

# Supplementary Materials – FT-Net: Learning Failure Recovery and Fault-tolerant Locomotion For Quadruped Robots

## Abstract

This document is structured as follows. Section I demonstrates the trajectory of zero moment point (ZMP) and center of pressure (COP) during the motion with the discussion of the results. Section II presents the theoretical background for the feature attribution analysis to showcase the interior mechanism when the failure occurs, including the SHAP and the reestablish net. Finally, in Section III, we conduct the comprehensive evaluation for the heuristic guidance including using other criteria such as feasible region stability [1]. Finally, several experiments are complemented for the illustration in the main article.

## I. ZERO MOMENT POINT (ZMP) AND CENTER OF PRESSURE (COP)

The computation of COP is outlined in the main article, which involves determining the weighted summation of the points of contact. Concerning ZMP, as detailed in references in [2], [3], the initial step entails computing the overall linear momentum  $\mathcal{P} = [\mathcal{P}_x, \mathcal{P}_y, \mathcal{P}_z]^T$  and angular momentum  $\mathcal{L} = [\mathcal{L}_x, \mathcal{L}_y, \mathcal{L}_z]^T$  relative to the world frame.

$$\mathcal{P} = \sum_{i=1}^N m_i \dot{\mathbf{c}}_i \quad \mathcal{L} = \sum_{i=1}^N \mathbf{c}_i \times (m_i \dot{\mathbf{c}}_i) + \mathbf{R}_i \mathbf{I}_i \mathbf{R}_i^T \omega_i \quad (1)$$

where  $c_i$ s is the COM positions of rigid bodies, and  $\mathbf{R}_i \mathbf{I}_i \mathbf{R}_i^T$  is the inertia tensor in the world frame. We can then obtain the ZMP projection position on the ground plane  $p_x, p_y$  as

$$p_x = \frac{Mgx - \dot{\mathcal{L}}_y}{Mg + \dot{\mathcal{P}}_z} \quad p_y = \frac{Mgy - \dot{\mathcal{L}}_x}{Mg + \dot{\mathcal{P}}_z}, \quad (2)$$

During the normal state, the robot employs a walking pattern, which is a trotting gait (Fig. S1) where the robot moves its legs in unison and makes contact with the ground in diagonal pairs [4], [5]. In this period, the Zero Moment Point (ZMP) displays minimal deviation from the Center of Pressure (COP), as depicted in Fig. S2. However, when the robot transitions into a faulty state, the computed ZMP could fall beyond the boundaries of the support polygon, as illustrated in Fig. S3(a)(iv). This happens only when the contact region is a line or a point so that the ground reaction force acts on the edge of the support polygon. In such cases, the ZMP is referred to as the fictitious ZMP (FZMP) [2], and a rotation about the contact edge is initiated and the robot is prone to overturn.

After the failure is triggered, the COP undergoes significantly broader and more intense movement throughout the motion. We further perform comparative experiments, in which the ZMP-base inverted pendulum (ZMP-IP in Fig. S3(a)(iv)) and COP-IP are respectively adopted for the derivation of heuristic rewards and training. The learning curves in Fig. S3(b) clearly indicate the superior performance of the COP-IP configuration with higher accumulated reward.

The aforementioned outcomes provide validation for our methodology, which employs COP to model the inverted pendulum compared to ZMP, resulting in enhanced performance. Essentially, according to the ZMP's definition, the overall moment with respect to ZMP is balanced, encompassing the influences of gravity, inertia force, and ground reaction force. Consequently, the angular acceleration of any rigid body with respect to ZMP remains zero. In contrast, at COP, only the moments generated by the ground reaction forces are self-counteracted, while the impacts of gravity and inertia forces persist. This distinction underlies our approach's choice of penalizing the angular acceleration of COM w.r.t. COP. This strategy aims to minimize the non-zero angular acceleration resulting from overturning moments. By doing so, it effectively guides the robot in acquiring skills in failure recovery and fault-tolerant locomotion.

## II. ATTRIBUTE ANALYSIS OF THE FAILURE EMERGENCE

### A. System physical variation attributes

In this section, we introduce an examination of the FT-Net's capability to accommodate the physical variation associated with failures. The control strategy will be adjusted by the latent output of the adaptor. From the perspective of the robotic system's physical attributes, we evaluate how well the low-dimensional latent vector, denoted as  $e_t^\varphi$ , represents the fault scenario during the emergence of a failure event. A network referred to as  $E_\varphi^R$  is trained to reestablish the latent information  $e_t^\varphi$  into a format compatible with the input of encoder  $E_\theta$ . The reestablishment net  $E_\varphi^R$  endeavors to estimate the mean  $\hat{\mu}_\theta$  and the standard deviation  $\sigma_\theta$  of the reestablished values. Given the real physical parameters  $\mu_\theta$  acquired from the simulation, it is trained by minimizing the loss

$$\mathcal{L}_{E_\varphi^R} = \frac{1}{2} (\log(\max(\sigma_\theta, \epsilon)) + \frac{(\hat{\mu}_\theta - \mu_\theta)^2}{\max(\sigma_\theta, \epsilon)}) + C, \quad (3)$$

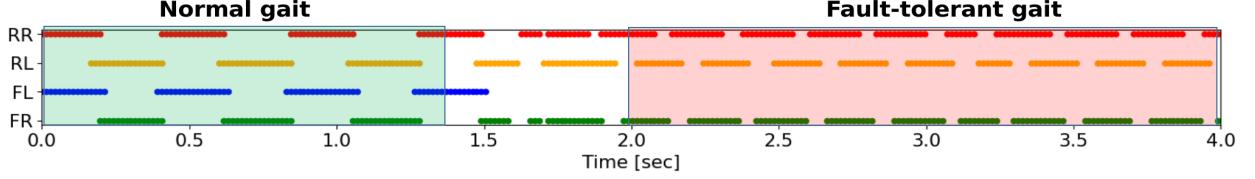


Fig. S1: Gait plot before and after the failure is triggered. The fault-tolerant gait is a self-adaptive gait depending on the fault scenarios.

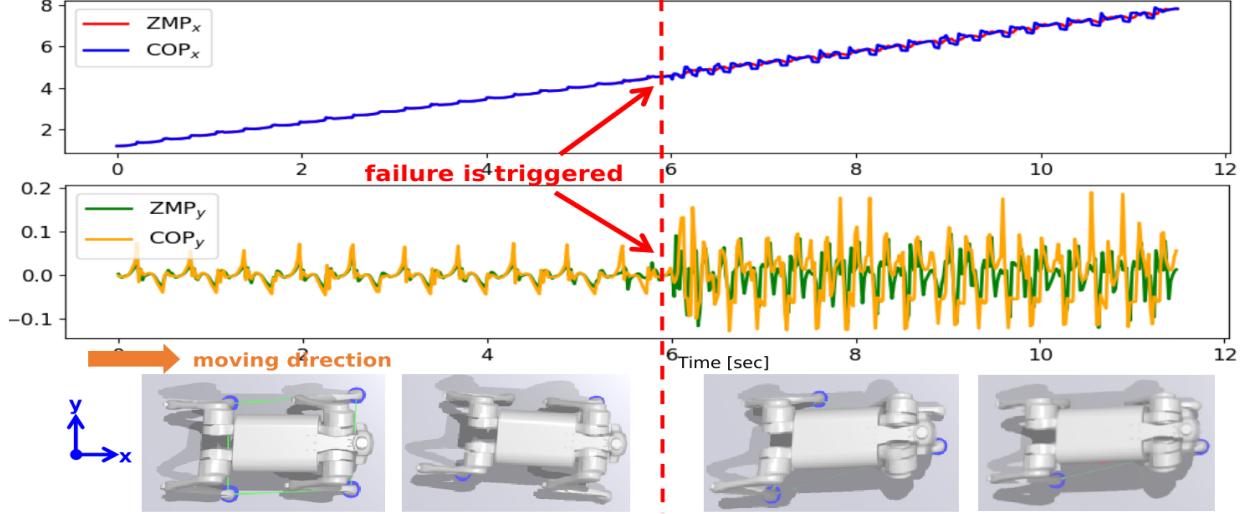


Fig. S2: Variation of ZMP and COP during the motion. ZMP and COP basically overlap in the normal trot gait but differ in the adaptive fault gait.

### B. Input contribution attributes

**SHAP** (SHapley Additive exPlanation) or Shapley value of each component within the proprioception history is calculated. This procedure fundamentally provides a way to explain the output of a neural network-based model using the game theory [6]. Within this theoretical framework, the predicted value  $e_t^\varphi$  given by the adaptor model  $E_\varphi$  corresponds to the aggregate outcomes or contributions  $\phi_i^j$  of all features, encompassing all the components of each timestep within proprioception history.

$$e_t^\varphi = \phi_0 + \sum_{i=1}^H \sum_{j=1}^D \phi_i^j z_i^j, \quad (4)$$

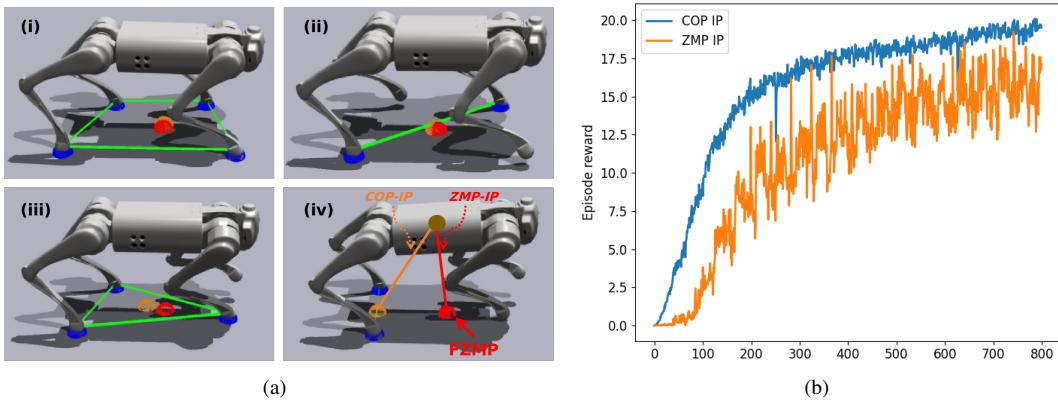


Fig. S3: Comparison related to ZMP and COP. (a) When the robot locomotes in the trot gait, the COP and ZMP basically coincide ((i), (ii)). When the robot locomotes in the adaptive faulty gait, the COP and ZMP differ significantly((iii), (iv)). (b) Learning curves of COP-IP and ZMP-IP models.

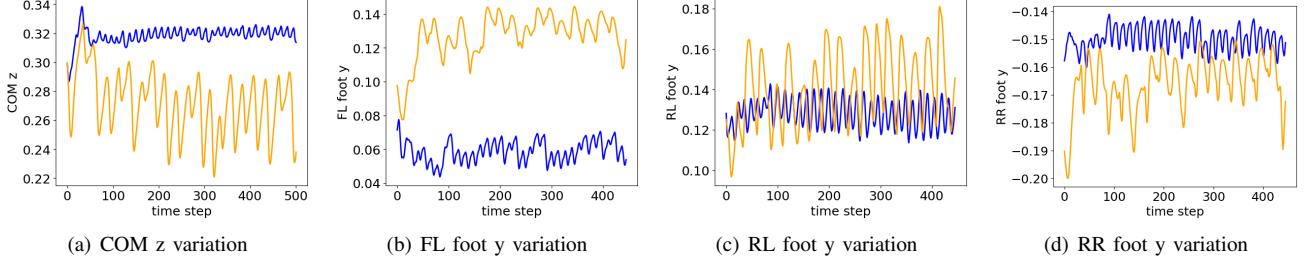


Fig. S4: Comparison of the COM position and foot position of normal feet between the policy with(blue) and without(orange) heuristic guiding terms when the Front Right (FR) leg is failed

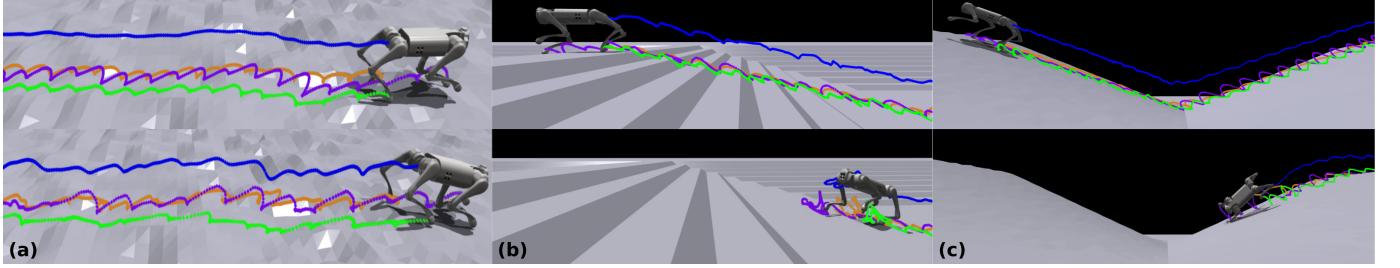


Fig. S5: Performance comparison between the control policy on (a) random surface, (b) stairs, (c) slope. It is worthwhile to note that (a) is within the terrain of the training range and (b) (c) terrain is unseen during training.  
(Upper) Policy with heuristic guiding terms; (Lower) Policy without heuristic guiding terms that are less stable.

where  $D$  is the dimension of every timestep's observation, and  $z_i^j \in \{0, 1\}^{H \cdot D}$  represents if the feature is observable (in our case they are all 1). The feature of  $j$ -th component in the  $i$ -th observation in the proprioceptive history buffer,  $\phi_i^j$  is eventually calculated as

$$\phi_i^j = \sum_{S \subseteq \mathcal{F} \setminus \{o_i^j\}} \frac{|S|! (H \cdot D - |S| - 1)!}{(H \cdot D)!} (f_{S \cup o_i^j}(S \cup o_i^j) - f_S(S)), \quad (5)$$

where  $\mathcal{F} \setminus \{o_i^j\}$  is the proprioception history buffer of each step except for  $o_i^j$  and  $f_S(S)$  is the prediction of the feature subset  $S$ . In the main article visualize the sum of the SHAP values of all observation components at each time instant, i.e.  $\sum_{j=1}^D \phi_i^j$ .

### III. EVALUATION OF FT-NET'S POLICY

#### A. Evaluation of FT-Net heuristic guiding terms

##### 1) Stability discussion on physical criteria:

To discuss and evaluate the effect of the heuristic guided reward which is formulated based on specific stability criteria, a series of experiments are devised to quantify the influence of this reward. Fig. S4 demonstrate that our guided reward terms facilitate the quadruped robot in learning to perform less unnecessary motion. This is evidenced by the decreased variation observed in both the z coordinates of the COM and the y coordinates of the feet. Besides, when the front right leg is malfunctioned, in the policy we trained with guided terms, the robot learns to adjust the position of the front left leg towards the center.

The experiments in Fig. S5 reflect that the policy with the heuristic reward terms can generalize the policy to recover from the faults arising in various terrains even on the terrains that the robot has not seen during the training stage Fig. S5(b) and S5(c)). As for the policy without heuristic terms, the robot is prone to take more aggressive action with longer step length. This behavior results in the easy overturn of the robot when the fault happens on uneven terrains. Even in the terrain featured in Fig S5(a), which is used to train our policies, the incorporation of heuristic terms leads to greater stability in both the foot placement and the COM behavior.

##### 2) Stability discussion on feasible region:

The feasible region in Orsolino et al.'s work [1] is re-implemented to evaluate our method. It finds a feasible region for COM ground projection within the support polygon by solving a linear programming problem, which guarantees that the locomotion is statically stable on rough terrains. Specifically, most scenarios of this criterion involve three or four stance legs, whereas our policy accommodates a maximum of three stance legs during faulty gait (even having two or just one stance leg). We project the swing legs along the world's Z axis to the ground to satisfy the condition of the feasible region's calculation and verify if the COM's ground projection is within the feasible region.

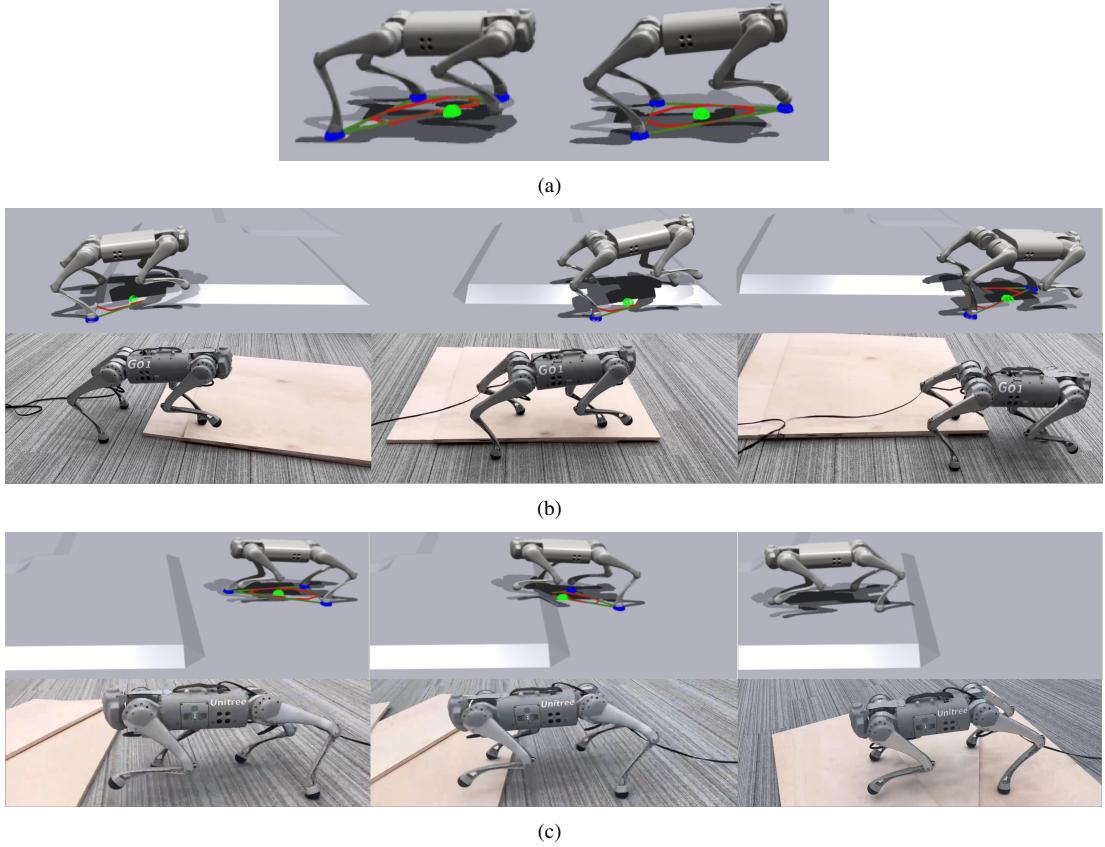


Fig. S6: (a) The robot with the unsafe movement when the COM projection falls outside (Left) and inside (Right) the feasible region (**Red**). They correspond to the policy without (Left) and with (Right) our heuristic terms. The robot traverses uneven terrains when the left front (b) and the right front (c) calf joint is failed. The upper is the simulation with the feasible region plotted.

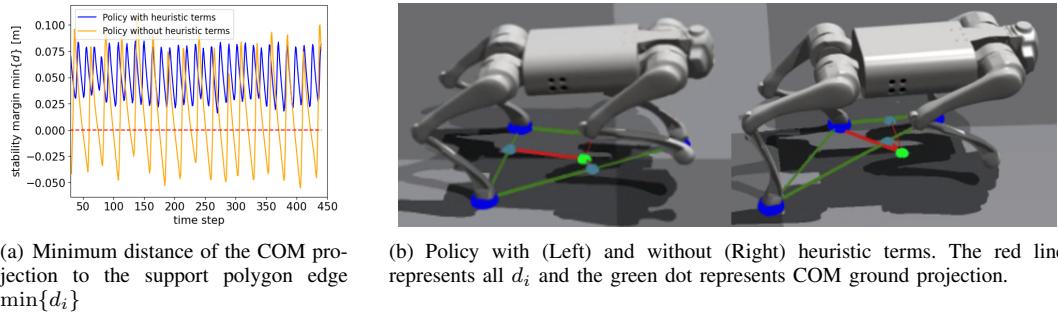


Fig. S7: Evaluation based on the stability margin.

In Fig. S6(b)(c), we present two experiments conducted on uneven terrain to validate our policy, using the feasible region criteria. In Fig. S6(b), we perform the experiment in which the robot traverses a raised platform with half of the body on the platform and the remainder still on the ground. It is observed that even though the feasible region keeps varying with the body orientation and feet contacts, the COM ground projection of the robot remains within the robot's feasible region. In Fig. S6(c), the robot successfully maintains stability while crossing a raised platform with its entire body situated on it. In contrast, even when performing an experiment on the flat plane, the policy without the heuristic terms, as shown in Fig. S6(a)(Left), can cause the COM ground projection to move outside of the feasible region, and this can easily give rise to overturns of the robot, while the COM projection is always kept inside the feasible region for the policy with heuristic terms (Fig. S6(a)(Right)).

### 3) Stability discussion on stability criteria:

To further evaluate the outcome that the above-mentioned unnecessary movement brings to the stability, we use an extra criterion, namely the minimum distance between the COM ground projection and the edges of the support polygon  $\min\{d_i\}$

( $d_i$  is the signed distance from the COM ground projection to the  $i$  th edge of the support polygon), which has been used in [3] as a stability margin to evaluate the stability of quadrupedal robots. If  $\min\{d_i\} > 0$ , it means the COM ground projection is inside the support polygon, and the greater this value the safer for the robot. If the  $\min\{d_i\} = 0$ , it means the COM ground projection is on the boundary of the support polygon and at risk of falling. If  $\min\{d_i\} < 0$ , it means the COM ground projection falls out of the support polygon.

Based on this criterion, the policy with the heuristic terms is safer than the policy without the heuristic terms, since  $\min\{d_i\}$  of the former is negative from time to time, which means it falls outside support polygon, as shown in Fig. S7(a) and Fig. S7(b).

### B. Comparison with explicit reward terms

To demonstrate the efficacy of our formulated heuristic guiding terms on improving the stability in the fault-tolerant control task, a comparative analysis is conducted between some existing methods and ours. The following methods are compared, which explicitly guide the robot to perform recovery tasks or maintain fault tolerance:

- 1) **Baseline** [7]: The policy is trained without any of the fault-tolerant guiding mechanisms.
- 2) **GCPO** [8]: Few rewards are hierarchically introduced representing soft to hard constraints to ensure stability. These rewards penalize the ZMP-COM distance in situations when the ZMP falls out of the support polygon. It also constrains foot contact to not fall out of the eligible area(constraint of the foot position).
- 3) **ACDR** [9]: A training is conducted that explicitly performs fault-tolerant tasks, with penalties imposed on actions that lead to the robot's falling. We implement the falling critical condition by judging whether the COP-based inverted pendulum's angle is greater than a threshold.
- 4) **FT-Net**: Our proposed method.

## IV. COMPLEMENTARY EXPERIMENTAL RESULTS

- 1) The experiment regarding fault-tolerant locomotion using Vicon velocity tracking, as shown in Fig. S8

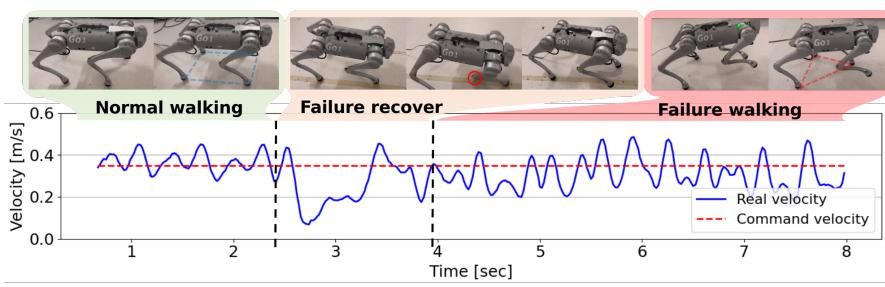


Fig. S8: Velocity tracking performance before and after the failure is triggered on the front right (FR) leg (**lower**). FR leg takes a longer duration than the RL leg to recover from the failure state and continue the tracking.

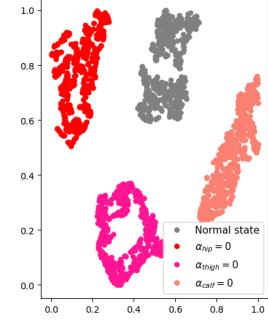


Fig. S9: t-SNE for encoded  $z_t$  vector with different representations of failures on the front left (FL) leg.

- 2) The classification of the failure scenarios: It is also observed that though the adaptor classifies the thigh and hip failure well (See Fig. S9 FL leg's example), they only show a minor adjustment in their gaits after the failure occurs (See linked video). Evident recovery behavior is observed solely in calf failure cases. For these reasons, we take the calf failure for demonstration in the main article experiments as they are more typical and representative.
- 3) The experiment regarding the computational time: The deployment of the controller is a multi-thread program, in which components are executed with different ROS frequencies. Notably, the most time-consuming operations are associated with the forward passes of two networks, namely the policy  $\pi$  and the adaptor  $E_\varphi$ . The time costs related to observation computation, hardware state reading, and sending position commands to the Unitree actuators are minimal. The forward time of these two networks is presented in Fig. S10 with the forward time cost of MLP  $\pi$  and CNN  $E_\varphi$  concentrate within  $[4ms, 9ms]$  and  $[2ms, 3.5ms]$ , respectively. Therefore the ROS frequencies of these two threads are set as the maximum forward time being  $100\text{ Hz}$  and  $250\text{ Hz}$ . The computation time cost of FT-Net is smaller than the existing RL-based [10] and model-based [11] controllers on a quadrupedal robot. Their frameworks do not even include the fault detection module.
- 4) The experiment regarding stand-still recovery: One of the robot's legs undergoes failure when it is in standing mode, as shown in the attached videos. When the failure occurs in one of the robot's legs, the robot topples over, with the diagonal leg of the failed leg lifting up. However, within approximately 1 second of the failure, the robot successfully regains stability by adjusting its center of mass towards the rear half of its body. This adjustment allows the robot to restore its body height, and reconfigure its body position and orientation as shown in Fig. S11. The foot contact forces are also automatically re-distributed from the four-legged mode to the three-legged mode.

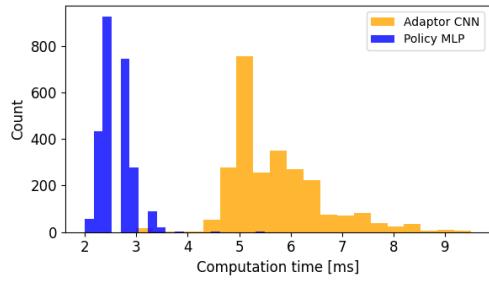


Fig. S10: Computation time of the two major networks that are deployed on the real robot.

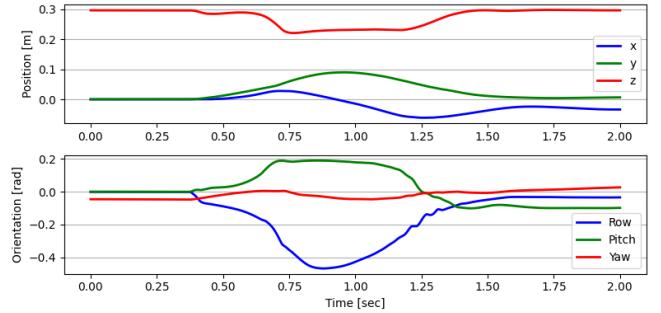


Fig. S11: Time history of the position and orientation for the failure recovery process.

## REFERENCES

- [1] R. Orsolino, M. Focchi, S. Caron, G. Raiola, V. Barasol, D. G. Caldwell, and C. Semini, “Feasible region: An actuation-aware extension of the support region,” *IEEE Transactions on Robotics*, vol. 36, no. 4, pp. 1239–1255, 2020.
- [2] M. Vukobratović and B. Borovac, “Zero-moment point—thirty five years of its life,” *International journal of humanoid robotics*, vol. 1, no. 01, pp. 157–173, 2004.
- [3] S. Kajita and B. Espiau, “Legged robot,” in *Springer handbook of robotics*. Springer Berlin/Heidelberg, Germany, 2008, pp. 361–389.
- [4] G. Bledt, M. J. Powell, B. Katz, J. Di Carlo, P. M. Wensing, and S. Kim, “Mit cheetah 3: Design and control of a robust, dynamic quadruped robot,” in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 2245–2252.
- [5] F. Farshidian, M. Neunert, A. W. Winkler, G. Rey, and J. Buchli, “An efficient optimal planning and control framework for quadrupedal locomotion,” in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 93–100.
- [6] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds. Curran Associates, Inc., 2017, pp. 4765–4774. [Online]. Available: <http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf>
- [7] N. Rudin, D. Hoeller, P. Reist, and M. Hutter, “Learning to walk in minutes using massively parallel deep reinforcement learning,” in *Conference on Robot Learning*. PMLR, 2022, pp. 91–100.
- [8] S. Gangapurwala, A. Mitchell, and I. Havoutis, “Guided constrained policy optimization for dynamic quadrupedal robot locomotion,” *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3642–3649, 2020.
- [9] W. Okamoto, H. Kera, and K. Kawamoto, “Reinforcement learning with adaptive curriculum dynamics randomization for fault-tolerant robot control,” *arXiv preprint arXiv:2111.10005*, 2021.
- [10] A. Kumar, Z. Fu, D. Pathak, and J. Malik, “Rma: Rapid motor adaptation for legged robots,” *arXiv preprint arXiv:2107.04034*, 2021.
- [11] Q. Liao, “Nmpc-wbc legged robot control stack and framework,” <https://github.com/qiayuanliao/leggedcontrol>, 2022.