IndoorSim-to-OutdoorReal: Learning to Navigate Outdoors without any Outdoor Experience

Google DeepMind

Joanne Truong[1,2], April Zitkovich[2], Sonia Chernova[1], Dhruv Batra[1,3], Tingnan Zhang[2], Jie Tan[2], Wenhao Yu[2]

[1] Georgia Tech, [2] Google DeepMind, [3] Meta AI

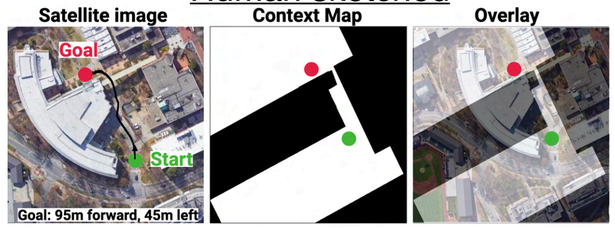
Georgia Institute of Technology Meta AI

Summary

- **Autonomous long-range navigation** (100 meters+) **in novel outdoor environments**
- Trained entirely in **simulation of indoor environments** (short-range, on average ~8 meters), **zero-shot sim-to-real transfer to the outdoors**
- Robot operates using **onboard sensors** + a inaccurate high-level map (**Context-Map**)

Context-Map

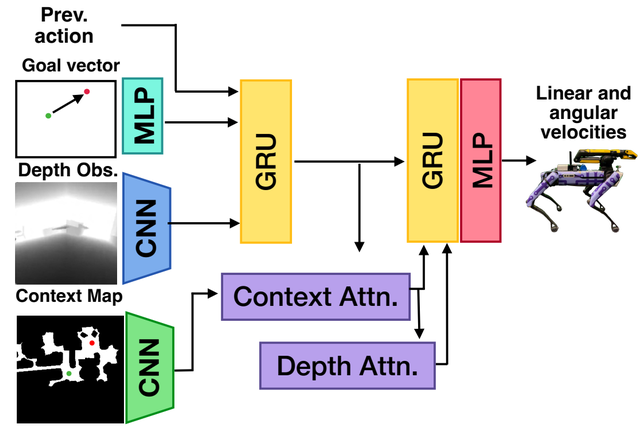
Human sketched



Automatically generated

- Does not need to be accurate or complete
- Real-world obstacles (trees, bushes, cars, pedestrians, etc.) are not drawn on map
- Provides the robot with a hint to the goal
- Abstracted binary map can be used for both indoor and outdoor navigation
- <u>In simulation:</u> Freely accessible via top-down maps of indoor environments
- <u>In the real-world:</u> Generated via human sketching (top row), or automatically generated from satellite map (bottom row)

Model Architecture



- Egocentric images to avoid obstacles
- Context-Maps for high-level guidance

Task: Point-Goal Navigation

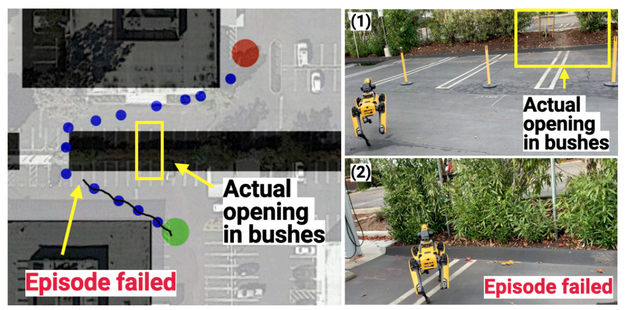
Navigate to a goal location in a novel environment

Real-World Results

Route #	Goal	Method	SR ↑	Distance Travelled (m) ↑
1	38m Forward, 16m Right	No-Context	0.0	16.6±0.1
		Context-Map	**100.0**	**63.4**±2.5
2	90m Forward, 30m Left	No-Context	0.0	9.7±3.4
		Context-Map	**100.0**	**112.2**±1.8
3	95m Forward, 45m Left	No-Context	0.0	5.1±0.3
		Context-Map	**100.0**	**129.8**±2.8

- 3 novel environments with various terrains (dirt slopes, grass, etc.)
- **Context-Map policy: 100% success rate**
- **Policy without context fails completely**
- Context-Map policy takes shortcuts not present in the map when vision senses free space

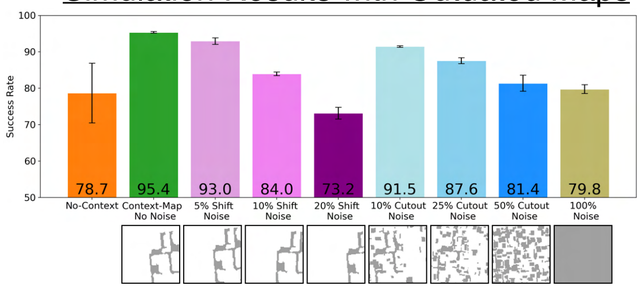
Classical planning + Context-Map baseline



- Classical approaches are sensitive to map input; cannot adapt to inaccuracies in Context-Map
- A perfect, always up-to-date map in the real-world is not realistic

Simulation Ablations

Simulation Results with Outdated Maps



- Simulation evaluation with varying degrees of shift and cutout noise to Context-Map
- Context-Map policy is robust to noise
- At 100% noise, policy regresses to No-Context policy behavior