FINAL REPORT

August 2023

A Comparative Study of Pedestrian Crossing Behavior and Safety in Baltimore, MD and Washington, DC Using Video Surveillance

Prepared for:

Urban Mobility & Equity Center
Morgan State University, CBEIS 327
1700 E. Cold Spring Lane, Baltimore, MD 21251

Principal Investigators:

Celeste Chavis, Ph.D., PE
Transportation & Urban Infrastructure Studies
Morgan State University
(443) 885-5061
celeste.chavis@morgan.edu

Kofi Nyarko, Ph.D.
Electrical & Computer Engineering
Morgan State University
(443) 885-3476
kofi.nyarko@morgan.edu

Cinzia Cirillo, Ph.D.
Civil Engineering
University of Maryland, College Park
(301) 405-6864
ccirillo@umd.edu

Contributors

The following graduate students provided significant contributions to this study:

- **Tasmeer Alam**, Doctoral Student, Morgan State University
- **Istiak Bhuyan**, Doctoral Student, Morgan State University
- **Md Mahmudul Huque**, Doctoral Student, University of Maryland, College Park
- **Daniel Stephens**, Masters Student, Morgan State University

Acknowledgments

This research was supported by the Urban Mobility & Equity Center at Morgan State University and the University Transportation Center Program of the U.S. Department of Transportation.

Special thanks to Omar Stevenson, Akim Mahadiow, and Martin Ndegwa of Cube Root Corporation for video collection and data.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

<table>
<tr><td>1. Report No.</td><td>2. Government Accession No.</td><td colspan="2">3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
A Comparative Study of Pedestrian Crossing Behavior and Safety in Baltimore, MD and Washington, DC Using Video Surveillance</td><td colspan="2">5. Report Date
August 2023</td></tr>
<tr><td colspan="2">6. Performing Organization Code</td></tr>
<tr><td colspan="2">7. Author(s)
Celeste Chavis, Kofi Nyarko, and Cinzia Cirillo</td><td colspan="2">8. Performing Organization Report No.</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
Morgan State University
1700 E. Cold Spring Lane
Baltimore, MD 21251</td><td colspan="2">10. Work Unit No.</td></tr>
<tr><td colspan="2">11. Contract or Grant No.
69A43551747123</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Organization Name and Address
US Department of Transportation
Office of the Secretary-Research
UTC Program, RDT-30
1200 New Jersey Ave., SE
Washington, DC 20590</td><td colspan="2">13. Type of Report and Period Covered</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code</td></tr>
<tr><td colspan="4">15. Supplementary Notes</td></tr>
<tr><td colspan="4">16. Abstract
Pedestrian safety is of paramount importance in urban transportation, and it is a key goal of initiatives such as Vision Zero. Washington, D.C., and Baltimore, Maryland, have experienced a significant number of traffic accidents in recent years, with pedestrians being particularly vulnerable. Video surveillance has long been established as a valuable method for analyzing pedestrian behavior. However, traditional manual analysis of video footage is costly, time-consuming, and prone to human errors. With the advancement of computer vision and machine learning technologies, such as the YOLO (You Only Look Once) algorithm, it has become possible to automate and streamline the analysis process. In this study, we applied the YOLOv8 algorithm to analyze video surveillance footage, allowing for efficient extraction of pedestrian data and the development of analytics for pedestrian behavior at signalized intersections. By leveraging computer vision and machine learning tools, we were able to process large volumes of video footage and obtain detailed insights into pedestrian behavior at intersections.
The application of computer vision and machine learning techniques, specifically the YOLOv8 algorithm, to analyze video surveillance footage has proven to be a valuable approach for studying pedestrian behavior at intersections in Washington, D.C., and Baltimore, Maryland. This preliminary study presents a method for tracking pedestrians and compares pedestrian volume and speed across five intersections.
By automating the analysis process, this study has provided comprehensive insights for using video footage to track pedestrian movements, contributing to the broader goal of improving pedestrian safety in urban environments. Future research endeavors should develop additional algorithms to improve pedestrian tracking and provide additional insights to vehicle and pedestrian interaction to enhance pedestrian safety.</td></tr>
<tr><td colspan="2">17. Key Words
pedestrian, machine learning, safety</td><td colspan="2">18. Distribution Statement
No restrictions</td></tr>
<tr><td>19. Security Classif. (of report)
Unclassified</td><td>20. Security Classif. (of this page)
Unclassified</td><td>21. No. of Pages
38</td><td>22. Price</td></tr>
</table>

Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation

Walking is the oldest form of transportation. Most trips, regardless of mode, begin and end with a walking component. Pedestrians represent the most vulnerable road users. Pedestrian safety is a paramount concern in modern urban transportation systems. As cities integrate technology to improve safety and operations in cities, there is an opportunity to understand pedestrian behavior better. Using video surveillance to determine the microscopic behavior of pedestrians along a corridor, this study develops a methodology for identifying, tracking, and classifying vehicles and pedestrians along roadway segments to evaluate pedestrian behavior and safety.

Jaywalking, or the act of crossing the street illegally outside designated crosswalks, is a common practice that poses significant dangers to pedestrians. Jaywalking disrupts the expected flow of traffic, making it challenging for drivers to anticipate pedestrian movements and potentially leading to collisions. Analyzing pedestrian crossing behavior is crucial for developing effective strategies to enhance pedestrian safety. By investigating factors such as crossing times, crossing speed, and compliance with traffic signals, transportation engineers and urban planners can gain insights into the critical challenges pedestrians face at intersections. Identifying these patterns can guide evidence-based decision-making in designing pedestrian-friendly infrastructure and implementing targeted safety interventions, ultimately reducing the number of accidents and enhancing overall road safety.

Every road user is a pedestrian at a certain point in their travel. Unfortunately, there were 6,205 pedestrian fatalities and approximately 76,000 pedestrian injuries nationwide in 2019 [1]. Among all traffic-related deaths, 17% accounted for pedestrians occurring mostly in urbanized areas (74%) from 6:00 pm to 8:59 pm. Intoxication of either the driver or pedestrian involved 46% of the fatalities and over 90% of fatalities during a pedestrian crossing the road [1]. In 2019, pedestrian fatalities comprised about 39% of all traffic fatalities in Washington, D.C., and 41% of traffic fatalities in Baltimore, Maryland [1]. In 2017, pedestrian fatalities comprised about 16 percent of all traffic fatalities, with Washington, D.C., experiencing the highest fatalities involving pedestrians at 35.5%. However, for non-fatal crashes, research has demonstrated consistent underreporting of crashes involving pedestrians since near-miss incidents often go unreported (4).

1.2 Scope of Work

This study use computer vision and machine learning technologies to track pedestrians and vehicles at intersections in two neighboring cities, Washington, D.C., and Baltimore, MD. Located 40 miles apart, these cities have very different socio-demographic profiles; see Table 1. Anecdotal evidence suggests that vehicle and pedestrian behavior in Baltimore, MD, and Washington, D.C., are very dissimilar, but the pedestrians are disproportionately represented in crashes [2]. Research has shown that pedestrian non-compliance increases with increases in delays and detours. Delays exceeding 40 seconds at signalized intersections and 20 seconds at unsignalized intersections may cause risk-taking behaviors [3]. Moreover, variations in behavior are hypothesized to exist within each city due to disparate land uses and demographics. The primary objective of this study is to develop a computer vision pipeline approach to identify and track pedestrians and conflicting vehicles at intersections in order to better understand the microscopic behavior of pedestrians and critical factors affecting pedestrian behavior.

Table 1: U.S. Census Quick Facts for Washington, D.C. and Baltimore, MD

Demographics	Washington, D.C.	Baltimore, MD
% White	45.6%	30.4%
% Black	46.4%	62.5%
Foreign born persons, 2014-2018	14.0%	8.1%
% with bachelor's degree or higher	57.6%	31.2%
Median household income	$82,604	$48,840
Persons per square mile	9856.5	7671.5
Population	705,749	620,770

Video surveillance has long been recognized as a valuable tool for studying pedestrian behavior at intersections. Traditional manual analysis of video footage is time-consuming and resource-intensive. In recent years, significant advancements in computer vision and machine learning techniques have revolutionized video data analysis, providing the means to extract valuable information efficiently and accurately [4].

1.3 Video-Based Tracking of Pedestrians and Vehicles

In this study, we leverage machine learning techniques, specifically the YOLO (You Only Look Once) algorithm [5], to analyze video surveillance data captured from selected intersections in both Baltimore, MD, and Washington, D.C. The locations varied by geometric configuration, land use, traffic volume, and socio-demographic characteristics. The locations include signalized and unsignalized intersections. Activity centers such as schools, retail, tourist attractions, and transit hubs were considered during site selection.

A computer vision pipeline approach was used to identify pedestrians and vehicles from video surveillance footage in order to extract key metrics to characterize pedestrian crossing behavior and associated traffic patterns. The pipeline consists of the following processing stages:

1. Data acquisition,
2. Pre-processing,
3. Background characterization and segmentation,
4. Object identification,
5. Object motion analysis, and
6. System analytics.

The YOLO algorithm can efficiently identify and track pedestrians and different types of vehicles, allowing for comprehensive and automated data extraction [5]. The first stage involves the acquisition of the video image sequences at a sufficiently high spatial resolution to facilitate the extraction of salient features. Video image sequences are filtered in the second stage to minimize signal and compression noise and optimize contrast across each frame. In the third stage, various methods are used to perform background subtraction and frame segmentation to create regions of interest (ROI) around pedestrians and vehicles. The fourth stage uses deep machine learning models to classify ROIs into various subclasses of pedestrians and vehicles. The fifth stage performs a temporal analysis of the motions of pedestrians and vehicles over a given window. The final stage uses this analysis to generate statistics of these motions that can be exported as a summarized report. This technology-driven approach reduces the time and resources required for analysis and enhances the accuracy and consistency of data processing.

The distribution of pedestrian speed was measured at each location. Temporal changes in walking speed will be explored. It is hypothesized that the average crossing speed in certain areas may vary during weekday rush hour versus weekend and off-peak periods. The use of video surveillance allows the team to create pedestrian speed profiles along the entire approach. By utilizing video surveillance data and applying machine learning algorithms, the study aims to better understand pedestrian crossing behavior and safety patterns in both cities. The analysis will offer insights into the differences and similarities between pedestrian behaviors in these urban environments, thereby supporting the formulation of targeted pedestrian safety measures. Ultimately, the findings of this comparative study will contribute to evidence-based decision-making in transportation planning and management, with the overarching goal of creating safer and more pedestrian-friendly cities.

2 Literature Review

2.1 Introduction

Pedestrian safety has been a main concern in the traffic safety area from the beginning of the modern era. According to a study by National Highway Traffic Safety Administration (NHTSA), 6,283 pedestrians deaths were reported in 2018, the highest annual total since 1990 and a 3 percent increase from the year 2017 [1]. In 2019, pedestrian fatalities had decreased by 2.7 percent from the previous year, while pedestrian injuries were reported 1.3 percent higher than in 2018 [6].

The Governors Highway Safety Association (GHSA) mentioned in their 2019 preliminary data report that pedestrian fatality is increasing disproportionately to other traffic fatalities [7]. Pedestrian fatalities as a proportion of total motor vehicle deaths increased from 12 percent to 17 percent from the year 2009 to 2018. Comparing 10 years (2009-2018) of data, they reported that pedestrian fatalities increased by 53 percent, whereas other traffic fatalities increased by 2 percent. As pedestrian safety is a concerning issue, numerous studies have been done to identify the factors and solutions to pedestrian crashes. Previous literature was reviewed to investigate the factors related to pedestrians' safety and understand the previous technologies to understand pedestrian behavior. The review includes research reports and scientific papers published in peer-reviewed journals, conferences, and databases that may contain relevant information. The review was summarized in the next part of this chapter, considering pedestrian crossing behavior, pedestrian detection techniques, pedestrian crash statistics, and previous safety studies in the two study areas.

2.2 Pedestrian Crossing Behavior

Since pedestrians are the most vulnerable road user group, pedestrian safety is always an important issue for transportation safety researchers. Pedestrian movements are considered the most complex and flexible as people are unpredictable and intelligent [8–10]. According to NHTSA, this unpredictable behavior is responsible for most of pedestrian crashes [8]. In Washington, DC, there were 2,600 pedestrian crashes from 2016 to 2020, resulting in 76 fatalities and 2,207 injuries [6]. And there were 3,607 pedestrian crashes in Baltimore City from 2016 to 2020, resulting in 174 fatalities and 3,288 injuries [6]. Interestingly, the highest number of crashes happened in 2016 for both cities, with 795 crashes in Baltimore and 620 crashes in Washington, DC. In terms of the causes of pedestrian crashes, the data shows that both cities have similar issues. The top three causes of pedestrian crashes in Washington, DC, were failure to yield (28%), driver inattention (18%), and pedestrian error (13%). In Baltimore City, the top three causes were driver inattention (26%), failure to yield (23%), and pedestrian error (16%). Numerous studies have been performed over time to investigate pedestrian behaviors that affect pedestrian safety. Different factors such as land use, intersection geometry, environmental condition, and demographics, have been investigated. Pedestrians show a variety of behaviors depending on the conditions, and previous research identified pedestrian behaviors as one of the important factors for pedestrian fatalities.

Most pedestrian-vehicle crashes happen during road crossings, both at intersections (signalized and unsignalized) and mid-block locations [2]. From previous research, pedestrian behavior can be grouped into five categories based on road crossing behavior: violation (intentional), error (knowledge deficiency), lapse (unintentional), aggressive behavior, and positive behavior [8]. The authors provided a framework to evaluate the behaviors mentioned above with a pedestrian behavior questionnaire tool, which can be used in pedestrian safety research under specific circumstances like the change of pedestrian behavior changes due to traffic infrastructure change. They found all these behaviors are responsible for pedestrian-vehicle crashes. Also, other important behaviors are walking speed, zone of comfort, accepted gap, and crossing manners, which are also complex and unpredictable. Different studies have been performed around the world to understand and identify these behaviors.

Walking speed is an important behavior of pedestrians in terms of safety. Walking speed can be affected by different factors like personal characteristics of pedestrians (demographics), trip purpose, route choice, trip length, infrastructure, and environmental characteristics such as grade of roadway and weather conditions [2]. Marked and unmarked crosswalks also play an important role in pedestrians walking speed. A study found that walking speed was more variable at unmarked crosswalks than marked crosswalks, and gradient and lighting were statistically significant variables for walking speed [11]. For the traffic signal design, pedestrian walking speed is a vital factor. Walking speed may vary during the peak and off-peak hours for the same location. Walking speed mostly varies by age and gender. In 2009, MUTCD includes 3.5 ft/s as walking speed for the signal design and included provision to use lower walking speed of pedestrians in the area walk slower [12].

The decision of when to cross the road is also an important factor for unsignalized intersections and mid-block crossing. Pedestrians need to judge the situation if they can find a proper chance to cross the street. This judgment process is governed by the gap acceptance theory. Fitzpatrick et al. studied how pedestrians determine if the gap between two incoming vehicles is good enough to cross the street safely [13]. Pedestrians do not always anticipate the gap effectively, and some pedestrians do not look at the oncoming vehicles with patience. This is often considered to be the most dangerous behavior when crossing [14, 15]. Zhuang and Wu studied the pedestrian crossing behavior at unmarked roadways [15]. Out of the 254 pedestrians surveyed, authors found a significant 65.7 percent did not even look for the vehicles after arriving at the curb. This behavior of waiting to safely cross the road might make a difference. From a literature review, Amini et al. found that road users adopt the crossing strategy by considering broad range of factors [16]. The authors mentioned gap acceptance, speed of an approaching vehicle, road characteristics, size of approaching vehicle, traffic volume, traffic behaviors, and situations, size of the city, visibility, and weather conditions as the decision-making factors for pedestrians to cross the roadway.

Another important pedestrian behavior that can affect pedestrian safety is the manner of crossing the road. This includes running, low walking speed, and using a cell phone while crossing. After a Chi-Square test, Rosenbloom et al. categorized not looking at oncoming vehicles as the most prevalent unsafe behavior [14]. Other behaviors mentioned by the author are a combination of not looking and not stopping, and not stopping before crossing. In the study by Zhuang and Wu, they found that 31.9 percent of pedestrians ran while interacting with the oncoming vehicles, and 11.4 percent stepped backward [15]. It was also observed that pedestrians adjust their walking or running speed according to the behavior of oncoming vehicles. The authors mentioned that pedestrians who ran while crossing usually cross the second half of the road at high speed. Going backward while crossing is mentioned as the more dangerous behavior, which goes against the driver's expectations. This could lead to potential pedestrian fatalities. From the focus group study, authors found that vehicle type could be another factor for pedestrian safety. Using a cell phone or listening to music as well as talking with a companion leads pedestrians to violate the rules unintentionally or forget to look around for the oncoming vehicles [8].

Pedestrian behaviors are unpredictable and cannot be effectively controlled by regulations. Its human behavior to violate traffic rules intentionally or unintentionally. This behavior of traffic violation can be responsible for additional crashes. People can also violate rules unintentionally for different factors like cell phone use, listening to music, and talking with a companion. Other researchers found that pedestrians are more likely to violate rules while walking individually rather than walking with a companion or in a group [17]. Authors mentioned that people would violate rules while crossing narrow roads (4 lanes) rather than a larger number of lanes with a median (7-8 lanes). A successful violation of traffic rules can inspire pedestrians to violate the rules in the same place [18]. These results also satisfy other researches where authors found out that pedestrian accidents often occur due to the disobeying of traffic rules by pedestrians [8, 15]. Detecting and understanding non-compliance behavior can be useful for safety analyses and developing safety countermeasures.

Reviewing the previous crash data and safety studies, it can be summarized that pedestrian characteristics like age, gender, and area characteristics are important factors. Pedestrians' behaviors are related to their characteristics, like age and gender. Researchers found that generally, females wait longer than males at the signalized intersections [19]. Another study found that waiting time is longer at the marked crosswalks if the pedestrian is older [20]. Male pedestrians are more likely to show unsafe behaviors in the roadway than females, which can lead to crashes [8]. Also, younger people are more inclined to intentional violation of traffic rules and unintentional risky behaviors. From these studies, it can be inferred that the personal characteristics of pedestrians may be important factors at roadway crossing. Different studies and crash statistics showed that pedestrian crashes are more common in urban areas [12, 13, 21]. Urban areas are generally more crowded than suburban areas, which means pedestrians are higher there. Large cities offer access to public transportation, limited or expensive parking, and sometimes lower car ownership which are the reasons for high pedestrians [3]. From previous studies, the author mentioned other reasons as traffic congestion, pedestrian facilities, shopping, entertainment, and service areas accessibility to pedestrians. Due to unpredictable behavior, it is very important to analyze the data correctly to understand pedestrian behavior. The next part of the chapter will summarize the existing data collection and analysis techniques for pedestrian safety.

2.3 Pedestrian Detection Techniques

Pedestrian detection is very important to understand pedestrian behaviors. Pedestrians have a higher dynamic range than vehicles, making it difficult to predict pedestrians' movement. The traditional method of studies on pedestrian safety and behavior analysis relies on collision data analysis and the use of judgment of traffic safety professionals, which is a challenging task. For studying pedestrian-vehicle interactions, solely relying on the collision data statistics may not always be sufficient due to data quantity and quality [22]. The conventional field-based method is time-consuming, labor intensive, and also has reliability issues as pedestrian movement is unpredictable, less organized, and more complex than vehicular movement [2, 22]. This reliability issue results from the unorganized pedestrian movement in higher-density areas.

With the advancement of modern technologies, transportation engineers adopted different technologies to collect and analyze data to understand pedestrian behaviors. Ridel et al. reviewed previous studies on pedestrian detection techniques where they mentioned sensors, lidar, cameras, and image processing as a few ways to collect data [23]. Using video sensors has some advantages over manual data collection. Data collection using video sensors is less expensive than conventional systems, can be used for office review and analysis, and archive permanently. Cameras also cover a wide field of view, thus covering more spaces and offering rich and detailed data on pedestrian movements [24]. However, manual video observations are also time-consuming, error-prone, and semi-automated processes also need manual operation, which is also laborious [11]. Though this manual process can be used for offline analysis, this has some shortcomings in manual data collection. Computer vision techniques can be used to overcome these shortcomings as this technique was developed to automatically detect and track moving objects. Zaki et al. used surrogate data to demonstrate the automated safety diagnosis of pedestrian safety issues using computer vision techniques [22, 25]. In another research, authors used computer vision technology to find the non-conforming behaviors both for spatial and temporal violation [25]. The spatial violation occurs when pedestrians cross intersections at undesignated regions, and temporal violation occurs when they cross intersections during improper light.

Several studies have leveraged development in computer vision technologies to study pedestrian behavior The evolution of machine learning and deep learning techniques has made detection easier and more accurate, but effective pedestrian detection and behavioral analysis is still a significant challenge. Vehicular traffic is generally easier to detect than pedestrian movement, as vehicles usually move in a predefined path. Pedestrians, on the other hand, move in unpredictable ways that aren't subject to environmental constraints In computer vision techniques, moving pedestrian detection is usually accomplished by matching pedestrians against predefined samples [24, 25]. The most difficult part of using computer vision techniques is pedestrian tracking, as pedestrians don't have a simple trajectory like vehicular traffic. Nighttime is another factor in pedestrian crashes, so it is important to understand pedestrian behavior at night. Computer vision technologies are limited in these instances as accurate detection is difficult to achieve in low light. Wang et al. proposed two models for human action recognition from video sequences which have the advantage of performing better by utilizing the information provided in the training set [26]. Previously, research initiatives were taken for long-term and short-term pedestrian behavior prediction. Studies showed that long-term predictions are more challenging due to pedestrians' quick, unpredictable movement, while short-time predictions can predict pedestrians' position up to 2.5 seconds [23]. In a simple way, pedestrian detection is a two-step process that includes feature extraction and classification. A typical flowchart of detection is shown in 1.

Gradient-based [28], shape-based [29], texture-based [30], motion-based [31], and part-based [3] features can be used in the pedestrian detection. Other difficulties of pedestrian detection are variations of posture and pose, clothing, different shapes, and variation in illumination. Few studies explore pedestrians' contour, posture, pose recognition, lateral speed, and body language to predict the pedestrian's intentions [23]. With the advancement of technology, researchers are exploring for new technologies to solve the detection problem of pedestrians. With these tools, transportation safety studies can use enormous amounts of detailed data, which is helpful in finding out the behavioral factors for safety.

2.4 Previous Studies and Crash Statistics of Washington, D.C., and Baltimore, MD

Washington, D.C., and Baltimore, MD, are two large cities located 40 miles apart. Due to the increasing traffic volume, density, and other land use factors, crash, and fatality rates are higher in urban areas than in

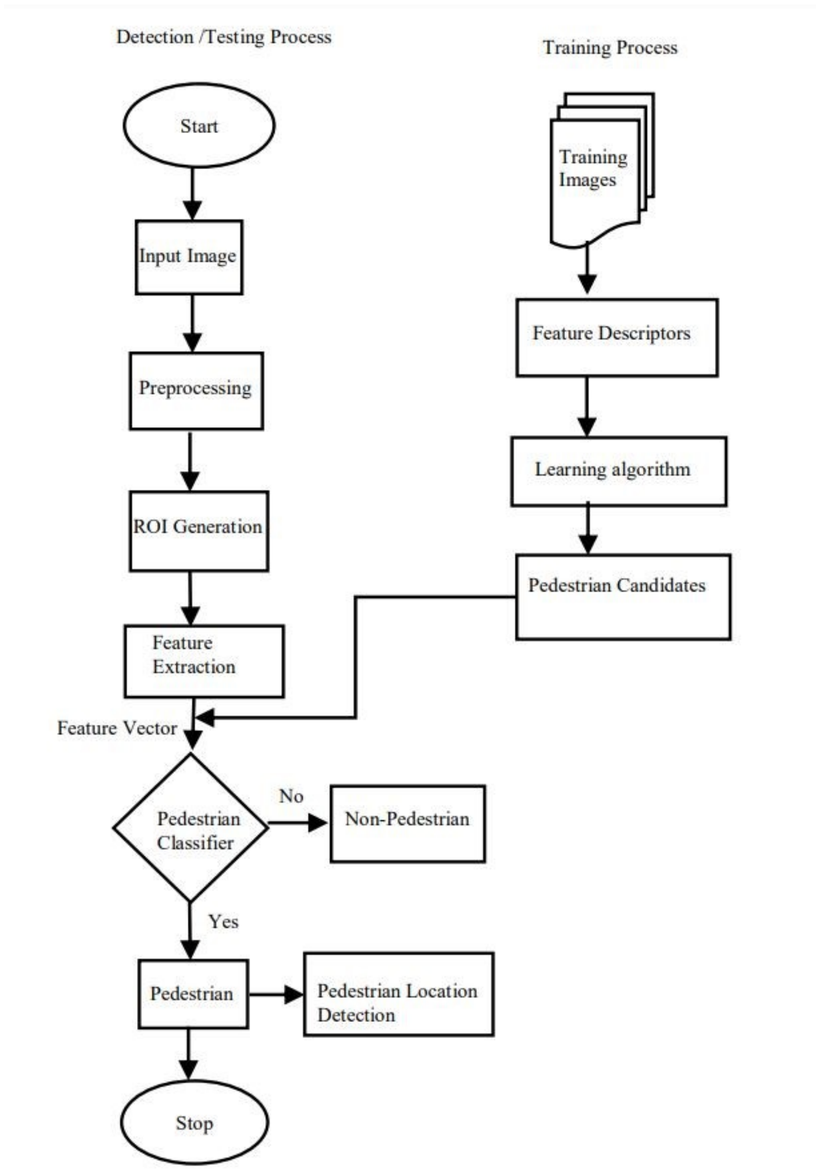


Figure 1: Flow chart of pedestrian detection [27]

rural areas. NHTSA reported that pedestrian fatalities had increased by 62 percent in urban areas in the period of the last 10 years from 2010 [6]. Analysis data for the ten largest cities of the U.S.A. showed that the total number of pedestrian fatalities has increased by 7% from 2017 to 2018 [1]. The report collected data from the Fatality Analysis Reporting System and showed changes in pedestrian fatalities over two years (2017 and 2018). Eight of the cities surveyed had an increased number of fatalities (see Figure 2).

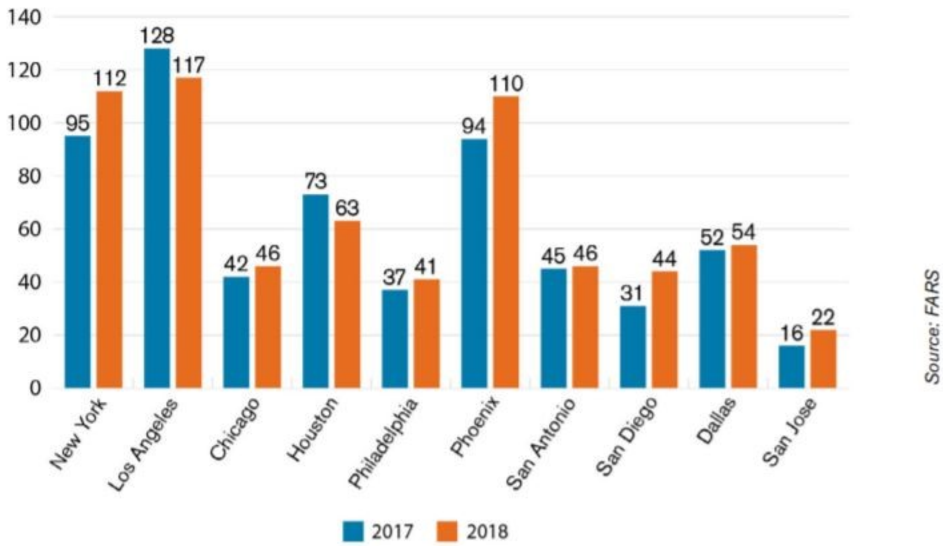


Figure 2: Pedestrian Deaths in the 10 largest U.S. cities: 2017-2018

In Maryland, urban areas are considered the most dangerous areas for pedestrians due to the land use characteristics and population density. According to NHTSA, 521 pedestrians were killed in 2019 in Maryland, which is 1.8 percent higher than the previous year [6]. In Maryland, pedestrian fatalities per 100,000 population were 2.1, whereas the national rate was 1.9 for the year 2018 [1]. Most of these crashes and fatalities occurred in urban areas. In another report of GHSA, Washington D.C, and Baltimore was mentioned as two important metropolitan regional jurisdictions of Maryland where over 80 percent of the pedestrian and bicycle crashes occur [7]. This study will explore the variation of the pedestrian travel behavior of two cities: Washington, D.C., and Baltimore, MD. As these cities have different socio-demographic profiles (see Table 1), vehicle and pedestrian behavior should be different.

In a study on pedestrian crashes in Washington, D.C., and Baltimore, MD, Preusser et al. analyzed pedestrian-involved crashes based on police reports to determine the crash patterns and identify countermeasures [32]. According to the study, the crash pattern changed in 1998 compared to the studies of the 1970s. Authors found a substantial decrease in "Midblock dart-dash" crashes by 22 percent and an increase in "Turning vehicle" crashes by 15 percent in the Washington, D.C. area. They concluded that traffic system changes (installing signals and reducing mid-blocks) have a significant influence on this pattern of changes in crashes. Though all these results reflect the changes in the traffic system, like increasing the controlled traffic from uncontrolled, it cannot be assumed that only vehicles are responsible for the crashes. This study found an increase in pedestrian fault in crashes compared to the 1970s data. They also found age and gender as important factors in the crashes. In another study on the Washington, DC area, Chavis et al. found pedestrians were at fault for crashes 26.7 percent of the time, which emphasizes the importance of the pedestrian behavior study to reduce pedestrian crashes [33].

Summarizing the studies on Washington D.C. pedestrian crashes, the main crash type was "Midblock dart-dash" (37%) in 1976, "Turning vehicle" (25%) in 2002, and "Motorist Left Turn-Parallel paths" (21.43%) in 2018 and the pedestrians' fault in the occurred crashes are increasing. This emphasizes the importance of studying pedestrian behavior studies to reduce fatalities. From the above pedestrian crash data and Washington, D.C. and Baltimore statistics, we can understand the importance of investigating pedestrian crossing behaviors. Urban planners and policymakers can benefit from this kind of study and use the results to enhance pedestrian safety.

Each research investigation examined the behavior of pedestrians when crossing intersections, albeit with variations in their specific methodologies. Certain studies focused solely on particular intersections or regions where pedestrian-vehicle accidents had seen an increase, while others explored the most heavily trafficked intersections within a city. One study specifically assessed how pedestrians navigate diverse types of intersections. The studies generally analyzed factors such as walking speed, jaywalking tendencies, and potential conflicts with automobiles. Although each study successfully identified a consistent pattern or correlation pertaining to pedestrian behavior, several studies revealed contradictory findings. Nevertheless, due to the distinct geographical settings in which each study was conducted, these divergent outcomes remain valid within their respective contexts.

3 Methodology

This section discusses the data sources, methods, and techniques for collecting and analyzing data. Various empirical data sources were used to obtain relevant information to overlay with the video processing data. In order to provide a comprehensive understanding of the process, this section will detail the different types of data sources utilized and the methods employed to collect and analyze the data, and then the analytical methods used to derive meaningful insights.

3.1 Study Area

Washington, D.C. has an area of approximately 58.34 square miles with a population of 702,455 that is 45.4% Black or African American, 35.8% White, and 11.3% Hispanic or Latino [34]. It has a 1,516-mile public road length, and 70% of it is a local urban road [6]. Fig 12 illustrates the camera locations in Washington, D.C.

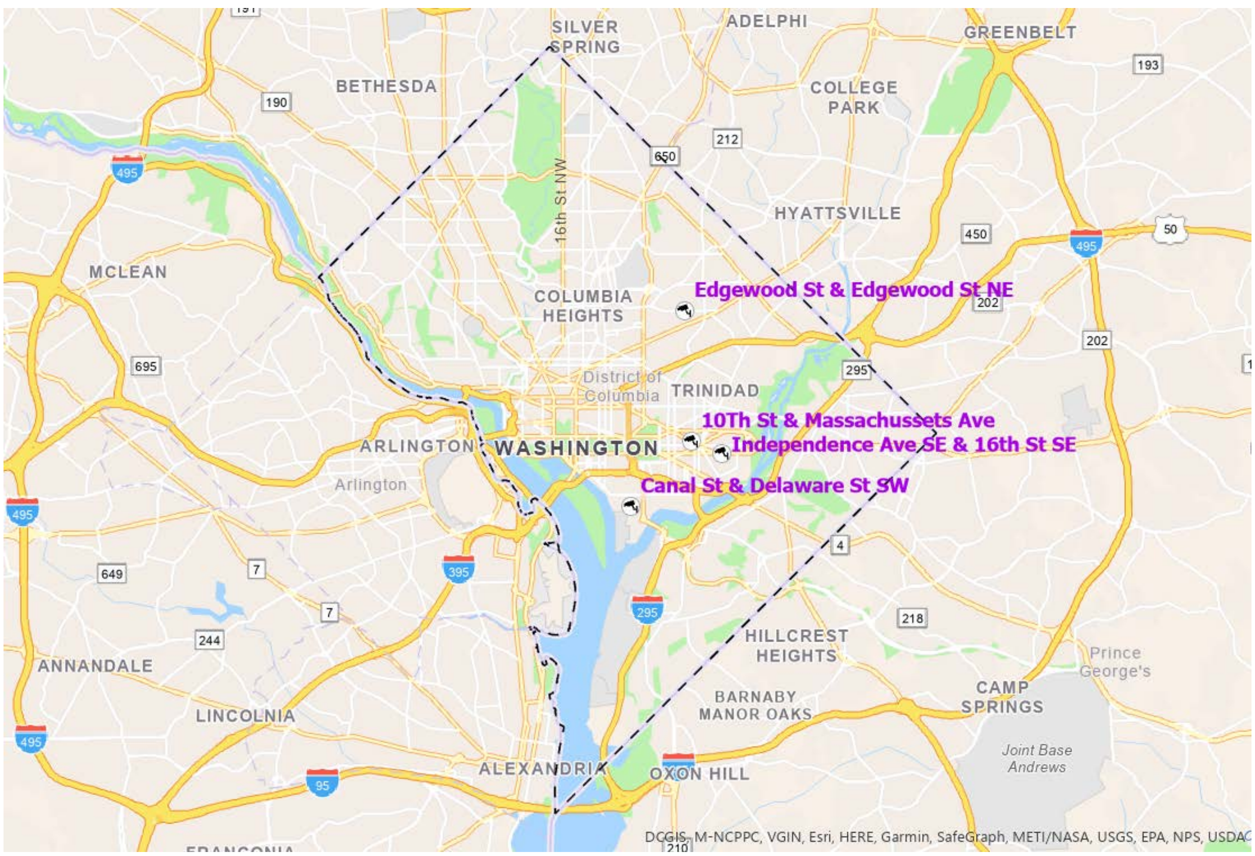


Figure 3: Washington D.C. Camera Locations

Baltimore, Maryland has 593,490 people and among them 62.8% are Black or African American, 28.3% White, and 6.1% Hispanic or Latino living in 92.05 square miles area [34]. The city has over 2,100 miles of roads that are maintained by the Department of Transportation [35]. Figure 4, shows the camera locations in Baltimore, MD.

Figure 4: Baltimore City Camera Locations

3.2 Data Sources

3.2.1 Video Data

CountCam, a vehicle counting system that relies on video footage, is used to collect video recordings with the help of CubeRoot, a specialized engineering consulting firm [36]. The CountCam houses a digital video camera and recorder with the capacity to store numerous hours of recorded material. In total, there were 52 hours of video footage recording for the four locations (see 3) for Washington DC and 48 hours of footage four locations in Baltimore (see 4). Capturing continuous High-Definition (HD) video data is more expensive, but the detail, accuracy, and possible automation can justify the expense [37]. The cameras were mounted on a pole on the selected sites and equipped with sufficient storage and a power bank to ensure uninterrupted footage; refer to Figure 5. The video footage collected allows us to analyze the pedestrian behavior at the intersection. The intersections were selected by focusing on accessible and busy locations with higher numbers of pedestrians crossing.

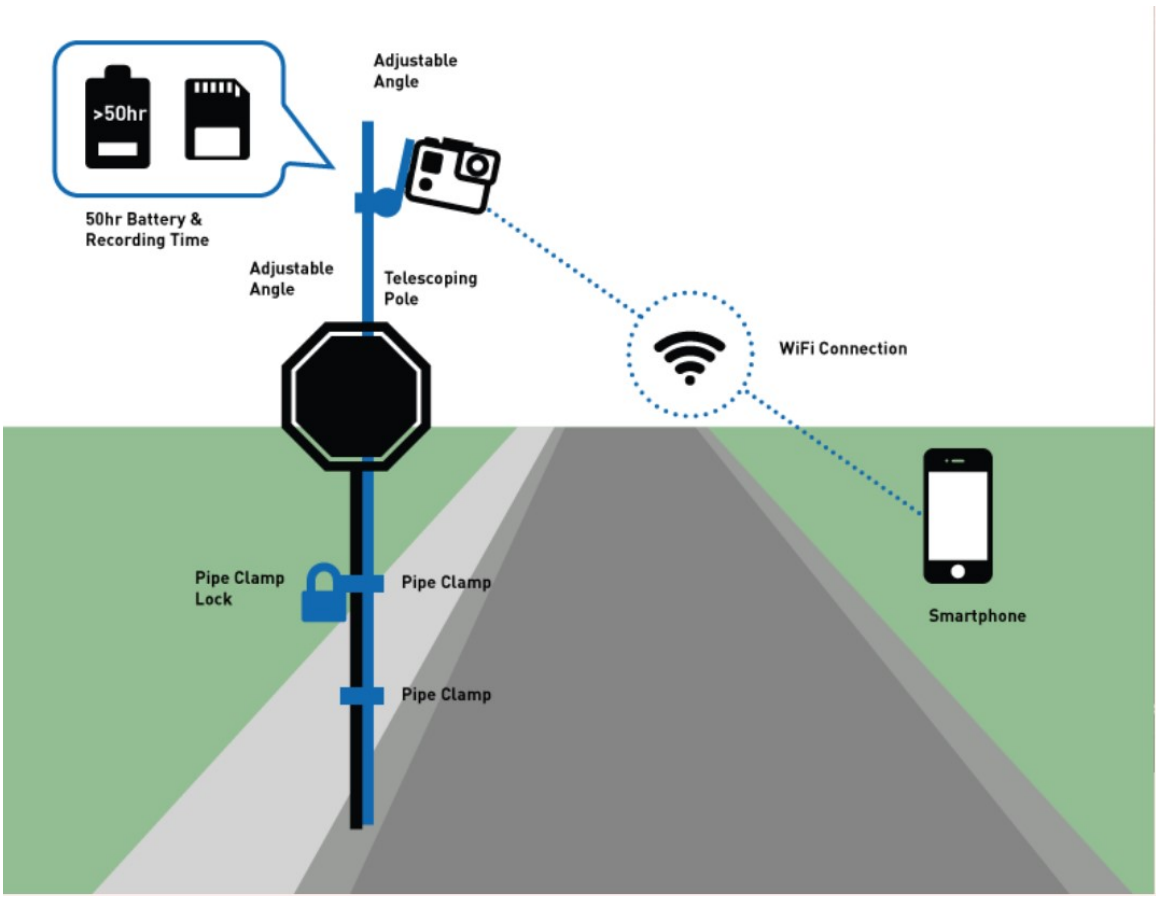


Figure 5: CountCam Setup ([36])

3.2.2 Sociodemographic Data

The American Community Survey (ACS), one-year estimates (2020) were used for the analysis. Following are the data tables used at the census tract level to develop the indicators of cluster analysis for site selection; see Figure 7.

- Population (B01001): Under 18 years old
- Population (B01001): Over 65 years old
- Race (B02001): White
- Race (B02001): African American
- Race (B02001): Asian
- Race (B02001): Hispanic
- Household (S1701): Below 100% Poverty Level
- Household (B08201): No Vehicle Access
- Household (B19013): Income below $35,000
- Household (B19013): Income Over $100,000

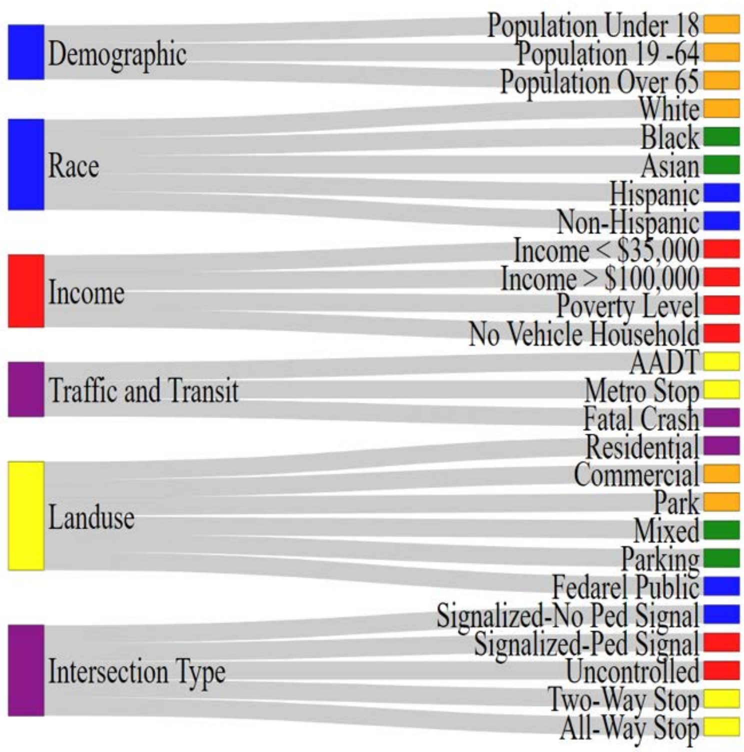


Figure 6: Cluster Indicators

3.2.3 Spatial Data

GIS datasets are used to demonstrate the spatial distribution and provide a geographic foundation for analysis. Census tracts (2019) are a commonly used boundary in this study for spatial analysis, both when paired with other spatial sources (e.g. demographic and points of interest data) and as the primary analysis unit for spatial clustering and regression. These boundaries can be obtained from the Maryland Open Data Portal and DC Open Data.

3.3 Site Selection

CubeRoot provided video footage for 109 locations in Washington, DC. The majority of the footage was too low resolution for the Computer Vision Machine Learning application. Out of 109, there were 24 locations that met the initial requirements. This study applied a hybrid methodology for selecting camera locations using spatial analysis combined with multi-criteria hierarchical clustering. Multi-criteria hierarchical clustering identifies distinctive clusters satisfying the socio-demographic variations and ensuring equity. The cluster analysis is a statistical analysis approach that presumes that the data analyzed often contains redundant information [38]. Hierarchical clustering is a recognized associativity analysis methodology used to determine variables or objects' inherent or natural groupings to summarize data into groups [39].

3.3.1 Hierarchical clustering

The hierarchical clustering technique is applied in this study to identify similar locations. It is an alternative approach to the K-means clustering for detecting groups within the dataset. The clusters can be illustrated in an attractive tree-based representation of the observations, called a dendrogram. The theoretical foundation of hierarchical clustering has the benefit of making no assumptions regarding the mutual independence of samples. Therefore, it does not require exploring all clustering possibilities. A distance metric establishes the similarity among members. The distance metric generates a similarity matrix in which data are cross-compared. Hierarchical clustering can be performed under two major approaches: (a) Agglomerative clustering and (b) Divisive clustering.

Agglomerative Clustering: Agglomerative clustering is also known as AGNES (Agglomerative Nesting). The AGNES method approaches clustering in a bottom-up manner considering each object as a single-element cluster (leaf) [40]. The algorithm grows bigger (nodes) by combining the two similar clusters at each step. This procedure is repeated until all the objects are members of one single cluster (root).

Divisive Clustering: Divisive hierarchical clustering, or DIANA works in a top-down manner. It is an inverse order of the AGNES. The algorithm begins with the root, including all objects in a single cluster. With each iteration, the most different cluster is split into two. The process is repeated until all objects are in their own cluster [40].

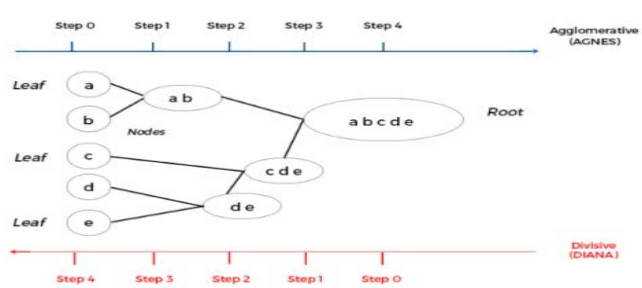


Figure 7: Agglomerative and Divisive Clustering

The hierarchical cluster analysis was performed using R programming language. The method of the analysis is as follows:

1. Required packages are: ***tidyverse, cluster, factoextra, dendextend***
2. Data Preparation
   (a) Missing value in the data is omitted
   (b) The data is standardized (i.e., scaled) to make all variables comparable.
3. Distance matrix was computed using the Euclidean distance.
4. Hierarchical cluster analysis was performed using the agglomeration methods ***"complete", "average", "single", "ward.D"***
5. The agglomerative coefficient was calculated using ***agnes*** function
6. Divisive clustering was performed using the ***diana*** function
7. The dendrogram was visualized and compared for both methods
8. The number of clusters was determined using the elbow method. The elbow method selects the number of clusters where the decrease in the between-cluster sum of squares (WSS) becomes less significant.

The cluster analysis conducted initially suggested the presence of six distinct clusters; see Figure 8. However, during the quality check process, it became evident that only four locations were suitable for further video processing. This decision was made as the video footage obtained from other locations did not meet the desired criteria for machine learning analysis. The Baltimore locations were selected using spatial analysis based on the cluster analysis results. The four selected intersections have similar attributes as the Washington D.C. locations. By focusing on these specific locations, the study ensured that the data used for analysis would be of higher quality and relevance, ultimately leading to more reliable and accurate findings.

Table 2 lists the four selected locations in Washington D.C. and four locations in Baltimore City. The primary objective of this research was to investigate pedestrian behavior at intersections, whether signalized or unsignalized intersections. To accomplish this, recordings were discreetly captured using high-definition, field-mounted video cameras. All video recordings were conducted during daylight hours and under clear weather conditions. It is worth noting that camera placement ensured the visibility of the entire crosswalk, including the pedestrian signals associated with each crosswalk. Due to time limitations, only the West Hamburg St and Scott St locations in Baltimore were included in the analysis.

Table 2: List of Washington DC and Baltimore Locations

SL No	City	Intersection Name	Intersection Control	Collection Date	Type of Day	Duration
1	D.C.	Independence Ave SE & 16th St SE	Two-Way Stop	4/23/2019	Weekday	5 AM - 8 PM
2	D.C.	10th St NW & Massachusetts Ave NW	Signalized, Ped Signal	5/9/2019	Weekday	5 AM - 8 PM
3	D.C.	Edgewood St NE & 8th St NE	Uncontrolled	4/9/2019	Weekday	6 AM -6 PM
4	D.C.	Canal St SW & Delaware Ave SW	All-Way Stop	5/14/2019	Weekday	4PM - 6PM
5	D.C.	Canal St SW & Delaware Ave SW	All-Way Stop	5/8/2019	Weekday	2PM - 10 PM
6	Baltimore	W Hamburg St & Scott St	All-Way Stop	4/5/2023	Weekday	7AM - 7PM
7	Baltimore	Harford Rd and Moravia Rd	Signalized	4/6/2023	Weekday	7AM - 7PM
8	Baltimore	Baltimore St & Charles St	Signalized	4/7/2023	Weekday	7AM - 7PM
9	Baltimore	Liberty Heights Ave & Garrison Blvd	Signalized	4/8/2023	Weekday	7AM - 7PM

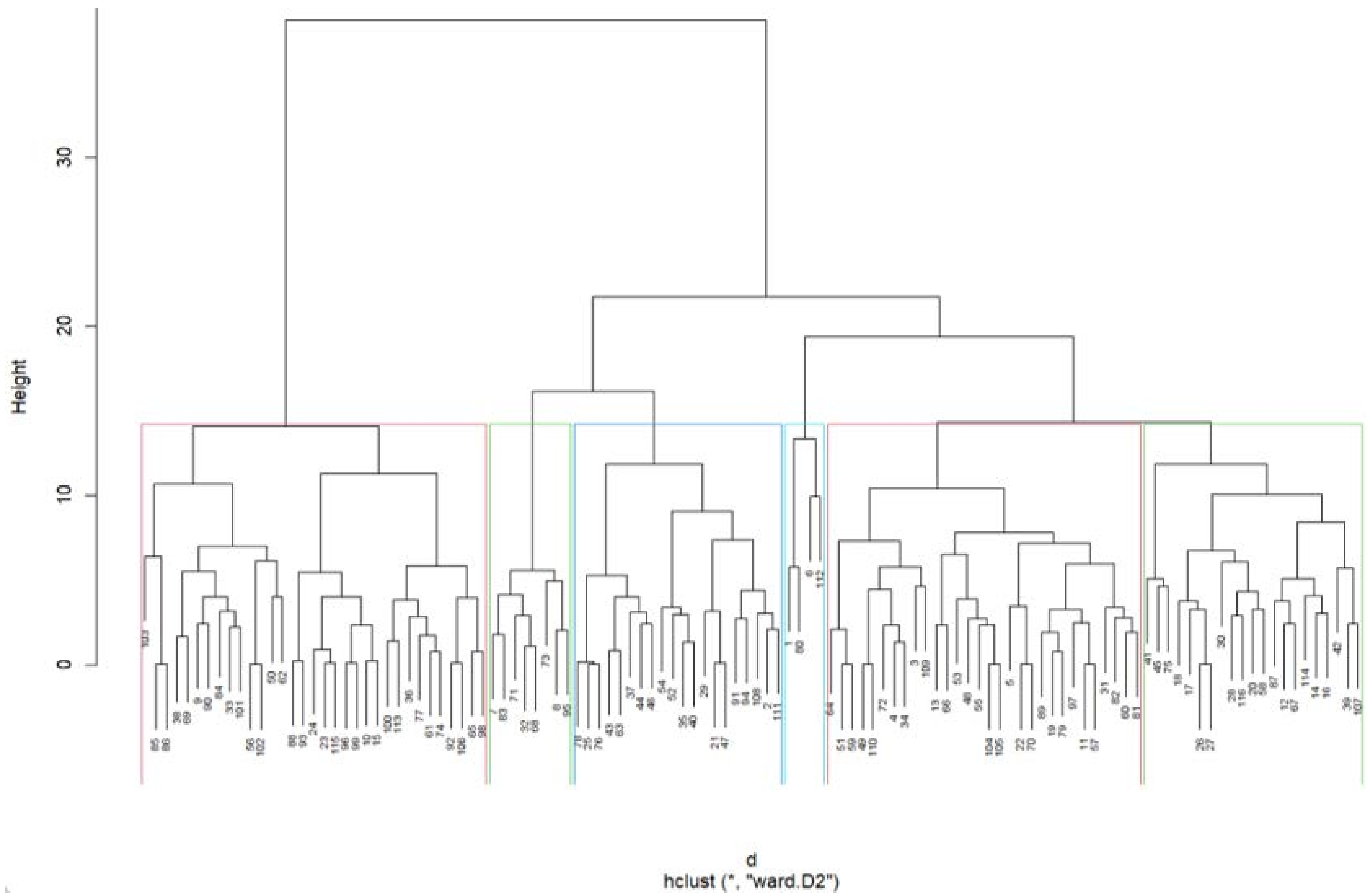


Figure 8: Cluster Dendrogram

3.4 Computer Vision Pipeline Methodology

Manual observation from a video recording is the most common method for obtaining information about pedestrian behaviors at intersections [13]. It is a time-consuming and costly process due to the labor-intensive nature of frame-by-frame observation and the potential for human error. In contrast, automating counting using machine learning (ML) techniques offers advantages such as faster processing, higher accuracy, and the ability to extract meaningful insights [4]. Automated video data analysis has become very popular with the advancement of object detection and tracking algorithms. As illustrated in Figure 9, our Computer Vision Pipeline methodology incorporates state-of-the-art object detection and object tracking algorithms to create a robust real-time moving vehicle and pedestrian detection, tracking, GPS location, and counting system.

3.4.1 OpenCV

OpenCV is an open-source, cross-platform library to develop real-time computer vision applications [41]. OpenCV can perform image processing, video capture, analysis, object recognition, etc. It supports multiple languages, including Python, Java, and C++. For this project, we used the OpenCV library to read and write images, capture and save videos, and process images for further analysis.

3.4.2 YOLOv7 for Object Detection

For our project, object detection is a crucial step, and it is the base on which object tracking and the rest of the methodology and process depend. For computers, detecting objects is a complex task. At first, it processes an input image or a single frame from a video and outputs features/information of objects on the image and their position (pixel coordinates). Then, an Object Detector detects an object in a frame, puts a bounding box around it, and classifies the object. Figure 10a illustrates this method. Among multiple Convolutional Neural Network (CNN) based object detectors, our computer vision pipeline uses the state-of-the-art real-time object detection algorithm called YOLO.

YOLO stands for "You Only Look Once," and YOLOv7 is the 7th version of YOLO Object detection models that uses deep CNN to perform object detection [42]. The original YOLO was first introduced in 2015 by Joseph Redmon in his research paper titled: "You Only Look Once: Unified, Real-Time Object Detection" [43]. Since then, it has produced a series of the best real-time object detectors in computer vision: YOLO, YOLOv2, YOLOv3, YOLOv4, YOLOv5, YOLOv6, YOLOv7 and YOLOv8. The YOLOv7 has proven its higher performance in a broad range of detection tasks [42]. In addition, Yolov7 has been implemented in multiple popular frameworks, including Tensorflow and Keras, which have also been used in the project.

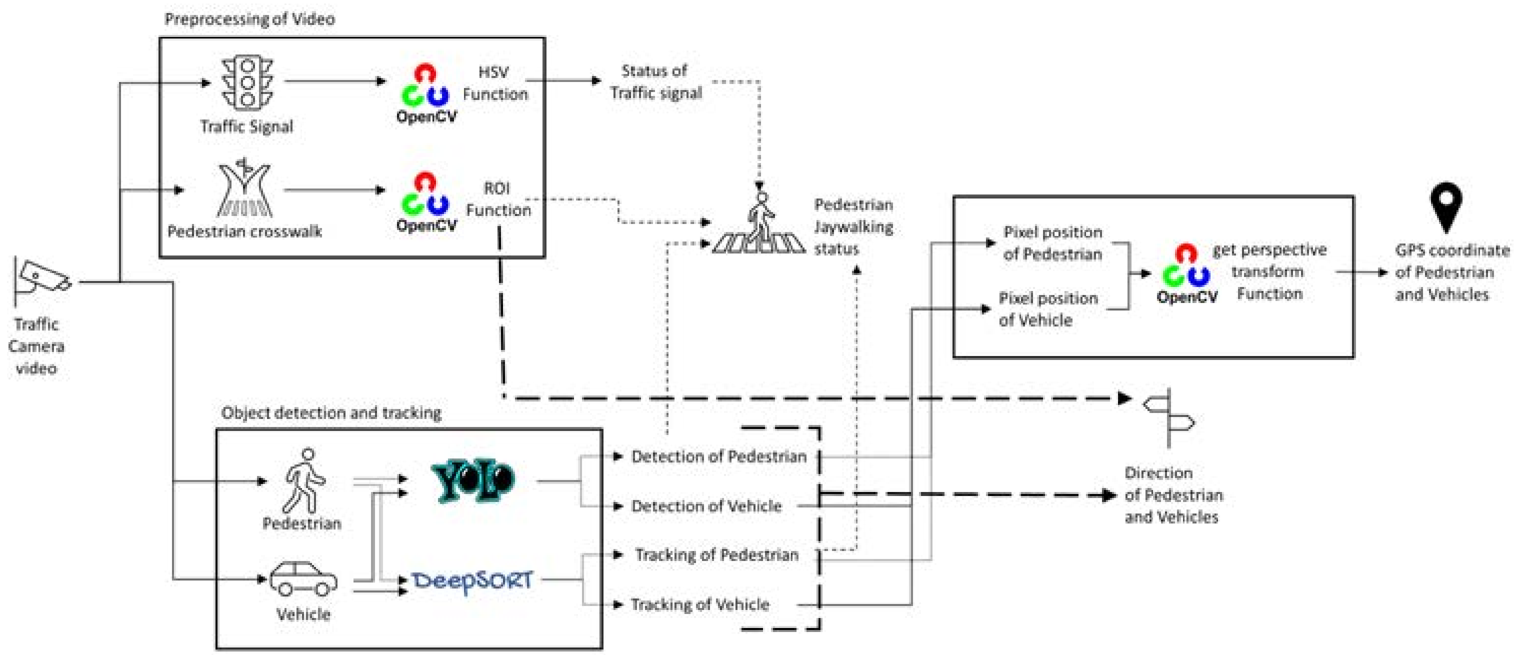
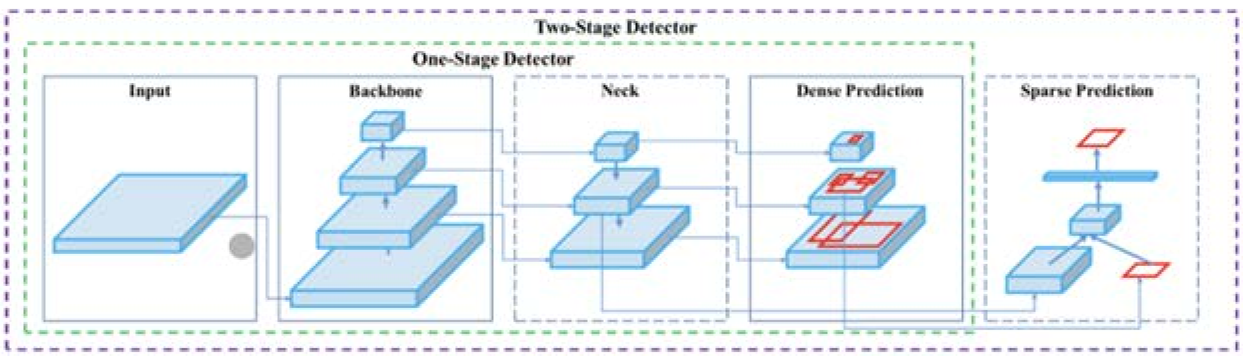


Figure 9: Computer Vision Pipeline Methodology

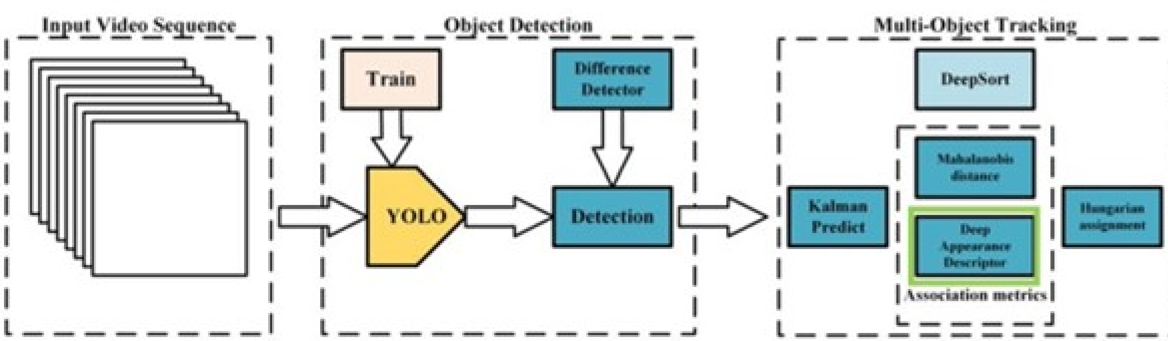
3.4.3 DeepSORT for Object Tracking

Object tracking is an imperative task in computer vision. An object tracker tracks a particular object across all frames of the entire video. An object tracker sits on top of an object detector, uses the bounding box and the classification from the object detector, and matches across all frames for tracking; see Figure 10b. For this project, we are using the state-of-the-art object tracking algorithm, DeepSORT.

DeepSORT model is an expansion of the popular SORT (Simple Online Real-Time Tracker) model. SORT is a simple framework that uses a Kalman filter for tracking [44]. On top of the SORT model, the DeepSORT model includes the appearance information for every detection. The appearance information is calculated using a CNN by computing a 128-dimensional feature vector.



(a) Object Detector



(b) Object Tracker

Figure 10: Deep SORT Object Detector and Tracker Framework (Image Source: [41])

3.4.4 Region of Interest (ROI) for Pedestrian Crossing Selection

Every traffic intersection has a different alignment of pedestrian crossing shown in the traffic video footage. Therefore, to accurately determine the pedestrian's behavior, we must correctly identify the pedestrian crossing. We used the ROI (Region of Interest) function of OpenCV for custom detection of Pedestrian crosswalks.

3.4.5 Traffic Signal Status Detection for Signalized Intersections

A signalized intersection has traffic lights installed at vantage locations at the intersection to control when drivers enter the intersection to assign right-of-way to conflict movements of traffic at the intersection. The traffic signal is determined using semi-supervised learning techniques. The HSV (hue, saturation, value) data of the traffic region is collected using unsupervised learning; the data is clustered into three different clusters - red, amber, and green. The clusters are then used to label the data, creating a supervised model. To test the model, a prediction is made on new traffic data and manually inspected to know which cluster corresponds to a particular color. In cases where wrong predictions are observed, the model is retrained on new data, the model used to train the data can be changed, or both. Once the model's predictions are assessed and confirmed, the model is saved using the Python pickle library and loaded into the main script.

3.4.6 Jaywalking Status

An important aspect of this project is determining if the pedestrians are jaywalking. To automate the process of determining jaywalking status, we needed four components:

1. Detection and location of objects in each frame from our object detector, YOLOv8
2. Tracker ID from our object tracker, DeepSORT
3. ROI of each pedestrian crossing
4. Traffic signal status

Combining these four components, we can now determine if pedestrians are jaywalking. If a detected object is a pedestrian (people) and if their position (location in each frame) is inside the location of a pedestrian crossing and if the traffic signal for that crossing is "green," then the pedestrian is "jaywalking."

3.4.7 Vehicle and Pedestrian Direction

To determine the direction of each object (both vehicle and pedestrian), we provided a 'marker' to each object whenever they were inside the pedestrian crossing. Each pedestrian crossing has a specific marker. Therefore, when an object's location is inside the pedestrian crossing, it will be assigned a specific marker. Later, we determine how many markers a single object has crossed during post-processing. Ideally, each object will not have more than two markers considering the pedestrian crossings are from a four-way intersection. Therefore, we can easily determine their direction from the order of the two markers of each object. We then add another label to track pedestrians who cross from one side of the road to another.

3.4.8 Time Stamp Extraction

With knowledge of the frame rate of the videos, the time can be computed for each frame. The initial time is extracted from the name of the video, and through computations using the Python datetime package, the time on each frame is generated.

3.4.9 GPS Coordinates

Another important task of this project was to determine the GPS coordinates for each object for further analysis. We have the pixel location of each object from the bounding box information provided by our object detector, YOLOv8. Then, we mapped the pixel coordinates to GPS coordinates using perspective transformation. Perspective transformation is a matrix operation that projects a set of points from one 2D plane to another. We performed this transformation using the OpenCV function, getPerspectiveTransform [45]. This function requires information (both pixel and GPS) about four corners of a quadrilateral and provides GPS coordinates of all the detected objects from their pixel coordinates.

3.4.10 Speed Trajectory Determination

The Haversine formula was applied to calculate distances between two geographical points on the surface. The Haversine formula is widely used for its accuracy in estimating distances over short to medium distances. The Haversine formula is a mathematical method for calculating the great-circle distance between two points on the Earth's surface, given their latitude and longitude coordinates. It is based on the Law of Haversines, which relates the sides and angles of spherical triangles. The formula is as follows ([46]):

$$
\begin{aligned}
a &= \sin^2\left(\frac{\Delta\phi}{2}\right) + \cos(\phi_1)\cdot\cos(\phi_2)\cdot\sin^2\left(\frac{\Delta\lambda}{2}\right) \\
c &= 2\cdot \operatorname{atan2}\left(\sqrt{a},\ \sqrt{1-a}\right) \\
d &= R\cdot c
\end{aligned}
\tag{1}
$$

where:

d is the distance between the two points in kilometers
$\Delta\phi$ is the difference in latitude between the two points
$\Delta\lambda$ is the difference in longitude between the two points
ϕ_1 and ϕ_2 are the latitudes of the two points, respectively
R is the Earth's mean radius (mean radius = 6,371 km)

The distance calculation procedure involves the following steps:

1. Convert latitude and longitude coordinates from degrees to radians.
2. Compute the differences $\Delta\phi$ and $\Delta\lambda$.
3. Apply the Haversine formula to calculate the great-circle distance between the two points.

To implement the distance calculation using the Haversine formula, Python programming language is used. The built-in functions and libraries for trigonometric calculations and conversion between degrees and radians were applied. It is important to note that the Haversine formula assumes a spherical Earth, which introduces some level of approximation. For high-precision applications over long distances, more complex models that account for the Earth's ellipsoidal shape may be necessary [46]. Additionally, the formula does not take into account factors such as altitude or variations in Earth's radius, which may affect distance calculations in specific scenarios.

3.5 Computer Visioning Pipeline Outputs

When the algorithm is run, Python OpenCV video capture breaks the video down into frames and works on one frame at a time. With each frame, the YOLO algorithm predicts objects on the frame. The selected objects to be detected are cars, trucks, persons, bicycles, and buses. Depicted in Figure 11 below is a frame with predictions and a boundary box around each object predicted.

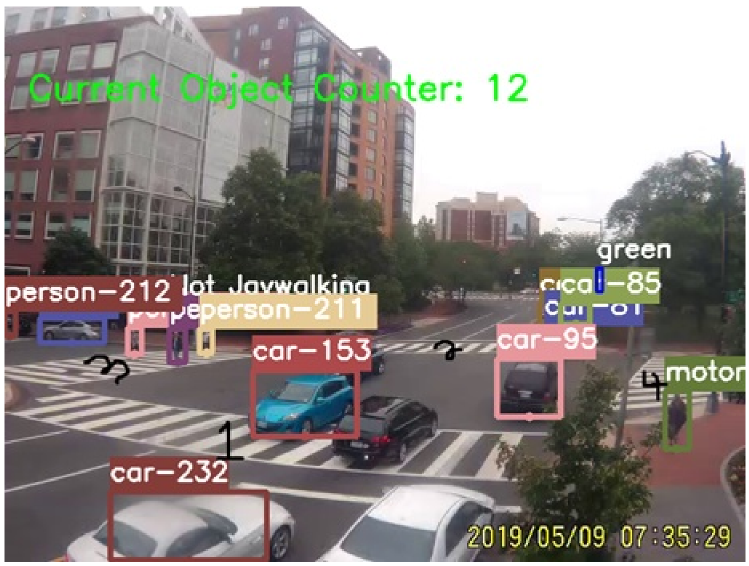


Figure 11: Predictions and Boundary Box

There are 12 regions of interest (ROI):

- ROI-1: Crosswalk marked 1
- ROI-11: The left side of crosswalk marked 1
- ROI-12: The right side of crosswalk marked 1
- ROI-2: Crosswalk marked 2
- ROI-21: The left side of crosswalk marked 2
- ROI-22:The right side of crosswalk marked 2
- ROI-3: Crosswalk marked 3
- ROI-31: The left side of crosswalk marked 3
- ROI-32: The right side of crosswalk marked 3
- ROI-4: Crosswalk marked 4
- ROI-41: The left side of crosswalk marked 3
- ROI-42: The right side of crosswalk marked 4

The pixels at the base of the boundary boxes are used as references for the objects in the images. Objects in the crosswalks marked 1-4 are marked accordingly, and all other places are marked 0.

The markers can be used to track the directions of vehicles. When the traffic light is detected as green:

- Pedestrians on the ROI 1 and 2 are marked Jaywalking

- Pedestrians on the ROI 3 and 4 are marked Not-Jaywalking

and when the light is detected as red:

- Pedestrians on the ROI 3 and 4 are marked Jaywalking
- Pedestrians on the ROI 1 and 2 are marked Not-Jaywalking.

For GPS tracking, 4 points on the stationary image with their pixel coordinates are used as a reference to estimate all the GPS references of other objects.

All video recordings were conducted during daylight hours and under clear weather conditions. It is worth noting that camera placement ensured the visibility of the entire crosswalk, including the pedestrian signals associated with each crosswalk. As shown in Figure 12 for the four locations in D.C., the GPS coordinates of fixed points along the intersection were used for reference. In addition, various site characteristics were documented for each study location. These characteristics encompassed the length of each study crosswalk in feet, the duration of "walk" and "flashing don't walk" pedestrian signal phases in seconds, the cycle length in seconds, the posted speed limit for the road being crossed in each crosswalk, and the presence of pedestrian push buttons at the intersections. From the trajectory of a pedestrian, the distance was calculated from different frames using the haversine formula. Then the speed between these frames was calculated from the distance and time.

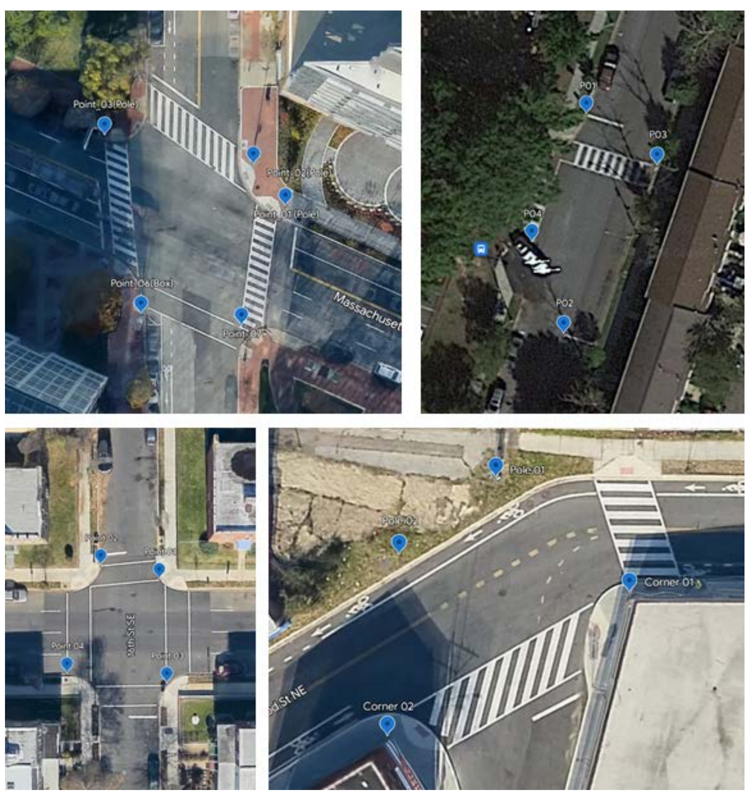


Figure 12: Satellite View of DC Locations

3.5.1 Processing CSVs

The data from the video processing is manipulated using tidyverse and dplyr packages. To begin, the directory containing the raw CSV files and the directory to save the cleaned CSV files were specified. A list of the CSV files in the directory was then created. The code was designed to loop through the list of CSV files and perform the following operations for each file:

First, the CSV file was read into a data frame. The columns of the data frame were then renamed to more meaningful names, and unnecessary columns were deleted. A function was defined to convert the time column to seconds. This function was applied to the time column and added a new seconds column to the data frame. The values in three columns were concatenated into a new column to create a unique identifier for each record. The TrSignal column was recorded from "red" to 0 and "green" to 1, while the Jaywalking column was recorded from "Jaywalking" to 1 and "No" to 0. The VClass column was recorded from "bicycle," "bus", "car", "motorbike," "person," and "truck" to 1, 2, 3, 4, 5, and 6, respectively. Lastly, the latitude and longitude values were aggregated to seconds using the UID column as the grouping variable.

The modified data frame was then written to a new CSV file that could be used for further analysis; see Figure 13. The parameters examined in this investigation comprised time, vehicle category, pedestrian presence, jaywalking classification, and displacement (latitude and longitude). Among these factors, latitude and longitude measurements were utilized to determine the distance traveled. The jaywalking classification was expressed in binary form, with "yes" and "no" as possible outputs. Timestamp data was employed to calculate the duration and velocity of road crossings. The duration represents the time taken by an individual to complete the act of crossing the road, while the velocity signifies the speed at which a person traverses the road. It was calculated by dividing the length of the crosswalk by the crossing duration.

Frame	Traffic Signal	Time	Tracker ID	Class	Jaywalking	location(x)	location(y)	GPS(lat)	GPS(lon)	Marker	Markers
3		1900-01-01 5:49:00	1	car	NA	421	314	38.88742939	-76.98222561	0	NA
4		1900-01-01 5:49:00	1	car	NA	421.5	314	38.88742943	-76.9822261	0	NA
5		1900-01-01 5:49:00	1	car	NA	421.5	314	38.88742943	-76.9822261	0	NA
6		1900-01-01 5:49:00	1	car	NA	423.5	313	38.88741865	-76.98222587	0	NA
7		1900-01-01 5:49:00	1	car	NA	424	313	38.88741869	-76.98222637	0	NA
8		1900-01-01 5:49:00	1	car	NA	424	313	38.88741869	-76.98222637	0	NA
9		1900-01-01 5:49:00	1	car	NA	423.5	313	38.88741865	-76.98222587	0	NA
10		1900-01-01 5:49:01	1	car	NA	423.5	313	38.88741865	-76.98222587	0	NA
11		1900-01-01 5:49:01	1	car	NA	423.5	313	38.88741865	-76.98222587	0	NA
12		1900-01-01 5:49:01	1	car	NA	422.5	313	38.88741856	-76.98222487	0	NA
13		1900-01-01 5:49:01	1	car	NA	423	313	38.8874186	-76.98222537	0	NA
14		1900-01-01 5:49:01	1	car	NA	423.5	313	38.88741865	-76.98222587	0	NA
15		1900-01-01 5:49:01	1	car	NA	424.5	314	38.88742962	-76.982229	0	NA
15		1900-01-01 5:49:01	6	car	NA	477	293	38.88692202	-76.98224561	0	NA
16		1900-01-01 5:49:01	1	car	NA	422.5	313	38.88741856	-76.98222487	0	NA
16		1900-01-01 5:49:01	6	car	NA	477	293	38.88692202	-76.98224561	0	NA
17		1900-01-01 5:49:01	1	car	NA	421.5	313	38.88741848	-76.98222387	0	NA
18		1900-01-01 5:49:01	1	car	NA	421.5	313	38.88741848	-76.98222387	0	NA
18		1900-01-01 5:49:01	7	car	NA	459.5	299	38.88716494	-76.98223357	0	NA
19		1900-01-01 5:49:01	1	car	NA	422.5	313	38.88741856	-76.98222487	0	NA
19		1900-01-01 5:49:01	7	car	NA	460.5	299	38.88716556	-76.98223519	0	NA
20		1900-01-01 5:49:02	1	car	NA	423	314	38.88742953	-76.98222755	0	NA
20		1900-01-01 5:49:02	7	car	NA	459.5	300	38.88719476	-76.98223669	0	NA
21		1900-01-01 5:49:02	1	car	NA	422	314	38.88742946	-76.98222658	0	NA
21		1900-01-01 5:49:02	7	car	NA	459.5	300	38.88719476	-76.98223669	0	NA
22		1900-01-01 5:49:02	1	car	NA	421.5	314	38.88742943	-76.9822261	0	NA
22		1900-01-01 5:49:02	7	car	NA	460	300	38.88719502	-76.98223745	0	NA
23		1900-01-01 5:49:02	1	car	NA	421.5	314	38.88742943	-76.9822261	0	NA
23		1900-01-01 5:49:02	7	car	NA	459.5	300	38.88719476	-76.98223669	0	NA
24		1900-01-01 5:49:02	1	car	NA	421.5	314	38.88742943	-76.9822261	0	NA
24		1900-01-01 5:49:02	7	car	NA	459.5	300	38.88719476	-76.98223669	0	NA
25		1900-01-01 5:49:02	1	car	NA	421.5	314	38.88742943	-76.9822261	0	NA

Figure 13: Sample Output CSV File

4 Comparative Analysis

4.1 Data Processing

After getting the intersection tracking file, the data was processed for further analysis. The datasets obtained from the video tracking contain information for every 30-minute interval as the raw video data were collected at 30-minute intervals. The video tracking process generated separate datasets for pedestrians and each modes of transportation. The pedestrian dataset includes tracker IDs, latitude and longitude of the pedestrians, timestamps, and information on whether they were using the crosswalk or not. For the analysis presented in this section, only the pedestrian dataset was used.

Since the tracking data was recorded every second, there was a possibility of erroneous tracking due to inherent inaccuracies. Considering that the typical pedestrian average speed is around 4 ft/s or 2.73 mph [47], and given the small distances and time intervals between consecutive frames, a constraint of 5 mph was applied to the final datasets. Any speeds exceeding 5 mph were deemed unreliable and thus discarded. Also, due to tracking inaccuracies, there is a chance of switching tracker IDs during different time frames or when pedestrians cross each other. So, another constraint was introduced to minimize this problem. Only tracker ID numbers with a minimum of 3 rows or time frames of data in the dataset were considered for analysis. The final dataset was prepared for the pedestrian behavior analysis by applying all these constraints.

For the intersections of the Washington D.C. area, three times of the day, morning peak, evening peak, and mid-day or off-peak time, were considered during the analysis. Based on the MWCOG travel demand model, the morning peak in this area is from 6:00 am to 9:00 am, evening peak hour is from 03:00 pm to 06:00 pm [48]. Traffic characteristics are different for the two peak hours and off-peak hours. Based on their data, the analysis presented in this section considered 8 am to 9 am as the morning peak, 5 pm to 6 pm as the evening peak, and 12 pm to 1 pm as the off-peak hour for the pedestrian behavior analysis.

In this section, descriptive statistics were described for the different intersections of the Washington D.C. and Baltimore area based on the data from pedestrian detection. The five intersections evaluated are as follows:

- 10th Street NW and Massachusetts Avenue NW,
- Canal Street SW and Delaware Street SW,
- Independence Avenue SE and 16th Street SE,
- Edgewood Street NE and 8th Street NE, and
- W Hamburg Streeet and Scott Street.

All of these intersections are stop-controlled intersections except of the intersection of 10th Street NW and Massachusetts Avenue NW. Due to time limitations, only one video in Baltimore, the intersection of W Hamburg St. and Scott St, was processed at the time of publication.

4.2 10th St NW and Massachusetts Avenue NW

Data was collected and processed for a duration of 24 hours at the intersection of 10th St NW and Massachusetts Ave NW; see Figure 14. This 4-approach signalized intersection experiences a high volume of pedestrian activity during both peak and off-peak hours. The total number of pedestrians detected at the intersection during specific time periods is as follows: 365 and 432 during the two 30-minute periods from 8 am to 9 am in the morning peak hours, 380 and 466 in the consecutive 30-minute periods of the evening peak hours, and 270 and 288 during the consecutive 30-minute periods from 12 pm to 1 pm in the off-peak hours. The substantial number of pedestrians confirms that this intersection is highly active regarding pedestrian presence.

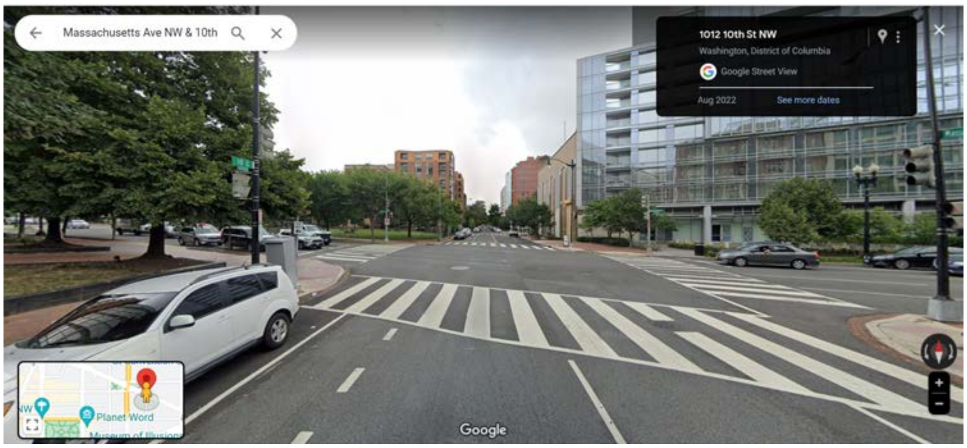
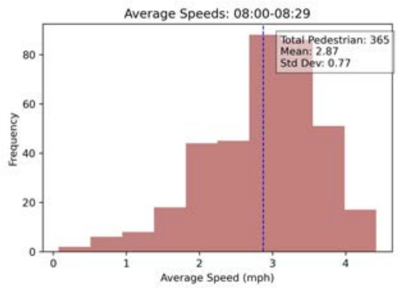
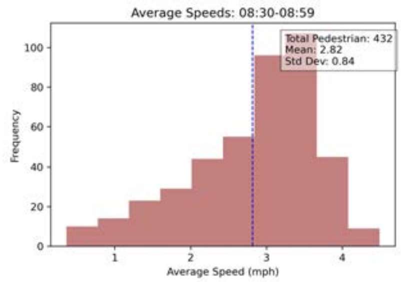


Figure 14: 10th St NW and Massachusetts Avenue NW Intersection Street View (Source: Google)

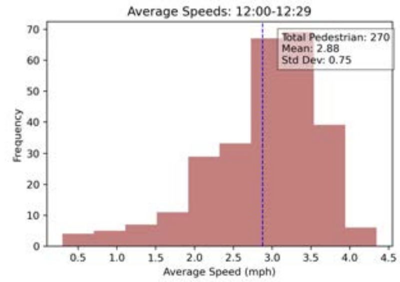
Figure 15 represents the average speeds for the different time periods of morning, evening peak hours and off-peak hours. The average speed of the pedestrians is 2.87 mph and 2.82 mph for the AM peak hour periods, 2.76 and 2.81 mph for the PM peak hour periods, and 2.88 and 2.72 for the off-peak hour periods. These average speeds align with the literature. The standard deviations range from 0.75 to 0.84 for these estimations. The greatest variation in speeds occurred during the off-peak hour.



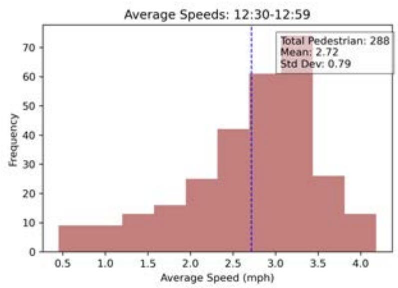
(a) 8:00-8:29 am



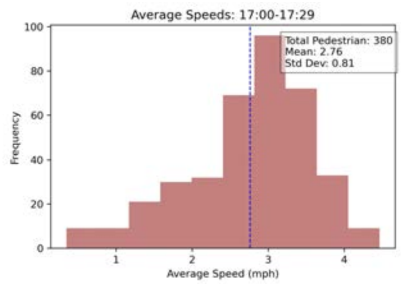
(b) 8:30-8:59 am



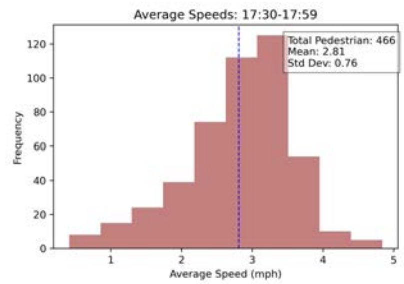
(c) 12:00-12:29 pm



(d) 12:30-12:59 pm



(e) 5:00-5:29 pm



(f) 5:30-5:59 pm

Figure 15: Pedestrian Speed Histogram for 10th St and Massachusetts Avenue

4.3 Canal Street SW and Delaware Street SW

The intersection of Canal Street SW and Delaware Street SW (Figure 16) is located three blocks away from Nationals Park, the home of the Washington Nationals Baseball Team. This intersection is controlled by stop signs and typically experiences low pedestrian traffic due to its location between residential areas. Therefore, typical peak hours and off-peak hours data would not contain a higher number of pedestrians. So, for analysis of this intersection, game day and non-game day data were considered. A relatively higher number of pedestrians are present in the evening from 7 pm to 10 pm, with the highest hour being from 7 to 8 pm.

A comparison was made between the average speed data during a game day (Figures 17a and 17b) and a non-game day (Figures 17c and 17d), specifically from 7 pm to 8 pm. As shown in Figure 17, the average speed is 1.41 mph and 1.48 mph for 7:00 to 7:29 and 7:30-7:59 pm, respectively, on a game day, and 1.10 mph and 1.78 mph for the same time interval on a non-game day. Both days exhibit a high standard deviation, which could be attributed to the lower number of pedestrians or a small sample size. Further data and analysis are necessary to understand why this intersection's average speed is lower.

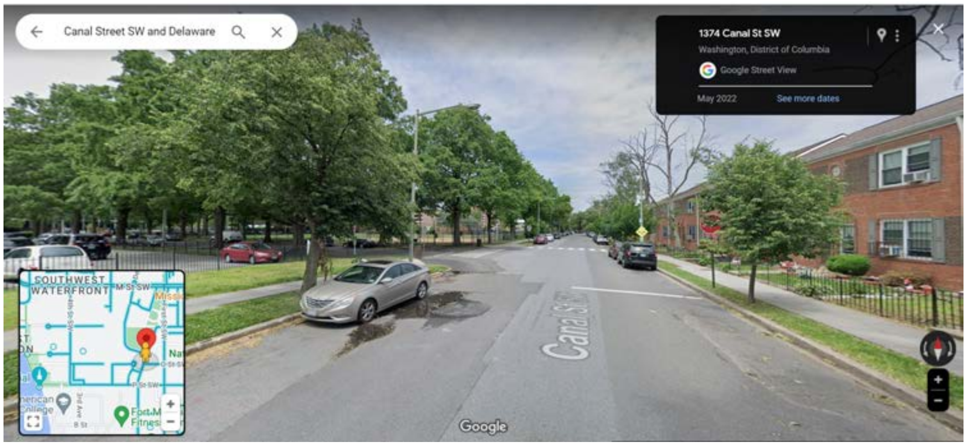
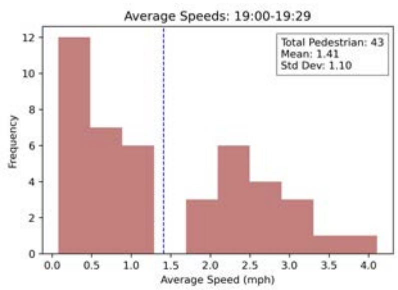
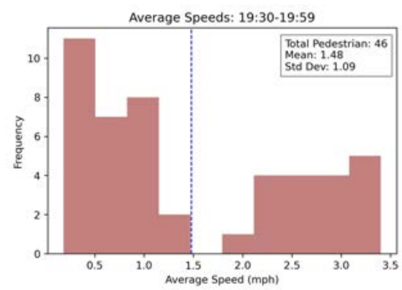


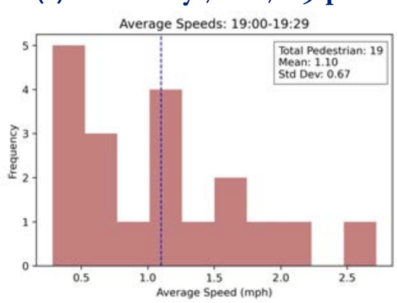
Figure 16: Canal Street SW and Delaware Street SW Intersection Street View (Source: Google)



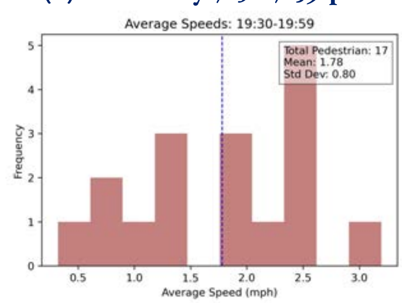
(a) Game Day 7:00-7:29 pm



(b) Game Day 7:30-7:59 pm



(c) Non-Game Day 7:00-7:29 pm



(d) Non-Game Day 7:30-7:59 pm

Figure 17: Pedestrian Speed Histogram for Canal Street SW and Delaware Street SW

4.4 Independence Avenue SE and 16th Street SE

As shown in Figure 18, this intersection is also a stop-controlled intersection that attracts low pedestrians. As it is a stop-controlled intersection, pedestrian numbers are low at this intersection. However, considering Figure 19 below, it is evident that the average speeds of pedestrians during different hours of the day were close to the standard value of 4 ft/s or 2.73 mph.

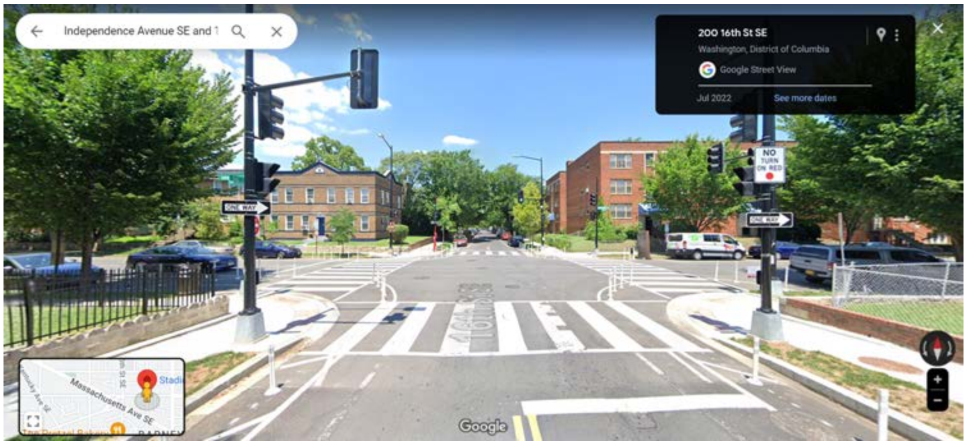
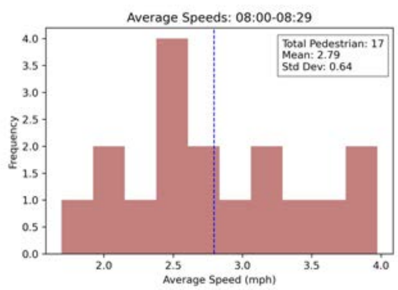
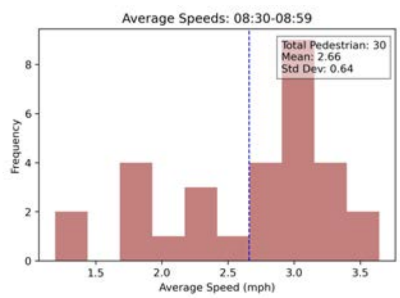


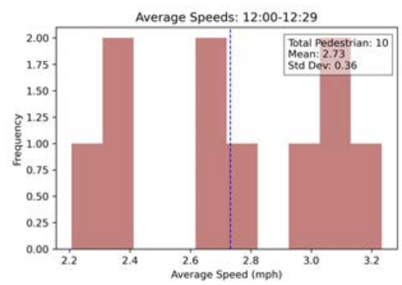
Figure 18: Independence Avenue SE and 16th Street SE Intersection Street View (Source: Google)



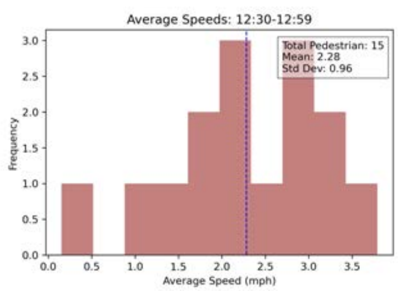
(a) 8:00-8:29 am



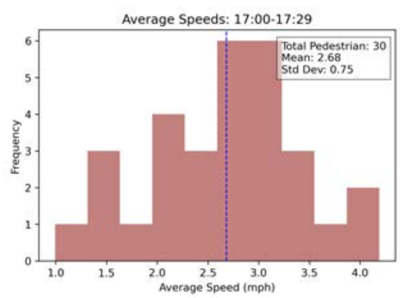
(b) 8:30-8:59 am



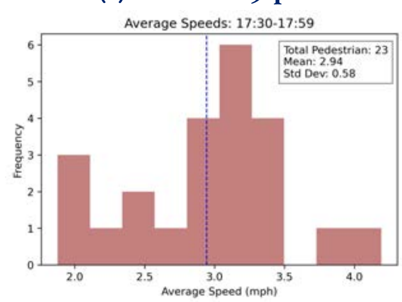
(c) 12:00-12:29 pm



(d) 12:30-12:59 pm



(e) 5:00-5:29 pm



(f) 5:30-5:59 pm

Figure 19: Pedestrian Speed Histogram for Independence Avenue SE and 16th Street SE

4.5 Edgewood Street NE and 8th Street NE

The intersection of Edgewood Street NE and 8th Street NE is stop-controlled; refer to Figure 20. Compared to the other two stop-controlled intersections, the total number of pedestrians detected during the morning and evening peak hours is higher. As presented in Figure 21, the average speed in the morning peak hours is 1.94 mph and 1.63 mph in the 30-min periods and 1.75 mph and 1.93 mph in the evening peak hours' 30-min periods. However, the average speeds during off-peak hours are lower, measuring 1.68 mph and 1.81 mph in the two 30-minute periods.

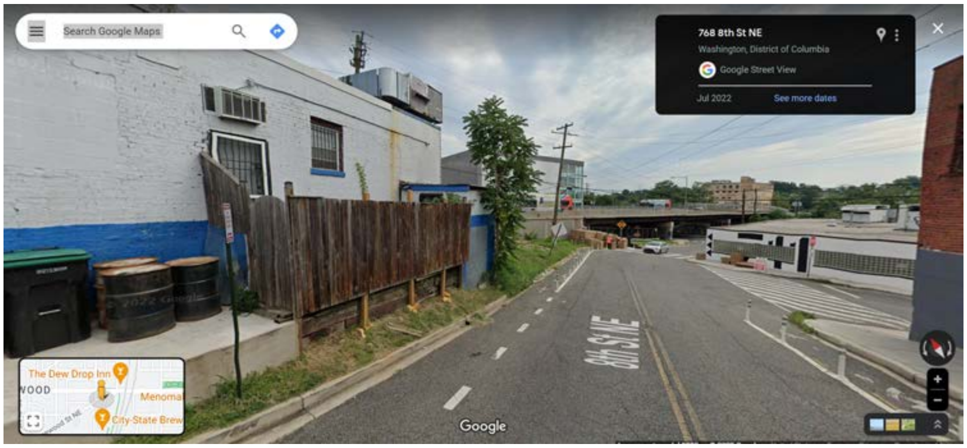
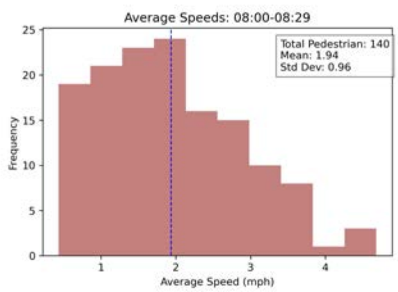
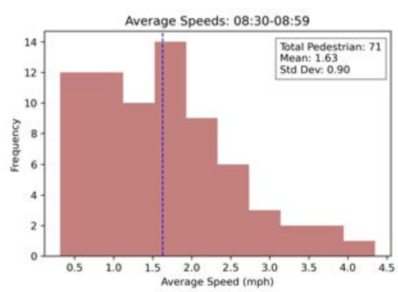


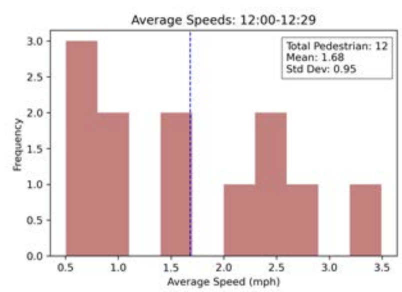
Figure 20: Edgewood Street NE and 8th Street NE Intersection Street View (Source: Google)



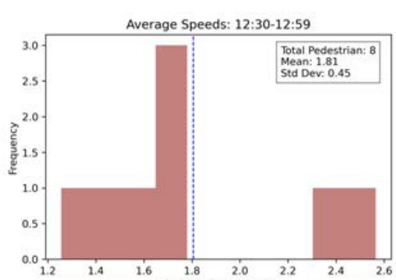
(a) 8:00-8:29 am



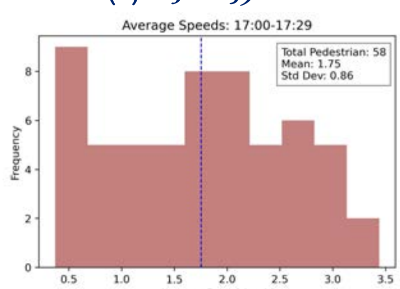
(b) 8:30-8:59 am



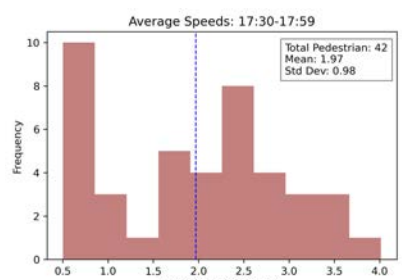
(c) 12:00-12:29 pm



(d) 12:30-12:59 pm



(e) 5:00-5:29 pm



(f) 5:30-5:59 pm

Figure 21: Pedestrian Speed Histogram for Edgewood Street NE and 8th Street NE

4.6 W Hamburg Street and Scott Street

The intersection of W Hamburg Street and Scott Street near George Washington Elementary School in Baltimore is controlled by stop signs; see Figure 22. The count of pedestrians identified at this junction (Figure 23) is notably minimal. During the morning peak hours, the average speeds are recorded at 2.63 mph and 3.32 mph within two 30-minute intervals. In the evening peak hours, the speeds are 3.06 mph and 2.69 mph during the respective 30-minute periods. Interestingly, even during off-peak hours, the average speeds remain nearly equivalent to those during rush hours, measuring 2.84 mph and 3.05 mph within the two 30-minute segments.

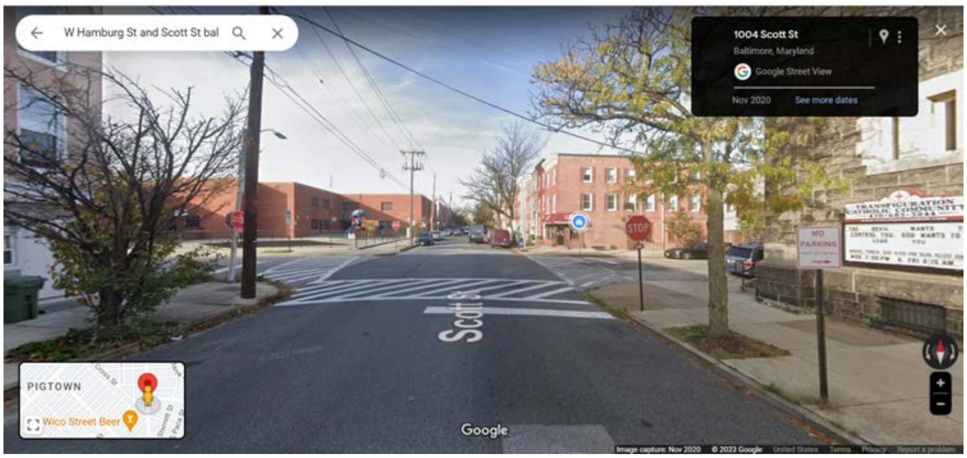
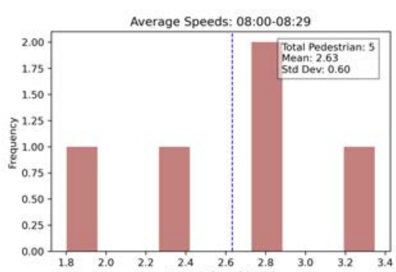
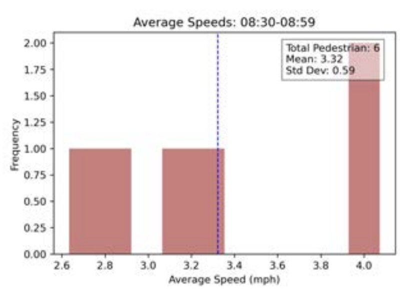


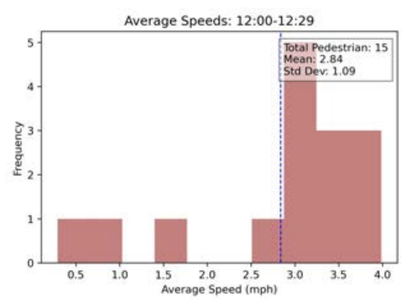
Figure 22: W Hamburg Street and Scott Street Intersection Street View (Source: Google)



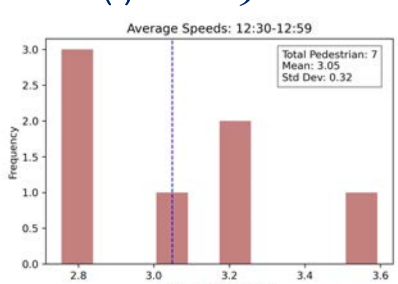
(a) 8:00-8:29 am



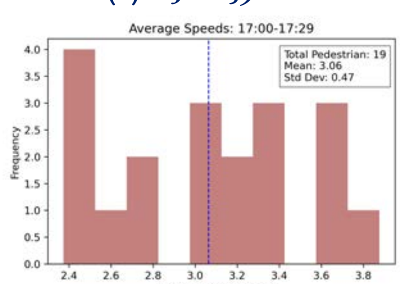
(b) 8:30-8:59 am



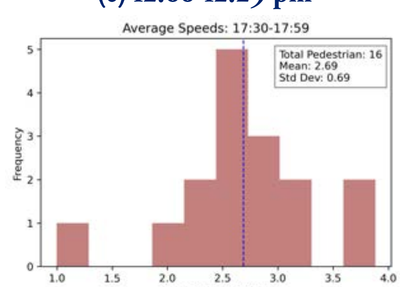
(c) 12:00-12:29 pm



(d) 12:30-12:59 pm



(e) 5:00-5:29 pm



(f) 5:30-5:59 pm

Figure 23: Pedestrian Speed Histogram for 10th St and Massachusetts Avenue

5 Conclusions

The previous section primarily concentrated on analyzing average pedestrian speed for different hours of the day, total detection, and detection of pedestrians on crosswalks. While these descriptive statistics provide some insights, additional analysis on pedestrian behavior could be conducted if more data were available, such as socio-demographic data and ground truth data for validation purposes. Furthermore, it would be beneficial to compare the accuracy of this detection technique with existing methods. By utilizing this detection technique, various modes of transportation were also identified with corresponding timestamp data for every intersection. Exploring the interaction between vehicles and pedestrians using this dataset could be a promising avenue for further pedestrian behavior analysis.

The preliminary analysis used a small subset of the data metrics collected during the computer visioning process. Video data is being analyzed on the additional three Baltimore locations and algorithms are being developed to look at the interaction of vehicles and pedestrians at intersections. We conclude this report by presenting some of the challenges and lessons learned while developing the computer vision pipeline methodology, and ideas for future work.

5.1 Challenges

The project faced several challenges that made the work more difficult and affected the accuracy and effectiveness of the computer vision pipeline. A key issue was variability in data quality and environmental conditions, which impacted algorithm accuracy. Foggy weather limited visibility, causing inaccuracies in object detection and tracking. Complex traffic scenes exacerbated these challenges, with object occlusions and overlaps hindering clear visibility. Distinguishing between similar-looking objects became challenging, leading to false positives in detections. Post-processing methods were needed to filter out these errors. Also, when multiple objects were in close proximity, their IDs could change, making it difficult to consistently associate behavior with specific objects. Camera positioning also played a critical role; some camera angles may not have captured important elements like traffic signals, impacting data analysis. Addressing these challenges requires robust algorithms, advanced pre-processing techniques, and careful consideration of the specific data quality issues and environmental factors. It often involves employing specialized approaches and optimization strategies to enhance the accuracy and reliability of the analysis.

5.2 Future Work

In our ongoing research, we are embarking on exciting new projects to enhance urban life and safety. One focus is improving pedestrian and road user detection in busy city environments. We aim to make this pipeline more accurate and efficient, especially in low light or visually complex settings. We will explore predicting pedestrian behavior, essentially forecasting people's next moves as in a chess match. This could be immensely valuable for managing traffic flow and preventing accidents, ensuring smooth and safe city operations. We are committed to making our pipeline universal, effective across diverse cities and regions globally. By testing and adapting our systems to varied urban environments and cultural contexts, we aim to ensure people worldwide can benefit from safer, more convenient cities. In pursuing these research avenues, we strive to create urban environments that are not only safer but also more responsive to the needs of diverse populations. Our goal is making cities globally better places to live, work, and enjoy life.

References

1. National Highway Traffic Safety Administration. *Traffic Safety Facts* tech. rep. October (2019). https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812826.
2. Daamen, W. & Hoogendoorn, S. P. Experimental research of pedestrian walking behavior. *Transportation research record* **1828,** 20–30 (2003).
3. Felzenszwalb, P. F., Girshick, R. B., McAllester, D. & Ramanan, D. Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32,** 1627–1645. ISSN: 00189162 (2014).
4. Emiyah, C., Nyarko, K., Chavis, C. & Bhuyan, I. *Extracting Vehicle Track Information from Unstabilized Drone Aerial Videos Using YOLOv4 Common Object Detector and Computer Vision* in *Proceedings of the Future Technologies Conference (FTC) 2021, Volume 2* (ed Arai, K.) (Springer International Publishing, Cham, 2022), 232–239. ISBN: 978-3-030-89880-9.
5. Bochkovskiy, A., Wang, C.-Y. & Liao, H.-Y. M. *YOLOv4: Optimal Speed and Accuracy of Object Detection* 2020. arXiv: 2004.10934 [cs.CV].
6. National Highway Traffic Safety Administration. *Traffic Safety Facts* tech. rep. December (2020). https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813060.
7. Governors Highway Safety Association. *Pedestrian Traffic Fatalities by State* tech. rep. (2019). https://www.ghsa.org/sites/default/files/2020-02/GHSA-Pedestrian-Spotlight-FINAL-rev1.pdf.
8. Deb, S. *et al.* Evaluating pedestrian behavior at crosswalks: Validation of a pedestrian behavior questionnaire for the U.S. population. *Accident Analysis and Prevention* **106,** 191–201. ISSN: 00014575. https://doi.org/10.1016/j.aap.2017.05.020 (2017).
9. Jian, L., Lizhong, Y. & Daoliang, Z. Simulation of bi-direction pedestrian movement in corridor. *Physica A: Statistical Mechanics and its Applications* **354,** 619–628 (2005).
10. Lavalette, C. B. D. *et al.* Pedestrian crossing decision-making : A situational and behavioral approach. *Safety Science* **47,** 1248–1253. ISSN: 0925-7535. http://dx.doi.org/10.1016/j.ssci.2009.03.016 (2009).
11. Li, S. *et al.* Automated collection of pedestrian data through computer vision techniques. *Transportation research record* **2299,** 121–127 (2012).
12. Fitzpatrick, K. *et al. Characteristics of Texas Pedestrian Crashes and Evaluation of Driver Yielding at Pedestrian Treatments* tech. rep. 2 (2014).
13. Fitzpatrick, K. *et al.* Improving pedestrian safety at unsignalized crossings. *National Cooperative Highway Research Program Report* **562** (2006).
14. Rosenbloom, T., Ben-Eliyahu, A. & Nemrodov, D. Children's crossing behavior with an accompanying adult. *Safety Science* **46,** 1248–1254 (2008).
15. Zhuang, X. & Wu, C. Pedestrians' crossing behaviors and safety at unmarked roadway in China. *Accident Analysis and Prevention* **43,** 1927–1936. ISSN: 00014575. http://dx.doi.org/10.1016/j.aap.2011.05.005 (2011).

16. Amini, R. E., Katrakazas, C. & Antoniou, C. Negotiation and decision-making for a pedestrian roadway crossing: A literature review. *Sustainability (Switzerland)* **11,** 1–24. ISSN: 20711050 (2019).

17. Koh, P. P. & Wong, Y. D. Gap acceptance of violators at signalised pedestrian crossings. *Accident Analysis and Prevention* **62,** 178–185. ISSN: 0001-4575. `http://dx.doi.org/10.1016/j.aap.2013.09.020` (2014).

18. Xu, Y., Li, Y. & Zhang, F. Pedestrians ' intention to jaywalk : Automatic or planned ? A study based on a dual-process model in China. *Accident Analysis and Prevention* **50,** 811–819. ISSN: 0001-4575. `http://dx.doi.org/10.1016/j.aap.2012.07.007` (2013).

19. Tiwari, G., Bangdiwala, S., Saraswat, A. & Gaurav, S. Survival analysis : Pedestrian risk exposure at signalized intersections. *Transportation Research Part F* **10,** 77–89 (2007).

20. Hamed, M. M. Analysis of pedestrians' behavior at pedestrian crossings. *Safety Science* **38,** 63–82. ISSN: 09257535 (2001).

21. Sisiopiku, V. P. & Akin, D. Pedestrian behaviors at and perceptions towards various pedestrian facilities: An examination based on observation and survey data. *Transportation Research Part F: Traffic Psychology and Behaviour* **6,** 249–274. ISSN: 13698478 (2003).

22. Zaki, M., Sayed, T., Tageldin, A. & Hussein, M. Application of computer vision to diagnosis of pedestrian safety issues. *Transportation Research Record,* 75–84. ISSN: 03611981 (2013).

23. Ridel, D., Rehder, E., Lauer, M., Stiller, C. & Wolf, D. A Literature Review on the Prediction of Pedestrian Behavior in Urban Scenarios. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC* **2018-Novem,** 3105–3112 (2018).

24. Ismail, K. A. *Application of computer vision techniques for automated road safety analysis and traffic data collection* PhD Dissertation (University of British Columbia, 2010).

25. Zaki, M., Sayed, T., Ismail, K. & Alrukaibi, F. Use of computer vision to identify pedestrians' nonconforming behavior at urban intersections. *Transportation Research Record,* 54–64. ISSN: 03611981 (2012).

26. Wang, Y., Member, S. & Mori, G. Human Action Recognition by Semi-Latent Topic Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31,** 1762–1774 (2009).

27. Bali, S. & Tyagi, S. A review of vision-based pedestrian detection techniques. *International Journal of Advanced Studies of Scientific Research* **3** (2018).

28. Dalal, N. & Triggs, B. *Histograms of oriented gradients for human detection* in *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005* **I** (2005), 886–893. ISBN: 0769523722.

29. Li, J., Huang, L. & Liu, C. People counting across multiple cameras for intelligent video surveillance. *Proceedings - 2012 IEEE 9th International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2012,* 178–183 (2012).

30. Wang, X., Han, T. X. & Yan, S. *An HOG-LBP human detector with partial occlusion handling* in *2009 IEEE 12th international conference on computer vision* (2009), 32–39.

31. Viola, P., Way, O. M., Jones, M. J. & Snow, D. Detecting Pedestrians Using Patterns of Motion and Appearance. *International Journal of Computer Vision* **63,** 153–161 (2005).

32. Preusser, D. F., Wells, J. K., Williams, A. F. & Weinstein, H. B. Pedestrian crashes in Washington, DC and Baltimore. *Accident Analysis and Prevention* **34,** 703–710 (2002).

33. Chavis, C., Lee, Y.-J. & Dadvar, S. *Analysis of Bicycle and Pedestrian Crash Causes and Interventions* tech. rep. (District Department of Transportation, 2018).

34. Bureau, U. C. *Data Releases* Census.gov. Section: Government. `https://www.census.gov/programs-surveys/acs/news/data-releases.html` (2023).

35. *About Us* Baltimore City Department of Transportation. `https://transportation.baltimorecity.gov/about-transportation` (2023).

36. *Cube Root | Home* `https://cuberootinc.com/` (2023).

37. Chavis, C., Bhuyan, I. & Cirillo, C. *Shared Bus-Bike Lane Safety Analysis: Assessing Multimodal Access and Conflicts* tech. rep. UMEC-013 (Urban Mobility & Equity Center, July 1, 2022). `https://rosap.ntl.bts.gov/view/dot/65819` (2023).

38. Everitt, B. S., Landau, S., Leese, M. & Stahl, D. *Cluster Analysis* 1st ed. ISBN: 978-0-470-74991-3 978-0-470-97781-1. `https://onlinelibrary.wiley.com/doi/book/10.1002/9780470977811` (2023) (Wiley, Jan. 7, 2011).

39. Johnson, R. A. & Wichern, D. W. in *Encyclopedia of Statistics in Quality and Reliability* _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470061572.eqr239 (John Wiley & Sons, Ltd, 2008). ISBN: 978-0-470-06157-2. `https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470061572.eqr239` (2023).

40. Analytics, P. *How to Perform Hierarchical Clustering using R | R-bloggers* `https://www.r-bloggers.com/2017/12/how-to-perform-hierarchical-clustering-using-r/` (2023).

41. *Home* OpenCV. `https://opencv.org/` (2023).

42. Wang, C.-Y., Bochkovskiy, A. & Liao, H.-Y. M. *YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors* 2022. arXiv: `2207.02696 [cs.CV]`.

43. Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. *You Only Look Once: Unified, Real-Time Object Detection* 2016. arXiv: `1506.02640 [cs.CV]`.

44. Wojke, N., Bewley, A. & Paulus, D. *Simple Online and Realtime Tracking with a Deep Association Metric* 2017. arXiv: `1703.07402 [cs.CV]`.

45. *OpenCV: Geometric Image Transformations* `https://docs.opencv.org/3.4/da/d54/group__imgproc__transform.html` (2023).

46. Gellert, W., Hellwich, M, Kästner, H & Küstner, H. *The VNR concise encyclopedia of mathematics* (Springer Science & Business Media, 2012).

47. Turner, S., Sandt, L., Toole, J., Benz, R. & Patten, R. Federal Highway Administration University Course on Bicycle and Pedestrian Transportation. Publication No. FHWA-HRT-05-133. *US Department of Transportation* (2006).

48. *Model Documentation - Travel Demand Modeling — Metropolitan Washington Council of Governments* Accessed: 2023-08-24. `https://www.mwcog.org/transportation/data-and-tools/modeling/model-documentation/`.