Learning-based Approach for Robot Motion Planning using Occupancy Grids

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

A. 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.

B. 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. [2] 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.

Another interesting work was by Gupta et. al [3] 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. [4] 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. [5] 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. [6] 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.

C. What is the 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 5x5 meters grid with an occupancy grid map as its input. The robot will use the map 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.

Furthermore, we will evaluate the performance of the proposed learning-based approach against traditional robotic algorithms. The evaluation will involve comparing the speed, accuracy, and adaptability of the two approaches under various scenarios, such as changing environments, obstacle avoidance, and real-time decision making.

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.

D. Why does your idea make sense intuitively?

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.

E. How does it relate to prior work in the area?

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. [7] 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[8] 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 transfer-ability.

F. Intermediate Results and Baselines

In this project, we focus on three baselines. Namely, a deep reinforcement learning (DRL) model [9] trained using Twin Delayed Deep Deterministic Policy Gradient [10]. This method purely uses visual data to make a differential-drive robot navigate in an environment with static obstacles. We also present two other methods that use classical Q-learning and Deep Q-Networks.

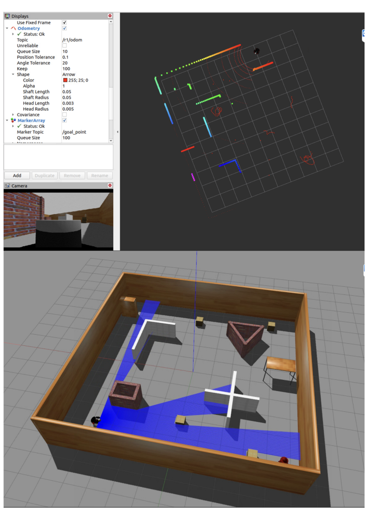

Figure 1. Twin Delayed Deep Deterministic Policy Gradient training run

The most promising result is obtained using the above-mentioned TD3 training method that takes in 3D LiDAR data and passes that as a state to the model. The model outputs a set of actions to be performed by the differential-drive robot. Our training process took about 18 hours on a GTX 1080 Ti GPU. This is partly due to the fact that 3D point-cloud data is very expensive to process on a CPU. The end-result however, still appears to be promising with good Sim2Real ability, as shown in figure 1. This can be improved upon by sampling a fixed number of scandots from the LiDAR scan data. The overall performance can also be improved upon by adding memory to the model by adding RNN layers and proprioception as an additional input.

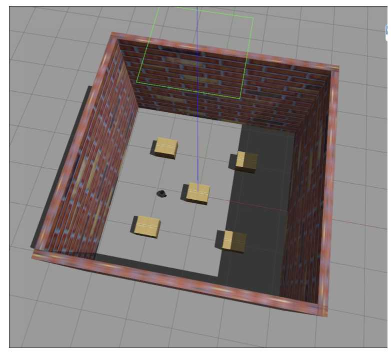

Figure 2. Deep Q-Network training episode

Our team also looked into classical RL training methods, Q-learning and Deep Q-networks. First observation made while training these models is that they train vision as a second-class citizen through creating a SLAM map and then perform policy optimization. Moreover, these models fail to generalize over different environment configurations and have poor Sim2Real ability. The training algorithm we developed, shown in figure 2, is much simpler for this case, which we shall use for our advanced algorithm.

G. Negative results and shortcomings

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 proprioception 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.

Regarding negative results, the differential-drive robot still managed to plan poor paths around dynamic obstacles as it did not account for past actions and motion data. This can be tackled by adding short-term memory, as discussed above.

H. Future Work and Timelines

We plan on to build upon our existing work with TD3 and adding short-term memory and changing the input data to use image depth instead of a 32-channel 3D LiDAR. This will help us considerably speed up training time while increasing the performance of the model. We shall also explore adding Proximal Policy Optimization (PPO) to improve stability. Overall, we can summarize the following tasks to be performed:

1. Improve simulation environment (Add more classes of obstacles)
2. Finalize dataset to be used (LiDAR versus stereo camera)
3. Finalize implementation of PPO to be used
4. Integrate PPO with the dataset and record the results

The corresponding timeline for the above tasks is as followed:

Target date	Task list
April 12th, 2023	Improve simulation environment
April 16th, 2023	Finalize dataset
April 20th, 2023	Finalize PPO implementation
April 23rd, 2023	Integrate PPO and run inference

Table 1. Timeline

References

[1] Ananye Agarwal, Ashish Kumar, Jitendra Malik, and Deepak Pathak. Legged locomotion in challenging terrains using egocentric vision, 2022. URL https://arxiv.org/abs/2211.07638. 1

[2] Zipeng Fu, Ashish Kumar, Ananye Agarwal, Haozhi Qi, Jitendra Malik, and Deepak Pathak. Coupling vision and proprioception for navigation of legged robots, 2021. URL https://arxiv.org/abs/2112.02094. 1

[3] Saurabh Gupta, Varun Tolani, James Davidson, Sergey Levine, Rahul Sukthankar, and Jitendra Malik. Cognitive mapping and planning for visual navigation, 2017. URL https://arxiv.org/abs/1702.03920. 1

[4] Devendra Singh Chaplot, Dhiraj Gandhi, Abhinav Gupta, and Ruslan Salakhutdinov. Object goal navigation using goal-oriented semantic exploration, 2020. URL https://arxiv.org/abs/2007.00643. 1

[5] Devendra Singh Chaplot, Dhiraj Gandhi, Saurabh Gupta, Abhinav Gupta, and Ruslan Salakhutdinov. Learning to explore using active neural slam, 2020. URL https://arxiv.org/abs/2004.05155. 1

[6] Somil Bansal, Varun Tolani, Saurabh Gupta, Jitendra Malik, and Claire Tomlin. Combining optimal control and learning for visual navigation in novel environments, 2019. URL https://arxiv.org/abs/1903.02531. 2

[7] 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. 3

[8] 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. 3

[9] 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. doi: 10.1109/LRA.2021.3133591. 3

[10] Scott Fujimoto, Herke Hoof, and David Meger. Addressing function approximation error in actor-critic methods. In *International conference on machine learning*, pages 1587–1596. PMLR, 2018. 3