaluation consists of four components: (i) quantitative benchmarking of motion tracking accuracy in the Isaac Gym simulation environment [40], (ii) assessment of robustness across diverse stance configurations, (iii) ablation studies examining key architectural decisions, and (iv) real-world validation on a Unitree G1 humanoid robot demonstrating unprecedented whole-body motion fidelity and precise position tracking. Together, these experiments validate both the technical performance of our approach and its practical applicability to real-world humanoid teleoperation.

4.1 Simulated Experiments

We conducted thorough quantitative experiments to evaluate **CLONE**'s ability to accurately track reference motions. Our simulated evaluations span four distinct settings: tracking reference motion, tracking diverse stance, ablation studies, and expert activation analysis.

We evaluated **CLONE** on motion tracking tasks from **CLONED** using five metrics: success rate **SR** (%), mean per-keybody position error (MPKPE) E_{mpkpe} (mm), root-relative mean per-keybody position error (R-MPKPE) $E_{r-\text{mpkpe}}$ (mm), average joint velocity error E_{vel} (mm/s), and hand orientation tracking error E_{hand} . Success rate (**SR**) represents the proportion of episodes where: (i) the robot maintains balance without falling, and (ii) the average per-keybody distance between the robot and reference motion remains below 1.5m across the three controlled joints. We defined the hand orientation tracking error as $E_{\text{hand}} = 1 - \langle \hat{\mathbf{q}}, \mathbf{q} \rangle^2$, where $\hat{\mathbf{q}}$ and \mathbf{q} represent the reference and robot hand quaternions.

We compared **CLONE** with two ablated baselines, **CLONE**[†], and **CLONE**^{*}. **CLONE**[†] employs an MLP as the student policy, resembling the OmniH2O baseline trained on our data and task. **CLONE**^{*} represents our **CLONE** model trained on OmniH2O data. Quantitative results in Tab. 1 reveal that both the MoE architecture and **CLONED** contribute significantly to accurate reference motion tracking.

Tracking Diverse Stances To assess **CLONE**'s robustness across varying postures, we evaluated its performance in tracking motions with head heights from 1.2m (standing) to 0.6m (deep squatting) in 0.1m decrements. We generated these reference motions by systematically editing sequences from the **CLONED** dataset, creat-

Table 1: Motion tracking evaluation on **CLONED** dataset. **CLONE**[†] uses an MLP as the student policy, **CLONE**^{*} is trained on the OmniH2O data.

Method	SR ↑	$E_{\text{mpkpe}} \downarrow$	$E_{r-\text{mpkpe}} \downarrow$	$E_{\text{vel}} \downarrow$	$E_{\text{hand-rot}} \downarrow$
CLONE [†]	100%	113.97	35.55	245.11	4.73
CLONE [*]	100%	102.20	41.07	309.65	4.61
CLONE	100%	87.84	33.30	227.17	3.61

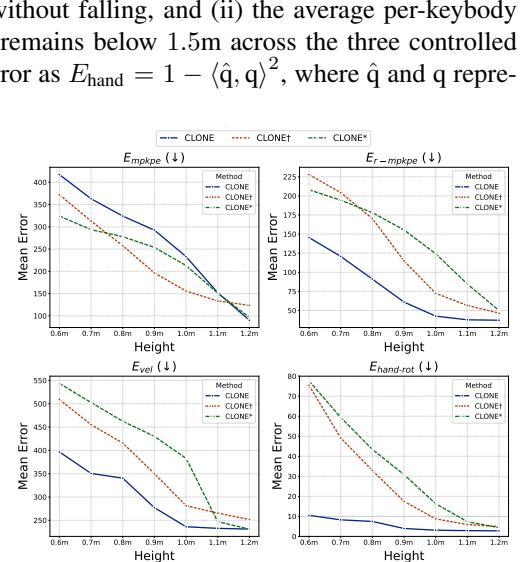


Figure 3: Motion tracking performance comparison between **CLONE** (blue solid), **CLONE**^{*} (green dashed), and **CLONE**[†] (red dashed). Lower values indicate better performance for all error metrics.

ing unseen poses that challenge teleoperation systems.

As shown in Fig. 3, **CLONE** demonstrates an interesting trade-off: while underperforming baselines in absolute position accuracy (MPKPE), it consistently outperforms them in local metrics (R-MPKPE, velocity error, and hand orientation). This pattern suggests that **CLONE** prioritizes faithful reproduction of reference stances—particularly for challenging postures—sometimes at the expense of global positioning. All methods show increased tracking errors at lower heights, confirming the challenge of teleoperating robots in squatting postures.

Ablation Studies We investigated the impact of key design choices, specifically history length and MoE parameters, through systematic ablation experiments reported in Tab. 2. Our results indicate that a configuration using 25 timesteps of history, three MoE layers, and four experts per layer yields optimal performance across most evaluation metrics. We observed that shorter history lengths and increased expert counts can produce marginally lower R-MPKPE values and larger global tracking errors, suggesting a trade-off between local and global motion fidelity.

Expert Activation Analysis To better understand the specialization within our mixture-of-experts architecture, we visualized expert activation weights across nine distinct motion types in Fig. 4. Results reveal clear specialization patterns where motions requiring similar skills activate specific experts. In the first layer, experts 1 and 2 are predominantly activated during standing motions, while experts 3 and 4 show stronger activation during squatting motions. Notably, all four experts in the first layer become activated during dynamic motions such as jumping and punching, suggesting collaborative processing of complex movements. Similar specialization patterns emerged in subsequent layers, albeit with reduced variance across different motion categories.

4.2 Real-World Validation

We evaluated **CLONE** on a physical Unitree G1 humanoid robot through both qualitative demonstrations and quantitative measurements. Our experiments focused on two key capabilities: (i) the accuracy of whole-body motion tracking for diverse skills, and (ii) the precision of our closed-loop error correction mechanism for global position tracking during extended teleoperation.

Whole-Body Motion Tracking As shown in Fig. 5, **CLONE** successfully enables real-time teleoperation across a diverse range of whole-body skills. The robot accurately tracks complex motions including arm waving, deep squatting, standing up from squatted positions, and even dynamic jumping. To our knowledge, this represents notable advancements in whole-body motion fidelity for real-time humanoid teleoperation, particularly for dynamic skills like jumping that require precise balance control and force application.

Global Position Tracking To quantitatively assess our system’s ability to maintain accurate global positioning over extended distances, we designed a controlled path-following experiment. We established fixed initial positions for both the operator and robot, marked on the floor for reference. The operator then walked along a straight path toward a pre-determined target position 8.90m away while teleoperating the robot. We measured the discrepancy between the robot’s final position and the ex-

Table 2: Ablation study on history length and architecture components.

Method	E_{mpkpe}	$E_{r-\text{mpkpe}}$	E_{vel}	$E_{\text{hand-rot}}$
(a) History Length Analysis				
History5	93.97	31.99	236.12	3.80
History50	135.60	41.33	286.66	12.23
History25(CLONE)	87.84	33.30	227.17	3.61
(b) Architecture Ablation				
CLONE ($L = 1$)	134.06	37.56	270.14	7.22
CLONE ($N = 8$)	89.21	30.90	251.10	4.26
CLONE	87.84	33.30	227.17	3.61

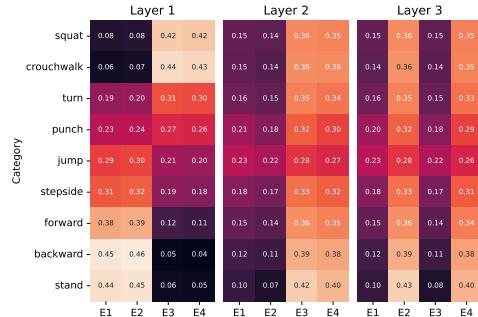


Figure 4: The activation status of each expert.

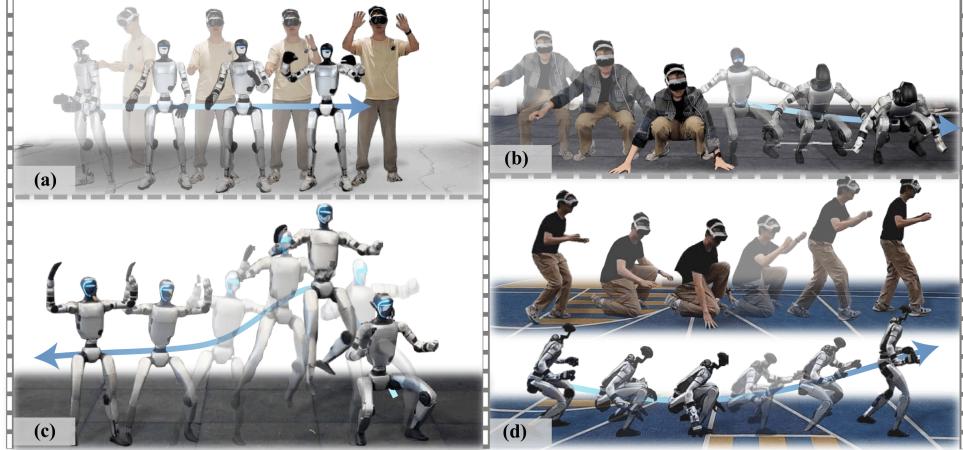


Figure 5: Qualitative results of motion tracking with **CLONE** on Unitree G1. Our method enables accurate tracking of diverse whole-body motions, including waving (a), squatting (b, d), and jumping (c).

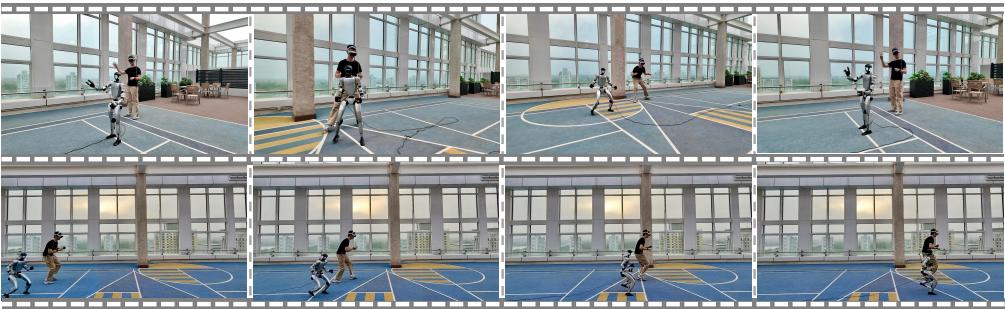


Figure 6: Qualitative results of long-horizon teleoperation. The humanoid accurately tracks both the human’s local pose and global translation over 15m and maintains its distance to its operator. expected position (8.90m from start) as the tracking error. Across ten repetitions of this experiment, our system achieved a mean tracking error of only 5.1cm, with a maximum error of 12.0cm. This high level of positional accuracy confirms that **CLONE**’s closed-loop error correction mechanism effectively compensates for drift and accumulated errors during extended teleoperation sessions.

Long-Horizon Mixed Navigation To further validate our system’s performance in more complex scenarios, we conducted extended teleoperation sessions incorporating multiple movement types. As visualized in Fig. 6, we recorded a continuous teleoperation sequence where the operator traversed a complex path spanning over 15m, incorporating diverse locomotion patterns including forward walking, turning, side-stepping, and returning to the original position.

Throughout this extended sequence, the robot consistently tracked the operator’s movements with high accuracy and returned to its starting position with minimal drift. This demonstration highlights **CLONE**’s robustness for extended teleoperation sessions involving both locomotion and whole-body motion control—a capability essential for practical real-world applications of humanoid robots.

5 Conclusion

In this work, we propose a closed-loop MoE-based teleoperation system that enables comprehensive control of humanoid robots while addressing accumulated tracking errors in long-horizon teleoperation. Our policy is trained on an augmented dataset refined through motion editing and MoCap, integrates LiDAR odometry to enhance position perception, enabling stepwise error correction. Additionally, our system empowers humanoids to perform a wide range of tasks, including complex manipulation, long-distance movement, and long-horizon tasks, thus facilitating the collection of diverse task demonstrations.

6 Limitation

While our approach demonstrates significant capabilities in humanoid teleoperation, several important limitations remain to be addressed in future work:

Control Precision with Minimal Input Our closed-loop error correction mechanism effectively reduces jitter and drift from accumulated errors during humanoid locomotion, enabling whole-body control with just three tracked points (head and hands) from a VR headset. However, this minimal input configuration inherently constrains the system’s stability in certain scenarios. Although our adversarial reward helps stabilize the lower body, fine-grained stability control remains challenging. Future work should explore additional sensing modalities or predictive algorithms to enhance stability while maintaining the simplicity of the interface.

Dynamic Motion Capabilities Despite our dataset augmentation strategies to increase motion diversity, the system exhibits reduced performance when executing highly dynamic movements such as jumping or single-leg hopping. These limitations stem from both the training data distribution and the inherent difficulty of balancing during such maneuvers. Addressing these constraints would require: (i) expanding the reference motion dataset to include more diverse dynamic sequences, (ii) developing specialized reward functions targeted at maintaining balance during rapid transitions, and (iii) potentially incorporating physics-based constraints into the policy to better model momentum dynamics.

References

- [1] Z. Fu, Q. Zhao, Q. Wu, G. Wetzstein, and C. Finn. Humanplus: Humanoid shadowing and imitation from humans. In *Conference on Robot Learning (CoRL)*, 2024.
- [2] Y. Ze, Z. Chen, W. Wang, T. Chen, X. He, Y. Yuan, X. B. Peng, and J. Wu. Generalizable humanoid manipulation with improved 3d diffusion policies. *arXiv preprint arXiv:2410.10803*, 2024.
- [3] T. He, W. Xiao, T. Lin, Z. Luo, Z. Xu, Z. Jiang, C. Liu, G. Shi, X. Wang, L. Fan, and Y. Zhu. Hover: Versatile neural whole-body controller for humanoid robots. *arXiv preprint arXiv:2410.21229*, 2024.
- [4] J. Mao, S. Zhao, S. Song, T. Shi, J. Ye, M. Zhang, H. Geng, J. Malik, V. Guizilini, and Y. Wang. Learning from massive human videos for universal humanoid pose control. *arXiv preprint arXiv:2412.14172*, 2024.
- [5] C. Lu, X. Cheng, J. Li, S. Yang, M. Ji, C. Yuan, G. Yang, S. Yi, and X. Wang. Mobile-television: Predictive motion priors for humanoid whole-body control. *arXiv preprint arXiv:2412.07773*, 2024.
- [6] Q. Ben, F. Jia, J. Zeng, J. Dong, D. Lin, and J. Pang. Homie: Humanoid loco-manipulation with isomorphic exoskeleton cockpit. In *Robotics: Science and Systems (RSS)*, 2025.
- [7] T. He, Z. Luo, X. He, W. Xiao, C. Zhang, W. Zhang, K. Kitani, C. Liu, and G. Shi. Omnih2o: Universal and dexterous human-to-humanoid whole-body teleoperation and learning. In *Conference on Robot Learning (CoRL)*, 2024.
- [8] J. Shi, X. Liu, D. Wang, O. Lu, S. Schwertfeger, F. Sun, C. Bai, and X. Li. Adversarial locomotion and motion imitation for humanoid policy learning. *arXiv preprint arXiv:2504.14305*, 2025.
- [9] T. He, Z. Luo, W. Xiao, C. Zhang, K. Kitani, C. Liu, and G. Shi. Learning human-to-humanoid real-time whole-body teleoperation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024.

- [10] X. Cheng, Y. Ji, J. Chen, R. Yang, G. Yang, and X. Wang. Expressive whole-body control for humanoid robots. In *Robotics: Science and Systems (RSS)*, 2024.
- [11] M. Ji, X. Peng, F. Liu, J. Li, G. Yang, X. Cheng, and X. Wang. Exbody2: Advanced expressive humanoid whole-body control. *arXiv preprint arXiv:2412.13196*, 2024.
- [12] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang. Fast-llo2: Fast direct lidar-inertial odometry. *IEEE Transactions on Robotics (T-RO)*, 38(4):2053–2073, 2022.
- [13] N. Mahmood, N. Ghorbani, N. F. Troje, G. Pons-Moll, and M. J. Black. AMASS: Archive of motion capture as surface shapes. In *Proceedings of International Conference on Computer Vision (ICCV)*, 2019.
- [14] S. Dafarra, K. Darvish, R. Grieco, G. Milani, U. Pattacini, L. Rapetti, G. Romualdi, M. Salvi, A. Scalzo, I. Sorrentino, et al. icub3 avatar system: Enabling remote fully immersive embodiment of humanoid robots. *Science Robotics*, 9(86):eadh3834, 2024.
- [15] K. Darvish, Y. Tirupachuri, G. Romualdi, L. Rapetti, D. Ferigo, F. J. A. Chavez, and D. Pucci. Whole-body geometric retargeting for humanoid robots. In *International Conference on Humanoid Robots (Humanoids)*, 2019.
- [16] A. Brygo, I. Sarakoglou, N. Garcia-Hernandez, and N. Tsagarakis. Humanoid robot teleoperation with vibrotactile based balancing feedback. In *Haptics: Neuroscience, Devices, Modeling, and Applications: 9th International Conference, EuroHaptics*, 2014.
- [17] L. Peternel and J. Babič. Learning of compliant human–robot interaction using full-body haptic interface. *Advanced Robotics*, 27(13):1003–1012, 2013.
- [18] J. Ramos and S. Kim. Dynamic locomotion synchronization of bipedal robot and human operator via bilateral feedback teleoperation. *Science Robotics*, 4(35):eaav4282, 2019.
- [19] X. Cheng, J. Li, S. Yang, G. Yang, and X. Wang. Open-television: Teleoperation with immersive active visual feedback. *arXiv preprint arXiv:2407.01512*, 2024.
- [20] J. Chagas Vaz, D. Wallace, and P. Y. Oh. Humanoid loco-manipulation of pushed carts utilizing virtual reality teleoperation. In *ASME International Mechanical Engineering Congress and Exposition*, 2021.
- [21] L. Penco, N. Scianca, V. Modugno, L. Lanari, G. Oriolo, and S. Ivaldi. A multimode teleoperation framework for humanoid loco-manipulation: An application for the icub robot. *IEEE Robotics & Automation Magazine*, 26(4):73–82, 2019.
- [22] S. Tachi, Y. Inoue, and F. Kato. Telesar vi: Telexistence surrogate anthropomorphic robot vi. *International Journal of Humanoid Robotics*, 17(05):2050019, 2020.
- [23] M. Moniruzzaman, A. Rassau, D. Chai, and S. M. S. Islam. Teleoperation methods and enhancement techniques for mobile robots: A comprehensive survey. *Robotics and Autonomous Systems*, 150:103973, 2022.
- [24] R. Huang, S. Zhu, Y. Du, and H. Zhao. Moe-loco: Mixture of experts for multitask locomotion. *arXiv preprint arXiv:2503.08564*, 2025.
- [25] S. Zhou, W. Zhang, J. Jiang, W. Zhong, J. GU, and W. Zhu. On the convergence of stochastic multi-objective gradient manipulation and beyond. In *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*, 2022.
- [26] C. Yang, K. Yuan, Q. Zhu, W. Yu, and Z. Li. Multi-expert learning of adaptive legged locomotion. *Science Robotics*, 5(49):eabb2174, 2020.

- [27] Z. Xie, S. Starke, H. Y. Ling, and M. van de Panne. Learning soccer juggling skills with layer-wise mixture-of-experts. In *ACM SIGGRAPH / Eurographics Symposium on Computer Animation (SCA)*, 2022.
- [28] W. Song, H. Zhao, P. Ding, C. Cui, S. Lyu, Y. Fan, and D. Wang. Germ: A generalist robotic model with mixture-of-experts for quadruped robot. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024.
- [29] G. Cheng, L. Dong, W. Cai, and C. Sun. Multi-task reinforcement learning with attention-based mixture of experts. *IEEE Robotics and Automation Letters (RA-L)*, 8(6):3812–3819, 2023.
- [30] C. Wang, L. Fan, J. Sun, R. Zhang, L. Fei-Fei, D. Xu, Y. Zhu, and A. Anandkumar. Mimicplay: Long-horizon imitation learning by watching human play. In *Conference on Robot Learning (CoRL)*, 2023.
- [31] Z. Lin, Y. Chen, and Z. Liu. Hierarchical human-to-robot imitation learning for long-horizon tasks via cross-domain skill alignment. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2024.
- [32] H. Shi, H. Xu, S. Clarke, Y. Li, and J. Wu. Robocook: Long-horizon elasto-plastic object manipulation with diverse tools. *Conference on Robot Learning (CoRL)*, 2023.
- [33] Y. Jiang, R. Zhang, J. Wong, C. Wang, Y. Ze, H. Yin, C. Gokmen, S. Song, J. Wu, and L. Fei-Fei. Behavior robot suite: Streamlining real-world whole-body manipulation for everyday household activities. *arXiv preprint arXiv: 2503.05652*, 2025.
- [34] D. Wisth, M. Camurri, and M. Fallon. Vilens: Visual, inertial, lidar, and leg odometry for all-terrain legged robots. *IEEE Transactions on Robotics*, 39(1):309–326, 2022.
- [35] G. Ou, D. Li, and H. Li. Leg-kilo: Robust kinematic-inertial-lidar odometry for dynamic legged robots. *IEEE Robotics and Automation Letters*, 2024.
- [36] F. G. Harvey, M. Yurick, D. Nowrouzezahrai, and C. Pal. Robust motion in-betweening. *ACM Transactions on Graphics (TOG)*, 39(4):60–1, 2020.
- [37] L. Ma, Z. Meng, T. Liu, Y. Li, R. Song, W. Zhang, and S. Huang. Styleloco: Generative adversarial distillation for natural humanoid robot locomotion. *arXiv preprint arXiv:2503.15082*, 2025.
- [38] Z. Wang, Y. Chen, T. Liu, Y. Zhu, W. Liang, and S. Huang. Humanise: Language-conditioned human motion generation in 3d scenes. In *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*, 2022.
- [39] 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)*, 40(4):1–20, 2021.
- [40] V. Makoviychuk, L. Wawrzyniak, Y. Guo, M. Lu, K. Storey, M. Macklin, D. Hoeller, N. Rudin, A. Allshire, A. Handa, et al. Isaac gym: High performance gpu-based physics simulation for robot learning. *arXiv*, 2021.

A Preliminaries

A.1 Formulation

We formulate the humanoid teleoperation as an Markov Decision Process (MDP) $\mathcal{M} = \{\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}\}$. \mathcal{S} includes proprioceptive states s and task-oriented observations o^{task} . The action space $\mathcal{A} \in \mathcal{R}^{29}$ represents the humanoid’s joint positions in our method. \mathcal{T} is the transition function conditioned on \mathcal{S} and \mathcal{A} . The reward functions \mathcal{R} is defined based on \mathcal{S}, \mathcal{A} . In this paper, we utilize the policy π to maximize the overall reward \mathcal{R} using the Proximal Policy Optimization (PPO) algorithm.

We formulate humanoid teleoperation as a MDP $\mathcal{M} = \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}$. The state space \mathcal{S} comprises proprioceptive states s and task-oriented observations o^{task} . The action space $\mathcal{A} \in \mathcal{R}^{29}$ represents the humanoid’s joint positions in our method. The transition function \mathcal{T} is conditioned on \mathcal{S} and \mathcal{A} , while the reward function \mathcal{R} is defined based on \mathcal{S} and \mathcal{A} . In this work, we leverage the policy π to maximize the cumulative reward \mathcal{R} using the Proximal Policy Optimization (PPO) algorithm.

A.2 LiDAR Odometry

LiDAR odometry is designed to accurately determine the robot’s current location. In this paper, we adopt FAST-LIO2 [12], which utilizes onboard LiDAR and IMU to estimate the humanoid’s state. FAST-LIO2 [12] leverages IMU data and LiDAR point clouds to construct and update a map in real time. It then registers the LiDAR point clouds with the map to estimate the robot’s current state. LiDAR odometry offers an effective approach for robot state estimation, which can be utilized to provide precise humanoid state estimation.

Previous teleoperation systems [7, 10, 11] often operate in an open-loop manner, primarily due to the absence of the humanoid’s global position. Consequently, stepwise tracking errors accumulate over time, resulting in significant drift during long-horizon tasks. In this work, we integrate the LiDAR odometry system into our humanoid whole-body control framework to estimate the actual global translation tracking results at each step in a closed-loop manner, as shown in Fig. A1.

Specifically, the teleoperator sends control signals acquired using the Apple Vision Pro to the server via WiFi, while the LiDAR odometry runs on the humanoid’s onboard computer and transmits the humanoid’s state to the server through a cable connected to a router. The policy is deployed on the server at 50 Hz, while the LiDAR odometry operates at 10 Hz. The humanoid’s position is transmitted to the policy via ROS2. By compensating for stepwise tracking errors and reducing globally accumulated tracking errors based on the discrepancy between the target and actual states, our system achieves closed-loop error correction. This step-by-step error compensation mechanism significantly enhances tracking accuracy in long-horizon tasks.

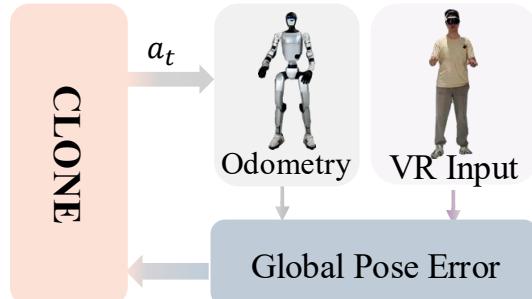


Figure A1: **System Illustration of CLONE.** CLONE utilizes the global poses of both the human and the humanoid to perform closed-loop error correction. The humanoid’s state is derived from LiDAR odometry, which uses IMU and LiDAR data as inputs to estimate the precise global pose.

B Reward Functions and Domain Randomization

Tab. A1 provides a detailed overview of the reward structure utilized in this study, while Tab. A2 outlines the domain randomization scheme employed.

Table A1: **Reward functions.** The details of the primary reward function used in our training process.

Term	Expression	Weight
Torque	$\ \tau\ _2^2$	-0.0001
Torque limits	$[\tau \notin [\tau_{\min}, \tau_{\max}]]_1$	-2
DoF position limits	$[\mathbf{p}_t \notin [\mathbf{p}_{\min}, \mathbf{p}_{\max}]]_1$	-625
DoF velocity limits	$[\dot{\mathbf{p}}_t \notin [\dot{\mathbf{p}}_{\min}, \dot{\mathbf{p}}_{\max}]]_1$	-50
Termination	termination ₁	$-e^4$
DoF acceleration	$\ \ddot{\mathbf{q}}_t\ _2^2$	$-1.1e^{-5}$
DoF velocity	$\ \dot{\mathbf{q}}_t\ _2^2$	-0.004
Lower-body action rate	$\ \mathbf{a}_t^{lower} - \mathbf{a}_{t-1}^{lower}\ _2^2$	-1.0
Upper-body action rate	$\ \mathbf{a}_t^{upper} - \mathbf{a}_{t-1}^{upper}\ _2^2$	-0.3
Feet air time	$T_{air} - 0.3$	2500
Stumble	$[(\mathbf{F}_{feet}^{xy} > 5 \times \mathbf{F}_{feet}^z)]_1$	$-1.25e^4$
Slippage	$\ \mathbf{v}_t^{feet}\ _2^2 \cdot [(\mathbf{F}_{feet} \geq 1)]_1$	-80
Feet orientation	$\ \mathbf{R}_z^{feet}\ $	-62.5
In the air	$[(\mathbf{F}_{feet}^{left}, \mathbf{F}_{feet}^{right} < 1)]_1$	-750
Orientation	$\ \mathbf{R}_z^{root}\ $	-50
DoF position	$\exp(-0.25\ \hat{\mathbf{p}} - \mathbf{p}\ _2)$	100
DoF velocity	$\exp(-0.25\ \hat{\mathbf{p}} - \dot{\mathbf{p}}\ _2^2)$	10
Extend body position	$\exp(-0.5\ \hat{\mathbf{q}} - \mathbf{q}\ _2^2)$	250
Body position (VR)	$\exp(-0.5\ \mathbf{q}_{vr} - \hat{\mathbf{q}}_{vr}\ _2^2)$	150
Body rotation	$\exp(-0.1\ \theta \ominus \hat{\theta}\)$	400
Body velocity	$\exp(-10.0\ \mathbf{v} - \hat{\mathbf{v}}\ _2)$	80
Body angular velocity	$\exp(-0.01\ \omega - \hat{\omega}\ _2)$	8
Body hand rotation	$(\theta_{hand} - \hat{\theta}_{hand})^2$	500
AMP	Sec. 3.3	500

Table A2: **Domain Randomization.** The details of the primary domain randomization used in our training process.

Term	Value
Friction	$\mathcal{U}(0.6, 2.0)$
Base CoM offset	$\mathcal{U}(-0.04, 0.04)m$
Link mass	$\mathcal{U}(0.7, 1.25) \times \text{default kg}$
P Gain	$\mathcal{U}(0.85, 1.15) \times \text{default}$
D Gain	$\mathcal{U}(0.85, 1.15) \times \text{default}$
Torque RFI	$0.05 \times \text{torque limit N} \cdot \text{m}$
Control delay	$\mathcal{U}(0.0, 20)\text{ms}$
Global step delay	$\mathcal{U}(0.0, 80)\text{ms}$
Rand born distance	$\mathcal{U}(0.0, 2.0)m$
Rand heading degree	$\mathcal{U}(-20.0, 20.0)\text{degree}$
Push robot	interval = 5s, $v_{xy} = 1.5\text{m/s}$
Terrain type	flat, rough, low obstacles [7]

C Implement Details

We train our policy in IsaacGym using a single A800 GPU. The teacher policy is trained for $1M$ iterations with 8192 parallel environments, while the student policy is trained for $600K$ iterations with 4096 parallel environments. In our implementation, the student policy consists of $L = 3$ MoE layers, each containing $N = 4$ experts. It uses a history length of $H = 25$ frames and activates the top $k = 2$ experts based on the highest weights determined by the router.

D Experiments

D.1 Ablation on The Number of MoE Layers and Experts

Additionally, we visualize the activation patterns of experts in Fig. 4, Fig. A2, and Fig. A4.

In Fig. A2, we observe that MoE layers with $N = 8$ experts activate only half of the experts in each layer, revealing that 8 experts are redundant for the current training data distribution, while 4

Category	Layer 1								Layer 2								Layer 3							
	E1	E2	E3	E4	E5	E6	E7	E8	E1	E2	E3	E4	E5	E6	E7	E8	E1	E2	E3	E4	E5	E6	E7	E8
squat	0.23	0.08	0.00	0.08	0.00	0.40	0.20	0.00	0.24	0.07	0.03	0.16	0.05	0.17	0.24	0.04	0.34	0.00	0.21	0.03	0.24	0.10	0.07	0.00
crouchwalk	0.29	0.06	0.00	0.04	0.00	0.40	0.19	0.02	0.29	0.02	0.02	0.14	0.05	0.12	0.28	0.07	0.32	0.00	0.16	0.02	0.30	0.09	0.11	0.01
turn	0.30	0.10	0.07	0.00	0.00	0.31	0.19	0.03	0.20	0.20	0.06	0.00	0.03	0.18	0.06	0.26	0.27	0.01	0.07	0.07	0.27	0.17	0.11	0.03
punch	0.13	0.24	0.09	0.00	0.09	0.23	0.14	0.07	0.25	0.06	0.10	0.03	0.15	0.19	0.13	0.10	0.28	0.01	0.07	0.05	0.32	0.13	0.13	0.01
jump	0.28	0.10	0.07	0.00	0.01	0.27	0.25	0.01	0.23	0.21	0.05	0.08	0.03	0.09	0.03	0.30	0.37	0.02	0.12	0.05	0.26	0.08	0.08	0.01
stepside	0.33	0.13	0.12	0.00	0.01	0.27	0.14	0.01	0.21	0.13	0.02	0.01	0.09	0.26	0.01	0.28	0.25	0.03	0.06	0.05	0.32	0.18	0.10	0.02
forward	0.33	0.13	0.11	0.00	0.00	0.29	0.13	0.01	0.19	0.17	0.02	0.00	0.06	0.19	0.07	0.30	0.29	0.02	0.04	0.05	0.30	0.17	0.10	0.03
backward	0.37	0.12	0.12	0.00	0.00	0.28	0.12	0.00	0.22	0.05	0.02	0.00	0.11	0.24	0.00	0.35	0.28	0.01	0.03	0.04	0.39	0.12	0.12	0.01
stand	0.36	0.08	0.08	0.00	0.00	0.37	0.12	0.00	0.46	0.00	0.09	0.00	0.08	0.28	0.02	0.07	0.26	0.01	0.00	0.03	0.45	0.08	0.17	0.01

Figure A2: Experts activation when $N = 8$

experts are sufficient. Furthermore, Fig. A4 shows that $\text{CLONE}^*(L = 1)$, which uses $L = 1$ layer, still learns to activate different experts. However, its tracking performance is inferior compared to CLONE .

Based on these results, we select CLONE with 3 MoE layers, each containing 4 experts, as the optimal configuration.

D.2 Quantitative Results of Long-Horizon Consistency

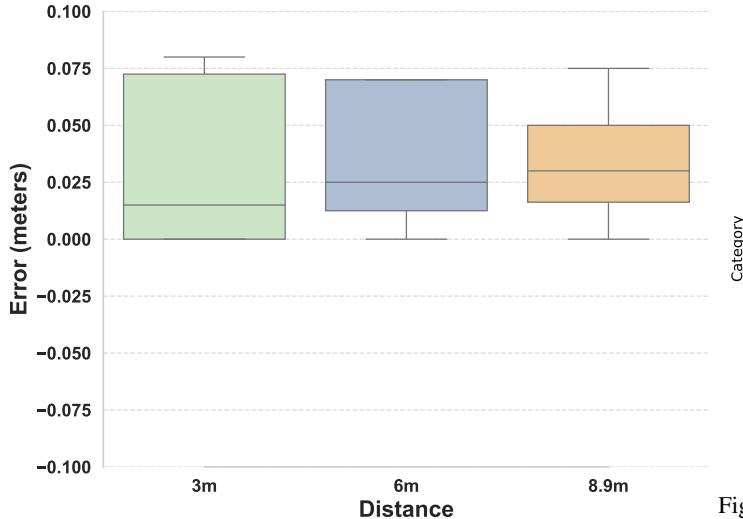


Figure A3: Long-Horizon Tracking Result of CLONE in Real-World.

Category	Layer 1			
	E1	E2	E3	E4
squat	0.10	0.40	0.43	0.07
crouchwalk	0.14	0.36	0.44	0.06
turn	0.47	0.05	0.32	0.15
punch	0.18	0.35	0.22	0.25
jump	0.47	0.04	0.13	0.36
stepside	0.48	0.02	0.16	0.34
forward	0.50	0.01	0.14	0.36
backward	0.50	0.00	0.11	0.39
stand	0.15	0.40	0.21	0.24

Figure A4: Experts activation when $L = 1$

In this section, we use the same experimental settings as described in Sec. 4.2. It is important to note that the test environment consists of glass walls to simulate the complex teleoperation scenarios encountered in daily life. The extended real-world experimental results shown in Fig. A3 demonstrate that our model maintains robust mean and maximum tracking errors within 8.90 m. Specifically, the mean tracking error is approximately 5.0 cm, while the maximum tracking error is around 12.0 cm. Although the maximum tracking error remains below 2% of the total tracking distance, both the mean and maximum tracking errors increase as the tracking distance grows. This increase is primarily attributed to the accumulation of odometry estimation errors over long distances and in challenging environments.