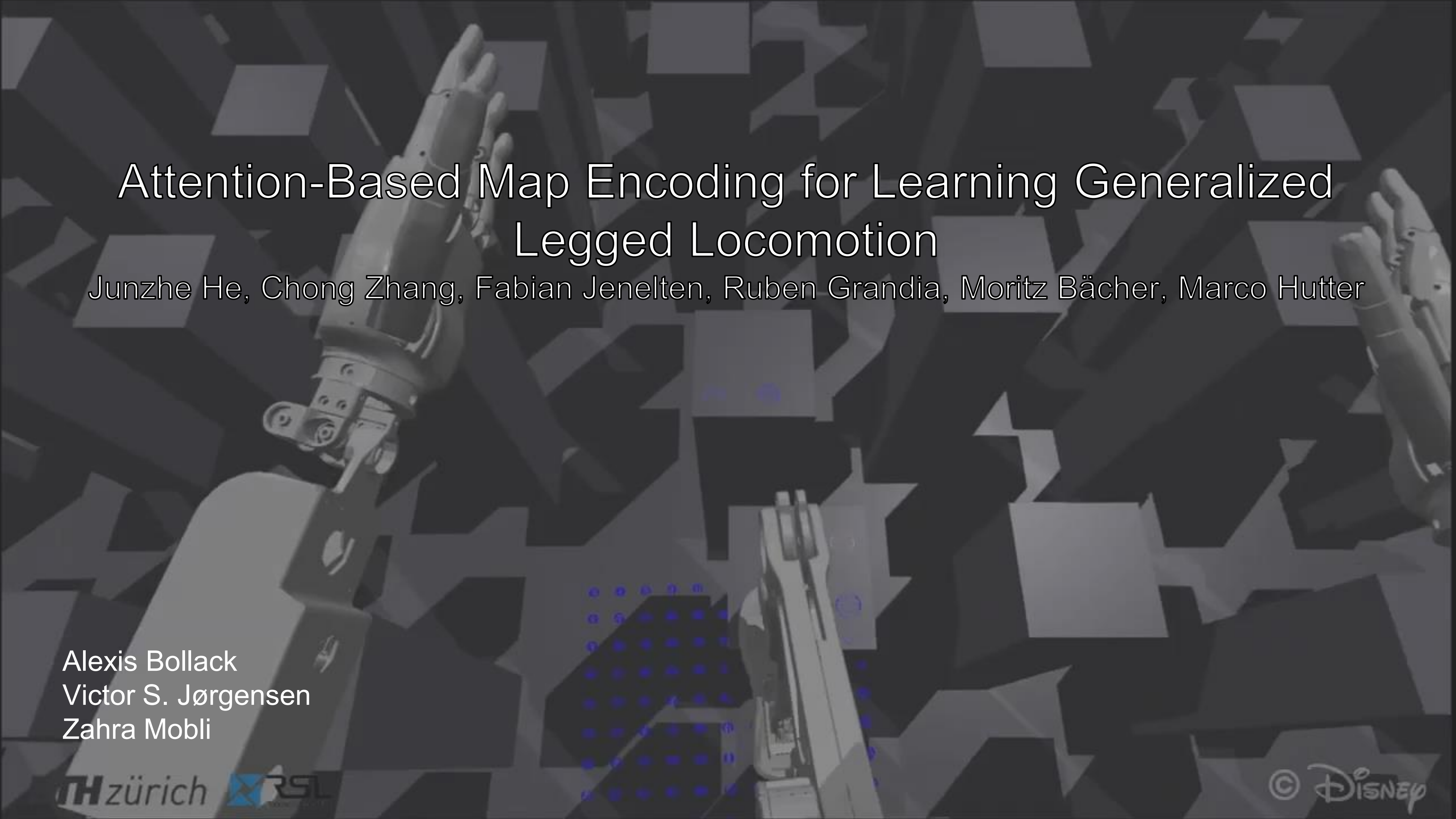
Attention-Based Map Encoding for Learning Generalized Legged Locomotion

Junzhe He, Chong Zhang, Fabian Jenelten, Ruben Grandia, Moritz Bächer, Marco Hutter



Alexis Bollack
Victor S. Jørgensen
Zahra Mobli

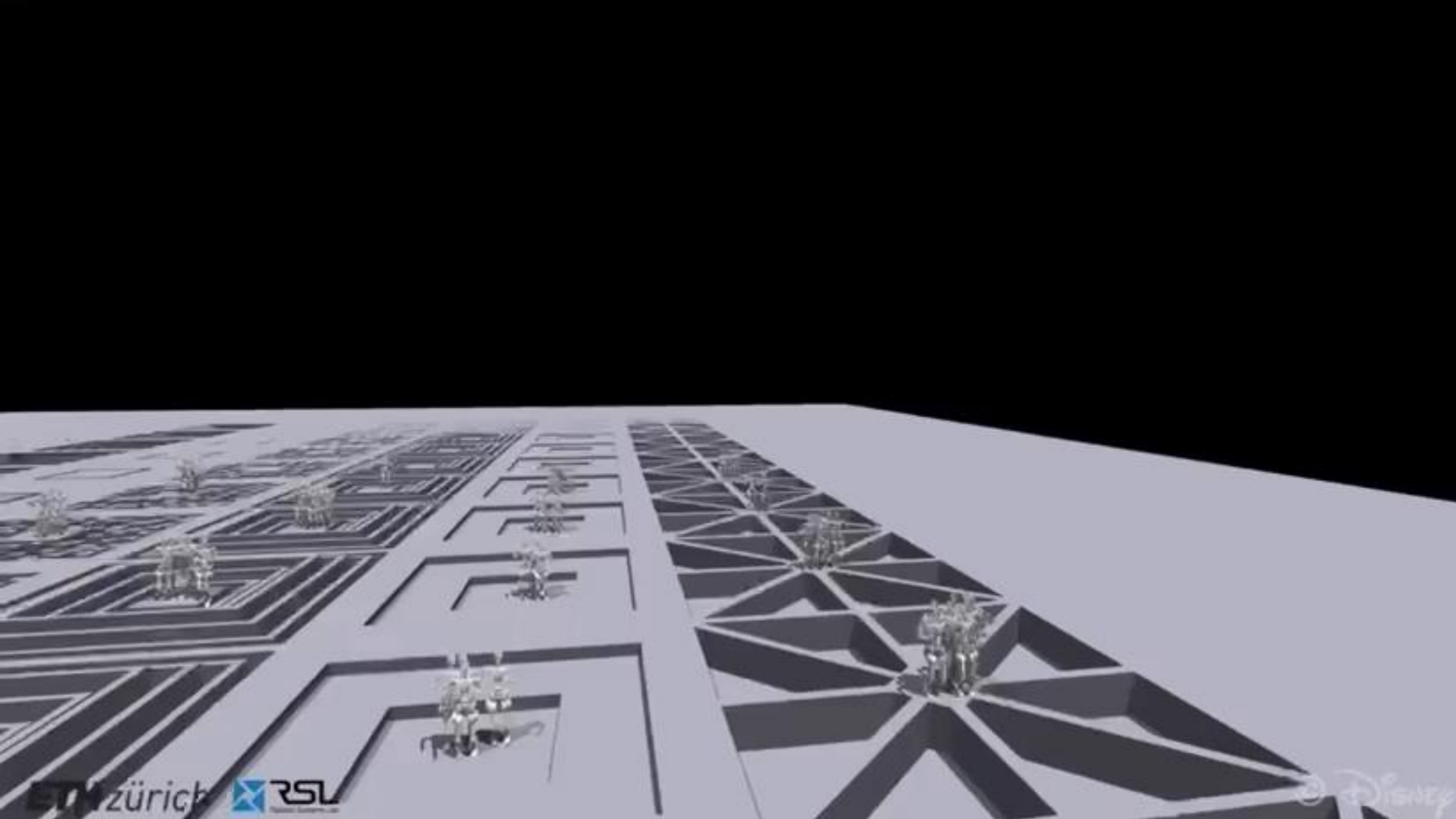
Summary:

- **Goal:**
  - Achieve **precise**, **robust**, and **generalized** dynamic locomotion across **diverse and challenging terrains** for **humanoid and quadruped robots**
- **Proposed Method:**
  - **End-to-End** Deep Reinforcement Learning Controller
  - **Attention**-Based Map Encoding
  - **Two-Stage** Curriculum Training Pipeline

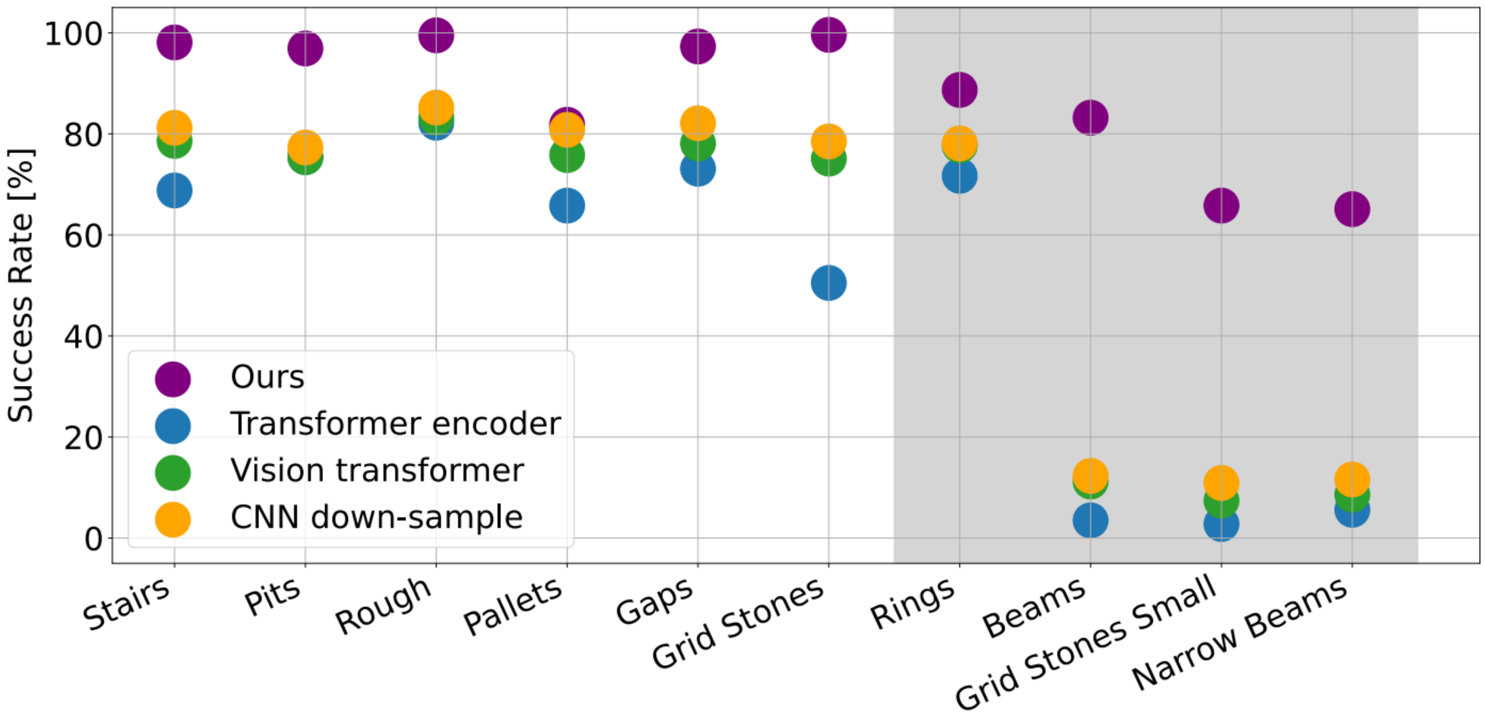
- **Key Properties:**
  - **Joint-level torque control**
  - **Adaptive multi-gait locomotion** depending on terrain type and commanded velocity
  - Use of **2.5D robot-centric height maps**

Multi Head Attention (MHA) - idea:

Aspect	LLM	Legged Robot
Goal	Focus on the most relevant words to understand context	Focus on the most relevant terrain for stable movement
Input	Sequence of word embeddings (meaning + position)	Proprioception + local terrain features
Process (MHA)	Attends to relationships between words across the sentence	Attends to relationships between body state and terrain
Output	Context-aware word representation	State-dependent map encoding (safe foothold selection)



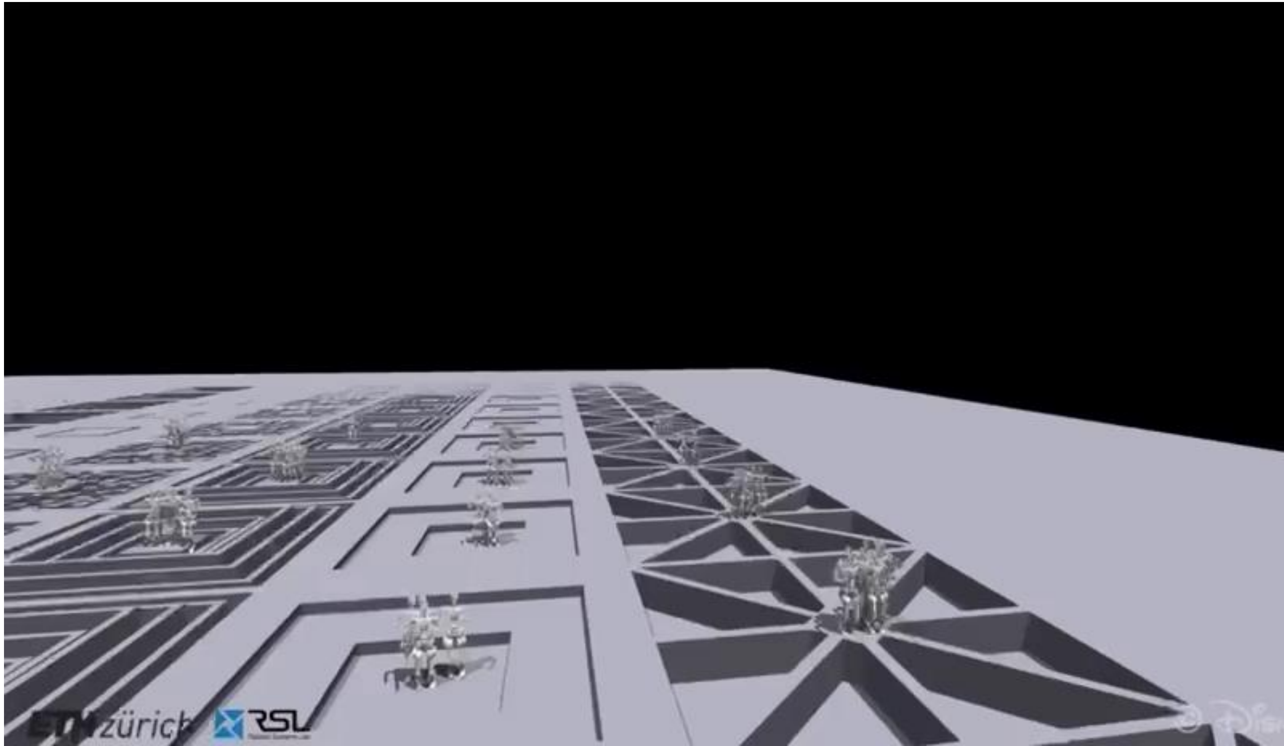
Ablation study – network structure



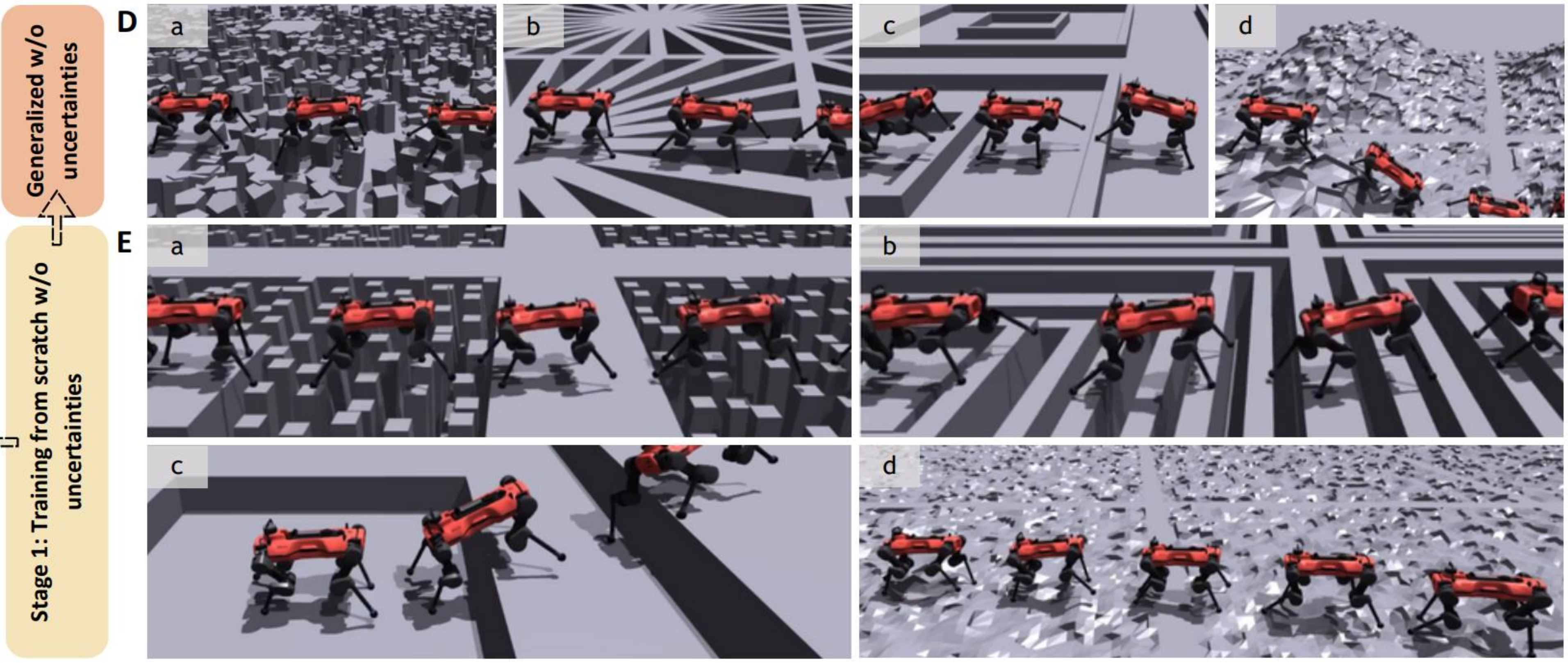
[1] N. Rudin et. al. (PMLR, 2022)
https://leggedrobotics.github.io/legged_gym/

Curriculum Based - Two Stage training policy

- Curriculum learning - “Learning To Walk in Minutes” *[1]*
- Privileged perception
- Learns specific skills

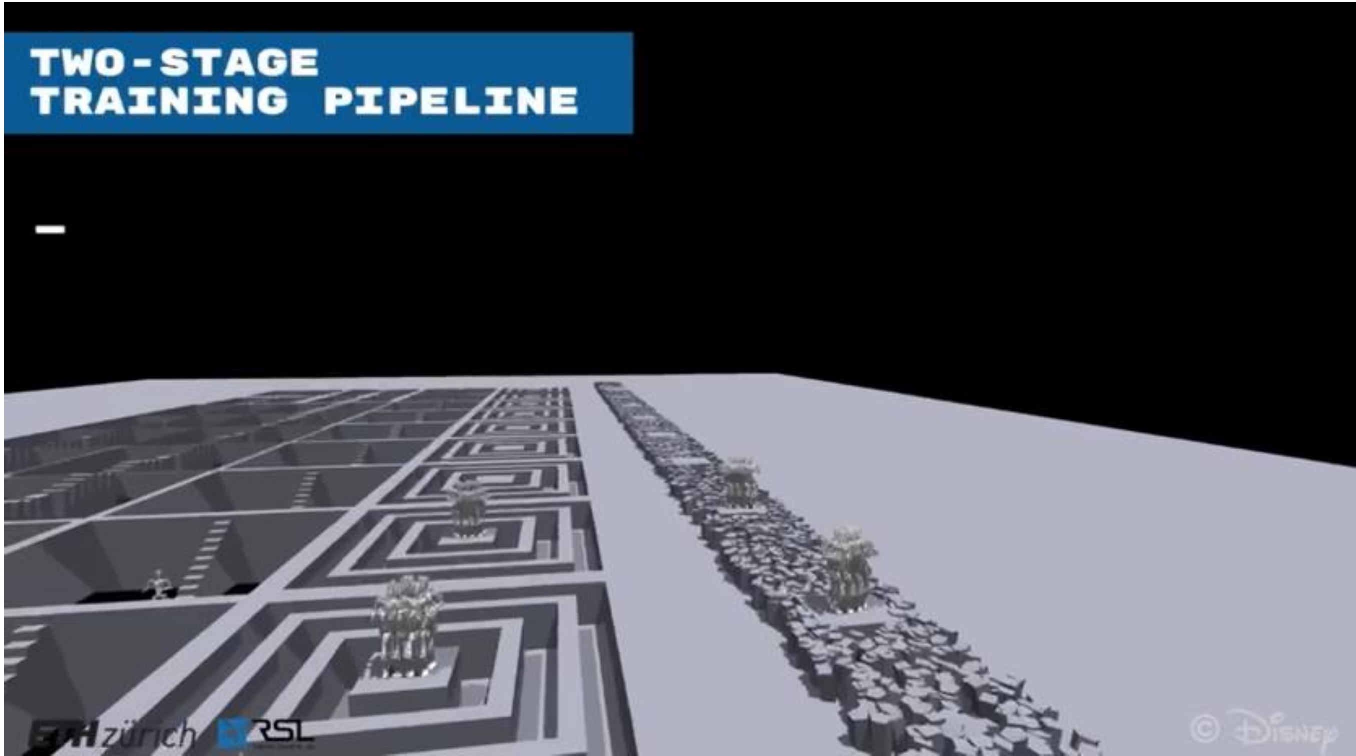


Generalization



Stage 2 – fine tuning

- Domain randomization to improve sim-to-real transfer
  - Mass
  - Friction coefficients
  - Artificial pushes
- Privileged critic

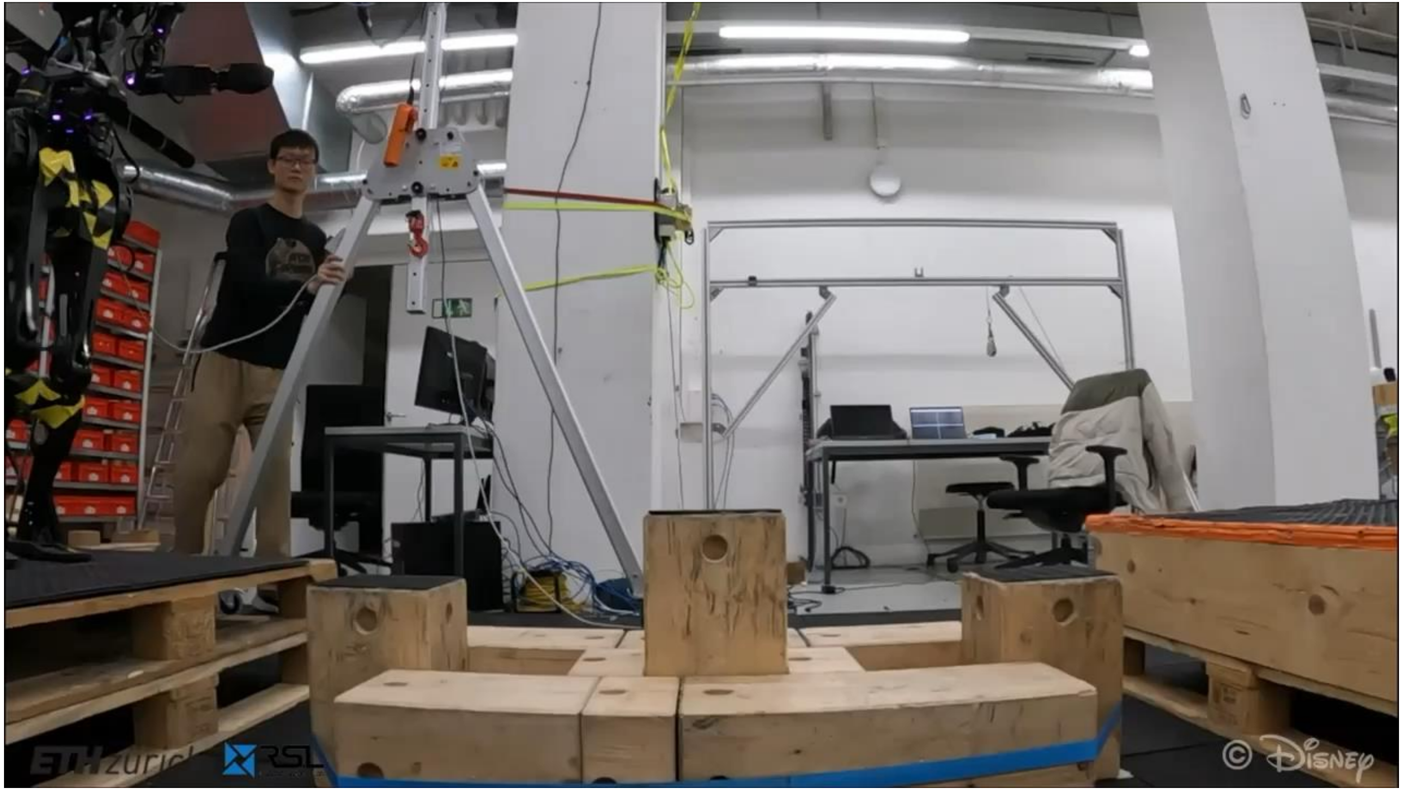


Ablation study – training

Result:

Result (cont.):

Result (cont.):



Citation Analysis:

The paper was recently published in **Science Robotics** on **27 Aug 2025** → it has **10 citations**

🔍 **RSL-RL: A Learning Library for Robotics Research** →

Cites *Attention-Based Map Encoding* as an example of advanced network architectures adopted in later works built on RSL-RL.

🔍 **Architecture Is All You Need: Diversity-Enabled Sweet Spots for Robust Humanoid Locomotion** →

Inspired by its vision–state fusion for robust foothold selection, they adopted a compact, local encoder that keeps perception separate from fast stabilizing feedback, enabling real-time performance.

Pros:

- ✓ Training is fast
- ✓ Works on multiple robot types
- ✓ Precise foot placement
- ✓ Robust to real-world issues
- ✓ Interpretable attention maps
- ✓ Whole-body agility

Cons:

- ✕ Training is slow
- ✕ Needs 2.5D height maps
- ✕ Focus only on locomotion
- ✕ GR-1 only performed indoors with motion capture

Thank you for your **attention** !!!

Any questions?

Ablation study – network structure

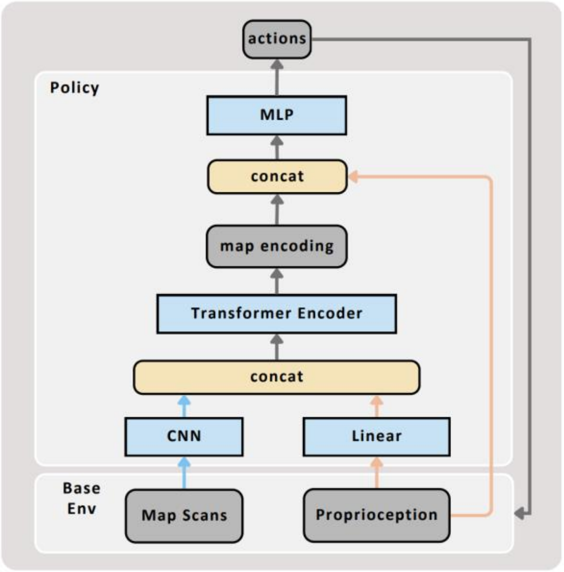


Fig. S1. Ablation Study on Network Structures - Transformer Encoder. We adopted the network structure proposed in [33]. The map scans are embedded with a CNN and the proprioception is embedded with a linear layer. The embedded map and proprioception features are then concatenated and sent to a transformer encoder to generate a map encoding.

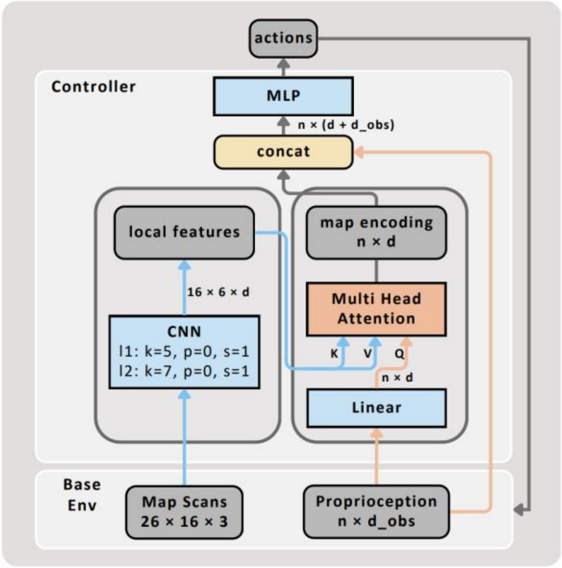


Fig. S2. Ablation Study on Network Structures - CNN downsampling. The CNN consists of two layers with kernel sizes (k) of 5 and 7, respectively. The padding and stride are set to be 0 and 1 for both layers.

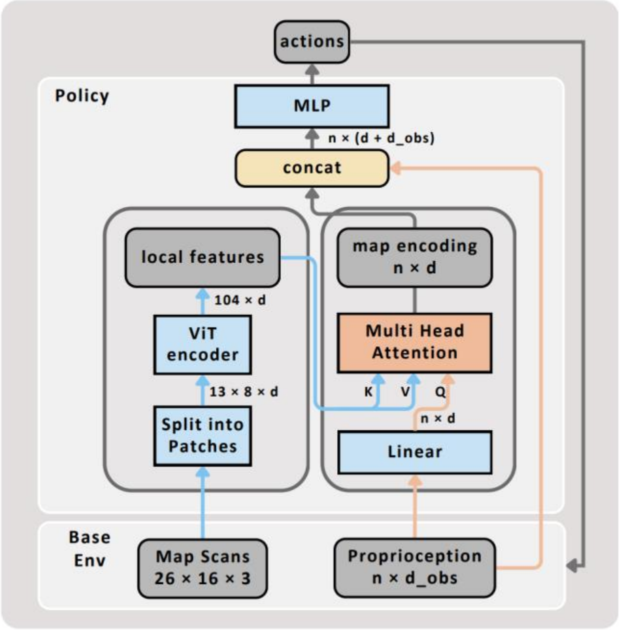


Fig. S3. Ablation Study on Network Structures - Vision Transformer. The map scans are split into 2×2 patches and sent to a vision transformer for feature extraction.

Table 2. Definition of reward terms for ANYmal-D and GR-1. The x-axis points forward and the z-axis points downward. The body, joint, and feet indices can be found in Appendix. The rewards marked in orange are only activated during stage 2 fine-tuning.

Reward Terms		functions	ANYmal-D		GR-1	
			weights	indices	weights	indices
Tasks	linear velocity tracking	$\exp(-\lVert\mathbf{v}^*_{xy,i} - \mathbf{v}_{xy,i}\rVert^2)$	5.0	$i = [i_{torso}]$	5.0	$i = [i_{torso}]$
	angular velocity tracking	$\exp(-\lVert\boldsymbol{\omega}^*_{z,i} - \boldsymbol{\omega}_{z,i}\rVert^2)$	3.0	$i = [i_{torso}]$	3.0	$i = [i_{torso}]$
	termination penalty	$-n_{termination}$	200	N/A	200	N/A
	collision penalty	$-n_{collision,i}$	1	$i = [i_{shank}]$		
Regulations	action rate	$-\lVert\mathbf{a}_{\mathbf{j}_t} - \mathbf{a}_{\mathbf{j}_{t-1}}\rVert^2$	5.0e-3	j = [0:12]	5.0e-3	j = [0:23]
	joint acceleration penalty	$-\lVert\ddot{\mathbf{q}}_j\rVert^2$	2.5e-7	j = [0:12]	1e-6	j = [15:18, 19:22]
	joint torques penalty	$\& -\lVert\boldsymbol{\tau}_j\rVert^2$	2.0e-5	j = [0:12]	1e-4 5e-5 5e-5	j = [15, 19] j = [3,9] j = [2,8]
	joint position limits	$-max(\lvert q_j\rvert - 0.9q_{lim,j}, 0)$	1.0	j = [0:12]	10	j = [0:23]
	joint velocity limits	$-max(\lvert\dot{q}_j\rvert - 0.9\dot{q}_{lim,j}, 0)$	1.0	j = [0:12]	0.1	j = [0:23]
	joint torque limits	$-max(\lvert\tau_j\rvert - 0.8\tau_{lim,j}, 0)$	0.2	j = [0:12]	2e-3	j = [0:23]
Styles	linear velocity penalty	$-\mathbf{v}^2_{z,i}$	1.0	$i = [i_{torso}]$		
	angular velocity penalty	$-\lVert\boldsymbol{\omega}_{xy,i}\rVert^2$	5.0e-2	$i = [i_{torso}]$	5.0e-2	$i = [i_{torso}]$
	contact forces penalty	$-max(\lVert\mathbf{F}_f\rVert - 700, 0)$	2.5e-5	f = [0:4]		
	foot slippage penalty	$-c_f * \lVert\mathbf{v}_f\rVert$	0.5	f = [0:4]	1.0	f = [0, 1]
	joint deviation penalty	$max(\lVert q_j - q_{0,j}\rVert^2 - 0.25, 0.0)$			0.5	j = [15:23]
	no fly	$-n_{zero\_contact}$			5.0	f = [0, 1]
	straight body	$-\lVert g_i\rVert^2$			3.0	$i = [i_{torso}, i_{pelvis}, i_{feet}]$
	standing joint positions penalty	$-\lVert\mathbf{q}^*_j - \mathbf{q}_j\rVert$	0.1	j = [0:12]		
	standing joint velocity penalty	$-\lVert\dot{\mathbf{q}}^*_j - \dot{\mathbf{q}}_j\rVert$	0.5	j = [0:12]	0.2	j = [0:23]