AUTONOMOUS NAVIGATION OF MOBILE ROBOT IN A DYNAMIC ENVIRONMENT USING DEEP REINFORCEMENT LEARNING

Dhivyadharshan Seetharaman

208814P

Master of Science in Industrial Automation

Department of Electrical Engineering

Engineering Faculty

University of Moratuwa

Sri Lanka

March 2023

AUTONOMOUS NAVIGATION OF MOBILE ROBOT IN A DYNAMIC ENVIRONMENT USING DEEP REINFORCEMENT LEARNING

Dhivyadharshan Seetharaman

208814P

Thesis submitted in partial fulfillment of the requirements for the degree Master of Science in Industrial Automation

Department of Electrical Engineering

Engineering Faculty

University of Moratuwa

Sri Lanka

March 2023

DECLARATION

I declare that this is my own work, and this thesis does not incorporate without Acknowledgment of any material previously submitted for a Degree or Diploma in any other University or institute of higher learning, and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgment is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date:

The above candidate has carried out research for the Master's thesis under my supervision. I confirm that the declaration made above by the student is true and correct.

Prof. Chandima D. Pathirana

Signature of the Supervisor: Date:

DEDICATION

I would like to dedicate this thesis to the people who aided me when I stumbled and lost my way. Especially, My parents and sister, school teachers, university professors, and lecturers.

ACKNOWLEDGMENTS

I would like to offer my sincerest gratitude and heartfelt indebtedness to my supervisor Prof.Chandima D.Pathirana, who has given me the opportunity to pursue my second-year research of MSC in Industrial Automation under his supervision and guidance. I still remember the day I joined the University of Moratuwa for my post-graduation with many knowledge gaps in my field. My supervisor, Prof.Chandima D.Pathirana, put endless efforts and deep patience into answering my numerous basic questions to fill my knowledge gaps and provided clarity on subjects by explaining the topic very deeply in this taught course. I am really obliged to him for putting additional effort into explaining countless topics in the subjects of Electrical Machines and Drives, Industrial Electronics, Control Systems Designs, and Mechatronics during my postgraduate diploma, which I missed gaining knowledge during my undergraduate period. Especially he guided me to approach the research problem in a step-by-step manner when I was trying to solve the problem with a quantum leap, and his counseling showed me the right direction when I was on the wrong way up. This research would have never been possible without his depth knowledge and proper guidance. His advice helped me approach the complex problem by dividing it into small problems in order to successfully accomplish the goal of this research within a given period.

I want to thank Prof. A.G. Buddhika P. Jayasekara, who untangled me when I was a jam at a randomization problem of python code. His advice helped me avoid getting stuck in local minima and confidently reach the global minima in my research. I would also like to thank my progress review panel members, Dr.A.M.H.S. Abeykoon and Dr.Velmanickam Logeeshan for their penetrating comments during progress review presentations which incentivized me to broaden my research and reshape it to generate valuable and sensible outputs. My special thanks go to Dr.Velmanickam Logeeshan for his friendly advice and guidance, which helped me enormously to mitigate psychological barriers during my research. I am sincerely obliged to the Department of Electrical Engineering for allowing me to access the required number of computers in the computer lab to train my Deep Reinforcement learning model for a more extended period during this research. I would like to thank all robotics lab-mates of the University of Moratuwa for their instructions and encouragement, special thanks to H.M.Ravindu T.Bandara, who has taught me significantly about the Deep Neural Networks and the deployment of the machine learning model.

Finally, no words can ever be strong enough to express my gratitude to my parents and my sister because of their great hope on me, which constantly stimulates me to take at least a few steps toward my goals. I would also like to thank God for giving me spiritual support in following my career path when I stumble.

ABSTRACT

Autonomous navigation of a mobile robot in a dynamic environment is a highly challenging application because the path to the goal frequently changes due to unpredictable movements of humans with different velocities. Deep Reinforcement completely trains a model in a simulator using a trial-and-error technique by exploring and collecting required data automatically and cheaply from the customized environment. This research develops a customized environment with robot and pedestrian models in OpenAI Gym, replicating humans' real-world measurements and motion patterns. A mathematical model was developed to encapsulate the navigation norms of humans' to teach the robot about socially compliant routes using a reward function in order to smooth the robot's navigation in a dynamic environment. A recently evolved algorithm, H-PPO, has been selected to train the model by considering the agent's hybrid action space, which consists of discrete actions parametrized by continuous values.

First, the model failed to learn due over fit in a simple environment, and then it learned when the task was randomized. The various approaches have been investigated to enhance the model's generalizability as much as possible in the simulator. Finally, the agent is trained in each environment separately. Despite this research has not considered the complex scenario as randomizing the whole environment initially, the developed model was able to scrutinize the performance of the recently evolved algorithm H-PPO in obstacle avoidance applications, and the developed model can learn obstacle avoidance in a dynamic environment by respecting social norms in the long-range motion for laboratory application. However, the success rate of the model trained later in a fully randomized 3-pedestrians environment was 86.67% out of 30 episodes of testing, which is higher than the previous research [1]. Further investigation has to be carried out in future work by adding memory ability to the model in order to enhance the performance, reduce the training time and mitigate the performance collapse during the learning phase.

Keywords: **Intelligent Mobile Robot, Navigation, OpenAI Gym, Deep Neural Network, Deep Reinforcement Learning, H-PPO algorithm, Machine Learning, Parametrized Action Space.**

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

API	Application Programming Interface
A3C	Asynchronous Advantage Actor-Critic Learning
CNN	Convolutional Neural Network
CADRL	Collision Avoidance with Deep Reinforcement Learning
DRL	Deep Reinforcement Learning
DNN	Deep Neural Network
DDPG	Deep Deterministic Policy Gradient
GNRON	Goal Not Reachable when Obstacles are Nearby
GCN	Graph Convolutional Network
GOSELO	Goal-Directed Obstacle and Self-Location Map for Robot Navigation using Reactive Neural Networks
GRU	Gated Recurrent Unit
H-PPO	Hybrid Proximal Policy Optimization
HyAR	Hybrid Action Representation
LiDAR	Light Detection and Ranging
LSTM	Long Short-Term Memory
MP-DQN	Multi-Pass Deep Q-Networks
MLP	Multi-Layer Perceptron
PPO	Proximal Policy Optimization
P-DQN	Paramterirzed Deep Q-Networks
PAMDPs	Parametrized Action Markov Decision Processes
RL	Reinforcement Learning
SLAM	Simultaneous Localization and Mapping
SOADRL	Social Obstacle Avoidance using Deep Reinforcement Learning
SA-CADRL	Socially Aware Collision Avoidance with Deep Reinforcement
TD	Temporal Difference

NOMENCLATURE

$R(\tau)$	Total Discounted Reward
T	Horizon (Episode Length)
θ	Heading angle of the robot
x_r , y_r	Co-ordinates of the robot
v_r	Speed of the robot
r	The outer radius of the robot
d_G	The distance between the center point of the robot and the goal
p_{ix} , p_{iy}	Co-ordinates of pedestrians
v_p	Constant speed of pedestrians
θ_i	Heading angle of pedestrians
d_{ci}	Collision distance between the robot and pedestrians
θ_{Rn}	Robot's norm-inducing angle
θ_{Pin}	Pedestrians' norm-inducing angle
a_t	The complete action
a	The chosen discrete action
x_a	The chosen continuous parameters for discrete action a
$\pi_{\theta_{old}}$	The old policy
$A_t^{\pi_{\theta_{old}}}$	Advantage function calculated by the old policy
ϵ	The clip ratio

LIST OF APPENDICES