Learning robust autonomous navigation and locomotion for **wheeled-legged robots**

Joonho Lee, Marko Bjelonic, Alexander Reske, Lorenz Wellhausen, Takahiro Miki, Marco Hutter

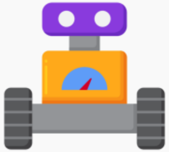
Legged Robots – Fall 2025

Till Beyer, Mattia Prandi, and Andrea Tarabay

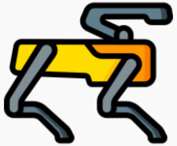
Why Wheeled- Legged Robots?

Urban last-mile delivery is hard

time pressure – obstacles – indoor/outdoor



Pure wheels are **fast** but can't handle **big obstacles**



Pure legs are **agile** but **slow and inefficient**

Wheeled-legged robots combine:

high-speed travel on moderate surfaces

\+ **agility** on challenging terrain

What are the gaps?

Gap 1: Hybrid Locomotion is hard

Gap 2: Navigation is not Dynamics-Aware

This Paper

Gap 3: Safety: Speed demands fast decisions

Gap 4: System integration is brittle

EPFL

What are the gaps?

Gap 1:
Hybrid
Locomotion
is hard

Prior Work

- ➢ Hand-tuned rules for drive/step
- ➢ Predefined gait sequences inspired by animal locomotion (trot/pace)
- ➢ Trajectory-optimization variants: compute-heavy.

Why this is limiting for wheeled-legged robots?

1. **Speed and cost of transport (COT)** depend on direction, terrain, and gait.
2. **Energy efficiency:** driving when feasible and stepping only when necessary lowers COT
3. **Gait switching** leads to lower cost of transport (COT)
4. **Prioritizing computational efficiency:** leads to suboptimal performance on the robot

What are the gaps?

Gap 1: Hybrid Locomotion is hard

Gap 2: Navigation is not Dynamics-Aware

This Paper

Gap 3: Safety: Speed demands fast decisions

Gap 4: System integration is brittle

What are the gaps?

Gap 2:
Navigation is not Dynamics-Aware

Prior Work

- ➢ overlook the characteristics of dynamic robots, leading to suboptimal plans.
- ➢ Kinematic Navigation plans by sampling-based planning on these estimated cost maps
- ➢ Cannot account for various dynamic characteristics of the robot

Consequently, they may result in frequent turning and stepping actions that can decrease efficiency.

What are the gaps?

Gap 1: Hybrid Locomotion is hard

Gap 2: Navigation is not Dynamics-Aware

This Paper

Gap 3: Safety: Speed demands fast decisions

Gap 4: System integration is brittle

What are the gaps?

Gap 1: Hybrid Locomotion is hard

Gap 2: Navigation is not Dynamics-Aware

This Paper

Gap 3: Safety: Speed demands fast decisions

Gap 4: System integration is brittle

What are the gaps?

This Paper

Builds an end-to-end autonomous system for wheeled-legged robots that *jointly* **learns hybrid locomotion** and **optimizes navigation** to the robot’s dynamics.

EXECUTIVE SUMMARY

1

Robot class & morphology:
Wheeled-legged quadruped with actuated wheels integrated on each leg.

2

Perception & Sensors:
Onboard:
- 3 LiDARs
- front stereo camera
- IMU
- joint encoders
- GPS antenna
- 5G router (networking)

3

Control Architecture:

Two-level learned stack:
Navigation
➢ **High Level Control (HLC)** → outputs base velocity targets

Locomotion
➢ **Low Level Control (LLC)** → tracks those commands and outputs joint position and wheel velocity commands.

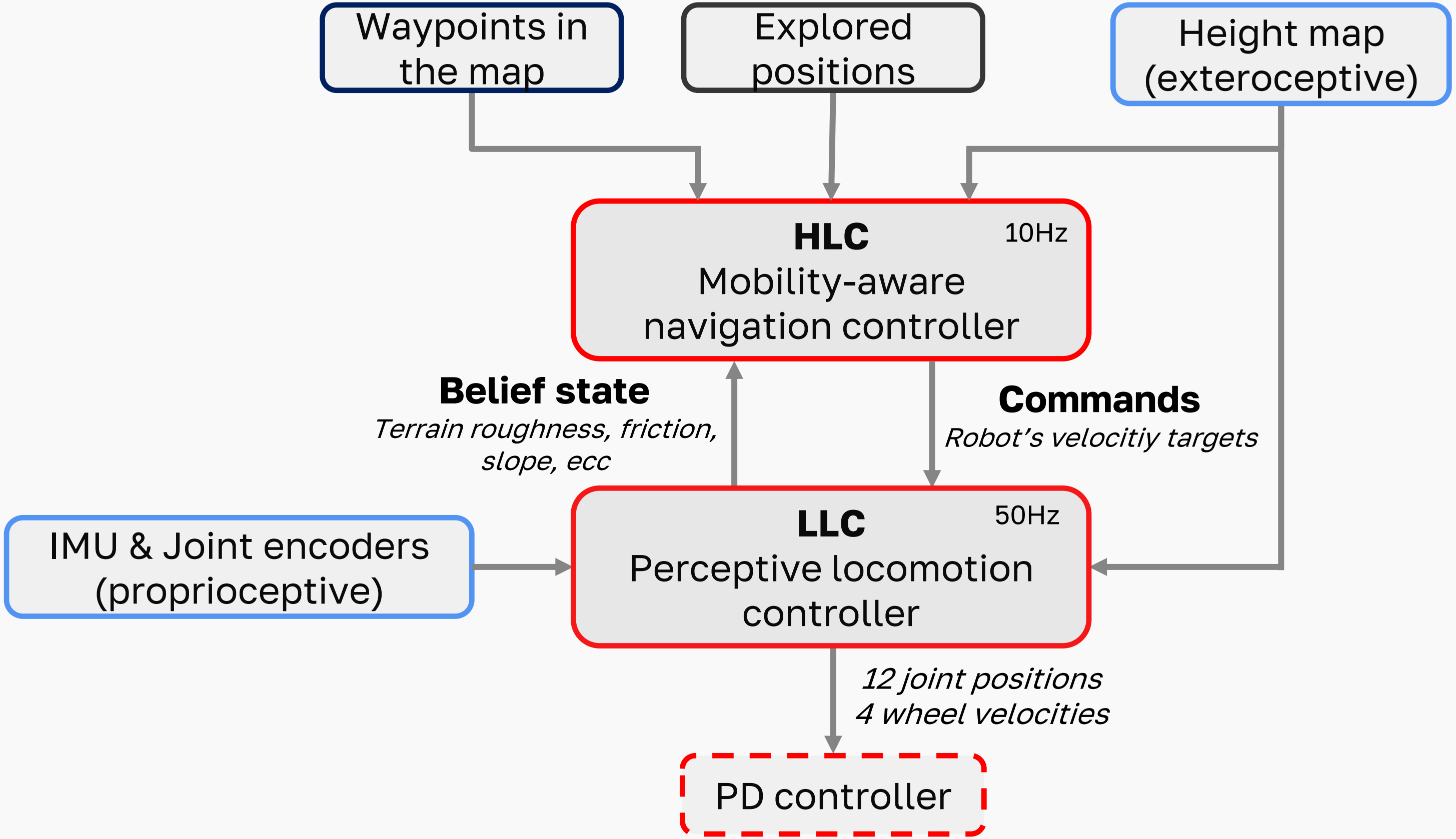
LLC action =
12 joint position commands +
4 wheel velocity commands;
joints run with PD control;
wheels in velocity control.

4

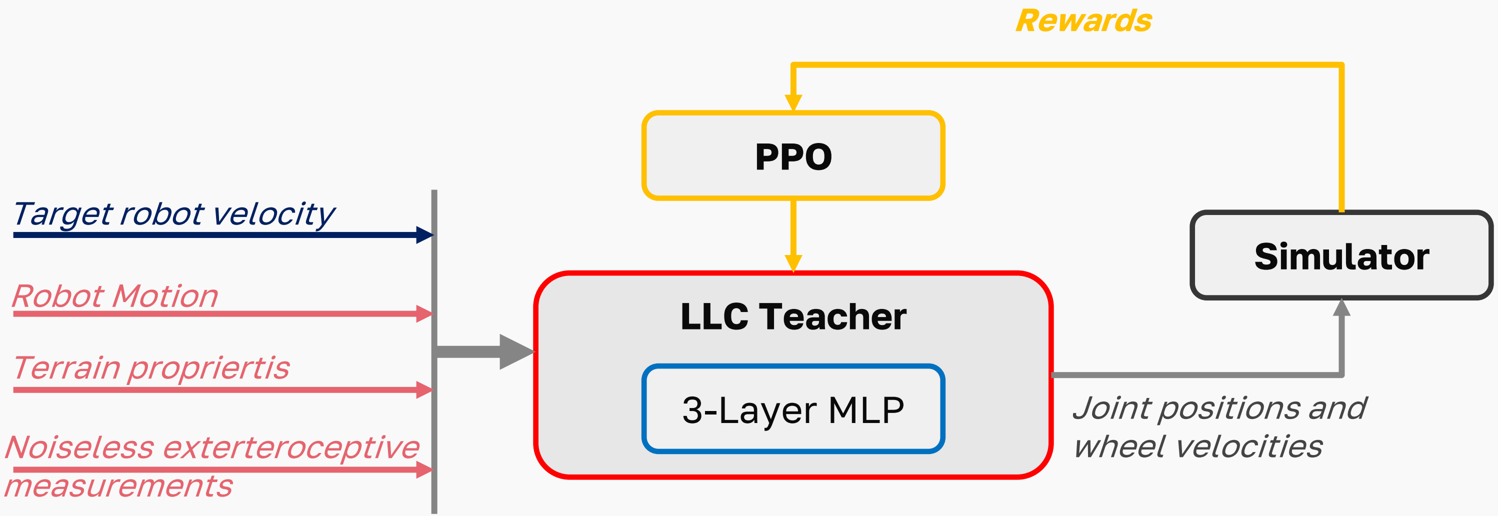
Design / learning methods
HLC – Navigation: hierarchical reinforcement learning that learns a local planner, follows the global graph, avoids obstacles, and outputs feasible base-velocity targets.
LLC – Locomotion: model-free imitation reinforcement learning that learns hybrid locomotion (drive/step, gait/posture) to track HLC commands via joint and wheel control.

Both controllers use **proximal policy optimization**.

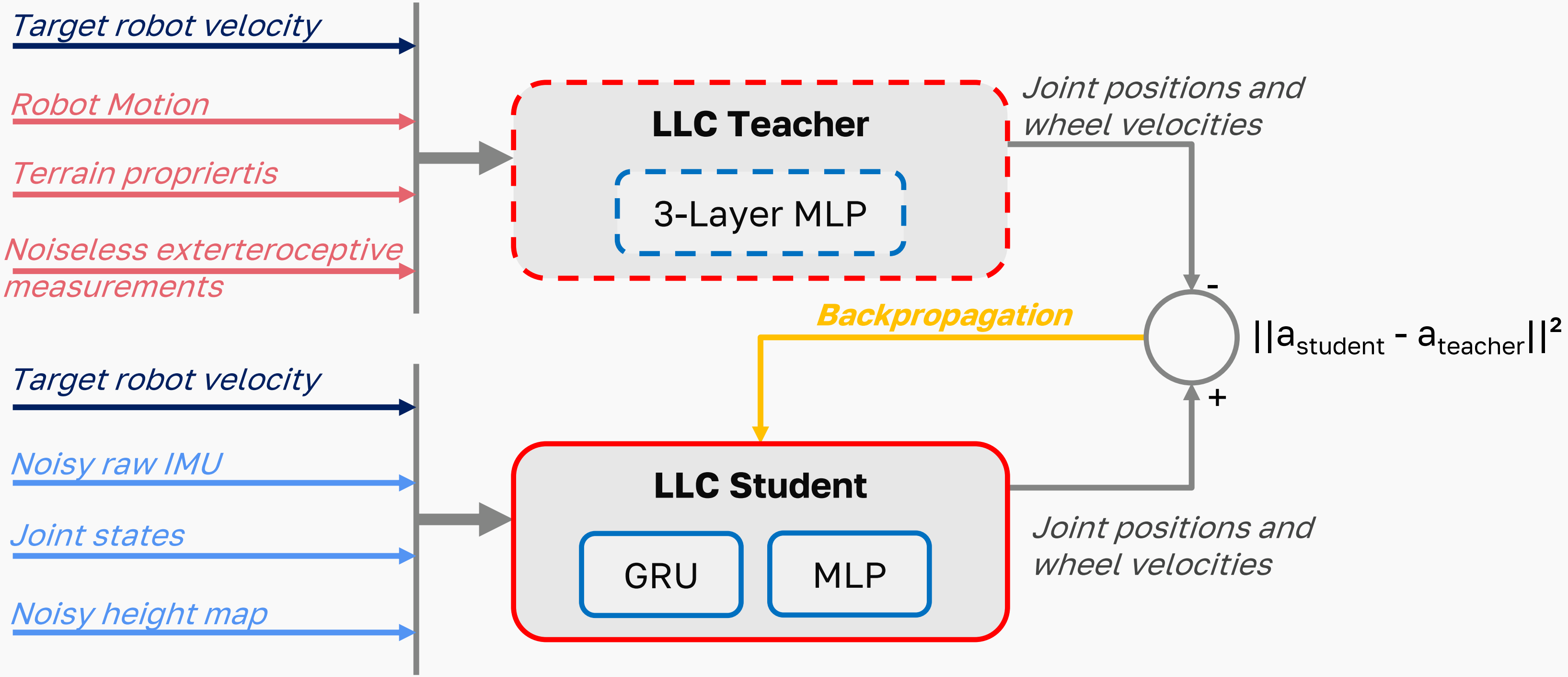
Control architecture



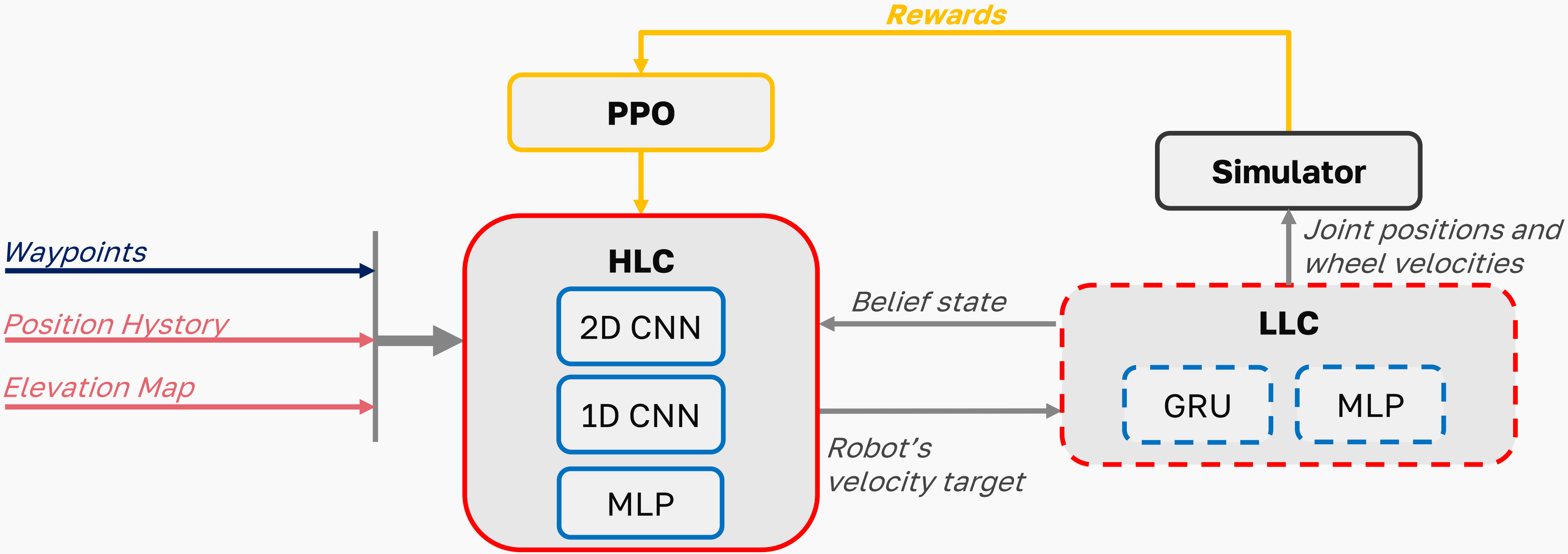
Training setup – LLC Teacher



Training setup – LLC Student



Training setup – HLC



Reward policy

High-level policy reward

$$r_h + w_l \cdot (r_l + r_r)$$

Low-level policy reward

$$r_l + r_r$$

Low-level Reward Policy

Linear velocity tracking reward

$$r_{lv} := \begin{cases} 2.0\exp(-2.0 \cdot \|v_{xy}^{body}\|^2), & \text{if } |v_{des}| < 0.05 \\ \exp(-2.0\|v_{xy}^{body} - v_{des}\|^2) + v_{des} \cdot v_{xy}^{body}, & \text{otherwise} \end{cases}$$

Angular velocity traking

$$r_{av} := \exp(-2.0(\omega_z^{body} - \omega_{des})^2)$$

Undesired body motion penalty

$$r_{bm} := -1.25(v_z^{body})^2 - 0.4|\omega_x^{body}| - 0.4|\omega_y^B|$$

Robot's body tilt penalty

$$r_{ori} = \arccos(R_b(3,3))^2$$

Robot's body height reward

$$r_h = max(0.0, |h_{base} - 0.55| - 0.05)$$

High-level Reward Policy

Goal reaching sparce reward

$$r_{h,goal} := \begin{cases} 1.0 & |p_{robot} - wp^1| < 0.75 \\ 0.0 & \text{otherwise} \end{cases}$$

Goal reaching dense reward
Only at the beginning

$$r_{h,dense} := \begin{cases} 1.0 & |e_{wp^1}| < 0.75 \\ clip(v \cdot \widehat{e_{wp^1}}, 0.0, v_{thres})/v_{thres} & \text{otherwise} \end{cases}$$

Exploration bonus

$$r_{h,exp} := \sum_{P_{buf}} C(s_t, wp_t^1, p_{buf}^i)$$

Near-goal stability reward

$$r_{h,stability} := \begin{cases} \exp\left(-2.0||v||^2\right) & |p_{robot} - wp^1| < 0.75 \\ 0.0 & \text{otherwise} \end{cases}$$

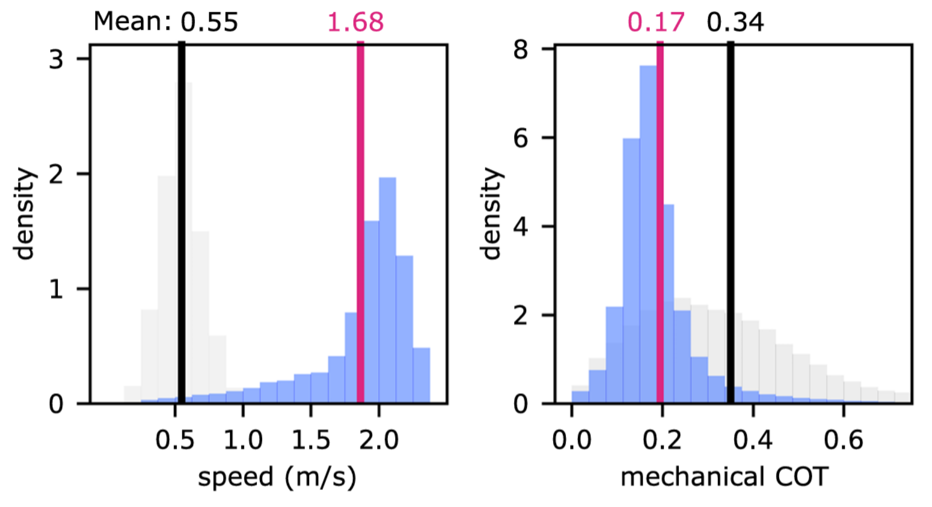
Regularization Reward policy

Joint torques penalty	$r_\tau := -\sum_{i \in joints} \|\|\tau_i\|\|^2$
High joint velocity and accelleration penalty	$r_s = -c_k \sum_{i=1}^{12} (\dot{q}_i{}^2 + 0.01\ddot{q}_i{}^2)$
Non-smooth trajectories penalty	$r_s = -c_k \sum_{i=1}^{12} ((q_{i,t,des} - q_{i,t-1,des})^2 + (q_{i,t,des} - 2q_{i,t-1,des} + q_{i,t-2,des})^2)$
Knee joint flipping in opposite direction penalty	$r_{jc} = \sum_{i=1}^{12} \begin{cases} -(q_i - q_{i,th})^2, & \text{if } q_i > q_{i,th} \\ 0.0 & \text{otherwise} \end{cases}$
Contact with environment penalty	$r_{bc} := -\|I_{c,body} \setminus I_{c,wheel}\|$
Survival reward	$r_{h,surv} := 1.0 \quad \text{while not terminated}$

Results: Autonomous deployment

Test: Follow 13 distinct goal points scattered across *Glattpark, Zurich*

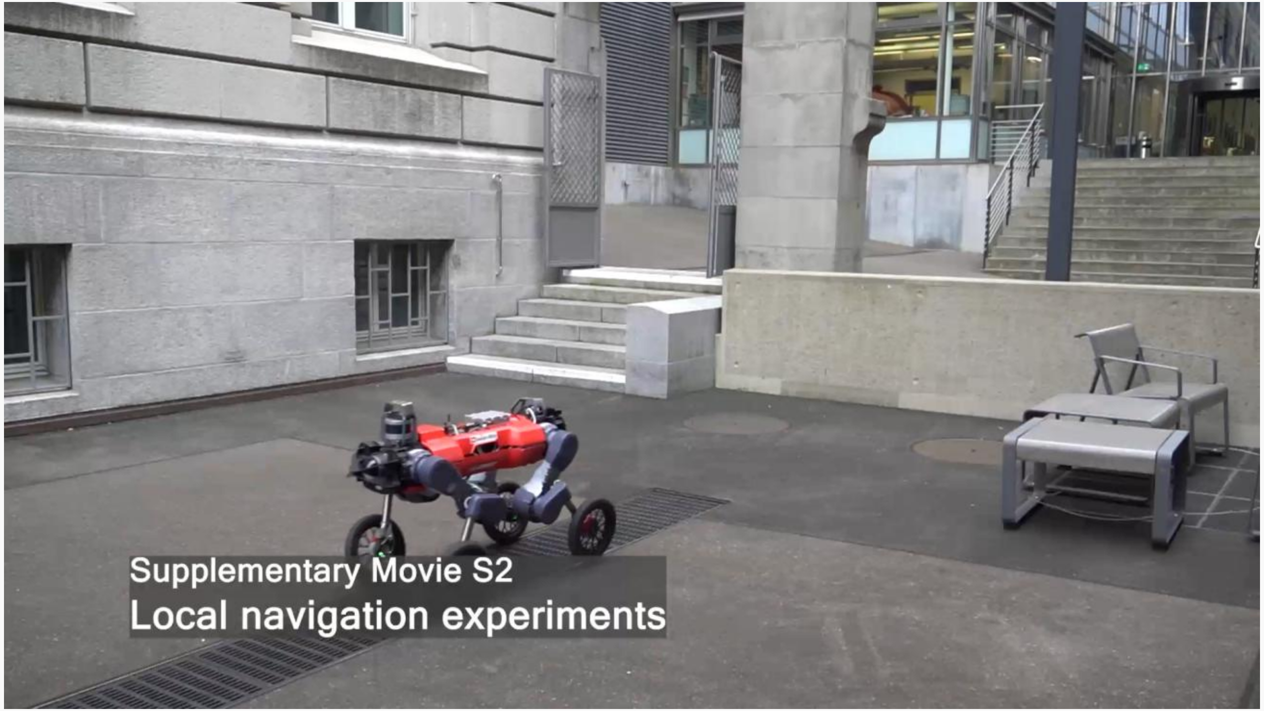
Mission: Visit location on map autonomously → Simulate *delivery*



Speed ↑, Cost of Transport ↓
→ **Higher Efficiency**

$$COT = \sum_{joints} \frac{\tau\dot{\theta}}{mg \cdot |v_{xy}|}$$

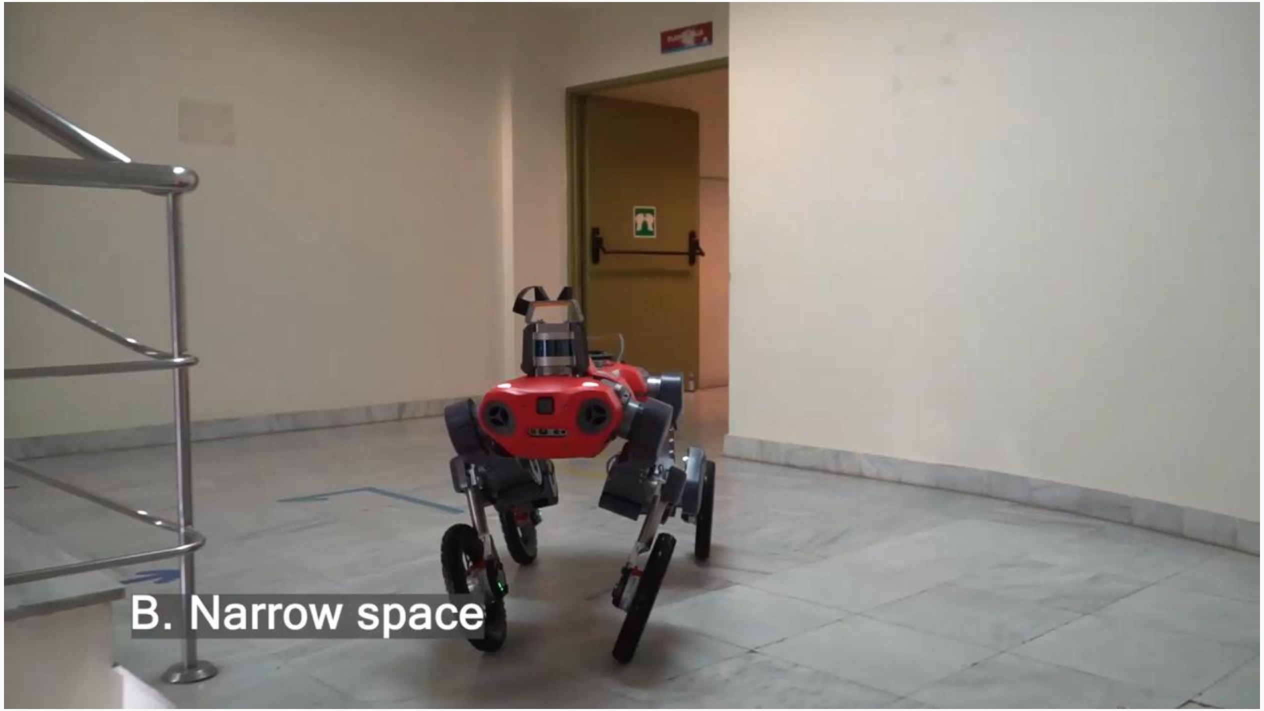
Results: Local Navigation



Manage blocked paths
→ Explore options thanks to **Position Memory**

Results: Local Navigation

Navigate tight spaces
→ precise, **real-time trajectory** adjustments

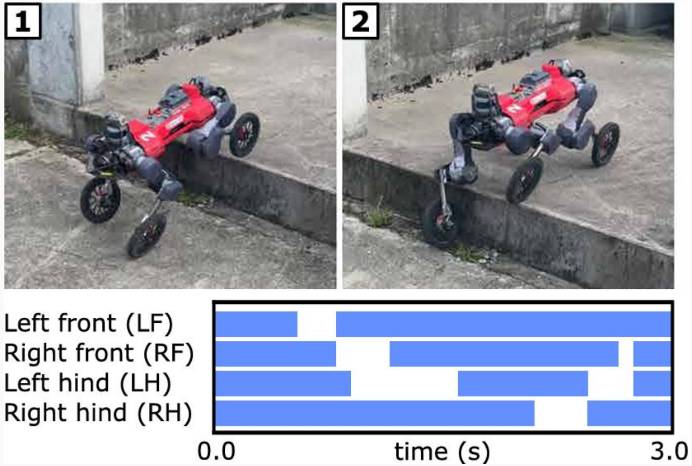


Results: Local Navigation

Human Safety
→ **No-Go** bubble

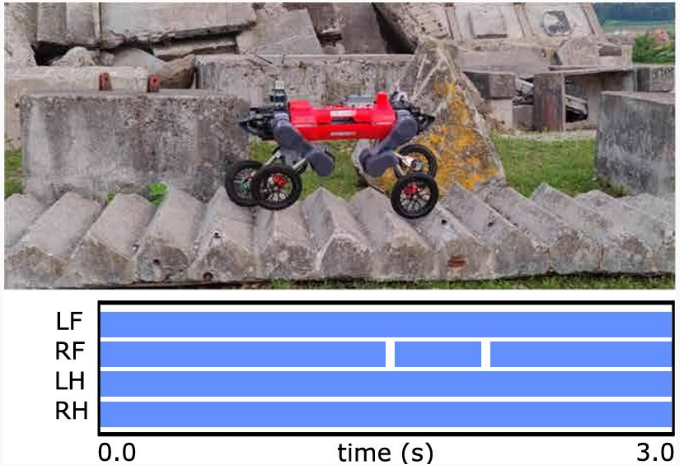
Results: Hybrid Locomotion

Large Step



- Combines **Creeping** and **Driving**
- Max. height direction-dependent

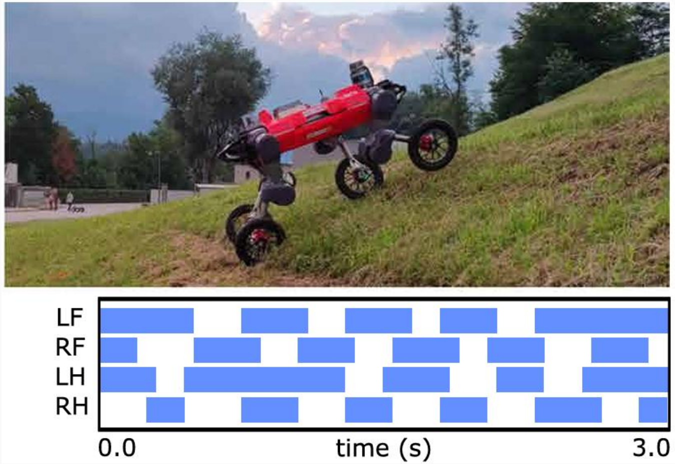
Bumpy Terrain



Adjust Legs to stabilize body
→ Active **Suspension**

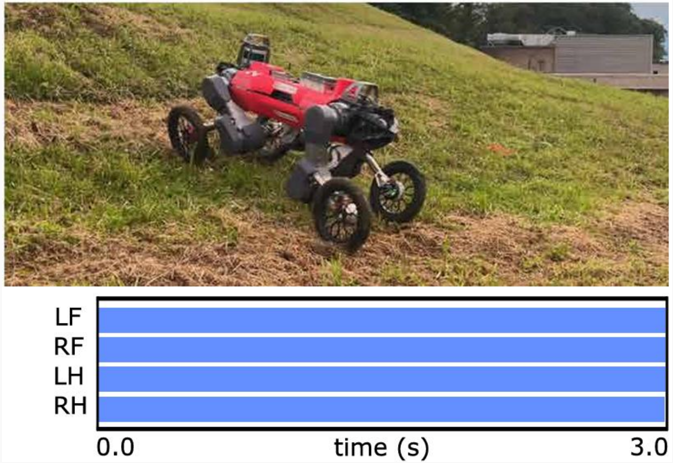
Results: Hybrid Locomotion

Uphill



Trotting gait
(if sufficiently fast and steep)

Downhill



Driving mode
→ **Crouching** improves stability

Wide **Exploration**

Energy Efficiency

Responsive → **Safety**

Fast

Limited **FoV**

Only **Geometric Map**

Offline **Point Cloud**

EPFL

Citations

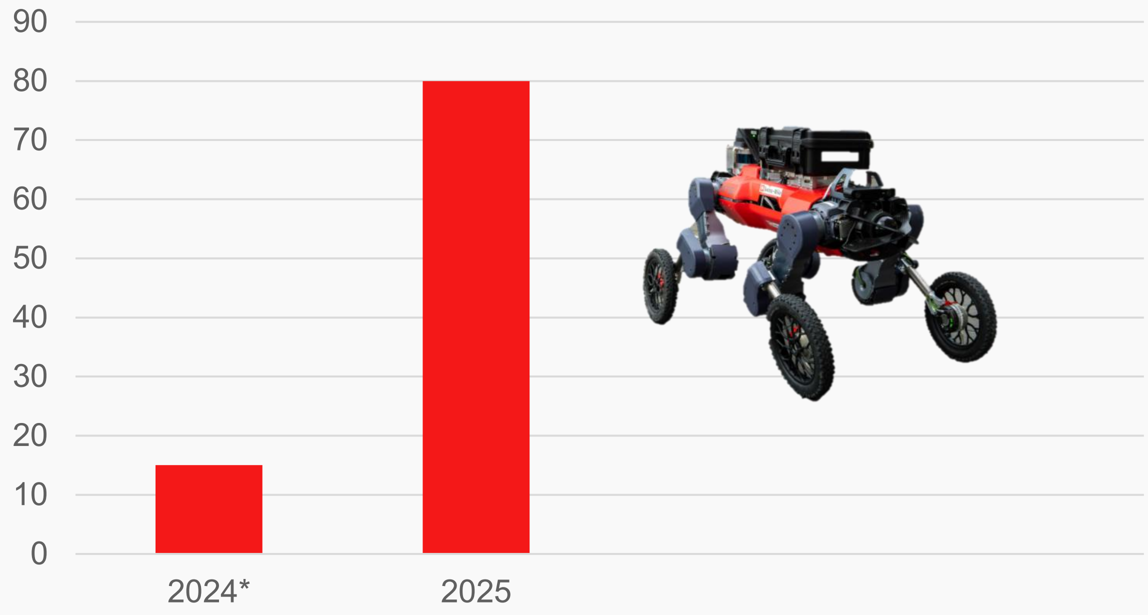
Wheel-Legged Locomotion

- Francesco Iotti, Alok Ranjan, Franco Angelini, Manolo Garabini, *OmniQuad: A wheeled-legged hybrid robot with omnidirectional wheels,* Mechanism and Machine Theory, Volume 214 (2025)
- Wei, Z., Ren, J., Guo, J. *et al.* SlidBot: *A Quadruped Robot with Passive Wheels for Roller Skating*. J Bionic Eng (2025)
- Guangrong Chen, Qingyu Meng, Yuxiang Lin, *Cooperative skating motion control for a quad-wheel-legged robot*, Results in Engineering, Volume 27 (2025)

Navigation, Localization, Mapping

- Cristóbal, E.D., Rivas, E.d.J.O., Treviño, L.M.T. (2026). *Artificial Vision-Guided Hexapod Robot for Autonomous Navigation and Object Localization.* In: Martínez-Villaseñor, L., Vázquez, R.A., Ochoa-Ruiz, G. (eds) Advances in Soft Computing. MICAI 2025
- D. Aditya et al., *Robust Localization, Mapping, and Navigation for Quadruped Robots*, 2025 European Conference on Mobile Robots (ECMR), Padova, Italy (2025)

Citations per year



** Publication year*


Thank you
for listening