MODELING PEDESTRIAN BEHAVIOR IN PEDESTRIAN-VEHICLE NEAR MISSES USING INVERSE REINFORCEMENT LEARNING ALGORITHMS

by

Payam Nasernejad

B.Sc., Sharif University of Technology, 2018

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Civil Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2021

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Modeling Pedestrian Behavior in Pedestrian-Vehicle Near Misses Using Inverse Reinforcement Learning Algorithms

submitted by Payam Nasernejad in partial fulfillment of the requirements for

the degree of Master of Applied Science

in Civil Engineering

Examining Committee:

Dr. Tarek Sayed, Professor, Department of Civil Engineering, UBC
Supervisor

Dr. Paul de Leur, Manager of Road Safety Programs at Insurance Corporation of British Columbia
Supervisory Committee Member

Abstract

Using simulation models to conduct safety assessments can have several advantages as it enables the evaluation of the safety of various design and traffic management options before actually making changes. However, limited studies have developed microsimulation models for the safety evaluation of active road users such as pedestrians. This can be attributed to the limited ability of the existing simulation models to capture the heterogeneity in pedestrian behavior and their complex collision avoidance mechanisms. Therefore, the objective of this thesis is to develop an agent-based framework to realistically model pedestrian behavior in near misses and to improve the understanding of pedestrian evasive action mechanisms in interactions with vehicles. Pedestrian-vehicle conflicts are modeled using single-agent and multi-agent approaches under the Markov Decision Process (MDP) and Markov Games (MG) frameworks, respectively. A continuous Gaussian Process Inverse Reinforcement Learning (GPIRL) approach is implemented to recover pedestrians' single-agent reward functions and infer their collision avoidance mechanisms in conflict situations. In the multi-agent framework, pedestrian-vehicle conflicts are modeled utilizing the Multi-Agent Adversarial Inverse Reinforcement Learning (MA-AIRL). Video data from a congested intersection in Shanghai, China is used as a case study. Trajectories of pedestrians and vehicles involved in traffic conflicts were extracted with computer vision algorithms. A Deep Reinforcement Learning (DRL) model is used to estimate optimal pedestrian single-agent policies in traffic conflicts. Moreover, the adversarial multi-agent IRL approach simulates road users' optimum evasive actions with an implementation of the multi-agent actor-critic using Kronecker-factored trust region (MACK). This algorithm enables multi-agent policy estimation using the rewards recovered from the discriminator of the adversarial neural network. The results show that the developed models predicted pedestrian trajectories and their evasive

action mechanisms (i.e., swerving maneuver and speed changing) in conflict situations with high accuracy. Moreover, the highly nonlinear structure of the reward function in the multi-agent framework enabled capturing more complex behavior of the road users in near misses and their collision avoidance mechanisms. This study is a crucial step in developing a safety-oriented microsimulation tool for pedestrians in mixed traffic conditions.

Lay Summary

Pedestrians could face hazardous situations with an elevated risk of injuries in interactions with motorized traffic. Understanding the important elements of safety from pedestrians' perspective and analyzing their reactions to oncoming danger can assist in creating a safer environment, particularly in dangerous accident-prone locations like intersections. Despite the importance of evaluating pedestrian safety in traffic conflicts with motorized traffic, the majority of the existing research has focused on driver maneuvers in microsimulation models. This could be attributed to the complexity of pedestrians' behavior in conflict situations, which can not be appropriately simulated with conventional models. This thesis presents two novel Inverse Reinforcement Learning (IRL) frameworks that enable understanding the underlying reasons and intentions behind the pedestrians' adopted strategies in conflicts with vehicles. Moreover, two novel single-agent and multi-agent simulation frameworks were developed to simulate the pedestrian behavior in interaction with conflicting vehicles.

Preface

Portions of the introduction in chapter 1, portions of the literature review in chapter 2, portions of data collection and video analysis in chapter 3, and a version of chapter 4 have been published [Nasernejad, P., Sayed, T., & Alsaleh, R. (2021). Modeling pedestrian behavior in pedestrian-vehicle near misses: A continuous Gaussian Process Inverse Reinforcement Learning (GP-IRL) approach. Accident Analysis & Prevention, 161, 106355.]. I was responsible for the investigation, analysis, data preparation, visualization, and writing the original draft of the manuscript. Dr. Sayed T. was the supervisory author involved in the conceptualization, methodology, data collection, investigation, and validation of the project. Alsaleh R. was involved in the software coding, conceptualization, methodology, investigation, and manuscript revision.

Portions of the introduction in chapter 1, portions of the literature review in chapter 2, portions of data collection and video analysis in chapter 3, and a version of chapter 5 have been submitted for publication. I was responsible for the investigation, analysis, data preparation, visualization, and writing the original draft of the manuscript. Dr. Sayed T. was the supervisory author involved in the conceptualization, methodology, data collection, investigation, and validation of the project. Alsaleh R. was involved in the software coding, conceptualization, methodology, investigation, and manuscript revision.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACKTR	Actor-Critic Using Kronecker-Factored Trust Region
AI	Artificial Intelligence
AIRL	Adversarial Inverse Reinforcement Learning
AV	Autonomous Vehicles
CA	Cellular Automata
CNN	Convolutional Neural Neatwork
CV	Conventional Vehicles
DBN	Dynamic Bayesian Networks
DRL	Deep Reinforcement Learning
ELU	Exponential Linear Unit
FV	Fixed Videography
GA	Genetic Algorithm
GAIL	Generative Adversarial Imitation Learning
GCL	Guided Cost Learning
GDP	Global Domestic Product
GHG	Greenhouse Gas
GP	Gaussian Process
GPIRL	Gaussian Process Inverse Reinforcement Learning
HD	Hausdorff Distance
HSFM	Headed Social Force Model
IMV	In-Motion Videography
K-FAC	Kronecker-Factored Approximate Curvature
L-BFGS	Limited memory Broyden-Fletcher-Goldfarb-Shanno
LSBRE	Logistic Stochastic Best Response Equilibrium
MA-AIRL	Multi-Agent Adversarial Inverse Reinforcement Learning
MACK	Multi-Agent Actor-Critic Using Kronecker-Factored Trust Region
MAE	Mean Absolute Error
MaxEnt IRL	Maximum Entropy Inverse Reinforcement Learning

MC	Marked Crosswalk
MDP	Markov Decision Process
MG	Markov Games
MSSF	Maximum Slope of Step Frequency
MV	Motorized vehicles
NGSIM	Next Generation Simulation
NMV	Nonmotorized vehicles
PARAMICS	Parallel Microscopic Simulation (Traffic Microsimulation Model)
PET	Post-Encroachment Time
PFG	Pedestrian Flashing Green
PGA	Pedestrian Gap Acceptance
PPF	Pedestrian Passing First
RTTC	Relative Time To Collision
SDG	Sustainable Development Goals
SFM	Social Force Model
SSA	Singular Spectrum Analysis
SSAM	Surrogate Safety Assessment Model
STCT	Swedish Traffic Conflict Technique
TCT	Traffic Conflict Technique
TTC	Time To Collision
UAV	Unmanned Aerial Vehicle
UMC	Unmarked Crosswalk
UN	United Nations
VISSIM	Verkehr In Städten - SIMulationsmodell" (German for "Traffic in cities - simulation model") (Traffic Microsimulation Model)
VPF	Vehicle Passing First
VY	Vehicle Yielding
WHO	World Health Organization

Acknowledgements

I would like to extend my appreciation to my supervisor Dr. Tarek Sayed, who has been my role model and has gone above and beyond to guide, support, and encourage me throughout this journey. His commitment and ethics have been an inspiration not only in my research but also to the individual that I would like to become. His patience and care have been heartwarming towards my progress in my research while being isolated during the pandemic. I would also like to thank my colleague from the transportation research group Rushdi Alsaleh for helping me in many aspects of this thesis and advising me during my projects.

Dedication

To my Parents

Chapter 1: Introduction

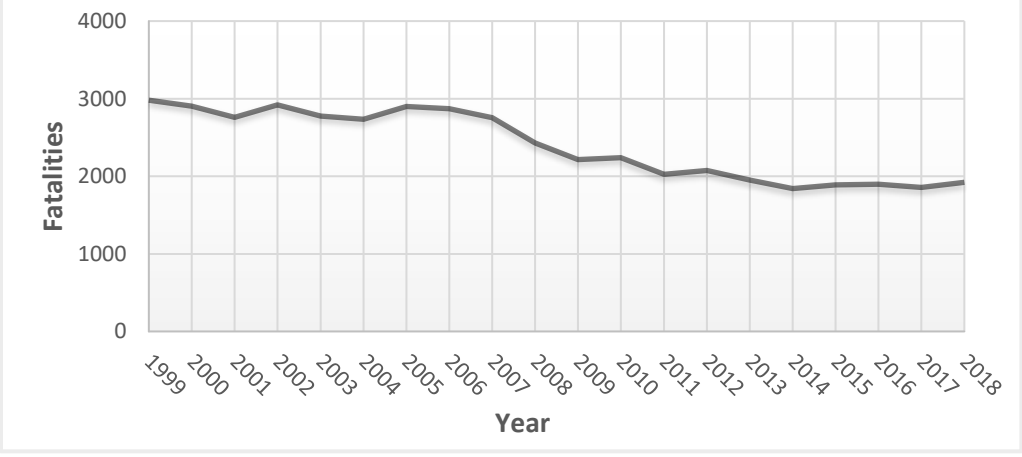
This chapter consists of four sections. The first section covers the background information of traffic safety problems and the importance of safety analysis. This section also describes the crash-based safety analysis and provides information on the surrogate safety measures, automated video analysis techniques, and traffic simulation models. The second section explains the problems and the gaps in safety studies that this thesis attempted to address. The third section presents the main objectives of this study, and the last section describes the thesis structure.

1.1 Background

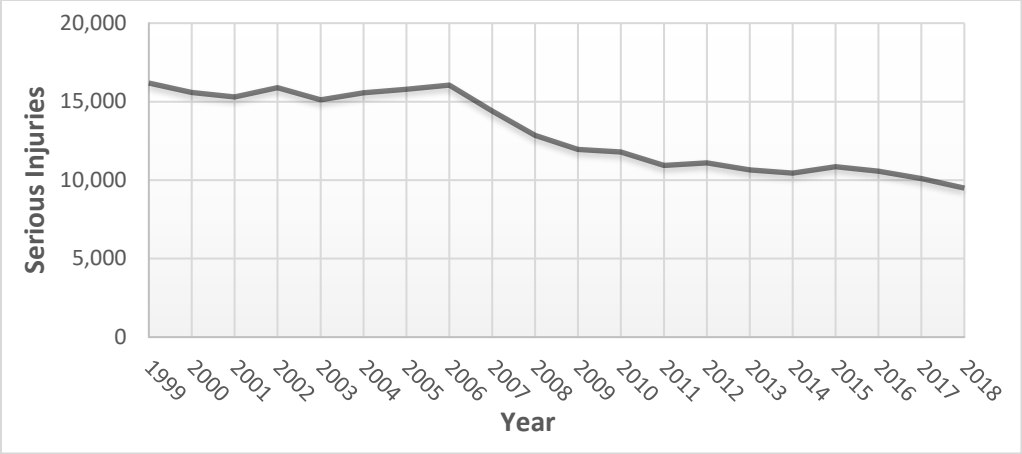
Road traffic collision is one of the leading causes of death among adults (WHO, 2018). Traffic fatalities and injuries are only direct consequences of traffic collisions for accident participants; moreover, family, friends, and the social sphere of the participants could also be impacted. Financial consequences, mental and physiological burdens will remain for many years after a collision. Moreover, long-lasting adverse effects of road crashes can impact the community, economically and psychologically. In the global status road safety report published by World Health Organization (WHO) (WHO, 2018), an annual estimate of 1.35 million traffic death was reported. Moreover, road traffic injuries is predicted to become the fifth leading cause of death by 2030 (WHO, 2018). Currently, collisions are the eighth leading cause of death worldwide for all age groups. Furthermore, the leading cause of death between the age ranges of 5 to 29 years is due to traffic injuries (WHO, 2018). Traffic safety improvement and collision reduction are far from the targets of the Sustainable Development Goals (SDG). According to SDG plans developed in September 2015 (SDG, 2015), it was pledged to “halve the number of global death and injuries from road traffic accidents” by 2020. Also, target 11.2 calls on delivering “access to safe,

affordable, accessible and sustainable transport systems for all, improving road safety" by 2030. The number of traffic injuries reported by the United Nations (UN) in 2020 shows a reduction from "18.7 per 100000 population in 2010 to 18.2 in 2016", which indicated less satisfactory progress for the safety goals to be met by the end of 2020 (SDG, 2020). The total cost of road injuries for the world economy is estimated to be around US$ 1.8 trillion (S. Chen, Kuhn, Prettner, & Bloom, 2019), which is approximately 0.12% of global domestic product (GDP) annual tax.

The extent of economic and human costs of road injuries in Canada are devastating. In Canada, the total number of fatalities due to traffic collisions increased from 1856 in 2017 up to 1922 in 2018, which indicates 3.6% rate of growth. Also, traffic crashes costs for Canadian society were estimated at CAD 40.7 billion, representing 2.1% of GDP in 2018 (ITF, 2020). Figure 1.1 illustrates the number of annual collision fatalities and serious injuries, according to the Transport Canada Collision Reports (*Canadian Motor Vehicle Traffic Collision Statistics*, 2018). The trend of road collision fatalities is far from Canada's road safety strategy (CCMTA, 2016).



(a)



(b)

Figure 1.1 Number of fatalities (a) and serious injuries (b) in Canada from 1999 to 2018

Active modes of transportation constitute a major portion of daily activities. Increasing the usage of nonmotorized transportation modes is one of the most effective solutions to traffic problems. Encouraging people to adopt walking and cycling modes can alleviate traffic congestion (Morar & Bertolini, 2013) and reduce the Green House Gas (GHG) emission in cities (Grabow et al., 2012). Moreover, it can enhance the communities' health conditions and reduce the pace of climate change (Rissel, 2009). However, an elevated risk of injuries in road crashes represents an obstacle for many active road users (Moudon, Lin, Jiao, Hurvitz, & Reeves, 2011). More than half of the globally recorded traffic fatalities comprise vulnerable road users. Specifically, pedestrians' fatalities account for 23% of road traffic death (WHO, 2018). According to (L. Wang et al., 2019), pedestrians are the most vulnerable road users in traffic accidents.

A similar situation regarding active road users' safety can also be seen in other countries. Crash reports in China show that more than 20% of all recorded traffic accidents in 2016 involved

pedestrians, which is associated with 44,129 pedestrian accidents (Tan et al., 2021), with pedestrians' mortality constituting 42% of all road traffic mortalities in China (L. Wang et al., 2019). In Canada, pedestrians represent 17.3% of all fatalities and 15.5% of serious injuries (*Canadian Motor Vehicle Traffic Collision Statistics*, 2018). Safety improvements would not be possible without a deep understanding of the motion dynamics and hazard factors. Therefore, there is a pronounced need for traffic safety studies to provide a deeper understanding of the important factors in pedestrians' safety level. A crucial step towards attracting more people to active modes of transportation is improving pedestrian safety. This way, not only the social/economic cost of pedestrians' fatalities in road collisions can be mitigated, but also society can benefit from active transportation advantages.

1.1.1 Crash-Based Safety Analysis

Traditionally, traffic safety studies are conducted relying on historical crash data. However, this approach has several drawbacks. First, relying on crash data is a reactive approach, where a significant number of collisions need to occur before a reliable safety evaluation can be done. Second, crash data usually suffer from limited availability, low reliability, and poor quality (Sayed, Ismail, Zaki, & Autey, 2012; Sayed & Zein, 1999). Moreover, historical collision records have other methodological issues, including under-reporting, small sample size, and unobserved heterogeneity (Mannering & Bhat, 2014). Third, reliance on crash data is morally controversial since no adjustment or treatment can be made before observing a sufficient number of collisions. Forth, collision records lack detailed information about the collision-contributing factors. Therefore, using historical crash data in safety diagnosis analysis can not provide a comprehensive

understanding of the situation. These limitations gave rise to more advanced assessment tools and evaluation methods in safety studies.

1.1.2 Surrogate Safety Measures (SSM)

The limitations of crash-based safety analysis have given rise to proactive approaches and alternative measures. Surrogate safety measures, such as Traffic Conflict Technique (TCT), can overcome the collisions fundamental issues and provide more information in a shorter time with less social cost (Chin & Quek, 1997; Sayed & Zein, 1999). Conflicts are defined based on road user interactions; therefore, different aspects of these interactions led to four types of specifications for conflicts, including evasive actions, spatial/temporal proximities, severity hierarchy, and counterfactual probability (Arun, Haque, Bhaskar, Washington, & Sayed, 2021). (Amundsen & Hyden, 1977) defined a conflict as "an observable situation in which two or more road users approach each other in space and time to such an extent that there is a risk of collision if their movements remain unchanged". (Parker Jr & Zegeer, 1989) considered evasive actions as an indication of conflicts and based their definition as "an event involving two or more road users, in which the action of one user causes the other user to make an evasive maneuver to avoid a collision". Different proximity-based indicators (e.g., Time-to-Collision (TTC) and Post Encroachment Time (PET)) and evasive-action-based indicators (step-frequency and step length) were proposed for conflict identification in literature (Tageldin, Sayed, & Shaaban, 2017). Despite the advantages, traditional conflict-based analysis suffers from other drawbacks mainly due to the manual data collection, including the cost of training observers, the subjectivity of the results, human errors, and misjudgments.

1.1.3 Automated Video Analysis

Automated data collection methods have been advocated in safety studies as a promising approach that can tackle the problems associated with manual data collection. Video analysis computer vision techniques facilitate the conflict detection process and provide objective and reliable data for safety applications. Moreover, computer vision techniques enable accurate extracting of the road users' trajectories which can be used for interaction analysis. The primary cost of video analysis is due to providing and installing high-quality cameras at study locations. However, rapid advancements in technology have reduced camera prices and eased the installment process. Nowadays, video data can be collected from locations and viewpoints which were not accessible for observers, such as aerial view of a congested intersection. The recorded video data can also be used for further investigations and validations in the future. The mentioned advantages have made automated computer vision an eminent method in data collection, traffic management, and safety diagnosis analysis (Y. Guo, Essa, Sayed, Haque, & Washington, 2019; Y. Guo, Sayed, & Zaki, 2017; Y. Guo, Sayed, Zaki, & Liu, 2016; Ismail, Sayed, Saunier, & Lim, 2009).

1.1.4 Simulation Models

Recently, there has been a growing interest in using simulation models for different road users such as vehicles (Katrakazas, Quddus, & Chen, 2017; Xu, Zheng, & Yang, 2021), pedestrians (Hussein & Sayed, 2015, 2017), bicycles (Alsaleh & Sayed, 2021b; Mohammed, Sayed, & Bigazzi, 2020), and motorcycles (Trinh, Sano, & Hatoyama, 2021). Simulation models promote an array of advantages: (1) The use of simulation models in traffic studies enables reliable and fast data collection that eliminates the need for actual field observations. (2) Simulation models enable early evaluation of different designs or treatments prior to implementation. (3) Simulation models

can also be used for safety assessment of a hypothetical situation, such as a mixed flow of Autonomous Vehicles (AV) and Conventional vehicles (CV) with different possible attributes of AVs.

1.2 Problem Statement

The majority of the simulation models are developed based on rules to prevent collisions between road users. These rules artificially control the road users' behavior at near misses without a comprehensive understanding of their preferences and collision avoidance mechanisms. As a result, the performance of the simulation models (e.g., VISSIM, PARAMICS) has been questioned by previous studies (Essa & Sayed, 2015a; Y. Guo et al., 2019). In normal situations where road users are not in conflict with any approaching traffic, they attempt to reach a comfortable velocity and movement pattern. In contrast, in dangerous situations such as conflicts, safety becomes a priority, and road users tend to perform evasive actions such as abrupt changes in speed or direction to steer clear of a collision. In this situation, the movement patterns could change drastically since evasive actions are different from the typical movements in normal circumstances. Therefore, conventional simulation models can not produce realistic results at conflict, and road users' performance at conflicts should be taken into deeper consideration.

Moreover, most of the simulation models used for traffic safety studies are mainly developed for vehicular traffic. Less attention was devoted to developing traffic safety simulation-based models for active road users. This can be attributed to the difficulty of modeling pedestrian behavior. Most of the developed microscopic pedestrian simulation models are extensions of the Cellular Automata (CA) model (Blue & Adler, 2001; Nagel & Schreckenberg, 1992) and the Social Force

Model (SFM) (Dias, Iryo-Asano, Nishiuchi, & Todoroki, 2018; Helbing & Molnar, 1995). However, these models suffer from several shortcomings. First, the models are based on methodologies that do not consider road user intelligence and rationality (Alsaleh & Sayed, 2020, 2021b). For example, the assumption of road users acting as particles influenced by forces resulted from interactions with other road users in the SFM is not realistic. Moreover, road users do not move based on predefined probability rules as the CA model assumes. Second, these models are challenging to infer road user behavior and intentions in conflict situations.

1.3 Research Objectives

As previously mentioned, it is challenging for commonly available simulation packages (e.g., VISSIM, PARAMICS) to simulate road users' interaction in conflict situations. Besides, rule-based models can not understand the pedestrians' intentions and rationality in their decision-making. Moreover, diversity of the human behavior and heterogeneity of the pedestrians' maneuvers in traffic conflicts are the main reasons that unbale the conventional approaches to appropriately simulate their behavior in conflict situations. However, understanding the pedestrians' evasive action mechanisms can provide insightful information for safety investigations. Therefore, this thesis aims to address the aforementioned problems and use advanced Artificial Intelligent (AI) techniques to model pedestrian-vehicle near misses in mixed traffic conditions.

The main objective of this thesis is to understand and simulate pedestrians' collision avoidance mechanisms in pedestrian-vehicle conflicts. Accordingly, two modeling frameworks have been proposed. The first simulation framework is a single-agent approach that utilizes the Markov

Decision Process framework to account for pedestrians' rationality in mixed traffic conditions. The second proposed model employs Markov Games to develop a multi-agent framework. The multi-agent conflict modeling approach enables understanding the preferences of both interacting road users (i.e., pedestrian and vehicle) in conflicts. These methods attempt to find an optimal policy for the simulating agents that can maximize the probability of a reasonable action similar to the actual road users. The policy is defined based on the utility of actions. In the proposed frameworks, states with higher utilities/rewards are preferred in decision-making. Both single-agent and multi-agent frameworks attempt to understand the road users' intentions from examples of their demonstrations using different Inverse Reinforcement Learning (IRL) algorithms. IRL enables determining the reward functions that can maximize the likelihood of the actual trajectories.

The single-agent framework adopted a Gaussian Process Inverse Reinforcement Learning (GPIRL) approach to recover the pedestrians' utility functions. While in the multi-agent framework, a Multi-Agent Adversarial Inverse Reinforcement Learning (MA-AIRL) algorithm is implemented to capture both pedestrian and vehicle underlying rewards in conflict situations. The reward structure allows overcoming the restrictions of the commonly used simulation models and enables realistic modeling of pedestrians and vehicles under mixed traffic conditions. In this work, two different single-agent and multi-agent Actor-Critic Deep Reinforcement Learning (DRL) approaches are applied to obtain road users' optimal policies and their evasive action mechanisms in pedestrian-vehicle conflicts. Moreover, this study attempts to evaluate and compare the proposed simulation frameworks using different metrics, including similarity measures, conflict indicators, and the performed collision avoidance mechanisms.

In summary, the main objectives of this thesis can be summarized as:

1. To understand pedestrians' evasive action mechanism in conflicts with vehicles using different IRL-based approaches to recover the underlying reward functions.
2. To obtain the road users' optimal policies to simulate the actual pedestrian-vehicle conflict situations.
3. To investigate the ability of the proposed simulation models in reproducing actual pedestrian trajectories.
4. To compare the performance of the single-agent and multi-agent simulation models in terms of accuracy of predicting safety-related metrics (i.e., conflict indicators and collision avoidance mechanism).

1.4 Thesis Structure

This thesis is comprised of six chapters organized as follows: Chapter one provides an introduction to the thesis by presenting background information of safety issues and safety evaluation methods, the importance of traffic simulation models, the thesis objective, and the thesis structure. Chapter two presents a comprehensive literature review of the topics and problems addressed in this thesis. Chapter three describes the field-measured data, video analysis techniques used in data collection and describes the database and the behavioral parameters used in the simulation models development. Chapter four provides a detailed explanation of the pedestrian-vehicle single-agent modeling framework, and the pedestrians' evasive action mechanisms are investigated. Chapter five elaborates on the development of the multi-agent framework to investigate pedestrian-vehicle interactions. In this chapter, the performance of the single-agent and multi-agent simulation

models is also compared. Lastly, a summary of the thesis, conclusions, limitations, and potential directions for future research are presented in chapter six.

Chapter 2: Literature Review

This chapter provides an outline of the theoretical frameworks and related research topics in the development of traffic microsimulation models. This literature review encompasses an overview of the key studies that have a considerable impact on the safety evaluation of traffic conflicts. This chapter consists of four sections. The first section begins with a review of traditional safety approaches, transitioning to the second section that provides information on the appearance of the traffic conflict techniques. The third section reviews the key studies in the advancement of traffic simulation models and primarily covers studies related to pedestrian-vehicle interactions. The fourth section is related to the Artificial Intelligence (AI) algorithms used in the previous traffic simulation models. This section covers an overview of agent-based modeling, MDP approach, deep reinforcement learning, and inverse reinforcement learning techniques used in previous traffic simulation models.

2.1 Traditional traffic Safety Analysis

Road safety assessments and traffic safety studies were traditionally undertaken using historical crash data. These data were used to evaluate network safety and to determine accident-prone locations. Collisions, as the outcome of hazardous situations, are important in safety assessments, and they have been used as the validation baseline of more advanced safety evaluation techniques. However, conducting safety studies relying on crash data has several well-known limitations (Arun et al., 2021; Mannering & Bhat, 2014; Saunier & Sayed, 2007; Sayed et al., 2012; Sayed & Zein, 1999). First, collisions are rare events; thus, to obtain a sufficient amount of data for a reliable statistical analysis, longer observation periods are required. Safety diagnosis studies aim to mitigate the safety issues to prevent collisions; hence waiting for collisions to occur ethically

contradicts the incentive of safety studies. Besides rarity, the randomness of the collisions can lead to statistical challenges. Since collision data are usually obtained from police reports, it can be exposed to subjectivity, under-reporting, and miss-reported information. Moreover, the driving behavior observed in collision records can not be generalized to the whole driving population, as reported collisions only cover the riskiest behaviors. Furthermore, this reactive approach can not yield any detailed behavioral information or influential factors leading to collision occurrence.

2.2 Traffic Conflicts

Due to the drawbacks of collision-based safety analysis, the Traffic Conflict Technique (TCT) has been introduced as an alternative approach. This concept was initially introduced by (Perkins & Harris, 1968) to identify traffic events that have higher occurrence rates compared to collisions. In this study, traffic conflicts between drivers are associated with the situations where one of the drivers is forced to perform an evasive action, such as breaking or sudden change in the direction. However, conflict identification based on the above definition can be subjective since not all braking actions are evasive mechanisms and could be attributed to a precautionary maneuver (Arun et al., 2021). Therefore, (Amundsen & Hyden, 1977) considered temporal/spatial proximity of the road users in traffic conflict identification (as mentioned in section 1.1.2). This definition has been widely accepted and applied in safety studies (Gettman, Pu, Sayed, Shelby, & Energy, 2008; Kraay, Van Der Horst, & Oppe, 2013).

The assumption that all traffic conflicts have the same severity is a naive approach that can lead to the misdiagnosis of hazardous situations. Thus, (Hydén, 1987) proposed a severity hierarchy for traffic conflicts that can demonstrate the nearness of different conflicts to actual collisions, as

shown in Figure 2.1. This hierarchy, which is in a pyramid shape, represents a spectrum of different severity levels of interactions, starting from actual collisions at the top going to the base with undisturbed passages. The width and height of each cross-sectional area correspond with the frequency and severity of that interaction type.

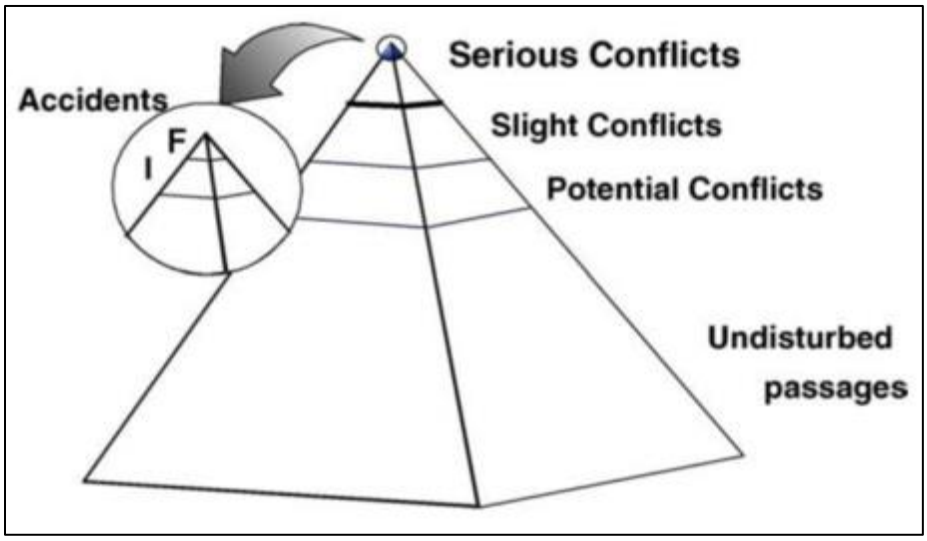


Figure 2.1 The safety pyramid (Source: (Hydén, 1987))

(Hydén, 1987) utilized a pyramid shape to describe the relationship between severity and frequency of the conflicts meaning that there is a continuous increase in the number of events in the direction of less severe conflicts. However, (Svensson, 1998) argued that if only interactions on the collision course are considered, then it is more likely that the severity hierarchy to be in a diamond shape. The diamond shape suggests that under a certain limit of severity, the road users are completely undisturbed by surrounding traffic. Therefore, it can be concluded that most of the interactions between road users have medium severity, i.e., the maximum frequency is at the center of the hierarchy (Figure 2.2).

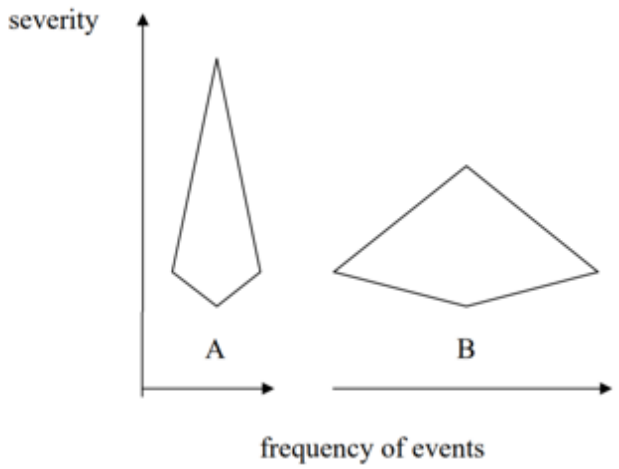


Figure 2.2 Examples of diamond hierarchy shapes for conflict frequency (Source: (Svensson, 1998))

2.2.1 Traffic conflicts Correlation with Collisions

Initially, safety studies were based on the number of collisions, and traffic conflicts have been introduced as an alternative to address the issues associated with collisions. Therefore, traffic conflicts are supposed to manifest similar safety results in comparison to collisions. Thus, the validity of traffic conflicts has been investigated by checking the correlation between the number of conflicts and collisions. For example, (Sayed & Zein, 1999) used data collected from 94 intersections (52 signalized and 42 unsignalized) to investigate the relationship between collisions and conflicts. The results showed a significant correlation between accidents and conflicts at signalized intersections, whereas unsignalized intersections showed a very weak correlation. (Sacchi & Sayed, 2016) developed safety performance functions using data collected from 49 signalized intersections in British Columbia to predict and compare the number of conflicts and collisions for different types of conflicts (i.e., left turn and rear-end). The results indicated a significant relationship between the predicted conflicts and collisions. Moreover, it was reported that the number of left-turn collisions increases more rapidly than the associated conflicts. (El-

Basyouny & Sayed, 2013) proposed a two-phase model to investigate the relationship between collisions and conflicts. This model employs a lognormal model in the first phase to predict the number of traffic conflicts, and the second phase uses a conflict-based negative binomial safety performance function to estimate the collisions. The results demonstrated a significant proportional relationship between conflicts and collisions.

2.2.2 Benefits and Drawbacks of Traffic Conflicts

The use of traffic conflict techniques in safety assessments brings about many benefits, which are summarized as follows:

1. High occurrence rate of traffic conflicts in comparison to collisions
2. Traffic conflicts can reduce the data collection time
3. Lower social and economic cost of data collection in comparison to collision
4. Resolving the ethical dilemma in the long waiting time for collision occurrence
5. Traffic conflicts can provide insightful information regarding the contributing factors to the collisions
6. Traffic conflicts can also facilitate before-and-after studies

Despite the advantages of traffic conflicts, several drawbacks have been mentioned in previous studies (Chin & Quek, 1997). These drawbacks are mainly stemming from the subjectivity in traffic conflict definition and data collection. Recognizing and differentiating the evasive actions from normal behavior could be challenging since not all braking and swerving maneuvers are of evasive nature. Moreover, manual data collections can be inaccurate due to human factors such as

lack of experience, inconsistency in observer's interpretation, and fatigue. However, errors associated with human factors can be mitigated using video analysis techniques (as mentioned in section 1.1.3).

2.3 Traffic Microsimulation Models

2.3.1 Earlier Road Safety Simulation Models

The simulation models in transportation safety were initially started by analyzing the driving characteristics. These earlier models are mainly developed based on vehicular traffic flow fundamentals, including speed, density, and gap acceptance. (Darzentas, Cooper, Storr, & McDowell, 1980) used empirical data to predict the number and severity of traffic conflicts at a nonurban T-junction. The simulation model incorporated behavioral parameters to examine the factors that can increase the likelihood of accident occurrence. (Yoshii & Kuwahara, 1995) developed a dynamic simulation model for oversaturated urban areas to estimate the dynamic traffic conditions (i.e., travel time). This study considered drivers' route choice decision-making and traffic density to simulate shockwave propagation at bottlenecks.

2.3.2 Advancement in Simulation Packages

For many years, the focus of traffic simulation models has always been on vehicular traffic. Accordingly, different modeling techniques and simulation packages such as VISSIM, PARAMICS, AIMSUN, Corridor Simulation (CORSIM), and Texas. Several studies utilized these models in safety analysis. For instance, (Archer & Young, 2010) used VISSIM to assess the impact of signal treatment on the intersections' safety level. They applied treatments including amber time extension, dilemma zone green extension, two scenarios of all-red extension, and a combination

of the dilemma zone green extension and all-red extension. The evaluation metrics used to compare the suggested treatments' effectiveness include the number of red-light violations and the number of conflicts with PET less than 4 seconds. The data used in the model calibration is collected from loop counters and video recordings. It was found that the combined treatment outperforms other scenarios in terms of reduction in the number of red-light violations and serious conflicts.

(Cunto & Saccomanno, 2008) exploited a four-step calibration process to improve a VISSIM-based microsimulation model. The calibration process consisted of heuristic selection of initial model inputs, statistical screening of inputs, relating significant inputs to safety performance, and using a genetic algorithm to obtain the best estimates. The required data for the validation data set were acquired from the Next Generation Simulation (NGSIM). The model attempted to estimate the vehicle's crash potential at signalized intersections. Accordingly, a crash potential index based on vehicle deceleration rate was used to indicate simulations' safety performance.

(Pirdavani, Brijs, Bellemans, & Wets, 2010) used PARAMICS to investigate the impacts of the vehicular speed limit on traffic safety. The microsimulation model investigated the effects of change in traffic volume and speed at signalized and unsignalized intersections. This paper's findings indicated that an increase in speed limit could result in safety deterioration.

(Caliendo & Guida, 2012) developed a microsimulation model to assess the safety level at unsignalized intersections. The task of traffic flow simulation was done using AIMSUN, and the Surrogate Safety Assessment Model (SSAM) was used to determine the critical conflicts. Critical

conflicts are identified by TTC and PET threshold values of 1.5 and 5 seconds, respectively. Results demonstrated a significant correlation between observed crashes and predicted conflicts.

(Fan, Yu, Liu, & Wang, 2013) proposed a two-step calibration process to improve the predictions of a VISSIM simulation model. The model was developed for safety assessment at freeway merge areas. In the first stage of the calibration process, the VISSIM model was modified to reach an acceptable correspondence between the field-measured and simulated volume, speed, and travel time. The second stage majorly dealt with the input parameters' calibration and adjustments of the traffic conflict indicators (e.g., TTC). The results showed that the calibration process enhanced the simulations' performance and the total conflicts' mean absolute percentage error declined from 71% to 19.9%.

(Essa & Sayed, 2015a) employed VISSIM simulation model with SSAM to investigate the level of correlation between simulated and field-measured conflicts at an urban signalized intersection. To increase the correlation, a two-step calibration procedure was proposed. In the first step, the model was adjusted to assure the consistency of the average delay times. In the second step, a Genetic Algorithm (GA) was implemented to calibrate important parameters identified by a sensitivity analysis. The same calibration technique was used by (Essa & Sayed, 2016) to improve the accuracy of a PARAMICS microsimulation model. This study also compared the results obtained from PARAMICS and VISSIM simulations. The comparison results indicated that at the TTC threshold of 1.5 seconds, the number of conflicts predicted by PARAMICS was more than the actual observations, whereas VISSIM predictions were lower than the actual number of conflicts. Nevertheless, if TTC threshold was increased to 3 s, then both PARAMIC and VISSIM

overestimated the number of conflicts. Also, further examinations were done by (Essa & Sayed, 2015b) to assess the transferability of the developed simulation models in different location studies.

(Dijkstra et al., 2010) investigated the relationship between simulated traffic conflicts and the field recorded data at junctions. This study used PARAMICS to model the microsimulation network. A TTC threshold of 2.5 seconds was used to identify the traffic conflicts. The results demonstrated a statistically strong relationship between the observed crash frequencies and the number of detected conflicts in the simulations.

2.3.3 Pedestrian simulation models

Initial investigations on pedestrian movement characteristics were conducted utilizing pedestrian flow fundamentals. The rudimentary models tried to account for the pedestrians' emerged behavior by considering their speed, density, and the preferred distance to obstacles in the walking environment. For example, (Chattaraj, Seyfried, & Chakroborty, 2009) investigated pedestrian flow attributes together with speed, density, and personal space in two different cultures, namely German and Indian. The length of the corridor was found to have no impact on pedestrians' movement. The results demonstrated that the minimum personal space for the German people is more than Indian people. In contrast, no meaningful difference between the analyzed groups had been observed for free flow speed.

(Ye, Chen, Yang, & Wu, 2008) analyzed pedestrian flow-density-speed relationships for four different walking facilities, including one-way and two-way passageways, descending stairways,

and ascending stairways. Assuming a constant pedestrian speed, the results showed that the passageways have the highest capacity compared to ascending and descending stairways. The observations indicated higher conflict potentials at locations with two opposite direction pedestrian flows. The results indicated that pedestrians tend to have lower speeds in two-way passageways compared to one-way passageways. The reason behind this phenomenon is the higher conflict occurrence rate at locations with pedestrians in opposite directional flows.

Over the decades, several microsimulation models have been developed for human motion prediction. However, these models are mainly developed for road users' path planning/navigation or obstacle avoidance without extensive evaluation for their traffic safety assessment validity. Cellular Automata (CA) (Blue & Adler, 2001; Nagel & Schreckenberg, 1992) and Social Force Models (SFM) (Helbing & Molnar, 1995) are two of the most frequently used approaches in pedestrian microsimulation.

2.3.3.1 Social Force Models

The Social Force Model (SFM) (Helbing & Molnar, 1995) is considered the pioneer of the pedestrian microsimulation models. This model is generally based on agents' stimuli and personal interests under external forces or repulsions. The social force models have been used in different applications, such as navigation, crowd panic evacuation, and group motion dynamics. For example, (Helbing, Farkas, & Vicsek, 2000) proposed a SFM to simulate pedestrian escape panics in life-threatening situations. This study suggested that the maximum chance of survival in escape panics happens when a combination of herding and individualistic behavior exist among people. Lack of individualism halts detecting available exits in the environment and herding behavior

assures detected escape ways are followed by the population. Similarly, (Parisi & Dorso, 2005) analyzed pedestrian evacuation dynamics with SFM. Three kinds of forces were considered in the model, namely desired force, social force, and granular force. This study's findings showed that the evacuation time could be expressed as a function of pedestrians' desired velocity.

(Zeng, Chen, Nakamura, & Iryo-Asano, 2014) employed a modified SFM to reproduce pedestrians crossing behavior at signalized crosswalks. The developed model considered the pedestrians' intended direction at the strategic level. Five types of forces influenced the simulated pedestrian decision-making, including driving force toward destination, repulsion force from crosswalk boundary, repulsive force from conflicting pedestrians, conflicting vehicles repulsions, and attractive force from signal control. The required data for model calibration was extracted manually at a crosswalk in japan.

(Zhou, Zhou, Pu, & Xu, 2019) extended SFM framework to simulate pedestrian behavior during Pedestrian Flashing Green (PFG) phase at signalized crosswalks. This study assumed that pedestrians' desired velocity changes according to the difference between required crossing time and PFG phase. Moreover, the model considered four repulsive and attractive forces from PFG signal, crosswalk boundary, surrounding pedestrian, and right-turning vehicles. The authors concluded that the presence of conflicting right-turning vehicles could significantly impact the pedestrian's crossing velocity. (Karamouzas, Heil, Van Beek, & Overmars, 2009) proposed an advanced SFM to simulate pedestrians' interaction with each other. This study considered psycho-social concepts of human interaction in the model to deliver realistic collision avoidance behavior. It was hypothesized that each agent (e.g., pedestrian) tries to minimize the number of interactions

with other agents. This minimization process, which led to energy saving, was done by adopting routes as early as possible for each individual. These modifications to SFM led to smooth and natural behavior in the simulated agents.

In the body of the literature, different adjustments and improvements have been applied to the SFM framework to account for the group motion dynamics. For example, (Farina, Fontanelli, Garulli, Giannitrapani, & Prattichizzo, 2017) introduced the Headed Social Force Model (HSFM), in which the pedestrians' heading is included to enhance SFM simulations. The pedestrian motion dynamics were described by two concepts; forces and torques. The former accounted for the translational dynamics, and the latter was adopted for the rotational dynamics. Furthermore, an additional force term was added to describe the group cohesion of the people moving together as a single unit.

(Moussaïd, Perozo, Garnier, Helbing, & Theraulaz, 2010) examined the interaction dynamics and organization of the human crowd motions. They utilized video recording data, including around 1500 pedestrian groups, to analyze social interactions and group walking patterns. The mathematical model is founded based on SFM fundamentals, meaning that pedestrian motions were described by a combination of driving and repulsive forces. It was found that different spatial patterns exist at different group densities. Pedestrians in groups with lower densities tend to walk side by side, while at higher densities, pedestrians tend to follow a V-like configuration to ease social interactions within the group.

2.3.3.2 Cellular Automata Models

Cellular Automata (CA) can be regarded as grid-based or discrete modeling framework. In the CA approach, cells and states are the main components that define agents' situations. Movement rules and cell occupancy conditions control the dynamic of the agents' interactions. Pedestrians' transition to adjacent cells is influenced by different factors like obstacles and the presence of other pedestrians nearby. In a general approach, CA models can be classified into three groups of lattice gas models, floor field models, and other field-based models (Li et al., 2019). One of the earliest pedestrian behavior simulations with CA was done by (Burstedde, Klauck, Schadschneider, & Zittartz, 2001). In this study, pedestrians' logic was reproduced by the floor field approach. The model structure comprised a two-dimensional grid, which was assembled by 40×40 cm^2 cells. The greatest number of occupancy in each cell was one pedestrian. This simulation framework was applied in two different situations, i.e., evacuation of people from a large room and the pedestrians' motion in a long corridor. (R.-Y. Guo, 2014) introduced a CA model with a finer discretization of space to simulate pedestrian evacuation from buildings. In this new approach, pedestrians occupied multiple cells, and in each time interval, they moved distances more than one cell. Accordingly, impacts of the discretization degree and pedestrians' velocity on the evacuation time were assessed. This study also investigated the effect of the exit width on the exit flow.

Even though CA models were one of the first introduced frameworks for pedestrian simulations, they are still widely used to this day. For example, (T. Liu, Yang, Wang, Zhou, & Xia, 2020) took advantage of the fuzzy logic method combined with cellular automata to simulate pedestrians' evacuation from a multi-exit room. The fuzzy logic method enables simulated pedestrians to determine the target exit based on the available exit options' density and distance. The developed

model is used to investigate the characteristics of the nearby obstacles and exits on pedestrian evacuation efficiency. The results demonstrated that increasing the exit width can decrease the evacuation time. As well, the number of people within the room and their initial location can change the evacuation time.

(Lu, Ren, Wang, Chan, & Wang, 2016) developed a CA model to simulate yielding and crossing behavior in pedestrian-vehicle interaction at unsignalized crosswalks. The calibration of the proposed model was done using field-observed collected data from two unsignalized crosswalks in Nanjing, China. The results showed that the model was able to predict pedestrian-vehicle interactions, pedestrian delay time, and the drivers yielding behavior. (Xie, Gao, Zhao, & Wang, 2012) employed CA modeling approach to investigate different interaction mechanisms that can occur between vehicles and pedestrians at signalized crosswalks. In this study, pedestrians' behavior is categorized into three types (i.e., careful, risker, and normal) and a floor-field model is employed to simulate pedestrians' behavior dynamics in interactions with vehicles.

2.3.4 Pedestrian-Vehicle Conflicts

Pedestrian-vehicle conflicts have always attracted attention in traffic safety. The importance of these interactions can not be overstated. Thus, researchers attempted to analyze pedestrian-vehicle conflict from different perspectives. (P. Chen, Zeng, Yu, & Wang, 2017) utilized Unmanned Aerial Vehicles (UAVs) to analyze pedestrian-vehicle conflicts at an urban intersection in China. They analyzed conflict spatially and temporally by considering Post Encroachment Time (PET) and Relative Time to Collision (RTTC) as safety surrogate measures. The authors compared two types of pedestrian-vehicle conflicts; first, Pedestrian Passing First (PPF) (the conflicts where vehicles

yield to pedestrians), second, Vehicle Passing First (VPF) (the conflicts where pedestrians yield to vehicles). The observations revealed that VPFs mainly occurred outside of the crosswalk boundary. Moreover, results demonstrated that VPFs are more dangerous than PPFs for pedestrians.

(Sheykhfard & Haghighi, 2020) took advantage of Fixed Videography (FV) and In-Motion Videography (IMV) to evaluate pedestrian-vehicle conflicts in urban and outskirt areas. This study used the linear regression model to analyze gap acceptance. Moreover, the logistic regression model was used to examine pedestrian crossing probability. The findings showed that as the number of pedestrians intending to cross the road increased, they were more likely to wait for larger gaps between vehicles, which in turn decreased the probability of accidents for pedestrians in groups. It was also concluded that with an increase in approaching vehicles' speed, the probability of pedestrians crossing decreased. The results indicated that besides vehicle speed, pedestrian distance to the vehicle and gender could be influencing pedestrian crossing behavior.

(Tageldin & Sayed, 2019) evaluated pedestrian conflicts in different traffic environments. They also compared the performance of an evasive action-based indicator (i.e., Maximum Slope of Step Frequency (MSSF)) and a traditional time proximity measure (TTC). Computer vision techniques were utilized to identify pedestrian conflicts from video recordings in five cities; Shanghai, New Delhi, New York, Doha, Vancouver. The authors employed the ordered-response logit model to assess the relevance of proposed conflict indicators with pedestrians' conflict severity. Eventually, the evasive action-based indicator had a better performance in less-organized traffic environments. On the other hand, TTC yielded better results in structured traffic environments.

(Tiwari, Bangdiwala, Saraswat, & Gaurav, 2007) used survival analysis statistical methodology to identify the importance of signal timing in pedestrian safety and risk of collision with motor vehicles at signalized intersections. This study deduced that pedestrians tended to violate the signal as signal waiting time increased. Additionally, results demonstrated that female pedestrians' mean waiting time was 27% more than males, which explained the higher frequency of traffic rules violations in the male population.

The previously mentioned studies were mainly focused on the pedestrians' behavior in pedestrian-vehicle conflicts. There is a limited number of studies targeting drivers' attributes in these kinds of conflicts. For instance, (Sheykhfard, Haghighi, Papadimitriou, & Van Gelder, 2021) used a binary logistic regression model to analyze the drivers' yielding behavior in Marked Crosswalks (MC) and Unmarked Crosswalks (UMC). The Swedish Traffic Conflict Technique (STCT) was employed to assess the severity of conflicts. Accordingly, observed conflicts were categorized into four groups: encounter, potential, slight, and serious conflicts. The results showed that the drivers' evasive maneuvers were more prevalent in more severe conflicts. The drivers most frequently used evasive maneuvers were deceleration and acceleration, lane changing, and stopping during conflicts. It was also concluded that in the UMCs, the probability of accidents between pedestrians and vehicles is higher than MCs.

(Zhao, Malenje, Wu, & Ma, 2020) developed a traffic flow model to analyze the interaction between pedestrians and vehicles at un-signalized mid-block crosswalks. These interactions were assumed as a combination of Pedestrian Gap Acceptance (PGA) and Vehicle Yielding (VY) behaviors. Subsequently, two logit models were established to define PGA and VY behaviors in

the simulation framework. The findings showed that the growth of pedestrians' volume increased the vehicular delay; meanwhile, it decreased the pedestrian waiting time. Furthermore, it was found that as the vehicles' volume grew, both vehicular delay and pedestrian waiting time increased.

2.4 Artificial Intelligence Advancements in Traffic Simulations

2.4.1 Agent-Based Modeling Approach

Heterogeneity and diversity of the pedestrians and vehicles behavior in mixed traffic conditions require a modeling approach that can understand the logic behind their decisions. Agent-based modeling (ABM) is a growing interest approach in transportation studies that can capture the emerged behavior of road users. Many advantages of ABM have been mentioned in previous studies (Erol, Levy, & Wentworth, 1998; Nicholas R Jennings, 2000; Nick R Jennings, Varga, Aarnts, Fuchs, & Skarek, 1993; Macal & North, 2005), some of which are listed as follows:

1. ABM provides a natural framework that enables describing a real-world complex phenomenon.
2. ABM framework enables adjusting different features of the agents and environment such as extent of rationality and learning capability. Moreover, this framework allows for additional rules to control the interactions between agents or between agents and the environment. This flexibility facilitates modeling complex behaviors.
3. ABM is a bottom-up approach that is capable of simulating aggregate behaviors by defining attributes of the system's constituent components.

4. ABM gives power to evaluate the consequences of hypothetical situations or proposed changes in the system before implementation.

(Hussein & Sayed, 2017) proposed an agent-based simulation model to replicate the pedestrian movements in interaction with other pedestrians in a bi-directional flow. The model relies on the behavior rules initially introduced in (Hussein & Sayed, 2015). Moreover, some additional rules were added to account for the pedestrian movements in a bi-directional flow. In this study, pedestrian groups were classified into flexible and strict. The flexible groups were characterized by individuals who were comfortable with momentary separation in the face of opposing traffic. In contrast, strict groups preferred to remain as a unit in conflicts. The parameters used in the simulation were: desired walking speed, preferred pedestrians' lateral distance, perception area, swerving distance, a threshold to distinguish flexible and district groups, prediction time, and opposing pedestrians' range. The model performance was evaluated based on actual trajectories extracted from video data. The simulated collision avoidance strategies were up to 95% similar to the actual strategies. The average speed error ranged between 5.1% and 13.3%, and the average distance error was less than 35 cm.

(Fujii, Uchida, & Yoshimura, 2017) simulated pedestrians as agents to analyze their interactions with cars and trams at crosswalks. Behavioral rules were defined to determine the pedestrians' walking speed and heading direction at each time step. These rules were categorized into five groups to account for five types of predefined pedestrian movements: free walk, collision avoidance, momentary stop, overtaking, and following. This study also used the proposed framework to examine the effects of pedestrians' existence on the left-turning cars' volume. (Rad,

de Almeida Correia, & Hagenzieker, 2020) implemented an ABM simulation in Any Logic modelling software. The developed simulation framework was used to investigate the influence of pedestrian's characteristics on their decision-making in the presence of AVs. Sixty participants of different demographics were asked to control the pedestrian agent in the simulation environment. The results showed that distance from the approaching vehicle, age, familiarity with AVs, and the communication between the AV and the pedestrian influenced the pedestrians' crossing behavior. However, the impact of gender on the crossing decision-making was not significant.

2.4.2 Markov Decision Process (MDP)

The previous studies utilized predefined rules to control the agents in the environment. Adjusting these rules is the sole path to improve the simulation results. An alternative method to traditional rule-based solutions is to add intelligence to agents' decision-making process. Markov Decision Process is a mathematical approach that enables agents to learn from their experience and improve their behavior. Different previous studies have used the MDP framework to simulate pedestrians and vehicles behavior. For example, (Hsu, Gopalswamy, Saripalli, & Shell, 2018) used an MDP framework to simulate pedestrian-vehicle interactions at unsignalized intersections. The MDP states consisted of the pedestrian's distance from the intersection, the pedestrian speed component towards the intersection, and the vehicle's distance from the intersection. The reward function penalized vehicles for entering the crosswalk area concurrently with the pedestrians.

(Brechtel, Gindele, & Dillmann, 2011) employed MDP in conjunction with Dynamic Bayesian Networks (DBN) to formulate the driving task. This study attempted to obtain the drivers' optimal decisions in the traffic environment. The reward function was designed so that the drivers were

encouraged to achieve a set of goals, outlined as follows: drive at the fastest possible speed without crashing into other vehicles, abiding the traffic regulations, economical and comfortable driving. The evaluation of the model is done by simulating the drivers' behavior in a two-lane highway scenario.

2.4.3 Inverse Reinforcement Learning (IRL) in Traffic Studies

In an MDP, agents are regarded as logical individuals executing their decisions by taking actions. The actions are prioritized based on a strategy. The strategy (policy) used by agents is built upon the received information from the environment and other interacting agents. As agents encounter different conditions in the environment, they leverage possible actions to move forward to the next states. More appealing future states are preferred in this framework, and the extent of the state's attractiveness is specified by reward functions. Reward functions can yield useful information on the agents' preferences. Moreover, the reward functions enable decision-making in unobserved situations. Defining the reward function for road users is a complicated task. In general, there are two approaches. The first is implementing a supervised learning algorithm to define the reward so that the intended task can be imitated. However, it has been shown that hand-tuned rewards can result in harmful and unintended behavior (Amodei et al., 2016). The second approach is Inverse Reinforcement Learning (IRL), which enables learning the reward functions based on examples of actual human behavior (Ng & Russell, 2000).

Simulating road user behaviors utilizing IRL approach was the subject of different recent studies. (Martinez-Gil et al., 2020) used an IRL algorithm based on maximum causal entropy to imitate pedestrians' navigation process. The action space for pedestrians contained two types of actions:

speed and velocity direction. Each of these actions was further divided into nine intensity levels. The agents in the environment had a target goal to reach. Moreover, two walls were assigned in the environment to obstruct pedestrians' paths. A discount factor equal to 0.9 was used to control the importance of future rewards. The features used to describe the agent states were defined based on several factors, including distance to the target, distance to the obstacles, speed, and angles. Eventually, a comparison of the IRL results with the Sarsa RL algorithm demonstrated better performance in emulating real pedestrian trajectories.

(Alsaleh & Sayed, 2020) employed two different IRL algorithms (i.e., Feature matching and Maximum Entropy) to simulate cyclists' behavior in shared space environments. The results indicated the superiority of the Maximum Entropy IRL algorithm in understanding and simulating the actual cyclist trajectories. In another study, (Alsaleh & Sayed, 2021b) utilized two linear and nonlinear (i.e., Gaussian Process) structure of reward function to understand the cyclist behavior in interaction with pedestrians. Moreover, this study developed a simulation framework using a Deep Reinforcement Learning approach to estimate the optimal policies. The comparison of the results indicated that the nonlinear Gaussian Process reward function could understand and account for the heterogeneity of the cyclist behavior better than the linear reinforcement learning approach. Recently, (Alsaleh & Sayed, 2021a) utilized the Multi-Agent Adversarial Inverse Reinforcement Learning (MA-AIRL) algorithm to develop a multi-agent pedestrian-cyclist simulation framework. The developed simulation tool implemented a multi-agent Actor-Critic deep reinforcement learning to simulate the pedestrian-cyclist interactions. The results of this study showed a significant improvement in the accuracy of the predicted trajectories compared to the Gaussian Process Inverse Reinforcement Learning approach.

Chapter 3: Data Collection and Analysis

This chapter provides information on the methodology used to extract pedestrian and vehicle trajectories required in the development of the simulation models. Moreover, an overview and analysis of the parameters used in this study to describe a conflict situation are presented. This chapter is divided into four sections. The first section describes the study location where the data was collected. The second section covers the algorithms utilized in the automated video analysis to track road users and identify traffic conflicts. The third section provides information on the final training and validation database extracted from the study location. The fourth section provides an overview of the road users' behavioral parameters and presents an analysis of the parameters used in the single-agent and multi-agent simulation frameworks.

3.1 Study Location

The data collection was conducted at the intersection of Wuning Road and Lanxi Road in Shanghai, China (Figure 3.1). The intersection is located near universities, commercial centers, and residential areas, making it a path for different road users. The intersection is characterized by high-density road user volumes, including vehicles, motorcycles, bicycles, and pedestrians. This highly congested four-leg intersection is a signalized intersection with permissive left turns, making pedestrians' movement phases occur concurrently with vehicles' left-turn phases. Four traffic signal heads are provided for vehicular traffic on each side of the intersection, and pedestrians' traffic signal poles are located at four corners of the intersection. The speed limit of the approaching roads is 60 kph. Pedestrians' non-compliant behaviors can be observed in different segments of the intersection. Examples of pedestrians' violations are jaywalking, walking out of

crosswalk boundary, initiating their movement prior to pedestrian green signal, using sidewalk area instead of crosswalks to wait for the green signal (Figure 3.2).

Figure 3.1 Camera view of the study location

The geometric design of the intersection consists of two perpendicular roads and four crosswalks. The major road has four lanes of vehicles in each direction, while the minor road has two. There are some important factors in the intersection's design, making the location prone to conflicts. Lack of furnishing zone, refuge island on the main roadway, and curb letdowns from pedestrian-through-zone to the crosswalk area are examples of the drawbacks in design characteristics. Congestion is a chronic problem not only throughout this large intersection but also on the sidewalks. The sidewalks are minimal in width and inappropriately designed for this intersection. Once sidewalks' capacity is exceeded by the crowd, it results in many pedestrians using the

roadways instead of sidewalk. Since pedestrians intending to cross the street can not block the narrow-designed sidewalks, they have to stand on the crosswalk area.

Traffic measurements show that the average hourly volume for vehicles, motorcycles, bicycles, and pedestrians are about 6200 vehicle/hr, 4900 motorcycle/hr, 1600 bicycle/hr, and 1200 pedestrian/hr in the study's intersection. This highly congested and large intersection rarely comes with an appropriate and clear gap for pedestrians to safely cross while the approaching vehicular traffic is in motion; consequently, they manage to accept the risks of proximal passage with surrounding vehicles. These risky movements make pedestrians use extreme evasive actions (i.e., sudden change in speed or abrupt direction change) to avoid potential accidents.

Temporal and spatial violations and pedestrian non-compliance behavior with traffic regulations are frequently observed in the intersection, leading to frequent pedestrian-vehicle conflicts. The high volume of the intersection and high occurrence rate of nonregulatory behavior (i.e., traffic violations) were the main distinguishing characteristics of the road users in the study location. Moreover, pedestrians' risky maneuvers were observed in their interactions with other road users (e.g., vehicles, bicycles). Video data was collected from the intersection for an hour during the daytime with a mild weather condition on April 15, 2014. The used high-resolution camera (1920 X 1088 pixels) was installed on a multi-story building on the northwest corner of the intersection, providing a clear vision of the scene. The frame frequency of the video footage is 25 Hz, meaning each second of the recorded data contains 25 frames.

(a) Early movement initiation

(b) Jaywalk

(C) Walking out of crosswalk boundaries

(d) Wait in the crosswalk area

Figure 3.2 Examples of pedestrians' nonregulatory behaviors

3.2 Automated Video Analysis

This section provides detailed information on the computer vision method used for road users' trajectories extraction. The video analysis procedure used in this thesis was developed at the University of British Columbia (Saunier & Sayed, 2006). This video analysis methodology was previously implemented and validated in (Tageldin, Zaki, & Sayed, 2017; Zaki, Sayed, Tageldin, & Hussein, 2013). A general view of the computer vision process is provided in Figure 3.3.

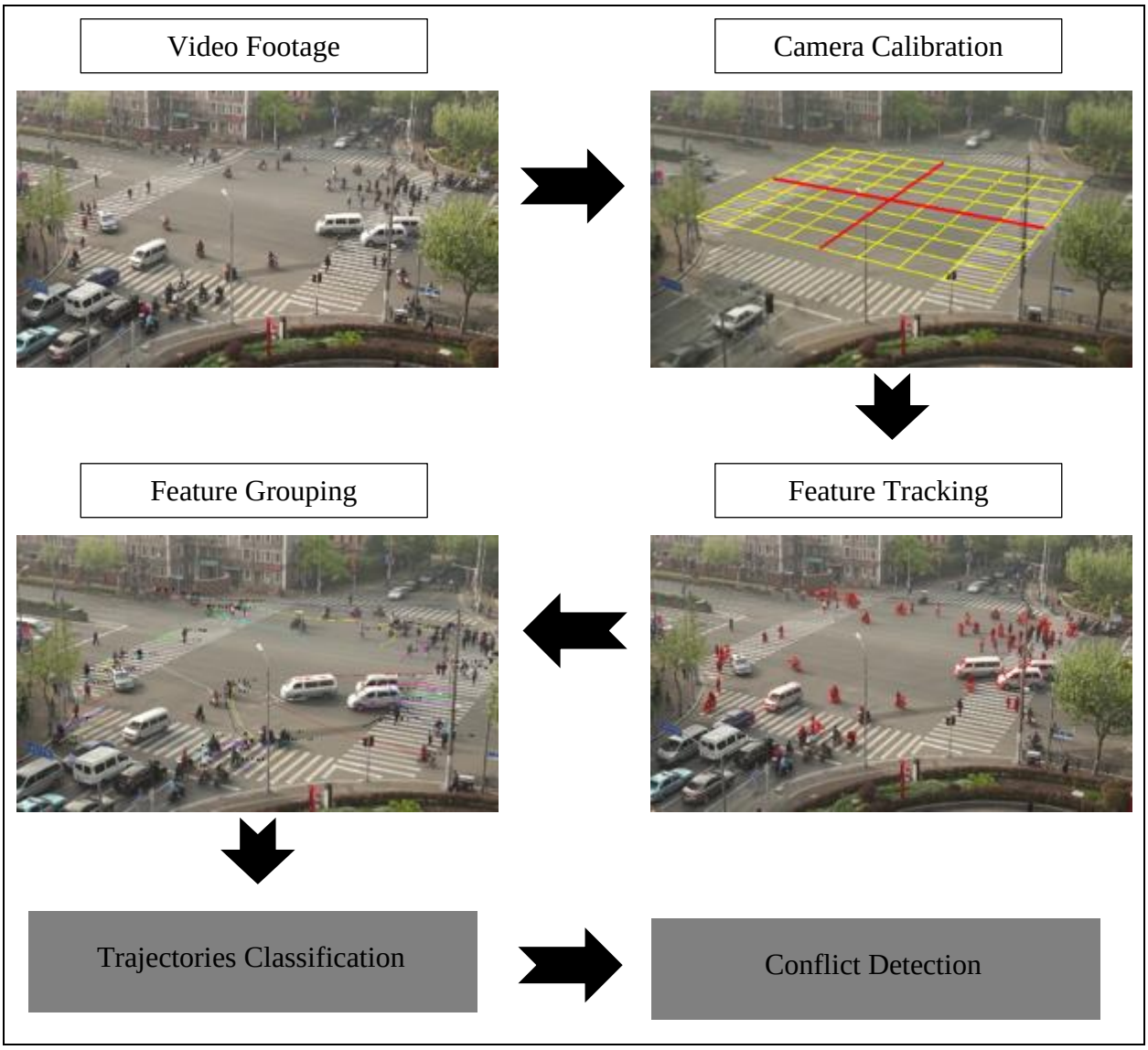


Figure 3.3 Video analysis process

3.2.1 Camera Calibration

The recorded video footage of the study location provides two-dimensional images of the real-world three-dimensional coordinates. The tracking algorithm analyzes the video data frame by frame to locate road users at each time step. The tracked locations on video frames are required to be transformed to three-dimensional coordinates. The camera calibration process defines a homography matrix to map the tracked positions on video frames onto the horizontal plane of the

intersection. This transformation matrix is influenced by camera position, orientation, skew angle, and lens distortions. Therefore, an optimization algorithm is used to accurately modify the transformation matrix. The calibration process uses four types of annotations to match the camera image features to an aerial orthographic image of the intersection. First, the optimization algorithm is adjusted based on corresponding points selected from orthographic and camera images. Second, the camera calibration is expected to annotate the actual distance between every two points; thus, the optimization process matches the distance between every two transformed points to the actual ones. In the third step, the camera calibration is modified to match the corresponding angles in the orthographic view and camera image. Therefore, actual right angles like intersection corners or actual parallel lines like lane markings are used in the optimization process. In the last step, the calibration process is adjusted to identify the third component of the actual three-dimensional space. To this end, normal vectors of the intersection plane can be used. The calibration process employs existing poles in the camera view to determine the normal vector. Eventually, a grid view of the 3D coordinates on the video image space depicts the accuracy of the transformation. More details of the camera calibration process can be found in (Ismail, Sayed, & Saunier, 2009, 2010, 2013). An illustration of the camera view and the orthographic image is provided in Figure 3.4.

Figure 3.4 Illustration of the camera view (left) and orthographic view (right) of the study location

3.2.2 Feature Tracking

The feature tracking algorithm is developed based on the Kanade-Lucas-Tomasi feature-tracking algorithm (Lucas & Kanade, 1981; Tomasi & Detection, 1991), in which the system can determine the features of the road users and the environment. The algorithm can differentiate between features in motion and stationary features. It is posited that the stationary features are associated with the environment. For example, the position of poles are fixed in the video frames; therefore, they will be recognized as a part of the background (environment) in the tracking algorithm. In contrast, moving features are assumed to be related to the road users (e.g., pedestrians and vehicles). Continuity of the detected features in video frames is the key to the moving objects tracking. Otherwise, the derived road user trajectories would be fragmented, which is not interpretable in the next steps. Regularity motion checks are then carried out to discard features with unreasonable accelerations or direction changes.

3.2.3 Feature Grouping

This step aims to connect the features belonging to a single road user and form its trajectory. It should also be noted that the detected objects have many distinguishable features. For example, vehicles are larger objects in comparison with pedestrians, motorcycles, and bicycles; also, due to the vehicle's rigidity, their extracted features exhibit identical patterns in different frames. The feature grouping algorithm uses several clues (e.g., spatial and temporal proximity and motion dynamic similarities) to identify features related to a unique object. Moreover, distances between features of each distinctive road user must lie under a specific threshold to be added to the corresponded group. This test ensures that different objects' features wouldn't be merged during the process. Also, road users are characterized by different features related to their different

segments. For instance, vehicles have different features indicating side mirrors and front bumpers. Thus, different features of a road user demonstrate similar motion vectors. These motion vectors are required to be within a specified threshold to be accounted as a part of a single group. The previous studies have shown the reliability and accuracy of the aforementioned tracking method (Saunier & Sayed, 2006).

3.2.4 Trajectories Classification

The previous steps provide continuous trajectories of the observed road users in the video data. Therefore, each extracted trajectory belongs to a specific category of road users, namely vehicle, pedestrian, motorcycle, and bicycle. Since the focus of this research is on conflicts between pedestrians and vehicles, the corresponding trajectories are required to be distinguished. The classification process enables the identification of moving objects with respect to their categories. This method mainly relies on road users' movement characteristics because the dominant motion characteristics of road users differ from each other. For example, pedestrian and cyclist motions exhibit oscillatory trends in their speed profile; however, the former renders higher frequency in the oscillatory patterns. On the other hand, vehicles demonstrate speed profiles comprised of linear segments associated with different accelerations along with their movement. The classification algorithm also considers other object-specific attributes such as size and maximum speed. The Singular Spectrum Analysis (SSA) is used to recognize the trajectories' time-series trends and the harmonic patterns. The classification process is described in detail in (Zaki & Sayed, 2013). This study also assessed the accuracy of the road users' classification technique. The results indicated accuracy values higher than 82%, which shows the robustness of the proposed framework.

3.2.5 Conflict Detection

Traffic conflicts between pedestrians and vehicles are of primary interest in this study. Thus, pedestrian-vehicle conflicts were distinguished from other types of interactions. The Post Encroachment Time (PET) is selected as the conflict indicator, as most of the pedestrian-vehicle interactions are crossing interactions. The PET between pedestrian and vehicle can be expressed as the elapsed time between the encroaching user departure from the possible conflict point and the next user arrival (Gettman & Head, 2003). Pedestrian-vehicle conflicts with PET less than 4 seconds are used in the analysis. Different PET threshold values have been suggested in the literature to determine traffic conflicts (PET<1 and PET<5 seconds for serious and slight conflicts (Almodfer, Xiong, Fang, Kong, & Zheng, 2016); and PET<3 sec for relatively risky conflicts (P. Chen, Zeng, & Yu, 2019)). The PET threshold value of 4 seconds was selected to account for different severity levels of conflicts in the study location. The PET value of 4 seconds was used in previous studies (e.g., (Y. Guo, Sayed, & Zheng, 2020)).

3.3 Training and Validation Data Sets

Pedestrian and vehicle trajectories that are involved in conflicts were extracted from the video scenes. The main criteria used for pedestrian-vehicle traffic conflict identification were the PET threshold and distraction status. Few conflict situations involved pedestrian distraction activities and therefore, distraction activities were excluded from the analysis. Eventually, a total number of 406 pedestrian and vehicle trajectories involved in conflicts were detected and obtained from the video footage. Pedestrian and vehicle trajectories were extracted at a frame rate of 1/25 second. The extracted trajectories are associated with 39138 data points. The extracted data were split into

two groups, the training and validation data sets. The training dataset consists of 310 pedestrian and vehicle trajectories, corresponding to 29414 data points (around 75% of the data points). The remaining trajectories, which is corresponded to 25% of the entire data points, are used for model testing. The same training and validation datasets are used in Chapter 4 and Chapter 5 for the development of the single-agent and multi-agent simulation models.

3.4 Behavior Analysis

The simulation models try to understand the surrounding environment from the road user's perspective. This way, the models can distinguish the essential parameters in road users' decision-making. Pedestrians and vehicles receive and analyze different aspects of a conflict to reach a suitable evasive action in response to approaching hazards. The intensity of the performed evasive actions can change according to existing nearby dangers. Explanatory variables can assist in quantifying the road users' perception of a conflict situation. The parameters used in this study to describe the road users' states and actions are presented in this section.

3.4.1 Analysis of the Previous Studies' Behavioral Parameters

In this thesis, different combinations of the behavioral variables have been tried to identify the most influential ones. These variables are consistent with the behavior components used in the body of literature. This section covers a summary of the important variables used in previously developed simulation models. These models focused on different road users and used different methods to imitate the actual trajectories. The following table provides the primary variables and associate road users, which were investigated in the body of literature.

Table 3.1 Overview of the behavioral parameters in previous studies

Study	Modeling Framework	Road User	Important Variables
(Alsaleh & Sayed, 2021b)	Gaussian Process Inverse Reinforcement Learning	Cyclist, Pedestrian	Longitudinal distance, Lateral distance, Cyclist speed, Interaction angle, Distance to destination
(Mohammed et al., 2020)	Generative Adversarial Imitation Learning	Cyclist	Cyclist speed, Speed difference, Longitudinal distance, Lateral distance, Path deviation from centerline, Direction angle with the approaching cyclist
(Y. Liu, Alsaleh, & Sayed, 2021)	Accelerated Failure Time Duration Model	Motorized vehicles (MV), Nonmotorized vehicles (NMV)	Speed of MV, Speed difference between MV and NMV, Yaw rate of NMV, Yaw rate of MV, NMV location, MV location, NMV type, Load state of NMV (whether NMV carries other passengers or goods)
(Alsaleh & Sayed, 2020)	Inverse Reinforcement Learning	Pedestrian, Cyclist	Longitudinal distance, Lateral distance, Road user speed, The speed difference of interacting road users, Road user acceleration, Interaction angle, Yaw rate
(Hussein & Sayed, 2017)	Agent Based Modeling	Pedestrian	Speed, Lateral distance, Perception area, Swerving distance, Prediction time, Opposing pedestrians range
(Dias et al., 2018)	Social Force Model	Pedestrian, Segway rider	Speed, Lateral distance, Longitudinal distance
(Zhou et al., 2019)	Social Force Model	Pedestrian	Crossing time, Speed, Direction
(Feliciani et al., 2017)	Floor Field Model, Gipps Model	Pedestrian, Vehicle	Acceleration, Deceleration, Road users' density and flow, Crosswalk width, Road users' speed, Accepted time gap
(Layegh, Mirbaha, & Rassafi, 2020)	Cellular Automata	Pedestrian	Pedestrian speed, Vehicle speed, Pedestrian relative position with respect to vehicle, Movement direction, Age, Gender
(Zeng, Chen, Yu, & Wang, 2017)	Hybrid approach combining route plan social force models	Pedestrian	Pedestrian density, Relative time and distance from surrounding pedestrians, Relative time and distance to the potential conflict point with vehicles
(P. Chen et al., 2019)	Social force model, speed and gap acceptance models	Vehicle, Pedestrian	Pedestrian speed, Vehicle speed, Movement direction, Time interval between road users
(Farah, Piccinini, Itoh, & Dozza, 2019)	Binary logistic-regression model	Vehicle, Cyclist	Lateral distance, Cyclist speed, Vehicle speed, Relative distance between cyclist and vehicle

The variables used to describe road users' states include pedestrian and vehicle speeds, interaction angle, speed difference, longitudinal and lateral distances, and distance to their destinations. In this study, road users' actions are described by acceleration and yaw rate. The variables used to describe road users' states and actions are elaborated in the following subsection.

3.4.2 Behavioral Parameters Extraction

The conflicting pedestrian and vehicle raw trajectories in traffic scenes contain their spatial and temporal information in ordered sequential form. A road user trajectory, denoted by T, is defined as a combination of finite spatial and velocity coordinates in 2-dimensional space.

$$T(t) = \{(X_1, Y_1, V_{X_1}, V_{Y_1}, \dots, X_i, Y_i, V_{X_i}, V_{Y_i}, \dots, X_n, Y_n, V_{X_n}, V_{Y_n})\} \quad (3.1)$$

where X_i and Y_i specify the road user location in the spatial coordinate system at time step i, and $V_{X_i}$and V_{Y_i} are the associated velocity components. The Savitzky-Golay smoothing filter was applied to eliminate the effect of noise in the extracted trajectories. In this study, different values of SG filter window length and polynomial degree parameters have been tested. This study applies a second-order-degree Savitzky-Golay filter and a 15 time-step (15/25 s) window size to smooth the data. The filtering setup was similar to previous studies and has been validated and showed reasonable accuracy in reducing the data noise (Bagdadi & Várhelyi, 2013; Zaki & Sayed, 2014; Zaki, Sayed, & Shaaban, 2014). Previous studies indicated that very high polynomial degrees can result in noisy outcomes (J. Wang et al., 2018). It is also noteworthy to mention that the smoothing parameters may affect the subsequent conflict results, particularly for the projection-based conflict indicators such as Time-to-collision (TTC). However, in this work, the PET is used for conflict identification, which is estimated based on the actual road user coordinates.

The road users' movement behavior can be described by their speed magnitude and direction in each time frame. Thus, the road users' evasive actions are identified based on changes in the current movement states. Abrupt swerving maneuvers and sudden changes in the speed magnitude can be an indication of a hazardous situation. The speed profiles (S) of pedestrians and vehicles are defined according to the magnitude of their velocity components each time step, as shown in equation (3.2). Road user acceleration profile (a) is an essential element in understanding its evasive action. The acceleration profile, which is the derivative of the velocity, can be computed using equation (3.3).

$$S_t = \left\| \vec{V} \right\| = l^2 - norm(V_X, V_Y) \tag{3.2}$$

$$a_t = \frac{dS}{dt} \tag{3.3}$$

where *t* is the time step.

Pedestrian's swerving maneuver can be described by the rate of change in the heading direction (φ), which is referred to as the yaw rate (r). The yaw rate is the angular velocity of a pedestrian, which indicates their rotation speed. This indicator captures the change in the road users' motion pattern, which is useful in quantifying pedestrians' evasive actions. Pedestrians' yaw rate profiles can be computed as given by equation (3.4) (Ayres, Wilson, & LeBlanc, 2004).

$$r(t) = \frac{d\varphi}{dt} \tag{3.4}$$

where *t* is the time step. Yaw rate can have positive and negative values; the former indicates spinning clockwise, and the latter shows counterclockwise rotations. Positive and negative values of acceleration indicate an increase and decrease of the road users' speed, respectively.

A distance measure variable is essential to account for the proximity of the involved road users in conflict situations. In this study, a distance vector ($\vec{d}$) for each agent (i.e., pedestrian and vehicle) is defined as a vector originating from that agent towards the other approaching road user with a magnitude of the distance between them. Hence, in a pedestrian-vehicle conflict, the magnitude of the distance vector is equal for both users, but the vectors are pointing to opposite directions.

The velocity vector ($\vec{v}$) for each agent contains information on the speed magnitude and direction. The components of each agent's distance vector with respect to the same agent's velocity vector can have extra clues accounting for the collision avoidance behavior. The vertical and horizontal components of the distance vector with respect to the velocity vector are referred to as the longitudinal distance and lateral distance, respectively. Therefore, the road users' longitudinal and lateral distance can be computed using equations (3.6) and (3.7). Figure 3.5 and Figure 3.6 demonstrate the lateral and longitudinal distances in pedestrian-vehicle interactions.

$$\theta = cos^{-1}\left(\frac{\vec{d}.\vec{v}}{\left\|\vec{d}\right\| \left\|\vec{v}\right\|}\right) \tag{3.5}$$

$$Longitudinal\ \mathrm{D}istance = \left\|\vec{d}\right\|.cos\,\theta \tag{3.6}$$

$$Lateral\ \mathrm{D}istance = \left| \left\| \vec{d} \right\| . \sin\theta \right| \tag{3.7}$$

where θ is the angle between the road user (i.e., pedestrian or vehicle) velocity vector and the distance vector ($\vec{d}$). A positive value of the pedestrian's longitudinal distance shows a situation where the vehicle is in front of the pedestrian, and a longitudinal distance with a negative value represents a conflict situation where the vehicle is behind the pedestrian (i.e., the conflict situation is resolved). This sign description can be extended to vehicles' longitudinal distance values with a difference that the situation is from the drivers' perspective. Hence, positive and negative values of vehicles' longitudinal distance represent the situations where the pedestrian is in front and behind the vehicle, respectively.

Pedestrians' reactions are not only influenced by their movement directions but also by their relative movement directions with respect to vehicles (Alsaleh & Sayed, 2021b). The pedestrians' movement directions alone do not give a complete estimation of nearby hazards recognized by the pedestrians. Similarly, vehicles' perception of safety can be influenced by the moving direction of the approaching pedestrian with respect to the vehicle's heading direction. The angular difference between the interacting road users accounts for the agents' relative direction arrangement (equation (3.8)). When a pedestrian and a vehicle are moving in the same direction (i.e., parallel velocity vectors), the angle difference gets close to zero. However, at the maximum angle difference (i.e., π), road users would be moving in opposite directions.

$$\Delta\varphi_{veh-ped} = \varphi_{vehicle} - \varphi_{pedestrian} \tag{3.8}$$

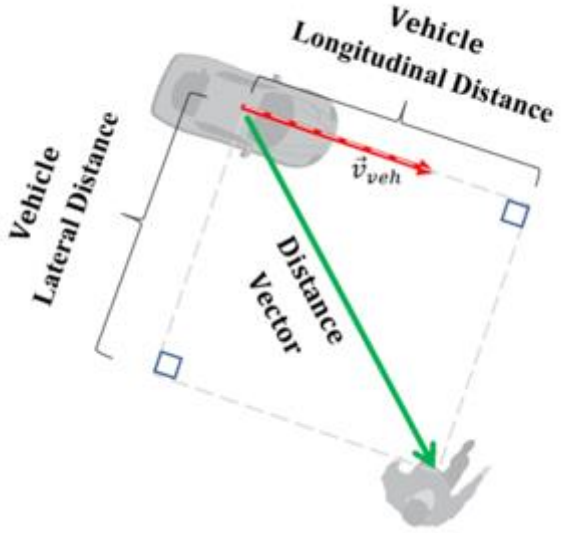


Figure 3.5 Vehicle's longitudinal and lateral distance illustration

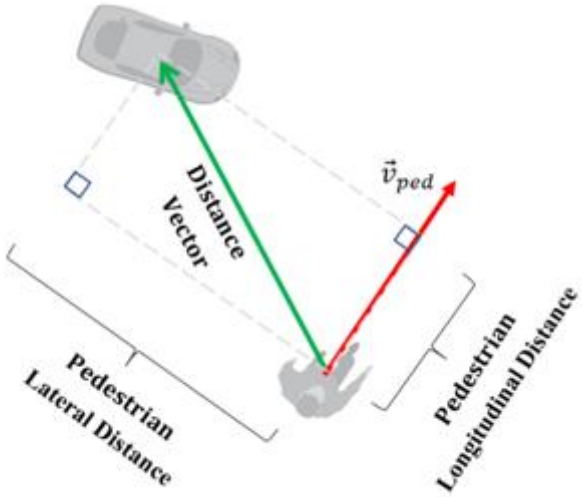


Figure 3.6 Pedestrian's longitudinal and lateral distance illustration

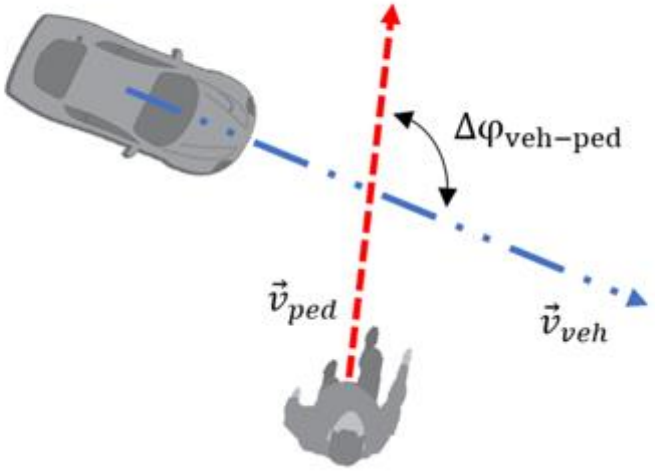


Figure 3.7 Angle difference illustration

Pedestrian and vehicle evasive actions are temporary movements to escape from a collision while moving towards their destinations. Had no conflicts occurred in the intersection, pedestrians and vehicles might have chosen to move directly to their destination point. Hence, distance to the destination is computed using instantaneous positions and final destination coordinates.

3.4.3 Analysis of the Behavioral Parameter

Pedestrian and vehicle trajectories that are involved in conflicts are considered in the analysis. Most of the pedestrian-vehicle conflicts analyzed in this study occurred at high mixed traffic densities. The main variables used to describe the road users' states in the conflict situation include road users' speed, the speed difference between conflicting road users, the angle difference between road users' heading direction, longitudinal distance, lateral distance, and distance to the destination. Furthermore, pedestrian and vehicle actions are specified by two parameters, namely

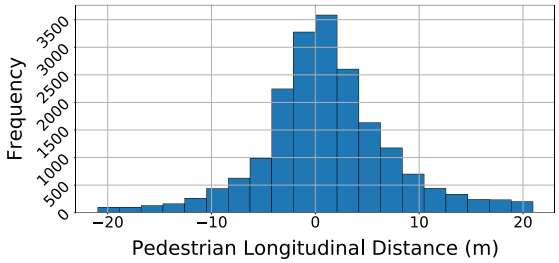
acceleration and yaw rate. The distributions of these behavioral parameters are provided in Figure 3.8, and the associated descriptive statistics are presented in Table 3.2.

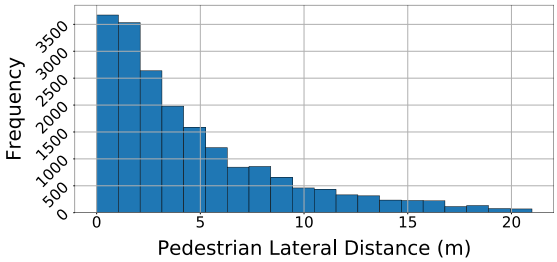
Analyzing the distributions revealed that vehicles tend to have higher velocities and acceleration rates in conflicts than pedestrians. However, the velocity difference distribution shows certain circumstances where pedestrians have higher velocities than the interacting vehicles (i.e., negative values of velocity difference). Both pedestrian and vehicle acceleration histograms are slightly skewed to the left, showing the road users' mild preference for increasing speed in conflict situations. The shape of the pedestrian yaw rate histogram is widely spread in comparison with the vehicle yaw rate, which confirms the higher reported standard deviation for pedestrian yaw rate. This observation shows pedestrians' tendency to adopt larger direction changes compared to vehicles.

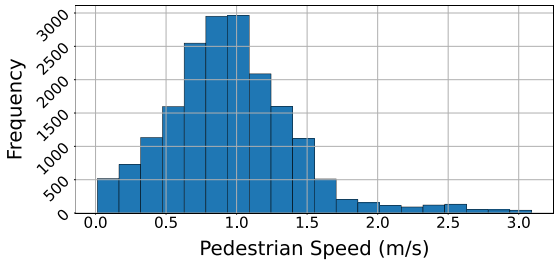
Table 3.2 Descriptive statistics of pedestrian and vehicle behavioral parameters

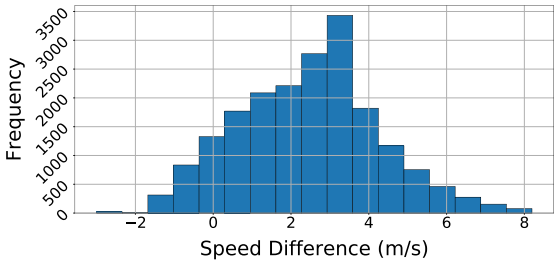
Variable	Mean (SD)*
Pedestrian Speed (m/s)	1.196 (0.466)
Pedestrian Acceleration (m/s^2)	0.104 (0.387)
Pedestrian Yaw Rate (rad/s)	-0.021 (0.381)
Speed Difference (m/s)	2.460 (1.808)
Pedestrian Lateral Distance (m)	4.633 (4.345)
Pedestrian Longitudinal Distance (m)	1.364 (6.509)
Angle Difference (rad)	1.625 (0.907)
Pedestrian Distance to Destination (m)	3.964 (4.779)
Vehicle Speed (m/s)	3.65 (1.907)
Vehicle Acceleration (m/s^2)	0.208 (0.603)
Vehicle Yaw Rate (rad/s)	-0.030 (0.274)
Vehicle Lateral Distance (m)	3.205 (3.157)
Vehicle Longitudinal Distance (m)	2.197 (7.817)
Vehicle Distance to Destination (m)	10.331 (9.203)

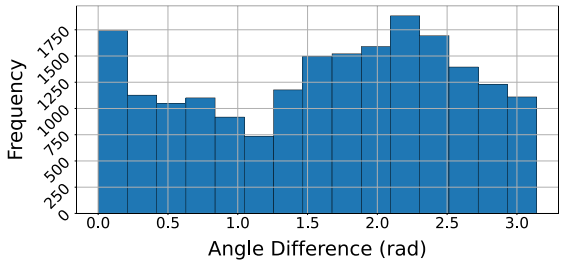
*SD = Standard Deviation

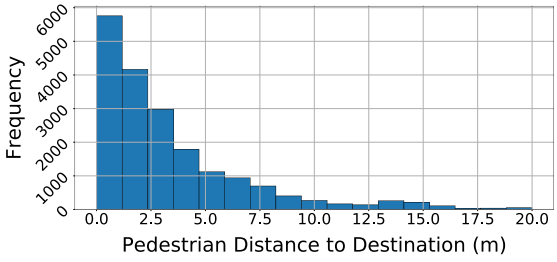

Frequency
Pedestrian Longitudinal Distance (m)

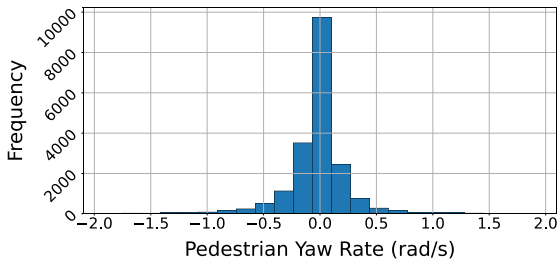

Frequency
Pedestrian Lateral Distance (m)

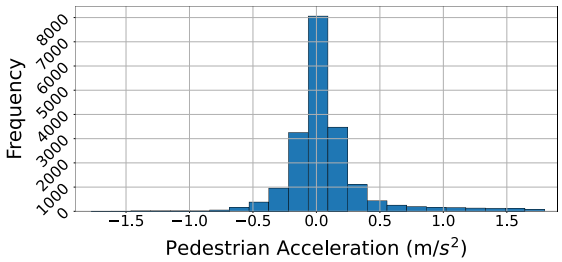

Frequency
Pedestrian Speed (m/s)

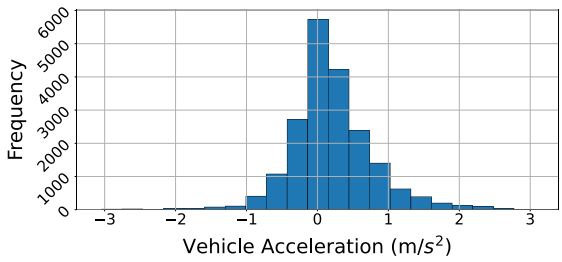

Frequency
Speed Difference (m/s)

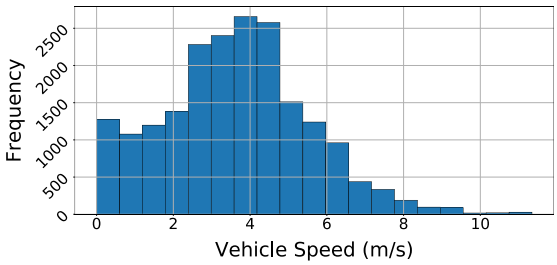

Frequency
Angle Difference (rad)


Frequency
Pedestrian Distance to Destination (m)


Frequency
Pedestrian Yaw Rate (rad/s)


Frequency
Pedestrian Acceleration (m/s^2)


Frequency
Vehicle Acceleration (m/s^2)


Frequency
Vehicle Speed (m/s)

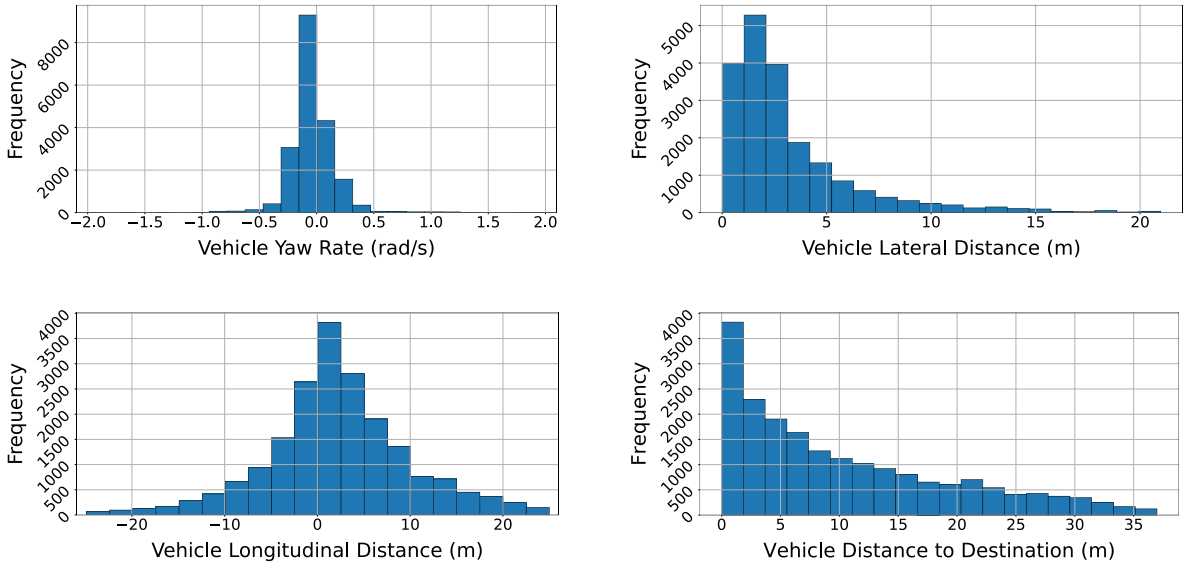


Figure 3.8 Pedestrian and vehicle behavioral parameters histogram

Chapter 4: Single-Agent Modeling of Pedestrian-Vehicle Interactions

This chapter presents the development of the single-agent modeling simulation framework for pedestrian-vehicles conflicts. The developed model utilizes the Markov Decision Process (MDP) modeling framework to simulate pedestrian-vehicle conflicts. In this approach, pedestrians are considered as rational, intelligent agents who can find an appropriate collision avoidance strategy in mixed traffic conditions. In this simulation, road users choose their actions following a policy that maximizes the underlying reward functions. The reward functions are extracted from actual demonstrations of the human behavior employing a Gaussian Process Inverse Reinforcement Learning approach. This technique enables capturing the complex nonlinear structure of the rewards. Furthermore, Asynchronous Advantage Actor-Critic Deep Reinforcement Learning is applied to obtain pedestrians' optimal policies to simulate their evasive action mechanisms in pedestrian-vehicle conflict situations.

This chapter contains six sections. The first section explains the theoretical framework of the MDP approach. The second section describes the development of the IRL models and presents a mathematical definition of the GPIRL approach. Section three describes the Deep Reinforcement Learning (DRL) framework utilized to extract pedestrian's optimal policies. In section four, a detailed explanation of the developed single-agent simulation tool is provided. Section five presents the results of the simulation framework, including reward functions, trajectory predictions, and collision avoidance strategies. This section also provides an evaluation of the simulated trajectories using similarity measures and conflict indicators (i.e., PET). Finally, a summary of the developed framework and results and further discussions are provided in section six.

4.1 Markov Decision Process (MDP)

Pedestrians' decision-making in conflict situations is based on the observed information about the conflict situation and the other nearby road users. Thus, the Markov Decision Process (MDP) framework attributes are suitable for pedestrian behavior modeling in conflict situations. In the MDP framework, the agent's behavior process is modeled as state-action pairs. The pedestrians' states describe their current situations and are presented by a set of variables including pedestrian longitudinal distance, pedestrian lateral distance, pedestrian speed, angle difference, pedestrians' distance to destination, and speed difference. The speed difference in this chapter is computed by subtracting the opposing vehicle velocity magnitude from the pedestrian velocity magnitude. The pedestrians' actions are described by the changes in their speeds (i.e., accelerations) and movement directions (i.e., yaw rates). The MDP can be described as a tuple $(S, A, P_{ss'}^{a}, R, \gamma)$, where $S = \{s\}$ is a set of states representing the situation of the agent on the environment each time frame., $A = \{a\}$ is a set of actions, $P_{ss'}^{a}$ denotes a set of transition probabilities, R is the reward function, and γ is the discount factor and is set to 0.975, similar to (Alsaleh & Sayed, 2020).

Figure 4.1 Agent environment interaction in MDP framework

In this framework, rewards are the incentives for agents to prioritize their actions. In each state, agents have a set of actions with specific rewards. Higher reward values are preferred and will specify future actions (Sutton & Barto, 2018). Reward functions link states to utilities. The total collected reward is the sum of the state rewards considering the present equivalent of all future values. A policy is a mapping from states to actions. To choose among possible actions in each state, the sum of the current and discounted ongoing rewards of that action should be estimated first. The value function denotes the rewards collection over time following a policy as given by equation (4.2). Thus, for any initial state S_0:

$$\pi : S \rightarrow A\,, \qquad \pi(a|s) = P[a_t = a|s_t = s] \tag{4.1}$$

$$V^{\pi}(s_0) = E[R(s_0) + \gamma R(s_1) + \gamma^2 R(s_2) + \cdots |\pi] \tag{4.2}$$

where π is the policy, and V denotes the value function.

Therefore, the value function comprises the immediate reward and the future rewards' current value. With similar logic, the bellman equation (Sutton & Barto, 2018) can be presented as given by equation (4.3). The Q-function represents the expected reward of taking action a in state s and following policy π, and it is given by equation (4.4).

$$V^{\pi}(s) = R(s) + \gamma \sum_{s'} P\big(s'\big|s, \pi(s)\big) V^{\pi}(s') = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R(s_t)|\pi\right] \tag{4.3}$$

$$Q^{\pi}(s, a) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R(s_t)|s_t = s, a_t = a\right] \tag{4.4}$$

where Q is the Q-function.

4.2 Inverse Reinforcement Learning (IRL)

In general, IRL methods give solutions to find the relation between action and underlying rewards. To specify road users' interactions in a simulated environment, reward functions are essential. The rewards are aimed to be defined so that a specific behavior can be imitated. Different solutions have been proposed for identifying the rewards. These approaches can be split into two groups; (1) mimicking the intended task by applying supervised-based learning to build a direct mapping from states to actions. (2) Inverse Reinforcement Learning (IRL) (Abbeel & Ng, 2004; Ng & Russell, 2000).

Utilizing manually-tuned and pre-specified rewards is a complicated task that has been shown to have downsides. Designing rewards manually can lead to unintended road user behavior and harmful side effects (Amodei et al., 2016). Instead, underlying reward functions can be obtained from actual demonstrations of human behaviors. IRL is the solution to the problem of inferring the rewards from expert demonstration ($\mathcal{D}$), where $\mathcal{D}$ includes the actual observed trajectories (τ). An expert trajectory includes a set of state-action pairs.

In this method, agents are aimed to perceive and act like actual road users. Therefore, simulated agents are required to understand the distinct traits of the actual pedestrian decision-making process. The reward functions can manifest the underlying intentions and clarify the incentive for agents' decision-making. As mentioned by (Abbeel & Ng, 2004) the reward function is the most succinct, robust, and transferrable representation of a task. Despite the advantages of IRLs, they are considered ill-defined problems for two reasons; first, there could be more than one policy that

can explain the expert demonstrations. Second, each of these optimal policies can also be explained by many rewards.

4.2.1 Maximum Entropy Inverse Reinforcement Learning

The actual pedestrians do not always follow the optimal policy and their behavior is near-optimal. Maximum Entropy Inverse Reinforcement Learning (MaxEnt IRL) is a general probabilistic framework introduced by (Ziebart, Maas, Bagnell, & Dey, 2008) to account for the suboptimal behavior of expert demonstrations. This method can also tackle the ambiguity problem of having multiple optimal policies to explain expert demonstrations. MaxEnt IRL defines the reward as a linear function of features. Under this method, the probability of selecting a trajectory is exponentially proportionate to the rewards collected along that trajectory. Therefore, the probability of taking a path ($p_\omega(\tau)$) can be expressed as given by equation (4.5).

$$p_\omega(\tau) \propto \exp\left(\sum_{t=1}^{T} r_\omega(s^t, a^t)\right)\left[\eta(s^1)\prod_{t=1}^{T} P(s^{t+1}|s^t, a^t)\right] \tag{4.5}$$

where ω denotes the reward function parameters, and $\eta(s^1)$ denotes the initial state distribution. The goal of the MaxEnt is to search for a set of reward parameters (ω) that can maximize the likelihood of the expert demonstrations.

$$\max_\omega \mathbb{E}_{\pi_E}[\log p_\omega(\tau)] = \mathbb{E}_{\tau\sim\pi_E}\left[\sum_{t=1}^{T} r_\omega(s^t, a^t)\right] - \log Z_\omega \tag{4.6}$$

where Z_ω is the partition function that is the integral of the exponential term in equation (4.5) over all possible trajectories.

4.2.2 Gaussian Process Inverse Reinforcement Learning (GPIRL)

Different studies have proposed various methods to obtain the rewards, such as assuming the reward as a linear combination of the state features (Ziebart et al., 2008). The Gaussian Process Inverse Reinforcement Learning is proposed by (Levine, Popovic, & Koltun, 2011). In this approach, rewards are described as a general nonlinear function of state features, which enables capturing complex road user behavior.

The goal of the current research is to find the reward function that imitates the pedestrians' trajectories. The IRL approach that models agent's reward function as continuous Gaussian Process (GP) is utilized in this study (Levine & Koltun, 2012). This algorithm uses the maximum entropy principle (Ziebart, 2010) to account for sub-optimal road user behavior. Therefore, for an agent starting from initial state x_0 taking action $\boldsymbol{a}$, equation (4.5) can be written as given in equation (4.7).

$$P(\boldsymbol{a}|x_0) = \frac{1}{Z}\exp\left(\sum_t r(x_t, a_t)\right) \tag{4.7}$$

where $r(x_t, a_t)$ is the reward collected from visited states along the trajectory length (t) and Z is the partition function. Computing Z is feasible for small-scale problems; however, in high-dimensional spaces, computing Z becomes challenging. To tackle this issue, Laplace approximation can be used (Levine & Koltun, 2012). Under this approximation, equation (4.7) can be rewritten to equation (4.8).

$$P(\boldsymbol{a}|x_0) = e^{r(\boldsymbol{a})}\left[\int e^{r(\tilde{\boldsymbol{a}})} d\tilde{\boldsymbol{a}}\right]^{-1} \tag{4.8}$$

where $\tilde{\boldsymbol{a}}$ shows all possible actions, $r(\boldsymbol{a})$ is the sum of the discounted rewards, and $r(\tilde{\boldsymbol{a}})$ is approximated using a second-order Taylor expansion of r around $\boldsymbol{a}$, as given by equation (4.9).

$$r(\tilde{\boldsymbol{a}}) \approx r(\boldsymbol{a}) + (\tilde{\boldsymbol{a}} - \boldsymbol{a})^T \frac{\partial r}{\partial \boldsymbol{a}} + \frac{1}{2} (\tilde{\boldsymbol{a}} - \boldsymbol{a})^T \frac{\partial^2 r}{\partial \boldsymbol{a}^2} (\tilde{\boldsymbol{a}} - \boldsymbol{a}) \tag{4.9}$$

Let $\boldsymbol{g}$ and $\boldsymbol{H}$ denote the gradient ($\frac{\partial r}{\partial \boldsymbol{a}}$) and the Hessian ($\frac{\partial^2 r}{\partial \boldsymbol{a}^2}$), respectively. Thus, the approximation to equation (4.8) can be written as given by equation (4.10).

$$P(\boldsymbol{a}|x_0) = e^{\frac{1}{2}\boldsymbol{g}^T \boldsymbol{H}^{-1} \boldsymbol{g}} + |-\boldsymbol{H}|^{\frac{1}{2}} (2\pi)^{-\frac{d_{\boldsymbol{a}}}{2}} \tag{4.10}$$

where $d_{\boldsymbol{a}}$ represents the number of action features. The log-likelihood of the expert's demonstration ($\boldsymbol{\mathcal{L}}$) is given by equation (4.11). The reward function can be obtained by maximizing the expert demonstration likelihood.

$$\boldsymbol{\mathcal{L}} = \frac{1}{2} \boldsymbol{g}^T \boldsymbol{H}^{-1} \boldsymbol{g} + \frac{1}{2} \log|-\boldsymbol{H}| - \frac{d_{\boldsymbol{a}}}{2} \log 2\pi \tag{4.11}$$

In this study, the reward function is described as a general GP nonlinear function of the state features. The GP tries to learn the true output u of the noisy output y so that the captured reward function reflects the expert demonstrations $\mathcal{D} = \{T_1, \dots, T_i, \dots, T_n\}$. Expert demonstrations consist of the collected pedestrian trajectories. The supporting states (X_u) contain the features of all states in the training dataset. The kernel hyperparameter is denoted by θ. The values for θ and u can be obtained by optimizing the expert demonstration likelihood given in equation (4.12).

$$P(\mathrm{D}, \boldsymbol{u}, \boldsymbol{\theta}|\boldsymbol{X_u}) = \left[\int_r \underbrace{P(\mathrm{D}|\boldsymbol{r})}_{IRL\ term} \underbrace{P(\boldsymbol{r}|\boldsymbol{u}, \boldsymbol{\theta}, \boldsymbol{X_u})}_{GP\ posterior} d\boldsymbol{r} \right] \underbrace{P(\boldsymbol{u}, \boldsymbol{\theta}|\boldsymbol{X_u})}_{GP\ probability} \tag{4.12}$$

The IRL term indicates the likelihood of the data and can be calculated using equation (4.13), as described by (Ziebart, 2010).

$$P(\mathrm{D}|\boldsymbol{r}) = \sum_i \sum_t (Q^r_{s_{i,t}\, a_{i,t}} - V^r_{s_{i,t}}) \tag{4.13}$$

The GP probability shows the prior probability of the selected u and θ. The posterior is the probability of reward $\boldsymbol{r}$ given $\boldsymbol{u}, \boldsymbol{\theta}, \boldsymbol{X_u}$. The GP posterior follows a Gaussian distribution with a mean of $\boldsymbol{K^T_{r,u} K^{-1}_{r,u} u}$ and covariance of $\boldsymbol{K_{r,r} - K^T_{r,u} K^{-1}_{u,u} K_{r,u}}$. The GP covariance is specified by GP kernel function k, as given by equation (4.14).

$$\boldsymbol{K_{i,j}} = k(f^i, f^j) = \beta\, exp\left(-\frac{1}{2}\sum_k \lambda_k \left[\left(f^i_k - f^j_k\right)^2 + 1_{i \neq j}\sigma^2\right]\right) \tag{4.14}$$

where λ, β are kernel function hyperparameters and f^i, f^j are two feature points. The IRL term in equation (4.12) cannot be solved with a closed-form expression. Alternatively, it can be approximated using the mean value of $\boldsymbol{r}$, which equals to $\boldsymbol{K^T_{r,u} K^{-1}_{r,u} u}$. Thus, the log-likelihood function can be written as given by equation (4.15).

$$log\, P(\mathcal{D}, \boldsymbol{u}, \boldsymbol{\theta}|\boldsymbol{X_u}) = log\, P\left(\mathcal{D}\middle|\boldsymbol{r} = \boldsymbol{K^T_{r,u} K^{-1}_{r,u} u}\right) + log\; P(\boldsymbol{u}, \boldsymbol{\theta}|\boldsymbol{X_u}) \tag{4.15}$$

Optimizing the log-likelihood given by equation (4.15) can yield the rewards. This optimization is done using the Limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method (D. C. Liu & Nocedal, 1989). Although the recovered reward is obtained from the expert's demonstrations, its functionality is not limited to the states visited in the training dataset, and it covers all the state space. Further details on the derivations are described in (Levine & Koltun, 2012; Levine et al., 2011).

4.3 Pedestrian Policy Estimation

Pedestrians' optimal policies (sequences of decisions/ actions) are estimated after recovering their GP reward function using the Deep Reinforcement Learning (DRL) framework. This study uses the Advantage Actor-Critic DRL method to estimate pedestrian policies (Mnih et al., 2016). The Actor-Critic method consists of two neural networks. The first one is the Critic, which estimates the value function using the recovered GP reward function based on the agent's actions. The second one is the Actor, which takes different actions in the environment. The role of the Critic neural network is to direct the Actor to a more rewarding policy. The Advantage Actor-Critic method uses the value function as the baseline to criticize selected actions and encourages a more rewarding policy. The advantage function in this approach can be formulated as given by equation (4.16).

$$A(s_t, a_t) = Q(s_t, a_t) - V(s_t) \tag{4.16}$$

where A is the value of the advantage function, which returns the difference between a specific action a_t from state s_t to the expected value of that state. Using the Bellman equation, Q-function can be expressed as a function of V. Thus, equation (4.16) can be rewritten as given by equation (4.17).

$$A(s_t, a_t) = \sum_{i=0}^{n-1} \gamma^i r(s_{t+i}, a_{t+i}) + \gamma^n V^\pi(s_{t+n}) - V^\pi(s_t) \tag{4.17}$$

This method updates the policy and value function after simulating all states of a trajectory. The reinforcement learning algorithm's objective is to maximize the expected returns, as presented in equation (4.18). Thus, the objective function gradient, considering the advantage function, can be reformulated as provided by equation (4.19) (Mnih et al., 2016).

$$J(\theta_\pi) = \mathbb{E}[\sum_{t=0}^{T-1} r_{t+1}] \tag{4.18}$$

$$\nabla J = \nabla_{\theta_\pi} log\, \pi(a_t|s_t\,; \theta_\pi) A(s_t, a_t: \theta_\pi, \theta_v) \tag{4.19}$$

where θ_π, θ_v are the policy and value function parameters, respectively. J is the objective function, T is the terminating state.

Actor and Critic architecture consists of two neural networks, including two hidden layers in each. The Exponential Linear Unit (ELU) is used as the activation function of nodes. Previous studies that utilized neural networks showed that using a single hidden layer network can work sufficiently for general approximations (Hornik, 1993; G.-B. Huang & Babri, 1998). However, recent studies indicated that neural networks with two hidden layers could show better performance compared to the single-hidden layer networks (Thomas, Petridis, Walters, Gheytassi, & Morgan, 2017; Thomas, Walters, Gheytassi, Morgan, & Petridis, 2016). Nevertheless, adding three or more hidden layers can significantly increase the computational cost. In this work, the selection of the number of hidden layers was based on previous studies that utilized Actor-Critic Deep Reinforcement Learning for policy estimation (Z. Huang, Zhou, Zhuang, & Zhou, 2017; Kim, Kaelbling, & Lozano-Pérez, 2019; Zhong, Lu, Gursoy, & Velipasalar, 2018).

The Critic and Actor networks can be considered as the value function and policy equivalents. In each action, the target is to compute the advantage of taking a specific action compared to that state's expected value (equation (4.17)). Therefore, the target of the Critic structure is to converge

to the $V(s_t)$. Accordingly, the Critic tries to modify the policy in order to minimize the difference of the selected action values with the expected ones. This difference, which is also known as the critic loss, can be written as given by equation (4.20).

$$l_{critic} = A(s_t, a_t)^2 \tag{4.20}$$

In the next step, the Actor updates the policy. Similarly, the Actor network tries to change the policy parameters so that it maximizes the collected rewards. The policy gradient theorem (Silver et al., 2014) can be used to optimize the agent's policy. Thus, the actor loss can be expressed as a minimization function with the objective given by equation (4.21).

$$l_{actor} = -log\ (\pi).\, l_{critic} \tag{4.21}$$

This iteration of updating policy and estimating the value function continues until the optimization converges. A diagram of this procedure is illustrated in Figure 4.2.

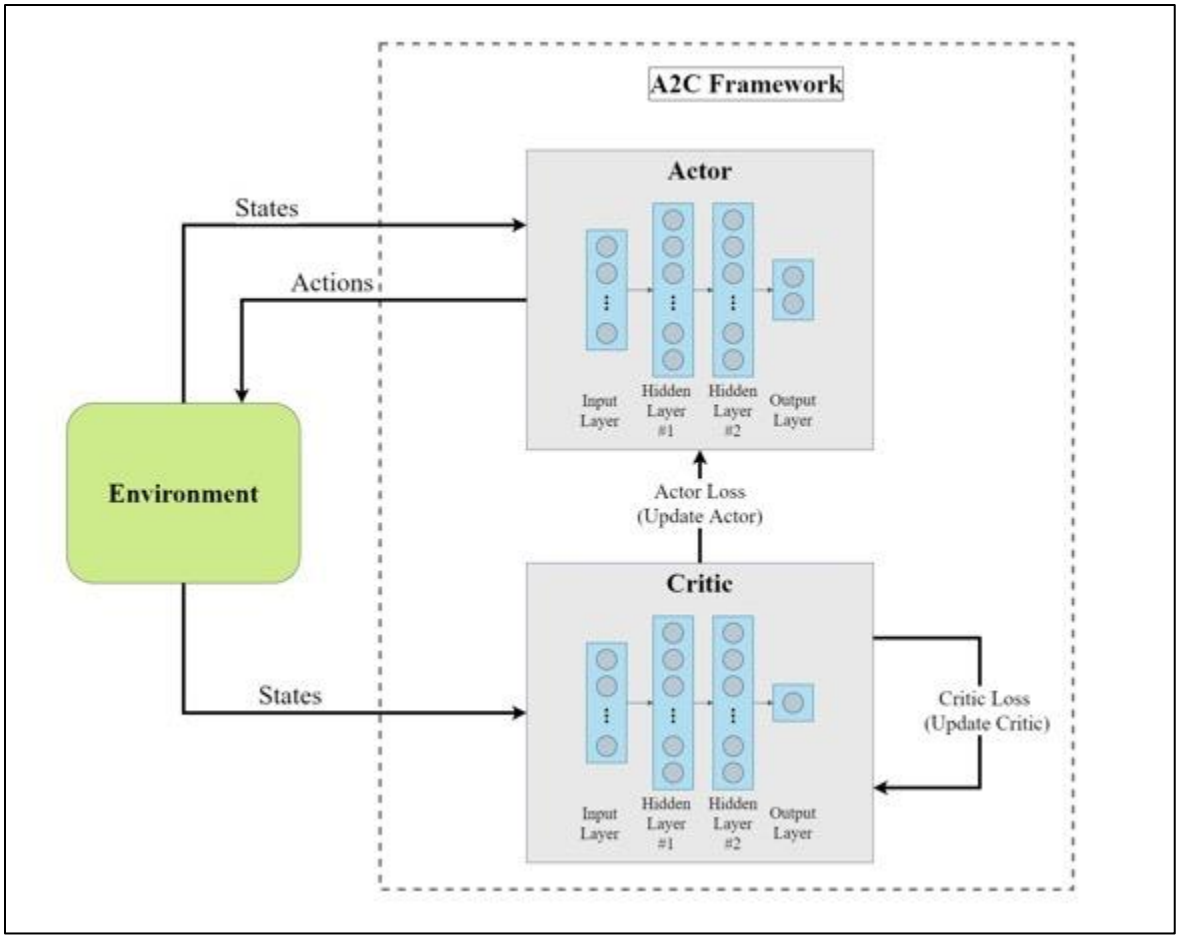


Figure 4.2 Illustration of the Actor-Critic simulation framework

4.4 Single-Agent Microsimulation Tool

In this research, a DRL-based simulation tool is developed to predict pedestrian trajectories in the validation dataset. Figure 4.3 illustrates the structure of the simulation tool. First, the environment of the mixed traffic is initiated with information about the vehicle movement over time. Then, the simulating agents (i.e., pedestrians) are initiated from their initial states. The pedestrian and vehicle trajectories of each interaction consist of their behavior around and during the conflict course, and these trajectories include their spatial coordinates, speeds, and heading directions. For each interaction, the initial states (i.e., initial spatial coordinates, speed, and heading direction) of road

users in the simulation tool are defined based on their trajectories in the validation dataset. The simulation tool then computes the states' features, including the longitudinal distance, lateral distances, speed differences between the conflicting road users, among others. Then, the simulating agents (i.e., pedestrians) select appropriate policy by sampling from the learned optimal A2C policy. In the developed simulation tool, the single-agent modeling approach is utilized. In this approach, pedestrians' movement and their evasive action mechanisms are considered for modeling, while considering the other interacting agent trajectories (i.e., vehicle) known over time.

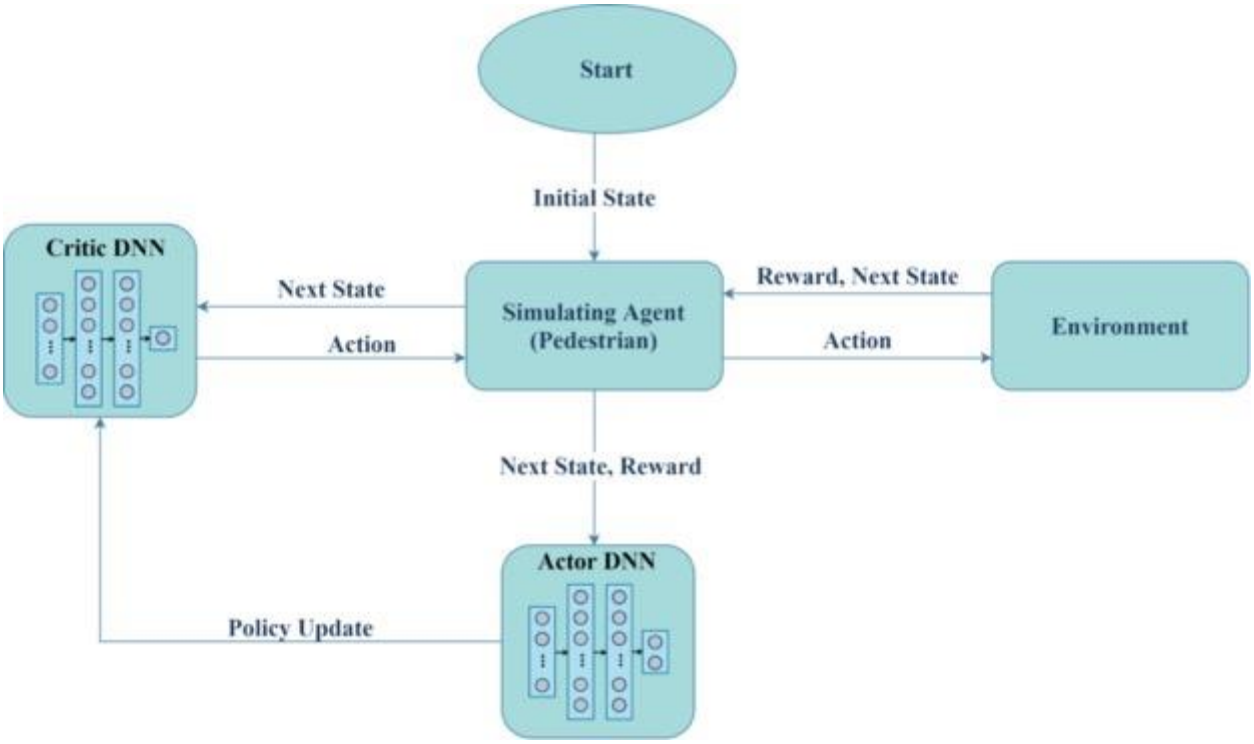


Figure 4.3 Single-agent DRL-based simulation tool

4.5 Results and Discussion

In this section, first, the recovered GP reward functions are presented. Then the simulated trajectories using the described simulation tool are evaluated. In this evaluation, the similarity of the predicted trajectories and the accuracy of the adopted collision avoidance mechanism are assessed.

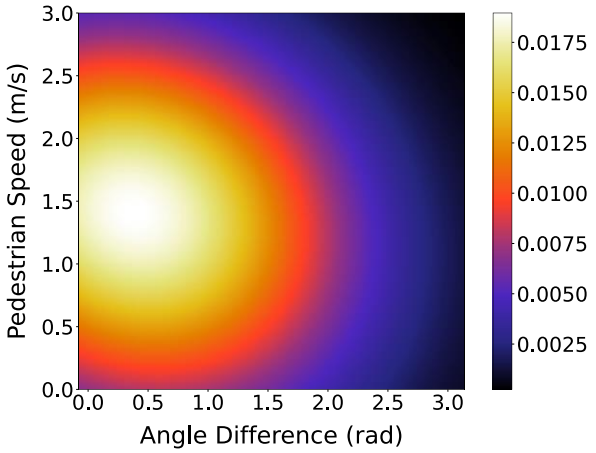
4.5.1 Reward Function Estimation

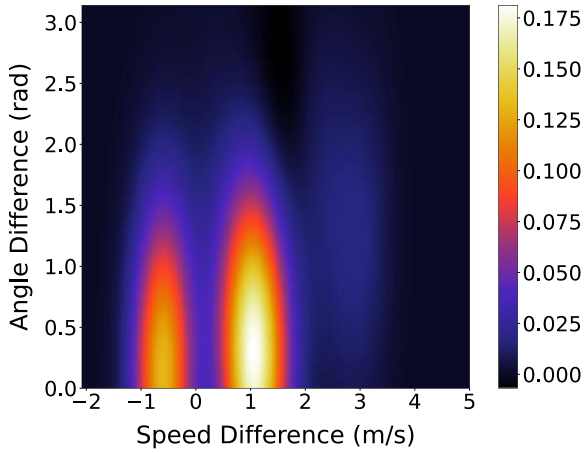
The estimated GP pedestrian reward function is presented in two-dimensional planer space, considering two-state features each time (Figure 4.4). The shape of the reward function can yield informative insights about the mechanism of the pedestrians' evasive actions and preferences. States with higher reward values are exponentially more preferred by pedestrians than the states with lower reward values. These states could be inferred as the highest preferred states that pedestrians try to reach by performing evasive actions in conflict situations. It is noteworthy to mention that the shape of the rewards clearly demonstrates the Gaussian Process.

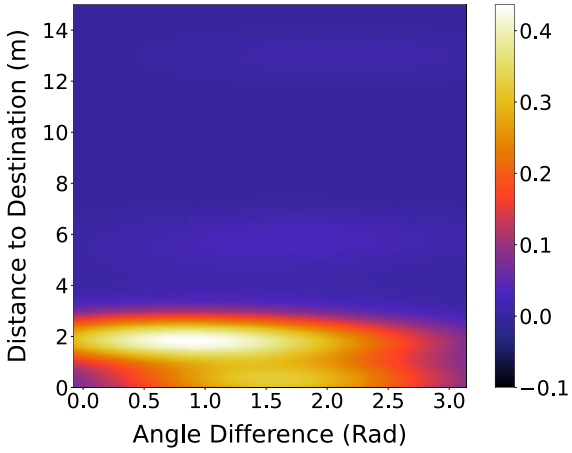
The lateral and longitudinal bivariate reward function shows that pedestrians prefer to maintain a safe lateral distance of about 2m while resolving the conflicts with vehicles. The GP reward function, which is presented using pedestrian longitudinal distance and angle difference state features, shows that the pedestrians prefer to implement a swerving maneuver to avoid a collision as the longitudinal distance decreases (i.e., pedestrians approach the conflict points). Moreover, after resolving the conflict situation (i.e., negative values of the longitudinal distance), pedestrians tend to return to their original movement direction. Other aspects of the pedestrian's evasive

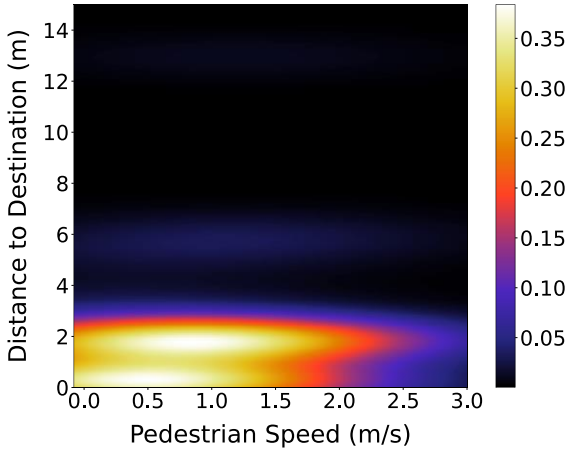
actions can be observed in their speed reward functions. The longitudinal distance and speed bivariate reward function shows two areas with higher rewards, which are located at positive and negative longitudinal distances. Lower speed values for the area with positive longitudinal distance suggest that pedestrians tend to have lower speeds when approaching the conflict compared to the post-conflict resolution state (i.e., area with negative longitudinal distance). This behavior shows pedestrians' inclination for deceleration to alleviate the crash risk.

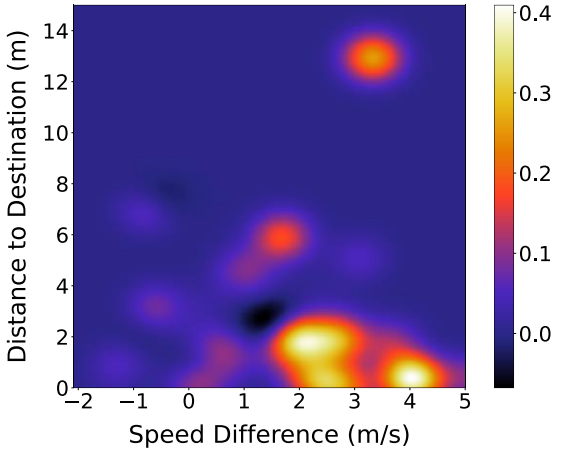
The GP bivariate reward function, which is presented using pedestrian longitudinal distance and velocity difference state features, demonstrates that the speed difference between pedestrians and vehicles tends to decrease as they approach each other (i.e., conflict point). However, at higher negative values of longitudinal distance, vehicles are more likely to increase their speeds, which lead to a higher difference with the pedestrians. Moreover, as illustrated in longitudinal distance and distance to target GP reward function, higher rewards are observed when pedestrians approach their destinations since areas with higher rewards are at lower values of the pedestrian distance to destination. The pedestrian reward function, illustrated using angle difference and speed, shows a higher preference for pedestrians towards lower speeds while maintaining an intermediate angle difference ($\pi/2$ rad). These preferences indicate that pedestrians prefer to decelerate to avoid the conflict situation instead of swerving under the current conflict situation.

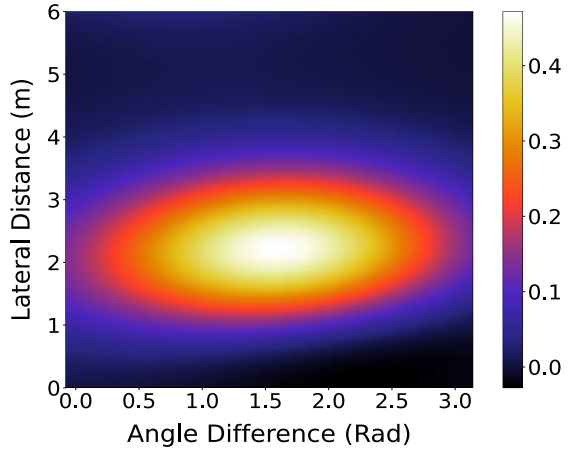

Pedestrian Speed (m/s)
Angle Difference (rad)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0.0025
0.0050
0.0075
0.0100
0.0125
0.0150
0.0175

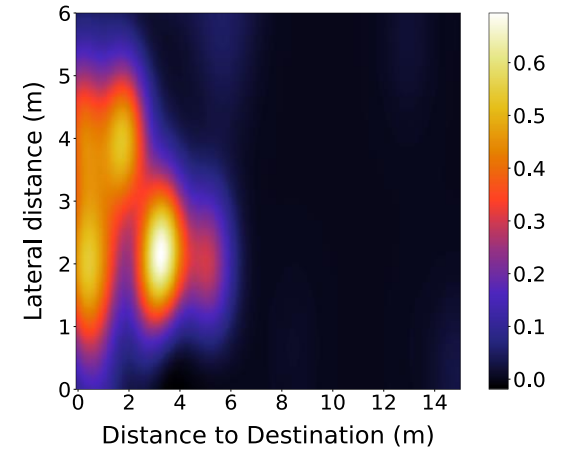

Angle Difference (rad)
Speed Difference (m/s)
−2
−1
0
1
2
3
4
5
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175

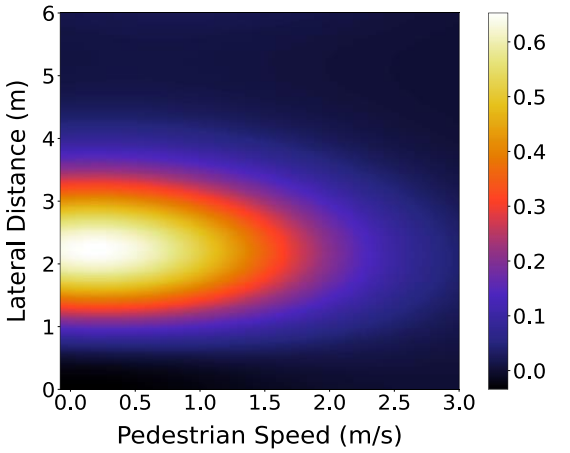

Distance to Destination (m)
Angle Difference (Rad)
0
2
4
6
8
10
12
14
−0.1
0.0
0.1
0.2
0.3
0.4

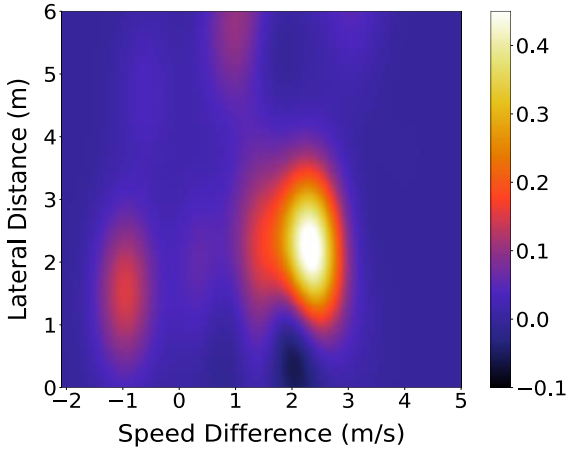

Distance to Destination (m)
Pedestrian Speed (m/s)
0.05
0.10
0.15
0.20
0.25
0.30
0.35

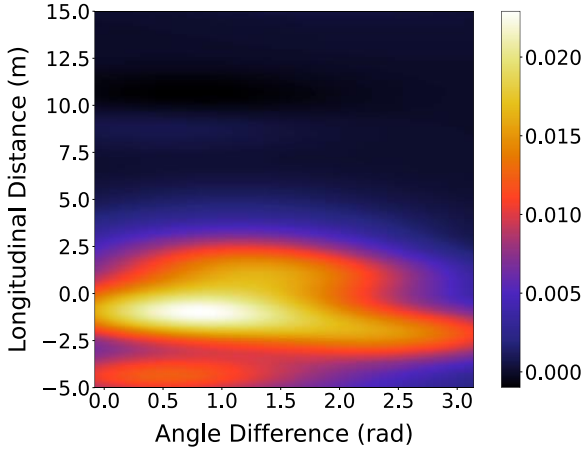

Distance to Destination (m)
Speed Difference (m/s)
−2
−1
0
1
2
3
4
5
0.0
0.1
0.2
0.3
0.4

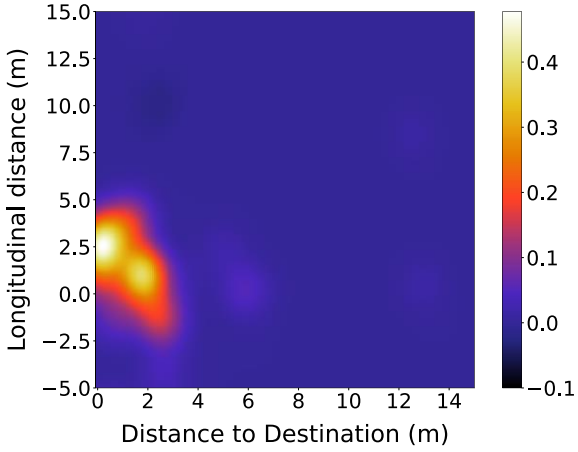

Lateral Distance (m)
Angle Difference (Rad)
0
1
2
3
4
5
6
0.0
0.1
0.2
0.3
0.4

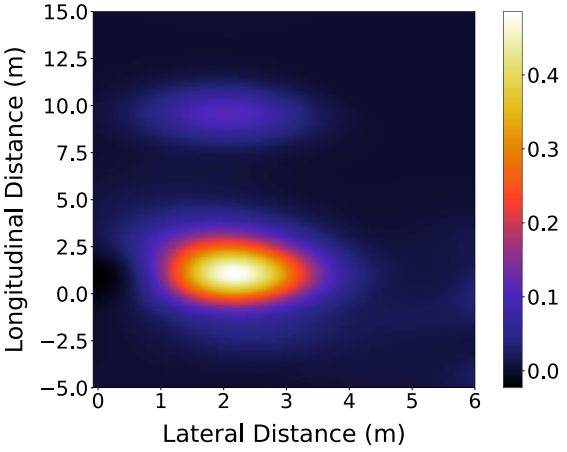

Lateral distance (m)
Distance to Destination (m)


Lateral Distance (m)
Pedestrian Speed (m/s)


Lateral Distance (m)
Speed Difference (m/s)


Longitudinal Distance (m)
Angle Difference (rad)


Longitudinal distance (m)
Distance to Destination (m)


Longitudinal Distance (m)
Lateral Distance (m)

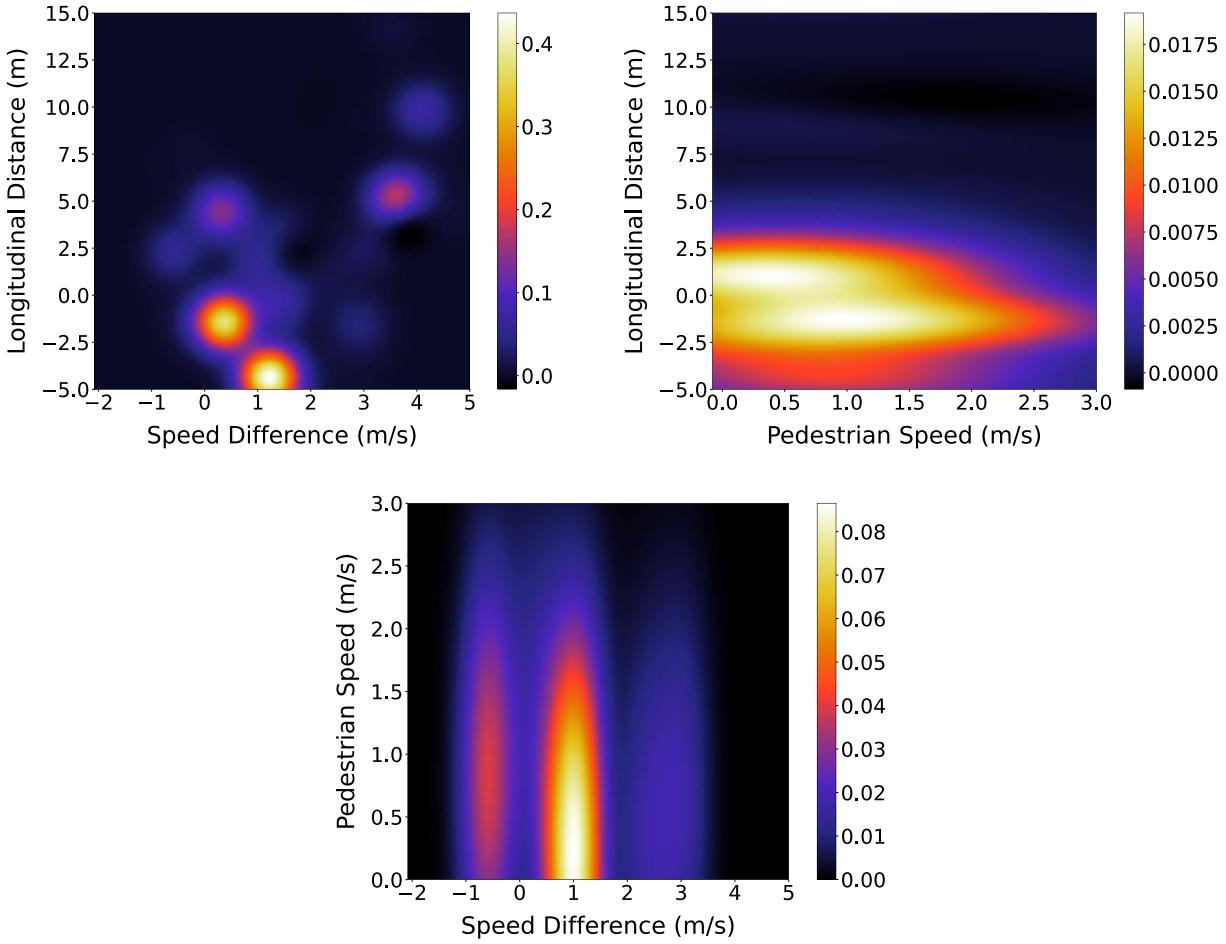


Figure 4.4 Gaussian Process pedestrian's reward function in pedestrian-vehicle conflicts

4.5.2 Evaluation Metrics

This study aims to develop a simulation model to replicate pedestrians and vehicles performance in conflicts. The accuracy of the produced results verifies the underlying rewards and policies. Therefore, the evaluation metrics are not limited to similarity measures, and additional measures are utilized to provide a comprehensive assessment of the model performance in terms of adopted evasive actions. To this end, the following evaluation approaches are used.

4.5.2.1 Similarity and Dissimilarity Measures

The simulated pedestrian trajectories' accuracy is assessed using the Mean Absolute Error (MAE) and the Hausdorff Distance (HD). The MAE is the average error between the simulation and actual pairs, which can be computed using equation (4.22).

$$\mathrm{MAE} = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \left| \mathrm{x}_{\mathrm{simulation,i}} - \mathrm{x}_{\mathrm{actual,i}} \right| \tag{4.22}$$

where $x_{simulation}$ and x_{actual} are the simulated and actual data points, respectively, and i is the time step-index.

The Hausdorff Distance (HD) (Huttenlocher, Klanderman, & Rucklidge, 1993) measures the error between real and simulated trajectories based on the mismatch's extent. The dissimilarity between trajectories matters in this method, and it considers the longest distance from each point in the actual trajectory to the simulated one. Unlike absolute error, which forms one-to-one mapping between the trajectories, HD builds many-to-many correspondences. The HD can be computed as given by equation (4.23).

$$\mathrm{HD}(A, B) = \max\left\{ \max_{a_i} \min_{b_i} d(a_i, b_i), \max_{b_i} \min_{a_i} d(a_i, b_i) \right\} \tag{4.23}$$

where $A = (a_1, \ldots, a_n)$ and $B = (b_1, \ldots, b_n)$ are the simulated and actual trajectories, respectively, d denotes the Euclidean distance.

4.5.2.2 Evaluation of the Collision Avoidance Strategies

The accuracy of the conflict severity indicator (i.e., PET) is assessed using the MAE measure. Moreover, Pearson's correlation is computed to test the strength of the linear dependencies between the simulated and the real trajectories PET. The collision avoidance strategy or evasive action mechanism is evaluated using the multiclass classification accuracy. The Confusion Matrix is used to evaluate the model's performance in identifying the type of the pedestrians' evasive action maneuvers. This method represents instances of actual and predicted classes in four groups. The true positive (tp) class, which occurs when the predicted positive value is true. Similarly, the true negative (tn) class presents the situation when the predicted negative value is true. Type 1 error or false positive (fp) situation occurs when the predicted positive value is false. On the other hand, a false predicted negative value is known as a type 2 error or false negative (fn). The pedestrians' evasive action patterns' accuracy can be assessed in the multiclass classification confusion matrix using equation (3.8) (Sokolova & Lapalme, 2009).

$$Accuracy = \frac{\sum tp + tn}{\sum tp + fn + fp + tn} \tag{4.24}$$

where tp, tn, fp, tp are the frequencies of each class.

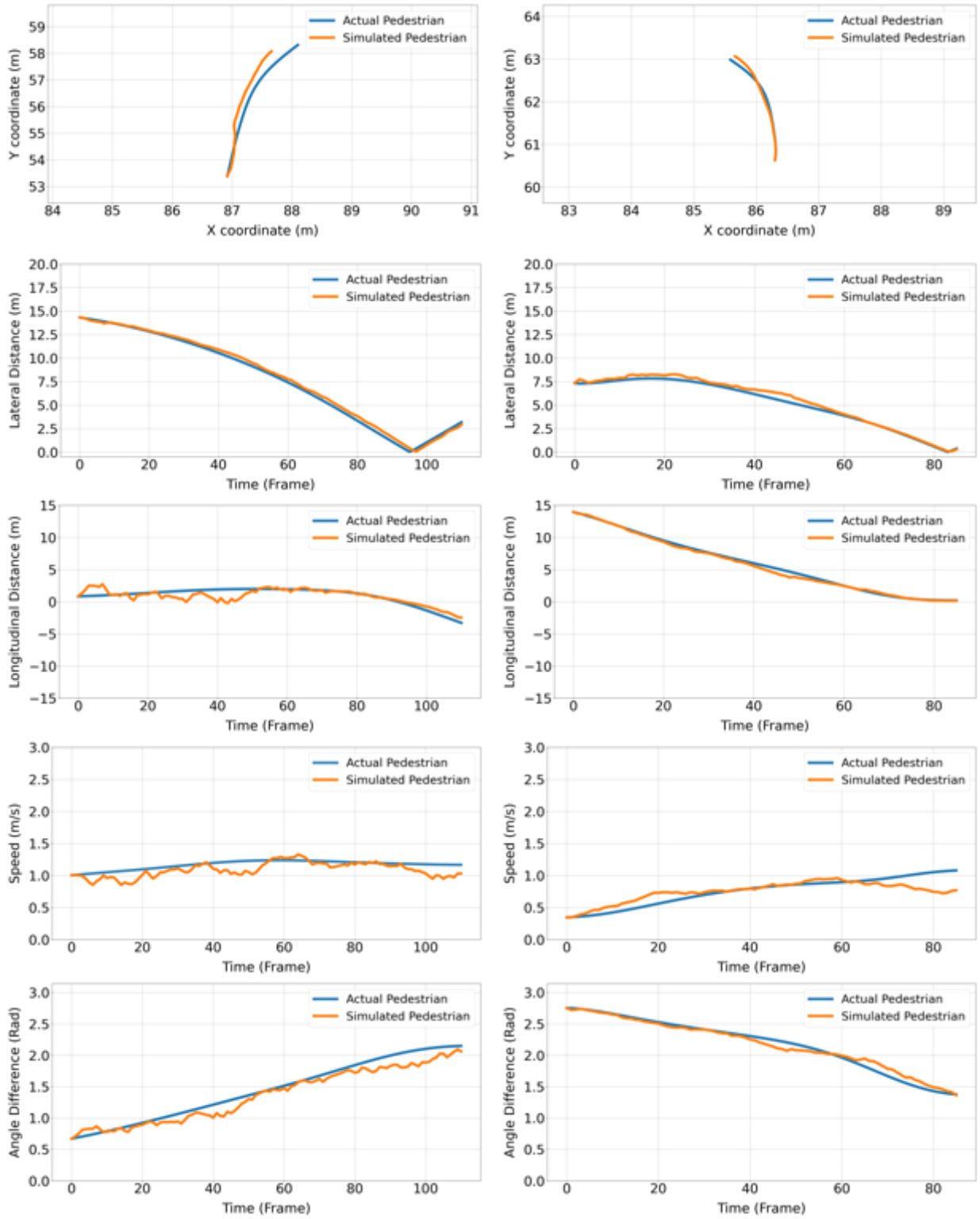
4.5.3 Trajectory Prediction

The trajectories used for performance evaluation and comparisons are the validation dataset trajectories, which consist of around 25% of the extracted trajectories (i.e., pedestrian trajectory ID from 156 to 203). The developed DRL-based simulation tool is used to obtain the optimum policy for trajectories, which can simulate the pedestrian evasive actions in conflicts with vehicles.

The validity of the simulated trajectories is assessed using the mean average error (MAE) and Hausdorff distance (HD) measures, as presented in Table 4.1. Generally, the model predicts pedestrian speed and angle difference more accurately than pedestrian position. Examples of the predicted pedestrian trajectories are presented in Figure 4.5. The figure presents actual and simulated pedestrians' longitudinal and lateral distances, velocity, and angle difference profiles. Generally, the model shows an acceptable performance in capturing the pedestrians' reactions in conflicts with vehicles in the mixed traffic condition. For example, Figure 4.5(a) and (b) show two examples of the simulated trajectories. Trajectory ID 176 illustrates a pedestrian swerving maneuver to the right (i.e., clockwise rotation), which is appropriately simulated by the model. Trajectory ID 201 shows a pedestrian adopted an acceleration evasive action following by a swerving maneuver to the left (i.e., counterclockwise rotation) to resolve the conflict situation. The simulation model predicts similar pedestrian evasive action mechanisms.

Table 4.1 Single-agent simulation error

Variable	Average MAE	Average HD
Lateral Distance (m)	0.61	1.03
Distance to Destination (m)	0.39	0.94
Longitudinal Distance (m)	0.75	1.40
Angle Difference (rad)	0.22	0.44
Speed (m/s)	0.34	0.65



(a) Trajectory ID 176 *(b) Trajectory ID 201*

Figure 4.5 Pedestrian's actual and predicted trajectories in pedestrian-vehicle conflicts

4.5.4 Traffic Conflict and Collision Avoidance Mechanism Evaluations

The changes in pedestrian speed (i.e., acceleration) and moving direction (i.e., yaw rate) are indicators of the evasive action mechanisms taken by pedestrians at near misses. Thus, predicting accurate road users' evasive action mechanisms is important for simulation model validity and safety evaluation. Therefore, the simulated trajectories' behaviors in the validation dataset (i.e., 48 conflict situations) are also evaluated based on the collision avoidance strategy. The road user evasive maneuvers were identified based on the acceleration and yaw rate profiles during the conflict course. Due to different pairs of pedestrians and vehicles with varying traveling conditions and speeds, the detection of evasive maneuvers was conducted manually according to the acceleration and yaw rate profiles and video scenes. A higher deceleration rate indicates a sudden change in speed in a short time. A high value of the yaw rate indicates a more significant change in the movement direction.

The confusion matrix is used to evaluate the model's performance in identifying the type of pedestrians' maneuvers at near misses. This method classifies each type of predicted evasive action/ maneuver into four categories. Table 4.2 and Table 4.3 present the model's prediction confusion matrix based on changing speed and swerving maneuver avoidance mechanisms, respectively. Generally, the results show that pedestrians tend to adopt an acceleration-based rather than a swerving maneuver in pedestrian-vehicle conflicts. The simulation model predicts pedestrians' swerving maneuver strategy more accurately than the speed's changing strategy. The developed simulation model predicted evasive pedestrian actions of changing speed and swerving maneuvers with accuracies of 75% and 77.1%, respectively.

The conflict indicator PET is computed and compared with corresponding values from the actual field-measured conflicts. The road users resolve conflict situations by adopting evasive actions. Thus, PET values can represent the similarity of the evasive actions in the actual and predicted models. The Pearson correlation coefficient can be used to examine the strength of the linear relationship between the real and the predicted PETs. The Pearson correlation coefficient between the predicted and actual PET values is 0.84. Moreover, the MAE is used to calculate the error between the simulated and actual PETs. The MAE between the real and simulated PETs is found to be 0.33 seconds. Figure 4.6 illustrates the predicted and actual PET values. In general, the figure shows that the GP-IRL model can obtain a reasonable estimate of the PET indicator at near misses, considering the complexity of pedestrian behavior at less organized mixed traffic situations. These results show the strong relationship between the severity of the simulated and the real trajectories PET.

Table 4.2 Acceleration-based evasive action Confusion Matrix

		Simulated pedestrian's behavior		
		Acceleration	Deceleration	No significant change
Actual Trajectories	Acceleration	21	1	3
	Deceleration	2	12	2
	No significant change	3	1	3

Table 4.3 Yaw Rate-based evasive action Confusion Matrix

		Simulated pedestrian's behavior		
		Positive yaw Rate	Negative yaw rate	No significant change
Actual Trajectories	Positive yaw Rate	12	1	1
	Negative yaw rate	2	17	1
	No significant change	3	3	8

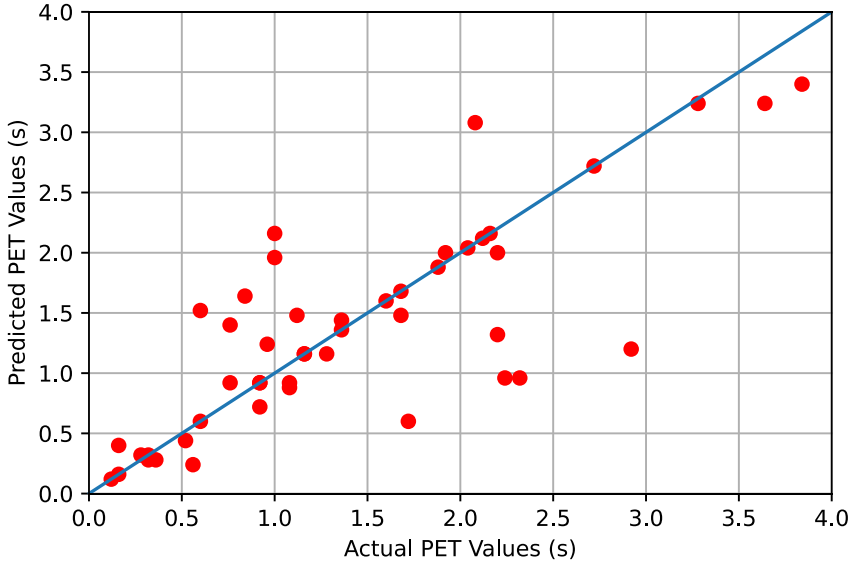


Figure 4.6 PET comparison between the actual and predicted values

4.6 Summary

In this chapter, a single-agent modeling framework was introduced. In this framework, a GPIRL algorithm is utilized to recover the pedestrians' reward function in conflict situations at mixed

traffic conditions. The Gaussian Process (GP) structure of the reward function allows describing the nonlinearity in road user behavior. A deep reinforcement learning advantage actor-critic method is employed to estimate pedestrians' optimal policies in conflicts with vehicles. Also, a simulation tool is developed to simulate pedestrians' trajectories using the obtained GP rewards and optimum policies. Finally, the simulated trajectories were evaluated using different similarity measures. Moreover, the exercised collision avoidance mechanisms in the simulations were compared to the actual ones.

The developed agent-based microsimulation model recognized pedestrians as rational decision-makers that acquire information from the environment to carry out ideal reactions. Overall, the results show that the developed agent-based microsimulation model can be used to model pedestrian trajectories and their evasive action mechanisms in conflict situations. The model predicted traffic conflicts (PET) with a strong linear correlation of 0.84 with the field-measured conflicts. Moreover, the model predicted pedestrian's evasive action mechanisms of swerving maneuver and speed changing with accuracies of 75% and 77.1%, respectively.

The main advantage of the developed simulation tool is in considering road users' intelligence and rationality and inferring their preferences in conflict situations using the recovered reward function. This enables a better understanding and analysis of road user behavior. However, it is challenging to infer pedestrian preferences in conflict situations using the commonly available simulation tools like VISSIM. Moreover, some concerns have been raised about using simulation tools such as VISSIM in analyzing traffic conflicts (Essa & Sayed, 2015a; Zheng, Sayed, Essa, &

Guo, 2019). The high sensitivity of the simulation model's outcomes (e.g., VISSIM) to parameters' calibration and the existence of the heuristic rules to prevent collisions between road users are the main shortcoming of most developed simulation packages, making them less reliable for safety investigations. The developed simulation tool in this study does not depend on predefined rules, and the model directly learns the reward function from actual demonstrations of human behavior. Lastly, the developed simulation tool can handle suboptimal human behavior, yielding more realistic results compared to conventional models.

Chapter 5: Multi-Agent modeling of Pedestrian-vehicle Interactions

This chapter presents the development of the multi-agent modeling framework to investigate pedestrian-vehicle interactions in mixed traffic conditions. In this framework, pedestrians and vehicles are simulated as intelligent agents under Markov Games (MG) approach that enables simultaneous modeling of multiple agents. The road users' underlying reward functions are obtained using a Multi-Agent Adversarial Inverse Reinforcement Learning (MA-AIRL) framework (Yu, Song, & Ermon, 2019). This method can simulate pedestrian and vehicle evasive actions using a continuous nonlinear reward structure. Moreover, the multi-agent optimum policies in conflict situations are obtained utilizing the multi-agent actor-critic using Kronecker-factored trust region (MACK) (Song, Ren, Sadigh, & Ermon, 2018). The optimum policies are then used to simulate the pedestrian and vehicle trajectories in conflict situations. The training and validation dataset used in the development and evaluation of the multi-agent framework is similar to the single-agent model, as described in Chapter 3.

This chapter is provided in seven sections. The first section describes the Markov Games modeling approach. The second section provides information on the evolution of the Adversarial Inverse IRL algorithms. Section three presents a detailed explanation of the MAIRL method. Section four describes the multi-agent actor-critic approach that has been used to obtain the optimal policies. Section five introduces a multi-agent simulation tool that enables the simultaneous estimation of pedestrian and vehicle trajectories. Section six presents the recovered multi-agent rewards, simulation results, and collision avoidance strategies. This section also compares the performance of the multi-agent framework with the single-agent approach. Finally, section seven provides a

summary of the simulation method, results, and further discussion regarding the performance of the developed simulation models (i.e., single and multi-agent frameworks).

5.1 Markov Games

The multi-agent simulation model employs Markov Game (MG) (Littman, 1994) approach to simulate pedestrian and vehicle interaction simultaneously. Markov games (S, A, P, γ, r, η) can be considered as a general form of MDPs, in which N agents interact with each other. In this framework, there is only one set of states for all agents. However, actions are defined separately for each agent. The general form of a Markov game consists of a set of states (S), and N set of action sets ($A_1, \dots, A_k, \dots, A_N$) for each agent in the environment. The transition process ($P(S)$) is a probability distribution over states that specifies the probability of a transition from state s to state s' taking actions ($a_1, \dots, a_N$). η is a subset of transition probabilities that specifies the initial state distribution. To compute the current value of the future rewards, a discount factor $\gamma \in [0,1)$ is used. In this multi-agent framework, a discount factor equal to 0.975 is used, similar to (Alsaleh & Sayed, 2020, 2021b). $r_{i,t}$ denotes the reward received by agent i in t timesteps ahead. In this chapter, notations are selected based on the game theory terminology. Bold variables indicate a vector containing all agent's associated variable (e.g., $\boldsymbol{r}$ denotes the reward vector), and subscript $-i$ refers to a vector composed of all other agents except for the i^{th} one. For instance, $\boldsymbol{r}$ can be represented by $(r_i, \boldsymbol{r}_{-i})$. In a Markov game, each agent has its own objective that is to maximize the expected sum of the discounted rewards through choosing the best policy. Expected rewards for agent i at time step t can be defined as equation (5.1) (Yu et al., 2019).

$$ExpRet_i^{\pi_i, \boldsymbol{\pi}_{-i}}(s_t, \boldsymbol{a}_t) = \mathbb{E}_{s^{t+1:T}, \boldsymbol{a}^{t+1:T}}\left[\sum_{l \geq t} \gamma^{l-t} r_i(s^l, \boldsymbol{a}^l) | s_t, \boldsymbol{a}_t, \boldsymbol{\pi}\right] \quad (5.1)$$

where $\boldsymbol{\pi}$ is the joint policy that creates a mapping between states to all agents' actions, π_i denotes the agent i policy, and $\boldsymbol{a}_t$ denotes the concatenation of all agents' actions.

5.2 Adversarial Inverse Reinforcement Learning

The single-agent model attempted to learn the rewards using an approximation of the MaxEnt IRL. Despite all benefits of the MaxEnt IRL, this method is not suitable in complex and high-dimensional spaces with unknown dynamics. Adversarial Inverse Reinforcement Learning (AIRL) was proposed as an alternative to MaxEnt for solving larger problems (Fu, Luo, & Levine, 2017). This method is developed based on Guided Cost Learning (GCL) (Finn, Levine, & Abbeel, 2016) and generative adversarial networks (Goodfellow et al., 2014).

A generative adversarial framework attempts to train a generator (G) and a discriminator (D) model, simultaneously. The generator is responsible for producing outputs similar to the actual data distribution. On the other hand, the discriminator attempts to distinguish the outputs of the generator from samples of the training data. Therefore, the objective of the discriminator network is to correctly classify the ground-truth data and outputs of the generator. The objective of the generative network is to capture the authentic distributions to maximize the probability of a mistake for the discriminator.

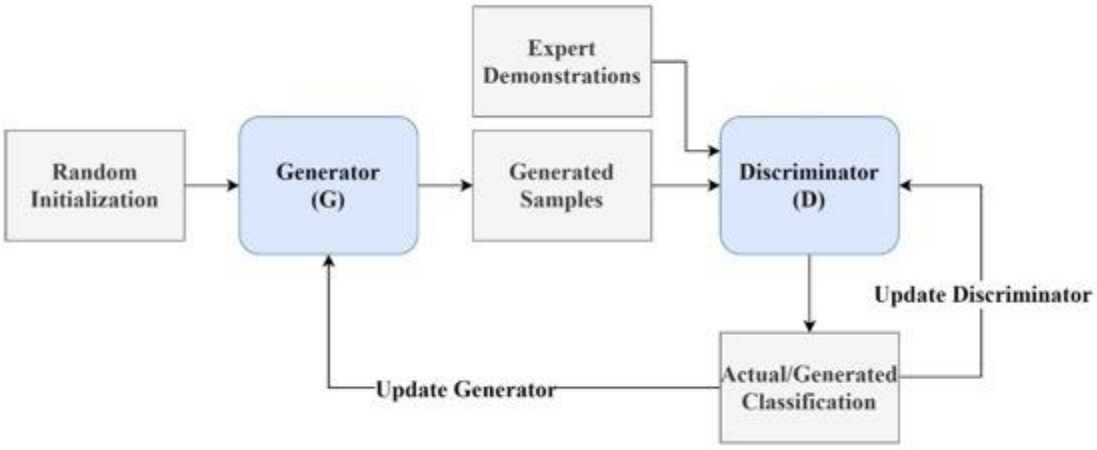


Figure 5.1 Illustration of the generative adversarial modeling framework

Using the generative adversarial modeling concept, (Ho & Ermon, 2016) proposed the Generative Adversarial Imitation Learning (GAIL) to estimate the policies. In this model, the discriminator is trained to differentiate the actual trajectories from the outputs of the policy. Therefore, the discriminator tries to maximize the probability of a valid classification, while the policy tries to produce outputs that the discriminator classifies as actual data. This algorithm is equivalent to a minimax problem that can be formulated as equation (5.2).

$$\min_{\theta} \max_{u} \mathbb{E}_{\pi_E}[\log D_u\,(s,a)] + \mathbb{E}_{\pi_E}[\log\big(1 - D_u(s,a)\big)] \tag{5.2}$$

where u and θ are the discriminator and policy parameters, respectively. GAIL lacks the ability to provide further information about the reward functions, as the discriminator always converges to 0.5. The idea of using generative adversarial networks in an IRL framework to obtain the rewards was proposed by (Finn, Christiano, Abbeel, & Levine, 2016; Fu et al., 2017). As oppose to GAILs, AIRL can provide rewards and policies, simultaneously. This algorithm assigns a particular form to the discriminator, as given by equation (5.3).

$$D_{\omega}(s,a,s') = \frac{exp\big(f_{\omega}(s,a,s')\big)}{exp\big(f_{\omega}(s,a,s')\big) + q(a|s)} \tag{5.3}$$

where $f_{\omega}(s,a,s')$ is the ω-parameterized reward function that indicates the reward of taking action a from state s to state s', and $q(a|s)$ denotes the adaptive sampling distribution. The adaptive algorithm tries to guide sampling towards areas with higher rewards to increase the similarity of the generated distributions and the actual observations. The sampled data distribution is then used as the input of the discriminator. Moreover, it has been shown that the partition function can be estimated by sampling distribution (Finn, Levine, et al., 2016). The objective of the discriminator can be defined similar to GAILs as a maximization function (equation (5.4)) to improve the binary real/fake classifications.

$$\max_{\omega} \log D_{\omega}(s,a,s') - \log[1 - D_{\omega}(s,a,s')] \tag{5.4}$$

The objective of the policy is to produce outputs that the generator classifies as the ground-truth data. Thus, the policy's objective can be defined against the objective of the discriminator, as given by equation (5.5).

$$\max_{\omega} \log[1 - D_{\omega}(s,a,s')] - \log D_{\omega}(s,a,s') \tag{5.5}$$

Eventually, the reward function is obtained from the optimal discriminator, and the optimal generator yields the optimal policy. To address the reward ambiguity problem (Ng, Harada, & Russell, 1999) (i.e., having multiple reward functions explaining one policy), (Fu et al., 2017) restricted the reward function ($f_{\omega,\phi}$) to a reward approximator (g_{ω}) as given by equation (5.6).

$$f_{\omega,\phi}(s,a,s') = g_{\omega}(s,a) + \gamma h_{\phi}(s') - h_{\phi}(s) \tag{5.6}$$

where h_ϕ is a shaping term parameterized by ϕ. In this approximation, rewards are not dependent on the dynamics of the environment as $g_\omega(s,a)$ is only influenced by the states. This state-only reward function can mitigate the ambiguity problem.

5.3 Multi-Agent Adversarial Inverse Reinforcement Learning

(Yu et al., 2019) generalized the AIRL solution concept to Markov games. In the multi-agent framework, it is assumed that in each state (s^t), agents select their optimal actions while other agents are fixed. Therefore, the algorithm iteratively optimizes each agent's policy. The transitions from states are defined in a Markov chain, in which the probability of future states for each agent is solely dependent on the current states. The probability of action a_i for agent i is defined by equation (5.7).

$$P_i\left(a_i \middle| \boldsymbol{a}_{-i} = \boldsymbol{z}_{-i}^{(k)}\right) = \frac{\exp\left(\lambda r_i(a_i, \boldsymbol{z}_{-i}^{(k)})\right)}{\sum_{a_i'} \exp\left(\lambda r_i(a_i', \boldsymbol{z}_{-i}^{(k)})\right)} \tag{5.7}$$

$$\boldsymbol{z}^{(k)} = \left(z_1, \dots, z_N\right)^{(k)} \tag{5.8}$$

where $\boldsymbol{z}^{(k)}$ denotes the state of the Markov chain at step k, z_i^k is the i-th player action at step k, and N is the total number of agents. $\lambda \in [0, \infty)$ is a parameter to control the agents' rationality. As λ decreases, agents tend to select random actions, whereas for higher values of λ agents' selection would be greedy. According to the previous studies (Yu et al., 2019), a λ value equal to one is selected. The objective of the policy in MA-AIRL is to produce trajectories with maximum similarity to the actual observations. (Yu et al., 2019) used Logistic Stochastic Best Response Equilibrium (LSBRE) to provide sampling distributions in the multi-agent framework. LSBRE is

defined as a sequence of joint stochastic policies at each timestep during the Markov game time horizon. The joint policy of the t-th Markov chain is given by equation (5.9).

$$\boldsymbol{\pi^t}(a_1, \dots, a_N) = P\left(\bigcap_i \left\{z_i^{t,(\infty)}(s^t) = a_i\right\}\right) \tag{5.9}$$

The training algorithm tries to improve the LSBRE distributions to maximize the likelihood of the expert trajectories. The probability of an LSBRE generated trajectory can be computed using equation (5.10).

$$p(\tau) = \eta(s^1) \cdot \prod_{t=1}^{T} \boldsymbol{\pi}^t(\boldsymbol{a}^t|s^t; \boldsymbol{\omega}) \cdot \prod_{t=1}^{T} P(s^{t+1}|s^t, \boldsymbol{a}^t) \tag{5.10}$$

In equation (5.10), $\boldsymbol{\omega}$ is the only parameter that can be modified to improve the LSBRE policies. Therefore, the likelihood problem can be written as equation (5.11).

$$\max_{\boldsymbol{\omega}} \mathbb{E}_{\tau \sim \boldsymbol{\pi}_E}\left[\sum_{t=1}^{T} \log \boldsymbol{\pi}^t(\boldsymbol{a}^t|s^t; \boldsymbol{\omega})\right] \tag{5.11}$$

This optimization can not yield the joint policy in a closed-form expression. (Yu et al., 2019) employed the pseudolikelihood technique (Besag, 1975) to approximate equation (5.11). Under this approximation, the joint likelihood can be estimated by the product of the agents' policies (π_i^t). Subsequently, the objective of the equation (5.11) can be simplified to equation (5.12).

$$\mathbb{E}_{\boldsymbol{\pi}_E}\left[\sum_{t=1}^{T}\sum_{i=1}^{N} \frac{\partial}{\partial \boldsymbol{\omega}} \log \pi_i^t(a_i^t|\boldsymbol{a}_{-i}^t, s^t; \omega_i)\right] \tag{5.12}$$

Also, optimizing the pseudolikelihood objective is equivalent to optimizing the surrogate loss, as given by equation (5.13) (Yu et al., 2019).

$$\mathbb{E}_{\boldsymbol{\pi}_E}\left[\sum_{t=1}^{T}\sum_{i=1}^{N}\frac{\partial}{\partial \boldsymbol{\omega}} r_i(s^t, \boldsymbol{a^t};\ \omega_i)\right] - \sum_{i=1}^{N}\frac{\partial}{\partial \boldsymbol{\omega}}\log Z_{\omega_i} \tag{5.13}$$

Equation (5.13) simplifies the multi-agent joint likelihood optimization to a function of single-agent rewards and partition functions. Similar to AIRL, MA-AIRL uses different adaptive samplers (q_θ) for different agents to compute the partition functions (Z_{ω_i}). Adaptive samplers play the role of the generators in the multi-agent algorithm that are trained to maximize the similarity between the generated distributions and the ground truth policies. Therefore, the algorithm can update the policies to maximize the likelihood of the actual trajectories, which in turn can fool the discriminators. For each agent, the structure of the discriminator is defined like AIRL discriminators' structure (equation (5.3)). Thus, the objective function of the discriminators and generators in MA-AIRL can be formulated as provided in equation (5.14) and equation (5.15), respectively.

$$\begin{aligned}\max_{\omega} \mathbb{E}_{\pi_E}\left[\sum_{i=1}^{N}\log\frac{\exp\left(f_{\omega_i}(s,\boldsymbol{a})\right)}{\exp\left(f_{\omega_i}(s,\boldsymbol{a})\right) + q_{\theta_i}(a_i|s)}\right] + \\ \mathbb{E}_{q_\theta}\left[\sum_{i=1}^{N}\log\frac{q_{\theta_i}(a_i|s)}{\exp\left(f_{\omega_i}(s,\boldsymbol{a})\right) + q_{\theta_i}(a_i|s)}\right]\end{aligned} \tag{5.14}$$

$$\begin{aligned}\max_{\theta} \mathbb{E}_{q_\theta}\left[\sum_{i=1}^{N}\log\left(D_{\omega_i}(s,\boldsymbol{a})\right) - \log\left(1 - D_{\omega_i}(s,\boldsymbol{a})\right)\right] \\ = \mathbb{E}_{q_\theta}\left[\sum_{i=1}^{N} f_{\omega_i}(s,\boldsymbol{a}) - \log\left(\ q_{\theta_i}(a_i|s)\right)\right]\end{aligned} \tag{5.15}$$

To address the reward ambiguity issue, the structure of equation (5.6) is used for the agents' reward function. Thus f_{ω_i} can be computed using equation (5.16).

$$f_{\omega_i,\phi_i}(s,a,s') = g_{\omega_i}(s,a) + \gamma h_{\phi_i}(s') - h_{\phi_i}(s) \quad (5.16)$$

where g_{ω_i} and h_{ϕ_i} are the i-th agent's reward estimator and shaping term. The algorithm updates the discriminators and the generators to the optimal point, where q_{θ_i} converges to the expert policy and g_{ω_i} approximates the ground-truth rewards.

5.4 Multi-Agent Actor-Critic with Kronecker Factors Deep Reinforcement Learning

This study employs a two-player multi-agent framework to simulate pedestrian-vehicle interactions at near misses. The agents are trained in a centralized setting where each player has information of the other player's states and actions; however, during the simulations, agents' decisions are solely dependent on their own states. In this study, the multi-agent environment setting is designed based on a competitive assumption in which agents try to maximize their own cumulative reward. In contrast, the cooperative environment can be used for the games where agents try to achieve a mutual objective, and accordingly, they collaborate to maximize a shared reward. In the developed multi-agent model, each agent tries to find the optimum policy so as to maximize its objective function (J), which is equal to the sum of the discounted rewards (equation (5.17)).

$$J(\theta_i) = \mathbb{E}_{\pi_{\theta_i}}\left[\sum_{l=t}^{T} \gamma^{l-t} r_i(s^l, \boldsymbol{a}^l)\right] \quad (5.17)$$

where θ_i denotes policy parameters of the i^{th} agent.

Traditionally, policy gradient methods have been used to optimize the policy parameters. Alternatively, natural policy gradient methods can improve the convergence during the training

phase (Kakade, 2001). However, this method can increase the training time as it uses second-order optimization. The Kronecker-factored approximate curvature (K-FAC) is a method proposed by (Martens & Grosse, 2015) that can efficiently approximate the natural gradient updates. Using K-FAC for actor-critic techniques, (Wu, Mansimov, Liao, Grosse, & Ba, 2017) proposed the actor-critic using Kronecker-factored trust region (ACKTR) method. This study utilizes a multi-agent version of ACKTR (a.k.a MACK) (Song et al., 2018) to train the adaptive samplers (i.e., agents' policies) according to the ground-truth policies.

In this framework, two deep neural networks are assigned to each agent, namely Actor and Critic. The Actor tries to estimate the associated agent's policy. The Critic tries to guide the policies by estimating the advantage function. The Critic network uses the state's expected reward as a comparison baseline for the Actor's estimated policies. Therefore, the difference between the value of the selected actions by the Actor and the expected return of the associated state provides a relative measure (i.e., advantage function) for the Critic to correct the Critic policies. Thus, the advantage function (A) of agent i, following policy π_{θ_i} can be written as equation (5.18).

$$A_{\zeta_i}^{\pi_{\theta_i}}(s_t, \boldsymbol{a}_t) = \sum_{j=0}^{k-1} \gamma^{l-t} r_i\left(s^{t+j}, \boldsymbol{a}^{t+j}\right) + \gamma^k V_{\zeta_i}^{\pi_{\theta_i}}\left(s^{t+k}, \boldsymbol{a}_{-i}^{t+k}\right) - V_{\zeta_i}^{\pi_{\theta_i}}(s^t, \boldsymbol{a}_{-i}^t) \qquad (5.18)$$

where V_{ζ_i} is the ζ-parameterized value function of agent i and π_{θ_i} is the θ-parameterized policy learned by the Actor. This algorithm employs K-FAC to approximate the natural policy gradients to provide update signals for Actors and Critics. Hence, the Kronecker-factored approximation enables updating both θ and ζ. An illustration of the multi-agent IRL algorithm used for pedestrian-vehicle interactions is provided in Figure 5.2.

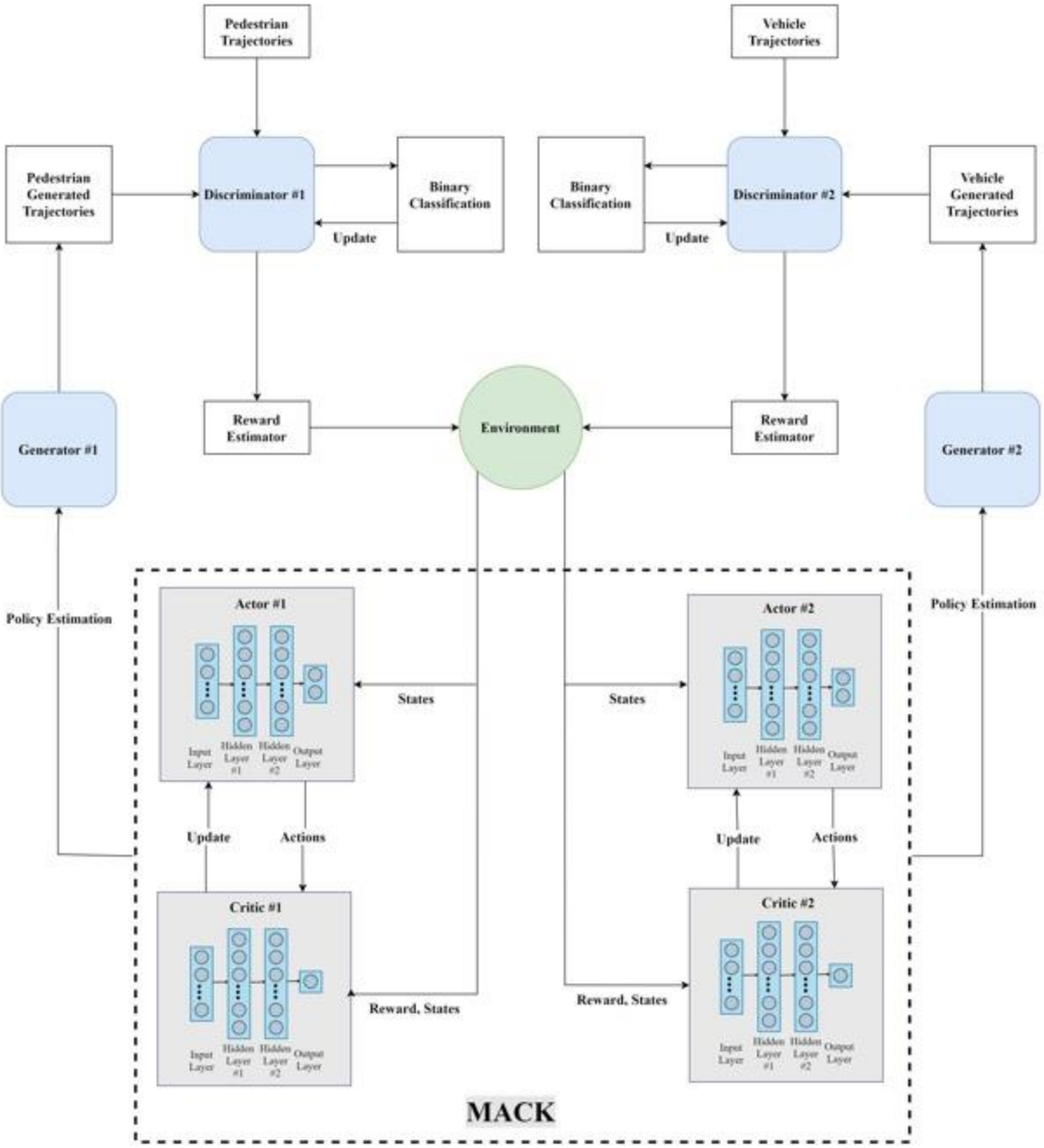


Figure 5.2 Pedestrian-vehicle multi-agent framework

5.5 Multi-Agent Microsimulation Tool

In this study, a multi-agent microsimulation tool is developed to predict pedestrian-vehicle interactions in conflict situations. In this simulation tool, pedestrian-vehicle multi-agent policies are obtained using the MACK DRL framework. The simulation environment initializes agents (i.e., pedestrian and vehicle) with information about their initial positionings, speeds, and heading directions to shape the conflict situation. The road users' initial situations are defined using their actual trajectories in the validation dataset. Then the simulation tool uses the initial information to compute the state features, including longitudinal and lateral distance, angle difference, speed, speed difference, and distances to targets. In the next step, the multi-agent framework employs the multi-agent optimal policy to predict the pedestrians' and vehicles' actions (i.e., acceleration and yaw rate). Finally, the simulation tool updates the state of the road users using the current step actions. For each trajectory, this process continues for the oncoming states until it reaches the terminating states (i.e., last time frame of the actual trajectory). Figure 5.3 illustrates a flowchart of the developed multi-agent simulation tool.

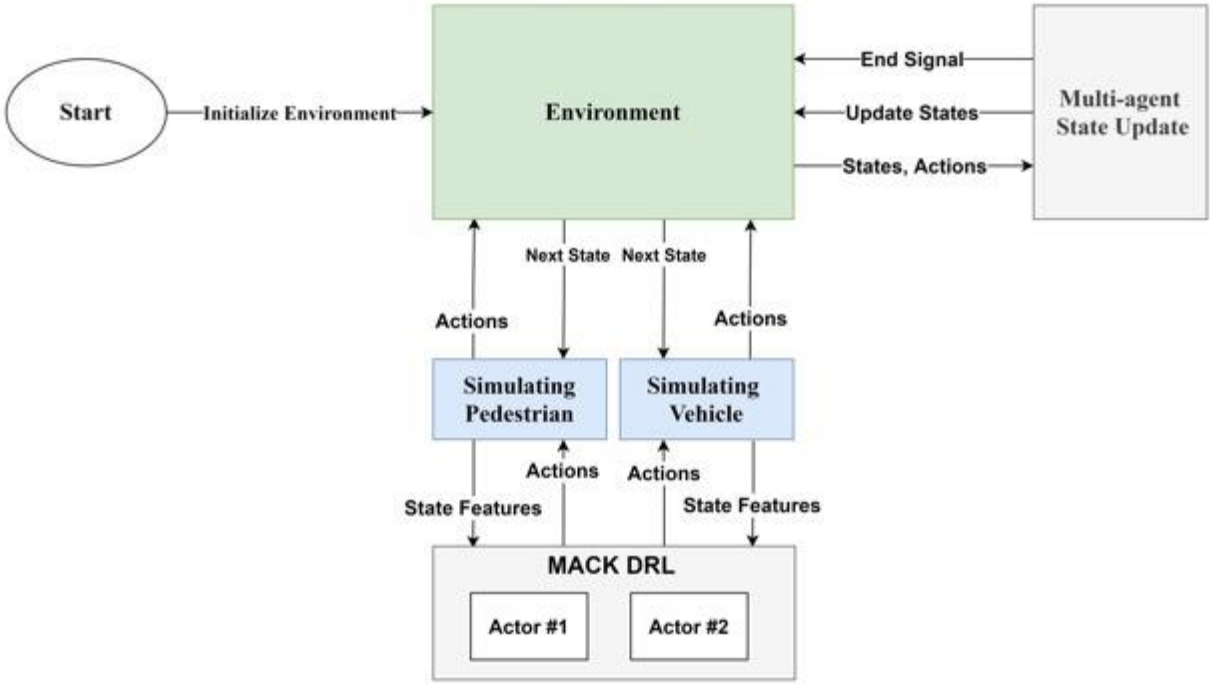


Figure 5.3 Multi-agent DRL-based simulation tool

5.6 Results and Discussion

This section presents the multi-agent simulation results, including simulated trajectories, conflicts indicators, and collision avoidance mechanisms. Furthermore, the accuracy of the simulated trajectories and performance of the model is evaluated and compared with the single-agent simulation framework. The evaluation metrics used in this chapter are the same as the single-agent framework, as previously explained in section 4.5.2.

5.6.1 Multi-Agent Reward Estimation and Comparison

The recovered pedestrian and vehicle reward functions estimated from applying the MA-AIRL algorithm are presented in Figure 5.4 and Figure 5.5, respectively. The reward functions are presented in two-dimensional state features space, while other dimensions (i.e., other state

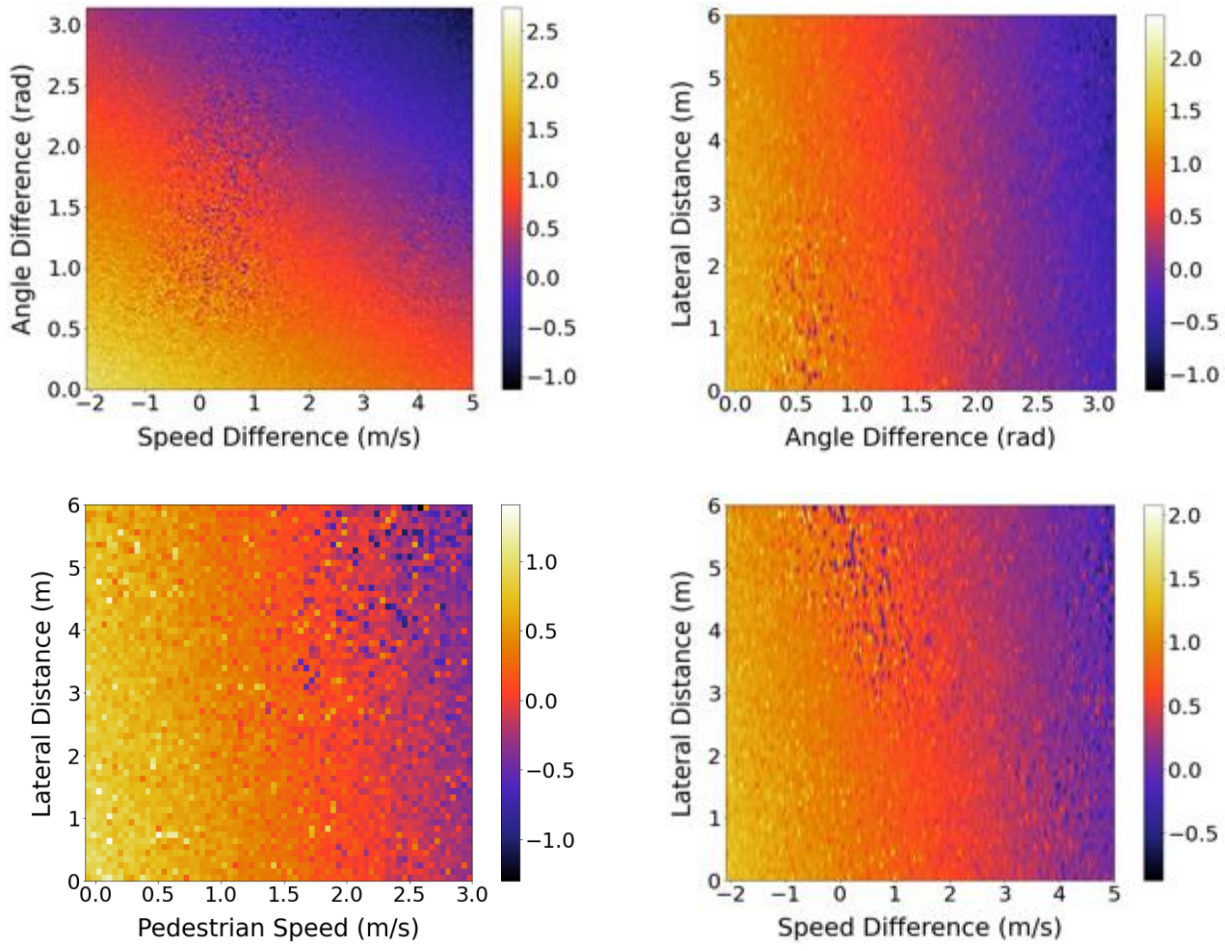
features) are fixed at their mean values. The multi-agent reward function for angle difference and speed shows that vehicles and pedestrians both have a high tendency for lower values of angle difference (i.e., road users moving in the same direction) at higher speeds. However, vehicles rewards show a wider range of preferences for angle difference compared to pedestrians. This behavior indicates a higher tendency for pedestrians to reduce the chance of head-on interactions with vehicles (i.e., angle difference close to π) at higher speeds, which might be attributed to the highly hazardous repercussions of the head-on interactions for vulnerable road users compared to vehicles. Since higher rewards can be obtained from a wider range of angle differences for vehicles, pedestrians are more likely to adopt swerving maneuvers to reach their preferred angle differences. The conflict resolving behavior of pedestrians and vehicles observed in angle difference selection is dependent on other interacting road users' heading direction. Therefore, the cumulative reward of each road user can be directly affected by the other road user's actions. Despite the differences in pedestrian and vehicle angle difference selection, their behaviors are in agreement with each other to reach lower angle difference values. Nevertheless, pedestrian and vehicle preferences are not always in agreement with each other, and some of the road users' rewards contradict each other (i.e., competitive behavior), which can be observed in situations where both road users have a tendency to cross the conflict point first.

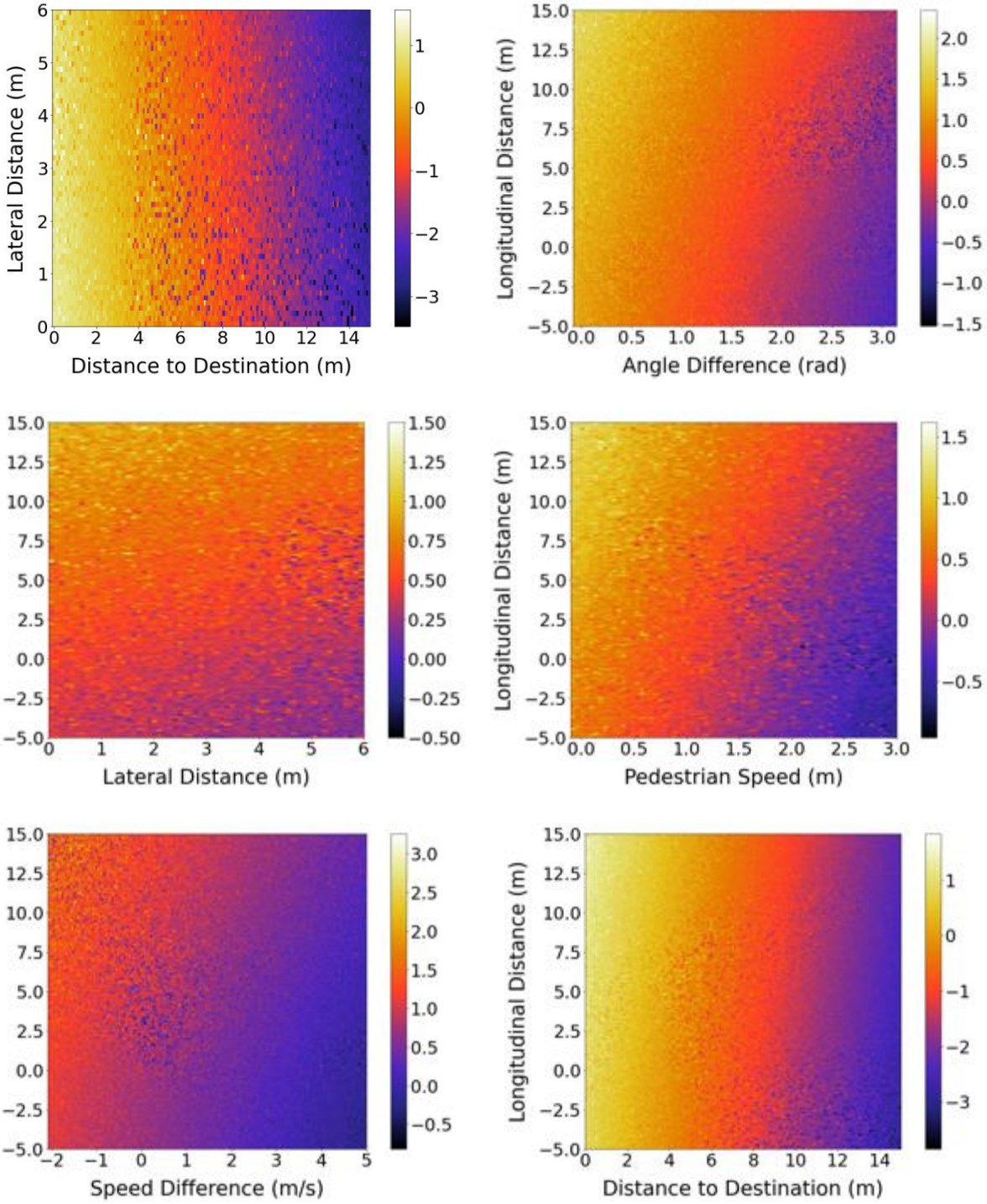
The competitive behavior of the road users to cross the conflict situation before the other interacting agent is clearly captured in the longitudinal distance versus speed difference multi-agent rewards. Pedestrian rewards show a tendency towards negative values of speed difference (i.e., pedestrian speed is higher than the conflicting vehicle) at negative values of longitudinal distance (i.e., post-conflict). However, an opposite preference can be observed in vehicles' utilities

for higher speed differences in post-conflict situations. Moreover, pedestrians are more likely to keep a positive speed difference before resolving the conflict situation (i.e., longitudinal distance is positive and close to zero). This can be interpreted as a situation where the approaching vehicle has higher velocity, and the pedestrian is taking precautions to yield to the vehicle at any point deemed necessary. Other aspects of this tradeoff between yielding or bypassing the approaching vehicle can be observed in the multi-agent speed and speed difference reward function. Comparing pedestrian and vehicle speed-speed difference multi-agent rewards suggests higher stochasticity of the pedestrian behavior for choosing the appropriate speed in the proximity of the conflict point, whereas vehicles show a clear tendency towards higher speed difference values. This stochasticity indicates pedestrians' cautious behavior that prefers to either yield to the approaching vehicle by reducing their speed or crossing the conflict point, being absolutely sure of the path safety.

The reward functions obtained using the GP-IRL algorithm generally demonstrate similar preferences with the MA-IRL algorithm for pedestrians' adopted collision avoidance behavior in different conflict situations. Nonetheless, some important aspects of the tradeoffs in road users' behavior are clearly neglected. For instance, the GP reward function, presented using the speed and speed difference state features, attempted to capture the stochasticity of the pedestrian behavior. However, due to the limitations of the GP function, the model had to specify separated high rewarding areas with gaps between them, while the multi-agent reward function could represent a better understanding of the behavior complexity in speed difference selection. Moreover, pedestrians' single-agent reward function, presented using lateral distance and speed, indicates a higher preference for lateral distances around 2 meters. However, as illustrated in pedestrians' multi-agent lateral distance and speed reward function, pedestrians can pass the

conflict point (i.e., negative longitudinal distances) with different lateral distances. Lastly, the pedestrian single-agent longitudinal distance and distance to destination reward function shows pedestrians' tendency towards the lower values of distance to destination while maintaining an intermediate longitudinal distance. However, the multi-agent reward function, which is presented using longitudinal distance and distance to destination, could capture the pedestrians' tendency to reach their destination without any unrealistic restrictions for longitudinal distance.




Lateral Distance (m)
Distance to Destination (m)
Longitudinal Distance (m)
Angle Difference (rad)
Longitudinal Distance (m)
Lateral Distance (m)
Longitudinal Distance (m)
Pedestrian Speed (m)
Longitudinal Distance (m)
Speed Difference (m/s)
Longitudinal Distance (m)
Distance to Destination (m)

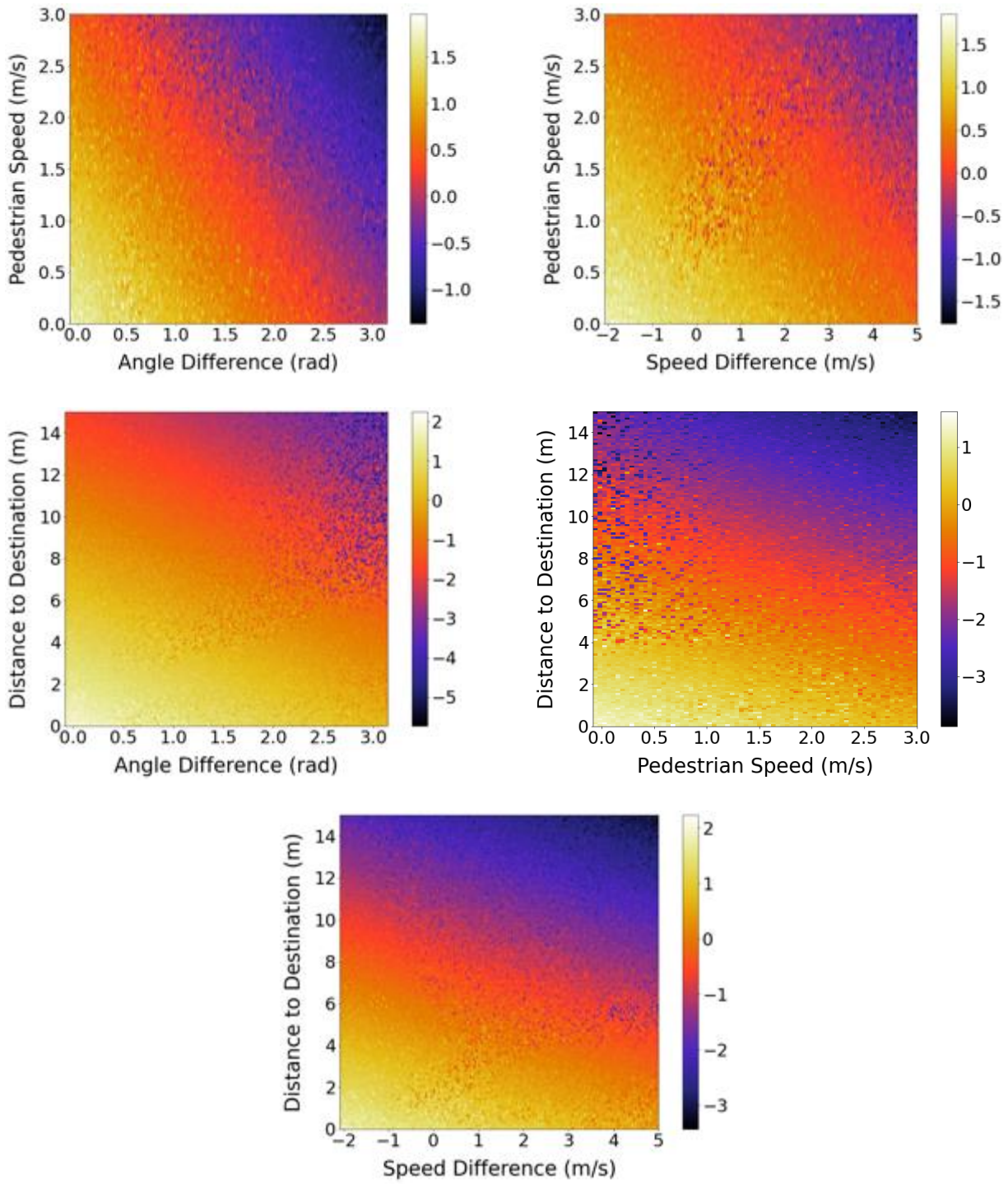
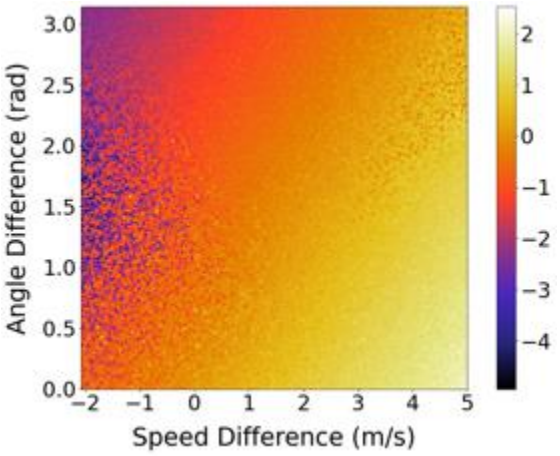
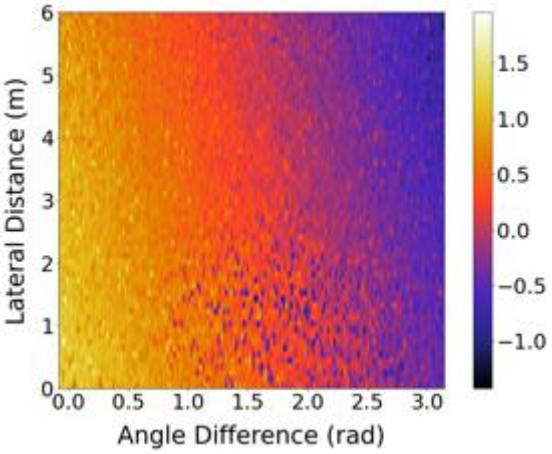
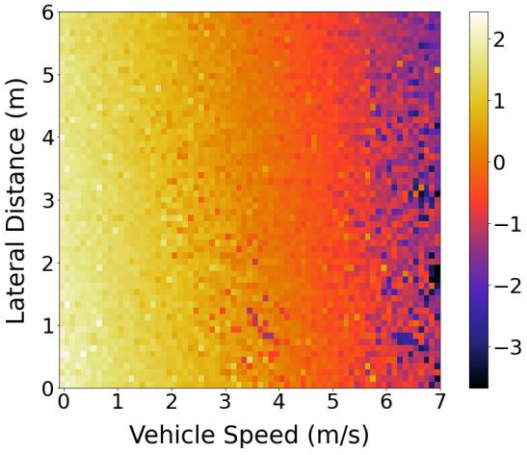
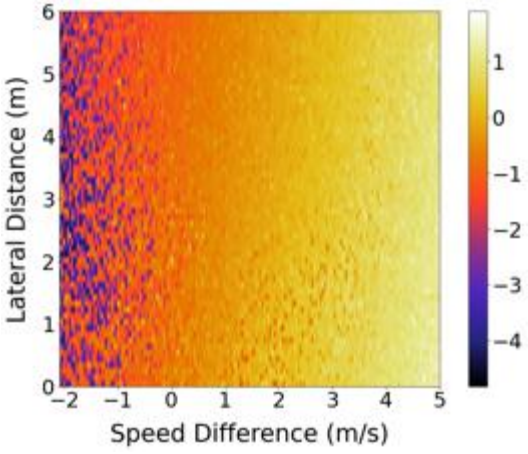


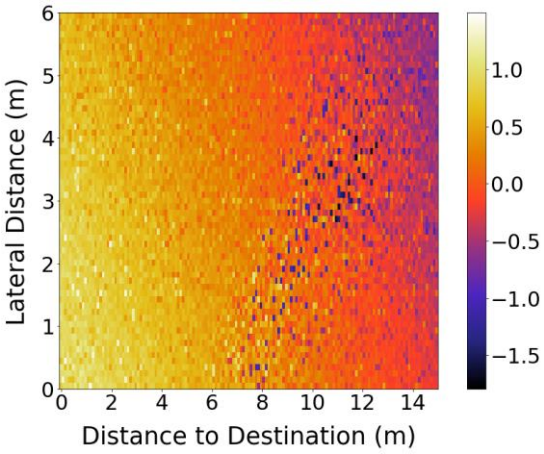
Figure 5.4 Pedestrian multi-agent bivariate reward function

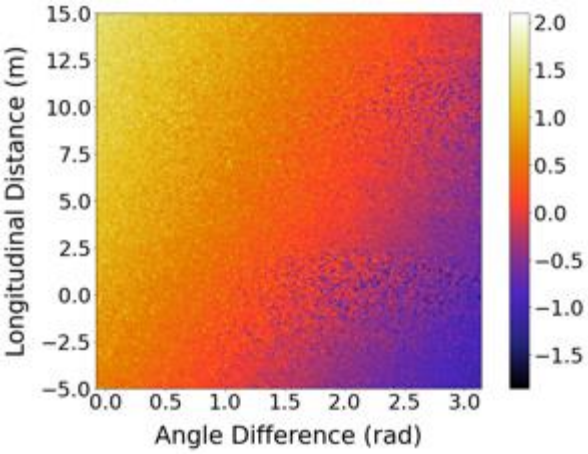

Angle Difference (rad)
Speed Difference (m/s)

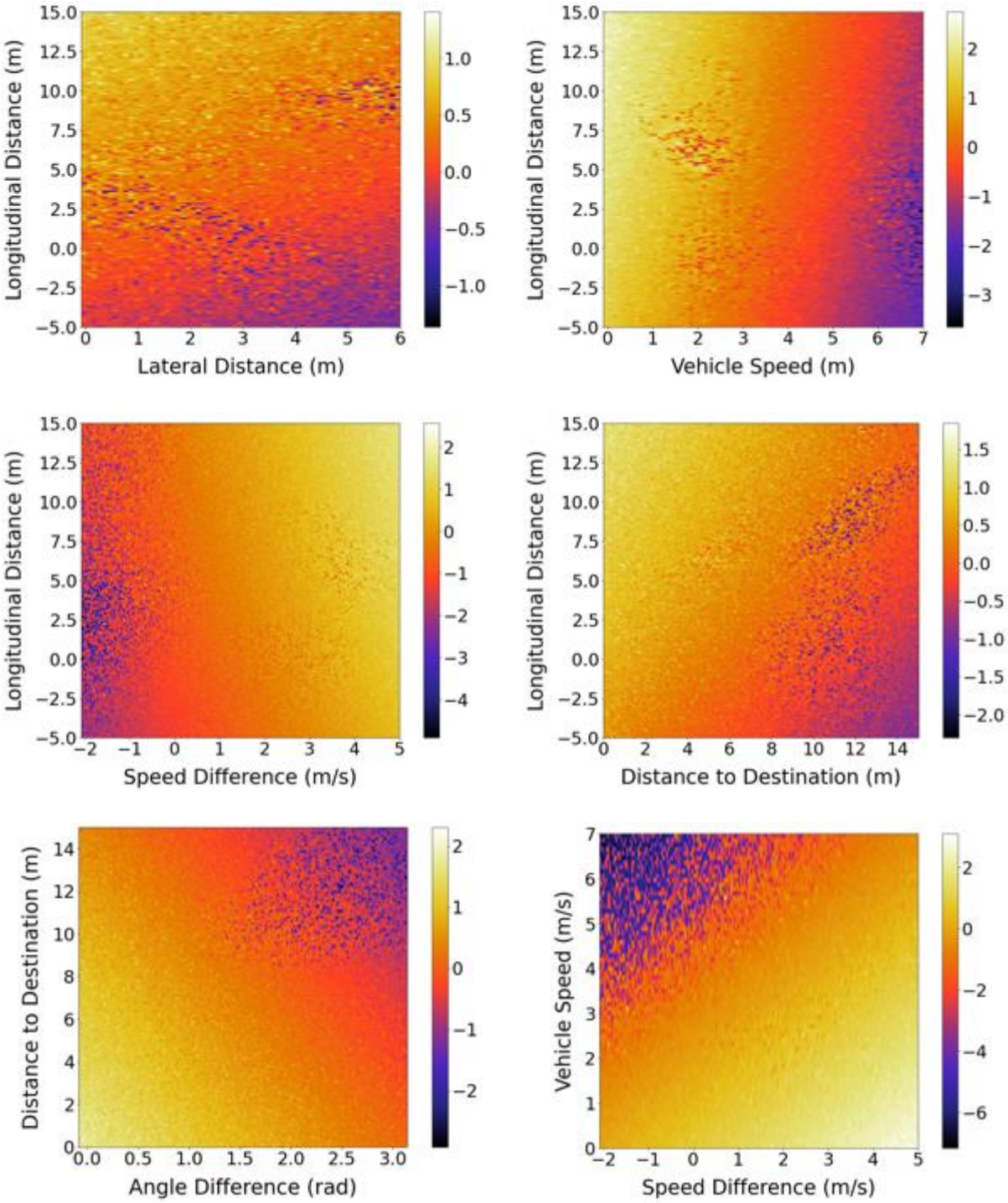

Lateral Distance (m)
Angle Difference (rad)


Lateral Distance (m)
Vehicle Speed (m/s)


Lateral Distance (m)
Speed Difference (m/s)


Lateral Distance (m)
Distance to Destination (m)


Longitudinal Distance (m)
Angle Difference (rad)


Longitudinal Distance (m)
Lateral Distance (m)
Longitudinal Distance (m)
Vehicle Speed (m)
Longitudinal Distance (m)
Speed Difference (m/s)
Longitudinal Distance (m)
Distance to Destination (m)
Distance to Destination (m)
Angle Difference (rad)
Vehicle Speed (m/s)
Speed Difference (m/s)

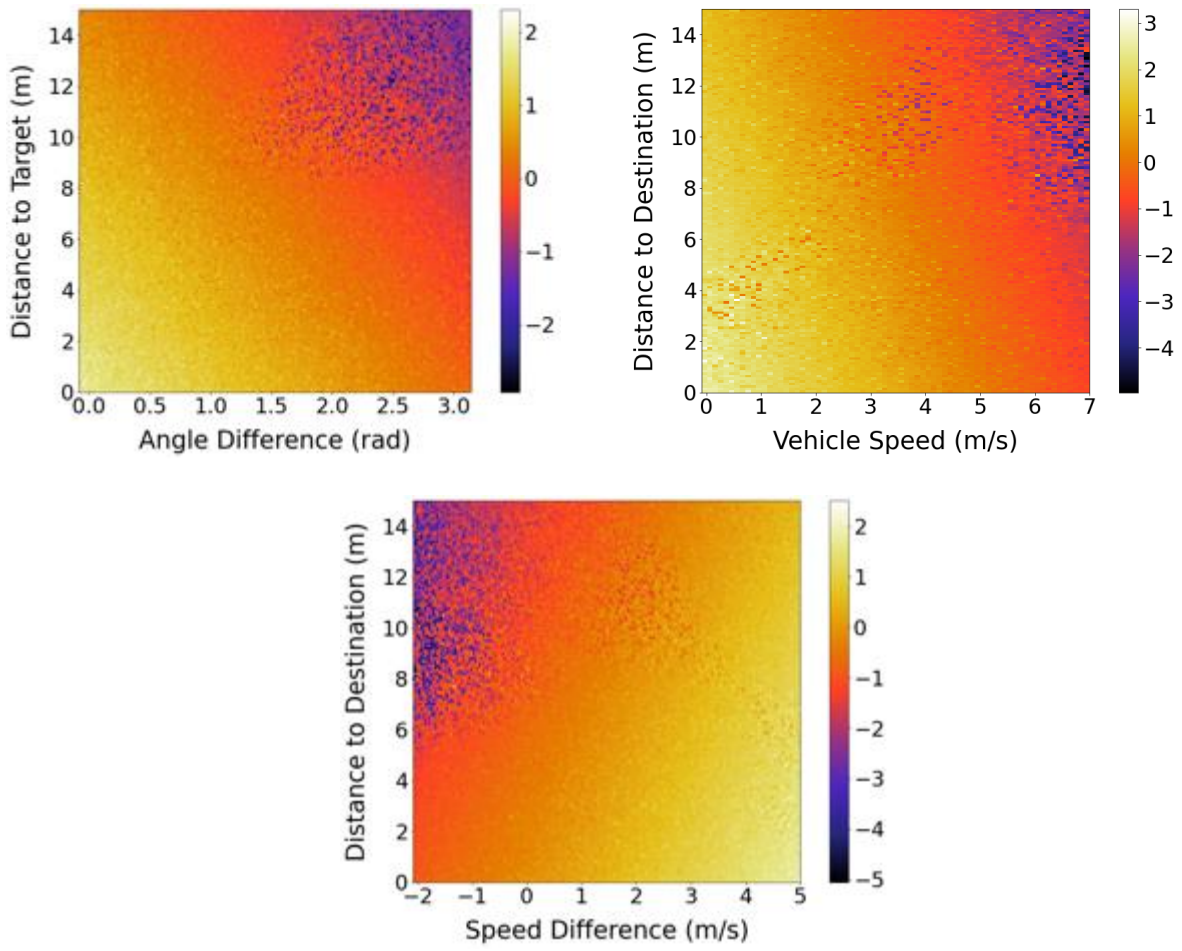


Figure 5.5 Vehicle multi-agent bivariate reward function

5.6.2 Trajectory Prediction and Comparison

The developed multi-agent simulation tool was utilized to predict the trajectories of the validation dataset. As opposed to the single-agent framework, which only considers pedestrians as simulating agents, the multi-agent model predicts both pedestrian and vehicle actions in the conflicts. The accuracy of the pedestrian and vehicle trajectories simulated by the multi-agent modeling approach are summarized in Table 5.1 and Table 5.2, respectively. Overall, the accuracy of the velocities and angle differences are higher than relative distances in both single-agent and multi-agent models. The multi-agent model's predictions significantly outperform the single-agent model. The

greatest improvement is related to the pedestrian's velocity, in which the multi-agent model performed 47% better than the single-agent model in terms of MAE. Besides, the MAE error for longitudinal distance, lateral distance, and angle difference improved about 17%, 10%, and 23% in the multi-agent model, respectively.

Table 5.1 Multi-agent pedestrian simulation errors

Pedestrian Variable	Average MAE	Average HD
Lateral Distance (m)	0.55	0.91
Distance to Destination (m)	0.22	0.67
Longitudinal Distance (m)	0.62	1.19
Angle Difference (rad)	0.17	0.29
Speed (m/s)	0.18	0.31

Table 5.2 Multi-agent vehicle simulation errors

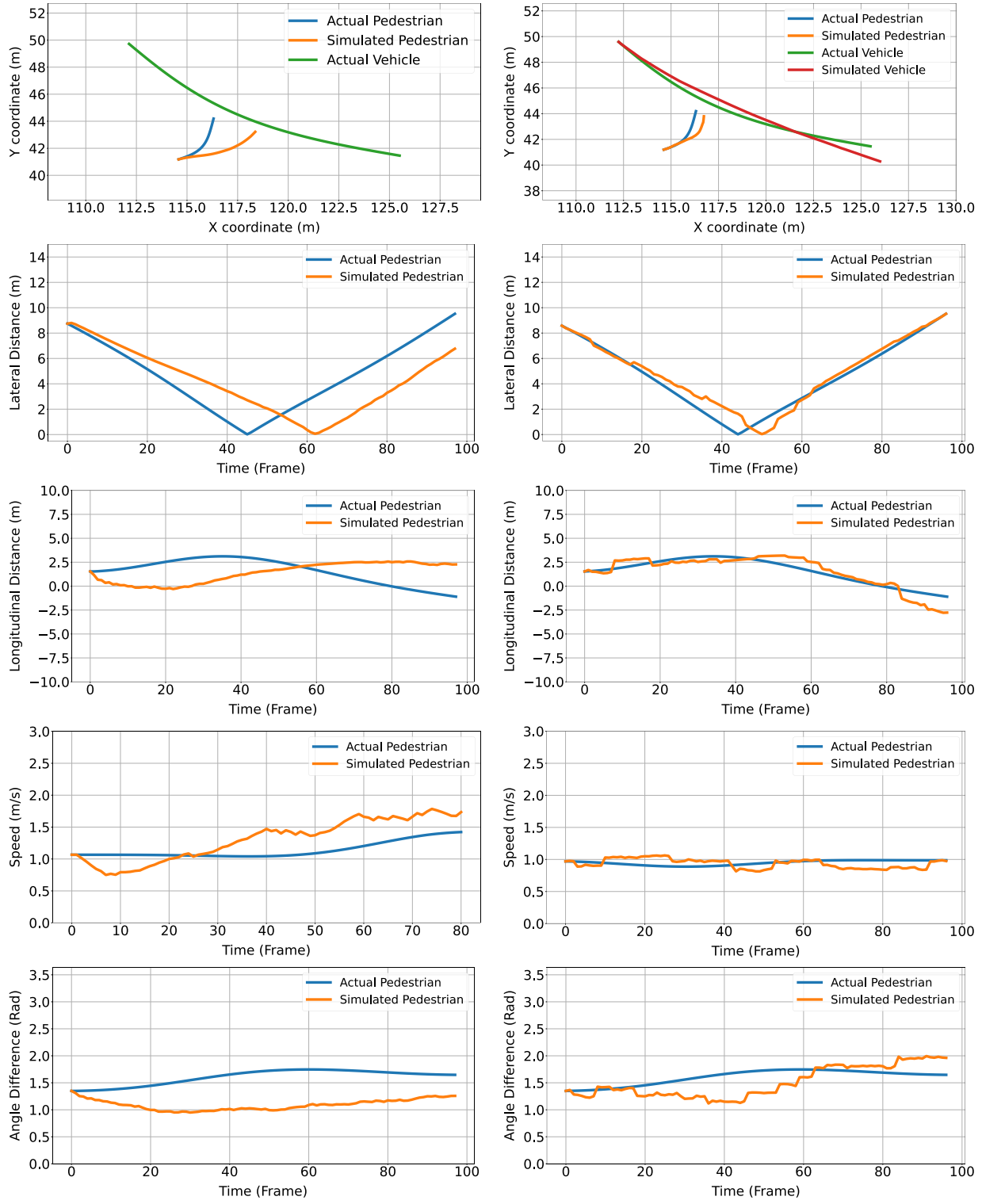
Vehicle Variable	Average MAE	Average HD
Lateral Distance (m)	0.83	1.11
Distance to Destination (m)	0.42	1.86
Longitudinal Distance (m)	0.51	1.18
Angle Difference (rad)	0.17	0.29
Speed (m/s)	0.35	0.72

Examples of the models' predictions are presented in Figure 5.6. Overall, results indicate that the multi-agent model can reflect the general collision avoidance behavior more accurately than the single-agent model. For instance, Figure 5.6(a) demonstrates a conflict situation where the pedestrian yielded to the approaching vehicle. In this interaction, swerving maneuver is the

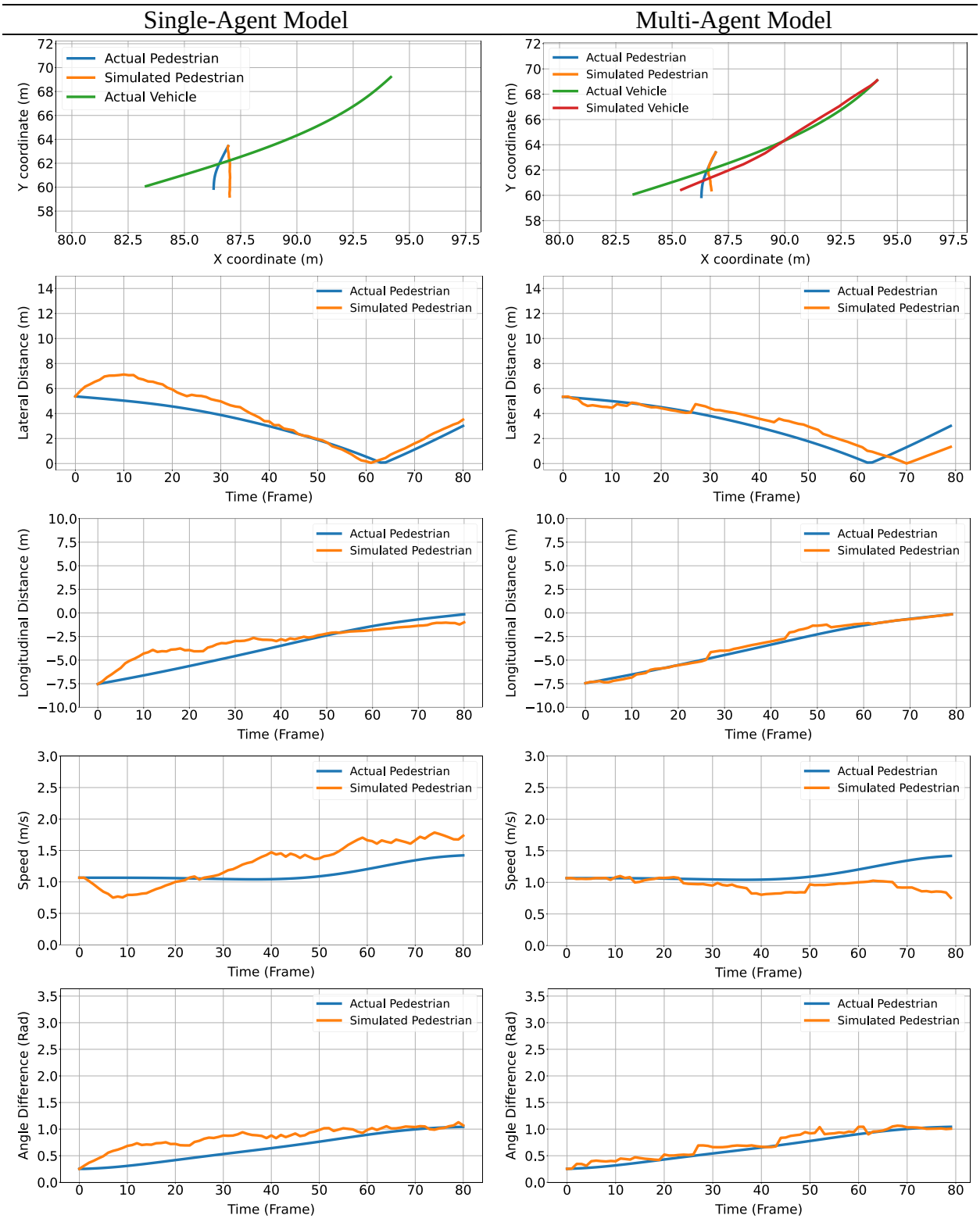
dominant emerged behavior of the pedestrian to provide enough space and time for the vehicle to clear the conflict zone. The multi-agent simulation tool understood the pedestrian yielding behavior and tried to resolve the conflict with similar evasive maneuvers to the actual pedestrian. However, the simulated pedestrian in the single-agent model started the reaction with an acceleration-based maneuver without clear swerving to the left. In this model, the simulated pedestrian had to finally decelerate since the pedestrian doesn't have enough speed to pass the conflict zone immediately. This misunderstanding of the single-agent model can increase the risk of collision. In accident-prone circumstances, there is a trade-off between different evasive maneuvers options. In this example, the single-agent model manifested a tendency to acceleration-based maneuver, whereas it wasn't the most efficient solution.

In another example, Figure 5.6(b) shows a situation where the vehicle yielded to a pedestrian by reducing speed near the conflict point. The actual pedestrian understood the yielding behavior of the driver and accepted the provided gap by crossing the conflict prior to the vehicle. To increase safety, the pedestrian implemented a counterclockwise swerving maneuver to provide enough gap distance in case of an unexpected behavior from the vehicle. Both single-agent and multi-agent models predicted the pedestrian intention to cross first. However, the single-agent model failed to capture the extra cautious swerving maneuver taken by the pedestrian, which is reflected in the angle difference figures. The more complicated structure of the rewards and policies in the multi-agent approach enabled capturing different aspects of the road users' collision avoidance behavior, which in turn led to a more precise reaction prediction.

Single-Agent Model	Multi-Agent Model



(a) *Trajectory ID 165*



(b) *Trajectory ID 203*

Figure 5.6 Pedestrian-vehicle interaction simulation predicted by Single-agent and multi-agent models

5.6.3 Collision Avoidance Mechanism Evaluation and Comparison

Pedestrian and vehicle collision avoidance behavior can be identified by two clues: a change in speed and a swerving maneuver (i.e., a change in heading direction). This study, rather than evaluating the simulated trajectories, analyzes the similarity of the exercised avoidance strategies. The collision avoidance strategies are categorized into two groups of acceleration-based and yaw-rate based maneuvers. Each of these subcategories is classified into three classes of positive, negative, and no change as elaborated in the single-agent framework. The pedestrian's collision avoidance classification confusion matrices are presented in Table 5.3 and Table 5.4. The multi-agent model led to a classification accuracy of 83.3% and 79.2% for the yaw-rate based and acceleration-based maneuvers, respectively. The results show an improvement of 6.2% in swerving evasive actions and 4.2% in the speed change maneuvers compared to the single-agent model. Also, results revealed pedestrians' slight tendency towards acceleration compared to deceleration.

Table 5.5 and Table 5.6 present the classification results for MA model vehicles' simulations. The simulation tool resulted in an accuracy of 81.3% for swerving maneuvers (i.e., yaw-rate based) and 83.3% for acceleration-based evasive actions. The results show that the vehicles are more likely to adopt an acceleration-based avoidance strategy rather than a swerving maneuver in conflict with pedestrians; the same pattern was observed in the model's simulations. Furthermore, the effectiveness of the employed evasive actions in ameliorating serious conflict was assessed by comparing the PET values extracted from field-measured and simulated trajectories. Figure 5.7 illustrates the scatter plot of the actual and predicted PET values for the multi-agent model. The multi-agent model led to persistently high precision, with an MAE value of 0.27 seconds for

predictions, whereas the MAE for the single-agent model was 0.33 seconds. Furthermore, the linear relationship between actual and predicted values was assessed utilizing the Pearson correlation coefficient, which led to a correlation of 0.89 for the multi-agent simulation model. Therefore, the superiority of the multi-agent model was also observed in replicating collision avoidance mechanisms.

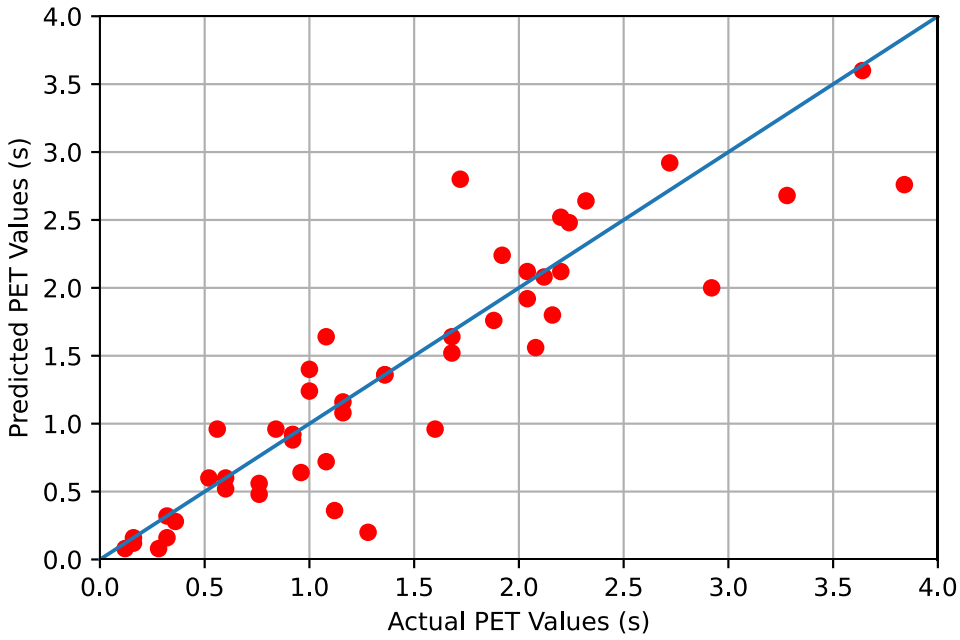


Figure 5.7 PET comparison between the actual and predicted values

Table 5.3 Pedestrians' acceleration-based maneuvers confusion matrix simulated by multi-agent framework

		Simulated pedestrian's behavior		
		Acceleration	Deceleration	No significant change
Actual Trajectories	Acceleration	21	1	3
	Deceleration	3	12	1
	No significant change	1	1	5

Table 5.4 Pedestrians' swerving maneuvers confusion matrix simulated by multi-agent framework

		Simulated pedestrian's behavior		
		Positive yaw Rate	Negative yaw rate	No significant change
Actual Trajectories	Positive yaw Rate	13	1	0
	Negative yaw rate	3	17	1
	No significant change	1	2	10

Table 5.5 Vehicles' acceleration-based maneuvers confusion matrix simulated by multi-agent framework

		Simulated vehicles' behavior		
		Acceleration	Deceleration	No significant change
Actual Trajectories	Acceleration	23	0	3
	Deceleration	2	10	2
	No significant change	1	0	7

Table 5.6 Vehicles' swerving maneuvers confusion matrix simulated by multi-agent framework

		Simulated vehicles' behavior		
		Positive yaw Rate	Negative yaw rate	No significant change
Actual Trajectories	Positive yaw Rate	12	1	2
	Negative yaw rate	2	18	2
	No significant change	0	2	9

5.7 Summary

In this chapter, a novel multi-agent simulation framework was introduced to model pedestrians' avoidance behavior in interaction with vehicles in mixed traffic conditions. The developed multi-agent simulation tool employed the MA-AIRL algorithm to recover the pedestrian and vehicles reward functions in conflict situations. In this framework, road users were modeled under the Markov Game approach that considers road users (i.e., pedestrian and vehicle) as intelligent individuals who can choose appropriate actions in conflict situations. Moreover, a multi-agent actor-critic deep reinforcement learning technique was applied to estimate the road users' optimal policies that enables simultaneous simulation of the pedestrian and vehicle trajectories. The proposed multi-agent framework was designed based on a competitive assumption that allows agents to prioritize their goals in the simulation environment.

The multi-agent simultaneous modeling of enabled realistic simulation of the exchanged actions and reactions between pedestrians and vehicles occurring in conflicts. Unlike the single-agent model that can not provide appropriate vehicle reactions in response to pedestrians' behavior, the

multi-agent approach can handle the equilibrium between road users' hastiness and cautiousness. The hasty behavior of the pedestrians is attributed to the situations where the vulnerable road users attempt to pass the conflict zone prior to vehicle arrival that can increase the risk of a conflict. Vehicles also demonstrated hasty behavior where they tried not to yield to the pedestrians. The cautious characteristic of the pedestrians' behavior is related to the situations where they try to yield to the approaching vehicles or accept the offered zone with enough precautions (e.g., providing enough gap distance). The results demonstrate that the multi-agent model can properly recognize this tradeoff for both simulating agents (i.e., vehicle and pedestrian) and provide appropriate reactions accordingly.

Comparison of the single-agent and multi-agent simulated trajectories revealed the superiority of the multi-agent framework. Simulation errors, adopted collision avoidance strategies, and conflict indicators (i.e., PET) in the multi-agent approach were consistently better than the single-agent framework. The superiority of the multi-agent framework could be attributed to two main reasons. First, the multi-agent framework employed artificial neural networks to describe the reward function, which gives much more flexibility compared to a specific gaussian process structure in understanding the agents' intentions. Second, the multi-agent actor-critic framework takes advantage of a more advanced K-FAC approximation technique that can enhance the performance of the optimal policies.

Chapter 6: Conclusions and Future Research

This chapter is divided into two sections. The first section presents a summary and conclusions of this thesis. The second section explains the limitations of this research and potential future studies.

6.1 Summary and Conclusions

The main objective of this thesis was to understand and simulate pedestrian evasive action mechanisms in mixed traffic conditions. A signalized intersection in the city of Shanghai, China was selected for this study. Video data were collected at the study location, and then computer vision video processing algorithms were employed to extract road user trajectories from the video scene. The obtained trajectories were utilized to compute elements of road user behavior. The primary variables used in this thesis to describe pedestrian and vehicle behavior in conflict situations were selected based on previous studies. These parameters include longitudinal distance to the other interacting agent, lateral distance to the other interacting agent, angular difference between road users' heading directions, speed, speed difference, and distance to the destination point.

In this study, two simulation frameworks were proposed to model pedestrian-vehicle conflicts at less organized mixed traffic environments. In the first simulation framework, pedestrians' decision-making process was modeled using an MDP approach. Moreover, a Gaussian Process Inverse Reinforcement Learning technique was applied to recover the pedestrian's single-agent reward functions, which enabled understanding pedestrian's preferences and rationales in conflicts with vehicles. In the single-agent framework, an actor-critic DRL approach was implemented to

obtain pedestrians' optimal policies and simulate their collision avoidance mechanisms in conflict situations.

The second simulation framework was developed based on the MG modeling approach. The MG approach enabled simultaneous simulation of pedestrian and vehicle decision-making in conflict situations. In this framework, a novel multi-agent Adversarial Inverse Reinforcement Learning technique was utilized to recover road users' multi-agent reward functions. Also, road users' optimal policies were obtained using a multi-agent version of ACKTR method. The developed multi-agent framework was designed based on a competitive approach that enabled agents to maximize their own objectives (i.e., total collected rewards) through interactions with other agents in the environment.

The developed simulation models (i.e., single-agent and multi-agent) directly obtained the reward functions from actual demonstrations of human behavior, which might not be optimal. The proposed simulation frameworks could handle the suboptimal behavior of the road users in conflict situations. Both simulation models demonstrated the ability of the IRL techniques in understanding the intentions of the road users in conflict situations. However, the results showed the superiority of the multi-agent model in simulating road users' trajectories based on different evaluation metrics, including similarity measures, conflict indicator (i.e., PET), and collision avoidance behavior. This could be attributed to the assumption of fixed trajectories for vehicles in the single-agent model. This assumption is likely to confuse the simulating pedestrian since it can not receive logical reactions from the approaching vehicle. Besides, the existence of the adversarial neural

networks enabled the multi-agent framework to better handle the complex behavior of the road users in comparison with the GP reward structure of the single-agent model.

6.2 Study Limitations and Future Research

There are several limitations to this study. First, the variables used to simulate pedestrians' evasive actions are mainly based on their behavioral characteristics and proximity measures with other involved road users in the conflict. Other variables that might increase the explained heterogeneity, including road user gender, age, and physical characteristics can be considered explicitly in future studies. Due to potential social demographic dependency of the evasive action patterns, a further investigation to evaluate the model's transferability in different locations and contexts is required.

The second limitation is that most of the traffic interactions in this work took place in relatively high densities, which makes it difficult to include density in the model. It should also be noted that some of the variables used in the modeling could implicitly account for the environment density, such as the speed of the opponent road user. It will be interesting for future studies to investigate the effects of change in road user density and compare the results using a more comprehensive dataset. Also, traffic conflicts might occur between more than two interacting road users (e.g., two pedestrians and one vehicle). Therefore, incorporating the effect of the adjacent pedestrians or vehicles, which can restrict the movement freedom, can add more complexity to the obtained reward functions and policies. Considering the group effect is an essential task for future research that can enhance the model’s performance in highly congested areas.

The third limitation to this study is related to the data collection methods. The video analysis computer vision algorithm used in the data collection was developed based on a feature-tracking method. However, applying more advanced computer vision algorithms such as Convolutional Neural Network (CNN) trackers can improve the quality of the extracted road users' trajectories, which can enhance the simulations' performance. Lastly, the conflicts selected for this study were based on a specific PET threshold of 4 seconds. However, it would be interesting for future research to develop models using different PET thresholds and other conflict indicators (e.g., TTC) and compare the results.

Bibliography

Abbeel, P., & Ng, A. Y. (2004). *Apprenticeship learning via inverse reinforcement learning.* Paper presented at the Proceedings of the twenty-first international conference on Machine learning.

Almodfer, R., Xiong, S., Fang, Z., Kong, X., & Zheng, S. (2016). Quantitative analysis of lane-based pedestrian-vehicle conflict at a non-signalized marked crosswalk. *Transportation research part F: traffic psychology and behaviour, 42*, 468-478.

Alsaleh, R., & Sayed, T. (2020). Modeling pedestrian-cyclist interactions in shared space using inverse reinforcement learning. *Transportation research part F: traffic psychology and behaviour, 70*, 37-57.

Alsaleh, R., & Sayed, T. (2021a). Markov-game modeling of cyclist-pedestrian interactions in shared spaces: A multi-agent adversarial inverse reinforcement learning approach. *Transportation Research Part C: Emerging Technologies, 128*, 103191.

Alsaleh, R., & Sayed, T. (2021b). Microscopic Modeling of Cyclists Interactions with Pedestrians in Shared Spaces: A Gaussian Process Inverse Reinforcement Learning Approach. *Transportmetrica A: transport science*, 1-28.

Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). Concrete problems in AI safety. *arXiv preprint arXiv:1606.06565*.

Amundsen, F., & Hyden, C. (1977). Proceedings of first workshop on traffic conflicts. *Oslo, TTI, Oslo, Norway and LTH Lund, Sweden*, 78.

Archer, J., & Young, W. (2010). Signal treatments to reduce the likelihood of heavy vehicle crashes at intersections: microsimulation modeling approach. *Journal of transportation engineering, 136*(7), 632-639.

Arun, A., Haque, M. M., Bhaskar, A., Washington, S., & Sayed, T. (2021). A systematic mapping review of surrogate safety assessment using traffic conflict techniques. *Accident Analysis & Prevention, 153*, 106016.

Ayres, G., Wilson, B., & LeBlanc, J. (2004). Method for identifying vehicle movements for analysis of field operational test data. *Transportation research record, 1886*(1), 92-100.

Bagdadi, O., & Várhelyi, A. (2013). Development of a method for detecting jerks in safety critical events. *Accident Analysis & Prevention, 50*, 83-91.

Besag, J. (1975). Statistical analysis of non-lattice data. *Journal of the Royal Statistical Society: Series D (The Statistician), 24*(3), 179-195.

Blue, V. J., & Adler, J. L. (2001). Cellular automata microsimulation for modeling bi-directional pedestrian walkways. *Transportation Research Part B: Methodological, 35*(3), 293-312.

Brechtel, S., Gindele, T., & Dillmann, R. (2011). *Probabilistic MDP-behavior planning for cars.* Paper presented at the 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC).

Burstedde, C., Klauck, K., Schadschneider, A., & Zittartz, J. (2001). Simulation of pedestrian dynamics using a two-dimensional cellular automaton. *Physica A: Statistical Mechanics and its Applications, 295*(3-4), 507-525.

Caliendo, C., & Guida, M. (2012). Microsimulation approach for predicting crashes at unsignalized intersections using traffic conflicts. *Journal of transportation engineering, 138*(12), 1453-1467.

Canadian Motor Vehicle Traffic Collision Statistics. (2018). Retrieved from https://tc.canada.ca/en/road-transportation/motor-vehicle-safety/canadian-motor-vehicle-traffic-collision-statistics-2018

CCMTA. (2016). *Canada's Road Safety Strategy 2025*. Retrieved from Canadian Councin of Motor Transportation Administrators: https://roadsafetystrategy.ca/web/road-safety-strategy/files/public/docs/RSS-2025-Report-July-2016-with-cover.pdf

Chattaraj, U., Seyfried, A., & Chakroborty, P. (2009). Comparison of pedestrian fundamental diagram across cultures. *Advances in complex systems, 12*(03), 393-405.

Chen, P., Zeng, W., & Yu, G. (2019). Assessing right-turning vehicle-pedestrian conflicts at intersections using an integrated microscopic simulation model. *Accident analysis & prevention, 129*, 211-224.

Chen, P., Zeng, W., Yu, G., & Wang, Y. (2017). Surrogate safety analysis of pedestrian-vehicle conflict at intersections using unmanned aerial vehicle videos. *Journal of advanced transportation, 2017*.

Chen, S., Kuhn, M., Prettner, K., & Bloom, D. E. (2019). The global macroeconomic burden of road injuries: estimates and projections for 166 countries. *The Lancet Planetary Health, 3*(9), e390-e398.

Chin, H.-C., & Quek, S.-T. (1997). Measurement of traffic conflicts. *Safety Science, 26*(3), 169-185.

Cunto, F., & Saccomanno, F. F. (2008). Calibration and validation of simulated vehicle safety performance at signalized intersections. *Accident analysis & prevention, 40*(3), 1171-1179.

Darzentas, J., Cooper, D. F., Storr, P., & McDowell, M. (1980). Simulation of road traffic conflicts at T-junctions. *Simulation, 34*(5), 155-164.

Dias, C., Iryo-Asano, M., Nishiuchi, H., & Todoroki, T. (2018). Calibrating a social force based model for simulating personal mobility vehicles and pedestrian mixed traffic. *Simulation Modelling Practice and Theory, 87*, 395-411.

Dijkstra, A., Marchesini, P., Bijleveld, F., Kars, V., Drolenga, H., & Van Maarseveen, M. (2010). Do calculated conflicts in microsimulation model predict number of crashes? *Transportation research record, 2147*(1), 105-112.

El-Basyouny, K., & Sayed, T. (2013). Safety performance functions using traffic conflicts. *Safety Science, 51*(1), 160-164.

Erol, K., Levy, R., & Wentworth, J. (1998). Application of agent technology to traffic simulation. *Proc. of Complex Systems, Intelligent Systems and Interfaces*.

Essa, M., & Sayed, T. (2015a). Simulated traffic conflicts: do they accurately represent field-measured conflicts? *Transportation Research Record, 2514*(1), 48-57.

Essa, M., & Sayed, T. (2015b). Transferability of calibrated microsimulation model parameters for safety assessment using simulated conflicts. *Accident analysis & prevention, 84*, 41-53.

Essa, M., & Sayed, T. (2016). A comparison between PARAMICS and VISSIM in estimating automated field-measured traffic conflicts at signalized intersections. *Journal of advanced transportation, 50*(5), 897-917.

Fan, R., Yu, H., Liu, P., & Wang, W. (2013). Using VISSIM simulation model and Surrogate Safety Assessment Model for estimating field measured traffic conflicts at freeway merge areas. *IET Intelligent Transport Systems, 7*(1), 68-77.

Farah, H., Piccinini, G. B., Itoh, M., & Dozza, M. (2019). Modelling overtaking strategy and lateral distance in car-to-cyclist overtaking on rural roads: A driving simulator experiment. *Transportation research part F: traffic psychology and behaviour, 63*, 226-239.

Farina, F., Fontanelli, D., Garulli, A., Giannitrapani, A., & Prattichizzo, D. (2017). Walking ahead: The headed social force model. *PloS one, 12*(1), e0169734.

Feliciani, C., Crociani, L., Gorrini, A., Vizzari, G., Bandini, S., & Nishinari, K. (2017). A simulation model for non-signalized pedestrian crosswalks based on evidence from on field observation. *Intelligenza Artificiale, 11*(2), 117-138.

Finn, C., Christiano, P., Abbeel, P., & Levine, S. (2016). A connection between generative adversarial networks, inverse reinforcement learning, and energy-based models. *arXiv preprint arXiv:1611.03852*.

Finn, C., Levine, S., & Abbeel, P. (2016). *Guided cost learning: Deep inverse optimal control via policy optimization.* Paper presented at the International conference on machine learning.

Fu, J., Luo, K., & Levine, S. (2017). Learning robust rewards with adversarial inverse reinforcement learning. *arXiv preprint arXiv:1710.11248*.

Fujii, H., Uchida, H., & Yoshimura, S. (2017). Agent-based simulation framework for mixed traffic of cars, pedestrians and trams. *Transportation Research Part C: Emerging Technologies, 85*, 234-248.

Gettman, D., & Head, L. (2003). Surrogate safety measures from traffic simulation models. *Transportation research record, 1840*(1), 104-115.

Gettman, D., Pu, L., Sayed, T., Shelby, S. G., & Energy, S. (2008). *Surrogate safety assessment model and validation*. Retrieved from

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., . . . Bengio, Y. (2014). Generative adversarial networks. *arXiv preprint arXiv:1406.2661*.

Grabow, M. L., Spak, S. N., Holloway, T., Stone Jr, B., Mednick, A. C., & Patz, J. A. (2012). Air quality and exercise-related health benefits from reduced car travel in the midwestern United States. *Environmental health perspectives, 120*(1), 68-76.

Guo, R.-Y. (2014). New insights into discretization effects in cellular automata models for pedestrian evacuation. *Physica A: Statistical Mechanics and its Applications, 400*, 1-11.

Guo, Y., Essa, M., Sayed, T., Haque, M. M., & Washington, S. (2019). A comparison between simulated and field-measured conflicts for safety assessment of signalized intersections in Australia. *Transportation Research Part C: Emerging Technologies, 101*, 96-110.

Guo, Y., Sayed, T., & Zaki, M. H. (2017). *Automated Safety Diagnosis of Powered Two Wheelers at a Shared Traffic Facility in China*. Retrieved from

Guo, Y., Sayed, T., Zaki, M. H., & Liu, P. (2016). Safety evaluation of unconventional outside left-turn lane using automated traffic conflict techniques. *Canadian Journal of Civil Engineering, 43*(7), 631-642.

Guo, Y., Sayed, T., & Zheng, L. (2020). A hierarchical bayesian peak over threshold approach for conflict-based before-after safety evaluation of leading pedestrian intervals. *Accident analysis & prevention, 147*, 105772.

Helbing, D., Farkas, I., & Vicsek, T. (2000). Simulating dynamical features of escape panic. *Nature, 407*(6803), 487-490.

Helbing, D., & Molnar, P. (1995). Social force model for pedestrian dynamics. *Physical review E, 51*(5), 4282.
Ho, J., & Ermon, S. (2016). Generative adversarial imitation learning. *arXiv preprint arXiv:1606.03476*.
Hornik, K. (1993). Some new results on neural network approximation. *Neural networks, 6*(8), 1069-1072.
Hsu, Y.-C., Gopalswamy, S., Saripalli, S., & Shell, D. A. (2018). *An mdp model of vehicle-pedestrian interaction at an unsignalized intersection.* Paper presented at the 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall).
Huang, G.-B., & Babri, H. A. (1998). Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions. *IEEE transactions on neural networks, 9*(1), 224-229.
Huang, Z., Zhou, S., Zhuang, B., & Zhou, X. (2017). Learning to run with actor-critic ensemble. *arXiv preprint arXiv:1712.08987*.
Hussein, M., & Sayed, T. (2015). A unidirectional agent based pedestrian microscopic model. *Canadian Journal of Civil Engineering, 42*(12), 1114-1124.
Hussein, M., & Sayed, T. (2017). A bi-directional agent-based pedestrian microscopic model. *Transportmetrica A: Transport Science, 13*(4), 326-355.
Huttenlocher, D. P., Klanderman, G. A., & Rucklidge, W. J. (1993). Comparing images using the Hausdorff distance. *IEEE Transactions on pattern analysis and machine intelligence, 15*(9), 850-863.
Hydén, C. (1987). The development of a method for traffic safety evaluation: The Swedish Traffic Conflicts Technique. *Bulletin Lund Institute of Technology, Department*(70).
Ismail, K., Sayed, T., & Saunier, N. (2009). *Automated pedestrian safety analysis using video data in the context of scramble phase intersections.* Paper presented at the Annual Conference of the Transportation Association of Canada, Vancouver, BC.
Ismail, K., Sayed, T., & Saunier, N. (2010). *Camera calibration for urban traffic scenes: practical issues and a robust approach.* Paper presented at the 89th Annual Meeting of the Transportation Research Board, Washington, DC.
Ismail, K., Sayed, T., & Saunier, N. (2013). A methodology for precise camera calibration for data collection applications in urban traffic scenes. *Canadian Journal of Civil Engineering, 40*(1), 57-67.
Ismail, K., Sayed, T., Saunier, N., & Lim, C. (2009). Automated analysis of pedestrian–vehicle conflicts using video data. *Transportation research record, 2140*(1), 44-54.
ITF. (2020). *Road Safety Annual Report*. Retrieved from https://www.itf-oecd.org/sites/default/files/canada-road-safety.pdf
Jennings, N. R. (2000). On agent-based software engineering. *Artificial intelligence, 117*(2), 277-296.
Jennings, N. R., Varga, L. Z., Aarnts, R. P., Fuchs, J., & Skarek, P. (1993). Transforming standalone expert systems into a community of cooperating agents. *Engineering Applications of Artificial Intelligence, 6*(4), 317-331.
Kakade, S. M. (2001). A natural policy gradient. *Advances in neural information processing systems, 14*.

Karamouzas, I., Heil, P., Van Beek, P., & Overmars, M. H. (2009). *A predictive collision avoidance model for pedestrian simulation.* Paper presented at the International workshop on motion in games.

Katrakazas, C., Quddus, M., & Chen, W.-H. (2017). A simulation study of predicting real-time conflict-prone traffic conditions. *IEEE Transactions on Intelligent Transportation Systems, 19*(10), 3196-3207.

Kim, B., Kaelbling, L. P., & Lozano-Pérez, T. (2019). *Adversarial actor-critic method for task and motion planning problems using planning experience.* Paper presented at the Proceedings of the AAAI Conference on Artificial Intelligence.

Kraay, J., Van Der Horst, A., & Oppe, S. (2013). Manual conflict observation technique DOCTOR (Dutch Objective Conflict Technique for Operation and Research).

Layegh, M., Mirbaha, B., & Rassafi, A. A. (2020). Modeling the pedestrian behavior at conflicts with vehicles in multi-lane roundabouts (a cellular automata approach). *Physica A: Statistical Mechanics and its Applications, 556*, 124843.

Levine, S., & Koltun, V. (2012). Continuous inverse optimal control with locally optimal examples. *arXiv preprint arXiv:1206.4617*.

Levine, S., Popovic, Z., & Koltun, V. (2011). *Nonlinear inverse reinforcement learning with gaussian processes.* Paper presented at the Advances in neural information processing systems.

Li, Y., Chen, M., Dou, Z., Zheng, X., Cheng, Y., & Mebarki, A. (2019). A review of cellular automata models for crowd evacuation. *Physica A: Statistical Mechanics and its Applications, 526*, 120752.

Littman, M. L. (1994). Markov games as a framework for multi-agent reinforcement learning. In *Machine learning proceedings 1994* (pp. 157-163): Elsevier.

Liu, D. C., & Nocedal, J. (1989). On the limited memory BFGS method for large scale optimization. *Mathematical programming, 45*(1), 503-528.

Liu, T., Yang, X., Wang, Q., Zhou, M., & Xia, S. (2020). A fuzzy-theory-based cellular automata model for pedestrian evacuation from a multiple-exit room. *IEEE Access, 8*, 106334-106345.

Liu, Y., Alsaleh, R., & Sayed, T. (2021). Modeling Lateral Interactions between Motorized Vehicles and Non-Motorized Vehicles in Mixed Traffic Using Accelerated Failure Duration Model. *Transportmetrica A: transport science*(just-accepted), 1-26.

Lu, L., Ren, G., Wang, W., Chan, C.-Y., & Wang, J. (2016). A cellular automaton simulation model for pedestrian and vehicle interaction behaviors at unsignalized mid-block crosswalks. *Accident Analysis & Prevention, 95*, 425-437.

Lucas, B. D., & Kanade, T. (1981). *An iterative image registration technique with an application to stereo vision*.

Macal, C. M., & North, M. J. (2005). *Tutorial on agent-based modeling and simulation.* Paper presented at the Proceedings of the Winter Simulation Conference, 2005.

Mannering, F. L., & Bhat, C. R. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic methods in accident research, 1*, 1-22.

Martens, J., & Grosse, R. (2015). *Optimizing neural networks with kronecker-factored approximate curvature.* Paper presented at the International conference on machine learning.

Martinez-Gil, F., Lozano, M., García-Fernández, I., Romero, P., Serra, D., & Sebastián, R. (2020). Using Inverse Reinforcement Learning with Real Trajectories to Get More Trustworthy Pedestrian Simulations. *Mathematics, 8*(9), 1479.

Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T., Harley, T., . . . Kavukcuoglu, K. (2016). *Asynchronous methods for deep reinforcement learning.* Paper presented at the International conference on machine learning.

Mohammed, H., Sayed, T., & Bigazzi, A. (2020). Microscopic modeling of cyclists on off-street paths: A stochastic imitation learning approach. *Transportmetrica A: transport science*, 1-39.

Morar, T., & Bertolini, L. (2013). Planning for pedestrians: a way out of traffic congestion. *Procedia-Social and Behavioral Sciences, 81*, 600-608.

Moudon, A. V., Lin, L., Jiao, J., Hurvitz, P., & Reeves, P. (2011). The risk of pedestrian injury and fatality in collisions with motor vehicles, a social ecological study of state routes and city streets in King County, Washington. *Accident Analysis & Prevention, 43*(1), 11-24.

Moussaïd, M., Perozo, N., Garnier, S., Helbing, D., & Theraulaz, G. (2010). The walking behaviour of pedestrian social groups and its impact on crowd dynamics. *PloS one, 5*(4), e10047.

Nagel, K., & Schreckenberg, M. (1992). A cellular automaton model for freeway traffic. *Journal de physique I, 2*(12), 2221-2229.

Ng, A. Y., Harada, D., & Russell, S. (1999). *Policy invariance under reward transformations: Theory and application to reward shaping.* Paper presented at the Icml.

Ng, A. Y., & Russell, S. J. (2000). *Algorithms for inverse reinforcement learning.* Paper presented at the Icml.

Parisi, D. R., & Dorso, C. O. (2005). Microscopic dynamics of pedestrian evacuation. *Physica A: Statistical Mechanics and its Applications, 354*, 606-618.

Parker Jr, M., & Zegeer, C. V. (1989). *Traffic conflict techniques for safety and operations: Observers manual*. Retrieved from

Perkins, S. R., & Harris, J. L. (1968). Traffic conflict characteristics-accident potential at intersections. *Highway Research Record*(225).

Pirdavani, A., Brijs, T., Bellemans, T., & Wets, G. (2010). *A Simulation-Based Traffic Safety Evaluation of Signalized and Un-Signalized Intersections.* Paper presented at the Proceedings of the 15th International Conference Road Safety on Four Continents.

Rad, S. R., de Almeida Correia, G. H., & Hagenzieker, M. (2020). Pedestrians' road crossing behaviour in front of automated vehicles: Results from a pedestrian simulation experiment using agent-based modelling. *Transportation research part F: traffic psychology and behaviour, 69*, 101-119.

Rissel, C. E. (2009). Active travel: a climate change mitigation strategy with co-benefits for health. *New South Wales public health bulletin, 20*(2), 10-13.

Sacchi, E., & Sayed, T. (2016). Conflict-based safety performance functions for predicting traffic collisions by type. *Transportation research record, 2583*(1), 50-55.

Saunier, N., & Sayed, T. (2006). *A feature-based tracking algorithm for vehicles in intersections.* Paper presented at the The 3rd Canadian Conference on Computer and Robot Vision (CRV'06).

Saunier, N., & Sayed, T. (2007). Automated analysis of road safety with video data. *Transportation research record, 2019*(1), 57-64.

Sayed, T., Ismail, K., Zaki, M. H., & Autey, J. (2012). Feasibility of computer vision-based safety evaluations: Case study of a signalized right-turn safety treatment. *Transportation Research Record, 2280*(1), 18-27.
Sayed, T., & Zein, S. (1999). Traffic conflict standards for intersections. *Transportation Planning and Technology, 22*(4), 309-323.
SDG. (2015). *Transforming our world: The 2030 agenda for sustainable development*. Retrieved from https://www.un.org/pga/wp-content/uploads/sites/3/2015/08/120815_outcome-document-of-Summit-for-adoption-of-the-post-2015-development-agenda.pdf
SDG. (2020). *The Sustainable Developement Goals Report*. Retrieved from https://unstats.un.org/sdgs/report/2020/progress-summary-for-SDG-targets/
Sheykhfard, A., & Haghighi, F. (2020). Assessment pedestrian crossing safety using vehicle-pedestrian interaction data through two different approaches: Fixed videography (FV) vs In-Motion Videography (IMV). *Accident analysis & prevention, 144*, 105661.
Sheykhfard, A., Haghighi, F., Papadimitriou, E., & Van Gelder, P. (2021). Analysis of the occurrence and severity of vehicle-pedestrian conflicts in marked and unmarked crosswalks through naturalistic driving study. *Transportation research part F: traffic psychology and behaviour, 76*, 178-192.
Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., & Riedmiller, M. (2014). *Deterministic policy gradient algorithms*.
Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information processing & management, 45*(4), 427-437.
Song, J., Ren, H., Sadigh, D., & Ermon, S. (2018). Multi-agent generative adversarial imitation learning. *arXiv preprint arXiv:1807.09936*.
Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*: MIT press.
Svensson, Å. (1998). *A method for analysing the traffic process in a safety perspective.* Lund University,
Tageldin, A., & Sayed, T. (2019). Models to evaluate the severity of pedestrian-vehicle conflicts in five cities. *Transportmetrica A: transport science, 15*(2), 354-375.
Tageldin, A., Sayed, T., & Shaaban, K. (2017). Comparison of time-proximity and evasive action conflict measures: case studies from five cities. *Transportation research record, 2661*(1), 19-29.
Tageldin, A., Zaki, M. H., & Sayed, T. (2017). Examining pedestrian evasive actions as a potential indicator for traffic conflicts. *IET Intelligent Transport Systems, 11*(5), 282-289.
Tan, Z., Che, Y., Xiao, L., Hu, W., Li, P., & Xu, J. (2021). Research of fatal car-to-pedestrian precrash scenarios for the testing of the active safety system in China. *Accident Analysis & Prevention, 150*, 105857.
Thomas, A. J., Petridis, M., Walters, S. D., Gheytassi, S. M., & Morgan, R. E. (2017). *Two hidden layers are usually better than one.* Paper presented at the International conference on engineering applications of neural networks.
Thomas, A. J., Walters, S. D., Gheytassi, S. M., Morgan, R. E., & Petridis, M. (2016). On the optimal node ratio between hidden layers: a probabilistic study. *International Journal of Machine Learning and Computing, 6*(5), 241.
Tiwari, G., Bangdiwala, S., Saraswat, A., & Gaurav, S. (2007). Survival analysis: Pedestrian risk exposure at signalized intersections. *Transportation research part F: traffic psychology and behaviour, 10*(2), 77-89.

Tomasi, C., & Detection, T. K. (1991). Tracking of point features. *Int. J. Comput. Vis*, 137-154.
Trinh, L. T., Sano, K., & Hatoyama, K. (2021). Modelling and simulating head-on conflict-solving behaviour of motorcycles under heterogeneous traffic condition in developing countries. *Transportmetrica A: transport science, 17*(4), 921-945.
Wang, J., Boltes, M., Seyfried, A., Zhang, J., Ziemer, V., & Weng, W. (2018). Linking pedestrian flow characteristics with stepping locomotion. *Physica A: Statistical Mechanics and its Applications, 500*, 106-120.
Wang, L., Ning, P., Yin, P., Cheng, P., Schwebel, D. C., Liu, J., . . . Zeng, X. (2019). Road traffic mortality in China: analysis of national surveillance data from 2006 to 2016. *The Lancet Public Health, 4*(5), e245-e255.
WHO. (2018). *Global status report on road safety 2018: summary*. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/277370/WHO-NMH-NVI-18.20-eng.pdf
Wu, Y., Mansimov, E., Liao, S., Grosse, R., & Ba, J. (2017). Scalable trust-region method for deep reinforcement learning using kronecker-factored approximation. *arXiv preprint arXiv:1708.05144*.
Xie, D.-F., Gao, Z.-Y., Zhao, X.-M., & Wang, D. Z. (2012). Cellular automaton modeling of the interaction between vehicles and pedestrians at signalized crosswalk. *Journal of Transportation Engineering, 138*(12), 1442-1452.
Xu, Y., Zheng, Y., & Yang, Y. (2021). On the movement simulations of electric vehicles: a behavioral model-based approach. *Applied Energy, 283*, 116356.
Ye, J., Chen, X., Yang, C., & Wu, J. (2008). Walking behavior and pedestrian flow characteristics for different types of walking facilities. *Transportation research record, 2048*(1), 43-51.
Yoshii, T., & Kuwahara, M. (1995). *SOUND: A traffic simulation model for oversaturated traffic flow on urban expressways.* Paper presented at the Preprint at 7th World Conference on Transportation Research, Sydney.
Yu, L., Song, J., & Ermon, S. (2019). *Multi-agent adversarial inverse reinforcement learning.* Paper presented at the International Conference on Machine Learning.
Zaki, M. H., & Sayed, T. (2013). A framework for automated road-users classification using movement trajectories. *Transportation research part C: emerging technologies, 33*, 50-73.
Zaki, M. H., & Sayed, T. (2014). Using automated walking gait analysis for the identification of pedestrian attributes. *Transportation Research Part C: Emerging Technologies, 48*, 16-36.
Zaki, M. H., Sayed, T., & Shaaban, K. (2014). Use of drivers' jerk profiles in computer vision–based traffic safety evaluations. *Transportation Research Record, 2434*(1), 103-112.
Zaki, M. H., Sayed, T., Tageldin, A., & Hussein, M. (2013). Application of computer vision to diagnosis of pedestrian safety issues. *Transportation research record, 2393*(1), 75-84.
Zeng, W., Chen, P., Nakamura, H., & Iryo-Asano, M. (2014). Application of social force model to pedestrian behavior analysis at signalized crosswalk. *Transportation Research Part C: Emerging Technologies, 40*, 143-159.
Zeng, W., Chen, P., Yu, G., & Wang, Y. (2017). Specification and calibration of a microscopic model for pedestrian dynamic simulation at signalized intersections: A hybrid approach. *Transportation Research Part C: Emerging Technologies, 80*, 37-70.

Zhao, J., Malenje, J. O., Wu, J., & Ma, R. (2020). Modeling the interaction between vehicle yielding and pedestrian crossing behavior at unsignalized midblock crosswalks. *Transportation research part F: traffic psychology and behaviour, 73*, 222-235.

Zheng, L., Sayed, T., Essa, M., & Guo, Y. (2019). *Do simulated traffic conflicts predict crashes? An investigation using the extreme value approach.* Paper presented at the 2019 IEEE Intelligent Transportation Systems Conference (ITSC).

Zhong, C., Lu, Z., Gursoy, M. C., & Velipasalar, S. (2018). *Actor-critic deep reinforcement learning for dynamic multichannel access.* Paper presented at the 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP).

Zhou, Z., Zhou, Y., Pu, Z., & Xu, Y. (2019). Simulation of pedestrian behavior during the flashing green signal using a modified social force model. *Transportmetrica A: transport science, 15*(2), 1019-1040.

Ziebart, B. D. (2010). Modeling purposeful adaptive behavior with the principle of maximum causal entropy.

Ziebart, B. D., Maas, A. L., Bagnell, J. A., & Dey, A. K. (2008). *Maximum entropy inverse reinforcement learning.* Paper presented at the Aaai.