ANALYSIS OF PEDESTRIAN-VEHICLE CONFLICTS AT SIGNALIZED INTERSECTIONS

PROJECT REPORT

Submitted by

SREERESHMI P M

REG. No.: TKM20CETE15

to

the APJ Abdul Kalam Technological University

in partial fulfilment of the requirements for the award of the Degree

of

Master of Technology

in

Transportation Engineering

DEPARTMENT OF CIVIL ENGINEERING

T.K.M. College of engineering, Kollam

July 2022

DECLARATION

I undersigned hereby declare that the project report “Analysis of Pedestrian-Vehicle Conflicts at Signalized Intersections”, submitted for partial fulfilment of the requirements for the award of degree of Master of Technology of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done by me under supervision of Prof. Meenu Tomson. This submission represents my ideas in my own words and where ideas or words of others have been included, I have adequately and accurately cited and referenced the original sources. I also declare that I have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

Kollam SREERESHMI P M

07-07-2022

DEPARTMENT OF CIVIL ENGINEERING

T.K.M. COLLEGE OF ENGINEERING, KOLLAM

CERTIFICATE

Certified that the report entitled '**ANALYSIS OF PEDESTRIAN-VEHICLE CONFLICTS AT SIGNALIZED INTERSECTIONS'** is the report of project presented by **SREERESHMI P M., Reg. No.: TKM20CETE15** during **2021- 2022** in partial fulfilment of the requirements for the award of the Degree of Master of Technology in Transportation Engineering of the APJ Abdul Kalam Technological University.

Guide:	**Co-ordinator:**	**Head of the department:**
Prof. Meenu Tomson Assistant Professor Dept. of Civil Engg. TKMCE, Kollam	Dr. Kavitha Madhu Associate Professor Dept. of Civil Engg. TKMCE, Kollam	Dr. Sajeeb R Professor Dept. of Civil Engg. TKMCE, Kollam

ACKNOWLEDGEMENT

I take this opportunity to express my deep sense of gratitude and sincere thanks to all who helped me to complete the project successfully.

I am deeply indebted to my guide, **Prof. Meenu Tomson**, Assistant Professor, Department of Civil Engineering for her excellent guidance, positive criticism and valuable comments.

I am greatly thankful to my project coordinators, **Dr. Kavitha Madhu,** Associate Professor, **Prof. Meenu Tomson,** Assistant professor and **Prof. Jijin A**, Assistant professor, Department of Civil Engineering for their constant supervision as well as for providing necessary information regarding the project.

I would like to extend my sincere thanks to **Dr. Sajeeb R**, Professor and Head of the Department of Civil Engineering.

Finally, I thank my parents and friends who directly and indirectly contributed to the successful completion of my project.

SREERESHMI P M

ABSTRACT

Pedestrian safety is a growing concern, especially in developing nations with heterogeneous traffic conditions. One of the common safety issues at signalized intersections is pedestrian-vehicle conflicts, as they can occasionally result in serious collisions. The investigation was carried out at four signalized intersections in Kollam, Kerala. The study focused on conflicts caused by pedestrian signal violations and permissive left-turn movements. The study used surrogate safety measures such as post encroachment time (PET), time to vehicle (TTV), time to accident (TTA), and deceleration to safety time (DST) to analyze conflicts between pedestrians and automobiles fastly and proactively. The conflicts were divided into highly severe, severe and normal conflicts using k-means clustering. A support vector machine (SVM) model was developed to predict the severity level of conflicts. The study revealed that pedestrians' signal-violating behaviour caused 57% of the highly severe conflicts at signalized intersections. The study proposed threshold values of the conflict indicators for each severity level. The developed severity prediction model achieved 97% accuracy in predicting the conflicts' severity level. The study's findings aid in evaluating pedestrians' safety at signalized intersections.

Keywords: *Pedestrian-vehicle conflicts, violation behaviour, permissive left-turn traffic, Surrogate safety measures, Severity prediction model*

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

DST	-	Deceleration To Safety Time
FN	-	False Negative
FP	-	False Positive
IVM	-	Import Vector Machine
PET		Post Encroachment Time
RBF	-	Radial Basis Function
RTTC	-	Relative Time To Collision
SSM	-	Surrogate Safety Measures
SVM	-	Support Vector Machine
TCT	-	Traffic Conflict Technique
TN	-	True Negative
TP	-	True Positive
TTA	-	Time To Accident
TTC	-	Time To Collision
TTV	-	Time To Vehicle

CHAPTER 1
INTRODUCTION

1.1. Background

Walking is one of the most crucial sustainable ways of transportation for creating liveable and healthy communities. Many road agencies worldwide are implementing rules that support non-motorised modes of transport and encourage users to switch from motorised means to other sustainable ways. Walking is a practical strategy to raise local populations' levels of physical activity, which can positively impact their health (Hussein et al., 2015). A pedestrian is anybody who travels at least partially by foot. However, one of the most at-risk road users is the pedestrian. They typically account for the highest fatal traffic collisions and are prone to the most significant safety issues.

According to a report from the United Nations World Health Organization from 2016, traffic accidents resulted in more than 1.25 million fatalities and more than 20 million people suffering serious injuries. (Hacohen et al., 2020). Among the total road fatalities occurring worldwide, 22% are pedestrian-related (Kumar et al., 2019). Road traffic fatalities are a problem in all emerging nations, including India. The number of road accidents increased by 0.46% in 2018 as compared to the previous year. A study from the Indian Ministry of Road Transport and Highways states that there were 1, 51,417 fatalities and 4, 69,418 injuries in road accidents in India in 2018. Pedestrians made up 15% of those killed in traffic incidents in 2018, which is higher than 13.8% in 2017 (Kumar & Ghosh, 2022). Therefore, pedestrian safety is a rising concern worldwide, particularly in developing nations with diverse traffic patterns.

The modal separation of time and space plays a significant role in maintaining pedestrian safety. At intersections, however, where people on foot are exposed to moving vehicles, this separation is frequently impracticable or impossible. Therefore, complicated interactions between pedestrians and vehicles occur at intersections. Even at signalized intersections, pedestrian-vehicle collisions are becoming more common due to the growth in urban traffic volume. The primary causes of pedestrian-vehicle conflicts at crossings are the intersection's geometry, pedestrians' crossing patterns, the timing of traffic signals, drivers' actions, visibility, and pedestrian interaction patterns

(Alhajyaseen et al., 2012; Kumar et al., 2019; Zhang et al., 2020). Conflicts are near-miss incidents and can cause major crashes in the worst-case scenarios. Consequently, the safety of pedestrians at junctions can be assessed by analysing the conflict scenario.

1.2. Problem Statement

Since pedestrians are much less protected than drivers, pedestrian-vehicle conflicts are a frequent cause of injury and death, even at signalized intersections. Vehicles are given more attention in the design of traffic facilities than pedestrians. Due to uncontrolled left-turns at the crossings and pedestrian signal violation behaviour, signalized intersections experience the majority of conflicts.

At signalized intersections in India, people frequently disobey traffic signals and cross at risky locations. The unsafe pedestrian crossing is often the result of poor traffic signal design, which encourages pedestrians to ignore the signals at crossings. Other potential causes of the signal violation include prolonged curb side waiting times, belligerent pedestrian behaviour, the absence of a pedestrian refuge island at the intersection, a lack of traffic police enforcement, etc. Given their lack of physical defence against the effects of collisions, pedestrians' reckless crossing behaviour puts their lives in serious danger. The imposition of a fine for pedestrian offences is not, however, permitted in India. Also, traffic police barely enforce any laws and guidelines pertaining to pedestrian offences. It is pretty obvious that pedestrians who cross the street illegally endanger approaching traffic at the crosswalks (Kumar & Ghosh, 2022).

Free left-turns are another major contributor to conflicts between pedestrians and moving vehicles at signalized intersections. Vehicles making a left-turn are required to stop for pedestrians at the crosswalk as well as for oncoming traffic. If left-turning vehicles misread the situation and fail to yield to pedestrians during the permitted left-turn period, the chance of a collision increases in such complex driving circumstances (Qi & Guoguo, 2017).

The traditional approach to evaluate pedestrian-vehicle conflicts is based on crash statistics. However, erroneous results are brought on by the lack of trustworthy crash records and poorer data quality. The traffic conflict technique (TCT), which is based on surrogate safety measures, is employed to get over the drawbacks of the conventional approach. Analysing the conflict scenario with this methodology is proactive and short-term. (Kumar et al., 2019).

There is very little research on conflicts between pedestrians and vehicles caused by free left-turns and signal violations. Most of the research has been conducted at signalized intersections without pedestrian signals. In the current scenario, the pedestrian signal is installed at each approach at nearly all signalized junctions. However, collisions occur at intersections as a result of pedestrian non-compliance. The presence of unfettered left-turns puts pedestrians in danger of accident even if they do not disobey the signal. Therefore, it is necessary to examine the severity and frequency of disputes resulting from these two issues. Modelling and predicting the severity of the conflicts at the junctions is also crucial since more safety precautions need to be put in place.

1.3. Objectives

The objectives of the study are as follows:

- To identify pedestrian-vehicle conflicts at signalized intersections
- To analyse pedestrian-vehicle conflicts using surrogate safety measures
- To propose threshold values of conflict indicators for different severity levels of conflicts arising at signalized intersections
- To develop a severity prediction model for the conflicts occurring at signalized intersections

1.4. Methodology

The methodology planned to achieve the mentioned objectives is described below:

- Identification of study locations: The appropriate sites for collecting video data have been selected. The chosen intersections are Kottarakkara, Nilamel, Ayathil and Ayoor, all of which are situated in Kollam, Kerala. These crosswalks had pedestrian signals and were four-legged signalized intersections.
- Data collection and extraction: A high-resolution camera was positioned in the multi-story building to record the interactions between people and automobiles. The video recording was examined using the Kinovea 0.9.5

programme and time, speed and distance measurements were extracted from the software.

- Estimation of conflict indicators: The pedestrian-vehicle conflicts caused by signal violations and by unfettered left-turn movements were identified from the footage. Then conflict indicators were calculated for each recognized conflict event.
- Severity level classification of conflicts: the identified conflicts were categorized into various severity categories based on the estimated values of conflict indicators. The k-means clustering technique was used with the SPSS version 25 programme to accomplish this.
- Development of severity prediction model for the conflicts: Using the support vector machine (SVM) method, a severity prediction model for the conflicts happening at the intersection was developed.
- Interpretation of the results

1.5. Scope

The study aims to highlight pedestrian-vehicle conflicts at signalized intersections caused by non-compliance behaviour of pedestrians and uncontrolled left-turn movement of vehicles. The study was conducted in the Kollam district at four-legged signalized junctions with pedestrian signals and unrestricted left-turns. The present study incorporates left-turning, right-turning and through vehicle-pedestrian conflicts. The current study does not consider how intersection layout affects conflicts between pedestrians and moving vehicles.

1.6. Organization of Report

The thesis is structured into five chapters. Chapter one describes the background of the study, problem statement, objectives and scope of the study.

Chapter two reports the literature review on the crossing behaviour of the pedestrians, pedestrian-vehicle conflicts, traffic conflict technique and severity level classification.

Chapter three discusses the research methodology. The data collection and data extraction are described in this chapter. It includes the estimation of different surrogate safety measures and severity level classification of conflicts. The development of the severity prediction model is also discussed in this section.

Chapter four deals with the results and discussions of the study. It reports the non-compliance behaviour of the pedestrians and the distribution of conflicts at each site. This section includes the classification of severity level and reports the threshold values for the conflict indicators. This section also constitutes the analysis of the severity prediction model developed based on the conflict indicators.

Chapter five presents the major findings from the study. It also describes the recommendations based on the study and the scope for future work.

CHAPTER 2

LITERATURE REVIEW

2.1. Crossing Behaviour of Pedestrians

Various studies have been conducted worldwide to evaluate pedestrian crossing behaviour. According to Brosseau et al. (2013) there are three categories into which pedestrian crossing behaviours can be divided: regular crossings, dangerous crossings and crossing violations. The manner in which pedestrians cross the street and the resulting conflicts were influenced by several factors. At intersections, conflicts were aroused from sudden changes in the speed of the pedestrian crossing. The length of the crosswalk, the speed required to complete the crossing before the pedestrian signal turns red, and the entering speed all have a big impact on the speed change decisions (Alhajyaseen & Iryo-Asano, 2017). The Age and group size of pedestrians had a considerable impact on how they cross the street (Marisamynathan & Perumal, 2014). The study of Kim (2019) suggested that older pedestrians were particularly prone to harm or death when involved in collisions. Aghabayk et al. (2021) investigated how people cross the street at crosswalks with and without signals. They recognised that electronic distractions had an impact on pedestrians' behaviour when crossing the street. Their research found that people who were walking while texting or conversing on their phones while crossing the street paid less attention to the road. Furthermore, they discovered that group crossings were less secure than individual crossings. Studies (Aghabayk et al., 2021; Tiwari et al., 2007) have shown that gender affects pedestrians' behaviour when crossing the street. Studies revealed that when crossing at signalized crosswalks, women were more cautious than men.

2.2. Pedestrian-Vehicle Conflicts

All the interactions between pedestrians and vehicles are not conflicts. The conflicts occurring at the intersection are not collisions. Conflict is described as a circumstance when two road users are so close to one another that a collision is almost certain to occur if their motions don't change (Chaudhari et al., 2021). The conflict zone may result in a conflict between the two road users if they continue to proceed in the same direction without making any evasive manoeuvres (such as sharp steering or sidestepping movements, acceleration, or stopping). The conflict zone is defined theoretically for each pedestrian during their crossing as a common space used by road users/vehicles approaching from distinct trajectories. The risk of collision with approaching automobiles is one that pedestrians in this area face frequently (Kumar et al., 2019). Several variables might affect vehicle-pedestrian conflict at crossings, including geometric design, control strategy, driving style, environmental variables like signal timing and pedestrians, etc. (Chen et al., 2019). As violating pedestrians were left exposed to motorised traffic without the protection of traffic signals, it was discovered that these behaviours were crucial for conflicts between pedestrians and vehicles at signalized junctions. For instance, it was found that there was a positive correlation between the quantity of traffic conflicts and the spatial violations of pedestrians, or when they crossed outside of the crosswalks. Additionally, the features of pedestrians had an impact on their crossing behaviours, which could amplify any anomalies in their movements. As stated by research (Qi & Guoguo, 2017), at intersections with protected/permissive left-turn (PPLT) control or permissive only control, pedestrians will proceed during the permissive phase with parallel via vehicular flow. To choose the right gap, the left-turning vehicle must first yield to oncoming traffic as well as pedestrians. In these circumstances, the likelihood of a pedestrian collision increases. Therefore, one of the primary causes of the conflicts is the unfettered left-turn movements at the crossings. Two multiple linear regression models were created by Kumar & Ghosh, (2020) to examine the impact of various variables on the quantity and vicinity of conflicts. The study's findings implied that an increase in conflicts was caused by a rise in the number of pedestrians each cycle, or high pedestrian flow at the intersection. Ni et al., (2016) recommended a behaviour-based methodology for the identification of conflicts that happen at intersections. This methodology enables the identification of conflicts in practically all vehicular-pedestrian interaction patterns. According to the evasive action taken by the road user, Chaudhari et al. (2021) categorized the interactions between pedestrians and vehicles into dangerous conflicts, conflicts, and normal conflicts.

2.2.1. *Conflicts due to Non-compliance behaviour of pedestrians*

If there is a lot of conflicting motor traffic and time separation is necessary for a safe crossing, traffic lights are provided at pedestrian crossings. In India, people frequently disobey traffic signals and cross at crossroads in a risky manner. But occasionally, whether a person is impacted by personal requirements (such as being late for a meeting) or by environmental factors (such as a downpour), he or she may take a chance to get where they're going as soon as possible. In other words, if they feel it is safe, they may opt to ignore the signals and cross the street before they are supposed to. A pedestrian may occasionally misunderstand the safe time gap, or a driver may not anticipate the violation, which causes an accident (Koh et al., 2014). Several research on pedestrian violations looked into the variables influencing pedestrian behaviour. According to Mukherjee & Mitra (2019), longer waiting times, the intended use of the journey, and inadequate facilities are the main causes of signal infractions. Kumar & Ghosh (2022) suggested that demographic characteristics like age and gender had a significant impact on how pedestrians violate the law. Their research revealed that men were more likely than women to ignore the signal when crossing. Also, younger pedestrians disregarded the signals more often than middle-aged and older adults. The duration of the pedestrian's wait is likewise positively correlated with the frequency of violations. Furthermore, the study found that pedestrians who have been waiting for a while may feel impatient and choose to break the law. Koh et al. (2014) recognised that environmental factors like the infrastructure of pedestrian crossing amenities, contextual factors like other pedestrians' presence, and passing motorised traffic volume all had a role in how often pedestrians violated the law. In accordance with studies (Kumar & Ghosh, 2022; Mukherjee & Mitra, 2019), pedestrians who violate the signal at signalized junctions engage in risky interactions with other road users and may even engage in conflicts with them.

2.2.2. *Conflicts due to free left-turn movements*

Numerous studies have revealed that left-turning movements at signalized junctions frequently endanger pedestrians (Alhajyaseen et al., 2012, 2013; Haroun, n.d.; Y. L. He et al., 2019; Qi & Guoguo, 2017; Society & Studies, 2013). Permissive left-turning manoeuvres at signalized crossings pose a serious threat to pedestrian safety. Pedestrians move at the permissive phase with the parallel through vehicular flow for crossings with protected/permissive left-turn (PPLT) control or permissive only (POL) control. The obligation for drivers to stop for pedestrians when making a free left-turn is not mandated by law. As a result, they typically do not give pedestrians the right of way when turning. Based on their individual perceptions of the speed and proximity of the closest opposing vehicles, pedestrians cross the road at these turns. Essentially, this means that they must cross the street at their own risk (Society &

Studies, 2013). If left-turn vehicles misinterpret the situation and fail to yield to the pedestrians, the probability of a pedestrian collision increases. The traffic conditions at the junctions are the main emphasis of the current signal design standards for left-turn operations. Few of them specifically address pedestrian safety issues while determining the best left-turn control mode at a junction (Qi & Guoguo, 2017).

2.3. Traffic Conflict Technique (TCT)

In general, conflict analysis and before-and-after studies employing crash data are the most popular methodologies in safety evaluation (Alhajyaseen et al., 2013). Police-reported crash data served as a basis for the traditional technique of analysing pedestrian safety. Safety evaluation that is only based on crash records, however, has a number of drawbacks, including underreporting of crashes, protracted observation periods, a reactive approach, etc. This yields false results (Kumar et al., 2019). Traffic accidents are unusual occurrences, and it takes a while to compile enough accident data to generate accurate predictions of the expected number of accidents. Since only the ultimate accident consequence is visible during the registration phase, the real accident process is frequently obscure(Laureshyn & Várhelyi, 2018).

Many researchers (Chen et al., 2017; Kadali & Vedagiri, 2016; Kathuria & Vedagiri, 2020; Kumar et al., 2019; Ni et al., 2016; Xin et al., 2021; Zhang et al., 2020) have argued in favour of the use of the traffic conflict technique (TCT) as an indirect, short-term, and proactive safety technique, and it is being used as an alternative to the historical approach for reliable and faster safety evaluation. Some conflicts in a traffic stream result in collisions, which can range in severity from minor to severe. The proactive safety measurement is a practical method for assessing the severity of conflicts between vehicles and pedestrians when crossing roads (Kadali & Vedagiri, 2016). This approach is based on traffic conflict indicators or surrogate safety measures. Near-crash indicators, or surrogate safety measures (SSM), are used to gauge how close the conflicts are to occurring both spatially and temporally. The fundamental idea behind surrogate safety indicators is that all traffic incidents involving proximity of any kind between two or more road users are connected to safety (Johnsson et al., 2018). According to Marisamynathan & Vedagiri (2020) SSM is predicated on the belief that interactions, which have a high probability of occurring, are what cause accidents.

Various studies have used surrogate safety measures to assess the conflict situation (Alhajyaseen, 2015; Chen et al., 2017; Johnsson et al., 2018; Kadali & Vedagiri, 2016; Kathuria & Vedagiri, 2020; Kumar et al., 2019; Kumar & Ghosh, 2020; Pawar et al., 2022; Shah & Vedagiri, 2017; Shekhar Babu & Vedagiri, 2018). The generally employed surrogate safety measures to analyse conflict situations are post encroachment time (PET), time to collision (TTC), time to vehicle (TTV), time to accident (TTA), and deceleration to safety time (DST_{veh} and DST_{ped}). When analysing pedestrian-vehicle collisions, PET is the most frequently employed surrogate safety measure. PET is defined as the interval of time between the first road user's departure from the possible conflict zone and the second road user's arrival (Almodfer et al., 2016; Chaudhari et al., 2021). Most serious conflicts result from post-encroachment times with shorter values. The road user travels at their usual speed with little chance of conflict when the PET value is more than 5 seconds. Consequently, PET less than 5s is regarded as a conflict, while PET less than 1.5s is characterised as a serious conflict (Chaudhari et al., 2021). TTC is yet another measure frequently employed in studies (Kathuria & Vedagiri, 2020; Ni et al., 2016) to analyse conflicts. When two road users are travelling at their current speeds and following the same routes, the time to collision (TTC) is the amount of time before a collision occurs. It is a continuous measure of time. In contrast, it is uncommon for a pedestrian-vehicle collision to occur when both a vehicle and a pedestrian are present at the same time at the intersection of their respective trajectories. To overcome this issue, (Chen et al., 2017) use the relative time to collision (RTTC) as the indicator to assess the severity of conflicts. If both road users maintain their present speeds, the RTTC is the time difference between the first road user and the second road user arriving at the potential conflict point. However, measuring the minimal TTC manually is quite challenging and resource-intensive. In order to assess the proximity of conflicts, Kumar et al. (2019) developed two indicators: vehicle TTA and pedestrian TTV, which are utilised independently for vehicles and pedestrians, respectively.

2.4. Severity Level Classification

Conflict severity cannot be assessed only on the individual values of conflict indicators. Consequently, a proper classification should be made based on the values of conflict indicators in order to obtain a clear image of the severity level of conflicts. For the classification of severity levels, distinct methods are used in various research.

The k-means clustering analysis was employed by Kumar et al. (2019) to categorise the conflict indicators into various degrees of severity. This method implements a logical classification in accordance with the pertinent significances and traits. The validation of clustering was done using silhouette plots. They divided the pedestrian-conflicts into 4 severity levels based on the values of conflict indicators. To categorise the severity levels, Xin et al. (2021) employed k-means fuzzy clustering. According to the values of SSMs, severity levels were divided into 5 grades. They discovered that the frequency and severity of conflicts involving pedestrians and right-turning automobiles varied on various parts of the crosswalk.

In order to categorise the severity levels based on PET and TTC values, Ni et al. (2016) used the support vector machine technique (SVM). Three different patterns of interactions between pedestrians and automobiles were established, each with a corresponding severity level. The pre-defined threshold values of the indicators for each interaction pattern are used to categorise severity levels as safe passage, critical event, and conflict.

Import vector machine (IVM) technology was employed by Kathuria & Vedagiri (2020) to categorise the severity degrees. The interaction patterns were divided into two categories by them. Safe passage, mild interaction, and critical interaction are the three severity levels for each pattern according to the IVM technique.

Utilizing the Swedish traffic conflict approach (STCT), Kumar & Ghosh, (2022) classified the severity levels. Based on the likelihood of a conflict occurring, the time to accident (TA) and conflict speed (CS) values were categorised into three general categories: potential conflicts (severity levels 20–23), slight conflicts (severity levels 24-26), and serious conflicts (severity levels 26–30). They discovered that the end of the pedestrian signal's red interval was when around 60% of significant conflicts took place.

2.5. Summary

Pedestrian crossing behaviours can be divided into three categories: regular crossings, dangerous crossings and crossing violations. Age and group size of pedestrians have a considerable impact on how they cross the street. Also found that electronic distractions had an impact on pedestrians' behaviour. All interactions between pedestrians and vehicles are not collisions. Conflict is a circumstance when two road users are so close to one another that a collision is almost certain to occur if their motions remains same. There are several variables that might affect vehicle-pedestrian conflict at crossings, including geometric design, control strategy and driving style. Collisions between pedestrians and vehicles occur more frequently at intersections. The primary causes of conflicts is signal violation behaviour of the pedestrian and unfettered left-turn movements at the crossings. In India, people frequently disobey traffic signals and cross at crossroads in a risky manner. Longer waiting times, the intended use of the journey, and inadequate facilities are the main causes of signal violations. Permissive left-turning manoeuvres at signalized crossings pose a serious threat to pedestrian safety. Drivers typically do not give pedestrians the right-of-way when turning. If drivers misinterpret the situation and fail to yield to the pedestrians, the probability of a collision increases. In general, conflict analysis and before-and-after studies employing crash data are the most popular methodologies in safety evaluation. The proactive safety measurement is a practical way to assess the severity of conflicts between vehicles and pedestrians when crossing roads. This approach is based on traffic conflict indicators or surrogate safety measures. When analysing pedestrian-vehicle collisions, PET is the most frequently employed surrogate safety measure. Most serious conflicts result from post-encroachment times with shorter values. Other conflict indicators that generally used to analyse the conflict situation are TTV, TTA and DST. Conflict severity cannot be assessed only on the individual values of conflict indicators. A proper classification should be made based on the combined values of these indicators. For the classification of severity levels, distinct methods are used in various research.

CHAPTER 3

METHODOLOGY

3.1. General

This chapter deals with the research methodology adopted for the study, which is shown in Figure 3.1. The study locations were identified in Kollam, Kerala and were four-legged, signalized and had pedestrian signals. Video data were collected from each intersection during peak hours using a high-resolution camera. The conflicts between pedestrians and vehicles occurring at the intersections were identified from the recorded video by manual observation and corresponding surrogate safety measures were calculated by extracting time, distance and speed measurements from Kinovea software version 0.9.5. Then the conflicts were classified into different severity levels using k-means clustering in SPSS version 25. Then, a severity prediction model was developed by support vector machine approach by training the data using the k-means clustering output to predict the severity level of conflicts.

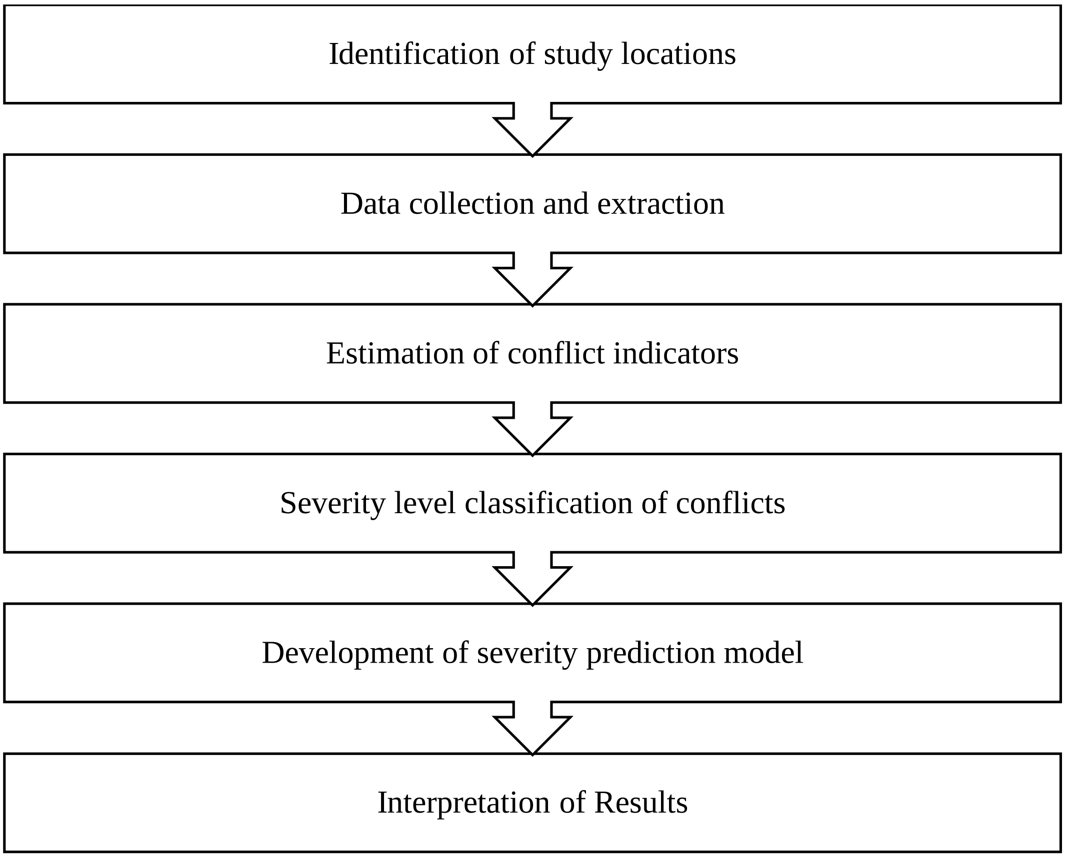


Figure 3.1. Study Methodology

3.2. Study Locations

Kottarakkara, Nilamel, Ayathil, and Ayoor were the four intersections in Kollam, Kerala that were chosen for investigation. These locations were four-legged and offered free left-turns at junctions. At the study sites, fixed signal control plans with pedestrian signals are put into practice. In the Table 3.1 below, the specifics of the study locations are displayed.

Table 3.1. Characteristics of Study Locations

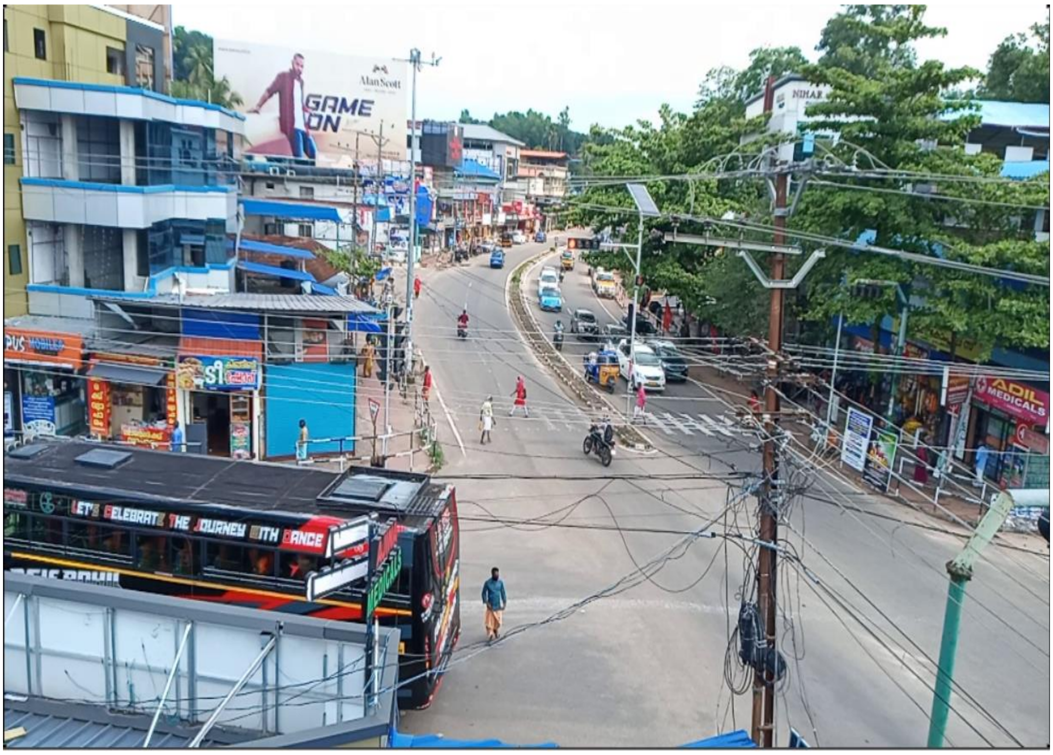
Site No.	Name of site	Target crosswalk	Pedestrian flow	Cycle length (s)	Pedestrian green interval (at one approach) (s)
1	Kottarakkara	Northbound	275/h	115	11
2	Nilamel	Northbound	160/h	115	10
3	Ayathil	Southbound	71/h	119	31
4	Ayoor	Northbound	178/h	114	8

3.3. Data Collection And Extraction

At the intended crosswalks of the intersection, a high-resolution camera was mounted on a multi-story building to record how people and cars interacted there. The video was recorded at 30 frames per second in 1080 P resolution. During morning the (8.30–10.00 am) and evening (4.00–5.30 pm) peak hours, three hours of data were gathered from each intersection. The sample images of the locations where the data were collected are shown in Figure 3.2.

With the aid of Kinovea software 0.9.5, data extraction was completed. It is a free video processing programme that is frequently used for sports analysis that demands the highest level of accuracy. An adequate level of accuracy in angular and linear measurements can be acquired using the semi-automated tracking programme Kinovea by digitising the x- and y-axis coordinates.

(a) Pulamon Junction, Kottarakkara



(b) Nilamel Junction

(c) Ayathil Junction

(d) Ayoor Junction

Figure 3.2. The Sample Pictures of Study Locations

Figure 3.3 displays a snapshot of the Kinovea software. The tracking in Kinovea is semi-automated since, when the automatic tracking fails to follow the intended path, the user can adjust the track point location using the outer rectangle. Control of the track is necessary because occasionally the tracked object is not properly contrasted and feature rich, which could lead to an inaccurate track route. Typically, it occurs while pedestrians are being tracked. The dynamic rectangles, i.e., feature region and search region, distinguish the tiny pedestrians due to their movement over the static background and also due to the colour contrast. However, this is only achievable if the video is captured in 1080 P with a frame rate of 30fps (Kathuria & Vedagiri, 2020). The time, distance, and instantaneous speed measurements were then retrieved using the software after conflicts were manually detected by watching the recorded video.

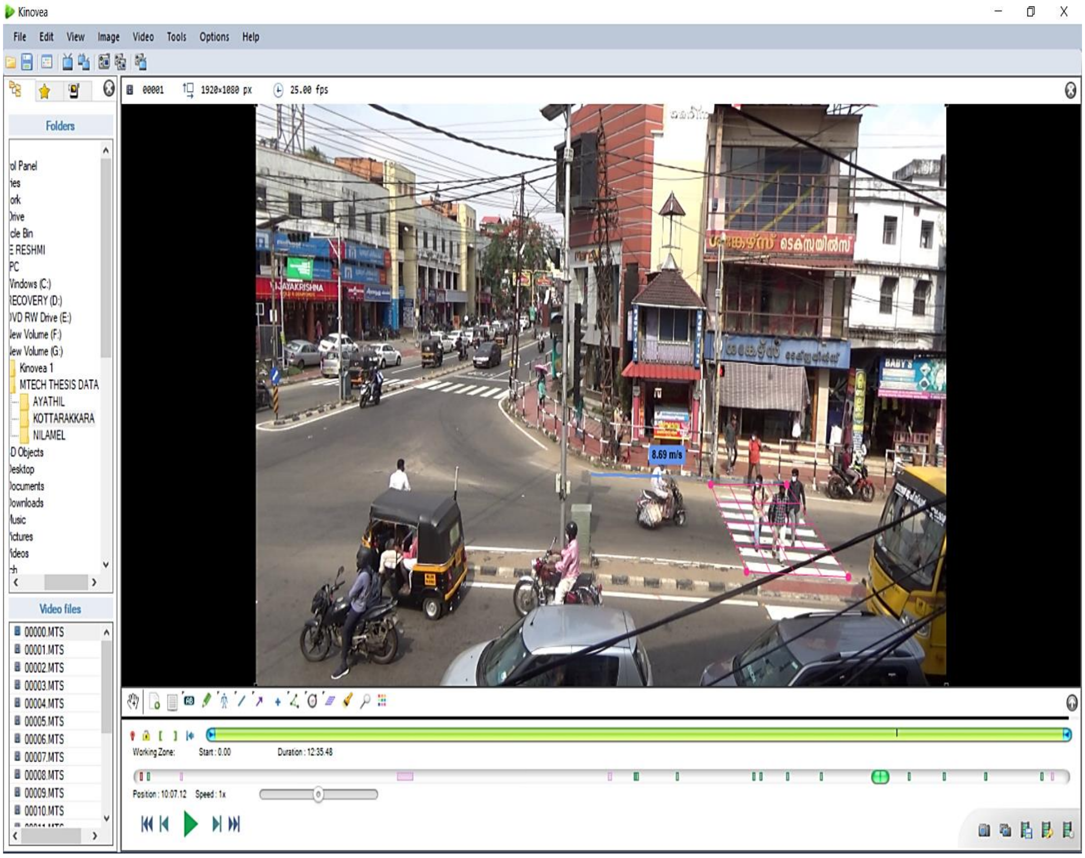


Figure 3.3. Snapshot of Kinovea software

3.3.1. *Identification of conflicts*

Videos of all the interactions between vehicles and pedestrians were reviewed. Following that, using the behaviour-based methodology recommended in the studies, conflicts between pedestrians and automobiles were manually recognised (Kumar et al., 2019; Kumar & Ghosh, 2022; Ni et al., 2016). They presented five scenarios, and if any of the following ones occurred, conflict was considered to have taken place:

- A pedestrian moved back, slowed down, or stopped and waited for a vehicle to pass
- A pedestrian sprinted or accelerated and passed the conflict point before a vehicle arrived;
- A vehicle braked, stopped, or changed path to avoid a pedestrian;
- A vehicle accelerated and passed before a pedestrian arrived at the conflict point;
- Neither pedestrians nor automobiles took any action, although there was a close distance and brief delay as they passed the conflict zone.

3.4. Estimation of Conflict Indicators

Conflict indicators or surrogate safety measures were then computed for each conflict situation after the conflicts had been identified. To analyse the conflict situation, the current study used five parameters: post encroachment time (PET), time to vehicle (TTV), time to accident (TTA), and deceleration to safety time (DST_{ped} and DST_{veh}).

3.4.1. *Post encroachment time (PET)*

It is the interval of time between when the first road user (vehicle or pedestrian) departed the potentially active conflict zone and when the second road user (vehicle or pedestrian) arrived at it. The schematic diagram for the calculation of PET is shown in Figure 3.4. Using the kinovea programme, a grid was placed that was the same size as the crosswalk since the crossing was regarded as the conflict zone for the purposes of calculating PET. It was noted what time the first road user left the zone and when the second road user entered the zone. The kinovea software employing a grid is shown in a screenshot in Figure 3.6. The following formula 3.1 is used to compute PET:

$$PET = T_L - T_E \tag{3.1}$$

Where T_L is the time when the first road user leaves the conflict zone and T_E is the time when the second road user enters the conflict zone.

3.4.2. *Time to vehicle (TTV)*

TTV is the time when pedestrian would reaches the conflict point. By calibrating a well-known reference dimension in the video, the distance and speed were extracted using the programme for the determination of TTV. Figure 3.5 shows the schematic diagram for the calculation of TTV. The Speed of the pedestrian was extracted by tracking the movement of the pedestrian. The following formula (3.2) is used to compute TTV:

$$\text{TTV} = \frac{L_p}{Vp} \quad (3.2)$$

Where L_p is the distance between the conflict point and the point at which the pedestrian began to decelerate and Vp is the speed of the pedestrian at the moment at which he/she decelerated.

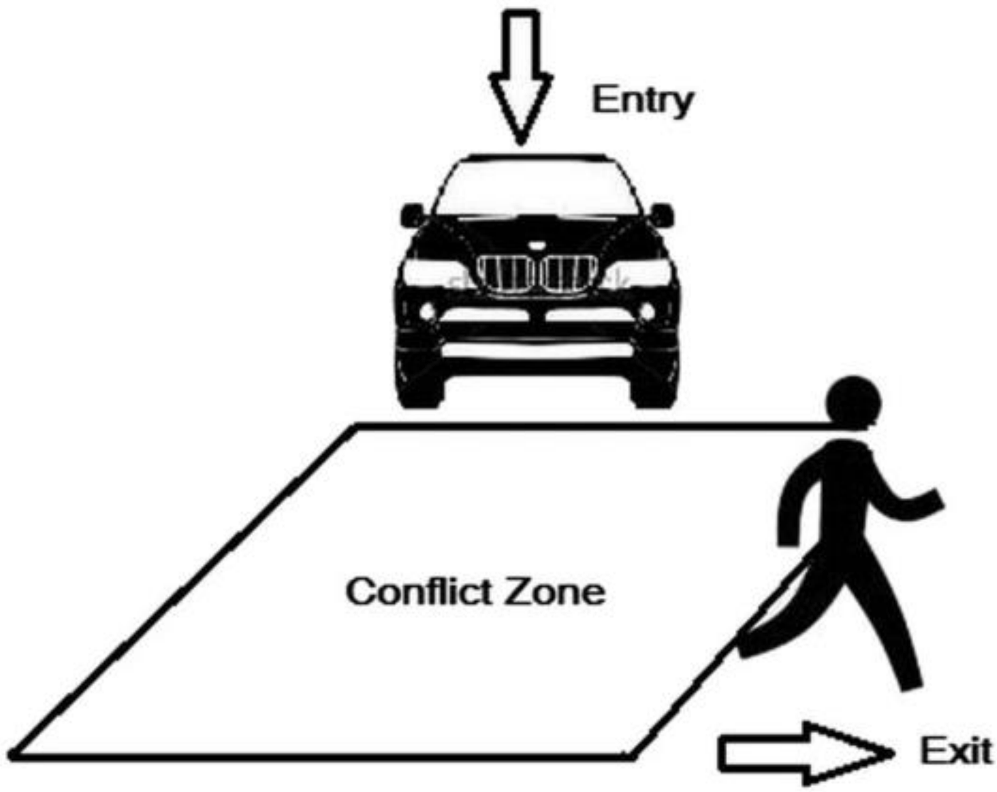


Figure 3.4. Schematic Diagram for the Calculation of PET (Kumar et al., 2019)

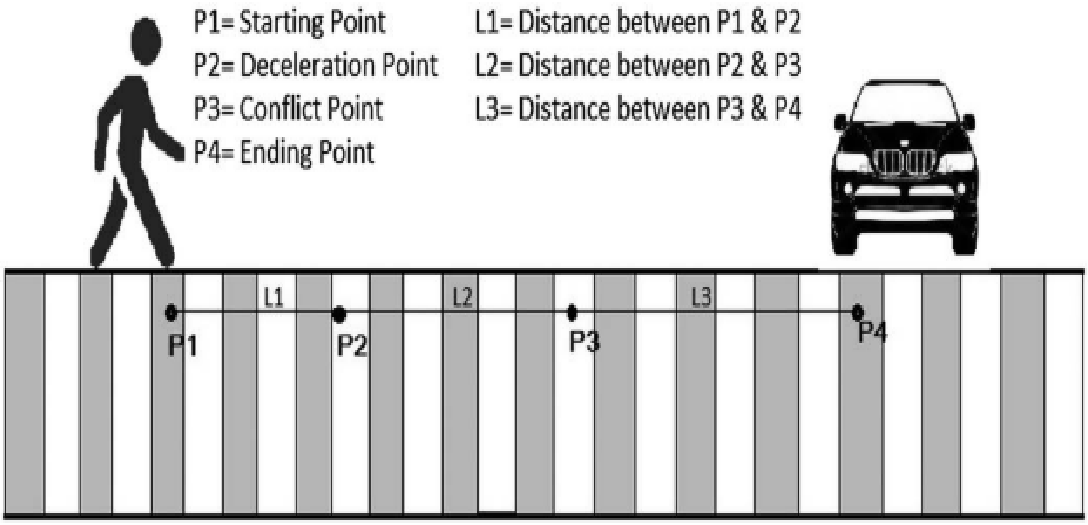


Figure 3.5. Schematic Diagram for the Calculation of TTV (Kumar et al., 2019)

3.4.3. *Time to accident (TTA)*

With the exception of using vehicle speed and positions rather than pedestrians, this metric is derived in the same manner as TTV. Figure 3.7 below illustrates how the software tracks the motion of moving vehicles to determine speed. The following equation (3.3) can be used to compute TTA:

$$\text{TTA} = \frac{L_v}{V_v} \tag{3.3}$$

3.4.4. *Deceleration to safety time (DST)*

DST is computed for both automobiles and pedestrians. It provides superior results for traffic events and indicates the average rate of deceleration of the vehicles and pedestrians during the crossing. During the interaction process, the kinovea software measures time and distance. The following formula (3.4) is used to determine DST:

$$\text{DST} = \frac{2L_p}{(T_3 - T_2)^2} \tag{3.4}$$

Where T_3 is the moment at which a pedestrian/vehicle decelerated, T_3 is the moment at which pedestrian/vehicle passes the conflict point and L_p is the distance between them.

Figure 3.6. Calculation of PET Using Grid in Kinovea

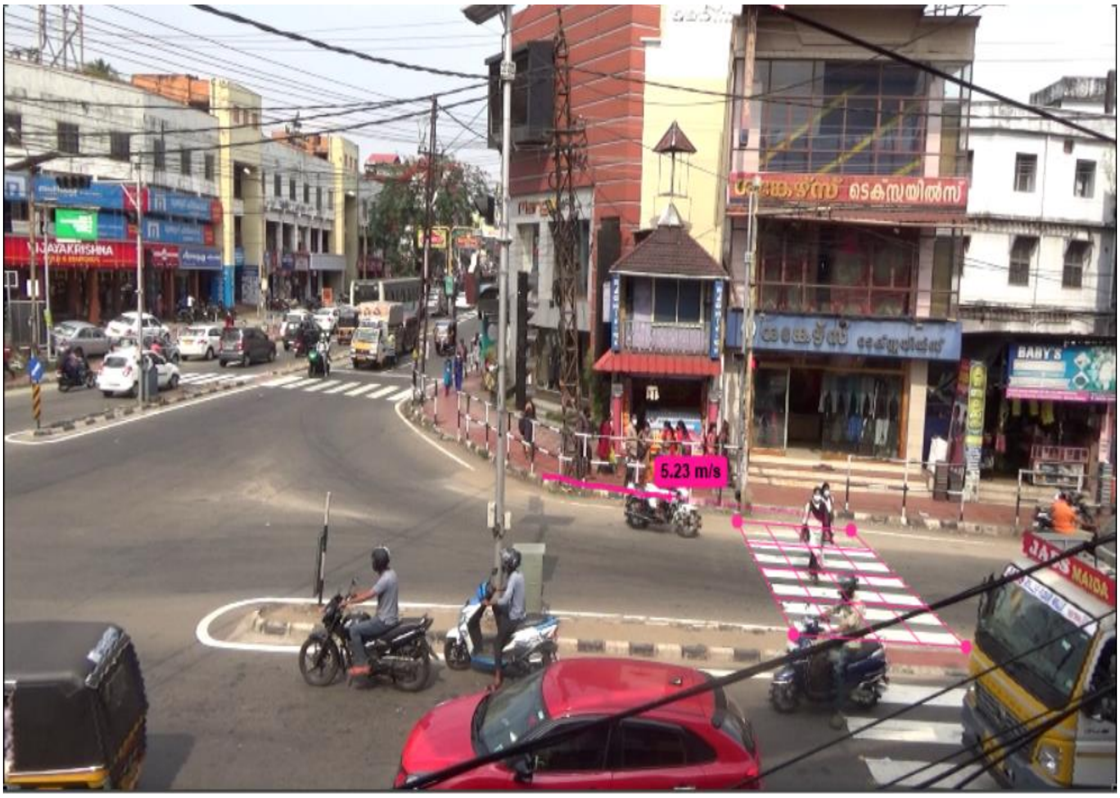


Figure 3.7. Tracking of Vehicle to Extract Speed in Kinovea

3.5. Severity Level Classification of Conflicts

Conflict indicators provide a general understanding of the risk a pedestrian is exposed to when crossing. However, the severity of the disputes has an impact on the risk assessment in reality. A conflict's severity indicates the likelihood that it will end in a crash. K-means clustering, which was used in conjunction with SPSS version 25, was used in the current study to classify the severity levels. Compared to other clustering approaches, it is effective at handling huge datasets with quantitative variables, making it a widely used clustering technique. The hierarchal clustering was performed prior to k-means clustering to find the optimal number of clusters. Estimated conflict indicator values for each conflict occurrence (PET, TTV, TTA, and DST) served as the input variables for the k-means clustering analysis.

3.5.1. *Hierarchal clustering*

The goal of hierarchical clustering, also known as hierarchical cluster analysis or HCA, is to create a set of clusters. These clustering techniques operate on the principle of agglomeration of members or states. With this approach, every observation starts in its own cluster, and as one advances up the hierarchy, pairs of clusters are combined (L. He et al., 2018). Each stage of agglomeration estimates the cost of combining members. This cost, which is expressed as an agglomerate coefficient, offers greater assistance in choosing the optimal number of clusters. To determine the ideal number of clusters, the change in the agglomeration coefficient at each stage is carefully examined (Maji et al., 2017). Hierarchal clustering was done by ward's method which uses squared Euclidian distance between the clusters for agglomeration of members. Ward's technique takes into account agglomeration with a minimum increase in average distance. The optimum number of clusters obtained from hierarchal clustering was used in k-means clustering.

3.5.2. *k-means clustering*

It is an effective tool for automatically forming coherent clusters out of data. This technique separates a dataset into categories based on how similar or different the patterns are. The k-means clustering is an appropriate classification technique that divides an unlabelled dataset into a large number of clusters with a high degree of similarity within each group (Kadir et al., 2018; Kumar et al., 2019). The process is started by giving the iterative algorithm a fixed set of centroids and a predetermined number of clusters to work with. Then, using a squared Euclidian distance, each data point to be clustered is matched with its closest centroid. The algorithm's goal is to identify the cluster centres as widely apart from one another as feasible and link each data point to the closest cluster centre. The dissimilarity metric in the k-means method is typically the Euclidean distance. The objective function J is described in following equation (3.5):

$$J = \sum_{i=1}^{k} \left(\sum_{k} \| X_k - C_i \|^2 \right) \quad (3.5)$$

where K is the number of clusters, C_i is the centres of clusters, and is X_k is the kth data point in ith cluster (Orhan et al., 2011). By calculating the average of all the points given to each cluster, the centroids are then updated. These stages are repeated until there is no substantial change in the assignment of the data points to each centroid (Saha et al., 2019).

3.6. Development of Severity Prediction Model

A SVM model was established to forecast the severity of the conflicts based on the estimated conflict indicator values. The support vector machine model, a relatively recent modelling approach, was proposed to address the classification and regression issues. It was discovered that the SVM model offers excellent potential for addressing classification issues (Li et al., 2012). In comparison to other machine learning methods, the support vector machine also provides superior accuracy for even smaller datasets (Wang et al., 2019). As a result, this strategy was chosen in the current study to create a prediction model.

Data is divided into a training set and a testing set before being specified for the SVM model. Every conflict in the training dataset includes a label that specifies the level of severity and its associated individual conflict indicator values (PET, TTV, TTA and DST). On the basis of the training data, the SVM model discovered the correlations between conflict severity and explanatory variables. Following that, the SVM model can forecast the severity level of each conflict based on values for conflict indicator values from the testing conflict dataset.

3.6.1. *Support vector machine (SVM) technique*

Support vector machines are supervised machine learning techniques that can be used to categorise data into distinct groups. The structural risk reduction concept and statistical learning theory are the foundations of the SVM model. The SVM model is able to transform the input vector X into a multidimensional feature space. The SVM model generates an optimal separation hyperplane in this higher dimensional space to divide the outcome into various groups while optimising the margin between the linear decision boundaries by selecting a non-linear mapping a priori.

The data is divided into a training set and a testing set when specifying the SVM model. Based on the training set, the SVM model generates a learning model, which is subsequently used to predict data from the testing set. The whole collection of conflict-related variables, including the severity of conflicts, are represented by the vectors $x_i \in N$ for i = 1, 2…, N. The severity of conflicts is represented by the hyperplane$y_i \in N$, which divides outcomes and training output. It is possible to express the hyperplane for separating outcomes as the collection of points X satisfying the equation (3.6):

$$W \cdot X - b = 0 \quad (3.6)$$

where "•" stands for the dot product, "W" is a normal vector that is perpendicular to the hyperplane, and "b" is a coefficient that stands for bias (Li et al., 2012). One of these hyperplanes has the greatest margin and is recognised as the optimal separating hyperplane. The margin vectors, or support vectors, are what specifically determine this hyperplane (Yu & Abdel-aty, 2013).

By combining linear classification with a technique known as the kernel trick, SVM efficiently accomplishes nonlinear classification tasks. It makes sense that choosing the hyperplane with the widest margin as the classifier yields the lowest misclassification of new data when establishing the decision boundary. The equation (3.7) for kernel function is shown below (Han et al., 2020):

$$K(x_{m,}x) = K(x^T mx) = \emptyset(x_m)^T \emptyset(x) \quad (3.7)$$

The selection of an appropriate kernel function will determine how well an SVM classifier performs. Predictive performance can be enhanced by selecting the ideal kernel function parameters, which can also prevent overfitting (Han et al., 2020). The sigmoid, linear, and radial basis function (RBF) kernels were used in the current investigation. Below formulae (3.8) to (3.10) represent various kernels (Hussain et al., 2011; Yu & Abdel-aty, 2013).

Linear kernel: $K(x_i, x_j) = x_i^T x_j$ (3.8)

Radial basis function (RBF) kernel: $K(x_i, x_j) = e^{(-\gamma(\|x_i - x_j\|)^2)}, \gamma > 0$ (3.9)

Sigmoid kernel: $K(x_i, x_j) = \tan h(\gamma x_i^T x_j + r)$ (3.10)

3.6.2. *Evaluation of SVM model*

Precision, recall, F1 score, and accuracy are the indices used to evaluate the SVM severity prediction model. These metrics were computed based on variables such as true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The number of samples that are accurately identified as true positives is referred to as TP. False positive (FP) samples are those that are incorrectly categorised despite being true negative ones. FN (false negative) is the number of actual negative samples that are wrongly classified. The number of actually negative samples that are correctly categorised is known as the TN (true negative). The equations for calculating the metrics using these four values are presented in equations (3.11) to (3.14) (Zhang et al., 2020). Precision, also known as positive predictive value (PPV), is the proportion of real positive samples to samples that have been labelled as positive. The percentage of real positive samples that are accurately categorised is called recall, which is often referred to as sensitivity. The F1 score is an integrated statistic that takes into account both recall value and precision value. The ratio of samples that were correctly identified over the entire data set, accounting for both positive and negative samples, is known as accuracy.

Precision = $\frac{\text{TP}}{\text{TP+FP}}$ (3.11)

Recall = $\frac{\text{TP}}{\text{TP+FN}}$ (3.12)

F1 score = $\frac{2*\text{Precision}*\text{Recall}}{\text{Precision+Recall}}$ (3.13)

Accuracy = $\frac{\text{TP+TN}}{\text{TP+FP+TN+FN}}$ (3.14)

By comparing the expected and observed values, accuracy may be evaluated. Accuracy sheds light on the SVM classifier's overall performance (Li et al., 2012).

CHAPTER 4
RESULTS AND DISCUSSIONS

4.1. Non-Compliance Behaviour of Pedestrians

A total of 2528 pedestrians were crossed during the three hours of observation at each study location. Among them, 77% of pedestrians ignored the signal and crossed the street while the light was still red. The percentage of pedestrians who crossed the roadway while the pedestrian signal was green was only 23%. This makes it abundantly evident that pedestrians are acting in a non-compliant manner. The frequency distribution of total pedestrian crossings is shown in the Figure 4.1. The frequency distribution of pedestrian crossings during the red and green pedestrian signal indications at each site is illustrated in Table 4.1. More than half of pedestrians ignore signals at every site, according to a comparison of the frequency distribution of each site. This suggests that pedestrians are just ignoring the pedestrian signal and crossing at intersections when they perceive a safe gap. This is a result of longer waiting times and a lack of punishment for signal violations. Highest percentage of pedestrians (83%) violating the signal are at site 2 a and relatively less percentage (54%) at site 3 according to the data collected.

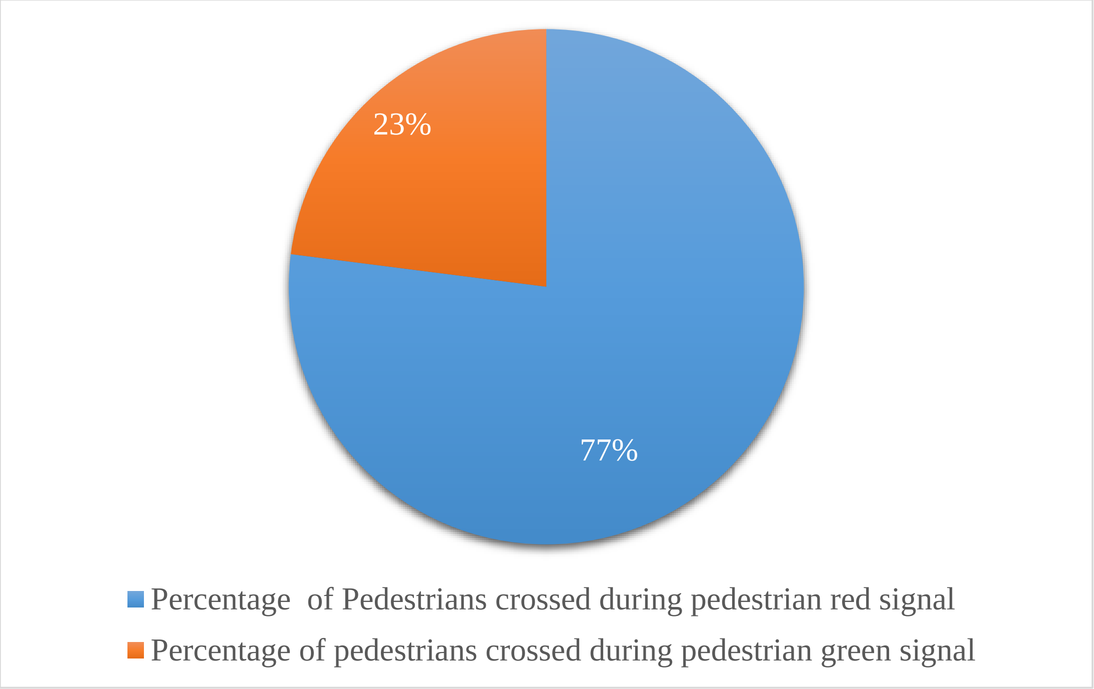


Figure 4.1. Frequency Distribution of Total Pedestrian Crossings

Table 4.1. Observed Frequency Distribution Pedestrian Crossings at Each Site

Site No.	No. of pedestrians crossed during red signal	No. of pedestrians crossed during green signal
1	905 (81.02%)	212 (18.98%)
2	508 (83%)	104 (17%)
3	142 (54%)	121 (46%)
4	392 (73.13%)	144 (26.87%)

4.2. Frequency Distribution of Conflicts

From all of the study sites, 374 conflicts were observed in total. The identified conflicts were categorized into two: conflicts due to signal violation behaviour of pedestrians and conflicts due to free left-turn movements. The pedestrian-vehicle conflicts caused by both right-turning and through moving vehicles were considered under the category of conflicts due to signal violations. According to the data gathered, 35% of conflicts are caused by free left-turn movements, while 65% are caused by signal violations. The distribution of total conflicts identified is shown in Figure 4.2. The frequency distribution of conflicts at each site is listed in Table 4.2. It was found that at all the study sites, conflicts due to signal violation shows higher percentage than conflicts due to free left-turn movements except at site 1.

Table 4.2. Frequency Distribution of Identified Conflicts at Each Site

Site No.	Conflicts due to signal violations	Conflicts due to free left-turn movements
1	72 (47.06%)	81 (52.94%)
2	64 (76.19%)	20 (23.81%)
3	42 (77.78%)	12 (22.22%)
4	57 (68.68%)	26 (31.32%)

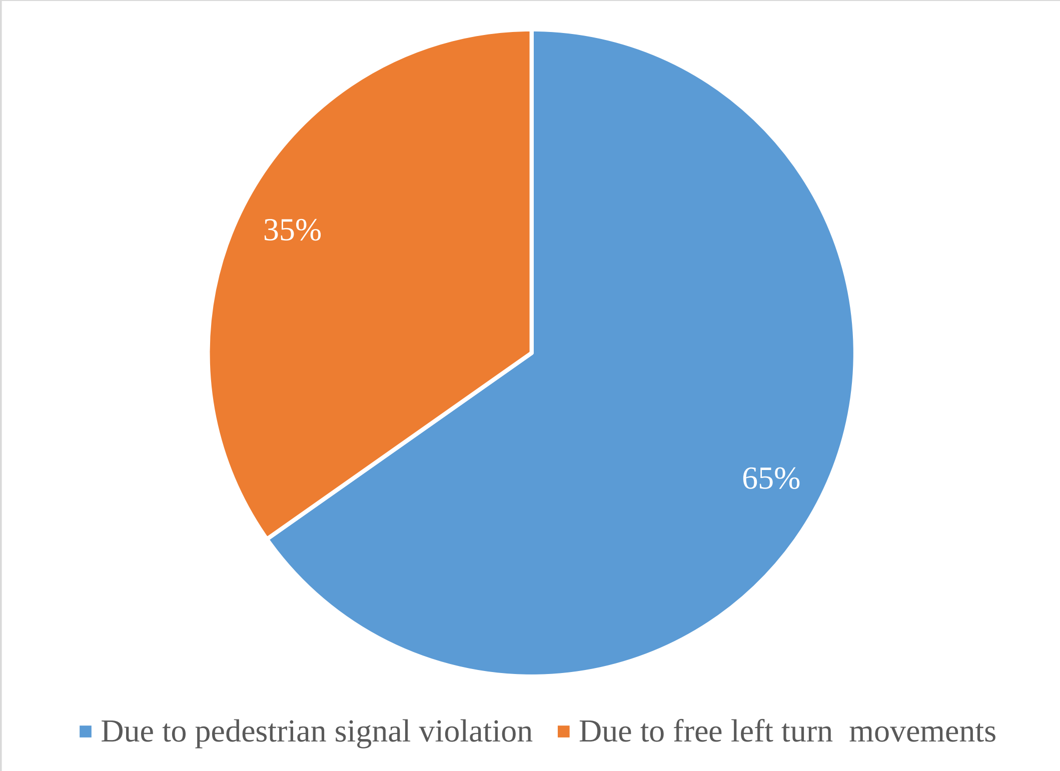


Figure 4.2. Distribution of total identified conflicts

4.3. Evaluation of Surrogate Safety Measures

Conflict indicators (PET, TTV, TTA, DST_{ped}, and DST_{veh}) or surrogate safety measures were calculated for the specified conflict events. The minimum, maximum, and average values of indicators for each location are listed in Table 4.3. The definitions and earlier research (Alhajyaseen, 2015; Kumar et al., 2019; Ni et al., 2016; Xin et al., 2021) indicate that the values of PET, TTV, and TTA are positively correlated with the severity of conflicts. The conflict becomes more intense as these values rise. While conflicts' severity and DST are negatively correlated. When DST values are high, the conflicts get worse.

At site 4, the average TTV and TTA values are lower, while the average DST_{ped} value is higher. However, the data show that site 3 has lower PET and higher DST_{veh}. Therefore, based on the individual values of the conflict indicators, it is impossible to define the levels of conflict severity. Consequently, cluster analysis was carried out in order to get a clear image of the different conflict severity levels.

Table 4.3. Estimated Values of Conflict Indicators at Each Site

Indicator	Type	Site 1	Site 2	Site 3	Site 4
PET	Minimum	0.20	0.36	0.32	0.40
	Average	1.31	1.48	1.28	1.31
	Maximum	4.76	4.98	4.64	4.60
TTV	Minimum	0.26	0.34	0.32	0.33
	Average	1.11	1.28	1.38	0.94
	Maximum	3.98	3.20	3.89	1.77
TTA	Minimum	0.47	0.28	0.57	0.24
	Average	1.80	1.59	1.30	1.28
	Maximum	4.60	3.70	3.93	2.90
DST_{ped}	Minimum	0.60	0.81	0.87	2.67
	Average	4.24	3.38	3.17	4.35
	Maximum	6.73	5.32	5.97	5.97
DST_{veh}	Minimum	0.47	0.63	0.84	0.87
	Average	2.66	2.84	3.30	3.29
	Maximum	5.77	5.37	5.38	5.41

4.4. Severity Level Classification of Conflicts

The severity levels of conflicts were classified using k-means clustering in SPSS. Hierarchical clustering on the data set was used to identify the appropriate number of clusters for k- means clustering. The optimum number of conflicts was decided based on the agglomeration plot as shown in Figure 4.3. By considering the elbow point in the graph, it was discovered that three clusters are sufficient. The detailed results of hierarchal clustering is given in Appendix A. All conflicts were divided into three clusters according to the values of the conflict indicators in the k-means clustering after the clustering centres had converged after nine iterations. From the final cluster centres, threshold values for each conflict indicator were obtained. The detailed results of the k-means clustering is given in Appendix B.

The threshold values for each indicator are shown in Table 4.4 along with their corresponding severity levels. The results of clustering indicated that severity level 1 represented the highly severe conflicts with the highest collision probability, severity level 2 indicated the severe conflicts, and severity level 3 indicated the normal conflicts with the lowest collision probability.

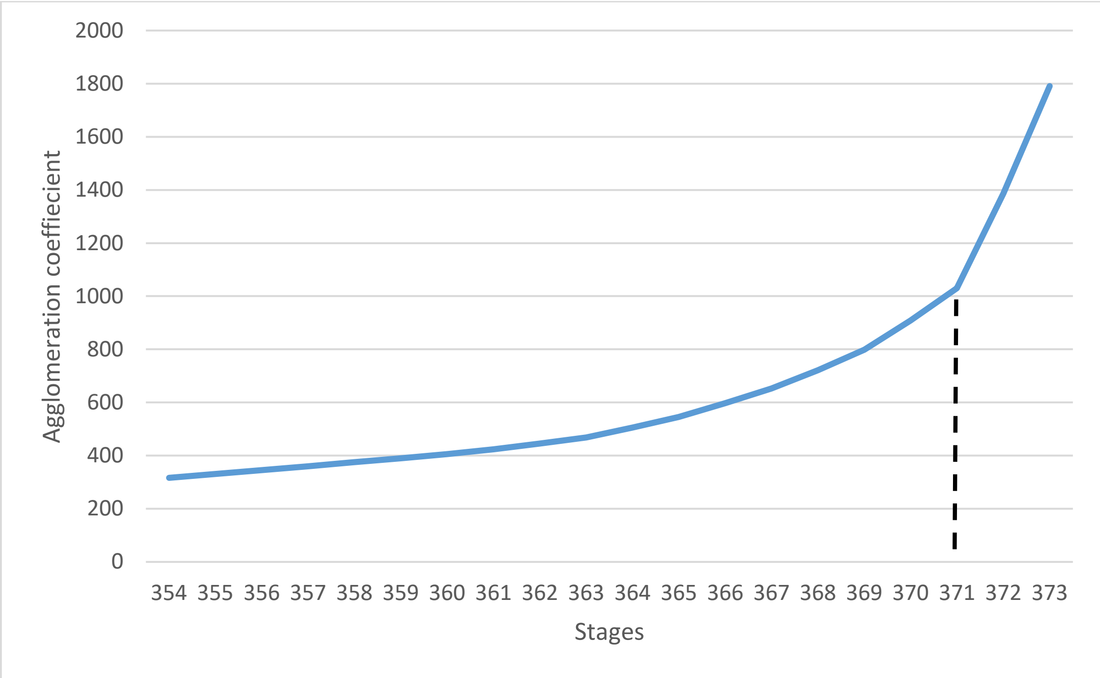


Figure 4.3. Agglomeration Coefficient Plot of Hierarchal Clustering

Table 4.4. Threshold Values of Conflict Indicators at Each Severity Level

Conflict indicators	Severity Level		
	1	2	3
PET	<0.88	0.88-2.19	>2.19
TTV	<1.08	1.08-1.31	>1.31
TTA	<1.28	1.28-1.90	>1.90
DST_{ped}	>3.99	3.99-3.56	<3.56
DST_{veh}	>3.48	3.48-2.42	<2.42

According to the findings, PET values below 0.88s signify highly serious conflicts, or dangerous circumstances at the time of crossing. The results of Ni et al. (2016), who determined that the PET threshold for severe conflict was 1 s, are in line with the obtained threshold value for highly severe conflicts. However, (Kumar et al., 2019) discovered 2.32s for highly serious conflicts. The result of 1.08s for the TTV threshold for extremely severe conflict reflects the finding of Kumar et al. (2019), who found that the value reached was 1.12s. In the current study, the TTA threshold limit in extremely serious conflict is 1.28. While the result achieved by Kumar et al. (2019) was 0.88s, which is considerably less than the current outcome. DST_{ped} threshold value for highly severe conflicts was 3.99 m/s^2, but Xin et al. (2021) reported 4.91s as the upper limit for severity level 1. When compared to the stated DST_{veh} threshold value of 3.48m/s^2, Kumar et al. (2019) obtained a substantially lower value (1.68 m/s^2) than the actual result.

In the current investigation, the upper threshold limits of typical conflicts (severity level 3) were 2.19s, 1.31s, 1.90s, 3.56m/s^2, and 2.42m/s^2 seconds, respectively, for PET, TTV, TTA, DST_{ped}, and DST_{veh}. Different upper limits for these indicators were reported by various researchers. According to Ni et al. (2016), the maximum PET threshold is 3s. Conflicts with PET values greater than 5s were not examined in the current investigation because the larger time intervals between these encounters are thought to be safe crossings (Almodfer et al., 2016; Chaudhari et al., 2021).

4.4.1. *Distribution of conflicts based on severity Level*

Based on the severity level determined by k-means clustering, the observed conflicts were categorised as highly severe, severe, and normal. The distribution of conflicts according to severity is shown in Figure 4.4.

75% of the conflicts fell under the serious category, while 13% of the conflicts were classified as being highly severe, as per the results. This implies that 88 % of all conflicts fell into the category of highly severe/severe conflicts with the highest likelihood of collision. Only 12% of the conflicts fell into the normal category. It was discovered that the majority of conflicts had high severity even at signalized junctions with pedestrian signals. This is undoubtedly a result of the intersection's pedestrians' signal-violating behaviour.

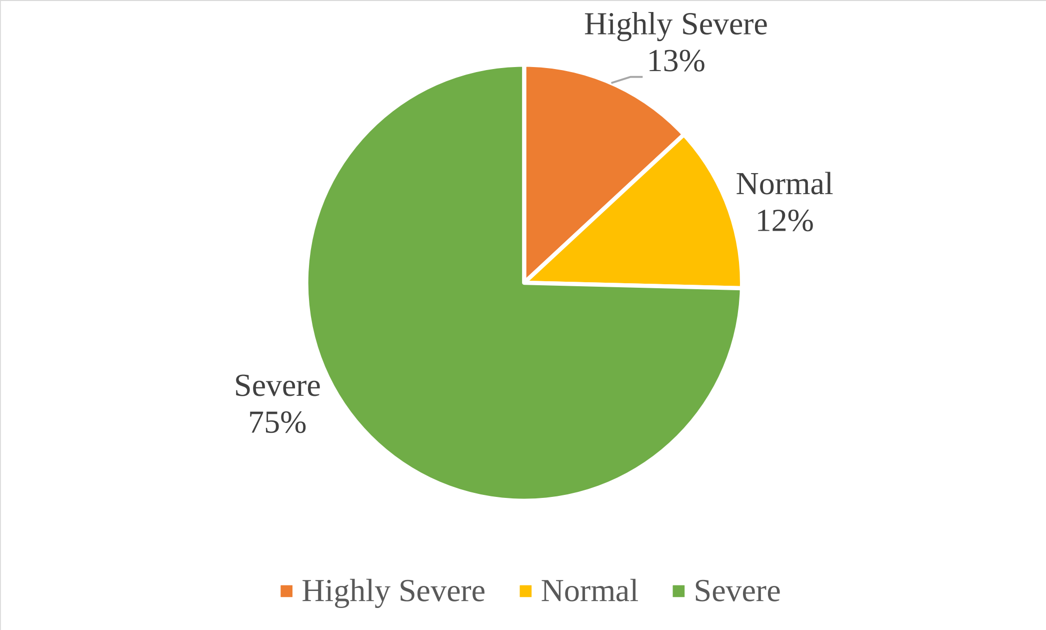


Figure 4.4. Distribution of Conflicts Based On Severity

4.4.2. *Distribution of highly severe conflicts*

The percentage of highly serious disputes brought on by pedestrian signal infractions and unrestricted left-turn movements that occur at each location is displayed in Table 4.5 below. Conflicts resulting from signal violations in the current study included those caused by right-turning vehicles as well as through-moving vehicles.

Table 4.5. Distribution of highly severe conflicts at each site

Site No.	Percentage of highly severe conflicts	
	Due to pedestrian signal violation (%)	Due to free left-turn movements (%)
1	44.44	55.56
2	58.33	41.67
3	66.67	33.33
4	70	30

Figure 4.5 depicts the distribution of total highly severe conflicts across every study sites. In accordance with the findings, 43% of highly severe conflicts are caused by free left-turns, and 57 % are caused by pedestrians violating traffic signals. By looking at the highly severe conflicts at each location, it can be seen that the majority of these

conflicts are caused by pedestrian signals being ignored. Only at site 1 do highly serious conflicts involving free left-turn movements outweigh those involving signal violations. The findings imply that pedestrian signal violations occur rather frequently at signalized intersections, which causes quite serious conflicts there. This might be as a result of the fact that disobeying pedestrians are exposed to conflicting traffic at junctions and that drivers fail to yield to them when it is green for vehicles to proceed. Pedestrians typically cross the road when there is a safe gap in traffic without taking into account the pedestrian signal indicator. Due to their errors in the perception of the distance, this may result in really serious conflicts. The main cause of the non-compliance behaviour is that pedestrian signal laws are not strictly enforced in India. Another potential cause of the non-compliance behaviour is the pedestrian's impatience in waiting for the signal. According to the research, installing pedestrian signals at junctions is not the sole way to lessen conflicts in Indian conditions because highly severe conflicts suggest a larger likelihood of collision. To increase pedestrian safety at intersections, preventative measures against signal infractions should be put into place. Although less frequent than conflicts resulting from signal infractions, conflicts caused by free left-turns nevertheless account for 43% of highly severe conflicts. The reason being that vehicles making a free left-turn must yield to both automobiles and pedestrians. They do not place more emphasis on pedestrians, which makes conflicts between vehicles and pedestrians more serious if the pedestrian is not careful.

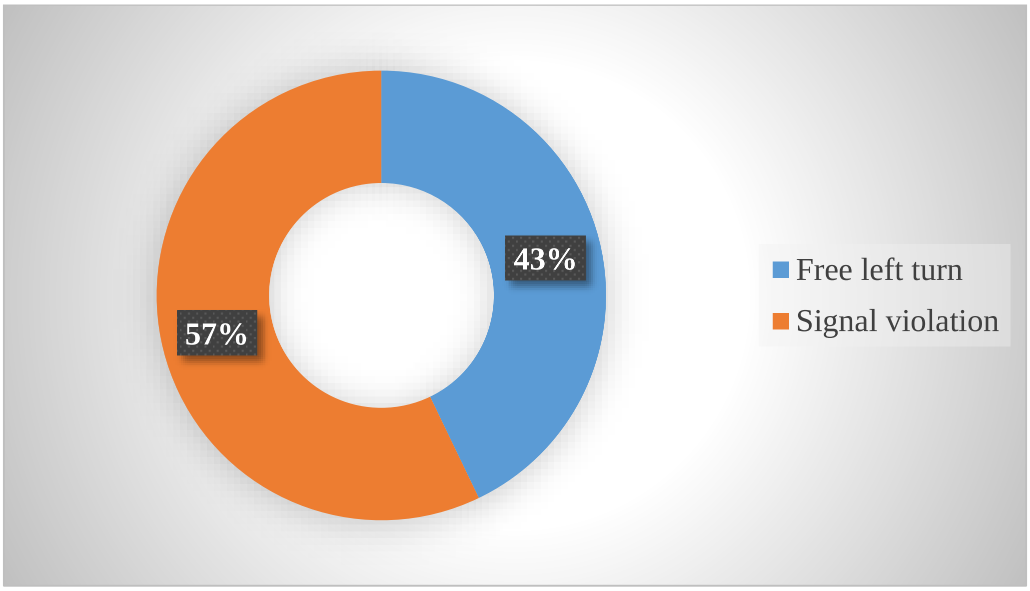


Figure 4.5. Distribution of highly severe conflicts

4.4.3. *Relation of conflicts with vehicle type*

This section emphasises how different vehicle types affect the severity and frequency of conflicts. It is a natural human propensity to act differently while approaching different types of vehicles, including two-wheelers (TW), cars, autos, light commercial vehicles (LCV), and heavy commercial vehicles (HCV), at intersections. Vehicles are also expected to operate differently during the entire engagement, in line with this expectation. The percentage of conflicts involving various vehicle types is shown in Figure 4.6. 33 % of the conflicts at the study locations are caused by cars, while 37% are caused by two-wheelers. Out of all the vehicle kinds, these two have the highest percentage.

The distribution of highly severe conflicts involving various types of vehicles that occurred at all of the study sites is depicted in Figure 4.7. The severity of the conflicts varied depending on the kind of vehicle. According to the statistics, a significant portion (47%) of highly serious conflicts involved two wheelers (TW). In comparison to interactions with other vehicle types, interactions between pedestrians and two-wheelers typically resulted in riskier situations. This might be as a result of two-wheelers travelling faster than other cars through intersections.

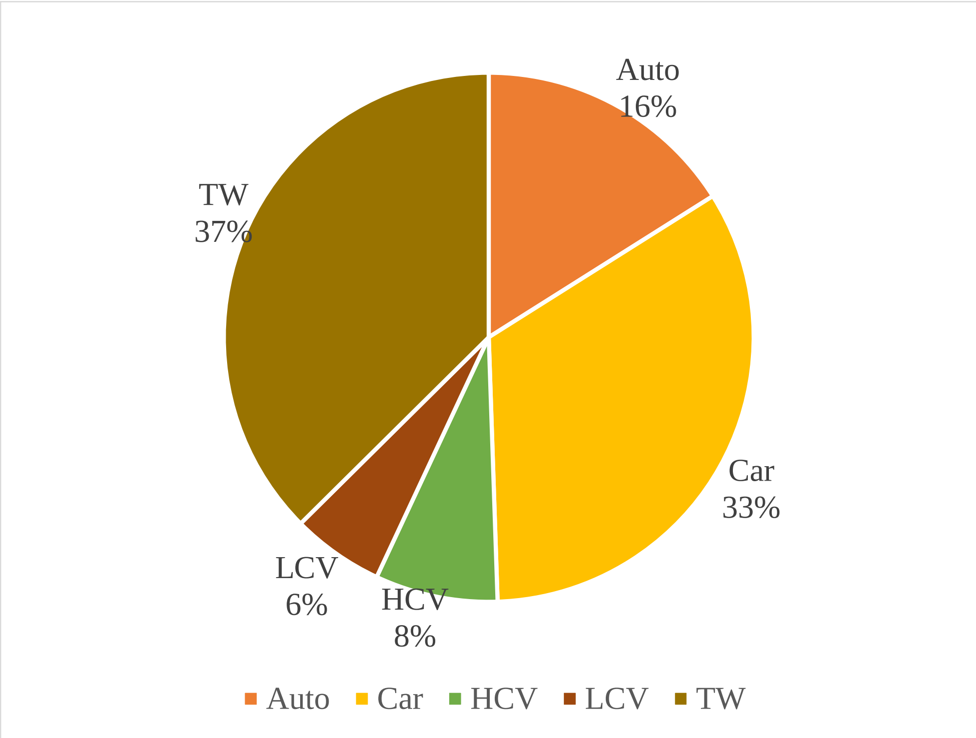


Figure 4.6. Frequency of Conflicts with Different Vehicle Types

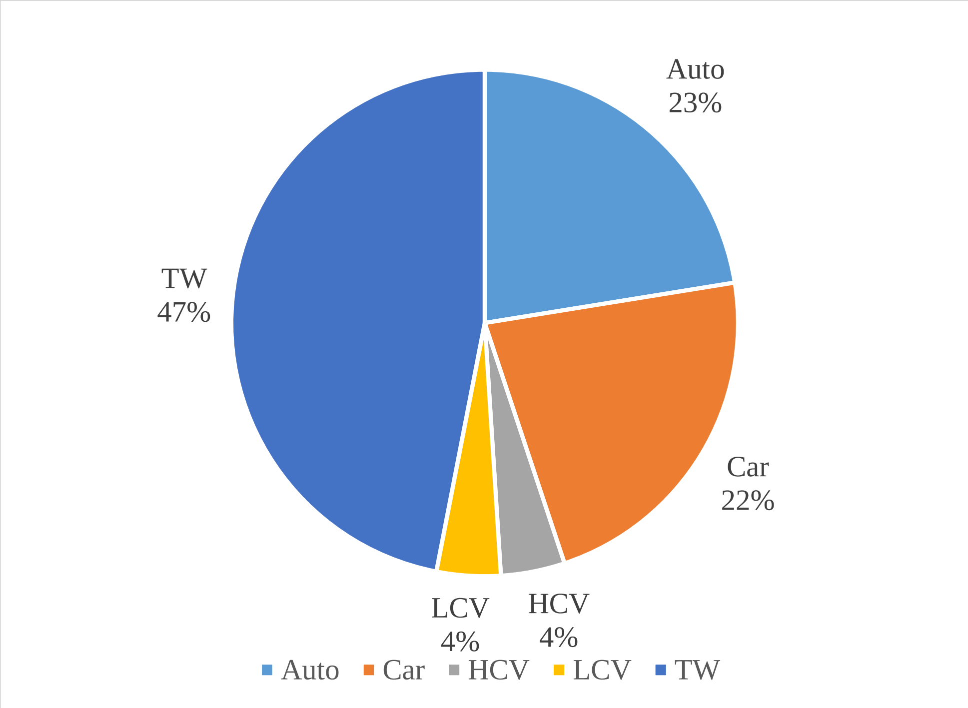


Figure 4.7. Distribution of Highly Severe Conflicts with Vehicle Types

4.5. Severity Prediction Model

Utilizing the support vector machine (SVM) approach, a severity prediction model was developed. All of the codes were implemented in Python using Google Colab. The python script for SVM is given in Appendix C. The estimated values of conflict indicators were used as the features or independent variables and the severity levels as target variable for developing the model. The SVM model was trained using the labelled data that emerged from k-means clustering. 80% of the conflicts were randomly chosen for training, and the remaining 20% were utilised as a test set to evaluate the performance of the modelling. Multi-collinearity was investigated prior to model calibration. There would be multi-collinearity among the predictors if the value of the variance inflation factor (VIF) is greater than or equal to 5 to 10 (Shrestha, 2020). There were no problems with multi-collinearity in any of the conflict indicator variables in the dataset. The results of the multi-collinearity test are given in Appendix D. Table 4.6 below displays the model's outcomes. To find the optimal performance model, different kernel functions like linear, radial basis function (RBF), and sigmoid were explored.

According to the findings, the SVM model with the RBF kernel is 97 % accurate in predicting the level of conflict that will occur at the intersection. Here, precision-recall is used to assess the classification quality in general, and the classification quality of each class in particular. The accuracy rate is a measure of how often the model can correctly identify exceedingly severe, severe, and normal conflicts. Recall shows the actual highly severe, severe, and normal conflicts that are correctly categorised in relation to all highly severe, severe, and normal conflicts, respectively.

Table 4.6. Results of Severity Prediction Model

Kernel	Precision			Recall			F1 score			Accuracy
	HS	S	N	HS	S	N	HS	S	N	
Linear	0.88	1.00	0.93	0.90	0.69	0.78	0.93	0.96	0.95	0.92
RBF	0.92	0.96	1.00	1.00	0.96	0.92	0.92	0.98	0.98	0.97
Sigmoid	0.00	0.00	0.75	0.00	0.00	1.00	0.00	0.00	0.85	0.75

HS: Highly severe, S: Severe, N: Normal

Additionally, by considering recall and precision, the F1 score can assess the classification's quality. SVM-RBF model provides the highest precision, recall, and F1 scores for classifying each category, making it the best model for predicting severity. When given conflict indicator values, this model can accurately forecast the amount of conflict severity. The degree of severity sheds light on the likelihood that a pedestrian and a vehicle would collide.

CHAPTER 5
CONCLUSION

5.1. General

The current investigation was carried out in Kollam, Kerala, at four specifically chosen signalised crossings with pedestrian signals. The interactions between pedestrians and automobiles were observed, and the conflict situations were identified manually from the recorded footage. The study concentrated on the conflicts caused by pedestrian signal violations and unrestricted left-turns at signalised intersections. The traffic conflict technique was used in the study to analyse the conflict scenario quantitatively. To evaluate how safe a situation might be in the event of a conflict, surrogate safety measures like PET, TTV, TTA, DST_{ped}, and DST_{veh} were calculated. These measures take into account the spatial and temporal proximity of conflicts. The indicators' maximum, minimum, and average values provided a general indication of the risk associated with pedestrian crossings. However, the actual risk assessment was correlated with conflict severity because conflict severity reflects the likelihood that a conflict may result in a collision. Based on these indicator values, k-means cluster analysis was used to categorise the conflicts into distinct severity categories. According to each severity level, the study suggested threshold values for the indicators. The study also developed an SVM severity prediction model by providing conflict indicator values to forecast conflict severity.

5.2. Key Findings

The major findings of the study are given below:

- The study revealed that a majority of pedestrians ignored the pedestrian signal at every study site, and 77% of pedestrians as a whole exhibited non-compliance behaviour when they crossed the road. This suggests that pedestrian signal violations at signalized intersections in India require further attention. Pedestrians cross the road when there is a safe gap by disobeying the signal because they are not prepared to wait for it.
- The study discovered that signal violation behaviour causes 65% of conflicts at intersections. The conflicts due to free left-turn movements also constitute 35% of total conflicts at the intersection.

- As the severity of the conflicts at the intersection cannot be determined manually, k-means clustering was used to categorise the conflicts into different severity levels based on the values of surrogate safety measures.
- According to the study, conflicts were classified into three categories: highly severe, severe, and normal conflicts.
- The study also suggested threshold values for the conflict indicators at each level of severity. A highly severe conflict with a high risk of collision is indicated by PET, TTV, and TTA values smaller than 0.88s, 1.08s, and 1.28s, respectively. DST_{ped} and DST_{veh} values above 3.99 m/s^2 and 3.48 m/s^2, respectively, also indicate the most severe conflicts.
- 88% of all identified conflicts fall into the highly severe or severe categories, according to the findings. This suggests that even with the installation of pedestrian signals, the safety of pedestrians at signalized junctions is undoubtedly a serious concern.
- Study results show that disobeying the pedestrian signal was to blame for 57% of highly severe conflicts at signalized intersections. The conflicts resulting from unrestricted left-turn movements account for 43% of the highly severe conflicts.
- Diverse vehicle types also resulted in differences in the frequency and severity of conflicts.
- The study found that two-wheelers were involved in a large number (47%) of highly severe conflicts. This may occur because two-wheelers are the majority of all vehicles at the intersection and also go at a faster rate than other types of vehicles.
- The study also calibrated a severity prediction model to predict the severity of conflicts that arise at intersections. The suggested SVM-RBF model obtains a total classification accuracy of 97%, demonstrating the model's superior predictive performance.

5.3. Recommendations

The riskiest places where complex interactions between pedestrians and vehicles occur are intersections. According to the study, installing pedestrian signals alone is not an excellent way to assure the safety of pedestrians at signalized intersections. It is essential to take steps to prevent pedestrians from violating the signal. Awareness programmes and online campaigns should be promoted to enhance pedestrian crossing behaviour. In India, most junctions lack intersection pedestrian signals with countdown timers. Installing countdown pedestrian signals would make it easier for people to cross the street safely by letting them know how long they have to finish the crossing. The lack of enforcement of laws and regulations regarding pedestrian offences will encourage pedestrians to disregard traffic signals. Hence, a penalty or fine should be imposed on pedestrians violating the signal. Crossing pedestrians are also put in danger by the unrestricted left-turns at crossings. Therefore, when developing the control mechanisms for left-turn movements, it is vital to consider the pedestrian demand at intersections.

5.4. Scope For Future Study

The current study manually recognised conflicts and determined conflict indicators. This manual data extraction process is time- and resource-intensive. Future research can therefore examine cutting-edge methods like trajectory data for analysing the interaction pattern between pedestrians and cars. Future research can include more intersections to generalise the findings. Further research can also examine the impact of intersection geometry on conflicts between pedestrians and vehicles.

REFERENCES

1. Aghabayk, K., Esmailpour, J., Jafari, A., & Shiwakoti, N. (2021). Observational-based study to explore pedestrian crossing behaviors at signalized and unsignalized crosswalks. *Accident Analysis and Prevention*, *151*(November 2020), 105990. https://doi.org/10.1016/j.aap.2021.105990
2. Alhajyaseen, W. K. M. (2015). The Integration of Conflict Probability and Severity for the Safety Assessment of Intersections. *Arabian Journal for Science and Engineering*, *40*(2), 421–430. https://doi.org/10.1007/s13369-014-1553-1
3. Alhajyaseen, W. K. M., Asano, M., & Nakamura, H. (2012). Estimation of left-turning vehicle maneuvers for the assessment of pedestrian safety at intersections. *IATSS Research*, *36*(1), 66–74. https://doi.org/10.1016/j.iatssr.2012.03.002
4. Alhajyaseen, W. K. M., Asano, M., & Nakamura, H. (2013). Left-turn gap acceptance models considering pedestrian movement characteristics. *Accident Analysis and Prevention*, *50*, 175–185. https://doi.org/10.1016/j.aap.2012.04.006
5. Alhajyaseen, W. K. M., & Iryo-Asano, M. (2017). Studying critical pedestrian behavioral changes for the safety assessment at signalized crosswalks. *Safety Science*, *91*, 351–360. https://doi.org/10.1016/j.ssci.2016.09.002
6. Almodfer, R., Xiong, S., Fang, Z., Kong, X., & Zheng, S. (2016). Quantitative analysis of lane-based pedestrian-vehicle conflict at a non-signalized marked crosswalk. *Transportation Research Part F: Traffic Psychology and Behaviour*, *42*, 468–478. https://doi.org/10.1016/j.trf.2015.07.004
7. Brosseau, M., Zangenehpour, S., Saunier, N., & Miranda-Moreno, L. (2013). The impact of waiting time and other factors on dangerous pedestrian crossings and violations at signalized intersections: A case study in Montreal. *Transportation Research Part F: Traffic Psychology and Behaviour*, *21*, 159–172. https://doi.org/10.1016/j.trf.2013.09.010
8. Chaudhari, A., Gore, N., Arkatkar, S., Joshi, G., & Pulugurtha, S. (2021). Exploring pedestrian surrogate safety measures by road geometry at midblock crosswalks: A perspective under mixed traffic conditions. *IATSS Research*, *45*(1), 87–101. https://doi.org/10.1016/j.iatssr.2020.06.001

9. Chen, P., Zeng, W., & Yu, G. (2019). Assessing right-turning vehicle-pedestrian conflicts at intersections using an integrated microscopic simulation model. *Accident Analysis and Prevention*, *129(May)*, 211–224. https://doi.org/10.1016/j.aap.2019.05.018
10. Chen, P., Zeng, W., Yu, G., & Wang, Y. (2017). Surrogate Safety Analysis of Pedestrian-Vehicle Conflict at Intersections Using Unmanned Aerial Vehicle Videos. *Journal of Advanced Transportation*, *2017*. https://doi.org/10.1155/2017/5202150
11. Hacohen, S., Shoval, S., & Shvalb, N. (2020). The paradox of pedestrian's risk aversion. *Accident Analysis and Prevention*, *142*(May 2019), 105518. https://doi.org/10.1016/j.aap.2020.105518
12. Han, D., Jung, J., & Kwon, S. (2020). Comparative Study on Supervised Learning Models for Productivity Forecasting of Shale Reservoirs Based on a Data-Driven Approach. *Applied sciences*
13. Haroun, A. (n.d.). Pedestrian Accidents With Left-Turning Traffic At Signalized Intersections : Characteristics , Human Factors And Unconsidered Issues University of Toronto. *5054*(416).
14. He, L., Agard, B., & Trépanier, M. (2018). Classification Of Public Transit Users With Smart Card Data Based On Time Series Distance Metrics And A Hierarchical Clustering Method. *Transportmetrica A : Transport Science A 9935*(May). https://doi.org/10.1080/23249935.2018.1479722
15. He, Y. L., Li, R. T., Li, L., Schwebel, D. C., Huang, H. L., Yin, Q. Y., & Hu, G. Q. (2019). Left-turning vehicle-pedestrian conflicts at signalized intersections with traffic lights: Benefit or harm? A two-stage study. *Chinese Journal of Traumatology - English Edition*, *22*(2), 63–68. https://doi.org/10.1016/j.cjtee.2018.07.007
16. Hussain, M., Wajid, S. K., Elzaart, A., Berbar, M., & Science, C. (2011). *A Comparison of SVM Kernel Functions for Breast Cancer Detection*. 1–6. https://doi.org/10.1109/CGIV.2011.31
17. Hussein, M., Sayed, T., Reyad, P., & Kim, L. (2015). Automated pedestrian safety analysis at a signalized intersection in New York city: Automated data extraction for safety diagnosis and behavioral study. *Transportation Research Record*, *2519*, 17–27. https://doi.org/10.3141/2519-03
18. Johnsson, C., Laureshyn, A., & De Ceunynck, T. (2018). In search of surrogate

safety indicators for vulnerable road users: a review of surrogate safety indicators. *Transport Reviews, 38(6), 765–785.* https://doi.org/10.1080/01441647.2018.1442888

19. Kadali, B. R., & Vedagiri, P. (2016). Proactive pedestrian safety evaluation at unprotected mid-block crosswalk locations under mixed traffic conditions. *Safety Science*, *89,* 94–105. https://doi.org/10.1016/j.ssci.2016.05.014
20. Kathuria, A., & Vedagiri, P. (2020). Evaluating pedestrian vehicle interaction dynamics at un-signalized intersections: A proactive approach for safety analysis. *Accident Analysis and Prevention*, *134*(August 2019), 105316. https://doi.org/10.1016/j.aap.2019.105316
21. Kim, D. (2019). The transportation safety of elderly pedestrians: Modeling contributing factors to elderly pedestrian collisions. *Accident Analysis and Prevention*, *131* (November2018), 268–274. https://doi.org/10.1016/j.aap.2019.07.009
22. Koh, P. P., Wong, Y. D., & Chandrasekar, P. (2014). Safety evaluation of pedestrian behaviour and violations at signalised pedestrian crossings. *Safety Science*, *70,* 143–152. https://doi.org/10.1016/j.ssci.2014.05.010
23. Kumar, A., & Ghosh, I. (2020). Analysis of Spontaneous Order of Pedestrian–Vehicle Conflicts at Signalized Intersections. *Transportation Research Record*, *2674*(11), 440–454. https://doi.org/10.1177/0361198120945994
24. Kumar, A., & Ghosh, I. (2022). Non-compliance behaviour of pedestrians and the associated conflicts at signalized intersections in India. *Safety Science*, *147*(September 2021), 105604. https://doi.org/10.1016/j.ssci.2021.105604
25. Kumar, A., Paul, M., & Ghosh, I. (2019). Analysis of Pedestrian Conflict with Right-Turning Vehicles at Signalized Intersections in India. *Journal of Transportation Engineering, Part A: Systems*, *145*(6), 04019018. https://doi.org/10.1061/jtepbs.0000239
26. Laureshyn, A., & Várhelyi, A. (2018). *The Swedish Traffic Conflict Technique – Observer's manual*.
27. Li, Z., Liu, P., Wang, W., & Xu, C. (2012). Using support vector machine models for crash injury severity analysis. *Accident Analysis and Prevention*, *45*, 478–486. https://doi.org/10.1016/j.aap.2011.08.016
28. Maji, A., Velaga, N. R., & Urie, Y. (2017). *Hierarchical clustering analysis framework of mutually exclusive crash causation parameters for regional road*

safetystrategies.*7300*(December). https://doi.org/10.1080/17457300.2017.1416485

29. Marisamynathan, & Perumal, V. (2014). Study on pedestrian crossing behavior at signalized intersections. *Journal of Traffic and Transportation Engineering (English Edition)*, *1*(2), 103–110. https://doi.org/10.1016/S2095-7564(15)30094-5
30. Marisamynathan, S., & Vedagiri, P. (2020). Pedestrian Safety Evaluation Of Signalized Intersections Using Surrogate Safety Measures. *35*(1), 48–56.
31. Mukherjee, D., & Mitra, S. (2019). A comparative study of safe and unsafe signalized intersections from the view point of pedestrian behavior and perception. *Accident Analysis and Prevention*, *132*(April), 105218. https://doi.org/10.1016/j.aap.2019.06.010
32. Ni, Y., Wang, M., Sun, J., & Li, K. (2016). Evaluation of pedestrian safety at intersections: A theoretical framework based on pedestrian-vehicle interaction patterns. *Accident Analysis and Prevention*, *96*, 118–129. https://doi.org/10.1016/j.aap.2016.07.030
33. Orhan, U., Hekim, M., & Ozer, M. (2011). EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. *Expert Systems with Applications*, *38*(10), 13475–13481. https://doi.org/10.1016/j.eswa.2011.04.149
34. Pawar, N. M., Gore, N., & Arkatkar, S. (2022). Examining Crossing Conflicts by Vehicle Type at Unsignalized T-Intersections Using Accepted Gaps: A Perspective from Emerging Countries. *Journal of Transportation Engineering, Part A: Systems*, *148*(6), 1–15. https://doi.org/10.1061/jtepbs.0000665
35. Qi, Y., & Guoguo, A. (2017). Pedestrian safety under permissive left-turn signal control. *International Journal of Transportation Science and Technology*, *6*(1), 53–62. https://doi.org/10.1016/j.ijtst.2017.05.002
36. Saha, R., Tariq, M. T., Hadi, M., & Xiao, Y. (2019). Pattern Recognition Using Clustering Analysis to Support Transportation System Management, Operations, and Modeling. *Journal of Advanced Transportation*, *2019*. https://doi.org/10.1155/2019/1628417
37. Shah, H., & Vedagiri, P. (2017). Evaluation of Pedestrian Safety at Unsignalised Intersection under Mix Traffic Condition Using Surrogate Safety Measures. *15th International Conference on Computers in Urban Planning and*

Urban Management (CUPUM), July 2017.

38. Shekhar Babu, S., & Vedagiri, P. (2018). Proactive safety evaluation of a multilane unsignalized intersection using surrogate measures. *Transportation Letters*, *10*(2), 104–112. https://doi.org/10.1080/19427867.2016.1230172
39. Shrestha, N. (2020). Detecting Multicollinearity in Regression Analysis. *American Journal of Applied Mathematics and Statistics*, *8*(2), 39–42. https://doi.org/10.12691/ajams-8-2-1
40. Society, E. A., & Studies, T. (2013). *Binary Probabilistic Models for Pedestrians' Crossing Behaviour and Risk at the Free Left Turn: Delhi, India Mariya KHATOON*. 1–12.
41. Tiwari, G., Bangdiwala, S., Saraswat, A., & Gaurav, S. (2007). *Survival analysis : Pedestrian risk exposure at signalized intersections*. *10*, 77–89. https://doi.org/10.1016/j.trf.2006.06.002
42. Xin, F., Wang, X., & Sun, C. (2021). Risk evaluation for conflicts between crossing pedestrians and right-turning vehicles at intersections. *Transportation Research Record*, *2675*(12), 1005–1014. https://doi.org/10.1177/03611981211031885
43. Yu, R., & Abdel-aty, M. (2013). Utilizing support vector machine in real-time crash risk evaluation. *Accident Analysis and Prevention*, *51*, 252–259. https://doi.org/10.1016/j.aap.2012.11.027
44. Zhang, S., Abdel-Aty, M., Cai, Q., Li, P., & Ugan, J. (2020). Prediction of pedestrian-vehicle conflicts at signalized intersections based on long short-term memory neural network. *Accident Analysis and Prevention*, *148*(March), 105799. https://doi.org/10.1016/j.aap.2020.105799

APPENDIX A: RESULTS OF HIERARICHAL CLUSTERING

Case Processing Summary[a,b]

Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
374	100.0	0	.0	374	100.0

a. Squared Euclidean Distance used

b. Ward Linkage

Ward Linkage

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	223	318	.006	0	0	32
2	272	273	.014	0	0	81
3	309	310	.026	0	0	53
4	243	244	.038	0	0	42
5	258	259	.056	0	0	85
6	220	222	.077	0	0	32
7	189	190	.099	0	0	194
8	191	192	.121	0	0	52
9	187	188	.143	0	0	69
10	195	197	.166	0	0	139
11	206	207	.190	0	0	125
12	63	64	.215	0	0	133
13	347	348	.241	0	0	115
14	270	271	.268	0	0	89
15	90	91	.297	0	0	184
16	164	165	.326	0	0	43
17	231	311	.356	0	0	37
18	325	326	.386	0	0	123
19	228	306	.416	0	0	56
20	249	251	.447	0	0	84
21	23	24	.480	0	0	241

22	236	237	.513	0	0	94
23	129	130	.546	0	0	127
24	208	209	.583	0	0	232
25	331	332	.620	0	0	192
26	44	45	.657	0	0	215
27	146	304	.696	0	0	79
28	13	108	.734	0	0	181
29	155	368	.775	0	0	207
30	8	9	.817	0	0	60
31	136	137	.860	0	0	170
32	220	223	.903	6	1	117
33	100	101	.947	0	0	124
34	365	367	.992	0	0	140
35	202	203	1.038	0	0	147
36	211	212	1.084	0	0	84
37	231	313	1.131	17	0	209
38	3	205	1.179	0	0	112
39	128	359	1.229	0	0	119
40	229	242	1.279	0	0	176
41	113	300	1.330	0	0	190
42	243	245	1.383	4	0	188
43	164	166	1.438	16	0	252
44	225	226	1.492	0	0	162
45	103	107	1.548	0	0	198
46	283	284	1.603	0	0	150
47	247	252	1.660	0	0	258
48	261	262	1.716	0	0	156
49	72	255	1.773	0	0	201
50	277	279	1.830	0	0	169
51	147	301	1.888	0	0	107
52	160	191	1.946	0	8	280
53	308	309	2.004	0	3	130
54	183	184	2.066	0	0	77
55	98	99	2.129	0	0	182
56	228	230	2.197	19	0	87
57	312	314	2.266	0	0	130
58	7	163	2.338	0	0	82
59	150	217	2.411	0	0	98
60	8	94	2.484	30	0	159
61	69	181	2.558	0	0	237

62	26	27	2.634	0	0	242
63	65	66	2.711	0	0	261
64	162	167	2.788	0	0	196
65	234	235	2.866	0	0	97
66	227	305	2.944	0	0	113
67	88	89	3.024	0	0	184
68	11	12	3.104	0	0	235
69	46	187	3.184	0	9	194
70	280	281	3.266	0	0	204
71	329	330	3.348	0	0	206
72	198	201	3.430	0	0	211
73	47	48	3.513	0	0	215
74	182	363	3.597	0	0	197
75	353	354	3.682	0	0	278
76	95	168	3.767	0	0	138
77	183	185	3.853	54	0	247
78	144	145	3.941	0	0	135
79	146	303	4.030	27	0	107
80	158	159	4.122	0	0	238
81	272	275	4.214	2	0	153
82	7	105	4.312	58	0	159
83	175	176	4.410	0	0	212
84	211	249	4.508	36	20	232
85	258	260	4.607	5	0	186
86	215	219	4.708	0	0	233
87	228	241	4.809	56	0	209
88	366	374	4.913	0	0	227
89	270	274	5.022	14	0	153
90	157	340	5.130	0	0	238
91	335	369	5.241	0	0	134
92	140	141	5.351	0	0	216
93	324	351	5.462	0	0	148
94	236	238	5.575	22	0	171
95	1	142	5.690	0	0	226
96	194	373	5.808	0	0	141
97	233	234	5.926	0	65	265
98	150	151	6.046	59	0	177
99	264	265	6.167	0	0	203
100	50	102	6.289	0	0	181
101	204	299	6.412	0	0	125

102	76	121	6.537	0	0	225
103	170	171	6.663	0	0	302
104	32	33	6.789	0	0	243
105	106	134	6.917	0	0	202
106	276	278	7.045	0	0	260
107	146	147	7.173	79	51	268
108	84	267	7.306	0	0	126
109	80	288	7.443	0	0	281
110	132	372	7.580	0	0	124
111	178	218	7.718	0	0	154
112	2	3	7.859	0	38	147
113	227	321	8.000	66	0	231
114	82	269	8.141	0	0	195
115	347	349	8.285	13	0	193
116	92	93	8.429	0	0	155
117	220	221	8.574	32	0	259
118	196	200	8.720	0	0	289
119	97	128	8.868	0	39	221
120	186	370	9.017	0	0	197
121	336	337	9.171	0	0	320
122	49	169	9.325	0	0	138
123	325	350	9.486	18	0	246
124	100	132	9.647	33	110	249
125	204	206	9.808	101	11	264
126	84	266	9.970	108	0	309
127	109	129	10.135	0	23	163
128	333	334	10.303	0	0	285
129	115	116	10.471	0	0	254
130	308	312	10.641	53	57	227
131	343	344	10.812	0	0	183
132	59	60	10.992	0	0	262
133	63	67	11.172	12	0	230
134	153	335	11.353	0	91	192
135	144	154	11.536	78	0	185
136	41	42	11.721	0	0	236
137	83	323	11.909	0	0	195
138	49	95	12.101	122	76	211
139	193	195	12.295	0	10	239
140	177	365	12.492	0	34	164
141	194	371	12.691	96	0	263

142	25	224	12.889	0	0	268
143	254	256	13.087	0	0	214
144	253	257	13.286	0	0	186
145	73	357	13.486	0	0	191
146	87	123	13.686	0	0	299
147	2	202	13.887	112	35	189
148	110	324	14.090	0	93	246
149	180	341	14.296	0	0	237
150	282	283	14.512	0	46	204
151	328	358	14.733	0	0	206
152	53	56	14.954	0	0	210
153	270	272	15.175	89	81	336
154	178	250	15.396	111	0	229
155	29	92	15.618	0	116	288
156	261	263	15.845	48	0	264
157	68	179	16.074	0	0	242
158	20	22	16.304	0	0	304
159	7	8	16.536	82	60	255
160	338	339	16.769	0	0	234
161	17	19	17.003	0	0	257
162	225	316	17.238	44	0	213
163	109	131	17.477	127	0	278
164	177	322	17.717	140	0	207
165	161	298	17.958	0	0	196
166	214	317	18.201	0	0	259
167	86	139	18.446	0	0	245
168	30	62	18.694	0	0	230
169	277	346	18.941	50	0	282
170	135	136	19.190	0	31	256
171	236	239	19.439	94	0	265
172	173	327	19.689	0	0	277
173	78	79	19.939	0	0	281
174	291	293	20.190	0	0	322
175	112	143	20.445	0	0	226
176	229	232	20.703	40	0	240
177	150	216	20.960	98	0	275
178	39	40	21.223	0	0	308
179	36	37	21.486	0	0	250
180	210	248	21.750	0	0	218
181	13	50	22.017	28	100	249

182	96	98	22.286	0	55	266
183	70	343	22.556	0	131	253
184	88	90	22.826	67	15	288
185	144	152	23.105	135	0	297
186	253	258	23.384	144	85	323
187	4	5	23.665	0	0	303
188	243	307	23.946	42	0	231
189	2	28	24.233	147	0	228
190	113	148	24.532	41	0	297
191	73	74	24.837	145	0	222
192	153	331	25.143	134	25	273
193	111	347	25.450	0	115	308
194	46	189	25.758	69	7	280
195	82	83	26.072	114	137	287
196	161	162	26.387	165	64	252
197	182	186	26.706	74	120	247
198	15	103	27.033	0	45	286
199	117	118	27.361	0	0	318
200	71	285	27.692	0	0	295
201	72	174	28.024	49	0	253
202	106	297	28.357	105	0	292
203	75	264	28.690	0	99	321
204	280	282	29.036	70	150	282
205	6	58	29.382	0	0	220
206	328	329	29.733	151	71	273
207	155	177	30.090	29	164	300
208	156	364	30.449	0	0	315
209	228	231	30.812	87	37	301
210	53	55	31.178	152	0	267
211	49	198	31.546	138	72	299
212	81	175	31.937	0	83	276
213	225	302	32.329	162	0	258
214	254	345	32.727	143	0	295
215	44	47	33.128	26	73	239
216	138	140	33.532	0	92	290
217	352	356	33.937	0	0	251
218	210	213	34.353	180	0	298
219	14	104	34.772	0	0	272
220	6	51	35.190	205	0	311
221	97	360	35.613	119	0	313

222	73	342	36.047	191	0	309
223	125	127	36.482	0	0	274
224	43	199	36.923	0	0	263
225	76	120	37.366	102	0	271
226	1	112	37.810	95	175	277
227	308	366	38.257	130	88	291
228	2	122	38.706	189	0	314
229	114	178	39.159	0	154	283
230	30	63	39.614	168	133	261
231	227	243	40.069	113	188	291
232	208	211	40.525	24	84	271
233	149	215	40.992	0	86	275
234	320	338	41.463	0	160	279
235	10	11	41.937	0	68	255
236	41	172	42.412	136	0	310
237	69	180	42.895	61	149	300
238	157	158	43.379	90	80	279
239	44	193	43.870	215	139	314
240	229	315	44.368	176	0	301
241	23	52	44.873	21	0	311
242	26	68	45.380	62	157	319
243	31	32	45.903	0	104	262
244	361	362	46.444	0	0	269
245	85	86	46.996	0	167	290
246	110	325	47.552	148	123	310
247	182	183	48.113	197	77	328
248	240	319	48.691	0	0	276
249	13	100	49.297	181	124	286
250	35	36	49.926	0	179	305
251	268	352	50.587	0	217	332
252	161	164	51.252	196	43	307
253	70	72	51.918	183	201	296
254	115	119	52.599	129	0	318
255	7	10	53.294	159	235	292
256	124	135	53.994	0	170	306
257	17	21	54.700	161	0	304
258	225	247	55.416	213	47	298
259	214	220	56.169	166	117	283
260	276	289	56.923	106	0	287
261	30	65	57.684	230	63	339

262	31	59	58.516	243	132	312
263	43	194	59.355	224	141	302
264	204	261	60.229	125	156	319
265	233	236	61.106	97	171	352
266	96	296	62.011	182	0	306
267	53	57	62.935	210	0	324
268	25	146	63.860	142	107	326
269	355	361	64.851	0	244	338
270	290	295	65.859	0	0	316
271	76	208	66.869	225	232	326
272	14	18	67.880	219	0	348
273	153	328	68.900	192	206	328
274	125	126	69.938	223	0	350
275	149	150	71.012	233	177	320
276	81	240	72.090	212	248	335
277	1	173	73.175	226	172	285
278	109	353	74.282	163	75	313
279	157	320	75.393	238	234	330
280	46	160	76.510	194	52	317
281	78	80	77.630	173	109	294
282	277	280	78.758	169	204	296
283	114	214	79.886	229	259	333
284	16	133	81.033	0	0	329
285	1	333	82.192	277	128	321
286	13	15	83.352	249	198	325
287	82	276	84.527	195	260	327
288	29	88	85.709	155	184	331
289	61	196	86.907	0	118	329
290	85	138	88.107	245	216	346
291	227	308	89.353	231	227	337
292	7	106	90.622	255	202	331
293	77	286	91.897	0	0	361
294	78	287	93.198	281	0	354
295	71	254	94.545	200	214	347
296	70	277	95.930	253	282	335
297	113	144	97.314	190	185	317
298	210	225	98.714	218	258	345
299	49	87	100.144	211	146	307
300	69	155	101.600	237	207	340
301	228	229	103.130	209	240	337

302	43	170	104.730	263	103	341
303	4	54	106.337	187	0	348
304	17	20	107.948	257	158	344
305	34	35	109.645	0	250	339
306	96	124	111.360	266	256	325
307	49	161	113.097	299	252	349
308	39	111	114.896	178	193	327
309	73	84	116.715	222	126	334
310	41	110	118.652	236	246	336
311	6	23	120.612	220	241	324
312	31	38	122.585	262	0	342
313	97	109	124.593	221	278	338
314	2	44	126.687	228	239	323
315	156	246	128.796	208	0	341
316	290	294	130.909	270	0	322
317	46	113	133.126	280	297	330
318	115	117	135.379	254	199	345
319	26	204	137.638	242	264	343
320	149	336	139.939	275	121	333
321	1	75	142.420	285	203	340
322	290	291	144.990	316	174	358
323	2	253	147.614	314	186	349
324	6	53	150.289	311	267	344
325	13	96	153.091	286	306	357
326	25	76	155.986	268	271	351
327	39	82	158.891	308	287	332
328	153	182	161.839	273	247	334
329	16	61	165.069	284	289	355
330	46	157	168.387	317	279	343
331	7	29	171.708	292	288	359
332	39	268	175.341	327	251	347
333	114	149	179.363	283	320	351
334	73	153	183.657	309	328	346
335	70	81	187.974	296	276	354
336	41	270	192.515	310	153	363
337	227	228	197.373	291	301	352
338	97	355	202.409	313	269	357
339	30	34	207.641	261	305	342
340	1	69	213.168	321	300	353
341	43	156	218.840	302	315	362

342	30	31	224.546	339	312	359
343	26	46	230.268	319	330	360
344	6	17	235.994	324	304	364
345	115	210	241.852	318	298	356
346	73	85	247.912	334	290	353
347	39	71	255.209	332	295	361
348	4	14	262.691	303	272	355
349	2	49	270.176	323	307	360
350	125	292	278.420	274	0	358
351	25	114	287.338	326	333	356
352	227	233	296.309	337	265	365
353	1	73	305.622	340	346	365
354	70	78	315.756	335	294	363
355	4	16	330.341	348	329	369
356	25	115	345.071	351	345	370
357	13	97	359.839	325	338	367
358	125	290	374.685	350	322	367
359	7	30	390.137	331	342	364
360	2	26	405.884	349	343	362
361	39	77	423.057	347	293	366
362	2	43	445.307	360	341	368
363	41	70	468.026	336	354	366
364	6	7	504.991	344	359	371
365	1	227	545.022	353	352	368
366	39	41	597.317	361	363	372
367	13	125	653.202	357	358	369
368	1	2	721.590	365	362	370
369	4	13	798.943	355	367	371
370	1	25	908.693	368	356	372
371	4	6	1030.624	369	364	373
372	1	39	1385.609	370	366	373
373	1	4	1790.569	372	371	0

APPENDIX B: RESULTS OF K-MEANS CLUSTERING

Initial Cluster Centers

	Cluster		
	1	2	3
PET	1.04	.40	4.76
TTV	2.30	.26	1.20
TTA	1.57	1.57	4.60
DTP	.81	6.73	3.72
DTV	2.93	2.93	.47

Iteration History[a]

	Change in Cluster Centers		
Iteration	1	2	3
1	2.972	2.453	2.621
2	.116	.425	.302
3	.061	.204	.172
4	.030	.058	.155
5	.022	.020	.114
6	.019	.065	.062
7	.009	.049	.000
8	.009	.054	.000
9	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 9. The minimum distance between initial centers is 6.229.

Final Cluster Centers

	Cluster		
	1	2	3
PET	1.17	.58	3.22
TTV	1.21	.95	1.40
TTA	1.55	1.01	2.26
DTP	3.65	4.34	3.47
DTV	2.89	4.08	1.95

APPENDIX C: PYTHON SCRIPT OF SVM

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
cell_df=pd.read_excel("/content/Untitled2 (datasheet) - SVM.xlsx")
normal_df= cell_df[cell_df["Conflict"]=="N"] [0:300]
severe_df=cell_df[cell_df["Conflict"]=="S"] [0:300]
highlysevere_df=cell_df[cell_df["Conflict"]=="HS"] [0:300]
feature_df=cell_df[['PET', 'TTV', 'TTA', 'DTP', 'DTV']]
X=np.asarray(feature_df)
y=np.asarray(cell_df["Conflict"])
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test=train_test_split(X,y,test_size=0.2)
from sklearn import svm
from sklearn.svm import SVC
kernels = ['RBF', 'Sigmoid','Linear']#A function which returns the corresponding SVC model
def getClassifier(ktype):
     if ktype == 1:
        # Radial Basis Function kernal
        return SVC(kernel='rbf', gamma="auto")
    elif ktype == 2:
        # Sigmoid kernal
        return SVC(kernel='sigmoid', gamma="auto")
    elif ktype == 3:
        # Linear kernal
        return SVC(kernel='linear', gamma="auto")
for i in range(4):
    # Separate data into test and training sets
    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2)# Train a SVC model using different kernal
    svclassifier = getClassifier(i)
    svclassifier.fit(X_train, y_train)# Make prediction
    y_pred = svclassifier.predict(X_test)# Evaluate our model
    print("Evaluation:", kernels[i], "kernel")
    from sklearn.metrics import classification_report
    print(classification_report(y_test,y_pred))
```

APPENDIX D: RESULTS OF MULTI-COLLINEARITY TEST

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.660[a]	.436	.428	.525

a. Predictors: (Constant), DTV_1, DTP_1, PET, TTV_1, TTA_1

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	78.330	5	15.666	56.792	.000[b]
	Residual	101.512	368	.276		
	Total	179.842	373			

a. Dependent Variable: Cluster Number of Case

b. Predictors: (Constant), DTV_1, DTP_1, PET, TTV_1, TTA_1

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-3.564	.630		-5.658	.000		
	PET	.486	.037	.627	13.095	.000	.670	1.494
	TTV	.488	.158	.243	3.101	.002	.251	3.989
	TTA	.540	.124	.369	4.354	.000	.213	4.689
	DTP	.452	.086	.404	5.262	.000	.260	3.852
	DTV	.400	.075	.440	5.337	.000	.226	4.426

a. Dependent Variable: Cluster Number of Case

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions (Constant)	PET	TTV_1	TTA_1	DTP_1	DTV_1
1	1	5.504	1.000	.00	.01	.00	.00	.00	.00
	2	.310	4.211	.00	.43	.00	.00	.00	.01
	3	.107	7.156	.00	.08	.06	.05	.01	.02
	4	.070	8.865	.00	.48	.10	.05	.01	.03
	5	.007	27.770	.00	.00	.27	.58	.34	.62
	6	.001	61.901	1.00	.00	.57	.31	.64	.32

a. Dependent Variable: Cluster Number of Case