

The Duke Humanoid: Design and Control For Energy-Efficient Bipedal Locomotion Using Passive Dynamics

Boxi Xia, Bokuan Li, Jacob Lee, Michael Scutari, Boyuan Chen

Abstract—We present the Duke Humanoid, an open-source 10-degrees-of-freedom humanoid, as an extensible platform for locomotion research. The design mimics human physiology, with symmetrical body alignment in the frontal plane to maintain static balance with straight knees. We develop a reinforcement learning policy that can be deployed zero-shot on the hardware for velocity-tracking walking tasks. Additionally, to enhance energy efficiency in locomotion, we propose an end-to-end reinforcement learning algorithm that encourages the robot to leverage passive dynamics. Our experimental results show that our passive policy reduces the cost of transport by up to 50% in simulation and 31% in real-world tests. Our website is <http://generalroboticslab.com/DukeHumanoidv1/>.

I. INTRODUCTION

Recent advances in robotics have enabled humanoids to perform impressive dynamic motions, such as walking, high-speed running, and backflipping [1–4]. However, many commercially available humanoids are closed systems with limited customizability and access to low-level sensor and actuator interfaces. Some recent efforts open-source the software [5, 6], while others offer open-source hardware but are limited in their ability to perform dynamic motions due to the use of high-reduction-ratio servo motors [7–9].

The demand for customizable and high-performance humanoid platforms in the research community is evident. Yet, there remains a scarcity of open-source humanoid hardware capable of performing dynamic locomotion. Such platforms are essential for enabling transparent access, customization, and investigation of low-level hardware interactions, all of which are critical to advancing humanoid research.

This paper introduces an open-source humanoid platform (Fig. 1), designed to address the lack of open-source hardware capable of dynamic motion. The platform is customizable, allowing researchers to integrate design changes and co-optimize hardware and control systems. Our hardware design mimics human physiology by referencing the ratios of the femur, tibia, foot length, and stance width. The humanoid has a nearly symmetric weight distribution, naturally supporting static straight-knee standing. Additionally, we demonstrate the platform’s capabilities by deploying a reinforcement learning (RL) walking policy, trained in simulation, to the physical system without fine-tuning. Both hardware and software are open-sourced.

With our platform, we further study a fundamental question in humanoid locomotion: How can robots achieve energy-efficient human-like walking? Many humanoids are

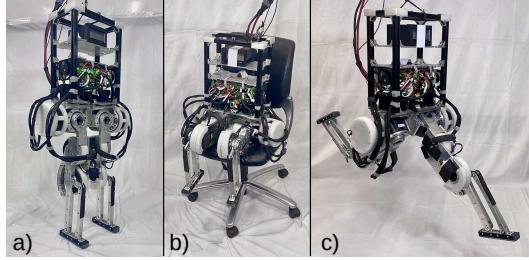


Fig. 1: **Duke Humanoid v1.0:** a) The frontal plane symmetry of the hip enables static standing with straight knees. b) and c) Additional poses demonstrating the robot’s range of motion.

designed with human-like morphology, which simplifies control but requires continuous active joint actuation, leading to inefficient energy use. This inefficiency results in frequent charging and battery replacement, leading to additional operational costs. Addressing this bottleneck requires both hardware innovations and improved control strategies.

Unlike most humanoids that rely on consistent active motor control in locomotion, humans leverage passive dynamics in their gait. Human walking requires minimal muscular effort for leg swinging by utilizing pendulum-like motion [10, 11]. Some bipedal robots use passive dynamics to achieve human-level walking efficiency while requiring little to no power input [12–15]. However, their mass distribution is specifically tuned for passive walking, and any customization, such as adding a payload, may compromise their ability to walk effectively and efficiently. Previous work [12] indicates that straight-knee designs promote efficient walking without active control. These designs can maintain static equilibrium in a standing pose by fully locking the stance leg, while allowing a slight unlock of the knee in the swing leg. Our mechanical design satisfies this criterion through our near-symmetrical hip arrangement.

In parallel, research on energy-efficient legged robot control has proposed learning gait patterns [16], using compliant reinforcement learning [17], and optimizing the center-of-mass trajectory [18]. Although end-to-end deep reinforcement learning has achieved state-of-the-art locomotion [19–21], reinforcement learning based controller design has not yet sufficiently explored incorporating passive dynamics for bipedal locomotion within their framework. While passive dynamics has been widely studied in traditional controller design, its potential application within learning-based controller remains relatively untapped.

In this paper, we aim to enhance the energy efficiency of humanoids through an early exploration of utilizing passive

*This work is supported by DARPA FoundSci HR00112490372, DARPA TIAMAT HR00112490419, ARL STRONG W911NF2320182 and W911NF2220113. All authors are from Duke University.

dynamics in the end-to-end reinforcement learning framework. Our key idea is to explicitly modulate the joint torques to switch between passive and active control during policy learning. The simulation result shows that our passive policy can achieve up to a 50% increase in energy efficiency at low-speed walking ($< 0.5\text{m/s}$) compared to a baseline RL policy. We validate the effectiveness of our passive action policy by deploying it on our physical platform, showing a 31% increase in energy efficiency.

Our main contributions are twofold: (1) we present an open-source humanoid platform for dynamic locomotion research, and (2) we demonstrate that the use of passive dynamics in the end-to-end reinforcement learning framework leads to enhanced energy efficiency in bipedal locomotion.

II. RELATED WORK

A. Humanoid Platforms

The robotics industry offers several humanoid platforms of various scales [1–4]. However, commercial humanoids often have limited customizability and accessibility due to closed-source hardware. This lack of openness makes it difficult for researchers to modify or experiment with the robot’s low-level systems, limiting opportunities for fundamental research. We address this limitation by offering an open-source platform. Academic labs have developed humanoids for research purposes [5, 6, 8, 22–27]. While these robots may not be as polished or robust as their commercial counterparts, they offer more flexibility in customizing hardware and control, enabling researchers to push the boundaries of hardware and algorithm designs.

Mid-sized humanoids strike a balance between functionality and cost. Compared to adult-sized robots, they are easier to design, manufacture, and handle. While miniature robots offer cost advantages, their size limits their ability to fully replicate human locomotion. Table I compares key parameters of several mid-sized industrial and academic humanoids. Among them, our platform is open-source in both hardware and software. Our design features 0.5 m legs, tall enough to access standard workbenches and tables. Similar to previous efforts [6, 23], we limit the ankle’s degrees of freedom (DoF) to one (plantar flexion/dorsiflexion) to reduce leg mass and inertia, allowing the motor to handle higher payloads while providing sufficient mobility for dynamic walking. Additionally, our humanoid provides a maximum torque surpassing that of most mid-sized humanoids, making it suitable for dynamic locomotion with higher torque and payload demands.

B. Control Strategies for Humanoid Locomotion

Model-Based Control Model-based control approaches like Zero-Moment Point (ZMP) [28] ensure stability by maintaining the center of pressure within the support polygon. Model predictive control (MPC) [29] optimizes future trajectories and control actions based on a dynamic model of the robot. Whole-body control (WBC) [30] coordinates the robot’s entire body motion while accounting for its dynamics and constraints. Although highly successful, model-based control

Robot	Symmetry*	Leg length [m]	Leg DoF	Mass [kg]	Max HFE [N·m]	Max KFE [N·m]
MIT [23]		0.55	5	24	68	136
Berkeley [5]		0.4	6	16	63	81
Unitree G1 [3]		0.6	6	35	88	120
HECTOR [6]		0.44	5	16	33.5	51.9
iCub [9]		0.4	6	24	40	40
COMAN [22]		0.44	6	55	55	40
Ours	✓	0.5	5	30	264	238

Symmetry*: hip arrangement symmetry across the frontal plane

TABLE I: Comparison of key parameters for mid-sized academics and industrial humanoids.

requires extensive modeling of the specific robot and often relies on numerous assumptions about its dynamics, limiting transferability to other humanoids without expert knowledge.

Learning-Based Control Learning-based control [31–34] has gained popularity due to its ability to operate with less prior knowledge of specific robots. One approach is to use imitation learning, which directly learns from expert gait demonstrations. However, this method requires access to expensive motion capture data and is prone to errors when the policy encounters situations that differ from the expert demonstrations. Another promising strategy is to use RL to learn control policies through trial and error by maximizing reward signals. We build our control framework based on this formulation due to its flexibility and minimal requirement of prior knowledge or expert demonstrations. When coupled with the recent development of massively parallel RL simulations [35, 36], this allows us to achieve effective control policies with rapid convergence.

C. Strategies for Energy-Efficient Locomotion

Hardware Strategies As discussed by Garcia et al. [12], hardware designs for passive dynamic walking often benefit from straight-knee configurations to enable efficient walking without active control. A key design requirement for passive walking is maintaining static equilibrium when standing on flat ground, which can be achieved by fully locking the stance leg while allowing the swing leg’s knee to slightly unlock. Various studies [13–15, 37] have demonstrated the effectiveness of straight-knee designs in achieving fully or mostly passive walking. Additionally, incorporating energy exchange components can enhance energy transmission. For instance, the DURUS robot [38] employs compression springs in the ankles, while the COMAN robot [22] uses compliant joints in the hip, knee, and ankle to store and release energy during locomotion.

Control Strategies Energy-efficient control strategies have also been explored. By removing the velocity tracking reward during post-disturbance training, a compliant reinforcement learning policy [17] leads to more energy-efficient push recovery. Similarly, COMAN [18] uses RL to optimize its center-of-mass trajectory for better energy efficiency.

Our work develops both hardware and software strategies to achieve energy-efficient locomotion. Our design features a near-symmetrical hip arrangement (front and back) with

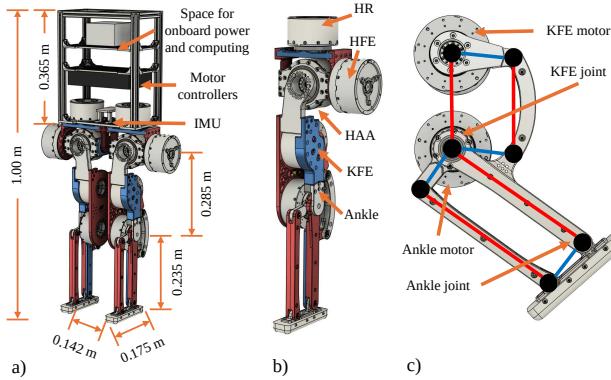


Fig. 2: Mechanical Design Overview: a) Major dimensions and the extensible body design. b) All joints in the left leg. c) Two parallel linkages in the knee and ankle.

straight knees, which satisfies the static standing requirement for passive walking. Furthermore, we explore directly optimizing passivity within an end-to-end deep reinforcement learning framework to improve energy efficiency.

III. METHOD

A. Hardware Design

The Duke Humanoid (Fig. 2) is a child-sized, open-source bipedal robot designed for dynamic locomotion. Standing 1 m from shoulder to foot and weighing 30 kg, it approximately matches the size of an average 8-year-old child [39]. This size was selected for two reasons: first, it allows the robot to interact comfortably with human-centric environment surfaces while standing, enabling future object manipulation tasks; second, its mid-size balances functionality, affordability, and customizability, especially given the high cost of commercially available geared motors.

Following a co-design philosophy for complex robotic platforms [40], we iteratively tuned and selected the mechanical design and actuators based on simulations. We designed preliminary humanoid models in Isaac Gym and trained walking policies via reinforcement learning, using the resulting joint torque and velocity data to guide actuator selection.

Mechanical Design Our robot features 10 degrees of freedom (DoFs), with 5 DoFs in each leg: 3 in the hip, 1 for knee flexion and extension (KFE), and 1 for ankle plantar flexion and dorsiflexion. The DoFs for the hip include hip rotation (HR), hip flexion and extension (HFE), and hip abduction and adduction (HAA). The center of mass of the hip is aligned in the frontal plane, allowing the robot to maintain balance and walk with straight knees.

The proportions of our leg design anatomically match that of a real human. The tibia measures 0.235 m, and the femur measures 0.285 m, resulting in a tibia-to-femur ratio of 0.825, which is comparable to the average human ratio of 0.78 [41]. In addition, the range of motion of our leg joints largely covers that of a real human, as shown in Table II.

The KFE joint is actuated by a motor via a parallel linkage. Similarly, the ankle motor, positioned at the KFE joint,

Joint	HR	HAA	HFE	KFE	Ankle
Human [deg]	[-50,40]	[-40,20]	[-30,110]	[0,150]	[-20,50]
Ours [deg]	[-90,60]	[-40,40]	[-90,90]	[0,110]	[-45,45]
Coverage	100%	100%	85%	73%	95%

TABLE II: Range of Motion: Our design mostly covers the range of motion of human leg joints [5].

connects to the ankle through another parallel linkage. As a result, the ankle joint is actuated in conjunction with the KFE and ankle motors, with a linear relationship between their movements. The parallel linkages relocate the motors higher in the leg, raising the leg’s center of mass and reducing the lower leg’s inertia, thereby reducing the torque required to actuate the KFE and hip motors. To save weight and simplify the design, the ankle does not include the inversion and eversion DoF. To compensate this DoF, the foot is designed with a linear contact surface to ensure consistent contact regardless of the roll angle.

Actuators Each joint is powered by a brushless DC motor with a planetary gearbox from Motorevo. The joint position is measured by a single-turn encoder at the back of the motor shaft. To simplify the design, the same brushless DC motor is used throughout, and varying reduction ratios are employed to satisfy the torque and velocity demands of each joint. Specifically, The HR, HAA, and KFE joints use 18:1 motors, providing $72 \text{ N} \cdot \text{m}$ rated torque and 20 rad/s maximum velocity; the HFE joints use 20:1 motors, providing $80 \text{ N} \cdot \text{m}$ rated torque and 18 rad/s max velocity; the ankles are powered by 10:1 motors, providing $40 \text{ N} \cdot \text{m}$ rated torques and 36 rad/s max velocity. This uniform motor selection provides flexibility, as the motors share identical rotors and mounting configurations, making them interchangeable. Motion control is achieved through EtherCAT communication at 2000 Hz to minimize latency.

Manufacturing The outer frame of the upper body is built using off-the-shelf aluminum extrusions, while the body plates and legs are custom-machined from aluminum 6061. This modular design allows for easy modifications. The upper body has enough space ($250 \times 120 \times 200 \text{ mm}$) reserved for future onboard power and computing to accommodate outdoor experiments. The body plates feature a grid of screw holes, allowing the attachment of custom 3D-printed electronic mounts. Additionally, the leg motors are covered with 3D-printed thermoplastic polyurethane (TPU) parts to protect against collisions.

Sensors An IMU (MicroStrain 3DM-CV7-AHRS), mounted on the base plate, provides angular velocity and orientation data. Two load cells are included at the toes, and two load cells are included at the heels. A Teesny 4.0 microcontroller interfaces with the IMU and the load cells and transmits the data to a computer through a serial connection.

B. Reinforcement Learning for Bipedal Locomotion

Task and Observation To demonstrate the locomotion capability of our platform, we develop a baseline control policy using RL. The task is to train the robot to follow

Reward Terms	Definition	Weight
Linear velocity	$\phi(\mathbf{v}_b^* - \mathbf{v}_b, -4 * [1, 1, 0.1])$	1
Angular velocity	$\phi(\omega_b^* - \omega_b, -8 * [0.1, 0.1, 1])$	0.5
Base height	$\phi(h_b^* - h_b, -2000)$	0.1
Orientation	$\max(\ \mathbf{g}_{b,xy}\ ^2, 0.1)$	-20
Joint acceleration	$\phi(\dot{\mathbf{q}}, -10^{-4})$	0.1
Joint torque	$\exp(G^{-1}\ \tau\ _1)$	0.05
Joint limit	$\ \mathbf{q} - \text{CLIP}(\mathbf{q}, \mathbf{q}_{\min}, \mathbf{q}_{\max})\ ^2$	-100
Feet air time	$\sum_{i=0}^{n_f} (t_{air,i} - 0.3)$	1
Contact pattern	$1 - n_f^{-1}\ c^* - c\ _1$	0.5
Feet orientation	$\phi(\mathbf{g}_{f,xy}, -8)$	0.1
Feet forward	$\phi(\mathbf{f}_{f,xy}, -8)$	0.1
Feet position	$\phi(\mathbf{p}_f^* - \mathbf{p}_f, [-1000, -1000, 0])$	0.1
Feet step height	$\ \max(h_f, h_f^*)\ _1$	0.5

TABLE III: Baseline RL policy reward functions.

a desired linear velocity \mathbf{v}^* and angular velocity ω^* . Our policy observation includes the base angular velocity \mathbf{v}_b , the base gravity vector \mathbf{g} projected onto the robot's base frame \mathbf{g}_b , joint positions \mathbf{q} , joint velocities $\dot{\mathbf{q}}$, and target body velocity commands, consisting of the linear velocities in the xy-plane and angular velocity around the z-axis, $[\mathbf{v}_{b,x}^*, \mathbf{v}_{b,y}^*, \omega_{b,z}^*]$.

Observations Additionally, the observation also includes a periodic phase signal ϕ encoded as $[\sin \phi, \cos \phi]$, and a desired foot contact pattern c^* derived from the periodic phase signal. We employ an asymmetric actor-critic architecture [42], where the critic network receives additional privileged information such as linear velocity (\mathbf{v}_b) and actual foot contact indicator (c). The actor and critic networks consist of three fully connected layers with hidden sizes of [512, 256, 128] with ELU activation functions [43].

Actions Our RL policy outputs \mathbf{a}_q which is converted to target joint positions ($\mathbf{q}^* = \mathbf{q}_0 + k_q \mathbf{a}_q$) at 50Hz. Where \mathbf{q}_0 are the default joint positions, k_q is a scaling factor for the action. The positions are smoothed with a first-order exponential low-pass filter before being converted to torques via a position PD controller, which operates at 200 Hz in simulation. In the real robot, the position PD controller runs at 2000 Hz to ensure smooth control. The PD gains are empirically determined to be $60 \text{ N} \cdot \text{m} \cdot \text{rad}^{-1}$ and $5 \text{ N} \cdot \text{m} \cdot \text{s} \cdot \text{rad}^{-1}$.

Reward The reward function (Table III) includes terms for tracking and regularization (body, joint, and gait). Here, $\phi(\psi, w) := \exp(\sum_{i=1}^n w_i \psi_i^2)$ represents the exponential of the sum of the squared quantity ψ weighted by w . \mathbf{v}_b , ω_b , h_b , and \mathbf{g}_b represent the base linear velocity, angular velocity, height, and projected gravity. \mathbf{q}_{\min} and \mathbf{q}_{\max} represent the minimum and maximum joint position limits. τ refers to joint actuation torque, $\dot{\mathbf{q}}$ to joint accelerations, and G to the robot's total gravity. t_{air} is the duration the feet remain airborne, and $n_f = 2$ is the number of feet. Finally, \mathbf{g}_f and \mathbf{f}_f are the feet's gravity and forward direction vectors, while \mathbf{p}_f and \mathbf{p}_f^* are the actual and desired feet positions, and h_f and h_f^*

denote the actual and desired step heights.

Training We trained 4096 parallel environments with Isaac Gym [44] for 30 min on a single NVIDIA RTX 4090 GPU.

C. Policy for energy-efficient locomotion

Energy efficiency in humanoid locomotion is important because the majority of the energy is used to actuate the joint during locomotion. Thus, in addition to our baseline RL policy, we integrate passive dynamics into the end-to-end reinforcement learning framework. We hypothesize that encouraging passive control in locomotion can enable a more energy-efficient controller. To achieve this, we propose explicitly modulating the joint torques to switch between passive and active modes. Specifically, we introduce a per-joint torque activation parameter, α , which scales existing position-based PD control. When $\alpha = 0$, the joint torque will be scaled to zero, making the joint completely passive, allowing its movement to be governed by passive dynamics. The modified control is expressed as:

$$\tau = \alpha(k_p(\mathbf{q}^* - \mathbf{q}) + k_d(\dot{\mathbf{q}}^* - \dot{\mathbf{q}}))$$

where k_p and k_d are stiffness and damping factors. \mathbf{q} and \mathbf{q}^* are the desired and actual joint position. $\dot{\mathbf{q}}^* = 0$ is the desired joint velocity (always 0), and $\dot{\mathbf{q}}$ is the actual joint velocity. The output of the new actor network is $[\mathbf{a}_q, \mathbf{a}_\alpha] \in \mathbb{R}^{2 \times \text{DoF}}$. A sigmoid function constrains $\mathbf{a}_\alpha \in [0, 1]$, where the constant k is empirically set to 10. $\alpha_0 \in [0, 1]$ is a constant that controls the minimum activeness of the scaling factor:

$$\alpha = \alpha_0 + (1 - \alpha_0) \cdot \text{sigmoid}(k\mathbf{a}_\alpha) = \alpha_0 + \frac{(1 - \alpha_0)}{1 + e^{-k\mathbf{a}_\alpha}}$$

To encourage the use of passive actions, we introduce a passive action reward, defined as $r_{\mathbf{a}_\alpha} = \|1 - \mathbf{a}_\alpha\|_1$, scaled by 0.005. In addition, to prevent the policy from getting stuck at small α values, we implement a linear curriculum for α_0 to decay initially from $\alpha_0 = 0.5$ to $\alpha_0 = 0$.

D. Bridging the Sim-to-Real Gap

Transferring trained RL policies from simulation to the real robot, particularly for complex platforms such as humanoids, remains challenging. Several factors contribute to the sim-to-real gap: Motor backlash can cause discontinuities in actuation, sensor noise introduces uncertainty in observations, and dynamics mismatches, such as differences in joint friction, link mass, and inertia, can change the underlying transition probabilities.

To bridge these gaps, we first identified critical joint dynamics parameters that are tunable in simulation, including damping, friction, and armature. For each joint, we conducted control frequency sweeps between 1.5-3 Hz to reduce discrepancies between simulation and real-world performance, specifically focusing on joint position, velocity, target position, and torque. This identification process enabled a closer match between the robot's simulated and real-world behavior.

We also applied domain randomization to increase the robustness of the simulation policy for real-world transfer.

Parameter	Unit	Range	Operator
Base COM	m	[-0.02, 0.02]	additive
Base mass	kg	[-0.5, 5]	additive
Link mass	kg	[0.95, 1.05]	scaling
Link inertia	kg · m ²	[0.95, 1.05]	scaling
Joint offset	rad	[-0.02, 0.02]	additive
Joint position	rad	[-0.03, 0.03]	additive
Joint velocity	rad/s	[-1.5, 1.5]	additive
Motor strength	-	[0.98, 1.02]	scaling
Observation delay	ms	[0, 20]	-
Angular velocity	rad/s	[-0.2, 0.2]	additive
Gravity vector	-	[-0.05, 0.05]	additive
Friction	-	[0.2, 1.2]	-

TABLE IV: Domain randomization parameters.

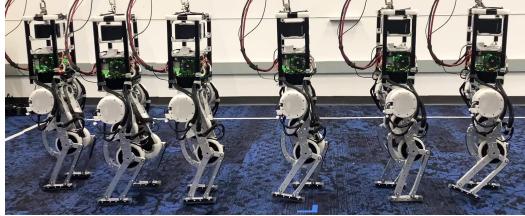


Fig. 3: Chronophotograph showing the Duke Humanoid walking using the baseline RL policy.

Table IV outlines the key parameters. Mass and inertia randomizations compensate for inconsistencies introduced by moving cables, electronics, loads, and mechanical model inaccuracies. Joint offset randomizations address variations in default joint positions after power-up and calibration. Position, velocity, and motor strength randomizations mitigate the effects of motor backlash, encoder noise, and actuation inconsistencies. Additionally, angular velocity and gravity vector randomizations simulate IMU noise, and friction randomizations consider locomotion on different surfaces.

IV. EXPERIMENTS AND RESULTS

This section demonstrates the performance of our baseline locomotion policy in a zero-shot sim-to-real setup on our humanoid platform. Additionally, we evaluate the real-world performance of our passive walking policy. To further validate the effectiveness of our passive walking policy, we compare its energy efficiency and gait characteristics against the baseline policy in simulation and reality.

A. Locomotion in the Real

We deployed our baseline policy zero-shot to our humanoid hardware. Fig. 3 shows a chronophotograph of the robot walking forward on a flat carpeted surface at a command velocity of 0.3 m/s.

As shown in Fig. 4, we quantitatively compared the target joint positions, the actual joint positions, and the joint velocities of the HFE and KFE joints between simulation and real-world performance. We chose to focus on these joints because they are critical for forward movements. We aligned the first steps of the simulation with the real data and compared the relative target joint positions. The maximum difference in the target joint position is 0.1 rad for HFE and 0.15 rad for KFE, indicating that the policy produces similar

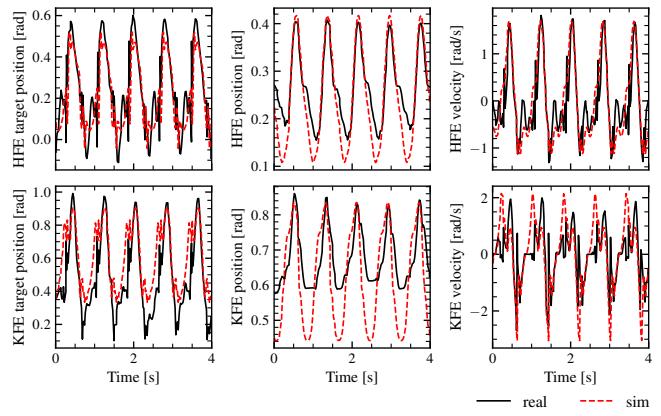


Fig. 4: **Comparison of Simulated and Real-World Baseline Walking:** Target joint positions (from RL policy), actual joint positions, and joint velocities for the HFE and KFE joints.

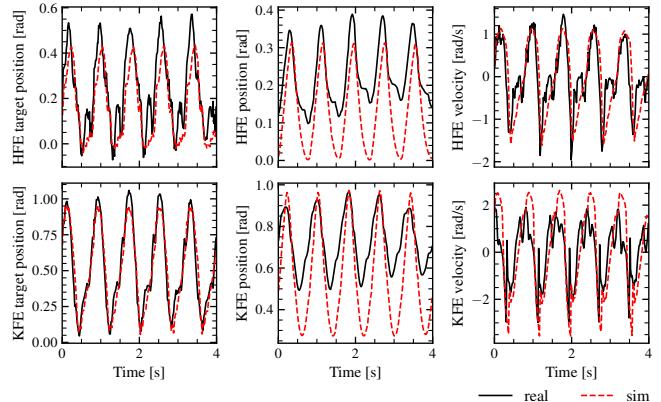


Fig. 5: **Comparison of Simulated and Real-World Passive Walking:** Target joint positions (from RL policy), actual joint positions, and joint velocities for the HFE and KFE joints.

actions in the simulation and the real world. For the actual joint positions, the maximum error is 0.1 rad for HFE and 0.2 rad for KFE, which is only significant during half of the gait cycle. Overall, the differences between the target joint positions and joint velocities are bounded, resulting in a similar behavior between the simulation and the real system.

B. Passive Policy for Energy-Efficient Locomotion

The trained passive policy behaves visually differently compared to the baseline policy at zero velocity. The baseline policy actively steps in place, whereas the passive policy remains motionless, enhancing energy efficiency.

Energy Efficiency A key metric for quantifying the energy efficiency of robot locomotion is the Cost of Transport (CoT). A lower CoT translates to higher energy efficiency. CoT is defined as the energy required for locomotion, normalized by the distance traveled and the robot weight [45]:

$$CoT = \frac{W_{input}}{m \cdot g \cdot d}$$

where $m \cdot g$ is the robot's total gravity, W_{input} is the energy input, and d is the measured distance traveled. Ideally, the W_{input} is obtained by measuring the electrical power usage

at the power supply, however, this term cannot be directly measured in simulation. In the context of legged robots, the majority of the power comes from joint actuation. Therefore, we approximate the energy input by integrating the motor's mechanical power.

To evaluate the effectiveness of our passive control algorithm, we trained a passive policy for walking (stance ratio of 0.6) and running (stance ratio of 0.38) in the simulation. We then compared the passive policy with its active counterpart across a range of command velocities, from 0.1 m/s to 1 m/s. As shown in Fig. 6, at low walking speeds, from 0.1 m/s to 0.5 m/s, the passive policy outperforms the baseline in CoT, achieving up to a 50% reduction at 0.1 m/s. The range of motion for both HFE and KFE is lower under the passive policy at low velocities. Base height (the IMU's distance from the ground) and foot heights are similar between the baseline and passive policies, indicating comparable gait patterns. The passive policy exhibits a larger base orientation error for walking, averaging 0.01 rad, suggesting that the robot maintains a forward tilt to take advantage of passive dynamics. For running, as shown in Fig. 7, the passive policy also outperforms the active baseline in running, albeit by a smaller margin of 10%. The range of motion for HFE is reduced compared to the baseline policy; the base height variability is increased, suggesting a shift in the center of mass. Other metrics remain comparable between the passive and baseline policy.

Robustness We evaluated the robustness of the learned policy by measuring the success rate of push recovery in simulation. In each trial, a random horizontal force sampled uniformly within a disk of 300 N is applied to the humanoid base body between 2-3 s for a duration of 0.1 s. Push recovery is considered successful if the robot does not fall or self-collide for 5 s. Fig. 8 compares the push recovery success rates of the baseline and passive policies trained for walking and running. The passive walking policy achieves similar success rates in push recovery as the baseline walking policy, while the passive running policy performs at a slightly reduced success rate at lower push magnitude, and slightly better in higher magnitude push recovery. The result indicates that the robustness of the control policy remains mostly unchanged with the addition of passive actions.

C. Passive Locomotion in the Real

We deployed the passive control policy zero-shot on our humanoid platform. The supplementary video provides a visual comparison of the walking performance baseline and passive policies. As shown in Fig. 5, the target and actual joint positions and velocities for the HFE and KFE joints suggest a reasonable agreement between simulation and real-world execution. This observation aligns with the sim-to-real performance observed for the active baseline (Fig. 4).

To visualize the joint actuation pattern, we show the single (left) leg dynamics of controlled joint activation and absolute torque during passive walking (Fig. 9). At a command velocity of 0.3 m/s, all hip joints exhibit periodic activation patterns. The KFE deactivates during leg swing, suggesting

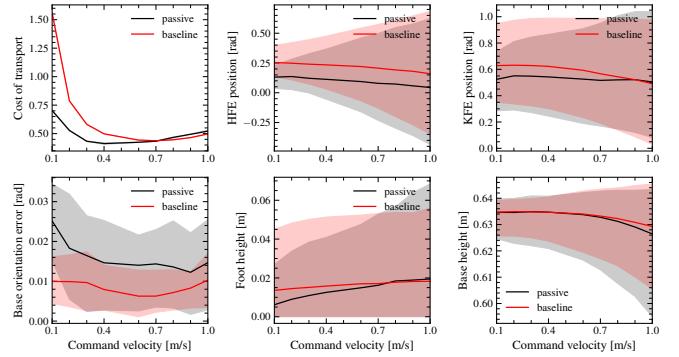


Fig. 6: **Comparison of baseline vs. passive policy for walking in simulation:** Red/grey shaded regions represent the 95% confidence interval. The passive policy exhibits up to 50% better energy efficiency at low speed, with reduced HFE and KFE positions, indicating less action in walking.

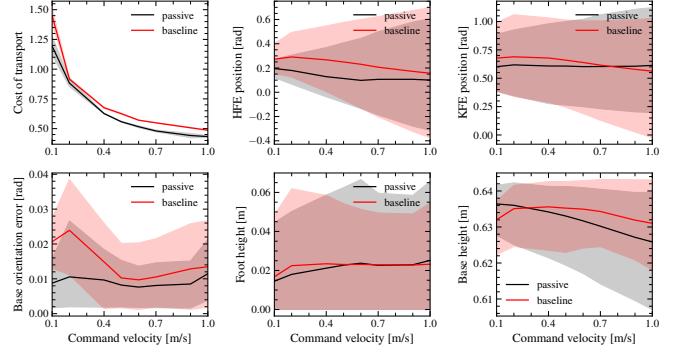


Fig. 7: **Comparison of baseline vs. passive policy for running in simulation:** Red/grey shaded regions represent the 95% confidence interval. The passive policy generally outperforms the baseline by 10% and exhibits a lower HFE position.

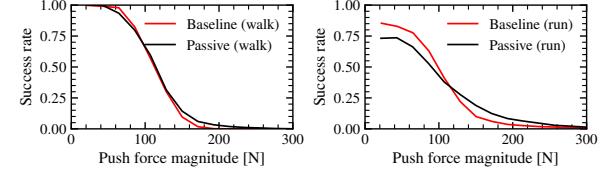


Fig. 8: **Push recovery comparison:** baseline vs. passive policy for walking (left) and running (right) in simulation.

the use of passive dynamics for leg swing. The ankle joint does not display a significant periodic activation pattern, likely because high joint friction reduces the lower leg's tendency to passively rotate. We confirmed similar periodic patterns for the right leg. Note that the joint actuation pattern does not exactly match the human pattern because the humanoid does not exactly match human morphology such as mass distribution.

We compared the real-world CoT of the passive policy with the baseline policy in three repeated trials at a command forward velocity of 0.3 m/s. The baseline policy achieved an average CoT of 1.13 ± 0.1 , and the passive policy improved the CoT to 0.77 ± 0.1 , resulting in a 31% CoT reduction. Both simulation (Fig. 6) and real-world results show that the passive policy outperforms the baseline in energy efficiency during low-speed walking.

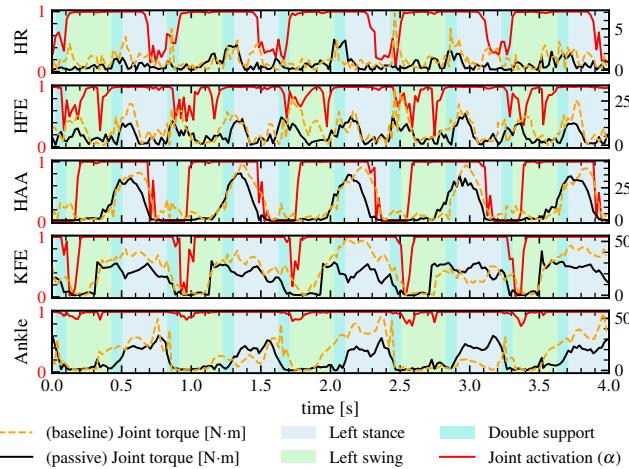


Fig. 9: Comparison of left leg absolute joint torque for real-world baseline vs. passive walking: KFE deactivation during swing in passive walking demonstrates the utilization of passive dynamics, contrasting with the generally higher torque of the baseline policy.

V. CONCLUSION, LIMITATION, AND FUTURE WORK

We introduce the Duke Humanoid v1.0, an open-source hardware and software research platform for studying humanoid locomotion. This paper details our design choices and the development of a baseline reinforcement learning policy for bipedal locomotion. Furthermore, we explore directly modulating joint activation within the end-to-end reinforcement learning framework to leverage passive dynamics for energy-efficient humanoid walking. Preliminary real-world experiments demonstrate that our approach achieved a 31% reduction in the cost of transport.

However, our platform currently lacks arms, limiting our ability to study whole-body control and mobile manipulation. In the future, we plan to equip the robot with arms and dexterous hands. Our robot is currently tethered, limiting its mobility. We plan on implementing onboard power to study more dynamic motions, such as running and jumping.

REFERENCES

- [1] B. Dynamics. (2024) Boston dynamics atlas. [Online]. Available: <https://bostondynamics.com/atlas>
- [2] Agility. (2024) Agility digit robot. [Online]. Available: <https://agilityrobotics.com>
- [3] Unitree. (2024) Unitree g1. [Online]. Available: <https://www.unitree.com/g1>
- [4] Tesla. (2024) Tesla bot. [Online]. Available: <https://www.tesla.com/AI>
- [5] Q. Liao, B. Zhang, X. Huang, X. Huang, Z. Li, and K. Sreenath, “Berkeley humanoid: A research platform for learning-based control,” *arXiv preprint arXiv:2407.21781*, 2024.
- [6] J. Li, J. Ma, O. Kolt, M. Shah, and Q. Nguyen, “Dynamic loco-manipulation on hector: Humanoid for enhanced control and open-source research,” *arXiv preprint arXiv:2312.11868*, 2023.
- [7] M. Lapeyre, P. Rouanet, J. Grizou, S. Nguyen, F. Depraetere, A. Le Falher, and P.-Y. Oudeyer, “Poppy project: open-source fabrication of 3d printed humanoid robot for science, education and art,” in *Digital Intelligence 2014*, 2014, p. 6.
- [8] G. Ficht, H. Farazi, A. Brandenburger, D. Rodriguez, D. Pavlichenko, P. Allgeuer, M. Hosseini, and S. Behnke, “Nimbro-op2x: Adult-sized open-source 3d printed humanoid robot,” in *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2018, pp. 1–9.
- [9] G. Metta, L. Natale, F. Nori, G. Sandini, D. Vernon, L. Fadiga, C. Von Hofsten, K. Rosander, M. Lopes, J. Santos-Victor *et al.*, “The icub humanoid robot: An open-systems platform for research in cognitive development,” *Neural networks*, vol. 23, no. 8-9, pp. 1125–1134, 2010.
- [10] A. D. Kuo, “The six determinants of gait and the inverted pendulum analogy: A dynamic walking perspective,” *Human movement science*, vol. 26, no. 4, pp. 617–656, 2007.
- [11] G. A. Cavagna, H. Thys, and A. Zamboni, “The sources of external work in level walking and running.” *The Journal of physiology*, vol. 262, no. 3, pp. 639–657, 1976.
- [12] M. Garcia, A. Chatterjee, and A. Ruina, “Efficiency, speed, and scaling of two-dimensional passive-dynamic walking,” *Dynamics and Stability of Systems*, vol. 15, no. 2, pp. 75–99, 2000.
- [13] H. Ohta, M. Yamakita, and K. Furuta, “From passive to active dynamic walking,” *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, vol. 11, no. 3, pp. 287–303, 2001.
- [14] S. Collins, A. Ruina, R. Tedrake, and M. Wisse, “Efficient bipedal robots based on passive-dynamic walkers,” *Science*, vol. 307, no. 5712, pp. 1082–1085, 2005.
- [15] M. Wisse and J. Van Frankenhuyzen, “Design and construction of mike; a 2-d autonomous biped based on passive dynamic walking,” in *Adaptive motion of animals and machines*. Springer, 2006, pp. 143–154.
- [16] Y. Yang, T. Zhang, E. Coumans, J. Tan, and B. Boots, “Fast and efficient locomotion via learned gait transitions,” in *Conference on robot learning*. PMLR, 2022, pp. 773–783.
- [17] A. Hartmann, D. Kang, F. Zargarbashi, M. Zamora, and S. Coros, “Deep compliant control for legged robots,” in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 11421–11427.
- [18] P. Kormushev, B. Ugurlu, D. G. Caldwell, and N. G. Tsagarakis, “Learning to exploit passive compliance for energy-efficient gait generation on a compliant humanoid,” *Autonomous Robots*, vol. 43, pp. 79–95, 2019.
- [19] J. Hwangbo, J. Lee, A. Dosovitskiy, D. Bellicoso, V. Tsounis, V. Koltun, and M. Hutter, “Learning agile and dynamic motor skills for legged robots,” *Science Robotics*, vol. 4, no. 26, p. eaau5872, 2019.

- [20] T. Miki, J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, “Learning robust perceptive locomotion for quadrupedal robots in the wild,” *Science Robotics*, vol. 7, no. 62, p. eabk2822, 2022.
- [21] J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, “Learning quadrupedal locomotion over challenging terrain,” *Science robotics*, vol. 5, no. 47, p. eabc5986, 2020.
- [22] N. G. Tsagarakis, S. Morfey, G. M. Cerdà, L. Zhibin, and D. G. Caldwell, “Compliant humanoid coman: Optimal joint stiffness tuning for modal frequency control,” in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 673–678.
- [23] A. SaLoutos, E. Stanger-Joncs, Y. Ding, M. Chignoli, and S. Kim, “Design and development of the mit humanoid: A dynamic and robust research platform,” in *2023 IEEE-RAS 22nd International Conference on Humanoid Robots (Humanoids)*. IEEE, 2023.
- [24] J. Carpentier and P.-B. Wieber, “Recent progress in legged robots locomotion control,” *Current Robotics Reports*, vol. 2, no. 3, pp. 231–238, 2021.
- [25] J. Englsberger, A. Werner, C. Ott, B. Henze, M. A. Roa, G. Garofalo, R. Burger, A. Beyer, O. Eiberger, K. Schmid *et al.*, “Overview of the torque-controlled humanoid robot toro,” in *2014 IEEE-RAS International Conference on Humanoid Robots*. IEEE, 2014, pp. 916–923.
- [26] Y. Asano, T. Kozuki, S. Ookubo, M. Kawamura, S. Nakashima, T. Katayama, I. Yanokura, T. Hirose, K. Kawaharazuka, S. Makino *et al.*, “Human mimetic musculoskeletal humanoid kengoro toward real world physically interactive actions,” in *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2016, pp. 876–883.
- [27] M. S. Ahn, *Development and Real-Time Optimization-based Control of a Full-sized Humanoid for Dynamic Walking and Running*. University of California, Los Angeles, 2023.
- [28] T. Sugihara, Y. Nakamura, and H. Inoue, “Real-time humanoid motion generation through zmp manipulation based on inverted pendulum control,” in *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292)*, vol. 2. IEEE, 2002, pp. 1404–1409.
- [29] N. Scianca, D. De Simone, L. Lanari, and G. Oriolo, “Mpc for humanoid gait generation: Stability and feasibility,” *IEEE Transactions on Robotics*, vol. 36, no. 4, pp. 1171–1188, 2020.
- [30] L. Sentis and O. Khatib, “A whole-body control framework for humanoids operating in human environments,” in *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006*. IEEE, 2006, pp. 2641–2648.
- [31] I. Radosavovic, T. Xiao, B. Zhang, T. Darrell, J. Malik, and K. Sreenath, “Real-world humanoid locomotion with reinforcement learning,” *Science Robotics*, vol. 9, no. 89, p. eadi9579, 2024.
- [32] X. Gu, Y.-J. Wang, X. Zhu, C. Shi, Y. Guo, Y. Liu, and J. Chen, “Advancing humanoid locomotion: Mastering challenging terrains with denoising world model learning,” *arXiv preprint arXiv:2408.14472*, 2024.
- [33] M. Seo, S. Han, K. Sim, S. H. Bang, C. Gonzalez, L. Sentis, and Y. Zhu, “Deep imitation learning for humanoid loco-manipulation through human teleoperation,” in *2023 IEEE-RAS 22nd International Conference on Humanoid Robots (Humanoids)*. IEEE, 2023.
- [34] X. B. Peng, P. Abbeel, S. Levine, and M. van de Panne, “Deepmimic: Example-guided deep reinforcement learning of physics-based character skills,” *ACM Trans. Graph.*, vol. 37, no. 4, pp. 143:1–143:14, Jul. 2018. [Online]. Available: <http://doi.acm.org/10.1145/3197517.3201311>
- [35] D. Makoviichuk and V. Makoviychuk, “rl-games: A high-performance framework for reinforcement learning,” https://github.com/Denys88/rl_games, May 2021.
- [36] 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 preprint arXiv:2108.10470*, 2021.
- [37] Y. I. K. Y. A. S. H. Fujimoto. (2008) A study of the leg-swing motion of passive walking. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/4543428>
- [38] J. Reher, E. A. Cousineau, A. Hereid, C. M. Hubicki, and A. D. Ames, “Realizing dynamic and efficient bipedal locomotion on the humanoid robot durus,” in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 1794–1801.
- [39] U. C. for Disease Control and Prevention. (2017) Clinical growth charts. [Online]. Available: https://www.cdc.gov/growthcharts/clinical_charts.htm
- [40] Q. Li, W. Zhang, and L. Chen, “Design for control-a concurrent engineering approach for mechatronic systems design,” *IEEE/ASME Transactions on Mechatronics*, vol. 6, no. 2, pp. 161–169, Jun. 2001. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/928731>
- [41] S. Aitken, “Normative values for femoral length, tibial length, and the femorotibial ratio in adults using standing full-length radiography,” *Osteology*, vol. 1, pp. 86–91, 05 2021.
- [42] D. Chen, B. Zhou, V. Koltun, and P. Krähenbühl, “Learning by cheating,” in *Conference on Robot Learning*. PMLR, 2020, pp. 66–75.
- [43] D.-A. Clevert, “Fast and accurate deep network learning by exponential linear units (elus),” *arXiv preprint arXiv:1511.07289*, 2015.
- [44] N. Rudin, D. Hoeller, P. Reist, and M. Hutter, “Learning to walk in minutes using massively parallel deep reinforcement learning,” 2021.
- [45] D. Zarrouk and R. S. Fearing, “Cost of locomotion of a dynamic hexapedal robot,” in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 2548–2553.