Dynamic obstacle avoidance in highly-constrained environments using vision-based RL

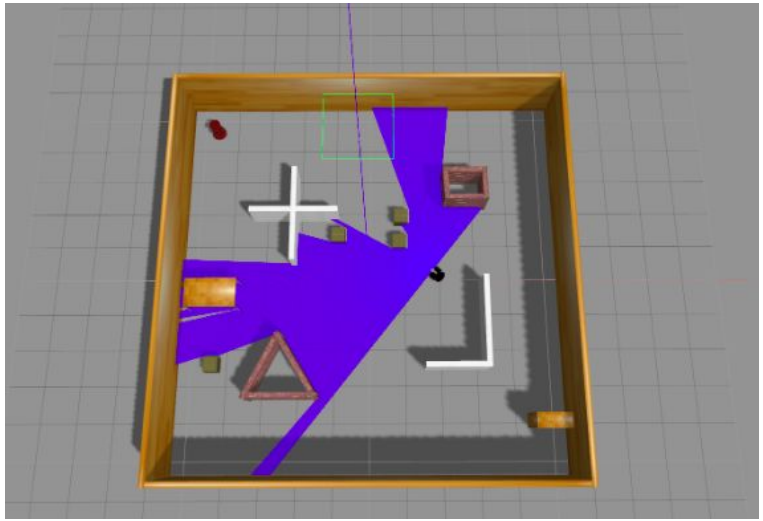
Atharva Pusalkar, Jash Shah, Madhu Korada, Onkar Thorat

Motivation

1. Current RL navigation methods do not work in dynamic environments
2. Leverage multiple vision sensors provided by Stretch RE1 robot.
3. Local optimum problem of reactive navigation is only partially mitigated. Current methods require human intervention for recovery.

Motivation

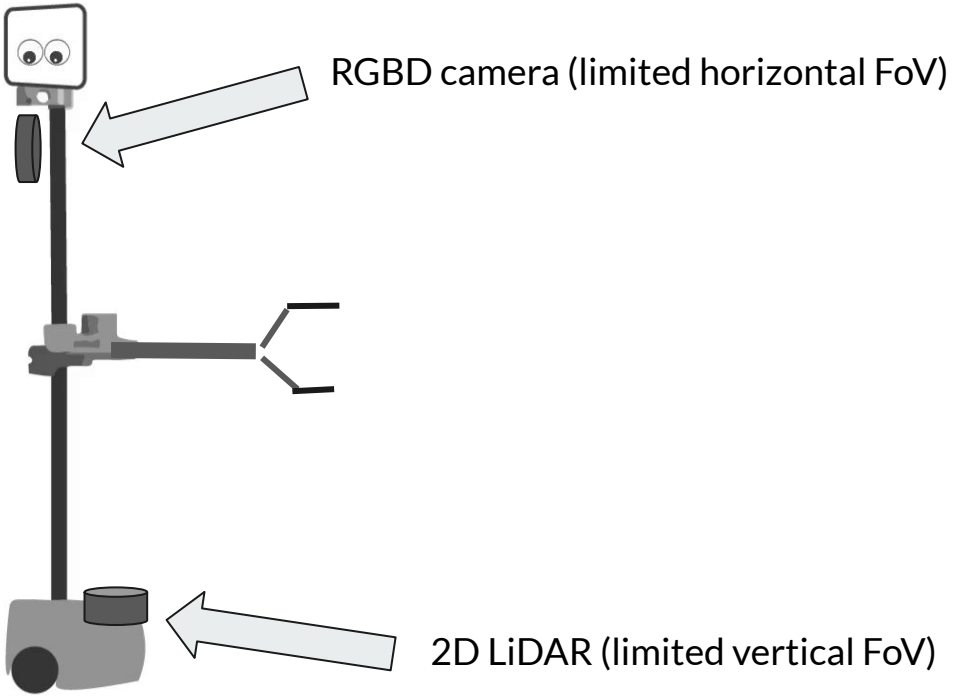
Current RL-based navigation does not work in dynamic environments. Current SOTA only works well with static obstacles



Motivation

Leverage multiple vision sensors provided by mobile manipulator

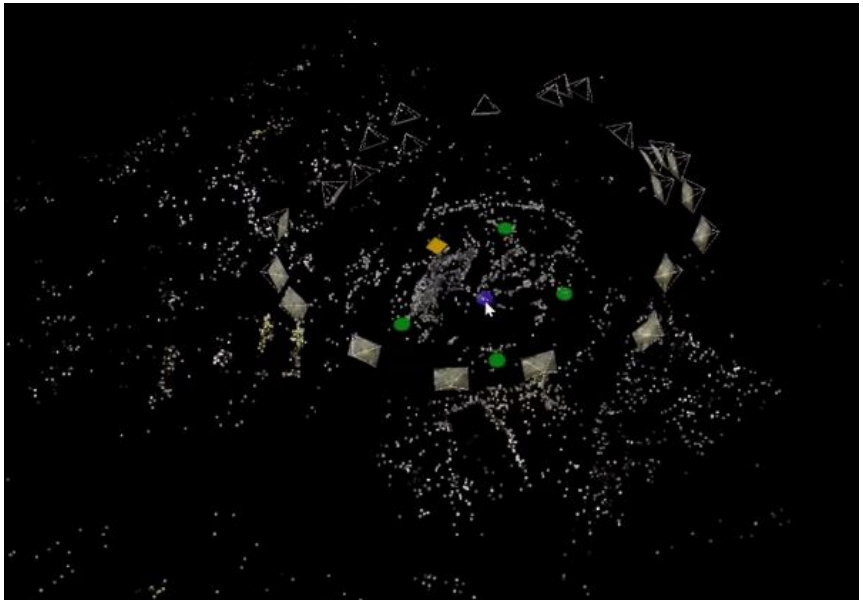
Stretch is one of the most widely used robots in CMU RI



Motivation

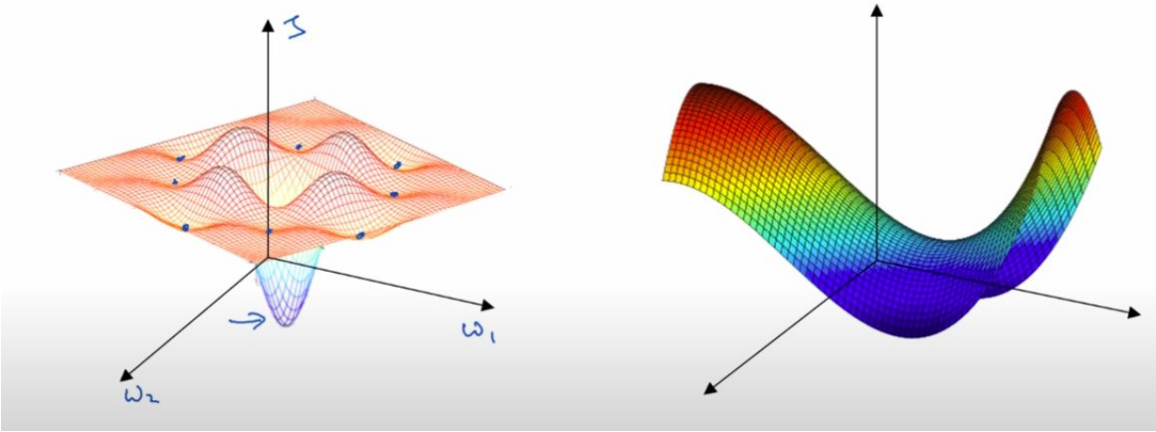
Vision is always used as supporting input because point cloud input is sparse

Solution => Sensor fusion

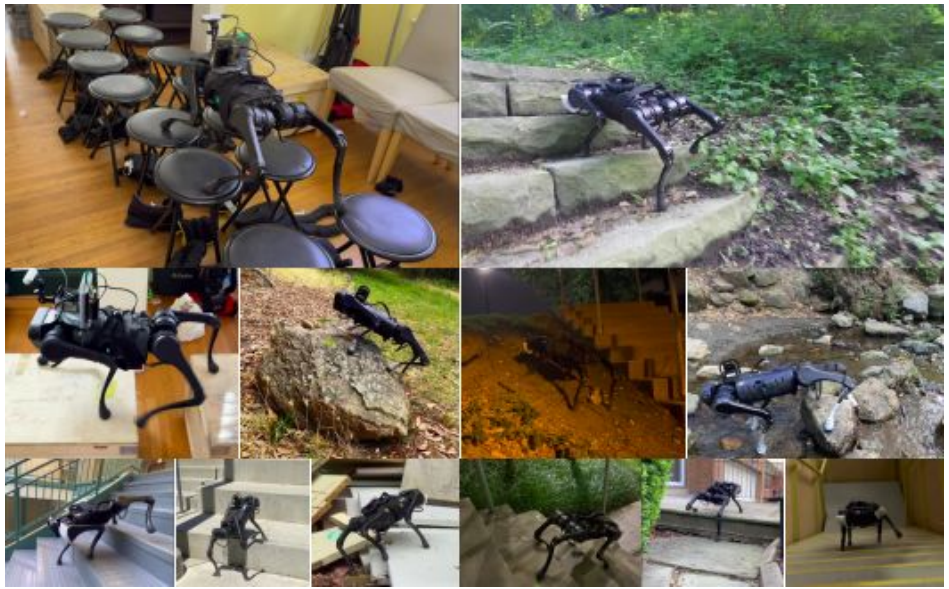


Motivation

Local optimum problem of reactive navigation needs to be mitigated



Prior Work

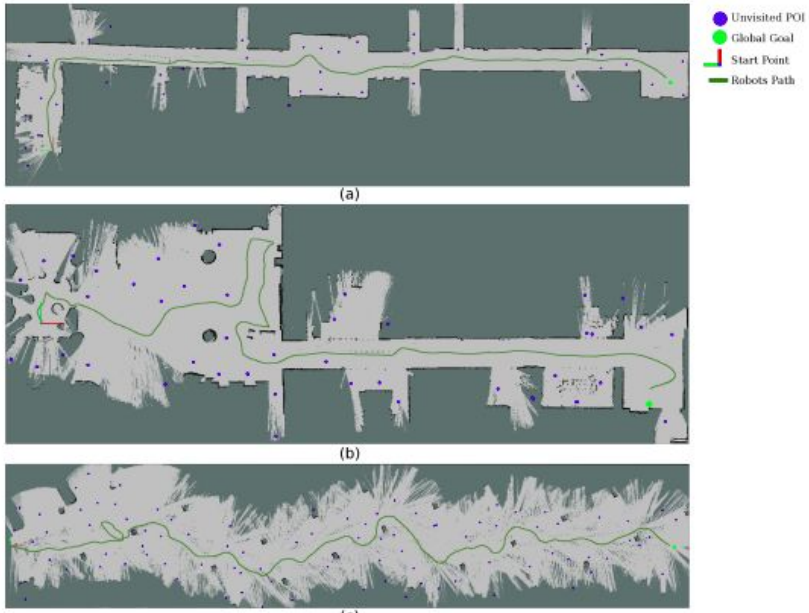


Agarwal, Ananye, et al. "Legged locomotion in challenging terrains using egocentric vision." *Conference on Robot Learning*. PMLR, 2023.

Works for legged locomotion

Robot fails in some scenarios due to terrain mismatch between simulation and real world

Prior Work



Cimurs et. Al, Goal-Driven Autonomous Exploration Through Deep Reinforcement Learning, 2021

Does not work for dynamic obstacles

Requires dense point cloud information

Uses depth data as input to the network

Baseline + Modifications

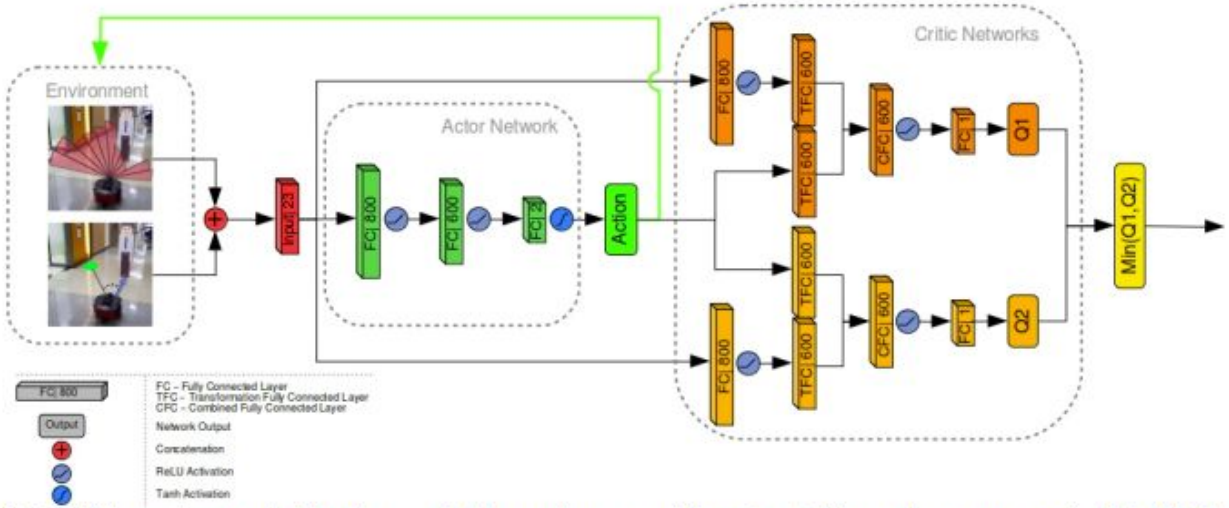


Fig. 3. TD3 network structure including the actor and critic parts. Layer type and the number of their respective parameters are described within the layers. TFC layers refer to transformation fully connected layers τ and CFC layer to combined fully connected layer L_c.

Cimurs et. Al, Goal-Driven Autonomous Exploration Through Deep Reinforcement Learning, 2021

Used monolithic actor from Vision Locomotion paper

Switched from TD3 to PPO

Our model

Vision

Robot state

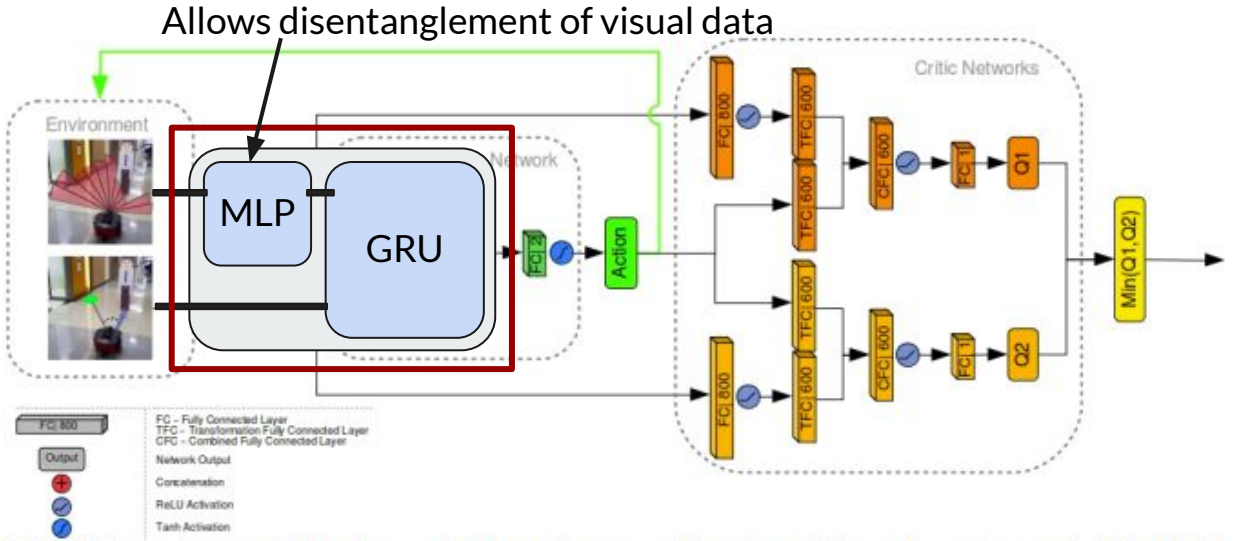


Fig. 3. TD3 network structure including the actor and critic parts. Layer type and the number of their respective parameters are described within the layers. TFC layers refer to transformation fully connected layers τ and CFC layer to combined fully connected layer L_c.

Our work

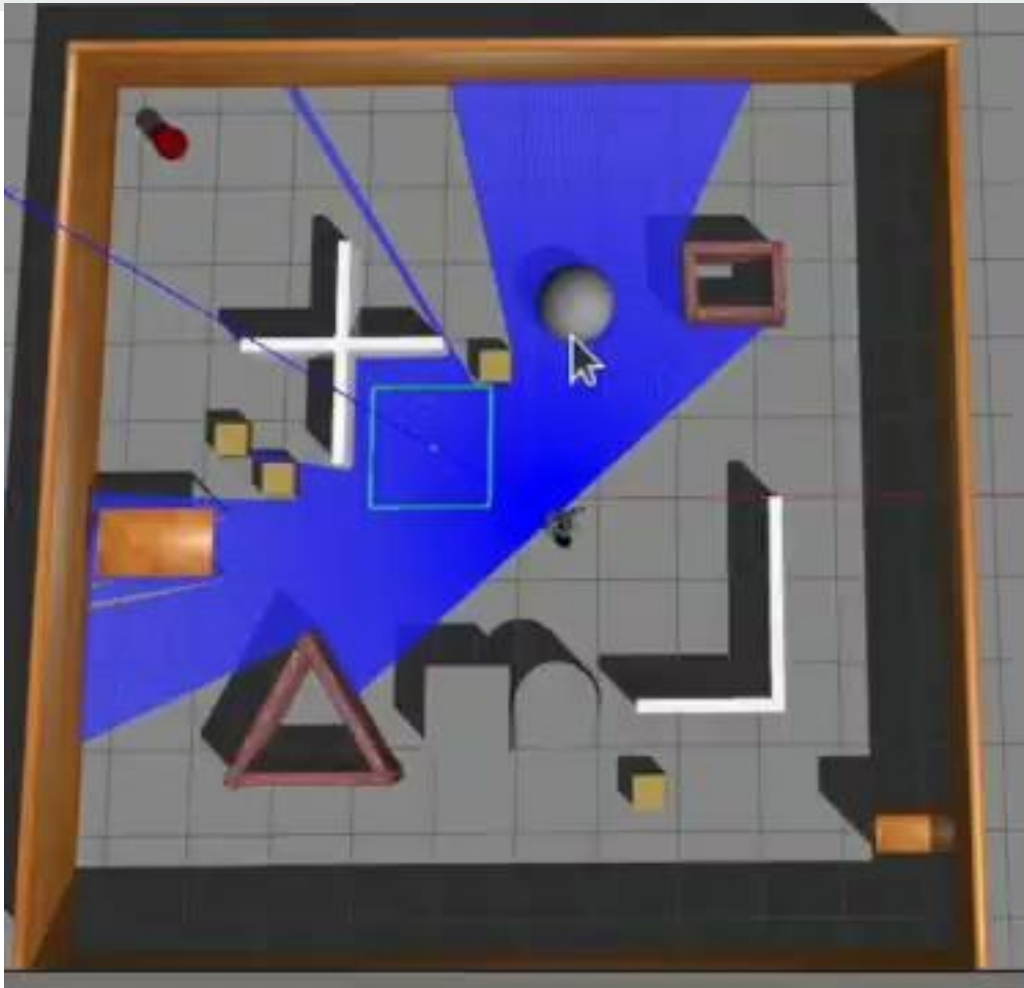
1. Implemented TD3 and sim environment for dynamic obstacle avoidance
2. Sensor fusion with limited FoV vision
3. Sim2Real using Stretch RE1
4. Demo in an unseen environment with narrow pathways

Method

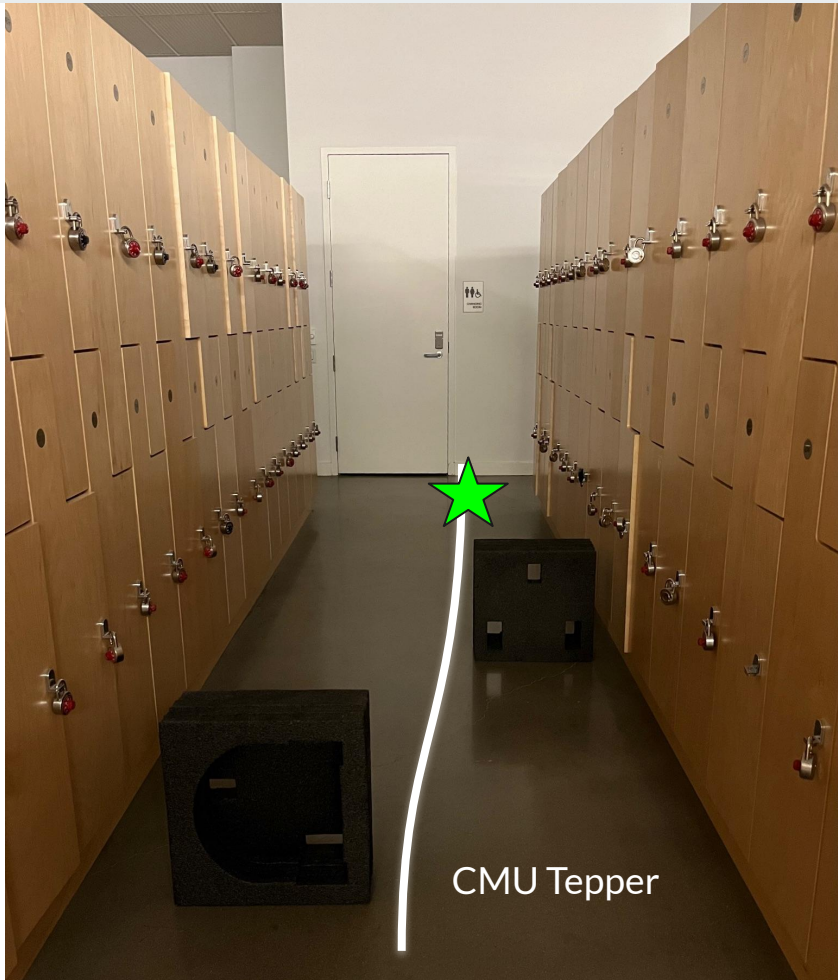
Custom reward function:

1. L2 Distance to goal
2. Linear and angular jerk penalty
3. Survival bonus
4. Velocity penalty: $\mathscr{V}$ - $|\omega|$
5. Laser minimum distance penalty: $x = \begin{cases} 1 - x & \text{if } x < 1 \\ 0, & \text{otherwise} \end{cases}$

Results (Sim)



Sim2Real



Summary

Model needs further tuning and training to further optimize the control policy

Can add custom noise models to input to better emulate real-world data

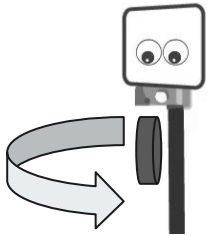
TD3 parameters need to be tuned to improve training stability

Future work

Switch to PyBullet (Assistive Gym, Z Erickson, et. al, ICRA 2020)

Add more noise models to depth data

Include head camera rotation in the pipeline (Stretch goal)



Thank You