Prediction of Intelligent Vehicle-Pedestrian Conflict in a Highly Uncertain Environment

by

Hesham M. Alghodhaifi

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical and Computer Engineering)
in the University of Michigan-Dearborn
2023

Doctoral Committee:

Associate Professor Sridhar Lakshmanan, Chair
Associate Professor Shan Bao
Professor Paul Richardson
Associate Professor Paul Watta

Hesham M. Alghodhaifi

halghodh@umich.edu

ORCID iD: 0000-0003-3312-7457

Dedicated to my family, friends, and colleagues.

Acknowledgments

Five incredible years have soared by since embarking on the mesmerizing odyssey of my PhD. This transformative voyage has been a symphony of ardor, wisdom, and boundless cognitive expansion. Now, as I bask in the triumph of reaching this fateful crossroads, my heart brims with immeasurable gratitude toward the indispensable souls who have embellished this enchanting expedition with their unwavering presence.

First, I extend my heartfelt gratitude to my esteemed advisor, Professor Sridhar Lakshmanan, whose unwavering guidance has accompanied me on this profound journey. Within the vast tapestry of knowledge imparted by Professor Sridhar, four invaluable lessons have illuminated my path: cultivating a researcher's mindset, fostering collaboration within a research team, maintaining meticulous organization and unwavering physical and mental vitality, and harmonizing the pursuit of professional accomplishment with the blessings of familial joy. The impact and inspiration Professor Sridhar has bestowed upon me transcend the confines of this dissertation, leaving an indelible imprint that will reverberate throughout my entire life.

Second, I would like to thank my committee members: Professor Paul Richardson, Professor Shan Bao, and Professor Paul Watta. I could not have gotten to where I am now without their assistance. They encourage me to think thoroughly about the situation and are always willing to assist. Their extensive expertise opens my eyes to the frontier of knowledge and connects my study to real-world implications. These contributions keep me motivated and proud of my work.

I would also like to thank everyone who helped make my dissertation possible, especially Professor Paul Richardson, who encouraged me throughout my degree; there are no words to express how grateful I am to Professor Paul. Furthermore, I want to express sincere thanks to Professor Selim Awad for his steadfast support and unrelenting encouragement during my whole time at the University of Michigan. His remarks have been a source of motivation for me, motivating me to exceed expectations and persist tenaciously throughout my PhD journey. I would also like to thank all my colleagues in the Electrical and Computer Engineering department at the University of Michigan-Dearborn.

Finally, I would like to thank my family for their unconditional support and love. My brother Naif has both encouraged me with kind words and inspired me with his incredible achievements. My lovely sisters and brothers, have inspired me and pushed me to achieve my goal. My wife's love and support know no boundaries, serving as a foundation that sustained me during the perilous path of my PhD studies. Her passion and support have been a source of strength for me, fuelling my commitment to overcome challenges and move forward. My beloved son, who brings boundless joy to my life, your tender presence has been a constant source of inspiration and motivation throughout my PhD journey. My mom and dad are everything to me. I’m really trying my best to make them proud and happy. I love you both.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ADAS Advanced Driver Assistance Systems

AVs Autonomous Vehicles

CNN Convolutional Neural Network

GAN Generative Adversarial Network

GAT Graph Attention

GCN Graph Convolutional Networks

GNN Graph Neural Network

GRU Gated Recurrent Units

HMM Hidden Markov Models

HSTGA Holistic Spatio-Temporal Graph Attention Trajectory Prediction for Vehicle-Pedestrian Interaction

LSTM Long Short-Term Memory Networks

MATF Multi-Agent Fusion Model

MHA Multi-Head Attention

MLP Multi-Layer Perceptron

NHTSA U.S. National Highway Traffic Safety Administration

PET Post-Encroachment Time

RNN Recurrent Neural Network

SAs Surrounding Agents

SSMs Surrogate safety measures

TA Target Agent

TTC Time-to-Collision

V2I Vehicle-to-Infrastructure

V2P Vehicle-to-Pedestrian

V2X Vehicle-to-Everything

VGMM Variational Gaussian Mixture Model

VRUs Vulnerable Road Users

WHO World Health Organization

Abstract

Ensuring that intelligent vehicles do not cause fatal collisions remains a persistent challenge due to pedestrians' unpredictable movements and behavior. The potential for risky situations or collisions arising from even minor misunderstandings in vehicle-pedestrian interactions is a cause for great concern. Considerable research has been dedicated to the advancement of predictive models for pedestrian behavior through trajectory prediction, as well as the exploration of the intricate dynamics of vehicle-pedestrian interactions. Such endeavors aim to enhance the ability of intelligent vehicles to respond appropriately to the presence of pedestrians in their surroundings. Predicting pedestrian movements in crowded spaces is complex due to pedestrian interactions' continuous and forward-looking nature. While current methods fail to account for temporal correlations among these interactions, recurrent neural networks have been used to model social interactions. However, they are limited in capturing spatiotemporal interactions. Graph Neural Networks (GNNs) have been introduced to address this limitation but do not fully capture real-world social interactions as they consider the impact between traffic participants as fixed or symmetric.

We propose a novel graph-based trajectory prediction model for vehicle-pedestrian interactions called Holistic Spatio-Temporal Graph Attention Trajectory Prediction (HSTGA) to address these limitations. HSTGA first extracts vehicle-pedestrian interaction spatial features using a multi-layer perceptron (MLP) sub-network and max pooling. Then, the vehicle-pedestrian interaction features are aggregated with the spatial features of pedestrians and vehicles to be fed into the LSTM. The LSTM is modified to learn the vehicle-pedestrian interaction adaptively.

Moreover, HSTGA models temporal interactions using an additional LSTM. Then, it models the spatial interactions among pedestrians and between pedestrians and vehicles using Graph Attention Networks (GATs) to combine the hidden states of the LSTMs. Then, the predicted trajectories of vehicles and pedestrians are used to design a vehicle-pedestrian conflict model. The conflict model is used to investigate safety-critical driving metrics, such as severity and near-miss metrics, as well as normal driving metrics, including vehicle speed, vehicle acceleration, vehicle heading angle, pedestrian speed, pedestrian acceleration, pedestrian heading angle, and distance distributions. Two safety indicators, namely Time-to-Collision (TTC) and Post-Encroachment Time (PET), are used to quantify the severity and near-miss conflicts between the vehicle and the pedestrian.

We evaluate the performance of HSTGA on three different scenario datasets, including complex unsignalized roundabouts with no crosswalks and unsignalized intersections. Results show that HSTGA outperforms several state-of-the-art methods in predicting linear, curvilinear, and piece-wise linear paths of vehicles and pedestrians and can also predict collisions between pedestrians and vehicles several seconds before they occur. Our approach provides a more comprehensive understanding of social interactions, enabling more accurate trajectory prediction for safe vehicle navigation.

Chapter 1

Introduction

1.1 Motivation

Driving in an urban environment (Fig.1.1) is a challenging task that is associated with heavy mixed traffic flows. In mixed traffic flow, vehicles and vulnerable road users such as pedestrians, bicycles, and tricycles share the same road. As a result, vehicle-pedestrian conflicts, vehicle-vehicle conflicts, and many other critical interactions may regularly occur. According to the U.S. National Highway Traffic Safety Administration (NHTSA) data, in 2020, 6,516 pedestrians died in traffic accidents, and almost 55,000 pedestrians were injured nationwide [1].

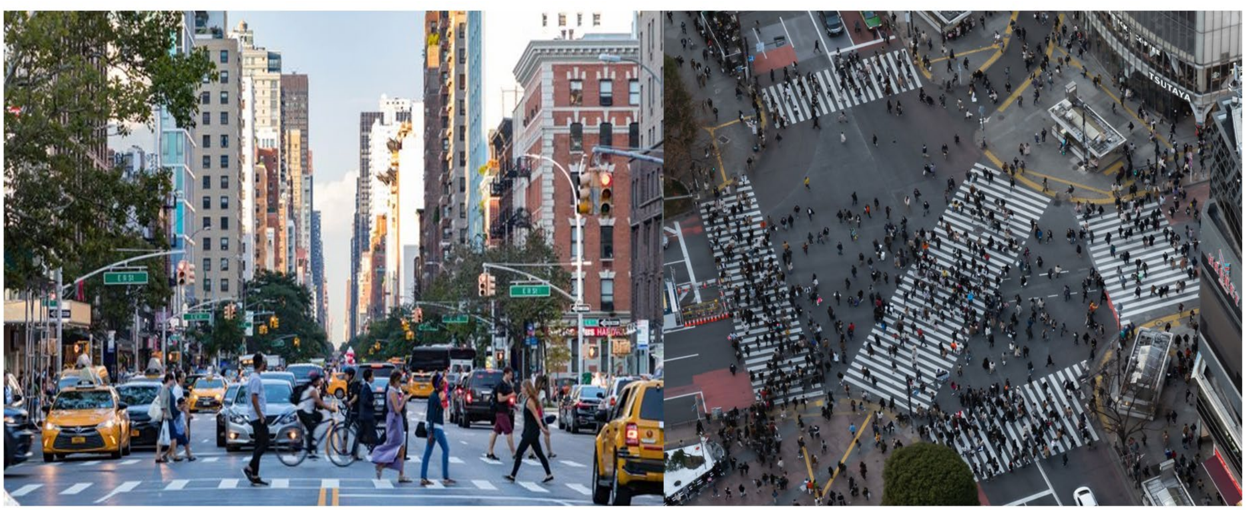


Figure 1.1 Urban Environment Scenarios [85], [86].

The conflict between pedestrians and vehicles (Fig. 1.2) is an important safety issue not just in the U.S. but everywhere in the world. This issue is even worse in developing countries. For example, in a report by the World Health Organization (WHO) in 2007, around 40% of road crash fatalities were pedestrians in countries such as Albania, Armenia, and Chile [2]. Moreover, in 2007, over 50% of road crash fatalities were pedestrians in countries such as Bangladesh, Ethiopia, and Mozambique [2]. According to the World Health Organization (WHO), almost every one out of five road deadly accidents is with a pedestrian [3]. This means that around 300,000 of the nearly 1.35 million individuals killed in traffic accidents each year are pedestrians [3]. Pedestrians are among the most vulnerable road users (VRUs) because they lack the physical protection to reduce accident consequences [4]. It is not surprising that pedestrians' conflicts with vehicles are most problematic in urban areas since pedestrians' activity is higher there. The problem of collision between vehicles and pedestrians has been the subject of deep study for a long time [5]-[12].

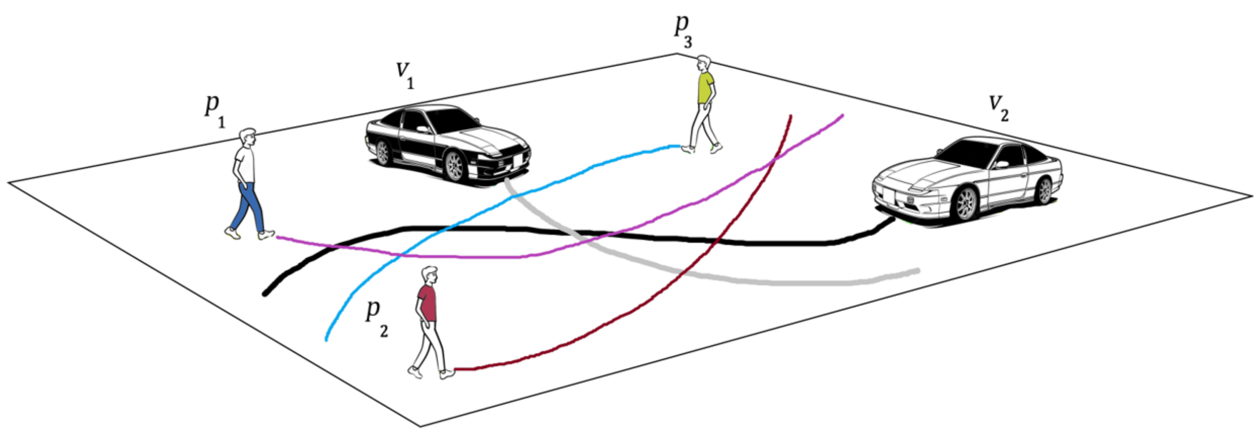


Figure 1.2 An Example of Vehicle-Pedestrian Conflicts.

The meaning of traffic conflict can vary among research publications. In [13], the authors noted that operational definitions of traffic conflict could generally be categorized into two types: those based on evasive actions and those based on spatial-temporal proximity. A situation involving two or more road users in which one user's activity induces another user to perform an evasive move to avoid a collision is characterized as an evasive action-based traffic conflict [14]. Pedestrian-vehicle conflicts can occur when an incoming vehicle must quickly stop or swerve to avoid a pedestrian or when a pedestrian must take evasive action to prevent a collision. This term focuses on either the driver's or the pedestrian's evasive actions. In contrast, proximity-based traffic conflict is characterized as a scenario in which two or more road users are so close in space and time that there is a chance of an accident if their movements do not alter [15]. This indicates that the closer the road users are to one other, either in time or space, the more likely an accident may occur. This is a conceptual (theoretical) definition that is operational since the time and space dimensions are quantifiable and can be monitored by traffic detectors [16].

Several published studies have dealt with pedestrian-vehicle conflict, but they were limited to studying the factors influencing conflicts, such as personal characteristics, traffic conditions, and environmental factors at crosswalks [16]. From the first perspective, personal characteristics like age, gender, and disability have been studied. For example, Liu and Tung [17] found that elderly pedestrians exposed themselves to a higher risk of road crossing than young pedestrians due to their decline in walking ability. Yagil [18] found that men are less aware than women of their conflicts with vehicles when they cross the street. Tom and Granié [19] explored gender differences in pedestrian rule compliance both at signalized and unsignalized crossroads. In terms of traffic conditions, studies have looked at things like traffic volume and vehicle speed. Cheng [20] proposed that high vehicle volume can lead to more severe pedestrian-vehicle conflicts because pedestrians' protracted waiting time exceeds their tolerance limit, whilst high vehicle speed increases the chance of pedestrian-vehicle crashes. Cheng created models to investigate the links between pedestrian waiting time, vehicle volume, time of pedestrian-vehicle conflict, vehicle speed, traffic delay, and pedestrian volume. Himanen and Kulmala [21] examined 799 pedestrian-vehicle conflict incidents and determined that the most relevant explanatory factors were the pedestrian distance from the curb, city size, number of people crossing concurrently, vehicle speed, and vehicle platoon size. Furthermore, environmental factors such as city size, signal settings, road width, and lane delineation have been thoroughly researched.

Traffic signals and crosswalks are the most important environmental factors in vehicle-pedestrian conflict because pedestrians and drivers are expected to obey restrictions at these landmarks. At a non-signalized marked crosswalk, pedestrians have the right-of-way according to traffic regulations in the U.S., Europe, Japan, China, and many other places, and vehicles are expected to give the right-of-way to pedestrians. However, due to lack of training, familiarity, etc. vehicles do not always give right-of-way to pedestrians [22], which makes the crossing of a non-signalized intersection more complex compared to those at signals. Many studies have tried to study whether particular traffic controls—such as a crosswalk, stop sign, or yield sign—have positive or negative effects on pedestrian safety. However, there is disagreement about the overall effectiveness of these traffic controls. On the one hand, intersections with traffic signals or stop signs are considered safer for pedestrians than intersections without these traffic controls [22]. Further, traffic signals help control traffic flow and prevent conflicts between vehicles and pedestrians and reduce the likelihood of accidents due to the less amount of time that people spend waiting at an intersection to cross. On the other hand, however, many studies have shown that traffic signals actually increase the risk of crashes because drivers tend to speed up to avoid traffic lights when there are no vehicles ahead. Some research suggests that traffic signals also increase traffic delays, which further exacerbates the speed-up issue and pedestrian risk. Furthermore, pedestrians cross a typical bidirectional crosswalk to the opposite side at signalized intersections in two stages (one direction, then the other), but at non-signalized crosswalks, the same crossing happens in four stages (lane by lane) [23]. Neither crossing behavior reflects the complex interaction between pedestrians and vehicles at uncontrolled intersections that includes eye contact, head nods, hand waves, and slowing down [24].

This dissertation investigated pedestrian-vehicle conflict in non-signalized and non-crosswalk scenarios, which makes it very distinct from previous related research. There is also another motivation, namely, autonomous vehicles.

In order to ensure safe and seamless travel, people make significant and intuitive decisions when driving. These decisions result from actions and interactions with individuals and vehicles in a scene. For example, people driving vehicles can monitor the area and see if a pedestrian intends to cross the road. That information is fundamental for drivers but also a valuable tool in determining the next steps that need to be taken, such as slowing down, speeding up, or stopping. The drivers navigate through traffic not only by monitoring the movements of pedestrians and vehicles but also by anticipating the interdependence between pedestrians and other objects, such as vehicles or traffic lights. In contrast, machines do not possess the ability to perceive a person's judgment through simple gestures and interactions. As a result, autonomous vehicles are very conservative compared to human drivers.

The automotive industry is currently focusing on pushing vehicle technology towards fully automated driving, with technology companies joining these efforts. Vehicle autonomy aims to make road traffic safer and more efficient by partially or wholly substituting the driver in the driving task [25]. The transition phase from manually-operated vehicles to fully automated vehicles is characterized by the degree of control over the driving task. Recent developments in the area of advanced driver assistance systems (ADAS) have shown vast improvements in the accessibility of autonomous driving. Many companies have raised their levels of autonomy over the last few years. Several projects are targeting SAE level 4 or higher. A list of the definition of

SAE levels of AVs is explained in [26] and is shown below.

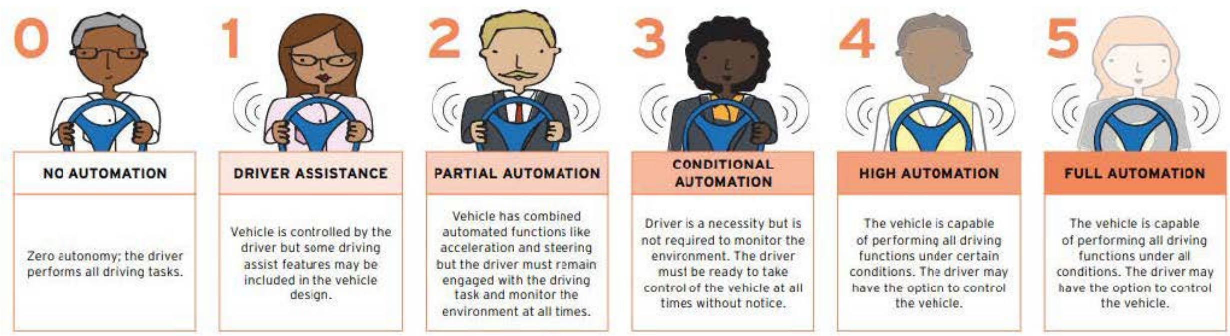
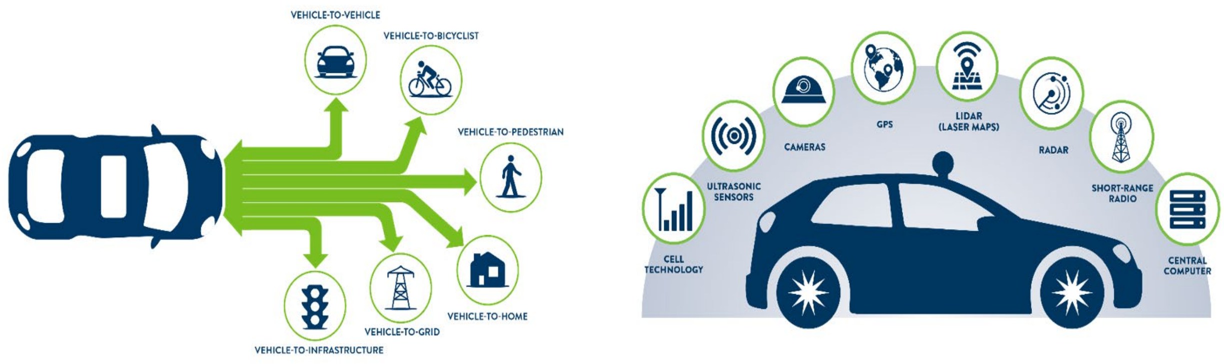


Figure 1.3 Summary of levels of driving automation for on-road vehicles [26], [27].

Since these systems cannot completely manage the driving task, the driver will partially have to support the system or take over control. Therefore, a lot of research is done to implement adequate take-over request strategies involving the driver as an additional sensor, for example, to extend the capabilities of the semiautonomous vehicles and to keep the driver in the loop [28], [29]. However, already existing systems (e.g., emergency stop assistants) and future systems will perform better than humans in more and more sub-areas of the driving task until they finally take over completely. This complex future artificial system, which perceives the environment and makes its own decisions, will be confronted by humans who are used to making decisions and actions in road traffic based on the perceived environment. If these two agents meet, situations will inevitably arise in which the automation's planned action contradicts the driver's desired action. With the advancements toward fully autonomous vehicles and the gradual removal of humans from the driving task, the social safety contract between vehicles and pedestrians is being reexamined [27]. Behavioral psychology studies have investigated the social aspects of driving and have shown the factors that can significantly impact road safety [24], [30]. These factors include pedestrian demographics, road conditions, social factors, and traffic characteristics [27]. Thus, a deep understanding of pedestrian crossing behavior, the factors influencing this behavior, and how the factors might be inter-connected is required. Initial efforts have already been made to develop intent prediction algorithms that estimate the future movements of both pedestrians and drivers [27].

Fully or partially autonomous vehicles (AVs) must develop situational and behavioral awareness to operate effectively in complicated and crowded environments with a wide range of interacting traffic participants. In particular, such vehicles must be aware of their locations and the locations and positions of other vehicles and pedestrians in the vicinity. Previous research has suggested that vehicle-to-everything (V2X) communication, also known as V2V, V2P, and V2I communications, could be used to develop and improve situational awareness in such vehicles. However, it has also been observed that these types of communication are not always reliable due to several factors, including occlusion, interference, and network congestion [9]. Moreover, it is also difficult for vehicles to communicate with pedestrians and other vulnerable road users (VRUs) due to the unsuitability of V2V and V2I communication systems for V2P applications. Therefore, an alternative approach to developing and improving situational awareness in AVs around pedestrians is desired. Of particular interest are technologies that will assist AVs in perceiving and predicting the behavior of the VRUs that are in their immediate vicinity [9]. V2I communications have the potential to transform how cities interact with autonomous vehicles (AVs) and make cities a safe environment for driving. This technology holds a lot of promise in autonomous driving because it allows AVs to connect with different infrastructure elements, such as traffic lights, road signs, and surveillance cameras, in real-time. This technology can more easily monitor the environment to obtain information about the movement of various traffic participants in their field of regard and communicate this information reliably to AVs. Moreover, this technology can assist AVs in predicting the trajectories of surrounding road users, which will be crucial in ensuring safe and efficient interactions between AVs and other road users. The following figure shows the fully and partially autonomous vehicle technology.



(a) (b)

Figure 1.4 Intelligent Vehicles: a) Connected Vehicle; b) Autonomous Vehicle [27].

In the realm of autonomous vehicles (AVs), the ability to anticipate the movement of pedestrians is of paramount significance, and the consequences of neglecting it could be catastrophic. This prediction enables AVs to chart safe routes while engaging in related driving tasks confidently. Unfortunately, the intricate nature of pedestrian motion creates significant challenges for long-term trajectory prediction. It is worth noting that pedestrians' movements are slower than those of vehicles but can change rapidly due to the complexities of human behavior. Furthermore, a pedestrian's gait can be subjective, depending on various factors such as personal characteristics, walking objectives, and the ever-changing environment. In this dissertation, we focus on predicting the trajectory of pedestrians when interacting with other pedestrians and vehicles. Trajectory prediction is crucial for autonomous vehicles because it allows them to estimate the movements of the surrounding road users several seconds into the future and make the right decision to avoid any critical conflicts. Achieving precise trajectory predictions requires the development of efficient algorithms that accurately model and replicate real-world scenarios.

Consequently, the design of such algorithms represents the most critical aspect of the task of accurate trajectory prediction.

Predicting the trajectories of the vehicles is comparably simpler due to their well-understood mechanics. Deep learning techniques, including LSTM models, have effectively achieved this objective [33]-[35]. However, predicting the movement of pedestrians is more difficult since their mechanics are not as readily modeled. Additionally, as mentioned earlier, pedestrians modify their "standard" actions when interacting with vehicles, mainly when they sense potential danger. A precise prediction of pedestrian and vehicle trajectories can assist autonomous vehicles or human drivers in making better decisions and avoiding obstacles and traffic buildup [36]. Furthermore, using intelligent technology to track vehicle and pedestrian movement can boost traffic safety and decrease traffic congestion [37], [38].

To achieve precise pedestrian trajectory prediction, it is imperative to obtain accurate measurements. This task, however, is quite difficult due to a number of factors that can introduce inaccuracies in the collected data. These factors include occlusions caused by large vehicles and illumination issues like shadows and glare [39], [40]. Additionally, pedestrians are physically smaller and lighter than most objects in their surroundings, and they can suddenly change their speed and direction, which further complicates trajectory prediction. This dissertation focuses on this challenging problem and aims to develop an efficient method for predicting pedestrian behavior via trajectory prediction. Accurate trajectory prediction assists autonomous vehicles in collision avoidance and can also be employed in smart intersections. The proposed method can also be extended to encompass the trajectory prediction of other vulnerable road users, such as bicycles, scooters, and others.

Researchers in the field of predictive analytics have looked deeply into the usage of deep learning for predicting pedestrian movements. Sequential models, which employ long short-term memory (LSTM) networks and LSTM-based generative adversarial networks (GANs) [41], [42], and non-sequential models, which use convolutional neural networks (CNNs) [43], [44], are the two most often used methodologies. LSTM networks are a form of recurrent neural network (RNN) that can handle sequential data such as time series or text. They are very effective for modeling long-term data relationships [41]. LSTM-based GANs is a generative model that generates realistic pedestrian trajectories through adversarial training [42]. CNNs, on the other hand, are a sort of feedforward neural network that is frequently employed for image identification [44]. They may also be used to handle non-image data, such as time series data as if it were a 2D picture. While these models show promise, they are not without flaws.

In recent years, there has been an increasing interest in developing LSTM-based methods for capturing dynamic interactions of pedestrians. These methods utilize pooling and attention mechanisms to represent latent motion dynamics of pedestrians in local neighborhoods or the whole scene. While pooling collects motion dynamics of nearby pedestrians, attention assigns different importance to each pedestrian to better understand crowd behaviors based on spatial interactions. However, the temporal continuity of interactions in the crowd has been neglected in previous works. Pedestrians need to consider others' historical movements to determine their current motion behavior to avoid potential collisions in the future, making temporal correlations of interactions important.

Many other studies on predicting pedestrian trajectories have been conducted. However, most of these studies fail to take into account one of the most important factors influencing pedestrian behavior: the presence of multiple surrounding vehicles and the interaction between these vehicles and pedestrians. Although some recent studies, such as the one by Eiffert et al. [46], have attempted to incorporate such influences, they only considered a single vehicle in the presence of pedestrians. Furthermore, previous research that predicts the trajectories of heterogeneous traffic agents, such as pedestrians, has tended to focus on vehicles or motorcycles [47]-[50]. Additionally, it is challenging to evaluate the accuracy of pedestrian trajectory predictions due to the absence of datasets containing annotations for both pedestrian crowds and vehicles. The widely used ETH [51] and UCY [52] datasets, for example, do not include annotations for automobiles and are hence unsuitable for evaluating this job. As a result, there is a need for more research that considers the impact of various surrounding vehicles and pedestrians and captures the spatiotemporal interactions between them on pedestrian behavior, and develops more accurate algorithms for this task. Moreover, diverse datasets that have many vehicles and pedestrians should be used to investigate pedestrian trajectory prediction accurately.

To address all the limitations mentioned above, in this dissertation, we build a novel Spatial-Temporal Graph Attention network called Holistic Spatio-Temporal Graph Attention Trajectory Prediction for Vehicle-Pedestrian Interaction (HSTGA), where the spatial and temporal interactions among pedestrians and among pedestrians and vehicles are encoded, respectively. Moreover, we use multiple datasets such as VCI-DUT [53], rounD [54], and uniD [55] datasets that have pedestrians and vehicles at the same time. This allows modeling the influence of pedestrian-vehicle conflict in the accurate prediction of pedestrian (and vehicle) trajectories.

In order to measure pedestrian-vehicle conflicts, proximity measures such as time to collision (TTC) [56] and post-encroachment time (PET) [57] are commonly used. TTC is the expected time for two road users to collide if they remain on the same trajectory, while PET is the time between the moment the first road user leaves the conflict zone and the moment the second road user reaches it. These measures are used to classify conflicts into different severity levels, with higher TTC and PET values indicating lower severity. Unlike TTC, PET does not consider a collision course criterion [16]. Pedestrian-vehicle conflicts are classified into discrete severity levels based on different thresholds of TTC and PET. For instance, a study analyzed pedestrian-vehicle conflicts and traffic violations using TTC and PET indicators, where conflicts were classified into four levels from high to low risk [58]. Other studies used TTC, PET, and vehicle deceleration to divide the severity of conflicts into three different levels [59]. These levels are serious, slight, and potential conflicts.

In this dissertation too, we adopt PET and TTC indicators to incorporate pedestrian-vehicle conflict in non-signalized non-crosswalk scenarios. The severity of conflict is divided into three categories: serious, slight, and potential according to the PET and TTC indicator levels.

This dissertation makes four contributions:

1. We develop a novel encoder-decoder interaction model called Holistic Spatio-Temporal Graph Attention Trajectory Prediction for Vehicle-Pedestrian Interaction (HSTGA). HSTGA models pedestrian-vehicle interactions at non-signalized and non-crosswalk scenarios using a trajectory-based model for long-horizon pedestrian and vehicle trajectory prediction.
2. We develop a vehicle-pedestrian interaction feature extraction model using a multi-layer perceptron (MLP) sub-network and max pooling.
3. We develop an LSTM network to learn the vehicle-pedestrian spatial interaction adaptively.
4. We predict pedestrian and vehicle trajectories by modeling the spatial-temporal interactions between pedestrian-pedestrian, vehicle-vehicle, and vehicle-pedestrian

using only the historical trajectories of pedestrians and vehicles. This approach reduces the information requirements compared to other learning-based methods.

5. We quantify pedestrian-vehicle conflicts at non-signalized non-crosswalk scenarios using safety indicators: PET and TTC.
6. We use the evolution of PET and TTC indicators over time to predict future collisions between vehicles and pedestrians well before (several seconds before) they occur.

The following publications have resulted from the underlying research performed to complete this dissertation:

1. H. Alghodhaifi, S. Lakshmanan, S. Baek, and P. Richardson, "Autonomy modeling and validation in a highly uncertain environment," in Proceedings of the 2018 Ground Vehicle Systems Engineering and Technology Symposium (GVSETS), 2018.
2. S. Lakshmanan, Y. Yan, S. Baek, and H. Alghodhaifi, "Modeling and simulation of leader-follower autonomous vehicles: environment effects," in Unmanned systems technology XXI, SPIE, 2019, pp. 116–123.
3. H. Alghodhaifi and S. Lakshmanan, "Safety model of automated vehicle-VRU conflict under uncertain weather conditions and sensors failure," in Unmanned Systems Technology XXII, SPIE, 2020, pp. 56–65.
4. H. Alghodhaifi and S. Lakshmanan, "Simulation-based model for surrogate safety measures analysis in automated vehicle-pedestrian conflict on an urban environment," in Autonomous Systems: Sensors, Processing, and Security for Vehicles and Infrastructure 2020, SPIE, 2020, pp. 8–21.

5. E. Cheek, H. Alghodhaifi, C. Adam, R. Andres, and S. Lakshmanan, “Dedicated short range communications used as fail-safe in autonomous navigation,” in Unmanned systems technology XXII, SPIE, 2020, pp. 159–177.
6. H. Alghodhaifi and S. Lakshmanan, “Autonomous vehicle evaluation: A comprehensive survey on modeling and simulation approaches,” IEEE Access, vol. 9, pp. 151531–151566, 2021.
7. H. Alghodhaifi and S. Lakshman, “HSTGA: Holistic Spatio-Temporal Graph Attention Trajectory Prediction for Vehicle-Pedestrian Interaction,” Under preparation. I will submit it to MDPI Sensors on May 15th, 2023.
8. H. Alghodhaifi and S. Lakshman, “A Novel Trajectory-based Approach for Intelligent Vehicle-Pedestrian Collision Avoidance System,” Under preparation. I will submit it to MDPI Imaging on May 20th, 2023.

1.2 Related Work

In this section, we review the existing work on trajectory prediction of vehicle and pedestrian and vehicle-pedestrian interactions with special emphasis on deep learning methods. Moreover, the vehicle-pedestrian conflicts-based approach is covered.

1.2.1 Pedestrian Trajectory Prediction Methods

Over the last few years, numerous techniques and algorithms have surfaced for predicting pedestrian trajectories, owing to their importance in creating a secure environment for autonomous vehicles and other applications. The research on this topic can be broadly classified into three groups [60], [61]:

1.2.1.1 Physics-based models.

1.2.1.2 Planning-based models.

1.2.1.3 Pattern-based models.

1.2.1.1 Physics-Based Models

Physics-based models leverage motion properties such as speed and location to predict future movements by applying physical laws. For example, Kim et al. utilized a Kalman filter and machine learning-based approach that used velocity-space reasoning to compute the desired velocity of pedestrians, which resulted in a good performance [62]. Zanlungo et al. proposed a social force-based model that predicted pedestrian locations while modeling walking behaviors using the social force paradigm and physical constraints. However, the model's performance could suffer when pedestrian density is low [63]. A. Martinelli et al. proposed a pedestrian dead reckoning method that relied on step length estimation [64]. Using the classification of walking behavior, the individual's step length was estimated and used to infer their position. Similarly, W. Kang et al. demonstrated a smartphone-based method for pedestrian position inference that used step length estimation-based inference, which was effective in indoor environments but accrued errors over long distances [65]. Additionally, Gao et al. developed a probabilistic method for indoor position estimation that relied on Wi-Fi signal fingerprints and smartphone signals, which improved accuracy and overcame signal changes [66]. However, most physics-based models rely on manually specified parameters or rules, which limits their application to scenarios such as predicting trajectories in a closed space. In contrast, our proposed graph attention model (HSTGA) learns trajectory patterns from historical trajectory profiles without relying on manually specified parameter values.

1.2.1.2 Planning-Based Models

In the realm of pedestrian trajectory prediction, planning-based models are typically geared toward reaching a specific destination. Ziebart et al. [67] devised a planning-based model that incorporated a distribution of destinations and utilized a Markov decision process to plan and predict trajectories, outperforming a variable-length Markov model in predicting 3 seconds trajectories [68]. Deo and Trivedi implemented a probabilistic framework called the variational Gaussian mixture model (VGMM) [69] that utilized trajectory clustering to predict pedestrian paths, outperforming a monolithic VGMM. Rehder et al. utilized deep neural networks in their planning-based approach, inferring a mixture density function for possible destinations to conduct goal-directed planning [70]. However, this method may not perform well in long-term horizon predictions. Dendorfer et al. proposed a two-phase strategy called goal-GAN, which estimated goals and generated predicted trajectories [71]. Yao et al. improved the performance of their model using a bidirectional multi-modal setting to condition pedestrian trajectory prediction on goal estimation [72]. Tran et al. separated their model into two sub-processes: a goal process and a movement process, enabling good performance in long-term trajectory prediction [73]. However, these models' reliance on guessing a pedestrian's future goals may hinder their performance in longer horizon predictions, unlike our proposed model, which does not speculate future goals or destinations, thus improving prediction accuracy and generalization ability.

1.2.1.3 Pattern-Based Models

In recent years, pattern-based models have gained popularity thanks to advances in deep learning. Most studies have focused on creating modules to learn about the interactions and social features between pedestrians, which contribute directly to individuals’ movements. One notable model is the social LSTM, proposed by Alahi et al., which can predict human trajectories in crowded spaces with high accuracy [41]. It captures social interactions by using a social pooling strategy to identify patterns, and it assumes that interactions between pedestrians can be captured with pooling layers in the model's architecture. In a comparable manner, authors in [74] implemented a distinct scaling technique to apprehend the impact of the surroundings on a particular pedestrian. Another model, social GAN, was introduced by Gupta et al., which used generative adversarial networks (GAN) to learn about interaction patterns between pedestrians and predict their trajectories [42]. This model predicted multiple possible future trajectories and chose the best one. Zhang et al. proposed the state refinement module SR-LSTM to decode implicit social behaviors among pedestrians [75], while Zhao et al. suggested the multi-agent tensor fusion model (MATF) to identify social and interactive relationships by aligning spatial encoding with agent encoding [76]. The multi-agent fusion model (MATF) synchronized the spatial encoding of scenes with the encoding of each agent present within the scene and then utilized a GAN model to acquire knowledge of patterns and make predictions.

Nikhil and Morris also presented a CNN-based model that was computationally efficient and allowed fast parallel processing with competitive performance [44]. Huang et al. extended the temporal correlation concept to produce more socially plausible trajectories [77]. Xu et al. devised a cutting-edge methodology hinged on deep neural networks that harnesses the intricate nature of social behaviors to anticipate pedestrian movements [78]. The researchers deftly employ encoding schemes to distinguish the varying degrees of influence exerted by different social interactions on the trajectory of passersby. Song et al. came up with a complex LSTM network that uses deep convolutional techniques [79]. This algorithm utilizes tensors to represent environmental features and incorporates a specially designed convolutional LSTM to predict sequences of trajectories. Quan et al. introduced an innovative perspective in trajectory forecasting using a model based on

Long Short-Term Memory (LSTM) [80]. Their approach featured a distinctive LSTM mechanism that accurately identified pedestrians' intentions and generated corresponding trajectory predictions.

Existing models require information from all pedestrians on the scene but don't consider the impact of surrounding vehicles and the interaction between these vehicles and pedestrians on pedestrian trajectory prediction. Our approach considers these factors and uses minimal information and a decentralized method, only utilizing the pedestrian's trajectory profile for whom the prediction is being made. The model assumes all other factors affecting the pedestrian's movement are unknown or uncertain, and it learns to adapt accordingly. This decentralized approach ensures that our model can provide high-quality predictions in various environments, not just crowded spaces, making it an ideal choice for practical pedestrian safety applications.

1.2.2 Vehicles-Pedestrians Interaction

Limited research has been done on the interaction between pedestrians and vehicles in trajectory prediction. Previous works have mostly focused on social interaction between pedestrians [41]-[43] and interaction with the environment [74], [81], [83]. However, the interaction between pedestrians and vehicles is an equally important factor that needs to be considered. Some researchers have tried to include vehicle information in pedestrian trajectory prediction, but their methods have limitations. Eiffert et al. [46] improved pedestrian trajectory prediction by encoding interactions between pedestrians and a single vehicle using a feature learning network called the "Graph pedestrian-vehicle Attention Network." However, this method only considered a single vehicle on the road, not multiple vehicles. On the other hand, Chandra et al. [47]-[49] and Carrasco et al. [50] proposed models that predict the trajectories of heterogeneous traffic agents, including pedestrians, but their primary focus was on vehicles and motorcycles rather than pedestrians. Therefore, there is still a need for more research on the interaction between pedestrians and vehicles in trajectory prediction.

1.2.3 Intelligent Vehicle Trajectory Prediction

The use of perception systems, cameras, and intelligent vehicular systems has made obtaining data from driving agents and the environment easier. However, relying solely on a traffic agent's trajectory history for prediction can lead to errors, especially in complex scenarios. Interaction-aware approaches that consider inter-agent interactions and behavior dependencies lead to higher prediction accuracy [84]. In this section, we center our attention on trajectory prediction studies that prioritize interaction awareness.

1.2.3.1 Interaction-Aware Trajectory Prediction

Numerous studies have endeavored to enhance interaction awareness for trajectory prediction approaches by modeling inter-agent correlations among all agents in a driving scene. The early literature on interaction-awareness employed traditional approaches, such as classical machine learning models, for example, Hidden Markov Models (HMM), Support Vector Machines (SVM), and Bayesian networks [87]-[90]. Nevertheless, these conventional methodologies exhibit suboptimal performance in long-term predictions, particularly for intricate scenarios, and are ill-suited for real-time analysis [91].

The employment of deep learning models, specifically Recurrent Neural Networks (RNNs), Temporal Convolutional Neural Networks (CNNs), and Graph Neural Networks (GNNs), has captured the interest of scholars owing to their effectiveness and versatility in various research fields, notably in predicting vehicle trajectories in complex settings. Additionally, the literature proposes a variety of techniques to model the inter-agent interactions for vehicle trajectory prediction. One such approach involves explicitly incorporating the trajectory history of the Target Agent (TA) and its Surrounding Agents (SAs) into the model [92]–[97], in order to consider the impact of SAs. For instance, Dai et al. [92] propose a two-group LSTM-based RNNs approach to model the interactions between the TA and each of its neighbors, and subsequently predict the future trajectory of the TA based on its trajectory history. Another approach, TrafficPredict, is introduced by Ma et al. [93], wherein a system architecture with two layers of LSTM recurrent units is designed to obtain the motion patterns of traffic participants and identify similar behavior among the same group of traffic participants, such as vehicles or bicycles. These methods have limitations as they fail to account for the effect of the environment and traffic regulations on the TA's behavior.

A potential alternative strategy for modeling social interactions among a large number of traffic participants in a given scenario involves the implementation of a social pooling mechanism [41], [42], [98]. This mechanism permits neighboring agents' LSTM units to share knowledge with one another. Alahi et al. [41] propose the S-LSTM method, which enables the recurrent units associated with SAs to connect with one another via the design of a pooling layer between each existing LSTM cell. This method facilitates the sharing of hidden states among the agents in an occupancy grid. To model pair-wise interactions between all SAs in a given scene, Gupta et al. [42] propose another pooling technique (S-GAN) based on Multi-Layer Perceptron (MLP) and max pooling. This method computes a global pooling vector for each TA based on relative coordinates between the TA and all of its SAs, as well as their hidden states. Deo et al. [98] introduce an encoder architecture, CS-LSTM, for vehicle trajectory prediction. This architecture employs convolution and max pooling operations over a spatial grid that depicts the TA's surroundings. However, the vehicles' extracted representations are independent of their states, and the local computations are inefficient. Messaoud et al. [99], [100] propose using the Multi-Head Attention (MHA) pooling method to address this issue. They generate a representation vector for each vehicle using an LSTM-based encoder. An MHA module is then employed to consider the inter-vehicle correlations between the target vehicle and its SAs within a spatial grid. It has been demonstrated that MHA reduces the number of local computations. Nevertheless, these methods' lack of efficiency in addressing complex spatiotemporal correlations among traffic participants is a significant drawback. Additionally, the performance of these methods can be affected by the distance used to generate the occupancy grid or the number of SAs considered.

1.2.3.2 Graph-Based Interaction Reasoning

Recently, the research area of trajectory prediction has seen a growing interest in graph-based interaction reasoning as an alternative approach to address the limitations of interaction-aware path prediction methods, as discussed in the previous section. Graph-based approaches have focused on modeling interactions between various agents within a driving scene as graphs, where nodes represent agents and edges represent inter-agent interactions. This allows for simultaneous consideration of spatial and temporal inter-agent correlations. For instance, Diehl et al. represented a highway-driving scene as a directed graph and compared the performance of GAT and GCN for traffic prediction, considering a fixed number of surrounding vehicles [101]. However, the homogeneous graph generated by the authors disregards the dynamics and types of vehicles. Li et al. generated a homogeneous undirected graph to represent inter-vehicle interactions and used graph convolutions to extract underlying features within the data [102]. The future trajectory of the vehicles was predicted using an LSTM-based decoder. Nonetheless, the method shares the aforementioned limitation. Azadani et al. utilized undirected spatiotemporal graphs to model inter-vehicle interactions and analyzed the trajectory history of target vehicles and their surrounding vehicles using graph and temporal gated convolutions [103]. The future trajectory of the vehicle agents was then predicted using temporal convolutions applied to the extracted latent representations.

In recent research, Wu et al. [104] have proposed an encoder-decoder architecture that takes into account temporal interdependencies using Multi-Head Attention (MHA) and spatial interactions with Graph Attention (GAT) modules. The resulting outputs from these separate modules are then aggregated and fed into a Long Short-Term Memory (LSTM)-based decoder. Similarly, Li et al. [105] have introduced the STG-DAT system, which comprises three key modules, namely feature extraction using Multilayer Perceptron (MLP), representation extraction using GAT as an encoder, and path generation employing Gated Recurrent Units (GRU) while considering the kinematic constraints.

Furthermore, Mo et al. [106] have developed a directed graph for various groups of agents, wherein individual encoders are considered for different types of agents within the driving scene, as the unique behavior of each agent type can affect their future trajectory patterns. In a similar vein, Sheng et al. [107] have constructed a distance-based weighted graph for the target agent (TA) and its surrounding vehicles. The spatial graph is then analyzed using Graph Convolutional Networks (GCN), while GRU units are utilized to generate the future trajectory of the vehicles. Moreover, Gao et al. [108] have created heterogeneous sub-graphs for each agent and a high-order graph to model the inter-agent interactions. However, the generated dense graph overlooks the spatial and edge features among the agents. These recent advancements in modeling temporal and spatial interactions among agents have shown promising results in predicting future trajectories in complex environments.

Prior research on trajectory prediction has yielded interaction-aware approaches that are customized for particular contexts and representations. These methods often overlook certain spatial and temporal considerations or rely on dense undirected graphs to depict the inter-agent interactions. Such graphs assume that every vehicle interacts with all other surrounding agents with equal impact. In contrast, our research introduces the HSTGA system, which adopts an asymmetric social interaction reasoning approach that utilizes sparse directed graphs for both vehicles and pedestrians. This innovative system aims to address the aforementioned challenges and enhance the accuracy of trajectory prediction.

1.2.4 Conflict Analysis of Vehicle-Pedestrian Interactions

Perkins and Harris [110] conducted a pioneering study aimed at devising a means to predict road accidents, which can provide valuable insights into causal factors related to traffic safety issues. The authors identified various potential accident situations, which they classified as traffic conflicts, and distinguished over twenty types of traffic conflicts that arise between road users, primarily by analyzing evasive actions such as swerving, stopping, and braking. In the context of traffic conflicts based on evasive actions, Johnsson et al. [111] discussed various surrogate measures of safety found in the literature, with few of these indicators focusing on aspects that are useful in investigating issues pertaining to Vulnerable Road Users (VRUs). The authors evaluated various safety indicators based on their ability to consider both injury risk and collision risk, taking evasive actions into account. Their study revealed that several indicators concentrate on braking as a critical indicator to define hazardous traffic situations while disregarding other types of evasive actions, such as running or swerving. The use of non-crash situations to examine road safety is rooted in the assumption that there is a relationship between the severity and frequency of traffic events [112]. Hyden [109] introduced the concept of a "safety pyramid," which refers to a hierarchical representation of traffic events, with the most severe events located at the top, typically defined as "accidents," followed by traffic conflicts, which are categorized into severe, slight, or potential conflicts based on their level of risk. Finally, the vast majority of traffic encounters are considered natural events. Fig. 1.5 presents the safety pyramid with the severity levels of traffic events [109].

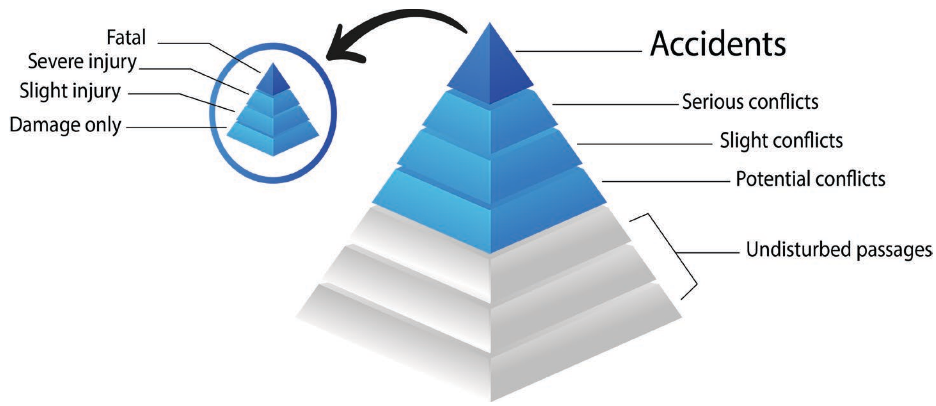


Figure 1.5 Conceptual Safety Pyramid [109].

Surrogate safety measures (SSMs) have been introduced to produce more analysis without relying on accident statistics alone. Moreover, the future goal is that these safety analysis measures will also be utilized to design autonomous driving algorithms. The term "surrogate" denotes that the indicators utilized in traffic safety analysis do not rely on crash databases. Instead, they are intended to be supplementary tools for historical records analysis. Throughout the 1970s and 1980s, several researchers proposed and developed various indicators. Several traffic safety indicators have been suggested and developed in recent years, including those associated with vulnerable road users (VRUs). Given the existence of numerous issues related to VRU safety analysis, such as underreporting of VRU crashes, transport modes have become increasingly cognizant of this matter [111]. Several studies have described, summarized, and compared a group of safety indicators. For instance, Laureshyn et al. [113], [114] provided an overview of nearness-to-collision and severity indicators. Zheng et al. [13] clarified traffic situations' temporal and spatial proximity characteristics. Whether a particular situation is classified as a traffic conflict is contingent on the proximity in the distance and/or time of the relevant road users. Ceunynck [115] presented previous research on the application of safety indicators and examined their frequency of use. He organized the safety indicators into groups based on the Time-To-Collision (TTC), the Post-Encroachment Time (PET), the deceleration families, other groups, and unspecified indicators. According to Ceunynck's [115] research, safety indicators originating from the TTC family are most commonly utilized in safety analysis, followed by those from the PET family.

According to Hayward's [116] definition, Time to Collision (TTC) refers to the duration required for two vehicles to collide while maintaining their current speed and trajectory. TTC at the onset of braking, referred to as TTCbr, denotes the available space for maneuvering at the moment of evasive action initiation. The minimum TTC attained during the approach of two vehicles on a collision course (TTCmin) indicates encounter severity. Lower TTCmin values indicate higher collision risk. Van der Horst [117] provides a comprehensive account of TTC calculation procedures. TTC is a crucial factor in various fields, such as smart transportation, robotics, and autonomous vehicle safety. Therefore, numerous researchers have dedicated their efforts to investigating and developing effective TTC estimation methods [119]-[122].

Tarko et al. [123] presented a white paper that underscored the significance of the observability or measurability of a surrogate measure in the traffic system. A measure that has garnered considerable attention in this regard is post-encroachment time (PET), as it offers a consistent and reliable means of measurement across different observers and locations. PET, which is defined as the time difference between the end of encroachment of the first vehicle and the entry of the second vehicle into the area of conflict, is a simple yet effective measure that facilitates the differentiation of crash from non-crash events [124]. In particular, a PET value of 0 indicates a collision, while non-zero values indicate crash proximity. Although PET does not capture the initial stage of the conflict or the actions of the involved drivers, it provides a measure of the closeness to a collision at the final stage. Post-encroachment time (PET) has gained significant attention in recent years, and numerous studies have investigated its potential applications [125]-[127].

Chapter 2

Vehicle-Pedestrian Interaction Problem

2.1 Problem Statement

Given the trajectories of pedestrians and vehicles in the past m frames, our goal is to predict their trajectories in the future h frames and then use these trajectories to design a conflict-based model for vehicle-pedestrian interaction. Both the prediction and the model are critical to predicting conflicts between vehicles and pedestrians well before they occur.

2.2 Trajectory Prediction

Anticipating the future trajectories of pedestrians is essential for the safe navigation of connected and autonomous vehicles and also for the safety of pedestrians crossing the road. Accurate prediction of pedestrians' future trajectories enables connected and autonomous vehicles to have more reaction time to take action, which is one of the key advantages of intelligent system technology. Overall, anticipating the future trajectories of pedestrians is crucial for ensuring all road users' safety, efficiency, and comfort. By effectively predicting and responding to pedestrian behavior, connected and autonomous vehicles can navigate the environment more effectively and reduce the risk of collisions.

The problem of predicting the future trajectories of pedestrians is challenging due to their stochastic behavior patterns. One of the challenges connected and autonomous vehicles face when interacting with pedestrians is that pedestrians can be unpredictable. For example, a pedestrian may suddenly step into the road without looking or may start running across the street to catch a bus. These types of behaviors can be difficult for human drivers to anticipate, let alone respond to quickly enough to avoid a collision. The problem of predicting pedestrians' future trajectories also includes many other factors that can impact their behavior, such as the movement of surrounding pedestrians and vehicles. Therefore, understanding the pedestrian-pedestrian, vehicle-pedestrian, and vehicle-vehicle interactions are critical in achieving accurate pedestrian trajectory prediction.

The pedestrian-pedestrian, vehicle-pedestrian, and vehicle-vehicle interactions in a dense environment is a multilayer task. At the same time step t, the spatial and temporal continuity of interactions must be captured. As seen in Fig. 2.1, three pedestrians p_1, p_2, and p_3 are crossing in front of two vehicles, v_1, and v_2. A serious conflict could readily arise in this kind of situation. We must correctly anticipate their trajectories by taking into account a variety of factors that affect their movements, such as timing, speed, visibility, distance, and human behavior, in order to determine whether there will be a crucial conflict. Timing is an important consideration because we need to know if vehicles and pedestrians are passing an area at the same time. A critical conflict is highly likely when both are crossing at the same moment. The risk is reduced if they cross at separate times, though.

Speed is also an important element because fast-moving vehicles can cover more ground in a shorter amount of time, which increases the possibility of a serious collision. Conversely, pedestrians who move slowly across the street or those who have already traversed it face less danger. A significant factor in avoiding collisions is visibility. A critical conflict is more likely to occur if the area is dimly lit or blocked by structures, trees, or other items. Another important factor is the distance between pedestrians and vehicles. This factor has a major impact on defining the type of interactions. A small distance between pedestrians and vehicles is an important indicator of a higher probability of critical conflicts. In many ADAS applications, the minimum distance is used widely as a safety indicator. The last factor is pedestrian behavior, which is one of the most challenging factors to predict. To ultimately avoid deadly conflicts, connected and autonomous vehicles must predict pedestrian behavior several seconds into the future. The future prediction of pedestrian behavior will give connected and autonomous vehicles more time to react and take appropriate action to avoid collisions. In general, predicting pedestrian behavior for a normal pedestrian is a difficult job. Pedestrians who are distracted are more likely to make errors or take unnecessary risks. This will also make behavior modeling and prediction more challenging. Therefore, one way to tackle this problem is to incorporate all aforementioned factors into the pedestrian behavior prediction task. Moreover, we also need to capture the spatial and temporal interactions between pedestrians themselves and between pedestrians and vehicles at the same time-step t to model and predict the behavior of pedestrians and vehicles accurately.

In order to comprehensively apprehend the complex dynamics present in Fig. 2.1, it is imperative to capture the intricate and interdependent spatial and temporal interactions between all objects therein. The movements of both pedestrians and vehicles are not only interdependent but also intricately linked, with each exerting a profound influence on the other. Without capturing the spatial and temporal interactions between pedestrians and vehicles, our understanding of the multifaceted dynamics of their movements and the factors that impel their behavior remains incomplete. This is precisely why connecting and automating vehicles necessitates a better understanding of the interactions that exist between these entities and how they can be harnessed to make more precise predictions about their trajectories, as well as the likelihood of conflicts or accidents. However, it is vital to recognize that considering solely the positions and velocities of pedestrians and vehicles at a given time is inadequate. This simplistic approach ignores the critical information that spatial and temporal interactions between these entities can provide. Thus, by capturing the spatial and temporal interactions between pedestrians and vehicles, we can gain a more in-depth understanding of how they will move and interact with one another in the future. Some specific examples of spatial and temporal interactions that need to be captured include pedestrian-vehicle, pedestrian-pedestrian, and vehicle-vehicle interactions.

Pedestrian-vehicle interactions exemplify the intricate and multifaceted ways in which pedestrians and vehicles interdependently influence each other's movements. A classic example is when a vehicle is rapidly approaching a group of pedestrians, where the pedestrians, in response, alter their movement, forcing the vehicle to adapt its trajectory and decelerate to avoid colliding with them. Such intricate interactions serve as an essential component in understanding the dynamics of pedestrian and vehicle movements. Similarly, pedestrian-pedestrian interactions play a significant role in shaping the complex dynamics of pedestrian movements. For instance, abrupt halts by a pedestrian will force other pedestrians behind them to make quick adjustments to avoid collisions. These interactions are often unpredictable, leading to the possibility of accidents, and hence need to be comprehensively captured to avoid any future mishaps. A vital aspect of capturing the complex movements of pedestrians and vehicles involves fully comprehending the nuances of vehicle-vehicle interactions, where the movements of two vehicles can have a profound impact on each other. For instance, consider a scenario where two vehicles confront each other while approaching an area where three pedestrians are crossing in different directions simultaneously. This presents a formidable challenge as both vehicles will need to make quick, precise adjustments, such as decelerating or changing direction, to avoid a calamitous collision. The sheer complexity of this interaction highlights the imperative need to accurately capture the spatial and temporal interactions between the vehicles in Fig. 2.1. Failing to do so could result in grave accidents and underscore the urgency to comprehensively address these interactions. By capturing these types of interactions, we can create more accurate models and simulations of pedestrian and vehicle movements and make better predictions about their trajectories and use these trajectories to forecast potential conflicts or accidents. This information can be used to develop strategies to reduce the likelihood of accidents and improve the safety of pedestrians and connected and autonomous vehicles in a given scenario.

As our cities grow and become more crowded, ensuring the safety of pedestrians has become an increasingly important concern. This is particularly true when it comes to predicting pedestrian trajectories, where the ability to accurately capture the spatial and temporal interactions between pedestrians and vehicles is crucial. At the heart of this challenge lies the need to capture these interactions at the same time step t. This is important because pedestrians and vehicles move in complex and dynamic ways, with each influencing the other's movement. Without this synchronized data, our ability to accurately predict pedestrian trajectories is severely compromised. Moreover, the intricate and complex dynamics of urban environments necessitate a deep understanding of the spatial and temporal interactions between pedestrians and vehicles. It's essential for each pedestrian and vehicle to plan its future moves at each time step t in a way that ensures safety and reduces any critical conflicts. As shown in Fig. 2.1, vehicles, v_1 and v_2, should be aware of the presence of pedestrians and adjust their speed and trajectory accordingly to avoid any potential collisions. Vehicles should also be prepared to stop if necessary to avoid hitting a pedestrian.

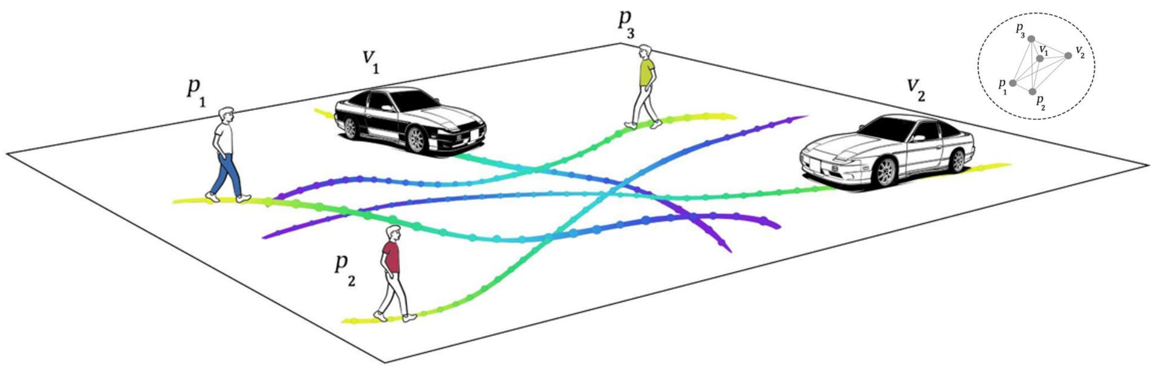


Figure 2.1 Illustration of pedestrian-vehicle interactions in the crowd. The same color on trajectories means the same time-step t.

The pedestrians, denoted by the labels p_1, p_2, and p_3, must maintain a heightened sense of awareness regarding the presence of vehicles and strategically plan their crossing maneuvers accordingly. This entails meticulously scanning their surroundings and preemptively evaluating the feasibility of their trajectories. The pedestrians must be prepared to dynamically adjust their planned paths should any exigent circumstances arise, aiming to circumvent any potential collision points. The safety and well-being of all the objects involved in this scenario hinge on their collective capacity to remain cognizant of each other's presence and adroitly adapt to their surroundings. Achieving this outcome necessitates a concerted effort by each object to attentively process the surrounding stimuli and rapidly execute the most prudent course of action.

For example, as shown in Fig. 2.2, the pedestrian p_1 needs to consider the historical movements of pedestrians p_2 and p_3 to determine his/her current and future motion behavior to avoid critical conflicts before they occur. Moreover, the pedestrian p_1 should also consider the historical movements of vehicles v_1 and v_2 to decide his/her future moves. At each time step t, the trajectory of pedestrian p_1 is impacted by the motion of other objects in the environment. For example, if one of the vehicles (v_1 or v_2) changes its velocity or direction of motion, it can impact the trajectory of pedestrian p_1. Similarly, if another pedestrian (p_1 or p_2) changes their speed or direction, it can also impact the trajectory of pedestrian p_1. Thus, predicting the trajectory of all other objects and ensuring the safety of pedestrian p_1 as he/she navigates a scenario with multiple objects is a challenging endeavor that involves the combination of several cognitive and physical abilities. Moreover, predicting the behavior of pedestrian p_1 under these conditions is a complex task.

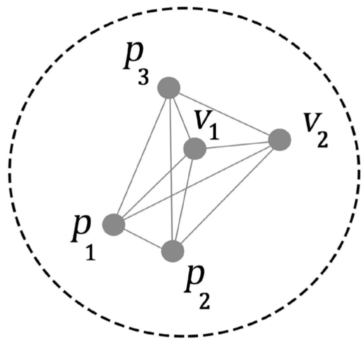


Figure 2.2 Pedestrian-pedestrian, vehicle-pedestrian, and vehicle-vehicle interactions.

To investigate the above problem, let's assume that prior image processing has already been applied to a raw video feed to extract the position and pose of individual pedestrians and vehicles in each video frame. We assume there are N pedestrians and M vehicles present in a video frame, represented as $p_1, p_2, \dots, p_N$ for pedestrians and $v_1, v_2, \dots, v_M$ for vehicles.

The state of pedestrians p_i ($i \in [1, N]$) and vehicles v_k ($k \in [1, M]$) at time-step, t is denoted as follows:

$$A_{obs}^t = [P_1^t,\ P_2^t, \dots,\ P_N^t] \tag{2.1}$$

$$B_{obs}^t = [V_1^t,\ V_2^t, \dots,\ V_M^t] \tag{2.2}$$

where P_i^t and V_k^t are the lateral and longitudinal positions with the heading angle of pedestrian i and vehicle k, respectively, at time-step t. The number of pedestrians i and vehicles k are variables in equations 2.1 and 2.2 because different datasets/ scenarios are used to evaluate this study. Eqs.

1 and 2 are the observed trajectories that will be used as input to our deep learning model. P_i^t and V_k^t are expressed as follows:

$$P_i^t = (x_i^t,\ y_i^t,\ \theta_i^t) \qquad (2.3)$$

$$V_k^t = (x_k^t,\ y_k^t,\ \theta_k^t) \qquad (2.4)$$

In Eqs. 2.3 and 2.4, $x_i^t,\ x_k^t,\ y_i^t, y_k^t, \theta_i^t$, and θ_k^t are the position coordinates and the heading angles of pedestrians and vehicles at each time-step t. The positions of the vehicles and pedestrians are relative to the world space.

Using the observed trajectories A_{obs}^t and B_{obs}^t in the past m frames at time-steps $t = 1, .., T_{obs}$, our goal is to predict the future trajectories A_f^t and B_f^t several seconds ahead in the future h frames at time-steps $t = T_{obs} + 1, \ldots, T_f$ as follows:

$$A_f^t = [P_1^{t+h},\ P_2^{t+h}, \ldots,\ P_N^{t+h}] \qquad (2.5)$$

$$B_f^t = [V_1^{t+h},\ V_2^{t+h}, \ldots,\ V_M^{t+h}] \qquad (2.6)$$

To forecast future trajectories through optimization and to learn pedestrians-vehicles spatiotemporal interaction, our goal is to learn the distribution $p(A_f|A_{obs})$, $p(B_f|B_{obs})$, $p(A_f|B_{obs})$, and $p(B_f|A_{obs})$.

2.3 Vehicle-Pedestrian Conflict-Based Model

Transportation is an important component of everyday living in our modern world. Vehicle-pedestrian interactions have become more regular and complicated as cities have grown in size. It is critical to have a clear knowledge of these interactions and be able to anticipate possible disputes in order to guarantee the safety and efficiency of the transportation system. Conflict-based models for vehicle and pedestrian interactions are proposed to investigate this type of interaction. A conflict-based model is a simulation framework that predicts vehicle and pedestrian conflict encounters in common spaces such as roads and junctions. The model considers a variety of variables, including both vehicles' and pedestrians' behavior and choices, real natural limitations, and traffic rules and regulations. By simulating potential conflicts, the model can help identify areas where safety risks are high and suggest measures to mitigate these risks.

The model can be used to assess how well various traffic control measures, including roundabouts, crosswalks, and traffic signals, promote secure and successful exchanges between vehicles and pedestrians. Understanding the conflicts between pedestrians and vehicles helps in designing and managing roadways, intersections, and pedestrian facilities. Urban planning is a significant area where conflict-based models are used. The intricate interactions that occur between vehicles and pedestrians in urban settings must be taken into account when designing a public transportation system.

The advancement of intelligent vehicles has also raised the significance of models that focus on conflict. In shared areas, intelligent vehicles must be able to communicate securely with infrastructure and potentially with pedestrians. Moreover, intelligent vehicle developers can assess these vehicles during the design and deployment in a simulation and real-world scenarios by using conflict-based models. Furthermore, they can create algorithms to guarantee the safe operation of these vehicles on public roads. Vehicle and pedestrian conflict-based models are essential tools for assuring the safety and effectiveness of the transportation system. Moreover, the vehicle-pedestrian conflict-based model is an essential block in providing valuable information for traffic experts, urban planners, and creators of intelligent vehicles. Thus, conflict-based modeling will make transit networks safer and more effective for all users, including vehicles and pedestrians.

To design an accurate conflict-based model for vehicle-pedestrian interactions, the movements of the vehicles and pedestrians at the same time-step should be fed into the conflict model. Feeding this type of data assists in investigating if there is any critical conflict between vehicles and pedestrians at each time step. The movements at each time-step can be represented as the trajectories of the vehicle and pedestrian at each time-step. Thus, designing the conflict-based model requires accurate prediction of vehicle and pedestrian trajectories under highly uncertain environments. Therefore, the conflict-based model has two stages namely, the trajectory prediction stage and the conflict-based stage.

Using the predicted trajectories A_f^t and B_f^t from the trajectory prediction stage, a conflict-based model is required to avoid any deadly conflicts several seconds into the future. Using the predicted trajectories from Eq. 2.5 and Eq. 2.6, we assume that we have a traffic network consisting of a road network for each agent such as a pedestrian or a vehicle. The road network of each agent is defined based on its predicted trajectory from the trajectory prediction stage. Two types of interactions are already considered in the trajectory prediction stage, namely vehicle-vehicle, and pedestrian-pedestrian interactions to predict the trajectories of the vehicle and pedestrian accurately and then design an efficient conflict-based model for vehicle-pedestrian interaction. The pedestrians' road networks and the vehicles' road networks interact in a scenario where pedestrians' trajectories and vehicles' trajectories can conflict. This type of scenario has no pedestrian crosswalk and no traffic lights. The two networks are expressed as shown below.

$$G^{ped}(A_f^t, L^{ped}) \quad (2.7)$$

$$G^{veh}(B_f^t, L^{veh}) \quad (2.8)$$

$G^{ped}(.)$ and $G^{veh}(.)$ in Eqs. 2.7 and 2.8 are the two graph road networks for pedestrians and vehicles at each time-step t. Aft and Bft denote the predicted trajectories from the trajectory prediction stage in Eqs. 2.5 and 2.6 represented as nodes at each time-step t. L^{ped} and L^{veh} are the edge set or the link set between nodes as shown below in Fig. 2.3. The edge or link set can provide extensive information to understand the interaction between pedestrians-pedestrian, vehicle-vehicle, and vehicle-pedestrian. For example, as shown in Fig. 2.3, the vehicle v_1 has four edges or links which means this vehicle interacts with four agents at the same time step. The four directed edges or links are namely, between vehicles v_1 and v_2, v_1 and p_1, v_1 and p_2, and between v_1 and p_3.

The goal is to investigate and predict vehicle-pedestrian conflict under a highly uncertain environment in the conflict region C_{ij} (Fig. 2.3). As shown in Fig. 2.3, the two vehicles v_1 and v_2 will potentially have a conflict with p_1, p_2, and p_3 under highly uncertain conditions. Furthermore, the spatiotemporal interactions between vehicle-pedestrian, vehicle-vehicle, and pedestrian-pedestrian should be captured during the trajectory prediction stage to be able to estimate the vehicle-pedestrian conflict in the conflict region C_{ij}. Figures 2.4, 2.5, and 2.6 present vehicle-pedestrian, vehicle-vehicle, and pedestrian-pedestrian interactions with the associated road networks for each object.

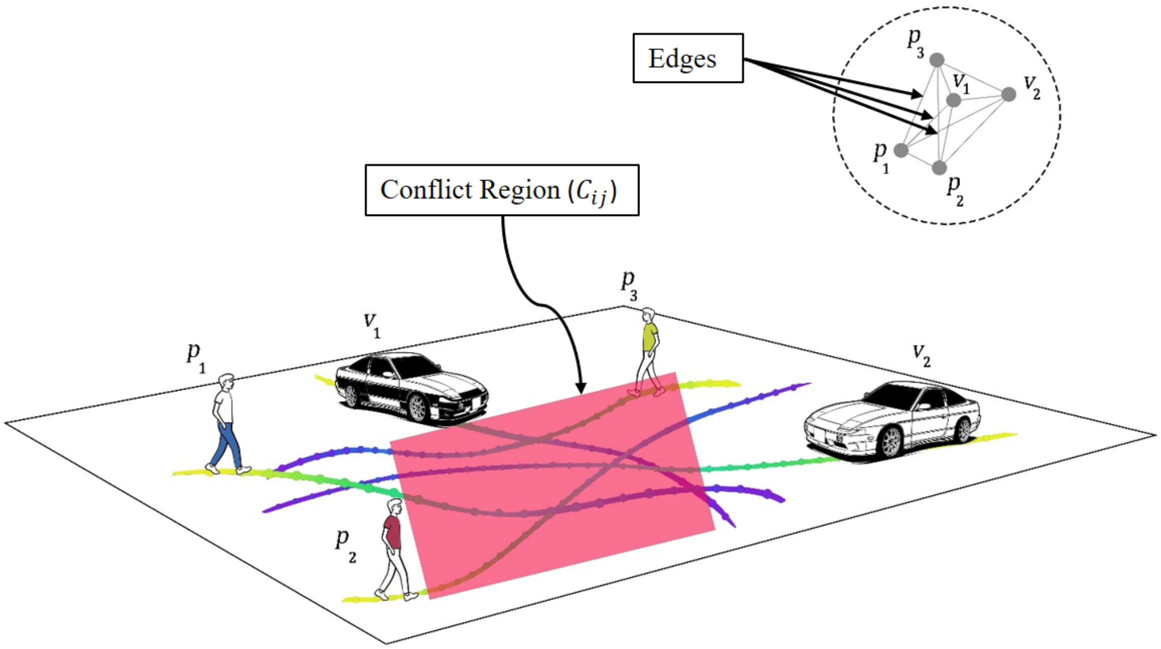


Figure 2.3 Pedestrian-vehicle conflict zone.

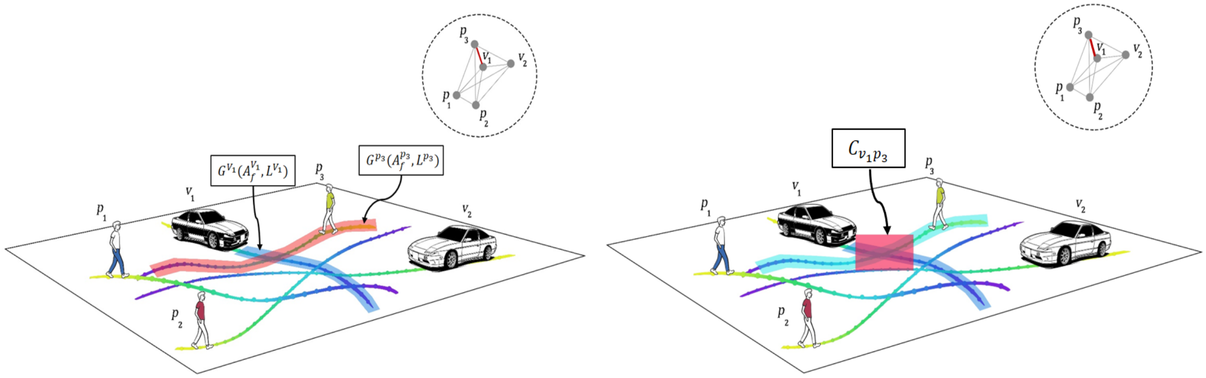


Figure 2.4 Pedestrian and vehicle road networks interact to create traffic graph networks.

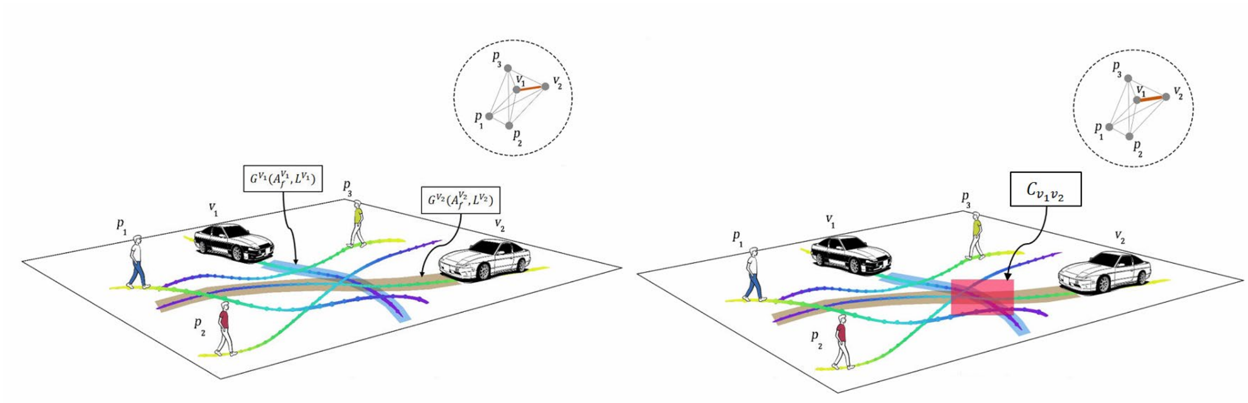


Figure 2.5 Two vehicles' road networks interact to create traffic graph networks.

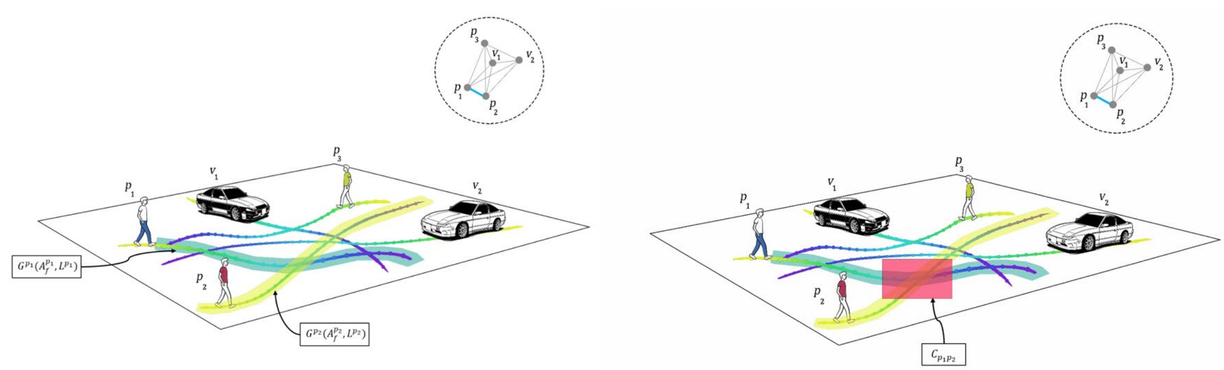


Figure 2.6 Two pedestrians' road networks interact to create traffic graph networks.

Our goal is to track pedestrians and vehicles at the same time step t and identify when and where the vehicle and pedestrian have a future conflict by calculating safety measures indicators at each time step t.

Chapter 3

Holistic Spatial-Temporal Graph Attention Trajectory Prediction for Vehicle-Pedestrian Interaction

3.1 Introduction

The development of advanced intelligent driving systems requires accurate prediction of pedestrian and vehicle trajectories in complex urban scenarios. This task becomes even more challenging in non-signalized and non-crosswalk scenarios where pedestrians and vehicles interact in a highly dynamic and complex manner. We introduce a novel encoder-decoder interaction model called Holistic Spatial-Temporal Graph Attention Trajectory Prediction for Vehicle-Pedestrian Interaction (HSTGA) to address this challenge.

HSTGA is a trajectory-based model that predicts long-horizon pedestrian and vehicle trajectories by modeling the social interactions between pedestrians themselves and between pedestrians and vehicles. Unlike other learning-based methods that require extensive amounts of data and detailed scene understanding, HSTGA only utilizes the historical trajectories of pedestrians and vehicles to predict their future movements. This approach significantly reduces the information requirements and computational costs while maintaining high prediction accuracy. Our model considers various factors that affect pedestrian-vehicle interactions, such as pedestrian crossing intentions, vehicle speeds, and spatial constraints. We model these interactions as a holistic spatiotemporal graph that captures the social interactions between pedestrians and vehicles. HSTGA uses directed graph attention mechanisms to dynamically weigh the importance of different nodes and edges in the graph based on their contextual relevance.

Overall, the proposed HSTGA model represents a significant step forward in accurately and efficiently predicting pedestrian and vehicle trajectories in non-signalized and non-crosswalk scenarios or any other scenario. This model can potentially enhance the safety and reliability of autonomous driving systems and improve the overall driving experience for pedestrians and drivers. The main contributions of this research are as follows:

1. We develop a novel encoder-decoder interaction model called Holistic Spatio-Temporal Graph Attention Trajectory Prediction for Vehicle-Pedestrian Interaction (HSTGA). HSTGA models pedestrian-vehicle interactions at non-signalized and non-crosswalk scenarios using a trajectory-based model for long-horizon pedestrian and vehicle trajectory prediction.
2. We develop a vehicle-pedestrian interaction feature extraction model using a multi-layer perceptron (MLP) sub-network and max pooling.
3. We develop an LSTM network to learn the vehicle-pedestrian spatial interaction adaptively.
4. We predict pedestrian and vehicle trajectories by modeling the spatial-temporal interactions between pedestrian-pedestrian, vehicle-vehicle, and vehicle-pedestrian using only the historical trajectories of pedestrians and vehicles. This approach reduces the information requirements compared to other learning-based methods.
5. We extensively evaluate the proposed model on diverse datasets featuring numerous challenging scenarios involving the interactions between vehicles and pedestrians.

3.2 Problem Definition

The problem statement is already discussed in chapter 2 section 2.2.

3.3 Methodology

This section provides a general overview of our multi-trajectory prediction model's key components and architecture design, HSTGA. We also delve into the specifics of each module within the framework.

3.3.1 HSTGA Overview

In order to anticipate the trajectories and interactions of pedestrians and vehicles within a given scene, a vehicle-pedestrian features extraction model and a Graph Attention Network (GAT) are employed in conjunction with two separate long short-term memory (LSTM) models, as depicted in Figure 3.1. The first LSTM, referred to as SLSTM, handles the individual trajectories of both vehicles and pedestrians. The GAT, situated between the SLSTM and the second model known as TLSTM, is responsible for capturing interactions between the two objects within the scene. Conversely, the TLSTM is designed to capture the temporal interactions between vehicles and pedestrians.

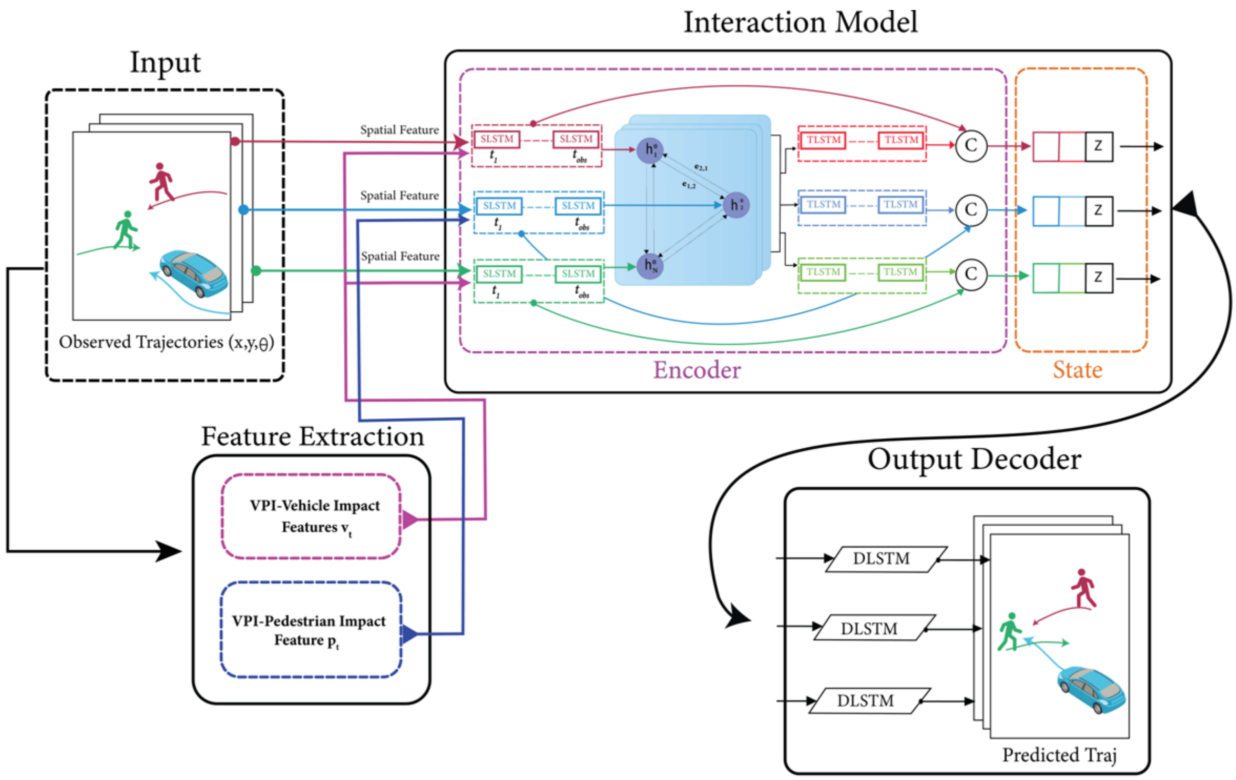


Figure 3.1 Illustration of the vehicle-pedestrian interaction model.

3.3.2 Vehicle-Pedestrian Interaction (VPI) Features Extraction

The interaction between vehicles and pedestrians is a significant factor in predicting their future trajectories. The process of extracting features related to vehicle-pedestrian interaction involves two sections, and every section has two stages, as depicted in Figure 3.2.

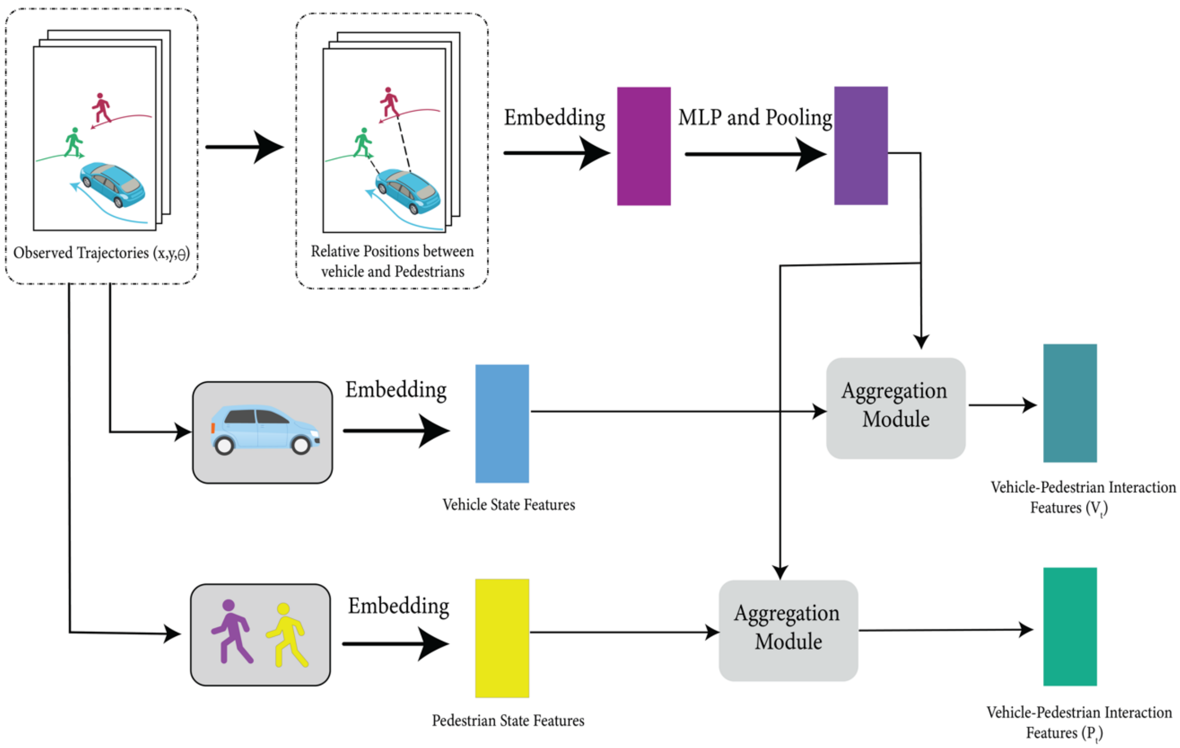


Figure 3.2 Vehicle-Pedestrian Interaction Feature Extraction Model.

The first section extracts the vehicle-pedestrian interaction feature when considering vehicle spatial influence on pedestrians. This section feature is then used with the pedestrian's motion state feature (spatial feature) to be fed to the SLSTM for each pedestrian. In the first stage of this section, the interaction weights between the vehicle and pedestrian are learned using their calculated relative positions. Next, a separate embedding module is used to extract the movement state of the vehicle. Finally, the two stages are combined to obtain the features related to vehicle-pedestrian interaction, which are then fed to the SLSTM for trajectory prediction. On the other hand, the second section extracts the vehicle-pedestrian interaction feature when considering the pedestrian spatial influence on vehicles. The resulting feature from this section is then fed with the vehicle's motion state (spatial feature) to the SLSTM. Stages one and two of both sections are discussed below.

In stage one, the vehicle-pedestrian interaction attention weights vp_{ij}^{t} between the i^{th} pedestrian and the j^{th} vehicle is calculated using Max pooling as shown in Eq. 3.1.

$$vp_{ij}^{t} = Pooling\left\{MLP\left(\phi\left(d_{ij}^{t};W_d\right);W_a\right)\right\}, i \in \{1,\dots,N\}, j \in \{1,\dots,M\} \quad (3.1)$$

$Pooling(\cdot)$ here is the pooling layer, and $MLP(\cdot)$ is the multi-layer perceptron sub-network with weight W_a. Moreover, $\phi(\cdot)$ is the embedding layer with weights. W_d. Finally, the relative position (d_{ij}^{t}) between the pedestrian and the vehicle is then calculated. Eq. 2.3 and 2.4 are used to calculate the relative position using the x and y coordinates and the heading angle θ as shown in Eq. 3.2.

$$d_{ij}^{t} = \left(x_j^{veh,t} - x_i^{ped,t}, y_j^{veh,t} - y_i^{ped,t}, \theta_j^{veh,t} - \theta_i^{ped,t}\right), \quad (3.2)$$

$$i \in \{1,\dots,N\},\ j \in \{1,\dots,M\}$$

In equation 3.2, N and M are the numbers of pedestrians and vehicles, respectively. To predict pedestrian trajectories accurately, we must consider the motion state of the j^{th} vehicle and then aggregate the vehicle-pedestrian interaction weights vp_{ij}^{t} and the vehicle motion states $m_j^{veh,t}$ of the vehicle to obtain the vehicle-pedestrian interaction features or vehicle impact. We calculate the vehicle's motion state using the equation below:

$$m_j^{veh\_ped,t} = \phi\left(\Delta V_j^t; W_{m^{veh\_ped}}\right), j \in \{1,\dots,M\} \quad (3.3)$$

In equation 3.3, $\phi(\cdot)$ component is the embedding with weights $W_{m^{veh\_ped}}$ and ΔV_j^t is the relative position of the j^{th} vehicle between the current and last time-step. The final step is aggregating the vehicle-pedestrian interaction weights vp_{ij}^{t} and the vehicle motion states $m_j^{veh\_ped,t}$ as follows:

$$v_i^t = AGG_{VPI}\left(m_j^{veh\_ped,t}, vp_{ij}^{t}\right), i \in \{1,\dots,N\}, j \in \{1,\dots,M\} \quad (3.4)$$

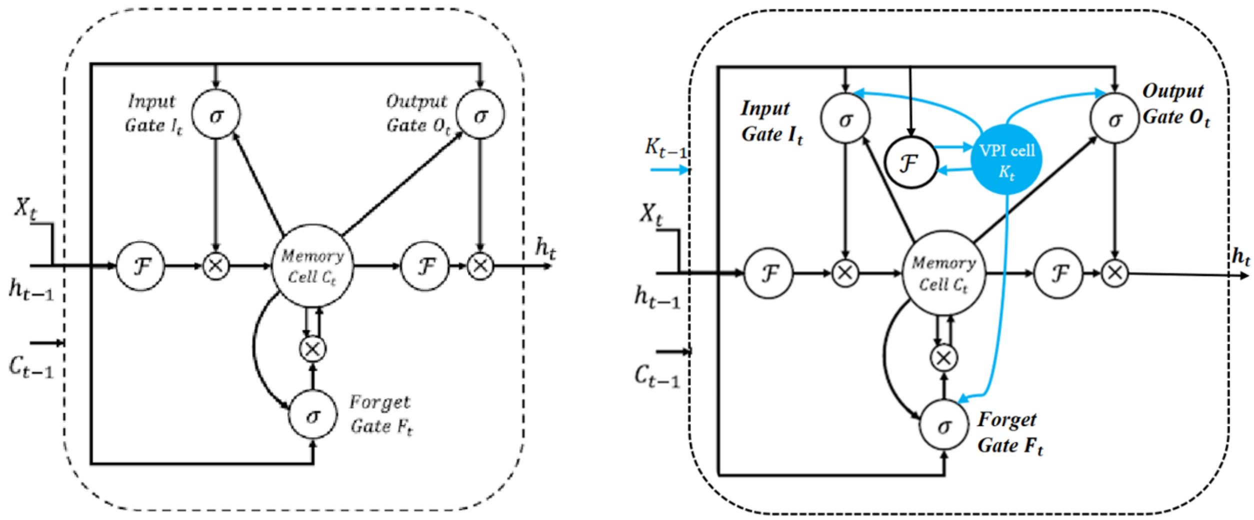
Eq. 3.4 is the vehicle-pedestrian interaction feature when considering the vehicle influence. This feature is then aggregated with the motion state of the individual pedestrian and fed to the SLSTM. For the vehicle-pedestrian interaction feature, when considering the pedestrian influence, the motion state of the pedestrian $m_i^{ped\_veh,t}$ should be calculated and then aggregated with the vehicle-pedestrian interaction weights vp_{ij}^t to get the following equation:

$$p_j^t = AGG_{VPI}\left(m_i^{ped\_veh,t}, vp_{ij}^t\right), i \in \{1, \dots, N\}, j \in \{1, \dots, M\} \qquad (3.5)$$

p_j^t is then aggregated with the motion state of the individual vehicle and fed to the SLSTM network.

3.3.3 Trajectory Encoding

LSTMs have been used widely to capture the motion state of pedestrians [41], [45], [77], [78], [130]. We build on this prior work. The precise predicting of forthcoming trajectories based solely on past trajectories poses a formidable challenge, primarily due to the inherent uncertainty that accompanies future trajectories, even when past trajectories are indistinguishable. To surmount this challenge, supplementary information cues, such as pedestrian intention, vehicle speed, and global scene dynamics, play a critical role in advancing the accuracy of future trajectory prediction, as these cues exhibit strong correlations with pedestrian trajectory prediction. To further augment the intrinsic interactions among these cues, we propose the integration of an additional memory cell and dynamic rescaling of the output gate in response to changes in vehicle-pedestrian spatial interaction. One LSTM is used for each pedestrian and each vehicle. The architecture of the proposed Long Short-Term Memory (LSTM) and a conventional LSTM are compared in Figure 3.3.



(a) LSTM (b) Our Proposed LSTM

Figure 3.3 (a). The structure of a standard LSTM neuron. (b) The structure of our proposed LSTM.

3.3.3.1 Pedestrian Trajectory Encoding

The implementation has two steps as follows:

1. We first calculate each pedestrian's relative position and pose to the previous time step.

$$\Delta x_i^t = x_i^t - x_i^{t-1} \quad (3.6)$$

$$\Delta y_i^t = y_i^t - y_i^{t-1} \quad (3.7)$$

For the relative pose:

$$\Delta \theta_i^t = \theta_i^t - \theta_i^{t-1} \quad (3.8)$$

2. The calculated relative positions and pose are then embedded into a fixed-length vector e_i^t for every time step, which is called the spatial feature of the pedestrian.

$$e_i^{ped,t} = \phi\,(\Delta x_i^t\,, \Delta y_i^t\,, \Delta\,\theta_i^t\,;\, W_{e^{ped}}) \quad (3.9)$$

$\phi(.)$ is an embedding function and W_e is the embedding weight. This vector $e_i^{ped,t}$ is then aggregated with the vehicle-pedestrian interaction feature v_i^t from equation 3.4 and then fed to the SLSTM network.

$$m_i^{ped,t} = SLSTM\left(m_i^{t-1}, e_i^t, v_i^t\,;\, W_{m^{ped}}\right) \quad (3.10)$$

$m_i^{ped,t}$ is the hidden state of the SLSTM at time step t; $W_{m^{ped}}$ is the weight of the SLSTM cell.

3.3.3.2 Vehicle Trajectory Encoding

The methodology for encoding vehicle trajectories is identical to the pedestrian section. The following two steps are followed:

1. We first calculate each vehicle's relative position and pose to the previous time step.

$$\Delta x_j^t = x_j^t - x_j^{t-1} \qquad (3.11)$$

$$\Delta y_j^t = y_j^t - y_j^{t-1} \qquad (3.12)$$

For the relative pose:

$$\Delta \theta_j^t = \theta_j^t - \theta_j^{t-1} \qquad (3.13)$$

2. The calculated relative positions and pose are then embedded into a fixed-length vector $e_j^{veh,t}$ for every time step, which is called the spatial feature of the vehicle.

$$e_j^{veh,t} = \phi\ (\Delta x_j^t\ , \Delta y_j^t\ , \Delta\ \theta_j^t\ ;\ W_{e^{veh}}) \qquad (3.14)$$

$\phi(.)$ is an embedding function and $W_{e^{veh}}$ is the embedding weight. This vector $e_i^{veh,t}$ is then aggregated with the vehicle-pedestrian interaction feature p_j^t from equation 11 and then fed to the SLSTM network.

$$m_j^{veh,t} = SLSTM\left(m_j^{t-1}\ , e_j^{veh,t}, p_j^t\ ;\ W_{m^{veh}}\right) \quad (3.15)$$

$m_j^{veh,t}$ is the hidden state of the SLSTM at time step t; $W_{m^{veh}}$ is the weight of the SLSTM cell.

3.3.4 Interaction Modeling and Prediction

Employing one LSTM with the VPI feature extraction model for each pedestrian and vehicle trajectory fails to capture the intricate and temporal interactions between humans and vehicles. To address this shortcoming and enable more information sharing across different pedestrians and vehicles in crowded environments, we propose treating pedestrians and vehicles as nodes on a directed graph and utilizing the recent advances in Graph Neural Networks (GNNs). By assigning varying levels of importance to different nodes, Graph Attention Network (GAT) models enable us to aggregate information from neighbors. Thus, we adopt GAT as the sharing mechanism in our approach. As demonstrated in Figure 3.4, pedestrians and vehicles are represented as nodes in the graph, and GAT serves as the sharing mechanism.

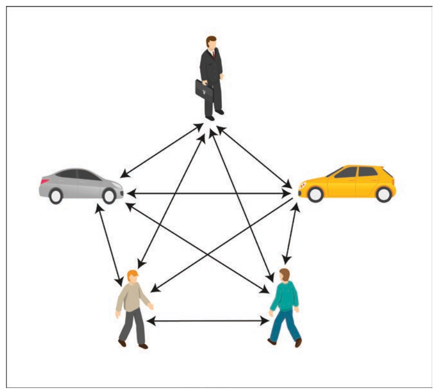


Figure 3.4 Interaction as a directed graph. Pedestrians and vehicles are nodes. The edges are the interaction between these objects.

Graph Attention Network (GAT) is designed to process graph-structured data and compute node features by attending to the features of their neighboring nodes based on a self-attention mechanism [131]. Multiple graph attention layers can be stacked to form the complete GAT model [131]. A single graph attention layer is illustrated in Figure 3.5.

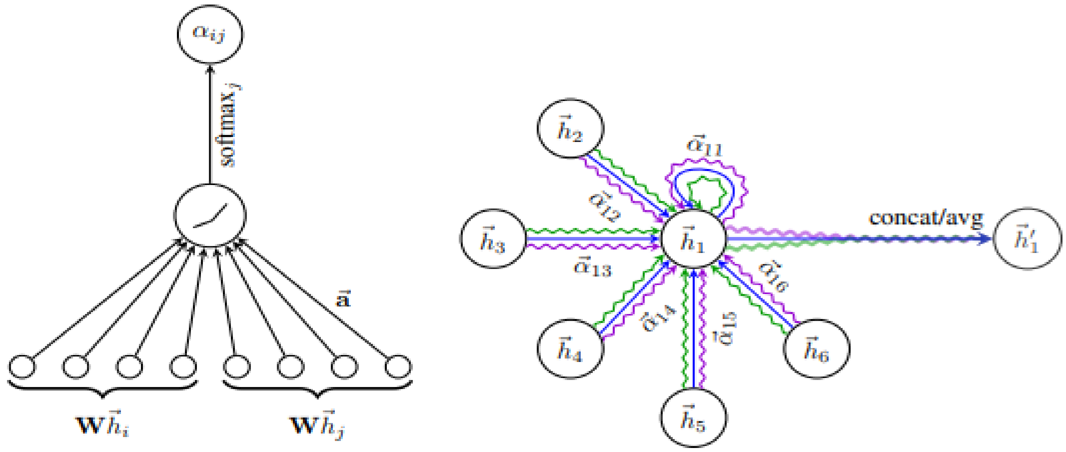


Figure 3.5 Graph Attention Network [131].

The input of the graph attention layer is $h = \{\overrightarrow{h_1}, \overrightarrow{h_2}, \dots, \overrightarrow{h_N}\}$ where $\overrightarrow{h_i} \in R^F$, N is the number of nodes, and F is the feature dimension of each node. The output is $\overrightarrow{h'} = \{\overrightarrow{h_1'}, \overrightarrow{h_2'}, \dots, \overrightarrow{h_N'}\}$ where $\overrightarrow{h_i'} \in R^{F'}$. $F'$and F can be unequal.

In the observation period of $m_i^{ped,t}$ where $t = 1, \dots, T_{obs}$ is fed to the graph attention layer. The coefficients in the attention mechanism of the node pair (i, j) can be computed by:

$$\alpha_{ij}^t = \frac{\exp\left(LeakyReLU(a^T[W\, m_i^{ped,t} \,\|\, W\, m_j^{veh,t}\,]\right)}{\sum_{k \in \mathcal{N}} \exp\left(LeakyReLU\,(a^T[W\, m_i^{ped,t} \,\|\, W\, m_j^{veh,t}\,]\right)} \tag{3.16}$$

where ‖ is the concatenation operation, $\{\cdot\}^T$represents transposition, α_{ij}^t is the attention coefficient of node j to i at timestep t, $\mathcal{N}$ represents the neighbors of node i on the graph. $W \in R^{F' \times F}$ is the weight matrix of a shared linear transformation that is applied to each node (F is the dimension of $m_i^{ped,t}$, F' is the dimension of output), and $a \in R^{2F'}$is the weight vector of a single-layer feedforward neural network. It is normalized by a softmax function with LeakyReLU.

After getting the normalized attention coefficients, the output of one graph attention layer for node i at t is given by:

$$\widehat{m}_i^{ped,t} = \sigma\left(\sum\nolimits_{j \in \mathcal{N}} \alpha_{ij}^t \boldsymbol{W} m_j^{veh,t}\right) \qquad (3.17)$$

where σ is a nonlinear function. Eq. 3.16 and Eq. 3.17 show how a single graph attention layer works. In our implementation, two graph attention layers are adopted. $\widehat{m}_i^t$ (the result after two graph attention layers) is the aggregated hidden state for pedestrian i at t, which contains the spatial influence of other pedestrians and vehicles.

To capture the temporal correlations between interactions, another LSTM called TLSTM is used as shown below:

$$g_i^{ped,t} = TLSTM\left(g_i^{ped,t-1}, \widehat{m}_i^{ped,t}, W_{g^{ped}}\right) \qquad (3.18)$$

$$g_j^{veh,t} = TLSTM\left(g_j^{veh,t-1}, \widehat{m}_j^{veh,t}, W_{g^{veh}}\right) \qquad (3.19)$$

where $\widehat{m}_i^{ped,t}$ is from Eq. 3.17. $W_{g^{ped}}$ is the TLSTM weight and is shared among all the sequences. In our proposed method, SLSTM is used to model the motion pattern of each pedestrian and each vehicle in the scene. Moreover, another LSTM called TLSTM is used to model the temporal correlations of interactions. These two LSTMs are part of the encoder structure. Then, these two LSTMs are utilized to fuse the spatial and temporal data.

At time-step T_{obs}, there are two hidden variables ($m_i^{ped,T_{obs}}$, $g_i^{ped,T_{obs}}$) from two LSTMs of each pedestrian. In our implementation, these two variables are fed to two different multilayer perceptrons ($\delta 1(\cdot)$ and $\delta 2(\cdot)$) before getting concatenated:

$$\overline{m}_i^{ped} = \delta_1\left(m_i^{T_{obs}}\right) \qquad (3.20)$$

$$\bar{g}_i^{ped} = \delta_2\left(g_i^{T_{obs}}\right) \qquad (3.21)$$

$$h^{ped} = \overline{m}_i^{ped} \parallel \bar{g}_i^{ped} \qquad (3.22)$$

Furthermore, at each time step T_{obs}, there are also two hidden variables ($m_j^{veh,T_{obs}}$, $g_j^{veh,T_{obs}}$) for each vehicle. Then, these two variables are fed to two different perceptrons ($\delta 1(\cdot)$ and $\delta 2(\cdot)$) before getting concatenated:

$$\bar{m}_j^{veh} = \delta_1\left(m_j^{veh,T_{obs}}\right) \quad (3.23)$$

$$\bar{g}_j^{veh} = \delta_2\left(g_j^{veh,T_{obs}}\right) \quad (3.24)$$

$$h^{veh} = \bar{m}_j^{veh} \parallel \bar{g}_j^{veh} \quad (3.25)$$

Using real-world data, our goal is to mimic pedestrians' and vehicles' motions and the interaction between them. Three parts are representing the intermediate state vector of our model namely, hidden variables of SLSTM, hidden variables of TLSTM, and the noise added (as shown in Fig. 3.1). The intermediate state vector is calculated as:

$$d_i^{ped,T_{obs}} = h_i^{ped} \parallel z \quad (3.26)$$

$$d_j^{veh,T_{obs}} = h_j^{veh} \parallel z \quad (3.27)$$

where z represents noise, $h_i^{ped}$and h_j^{veh} are from Eq. 3.22 and Eq. 3.25. The intermediate state vectors $d_i^{ped,T_{obs}}$and $d_j^{veh,T_{obs}}$ then act as the initial hidden state of the decoder LSTM (termed as DLSTM). The pedestrian and vehicle predicted relative positions are shown below:

$$d_i^{ped,T_{obs}+1} = DLSTM\left(d_i^{ped,T_{obs}}, e_i^{ped,T_{obs}}; W_{d^{ped}}\right) \quad (3.28)$$

$$d_j^{veh,T_{obs}+1} = DLSTM\left(d_j^{veh,T_{obs}}, e_j^{veh,T_{obs}}; W_{d^{veh}}\right) \quad (3.29)$$

$$\left(\Delta x_i^{ped,T_{obs}+1}, \Delta y_i^{ped,T_{obs}+1}, \Delta\theta_i^{ped,T_{obs}}\right) = \delta_3\left(d_i^{ped,T_{obs}}\right) \quad (3.30)$$

$$\left(\Delta x_j^{veh,T_{obs}+1}, \Delta y_j^{veh,T_{obs}+1}, \Delta\theta_j^{veh,T_{obs}}\right) = \delta_3\left(d_j^{veh,T_{obs}}\right) \quad (3.31)$$

where W_d is DLSTM weight, $\delta_3(\cdot)$ is a linear layer, $e_i^{ped,T_{obs}}$ is from Eq. 3.9. After getting the predicted relative position at time-step $T_{obs}+1$, the subsequent inputs of DLSTM are calculated based on the last predicted relative position according to Eq. 3.9. And it's easy to convert relative positions to absolute positions for computing loss. The variety loss from [42] works as follows: for each pedestrian and each vehicle, the model produces multiple predicted trajectories by randomly sampling z from $\mathcal{N}(0,1)$ (the standard normal distribution). Then it chooses the trajectory that has the smallest distance to ground truth as the model output to compute the loss:

$$L_{variety}^{ped} = \min_{k^{ped}} \left\| Y_i - \hat{Y}_i^{k^{ped}} \right\|^2 \qquad (3.32)$$

$$L_{variety}^{veh} = \min_{k^{veh}} \left\| Y_j - \hat{Y}_j^{k^{veh}} \right\|^2 \qquad (3.33)$$

In Eq. 3.32, Y_i is the ground-truth trajectory of pedestrian i, $\hat{Y}_i^{k^{ped}}$ is the trajectory produced by our model, and k^{ped} is a hyperparameter. By considering only the best trajectory, this loss encourages the network to cover the space of outputs that conform to the past trajectory. In Eq. 3.33, Y_j is the ground-truth trajectory of vehicle j, $\hat{Y}_j^{k^{veh}}$ is the trajectory produced by our model, and k^{veh} is a hyperparameter.

3.4 Implementation Details

In our implementation, each LSTM consists of only one layer. In Eqs. 3.9 and 3.14, the dimension of $e_i^{ped,t}$ and $e_j^{veh,t}$ are set to 256 and the dimension of $m_i^{ped,t}$ and $m_j^{veh,t}$ in Eqs. 3.10 and 3.15 are set to 64. The weight matrix W (Eq. 3.16) for the first graph attention layer has a shape of 32×32, while for the second layer, it has a shape of 32 × 64. The dimension of the attention coefficient matrix a in Eq. 3.16 is set to 32 for the first graph attention layer and 64 for the second layer. Batch normalization is applied to the input of the graph attention layer. In Eqs. 3.18 and 3.19, the dimension of $g_i^{ped,t}$ and $g_j^{veh,t}$ is set to 32. The activation function $\delta_1(.)$ (Eqs. 3.20 and 3.23) contains three layers with ReLU activation functions. The number of hidden nodes in these layers is 32, 64, and 24, respectively. Similarly, the activation function δ_2 (.) (Eqs. 3.21 and 3.24) contains three layers with ReLU activation functions, and the number of hidden nodes is 32, 64, and 16, respectively. The dimension of z in Eq. 3.26 and Eq. 3.27 is set to 16. We trained the network using the Adam optimizer with a learning rate of 0.01 and a batch size of 64.

3.5 Experiments

3.5.1 Dataset

Datasets play a crucial role in developing and assessing deep learning models. For example, researchers frequently employ the widely-used ETH [51] and UCY [52] datasets to evaluate the efficacy of pedestrian trajectory prediction. However, these datasets are not specifically designed for urban traffic scenarios. We employed the VCI-DUT [53] and inD datasets [128] to overcome this limitation to train and evaluate our proposed HSTGA model. These datasets contain large numbers of real-world vehicle-pedestrian trajectories, encompassing various human-human, human-vehicle, and vehicle-vehicle interactions. Additionally, we compared our model against state-of-the-art pedestrian trajectory prediction models on several pedestrian datasets, including ETH, UCY, and Stanford Drone Dataset (SDD) [129].

The VCI-DUT dataset comprises real-world pedestrian and vehicle trajectories collected from two locations on China's Dalian University of Technology (DUT) campus, as depicted in Figure 3.6. The first location consists of a pedestrian crosswalk at an intersection without traffic signals, where the right-of-way is not prioritized for either pedestrians or vehicles. The second location is a relatively large shared space near a roundabout where pedestrians and vehicles have free movement. Similar to the CITR dataset, the recordings were captured using a DJI Mavic Pro Drone equipped with a downward-facing camera, which was positioned high enough to go unnoticed by pedestrians and vehicles. The footage has a resolution of 1920×1080 with a frame rate of 23.98 fps. The dataset primarily comprises trajectories of college students leaving their classrooms and regular cars passing through the campus. The dataset comprises 17 clips of crosswalk scenarios and 11 clips of shared space scenarios, including 1793 trajectories. Some of the clips involve multiple VCIs, i.e., more than two vehicles simultaneously interacting with pedestrians, as illustrated in Figure 3.6.

Figure 3.6 VCI-DUT Dataset with trajectories of vehicles (red dashed line) and pedestrians (colorful solid lines). Upper: Intersection. Lower: Roundabout [53].

The second dataset utilized in this study is the inD dataset, as depicted in Figure 3.7. This new dataset contains naturalistic vehicle trajectories captured at intersections in Germany. Traditional data collection methods are prone to limitations such as occlusions; however, by using a drone, these obstacles are overcome. Traffic at four distinct locations was recorded, and the trajectory for each road user was extracted, along with their corresponding type. State-of-the-art computer vision algorithms were used to obtain positional errors, typically less than 10 centimeters. The inD dataset is applicable to numerous tasks, including road user prediction, driver modeling, scenario-based safety validation of automated driving systems, and data-driven development of highly automated driving (HAD) system components.

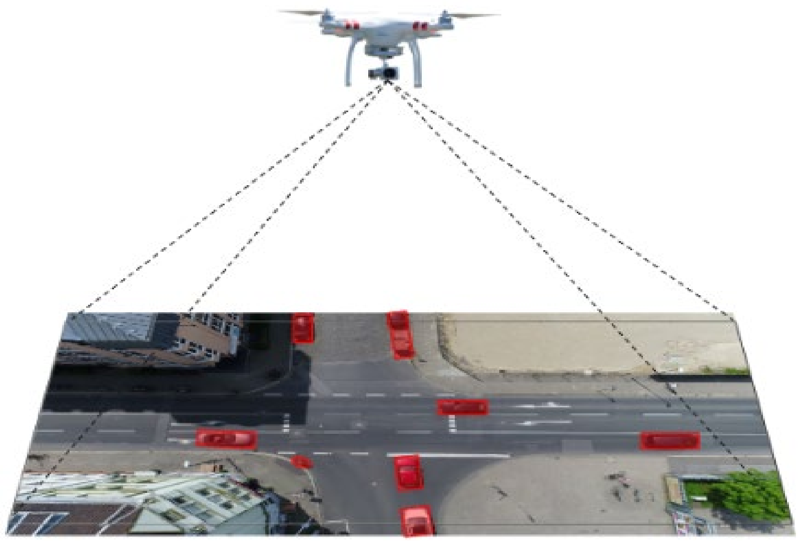


Figure 3.7 inD dataset [128].

3.5.2 Evaluation Metrics

Following the same as prior works [41], [45], [77], [78], [130], we use two error metrics to report prediction errors:

- *Average Displacement Error (ADE):* The mean distance between the actual and predicted trajectories over all predicted time steps, as specified by Equation 3.34.
- *Final Displacement Error (FDE):* The mean distance between the actual and predicted trajectories at the last predicted time-step, which is expressed in Equation 3.35.

$$ADE_{ped} = \frac{\sum_{i\in N}\sum_{t=T_{obs}+1}^{T_f}\left|\left|Y_i^{ped,t} - \hat{Y}_i^{ped,t}\right|\right|^2}{N \times (T_f - T_{obs})} \tag{3.34}$$

$$FDE_{ped} = \frac{\sum_{i\in N}\left|\left|Y_i^{ped,t} - \hat{Y}_i^{ped,t}\right|\right|^2}{N}, t = T_f \tag{3.35}$$

In Eqs. 3.34 and 3.35, N is the number of pedestrians. To find the ADE and FDE for vehicles, N will be replaced by M, which is the number of vehicles.

3.6 Results and Analysis

3.6.1 Quantitative Results

Our model is trained and evaluated on the VCI-DUT, and inD datasets. Moreover, we evaluated our model on additional datasets such as ETH, UCY, HOTEL, ZARA1, and ZARA2. The ADE and FDE results (in meters) for 12 time-step predictions are shown in Table 3.1; lower results are better. The bold font represents the best results. The proposed model outperforms the previous approaches Social-LSTM [41], Social Attention [133], Social-GAN [132], CIDNN [78], STGAT [77], and Step Attention [61] on both ADE and FDE. The proposed model outperforms all previous approaches as shown in Table 3.1. The results demonstrate that the use of human-human, human-vehicle, and vehicle-vehicle information improves the accuracy of pedestrian trajectory prediction.

Table 3.1 Quantitative results of all the baselines and our model. Two evaluation metrics namely, ADE and FDE are presented (lower results are the best).

Metric	Dataset	LSTM	S-LSTM [41]	SocialAttention [133]	CIDNN [78]	SGAN [132]	STGAT [77]	HSTGA (Ours)
ADE	ETH	0.70/1.09	0.73/1.09	1.04/1.39	0.89/1.25	0.60/0.87	0.56/0.65	**0.42/0.53**
ADE	HOTEL	0.55 / 0.86	0.49 / 0.79	1.95 / 2.51	1.25 / 1.31	0.48 / 0.72	0.27 / 0.35	**0.22/0.31**
ADE	UNIV	0.36 / 0.61	0.41 / 0.67	0.78 / 1.25	0.59 / 0.90	0.36 / 0.60	0.31 / 0.51	**0.27/0.44**
ADE	ZARA1	0.25 / 0.41	0.27 / 0.47	0.59 / 1.01	0.29 / 0.50	0.21 / 0.34	0.21 / 0.34	**0.19/0.31**
ADE	ZARA2	0.31 / 0.52	0.33 / 0.56	0.55 / 0.88	0.28 / 0.51	0.27 / 0.42	0.20 / 0.29	**0.20/0.27**
FDE	ETH	1.45 / 2.41	1.48 / 2.35	1.83 / 2.39	1.89 / 2.32	1.19 / 1.62	1.10 / 1.12	**0.96/1.03**
FDE	HOTEL	1.17 / 1.91	1.01 / 1.76	2.97 / 2.91	2.20 / 2.36	0.95 / 1.61	0.50 / 0.66	**0.44/0.52**
FDE	UNIV	0.77 / 1.31	0.84 / 1.40	1.56 / 2.54	1.13 / 1.86	0.75 / 1.26	0.66 / 1.10	**0.55/0.98**
FDE	ZARA1	0.53 / 0.88	0.56 / 1.00	1.24 / 2.17	0.59 / 1.04	0.42 / 0.69	0.42 / 0.69	**0.41/0.62**
FDE	ZARA2	0.65 / 1.11	0.70 / 1.17	1.09 / 1.75	0.60 / 1.07	0.54 / 0.84	0.40 / 0.60	**0.38/0.61**

Table 3.2 presents a comparative analysis of the factors that influence pedestrian trajectory in LSTM-based models and our proposed method. We investigate the influence of the social interaction (SI), the pedestrian-vehicle interaction (VPI), and different inputs including the relative position (RP), the relative velocity (RV), and learning the vehicle-pedestrian interaction adaptively (LIA).

Table 3.2 Interaction and Influencing Factors of LSTM-based Models.

Model Name	Dataset	ADE	FDE	Influencing Factors						
LSTM	ETH	0.70/1.09	1.45/2.41	-	-	-	-	-	-	-
S-LSTM [41]	ETH	0.73/1.09	1.48/2.35	SI	-	RP	-	-	-	-
SocialAttention [133]	ETH	1.04/1.39	1.83 / 2.39	SI	-	RP	-	-	-	-
CIDNN[78]	ETH	0.89/1.25	1.89/2.32	SI	-	RP	-	-	-	-
SGAN [132]	ETH	0.60/0.87	1.19 / 1.62	SI	-	RP	RV	-	-	-
STGAT [77]	ETH	0.56/0.65	1.10 / 1.12	SI	TI	RP	RV	-	-	-
HSTGA (Ours)	ETH	**0.42/0.53**	**0.96/1.03**	SI	TI	RP	RV	LIA	-	HA

*SI: spatial interaction, TI: temporal interaction, RP: relative position, RV: relative velocity, VPI: vehicle-pedestrian interaction, VVI: vehicle-vehicle interaction, HA: heading angle, LIA: learning vehicle-pedestrian interaction adaptively using LSTM.

In Table 3.3, we demonstrate the evaluation outcomes of our method on the VCI-DUT and inD datasets, and compare them with baseline techniques, including the state-of-the-art DNN-based pedestrian prediction methods.

• Constant Velocity (CV) [134]: The pedestrian is assumed to travel at a constant velocity.

• Social GAN (SGAN) [132]: A GAN architecture using a permutation invariant pooling module to capture pedestrian interactions at different scales.

• Multi-Agent Tensor Fusion (MATF) [76]: A GAN architecture using a global pooling layer to combine trajectory and semantic information.

•Off the Sidewalk Predictions (OSP) [134]: The probabilistic interaction model introduced in [330].

Table 3.3 Quantitative results on DUT and inD datasets.

Metric	Dataset	CV [134]	SGAN [132]	MATF-S [76]	OSP [134]	**HSTGA (our)**
ADE	DUT	0.39	0.62	1.65	0.22	**0.11**
FDE	DUT	0.38	0.66	1.87	0.30	**0.16**
ADE	inD	0.50	0.98	1.01	0.42	**0.23**
FDE	inD	0.50	1.09	1.12	0.50	**0.29**

As shown in Table 3.3, the proposed method HSTGA outperforms previous works in both the shared spaces of DUT and un-signalized intersections of inD.

3.6.2 Qualitative Results

To qualitatively analyze the performance of the HSTGA model, we plot the predicted trajectories against the ground truth. The following scenarios show the qualitative results of our model where pedestrians interact with vehicles in a very challenging environment.

Scenario 1:

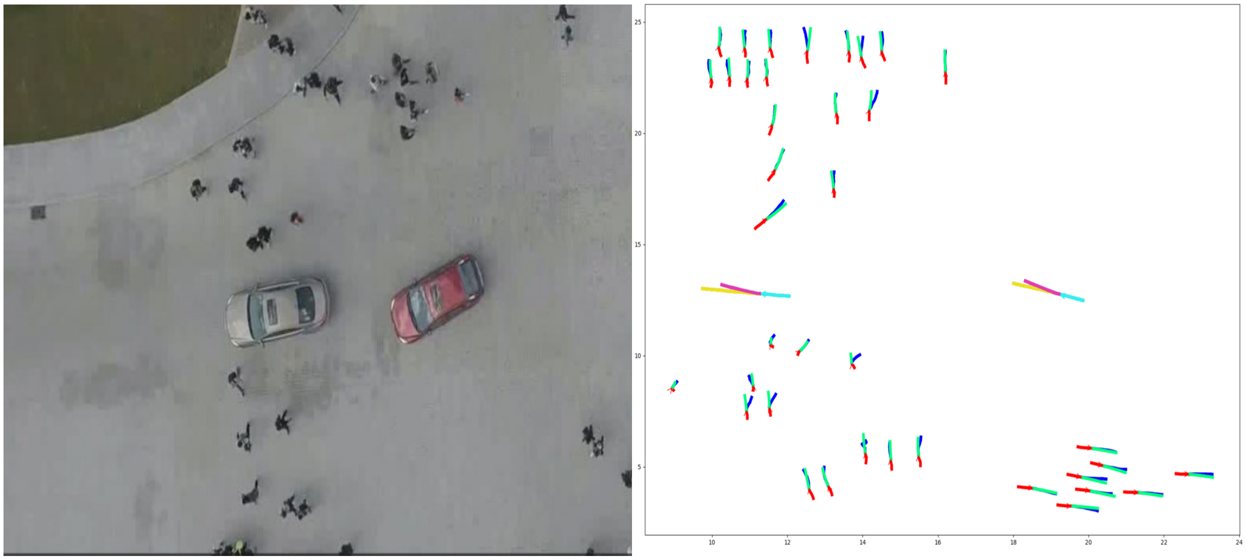


Figure 3.8 The output trajectories of the model on the DUT scenario. Left: Visual of the scene; Right: Trajectory model and prediction – Pedestrians: Red (observed trajectory), blue (ground truth), green (predicted trajectory). Vehicles: Turquoise (observed trajectory), yellow (ground truth), and pink (predicted trajectory).

Scenario 2:

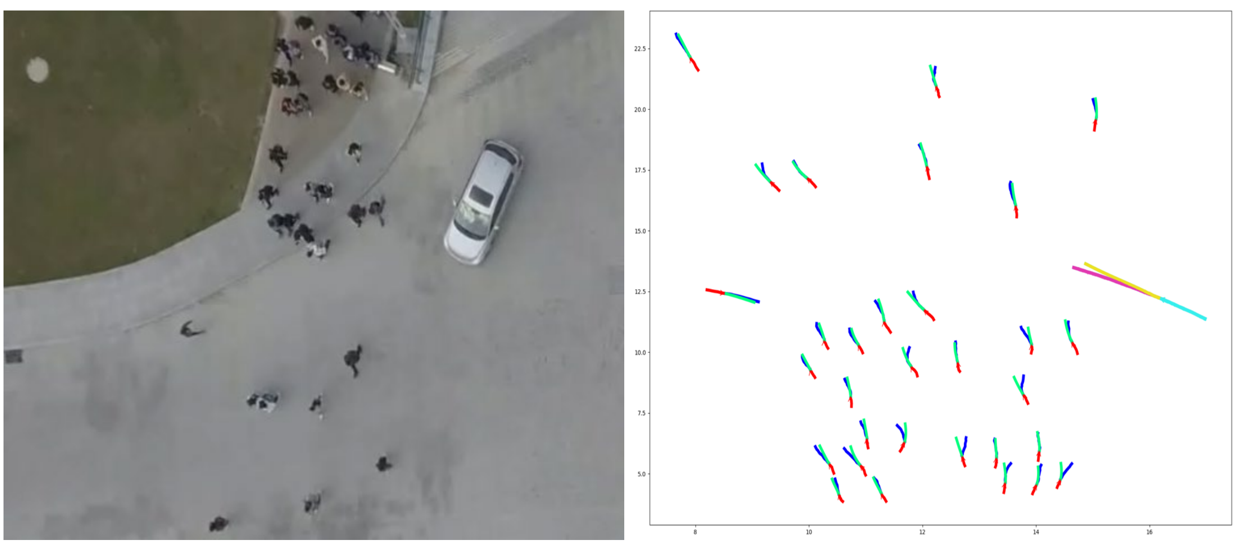


Figure 3.9 The output trajectories of the model on the DUT scenario. Left: Visual of the scene; Right: Trajectory model and prediction – Pedestrians: Red (observed trajectory), blue (ground truth), green (predicted trajectory). Vehicles: Turquoise (observed trajectory), yellow (ground truth), and pink (predicted trajectory).

Scenario 3:

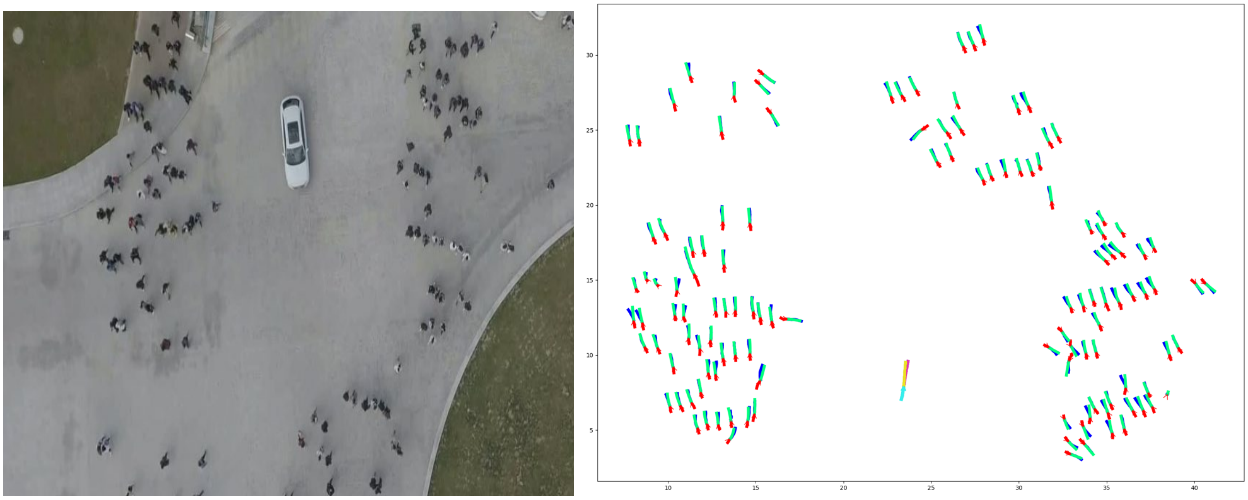


Figure 3.10 The output trajectories of the model on the DUT scenario. Left: Visual of the scene; Right: Trajectory model and prediction – Pedestrians: Red (observed trajectory), blue (ground truth), green (predicted trajectory). Vehicles: Turquoise (observed trajectory), yellow (ground truth), and pink (predicted trajectory).

Scenario 4:

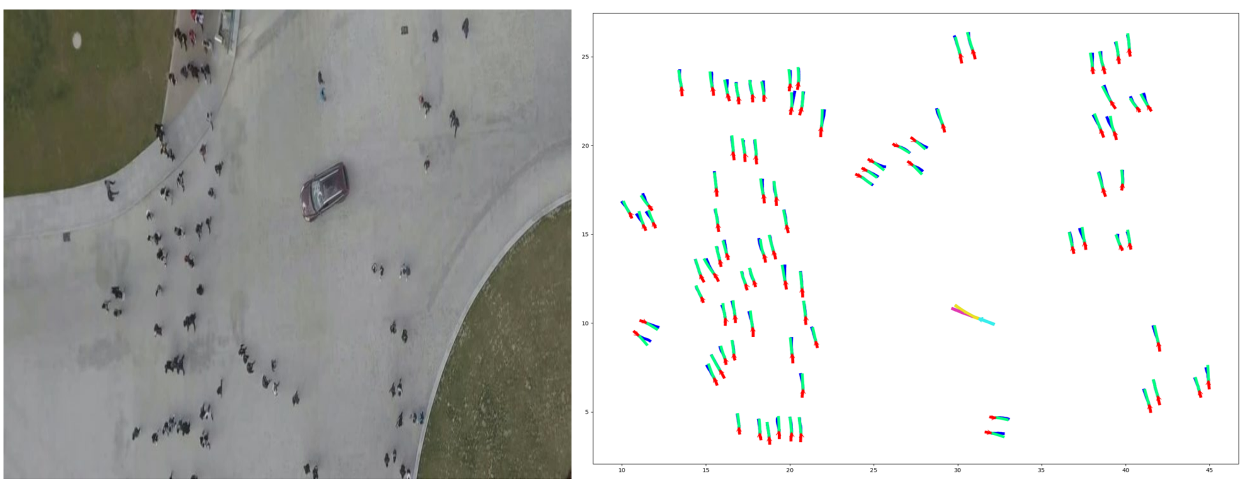


Figure 3.11 The output trajectories of the model on the DUT scenario. Left: Visual of the scene; Right: Trajectory model and prediction – Pedestrians: Red (observed trajectory), blue (ground truth), green (predicted trajectory). Vehicles: Turquoise (observed trajectory), yellow (ground truth), and pink (predicted trajectory).

Scenario 5:

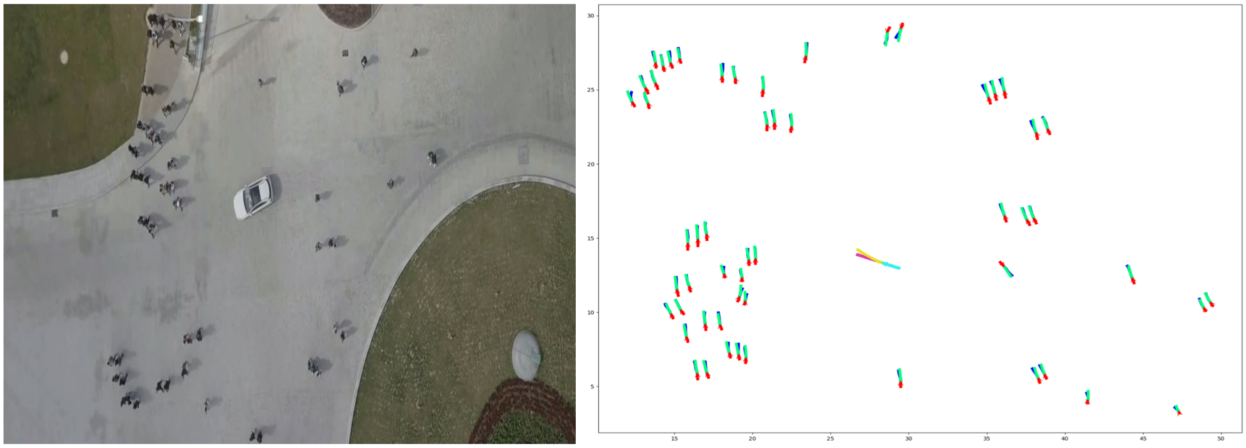


Figure 3.12 The output trajectories of the model on the DUT scenario. Left: Visual of the scene; Right: Trajectory model and prediction – Pedestrians: Red (observed trajectory), blue (ground truth), green (predicted trajectory). Vehicles: Turquoise (observed trajectory), yellow (ground truth), and pink (predicted trajectory).

Our investigation involved a rigorous evaluation of the predictive capacities of our model, which entailed the projection of future outcomes across a range of distinct time steps. Specifically, we examined the predictive accuracy at time steps 8, 12, 14, 16, 18, 20, 22, and 24 steps ahead. These chosen time steps were critical in assessing the model's efficacy in forecasting future events. To illustrate our findings, we present a series of figures that offer a comprehensive depiction of the obtained results for each designated time step. Importantly, the data presented in these figures pertain specifically to the 5th scenario, ensuring a focused and contextually relevant analysis.

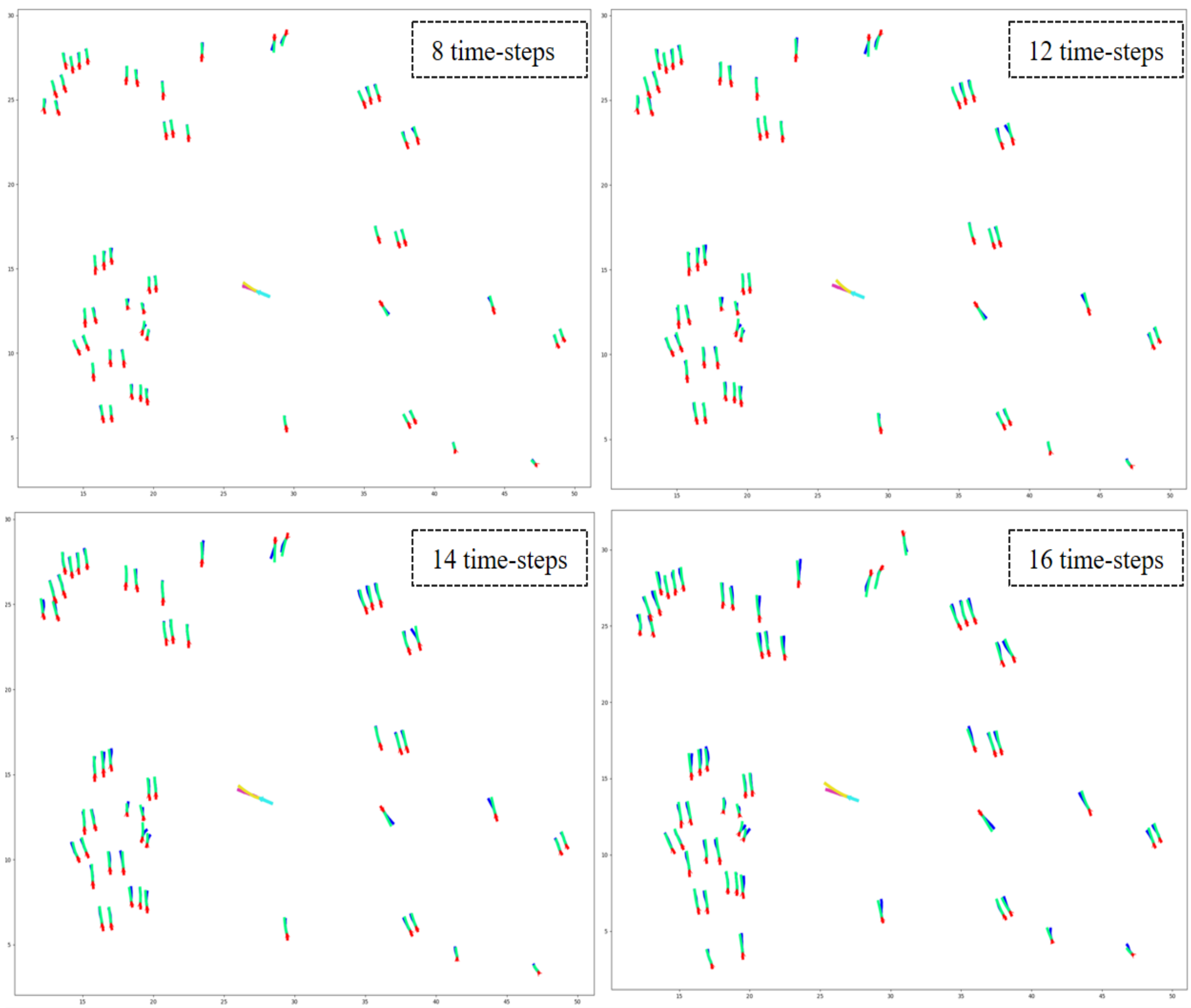


Figure 3.13 Predicted trajectories at 8, 12, 14, and 16-time steps.

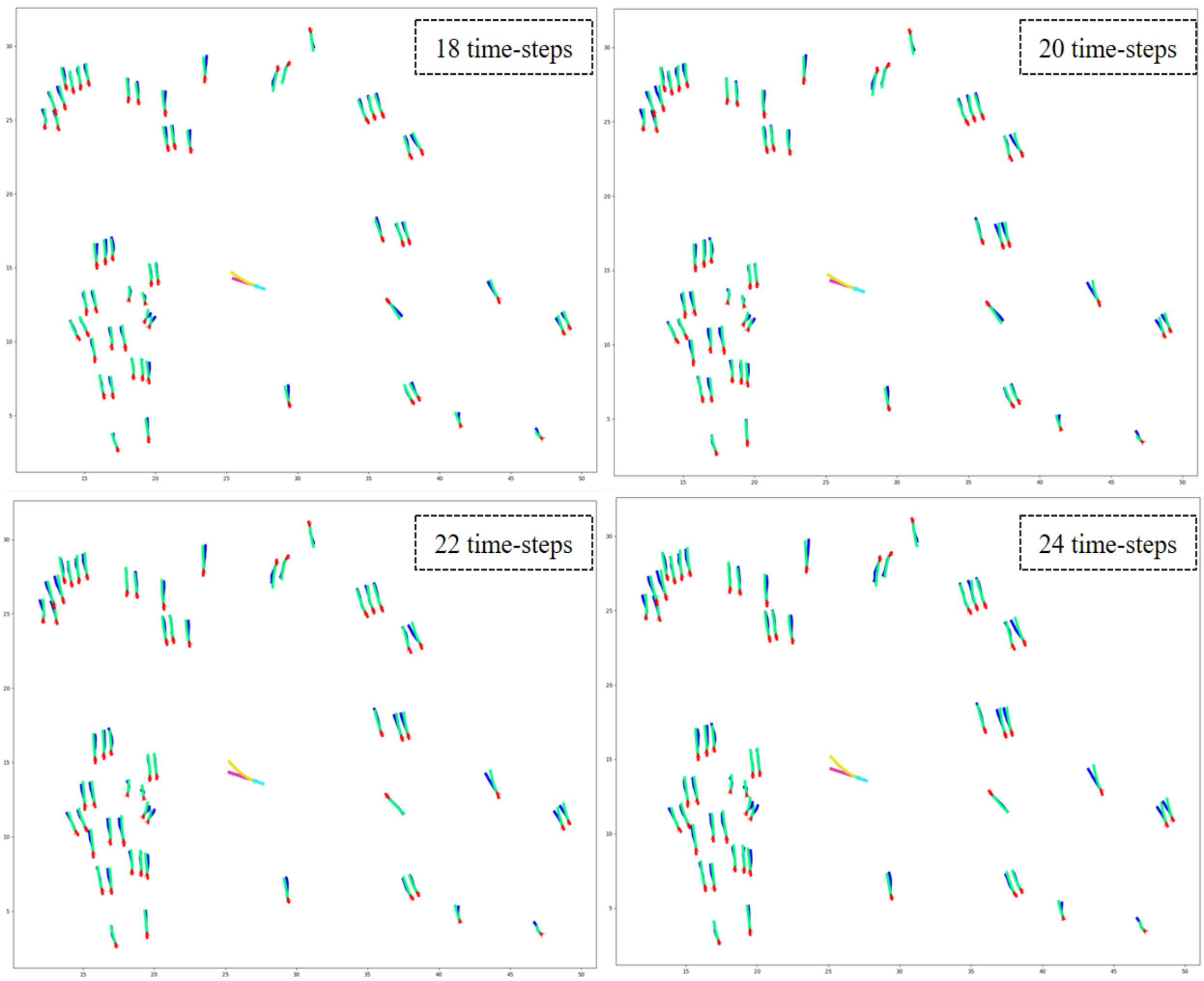


Figure 3.14 Predicted trajectories at 18, 20, 22, and 24-time steps.

Using the output trajectories, we used the predicted trajectories to see if the pedestrian and vehicle would have a critical interaction. We investigate this in Chapter 4.

3.7 Summary

In this study, we propose a novel encoder-decoder interaction model named Holistic Spatio-Temporal Graph Attention Trajectory Prediction for Vehicle-Pedestrian Interaction (HSTGA). HSTGA aims to predict long-horizon pedestrian and vehicle trajectories by modeling the pedestrian-vehicle interactions in non-signalized and non-crosswalk scenarios. The proposed model uses a trajectory-based approach to capture the complex interactions between pedestrians and vehicles. HSTGA integrates a holistic spatiotemporal graph attention mechanism that learns the attention weights of spatial and temporal features of pedestrians and vehicles. The proposed method outperforms state-of-the-art pedestrian trajectory prediction models on various benchmark datasets, highlighting the effectiveness of the HSTGA model.

In order to effectively capture the interaction features between pedestrians and vehicles, a vehicle-pedestrian interaction feature extraction model utilizing a multi-layer perceptron (MLP) sub-network and max pooling has been proposed. The MLP sub-network is responsible for extracting the features of both pedestrians and vehicles, while the max pooling operation aggregates these features into a singular vector. The extracted features are then inputted into an LSTM network to predict the trajectories of both pedestrians and vehicles. This feature extraction model enhances the model's ability to capture the intricate interactions between pedestrians and vehicles, resulting in heightened prediction accuracy.

Compared to other methods, the proposed approach reduces both computational and data requirements, rendering it suitable for real-time applications. The MLP sub-network extracts features in parallel, reducing the overall time complexity of the model. The max pooling operation combines the features of pedestrians and vehicles into a single vector, thereby decreasing the number of input parameters required for the LSTM network. Furthermore, the proposed approach solely utilizes the historical trajectories of pedestrians and vehicles, thus eliminating the need for external data sources. Extensive evaluations conducted on diverse datasets containing numerous challenging scenarios involving the interactions between vehicles and pedestrians demonstrate the effectiveness and efficiency of the proposed approach.

Chapter 4

Vehicle-Pedestrian Conflict Model

4.1 Introduction

Transportation has become an essential aspect of our everyday life. As cities continue to grow, interactions between vehicles and pedestrians have become increasingly complex and frequent. To ensure safety and efficiency in the transportation system, it is crucial to have a clear understanding of these interactions and be able to anticipate possible conflicts. Conflict-based models offer an approach to studying these interactions by simulating potential vehicle-pedestrian conflicts in shared spaces such as roads and intersections. These models consider various factors such as natural limitations, traffic rules, and the behavior and choices of vehicles and pedestrians. Identifying high-risk areas and proposing measures to mitigate them can significantly improve transportation safety.

Conflict-based models may be used to assess the performance of various traffic management devices such as roundabouts, crosswalks, and traffic signals in fostering safe and successful vehicle-pedestrian interactions. Urban planning is an important field where conflict-based models are employed. When building a public transportation system, the complex interactions that occur between vehicles and pedestrians in metropolitan contexts must be considered.

The rise of intelligent vehicles has further underscored the importance of conflict-based models. In shared areas, intelligent vehicles must be able to communicate securely with infrastructure and potentially with pedestrians. Developers of intelligent vehicles can use conflict-based models to assess these vehicles during design and deployment in simulation and real-world scenarios. The utilization of conflict-based model simulation presents a distinctive advantage in the evaluation of intelligent vehicles, particularly in its ability to capture and test rare events with enhanced flexibility. Unlike traditional testing methods, which often rely on predetermined scenarios and well-defined parameters, the conflict-based model simulation introduces a dynamic and adaptable framework that can replicate a wide range of complex real-world situations. By incorporating a multitude of potential conflicts, including unexpected behavior from other road users, this simulation approach facilitates the evaluation of intelligent vehicles in scenarios that are infrequent yet crucial for assessing their robustness and safety. The inherent flexibility of the conflict-based model simulation thus ensures a more comprehensive and realistic evaluation of intelligent vehicles' capabilities, enabling researchers and developers to better understand and address rare events that may significantly impact their performance on the road. Additionally, they can create algorithms to ensure the safe operation of these vehicles on public roads.

Vehicle and pedestrian conflict-based models are vital tools in ensuring the safety and effectiveness of the transportation system. Furthermore, the vehicle-pedestrian conflict-based model provides critical information to traffic experts, urban planners, and intelligent vehicle creators. Therefore, conflict-based modeling can improve transit networks' safety and effectiveness for all users, including pedestrians and vehicles.

To create an accurate conflict-based model for vehicle-pedestrian interactions, the movements of vehicles and pedestrians at the same time-step must be input into the conflict model. This data assists in investigating any critical conflicts between vehicles and pedestrians at each time step. The movements at each time step can be represented as vehicle and pedestrian trajectories. Designing the conflict-based model requires accurate vehicle and pedestrian trajectory prediction under highly uncertain environments. Therefore, the conflict-based model has two stages, the trajectory prediction stage, and the conflict-based stage.

4.2 Problem Statement

The problem statement is already discussed in chapter 2 section 2.3.

4.3 Research Methodology

Historically, researchers have utilized road accident data and surrogate safety measures (SSM) to examine and evaluate the safety of diverse road amenities. In this study, a conflict-based approach is employed to investigate the conduct of pedestrian-vehicle conflict through trajectory data. The demeanor of this conflict is contingent on which road user traverses the conflict zone first. If a vehicle passes first, the pedestrian may withdraw, slow down, or stop until the vehicle has passed, or may alter its trajectory and move to the rear of the vehicle. Conversely, the vehicle may hasten and traverse the conflict zone before the pedestrian's arrival. In the event that a pedestrian passes first, the vehicle may decelerate, stop, or alter its path, whereas the pedestrian may accelerate and traverse the conflict zone before the vehicle's arrival.

Following our prior works in [27], and [138]-[142], we propose a conflict-based model to investigate these behaviors and extract valuable data, such as normal and safety-critical driving metrics. The normal and safety-critical metrics such as vehicle speed, pedestrian speed, time-to-collision (TTC), and post-encroachment time (PET) are important indicators to help categorize events into different patterns. Whereas minimum values of TTC and PET were used to get severity levels to evaluate the safety between a vehicle and a pedestrian. Finally, the TTC and PET thresholds are used to design Autonomous Emergency Braking (AEB) algorithm. Fig. 4.1 illustrates the architecture of the proposed model.

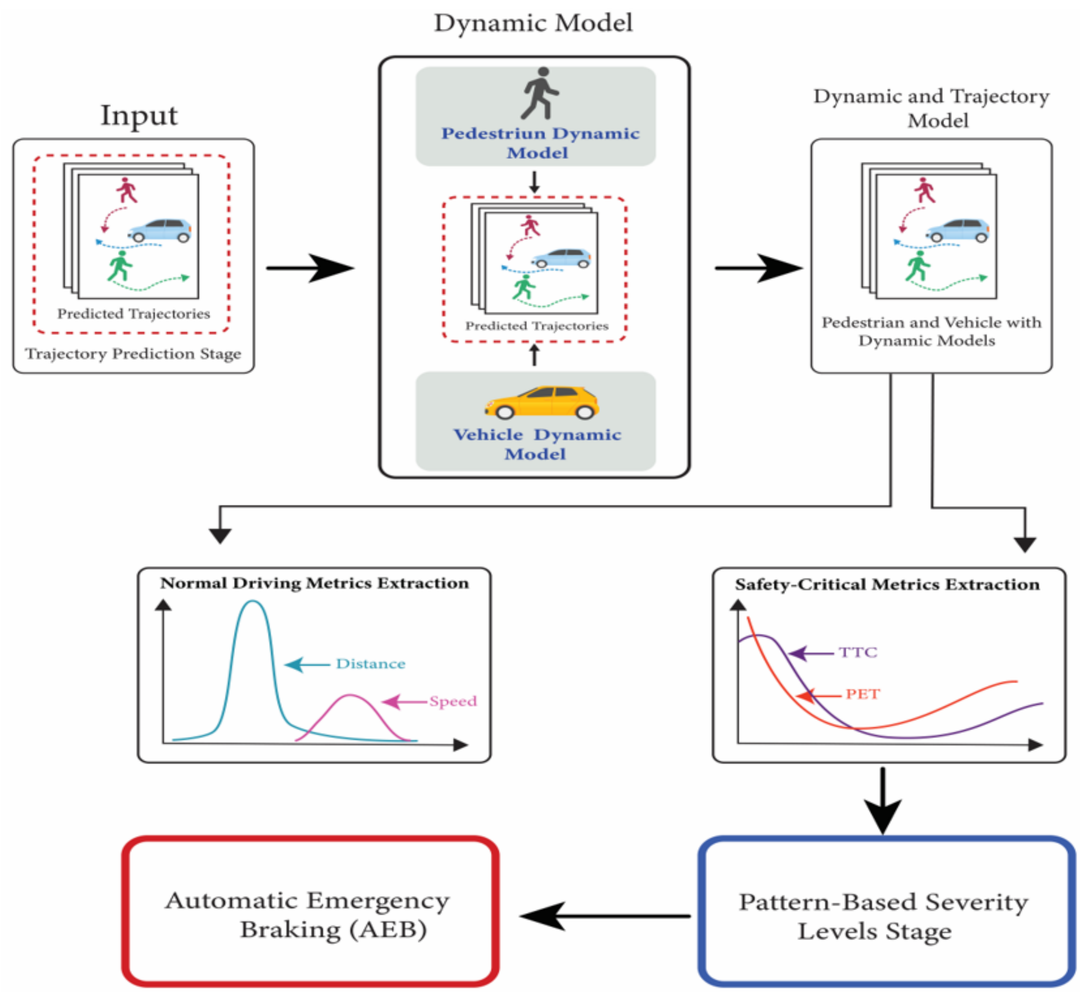


Figure 4.1 Vehicle-pedestrian conflict-based model.

4.4 Data Acquisition

As shown in Figure 4.1, the input to the conflict-based model is the predicted trajectories from our HSTGA deep learning model in Chapter 3. We train and evaluate our HSTGA model using the VCI-DUT [53] and inD datasets [128]. These datasets contain thousands of real-world vehicle-pedestrian trajectories and cover rich human-human, human-vehicle, and vehicle-vehicle interactions. Moreover, we also evaluate our HSTGA model using many pedestrian datasets such as ETH [51], UCY [52], and Stanford Drone Dataset (SDD) [129]. More details on datasets are listed in Chapter 3, section 3.5.1.

4.5 Interactive Pedestrian and Vehicle Motion

4.5.1 Pedestrian Dynamic Model

For pedestrian dynamics, we only use the trajectory data to control the pedestrian dynamics.

4.5.2 Vehicle Dynamic Model

The kinematic Bicycle Model is used to model the vehicle dynamics. Then, the vehicle-predicted trajectory data from our HSTGA model is then integrated into the vehicle dynamics model. Specifically, the x and y coordinates of the vehicle's position at each time step are extracted from the vehicle trajectory data and used to update the vehicle's position. The vehicle speed is also calculated using the trajectory data by computing the distance traveled between consecutive time steps and dividing it by the time step. This speed is then used in the vehicle dynamics calculations to update the vehicle's position and orientation in the next time step.

4.6 Results and Analysis

To assess the accuracy and real-world representation of the model being proposed, we carefully analyze a set of statistical measures. These metrics cover a range of driving behaviors, including both everyday driving measures and those related to safety-critical situations. To provide a clear understanding of each metric, we will introduce their detailed definitions in the upcoming sections.

4.6.1 Safety-Critical Driving Metrics

The primary obstacle to the advancement of intelligent vehicles, particularly autonomous vehicles, lies in the management of safety-critical driving scenarios encountered in actual environments. As such, the conflict-based model necessitates the ability to accurately recognize these infrequent, high-risk occurrences. This segment will evaluate the effectiveness of our conflict-based model in generating safety-critical metrics, exclusively encompassing proximity-based indicators like Time-to-Collision (TTC) and Post-Encroachment Time (PET).

4.6.1.1 Time-to-Collision (TTC) Estimation

The computational methodology detailed in references [135] and [136] has been selected for the purpose of calculating the Time-to-Collision (TTC) between the vehicle and the pedestrian. This approach has been identified as an advanced technique for determining the TTC as compared to the previous methodology proposed in reference [137]. The improved method incorporates the understanding that, during a collision between two objects, one of the objects' corners will make initial contact unless the impact occurs perpendicularly, in which case both corners will make contact simultaneously. Therefore, by evaluating the intersection points that align with the corners of both objects, the first point of intersection can be computed. The intersection point is determined using the subsequent formula:

$$x_{+} = \frac{(y_2 - y_1) - (x_2 \cdot tan\theta_2 - x_1 \cdot tan\theta_1)}{tan\theta_1 - tan\theta_2} \tag{4.1}$$

$$y_{+} = \frac{(x_2 - x_1) - (y_2 \cdot cot\theta_2 - y_1 \cdot cot\theta_1)}{cot\theta_1 - cot\theta_2} \tag{4.2}$$

The variables x_{+} and y_{+} denote the conflicts or intersections coordinates while x_1, y_1, x_2, and y_2 denote the corner coordinates. The variables q_1 and q_2 represent the respective directions of the vehicle and the pedestrian. As shown in Fig. 4.2, we represent the vehicle and pedestrian with a rectangular shape.

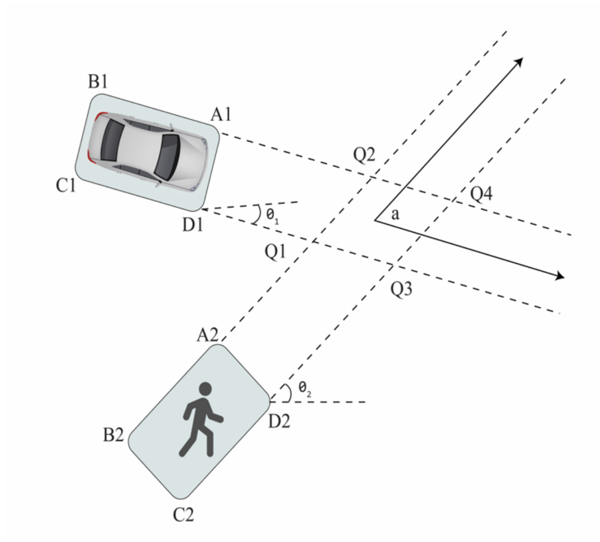


Figure 4.2 Time-to-Collision Estimation.

Using an example conflict between a vehicle and pedestrian from the following scenario (Figure 4.3), we build the vehicle-pedestrian conflict-based simulation model using the predicted trajectories from our HSTGA model of the vehicle in the red circle and the pedestrian in the yellow circle. Moreover, the vehicle dynamic is integrated with the predicted trajectories to mimic the real-world motion of the vehicle. The predicted trajectories for 16-time steps are used to generate the future conflict between the vehicle and the pedestrian.

Figure 4.3 Vehicle-pedestrian conflict scenario. The red circle is the vehicle and the yellow circle is the pedestrian.

The following figure shows the vehicle-pedestrian conflict–based simulation that is created as discussed in section 4.5.

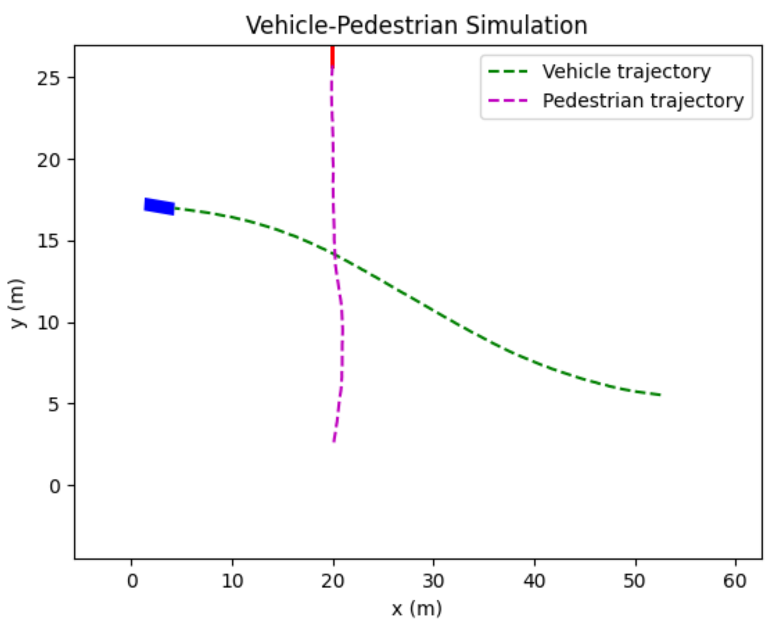


Figure 4.4 Vehicle-pedestrian conflict-based simulation.

The time-to-collision plot (Fig. 4.5) is generated using the above technique (Fig. 4.2) and using the vehicle-pedestrian simulation shown in Fig. 4.4.

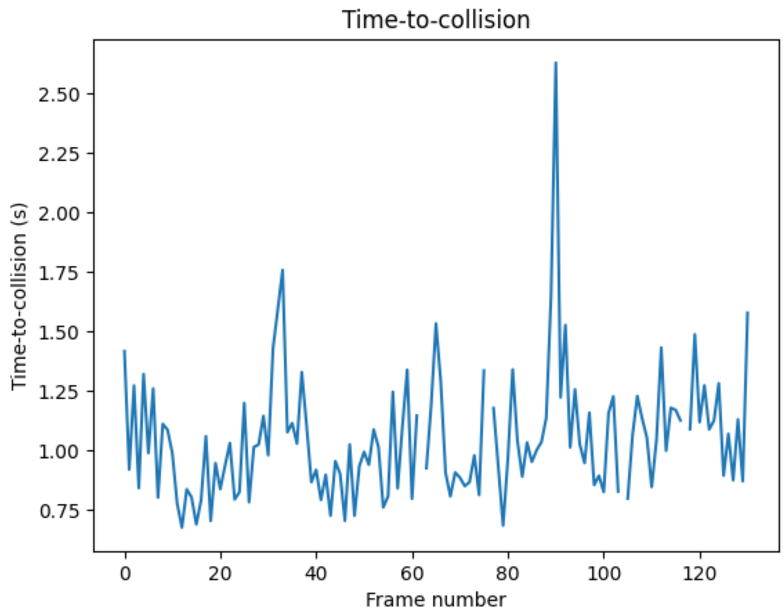


Figure 4.5 Time-to-Collision.

Below is the histogram depicting the time-to-collision for the vehicle and pedestrian illustrated in Figure 4.3. The histogram reveals their close proximity, indicating a potential future collision between the two entities.

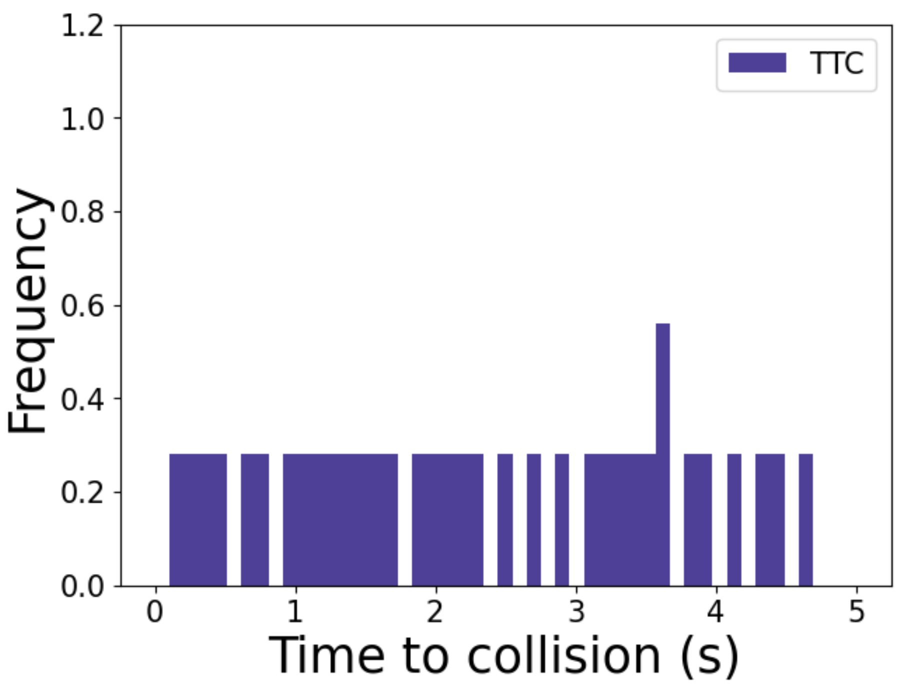


Figure 4.6 Time-to-collision histogram.

4.6.1.2 Post-Encroachment Time (PET) Estimation

Using the following technique as shown in Figure 4.7, we plot the PET histogram between the vehicle and pedestrian in Figure 4.8.

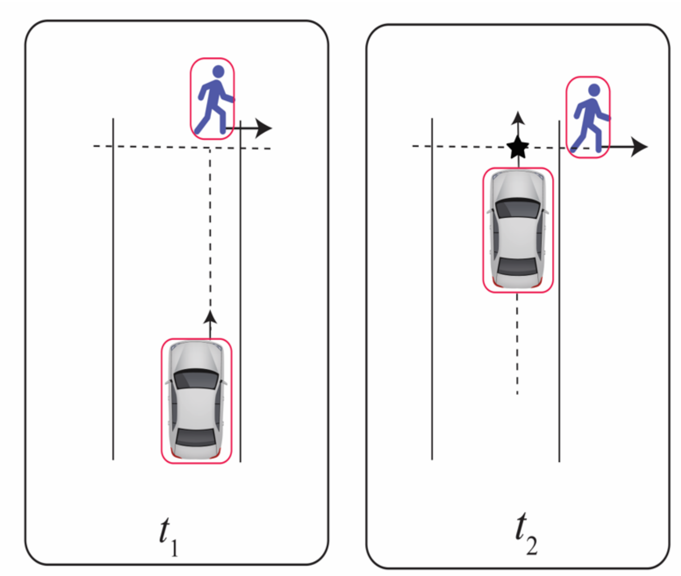


Figure 4.7 PET Estimation.

The PET histogram is presented in Figure 4.8. The greater the proximity and the shorter the PET (Post-Encroachment Time), the higher the level of danger in the situation.

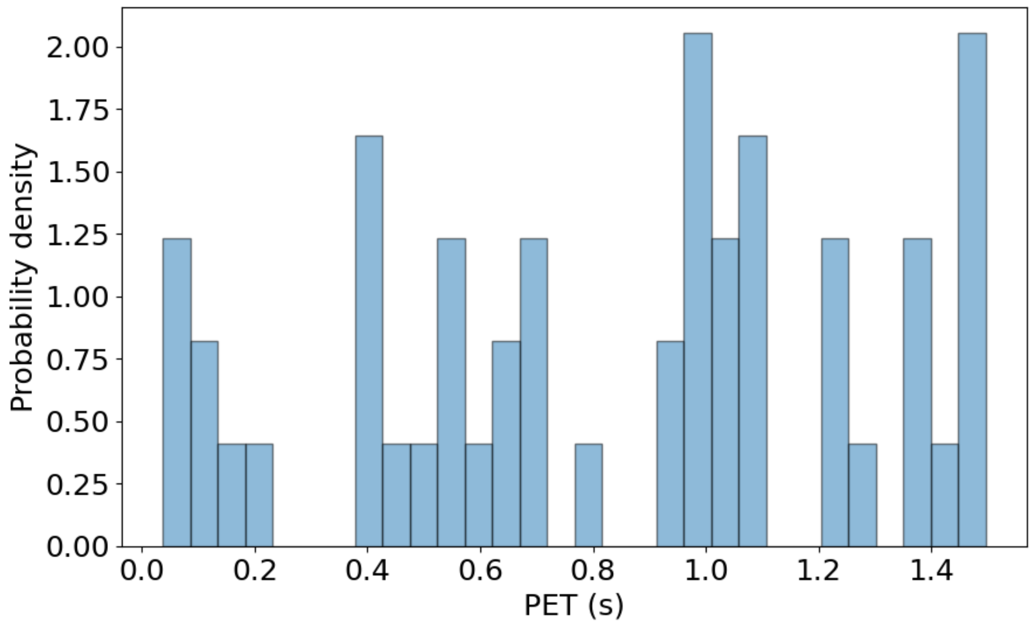


Figure 4.8 PET Histogram.

4.6.2 Normal Driving Metrics

During the simulation of the conflict-based model, various standard driving metrics are recorded, including the speeds of the vehicle and pedestrian, the accelerations of both entities, their respective heading angles, and the distance between them. These metrics provide crucial information regarding the dynamics of the simulated scenario as shown in the following figures. Due to the complex nature of the scenario we set out to investigate, the vehicle and pedestrian change their speeds and accelerations rapidly and frequently. This dynamic is primarily attributed to their close proximity to each other, demanding swift adjustments to their respective speeds and rates of acceleration.

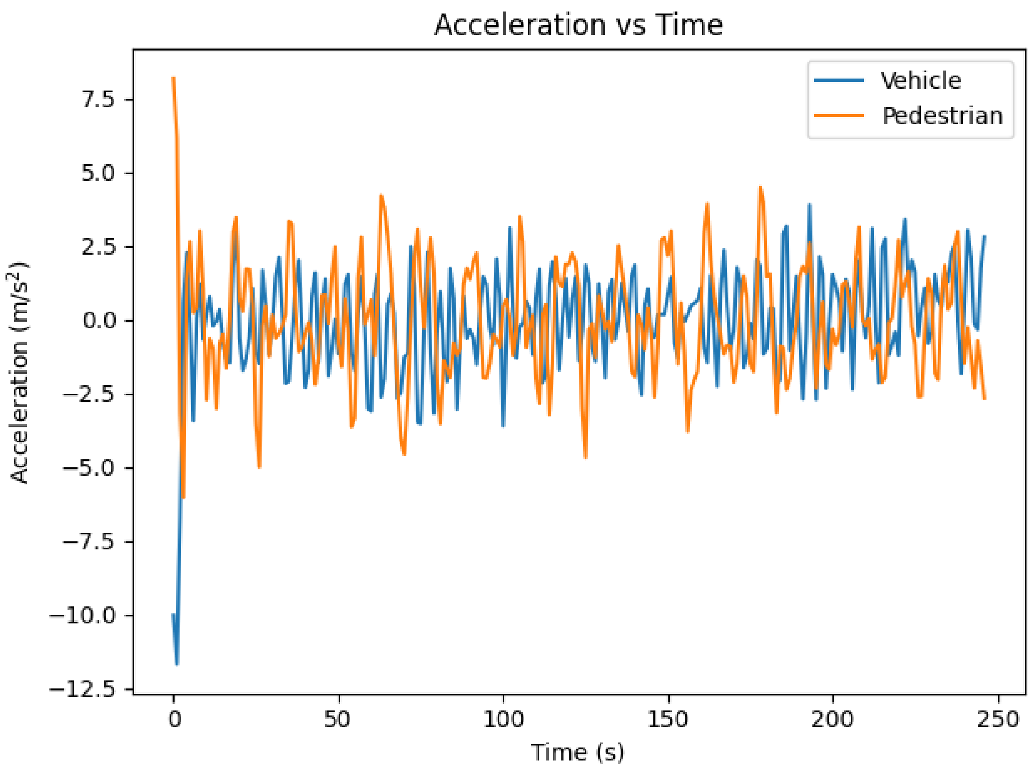


Figure 4.9 Acceleration Profile.

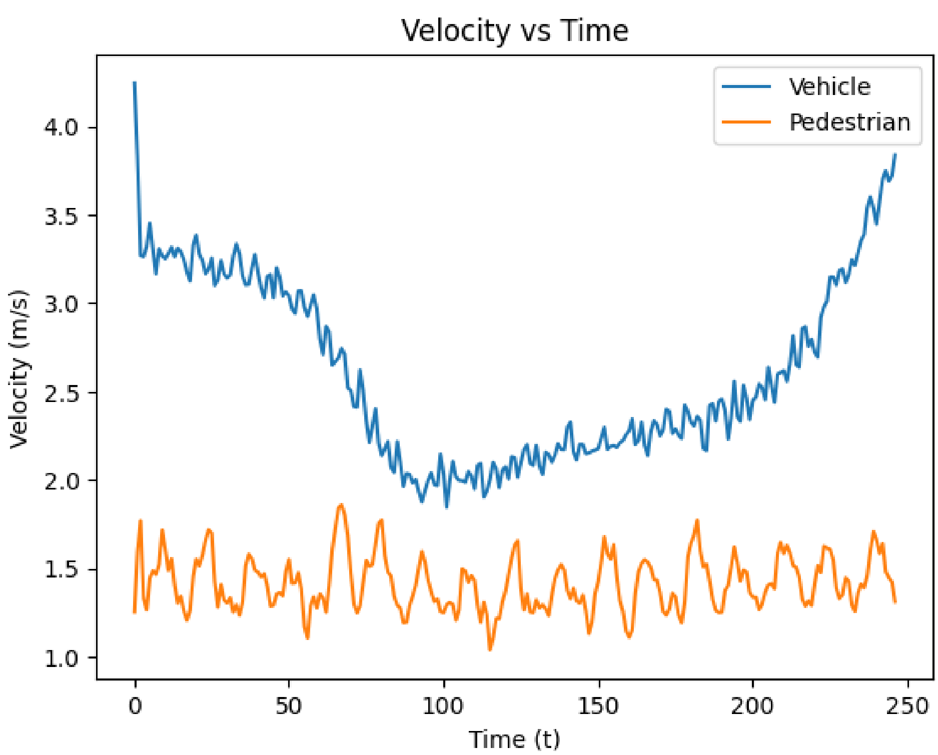


Figure 4.10 Velocity Profile.

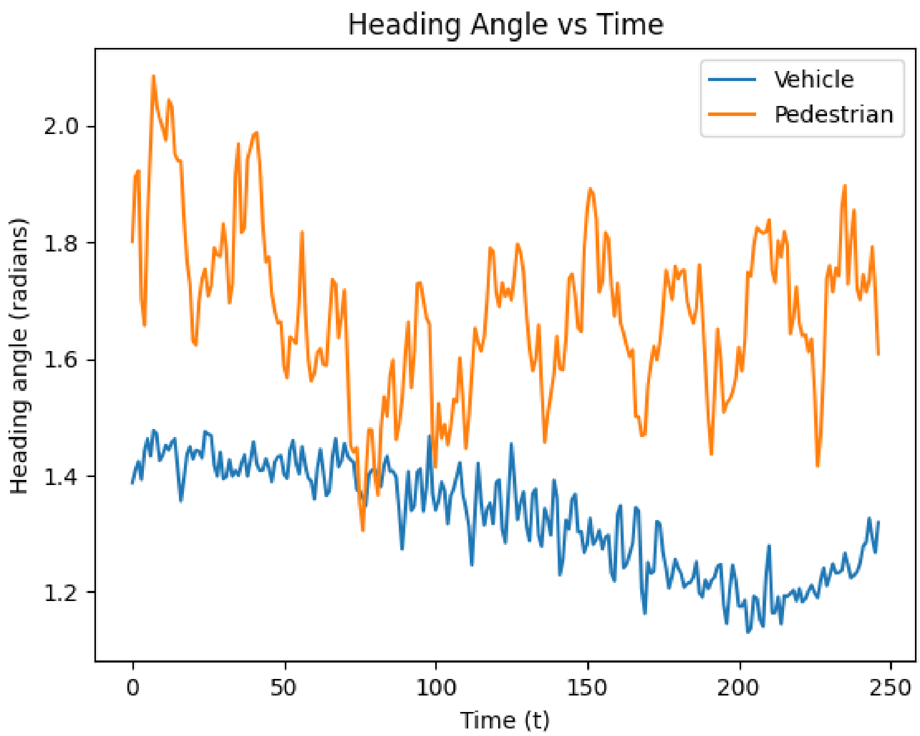


Figure 4.11 Heading Angle Profile.

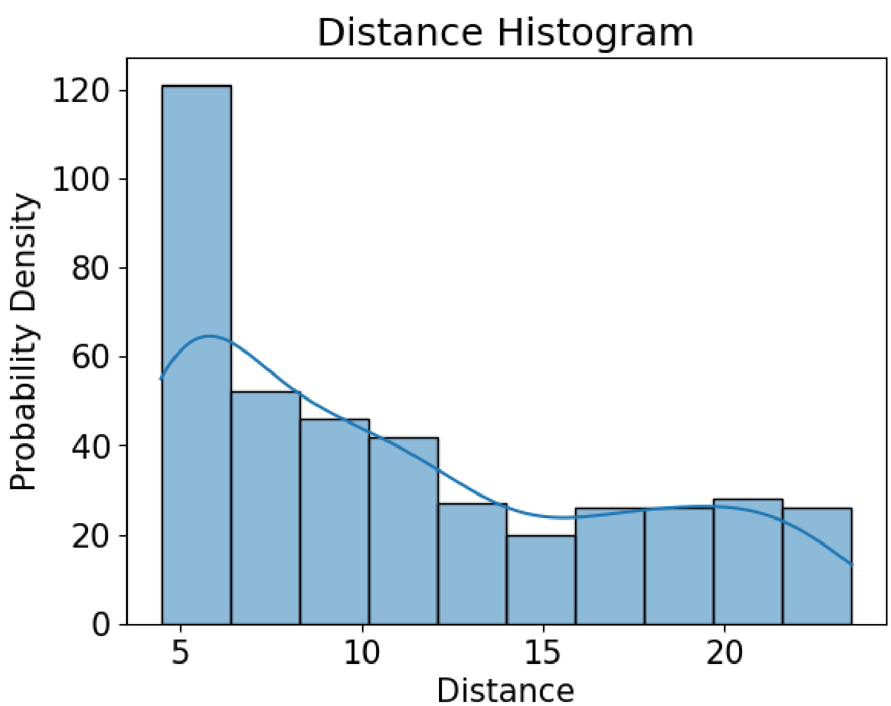


Figure 4.12 Distance histogram.

4.7 Summary

We propose a conflict-based model to study driving behaviors and extract valuable data, including normal and safety-critical driving metrics. These metrics include vehicle speed, pedestrian speed, time-to-collision (TTC), and post-encroachment time (PET), which serve as essential indicators to categorize events into various patterns. The minimum values of TTC and PET are used to evaluate the severity of the interaction between the vehicle and the pedestrian. The proposed model aims to enhance the safety of the driving system by designing an Autonomous Emergency Braking (AEB) algorithm based on TTC and PET thresholds. Our future works will focus more on testing and improving the AEB algorithm.

By utilizing the conflict-based model, our approach offers a comprehensive understanding of the interaction between vehicles and pedestrians. The extracted metrics provide valuable insights into driving behaviors and aid in identifying normal and critical scenarios. The proposed approach's TTC and PET thresholds ensure the safety of pedestrians and drivers and reduce the likelihood of accidents. The AEB algorithm designed based on these metrics can assist in the development of advanced driving assistance systems, enhancing the overall driving experience. The proposed model's architecture is highly adaptable, making it suitable for real-time applications, such as autonomous driving systems, where prompt action is critical.

Chapter 5

Conclusion and Future Work

This research presents a novel encoder-decoder interaction model, named Holistic Spatio-Temporal Graph Attention Trajectory Prediction for Vehicle-Pedestrian Interaction (HSTGA), which aims to predict long-horizon pedestrian and vehicle trajectories in non-signalized and non-crosswalk scenarios. The proposed model utilizes a trajectory-based approach to capture the complex interactions between pedestrians and vehicles, integrating a holistic spatiotemporal graph attention mechanism that learns the attention weights of spatial and temporal features of pedestrians and vehicles. The HSTGA model outperforms state-of-the-art pedestrian trajectory prediction models on various benchmark datasets, demonstrating its effectiveness.

To effectively capture the interaction features between pedestrians and vehicles, a vehicle-pedestrian interaction feature extraction model has been proposed. This model employs a multi-layer perceptron (MLP) sub-network and max pooling to extract features of both pedestrians and vehicles, which are then aggregated into a singular vector and inputted into an LSTM network to predict the trajectories of both pedestrians and vehicles. This feature extraction model enhances the model's ability to capture the intricate interactions between pedestrians and vehicles, resulting in heightened prediction accuracy.

Compared to other methods, the proposed approach reduces both computational and data requirements, making it suitable for real-time applications. The MLP sub-network extracts features in parallel, reducing the overall time complexity of the model. The max pooling operation combines the features of pedestrians and vehicles into a single vector, thereby decreasing the number of input parameters required for the LSTM network. Furthermore, the proposed approach solely utilizes the historical trajectories of pedestrians and vehicles, thus eliminating the need for external data sources. Extensive evaluations conducted on diverse datasets containing numerous challenging scenarios involving the interactions between vehicles and pedestrians demonstrate the effectiveness and efficiency of the proposed approach.

In the past, road accident data and surrogated safety measures (SSM) have been utilized by researchers to evaluate the safety of various road amenities. In this study, we employ a conflict-based approach to investigate pedestrian-vehicle conflicts through trajectory data. The behavior of these conflicts depends on which road user traverses the conflict zone first. If a vehicle passes first, the pedestrian may withdraw, slow down, or stop until the vehicle has passed, or may alter its trajectory and move to the rear of the vehicle. Conversely, the vehicle may hasten and traverse the conflict zone before the pedestrian's arrival. If a pedestrian passes first, the vehicle may decelerate, stop, or alter its path, whereas the pedestrian may accelerate and traverse the conflict zone before the vehicle's arrival.

We propose a conflict-based model to investigate these behaviors and extract valuable data, such as normal and safety-critical driving metrics. These metrics, including vehicle speed, pedestrian speed, time-to-collision (TTC), and post-encroachment time (PET), are important indicators used to categorize events into different patterns. The minimum values of TTC and PET are used to determine severity levels to evaluate the safety between a vehicle and a pedestrian. Finally, the TTC and PET thresholds are used to design an Autonomous Emergency Braking (AEB) algorithm. The utilization of TTC and PET values proved to be valuable indicators in detecting critical conflicts. The AEB algorithm successfully managed to govern the ego vehicle with the aid of TTC values, resulting in a favorable outcome.

Our future work will involve improving and implementing advanced models for predicting pedestrian and vehicle trajectories in non-signalized and non-crosswalk scenarios. Our future work will also involve improving the conflict-based models to investigate pedestrian-vehicle conflicts and extract valuable data to evaluate the safety between a vehicle and a pedestrian. Additionally, we will improve the design and the implementation of the Autonomous Emergency Braking (AEB) algorithms to govern the ego vehicle and prevent critical conflicts, making them suitable for real-time applications.

Bibliography

[1]. “Pedestrian Safety | NHTSA.” https://www.nhtsa.gov/road-safety/pedestrian-safety (accessed Apr. 06, 2023).

[2] “Global status report on road safety 2018.” https://www.who.int/publications-detail-redirect/9789241565684 (accessed Apr. 08, 2023).

[3] “Global Status Report on Road Safety - Time for Action,” WHO | Regional Office for Africa, Apr. 07, 2023. https://www.afro.who.int/publications/global-status-report-road-safety-time-action (accessed Apr. 08, 2023).

[4] “Pedestrian Safety Campaign.” http://txdot.gov/en/home/safety/traffic-safety-campaigns/pedestrian-safety.html (accessed Apr. 16, 2023).

[5] Y. Lu, J. Shen, C. Wang, H. Lu, and J. Xin, “Studying on the design and simulation of collision protection system between vehicle and pedestrian,” International Journal of Distributed Sensor Networks, vol. 16, no. 1, p. 1550147719900109, Jan. 2020, doi: 10.1177/1550147719900109.

[6] J. R. Crandall, K. S. Bhalla, and N. J. Madeley, “Designing road vehicles for pedestrian protection,” BMJ, vol. 324, no. 7346, pp. 1145–1148, May 2002, doi: 10.1136/bmj.324.7346.1145.

[7] G. Stcherbatcheff, C. Tarriere, P. Duclos, A. Fayon, C. Got, and A. Patel, “Simulation of Collisions Between Pedestrians and Vehicles Using Adult and Child Dummies,” SAE International, Warrendale, PA, SAE Technical Paper 751167, Feb. 1975. doi: 10.4271/751167.

[8] A. Ganichev and O. Batishcheva, “Evaluating the conflicts between vehicles and pedestrians,” Transportation Research Procedia, vol. 50, pp. 145–151, Jan. 2020, doi: 10.1016/j.trpro.2020.10.018.

[9] A. Tahmasbi-Sarvestani, H. Nourkhiz Mahjoub, Y. P. Fallah, E. Moradi-Pari, and O. Abuchaar, “Implementation and Evaluation of a Cooperative Vehicle-to-Pedestrian Safety Application,” IEEE Intelligent Transportation Systems Magazine, vol. 9, no. 4, pp. 62–75, 2017, doi: 10.1109/MITS.2017.2743201.

[10] T. Gandhi and M. M. Trivedi, “Pedestrian Protection Systems: Issues, Survey, and Challenges,” IEEE Transactions on Intelligent Transportation Systems, vol. 8, no. 3, pp. 413–430, Sep. 2007, doi: 10.1109/TITS.2007.903444.

[11] R. Ezzati Amini, K. Yang, and C. Antoniou, “Development of a conflict risk evaluation model to assess pedestrian safety in interaction with vehicles,” Accident Analysis & Prevention, vol. 175, p. 106773, Sep. 2022, doi: 10.1016/j.aap.2022.106773.

[12] S. Bai, D. D. Legge, A. Young, S. Bao, and F. Zhou, “Investigating External Interaction Modality and Design Between Automated Vehicles and Pedestrians at Crossings,” in 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), Sep. 2021, pp. 1691–1696. doi: 10.1109/ITSC48978.2021.9564867.

[13] L. Zheng, K. Ismail, and X. Meng, “Traffic conflict techniques for road safety analysis: open questions and some insights,” Can. J. Civ. Eng., vol. 41, no. 7, pp. 633–641, Jul. 2014, doi: 10.1139/cjce-2013-0558.

[14] M. R. Parker, Traffic Conflict Techniques for Safety and Operations: Observers Manual. Federal Highway Administration, 1989.

[15] Amundsen and Hydén, Proceedings of the 1st Workshop on Traffic Conflicts, Oslo, Norway, 1977.

[16] R. Almodfer, S. Xiong, Z. Fang, X. Kong, and S. Zheng, “Quantitative analysis of lane-based pedestrian-vehicle conflict at a non-signalized marked crosswalk,” Transportation Research Part F: Traffic Psychology and Behaviour, vol. 42, pp. 468–478, Oct. 2016, doi: 10.1016/j.trf.2015.07.004.

[17] Y.-C. Liu and Y.-C. Tung, “Risk analysis of pedestrians’ road-crossing decisions: Effects of age, time gap, time of day, and vehicle speed,” Safety Science, vol. 63, pp. 77–82, Mar. 2014, doi: 10.1016/j.ssci.2013.11.002.

[18] D. Yagil, “Beliefs, motives and situational factors related to pedestrians’ self-reported behavior at signal-controlled crossings,” Transportation Research Part F: Traffic Psychology and Behaviour, vol. 3, no. 1, pp. 1–13, Mar. 2000, doi: 10.1016/S1369-8478(00)00004-8.

[19] A. Tom and M.-A. Granié, “Gender differences in pedestrian rule compliance and visual search at signalized and unsignalized crossroads,” Accident Analysis & Prevention, vol. 43, no. 5, pp. 1794–1801, Sep. 2011, doi: 10.1016/j.aap.2011.04.012.

[20] G. Cheng, Y. Wang, and D. Li, “Setting Conditions of Crosswalk Signal on Urban Road Sections in China,” ScholarMate, 2013. https://www.scholarmate.com/A/Evu6ja (accessed Apr. 18, 2023).

[21] V. Himanen and R. Kulmala, “An application of logit models in analysing the behaviour of pedestrians and car drivers on pedestrian crossings,” Accident Analysis & Prevention, vol. 20, no. 3, pp. 187–197, Jun. 1988, doi: 10.1016/0001-4575(88)90003-6.

[22] Troutbeck, R. J., and W. Brilon. "Unsignalized Intersection Theory. in Traffic Flow Theory." Washington, DC: US Federal Highway Administration (1996): 8-1.

[23] W. K. M. Alhajyaseen and H. Nakamura, “Quality of pedestrian flow and crosswalk width at signalized intersections,” IATSS Research, vol. 34, no. 1, pp. 35–41, Jul. 2010, doi: 10.1016/j.iatssr.2010.06.002.

[24] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, “Understanding Pedestrian Behavior in Complex Traffic Scenes,” IEEE Transactions on Intelligent Vehicles, vol. 3, no. 1, pp. 61–70, Mar. 2018, doi: 10.1109/TIV.2017.2788193.

[25] J. Van Brummelen, M. O’Brien, D. Gruyer, and H. Najjaran, “Autonomous vehicle perception: The technology of today and tomorrow,” Transportation Research Part C: Emerging Technologies, vol. 89, pp. 384–406, Apr. 2018, doi: 10.1016/j.trc.2018.02.012.

[26] “J3016_202104: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles - SAE International.” https://www.sae.org/standards/content/j3016_202104/ (accessed Apr. 19, 2023).

[27] H. Alghodhaifi and S. Lakshmanan, “Autonomous Vehicle Evaluation: A Comprehensive Survey on Modeling and Simulation Approaches,” IEEE Access, vol. 9, pp. 151531–151566, 2021, doi: 10.1109/ACCESS.2021.3125620.

[28] V. Melcher, S. Rauh, F. Diederichs, H. Widlroither, and W. Bauer, “Take-Over Requests for Automated Driving,” Procedia Manufacturing, vol. 3, pp. 2867–2873, Jan. 2015, doi: 10.1016/j.promfg.2015.07.788.

[29] M. Kuehn, T. Vogelpohl, and M. Vollrath, “TAKEOVER TIMES IN HIGHLY AUTOMATED DRIVING (LEVEL 3)”.

[30] A. Rasouli and J. K. Tsotsos, “Autonomous Vehicles That Interact With Pedestrians: A Survey of Theory and Practice,” IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 3, pp. 900–918, Mar. 2020, doi: 10.1109/TITS.2019.2901817.

[31] A. Hbaieb, S. Ayed, and L. Chaari, “Internet of Vehicles and Connected Smart Vehicles Communication System Towards Autonomous Driving,” In Review, preprint, Jun. 2021. doi: 10.21203/rs.3.rs-493419/v1.

[32] K. Messaoud, N. Deo, M. M. Trivedi, and F. Nashashibi, “Trajectory Prediction for Autonomous Driving based on Multi-Head Attention with Joint Agent-Map Representation,” in 2021 IEEE Intelligent Vehicles Symposium (IV), Jul. 2021, pp. 165–170. doi: 10.1109/IV48863.2021.9576054.

[33] L. Rossi, A. Ajmar, M. Paolanti, and R. Pierdicca, “Vehicle trajectory prediction and generation using LSTM models and GANs,” PLOS ONE, vol. 16, no. 7, p. e0253868, Jul. 2021, doi: 10.1371/journal.pone.0253868.

[34] L. Xin, P. Wang, C.-Y. Chan, J. Chen, S. E. Li, and B. Cheng, “Intention-aware Long Horizon Trajectory Prediction of Surrounding Vehicles using Dual LSTM Networks,” in 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI: IEEE, Nov. 2018, pp. 1441–1446. doi: 10.1109/ITSC.2018.8569595.

[35] W. Xu et al., “GlobalInsight: An LSTM Based Model for Multi-Vehicle Trajectory Prediction,” in ICC 2020 - 2020 IEEE International Conference on Communications (ICC), Jun. 2020, pp. 1–7. doi: 10.1109/ICC40277.2020.9149261.

[36] T. Tettamanti, I. Varga, and Z. Szalay, “Impacts of Autonomous Cars from a Traffic Engineering Perspective,” Periodica Polytechnica Transportation Engineering, vol. 44, no. 4, Art. no. 4, Oct. 2016, doi: 10.3311/PPtr.9464.

[37] G. Solmaz et al., “Learn from IoT: Pedestrian Detection and Intention Prediction for Autonomous Driving,” in Proceedings of the 1st ACM Workshop on Emerging Smart Technologies and Infrastructures for Smart Mobility and Sustainability, in SMAS ’19. New York, NY, USA: Association for Computing Machinery, Oct. 2019, pp. 27–32. doi: 10.1145/3349622.3355446.

[38] S. Sorooshian, S. Khademi Sharifabad, M. Parsaee, and A. R. Afshari, “Toward a Modern Last-Mile Delivery: Consequences and Obstacles of Intelligent Technology,” Applied System Innovation, vol. 5, no. 4, Art. no. 4, Aug. 2022, doi: 10.3390/asi5040082.

[39] A. Shetty, M. Yu, A. Kurzhanskiy, O. Grembek, H. Tavafoghi, and P. Varaiya, “Safety challenges for autonomous vehicles in the absence of connectivity,” Transportation Research Part C: Emerging Technologies, vol. 128, p. 103133, Jul. 2021, doi: 10.1016/j.trc.2021.103133.

[40] S. Iftikhar, Z. Zhang, M. Asim, A. Muthanna, A. Koucheryavy, and A. A. Abd El-Latif, “Deep Learning-Based Pedestrian Detection in Autonomous Vehicles: Substantial Issues and Challenges,” Electronics, vol. 11, no. 21, Art. no. 21, Jan. 2022, doi: 10.3390/electronics11213551.

[41] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, “Social LSTM: Human Trajectory Prediction in Crowded Spaces,” in 2016 IEEE Conference on Computer Vision

and Pattern Recognition (CVPR), Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 961–971. doi: 10.1109/CVPR.2016.110.

[42] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese, and A. Alahi, "ETH with Generative Adversarial Networks." arXiv, Mar. 28, 2018. Accessed: May 02, 2023. [Online]. Available: http://arxiv.org/abs/1803.10892.

[43] A. Mohamed, K. Qian, M. Elhoseiny, and C. Claudel, "Social-STGCNN: A Social Spatio-Temporal Graph Convolutional Neural Network for Human Trajectory Prediction." arXiv, Mar. 24, 2020. Accessed: May 02, 2023. [Online]. Available: http://arxiv.org/abs/2002.11927.

[44] N. Nikhil and B. T. Morris, "Convolutional Neural Network for Trajectory Prediction." arXiv, Nov. 26, 2018. Accessed: May 02, 2023. [Online]. Available: http://arxiv.org/abs/1809.00696.

[45] C. Zhang and C. Berger, "Learning the Pedestrian-Vehicle Interaction for Pedestrian Trajectory Prediction." arXiv, Apr. 25, 2022. Accessed: May 02, 2023. [Online]. Available: http://arxiv.org/abs/2202.05334.

[46] S. Eiffert, K. Li, M. Shan, S. Worrall, S. Sukkarieh, and E. Nebot, "Probabilistic Crowd GAN: Multimodal Pedestrian Trajectory Prediction using a Graph Vehicle-Pedestrian Attention Network," IEEE Robot. Autom. Lett., vol. 5, no. 4, pp. 5026–5033, Oct. 2020, doi: 10.1109/LRA.2020.3004324.

[47] R. Chandra, U. Bhattacharya, A. Bera, and D. Manocha, "TraPHic: Trajectory Prediction in Dense and Heterogeneous Traffic Using Weighted Interactions," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2019, pp. 8475–8484. doi: 10.1109/CVPR.2019.00868.

[48] R. Chandra, U. Bhattacharya, C. Roncal, A. Bera, and D. Manocha, "RobustTP: End-to-End Trajectory Prediction for Heterogeneous Road-Agents in Dense Traffic with Noisy Sensor Inputs." arXiv, Jul. 20, 2019. Accessed: May 02, 2023. [Online]. Available: http://arxiv.org/abs/1907.08752.

[49] R. Chandra et al., "Forecasting Trajectory and Behavior of Road-Agents Using Spectral Clustering in Graph-LSTMs." arXiv, Aug. 05, 2020. Accessed: May 02, 2023. [Online]. Available: http://arxiv.org/abs/1912.01118.

[50] S. Carrasco, D. F. Llorca, and M. Á. Sotelo, "SCOUT: Socially-COnsistent and UndersTandable Graph Attention Network for Trajectory Prediction of Vehicles and VRUs." arXiv, May 29, 2021. Accessed: May 02, 2023. [Online]. Available: http://arxiv.org/abs/2102.06361.

[51] S. Pellegrini, A. Ess, K. Schindler, and L. van Gool, “You’ll never walk alone: Modeling social behavior for multi-target tracking,” in 2009 IEEE 12th International Conference on Computer Vision, Sep. 2009, pp. 261–268. doi: 10.1109/ICCV.2009.5459260.

[52] A. Lerner, Y. Chrysanthou, and D. Lischinski, “Crowds by Example,” Computer Graphics Forum, vol. 26, no. 3, pp. 655–664, 2007, doi: 10.1111/j.1467-8659.2007.01089.x.

[53] D. Yang, L. Li, K. Redmill, and Ü. Özgüner, “Top-view Trajectories: A Pedestrian Dataset of Vehicle-Crowd Interaction from Controlled Experiments and Crowded Campus.” arXiv, Apr. 19, 2019. Accessed: May 02, 2023. [Online]. Available: http://arxiv.org/abs/1902.00487.

[54] R. Krajewski, T. Moers, J. Bock, L. Vater, and L. Eckstein, “The rounD Dataset: A Drone Dataset of Road User Trajectories at Roundabouts in Germany,” in 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), Sep. 2020, pp. 1–6. doi: 10.1109/ITSC45102.2020.9294728.

[55] J. Bock, L. Vater, R. Krajewski, and T. Moers, “Highly Accurate Scenario and Reference Data for Automated Driving,” ATZ Worldw, vol. 123, no. 5, pp. 50–55, May 2021, doi: 10.1007/s38311-021-0668-8.

[56] C. Hydén, The Development of a Method for Traffic Safety Evaluation: The Swedish Traffic Conflicts Technique. Ch. Hydén, 1987.

[57] A. Várhelyi, “Drivers’ speed behaviour at a zebra crossing: a case study,” Accident Analysis & Prevention, vol. 30, no. 6, pp. 731–743, Nov. 1998, doi: 10.1016/S0001-4575(98)00026-8.

[58] K. Ismail, T. Sayed, and N. Saunier, “Methodologies for Aggregating Indicators of Traffic Conflict,” Transportation Research Record, vol. 2237, no. 1, pp. 10–19, Jan. 2011, doi: 10.3141/2237-02.

[59] S. Malkhamah, M. Tight, and F. Montgomery, “The development of an automatic method of safety monitoring at Pelican crossings,” Accident Analysis & Prevention, vol. 37, no. 5, pp. 938–946, Sep. 2005, doi: 10.1016/j.aap.2005.04.012.

[60] A. Rudenko, L. Palmieri, M. Herman, K. M. Kitani, D. M. Gavrila, and K. O. Arras, “Human Motion Trajectory Prediction: A Survey,” The International Journal of Robotics Research, vol. 39, no. 8, pp. 895–935, Jul. 2020, doi: 10.1177/0278364920917446.

[61] E. Zhang, N. Masoud, M. Bandegi, J. Lull, and R. K. Malhan, “Step Attention: Sequential Pedestrian Trajectory Prediction,” IEEE Sensors Journal, vol. 22, no. 8, pp. 8071–8083, Apr. 2022, doi: 10.1109/JSEN.2022.3158271.

[62] S. Kim et al., “BRVO: Predicting pedestrian trajectories using velocity-space reasoning,” The International Journal of Robotics Research, vol. 34, no. 2, pp. 201–217, Feb. 2015, doi: 10.1177/0278364914555543.

[63] F. Zanlungo, T. Ikeda, and T. Kanda, “Social force model with explicit collision prediction,” EPL, vol. 93, no. 6, p. 68005, Mar. 2011, doi: 10.1209/0295-5075/93/68005.

[64] A. Martinelli, H. Gao, P. D. Groves, and S. Morosi, “Probabilistic Context-Aware Step Length Estimation for Pedestrian Dead Reckoning,” IEEE Sensors Journal, vol. 18, no. 4, pp. 1600–1611, Feb. 2018, doi: 10.1109/JSEN.2017.2776100.

[65] “SmartPDR: Smartphone-Based Pedestrian Dead Reckoning for Indoor Localization | IEEE Journals & Magazine | IEEE Xplore.” https://ieeexplore.ieee.org/document/6987239 (accessed May 05, 2023).

[66] “Indoor Trajectory Prediction Algorithm Based on Communication Analysis of Built-In Sensors in Mobile Terminals | IEEE Journals & Magazine | IEEE Xplore.” https://ieeexplore.ieee.org/document/9350660 (accessed May 05, 2023).

[67] B. D. Ziebart et al., “Planning-based prediction for pedestrians,” in 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, MO, USA: IEEE, Oct. 2009, pp. 3931–3936. doi: 10.1109/IROS.2009.5354147.

[68] A. Galata, N. Johnson, and D. Hogg, “Learning Variable-Length Markov Models of Behavior,” Computer Vision and Image Understanding, vol. 81, no. 3, pp. 398–413, Mar. 2001, doi: 10.1006/cviu.2000.0894.

[69] N. Deo and M. M. Trivedi, “Learning and predicting on-road pedestrian behavior around vehicles,” in 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Oct. 2017, pp. 1–6. doi: 10.1109/ITSC.2017.8317865.

[70] E. Rehder, F. Wirth, M. Lauer, and C. Stiller, “Pedestrian Prediction by Planning using Deep Neural Networks.” arXiv, Jun. 20, 2017. Accessed: May 05, 2023. [Online]. Available: http://arxiv.org/abs/1706.05904.

[71] P. Dendorfer, A. Ošep, and L. Leal-Taixé, “Goal-GAN: Multimodal Trajectory Prediction Based on Goal Position Estimation.” arXiv, Oct. 02, 2020. doi: 10.48550/arXiv.2010.01114.

[72] Y. Yao, E. Atkins, M. Johnson-Roberson, R. Vasudevan, and X. Du, “BiTraP: Bi-directional Pedestrian Trajectory Prediction with Multi-modal Goal Estimation.” arXiv, Nov. 16, 2020. doi: 10.48550/arXiv.2007.14558.

[73] H. Tran, V. Le, and T. Tran, “Goal-driven Long-Term Trajectory Prediction.” arXiv, Nov. 05, 2020. doi: 10.48550/arXiv.2011.02751.

[74] H. Xue, D. Q. Huynh, and M. Reynolds, “SS-LSTM: A Hierarchical LSTM Model for Pedestrian Trajectory Prediction,” in 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Mar. 2018, pp. 1186–1194. doi: 10.1109/WACV.2018.00135.

[75] P. Zhang, W. Ouyang, P. Zhang, J. Xue, and N. Zheng, “SR-LSTM: State Refinement for LSTM towards Pedestrian Trajectory Prediction.” arXiv, Mar. 07, 2019. doi: 10.48550/arXiv.1903.02793.

[76] T. Zhao et al., “Multi-Agent Tensor Fusion for Contextual Trajectory Prediction.” arXiv, Jul. 28, 2019. doi: 10.48550/arXiv.1904.04776.

[77] Y. Huang, H. Bi, Z. Li, T. Mao, and Z. Wang, “STGAT: Modeling Spatial-Temporal Interactions for Human Trajectory Prediction,” in 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Oct. 2019, pp. 6271–6280. doi: 10.1109/ICCV.2019.00637.

[78] Y. Xu, Z. Piao, and S. Gao, “Encoding Crowd Interaction with Deep Neural Network for Pedestrian Trajectory Prediction,” in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT: IEEE, Jun. 2018, pp. 5275–5284. doi: 10.1109/CVPR.2018.00553.

[79] “Pedestrian Trajectory Prediction Based on Deep Convolutional LSTM Network | IEEE Journals & Magazine | IEEE Xplore.” https://ieeexplore.ieee.org/document/9043898 (accessed May 05, 2023).

[80] R. Quan, L. Zhu, Y. Wu, and Y. Yang, “Holistic LSTM for Pedestrian Trajectory Prediction,” IEEE Trans. on Image Process., vol. 30, pp. 3229–3239, 2021, doi: 10.1109/TIP.2021.3058599.

[81] A. Sadeghian, V. Kosaraju, A. Sadeghian, N. Hirose, S. H. Rezatofighi, and S. Savarese, “SoPhie: An Attentive GAN for Predicting Paths Compliant to Social and Physical Constraints.” arXiv, Sep. 20, 2018. doi: 10.48550/arXiv.1806.01482.

[82] V. Kosaraju, A. Sadeghian, R. Martín-Martín, I. Reid, S. H. Rezatofighi, and S. Savarese, “Social-BiGAT: Multimodal Trajectory Forecasting using Bicycle-GAN and Graph Attention Networks.” arXiv, Jul. 16, 2019. doi: 10.48550/arXiv.1907.03395.

[83] H. Manh and G. Alaghband, “Scene-LSTM: A Model for Human Trajectory Prediction.” arXiv, Apr. 15, 2019. doi: 10.48550/arXiv.1808.04018.

[84] M. N. Azadani and A. Boukerche, "STAG: A novel interaction-aware path prediction method based on Spatio-Temporal Attention Graphs for connected automated vehicles," Ad Hoc Networks, vol. 138, p. 103021, Jan. 2023, doi: 10.1016/j.adhoc.2022.103021.

[85] A. Plitt, "New York City's streets are 'more congested than ever': report," Curbed NY, Aug. 15, 2019. https://ny.curbed.com/2019/8/15/20807470/nyc-streets-dot-mobility-report-congestion (accessed May 06, 2023).

[86] "Pedestrian scramble," Wikipedia. May 02, 2023. Accessed: May 06, 2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Pedestrian_scramble&oldid=1152818953.

[87] G. Agamennoni, J. I. Nieto, and E. M. Nebot, "A bayesian approach for driving behavior inference," in 2011 IEEE Intelligent Vehicles Symposium (IV), Jun. 2011, pp. 595–600. doi: 10.1109/IVS.2011.5940407.

[88] M. Brand, N. Oliver, and A. Pentland, "Coupled hidden Markov models for complex action recognition," in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Jun. 1997, pp. 994–999. doi: 10.1109/CVPR.1997.609450.

[89] T. Gindele, S. Brechtel, and R. Dillmann, "A probabilistic model for estimating driver behaviors and vehicle trajectories in traffic environments," in 13th International IEEE Conference on Intelligent Transportation Systems, Sep. 2010, pp. 1625–1631. doi: 10.1109/ITSC.2010.5625262.

[90] M. Liebner, M. Baumann, F. Klanner, and C. Stiller, "Driver intent inference at urban intersections using the intelligent driver model," in 2012 IEEE Intelligent Vehicles Symposium, Jun. 2012, pp. 1162–1167. doi: 10.1109/IVS.2012.6232131.

[91] "A survey on motion prediction and risk assessment for intelligent vehicles | ROBOMECH Journal | Full Text." https://robomechjournal.springeropen.com/articles/10.1186/s40648-014-0001-z (accessed May 07, 2023).

[92] "Modeling Vehicle Interactions via Modified LSTM Models for Trajectory Prediction | IEEE Journals & Magazine | IEEE Xplore." https://ieeexplore.ieee.org/document/8672889 (accessed May 07, 2023).

[93] Y. Ma, X. Zhu, S. Zhang, R. Yang, W. Wang, and D. Manocha, "TrafficPredict: Trajectory Prediction for Heterogeneous Traffic-Agents." arXiv, Apr. 09, 2019. Accessed: May 07, 2023. [Online]. Available: http://arxiv.org/abs/1811.02146.

[94] W. Ding and S. Shen, "Online Vehicle Trajectory Prediction using Policy Anticipation Network and Optimization-based Context Reasoning." arXiv, Mar. 03, 2019. doi: 10.48550/arXiv.1903.00847.

[95] M. Koschi and M. Althoff, “Set-Based Prediction of Traffic Participants Considering Occlusions and Traffic Rules,” IEEE Transactions on Intelligent Vehicles, vol. 6, no. 2, pp. 249–265, Jun. 2021, doi: 10.1109/TIV.2020.3017385.

[96] W. Ding, J. Chen, and S. Shen, “Predicting Vehicle Behaviors Over An Extended Horizon Using Behavior Interaction Network.” arXiv, Jun. 03, 2019. doi: 10.48550/arXiv.1903.00848.

[97] N. Deo and M. M. Trivedi, “Multi-Modal Trajectory Prediction of Surrounding Vehicles with Maneuver based LSTMs,” in 2018 IEEE Intelligent Vehicles Symposium (IV), Jun. 2018, pp. 1179–1184. doi: 10.1109/IVS.2018.8500493.

[98] N. Deo and M. M. Trivedi, “Convolutional Social Pooling for Vehicle Trajectory Prediction.” arXiv, May 14, 2018. doi: 10.48550/arXiv.1805.06771.

[99] K. Messaoud, I. Yahiaoui, A. Verroust-Blondet, and F. Nashashibi, “Attention Based Vehicle Trajectory Prediction,” IEEE Transactions on Intelligent Vehicles, vol. 6, no. 1, pp. 175–185, Mar. 2021, doi: 10.1109/TIV.2020.2991952.

[100] K. Messaoud, I. Yahiaoui, A. Verroust-Blondet, and F. Nashashibi, “Non-local Social Pooling for Vehicle Trajectory Prediction,” in 2019 IEEE Intelligent Vehicles Symposium (IV), Jun. 2019, pp. 975–980. doi: 10.1109/IVS.2019.8813829.

[101] F. Diehl, T. Brunner, M. T. Le, and A. Knoll, “Graph Neural Networks for Modelling Traffic Participant Interaction.” arXiv, May 07, 2019. doi: 10.48550/arXiv.1903.01254.

[102] X. Li, X. Ying, and M. C. Chuah, “GRIP: Graph-based Interaction-aware Trajectory Prediction,” in 2019 IEEE Intelligent Transportation Systems Conference (ITSC), Oct. 2019, pp. 3960–3966. doi: 10.1109/ITSC.2019.8917228.

[103] M. N. Azadani and A. Boukerche, “An Interaction-Aware Vehicle Behavior Prediction for Connected Automated Vehicles,” in ICC 2022 - IEEE International Conference on Communications, May 2022, pp. 279–284. doi: 10.1109/ICC45855.2022.9839048.

[104] Y. Wu et al., “HSTA: A Hierarchical Spatio-Temporal Attention Model for Trajectory Prediction,” IEEE Transactions on Vehicular Technology, vol. 70, no. 11, pp. 11295–11307, Nov. 2021, doi: 10.1109/TVT.2021.3115018.

[105] J. Li, H. Ma, Z. Zhang, J. Li, and M. Tomizuka, “Spatio-Temporal Graph Dual-Attention Network for Multi-Agent Prediction and Tracking.” arXiv, Feb. 17, 2021. doi: 10.48550/arXiv.2102.09117.

[106] "Multi-Agent Trajectory Prediction With Heterogeneous Edge-Enhanced Graph Attention Network | IEEE Journals & Magazine | IEEE Xplore." https://ieeexplore.ieee.org/abstract/document/9700483 (accessed May 07, 2023).

[107] Z. Sheng, Y. Xu, S. Xue, and D. Li, "Graph-Based Spatial-Temporal Convolutional Network for Vehicle Trajectory Prediction in Autonomous Driving," IEEE Trans. Intell. Transport. Syst., vol. 23, no. 10, pp. 17654–17665, Oct. 2022, doi: 10.1109/TITS.2022.3155749.

[108] J. Gao et al., "VectorNet: Encoding HD Maps and Agent Dynamics from Vectorized Representation." arXiv, May 08, 2020. doi: 10.48550/arXiv.2005.04259.

[109] C. Hyden, "The Development Of A Method For Traffic Safety Evaluation: The Swedish Traffic Conflicts Technique," Bulletin Lund Institute Of Technology, Department, No. 70, 1987, Accessed: May 07, 2023. [Online]. Available: https://trid.trb.org/view/239059.

[110] S. R. Perkins, "Traffic conflict characteristics-accident potential at intersections".

[111] C. Johnsson, A. Laureshyn, and T. De Ceunynck, "In search of surrogate safety indicators for vulnerable road users: a review of surrogate safety indicators," Transport Reviews, vol. 38, no. 6, pp. 765–785, Nov. 2018, doi: 10.1080/01441647.2018.1442888.

[112] H.-C. Chin and S.-T. Quek, "Measurement of traffic conflicts," Safety Science, vol. 26, no. 3, pp. 169–185, Aug. 1997, doi: 10.1016/S0925-7535(97)00041-6.

[113] A. Laureshyn, Å. Svensson, and C. Hydén, "Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation," Accident Analysis & Prevention, vol. 42, no. 6, pp. 1637–1646, Nov. 2010, doi: 10.1016/j.aap.2010.03.021.

[114] A. Laureshyn, T. De Ceunynck, C. Karlsson, Å. Svensson, and S. Daniels, "In search of the severity dimension of traffic events: Extended Delta-V as a traffic conflict indicator," Accident Analysis & Prevention, vol. 98, pp. 46–56, Jan. 2017, doi: 10.1016/j.aap.2016.09.026.

[115] T. De Ceunynck, "Defining and applying surrogate safety measures and behavioural indicators through site-based observations," 2017.

[116] J. Hayward, "Near-Miss Determination Through Use Of A Scale Of Danger," Highway Research Record, 1972, Accessed: May 08, 2023. [Online]. Available: Https://Www.Semanticscholar.Org/Paper/Near-Miss-Determination-Through-Use-Of-A-Scale-Of-Hayward/D7dc3003871814c9ac84d9c75456aa8089fc70fd.

[117] A. R. A. Van der Horst, "A time-based analysis of road user behaviour in normal and critical encounters," 1990. Accessed: May 08, 2023. [Online]. Available: https://repository.tudelft.nl/islandora/object/uuid%3A8fb40be7-fae1-4481-bc37-12a7411b85c7.

[118] M. Wessels, C. Zähme, and D. Oberfeld, “Auditory Information Improves Time-to-collision Estimation for Accelerating Vehicles,” Curr Psychol, Jul. 2022, doi: 10.1007/s12144-022-03375-6.

[119] D. Oberfeld, M. Wessels, and D. Büttner, “Overestimated time-to-collision for quiet vehicles: Evidence from a study using a novel audiovisual virtual-reality system for traffic scenarios,” Accident Analysis & Prevention, vol. 175, p. 106778, Sep. 2022, doi: 10.1016/j.aap.2022.106778.

[120] M. Kim, P. Ladosz, and H. Oh, “Monocular vision-based time-to-collision estimation for small drones by domain adaptation of simulated images,” Expert Systems with Applications, vol. 199, p. 116973, Aug. 2022, doi: 10.1016/j.eswa.2022.116973.

[121] X. Nie, Y. Liang, and K. Ohkura, “Autonomous highway driving using reinforcement learning with safety check system based on time-to-collision,” Artif Life Robotics, vol. 28, no. 1, pp. 158–165, Feb. 2023, doi: 10.1007/s10015-022-00846-8.

[122] S. Fakhoury and K. Ismail, “Improving Pedestrian Safety Using Ultra-Wideband Sensors: A Study of Time-to-Collision Estimation,” Sensors, vol. 23, no. 8, Art. no. 8, Jan. 2023, doi: 10.3390/s23084171.

[123] Tarko, Andrew, Gary Davis, Nicolas Saunier, Tarek Sayed, and Simon Washington. "Surrogate measures of safety. White paper." Transportation Research Board, Washington, DC (2009).

[124] L. N. Peesapati, M. P. Hunter, and M. O. Rodgers, “Can post encroachment time substitute intersection characteristics in crash prediction models?,” Journal of Safety Research, vol. 66, pp. 205–211, Sep. 2018, doi: 10.1016/j.jsr.2018.05.002.

[125] A. Ansariyar, “Investigating the Car-Pedestrian Conflicts Based on an Innovative Post Encroachment Time Threshold (PET) Classification.” Rochester, NY, Dec. 05, 2022. doi: 10.2139/ssrn.4377745.

[126] A. Ansariyar, A. Ardeshiri, and M. Jeihani, “Investigating the collected vehicle-pedestrian conflicts by a LIDAR sensor based on a new Post Encroachment Time Threshold (PET) classification at signalized intersections,” Advances in Transportation Studies, vol. 61, pp. 103–118, Apr. 2023.

[127] A. B. Paul, N. Gore, S. Arkatkar, and G. Joshi, “Investigating and modeling the influence of PET-types on crossing conflicts at urban unsignalized intersections in India,” International Journal of Injury Control and Safety Promotion, vol. 0, no. 0, pp. 1–16, Nov. 2022, doi: 10.1080/17457300.2022.2147194.

[128] J. Bock, R. Krajewski, T. Moers, S. Runde, L. Vater, and L. Eckstein, “The inD Dataset: A Drone Dataset of Naturalistic Road User Trajectories at German Intersections,” in 2020 IEEE Intelligent Vehicles Symposium (IV), Oct. 2020, pp. 1929–1934. doi: 10.1109/IV47402.2020.9304839.

[129] A. Robicquet, A. Sadeghian, A. Alahi, and S. Savarese, “Learning Social Etiquette: Human Trajectory Understanding In Crowded Scenes,” in Computer Vision – ECCV 2016, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2016, pp. 549–565. doi: 10.1007/978-3-319-46484-8_33.

[130] H. Su, J. Zhu, Y. Dong, and B. Zhang, “Forecast the plausible paths in crowd scenes,” in Proceedings of the 26th International Joint Conference on Artificial Intelligence, in IJCAI’17. Melbourne, Australia: AAAI Press, Aug. 2017, pp. 2772–2778.

[131] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, “Graph Attention Networks.” arXiv, Feb. 04, 2018. doi: 10.48550/arXiv.1710.10903.

[132] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese, and A. Alahi, “Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks.” arXiv, Mar. 28, 2018. doi: 10.48550/arXiv.1803.10892.

[133] A. Vemula, K. Muelling, and J. Oh, “Social Attention: Modeling Attention in Human Crowds.” arXiv, Oct. 29, 2018. doi: 10.48550/arXiv.1710.04689.

[134] C. Anderson, R. Vasudevan, and M. Johnson-Roberson, “Off The Beaten Sidewalk: Pedestrian Prediction In Shared Spaces For Autonomous Vehicles.” arXiv, Jun. 01, 2020. Accessed: May 12, 2023. [Online]. Available: http://arxiv.org/abs/2006.00962.

[135] D. Tøttrup, S. L. Skovgaard, J. le F. Sejersen, and R. Pimentel de Figueiredo, “A Real-Time Method for Time-to-Collision Estimation from Aerial Images,” Journal of Imaging, vol. 8, no. 3, Art. no. 3, Mar. 2022, doi: 10.3390/jimaging8030062.

[136] “An Improved Method to Calculate the Time-to-Collision of Two Vehicles | SpringerLink.” https://link.springer.com/article/10.1007/s13177-012-0054-4 (accessed Jun. 13, 2023).

[137] R. Miller and Q. Huang, “An adaptive peer-to-peer collision warning system,” in Vehicular Technology Conference. IEEE 55th Vehicular Technology Conference. VTC Spring 2002 (Cat. No.02CH37367), May 2002, pp. 317–321 vol.1. doi: 10.1109/VTC.2002.1002718.

[138] H. Alghodhaifi and S. Lakshmanan, “Simulation-based model for surrogate safety measures analysis in automated vehicle-pedestrian conflict on an urban environment,” in Autonomous Systems: Sensors, Processing, and Security for Vehicles and Infrastructure 2020, SPIE, 2020, pp. 8–21.

[139] S. Lakshmanan, Y. Yan, S. Baek, and H. Alghodhaifi, “Modeling and simulation of leader-follower autonomous vehicles: environment effects,” in Unmanned Systems Technology XXI, SPIE, May 2019, pp. 116–123. doi: 10.1117/12.2520485.

[140] E. Cheek, H. Alghodhaifi, C. Adam, R. Andres, and S. Lakshmanan, “Dedicated short range communications used as fail-safe in autonomous navigation,” in Unmanned Systems Technology XXII, SPIE, Apr. 2020, pp. 159–177. doi: 10.1117/12.2558925.

[141] H. Alghodhaifi, S. Lakshmanan, S. Baek, and P. Richardson, “AUTONOMY MODELING AND VALIDATION IN A HIGHLY UNCERTAIN ENVIRONMENT,” 2018.

[142] H. Alghodhaifi and S. Lakshmanan, “Safety model of automated vehicle-VRU conflict under uncertain weather conditions and sensors failure,” in Unmanned Systems Technology XXII, SPIE, May 2020, pp. 56–65. doi: 10.1117/12.2558756.