Vision-Based Road Conditions Alert Systems in Connected Vehicle Environment for Accident-Prone Roads.

by

Haibo Cui

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

Master of Science

in

Signal and Image Processing

Department of Electrical and Computer Engineering
University of Alberta

Abstract

Work zones, being a critical component of roadway transportation systems, can benefit greatly from computer vision-enabled roadway infrastructures, specifically in connected vehicle (CV) environments. Connected infrastructures, such as roadside units (RSU) and on-board units (OBU), can greatly improve the environmental awareness and safety of CVs driving through a work zone. In this regard, the contribution of this thesis lies in developing a vision-based approach to generate work zone safety messages in real-time, utilizing video streams from roadside monocular traffic cameras that can be used by CV work zone safety apps on mobile devices to reliably navigate through a work zone. A monocular traffic camera calibration method is proposed to establish an accurate mapping between the image plane and Global Position System (GPS) space. Real test scenarios show that our algorithm can precisely and effectively locate work zone boundaries from a monocular traffic camera in real-time. We demonstrate the capabilities and features of our system through real-world experiments where the driver cannot see the work zone. End-to-end latency analysis reveals that the vision-based work zone safety warning system satisfies the active safety latency requirements. This vision-based work zone safety alert system ensures the safety of both the worker and the driver in a CV environment.

Winter roads that are covered by snow or ice, as seen in Alberta, can cause severe traffic accidents. Current winter road surface conditions (RSC) monitoring methods often generate incomplete RSC maps in city center areas. Cameras mounted on CVs and traffic cameras can be used as sensors to detect RSC. In this case, the contribution of this thesis focuses on developing automated RSC classification applications using CVs and traffic cameras in Alberta. Three state-of-the-art machine learning algorithms are trained and tested on RSC datasets. The pipeline of automated RSC classification applications in a CV environment is proposed. Comparisons of our methods versus current methods in real-world scenarios reveal our method can provide more detailed RSC maps in city center areas and narrow roads. Our RSC methods ensure the safety of drivers on winter roads.

Preface

The work presented in Section 2.1 and Chapter 3 of this thesis has been accepted for presentation and publication at the 102^{nd} Annual Meeting of the Transportation Research Board, Washington, DC, January 2023, and has been accepted for publication to the Transportation Research Record. I was responsible for the system development and training, data collection and analysis as well as manuscript composition. Dr. Tony and Dr. Mahdi were the supervisory authors and were involved with concept formation and manuscript composition. Kaizhe Hou, Jiarui Zhang, Siqi Yan, and Dr. Mudasser Seraj, and YingKe Wang assisted with the data collection, system development and contributed to manuscript edits.

- Haibo Cui, Kaizhe Hou, Jiarui Zhang, Siqi Yan, Mudasser Seraj and Tony Qiu (accepted), "Vision-Based Work Zone Safety Alert System in a Connected-Vehicle Environment", *Proceedings of Transportation Research Board 102nd Annual Meeting.*

- Haibo Cui, Kaizhe Hou, Jiarui Zhang, Siqi Yan, Mudasser Seraj, Yingke Wang, Mahdi Tavakoli and Tony Qiu (accepted), "Vision-Based Work Zone Safety Alert System in a Connected-Vehicle Environment", *Transportation Research Record: Journal of the Transportation Research Board.*

I also participated in the following important projects.

- Tony Qiu. Digitized Infrastructure as Service with ATS traffic Canada. 2020.

- Tony Qiu. Vehicle as Sensor with IBI Group. 2020.

Acknowledgements

I would like to take this opportunity to thank my supervisors Dr. Zhi-Jun (Tony) Qiu and Dr. Mahdi Tavakoli for their invaluable guidance, encouragement, and continuous support throughout the journey of my M.Sc. They were with me through any situation during my stay at the University of Alberta. I am grateful to them for helping me to gain confidence, think critically, and gain profound academic understanding. I thoroughly enjoyed working with them, and I hope to work with them again in the future.

I also want to thank Dr. Mudasser Seraj, Jiarui Zhang, Siqi Yan, and Kaizhe Hou for their guidance, effort, and support during the project which was the main part of this thesis.

My sincere appreciation also goes to ATS traffic Canada and IBI Group for their assistance and project sponsorship.

Special thanks to the University of Alberta and iSmartWays for their valuable help in collecting the data of this research.

I would also like to thank the members of my M.Sc. defense committee for their time and constructive suggestions to improve the quality of the thesis.

I am indebted to my parents and my friends for their continuous support and encouragement. Their unwavering support facilitated my dedication to my studies.

Table of Contents

List of Tables

List of Figures

Abbreviations

5G Fifth Generation Wireless.

CCTV Closed-Circuit Television.

CNN Convolutional Neural Network.

CPM Cooperative Perspective Message.

CPU Central Processing Unit.

C-V2X Cellular Vehicle-to-Everything.

CV Connected Vehicle.

DCGAN Deep Convolution Generative Adversarial Networks.

DL Deep Learning.

DSRC Dedicated Short Range Communication.

FN False Negative.

FP False Positive.

GAN Generative Adversarial Networks.

GNSS Global Navigation Satellite System.

GPS Global Position System.

GPU Graphic Processing Unit.

HMI Human Machine Interface.

ILSVRC ImageNet Large Scale Visual Recognition Competition.

LiDAR Light Detection and Ranging.

LT Linear Transformation.

LTE Long-Term Evolution.

MAP Map Data Message.

MEC Multi-Edge Computing.

ML Machine Learning.

MS-COCO Microsoft Common Objects in Context.

NHTSA National Highway Traffic Safety Administration.

OBU On Board Unit.

PING Packet Internet or Inter-Network Groper.

PSMs Personal Safety Messages.

PT Perspective Transformation.

R-CNN Region-Convolutional Neural Network.

RFS Rain-Fog-Snow.

RMSE Root Mean Square Error.

RNN Recurrent Neural Network.

RSC Road Surface Condition.

RSU Road Side Unit.

RTT Round Trip Time.

RWIS Road Weather Information System.

SAE Society of Automotive Engineers.

SGD Stochastic Gradient Descent.

SPAT Signal Phase and Timing Message.

SSD Single Shot Multibox Detector.

STC Smart Traffic Cone.

SVM Support Vector Machine.

TIM Traveller Information Message.

TN True Negative.

TP True Positive.

V2C Vehicle to Cloud.

V2I Vehicle to Infrastructure.

V2V Vehicle to Vehicle.

V2X Vehicle to Everything.

VGG Visual Geometry Group.

Wi-Fi Wireless Fidelity.

WZAD Work Zone Activity Data.

WZDx Work Zone Data Exchange.

YOLO You Only Look Once.

YOLOv5 You Only Look Once-Version 5.

Chapter 1

Introduction

1.1 Background

Roads that are covered by snow or have work zones can cause accidents according to National Highway Traffic Safety Administration (NHTSA) [1] and Alberta Transportation [2–6]. Every day, at least one traffic-related injury occurs in one of 70 work zones in North America [1]. Employees on highway construction and maintenance projects are frequently placed in close proximity to moving traffic. Despite the fact that various safety procedures are frequently taken to safeguard employees, these precautions may occasionally be insufficient due to a range of environmental and human variables, such as inattentive driving, severe weather, and poor road conditions. Larger vehicles, such as trucks or buses, typically require more room to merge, putting neighbouring cars and pedestrians in danger [7, 8]. Furthermore, cars behind trucks lose their ability to determine the precise position of lane merges as well as the appropriate speed to maintain, increasing the chance of rear-end collisions [9].

Drivers are constantly put in danger due to bad winter road conditions. When there is snow or ice on the roads, driving becomes more difficult and dangerous. Winter weather is recognized to be a major contributor to an increased chance of collisions due to factors such as lower friction on the road surface [10]. According to Alberta Transportation [2–6], slush, snow, or ice was involved in 27% of total casualty collisions in total from 2015 to 2019. Monitoring the status of roads is critical for those who manage winter roads as well as the public. Transportation officials must organize and coordinate several efforts to keep roads as clear of snow and ice as possible so that automobiles may use the road network safely. Road Surface Condition (RSC) is a measure that transportation authorities frequently use to identify the current state of the road in terms of snow or ice coating, as well as a communication method.

Given that the majority of work zone-related accidents may be prevented with early vehicle alerts, work zone safety research in a connected vehicle (CV) setting is a developing topic of research [11]. Drivers can benefit from greater situational awareness regarding future risks or situations by leveraging wireless communications such as Vehicle to Infrastructure (V2I), Vehicle to Vehicle (V2V), and Vehicle to Everything (V2X) [12]. In recent studies including Han et al. and Schonrock et al. [13, 14] of giving early work zone safety alerts to drivers, to locate the work zone borders in a CV environment, special equipment such as smart traffic cones and wearable localization devices are necessary. Such sophisticated technology is difficult and expensive to deploy. For example, it is doubtful that all personnel or traffic cones would be equipped with a localization device akin to a GPS sensor.

In terms of winter roads problems, recent studies including Carrillo et al., Wu et al., Pan et al., Ramanna et al. [15–18] have concentrated on the use of machine learning methods to automatically categorize and monitor RSC by images collected from cameras in vehicles or traffic cameras. However, little research has been conducted in Alberta on machine-learning approaches for automatically classifying RSC images in a CV environment. In Alberta, the major RSC monitoring methods are using the data from stationary and mobile Road Weather Information Systems (RWIS) stations [19], resulting in incomplete RSC maps, especially in city center areas.

1.2 Research Problem Statement

One of the problems of current research in the study of work zone alert systems in a CV environment is the work zone localization equipment. For example, in work zone alert systems developed by Han et al. [13], every construction worker attached a sensor that could provide accurate location in real-time. Although work zone alert systems have proved to be effective in some ways, installing such work zone alert systems in real-world scenarios can be troublesome and expensive. Thus, a research problem is how to safely navigate CVs through the work zone in a CV environment without using too many localization sensors.

In addition, some rear-end crashes happened near work zones because the driver's vision was blocked by front vehicles, such as large trucks [9]. If there is a way to tell the driver there is a work zone ahead, these kinds of accidents can be avoided. Sometimes large vehicles like buses or trucks need very early warnings of work zones ahead since these large vehicles require more room to change traffic lanes [7, 8]. It is safer to give bus or truck drivers safety alert messages before they can even see the work zone.

We directed our research to design a way to give drivers a warning that is not reliant on the driver's vision. Sometimes the work zone appeared on Google Maps or the 511 Alberta website [20] after the work zone had been there for hours or days. The drivers were endangered on these roads without knowing there was a work zone ahead. Thus, our research also focused on real-time work zone detection and warning broadcasting.

In terms of the monitoring problems of winter road conditions, we directed our research to use CVs and traffic cameras as sensors for detecting and classifying different RSCs since current methods usually result in an incomplete RSC map. 511 Alberta is the main source citizens can access and know which roads are covered by snow or ice. The main data used by 511 Alberta is from Road Weather Information Systems (RWIS) stations that are located in rural areas [19]. Thus the RSC maps provided by 511 Alberta often have detailed RSC reports on highways, yet no or little RSC reports on city center areas and some narrow, accident-prone roads. Traffic cameras and cameras mounted on vehicles can be used as sensors to collect RSC information, and with low latency communication like Dedicated Short Range Communications (DSRC), Cellular Vehicle-to-Everything (C-V2X), as well as Fifth-Generation Wireless (5G), drivers can receive RSC reports in real-time. We directed our research to RSC maps generated by CVs and traffic cameras.

The relationship between the work zone warning system and the RSC warning system is that cameras can be used as sensors in both systems. In specific, a traffic camera in Edmonton can be used to detect work zones in the summer months and classify RSC in the winter months. In a CV environment, the work zone or RSC information detected by the traffic camera can be broadcast from RSU to OBU to give drivers early warnings.

1.3 Research Objectives and Scopes

Considering the issues discussed in the previous section, the research needs to include localizing work zone items without a Global Positioning System (GPS) sensor, real-time work zone warning broadcasting, and RSC maps generated by CVs and traffic cameras. Reflecting on these major focuses, this research intends to design a video-based road conditions (work zone and snow) alert system in accident-prone roads under a CV environment. The overall objective of the research can be broken down into two major objectives: (a) Designing the work zone safety alert system in a CV environment. In contrast to earlier studies like Han et al. and Schonrock et al. [13, 14], this model will require no GPS sensor attached to construction workers and work zone items. We will select a real-time object detection algorithm to train our dataset containing work zone images in a real-world scenario. A few work zone localization strategies will be designed under different types of roads. The work zone detection and localization accuracy will also be studied. Additionally, this study will explore ways to broadcast lane closure information as well as work zone location to drivers via C-V2X. We will also present latency tests from a system level in a CV environment using C-V2X. The effect of different weather and camera resolution on our method will also be discussed; (b) Using CVs and traffic cameras as sensors to detect snow on roads and give RSC reports based on three main categories, including bare, partly covered, and fully covered according to 511 Alberta. State-of-the-art machine learning models will be studied and selected to train datasets that have RSC images in Alberta. The advantages and drawbacks of our methods will also be explored based on the comparisons with current in-use RSC monitoring methods by 511 Alberta.

The research scope will be restricted to roads in Alberta, Canada. The test site for work zone warning systems will be approximately 1.5 kilometres (km) long on 118 street northwest near the University of Alberta South Campus. The studied roadway has a static speed limit of 40 kilometres per hour (kph) and multiple traffic cameras, as well as many Road Side Units (RSUs) installed. The data collected on the studied roadway will be used to train the machine learning algorithm, verify the effectiveness of our work zone warning systems, and study the effect of different weather and camera resolutions. As for winter RSC classification problems, highways in Alberta will be used to collect data to train and test machine learning algorithms. Live traffic camera images from the 511 Alberta website will also be used as data sources to form datasets to train and test algorithms.

1.4 Contributions

1.4.1 Research Contributions

The research contributions of this thesis include: (1) Designing a work zone warning system using a monocular traffic camera as an input data source. (2) Proposing two work zone localization methods using a monocular traffic camera. (3) Investigating the performance of localization methods using different traffic cameras. (4) Proposing two automated RSC monitoring methods. (5) Investigating the performance of automated RSC monitoring methods in different locations.

1.4.2 Practical Contributions

The practical contributions of this thesis include: (1) Creating a work zone dataset with pixel-wise labels on traffic cones, traffic barrels, traffic barricades, vehicles and construction workers. (2) Implementing the communication framework for the work zone warning system in a CV environment. (3) Constructing a RSC dataset with image-wise labels on three categories, bare pavement, part snow-covered and full snow-covered. (4) Implementing the automated RSC methods in different locations.

1.5 Organization of the Thesis

As stated previously, this research is focused on two major issues, including work zones and snow on road detection, localization, and warning broadcasting, and the thesis is formed keeping those issues in mind. Chapter 2 focuses on the literature review of recent studies. We first discuss work zone detection methods, work zone safety alert systems in a CV environment, and standards for work zone information broadcasting conducted by researchers in recent years, then investigate various research applying machine learning algorithms to the problems of winter RSC classifications. Chapter 3 starts by discussing the background of work zone issues faced by governments and the public. The data collection is covered in terms of dataset collecting and labelling, training and testing the machine learning algorithm, calibrations of our localization methods, and localization error analysis. Then, a case study on 118 Street NW, Edmonton, Alberta is explored with a detailed experiment setup, design, and results. The limitations of our method are studied under different weather conditions and camera resolutions. The lane closure information broadcasting is then explored by different work zone standard messages and field tests as well as latency tests using C-V2X. Chapter 3 ends by summarizing the findings of our work zone warning systems in a CV environment. Chapter 4 begins by discussing the current issues and dangers caused by slippery winter roads and methods we can use to ensure the safety of drivers.

Three state-of-the-art machine learning algorithms are then selected for training and testing. Then we explain the location, device, and methods we used to establish our winter RSC datasets in Alberta. The training and testing on different machine learning algorithms are explored, with performance comparisons in terms of accuracy and processing time. The applications of our methods using CVs and traffic cameras are then explained and tested on roads in Alberta. Chapter 4 ends by summarizing the findings of automated RSC classifications using machine learning algorithms. Chapter 5 is the last chapter of this thesis, covering conclusions, limitations, and the direction of future research.

References

[1] N. H. T. S. Administration *et al.*, "2015 motor vehicle crashes: Overview," *Traffic safety facts: research note*, vol. 2016, pp. 1–9, 2016.

[2] *Alberta traffic collision statistics 2015*, 2017. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/97980388-9327-46fc-b01f-9451d5e076b6/download/ar2015.pdf.

[3] *Alberta traffic collision statistics 2016*, 2018. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/edf10f2d-9056-4d34-8470-9378f590c80d/download/ar2016.pdf.

[4] *Alberta traffic collision statistics 2017*, 2020. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/15f070d4-e0ee-4e49-b107-95fb036ee7d1/download/trans-alberta-traffic-collision-statistics-2017.pdf.

[5] *Alberta traffic collision statistics 2018*, 2021. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/982e6d4f-64d5-4167-81ca-b8c10d76fa59/download/trans-alberta-traffic-collision-statistics-2018.pdf.

[6] *Alberta traffic collision statistics 2019*, 2022. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/30344114-e6b3-459d-a74d-850b6b0d7c07/download/trans-alberta-traffic-collision-statistics-2019.pdf.

[7] S. D. Schrock, G. L. Ullman, A. S. Cothron, E. Kraus, and A. P. Voigt, "An analysis of fatal work zone crashes in texas," *Report FHW A/TX-05/0-4028*, vol. 1, 2004.

[8] J. Daniel, K. Dixon, and D. Jared, "Analysis of fatal crashes in georgia work zones," *Transportation Research Record*, vol. 1715, no. 1, pp. 18–23, 2000.

[9] N. J. Garber and M. Zhao, "Distribution and characteristics of crashes at different work zone locations in virginia," *Transportation Research Record*, vol. 1794, no. 1, pp. 19–25, 2002.

[10] I. Juga, P. Nurmi, and M. Hippi, "Statistical modelling of wintertime road surface friction," *Meteorological Applications*, vol. 20, no. 3, pp. 318–329, 2013.

[11] A. Dehman and B. Farooq, "Are work zones and connected automated vehicles ready for a harmonious coexistence? a scoping review and research agenda," *Transportation research part C: emerging technologies*, vol. 133, p. 103 422, 2021.

[12] X. Wu *et al.*, "Vehicular communications using dsrc: Challenges, enhancements, and evolution," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 399–408, 2013.

[13] W. Han, E. White, M. Mollenhauer, and N. Roofigari-Esfahan, "A connected work zone hazard detection system for roadway construction workers," in *Computing in Civil Engineering 2019: Smart Cities, Sustainability, and Resilience*, American Society of Civil Engineers Reston, VA, 2019, pp. 242–250.

[14] R. Schönrock, F. Wolf, and T. Russ, "Smart traffic cone-dynamic detection and localization of traffic disruptions," in *FAST-zero'15: 3rd International Symposium on Future Active Safety Technology Toward zero traffic accidents, 2015*, 2015.

[15] J. Carrillo, M. Crowley, G. Pan, and L. Fu, "Comparison of deep learning models for determining road surface condition from roadside camera images and weather data," in *Transportation Association of Canada and ITS Canada 2019 Joint Conference and Exhibition*, 2019.

[16] M. Wu and T. J. Kwon, "An automatic architecture designing approach of convolutional neural networks for road surface conditions image recognition: Tradeoff between accuracy and efficiency," *Journal of Sensors*, vol. 2022, 2022.

[17] G. Pan, M. Muresan, R. Yu, and L. Fu, "Real-time winter road surface condition monitoring using an improved residual cnn," *Canadian Journal of Civil Engineering*, vol. 48, no. 9, pp. 1215–1222, 2021.

[18] S. Ramanna, C. Sengoz, S. Kehler, and D. Pham, "Near real-time map building with multi-class image set labeling and classification of road conditions using convolutional neural networks," *Applied Artificial Intelligence*, vol. 35, no. 11, pp. 803–833, 2021.

[19] *Emerging transportation technologies*. [Online]. Available: https://www.alberta.ca/emerging-transportation-technologies.aspx.

[20] *511 alberta*. [Online]. Available: https://511.alberta.ca/#:Alerts.

Chapter 2

Literature Review

In this Chapter, work zone-related studies will be discussed first, then RSC classification-related studies will be investigated.

2.1 Work Zone-Related Studies

In this section, we will look at the current literature on work zone detection methods, work zone safety-related research in a CV environment, and work zone safety message requirements.

2.1.1 Work Zone Detection Methods

Existing work zone detection research is primarily concerned with identifying the existence or absence of a work zone. Their detection results are generally vague and lack work zone boundary information. Abodo et al. (2018) find work zones by employing a CNN to identify work zone photos [21]. Seo et al. (2015) use the vehicle's camera to recognize traffic signs [22]. This approach then identifies work zone signs and utilizes them to indicate the starting and ending point of a work zone road segment. Mathibela et al. (2012) also identify traffic signs and traffic cones [23]. This method employs the detection findings as characteristics to calculate the likelihood that the vehicle is in a work zone. Kunz et al. (2017) create a Bayesian network for detecting a work zone [24]. The Bayesian network uses identified traffic objects and vehicle statuses as input to forecast the likelihood of a work zone at various distance bins. None of these studies infer detailed geometric properties like work zone areas and borders. Knowing the work zone geometry is crucial for CVs to drive safely.

Graf et al. (2012) investigate a more limited example in which temporary lane markers assist drivers through a work zone [25]. This method monitors the lanes even if both temporary and original lane markers are present. As a result, it allows the vehicle to follow the lanes in a work zone. However, such a strategy is reliant on lane marking regulations and is unable to deal with regular work zones that lack temporary lane markers. Shi et al. (2021) [26] develops work zone detection, which detects and locates the boundaries of a work zone. They give many baseline implementations utilizing various sensor combinations, such as camera and LiDAR. Several cutting-edge deep learning-based object detection strategies, such as the Region-Convolutional Neural Network (R-CNN) [27], Fast RCNN [28], Faster R-CNN [29], Single Shot MultiBox Detector (SSD) [30], and You Only Look Once-Version 5 (YOLOv5) [31], are used for real-time applications, taking advantage of high-performance GPU-enabled computing devices. The YOLOv5 outperforms all previous state-of-the-art object detection deep learning models [31, 32], with higher detection accuracy and lower detection time. Unfortunately, all of the studies mentioned in this subsection do not have the capability of determining the precise location of a work zone in terms of latitude and longitude in real-time.

In order to address these research gaps in work zone detection methods, we introduce a YOLO-based work zone detection and localization method in Chapter 3. We contribute to the existing work zone detection methods by precisely positioning a work zone area, and then broadcasting work zone safety messages to nearby CVs in a CV environment.

2.1.2 Work Zone Safety-Related Studies in a CV Environment

Researchers began to think about how CVs may be combined with work zone safety as they became more prevalent. Vehicles that are connected can communicate with the driver, other vehicles on the road (V2V), roadside infrastructure (V2I), and the cloud server (V2C) via various communication methods through DSRC, Wi-Fi, and cellular communication technologies [33, 34]. DSRC technology has reduced connection latency less than existing Wi-Fi and Cellular LTE technologies, making it a far faster two-way communication alternative for information sharing [35].

Vehicles may now be viewed as integrated components of a system rather than as independent actors on the road, thanks to the incorporation of communication technologies. When the effects of CVs on safety performance are investigated in a work zone context, Abdulsattar et al. (2018) [36] reveals that V2V/V2I communication can increase work zone safety performance at low traffic flow rates. Genders et al. (2015) [34] also investigated how employing CVs in a network with work zones affected traffic safety. The Michigan Department of Transportation and 3M [37] constructed the first connected work zone in the United States. In this connected work zone setting, orange barrels with 2D barcodes were supplied by 3M, and the CV's infrared devices sent information to the vehicle and the driver by reading the barcode. Han et al. (2019) [13] designs a connected work zone alert system with wearable localization devices that can be placed on workers in work zones. This method monitors the potential danger between CVs and workers by calculating the collision risks from both CVs and workers' trajectories. Schonroack et al. (2015) [14] developed a traffic cone with GPS and communication sensors that can be placed at a work zone boundary to give the location of a work zone in a CV environment. In this approach, the location of this special traffic cone is sent to a central server and then broadcast to nearby CVs. Mishra et al. (2021) [38] developed work zone alert systems from work zone intrusion technologies and Qiao et al. (2017) [39] used a cell phone app to give drivers early warnings of work zones. The development of location sensing technologies and CVs have made it possible to gather data from all parties involved in a work zone, including vehicles, equipment, and workers on foot, and use it to ensure the safety of the work zone. However, current research in the localization of work zones for a CV environment usually requires localization devices, including GPS sensors attached to construction equipment, traffic cones, or workers, which makes it costly and difficult to set up.

In summary, Table 2.1 shows comparisons of different warning systems in a CV environment and three commercially available warning systems analyzed by Mishra et al. [38] are compared and explained in Table 2.2. These methods utilize pressure, radar sensors, or additional equipment to recognize incoming traffic. Unfortunately, such equipment usually costs a lot and has poor mobility. In comparison, our method uses a camera for object detection, which has two main advantages. First, it can recognize various types of objects than traffic, including cones, barrels, vehicles, people, etc. Due to the nature of computer vision, we can feed our YOLOv5 models with as many object types as desired. Second, Our method has a lower expense and higher mobility than other proposed method devices. The only device we need onsite is the camera. The device is easy to obtain, move around, and adapt to work zones with different shapes and environments.

To overcome the limitations of current research in localizing work zones, we provide a framework in Chapter 3 to increase work zone safety using a vision-based deep learning technique, assuming that workers, construction equipment, and traffic cones do not have a localization device.

2.1.3 Standards for Work Zone Safety Messages

The SAE J2945 standard [41] defines the work zone safety messages for safety data communication between the work zone and other associated components (e.g., vehicle and traffic signals). Although the work zone safety messages standard is defined by SAE J2945, the format, and structure of the message, data frames, and data components for sharing data between work zones and vehicles, as well as between work zones and infrastructure, are defined by SAE J2735 [42]. In contrast, the SAE J2945 considers all of the data items established in the SAE J2735. Work zone safety messages data items are shown in Table 2.3.

Work Zone Activity Data (WZAD) – Data Dictionary Report [43], developed as part of the Federal Highway Administration's Work Zone Data Initiative Project, defines and standardizes digital descriptions of work zone activities, allowing transportation authorities and third-party providers to describe and communicate work zone information. Work zone activities data structure based on WZAD is shown in Figure 2.1.

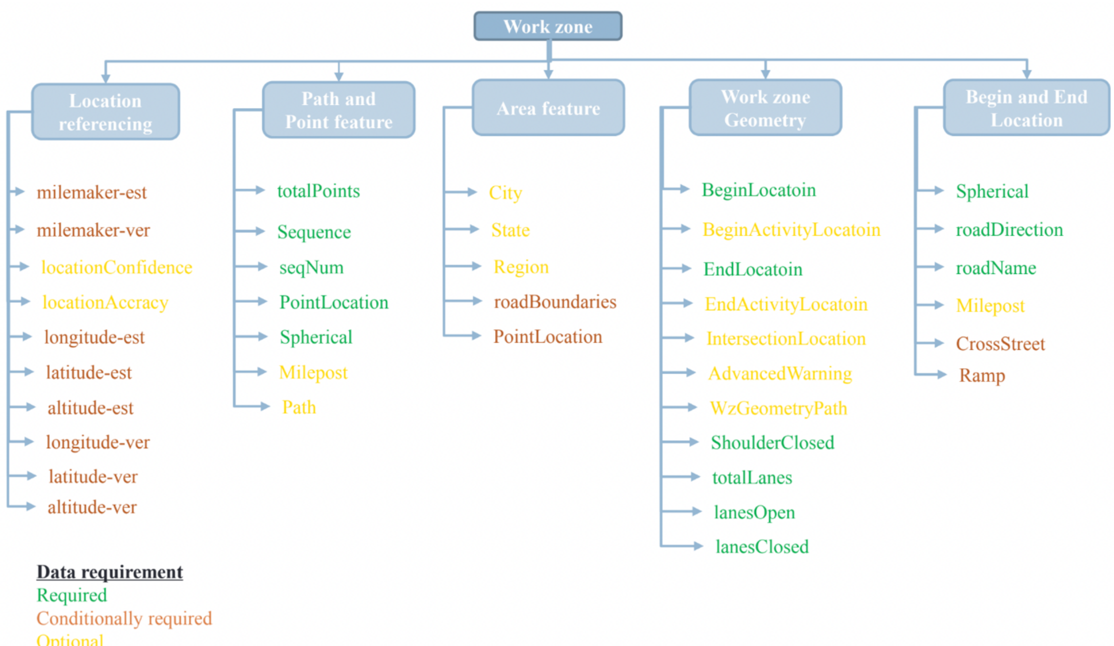


Figure 2.1: Work Zone Activities Data Structure Based on WZAD.

2.2 RSC Classification-Related Studies

Considerable research had been conducted using deep-learning models to classify weather conditions. Elhoseiny et al. [44] used ImageNet (a large dataset containing usual objects) [45] to train a Convolutional Neuron Network (CNN) to categorize weather photos as sunny or cloudy. Lu et al. [46] propose a collaborative learning strategy that uses innovative weather variables to categorize a single outdoor photograph as sunny or cloudy. To extract characteristics, a CNN was employed, which was then input into an SVM framework to generate individual weather features. A data augmentation strategy was also applied to supplement the training data. Lin et al. [47] constructed a multi-class benchmark dataset with six common categories for sunny, overcast, rainy, snowy, hazy, and thunder weather. To identify visual concurrency on area pairs of weather categories, a region selection and concurrency method were presented. A deep-learning framework was used to test this model. Carrillo et al. [15] created a CNN with a basic architecture from scratch and compared its results in terms of classification accuracy to automatically categorize winter RSC pictures (from stationary RWIS cameras) to other pre-trained CNN models. This experiment was conducted with three categories of RSC: dry/wet, partly snow-covered, and totally snow-covered. The results supported CNN's efficacy in detecting RSC using imaging, as all CNN models provided high classification accuracy, with the author's model being the best. However, the authors cautioned that the results might only be informative for their unique application.

Zhu et al. [48] developed severe weather characteristics and recognition models from a large-scale extreme weather dataset in which over 16,000 extreme weather photos with complex scenarios were classified into four classifications including sunny, rainstorm, blizzard, and fog. Pre-training and fine-tuning are used to create an extreme weather identification model. Rain-Fog-Snow (RFS) Dataset is a new open-source weather conditions dataset published by Guerra et al. [49] that consists of photos illustrating three types of weather: rain, snow, and fog. A unique approach that uses superpixel delimiting masks as a kind of data augmentation has also been suggested, yielding respectable results in comparison to 10 CNNs. Li et al. [50] proposed a method for data augmentation based on generative adversarial networks

(GAN). It can augment and complete picture data diversity. The author built a system that uses a deep convolution generative adversarial networks (DCGAN) model as a generator to generate pictures to balance the unbalanced data and a CNN model as a classifier to check the classification results. The author also presented an assessment approach on three benchmark datasets as a comparison experiment to validate the performance of DCGAN. The empirical results showed that high-quality weather images can be created on weather data sets using DCGAN. Weather recognition was considered by Zhao et al. [51] as a multi-label classification challenge, in which a picture was assigned many labels based on the exhibited weather conditions. Then, a multi-label classification strategy based on CNN-RNN was presented. To extract the most associated visual information, the CNN was enhanced with a channel-wise attention model. The Recurrent Neural Network (RNN) analyzed the information further and discovered the correlations between weather types.

Wu et al. [16] suggested a novel method for autonomously designing RSC CNN architecture without sacrificing classification accuracy. The suggested method employed a weighted sum method, which allowed for the selection of the relative relevance level between accuracy and efficiency. The findings of this study bridged a gap in existing CNN design approaches that do not account for the tradeoff between accuracy and efficiency, while also offering insight into the impact of different architectures on CNN model performance. Ramanna et al. [18] used cutting-edge CNNs to categorize photos captured by street and highway cameras across North America. To identify photos by road condition, road camera images were used in studies with several deep learning frameworks. These studies employed photos labelled as dry, wet, snow/ice, poor, and offline as training data. The trials evaluated the suitability of six CNNs in various configurations. The identified photos were then utilized to create a map of real-time road conditions across North America at various camera sites. By Pan et al. [17, 52], photographs from stationary weather/traffic cameras or in-car electronics were used to train and fine-tune four state-of-the-art CNN models. The results of their studies demonstrated that CNN was a promising method for addressing the RSC recognition issues and can be useful in assisting winter road maintenance decision-making. Similar findings were also revealed by a number of other researchers including Linton et al. [53, 54], Kuehnle et al. [55] and Zhang et al. [56]. However, there is no public RSC dataset collected in Alberta and little automated RSC classification application in a CV environment in Alberta.

We contribute to the existing RSC classification methods by proposing a RSC monitoring pipeline using onboard and traffic cameras as input data, and ML algorithms trained by datasets collected in Alberta and Ontario to automatically classify RSC in Chapter 4.

Author	**Real-Time**	**Lane Level Localization**	**Special Localization Device**	**Summary**
Han et al. [13]	Yes	Separate a single line into four zones for detection.	Wearable localization devices that can be placed on workers in work zones.	Using risk score to calculate the collision risks from both CVs and workers' trajectories. The algorithm can also be used in the vehicle turning and workers from different categories have their own risk score.
Schonroack et al. [14]	Yes	By extending smart traffic cone (STC) using the GNSS model, the position accuracy is in the sub-meter range.	Traffic cone with GPS and communication sensors.	The location of these 16 special traffic cones is sent to a central server and then broadcast to nearby CVs.
Qiao et al. [39]	Yes	N/A	No extra device needed.	Pre-installed application detects the approaching construction zone based on geo-location in a phone to give drivers early warnings of work zones.
Islam et al. [40]	Yes	Not at lane level.	Site camera and personal safety messages(PSMs) devices carried by roadside users.	Generate PSMs using real-time video streams collected from a traffic camera. Train the YOLOv3 model with the video data from a roadway section to detect pedestrians.
Genders et al. [34]	N/A	N/A	N/A	Build a control simulation of a network with a work zone to simulate various market penetrations of 20%, 40%, 60%, 80%, and 100% connected vehicles to determine their effect on the safety of the network.

Table 2.1: Comparisons of Recent Studies

Method	**Real-Time**	**Lane Level Localization**	**Special Localization Device**	**Note**
Intellicone	Yes	Not at lane level.	Need traffic cone mounted sensor and portable site alarm.	Results are inconsistent and not available in the US.
Advanced Warning and Risk Evasion	Yes	Lane intrusion system (not yet available for evaluation).	Onsite sentry needed.	Warning equipment is worn by personnel and sentry in front of the work zone.
Worker Alert System	Delayed	Not at lane level.	A pneumatic trip hose sensor with a signal transmitter, a Portable Alarm case, personal safety device.	Workers and pedestrians are alerted when sensors on the ground detect vehicle pressure.

Table 2.2: Commercially Available Methods

Data Element	**Purpose**
Road Segment	Store complete description including road Geometry allowed navigational paths, and any current disruptions such as a work zone or incident event.
Road Sign ID	Used to provide a precise location of one or more roadside signs.
Traveler Data	Used to send a single message in traveller information message. It uses the ITIS encoding system to send well-known phrases but allows limited text for local place names.
Public Safety and Road Worker Activity	Used to describe the type of activity a worker or workers are engaged in.
Speed Limit Type	Relates the type of speed limit to which a given speed refers.
Personal Safety Message	Used to broadcast safety data regarding the dynamic state of various types of Vulnerable Road Users (VRU), such as pedestrians, cyclists, or road workers.

Table 2.3: Data Elements for Work Zone Safety Messages Based on SAE J2945 and SAE J2735.

References

[13] W. Han, E. White, M. Mollenhauer, and N. Roofigari-Esfahan, "A connected work zone hazard detection system for roadway construction workers," in *Computing in Civil Engineering 2019: Smart Cities, Sustainability, and Resilience*, American Society of Civil Engineers Reston, VA, 2019, pp. 242–250.

[14] R. Schönrock, F. Wolf, and T. Russ, "Smart traffic cone-dynamic detection and localization of traffic disruptions," in *FAST-zero'15: 3rd International Symposium on Future Active Safety Technology Toward zero traffic accidents, 2015*, 2015.

[15] J. Carrillo, M. Crowley, G. Pan, and L. Fu, "Comparison of deep learning models for determining road surface condition from roadside camera images and weather data," in *Transportation Association of Canada and ITS Canada 2019 Joint Conference and Exhibition*, 2019.

[16] M. Wu and T. J. Kwon, "An automatic architecture designing approach of convolutional neural networks for road surface conditions image recognition: Tradeoff between accuracy and efficiency," *Journal of Sensors*, vol. 2022, 2022.

[17] G. Pan, M. Muresan, R. Yu, and L. Fu, "Real-time winter road surface condition monitoring using an improved residual cnn," *Canadian Journal of Civil Engineering*, vol. 48, no. 9, pp. 1215–1222, 2021.

[18] S. Ramanna, C. Sengoz, S. Kehler, and D. Pham, "Near real-time map building with multi-class image set labeling and classification of road conditions using convolutional neural networks," *Applied Artificial Intelligence*, vol. 35, no. 11, pp. 803–833, 2021.

[21] F. Abodo, R. Rittmuller, B. Sumner, and A. Berthaume, "Detecting work zones in shrp 2 nds videos using deep learning based computer vision," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, IEEE, 2018, pp. 679–686.

[22] Y.-W. Seo, J. Lee, W. Zhang, and D. Wettergreen, "Recognition of highway workzones for reliable autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 708–718, 2014.

[23] B. Mathibela, M. A. Osborne, I. Posner, and P. Newman, "Can priors be trusted? learning to anticipate roadworks," in *2012 15th International IEEE Conference on Intelligent Transportation Systems*, IEEE, 2012, pp. 927–932.

[24] P. Kunz and M. Schreier, "Automated detection of construction sites on motorways," in *2017 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, 2017, pp. 1378–1385.

[25] R. Graf, A. Wimmer, and K. C. Dietmayer, "Probabilistic estimation of temporary lanes at road work zones," in *2012 15th International IEEE Conference on Intelligent Transportation Systems*, IEEE, 2012, pp. 716–721.

[26] W. Shi and R. R. Rajkumar, "Work zone detection for autonomous vehicles," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, IEEE, 2021, pp. 1585–1591.

[27] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587.

[28] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448.

[29] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, vol. 28, 2015.

[30] W. Liu *et al.*, "Ssd: Single shot multibox detector," in *European conference on computer vision*, Springer, 2016, pp. 21–37.

[31] G. Jocher, *Yolov5. code repository https://www. github. com/ultralytics/yolov5*, 2022.

[32] J. Redmon and A. Farhadi, "Yolo9000: Better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271.

[33] C. for Advanced Automotive Technology. "Connected and automated vehicles." (2018), [Online]. Available: http://autocaat.org/Technologies/Automated_and_Connected_Vehicles/ (visited on 09/12/2022).

[34] W. Genders and S. N. Razavi, "Impact of connected vehicle on work zone network safety through dynamic route guidance," *Journal of Computing in Civil Engineering*, vol. 30, no. 2, p. 04 015 020, 2016.

[35] K. C. Dey, A. Rayamajhi, M. Chowdhury, P. Bhavsar, and J. Martin, "Vehicle-to-vehicle (v2v) and vehicle-to-infrastructure (v2i) communication in a heterogeneous wireless network–performance evaluation," *Transportation Research Part C: Emerging Technologies*, vol. 68, pp. 168–184, 2016.

[36] H. Abdulsattar, A. Mostafizi, and H. Wang, "Surrogate safety assessment of work zone rear-end collisions in a connected vehicle environment: Agent-based modeling framework," *Journal of Transportation Engineering, Part A: Systems*, vol. 144, no. 8, 2017.

[37] J. Lombardo. "How connected vehicles make work zones safer." (2017), [Online]. Available: https://www.forconstructionpros.com/asphalt/article/20867242/how-connected-vehicles-make-work-zones-safer (visited on 09/12/2022).

[38] S. Mishra, M. M. Golias, D. Thapa, *et al.*, "Work zone alert systems," Tennessee. Department of Transportation, Tech. Rep., 2021.

[39] F. Qiao, R. Rahman, Q. Li, and L. Yu, "Safe and environment-friendly forward collision warning messages in the advance warning area of a construction zone," *International journal of intelligent transportation systems research*, vol. 15, no. 3, pp. 166–179, 2017.

[40] M. Islam, M. Rahman, M. Chowdhury, G. Comert, E. D. Sood, and A. Apon, "Vision-based personal safety messages (psms) generation for connected vehicles," *IEEE transactions on vehicular technology*, vol. 69, no. 9, pp. 9402–9416, 2020.

[41] V. V. A. T. Committee, *Vulnerable road user safety message minimum performance requirements*, Mar. 2017.

[42] V. C. T. Committee, *Dedicated short range communications (dsrc) message set dictionary*, Mar. 2016.

[43] O. Polly, D. Stephens, and R. Ostroff, "Work zone activity data (wzad) – data dictionary report (version 2)," 2019.

[44] M. Elhoseiny, S. Huang, and A. Elgammal, "Weather classification with deep convolutional neural networks," in *2015 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2015, pp. 3349–3353.

[45] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, Ieee, 2009, pp. 248–255.

[46] C. Lu, D. Lin, J. Jia, and C.-K. Tang, "Two-class weather classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3718–3725.

[47] D. Lin, C. Lu, H. Huang, and J. Jia, "Rscm: Region selection and concurrency model for multi-class weather recognition," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4154–4167, 2017.

[48] Z. Zhu, L. Zhuo, P. Qu, K. Zhou, and J. Zhang, "Extreme weather recognition using convolutional neural networks," in *2016 IEEE International Symposium on Multimedia (ISM)*, IEEE, 2016, pp. 621–625.

[49] J. C. V. Guerra, Z. Khanam, S. Ehsan, R. Stolkin, and K. McDonald-Maier, "Weather classification: A new multi-class dataset, data augmentation approach and comprehensive evaluations of convolutional neural networks," in *2018 NASA/ESA Conference on Adaptive Hardware and Systems (AHS)*, IEEE, 2018, pp. 305–310.

[50] Z. Li, Y. Jin, Y. Li, Z. Lin, and S. Wang, "Imbalanced adversarial learning for weather image generation and classification," in *2018 14th IEEE International Conference on Signal Processing (ICSP)*, IEEE, 2018, pp. 1093–1097.

[51] B. Zhao, X. Li, X. Lu, and Z. Wang, "A cnn–rnn architecture for multi-label weather recognition," *Neurocomputing*, vol. 322, pp. 47–57, 2018.

[52] G. Pan, L. Fu, R. Yu, and M. I. Muresan, "Winter road surface condition recognition using a pre-trained deep convolutional neural network," Tech. Rep., 2018.

[53] M. A. Linton and L. Fu, "Connected vehicle solution for winter road surface condition monitoring," *Transportation Research Record*, vol. 2551, no. 1, pp. 62–72, 2016.

[54] M. A. Linton and L. Fu, "Winter road surface condition monitoring: Field evaluation of a smartphone-based system," *Transportation research record*, vol. 2482, no. 1, pp. 46–56, 2015.

[55] A. Kuehnle and W. Burghout, "Winter road condition recognition using video image classification," *Transportation research record*, vol. 1627, no. 1, pp. 29–33, 1998.

[56] C. Zhang, "Application of convolutional neural network (cnn) models for automated monitoring of road pavement and winter surface conditions using visual-spectrum and thermal video cameras," 2021.

Chapter 3

Vision-Based Work Zone Safety Alert System in a Connected-Vehicle Environment

3.1 Introduction

Transportation authorities and the public are concerned about safety in work zones on highways. According to the National Highway Traffic Safety Administration (NHTSA) [1], there is at least one traffic-related injury in 70 work zones daily. Additionally, NHTSA shows that work zone collisions have greater mortality rates than crashes outside of them. It is common for employees on highway construction and maintenance projects to be in close proximity to moving traffic. Although many safety precautions are routinely taken to protect workers, these precautions may be insufficient owing to a variety of environmental and human variables, such as distracted driving, bad weather, and poor road conditions.

Work zones typically feature advanced warning zones with visual warning signs to alert oncoming vehicles. Static signs are the most prevalent type of warning system. Furthermore, dynamic warning systems are frequently utilized to improve traffic flow in work zones [57]. Because diverse operating characteristics of arriving vehicles and their relative positioning near a work zone are rarely considered in the design phase of these alerts, they typically fail to adequately aid drivers. Larger vehicles, such as trucks or buses, usually require more space to merge, endangering adjacent vehicles and pedestrians [7, 8]. Furthermore, vehicles behind trucks lose their ability to discern the precise location of lane merges and the proper speed to maintain, increasing the likelihood of rear-end collisions. Rear-end collisions are the most prevalent kind of crash in a work zone's advanced warning area, according to Garber et al. (2002) [9]. According to Garber et al. (2002) [9] and Nemeth et al. (1978) [58], the advanced warning region accounts for 10%-35% of all work zone collisions. Furthermore, comparative accident evaluations conducted by Hall and Lorenz (1989) and Rouphail et al. (1988) [59, 60] demonstrated that rear-end collisions are more likely in work zones than in non-work zones.

Given that the majority of work zone-related events can be avoided with early alerts to vehicles, safety research relating to work zones in a CV environment is an emerging area of research [11]. Through vehicle-to-infrastructure (V2I) communication, enabling a low latency communication channel, such as dedicated short-range communication (DSRC) [12] and upcoming 5G technology, may dramatically improve work zone safety. The drivers of CVs are alerted to potential collision hazards using work zone early warning messages provided by the DSRC or 5G-enabled V2I communication. Through communicated proactive decision-making aids using in-vehicle displays, also known as the Human Machine Interface (HMI), drivers can benefit from increased situational awareness about upcoming hazards or conditions by utilizing wireless communications such as: V2I, vehicle-to-vehicle (V2V), and vehicle-to-everything (V2X). The transmission of early warnings is one benefit that CVs can use. For instance, early in-vehicle lane closure alerts can be used to meet the demands of heavy truck drivers who need to move to the available lane in the work zone, well in advance of the lane closure area [61]. Significantly, a survey conducted by Benekohal et al. (1995) [62] revealed that over half of truck drivers preferred that warning signs for work zones be put up to 3 to 5 miles in advance of the work zone.

In current research studies of work zone alert systems in a CV environment, special equipment including smart traffic cones (traffic cones mounted with a GPS sensor) and wearable localization devices (mobile tags that can be placed on the worker with a localization sensor) are required to locate the work zone boundaries [13, 14]. Such specialized equipment is difficult to deploy and expensive. For instance, all workers or traffic cones are unlikely to attach a localization device similar to a GPS sensor.

To solve the research question of how connected infrastructures can help vehicles to navigate through a work zone, the objective of this study is to provide CVs with work zone safety alerts in real-time. We provide a system for generating work zone safety alerts, utilizing real-time video feeds gathered from traffic cameras, in order to get around these restrictions and improve work zone safety. The contribution of this study is the real-time generation (every 100 milliseconds) of safety warnings and work zone safety messages in accordance with the Society of Automotive Engineers (SAE) J2945 standards [41]. The roadside infrastructure generates and broadcasts work zone safety alerts to approaching CVs using work zone information (such as location). No localization device attached to workers or work zone boundaries (traffic cones) is necessary for our work zone boundaries localization method. We evaluate the accuracy of generated work zone safety messages by comparing them with field-collected ground truth data. Furthermore, we validate our vision-based approach at the system level in the real-world road environment field test.

The remainder of the chapter is organized as follows. We begin by discussing the real-time vision-based work zone safety message-generating method, which is not dependent on work zone localization devices. The evaluation of the vision-based work zone safety message generation is then presented. Then, we discuss the system-level validation utilizing the vision-based system's generated work zone safety alerts. Finally, we go over the study's findings and potential directions for further research.

3.2 Work Zone Safety Alerts Systems

We created a system that used a real-time camera feed to produce work zone safety messages by YOLOv5 and provided a safety warning in the event of a probable vehicle-work zone collision. The primary data elements of the work zone safety messages were the work zone geometry, including the work zone starting point, ending point, and lane closures as given in Figure 2.1. Furthermore, positional accuracy was determined by the precision of work zone positioning information (i.e., longitude and latitude). Thus, after correctly identifying a work zone, the localization of the work zone must also be correct in order to create work zone safety messages. To locate a work zone, we created a mathematical technique for converting an image's pixel coordinates to global coordinates. After determining a work zone's position, we built the work zone safety messages in accordance with the SAE J2945 standard and WZAD outlined in the preceding section.

In our system, the monocular traffic camera served as our primary data source. The ethernet, which is a wired communication method, transmitted real-time video data to the central server. The video data was then processed by the central server using YOLOv5 to identify traffic cones, which are commonly used to define work zone boundaries, and to convert pixel coordinates to global coordinates. The RSU then received the work zone position information by ethernet and transmitted it to the CV's OBU by DSRC. The HMI developed in this system was a work zone warning app that could be used on a driver's cellphone or tablet. This app, connected to OBU's WIFI, was meant to show the driver real-time work zone safety messages. The communication topology in this system and the experiment setup are shown in Figure 3.1 and Figure 3.2 respectively.

Localization for work zone boundaries (traffic cones) is vital to generate work zone safety alerts. But normally, monocular traffic cameras have poor localization ability compared to LiDAR sensors. Inspired by other research including Islam et

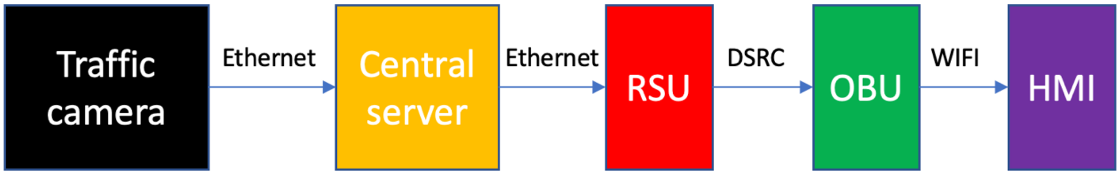


Figure 3.1: Communication Topology in the Vision Based Work Zone Alert System.

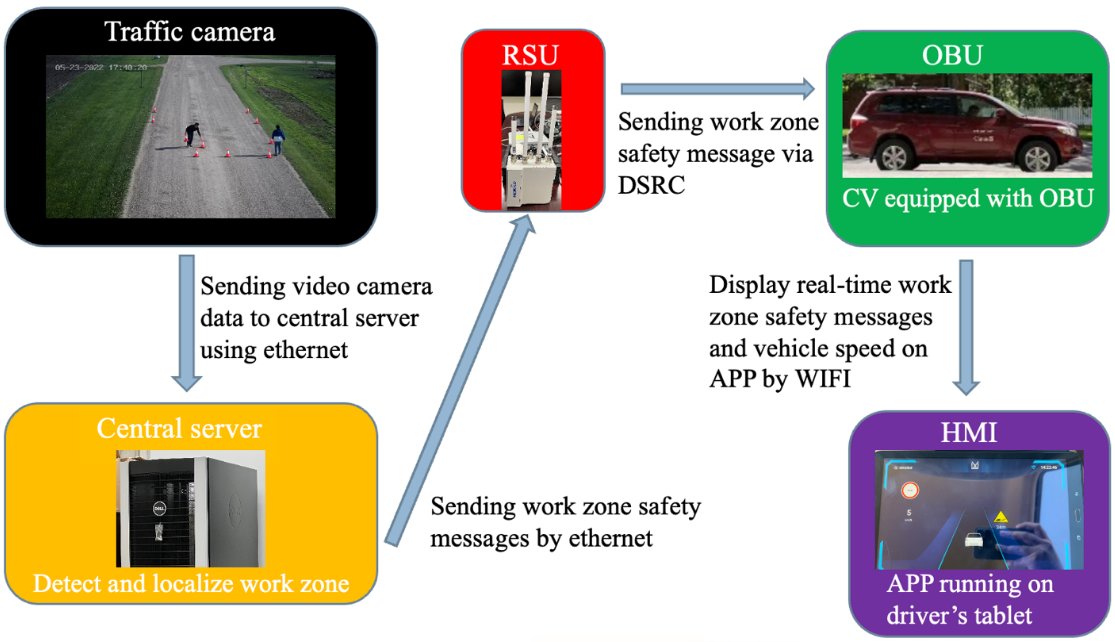


Figure 3.2: Experiment Setup for the Vision-Based Work Zone Alert System in CV Environment.

al. and Wen et al. on object localization from monocular traffic cameras [40, 63], we developed a low-cost monocular traffic camera calibration method for work zone boundaries localization without knowing the camera's intrinsic parameters. First, the traffic camera image plane was converted into a top-down view by perspective transformation (PT) in Equation 3.1. PT is a technique to obtain a different view from an image. We developed Equation 3.2 to convert from pixel coordinates into GPS coordinates by linear transformation (LT). The methodology to locate a work zone from monocular traffic cameras is shown in Figure 3.3.

Perspective transformation (PT): convert from pixel coordinates (p_x,p_y) in a traffic camera image plane into pixel coordinates (p'_x,p'_y) in the top-down view image

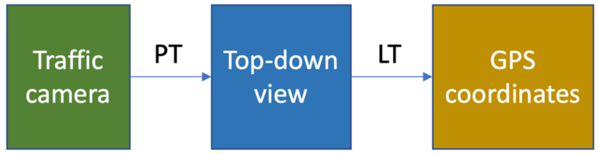
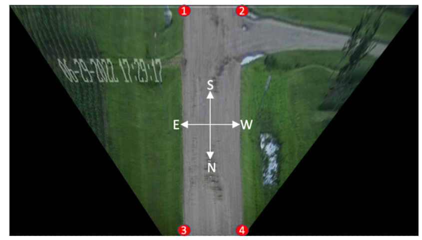


Figure 3.3: Method Used to Transform Pixel Coordinates to GPS Coordinates.

(a) The traffic camera image plane before PT.



(b) The traffic camera image plane after PT.

Figure 3.4: Perspective Transformation (PT).

plane by transformation matrix M shown in Equation 3.1. Where M is calculated from four pairs of matching points (shown in Figure 3.4) by python using the function cv2.getPerspectiveTransform. Shown in Figure 3.4a, the four red dots are the intersection points of the road margin and vision margin and they are chosen to be matching points. First, the pixel coordinates of the four matching points were recorded in the traffic camera plane and then we changed the pixel coordinates in the x-axis (horizontal direction) of the lower two matching points so that the four matching points formed a rectangle. This method assumes that the road is straight and the camera is horizontally placed. The new pixel coordinates of the four matching points were recorded. Then M was calculated from the original and the new pixel coordinates of the four matching points.

$$[p_x', p_y', 1]^T = M \cdot [p_x, p_y, 1]^T \tag{3.1}$$

Linear transformation (LT): convert from pixel coordinates (p'_x,p'_y) in top-down view image plane into GPS coordinates (longitude,latitude) by transformation matrix L shown in Equation 3.2. Where L is calculated from calibration points with measured

(p'_x,p'_y) in the top-down view image plane and GPS coordinate (longitude,latitude) by python using the function np.linalg.lstsq. The detailed steps to calculate L will be covered in Section 3.3.3.

$$[longitude, latitude] = [p_x', p_y', 1] \cdot L \tag{3.2}$$

The road touching point of a traffic cone can be approximated by the bottom center of the detected bounding box from YOLOv5 illustrated by the black point in Figure 3.5.

Figure 3.5: Road Touching Point of a Traffic Cone.

To locate a traffic cone, we took the road touching point (p_x,p_y) from YOLOv5 to obtain the pixel coordinate in top-down view (p'_x,p'_y) by PT shown in Equation 3.1, and then used (p'_x,p'_y) to get the GPS coordinate (longitude,latitude) by LT shown in Equation 3.2. Then work zone safety alerts were generated from the GPS coordinates of traffic cones used to form work zone boundaries.

3.3 Data Collection

Field data is collected for: 1) training and testing of YOLOv5 to ensure a high level of traffic cone detection accuracy; 2) calibrating the distortion in the top-down view from PT to valid the transformation M; 3) calibrating the linear transformation matrix L in Equation 3.2 to transform pixel coordinates to GPS coordinates; 4) localization error analysis on vision-based work zone boundary localization method to measure the localization accuracy.

3.3.1 Training and Testing YOLOv5

Phase One

In phase one of the experiments, we focused on training and testing YOLO to detect small traffic cones shown in Figure 3.5.

According to the official documentation of YOLOv5 [31], the recommended number of training images to detect an object is 1200 images. We placed traffic cones randomly in the vision range of traffic cameras at 118 Street NW, Edmonton under different weather conditions. The traffic cameras captured images to form a dataset to train and test YOLOv5. We annotated each traffic cone in every image of the dataset with a bounding box to generate ground truth data. Then, some dataset preprocessing techniques were applied to decrease training time and increase performance by applying image transformations to all images in this dataset. The preprocessing techniques [31] used in this dataset were: 1) auto-orientation, to correct a mismatch between the annotation and the image; 2) reducing the size of the image to train faster throughout the training phase.

Dataset augmentation techniques were then used to create new training examples for YOLOv5 to learn from, by generating augmented versions of each image in the training dataset. Figure 3.6 shows the dataset augmentations used in this dataset. The dataset augmentation techniques [31] used in this dataset were: 1) horizontal flip: to help YOLOv5 be insensitive to subject orientation; 2) crop: to add variability to positioning and size to help YOLOv5 be more resilient to subject translations and camera position; 3) brightness: to add variability to image brightness to help YOLOv5 be more resilient to lighting and camera setting changes; 4) blur: to add random Gaussian blur to help YOLOv5 be more resilient to camera focus; 5) noise: add noise to help YOLOv5 be more resilient to camera artifacts; 6) cutout: add cutout to help YOLOv5 be more resilient to object occlusion; 7) bounding box brightness: add variability to bounding box brightness to help YOLOv5 be more resilient to lighting on work zone boundary object (traffic cone). The dataset after augmentations had 1,200 images for training, 150 images for validation and 150 images for testing.

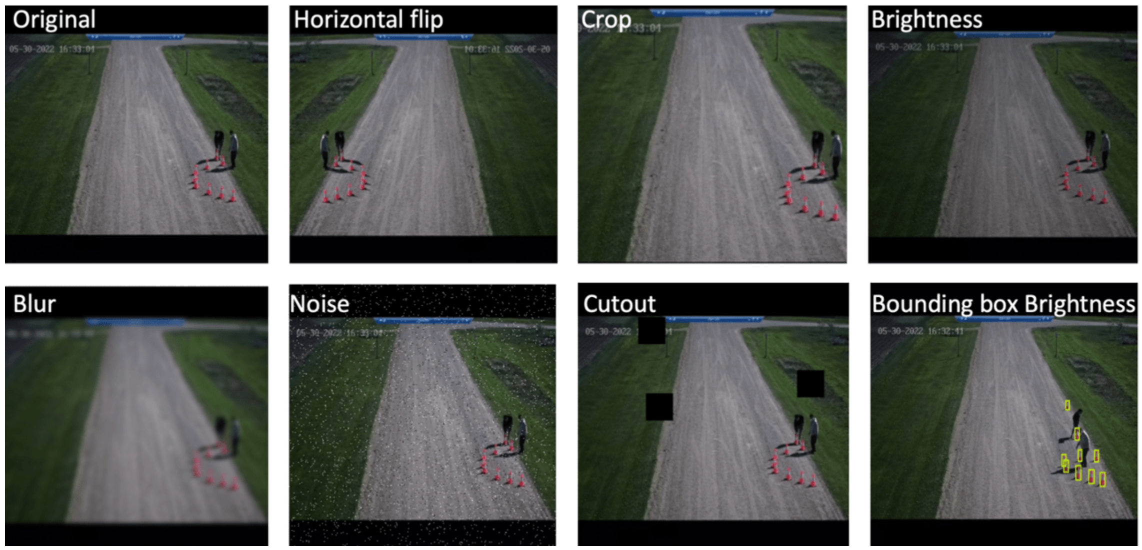


Figure 3.6: Dataset Augmentations Used in This Dataset.

The work zone boundary detection accuracy was measured using the following parameters and metrics: 1) True Positive (TP): YOLOv5 successfully recognized the presence of a traffic cone; 2) False Positive (FP): YOLOv5 incorrectly recognized the presence of a traffic cone; and 3) False Negative (FN): YOLOv5 failed to recognize the presence of a traffic cone. Using the above definitions, two key parameter values were calculated, and the following definitions were used to determine the parameter values: 1) precision given by Equation 3.3 is the fraction of correct recognition instances out of total recognitions; and 2) recall given by Equation 3.4 is the fraction of correct recognition instances retrieved over total expected recognitions. The training result regarding precision and recall is shown in Figure 3.7.

$$Precision = \frac{TP}{TP + FP} \tag{3.3}$$

$$Recall = \frac{TP}{TP + FN} \tag{3.4}$$

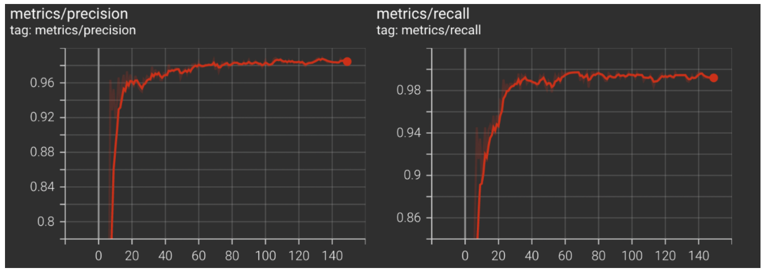
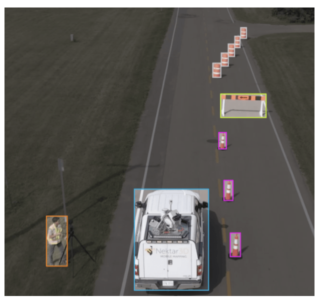


Figure 3.7: Training Result on YOLOv5 to Detect Traffic Cones.

The YOLOv5 used the 416×416 input images and reached 98% precision and 99% recall on our validation dataset. We trained the network from a checkpoint that is pre-trained in the MS-COCO dataset [64] using our training dataset. All the detection results indicated the overall result is above 90%, which is an adequate detection accuracy for any safety-critical work zone detection applications.

Phase Two

In phase two of the experiments, we repeated the training and testing procedure in phase one, but this time YOLO is trained on our improved dataset with labelling (Figure 3.8a) on not only small traffic cones, but also all common work zone objects like traffic barrels, traffic barricades, construction workers, as well as vehicles. We used drones to capture images of real-world work zone setup by ATS traffic shown in Figure 3.8b. The drones were controlled to fly over from the starting point to the ending point of the work zone so that the camera on the drones could see everything within the work zone. We only labelled objects that were not too small on the images since YOLO has difficulty detecting and distinguishing extremely small objects in images according to YOLO official release documents [31]. For example, we did not label some traffic cones that were too far away from the camera where its bounding box length or width is less than 3 pixels on a 640-pixel by 640-pixel image. The whole dataset was split into three parts including a training set, validation set, and test set. Then all images in the dataset were resized to 640 pixels by 640 pixels before some data augmentation techniques including horizontal flip and brightness variation introduced in phase one were applied. The dataset after augmentations had 1,600 images for training, 200 images for validation and 200 images for testing. Yolo was later trained on these data using 16 batches and 300 epochs with Stochastic Gradient Descent (SGD) optimizer of 0.01 learning rate, and took 2 hours to complete training. Yolo reached 98% precision and 95% recall on our validation dataset. Google Colab was used to train the model since we had driver issues on our lab computer with NVIDIA RTX3090. We utilized a Roboflow.ai [65] notebook that is based on YOLOv5 and employed pre-trained weights trained by the COCO dataset.



(a) Example of data labelling.

(b) Example of work zone image.

Figure 3.8: Dataset Images.

Figure 3.9 shows various performance measures for both the training and validation sets in each epoch during the whole training process and depicts three forms of loss: box loss, objectness loss, and classification loss. The box loss quantifies how effectively the algorithm can detect an item's center and how well the anticipated bounding box covers an object. Objectness loss is simply a measure of the likelihood of finding an item in a particular zone of interest. Classification loss indicates how successfully the algorithm predicts the proper class of an item. Around epoch 100, accuracy, recall, and mean average precision stop improving rapidly and the box, objectness, and classification losses evaluated in the validation set stop declining dramatically. We chose the best weights by the evaluation in the validation set. The confusion matrix in Figure 3.10 is also an important metric to measure performance, where the values in the diagonal show that the model predicted correctly while other values indicate that model made wrong predictions. We can see that the model made more wrong predictions in detecting traffic cones than other objects, which is expected since small object detection is one of the difficulties that YOLO has.

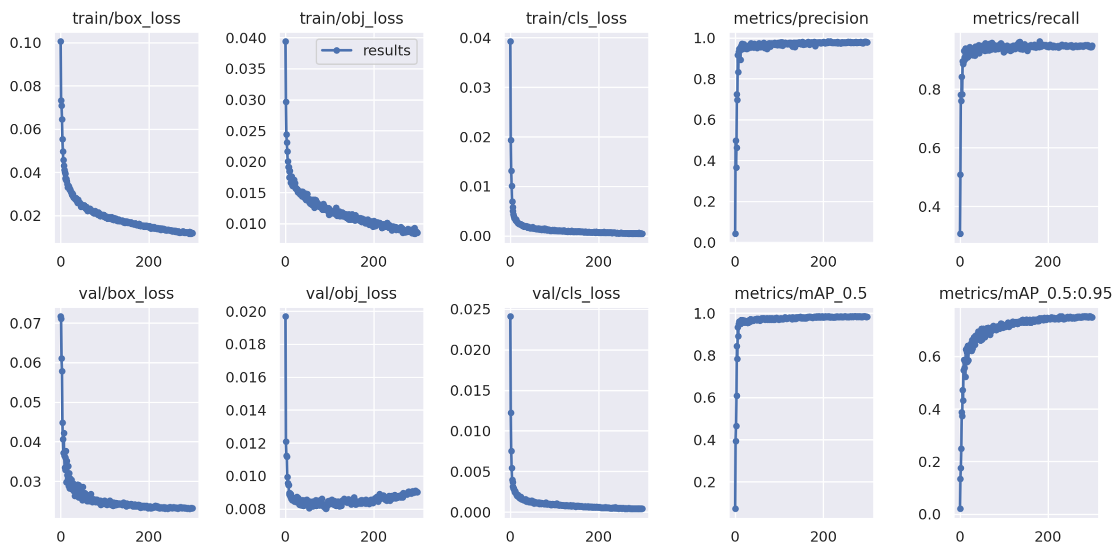


Figure 3.9: YOLO Training Results.

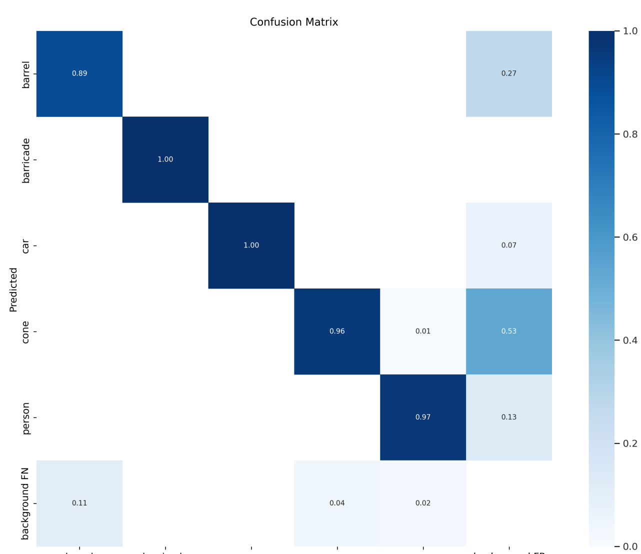
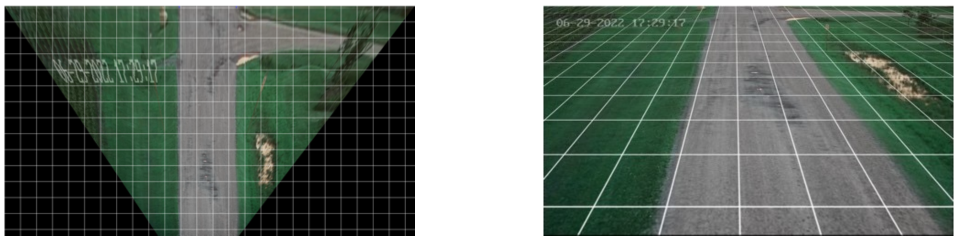


Figure 3.10: Confusion Matrix.

(a) The top-down view image obtained by PT.

(b) The ground truth top-down view image.

Figure 3.11: Global Distortion of PT.



(a) The top-down view image with grid lines.

(b) The image after performing inverse PT.

Figure 3.12: Local Distortion of PT.

3.3.2 Calibration of the Distortion in the Top-Down Image

This subsection is to check if the correct perspective transformation matrix was obtained. First, we compared our top-down view image obtained from PT with the satellite image at the same road as ground truth, to check the distortion globally in our top-down view image obtained from PT, illustrated in Figure 3.11. Then, we added grid lines to our top-down view image obtained from PT and performed an inverse PT to check the distortion locally, shown in Figure 3.12. The results reveal that the global and local distortion in the vertical road of the top-down view from PT is acceptable as there is no visible distortion in Figure 3.11 and Figure 3.12. If the perspective transformation matrix was incorrect, the road margin would not be vertical.

3.3.3 Calibration of the Linear Transformation Matrix

The accuracy of GPS sensors is important to establish high-precision localization from the image plane. However, a normal commercial GPS sensor is unsuitable for our use, as it has a position error of about 20m. In our experiment, we used the GNSS [66] sensor that has a position error within 1m from the OBU as GPS ground truth. The calibration procedure for calibration of the linear transformation (LT) matrix L in Equation 3.2 is as follows. We measured each point's GPS coordinates (longitude,latitude) with the GNSS sensor on the road, and this point's pixel coordinate (p_x,p_y). Then we performed PT to get the transformed pixel coordinate (p'_x,p'_y) from (p_x,p_y). We recorded each point's (p'_x,p'_y) and (longitude,latitude) to calculate LT matrix L by python using the function np.linalg.lstsq. Figure 3.18 shows example data. The calibration results are shown in Figure 3.13 where the blue points are GPS ground truth data and the orange points are GPS coordinates calculated from Equation 3.2 using LT matrix L.

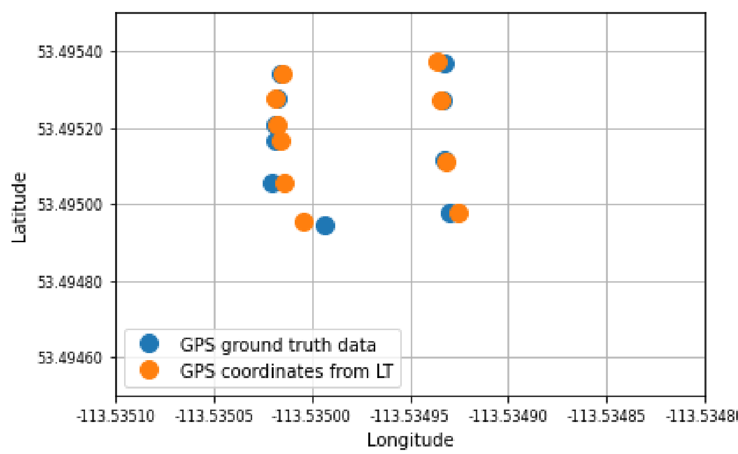


Figure 3.13: Calibration Results of LT.

3.3.4 Localization Error Analysis

The performance of a vision-based work zone alert system depends on the accurate localization of a work zone boundary from the traffic camera. The experiment setup is shown in Figure 3.14.

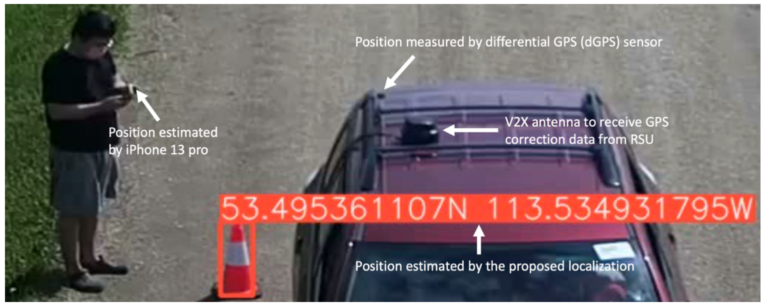


Figure 3.14: Experiment Setup to Measure Localization Error.

We used the OBU's GNSS data as GPS ground truth data in this experiment to calculate the localization error of our vision-based system. We also compared our localization performance with a GPS app on a cellphone that the work zone site manager would use to provide the work zone location.

On each section of the road, we placed a traffic cone and moved the OBU's GNSS sensor next to the traffic cone to obtain GPS ground truth data for 15 minutes under good reception conditions. From the vision-based work zone boundary localization method, we can obtain the GPS data estimated by the proposed localization. Our team member stood next to the traffic cone recording the data from the GPS app on a cell phone. We then calculated the root mean square error (RMSE) based on Equation 3.5. Here G_i is the GPS ground truth data, L_i is the location estimated by the proposed localization, and N is the number of data points.

$$RMSE = \sqrt{\sum_{i=1}^{N}(G_i - L_i)^2/N} \tag{3.5}$$

Testing points used in this experiment are green points shown in Figure 3.15. We calculated RMSE on each testing point and obtained localization error analysis results shown in Figure 3.16. RMSE is the localization performance evaluation metric in this experiment.

From Figure 3.16, we can see our proposed localization method outperformed the GPS app on the cellphone. Our proposed method had an average RMSE of 0.40m and

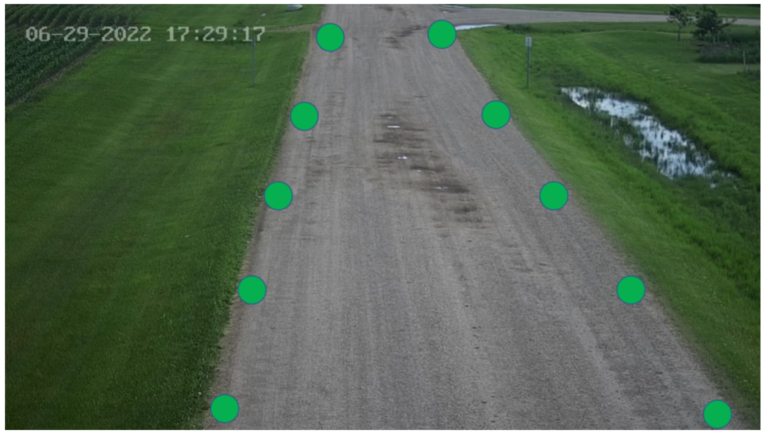
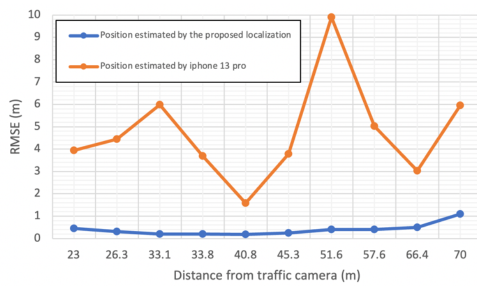
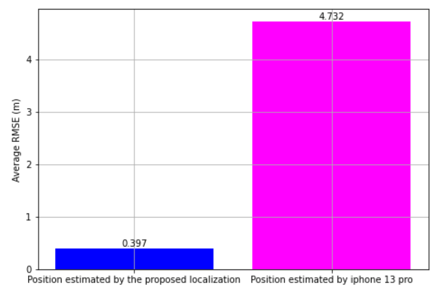


Figure 3.15: Localization Error Analysis Testing Points.



(a) The RMSE comparison between our method and iPhone 13pro on each testing point.



(b) The average RMSE comparison calculated from all testing points.

Figure 3.16: RMSE Comparison.

a maximum RMSE of 1.1m, which means a work zone boundary can be accurately located within half of a lane width (1.75m).

3.4 Alternative Localization Method

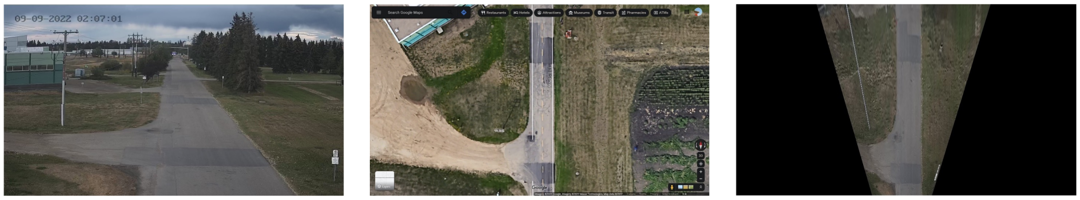
In this section, an alternative localization method was proposed and followed by a detailed localization error analysis with visualization on Google Maps.

3.4.1 Perspective Transformation and Linear Transformation Matrix Calculations

Google Maps was used to calculate the PT and LT matrices. First, the pixel coordinates of four pairs of matching points on the traffic camera image plane (Figure 3.17a) and the Google Maps satellite image plane (Figure 3.17b) were recorded. Then python was used to calculate the PT matrix. To solve the equation of the PT matrix, at least four pairs of matching points were required. More pairs of matching points would help us to reduce the distortions. The image plane after applying the PT matrix was shown in Figure 3.17c. Visually compared with Figure 3.4, where no ground truth top-down view image was used, calculating the PT matrix with the Google Maps satellite image as ground truth (Figure 3.17) did reduce some distortions, although distortions were more obvious at a far distance from the traffic camera.

The LT matrix was calculated by collecting the pixel coordinates on the Google Maps satellite image plane and the GPS coordinates from Google Maps of those matching points we chose when we solved the PT matrix. Again, at least four pairs of matching points were required to solve the LT matrix. The matching points used to calculate the LT matrix did not have to be the same matching points used to solve the PT matrix. To show an example, Figure 3.19 shows the pixel coordinate in the traffic camera image plane after PT and the GPS coordinates from Google Maps of a matching point (shown by a green circle point) used to calculate the LT matrix. The GPS coordinate was collected by placing the pointer on the matching point and right clicked the mouse. This localization method was based on the assumption that there were at least four distinguishable points that were not on the same line from the Google Maps satellite image. Otherwise, we would have to collect GPS data on site as we did in Chapter 3.3.3.

This alternative localization method had a few advantages. First, onsite GPS data collection was not necessary to calculate the LT matrix since we could collect GPS data directly from Google Maps. Second, this method was very low-cost and fast. It required no GPS or GNSS sensor, and the process of collecting data from Google Maps was much faster than collecting data using the GNSS sensor from OBU. It usually took half an hour to collect all the data needed to calculate the LT matrix, whereas we could normally finish data collection within ten minutes, including five



(a) The traffic camera image plane. (b) The Google Maps satellite image plane. (c) The traffic camera image plane after PT.

Figure 3.17: Process of Satellite Image Based PT

minutes of finding matching points plus five minutes of collecting GPS coordinates for these matching points, with the help of Google Maps. Third, since we used the Google Maps satellite image as a ground truth to calculate the PT matrix, the distortions were smaller, which would benefit the localization accuracy. This localization method is recommended to use on highways where onsite data collection is difficult or dangerous, and the roads have many lane markings.

	latitude	longitude
0	53.495209	-113.535019
1	53.495209	-113.535019
2	53.495209	-113.535019
3	53.495209	-113.535019
4	53.495209	-113.535019
...	...	...

(a) The pixel coordinate of a matching point in the top-down view. (b) The GPS coordinate of the same matching point from the GNSS sensor.

Figure 3.18: Example Data to Calculate LT (method 1).

(a) The pixel coordinate of a matching point in the top-down view. (b) The GPS coordinate of the same matching point from Google Maps.

Figure 3.19: Example Data to Calculate LT (method 2).

3.4.2 Localization Error Analysis with Visualizations

As we did the localization error analysis in Section 3.3.4, we chose eight points on the traffic camera image plane to be the testing points ranging from 30 meters to

80 meters from the traffic camera shown in Figure 3.20a. At the beginning of this experiment, we opened the live video stream from the traffic camera and clicked the record button so that localization error analysis could be conducted offline. Then we placed a traffic barrel on each testing point (Figure 3.20b) and our localization method would output the GPS coordinate of the traffic barrel. We also used the GNSS sensor on OBU to collect GPS coordinates of those testing points, in addition, to collecting GPS data from Google Maps, since we wanted to explore the difference between these data sources at the same point. By Equation 3.5, RMSE can be calculated. We calculated the RMSE between our method and GPS data from the GNSS sensor and also obtained the RMSE between our method and GPS data shown in Figure 3.21a. We found out that RMSE had a trend of increasing as the testing points were further away from the traffic camera, which indicated our localization method generally worked well at a relatively near distance and did not perform as well at a far distance. To be more specific, we can see that testing points that were within 40 meters of the traffic camera had a RMSE of less than 0.5 meters when comparing our method and GPS data from Google Maps. RMSE certainly increased on testing points that were 70 meters further away from the traffic camera. The average RMSE of all testing points is 0.9 meters using Google Maps as a ground truth and 1.8 meters using the GNSS sensor as a ground truth. We also calculated the localization noise produced from our localization method. The localization noise was defined by finding the area of distribution of all points generated by our localization method at each testing point. We defined the area of distribution as the area of the smallest circle that could cover all points from our method. The localization noise plot is shown in Figure 3.21b, showing all points had less than 0.5 square meters area of distribution, except for the furthest point that had 0.7 square meters. This experiment showed that longer distances from the traffic camera had some negative impacts on our localization method in terms of both RMSE and localization noise. Figure 3.22 shows a few visualizations by plotting all GPS data on Google Maps, where red dots

(a) Eight testing points for localization error analysis are shown in green circle points.

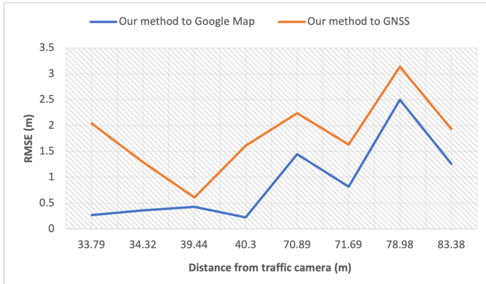
(b) This is one example of a traffic barrel placed on a testing point with YOLO detection.

Figure 3.20: Testing Points for Localization Error Analysis

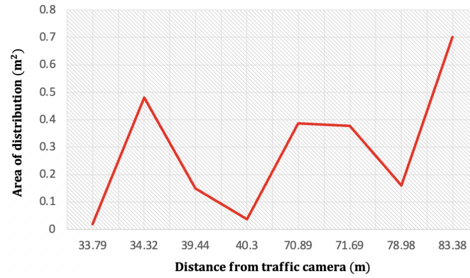
were from our localization method, blue dots were from Google Maps, and purple dots were from the GNSS sensor. We can clearly see the area of distribution from our method is larger at the furthest testing point than at other testing points. The area of distributions from the GNSS sensor was basically the same at different testing points. We also plotted all GPS data from all testing points in a scatter plot shown in Figure 3.23 and found out that the GPS data from the GNSS sensor generally had about 0.5 meters offset in the west direction compared with GPS data from Google Maps. Since the lane information, including which lane was closed by the work zone, was generated by a map matching algorithm with GPS data of the work zone and GPS data of the traffic lanes as an input, this localization method is recommended to use when the GPS data of the traffic lanes is collected from Google Maps instead of GNSS sensor. The lane information may not be very accurate when using the GPS data of the traffic lanes collected from the GNSS sensor. For example, the work zone warning app may display two traffic lanes closed by the work zone when there is actually one traffic lane closed by the work zone.

3.5 Field Testing

We conducted a case study on 118 street, NW, located in Edmonton, Alberta, Canada, to evaluate the performance of our vision-based work zone safety alert system in a CV environment.

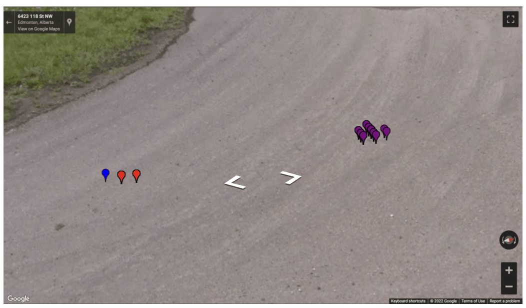


(a) The RMSE plot.

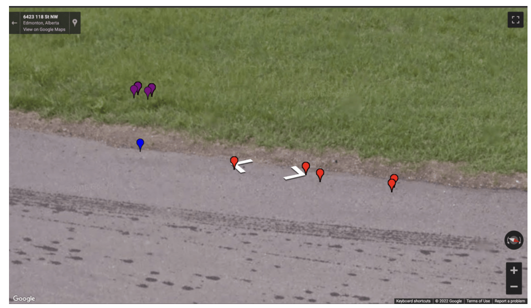


(b) Localization noise.

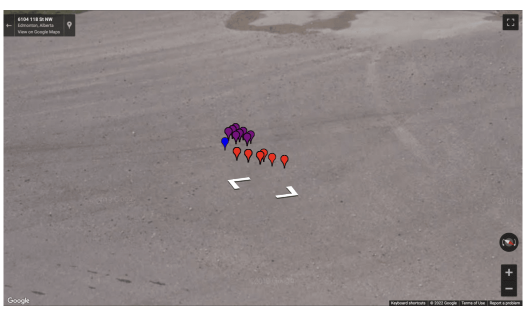
Figure 3.21: Results of Localization Analysis



(a) Visualization on the nearest testing point.



(b) Visualization on the furthest testing point.



(c) Visualization on the testing point that both RMSE values were less than 1 meter.

(d) Visualization on the testing point that both RMSE values were greater than 2 meters.

Figure 3.22: Visualization on Testing Points.

3.5.1 Experiment Setup

In this experiment, we used a video camera that connected to the central server with NVIDIA RTX3090 and 24GB video memory. The central server ran the work zone boundary detection algorithm and sent work zone starting point GPS coordinates, to the DSRC-enabled RSU. After receiving work zone static information from the

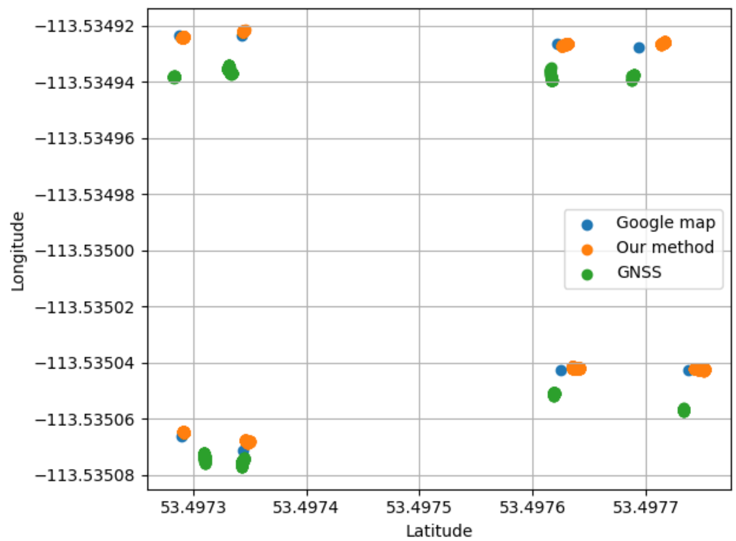


Figure 3.23: The Scatter Plot of All GPS Data.

central server, the DSRC-enabled RSU broadcast this information to the nearby CV equipped with a DSRC-enabled On-Board Unit (OBU). Then each CV within the range of DSRC-enabled RSU (400 meters) received work zone safety alerts.

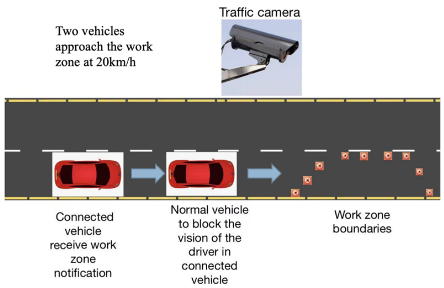
3.5.2 Experiment Design

In this experiment, we used a normal vehicle to block the CV, so that the driver in the CV could not see the work zone in front of the normal vehicle. The two vehicles are approaching the work zone. Figure 3.24 shows the experiment design. Since the driver in the CV cannot see the work zone, the work zone warning app plays a vital role to inform the driver of the position of the work zone.

3.5.3 Experiment Results

The work zone boundary localization results in this experiment are shown in Figure 3.25. From these results, we generated work zone safety alert messages to the CV driven through the work zone.

In the first experiment, we recorded the CV trajectory by OBU without the work zone warning app enabled. The CV trajectory, speed and work zone boundary loca-

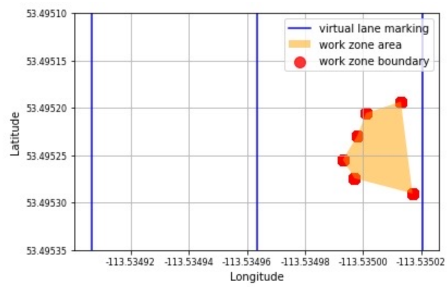


(a) Experiment Design and Planning

(b) The actual picture at the experiment site.

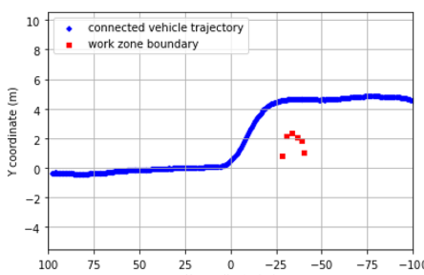
Figure 3.24: Experiment Design for Real World Testing.

(a) The work zone in the traffic camera view with virtual lane marking added from Google Maps

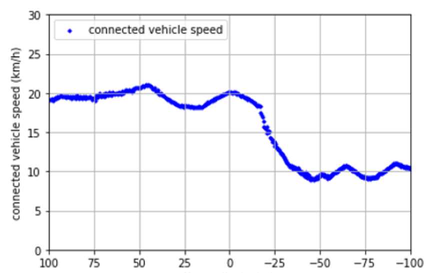


(b) The localization results show the precise location of the work zone.

Figure 3.25: Work Zone Localization



(a) How CV is driven through the work zone (from the positive x-axis to the negative x-axis) without the work zone warning app enabled.



(b) CV's speed before and after work zone as well as where CV slowed down.

Figure 3.26: CV's Behavior Without the Work Zone Warning app

tion estimated by the proposed localization are shown in Figure 3.26.

From Figure 3.26, we can see that the CV performed a lane change 25 meters in front of the work zone starting point. Changing lanes at such a close distance can cause potential danger to both the driver and the worker in the work zone. The lane-changing action of the CV depended on the front vehicle since the driver of the CV can only see the work zone when there is no vehicle in front of the CV.

In the second experiment, we repeated the same experiment but with the work zone warning app enabled. Since the traffic camera can see the work zone, the central server ran the work zone detection and localization algorithm, and then sent the work zone static information to the RSU. The RSU then broadcast the work zone static information through the DSRC channel to the OBU on the CV. The OBU then calculated the real-time distance between the CV and the work zone and displayed the distance on the work zone warning app through OBU's Wi-Fi.

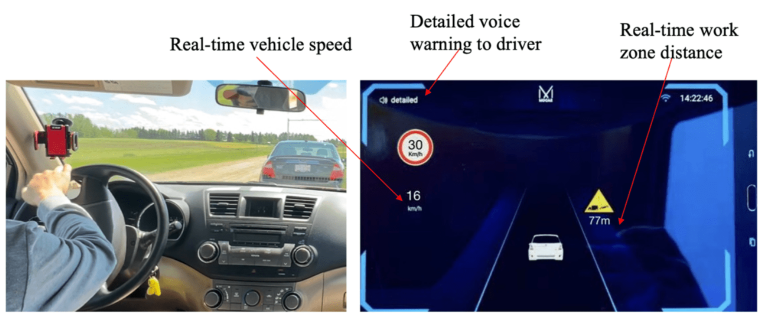
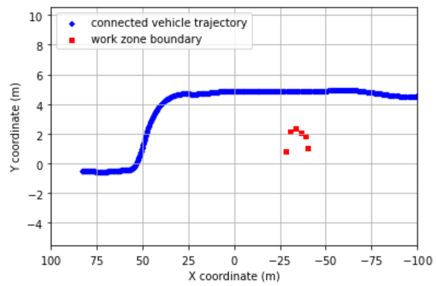


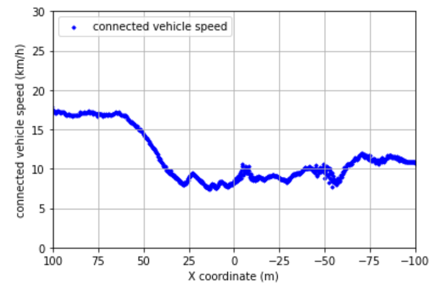
Figure 3.27: The Work Zone Warning app.

Figure 3.27 shows the moment when the driver performed a lane change based on the work zone warning app's safety alert. At this moment, the driver could not see the work zone, but still knew the precise location of the work zone with the help of the work zone warning app. In the work zone warning app, we set the work zone warning sound to trigger when the distance between the CV and the work zone starting point is less than 100m since the speed limit on this testing road is 40km/h. The driver can change the setting in the work zone warning app to trigger the warning sound at a farther distance (for example, 1000m for a highway application). The CV trajectory, speed and work zone boundary location estimated by the proposed localization are shown in Figure 3.28.

From Figure 3.28, we can see that the CV performed a lane change at 80 meters in



(a) How CV is driven through the work zone (from the positive x-axis to the negative x-axis) with the work zone warning app enabled.



(b) CV's speed before and after work zone as well as where CV slowed down.

Figure 3.28: CV's Behavior With the Work Zone Warning app

front of the work zone starting point. With the help of the work zone warning app, the driver in the CV changed lanes and slow down much earlier, which ensured the safety of both the driver and the workers in the work zone.

3.6 Effect of Different Weather Conditions and Low Resolution

To explore the effect of different weather conditions and low camera resolution, we used a 1280-pixel by 720-pixel test video (Figure 3.29a) where a construction worker was walking inside the work zone, placed at 50 meters from the traffic camera, as an input to YOLO and our localization method. Shown in Figure 3.29b, Figure 3.29d, Figure 3.29e and Figure 3.29f, negative 100% exposure and negative 100% brightness were applied to the test video before some weather effects were added by a video editing software and the underlying pixel integrity data was still there. This step was to test how well YOLO can detect work zone items under nighttime fog, rain, snow, and thunderstorm. There were several reasons we did not conduct this experiment under real-world conditions. For example, it is unsafe for both drivers and our team members to collect work zone data during a thunderstorm at night. We also tested YOLO's performance on an extremely low camera resolution setting

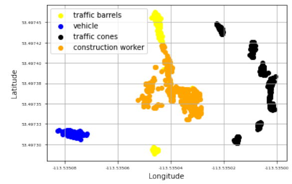
(254-pixel by 144-pixel) shown in Figure 3.29c. The localization results for different test videos are shown in Figure 3.30 by using the framework developed by Mike et al. [67]. It is obvious to see that localization noise increased compared with the original test video. The version to simulate nighttime rain had the largest impact on YOLO's detection, where YOLO missed 3 traffic cones and 1 vehicle completely, and missed the detection of the walking construction worker many times (Figure 3.30d). Looking into the annotation videos on nighttime fog, rain, snow, and thunderstorm, where each object was annotated on the videos when detected by YOLO, we found out that YOLO generally can not obtain a stable detection on small objects such as a traffic cone. Lowering the resolution of the test video did not seem to largely affect YOLO's performance. YOLO still could detect and locate all objects within lane level accuracy although noise certainly increased for small objects. This indicated our method can be applied to most low-resolution traffic cameras in Edmonton, where the live traffic camera only supports 600 pixels by 400 pixels video stream. Lowering the resolution of traffic cameras generally would increase the stability of live video streams, for instance, more stable video frames, and decrease the latency since most traffic networks in Edmonton do not support high-resolution video transmission.

(a) original (b) fog (c) low resolution

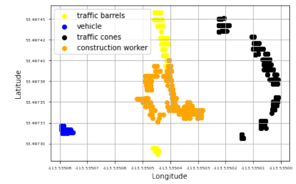
(d) rain (e) snow (f) thunderstorm

Figure 3.29: Test Videos from Different Weather Conditions and Camera Resolution.

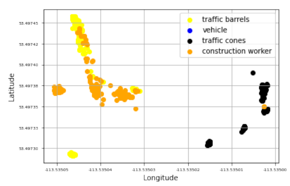
(a) original



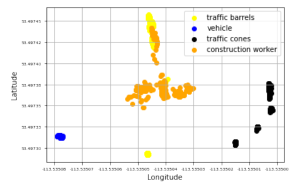
(b) fog



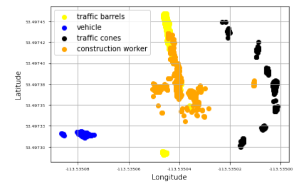
(c) low resolution



(d) rain



(e) snow



(f) thunderstorm

Figure 3.30: Tracking Results from Different Weather Conditions and Camera Resolution.

3.7 Lane Closure Information Broadcasting

We used standard map data message (MAP) and Traveler Information Message (TIM) defined by the SAE J2735 standard to represent the work zone lane geometry and a map-matching algorithm was used to determine the lane closure information based on the coordinate of each traffic cone, which will be explained in this section.

3.7.1 MAP and TIM Definitions

As described in SAE J2735, the necessary fields shown in Figure 3.31a are required to serve as the basic reference to other messages such as Signal Phase and Timing Message (SPAT) and TIM which will reference back to Intersection/lane IDs in the MAP message. By sending a MAP message, it is possible to broadcast the static work zone lane geometry to a CV nearby. The CV would be able to know which lane it is driving, but not know if the lane is open or closed.

SAE J2735 also defines traveller information message (TIM) which contains information related to road conditions. Figure 3.31b shows the essential fields for TIM. It can be used to broadcast work zone location with respect to the lanes defined in the MAP message. The work zone definition in TIM is similar to Work Zone Data

- Geographic description (multiple)
  - Intersection ID
  - Reference Point
  - Lane width
  - Number of lanes
  - List of lanes (multiple)
    - Lane ID
    - Node points
    - Connects to

(a) MAP Data Frame.

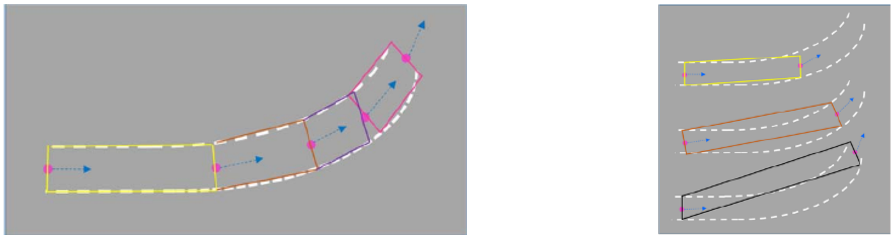
- Traveler dataframe
  - Frame type (advisory/road sign)
  - Start time
  - Duration
  - Geographical path (multiple)
    - Road ID
    - Lane width
    - Directionality
    - Closed
    - Path (XYZ/LLH offset system)
    - Geometry (projected circle)
  - Content
    - Advisory
    - Work zone
    - Generic sign
    - Speed limit

(b) TIM Data Frame.

Figure 3.31: MAP and TIM.

Exchange (WZDx) format. WZDx Specification [68] allows infrastructure owners and operators (IOOs) to make standardized work zone data available to third parties. The goal is to make public road traffic safer and more efficient by providing ubiquitous access to data on work zone activities. It is possible to choose a path manually/automatically in GPS coordinates and indicate road closure and generate the corresponding TIM to be broadcast by RSU. TIM messages can be broadcast simultaneously with the MAP message.

Shown in Figure 3.32, lane geometry can be defined by center line segments and road width based on SAE J2735 MAP message. To be more specific, the center line can be collected by: 1) a road survey; 2) GNSS sensor waypoints; 3) manually selecting points from Google Maps; 4) our localization method. Lane-level localization can be achieved using a map-matching algorithm by determining if the ego vehicle is located within any bounding box that belongs to a particular lane. The displacement between the bounding boxes affects the accuracy of lane geometry representation as depicted in Figure 3.32b.



(a) Lane geometry defined by center line segments and road width.

(b) Displacement between bounding boxes affects lane geometry accuracy.

Figure 3.32: Lane Geometry Represented by MAP.

3.7.2 Map Matching and Projecting

It is important to determine which lane is affected by road construction, and what is the exact starting and ending points. Using a map-matching algorithm to find the lane ID, and project the cone to the center line allows us to know the exact start/ending point of a work zone as shown in Figure 3.33. This is a naive approach assuming lane-level localization accuracy of traffic cone shown in Equation 3.6 to Equation 3.9, where $\mathrm{p}_{\mathrm{start}}$, and $\mathrm{p}_{\mathrm{end}}$, as well as lane width (l), are known information about lane geometry. So the purpose of these equations is to find if traffic cones are in any specific traffic lanes.

Shown in Figure 3.34, in TIM message the work zone geometry is represented as a series of points on the center lines projected by the cone's GPS coordinates (red). The starting point (green) of each work zone is an anchor point, and the remaining points (purple) are the offsets to the anchor point. For each segment, the minimum projection length ($\mathrm{d}_{0,1}$) is used for computing the starting point, and the maximum projection length ($\mathrm{d}_{1,2}$) is used for computing the ending point. By rendering the MAP and TIM message data, Figure 3.35 could be obtained showing corresponding legends.

$$\theta = arctan(\frac{p_{start_y} - p_{end_y}}{p_{end_x} - p_{start_x}}) \tag{3.6}$$

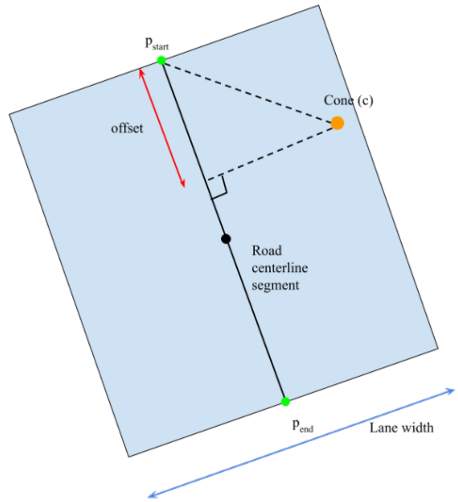


Figure 3.33: Map Matching.

$$\begin{bmatrix} p'_{start_x} & p'_{end_x} & c'_x \\ p'_{start_y} & p'_{end_x} & c'_y \end{bmatrix} = \begin{bmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{bmatrix} \begin{bmatrix} p_{start_x} & p_{end_x} & c_x \\ p_{start_y} & p_{end_y} & c_y \end{bmatrix} \tag{3.7}$$

$$\text{Match iff. } p'_{start_x} < c'_x < p'_{end_x} \tag{3.8}$$

$$\text{AND } -\frac{l}{2} < c'_y < \frac{l}{2} \tag{3.9}$$

3.7.3 Field Testing

We conducted field tests in a CV environment using C-V2X communication technology. We used an OBU to function as a RSU to broadcast MAP and TIM regarding work zone and lane information because the RSU could not be installed at that time. Figure A.1 and Figure A.2 in Appendix A show the communication topology and devices used in this field test respectively. Currently, the map shown in Figure A.3 in Appendix A is created using manually chosen points from OpenStreetMap [69] for simplicity. This method cannot be used for generating long-distance work accurately. The resulting GPS coordinates of each node in Extensible Markup Language (XML) format are used to generate MAP messages. MAP and TIM were broadcast at 10 hertz to nearby CVs. We developed a prototype of a work zone warning app with lane information running on a laptop placed in a CV to serve as an HMI for the

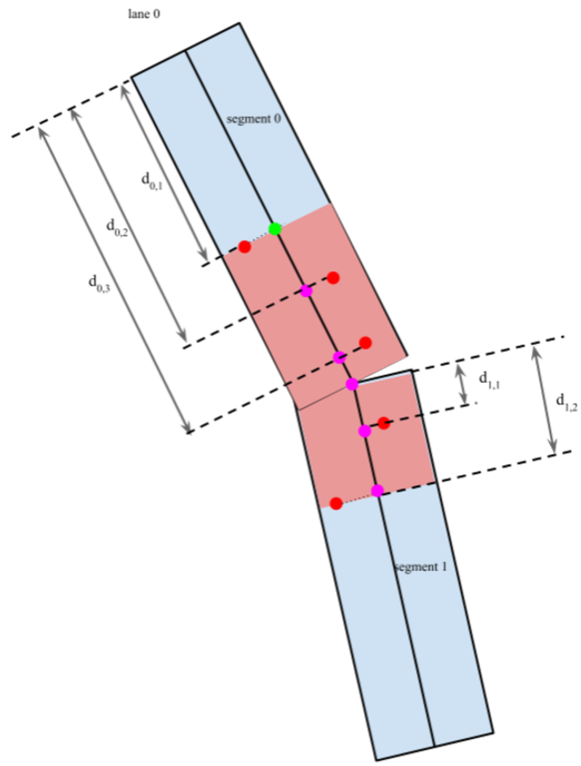


Figure 3.34: TIM Work Zone.

driver. This work zone warning app could show speed limit information, traffic lanes information and the work zone location as well as which lane was closed by the work zone. We set the speed limit from 40km/h to 10km/h when the work zone was less than 100 meters from the driver. We recorded the data, and the results are shown in Figure A.4, Figure A.5 and Figure A.6 in Appendix A, where we can see YOLO detection from the traffic camera, MAP and TIM visualization, the work zone warning app, as well as the driver's vision. Since one of the strong features in our system is real-time detection, localization, and warning, we also recorded results when removing traffic cones from the road. The results are shown in Figure A.7 and Figure A.8 in Appendix A. We can clearly see that the work zone region shown in MAP and TIM were decreasing and the work zone warning app would not give any warning when all traffic cones were removed. We repeated the experiments a few times to test stability. The mobile version of the work zone warning app is being developed now.

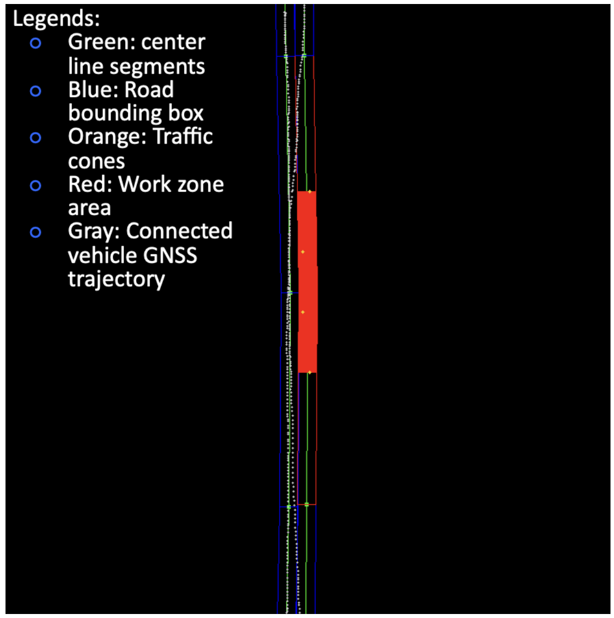


Figure 3.35: Render MAP and TIM.

3.7.4 System Level End-to-End Latency Tests

The data flow in this work zone warning system is shown in Table 3.1. The latency includes computational delay and communication delay. The computation delay depends on the hardware specification of the central server used to run the work zone detection and localization algorithm. For our experimental setup, we used a workstation with NVIDIA RTX3090 having 24GB GPU memory to run the work zone detection and localization algorithm. The communication network delay is the time difference between sending a message from one device and receiving the same message on another device. The communication latency (except for C-V2X) was measured by Packet Internet or Inter-Network Groper (PING) function to determine the Round-Trip Time (RTT) of a message. Each latency test was repeated 9 times and 100 packages were used to calculate RTT in each latency test to reduce error and improve accuracy. The tests used Cellular Vehicle to Everything (C-V2X) RSU and OBU to broadcast work zone information. The communication latency on C-V2X was measured by the time difference between the same message (sent and received) recorded by system log files on 2 OBUs close to each other (about 50m), and the test was repeated 9 times and 300 packages were used in each test. This assumes that V2I has approximately the same latency as V2V. The computation latency was measured by processing time on each image frame and 943 image frames were used. The actual communication latency between the traffic camera and the server would be higher than 2.1 microseconds since the PING function used a package size of only 64 bytes. A more suitable method will be designed to measure the actual latency between the traffic camera and the server in future research.

Table 3.1 shows the end-to-end latency results in every part of the work zone alert system and the total latency combines both computational and communication latency. The overall latency is below 100ms, which satisfies the latency requirement for safety-critical applications.

End To End Latency	**Latency Type**	**Average Latency**	**Standard Deviation**
Traffic camera sending live video data to central server	Communication latency on wired network	2.1ms	0.3ms
Work zone detection and localization algorithm results generation	Computation latency on central server	5.4ms	1.2ms
Central server sending work zone information to RSU	Communication latency on wired network	1.5ms	0.2ms
RSU broadcasting work zone information to OBU	Communication latency on C-V2X	14.0ms	4.3ms
OBU sending vehicle's speed and work zone information to Samsung pad running work zone warning app	Communication latency on OBU's Wi-Fi	17.7ms	6.6ms
Overall latency	Overall latency	40.7ms	12.6ms

Table 3.1: End-to-End Latency Test Results.

3.8 Conclusions

The goal of this chapter is to create a real-time vision-based work zone recognition and localization approach, that will increase both the driver and worker's safety in work zones. By broadcasting the safety alert (MAP and TIM) in real-time (every 100ms) from RSUs to CVs within its communication range, our work zone safety messages were used to alert the driver. When compared to commercially available smartphones that the work zone site manager utilized to give work zone location, our localization error study demonstrates that the vision-based work zone localization approach can estimate the position more precisely in terms of RMSE. The results from localization analysis revealed that we can accurately locate the work zone (placed within 80 meters of the traffic camera) within a lane. Furthermore, we tested the work zone alert system in a scenario where the driver cannot see the work zone. The work zone warning app generates safety alerts based on the potential forward collision risks between the work zone and the CVs. The results from the CV trajectory and work zone location demonstrate that the work zone warning app can inform the driver of the work zone location even if the driver cannot see the work zone and let the driver slow down much earlier. This cooperative perception ensures the safety of both the worker and the driver. The overall work zone alert system latency is below 100ms, which satisfies the latency requirement in a CV environment. The effect of different weather conditions and camera resolutions was also studied. The results indicated our localization method can be applied to most low-resolution traffic cameras in Edmonton.

References

[1] N. H. T. S. Administration *et al.*, "2015 motor vehicle crashes: Overview," *Traffic safety facts: research note*, vol. 2016, pp. 1–9, 2016.

[7] S. D. Schrock, G. L. Ullman, A. S. Cothron, E. Kraus, and A. P. Voigt, "An analysis of fatal work zone crashes in texas," *Report FHW A/TX-05/0-4028*, vol. 1, 2004.

[8] J. Daniel, K. Dixon, and D. Jared, "Analysis of fatal crashes in georgia work zones," *Transportation Research Record*, vol. 1715, no. 1, pp. 18–23, 2000.

[9] N. J. Garber and M. Zhao, "Distribution and characteristics of crashes at different work zone locations in virginia," *Transportation Research Record*, vol. 1794, no. 1, pp. 19–25, 2002.

[11] A. Dehman and B. Farooq, "Are work zones and connected automated vehicles ready for a harmonious coexistence? a scoping review and research agenda," *Transportation research part C: emerging technologies*, vol. 133, p. 103 422, 2021.

[12] X. Wu *et al.*, "Vehicular communications using dsrc: Challenges, enhancements, and evolution," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 399–408, 2013.

[13] W. Han, E. White, M. Mollenhauer, and N. Roofigari-Esfahan, "A connected work zone hazard detection system for roadway construction workers," in *Computing in Civil Engineering 2019: Smart Cities, Sustainability, and Resilience*, American Society of Civil Engineers Reston, VA, 2019, pp. 242–250.

[14] R. Schönrock, F. Wolf, and T. Russ, "Smart traffic cone-dynamic detection and localization of traffic disruptions," in *FAST-zero'15: 3rd International Symposium on Future Active Safety Technology Toward zero traffic accidents, 2015*, 2015.

[31] G. Jocher, *Yolov5. code repository https://www. github. com/ultralytics/yolov5*, 2022.

[40] M. Islam, M. Rahman, M. Chowdhury, G. Comert, E. D. Sood, and A. Apon, "Vision-based personal safety messages (psms) generation for connected vehicles," *IEEE transactions on vehicular technology*, vol. 69, no. 9, pp. 9402–9416, 2020.

[41] V. V. A. T. Committee, *Vulnerable road user safety message minimum performance requirements*, Mar. 2017.

[57] E. Radwan, R. Harb, S. Ramasamy, *et al.*, "Evaluation of safety and operational effectiveness of dynamic lane merge system in florida: Final report, april 2009.," 2009.

[58] Z. A. Nemeth and D. J. Migletz, "Accident characteristics before, during, and after safety upgrading projects on ohio's rural interstate system," *Transportation Research Record*, vol. 672, pp. 19–23, 1978.

[59] J. Hall and V. Lorenz, "Characteristics of construction-zone accidents," *Transportation Research Record*, vol. 1230, pp. 20–27, 1989.

[60] N. M. Rouphail, Z. S. Yang, and J. Fazio, "Comparative study of short-and long-term urban freeway work zones," *Transportation Research Record*, vol. 1163, pp. 4–14, 1988.

[61] O. Raddaoui, M. M. Ahmed, and S. M. Gaweesh, "Assessment of the effectiveness of connected vehicle weather and work zone warnings in improving truck driver safety," *IATSS research*, vol. 44, no. 3, pp. 230–237, 2020.

[62] R. F. Benekohal, E. Shim, and P. T. Resende, "Truck drivers' concerns in work zones: Travel characteristics and accident experiences," *Transportation Research Record*, vol. 1509, pp. 55–64, 1995.

[63] T. Wen, Z. Xiao, K. Jiang, M. Yang, K. Li, and D. Yang, "High precision target positioning method for rsu in cooperative perception," in *2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP)*, IEEE, 2019, pp. 1–6.

[64] T.-Y. Lin *et al.*, "Microsoft coco: Common objects in context," in *European conference on computer vision*, Springer, 2014, pp. 740–755.

[65] *How to train yolov5 on custom objects*. [Online]. Available: https://colab.research.google.com/drive/1gDZ2xcTOgR39tGGs-EZ6i3RTs16wmzZQ.

[66] T. U. goverment. "Other global navigation satellite systems (gnss)." (2021), [Online]. Available: https://www.gps.gov/systems/gnss/ (visited on 09/12/2022).

[67] M. Brostrom, *A collection of real-time multi-camera multi-object sota trackers using yolov5*. [Online]. Available: https://github.com/mikel-brostrom/Yolov5_StrongSORT_OSNet.

[68] *Usdot-jpo-ode/wzdx: The work zone data exchange (wzdx) specification aims to make harmonized work zone data provided by infrastructure owners and operators (ioos) available for third party use, making travel on public roads safer and more efficient through ubiquitous access to data on work zone activity.* [Online]. Available: https://github.com/usdot-jpo-ode/wzdx.

[69] [Online]. Available: https://www.openstreetmap.org/#map=15/53.4961/-113.5359.

Chapter 4

Vision-Based Road Surface Conditions Classifications with Applications in a Connected-Vehicle Environment

4.1 Introduction

Drivers frequently face risks due to poor road conditions. Driving is difficult and dangerous when there is snow or ice on the roadways. Winter weather is known to be one of the main contributors to an increased likelihood of crashes [70, 71] because of things like reduced friction on the road surface [10]. Additionally, if the road surface has not been cleaned promptly and adequately after a period of snowfall, the chance of death also rises. Every winter, Canada engages with road repair operations on thousands of kilometres of urban and rural highways. Despite the fact that Canada is a country with expertise in keeping its roads operational throughout the winter, there is still space for development in terms of road safety. According to Alberta Transportation [2–6], slush, snow, or ice was involved in 214 fatal collisions and 14,505 non-fatal injury collisions, causing 14,719 (27% of total) casualty collisions in total from 2015 to 2019. The number of collisions caused by slush, snow, or ice on the road each year is shown in Figure 4.1, and we can see a trend of more collisions occurring over the years.

For those who maintain winter roads as well as the public, monitoring the state of roads is crucial. Multiple efforts must be arranged and coordinated by transportation offices to maintain roads as free of snow and ice as possible, so that cars may utilize the road network safely. Road Surface Condition (RSC) is a metric often used by transportation agencies to determine the present status of the road in terms of snow or ice covering, as well as a communication mechanism. Winter maintenance staff can deliver the appropriate types of maintenance treatments and quantities of deicing products at the appropriate times thanks to real-time, trustworthy road surface condition (RSC) data, which results in considerable cost and salt savings. RSC monitoring has generally been carried out either manually by highway agencies and maintenance providers, or by employing RWIS stations [72]. Very high geographical resolution and extra qualitative information are provided by manual patrolling, but it has the shortcomings of being arbitrary, labour-intensive, and time-consuming. RWIS stations, on the other hand, offer continuous information on a variety of weather and road conditions, but they are expensive and can only be put in a small number of places, limiting their spatial coverage.

Emerging research has been focusing on using machine learning algorithms to automatically classify and monitor RSC. A subset of Machine Learning (ML) methods called Deep Learning (DL) approaches was created by combining sophisticated algorithms with mathematics [73]. Computer vision is one of the areas where DL approaches have shown impressive results, demonstrating cutting-edge accuracy in tasks like picture classification, object identification, and semantic segmentation. A few studies have recently assessed the application of DL to automatically classify photos with the goal of estimating RSC during the winter, with astounding results over RSC images from traffic cameras and dash cameras [15–18]. Less research has, however, examined DL techniques for calculating RSC from pictures captured by in-vehicle cameras in Alberta. The RSC monitoring process has recently been automated thanks to the development of new technologies including CCTV cameras, in-car video recorders, smartphone-based systems, and high-end imaging systems with DL. However, it has been discovered that the operating conditions and classification precision of these systems are still constrained.

The remainder of this chapter is organized as follows. We begin by discussing the structures and details of ML algorithms selected for RSC classification. Then the data collection, as well as the data labelling process, are detailed, where we collected and labelled over 15,000 images collected in Alberta. The assessment based on classification accuracy and processing time was performed after. The best DL model was then selected to be used in applications of automated RSC classification in a CV environment, with validation of the current RSC methods.

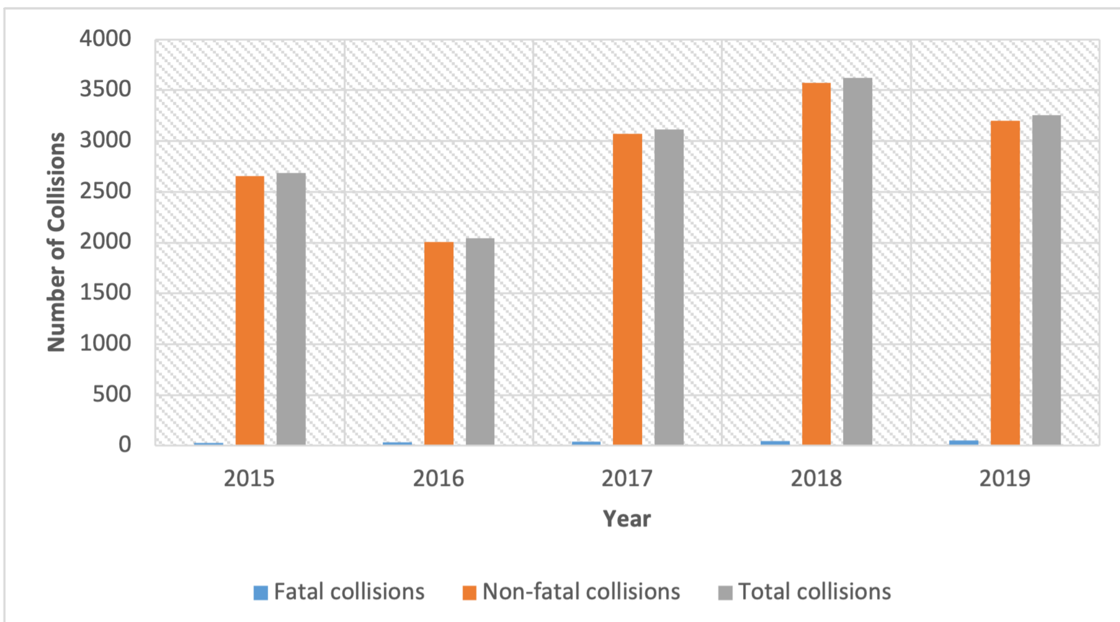


Figure 4.1: Number of Collisions Caused by Slush, Snow, or Ice [15–18]

4.2 Model Selection

A subset of Machine Learning (ML) methods called Deep Learning (DL) approaches was created by combining sophisticated algorithms with mathematics [73]. The recent availability of powerful Graphic Processing Units (GPUs) to train DL models, is also credited with the spectacular successes of DL methods across a variety of applications. Computer vision is one of the areas where DL approaches have shown impressive results, demonstrating cutting-edge accuracy in tasks like picture classification, object detection, and semantic segmentation.

One of the deep learning models that have proven particularly effective for picture categorization is CNN. As the input image progresses through the model's layers, a vector of probabilities is produced, with the highest probability corresponding to the RSC category that the image most possibly belongs to.

The numerous layers that make up DL models each take the output of the layer before them as input. In our scenario, we anticipate the DL models to receive photos from cameras installed on vehicles as input and output the corresponding RSC category. The objective of DL models is to make the model generate a mapping from input observations into the desired output.

We took three DL models into consideration, which Xception [74] and VGG [75] demonstrated state-of-the-art performance over the ImageNet [45] image classification benchmark, and Juan's CNN model [15] achieved over 90% accuracy on the RSC classification problems on their dataset captured by traffic cameras in Ontario. Then we compared their performance for RSC classification over the dash camera photos dataset. The ImageNet dataset, which contains pictures of common items, was previously used to train Xception and VGG models. Incorporating Xception and VGG models into the RSC image classification problem is the objective. The assumption is that Xception and VGG models that were initially trained on massive datasets have learned to recognize fundamental patterns that are helpful for classifying things across many domains. The various architectures employed in this research are briefly described in this section.

An alternative strategy called transfer learning would be to employ a CNN model that has previously been trained with demonstrated performance rather than developing a brand-new CNN model, which frequently involves a substantial quantity of data and computational time [76, 77]. The usage of such characteristics would enable us to achieve higher accuracy than any method that would solely rely on the available data. Such a model would have already learned features that are relevant to most computer vision challenges. We applied transfer learning on Xception and VGG since their authors posted the pre-trained weights, while Juan's CNN model was trained from scratch because the pre-trained weight was not released.

Francois Chollet [74] introduced the Xception network, which replaced Inception modules with depth-wise separable convolutions and residual connections. The data flows through an entering flow, a middle flow, and ultimately an exit flow in the Xception architecture. There are eight iterations of this technique. The top layers were removed and replaced with a dropout layer, and the fully connected layer was replaced with a softmax layer for the three different classes of road conditions in order to modify this network for our goal.

VGG, which was created by Simonyan and Zisserman [75], came in second in the 2014 ILSVRC (ImageNet Large Scale Visual Recognition Competition). It was one of the first networks to demonstrate the effectiveness of combining a highly layered architecture with small convolutional filters. Deep feed-forward architecture with no residual connections characterizes VGG. This is made up of two completely connected layers at the conclusion of two linearly connected convolutional layers with max-pooling after the second or third layer. For the RSC problems, the top layers were removed and swapped out for a dropout layer, and the fully connected layer was switched out for a softmax layer.

A simple CNN model created by Juan Carrillo [15] was influenced by important DL research like AlexNet [78] and VGG [75]. When compared to the features of the other model Xception, this model has fewer layers and fewer parameters, including layers for data reduction (max-pooling), model regularization (dropout), and vector concatenation.

4.3 Data Collections

Data was collected using video cameras at the front of the leading and following trucks. The route of the trials was between Edmonton and Calgary, and 21 trials were conducted in total as shown in Figure 4.2. The trials started on September 12, 2021, and ended on January 30, 2022. The trucks travelled around 23,115 km. 9 drivers participated in these trials. More than 585 videos were taken for the lead truck, and over 570 videos were taken for the follower truck. Each video is approximately 20 minutes long.

The three-class descriptions shown in Figure 4.3 are used to manually label each image as ground truth, and automatic image processing is made to classify each image in accordance with those descriptions. The Transportation Association of Canada's route reporting nomenclature, which is commonly used to notify the public about RSC information about maintenance routes, is used to group RSCs into three categories in the three-class description. All collected images (15,855 in total) were manually labelled into three categories; bare pavement (5,391 or 34%), partially snow-covered (5,011 or 31%) and fully snow-covered (5,453 or 34%). Examples of each category are shown in Figure 4.3.

4.4 Training and Testing Deep Learning Models

90% of the total number of images are used to train and evaluate each of the three models. The remaining 10% is kept as a test set, and we only utilized it at the very end to report accuracy. 20% of the training data are taken for validation and the charting of the accuracy and loss functions. so we had 11,415 images for training, 2,854 images for validation, and 1,586 images for testing. Dataset split was done randomly to avoid bias. To reduce the classification error, we employ the Backpropagation Algorithm with Stochastic Gradient Descent (SGD) as an optimizer. The learning rate is maintained at 0.001. Other settings for the SGD optimizer, including

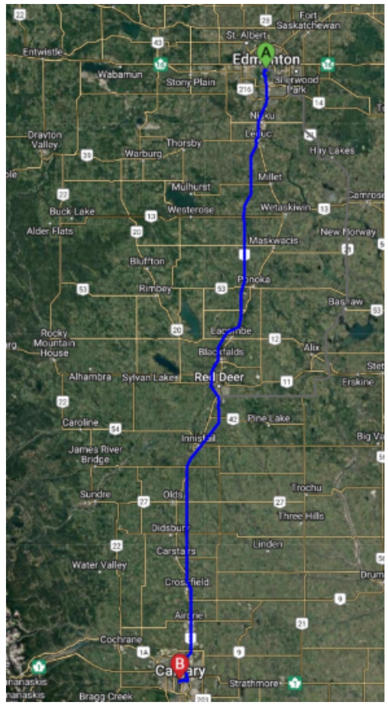


Figure 4.2: Locations of Data Collection.

(a) Bare pavement. (b) Part snow coverage. (c) Full snow coverage.

Figure 4.3: Definitions of Different Types of RSC

momentum = 0.9 and Nesterov momentum enabled, are made in accordance with suggested values in the literature [15, 74, 75].

The training set is processed through each DL model 10 times (epochs) in order to minimize the classification error. Additionally, the images are fed in batches of 16 to save memory. With the exception of the model created by Juan [15], which needs to be trained entirely from scratch, Xception and VGG need to be trained by applying transfer learning. Transfer learning entails applying features learned on one problem to a new, similar problem. Transfer learning is typically used for tasks where the training dataset contains insufficient data to train a full-scale model from scratch. In our case, we have enough RSC data (about 15,000 images) to train Juan's CNN since it is a relatively small model with much fewer parameters, but we may not have enough RSC data to train Xception and VGG since they are much more complex and large model with ten times more parameters than Juan's CNN. We decided to use pre-trained weights to train Xception and VGG by transfer learning since we assumed if we trained Xception and VGG from scratch, the classification accuracy would not be higher than 50%. The following workflow is the most common incarnation of transfer learning in the context of deep learning: 1) Take layers from an earlier trained model; 2) Freeze them to prevent any of the information they contain from being destroyed during future training rounds; 3) On top of the frozen layers, add some new, trainable layers. They'll figure out how to turn old features into predictions on a new dataset; 4) Run the new layers through their paces on our dataset.

Data collection, labelling, and experimentation were all important parts of this project. One of the most difficult aspects of this work was labelling millions of raw images of road and weather conditions in a variety of scenery (urban, rural), sky conditions (clear, overcast), illumination (day, night, twilight), and quality to produce a reliable set of training images were to have about the same numbers of samples for each type of road condition, including bare pavement, part snow coverage and full snow coverage to ensure there is no bias in the dataset. Another challenge was to account for model complexity and memory usage, as well as classification accuracy, at each stage of the dataset labelling and classification process.

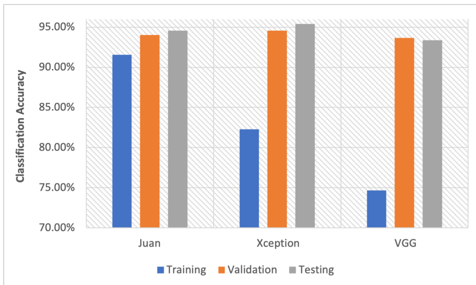
In this work, the following metrics were used in Equation 4.1 and Equation 4.2. In Equation 4.1, TP stands for True Positives, FP stands for False Positives, TN stands for True Negatives and FN stands for False Negatives. For example, the false positive rate is defined as the proportion of cases where road surfaces are covered in snow/ice but are classified as bare pavement. We used categorical cross entropy [79–81] as the loss function for our RSC classification problem shown in Equation 4.2, where M is the number of classes, the log is the natural log, and y is the binary indicator (0 or 1) if class label c is the correct classification for observation o, and p is the predicted probability observation o is of class c. We desired to have a higher accuracy value and a lower loss function value for a better training result. Over the course of model training, we expected an increasing accuracy value and a decreasing loss function value.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.1}$$

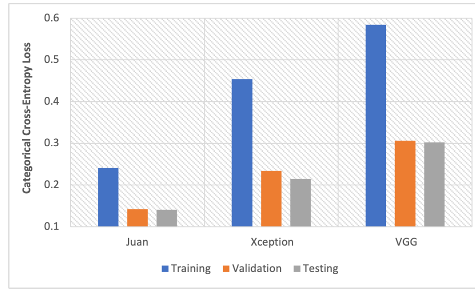
$$Loss = -\sum_{c=1}^{M} y_{o,c} \log(p_{o,c}) \tag{4.2}$$

Juan's CNN model, Xception and VGG are all trained and validated on 14,269 images (80% for training and 20% for validation) where 33% are bare pavement, 31% are part snow coverage and 34% are full snow coverage, and then tested on 1,586 images where 34% are bare pavement, 31% are part snow coverage and 34% are full snow coverage. All three categories had about the same proportions in training, validation and testing datasets. Figure 4.4 shows the summary of training results for the models when transfer learning was applied on Xception and VGG, except for Juan's CNN model was trained from scratch. We expected training accuracy to be higher than validation accuracy and testing accuracy. But our results showed validation accuracy was higher than training accuracy, which was unusual. Testing accuracy was about the same as validation accuracy, which was reasonable. To be sure there was no overlap between the training dataset and the testing dataset, another training was done using only 30% of total images for training, another 30% of total images for validation and another 30% of total images for testing on Juan's CNN. Training accuracy was 86.23% and validation accuracy was 93.63% and testing accuracy was 92.89% after 20 epochs. Still, validation accuracy was higher than training accuracy. This unusual behaviour can be caused by dropout layers in the CNN. Dropout layers are used to avoid overfitting. Some neural nodes were randomly turned off during the training process but all neural nodes were turned on during the validation and testing process. So CNN can have better validation accuracy. Another possible explanation is the CNN was underfitting after the training. To investigate this unusual behaviour, dropout layers were removed before all CNNs went through the same training process again. The results are shown in Figure 4.5. Training accuracy is higher than validation accuracy, which is reasonable and usual. Also, we achieved better accuracy without dropout layers. Based on our experiments, the classifiers were able to distinguish between images of bare pavement, part snow coverage, and full snow coverage. The main significance of these findings is that they demonstrated that we can use modern CNN frameworks to achieve results of over 90% accuracy. Another interesting fact is that these results were obtained with no pre-processing on the images other than re-scaling (all images in this dataset were re-scaled to 224-pixel by 224-pixel size to save memory and maintain features).

We also recorded the time to train each DL model and the processing time for each DL model to classify one 224-pixel by 224-pixel size RSC image on a commercially available CPU (11th Gen Intel i7-11700K) as shown in Figure 4.6. It is very obvious that Juan's CNN model took significantly less time to train while achieving competitive accuracy. We chose Juan's CNN model for our automated RSC classification application in the CV environment because Juan's CNN model took only 16 microseconds to process one image while others took over 50 microseconds. The reason we recorded processing time on CPU is that all laptops have CPUs, but not all of them have GPUs that can process DL models much faster than CPUs. We would like to use any laptop that can be placed on CVs to run the RSC classification algorithm and send RSC reports through V2V and V2I communication technologies by DSRC in real-time. We could also stream RSC videos recorded by dash cameras to our lab computers with GPUs to generate RSC reports with the RSC classification algorithm in real-time.

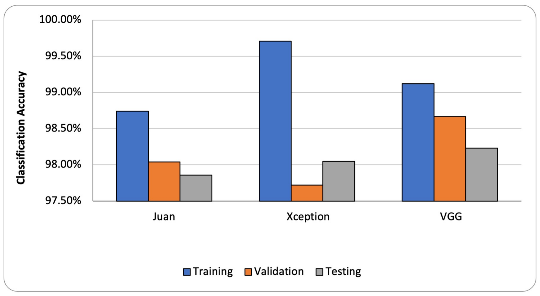


(a) Models accuracy summary.

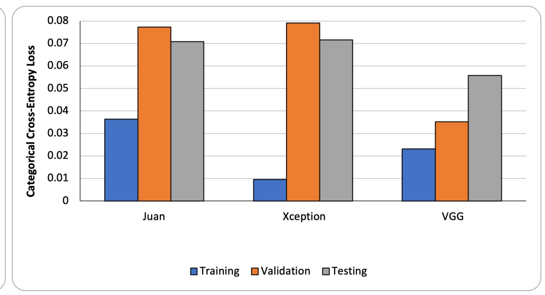


(b) Models loss summary.

Figure 4.4: Training Results Summary with Dropout Layers

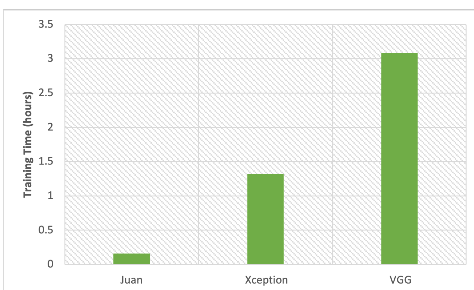


(a) Models accuracy summary.

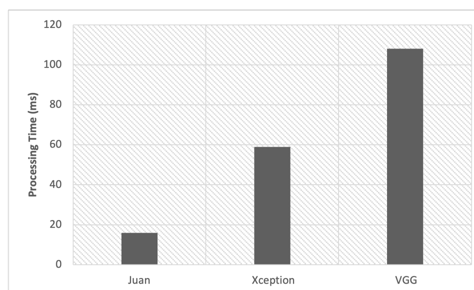


(b) Models loss summary.

Figure 4.5: Training Summary without Dropout Layers



(a) Models training time.



(b) Models processing time.

Figure 4.6: Training and Processing Time Summary

4.5 Applications in a CV Environment

In a CV environment, any CV with a camera and laptop or smartphone to run the RSC classification algorithm can use V2V and V2I communication technology to send other CVs information about RSC data (Figure 4.7) in real-time. If the CV with a camera is not in a CV environment, or the CV with a camera does not have a device such as a laptop or a smartphone to run the RSC classification algorithm, we can upload the images taken by a dash camera on the vehicle as well as the GPS data to the workstation or cloud server with a GPU and CPU to run the RSC classification algorithm and submit the RSC report to any road condition monitoring website, for instance, 511 Alberta, to generate an RSC map (Figure 4.9). The whole pipeline of this application is shown in Figure 4.8.

We used the images and GPS data recorded on November 25th, 2021 to generate an RSC map shown in Figure 4.9 and compared it with the RSC map from 511 Alberta website [82] (Figure 4.10) around the same time. Both RSC maps used the same colours to represent the three RSC, where black stands for bare wet or dry, yellow stands for partly covered by snow or ice, and pink stands for fully covered by snow or ice. We can see that our RSC results on Highway 16 mostly coincide with the RSC results from 511 Alberta, which indicates the RSC on Highway 16 was bare wet or dry. One of the advantages of our RSC map is that we had very detailed RSC on narrow roads, which tend to be accident-prone areas, for example, Liberty Road and Sherwood Drive, where the RSC map from the 511 Alberta website gave no RSC report. We discovered part of Liberty Road was partly covered by snow or ice, which was dangerous for driving. According to CTV News Edmonton [83], 129 crashes were reported to Edmonton police between 5 a.m. and noon on November 25th, 2021. Some crashes could have been avoided with our automated RSC systems by sending RSC data through V2V and V2I in real-time.

Time	RSC	Latitude	Longitude
2021/11/25 13:06:47	Dry/wet	53.5727595	-113.2850251
2021/11/26 13:06:48	Dry/wet	53.5727414	-113.2850255
2021/11/27 13:06:49	Dry/wet	53.5727222	-113.2850278
2021/11/28 13:06:50	Dry/wet	53.5727045	-113.2850325
2021/11/29 13:06:51	Dry/wet	53.572684	-113.2850468

Figure 4.7: Examples of RSC Data

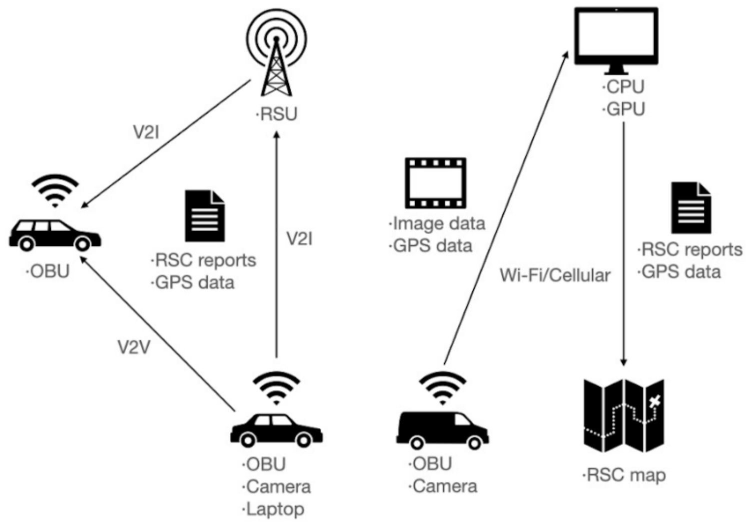


Figure 4.8: Automated RSC Monitoring Pipeline

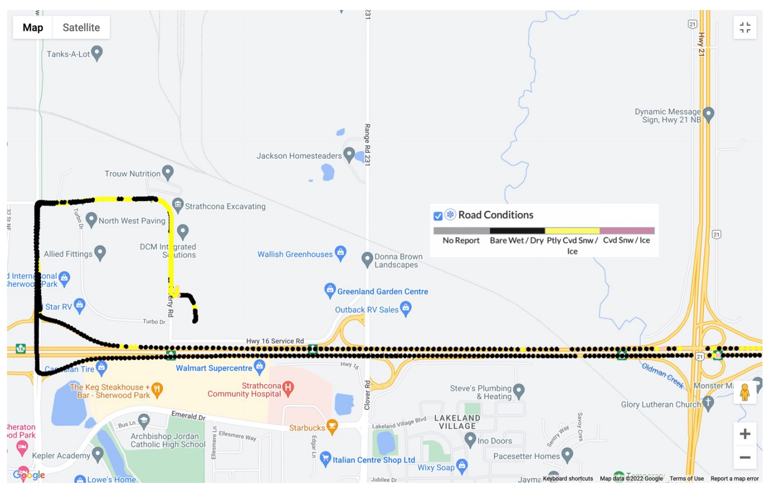


Figure 4.9: Generated RSC Map

4.6 RSC Classifications from Traffic Cameras

We trained and tested Juan's CNN on a dataset that had traffic camera images in Ontario labelled by researchers in iTSS lab [84] at the University of Waterloo, and labelled traffic camera images in Alberta. I mainly collected traffic images from the 511 Alberta website. The images were categorized into three groups based on different RSC conditions. Images with no visible snow on roads were labelled as 'bare pavement' or 'dry/wet', images with part snow-covered road (20% to 60% of the road area were covered by snow) were labelled as 'part snow coverage', and images with full

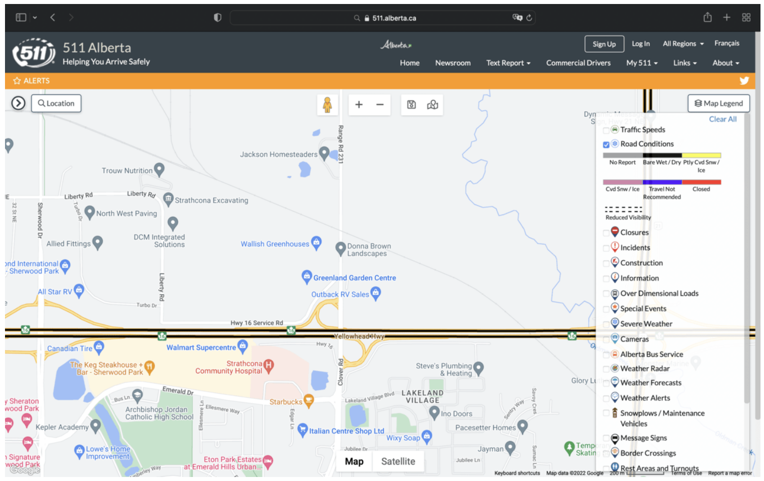
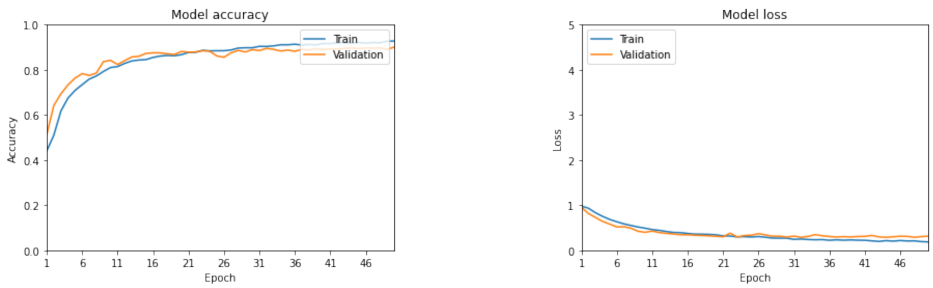


Figure 4.10: RSC Map from 511 Alberta

snow-covered roads (80% to 100% of the road area were covered by snow) were labelled as 'full snow coverage'. Figure 4.11 shows a few sample images from the dataset. The dataset had over 20,000 images in total and 70% were used in training, 20% for validation, and 10% for testing. There were 2,1142 images in total, where 44% were 'bare pavement', 43% were 'part snow coverage', and 11% were 'full snow coverage'. We resized images to 224 pixels by 224 pixels before training. The batch size was chosen to be 4 to save memory and the epoch was set to 50. Training and testing results were shown in Figure 4.12, where the final accuracy was 92.79%, 90.07%, and 90.70% on training, validation, and testing datasets respectively. Since Xception also achieved decent accuracy on the dataset that was full of images collected by onboard cameras, we also trained and tested Xception with the same parameters. Figure 4.13 shows the training and testing results for Xception, where the final accuracy was 80.49%, 82.43% and 81.51% on training, validation, and testing dataset respectively. The processing time was approximately the same as before, where Juan's CNN took about 20 ms and Xception took about 60 ms per image classification. A prototype of an automated RSC monitoring app from live traffic cameras on the 511 Alberta website was created. This app could fetch multiple live images from different traffic cameras and output RSC classification results as well as when the image was taken and its location information. The location information was extracted by reading the text at the bottom of the image from the 511 Alberta website. Figure 4.14 shows a few demonstrations of this prototype. We also compared our RSC map with the RSC map from the 511 Alberta website at the same time on October 24, at noon. The results are shown in Figure 4.15, where black, yellow and pink represent dry/wet, part snow coverage, and full snow coverage respectively. We can clearly see that our method provided a more detailed RSC map in the downtown area where 511 Alberta gave no RSC report. This is because 511 Alberta mainly relies on RWIS stations to provide RSC reports, and RWIS stations are usually located in rural areas. Figure 4.16 gives a comparison of locations of traffic cameras versus RWIS stations in Edmonton. According to statics from 511 Alberta and Alberta transportation [19], there are more than 300 traffic cameras in Alberta while less than 150 RWIS stations and mobile RWIS stations are available. RSC reports generated from traffic cameras located in urban areas by our method were mostly the same as 511 Alberta. In terms of limitations, there were a few cases where our method misclassified RSCs. For example, the RSC report should have been 'dry/wet' but our method gave 'part snow coverage' shown in Figure 4.17a. One possible explanation is that CNN was confused by white trucks on the road and classified white trucks as snow. Another example is shown in Figure 4.17b, where our method gave 'part snow coverage' but the result should have been 'dry/wet'. The reasonable explanation for this can be that CNN thought the roads shown in the figure are one road instead of many and classified the snow on dams as snow on the road. Incorrect reports like these can be solved by adding more data collected from the cameras that our method made mistakes and re-train the CNN. The location of roads that are covered by snow can be broadcast from RSU to nearby CVs in a CV environment.

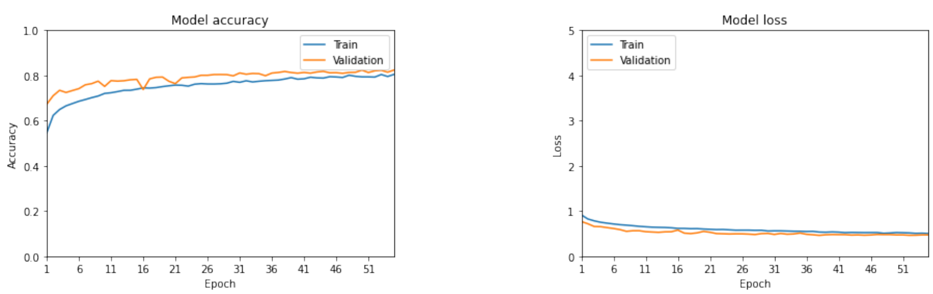
(a) Bare pavement. (b) Part snow coverage. (c) Full snow coverage.

Figure 4.11: Sample Images



(a) The plot of training accuracy. (b) The plot of training loss.

Figure 4.12: Training Juan's CNN with Dropout Layers

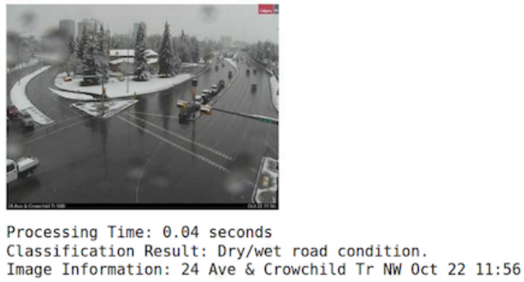


(a) The plot of training accuracy. (b) The plot of training loss.

Figure 4.13: Training Xception with Dropout Layers

4.7 Conclusions

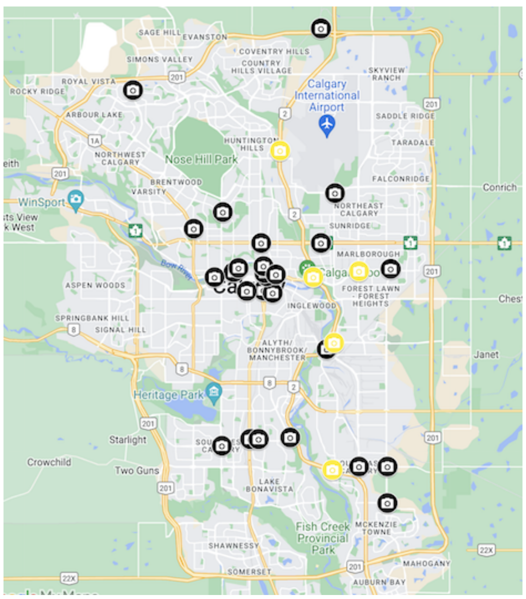
The purpose of this chapter is to use vehicles or CVs as RSC sensors, to give RSC safety alert warnings to other vehicles or RSC monitoring websites such as 511 Alberta in order to avoid crashes. We used images from dash cameras to classify RSC. Three state-of-the-art deep learning models were selected for training on our dataset with over 15,000 labelled images and then compared with each other in terms of accuracy



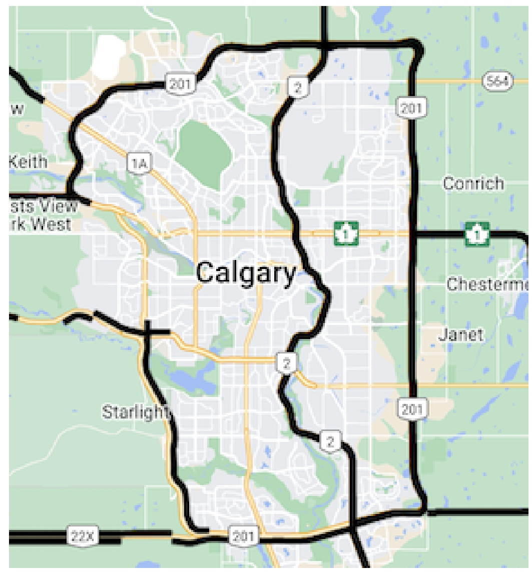
(a) RSC on 24 Ave.

(b) RSC on 34 Ave.

Figure 4.14: RSC Reports app.



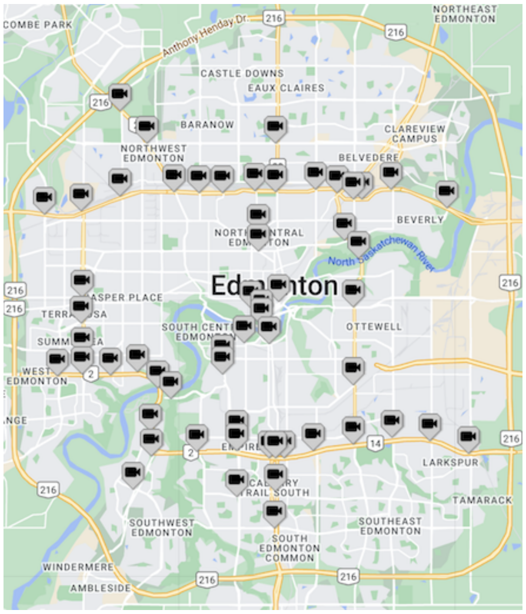
(a) RSC map by our method.



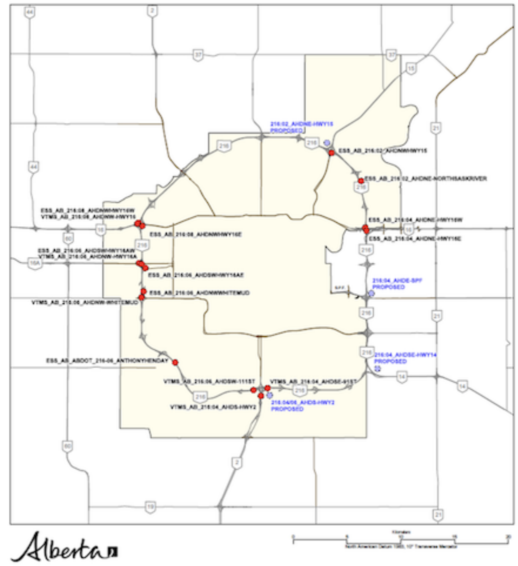
(b) RSC map by 511 Alberta.

Figure 4.15: RSC Map Comparison.

and processing time. From the training results, Juan's CNN was chosen to automatically classify RSC on November 25th, 2021 because of its fast processing time and superior accuracy. The pipeline to generate RSC data and map in a CV environment or non-CV environment was developed. In our RSC map generated on November 25th, 2021, most of the RSC on Highway 16 agreed with the results from 511 Alberta, which validated the reliability of our automated RSC classification systems. By comparing our RSC map with the RSC map from 511 Alberta, a few advantages were found in our method, including detailed RSC on narrow, accident-prone roads where 511 Alberta gave no RSC report. RSC classifications from traffic cameras were also



(a) Locations of traffic cameras are labelled by 'camera' icons.



(b) Locations of RWIS stations are labelled by red points.

Figure 4.16: Traffic Cameras vs. RWIS Stations.

(a) The CNN thought white trucks were snow.

(b) The CNN thought dams were part of roads.

Figure 4.17: Incorrect RSC Reports.

achieved by training CNNs using datasets that had traffic camera images in Ontario and Alberta. Compared with 511 Alberta, our method provided more detailed RSC reports in the downtown area. For future work and research, we will develop an HMI such as an RSC warning app to give drivers RSC safety alerts in real-time. We will also try to augment our dataset and train deep learning models again to analyze performance differences. Other state-of-the-art deep learning models, which were not considered in this chapter, will also be trained and compared.

References

[2] *Alberta traffic collision statistics 2015*, 2017. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/97980388-9327-46fc-b01f-9451d5e076b6/download/ar2015.pdf.

[3] *Alberta traffic collision statistics 2016*, 2018. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/edf10f2d-9056-4d34-8470-9378f590c80d/download/ar2016.pdf.

[4] *Alberta traffic collision statistics 2017*, 2020. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/15f070d4-e0ee-4e49-b107-95fb036ee7d1/download/trans-alberta-traffic-collision-statistics-2017.pdf.

[5] *Alberta traffic collision statistics 2018*, 2021. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/982e6d4f-64d5-4167-81ca-b8c10d76fa59/download/trans-alberta-traffic-collision-statistics-2018.pdf.

[6] *Alberta traffic collision statistics 2019*, 2022. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/30344114-e6b3-459d-a74d-850b6b0d7c07/download/trans-alberta-traffic-collision-statistics-2019.pdf.

[10] I. Juga, P. Nurmi, and M. Hippi, “Statistical modelling of wintertime road surface friction,” *Meteorological Applications*, vol. 20, no. 3, pp. 318–329, 2013.

[15] J. Carrillo, M. Crowley, G. Pan, and L. Fu, “Comparison of deep learning models for determining road surface condition from roadside camera images and weather data,” in *Transportation Association of Canada and ITS Canada 2019 Joint Conference and Exhibition*, 2019.

[16] M. Wu and T. J. Kwon, “An automatic architecture designing approach of convolutional neural networks for road surface conditions image recognition: Tradeoff between accuracy and efficiency,” *Journal of Sensors*, vol. 2022, 2022.

[17] G. Pan, M. Muresan, R. Yu, and L. Fu, “Real-time winter road surface condition monitoring using an improved residual cnn,” *Canadian Journal of Civil Engineering*, vol. 48, no. 9, pp. 1215–1222, 2021.

[18] S. Ramanna, C. Sengoz, S. Kehler, and D. Pham, “Near real-time map building with multi-class image set labeling and classification of road conditions using convolutional neural networks,” *Applied Artificial Intelligence*, vol. 35, no. 11, pp. 803–833, 2021.

[19] *Emerging transportation technologies*. [Online]. Available: https://www.alberta.ca/emerging-transportation-technologies.aspx.

[45] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE conference on computer vision and pattern recognition*, Ieee, 2009, pp. 248–255.

[70] C. M. V. T. C. Statistics, "Canadian council of motor transport administrators," *Ottawa, Ontario*, 2013.

[71] C. Pennelly, G. W. Reuter, and S. Tjandra, "Effects of weather on traffic collisions in edmonton, canada," *Atmosphere-Ocean*, vol. 56, no. 5, pp. 362–371, 2018.

[72] A. Lo, "Alberta's road weather information system (rwis) deployment–an innovative way to outsource the rwis contract," in *2005 Annual Conference of the Transportation Association of CanadaTransportation Association of Canada (TAC)*, 2005.

[73] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[74] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[75] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[76] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[77] O. Russakovsky *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.

[78] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[79] G. Cybenko, D. P. O'Leary, and J. Rissanen, *The Mathematics of Information Coding, Extraction and Distribution*. Springer Science & Business Media, 1998, vol. 107.

[80] J. Brownlee, *Probability for machine learning: Discover how to harness uncertainty with Python*. Machine Learning Mastery, 2019.

[81] K. P. Murphy, *Machine learning: a probabilistic perspective*. MIT press, 2012.

[82] 5. Alberta, *Roads from edmonton to entwistle and red deer are covered in ice due to freezing rain this morning. rcmp are responding to multiple mvcs in the area, and snow plows are sanding. if travel is necessary, slow down and leave extra room. (6:17am)*, Nov. 2021. [Online]. Available: https://twitter.com/511alberta/status/1463859460762320897?s=46&t=PC3z6CXG8o_YUbU7l1uNoA.

[83] A. Antoneshyn, *Freezing rain, black ice create dangerous conditions throughout edmonton region*, Nov. 2021. [Online]. Available: https://edmonton.ctvnews.ca/freezing-rain-black-ice-causing-dangerous-road-conditions-around-edmonton-1.5680787.

[84] *Itss lab*, Oct. 2022. [Online]. Available: https://itsslab.com/.

Chapter 5

Conclusions, Recommendations, & Future Work

5.1 Conclusions

This research was focused on using vision-based machine learning algorithms in a CV environment to monitor road conditions, including work zone detection and localization, as well as snow detection and classification.

YOLO was used to recognize a variety of work zone items including traffic cones, traffic barrels, traffic barricades, construction workers, etc. Localization was achieved by applying perspective transformation and linear transformation. Two localization methods were proposed and tested. Recommendations and cautions to use each method were also discussed. In our localization error analysis, our method can effectively locate work zone items that are placed within 80 meters of the traffic camera. We used MAP and TIM messages to broadcast work zone information such as work zone location and which lane was closed by the work zone, to nearby CVs using C-V2X. The development of the work zone warning app was detailed. We first developed a work zone warning app without traffic lane information, then with the help of MAP and TIM, a prototype of the work zone warning app with detailed traffic lane information was tested in different scenarios. We also studied the effects of different weather conditions and camera resolutions on our detection and localization method. We found out that our method could work fine using an extremely low-resolution traffic camera. A study case was performed on 118 Street to validate the system's stability as well as latency. Results revealed that our system can ensure the safety of both drivers and construction workers. I believe a lot of lives can be saved with our system in the future since the driver will receive work zone safety alerts before the driver can even see the work zone.

Three state-of-the-art CNNs were trained and tested on our own RSC dataset collected from cameras mounted on vehicles. The dataset was collected in Edmonton and Calgary and labelled into three categories, including bare pavement, part snow-covered, and full snow-covered. We selected the CNN with superior performance and fast processing time to be used in our automated RSC application in a CV environment. We compared our RSC method with the method that is currently in use by the Alberta government and found that our method had the advantage of a more detailed RSC report on narrow, accident-prone roads. A pipeline of automated RSC classification in a CV environment was proposed. With the help of DSRC, C-V2X, and 5G communication technology, RSC reports with GPS information can be shared in real-time with nearby CVs, further ensuring driver safety under slippery roads. Images from traffic cameras were also used to train CNNs and an automated RSC classifications prototype was developed and tested. The prototype used traffic camera images from the 511 Alberta website as input and output RSC reports and RSC maps. Compared with the RSC map from 511 Alberta, our method had a more detailed RSC map in the downtown area. Our methods including automated RSC classifications from CVs and traffic cameras had proved to be reliable complements to current RSC monitoring methods, which mainly rely on stationary and mobile RWIS stations.

This study mostly relied on cameras as input sources so it had limitations of poor performance at night. It is recommended to use Lidar sensors to complement this drawback. In addition, our work zone localization method was effective if the work zone items were placed within 80 meters of traffic cameras. But in real-world scenarios, work zones can be kilometres long on highways where our method cannot cover the full work zone range. Feasible solutions to solve this problem can be using a drone to detect and locate work zones, and we can use multiple traffic cameras to enlarge the range of detection and localization. For RSC classification problems from CVs and traffic cameras, it also had poor performance at night. Our automated RSC monitoring systems can be perfected by fusing the data with RWIS station and On-Board Diagnostic systems (OBD) from vehicles.

5.2 Future Work

For future works and plans, we will first complete the development of mobile versions of the work zone warning apps. The work zone boundary detection accuracy was measured by recognizing traffic cones instead of boundary detection. It would make more sense to evaluate the boundary identification accuracy where the missing detection of a traffic cone does not affect the boundary detection results in many cases, but sometimes it does. The false positive case could have a high impact on boundary detection in certain cases and many have little impact in some other cases. We will follow the pipeline developed by Shi et al. [26] where they evaluated the overlap area between the polygons that enclose the entire work zone area measured by ground truth and vision-based methods. Although they did not specify which sensors were used to measure the ground truth of the polygon, we can use GNSS sensors in our laboratory to reproduce the experiment. We will develop systematic ways to quantify and minimize the distortions generated by perspective transformation. For the case study on 118 street, we will repeat the experiments as many times as possible to ensure effectiveness and stability. We used DSRC and C-V2X communication technologies in the CV environment, and 5G will be considered and utilized in the future to further reduce the latency. In our work zone warning applications, we used server-based structures to transmit data from the traffic camera to the central server and then to the RSU. This structure can be simplified by a Multi-Edge Computing (MEC) device such as

Islam et al. [40] used in their paper. They placed the computing device right next to the traffic camera and the RSU. As a result, this MEC device dramatically reduced the latency. We will also explore the possibilities of using cooperative perspective messages (CPM) to detect and locate moving objects, such as construction workers from onboard cameras mounted on CVs. A few types of research are being studied and will be reproduced in the future including Rauch et al. [85–87], Gunther et al. [88] and Shan et al. [89].

In terms of classifications of winter road surface conditions, we will fuse the data with CVs and traffic cameras with RWIS stations as Juan et al. [15] and Diaby et al. [90] did in their papers. These fusion methods will perfect our methods and fix the issues of poor performance at night. We will continue to collect data from CVs and traffic cameras this winter in Alberta to enlarge the datasets. We will also complete the development of automated RSC monitoring apps using input data from CVs and traffic cameras. More testing will be conducted this winter to verify the effectiveness and stability of our methods. CNN will be trained on mixed datasets to see if we can achieve better results. 5G and MEC devices will be used in the future to reduce the latency of transmitting RSC reports from one CV to another.

References

[15] J. Carrillo, M. Crowley, G. Pan, and L. Fu, “Comparison of deep learning models for determining road surface condition from roadside camera images and weather data,” in *Transportation Association of Canada and ITS Canada 2019 Joint Conference and Exhibition*, 2019.

[26] W. Shi and R. R. Rajkumar, “Work zone detection for autonomous vehicles,” in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, IEEE, 2021, pp. 1585–1591.

[40] M. Islam, M. Rahman, M. Chowdhury, G. Comert, E. D. Sood, and A. Apon, “Vision-based personal safety messages (psms) generation for connected vehicles,” *IEEE transactions on vehicular technology*, vol. 69, no. 9, pp. 9402–9416, 2020.

[85] A. Rauch, F. Klanner, and K. Dietmayer, “Analysis of v2x communication parameters for the development of a fusion architecture for cooperative perception systems,” in *2011 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, 2011, pp. 685–690.

[86] A. Rauch, F. Klanner, R. Rasshofer, and K. Dietmayer, “Car2x-based perception in a high-level fusion architecture for cooperative perception systems,” in *2012 IEEE Intelligent Vehicles Symposium*, IEEE, 2012, pp. 270–275.

[87] A. Rauch, S. Maier, F. Klanner, and K. Dietmayer, “Inter-vehicle object association for cooperative perception systems,” in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, IEEE, 2013, pp. 893–898.

[88] H.-J. Günther, B. Mennenga, O. Trauer, R. Riebl, and L. Wolf, “Realizing collective perception in a vehicle,” in *2016 IEEE Vehicular Networking Conference (VNC)*, IEEE, 2016, pp. 1–8.

[89] M. Shan *et al.*, “A novel probabilistic v2x data fusion framework for cooperative perception,” *arXiv preprint arXiv:2203.16964*, 2022.

[90] I. Diaby, M. Germain, and K. Goïta, “Evidential data fusion for characterization of pavement surface conditions during winter using a multi-sensor approach,” *Sensors*, vol. 21, no. 24, p. 8218, 2021.

Bibliography

[1] N. H. T. S. Administration *et al.*, "2015 motor vehicle crashes: Overview," *Traffic safety facts: research note*, vol. 2016, pp. 1–9, 2016.

[2] *Alberta traffic collision statistics 2015*, 2017. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/97980388-9327-46fc-b01f-9451d5e076b6/download/ar2015.pdf.

[3] *Alberta traffic collision statistics 2016*, 2018. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/edf10f2d-9056-4d34-8470-9378f590c80d/download/ar2016.pdf.

[4] *Alberta traffic collision statistics 2017*, 2020. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/15f070d4-e0ee-4e49-b107-95fb036ee7d1/download/trans-alberta-traffic-collision-statistics-2017.pdf.

[5] *Alberta traffic collision statistics 2018*, 2021. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/982e6d4f-64d5-4167-81ca-b8c10d76fa59/download/trans-alberta-traffic-collision-statistics-2018.pdf.

[6] *Alberta traffic collision statistics 2019*, 2022. [Online]. Available: https://open.alberta.ca/dataset/25020446-adfb-4b57-9aaa-751d13dab72d/resource/30344114-e6b3-459d-a74d-850b6b0d7c07/download/trans-alberta-traffic-collision-statistics-2019.pdf.

[7] S. D. Schrock, G. L. Ullman, A. S. Cothron, E. Kraus, and A. P. Voigt, "An analysis of fatal work zone crashes in texas," *Report FHW A/TX-05/0-4028*, vol. 1, 2004.

[8] J. Daniel, K. Dixon, and D. Jared, "Analysis of fatal crashes in georgia work zones," *Transportation Research Record*, vol. 1715, no. 1, pp. 18–23, 2000.

[9] N. J. Garber and M. Zhao, "Distribution and characteristics of crashes at different work zone locations in virginia," *Transportation Research Record*, vol. 1794, no. 1, pp. 19–25, 2002.

[10] I. Juga, P. Nurmi, and M. Hippi, "Statistical modelling of wintertime road surface friction," *Meteorological Applications*, vol. 20, no. 3, pp. 318–329, 2013.

[11] A. Dehman and B. Farooq, "Are work zones and connected automated vehicles ready for a harmonious coexistence? a scoping review and research agenda," *Transportation research part C: emerging technologies*, vol. 133, p. 103 422, 2021.

[12] X. Wu *et al.*, "Vehicular communications using dsrc: Challenges, enhancements, and evolution," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 399–408, 2013.

[13] W. Han, E. White, M. Mollenhauer, and N. Roofigari-Esfahan, "A connected work zone hazard detection system for roadway construction workers," in *Computing in Civil Engineering 2019: Smart Cities, Sustainability, and Resilience*, American Society of Civil Engineers Reston, VA, 2019, pp. 242–250.

[14] R. Schönrock, F. Wolf, and T. Russ, "Smart traffic cone-dynamic detection and localization of traffic disruptions," in *FAST-zero'15: 3rd International Symposium on Future Active Safety Technology Toward zero traffic accidents, 2015*, 2015.

[15] J. Carrillo, M. Crowley, G. Pan, and L. Fu, "Comparison of deep learning models for determining road surface condition from roadside camera images and weather data," in *Transportation Association of Canada and ITS Canada 2019 Joint Conference and Exhibition*, 2019.

[16] M. Wu and T. J. Kwon, "An automatic architecture designing approach of convolutional neural networks for road surface conditions image recognition: Tradeoff between accuracy and efficiency," *Journal of Sensors*, vol. 2022, 2022.

[17] G. Pan, M. Muresan, R. Yu, and L. Fu, "Real-time winter road surface condition monitoring using an improved residual cnn," *Canadian Journal of Civil Engineering*, vol. 48, no. 9, pp. 1215–1222, 2021.

[18] S. Ramanna, C. Sengoz, S. Kehler, and D. Pham, "Near real-time map building with multi-class image set labeling and classification of road conditions using convolutional neural networks," *Applied Artificial Intelligence*, vol. 35, no. 11, pp. 803–833, 2021.

[19] *Emerging transportation technologies*. [Online]. Available: https://www.alberta.ca/emerging-transportation-technologies.aspx.

[20] *511 alberta*. [Online]. Available: https://511.alberta.ca/#:Alerts.

[21] F. Abodo, R. Rittmuller, B. Sumner, and A. Berthaume, "Detecting work zones in shrp 2 nds videos using deep learning based computer vision," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, IEEE, 2018, pp. 679–686.

[22] Y.-W. Seo, J. Lee, W. Zhang, and D. Wettergreen, "Recognition of highway workzones for reliable autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 708–718, 2014.

[23] B. Mathibela, M. A. Osborne, I. Posner, and P. Newman, "Can priors be trusted? learning to anticipate roadworks," in *2012 15th International IEEE Conference on Intelligent Transportation Systems*, IEEE, 2012, pp. 927–932.

[24] P. Kunz and M. Schreier, "Automated detection of construction sites on motorways," in *2017 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, 2017, pp. 1378–1385.

[25] R. Graf, A. Wimmer, and K. C. Dietmayer, "Probabilistic estimation of temporary lanes at road work zones," in *2012 15th International IEEE Conference on Intelligent Transportation Systems*, IEEE, 2012, pp. 716–721.

[26] W. Shi and R. R. Rajkumar, "Work zone detection for autonomous vehicles," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, IEEE, 2021, pp. 1585–1591.

[27] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587.

[28] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448.

[29] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, vol. 28, 2015.

[30] W. Liu *et al.*, "Ssd: Single shot multibox detector," in *European conference on computer vision*, Springer, 2016, pp. 21–37.

[31] G. Jocher, *Yolov5. code repository https://www. github. com/ultralytics/yolov5*, 2022.

[32] J. Redmon and A. Farhadi, "Yolo9000: Better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271.

[33] C. for Advanced Automotive Technology. "Connected and automated vehicles." (2018), [Online]. Available: http://autocaat.org/Technologies/Automated_and_Connected_Vehicles/ (visited on 09/12/2022).

[34] W. Genders and S. N. Razavi, "Impact of connected vehicle on work zone network safety through dynamic route guidance," *Journal of Computing in Civil Engineering*, vol. 30, no. 2, p. 04 015 020, 2016.

[35] K. C. Dey, A. Rayamajhi, M. Chowdhury, P. Bhavsar, and J. Martin, "Vehicle-to-vehicle (v2v) and vehicle-to-infrastructure (v2i) communication in a heterogeneous wireless network–performance evaluation," *Transportation Research Part C: Emerging Technologies*, vol. 68, pp. 168–184, 2016.

[36] H. Abdulsattar, A. Mostafizi, and H. Wang, "Surrogate safety assessment of work zone rear-end collisions in a connected vehicle environment: Agent-based modeling framework," *Journal of Transportation Engineering, Part A: Systems*, vol. 144, no. 8, 2017.

[37] J. Lombardo. "How connected vehicles make work zones safer." (2017), [Online]. Available: https://www.forconstructionpros.com/asphalt/article/20867242/how-connected-vehicles-make-work-zones-safer (visited on 09/12/2022).

[38] S. Mishra, M. M. Golias, D. Thapa, *et al.*, "Work zone alert systems," Tennessee. Department of Transportation, Tech. Rep., 2021.

[39] F. Qiao, R. Rahman, Q. Li, and L. Yu, "Safe and environment-friendly forward collision warning messages in the advance warning area of a construction zone," *International journal of intelligent transportation systems research*, vol. 15, no. 3, pp. 166–179, 2017.

[40] M. Islam, M. Rahman, M. Chowdhury, G. Comert, E. D. Sood, and A. Apon, "Vision-based personal safety messages (psms) generation for connected vehicles," *IEEE transactions on vehicular technology*, vol. 69, no. 9, pp. 9402–9416, 2020.

[41] V. V. A. T. Committee, *Vulnerable road user safety message minimum performance requirements*, Mar. 2017.

[42] V. C. T. Committee, *Dedicated short range communications (dsrc) message set dictionary*, Mar. 2016.

[43] O. Polly, D. Stephens, and R. Ostroff, "Work zone activity data (wzad) – data dictionary report (version 2)," 2019.

[44] M. Elhoseiny, S. Huang, and A. Elgammal, "Weather classification with deep convolutional neural networks," in *2015 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2015, pp. 3349–3353.

[45] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, Ieee, 2009, pp. 248–255.

[46] C. Lu, D. Lin, J. Jia, and C.-K. Tang, "Two-class weather classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3718–3725.

[47] D. Lin, C. Lu, H. Huang, and J. Jia, "Rscm: Region selection and concurrency model for multi-class weather recognition," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4154–4167, 2017.

[48] Z. Zhu, L. Zhuo, P. Qu, K. Zhou, and J. Zhang, "Extreme weather recognition using convolutional neural networks," in *2016 IEEE International Symposium on Multimedia (ISM)*, IEEE, 2016, pp. 621–625.

[49] J. C. V. Guerra, Z. Khanam, S. Ehsan, R. Stolkin, and K. McDonald-Maier, "Weather classification: A new multi-class dataset, data augmentation approach and comprehensive evaluations of convolutional neural networks," in *2018 NASA/ESA Conference on Adaptive Hardware and Systems (AHS)*, IEEE, 2018, pp. 305–310.

[50] Z. Li, Y. Jin, Y. Li, Z. Lin, and S. Wang, "Imbalanced adversarial learning for weather image generation and classification," in *2018 14th IEEE International Conference on Signal Processing (ICSP)*, IEEE, 2018, pp. 1093–1097.

[51] B. Zhao, X. Li, X. Lu, and Z. Wang, "A cnn–rnn architecture for multi-label weather recognition," *Neurocomputing*, vol. 322, pp. 47–57, 2018.

[52] G. Pan, L. Fu, R. Yu, and M. I. Muresan, "Winter road surface condition recognition using a pre-trained deep convolutional neural network," Tech. Rep., 2018.

[53] M. A. Linton and L. Fu, "Connected vehicle solution for winter road surface condition monitoring," *Transportation Research Record*, vol. 2551, no. 1, pp. 62–72, 2016.

[54] M. A. Linton and L. Fu, "Winter road surface condition monitoring: Field evaluation of a smartphone-based system," *Transportation research record*, vol. 2482, no. 1, pp. 46–56, 2015.

[55] A. Kuehnle and W. Burghout, "Winter road condition recognition using video image classification," *Transportation research record*, vol. 1627, no. 1, pp. 29–33, 1998.

[56] C. Zhang, "Application of convolutional neural network (cnn) models for automated monitoring of road pavement and winter surface conditions using visual-spectrum and thermal video cameras," 2021.

[57] E. Radwan, R. Harb, S. Ramasamy, *et al.*, "Evaluation of safety and operational effectiveness of dynamic lane merge system in florida: Final report, april 2009.," 2009.

[58] Z. A. Nemeth and D. J. Migletz, "Accident characteristics before, during, and after safety upgrading projects on ohio's rural interstate system," *Transportation Research Record*, vol. 672, pp. 19–23, 1978.

[59] J. Hall and V. Lorenz, "Characteristics of construction-zone accidents," *Transportation Research Record*, vol. 1230, pp. 20–27, 1989.

[60] N. M. Rouphail, Z. S. Yang, and J. Fazio, "Comparative study of short-and long-term urban freeway work zones," *Transportation Research Record*, vol. 1163, pp. 4–14, 1988.

[61] O. Raddaoui, M. M. Ahmed, and S. M. Gaweesh, "Assessment of the effectiveness of connected vehicle weather and work zone warnings in improving truck driver safety," *IATSS research*, vol. 44, no. 3, pp. 230–237, 2020.

[62] R. F. Benekohal, E. Shim, and P. T. Resende, "Truck drivers' concerns in work zones: Travel characteristics and accident experiences," *Transportation Research Record*, vol. 1509, pp. 55–64, 1995.

[63] T. Wen, Z. Xiao, K. Jiang, M. Yang, K. Li, and D. Yang, "High precision target positioning method for rsu in cooperative perception," in *2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP)*, IEEE, 2019, pp. 1–6.

[64] T.-Y. Lin *et al.*, "Microsoft coco: Common objects in context," in *European conference on computer vision*, Springer, 2014, pp. 740–755.

[65] *How to train yolov5 on custom objects*. [Online]. Available: https://colab.research.google.com/drive/1gDZ2xcTOgR39tGGs-EZ6i3RTs16wmzZQ.

[66] T. U. goverment. “Other global navigation satellite systems (gnss).” (2021), [Online]. Available: https://www.gps.gov/systems/gnss/ (visited on 09/12/2022).

[67] M. Brostrom, *A collection of real-time multi-camera multi-object sota trackers using yolov5*. [Online]. Available: https://github.com/mikel-brostrom/Yolov5_StrongSORT_OSNet.

[68] *Usdot-jpo-ode/wzdx: The work zone data exchange (wzdx) specification aims to make harmonized work zone data provided by infrastructure owners and operators (ioos) available for third party use, making travel on public roads safer and more efficient through ubiquitous access to data on work zone activity.* [Online]. Available: https://github.com/usdot-jpo-ode/wzdx.

[69] [Online]. Available: https://www.openstreetmap.org/#map=15/53.4961/-113.5359.

[70] C. M. V. T. C. Statistics, “Canadian council of motor transport administrators,” *Ottawa, Ontario*, 2013.

[71] C. Pennelly, G. W. Reuter, and S. Tjandra, “Effects of weather on traffic collisions in edmonton, canada,” *Atmosphere-Ocean*, vol. 56, no. 5, pp. 362–371, 2018.

[72] A. Lo, “Alberta’s road weather information system (rwis) deployment–an innovative way to outsource the rwis contract,” in *2005 Annual Conference of the Transportation Association of CanadaTransportation Association of Canada (TAC)*, 2005.

[73] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[74] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[75] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.

[76] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[77] O. Russakovsky *et al.*, “Imagenet large scale visual recognition challenge,” *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.

[78] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[79] G. Cybenko, D. P. O’Leary, and J. Rissanen, *The Mathematics of Information Coding, Extraction and Distribution*. Springer Science & Business Media, 1998, vol. 107.

[80] J. Brownlee, *Probability for machine learning: Discover how to harness uncertainty with Python*. Machine Learning Mastery, 2019.

[81] K. P. Murphy, *Machine learning: a probabilistic perspective.* MIT press, 2012.

[82] 5. Alberta, *Roads from edmonton to entwistle and red deer are covered in ice due to freezing rain this morning. rcmp are responding to multiple mvcs in the area, and snow plows are sanding. if travel is necessary, slow down and leave extra room. (6:17am)*, Nov. 2021. [Online]. Available: https://twitter.com/511alberta/status/1463859460762320897?s=46&t=PC3z6CXG8o_YUbU7l1uNoA.

[83] A. Antoneshyn, *Freezing rain, black ice create dangerous conditions throughout edmonton region*, Nov. 2021. [Online]. Available: https://edmonton.ctvnews.ca/freezing-rain-black-ice-causing-dangerous-road-conditions-around-edmonton-1.5680787.

[84] *Itss lab*, Oct. 2022. [Online]. Available: https://itsslab.com/.

[85] A. Rauch, F. Klanner, and K. Dietmayer, "Analysis of v2x communication parameters for the development of a fusion architecture for cooperative perception systems," in *2011 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, 2011, pp. 685–690.

[86] A. Rauch, F. Klanner, R. Rasshofer, and K. Dietmayer, "Car2x-based perception in a high-level fusion architecture for cooperative perception systems," in *2012 IEEE Intelligent Vehicles Symposium*, IEEE, 2012, pp. 270–275.

[87] A. Rauch, S. Maier, F. Klanner, and K. Dietmayer, "Inter-vehicle object association for cooperative perception systems," in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, IEEE, 2013, pp. 893–898.

[88] H.-J. Günther, B. Mennenga, O. Trauer, R. Riebl, and L. Wolf, "Realizing collective perception in a vehicle," in *2016 IEEE Vehicular Networking Conference (VNC)*, IEEE, 2016, pp. 1–8.

[89] M. Shan *et al.*, "A novel probabilistic v2x data fusion framework for cooperative perception," *arXiv preprint arXiv:2203.16964*, 2022.

[90] I. Diaby, M. Germain, and K. Goïta, "Evidential data fusion for characterization of pavement surface conditions during winter using a multi-sensor approach," *Sensors*, vol. 21, no. 24, p. 8218, 2021.

Appendix A: First Appendix

This appendix contains figures in Section 3.7.3

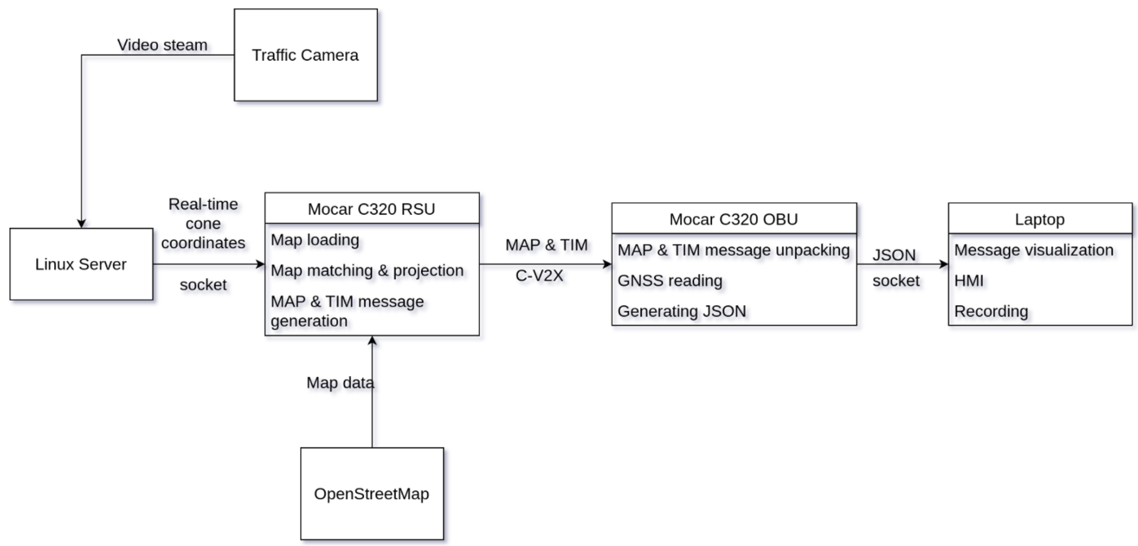


Figure A.1: System Diagram.

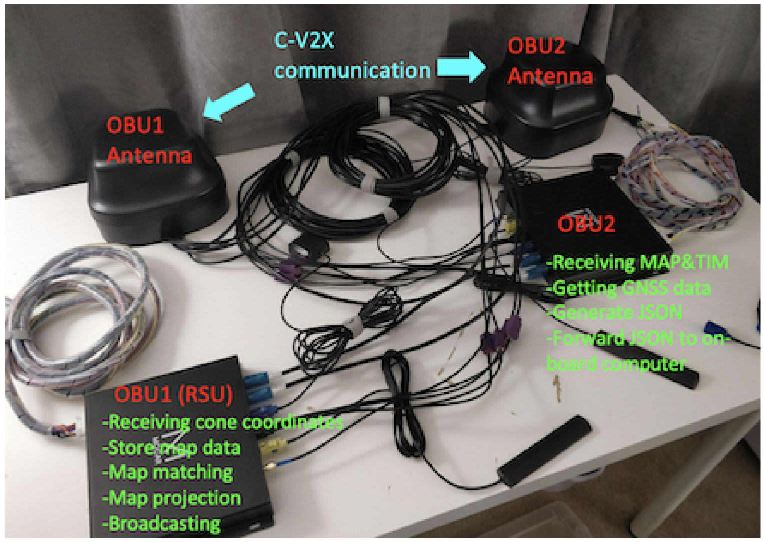
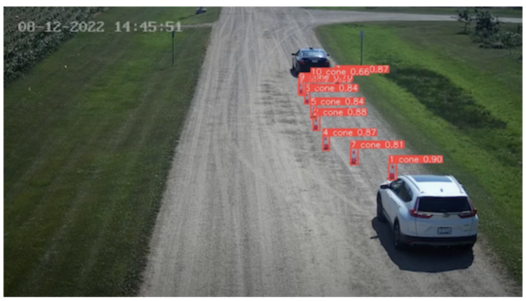
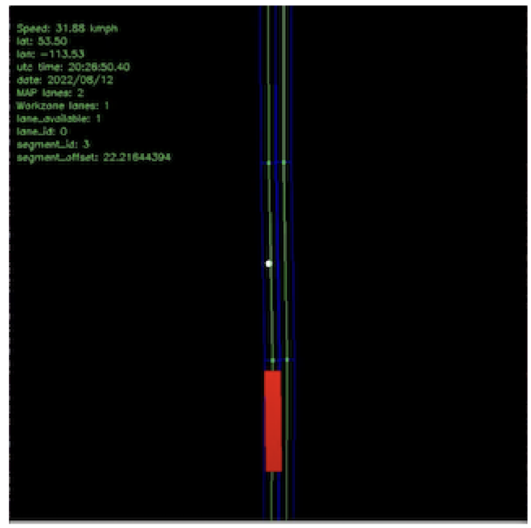


Figure A.2: Hardware.

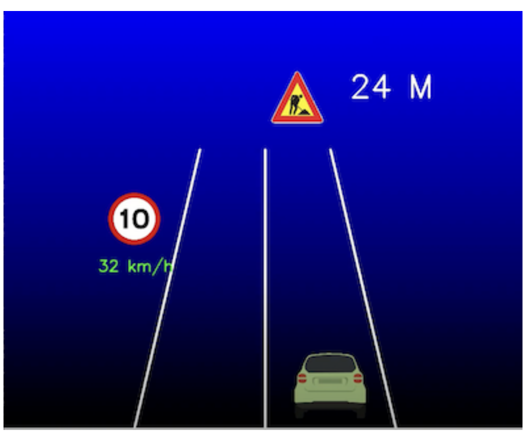
Figure A.3: MAP Data Collection.



(a) YOLO detection.



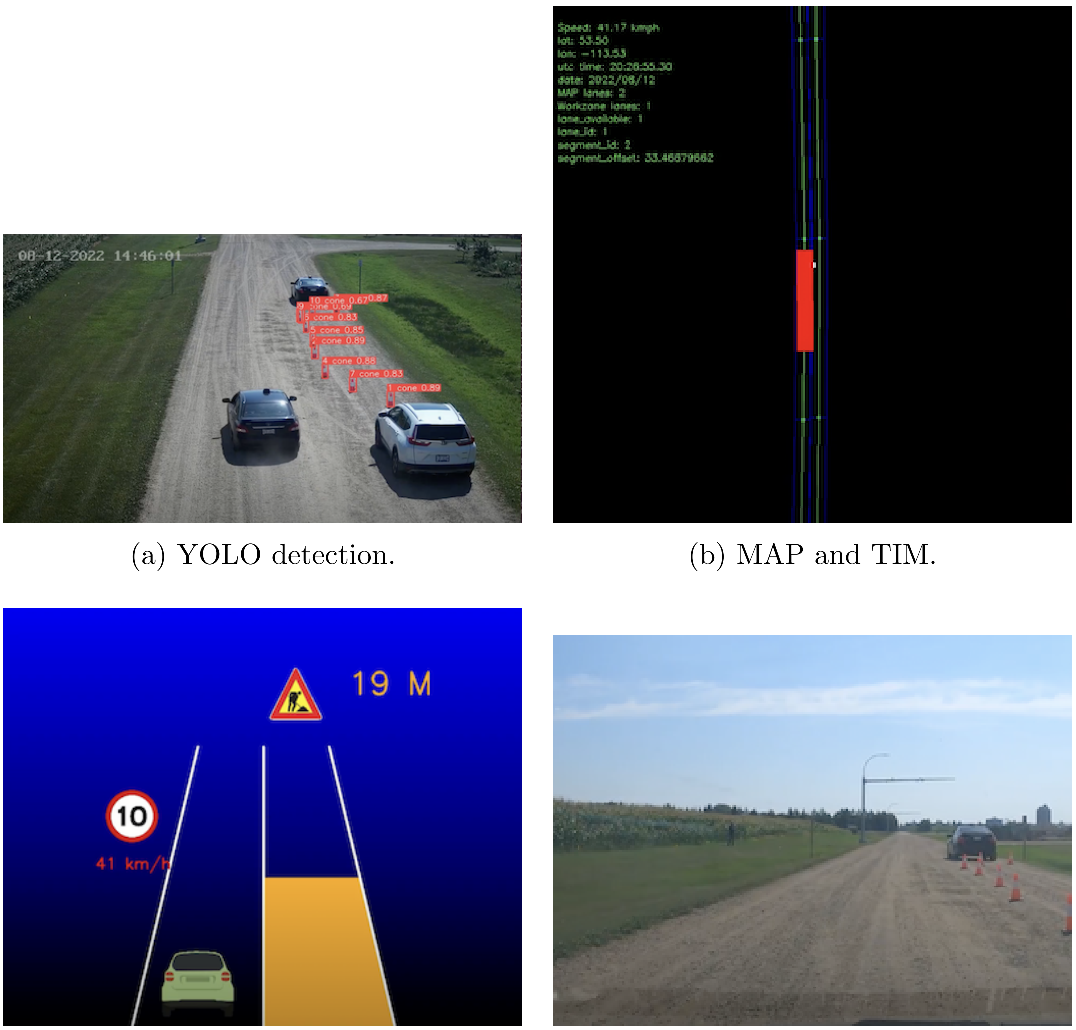
(b) MAP and TIM.



(c) Warning app.

(d) Driver's vision.

Figure A.4: Approaching the Work Zone.



(a) YOLO detection.

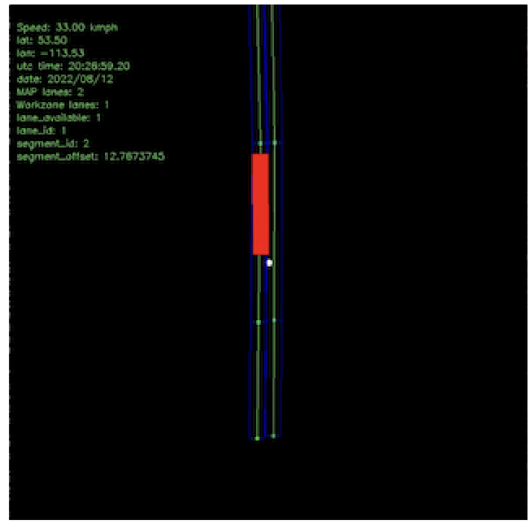
(b) MAP and TIM.

(c) Warning app.

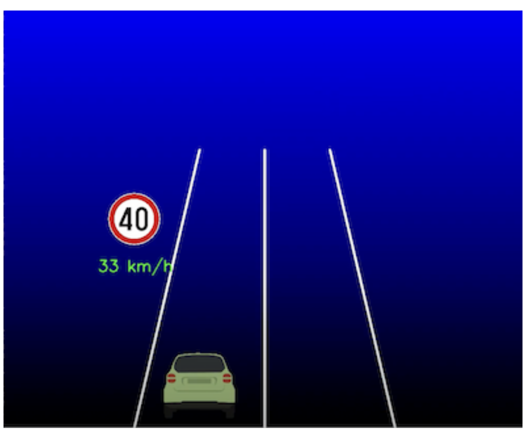
(d) Driver's vision.

Figure A.5: Entering the Work Zone.

(a) YOLO detection.



(b) MAP and TIM.

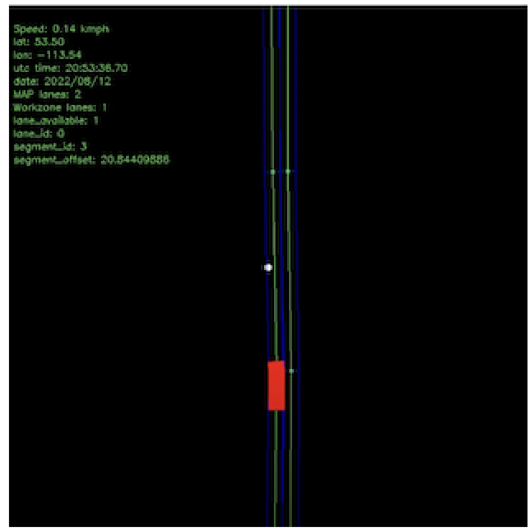


(c) Warning app.

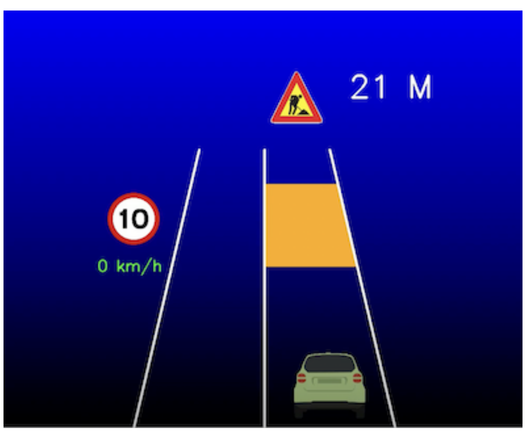
(d) Driver's vision.

Figure A.6: Leaving the Work Zone.

(a) YOLO detection.



(b) MAP and TIM.

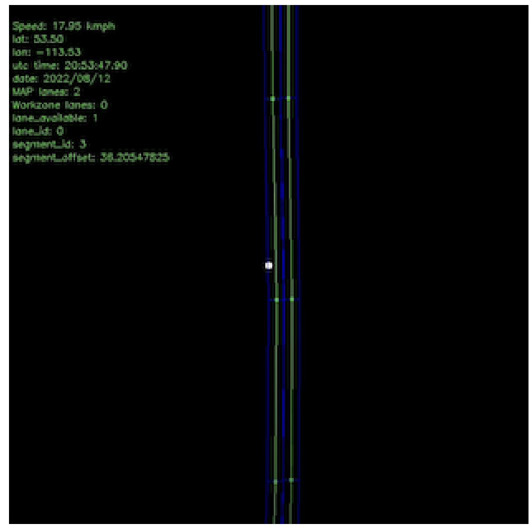


(c) Warning app.

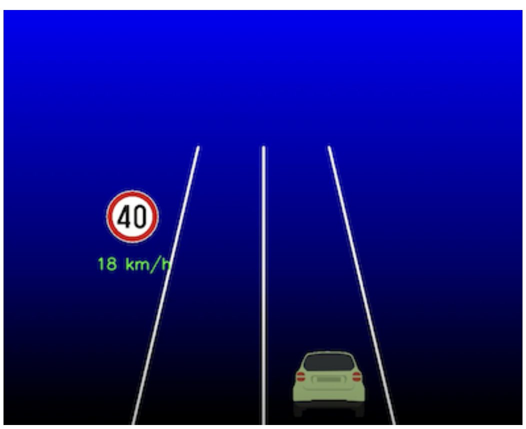
(d) Driver's vision.

Figure A.7: Half of Traffic Cones were Removed.

(a) YOLO detection.



(b) MAP and TIM.



(c) Warning app.

(d) Driver's vision.

Figure A.8: All Traffic Cones were Removed.