

Extreme Parkour with Legged Robots

Group 18

Julie CARRETTI

Rose KAPPS

Geoffroy RENAUT

Extreme Parkour



Popular athletic sport that involves humans traversing obstacles in a highly dynamic manner

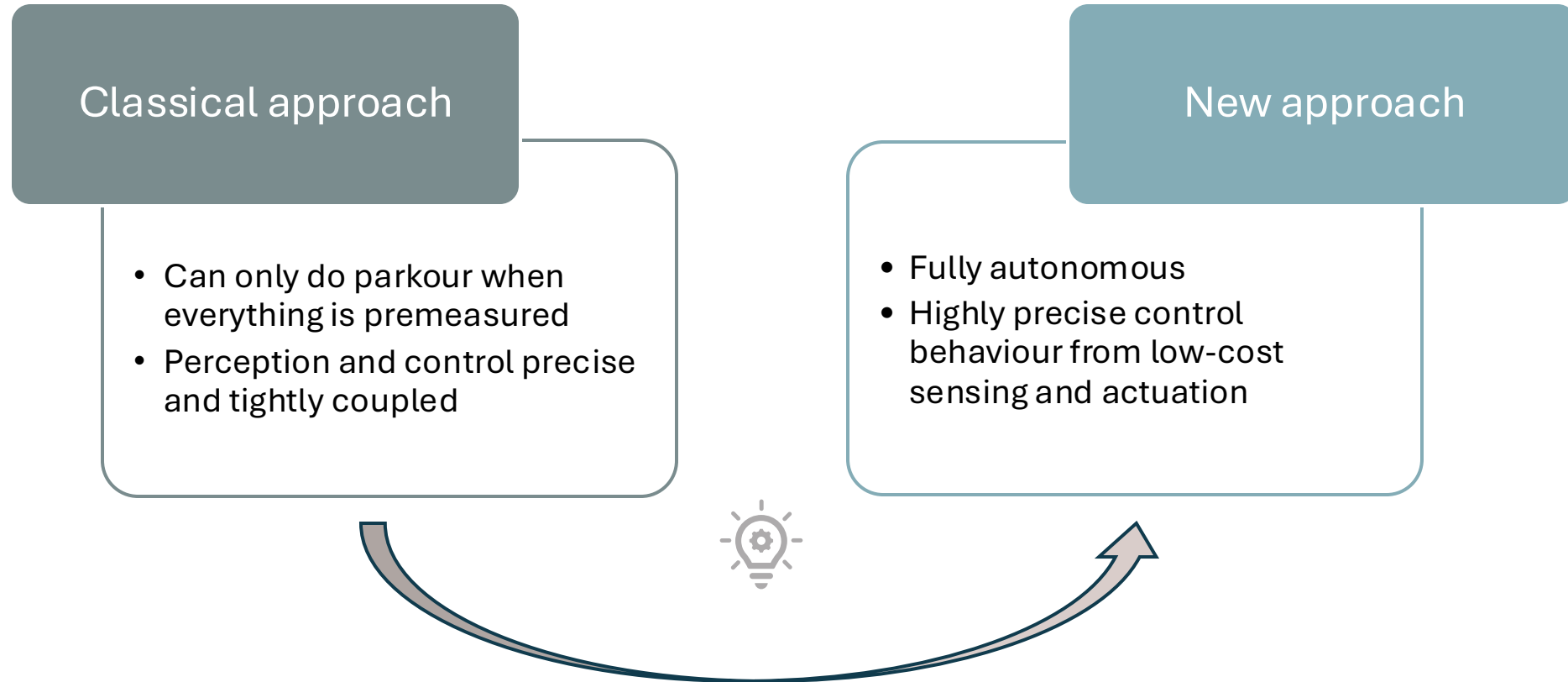


In robotics → hardware as well as software challenge

How a low-cost robot with imprecise actuation can perform precise athletic behaviors in an autonomous manner?



Main idea and contribution of the article



What if any obstacle were to move? What if the robot is asked to perform a new parkour course?

Summary of key aspects



Type of robot

Quadruped low-cost robot with imprecise actuation (Unitree A1)



Type of control

End-to-end control based on motor commands generated through a neural network



Design method

Single neural network policy trained in simulation with large-scale RL



Gaits

Walk, run, long jumps, high jumps, handstands, and ramp navigation



Sensors

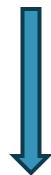
Single Front-facing Intel RealSense D435 depth camera

Challenges

Make highly precise control behaviour from low-cost sensing and actuation

Freedom to automatically adjust heading direction

Combining each parkour behaviours which are very different in nature withing a single neural network



2 KEY TECHNICAL INNOVATIONS

A novel dual distillation method

To infer precise motor commands and adjust heading based on obstacle detection in real time

A unified reward design

To simplify reward functions for different behaviours, enabling the emergence of behaviours like high jumps and handstands through reinforcement learning.

Detailed Presentation : Rewards

$$r_{tracking} = \min(\langle \mathbf{v}, \hat{\mathbf{d}}_w \rangle, v_{cmd}) \quad \hat{\mathbf{d}}_w = \frac{\mathbf{p} - \mathbf{x}}{\|\mathbf{p} - \mathbf{x}\|}$$

$$r_{clearance} = - \sum_{i=0}^4 c_i \cdot M[p_i]$$

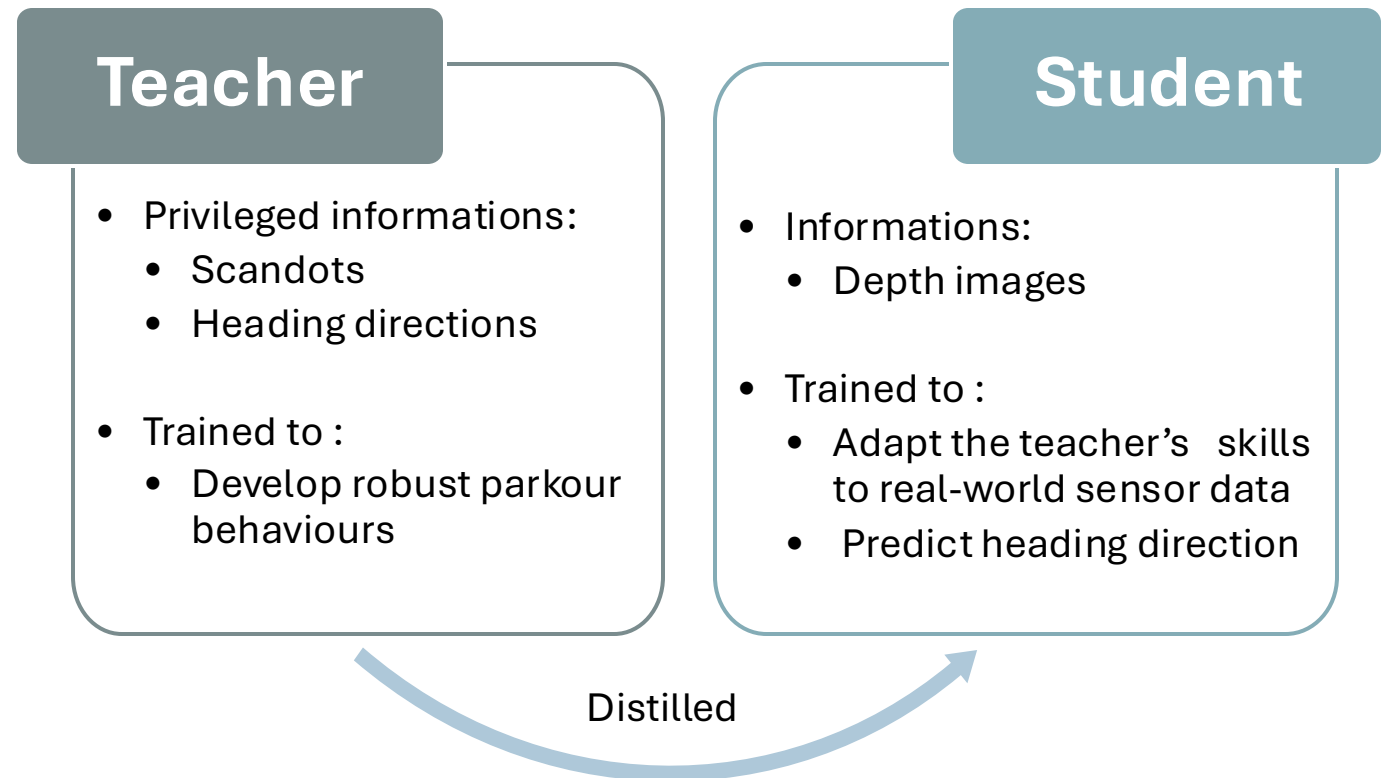
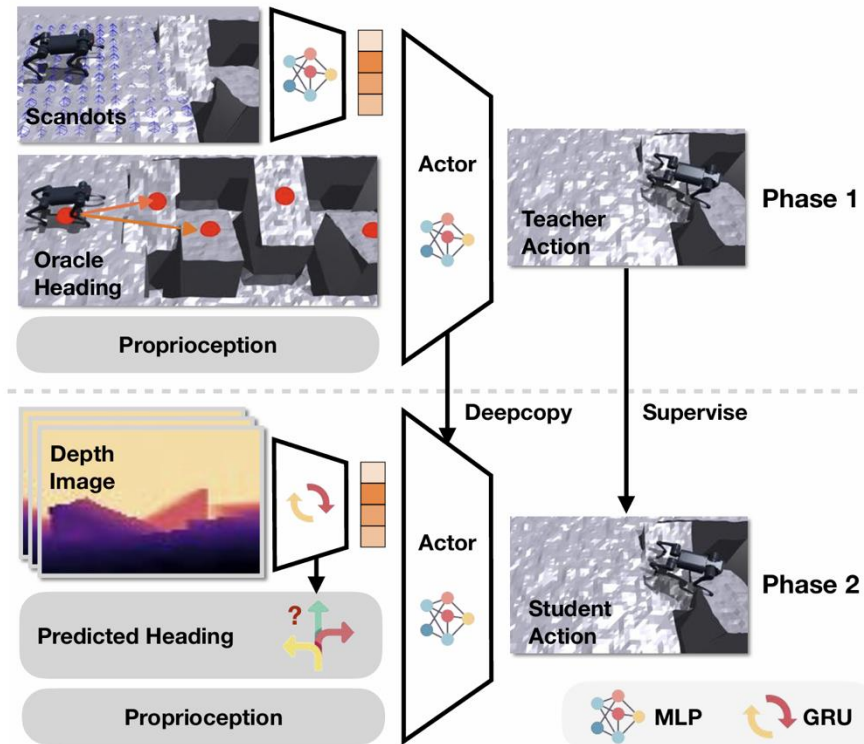
$$r_{stylized} = W \cdot [0.5 \cdot \langle \hat{\mathbf{v}}_{fwd}, \hat{\mathbf{c}} \rangle + 0.5]^2$$

- Gives a higher reward if the speed is in the same direction as the waypoint
- Cannot exceed the desired speed.

- Gives a punition for each robot leg « too close » from an edge.

- Gives a reward for being close to a desired body orientation.
- Allows the robot to handstand.

Dual distillation method



Dual distillation – Phase 1 : Teacher

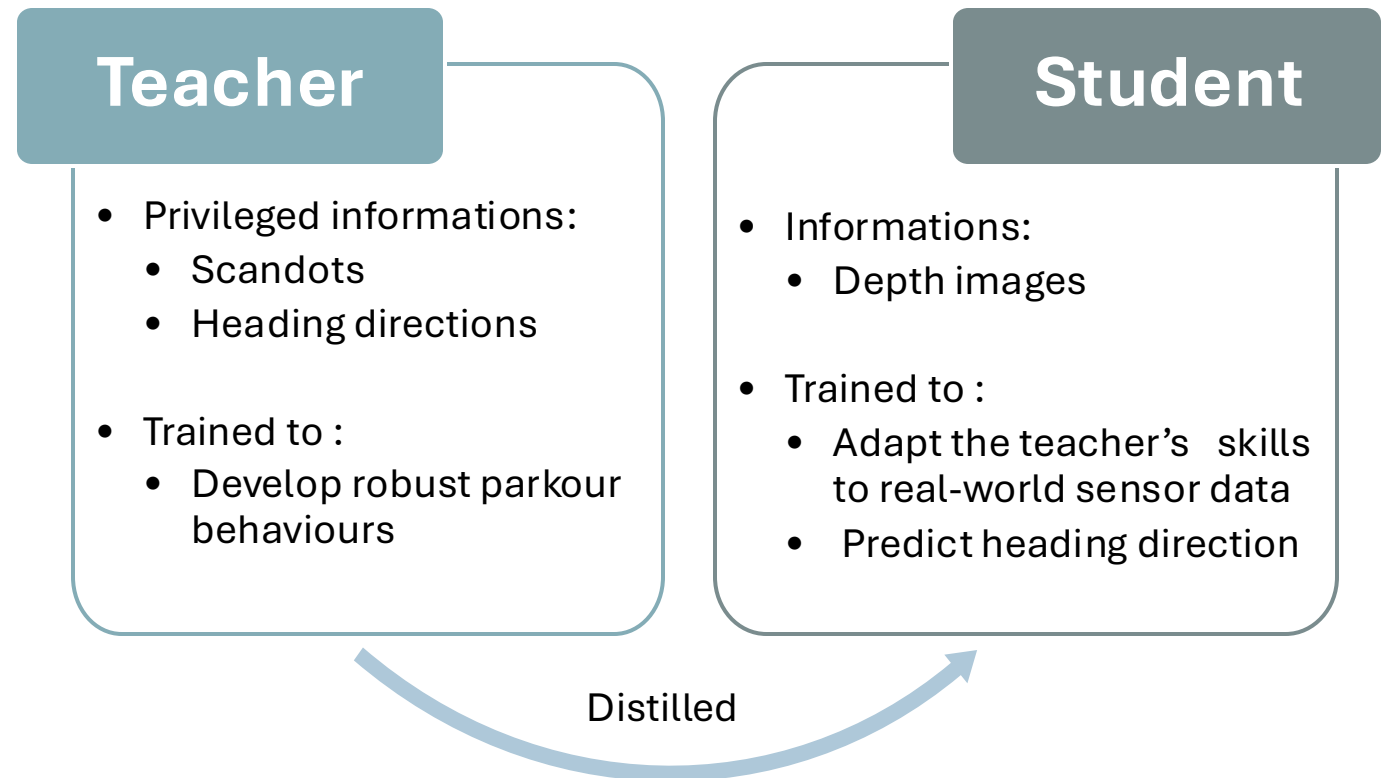
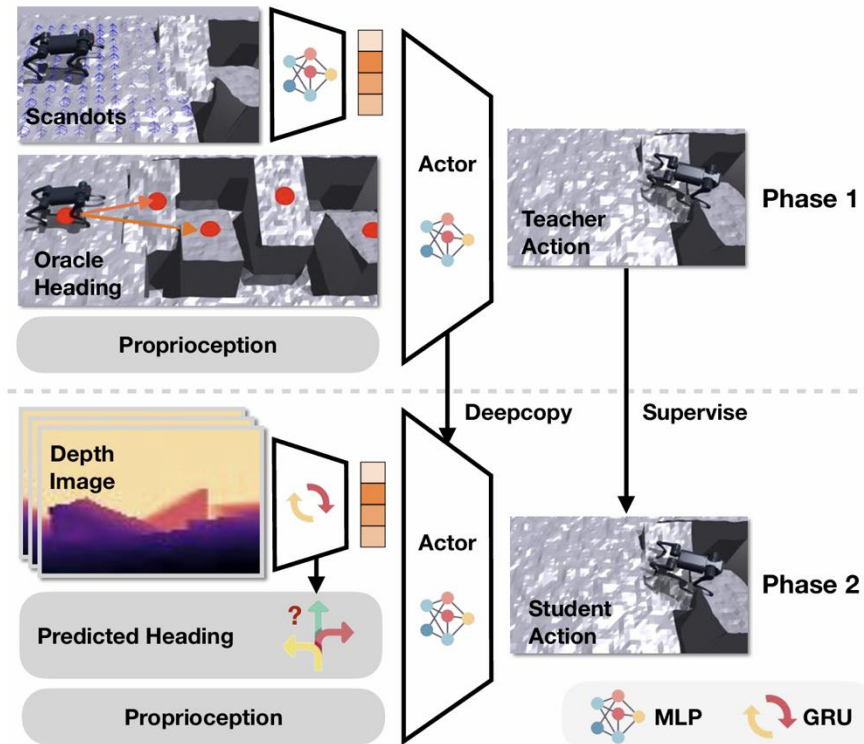
The policy is trained with privileged information : Scandots and heading directions

Use of ROA's (Regularized Online Adaptation) technique : the robot learn to adapt his behaviour in real time to handle changes in its environment.

The robot is trained with a set of increasingly difficult terrains :

- Robot successfully traverses more than half of the obstacle → promoted to a harder terrain
- Robot fails to traverse more than half of the obstacle → demoted to an easier terrain

Dual distillation method



Dual distillation – Phase 2 : Student

The policy has only access to depth images.

To help the robot learn to estimate its heading direction, they used a Mixture of Teacher and Student (MTS) approach:

$$obs_{\theta} = \begin{cases} \theta_{pred}, & \text{if } |\theta_{pred} - \hat{\mathbf{d}}_w| < 0.6 \\ \hat{\mathbf{d}}_w, & \text{otherwise} \end{cases}$$

obs_{θ} = yaw angle the policy observes

θ_{pred} = yaw angle from the prediction

$\hat{\mathbf{d}}_w$ = yaw angle from the oracle

The policy keeps the robot's prediction if it is reasonable compared to the teacher's, otherwise, it retains the teacher's prediction.

Results



High jump obstacles :
2x robot's height



Long jump obstacles :
2x robot's length



Handstand gait

Who Cites the Article?

Eight citations in other papers, for example :

« However, solely explicitly estimating terrain using laggy and **noisy depth cameras is not sufficiently reliable, especially near edges** where the risk of stepping off is high »

S. Luo, S. Li, R. Yu, Z. Wang, J. Wu and Q. Zhu, "PIE: Parkour With Implicit-Explicit Learning Framework for Legged Robots," in *IEEE Robotics and Automation Letters*

« In contrast to most systems used for handshaking, **quadruped robots are showing impressive abilities to traverse challenging environments [22], [23], run at high speeds [24], jump in difficult terrains [25], and locomote over dynamic parkour obstacles [26]–[28].** »

A. Chappuis, G. Bellegarda and A. Ijspeert, "Learning Human-Robot Handshaking Preferences for Quadruped Robots," 2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN)

« Regardless, reinforcement learning controllers are frequently used to generate forceful interactions such as during walking, running, parkour, fall recovery, object re-orientation, and others [13]–[18]. »

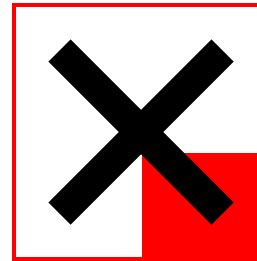
T. Portela, G. B. Margolis, Y. Ji and P. Agrawal, "Learning Force Control for Legged Manipulation," 2024 IEEE International Conference on Robotics and Automation (ICRA)

Analysis of Pros and Cons



PROS

- The robot achieve quite impressive results compared to its price
- The deployable policy can be trained in less than 20 hours.



CONS

- The noisy depth camera is not reliable enough to walk near an edge

Possible Exam Questions

- **What is the purpose of the teacher-student approach in training the robot's neural network. How does this method help the robot adapt to real-world conditions? (slides 8-11)**

→ The teacher-student approach transfers knowledge from a teacher model, trained with privileged information in simulation, to a student model, which only uses real-world sensor data like depth images. This helps the robot adapt to real-world conditions by learning to perform optimally without relying on simulation-only data.

- **How does the reward function promote the emergence of new behaviours ? (slide 7)**

→ The reward function includes components for efficient movement, edge avoidance, and specific postures. This helps the robot adapt to real-world conditions by teaching it to perform well without relying on simulation-only data.