Target Driven Indoor Navigation for Robots

Tariq Patanam, Eli Shayer, Younes Bensouda Mourri
CS 231N | Stanford University

Introduction

With the rise of autonomous vehicles, not only are self-driving applications promising but so too are indoor navigation problems. Deep reinforcement learning such as actor critic networks are currently being researched, but in this project we turn our attention to a relatively unexplored area: deep imitation learning. Instead of the agent making mistakes and learning through cost functions, imitation learning has an expert teach high value actions to the robot before it explores. Imitation learning is promising because it mimics a crucial aspect of the human learning process: teaching.

Objective

Our goal is to teach a robot to navigate indoors from any starting location to any target on the shortest possible path using only visual data.

Data Set and Problem Statement

We used the THOR dataset which, provides a framework to navigate several scenes. Each scene consists of a graph of locations within a room, along with the corresponding image observed from that location (see above). The state of the robot at a given time is its robot and its orientation, one of the four cardinal directions. For each state, the information available via THOR includes the visual input that the robot receives at that location as well as features of the images derived from ResNet. The goal, for a given starting location and ending location, is to efficiently move through the scene without making any collisions along the way. We seek to address this problem through imitation learning.

Methods

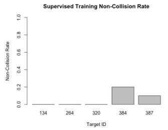
We implemented supervised imitation learning algorithm and the DAgger algorithm. The DAgger Algorithm goes as follows:

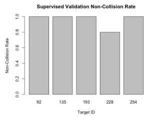
Initialize $\mathcal{D} \leftarrow \emptyset$.
Initialize $\hat{\pi}_1$ to any policy in Π.
for $i = 1$ **to** N **do**
  Let $\pi_i = \beta_i \pi^* + (1 - \beta_i)\hat{\pi}_i$.
  Sample T-step trajectories using π_i.
  Get dataset $\mathcal{D}_i = \{(s, \pi^*(s))\}$ of visited states by π_i and actions given by expert.
  Aggregate datasets: $\mathcal{D} \leftarrow \mathcal{D} \bigcup \mathcal{D}_i$.
  Train classifier $\hat{\pi}_{i+1}$ on $\mathcal{D}$ (or use online learner to get $\hat{\pi}_{i+1}$ given new data $\mathcal{D}_i$).
end for
Return best $\hat{\pi}_i$ on validation.

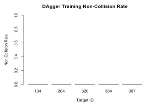
In other words, at every step we follow the expert policy with a certain probability and the trained policy with a certain probability. The rate at which we use the expert policy decays over time, as we train the policy on an increasingly large dataset aggregated over the training iterations.
In the supervised imitation algorithm, we use the expert policy to train only on the locations that would ideally be chosen by the robot, which is an optimistic . This

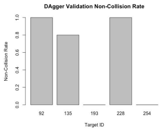
Results and Analysis

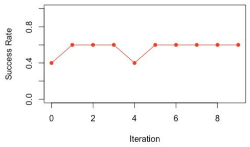
We expect the DAgger Algorithm to outperform the supervised learning algorithm because supervised learning is unable to recover from steps off of the expert policy steps that it learned.

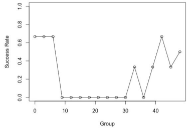


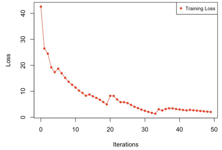












Future Work

We can considerably improve our success rates by changing the architecture of our neural network to a pre-trained ResNet instead of our current four layer fully connected model. We can further improve our success rate through retraining parts of the ResNet model, implementing SIFTDagger and by training on more scenes and targets.

References

- Stephane Ross, Geoffrey J. Gordon, and J. Andrew Bagnell. A Reduction of Imitation Learning and Structured Prediction to No-Regret Online Learning 2010.
https://www.cs.cmu.edu/sross1/publications/Ross-AIStats11-NoRegret.pdf
- Zhu et. al. Target-driven Visual Navigation in Indoor Scenes using Deep Reinforcement Learning 2016. http://stanford.edu/~yukez/papers/icra2017.pdf
- Jonathan Ho, Stefano Ermon. Generative Adversarial Imitation Learning 2016. https://arxiv.org/pdf/1606.03476.pdf
- Stefan Schaal. Is imitation learning the route to humanoid robots?. http://www.bcp.psych.ualberta.ca/mike/PearlStreet/PSYCO354/pdfstuff/Readings/Schaal1.pdf