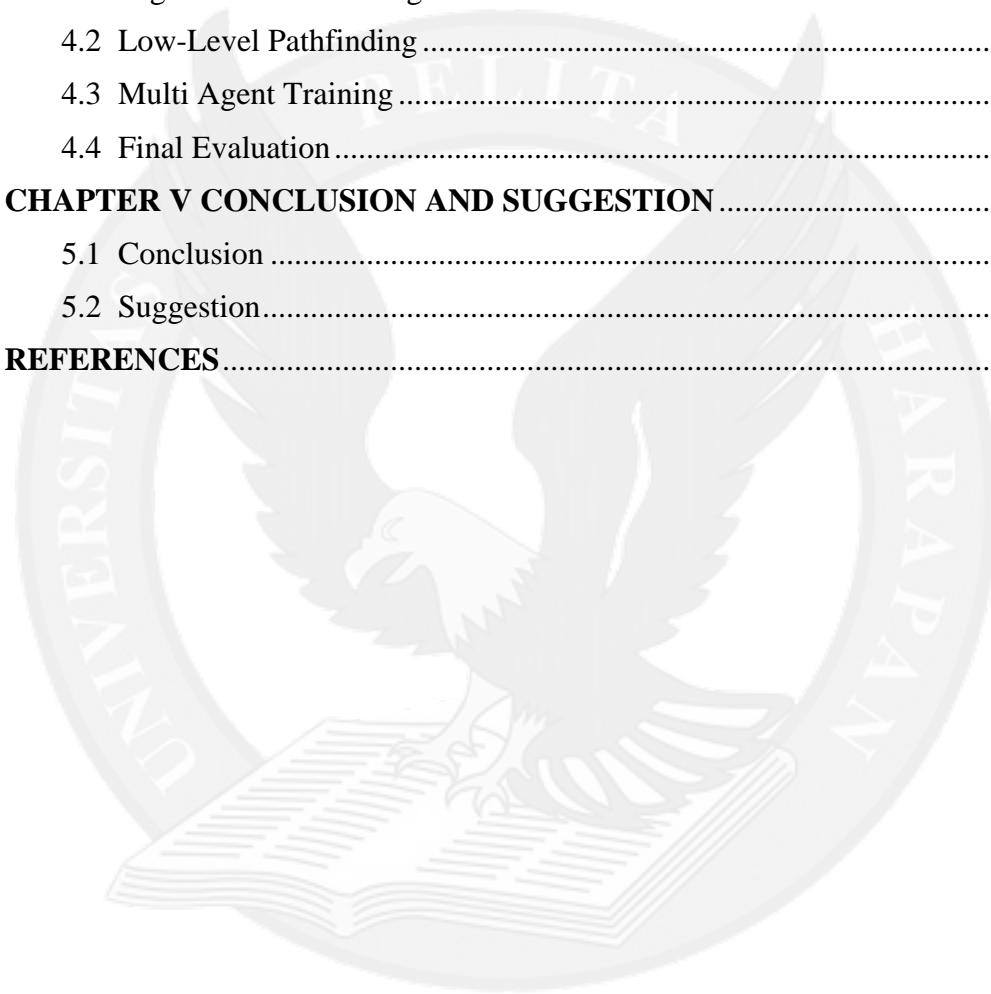


# TABLE OF CONTENTS

<b>COVER</b>	
<b>ORIGINALITY STATEMENT OF THE THESIS</b>	
<b>THESIS SUPERVISOR'S APPROVAL</b>	
<b>THESIS EXAMINATION COMMITTEE</b>	
<b>ABSTRACT</b> .....	v
<b>ABSTRAK</b> .....	v
<b>ACKNOWLEDGEMENTS</b> .....	vii
<b>TABLE OF CONTENTS</b> .....	ix
<b>LIST OF FIGURES</b> .....	xi
<b>LIST OF EQUATIONS</b> .....	xii
<b>LIST OF ABBREVIATIONS</b> .....	xiii
<b>CHAPTER I INTRODUCTION</b> .....	1
1.1 Background.....	1
1.2 Problem Identification .....	5
1.3 Problem Limitations.....	7
1.4 Problem Definition.....	8
1.5 Research Purpose .....	8
1.6 Outline of the Thesis .....	9
<b>CHAPTER II THEORITICAL REVIEW</b> .....	12
2.1 AGV and AMR.....	12
2.2 Warehouse Layouts.....	14
2.3 Centralized and Decentralized Task Assignments.....	14
2.3 Multi Agent Pathfinding Problems .....	15
2.4 Conflict-Based Search .....	18
2.5 Reinforcement Learning .....	21
2.6 Multi Agent Reinforcement Learning.....	25
<b>CHAPTER III RESEARCH METHODOLOGY</b> .....	31
3.1 Overview .....	31
3.2 High-Level Pathfinding .....	33
3.2.1 Warehouse Layout Creation .....	34
3.2.2 CBS Pathfinding.....	34

3.3 Low-Level Pathfinding .....	36
3.3.1 PyGame Environment Creation.....	36
3.3.2 OpenAI Gym Environment Conversion.....	38
3.3.3 Model Training.....	38
3.4 Multi Agent Training .....	41
3.5 Evaluation and Comparison.....	41
<b>CHAPTER IV RESULTS AND DISCUSSION .....</b>	<b>43</b>
4.1 High-Level Pathfinding .....	43
4.2 Low-Level Pathfinding .....	44
4.3 Multi Agent Training .....	46
4.4 Final Evaluation .....	47
<b>CHAPTER V CONCLUSION AND SUGGESTION .....</b>	<b>52</b>
5.1 Conclusion .....	52
5.2 Suggestion.....	52
<b>REFERENCES.....</b>	<b>54</b>



## LIST OF FIGURES

Figure 1.1 Forklift AGV with Stabilizer Pad .....	1
Figure 1.2 Centralized approach and decentralized approach .....	4
Figure 2.1 Car and train analogy as AMR and AGV .....	13
Figure 2.2 Visual representation of conflict interaction.....	18
Figure 2.3 Path planning concept.....	25
Figure 2.4 MAPPO workflow diagram.....	26
Figure 2.5 Average returns comparing PS-TRPO and DQN/DDPG .....	30
Figure 3.1 Block diagram of the simulation process .....	31
Figure 3.2 Warehouse layout example.....	34
Figure 3.3 Conflict detection example.....	35
Figure 3.4 Specialized track for single agent training.....	37
Figure 3.5 Radar placement on the agent.....	37
Figure 3.6 Multi-layer Perceptron data .....	39
Figure 3.7 Network structure for both actor and critic network .....	40
Figure 4.1 Extracted image examples .....	43
Figure 4.2 Agent starting, goal, and direction data .....	43
Figure 4.3 Final training log data.....	44
Figure 4.4 Agent visualization snippets .....	45
Figure 4.5 Agent positions in multi agent environment.....	46
Figure 4.6 Agents are moving in the track.....	46
Figure 4.7 Agents not occupying the same space .....	47
Figure 4.8 Box plot for high-level task data .....	48
Figure 4.9 Box plot for comparison task data.....	49
Figure 4.10 Box plot for low-level task data .....	50
Figure 4.11 Data comparison of proposed model and base CBS model.....	51

## LIST OF EQUATIONS

Equation 1 Markov decision process expected return formula.....	20
Equation 2 MAPPO critic value function .....	25
Equation 3 Generalized Advantage Estimation function .....	25
Equation 4 MAPPO actor policy value function.....	25



## LIST OF ABBREVIATIONS

AGV	Automated Guided Vehicle
AMR	Autonomous Mobile Robot
CBS	Conflict-Based Search
RL	Reinforcement Learning
MARL	Multi Agent Reinforcement Learning
MAPF	Multi Agent Pathfinding
ICTS	Increasing Cost Tree Search
EPEA*	Enhanced Partial Expansion A*
PPO	Proximal Policy Optimization
MAPPO	Multi Agent Proximal Policy Optimization
DDPG	Deep Deterministic Policy Gradient
DQN	Deep Q Network
TRPO	Trust Region Policy Optimization
MDP	Markov Decision Process
GAE	Generalized Advantage Estimation