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

Atharva Pusalkar
Carnegie Mellon University
apusalka@andrew.cmu.edu

Jash Shah
Carnegie Mellon University
jbshah@andrew.cmu.edu

Madhu Korada
Carnegie Mellon University
mkorada@andrew.cmu.edu

Onkar Thorat
Carnegie Mellon University
othorat@andrew.cmu.edu

Abstract

Robot navigation in dynamic environments is a challenging problem that requires a comprehensive and integrated approach. In this project, we propose a deep reinforcement learning model utilizing the Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm, which combines visual data with 2D lidar point cloud data to facilitate goal point navigation for a mobile robot operating in environments containing both static and dynamic obstacles. The model was trained within a custom Gazebo simulation environment and then successfully deployed on Stretch RE1 robot, which was then tested for its ability to navigate to a goal point in an unseen environment characterized by narrow spaces. The robot demonstrated success in accomplishing this task. Nonetheless, we recognize the limitations of our experiments, as they were conducted in a restricted number of real-world environments, and the model may not perform optimally in more complex settings.

1 Motivation

The advancement of robotics has played a critical role in enhancing various industries, from manufacturing to healthcare. With the increasing demand for robots to operate in complex and dynamic environments, the conventional approach to robot navigation, which employs independent modules for simultaneous localization and mapping (SLAM), motion planning, and control, may not suffice. In such dynamic settings, robots must possess a high degree of adaptability and efficiency to interact with humans and other dynamic agents. Although the current navigation approach may work well in static and unchanging environments, it fails to address the challenges that arise in dynamic settings. Dynamic environments introduce uncertainty, unpredictability, and real-time constraints that require a more comprehensive and integrated approach to robot navigation. Therefore, there is a pressing need to develop an approach that enables robots to navigate the real world seamlessly, allowing them to coexist and collaborate with humans and other dynamic agents.

The motivation for this project stems from the need to make robots more effective and efficient in navigating through complex and dynamic environments, ultimately making human lives easier and more convenient. By developing a more holistic and integrated approach to robot navigation, we can broaden the range of applications for robots, from healthcare to manufacturing. For instance, in the healthcare sector, robots can assist with surgeries and rehabilitation exercises, reducing the workload on medical personnel and improving patient outcomes. In manufacturing, robots can increase productivity and reduce the risk of workplace accidents. Therefore, this project's objective is to develop a more comprehensive and integrated approach to robot navigation that enables robots to operate efficiently in dynamic environments. This will involve the use of advanced techniques, such as deep learning, reinforcement learning, and computer vision, to create a robust and adaptable system for robot navigation. Through this project, we aim to pave the way for robots to be used in a wider range of applications and improve the quality of life for people worldwide.

2 Prior Work

The main work in this domain was by Agarwal et. al [1] where they propose a method for extracting terrain information from egocentric visual data, which can be used to improve the robot's gait and stability in difficult terrain. This was followed by Fu et. al. [7] in which the visual perception makes use of a stereo camera system to provide depth information about the environment and the proprioceptive feedback component makes use of joint angle sensors and accelerometers to provide information about the robot's body position and motion. A deep neural network is then used to combine the visual and proprioceptive information and generate commands for the robot's leg joints where the neural network is trained using reinforcement learning.

The paper by Cimurs et. al [6] proposes a novel approach for autonomous exploration in unknown environments using deep reinforcement learning (DRL). The authors focus on addressing the challenge of exploring new and complex environments where the agent does not have prior knowledge or a map of the surroundings. The proposed method involves training an agent to perform a goal-driven exploration task using a DRL algorithm known as Soft Actor-Critic (SAC).

Another interesting work was by Gupta et. al [9] in which they propose a cognitive mapping and planning approach that allows the robot to build a spatial map of the environment and plan efficient paths to navigate through it. This approach combines two main components - a topological map and a planner. The topological map is a high-level representation of the environment that captures the spatial relationships between different locations. The authors use a graph-based representation, where nodes represent distinct locations and edges represent the connections between them. The map is built incrementally as the robot explores the environment, using a combination of visual and proprioceptive sensing. On the other hand, the planner is responsible for generating efficient paths through the environment. The authors use a graph-search algorithm that operates on the topological map to find the shortest path between the robot's current location and its goal. The planner takes into account the robot's current location, the goal location, and any obstacles or constraints in the environment.

Chaplot et. al. [3] present a method for navigation of robots towards specific objects in an indoor environment. Their approach combines semantic exploration with object recognition techniques and consists of two parts: a semantic exploration module and an object recognition module. The semantic exploration module uses a combination of topological and geometric information to explore the environment in a goal-oriented manner, by prioritizing areas that are likely to contain objects of interest. The object recognition module uses a convolutional neural network to recognize specific objects of interest in the environment.

Chaplot et. al. [4] also present a method for autonomous exploration and mapping of environments by a robot using a combination of neural networks and SLAM techniques. The Active Neural SLAM algorithm consists of two parts: a neural network that predicts the reward of exploring a particular area, and a SLAM algorithm that builds a map of the environment. The neural network takes as input the robot's current state and a partial map of the environment, and outputs a reward value for each unexplored area. The SLAM algorithm uses the robot's sensory data to build a map of the environment, which is used to update the neural network's input and refine the reward predictions.

Bansal et. al. [2] propose a method that combines optimal control and learning to enable the robot to navigate through complex and previously unseen environments. The method consists of two parts: an optimal control policy and a reinforcement learning algorithm. The optimal control policy is used to generate initial trajectories for the robot based on a model of the environment. The reinforcement learning algorithm is used to refine these trajectories based on the robot's actual experiences in the environment.

3 Proposed Idea

The proposed idea is to explore the potential of using a learning-based approach to automate the local planner for a mobile robot, thereby making the local planner more adaptable and robust to handle complex scenarios in the environment. In particular, we aim to investigate the effectiveness of this approach in comparison with traditional robotic algorithms currently used for local planner design.

To achieve this objective, we will employ a mobile robot that will learn to traverse in a small area with visual data as its input. The robot will learn to identify obstacles, plan its path, and avoid collisions in real-time. The current robotic algorithms used for local planner design work well in simple and static environments. However, they face significant challenges in dynamic environments where there is an increased level of uncertainty, variability, and unpredictability.

We will leverage recent advancements in machine learning, such as deep learning and reinforcement learning, to develop an adaptable and robust local planner that can handle complex scenarios in the environment. This learning-based approach will enable the robot to learn from experience, and through continuous training, the local planner will improve its ability to plan paths that avoid obstacles and reach the desired destination efficiently.

The results of this research will have significant implications for the field of robotics, particularly in developing adaptable and robust local planners for mobile robots. By improving the local planner's efficiency, we can enhance the robot's ability to operate in complex and dynamic environments, such as disaster zones, warehouses, and manufacturing plants, where they can assist humans in carrying out dangerous and repetitive tasks. Ultimately, this research will contribute to advancing the capabilities of mobile robots and improving their functionality, leading to greater convenience and safety for people worldwide.

3.1 Intuitive sense of our idea

The intuition behind the proposed idea is grounded in the limitations of traditional robotic algorithms, which may not be able to handle all possible scenarios that a mobile robot may encounter in a complex and dynamic environment. In such environments, the level of uncertainty, variability, and unpredictability can make it challenging for traditional algorithms to plan paths and avoid obstacles effectively. Consequently, there is a need for a more adaptable and robust approach that can handle complex scenarios and improve the efficiency of mobile robots.

A learning-based approach can provide a solution to this problem by leveraging recent advancements in machine learning, such as deep learning and reinforcement learning. By allowing the robot to learn from experience, a learning-based approach can adapt to changes in the environment and improve its performance over time. This approach can provide a more generalized solution that can be applied to various environments, making it more versatile than traditional algorithms.

Moreover, a learning-based approach can enable the robot to learn from data generated by its own sensors, which is particularly important in environments where there is an increased level of uncertainty and unpredictability. The robot can use this data to develop a better understanding of the environment and improve its decision-making capabilities in real-time.

Furthermore, a learning-based approach can help address the problem of handcrafting rules and features for traditional algorithms, which can be time-consuming and challenging in complex environments. In contrast, a learning-based approach can learn the relevant features and rules from the data and adapt to the environment, making it more efficient and effective in handling complex scenarios.

The proposed research aims to investigate the effectiveness of a learning-based approach to improve the local planner of a mobile robot. By using a learning-based approach, we hope to develop a more adaptable and robust local planner that can handle complex scenarios and improve the efficiency of mobile robots in various environments. The results of this research will have significant implications for the field of robotics, particularly in developing more versatile and efficient mobile robots that can assist humans in carrying out dangerous and repetitive tasks, such as disaster response, warehousing, and manufacturing.

3.2 Relation with prior work

Some of the earliest work to automate planning tasks involved behavior cloning using imitation learning. The idea is to fit a deep neural network on a sample input and the corresponding output data for a tele-operated robot. The inference is then run on real-time data from the same source. However, this method fails to generalize in a non-familiar environment. Other methods that were previously used involved deep reinforcement learning using Q-networks but require a large amount of training data.

Some novel work includes A. Giusti et. al. [8] that uses a combination of CNN and LSTMs to navigate mobile robots along forest trails. This method does generalize well beyond environments such as forest trails and the experiments were conducted in a simulated environment which proved difficult Sim2Real transfer. As discussed before Chaplot et. al[5] presents a method for autonomous navigation using semantic exploration that allows a robot to efficiently estimate the relative arrangement of objects in view. However, this method does not take into account dynamic scenarios such as collisions and expects the input representation of the environment to be perfect without any ambient noise. Moreover, it only takes feedback from the reward policy and does not penalize the model against such dynamic conditions. Finally, the model inference is also executed in a simulated environment which also limits the Sim2Real transferability.

4 Results

4.1 Our Implementation

We focus on a deep reinforcement learning (DRL) model, inherited from [6], trained using Twin Delayed Deep Deterministic Policy Gradient (TD3). Our method uses visual data fused with 2D lidar point cloud data to make a differential drive robot navigate in an environment with static and dynamic obstacles. We experimented with various approaches to make the robot more susceptible to dynamic obstacles. We figured out that fusing 2D lidar point cloud data with visual data significantly improved our performance.

Our vision data inherently has a smaller field of view and it would be difficult to handle the dynamic obstacles coming from robot's blind spots. To improve the robot's ability to detect and navigate around dynamic obstacles, we experimented with different approaches, ultimately deciding to fuse 2D lidar point cloud data with visual data to create a more comprehensive input representation for the DRL model. A multi-layer perceptron (MLP) was designed to fuse the 2D point cloud data with the visual data, which was then fed into the TD3 algorithm for training.

During the training process, we used a reward function that included several components, such as the L2 distance to the goal, linear and angular jerk penalties, a survival bonus, a velocity penalty, and a laser minimum distance penalty. The L2 distance to the goal component provided a reward proportional to the robot's distance from the goal, incentivizing the robot to reach its target efficiently. The linear and angular jerk penalties aimed to reduce the robot's sudden and abrupt movements, promoting smooth and controlled motion. The survival bonus provided a constant reward to ensure the robot's survival and continued exploration of the environment. The velocity penalty component penalized the robot's velocity based on its angular velocity, encouraging the robot to navigate at a constant speed. Finally, the laser minimum distance penalty component penalized the robot's proximity to obstacles detected by its lidar sensors, incentivizing the robot to avoid collisions with obstacles. Thus, our custom reward function included:

1. L2 Distance to goal

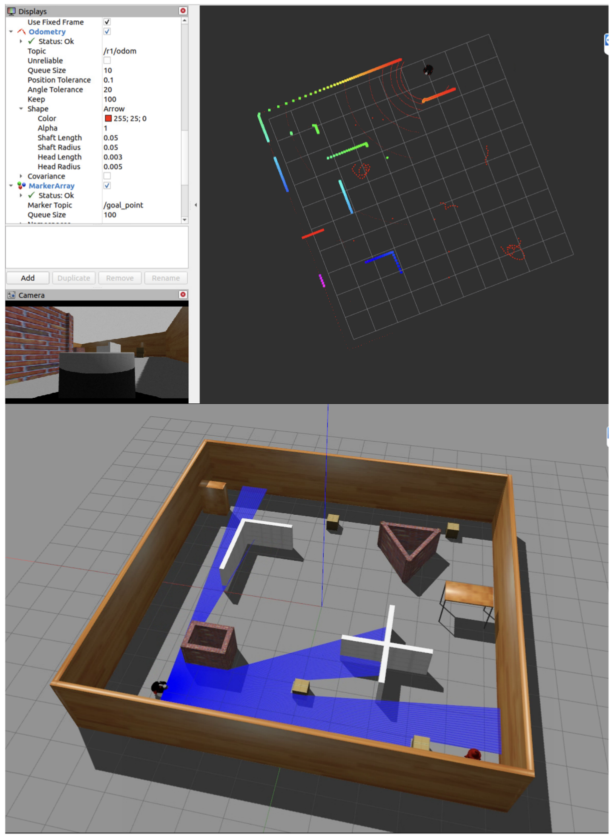

Figure 1: Twin Delayed Deep Deterministic Policy Gradient training run

2. Linear and angular jerk penalty
3. Survival bonus
4. Velocity penalty:

$$v - |\omega|$$

5. Laser minimum distance penalty:

$$x = \begin{cases} 1 - x, & \text{if } x < 1 \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

Our approach showed significant improvements in the robot's performance, particularly in the presence of dynamic obstacles. By fusing 2D lidar point cloud data with visual data, we were able to improve the robot's ability to detect obstacles and navigate around them efficiently. We tested the trained model on a Stretch RE1 robot equipped with a realsense camera and 2D lidar, deploying it in an unseen environment with narrow spaces. The robot was able to successfully navigate to the goal point, even with the presence of static and dynamic obstacles. Our approach demonstrated its effectiveness in improving the robot's ability to handle dynamic obstacles and navigate in complex environments.

We deployed our trained model on a Stretch RE1 robot which has a realsense camera and 2d lidar. We have tested the robot to go to a goal point in an unseen environment with narrow spaces. We initialized the robot at random locations and the robot was able to go to the goal point. We tested with static and dynamic obstacles.

4.2 Negative results and shortcomings

Despite promising results in simulation and controlled static environments, our deep reinforcement learning model might face challenges when tested in real-world scenarios with diverse dynamic obstacles. Due to time and resource constraints, we were unable to conduct extensive testing in multiple environments with varying obstacle configurations. This limitation is a common challenge in robotics research, where the real-world testing of autonomous systems is often constrained by factors such as logistical constraints.

The first algorithm that we trained using a TD3 model took 19 hours to train. This is partly due to low CPU power and as well as the prediction NN model being shallow. We can make the model converge faster by adding robot state and as well as more recurrent layers to add short-term memory. The team also plans on to change the loss function and optimizers to make the model converge faster. Moreover, the input observation data (point-cloud) can be down-sampled through scandots to save on compute time.

5 Summary and Conclusion

In conclusion, this project aimed to develop a more comprehensive and integrated approach to robot navigation that enables robots to operate efficiently in dynamic environments. The proposed learning-based approach using deep learning and reinforcement learning was explored to automate the local planner for a mobile robot. The results of the research indicate that the TD3 training method using 2D LiDAR data shows promising results with good Sim2Real ability. This research has significant implications for the field of robotics, particularly in developing adaptable and robust local planners for mobile robots, which can assist humans in carrying out dangerous and repetitive tasks in complex and dynamic environments. By improving the local planner's efficiency, this research contributes to advancing the capabilities of mobile robots and improving their functionality, leading to greater convenience and safety for people worldwide. Further improvements were made by adding memory to the model and sampling a fixed number of scandots from the LiDAR scan data.

6 Future directions

Our project has several avenues for future work to further improve the performance of the developed model. Firstly, we plan to switch to PyBullet or Habitat, to test the model in a more realistic and diverse environment. This will allow us to test the model's ability to navigate in different environments with varying obstacle configurations and improve the model's robustness to dynamic obstacles. Additionally, we plan to use Assistive Gym or Habitat Sim that provides a diverse set of tasks and configurable environments for robot learning. This will allow us to train the model on a wider range of tasks and environments, further improving its adaptability.

Secondly, we plan to include more noise models to depth data to make the model more robust to noisy and incomplete sensor data. This will help the model to better handle uncertainties in real-world environments and make it more reliable in different scenarios.

Thirdly, we plan to include head camera rotation in the pipeline as a stretch goal. This will enable the robot to have a wider field of view, which will help it to detect dynamic obstacles that may be outside of its current field of view.

7 Videos/Images of Results

As illustrated in Figure 2, the robot commences its navigation from the initial location and successfully reaches the goal, as depicted in Figure 3. Additional images and videos can be found on our website, the link to which is provided in the subsequent section.

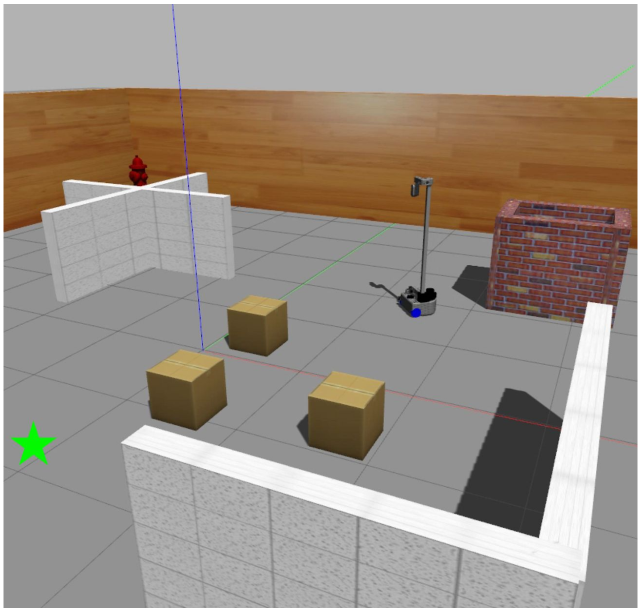

Figure 2: Robot at Start Location in Gazebo

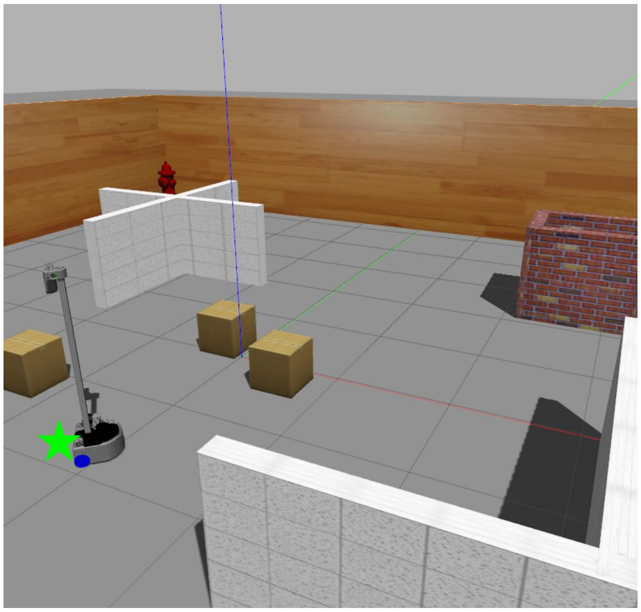

Figure 3: Robot at End Location in Gazebo

8 Link to code and website

Repository link: `https://github.com/VLR-Project-Team/VLR-Project`.

Website Link: `https://vlr-project-team.github.io/vision-rl-navigation/`.

References

[1] Ananye Agarwal, Ashish Kumar, Jitendra Malik, and Deepak Pathak. Legged locomotion in challenging terrains using egocentric vision, 2022.

[2] Somil Bansal, Varun Tolani, Saurabh Gupta, Jitendra Malik, and Claire Tomlin. Combining optimal control and learning for visual navigation in novel environments, 2019.

[3] Devendra Singh Chaplot, Dhiraj Gandhi, Abhinav Gupta, and Ruslan Salakhutdinov. Object goal navigation using goal-oriented semantic exploration, 2020.

[4] Devendra Singh Chaplot, Dhiraj Gandhi, Saurabh Gupta, Abhinav Gupta, and Ruslan Salakhutdinov. Learning to explore using active neural slam, 2020.

[5] Devendra Singh Chaplot, Dhiraj Prakashchand Gandhi, Abhinav Gupta, and Russ R Salakhutdinov. Object goal navigation using goal-oriented semantic exploration. *Advances in Neural Information Processing Systems*, 33:4247–4258, 2020.

[6] Reinis Cimurs, Il Hong Suh, and Jin Han Lee. Goal-driven autonomous exploration through deep reinforcement learning. *IEEE Robotics and Automation Letters*, 7(2):730–737, 2022.

[7] Zipeng Fu, Ashish Kumar, Ananye Agarwal, Haozhi Qi, Jitendra Malik, and Deepak Pathak. Coupling vision and proprioception for navigation of legged robots, 2021.

[8] Alessandro Giusti, Jérôme Guzzi, Dan C Cireşan, Fang-Lin He, Juan P Rodríguez, Flavio Fontana, Matthias Faessler, Christian Forster, Jürgen Schmidhuber, Gianni Di Caro, et al. A machine learning approach to visual perception of forest trails for mobile robots. *IEEE Robotics and Automation Letters*, 1(2):661–667, 2015.

[9] Saurabh Gupta, Varun Tolani, James Davidson, Sergey Levine, Rahul Sukthankar, and Jitendra Malik. Cognitive mapping and planning for visual navigation, 2017.