A Data-Driven Approach to Creating a Traffic Sign Asset Inventory using Remote Sensing Technology

by

Lloyd Andrew Maurice Karsten

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Transportation Engineering

Department of Civil and Environmental Engineering
University of Alberta

ABSTRACT

Traffic signs play a critical role in the safety and efficiency of any roadway, but with limited information on the current status of traffic sign inventories, the placement and condition of traffic signs goes largely unchecked. Therefore, the collection of a complete traffic sign inventory (TSI) is needed to ensure traffic signs meet the needs of the current driving population. However, the size of current global traffic sign networks makes applying traditional survey methods to the collection of a TSI difficult, if not economically infeasible. Therefore, there is room for technological and methodological advancement to create a new TSI process to inventory and analyze traffic signs. This thesis proposes the use of light detection and ranging (LiDAR) and video-log imaging to conduct an automated extraction of a TSI. The details of traffic sign location, orientation, placement, and panel classification define the fundamental components of a TSI.

Traffic signs are extracted through a Gaussian mixture model, density clustered, and filtered for flatness before measuring the vertical and horizontal orientation through principal component analyses. Traffic sign placement is dependent on the location of lane markings. Therefore, the road surface near each traffic sign is extracted, rasterized, and intensity manipulated to determine the linear lane marking intensity edges. The markings are used to determine the lateral and vertical placement of the traffic signs. Sign classification is determined from video-log images by applying a trained GoogLeNet convolutional neural network. This completes the traditional TSI and creates a platform with which to analyze the efficacy of traffic sign placements. Additionally, the detail provided by LiDAR scans allows for the measurement of the visibility of the traffic signs, which is unavailable through traditional surveying methods. This is used to assess the time available to drivers to read and react to traffic signs placed along

the segment. A 4-km test segment was utilized to assess the accuracy of the proposed method, providing an Eastbound TSI for 30 traffic signs.

The intensity-based extraction of traffic signs had a precision, recall, and F1-Score of 98.3%, 92.06%, and 95.08%, respectively. The extraction of lane markings had a precision, recall, and F1-Score for the left-lane markings of 100%, 89.36%, and 94.38%, respectively. The corresponding values for right-lane markings were estimated as 93.47%, 86%, and 89.58%, respectively. The image classifier had a sample of 13,604 training images spanning 155 traffic sign classes and 10 false positive classes within Alberta. The structure is trained within a half hour on GPU and ~8 hours on CPU and produces 83.6% accuracy on the validation set. This translates to a Top-1 and Top-2 classification error of 10.35% and 3.24%, respectively. However, when applied to the original video-log images, the sliding window procedure used to apply the trained classifier to cropped image samples creates the opportunity for misclassifications across the video-log image. This reduces the accuracy of the classifier to 53.3%. Finally, the visibility assessment considered day and night-time conditions as well as the impact of consecutive placement (i.e. driver's attention fixated on only the nearest sign). This provides a discussion of how available visibility affects different driving populations and which traffic signs are most susceptible to being missed while driving.

This thesis presents a method for the expedited accurate extraction of a TSI, including location, orientation, placement, classification, and visibility. Utilizing the detail available from high-density LiDAR scans, the extractions are completed with a high degree of accuracy and with time benefits over the traditional manual methods. The contributions of this thesis include (i) proposing a method for the extraction of a TSI, (ii) assessing sign visibility, and (iii) creating Canada's first traffic sign image database. This sets the stage for continued research into the extraction of TSIs, the continued development of a traffic sign image database for Canada, and guidance for industry professionals considering using LiDAR or video-logs for creating a TSI.

PREFACE

Work presented in this thesis is either published or is under-review for publication.

PEER-REVIEWED CONFERENCE PUBLICATIONS

1. **Karsten**, **Gargoum**, and El-Basyouny (2019) "*LiDAR-based Assessment of Highway Traffic Sign Visibility*". 98th Transportation Research Board Annual Meeting, Accepted October 2018.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor, Dr. Karim El-Basyouny, for the opportunity to begin my research career during my undergraduate degree and continued guidance throughout my graduate program. The opportunity to explore my research interests and develop my career aspirations will forever impact my career as a Civil Engineer, and for that I am grateful.

I would also like to thank the examination committee, Dr. Bipro Dhar, Dr. Douglas Tomlinson, and Dr. Karim El-Basyouny, and the committee chair, Dr. Evan Davies, for their time and insightful questions and comments during the review of my thesis.

Furthermore, I would like to thank Natural Sciences and Engineering Research Council of Canada for their financial support of my research during the latter half of my time as a graduate student. I would like to extend thanks to Alberta Transportation as well for providing the data used within this thesis, because without it my exploration into this research would never have begun.

The start of my career as a Transportation Engineer was heavily influenced by the current and former members of Institute of Transportation Engineers (ITE) U of A student chapter and the Northern Alberta Chapter of the Canadian ITE. I would like to thank them for their continued support and eagerness to bring the world of transportation engineering into the student eye.

I would like to thank Suliman Gargoum for his support, guidance, and the occasional push to challenge myself to be a better person and better researcher. I would like to thank my friends for their encouragement and for providing a listening ear when I need it.

I want to thank my Mom and Dad for their endless love and for instilling in me the idea that, with a little hard work, you can accomplish anything you set your mind to. I would not have been able to work through the life changing decisions and countless hours of classes and research without your unconditional support. I would also like to thank my three siblings for their support. Although I know you are all still not completely sure what I do on a regular basis, I have greatly appreciated your support and I look forward to the honor of supporting you throughout your careers as well.

Finally, words cannot describe how thankful I am to have the support of my girlfriend, Gabbie, for her unending love and support throughout my graduate degree. You believed in me even when I did not believe in myself, and for that I am forever grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Expansion
AASHTO	American Association of State Highway and Transportation Officials
BTSC	Belgian Traffic Sign Classification
CTSP	candidate traffic sign points
CPU	central processing unit
CNN	convolutional neural network
CDF	cumulative density function
DBM	deep Boltzmann Machine
DNN	deep neural network
DBSCAN	density based clustering for applications with noise
EDEC	edge detection and edge constraint
FRMP	fake road marking points
FN	false negatives
FP	false positives
GMM	Gaussian mixture model
GMMEM	Gaussian mixture model with expectation maximization
GRF	geo-referenced feature
GTSRB	German Traffic Sign Recognition Benchmark
GNSS	global navigation satellite system
GPU	graphical processing unit
HOG	histogram of oriented gradients
IMU	inertial measurement unit
LISA	Laboratory for Intelligent and Safe Automobiles
LiDAR	light detection and ranging
MUTCD	Manual for Uniform Traffic Control Devices
MLD	minimum legibility distance
MLS	mobile laser scanning
MCDNN	multi-column deep neural network
MSTV	multi-scale tensor voting
NCHRP	National Cooperative Highway Research Program
PC	passenger car
PCQI	patch-based quality index
PRT	perception reaction time
PCA	principal component analysis
PDF	probability density function
ReLU	rectified linear units
SSD	stopping sight distance
SVM	support vector machine
TSI	traffic sign inventory
TAM	transportation asset management
TN	true negative
TP	true positive

1 INTRODUCTION

1.1 BACKGROUND & MOTIVATION

The key to the effectiveness of any transportation infrastructure, whether it be for pedestrian, cyclist, transit, commuter, or commercial vehicle traffic, is the inventory and maintenance of transportation assets. Otherwise known as Transportation Asset Management (TAM), this process is applied to maintain a cost-effective, safe, and efficient transportation network [1]. Fundamental to any TAM system is reliable, accurate asset management data. Combined with an effective data management system, a complete asset management portfolio allows for a better understanding of the transportation network, resulting in decreased lifecycle cost and increased level of service (i.e., performance) of assets for users [2]. This ensures the installed assets adhere to the standards set by transportation officials (i.e., to assess the performance of assets) and provide the supply necessary to meet the demands of current and future populations.

Specific to transportation engineering, an efficient asset management network would include descriptive detail of the network-level assets, including signal heads, light fixtures, traffic signs, bridge structures, roadway geometric features, etc. However, no asset is more fundamental to the driving environment than traffic signs. Signs are a physically small but incredibly important asset of any transportation network, utilized by engineers to communicate critical information to road users. Traffic signs are used to provide information about traffic control, upcoming geometric changes, navigation, and work-zone traffic guidance. These signs ensure the safe and efficient use of any transportation network. As such, a comprehensive TAM inventory of traffic signs is of utmost importance.

Canada has 1.13 million lane-kilometres of two-lane equivalent roadway, all of which require traffic signs to inform the driving population of the road environment ahead [3]. This results in a significant quantity of traffic sign assets that must be placed and maintained to meet the needs of the driving population. Although the intention of previous legislation was to unify traffic sign and signal design, many jurisdictions utilize their own transportation infrastructure guidelines. In North America, the Manual on Uniform Traffic Control Devices (MUTCD) governs the use of traffic signs on highways and provides general recommendations on where traffic signs should be located [4]. Given the vast quantity of signs required to guide global traffic movements, municipalities look to traffic sign inventories to inform them of the current location, status, and history of the traffic signs along any given road segment.

Due to their extreme volume, inventories are not commonly available, as is the case for traffic sign assets present in Canada. A study was conducted by the United Kingdom Department of Transport which discusses the growth seen by developing countries in the volume of their traffic sign assets. Their study estimated the volume of traffic signs placed in the United Kingdom in 2013 [5], combining data from 19 local authorities to create a model to estimate the number of traffic signs in the country. Compared against a previous study in 1993, the number of traffic sign assets available on UK road networks increased by over 110% for the sign groups that could be compared; placing increased stress on their regional asset management systems. Although changes within Canada and the United Kingdom are not directly comparable, the total built-up area for Canadian census metropolitan areas has increased 157% from 1971 to 2011, highlighting the significant increase in land used by people in Canada which are connected by roadways [6].

Assessing different elements of the expansive road network and ensuring these elements meet design requirements that satisfy the needs of drivers is an integral step to creating a safe and efficient driving environment. In the case of traffic signs, the amount of information on a traffic sign, its location on the highway, and whether it meets legibility requirements or not are extremely important attributes that determine a sign's ability to efficiently convey critical information to the drivers in time for them to take appropriate action. These attributes become even more critical when considering statistics showing that the median age in developed regions around the world is expected to increase to as high as 45 years by 2050, increasing the number of drivers with potentially limited capabilities [7].

Additionally, in recent years, roadway design has shifted the responsibility of improving road safety from the driving population towards designing more forgiving highways, where road infrastructure is designed to accommodate the vulnerability of the human body as well as human fallibility [8]. To this end, the assessment of elements is especially important considering that "*Although the MUTCD specifies the general location of large roadside signs, the highway design engineer has a significant degree of latitude in the exact placement of any given sign.*" [9]. The design engineer may place a sign based on their judgement, providing an element of human input not accounted for in design guides. Given the combination of the expected population shift and the human component to traffic sign installation, the placement and visibility of traffic signs along a highway must be constantly assessed to ensure that drivers are able to see and react to those signs in a timely manner.

The current practice of collecting measurements on traffic sign placement can be primarily manual, as shown in Figure 1.1. This process is applied to each traffic sign along a network, requiring a field technician to stop along the roadway, likely blocking traffic from using the right-most lane to protect the technician and measure the location, orientation, and lateral and vertical placement of each sign. Not only is this time consuming and inefficient on a large scale, this process also places the technician at risk. This requires personnel on foot conducting traffic sign measurements for inventory purposes, putting their safety at risk as they stop in or disrupt traffic to conduct these measurements. This is recognized in Sign Click [10], an in-house software developed for the Kentucky Department of Transport that uses manual input of a sign's condition, mounting type, etc., gathered from inspections. However, attempting to mitigate the danger of field-collected traffic sign inventories (TSI), Sign Click specifies that "*Data should be collected safely … [and] shoulders should be utilized as much as possible. If the road has no shoulders then either a median or the right lane must be used as to let the flow of traffic continue normally. On two lane roads, make sure that your vehicle is visible to the drivers and allow them to pass when oncoming traffic is clear.*"

Figure 1.1 – Manual Traffic Sign Data Collection [11]

As mentioned previously, manual measurement processes are time consuming, greatly reducing the frequency at which traffic signs can be inventoried on a network level. Attempts have been made to digitize TSI processes with the introduction of video and photo-logs and site-visit GPS stamps at sign locations. For example, Barg et al. developed CityPoints, a web application for the development of a Google Street View based TSI [12]. The application is embedded onto the Google Street View application, allowing users to manually classify traffic signs and calculate their global coordinates from Google's extensive roadway image database. The application uses multiple angles of a traffic sign to triangulate the coordinates, which are verified for accuracy using field-collected data. GPS readings could only guarantee a positional accuracy of 5 metres, but this is further reduced in dense urban areas. Results were overlain on a satellite image to visually compare positional accuracy, with the CityPoints approach offering greater consistency in sign position than the GPS results.

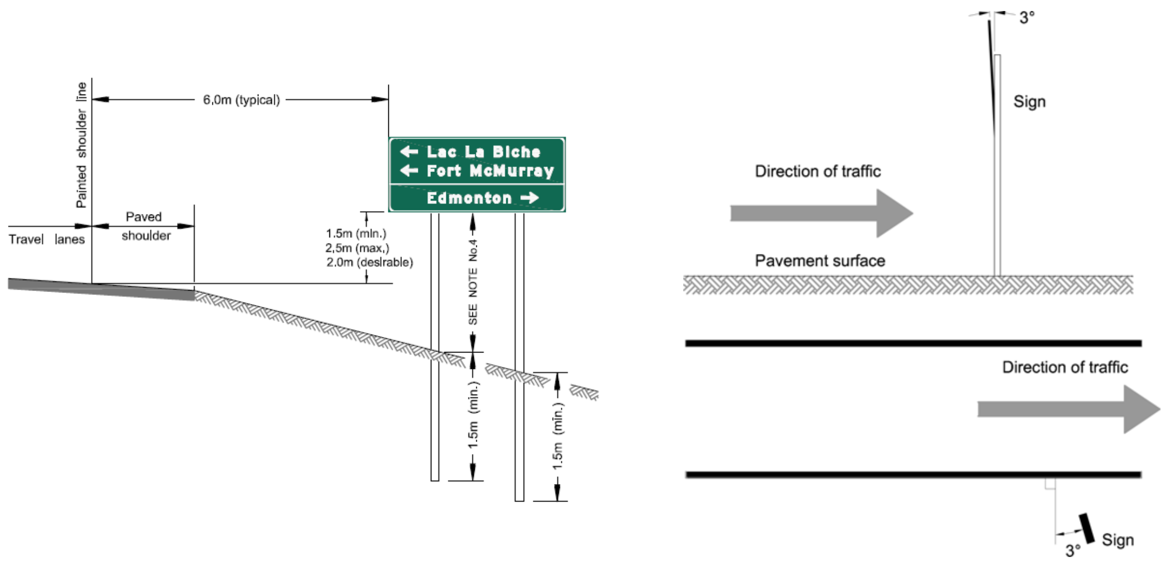
Similarly, Balali et al. [13] conducted an automated TSI from Google Street View images to remove the manual component of the assessment altogether. By interacting with the Google Street View API, a TSI was created by classifying traffic sign images into one of four categories: regulatory, warning, stop, or yield signs. Traffic signs are classified using a histogram of oriented gradients (HOG)+Color with a linear support vector machine (SVM) classifier based on previous work by the authors. They apply a multi-scale sliding window that passes over the entire image. Additionally, traffic signs are only considered detected if they are similarly classified from at least three different views. The automated TSI procedure resulted in 100% and 83.93% detection accuracy for warning and regulatory signs, respectively, and 91.96% classification accuracy.

Although the image or video-based TSI processes remove a large portion of the field-work required to collect a complete inventory, there are a few drawbacks. Firstly, the image or video-based collection of a TSI is still incredibly time consuming when conducted manually. Secondly, it is quite difficult to extract or measure additional information from the images or videos. The previous work into image or video-based methods recognizes the need for advancements into the available TSI methods, moving away from manual field measurements towards safer, more accurate methods. Undoubtably however, there is still extensive room for improvement in the creation of TSIs, both in efficiency and in the level of detail collected. The process of extracting a TSI needs a streamlined approach to conduct efficient, accurate measurements and assessments of traffic sign placements. Extracting a TSI would greatly benefit from a hands-off approach, wherein the survey of traffic signs would be collected and measured remotely. This would reduce the required field time for collecting traffic sign information and makes the surveying process safer for technical staff and other road users.

1.2 Creating a Complete Traffic Sign Inventory

To understand how the TSI can be improved, the components of a single traffic sign within a TSI should be detailed. The fundamental focuses of previous TSI procedures were locating and classifying the traffic signs along a segment. Although this information is critical to the general mapping of traffic sign assets, this leaves practitioners without information about sign placement, including orientation, lateral offset, vertical offset, and height. These characteristics are fundamental to the understanding of how signs impact traffic, hence their inclusion in regional design guides such as the MUTCD [4], Roadside Design Guide [9], and the Alberta Highway Guide and Information Traffic Sign Manual [14].

As stated in Section 1A.04 of the MUTCD [4] "*To aid in conveying the proper meaning, the traffic control device should be appropriately positioned with respect to the location, object, or situation to which it applies. The location and legibility of the traffic control device should be such that a road user has adequate time to make the proper response in both day and night conditions*." Therefore, in order to assess the correct placement of a traffic sign to maintain a safe and efficient roadway network, a TSI should consider the placement standards, as shown in Figure 1.2 (a) for a highway guide sign installation. Additionally, traffic sign placement is dependent on the local angle of the signs relative to the traffic stream. Both ground-mounted and overhead signs are tilted away from drivers to retain traffic sign visibility and reduce glare from vehicle headlights. The Alberta Highway Guide and Information Sign Manual [14] provides general orientation specifications for traffic signs - tilted 1°-3° from perpendicular unless otherwise stated. This is described in further detail in Figure 1.2 (b).



(a) (b)

Figure 1.2 – Alberta Traffic Sign Guide Lateral, Vertical, and Angular Placement [14]

Placement characteristics are critical because the reflectivity of traffic signs is directly related to the sign's position [15]. The correct positioning of shoulder-mounted and overhead signs ensures the correct amount of light is reflected from the headlights back towards the driver to maximize the conspicuity of traffic signs. Failure to mount signs in the correct position and orientation can result in blinding oncoming traffic or reduced conspicuity of the traffic signs. Both scenarios are detrimental to the safety of passing traffic and to the effectiveness of the traffic signs. Monitoring the placement conditions within a TSI ensures mounted traffic signs meet geometric design standards and thus, the needs of the driving population.

The fundamental TSI characteristics outlined by the previously mentioned geometric parameters define the characteristics of a standard TSI. In addition to the standard TSI, the measurement of traffic sign visibility is an additional indicator of sign placement efficacy. In case of any environmental irregularities such as steep cross slopes and cluttered clear zones, the placement of traffic signs is left up to engineering judgement [8], [16]. Given the variability of alignment and geography along a roadway network, situational circumstances and engineering judgement can lead to inconsistent placement of traffic signs, which might impact effectiveness. Therefore, the measurement of the first location at which a sign is visibile will determine if the placement meets the needs of the driving population.

Legibility, a subset of visibility, is defined as the maximum distance at which the smallest detail on a sign can be perceived [17]. This is defined as 30 feet per inch, or 3.6 metres per cm, of letter height in the MUTCD [4]. The MUTCD does not provide any specific guidance for symbols, but Castro and Horberry [17] defined the legibility distance for symbols as 250 metres. Realizing that a single measure might not be truly representative of different sizes of symbols, other research provides a more comprehensive legibility distance of 6.9 m per cm of symbol height [18]. Figure 1.3 highlights the legibility distance for a vehicle with a traffic sign on the left side of the travel lane.

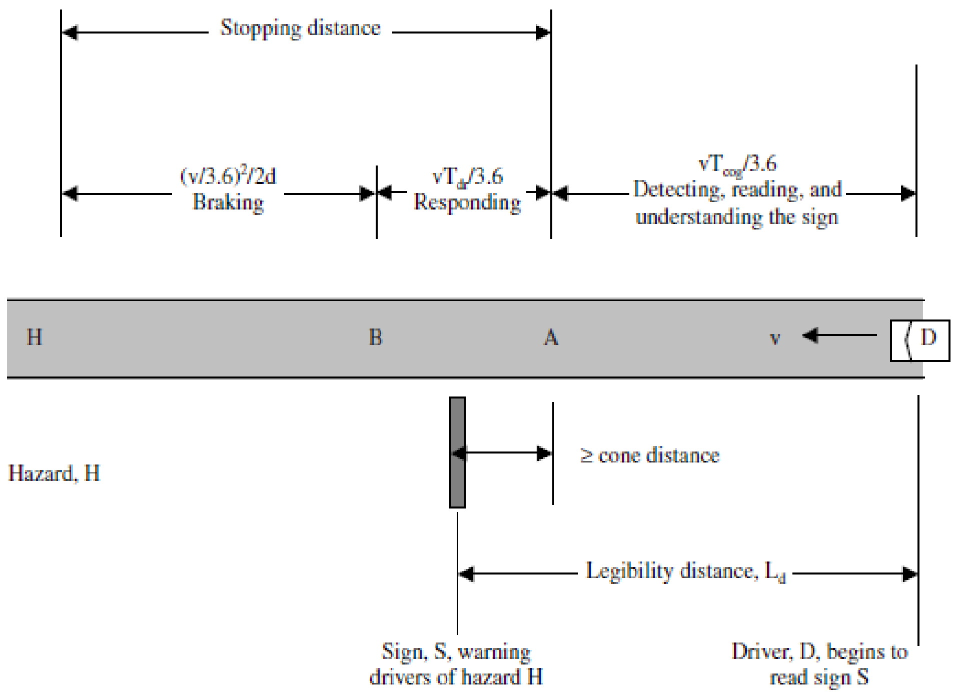


Figure 1.3 – Legibility Distance of Traffic Signs [17]

The addition of visibility completes the information required from a TSI, but a full inventory of traffic signs is difficult to complete from manual, image, or video-based inventory methods. One technology that could be used to efficiently create a TSI is mobile Light Detection and Ranging (LiDAR) technology. Using information on reflection time, energy, and intensity, the positional information of points on specific objects is computed. LiDAR scanning creates 360°-virtual point cloud data of the highway. To collect LiDAR data, Mobile Laser Scanning (MLS) equipment is mounted on a vehicle that collects georeferenced point clouds while travelling at highway speeds. This causes minimal disruption to traffic and increases the efficiency of the data collection process. Creating TSIs using mobile LiDAR has gained traction in global research in recent years [19]–[21].

The use of mobile LiDAR for asset management in the transportation industry is also supported by the National Cooperative Highway Research Program (NCHRP). The Guidelines for the Use of Mobile LiDAR in Transportation Applications [22] was published to provide suggestions for use of mobile LiDAR in the current transportation industry. This includes the application of LiDAR to all components of a project, including planning, development, construction, operations, asset management, maintenance, safety, and future research and advancement. Furthermore, guidelines are provided for data collection methods and translating results into industry products, key to the successful integration of this up-and-coming technology into existing asset management and analysis workflows [22]. Therefore, the industry-recognized method of mobile LiDAR scanning may provide the efficiency and accuracy needed to complete a TSI.

1.3 Research Problem Statement

By creating a complete TSI, the regional transportation authority can update its understanding of the inventory and placement of traffic signs serving today's traffic. This information can be incorporated into regional maintenance plans to ensure traffic sign panels are up-to-date and effective for the current driving population. Additionally, the availability of a TSI eases the burden on industry professionals or contracting firms when conducting work on regional highways. An inventory allows for better contractor preparedness before going on site, providing information on traffic signs to safely plan work zones.

To discuss the application of LiDAR for the extraction of a TSI, this thesis is broken into four primary components; namely, traffic sign extraction, lane marking extraction, image-based traffic sign classification, and traffic sign visibility measurement. The traffic sign extraction component locates the signs along a mobile LiDAR scanned segment, clustering candidate sign panel points and measuring sign panel orientation. The lane marking extraction component is used to determine traffic sign placement, as the placement is measured from the nearest lane marking. The traffic sign classification is image-based, utilizing a neural network structure to classify traffic signs as one of the 155 available traffic sign classes. The development of the traffic sign dataset used to train the image classifier is the first of its kind in Canada. Finally, the traffic sign visibility component serves to measure the first point along a segment at which each traffic sign is visible. To assess the applicability of mobile LiDAR to the creation of a TSI, the measurements and extractions are validated, where possible, using manual methods. The algorithm development process and TSI extraction outcomes are highlighted in Figure 1.4.

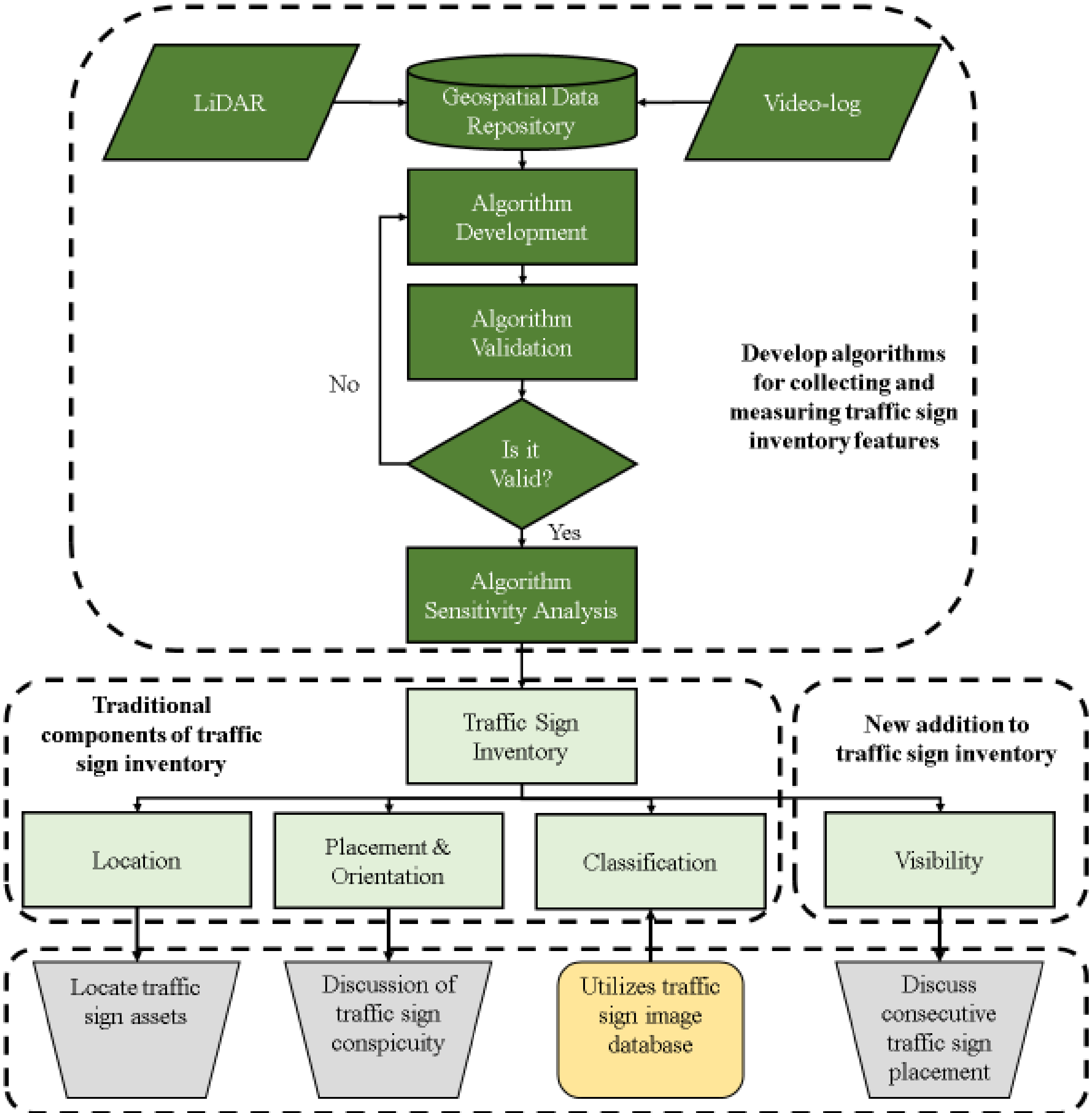


Figure 1.4 – Algorithm Development and TSI Extraction Workflow

In summary, this research collects a sample traffic sign inventory within Alberta, Canada, applying the fusion of mobile LiDAR scanning and video-log imaging through the following research objectives:

1. Creating a traffic sign inventory by extracting information on sign location (i.e., latitude and longitude coordinates), sign placement (i.e., vertical and lateral distance), and sign orientation (i.e., yaw and pitch of traffic sign panel). The addition of the orientation and placement measurements to the traffic sign inventory allow practitioners to assess the general traffic sign placement in greater detail. Sign conspicuity is highly dependent on these parameters, and their accurate and timely measurement from a reliable data source allows for the discussion of adherence to placement standards and the effects of non-compliance.
2. A link between LiDAR and video-log imaging in the absence of information; for traffic sign classification. The link between LiDAR and video-log images creates the opportunity to provide LiDAR-based assessments with additional detail about the surrounding area from another data source. In this case, this link will be utilized to classify the traffic signs which exist along a test segment, otherwise difficult from the LiDAR data alone.
3. Traffic sign visibility assessment (i.e., the distance of first available visibility). The visibility assessment is a new addition to the creation of a TSI, providing practitioners with an understanding of the locations along a roadway at which different traffic signs are visible to the driving population. This allows for an assessment of traffic sign visibility compliance and a discussion of the effects of consecutive sign placement on sign visibility.
4. The creation of Canada's first traffic sign dataset. To conduct a neural network-based classification of traffic sign images, this thesis creates the first traffic sign image database within Canada. Spanning 155 traffic sign classes with 12,315 images, this database creates the opportunity for continued research into traffic signs within the Canadian context and for applications to the transportation industry in Canada.

1.4 Thesis Structure

The research is split into four main components: sign extraction, lane marking extraction, sign classification, and visibility. This is done to first locate the traffic signs and determine their orientation. The lane markings are then extracted as they are needed for placement measurements. The classification is considered separately from the placement measurements and utilizes an image-based approach. Finally, the visibility assessment, which is dependant on the sign extraction, completes the creation of the TSI. Consequently, the thesis starts with Chapter 1 which provides an introduction into the need for a TSI and its components. Additionally, the introduction has described the setting of the research presented in this thesis.

Chapter 2 conducts a review of the available literature for the four components of the TSI - LiDAR-based sign extraction, lane marking extraction, and visibility assessment. Additionally, sign classification is reviewed through LiDAR-based and image-based methods.

Chapter 3 describes the methodology utilized within this thesis as applied to the four components. The traffic sign and lane marking extractions both utilize the intensity component of the LiDAR data to locate these assets. Their extraction is utilized to determine traffic sign orientation and placement. Sign panel classification is conducted through an image-based process, where a manually created set of 13,604 training images across 155 traffic sign image classes and 10 false positive image classes is used to train a convolutional neural network image classifier. Finally, the traffic sign visibility assessment creates sightlines between traffic signs and trajectory points to determine the maximum distance at which each traffic sign is visible.

Chapter 4 contains results and discussions of the application of the proposed methods to create a TSI. Where applicable, confusion matrices are used to quantify the accuracy of asset extractions, with discussions of the cause of false or missed extractions. This chapter also briefly discusses processing time the processing time of the proposed methods, discussing the time advantage of a remote sensing-based method. Finally, Chapter 5 completes this thesis with a discussion of the contributions, limitations, and future research for the processes outlined in this thesis.

2 LITERATURE REVIEW

Traffic sign regulations were first introduced with the International Convention on Motor Traffic in 1909 [23]. This Convention was commissioned to regulate the essentials of road travel, including the construction of motor vehicles, international travel, and traffic signs and signals. However, this Convention and a series of subsequent international updates throughout the first half of the 1900s still did not fully regulate traffic signs and signals. With the increase of international travel, economic cooperation, and the growing need for a unified set of international standards, the Convention of Road Signs and Signals was held in Vienna, Austria in 1968 to provide uniformity amongst traffic signs and increase international traffic safety [23]. Annex 1 of the Convention defines eight classes of traffic signs, including (A) danger warning signs, (B) priority signs, (C) prohibitory or restrictive signs, (D) mandatory signs, (E) special regulation signs, (F) information, facilities, or services signs, (G) direction, position, or indication signs, and (H) additional panel signs. Since 1968, the Convention has seen further updates to continue unifying the regulations surrounding traffic signs, signals, and pavement markings, with the most recent revision dating 2006 [23].

The components of a TSI can be described as local geometric characteristics (i.e., size and shape) and global geometric characteristics (i.e., lateral and vertical placement) required to be measured from the nearest lane marking for highway applications. These characteristics will be reviewed considering a LiDAR-based extraction. Additionally, image-based traffic sign classification will be reviewed to determine the state-of-the-art in image recognition. Finally, the additional application of traffic sign visibility will be considered for the sign inventory. The flexibility of large-scale survey data like the LiDAR data used in this study allows for the consideration of traffic sign visibility measurements.

2.1 TRAFFIC SIGN EXTRACTION

The current industry practice for traffic sign inventory is tedious and manual. Certain processes have been created with technological advancement, including the use of still images and video feeds of a vehicle in motion, but are still limited in their ability to conduct recurrent traffic sign inventory. Additionally, the previously highlighted examples require manual measurement of additional attributes about the traffic sign, like placement and dimensions. LiDAR data, with the capability of millimetre-level accuracy and widespread data collection capabilities, represents an alternative to inventory signs. The following details the current state of research regarding LiDAR-based sign extraction.

Gargoum et al. [20] proposed a traffic sign extraction methodology for LiDAR data collecting using a REIGL VMX-450 dual scanner. The multistep process includes (i) retro-intensity filtering, (ii) 3D density based spatial clustering for applications with noise (DBSCAN), and (iii) geometric filtering. DBSCAN was applied using a minimum hit count of 17 points and a spacing of 1.0 metres. Traffic signs were filtered by height considering the regional minimum sign height. Furthermore, GIS software was used to create buffer zones both along the pavement surface and at a large offset from the centreline of the road. The proposed method was tested on three different highways totaling 12 kilometres, with a minimum detection rate of 93.4%.

Ai and Tsai [24] conducted LiDAR traffic sign extraction through their proposed multi-filter process. The LiDAR data was collected using a REIGL LMS-Q120i with a scan collection rate of 10,000 points per second. The LiDAR data is subject to a retro-intensity filter, an elevation filter, and a lateral offset filter before finally being clustered by distance and filtered for hit count. Filters were assessed for maximum detection accuracy through sensitivity analysis, resulting in optimized thresholds for urban and interstate scanning scenarios. The optimized parameters are summarized in Table 2.1. Overall, this method reported a 91.4% detection rate with seven false positives (FP).

Table 2.1 – Geometric and Intensity Threshold Values [24]

Parameter	**Interstate**	**Local Urban**
Retro-intensity [%, (16-bit)]	0.70 (45,874 as 16-bit*)	0.65 (42,597 as 16-bit*)
Traffic Sign Elevation [metres]	2.13	1.83
Minimum Lateral Distance [metres]	2.13	0.61
Maximum Lateral Distance [metres]	17.98	6.71
Minimum Hit Count (at 96.5 km/h operating speed)	10 points	20 points

**NOTE: intensity percentage values were converted to 16-bit intensity value to compare against intensity thresholds noted in other publications.*

Soilán et al. [21] proposed a traffic sign extraction methodology consisting of (i) point cloud preprocessing, (ii) ground and non-ground extraction and intensity filtering, (iii) geometric filtering and distance-based clustering, and (iv) projection onto RGB images for traffic sign class recognitions. The point cloud was filtered at 20 metres from the trajectory, as this defines the limit of the region of interest for urban traffic sign extraction conducted here. Points are then rasterized at 0.5-metre cell spacing, with rasters created for point count, height, average intensity, accumulated height, and vertical variance. An additional raster (I_{height}) consisting of the accumulated height weighted with the vertical variance is created. The $\mathrm{I_{height}}$ and intensity image results were combined through a Boolean AND operation in a coarse ground filtering process leaving only high intensity non-ground points. The remaining points represent traffic sign candidates, but also contain reflective poles, walls, and facades. Therefore, a Gaussian mixture model (GMM) with two components is applied to segment the higher intensity points to extract the traffic signs. The resulting points are clustered based on distance using DBSCAN, where *MinPts* = 25 and *Eps* = 0.2 metres.

To determine the geometric orientation of each candidate traffic sign cluster, principal component analysis (PCA) is applied to filter cluster flatness. Finally, a simple height filter is applied to remove clusters less than 25 cm tall, as these cannot represent traffic signs. Additional attributes are determined to describe the position of each candidate traffic sign, including (i) traffic sign position (centroid), (ii) traffic sign height (measured from centroid to ground), (iii) the distance between scanning trajectory and traffic sign, (iv) the angle between the traffic sign and the scanning trajectory, and, (v) if the sign is pole mounted, the inclination of the pole in profile and plan views of the traffic sign.

Riveiro et al. [25] outline three main components for traffic sign detection: (i) segmentation, (ii) clustering, and (iii) feature recognition. First, an intensity map (i.e., a projection of the LiDAR points onto the horizonal plane) is created. Then, coarse thresholding is applied to extract the highly reflective surfaces (i.e., traffic signs) from the intensity map. These pixels are mapped back into LiDAR points to segment the signs from the point cloud. Finally, an optimized threshold is computed using the segmented LiDAR points and an optimal value is obtained using a GMM. The GMM is fit comprising of four Gaussian curves, separating the traffic sign panels and poles detected from the rough intensity segmentation. The two Gaussian curves with the highest areas are taken to be the traffic sign panel points, with the optimized threshold set as the mean minus two times the standard deviation for the lower of the two Gaussian curves. This encompasses 95% of the points in the lower Gaussian distribution and retains points belonging to the higher Gaussian distribution. To remove other high intensity objects (car lights, lamp posts, etc.) after filtering, DBSCAN and curvature analysis are also conducted. DBSCAN was conducted using a minimum number of 50 points and a search radius of 0.2 metres - to avoid clustering license plates. Geometric-based filter was applied to remove points belonging to road structures (car lights, lamp posts, other reflective objects).

Deng and Zhuo [26] proposed a traffic sign detection framework which considers the use of colorized LiDAR scans. Planar objects are detected through the creation of aggregation-based feature vectors, where RGB, HSV, and CIELab color spaces are combined with LiDAR intensity. Using *a priori* knowledge of traffic sign 3D planar geometry and correcting for perspective deformation, sign recognition is completed using a linear SVM with HOG features classifier. The authors noted difficulty with traffic sign detection in real-world situations where occlusion, lighting changes, perspective distortion, and weather may obscure signs.

Arcos-García et al. [27] proposed a traffic sign detection procedure that considers the fusion of LiDAR scanned point clouds and recurrent driver-perspective images. The point cloud is first pre-processed by (i) removing points greater than 15 metres from trajectory (not of interest for traffic sign extraction), (ii) conducting ground extraction, and (iii) conducting ground region growing, where ground seeds are selected using a nearest neighbor search of the trajectory, where voxels with vertical mean and variance differences less than 0.1 m and 0.05 m (empirically tuned) compared to kNN trajectory voxels are clustered and used as seeds to grow the “ground region”.

The non-ground region is then used to begin extracting traffic signs. The authors propose an unsupervised classification algorithm based on a GMM with two components. All non-ground points are assigned to one of two GMM classes and the class with the highest mean (i.e., closest to 65,535; maximum intensity) is selected for further processing. The remaining points are grouped using DBSCAN, followed by principal component analysis to filter for cluster flatness, and finish with a cluster height filter identical to Soilán et al. [21].

Wen et al. [28] proposed a traffic sign detection process through terrain point filtering, linear structured objects detection, and reflectance intensity-based sign surface filtering. Terrain points are gridded in XY, then for each non-empty grid, points are selected within a percentile for elevation. The representative points within the grid cells neighboring the current cell are used to fit a local plane, assigning the points outside of a bounding box around the plane as “off-terrain”. This is repeated by assigning points to one of four sub-grid cells and segmenting until the finest level is reached.

Euclidean clustering is applied to segment off-terrain points into clusters, followed by an eigen-decomposition of the covariance matrix to calculate the principle directions within each cluster. Linear structures are identified considering $\lambda_1 \gg \lambda_2 \cong \lambda_3$ and $\lambda_1/\lambda_2 > 10$, where λ_i is the principal component in the i^{th} direction. Additionally, clusters that contain less than 50 points are removed. Traffic signs are selected by their retroreflective properties as clusters with intensity greater than 60,000. Road surface and road boundary are extracted through a curb-based method, which considers local difference in elevation to differentiate curbs from the road surface and other features.

Traffic sign placement is measured as the traffic sign height above ground, the distance from traffic sign to road boundary, the orientation of the traffic sign with respect to the road direction, the inclination (i.e. vertical angle) of the traffic sign board, the horizontal pitch of the traffic sign, and the planarity and curvature of the traffic sign. The scanning was conducted using a RIEGL VMX-450, reported positional accuracy of ± 0.231, ± 0.287, ± 0.442 metres in x, y, and z, respectively. Additionally, the vertical and horizontal placement and rotational accuracies are ± 0.063 and ± 0.076 metres and $\pm 52''$ and $\pm 37''$, respectively. Difficulties in accuracy were noted by the authors as follows: traffic signs were not extracted from point cloud due to incomplete scans, or if the scans were conducted in the opposite direction of a traffic sign, thereby only collecting the back face (i.e. not the reflective side, not picked up in high intensity filter).

Yu et al. [29] devised a traffic sign extraction methodology using a bag-of-visual-phrases representation of 3D points, considering both single features and distributed features in local point cloud regions. First, an upward region-growing algorithm segments the point cloud into ground and non-ground points. Points are voxelized and a set of point cloud segmentation training data is then segmented into supervoxels using the voxel cloud connectivity segmentation algorithm.

A deep Boltzmann machine (DBM) feature encoder is proposed to generate high-order features of local point cloud regions. Each non-ground query object is run through the trained DBM and assessed for similarity to the trained traffic sign point features. The three-dimensional candidate traffic sign clusters are transformed into traffic sign image coordinates to extract image information. An additional Gaussian-Bernoulli DBM is applied to conduct image-based traffic sign recognition, with 161,792 augmented image samples are used to train the Gaussian-Bernoulli DBM model. The proposed methodology resulted in 94.6% correctness, providing a unique alternative to previously proposed traffic sign methodologies. However, this method noted the need for exceptional computational power.

2.2 LANE MARKING EXTRACTION

The extraction of lane markings from remote sensing data, mainly LiDAR fused with driver point-of-view images, has become increasingly popular with the recent interest in driver assistance systems. In that context, knowledge of the location of lane markings allows for the driver assistance applications like lane keep assist, lane departure warning, and semi/fully autonomous driving systems. However, in the context of transportation infrastructure, the knowledge of the location of lane markings allows for the assessment of several transportation planning and operations requirements. This includes the assessment of lane width, lane marking visibility, and for this research, traffic sign offset. As per the Alberta Highway Guide and Information Sign Manual [14], minimum lateral and vertical traffic sign offsets are required for the safe and efficient placement of traffic signs - measured from the nearest lane marking.

Guan et al. [30] propose an image-based lane marking extraction method, where the reflectivity of the lane markings is manipulated through multiple intensity thresholds and statistical segmentation along the road cross-section. The point cloud is segmented using curb extraction to detect the road surface, and the road surface is then split into cross sections longitudinally along the direction of the roadway. Then, georeferenced feature (GRF) images are developed from the cross sections using a modified inverse distance weighting method, where pixel intensity values are determined by distance and intensity-weighting. Intensity-weighting is used locally (i.e., cross section level) and globally (i.e., segment level), where intensity value weights are determined through analyses of the GRF image point density and intensity histograms, effectively equalizing the intensity histograms. The intensity of the cross section is fit to a Gaussian normal distribution and segmented into bins, where each bin contains one standard deviation of the fit distribution. A binary intensity threshold technique is then applied to each intensity bin, segmenting each bin into foreground (i.e., lane markings) and background (i.e., road surface). Finally, the road markings are further refined through morphological closing operations.

Riveiro et al. [31] propose a scan-line and image based lane marking extraction procedure, where the roadway is segmented using the scan-lines and converted into images for the application of image processing techniques. Scan-lines (defined as "scanner cycles" in this paper) are assessed with PCA to a local neighborhood of points (10 points) to determine the "traversal limits" of the roadway. The roadway is further segmented into "strips" (i.e., cross sections) consistently 18 metres in length, collected with 50% overlap to ensure that zebra crossing paint lines are not segmented within separate cross sections. Each strip is rasterized by determining the best-fit plane to the points and assigning the pixel value to the nearest point in the cross section. Binary images are created through intensity thresholding as in Guan et al. [30]. Morphological operations of median filter (to reduce salt and pepper noise) and closing (to fill in holes within the zebra marking) are applied. The Canny edge detection process is applied to determine the edges of the binarized lane marking images, finalized with the application of the Hough Transform to detect linear segments within the extracted edges. The extracted Hough lines are meant to represent the edges of the lane markings as these are linear for zebra crossings.

Kumar et al. [32] proposed an intensity-based road marking extraction method, which utilizes automated distance-based intensity thresholding. Using the road boundary extracted during their previous study, road marking extraction is developed based on the assumption that road markings exhibit higher intensity than their surrounding road surface elements. For the extracted road surface, a series of cross sections are created and rasterized using an optimized cell size. Raster cell values are based on average values of intensity and range and normalized based on their global minimum and maximum. Range-based intensity thresholding is then applied to each intensity raster image. The trajectory of the scanning vehicle is used to select a range value to be applied for multiple values of thresholding. This results in multiple 'blocks' along the raster image at which intensity thresholding will be applied unique to that block. The initial threshold value (T_1) is set empirically and used to estimate the threshold values within all four blocks ($T_{l1}, T_{l2}, T_{l3}, T_{l4}$), assuming a two-lane road (i.e., two blocks per lane). The raster cell size was optimized through a comparison of extraction results along a 10-metre segment. Average lane marking width and length was compared to the design standards, choosing the cell size that produces the closest width and length.

Binary morphological operations and *a priori* knowledge of road marking dimensions are applied to resulting extractions to complete the road markings and reduce road surface noise. The resulting raster images are converted to binary images and linearly dilated along the direction of the lane markings to fill any holes. Connected component labelling is applied to the resulting cells and filtered using length and width thresholds. Raster cells are then eroded to revert to the original shape and size of the road markings. The erosion uses the same angled linear structural elements as used in the dilation operation. This procedure resulted in object and point detection completeness of 90.91% and 88.43%, respectively.

Guan et al. [33] utilize interpolated GRF images from three-dimensional points, utilizing weighted raster neighborhoods and tensor voting to extract road markings. This work is based on the theory that LiDAR scanning intensity is dependent on a point's distance from the scanner and incidence angle. First, a curb-based segmentation of the road surface was conducted and subsequently transformed into two-dimensional georeferenced images. These are subdivided into square sub-images and a weighted neighboring difference histogram is applied. For each sub-image, this analyzes the intensity differences within each pixel neighborhood and automatically determines a locally optimal threshold for candidate road marking pixel extraction.

$$\mathrm{d_m(i,j)} = \frac{\sum_{(u,v)\in N(i,j)} \left|\mathrm{I_{uv}^M} - \mathrm{I_{ij}^M}\right| * \left(\mathrm{I_{uv}^M} - \mathrm{I_{ij}^M}\right)}{\sum_{(u,v)\in N(i,j)} \left|\mathrm{I_{uv}^M} - \mathrm{I_{ij}^M}\right|} \qquad 1$$

where *N(i,j)* is the subset of neighborhood pixels. The value of $d_m(i,j)$ determines the location of road markings, where if $d_m(i,j)$ is close to zero then the pixel is at the centre of the road marking or road surface. If $d_m(i,j)$ is negative, then the pixel is at an internal boundary of a road marking. And if $d_m(i,j)$ is positive then the pixel is at an external boundary of a road marking.

Then, the candidate road marking pixels are further refined by applying two-dimensional multi-scale tensor voting (MSTV), which suppresses pixel noise while preserving the lane markings. A multi-threshold technique is applied to the tensor size map during MSTV iterations to remove pixels with low tensor size. Finally, lane markings are clustered together through a region growing process that segments potential lane markings into their respective parts and removes any incorrect lane markings.

Yan et al. [34] propose a scan-line-based procedure for the extraction of road markings. LiDAR scans are preprocessed, analyzed to extract road points, and then analyzed to extract road markings. First, the point cloud is organized into a series of scan lines using each point's timestamp and scanner angle. Then, the roadway is extracted through a review of the height difference between the point cloud and the scanning vehicle's trajectory and an analysis of a moving least squares line. The scan lines are then analyzed using a moving window median filter of its intensity to minimize intensity noise. The Edge Detection and Edge Constraint (EDEC) method is then applied to detect road markings, followed by refinement based on local segments and linearity features to reduce false positive extractions.

Lane markings with fake road marking points (FRMP) are analyzed perpendicular to their principal axis, where stand-alone FRMPs should produce linear segments. The neighborhood surrounding each potential lane marking point p is assessed for linearity, where linearity is defined in Equation 2 and λ_1 and λ_2 are the first and second principal components of the perpendicular section (i.e., eigenvalues).

$$\mathrm{L}_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_1} \qquad 2$$

If L_λ exceeds 0.9, the point p and its neighborhood are considered FRMPs and are removed. The final extraction was found to be 96% complete and 93% correct when the raster of the extraction was compared to the raster of the manually extracted points.

Yu et al. [35] propose a method to extract road markings from 3D point cloud rather than 2D images. First, the road surface is extracted from the point cloud using curb-based segmentation. Then, road markings are detected through a two-fold procedure (i) cross-sectional intensity thresholding and (ii) spatial density filtering. First, the extracted road surface is segmented into cross sectional blocks of equal length, and each block is split into four regions parallel to the direction of travel. Then, a multi-thresholding algorithm is applied to separate the blocks into road marking and road surface points. As a result of the reflective properties of the bituminous components of the pavement, additional noise exists alongside the desired road markings. Therefore, a spatial density filter is applied to the extracted road marking points to remove road surface points that exhibit similar intensity characteristics to the road markings. Finally, the resulting cleaned road markings are classified through distance-based clustering and voxel-based normalized cut segmentation, identifying the road markings as boundary lines, stop lines, rectangular lines, arrow markings, and other road marking types.

2.3 Traffic Sign Image Classification

As outlined in Section 1.1, the collection of images for traffic sign inventory is a commonly considered alternative to manual image classification and serves as a valuable addition to the LiDAR-based traffic sign extractions that were previously detailed. Although mobile LiDAR data excels in the accurate location and measurement of objects within a scanned area, drawing conclusions about an object strictly from visual information is dependent on the density of the point cloud collected. This is typically not a concern in urban areas as the lower speed limits allow the same scan rate to collect a higher density of points on the same object. However, at highway speeds, the density of points on any given object decreases as the exposure time of each object to the scanner decreases. Therefore, information on the color of scanned objects may be sparse, which is where the fusion of LiDAR scans and color images can fill in this information. The mobile LiDAR data used in this study did not have any assigned color attributes, making the fusion with video-log images even more important to the completeness of a traffic sign inventory.

Computer vision has been a longstanding problem and popular topic within global research since the late 1960's [36], but recent advancements in computer hardware technology and the rise of semi-autonomous and autonomous vehicles have both heavily contributed to the recent resurgence in computer vision research. The challenge of general object classification from images has been of focus in the industry, with research teams developing vast image databases resulting in the annual ImageNet object classification challenge from 2013 onwards. Ultimately, the development of these databases resulted in purposefully designed modelling techniques for object recognition, including images of traffic signs.

As a result of the Vienna Convention and resulting updates, traffic signs across the Europe are common to many European Union countries [23]. Alternatively, the Laboratory for Intelligent and Safe Automobiles (LISA) dataset is an annotated dataset of driving perspective, where images that contain a specific traffic sign are annotated as such [37]. This focuses on traffic sign extraction, given the American driving perspective, and although some of the traffic signs used in the United States are similar to those used in Canada, there is a lack of specific regions of interest around only the traffic sign making this dataset difficult to use for this application. However, if a similar dataset existed for the Canadian context, the perspective provided by these images would be comparable to the perspective of the images collected during LiDAR scanning.

To conduct traffic sign classification from images, the characteristics of a traffic sign must be defined, either visually (color, shape, etc.) or through a comparison against examples of traffic signs (image logs, videos, etc.). Initial attempts at traffic sign recognition considered the color and shape of traffic signs, commonly using color transformations and image gradients to roughly classify traffic sign super-classes [21], [25]. However, with the recent resurgence of convolutional neural network (CNN) image classifiers, having a complete image dataset available alongside a stable CNN allows for greatly improved image classifiers capable of high levels of detail. For use in CNN classifiers, several traffic sign datasets have been made available. The problem of image-based traffic sign recognition has been longstanding [36], and the use of geometric attributes and two-dimensional representations of the potential traffic signs have been used in preliminary attempts to classify traffic signs from three-dimensional point cloud data.

Riveiro et al. [25] classified rasterized LiDAR traffic sign clusters into super-classes based on the goodness of fit of a sign's shape to a set of polynomials. Super-classes of danger, give way, prohibited or obligation, and indication, taken to be directly related to traffic sign shape classes of triangle, inverted triangle, circular, and rectangular. The curvature of LiDAR clusters was calculated using PCA, only retaining planar clusters as candidate traffic signs. Using the third principal component (which describes the direction of least variance, i.e., the face of the sign), all clusters were rotated to align them for rasterization. The shape of the signs is estimated based on the point distribution within raster images. Pixel size is chosen depending on the scanning location (urban areas are lower speed, therefore a higher resolution of points requiring a lower pixel resolution).

The gradients of the raster images are used to analyzed individual traffic sign clusters, splitting multiple traffic signs if they are present (common with pole mounting in urban areas) and assessing the contour of detected traffic signs. Polynomials are fit to the edge of the traffic sign raster images, allowing traffic signs to be distinguished by the degree of their polynomial as rectangular (degree 0), triangular (degree 1), rhomboid (degree 1 with horizontal axis of symmetry), circular (degree 2), and octagonal (degree 2 with additional constraints). The authors noted that distinguishing between circular and octagonal signs was often difficult due to their structural similarity. Data was collected with an Optech LYNX scanner, resulting in final completeness and correctness rates of 92.11% and 93.96%, respectively.

Soilán et al. [21] conduct image-based traffic sign recognition using color transformations, where feature extraction is conducted using a Histogram of Oriented Gradients and passed through a Support Vector Machine classifier for super-class recognition. Seven super-classes define the different traffic sign types; prohibition, danger, give way, no entry, stop, indication and obligation. For each input image, red and blue color maps are computed. Transformed from RGB to the Hue Luminance Saturation color space. Regularized logistic regression is trained to correctly classify color in the bitmap images in a variety of lighting conditions. Shape is classified for each image using HOG and SVM. Each traffic sign is also classified within its superclass for the general classes (i.e., prohibition, danger, indication, and obligation). The CIELab color space was used within an additional HOG and SVM pair for classification. Negative training images are provided from random background (i.e., surrounding environment) images and image samples from other sub-classes in the same superclass. In the urban and rural contexts, this method had 97.2% and 94.57% precision and 81.4% and 82.19% recall, respectively.

However, these classification procedures only provide unfinished classifications, partially due to the lack of available training data against which to compare potential traffic signs. In 2011 and 2014, respectively, the German Traffic Sign Recognition Benchmark (GTSRB) [38] and the Belgian Traffic Sign Classification dataset (BTSC) [39] were collected to aid traffic sign recognition research. The GTSRB contains 50,000 images labelled across 43 classes and was used in a traffic sign recognition competition won by Cireșan et al. [40]. The BTSC contains 4591 training images and 2534 test images across 62 traffic sign classes. Both datasets have been used to train and evaluate the performance of traffic sign recognition computer vision algorithms.

Cireșan et al.[40] trained a multi-column deep neural network (MCDNN), where multiple DNN models are trained and the output activations are averaged to produce the final classification result. Each deep neural network (DNN) contains nine layers, with 25 separate DNN columns trained based on one of five different image contrast enhancement and normalization techniques each randomly initialized five times. With 39209 training images and 12630 testing images the 25 column MCDNN obtained 99.46% accuracy. The five input image types are:

1. Original image; no transformation.
2. Image adjustment; increases image contrast by saturating the top and bottom 1% of pixel intensities.
3. Histogram equalization; pixel intensity transformation, which results in a roughly uniform pixel intensity histogram.
4. Adaptive histogram equalization; local version of histogram equalization, where a non-overlapping window 6x6 in size is passed over the image and the pixels are contrast enhanced – also producing a roughly uniform histogram.
5. Contrast normalization; enhanced edges within an image using a 5x5 difference of Gaussians filter.

Arcos-García et al. [27] utilized a mixture of GTSRB, BTSC, and Spanish traffic signs to train a traffic sign classifier in Spain. The combination of the three labelled image sets contained 44130 training images and 15345 validation images for 83 classes. The image sample was an imbalanced dataset, with 9/83 categories in the training set and 21/83 categories in the validation set containing less than 10 samples. 17/83 categories contain more than 100 training samples. The deep neural network used to classify traffic sign images is designed around convolutional and spatial transformer layers. The spatial transformer network layers are used to ensure the CNN is spatially invariant to the input images. Spatial transformer networks are applied to circumvent the need for additional data augmentation, normalization, or training supervision. Images are contrast normalized within the CNN using contrast normalization layers, which apply Gaussian kernels in local regions.

The authors detect traffic signs within their video-log images by transforming their LiDAR extracted traffic signs into the image coordinate space. Once every LiDAR point has been projected onto an image, the bounding box, which is inflated 25% to account for calibration errors and to add background detail, is extracted. Both the LiDAR scans and the video-log were collected using a LYNX Mobile Mapper, based on the REIGL VMX-450 scanner. This creates a region-of-interest within the image, which serves as the input to their trained NN classifier.

The training dataset contains images obtained from mobile scanning, combined with the German [38] and Belgian [39] datasets, which contain traffic signs that are similar to those in Spain. Only image classes with seven or more samples were used in the initial dataset. Traffic signs collected in Spain comprise of 897 training images and 452 validation images over 43 classes. Combination with GTRSB and adapted BTSC resulted in 44,130 training images and 15,345 validation images.

2.4 TRAFFIC SIGN VISIBILITY ASSESSMENT

The assessment of legibility distance for traffic signs has been sparingly studied in the literature, with most studies simulating occlusions that occur along traffic sign sightlines. The literature considers occlusion management scenarios, where different types of occlusions are tracked, and dynamic occlusion scenarios, where occlusions are tracked when passing heavy vehicle traffic.

Dahlstedt and Svenson [41] analyzed different traffic sign characteristics for their effects on legibility distance. Traffic signs were placed along a segment and drivers were instructed to audibly note when they saw traffic signs, and on which side of the road. A passenger in the back seat of the vehicle then recorded the distance to the sign by reading distance markers only visible to them. This procedure required 1-5 hours of driving per subject for the 12 subjects and resulted in 72 measurements of traffic sign legibility distance.

Baek [42] examined legibility distance by analyzing distance-tagged digital photologs of real highway environments. Images were collected by a survey vehicle equipped with an array of high-resolution cameras. Line of sight was manually assessed for the collected traffic sign images, with the author considering traffic sign occlusions as binary events; either a traffic sign is significantly obstructed, or there is no obstruction. Line of sight measurements in the unobstructed case were limited by the resolution of the images. Line of sight was determined for 3142 traffic signs, with some traffic signs being skipped because the fell out of frame on horizontal curves or line of sight could not be measured in urban close-quarters. The authors determined that road classification, sign size, and type of obstruction had a significant impact on the line of sight to obstructed signs.

Huang et al. [43] utilize a point cloud to extract traffic signs and analyze their occlusion from trajectory points. The proposed method conducts traffic sign extraction considering intensity and geometric properties. Then, the degree of occlusion of the extracted traffic signs is determined by iteratively finding the visible points from driver locations along the scanning trajectory. Traffic signs are linked to upstream trajectory points from the traffic sign based on local safe sight distance requirements, which amounts to a smaller distance in the urban centres analyzed in this study. The occlusion of traffic signs was measured with a precision and recall of 73.81% and 91.06% on average, respectively.

Wu et al. [44] extended their previous work on traffic sign detection from LiDAR data to evaluate the visibility of these traffic signs. The proposed framework creates feature vectors for each traffic sign whereby details are determined from the LiDAR data and from video-log imagery. The LiDAR data is utilized to determine horizontal and vertical orientation, flatness, and local distances of the traffic sign. The images are utilized to make pixel-wise and pixel-region measures of the traffic sign, determining sign panel area, and differences in the sign panel's edge strength, mean color, and color histograms as compared to the background of the image. These feature vectors are used to calculate the visibility of the traffic signs and compare them against subjective measures from the images. Signs were categorized as being invisible or having low, medium, or high visibility. The comparison between the calculated and subjective visibility measures were within 5% of one-another.

There were also some examples where legibility distance was utilized but not studied for the specific purpose of measurement. Al-Kaisy et al. [45] analyzed different factors that result in occlusions by heavy vehicles within the legibility distance of a standard passenger car (PC). A parametric analysis was conducted considering characteristics of the traffic sign and the driving environment to determine the probability of a traffic sign being occluded by a heavy vehicle and the probability of a PC missing the traffic sign under variable traffic conditions. The authors considered legibility distance between a travelling vehicle and a traffic sign within their simulated environment.

Discetti and Lamberti [46] assessed stopping sight distance (SSD) along select horizontal curves, and when SSD could not be improved through changes to roadside furniture, a traffic sign was placed with an adjusted speed to ensure adequate SSD was available for those curves. For the added speed signs, traffic sign sight distance was measured with an optical laser and verified with a pilot vehicle.

Unlike the previously mentioned publications, Nassar and Al-Kaisy [47] considered legibility distances to measure occlusions for guidance signs within buildings. The authors determined the percentage of the sightline which was occluded and the probability of missing the sign; similar to the procedure by Al-Kaisy et al. [45]. The line of sight to each sign was measured within a simulated environment, where the position of building occupants was changed with each time step and the line of sight recalculated to locate line of sight occlusions.

2.5 SUMMARY

Based on the reviewed literature, the intensity-based manipulation of the point cloud to extract candidate traffic signs is commonly conducted through both heuristic thresholds and statistical measures. Candidate traffic sign points are then combined based on their local neighborhood, commonly through distance and density-based clustering to create meaningful groups of points. Finally, candidate traffic sign clusters are further processed through the creation of cluster descriptors, typically through principal component analysis. This defines the local orientation of a candidate traffic sign cluster and is useful in the false-positive filtering process. Sign placement is calculated through a comparison against the local lane markings. For their extraction, common to the bulk of the reviewed literature is the extraction of the road surface, the manipulation of the intensity attribute of LiDAR data, and the linearity of lane markings. Previous literature focuses on one of two extraction techniques – point-based or raster and image based. Both approaches are equally powerful, however the previous research surrounding image processing provides a proven background of intensity classification research for lane marking extraction. For this work, this thesis contributes the creation of a complete TSI, determining the position of signs along the segment and measuring their placement and orientation. Both attributes are used in discussions about sign conspicuity in Chapter 4.

The fusion with image-based processes is commonly used to provide additional detail to the candidate traffic sign clusters. The power of image processing makes the classification of traffic signs from images a realistic endeavour and a key component to the completion of a traffic sign inventory. For the purpose of traffic sign image classification, neural network structures are the most common and the most accurate methods for traffic sign image recognition. However, the literature highlights varied approaches within each of the proposed neural network structures. The literature examples provide robust image classifications with marginal differences in accuracy. This thesis will discuss the application of an industry recognized neural network structure for image-based traffic sign classification. For this work, this thesis develops the first-of-its-kind training dataset for Alberta, Canada to conduct CNN-based image classification.

Finally, in contrast to the literature, this study develops an algorithm to automatically quantify real-world visibility distances from LiDAR data to determine the maximum distance at which traffic signs are visible. This is compared against required legibility distances and used to discuss the effects of consecutive traffic sign placement on sign visibility. This thesis contributes the visibility measurement of traffic signs – a new addition to the literature which considers the maximum distance at which traffic signs along a segment are visible.

3 METHODOLOGY

Given the sheer size of the LiDAR data collected for highway network scanning, the extraction of a single asset (traffic sign) becomes a big data management problem. The proposed procedure is a process of feature engineering, where meaning is created from the LiDAR dataset through data discretization and the creation of geometric and intensity-based descriptors. Before describing the proposed methodology, the following section will describe the data used within this thesis.

3.1 DATA DESCRIPTION

To assess the application of LiDAR data for traffic sign asset management, highway segments were scanned with a state-of-the-art LiDAR scanner. The LiDAR data was collected for Alberta Transportation by a third-party subcontractor using RIEGL's VMX 450 Laser Scanning System. The RIEGL VMX-450 system uses two VQ-450 scanners with IMU/GNSS units (Inertial Measurement Unit/Global Navigation Satellite System) to collect LiDAR data. The laser scanners are symmetrically configured on the left and right sides, pointing toward the rear of the vehicle at a heading angle of approximately 145°. The VQ-450 scanner has a scan rate of up to 1.1 million points per second and a scan speed of 400 lines per second [48]. The density of the points on a scanned object depends on the range, and the speed of the data collection truck. Provincial surveys conducted at 95 km/h result in LiDAR point densities on the pavement surface of 150-1000 points/m^2. It is worth noting that the MLS system can be mounted on any vehicle to conduct the surveys.

Figure 3.1 – RIEGL VMX-450 Dual Scanner with 360° Field of View Cameras [49]

The operation of a LiDAR scanner such as the one in Figure 3.1 results in a collection of individual points, which when assembled, collectively represent a point cloud. A point cloud is an accumulation of the points collected when a light pulse from a scanner reflects off an object and returns to the scanner. The number of points in any given point cloud is directly proportional to the scan rate of the LiDAR scanner used and the speed at which the scanning vehicle travels. The RIEGL VMX-450 [48] consists of two scanning heads and can collect 1.1 million points per second at a scan rate of up to 1.1 MHz, resulting in the dense point cloud in Figure 3.2.

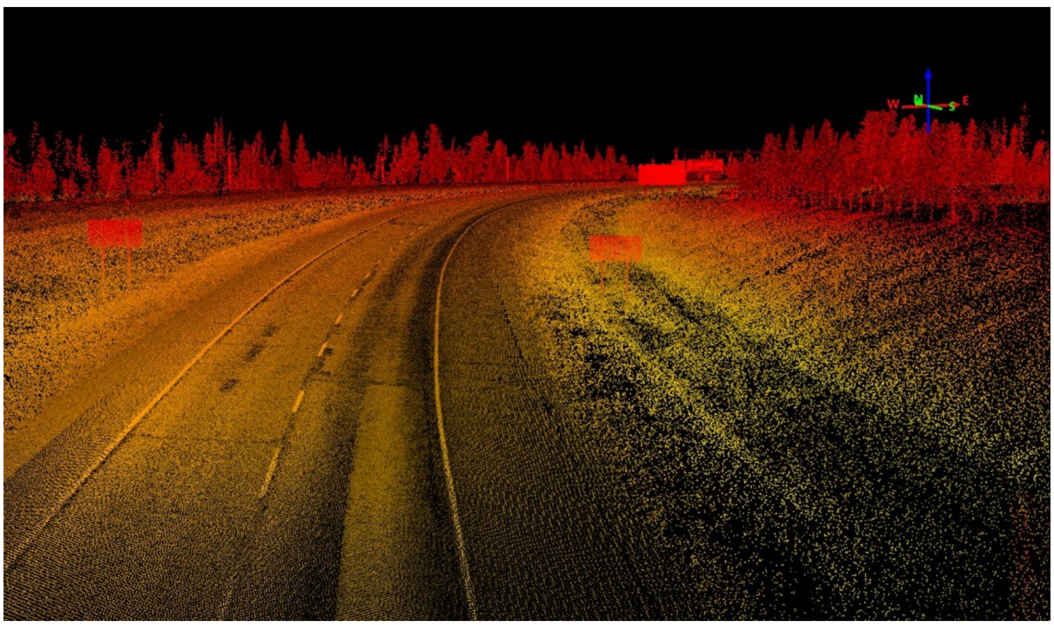


Figure 3.2 – Sample of LiDAR Data: Highway Speed RIEGL VMX-450

Part of the attractiveness of the application of LiDAR is the relative accuracy capable by larger LiDAR scanners – of which the REIGL VMX-450 is an example. The RIEGL dual scanner reports an 8mm absolute accuracy [48]. The NCHRP Report 748 [22] suggests different point density and accuracy measurements for different extraction and measurement procedures from LiDAR data. Engineering surveys and post-construction quality control requiring $> 100\ \mathrm{pts/m^2}$ resolution and $< 0.05\mathrm{m}$ accuracy. Along the test segments within this thesis the VMX-450 has a point density of roughly $400\ \mathrm{pts/m^2}$ along the road segment. Therefore, to retain the quality of the TSI conducted in this study, high density and high accuracy mobile scanners offer an accurate inventory.

The resulting dense point clouds consist of the temporal, geographic, and additional attributes as per the American Society for Photogrammetry and Remote Sensing's LAS Specification Version 1.2 [50]. Table 3.1 contains information regarding the Point Data Format of information used within this thesis. Additional attributes are available in the LAS 1.2 format but are not used because they are incomplete or unavailable. It should be noted that LAS 1.4 format is the latest point cloud format currently available, but LAS 1.2 was the latest version available during scanning.

Table 3.1 – LiDAR Point Data Record Format [50]

Data Type	**Format**	**Size**
X	Long	4 bytes
Y	Long	4 bytes
Z	Long	4 bytes
Intensity	Unsigned short	2 bytes
Scan Angle Rank	Char	1 byte
Point Source ID	Unsigned short	2 bytes

Alongside the collection of LiDAR data used for this research Alberta Transportation collected a series of roadside video-log images (henceforth referred to as "video-log") that show the driving perspective of a vehicle on its left and right-hand sides. The camera array collects a video-log simultaneously with the LiDAR scanning, and the video-log is geo-referenced using the camera's latitude and longitude coordinates. Samples of the video-log, shown in Figure 3.3, display the perspective of a specific side of the driving vehicle and includes collection date, kilometre marker, location, and latitude and longitude coordinates embedded at the top of the image. The video logs supplement the LiDAR analysis to provide further information about the sign itself (i.e., size, color, lettering, etc.) and visual proof of the detection rate (i.e., false positives, false negatives, etc.).

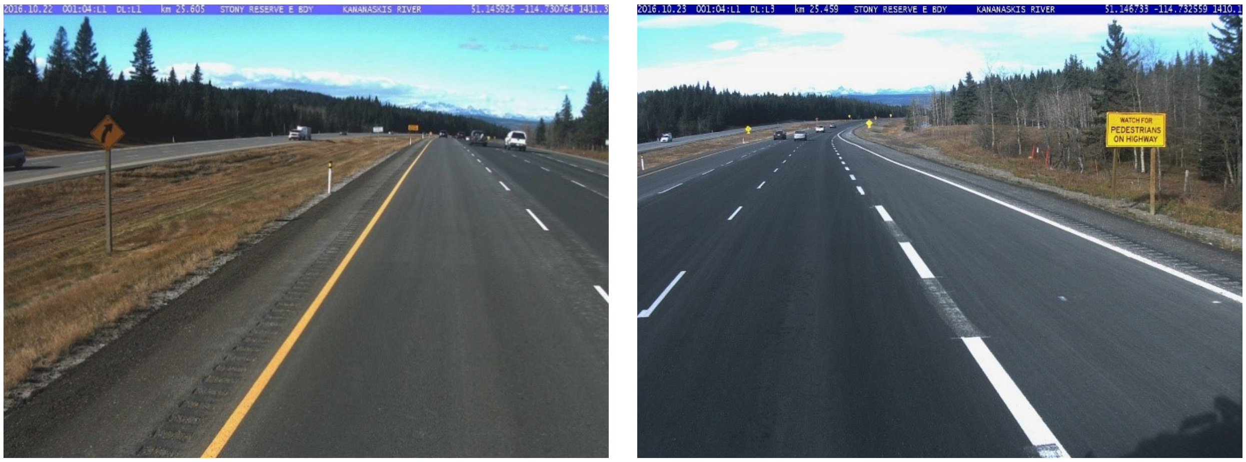


Figure 3.3 – Sample Video-log of Highway 1, Alberta, Canada

This research will consider a portion of Highway 1 extending a length of 4-km in the Eastbound direction which lies in the South-Western part of the Province of Alberta. The segment is part of a multi-lane divided highway located southwest of the city of Calgary. The speed limit on the segment is 110 km/h. Highway 1 is a critical highway within Alberta's economy, serving inter- and intra-provincial commercial, tourist, and leisure traffic travelling into the Rocky Mountains year-round. The segment is bounded by rivers, lakes, rock faces, and thick tree-lines past the clear zone and can greatly vary in horizontal and vertical alignment.

3.2 LiDAR Data Viewing & Processing

All algorithms outlined in this thesis were scripted using MathWorks' MATLAB version 2018a [51]. The volume of data present in LiDAR scanning requires the big data processing capabilities present in MATLAB's programming environment. For visualization, LiDAR data is passed into Applied Imagery's Quick Terrain Reader [52]. Quick Terrain Reader is also compatible with Google Earth's .KML files, which were used to visually debug extraction results.

3.3 Traffic Sign Extraction

3.3.1 Intensity-based Extraction

The fundamental difference between traffic signs and their surroundings is that they have retro-reflective properties. In the point cloud context, this results in points with high intensity values on the traffic signs. As suggested by Soilán et al. [21] and Riveiro et al. [25], a Gaussian Mixture Model with Expectation Maximization (GMMEM) was used to determine the probability that any point's intensity matches the intensity distributions of the entire point cloud. The value of fitting a GMM is that the algorithm remains open to variation, ensuring that high intensity points are extracted regardless of the test segment in question. The intensity histogram of the entire point cloud for the test segment is shown in Figure 3.4. Figure 3.4 (a) shows the entire intensity histogram with a peak around 16000, and Figure 3.4 (b) shows only the high intensity region (intensity > 35000) in the LiDAR segment which contains the traffic signs.

(a) (b)

Figure 3.4 – Intensity Histogram of Test Segment

Fitting a two-component GMM through expectation maximization to the intensity histogram requires two steps: *the expectation step* and *the maximization step*. The following description of the calculations behind the GMM process is based on Hastie and Tibshirani [53]. Starting with the intensity histogram, the two Gaussian distributions are initialized with guesses for the means ($\hat{\mu_i}$), standard deviations ($\hat{\sigma_i}^2$), and weighting parameters ($\hat{\pi}$) for the two distributions. For k number of Gaussian distributions, the probability of a value (x) belonging to one distribution is given by the linear combination of the Gaussians – as given in Equation 3.

$$p(x|\mu, \sigma) = \sum_{i \in [0,k)} \pi_i \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-(x-\mu_i)^2/2\sigma_i^2} \qquad 3$$

In the expectation step, the initial guesses of the mean and standard deviation for each Gaussian are used to determine the probabilities. This determines which Gaussian curves are "responsible" for the values of all points using Equation 4. For example, if a point has an 80% probability of belonging to Gaussian curve 1 and a 20% probability of belonging to Gaussian curve 2, then the first and second Gaussian curves are 80% and 20% responsible for the value of the point, respectively.

$$r_{ic} = \frac{\pi_c \, N(x_i, \mu_c, \sigma_c)}{\sum_{j \in [0,k)} \pi_j \, N\left(x, \mu_j, \sigma_j\right)} \tag{4}$$

The maximization step then takes the responsibilities from the expectation step to recalculate the means, standard deviations, and weighting values of the Gaussian curves. Using Equations 5 and 6, the new mean and standard deviation redefine the Gaussian curves, and this process is repeated until convergence. Equation 7 recalculates the weighting parameters for the Gaussian curves.

$$\mu_c^{new} = \frac{\sum_i r_{ic} x_i}{\sum_i r_{ic}} \tag{5}$$

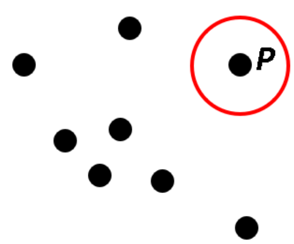
$$\mu_c^{new} = \frac{1}{\sum_i r_{ic}} \sum_i r_{ic} (x_i - \mu_c^{new})^2 \tag{6}$$

$$\pi_c = \frac{\sum_i r_{ic}}{n} \tag{7}$$

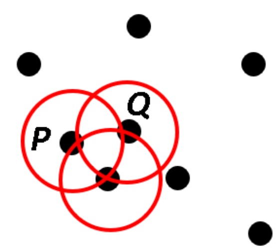
With the GMM fit to the intensity histogram, the potential traffic sign points are extracted based on the Gaussian curve with the higher mean. Points that have an intensity greater than one standard deviation below the mean are considered candidate traffic sign points. One standard deviation is chosen to minimize high intensity noise present in the dataset, such as passing vehicles and highway markers, while maintaining candidate traffic sign points.

3.3.2 TRAFFIC SIGN CLUSTERING

To determine geometric features about the extracted set of candidate traffic sign points and to conduct false-positive filtering, the extracted points are clustered using DBSCAN [54]. DBSCAN is utilized by [20], [21], [24], [25], [27], [28] and is conducted to place individual candidate traffic sign points into grouped regions of interest, removing high intensity noise (e.g. street lights, license plates, passing vehicles, etc.) kept in the GMM process, and providing a new candidate point format for further analysis. DBSCAN is an unsupervised distance-based clustering technique designed to group data irrespective of shape using minimal information about the input dataset. The minimum number of points per cluster (*MinPts*) and the maximum intra-point spacing (*Eps*) must be predefined, and DBSCAN segments the data into density-reachable and density-connected clusters and noise [54]. Cluster and noise classifications are shown in Figure 3.5, where clusters contain points that are density-reachable and density-connected with respect to the original parameters *MinPts* and *Eps*. If a point is not density-reachable from any cluster, the point is labelled as noise. An example is visually detailed in Figure 3.5, where *Eps* is shown by the red circles and *MinPts* is defined as 3 points. In Figure 3.5 (a), the no other point is density-reachable from Point *P*. Therefore, Point *P* would be classified as noise during this step. In Figure 3.5 (b), Point *Q* is density reachable from Point *P*, connecting the points that are also density-reachable. The number of points that are density-reachable from *P* is also greater than the *MinPts*, thereby assigning these points to a cluster in this step. For further detail on this method, please consult Ester et al. [54].



(a)



(b)

Figure 3.5 – DBSCAN Cluster Definition

These parameters are assigned based on the desired attributes of the point clusters. DBSCAN was used for traffic sign extraction in a few literature examples, as outlined in Table 3.2. The difference in DBSCAN attributes are related to the environment in which scanning occurs and the scanner used. As the compared studies all used comparable scanners, capable of up to 1.1 million points per second, the difference in clustering attributes are likely due to the speed at which scans were conducted. Highway LiDAR scanning is conducted near 100 km/h [24], whereas urban scans can be conducted at speeds as low as 30 km/h. Therefore, the same object scanned in the highway environment will have a lower point density when compared to the same object scanned in an urban environment.

Table 3.2 – DBSCAN Traffic Sign Extraction Attributes

Authors	Area Type	Minimum Points (#)	Epsilon (metres)
Gargoum et al. [20]	Highway	17	1
Riveiro et al. [25]	Urban	50	0.2
Soilán et al. [21]	Highway and Urban	25	0.2

Although the literature provides different recommendations for DBSCAN clustering attributes for traffic sign extraction, the extractions conducted by Gargoum et al. [20] utilize the same LiDAR scanned datasets for the Province of Alberta. The DBSCAN clustering attributes proposed in [20] will be used for this study. With the completion of traffic sign point clustering, the list of possible traffic signs can be further filtered to reduce false-positive extractions. The traffic sign clusters exhibit little unique properties by themselves, requiring additional processing to create additional traffic sign descriptors. Previously used in point cloud processing by [19], [21], [25], [27], [31], [35], [55], PCA is a common data exploration method that determines the structure of an input dataset. The following description of the calculations behind PCA are based on Smith [56]. For a single set of three-dimensional geometric data (as in a single candidate traffic sign cluster), $p = \{x, y, z\}_m$ for *m* points, the data is adjusted by removing the mean.

$$p_m{}^i = x^i - \frac{1}{m}\sum^{m} p^i \qquad 8$$

Then, the problem of determining the principal components can be formalized by determining the 3D vector, which when the data is projected on that vector has maximum variance. The covariance of the mean-removed points is determined across all *n* dimensions, where:

$$\mathrm{cov}(x, y) = \frac{\sum_{i=1}^{n}\left(x_m^i - \bar{x}_m\right)\left(y_m^i - \bar{y}_m\right)}{n - 1} \qquad 9$$

$$C = \begin{pmatrix} \mathrm{cov}(x,x) & \mathrm{cov}(x,y) & \mathrm{cov}(x,z) \\ \mathrm{cov}(y,x) & \mathrm{cov}(y,y) & \mathrm{cov}(y,z) \\ \mathrm{cov}(z,x) & \mathrm{cov}(z,y) & \mathrm{cov}(z,z) \end{pmatrix} \qquad 10$$

Finally, PCA requires an eigenvalue decomposition of the covariance matrix. The result is three principal components (λ_i) for the three input dimensions described by their magnitude (eigenvalues) and direction (eigenvectors) in Figure 3.6. For traffic signs, the principal direction is governed by the largest dimension of the traffic sign. A horizontal guide sign's principal component exists along it's horizontal axis, whereas a maximum speed sign's principal component exists along it's vertical axis. For traffic signs that are symmetrical, the first and second principal components should be roughly equal, but irrespective of a traffic sign's size, the third principal component will always be orthogonal to the traffic sign panel.

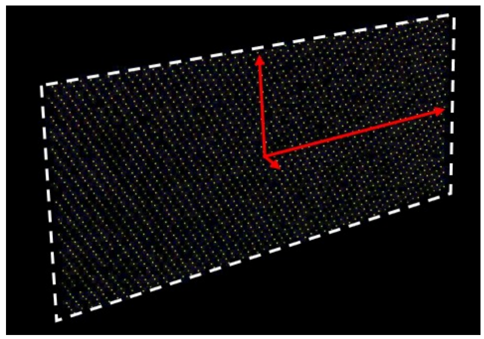


Figure 3.6 – Three- Dimensional Principal Component Analysis

For the three-dimensional data within a point cloud this results in three directional vectors with which to determine the orientation of any given point cluster. Geometrically, the clusters defining traffic signs can be simplified to clusters that exhibit flatness. The flatness of a cluster of points is defined in Equation 11 [57]. For a cluster of point to be considered flat, as is required for a traffic sign, the flatness of a cluster must be less than 1/3 [57].

$$\mathrm{F} = \frac{\sqrt{(\lambda_2)} - \sqrt{(\lambda_3)}}{\sqrt{(\lambda_1)}} \tag{11}$$

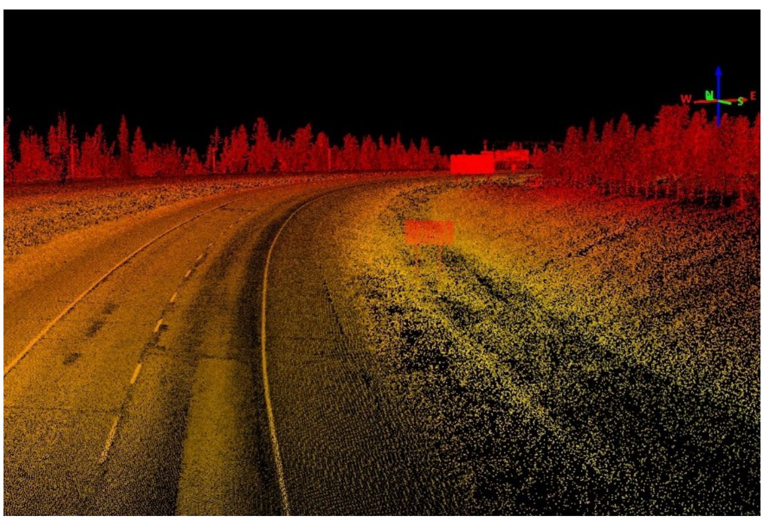
Additionally, the local angle of the traffic sign cluster was calculated to determine the traffic sign orientation relative to the travel lane. The perpendicularity of the traffic sign relative to the nearest travel lane was determined by comparing the 3rd principal component of a traffic sign cluster to the trajectory of the scanning vehicle. The result is the lateral and vertical angle of the traffic sign relative to the scanner travel direction and the horizontal normal, respectively. For a traffic sign T_i with principal components λ_1, λ_2, and λ_3, the nearest trajectory vector $T_v = [x_v, y_v, z_v]$, and the normal vector to the surface at the nearest trajectory vector, $N_v = [x_{N_v}, y_{N_v}, z_{N_v}]$, the horizontal and vertical angles of the traffic sign, α and γ, respectively, are given by Equations 12 and 13.

$$\alpha = \tan^{-1}(T_v, \lambda_3) \qquad 12$$

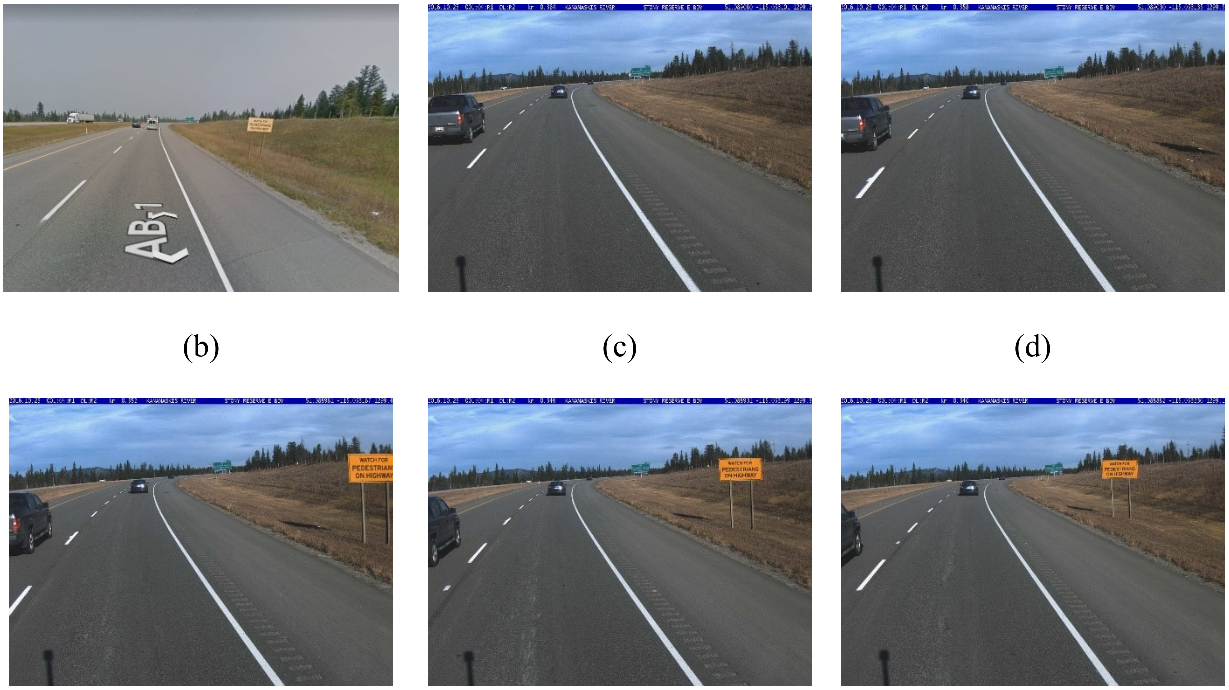
$$\gamma = \tan^{-1}(N_v, \lambda_3) - 90 \qquad 13$$

3.3.3 *COMPARISON TO VIDEO-LOG IMAGES*

Video-log images are collected simultaneously alongside LiDAR scanning. This includes a log of the images, their collection location, angle, and filename. This can be used to locate the images relative to points extracted from the LiDAR data. For the extracted traffic signs, the nearest trajectory point is located for each potential traffic sign. This trajectory point is compared against the video-log of the collected images, using the heading angle of the camera to approximate which image corresponds to each potential traffic sign cluster. The image frames shown in Figure 3.7 (c) - (g) outline an additional step taken to ensure the images contain the traffic sign of interest. This is done to provide the image classification procedure, described later, with additional samples of the traffic sign to verify the classification. Figure 3.7 is a sample extraction, with the LiDAR position and traffic sign shown in Figure 3.7 (a) and (b) illustrate a LiDAR traffic sign and an image of the sign in Google Earth [58], respectively.



(a)



(b) (c) (d)

(e) (f) (g)

Figure 3.7 – LiDAR to Image Extraction Sample

Once the set of nearest images have been collected, the classification of the traffic sign within the image comes from the maximum classification accuracy of a sliding window passed over the input images. The classification that is most consistent across multiple images is assigned as the classification of the current traffic sign.

3.4 LANE MARKING EXTRACTION

The literature highlights images of the point cloud as a popular basis for lane marking extraction – leveraging the recent influx of image processing research to aid the intensity-based lane marking extraction. Therefore, this research explores a lane marking extraction method considering the advantages of using a two-dimensional representation of a LiDAR-scanned roadway. To this end, the LiDAR point cloud was voxelized to create expanded regions of interest along the roadway used for the extraction of cross sections. Voxelization splits scanned point cloud data into cubes of equal size, resulting in an alternative representation of the point cloud with a fraction of the point density. All points $P(x, y, z)$ in a point cloud are assigned to a voxel $v(i\ , j, k)$ depending on the dimensions of the point cloud and the desired size of the voxels. If the length, width, and height of a voxel are defined as d_x, d_y, and d_z, and the origin of the point cloud is denoted as x_o, y_o, and z_o, the voxel coordinates of a point are determined as follows:

$$\mathrm{i} = \mathrm{int}\left(\frac{\mathrm{x} - \mathrm{x_o}}{\mathrm{d_x}}\right) \quad 14$$

$$\mathrm{j} = \mathrm{int}\left(\frac{\mathrm{y} - \mathrm{y_o}}{\mathrm{d_y}}\right) \quad 15$$

$$\mathrm{k} = \mathrm{int}\left(\frac{\mathrm{z} - \mathrm{z_o}}{\mathrm{z_x}}\right) \quad 16$$

Next, to refine the point cloud into constituent "object" and "non-object" point classes, the point cloud is subject to a ground and non-ground separation. This process is built upon the assumption that, for a section of given LiDAR scan, the ground will always be the lowest point in that section and any other point is a non-ground object. Algorithmically, given a column of voxels V_c in location $\{I, J, k\}$ where $k > 0$ and $\{I, J, k\} \in V_c$, the ground voxel g_v is defined by Equation 17. The non-ground voxels are then defined as $\{I, J, k\} \notin g_v$.

$$\mathrm{g_v} = \min\{\mathrm{I}, \mathrm{J}, \mathrm{k}\} \quad 17$$

Road surface extraction is common in lane marking extraction, with [31] conducting raster-based extraction and [30], [35] conducting curb or elevation-based extraction. This work will consider only rural highway segments, leaving curb extraction unavailable for road surface extraction. To conduct elevation-based road surface extraction, the voxels calculated to discretize the point cloud and create the raster cross sections can be used to extract the road surface. For the extraction of a road surface, the elevation of the points within each voxel provides a description of the local change in elevation. For the purpose of road surface extraction, the standard deviation of elevation is calculated for each voxel as in Equation 18, where σ is the standard deviation, z_i is the elevation of point i, n is the number of points in voxel v, and $\overline{z^v}$ is the mean of elevation. For a set of $\{i, j\}$ voxel indices within a LiDAR scan, the voxels are segmented to extract the road surface using the heuristic filter in Equation 19.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} \mathrm{z_i^v} - \overline{\mathrm{z^v}}}{\mathrm{n}}} \quad 18$$

$$V_{std}(i,j) = \begin{cases} \text{other}, & V_{std}(i,j) > 0.01 \\ \text{candidate road surface}, & V_{std}(i,j) \leq 0.01 \end{cases} \quad 19$$

The largest cluster of candidate road surface voxels is taken to be the road surface. To extract roadway cross sections, a parametric representation is used based on the voxel coordinates and the trajectory location of the scanning vehicle. For a trajectory point $V_T = i_T, j_T, k_T,$ cross section width (w) in metres, cross section depth (d) in metres and the vectors parallel and perpendicular to the scanning vehicle's direction of travel (v_{Par}, v_{Per}), cross section number x is defined as:

$$V_1(x) = V_T(x) + v_{per}(x) * w/\text{voxel size} \quad 20$$

$$V_2(x) = V_T(x) - v_{per}(x) * w/\text{voxel size} \quad 21$$

$$P_w = V_1 + t_w(V_1 - V_2) \in t_w = 0\text{: step increment: } 1 \quad 22$$

$$XS(x) = P_w + t_d * (d/\text{voxel size}) \quad 23$$

This results in detailed rasters of the roadway cross section, where the pixel value is assigned as the average intensity within each voxel. The cross sections are split into 5-cm pixels, where the pixel size was chosen to retain detail about the local intensity changes within the cross sections while still discretizing the points in the cross section into a manageable format. The roadway is now described by two-dimensional raster images, allowing for the application of image processing techniques to begin extracting the lane markings. Lane markings differ from their surroundings in that they exhibit greater intensity and follow linear patterns along the road segment. Figure 3.8 shows an example of the rasterization.

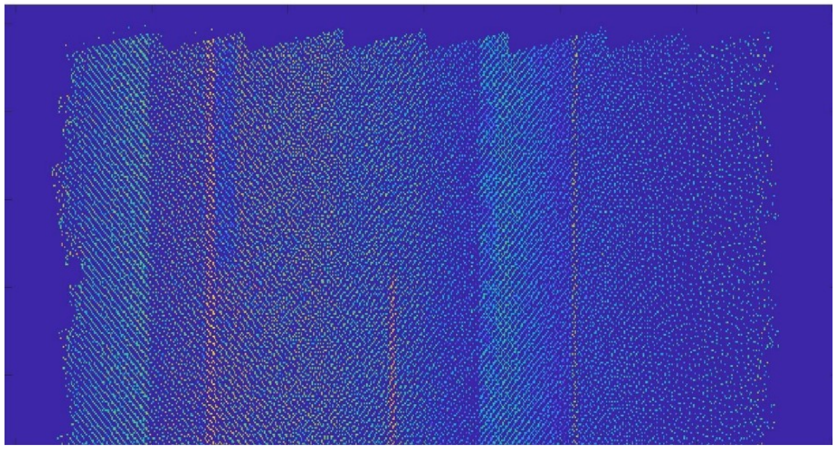


Figure 3.8 – Raster Image of Cross Section

The raster image is then completed by applying the morphological closing operation to fill gaps between raster pixels. The closing operation applies a structural object to erode the dilation of the input raster [59]. Structural objects can be applied as square, rectangular, circular, or of arbitrary sizing. In the erosion operation, the structural element is centred on each background pixel and the image's pixel values within that structural element are collected. If at least one foreground pixel exists within the structural element, the current background pixel is set to the value of the foreground pixel. Similarly, in the dilation operation, the structural element is centred on each foreground pixel and the image's pixel values within that structural element are collected. If any pixel within the structural element is not a foreground pixel, then the current foreground pixel is set as a background pixel. The results of the complete erosion operation are outlined in Figure 3.9.

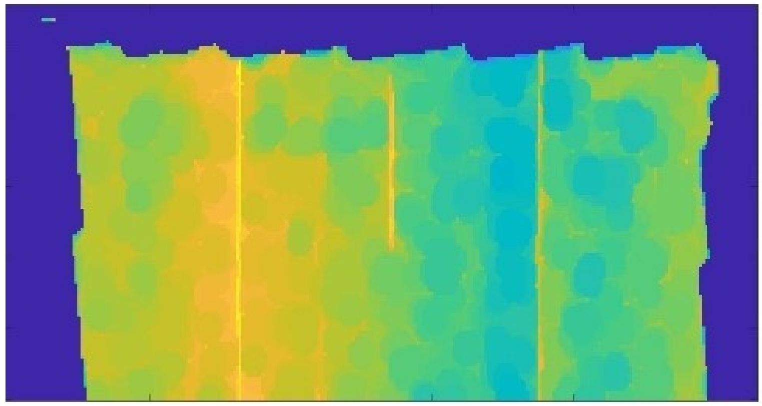


Figure 3.9 – Erosion Operation Applied to Raster Image

With the road surface split into cross sections, the focus is now on further reducing the point cloud to retain only the lane marking points. The primary differentiation between lane marking points and the pavement is its intensity value; also considered by every publication reviewed previously. However, as seen in Figure 3.9, the intensity of the lane markings along a cross section are not necessarily uniform due to uneven wear and tear and their location from the LiDAR scanner. Therefore, contrast enhancement will be used to make the lane marking points more discernible. This is typically conducted through gamma adjustment of the original image's cumulative density function (CDF). The procedure for contrast enhancement will be based on the work by Huang et al. [60]. For the 65,536 unique intensity values available in 16-bit color space, the AGCWD process starts by defining the probability density function (PDF) of the input image. For images, the PDF can be approximated using a histogram, $h[i]$, of the image, normalized such that the area under the histogram is equal to 1. Therefore, for an image with *N* pixels the PDF is:

$$\mathrm{P}[\mathrm{i}] = \frac{1}{\mathrm{N}}\,\mathrm{h}[\mathrm{i}] \qquad 24$$

The PDF of the input image is then modified using characteristics of the PDF and a defined weighting parameter. This is conducted to account for common fluctuations within the CDF due to environmental variations [60]. The weighted PDF is given in Equation 25, where $\mathrm{PDF_{min}}$ and $\mathrm{PDF_{max}}$ correspond to the minimum and maximum values of the PDF for the input image, and *w* is a weighting parameter.

$$\mathrm{PDF_w} = \mathrm{PDF_{max}} \left(\frac{\mathrm{P[i]} - \mathrm{PDF_{min}}}{\mathrm{PDF_{max}} - \mathrm{PDF_{min}}} \right)^{\mathrm{w}} \quad 25$$

The gamma correction process is then applied based on the CDF of the input image, where the gamma correction parameter and final image gamma correction are defined by Equations 26 - 28, where $l \in [l_{min}, l_{max}]$ and *T(l)* is the gamma adjusted intensity values.

$$\mathrm{CDF_w} = \sum \frac{\mathrm{PDF_w}}{\sum \mathrm{PDF_w}} \quad 26$$

$$\gamma = 1 - \mathrm{CDF\_w} \quad 27$$

$$\mathrm{T(l)} = 65535 \left(\frac{\mathrm{l}}{65535} \right)^{\gamma} \quad 28$$

The MATLAB script written for contrast enhancement was adapted from [61] for its application to the raster images in this study. The improvement in contrast makes the lane markings present within the rasters more prominent compared to their surroundings. Finally, gaussian smoothing was used to alleviate some of the aliasing caused by the rasterization process – applied using a standard deviation of 0.5 for the Gaussian curve. However, contrast enhancement creates an interesting challenge when assessing image quality, as a comparison against the initial image as the ideal reference image is not always reasonable. The contrast enhanced image is often of better quality than the original image. To determine the quality of the contrast enhancement, the patch-based contrast image quality index (PCQI) is used to illustrate the result of the contrast enhancement and determine the optimal weighting to apply to the raster's PDF [62]. The PCQI mimics the human visual perception of contrast to determine when an image's contrast has been enhanced without the need for a perfect reference image. The greatest quality improvements come from contrast adjustments with a high mean PCQI, where changes to the input image can be discerned based on using a PCQI heatmap.

With the contrast of the lane markings enhanced, the next stage of their extraction requires a definition of the edges that define the lane marking. The following description of the Canny edge detector is based on Forsyth and Ponce [63]. The Canny edge detector utilizes local intensity gradients within an image to determine pixel edges. The edge detection process consists of five stages: (i) Gaussian filtering, (ii) intensity gradient calculation, (iii) non-maximal suppression, (iv) two-fold edge thresholding, and (v) tracking edges by hysteresis. A Gaussian kernel is applied to smooth noise within the image, where the kernel is defined in Equation 29, where σ is the standard deviation of the Gaussian curve.

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \qquad 29$$

The effect of the image smoothing is highly dependent on the size of the Gaussian kernel and the chosen standard deviation. After smoothing the image, the intensity gradient and orientation for each pixel is determined using Equations 30 and 31.

$$M = \sqrt{\left(G_x^2 + G_y^2\right)} \qquad 30$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right) \qquad 31$$

The next stage, non-maximal suppression, scans the pixels within an image and removes all edge pixels that are less than the set threshold. For each pixel, the gradient at that pixel is used to identify the location of an edge, where pixel edges are located perpendicular to the gradient of the pixels. By comparing against adjacent pixels, if the current pixel is maximum and its magnitude is greater than the defined threshold, this pixel is marked as an edge.

A two-fold threshold is applied to select candidate edges by preserving high value edges and remove weak edges caused by color variation or noise. The threshold is applied by considering the pixel intensity relative to the threshold, where:

a. if pixel value > high threshold = strong edge;
b. if pixel value < high threshold and pixel value > low threshold = weak edge; and,
c. if pixel value < low threshold, suppress value.

The high and low threshold values are determined empirically dependent on the image content, where the lower threshold is 40% of the upper threshold value. Finally, edges are tracked through the process of hysteresis. The edge detection process is finished by connecting the extracted edges, suppressing all weak edges that are not connected to strong edges.

However, due to natural changes in intensity along the roadside and road surface, edges other than the lane markings are typically also present. Therefore, the Hough transform [63] is used to search for linear objects within the series of edges. This process is meant to locate prominent linear edges caused by the lane markings.

The Hough transform creates a two-dimensional accumulator array from an input image using Equation 32. For each pixel in an image, its neighborhood is analyzed to determine if this pixel belongs to a linear edge and the values of r and θ are calculated. The accumulator values are combined into bins, displayed in Figure 3.10, where the bins with higher values correspond to lines with the best representation within an image. Multiple prominent linear edges within an image by specifying a series of accumulator peaks can now be extracted by locating the edges of multiple lane markings within a raster image.

$$\mathrm{x}\cos(\theta) + \mathrm{y}\sin(\theta) + \mathrm{r} = 0 \qquad 32$$

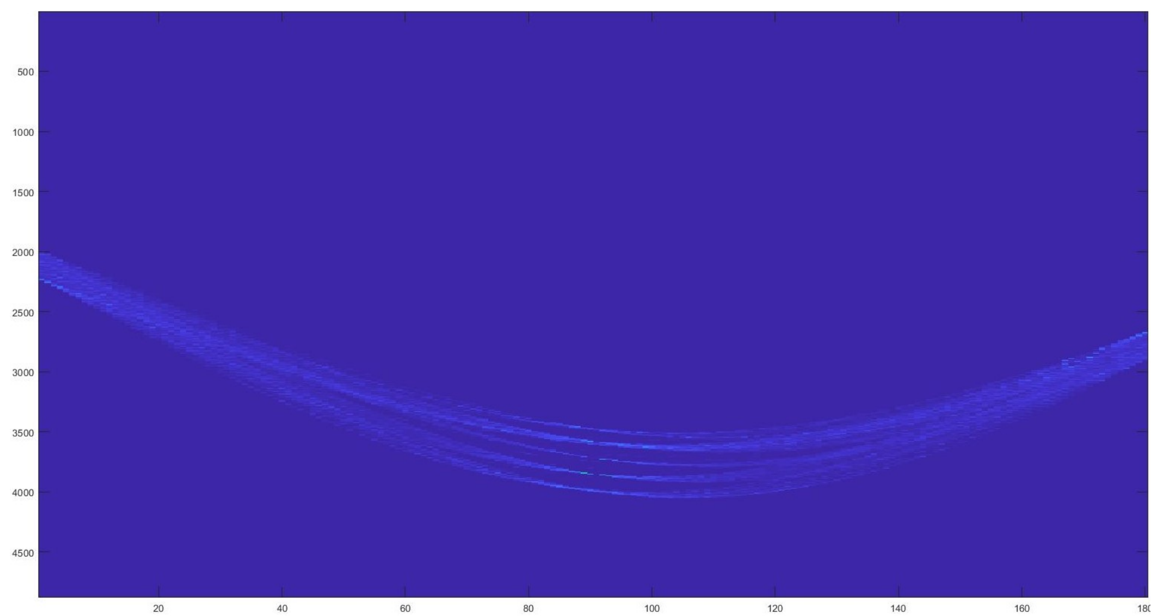


Figure 3.10 – Hough Transform Line Detection

With the addition of lane markings local to each traffic sign, the mounting height and lateral placement of the traffic sign can be calculated from the bottom and nearest edges of the traffic sign to the nearest lane marking. For a given point cloud containing a traffic sign cluster defined by the set of n points $\{x_{TS}, y_{TS}, z_{TS}\} \in P^{TS}$, a traffic sign point on the edge nearest to the lane marking $P_n^{TS} = \{x_n, y_n, z_n\}$, the nearest lane marking as defined by vector $\overrightarrow{P^{LM}} = [x_{LM}, y_{LM}, z_{LM}]$, and a lane marking point $P^{LM} = \{x_{LM}, y_{LM}, z_{LM}\}$, the lateral and vertical placement of the traffic sign, d_L and d_V, respectively, are defined in Equations 33 and 34. The $\| \ \|$, $\times$, and $\min\{ \ \}$ operators refer to the vector norm, cross product, and minimum calculations, respectively.

$$\mathrm{d_L} = \frac{\left\| (\mathrm{P_n^{TS}} - \mathrm{P^{LM}}) \times \overrightarrow{\mathrm{P^{LM}}} \right\|}{\left\| \overrightarrow{\mathrm{P^{LM}}} \right\|} \qquad 33$$

$$\mathrm{d_V} = \min\left\{ \mathrm{z_{TS}^i} - \mathrm{z_{LM}} \in \mathrm{i} = 1{:}\,\mathrm{n} \right\} \qquad 34$$

The lateral placement is thereby measured from the inside edge of the sign and the vertical placement measured from the bottom edge of the sign. Traffic sign placement standards are defined to maintain a combination of "*safety, visibility, and practicality*" [64], making both lateral and vertical placement attributes fundamental to the maintenance of a safe and efficient roadway. With the location, placement, and orientation measurements complete, professionals and municipalities can assess the correctness and effectiveness of current traffic sign placements. The final stage of a traditional TSI requires the classification of the traffic signs to be determined. The following section will discuss the fusion of LiDAR and video-logs for the purpose of traffic sign classification.

3.5 TRAFFIC SIGN CLASSIFICATION

3.5.1 NEURAL NETWORK STRUCTURE

NCHRP Report 748, while considering the application of mobile LiDAR mapping for transportation asset management, also recommends the application of image from traffic signs as they provide additional context as to the placement and surroundings of each traffic sign [22]. However, the computational determination of what objects make up an image is not a simple task. To this end, the CNN structure is commonly used in image processing due to the similarity between convolutional kernels and the human brain's natural visual cortex [65].

A neural network structure is typically described by a directed acyclic graph, where the neurons of the neural network are represented by the nodes within the graph. The simplest form of this neuron arrangement is given in Figure 3.11, where a series of inputs and weights are combined in a hidden layer before being passed to the output. A series of inputs are defined, weights are applied, and the result is summed to determine the value of this neuron. Before being output, the value is passed through an activation function to align the neuron value within an expected range (0-1) and determine whether this neuron "fires" to the output.

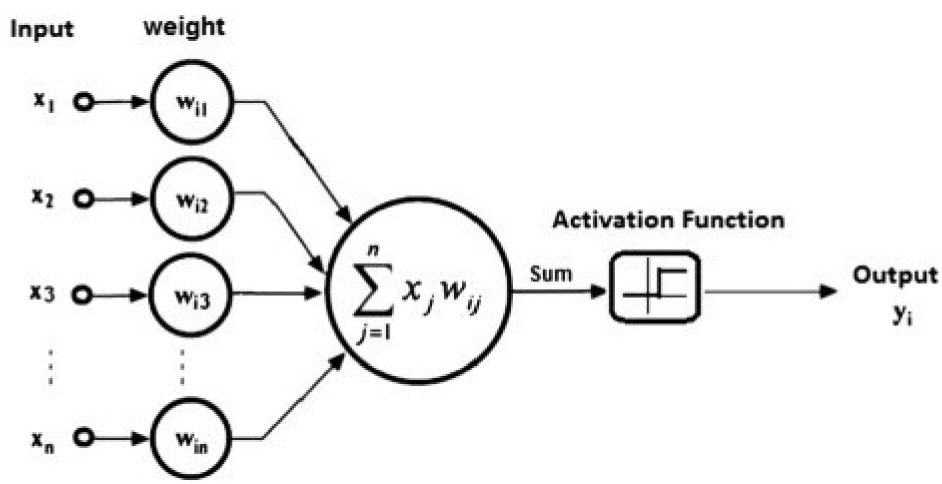
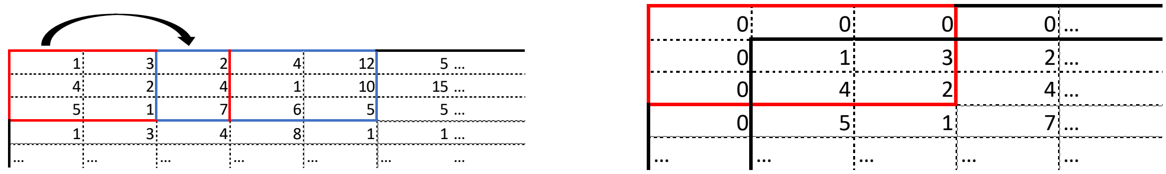


Figure 3.11 – Neuron Structure for Neural Networks [66]

The neural network structure can then be expanded with additional hidden layers or different activation functions to change the response to the input data. The CNN structure, which is based on the concept of image convolution, is commonly used in applications for image processing. A simple CNN consists of an input layer, an output layer, and multiple hidden layers

– with the hidden layers commonly consisting of convolutional, pooling, normalization, and fully connected layers [67]. Before describing those common neural network layers, the definitions of stride and padding serve to describe the details of the convolutions that occur in different CNN layers. Stride is the pixel-wise distance between adjacent convolutional kernels. For example, if a convolution layer is being applied with stride 2, the first and second convolutional kernels will be centred on pixels 1 and 3, respectively. This is illustrated in Figure 3.12 (a). Additionally, padding is used to broaden the extents of an input image to which a kernel will be applied. If a kernel is to be applied to all individual pixels in an image, the padding assigns a value to the *p* padding pixels outside of the extents of the input image to ensure that the values outside of the input image still exist. Figure 3.12 (b) shows an image with padding of 1 with a convolutional region centred on the first true pixel in the image.



(a) (b)

Figure 3.12 – Examples of (a) Stride and (b) Padding within a Neural Network

Convolution is the basis of the CNN structure and used to detect local and global features, including contrast, edges, and other geometric properties of the image. Each convolutional layer convolves a filter of size N_x x N_y across the extents of an input image, performing a dot product combination during convolution, which is carried forward through the neural network. Each convolution kernel contains a set of N^2 neurons, which contain trainable biases and weights. Figure 3.13 shows an example of convolution, with an input map of size m x n, a convolutional kernel of size 3 x 3, and a padding size of zero. The output map after convolution will be of size $m - 1$ x $n - 1$, where the value of each entry in the output map is the dot product of the input map and the kernel matrix.

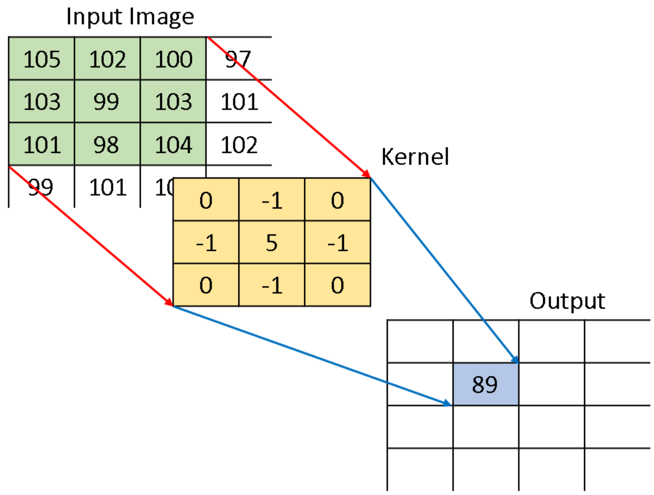


Figure 3.13 – Image Convolution

Activation functions are used to normalize the neural output of the hidden layers prior to being output, taking input values from a previous layer to determine if the neuron is activated (i.e. fired). Rectified Linear Units (ReLU) is the most common activation function, where values < 0 are set to zero, and values > 0 follow the linear function $y = cx$. More formally, this is conducted by applying the function $f(x) = max(0, x)$, where x is the neuron input from the previous layer. The application of a non-linear activation function enhances a network's decision-making capabilities without affecting the trainable parameters of previous convolutional layers [27]. An example is provided in Figure 3.14, where the application of the ReLU layer to an input feature map retains only the features which are positive.

3	-4	-8	-3	-5
4	0	-10	0	0
7	0	-9	1	0
4	0	1	3	0
0	2	0	1	0
-1	3	2	0	0

Apply ReLU →

3	0	0	0	0
4	0	0	0	0
7	0	0	1	0
4	0	1	3	0
0	2	0	1	0
0	3	2	0	0

Figure 3.14 – Rectified Linear Unit Layer

The pooling layers apply a discretization function, which passes a kernel across an input image in non-overlapping regions. This calculates the maximum or average value to retain the significant details while minimizing the complexity of the following layer and the number of trainable parameters. The application of pooling layers improves network generalization and controls overfitting by reducing input maps to its invariant features. By applying a non-overlapping pooling kernel of size P_x x P_y to an input map of size m x n, the output map is down-sampled to a size of m/P_x x n/P_y but retains the depth of the input map. Figure 3.15 shows the application of 2 x 2 maximum and average pooling layers to a 4 x 4 input map. The maximum and average values of the non-overlapping regions are assigned to the new pooling layer which is of size m/P_x x n/P_y.

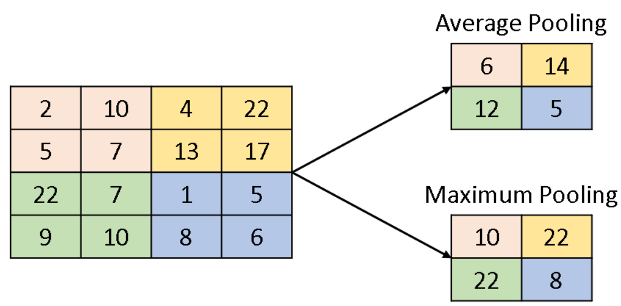


Figure 3.15 – Pooling Layers

Drop-out layers are used to prevent over-fitting in neural network training by randomly “dropping out” (i.e. turning off) neurons in a neural network, by setting the drop-out of a network to 0.5; during each iteration a random sample of 50% of the network’s neurons are turned off. This is illustrated in Figure 3.16 through the connections in a simple neural network before and after dropout is applied. Just under half of the neurons in this network are turned off during drop-out but retaining their connections further into the network.

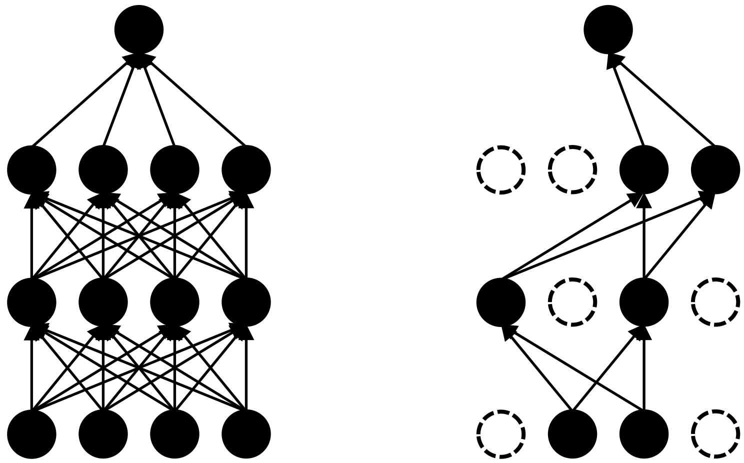


Figure 3.16 – Neural Network Dropout

Finally, the fully connected layer connects every neuron from previous layer to every neuron in another layer. The convolutional layers previously in the network discern specific features about the image, and the fully connected layer serves to aggregate the features into a one-dimensional feature vector – commonly used at the end of a network to classify the input image based on the network's activations.

To combine these layers into a functional image processing network, LeCun et al. [68] pioneered an image processing CNN structure of stacked convolutional layers followed by at least one fully-connected layer. LeCun et al.'s CNN requires 32x32 pixel images as input, which are then fed through a 7-layer network, as shown in Figure 3.17, for the purpose of classifying hand-written text.

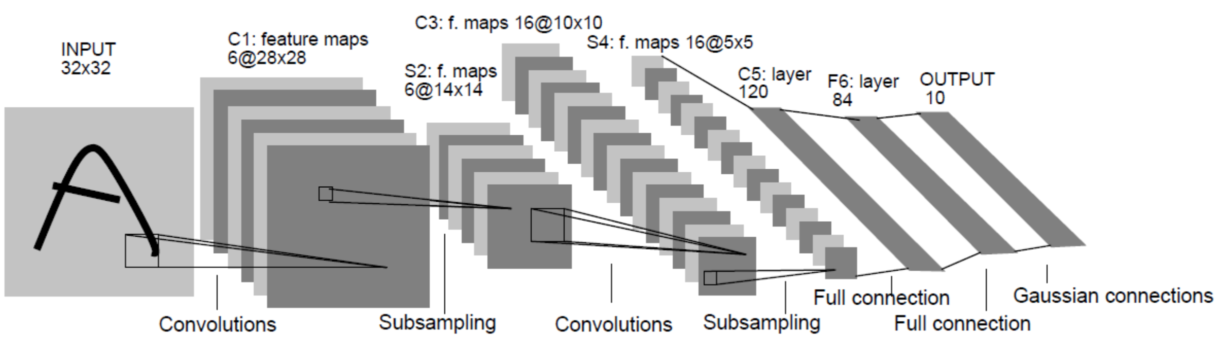


Figure 3.17 – LeNet-5 CNN Structure [68]

The layers within Figure 3.17 show the series of convolutional (C#), subsampling (S#), and fully-connected (F#) layers used to extract low, mid, and high-level features within the input image. The layers between the input and output layers are stacked such that each layer calculates features based on the result of the previous layer. The layers utilize 5x5 and 2x2 neighborhood regions to discretize the previous layer's result, occasionally applying trainable bias coefficients and activation functions to determine the effect that each neuron has on the network's overall feature extraction [68].

Reminiscent of the LeCun et al. [68] CNN structure, the GoogLeNet network (also known as Inception V1) builds upon the LeNet-5 structure of convolutions and connections with depth and varied kernel and regularization structures. GoogLeNet represents Google's winning entry into the ImageNet Large Scale Visual Recognition Competition in 2014, classifying images from the ImageNet database with a Top-5 error of 6.67% [67]. The predictive power and speed of the Inception structure comes from the use of dimension reduction, where 1x1 convolutions are applied to reduce the complexity of the input before continuing onto the more computationally expensive 3x3 or 5x5 convolutions [67]. A single Inception module is illustrated in Figure 3.18 to highlight the dimensionality reduction previously mentioned. Figure 3.19 illustrates the entire GoogLeNet structure created with stacked Inception modules.

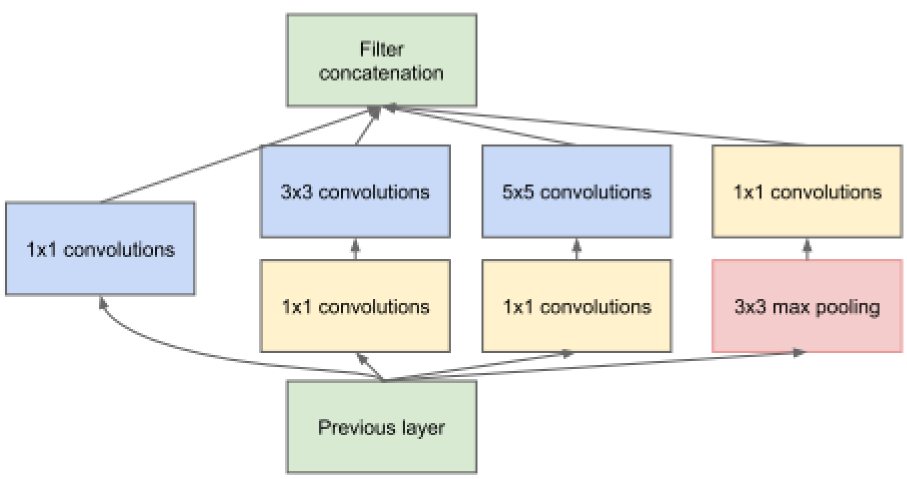


Figure 3.18 – GoogLeNet Inception Module [67]

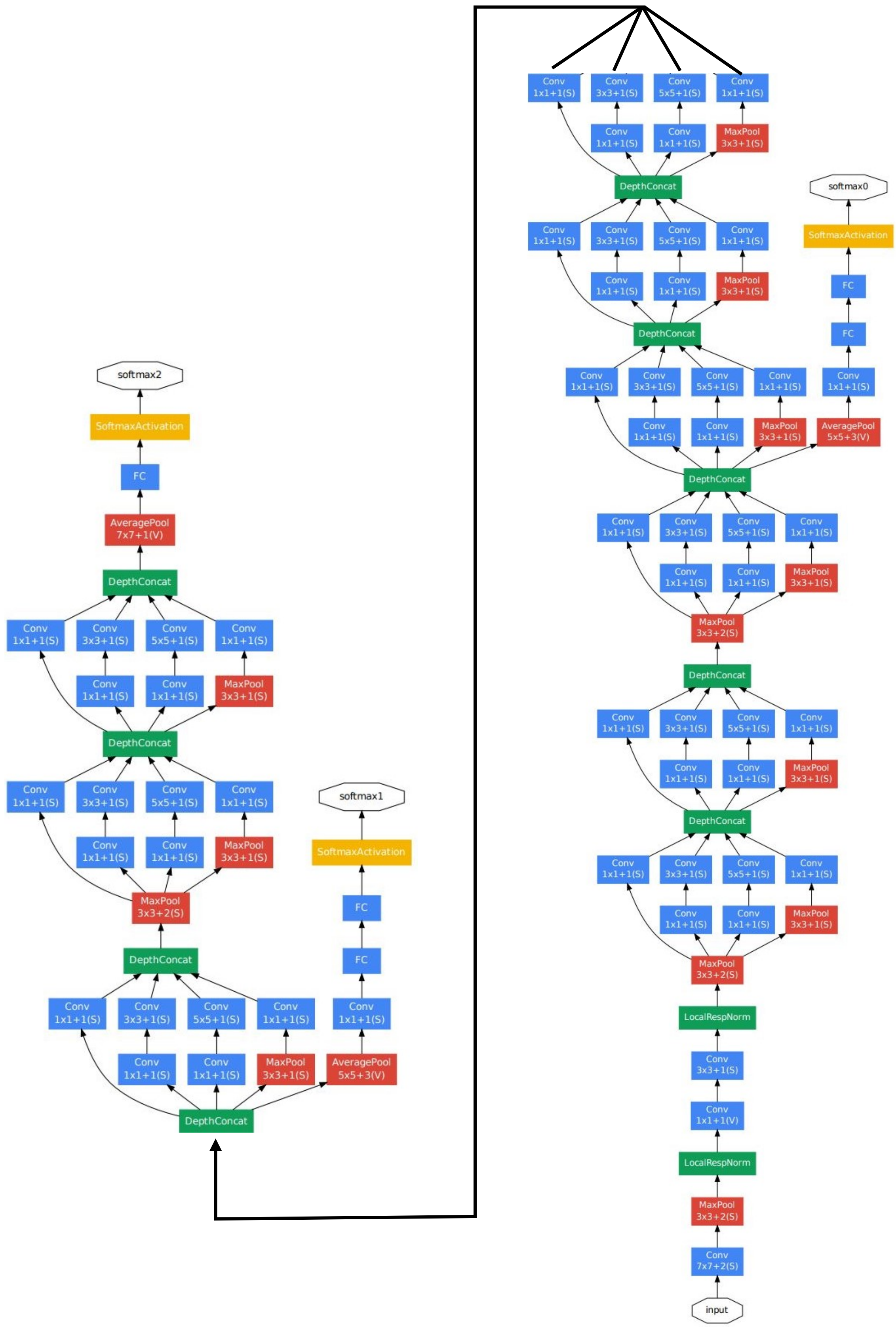


Figure 3.19 – GoogLeNet Convolutional Neural Network Model, Adapted from [67]

3.5.2 Image Classifier Training Data

Critical to the success of any neural network classifier is the availability of a thorough training dataset. Alberta Transportation publishes a traffic sign database for all of the traffic signs placed in Alberta, including their panel message, size, color, and letter type/sizing [69]. The sign panels are collected as individual images of the “perfect” traffic sign and describe the design parameters of the traffic sign panels. Although descriptive, these images are not representative of traffic signs placed on real highways. True traffic signs may be weathered and exist in variable lighting, weather, and placement conditions. Therefore, a dataset representative of the effects of real-world placement conditions on traffic signs needed to be created. To this end, video-log images collected during LiDAR scanning in other parts of Alberta were manually cropped into training images of size 224×224. To ensure diversity in condition, lighting, and location of the traffic signs cropped from video-log images, multiple segments from different divided and undivided highways in Alberta were used. The ambient lighting based on time of day and season, condition based on regional weather patterns and traffic sign age, and overall traffic environment all factor into the condition and conspicuity of the traffic signs within the images. The variability in ambient lighting, placement, and traffic sign condition are illustrated through different examples of a curve left warning sign (WA-3-L) in Figure 3.20. The random cropping of the traffic signs ensures the neural network trained for traffic sign classification is not over-trained to “perfect” representations of the traffic sign.

(a) (b) (c) (d)

Figure 3.20 – Classifier Training Image Variation

Within Alberta, traffic signs fall into one of three categories: Regulatory (RA, RB, RC), Warning (WA, WB, WC, WD), and Guide and Information (IA, IB, IC, ID, IF) signs [69]. In total, 155 classes of traffic signs along Albertan highways were collected resulting in 12,315 total images. Additionally, 1,289 images across 10 classes of false positives were. The false positives were added to the dataset to ensure the neural network accurately recognizes non-traffic sign objects rather than mistaking them for different traffic signs. Given the variable condition and placement of traffic signs along any given highway network, the collection of traffic sign samples provides an array of different placement conditions within the collected 224 x 224 sized sample images. Of the 13,604 available images, 70% were used for training and 30% were used for validation. The mean sample size across all classes was 57.74 images per class for the training dataset. The images were collected based on their real-world placement along a random selection of major and minor Albertan rural highways, thereby creating an uneven collection of samples for each of the traffic sign classes. Figure 3.21 shows the number of collected samples for each traffic sign class.

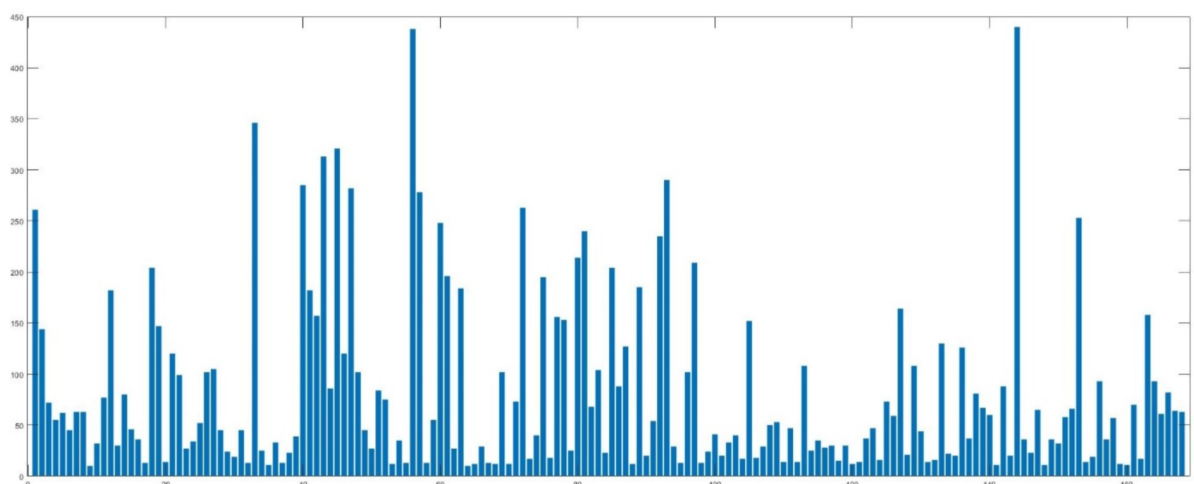


Figure 3.21 – Classifier Training Image Count

The maximum speed signs (RB-1) were split into their respective speeds (i.e. 70 km/h sign is labelled RB-1-70). Although they are technically grouped into one variable traffic sign structure (where the speed changes on the sign but the lettering and sign sizing options stay the same), the unique classification of different speed signs was deemed a useful addition for future use.

With the image classifier trained, the traditional TSI structure of traffic sign location, classification, and placement measurements are completed. However, this research serves to contribute an additional measure of traffic sign placement efficacy, namely a traffic sign's visibility.

3.6 TRAFFIC SIGN VISIBILITY ANALYSIS

3.6.1 TRAFFIC SIGN VISIBILITY MEASUREMENT

Historically, the process of traffic sign visibility measurement has been conducted manually with in-vehicle observers [41] or with optical laser measurements [46]. However, given the high detail of LiDAR scans, these scans were utilized to efficiently assess traffic sign visibility in this paper. A critical difficulty with the application of LiDAR data to a visibility assessment is LiDAR's susceptibility to noise. If the scanning was conducted in heavier traffic, the LiDAR scans often contain points detailing the vehicles the scanner passes and is passed by during scanning. To ensure the visibility assessment, which determines the maximum unobstructed distance at which a traffic sign is visible, is not impacted by the points comprising of passing vehicles, the vehicles travelling in the same direction (i.e. Eastbound) as the scanning vehicle are removed from the dataset. For the Eastbound (i.e. scanning direction) vehicles from the point cloud, the road surface is used to localize the region in which the scanning direction vehicles can exist. The point cloud is segmented to remove the voxels that exist above the road surface and applies a vertical filter of four metres, a conservative height based on the minimum bridge clearance standard in Alberta [70]. This ensures that points are only segmented along the section and do not remove overhead structures like bridges and powerlines. These points are then grouped based on local connectivity, and segmented if they contain an above average number of points. This ensures that other objects key to the surrounding roadway (i.e. highway markers, overhanging traffic signs, small barriers, etc.) remain a part of the point cloud. Although LiDAR datasets have been used to assess stopping and passing sight distance for drivers [71], [72], limited work exists that extrapolates LiDAR-based sight distance assessments to roadside assets. Therefore, this thesis develops a novel voxel-based procedure to automatically assess maximum visibility for all traffic signs along a road segment.

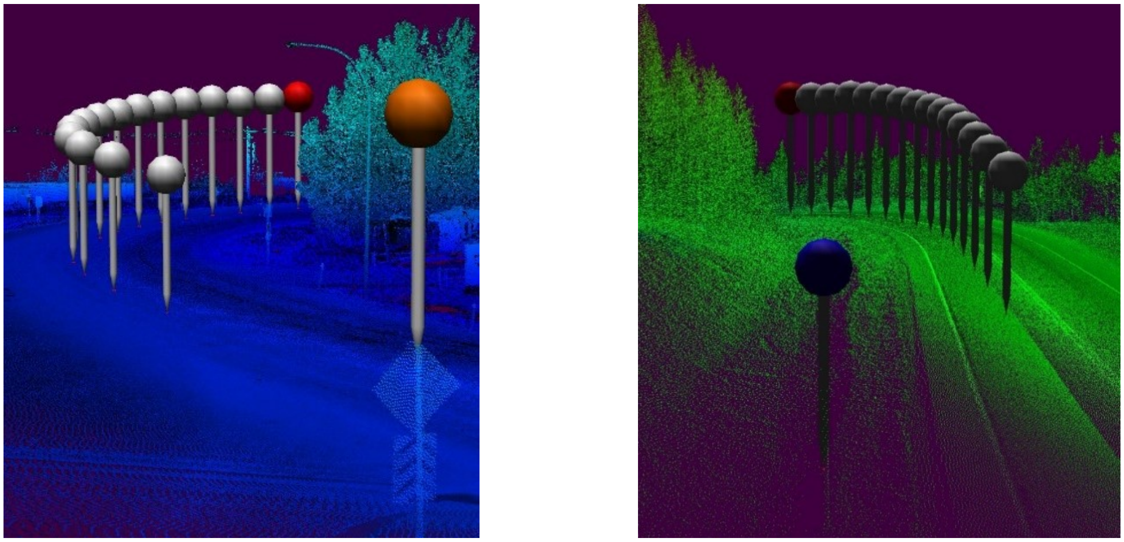
For an observer-target pair, the sightline connecting their voxel coordinates is connected and assessed for the existence of potential occlusions. If an occluding voxel is located along this specific sightline, the sightline is considered obstructed. Sightlines are created through a parametric representation wherein all voxels along a sightline are enumerated. For the observer voxel $V_O = i_O, j_O, k_O$ and the target voxel $V_T = i_T, j_T, k_T$, the points P along a sightline are defined by:

$$\mathrm{P} = \mathrm{V_O} + \mathrm{t}(\mathrm{V_T} - \mathrm{V_O}) \in \mathrm{t} = 0\text{: step: }1 \qquad 35$$

The level of detail extracted along the sightline is determined by the "step", which determines the number of increments of the parameter t analyzed along the sightline. Therefore, for a maximum distance of 4 km along a segment and a voxel size of 0.2 metres, a step size of 0.00005 ($4000\text{m}/0.2\text{m}$) is used to extract one voxel per step. This will create duplicate voxel enumerations at closer distances but ensures no voxels are missed along the sightline at larger distances.

For this study, the observers for the visibility assessment are the traffic signs extracted according to Section 3.1. Since, the aim of this assessment is to extract the furthest distance from which each traffic sign is visible, the centroids of signs are used as observer points. The target points are a set of trajectory points that run parallel to the road's profile and that are spaced at 10-metre increments along the segment and which are located upstream of the sign of interest. To account for American Association of State Highway and Transportation Officials (AASHTO) eye-height requirements [16] for the average driver in a travelling vehicle, the trajectory points were elevated by 1.05 metres.

Figure 3.22 shows the sightlines and the visibility distance as they are measured from the observers to the targets. Voxels along the sightlines were enumerated from each observer to all targets. These sightlines are then assessed for visibility by comparing the set of enumerated sightline voxels with the entire set of data, defining any sightline voxel that contains points as being occluded. Finally, the maximum visibility of a traffic sign can be measured considering the number of trajectory points visible upstream from each traffic sign. The available visibility distance for a traffic sign will be considered as the first break in visible trajectory points, as shown by the red pins in Figure 3.22 (a) and (b).



(a) (b)

Figure 3.22 – Legibility Distance Measurement

3.6.2 MINIMUM LEGIBILITY DISTANCE MEASUREMENT

The minimum legibility distance (MLD) for each traffic sign was determined for each extracted traffic sign through a comparison against the Alberta Traffic Sign Catalogue for traffic sign panel types [69]. In this database, the smallest detail present on the traffic sign governs the MLD, where 1 cm of letter height equals 3.6 metres of legibility distance [4], and 1 cm of symbol height equals 6.9 metres of legibility distance [18]. Examples of MLD presented in the catalogue are shown in Table 3.3. The addition of the visibility completes the TSI conducted within this thesis. Traffic signs along each segment can now be defined by their global position, relative placement and orientation measurements, classification, and upstream visibility. To determine the efficacy of the application of LiDAR and video-log data for the extraction of a TSI, the following section calculates and discusses the accuracy of these extractions and measurements.

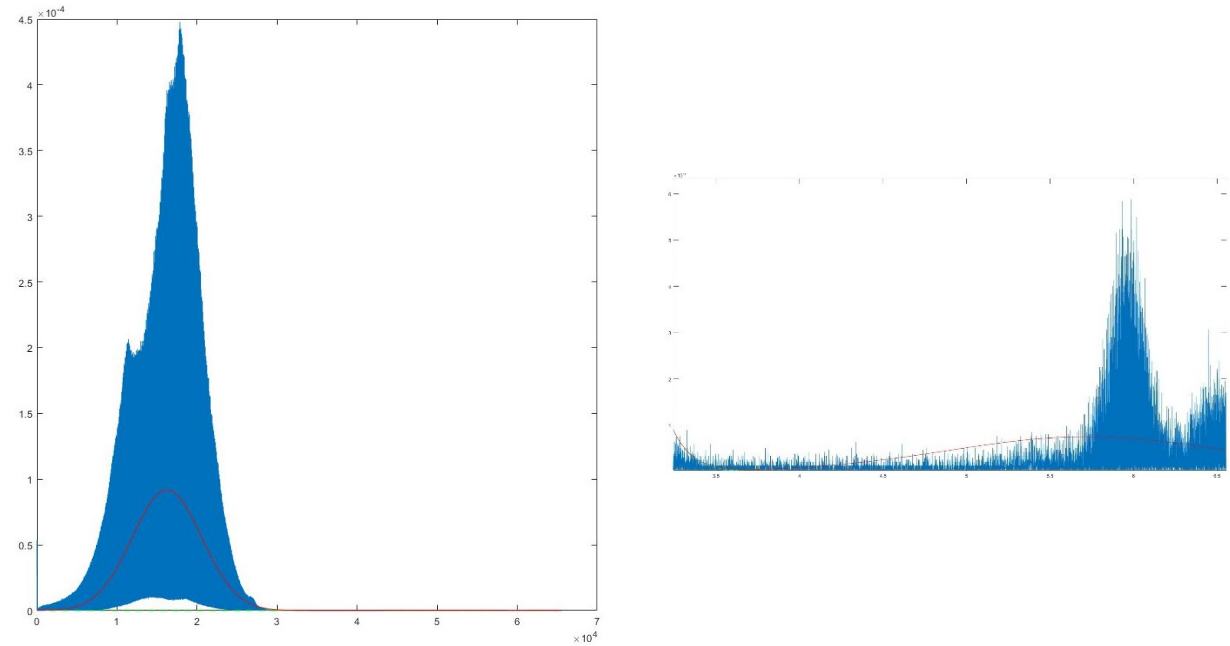
Table 3.3 – Minimum Legibility Distance

Classification	**Image** [69]	**Letter Height (mm)**	**Symbol Height (mm)**	**Legibility Index (m/cm)**	**Daytime Legibility Distance (m)**
IA-201		203	-	3.6	73.08
WA-3-L		-	600	6.9	414

4 RESULTS AND DISCUSSION

4.1 TRAFFIC SIGN EXTRACTION

A summary of the traffic sign inventory is available by applying the procedures in Chapter 3, with the completed inventory outlined in the Appendix. To assess the accuracy of this inventory, a comparison was conducted against available traffic sign information. An inventory of traffic signs with lateral placement, orientation, and classification is not available, so there is no immediate point of comparison for the measured attributes. However, the extraction of traffic sign locations was verified through a visual comparison against the objects located at the potential traffic sign coordinates on Google Earth [58].



(a) (b)

Figure 4.1 – Histogram with Gaussian Mixture Model Overlay

For the test segment, the retention of points within a standard deviation of the higher gaussian curve's mean, as illustrated in Figure 4.1 (a) and (b), resulted in 17,443 points belonging to traffic signs, passing vehicles, and roadway markers being collected. The DBSCAN clustering process is then applied where clusters have a maximum interpoint spacing of one metre and clusters must contain at least 17 points. After clustering the number of candidate traffic sign points (CTSP) reduces to 16,473, retaining all CTSPs and the sides of commercial trucks - a common occurrence along highways that were collected during scanning. The CTSPs are reduced to 73 clusters, which after the application of flatness filtering is reduced to 59 traffic sign objects. Figure 4.2 highlights the traffic sign extraction results on the LiDAR segment.

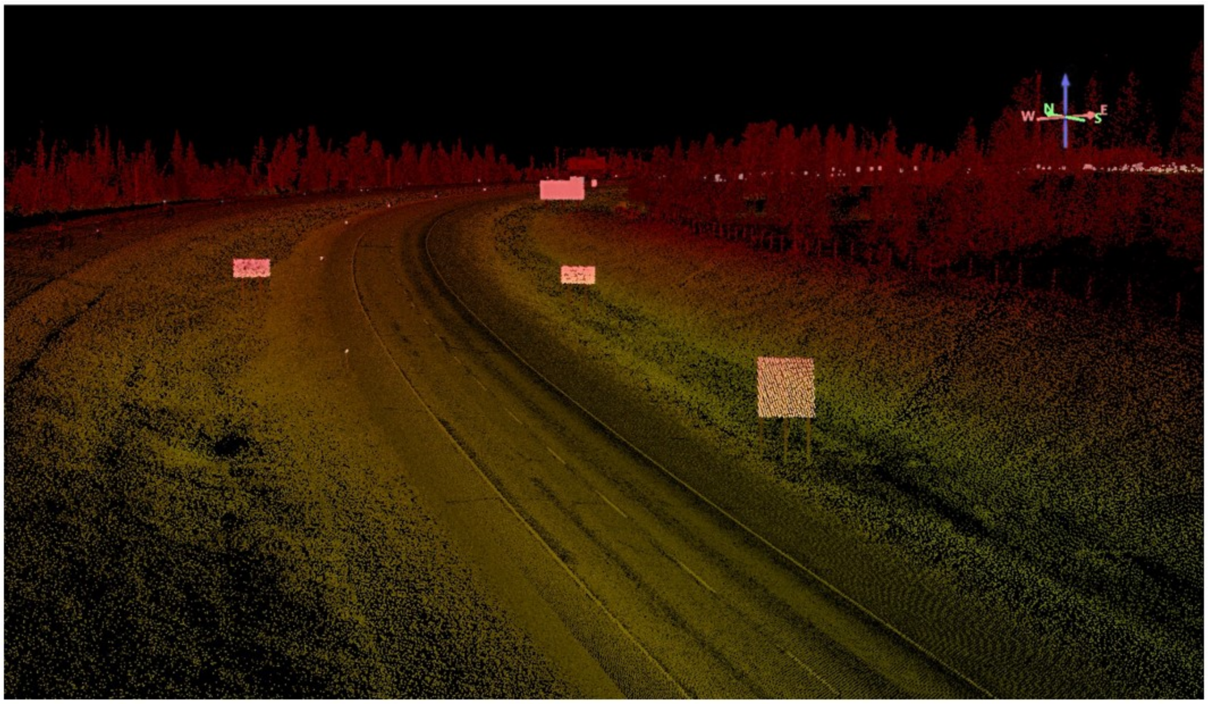


Figure 4.2 – LiDAR Point-of-View Traffic Sign Extraction

For the test segment, Table 4.1 details the extraction accuracy for the traffic signs along the test segment. The confusion matrix outlines the different types of correct and incorrect traffic sign classifications, where true positives (TP) constitute traffic signs that were classified as traffic signs, and false positives (FP) constitute non-traffic sign objects that were falsely classified as traffic signs. Similarly, false negatives (FN) constitute traffic signs that were falsely classified as non-traffic sign objects, and true negatives (TN) are not traffic signs that were correctly classified as not traffic signs. The number of TPs. FPs, and FNs is determined through a comparison against existing knowledge of traffic sign assets. For this thesis, as surveyed traffic sign information was unavailable, the locations of extracted traffic sign were verified visually from the LiDAR scans and through a comparison to Google images.

The confusion matrix results are then expanded to assess the overall extraction results using Equations 36 - 38. The Precision outlines a method's positive predictive rate, determining the proportion of the extracted positives represent true positive extractions. The Recall outlines a method's sensitivity, determining the proportion of false negatives that occur alongside the number of extracted true positives. Finally, the F1–Score calculates the harmonic mean of the Precision and Recall to describe the overall predictive power of the proposed method. This can be utilized in the future as a comparison metric for predictive power against other proposed methods.

Table 4.1 – Confusion Matrix for Extraction Accuracy

<table>
<tr><td colspan="2" rowspan="2">N = #</td><td colspan="2">Predicted Class</td></tr>
<tr><td>Not Traffic Sign</td><td>Traffic Sign</td></tr>
<tr><td rowspan="2">Actual Class</td><td>Not Traffic Sign</td><td>-</td><td>1 (FP)</td></tr>
<tr><td>Traffic Sign</td><td>5 (FN)</td><td>58 (TP)</td></tr>
</table>

$$\text{Precision} = \frac{\text{tp}}{\text{tp} + \text{fp}} \tag{36}$$

$$\text{Recall} = \frac{\text{tp}}{\text{tp} + \text{fn}} \tag{37}$$

$$\text{F1} - \text{Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{38}$$

The precision, recall, and accuracy of the traffic sign detection are 98.3%, 92.06%, and 95.08%, respectively. In other words, this procedure has false negative and false positive rates of 0.5 and 0.25 traffic signs per kilometre, respectively.

The single false positive, illustrated in Figure 4.3, is the metallic-sided trailer attached to a semi-truck. This type of noise is common along segments in Alberta with commercial vehicle traffic present at all hours of the day. Unfortunately, the high scanning density of the VMX-450 still provides a large quantity of points detailing the high-intensity trailer, and the side of the trailer is flat enough to pass through the filtering process. This risk of collecting commercial vehicles in LiDAR scans can be somewhat alleviated by scanning during night-time hours; however, the video-log camera array will not be useful in that scenario.

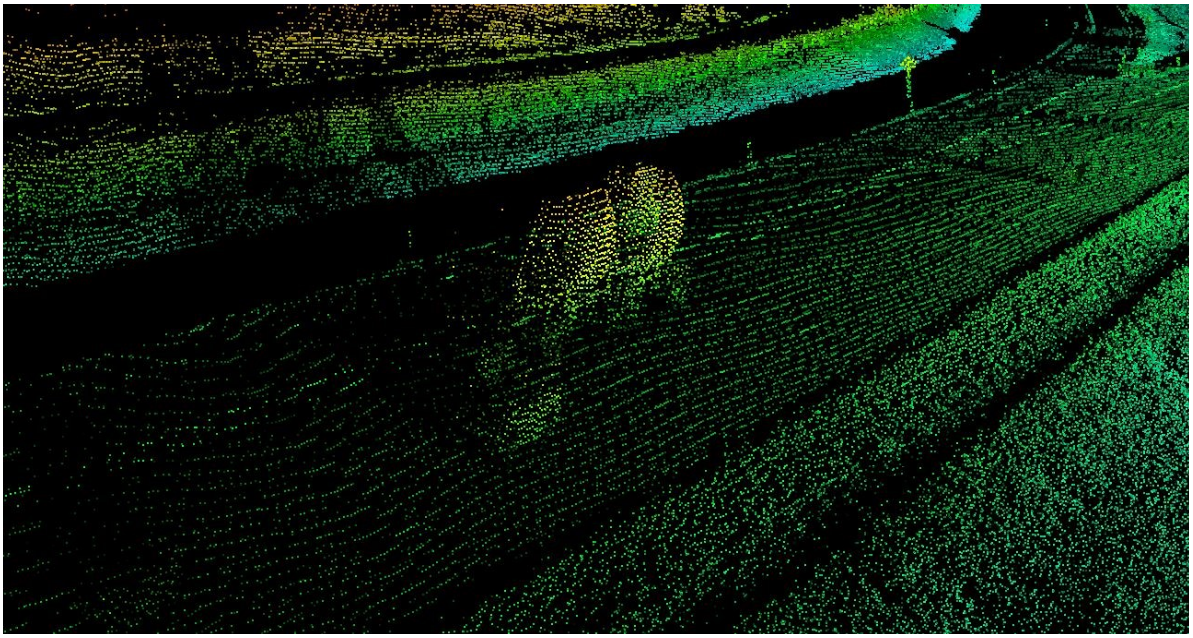
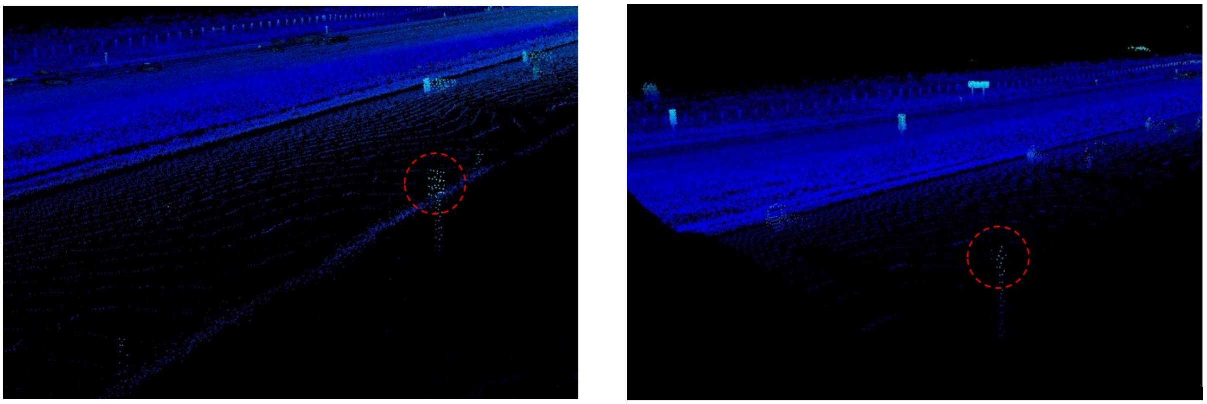


Figure 4.3 – False Positive Traffic Sign

The five false negatives are due to the reduction in point density at greater distances, with examples illustrated in Figure 4.4. The false negatives were traffic signs along the segment and not signs placed for advertising or other non-transportation purposes. As Highway 1 is divided along the analyzed segment, the East and Westbound traffic lanes are separated by a large grass median shown in Figure 4.4 (a). If only the traffic signs in the scanning direction (Eastbound) are considered, as this is the direction with the highest point density and the traffic signs pertinent to traffic in that direction, the precision, recall, and F1–Score are 98.3%, 100%, and 99.14%, respectively.



(a) (b)

Figure 4.4 – Examples of Missed Traffic Signs

The Gaussian Mixture Model extraction of traffic signs from a LiDAR point cloud relies on their reflective nature, which causes high-intensity return values. If all traffic signs retain the maximum retro-reflectivity they have when they are new, then they will remain within one standard deviation of the higher Gaussian curve. However, if traffic signs are aged (i.e. faded) or dirty, the retro-reflectivity may fall outside of the Gaussian curve. Ai and Tsai [24] encountered this during their study, with investigation revealing that these signs displayed intensity values of 29,490 (converted from 0.45, the normalized value original provided in the study). The lower intensity values were due to an aged parking sign with a deteriorated panel.

The next stage in the extracted TSI is an assessment of compliance of the measured traffic sign orientation to the angular placement standards. As per Section 1.2, the orientation of the traffic sign should be 3° down and away from the travel direction. The negative angles of horizontal orientation correspond to signs that are turned towards from the travel direction, and vice-versa for positive angles. Similarly, the vertical orientation of each traffic sign is calculated, where the negative angles correspond to signs that are tilted backwards (i.e. top side away from traffic), and vice-versa for the positive angles.

Table 4.2 – Traffic Sign Orientation

Traffic Sign ID	**Horizontal Orientation (°)**	**Vertical Orientation (°)**	**Traffic Sign ID**	**Horizontal Orientation (°)**	**Vertical Orientation (°)**
2	-5.36	-3.84	27	-2.72	-0.88
3	-8.53	-1.87	29	-8.52	-1.53

Traffic Sign ID	Horizontal Orientation (°)	Vertical Orientation (°)	Traffic Sign ID	Horizontal Orientation (°)	Vertical Orientation (°)
5	-0.64	-12.07	30	-8.60	-0.35
8	-8.51	-0.47	32	-56.66	-1.30
9	0.23	-2.06	34	0.69	-0.01
10	-7.94	-0.61	35	-8.61	-2.43
11	-5.77	-3.75	37	-11.71	-1.50
12	-5.98	-0.41	38	4.07	-4.15
13	-8.93	-2.08	45	-19.36	-0.51
14	-9.01	-8.41	47	-1.11	-3.00
15	-6.47	-3.56	49	3.47	-1.83
16	-12.21	-5.75	51	1.42	-2.52
20	-0.77	-1.20	57	5.02	-3.16
23	-3.15	-0.48	58	-0.40	-5.81
24	1.48	-0.06	59	2.20	-4.75

In practice, placing traffic signs at an exact angle orientation is difficult to accomplish. Therefore, it is expected that there will be some variation from the design standard and signs will not be exactly oriented vertically and horizontally at 3°. Nonetheless, only eight signs were faced away from traffic, with all other signs faced towards the travel lane. Additionally, all signs were tilted away from the travel lane, deviating from the standard of a 3° tilt towards traffic.

This is critical because the conspicuity of traffic signs during nighttime conditions is dependent on their placement and orientation relative to the traffic lane. The angle formed between the traffic sign and an approaching vehicle's headlights dictates the conspicuity of the traffic sign. Under normal orientation conditions, the angle between a vehicle's headlights and the traffic sign panel is small enough such that light is reflected back to the driver to illuminate the sign compared to the surrounding environment, making the traffic sign and it's details conspicuous and legible [73]. However, if the headlight beam and traffic sign are perpendicularly oriented, the sign is at its brightest. In some cases so bright that the details of sign cannot be distinguished from the rest of the traffic sign (i.e., specular reflection) [73] or even blinding.

Similarly, the vertical orientation of traffic signs plays an equal role in maintaining a traffic sign's conspicuity. Sign 5 shows the greatest vertical deflection away from the direction of travel as shown by the side view in Figure 4.5. Although the reason for this sign's condition is unclear, the placement of this traffic may jeopardize its conspicuity, particularly in nighttime conditions. The light from travelling vehicles would reflect upwards off the traffic sign panel instead of back towards the vehicle to make the sign visible. Therefore, drivers may not be able to read the sign until they are too close, possibly resulting in missing or only partially reading or understanding the sign's message. The remaining signs are all pitched backwards as well, which should aid to maintain sign conspicuity by ensuring drivers are not blinded by their headlights. However, Signs 8, 10, 12, 23, 24, 27, 30, 34, and 45 are all within a degree of horizontal and may contribute to specular reflection from the traffic signs.

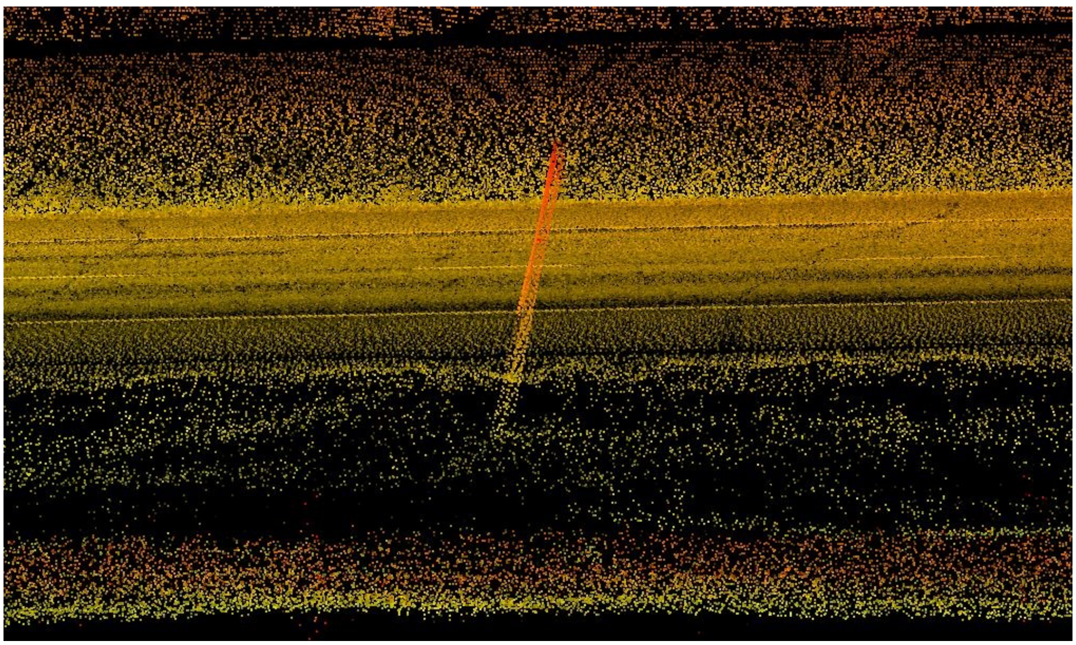


Figure 4.5 – Vertical Orientation of Traffic Sign 5

However, not all non-compliant traffic signs are truly non-compliant. Sign 32 is severely horizontally deflected from the travel direction but is a compliant traffic sign. This sign represents a special case, where this No U-Turn (RB-16) sign is used to warn traffic that the median turnaround cannot be used by local traffic to turn around. Therefore, although the proposed procedure provides a means with which to measure traffic sign orientation, the ultimate decision still requires a level of manual intervention by a trained professional as to the compliance of these measurements to sign placement standards.

Traffic sign orientation is fundamental to assessing and maintaining traffic sign conspicuity, but orientation is only part of the necessary inventory. Therefore, the next section will discuss sign placement measurements and the impact on the conspicuity of traffic signs.

4.2 Lane Marking Extraction

The lane marking extraction procedure for the measurement of lateral and vertical placement measurements is considered in stages within this research. Starting with the segmentation of the point cloud into ground and non-ground voxel classifications, the ground voxels are then further segmented to retain only the road surface voxels. This is once more segmented, resulting in cross sections that contain only road surface information and allow for future intensity-based procedures to locate lane markings. The discretization of the point cloud through multiple segmentation stages ensures the higher intensity surrounding vegetation and traffic signs do not interfere with the intensity adjustments used to extract the lane markings. The example in Figure 4.6 highlights the clear divide in the ground and non-ground segments of the point cloud.

Figure 4.6 – Example of Ground and Non-Ground Segmentation

As this process was conducted in voxel space, the segmentation of non-ground objects from the remainder of the point cloud is at the edges between these voxels. The highlighted non-ground points within the image include traffic signs and highway markers in the foreground and trees bounding the segment in the background – all objects that could be considered non-ground objects. The extraction of ground points is critical to the contrast enhancements utilized in lane marking extraction later. However, before the lane marking extraction can begin, an additional road surface extraction step is necessary to reduce the point cloud to road surface points only. This reduces intensity noise like low-lying foliage within the cross sections used for lane marking extraction. An example of this extraction is shown in Figure 4.7 for the current test segment. The results show the road surface is under-segmented, creating artifacts at the road edges. This results in additional ground surface being included as part of the road surface, but due to their small size, this does not negatively impact the creation of cross sections that follows.

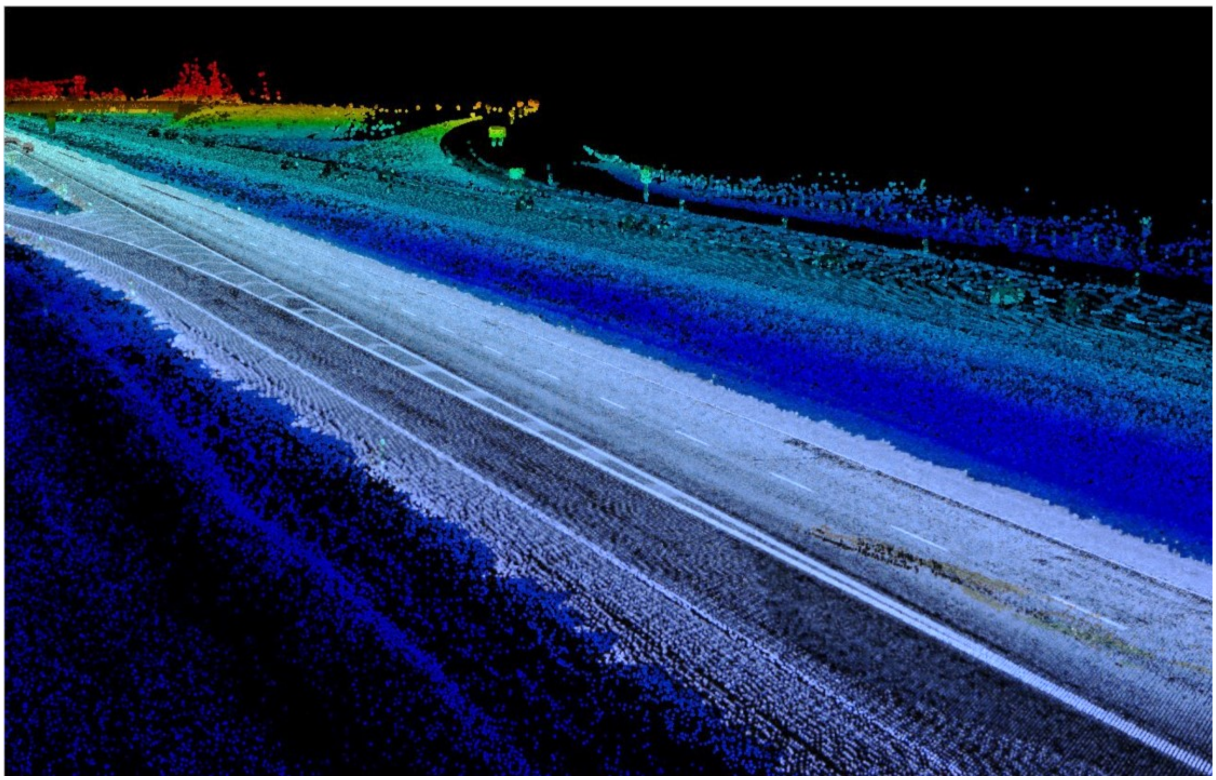


Figure 4.7 – Road Surface Extraction Results

Finally, the voxels that define the road surface are segmented into cross sections nearest each traffic sign. This is conducted to reduce overall processing time, applying the lane marking extraction procedure only to critical cross sections rather than to all the cross sections along the segment. The ACGWD is dependent on the assigned weighting parameter used to smooth a cross section raster's probability density function, which was determined experimentally as 0.8 using PCQI [62]. The PCQI heatmap is also provided in Figure 4.8, which illustrates the intensity reduction of the brighter lane marking and the intensity enhancement of the dimmer lane marking to improve the overall contrast of the lane markings within the entire image.

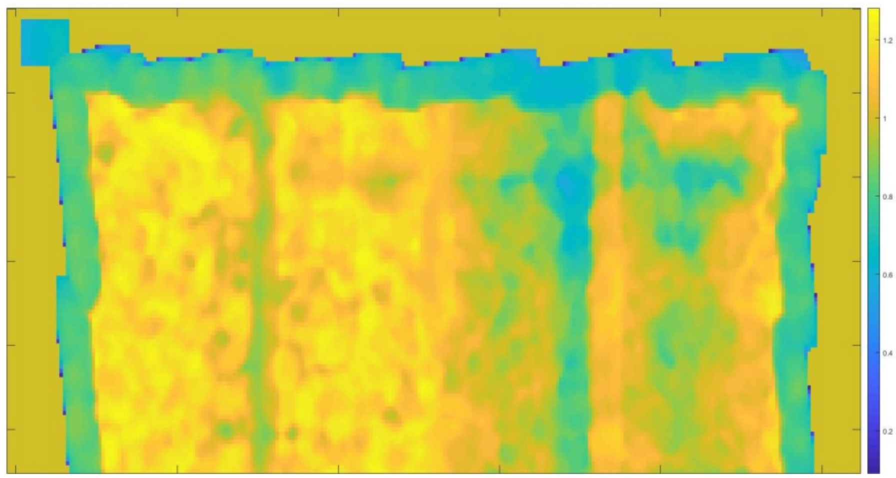
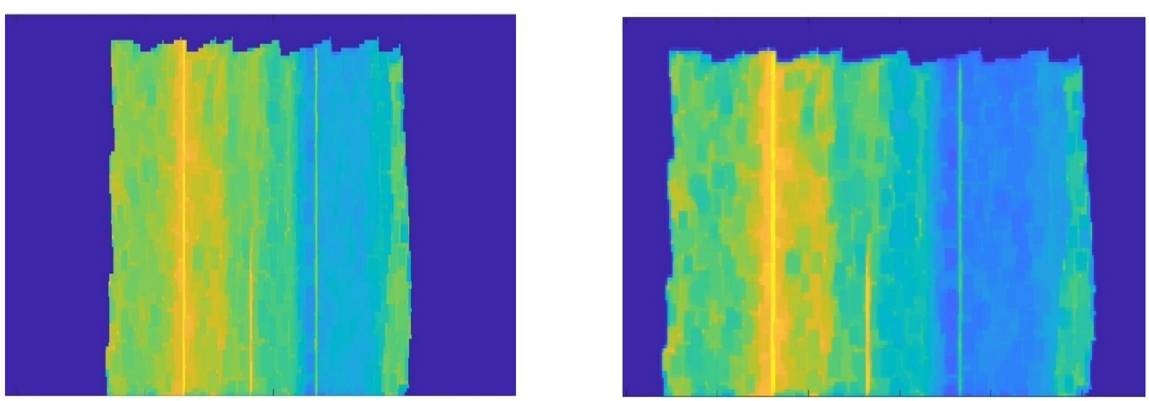


Figure 4.8 – Patch-based Contrast Quality Index of Cross Section

Finally, the aliasing artifacts caused by the rasterization process are smoothed through the application of a Gaussian filter with 0.5 standard deviation. An example of the contrast enhanced and smoothed cross section is illustrated in Figure 4.9 (a) and (b), respectively, effectively increasing the local contrast of the lane markings relative to the road surface.



(a) (b)

Figure 4.9 – Contrast Enhanced Lane Markings

The Canny edge detection is then applied to the raster image to determine the position of the lane markings. The Canny detector requires that the edge threshold be defined, which was empirically chosen as 0.15. An example of the edge detection result is illustrated in Figure 4.10, with the colored line segments showing the result of the Hough transform used to locate the linear edges within the raster.

Figure 4.10 – Lane Marking Edge Detection

With the lane marking extraction process complete, the results were manually assessed for completeness. For each extracted cross section, a lane marking extraction is considered complete if the left solid edge and right solid edge are all extracted, because only the left and right edge markings are required for the lateral and vertical placement measurements. The completeness of the lane marking extraction was measured for the same segment along which the traffic signs were extracted, with results in the confusion matrix in Table 4.3. There should be an equal number of left and right lane marking extractions, with the total number of lane markings less than or equal to the number of traffic signs. There may be fewer cross sections than there are traffic signs if two traffic signs exist on opposite sides of the travel lane, thereby measuring placement to the same cross section.

Table 4.3 – Lane Marking Extraction Confusion Matrix

N = #		**Predicted Class (Left)**		**Predicted Class (Right)**	
		Not Lane Marking	*Lane Marking*	*Not Lane Marking*	*Lane Marking*
Actual Class	*Not Lane Marking*	-	0 (FP)	-	3 (FP)
	Lane Marking	5 (FN)	42 (TP)	4 (FN)	43 (TP)
Precision (%)		100.00		93.47	
Recall (%)		89.36		86.00	
F1–Score (%)		94.38		89.58	

The primary causes of false negatives within the lane marking extractions include lower density within the point cloud and faded lane markings. Additionally, the point cloud density in adjacent travel lanes can be occluded or missing if a vehicle passes during scanning. The reduction of point cloud density along the right-hand side is illustrated by the voxel point density heat-map in Figure 4.11. The point cloud appears to be segmented along the trajectory of the scanning vehicle (i.e., the bright colored strip in middle), with a higher voxel point density along the left side and lower on the right side. The reduced point density causes segmentation in the lane markings on the right side, thus reducing the accuracy of the lane marking extractions.

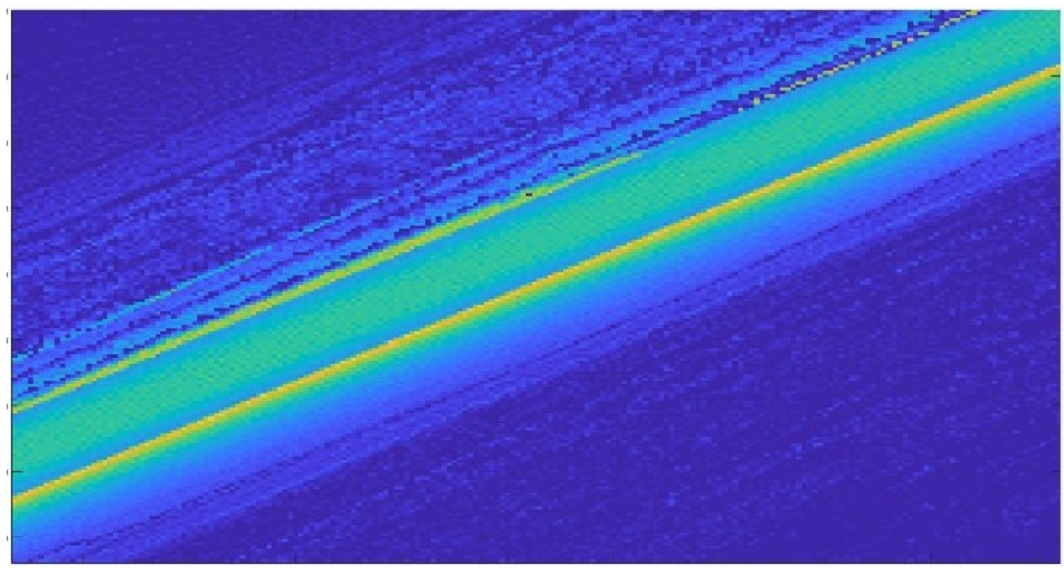


Figure 4.11 – Voxel Point Density Heatmap

The addition of lane markings allows for the completion of the traffic sign placement assessment. The lateral and vertical placement is measured from the nearest lane marking as dictated by the Alberta Highway Guide and Information Sign Manual [14]. For rural highways, a traffic sign is typically laterally placed 6.0 metres and vertically placed between 1.5 and 2.5 metres (2.0 metres ideal) from the nearest lane marking [14]. As the lane markings can only be assessed in the scanner's direction of travel, only the traffic signs in that direction are assessed for compliance with these placement standards in Table 4.4. Traffic signs are considered non-compliant if their lateral placement is below 6.0 metres or if the vertical placement is outside of the defined compliant range.

Table 4.4 – Compliance with Traffic Sign Placement Standards

Traffic Sign ID	Lateral Placement	Vertical Placement	Traffic Sign ID	Lateral Placement	Vertical Placement
2	5.40	1.56	27	6.08	1.80
3	5.16	1.37	29	5.17	2.09
5	3.87	1.51	30	3.16	2.05
8	5.72	1.70	32	4.47	0.73
9	2.50	2.04	34	6.06	2.78
10	13.55	2.08	35	4.70	1.74
11	6.79	1.56	37	19.66	1.26
12	10.14	1.53	38	2.18	1.36
13	2.01	1.44	45	0.57	6.21
14	1.51	1.62	47	5.22	1.71
15	9.66	1.48	49	2.56	2.11
16	6.98	1.80	51	-999	-999
20	5.71	2.16	57	4.70	1.58
23	4.10	1.67	58	4.68	1.16
24	0.86	2.02	59	5.01	1.27

Similar to the orientation of traffic signs, when mounting traffic signs on their respective poles a level of error is expected due to the manual process of installing these signs. Therefore, traffic signs are not expected to have an exact lateral placement of six metres or vertical placement at the ideal values. Nonetheless, 60% of the Eastbound traffic signs were placed too close to the travel lane, 6.67% were placed too high, and 23.3% were placed too low.

The lateral and vertical placement of traffic signs is critical as these factors directly relate to the visual angle between signs and drivers along the roadway, thereby affecting the sign's conspicuity. Under nighttime conditions, Zwahlen [74] found that for a 10°, 20°, and 30° visual angle from the central visual area, the upstream distance a retroreflective object is detected at decreases to approximately half, one-third, and one-quarter, respectively. Therefore, a traffic sign placed laterally and vertically further away from the central visual area requires drivers to be much closer to the sign before they can detect it. The conspicuity of the traffic sign decreases with distance, reducing the time available for drivers to read and react to the traffic sign.

Although the reflectivity of traffic signs cannot be directly assessed from the LiDAR data [22], the assessment of placement conditions allows for the identification of deviations from design conditions. Traffic signs that deviate from design conditions should be under scrutiny to ensure they meet reflectivity requirements. Previously, the assessment of placement was tedious, but the addition of this method allows us to determine compliant and non-compliant signs and streamline additional traffic sign maintenance.

The eight traffic signs that are placed too close to the travel lane represent conditions wherein traffic signs may run into the problem of specular reflection [73], which could result in illegible traffic sign details or the blinding of drivers. In contrast, Signs 10, 12, and 37 represent conditions where the lateral placement may reduce the traffic sign's conspicuity due to the reduction of reflected light caused by further placement. However, Signs 12 and 37 belong to exit lanes and their placement is thereby directed at exiting traffic – not directly impacting eastbound through traffic.

The mounting height of traffic signs have similar impacts, wherein a high mounted sign may be too far outside a driver's peripheral field-of-view to be noticed. Vice-versa, mounting a sign too low may cause it to be missed by commercial vehicle drivers or increase the visual workload on the regular driving population. Sign 34 was mounted higher than the design specification, but this may be due to the design engineer's judgement. This exit sign relays information about the exit lane immediately followed by the sign. Therefore, the sign may be mounted higher to accommodate its visibility at a larger distance upstream in both travel lanes.

Additionally, some special placement conditions were noted within the set of traffic signs. The cross section at Sign 37 did not successfully extract the lane marking closest to the sign, thereby making this calculation potentially inaccurate. Sign 45 is a bridge-mounted clearance sign, therefore the lateral and vertical placement criterial do not apply. Signs 13 and 38 are exit signs within the gore area and have lowered than required lateral placement measurements. However, the gore area can be space limited, thereby limiting the possibility of compliance to these placement standards. Signs 12 and 37 are signs of concern for already exited traffic. Their placement measures are based on the lane marking within the through travel lane, meaning that these placements are potentially incorrect. Visually however it can confirm that their placement serves traffic that has taken the exit. Finally, the lane markings at the cross section for Sign 51 were missed altogether, thereby having no points from which to measure. The placeholder values of -999 were used to indicate that these values will need to be measured through a different process.

The placement and orientation work in unison to create traffic sign placements that ensure the greatest conspicuity. Especially during nighttime conditions when overall visibility is decreased, traffic signs serve to inform and guide the driving population [73]. If the placement of traffic sign is non-compliant with design standards, the traffic signs may hinder rather than help drivers if they are blinding, attribute to high visual workload, or if they are not visible at all.

4.3 TRAFFIC SIGN CLASSIFICATION

The collection of a traffic sign database for the purpose of neural network image classification, like those found in the GTSRB and BTSC, is the first of its kind in the Canadian context. This work represents a sample of natural traffic sign placement in the Albertan context, key to the classification of traffic sign conditions experienced by Canadian drivers.

The increased capabilities of computer hardware in recent decades initiated the resurgence of neural network models for classification as the neural network training times significantly decreased. Therefore, the training time for the Inception neural network structure is of utmost importance for the assessment of the compatibility of current neural network structures to other image classification procedures. Including pooling layers, the GoogLeNet network contains 27 layers in total [67]. The network was trained with stochastic gradient descent with momentum across eight epochs, a mini-batch size of 128 images, and a learning rate of 0.001. Dropout was also utilized, where a random 40% of the neurons are dropped within each iteration to attempt to reduce overfitting. The initial weights of the convolutional and fully connected layers are randomly generated from a Gaussian distribution with mean 0 and standard deviation of 0.01.

The network training times are summarized in Table 4.5. Training of the CNN structure was conducted on two platforms: a 4-core Intel i7-7700 at 4.2 MHz with 16GB DDR4 RAM and a Nvidia GTX 1070 GPU with 8GB video RAM, and an 4-core Intel i7-4790 CPU at 3.6 GHz with 32 GB DDR4 RAM. The two PCs are used for GPU-based and CPU-based training of the CNN respectively. The results indicate the direct advantage of having a GPU available for network training, with the GPU requiring 16.45 times less training time. This improves the overall workflow when training a neural network for any application, allowing for faster responses by the user to original network structure when making improvements. Although this is still manageable with CPU training, the time in-between training attempts greatly decreases the efficiency of training and adjusting a classifier. The five GPU trained classifiers were randomly initialized, resulting in an average training time and mean average precision of roughly 29 minutes and 83.39% respectively.

Table 4.5 – Comparison of Neural Network Model Training

Iteration	Training Time (s)	Accuracy
1	00:29:12	84.16
2	00:28:57	83.33
3	00:29:09	82.96
4	00:29:11	83.06
5	00:29:13	83.45
Mean	00:29:10	83.39
CPU	08:00:00	83.60

The classification capability of the trained network was determined through two conditions i) based on a test set from the manually classified Alberta traffic sign database, and ii) based on the results of the classification results along the test segment video-log images. The two test sets are differentiated because the lower sample size of traffic signs along the segment does not completely describe the accuracy of the classifier, but it does allow for the assessment of missed classifications. Furthermore, the training dataset is only partial and there might be traffic signs placed along the test segment that were not trained in the image classifier. This allows for a comprehensive assessment of accuracy and the continued expansion of the training dataset. Figure 4.12 provides the confusion matrix for the classification of the validation images.

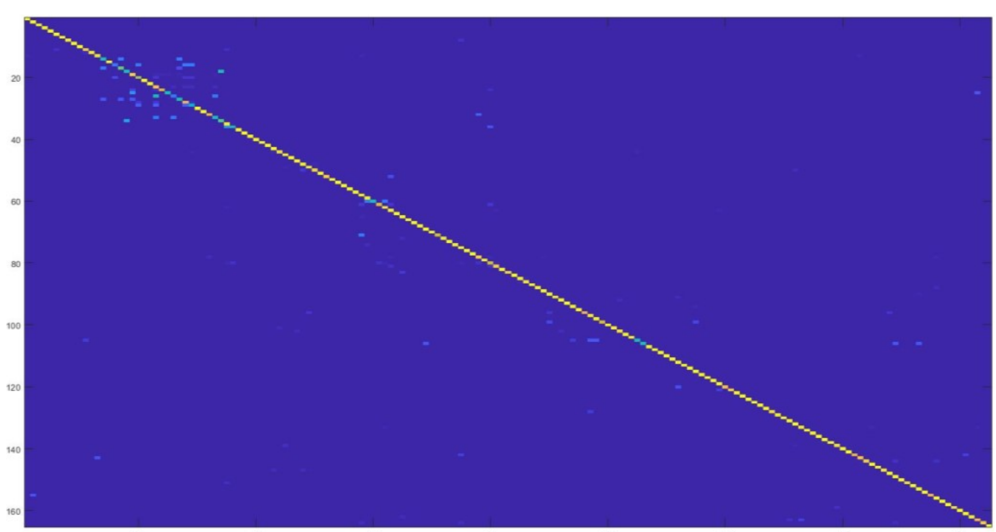


Figure 4.12 – Confusion Matrix for Classification

The confusion matrix visually describes any "confusion" in the classification attempts, where a perfect classifier would show 100% of the classifications along the diagonal (i.e., all the images were correctly classified as its class). Any "confusion" that occurs when classifying images is colorized in the off-diagonal entries, where pixel brightness corresponds to the number of classifications within that pair of Validation and Actual classes.

The network resulted in a Top-1 error rate of 10.35% and a Top-2 error rate of 3.24%. Of the 165 total classes, five classes were not identified in the Top-1 classifications, but only two were not identified in the Top-2 classifications. Therefore, 163 of the 165 classes of true positives and false positives were correctly identified in the first two predictions by the classifier.

The two classes that were not included in the Top-2 were "*Disability Access*" (IC-14) and "*Museum-Tab*" (IC-20-T). These misclassifications are largely due to their similarity to other "*information*" traffic signs classes with which they are typically mounted. The "*Take a Rest*" (IC-241) and other information signs (IC-#) are commonly combined on the same sign panel near exits, reducing the uniqueness between instances of training images and furthering the possibility of misclassification. As the training dataset only represents a small sample of traffic signs in Alberta this is not unexpected and strengthens the need for additional training samples, especially for the classes with sample sizes much lower than the mean sample size.

The GoogLeNet structure provides the depth needed to conduct accurate classifications across many classes. The network structure attempts to reduce overfitting by utilizing dropout, but for the entire test set, IA-201 was the Top-2 classification 54.5% of the time, and IA-201 and IA-202 were the Top-3 classification 18.7% and 53.3% of the time for all classes, respectively. This is a side-effect of the small training sample size in total and the variety available in the two classes. More specifically, the two classes exhibit large variation in their backgrounds, causing the tree-lines and general foliage present within these classes to be a common classification outcome of the network whenever faced with these features in an image. The image classifier was also applied to the nearest video-log images to each traffic sign, as presented in Section 3.3.3. The classifier's accuracy is analyzed based on a manual classification of the traffic signs that exist in the same direction of travel. The image classifications are provided in Table 4.6.

Table 4.6 – Traffic Sign Classification from Video-log Images

Traffic Sign ID	Predicted Classification	Actual Classification
2	IA-210 (quadruple-direction)	ID-204 (First Nation)
3	RB-35 (info-slow-traffic-keep-right)	RB-35 (info-slow-traffic-keep-right)
5	WA-10B (ramp ahead advisory speed)	*
8	WA-10B (ramp ahead advisory speed)	WC-2A (watch-for-pedestrians-on- highway)
9	WA-10B (ramp ahead advisory speed)	WC-2A (watch-for-pedestrians-on- highway)
10	RB-1-100	IF-203 (next-exit)
11	RB-1-110	**
12	RB-1-110	*
13	RB-1-110	*
14	RB-1-110	*
15	RB-1-110	RC-104 (do-not-cross-median)
16	RB-35 (info-slow-traffic-keep-right)	RC-104 (do-not-cross-median)
20	IF-203 (next-exit)	IF-203 (next-exit)
23	IC-5 (picnic-table)	IC-216A
24	RB-1-110	IF-203 and IC-212
27	IC-57 (trailer and tent)	IC-216A (major-attraction-for-conventional-highway)
29	WA-26 (bridge)	WA-26 (bridge)
30	WA-26 (bridge)	WA-26 (bridge)
32	IC-5 (picnic-table)	RB-16 (no-u-turn)
34	WA-10B (ramp ahead advisory speed)	IF-203 and IC-212
35	WA-10B (ramp ahead advisory speed)	WA-10B (ramp ahead advisory speed)
37	IF-205 (exit with number)	WA-10A (ramp-advisory-speed)
38	IF-205 (exit with number)	IF-205 (exit with number)
45	WA-27 (bridge clearance)	WA-27 (bridge clearance)
47	WA-10B (ramp ahead advisory speed)	WC-111 (caution-logging-trucks-next-x-km)
49	WA-16-R (merge-right-arrow)	WA-16-R (merge-right-arrow)
51	WD-101 (construction-with-arrow)	ID-411 (caring-for-alberta's-highways)
57	RB-1-100	IA-205 (double-distance)
58	IA-203 (single-distance-and-arrow)	IB-1 (transcanada-highway-1-marker)
59	RB-1-100	*

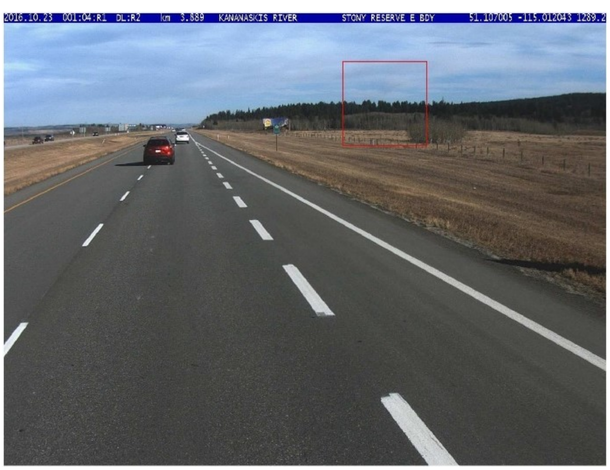
** Traffic sign does not exist in video-log or on Google Earth*

*** Sign post visible but panel missing in video-log and Google Earth*

For the thirty eastbound traffic signs, considering only the signs that were classes available in the classifier to begin with, the precision and recall were 53.3% and 100%, respectively. Six of eighteen classes were new to the classifier, solidifying the need for continued development of the training dataset. This is primarily due to the sliding window used to scan the input images for traffic signs and scans the surrounding area. Examples are provided in Figure 4.13. The sliding window passes over sequences in the image that mimic background commonly seen in other traffic sign classes, resulting in the network relying on these features rather than the traffic sign's features to conduct the classification. This suggests that additional work is needed in defining additional false positive classes like skylines and tree-lines.

(a)

(b)



(c)

(d)

Figure 4.13 – Sliding Window Misclassifications

In the absence of the traffic sign within the image, as in the case of Signs 12-14 and 59, the lane markings present across most of the video-log images are typically confused with the RB-1-100 and RB-1-110 (i.e., maximum speed of 100 or 110 km/h; black and white signs) classifications. This misclassification also occurred for Signs 10, 24, and 57. For the application of the classifier to the segment's video-log images, the RB-1-110 and RB-1-100 signs appeared most often as false classifications. The bounding-box examples provided in Figure 4.13 show that when the classifier is passed over areas with heavy trees, skylines, and powerlines, the RB-1-# signs were still obtained with high accuracy. This may be explained by their respective sample sizes of 307 and 195, which is well above the average for all classes and may have attributed to over-training. Additionally, the RB-1-# signs feature a prominent white linear edge around its borders that could be visually confused with a white lane marking.

The remaining traffic signs were misclassified under other circumstances. The true classification of Signs 8 and 9 were unavailable within the training data. However, these traffic signs were mis-classified as a similarly colored traffic sign. Sign 5 did not exist within the video-log images, but the following downstream traffic sign was present and correctly classified. Therefore, although the downstream sign was correctly classified, the traffic sign of interest was missing and thereby misclassified.

As previously discussed, the training image dataset used to train the image classifier does not encompass all the traffic sign classes used by Alberta Transportation. Recognized as confused classifications, the following represents traffic sign classes that are located on the test segment(s) that were not included in the training dataset:

- *First Nations (ID-204)*
- *Do Not Cross Median (RC-104)*
- *Trans-Canada Highway 1 Marker (IB-1)*
- *Caution Logging Trucks Next 'X' KM (WC-111)*
- *Major Attraction for Conventional Highway (IC-216A)*
- *Tourist Attraction (IC-212)*

Compared to the 50,000 traffic sign samples found in the GTSRB, this work is still in progress and only represents a proportion of the traffic sign samples in Alberta. However, the classifier made accurate predictions on the validation dataset and provided valuable classifications when compared to the video-log images. This dataset was collected to make image-based traffic sign classification more accessible to municipalities, industry professionals, and researchers. The assessment of the image-based traffic sign classification highlights its application to the creation of a traffic sign inventory. The final contribution of this thesis is presented in the following section, utilizing the depth of LiDAR scans for the extension of a traffic sign inventory to include measurements of traffic sign visibility.

4.4 VISIBILITY ASSESSMENT

The extracted inventory of traffic signs is extended to determine their visibility along the segment. Fundamental to the legibility of traffic signs, this assessment serves to discuss the application of LiDAR to visibility assessment and the impacts of traffic signs with limited visibility. For this assessment the eastbound traffic signs described in Sections 4.1 and 4.2 will be used. The test segment has 16 vehicles along the eastbound approach that could potentially occlude visibility measurements. The proposed method successfully removed vehicles with precision and recall of 93.75% and 100%, respectively, with an example of the removed vehicles shown in Figure 4.14 as brightened points. The rear-end of one vehicle along the segment remained within the point cloud due to its segmentation from the rest of the vehicle.

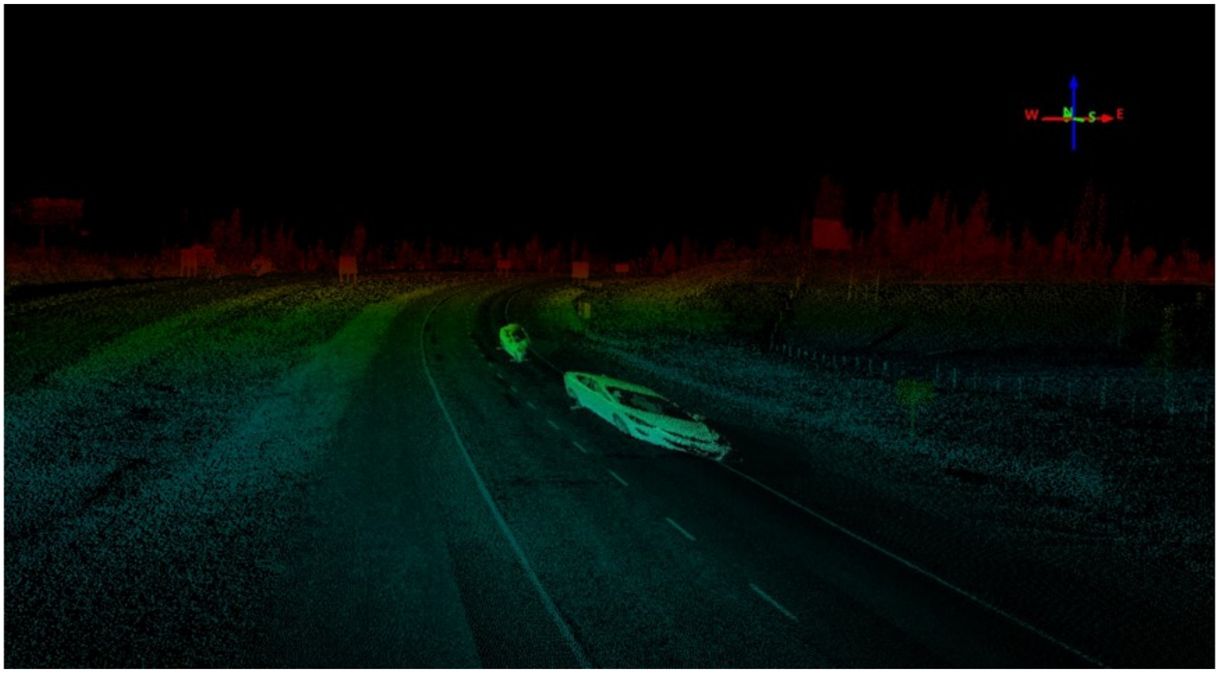


Figure 4.14 – Vehicle Removal from Scanner Travel Lane

Table 4.7 shows the results of the traffic sign visibility assessment using the proposed method, where traffic sign legibility was manually calculated by considering the smallest detail on the sign provided by the Alberta Traffic Sign Catalogue [69]. It should be noted that the signs that could not be classified in the previous section due to missing validation information were not assigned legibility values.

Table 4.7 – Visibility Measurements of Segment Traffic Signs

Sign	Lat	Long	Visibility (m)		Class	Legibility (m)	Speed Limit (km/h)	Time (s)	
			Day	Night				Day	Night
2	51.08604	-115.054	75.369	41.871	ID-204	135.00	110	2.466	1.370
3	51.08604	-115.054	159.022	88.345	RB-35	54.72	110	5.204	2.891
5	51.08604	-115.054	296.191	164.550	*	0.00	110	9.693	5.385
8	51.08801	-115.054	130.976	72.764	WC-2A	72.00	110	4.286	2.381
9	51.08604	-115.054	353.973	196.651	WC-2A	72.00	110	11.584	6.435
10	51.08862	-115.053	223.414	124.119	IF-203	73.08	110	7.311	4.062
11	51.08879	-115.053	360.103	200.057	**	0.00	110	11.785	6.547
12	51.08993	-115.052	380.123	211.179	*	0.00	110	12.440	6.911
13	51.08961	-115.053	421.375	234.097	*	0.00	110	13.790	7.661
14	51.09202	-115.050	188.881	104.933	*	0.00	110	6.181	3.434

Sign	Lat	Long	Visibility (m)		Class	Legibility (m)	Speed Limit (km/h)	Time (s)	
			Day	Night				Day	Night
15	51.09091	-115.051	501.102	278.390	RC-104	54.72	110	16.399	9.110
16	51.09047	-115.052	563.033	312.796	RC-104	54.72	110	18.426	10.236
20	51.09145	-115.051	1046.719	581.510	IF-203	91.44	110	34.256	19.031
23	51.09276	-115.048	1070.031	594.461	IC-216A	91.44	110	35.019	19.455
24	51.09251	-115.049	1226.438	681.354	IF-203 & IC-212	91.44	110	40.137	22.298
27	51.09359	-115.046	1168.379	649.099	IC-216A	91.44	110	38.237	21.243
29	51.09407	-115.045	1243.076	690.597	WA-26	90.00	110	40.682	22.601
30	51.09436	-115.044	1182.589	656.994	WA-26	90.00	110	38.702	21.501
32	51.09431	-115.044	1270.806	706.003	RB-16	414.00	110	41.590	23.105
34	51.09335	-115.047	1536.846	853.803	IF-203 & IC-212	91.44	110	50.296	27.942
35	51.09451	-115.044	1332.847	740.470	WA-10B	64.08	110	43.620	24.233
37	51.09978	-115.030	464.937	258.298	WA-10A	64.08	110	15.216	8.453
38	51.09490	-115.043	1487.700	826.500	IF-205	109.80	110	48.688	27.049
45	51.09301	-115.047	2231.838	1239.910	WA-27	72.00	110	73.041	40.578
47	51.09412	-115.045	2108.632	1171.462	WC-111	72.00	110	69.009	38.338
49	51.09402	-115.045	2306.068	1281.149	WA-16-R	465.75	110	75.471	41.928
51	51.09412	-115.045	2380.262	1322.368	ID-411	45.00	110	77.899	43.277
57	51.09596	-115.040	2319.960	1288.866	IA-205	73.08	110	75.925	42.181
58	51.10269	-115.023	984.414	546.897	IB-1	207.00	110	32.217	17.898
59	51.09407	-115.045	2784.057	1546.699	*	0.00	110	91.114	50.619

NOTE: Visibility is the distance at which a driver can first see a traffic sign; Legibility is the distance at which a driver can first read a traffic sign.

The assessment of visibility can be used to highlight locations at which the theoretical legibility distance (i.e., the distance at which a driver can first read a sign) greatly exceeds the available visibility (i.e., the distance at which a driver can first see a sign). These locations are of interest because, although the traffic signs are still visible in advance, the distance at which they are visible may not be enough for the average driver to perceive them since they are not visible for the entirety of the legibility distance.

The last column of Table 4.7 converts the available visibility measurements into available time based on the posted speed limits. For instance, the daytime legibility for Sign 5 was determined by dividing the visibility (196.65 metres) by the segment's posted speed limit (110 km/h, or 30.56 m/s) to convert the visibility into an available time of 11.58 seconds. This means that the traffic sign is only visible for 11.58 seconds. In that time, a driver is expected to perceive and read the sign. However, traffic signs are less visible during nighttime conditions, with the lighting to see a sign reduced to only what is visible within a vehicle's headlights [16]. Zwahlen and Schnell [18] determined that nighttime legibility distances decreased by 1.8 times compared to daytime conditions. Therefore, the 5th column in Table 4.7 assumes a reduction in the available visibility distances for comparison against the legibility distances in nighttime conditions. This is critical as any traffic signs that are already limited in visibility during the daytime are even less effective during nighttime conditions. If the available time to perceive a traffic sign is less than what is required for the average population, this increases the risk of only partially recognizing the sign or missing it altogether.

During daytime conditions Sign 2 has low visibility due to its placement towards the start of the segment. As the LiDAR data does not extend further upstream of this sign, its visibility is prematurely cut off by the edge of the segment within the LiDAR scan. Otherwise, the remainder of the traffic signs along the segment post high visibility, which is expected of signs posted along high-speed, high-volume highway segments. This is based on a comparison to the AASHTO standard of 2.5 seconds for preception reaction time (PRT) [16], with all of the traffic signs in Table 4.7 displaying visibility measures exceeding the minimum time required by the driver to perceive the traffic sign.

Notably Signs 3, 8, 10, and 14 all have visibility times below five seconds during nighttime conditions. This is important because, with an aging population, the visual acuity of the average driver typically decreases [17] and PRT of the average driver will increase. Hence, for the same traffic signs, the required visibility of older drivers increases compared to younger drivers because they are slower to perceive, read, and react to traffic signs in their periphery [75]. These characteristics are worsened in complex scenarios, either through visual complexity, decreased sign luminance, or small letter or symbol size [75]. Therefore, when considering a PRT of 5.0 seconds for an aged population, Signs 3, 8, 10, and 14 fail to provide enough perception time.

This is particularly important when considering the types of signs which were not visible within the required PRT. During daytime conditions, a "*Watch for Pedestrians on Highway*" sign (WC-2A) has the least available perception time. During nighttime conditions, a "*Slow Traffic Keep Right*" (RB-35) and a "*Watch for Pedestrians on Highway*" (WC-2A) sign have available perception times less than 2.5 seconds. If a driver is distracted, or if their perception reaction time falls significantly above the average of 2.5 seconds, drivers could impede traffic in the left lane or be less attentive to hitchhiking pedestrians along the highway's shoulders.

The visibility measurements provide the absolute maximum visibility distance and time available for the traffic signs along this segment. However, just because a traffic sign is visible does not mean it has a driver's attention. To further this discussion, each target point along the roadway is analyzed to determine which traffic sign is the nearest to the driver. It is assumed that the traffic sign nearest to each driver will be the sign of focus for the driver. Therefore, Table 4.8 highlights the decreased visibility of each traffic sign assuming they can only be visible consecutively along the highway. It should be noted that if traffic signs are located side-by-side across the travel lanes, the consecutive placement only illustrates one of the two signs. This is the case for Signs 9, 13, 16, 30, 38, and 59, all of which were present adjacent to another sign and are not included in Table 4.8.

Table 4.8 – Visibility Measurements of Signs based on Consecutive Placement

Sign	**Visibility (m)**		**Legibility (m)**	**Time (s)**	
	Day	**Night**		**Day**	**Night**
2	75.369	41.871	135.00	2.466	1.370
3	108.962	60.534	54.72	3.566	1.981
5	155.969	86.649	0.00	5.104	2.835
8	71.488	39.715	72.00	2.339	1.299
10	184.099	102.277	73.08	6.025	3.347
11	182.107	101.170	0.00	5.959	3.311
12	183.896	102.164	0.00	6.018	3.343
14	128.881	71.600	0.00	4.217	2.343
15	162.483	90.268	54.72	5.317	2.954
20	492.891	273.828	91.44	16.130	8.961
23	270.947	150.526	91.44	8.867	4.926
24	124.850	69.361	91.44	4.086	2.270
27	188.368	104.649	91.44	6.164	3.424
29	191.613	106.452	90.00	6.270	3.483
32	98.085	54.491	414.00	3.210	1.783
34	82.672	45.928	91.44	2.705	1.503
35	59.697	33.165	64.08	1.953	1.085
37	253.539	140.855	64.08	8.297	4.609
45	371.609	206.449	72.00	12.161	6.756
47	126.683	70.379	72.00	4.145	2.303

Sign	Visibility (m)		Legibility (m)	Time (s)	
	Day	Night		Day	Night
49	203.029	112.794	465.75	6.644	3.691
51	115.761	64.312	45.00	3.788	2.104
57	349.139	193.966	73.08	11.426	6.347
58	96.155	53.419	207.00	3.146	1.748

When considering the consecutive placement of traffic signs, their available visibility distances and times decrease substantially. Three signs do not meet the 2.5 second average PRT under consecutive placement conditions during the daytime, and eleven signs do not meet the 2.5 second average PRT under consecutive placement conditions during the nighttime. Therefore, if a driver suffers from higher than average PRT due to aging, or is driving distracted or fatigued, a series of signs along the segment may be missed or only partially recognized. To avoid overloading drivers with a limited amount of time dedicated to each individual sign, agencies need to perform traffic sign visibility assessments to ensure that all essential signs are visible, and sign placement overlaps are limited to scenarios with low complexity and high visibility.

Canada's median population age is expected to continue increasing, creating the possibility that inadequate placement of traffic signs may result in increased frequency of emergency braking and maneuvering scenarios, wherein drivers have to react to scenarios later than usual (due to limited time to read, perceive, and react to traffic signs). Similar concerns can be applied to the heavy tourist population in Alberta. A large proportion of Alberta tourism is focused on the Rocky Mountain Range on the west border of the province, both during summer and winter months. With drivers from all parts of the world and with different language backgrounds, the efficacy of a traffic sign network determines the ease with which the tourist population can safely navigate Alberta's highway networks. As they are likely unfamiliar with Alberta's highways, the guide and warning signs are the most important for ensuring a tourist's safe travel. As an additional measure of traffic sign efficacy, the visibility measurement from LiDAR data allows for additional validation to confirm that traffic signs placed based on historical design standards do not meet the needs of current and future driving populations.

Although this thesis presents progress towards the extraction of a complete TSI, the implementation is based on the efficiency of the proposed process as compared to other methods.

Therefore, the next section serves to discuss the processing time needed for the proposed method to extract the TSI.

4.5 PROCESSING TIME

With the introduction of large-scale data processing, a discussion of the time required to implement the process is important before being implemented within the industry. Table 4.9 summarizes the processing time from an average of three runs of the different procedural stages in this thesis.

Table 4.9 – Processing Time for Procedural Stages

Procedure Stage	**Processing Time (seconds/km)**			
	1	**2**	**3**	**Mean**
LiDAR Preprocessing	15.98	16.80	16.83	16.53
Ground/Non-ground and Vehicle Removal	18.21	18.40	18.36	18.32
Traffic Sign Extraction with Orientation	46.56	45.96	46.14	46.22
Lane Marking Extraction with Placement	3.58	3.40	3.44	3.47
Visibility	24.30	18.79	18.78	20.62
Image Classification	780.16	795.00	804.60	793.25
Total	888.79	898.37	908.14	898.43

The unit processing time measurements were calculated considering the 4-km highway test segment, containing approximately 26 million points. If this process is to be scaled across a series of segments for the assessment of the entirety of an Alberta highway section, the proposed method posted a data collection rate of 15 minutes per kilometre. It should be noted that although this does not include the time required to determine the directionality of the signs and determine the legibility distances, this still represents an improvement over manual or field processing techniques.

However, Mobile LiDAR technology is seen as unattainable to certain municipalities or contractors due to its capital cost, operational cost, and the required workflow changes to incorporate the new technology into their everyday decision-making. To shed light on the possible cost savings associated with utilizing a LiDAR scanner, Yen et al. [76] conducted a detailed cost-benefit analysis of the application of mobile LiDAR scanning to highway infrastructure. Several options of contracting, renting, and ownership of mapping and survey-grade LiDAR systems were considered and compared against the local department of transportation TAM inventory annual average cost. Depending on the desired ownership, overall cost savings were between 1.3 – 6.1 million of three data collection cycles over a six-year lifecycle. Therefore, although the initial cost of contracting, renting, or owning a mobile LiDAR scanner is daunting, the benefits of utilizing this technology for TAM are easily realized.

Additionally, Ai and Tsai [24] conducted a productivity assessment to determine the time saved through the application of a LiDAR-based methodology to traffic sign extraction. A manual processing of traffic sign inventory was conducted in 12.8 hours, or 34 minutes per kilometre. In comparison, the LiDAR-based methodology achieved the same results at a rate of 8.1 minutes per kilometre, representing a 76% difference as time saved. However, some additional information was collected in their manual field survey, such as traffic sign panel classification and text, suggesting that the true time-savings may be slightly lower. In comparison, the proposed method is slower than that of Ai and Tsai but also includes additional lane marking placement, classification, and visibility assessments. Additionally, the proposed method still shows a 50% improvement in time spent compared to the manual method. The bulk of the time spent in the proposed method was during the sliding window implementation of the image classifier, accounting for 87% of the total time needed.

5 CONCLUSIONS

Traffic signs are a fundamental component to any roadway environment, providing critical driving information to the surrounding traffic. To ensure traffic signs are placed correctly, industry practitioners rely on inventories of transportation assets to understand information about transportation networks. However, the vast networks of roadways in Canada contain a significant quantity of transportation assets, requiring a thorough collection of transportation assets to quantify the placed infrastructure. Therefore, the collection of a complete traffic sign inventory (TSI) is needed to ensure traffic signs meet the needs of current and future driving populations. However, the size of current global traffic sign networks makes applying traditional survey methods to the collection of a TSI difficult, if not economically infeasible. There is room for technological and methodological improvement towards the creation of an efficient TSI extraction process. This problem has been longstanding, with efforts focused on image-based TSI approaches. Although these approaches provide approximate sign location, no other information about the sign placement, orientation, or condition is present. In an effort to update the TSI extraction process, this research proposes the application of Light Detection and Ranging (LiDAR), coupled with video-log images, to create a complete TSI. The contributions in this thesis represent progress towards an automated TSI but contain limitations in their implementation. The limitations of the TSI methods outlined in this thesis will be discussed in the following section.

5.1 RESEARCH SUMMARY

This research proposes a LiDAR-based framework and discusses the successes and limitations of the application of such a process. The TSI framework is analyzed along a segment of Highway 1 in Alberta, Canada, with the four main components of this thesis: (i) the LiDAR-based extraction of traffic signs, determining their georeferenced location and local orientation, (ii) the LiDAR-based extraction of lane markings for placement measurements for each traffic sign, (iii) the classification of traffic signs from video-log images, utilizing the first traffic sign inventory in Canada to classify signs, and (iv) the measurement of the maximum traffic sign visibility. As mentioned, the LiDAR data utilized in this research is collected by the high-density REIGL VMX-450, which creates detailed point clouds of the scanning environment. To assess its application, the four stages of this thesis are individually assessed for extraction accuracy.

The difficulty with the accuracy assessments is that they are largely manual and that there is no set of reference data for traffic sign location, placement, or classification. The sign location and classification can be assessed using publicly available mapping [58] and the Alberta Traffic Sign Dataset [69]. The placement conditions can still be assessed for compliance to traffic sign placement standards published for Alberta [14]. For all traffic signs along the test segment, the intensity-based extraction of traffic signs had a precision, accuracy, and recall of 98.3%, 92.06%, and 95.08%, respectively. There were five false negatives caused by traffic signs across the divided highway. These signs had lower point density due to their distance from the scanner and were missed through the intensity extraction and clustering processes. Therefore, only signs relevant to the scanning direction (i.e., eastbound) were assessed for accuracy. In the eastbound direction along Highway 1 in Alberta, Canada, the intensity-based extraction of traffic signs had a precision, accuracy, and recall of 98.3%, 100%, and 99.14%, respectively. These signs were then assessed for their orientation compared to the placement standards. Certain signs were found to severely deviate from the horizontal and vertical orientation standards, providing the discussion of the potential effects this has on traffic sign conspicuity. This is especially important at night as too little or too much orientation can result in reduced sign conspicuity, causing the sign to be illegible from poor lighting or from specular reflection from driving vehicle's headlights.

The assessment moved to measure the lateral and vertical placement of the extracted traffic signs. As required in the Highway Guide and Information Sign Manual [14], the placements are measured from the nearest lane markings for highway segments. The extraction of lane markings was assessed by precision, recall, and accuracy for the left- and right-lane markings as 100%, 89.36%, and 94.38%, respectively and 93.47%, 86%, and 89.58%, respectively. The placement measurements combined with the orientation measurements from the traffic sign extraction, serve to complete the discussion of traffic sign conspicuity. Following the previous discussion, the lateral and vertical placement play equally important roles in best utilizing the available retro-reflectivity of placed traffic signs. Signs placed too far outside a driver's field of view reduce their conspicuity, creating the possibility that older drivers or distracted drivers cannot read or miss the traffic signs altogether.

As per the recommendations of NCHRP 748 Section 9.1.5. [22], the assessment of traffic sign classification is assisted by the addition of video-log images along the test segment(s). To this end, a training set of traffic sign classes in Alberta were manually collected to train an image classifier. A total of 13,604 images are collected spanning 155 traffic sign classes and 10 false positive classes. To determine the applicability of a traffic sign classifier, the industry-recognized GoogLeNet structure is used to conduct image classification. The structure is trained within a half hour on a graphical processing unit (GPU) and ~8 hours on the central processing unit (CPU) and produces 83.39% accuracy on the validation set. This translates to a Top-1 and Top-2 classification error of 10.35% and 3.24%, respectively. Two traffic sign classes were not correctly classified within their Top-2 classifications.

When applied to the original video-log images, the sliding window procedure used to apply the trained classifier to cropped image samples creates the opportunity for misclassifications across the input image. This reduces the accuracy of the classifier to 53.3%. However, the results still indicate that the GoogLeNet structure is applicable to the classification of the traffic signs from images and completes the extraction of a traditional TSI. When applied to the video-log images, six new traffic sign classes were discovered. This reinforces the need for the expansion of the training dataset to ensure the image classifier can adequately classify all traffic signs along Canadian roadways.

The detail provided by LiDAR scans is also used to extend the TSI to consider the upstream visibility of extracted traffic signs. This allows for the determination of maximum visibility of traffic signs and the assessment of the efficacy of current traffic sign placements. In daytime conditions, the traffic signs along the segment all met the average perception reaction time (PRT) of 2.5 seconds as set by AASHTO [16]. However, one of the traffic signs had available visibility times below five seconds, creating the possibility for this sign to be only partially read or missed by an older or distracted driving population. The visibility of traffic signs was also assessed during nighttime conditions, where the available visibility is decreased 1.8 times. In this context, two signs had visibility times below 2.5 seconds of PRT and four signs had visibility times below five seconds of PRT. The two signs below 2.5 seconds of PRT are "*Slow Traffic Keep Right*" (RB-35) and "*Watch for Pedestrians on Highway*" (WC-2A). Therefore, if a driver is distracted, or if their perception reaction time falls significantly above the average of 2.5 seconds, drivers could impede traffic in the left lane or be less attentive to hitchhiking pedestrians along the highway's shoulders.

Additionally, the visibility of traffic signs was considered given that, although a traffic sign might be visible if not the most immediate object in view, this sign may not be noticed until further downstream. Traffic signs were assessed for consecutive placement and three signs do not meet the 2.5 second average PRT during the daytime, while eleven signs do not meet the 2.5 second average PRT at night. Therefore, if a driver suffers from higher than average PRT due to aging or is driving distracted or fatigued, a series of signs along the segment may be missed or only partially recognized by a driver. To avoid overloading drivers with a limited amount of time dedicated to each individual sign, agencies need to perform traffic sign visibility assessments to ensure that all essential signs are visible, and sign placement overlaps are limited to scenarios with low complexity and high visibility.

The traffic sign visibility assessment serves to outline signs of concern which may not appropriately serve the driving population. Older drivers, distracted drivers, tourist populations, and nighttime conditions create different requirements of traffic signs and the visibility assessment serves to determine signs that do not meet the needs of current or future driving populations.

This thesis set out to create a complete TSI. As this process is historically manual or partial, the breadth of information available from the combination of LiDAR data and video-log images is utilized to detail the tested roadway. The four-primary tasks of traffic sign extraction, lane marking extraction, image-based traffic sign classification, and traffic sign visibility measurement. With a review of the current state of research literature, this thesis proposes a LiDAR and video-log based method with which to complete a TSI. The contributions of this thesis are outlined in the following section of this thesis.

5.2 RESEARCH CONTRIBUTIONS

The contributions in this research create a framework for the extraction of a TSI based on the understanding that the traffic sign is a key component along any highway network. There have been previous attempts to streamline the TSI extraction process through image-based and manual inventory counts. Although these methods can be accurate, they are also incredibly time consuming and thereby inefficient to implement. Additionally, these methods are limited in that they only allow for the extraction of some of the attributes, and the measurement of traffic sign characteristics (i.e., size, placement, and visibility) are very unsafe if not impossible. Therefore, LiDAR is suggested as a method to build traffic sign inventory, allowing for the extraction, measurement, and assessment of traffic signs along the segment. The density and accuracy of LiDAR data allow for the accurate survey of highway segments at highway speed, resulting in reduced impact on the transportation network and keeping survey staff and the public safe. The following sections will discuss the contributions of this thesis, separated into academic and practical contributions.

The contributions of interest to the academic community include the proposed method for the extraction of a complete TSI, the assessment of traffic sign visibility, and the creation of a Canadian traffic sign image database. The complete TSI provides the addition of orientation and placement to the traffic sign extractions found in the literature. The utilization of intensity-based methods for traffic sign and lane marking extraction further solidify their use in future extraction efforts. Additionally, the raster format used to locate the lane markings also proved successful and was very time efficient, thereby further solidifying its future use as well. The local orientation and the lateral and vertical placement allow for the comparison to design standards for traffic sign placement. If design standards or placement conditions are to be assessed individually, the proposed method allows for the large-scale measurement of these attributes quickly and efficiently.

The visibility assessment is a new addition to the literature, measuring the maximum distance at which traffic signs are visible. This highlights an additional use for the LiDAR data which was previously unexplored and creates the opportunity for continued research into the time available for drivers to read signs. This can be extended to consider different traffic sign placement conditions and their potential effects on the workload of drivers along any given segment.

Finally, this research creates a Canadian image database of traffic signs, which can be utilized for continued research into image-based traffic signs applications. The database was used to successfully classify signs within the video-log images and creates an inventory of varied placement conditions within Alberta. This allows the academic field to continue research into the creation of TSIs, autonomous vehicles research, or as an additional benchmark for image classifiers.

The practical contributions in this thesis of interest for industry applications revolve around the accuracy, time, and cost savings of the proposed method for the creation of a TSI. This thesis highlights that the creation of a TSI in a timely manner is possible utilizing only

LiDAR scans. The extraction of a TSI was completed accurately and provides relevant discussions of the challenges associated with this work. Industry professionals may have their own LiDAR or image databases already created with no means of automatically assessing them. The methods proposed in this research allow practitioners to create additional value from their previously collected data. Additionally, for professionals considering its industry application, the discussions within this thesis assist in guiding practitioners through potential limitations and concerns.

The implementation of a TSI extraction method is largely dependent on the time and cost savings associated with the use of the new method. Overall, the complete assessment of a TSI requires 15 minutes per kilometre of LiDAR roadway. The greatest bottleneck within the proposed procedures is in the sliding window procedure used to create potential traffic sign regions-of-interest within the video-log images. The proposed procedure still provides an accurate extraction of a complete TSI with significant time savings and completeness compared to other manual and image-only based processes. Therefore, the LiDAR and video-log based procedure proposed in this research is applicable to the collection of a complete TSI.

The image database created for the Canadian context can also be utilized by industry professionals to build their own traffic sign inventory applications. The database provides practitioners with information on the placement conditions commonly seen by signs and allows for additional development of industry professional's in-house image processing models. The GoogLeNet model used in this thesis is readily available and could be replicated by professionals for future use.

It is understood that the initial cost of LiDAR can be an inhibiting factor for industry professionals and municipalities considering its application. However, the assessment of different LiDAR ownership strategies by the NCHRP suggest that direct ownership of a LiDAR scanner lead to the highest cost savings for infrastructure assessments like the TSI [22]. Therefore, the contributions carried out in this thesis exist as a candidate for future implementations of a TSI extraction process. It should also be noted that the LiDAR data used within this study is not limited only to TSI-based analyses, including applications to other transportation inventory [34], [72], [77] and safety assessments [55], [71]. The case study within this thesis was applied to rural highway segments, but similar analyses could be conducted in an urban setting or for the assessment of transport signs within the aviation and rail industries [17].

In the North American context, LiDAR scanning is still a juvenile technology for industry applications, but the application of LiDAR complements an effective data management strategy for transportation asset management. As described by Hessle [2]:

> "*Effective Data Management will result in the availability of a stable, high quality information resource across the whole organisation, and in more reliable, better-understood data. This, in turn, will provide hard financial benefits and quantifiable improvements in service delivery, as well as soft benefits, which although less easy to quantify and to cost, are an important part of the value of Data Management. Public organisations are accountable for their decisions, and good Data Management provides the necessary audit path to demonstrate the basis for decision-making. Data that are meaningful and relevant can form the basis of sound decision making. This can lead to cost savings. Good Data Management reduces Data Duplication and Data Redundancy, which again saves money. Well-managed data may also have commercial value - something that many Road Administrations have yet to fully appreciate and capitalise on.*"

Ultimately, the design of roadways is based on standards created in a historical context, and these designs may no longer apply to the current driving population as they once did. Therefore, it is the job of current engineers to assess the efficacy of traffic sign placements to ensure they meet the needs of the driving population. The completion of a traffic sign inventory as part of a transportation asset management strategy provides the data needed to conduct assessments of traffic sign placement standards. As the industry is moving towards more responsible design environments, where collisions on a roadway are attributed to a failure of the design to accommodate the driver rather than a failure of the driver, the contributions of this thesis propose a framework for the assessment of traffic sign installations to meet the needs of current and future driving populations.

5.3 LIMITATIONS

Each section will discuss both the limitations regarding the application of LiDAR or video-logs to the problem of TSI and the limitations of the methods chosen to extract these features from the data. This includes limitations associated with the accuracy of the chosen methods and potential limitations of their reimplementation in other settings.

5.3.1 *Traffic Sign Extraction*

Fundamental to the extraction of traffic signs from the LiDAR data is their high intensity relative to the rest of the data collected within each scan. Gaussian curves were fit to segment the LiDAR data, and the extraction of a correct Gaussian curve is directly impacted by the surroundings collected in a LiDAR scan. This work has been conducted on highway segments, where high intensity noise is limited to the surrounding roadway and vegetation. However, in urban environments with higher traffic volumes and higher infrastructure density, the high intensity Gaussian curve may be skewed to include a larger proportion of noise versus the traffic sign points. Therefore, when extracting the candidate traffic sign points the sample that needs to be clustered and filtered increases. This increases the chance of false positive extractions within the dataset and increases the computational load required when clustering the candidate traffic sign points. However, this could be alleviated by applying an additional Gaussian Mixture Model, additionally segmenting high intensity points to remove only traffic sign panels.

Once the candidate traffic sign points have been extracted from the LiDAR data, a key step to extracting candidate traffic signs is through clustering the candidate points into sets of candidate traffic sign clusters. The DBSCAN attributes for clustering must be manually set and may not translate to a different dataset collected by the same scanner. DBSCAN attributes will likely change even further using a different mobile scanner or a lower scan rate terrestrial scanner (e.g. Velodyne VLP-16 [78]). It is recommended to verify the applicability of the chosen parameters to any new dataset, and if necessary, update these parameters to better suit the available data.

Additionally, the process of principal component analysis (PCA) is limited in that it is indiscriminate to the set of input data. If the result of the clustering process contains noise (e.g., a traffic sign is clustered with another traffic sign or with a metallic post), PCA will determine the principal axes with the additional noise included in the calculation. Therefore, the calculated principal axes will include the noise and there is no indication that this has occurred. It is imperative that the DBSCAN attributes are appropriately tuned to the input dataset to avoid errors in the principal axes and volume calculations.

If the segment is too long or contains too many high-intensity points the proposed procedure is limited by the available computer hardware. As DBSCAN clustering with a distance matrix is an $O(n^2)$ complex problem, a larger dataset demands higher available computational power (i.e., RAM). For the hardware available described in Section 4.3, the maximum allowable pre-clustering data size is approximately 65,000 points.

The vertical angle of the traffic signs is measured based on the normal vector from the road surface, which makes this measurement dependent on the quality of the road surface. If there are any inundations or pot holes along the road surface, the accuracy of this measurement may be affected.

Once the traffic signs have been extracted, the accuracy of each determines the applicability of the proposed methodology to real-world scenarios. However, there are currently limitations in the verification of traffic sign position. No traffic sign database is available to assess the accuracy of the extracted traffic signs. However, visual inspection using Google Maps or Earth applications allow for validation from a trusted third-party source. In the future, these extractions should be further verified with manually surveyed traffic sign locations. However, the traffic sign extractions could be used to create the missing inventory of traffic signs. Any extractions thereafter of either traffic signs or other assets along a segment can then be verified using the same LiDAR scan.

Mentioned throughout this thesis is the focus on the only the traffic signs in the direction of the scanning vehicle. Due to the density of the point cloud closer to the scanner, objects along the segment are best described when immediately adjacent to the scanner. Therefore, the proposed procedures are best suited for applications per scanning direction. To ensure the TSI extraction encompasses all traffic signs in both directions, this process should be repeated for the other travel direction as well to complete the TSI along this segment.

Finally, there is a limitation associated with the application of LiDAR to traffic sign extraction which is present across all field-based inventory processes. If traffic signs are broken or missing from the LiDAR data, they cannot be extracted. This may be beneficial if one has multiple scans of the same location from different times, as this would allow for the detection of missing traffic signs. However, for the extraction of the initial inventory, signs missing from the LiDAR data will result in an incomplete traffic sign inventory.

5.3.2 LANE MARKING EXTRACTION

The proposed procedure is based on the voxel discretization structure to conduct ground and non-ground segmentation, and road surface extraction for the purpose of lane marking extraction. The ground extraction procedure uses the lowest voxels along the segment to separate ground and nonground objects. In the presence of shadows within the point cloud, the ground segmentation may include portions of the non-ground segment. Additionally, the sizing of the voxels may cause under-segmentation of the non-ground objects, resulting in non-ground points being present in the set of ground-segmented points. With increased voxel size, the under-segmentation of ground objects worsens as more points are included per voxel. A decrease in voxel size can also attribute to under-segmentation of the ground objects as the likelihood decreases of a non-ground object's voxels being stacked atop one another.

The effectiveness of the road-surface extraction voxel parameter is dependent on the voxel size. To understand the sensitivity of the road surface extraction, the voxel size was altered while retaining the same road surface extraction parameter. For a larger voxel size, the chosen filtering parameter is less effective along the vertical curved sections of roadway. Due to the higher number of points within each voxel at a larger voxel size, the increased voxel size has a higher standard deviation of elevation for voxels along the vertical curve. Therefore, if the voxel size is to be changed, the heuristic parameter to extract the road surface should be adjusted.

To utilize the intensity difference between the lane markings and the road surface for the purpose of lane marking extraction, the raster image provides a convenient format with which to analyze the lane markings. However, the raster format discretizes points thereby sacrificing some of the "uniqueness" associated with having point-scale measurements of intensity for processing simplicity. Additionally, the raster format and the closing operation cause aliasing, which introduces additional edge noises. The closing operation is required to create a complete raster image, as parts of the cross section contain holes due to the point spacing on the roadway. However, the closing operation creates intensity edges when it determines the values of the missing pixels. To utilize the predictive power of the image processing methods currently available, the individuality of the points was sacrificed to allow for the classification of lane markings.

The primary focus was the lane markings neighboring the shoulders of the test segment as these markings are critical to the assessment of traffic sign placement. The cross sections are chosen based on their proximity to the traffic signs along the segment. The size of each cross section is fixed, and if part is segmented by the cross section then these may not be adequately located. This is of concern at accesses and egresses, where the markings may already be segmented due to changes in the roadway geometry.

The placement of the traffic signs is determined relative to the lane markings. Therefore, if a lane marking is missing the measurement of the traffic sign placement cannot be completed. This was the case for Sign 51, where the missing lane markings resulted in incomplete placement measurements for that sign.

Furthermore, the determination of the exact location of the markings is conducted based on the Canny edge detection procedure, which uses a visually chosen parameter to determine the locations of valuable edges within images. The application of this procedure to different LiDAR segments would require this parameter to be visually re-tuned to maintain the accuracy of the assessment.

Like the traffic signs, the lane markings have a higher probability of being extracted in the travel direction of the scanning vehicle, and it is suggested that on and off ramps are also scanned to improve the detection of both traffic sign and lane marking assets along the ramps. Additionally, the road surface and lane markings may be segmented due to occlusion cause by the scanning vehicle passing or being passed by another vehicle. These shadows cause a lower density of points where the pavement surface was occluded by the other vehicle, reducing the points available to determine the edge of the lane markings. This may result in the Hough transform occasionally missing the linear lane marking edges. Therefore, it is recommended that LiDAR scanning be conducted with as little surrounding traffic as possible to reduce the chance of occlusions along the segment.

Finally, a limitation within the LiDAR data itself is regarding the condition of the lane markings along the scanned road segments. If lane markings are faded or missing, they may not be discernible as lane markings within the LiDAR. The edge detection scheme is less effective if the intensity of the lane markings is too close to the pavement surface, resulting in unclear edges from which to extract the lane markings. The placement measurements are directly dependent on the completeness of the lane marking extraction. In comparison, this lane marking extraction method is less accurate than those presented in previous literature. However, the location in which the LiDAR scanning takes place plays a large role in the success of the lane marking extraction. Alberta, Canada is known for its harsh winters, which take a toll on the quality of the lane markings in the province. Not only does the ice, snow, and hail wear away at the markings, the province of Alberta uses fleets of snow plow trucks to clear highways of snow and ice. This causes further forced degradation of the lane markings [79]. In comparison to Spain [77] and

Coastal China [30], the general conditions experienced by Canadian lane marking infrastructure are much harsher and can result in a higher percentage of incomplete lane markings.

5.3.3 Traffic Sign Classification

The procedure for the classification of traffic signs from the video-log images is built upon the success of the GoogLeNet convolutional neural network. However, a neural network is only ever as good as it's training data, with a high accuracy neural network requiring a breadth of training examples to ensure it learns as many real-life scenarios as possible. In comparison to other neural network training databases [38], [39], the volume of training data used in this research can be improved. However, even given the limited sample of training images available, this sample still demonstrates the power of neural network-based image classification in the Albertan context. Additional collection of traffic sign images should be conducted to collect the remaining traffic sign classification types and to make the samples that describe each traffic sign class as more robust. Given that Alberta experiences drastically varied weather conditions year-round, the dataset could also be improved to include different lighting and weather conditions. LiDAR scanning is only conducted during dry conditions, but images of the roadway network in varied conditions could still be collected by other cameras to be added to the training dataset.

The immediate implementation of a neural network-based image classifier may be limited depending on the available computational hardware. The network was trained on a stand-alone CPU requiring approximately eight hours to complete training. This is greatly dependent on the complexity of the network and the size of the training sample, but if there is to be continued development on the training dataset then the training time will only increase. If an industry professional or municipality is seriously considering training their own neural network-based image classifier, the use of graphical processing units (GPUs) should be under immediate consideration given their improved processing speed and efficiency during training. However, training using only a CPU shows that, with limited resources, accurate neural network classification is still a possibility.

Additionally, the application of this neural network is location dependent. The training images used in this network consider a sample of the contrast, intensity, and placement conditions relevant to the lighting and weather conditions present in Alberta. Although the network model can still be applied to other traffic sign classification scenarios within Canada and the United States, the set of training images will need to be adjusted to add or replace relevant traffic signs used in the new study region. For example, any application of traffic sign classification in the United States will need to replace the maximum speed (Alberta Traffic Sign class RB-1) with maximum mile-per-hour speed signs.

Finally, although the classifier was accurate on the validation dataset, the sliding window procedure severely reduced the accuracy of the classifier when applied to the full-sized video-log images. The sliding window travels across the entire image input image, requiring valuable time during the TSI extraction process. The misclassifications decrease the accuracy of the assessment due to surrounding objects in the sliding window, however this can be improved upon with additional positive and negative training samples. Additionally, the final traffic sign classification is assigned based on the highest classification in the video-log image. Therefore, if multiple signs are present in one image the actual traffic sign may be misclassified for a different traffic sign.

5.3.4 VISIBILITY ASSESSMENT

To ensure the visibility assessment would comprise of only pertinent ground and non-ground objects along the road segment, a vehicle removal process based on the existence of road-surface segmentation was utilized to clean the dataset prior to analysis. However, if the road surface extraction is incomplete, then the vehicles may be improperly or incompletely removed. Additionally, if there are non-ground or noise artifacts within the set of ground points these may result in incomplete vehicle removals.

Fundamentally, the greatest limitation in the visibility assessment is that a manual process is still required to determine the direction traffic signs face (i.e., which direction of travel the traffic signs are serving). Although the assessment could be completed including the traffic signs along the segment, the visibility measurements for signs belonging to the other direction of travel would be inaccurate and waste valuable computational resources to do so. Currently the distinction between traffic signs mounted for the east and westbound directions can only be determined through a manual assessment of extracted traffic signs.

A limitation of the visibility assessment is in the determination of occluding objects, in that any obstructing voxel along the sightline is considered an occlusion without consideration for what kind of object caused the occlusion. This could cause smaller objects to be considered an occlusion, resulting in an inaccurate measurement of visibility. Additionally, manual processing is required to calculate the legibility distance of traffic signs by determining which symbol and letter sizes correspond with each of the traffic signs. This information does not exist otherwise, further necessitating the collection of a traffic sign inventory. The assessment of legibility distance is also based on the values determined by fundamental transportation literature which previously determined the legibility requirements of a driving population. These values may no longer be entirely representative of the driving population as current legibility standards were researched in the 1960's and 1970's [64]. Therefore, the accurate implementation of these processes to measure traffic sign legibility would benefit from an additional study verifying or updating the legibility requirements of today's driving population.

Additionally, the visibility assessment is limited at the start of each LiDAR segment. As shown by Signs 1 and 2 in the visibility assessment, their placement close to the beginning of the segment prematurely limits their visibility to what is available within the LiDAR scan. This could be alleviated by considering multiple segments along the same highway, providing a complete assessment of those signs at the start of the segment.

Finally, an important factor that could have significant impacts on traffic sign visibility, but is not considered in the visibility analysis, is the condition of the traffic sign. The prior assessments were conducted assuming traffic signs that were in good condition. However, even signs placed in compliance with the MUTCD can be obscured due to (i) low contrast between a traffic sign and the background, (ii) too much information on one sign or a high density of signs in a limited area, (iii) dirty signs, (iv) poorly placed or obscured signs, and (v) missing information [17]. Figure 5.1 (a) and (b) are examples of traffic signs posted along other highways in Alberta, which in clear daylight are already difficult to read due to severe wear. This would be exacerbated in nighttime conditions, becoming more difficult for older drivers to locate.

(a) (b)

Figure 5.1 – Traffic Signs in Poor Condition [58]

Although this procedure is currently limited by the manual component required to complete the visibility assessment, the lack of additional TSI information makes this a required process within the initial adoption of this procedure. Therefore, although this currently stands as a limitation, with the collection of initial TSIs, this component would become obsolete.

The proposed method does contain some limitations, as listed within this section. However, through continued research towards the advancement of TSI extraction processes, these limitations can be mitigated or eliminated. The following section serves to outline areas of future research towards the proposed TSI extraction process.

5.4 FUTURE RESEARCH

The discussion of future research topics is introduced with general research which could improve the TSI extraction process. These topics are general extensions or substitutions to the proposed methodology and are discussed separately from future research into individual components of the methodology. Thereafter, the four components of traffic sign extraction, lane marking extraction, traffic sign classification, and sign visibility are individually discussed for future research contributions.

First and foremost, a topic of future research of interest to municipalities is the application of 3D representations of regions based on images – an alternative method to creating point clouds. The comparison of an image-based and LiDAR-based point cloud for the purpose of conducting a TSI would provide an interesting assessment of the trade-offs of cost, efficiency, and accuracy from both methods. Research has been conducted for the location of traffic signs but a comparison of both methods at the same locations would be needed for a thorough comparison of the two methods.

Specific to the extraction pipeline presented in this research, extracting a TSI from LiDAR can benefit from assessing the impact of point density on the TSI. As high-end dual scanners are not immediately feasible acquisitions by all municipalities and industry professionals, an assessment of how a reduced point density affects extracting a TSI would provide practitioners with an understanding of the expected accuracy of the assessment when using different scanners. For example, Figure 5.2 displays a scan collected by a Velodyne VLP-16 [78]. The VLP-16 can collect roughly 300,000 points per second, but with a reduced range of 100 metres and a scan rate of 5-20 Hz, the collection of points at speed results in an increased scan-line spacing (i.e., the spacing between points collected in a single rotation of the scanner, resulting in separated lines). However, details regarding the surrounding environment are still visible within the scan making this application a potentially cost-effective means of collecting a partial TSI.

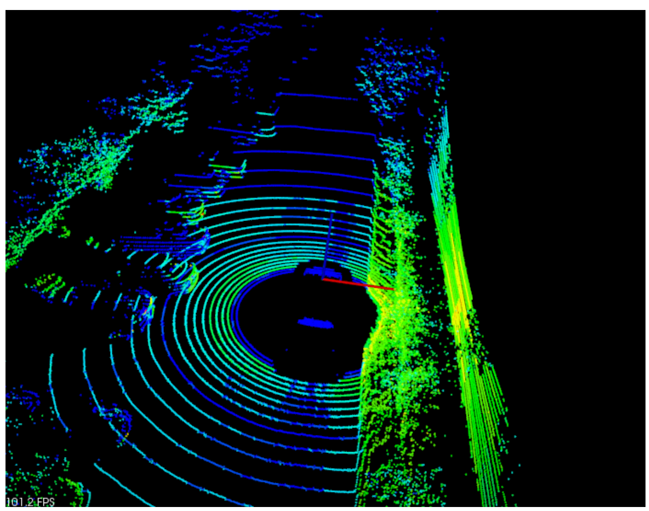


Figure 5.2 – Sample of LiDAR Data - Low Speed Velodyne VLP-16 Sample [80]

Ai and Tsai [24] studied LiDAR-based traffic sign extraction using a scanner with a scan rate of 10,000 points per second, significantly lower than the RIEGL VMX-450. They still report accurate traffic sign extraction, thereby begging the question of whether lower scan rates or lower scan rate scanners could be used to conduct the inventory in a more cost-effective manner.

Finally, the contributions of this research present the entire TSI extraction effort as individual processing pipelines. Future research could consider the classification of all LiDAR data rather than through individual efforts, like the pixel-wise classification research currently taking place for image segmentation. Instead of analyzing the LiDAR for individual features, a point-based classifier could look at providing a classification value to each LiDAR point. Work in this field would apply not only to conducting a TSI and would extend to a full-scale infrastructure inventory.

The extraction pipelines for each step in the TSI outlined in this thesis were discussed based on their limitations, showing room for improvement either methodologically or ideologically. The individual extraction pipelines are further discussed in this section to recognize areas of future research.

5.4.1 Traffic Sign Extraction

Although some of the procedures within this research have been adapted from previous work, the additional contributions of this research are as follows: complete inventory of traffic signs, including the location of the sign panels, their placement, and their visibility analysis within the Albertan context. This allows for a two-part assessment of any highway network: (i) the collection of traffic sign inventory to update a transportation infrastructure inventory, and (ii) the assessment of traffic sign efficacy.

The collection of traffic sign points is fundamentally based on the DBSCAN clustering process. The DBSCAN parameters used in this research are based on the analysis from a previous study using the same dataset. However, there is still the possibility of random variation within datasets, which may reduce the completeness of the potential traffic sign points with the current DBSCAN attributes. Therefore, a sensitivity assessment of both DBSCAN parameters would further tune the clustering algorithm to the input dataset, ensuring the potential traffic sign clusters are as complete as possible. This is likely only required under certain scanning conditions (i.e., using a specific scanner or collecting scans at a specific travel speed) and once verified can be applied to other scans collected under similar conditions.

To improve the false positive filtering process outlined in this research and to expand the details extracted in a TSI, this research could be expanded to include the mounting conditions for each traffic sign. This would provide practitioners with additional information on the placement conditions of traffic signs and allow for a better understanding of how the sign is mounted relative to rest of the objects in the LiDAR scan. Additionally, the application of breakaway posts and overall post size or radius contribute to traffic safety and their extraction may be beneficial in the assessment of the potential likelihood and severity associated with colliding with different traffic signs. However, LiDAR scanning passes may need to be conducted at lower speed to ensure a higher point density is collected along the traffic sign posts.

Finally, although traffic sign retro-reflectivity cannot be determined from LiDAR scans as per NCHRP 748 Section 9.1.1 [22], the relative intensity can be compared within single LiDAR scans. Therefore, future research could explore the relative differences between traffic signs within a scan. If a traffic sign exhibits lower intensity than the other traffic signs within a LiDAR segment it can be flagged as potentially dirty or aged.

5.4.2 LANE MARKING EXTRACTION

The lane marking extraction is conducted using a voxel-based approach, determining ground and road-surface points prior to extracting cross sections at which to determine the position of the lane markings. Therefore, improvements to this process should start with the underlying ground and road-surface segmentations. The ground extraction procedure could be expanded to consider additional or alternative measures of voxel attributes as they belong to the ground surface. By considering the regions around the currently extracted ground points, the process could consider the properties of points relative to other ground points for verification purposes. This serves to improve the ground extraction to reduce the under-segmented non-ground points that exist in the ground point set. This would reduce the possibility of noise being present in the cross sections and improve the vehicle removal process used in the visibility analysis. Similarly, the improvement of the road-surface extraction could consider local regions of change rather than considering the global standard deviation of elevation parameters. This guarantees the completeness of the cross sections used in the lane marking extraction.

When detecting the lane markings, the proposed procedure depends on the development of raster images and the detection of the lane marking intensity edges within those images. Further research could consider alternative pixel processing algorithms to maintain and enhance the intensity on the lane markings whilst reducing the noise specific to the road surface. Additionally, the Canny edge detector requires defining the desired edge strength, which is currently determined visually. Future research could consider the quantitative sensitivity of this parameter, perhaps through the application of reference values of the correct edges within the raster images.

As previously mentioned, the point-based classification would be a strong substitute for future research into the extraction of lane markings. Lane markings exist in varied conditions, sizes, and shapes, and training a point-based classifier to determine local and global differences between the lane markings and other points along the segment would improve the completeness of the lane markings.

5.4.3 TRAFFIC SIGN CLASSIFICATION

The process of image-based traffic sign classification hinges on the quality and completeness of the training dataset and the capabilities of the image classifier. Fundamentally, the accuracy of the image classifier can be improved by one of two changes: increase the amount of training data available or improve upon the neural network model.

Improvements to the neural network classification can come from one of two areas: change the overall structure of the model or fine-tune the individual components of the model. Although the GoogLeNet model performed well, filter sizes for the convolutional and pooling layers or the weight and bias parameter initializations could be altered to improve the performance of the model. Additional changes could also be made to the network structure, with the simplest change being the additional model depth. Alternatively, different model structures could also be considered, including residual networks [81] or multi-column networks (where multiple neural network structures under different training conditions are combined to conduct classification [40]).

The second change to the image classifier is in the increase in the number of training samples to increase provincial and federal coverage to provide additional samples of roadways in the Canadian driving context. The training dataset is based on a set of the available video-log images in Alberta, and with continued LiDAR scanning in different regions, under different conditions, and at different times, the training dataset could be updated.

The primary point of concern for the expansion of this dataset would be to ensure each class has roughly the same number of training image and that they contain reasonable environmental variety. There should be enough samples within each class to reduce the variance between class sample sizes and there should be additional focus on more false positive classes. Tree-lines and skylines were commonly mis-classified and these need to be added to the training data. If more video-log images are not available, traffic sign samples could be partially inflated with image augmentation. This will increase the variation in the input dataset and provide additional training samples, but it is still dependent on the original input images.

Traffic signs mounted in Alberta are still occasionally proprietary and the guaranteed application of this dataset to traffic signs across Canada requires samples from each province. This ensures the database contains samples from these provinces, both to include the different proprietary traffic signs and to ensure the surrounding environmental characteristics are trained into any new classifier.

To avoid the issues caused by the sliding window detection process, a traffic sign detection procedure could be included in this processing pipeline in the future. By training a traffic sign detector with the bounding boxes of traffic signs in an image, the detector can feed the classifier only with traffic sign images to better utilize the accuracy possible from the classifier.

5.4.4 VISIBILITY ASSESSMENT

The environmental detail provided in a LiDAR survey supplements the traditional TSI with an additional traffic sign efficacy metric – the visibility measurement of each traffic sign. The first change to the visibility assessment would be to expand the line-based visibility assessment to consider the driver's viewshed for any given trajectory point. This could be improved upon in further research by applying a viewshed to the traffic sign panel to determine what percentage of the traffic sign is occluded.

The assessment of visibility could be extended to consider the different kinds of occluding objects along the highway segments. The occlusion by different objects and roadway geometry may signal issues with the highway design or use of a transportation asset, creating the need for additional assessment. This assessment could also be extended to allow for the assessment of general visibility for any asset.

The determination of traffic sign directionality is currently manual. The TSI could be expanded to include the mounting conditions as mentioned in the Future Research for the traffic sign extraction. If the mounting conditions of each traffic sign exists, the post will always be facing away from the direction of travel, thereby definition the direction of the traffic sign.

Already noted as a limitation, the placement of traffic signs towards the start of the segment prematurely limits the measurement of visibility of those signs. Future research could consider the combination of multiple segments to ensure the visibility for traffic signs at the beginning of the segment are correct. However, this research should consider the computational expense of combining multiple LiDAR segments, as each segment contains 30 million points on average.

Although this thesis provides a discussion on the consequences of occlusions on traffic sign visibility, no suggestion is made on how the traffic sign placement should change in order to circumvent the occlusion. Future research should investigate recommendations to change highway designs or to change traffic sign placement when occlusions are detected.

Finally, the assessment of traffic sign visibility can be applied to a discussion of changes to the traffic sign placement conditions for an aging driving population. Within the three-dimensional context of the point clouds, traffic signs placements could be altered, and signs could be resized and reassessed for visibility. This would better inform the use and design of more effective traffic control devices [17]. For example, Dissanayke and Lu [82] increased the size of stop signs and noted an increase in sign legibility for older drivers, allowing them to begin their deceleration at further, safer distances. Furthermore, Campbell et al. [83] recommended the standardizations of traffic sign symbols to improve sign legibility. These include (i) minimizing symbol complexity; (ii) maximizing distance between sign elements; (iii) using representational symbols; (iv) using solid figures; (v) standardizing the design of arrowheads, human figures, and vehicles; (vi) providing maximum contrast for symbols on signs; and (vii) using large font as often as possible [83]. Although some of these recommendations already exist in signs placed by MUTCD standards [4], making compensatory changes such as maintaining traffic sign contrast and increasing letter/symbol size comes at an increased cost per traffic sign. However, Castro and Horberry [17] note a lack of research into the cost-benefit relationship of these changes, how they might be utilized by other drivers, and how this might better inform future traffic sign installations. This should be included in the assessment of sign panel changes as well.

REFERENCES

[1] J. Elliott *et al.*, "Implementation Guide for an ISO 55001 Asset Management System," 2016.

[2] M. Hessle, "Data Management for Road Administrations: A Best Practice Guide," 2003.

[3] Government of Canada, "The Canadian Transportation System," 2018. [Online]. Available: https://www144.statcan.gc.ca/tdih-cdit/cts-rtc-eng.htm. [Accessed: 05-Nov-2018].

[4] Federal Highway Administration, "Manual on Uniform Traffic Control Devices for Streets and Highways," pp. 347–432, 2009.

[5] Department for Transport, "Estimating the number of traffic signs in England," 2013.

[6] Statistics Canada, "The changing landscape of Canadian metropolitan areas," 2016. [Online]. Available: https://www150.statcan.gc.ca/n1/pub/16-201-x/16-201-x2016000-eng.htm. [Accessed: 31-Dec-2018].

[7] United Nations, "World Population Ageing," 2015.

[8] E. Hauer, "An exemplum and its road safety morals," *Accid. Anal. Prev.*, vol. 94, no. January, pp. 168–179, Sep. 2016.

[9] AASHTO, *Roadside Design Guide*. 2011.

[10] E. Green, "Development of a Traffic Sign Inventory System (Sign Click)," 2006.

[11] GIS Mid-West Inc., "Street Sign Inventory and Data Collection," 2018. [Online]. Available: https://www.mid-westgis.com/services/sign-inventory/. [Accessed: 26-Sep-2018].

[12] D. Barg *et al.*, "A Web Application Approach to Street Sign Inventory Development," 2011.

[13] V. Balali, A. Ashouri Rad, and M. Golparvar-Fard, "Detection, classification, and mapping of U.S. traffic signs using google street view images for roadway inventory management," *Vis. Eng.*, vol. 3, no. 1, pp. 1–18, 2015.

[14] Alberta Infrastructure and Transportation, "Highway Guide and Information Sign Manual," 2006.

[15] R. W. McNees and H. D. Jones, "Legibility of Freeway Guide Signs as Determined by Sign Materials," *Transp. Res. Rec.*, vol. 1142, 1987.

[16] AASHTO, "Policy on geometric design of highways and streets," Washington, DC, 2011.

[17] C. Castro and T. Horberry, *Human Factors of Transport Signs*. Taylor & Francis Group, LLC, 2004.

[18] H. T. Zwahlen and T. Schnell, "Legibility of Traffic Sign Text and Symbols," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1692, no. 1, pp. 142–151, 1999.

[19] C. Ai and Y. J. Tsai, "An automated sign retroreflectivity condition evaluation methodology using mobile LIDAR and computer vision," *Transp. Res. Part C Emerg. Technol.*, vol. 63, pp. 96–113, 2016.

[20] S. Gargoum, K. El-Basyouny, J. Sabbagh, and K. Froese, "Automated Highway Sign Extraction Using Lidar Data," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2643, pp. 1–8, 2017.

[21] M. Soilán, B. Riveiro, J. Martínez-Sánchez, and P. Arias, "Traffic sign detection in MLS acquired point clouds for geometric and image-based semantic inventory," *ISPRS J. Photogramm. Remote Sens.*, vol. 114, pp. 92–101, 2016.

[22] M. J. Olsen *et al.*, *Guidelines for the Use of Mobile LIDAR in Transportation Applications*

(NCHRP Report 748). 2013.
[23] United Nations Economic Commission for Europe, "Convention on Road Signs and Signals," 2006.
[24] C. Ai and Y. J. Tsai, "Critical Assessment of an Enhanced Traffic Sign Detection Method Using Mobile LiDAR and INS Technologies," vol. 141, no. 2011, pp. 1–12, 2011.
[25] B. Riveiro, L. Diaz-Vilarino, B. Conde-Carnero, M. Soilan, and P. Arias, "Automatic Segmentation and Shape-Based Classification of Retro-Reflective Traffic Signs from Mobile LiDAR Data," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 9, no. 1, pp. 295–303, 2016.
[26] Z. Deng and L. Zhou, "Detection and Recognition of Traffic Planar Objects Using Colorized Laser Scan and Perspective Distortion Rectification," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 5, pp. 1485–1495, 2018.
[27] Á. Arcos-García, M. Soilán, J. A. Álvarez-García, and B. Riveiro, "Exploiting synergies of mobile mapping sensors and deep learning for traffic sign recognition systems," *Expert Syst. Appl.*, vol. 89, pp. 286–295, 2017.
[28] C. Wen *et al.*, "Spatial-Related Traffic Sign Inspection for Inventory Purposes Using Mobile Laser Scanning Data," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 1, pp. 27–37, 2016.
[29] Y. Yu, J. Li, C. Wen, H. Guan, H. Luo, and C. Wang, "Bag-of-visual-phrases and hierarchical deep models for traffic sign detection and recognition in mobile laser scanning data," *ISPRS J. Photogramm. Remote Sens.*, vol. 113, pp. 106–123, 2016.
[30] H. Guan, J. Li, Y. Yu, C. Wang, M. Chapman, and B. Yang, "Using mobile laser scanning data for automated extraction of road markings," *ISPRS J. Photogramm. Remote Sens.*, vol. 87, pp. 93–107, 2014.
[31] B. Riveiro, H. González-Jorge, J. Martínez-Sánchez, L. Díaz-Vilariño, and P. Arias, "Automatic detection of zebra crossings from mobile LiDAR data," *Opt. Laser Technol.*, vol. 70, pp. 63–70, 2015.
[32] P. Kumar, C. P. McElhinney, P. Lewis, and T. McCarthy, "Automated road markings extraction from mobile laser scanning data," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 32, pp. 125–137, 2014.
[33] H. Guan, J. Li, Y. Yu, Z. Ji, and C. Wang, "Using Mobile LiDAR Data for Rapidly Updating Road Markings," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2457–2466, 2015.
[34] L. Yan, H. Liu, J. Tan, Z. Li, H. Xie, and C. Chen, "Scan line based road marking extraction from mobile LiDAR point clouds," *Sensors (Switzerland)*, vol. 16, no. 6, pp. 1–21, 2016.
[35] Y. Yu, J. Li, H. Guan, F. Jia, and C. Wang, "Learning hierarchical features for automated extraction of road markings from 3-D mobile LiDAR point clouds," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 8, no. 2, pp. 709–726, 2015.
[36] S. Papert, "The Summer Vision Project," 1966.
[37] A. Møgelmose, M. Trivedi, and T. Moeslund, "Vision based Traffic Sign Detection and Analysis for Intelligent Driver Assistance Systems: Perspectives and Survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, pp. 1484–1497, 2012.
[38] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition," *Neural Networks*, vol. 32, pp. 323–332, Aug. 2012.

[39] R. Timofte, K. Zimmermann, and L. Van Gool, “Multi-view traffic sign detection, recognition, and 3D localisation,” *Mach. Vis. Appl.*, vol. 25, no. 3, pp. 633–647, 2014.
[40] D. Cireșan, U. Meier, J. Masci, and J. Schmidhuber, “Multi-column deep neural network for traffic sign classification,” *Neural Networks*, vol. 32, pp. 333–338, 2012.
[41] S. Dahlstedt and O. Svenson, “Detection and reading distances of retroreflective road signs during night driving,” *Appl. Ergon.*, vol. 8, no. 1, pp. 7–14, 1977.
[42] C. Baek, “Determining Visibility Distance of Signs Installed on the Roadside Using Videologs,” North Carolina State University, 2003.
[43] P. Huang, M. Cheng, Y. Chen, H. Luo, C. Wang, and J. Li, “Traffic Sign Occlusion Detection Using Mobile Laser Scanning Point Clouds,” *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 9, pp. 2364–2376, 2017.
[44] S. Wu, C. Wen, H. Luo, Y. Chen, C. Wang, and J. Li, “Using mobile LiDAR point clouds for traffic sign detection and sign visibility estimation,” *Int. Geosci. Remote Sens. Symp.*, vol. 2015–Novem, pp. 565–568, 2015.
[45] A. Al-Kaisy, J. Bhatt, and H. Rakha, “Assessing the occlusion effect of large-dimension vehicles on traffic signs at multilane highways,” *J. Transp. Eng.*, vol. 130, no. 5, pp. 648–657, 2004.
[46] P. Discetti and R. Lamberti, “Traffic Sign Sight Distance for Low-Volume Roads,” *Transp. Res. Rec. J. Transp. Res. Board*, no. 2203, pp. 64–70, 2011.
[47] K. Nassar and A. Al-Kaisy, “Assessing sign occlusion in buildings using discrete event simulation,” *Autom. Constr.*, vol. 17, no. 7, pp. 799–808, 2008.
[48] RIEGL Laser Measurement Systems, “RIEGL VMX-450 Datasheet,” 2014.
[49] C. Briesea *et al.*, “Analysis of Mobile Laser Scanning Data and Multi-View Image Reconstruction,” *XXII ISPRS Congr. Tech. Comm. V*, vol. 39-B5, no. July, pp. 163–168, 2012.
[50] American Society for Photogrammetry & Remote Sensing, “LAS Specification, Version 1.2,” pp. 1–28, 2008.
[51] The Mathworks, “MATLAB.” Natick, Massachusetts, United States, 2018.
[52] Applied Imagery, “Quick Terrain Reader.” Bethesda, Maryland, United States, 2018.
[53] T. Hastie and R. Tibshirani, “Gaussian Mixture Models,” 2008. [Online]. Available: http://statweb.stanford.edu/~tibs/stat315a/LECTURES/em.pdf. [Accessed: 11-Dec-2018].
[54] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise,” in *KDD’96 Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 1996, pp. 226–231.
[55] M. Soilán, B. Riveiro, A. Sánchez-Rodríguez, and P. Arias, “Safety assessment on pedestrian crossing environments using MLS data,” *Accid. Anal. Prev.*, vol. 111, pp. 328–337, 2018.
[56] L. Smith, “A Tutorial on Principal Component Analysis,” 2002. [Online]. Available: http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf. [Accessed: 12-Dec-2018].
[57] A. Gressin, C. Mallet, J. Demantké, and N. David, “Towards 3D lidar point cloud registration improvement using optimal neighborhood knowledge,” 2014.
[58] “Google Earth,” 2018. [Online]. Available: https://www.google.com/earth/. [Accessed: 30-Jul-2018].
[59] R. Fisher, S. Perkins, A. Walker, and E. Wolfart, “Closing,” 2003. [Online]. Available:

https://homepages.inf.ed.ac.uk/rbf/HIPR2/close.htm.
[60] S. C. Huang, F. C. Cheng, and Y. S. Chiu, "Efficient contrast enhancement using adaptive gamma correction with weighting distribution," *IEEE Trans. Image Process.*, vol. 22, no. 3, pp. 1032–1041, 2013.
[61] "Efficient Contrast Enhancement Using Adaptive Gamma Correction With Weighting Distribution," 2018. [Online]. Available: https://github.com/mss3331/AGCWD. [Accessed: 19-Nov-2018].
[62] S. Wang, K. Ma, H. Yeganeh, Z. Wang, and W. Lin, "A Patch-Structure Representation Method for Quality Assessment of Contrast Changed Images," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2387–2390, 2015.
[63] D. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*. 2003.
[64] P. M. Garvey and B. Kuhn, "Highway Sign Visibility," *Automob. Transp. - Traffic, Streets Highw.*, no. Li, pp. 1–17, 2011.
[65] M. Matsugu, K. Mori, Y. Mitari, and Y. Kaneda, "Subject independent facial expression recognition with robust face detection using a convolutional neural network," *Neural Networks*, vol. 16, no. 5–6, pp. 555–559, 2003.
[66] S. Saxena, "Artifical Neuron Networks (Basics) - Introduction to Neural Networks." [Online]. Available: https://becominghuman.ai/artificial-neuron-networks-basics-introduction-to-neural-networks-3082f1dcca8c. [Accessed: 04-Dec-2018].
[67] C. Szegedy *et al.*, "Going deeper with convolutions," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9.
[68] L. Yann, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
[69] S. Lai, "Alberta Sign Catalogue and Images," 2018. [Online]. Available: https://www.transportation.alberta.ca/1842.htm. [Accessed: 30-Jul-2018].
[70] A. Transportation, "Highway Geometric Design Guide," 1999.
[71] S. A. Gargoum, M. H. Tawfeek, K. El-Basyouny, and J. C. Koch, "Available sight distance on existing highways: Meeting stopping sight distance requirements of an aging population," *Accid. Anal. Prev.*, vol. 112, no. December 2017, pp. 56–68, 2018.
[72] S. Gargoum, K. El-Basyouny, and J. Sabbagh, "Assessing Stopping and Passing Sight Distance on Highways Using Mobile LiDAR Data," *J. Comput. Civ. Eng.*, vol. 32, 2018.
[73] G. F. King and H. Lunenfeld, "Development of Information Requirements and Transmission Techniques for Highway Users," 1971.
[74] H. T. Zwahlen, "Conspicuity of Suprathreshold Reflective Targets in a Driver's Peripheral Visual Field At Night," *Transp. Res. Rec.*, no. 1213, pp. 35–46, 1989.
[75] D. J. Mace, "Sign legibility and conspicuity," in *Transportation in an Aging Society: Improving Safety and Mobility for Older Persons*, 1988, vol. 2, pp. 270–293.
[76] K. S. Yen, T. A. Lasky, and B. Ravani, "Cost-Benefit Analysis of Mobile Terrestrial Laser Scanning Applications for Highway Infrastructure," *J. Infrastruct. Syst.*, vol. 20, no. 4, p. 04014022, 2014.
[77] A. Holgado-Barco, B. Riveiro, D. González-Aguilera, and P. Arias, "Automatic Inventory of Road Cross-Sections from Mobile Laser Scanning System," *Comput. Civ. Infrastruct. Eng.*, vol. 32, no. 1, pp. 3–17, 2017.
[78] Velodyne, "VLP-16 Data Sheet," 2017.
[79] M. Lali, "Highway Pavement Marking Guide," 2003.
[80] Point Cloud Library, "Documentation." [Online]. Available:

http://pointclouds.org/documentation/tutorials/hdl_grabber.php. [Accessed: 10-Oct-2018].
[81] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.
[82] S. Dissanayke and J. J. Lu, "Effect of larger stop signs on older drivers," in *80th Annual Meeting, Transportation Research Board*, 2001.
[83] J. L. (Batelle) Campbell *et al.*, *Human Factors Guidelines for Road Systems*. 2012.

APPENDIX

The Segment Traffic Sign Measures serves to detail the entire segment's traffic sign measurements, including the lateral placement, vertical placement, horizontal orientation, vertical orientation, and direction. Direction includes the cardinal directions and, if applicable, the designation "R" for signs related to on or off ramps. This is to detail the traffic signs which existed along the segment. It should be noted that the placement and orientation measures are taken from the Eastbound lanes and may not be correct for Westbound signs.

The Training Dataset Image Classes outlines the training dataset image classes, including the full number of images, also split into the number of training and testing images. These are utilized in the GoogLeNet convolutional neural network for image classification.

SEGMENT TRAFFIC SIGN MEASURES

Sign ID	Lateral Placement	Vertical Placement	Horizontal Orientation	Vertical Orientation	Direction
1	31.46	0.49	11.38	-0.88	W
2	5.40	1.56	-5.36	-3.84	E
3	5.16	1.37	-8.53	-1.87	E
4	15.62	2.06	-3.64	-3.55	W
5	3.87	1.51	-0.64	-12.07	E
6	38.17	3.66	10.72	-1.61	W
7	45.00	3.53	4.11	-0.58	W
8	5.72	1.70	-8.51	-0.47	E
9	2.50	2.04	0.23	-2.06	E
10	13.55	2.08	-7.94	-0.61	E
11	6.79	1.56	-5.77	-3.75	E
12	10.14	1.53	-5.98	-0.41	E
13	2.01	1.44	-8.93	-2.08	E
14	1.51	1.62	-9.01	-8.41	E
15	9.66	1.48	-6.47	-3.56	E
16	6.98	1.80	-12.21	-5.75	E
17	12.69	1.97	4.58	-0.41	W
18	15.19	2.04	-2.09	-1.38	W
19	36.14	1.79	-4.23	-4.47	W
20	5.71	2.16	-0.77	-1.20	E
21	12.04	2.23	-0.81	-3.66	W
22	31.26	2.11	19.37	-1.84	W
23	4.10	1.68	-3.15	-0.48	E
24	0.86	2.02	1.48	-0.06	E

Sign ID	Lateral Placement	Vertical Placement	Horizontal Orientation	Vertical Orientation	Direction
25	15.66	2.02	5.74	-1.81	W
26	34.38	2.67	4.79	-0.82	W
27	6.08	1.80	-2.72	-0.88	E
28	14.16	1.79	-8.64	-1.43	W
29	5.17	2.09	-8.52	-1.53	E
30	3.16	2.05	-8.60	-0.35	E
31	28.55	1.97	18.70	-0.63	E
32	4.47	0.73	-56.66	-1.30	W
33	34.48	1.65	0.85	-0.31	E
34	6.06	2.78	0.69	-0.01	E
35	4.70	1.74	-8.61	-2.43	E
36	24.95	2.87	-22.99	-11.53	F
37	19.66	1.26	-11.71	-1.50	E
38	2.18	1.36	4.07	-4.15	E
39	53.40	2.46	-2.57	-2.10	R
40	39.26	2.68	-19.48	-2.37	R
41	35.03	2.40	8.96	-1.52	W
42	81.71	7.43	-7.95	-2.66	R
43	92.56	6.66	-29.32	-10.14	R
44	23.55	8.71	37.56	-0.39	R
45	0.57	6.21	-19.36	-0.51	E
46	20.88	6.51	0.91	-0.41	W
47	5.22	1.71	1.11	-3.00	E
48	123.47	8.03	27.85	-5.56	R
49	2.56	2.11	3.47	-1.83	E
50	71.01	3.61	15.96	-2.32	R
51	-999.00	-999.00	1.42	-2.52	E
52	14.87	1.85	-15.17	-11.94	R
53	-999.00	-999.00	3.88	-4.28	W
54	-999.00	-999.00	7.90	-2.86	R
55	40.17	2.00	9.33	-3.57	W
56	30.08	3.64	0.71	-2.57	W
57	4.70	1.57	5.02	-3.16	E
58	4.68	1.22	-0.40	-5.81	E
59	5.01	1.27	2.20	-4.75	E

Training Dataset Image Classes

Label	Full	Training	Validate	Label	Full	Training	Validate
IA-201 (single-item)	261	183	78	WA-10A (ramp-advisory-speed)	214	150	64
IA-202 (single-distance)	143	100	43	WA-10B (ramp-ahead-advisory-speed)	240	168	72
IA-203 (single-distance-and-arrow)	57	40	17	WA-112-R (free-flow-right)	68	48	20
IA-204 (double-direction)	72	50	22	WA-113 (high-loads-exit)	104	73	31
IA-205 (double-distance)	55	39	16	WA-14 (t-intersection)	23	16	7
IA-207 (triple-direction)	62	43	19	WA-16-R (merge-right-arrow)	204	143	61
IA-208 (triple distance)	45	31	14	WA-16X-O (merge)	88	62	26
IA-210 (quadruple-direction)	63	44	19	WA-17 (barrier)	127	89	38
IB-119 (cowboy trail)	10	7	3	WA-21 (downgrade)	12	8	4
IB-121 (caring-for-albertas-highways)	32	22	10	WA-26 (bridge)	185	130	55
IB-1B (transcanada-highway-16-marker)	63	44	19	WA-26B	20	14	6
IB-2 (highway-marker)	77	54	23	WA-27 (bridge clearance)	54	38	16
IB-2 (highway-direction)	182	127	55	WA-3-L (curve-left)	235	165	70
IC-1 (gas)	30	21	9	WA-3-R (curve-right)	290	203	87
IC-10 (hospital)	80	56	24	WA-33-L (lane-ends-left-single)	29	20	9
IC-14 (disability-access)	46	32	14	WA-33-R (lane-ends-right-single)	13	9	4
IC-2 (food)	36	25	11	WA-33X-R (lane-ends-right-double)	102	71	31
IC-20-T (museum-tab)	13	9	4	WA-36 (chevron-barrier)	209	146	63
IC-213A	204	143	61	WA-36-L (object-on-left)	13	9	4
IC-241 (rest-area-next-exit)	147	103	44	WA-5-R (right-reverse-curve)	24	17	7
IC-251 (roadside-turnout-1-km)	14	10	4	WA-7 (advisory-speed)	41	29	12
IC-255 (historical-point-of-interest)	36	25	11	WA-8A (bidirectional-checkers)	20	14	6
IC-5 (picnic-table)	120	84	36	WA-9 (chevron-curve)	33	23	10
IC-57 (trailer-and-tent)	99	69	30	WB-1 (stop-ahead)	40	28	12
IC-64 (day-hiking)	27	19	8	WB-4 (signal-ahead)	17	12	5
IC-67 (fishing)	34	24	10	WB-5A (signal-ahead-flashing)	152	106	46
IC-7 (lodging)	52	36	16	WC-10 (snowmobile-crossing)	18	13	5
IC-73 (garbage-disposal)	102	71	31	WC-109 (logging-trucks-turning)	29	20	9
IC-74 (washrooms)	105	74	31	WC-112 (wind-gusts)	22	15	7
IC-76 (equestrian)	45	31	14	WC-13 (deer)	50	35	15
IC-82 (shooting-range)	24	17	7	WC-23 (slippery-when-cold)	53	37	16

IC-85-T (ecostation)	19	13	6	WC-314-L (logs-may-swing-into-lane-left)	14	10	4
IC-9 (viewpoint)	45	31	14	WC-8-L (heavy-vehicle-traffic)	47	33	14
IC-D-T (distance-km-or-m)	13	9	4	WC-9-O (school-bus-stop-ahead)	14	10	4
ID-33 (photo-radar)	346	242	104	WD-101 (construction-with-arrow)	108	76	32
ID-33A (red-light)	25	18	7	WD-102 (begin-detour)	25	18	7
ID-33B-T (speed)	11	8	3	WD-104 (barricade-ahead)	35	25	10
ID-502 (speed-fine-doubles)	33	23	10	WD-111 (be-prepared-to-stop)	28	20	8
ID-503 (speed-fines-double)	13	9	4	WD-154 (end-construction)	30	21	9
ID-504 (speed-fine-doubles)	23	16	7	WD-192 (construction-next-x-km)	15	11	4
IF-201 (next-exits)	285	200	85	WD-A-22 (bump-ahead-symbol)	30	21	9
IF-202 (advance guide sign)	182	127	55	WD-A-23-R (road-narrows-right)	12	8	4
IF-203 (next-exit)	157	110	47	WD-A-33-L (road-narrows-lane-ends)	14	10	4
IF-204 (exit-single-arrow)	313	219	94	WD-A-41 (shovel-worker)	37	26	11
IF-204A (exit-direction-guide)	86	60	26	WD-A-48-R (truck-entrance-right)	47	33	14
IF-205 (exit without number)	39	27	12	WD-A-51-L (diversion-two-lanes-left)	16	11	5
IF-205A (exit with number)	321	225	96	access-distance	73	51	22
IF-207 (overhead-lane-one-arrow)	120	84	36	advance-exit-lanes	59	41	18
IF-207A (overhead-lane-two-arrow)	282	197	85	advertising	164	115	49
IF-207B (overhead guide exit only)	102	71	31	building-a-better-alberta	21	15	6
IF-208 (diagrammatic)	45	31	14	city-of-calgary-welcome	32	22	10
LED	27	19	8	dangerous-goods-info	108	76	32
RA-1 (reg-stop-sign)	84	59	25	exit-backwards-arrow	14	10	4
RA-102 (ped-crossing-flashing)	75	53	22	exit-bypass	16	11	5
RA-2 (yield)	12	8	4	exit-junction-distance	130	91	39
RA-4-L (ped-crossing-symbol)	35	25	10	guide-lane-ends-arrow	20	14	6
RA-6 (railway)	13	9	4	guide-lane-ends-diagram	126	88	38
RB-1-100	438	307	131	high collision location	37	26	11
RB-1-110	278	195	83	highway marker obsolete	81	57	24
RB-1-50	13	9	4	highway-marker	67	47	20
RB-1-60	55	39	16	impaired-driver	60	42	18
RB-1-70	248	174	74	info-city-centre	11	8	3
RB-1-80	196	137	59	junction-direction	88	62	26
RB-1-90	27	19	8	lane-ends-distance	20	14	6
RB-11-L (no-left-turn)	184	129	55	lane-marking	440	308	132

RB-11-R (info-no-right-turn)	10	7	3	max-speed-fines-double	36	25	11
RB-16 (no-u-turn)	12	8	4	no-slow-moving-vehicles	65	46	19
RB-209 (engine-retarder-brakes)	44	31	13	no-slow-vehicles-with-exit	11	8	3
RB-21-L (one-way-left-arrow)	29	20	9	pole	253	177	76
RB-23 (no-entry)	23	16	7	ramp-speed-truck-tipping	14	10	4
RB-24 (two-way-traffic)	25	18	7	right-lane-must-exit	19	13	6
RB-25 (keep-right)	102	71	31	sign-post	93	65	28
RB-31 (do-not-pass)	12	8	4	sky	12	8	4
RB-35 (slow-traffic-keep-right)	73	51	22	slow-down-now	11	8	3
RB-41-R (right-turn-only)	263	184	79	street-sign	70	49	21
RB-45 (through-only)	17	12	5	through-lane	17	12	5
RB-46-L (double-left-turn)	40	28	12	to-location-of-interest	158	111	47
RB-5 (maximum speed ahead)	195	137	58	traffic-light	93	65	28
RB-55 (no-stopping)	18	13	5	tree	61	43	18
RB-61 (dangerous-goods)	156	109	47	vehicle	82	57	25
RB-69 (dangerous goods route)	153	107	46	warning-road-may-flood	63	44	19
RB-75 (vehicle-inspection-station)	64	45	19				
RC-4-R (signal-stop-line)	36	25	11				
WA-106 (rumble-strips)	25	18	7				