DEVELOPMENT OF LIDAR-BASED METHODS FOR TRAFFIC MONITORING AND SURROGATE SAFETY ANALYSIS

Ehsan Nateghinia

Department of Civil Engineering

McGill University, Montreal

January 2024

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy in Engineering.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Automated traffic and surrogate safety monitoring at urban intersections commonly use visual-spectrum video-based systems, which offer flexibility and low cost. However, video-based systems face challenges in low-light conditions and accurate distance measurement, requiring manual calibration for precise trajectories mapping in x-y coordinates. Recently, 3D LiDAR-based methods have emerged as an alternative, overcoming these limitations. LiDAR operates effectively in low light and provides direct and calibration-free measurements in x-y-z coordinates. LiDAR sensors yield more accurate spatial data with higher resolution and range than video-based systems. Additionally, LiDAR's point cloud data can accurately represent road users' shapes in three dimensions, enhancing the precision of distance and size measurements crucial for traffic and road safety applications.

The general objective of this thesis is to develop and evaluate various LiDAR-based methods for automated traffic monitoring and surrogate safety analysis. First, a supervised learning methodology is developed using point cloud data from low-resolution and high-resolution rotational LiDAR sensors. This methodology includes background modeling, foreground detection, clustering, road user classification, and trajectory construction. The proposed methodology is calibrated and tested at various urban intersections. The detection accuracy of low- and high-resolution LiDARs is 89.9% and 94.2%, respectively, with road user classification rates of 0.91 and 0.95. The average absolute percent difference of high- and low-resolution LiDAR counts compared to manual video counts is 6% and 13%, respectively. High-resolution LiDAR shows notable potential for urban intersection traffic monitoring.

Second, a novel method is developed for determining surrogate safety indicators, such as Time-to-Collision and Post-Encroachment Time, using shape and trajectory data from the developed 3D LiDAR systems in the first study. This approach utilizes the proximity of road users' point clouds, offering an alternative to conventional trajectory-based analysis. For pedestrian-vehicle and cyclist-vehicle interactions with a safety threshold of under 10 seconds, the shape-based results align with the centroid-based method using buffer sizes of 2 meters and 2.5 meters, respectively. However, using the same setting, the centroid-based method significantly overreports critical conflicts occurring in under 1.5 seconds between vehicles and vulnerable road users.

Third, an unsupervised methodology is proposed for safety monitoring and trespassing detection at a road/railway level crossing, utilizing a low-resolution 3D LiDAR sensor. The methodology includes road user detection, clustering, and tracking. The classification uses the shape, speed, and geo-location of road users. The average absolute percentage deviation for counting motorized road users is 5% and 3%, and for non-motorized road users is 10% and 14% on two test days. Overall, the system demonstrated high performance in detecting trespassing.

Finally, a novel methodology for computing bicycle-flow parameters is proposed, leveraging a LiDAR system composed of two single-beam sensors. A machine-learning model is implemented to enhance the accuracy of speed measurements. The LiDAR methodology calculates headway and spacing between consecutive cyclists using timestamped detections and speed measurements. A comparison is made with ground truth video data to assess the accuracy of the proposed methodology. The Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) of a Neural Network employed for speed estimation are measured at 0.61 m/s and 7.1% on the test set.

RÉSUMÉ

La surveillance automatisée de la circulation et la sécurité routière aux intersections urbaines reposent généralement sur des systèmes vidéo utilisant le spectre visible, qui offrent flexibilité et faible coût. Cependant, ces systèmes rencontrent des défis dans des conditions de faible luminosité et pour la mesure précise des distances. De plus, ils nécessitent une calibration manuelle pour une cartographie précise des trajectoires en coordonnées x-y. Récemment, les méthodes basées sur le LiDAR 3D ont émergé comme une alternative prometteuse, surmontant ces limitations. Le LiDAR fonctionne efficacement dans des conditions de faible luminosité et fournit des mesures directes et précises en coordonnées x-y-z sans nécessiter de calibration. Les capteurs LiDAR produisent des données spatiales plus précises avec une résolution et une portée supérieures à celles des systèmes vidéo. De plus, les données de nuage de points générées par le LiDAR peuvent représenter avec précision les formes des usagers en trois dimensions, améliorant ainsi la précision des mesures de distance et de dimensions, cruciales pour les applications en sécurité routière et gestion du trafic.

L'objectif général de cette thèse est de développer et d'évaluer des méthodes alternatives basées sur le LiDAR pour la surveillance automatisée du trafic et l'analyse de sécurité routière substitutive. Tout d'abord, une méthodologie d'apprentissage supervisé est développée en utilisant les données de nuage de points provenant de capteurs LiDAR rotatifs à basse résolution et à haute résolution. Cette méthodologie comprend la modélisation de l'arrière-plan, la détection du premier plan, le regroupement, la classification des usagers de la route et la construction de trajectoires. La méthodologie proposée est calibrée et testée à différentes intersections urbaines. La précision de détection des LiDAR à basse et haute résolution est respectivement de 89,9 % et 94,2 %, avec des taux de classification des usagers de la route de 0,91 et 0,95. La différence moyenne en pourcentage absolu entre les comptages LiDAR haute résolution et basse résolution par rapport aux comptages vidéo manuels est de 6% et 13%, respectivement. Le LiDAR à haute résolution présente un potentiel important pour la surveillance du trafic aux intersections urbaines.

Deuxièmement, une nouvelle méthode est développée pour déterminer des indicateurs de sécurité substituts, tels que le temps de collision et le temps post-empiètement, en utilisant les données de forme et de trajectoire des systèmes LiDAR 3D développés lors de la première étude. Cette approche utilise la proximité des nuages de points des usagers de la route, offrant une alternative

à l'analyse conventionnelle basée sur la trajectoire. Pour les interactions piéton-véhicule et cycliste-véhicule avec un seuil de sécurité de moins de 10 secondes, les résultats basés sur la forme correspondent à ceux obtenus par la méthode basée sur le centroïde en utilisant des tailles de tampon de 2 mètres et 2,5 mètres, respectivement. Cependant, en utilisant le même paramètre, la méthode basée sur le centroïde surestime significativement les conflits critiques, survenant en moins de 1,5 seconde, entre les véhicules et les usagers vulnérables de la route.

Troisièmement, une méthodologie non supervisée est proposée pour la surveillance de la sécurité et la détection des intrusions à un passage à niveau routier/ferroviaire, en utilisant un capteur LiDAR 3D de basse résolution. Un algorithme non supervisé est développé pour détecter les intrusions à un passage à niveau routier/ferroviaire à Montréal, Canada. La méthodologie comprend la détection, le regroupement et le suivi des usagers de la route. La classification utilise la forme, la vitesse et la géolocalisation des usagers de la route. L'écart absolu moyen en pourcentage pour le comptage des usagers de la route motorisés est de 5 % et 3 %, et est de 10 % et 14 % pour les usagers de la route non motorisés pour les deux jours de test. Dans l'ensemble, le système a démontré une haute performance dans la détection des intrusions.

Enfin, une nouvelle méthodologie pour calculer les paramètres du flux cycliste est proposée, utilisant un système LiDAR composé de deux capteurs à faisceau unique. Un modèle d'apprentissage automatique est mis en œuvre pour améliorer la précision des mesures de vitesse. La méthodologie LiDAR calcule l'intervalle et l'espacement entre les cyclistes consécutifs en utilisant des détections horodatées et des mesures de vitesse. Une comparaison est faite avec les données vidéo de validation pour évaluer la précision de la méthodologie proposée. L'erreur quadratique moyenne et l'erreur de pourcentage absolue moyenne d'un réseau neuronal utilisé pour l'estimation de la vitesse sont mesurées à 0,61 m/s et 7,1 % sur l'ensemble de test.

ACKNOWLEDGMENTS

I would like to thank my supervisor, Luis Miranda-Moreno, for his guidance throughout my Ph.D. Without his knowledge and steady support, the completion of my thesis would not have been possible.

I owe thanks to a very special person, my wife, Hasti, whose unwavering love and support have been my guiding light and source of strength throughout this journey.

There are several other people and agencies I would like to acknowledge, each of which played an important role in my research:

- Asad Lesani and David Bietel for their collaboration in Chapter 5 of this thesis.
- Ce Zhang, Wooseok Do, Alejandro Pérez Villaseñor, Bismarck Ledezma-Navarro, Mohamed Elagaty, Adham Badran, and other McGill students who helped me with the data collection.
- Trottier Foundation and the Natural Sciences and Engineering Research Council (NSERC) for financial support during my Ph.D.
- Centre interuniversitaire de recherche sur les reseaux d'entreprise, la logistique et le transport (CIRRELT) for their generous awards.
- Transport Canada for funding one of my research projects.
- The City of Montreal for their support in data collection permission.

CONTRIBUTION OF AUTHORS

JOURNAL PUBLICATIONS

- **Nateghinia, E.**, Beitel, D., Lesani, A., & Miranda-Moreno, L. F. (2022). A LiDAR-based methodology for monitoring and collecting microscopic bicycle flow parameters on bicycle facilities. Transportation, 1-25.

JOURNAL UNDER REVIEW:

- Submitted to Transportation Research Part C: **Nateghinia, E.**, & Miranda-Moreno, L. F. (2024). A 3D LiDAR-Based Supervised Methodology for Automated Traffic Monitoring and Data Collection at Urban Intersections with High-Mixed Traffic.
- Submitted to Accident Analysis and Prevention: **Nateghinia, E.**, & Miranda-Moreno, L. F. (2024). A 3D-LiDAR-Based Methodology for Surrogate Safety Analysis at Intersections with High Non-Motorized Traffic.
- Submitted to International Journal of Transportation Science and Technology: **Nateghinia, E.**, & Miranda-Moreno, L. F. (2024). Development of an Unsupervised 3D LiDAR-Based Methodology for Automated Safety Monitoring of Railway Facilities.

PEER REVIEWED CONFERENCE PAPERS

- **Nateghinia, E.**, & Miranda-Moreno, L. (2022). A 3D-LiDAR-based Methodology for Extracting Data for Surrogate Safety Analysis at Intersections. In Proceedings of the Transportation Research Board 101st Annual Meeting, Washington, DC, USA.
- **Nateghinia, E**., & Miranda-Moreno, L. (2020). Development of a LiDAR-based System Prototype for Detection, Classification and Tracking of Road Users at Railway Facilities. In Proceedings of the Transportation Research Board 99th Annual Meeting, Washington, DC, USA.
- **Nateghinia, E.**, Beitel, D., Lesani, A., & Miranda-Moreno, L. F. (2020). Measuring Instantaneous Cyclist Speed using 1D LiDAR. In Proceedings of the Transportation Research Board 99th Annual Meeting, Washington, DC, USA.

GLOSSARY OF TERMS

Traffic Monitoring: The systematic observation of road users on transportation networks, including streets, highways, intersections, sidewalks, and bicycle facilities, utilizing sensing technology such as LiDAR, cameras, radar, infrared, Bluetooth, etc. A traffic monitoring programs aims to collect and study traffic flow and congestion and informs traffic management and transportation planning decisions.

Safety Monitoring: A systematic surveillance of transportation facilities that aims at preventing accidents and injuries by analyzing road users' behavior and identifying potentially dangerous interactions.

Light Detection and Ranging (LiDAR): A remote sensing technology that periodically emits laser light and measures distance to an object based on the receipt time of the reflected light from the object's surface, a process known as time of flight.

Roadside LiDAR: A traffic monitoring data collection technique involving the installation of a LiDAR sensor on the roadside, as opposed to mounting it on a vehicle, commonly used for applications such as Autonomous Vehicles (AVs).

1D LiDAR: A LiDAR sensing method that utilizes a single laser channel capable of measuring one distance value to an object present at or crossing in front of its line of sight.

3D LiDAR: An extended LiDAR sensing method that utilizes multiple laser channels arranged vertically and horizontally, capable of measuring multiple distance values in the azimuth and elevation planes. This technique creates a three-dimensional point cloud of objects observed by the LiDAR.

Point Cloud: A set of data points in three-dimensional space, often generated by 3D scanning technologies such as LiDAR, stereo cameras, and radars, representing the surfaces of objects or environments.

Background Modelling: A computational technique used to identify and model the stationary elements within data captured from an environment, such as video frames or continuously collected LiDAR point clouds.

Foreground Detection: A computational technique used to identify and model the dynamic elements within data captured from an environment by comparing the current observation with a previously built background model. This process aims to isolate moving objects (road users) from the stationary background in LiDAR point cloud or video frames.

Point Cloud Clustering: A category of clustering methods that groups data samples, typically 3D points with x-y-z coordinates, into clusters of neighboring points with similar characteristics. This technique segments and identifies meaningful structures defined as objects (road users) within point cloud data.

Road User Classification: A process that applies supervised or unsupervised methods to features extracted from LiDAR point clouds of individual road users. This process categorizes each road user into groups such as pedestrians, cyclists, passenger cars, and trucks.

Un-Supervised Learning: A classification method that categorizes road users' point clouds based on inherent patterns and structures within the data, without the need for predefined labels. Algorithms within unsupervised learning autonomously identify similarities and differences among road users.

Supervised Learning: A classification method that utilizes a set of features extracted from road users' point clouds, which are labeled with their actual road user class.

Road User Trajectory: A sequence of x-y coordinates representing the position of a road user's representative point at consecutive timestamps, along with timestamped velocity information in the x-y direction. This data provides a detailed record of the road user's movement over time.

Data Association: An algorithm that establishes connections between different detections of the same road user across consecutive frames.

Kalman Filter: An algorithm used in road user tracking to predict and update the positions and velocities of road users based on LiDAR detections at a given timestamp.

Traffic Conflict: An interaction in time and space between two or more road users that can create a risk of collision.

Surrogate Safety Measure: An alternative to directly measuring crash frequency, it’s typically defined by observing conflicts in interactions between road users or dangerous road user behaviors like speeding.

Time-to-Collision (TTC): TTC measures the time duration before a potential collision between two road users, considering their current motion patterns, speed, and acceleration remain unchanged.

Post-Encroachment Time (PET): PET is a critical surrogate safety indicator that heavily relies on spatial data and the distance between road users. Defined as the temporal gap between two road users traversing an intersection, PET quantifies the time between one road user leaving a designated area and another entering it.

CHAPTER 1.

INTRODUCTION

CHAPTER 1: INTRODUCTION

1.1 Background

Intelligent Transportation Systems (ITS) integrates cutting-edge computing, sensing, and communication technologies to build a connected transportation framework. ITS supports real-time automated traffic monitoring and data collection, adaptive response to dynamic conditions, and proactive safety approaches (1). The main components of ITS include sensing technologies referred to as sensors, communication systems, traffic control and management systems, connected and autonomous vehicles, and safety-related devices such as collision avoidance systems (2, 3).

Through active and passive frameworks, ITS leverages various sensing technologies to collect traffic-related data, including traffic volume, average speed, and travel times (4). Active data collection relies on users to initiate the process. These resources include data collected through the user's GPS navigation system while traversing the transportation network, mobile applications that allow users to participate in a data collection program, Bluetooth/Wi-Fi, and pedestrian traffic signal activation (5). On the other hand, passive data collection methods do not require user involvement. Passive frameworks' data collection technologies range from large-scale solutions like GPS-enabled devices (6, 7) to micro-monitoring systems like pedestrian counters.

The primary technologies employed in passive data collection include, but are not limited to, sensor-based methods such as traffic cameras, radars, inductive loop detectors, pneumatic tubes, passive infrared sensors, weight-in-motion sensors, environmental sensors, and GPS-enabled devices. The sensors vary in capabilities, each offering one or more of these critical components: road user detection and counting, classification, and tracking.

Inductive loops are placed under arterial or highways to estimate the traffic volume for travel demand model calibration or estimating Average Annual Daily Traffic. Pneumatic tubes are suitable for collecting cyclist traffic flow (8). Passive infrared sensors are installed at crosswalks for pedestrian counting (9). Radar-based systems measure speed and spacing between vehicles and provide real-time data on average highway speed, congestion, and travel time (10). The cameras-based systems are primarily used at intersections for traffic and safety studies (11) and highways for traffic flow extraction (12). Other camera-based monitoring systems may incorporate thermal or stereo cameras. Generally, these two systems have garnered less attention in the context of traffic monitoring compared to visual-spectrum cameras. This lower implementation rate is primarily attributed to their higher cost or the need for more significant processing power (13-15).

In ITS, Light Detection and Range sensors, referred to as LiDAR, were first frequently associated with Autonomous Vehicles (AV) (16). Due to their success as part of AVs' vision system, their application as traffic data collection technology emerged. LiDAR sensors are now tested at fixed locations for traffic data collection, commonly known as roadside LiDAR systems (17, 18). Roadside LiDAR systems have gained attention as a potential alternative or complement to camera-based systems. LiDAR, similar to cameras, can detect, classify, and track road users in complex transportation facilities. One distinctive advantage of LiDAR technology over cameras is its ability to precisely measure distances in a spherical coordinate system, which is convertible to Cartesian coordinates. This capability provides a comprehensive three-dimensional representation of the traffic data collection environment that can quantify road users' movement (19).

LiDAR sensors exhibit significant differences in resolution and field of view compared to camera-based systems. These characteristics have potential advantages and disadvantages. The differences in LiDAR's parameters allow for the selection of the lower resolution for active users' data collection programs such as pedestrian counting (6) or cyclist monitoring (7). On the other hand, the extendibility of results from one research to another is often not as straightforward as a camera-based system. More importantly, labeled samples in one study may not be fully compatible with another study unless the characteristics of the LiDAR sensors in the two studies are similar.

This thesis adopts the direction of developing and evaluating a LiDAR-based methodology for automated monitoring of real-world traffic scenarios, targeting objectives such as monitoring traffic at intersections, monitoring cyclists in bike lanes, and evaluating LiDAR efficacy in surrogate safety analysis. The directional approach of this research encompasses several key steps: a comparative analysis of LiDAR sensors to identify the impact of LiDAR resolution for specific tasks, the development of an unsupervised and supervised automated methodology for LiDAR data processing, an extensive data collection program tailored to each application, a large scale semi-automated routine for road users labeling in LiDAR point cloud data, and testing and performance evaluation of the proposed LiDAR-based systems.

1.2 Research Motivation

1.2.1 Traffic monitoring

Most traffic data collection efforts and resources primarily target vehicular traffic volume data on highways and arterials, encompassing traffic volume, travel time, average speed, and congestion levels. However, in urban environments, where the scenario includes vehicular traffic and active road users, the data collection process often lacks the same level of attention and detail. Urban intersections, in particular, present a complex environment where different modes of transportation intersect. The challenges in these settings are multifaceted, involving factors such as the density of pedestrians, vehicle flow, and cyclist movements. Traditional traffic monitoring systems, mainly focused on vehicular data, fail to capture the full spectrum of the traffic complexity in urban transportation facilities such as intersections, bike lanes, sidewalks, and level crossings.

Traffic monitoring at intersections heavily relied on the use of camera-based systems. There have been various applications of camera-based systems for achieving the same objective. However, there are certain limitations to the video-based approach to traffic monitoring at intersections, including degraded performance in low-light conditions, the necessity for manual geometric calibration at each intersection, and the inability to provide accurate distance measurements and 3D dimensions of road users (20-22). Alternatively, the LiDAR system can be utilized to overcome the shortcomings of camera-based systems. Using LiDAR-based systems for traffic monitoring and data collection is a relatively new approach and is seen as an emerging technology. This suggests that research focusing on implementing and evaluating LiDAR-based systems for these applications will likely expand rapidly.

Few key studies have implemented fixed-location roadside LiDAR for traffic monitoring. Tarko et al. developed a LiDAR-based system for intersection monitoring, employing a super high-resolution LiDAR with 64 laser channels (23). The system is calibrated with an unsupervised algorithm, where road users are obtained through background subtraction and classified based on their speed and dimensions. The classification stage considers the part of the intersection where the road user has been detected, such as the sidewalk, crosswalk, or street. The accuracy of counting and classifying 504 road users was 98% (23). Zhao et al. developed a 16-laser channels LiDAR-based traffic monitoring system (18). The system utilizes a methodology involving background filtering, clustering, and classification to detect and track pedestrians and vehicles at intersections. The classification is performed using a neural network with three features: the total number of points, the 2D distance of the user to the LiDAR sensor, and the direction. The reported accuracy of classification is approximately 93%. However, other road users are not discussed (18).

In addition to traffic monitoring at intersections, monitoring and extracting traffic flow parameters of active transportation users at cyclist facilities is highly important. Understanding microscopic bicycle flow relationships is essential. Yet, limited research focuses on real-time automated extraction of basic bicycle flow parameters (24, 25). Existing point-based monitoring systems primarily count cyclists but lack measurement of vital microscopic flow parameters like speed, density, headway, and spacing. Various technologies, including Radar, Pneumatic tubes, and inductive loop sensors, are implemented for cyclist monitoring applications to measure cyclist speed (26-28). GPS data has also been used to estimate cyclist speed and delay at intersections (29). Gathering GPS trajectory data allows researchers to study traffic flow fundamental diagrams for cyclists (30). Using image processing and Artificial Intelligence, video-based systems track cyclists and estimate their speed from trajectory data. The feasibility of video analysis for bicycle data collection, including counts, density, and speed, has been demonstrated (31, 32). Studies have investigated cyclist maneuvers and interactions using video-derived trajectories (33).

1.2.2 Surrogate traffic safety

Traditionally, the diagnosis of safety issues and the recommendation of appropriate countermeasures use historical crash data. While studying collision frequencies provides direct insight into road safety, proactive approaches are essential.

Camera-based systems are among the most common automatic traffic monitoring and safety analysis systems. The applications of video-based systems in road safety analysis using conflict analysis and surrogate safety indicators are studied in the literature. Fu et al. studied pedestrians' safety in interactions with vehicles at non-signalized crosswalks by analyzing vehicle yielding and pedestrian crossing decisions obtained from a video-based monitoring system (34) and secondary interactions with cars exiting the intersection (35). Zangenehpour et al. showed that cycle-track intersections appear safer than those without cycle tracks. They employed trajectories obtained from an automatic video-based monitoring system to investigate the interactions between cyclists and turning vehicles using PET measure (36). Seyed et al. used a computer vision algorithm to automatically detect vehicle-bicycle conflicts and rank them based on the severity of interactions using the Time-To-Collusion safety indicator (37).

A systematic review of the literature on the application of LiDAR in surrogate safety-based conflict analysis reveals that the primary body of work employing LiDAR-based monitoring systems for this application has been conducted by one research group. This conclusion is drawn from several recent survey articles reviewing surrogate safety analysis methods (38-40). This group has developed a methodology using a low-resolution LiDAR system for roadside traffic monitoring (41-44). These works study the interactions between pedestrian and vehicular traffic using the trajectory data obtained from the LiDAR system. In another study, a high-resolution LiDAR-based system was installed and tested at three intersections, and the trajectories of road users were incorporated to identify near-miss events and compute time-to-collision indicators (23).

However, it is essential to note that LiDAR systems are extensively used as a component of the vision systems in Autonomous Vehicles. In this context, studies have specifically focused on the performance of such LiDAR systems in collision scenarios and reconstructing accidents (45).

In addition to the safety analysis of road users at intersections, there is a growing concern for vulnerable road users' interaction with trains at railroad grade crossings. In Canada, pedestrian-related collisions on railway facilities account for 59% of all railway-related deaths (46). Efficient rail facility monitoring depends on solutions incorporating alternative technologies such as cameras, thermal cameras, stereo cameras, ultrasonic sensors, active or passive infrared sensors, radars, and LiDAR sensors to detect trespassing incidents (47-50). The application of radar and cameras proves to be a challenge in monitoring a large coverage area for an extended period. Specifically, they are prone to high false alarms (50, 51).

A limited number of studies have applied LiDAR to railway safety. A 2D LiDAR, installed horizontally for level crossing monitoring, achieved a 99.25% detection rate but had limitations, including missing pedestrians behind vehicles and lacking classification and tracking capabilities (52). In contrast, 3D LiDAR systems offer a more comprehensive monitoring range. One study implemented a 3D LiDAR system to monitor a level crossing. However, the results primarily consisted of visualized 3D point clouds, and the study lacked a comparative analysis (49).

1.2.3 Research needs

The research motivations underlying this work are based on the recognized limitations within existing literature, including:

1. **LIDAR sensing technologies in urban traffic applications:** There is a potential need to explore LIDAR sensing technologies for traffic monitoring and road safety applications in urban settings with high pedestrian and cyclist volumes, particularly in mixed traffic conditions. The current implementation of LiDAR systems is limited in terms of evaluating system performance under different traffic conditions. Conventional visual-spectrum video cameras and computer vision methods face challenges addressing occlusion issues in these high-density scenarios. Additionally, an accurate calibration of camera-based systems is essential for their application to accurately reconstruct road user trajectories and movement patterns at intersections.
2. **Impact of sensor resolution:** The current state of LiDAR application in traffic and safety monitoring suggests an insufficient investigation into the influence of sensor resolution on detection and classification using substantial datasets. Alternative sensor characteristics, including the number of channels, may significantly affect outcomes, and this aspect deserves a more comprehensive exploration.
3. **Diversity of approaches:** There is a noticeable gap in the literature regarding comparative studies of different approaches, such as unsupervised and supervised methods, in LiDAR-based traffic monitoring methodologies. A systematic investigation into these methods is yet to be extensively undertaken, which could potentially yield novel insights and advancements in this field. In contrast to camera-based systems, LiDAR data, especially in applications like traffic monitoring at intersections, are not as readily available. Only a few LiDAR datasets have been compiled, often using costly high-resolution LiDAR systems, primarily for applications in Autonomous Vehicles.
4. **Machine learning methods:** The diverse array of machine learning methods available for detection and classification tasks. A thorough examination of various machine learning, including unsupervised and supervised learning techniques, is essential for identifying the most effective and efficient solutions for traffic and safety monitoring challenges.

5. **Exploration of 1D LiDAR sensors**: LiDAR systems, specifically 1D and 2D LiDAR, offer advantages over video analysis, including lower data volume, simplified data processing, reduced computational power requirements, and cost-effectiveness for real-time applications. The potential of 1D LiDAR sensors in applications related to monitoring active transportation users is noteworthy, particularly in scenarios where cost-effective solutions are essential. This avenue of research holds promise for expanding the range of feasible and economical sensor solutions in relevant contexts.
6. **Application of 3D LiDAR for safety monitoring at intersections**: There has been very little research and practical implementation of LiDAR systems for surrogate safety analysis of traffic conflict. Specifically, the performance of LiDAR systems in identifying conflicts at various intersections, particularly under medium to high traffic conditions, remains understudied. It is necessary to investigate how using 3D shape data of road users, instead of using the trajectory of centroids, influences the outcome of such studies. Evaluating the impact of this 3D data could provide deeper insights into the effectiveness of LiDAR in complex traffic environments and significantly contribute to improving intersection safety monitoring.
7. **Application of 3D LiDAR for safety monitoring at level crossing:** In level crossing monitoring applications, accurate detection, especially with a low false alarm rate, is paramount, as false alerts can disrupt train operations. While alternative systems offer advantages, certain limitations are noteworthy. Visual spectrum camera-based systems are significantly affected by low-light conditions. Unlike typical traffic data collection programs, monitoring the operation and safety of railway facilities is a 24-hour task, including prioritizing nighttime operations. Infrared, ultrasonic, and radar sensors struggle with accurate user classification and are vulnerable to adverse weather conditions and precipitation, requiring long-term implementation and high maintenance considerations. Moreover, the larger coverage area for railway monitoring compared to intersections poses challenges for traditional camera-based systems. 3D LiDAR sensors, with their extensive field of view, are well-suited for such applications. Only a few studies are exploring LiDAR technology's application in railway safety, but existing literature highlights significant benefits over conventional methods like camera-based systems.

1.3 Objectives

This research aims to develop LIDAR-based methodologies using supervised and unsupervised algorithms and evaluate their performance for traffic monitoring and road safety applications in urban environments characterized by mixed traffic conditions. The study intends to assess the capability of LiDAR systems in accurately detecting, classifying, and tracking various road user types, including vehicles, cyclists, and pedestrians.

The specific objectives of this research are to:

1. Develop and validate a supervised methodology for processing three-dimensional LiDAR data to detect, classify, and track road users in mixed and high-density traffic conditions, including non-signalized and signalized intersections.
2. Extend the LiDAR-based traffic monitoring system methodology to surrogate safety applications leveraging point cloud proximity measures in combination with established methods based on road users' trajectory analysis.
3. Develop an unsupervised methodology for processing low-resolution three-dimensional LiDAR data and evaluate its performance for monitoring traffic and safety at road-railway level crossings.
4. Propose and assess a method using a one-dimensional LiDAR-based system for monitoring bicycle facilities and measuring microscopic traffic flow parameters, including speed, density, and cyclist volume.

1.4 Original Contributions

This thesis introduces a comprehensive set of methodologies utilizing LiDAR sensors for automated traffic and road safety monitoring. These methodologies are categorized into two primary components. The first involves developing road user detection techniques for 1D and 3D LiDAR and enhancing classification and tracking capabilities specifically for 3D-LiDAR sensors. The second component focuses on developing surrogate safety measures derived from LiDAR-based trajectories and point cloud proximity analysis, such as traffic conflict techniques.

This thesis contributes to the existing literature by addressing some of the shortcomings by:

1. Developing a supervised learning methodology leveraging 3D LiDAR technology for traffic monitoring. The research method introduces scalable 3D processing techniques for detecting, classifying, and tracking road users. The system's methodology is extensively tested at various intersections. The methodology is designed and implemented for two LiDAR sensors with low-resolution and medium resolution. A semi-automated routine for labeling point clouds is introduced as part of developing a supervised learning methodology for LiDAR data processing. This routine leverages the unique characteristics of traffic monitoring at intersections to sample and label pedestrian samples on the sidewalk, cyclist samples on bike lanes, and car samples on the street.
2. Developing a methodology for surrogate safety analysis based on 3D LiDAR system output. The method involves introducing a reliable approach to compute time-to-collision and post-encroachment time using the proximity of the LiDAR point cloud instead of the centroid of the road users' trajectory.
3. Developing an unsupervised learning methodology for processing point cloud data of a low-resolution 3D LiDAR technology with an application in railway level crossing safety monitoring. The system is tested in a level crossing with a high volume of pedestrians.
4. Developing a methodology for processing and analyzing the one-dimensional LiDAR Data for cyclist monitoring applications. This methodology estimates microscopic cyclist flow parameters, including volume, speed, density, and headway.

1.5 General Literature Review of Traffic Data Collection Technologies

A traffic data collection program encompasses diverse sensing technologies tailored to various needs and applications in transportation engineering, planning, and travel demand modeling. Some important traffic data collection programs focus on gathering data for continuous traffic monitoring, pedestrian and cyclist counting (6, 7), safety monitoring and analysis, and automated parking management (53). Other programs are designed to collect trip-level data, such as methods for tracking travel times (29), conducting Origin-Destination (OD) surveys (54), and collecting data related to public transit.

Automated techniques and solutions are fundamental in traffic monitoring and surrogate safety analysis within transportation engineering. Data extraction varies based on application, employing real-time algorithms or post-processing methods. A wide range of alternative technologies and methodologies has been explored in academic literature to fulfill these objectives.

The body of literature encompassing traffic monitoring and data collection techniques is extensive. Numerous comprehensive review articles provide in-depth overviews of the available methods, as well as their impacts and limitations. These sources serve as valuable references for a thorough understanding of the field (3, 55-58).

However, this section provides a general overview of the existing framework for data collection for traffic monitoring and safety analysis. Subsequently, it presents an overview of the applications of LiDAR sensing technology in civil and transportation engineering and an overview of some of LiDAR data processing techniques in a general context. The detailed literature review fundamental to each chapter of this thesis is reported in the corresponding chapter. It is important to note that there may be minor overlaps in content, particularly in the individual literature reviews of each chapter. This is to ensure consistency and coherence across the thesis, and as such, occasional reiteration of specific methods or works may occur.

Overview of data collection methods for traffic monitoring

Several factors are used to categorize traffic data collection methods. In the early stages, traffic data collection heavily relied on individuals counting vehicles, pedestrians, or cyclists. This method was direct but limited by human capacity and prone to errors. Several in-field mechanisms, such as tally counters or software, were developed to facilitate and increase the reliability of the manual counting process (59).

With the advancement of technology, traffic data collection programs have shifted toward semi-automated or fully automated methods and technologies. Fully automated techniques leverage computing resources and a variety of sensors, such as inductive loops, pneumatic tubes, infrared sensors, Radar, cameras, and LiDARs, to achieve a more sophisticated traffic data inventory.

Data collection technologies are typically categorized according to the type of data they provide, the primary purpose of the data collection, the location where data is gathered, the duration of the data collection program, the granularity of the data, and, most crucially, the coverage area of the specific technologies.

The primary locations for data collection are typically highways and urban areas, with a particular emphasis on intersections. The scope of coverage for these techniques is categorized into point-based (fixed location) systems, point-to-point (between points) systems, and area-wide traffic monitoring systems. In point-to-point and area-wide coverage, the choice of technology dictates the extent and scalability of the area under study.

- **Point-based technology:**

The primary objective of point-based technology is to detect and count road users passing through its limited coverage area. These systems are designed to collect traffic flow information at specific, targeted locations. It is necessary to install multiple point-based systems across various locations to achieve network-wide data coverage.

The choice of sensing technology for a point-based system largely depends on its intended application. In the context of monitoring motorized traffic, commonly used technologies include loop detectors, pneumatic tubes, and Radar, which are particularly effective on highways and arterial roads (26).

Furthermore, point-based technology is essential for monitoring non-motorized traffic, particularly in urban settings. For cyclist traffic monitoring, adapted and customized technologies such as loop detectors and pneumatic tubes are used to capture and analyze bicycle movements effectively (27).

Similarly, pedestrian counting systems often employ passive infrared sensors and LiDAR (6, 7). These technologies are selected for their ability to detect and count pedestrians in a variety of environmental conditions, ensuring reliable data collection essential for understanding and managing urban traffic flow.

Camera-based traffic monitoring systems can monitor both motorized and non-motorized traffic. They effectively capture vehicle movements in motorized traffic, supporting highway counting and speed analysis (15). The same systems adeptly monitor pedestrians, cyclists, and vehicles in urban settings, distinguishing between different road users (32, 60). Because of this adaptability, camera-based systems are invaluable in various traffic monitoring applications.

- **Between-point systems:**

The main focus of the technologies used in between-point systems is to collect mesoscopic traffic data such as point-to-point travel time, average speed, and an estimated travel demand for a limited pair of origin destinations.

Bluetooth and Wi-Fi technologies capture and trace the MAC address of specific users within an extended traffic study network (61). A critical consideration when implementing a point-to-point system is to expand the coverage area by embedding probe devices throughout the transportation network (62). Based on license plate recognition, camera monitoring systems can also trace road users anonymously across an extensive network of highway cameras (63). The primary consideration for this type of camera system is to avoid associating plate information with specific trip characteristics, such as origin, destination and stops.

- **Area-wide systems:**

The third category of data collection methods involves technologies suitable for area-wide monitoring. Key technologies that enable extensive coverage in traffic monitoring encompass GPS-enabled systems (64) and data from mobile phone service providers (65). These technologies can track users' real-time geo-location throughout the transportation network. Additionally, large-scale wireless networks and service providers are categorized as area-wide systems in the context of traffic monitoring (65).

The data collection methods utilizing these technologies are further categorized into passive and active approaches, determined by the road user's involvement in data collection. A technique is considered active if the road user is assigned a vehicle probe device or participates in a data collection program using a smartphone application. In contrast, passive methods do not require direct user involvement; instead, they anonymously log road users' trips on the network through appropriate receivers linked to the system's technology (66).

- **Vision-based systems:**

Technologies utilizing various vision sensory systems are frequently employed as fixed-location systems for traffic monitoring. This category includes camera-based systems, modern Radar, and LiDAR. Although these technologies operate on fundamentally different principles, their data processing often involves computer vision and machine learning algorithms.

Video-based computer vision technologies and methods are particularly prevalent in traffic applications. These are implemented for traffic monitoring on highways and intersections, as well as for monitoring active transportation, such as cycling facilities and sidewalks (56, 67-69).

Modern radars, especially those based on phased array technology, can perform 3D scanning in Cartesian coordinates and measuring speed. However, this advancement is relatively recent, and current implementations are primarily focused on integrating these radars as components in the vision systems of autonomous vehicles (AVs) (70, 71).

LiDAR-based systems are increasingly recognized in traffic monitoring, driven by the development of smaller, more cost-effective units and their extensive use in autonomous vehicles (AVs). LiDAR sensors, known for their accuracy in distance measurement, eliminate manual calibration and do not require the geometric computations necessary for cameras, thus providing distortion-free data. However, the relatively higher cost of LiDAR sensors than cameras has led to a scarcity of labeled data for extensive research. LiDAR is an emerging technology in transportation, with applications in both traffic monitoring and Autonomous Vehicles. So far, a few studies have attempted to develop LiDAR-based methodologies for traffic monitoring (72, 73).

LiDAR, 3D radars, and camera-based systems share common characteristics in traffic monitoring applications. They can produce high-resolution measurements of their surroundings. Designed to operate in real-time, they adeptly detect various road users. LiDAR and 3D radars accurately record distance and speed, respectively. Most importantly, all three technologies are often integrated with computer vision and machine learning algorithms, offering extensive opportunities for development and research. Key components of a vision-based methodology applicable to these systems include road user detection, classification, and tracking. However, camera-based systems have an advantage due to the vast access to pre-labeled data (images) from real-world traffic scenarios.

Review of data collection methods for safety studies

The Federal Highway Association's Traffic Conflict Technique (TCT) manual identifies six conflict categories, each potentially containing multiple conflict types: same direction, opposing left turn, cross-traffic, right-turn-on-red, pedestrian, and secondary (an evasive maneuver endangering the third user) (74). Finding these traffic events needs a traffic conflict survey at the site. A traffic conflict survey could have several objectives, such as safety diagnosis, identifying hazardous sections, and before/after study to investigate the performance of the safety program (75). Methods for collecting traffic conflict data encompass field observation, manual video observation, automated video analysis, and a vehicle-equipped approach (76).

Traditional traffic conflict techniques relied on manual surveys conducted by observers. As part of those studies, numerous manuals have been developed to instruct observers on conflict types and characteristics (74, 75). Manual conflict surveys involve an observer in the field or the manual analysis of recorded video footage. In-field observations tend to be more accurate than video ground truth because observers can account for additional factors, such as weather and road conditions, which are often challenging to discern in video footage. Nevertheless, in-field observation demands a significant amount of attention, and if observers fail to track road users' actions, there is no opportunity for a second chance to rectify that mistake (75). The primary challenge of manual conflict surveys lies in the high costs associated with employing and training observers. Additionally, the time required for the manual analysis of recorded videos and the difficulties in recognizing and differentiating actual conflicts contribute to this approach's complexity.

Advances in computer-vision techniques have enabled automated video analysis systems to serve as an alternative to manual observation surveys. This video analysis tool generates timestamped trajectories of road users, allowing the extraction of surrogate safety indicators through the study of these trajectories over time. The capabilities of automated video processing tools offer opportunities for large-scale conflict analysis (76).

Surrogate safety indicators have been extensively discussed in the literature. Sohel Mahum et al. reviewed and categorized them into four types: temporal proximity-based conflict indicators, distance-based proximal indicators, deceleration-based indicators, and other indicators (77). Temporal indicators include time-to-collision (TTC), post-encroachment time (PET), and crash index (CI) (77).

Beyond surrogate safety indicators, trajectory data can offer additional insights into the severity of conflicts. Assessing the severity of a traffic interaction between two road users involves evaluating collision and injury risks (78). Various factors, including proximity in time, proximity in space, and the speed of the involved road users, can determine collision risk. Conversely, factors like speed differences, mass differences, the relative angle of conflicts, and the vulnerability of the involved road users can control the risk of injury (78, 79).

The capacity of surrogate safety indicators to identify traffic conflicts and assess their severity has created a potential application for employing computer vision techniques in trajectory extraction and road user classification (80). Intersections, as a critical aspect of transportation facilities, have been subject to extensive safety study and examination. These studies are primarily motivated by the high frequency of diverse conflict types, particularly those involving vulnerable road users. Several computer vision-based techniques have been developed for detecting, tracking, and classifying road users at intersections.

Saunier et al. proposed a vision-based system for estimating collision probabilities using a probabilistic time-to-collision method (81). More than 300 severe interactions and collisions were automatically analyzed and categorized into four groups: head-on, rear-end, side, and parallel.

Laureshyn and Aliaksei extensively discussed the application of an automated vision system for analyzing the behaviors of road users (82). The extracted trajectories are utilized to identify potential conflicts between any pair of road users, focusing on cyclists in intersections and roundabouts. Concerns arise regarding the system's accuracy in cyclist detection, where the cyclist detection rate ranges from 18% to 72%, and the false-positive rate varies from 20% to 90%.

Pedestrian-vehicle conflicts have been studied using conflict data obtained from video (83). The automated vision system detects, tracks, and classifies pedestrian and motorized users. Then, from the trajectories, it identifies hazardous interactions and estimates the severity of conflicts using several surrogate safety indicators (83).

Vision-based systems have been employed for cyclist and pedestrian safety studies (35, 36, 84). One study focused on interactions between pedestrians and vehicles exiting intersections based on a distance-velocity model (35). Another study investigated cyclist maneuvers using trajectory data obtained from automatic video processing at locations with cycling network discontinuity (84). The third study utilizes cyclists' trajectory data to evaluate the safety effect of cycle tracks at signalized intersections (36).

Despite the popularity of automated video analysis techniques for surrogate safety, video-based systems have some limitations, including underperformance in adverse weather conditions and low visibility. Additionally, computer vision algorithms are often computationally intensive for real-time applications on affordable embedded systems. Obtaining distance measures requires a manual calibration process, and, more importantly, these measures may be less accurate if the camera is not set up under optimal conditions or if the calibration is not performed accurately. With that being said, video-camera systems remain appealing due to their flexibility for installation and low economic cost.

A review of the application of LiDAR systems in civil and transportation engineering

This section offers a concise overview of LiDAR sensing technology applications in civil engineering, specifically focusing on ITS. These systems, developed initially for various engineering concepts, have been adapted in multiple studies and applications, primarily for traffic data collection. The LiDAR-based systems discussed here are distinct from the conventional roadside LiDAR implementations covered in the previous section.

LiDAR systems have extensive applications in Airborne, Terrestrial, and Mobile Laser Scanning systems. In an Airborne Laser Scanning (ALS) system, a high-range LiDAR is installed beneath the aircraft, capturing reflectivity and distance to the Earth's surface. ALS classifies Earth's surface into various geographical regions and land covers (85, 86). While utilizing ALS for traffic monitoring proves costly, a feasibility simulation of this application has been discussed in (87). Additionally, two vehicle detection and extraction algorithms have been examined based on the segmentation of airborne laser points (88).

In a 3D Terrestrial Laser Scanning (TLS) system, a two-dimensional LiDAR is fixed to a tripod equipped with an automated or manually operated rotating system. The TLS system is utilized for land surveying and three-dimensional modeling of urban structures and road infrastructures (89, 90). However, in one study, a TLS system is employed for accident reconstruction and investigation in a road safety analysis (91).

In a Mobile Laser Scanning (MLS) system, 2D or 3D LiDAR sensors are installed on a dedicated vehicle equipped with navigation devices (92, 93). The MLS system models the 3D structure of urban facilities and constructs a road inventory database by extracting road widths, curbs, slopes, and curve radii (93, 94). The 3D segmented models of transportation facilities are utilized to assess the safety level of roads, including detecting or investigating hazardous sections (92).

Examples of road inventory databases collected by MLS include traffic sign recognition (95), traffic light detection, and streetlight pole detection (96). In traffic sign recognition applications, geometric features from the point cloud and images from the camera are used for recognition after sign detection by LiDAR (95). Due to the high reflectivity of road marking elements, the intensity values of 3D LiDAR point clouds are used to identify and extract them (97). Other tasks of an MLS system may include pavement analysis, crack detection, and pothole detection (94, 98). Utilizing an MLS system for parking occupancy detection involves detecting street boundaries and parked cars along the street side (99).

Autonomous Vehicle (AV) technology, another applied field of research, utilizes 3D LiDAR sensors mounted on a vehicle. An AV may incorporate more than one, and in some cases, up to five 3D LiDARs—one positioned on the top of the vehicle and one in each corner LiDAR sensors, collaborating with cameras and radars, construct a robust visual perception system for the AV, offering an accurate understanding of the objects and environment surrounding it. Like an MLS system, an AV builds a 3D model of the street environment to accurately detect intersections, road alignment, and traffic signs. The most critical task of an AV's vision system is detecting pedestrians, cyclists, and cars, contributing significantly to the safety of both non-motorized and motorized transportation users (16-18).

A review of algorithms for LiDAR data processing

The first step of the LiDAR processing systems involves background modeling, which is achievable through various approaches. The first approach treats the 2D matrices of distance values as images. The background is constructed by averaging frames, allowing for the use of sophisticated algorithms like the Gaussian Mixture Model (100).

The second approach involves converting each distance value to the 3D Cartesian coordinate system and storing an initial set of point clouds in the Cartesian coordinates as a reference to the background model. Subsequently, a 3D segmentation clusters the point cloud and determines whether each point corresponds to the background. Background decision-making involves assessing each cluster's point density, where clusters with higher point density are labeled background (101).

Alternative methodologies uniformly divide 3D space into smaller segments, counting observed points within each segment in an initial set of samples. Segments are classified as background if the frequency of points within that segment exceeds a specified threshold. When segmentation occurs in a spherical coordinate system, the resulting segments are spherical volumes (102-104). Conversely, when segmentation is conducted in the Cartesian coordinate system, the segments are called cubic volumes or voxels (105-107). The Cartesian coordinate system typically produces more segments within the same coverage area. Nevertheless, fewer segments are preferred due to the fixed and limited resolution of LiDAR sensors, and most of the segments are not observed by the LiDAR. The results of background modeling are used to generate a point cloud of foreground points.

The second stage of the conventional LiDAR system entails the application of a density-based clustering algorithm to group foreground points into small point clouds, each representing a potential object (108, 109). A classifier algorithm such as Support Vector Machines (SVM), Neural Networks (NN), or other learning algorithms can be used for object classification, but these supervised learning algorithms need labeled data in the training phase. Unlike image datasets, segmenting and labeling point clouds is still challenging. However, some available datasets provide such information, and the number of datasets is expected to increase. Among all the labeled LiDAR datasets, the KITTI vision benchmark seems to be used as a benchmark. This dataset has been obtained from multiple sensors installed on a vehicle and includes stereo sequences, 3D LiDAR point clouds, GPS/IMU data, and labels of objects (110). The point clouds are recorded using a 64-channel Velodyne LiDAR; therefore, using it for other sensors, especially lower resolution, decreases accuracy and requires extra validation and calibration.

Feature extraction is one of the steps before using classification. Before that, the point clouds are segmented into grids with various sizes where the statistical features of point clouds located inside each occupied cell are extracted. These features include the number of points in each small cell, the average and variance of intensity, the spatial mean and variance of point clouds, and other geometric features (111). In (112), regional shape descriptors are introduced for vehicle detection in point cloud data. The 3D shape context is the first feature that is extracted by computing the statistical characteristics (e.g., histogram) of the point cloud of an object in multiple spherical bins. Then, the harmonic shape contexts are obtained by using these 3D shape contexts (112). The heavy computations of the regional and geometric features may affect the system's real-time performance. Therefore, a histogram-based descriptor on the object and point levels could be used for feature extraction. An object classification method is implemented by applying SVM on histogram features extracted from point clouds (113).

Since background modeling is not feasible in mobile laser scanning applications, sophisticated classification algorithms are used to segment and classify objects in point clouds directly. Several Convolutional Neural Networks (CNN) are trained with KITTI or a similar dataset, and then the trained models are used to classify unseen point clouds. Feature extraction is an essential part of these approaches, where point clouds are segmented into small grids, and a set of features is computed from the points within each grid.

Vote3Deep is a Convolutional Neural Network that uses a voting scheme applied to the statistical features of each grid (114). The VoxelNet architecture uses CNN applied to voxel grids, which has a unique feature extraction method (115). PointFlowNet uses two consecutive frames to extract the features based on the motion and flow in the point clouds and then applies a CNN to its unique feature sets. Therefore, it can detect objects and also estimate their movement in the point clouds (116). The human detection and tracking method implemented in (117) applies a segmentation method to the point clouds and then estimates the motion and speed of each cell by a real-time speed estimation algorithm. Finally, for each segment, it applies an SVM classifier to the extracted feature set.

The other category of methods has fused 3D LiDAR sensors and cameras and has developed an algorithm to handle RGB-D data provided by these two sensors. This fusion will help to improve the results of using each data individually. PointNet applies a CNN to RGB images, matches point clouds to the detected objects, and builds a segmented object point. Finally, it estimates a 3D box encompassing the object in the 3D point clouds and its equivalent in the 2D image (118). PointFusion is another study that applies PointNet to the point cloud data and ResNet to images to extract suitable features. Then, it uses a dense fusion model and a global fusion model to combine PointNet and ResNet outputs. Finally, a neural network is used to estimate the 3D box coordinates (119). A 3D vehicle detection system, implemented in (120), applies CNN to the image to detect cars, then matches the vehicles seen in the image with their corresponding point cloud and computes their geometric features from LiDAR data. Finally, another CNN is applied to 3D car point clouds and estimates the 3D box for each car. MV3D network implemented a distinct architecture for 3D object detection (121). It applied a CNN to image frames, but instead of using 3D points, it applies two different CNN to LiDAR front-view and LiDAR bird-view (from top) images. The bird-view images consist of multiple height, density, and intensity maps extracted from 3D point clouds.

1.6 Organization of the Document

This thesis is structured into six chapters, starting with the introduction. As a manuscript-based thesis, chapters two through five consist of individual articles, with the author of this thesis serving as the primary author for each. These papers are published in journals or presented at conferences and prepared to be submitted to a journal.

Chapter 2 presents the article: A 3D LiDAR-Based Supervised Methodology for Automated Traffic Monitoring and Data Collection at Urban Intersections with High-Mixed Traffic.

Chapter 3 features the article: A 3D-LiDAR-Based Methodology for Surrogate Safety Analysis at Intersections with High Non-Motorized Traffic.

Chapter 4 discusses the article: Development of an Unsupervised 3D LiDAR-Based Methodology for Automated Safety Monitoring of Railway Facilities.

Chapter 5 introduces the journal article: A LiDAR-Based Methodology for Monitoring and Collecting Microscopic Bicycle Flow Parameters on Bicycle Facilities.

Chapter 6 summarizes the achieved objectives, provides concluding remarks, and outlines potential future work.

References

1. Qureshi KN, Abdullah AH. A survey on intelligent transportation systems. Middle-East Journal of Scientific Research. 2013;15(5):629-42.

2. Dimitrakopoulos G, Demestichas P. Intelligent transportation systems. IEEE Vehicular Technology Magazine. 2010;5(1):77-84.

3. Zhang J, Wang F-Y, Wang K, Lin W-H, Xu X, Chen C. Data-driven intelligent transportation systems: A survey. IEEE Transactions on Intelligent Transportation Systems. 2011;12(4):1624-39.

4. Welch TF, Widita A. Big data in public transportation: a review of sources and methods. Transport reviews. 2019;39(6):795-818.

5. Li X, Wu Y-J. Real-time estimation of pedestrian volume at button-activated midblock crosswalks using traffic controller event-based data. Transportation research part C: emerging technologies. 2021;122:102876.

6. Lesani A, Nateghinia E, Miranda-Moreno LF. Development and evaluation of a real-time pedestrian counting system for high-volume conditions based on 2D LiDAR. Transportation research part C: emerging technologies. 2020;114:20-35.

7. Nateghinia E, Beitel D, Lesani A, Miranda-Moreno LF. A LiDAR-based methodology for monitoring and collecting microscopic bicycle flow parameters on bicycle facilities. Transportation. 2022:1-25.

8. Hassanpour A, Bigazzi A. What is on the Bicycle Paths? A Detailed Vehicle Taxonomy with Mode Share Data for Off-Street Paths in Metropolitan Vancouver, Canada. Transportation Research Record. 2023:03611981231165017.

9. Saad WK, Hashim Y, Jabbar WA. Design and implementation of portable smart wireless pedestrian crossing control system. IEEE Access. 2020;8:106109-20.

10. Lu C, Dong J. Estimating freeway travel time and its reliability using radar sensor data. Transportmetrica B: Transport Dynamics. 2018;6(2):97-114.

11. Shirazi MS, Morris BT. Looking at intersections: a survey of intersection monitoring, behavior and safety analysis of recent studies. IEEE Transactions on Intelligent Transportation Systems. 2016;18(1):4-24.

12. Chakraborty P, Adu-Gyamfi YO, Poddar S, Ahsani V, Sharma A, Sarkar S. Traffic congestion detection from camera images using deep convolution neural networks. Transportation Research Record. 2018;2672(45):222-31.

13. Jokela M, Kutila M, Le L, editors. Road condition monitoring system based on a stereo camera. 2009 IEEE 5th International conference on intelligent computer communication and processing; 2009: IEEE.

14. Fu T, Miranda-Moreno L, Saunier N. Pedestrian crosswalk safety at nonsignalized crossings during nighttime: Use of thermal video data and surrogate safety measures. Transportation research record. 2016;2586(1):90-9.

15. Fu T, Stipancic J, Zangenehpour S, Miranda-Moreno L, Saunier N. Automatic traffic data collection under varying lighting and temperature conditions in multimodal environments: Thermal versus visible spectrum video-based systems. Journal of advanced transportation. 2017;2017.

16. Royo S, Ballesta-Garcia M. An overview of lidar imaging systems for autonomous vehicles. Applied sciences. 2019;9(19):4093.

17. Sun P, Sun C, Wang R, Zhao X. Object detection based on roadside LiDAR for cooperative driving automation: a review. Sensors. 2022;22(23):9316.

18. Zhao J, Xu H, Liu H, Wu J, Zheng Y, Wu D. Detection and tracking of pedestrians and vehicles using roadside LiDAR sensors. Transportation research part C: emerging technologies. 2019;100:68-87.

19. Nateghinia E, Miranda-Moreno L. A 3D-LiDAR-based Methodology for Extracting Data for Surrogate Safety Analysis at Intersections. Presented at Transportation Research Board 101st Annual Meeting; 2022.

20. Kamijo S, Matsushita Y, Ikeuchi K, Sakauchi M. Traffic monitoring and accident detection at intersections. IEEE transactions on Intelligent transportation systems. 2000;1(2):108-18.

21. Muffert M, Milbich T, Pfeiffer D, Franke U, editors. May I enter the roundabout? A time-to-contact computation based on stereo-vision. 2012 IEEE Intelligent Vehicles Symposium; 2012: IEEE.

22. Barth A, Pfeiffer D, Franke U, editors. Vehicle tracking at urban intersections using dense stereo. 3rd Workshop on Behaviour Monitoring and Interpretation, BMI; 2009.

23. Tarko AP, Ariyur KB, Romero MA, Bandaru VK, Lizarazo CG. TScan: Stationary LiDAR for Traffic and Safety Studies—Object Detection and Tracking. 2016.

24. Jin S, Qu X, Zhou D, Xu C, Ma D, Wang D. Estimating cycleway capacity and bicycle equivalent unit for electric bicycles. Transportation Research Part A: Policy and Practice. 2015;77:225-48.

25. Xu C, Li Q, Qu Z, Tao P. Modeling of speed distribution for mixed bicycle traffic flow. Advances in mechanical engineering. 2015;7(11):1687814015616918.

26. Jeng SL, Chieng WH, Lu HP. Estimating Speed Using a Side-Looking Single-Radar Vehicle Detector. IEEE Transactions on Intelligent Transportation Systems. 2014;15(2):607-14.

27. Nordback K, Kothuri S, Phillips T, Gorecki C, Figliozzi M. Accuracy of bicycle counting with pneumatic tubes in Oregon. Transportation research record. 2016;2593(1):8-17.

28. Nordback K, Janson BN. Automated Bicycle Counts: Lessons from Boulder, Colorado. Transportation research record. 2010;2190(1):11-8.

29. Strauss J, Miranda-Moreno LF. Speed, travel time and delay for intersections and road segments in the Montreal network using cyclist Smartphone GPS data. Transportation research part D: transport and environment. 2017;57:155-71.

30. Wang Y, Yang X, Liang H, Liu Y. A review of the self-adaptive traffic signal control system based on future traffic environment. Journal of Advanced Transportation. 2018;2018.

31. Zaki MH, Sayed T. Automated cyclist data collection under high density conditions. IET intelligent transport systems. 2016;10(5):361-9.

32. Zaki M, Sayed T, Cheung A. Computer Vision Techniques for the Automated Collection of Cyclist Data. Transportation Research Record: Journal of the Transportation Research Board. 2013;2387(1):10-9.

33. Mohammed H, Bigazzi AY, Sayed T. Characterization of bicycle following and overtaking maneuvers on cycling paths. Transportation research part C: emerging technologies. 2019;98:139-51.

34. Fu T, Miranda-Moreno L, Saunier N. A novel framework to evaluate pedestrian safety at non-signalized locations. Accident Analysis & Prevention. 2018;111:23-33.

35. Fu T, Hu W, Miranda-Moreno L, Saunier N. Investigating secondary pedestrian-vehicle interactions at non-signalized intersections using vision-based trajectory data. Transportation research part C: emerging technologies. 2019;105:222-40.

36. Zangenehpour S, Strauss J, Miranda-Moreno LF, Saunier N. Are signalized intersections with cycle tracks safer? A case–control study based on automated surrogate safety analysis using video data. Accident Analysis & Prevention. 2016;86:161-72.

37. Sayed T, Zaki MH, Autey J. Automated safety diagnosis of vehicle–bicycle interactions using computer vision analysis. Safety Science. 2013;59:163-72.

38. Wang C, Xie Y, Huang H, Liu P. A review of surrogate safety measures and their applications in connected and automated vehicles safety modeling. Accident Analysis & Prevention. 2021;157:106157.

39. Arun A, Haque MM, Bhaskar A, Washington S, Sayed T. A systematic mapping review of surrogate safety assessment using traffic conflict techniques. Accident Analysis & Prevention. 2021;153:106016.

40. Arun A, Haque MM, Washington S, Sayed T, Mannering F. A systematic review of traffic conflict-based safety measures with a focus on application context. Analytic methods in accident research. 2021;32:100185.

41. Wu J, Xu H, Zheng Y, Tian Z. A novel method of vehicle-pedestrian near-crash identification with roadside LiDAR data. Accident Analysis & Prevention. 2018;121:238-49.

42. Lv B, Sun R, Zhang H, Xu H, Yue R. Automatic vehicle-pedestrian conflict identification with trajectories of road users extracted from roadside LiDAR sensors using a rule-based method. Ieee Access. 2019;7:161594-606.

43. Bhattarai N, Zhang Y, Liu H, Pakzad Y, Xu H. Proactive Safety Analysis Using Roadside LiDAR Based Vehicle Trajectory Data: A Study of Rear-End Crashes. Transportation Research Record. 2023:03611981231182704.

44. Bhattarai N, Zhang Y, Liu H, Xu H. Crash frequency prediction based on extreme value theory using roadside lidar-based vehicle trajectory data. Accident Analysis & Prevention. 2023;193:107306.

45. Beck J, Arvin R, Lee S, Khattak A, Chakraborty S. Automated vehicle data pipeline for accident reconstruction: new insights from LiDAR, camera, and radar data. Accident Analysis & Prevention. 2023;180:106923.

46. Transportation Safety Board of Canada. Railway Occurrence Data (2013-2022) Canada, 2022 [Available from: http://www.bst-tsb.gc.ca/eng/stats/rail/data-1.html.

47. Fakhfakh N, Khoudour L, Bruyelle J-L, El-Koursi E-M. Intelligent Surveillance System Based on Stereo Vision for Level Crossings Safety Applications. Recent Developments in Video Surveillance: InTech; 2012.

48. Salmane H, Khoudour L, Ruichek Y. A video-analysis-based railway–road safety system for detecting hazard situations at level crossings. IEEE Transactions on Intelligent Transportation Systems. 2015;16(2):596-609.

49. Amaral V, Marques F, Lourenço A, Barata J, Santana P. Laser-based obstacle detection at railway level crossings. Journal of Sensors. 2016;2016.

50. Horne D, Findley DJ, Coble DG, Rickabaugh TJ, Martin JB. Evaluation of radar vehicle detection at four quadrant gate rail crossings. Journal of Rail Transport Planning & Management. 2016;6(2):149-62.

51. daSilva MP, Baron W, Carroll AA. Railroad infrastructure trespassing detection systems research in Pittsford, New York. United States. Federal Railroad Administration; 2012.

52. Hsieh H-H, Hsu C-Y, Ke P-Y, Liu G-S, Lin C-P, editors. Appling Lidar-based obstacle detection and wireless image transmission system for improving safety at level crossings. Security Technology (ICCST), 2015 International Carnahan Conference on; 2015: IEEE.

53. Thornton DA, Redmill K, Coifman B. Automated parking surveys from a LIDAR equipped vehicle. Transportation research part C: emerging technologies. 2014;39:23-35.

54. Canada S. Travel Survey of Residents of Canada (TSRC). 2017.

55. Abdulhafedh A. Road traffic crash data: an overview on sources, problems, and collection methods. Journal of transportation technologies. 2017;7(2):206-19.

56. Datondji SRE, Dupuis Y, Subirats P, Vasseur P. A survey of vision-based traffic monitoring of road intersections. IEEE transactions on intelligent transportation systems. 2016;17(10):2681-98.

57. Won M. Intelligent traffic monitoring systems for vehicle classification: A survey. IEEE Access. 2020;8:73340-58.

58. D'Alconzo A, Drago I, Morichetta A, Mellia M, Casas P. A survey on big data for network traffic monitoring and analysis. IEEE Transactions on Network and Service Management. 2019;16(3):800-13.

59. Shoup G, Remias SM, Hainen AM, Grimmer G, Davis AD. Characterizing Reliability of Manual Intersection Turning Movement Counts Using Modern Data Collection Technology. 2013.

60. Kocamaz MK, Gong J, Pires BR, editors. Vision-based counting of pedestrians and cyclists. 2016 IEEE Winter Conference on Applications of Computer Vision (WACV); 2016: IEEE.

61. Lesani A, Miranda-Moreno L. Development and testing of a real-time WiFi-bluetooth system for pedestrian network monitoring, classification, and data extrapolation. IEEE Transactions on Intelligent Transportation Systems. 2018;20(4):1484-96.

62. Jackson S, Lesani A, Miranda-Moreno L. Towards a WIFI system for traffic monitoring in different transportation facilities. 2014.

63. Hadavi S, Rai HB, Verlinde S, Huang H, Macharis C, Guns T. Analyzing passenger and freight vehicle movements from automatic-Number plate recognition camera data. European Transport Research Review. 2020;12(1):1-17.

64. Stipancic J, Miranda-Moreno L, Saunier N. Vehicle manoeuvers as surrogate safety measures: Extracting data from the gps-enabled smartphones of regular drivers. Accident Analysis & Prevention. 2018;115:160-9.

65. Bottero M, Dalla Chiara B, Deflorio FP. Wireless sensor networks for traffic monitoring in a logistic centre. Transportation Research Part C: Emerging Technologies. 2013;26:99-124.

66. Bifulco GN, Galante F, Pariota L, Spena MR, Del Gais P. Data collection for traffic and drivers' behaviour studies: a large-scale survey. Procedia-social and behavioral sciences. 2014;111:721-30.

67. Messelodi S, Modena CM, Zanin M. A computer vision system for the detection and classification of vehicles at urban road intersections. Pattern analysis and applications. 2005;8:17-31.

68. Sivaraman S, Trivedi MM. Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis. IEEE transactions on intelligent transportation systems. 2013;14(4):1773-95.

69. Pourhomayoun M. Automatic traffic monitoring and management for pedestrian and cyclist safety using deep learning and artificial intelligence. 2020.

70. Fang X, He M, Huang D, Zhang Z, Ge L, Xiao G. JTEA: A Joint Trajectory Tracking and Estimation Approach for Low-Observable Micro-UAV Monitoring with 4D Radar. IEEE Transactions on Instrumentation and Measurement. 2023.

71. Zhang Z, Lai H, Huang D, Fang X, Zhou M, Zhang Y. RETA: 4D Radar-Based End-to-End Joint Tracking and Activity Estimation for Low-Observable Pedestrian Safety in Cluttered Traffic Scenarios. IEEE Transactions on Intelligent Transportation Systems. 2023.

72. Lan X, Wang C, Lv B, Li J, Zhang M, Zhang Z. 3D Point Cloud Stitching for Object Detection with Wide FoV Using Roadside LiDAR. Electronics. 2023;12(3):703.

73. Gao Y, Zhou C, Rong J, Wang Y. High-accurate vehicle trajectory extraction and denoising from roadside LIDAR sensors. Infrared Physics & Technology. 2023;134:104896.

74. Parker Jr M, Zegeer CV. Traffic conflict techniques for safety and operations: Observers manual. United States. Federal Highway Administration; 1989.

75. Almqvist S, Ekman L. The Swedish Traffic Conflict Technique Observers Manual. Department of Technology and Society Traffic Engineering, Lund University. 2001.

76. Zheng L, Ismail K, Meng X. Traffic conflict techniques for road safety analysis: open questions and some insights. Canadian journal of civil engineering. 2014;41(7):633-41.

77. Mahmud SS, Ferreira L, Hoque MS, Tavassoli A. Application of proximal surrogate indicators for safety evaluation: A review of recent developments and research needs. IATSS research. 2017;41(4):153-63.

78. Johnsson C, Laureshyn A, De Ceunynck T. In search of surrogate safety indicators for vulnerable road users: a review of surrogate safety indicators. Transport Reviews. 2018;38(6):765-85.

79. Laureshyn A, Svensson Å, Hydén C. Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation. Accident Analysis & Prevention. 2010;42(6):1637-46.

80. St-Aubin P, Saunier N, Miranda-Moreno L. Large-scale automated proactive road safety analysis using video data. Transportation Research Part C: Emerging Technologies. 2015;58:363-79.

81. Saunier N, Sayed T, Ismail K. Large-scale automated analysis of vehicle interactions and collisions. Transportation Research Record. 2010;2147(1):42-50.

82. Laureshyn A. Application of automated video analysis to road user behaviour: Lund University; 2010.

83. Ismail K, Sayed T, Saunier N, Lim C. Automated analysis of pedestrian–vehicle conflicts using video data. Transportation research record. 2009;2140(1):44-54.

84. Niaki MSN, Saunier N, Miranda-Moreno LF. Is that move safe? Case study of cyclist movements at intersections with cycling discontinuities. Accident Analysis & Prevention. 2019;131:239-47.

85. Wehr A, Lohr U. Airborne laser scanning—an introduction and overview. ISPRS Journal of photogrammetry and remote sensing. 1999;54(2-3):68-82.

86. Yan WY, Shaker A, El-Ashmawy N. Urban land cover classification using airborne LiDAR data: A review. Remote Sensing of Environment. 2015;158:295-310.

87. Yao W, Hinz S, Stilla U, editors. Traffic monitoring from airborne LIDAR–Feasibility, simulation and analysis. XXI Congress, proceedings International archives of photogrammetry, remote sensing and spatial geoinformation sciences, Beijing, China; 2008: Citeseer.

88. Borcs A, Benedek C. Extraction of Vehicle Groups in Airborne Lidar Point Clouds With Two-Level Point Processes. IEEE Transactions on Geoscience and Remote Sensing. 2015;53(3):1475-89.

89. Lemmens M. Terrestrial laser scanning. Geo-information: Springer; 2011. p. 101-21.

90. Pu S, Vosselman G. Knowledge based reconstruction of building models from terrestrial laser scanning data. ISPRS Journal of Photogrammetry and Remote Sensing. 2009;64(6):575-84.

91. Pagounis V, Tsakiri M, Palaskas S, Biza B, Zaloumi E, editors. 3D laser scanning for road safety and accident reconstruction. Proceedings of the XXIIIth international FIG congress; 2006.

92. Williams K, Olsen M, Roe G, Glennie C. Synthesis of transportation applications of mobile LiDAR. Remote Sensing. 2013;5(9):4652-92.

93. Paparoditis N, Papelard J-P, Cannelle B, Devaux A, Soheilian B, David N, et al. Stereopolis II: A multi-purpose and multi-sensor 3D mobile mapping system for street visualisation and 3D metrology. Revue française de photogrammétrie et de télédétection. 2012;200(1):69-79.

94. Haiyan G, Li J, Yongtao Y, Chapman M, Cheng W. Automated Road Information Extraction From Mobile Laser Scanning Data. IEEE Transactions on Intelligent Transportation Systems. 2015;16(1):194-205.

95. Tan M, Wang B, Wu Z, Wang J, Pan G. Weakly supervised metric learning for traffic sign recognition in a LIDAR-equipped vehicle. IEEE Transactions on Intelligent Transportation Systems. 2016;17(5):1415-27.

96. Yu Y, Li J, Guan H, Wang C, Yu J. Semiautomated Extraction of Street Light Poles From Mobile LiDAR Point-Clouds. IEEE Transactions on Geoscience and Remote Sensing. 2015;53(3):1374-86.

97. Yu Y, Li J, Guan H, Jia F, Wang C. Learning Hierarchical Features for Automated Extraction of Road Markings From 3-D Mobile LiDAR Point Clouds. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2015;8(2):709-26.

98. Tsai Y-C, Chatterjee A. Pothole Detection and Classification Using 3D Technology and Watershed Method. Journal of Computing in Civil Engineering. 2018;32(2).

99. Xiao W, Vallet B, Schindler K, Paparoditis N. Street-side vehicle detection, classification and change detection using mobile laser scanning data. ISPRS Journal of Photogrammetry and Remote Sensing. 2016;114:166-78.

100. Zhang T, Jin PJ, Ge Y. Weighted bayesian gaussian mixture model for roadside lidar object detection. Ithaca: Cornell University Library, arXiv org. 2022.

101. Douillard B, Underwood J, Kuntz N, Vlaskine V, Quadros A, Morton P, et al., editors. On the segmentation of 3D LIDAR point clouds. 2011 IEEE International Conference on Robotics and Automation; 2011: IEEE.

102. Vranic DV, Saupe D, editors. Description of 3D-shape using a complex function on the sphere. Proceedings IEEE International Conference on Multimedia and Expo; 2002: IEEE.

103. Shin M, Seo Y, Bae J, Choi YS, Kim H, Byun H, et al. BallGAN: 3D-aware Image Synthesis with a Spherical Background. arXiv preprint arXiv:230109091. 2023.

104. Saupe D, Vranić DV, editors. 3D model retrieval with spherical harmonics and moments. Joint Pattern Recognition Symposium; 2001: Springer.

105. Mahmoud A, Hu JS, Waslander SL, editors. Dense voxel fusion for 3D object detection. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision; 2023.

106. Schwarz K, Sauer A, Niemeyer M, Liao Y, Geiger A. Voxgraf: Fast 3d-aware image synthesis with sparse voxel grids. Advances in Neural Information Processing Systems. 2022;35:33999-4011.

107. Xiang Y, Choi W, Lin Y, Savarese S, editors. Data-driven 3d voxel patterns for object category recognition. Proceedings of the IEEE conference on computer vision and pattern recognition; 2015.

108. Ahmed SM, Chew CM, editors. Density-based clustering for 3d object detection in point clouds. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2020.

109. Xiong S, Li B, Zhu S. DCGNN: A single-stage 3D object detection network based on density clustering and graph neural network. Complex & Intelligent Systems. 2023;9(3):3399-408.

110. Geiger A, Lenz P, Stiller C, Urtasun R. Vision meets robotics: The KITTI dataset. The International Journal of Robotics Research. 2013;32(11):1231-7.

111. Wang DZ, Posner I, editors. Voting for Voting in Online Point Cloud Object Detection. Robotics: Science and Systems; 2015.

112. Frome A, Huber D, Kolluri R, Bülow T, Malik J, editors. Recognizing objects in range data using regional point descriptors. European conference on computer vision; 2004: Springer.

113. Himmelsbach M, Luettel T, Wuensche H-J, editors. Real-time object classification in 3D point clouds using point feature histograms. 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems; 2009: IEEE.

114. Engelcke M, Rao D, Wang DZ, Tong CH, Posner I, editors. Vote3deep: Fast object detection in 3d point clouds using efficient convolutional neural networks. 2017 IEEE International Conference on Robotics and Automation (ICRA); 2017: IEEE.

115. Zhou Y, Tuzel O, editors. Voxelnet: End-to-end learning for point cloud based 3d object detection. Proceedings of the IEEE conference on computer vision and pattern recognition; 2018.

116. Behl A, Paschalidou D, Donné S, Geiger A. PointFlowNet: Learning Representations for 3D Scene Flow Estimation from Point Clouds. arXiv preprint arXiv:180602170. 2018.

117. Yan Z, Duckett T, Bellotto N, editors. Online learning for human classification in 3d lidar-based tracking. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2017: IEEE.

118. Qi CR, Liu W, Wu C, Su H, Guibas LJ, editors. Frustum pointnets for 3d object detection from rgb-d data. Proceedings of the IEEE conference on computer vision and pattern recognition; 2018.

119. Xu D, Anguelov D, Jain A, editors. Pointfusion: Deep sensor fusion for 3d bounding box estimation. Proceedings of the IEEE conference on computer vision and pattern recognition; 2018.

120. Du X, Ang MH, Karaman S, Rus D, editors. A general pipeline for 3d detection of vehicles. 2018 IEEE International Conference on Robotics and Automation (ICRA); 2018: IEEE.

121. Chen X, Ma H, Wan J, Li B, Xia T, editors. Multi-view 3d object detection network for autonomous driving. Proceedings of the IEEE conference on Computer Vision and Pattern Recognition; 2017.

CHAPTER 2.

A 3D LIDAR-BASED SUPERVISED METHODOLOGY FOR AUTOMATED TRAFFIC MONITORING AND DATA COLLECTION AT URBAN INTERSECTIONS WITH HIGH-MIXED TRAFFIC

CHAPTER 2: A 3D LIDAR-BASED SUPERVISED METHODOLOGY FOR AUTOMATED TRAFFIC MONITORING AND DATA COLLECTION AT URBAN INTERSECTIONS WITH HIGH-MIXED TRAFFIC

2.1 Abstract

This study aims to develop and evaluate a novel 3D LiDAR-based supervised learning methodology for automated traffic monitoring and data collection at urban intersections with high mixed traffic conditions characterized by high volumes of pedestrians and cyclists. Two alternative sensor resolutions (16 and 32 channels) are evaluated, and LiDAR data are utilized for training and evaluation. The critical components of the proposed methodology include LiDAR data processing and background modeling based on the Gaussian mixture model, road user detection and clustering, feature extraction and classification, and tracking. These processes are primarily automated, leveraging point cloud data processing and machine learning algorithms for advanced data analysis and interpretation. The methodology is designed to accommodate high- and low-resolution LiDAR sensors, ensuring suitability for different LiDAR systems.

An extensive performance evaluation procedure was conducted to assess the various components of the proposed methodology and key sensor parameters. The background modeling framework demonstrates a 94.2% accuracy when applied to data from high-resolution LiDAR and 89.8% for lower-resolution systems. The XGBoost classifier achieves the best performance of the alternative classifiers, with an average classification rate of 0.95 for high-resolution LiDAR systems and 0.91 for lower-resolution systems. The Average Displacement Error (ADE) for road user tracking with the LiDAR system is 0.37 meters for high-resolution systems and 0.40 meters for low-resolution systems. The weighted average absolute percentage difference of road user counts is 6% and 13% for high and lower resolution, respectively.

Overall, the results demonstrate the accuracy of our methodology and the impact of sensor resolution. The performance of the sensors varies across different intersection settings; notably, in intersections with an extended area, the performance is influenced by the sensors' limited vertical field of view.

***Keywords**:* Urban Intersection Monitoring, 3D LiDAR Monitoring, Supervised LiDAR Method, Alternative Technologies

2.2 Introduction

Urban intersections represent a microcosm of urban mobility challenges and opportunities. The efficiency of the transportation network is tested at urban intersections where different modes of transportation utilize the infrastructure for making daily trips. An urban intersection is of utmost importance from a transportation safety perspective, where many interactions between road users, especially between motorized and vulnerable road users, happen daily. Thus, there is a growing need for advanced and efficient traffic and safety monitoring systems at urban intersections.

The current approach to traffic monitoring at urban intersections combines traditional methods with new technologies. The choice of methodology depends on the data collection program's requirements, overall cost, duration, and available technologies. Traditional approaches to traffic monitoring, such as inductive loops, are still widely used. Although primarily designed for monitoring traffic on highways or road sections, correctly implementing inductive loops can provide valuable insights into traffic conditions at a given intersection. Inductive loops collect data on various aspects, including the count of vehicles entering or exiting, speed, link travel time, and vehicle classification based on length (1). They also gather information on vehicle queues, which is essential for traffic control operations and signal timing (2).

Radar-based technologies effectively monitor the speeds of vehicles approaching an intersection across various lanes. Understanding the speed profile of vehicular traffic on connected links is crucial for traffic signal control and safety analysis. The accuracy of radar in collecting speed data, and consequently acceleration and deceleration rates, is pivotal for analyzing driving behavior at stop sign-controlled intersections (3). One of the critical advantages of radar is its robust performance in adverse weather conditions. Additionally, modern Radar systems integrated with machine learning algorithms can classify different types of vehicles (4).

Loop detectors and radars are primarily utilized for monitoring traffic on street sections rather than for in-depth analysis at intersections. They fall short in tracking the movement of vehicles through intersections, counting turning movements, or observing interactions between road users and the infrastructure. Crucially, these technologies do not encompass active transportation monitoring, a vital component in understanding urban traffic patterns. Consequently, there has been a move towards introducing or integrating alternative technologies to capture more sophisticated traffic data to address these limitations.

Bluetooth and Wi-Fi technologies are designed to overcome some of the issues previously mentioned. A Bluetooth system continuously captures the MAC address of any device with an active Bluetooth module and traces these addresses using an extensive network of Bluetooth readers for traffic monitoring (5). Placing Bluetooth sensors on each approach at urban intersections, especially over longer distances, helps understand road users' origin-destination patterns (6). Bluetooth systems can collect cyclist and pedestrian data at urban transportation facilities (7). However, there are limitations, such as the inability to observe detailed intersection activities, road user interactions, and precise speed measurements. The effectiveness of these systems also varies based on the penetration rates of Bluetooth-enabled devices in different areas. Consequently, the data from these systems must be carefully balanced and typically require scaling up with reference counts collected using loop detectors.

Camera-based systems, widely implemented for traffic monitoring at urban intersections, address some shortcomings of the technologies mentioned earlier. These systems and their installations can be customized for specific traffic data collection tasks. Camera-based systems are adept at detecting, classifying, and tracking road users (8). An essential application of these systems is monitoring intersections to extract turning movement counts and origin-destination traffic volumes (9, 10). With advancements in the resolution of visual-spectrum cameras, these systems are now more suitable for monitoring pedestrian and cyclist traffic (11). For traffic safety applications, a camera-based system can be installed to actively monitor for incident detection and management (12) and for data collection to conduct comparative analytics of road user trajectories for studying their interactions and potential conflicts (13).

While camera-based systems have proven valuable in many traffic studies and applications, they have several key limitations. First, the performance of a camera-based system is downgraded by environmental changes, such as lighting conditions and weather. Additionally, the field of view and range of coverage for a system relying on a single camera is limited. The system's performance is closely linked to the quality of the video sequences it captures. Higher-quality videos often lead to increased costs due to the need for higher-resolution cameras and greater processing power.

The camera-based system collects data in a 2D pixel/image coordinate system, where units are in pixels and images are captured from the camera's perspective. Consequently, for traffic monitoring applications that require accurate distance measurements, image data must be converted into real-world coordinates using geometric calibration. This process involves mapping the 2D image coordinates to real-world 3D coordinates, taking into account the camera's position, angle, and lens properties (14). This is a complex operation that would require knowledge of specific camera optic systems. Any change in the system field of view, angles, or sudden vibration would reset the geometry calibration and require another manual calibration.

As an alternative to video cameras, LiDAR-based methods present a potential solution to several limitations associated with camera-based systems. LiDAR sensors, depending on their resolution and range, can offer more precise data in the x-y-z coordinate system. Each LiDAR scan generates a point cloud of the scanned area, encompassing specific attributes. From this point cloud, the individual point clouds representing each road user can be extracted, allowing for the further extraction of additional physical characteristics such as length, width, and area.

A LiDAR system is not affected by low light conditions, such as at night, making it suitable for long-term monitoring. While adverse weather conditions impact LiDAR, its level of susceptibility is generally lower compared to camera-based systems (15).

The LiDAR sensor's ability to provide precise distance data eliminates the need for manual calibration. Furthermore, the distance-based nature of LiDAR's measurement excludes the geometric computation step required for cameras, making its data distortion-free. Despite recent interest in LiDAR, documented research studies in this field are scarce. The limited research on LiDAR applications can be attributed to the high cost of LiDAR sensors compared to cameras, resulting in a lack of labeled data. Additionally, producing labeled data in a point cloud setting is much more challenging than labeling objects in images.

LiDAR is considered an emerging technology in transportation, with applications in both traffic monitoring and Autonomous Vehicles. Currently, only a handful of studies have utilized LiDAR as a fixed-position monitoring technology for traffic monitoring and data collection (16, 17).

This research proposes a supervised methodology for 3D LiDAR-based traffic monitoring at intersections with mixed traffic. The methodology includes the development of point cloud data processing and machine learning algorithms for object detection, clustering of road users, classification, and tracking.

The structure of the paper is as follows: initially, a literature review tailored to traffic monitoring at intersections, focusing on camera-based systems and exploring LiDAR-based systems as an alternative technology, is presented. Next, the system's methodology, including data collection, is discussed in detail. Finally, the performance of the proposed methods for LiDAR data processing is evaluated.

2.3 Literature Review

Monitoring traffic flow at intersections is crucial for analyzing network flow at microscopic levels. While traditional counting systems, such as inductive loops, excel on highways and streets, they prove inadequate for intersection monitoring as they cannot capture the direction of movements. Video-based systems have been developed to detect and track road users, especially vehicles at intersections (18). The camera monitoring system proposed by Kamijo et al. utilized traditional image processing techniques and achieved a success rate of 93%–96% in vehicle detection and tracking. However, the study only focused on vehicle detection and tracking. The system is also evaluated for detecting bumping accidents in a few cases (18).

The real-time vision-based system developed by Messelodi et al. utilized a monocular camera (19). Moving objects are detected with segmentation and motion analysis. Detection of moving objects involves segmentation and motion analysis. A two-stage classification method, initially based on volume and subsequently on feature matching, determines the type of road users. It is noteworthy that pedestrians are not treated as a single class. The classification rates of bicycle, motorcycle, car, van, and bus are 72.5%, 89.6%, 95.9%, 58.7%, and 100% respectively (19).

A stereo-camera system was proposed by Muffert et al. to identify dangerous conflicts as vehicles enter the roundabout (20). Objects are detected using the disparity map technique and are clustered into road users, specifically vehicles, using the DBSCAN algorithm. The time-to-contact indicator for multiple scenarios is then computed from the extracted trajectories. However, the number of validated samples and the system's accuracy in detection and tracking have not been discussed (20). The use of stereo vision mounted on a car for detecting and tracking other vehicles at urban intersections is discussed in (21). The system underwent testing as a vehicle equipped with a stereo camera passed several intersections, and the results of tracking other cars were visually assessed.

Wang et al. proposed a vision-based system that utilizes the fisheye view of the camera to monitor and measure traffic volume at an intersection (22). The system achieves real-time vehicle detection, tracking, and counting through feature point tracking. The system was deployed at three intersections, showing variations in tracking accuracy and counting across different sites. The counting performance indicates a promising accuracy of 98%. However, the system has not been tested to estimate origin-destination or extract of safety indicator studies (22).

The development of the origin-destination trip table estimation using a video-based monitoring system is discussed in (23). The system leverages the particle filter to obtain trajectories and counts the turning movements in two intersections. Nevertheless, the system lacked calibration for pedestrian detection and tracking.

LiDAR, initially utilized for remote sensing applications, is an emerging technology showing promising performance in areas such as topographic mapping, forestry (24), agriculture (25), and 3D scanning of buildings for urban planning (26). With advancements in autonomous vehicles (AVs), LiDAR sensors have become a crucial component of their vision systems. In these systems, LiDAR contributes to tasks like environmental mapping of real-world objects, such as traffic signs (27) and obstacle detection (28), while also complementing other parts of the system, such as cameras (29).

Using LiDAR-based systems for traffic monitoring and data collection is relatively new and is seen as an emerging technology. This suggests that the body of research focusing on implementing and evaluating LiDAR-based systems for these applications is likely to expand rapidly. However, few key studies have utilized LiDAR implementation for traffic monitoring.

Tarko et al. developed a LiDAR-based system for intersection monitoring, employing Velodyne's high-resolution LiDAR with 64 laser channels (30). The system is calibrated with an unsupervised algorithm, where road users are obtained through background subtraction and classified based on their speed and dimensions. The classification stage considers the part of the intersection where the road user has been detected, such as the sidewalk, crosswalk, or street. The system, installed at three intersections, identified one conflict with a time-to-collision of less than 1.5 s and falsely detected two non-conflicts while accurately counting and classifying 98% of 504 road users (30).

Zhao et al. developed a 16-laser channels LiDAR-based traffic monitoring system (31). The system utilizes a methodology involving background filtering, clustering, and classification to detect and track pedestrians and vehicles at intersections. The classification is performed using a neural network with three features: the total number of points, the distance of the user to the LiDAR sensor, and the direction. The reported accuracy of classification is approximately 93%. However, other road users have not been discussed (31).

Despite the existing research, several gaps exist in implementing and developing LiDAR systems for ITS applications. This research's proposed 3D LiDAR-based methodology offers advantages over those implemented in (30, 31). The first study (30) developed an unsupervised method, and the second study (31) introduced a semi-supervised method applicable to two broad classes: pedestrians and vehicles.

This research presents an efficient semi-automated labeling approach to addressing the challenges of developing a labeled dataset for two LiDAR sensors with 16 and 32 laser channels, respectively. It also involves calibrating automated methods for detecting, tracking, and classifying all road users, including pedestrians, cyclists, cars, and trucks. Secondly, compared to (30), using lower-resolution LiDAR sensors reduces the system's cost, rendering it economically feasible for practical applications. Lastly, the system is tested and evaluated in various traffic conditions.

2.4 3D Rotational LiDAR Sensor and Definition

This section introduces LiDAR-based technology and its associated definitions, focusing on the sensor data characteristics vital for our proposed traffic data collection and monitoring methodology at intersections with mixed traffic types. The core activities involve detecting, classifying, and tracking various road users, including passenger cars, trucks, cyclists, and pedestrians. These primary objectives are essential for accurate and comprehensive traffic analysis in such environments.

The proposed methodology is built for rotational LiDAR systems with *N* number of channels. In this research, the proposed LiDAR system utilizes two rotational LiDAR sensors produced by Velodyne Lidar: VLP-16 and VLP-32c LiDARs (32, 33). The VLP-16 and the VLP-32c sensors have 16 and 32 laser channels ($n_{channel} = 16\ or\ 32$), respectively. The LiDAR channels are vertically separated and rotate with an adjustable speed (ω) ranging from 5 to 20 rotations per second. Both sensors have a 360° horizontal field-of-view (α_{FOW}). The vertical field of views (γ_{FOW}) of VLP-16 and VLP-32c are 30° and 40°, respectively. The rotational LiDAR sensor scans the surroundings and generates a comprehensive three-dimensional (3D) point cloud representing the observed environment.

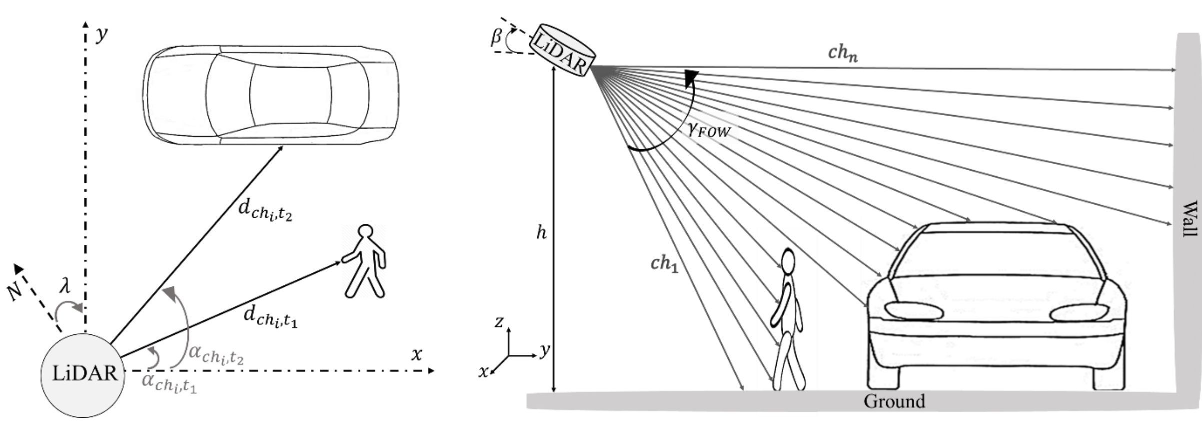
The LiDAR sensor output includes pairs of distance-reflection measurements. These measurements are associated with three distinct indices: the channel ID (ch_i) which spans from 1 to 32 (or 16 for VLP-16), the azimuth value (α_{ch_i,t_j}) ranging from 0° to 360° and the timestamp (t_j). The sensors measure distance (d_{ch_i,t_j}) up to 200m (or 100m for VLP-16) with a precision within ±3cm.

The VLP-16 has a vertical angular resolution of 2°. The angular resolution of VLP-32c is not evenly distributed across its vertical channels. Among the 31 vertical angular gaps between LiDAR channels in VLP-32c, there are 17 gaps of 0.33°, four gaps of less than 1°, four gaps of less than 2°, and six gaps of more than 2°. The horizontal resolution of both sensors ranges from 0.1° to 0.4° depending on the rotational speed. Each laser channel in both LiDAR sensors records 18,000 distance points per second. When the rotational speed is adjusted at 10Hz, these 18,000 distance points are evenly distributed across ten rotations, yielding 1,800 distance points measured per rotation and laser channel. As a result, the horizontal resolution is fixed at $360°/1800 = 0.2°$, and the LiDAR sensors observe $n_{channels}$×1800 points per rotation.

Table 2-1 summarizes all the characteristics mentioned above for both sensors, the notations of the variables introduced in this manuscript, and the distinction between the two LiDAR sensors. Notably, VLP-32c presents two distinctive characteristics: the non-uniform vertical angular resolution (δ_γ) and the non-uniform horizontal angular offset (σ_α), both varying from one channel to another.

Table 2-1 Key parameters of the two LiDAR systems

Parameter	Notation	VLP-16 Velodyne	VLP-32c Velodyne
Number of channels	$n_{channels}$	16	32
Distance measured by $\boldsymbol{i^{th}}$ channels at time $\boldsymbol{t_j}$	d_{ch_i,t_j}	$up\ to\ 100m$	$up\ to\ 200m$
Distance resolution	δ_d	$3cm$	$3cm$
Horizontal field-of-view	α_{FOW}	$360°$	$360°$
Vertical field-of-view	γ_{FOW}	$30°$	$40°$
Azimuth (horizontal) of the $\boldsymbol{i^{th}}$ channels at time $\boldsymbol{t_j}$	α_{ch_i,t_j}	$[0° - 360°]$	$[0° - 360°]$
Vertical angle of he $\boldsymbol{i^{th}}$ channels	γ_{ch_i}	$[-15°, +15°]$	$[-25°, +15°]$
Sampling rate (number of points per second)	SR	288,000	576,000
Vertical angular resolution	δ_γ	$2°$	Variable
Horizontal angular offset	σ_α	$0°$	Variable
Horizontal angular resolution	δ_α	$0.1°$– $0.4°$ (Adjusted as $0.2°$)	
Rotational speed	ω	$5 - 20Hz$ (Adjusted as $10Hz$)	
Frame rate	FR	$5 - 20fps$ (Adjusted as $10fps$)	
Height of the LiDAR's installation	h	depends on the intersection	
Title angle of the LiDAR's installation	β	depends on the intersection	
Rotation angle of the LiDAR's installation relative to the North	λ	depends on the intersection	



a- Top-view of the system

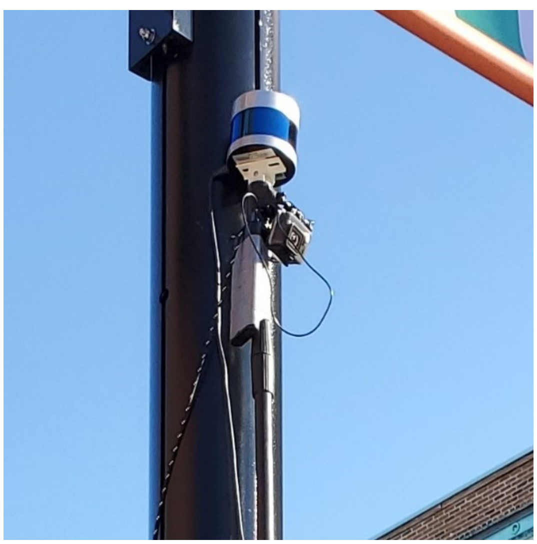
b- Side-view of the system

Figure 2-1 Setup and operation of the 3D LiDAR system

Figure 2-1 (a) illustrates the top view of a LiDAR in a hypothetical scenario emphasizing the operation of channel i at two different horizontal angles, corresponding to two different timestamps (t_1 and t_2). In this example, d_{ch_i,t_j} and α_{ch_i,t_j} are the distance and azimuth of channel i at time t_j. **Figure 2-1 (b)** illustrates the side-view of the VLP-16 sensor in the same hypothetical scenario where 16 channels operate simultaneously. To maximize the coverage area, the sensor is tilted downward along the x-axis by an angle of β and rotated around the z-axis by an angle of λ.

2.5 Data Collections

This section details the data collection process for developing and testing the proposed methodology. This research utilizes an integrated system comprising a 3D-LiDAR sensor (16 and 32 channels), a camera for ground truth data collection, a Raspberry Pi, a memory card, and batteries. These components were assembled and attached to a telescopic mast for easy and secure installation on existing structures like lamp posts. The integrated LiDAR systems, illustrated in **Figure 2-2,** are installed at thirteen urban intersections in Montreal, Canada.



a-LiDAR and camera installation

b- Processing unit and battery box installation

Figure 2-2 Integrated LiDAR and hardware components for data collection

The selection of the intersections is subject to a few factors, including the level of non-motorized and motorized traffic volumes, diverse intersection sizes, intersection traffic control type (stop sign vs. traffic light), and availability of a pole for attaching and installing the LiDAR system.

Table 2-2 provides the key characteristics of the selected intersections and the LiDAR installation configurations at each of these intersections. The data collection spanned from 2018 to 2021. At each intersection, data were collected for four hours, with LiDAR's frame rate consistently set at ten frames per second.

Table 2-2 Key Characteristics of LiDAR system installation at various intersections

Site ID	Intersection Name	LiDAR' s Channel	Start Time of Data Collection Period	LiDAR Height (m)	Tilt Angle	Rotation Angle Toward North	Intersection Area (m2)	LiDAR Distance to Intersection' s Boundary (m)	LiDAR Distance to Intersection' s Center (m)
101	Sainte Famille - Milton	16	09:20	4.2	-17.0º	-9º	123	7.0	13.1
102	Papineau – Sherbrooke E	16	08:40	5.2	-20.9º	71º	353	6.9	17.1
103	Atwater – Sherbrooke W	16	09:35	4.7	-14.6º	97º	444	6.1	21.5
104	De La Roche – Marie Anne E	16	14:50	4.6	-11.4º	-3º	124	4.2	11.5
105	Coloniale – Rachel E	16	09:30	5.1	-17.5º	117º	102	5.2	12.3
106	Girouard - Monkland	16	09:30	4.5	-21.3º	-117º	177	6.2	15.6
107	University - Milton	16	11:20	4.5	-19.9º	-164º	79	4.2	10.7
108	Hutchison - LaurierE	32	10:54	4.5	-7.1º	163º	302	6.3	19.2
109	Sainte Famille – Prince Arthur W	32	09:12	4.0	-9.9º	69º	134	7.2	14.0
110	Parc - PineW	32	12:00	4.6	-3.3º	-99º	684	27.1	40.1
111	Saint Denis – Saint Joseph E	32	10:35	4.2	-12.9º	-17º	409	5.4	18.5
112	Parthenais – Rachel E	32	09:15	3.9	-12.0º	40º	117	5.9	13.4
113	University - Milton	32	15:25	4.2	-19.6º	-100º	74	4.0	8.7

The sensor setup parameters, including height (h), tilt angle (β), and rotation angle around the z-axis (λ), are selected to maximize coverage area and are determined by the intersection size and LiDAR's distance to the intersection. In this study, the height of the LiDAR installation ranges from 3.9m to 5.2m. The tilt angle of the 16-channel LiDAR installation spans from -11.4° to -21.3°, whereas the tilt angle of the 32-channel LiDAR installation ranges from -3.3° to -19.6°. The 32-channel LiDAR offers the flexibility of an installation at a lower height and tilt angle (aligned more horizontally).

The rotation angle along the z-axis (λ) is to orient the LiDAR measurements toward the true North. **Figure 2-3** illustrates a hypothetical setup in which the system is installed at the Southwest corner of an intersection with its northbound leg fully aligned with the North direction. The LiDAR is directly toward the Northeast corner of the intersection. Consequently, an approximate rotation angle (λ) of 45° is necessary to align the LiDAR measurement with the intersection. This rotation is crucial since the LiDAR point cloud needs to be overlayed and compared with the GIS shapefile of the intersection's road elements, including the intersection area, crosswalks, sidewalks, streets, and bike lanes.

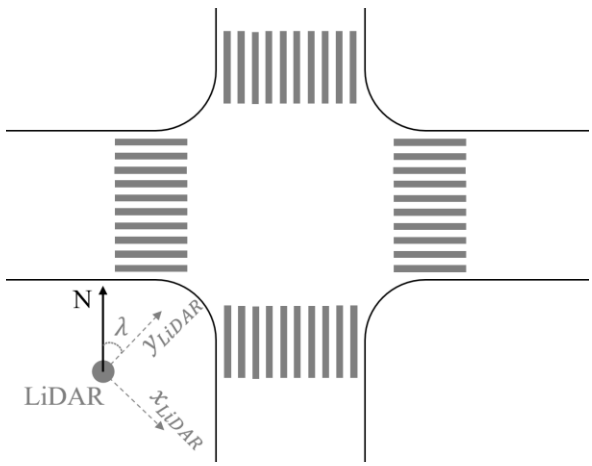


Figure 2-3 Rotation angle (λ) along the z-axis

2.6 Methodology

The main components of the methodology developed for LiDAR-based intersection traffic monitoring and analysis are:

1. LiDAR data preparation
2. Spatial data calibration
3. Background modeling
4. Road user detection and clustering
5. Feature extraction
6. Road user sampling and labeling
7. Road user classification
8. Road user tracking

Figure 2-4 illustrates the flowchart outlining the system's methodology. Except for the Spatial Data Calibration step, which is carried out manually, the entire process is automated using point cloud data processing and machine learning algorithms. Data acquisition and preparation convert binary reading transmitted from LiDAR sensors to a set of frames containing distance, azimuth, channel angle values, and point clouds with X-Y-Z coordinates.

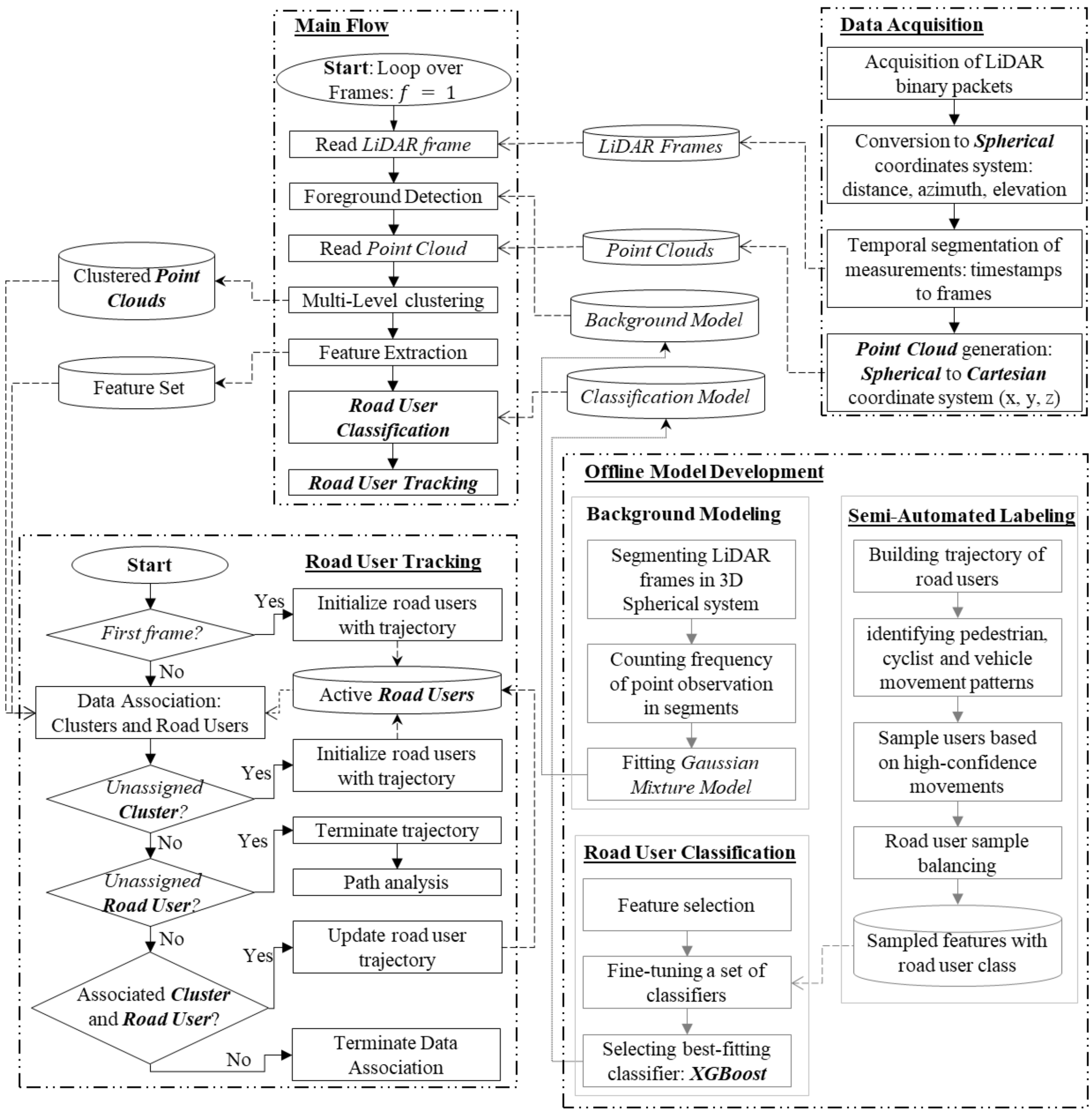


Figure 2-4 Flowchart of the system's algorithm

An initial set of frames from the data preparation is utilized to construct the background model. Subsequently, the remaining data is processed frame-by-frame, allowing the proposed methodology to function in real-time. This process involves foreground detection and clustering foreground point clouds into smaller groups (road users). The clustered x-y-z coordinates of each road user are then transformed into a vector of features, such as the road user's length and width.

The set of feature vectors is used to train a supervised learning algorithm (classification model) that can predict the road user's class as pedestrian, cyclist, car, or truck. Finally, a road user tracking algorithm is implemented to associate and construct the trajectories of each road user. The details of the methodology's components are discussed in this section.

2.6.1 LiDAR data preparation

The LiDAR sensor streams data in a binary format, which is transformed into distance, reflection, azimuth, channel ID, and timestamp values. Distance measurements at each rotation (d_{ch_i,t_j}) are stored as a 2D matrix where the rows represent the channel-ID (ch_i) and the columns represent the azimuth index. If the location of the sensor is assumed to be the center of the coordinate system, then the x-y-z coordinates of a single LiDAR measurement are computed as Equation (2-1):

$$v_{i,j} = \begin{bmatrix} x_{i,j} \\ y_{i,j} \\ z_{i,j} \end{bmatrix} = \begin{bmatrix} d_{ch_i,t_j} \times \cos(\gamma_{ch_i}) \times \sin\left(\alpha_{ch_i,t_j}\right) \\ d_{ch_i,t_j} \times \cos(\gamma_{ch_i}) \times \cos\left(\alpha_{ch_i,t_j}\right) \\ d_{ch_i,t_j} \times \sin(\gamma_{ch_i}) \end{bmatrix} \tag{2-1}$$

Here, $v_{i,j}$ is the x-y-z coordinate, i is the channel ID, j is the column ID, α_{ch_i,t_j} is the azimuth, d_{ch_i,t_j} is the distance to the object, and γ_{ch_i} is the laser channels' vertical angle each (**Table 2-1**).

If the sensor is installed at the height of h and tilted down with β, then the actual x-y-z coordinates of each point are obtained by applying a rotation around x-axis, R_x, and translating in the direction of z-axis by adding the height, as Equation (2-2):

$$\hat{v}_{i,j} = \begin{bmatrix} \hat{x}_{i,j} \\ \hat{y}_{i,j} \\ \hat{z}_{i,j} \end{bmatrix} = R_x \times v_{i,j} + \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\beta) & -sin(\beta) \\ 0 & sin(\beta) & cos(\beta) \end{bmatrix} \times \begin{bmatrix} x_{i,j} \\ y_{i,j} \\ z_{i,j} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix} \tag{2-2}$$

The sensor is also rotated by an angle, λ, along the z-axis (see **Figure 2-3**). In principle, the front of the sensor (the measurements corresponding to the horizontal angle of 0 or $\alpha_{ch_i} = 0$), face the center of the intersection. Then, to align and overlay the LiDAR point cloud with the intersection layout (polygon maps in GIS format), the same rotation is applied as Equation (2-3):

$$\hat{\hat{v}}_{i,j} = \begin{bmatrix} \hat{\hat{x}}_{i,j} \\ \hat{\hat{y}}_{i,j} \\ \hat{\hat{z}}_{i,j} \end{bmatrix} = R_z \times \hat{v}_{i,j} = \begin{bmatrix} cos(\lambda) & sin(\lambda) & 0 \\ -sin(\lambda) & cos(\lambda) & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} \hat{x}_{i,j} \\ \hat{y}_{i,j} \\ \hat{z}_{i,j} \end{bmatrix} \tag{2-3}$$

2.6.2 Spatial data calibration

The spatial data calibration of LiDAR point clouds is the transformation of the x-y-z coordinates of the LiDAR point cloud ($\hat{\hat{v}}_{i,j}$ in Equation (2-3)) into the World Geodetic System 1984 (WGS-84 – EPSG:4326), used in the Global Positioning System (GPS).

First, the x-y coordinates (Longitude-Latitude) of the LiDAR setup point (in WGS 84 format) are converted to one of the North American Datum 1983 map projections for Montreal (NAD-83/MTM Zone 8 – EPSG:32188). Second, the x-y coordinates of the LiDAR point cloud are added with the projected coordinates of the LiDAR setup point. Lastly, the LiDAR points are projected to the WGS-84 map base projection system. This procedure produces a projected point cloud that can be easily mapped and matched with the intersection's elements.

Figure 2-5 (a & b) illustrates the polygons defined in one of the intersections where the LiDAR system was installed. Using Google My Map, the boundaries of all road elements are precisely reconstructed. The location of the LiDAR and camera setup is defined and visualized by a point layer. The road elements defined by a polygon geometry type are streets, bike lanes, sidewalks, crosswalks, curbs, and the inner area of the intersection. The union of all road elements at one intersection forms the coverage area. Any point measured outside the coverage area is eliminated by overlaying the projected point cloud of LiDAR measurements with the polygon layer of the coverage area.

Additionally, the x-y trajectories of the road users are converted to WGS-84 and overlayed by polygons of road elements in the intersection. As a result, each point in the trajectory is tagged with a geospatial label that helps to follow road users' movements in the intersection. The geospatial labels of a trajectory's points help study the traffic movements of road users, especially active/vulnerable road users.

a) Map View b) Satellite View

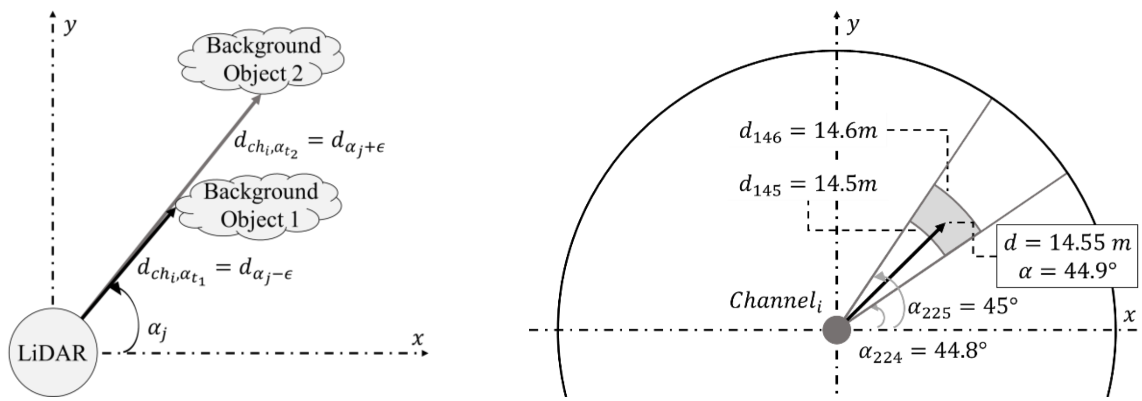
Figure 2-5 Geo-spatial calibration of all road elements in an intersection

2.6.3 Background modeling

As part of the methodology, a 3D background modeling is implemented based on Gaussian Mixture Models (GMMs) introduced by Staufer et el. (34). In the 3D background model, the distribution of the distance measured by each pixel (i^{th} laser channel at j^{th} azimuth) can be associated with a multimodal Gaussian distribution (more than one background point). Building a 3D Gaussian mixture background model in a point cloud environment is time-intensive. This section proposes an alternative and fast algorithm to the identical problem.

Figure 2-6 (a) illustrates a scenario in which the LiDAR is observing two objects around an azimuth (α_j) at two consecutive frames. In this example, the i^{th} laser channel of the LiDAR sensor reports d_1 at azimuth $\alpha_j - \epsilon$ in the first frame and reports d_2 at azimuth $\alpha_j + \epsilon$ in the second frame. The azimuth difference, ϵ, is smaller than the horizontal resolution (δ_α) of the LiDAR which is 0.2°. In most cases, when there is a slight deviation in the azimuth direction, the LiDAR beam still reaches the same object, and the observed distance data remains fixed ($d_1 \cong d_2$).

However, **Figure 2-6 (a)** shows an example that more than one distance value can be expected from the same pixel (the j^{th} azimuth of the i^{th} channel). In this example, there is an edge (moving from one object's surface to another), and the LiDAR beam reaches both sides of it, $d_{ch_i,\alpha_{t_1}}$ and $d_{ch_i,\alpha_{t_2}}$, at time t_1 and t_2. The edge of street poles, signs, walls, and trees' leaves create similar patterns in their point cloud. Most importantly, shaking and vibration of the LiDAR can result in the same pattern.



a- The distance measurement at azimuth α_j b- Sample LiDAR segmentation

Figure 2-6 Sample LiDAR measurement and segmentation of channel *i*

The coverage area of each laser channel per rotation is described as a circular plane within a polar coordinate system, where the radius and angle correspond to the LiDAR's distance and azimuth (**Figure 2-6 (b)**). The azimuth range is $360°$ and the distance range (d_{max}) depends on the intersection size. In this study, the distance range is capped at 50 meters, except for intersections 103 and 110.

For the 3D background modeling, each plane is partitioned into small circular segments, where the angular and radius dimensions of this partitioning are set as $0.2°$ and $0.1m$ (matching the sensor's angular and distance resolution), respectively. As a result, each circular plane is divided into 1800×500 segments ($(360°/0.2°) \times (50m/0.1m)$), forming a 2D matrix called the background matrix. There are 16 or 32 background matrices, one per LiDAR channel.

In the first step, an initial set of 3000 frames collected in the first five minutes is sampled from the data for the 3D background modeling. Azimuth and distance values are discretized for each frame and converted to azimuth and distance indices (α_{id} and d_{id}). These indices correspond to the segment in which the LiDAR measurement belongs. **Figure 2-6 (b)** illustrates the operation of the sensor when the i^{th} channel is reporting a distance of $14.55m$ at an azimuth of about $44.9°$. In this case, the azimuth value belongs to the segment between $44.8°$-$45°$ angles and $14.5m$-$14.6m$

radiuses. The azimuth and distance indices of the corresponding background segment are 224 and 145 ($\alpha_{id} = 44.8°/0.2° = 224$ and $d_{id} = 14.5m/0.1m = 145$).

The frequency of observation for each circular segment over the initial set of frames is tallied and saved as a 3D matrix called BG_{freq} , where its dimension is $n_{channel} \times [1800 \times 500]$. For example, if $d_{2,224}$ (the 224th horizontal measurement of the 2nd channel) is equal to $14.55m$, then in the 2nd background frequency matrix (BG_{freq}), the value of the segment $[224{,}145]$ is increased by one.

Eventually, the background frequency matrices show the frequency of observations belonging to each segment over the initial set of frames. For example, $BG_{freq}[2, 224, :]$ is a vector containing the observation frequency of every segment the 2nd LiDAR channel reaches at an azimuth equal to $44.8°$ and along the distance/radius axis.

Segments lacking observations are excluded (empty space). The remaining segments are clustered into groups of segments based on Euclidean distance between them along the radius axis. The maximum number of clusters of pixels per channel-azimuth pair is restricted to five. All clusters characterized by high frequency build the 3D background model, and the remaining clusters are omitted from the background model (temporary objects in the initial set).

Ultimately, the 3D matrix of observations (denoted as BG_{freq} with $n_{channel} \times [1800 \times 500]$ pixels) is transformed into the 3D background matrix of distances (denoted as BG_{dist} with $[n_{channel} \times 1800] \times 5$ cells), where its contents are frequent distance values (clusters' centroid). In cases where a channel-azimuth pair has, for instance, only three individual clusters, the last two cells are filled with zero and excluded from the foreground detection step.

2.6.4 Road user detection and clustering

Table 2-3 presents road user detection and clustering steps. The process of object detection is a matrix comparison between the new LiDAR distance measurement and the background models. A cell (pixel) is part of a foreground object if the difference between the measurement and the background model is bigger than a pre-defined threshold. Otherwise, it corresponds to part of the background (**Table 2-3: 7-9**). The threshold is defined based on the standard deviation of each

Gaussian mode in the Gaussian Mixture Model (GMM). The output of this step is a two-dimensional binary matrix of zero/one values (FG_{mask}), where one corresponds to foreground pixel and zeros are the background measurements. Then, a de-noising filter is applied to the foreground mask frame (**Table 2-3: 10**). If no other detections are nearby, the pixel is considered a false detection and should be removed from the foreground mask.

Since the point cloud of the foreground object represents the road users (pedestrians, cyclists, or vehicles), they have spatial distribution in the x-y-z coordinate system. Therefore, a multi-level spatial clustering algorithm based on Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is applied to the x-y-z coordinates of all the foreground points (35). The DBSCAN is a powerful clustering method for spatial data and can filter out outliers. DBSCAN requires three hyperparameters: a distance metric, a minimum distance threshold between clusters, and a minimum number of samples for a cluster. The fine-tuning of these three hyperparameters occurred during the clustering algorithm's development and evaluation stages.

The clustering algorithm, incorporated into the LiDAR system, combines three levels of DBSCAN with varying inputs to cluster road users effectively. The three-level clustering algorithm involves clustering pedestrians and cyclists characterized by smaller point clouds and clustering cars and trucks distinguished by larger point clouds.

A DBSCAN is applied with a Squared Euclidean distance metric in level I, utilizing a 0.4m distance threshold and a minimum sample size of six. This stage is designed to cluster pedestrians and cyclists, which are smaller point clouds in closer proximity compared to vehicles. Levels II and III apply two DBSCANs with the Euclidean distance metric, with the objective of clustering cars and trucks. These levels incorporate minimum sample sizes of 4 and 8 and minimum distance thresholds of 1m and 1.75m, respectively. Finally, the three clustering labels are combined to assign a singular cluster label to the point cloud. Clusters in Level I are considered for a potential merging if they share common clusters in both Level II and Level III.

After assigning a cluster label to each point, the point clouds of each cluster are separated and further analyzed to calculate geometric features. Initially, a convex hull polygon is constructed around j^{th} point cloud, with its centroid, $C_{cvxh_j} = \left[x_{cvxh_j}, y_{cvcx_j}\right]$, representing the object's position at each frame. Additionally, the object's height is determined as the 95th percentile of the z-coordinate of the points in the cluster.

Table 2-3 Road user detection, clustering, and feature extraction

Steps	Operation/Process
1	$n_{channels} = 16\ or\ 32$
2 3 4 5 6 7 8 9 10	for *lidar frame* in the *database*: read *azimuth vector* as a_V where its dimension is (1×1800) read *distance matrix* as d_M where its dimension is $(n_{channels} \times 1800)$ convert *azimuth values* [0.0°–360°) to *indices*: $a_Q = \lfloor a_V / 0.2° \rfloor$ convert *distance values* [0m–50m) to *indices*: $d_Q = \lfloor d_M / 0.1\text{m} \rfloor$ read the background model: $BG_{dist}^{current} = BG_{dist}[:, 1{:}\,n_{channels}, a_q]$ compute the Euclidean distance matrix: $\Delta d = d_Q - BG_{dist}^{current}$ detects *foreground mask*: $FG_{mask} = \Delta d > Threshold,$ apply a 3×9 *denoising filter* to remove isolated detections and noises
11 12	convert *foreground mask* to point cloud with x-y-z- coordinates apply a three-level clustering algorithm based on DBSCAN
13 14 15 16 17 18	for j^{th} *clusters* of the current frame: extract the x-y coordinate of the cluster centroid: $C_{cluster\,j}$ build the convex hull polygon and extract the centroid of the polygon run Singular Value Decomposition (SVD) compute the *length, width, height,* and *orientation* of the road users build the rectangle boundary surrounding the road user.

Road users, especially vehicles, have a rectangular shape. However, to determine the length and width of these road users, it is initially assumed that they have an oval shape in the x-y plane. Therefore, the Singular Value Decomposition (SVD) is used to calculate the diameters and orientation of this oval, utilizing the concept of Eigenvectors (36). First, the covariance matrix of the x-y coordinates of the j^{th} clustered point cloud is calculated, as Equation (2-4):

$$Cov_j = \left[X_j - \bar{X}_j \quad , \quad Y_j - \bar{Y}_j\right]^T \cdot \left[X_j - \bar{X}_j \quad , \quad Y_j - \bar{Y}_j\right] = \begin{bmatrix} cov(X,X) & cov(X,Y) \\ cov(Y,X) & cov(Y,Y) \end{bmatrix} \qquad (2\text{-}4)$$

For a clustered point cloud, the eigenvalues (λ_1 and λ_2) and eigenvectors ($\overrightarrow{u_1}$ and $\overrightarrow{u_2}$) of the covariance matrix of the x-y coordinates are computed using Singular Value Decomposition (SVD). These eigenvalues and eigenvectors are utilized to calculate the object's length (l_j), width (w_j), orientation (θ_j) using Equations (2-5 & 2-6):

$$\begin{cases} l_j = \max\left(\sqrt{\lambda_1}\|\overrightarrow{u_1}\|_2, \sqrt{\lambda_2}\|\overrightarrow{u_2}\|_2\right) \\ w_j = \min\left(\sqrt{\lambda_1}\|\overrightarrow{u_1}\|_2, \sqrt{\lambda_2}\|\overrightarrow{u_2}\|_2\right) \end{cases} \tag{2-5}$$

$$\theta_j = \arctan\left(u_1^y / u_1^x\right) \tag{2-6}$$

The object's area and volume are calculated from the length, width, and height. The scaled eigenvectors, transferred to the point cloud centroids, are utilized to construct the rectangle boundary surrounding the road user by forming its four corners, as described by Equation (2-7):

$$P = \begin{Bmatrix} P_{c_1} = C_{cvxh_j} + \left(\sqrt{\lambda_1} \times \|\overrightarrow{u_1}\|_2 + \sqrt{\lambda_2} \times \overrightarrow{u_2}\right) \\ P_{c_2} = C_{cvxh_j} + \left(\sqrt{\lambda_1} \times \|\overrightarrow{u_1}\|_2 - \sqrt{\lambda_2} \times \overrightarrow{u_2}\right) \\ P_{c_3} = C_{cvxh_j} + \left(-\sqrt{\lambda_1} \times \|\overrightarrow{u_1}\|_2 + \sqrt{\lambda_2} \times \overrightarrow{u_2}\right) \\ P_{c_4} = C_{cvxh_j} + \left(-\sqrt{\lambda_1} \times \|\overrightarrow{u_1}\|_2 - \sqrt{\lambda_2} \times \overrightarrow{u_2}\right) \end{Bmatrix} \tag{2-7}$$

2.6.5 Feature extraction

This section discusses the steps for feature extraction, a crucial step for training a machine learning model for road user classification presented in the following sections. The main objective of the feature extraction is to transform the raw x-y-z point cloud of road users into an informative and discriminative feature set for the classification model. The features extracted from sampled clustered point clouds (sampled road users) build the classification dataset.

Section 2.6.4 presents the steps for clustering and computing physical properties of road users' point clouds, including length (l), width (w), height (h), area (h) and volume (v). Theoretically, these physical features are invariable for a road user crossing in any of the intersections, whether they are being monitored with a 16-channel LiDAR or a 32-channel LiDAR. The only difference is that variations in the LiDAR setup and installation or sensor resolution can impact the accurate estimation of the road user's shape. The physical properties are fundamental features included in the feature set.

Moreover, a set of features for road users in different LiDAR setups are extracted from the point cloud of road users. The features of this group are not indifferent to various LiDAR setups or types. These features include the number of points in the road user's point cloud (N_p), point density (ρ_p), the number of LiDAR channels that are engaged with the road user (N_{ch}), and the number of points per LiDAR channel (N_{p-ch}).

Additionally, a third group of features is added to the feature set that represents the spatial information of the road users, including the coordinates of the centroid of the point cloud ($C_{pcl} = [x_{pcl}, y_{pcl}]$), and the average distance of the points to the LiDAR sensor ($\bar{d}$).

Lastly, a rectangle boundary is estimated for each road user. The centroid of the rectangle is the same as the centroid of the surrounding convex hull polygon ($C_{cvxh} = [x_{cvxh}, y_{cvxh}]$). The deviation of the rectangle's centroid (which also is the centroid of the convex hull) from the centroid of the point cloud ($C_{pcl} = \left[x_{pcl}, y_{pcl}\right]$) is another feature added to the feature set and is defined as $\sigma_C = \sqrt{\left(x_{cxvh} - x_{pcl}\right)^2 + \left(y_{cxvh} - y_{pcl}\right)^2}$. For smaller road users, such as pedestrians and cyclists, σ_C leans toward zero, while for bigger road users, it increases by the size of road users.

To differentiate between feature vectors produced by different LiDAR systems, a binary dummy variable (α_L) is being added to the feature set, which is 0 when the system is embedded with a 16-channel LiDAR sensor and 1 when it is embedded with a 32-channel LiDAR system.

The feature vector is defined as f in Equation (2-8), which also includes an offset, the average reflection of the point cloud ($\bar{r}$), and the orientation angels of road users (θ_1 and θ_2):

$$f = \left[1, l, w, h, A, V, N_p, \rho_p, N_{ch}, N_{p-ch}, x_c, y_c, \bar{r}, \bar{d}, \sigma_C, \theta_1, \theta_2, \alpha_L\right] \tag{2-8}$$

Figure 2-7 illustrates a box-and-whisker plot representation of features extracted from clustered point clouds of road users. These plots summarize key statistics of each feature organized by road user class and the type of LiDAR sensor used for data collection. The LiDAR features extracted from a point cloud, such as the number of points, number of channels, number of points per channel, and point density, are significantly higher in the plots corresponding to 32-channel LiDAR. Additionally, the average length, width, height, area, and volume are also higher for critical road user classes such as cars, which indicates that the accuracy of shape estimation is better in the 32-channel LiDAR system.

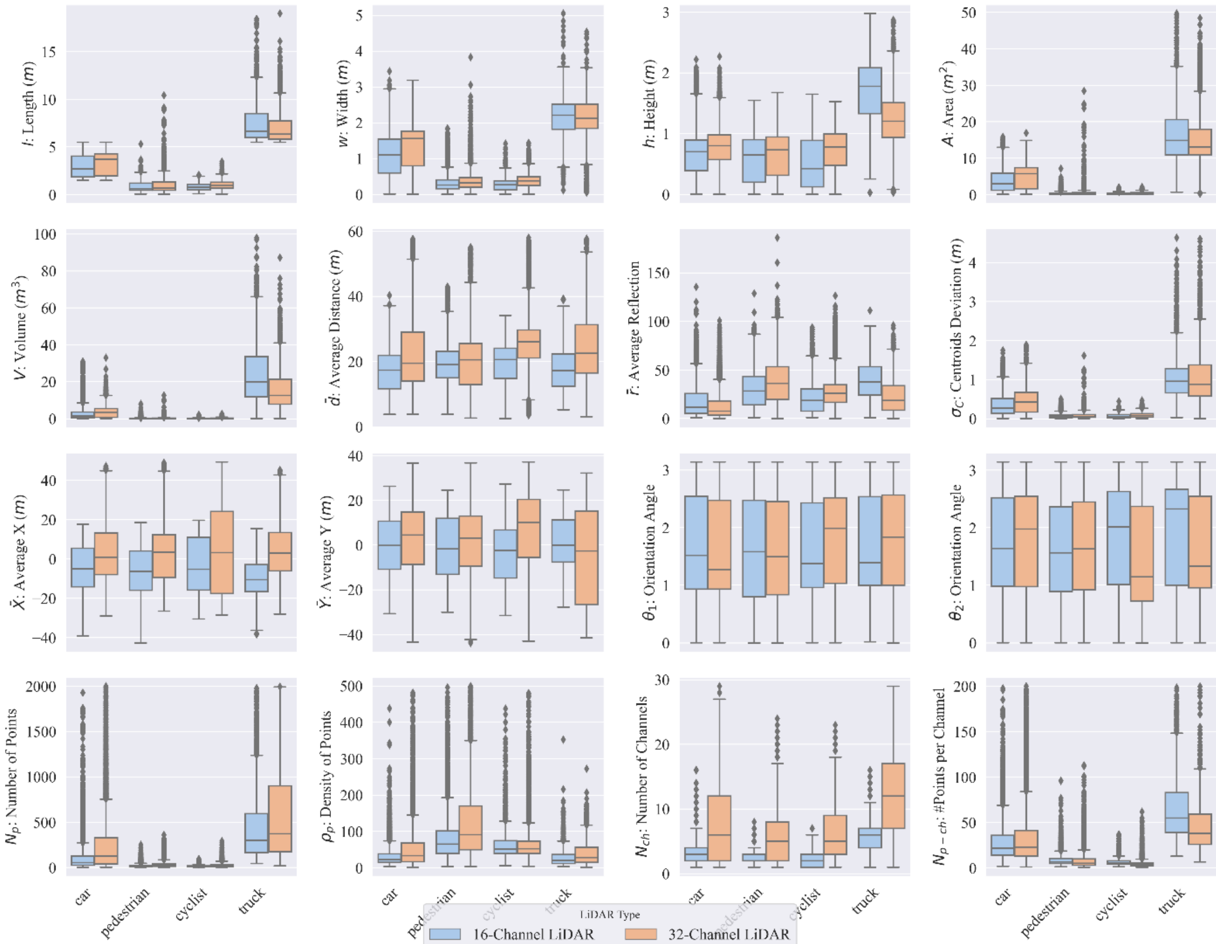


Figure 2-7 Distribution of road users' features collected by 16- and 32-channel LiDARs

2.6.6 Road user sampling and labeling

In addition to feature extraction from clustered point clouds, sampling and labeling are other essential steps for building a road user classification model. Such classification assigns a road user class to each clustered point cloud. The road user classes are pedestrians, cyclists, passenger cars (or equivalent), and trucks (any size).

To ensure accurate classification, a substantial number of labeled samples is essential. However, labeling point cloud samples on a large scale poses significant challenges. Manually labeling extensive point clouds is labor-intensive, especially when identifying and labeling moving objects in snapshots containing 28,800 to 57,600 points per frame. The task is further complicated by point clouds' three-dimensional complexity, object ambiguity, absence of color and texture information, and inconsistent point density. As an alternative, specialized commercial software can be employed, but this often incurs high costs and is less effective for large-scale labeling than image labeling software in camera-based systems.

The application of fixed roadside LiDAR setups for traffic monitoring at intersections presents an opportunity for a semi-automated labeling process. The authors' previous work involved applying an unsupervised learning methodology to monitor traffic at level crossings and extract road users' trajectories (37). As a preliminary step, this research applies the same unsupervised methodology to a sample of LiDAR data from every intersection. Then, the trajectory data are overlaid with the geo-boundaries of intersections, streets, crosswalks, sidewalks, and bike lanes, similar to those illustrated in **Figure 2-5**. Road user classes are assigned to each trajectory based on their movement patterns while observing the speed to exclude unrealistic movements.

For each user's trajectory, a confidence factor is assigned. If the pre-label is 'pedestrian,' the confidence level is high if the trajectory originates from a sidewalk, crosses a crosswalk, and then enters another sidewalk. For cars and trucks, the confidence is considered high if the trajectory starts from one street, crosses an intersection (including crosswalks), and continues onto another street. For cyclists, the confidence factor is high if the trajectory starts and ends on bike lanes and includes passing through a defined 'bike crossing' area, which connects one bike lane to another. Any non-deterministic observations are excluded from the sample pool.

A certain number of frames are selected randomly from each intersection to build a pool of labeled samples. The 100 minutes of four hours of collected data at each intersection are used for sampling. On average, one frame is selected every 2.5 seconds, randomly sampling 2,400 frames from the 100 minutes of data.

For each selected frame, road users are either sampled or discarded based on the geo-spatial location of their point clouds and the associated confidence factor. For example, a pedestrian's point cloud in a selected frame is retained if only it has a high confidence factor and is located on the sidewalk. Similarly, the point cloud is chosen for cyclists if it has a high confidence factor and is situated on a cycle-track or bike lane; otherwise, it is discarded. For cars and trucks, samples are chosen based on whether they are located on the street or near the center of the intersection.

Additionally, if the movement pattern of a pedestrian strictly includes transitioning from a sidewalk to a crosswalk and back to a sidewalk, then samples are also taken from the point cloud on crosswalks, albeit with a reduced confidence factor. Similarly, samples are drawn from the point cloud on bike crossings for cyclists whose movement pattern exclusively consists of transitioning from a bike lane to a bike crossing and back to a bike lane.

The distinction between cars and trucks is made using the 85th percentile of road users' length and width measurements in their trajectories. A user is classified as a truck and included in the sample if the length exceeds 6 meters and the width exceeds 2 meters. Conversely, a user is classified as a car and included if the length is less than 5 meters and the width is less than 2 meters; other samples that do not meet these criteria are discarded.

Table 2-4 presents the results of sampling and labeling. Out of 33,600 sampled frames from 13 data collection sites, 160,980 road users are detected, clustered, and consequently labeled. The sample set has 90,762 pedestrians, 11,614 cyclists, 53,090 cars, and 5,514 trucks. It is expected to observe a higher volume of pedestrians and vehicles in the sample set. The "Pedestrian" and "Car" classes are down-sampled to 18,000 and 25,000 road users to overcome the sampling imbalance and keep high-confidence samples. The balanced feature set contains **60,000** samples. It is worth mentioning that these 60,000 samples do not represent unique road users, as one user can be sampled in multiple frames.

Table 2-4 Road user sampling distribution and labeling

Labeling Type	Road User Class Code	Road User Class Name	Number of Road Users (Original Pool)	Number of Road Users (Balanced Pool)	16-CH LiDAR	32-CH LiDAR
Aggregated	0	Pedestrian	90762	18000	9001	8999
	1	Cyclist	11614	11500	3221	8279
	2	Car	53090	25000	12941	12059
	3	Truck	5514	5500	2891	2609
	-	**All Classes**	**160980**	**60000**	**28054**	**31946**

This approach has its pros and cons. A notable disadvantage is that only select samples per frame are labeled, leaving some samples in each frame unlabeled if they have low confidence scores. This limitation restricts using the labeled dataset for feature classification rather than applying deep neural networks to entire frames. However, this research adopts the former approach and thus is not adversely impacted by this limitation. On the other hand, this method allows for fast labeling by drawing conclusions based on solid evidence of the road user's presence in designated areas. Moreover, the methodology is scalable and can be extended to new collection sites.

2.6.7 Road user classification

Various classification algorithms are trained on the sampled feature set. These classification algorithms are Neural Network Multi-Layer Perceptron (NN-MLP) (38), Support Vector Machine (SVM) (39), K-Nearest Neighbor (KNN) (40), Decision Tree (DT) (41), Random Forest (RF) (42), Logistic Regression Classifier (43), Stochastic Gradient Descent (SGD) (44), Gradient Boosting Machine (GBM) (45), AdaBoost (AB) (46), Light Gradient-Boosting Machine (LGBM) (47), and eXtreme Gradient Boosting (XGBoost) (48).

The classification dataset is split into training, validation, and test sets to select the best-fitted classification model. The ratio of this split is 70%, 5%, and 25%, respectively. Each model is calibrated on the validation set to tune the best values for hyperparameters, trained with the training set to learn the optimal weights and coefficients of the model, and evaluated on the test set to report the performance of the classification model and choose the one with the best performance.

Once calibrated and trained, the best-fitted classification model is integrated into the LiDAR system. Afterward, the feature extraction process generates a feature vector for each observed road user (discussed in Equation (2-8)). The classification model takes in the road user feature vector and predicts the road user class.

2.6.8 Road user tracking

The next step of the LiDAR system's methodology is analyzing the movement patterns of road users in an intersection. Road user tracking involves associating point clouds corresponding to the same road user across different frames. The outcome of this process is a trajectory comprising the centroids of the point clouds. This trajectory offers insights into the movement pattern of road users within various sections of the intersection, their velocity and acceleration, their interaction with other road users, and their origin and destination while crossing the intersection.

Table 2-5 describes the proposed algorithm for the LiDAR system's road user tracking. The algorithm remains inactive (idle) until the first detection of a road user. Upon initiation, new *trackers* are set up for road users entering or appearing for the first time within the coverage area (**Table 2-5: 1-2**). Each tracker is initialized with the current position of the road user, and its velocity and acceleration vectors are both set to $\vec{0}$.

Table 2-5 Road user tracking

Step	Operation/Process
1 2	**for** the i^{th} road user at the *first lidar frame* in the *database*: initiate a *road user's tracker* with the observed position, zero velocity, zero acceleration, and a Kalman Filter.

3
4
5
6
7
	for *frame* k in the list of *consecutive lidar frames* in the *database*:
	    predict the *positions* of the N *road users* using KF: P_k
	    extract the *positions* of the M *observed road users*: O_k
	    calculate the cost matrix for any pair of *observation-prediction*: $C(O_i, P_j)$
	    solve Data Association for the cost matrix $C_{M\times N}$.

8	**for** every pair of *associated observation-prediction* $(O_i, P_j,)$: updatej^{th} *tracker* and its KF's state variables with i^{th} *observation*.
9	**for** every *unassigned observation:* initiate a *road user's tracker*.
10	**for** every *unassigned tracker:* terminate trackers if not associated with an observation for δf frames.

The tracking component of the proposed LiDAR system integrates a six-state Kalman Filter (KF) to predict road users' x-y trajectory (49). The main variables of the Kalma Filter are the vector of state variables ($X = \left[x, V_x, a_x, y, V_y, a_y\right]^T$) and vector of observed variables (Z). The main parameters of the Kalman Filter are the state transition matrix (A) and measurement matrix (C). The transition matrix (A) incorporates kinematic equations for constant acceleration movement in the Cartesian coordinates system, as Equation (2-9):

$$A = \begin{bmatrix} 1 & \Delta\text{t} & \Delta t/2 & 0 & 0 & 0 \\ 0 & 1 & \Delta\text{t} & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \Delta\text{t} & \Delta t/2 \\ 0 & 0 & 0 & 0 & 1 & \Delta\text{t} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{2-9}$$

Here Δt represents the time interval between consecutive LiDAR measurements. The estimated state vector (X^{+}) for the road user in the next frame is derived by applying the transition matrix (A) to the current state vector (X), expressed as $X^{+} = AX$. Since the LiDAR detection component solely captures the position of the road user, the observation vector is a two-dimensional array denoted as $Z = [x, y]^{T}$.

The prediction of the road user's position in the next frame (P^{+}) is derived by applying the observation matrix (C) to the estimated state vector, expressed as $P^{+} = CX^{+}$. The estimated velocity vector of the road user ($V_{x,y}$) is derived by applying the velocity observation matrix (C_V) to the estimated state vector, expressed as $V_{x,y} = C_V X$. The observation matrices of position and velocity vectors are defined as in Equations (2-10 & 2-11):

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{2-10}$$

$$C_V = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{2-11}$$

Before processing a LiDAR frame (e.g., frame k), the road user trackers from the preceding frame ($k - 1$) predict the coordinates of the road users' positions at the succeeding frame. The prediction set of N active trackers is defined as $P_k = \left\{P_{j,k} = \left[x_{p_j}, y_{p_j}\right]\right\}_{j=1:N, frame=k-1}$, where $P_{j,k}$ represents the projected position by the tracker j at the frame k.

At the current frame (k), the LiDAR system captures the point clouds of road users and builds the observation set of M road users, which is denoted as $O_k = \{O_{i,k} = [x_{o_i}, y_{o_i}]\}_{i=1:M, frame=k}$, where $O_{i,k}$ represents the observed position of road user i at the frame k.

Data association is the next crucial step in the road user tracking component. A data association algorithm establishes connections between different detections of the same road user across consecutive frames, mitigating known issues of multiple road user tracking, such as temporary disappearances of road users due to occlusion.

A data association method employing the Munkres algorithm (50) is integrated into the tracking step (**Table 2-5: 3-10**). This technique associates the pairs of users from two successive frames by minimizing a customized cost function that utilizes three different distance matrices between observations (rows of the matrices) and predictions (columns of the matrices). The base matrix, denoted as $d_c(O_i, P_j)$, is the distance between each pair of predicted and observed road users' positions in the x-y plane (51). To leverage the LiDAR's ability to observe the point cloud of the road users, the second matrix, denoted as $d_{PCL}(O_i, P_j)$, is defined as the distance between the current and previous frames' point cloud. The point clouds of the prior frame are projected to the current frame using the two-dimensional velocity vector estimated by the Kalman Filter. The third metric is characterized by the distance between the feature vector of the road users, denoted as $d_{\vec{f}}(O_i, P_j)$, as Equation (2-12):

$$d_{\vec{f}}(O_i, P_j) = \beta \times \frac{1}{n_f} \sum_{k=1}^{n_f} \frac{1}{\alpha} \times \left[\left(1 - \exp\left(- \frac{\left| f_{O_i}^k - f_{P_j}^k \right|}{\min\left(f_{O_i}^k, f_{P_j}^k \right)} \right) \right) + \left(1 - \exp\left(- \frac{\left| f_{O_i}^k - f_{P_j}^k \right|}{\max\left(f_{O_i}^k, f_{P_j}^k \right)} \right) \right) \right] \quad (2\text{-}12)$$

where k is the feature index, n_f is the number of features, and α denotes the maximum distance between two features or between a given feature and infinity, as specified in Equation (2-13). Both n_f and α scale their right-side expression to 0 and 1. The feature distance matrix, $d_{\vec{f}}(O_i, P_j)$, is scaled up by $\beta = \sum \sum d_c(O_i, P_j)$ to be aligned within the same range as d_c.

$$\alpha = \left(1 - \exp\left(- \frac{|\infty - x|}{\min(\infty, x)} \right) \right) + \left(1 - \exp\left(- \frac{|\infty - x|}{\max(\infty, x)} \right) \right) \cong 1.632 \quad (2\text{-}13)$$

The overall cost matrix, denoted as $C(O_i, P_j)$, between the pair of observation and prediction is determined by summing the three distance matrices, as Equation (2-14):

$$C(O_i, P_j) = d_c(O_i, P_j) + d_{PCL}(O_i, P_j) + \left(\beta \times d_{\vec{f}}(O_i, P_j) \right) \quad (2\text{-}14)$$

First, for the pair of observation-prediction (O_i, P_j) identified by the data association, the Kalman Filter of j^{th} tracker undergoes an update with the observed position of i^{th} road user. Each observed road user not assigned to a tracker from the previous frame is treated as a new user, leading to a tracker initialization. Similarly, each tracker not associated with a new observation is flagged as '*not observed*' and terminated if the '*not observed*' status persists for δf frames. δf serves as a threshold indicating the system's tolerance for the absence of new observations for a given tracker.

It is dynamically adjusted based on estimated speed and distance from the coverage area. However, to prevent data association errors, δf is constrained to a range of 5 to 15 frames.

Following the road users' departure from the intersection, their trajectory is terminated, converted to the *WGS-84* projection system, and overlayed with the polygons of the intersection elements. Therefore, the geospatial information assigned to a user, such as road users' origin, destination, and crossing section, is instrumental in distinguishing pedestrians from other types of road users.

Second, all the road user classes assigned to each point of the road user's trajectories are merged to create a set of road user classes. The most repeated class in this set is chosen as the representative class of the road user. The results of the above two steps are combined to assign a final class to the road user. The decision is made by considering the outcomes of the second step while ensuring accurate identification of pedestrians, as discussed in the first step.

2.7 Performance Measure and Evaluation

The proposed LiDAR-based methodology incorporates several algorithms requiring tuning, calibration, training, and evaluation. This section is dedicated to assessing the performance of the individual components within the proposed methodology and evaluating the impact of low (16-channel) and high-resolution (32-channel) LiDAR sensors.

2.7.1 LiDAR road user detection

The first component of the evaluation focuses on road user detection, which includes background modeling, foreground detection, and clustering algorithms. This evaluation also considers the performance of the LiDAR setup. Generally, if a road user crosses a crosswalk and enters the intersection, its detection is subject to assessment. Therefore, if the LiDAR is installed in such a way that it creates a blind spot in a particular area, any resulting missed detections are factored into the analysis.

Table 2-6 presents the aggregated accuracy of road user detection and clustering by the two LiDAR systems at 13 intersections. 150 frames of each intersection were randomly sampled and manually compared with synced images from video data. The accuracy of road user detection increases by upgrading the LiDAR system with a higher-resolution LiDAR sensor.

The 16-channel LiDAR system accurately detected 3,818 road users out of 4,254 (89.8%) and clustered them into individual point clouds. On the other hand, the 32-channel LiDAR system correctly identified and clustered 3,608 road users out of 3,830 (94.2%). Specifically, the detection and clustering rates have increased from 88.7% to 93.0% for pedestrians, 88.5% to 93.6% for cyclists, and 90.3% to 95.3% for cars. These results are categorized based on road user classes manually observed in the video data. The accuracy of road user classification is discussed in the next section.

Table 2-6 Detection accuracy of LiDAR systems installed at urban intersections

LiDAR Type	*Performance Measure*	*Pedestrian*	*Cyclist*	*Car*	*Truck*	Total
16-Channel LiDAR	Ground Truth	1,848	532	1,705	169	4,254
	LiDAR Detection	1,639	471	1,539	169	3,818
	Detection Accuracy	**88.7%**	**88.5%**	**90.3%**	**100.0%**	**89.8%**
32-Channel LiDAR	Ground Truth	1,535	690	1,499	106	3,830
	LiDAR Detection	1,427	646	1,429	106	3,608
	Detection Accuracy	**93.0%**	**93.6%**	**95.3%**	**100.0%**	**94.2%**

Table 2-7 Detection accuracy of LiDAR systems installed at urban intersections

LiDAR Type	*Intersection ID*	*Intersection Name*	*Pedestrian*	*Cyclist*	*Car*	*Truck*	*Total*
16-Ch	101	Sainte Famille - Milton	88.6%	87.8%	93.5%	100.0%	92.0%
	102	Papineau – Sherbrooke E	83.3%	80.0%	88.4%	100.0%	89.1%
	103	Atwater – Sherbrooke W	81.2%	77.8%	87.5%	100.0%	90.4%
	104	De La Roche – Marie Anne E	90.8%	89.9%	93.9%	100.0%	92.8%
	105	Coloniale – Rachel E	90.2%	90.4%	92.6%	100.0%	92.6%
	106	Girouard - Monkland	87.2%	87.8%	94.1%	100.0%	91.0%
	107	University - Milton	91.5%	90.5%	93.2%	100.0%	91.8%
32-Ch	108	Hutchison - LaurierE	94.1%	95.9%	95.0%	100.0%	96.6%
	109	Sainte Famille – Prince Arthur	95.2%	92.9%	98.4%	100.0%	97.3%
	110	Parc - PineW	89.6%	93.5%	96.0%	100.0%	93.5%
	111	Saint Denis – Saint Joseph E	92.0%	92.4%	92.9%	100.0%	94.4%
	112	Parthenais – Rachel E	94.4%	96.9%	98.5%	100.0%	96.7%
	113	University - Milton	93.6%	92.7%	96.9%	100.0%	96.4%

Table 2-7 presents the results of road user detection for each intersection. The LiDAR systems were installed at a similar location but on different dates and times. Both sensors were installed at the *Rue University–Rue Milton* intersection (site IDs #107 and #113), where the accuracy of the 16- and 32-channel LiDAR systems was 91.8% and 96.4%, respectively.

Intersections identified by IDs 102, 103, 110, and 111 demonstrate a below-average performance in detecting and reconstructing the point clouds of road users. The average intersection and crosswalk area of these four intersections is 756m^2, significantly higher than the rest, whose average is 312m^2 (**Table 2-2**). Therefore, the LiDAR systems had to be installed relatively farther from the center of the interactions (**Table 2-2**). With an increase in the LiDAR distance to the intersection, the gap between LiDAR channels increases.

2.7.2 Road user classification – base scenario

This section aims to assess the effectiveness of the road user classification component of the LiDAR system. The primary focus of this evaluation involves selecting and calibrating the optimal classification algorithm from the models introduced in **Section 2.6.7**. Each classification model is calibrated and trained using 16-channel LiDAR samples and 32-channel LiDAR samples separately.

In the base scenario, the classification feature set is split into 70% training, 5% validation, and 25% test samples. The hyperparameters of each classification model are fine-tuned using training and validation sets. Afterward, the classification models learn the optimal coefficients using the training set. Finally, each calibrated and trained model is applied to the unseen samples of the test set, and their Correct Classification Rates (CCR) are compared to choose the most fitting model.

Table 2-8 presents the results of the road user classification in the base scenario. The CCR on the test set is the primary indicator for choosing the best classification model. The XGBoost classifier outperforms every other model. The CCR of XGBoost for datasets with samples from the 16-channel and 32-channel was 0.91 and 0.95, respectively. In general, the performance of each classifier is higher on the dataset comprised of samples from 32-channel LiDAR. This observation aligns with the expectation that higher-resolution LiDAR can better reconstruct the point cloud of road users.

Table 2-8 Base scenario – correct classification rates of models

Model Name↓	Train		Test	
LiDAR Type →	16-CH	32-CH	16-CH	32-CH
XGB	0.95	0.98	**0.91**	**0.95**
NN-MLP	0.90	0.91	0.88	0.90
SVM	0.73	0.74	0.74	0.75
KNN	0.83	0.85	0.82	0.85
DT	0.94	0.97	0.87	0.89
RF	0.85	0.81	0.85	0.82
LRC	0.80	0.79	0.80	0.81
SGDC	0.70	0.70	**0.70**	**0.72**
GBC	0.88	0.90	0.89	0.90
LGBM	0.90	0.92	0.88	0.90

Table 2-9 presents the detailed CCR of the XGBoost classifier for each road user class on the test datasets. The "*Cyclist*" and "*Pedestrian*" classes have a lower CCR than the rest. In the case of higher resolution LiDAR, 133 of 2,240 "*Pedestrian*" samples are classified as "*Cyclist,*" and 132 of 843 "*Cyclist*" samples are classified as "*Pedestrian.*"

Table 2-9 Base scenario – correct classification rates per each road user class

		Training Set		Test Set	
Class Name	LiDAR→ Label ↓	16-CH	32-CH	16-CH	32-CH
Pedestrian	0	0.95	0.97	0.89	0.92
Cyclist	1	0.95	0.97	0.85	0.93
Car	2	0.95	0.98	0.94	0.98
Truck	3	0.95	0.98	0.94	1.00
All Classes	-	0.95	0.98	0.91	0.95

2.7.3 Road user classification – alternative scenarios for performance evaluation

Two scenarios are devised to evaluate the robustness of the selected classification model (XGBoost). The first scenario (I) evaluates the importance of feature selection and the impact of different groups of features on the model's performance. For a better comparison, the features are grouped into four following categories:

- *Physical Features*: length, width, height, area, and volume (denoted as $[l, w, h, A, V]$).
- *LiDAR Features*: number of points, density of points, number of channels, and number of points per channel (denoted as $[N_p, density_p, N_{ch}, N_{p-ch}]$).
- *Spatial Attributes*: the x-y coordinate of road users' centroid (denoted as $[x_c, y_c]$).
- *Remaining Features*: average distance, average reflection, orientation angles, centroids' deviance (denoted as $[\bar{r}, \bar{d}, \theta_1, \theta_2, \ d_{C_{cxvh}, C_{pcl}}]$).

Table 2-10 summarizes the CCR of XGBoost on the given dataset (75%/25% training-test split) in Scenario I, where the classification model is trained with the original feature set except for each group of features specified in the first column. The CCR on the 32-channel dataset is 0.95. By excluding physical features, the CCR decreases by -0.06 and drops to 0.89, indicating that the features of the physical group significantly impact the model's training. By excluding the group of LiDAR features, spatial attributes, and the fourth group, the CCR decreases by -0.03, -0.03, and -0.02, respectively. The reductions in CCR resulting from removing the last three groups of features are below 0.03.

Nevertheless, the classifier trained on these features (excluding the physical group) demonstrated a CCR of 0.89, inferring that the union of the last three groups proves resourceful for the classification. The importance of physical features is greater for the classifier trained with samples from a lower-resolution LiDAR. The exclusion of the physical features decreases the CCR by -0.09 for the feature sets with only 16-channel samples.

Table 2-10 Scenario I – Feature importance and their impact on CCR of the test set

Feature Set		16-CH	32-CH
CCR with all Features		0.91	0.95
CCR Reduction by excluding	Physical Features	**-0.09**	**-0.06**
	LiDAR Features	-0.03	-0.03
	Spatial Features	-0.03	-0.03
	Other Features	-0.02	-0.02

Scenario II discusses the results of a Leave-One-Out analysis in which, at each step, all the samples of one intersection are separated and defined as the test set, while the samples from the other intersections collected by the same LiDAR sensor form the training set. Therefore, no sample from the test intersection is observed during the training phase, which makes it challenging for the classification model to predict road user classes accurately.

Table 2-11 reports the results of the Leave-One-Out analysis. The sample sets include the original features set except for Spatial attributes (the coordinates of the road users' centers: $(x_c,\ y_c)$) since these attributes are specific to the intersections. The average CCR of each scenario is compared against the CCR of the model trained without Spatial features (from **Table 2-10**). Generally, the average performance of the 16-channel and 32-channel LiDAR systems is reduced by 0.08 and 0.06, respectively. Specifically, the Correct Classification Rate drops significantly when an intersection differs (with a larger area) from the others (Site IDs 102, 103, 110, and 111).

Table 2-11 Scenario II – Leave-One-Out analysis

Test Site ID	Intersection Name	Area (m^2)	Distance to Intersection (m)	16-CH	32-CH
101	*Sainte Famille - Milton*	278	13.1	0.81	-
102	*Papineau – Sherbrooke E*	**690**	17.1	**0.77**	-
103	*Atwater – Sherbrooke W*	**765**	21.5	**0.78**	-
104	*De La Roche – Marie Anne E*	268	11.5	0.83	-
105	*Coloniale – Rachel E*	246	12.3	0.79	-
106	*Girouard - Monkland*	412	15.6	0.83	-
107	*University - Milton*	296	10.7	0.83	-
108	*Hutchison - LaurierE*	471	19.2	-	0.87
109	*Sainte Famille – Prince Arthur W*	**355**	14	-	0.87
110	*Parc - PineW*	**899**	**40.1**	-	**0.85**
111	*Saint Denis – Saint Joseph E*	669	18.5	-	**0.80**
112	*Parthenais – Rachel E*	284	13.4	-	0.86
113	*University - Milton*	231	8.7	-	0.88
Average CCR Leave-One-Out				**0.80**	**0.86**
Original CCR without Spatial Features				0.88	0.92
Difference between Average CCR and Original CCR				**-0.08**	**-0.06**

2.7.4 Road user tracking performance evaluation

This section discusses an evaluation of the performance of the road user tracking component of the methodology, focusing on both low- and high-resolution LiDAR sensors. In order to assess the overall efficacy of the proposed methodology in traffic monitoring contexts, a ground truth dataset comprising road user counts within the first 30-minute segment of data collection at each intersection is being compiled manually from synchronized recorded video. The number of trucks is significantly lower than cars. Therefore, trucks and cars are combined and reported as vehicles.

Table 2-12 presents the results of manual counts and LiDAR trajectory counts. A comparison is conducted by calculating the Percentage Difference (PD), which is defined as the difference between the LiDAR and the manual counts and divided by the manual counts. The PD manifests as a negative value in scenarios where the LiDAR system undercounts. In contrast, a positive PD signifies an overcounting by the system. Undercounting is often linked to errors in the detection or classification of road users. Overcounting can be associated with road user tracking failure in constructing the entire trajectory and further splitting it into two road users for a short period.

A substantial volume of vehicular traffic was observed in intersections 102, 103, 110, and 111. This observation was particularly evident at intersections 102 and 103, where a considerable count of vehicles was recorded within thirty minutes. It is pertinent to mention that the data collection employing the higher resolution LiDAR system was conducted during the fall of 2021 when the pandemic notably influenced traffic volumes. Additionally, a few of the intersections were selected due to their higher volume of cyclists (IDs: 105, 106, 107, and 111).

Generally, the lower-resolution LiDAR tends to overcount cyclists. This issue often arises due to the clustering of passenger cars. When a passenger car is partially in the LiDAR’s blind spot, its front and rear may be split into two distinct point clouds, resembling those of cyclists. Additionally, pairs of pedestrians are sometimes misclassified as cyclists. However, it is worth mentioning that the total number of cyclists observed at some intersections was low, resulting in a large absolute percentage difference when misclassifying another road user as a cyclist.

The last two rows of the table report the Weighted Average Absolute Percentage Deviation for the two systems, where the absolute percentage error for each road user class at each intersection is weighted by the total volume of road users of the corresponding intersection divided by the total per each LiDAR system.

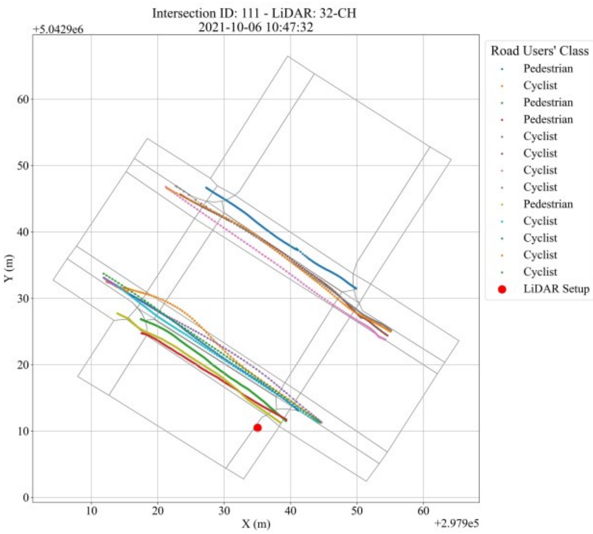
For pedestrians, the WAAPD of the lower and higher resolution LiDARs is 7% and 5%, respectively. For cyclists, the WAAPD is 23% for the lower resolution LiDAR and 7% for the higher resolution. For vehicles, the WAAPD is 10% for the lower resolution LiDAR and 6% for the higher resolution. The average of WAAPD over three categories is 13% for the lower resolution and 6% for the higher resolution.

Table 2-12 Aggregate count validations in the first 30-minute interval per intersection

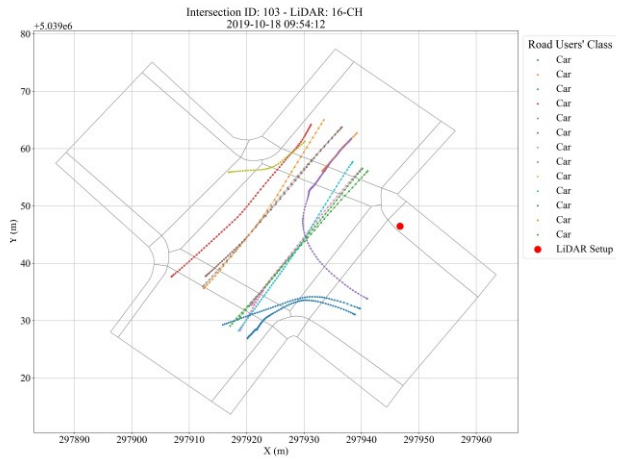
Site ID - LiDAR Type	Total	Vehicles	Pedestrians	Cyclists	Total	Vehicles	Pedestrians	Cyclists	Total	Vehicles	Pedestrians	Cyclists
	Ground Truth – Manual Counts				LiDAR Trajectory Counts				Percentage Error			
101 – 16-ch	446	202	158	86	430	206	154	70	-4%	2%	-3%	**-19%**
102 – 16-ch	1,847	1,556	220	71	2,062	1,754	230	78	12%	13%	5%	10%
103 – 16-ch	1,309	1,154	134	21	1,184	1,043	112	29	-10%	-10%	**-16%**	38%
104 – 16-ch	379	224	99	56	362	199	97	66	-4%	**-11%**	-2%	18%
105 – 16-ch	597	295	186	116	583	288	162	133	-2%	-2%	**-13%**	15%
106 – 16-ch	524	340	150	34	575	370	162	43	10%	9%	8%	26%
107 – 16-ch	664	265	284	115	691	261	274	156	4%	-2%	-4%	36%
108 – 32-ch	376	211	129	36	373	214	128	31	-1%	1%	-1%	**-14%**
109 – 32-ch	290	131	135	24	291	134	129	28	0%	2%	-4%	17%
110 – 32-ch	1,094	871	153	70	1,162	924	166	72	6%	6%	8%	3%
111 – 32-ch	1,167	930	106	131	1,223	993	100	130	5%	7%	-6%	-1%
112 – 32-ch	309	155	44	110	291	147	44	100	-6%	-5%	0%	-9%
113 – 32-ch	484	167	158	159	453	159	150	144	-6%	-5%	-5%	-9%
Total – 16-ch	5,766	4,036	1,231	499	5,887	4,121	1,191	575	-	-	-	-
Total – 32-ch	3,720	2,465	725	530	3,793	2,571	717	505	-	-	-	-
								WAAPD 16-ch	**13%**	**10%**	**7%**	**23%**
								WAAPD 32-ch	**6%**	**6%**	**5%**	**7%**

Figure 2-8 illustrates two snapshots of each LiDAR system's performance, featuring several road users simultaneously. The trajectories are superimposed onto a base map of each intersection, calibrated, and segmented into its various elements using the *NAD-83/MTM Zone 8 – EPSG:32188* projection. **Figure 2-8 (a)** displays the paths of both cyclists and pedestrians during a bike signal

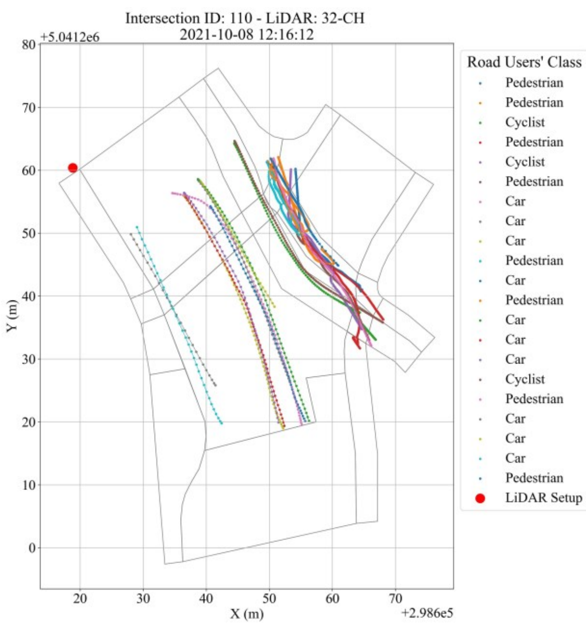
phase, highlighting a group of cyclists traversing the intersection in the designated bike crossing area. **Figure 2-8 (c)** shows the busiest intersection captured during the data collection period with the higher resolution LiDAR system.



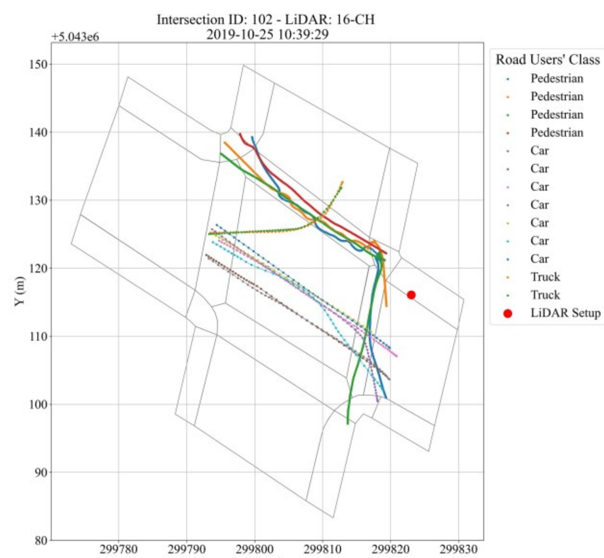
a) 32-channel LiDAR system



b) 16-channel LiDAR system



c) 32-channel LiDAR system



d) 16-channel LiDAR system

Figure 2-8 Snapshots of LiDAR systems' outputs

The trajectories of road users are constructed and continuously updated by the tracking algorithm, which necessitates an evaluation of the tracking component itself. In this tracking framework, each road user is associated with two sets of x-y coordinates: the observed x-y coordinates, representing the centroids of the clustered point cloud for each user, and the predicted x-y coordinates, estimated using the Kalman Filter. The primary goal is accurately predicting the road user's next position, ensuring it closely matches the subsequent observation. This level of accuracy is crucial for the data association algorithm to link each observation reliably with the correct prediction.

The Average Displacement Error (ADE) is a key metric for evaluating performance in object tracking problems, as outlined in reference (52). This metric measures the deviation between the model's predicted positions and the observed positions of road users. For a specific road user k, ADE_k is calculated as Equation (2-15):

$$ADE_k = \frac{1}{N_k} \sum_{t=t_i}^{t_i+N_k\Delta t} \sqrt{\left(x_{P_i} - x_{O_i}\right)^2 + \left(y_{P_i} - y_{O_i}\right)^2} \qquad (2\text{-}15)$$

where t_i is timestamp, δt is the sampling time of LiDAR ($\delta t = 0.1$), N_k is the number of frames that LiDAR is observing the road user, $\left[x_{O_i}, y_{O_i}\right]$ correspond to the observed and $\left[x_{P_i}, y_{P_i}\right]$ is the prediction position of the road user.

The Prediction Accuracy (PA) is the percentage of predictions that fall within specific tolerances of their respective observed positions. For each point along their trajectories, the distances between the observed and predicted coordinates are evaluated against reference values unique to each road user. These reference values are set as the 50th percentiles of the road users' displacements per frame while the LiDAR observed them.

Table 2-13 reports the overall performance of the LiDAR systems based on ADE and PA metrics. This report compiles data from the first 30 minutes of processed LiDAR data at each intersection, subsequently aggregated for each type of LiDAR system. This report is auto-generated, and the observation and prediction of the LiDAR system are compared. No manual ground truth is used to generate this report.

The LiDAR systems constructed trajectories for over 9,680 road users, encompassing more than 660,000 data points. On average, road users are captured in 69 LiDAR frames, equivalent to 6.9 seconds. The table also reports the percentage of unobserved data points within the trajectory files for each user. Generally, the higher resolution LiDAR system exhibits an 11% rate of missing trajectory points, compared to 15% for the lower resolution system. These percentages are with respect to those counted trajectories. Some users may have trajectories that are not detected (missing detection) and can not be summarized in this table.

The ADE for the 32-channel and 16-channel LiDAR systems are 0.37m and 0.40m, respectively. Notably, the ADE for pedestrians is significantly lower than for other road users, which can be attributed to their smaller size and slower speed in comparison to the other three user categories. Furthermore, the ADE for cyclists is higher, typically ranging from 0.33m to 0.35m. This increase is mainly attributable to cyclists' higher speeds than pedestrians. The ADE for cars and trucks is 0.46m and 0.43m at intersections where higher resolution LiDAR was installed and 0.50m and 0.42m, respectively when lower resolution LiDAR was installed.

The weighted averages of the PA (50th percentile) are 0.78 and 0.80 for the lower and higher resolution LiDAR systems, respectively. This indicates that for 80% of the trajectory points, the predicted positions of the road users were as close as 50% of their average displacement.

Table 2-13 Performance metrics of road user tracking – 30-minute period

Performance Metric	**User Class → LiDAR ↓**	**Car**	**Truck**	**Cyclist**	**Pedestrian**	**Total or Weighted Average**
LiDAR counts	16	3,812	309	575	1,191	5,887
	32	2,351	220	505	717	3,793
Number of observed trajectory points	16	162,082	15,816	42,548	150,835	371,281
	32	121,127	11,561	52,629	109,670	294,987
% unobserved trajectory points	16	15%	15%	16%	11%	15%
	32	10%	9%	9%	13%	**11%**
ADE (m)	16	0.50	0.42	0.35	0.11	0.40
	32	0.46	0.43	0.33	0.11	**0.37**
PA (50^{th} percentile)	16	0.74	0.72	0.78	0.82	0.76
	32	0.76	0.75	0.83	0.83	**0.78**

2.8 Conclusion and Future Work

This research introduces a supervised learning approach for monitoring traffic at intersections, employing LiDAR technology with VLP-16 and VLP-32c sensors. An extensive data collection initiative led to the development and installation of LiDAR proof-of-concept systems at various intersections in Montreal. The methodology implemented for the LiDAR system includes background modeling, road user detection, classification, and tracking. The 3D background modeling involves the creation of a customized Gaussian Mixture Model in a three-dimensional spherical coordinate system. Road users are clustered using multi-level algorithms based on DBSCAN. A classification model, developed based on XGBoost, is calibrated with ground truth data and then applied to the clustered point clouds. Lastly, based on Kalman Filters, the tracking component identifies road users within the coverage area and constructs their trajectories.

The empirical results from the outputs of the system are insightful. The 16-channel LiDAR system achieved an accuracy rate of 89.8% in detecting and clustering road users, while the 32-channel system surpassed this with a 94.2% accuracy rate. This improvement was especially notable in detecting pedestrians, cyclists, and cars.

The XGBoost classifier proved to be the most effective model, showing high Correct Classification Rates (CCR) across various road user categories. The CCR of all classes is 0.95 and 0.91 for high and low-resolution LiDARs, respectively. However, the study also revealed challenges in classifying pedestrians and cyclists. In the case of pedestrians, the CCR of the classifier is 0.92 and 0.89 for high and low-resolution LiDARs, respectively. In the case of cyclists, the CCR of the classifier is 0.93 and 0.85 for high and low-resolution LiDARs, respectively.

This research identified undercounting and overcounting errors, underscoring the complexities of accurately tracking and classifying road users. The low-resolution LiDAR mainly exhibits an overcounting of cyclists. The weighted average absolute percentage difference is 6% and 13% for high and lower resolution, respectively.

Additionally, the Average Displacement Error of the tracking component reports an average of 0.37m for the 32-channel system and 0.40m for the 16-channel system. On average, trajectories captured by high-resolution LiDAR experience an 11% unobserved state, while those captured by low-resolution LiDAR experience a 15% unobserved state, during which no detection is made and only prediction is performed.

In conclusion, the effectiveness of LiDAR systems, especially the higher performance of the 32-channel system in detection accuracy and classification, establishes a foundation for future developments in traffic analysis. The proposed methodology can be adapted to process and analyze data in real time with a powerful computing machine. This research underscores the potential of LiDAR technology in providing detailed and accurate traffic monitoring and its potential for improvements in urban traffic monitoring.

Acknowledgments: Funding for this project is provided in part by the Natural Sciences and Engineering Research Council (NSERC).

Conflict of Interest: The authors confirm that there are no known conflicts of interest associated with this publication.

References

1. Robinson S. The development and application of an urban link travel time model using data derived from inductive loop detectors. Department of Civil and Environmental Engineering, Imperial College London. 2005.

2. Sassella A, Abbracciavento F, Formentin S, Bianchessi AG, Savaresi SM, editors. On queue length estimation in urban traffic intersections via inductive loops. 2023 American Control Conference (ACC); 2023: IEEE.

3. Wang J, Dixon KK, Li H, Ogle J. Normal deceleration behavior of passenger vehicles at stop sign–controlled intersections evaluated with in-vehicle Global Positioning System data. Transportation research record. 2005;1937(1):120-7.

4. Ulrich M, Gläser C, Timm F, editors. Deepreflecs: Deep learning for automotive object classification with radar reflections. 2021 IEEE Radar Conference (RadarConf21); 2021: IEEE.

5. Pu Z, Cui Z, Tang J, Wang S, Wang Y. Multimodal traffic speed monitoring: A real-time system based on passive Wi-Fi and Bluetooth sensing technology. IEEE Internet of Things Journal. 2021;9(14):12413-24.

6. Park S, Saeedi A, Kim DS, Porter JD. Measuring intersection performance from bluetooth-based data utilized for travel time data collection. Journal of Transportation Engineering. 2016;142(5):04016014.

7. Gong Y, Abdel-Aty M. Using machine learning to estimate pedestrian and bicyclist count of intersection by Bluetooth low energy. Journal of transportation engineering, Part A: Systems. 2022;148(1):04021101.

8. Messelodi S, Modena CM, Zanin M. A computer vision system for the detection and classification of vehicles at urban road intersections. Pattern analysis and applications. 2005;8:17-31.

9. Sivaraman S, Trivedi MM. Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis. IEEE transactions on intelligent transportation systems. 2013;14(4):1773-95.

10. Datondji SRE, Dupuis Y, Subirats P, Vasseur P. A survey of vision-based traffic monitoring of road intersections. IEEE transactions on intelligent transportation systems. 2016;17(10):2681-98.

11. Pourhomayoun M. Automatic traffic monitoring and management for pedestrian and cyclist safety using deep learning and artificial intelligence. 2020.

12. Yang D, Wu Y, Sun F, Chen J, Zhai D, Fu C. Freeway accident detection and classification based on the multi-vehicle trajectory data and deep learning model. Transportation research part C: emerging technologies. 2021;130:103303.

13. Adewopo VA, Elsayed N, ElSayed Z, Ozer M, Abdelgawad A, Bayoumi M. A review on action recognition for accident detection in smart city transportation systems. Journal of Electrical Systems and Information Technology. 2023;10(1):57.

14. Gunawan AA, Tanjung DA, Gunawan FE. Detection of vehicle position and speed using camera calibration and image projection methods. Procedia Computer Science. 2019;157:255-65.

15. Li Y, Duthon P, Colomb M, Ibanez-Guzman J. What happens for a ToF LiDAR in fog? IEEE Transactions on Intelligent Transportation Systems. 2020;22(11):6670-81.

16. Lan X, Wang C, Lv B, Li J, Zhang M, Zhang Z. 3D Point Cloud Stitching for Object Detection with Wide FoV Using Roadside LiDAR. Electronics. 2023;12(3):703.

17. Gao Y, Zhou C, Rong J, Wang Y. High-accurate vehicle trajectory extraction and denoising from roadside LIDAR sensors. Infrared Physics & Technology. 2023;134:104896.

18. Kamijo S, Matsushita Y, Ikeuchi K, Sakauchi M. Traffic monitoring and accident detection at intersections. IEEE transactions on Intelligent transportation systems. 2000;1(2):108-18.

19. Messelodi S, Modena CM, Zanin M. A computer vision system for the detection and classification of vehicles at urban road intersections. Pattern analysis and applications. 2005;8(1-2):17-31.

20. Muffert M, Milbich T, Pfeiffer D, Franke U, editors. May I enter the roundabout? A time-to-contact computation based on stereo-vision. 2012 IEEE Intelligent Vehicles Symposium; 2012: IEEE.

21. Barth A, Pfeiffer D, Franke U, editors. Vehicle tracking at urban intersections using dense stereo. 3rd Workshop on Behaviour Monitoring and Interpretation, BMI; 2009.

22. Wang W, Gee T, Price J, Qi H, editors. Real time multi-vehicle tracking and counting at intersections from a fisheye camera. 2015 IEEE Winter Conference on Applications of Computer Vision; 2015: IEEE.

23. Lee S-M, Baik H, editors. Origin-destination (od) trip table estimation using traffic movement counts from vehicle tracking system at intersection. IECON 2006-32nd Annual Conference on IEEE Industrial Electronics; 2006: IEEE.

24. Dassot M, Constant T, Fournier M. The use of terrestrial LiDAR technology in forest science: application fields, benefits and challenges. Annals of forest science. 2011;68:959-74.

25. McCoy MD, Asner GP, Graves MW. Airborne lidar survey of irrigated agricultural landscapes: an application of the slope contrast method. Journal of Archaeological Science. 2011;38(9):2141-54.

26. Shiravi S, Zhong M, Beykaei SA, editors. Accuracy assessment of building extraction using LIDAR data for urban planning transportation applications. Conference of the Transportation Association of Canada; 2012.

27. Weng S, Li J, Chen Y, Wang C, editors. Road traffic sign detection and classification from mobile LiDAR point clouds. 2nd ISPRS International Conference on Computer Vision in Remote Sensing (CVRS 2015); 2016: SPIE.

28. Asvadi A, Premebida C, Peixoto P, Nunes U. 3D Lidar-based static and moving obstacle detection in driving environments: An approach based on voxels and multi-region ground planes. Robotics and Autonomous Systems. 2016;83:299-311.

29. Elsisi M. Improved grey wolf optimizer based on opposition and quasi learning approaches for optimization: Case study autonomous vehicle including vision system. Artificial intelligence review. 2022;55(7):5597-620.

30. Tarko AP, Ariyur KB, Romero MA, Bandaru VK, Lizarazo CG. TScan: Stationary LiDAR for Traffic and Safety Studies—Object Detection and Tracking. 2016.

31. Zhao J, Xu H, Liu H, Wu J, Zheng Y, Wu D. Detection and tracking of pedestrians and vehicles using roadside LiDAR sensors. Transportation research part C: emerging technologies. 2019;100:68-87.

32. Velodyne. VLP-16c LiDAR (Puck) United States, 2018 [Available from: https://velodynelidar.com/products/puck/.

33. Velodyne. VLP-32c LiDAR (Puck Hi-Res) United States, 2018 [Available from: https://velodynelidar.com/products/puck-hi-res/.

34. Stauffer C, Grimson WEL, editors. Adaptive background mixture models for real-time tracking. Proceedings 1999 IEEE computer society conference on computer vision and pattern recognition (Cat No PR00149); 1999: IEEE.

35. Ester M, Kriegel H-P, Sander J, Xu X, editors. A density-based algorithm for discovering clusters in large spatial databases with noise. Kdd; 1996.

36. Golub GH, Reinsch C. Singular value decomposition and least squares solutions. Linear Algebra: Springer; 1971. p. 134-51.

37. Nateghinia E, Miranda-Moreno L. Development of a LiDAR-Based System Prototype for Detection, Classification, and Tracking of Road Users at Railway Facilities. Presented at Transportation Research Board 99th Annual Meeting; 2020.

38. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. nature. 1986;323(6088):533-6.

39. Cortes C, Vapnik V. Support-vector networks. Machine learning. 1995;20:273-97.

40. Cover T, Hart P. Nearest neighbor pattern classification. IEEE transactions on information theory. 1967;13(1):21-7.

41. Breiman L. Classification and regression trees: Routledge; 2017.

42. Breiman L. Random forests. Machine learning. 2001;45:5-32.

43. Hastie T, Tibshirani R, Friedman JH, Friedman JH. The elements of statistical learning: data mining, inference, and prediction: Springer; 2009.

44. Zadrozny B, Elkan C, editors. Transforming classifier scores into accurate multiclass probability estimates. Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining; 2002.

45. Friedman JH. Greedy function approximation: a gradient boosting machine. Annals of statistics. 2001:1189-232.

46. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of computer and system sciences. 1997;55(1):119-39.

47. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems. 2017;30.

48. Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H, et al. Xgboost: extreme gradient boosting. R package version 04-2. 2015;1(4):1-4.

49. Kalman RE. A new approach to linear filtering and prediction problems. 1960.

50. Munkres J. Algorithms for the assignment and transportation problems. Journal of the society for industrial and applied mathematics. 1957;5(1):32-8.

51. Neira J, Tardós JD. Data association in stochastic mapping using the joint compatibility test. IEEE Transactions on robotics and automation. 2001;17(6):890-7.

52. Rudenko A, Palmieri L, Herman M, Kitani KM, Gavrila DM, Arras KO. Human motion trajectory prediction: A survey. The International Journal of Robotics Research. 2020;39(8):895-935.

Link Between Chapter 2 and Chapter 3

Chapter 2 presented a novel methodology based on 3D LiDAR sensors for traffic monitoring at urban intersections. The LiDAR system was installed at various signalized and non-signalized intersections in Montreal, Canada. The proposed methodology's objective includes road user detection, clustering, classification, and tracking, each crucial for the comparative study of Chapter 3. The results of Chapter 2 form a comprehensive data set entered around the road users crossing an urban intersection. This data set included road users' trajectory in x-y coordinates, velocity in x-y coordinates, road users' class as pedestrian, cyclist, car, or truck, and most importantly, the 3D LiDAR point cloud representing each road user.

Chapter 3 examines surrogate safety analysis, an alternative to traditional crash-based methods. This chapter adopts a conflict-based approach for safety studies, defining a conflict as an interaction between two road users that occurs closely in time and space. Traditionally. The video-based trajectory of road users has been widely used to study these interactions. However, trajectory-based methods are subject to sensitivity and noise. Chapter 3 offers a novel method for studying these interactions using the point cloud of road users generated by the methodology presented in Chapter 2 when it is applied to urban intersections. The point cloud represents the shape of the road users and provides valuable insight into road user interactions, such as an accurate measure of their proximity in space, leading to a more precise measurement of their proximity in time. These measures are used to calculate surrogate safety indicators reliably, such as Time-to-Collision and Post-Encroachment Time.

CHAPTER 3.

A 3D-LIDAR-BASED METHODOLOGY FOR SURROGATE SAFETY ANALYSIS AT INTERSECTIONS WITH HIGH NON-MOTORIZED TRAFFIC

CHAPTER 3: A 3D-LIDAR-BASED METHODOLOGY FOR SURROGATE SAFETY ANALYSIS AT INTERSECTIONS WITH HIGH NON-MOTORIZED TRAFFIC

3.1 Abstract

Surrogate safety techniques utilize road users' trajectories to identify conflicts at urban intersections. Recent advancements in computer vision have enabled the large-scale collection of trajectory data. However, camera-based systems require calibration to convert two-dimensional pixel trajectories into x-y coordinates. Additionally, using trajectories would require defining a buffer size with an R radius to compensate for the physical shape of the road users. As an alternative, LiDAR-based systems are introduced to capture road user trajectories and shapes without calibration.

This paper leverages road users' shape data collected by LiDAR and corresponding trajectories for surrogate safety assessments. A shape-based method is developed to calculate Time-to-Collision (TTC) and Post-Encroachment Time (PET). The proposed method is then compared with traditional centroid-based methods, choosing six different buffer sizes from 1m to 4m.

The total numbers of post-encroachment time and time-to-collision conflicts between the two methods are compared. The comparison of the results with respect to the shape-based method highlights that each conflict type requires a custom value of R in the trajectory-based method. For example, pedestrians' and cyclists' interactions with vehicles require a buffer size of 2m and 2.5m, respectively, to mirror the results of the shape-based method, while for vehicle-vehicle conflicts, this extends to a range of 3m to 3.5m. However, predefined buffer sizes extend the road users' shape in both directions, capturing irrelevant interactions. Most importantly, the results show that centroid-based methods are susceptible to the size of buffer for critical conflicts with TTC or PET under 1.5 seconds, whereas for pedestrians and cyclists, choosing the proper buffer size becomes challenging for identifying critical conflict. On the other hand, the shape-based method only extends road users along their length and limits their width to their actual size, thereby eliminating potential false detections of conflicts. This method offers a precise arrival and departure time for each road user, which is essential to calculate surrogate safety metrics accurately.

***Keywords**:* LiDAR-based Surrogate Safety Analysis, Time-to-Collision, Post-Encroachment Time, Alternative Technology

3.2 Introduction

The interactions of road users at intersections can result in severe collisions and dangerous interactions involving vulnerable road users such as pedestrians and cyclists. Every year, thousands of people are seriously injured or killed in preventable crashes in urban and rural traffic facilities in Canada. According to the National Collision Database published by Transport Canada, from 2012 to 2021, in fatal or injury collisions happening in Canada, 1,461,144 individuals were injured, and 18,722 lost their lives. Of the total fatalities, 66.2% were vehicle occupants, 16.7% were pedestrians, 2.6% were cyclists, and 11.0% were motorcyclists. Furthermore, 26.6% of total fatalities happened at intersections, and more importantly, 39.2% of pedestrian and 46.25% of cyclists' fatalities are results of collisions at intersections (1).

Various countermeasures, such as traffic controls, geometry changes, and marking, are available to address safety issues and prevent injuries at urban intersections. These countermeasures include changes in traffic control (e.g., pedestrian signal installation, an all-red clearance interval, a left-turn phase), geometrics design modifications such as median and curb extensions, and installation of bicycle facilities (2). Traditionally, diagnosing safety issues and recommending appropriate countermeasures is based on historical crash data and crash-based methods. However, the use of collision data poses some challenges documented in the literature, including the underreporting issue of crash data, the need for long observations (a few years of crash data), crash data location accuracy, accuracy in reporting crash severity, and, more importantly, the reactive nature of crash-based methods, meaning a crash needs to happen before treating (3).

As an alternative approach, the application of surrogate safety methods has become more prevalent in recent years. This is partly thanks to the advancements in computer vision and machine learning, which enables the collection and processing of video data to investigate road user behaviors and conflict indicators for surrogate safety analysis. Computer vision techniques and tools for surrogate safety analysis using video data have allowed transportation safety engineers to implement proactive safety approaches that do not depend entirely on crash data.

Despite the advantages, a few limitations are associated with video data and automatic video data processing. First, the performance of the visual spectrum camera-based system is degraded in low-light conditions. Moreover, the camera-based system requires manual geometric calibration for each site's setup to ensure it correctly estimates the x-y coordinates of road users. More importantly, regular cameras do not directly provide the road users' 3D dimensions (shapes), which is critical for accurately computing surrogate safety indicators. Camera-based surrogate safety methods rely on utilizing road users' trajectories and defining a buffer around the centroid of road users to search for possible conflicts. However, determining the buffer size is a challenging task and can vary from one study to another, depending on the criteria defined for the surrogate safety indicator and the road users involved in the study. For example, if the buffer size is set to 2 meters to avoid extending beyond the width of vehicles, the effective length of cars would be reduced to 2 meters, which is unrealistic for surrogate safety analysis.

LiDAR technologies have shown advantages over traditional visual-spectrum camera-based systems as an alternative technological solution in the literature. One of the key benefits is their capability to monitor large areas effectively, which is crucial for comprehensive traffic analysis. LiDAR systems perform reliably under various lighting conditions, particularly in low-light environments such as nighttime. Another significant advantage of LiDAR is its inherent ability to measure distances accurately. This attribute enables LiDAR to precisely quantify space by reconstructing 3D environments as point clouds in x-y-z coordinates, providing a detailed and accurate representation of the traffic scenario.

Recent advancements in the development of medium-range rotational LiDAR have created new applications in Intelligent Transportation Systems (ITS) and Autonomous Vehicles (AV) (4). Currently, the application of LiDAR-based systems in surrogate safety analysis remains limited. Wu et al. implemented a low-resolution LiDAR system to identify near-miss pedestrian-vehicle interactions at an intersection, observing 258 vehicles and 36 pedestrians (5). In contrast, Tarko et al. utilized a super high-resolution LiDAR sensor for their surrogate safety analysis (6). However, both studies primarily employed the LiDAR-based x-y trajectories of road users as a substitute for video-based trajectories in their analysis.

To address some of these gaps, our study broadens the investigative scope by evaluating the impact of LiDAR-sensor resolution as well as the analysis of road environments with mixed traffic modes, facilitating the study of interactions among cars, pedestrians, and cyclists. Furthermore, this research proposes timestamped 3D point clouds to calculate two surrogate safety indicators: Time-

To-Collision (TTC) and Post-Encroachment Time (PET). The corresponding 3D point cloud of road users at each timestamp is converted to a minimum rotated rectangle to cover the point cloud fully in 2D space. Then, the length and shape of road users are estimated by the series of length and width. A shape-based approach is introduced to calculate surrogate safety indicators. The results are extensively compared with the trajectory-based approach.

3.3 Literature Review

Camera-based systems are among the most common and effective automatic traffic monitoring and safety analysis systems. Alternative methods and applications of video-based systems for conflict analysis and surrogate safety have been proposed in the literature. St-Aubin built an automated video-based traffic data collection to extract surrogate safety indicators such as time-to-collision and post-encroachment time (7). Zangabepur developed a fully automated video-based system for surrogate safety analysis, emphasizing the safety monitoring of cyclists (8). Fu et al. studied pedestrians' safety in interactions with vehicles at non-signalized crosswalks by analyzing vehicle yielding and pedestrian crossing decisions obtained from a video-based monitoring system (9) and secondary interactions with cars exiting the intersection (10). Zangenehpour et al. showed that intersections with cycle tracks appear safer than those without. They employed trajectories obtained from an automatic video-based monitoring system to investigate the interactions between cyclists and turning vehicles using PET measure (11). Seyed et al. used a computer vision algorithm to automatically detect vehicle-bicycle conflicts and rank them based on the severity of interactions using the Time-To-Collusion safety indicator (12).

There are several limitations associated with using video-based monitoring systems for traffic analysis. First, implementing computer vision algorithms in real-time demands powerful computing resources, which can increase the cost of such systems. Additionally, the effectiveness of image-based systems that rely on the visual spectrum often diminishes in low light conditions. A more critical issue is the heavy reliance of video data on geometric calibration, tailored to each camera's specific characteristics and the location of the study area. This calibration is vital for accurately converting trajectory data from pixel to x-y coordinates. Inaccurate calibration can lead to errors in determining the precise location of road users within a Cartesian coordinate system and in estimating their speeds, both of which are essential for effective surrogate safety analysis.

LiDAR has become widely used for scanning the environment in civil engineering. The applications of LiDAR include the modeling of 3D structures in urban facilities and building road inventory databases, including traffic signs and traffic light detection and recognition (13), street light pole detection (14), extracting road markings (15), pavement analysis, and crack and pothole detection (16). The 3D segmented models of the transportation facilities can be used for evaluating the safety factors at roads and intersections, including obstacle detection, accident detection or investigation, and detecting hazardous sections of the streets. (4).

There are a few research studies regarding the application of automatic LiDAR-based monitoring systems in traffic safety and surrogate safety analysis. Wu et al. implemented a low-resolution LiDAR system to identify near-miss pedestrian-vehicle interactions at an intersection, observing 258 vehicles and 36 pedestrians (5). In contrast, Tarko et al. utilized a super high-resolution LiDAR sensor for their surrogate safety analysis (6). They implemented an automatic system using 64-channel LiDAR for multi-object tracking at intersections and the extracted trajectories of road users. The extracted trajectories were given to computer software to compute surrogate safety indicators (6). These studies conceptualize using LiDAR to extract accurate trajectories for surrogate safety studies. Their analysis used LiDAR-based x-y trajectories of road users as a substitute for video-based trajectories. However, LiDAR technology has shown significant potential in estimating the shape of road users, which is of vital importance in the context of surrogate safety analysis of conflicts. This paper develops a method to utilize 3D point clouds from LiDAR data for calculating Time-to-Collision (TTC) and Post-Encroachment Time (PET) and compares the results with traditional trajectory-based methods.

3.4 LiDAR System Overview

This research proposes a 3D-LiDAR methodology for surrogate safety analysis built on the detection and classification methodology proposed by (17). Two LiDAR systems are tested, including a lower resolution LiDAR sensor with 16 laser channels (Velodyne Lidar: VLP-16) and a higher resolution LiDAR sensor with 32 laser channels (Velodyne Lidar: VLP-32c) (18, 19).

The proposed methodology using 3D LiDAR systems consists of two primary steps (**Figure 3-1**). In the first step, computational algorithms, data processing, and machine learning models are applied to the raw LiDAR data collected at intersections. This step yields road users' trajectories and corresponding clustered point clouds at each frame (timestamped at 0.1-second intervals). Additionally, essential information such as road users' speed and class are either estimated or predicted from this processed data.

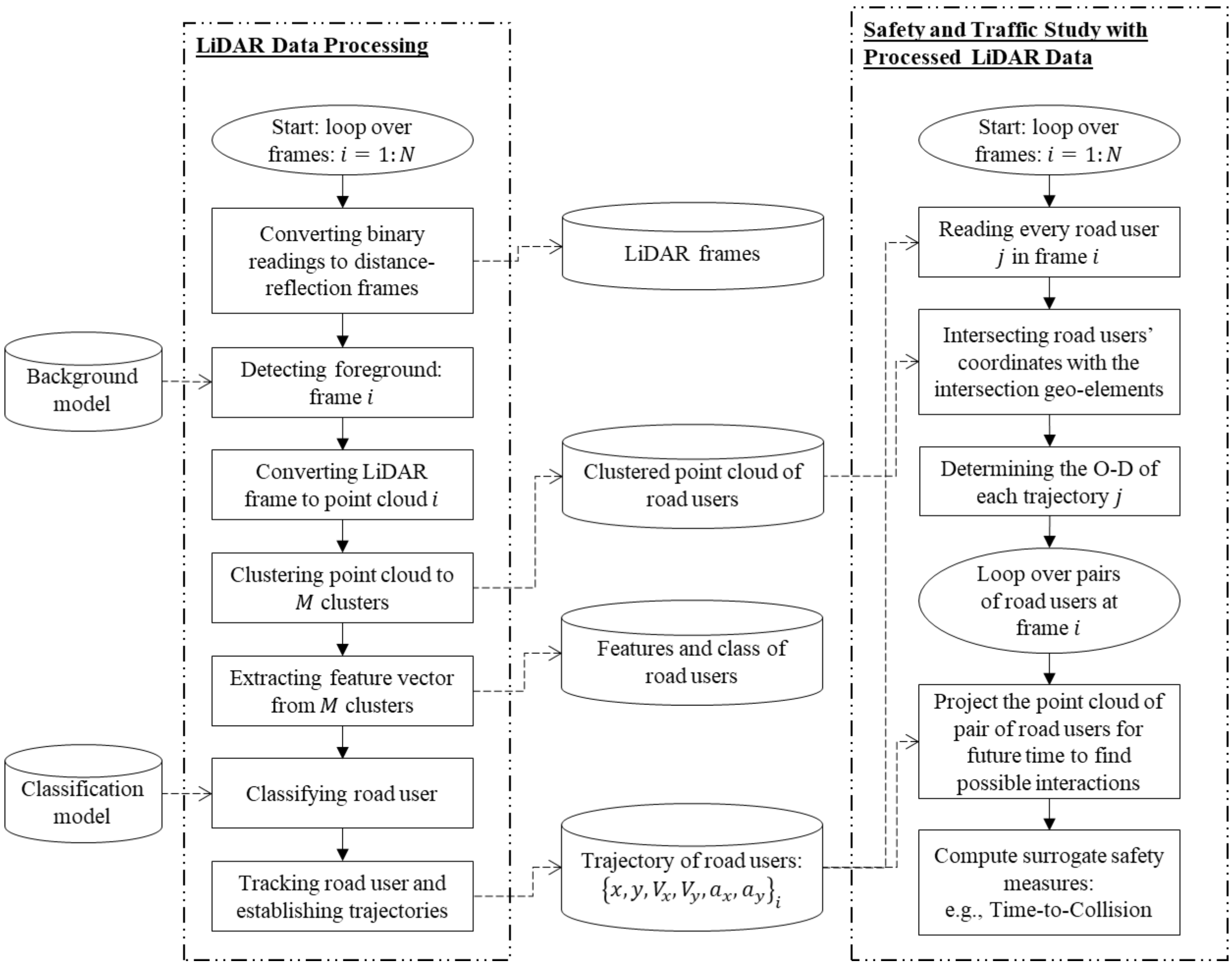


Figure 3-1 LiDAR system overview for traffic and safety monitoring at intersections

The second component of the methodology, the focal point of this research work, involves processing the road users' data, such as road user class, timestamped point clouds, and timestamped x-y trajectories, along with velocity values, for traffic and safety studies. This step incorporates safety analysis based on surrogate safety measures derived from the interactions of road users with each other in a point cloud space. A shape-based approach is proposed to compute surrogate safety indicators such as TTC and PET.

This section first provides an overview of the applied methodology for extracting road user trajectories and clustered point clouds (17). It subsequently discusses a preprocessing routine designed for cleaning and validating the trajectory data. Finally, the computation of essential surrogate safety indicators, including TTC and PET, is elaborated upon using LiDAR point cloud data in conjunction with trajectory data.

3.4.1 LiDAR data processing for road user extraction

The main component of the methodology for LiDAR data processing is illustrated in **Figure 3-1**. The left side box in the diagram features the method for processing raw data developed in the authors' previous work (17). In that framework, LiDAR binary data are first converted to LiDAR frames of distance values. Then, a 3D background model, based on a mixture of Gaussian, is built from a set of initial LiDAR frames. The background model is compared against every new LiDAR frame to identify foreground pixels. Next, the entire set of foreground pixels is converted to a three-dimensional point cloud, and a Density-based Spatial Clustering algorithm is applied to them. A feature set of physical, LiDAR, and spatial attributes is extracted from each clustered point cloud. XGBoost (eXtreme Gradient Boosting) is used to classify point clouds to pedestrians, cyclists, cars, or trucks. The LiDAR system utilizes data association and a Kalman Filter to track the point cloud's centroid of every road user and to establish the trajectory of the road users while they pass through the covered area at an intersection (17).

3.4.2 Trajectory preprocessing

This paper introduces a methodology for surrogate safety analysis that leverages the point cloud features of the LiDAR system rather than relying exclusively on the x-y coordinates of the trajectories. The raw output from the LiDAR system comprises timestamps, x-y coordinates of road users' centroid, velocity components in x-y coordinates, and a series of point clouds for each road user. Additionally, the LiDAR system classifies road users at each timestamp into categories such as pedestrians, cyclists, cars, or trucks. Point cloud data is crucial as it provides a more detailed and three-dimensional perspective of road user interactions, which is challenging to achieve with camera-based systems due to their limitations in accurately capturing depth and dynamic spatial relationships. For a robust and scalable implementation, each user's point cloud is transformed into a polygon encompassing the road user. This representative polygon is then utilized to assess interactions among road users. This conversion reduces processing time and opens up opportunities for the utilization of spatial data processing techniques.

Road users, typically cars, trucks, and cyclists, have a predominantly rectangular body shape. This assumption of a rectangular body shape can also be reasonably extended to pedestrians. To this end, the Minimum Area Bounding Rectangle (MABR) technique is utilized to fit a rotated rectangle to the point cloud of each road user (20). The LiDAR point cloud size for a road user varies between frames, influenced by their movement across different laser channels. For instance, if ten channels capture a user at a given timestamp, the point cloud will be larger compared to when only five channels observe the same user. As a result, a series of lengths and widths corresponding to the point clouds at various timestamps are extracted for each road user. The 85th percentile of these dimensions is used to define the user's length and width. This standardized measurement is applied consistently to the central point of each road user only in those frames where the original dimension significantly differs from the reconstructed one. This process ensures a more accurate and dynamic representation of road users' spatial presence and movement.

Subsequent data processing steps concentrate on refining road user trajectories and identifying outliers. Accurate and reliable trajectory data are crucial for surrogate safety analysis, as they provide the basis for calculating road users' speed and direction at each timestamp. These elements are then integrated into the polygon representing road users to forecast their future positions for safety analysis. If the trajectory data are of poor quality, it significantly undermines the analysis. Like other technologies, including camera-based systems and GPS-based data collection, trajectory data from the 3D LiDAR system exhibit noise and variations. The variations in LiDAR data trajectory arise from changes in the road users' point cloud shapes due to the resolution of laser channels and their blind spots. Centroids are located within the captured point cloud segment, but partial observations may introduce trajectory point deviations. The process involves steps that are selectively applied to correct anomalies in trajectories. For a given trajectory, the LiDAR system goes into prediction mode if it does not observe a road user from the previous. As a result, gaps may appear in the observed trajectory points. These gaps are filled by averaging the corresponding prediction coordinates with interpolations between adjacent observations in time.

Each trajectory is then aligned with a reference line created using a high confidence smoothing Kalman filter. This "high confidence" approach normalizes the observation covariance by its 2nd-degree norm to prevent over-smoothing and significant directional changes in the trajectories. Once the reference line is established, each trajectory point is projected onto the closest point on the reference line. This step ensures that the average displacement of a road user between consecutive frames is consistent with the trends in the original trajectory. Points significantly distant from the reference line are identified as outliers and replaced by a combination of predictions and interpolations based on the nearest observations in time.

After aligning trajectory points with their reference lines, some may exhibit backward movement, indicating a change in direction. A directional sliding window algorithm is implemented to identify these outliers (21, 22). The process begins by calculating the movement angle (θ_i) for each road user. Subsequently, within a sliding window of n consecutive frames (e.g., 10), the average angle ($\bar{\theta}$) for each local window is determined using the following formula:

$$\bar{\theta} = \text{arctan2}\left(\frac{1}{n}\sum_{i=1}^{n} \sin\theta_i , \frac{1}{n}\sum_{i=1}^{n} \cos\theta_i\right) \tag{3-1}$$

An angular distance between samples of the same sliding window to the average is defined as d_{θ_i}:

$$d_{\theta_i} = \left|\text{arctan2}\left(\sin(\bar{\theta} - \theta_i), \cos(\bar{\theta} - \theta_i)\right)\right| \tag{3-2}$$

A road user's direction is classified as an outlier if its angular distance exceeds the 50th percentile value of angular distances plus half the inter-percentile range, calculated as the difference between the 85th and 15th percentile. This outlier range is determined by the formula in Equation (3-3):

$$d_{\theta_i} > d_{50\%} + \frac{1}{2} \times (d_{85\%} - d_{15}) \tag{3-3}$$

An additional metric is implemented to complement the result of the first method. This approach converts users' angular directions to x-y coordinates on a unit circle. The Euclidean distances between each sample's direction and the samples' average direction vector are used to detect outliers similar to Equation (3-3). A sample is considered an outlier if concurrently identified as such by both methods. Finally, a smoother Kalman filter is applied to the corrected x-y coordinates of each trajectory, refining the velocity vector and direction of road users. This filter incorporates a normalized observation covariance matrix, scaled by a factor of 2, to enhance accuracy while closely mirroring the original movement pattern.

3.5 Surrogate Safety Measures based on LiDAR Trajectory Data

This section outlines the systematic approach employed in conducting surrogate safety analysis at selected intersections using LiDAR's trajectory and point cloud data of road users. An overview of the safety study framework utilizing processed LiDAR data is presented in the right-side box of the system diagram in **Figure 3-1**. The processing of LiDAR data for complex surrogate safety analysis requires an offline method that utilizes the completed trajectories of road users.

Following LiDAR data processing through various algorithms and models, distinct data sets are generated and stored. Except for the set of LiDAR frames (distance, reflection, and foreground mask), the surrogate safety component utilizes the remaining three datasets, including clustered point clouds of road users, road users' feature arrays (e.g., length, width), and arrays of road users' trajectories, observed and smoothed by the Kalman Filter. The trajectory arrays contain crucial information, including x-y coordinates representing the approximate centers of road users at each timestamp, two-dimensional velocity, and acceleration vectors.

Generally, LiDAR's trajectory data can be applied to any surrogate safety measures traditionally derived from video-based trajectory data. However, the scope of this research is to utilize road users' shape characteristics as an alternative method to compute two common surrogate safety indicators: Time-to-Collision and Post-Encroachment Time.

3.5.1 Time-to-Collision (TTC)

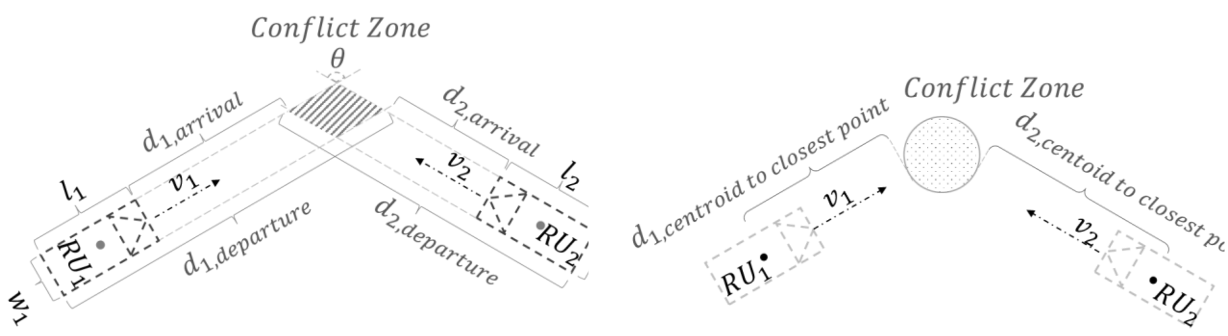
TTC serves as a critical surrogate safety indicator closely associated with actual accidents (23). TTC measures the time duration before a potential collision between two road users, considering their current motion patterns, speed, and acceleration remain unchanged. Extracting TTC data necessitates trajectory information, typically sourced from video-based systems. Developing methodologies based on alternative technologies, such as LiDAR, and enhancing methods based on camera systems, such as Unmanned Aerial Vehicle (UAV) video data, for surrogate safety assessment remains a primary focus in this field (24, 25). **Figure 3-2** depicts two interactions involving road user pairs, potentially representing TTC conflicts. In these images, RU_i denotes road user i, v_i represents their speed at the event time and l_i and w_i indicate the length and width of the respective user.

Figure 3-2 (a) shows two vehicles approaching a conflict zone at a non-direct angle (θ) to each other. In real-case scenarios, knowing the vehicle length and width allows for determining the boundary of the conflict area. Furthermore, considering their speeds, a few time measures can be calculated: the times of arrival, which is when the road user's front edge enters the closest side of the conflict zone, and the time of departure, noted as the moment when the back of the road user exists the farthest edge of the conflict area.

The existence of a TTC conflict can be determined by comparing the arrival and departure times of the two road users in the conflict zone. If the arrival time of one road user occurs between the arrival and departure times of the other, a TTC conflict exists. If the first road user arrives and remains in the zone until the second road user arrives, the TTC is calculated as ttc_{21} (Equation (3-4)). Conversely, if the second road user arrives and stays until the first road user arrives, the TTC is determined as ttc_{12} (Equation (3-5)).

$$if\ \frac{d_{1,arrival}}{v_1} < \frac{d_{2,arrival}}{v_2} < \frac{d_{1,departure}}{v_1};\ \text{then}\ ttc_{21} = \frac{d_{2,arrival}}{v_2} \tag{3-4}$$

$$if\ \frac{d_{2,arrival}}{v_2} < \frac{d_{1,arrival}}{v_1} < \frac{d_{2,depature}}{v_2};\ \text{then}\ ttc_{12} = \frac{d_{1,arrival}}{v_1} \tag{3-5}$$



a) Interaction between two road users with known shapes and dimensions

b) Interaction between two road users with known centroids in trajectory data

Figure 3-2 TTC conflict comparison based on data availability for position and dimensions

In practical applications, high-resolution trajectory data, typically obtained through cameras, are utilized to determine a potential TTC conflict. However, this trajectory data often lacks accurate information about the physical shape and dimensions of the road users**. Figure 3-2 (b)** presents a hypothetical scenario involving two road users. The only data points available for this scenario are the road users' centroids and velocities. Furthermore, the conflict zones are not delineated due to the absence of detailed spatial information, rendering the determination of precise arrival and departure times non-deterministic.

TTC calculation with centroid (no shape)

The derivation of the formula for the trajectory-based method is examined next. **Figure 3-3** depicts a potential conflict scenario between two road users within trajectory data.

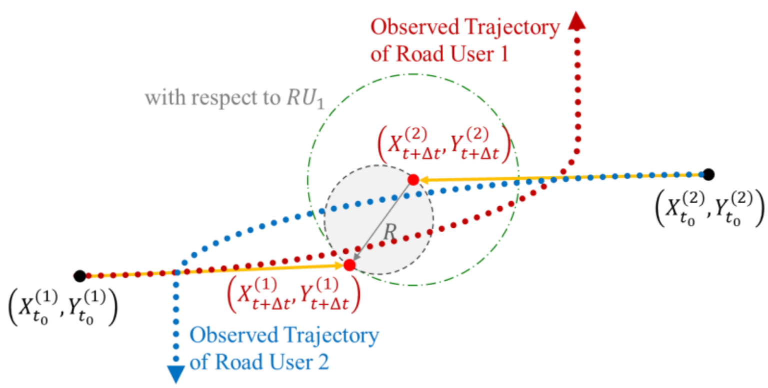


Figure 3-3 Illustration of a TTC conflict between two road users' trajectories

For the TTC calculation, first, it is assumed that the future positions of the road users are predicted using a constant velocity model for each frame. Acceleration data are not utilized due to their sensitivity; even minor noise or variations can significantly alter the predicted path of a road user. Furthermore, evasive maneuvers are immediately observed in the next frame's velocity vector since this process is applied to every frame. Hence, at time t, the position of the given road user i for any future timestamp Δt is estimated as Equation (3-6):

$$\begin{bmatrix} x^{(i)}_{t+\Delta t} \\ y^{(i)}_{t+\Delta t} \end{bmatrix} = \begin{bmatrix} x^{(i)}_{t} \\ y^{(i)}_{t} \end{bmatrix} + \begin{bmatrix} V^{(i)}_{x} \\ V^{(i)}_{y} \end{bmatrix} \times \Delta t \tag{3-6}$$

For two trajectories of different road users to exhibit a conflict, they must come closer than a specified threshold, denoted as R. A conflict is identified if, at a future time point $t + \Delta t$, the positions of both road users are within a conflict buffer having a diameter of R (depicted as the smaller circle in grey). The center of this conflict area is equidistant between the two road users, calculated as the average of their positions at time $t + \Delta t$. For a specified distance threshold (diameter), R, the positions of the road users at the boundary of the conflict area at the time t_2 (where $t_2 = t_1 + \Delta t$) is determined by Equation (3-7):

$$\left(X_{t_2}^{(1)} - X_{t_2}^{(2)}\right)^2 + \left(Y_{t_2}^{(1)} - Y_{t_2}^{(2)}\right)^2 = R^2 \tag{3-7}$$

Incorporating the predicted positions from Equation (3-6) into Equation (3-7), the only unknown variable in this quadratic Equation is $\Delta t = t_2 - t_1$:

$$\left(\left(V_x^{(1)} - V_x^{(2)}\right)\Delta t + \left(X_{t_1}^{(1)} - X_{t_1}^{(2)}\right)\right)^2 + \left(\left(V_y^{(1)} - V_y^{(2)}\right)\Delta t + \left(Y_{t_1}^{(1)} - Y_{t_1}^{(2)}\right)\right)^2 - R^2 = 0 \tag{3-8}$$

This Equation can be solved for Δt. If there is no solution, it indicates that the two road users do not have a TTC conflict within a distance of R. Conversely, a TTC conflict for the interaction is identified if the Equation yields one or two solutions. In such cases, the single solution or the lesser of the two solutions is reported as the actual TTC.

Selecting an appropriate value for R is a challenging aspect of this analysis. R is intended to represent the dimensions of road users, which vary significantly across different transportation modes. A further complication in determining R arises from the size of the road users as observed by the monitoring system, camera, or LiDAR. For instance, when the system captures a vehicle's side view, the centroid of this observation (and thus the trajectory) can accurately represent the center of the road user. In such cases, a larger value of R encompassing the length and width of a vehicle is more appropriate. However, when a vehicle is observed from any other angle, the center does not accurately represent the centroid of the road user. Utilizing a larger R in these scenarios might erroneously include bypassing road users, leading to inaccurate conflict detection.

TTC using road users' shapes in LiDAR point cloud

This research introduces an innovative approach as an alternative to the traditional method of extracting TTC data relying solely on the x-y coordinates of the centroids of road users in trajectory data. This proposed method utilizes road users' detected and reconstructed shapes to assess their proximity to others. Notably, this approach eliminates the need for distance thresholds such as R and related assumptions. This section then outlines the proposed steps to develop an efficient algorithm for extracting conflict data considering road user shapes.

As previously discussed in the data preprocessing section, the first step involves converting the clustered point cloud of each road user into x-y coordinates on a Cartesian plane by discarding the z-component from each point. A minimum rotated rectangle encapsulates each point cloud on the 2D Cartesian plane. Notably, the orientation of this rotated rectangle is adjusted from the algorithm's initial output to align with the direction of the road user's movement, as observed in the velocity data.

The geometric representation of each road user is encapsulated within a rectangle centrally aligned along the user's smoothed trajectory. The rectangle is defined by four sequentially ordered corners, starting with the front left ($p_{frontLeft}$), followed by the front right ($p_{frontRight}$), then the back right ($p_{backRight}$), and concluding with the back left ($p_{backLeft}$). This specific ordering of vertices is instrumental for projecting the road users into future frames at different timestamps and maintaining the algorithms' consistency. Each rectangle represents the road user's polygon. P_{RU}, and is constructed based on the central point p_{center}, from the smoothed trajectory, and the dimensions of the road user defined as length L and width W. The polygon is oriented using two directional vectors: $\vec{d}$ (indicating the direction of length), and $\vec{p}_{CCW}$, a counterclockwise perpendicular vector to $\vec{d}$. The mathematic representation of the road user's polygon is expressed in Equation (3-9):

$$P_{RU} = \begin{bmatrix} p_{frontRight} \\ p_{frontLeft} \\ p_{backRight} \\ p_{backLeft} \end{bmatrix} = \begin{bmatrix} (x_1, y_1) \\ (x_2, y_2) \\ (x_3, y_3) \\ (x_4, y_4) \end{bmatrix} = p_{center} + \frac{1}{2} \times \begin{bmatrix} L.\vec{d} + W.\vec{p}_{CCW} \\ L.\vec{d} - W.\vec{p}_{CCW} \\ -L.\vec{d} - W.\vec{p}_{CCW} \\ -L.\vec{d} + W.\vec{p}_{CCW} \end{bmatrix} \tag{3-9}$$

where $\vec{d}$, is defined as the normalized velocity vector, $\vec{d} = (d_x, d_y) = \frac{1}{\|\vec{v}\|}(v_x, v_y)$, and $\vec{p}_{CCW}$ clockwise perpendicular vector to the direction, as $\vec{p}_{CCW} = (d_y, -d_x)$.

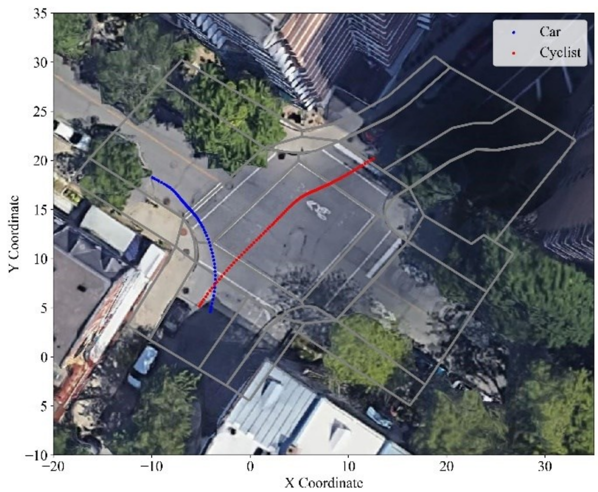
Identifying potential TTC conflicts involves projecting the polygons of all road users in a given frame into the future, employing a constant velocity model similar to that described in Equation (3-6). A TTC conflict is detected when polygons representing two road users intersect within any future space-time window. An extended path is defined for each road user in the frame to enhance the algorithm's efficiency and avoid extensive distance calculations between each pair of road user polygons. This extended path maintains the position of the rear points of the road users while projecting the front points forward by a time increment, δt. This strategy focuses on potential overlaps of road user paths in the future projection, thus efficiently identifying TTC conflicts.

In determining road user interactions, it is crucial to establish the value of δt to create a tangible path. The safety analysis based on TTC primarily focuses on interactions within a 10-second timeframe, paying particular attention to those within 1.5 to 5 seconds due to their critical nature. Therefore, δt is set at 10 seconds. For each road user, a consistent polygonal path is established in every frame and updated continuously with each new frame based on observed changes in position and velocity data (p_{center} and $\vec{v}$). The construction of this polygon path is denoted as $\hat{P}_{RU}$:

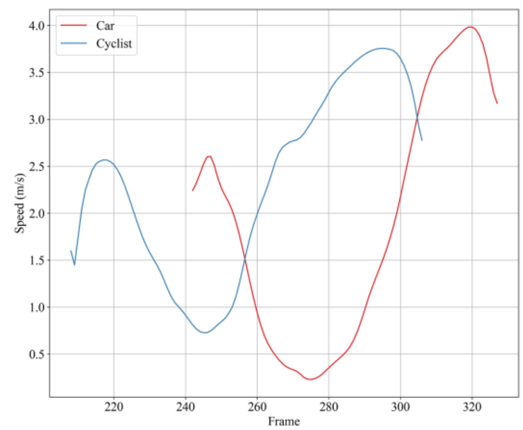
$$\hat{P}_{RU} = \begin{bmatrix} p_{frontRight} \\ p_{frontLeft} \\ p_{backRight} \\ p_{backLeft} \end{bmatrix} = \begin{bmatrix} (x_1, y_1) \\ (x_2, y_2) \\ (x_3, y_3) \\ (x_4, y_4) \end{bmatrix} = p_{center} + \frac{1}{2} \times \begin{bmatrix} L.\vec{d} + W.\vec{p}_{CCW} \\ L.\vec{d} - W.\vec{p}_{CCW} \\ -L.\vec{d} - W.\vec{p}_{CCW} \\ -L.\vec{d} + W.\vec{p}_{CCW} \end{bmatrix} + \begin{bmatrix} \delta t.(v_x, v_y) \\ \delta t.(v_x, v_y) \\ 0 \\ 0 \end{bmatrix} \quad (3\text{-}10)$$

Figure 3-4 illustrates a real TTC conflict between a car traveling Southbound and making a right turn and a Westbound cyclist. **Figure 3-4 (a)** displays the trajectories of these road users during their interaction. Additional insights into this interaction are provided in **Figure 3-4 (b)**, where the speed profiles of both users at the same frames are compared. Notably, the passenger car decelerates to yield the right of way to the cyclist, mitigating the risk of a potentially severe conflict.

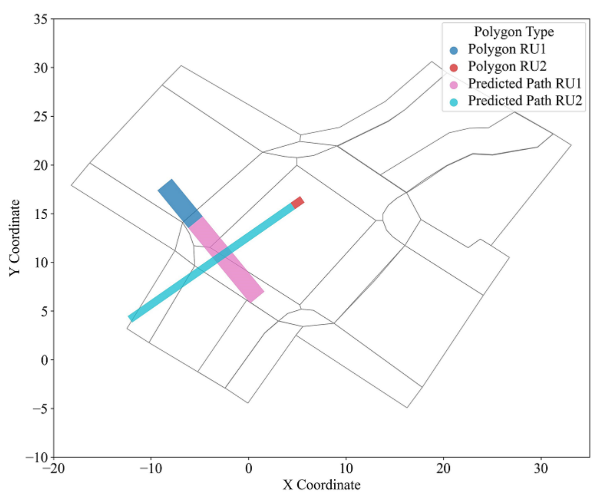
Figure 3-4 (c) illustrates the bounding polygons (P_{RU}) and the predicted polygon paths ($\hat{P}_{RU}$) of both road users at the frame where the minimum TTC is recorded. The smaller polygons represent the bounding rectangles of the road users at that specific frame, whereas the extended polygons depict their projected paths over the next 10 seconds. This visualization indicates that a collision will occur if both users maintain their current movement patterns and speeds.



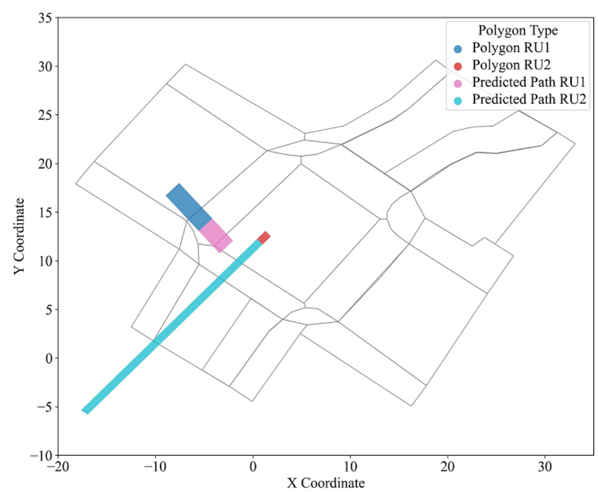
a- trajectory of two road users in a conflict



b- smoothed speed profile of the road users



c- polygon and path polygon of the road users at the frame with minimum TTC



d- polygon and path polygon of the road users at 2 seconds after minimum TTC

Figure 3-4 A sample of TTC conflict between a car and a cyclist

Conversely, **Figure 3-4 (d)**, captured two seconds after the minimum TTC event, shows the car decelerating to yield to the cyclist. This behavior is evidenced by the altered polygon path of the passenger car (highlighted in pink). The intersection (or lack thereof) of these polygon paths in **Figure 3-4 (c & d)** indicates potential TTC conflicts. Following this, the arrival and departure times of both users to and from the conflict area are calculated to ascertain the existence of an actual TTC conflict, as outlined in Equations (3-4) and (3-5). The movement is assumed to be with constant velocity; therefore, the arrival and departure times are calculated based on the closest and farthest distance of the road user's original polygon at the current from the conflict area.

3.5.2 Post-Encroachment Time (PET)

PET is a critical surrogate safety indicator that heavily relies on spatial data and the distance between road users. Defined as the temporal gap between two road users traversing an intersection, PET quantifies the time between one road user leaving a designated area and another entering it. As depicted in **Figure 3-2 (a & b)**, the situation is identified as a TTC conflict if two road users reach the conflict area simultaneously. However, the likelihood of such a conflict is relatively low due to evasive maneuvers typically employed by road users. Conversely, if road users arrive at different times, PET can be effectively calculated.

A primary challenge in computing PET lies in defining the conflict area, which is the zone visited by two road users at different times. Similar to TTC, trajectory data lack specifics regarding the shapes of the road users, necessitating an approach that establishes a distance threshold or buffer area, denoted as R, for PET. The first step in identifying a PET conflict involves calculating the minimum Euclidean distance between each pair of data points from two selected road users. The road users are chosen if they overlap in time. Similar to TTC, PET values of 10 seconds or less are the main focus of safety studies. Therefore, for a specific road user, any other users observed from 10 seconds before to 10 seconds after the last observation of that user is selected. The minimum distance of data point between two users is defined as:

$$MinDistance = \min_{P_i \in RU_i, P_j \in RU_2} \text{Euclidean}(P_i, P_j) \tag{3-11}$$

If this minimum distance is less than or equal to the threshold R ($MinDistance \leq R$), a PET conflict is considered to exist. Subsequently, each pair of data points with a distance less than R is selected, and the time difference between these points is computed. PET is defined as the minimum of these time differences among the chosen pairs of points, subject to the condition that their Euclidean distance is within R:

$$PET = \min_{P_i \in RU_i, P_j \in RU_2} |t_{P_i} - t_{P_j}|; \quad where \text{ Eucildean}(P_i, P_j) \leq R \tag{3-12}$$

PET using road users' shapes in LiDAR point cloud

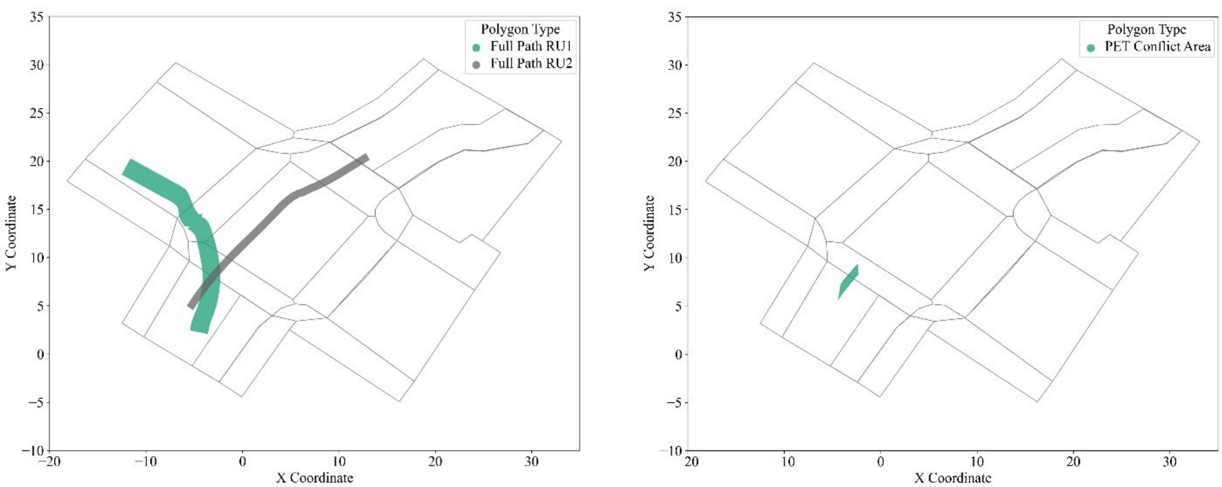
The shape data of road users, estimated by the LiDAR system, can be effectively utilized to identify and compute PET. For this purpose, a unique convex polygon path is constructed for each road user, encompassing all rectangular polygons representing the user at various timestamps. This path extends from the initial polygon at the user's first frame to the last observed polygon.

Each polygon is a rectangle determined by the coordinates of its corners. The order of points is carefully selected from individual polygons of different timestamps to ensure the integrity and convexity of the road user's polygon path. The construction typically starts with the back-left corner at the first timestamp ($p_{t_1-backLeft}$), followed by the front-left corners of consecutive frames ($p_{t_1-frontLeft}$ to $p_{t_n-frontLeft}$). After reaching the last left-side corner, the sequence continues with the front-right corners of consecutive timestamps ($p_{t_n-frontRight}$ to $p_{t_1-frontRight}$), concluding at the back-right corner of the first timestamp ($p_{t_1-backRight}$). Notably, only the back corners from the first timestamp are used, with the remaining points being the front-center points of each rectangle at different timestamps.

This methodical construction of a convex polygon path accurately represents the road user's shape and movement trajectory over time, which is essential for precisely determining PET in various traffic scenarios. The mathematical construction of a road user's polygon path ($P_{Path-RU}$) is:

$$P_{Path-RU} = \begin{bmatrix} p_{t_1-backLeft} \\ p_{t_1-frontLeft} \\ \vdots \\ p_{t_n-frontLeft} \\ p_{t_n-frontRight} \\ \vdots \\ p_{t_1-frontRight} \\ p_{t_1-backRight} \end{bmatrix} \tag{3-13}$$

Upon the completion of each road user's full polygon path, a comparison is made between pairs of road users within a 10-second interval before and after their trajectories overlap. Initially, the polygon of the current road user is overlaid with those of the selected road users. If a shared area exists, a PET conflict is identified and becomes a candidate for calculation. **Figure 3-5** illustrates two road users (same users in **Figure 3-4**) involved in a PET conflict. **Figure 3-5 (a)** displays both users' constructed convex polygon paths, with the cyclist's path being notably narrower, as expected. **Figure 3-5 (b)** demonstrates the result of overlaying the road users' polygons, indicating a small intersecting area fully traversed by both users.



a- full polygon path of two road users

b- PET conflict area

Figure 3-5 A sample of PET conflict between a car and a cyclist

Once this conflict polygon is identified, the closest and farthest distances from any point on the road users' polygons (four corner points) to this area are calculated to determine each road user's arrival and departure time. PET is defined as follows:

$$PET = t_{arrival,RU2} - t_{departure,RU1} \tag{3-14}$$

In this context, RU_1 is strictly defined as the road user who arrives first at the conflict area.

This section discusses an overview of calculating two surrogate safety indicators, TTC and PET, using trajectory data and integrating road users' shape data. The primary distinction lies in the reliance on methods based on trajectory data on a minimum distance threshold. This threshold is crucial for identifying conflicts and calculating the corresponding time metrics associated with such road users' conflicts. The subsequent section will provide an overview of the LiDAR data and the intersection employed in this study. The following section will set the stage for a comparative analysis of these two distinct approaches in identifying and calculating TTC and PET.

3.6 LiDAR Data

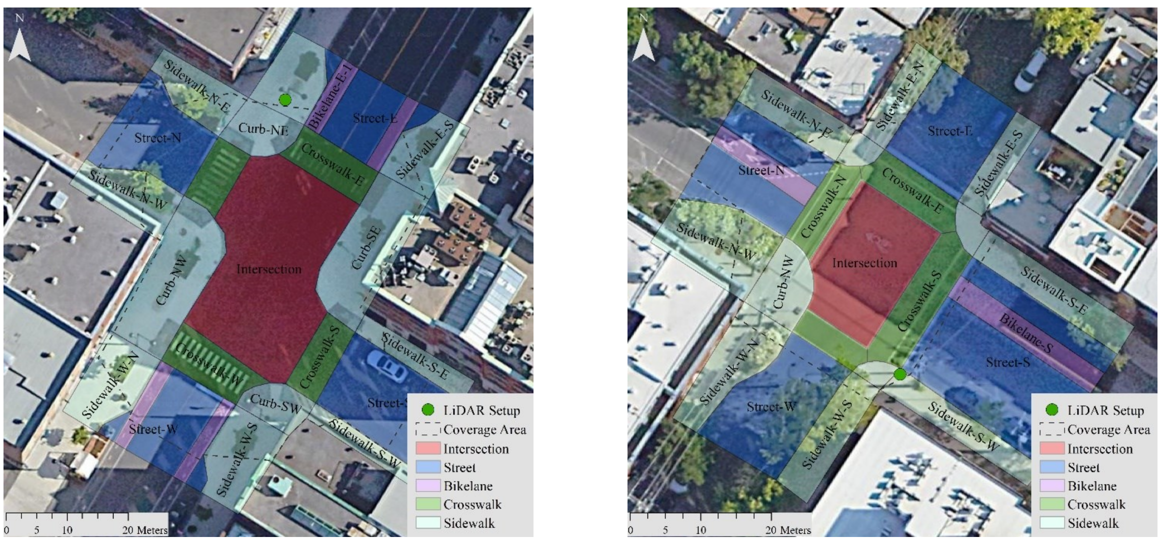
A comprehensive dataset encompassing road users' trajectories, point clouds, and classification data has been collected from 10 intersections. For each intersection, two hours of data are utilized to perform a comparative analysis of the proposed methodology, analyzing both TTC and PET based on road users' shape data instead of solely using trajectory data. **Table 3-1** provides an overview of the data collected using the LiDAR system at these ten intersections. These intersections were chosen for their diversity in terms of vehicular traffic volume and mixed-mode traffic. All selected locations are situated in Montreal, Canada.

Table 3-1 Summary of road user data collection by the LiDAR systems at different intersections

Site ID	Site Name	LiDAR Channel	Traffic Control	Intersection and Crosswalk Area (m2)	Vehicular Traffic			Cyclist	Pedestrian
					Left Turn	Right Turn	Through		
1	SainteFamille - Milton	16	All-way Stop	277.6	47	195	463	212	416
2	Delaroche - MarieAnneE	16	Traffic Light	267.9	14	36	967	388	471
3	Colonial - RachelE	16	Traffic Light	245.7	87	146	769	436	461
4	Girouard - Monkland	16	Traffic Light	411.8	332	185	1,041	142	**626**
5	University - Milton	16	Traffic Light	295.9	77	91	793	**594**	**1,374**
6	Hutchison - LaurierE	32	All-way Stop	470.9	206	202	652	169	**651**
7	SainteFamille - PrinceArthurW	32	All-way Stop	354.5	48	132	363	96	525
8	Parthenais - Rachele	32	Traffic Light	283.7	89	161	396	344	173
9	SainteFamille - Milton	32	All-way Stop	278.3	48	106	339	166	338
10	University - Milton	32	Traffic Light	231.3	18	77	685	**593**	**625**
			Total		966	1,331	6,468	3,140	5,660

The data was captured using two distinct types of LiDAR sensors: one with a lower resolution of 16 laser channels and the other with a higher resolution featuring 32 channels. Intersections 1 and 9 are at the same locations but were monitored using different LiDAR sensors. This is true for intersections 5 and 10 as well.

As an integral part of this work to enhance the surrogate safety study, the boundaries of various elements within each intersection were meticulously mapped in a Geographic Information System (GIS). These elements encompass intersections, streets, bike lanes, crosswalks, and sidewalks. **Figure 3-6** showcases some samples of these intersections, illustrating the detailed construction of these geo-elements. Each road user's trajectory is intersected with the mapped elements, and an origin-destination label is consequently assigned to each trajectory. This process is instrumental in monitoring specific movements and facilitates reporting based on the type of turning movement. As indicated in **Table 3-1**, vehicular traffic data are categorized and reported according to turning movements.



a) ID 6 - Hutchison-Laurier E (32-Channel) b) ID 4 – Girouard-Monkland (16-Channel)

Figure 3-6 GIS calibration of two samples of intersection

3.7 Comparative Analysis: Centroid-based vs Shape-based Method

This section delves into the comparative analysis of TTC and PET derived from centroid- and shape-based methods. In the shape-based method, road users are represented as polygons, while the centroid-based method represents them as points along their trajectory data. The analysis focuses primarily on the key differing variable between the two models: the distance threshold (R) in the centroid-based approach. For this purpose, point-based metrics are computed using six different R values: 1m, 2m, 2.5m, 3m, 3.5m, and 4m, selected based on a combination of the length and width of various road users. Other tested values, such as 0.5m and 5m, were deemed unsuitable for the analysis and thus excluded to enhance the clarity of the comparison.

3.7.1 Structure of surrogate safety data

In the dataset prepared for safety analysis, conflicts that both users are vulnerable road users, pedestrians and cyclists, are excluded. Thus, the focus is on conflicts involving cars or trucks with other road users, forming seven distinct categories. The database categorizes conflicts as involving two parties: Road User 1 (always a car or truck) and Road User 2 (which may be a car, truck, pedestrian, or cyclist).

In identifying TTC conflicts, interactions between two road users are analyzed over consecutive frames. A TTC value is assigned to each frame where a potential conflict is detected. The definitive TTC value for any given interaction is then determined by taking the 15th percentile of the TTC series for the involved pair of users. Typically, TTC conflicts are calculated using directional speed, making them sensitive to abrupt changes in direction. Therefore, conflicts observed in only one or two frames are excluded from consideration to ensure accuracy.

The conflicts are categorized based on the conflict type derived from conflict angles. The interaction is considered a rear-end conflict for conflict angles less than 15°. Angles between 15° and 75° indicate angular, and those between 75° and 90° are classified as side-impacts. This categorization applies to every other angle except those between 165° and 180°, which are categorized as head-on conflicts.

To ensure a harmonized structure when combining TTC and PET data from different intersections, only conflicts occurring within the intersection, on crosswalks, or on streets within 1 meter of the crosswalks are reported. This approach was chosen because some streets were not fully covered at some intersections in the data collection process. A preliminary analysis revealed that 55% of conflicts were rear-end types occurring near intersections due to acceleration-deceleration maneuvers. Thus, focusing on the aforementioned selected areas helps achieve a balanced comparison.

Finally, road users' interactions are categorized into three groups based on their TTC or PET values, a common practice in road safety. Interactions with a duration of less than 10 seconds are reported as the baseline for safety analysis. Additionally, interactions with PET or TTC that are less than 5 seconds and 1.5 seconds are categorized as conflicts and serious conflicts, respectively.

3.7.2 TTC

Table 3-2 summarizes the safety state of the ten selected intersections in a 2-hour time window. For readability purposes, only the shape-based results are reported. However, the same measures are also calculated for the centroid method. The direct comparison of the methods follows this table. Interactions are categorized by the types of road users involved. TTC events are extracted using a LiDAR system-based approach that focuses on the shape of road users, with cars and trucks collectively considered as 'vehicles.' In every instance, the first road user is identified as a vehicle.

Table 3-2 Summary of TTC conflicts using shape-based approach (first road user vehicle)

ID	# TTC < 10 sec	**Second Road User**								
		Pedestrian			**Cyclist**			**Vehicle**		
		# TTC < 5s	# TTC < 1.5s	15th TTC	# TTC < 5s	# TTC < 1.5s	15th TTC	# TTC < 5s	# TTC < 1.5s	15th TTC
1	177	41	6	4.46	26	8	4.08	46	10	3.57
2	75	44	16	3.17	9	6	2.64	9	4	2.45
3	123	31	8	3.68	31	14	2.79	45	**28**	1.53
4	269	75	14	3.99	13	4	3.74	122	**61**	2.60
5	197	70	19	3.94	46	9	3.59	32	15	2.94
6	555	151	18	4.56	39	5	3.64	173	**32**	4.09
7	120	39	2	4.65	5	0	3.80	25	2	4.88
8	30	6	0	4.75	9	2	4.43	6	4	2.04
9	67	23	5	4.17	6	0	3.99	16	4	4.62
10	109	24	7	3.89	55	16	2.84	6	2	3.92
All	**1,722**	**504**	**95**	**4.13**	**239**	**64**	**3.55**	**480**	**162**	**3.26**

A review of this data compared to traffic volumes from **Table 3-1** indicates a correlation between traffic flow and conflict frequency. Notably, data for the latter five intersections was gathered in 2021, a period marked by pandemic-related changes in travel patterns. The first five intersections (IDs 1-5) are collected using a 16-channel LiDAR susceptible to slight overcount and misclassification of cyclists. The last five intersections (IDs 6-10) are collected using a 32-channel LiDAR sensor.

The average 15th percentile TTC values for interactions between vehicles and pedestrians, cyclists, or other vehicles are calculated as 4.13 seconds, 3.55 seconds, and 3.26 seconds, respectively.

On average, intersections 3, 4, 5, and 6 exhibit a higher frequency of critical events during the two-hour analysis period for each intersection. **Table 3-3** delves into the specifics of these critical conflicts, including the volume of the second user involved in each conflict. Notably, at intersections 3 and 4, most vehicular conflicts are head-on, attributable to the limited entry width of their respective legs when performing a left turn. The heightened instances of critical conflicts involving pedestrians and cyclists can be linked to the observed high volume of road users.

Table 3-3 Distribution of critical TTC conflicts at selected intersections with higher TTC rates

Intersection ID	2nd Road User	Traffic Volume	Total TTC<1.5s	Rear-End	Angular	Head-On	Side-Impact
3	Cyclist	**436**	**14**	6	5	1	2
	Pedestrian	461	8	2	2	1	3
	Vehicle	**1,002**	**28**	4	**10**	**12**	2
4	Cyclist	142	4	1	2	1	0
	Pedestrian	**626**	**14**	4	4	0	6
	Vehicle	**1,558**	**61**	5	**27**	**21**	8
5	Cyclist	594	9	4	1	1	3
	Pedestrian	**1,374**	**19**	1	12	1	5
	Vehicle	961	15	5	3	1	6
6	Cyclist	169	5	1	2	0	2
	Pedestrian	**651**	**18**	3	5	0	10
	Vehicle	**1,060**	**32**	8	18	3	3

Table 3-4 thoroughly compares two methods used for calculating TTC and identifying TTC conflicts, focusing on the pairs of involved road users. This is due to the varying distances between the centroids of their trajectories in different pairings. The methods are compared in terms of the total number of TTC interactions, conflicts with TTC lower than five seconds, conflicts with TTC lower than 1.5 seconds, and the average of the 15th percentiles of TTC, which are influenced by the size of the buffers. An exact match does not necessarily represent the best performance user-

wise; however, surrogate safety indicators are often used at an aggregated level in before-after studies. Therefore, the total number of conflicts and the average of TTC or PET are of importance.

Table 3-4 A comparison between the shape-based approach vs centroid-based approach

TTC Metric	Road User 1	Road User 2	Shape-Based Approach	Centroid Approach – Buffer Size:					
				1m	2m	2.5m	3m	3.5m	4m
Number of Interactions with TTC<10 sec	Car	Pedestrian	**708**	441	**645**	**820**	1,097	1,392	1,791
		Cyclist	**290**	148	215	**273**	387	496	604
		Car	**561**	235	309	392	**525**	766	964
		Truck	**78**	28	40	50	**66**	**93**	122
	Truck	Pedestrian	*58*	27	39	*47*	*72*	94	112
		Cyclist	*15*	5	9	*10*	*15*	23	28
		Truck	12	4	5	6	6	7	9
	All		1,722	888	1,262	1,598	2,168	2,871	3,630
Number of Conflicts with TTC< 5 sec	Car	Pedestrian	**465**	233	**404**	**580**	845	1,141	1,514
		Cyclist	**228**	93	160	**223**	338	455	569
		Car	**418**	120	175	253	**393**	625	833
		Truck	**55**	15	22	31	**44**	**75**	105
	Truck	Pedestrian	39	10	21	29	43	61	86
		Cyclist	11	4	8	9	13	21	25
		Truck	7	3	4	4	5	6	8
	All		**1,223**	478	794	**1,129**	**1,681**	2,384	3,140
Number of Conflicts with TTC< 1.5 sec	Car	Pedestrian	**85**	11	**47**	**146**	312	540	826
		Cyclist	**62**	10	**38**	**81**	171	268	352
		Car	**139**	7	27	74	**184**	396	584
		Truck	19	0	0	6	15	36	60
	Truck	Pedestrian	10	2	2	4	12	17	32
		Cyclist	2	0	0	2	5	11	16
		Truck	4	0	0	0	0	0	3
	All		321	30	114	313	699	1,268	1,873
Average of TTC (15th percentile)	Car	Pedestrian	**4.22**	5.07	**4.36**	**3.86**	3.31	2.87	2.56
		Cyclist	**3.42**	4.35	**3.76**	**3.12**	2.43	2.04	1.82
		Car	**3.48**	5.22	4.61	4.03	**3.18**	2.51	2.12
		Truck	**3.51**	5.49	4.93	4.34	**3.84**	2.82	2.34
	Truck	Pedestrian	4.01	5.43	4.72	4.56	4.39	4.05	3.36
		Cyclist	3.72	3.98	3.56	2.92	2.75	2.28	1.95
		Truck	4.20	5.20	4.73	4.86	3.65	3.38	2.72
	All		**3.79**	4.96	4.38	**3.96**	**3.37**	2.85	2.41

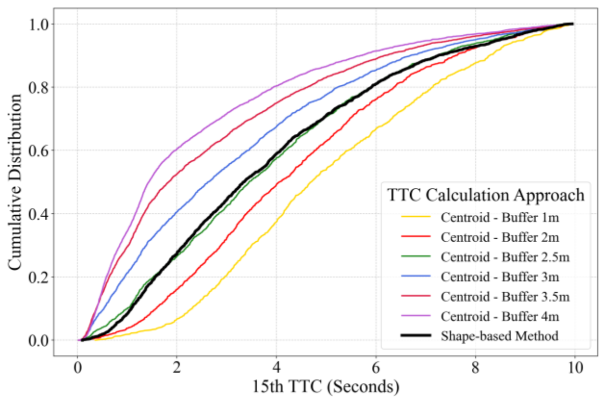
For car-pedestrian conflicts, the total number of TTC events and conflicts in the shape-based approach closely aligns with the centroid-based methods using 2-meter and 2.5-meter buffers. Moreover, the average TTC from the shape-based method matches the average TTC when using a 2m buffer. However, in the case of critical conflicts, TTC < 1.5 seconds, using a 2-meter buffer underreports, and a 2.5-meter buffer overreports the number of conflicts, highlighting the sensitivity of centroid-based approaches to identifying critical conflicts between cars and pedestrians.

In the case of car-cyclist conflicts, the number of interactions and conflicts with a TTC under 5 seconds, and the average 15th percentile TTC is more closely aligned with a buffer size of 2.5 meters. However, the number of serious conflicts does not fully align with either a 2.5-meter or a 2-meter buffer. Utilizing a 2.5-meter buffer reduces the average TTC for this pair of users, resulting in more critical conflicts.

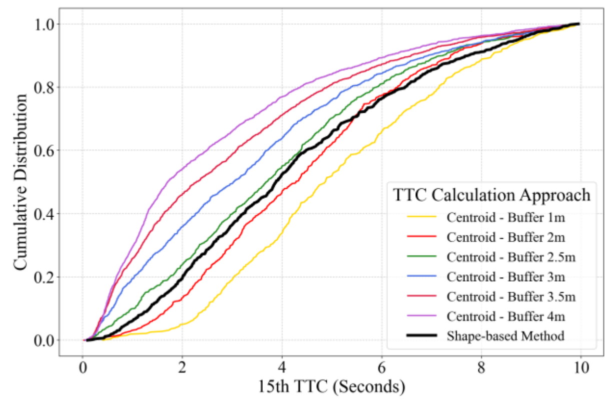
For car-car conflicts, centroid-based methods with a 3-meter buffer provide results that are more similar to shape-based methods. However, using this buffer size reduces the average TTC, indicating a shorter distance to reach a conflict point at a constant speed and, consequently, a higher number of critical TTC conflicts (<1.5 seconds). Generally, as the buffer size in the centroid-based approaches increases, TTC significantly decreases, leading to the detection of more critical conflicts.

The 2-meter buffer significantly underestimates the critical conflicts in truck-pedestrian and truck-cyclist interactions. Instead, the results of 2.5-meter and 3-meter buffers are more similar to the shaped-based method.

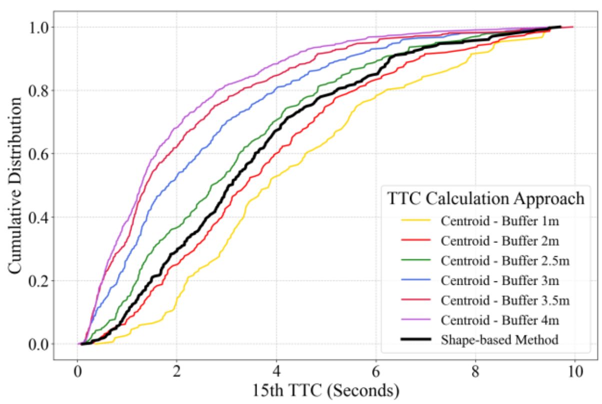
Figure 3-7 illustrates the cumulative distribution of the 15th percentile TTC, comparing the approaches used to identify TTC interactions. The shape-based method, which is based on estimated length and width, is highlighted in black. In contrast, each other color represents one of the buffer sizes. **Figure 3-7 (a)** shows that the centroid method with a 2.5-meter buffer aligns closely with the shape-based method’s combined cumulative distribution for all road users. For pedestrians and cyclists, the images support the earlier discussion of choosing a 2 to 2.5-meter buffer. However, the gap around TTC=1.5 seconds between the shape-based, the 2-meter buffer, and the 2.5-meter buffer is noticeable, resulting in differences in the number of critical conflicts.



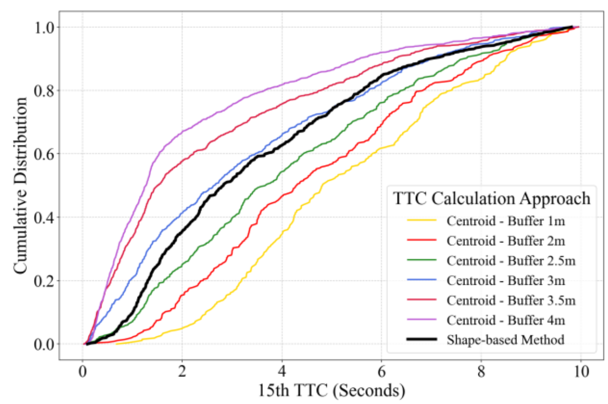
a) Every TTC interaction



b) Pedestrian-Vehicle



c) Cyclist-Vehicle

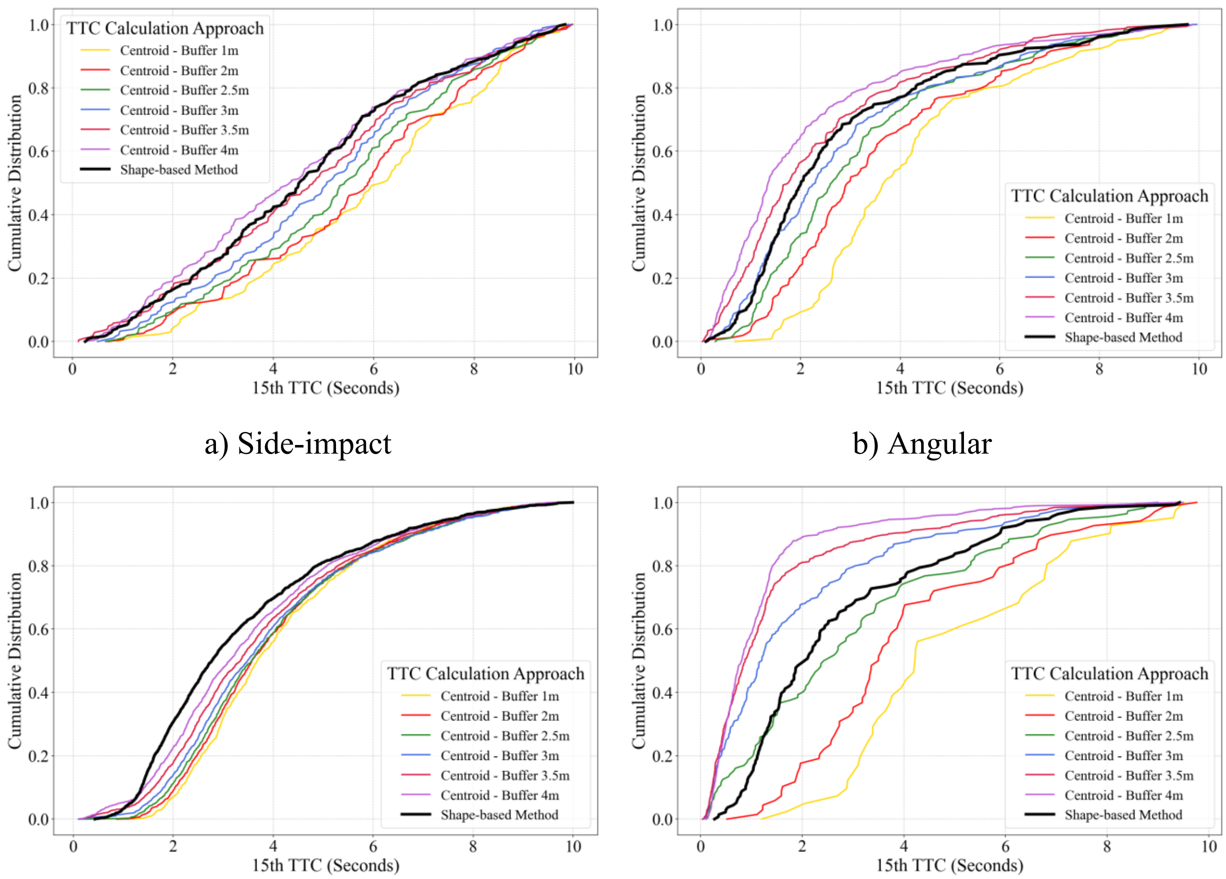


d) Vehicle-Vehicle

Figure 3-7 A comparison of different approaches for extracting TTC conflicts and interactions

Vehicles observed at intersections predominantly involve angular and side-impact conflicts. A more accurate comparison can be achieved by revisiting the entire pool of TTC events, including those with conflict centers outside the intersection's boundary. **Figure 3-8** compares methods for vehicle-vehicle conflicts based on conflict angles. The shape-based model (as depicted in **Figure 3-8 (a)**) corresponds closely with a 3.5-meter buffer for side-impact conflicts. Practically, the minimum distance between the centers of two vehicles in a side-impact conflict is half the length of one vehicle plus the width of another, typically between 3 to 3.5 meters. In angular conflicts, the center points of the vehicles can be closer than in side impacts, implying a conflict radius smaller than three meters, as supported by **Figure 3-8 (b)**. As shown in **Figure 3-8**, a buffer exceeding four meters aligns with the average vehicle length for rear-end conflicts.

However, comparing head-on conflicts presents challenges, as larger buffers can inadvertently encompass passing vehicles, given their extension in the perpendicular direction. Considering the average vehicle width is under two meters, utilizing a 3-meter buffer results in a conflict zone extending approximately two meters beyond each side of the vehicle's body. Therefore, the biggest buffer size based on the shape-based method is in the range of two meters.



a) Side-impact b) Angular

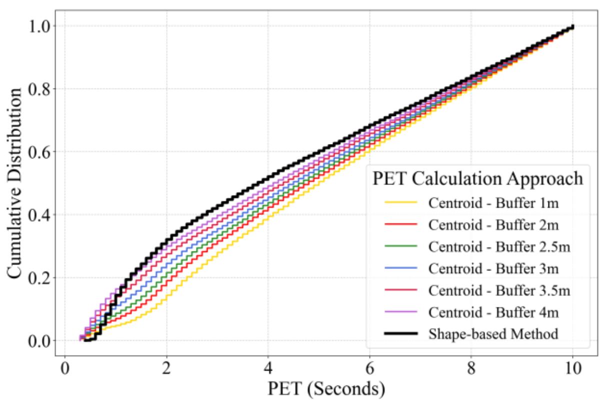
c) Rear-end d) Head-on

Figure 3-8 A comparison of results for various vehicle-vehicle conflict types

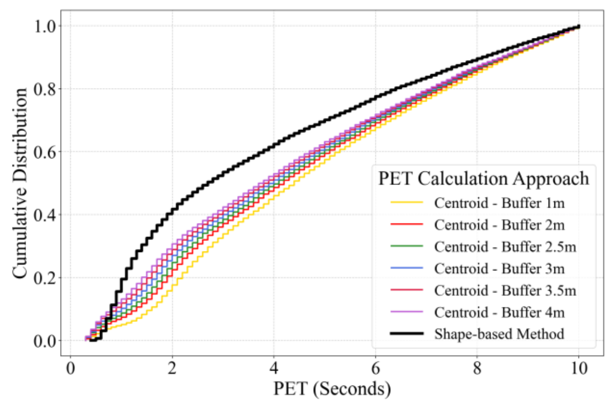
3.7.3 PET

This section underscores the main distinctions between the centroid and polygon methods in identifying and calculating PET. The correlation between buffer size and the combination of road users' body dimensions in different conflict angles is anticipated to mirror those observed in TTC interactions. However, the PET comparison offers specific advantages. First, PET involves a spatial intersection of road users at different times, which is more probable than TTC. Additionally, the frequency of PET events tends to be higher than that of TTC events and is generally less prone to noise, enabling a more robust comparison.

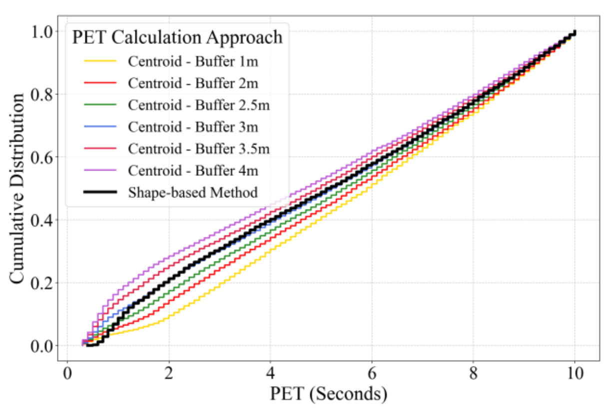
Figure 3-9 (a) illustrates the cumulative PET across different methods. The graph for the proposed shape-based method is noticeably higher than others, indicating a lower average PET than centroid-based options. The PET metric is predominantly skewed toward rear-end events. In the shape-based approach, considering vehicle lengths generally exceeding 4 meters, particularly with trucks and buses, the following vehicle's distance to the leading vehicle's location is often less than the conflict area defined in centroid-based methods. A shorter travel distance translates to a quicker arrival time and a lower PET, a pattern observable in **Figure 3-9 (b)**, which exclusively plots vehicle-vehicle rear-end PET interactions.



a) Cumulative PET – all events with PET less than 10 seconds



b) Cumulative PET – rear-end conflict of vehicles



c) Cumulative PET – all events except rear-end vehicle interactions

Figure 3-9 Comparative analysis of PET calculation approaches based on different criteria

Figure 3-9 (c) presents the outcome of excluding vehicular rear-end interactions from the PET events pool and limiting the conflict area to the zones discussed for TTC (intersection, crosswalk, and street areas within 1 meter of the crosswalk). These adjustments result in findings that align more closely with the average TTC trends.

Table 3-5 presents an in-depth comparison of PET calculations, considering the type of the second road user. The first user is either mostly a car or sometimes a truck. This comparison includes conflicts where PET is less than 1.5 seconds. Applying a 2-meter distance threshold aligns closely with the shape-based method regarding the number of conflicts and the average PET for vehicle-pedestrian conflicts. The optimal distance threshold for vehicle-cyclist conflicts is between 2.5 and 3 meters. Notably, cyclists are often involved in rear-end conflicts with vehicles, necessitating a larger buffer size. Of the 706 cyclist conflicts determined by the shape-based method, 408 are rear-end conflicts with a car.

Table 3-5 Comparative summary of PET calculation methods per road user class

PET Metric	Second Road User	Shape-Based Approach	Centroid-based Approach – Buffer Size:					
			1m	2m	2.5m	3m	3.5m	4m
Number of Conflicts PET<5 sec	Pedestrian	**798**	638	**835**	963	1,179	1,409	1,632
	Cyclist	**706**	437	528	**632**	**713**	815	930
	Car	**1,148**	545	782	**974**	**1,304**	1,643	1,904
	Truck	**276**	80	124	152	207	**287**	344
Number of Conflicts PET<1.5 sec	Pedestrian	**133**	59	**164**	266	476	725	982
	Cyclist	**415**	215	291	**380**	**453**	552	657
	Car	**329**	47	115	192	**340**	564	771
	Truck	**86**	5	16	18	39	**76**	108
Mean PET (sec)	Pedestrian	**5.4**	5.8	**5.4**	5	4.6	4.2	3.8
	Cyclist	**3.9**	4.6	4.3	**4**	**3.8**	3.6	3.3
	Car	**5.1**	5.8	5.6	5.4	**5.1**	4.8	4.7
	Truck	**5.1**	6.3	5.9	5.8	5.5	**5.2**	4.9
Median PET (sec)	Pedestrian	**5.4**	5.9	**5.4**	5	4.5	3.8	3
	Cyclist	**2.9**	4.6	4.2	3.6	**3.2**	2.4	1.7
	Car	**5.1**	6	5.7	5.4	**5.1**	4.7	4.4
	Truck	**5**	6.3	6	5.8	5.5	**5.1**	4.8

The relevant distance threshold spans from 2.5m to 3m in car-vehicle interactions. Several factors influence this range: firstly, rear-end conflicts are omitted, effectively reducing the distance between centroids. Secondly, the distance between centroids can be reduced to zero for PET analysis, as the users are not simultaneously present in the conflict area. However, considering critical PET and mean PET values, the findings align more closely with a 3-meter threshold.

3.8 Conclusion and Future Work

This research proposes a shape-based methodology to compute TTC and PET surrogate safety measures and implements a comparative analysis between the proposed approach and the traditional centroid-based method. The centroid-based approach, involving the determination of an appropriate distance threshold or buffer size (R), poses challenges for robust surrogate safety assessments. The effectiveness of this method relies on accurately defining R. In this approach, the length and width of road users are treated uniformly, leading to situations where bigger values of R might inadvertently capture interactions with bypassing road users rather than direct conflicts. In extreme cases, using large values for distance thresholds may result in the identification of conflicts between vehicles and pedestrians walking on the sidewalk.

In contrast, the proposed rectangular shape-based approach leverages the dimensional attributes of road users –length and width – as measured by a 3D LiDAR system. This system provides the trajectory of the road users and generates a clustered point cloud for each one. The point cloud data is utilized to estimate the length and width of road users. The 85th percentile of the length and width measurements from a series of observed dimensions is used to estimate the shape. The shape-based method circumvents the need for calibrating buffer size R, inherently providing a more precise description of road user interactions. However, the effectiveness of this approach is contingent upon the performance of the LiDAR system, particularly its accuracy in estimating the shapes of road users across consecutive frames.

TTC and PET were analyzed to compare the performance of the two methods. For the trajectory approach, six different runs are completed with various values for R, including 1, 2, 2.5, 3, 3.5, and 4m. The comparison is limited to intersections, crosswalks, and portions of streets within a one-meter distance from crosswalks. This decision helps capture road users' interactions and movement as much as possible without introducing bias by excluding the high rear-end vehicular movements.

Regarding TTC, four metrics are used for comparison: interactions with TTC under 10 seconds, conflicts with TTC under 5 seconds, critical conflicts with TTC under 1.5 seconds, and the average of the 15th percentile TTC. In pedestrian-car interactions, results using a 2-meter distance threshold align more with the shape-based method. For cyclist-car interactions, the threshold extends to 2.5 meters. In car-car interactions, a 3-meter threshold aligns better match with the shape-based method. However, the centroid-based approach exhibits higher sensitivity for identifying critical conflicts. For pedestrian-car interactions, using a 2-meter buffer results in 47 conflicts, and using a 2.5-meter buffer results in 146 conflicts, while the shape-based method estimated 85 conflicts. For cyclist-car interactions, using a 2.5-meter buffer, 81 critical conflicts are identified, whereas the shape-based method detects 62 conflicts. Generally, an increase in buffer size decreases the TTC, affecting those conflicts with TTC slightly higher than 1.5 seconds.

Analyzing PET, all rear-end vehicular conflicts are excluded to ensure a balanced representation of road users' interactions. The methods are evaluated based on conflicts with PET under 5 seconds, critical conflicts with PET under 1.5 seconds, and the mean and median of PET. For pedestrian-vehicle conflicts, a centroid method with a 2-meter threshold suggested by TTC identifies 164, while the shape-based method identifies only 133 critical conflicts. In cyclist-vehicle conflicts, the results of the shape-based methods fall within a 2.5 to 3-meter threshold. However, a 3-meter buffer identifies more than 19% of critical conflicts.

A fundamental insight in trajectory-based methods is the impact of increasing radius on conflict detection. Firstly, a larger radius tends to identify unnecessary TTC or PET conflicts with users on divergent paths. Secondly, a larger radius implies users need to travel a shorter distance to reach conflict points, systematically reducing the value of TTC. For example, a 0.5-meter difference in buffer size for a pedestrian walking with an average speed of 1.5 m/s is translated to 0.34 seconds; therefore, using a buffer with a 0.5-metter larger radius reassigns conflicts from one category to another, increasing the sensitivity of the surrogate safety assessment.

On the other hand, a shape-based method extends only one dimension (length) while limiting the other dimension (width). This characteristic naturally leads to more accuracy, as it mirrors the actual road users' interactions and avoids categorizing every proximity as a conflict. For example, under this method, a vehicle moving through an intersection while a pedestrian crosses a parallel crosswalk or a cyclist parallelly passing a vehicle would not be classified as conflicts, even though such scenarios are fairly common at intersections.

Future research will involve a more comprehensive comparison, expanding the data collection in duration and location to various intersections and roadways. Applying these methodologies in varied traffic conditions and diverse urban settings could validate their effectiveness across different environments.

Acknowledgments: Funding for this project is provided in part by the Natural Sciences and Engineering Research Council (NSERC).

Conflict of Interest: The authors confirm that there are no known conflicts of interest associated with this publication.

References

1. Transport Canada. National Collision Database Online (2002-2021) 2022 [Available from: https://wwwapps2.tc.gc.ca/Saf-Sec-Sur/7/NCDB-BNDC/p.aspx?l=en.

2. Federal Highway Administration - Office of Safety. Toolbox of Countermeasures and Their Potential Effectiveness for Intersection Crashes. In: Transportation Do, editor. 2009.

3. Miranda-Moreno LF, Morency P, El-Geneidy AM. The link between built environment, pedestrian activity and pedestrian–vehicle collision occurrence at signalized intersections. Accident Analysis & Prevention. 2011;43(5):1624-34.

4. Williams K, Olsen M, Roe G, Glennie C. Synthesis of transportation applications of mobile LiDAR. Remote Sensing. 2013;5(9):4652-92.

5. Wu J, Xu H, Zheng Y, Tian Z. A novel method of vehicle-pedestrian near-crash identification with roadside LiDAR data. Accident Analysis & Prevention. 2018;121:238-49.

6. Tarko AP, Ariyur KB, Romero M, Bandaru VK, Jimenez CL. Guaranteed LiDAR-aided multi-object tracking at road intersections: USDOT Region V Regional University Transportation Center final report. 2016.

7. St-Aubin P. Driver behaviour and road safety analysis using computer vision and applications in roundabout safety: Ecole Polytechnique, Montreal (Canada); 2016.

8. Zangenehpour S, Miranda-Moreno LF, Saunier N. Automated classification based on video data at intersections with heavy pedestrian and bicycle traffic: Methodology and application. Transportation research part C: emerging technologies. 2015;56:161-76.

9. Fu T, Miranda-Moreno L, Saunier N. A novel framework to evaluate pedestrian safety at non-signalized locations. Accident Analysis & Prevention. 2018;111:23-33.

10. Fu T, Hu W, Miranda-Moreno L, Saunier N. Investigating secondary pedestrian-vehicle interactions at non-signalized intersections using vision-based trajectory data. Transportation research part C: emerging technologies. 2019;105:222-40.

11. Zangenehpour S, Strauss J, Miranda-Moreno LF, Saunier N. Are signalized intersections with cycle tracks safer? A case–control study based on automated surrogate safety analysis using video data. Accident Analysis & Prevention. 2016;86:161-72.

12. Sayed T, Zaki MH, Autey J. Automated safety diagnosis of vehicle–bicycle interactions using computer vision analysis. Safety Science. 2013;59:163-72.

13. Tan M, Wang B, Wu Z, Wang J, Pan G. Weakly supervised metric learning for traffic sign recognition in a LIDAR-equipped vehicle. IEEE Transactions on Intelligent Transportation Systems. 2016;17(5):1415-27.

14. Yu Y, Li J, Guan H, Wang C, Yu J. Semiautomated Extraction of Street Light Poles From Mobile LiDAR Point-Clouds. IEEE Transactions on Geoscience and Remote Sensing. 2015;53(3):1374-86.

15. Hsieh H-H, Hsu C-Y, Ke P-Y, Liu G-S, Lin C-P, editors. Appling Lidar-based obstacle detection and wireless image transmission system for improving safety at level crossings. Security Technology (ICCST), 2015 International Carnahan Conference on; 2015: IEEE.

16. Haiyan G, Li J, Yongtao Y, Chapman M, Cheng W. Automated Road Information Extraction From Mobile Laser Scanning Data. IEEE Transactions on Intelligent Transportation Systems. 2015;16(1):194-205.

17. Nateghinia E, Miranda-Moreno L. A 3D-LiDAR-based Methodology for Extracting Data for Surrogate Safety Analysis at Intersections. Presented at Transportation Research Board 101st Annual Meeting; 2022.

18. Velodyne. VLP-16c LiDAR (Puck) United States, 2018 [Available from: https://velodynelidar.com/products/puck/.

19. Velodyne. VLP-32c LiDAR (Puck Hi-Res) United States, 2018 [Available from: https://velodynelidar.com/products/puck-hi-res/.

20. Chaudhuri D, Samal A. A simple method for fitting of bounding rectangle to closed regions. Pattern recognition. 2007;40(7):1981-9.

21. Mardia KV, Jupp PE, Mardia K. Directional statistics: Wiley Online Library; 2000.

22. Fisher NI. Statistical analysis of circular data: cambridge university press; 1995.

23. Sacchi E, Sayed T. Conflict-based safety performance functions for predicting traffic collisions by type. Transportation Research Record. 2016;2583(1):50-5.

24. Bhattarai N, Zhang Y, Liu H, Xu H. Crash frequency prediction based on extreme value theory using roadside lidar-based vehicle trajectory data. Accident Analysis & Prevention. 2023;193:107306.

25. Yang D, Ozbay K, Xie K, Yang H, Zuo F, Sha D. Proactive safety monitoring: A functional approach to detect safety-related anomalies using unmanned aerial vehicle video data. Transportation research part C: emerging technologies. 2021;127:103130.

Link Between Chapters 2 and 3 and Chapter 4

Chapter 2 discussed the development of a supervised 3D LiDAR methodology for traffic monitoring at urban intersections. In this Chapter, a set of labeled road users was sampled from various intersections to train a machine-learning algorithm for road user classification. Chapter 3 utilized the road user dataset created in Chapter 2 for surrogate safety analysis at urban intersections using LiDAR-based shape data of road users.

Chapter 4 develops an unsupervised 3D LiDAR methodology for safely monitoring at a railroad-grade crossing. The methodology includes road user detection, clustering, and tracking. Road user classification is made based on their shape, speed, and geo-location. The speed and shape of the users are estimated from their trajectory and point cloud, respectively.

The methodologies in chapters 2 and 4 differ in background modeling, road user detection techniques, and road user classification. The unsupervised method utilizes a low-resolution LiDAR. Chapter 2 extends the methodology to a low-resolution and a higher-resolution LiDAR system installed at various urban intersections.

The unsupervised learning methodology implemented in Chapter 4 provided insight into developing a semi-automated road user labeling utilized in Chapter 2.

CHAPTER 4.

DEVELOPMENT OF AN UNSUPERVISED 3D LIDAR-BASED METHODOLOGY FOR AUTOMATED SAFETY MONITORING OF RAILWAY FACILITIES

CHAPTER 4: DEVELOPMENT OF AN UNSUPERVISED 3D LIDAR-BASED METHODOLOGY FOR AUTOMATED SAFETY MONITORING OF RAILWAY FACILITIES

4.1 Abstract

Railway safety (e.g., at grade crossings, platforms, or rail tracks) is a primary concern for transportation authorities. Unfortunately, preventable railway collisions claim the lives of hundreds annually, often involving individuals crossing illegally at highway-railway grade crossings or trespassing at unauthorized railroad facilities. Transportation authorities often deploy a range of engineering countermeasures to mitigate the frequency or risk of such events. These countermeasures include technological solutions that automatically activate warning systems, barriers, or gates to alert and deter road users from unlawfully entering restricted railway facilities. For the safety monitoring of such facilities, alternative sensing technologies such as video-based computer-vision systems have been evaluated and, in some cases, utilized in practice. Despite their merits, implementing automated LiDAR-based detection and tracking methods has yet to be explored in railway safety applications. This research aims to introduce and assess an unsupervised 3D-LiDAR-based methodology for monitoring rail-road level facilities. This study's core is the implementation of an unsupervised learning algorithm designed to detect, track, and classify road users using point clouds gathered by a 3D-LiDAR sensor. The proposed methodology demonstrates encouraging results when monitoring rail-road level crossings. The aggregate average absolute percentage deviation (AAPD) for motorized road users and counting motorized road users stands at 5% and 3%, for non-motorized road users at 10% and 14% on two separate test days, each featuring distinct system installations.

***Keywords**:* 3D LiDAR Sensor, Level Crossing Monitoring, Trespassing Detection, Unsupervised LiDAR Algorithm, Alternative Technologies

4.2 Introduction

Safety at railway facilities is paramount for transportation authorities and the railway industry. Interactions between road users and trains at railway facilities, such as road-rail level crossings, railway tracks, and terminals, can lead to severe collisions, especially in scenarios involving conflicts between pedestrians and trains. Each year, preventable train accidents contribute to hundreds of deaths and thousands of serious injuries. Train-pedestrian collisions are generally classified into two main categories based on pedestrian intent: level crossing and trespasser accidents. Pedestrians crossing railway tracks at designated shared areas, such as grade crossings, are categorized as level crossing users.

Conversely, trespassers cross the railway right of way and access unauthorized and restricted railway areas. From 2013 to 2022, trespassing was the predominant cause of railway-related fatalities, constituting more than 59% and 64% in Canada and the US, respectively. According to the Transportation Safety Board (TSB) of Canada, from 2013 to 2022, 1,581 railway-road level crossing collisions resulted in 200 deaths and 256 serious injuries. Additionally, 654 trespasser incidents resulted in 409 deaths and 190 serious injuries (1). During the same period, the Federal Railroad Administration reported 21,385 crossing collisions in the US, resulting in 2,507 fatalities, 8,480 injuries, and 9,635 trespasser incidents, with 5,062 deaths and 4,909 injuries (2).

Several countermeasures are often implemented across railway facilities to address safety issues and prevent or mitigate railway-related injuries. These countermeasures often fall into three categories: educational, enforcement, and engineering interventions. Educational and enforcement interventions aim to increase the responsibility and awareness of people residing or working near railway facilities (3). Engineering countermeasures are divided into two forms: geometry-related interventions and technological solutions. Geometric engineering interventions involve redesigning or modifying railway facilities to improve pedestrian safety by deterring unauthorized access (4). Technological solutions are pivotal in monitoring rail facilities by detecting, warning, and restricting individuals' access through signals, sounds, and gates. In a standard configuration, a detection system verifies unauthorized road users' presence, activates warning devices, and sends feedback to authorities (5). Such system incorporates alternative technologies such as cameras, thermal cameras, stereo cameras, ultrasonic sensors, active or passive infrared sensors, and radars (6-9).

Despite the distinct advantages of alternative systems, certain limitations must be highlighted. For instance, the visual spectrum camera-based system is adversely affected by low-light conditions. Moreover, deploying a camera-based system requires manual geometric calibration at each site to accurately estimate road users' x-y coordinates. While thermal and stereo cameras may address these issues, their implementation comes with elevated costs, and the real-time operation of these systems demands powerful processing units. Infrared, ultrasonic, and radar sensors, on the other hand, face challenges in accurately classifying road users and are significantly influenced by adverse weather conditions and precipitation (2).

Recently, the literature has underscored the essential advantages of LiDAR technologies over traditional monitoring technologies, particularly camera-based systems. These benefits include the capability to monitor large areas and deliver accurate performance irrespective of lighting conditions (10). Recent advancements in the development of medium-range rotational LiDAR sensors have opened up new applications, especially in the domains of Intelligent Transportation Systems (ITS) and Autonomous Vehicles (AV) (11). However, there are still few traffic monitoring and safety applications and no documented applications in railway facilities. Large coverage areas of interest in rail facilities often challenge traditional camera-based monitoring systems. The extensive field of view of 3D LiDAR sensors, especially rotational LiDARs, makes them well-suited for these applications.

However, the availability of open-source LiDAR datasets suitable for developing supervised algorithms remains limited. Existing datasets, such as the KITTI dataset (12), use high-resolution LiDAR sensors, which are costly applications related to road facility monitoring. Besides, implementing a supervised learning algorithm for a point cloud of high-resolution LiDAR requires complex features, and extracting those features in real-time applications is computationally intensive (13). In this context, non-supervised methodologies present an alternative approach for processing LiDAR point clouds. The unsupervised methods rely on some key physical features such as size to perform classification and, therefore, do not require sophisticated feature extraction methods that are often time-intensive.

This work develops and evaluates an unsupervised 3D LiDAR-based methodology for monitoring non-motorized and motorized road users at railroad-level facilities, specifically grade crossings. The proposed method utilizes a rotational 3D LiDAR sensor with a 360-degree horizontal field of view to collect data at a railroad-level crossing. An unsupervised learning algorithm is developed to detect, classify, and track road users within the 3D LiDAR point cloud. Notably, the unsupervised methodology eliminates the need for labeled data, thereby reducing the need for sample data generation and labeling.

The ultimate objective of a safety monitoring system at a railway level crossing is to be integrated with the level crossing's control mechanism and provide real-time feedback to the control room. However, developing and validating the methodology before this integration is essential to ensure its reliability and effectiveness in real-world applications. This work focuses on collecting LiDAR data and developing such a methodology. The methodology's performance in safety monitoring and trespassing detection is evaluated through its effectiveness in road user detection, classification, and tracking.

4.3 Literature Review

A substantial body of railway safety literature is centered on investigating the risk factors and assessing the impact of safety countermeasures. The mechanism of non-technological interventions relies on access deterrence, enhancing general safety terms such as sight distance, and designing criteria for railway facilities. Comprehensive discussions on these criteria can be found in resources such as *Transport Canada*'s "*Grade Crossings-Handbook*" (14) and the *Federal Railway Administration's* (FRA) "*Railroad-Highway Grade Crossing Handbook*" (15).

While non-technological or physical countermeasures are commonly implemented, their effectiveness is not extensively studied. Silla and Luoma evaluated fencing, landscaping, and prohibitive signs as countermeasures to reduce illegal railroad crossings in Finland (16). These three countermeasures were installed near the authorized crossing locations (not more than 300 m away), each for 11, 10, and 17 days, respectively. Analysis of recorded videos revealed a significant reduction in trespassers: 94.6% with fencing, 91.3% with landscaping, and 30.7% with a prohibitive sign.

As explored in train-pedestrian collision literature, the technologies used for pedestrian or trespasser detection include conventional visible spectrum cameras, thermal or stereo cameras, ultrasonic sensors, Radar and ultra-wideband Radar, active and passive infrared, and LiDAR.

Video monitoring is the most common approach for perimeter monitoring. The *Federal Railroad Administration (FRA)* conducted technical research by installing a video surveillance system on a railway bridge in the United States (5). The system comprises a camera, a motion detector, and an audio warning device. Over the three-year installation period, out of 3,726 alarm-causing events, the system correctly detected 335 events and falsely reported 633 events. The remaining detections were related to other incidents, such as animal crossings or railway vehicles. Zhang et al. implemented a camera-based system in the United States to detect trespassers at a grade crossing with gates and stop signs (17). During the study, their system identified two near-miss events involving train-pedestrian conflicts.

Salmane et al. assessed a camera installation at a grade crossing in France to evaluate dangerous interactions caused by level crossing users (7). The activity monitoring system detects and tracks road users and classifies their trajectories into three scenarios: 1) a user is present at the level crossing area; 2) a user performs a zigzag maneuver when the gates are closed at the level crossing; 3) a vehicle has stopped on the railways since the leading vehicle has stopped and is blocking the railway track.

Ohta used two stereo cameras to expand the coverage area for monitoring non-motorized traffic users crossing the railway (18). Their proposed system detected 100% of the 2327 object crossings in various daylight conditions. In another study, Fakhfakh et al. installed an intelligent stereo-vision system at several grade crossings (6). They implemented an unsupervised stereo-matching algorithm to detect and separate objects from the background and to estimate a 3D model of the object's body. Their system demonstrated a recall of 96.14% and a precision of 97.34%.

A few studies have tested monitoring systems that leverage multiple technologies. One study integrated stereo and thermal cameras for monitoring platforms (a dedicated pedestrian area where people wait to board a train) to identify abnormal activities and trespassing on tracks (19). García et al. installed an array of active infrared emitters on one side and various receivers on the other side of a railway site to monitor the rail tracks. They tested their system under different weather conditions (20). Additionally, they added a camera and an ultrasonic sensor to the infrared system, enhancing the detection rate by combining their output (21).

Hilleary and Omar conducted a test using a Radar sensor as a substitute for inductive loops to detect vehicles in highway-rail grade crossings (22). Upon observing a road user, their system keeps the exit gates open to allow cars to leave the level crossing area. The Radar system demonstrated an advantage over the inductive loop due to a lower false detection rate (22). However, the Radar system has a critical limitation; it performs poorly under high precipitation conditions. In a separate study, Horne et al. installed a dual Radar system at railroad crossings (9). Their system provided feedback to the four-quadrant gate control system and triggered seven false alarms out of 477 vehicle crossings.

One application of LiDAR sensors in civil engineering is modeling the 3D structure of transportation facilities (such as intersections and roads) and building a road inventory database. Tan et al. implemented a LiDAR-based system designed explicitly for traffic signs and traffic light detection and recognition (23). Yu et al. processed LiDAR point cloud for street light pole detection (24) and road markings extraction (25). Haiyan et al. developed a LiDAR system for pavement distress type detection, effectively identifying and classifying pavement cracks or potholes (26). Moreover, the 3D model generated by LiDAR technology proves valuable in evaluating safety factors at intersections, roads, and railroads. These factors include obstacle detection, accident investigation, and identification of hazardous street sections (11).

The literature on processing a point cloud collected by 3D LiDAR sensors is not restricted to transportation applications. This literature includes unsupervised and supervised learning methods. Supervised approaches may have an edge as they use pre-labeled data for object classification. However, the available datasets are limited and impractical when the type of LiDAR sensor changes from one system to another, as point clouds collected by one LiDAR sensor can differ significantly from another. More importantly, supervised methods could be computationally more expensive, posing a challenge in real-time applications.

Nonetheless, there are some steps that both approaches have in common. For instance, the first step of some processing methods involves segmenting a LiDAR point cloud into smaller 3D cells and extracting statistical features from each cell. This feature set includes the number of points, the average and variance of intensity, the spatial mean and variance of segmented cloud points, and geometric features (27).

Frome et al. introduced regional shape descriptors for vehicle detection in point cloud data (28). One of these shape descriptors is the 3D shape context, which is extracted by computing the statistical characteristics (e.g., histogram) of the point cloud of an object in multiple spherical bins (28). Himmelsbach et al. developed an object classification method by applying the Support Vector Machine (SVM) classifier to the histogram features extracted from the point clouds (29). Yan et al. implemented a human detection and tracking algorithm. This algorithm initially segments the point cloud, estimates each cell's motion and speed, and classifies the feature set using SVM (13).

Only a few studies have implemented a LiDAR-based monitoring system for railway safety applications. Hsieh et al. installed a 2D LiDAR sensor horizontally, functioning as an array of infrared sensors for level crossing monitoring (30). Although the detection rate of their system was 99.25%, the system could miss detecting pedestrians passing behind a vehicle because of its specific setup. Besides, classification and tracking are not feasible using a 2D LiDAR system.

One advantage of 3D LiDAR sensors is their ability to monitor a wider area compared to 2D LiDAR sensors. Amaral et al. implemented a 3D LiDAR system for monitoring a level crossing (8). They validated the performance of their system with a few samples of point clouds collected from a level crossing while people were walking on the railroad. However, the results only included visualizing a few plotted examples of 3D point clouds of LiDAR detections.

Hisamitsu et al. implemented a 3D LiDAR system on a grade crossing (31). The study area was relatively small and limited due to employing a LiDAR sensor with a 30º×60º field of view and a 30-meter distance range. They extensively installed the proposed system at level crossing areas in Japan. Its mechanism is similar to that of Amaral et al. (8). However, the experimental results of the system are not available for comparison purposes.

This study's proposed 3D LiDAR-based methodology has several advantages over the previous works implemented in (8) and (31). First, the current study uses a rotational 3D LiDAR sensor that provides an extensive field of view. Due to its unique installation, the LiDAR sensor covers an area of up to 50 meters with an effective horizontal field of view of more than 180 degrees. Second, the proposed 3D LiDAR methodology provides traffic volume information while monitoring the railway facilities. Third, it delivers road use detection and tracking, an advantage over previous studies that relied solely on road user detection. This study offers a more rigorous performance evaluation of the developed system.

4.1 Hardware Components

Figure 4-1 illustrates the hardware components of the 3D LIDAR data collection system prototype that was installed at a level crossing in Montreal, Canada. The integrated hardware components for data collection include a 3D-LiDAR sensor, a camera (for collecting the ground truth data), a Raspberry Pi, a memory card, and battery packs. Although the suggested methodology and the 3D LiDAR sensor are not integrated with existing communication and alarm devices for practical application, such integration could alert road users when a train is approaching and, conversely, notify train operators when road users are illegally entering railway tracks.

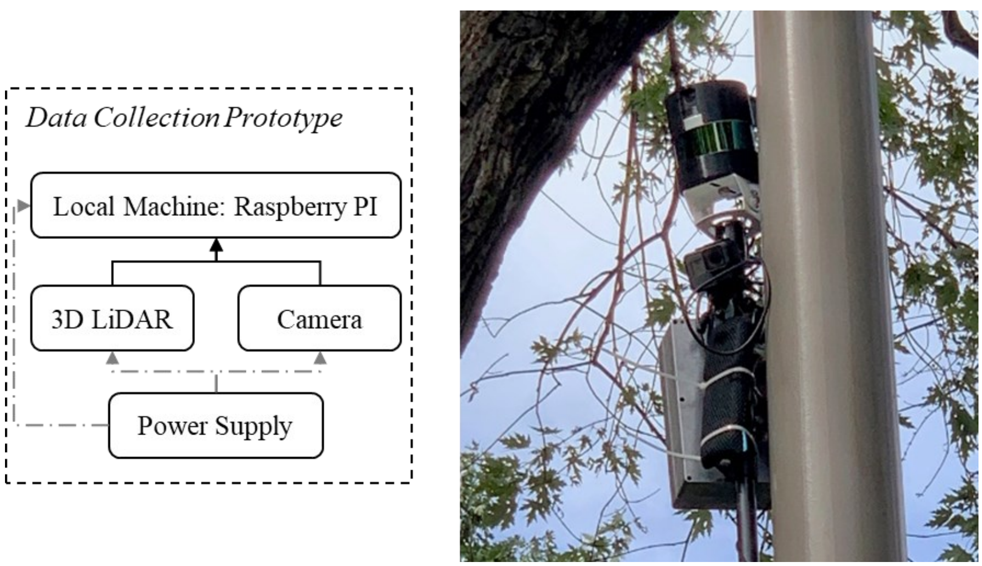


Figure 4-1 The 3D LiDAR data collection system prototype

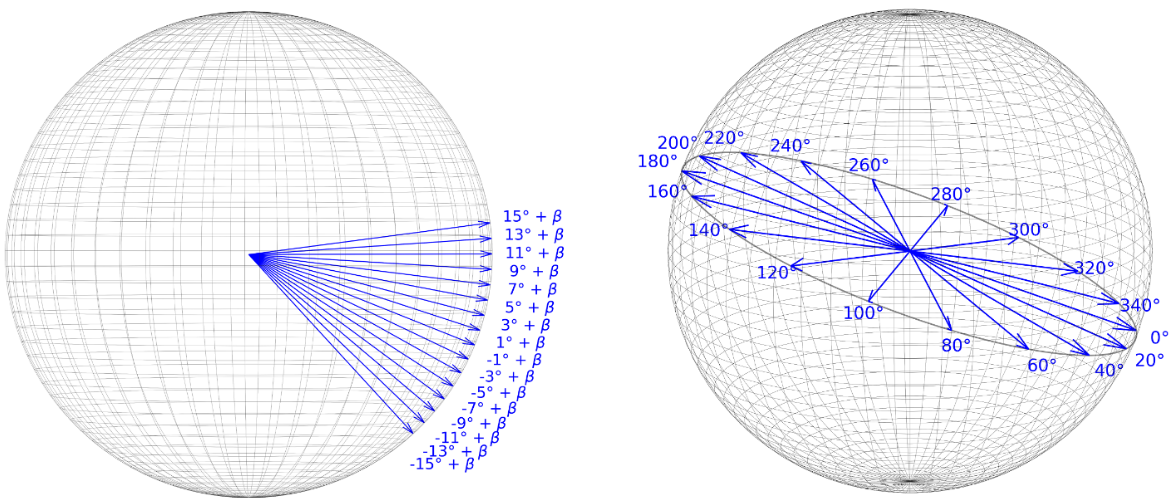
Velodyne's rotational LiDAR, VLP-16, is utilized to develop the methodology. This sensor is a lower-resolution LiDAR with relatively low cost compared to higher-resolution LiDAR sensors. **Table 4-1** presents the parameters of the 3D LiDAR sensor. The algorithm designed in the methodology section of this paper works with any rotational or solid-state LiDAR sensor that produces a two-dimensional distance matrix, with its vertical and horizontal indices corresponding to the LiDAR's field-of-view. Therefore, the proposed LiDAR methodology can adapt to changes in any of these parameters, allowing the developed unsupervised algorithms to be employed when selecting any alternative 3D LiDAR sensor.

The VLP-16 LiDAR sensor has 16 laser channels ($n_{channels} = 16$), arranged vertically, and rotates with an adjustable speed (ω) set to 10 Hz (rotations per second), enabling 3D scanning of its surrounding environment. The sensor has a 30°×360° ($\gamma_{FOW}\times \alpha_{FOW}$) field-of-view. The LiDAR sensor output includes distance and reflection measurements, each linked to a channel angle (γ_{ch_i}), corresponding to a channel ID (ch_i) ranging from 1 to 16, an azimuth value (α_{ch_i}) spanning 0° to 360°, and a timestamp (t_j). The sensor measures distance (d_{ch_i}) up to 100 m and within ±3 cm. Its vertical angular resolution ($\delta\gamma$) is 2°, while its horizontal resolution ($\delta\alpha$) is set to 0.2°.

Table 4-1 Key parameter of the 16-channel LiDAR data collection system prototype

Parameter	Notation	VLP-16 Velodyne
Number of channels	$n_{channels}$	16
Vertical field-of-view	γ_{FOW}	30°
Horizontal field-of-view	α_{FOW}	360°
Rotational speed	ω	$10\ Hz$
Time interval	δt	$0.1\ seconds$
Distance by channel i	d_{ch_i}	$[0m, 100m]$
Azimuth (horizontal angle) of channel i	α_{ch_i}	$[0°, 360°]$
Elevation (vertical angle) of channel i	γ_{ch_i}	$[-15°, +15°]$
Vertical angular resolution	$\delta\gamma$	2°
Horizontal angular resolution	$\delta\alpha$	0.2°
Distance resolution	δd	$3\ cm$
Installation height	h	$4.2m\ or\ 4.7m$
Installation title angle	β	$-18.3°\ or -12.5°$

Figure 4-2 (a & b) illustrates the horizontal and vertical operation of the 16-channel LiDAR sensor from different perspectives. The 16 laser channels simultaneously measure the distance to the surrounding objects. Tilting down the sensor by an angle (β) extends and maximizes the coverage area of the sensor. With this angle, some channels cover the ground and the bodies of road users, and the rest capture the top of the vehicles, trains, and far-away objects. The optimal angle (β) depends on the application, location, distance from the sensor to the coverage area, and installation height. **Figure 4-2 (b)** illustrates the operation of one of the LiDAR channels while tilted down at 20 different angles. In practice, each LiDAR channel captures data at 1,800 azimuths per rotation.



a) 16 laser channels - titled vertical plane   b) A single laser channel - tilted horizontal plane

Figure 4-2 The 3D LiDAR sensor's measurements in the Spherical coordinate system

4.2 Methodology of Unsupervised Algorithms

The three main components of the unsupervised algorithm developed for level crossing monitoring include:

- Point cloud data preparation
- 3D background modeling
- Road user detection, tracking, and classification

4.2.1 Point cloud data preparation

The VLP-16 LiDAR sensor, equipped with 16 laser channels, measures distance to the surrounding objects and encodes the results in binary format, represented by sequences of 0/1 bytes. The scanning data are captured in spherical coordinates, wherein each sensor's reading contains numerical values for distance, azimuth (horizontal angle), and elevation (vertical angle). According to the LiDAR sensor's product specification sheet, the binary readings are decoded and converted to spherical vectors of radius, azimuth, and elevation (32). In addition to the distance data, the reflectivity of the object's surface is reported as a 2-byte intensity variable with a maximum of 255. Each measurement is labeled with a 4-byte timestamp.

The spherical vectors are converted to the Cartesian vectors of x-y-z coordinates for further analysis. If the location of the sensor is assumed to be the center of the coordinate system, then the array of x-y-z coordinates of a single measurement captured by i^{th} channel of the LiDAR sensor is computed as Equation (4-1):

$$P_{ch_i} = \begin{bmatrix} x_{ch_i} \\ y_{ch_i} \\ z_{ch_i} \end{bmatrix} = \begin{bmatrix} d_{ch_i} \times \cos(\gamma_{ch_i}) \times \sin(\alpha_{ch_i}) \\ d_{ch_i} \times \cos(\gamma_{ch_i}) \times \cos(\alpha_{ch_i}) \\ d_{ch_i} \times \sin(\gamma_{ch_i}) \end{bmatrix} \tag{4-1}$$

where P_{ch_i} is the vector of the x-y-z coordinates, and d_{ch_i}, α_{ch_i} and γ_{ch_i} are the distance, azimuth, and elevation of the laser channel i. The elevation, γ_{ch_i}, is the vertical angle associated with each laser channel and is derived as $\gamma_{ch_i} = 2 \times (i - 8) - 1$, where i is the channel ID.

Generally, the sensor is positioned at the height of h and tilted down by an angle β. The actual x-y-z coordinates of the observed point, P'_{ch_i}, are determined by applying a rotation around the x-axis, R_x, and a translation vector, $\overrightarrow{h_L}$, in the direction of the z-axis (Equation (4-2)):

$$P'_{ch_i} = \begin{bmatrix} x'_{ch_i} \\ y'_{ch_i} \\ z'_{ch_i} \end{bmatrix} = R_x P_{ch_i} + \overrightarrow{h_L} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\beta) & -sin(\beta) \\ 0 & sin(\beta) & cos(\beta) \end{bmatrix} \begin{bmatrix} x_{ch_i} \\ y_{ch_i} \\ z_{ch_i} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix} \tag{4-2}$$

The relative position of the point with respect to the pole to which the LiDAR sensor is mounted is determined by $[x'_{ch_i}, y'_{ch_i}]$, and the height of the point is determined by z'_{ch_i}.

A frame is defined as the duration the LiDAR sensor requires to complete a full rotation. The duration for a full rotation is 0.1 seconds, the inverse of the rotational speed ($1/\omega = 1/10Hz$). The set of points observed by every LiDAR channel in each single rotation (frame) forms a 3D point cloud of the surrounding environment. The point cloud at the frame f_j, corresponding to the time interval $[t_j, t_j + 0.1]$, is defined as Equation (4-3):

$$PCL_{f_j} = \left\{P'_{ch_i,\alpha_{ch_i}}\right\} \; where \; ch_i \in [1, 16] \; and \; \alpha_{ch_i} \in [0°, \; 360°] \tag{4-3}$$

4.2.2 3D background modeling

The second step in this research methodology for raw LiDAR data processing involves implementing the background modeling algorithm. This step concentrates on generating a point cloud representing background objects at the level crossing site. The algorithm's main objective is to identify and classify objects, excluding road users, as part of the background. The foreground objects exclusively consist of road users, including trains, vehicles, and vulnerable users. The detection of road users relies on comparing the observed point cloud with the constructed background model. Therefore, accurately estimating a robust and well-structured background model is crucial for ensuring the overall effectiveness of the model.

In the background modeling process, the three-dimensional space centered at the LiDAR position is discretized into smaller segments. The discretization could involve dividing the space into cubic or spherical volumes based on the coordinate system format.

A spherical volume (range-view) is characterized by its radius (distance to LiDAR), azimuth angle, and elevation angle (10). For example, segmentation in the spherical system could involve dividing the radius into small sections of 0.2m and azimuth angle into small segments of 0.2° width. The elevation angle is restricted to the number of LiDAR channels (16 channels). In this scenario, an area up to 40m away from the LiDAR sensor, with a vertical and horizontal coverage of 30° and 180°, respectively, is segmented into 2,700,000 spherical volumes (200×15×900).

A voxel is a small cubic volume in the Cartesian coordinate system, where each Cartesian axis is discretized uniformly (33). For example, if the dimensions of the observed area that the LiDAR sensor aims to monitor are 40m×40m×5m, with a discretization factor of 0.2m, the partitioned 3D space would contain 1,000,000 voxels (200×200×25).

Each segmentation approach comprehensively reconstructs the observed area's three-dimensional space. The background modeling in this research methodology is based on the latter approach, involving discretization into cubic volumes. The dimension of the coverage area in the selected level crossing is 40m×40m×5m. The discretization factor, d_f, is set to 0.2m. Therefore, the 3D discretized space comprises one million voxel grids.

For cubical voxelization, the LiDAR data is converted to x-y-z coordinates. **Figure 4-3 (a)** illustrates a hypothetical sub-space divided into small voxels. Each voxel (v_k) is identified by an ID, k, and its discretized coordinates, $\left[x_{d_k}, y_{d_k}, z_{d_k}\right]$, calculated as Equation (4-4):

$$\left[x_{d_k}, y_{d_k}, z_{d_k}\right] = \left[x'_k/d_f, y'_k/d_f, z'_k/d_f\right] \tag{4-4}$$

The background modeling algorithm utilizes an initial set of 3,000 individual point clouds captured within a five-minute interval. To adapt to dynamic changes, particularly setup vibrations, the background model undergoes updates every 15 minutes, ensuring its continuous accuracy and relevance.

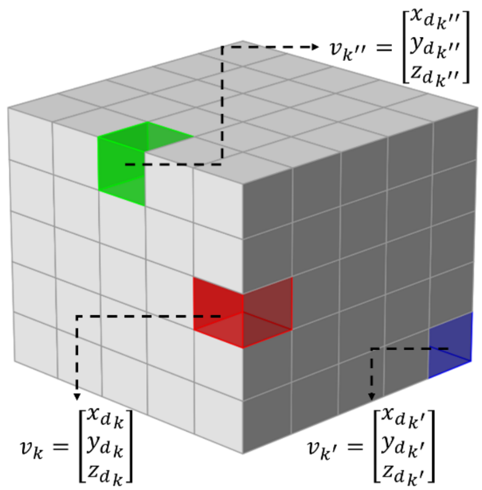
Figure 4-3 (b) provides an overview of the background modeling. The left-side table presents discretized voxel grids in terms of x_{d_k}, y_{d_k} and z_{d_k}. For each voxel, v_k, the algorithm counts and stores the number of points observed in the initial set of 3,000 frames (point clouds) as N_{v_k}. A voxel is labelled as background if N_{v_k} exceeds an upper threshold, denoted as $\widetilde{N}_U$. This criterion ensures frequently observed voxels in the initial point cloud set are labeled as background.

Conversely, a voxel is excluded from consideration for the background if N_{v_k} is below a lower threshold, denoted as $\widetilde{N}_L$. An iterative algorithm utilizing K-Nearest Neighbor (KNN) is implemented to classify voxels with N_{v_k} in the range of $\widetilde{N}_L$- $\widetilde{N}_U$. If the adjacent voxels to the voxel under review are labeled as background, the voxel is classified as background.

The background modeling algorithm utilizes an initial set of 3,000 individual point clouds captured within a five-minute interval. To adapt to dynamic changes, particularly sensor vibrations, the background model undergoes updates every 15 minutes, ensuring its continuous accuracy and relevance.

$$BG_{v_k} = \begin{cases} 0, & N_{v_k} < \widetilde{N}_L \\ 1, & N_{v_k} > \widetilde{N}_U \\ KNN, & \widetilde{N}_L \le N_{v_k} \le \widetilde{N}_U \end{cases} \tag{4-5}$$

The upper and lower threshold values are determined in a calibration process. Since the initial set contains 3,000 frames (point clouds), the upper threshold is calibrated as 70% of the number of frames ($\widetilde{N}_U = 2100$), and the lower threshold is calibrated as 20% of 3,000 ($\widetilde{N}_L = 600$).



a) Voxel: cubic volume

Number of Points in Voxels

v_k: x_{d_k}	y_{d_k}	z_{d_k}	N_{v_k}	BG_{v_k}
2	2	2	⋮	⋮
⋮	⋮	⋮	⋮	⋮
22	10	2	$N_{v_k} > \widetilde{N}_U$	1
22	10	4	$N_{v_k} > \widetilde{N}_U$	1
22	10	6	$\widetilde{N}_L \le N_{v_k} \le \widetilde{N}_U$	KNN
22	10	8	$\widetilde{N}_L \le N_{v_k} \le \widetilde{N}_U$	KNN
22	10	10	$N_{v_k} < \widetilde{N}_L$	0
22	10	12	$N_{v_k} < \widetilde{N}_L$	0
⋮	⋮	⋮	⋮	⋮
400	400	50	0	0

Background Voxels

v_k: x_{d_k}	y_{d_k}	z_{d_k}	BG_{v_k}
⋮	⋮	⋮	⋮
22	10	2	1
22	10	4	1
22	10	6	<u>1</u>
⋮	⋮	⋮	⋮

b) Background modeling

Figure 4-3 The 3D segmentation and background modeling

The right-side table in **Figure 4-3 (b)** illustrates the background model where voxels with BG_{v_k} equal to 0 are discarded. This effectively reduces the size of the voxel grid in the background model since the majority of voxels represent open space and their corresponding N_{v_k} is 0. In every frame, the LiDAR sensor captures approximately 28,800 points (16 channels × 1800 azimuth indices). Therefore, the size of the background model is reduced from one million voxels to a subset of approximately 28,800 voxels. Increasing the discretization factor, d_f, further reduces the size of the background model and diminishes the resolution, impacting the accuracy of road user detection.

4.2.3 Road user detection, tracking, and classification

This section discusses the procedures for frame-by-frame processing of raw LiDAR data. Following one full rotation, the LiDAR sensor's measurements are stacked to form a data frame including distance, reflection, azimuth, elevation, and timestamp corresponding to each frame. This section is structured in three subsections: 1) road user detection and clustering, 2) road user tracking, and 3) road user classification.

Road user detection and clustering

Table 4-2 outlines the steps of road user detection, tracking, and classification. Initially, the current point cloud is discretized, and a voxel ID, k^*, is assigned to each observed point. A voxel is indexed and accessible through three discretized coordinates: $[x_{d_{k^*}}, y_{d_{k^*}}, z_{d_{k^*}}]$. When the corresponding voxel of the point belongs to the background, the point is excluded from the set of foreground points. Conversely, if the corresponding voxel does not exist in the background model, the point is included in the preliminary foreground point cloud.

In the 3D space segmented into voxel grids, each voxel is adjacent to six neighboring voxels (sharing a face with the central voxel) and 20 diagonal neighbors (sharing only a vertex). With a discretization factor of 0.2m, the maximum distance between two points in neighboring voxels is 0.4m. For every point in the preliminary foreground point cloud, if there is no other foreground point in the same voxel and the six neighboring voxels, the point is considered noise or false detection and is discarded from the initial foreground point cloud (**Table 4-2**: 1-2).

Once the potential foreground points are detected and separated as a point cloud of x-y-z coordinates, a spatial-based clustering method is employed to cluster and construct the point cloud of each individual road user (**Table 4-2**: 3-4). The Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is calibrated to cluster the 3D point cloud associated with the foreground objects at the level crossing (34).

DBSCAN functions without necessitating the number of clusters as an input variable. Instead, it requires the specification of a minimum number of samples and a minimum distance to prevent the fragmentation of objects into smaller clusters. Based on the collected LiDAR data, the point cloud of a pedestrian at its maximum distance to the sensor comprises four or more samples. In contrast, cars, trucks, and trains have considerably larger point clouds compared to pedestrians. Consequently, the minimum number of samples in the DBSCAN algorithm is configured to be four. Objects with fewer samples than this threshold are classified as noise and removed from the foreground objects.

The position of the road user is determined by the average of the x and y coordinates of every point assigned to a cluster, which is equivalent to the cluster's centroid. Additionally, the maximum z-coordinate of all points assigned to a cluster determines the road user's height.

The Singular Value Decomposition (SVD) algorithm is applied to extract road users' shape, length, and width from the clustered point clouds (35). The SVD is applied to the covariance matrix of x-y coordinates of a clustered point cloud (j) as Equation (4-6):

$$cov(X_j, Y_j) = \left[X_j - \bar{X}_j \quad , \quad Y_j - \bar{Y}_j\right]^T . \left[X_j - \bar{X}_j \quad , \quad Y_j - \bar{Y}_j\right] \tag{4-6}$$

The cluster's shape is determined using the eigenvalues (λ_1 and λ_2) and eigenvectors ($\vec{u}_1$and $\vec{u}_2$) of the covariance matrix. The object's length ($length_j$) and width ($width_j$) are derived by taking the square root of these eigenvalues multiplied by each eigenvector's magnitude (Equation (4-7)). The cluster orientation in the x-y plane is defined as: $\theta_j = \arctan\left(u_{1_y}/u_{1_x}\right)$.

$$length_j = \sqrt{\lambda_1} \times \|\vec{u}_1\|_2 \,\&\, width_j = \sqrt{\lambda_2} \times \|\vec{u}_2\|_2 \; where: \lambda_1 \geq \lambda_2 \tag{4-7}$$

Table 4-2 Unsupervised LiDAR data processing algorithm for a given frame

1	convert from spherical to cartesian *point cloud* and apply rotation and translation
2	discretize the *point cloud* and detect the *foreground point cloud*
3	cluster *foreground point cloud*: apply DBSCAN and exclude small clusters
4	for every *cluster* created by DBSCAN: fit a *convex hull*, apply SVD, and extract features
5	if there is no active *tracker* from the previous timestamp: initialize road user's tracker with position, zero velocity and *Kalman Filter*
6	solve *Data Association* algorithm for the distance cost function
7	update associated *tracker*'s KF with *observations* and predict their next *positions*
8	create new *tracker*s for *unassigned observations* and terminate *unassigned trackers*
9	for an *assigned tracker:* update speed, over compute *velocity* based on current and previous *observations* overlay *road users' point cloud* with the level crossing's geospatial polygon run the multiple classification criteria

Road user tracking

Table 4-2 (5-9) outlines the road user tracking and classification process. A new road user, detected and clustered for the first time, is assigned a dynamic object, *tracker*. This *tracker* retains pertinent information, including the coordinates of the road user's current and past positions and features extracted from the road user's point cloud. The *tracker* utilizes a prediction algorithm to forecast the road user's position in the next frame or timestamp. The projected position of road users from the previous frame is compared with the observed position of road users in the current frame to track their movement through consecutive frames.

The data association (observation-prediction association) utilizes a linear assignment solver that minimizes the general cost of associating road users between two consecutive frames (36). The cost function is formulated as the sum of the distance between positions of the associated observed and predicted pairs (Equation (4-8)):

$$Cost_{association} = \sum_{(O_i,P_j)}^{N_{pairs}} \sqrt{\left(x_{O_i} - x_{P_j}\right)^2 + \left(y_{O_i} - y_{P_j}\right)^2} \tag{4-8}$$

here, (O_i, P_j) represents the set of associated observation-prediction pairs, N_{pairs} is the total number of *tracker*s that are assigned to the current set of observations, and $\sqrt{.}$ is the Euclidean distance function.

The prediction component of the *tracker* utilizes the Kalman Filter constructed with four state variables. The transition equation of the Kalman Filter is characterized as (Equation (4-9)):

$$\begin{bmatrix} x \\ V_x \\ y \\ V_y \end{bmatrix} = \begin{bmatrix} 1 & \delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \delta t \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^- \\ V_x^- \\ y^- \\ V_y^- \end{bmatrix} \tag{4-9}$$

here, $(.)^-$ denotes the state variables from the previous timestamp, $[V_x, V_y]$ is the velocity vector, and $\delta t = 0.1\ seconds$ is the time interval.

Once the association is completed, the instantons velocity of the road user, j, is updated as Equation (4-10):

$$\begin{bmatrix} V_{x_j} \\ V_{y_j} \end{bmatrix} = \delta t^{-1} \begin{bmatrix} x_{O_j} - x_{O_j}^{-} \\ y_{O_j} - y_{O_j}^{-} \end{bmatrix} \; m/s \tag{4-10}$$

Consequently, the average velocity of the road user, $\vec{\bar{V}}_j$, at its n_j^{th} observation is updated using the moving average filter as Equation (4-11):

$$\vec{\bar{V}}_j = \begin{bmatrix} \bar{V}_{x_j} \\ \bar{V}_{y_j} \end{bmatrix} = n_j^{-1} \left((n_j - 1) \begin{bmatrix} \bar{V}_{x_j}^{-} \\ \bar{V}_{y_j}^{-} \end{bmatrix} + \begin{bmatrix} V_{x_j} \\ V_{y_j} \end{bmatrix} \right) m/s \tag{4-11}$$

The average speed of the road user is computed using Equation (4-12):

$$\bar{S}_j = \left\| \vec{\bar{V}}_j \right\|_2 = \sqrt{\bar{V}_{x_j}^{\;2} + \bar{V}_{y_j}^{\;2}} \tag{4-12}$$

Any road user not paired with a *tracker* from the previous timestamp is treated as a newly observed user and assigned an initialized *tracker*. Similarly, any *tracker* not associated with an observation is flagged, and the *tracker* is terminated if it remains unassociated for five frames.

Road user classification

At each step, the convex polygon surrounding the road users' point cloud in the x-y plane is overlayed with the geospatial boundaries of level crossing elements, including the level crossing area, railway tracks, streets, crosswalks, and sidewalks. **Figure 4-4** illustrates the manual calibration of geospatial boundaries of the selected level crossing. The results of this geospatial analysis are used as part of the road user classification algorithm.

An unsupervised multi-criteria method is developed for classifying road users as trains, non-motorized (pedestrian and cyclist) road users, and motorized road users, including cars and trucks. The unsupervised classification method is not trained with labeled samples of road users. Instead, pertinent information such as length, width, speed, and geospatial data of road users at each timestamp is utilized for classification.

The geospatial data includes every section intersected by the road user's polygon and the proportionate area occupied by the road user in each corresponding geospatial section. For example, the polygon of a pedestrian, road user k, who is beginning to enter the crosswalk from the sidewalk intersects with both road sections. Therefore, two proportionate areas for this road user are reported: $a_{Sidewalk_k}$ and $a_{Crosswalk_k}$. The geospatial set corresponding to the road user k is denoted as G_k, and the set of proportionate areas is denoted as a_{G_k}.

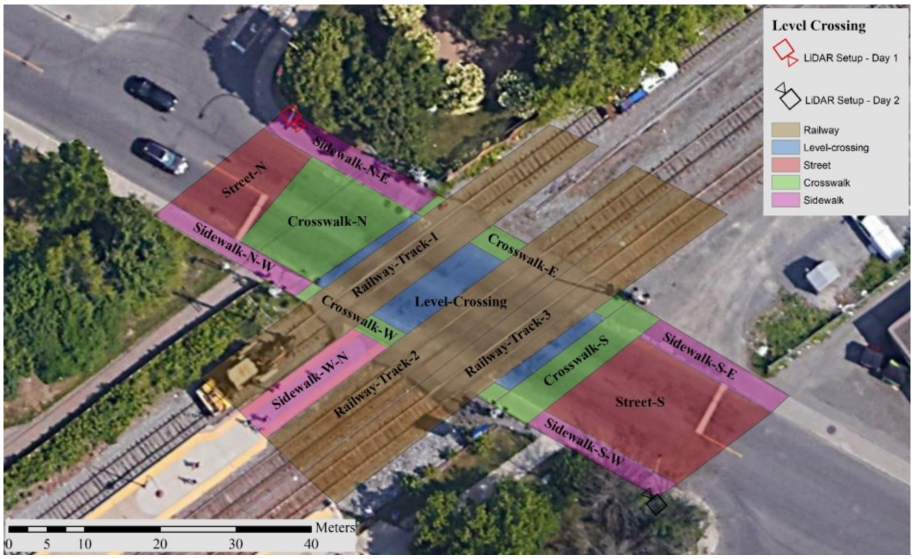


Figure 4-4 Geospatial boundaries of road sections in the level crossing

There are a few constraints on classifying road users. The steps for unsupervised road user classification can be summarized as follows:

1. If the geospatial set of the road user contains any sidewalks, then the user is classified as a pedestrian.
2. If the geospatial set of the road user contains any of the two crosswalks, East or West, the classification as a pedestrian or a train is determined by the length of the road user.
3. If the geospatial set includes a railway track, the classification as a train or not is determined by the length of the road user.
4. If the geospatial set does not include a railway, then the road user's class is not a train, and a long road user is classified as a truck.
5. If the road user's speed exceeds 10 km/hour, then the user is not a pedestrian, and the length of the road user determines the classification as a cyclist or otherwise.

6. For any other cases, the speed, length, and proportionate areas are used for road user classification. The multi-criteria classification based on the length and width of the road user is formulated as follows:
   - if $lenght_k < 1m$: *pedestrian/cyclist*,
   - else if $1m < lenght_k < 2m$: *a group of pedestrians/cyclists* or *cars,* check G_k, a_{G_k}, and $\bar{S}_k$,
   - else if $2m < lenght_k < 5m$: *car*,
   - else if $5m < lenght_k < 7m$: *car or truck*, check the object's width:
     - if $width_k < 2m$: *car*; else: *truck*,
   - else if $length_k > 7m$: *check geospatial:*
     - if G_k includes railway tracks and $a_{railway_k} \geq 90\%$: *train*; else *truck.*

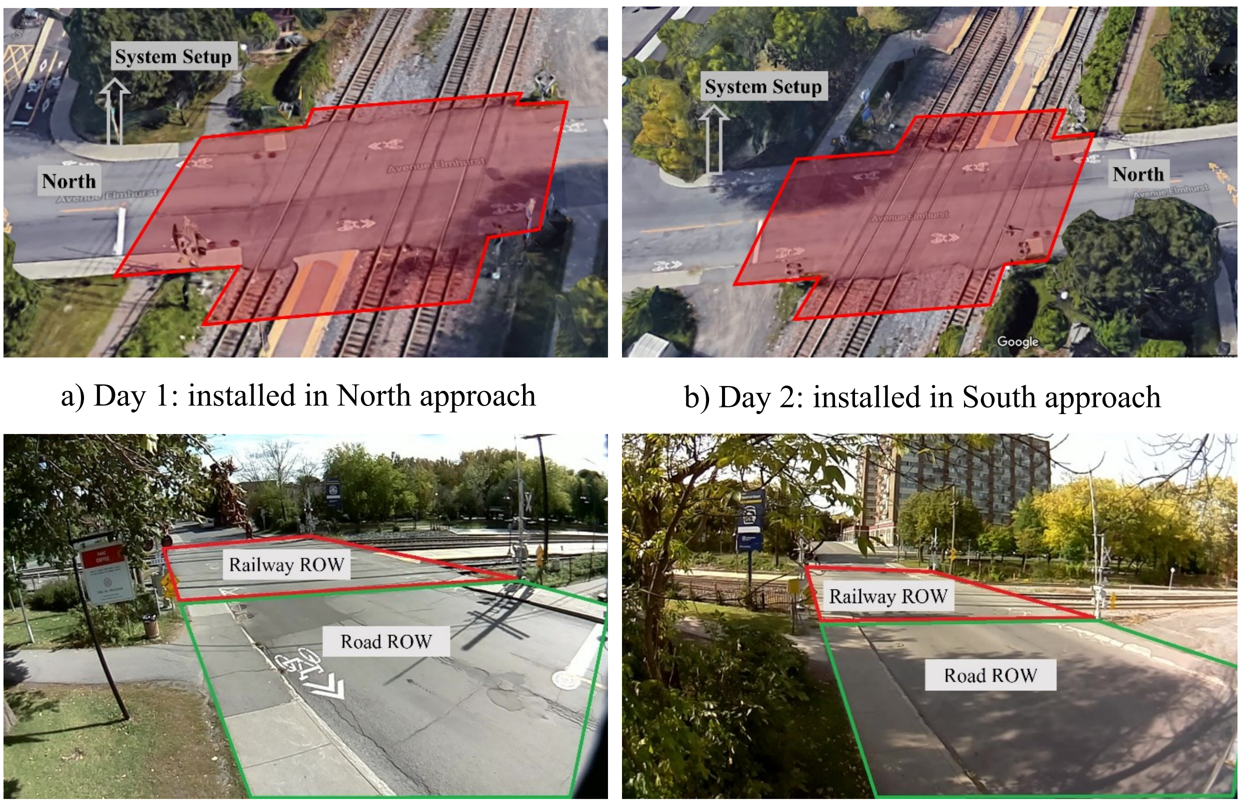
4.3 Performance Evaluation

Trespassing events on railway facilities are rare; therefore, a performance evaluation based solely on these events is unsatisfactory. Instead, the proposed methodology and its components are evaluated regarding the detection, classification, and tracking of every road user passing through the coverage area. The ground truth data is generated manually from video footage collected simultaneously. Additionally, the performance of the methodology in trespassing detection is also evaluated.

4.3.1 Application - case study

The selected study location is a level crossing with high traffic volumes in Montreal, positioned at the railway intersection with Avenue Elmhurst. The chosen level crossing has four gates: two long-arm gates designed for cars and pedestrians and two short-arm gates exclusively for pedestrians. The long gates obstruct the vehicular lane entering one approach and its adjacent sidewalk, while the short gates block the sidewalk near the exiting vehicular lane. The level crossing is near a train station and experiences a high volume of pedestrian activities. The train station includes a platform between two railway tracks connected to a crosswalk on the west side of the level crossing. As a result, users need to cross one of the railway tracks to access the platform.

Figure 4-5 (a & b) displays the aerial view of the level crossing using Google's satellite imagery. The LiDAR was installed on two different days at this level crossing. On the first day, it was placed in the north vehicular approach, 23m from the center of the level crossing, with an installation height of 3.8m and a downward tilt of -16.3°. On the second day, it was installed at the south vehicular approach and 31m from the center of the level crossing, with an installation height of 4.4m and a downward tilt of -12.5°. As the distance from the LiDAR to the center of the level crossing increases, the installation height increases, and the tilting angles become more horizontal.



a) Day 1: installed in North approach

b) Day 2: installed in South approach

c) Day 1: collected video footage

d) Day 2: collected video footage

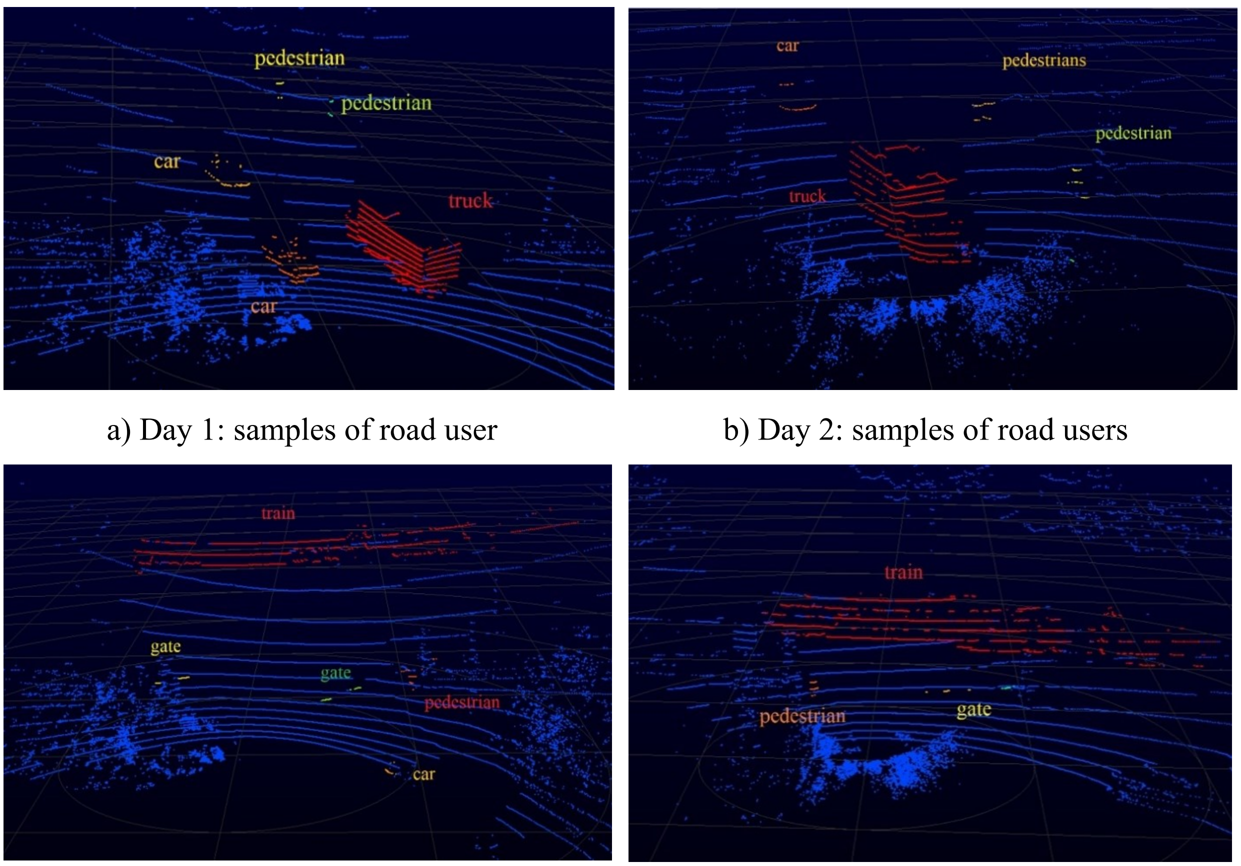
Figure 4-5 Aerial view and map of the level crossing (Google Maps and Google Earth)

In this study, the placement of the LiDAR was determined by the availability of nearby poles around the railway facilities. The LiDAR was installed at a distance from the level crossing to ensure the safe operation of trains. For a permanent setup, it is advisable to designate a dedicated pole for the sensor. Such a setup guarantees optimal performance by achieving a balanced resolution across the coverage area. The LiDAR sensor has a broad field of view spanning from 0° to 359.9°, with the capability to measure distances up to 100 meters. The red polygons (**Figure**

4-5 (a & b)) delineate the chosen coverage area of the LiDAR, where it detects and tracks road users. **Figure 4-5 (c & d)** presents a sample frame from the collected video footage for each day of data collection. Red polygons on both images show the railway's right-of-way, while green polygons outline the road's right-of-way. Red polygons include the level crossing area and the East and West crosswalks in **Figure 4-4**.

4.3.2 Road user detection

Figure 4-6 (a & b) illustrates two samples while multiple road users are detected.



a) Day 1: samples of road user

b) Day 2: samples of road users

c) Day 1: frame with a train sample

d) Day 2: frame with a train sample

Figure 4-6 Road user detection and clustering illustration

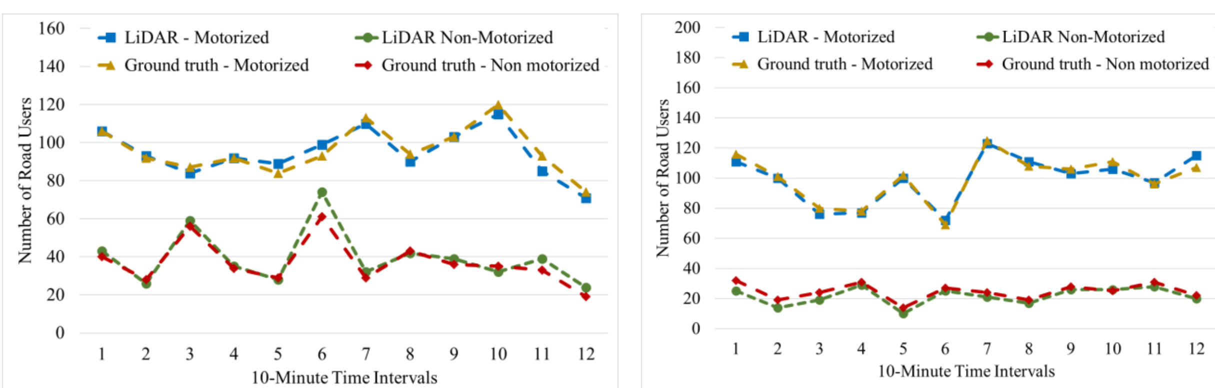
The detection accuracy of the LiDAR methodology is evaluated by manually comparing the number of road users per frame with ground truth video frames for a randomly selected set of 200 frames per data collection period. The detection rate of the proposed methodology varies based on road user class. Within its coverage range, the methodology applied to the low-resolution LiDAR's point cloud data detected 93.3% and 91.2% of pedestrians and cyclists on the first and second days, respectively. Furthermore, it achieved a car detection rate of 95.3% on the first day and 94.1% on the second day. Additionally, the detection rate for trucks and trains was 100% on both days.

Errors can occur in foreground detection or clustering of the road users' point cloud. In such cases, two road users may be clustered into a single point cloud, or one large road user may be split into two or more clusters. DBSCAN performs effectively not only for small clusters but also for larger ones. **Figure 4-6 (c & d)** illustrates two large point clouds representing samples of two trains; both were accurately detected and classified.

4.3.3 Count results

Automated counts of road users detected by the LiDAR are cross-referenced with ground truth counts. Ground truth data is obtained through manual verification of the video data recorded over a 2-hour duration for each day of data collection. This ground truth data is counted and aggregated in 10-minute intervals. LiDAR counts are similarly aggregated within the same time intervals. Overall, there are 12 ten-minute time intervals per day of data collection.

Figure 4-7 (a & b) shows the motorized and non-motorized traffic flows over these 12 ten-minute intervals for both days. Although the flow diagrams of the proposed methodology are very close to ground truth, there are cases of overcounting or undercounting. The overcounting occurs when a trajectory of the road users is split into two trajectories because of occlusion, and the undercounting occurs when two or more road users form one cluster and have one trajectory.



a) Data collection – Day 1 b) Data collection – Day 2

Figure 4-7 LiDAR and ground truth counts in 10-minute intervals

Table 4-3 presents the performance of the proposed methodology in various aspects. In the ground truth data, 443 vulnerable and 1,151 motorized road users were observed on the first day, and 296 vulnerable and 1,144 were observed on the second. For evaluating the performance of the methodology, the Average Absolute Percentage Deviations (AAPD) is calculated as Equation (4-13):

$$AAPD = 100 \times N_i^{-1} \times \sum_{i=1}^{N_i} \left| N_{GT_i} - N_{LiDAR_i} \right| / N_{GT_i} \quad (4\text{-}13)$$

where N_{GT_i} is the number of road users in i^{th} time interval in the ground truth set, N_{LiDAR_i} is the number of road users counted by the LiDAR in the same period, and N_i is the number of time intervals.

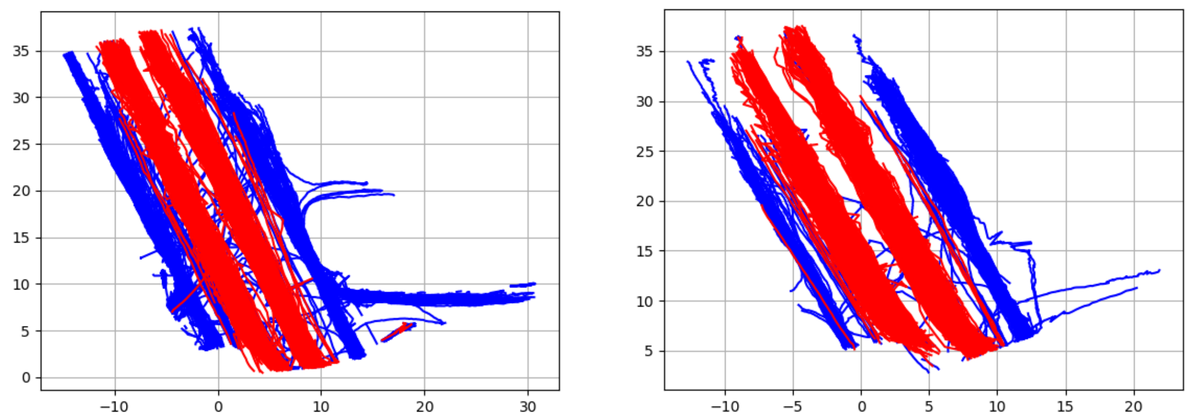
Table 4-3 Summary of the count results

Date of Installation	First Day: 2018-09-27		Second Day: 2018-10-05	
Count Type	Ground truth (video)	LiDAR-based methodology	Ground truth (Video)	LiDAR-based methodology
Duration of Analysis	2 hours		2 hours	
Number of observed and detected road users				
Pedestrian or Cyclist	443	473	296	260
Car	1,075	1,047	1,144	1,105
Truck	76	90	55	86
Train	7	7	9	9
Non-Motorized Users	443	473	296	260
Motorized Users	1,151	1,137	1,199	1,191
Average Absolute Percentage Deviation (AAPD)				
Motorized	-	5%		3%
Non-Motorized	-	10%		14%

The AAPD of counting motorized and non-motorized users on the first day are 3% and 10%, respectively, and on the second day are 5% and 14%. On the first day, the LiDAR overcounted non-motorized road users due to its proximity to the main crossing point for pedestrians. The LiDAR undercounted non-motorized road users on the second day due to its distance from the major intersection.

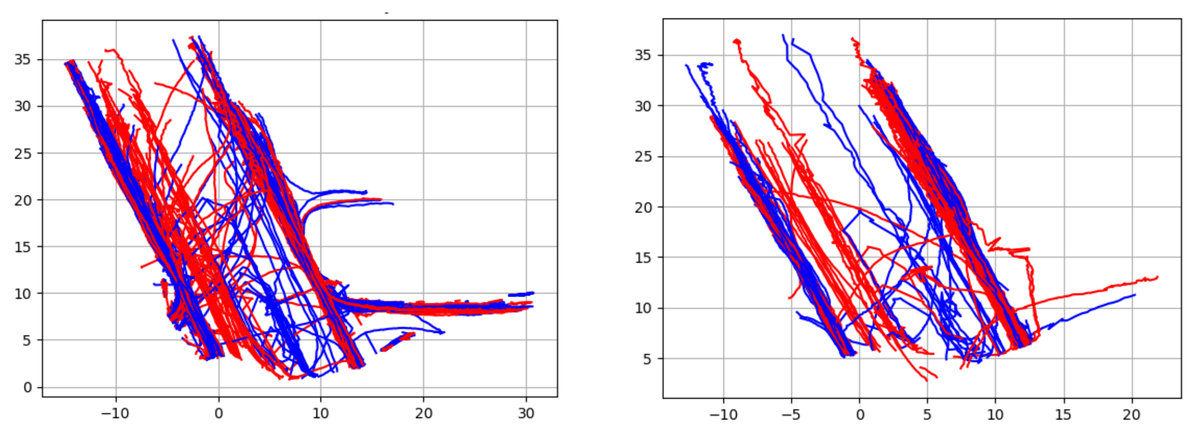
4.3.4 Road users’ trajectories

The trajectories of the road users provide a better understanding of the road users’ activities and movement patterns in the coverage area. These trajectories can be utilized for surrogate safety analysis and identifying near-miss events. **Figure 4-8 (a & b)** illustrates the non-motorized users’ trajectories (blue lines) versus motorized users (red lines). **Figure 4-9 (a & b)** shows the trajectories of the non-motorized road users, including pedestrians and cyclists, in two different directions. It is worth noting that a platform connected to the center of the level crossing provides road users access to the train station. Therefore, pedestrian trajectories appear out of place as some pedestrians have crossed the level crossing center area instead of nearby crosswalks.



a) First Day b) Second Day

Figure 4-8 The trajectories of motorized (red) versus non-motorized (blue) road users



a) First Day b) Second Day

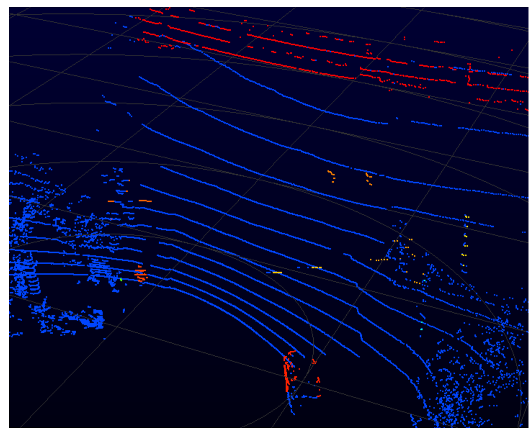
Figure 4-9 The trajectories of vulnerable road users

4.3.5 Interaction between trains and vulnerable road users

In total, four hours of LiDAR data were collected each day. The initial two hours were utilized to calculate specific performance measures. For safety analysis of the level crossing, the four hours of data each day are processed using the LiDAR methodology. Within these four hours, the LiDAR accurately identified 14 trains on the first day and 17 trains on the second.

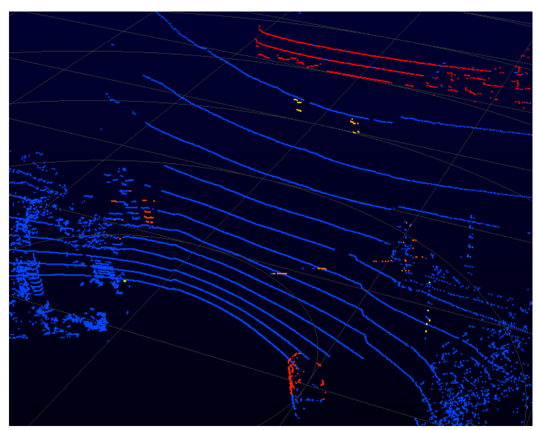
The activities of road users during train crossing are monitored and cross-referenced with the LiDAR results. During the analysis, the LiDAR methodology correctly identified three conflicts, including an event involving two pedestrians and a train approaching the level crossing.

a) two pedestrians entering the level crossing



b) point cloud of the same conflict in (a)

c) two pedestrians crossing the railway



d) point cloud of the same conflict in (c)

Figure 4-10 Trespassing detection of two pedestrians

In **Figure 4-10 (a)**, two pedestrians can be observed passing through the gates and entering the railway right of way as a train crosses. **Figure 4-10 (b)** displays the output of the 3D LiDAR, accurately detecting two pedestrians, the train, and other road users. **Figure 4-10 (c & d)** shows the same pedestrians crossing the railway while the train still passes through the level crossing. This sequence of events suggests that the pedestrians intended not to access the train station using the middle platform but to illegally cross the railway tracks while the train had the right of way.

4.4 Conclusion and Future Work

This research paper proposes and tests a methodology based on 3D LiDAR sensors for traffic and safety monitoring of railroad facilities. This paper describes the proposed methodology's components, including hardware and algorithms. The hardware components for data collection integrate a low-resolution rotational LiDAR sensor. The algorithm comprises background modeling, object detection, clustering, and tracking.

The background is modeled as the observation frequency in the discretized cubic volumes (voxels) in the three-dimensional Cartesian coordinate system. To improve the accuracy of background reconstruction, a K-Nearest Neighbor algorithm is applied to the voxels with a lower observation frequency. For road user detection, the Voxelized point cloud of each frame is compared with the voxel grid of the background model. Subsequently, a density-based spatial clustering algorithm is applied to group points into distinct road users. Various physical attributes, including x-y coordinates of the cluster center, length, width, height, area, and volume, are derived from each clustered point cloud.

Each clustered point cloud is designated as a road user, and a tracker is assigned to each. The tracker observes the road user's position at each frame, estimates its speed, and predicts its next position. Road user tracking aims to establish associations between observations of the same road user across consecutive frames.

A non-supervised classification algorithm is introduced to classify road users into pedestrian/cyclist (combined), car, truck, and train. This algorithm relies on road users' speed, dimensions, and geospatial attributes to predict their class. For example, a road user whose point cloud is on the sidewalk is consistently classified as a pedestrian. In another example, a road user exceeding 7 meters might fall into either the train or truck category. Therefore, its classification is determined by examining its point cloud with regard to the railway polygon. If the point cloud is entirely within the railway polygon, it is clustered as a train; otherwise, it is classified as a truck.

The proposed LiDAR system was installed at a level crossing in Montreal, Canada, on two days. The performance of the 3D LiDAR-based methodology, employing unsupervised methods, is promising. The detection rate for vulnerable road users (pedestrians and cyclists) ranges from 91.2% to 93.3%, and for cars, it ranges from 94.1% to 95.4%. Trucks and trains, given their relatively large point cloud, present a detection rate of 100%.

In addition to frame-by-frame comparison, the performance of the LiDAR system over two hours on both days is assessed. The LiDAR system's counts are compared against manually collected ground truth data obtained from recorded videos. The Average Absolute Percentage Deviation (AAPD) for counting motorized road users is 5% and 3%, while for non-motorized road users, the AAPD ranges between 10% and 14% across two different system setups.

The proposed methodology could empower railway systems to automate the monitoring of vulnerable road users' activities at level crossings or railway facilities. An activity-recognition algorithm can find whether a person is crossing the area or stationery while a train is approaching the facilities. During the data collection period, the LiDAR system correctly identified three conflicts, including the trespassing of two pedestrians.

As part of future work, semi- or fully-supervised methods will be developed and compared with the proposed unsupervised method. Although these alternative methods could perform better than the proposed method, generating labeled (annotated) data for training would be highly time-consuming. Other future work may involve the long-term installation of the LiDAR system and the integration of a warning system. Evaluating the proposed methodology on a more extensive set of railway facilities could enhance its assessment. Exploring the integration of computer vision with LiDAR systems could open up promising avenues for future research. Such integration has the potential to improve both system redundancy and accuracy. Additionally, evaluating alternative surrogate safety measures will be explored for safety applications. The evaluation of alternative sensor resolutions, including those with more channels, will be conducted to assess their impact on classification and tracking performance.

Acknowledgments

Transport Canada provided funding for this project.

Conflicts of Interest

The authors confirm that there are no conflicts of interest associated with this publication.

References

1. Transportation Safety Board of Canada. Railway Occurrence Data (2013-2022) Canada, 2019 [Available from: http://www.bst-tsb.gc.ca/eng/stats/rail/data-1.html.

2. Velodyne. VLP-32c LiDAR (Puck Hi-Res) United States, 2018 [Available from: https://velodynelidar.com/products/puck-hi-res/.

3. Havârneanu GM, Burkhardt J-M, Paran F. A systematic review of the literature on safety measures to prevent railway suicides and trespassing accidents. Accident Analysis & Prevention. 2015;81:30-50.

4. Fitzpatrick K, Warner J, Brewer M, Bentzen BL, Barlow JM, Sperry B. Treatments Used at Pedestrian Crossings of Public Transit Rail Services. 2014.

5. daSilva MP, Baron W, Carroll AA. Railroad infrastructure trespassing detection systems research in Pittsford, New York. United States. Federal Railroad Administration; 2012.

6. Fakhfakh N, Khoudour L, Bruyelle J-L, El-Koursi E-M. Intelligent Surveillance System Based on Stereo Vision for Level Crossings Safety Applications. Recent Developments in Video Surveillance: InTech; 2012.

7. Salmane H, Khoudour L, Ruichek Y. A video-analysis-based railway–road safety system for detecting hazard situations at level crossings. IEEE Transactions on Intelligent Transportation Systems. 2015;16(2):596-609.

8. Amaral V, Marques F, Lourenço A, Barata J, Santana P. Laser-based obstacle detection at railway level crossings. Journal of Sensors. 2016;2016.

9. Horne D, Findley DJ, Coble DG, Rickabaugh TJ, Martin JB. Evaluation of radar vehicle detection at four quadrant gate rail crossings. Journal of Rail Transport Planning & Management. 2016;6(2):149-62.

10. Soilán M, Sánchez-Rodríguez A, del Río-Barral P, Perez-Collazo C, Arias P, Riveiro B. Review of laser scanning technologies and their applications for road and railway infrastructure monitoring. Infrastructures. 2019;4(4):58.

11. Williams K, Olsen M, Roe G, Glennie C. Synthesis of transportation applications of mobile LiDAR. Remote Sensing. 2013;5(9):4652-92.

12. Geiger A, Lenz P, Stiller C, Urtasun R. Vision meets robotics: The KITTI dataset. The International Journal of Robotics Research. 2013;32(11):1231-7.

13. Yan Z, Duckett T, Bellotto N, editors. Online learning for human classification in 3d lidar-based tracking. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2017: IEEE.

14. Transport Canada. GRADE CROSSING – HANDBOOK. In: Rail Safety Branch, editor. 2016.

15. Ogden BD. Railroad-highway grade crossing handbook. 2007.

16. Silla A, Luoma J. Effect of three countermeasures against the illegal crossing of railway tracks. Accident Analysis & Prevention. 2011;43(3):1089-94.

17. Zhang Z, Trivedi C, Liu X. Automated detection of grade-crossing-trespassing near misses based on computer vision analysis of surveillance video data. Safety Science. 2017.

18. Ohta M. Level crossings obstacle detection system using stereo cameras. Quarterly Report of RTRI. 2005;46(2):110-7.

19. Oh S-C, Lee H, editors. Performance analysis of vision based monitoring system for passenger's safety on railway platform. ICCAS 2010; 2010: IEEE.

20. García JJ, Ureña J, Mazo M, Espinosa F, Hernández Á, Losada C, et al. Sensory system for obstacle detection on high-speed lines. Transportation Research Part C: Emerging Technologies. 2010;18(4):536-53.

21. Diaz J, García JJ, Hernández A, Losada C, García E, editors. Advanced multisensorial barrier for obstacle detection. Intelligent Signal Processing, 2007 WISP 2007 IEEE International Symposium on; 2007: IEEE.

22. Hilleary TN, Omar T. A Radar Vehicle Detection System for Four-Quadrant Gate Warning Systems and Blocked Crossing Detection. 2012.

23. Tan M, Wang B, Wu Z, Wang J, Pan G. Weakly supervised metric learning for traffic sign recognition in a LIDAR-equipped vehicle. IEEE Transactions on Intelligent Transportation Systems. 2016;17(5):1415-27.

24. Yu Y, Li J, Guan H, Wang C, Yu J. Semiautomated Extraction of Street Light Poles From Mobile LiDAR Point-Clouds. IEEE Transactions on Geoscience and Remote Sensing. 2015;53(3):1374-86.

25. Yu Y, Li J, Guan H, Jia F, Wang C. Learning Hierarchical Features for Automated Extraction of Road Markings From 3-D Mobile LiDAR Point Clouds. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2015;8(2):709-26.

26. Haiyan G, Li J, Yongtao Y, Chapman M, Cheng W. Automated Road Information Extraction From Mobile Laser Scanning Data. IEEE Transactions on Intelligent Transportation Systems. 2015;16(1):194-205.

27. Wang DZ, Posner I, editors. Voting for Voting in Online Point Cloud Object Detection. Robotics: Science and Systems; 2015.

28. Frome A, Huber D, Kolluri R, Bülow T, Malik J, editors. Recognizing objects in range data using regional point descriptors. European conference on computer vision; 2004: Springer.

29. Himmelsbach M, Luettel T, Wuensche H-J, editors. Real-time object classification in 3D point clouds using point feature histograms. 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems; 2009: IEEE.

30. Hsieh H-H, Hsu C-Y, Ke P-Y, Liu G-S, Lin C-P, editors. Appling Lidar-based obstacle detection and wireless image transmission system for improving safety at level crossings. Security Technology (ICCST), 2015 International Carnahan Conference on; 2015: IEEE.

31. Hisamitsu Y, Sekimoto K, Nagata K, UEHARA M, OTA E. 3-D laser radar level crossing obstacle detection system. IHI Engineering Review. 2008;41(2):51-7.

32. Velodyne. VLP-16c LiDAR (Puck) United States, 2018 [Available from: https://velodynelidar.com/products/puck/.

33. Ye M, Xu S, Cao T, editors. Hvnet: Hybrid voxel network for lidar based 3d object detection. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2020.

34. Ester M, Kriegel H-P, Sander J, Xu X, editors. A density-based algorithm for discovering clusters in large spatial databases with noise. Kdd; 1996.

35. Golub GH, Reinsch C. Singular value decomposition and least squares solutions. Linear Algebra: Springer; 1971. p. 134-51.

36. Crouse DF. On implementing 2D rectangular assignment algorithms. IEEE Transactions on Aerospace and Electronic Systems. 2016;52(4):1679-96.

Link Between Chapters 2, 3, and 4 and Chapter 5

Chapter 2 presented a supervised 3D LiDAR-based methodology for traffic monitoring at urban intersections with high-mixed traffic. The system utilizes low-resolution and high-resolution 3D LiDAR sensors.

Chapter 3 extended the work of Chapter 2 and introduced a new method based on the shape of road users for calculating surrogate safety indicators such as time-to-collision or post-encroachment time. The road users' point cloud data in Chapter 3 was produced by low and high-resolution LiDAR sensors.

Chapter 4 presented an unsupervised 3D LiDAR-based methodology for safety monitoring at a railroad-grade crossing utilizing a low-resolution LiDAR sensor.

All previous three studies utilized 3D rotational LiDAR. Although these 3D LiDAR are powerful at 3D scanning the urban environment and quantifying road users in space, they are often costly, making them especially useful for complex scenarios such as intersection and level crossing monitoring.

Chapter 5 takes a unique direction in developing a methodology for processing 1D LiDAR data. The main objective of this chapter is to apply a low-cost, low-resolution, fast, and scalable LiDAR system for collecting cyclist traffic flow information at bike lanes, cycle tracks, and other cyclist facilities. The proposed system in Chapter 5 and those presented in Chapters 2-4 provide an opportunity for multi-model data collection in urban areas.

CHAPTER 5.

LIDAR-BASED METHODOLOGY FOR MONITORING AND COLLECTING MICROSCOPIC BICYCLE FLOW PARAMETERS ON BICYCLE FACILITIES

CHAPTER 5: A LIDAR-BASED METHODOLOGY FOR MONITORING AND COLLECTING MICROSCOPIC BICYCLE FLOW PARAMETERS ON BICYCLE FACILITIES

5.1 Abstract

Research on microscopic bicycle flow parameters (speed, headway, spacing, and density) is limited given the lack of methods to collect data in large quantities automatically. This paper introduces a novel methodology to compute bicycle flow parameters based on a LiDAR system composed of two single-beam sensors. Instantaneous mid-block raw speed for each cyclist in the traffic stream is measured using LiDAR sensor signals at seven bidirectional and three unidirectional cycling facilities. A Multilayer Perception Neural Network is proposed to improve the accuracy of speed measures. The LiDAR system computes the headway and spacing between consecutive cyclists using time-stamped detections and speed values. Estimation of density is obtained using spacing. For model calibration and testing, 101 hours of video data collected at ten mid-block sites are used. The performance of the cyclist speed estimation is evaluated by comparing it to ground truth video. When the dataset is randomly split into training and test sets, the RMSE and MAPE of the speed estimation method on the test set are 0.61m/s and 7.1%, respectively. In another scenario, when the model is trained with nine of the ten sites and tested on data from the remaining site, the RMSE and MAPE are 0.69m/s and 8.2%, respectively. Lastly, the relationships governing hourly flow rate, average speed, and estimated density are studied. The data were collected during the peak cycling season at high-flow sites in Montreal, Canada; However, none of the facilities reached or neared capacity.

Keywords: LiDAR Sensor, Microscopic Cyclist Flow Parameters, Cyclist Speed, Automated Extraction, Alternative Technologies

5.2 Introduction

The increase of bicycle usage in cities has heightened the need for improved automated data collection, metrics, and methods that provide an understanding of traffic flow characteristics, as had been done for vehicular traffic. The planning, design, and operations of bicycle facilities are essential for making urban cycling a more efficient and attractive mode of transportation (1). Planning and designing cycling infrastructure requires automated data collection technologies and methods to determine the performance of the existing network (2, 3). As bicycle ridership grows, performance indicators that evaluate traffic conditions at bicycle facilities become increasingly crucial in the planning process (4). Average Annual Daily Bicycle (AADB) traffic is among the basic performance metrics for evaluating bicycle facilities. AADB is estimated based on volume or count data. Many methods are proposed to determine the AADB using short- and long-term counts at bicycle facilities in recent years. Despite the importance of traffic volume data and AADB estimation, microscopic traffic flow parameters are required to understand cycling facilities' performance and improve traffic operations; this includes speed, spacing, headway, and density (5). The relationships governing cycling volume, speed, and density would be better understood by monitoring and automatically gathering cycling traffic flow parameters. This could help identify when facilities reach capacity, evaluate the impacts of new bicycle infrastructure, or evaluate improvements on existing facilities (6, 7). Microscopic bicycle traffic flow data could also help to calibrate microsimulation models (8) and provide information for signal timing (6, 9).

Moreover, there is currently a need for real-time cyclist microscopic traffic parameters to improve bicycle flow operations in, for instance, green-wave signalization. As bicycle ridership continues to grow and e-bikes gain prevalence, facilities will experience more speed heterogeneity. Two-wheeled facilities will transport significantly higher flows of users than conventional motorized vehicle lanes (10). Thus, the ability to monitor the performance of cycling facilities will become critical.

Knowledge of microscopic bicycle flow relationships is pertinent in the fields of transportation planning and engineering. However, few studies have attempted to develop data collection tools and methods to automatically extract basic bicycle flow parameters in real-time (11, 12). Point-based monitoring systems are standard for counting cyclists; however, none are reported to measure additional microscopic flow parameters such as instantaneous speed, density, headway, and spacing.

In response to these shortcomings in both research and practice, this research proposes a methodology to compute traffic flow parameters at bicycle facilities using distance measurements from a two single-beam LiDAR system. Based on distance measures, an algorithm to estimate instantaneous mid-block cyclist speed in real-time is proposed and evaluated using ground truth data manually calculated from video footage. The accuracy of bicycle traffic measurements is evaluated using 101 hours of video data collected at ten different cycling facilities in Montreal, Canada. Alternative regression methods are used for speed correction, including Ridge regression, Multilayer Perception Neural Network, and Decision Tree. Several features are extracted from the distance signals of a two single-beam LiDAR system and employed for tuning these three classifiers.

5.3 Literature Review

Cycling microscopic flow parameters may be more complex to model than motor-vehicle traffic parameters due to the heterogeneity of individual factors such as cyclist age and bicycle type (6, 7). In addition, external factors such as road grade, weather, and bicycle flow (13) can influence speeds.

Vehicular traffic flow analysis is widely discussed (14, 15). These studies often involve developing fundamental diagrams using flow, density, and speed variables. A traditional model, utilizing flow-density and speed-density diagrams, was first introduced by Greenshields et al. and then expanded by Aerde and Rakha (16, 17). Collecting cyclist traffic factors, including cyclist speed, allows for developing the fundamental diagrams of cyclist facilities.

Several technologies used for bicycle counting or vehicle speed measures could be adapted to measure point-based cycling speed. Few studies have used Radar to measure cyclist speed. Jeng proposed a vehicle and cyclist speed estimation system equipped with a 2D range-Doppler frequency-modulated continuous-wave radar (18). Pneumatic tubes (19) and inductive loop sensors (20) are the most common technologies used in cyclists detection and counting systems. The two tubes or loops must be placed precisely with a specific separation length to capture cyclist speed. Then the system can obtain cyclists' speed by dividing the tube separation length by the measured time between signal pulses from the two tubes. However, a low response time and a low sampling rate of pneumatic tubes limit the accuracy of speed measurements.

Using GPS trajectory and speed data, two cyclist speed studies focused on internal factors, such as gender and ability, and grouped cyclists into different categories. Berjisian investigated cyclist cruising speed by applying three time-series clustering methods on bicycle travel data collected by GPS to identify cruising, idling, acceleration, and deceleration events (7). Strauss used bicycle trip GPS data, generated from a smartphone application, to estimate individual cyclists' speed on segments and intersections and delay at intersections in Montreal (4). Collecting GPS trajectories data could enable researchers to study the fundamental diagrams of traffic flow for cyclists (11, 12, 21-24). Although GPS trajectory data has the advantage of monitoring cyclists across the network, it has a couple of important disadvantages. Firstly, penetration rates are very low; thus, only a few cycling speeds are measured. Secondly, because of the low sampling rate, these technologies are not well-suited for real-time applications.

Video-based systems use image processing or artificial intelligence algorithms to detect road users and track their trajectories. An estimate of road user speed is calculated from the time-stamped trajectory data. This process can be adapted to monitor cyclists and measure cyclists' speed. Figliozzi developed and applied a methodology to manually estimate bicyclist acceleration and speed from video data for traffic signal timing applications (6). Zaki demonstrated that video analysis for bicycle data collection in high-density environments is feasible (25, 26). Bicycle counts, density, and speed are automatically obtained in a bicycle facility, and results are validated using manual methods from video footage. In a previous study, Zaki validated the speed of 70 cyclists with an average speed of 3.9 m/s, and found RMSE of 0.38 m/s or approximately 10% (26).

Hoogendoorn proposed a video-based composite headway modeling system for cyclists traversing a crosswalk (27). However, the classification of the crossing road users at a crosswalk into cyclists and pedestrians is not discussed. In recent work, Mohammed et al. studied cyclist maneuvers, such as following and overtaking, using trajectories extracted from video footage, and utilized cyclist interactions for microsimulation modeling (28). Their method obtains 95% accuracy in retrieving cyclist trajectories and shows promising outcomes in maneuver analysis, counting, and speed estimation.

Recent video-based systems have demonstrated encouraging outcomes in collecting cyclist flow parameters. However, as the complexity of video-based trajectories analysis increases, the efficiency and cost of the system and the scalability and reproducibility of the process limit the applicability of this approach. A camera-based system may require a complete calibration process when installed in a new facility; in some cases, the calibration must be completed on the video processing module. In general, the efficacy of a video-based system for cyclist monitoring applications is affected by several criteria such as budget, coverage area, hours of operation (daytime/nighttime), etc. A highly accurate camera-based system performing in real-time requires powerful CPU/GPU modules, data storage, and transmission units. These components add high cost in a large-scale implementation.

Although recent advances are made in bicyclist data collection, some research gaps emerge from the literature. Firstly, few studies have explored technologies and methods to automate the measurement of cyclist speed in real-time on cycling facilities. Secondly, only one known project (25, 26, 28) has validated speed measures using video-based systems. Lastly, accurate video-based systems operating in real-time are expensive and require calibration at each site.

Pedestrian and bicycle counting systems that use emerging laser technology, Light Detection and Ranging (LiDAR), have shown promising results. A pedestrian counting system, designed and proposed in our recent research work (29), can collect pedestrian data on different types of facilities such as sidewalks or crosswalks. Based on this past work, there is potential for LiDAR sensors to be used for bicycle traffic monitoring and data collection of microscopic parameters. LiDAR systems are developed to classify and count pedestrians or cyclists (29, 30); however, no known method is applied to measure cyclist speed.

LiDAR systems, specifically 1D and 2D LiDAR, have some potential advantages over video analysis. Firstly, the amount of data at each sampling time is lower than with video, and the algorithms employed for processing these data are less complicated. Secondly, for on-site, real-time applications, the embedded processor would require significantly less computational power than with video and thus would be considerably cheaper. Therefore, an integrated LiDAR system for counting and speed monitoring can be more competitive than video-based systems. A monitoring system based on 1D LiDAR requires less energy, is less costly than video-based systems, and can be implemented for real-time applications. A real-time system would eliminate the cost of transmitting and storing data for offline processing. Thus, the system could be installed permanently with a relatively low operating cost. Additionally, real-time bike speed and density information could help optimize flow and ensure green wave operation using smart signal actuation and timing.

Vehicle traffic flow relationships are well understood. An automated roadside system for measuring cyclist traffic flow parameters in real-time would allow for an increased understanding of cycling speed, flow, and capacity on cycling facilities. It would enable optimization of traffic signal timing on high-flow urban bicycle facilities.

5.4 Methodology

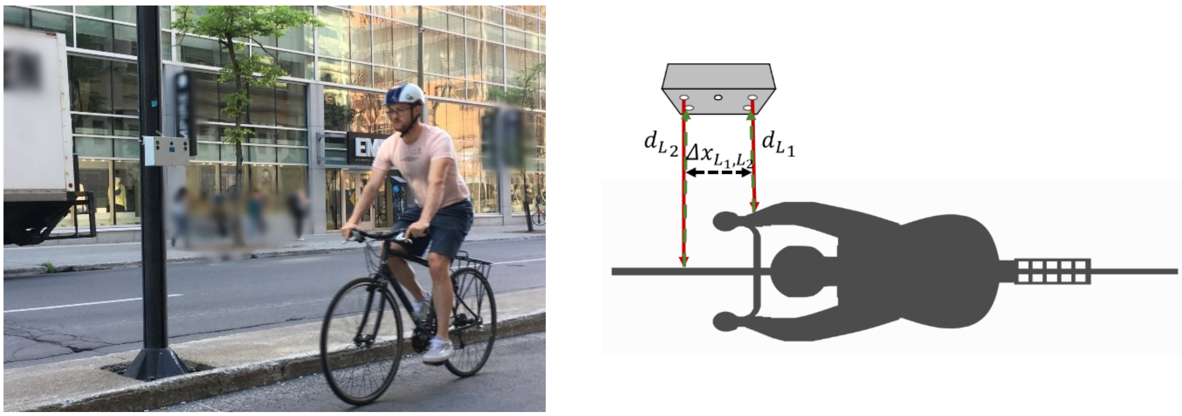
To achieve the objectives of this research, a methodology consisting of four steps is proposed:

1. LiDAR system development and deployment,
2. Cyclist detection,
3. Computation of bicycle speeds, error correction, and speed validation,
4. Estimation of headway, spacing, and density.

Additional details on each of these steps are provided below.

5.4.1 LiDAR system development and deployment

As a first step, a LiDAR system is developed with two single-beam LiDAR sensors. The system hardware consists of a processor, a low-power ARM microcontroller with a 1GHz processor and a 1GB RAM, and two single-beam LiDAR sensors from LidarLite. The LiDAR sensor measures distance up to 40m with a resolution of 3cm and a sampling rate of 500Hz (2ms). All the components are placed in a waterproof enclosure and powered by a 12V-2A battery pack. The two 1D LiDAR sensors are spaced by a given distance (e.g., 0.20m), which allows for speed measurement of passing cyclists. **Figure 5-1 (a)** shows a sample installation where the 1D LiDAR system is installed at a cycle track in Montreal. **Figure 5-1 (b)** shows the top view of the system operation. The two single-beam lasers emit two signals (red arrows), L_1 and L_2, and the detector of each laser receives the reflected signals (green arrows) from cyclists, then the laser scanners calculate the distance to the cyclists, d_{L_1} and d_{L_2}, from their signal's time of flight. The distance between two single-beam LiDAR, which is also the distance between two emitted light beams or two reflected light beams, is called $\Delta x_{L_1,L_2}$ (equivalent to 0.20m). Additionally, the installation height of the system (with respect to the pavement) is h and ranges from 1.1m to 1.3m.



a) Sample Installation **b) Top view of the system operation**

Figure 5-1 System setup in an actual installation

5.4.2 Cyclist detection

The output from each of the LiDAR sensors is a pulsed distance signal. If the sensor is facing a solid object (e.g., a wall), then the base value of the distance signal (background value) is a fixed value equal to the distance from the sensor to the wall. Alternatively, if the sensor faces an open space or an object farther than its maximum distance range, the base value is 0.

For each of the two sensors, a lower band (d_{min} – distance to the nearest side of the bike path) and an upper band (d_{max} – distance to the farthest side of the bike path) distance values are defined. The value of the distance signal (d_{L_i}; $i \in \{1,2\}$) is set to a pre-defined distance value when it is not within these bands. The pre-defined distance value, d_{base}, is lower than d_{min}:

$$d_{L_i} \leftarrow d_{base}, when: d_{L_i} \leq d_{min} \; or \; d_{L_i} \geq d_{max} \quad (5\text{-}1)$$

This logic ensures that the signal value of d_{L_1}or d_{L_2} is either d_{base} or equivalent to the distance to a road user moving in front of the system on the bike path. The pulsed distance signal has a dynamic duty cycle that depends on the headway between the cyclists. The width and amplitude of each pulse could indicate the speed and distance from the sensor to the corresponding cyclist. The amplitude series of each pulse that represents the distance to the cyclist is not a unique value and has a small and bounded range that indicate the un-even surface of the cyclist's upper body.

Using pulses of distance signals, a multi-step clustering algorithm is implemented for cyclist detection and counting. In the first step, a binary signal ($m_{L_1 \,\&\, L_2}$), indicating the presence of a cyclist in front of either of the LiDAR sensors, is generated by using:

$$m_{L_1 \,\&\, L_2} \leftarrow 1, when\colon d_{L_1} > d_{base}\ or\ d_{L_2} > d_{base}; m_{L_1 \,\&\, L_2} \leftarrow 0\colon otherwise \tag{5-2}$$

$m_{L_1 \,\&\, L_2}$ is 1 when at least one of the sensors reads a distance value other than the base distance. Using this time-stamped binary signal, each pulse in $m_{L_1 \,\&\, L_2}$ is separated as a candidate cluster for the second step, where each cluster is evaluated to be split into smaller clusters, unchanged or removed.

The size of each cluster is equal to the width of the corresponding pulse obtained from $m_{L_1 \,\&\, L_2}$. Noises of distance measurements, corresponding to the cases where LiDARs falsely report a distance value while there is no object in front of them, are eliminated by filtering out the cluster with a small size. Since the installation height of the system ranges from 1.1m to 1.3m, the laser beams measure the distance to the upper body of each cyclist. Assuming that the minimum width of a cyclist's upper body is 15cm, each cyclist needs to travel the line of sight of two LiDARs (2×15cm) and the distance between two LiDARs (20cm) to pass both sensors. A cyclist traveling quickly (i.e., 10m/s) will spend approximately 50ms in the detection area and will therefore be detected and measured in about 25 samples. Thus, a lower threshold for cyclist cluster removal is set as 15 samples.

Afterward, the distance signal, d_{i,L_j}, corresponding to the $i-th$ cluster in $m_{L_1 \,\&\, L_2}$, are extracted from $j-th$ LiDAR, and the binary signal, m_{i,L_j} is computed as:

$$m_{i,L_j} \leftarrow 1, when\colon d_{i,L_j} > d_{base}; m_{i,L_j} \leftarrow 0\colon otherwise \tag{5-3}$$

If there is only one pulse in m_{i,L_j}, it will be considered as one cyclist. If there is more than one pulse, a set of rules is applied to split the pulses into multiple clusters or merge them as one cluster. First, if the difference in the average amplitude of the pulses in a single cluster in m_{i,L_j} is more than 0.5m, the cluster will be split into multiple pulses of similar amplitude within each pulse.

Second, for all the pulses in one cluster, the duty cycle of pulses (time gap between pulses) is used for the merging-splitting process. A threshold of 25 samples between pulses is chosen as a reasonable criterion for the merging of clusters. If two cyclists with a headway of 1 meter are traveling with a speed of 20m/s, then the time difference in their pulses will be 1/20=0.05s (equivalent to 25 samples). This scenario represents an extreme situation where two unique cyclists are traveling with very little headway at an extremely high speed. Therefore, if two pulses are within 25 samples of each other, they will be merged as one cluster.

Afterward, clusters of both sensors are matched based on their appearance, chronology, and pulse width. Finally, a cyclist sample with one cluster (distance pulse) per LiDAR is detected and counted (see **Figure 5-2**).

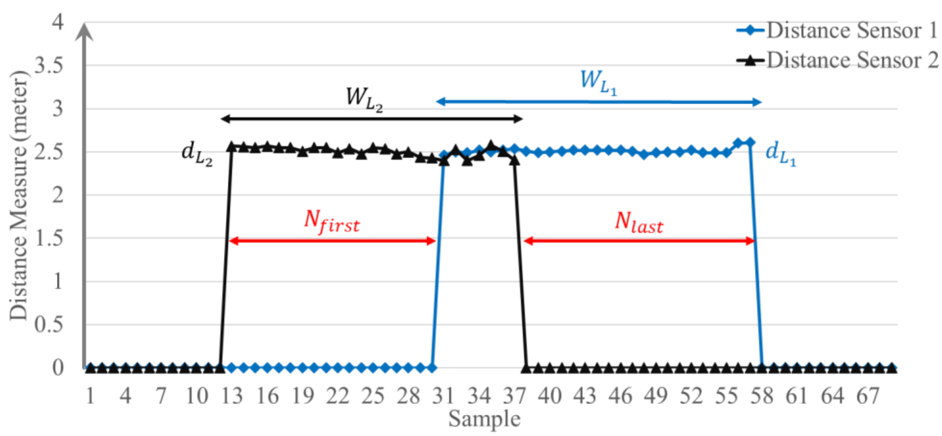


Figure 5-2 Sample of sensor distance measures (in meters) for a single cyclist

5.4.3 Computation of bicycle speed

Figure 5-2 illustrates the distance signals from the two LiDAR sensors when a cyclist is traversing the coverage area and detected by both sensors. In the case when no cyclist is present, the output of the sensor is set to zero (d_{base}). When a cyclist appears, the output corresponds to the distance to that cyclist. The x-axis represents consecutive samples in time, and the y-axis represents the distance from the LiDAR system to the cyclist.

In this sample plot, the cyclist is first detected by sensor 2. The order of detection of the two sensors determines the cyclist's direction and the sign of its velocity vector. The theoretical cyclist speed can be measured by (5-4).

$$\widehat{U}_{raw} = \frac{\Delta x_{L_1,L_2}}{\Delta T_L} = \frac{0.2}{\Delta T_L}\left(\frac{m}{s}\right) \quad (5\text{-}4)$$

where:

- $\widehat{U}_{raw}$ is the raw speed that the LiDAR system can measure.
- $\Delta x_{L_1,L_2}$ is the distance between the two sensors, which is fixed and equal to 0.2m. Note that the spacing between the two sensors is kept low to build a relatively compact system that can easily be installed everywhere. A spacing significantly smaller than 0.2m would reduce the accuracy of extracting features from the distance signals.
- ΔT_L is the difference in time between the two sensors detecting the cyclist. The time difference (ΔT_L) is computed by multiplying the number of samples between two sensor detections by the sampling time of the sensors. The nominal sampling rate of the LiDAR sensors is 500Hz, meaning that the standard sampling time (T_s) is two milliseconds.
- N_{first} is defined as the difference in samples between the first detection of the cyclist by 1st and 2nd sensors (see **Figure 5-2**). Similarly, N_{last} is defined as the difference in samples between the last detection of the cyclist by 1st and 2nd sensors, respectively.

Ideally, N_{first} should be equal to N_{last}, and $\Delta T_L = N_{first} \times T_s$, where T_s is the sampling time; therefore, $\widehat{U}_{raw}$ is obtained by Equation (5-4). However, as the laser beams diverge, the distance between the two beams differ from 0.2m, and ΔT_L may not precisely correspond to a 0.2m displacement. The error associated with an inaccurate ΔT_L can lead to speed estimation error. Equation (5-1) provides a raw speed measurement. In the following step, alternative error correction techniques are applied to the raw speed measure.

5.4.4 Error correction and speed validation

Alternative regression speed estimation methods with multiple regressors are proposed to improve raw measurements' accuracy by correcting minor uncertainties in measurements. First, the factors used to correct the error are defined, including N_{first} and N_{last}, defined before, as well as the other regressors extracted from the two distance signals (d_{L_1} and d_{L_2}):

- For each cyclist, $\bar{d}_{L_1}$ and $\bar{d}_{L_2}$ represent the average magnitude of the signal or distance value from the LiDAR sensor 1 or 2 to that cyclist (as illustrated in **Figure 5-2**). The slight variations in distance measurement are related to the uncertainty caused by beam divergence, as cyclists closer to the sensor are less affected. Therefore, taking the average distance of the cyclist into account could reduce the speed estimation error.
- For each cyclist, W_{L_1} and W_{L_2} are the width of the distance signals recorded by LiDAR sensors 1 and 2, respectively. These values represent the number of samples for which each sensor has observed that cyclist. The faster the cyclist passes the coverage area, the smaller the number of samples is recorded by the LiDAR sensor. The cyclist's body shape also affects W_{L_i}, therefore the estimated speed by video validation ($\widehat{U}_V$) and W_{L_i} are not expected to be highly correlated. Nonetheless, the signal widths are considered in the regression model.

Since the proposed system uses two identical 1D LiDAR sensors placed very close to each other, a cyclist should have similar patterns in both. However, slight variations in the signal pattern could occur. Nonetheless, $\bar{d}_{L_1}$ is highly correlated with $\bar{d}_{L_2}$ (99.7% correlation) and W_{L_1} with W_{L_2} (96.7% correlation). Including each of these pairs makes the regression model inconsistent. Therefore, the average of each pair will be used, which are $\overline{W}_L = (W_{L_1} + W_{L_2})/2$ and $\bar{d}_L = \left(\bar{d}_{L_1} + \bar{d}_{L_2}\right)/2$.

In this study, two different approaches are used to build the regression model. In the first approach, the direct model, the estimated speed from video validation is the dependent variable of the regression model. Since the cyclist speed has an inverse correlation with the time difference (N_{first} or N_{last}), the inverse of these factors is included in the regression model. The regression function of the direct model is written as (5-5):

$$\widehat{U}_L = f\left(N_{first}^{-1}, N_{last}^{-1}, \bar{d}_L, \overline{W}_L \middle| \Theta\right) \tag{5-5}$$

where Θ is the model coefficients which are estimated using the manually validated data. In the second approach, the validated video time difference is the dependent variable instead of the validated speed. Consequently, N_{first} and N_{last} are included in the regressor vector instead of their inverse. The regression function to estimate the time difference is written as (5-6):

$$\widehat{\Delta T}_V = f\left(N_{first}, N_{last}, \bar{d}_L, \overline{W}_L^{-1} \middle| \Theta\right) \tag{5-6}$$

The estimated speed of the cyclist ($\widehat{U}_L$) has an inverse relationship with the dependent variable of the regression model ($\widehat{\Delta T}_V$), therefore, the second model is called the inverse model. The estimated speed is obtained by (5-7).

$$\widehat{U}_L = \frac{d_V}{\widehat{\Delta T}_V} = \frac{7m}{\widehat{\Delta T}_V} \tag{5-7}$$

In this study, two identical LiDAR-based systems are prepared and installed in different locations on different days. The bike facilities used for testing include bidirectional cycle tracks, bidirectional and unidirectional painted bike lanes. A portion of each dataset is manually validated by analyzing the video footage. Then all the validated data are merged to create a comprehensive cyclist dataset. To consider the effect of the bike facilities that each sample is taken from, a fixed effect vector containing dummy variables is used as (5-8):

$$\alpha_k = \begin{bmatrix} \alpha_{k,1} \\ \alpha_{k,2} \\ \vdots \\ \alpha_{k,m-1} \end{bmatrix}_{(m-1)\times 1} \tag{5-8}$$

where k is the sample number and m is the number of bike facilities contained within the dataset. If the k^{th} sample of the dataset belongs to the j^{th} bike facility, then α_{jk} is set as one; otherwise, it is zero. In fact, α_{jk} is a dummy variable that indicates which sample belongs to which site.

In total, there are $m - 1$ dummy variables: when all the dummy variables are zero, it implies that the sample belongs to m^{th} site. The regressors' vector is written as (5-9). The dimension of this vector is $5 + (m - 1)$.

$$x = \begin{bmatrix} 1 \\ g_1(N_{first}) \\ g_2(N_{last}) \\ g_3(\bar{d}_L) \\ g_4(\bar{W}_L) \\ \alpha_k \end{bmatrix}_{(5+m-1)\times N} \tag{5-9}$$

where g_i depends on the type of model in use which could be direct or inverse. For example, if the direct model is used, then $g_1(\cdot) = g_2(\cdot) = (\cdot)^{-1}$, and g_3 and g_4 are identity functions, and if the inverse approach is used, $g_3(\cdot) = (\cdot)^{-1}$ and g_1, g_2 and g_3 are identity functions. Different regression models, including Ridge Regression, Multilayer Perception (a feedforward artificial neural network), and Decision Tree, are implemented in this study. The hyperparameters of each of these models are tuned by a k-fold cross-validation technique.

5.4.5 Estimation of headway, spacing, and density

Additional traffic parameters such as flow rate, headway, spacing, and density are computed using time-stamped information and the cyclists' speed. The flow rate, q, is obtained by summarizing the count results in different time intervals (5-10).

$$q = \frac{m}{\Delta T} \tag{5-10}$$

where m is the flow or number of cyclists that pass in front of the LiDAR detector at the fixed position (installed mid-block) during the time interval of ΔT in hours.

The headway between any two consecutive cyclists is computed simply by calculating the difference between the two detection times. In addition to the flow rate and headway, the density, k, can be calculated as (5-11):

$$k = \frac{n}{L} \tag{5-11}$$

where n is the number of cyclists traveling along a bike facility segment of length L at a specific time. Equation (5-11) estimates density as the number of cyclists per unit of length, typically represented as cyclists/km. However, accurately calculating density from (5-11) requires continuous spatial monitoring of a relatively long section of the bike path and counting the number of cyclists in that section. Choosing small values of L causes high temporal variations in n as the number of cyclists traveling along a short segment will vary widely for different time intervals. Alternatively, choosing a large value for L requires a monitoring system with a large field-of-view or coverage area, likely a system with multiple cameras. The cost of such a system may be prohibitively high. Instead of calculating the density by counting cyclists along a segment of length L, the density can be estimated by measuring the space between two consecutive cyclists as (5-12):

$$k = \frac{1}{\bar{S}} \tag{5-12}$$

where $\bar{S}$ is the average spacing between every two consecutive cyclists in any time interval: $\bar{S} = \sum_{i=1}^{m} S_i / m$. Here, $\sum_{i=1}^{m} S_i$ is the summation of the spacing between m cyclists traveling along the bike path at a specific time interval (ΔT). The proposed system is adapted to estimate spacing between two cyclists using headway and estimated speed information.

Figure 5-3 shows a scenario in which three cyclists are passing through the sensor line-of-sight. The leading cyclist (*L*) has crossed the line-of-sight at time t_L and with speed U_L, the current cyclist (*C*), the middle one, is passing the sightline at time t_C and with speed U_C, and the following cyclist (*F*) will pass it at time t_F and with speed U_F. The headway or gap time is obtainable by computing the time difference of every two successive cyclist's observations in the same direction. In this case, the headway to the leading cyclist is $\Delta t_L = |t_C - t_L|$ and the headway to the following cyclist is $\Delta t_F = |t_F - t_C|$.

The spacing is another important factor that can represent level-of-service in a bike facility. The information of the headway and speed of the cyclists are used to extract the spacing of two consecutive cyclists. When the current cyclist arrived at the line-of-sight, the leading cyclist has traveled for Δt_L seconds with the speed of U_L m/s. Therefore, the spacing between the current cyclist and the leading cyclist at the time of current cyclist observation is $S_{C,L}^{(1)} = \Delta t_L \times U_L$.This is with respect to the time that the current cyclist is observed. On the other hand, when the leading cyclist is detected, it is expected that the current cyclist arrived after Δt_L seconds while traveling with the speed of U_C m/s. Therefore, the spacing between the current cyclist and the leading cyclist is $S_{C,L}^{(2)} = \Delta t_L \times U_C$. The actual value of the spacing is between $S_{C,L}^{(1)}$ and $S_{C,L}^{(2)}$. The best value that can represent this spacing is the average of these two values, $S_{C,L} = (S_{C,L}^{(1)} + S_{C,L}^{(2)})/2 = (\Delta t_L \times U_L + \Delta t_L \times U_C)/2 = \Delta t_L \times (U_L + U_C)/2$. Similarly, the spacing between the current cyclist and the following cyclist is computed by $S_{C,F} = \Delta t_F \times (U_C + U_F)/2$. The spacing between the current and the leading cyclists ($S_{C,L}$) are used to compute $\overline{S}$ in (5-12).

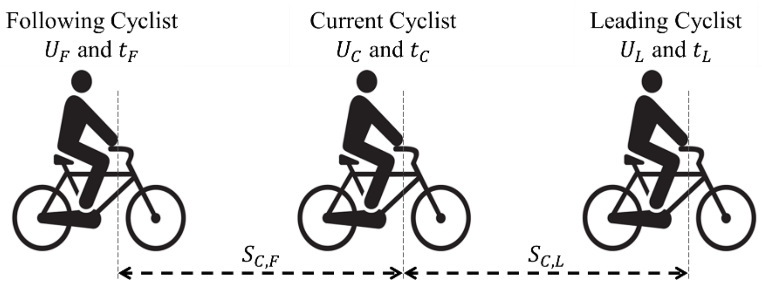


Figure 5-3 Three consecutive cyclists traveling in the same direction

Since the proposed LiDAR system, composed of two single-beam sensors, can detect cyclists' direction, the traffic characteristics, including flow rate, density, speed, headway, and spacing, are also obtained per direction.

5.4.6 Ground-truth speed data generation

Manual estimation of the cyclist speed is evaluated and compared against the automated measurements. The LiDAR system is equipped with a camera (placed between two single-beam LiDARs). Video footage, used for manual estimation of the ground-true speeds, is collected simultaneously as the LiDAR data. The manual speed is an average speed along a section defined by two screen lines marked on the surface with white chalk (see **Figure 5-4 (a)**). The chalk is drawn when the LiDAR and video systems are installed. The length between the two screen lines is 7.0m. The time it took for all the cyclists to travel across the 7.0m distance is timed by manually counting the number of video frames from the frame when the front wheel reaches the first line to the frame when the front wheel reaches the second line. The manual speed is obtained from (5-13):

$$\widehat{U}_V = \frac{d_V}{\Delta T_V} = \frac{d_V}{\Delta f_V / fps} \qquad (5\text{-}13)$$

where:

- d_V is the distance traveled in the video footage (in this case, 7.0m),
- ΔT_V is the time required for the cyclist to travel this distance,
- Δf_V is the number of frames that the camera captures during ΔT_V,
- fp$s$ is the frame rate of the camera (in this case, 30 frames per second).

In some cases, due to the proximity of the sensor installation to the bike lane, using the 7-meters lines is not possible; therefore, 6-meters and 5-meters lines are also drawn for manual speed estimation (see **Figure 5-4 (b)**).

(a) Boulevard de Maisonneuve

(b) University Street

Figure 5-4 Sample snapshots of the sites and the manual speed estimation

Manual validation has inherent accuracy limitations. Manual counting of the number of frames that the camera captures during the ΔT_V can introduce a small error in the ground truth speed measure. For example, when the cyclist's speed is 5m/s, it takes 1.4s or 42 frames to cross both lines. By undercounting one video frame ($\Delta f_V = 41$), the estimated speed will be 5.12m/s, an absolute percentage difference of 2.4%. Over-counting one video frame ($\Delta f_V = 43$) would result in an absolute percentage difference of -2.3%. As most cyclists cross the drawn white lines between two video frames, each frame count, for validation purposes, is subject to at most one frame of error. A quality control procedure is defined and performed for manual video validation.

5.5 Evaluation of System Performance

5.5.1 Data collection and validation

Cyclists are monitored at ten different sites over a ten-day in 2018, using two time-synchronized systems: LiDAR and camera. All study sites are located mid-block along a bike facility. Most sites have high bicycle flow; during peak periods, total bicycle flow can exceed 1000 cyclists/hour on some bidirectional facilities. **Table 5-1** provides descriptive statistics from these sites. The data collection started early in the morning, between 7 am to 8 am on each day.

For manual speed validation using video footage, a computer software is designed to play video frame-by-frame at a slow frame rate. The user can identify the exact time a cyclist arrives at the coverage area, demarcated by white lines drawn on the pavement (see **Figure 5-4**). For each day, a set of samples are chosen for validation. Then for each sample, the video is analyzed by the user. Finally, and for quality assessment of the manual validation, some of the validated speeds are randomly selected, and another user verifies their results through the same procedure. In case of a mismatch, the samples of that bike facility are re-validated and modified.

Table 5-1 summarizes the descriptive statistics of the collected and chosen data. In 101 hours of data collection, the LiDAR system detects 28,053 cyclists. In total, 2,385 samples are validated from ten bike facilities. The average cyclist speed for each site varies from 4.6m/s to 6.2m/s. Three types of bicycle-dedicated sites are analyzed. Of these ten sites, six are bidirectional cycle tracks (type 1 – protected bike lanes that are separated from vehicular traffic and pedestrians by raised median), one is a bidirectional bike lane (type 2 – separated from vehicular traffic and pedestrians by pavement marking), and three are unidirectional bike lanes (type 3 – separated from vehicular traffic and pedestrians by pavement marking). In all these bicycle facilities, pedestrians are not allowed.

For the period of validation analysis of each site, the number of LiDAR detections is obtained from the proposed system. Then manually validated samples are matched with automated LiDAR detections one by one. This process helps to find various error factors such as overcount, undercount, and correct detections that are used to compute *Recall* (5-14), *Precision* (5-15), and *F1-Score* (5-16) for each site.

$$Recall\ (accuracy) = \frac{Correct\ Detections}{Manually\ Validated\ Samples} \tag{5-14}$$

$$Precision = \frac{Correct\ Detections}{LiDAR\ Detections} \tag{5-15}$$

$$F1 - Score = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{5-16}$$

Recall (accuracy) rate of the LiDAR system varies from 94.3% to 99.6%, *Precision* varies from 95.8% to 100%, and *F1-Score* varies from 95.3% to 99.8%. The weighted average values of *Recall*, *Precision*, and *F1-Score* are 97.3%, 98.4%, and 97.8%, respectively.

Table 5-1 Summary of descriptive statistics and validation results of the evaluation sites

Site ID	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Site Type	1	1	1	1	1	2	3	3	3	1
Duration (hour)	4.5	1.9	12.0	12.5	9.0	12.0	12.5	10.6	12.2	13.5
Total Cyclists (LiDAR)	928	740	6717	3939	2470	3477	1458	1676	2731	3917
Average Hourly Flow	207	393	560	316	275	290	117	157	224	290
Video Screen Line Space (m)	7	7	5	5	6	6	5	6	7	7
Manually Validated Samples	278	220	286	269	169	291	87	235	264	286
Average Speed (m/s)	5.93	5.87	5.5	5.8	4.6	6.2	5.9	5.6	4.6	5.3
Average Speed (km/hr)	21.4	21.1	19.6	20.8	16.6	22.5	21.4	20.3	16.7	18.9
Standard Dev. Speed (m/s)	1.1	1.1	1.0	1.4	1.0	1.0	0.9	1.2	0.9	1.3
LiDAR Detections (Validation)	277	217	282	260	166	289	85	236	257	288
Over Counts	0	2	5	6	2	2	3	10	2	5
Under Counts	1	5	9	15	5	4	5	9	9	3
Correct Detections	277	215	277	254	164	287	82	226	255	283
Positive Direction Samples	247	71	118	63	59	50	0	0	0	165
Negative Direction Samples	30	144	159	191	105	237	82	226	255	118
Direction Error	0	0	2	0	0	0	0	0	0	0
Recall (%)	99.6	97.7	96.9	94.4	97.0	98.6	94.3	96.2	96.6	99.0
Precision (%)	100.0	99.1	98.2	97.7	98.8	99.3	96.5	95.8	99.2	98.3
F1-Score (%)	99.8	98.4	97.5	96.0	97.9	99.0	95.3	96.0	97.9	98.6

Types of the Cyclist Facility:
1: Cycle Track, Bidirectional; 2: Bike Lane, Bidirectional; 3: Bike Lane, Unidirectional

False-positive errors could be road users, such as pedestrians, skateboard riders, scooter riders, or a person who uses a wheelchair traveling on a bike path. False-negative errors are short cyclists, a cyclist blocked in an overtaking maneuver, or system failure in clustering a group of cyclists.

5.5.2 Testing scenarios for speed modeling

Two different scenarios are performed to evaluate the system performance. In the first evaluation scenario, the dataset is randomly partitioned into training, validation, and test sets. Of 2,320 validated samples, 25 percent are chosen as the test set before any modeling and are only revealed when the system performance is tested in terms of estimation accuracy over an unseen dataset. In the second scenario, the regression model is built and tested under ten different conditions in a leave-one-out procedure wherein each stage, the whole dataset of one of the sites is selected as the test set, and the data of the nine other sites are merged and used for tuning and fitting the regression models.

In the first scenario, since all the bike facilities have samples in the training and test sets, the fixed effect factor regression model with dummy variables is implemented. The system is tested to determine how the regression model will perform on the test samples while it is tuned with different training samples from the same facilities. On the other hand, in the second scenario, the regular regression model without fixed factors is implemented because there is only data from one site in the test set, and the site's characteristics of the test set are not seen in the training stage. In this case, the system is tested to determine how the system will perform when installed on a new bike facility when the regression model is built by data from several other sites.

The fitting data, 75 percent of the total in the 1st scenario or nine out of the ten sites in the 2nd scenario, is split into training and validation. The training set is used to fit the optimal coefficients, and the validation set is used to tune the hyperparameters of each regression model. The Ridge regression hyperparameter is λ, the MLP regression hyperparameters are the learning rate and the number of neurons in the hidden layer, and the Decision Tree has many hyperparameters of which the most important is the tree depth. A five-fold cross-validation technique is used to split the fitting data into the training and validation sets to minimize random sample selection bias for the training and validation sets.

Table 5-2 shows the results of each regression model implemented with its tuned hyperparameters in the first scenario. The λ in the Ridge regression is tuned to 1 and 0.01 in inverse and direct approaches, respectively. The number of neurons in the hidden layer is set to 3 and 4 in inverse and direct approaches, and the learning rate is tuned to 0.0001 in both. The maximum depth of the decision trees in both methods is set to 5.

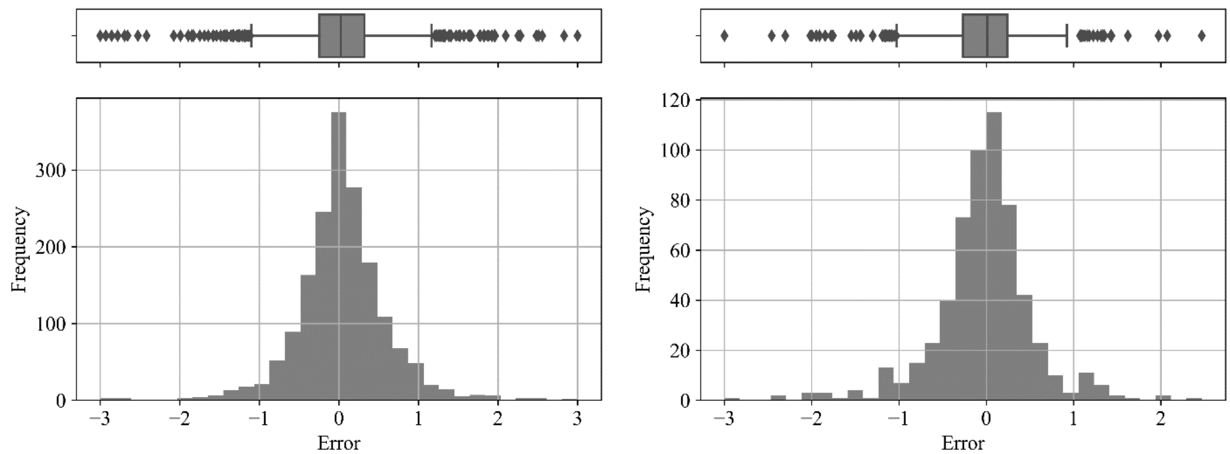
Table 5-2 Error measures for the first evaluation scenario

Regression Method	**Regression Models**	**RMSE (m/s)**		**MAE (m/s)**		**MAPE (%)**		**R-Square**	
		Train	Test	Train	Test	Train	Test	Train	Test
Inverse Approach	Ridge Regression	0.58	0.58	0.4	0.39	7.1	7.0	0.71	0.72
	MLP Regression	0.57	0.57	0.39	0.38	7.0	6.9	0.72	0.73
	Decision Tree	0.59	0.58	0.41	0.42	7.2	7.6	0.67	0.67
Direct Approach	Ridge Regression	0.59	0.62	0.42	0.42	7.6	7.7	0.68	0.67
	MLP Regression	0.58	0.63	0.41	0.4	7.3	7.4	0.69	0.64
	Decision Tree	0.56	0.6	0.41	0.43	7.4	7.8	0.72	0.68
Theoretical Cyclist Speed in Equation (5-4)		1.43		0.77		10.9		-	

MLP regression in the inverse approach, where the travel time is the predicted variable, outperforms other regression models. The Root Mean Square Error (RMSE) of this model is 0.57m/s, the Mean Absolute Error (MAE) is 0.38m/s, and the Mean Absolute Percentage Error (MAPE) is 6.9%. A cyclist traveling at 5m/s will have an MAE of 0.34m/s, and a cyclist traveling at 10m/s will have an MAE of 0.69m/s.

The RMSE, MAE, and MAPE of the theoretical cyclist speed (Equation (5-4)) are computed and given in the last row. The MLP regression method in the inverse approach improve RMSE by 0.86m/s (3.1km/hr), MAE by 0.39m/s (1.4km/hr), and MAPE by 4%.

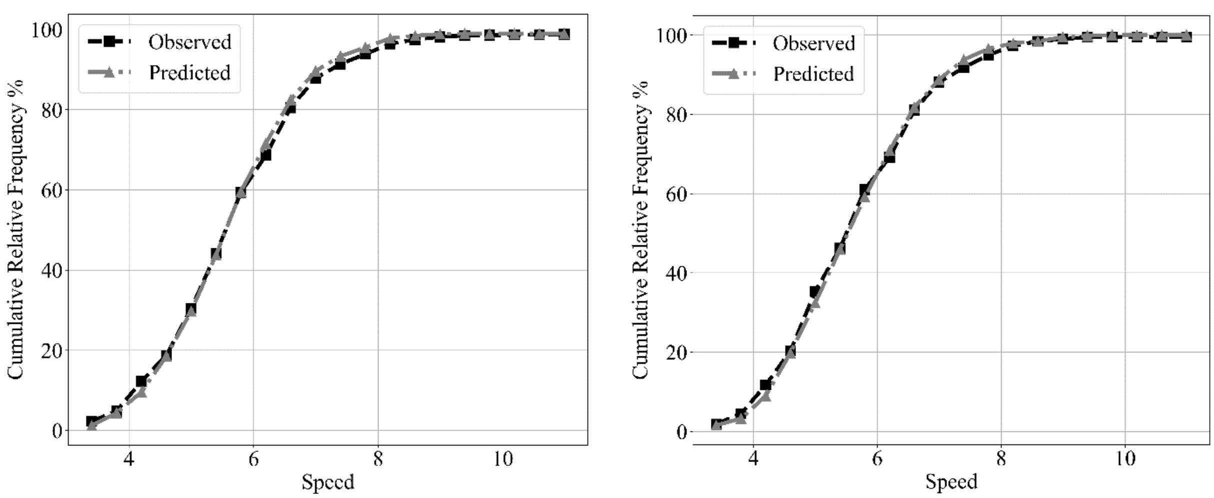
Figure 5-5 illustrates the histogram of speed estimation error (residual) on the training and test sets for the MLP regression model using the inverse approach. The center of the histogram is at zero, and most of the residuals are within -0.5m/s and +0.5m/s. The positive error value shows that the estimated speed is greater than the observed speed and vice versa. The average residuals of the training and test sets are 0.04m/s and -0.04m/s, respectively, and the standard deviation is 0.57m/s which is equal to the RMSE value of the training and test set respectively in **Table 5-2**. The box plots in **Figure 5-5**. show that few samples have an absolute error greater than 1m/s.



(a) Training set

(b) Test set

Figure 5-5 Histogram of error



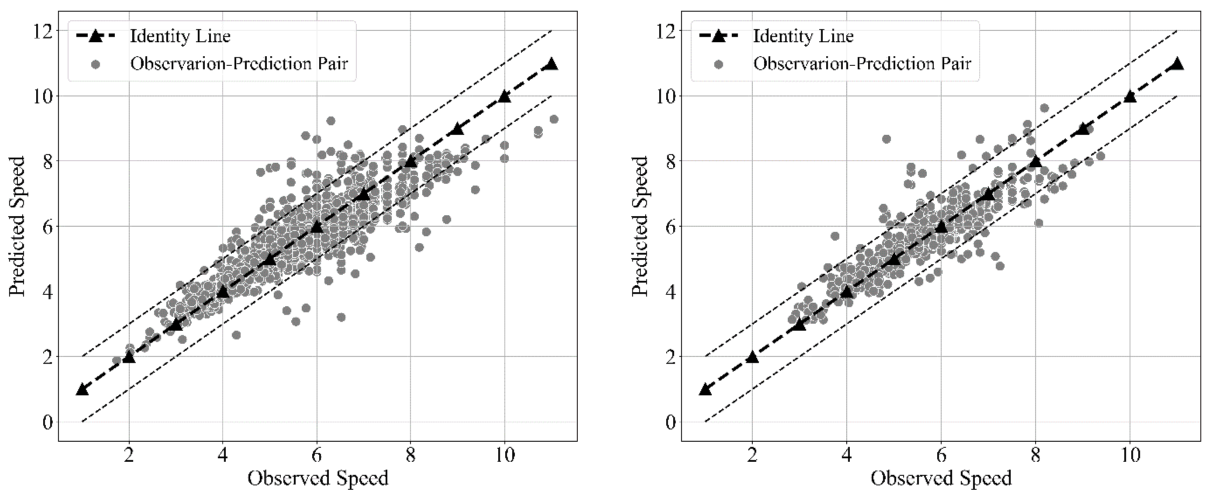
(a) Training set

(b) Test set

Figure 5-6 Aggregate level cumulative relative frequency of actual and estimated speed values

Figure 5-6 illustrates the cumulative relative frequency of observed speed (ground truth obtained by manually validating the video footages) and estimated speed (obtained by MLP regression model) for the training and test sets. The two plots indicate that the observed and estimated speed distributions are similar: the observed and predicted cumulative frequency diagrams are almost superimposed. For speeds less than 5m/s, the observed speed is greater than the estimated (the observed curve is to the left side of the predicted curve). The region bounded by the two plots corresponds to the negative part of the histogram error from **Figure 5-5**. For the speeds above 6m/s, the predicted speed is greater than the observed speed on average. This region corresponds to the positive part of the histogram error.

The disaggregated performance of the MLP regression model for both training and test sets is illustrated in **Figure 5-7**. The scatter plots are oriented around the identity line. Note that few samples are more than 1m/s from the identity line. The R-Square values for training and test sets are 0.72 and 0.73, respectively (see **Table 5-2**).



(a) Training set (b) Test set

Figure 5-7 Estimated and actual speed for both training and test sets

In the second scenario, the modeling and testing are repeated ten times, each time one of the sites represents the test set and is entirely excluded from the training set. This scenario helps determine the LiDAR system's feasibility of estimating cyclists' speed at new sites without additional model calibration. **Table 5-3** represents the results of this analysis in terms of RMSE and MAE over the test set. The regression models are tuned with the same hyperparameters as in **Table 5-2**. The MLP regression, implemented with the inverse regression approach, has the best performance (see 2nd and 8th rows of **Table 5-3**), where RMSE varies from 0.47m/s to 0.82m/s and MAPE varies from 6.8% to 9.5%. The average RMSE and MAPE of the MLP regression (with inverse dependent variable) over these ten conditions are 0.62m/s and 7.65%, respectively. MLP regression is more robust than Ridge regression with respect to the MAPE metric over ten sites. The standard deviation of MAPE of MLP regression is 0.78%, and the standard deviation of Ridge regression is 1.03%. This result shows the importance of having diverse cyclist data in the training set so that the system can be used for different facilities.

Table 5-3 System performance in 2nd scenario

Error Measure	Approach	Model	Site Under Test in Leave-One-Out Procedure									
			1	2	3	4	5	6	7	8	9	10
RMSE	Inverse	Ridge	0.66	0.61	0.66	0.67	0.61	0.65	0.83	0.58	0.49	0.49
		MLP	0.62	0.62	0.65	0.67	0.58	0.63	0.82	0.58	0.47	0.56
		Decision Tree	0.72	0.60	0.62	0.80	0.56	0.66	0.84	0.57	0.55	0.54
	Direct	Ridge	0.65	0.61	0.65	0.73	0.56	0.71	0.78	0.61	0.53	0.52
		MLP	0.65	0.63	0.68	0.79	0.59	0.65	0.74	0.60	0.54	0.52
		Decision Tree	0.69	0.58	0.62	0.84	0.53	0.66	0.78	0.57	0.57	0.53
MAPE (%)	Inverse	Ridge	7.9	7.2	8.3	8.4	7.8	8.0	10.2	6.9	7.1	6.6
		MLP	7.0	7.3	8.0	8.1	7.5	7.8	9.5	6.8	7.0	7.5
		Decision Tree	8.4	7.5	8.6	9.5	7.4	8.1	10.6	7.6	8.6	7.0
	Direct	Ridge	7.5	7.3	9.1	9.6	7.7	8.4	9.6	7.5	9.0	7.6
		MLP	7.4	7.5	9.6	9.9	8.0	7.7	8.9	7.4	9.3	7.6
		Decision Tree	8.1	7.2	8.6	10.3	7.6	8.0	10.2	7.6	9.4	7.6

5.5.3 MLP regression model

The MLP regression models with the inverse approach offer the best performance based on the results presented in **Table 5-2** and **Table 5-3**. **Figure 5-8** shows the architecture of the implemented neural network with fixed factors. The input layer includes all the features in addition to the intercept value, which is one. The hidden layer has three neurons that are activated by a *Rectified Linear Unit* (*ReLU*) function. These neurons learn and compute a weighted combination of the features and provide the input for the output layer. The weights of the hidden layer are $u_{j,i}$ and the weights of the output layer are $v_{i,1}$. The outputs of the hidden layer neurons are computed in (5-17).

The k is the sample ID, the number of features including intercept is 14, and the number of hidden layer neurons is 3. The $x_j^{(k)}$ is the j^{th} feature of the k^{th} sample: $x_1 = N_{first}$, $x_2 = N_{last}$, …, $x_{14} = 1$. The $u_{j,i}$ is the weight of the link that connects j^{th} feature to the i^{th} hidden layer neuron. The value of the $u_{j,i}$ of the neural network tuned for cyclist speed estimation are presented in the first three rows of **Table 5-4**. The output of the i^{th} neuron is calculated by $ReLU(y_i)$ where the activation functions, $ReLU(x)$, is x when $x \geq 0$ and is 0 when $x < 0$.

$$y_i^{(k)} = \sum_{j=1}^{14} u_{j,i} \times x_j^{(k)}; i = 1,2,3. \quad (5\text{-}17)$$

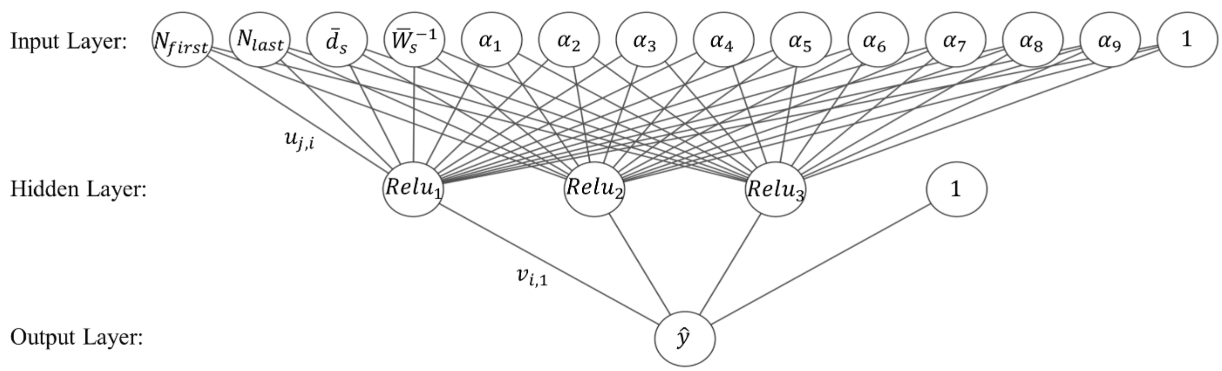


Figure 5-8 The Architecture of the implemented neural network

The output of the hidden layer is the input of the output layer, where a weighted combination of the input is computed as the predicted value. The dependent variable, $\hat{y}^{(k)}$, is computed in (5-18).

$$\hat{y}^{(k)} = \sum_{i=1}^{3} v_{i,1} \times ReLU\left(y_i^{(k)}\right) + v_{4,1} \times 1 \quad (5\text{-}18)$$

$v_{j,1}$ is the weight of the link that connects i^{th} hidden neuron to the output neuron, and $v_{4,1}$ is the weight of the intercept. The values of $v_{j,1}$ of the fitted neural network for cyclist speed estimation are presented in the last row of **Table 5-4**.

In this model, since the inverse approach is used, the number of video frame differences is used as the dependent variable, therefore, $\hat{y}$ is the predicted value of frame difference. The corresponding value of time difference is calculated by $\widehat{\Delta T}_L = \hat{y}/fps = \hat{y}/28.8$, and the estimated speed by LiDAR sensor ($\hat{S}_L$) is obtained from (5-19).

$$\hat{S}_L = d_V/(\hat{y}/FrameRate) = 7/(\hat{y}/28.8) = (7 \times 28.8) \times (\hat{y})^{-1} \quad (5\text{-}19)$$

Table 5-4 Coefficients of the implemented neural network

	N_{first}	N_{last}	$\bar{d}_s$	$\bar{W}_s^{-1}$	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	1
$u_{j,1}$	1.196	0.834	0.682	0.747	-0.195	-0.285	0.517	1.156	0.903	0.218	0.736	0.745	1.240	1.373
$u_{j,2}$	0.213	-0.383	0.322	0.378	1.126	0.425	0.094	0.484	-0.705	1.184	-0.795	-0.188	-0.630	-0.565
$u_{j,3}$	-0.388	-0.041	0.292	-0.272	0.362	0.540	0.574	0.034	-0.972	0.787	-0.642	0.052	-0.301	-0.312

	y_1	y_2	y_3	1
$v_{i,1}$	1.245	-0.972	-0.840	0.6680

These weights and formulas can be stored in the low-power ARM microcontroller used in this study. Since the proposed neural network is a feedforward network and computes the output using equations (5-11, 5-12 and 5-13), the system can provide real-time count, headway, spacing, and cyclists' speed information.

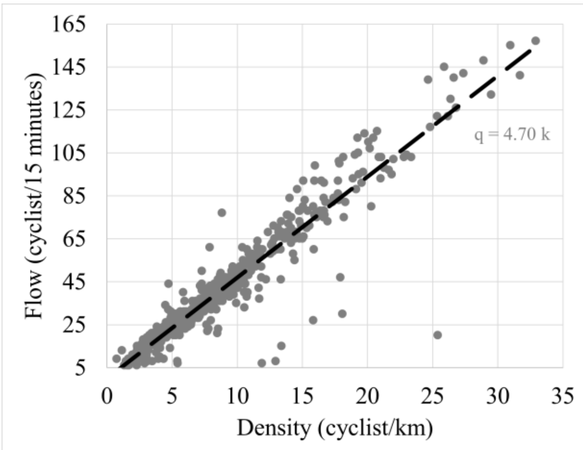
5.6 Traffic Flow Parameters Outcomes

The 1D LiDAR system monitors cyclist facilities and extracts several cyclist traffic factors: flow, density, speed, headway, and spacing. These traffic factors can help in bicycle planning and traffic management.

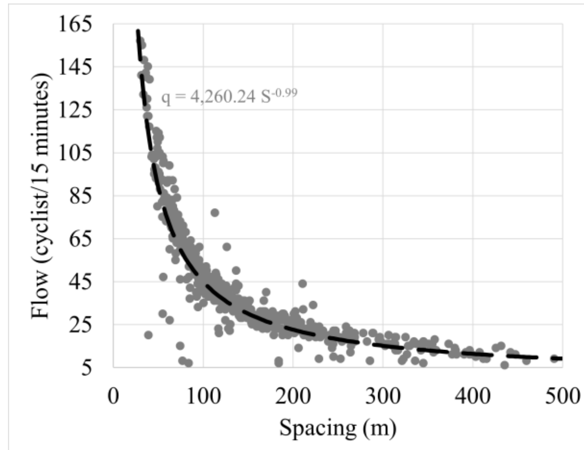
5.6.1 Traffic flow parameters

The traffic data of cyclist facilities, discussed in **Table 5-1** and aggregated in 15-minute time intervals, facilitate the examination of the fundamental traffic flow relationships. For every 15-minutes interval, the flow rate is the total number of cyclists passing mid-block through the cycling facility, the density is the inverse of the average spacing (headway) of the cyclists (see Equation (5-12)), and the speed factor is the average speed of all cyclists observed in that interval.

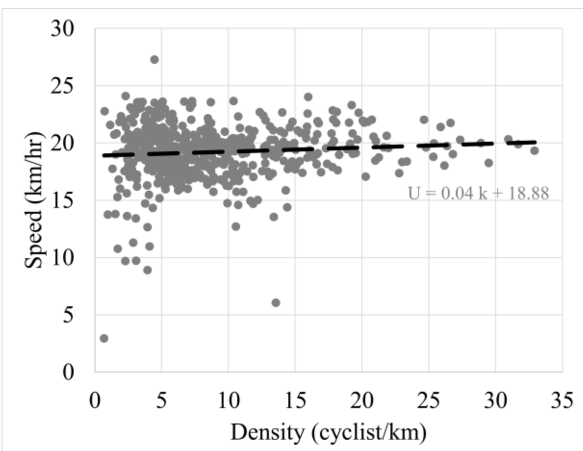
A set of indicators (flow, density, spacing, and speed) is generated for each time interval. Then the seven bi-directional cyclist facilities (see **Table 5-1**) in the collected dataset are aggregated to generate four diagrams: flow-density (**Figure 5-9 (a)**), flow-spacing (**Figure 5-9 (b)**), speed-density (**Figure 5-9 (c)**), and speed-flow (**Figure 5-9 (d)**). A dashed line is added to each figure to show the general trend of the data points. Although an equation is developed for each of these trendlines, calibrating a function for each pair of traffic factors requires extensive data collection from several locations on the cyclist network.



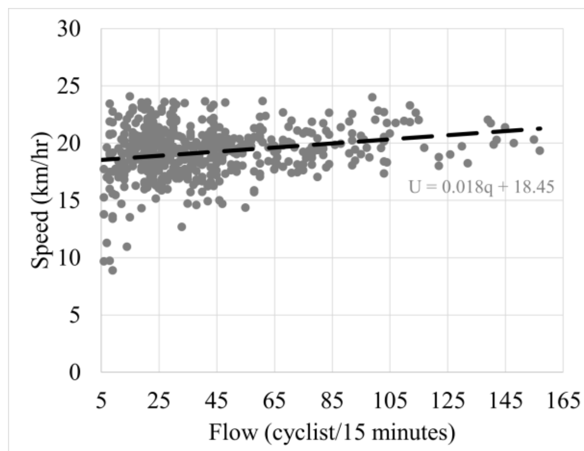
a) Flow-Density



b) Flow-Spacing



c) Speed-Density



d) Speed-Flow

Figure 5-9 Fundamental diagrams of cyclist traffic flow

The fundamental diagrams of these seven cyclist facilities illustrated in **Figure 5-9 (a-d)** demonstrate the free flow (uncongested) regime with a high level of service. The flow-density relationship (**Figure 5-9 (a)**) supports that the data points represent the free-flow regime where the flow increases proportionally with density. Based on traditional flow diagrams of vehicles, it is expected to have a maximum flow at the capacity condition after which, the flow decreases with density. However, this chart shows that the capacity is not reached for the seven cyclist facilities that are studied. Similarly, the speed-flow chart (**Figure 5-9 (d)**) is expected to shift to a downtrend after reaching the capacity value; however, it remains flat, demonstrating only the free-flow regime of cyclist traffic.

Although the data are collected during the peak cycling season for all daytime hours (including the entire AM and PM peak periods) at high-demand cyclist facilities in Montreal, Canada, none of the facilities reached or neared capacity. Therefore, the congested regime is not observed. The highest flows recorded per direction were 160 cyclists per 15 minutes or 640 cyclists per hour. The capacity of a 3-meter-wide bidirectional cycling facility is at least 1280 cyclists per hour and likely much greater.

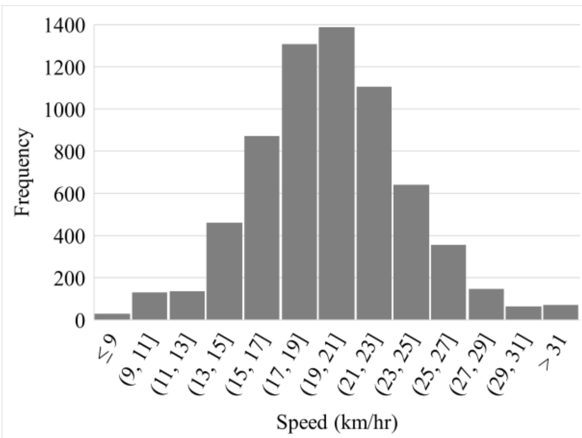
There are a few irregularities in the bottom-right of **Figure 5-9 (a)** or bottom-left of **Figure 5-9 (b)**. These samples correspond to a platoon of cyclists traveling on the bike path with low average spacing at a time interval with relatively low cyclist traffic. As a result, the density, which is the inverse of average spacing, is higher than normal while the flow remains low.

5.6.2 Speed analysis

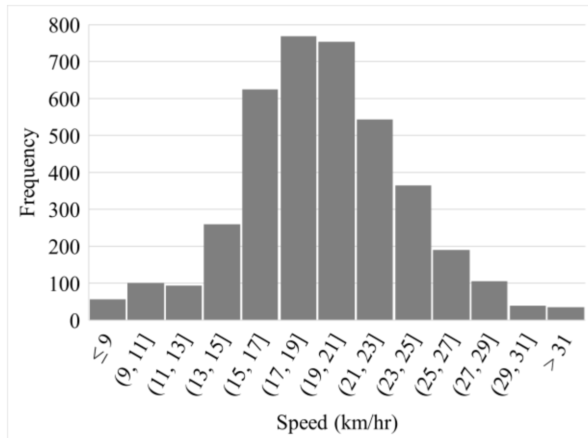
Of the sites listed in **Table 5-1**, two locations are chosen to explore the factors extracted from the LiDAR system based on their high cycling volume. At both locations, cyclist speed distributions, flow, and density are analyzed. The 3rd site has the highest hourly cyclist volume; in 12 hours of data collection, 6,717 cyclists passed the line-of-sight of the LiDAR system. This site is located mid-block, on a bidirectional cycle track in downtown Montreal (De Maisonneuve Boulevard near Metcalfe Street), and the data were collected on Thursday, September 13, 2018. The 4th site, located mid-block on the same cycle track as site three (de Maisonneuve Boulevard near Saint-Denis Street, 1.8km to the east of site three), also has high flow; 3,939 cyclists were recorded in 12 hours on Friday, September 14, 2018.

Figure 5-10 (a & b) illustrates the cyclist speed distribution of the 3rd and 4th sites, respectively. As expected, the speed distributions at both sites appear as normal distributions. Since both sets include many samples, the Kolmogorov-Smirnov test is applied to the speed distributions; both distributions have been proven not to be statistically significant, with the hypothesis that their samples are not taken from a normal distribution.

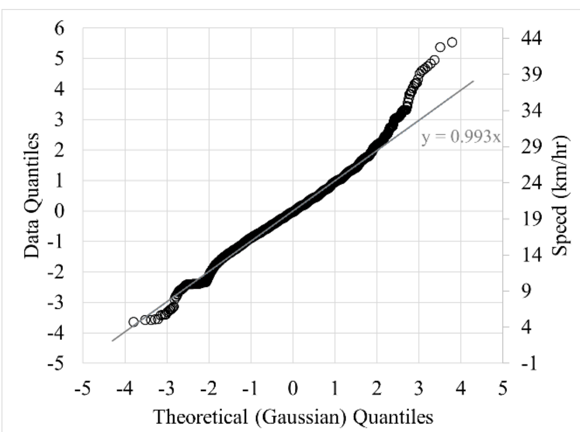
Figure 5-10 (c & d) shows the normal Q-Q plot of the two distributions in **Figure 5-10 (a & b)**. The normal Q-Q chart visually describes if a set of samples is taken from a normal distribution. For speeds between 9km/hr and 30km/hr, the data points follow very close to a normal distribution; the data are tangent on the expected line (zero mean). On the other hand, the samples with a speed value outside the range (9km/hr to 30km/hr) stray from the desired line (zero mean), though their frequency is very low with respect to the entire sample.



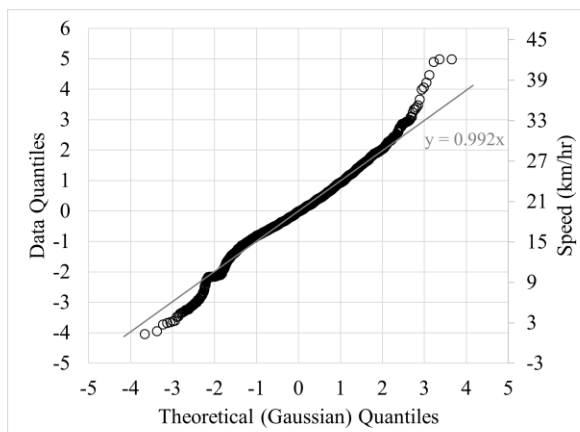
a) Cyclists speed distribution of the 3rd site



b) Cyclists speed distribution of the 4th site



c) Normal Q-Q plot of speed of the 3rd site



d) Normal Q-Q plot of speed of the 4th site

Figure 5-10 Histogram of bidirectional estimated speed (m/s)

5.7 Conclusion

Despite the importance of cycling in urban mobility, there is a lack of automated real-time embedded systems and methods to monitor bicycle traffic flow parameters (volume, speed, density, etc.) continuously and automatically and investigate the performance of cycling facilities.

This study proposes a LiDAR system composed of two single-beam sensors and algorithms to collect and estimate bicycle microscopic traffic flow data. The performance of the proposed monitoring system for measuring cyclist speed and related parameters is evaluated by comparing automated traffic flow parameters with ground truth from manually validated video footage. The LiDAR system generates time-stamped records of each passing cyclist from which hourly flow rate, speed, headway, spacing, and cyclist density are extracted automatically. Among other applications, the system can collect data for evaluating the performance of cycling facilities and improving bicycle traffic operations.

Three different regression models with four regressors are implemented in this study to correct sensor measures and automatically estimate cyclist speed. The same optimal functional form of the models is tested in two scenarios, each producing different model coefficients. In the first scenario, the data of the ten sites are merged and then split 75% and 25% for the training and test sets, respectively. In this scenario, implementing MLP regression gives high performance with MAPE of 7.1% and RMSE of 0.61m/s over the test set. The coefficient of determination (R^2) of this estimation is 0.78 over the test set. In the second scenario, by following a leave-one-out strategy, the data of one site is used for testing and the nine remaining sites for training. The performance is similar to the first scenario, RMSE varies from 0.43m/s to 1.21m/s, and MAPE ranges from 5.9% to 11.3% over ten different selections available for the test set. The methodology proposed in this study is developed using off-site, post-processing of the LiDAR signals collected at ten study sites. Although the methodology is developed using off-site post-processing, a relatively low-cost and low-power ARM microcontroller would be capable of estimating bicycle traffic flow parameters on-site, continuously in real-time.

Future work includes a deeper analysis of traffic-flow performance on more congested cycling facilities. This involves a similar exploration of flow, speed, and density measures for cyclists on different types of cycling infrastructure such as unidirectional bicycle lanes and bicycle paths, in real-time. The impact of cycling facility attributes, such as lane width, intersection spacing, and intersection attributes on cycling flow parameters will be tested. These results will allow for modeling the capacity of different types of cycling infrastructure. Other applications of the proposed system could also be explored, such as its integration in warning signs and traffic signals to protect or prioritize cyclists.

5.8 Declaration

Acknowledgments: Funding for this project is provided in part by the Natural Sciences and Engineering Research Council (NSERC).

Authors' contributions: The authors confirm contribution to the paper as follows: study conception and design: E. Nateghinia, A. Lesani, D. Beitel, and L. F. Miranda-Moreno; data collection: A. Lesani, and D. Beitel; data processing and system modeling: E. Nateghinia; analysis and interpretation of results: E. Nateghinia, and D. Beitel; draft manuscript preparation: E. Nateghinia, D. Beitel, and L. F. Miranda-Moreno. All authors reviewed the results and approved the final version of the manuscript.

Data Availability: The cyclist's hourly flow and speed data are available from the corresponding author upon request.

Code Availability: The code of this study is available from the corresponding author upon request.

Conflict of Interest: The authors confirm that there are no known conflicts of interest associated with this publication.

Conference Presentation: An early version of this research was presented at the 99th TRB Annual Meeting in January 2020.

References

1. Milakis D, Cervero R, Van Wee B, Maat K. Do people consider an acceptable travel time? Evidence from Berkeley, CA. Journal of Transport Geography. 2015;44:76-86.

2. Lusk AC, Furth PG, Morency P, Miranda-Moreno LF, Willett WC, Dennerlein JT. Risk of injury for bicycling on cycle tracks versus in the street. Injury prevention. 2011;17(2):131-5.

3. Strauss J, Miranda-Moreno LF, Morency P. Cyclist activity and injury risk analysis at signalized intersections: A Bayesian modelling approach. Accident Analysis & Prevention. 2013;59:9-17.

4. Strauss J, Miranda-Moreno LF. Speed, travel time and delay for intersections and road segments in the Montreal network using cyclist Smartphone GPS data. Transportation research part D: transport and environment. 2017;57:155-71.

5. Johnson PC. Bicycle Level of Service: Where are the Gaps in Bicycle Flow Measures? : Portland State University; 2014.

6. Figliozzi M, Wheeler N, Monsere CM. Methodology for estimating bicyclist acceleration and speed distributions at intersections. Transportation research record. 2013;2387(1):66-75.

7. Berjisian E, Bigazzi A. Extraction and Investigation of Bicycle Cruising Speeds from GPS Data Using Time Series Clustering Transportation Research Board 98th Annual Meeting; Washington D.C.2019.

8. Twaddle H, Grigoropoulos G. Modeling the Speed, Acceleration, and Deceleration of Bicyclists for Microscopic Traffic Simulation. Transportation Research Record: Journal of the Transportation Research Board. 2016(2587):8-16.

9. Goñi-Ros B, Yuan Y, Daamen W, Hoogendoorn SP. Empirical analysis of the macroscopic characteristics of bicycle flow during the queue discharge process at a signalized intersection. Transportation research record. 2018;2672(36):51-62.

10. Bai L, Liu P, Chan CY, Li Z. Estimating level of service of mid-block bicycle lanes considering mixed traffic flow. Transportation research part A: policy and practice. 2017;101:203-17.

11. Jin S, Qu X, Zhou D, Xu C, Ma D, Wang D. Estimating cycleway capacity and bicycle equivalent unit for electric bicycles. Transportation Research Part A: Policy and Practice. 2015;77:225-48.

12. Xu C, Li Q, Qu Z, Tao P. Modeling of speed distribution for mixed bicycle traffic flow. Advances in mechanical engineering. 2015;7(11):1687814015616918.

13. Beitel D, Domínguez-Michelen M, Miranda-Moreno L. The Relationship Between Macroscopic Cycling Traffic Parameters on Cycle Tracks. Transportation Research Board 98th Annual Meeting; Washington D.C. 2019.

14. May AD. Traffic Flow Fundamentals. Englewood Cliffs, NJ: Prentice Hall; 1990.

15. Highway Capacity Manual. Washington, DC2000.

16. Greenshields BD, Bibbins JR, Channing WS, Miller HH. A study of traffic capacity. Highway Research Board proceedings. 1935;1935.

17. Van Aerde M, Rakha H, editors. Multivariate calibration of single regime speed-flow-density relationships [road traffic management]. Pacific Rim TransTech Conference 1995 Vehicle Navigation and Information Systems Conference Proceedings 6th International VNIS A Ride into the Future; 1995; Seattle, WA: IEEE.

18. Jeng SL, Chieng WH, Lu HP. Estimating Speed Using a Side-Looking Single-Radar Vehicle Detector. IEEE Transactions on Intelligent Transportation Systems. 2014;15(2):607-14.

19. Nordback K, Kothuri S, Phillips T, Gorecki C, Figliozzi M. Accuracy of bicycle counting with pneumatic tubes in Oregon. Transportation research record. 2016;2593(1):8-17.

20. Nordback K, Janson BN. Automated Bicycle Counts: Lessons from Boulder, Colorado. Transportation research record. 2010;2190(1):11-8.

21. Zhao Y, Zhang HM. A unified follow-the-leader model for vehicle, bicycle and pedestrian traffic. Transportation research part B: methodological. 2017;105:315-27.

22. Liang X, Xie M, Jia X. New microscopic dynamic model for bicyclists' riding strategies. Journal of Transportation Engineering, Part A: Systems. 2018;144(8):04018034.

23. Guo N, Jiang R, Wong S, Hao QY, Xue SQ, Hu MB. Bicycle flow dynamics on wide roads: Experiment and modeling. arXiv preprint arXiv:06084. 2019.

24. Romanillos G, Gutierrez J. Cyclists do better. Analyzing urban cycling operating speeds and accessibility. International journal of sustainable transportation. 2020;14(6):448-64.

25. Zaki MH, Sayed T. Automated cyclist data collection under high density conditions. IET intelligent transport systems. 2016;10(5):361-9.

26. Zaki M, Sayed T, Cheung A. Computer Vision Techniques for the Automated Collection of Cyclist Data. Transportation Research Record: Journal of the Transportation Research Board. 2013;2387(1):10-9.

27. Hoogendoorn S, Daamen W. Bicycle headway modeling and its applications. Transportation research record. 2016;2587(1):34-40.

28. Mohammed H, Bigazzi AY, Sayed T. Characterization of bicycle following and overtaking maneuvers on cycling paths. Transportation research part C: emerging technologies. 2019;98:139-51.

29. Lesani A, Nateghinia E, Miranda-Moreno LF. Development and evaluation of a real-time pedestrian counting system for high-volume conditions based on 2D LiDAR. Transportation research part C: emerging technologies. 2020;114:20-35.

30. Saleh K, Hossny M, Hossny A, Nahavandi S, editors. Cyclist detection in LIDAR scans using faster R-CNN and synthetic depth images. IEEE 20th International Conference on Intelligent Transportation Systems (ITSC); 2017; Yokohama, Japan.

CHAPTER 6.

CONCLUSION AND FUTURE WORK

CHAPTER 6: CONCLUSION AND FUTURE WORK

6.1 General Conclusion and Summary of Results

The primary motivation of this thesis was to propose and test novel methodologies for traffic monitoring and surrogate safety analysis based on alternative LiDAR-sensing technologies, including high-resolution 3D LiDAR, low-resolution 3D LiDAR, and dual-beam 1D sensors. This research focused on the methodological development and evaluation of the performance of the LiDAR-based methodologies integrating various computer data processing and machine learning algorithms. For this purpose, multiple applications are documented using ground (real) data collected in bicycle facilities and urban intersections with mixed traffic conditions.

6.1.1 LiDAR-based methodology for traffic monitoring at urban intersection

Chapter 2 presented a novel methodology for processing 3D LiDAR point cloud data for traffic monitoring at urban intersections with high-mixed traffic. The method is developed to be adaptive to two different resolution LiDAR sensors. It integrates algorithms for background modeling, foreground detection, and clustering foreground objects into road users' point clouds. The background modeling and foreground detection are implemented in the spherical coordinate system. The foreground frames in the spherical coordinate system are converted to x-y-z in the Cartesian coordinate system, where spatial clustering is applied to construct the road users. A set of features based on various attributes of a point cloud, such as physical, LiDAR, and spatial attributes, are extracted and used for road user classification. A semi-automated method, leveraging the unique geo-boundary of urban intersections, is implemented to resolve the main challenge of data labeling. A classification algorithm based on the XGBoost tool is integrally implemented for road user classification given the feature vector and labeled point clouds. Additionally, the proposed methodology employs a Kalman Filter and a data association to collect the trajectories of the road users.

Each LiDAR sensor was installed at seven urban intersections, some of which experience high-mixed traffic. The correct classification rates for the high and low-resolution LiDAR are 95% and 91%, respectively. Specifically, pedestrians and cyclists were correctly classified at a rate of 92%-93% in the high-resolution setting. In the low-resolution setup, the correct classification rates for pedestrians and cyclists were between 85% and 89%. As for cars, they were correctly classified at a rate of 98% in the high-resolution setting and 94% in the low-resolution setup. The trajectory counts obtained from the LiDAR are compared against manual counts for 30 minutes at each intersection in terms of Weighted Average Absolute Percent Difference (WAAPD). For the higher resolution LiDAR, WAAPD between LiDAR and ground-truth counts is 5%, 7%, and 6% for pedestrians, cyclists, and vehicles, respectively. In contrast, the corresponding WAAPD values for the lower resolution LiDAR are 7%, 23%, and 10%. Overall, the high- and low-resolution LiDAR's WAAPD shows a 6% and 13% deviation from ground truth manual counts, respectively.

One of the main challenges of the system is to differentiate between some of the cyclists and some of the vehicles partially observed, leading to overcounting cyclists and undercounting vehicles. The two groups exhibit highly similar narrow-point clouds.

6.1.2 LiDAR-based methodology for surrogate safety analysis

Chapter 3 expanded the methodology developed in Chapter 2 into road safety analysis by developing a method for surrogate safety analysis based on 3D LiDAR sensing technologies. This includes the development of a method for calculating surrogate safety indicators, such as Time-to-Collision (TTC) and Post-Encroachment Time (PET), based on road users' shape data. This approach is compared with traditional methods that rely on the trajectory's centroid. The point clouds representing road users are used to fit a minimum rotated rectangle, reconstructing the road users' shape, length, and width. The orientation of the road users is adjusted toward the direction of travel in 2D space. The corner points of the modified rectangle are utilized to project the future positions of road users a few seconds ahead (e.g., 10 seconds). A polygonal path is constructed for each user, extending from their current position to a projected future position. During each frame, if the polygon paths of two users intersect, the times of arrival and departure from the shared area are used to calculate the TTC. For calculating Post-Encroachment Time (PET), the combined polygon of a user's rectangle is built and intersected with other users' polygon path.

The comparative analysis demonstrated that applying a uniform buffer size for all types of road user interactions does not replicate the results achieved with the polygon-based method. For non-critical conflicts, a 2-meter buffer size for pedestrian-vehicle and a 2.5-meter to 3-meter buffer size for cyclist-vehicle interactions align with the shape-based method's TTC and PET results. However, for critical conflicts, these buffer sizes tend to estimate a lower value for TTC and PET and, therefore, overreport critical conflicts occurring in under 1.5 seconds. For vehicle-vehicle interactions, the buffer size that aligns with the results of the shape-based method varies depending on the conflict angle, ranging from 3 meters to 3.5 meters. However, these buffer sizes would result in a lower estimation of TTC and PET and, therefore, overreporting of critical conflicts.

6.1.3 Unsupervised methodology for a LiDAR-based level crossing monitoring

Chapter 4 presents an unsupervised method for the automated safety monitoring of railroad level crossings. A low-resolution LiDAR system was installed at a level crossing, with data collected on two different days. The methodology for processing LiDAR point cloud data includes background detection in an x-y-z coordinate system, achieved by voxelizing the 3D space. This is followed by clustering the point cloud in three-dimensional space and tracking their centroid in two-dimensional space using a Kalman Filter with four state variables. An unsupervised, rule-based classification method, considering the shape and speed of road users, was designed to identify the road user class of each trajectory, classifying them into trains, trucks, cars, and a combined group for pedestrians and cyclists. This classification method was developed to circumvent the unavailability of labeled LiDAR point cloud data for this specific low-resolution LiDAR sensor. Additionally, the system's processing time was lower since the classification was applied once per trajectory, and extra features were not extracted.

Road users at level crossings were monitored for potential trespassing incidents when a train was either approaching or present at the crossing. The system successfully identified two critical conflicts and a near-miss event involving two pedestrians and an approaching train. The average absolute percentage deviation of the model in counting road users over 2 hours was reported at 5% and 3% for motorized vehicles and 10% and 13% for non-motorized road users on two separate days. Although the results suggest better performance than the aggregated results of the supervised methodology for the same low-resolution LiDAR in Chapter 2, there are a few notable points. First, the performance of the supervised method is reported as an average across several intersections, with some intersections showing better results and others showing lower performance where the LiDAR was installed at intersections with extremely high traffic volumes. Second, the unsupervised methodology does not distinguish between pedestrians and cyclists.

Additionally, the monitored level crossing had a significantly simpler design, with only two streets, no turning movements, and pedestrians traveling only on two crosswalks and two sidewalks. The sensors were installed at a relatively appropriate distance from the level crossing, allowing optimal coverage. The supervised method tends to overcount cyclists, and the unsupervised method overcounts non-motorized users using a low-resolution LiDAR. This discrepancy is primarily due to the partial observation of vehicles, especially passenger cars, that are occasionally mixed with cyclists.

6.1.4 1D LiDAR-based methodology for cyclist traffic monitoring

Chapter 5 introduces a novel approach for large-scale cyclist traffic data collection, utilizing a low-cost LiDAR solution. A system comprising two 1D LiDAR sensors was installed at ten cyclist facilities. These sensors were positioned perpendicular to the cyclists' direction of travel, enabling the detection of cyclists and the estimation of their average speed as they passed through the system's line of sight. The LiDAR data processing methodology involves cyclist detection based on the system's distance measurements. Additionally, a neural network was designed, which utilizes a feature vector as input, to estimate the speed of cyclists accurately. In addition to counting cyclists, Chapter 5 details the computation of other traffic parameters such as flow rate, headway, spacing, and density. These are calculated using time-stamped data along with the cyclists' speed. The accuracy of cyclist counting varies significantly across different cyclist facilities, ranging from 94.3% to 99.6%. Additionally, the Root Mean Square Error (RMSE) in speed estimation fluctuates between 0.43m/s and 1.21m/s. The Mean Absolute Percentage Error (MAPE) for speed estimation also shows variation across ten locations, ranging from 5.9% to 11.3%.

In conclusion, this research presents alternative methodologies and applications leveraging LiDAR sensing technologies to monitor mixed-traffic transportation facilities. The suitability of each methodology can be determined based on various factors, such as the type of application (real-time vs. post-processing), economic considerations (sensor resolution), and specific data requirements. Overall, LiDAR technologies demonstrate significant potential for traffic monitoring and data collection, particularly beneficial in environments with high volumes of non-motorized traffic.

It is important to acknowledge certain limitations inherent in this research, which are discussed in subsequent sections. These discussions also pave the way for future work, offering directions for further advancements in this field.

6.2 Limitation

The proposed LiDAR-based methodologies have proven to be valuable for traffic monitoring and road safety applications. However, their implementation in practice is not without challenges.

One primary challenge is the cost. LiDAR, as an emerging technology, is significantly more expensive than its counterparts, notably camera-based systems. The price of a single 3D LiDAR unit ranges from a few thousand to hundreds of thousands of dollars. This high cost restricts the prospects of large-scale deployment or deployment of multiple LiDAR at one intersection and limits the scope of research and development of methodologies for LiDAR data processing.

Another limitation closely linked to cost concerns the resolution and field of view of LiDAR sensors. For effective traffic monitoring in urban intersections with high volumes of mixed traffic, high-resolution LiDARs are preferable, as they improve outcome accuracy, a fact demonstrated in this research. However, enhancement in either the field of view or resolution typically leads to proportional increases in costs, posing significant budgetary challenges for urban traffic control projects. On the other hand, while low-resolution LiDAR sensors are more cost-effective, they face unique challenges in busy urban intersections. These challenges include the difficulty of accurately capturing the fine details of complex intersection scenarios, which are crucial for effective traffic monitoring and safety analysis.

A critical factor in the deployment of the most widely used type of LiDAR, 3D rotational LiDARs, is the vertical field of view and its impact on system performance. In applications like autonomous vehicles or mobile laser scanning systems, LiDARs are typically mounted horizontally at a lower height on a vehicle. This positioning allows the laser channels to monitor the area around them uniformly. In contrast, the applications discussed in this thesis require LiDAR systems to be installed at a higher elevation to monitor urban intersections effectively. Such installations necessitate a downward tilt in orientation, significantly altering the coverage area and the uniformity of the angular space between vertical channels. This adjustment introduces new blind spots, making installation and effective deployment more challenging. Therefore, optimizing the installation of LiDAR systems is essential to ensure their successful implementation in urban traffic control projects.

One of the limitations this thesis faced was the lack of LiDAR data availability in traffic monitoring. The open-source data for LiDAR are comparatively less abundant. The current LiDAR dataset is primarily built for Autonomous vehicles and utilizes super high-resolution LiDAR (1, 2). This scarcity is partly due to the varied characteristics of LiDAR sensors, which can differ significantly from one system to another, thereby complicating the standardization of research and methodology development. Using these datasets for a lower-resolution LiDAR system might not yield the best performance due to the discrepancy in data quality. Furthermore, the point density varies significantly between different LiDAR systems, adding to the challenge of standardizing data across various models.

The limitations discussed are general issues researchers can encounter when working with LiDAR sensors. This thesis also discovered those limitations primarily. In addition, specific limitations are associated with the methodologies proposed in this research.

First, implementing the supervised learning algorithm necessitates a diverse sample set of road users, which must be manually extracted from the point cloud data. However, this process is highly time-intensive and requires significant budget allocation. As an alternative, a semi-automated labeling technique has been developed in a unique approach, which samples users based on their movement patterns and geo-location at intersections. For example, it samples pedestrians only if they are on a sidewalk or exhibit a sidewalk-crosswalk-sidewalk movement pattern. However, this approach still has gaps, particularly in distinguishing between cyclists and pedestrians. Additionally, the sampling and labeling of cyclists present challenges at intersections without dedicated bike lanes.

The results indicate a degree of overcounting across some intersections. This is observed either as a misclassification between two particular road user classes or the degraded performance of the clustering algorithm. The first issue arises when the point clouds of two road users become visually indistinguishable. For example, a partial car and a cyclist can both have very narrow point clouds. The latter issue often arises from a road user (primarily a vehicle) being divided into two distinct point clouds due to blind spots in the vertical laser channels when tilted. Additional validation could be conducted to improve each system component or the system installation procedure.

There have been a few limitations in applying LiDAR-extracted data of road users' shape and trajectory to calculate surrogate safety indicators, such as TTC and PET. The accuracy of classification and tracking in the LiDAR data affects the results of surrogate safety assessments. However, this issue impacts both centroid-based and shape-based methods. A pre-processing technique is implemented to overcome the tracking issue, and the same geo-locating technique outlined in Chapter 2 is used to flag road users with incorrect classifications. This process has the potential to be expanded and improved.

6.3 Future Work

Future work can be structured in a few directions. The first direction involves developing 3D methodologies incorporating emerging machine learning algorithms, including deep learning, to streamline LiDAR data processing from object detection to classification and clustering. This path presents a significant challenge, primarily due to the requirement of labeling the entire LiDAR point cloud. In the development of a 3D deep learning methodology using LiDAR sensors, it is crucial to reevaluate the sensor and its installation criteria. Specifically, understanding whether tilting the sensor downward would benefit such a system is essential. This is because most available datasets are prepared with the LiDAR installed horizontally.

The second avenue for future research involves a data fusion approach to traffic monitoring in urban environments. Specifically, combining LiDAR with camera-based systems could harness the strengths of the accurate spatial representation from LiDAR and the road user detection and classification capabilities of computer vision systems. Such a fusion could potentially improve the performance of the system. This improvement is attributed to the fact that road user classification in camera-based systems is advancing at a faster rate than in LiDAR-based systems, while LiDAR remains the most reliable source for geo-locating road users at urban intersections. Therefore, the accuracy of classification and tracking would significantly increase. This combination also presents a unique opportunity to rectify LiDAR point cloud labeling using image labels on a large scale, which could also enhance the performance of standalone LiDAR systems.

The third potential development area involves extending the methodology to a traffic monitoring system that utilizes two or more LiDAR sensors, ideally of the exact resolution. While using a super high-resolution LiDAR is cost-restricted, integrating multiple low-resolution LiDARs can be a more cost-effective alternative. This approach could address the partial field of view observation inherent in single LiDAR systems at large intersections. For example, the partial observation of a passenger car entering the coverage area from a distance might be indistinguishable from that of a cyclist. Dual LiDAR sensors could provide a more comprehensive view, mitigating such ambiguities. Moreover, this would add processing complexity but not significantly, as the LiDAR data are converted to point coordinates (x-y-z) and processed simultaneously.

Additional research is required to evaluate the performance of LiDAR systems under adverse weather conditions, particularly when compared with camera-based systems operating in the visual and thermal spectrum. One productive direction for future research could be to collect data under adverse weather conditions (e.g., rain, snow, fog) and assess the performance of the current methodology. The method could be expanded by implementing a systematic approach to manage and analyze extremely noisy data, thereby enhancing the robustness and reliability of LiDAR systems in varied environmental conditions.

Several potential directions for future work concerning the proposed shape-based surrogate safety methodology can be highlighted. Firstly, a comparative study at the individual conflict level is necessary. This would support and expand upon the comparative analysis presented in this thesis, contrasting shape-based versus centroid-based methods for traffic conflict identification and safety assessment. Additionally, complex movement patterns could be implemented and applied to the polygon to project road user positions along a non-linear path using acceleration or motion patterns. The effectiveness and reliability of such a method are subject to evaluation. Furthermore, PET results reveal promising applications in estimating headway and gaps in car-following scenarios. Accordingly, the LiDAR system could be strategically installed at selected locations to study these aspects in greater detail.

Finally, concerning the 1D-LiDAR methodology, there are also opportunities for further research. One key area is expanding the data collection program to high cyclist volume locations. This would enable more extensive data collection, facilitating detailed analysis of cyclist traffic volume and arrival rates, especially near critical points in cyclist facilities. Additionally, this system can be implemented to further investigate microscopic traffic parameters and patterns in alternative bicycle facilities.

REFERENCES

1. Liao Y, Xie J, Geiger A. KITTI-360: A novel dataset and benchmarks for urban scene understanding in 2d and 3d. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2022;45(3):3292-310.

2. Mei J, Zhu AZ, Yan X, Yan H, Qiao S, Chen L-C, et al., editors. Waymo open dataset: Panoramic video panoptic segmentation. European Conference on Computer Vision; 2022: Springer.

APPENDIX A: Map of Data Collection Intersections – Chapter 2

Figure A-1 presents a map showing the location of LiDAR installations with the intersections featuring the 16-channel LiDAR marked in red and those with the 32-channel LiDAR highlighted in green.

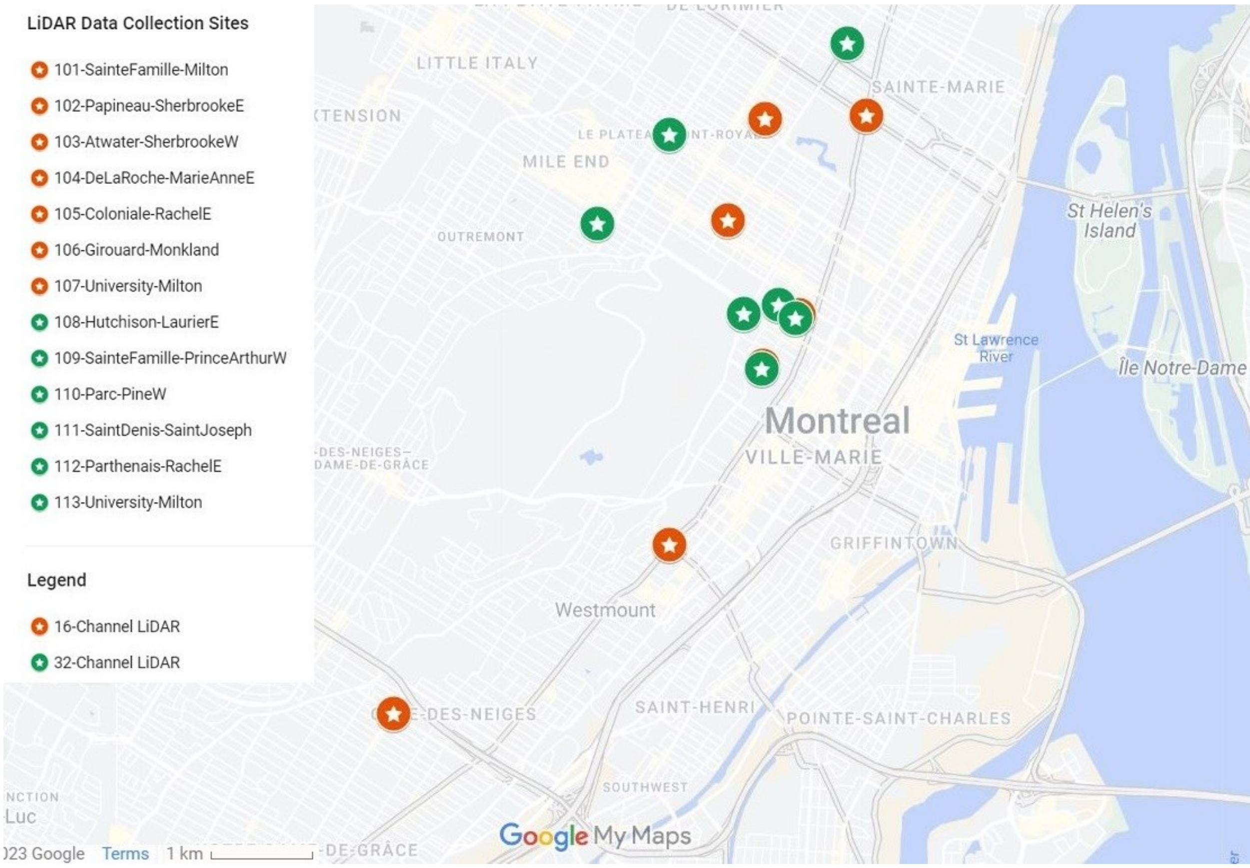


Figure A-1 A Map of intersection for data collection with the LiDAR systems

APPENDIX B: LiDAR Channels Gap Analysis

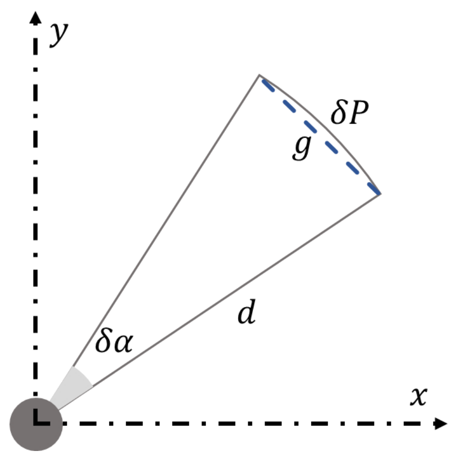
A LiDAR system features laser channels with horizontal and vertical resolutions. **Figure B-1 (a & b)** illustrates potential gaps. The vertical gap arises from the vertical resolution, indicating the angle difference between neighboring laser channels. Meanwhile, the horizontal gap results from the horizontal resolution, representing the angle difference between readings of the same laser channel at two consecutive timestamps.

The gap (g) between LiDAR measurements depends on the angular resolution ($\delta\alpha$) and the distance measured by the LiDAR. The angular resolution ($\delta\alpha$) can be either horizontal or vertical. At a given distance (d), the gap (g) can be determined using $g = 2\sin(\delta\alpha/2) \times d$. Moreover, the arc of this gap (δP) is calculated as a portion of the circle's circumference using $\delta P = (2\pi d) \times (\delta\alpha/360°) = ((\delta\alpha \times \pi)/180) \times d$.

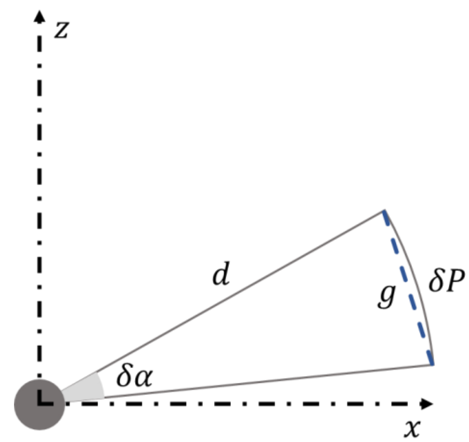
The horizontal resolution of both LiDAR sensors is 0.2° ($\delta\alpha = 0.2°$). Therefore, the gap and arc are approximately equal, and both are a function of distance: $g \cong \delta P = 6.98 \times 10^{-3} d$. At a distance of 10 meters, the gap is 3.49 cm, and at a distance of 50 meters, the gap is 17.45 cm.

The vertical angular resolution and gap vary between the two LiDAR sensors. The low-resolution LiDAR, VLP-16, maintains a consistently distributed vertical angular resolution of 2°. Consequently, at a distance of 10 meters, the vertical gap between two laser channels equals $2\sin(1) \times d = 34.9\ cm$, and at 50 meters, the gap extends to 1.7 cm. However, to address this gap, the low-resolution LiDAR sensor is tilted downward at an angle (β) and installed closer to the intersection.

The vertical angular resolution of the higher resolution LiDAR, VLP-32c, is not evenly distributed across its vertical channels. Among the 31 vertical angular gaps between LiDAR channels in VLP-32c, there are 17 gaps of 0.33°, four gaps of less than 1°, four gaps of less than 2°, and six gaps of more than 2°. For laser channels with a resolution of 0.33°, the gap at 10 meters is 5.76 cm, and for laser channels with a resolution of 1°, the gap at 10 meters is 17.45 cm. The installation of the higher resolution LiDAR is such that LiDAR channels with a resolution greater than 2° do not face the intersection and are only used to scan objects at higher heights.



a) Horizontal gap



b) Vertical gap

Figure B-1 Horizontal and vertical gap between laser channels

APPENDIX C: Road User Classification

Table C-1, **Table C-2**, and **Table C-3** presents the details results of the road user classification performance in Chapter 2. They report classification of road users on the test set per each LiDAR type and the combination of both.

Table C-1 Confusion matrix of XGBoost applied to the test set of 16-channel LiDAR

Labels		**LiDAR Observations**				
	Class Labels	0	1	2	3	total
Ground truth Observations	0	1958	133	149	0	2240
	1	132	703	8	0	843
	2	166	0	3062	105	3333
	3	0	0	56	684	740
	total	2256	836	3275	789	7156

Table C-2 Confusion matrix of XGBoost applied to the test set of 32-channel LiDAR

Labels		**LiDAR Observations**				
	Class Labels	0	1	2	3	total
Ground truth Observations	0	2055	171	84	2	2312
	1	218	1910	1	0	2129
	2	146	15	2977	10	3148
	3	0	0	28	661	689
	total	2419	2096	3090	673	8278

Table C-3 Confusion matrix of XGBoost applied to the combined test set of both LiDARs

Labels		**LiDAR Observations**				
	Class Labels	0	1	2	3	total
Ground truth Observations	0	3883	321	182	4	4390
	1	298	2535	12	0	2845
	2	287	24	6038	42	6391
	3	0	0	58	1327	1385
	total	4468	2880	6290	1373	15011

APPENDIX D: Kalman Filter Implementation

The following equation presents the implemented Kalman Filter as part of tracking component or smoothing of trajectories in Chapter 2, Chapter 3 and Chapter 4.

State Prediction:

$$X^{-} = AX + BU$$

$$P^{-} = APA^{T} + BQB^{T}$$

State Correction:

$$S = CPC^{T} + R$$

$$K = P^{-}C^{T}S^{-1}$$

$$X = X^{-} + K(Z - CX^{-})$$

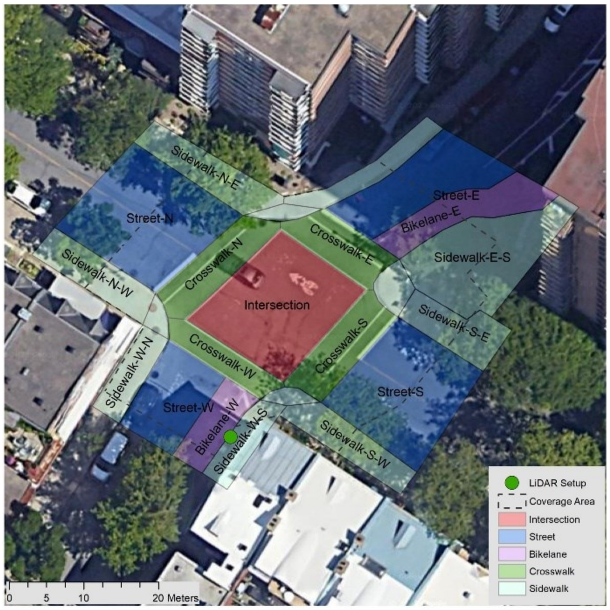
$$P = \left(I_{n_s} - KC\right)P^{-}$$

Output Prediction:

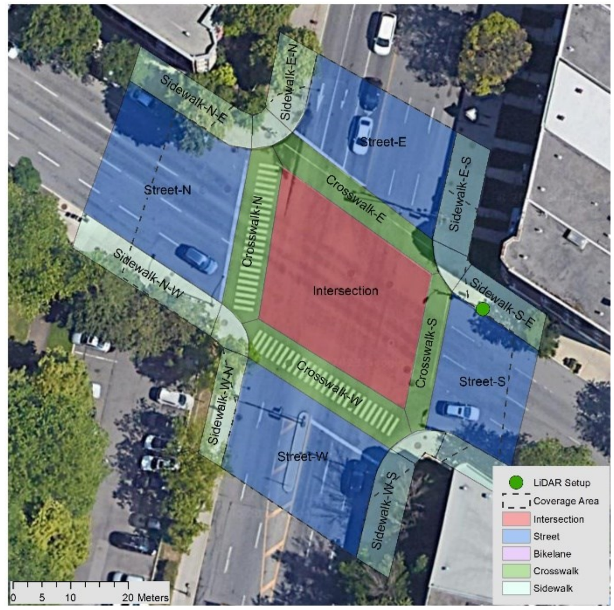
$$X^{+} = AX + BU$$

$$Z^{+} = CX^{+}$$

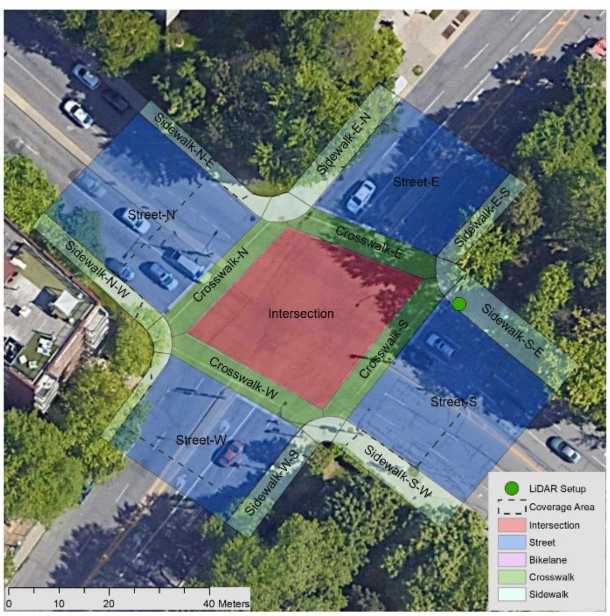
APPENDIX E: Intersection Segmentation and GIS Calibration



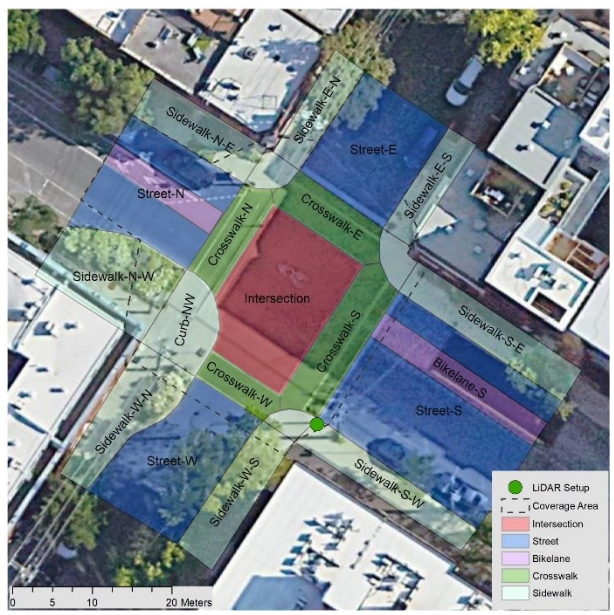
ID #101– Sainte Famille - Milton



ID #102 – Papineau – Sherbrooke E

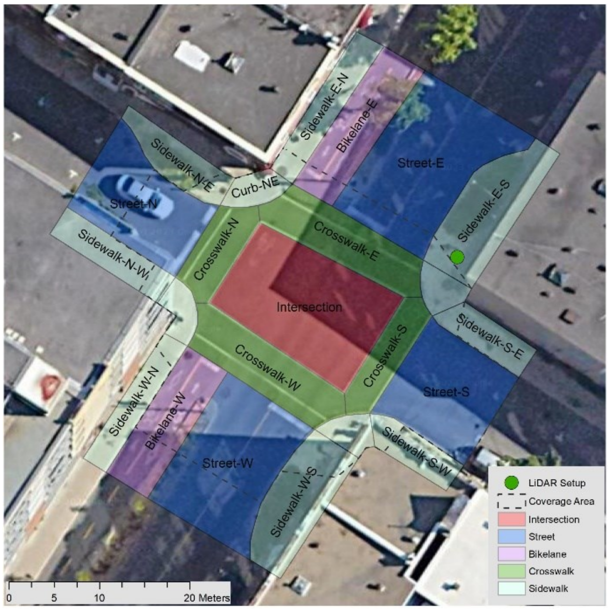


ID #103 – Atwater – Sherbrooke W



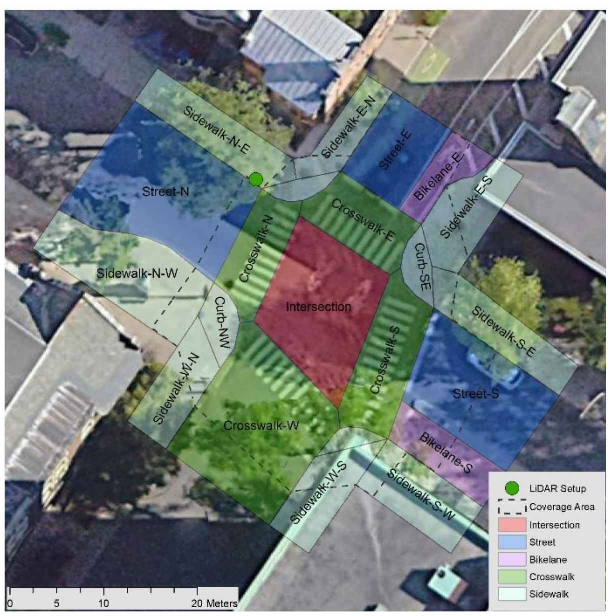
ID #104 – De La Roche – Marie Anne E

Figure E-1 The GIS calibration of intersections with 16-Channel LiDAR system (part 1)



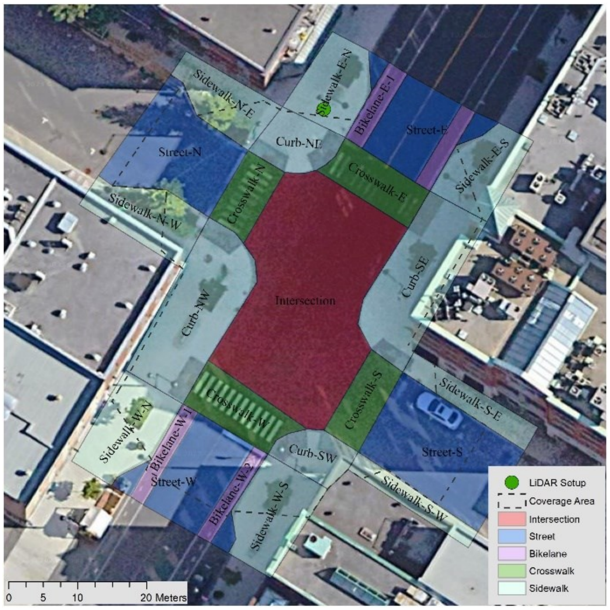
ID #105 – Coloniale – Rachel E

ID #106 – Girouard - Monkland

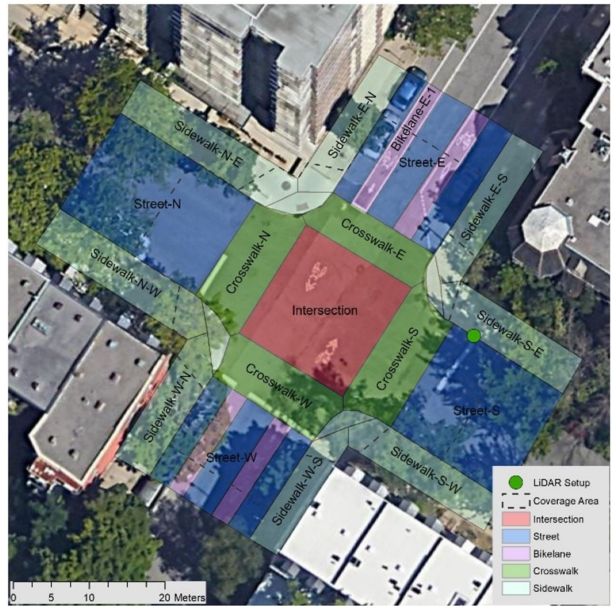


ID #107 – University - Milton

Figure E-2 The GIS calibration of intersections with 16-Channel LiDAR system (part 2)



ID #108 – Hutchison – Laurier E

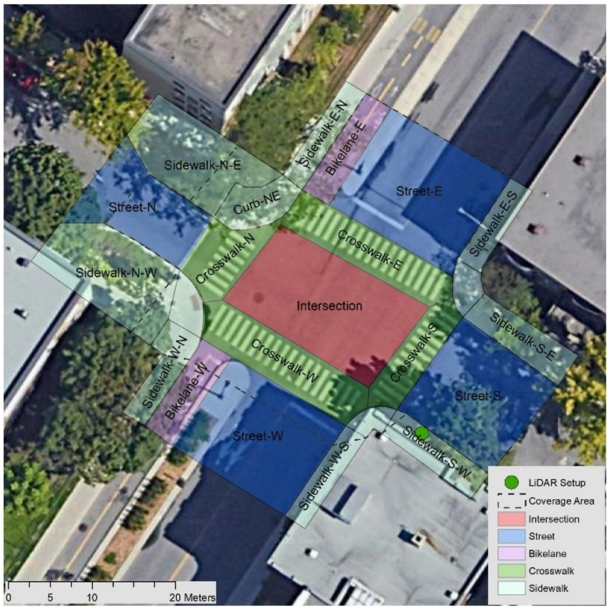


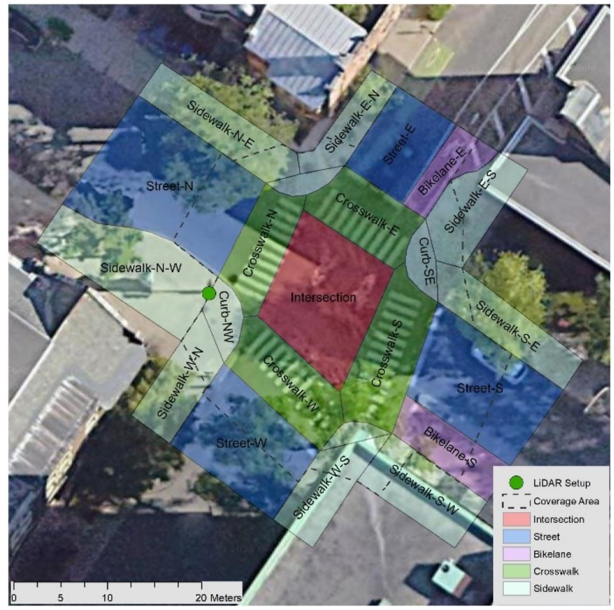
ID #109 – Sainte Famille – Prince Arthur W

ID #110 – Parc – Pine W

ID #111 – Saint Denis – Saint Joseph E

Figure E-3 The GIS calibration of intersections with 32-Channel LiDAR system (part 1)





ID #112 – Parthenais – Rachel E

ID #113 – University - Milton

Figure E-4 The GIS calibration of intersections with 32-Channel LiDAR system (part 2)