EPFL

# Legged Locomotion in Challenging Terrains using Egocentric Vision , Agarwal et al. , 2022


Valmir Syla

Guillaume Spahr

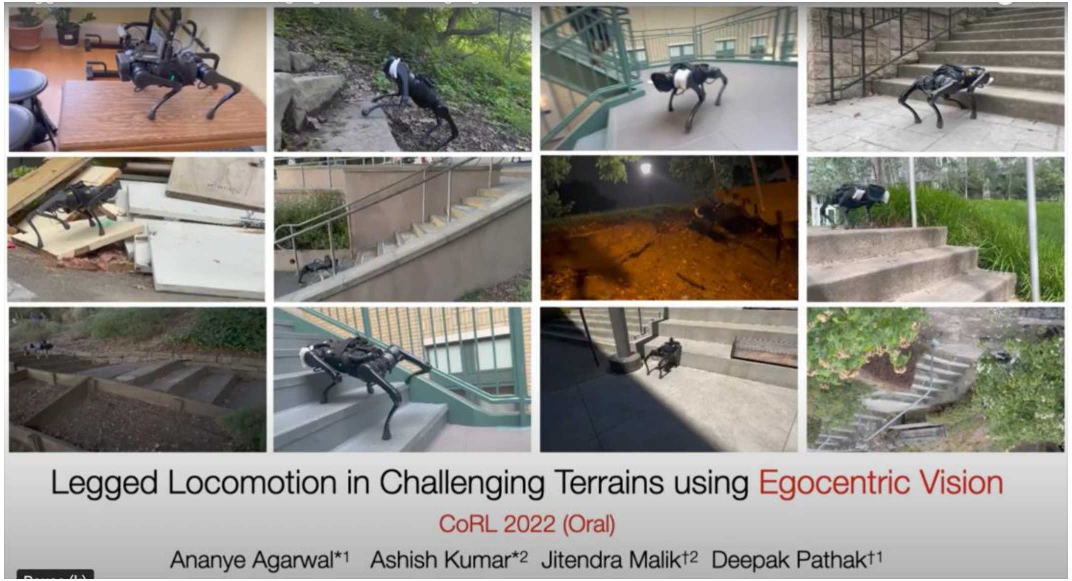
Mischa Mez


November 2024


École polytechnique fédérale de Lausanne

# Introduction

- The article introduces **an end-to-end quadrupedal locomotion system** that uses a **single front-facing depth camera** for visual navigation.
- The system is designed to enable a small quadruped robot to **traverse challenging terrains** like stairs, curbs, stepping stones, and gaps **without relying on elevation maps or pre-programmed gait patterns**.
- Using **a combination of reinforcement learning and supervised learning**, the robot can develop adaptive gait patterns and **remember past visual information to navigate complex terrain.**



# Robot article key points

- **Robot Type:** Medium-size quadruped with 12 actuated joints
- **Control Type:** Joint angle control with position commands sent by a policy network, converted to torques using a PD controller.
- **Learning Methodology**: Two-Phase Training

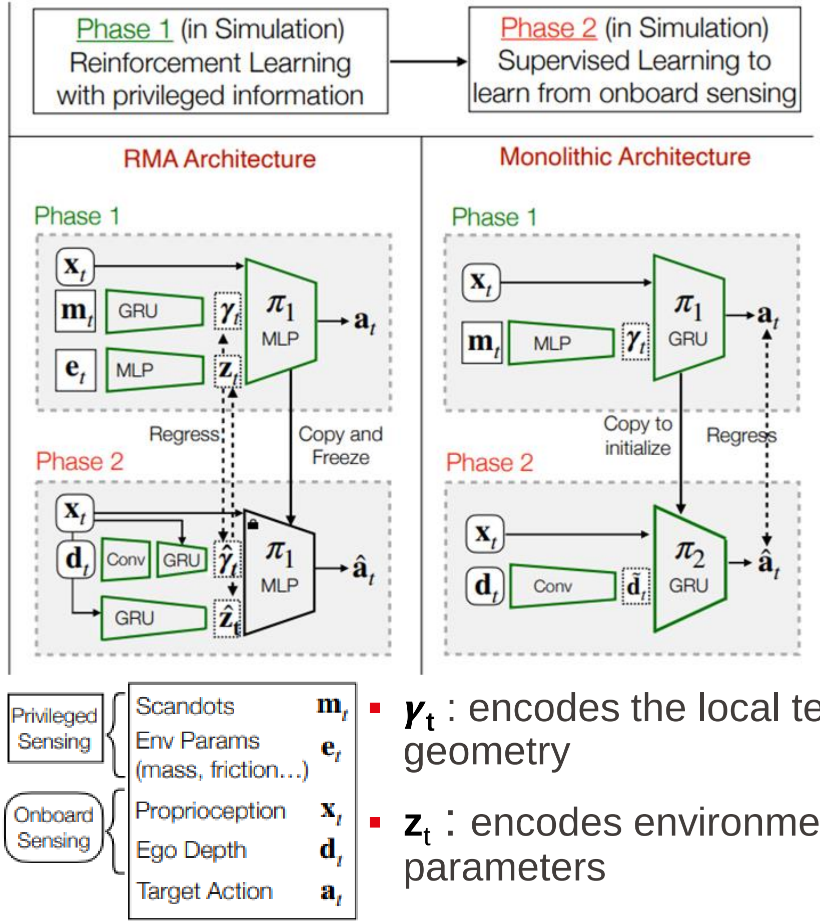
  1. **Reinforcement Learning (RL)** in simulation using simplified depth representation (“scandots”) to learn locomotion without predefined gaits

  2. **Supervised Learning** with real depth data from the onboard camera to refine joint angles for deployment
- **Type of Gait:** Emergent, adaptive gait suitable for walking and obstacle traversal
- **Sensors**: Single front-facing Intel RealSense depth camera and proprioceptive feedback, with the depth data used to remember recent terrain for foot placement
- **Environments:** Indoor and outdoor terrains, including stairs, ramps, rocky trails, and obstacles like stepping stones

- **Performance:** The robot achieves high success rates in various challenging terrains, however there are some instances where the reality gab still occur

# Models

- Two models based on **gated recurrent unit (GRU)** used for **short term memory**
- **Monolithic Architecture:**
  - Simpler
  - Implicit track of $\boldsymbol{\gamma}_t$ and $\mathbf{z}_t$ in the weights → hard to disentangle
  - GRU as the controller
- **RMA (Rapid Motor Adaptation) Architecture**
  - Direct access to $\boldsymbol{\gamma}_t$ and $z_t$
  
  → possibility of swapping sensors
  - MLP as the controller
  - No need to retrain the controller




- $\boldsymbol{\gamma}_t$ : encodes the local terrain geometry
- $\mathbf{z}_t$ : encodes environment parameters

# Training phases

- **During phase 1, use of PPO algorithm with specific rewards:**
  - Absolute work penalty : minimize energy consumption
  - Command tracking penalty
  - Behavioral and environnemental penalties
  - Survival Bonus
- **During phase 2, supervised learning to distil the phase 1 policy :**
  - **Monolithic Architecture:** Minimize RMS of the difference between predicted and ground thruth actions
  - **RMA Architecture:** Minimize RMS of the difference between predicted and ground thruth $\boldsymbol{\gamma}_t$ and $\mathbf{z}_t$
- Trained on **IsaacGym (IG)** simulator on large terrain maps with 100 sub-terrains arranged in a 20x10 grid.
- Different sets of terrain of **varying difficulty** level.
- Two other models trained to be used as **baselines**:
  - **Blind** : visual observations removed to only rely on proprioception informations
  - **Noisy** : trained with noiseless elevation maps and distilled with large noise added to the elevation map

# Results

- **Results of the 2 models and the baselines (Noisy and Blind)**

| Terrain | Average $x$-Displacement (↑) | | | | Mean Time to Fall (s) | | | |
|---|---|---|---|---|---|---|---|---|
| | RMA | MLith | Noisy | Blind | RMA | MLith | Noisy | Blind |
| Slopes | 43.98 | 44.09 | 36.14 | 34.72 | 88.99 | 85.68 | 70.25 | 67.07 |
| Stepping Stones | 18.83 | 20.72 | 1.09 | 1.02 | 34.3 | 41.32 | 2.51 | 2.49 |
| Stairs | 31.24 | 42.4 | 6.74 | 16.64 | 69.99 | 90.48 | 15.77 | 39.17 |
| Discrete Obstacles | 40.13 | 28.64 | 29.08 | 32.41 | 85.17 | 57.53 | 59.3 | 66.33 |
| Total | 134.18 | 135.85 | 73.05 | 84.79 | 278.45 | 275.01 | 147.83 | 175.06 |

- **Results in specific scenarios:** We observe a clear improvement between the blind and the proposed model for each case except for the "downstairs" case where the blind model succeeds. The proposed model has great recovery skills and is able to climb obstacles almost high as the robot

Upstairs
17cm high, 30cm deep

| | Success | #Stairs |
|---|---|---|
| Ours | 100% | 13 |
| Blind | 0% | 2.2 |

Downstairs
17cm high, 30cm deep

| | Success | #Stairs |
|---|---|---|
| Ours | 100% | 13 |
| Blind | 100% | 13 |

Stepping Stones
30cm wide, 15cm apart

| | Success | #Stones |
|---|---|---|
| Ours | 94% | 9.4 |
| Blind | 0% | 0 |

Gaps
26cm apart

| | Success |
|---|---|
| Ours | 100% |
| Blind | 0% |

# Influence, Adoption, and Critique of the Paper

- CoRL 2022 Best Systems Paper Award
- Cited 188 times (MIT, Carnegie University, Sandford, ETH, DeepMind, etc…)
- Overall remarks: Advances in Egocentric Vision, Hybrid Learning
- Neural Volumetric Memory for Visual Locomotion Control, Ruihan Yang et al. 2023 : exploiting further the 3D structure of the environment
- Learning Robust and Agile Legged Locomotion Using Adversarial Motion Priors, Jinze Wu et al. 2023 : critic on low/moderate speed over challenging terrains, legged movements are unnatural and jerky

EPFL

# PROS & CONS



- PROS:
  - First end-to-end learning-based system enabling quadruped robots to navigate complex terrains
  - Effective use of egocentric vision, not elevation map (prone to noise), depth camera-based
  - Two-Phase Training Efficiency (RL + SL)
  - Minimal compute requirements
  - Robustness Across Terrains

- CONS:
  - Limited Field of View and Sensory Data (narrow perspective)
  - Challenges with Simulation-to-Real World Transfer (needs sometimes further retraining or fine-tuning, recreate specific failure scenarios)
  - Cannot interact with its surroundings in meaningful ways (possible future work on articulated arm to enable interaction)

# Possible exam questions

- **Q:** *How does the system in the paper enable the robots to retain past information for foot planning?*
- **A:** The system uses a recurrent neural network (RNN) to store recent visual and proprioceptive data, allowing the robot to retain information about terrain it has already passed. This enables accurate foot placement, even when the terrain is no longer visible in its current view.

- **Q:** *What were the key elements that the paper introduced that differentiates it from previous approaches in robotic locomotion?*
- **A:** Egocentric vision over elevation maps, simpler hardware, reduced noise, less computation, but limited camera view, terrain changes may be missed.