

HyperDrive: Autonomous Self Driving Car in an Urban Setting using Deep Reinforcement Learning

Project Team

Sana Haider p18-0011
Syed Asad Zaman p18-0034

Session 2018-2022

Supervised by

Mr. Waqas Ali



Department of Computer Science

**National University of Computer and Emerging Sciences
Peshawar, Pakistan**

December, 2021

Student's Declaration

We declare that this project titled “*HyperDrive: Autonomous Self Driving Car in an Urban Setting using Deep Reinforcement Learning*”, submitted as requirement for the award of degree of Bachelors in Computer Science, does not contain any material previously submitted for a degree in any university; and that to the best of our knowledge, it does not contain any materials previously published or written by another person except where due reference is made in the text.

We understand that the management of Department of Computer Science, National University of Computer and Emerging Sciences, has a zero tolerance policy towards plagiarism. Therefore, We, as authors of the above-mentioned thesis, solemnly declare that no portion of our thesis has been plagiarized and any material used in the thesis from other sources is properly referenced.

We further understand that if we are found guilty of any form of plagiarism in the thesis work even after graduation, the University reserves the right to revoke our BS degree.

Sana Haider

Signature: _____

Syed Asad Zaman

Signature: _____

Verified by Plagiarism Cell Officer

Dated:

Certificate of Approval



The Department of Computer Science, National University of Computer and Emerging Sciences, accepts this thesis titled *HyperDrive: Autonomous Self Driving Car in an Urban Setting using Deep Reinforcement Learning*, submitted by Sana Haider (p18-0011), and Syed Asad Zaman (p18-0034), in its current form, and it is satisfying the dissertation requirements for the award of Bachelors Degree in Computer Science.

Supervisor

Mr. Waqas Ali

Signature: _____

Ms. Mashal Khan

FYP Coordinator

National University of Computer and Emerging Sciences, Peshawar

Dr. Hafeez Ur Rehman

HoD of Department of Computer Science

National University of Computer and Emerging Sciences

Acknowledgements

Your abstract here

Sana Haider
Syed Asad Zaman

Abstract

Your abstract here

Contents

- 1 Introduction 1**
 - 1.1 my section 1
 - 1.1.1 my subsection 2
- 2 Review of Literature 7**
- 3 Review of Literature 9**
 - 3.1 Reinforcement Learning 9
- 4 Conclusions and Future Work 13**
- References 15**

List of Figures

List of Tables

Chapter 1

Introduction

The chapter introduces the basic concepts of blah blah blahThe chapter introduces the
basic concepts of blah blah blahThe chapter introduces the basic concepts of blah blah
blahThe chapter introduces the basic concepts of blah blah blahThe chapter introduces
the basic concepts of blah blah blahThe chapter introduces the basic concepts of blah
blah blahThe chapter introduces the basic concepts of blah blah blahThe chapter intro-
duces the basic concepts of blah blah blahThe chapter introduces the basic concepts of
blah blah blah. [?]

1.1 my section

[illegible]

dfik alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslk-
fjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsd-
flk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslk-
fjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsd-
flk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkf-
jalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsd-
flk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk

1.1.1 my subsection

alakkjsf alksdjf lakssjfd lajfd lkasjf dlkasj sfdlkajs fldkjasldf jlaskdfj alkjaslkfjalkssjflkajs-
dfik alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkf-
jalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsd-
flk alkkjsf alksdjf lakssjfd lajfd lkasjf dlkasj sfdlkajs fldkjasldf jlaskdfj alkjaslkfjalkssjflka-
jsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslk-
fjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajs-
dfik alkkjsf alksdjf lakssjfd lajfd lkasjf dlkasj sfdlkajs fldkjasldf jlaskdfj alkjaslkfjalkssjflka-
jsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslk-
fjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajs-
dfik alkkjsf alksdjf lakssjfd lajfd lkasjf dlkasj sfdlkajs fldkjasldf jlaskdfj alkjaslkfjalkssjflka-
jsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslk-
fjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajs-
dfik alkkjsf alksdjf lakssjfd lajfd lkasjf dlkasj sfdlkajs fldkjasldf jlaskdfj alkjaslkfjalkssjflka-
jsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslk-
fjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajs-

dflk alkkjsf alksdjf lakssjfd lajfd lkasjf dlkasj sfdlkajs fldkjasldf jlaskdfj alkjaslkfjalkssjflka-
jsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslk-
fjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkaj-
sdflk alkkjsf alksdjf lakssjfd lajfd lkasjf dlkasj sfdlkajs fldkjasldf jlaskdfj alkjaslkfjalkssjflka-
jsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslk-
fjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkaj-
sdflk

alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflka-
jdsflk alkkjsf alksdjf lakssjfd lajfd lkasjf dlkasj sfdlkajs fldkjasldf jlaskdjf alkjaslkfjalkssjflka-
jdsflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslk-
fjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsd-
dflk alkkjsf alksdjf lakssjfd lajfd lkasjf dlkasj sfdlkajs fldkjasldf jlaskdjf alkjaslkfjalkssjflka-
jdsflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslk-
fjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsd-
dflk alkkjsf alksdjf lakssjfd lajfd lkasjf dlkasj sfdlkajs fldkjasldf jlaskdjf alkjaslkfjalkssjflka-
jdsflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslk-
fjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsd-
dflk

[illegible]

alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflkajsdflk alkjaslkfjalkssjflka-

Chapter 2

Review of Literature

This chapter provides a comprehensive literature review of the related work to blah blah

Chapter 3

Review of Literature

In the past few years, deep reinforcement learning become a very focus topic due to its success in games like Atari games, GO, and chess. Especially after the success of the project AlphaGO[4] in 2016 and AlphaZero[3] in 2017 by Google DeepMind reinforcement learning become a hot topic. Besides, a fairly recent project known as AlphaFold[1] is used to predict the protein structure has opened doors for reinforcement learning to be applied on more complex real-world problems like goal-oriented autonomous self-driving in a complex dynamic dense environment that includes route planning also. From the review of the following research papers, the authors have concluded that there is not any neural network or reinforcement architecture modeled in CARLA yet that has addressed the problem of route planning in a dynamic dense urban environment

3.1 Reinforcement Learning

When we talked about reinforcement learning it somehow works like human psychology i.e, learning from previous experiences and taking the future action based upon those experiences. In the past, RL was limited to the domain where the features are given to the RL model which was handcrafted or had a low dimensional search space. Due to this limitation, RL wasn't able to be applied to a problem where the search space was limited. For that purpose, a novel artificial agent was designed which was named as **deep**

Q-network (DQN)[2]. This agent has trained on Atari 2600 games. The only input that was provided to this agent was pixels of the game played and the score input. Surprisingly, the agent was able to achieve a professional level of success. DQN was a neural network specifically a deep neural network combined with reinforcement learning whose goal was to maximize the future reward. The limitation of DQN (mentioned in the paper) was that it was limited to Atari games only and didn't apply to games with higher search and state space and to 3D games.

DQN was able to achieve some notable success when it comes to the games where the search space of the possible moves (b^d b is game's breadth and d is depth (game length)) was limited but when it comes to the games where the search space is pretty much large for example Chess ($b = 35, d = 80$) and especially Go ($b = 250, d = 150$) than typical reinforcement learning techniques become exhausted. Another difficulty is in the evaluation of board positions and moves. So to solve the problem of effective search space two general principles can be used:

- Depth of the search can be reduced by position evaluation truncating the tree at state s and replacing the subtree below s .
- Breadth of the search may be reduced by sampling actions from a policy $p(a|s)$ that is a probability distribution over possible moves a in position s .

The approach used for **AlphaGo**[4] was that a deep neural network was trained on the novel combination of supervised learning from human expert games and for self-play games, reinforcement learning was used. The tree search algorithm that was used in **AlphaGo** was **Monte Carlo Tree Search (MCTS)**. AlphaGo was able to achieve the groundbreaking success of a **99.8% winning rate**. By defeating the European Champion by **5-0** and worldwide champion by **4-1**. This was the first time in the history of a full-sized game of Go that a machine was able to defeat human experts which were almost thought impossible at least for a decade. The main achievement of AlphaGo was that it has provided the solution for the games with higher search space using DQN and MCTS, which open doors for solving much more complex real world problem like autonomous self-driving in a complex environment. The limitation of this project was that AlphaGo

was taking too much time while making decision which make it sometimes inefficient.

With the advancement in AI, It is now heading to a place where an A. I agent capable of learning from themselves i.e, learn from its past experiences. Just like a small kid when he/she tried to walk for the first time. At first, it stumbles a lot but a point came where it learn how to walk without stumbling. The same thing is mimicked in RL-based agents like AlphaZero. **AlphaZero[3]** is a fully generic algorithm which achieved notable success in games like **Chess, Go & Shogi**. Unlike, AlphaGo where the agent has to create a tree, evaluate board positions, and then solve the problem using the MCTS. AlphaZero, play the game just by knowing the rules of the game it doesn't require any additional domain knowledge using. And the interesting part is that it achieves this by doing self-play and reinforcement learning only. This shows that general-purpose RL can achieve a **superhuman** level performance in much more complex real-world problems.

Chapter 4

Conclusions and Future Work

conclusions here

Bibliography

- [1] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.
- [2] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *nature*, 518(7540):529–533, 2015.
- [3] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, 362(6419):1140–1144, 2018.
- [4] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. Mastering the game of go without human knowledge. *nature*, 550(7676):354–359, 2017.