NRC·CNRC

Technical Report

Enhancing Situational Awareness at Grade Crossings Using Machine Vision Technologies

Prepared for: Invest Ottawa (Area X.O)
7 Bayview Station Rd.
Ottawa, ON, K1Y 2C5

Prepared by: Md Atiqur Rahman, Samy Metari, Abdelhamid Mammeri, Mark Croken

Automotive and Surface Transportation
National Research Council Canada

13 June 2024

Project: A1-022649
Report number: AST-2024-0035
Version: 1.0

National Research Council Canada | Conseil national de recherches Canada

Canada

Notice of Disclosure

Change Control

Version	Date	Description	Author
0.1	09 May 2024	Initial draft release	Md Atiqur Rahman, Samy Metari, Abdelhamid Mammeri, Mark Croken
0.2	28 May 2024	Final draft release	Md Atiqur Rahman, Samy Metari, Abdelhamid Mammeri, Mark Croken
0.3	11 June 2024	Revised final draft release	Md Atiqur Rahman, Samy Metari, Abdelhamid Mammeri, Mark Croken
1.0	13 June 2024	Initial release	Md Atiqur Rahman, Samy Metari, Abdelhamid Mammeri, Mark Croken

Prepared by:

Rahman, MdAtiqur

Digitally signed by Rahman, MdAtiqur
DN: cn=Rahman, MdAtiqur, c=CA, o=GC, ou=NRC-CNRC, email=mdatiqur.rahman@cnrc-nrc.gc.ca
Date: 2024.06.13 16:06:48 -04'00'

Md Atiqur Rahman
Research Associate, Automotive and Surface Transportation

Metari, Samy

Digitally signed by Metari, Samy
DN: cn=Metari, Samy, c=CA, o=GC, ou=NRC-CNRC, email=samy.metari@cnrc-nrc.gc.ca
Date: 2024.06.13 20:13:17 -04'00'

Samy Metari
Research Council Officer, Automotive and Surface Transportation

Abdelhamid Mammeri

Digitally signed by Abdelhamid Mammeri
DN: cn=Abdelhamid Mammeri, c=CA, o=National Research Council Canada, ou=Automotive and Surface Transportation, email=abdelhamid.mammeri@nrc-cnrc.gc.ca
Date: 2024.06.13 16:33:55 -04'00'

Abdelhamid Mammeri
Team Leader, Intelligent Transportation Systems, Automotive and Surface Transportation

Croken, Mark

Digitally signed by Croken, Mark
Date: 2024.06.14 07:20:38 -04'00'

Mark Croken
Research Council Officer, Automotive and Surface Transportation

Reviewed by:

PrestonThomas, Jonathan

Digitally signed by PrestonThomas, Jonathan
DN: cn=PrestonThomas, Jonathan, c=CA, o=GC, ou=NRC-CNRC, email=jon.preston-thomas@canada.ca
Date: 2024.06.13 17:26:14 -04'00'

Jon Preston-Thomas, P.Eng.
Principal Engineer, Automotive and Surface Transportation

Approved by:

Marsh, Philip

Digitally signed by Marsh, Philip
DN: cn=Marsh, Philip, c=CA, o=GC, ou=NRC-CNRC, email=philip.marsh@canada.ca
Date: 2024.06.17 15:27:28 -04'00'

Philip Marsh, P.Eng.
Director R&D, Transportation Engineering Centre

Acknowledgement

The authors would like to thank the following people for their help throughout this project:

- **Divyanshu Kamboj**, Invest Ottawa (Area X.O): for contributing input into the project deliverables and aiding with the client data transfer
- **Omar Abdeddaiem,** student intern: for contributing to data annotation and data validation.
- **Niloofar Hooshyaripour,** student intern: for contributing to data annotation.
- **Reza Sadeghian,** student intern: for contributing to data annotation.
- **Ismael Ouadria,** student intern: for contributing to data annotation.
- **Kristin Scott,** student intern: for contributing to data annotation.

Executive Summary

This project aimed to enhance safety at Canadian grade crossings by identifying and mitigating the root causes of collisions through the use of sophisticated artificial intelligence and machine vision technologies. Initially, a comprehensive literature review was conducted to identify the primary factors contributing to grade crossing collisions and to highlight specific unsafe events, such as vehicle U-turns, queue buildup on the crossing, and intrusion into the railway right of way.

The methodology included installing two thermal cameras and two PTZ cameras at a public grade crossing to collect data over four months. This data was then analyzed and annotated to develop deep learning algorithms for detecting and tracking vehicles and vulnerable road users (VRUs). The detection pipeline used raw video feeds from both RGB and thermal cameras to identify vehicles and VRUs, tracked their movements to create trajectories, and analyzed these trajectories to detect potentially unsafe events at or near the grade crossing.

The detection results showed excellent performance, with RGB cameras excelling in vehicle detection during daytime and thermal cameras performing the best at night. VRU detection was particularly effective with thermal images due to the distinctive heat signatures captured. The algorithms developed for detecting unsafe events, such as U-turns, queue buildup on the crossing, and jaywalking/intrusions, were highly successful, making the overall pipeline a solid proof of concept.

This comprehensive approach aims to develop a real-time system capable of recognizing unsafe events and broadcasting warnings through vehicle-to-everything (V2X) communications to approaching vehicles and road users in order to increase their situational awareness, thereby preventing accidents and enhancing safety at grade crossings.

Table of contents

List of tables

List of figures

1 Introduction

1.1 Background

There are about 14,000 public and 9,000 private grade crossings along over 26,000 miles of federally regulated rail tracks in Canada. From 2012 to 2022, the Transportation Safety Board of Canada (TSB) recorded a total of 11,426 rail accidents [1]. Among these accidents, approximately 16% (1,773 accidents) took place at-grade crossings, resulting in 233 fatalities and 288 serious injuries [1]. This data highlights that grade crossings are the second most hazardous type of rail accident in Canada, surpassed only by train derailments. Consequently, it is crucial to prioritize the enhancement of rail crossing safety to establish a safer rail transportation system.

This calls for an exploration of emerging and innovative technologies to enhance public safety at railway crossings, especially focusing on solutions that have the potential to increase the situational awareness of road users approaching at-grade crossings. The emergence of digital and cutting-edge technologies, such as multi-modal vision sensors including color (aka RGB), thermal, Light Detection and Ranging (LiDAR) and Radio Detection and Ranging (RADAR) cameras, sophisticated machine vision algorithms empowered by artificial intelligence (AI) techniques, as well as connected and automated vehicles has started to disrupt the ground transportation sector. A recent study by Deloitte [2] suggests that within the next 20-25 years, vehicles capable of performing most or all of the driving tasks will become commonplace, with many of them equipped with vehicle-to-everything (V2X) communication abilities, thereby causing a paradigm shift in ground transportation.

Therefore, by harnessing these technological advancements, it will become possible to alert road users about their proximity to rail crossings, the presence and speed of oncoming trains, and potentially hazardous conditions ahead at the crossing. This proactive approach will significantly enhance the situational awareness of road users as they approach a grade crossing. Consequently, this heightened awareness will ensure a safer passage through crossings, effectively reducing the occurrence of tragic accidents at these locations.

1.2 Project objectives

The overarching objective of this project was to improve safety at grade crossings by leveraging machine vision technologies to prevent collisions. In particular, the project focused on developing sophisticated algorithms using deep learning and AI techniques, drawing from real-world data collected at a public grade crossing. The intention was to automatically detect and recognize vehicles and vulnerable road users (VRUs), as well as identify potentially dangerous situations involving them at or near grade crossings. By integrating this capability with secure V2X communications, downstream applications could be enabled to warn vehicles and VRUs approaching at-grade crossings, aiming to lower the risk of serious accidents at grade crossings.

To that effect, the specific objectives of the project included the following:

- Conduct a comprehensive literature review to identify the root cause and contributing factors of grade crossing collisions in Canada.
- Build a large-scale, representative and multi-modal dataset containing labeled samples of vehicles, VRUs, and potentially unsafe events involving vehicles/VRUs based on data captured from a real-world public grade crossing.
- Analyze the collected data to validate the findings from the literature review, while also gaining insights into traffic dynamics and unsafe events involving vehicles and VRUs at, through, and around the grade crossing.
- Develop sophisticated machine vision algorithms to automatically detect and recognize vehicles, and VRUs, as well as identify potentially unsafe events involving them at, or near the grade crossing.

2 Review of grade crossing collisions and inspections

This section presents a study of published and publicly available reports on grade crossing collisions and inspections in Canada. The main objectives of this literature review encompassed the following:

- Review the latest evidence on the root cause and contributing factors of collisions from collision investigation reports, while also identifying primary safety concerns at-grade crossings from crossing inspection reports.
- Recognize the potential prospects for novel and emerging technologies, particularly secure V2X communications and advanced machine vision systems empowered by sophisticated AI techniques, to increase the situational awareness of road users in order to help:
    - mitigate collision risks between vehicles and trains, and
    - improve crossing safety of VRUs such as pedestrians, and cyclists.
- Summarize the findings to derive use cases for deploying emerging communications and machine vision technologies at-grade crossings.

The literature review focused solely on publicly available investigation reports on grade crossing collisions from the TSB as well as publicly available Notices and Orders issued by Transport Canada (TC) concerning safety at-grade crossings.

2.1 Canadian crossing collision reports – causation factors and potential countermeasures

This section provides information on the Canadian grade crossing collisions reviewed. For the purpose of this study, a total of 54 investigation reports that were completed and published by TSB since 1994 were reviewed. Of the 54 incidents, a brief summary is provided in sections 2.1.1 to 2.1.23, for those where the accident may have been avoided or the collision risks may have been mitigated by the use of the technology enablers as mentioned in the Introduction section. For each incident summarized, a commentary is also included in italics on how the technology enablers could have been leveraged to mitigate the collision risks.

Section 2.1.24 then provides a summary of the 23 crossing collisions. Table 1 in that section presents a summary of the collisions reviewed, and Table 2 there presents the causation factors to provide insights into the trends.

2.1.1 TSB investigation report # R21H0087 [3]

Background information: A cutaway van was struck by a VIA Rail passenger train at Smith Falls subdivision on 30 June 2021 at 1216 eastern daylight time (EDT) when the driver of the vehicle attempted to steer around the fully descended gates at a grade crossing. The vehicle's driver was fatally injured.

Causation: Since the crossing was equipped with *grade crossing warning devices* (GCWDs), the sightline was not maintained. This, when combined with the factors such as the acute crossing angle (38°) and the limited visibility to the sides of and behind the vehicle cab, caused the train to remain out of sight of the driver until it was too close for the driver to avoid the accident. Moreover, the activated GCWDs were found to be inconspicuous from a distance – the fully descended crossing gates blended into the background and the gate lights had reduced visibility due to the bright daylight conditions. Added to these factors was the driver's familiarity with the crossing where he had likely seldom encountered any trains, thereby had likely formed the expectation that there would not be any train at the crossing. A report from the University of Calgary [4] that was prepared for TC describes this is a human behavior that contributes to non-compliance of drivers with grade crossing stop signs.

Comments: *The use of advanced train detection technology coupled with secure V2X communications could have increased the situational awareness of the driver by having the status of the crossing and the approaching train displayed on a V2X onboard unit, thereby likely mitigating the risk of the collision.*

2.1.2 TSB investigation report # R20D0013 [5]

Background information: On 18 February 2020 at approximately 0922 eastern standard time (EST), a 70-year-old man, while taking a road test with an evaluator onboard, approached a grade crossing under reduced visibility conditions due to heavy snowfall and an overcast sky. The driver had significantly reduced lower binocular visual field. The car entered onto the railway right-of-way and stopped on the tracks while the GCWDs were activated. Constant instructions from the evaluator and a firefighter participating in a drill on the other side of the crossing as well as the horn of the vehicle behind were of no help for the driver to move forward and clear the tracks. The car was eventually struck by the commuter train EXO 182. The driver died while the other passenger of the car sustained serious injuries.

Causation: The report mentioned that the driver was likely overwhelmed by the many visual and auditory stimuli that he was subjected to simultaneously, including the flashing lights and moving gates of the activated GCWDs, the nearby firetrucks, the constant instructions and indications from the evaluator inside the car and the firefighter outside the car, as well as the horn of the vehicle behind. With a significantly reduced lower binocular visual field, the driver was likely confused with all these stimuli, especially under the reduced visibility conditions that prevailed during the incidence.

Comments: *With secure V2X communications systems installed at the grade crossing, a V2X onboard unit in the vehicle would have received the status of the crossing from as far as 1 km away* [6]*, well before the activated GCWDs would be conspicuous. This would have increased the situational awareness of the driver, especially under the reduced visibility conditions reported in this specific incident. This extra time, coupled with the fact that the V2X onboard unit likely needing less attention to comprehend the crossing status than the multiple external stimuli, would have likely helped the visually-challenged driver to react to the situation in a timelier fashion, thereby reducing the likelihood of the collision. Even if the driver would have found these future warnings confusing, the evaluator would have far more time to safely address the situation.*

2.1.3 TSB investigation report # R19T0191 [7]

Background information: On 13 November 2019 at about 1444 EST, 11 pedestrians including 6 adults and 5 children waited on the sidewalk of a double-track crossing equipped with GCWDs for a slowly

moving Canadian National Railway (CN) freight train 568 to clear the crossing. Immediately after the freight train had cleared the crossing, a faster Metrolinx train (GO 3919) approached the crossing on the farther track from the other end. Despite being aware of the activated GCWDs, two pairs of adult and children attempted to cross, with one pair making it to the other end, and the other pair getting struck by the GO 3919 train sustaining serious injuries.

Causation: The investigation mentioned that the adult pedestrians attributed the activated GCWDs solely to the already passed train CN 568 and did not recognize that it could also indicate the approach of another train on the other track. Furthermore, due to the slowly moving CN 568 to the west, they were not able to see the GO 3919 approaching from the east on the farther track. Additionally, just before the accident, the group had spent approximately 50 min outside. The children were growing cold, causing restlessness among the adults. As a result, there was a sense of urgency to return to their destination, further heightening the adults' determination to cross the grade crossing.

Potentially Hazardous Activities: The investigation also discovered the following potentially hazardous events at the subject crossing based on a 10 day video recording between 20 June 2020 till 29 June 2020. The safety risks associated with these activities are equally relevant to other grade crossings, especially in urban settings.

1. Due to the prolonged use of the crossing for switching activities by CN; pedestrians, cyclists and motorists were observed to exhibit a tendency to trespass into the crossing while the GCWDs were activated.
2. Vehicles were observed performing U-turns, some within 30 m (98 ft) of the crossing, which is a violation of the Highway Traffic Act.
3. Occasionally, due to traffic lights at a road intersection a short distance ahead, queues were observed to build up onto the rail crossing as motorists waited for the traffic lights.
4. Pedestrians were found to enter/exit the railway right-of-way at the crossing without authority.

Comments: *Using advanced AI technology utilizing multi-modal vision sensors deployed on the roadside infrastructure at the crossing, the pedestrians waiting at the crossing could have been reliably detected. This information, combined with approaching train location, number, speed and estimated arrival times could then be displayed on a large display via secure V2X communications, which would have likely increased the situational awareness of the pedestrians in real-time, thereby reducing the likelihood of the collision.*

2.1.4 TSB investigation report # R18V0127 [8]

Background information: On 26 May 2018 at about 1735 pacific daylight time (PDT), a pedestrian on a motorized wheelchair was struck by a CN train Q10521-21 travelling westward through Chilliwack, British Columbia, when the motorized wheelchair became immobilized on the crossing surface. The accident left the pedestrian with fatal injuries and resulted in serious injuries to one of the two motorists who attempted to rescue the pedestrian from the collision.

Causation: As the pedestrian approached the crossing, he stopped the motorized wheelchair in the vicinity of the crossing a number of times. When the train neared the crossing, the pedestrian, who was stopped with the rear caster wheels on the south rail clear off the flangeway, moved the wheelchair in the

opposite direction causing both rear caster wheels to fall into the flangeway gap of the south rail, eventually immobilizing the wheelchair.

Comments: *With the use multi-modal machine vision sensors combined with advanced AI technologies deployed on the roadside infrastructure at the crossing, such hazardous scenarios could be reliably detected. Upon detection, a V2X message could have been broadcast to alert the approaching train of the hazardous situation, allowing the train to apply emergency brakes as far as 1km away from the crossing, thus decreasing the likelihood or the impact of the collision.*

2.1.5 TSB investigation report # R18T0006 [9]

Background information: On 9 January 2018 at about 0940, a snowplow, while clearing snow from the sidewalk, approached an anti-whistling grade crossing having 4 sets of tracks with a total clearance distance of 85 ft. Unaware of the oncoming train, the snowplow operator entered into the crossing just before the GCWDs were activated and continued onto the track before being struck by a CN freight train. The operator was fatally injured as a result of the collision.

Causation: When the GCWDs were activated, the gate lights at the southeast quadrant where the snowplow was positioned were approximately 90° from the operator's forward view, while the gate lights on the northwest quadrant were pushed to the operator's peripheral view as he continued onto the track (Figure 1 left). This was further aggravated by the restricted sightline of the operator due to the A-pillar and the screening on the side windows of the cabin (Figure 1 right). Added to these were factors such as the background noise of the snowplow operation and the switching of the operator's focus between clearing snow from the forward at the sidewalk to dumping snow off to the right-hand side. All of these combined likely reduced the likelihood of the operator detecting the activated GCWDs and the approaching train, or hearing the horn of the approaching train.

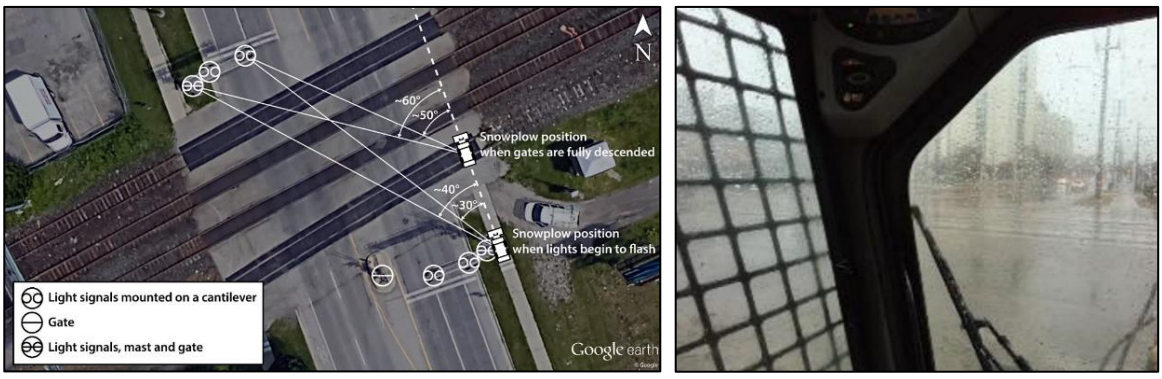

Figure 1: Location of the snowplow and viewing angles from the cab to the GCWDs (left image); (right image) Photo taken from a driver's seated eye position in the snowplow showing the view toward the northwest quadrant of the crossing when the snowplow was positioned adjacent to the southeast quadrant crossing warning system mast (the location of the snowplow when the gates began to descend); (Source: TSB)

Comments: *With secure V2X communications and advanced train detection technology deployed on the roadside infrastructure at the crossing, a V2X onboard unit in the snowplow showing the train approach speed and estimated arrival time could have enhanced the situational awareness of the snowplow*

operator. This is particularly true for the sidewalk snowplows, as these are typically low-ground tight-cabin vehicles, thereby possessing higher risks of blocking the operator's view as mentioned in this specific incidence.

2.1.6 TSB investigation report # R17H0015 [10]

Background information: At approximately 0732 EST on February 13, 2017, a school bus at the Town Line Road public crossing in Mile 121.36 of the Belleville Subdivision, near Colborne, Ontario, was struck by Canadian Pacific Railway (CP) freight train 142-12, which was traveling eastward. The school bus was immobilized at the crossing, which was equipped with flashing lights and a bell. Before the collision occurred, the bus driver and two occupants had already exited the bus and were positioned at a safe distance. The collision resulted in the destruction of the bus and the signal mast on the north side, while the locomotive sustained minor damage. Fortunately, no injuries were reported.

Causation: In accordance with the regulations in the Province of Ontario, school buses are required to come to a complete stop before railway crossings, even if the warning systems are not active. Adhering to this protocol, the school bus driver stopped the bus within 15 ft of the crossing, engaged the parking brake, opened the window and door to check for approaching trains, closed them, and released the parking brake. The driver then proceeded cautiously onto the crossing, slowing down to check for trains once again. Unfortunately, upon entering the crossing, the bus was unable to accelerate due to the rear tires lacking traction on the 1.95% inclined road, compounded by accumulated snow. Consequently, the bus slid sideways towards the road's edge where the snow was deeper, ultimately becoming immobilized. Unaware of the fact that the railway emergency contact information was posted at the crossing, the bus driver called the dispatcher of the bus company. The dispatcher dropped all calls, verified the location of the bus based on the bus GPS data, followed by consultation with a crossing reference document in an effort to find out the exact location of the crossing, the railway track owner and the crossing's emergency contact information. However, by this time the CP freight train 142 approached the crossing and struck the empty bus.

***Comments*:** *Immediate notification to the crossing's emergency contact could have avoided this collision. To this end, using machine vision sensors installed on the crossing infrastructure and with the utilization of advanced AI technologies, the mentioned scenario could have been reliably detected. Upon detection, an automated call to the crossing's emergency contact could have increased the chances of avoiding the collision, especially since a significant period of time elapsed between the school bus becoming immobilized on the crossing and the train arrival.*

2.1.7 TSB investigation report # R16M0026 [11]

Background information: At around 0143 atlantic daylight time on July 27, 2016, a tragic incident occurred on the Springhill Subdivision when CN freight train Q-12111-26 collided with a pedestrian in a wheelchair at the Robinson Street public crossing (Mile 124.43) in Moncton, New Brunswick. The crossing was properly equipped with flashing lights, a bell, and gates. Regrettably, the pedestrian suffered fatal injuries as a result.

Causation: The incident occurred when the pedestrian's motorized wheelchair became immobilized on the crossing. The paved surface of the crosswalk did not extend across the entire width of the sidewalk on the east side, creating an empty space. It is likely that the wheelchair was directed toward the reflective line on the sidewalk to avoid hitting the post of the grade crossing warning system. The edge of the sidewalk led directly to the empty space in the pavement. Since there were no reflective line markings on the newly paved section of the sidewalk, the crosswalk did not have sufficient visual cues for the pedestrian to navigate it safely at night. When the wheelchair's right caster wheel fell into the empty space in the sidewalk, the wheelchair got stuck in the flangeway gap, immobilizing the pedestrian.

Comments: *This incident bears a close resemblance to the one described in section 2.1.4. Therefore, any remarks or comments made regarding that incident are equally applicable to this one.*

2.1.8 TSB investigation report # R13T0192 [12]

Background information: On 18 September 2013 at about 0832 EDT, a double-decker OC Transpo bus left Fallowfield Station on the OC Transpo Transitway, traveling at approximately 43 mph. As it approached Mile 3.30 of VIA's Smiths Falls Subdivision, a VIA Rail Canada Inc. passenger train entered the crossing. The crossing lights, bells, and gates were activated at the time. The northbound bus, with the brakes applied, was moving at around 5 mph when it collided with the train. The collision resulted in the front of the bus being torn off. The train, consisting of 1 locomotive and 4 passenger cars, derailed but remained upright. The bus had 6 fatalities, 9 serious injuries, and around 25 minor injuries among its occupants. No injuries were reported among the train occupants.

Causation: According to the report's findings, the bus driver was driving at approximately 67.6 km/h, which exceeded the posted speed limit of 60 km/h, just before applying the brakes. This excessive speed resulted in an increased stopping distance requirement. Initially, the driver did not fully engage the brakes, further lengthening the stopping distance. The report suggests that the driver was likely distracted by multiple factors, including looking at the video monitor, conversing with passengers, and feeling the need to make an announcement to passengers on the upper deck as they were standing. Additionally, the report highlights that the view of the activated automatic warning devices was obstructed by trees, shrubs, foliage, and roadway signage along the Transitway until the bus was 122.5 m away from the crossing.

Comments: *A V2X onboard unit in the bus, capable of displaying the status of the crossing and the approaching train as far as 1 km away from the crossing, could have mitigated the visibility issue with the activated GCWDs as mentioned in the report, and would likely have enhanced the driver's situational awareness.*

2.1.9 TSB investigation report # R13E0015 [13]

Background information: On 24 January 2013 at 0856 central standard time, when a snow-clearing road grader had stopped on a public crossing equipped with a passive warning system, likely to adjust the grader blades, a CN freight train carrying dangerous goods struck the grader causing fatal injuries to the grader operator, heavy damage to the grader, while releasing approximately 106,000 L of crude oil.

Causation: The report indicates that the grader operator being likely focused on resetting the grader's blades for snow clearing in the vicinity of the crossing was contributory to him not being able to detect the approaching train. Moreover, the effectiveness of the locomotive horn was reduced mainly because, the locomotive was being operated in the long hood orientation (i.e., reverse direction), with the horn placed in the middle of the locomotive in a recessed location. Additionally, the background noise from the engine and the fans inside the closed cab of the grader further reduced the sound level of the horn, resulting in an audible warning of that lasted for less than 2 s.

Comments: *As commented on for a similar incident mentioned in Section 2.1.5, having a V2X onboard unit installed on the grader, which displays the train's status, could potentially have improved the grader operator's situational awareness about the crossing. This is particularly true as the operator was situated inside a closed cabin and encountered difficulties in effectively monitoring the visual and auditory signals originating from the external environment, which are currently essential for detecting the presence of a train.*

2.1.10 TSB investigation report # R13D0001 [14]

Background information: On January 9, 2013 at approximately 09:50 EST, VIA Rail Canada Inc.'s passenger train No. 601 collided with a westbound vehicle that was following another vehicle while traversing the public crossing called rang de la Deuxième-Chaloupe, located at Mile 98.79 near Joliette, Quebec. The first vehicle narrowly escaped the accident as the driver drove past the crossing mainly because he was not able to stop safely at the sightline distance. However, the driver of the second vehicle, following closely behind the first vehicle at a similar speed, proceeded onto the crossing and was subsequently struck by the train. Both occupants of the vehicle lost their lives in the collision.

Causation: The thick fog surrounding the crossing significantly diminished the visibility, limiting the vehicle drivers' ability to spot both the oncoming train and the activated crossing signals before reaching the crossing. Despite the train horn situated in the middle of the locomotive being sounded, it failed to effectively notify both the driver of the initial vehicle and the subsequent vehicle about the train's presence. After the first vehicle crossed the intersection, the driver of the second vehicle proceeded onto the crossing and tragically collided with the train.

Comments: *As commented on for a similar incident mentioned in section 2.1.2 in which heavy snow reduced the visibility, with secure V2X communications deployed at the crossing, a V2X onboard unit inside the vehicle displaying the status of the crossing and the approaching train from as far as 1 km could have potentially enhanced the situational awareness of the two vehicle drivers even under the foggy conditions, thereby reducing the likelihood of the collision.*

2.1.11 TSB investigation report # R11T0175 [15]

Background information: On July 29, 2011 at approximately 10:40 EDT, VIA Rail Canada Inc.'s passenger train No. 71, traveling in a westerly direction at a speed of 78 mph on the south main track of CN's Chatham Subdivision, collided with a pickup truck at the Pratt Siding Road crossing equipped with passive warning signs situated at Mile 30.62 near Glencoe, Ontario. The impact caused the locomotive and all four coaches to derail, with some coaches coming to rest where they were obstructing the north

main track. The pickup truck was completely destroyed, and the sole occupant succumbed to his injuries in the hospital. Six passengers of VIA Rail Canada Inc. sustained minor injuries.

Causation: As the report indicates, the pickup truck did not stop, but rather only slowed down at the crossing stop sign, and applied the brakes just prior to reaching the crossing upon noticing the train, eventually skidding onto the crossing before being struck by the train. Although the sightlines at the stop sign were adequate, the driver was not able to get an advance view of the train due to the presence of buildings and vegetation along the tracks and in the fields past the sightline distance. Moreover, the locomotive horn may not have alerted the driver since the auditory warning of the horn was reduced due to the train travelling at 80 mph and the horn being placed near the mid-point of the locomotive.

Comments: *Again, with secure V2X communications and advanced train detection technology deployed at the crossing, the driver could have received an advance warning of the approaching train on a V2X onboard unit, which could have compensated for the lack of the driver's advance view of the train due to the blocked view past the sightline distances. This would have increased the situational awareness of the driver, thus reducing the likelihood of the collision.*

2.1.12 TSB investigation report # R09V0219 [16]

Background information: VIA passenger train 198 struck a vehicle at the Dorman Road crossing, located at Mile 75.68 of the Victoria Subdivision on 14 October, 2009 at 1514 PDT. Following this accident, two occupants of the vehicle died while the third suffered serious injuries.

The level crossing signal lights and bells were activated once the train approached the crossing. At the same time, a southbound vehicle was turning west onto Dorman Road. Seeing the train at the last minute, the vehicle driver braked and stopped at the level crossing. The vehicle driver backed up a short distance and then drove forward again. It was then that the vehicle was struck by the train.

Causation: The driver likely did not notice the active level crossing signal until the vehicle began to turn. Given the orientation of the traffic lights for southbound drivers entering Dorman Road, the closest signal to the approach was parallel to the railroad tracks in the driver's peripheral view and was therefore, less apparent. Once the driver became aware of the flashing lights, the distance of 14 m separating the intersection from the stop line of the crossing was insufficient to allow the driver to stop the vehicle before it entered the level crossing.

Comments: *A V2X onboard unit showing the status of the crossing and the approach of the train would likely have increased the situational awareness, especially since the visual cues of the activated crossing signal was less conspicuous to the driver as stated in this specific incident.*

2.1.13 TSB investigation report # R08T0158 [17]

Background information: On 15 July 2008, at approximately 1525 EDT, VIA passenger train No. 60, traveling eastbound on the north main track of the CN Kingston Subdivision, struck a loaded tractor-trailer that was immobilized at the public crossing Quabbin in Mallorytown, Ontario. The train could not be stopped before hitting the tractor-trailer, even after activating the train brakes. The trailer and the equipment it was carrying were destroyed. The tractor driver was uninjured as he had left the cab just before the impact. The train driver and four passengers on board were slightly injured.

Causation: The collision occurred when the passenger train struck a lowboy truck which was immobilized on the crossing that had an excessively steep approach. The lowboy truck was preceded by a pilot vehicle whose driver did not notice that the crossing was unsuitable for a low ground-clearance vehicle, since there was no advanced signs or warning systems for such crossings. Although the truck remained stuck at the crossing for about 7 min before the train's arrival, none of the vehicle drivers noticed the emergency contact posted on the signal bungalow and therefore, were not able to inform the train company of the emergency situation.

Comments: *As commented for a very similar incident described in Section 2.1.6, where a school bus was immobilized on the crossing surface for quite a while before being struck by a train, automated detection of the event via the use of advanced AI technologies deployed at the crossing could have allowed to make an automated call to the railway company's emergency contact. This would have likely reduced the risk of the collision, especially since the train arrived 7 min after the vehicle got immobilized on the crossing.*

2.1.14 TSB investigation report # R08M0002 [18]

Background information: On 19 January 2008, at approximately 1105 EST, Chemin de fer de la Matapédia et du Golfe inc. freight train 403, traveling westbound on the Mont-Joli subdivision, struck a northbound minivan at the Route 291 level crossing, in the municipality of Saint-Arsène (Quebec). The minivan spun around and struck the signal mast positioned in the northwest quadrant of the crossing. Two occupants were fatally injured and a third person sustained serious injuries.

Causation: The minivan driver was alerted by the front seat passenger of the presence of the train, but it was already too late to avoid the collision. The train struck the minivan because the driver was unable to stop his vehicle before it entered the track, or accelerate to clear the track before the train passed. As the crossing lights, mounted on the mast below the warning cross, were pointed towards the roads approaching the crossing, it is possible that the driver unconsciously determined that the lights were not working, and believed that there was no train approaching. The noise that could be heard inside the minivan, coupled with the fact that the noise of the whistle and the train was muffled by residences built along the road, deprived the driver of additional clues that could have warned him of the imminent arrival of the train.

Comments: *A V2X onboard unit showing the status of the crossing and the approaching train would likely have increased the situational awareness of the driver approaching the grade crossing, thereby reducing the risk of this collision.*

2.1.15 TSB investigation report # R07D0111 [19]

Background information: VIA Rail Canada Inc. passenger train No. 35, traveling westbound at 62 mph on the south main track of the CN Kingston Subdivision, struck an empty tractor-trailer that was immobilized on the 3rd Avenue public crossing at Mile 23.57, near Pincourt/Terrasse-Vaudreuil, Quebec. The tractor-trailer was destroyed and the locomotive was damaged. The driver of the tractor-trailer suffered minor injuries. The train did not derail and the track was not damaged.

Causation: The VIA train No. 35 struck a tractor-trailer that had become stuck on the crossing after the rear axle of the trailer became stuck in a snow bank. The configuration, the reduced width and the icy surface of the roadway contributed to the immobilization of the tractor-trailer. Transport Quebec and CN guidelines for snow removal may be acceptable for most roads and crossings, but they were not appropriate for this crossing due to its particular configuration and the snow accumulations. Although the tractor-trailer remained immobilized on the crossing for about 4 min and 40 s, the VIA 35 crew did not receive an emergency radio message until the train was about 30 s from the crossing, and as such, the collision could not be avoided.

Comments: *As commented on in Sections 2.1.13 and 2.1.6 for similar incidents, the time duration of 4 min and 40 s for which the tractor-trailer remained immobilized on the crossing before being struck by the train, was sufficient for an automated emergency call to be placed to notify VIA 35, upon detection of the immobilized vehicle via the use of advanced AI technologies based on multi-modal vision sensors deployed on the crossing.*

2.1.16 TSB investigation report # R05E0008 [20]

Background information: On 31 January 2005 at 1310 mountain standard time (MST), VIA Rail Canada Inc. passenger train No. 1, was struck by a logging truck traveling south on Secondary Highway 751 at the public level crossing at Mile 92.26 of the CN Edson Subdivision. The crossing was protected by flashing lights, signals and bells, and by advance warning signals which were placed approximately 104 m from the crossing. The collision caused the derailment of the train's two locomotives and nine cars. The driver of the truck suffered serious injuries and was taken to hospital. Some 6,500 L of fuel spilled from the lead locomotive.

Causation: The logging truck driver apparently did not notice the warning signals or identify the approaching train until the collision became impossible to avoid. The driver's physiological condition (he was affected by hyperglycemia, fatigue and dehydration) likely impaired his ability to recognize the warning signals and the approaching train, and to react accordingly. The driver was in a physical and mental condition that made him unfit to drive a vehicle. In addition to the driver's physiological state, there were a number of distractions in the cabin that might have prevented him from fully concentrating on the road.

Comments: *A V2X onboard unit in the vehicle displaying the status of the crossing signals and the approaching train would likely have enhanced the situational awareness of the driver about the crossing, especially since the driver had difficulties processing external stimuli and cues to recognize the crossing status.*

2.1.17 TSB investigation report # R04C0110 [21]

Background information: On 24 October 2004 at 0138 mountain daylight time, CP freight train 2nd 269-23, traveling south in heavy foggy conditions towards Lethbridge, Alberta, was struck by the second truck in a convoy of three eastbound loaded cattle-liners as they passed the Highway 23 crossing at Mile 69.33 of the Aldersyde Subdivision. As a result of the collision, a tank car loaded with anhydrous ammonia derailed and sustained damage, and five other tank cars were damaged. No product was spilled. As a

result of the collision, the train separated, the road tractor was destroyed and a fire broke out. The driver of the truck suffered serious injuries and his assistant lost his life. The railway track was damaged for a distance of approximately 2,060 ft (655 m).

Causation: The severe foggy conditions prevented the observation of the railway crossing ahead signal, the railway crossing warning system and the reflective signs affixed to the cars of the train that occupied the crossing. The urge to stay with the convoy to be able to benefit from the experience of the other drivers may have led the driver of the second truck to continue driving at a speed inappropriate for the prevailing visibility conditions. Moreover, the truck was not inspected before the trip to ensure the effectiveness of it's trailer brakes, and was found to have no effective brakes as concluded by the inspectors. After approaching the crossing at a speed of 20 km/h, the truck could not have been stopped before the crossing, even if the driver had been alert and the trailer brakes had worked with all due efficiency on a road surface in good condition, given the visibility conditions (approximately 9 m) that existed at the time of the accident.

Comments: *A V2X onboard unit inside the truck showing the status of the crossing and the information about the approaching train as far as 1 km away from the crossing would have compensated for the poor visibility conditions that were prevailing in this specific incident, thereby increasing the chance of avoiding the collision.*

2.1.18 TSB investigation report # R04H0009 [22]

Background information: On 28 June 2004 at 1836 EDT, VIA Rail Canada Inc. passenger train No. 49, travelling westward at 93 mph, struck an empty 10-ton dump truck at the public crossing at Mile 17.88 of the Smiths Falls Subdivision, near Munster, Ontario. The truck was destroyed, and the occupant was fatally injured. VIA crew members and train passengers were not injured.

Causation: Absence of visual or audible cues to alert the driver of the presence of the approaching train caused the collision. The sharp angle of the crossing, combined with the inability to look back through the rear window since it was a dump truck, prevented the driver from seeing the approaching train. Due to the normal level of ambient noise in the truck cab, the driver was only able to hear the train horn for about 1 s before the train entered the crossing.

Comments: *Increasing the situational awareness of the driver via the use of V2X onboard unit displaying the status of the train and the crossing could have compensated for the blocked view of the driver and likely have reduced the risk of the collision.*

2.1.19 TSB investigation report # R02T0149 [23]

Background information: On 13 May 2002, at approximately 0915 EDT, a VIA Rail Canada passenger train (No. 52) struck an immobilized tractor-trailer at the public crossing at Mile 181.71, Kingston, Ontario. Though the train crew had applied the brakes, the train was unable to stop before the collision. While the truck occupants exited the tractor and escaped uninjured, the locomotive engineer suffered minor injuries.

Causation: The TSB report indicates that the accident happened at a crossing where low clearance trailers had a high probability of grounding out. Also, the geometry of the crossing, including the

intersection of the two superelevated tracks through the road's vertical curvature and the irregularity of the surface profile, created a hazard for low-clearance trailers. It is worth mentioning that there was no indication that the crossing would present a risk, though there were bump signs installed near the grade crossing.

Comments: *Although an emergency number was stenciled on the crossing sign, none of the vehicle occupants noticed the number and as such, no emergency call was made to the railway company. As mentioned in similar incidents described in Sections 2.1.6, 2.1.13, 2.1.15, upon detection of the vehicle stuck on the crossing surface via the use of advanced AI technologies based on multi-modal vision sensors deployed on the crossing, an automated call to the emergency number could have been placed, thus reducing the likelihood of this type of incident.*

2.1.20 TSB investigation report # R00C0159 [24]

Background information: On 19 December 2000, at approximately 2037 MST, two consecutive accidents occurred at a passive crossing of the Waterways Subdivision involving a freight train (Nº 590-19) from Athabasca Northern Railway Ltd. and two ground vehicles, when the freight train was travelling southward over the secondary Highway 881 crossing at Mile 138.07. There was no illumination in the neighborhood of the crossing when the first accident happened, and the 21st car of the train was struck by a semi-trailer truck. The second fatal accident occurred approximately 3 h later (while the train was still occupying the crossing) and there were no warning devices/signage to show that the crossing was still occupied by the train. The second accident occurred between the standing freight train and a truck and left the driver of the truck fatally injured.

Causation: The investigation revealed that the illumination conditions in the general vicinity of the passive crossing were not sufficient for drivers to detect the presence of the train on the crossing, which contributed to both accidents. The report indicates that though the driver of the second truck was on-duty for a long number of hours (76.5 including a one day of rest) during the week that preceded the accident, it does not really mention that fatigue was the primary cause behind any of the two accidents. However, It was indicated that long working hours might affect driver's performance.

Comments: *The use of emerging train detection technology together with V2X communications could have mitigated the risk of this accident, as the status of the train could be displayed on a V2X onboard unit in the vehicle, thereby compensating for the visibility issues due to the poor lighting conditions that prevailed in this specific incident.*

2.1.21 TSB investigation report # R99T0298 [25]

Background information: On 23 November 1999, at approximately 1845 EST, a CN freight train (No. M-321-21-22), travelling westward on the north main track, struck an abandoned tractor-trailer at a farm level crossing at Mile 292.59 of the CN Kingston Subdivision and dragged it for approximately 2,000 ft along the track. A VIA Rail passenger train No. 68 (VIA 68), travelling eastward on the south main track, struck the debris and derailed, just before the freight train had come to a halt. Although there were no fatalities (only minor injuries), some important damages were reported, including 6800 L of diesel that spilled from the VIA locomotive and burned, and 4550 L that spilled from the lead CN locomotive.

Causation: The report concluded that the truck driver did not take appropriate actions when his tractor-trailer got stuck in the wooden ties of the track. It was also reported that if the truck driver had been aware about the emergency communications tools available onsite, the accident would never have happened. The report also indicated that the signage and maintenance of many existing private and farm crossings were of a lower standard than those of public road crossings.

Comments: *This incident is very similar to the ones reported in sections 2.1.6, 2.1.13, 2.1.15, 2.1.19, and therefore, the comments made on those incidents equally apply to this one.*

2.1.22 TSB investigation report # R99S0100 [26]

Background information: On 9 November 1999, at approximately 0900 EST, a VIA Rail Canada Inc. passenger train (No 85) fatally collided with a dump truck on Fourth Line Road in the community of Limehouse, Ontario. The locomotive and four passenger coaches derailed on the main track, some of them were extensively damaged, and the truck driver sustained fatal injuries.

Causation: The report indicates that the visual cues available to the truck driver were not sufficient to see the approaching train through the window and were partially obstructed by the roof pillar and side mirror.

Comments: *The use of advanced train detection technology coupled with V2X communication systems deployed on the crossing would have likely increased the situational awareness of the driver by having the status of the train displayed on a V2X onboard unit as far as 1 km away from the crossing, thereby mitigating the risk of such collisions.*

2.1.23 TSB investigation report # R99H0009 [27]

Background information: At approximately 0804 EDT on 14 July 1999, a VIA Rail Canada Inc. train (No. 2) struck the rear portion of an empty tractor-trailer at a private crossing near Hornepayne, Ontario. Three people were seriously injured and eight were taken to the Hornepayne community hospital.

Causation: The report indicates that because of the noise inside the truck's cab, the driver did not hear the train whistle nor observed the approaching train. The report also stated that if the driver had stopped at the stop sign and looked down the track, he would have been able to spot the train, considering the visual cues and the presence of a stop sign at the Becker crossing.

Comments: *Use of V2X communications to notify the driver about the status of the train would have provided greater visibility to the driver, and reduced the likelihood of this type of accidents.*

2.1.24 Summary of the crossing collisions

The 23 crossing collisions reviewed are summarized in Table 1. The collision reports demonstrated a variety of contributing factors. Table 2 summarizes eight contributing factors involved in the collisions, while also listing the total number of fatalities and serious injuries incurred in the collisions where the corresponding factor had a role to play. Caution is advised when interpreting the casualty numbers associated with the factors, as the Table indicates the potential involvement of multiple factors in a collision.

Table 2 shows that human factors (e.g., mis-judgement of the situation at hand, negligence, failing to follow instructions etc.) were involved with other factors in almost all (18 of 23) of the collision reports. The exceptions were for a few where crossing geometry (4 of 23), track flangeway gap (2 of 23), coupled with weather elements (e.g., snow/ice, 2 of 23), were exclusively contributory to the incidents. In some collisions reviewed, the fact that the driver's view was blocked due to vehicle design (e.g., cutaway van, dump truck) or some transitory obstructions (e.g., roadway signage blocking the view of the crossing signal from a specific spot on the road approach) was also cited to be contributory, though these did not cause any instance of sightline blocking. Visibility issues due to adverse weather and illumination conditions were involved along with other factors in 4 of the 23 collisions reviewed. Interestingly, in 4 out of the 23 collisions reviewed, the level of auditory warning from the locomotive horn was found to be insufficient and was cited as a contributing factor.

Table 1: Summary of the crossing collisions reviewed. Best viewed when zoomed in.

Report section	TSB report #	Crossing location	Province	Year	Mon.	Local time	Crossing GPS coordinates	# Tracks	GCWDs	GCWD activation prior to train arriving at crossing	Road-way layout	Side-walks	Crossing angle	Traffic posted speed limit	Vehicle speed(s)	Involved vehicle(s)	Train	Train speed at time of collision	Weather	Comment
2.1.1	R21H0087	Rural	ON	2021	6	1216	45.198527 -75.816216	single	flashing light bell gates	35 s (18 s due to hill)	2-lane asphalt	no	38°	80 km/h	20 km/h	2017 GMC 3500 Cutaway van	Passenger	85 mph (137 km/h)	27.9°C	The vehicle then started to descend the hill, which had an average gradient of 4% before levelling out about 475 ft from the crossing.
2.1.2	R20D0013	Urban	QC	2020	2	922	45.546107 -73.697588	double	flashing light bell gates	23 s	2-lane asphalt	yes	15°	30 km/h	0 km/h	2016 4-door passenger sedan	Passenger	42 mph (68 km/h)	−10°C overcast, snowing heavily	
2.1.3	R19T0191	Urban	ON	2019	11	1444	43.459226 -80.482799	double	flashing light bell gates	Not reported	4-lane asphalt	yes	<80°	50 km/h	n/a	Pedestrians	Passenger	not reported	-4.9°C	Pedestrians positioned on track >5 min prior to impact.
2.1.4	R18V0127	Urban	BC	2018	5	1735	49.166952 -121.935351	single	flashing light bell gates	Not reported	2-lane asphalt	yes	74°	50 km/h	n/a	Motorized wheelchair	Freight	47 mph (76 km/h)	18.6°C	Wheelchair positioned on track ~ 1 min prior to impact.
2.1.5	R18T0006	Urban	ON	2018	1	940	42.983416 -81.23817	quad	flashing light bell gates	29 s	4-lane asphalt island median	yes	<90°	40 km/h	1.4 km/h	Bobcat S130 skid-steer loader equipped with a blower	Freight	44 mph (70 km/h)	−4°C	Vehicle 4.6 m beyond the crossing mast when the gates began to descend.
2.1.6	R17H0015	Rural	ON	2017	2	732	43.988819 -77.902564	single	flashing lights bell	33 s	2-lane asphalt	no	69° N 75° S	50 km/h	0 km/h	2008 International PB10500 school bus (immobilized)	Freight	52 mph (84 km/h)	−3°C snowing	
2.1.7	R16M0026	Urban	NB	2016	7	143	46.089923 -64.780313	single	flashing light bell gates	Not reported	2-lane asphalt	yes	34°	50 km/h	n/a	Motorized wheelchair	Freight	30 mph (48 km/h)	19°C mostly overcast	Train engineer viewed wheelchair 27 s prior to impact.
2.1.8	R13T0192	Urban	ON	2013	9	848	45.302801 -75.734072	single	flashing light signals bell gates	49 s	2-lane asphalt (dedicated transitway)	no	50°	60 km/h	4.8 mph	ADL E500 double-decker transit bus	Passenger	43 mph (70 km/h)	14°C sunny, clear visibility	Vehicle initiated emergency brake prior from crossing at speed of 67.6 km/h. Crew members first noticed the bus travelling northward toward the crossing when the train was approximately 600 ft (183 m) from the crossing. Gates fully horizontal for at least 26 s before the accident
2.1.9	R13E0015	Rural	SK	2013	1	856	53.010506 -108.931129	single	crossbucks stop sign (train horn/bell)	Not reported	2-lane crushed stone	no	<90°	50 km/h	stationary on track	2009 Volvo Model G970 road grader	Freight	37 mph (60 km/h)	-17°C overcast, good visibility	Locomotive horn was sounded 4 times starting approximately 1,780 ft and ending 718 ft from crossing.
2.1.10	R13D0001	Rural	QC	2013	1	950	46.058696 -73.396638	single	flashing light signals bell	~8.25 s [A]	2-lane asphalt	no	35°/145°	80 km/h	50 km/h 60 km/h	#1 Passenger vehicle (leading) #2 SUV (struck)	Passenger	60 mph (97 km/h)	-3°C heavy fog, road wet	Visibility approximately 150 to 200 ft in the vicinity of the crossing.
2.1.11	R11T0175	Rural	ON	2011	7	1040	42.715926 -81.751986	double	SRCS stop sign crossing ahead No stop line	~22 s [B]	gravel (paved near crossing)	no	90°	50 km/h	not stated	Pick-up truck	Passenger	80 mph (129 km/h)	29.5°C	
2.1.12	R09V0219	Urban	BC	2009	10	1514	49.198055 -123.984844	single	flashing signals bells cross bucks	25 s	2-lane asphalt (arterial)	no	90°	50 km/h	not stated	1990 Sedan	Passenger	45 mph (72 km/h)	15°C	
2.1.13	R08T0158	Rural	ON	2008	7	1525	44.479969 -75.885458	double	flashing lights bell gates	Not reported	2-lane asphalt	no	65°	50 km/h	stationary on track	1997 Mack CH613 48-foot, model TC3 (lowboy) (tractor/trailer combination 60 ft)	Passenger	83 mph (134 km/h)	21°C part cloudy, vis 24 km	Train crew observed obstacle prior to event. Train travelled 2,409 ft from the time when the brakes were applied to the point where it struck the truck

2.1.14	R08M0002	Rural	QC	2008	1	1105	47.911869 -69.431832	single	flashing light bell (cantilever)	33 s	2-lane asphalt (arterial)	no	90°	50 km/h	50 km/h (est)	Minivan	Freight	51 mph (82 km/h)	-6°C bright, thin cloud, visibility good	
2.1.15	R07D0111	Urban	QC	2007	12	1505	45.387168 -73.991254	double	flashing light bell gates	Not reported	2-lane asphalt	no	<90°	50 km/h	stationary on track	1999 Volvo VNL64 w/ Manax trailer (53 ft)	Passenger	74 mph (119 km/h)	-11°C clear, snow cover/icy road	Truck immobilized on crossing 3 min 40 s before collision. Rear undercarriage of the trailer got stuck on a large snow bank. Winter storm previous day (32 cm)
2.1.16	R05E0008	Rural	AB	2005	1	1310	53.644212 -115.58228	double	flashing lights bell gates	23 s	2-lane asphalt	no	90°	100 km/h (reduced 50 km/h at 384 m from crossing)	~50 km/h	1994 Freightliner w/40-foot log trailer (Peerless Page HRL Tri-Dem)	Passenger	70 mph (113 km/h)	-5°C clear, sunny	Locomotive horn for approximately 15 s before entering the crossing. The bell was also sounded.
2.1.17	R04C0110	Rural	AB	2004	10	138	50.572085 -113.538025	single	flashing light signals bell	Not reported	2-lane gravel	no	82°	100 km/h	unknown	Class 8 Cattle-liner truck (loaded)	Freight	45 mph (72 km/h)	-4°C thick fog, visibility 9 m	Yellow AWS sign (287 m prior to crossing) X painted on pavement (287 m prior to crossing)
2.1.18	R04H0009	Rural	ON	2004	6	1836	45.118361 -75.870432	single	reflective crossing sign	locomotive horn (whistle) sounded for 10 s before crossing	2-lane gravel	no	52°	50 km/h	16 km/h	1989 Ford LNT 8000 dump truck	Passenger	93 mph (150 km/h)	20°C partly cloudy, clear visibility	
2.1.19	R02T0149	Rural	ON	2002	5	915	44.230116 -76.632199	double	flashing lights bell gates	Not reported	2-lane gravel	no	80°	50 km/h	stationary on track	1994 Freightliner, model FL 120 w/ 48-foot, low-boy trailer (TC3)	Passenger	77 mph (124 km/h)	8°C overcast, lt rain, vis 11.2. km	Train had travelled 2,409 ft from the time when the emergency brakes
2.1.20	R00C0159	Rural	AB	2000	12	2037	55.006129 -111.731985	single	flashing lights bell	Not reported	2-lane asphalt	no	67°	80 km/h	ditched 68 km/h	Truck#1: 1981 Kenworth, w/trailer Truck#2: 1997 Freightliner w/B-train	Freight	23 mph (37 km/h)	-20°C clear, snow cover/icy road	truck #1: 200 m from the crossing saw train/braked/ditched truck Warning railway track now in use sign (325 m prior to crossing) Advance warning sign (205 m prior to crossing)
2.1.21	R99T0298	Rural	ON	1999	11	1845	43.882047 -78.704529	double	none	Not reported	1-lane dirt	no	n/a	n/a	stationary on track	Class 8 Highway tractor w/ trailer	Freight Passenger	59 mph (95 km/h)	12.5°C	
2.1.22	R99S0100	Rural	ON	1999	11	900	43.630565 -79.995784	single	flashing lights bell signage	24 s	2-lane asphalt	no	78°	80 km/h	~80 km/h	1975 International Paystar dump truck	Passenger	63 mph (101 km/h)	10°C clear visibility	
2.1.23	R99H0009	Rural	ON	1999	7	635	49.201136 -84.669255	single	stop sign danger high speed trains (3 m before crossing)	Not reported	1-lane dirt	no	<90°	30 km/h	15 km/h	1990 International Sleeper w/ log trailer (~50ft)	Passenger	53 mph (85 km/h)	25°C clear visibility	Average clearance time 23 s (loaded) and 17 s (empty) trucks (w/ vehicle stopped 8 m from the track until it was clear at a point 8 m on the opposite side of the track)

[A] NRC calculation of time using TSB reported GCWS visibility at 600 ft
[B] NRC calculation of time using TSB reported GCWS visibility at1000 ft

Table 2: Summary of the factors contributing to crossing collisions. Caution is advised when interpreting the casualty numbers, as multiple factors were potentially involved in a collision.

Collision contributing factors	No. of fatal injuries	No. of serious injuries	Reference to the collisions
Human factors	21	18	2.1.1, 2.1.2, 2.1.3, 2.1.5, 2.1.8, 2.1.9, 2.1.10, 2.1.11, 2.1.12, 2.1.13, 2.1.14, 2.1.16, 2.1.17, 2.1.18, 2.1.20, 2.1.21, 2.1.22, 2.1.23
Crossing geometry (e.g., steep approach gradient, uneven crossing surface, acute crossing angle)	2	0	2.1.1, 2.1.6, 2.1.7, 2.1.13, 2.1.15, 2.1.19
Track flangeway gap	2	1	2.1.4, 2.1.7
Visibility issues due to adverse weather conditions (e.g., heavy snowfall, dense fog) and poor lighting conditions	5	2	2.1.2, 2.1.10, 2.1.17, 2.1.20
Weather elements (e.g., snow, ice)	0	0	2.1.6, 2.1.15
Driver's view blocked due to vehicle type (e.g., cutaway van, dump truck), and other obstructions (e.g., building, roadway signage). This does not include any instance of sightline blocking.	10	9	2.1.1, 2.1.5, 2.1.8, 2.1.18, 2.1.22
Audibility issues with locomotive horn	5	1	2.1.9, 2.1.11, 2.1.14, 2.1.18
Driver health issues (both physical and mental)	1	1	2.1.2, 2.1.16

2.2 Canadian crossing inspection reports – Notices and Orders

This section provides information on the *Notices and Orders* [28] issued by TC to railway companies, railroad authorities, or individuals following the identification of conditions or hazards during routine rail safety inspections. These conditions or hazards had the potential to cause harm to individuals, damage property, or could have had an impact on the environment. A *Notice and Order* was (and is) revoked once the immediate threat to safety due to the identified condition or hazard was (and is) removed to the satisfaction of a rail safety inspector.

For the purpose of this study, a total of 92 *Notices and Orders* were reviewed that were issued between 2009 and February 2024. A brief summary of those that were related to safety concerns at-grade crossings are presented below.

2.2.1 Issued to: CP; issue date: 2022-12-09; revocation date: 2022-12-19

Condition or hazard observed: Due to a CP lead locomotive stopping on a public crossing and blocking it for a prolonged period of time, pedestrians were observed to take risks by crossing between the stopped railcars, which could have led to a person being injured.

Order: CP was ordered to physically verify that the train was clear of any pedestrians before allowing the train to proceed.

2.2.2 Issued to: CP; issue Date: 2022-06-21; revocation Date: 2022-06-23

Condition or hazard observed: Vehicles, especially larger vehicles, were observed to queue up on the crossing at Mile 95.92 of the CP Cascade Subdivision due to the limited vehicular storage space and a narrow road approach. This posed an immediate risk of vehicles being struck by railway equipment or oncoming traffic.

Order: CP received an order to provide a stop and then proceed on the crossing until evidence of vehicles queuing over the crossing no longer existed.

2.2.3 Issued to: Municipality of Merrickville-Wolford and the United Counties of Leeds and Grenville; issue date: 2021-09-08; revocation date: 2021-12-03

Condition or hazard observed: Vehicles, especially larger vehicles, were observed to queue up on the crossing on Kilmarnock Road due to the limited vehicular storage space and a narrow road approach. This posed an immediate risk of vehicles being struck by railway equipment or oncoming traffic.

Order: The Municipality of Merrickville-Wolford and the United Counties of Leeds and Grenville was issued an order to limit access to the road section leading to the crossing, including the adjacent intersection of County Road 17 (Jasper Road), exclusively for vehicles shorter than 10 m (e.g., medium single unit trucks), until any evidence of vehicles forming a queue over the crossing no longer existed.

2.2.4 Issued to: CN; issue date: 2016-02-19; revocation date: 2017-07-07

Condition or hazard observed: Prolonged blocking of a public crossing by CN trains at Mile 8.8 of the CN Halton subdivision at Goreway Drive in Brampton, Ontario, aggravated by CN’s reluctance to promptly address the issue, resulted in dangerous driving behaviour at the crossing. Additionally, this situation could also cause delays for emergency vehicles passing through the crossing, potentially causing harm to an individual.

Order: CN was ordered to guarantee that every train arriving at Brampton Intermodal Terminal was granted an unobstructed route, with all switches aligned directly to their designated yard track. This was to enable the trains to proceed seamlessly until the crossing was clear. Additionally, CN was to provide to TC a weekly download of the event recorder data from the GCWDs at the subject crossing.

2.2.5 Issued to: CN; issue date: 2015-07-02; revocation date: 2015-08-28

Condition or hazard observed: There was an imminent threat of individuals getting struck by trains in the Souligny area of the Longue-Pointe Spur location at Quebec as a large number of intruders were observed to cross the tracks, slipping through the holes in the fencing.

Order: CN received an order to reintroduce the whistling policy at the following six grade crossings in the Rue Souligny area of the Longue Pointe Spur:

- Avenue Georges V (4.98)
- Avenue Hector (5.08)
- Boulevard Pierre Bernard (5.58)
- Rue Des Ormeaux (5.78)
- Avenue Lebrun (5.94)
- Rue Honoré Beaugrand (6.39)

2.2.6 Issued to: CP; issue date: 2015-1-19; revocation date: 2015-12-03

Condition or hazard observed: Due to high density of pedestrian traffic at the location of a public grade crossing situated at 3rd Avenue – CP Mile 0.10, there was a heightened risk of pedestrians exhibiting potentially dangerous crossing activities, including going through the crossing while it was deployed, running in front of approaching trains, or even crawling beneath the railway equipment of stopped trains on the crossing. The situation was further exacerbated by trains stopping on and past the crossing island circuit due to crew changes, thereby causing the GCWDs to remain activated for a prolonged period of time.

Order: CP was ordered to ensure that no eastbound railway equipment blocked the grade crossing except for an emergency, while also making sure that crew changes took place at the 2nd Avenue – CP Mile 0.21 Thompson Subdivision in a manner that it did not cause the crossing to be activated continuously. Additionally, no eastbound railway equipment was to be operated past this crew change location after stopping there until there was a DTMF crossing gate activation following a wait of at least 25 s since the crossing gate had become parallel to the ground. Moreover, CP and the city of Kamloops, BC were ordered to deploy at least two flag persons on each side of the tracks at the grade crossing holding stop signs when the GCWDs were activated to make sure that no pedestrian could pass through the grade crossing.

2.2.7 Issued to: Resort Municipality of Whistler, BC; issue date: 2013-05-07; revocation date: 2013-05-30

Condition or hazard observed: Queuing over a public crossing at Squamish Subdivision was observed due to a newly added 3-way stop. This significantly increased the risk of a vehicle-train collision, particularly since the sightlines for traffic approaching the crossing travelling north-west were not clear.

Order: The resort municipality of Whistler, BC was ordered to cease the use of the 3-way road crossing and not allow any motorists, pedestrians, and cyclists to use the railway crossing from 7:00 am until 6:30 pm throughout the week except for statutory holidays, unless a qualified flag-person was provided at the south vehicle stop bar at the crossing. The flag person, upon activation of the GCWDs, would prevent

traffic, cyclists, and pedestrians from advancing onto the crossing from the south until the train had completely cleared the crossing.

2.3 Rail safety use cases

By examining the crossing collision scenarios outlined in Section 2.1 and the hazardous conditions observed at-grade crossings as outlined in section 2.2, the primary instances of risky situations and potentially dangerous actions exhibited by motorists, cyclists, and pedestrians as they traversed through the railway crossings were identified. This analysis serves to mirror actual rail safety concerns in the real world, thus shaping the rail safety use cases and also establishing the basis for data collection, analysis, and assessment undertaken in this project.

Sections 2.3.1 to 2.3.6 below present six rail safety use cases, highlighting how various technology enablers, including secure V2X communications, advanced train detection, and AI-empowered machine vision sensors deployed on roadside infrastructure at-grade crossings, could reduce collision risks, increase situational awareness of road users, and thereby enhance safety at-grade crossings.

2.3.1 Pedestrian approaching/waiting at a crossing

Use case scenario: This pedestrian-approaching/waiting-at-a-crossing use case scenario was derived directly from the crossing collision described in section 2.1.3, specifically addressing the circumstance where pedestrians approach a grade crossing, or wait at a crossing for the train(s) to clear the crossing. Given that virtually all collisions involving a moving train and pedestrians typically result in fatalities and/or serious injuries, as was the case in the referenced collision, enhancing the pedestrians' situational awareness using technology enablers will contribute positively to enhancing pedestrian safety at-grade crossings.

Technology enablers to enhance crossing safety: *Pedestrians approaching or waiting at-grade crossings can be detected and tracked using advanced AI technologies based on multi-modal vision sensors installed on the roadside infrastructure at the crossing, which can adapt to varying weather and lighting conditions. Upon detection of pedestrians, information about the number of approaching trains, train speed, estimated arrival time as obtained through advanced train detection systems can be presented on a large display board located on sidewalks via secure V2X communications. Alternatively, the information can be directly transmitted to pedestrians using vehicle-to-pedestrian (V2P) technology. The goal is to provide real-time updates about the crossing dynamics to pedestrians in an effort to enhance their situational awareness.*

2.3.2 Vehicle/assistive device immobilized on the crossing

Use case scenario: This use case scenario involves a situation where a vehicle or a user of an assistive device (such as a motorized wheelchair) becomes stuck on the crossing surface due to various reasons as mentioned in the TSB investigation reports outlined in sections 2.1.4, 2.1.6, 2.1.7, 2.1.13, 2.1.15 and 2.1.19. This occurrence gives rise to a potentially dangerous situation.

Technology enablers to enhance crossing safety: *With the deployment of secure V2X communications and advanced machine vision sensors equipped with sophisticated AI technologies on the roadside infrastructure at-grade crossings, hazardous scenarios such as a wheelchair/vehicle*

immobilized on the crossing surface could be reliably detected under varying weather and lighting conditions. This critical piece of information could then be leveraged in an effort to avoid a potential collision as follows.

- ***When no trains are approaching:***
  - *An automated call/notification could be made to the emergency contact available at the crossing upon detection of such a hazard using the technology enablers mentioned above. This could be potentially life-saving when there is no one around to help a pedestrian with an assistive device immobilized on the crossing (e.g., night time, or a remote location). Moreover, as reported in most of the referenced incidents, the vehicles were struck by the train even though the train was not approaching immediately, mainly due to the vehicle operators' lack of knowledge about the emergency contact information. Therefore, such collisions could potentially have been avoided if an automated call was made to the emergency contact upon automated detection of such events.*

  - *A V2X SOS message* [29] *could be broadcast to draw the attention of drivers of nearby vehicles (even pedestrians via V2P) to the emergency in order to allow them to come to the rescue of the trapped pedestrian/vehicle.*

- ***When one or more trains are approaching:***
  - *While the existing framework does not incorporate the train in the communication loop, it makes a compelling argument to incorporate the approaching train in the loop so that this hazardous situation could be communicated directly to train operators as far as 1 km away* [6] *from the crossing via V2X communications. This would allow the emergency brakes to be engaged well in advance of the hazard being visually detected by the locomotive engineer, thereby increasing the chance of avoiding an imminent collision.*

2.3.3 Queue buildup on the crossing

Use case scenario: This use case is where vehicles (rail, road or both) form a queue on the crossing surface. Queuing of vehicles on both road approaches, particularly on the egress of the crossing, can potentially lead to queuing on the crossing surface. This is particularly likely to happen when there is limited storage space available for the vehicles, or when long vehicles such as bendy buses and trailer trucks are involved. References to this use case can be found in sections 2.1.3, 2.2.2, 2.2.3, and 2.2.7.

Technology enablers to enhance crossing safety: *By utilizing advanced AI technologies based on multi-modal vision sensors installed at roadside infrastructure at a crossing, queuing of vehicles can be detected on both road approaches, particularly on the egress. The primary objective is to alert approaching vehicles via the use of secure V2X communications about the potential hazard of forming a queue on the crossing surface. Additionally, this information has the potential to be utilized proactively to pre-empt any appropriate traffic signals along the road(s) leading to the ingress, in an effort to reduce the number of vehicles joining the queue on the egress.*

2.3.4 U-turn on the crossing

Use case scenario: This use case was established on the basis that performing a U-turn on or within 30 m of a railway crossing creates a potentially dangerous situation as indicated in the TSB investigation report in section 2.1.3. This action is also against the law as stated in the *Highway Traffic Act* [30].

Technology enablers to enhance crossing safety: *Vehicle detection and tracking via the use of advanced AI technologies based on roadside infrastructure mounted vision sensors makes it possible to detect U-turns on the crossing and notify the approaching vehicles of the potentially hazardous situation via V2X communications.*

2.3.5 Pedestrians trespassing onto railway right-of-way

Use case scenario: This use case is where pedestrians illegally access or cross the railway tracks at a grade crossing when the crossing is configured to allow rail traffic. Trespassing in these circumstances is highly dangerous and can lead to severe injuries or fatalities as mentioned in the TSB investigation report in section 2.1.3, or in the TC Notices and Orders in sections 2.2.1, 2.2.5 and 2.2.6.

Technology enablers to enhance crossing safety: *With the use of secure V2X communications and advanced AI technologies based on multi-modal vision sensors deployed on roadside infrastructure at a crossing, trespassing events can be reliably detected under varying weather and lighting conditions. This detection capability can then be leveraged to –*

- *promptly alert the trespassers through loudspeakers in situations where a potentially dangerous scenario is imminent, such as the approach of a train;*
- *help railway companies and road authorities evaluate if a whistling policy should be re-established based on the frequency and nature of trespassing (e.g., during train approach, or other times), in cases where the crossing has an anti-whistling policy established.*

2.3.6 Trains obstructing crossing for prolonged duration

Use case scenario: According to the Grade Crossing Regulations (GCR), it was prohibited to obstruct a public crossing for more than 5 min by leaving railway equipment standing on the crossing when vehicular or pedestrian traffic is waiting to cross it [31]. Blocking a crossing for a prolonged duration, or even for a short duration if it happens regularly, possesses many safety concerns including obstructing the passage of emergency vehicles as well as prompting pedestrians to take hazardous means to traverse through the crossings as mentioned in the TC Notice and Orders outlined in sections 2.2.1 and 2.2.4.

Technology enablers to enhance crossing safety: *Using advanced AI technologies based on multi-modal vision sensors installed at roadside infrastructure at a crossing, events such as a stopped train blocking a crossing for more than the GCR allowed time limits can be reliably detected. This information can then be broadcast to approaching vehicles via secure V2X communications in an effort to enhance the situational awareness of their drivers, thereby allowing them, and more importantly emergency vehicle drivers, to detour whenever possible.*

2.4 Summary

In section 2 the authors provided a comprehensive study of the grade crossing collisions investigated by TSB since 1994, as well as the notices and orders issued by TC concerning grade crossing safety for the same period. The primary goal of this review was to pinpoint the underlying causes and contributing factors that led to the crossing collisions and the potential safety issues that prevailed at rail crossings, with a view to identifying potential opportunities for utilizing emerging technologies to reduce collisions and improve safety at grade crossings. Additionally, the review was leveraged to guide the formation of rail safety scenarios for at-grade crossings that could potentially benefit from the integration of emerging technologies including secure V2X communications, and sophisticated AI technologies utilizing multi-modal vision sensors. This proactive approach is likely to reduce fatal accidents on at-grade crossings, thereby enhancing crossing safety for motorists, pedestrians and train operators.

3 Collection, processing and analysis of data

To validate the six rail safety use case scenarios developed in sections 2.3.1 to 2.3.6, and better understand the traffic dynamics at or near grade crossings, and more importantly, to facilitate the development of data-driven machine vision algorithms for detection and recognition of vehicles and VRUs, real-world data from a public grade crossing was collected, processed, and analyzed. The following sub-sections provide details about the data collection, data cleaning, and data annotation processes used for this purpose. The authors then present an initial analysis of the dataset produced with a focus on the potentially unsafe events captured in the dataset.

3.1 Data collection

Data was collected at the public grade crossing on March Rd. in Ottawa, located at Mile 2.1 in the Renfrew Subdivision. This crossing, characterized by a single track and six road vehicle lanes running north-south, was equipped with active GCWDs such as gates, flashing lights, and bells. On average, it sees 1 train and 38,419 vehicles passing through daily, resulting in a cross-product of 38,419 [32]. The selection of this crossing was influenced in part by its moderately high ranking on Transport Canada's grade crossing risk assessment tool called GradeX [32], where it ranked 2,916 out of 24,967 crossings (with a lower rank indicating higher risk). Additionally, its location on a high-speed road with an average speed of 80 km/h made it a suitable location for the project.

The crossing's ingress and egress views were both captured using a variety of vision sensors, including RGB and thermal cameras, as illustrated in Figure 2. These cameras were mounted on roadside infrastructure west of the road. The ingress and egress camera views were based on a vehicle traveling north in the east lanes. However, the camera views were quite broad and provided good enough resolution to cover the ingress and egress for south-bound traffic in the west lanes too.

The decision to use multiple sensor types stemmed from their ability to provide distinct and complementary information, thereby aiding in capturing a comprehensive picture of the environment. For instance, RGB cameras excel at capturing detailed objects and scenes, while thermal cameras, unaffected by lighting conditions, provide effective night vision. In the remainder of this report, RGB cameras will be denoted by PTZ due to their pan, tilt and zoom capability, and the two RGB cameras will be referred to as PTZ_ingress and PTZ_egress. The two corresponding thermal cameras will be referred to as Thermal_ingress and Thermal_egress.

Figure 2: Ingress view of the crossing captured by RGB camara (top-left) and thermal camera (top-right); egress view of the crossing captured by RGB camera (bottom-left) and thermal camera (bottom-right). Northbound road traffic is in the far lanes and moves from right to left, and southbound road traffic is in the near lanes and moves from left to right.

To comprehensively capture traffic dynamics at or near the crossing, along with variations in lighting and weather, data was collected at various times of the day in two different months – July and September 2023. Table 3 provides the recording schedules and details about the collected video feeds. Camera frame rates in the table are in frames/s (FPS). In total, 3 batches of data totaling 154 h of video were captured, with each camera contributing 38.5 h of video.

3.2 Dataset building

The collected videos were filtered, processed and annotated to build a representative and multi-modal dataset containing labeled samples of vehicles and VRUs including cars, buses, trucks, pedestrians and cyclists, captured by RGB and thermal cameras. The instances of potentially unsafe events involving these road users were also included in the dataset. Such a dataset was required to develop machine vision algorithms to automatically detect and track vehicles and VRUs with a view to analyzing their behavior to be able to identify potentially unsafe events at or near a grade crossing. Sections 3.2.1 and 3.2.2 delve into some details about how this dataset was constructed.

3.2.1 Data filtering and preprocessing

The collected videos went through several filtering and pre-processing steps to ensure that the data were free of outliers before it could be fed through the rest of the dataset building pipeline. In particular, the following filtering and preprocessing steps were applied:

1. Continuous videos were chunked into 1-min duration files with the frames extracted in JPEG format.
2. To avoid redundancy in data while also retaining relevant information, videos were temporally sub-sampled. To this end, the collected videos were first manually reviewed to include only those segments capturing potentially unsafe events at or near the crossing involving vehicles and VRUs.
3. The above step was followed by selecting every 5th minute for the first batch of videos and every 10th minute for the second batch within each hour. The third batch was sub-sampled to retain only those segments containing VRUs to make the dataset more balanced.
4. To remove irrelevant background information, frame sequences without vehicles and VRUs were dropped.

The above steps resulted in a dataset consisting of 825 min of video data from the four cameras across the three batches as shown in Table 4.

Table 3: Details of the raw video feeds collected.

Batch	Date	Time	Duration (h)	Camera	Resolution	Frame rate (FPS)
1	17-Jul-2023	05:00 – 18:00	14	PTZ_ingress	1920x1080	15
				PTZ_egress		
				Thermal_ingress	640x480	
				Thermal_egress		
2	26-Jul-2023	10:30 – 23:59	14.5	PTZ_ingress	1920x1080	
				PTZ_egress		
				Thermal_ingress	640x480	
				Thermal_egress		

3	20-Sep-2023	16:42 – 20:42	4	PTZ_ingress	1280x720	30
				PTZ_egress	1920x1080	15
				Thermal_ingress	640x480	30
				Thermal_egress		15
	21-Sep-2023	17:37 – 21:37	4	PTZ_ingress	1280x720	30
				PTZ_egress	1920x1080	15
				Thermal_ingress	640x480	30
				Thermal_egress		15
	22-Sep-2023	17:40 – 21:40	4	PTZ_ingress	1280x720	30
				PTZ_egress	1920x1080	15
				Thermal_ingress	640x480	30
				Thermal_egress		15

Table 4: Total video minutes retained per camera per batch after the filtering process.

Camera	**PTZ_ingress**			**PTZ_egress**			**Thermal_ingress**			**Thermal_egress**			**Total duration (min)**
Batch	1	2	3	1	2	3	1	2	3	1	2	3	
Duration (min)	86	33	125	84	28	92	74	28	78	96	21	80	825

3.2.2 Data annotation

Data annotation in this study involved labeling data with ground-truth information, which served as human supervision during the development of machine learning models. To this end, video frames in the filtered and pre-processed dataset were labelled by human annotators using an image annotation tool called DarkLabel [33]. The choice of this annotation tool was motivated by several useful features, particularly its semi-automatic annotation capabilities, support for video input, quick and easy visualizations, and most importantly, the ability to work locally without requiring data to be uploaded to any third-party server. Below is an outline of the annotation process.

1. For each video frame, a tight bounding box was drawn around each vehicle/VRU present in the frame to record its precise location within the image, its class information (i.e., the type of vehicle/VRU), as well as its tracking ID (an identification number for this specific object that is unique throughout the duration the object is visible in any of the cameras). Figure 3 illustrates the process.
2. Fully occluded vehicles/VRUs were not annotated while partially occluded ones were annotated by estimating the extent of their full body.
3. Videos from each camera were annotated by several annotators with each annotator annotating different video min. This was done to minimize the chances of any bias in the annotations.
4. Once the annotation of the whole dataset was completed, a manual review of the annotations was performed on a random selection of 5% of the videos. Had there been an anomaly rate of at

least 20% in the review set, a full manual review of all annotations would have been conducted. However, this was not the case.

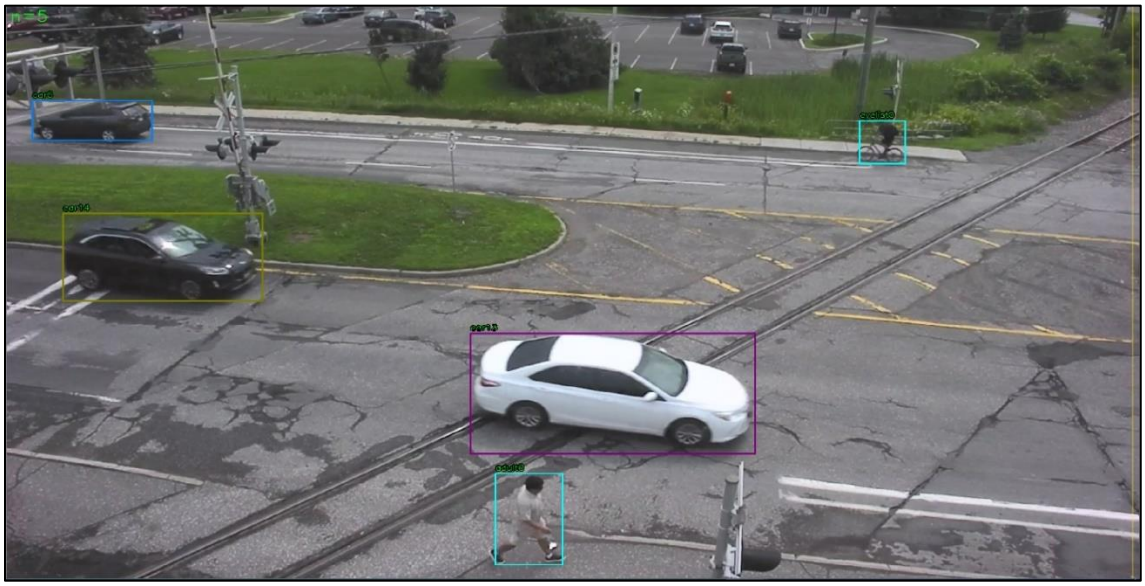

Figure 3: Screenshot of DarkLabel annotation tool showing the results of a completed annotation process. Bounding boxes were drawn around the vehicles and VRUs to record their type and precise location within the image.

3.3 Data analysis

To better comprehend the scale and diversity of the dataset, as well as gain insights into the rail safety use cases based on real-world evidence, a thorough analysis of the collected data was conducted. This analysis is presented in sections 3.3.1 and 3.3.2.

3.3.1 Analysis of vehicles and VRUs

This section presents detailed statistics about the vehicle and VRU categories observed in the collected dataset, including their distribution across different traffic conditions, as well as the distribution of their sizes across different cameras.

Table 5 displays the total count of frames and the various vehicles and VRUs that were annotated from each camera in the collected dataset. Each ‘Count’ column provides a separate count of the respective vehicle/VRU category observed by each respective camera, whereas each ‘Total’ column indicates total instances annotated for that category from the respective camera.

In the table, PTZ_ingress recorded the highest number of vehicles and VRUs, with more annotated frames compared to the other cameras. In total, 233,117 frames were annotated across the four cameras. These annotations included 7,241 separate cars, 75 separate buses, 150 separate trucks, 393 separate persons, and 116 separate cyclists. Consequently, the dataset comprised a total instance of 333,645

cars, 3,718 buses, 7,367 trucks, 63,855 persons, and 11,180 cyclists. For the purpose of the dataset, a vehicle/VRU was considered separate as long as it remained visible within any of the four camera views.

Table 5: Number of frames, vehicles and VRUs annotated from each camera in the collected dataset. The 'Count' and 'Total' columns for each vehicle/VRU category are described when the table is introduced.

Camera	# Frames annotated	# Cars		# Buses		# Trucks		# Persons		# Cyclists	
		Count	Total	Count	Total	Count	Total	Count	Total	Count	Total
PTZ_ingress	107,663	5,420	161,883	49	1,930	77	3,707	224	22,686	77	3,825
PTZ_egress	49,387	4,682	69,197	42	618	76	1,247	254	18,450	73	2,762
Thermal_ingress	28,216	3,502	42,781	26	435	51	822	113	8,586	39	1,116
Thermal_egress	47,907	3,883	59,784	40	735	104	1,591	121	14,133	67	3,477
Total across all cameras	**233,117**	**7,240**	**333,645**	**75**	**3,718**	**150**	**7,367**	**393**	**63,855**	**116**	**11,180**

To understand the traffic patterns at the crossing, Figure 4 illustrates the average number of vehicles and VRUs passing through the crossing per min for 1 h, during the selected hours shown in the figure. The VRUs are separated into persons and cyclists. Rush hour commutes are notably in the 0700 hour (i.e., up to, but excluding 0800), the 0800 hour, 1200 hour, 1500 hour, 1600 hour, and 1700 hour, with the average traffic volume exceeding 30 vehicles/VRUs per min. Among these, the 0800 hour stands out with the highest traffic volume, averaging approximately 48 vehicles/VRUs per min, followed by the 1600 hour registering approximately 44 vehicles/VRUs per min on average.

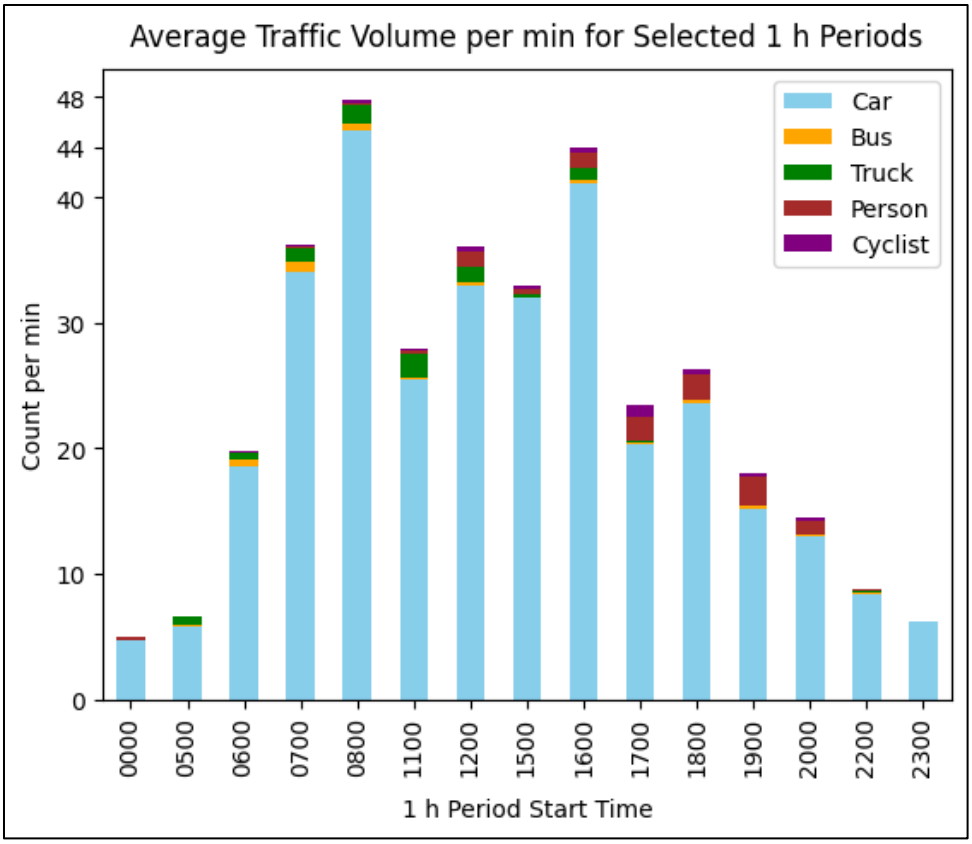

Figure 4: Average number of vehicles and VRUs passing through the crossing per min for selected 1 h periods of interest.

The left side of Figure 5 illustrates the length distributions (in pixels) of various vehicles and the height distributions (again in pixels) of different VRUs (again divided into persons and cyclists) observed by the PTZ cameras. On the right, the same distributions are depicted for the thermal cameras. VRUs, being shorter, dominate the lower tail of the distributions ranging from approximately 30 to 250 pixels for the PTZ cameras, and 18 to 120 pixels for the thermal cameras. On the other hand, buses and trucks, due to their larger sizes, occupy the higher tail of the distributions having a range of 120 to 1,890 pixels for the PTZ and 80 to 640 pixels for the thermal cameras. Cars exhibit wide variability in their lengths with the peak of the distribution between 150 to 250 pixels for the PTZ cameras and 100 to 150 pixels for the thermal cameras. Comparatively, in the thermal camera distributions, vehicles and VRUs appear smaller, which can be attributed to the lower resolution of thermal cameras compared to PTZ cameras as mentioned in Table 3.

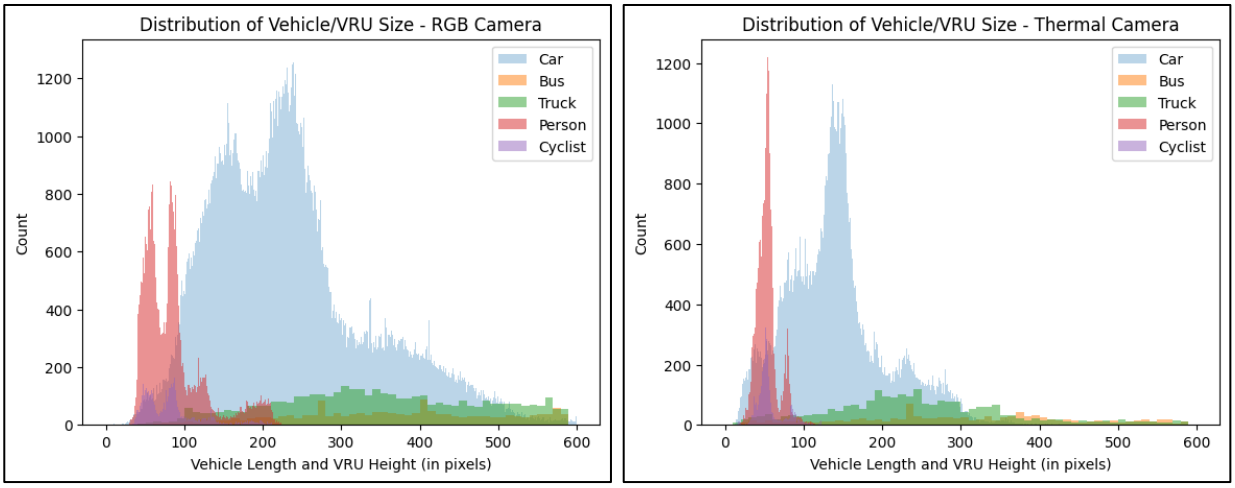

Figure 5: Distribution of vehicle and VRU sizes for PTZ cameras (left) and thermal cameras (right).

3.3.2 Analysis of the rail safety use cases

The collected dataset allowed for the validation of the rail safety use cases developed in Sec. 2.3 based on real-world evidence. The different potentially unsafe events that were captured in the collected 825 min dataset included U-turn on the crossing, queue buildup on the crossing, vehicles stopped on the crossing, vehicles stopped near the crossing (referred to as stopped elsewhere), as well as instances of jaywalking and pedestrian intrusion into the railway right-of-way. Table 6 shows the numbers of the six different potentially unsafe events captured by each of the four cameras over the entire data collection period.

Table 6: Counts of potentially unsafe events observed in the different camera views. Caution should be applied when interpreting these numbers as the same event was likely captured in multiple camera views.

Camera	U-turn	Queuing	Stopped on crossing	Stopped elsewhere	Jaywalking	Intrusion
PTZ_ingress	5	28	1	19	83	16
PTZ_egress	5	20	1	18	72	12
Thermal_ingress	4	18	0	12	54	12
Thermal_egress	4	17	1	15	78	16
Total Unique Events	**5**	**29**	**1**	**19**	**100**	**35**

In total, 5 instances of U-turn on the crossing, 29 instances of queuing, 1 instance of vehicle stopping on the crossing surface, 19 instances of vehicle stopping near the crossing, as well as 100 jaywalking and 35 intrusion events were captured in the dataset. Figure 6 shows two examples (top – queuing on the crossing; bottom – pedestrian intrusion into the railway right-of-way) of the six event types.

Figure 6: Examples of some of the potentially unsafe events captured by both PTZ and thermal cameras. The top row depicts a scenario where cars queue up on the egress of the crossing, extending past the crossing into the ingress area. The bottom row shows a pedestrian (encircled in red) intruding into the railway right of way.

Among the different events involving vehicles, queue buildup on either ends of the crossing was observed the most in the collected dataset. This was particularly concerning, especially queuing on the egress as mentioned in section 2.3.3, due to the potentially dangerous situation this can lead to. For example, queuing on the egress can potentially trap vehicles between the gate arms for a prolonged time – an example is shown in the top row of Figure 6 – thereby increasing the chances of collision with a train.

To further investigate into the queuing events, Figure 7 illustrates the queuing durations for selected 1 h periods on the four days the events took place – one in the morning rush hour and five in the afternoon rush hour. As was shown previously in Figure 4, the majority of the queuing events occurred during rush hour commutes, particularly in the afternoon, and so those are the hours shown in Figure 7. For the specific crossing studied in this project, the orange line indicates the time duration from the activation of the crossing arms to the arrival of a train on the crossing surface. The red line indicates the time duration from the activation of the crossing lights/bells to the arrival of a train on the crossing surface.

The 1600 hour (for 1600 up to but excluding 1700) registered the highest number of events, totaling 7. While the average queuing duration stood at 9 s, a significant portion of the events extended beyond 15 s, with the longest lasting up to 26 s.

Given that the time from the activation of the crossing arms until the arrival of a train at the crossing surface was approximately 16 s, and the time from the activation of lights/bells until the arrival of a train at the crossing surface was approximately 34 s for this specific crossing, and considering that the duration of such queuing events was entirely dictated by instantaneous traffic dynamics, it is evident that these events increasingly jeopardize public safety.

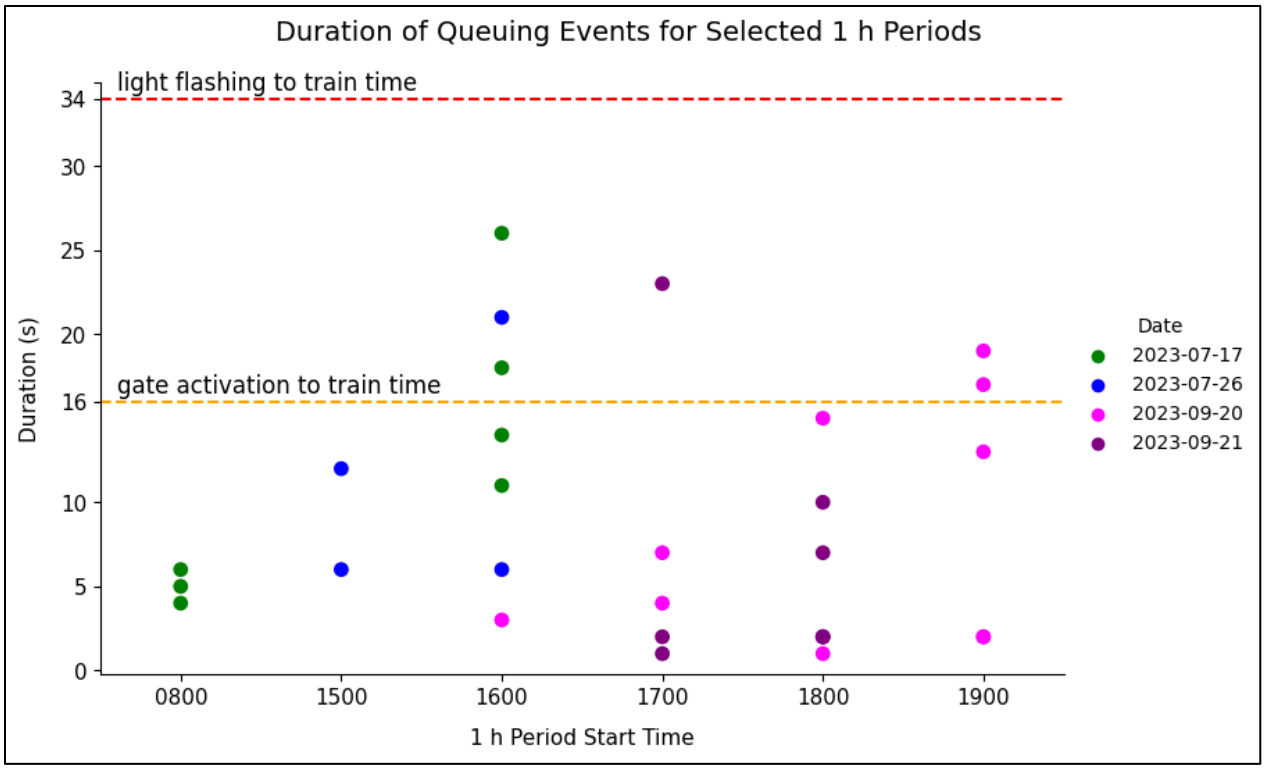

Figure 7: Duration of the queuing events for selected 1 h periods in the dataset. More details about the orange line and red line are provided in the text where this figure is described.

3.4 Summary

Section 3 provides details about the collected data and the process for building a multi-modal dataset containing labeled samples of vehicles and VRUs for machine learning model development. It also presents key statistics from the collected dataset, including the distribution of the vehicles/VRUs across different 1 h periods of the day, as well as the distribution of their sizes across the different cameras. Additionally, it includes an analysis of the rail safety use cases based on real world evidence from the collected data.

4 Methodology

Inspired by recent research in human activity detection from videos [34] [35], the overall pipeline for detection of vehicles and VRUs (divided into persons and cyclists in this study), along with identifying potentially unsafe events is illustrated in Figure 8. The pipeline comprises three main components:

i. Detection and recognition of vehicles and VRUs at, through, and around the crossing.
ii. Tracking of detected vehicles and VRUs as they traverse through the crossing.
iii. Identifying potentially unsafe events based on the tracked vehicles and VRUs.

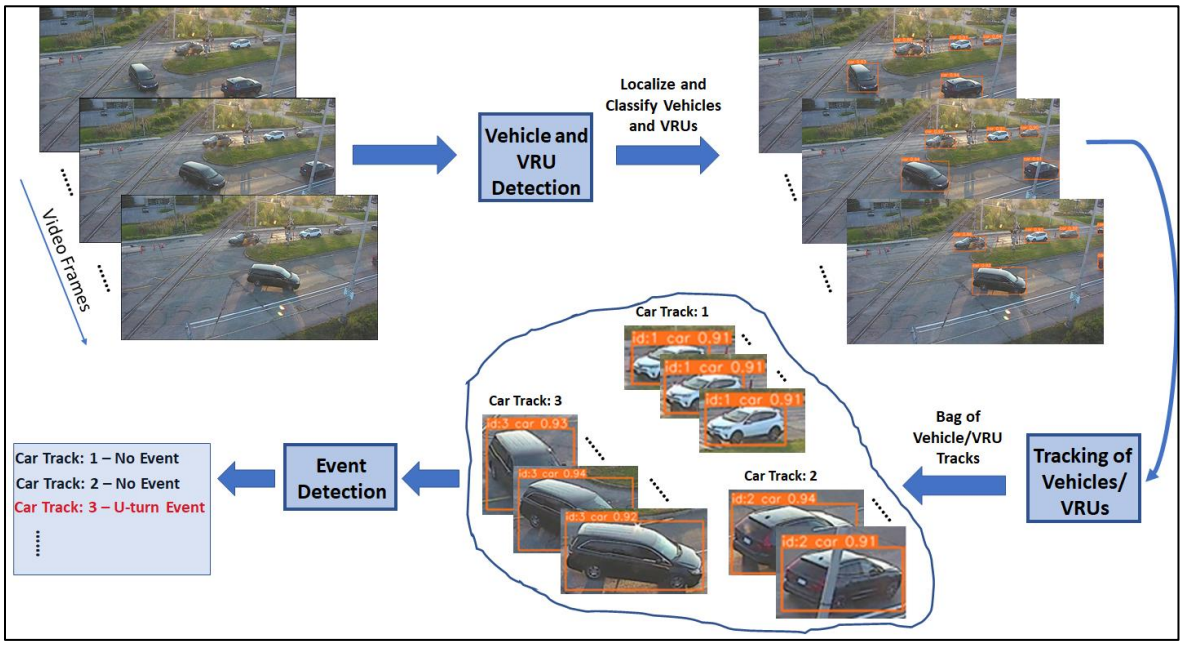

Figure 8: Overall pipeline for detection of vehicles/VRUs and the identification of potentially unsafe events. Best viewed when zoomed in.

As illustrated in the figure, the input to the pipeline is raw video frames which are processed by a state-of-the-art object detection algorithm to classify vehicles/VRUs, as well as localize them within the individual video frames. Each detected vehicle/VRU is then tracked over time using a multi-object tracker. The objective is to link the detected vehicles/VRUs across the sequence of frames in the input video, resulting in a collection of their trajectories. Finally, after some pre-processing steps, the discovered trajectories are fed as input to a set of event detection algorithms, each responsible for identifying a specific unsafe event (e.g., U-turn, queuing, etc.).

The different components of this pipeline are elaborated on in sections 4.1 to 4.4.

4.1 Vehicle and VRU detection

The state-of-the-art object detection algorithm called YOLOv8 [36] was trained leveraging the collected dataset to detect and classify the different vehicles and VRUs from both the PTZ (RGB) and thermal camera images. Details of the model development, model training, and model testing are outlined below.

4.1.1 Model development

Two different object detection models were trained based on YOLOv8 – one using the PTZ camera images, and the other using the thermal camera images. For each camera type, the frames in the collected dataset were split into three distinct sets – training, validation, and test. The training set was used during model development while the validation set was leveraged for best model selection and hyperparameter tuning. The test set was not used during training or validation, as it was reserved for final performance evaluation of the developed models. Additionally, the test set video minutes, wherever possible, were selected to be consistent across the different cameras in order to allow for evaluation across the different sensor modalities.

To get a more uniform distribution across traffic dynamics, scene variations, lighting conditions, and more importantly, different vehicle and VRU classes, video minutes were sampled following the steps outlined below to obtain the training, validation, and test splits.

1. Each set included data from rush hour and non-rush hour traffic, as well as day and night time traffic. To address some of the class imbalance issues in the collected dataset as revealed in Table 5 (e.g., too many car samples than other vehicle types and VRUs), data for the training set were mainly sampled from those video minutes that included the under represented classes.
2. To prevent leakage of information from the training set to the validation/test set, consecutive minutes in an hour were designated for each set while making sure that there was a discontinuity in time between any two sets.
3. Following standard practice, approximately 70%, 10%, and 20% of the sampled video minutes were selected for training, validation, and test sets, respectively.

4.1.2 Training and inference details

For both the PTZ and thermal camera models, the x version of YOLOv8 (aka YOLOv8x) pre-trained on the COCO dataset was fine-tuned. To this end, the last classification layer of the pre-trained model was changed from 80 (COCO dataset has 80 classes) to match the number of vehicle and VRU classes in the collected dataset, to be specific 5 (car, bus, truck, person and cyclist). Moreover, the backbone layers of YOLOv8x were kept fixed during training, since these layers mainly extract generic and meaningful representations of an image. To generate diverse variations in input, several data augmentation techniques were applied on the input images during training including random scaling, rotation, and translation. However, inference was performed on the actual images without any data augmentation. Moreover, evaluation was performed based only on the region of interest (ROI) that covered the road and sidewalk areas in the images.

During training, the batch size was set to 8 for the PTZ camera images and 16 for thermal camera images, while during testing, it was set to 1 for both cameras. An initial learning rate of 0.001 was used to warm start the training process which was reduced to 0.0001 after 5 epochs. Both models were trained until full convergence while model selection was performed based on model performance on the

validation set. This meant that the best performing model on the validation set was chosen for final performance evaluation.

4.2 Tracking of detected vehicles and VRUs

The detection results from the YOLOv8 algorithm were fed as input into a multi-object tracker called ByteTrack [37]. The main goal of tracking the detected vehicles/VRUs was to obtain their individual trajectories over time so that those could be analyzed by the event detection algorithms. The novelty of ByteTrack stems from the fact that it processes both high and low confidence detections, effectively linking objects across frames even when detection confidence varies widely, thus recovering more trajectories for vehicles and VRUs. This resulted in more robust and accurate tracking, especially in challenging scenarios with occlusions and varying object appearances, phenomena that are typically present in vehicle/VRU detection in the wild.

The low and high confidence thresholds for ByteTrack were set to 0.1 and 0.3, respectively. The patience value for keeping the lost trajectories alive was set to 30, meaning that the trajectory of any object that was not observed for more than 30 consecutive frames was discarded.

4.3 Detection of potentially unsafe events

This section describes the detection and recognition of potentially unsafe events involving vehicles and VRUs at a grade crossing. Referring back to Figure 8, the event detection algorithms were based on the analysis of vehicle/VRU trajectories available from the tracking of the detected vehicles/VRUs, while taking into account the configuration of intersections (danger zone, jaywalking/intrusion areas, etc.).

Two sets of event detection algorithms were developed using algorithmics and computer vision techniques. The first set was related to the vehicle events (i.e., U-turn, queuing, stopped on crossing, and stopped elsewhere), while the second set encapsulated the VRU events (i.e., jaywalking and intrusion). Each algorithm supported both cameras (i.e., PTZ and thermal) and both views (i.e., ingress and egress). These algorithms are briefly summarized in sections 4.3.1 to 4.3.4 .

4.3.1 U-turn event

The U-turn event was detected by analyzing the vehicle trajectories. Figure 9 shows that, depending on the viewpoint of the camera, the U-turn event could have a partial (yellow) or complete (blue) trajectory.

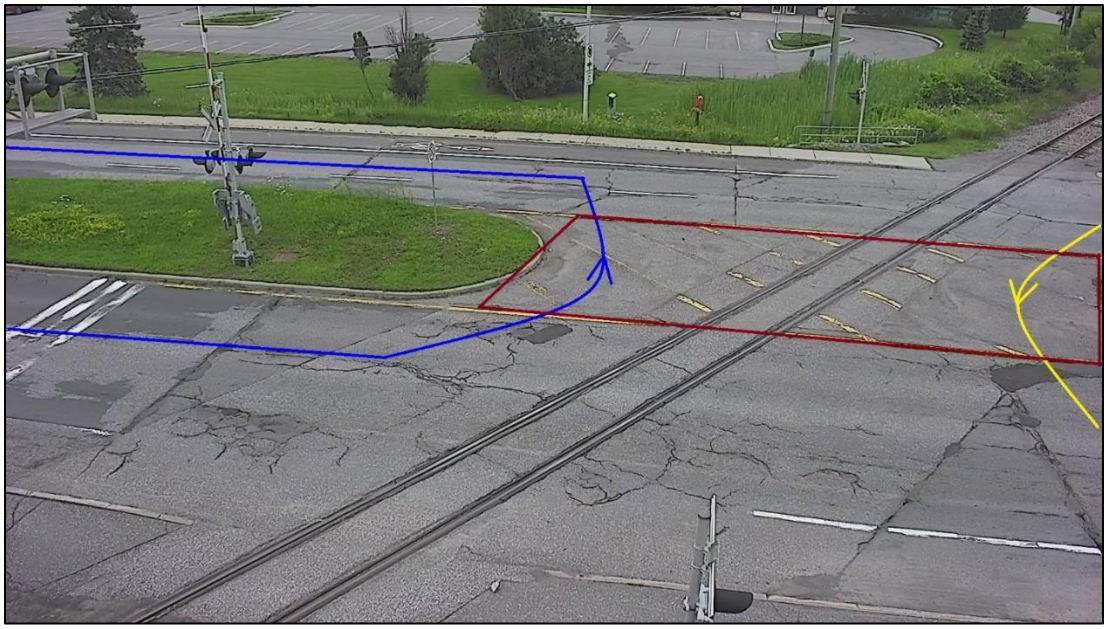

Figure 9: PTZ_egress: partial U-turn trajectory (yellow); full U-turn trajectory (blue); and U-turn area (dark red border).

The algorithm was developed to detect and recognize both types of U-turns. The algorithm analyzed vehicle trajectories to identify those making a full U-turn as well as vehicles passing through the U-turn area (see Figure 9, ROI highlighted in dark red) and those making a partial U-turn. These vehicle trajectories were obtained from deep-learning based vehicle/VRU detection and tracking models as described in sections 4.1 and 4.2.

4.3.2 Stopped on the crossing and stopped elsewhere events

Stopped on the crossing and stopped elsewhere events share the core idea of detecting and recognizing a stopped or slowed vehicle on the road. This algorithm used two important parameters to quantify the vehicle's motion between two successive camera frames and to determine the period of time needed to recognize that a vehicle was stopped or slowing down. For example, to detect a completely stationary vehicle, the motion of the vehicle between two successive frames should be equal to zero pixels. If the camera frame rate is 15 FPS and a vehicle needs to be stopped for at least 2 s before the algorithm reports a stopped vehicle event, the second parameter of the algorithm will be set to 30 frames. The algorithm should have detected a stopped vehicle over 30 consecutive frames, which corresponded to a time period of 2 s. These two parameters are very important especially in a V2X communication system. The stopped elsewhere and stopped on the crossing algorithms are differentiated based on their corresponding ROIs. In Figure 10, a vehicle stopped inside the red ROI was reported as stopped on the crossing event while a vehicle stopped inside one of the blue ROIs was recognized as stopped elsewhere.

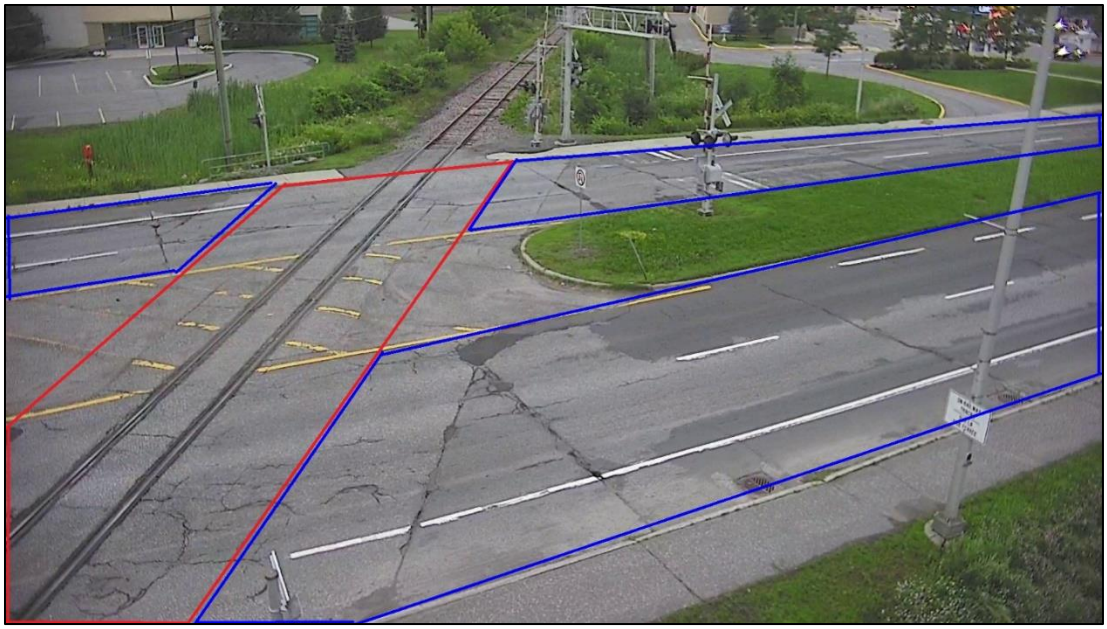

Figure 10: PTZ_ingress: stopped on the crossing ROI (red) and stopped elsewhere ROIs (blue).

4.3.3 Queuing event

The queuing event occurred in one or more traffic lanes depending on the configuration of the grade crossing. Figure 11 shows the crossing has four traffic lanes, with lanes 1 and 2 being the two northbound lanes, and lanes 3 and 4 the two southbound lanes. The detection of queuing events was performed in two parallel steps. The algorithm identified all stopped or slowed vehicles in the scene using the same algorithm core described in section 4.3.2 for the stopped on the crossing and stopped elsewhere events. At the same time, the algorithm analyzed each traffic lane and reported a queuing event if at least two stopped or slowed vehicles were detected in the same lane. Note that multiple events were able to occur at the same time. For example, a vehicle could have reported as queuing and stopped on the crossing. The queuing algorithm had the potential to share the number and position of vehicles involved in the queue through a V2X communication system. The vehicle positions needed to be converted from camera coordinates to real world coordinates. Cameras should be calibrated to be able to perform this conversion effectively in any future projects.

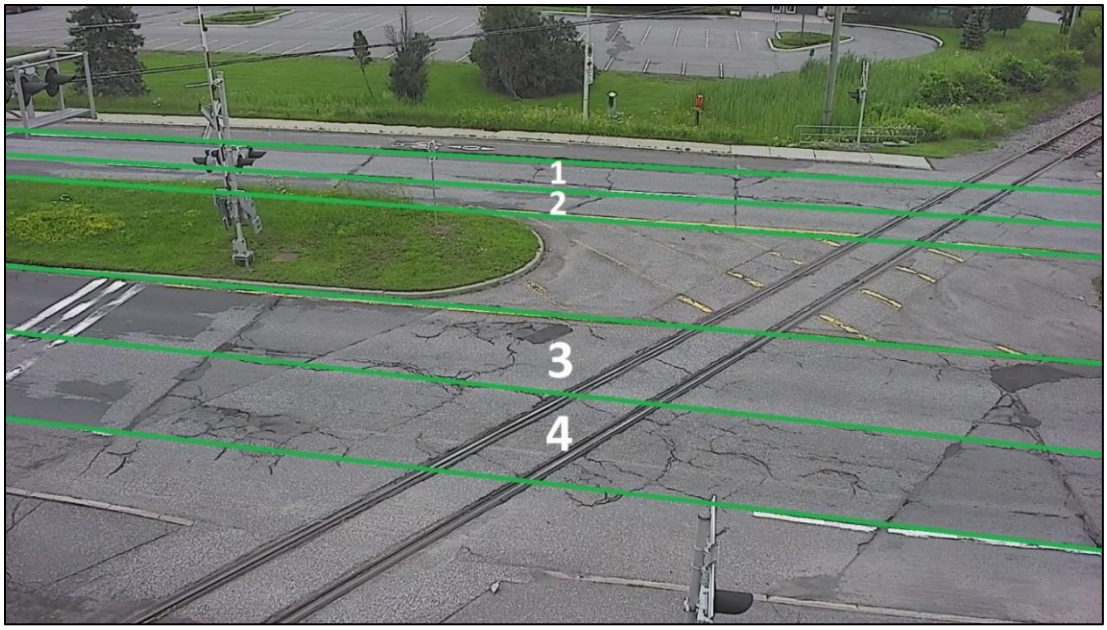

Figure 11: PTZ_egress: grade crossing intersection with four traffic lanes.

4.3.4 Jaywalking and intrusion events

Jaywalking and intrusion events were related to VRUs (i.e., pedestrians and cyclists). The algorithm analyzed the VRU trajectories using the jaywalking (green) and intrusion (red) ROIs shown in Figure 12. The algorithm reported a jaywalking or intrusion event when a VRU crossed a jaywalking or intrusion ROI, respectively, through a specific number of frames. If the jaywalking or intrusion event should have been reported as soon as the VRU crossed the ROI, the number of frames was set to one.

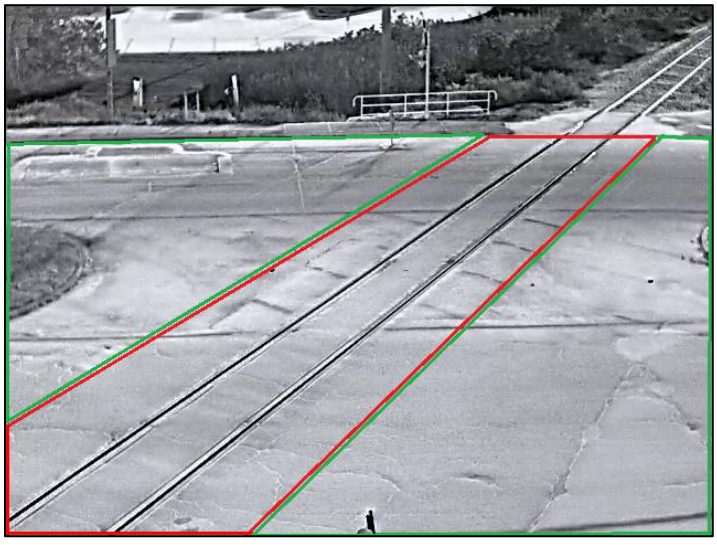

Figure 12: Thermal_egress: jaywalking (green) and intrusion (red) ROIs.

4.4 Performance metrics

The performance of the vehicle/VRU detection models and the algorithms for identifying unsafe events was evaluated based on several standard metrics used for general object detection and classification problems. Each performance metric is described below.

4.4.1 Precision and recall

Precision is defined as the fraction of instances predicted as positive that are in fact positive. It is measured using a scale of 0 (i.e., all predictions are incorrect) to 1 (i.e., all predictions are correct). In the context of vehicle/VRU detection, precision indicates how accurate the models' predictions are in terms of the detected vehicle/VRU instances, while for identifying unsafe events, precision refers to the accuracy of the detected events.

Recall measures what fraction of the positive instances is correctly predicted as positive and is also measured on the scale of 0 (i.e., none of the positive instances are correctly predicted) to 1 (i.e., all positive instances are correctly predicted). For vehicle/VRU detection, recall indicates how complete the developed models are in terms of detecting all instances of vehicles/VRUs in the dataset. On the other hand, for unsafe event identification, recall indicates the fraction of the total events in the dataset that the algorithms are able to detect correctly.

For predictive algorithms such as the ones used to train vehicle/VRU detection models, precision and recall are computed at a specific confidence threshold *k* such that predictions with confidence score above *k* are considered positive. Mathematical definitions of precision and recall are as follows:

$$Precision@k = \frac{TP@k}{TP@k\ + FP@k}; \qquad Recall@k = \frac{TP@k}{TP@k + FN@k};$$

Here, $TP@k$, $FP@k$, and $FN@k$ refer to true positive counts, false positive counts, and false negative counts, respectively, at the given confidence threshold *k*.

Since vehicle/VRU detection required both classifying and localizing vehicles/VRUs, a prediction was considered true positive (TP) only if the predicted vehicle/VRU type was correct as well as the intersection-over-union (IoU) score (the ratio of intersection to union between the ground-truth and the predicted object regions in an image) was above a predefined threshold. The IoU threshold is commonly set to 0.5 for most applications involving vehicle/VRU detections [38], [39], and that was used for the study presented here.

4.4.2 F1-score

The F1-score combines precision and recall into a single metric to better measure the algorithm performance. It is defined as below.

$$\text{F1-score} = \frac{2*precision*recall}{precision+recall}$$

Similar to precision and recall, the F1-score ranges between 0 and 1 with higher metric values referring to better performance.

4.4.3 Confusion matrix

A confusion matrix is used to evaluate the performance of classification and detection algorithms. It is essentially an $N \times N$ matrix, where N is the total number of classes in the dataset. The confusion matrix for each class displays the fraction of observations correctly classified as belonging to that class, as well as the fractions of observations incorrectly classified as belonging to other class(es). Table 7 shows the general structure of the confusion matrix for a two-class classification problem.

Table 7: Confusion matrix for a two-class classification problem.

Ground truth (Actual observations)	**Total number of observations**	**Detected instances**	
		Positive class	**Negative class**
	Positive class	True positive (TP)	False Negative (FN)
	Negative class	False Positive (FP)	True Negative (TN)

4.4.4 Mean average precision (mAP)

Mean average precision (mAP) is a commonly used performance metric for the object detection task. It provides the mean of the average precision (AP) values of the different classes present in the dataset. AP, on the other hand, is defined as follows. mAP is used as the single most significant performance indicator over the whole dataset for general object detection task.

$$AP = \sum_{k=1}^{k=N-1} [(r@k - r@(k-1)) \times \max(p@k, p@(k-1))]$$

$$p@k = Precision\ at\ confidence\ threshold\ k$$
$$r@k = Recall\ at\ confidence\ threshold\ k$$
$$N = Number\ of\ possible\ confidence\ thresholds\ varied\ from\ 1\ to\ 0$$

4.4.5 Precision-recall curve

A precision-recall curve plots precision against recall for all possible confidence thresholds varied from 1 to 0. The higher the area under the precision-recall (PR) curve, the higher the precision and recall. Since precision and recall vary with the confidence threshold, the PR-curve provides an intuitive measure to select a suitable operating threshold that maximizes both precision and recall.

4.5 Summary

Section 4 first provides an outline of the entire solution methodology used to address the problem at hand. It then delves into the necessary details for training deep learning models to detect vehicles/VRUs, followed by an overview of the tracking algorithm, as well as the algorithms developed to address the potentially unsafe events. Finally, it describes the different metrics used to evaluate the performance of the developed solution.

5 Results

In the following sub-sections, results for an automated detection and recognition of vehicles and VRUs at, through, and around one at-grade crossing are presented. Additionally, the performance of the algorithms developed to identify potentially unsafe events involving vehicles and VRUs as they traverse through a grade-crossing are presented.

5.1 Detection and recognition of vehicles and VRUs

This section presents the results of vehicle and VRU detection based on the collected dataset. Section 5.1.1 first presents results for vehicle detection, while Section 5.1.2 through 5.1.5 presents analysis of model performance for vehicle and VRU detection and tracking under various operational conditions including variation in lighting conditions, road traffic, and video frame rate while also discussing the impacts of different image artifacts on model performance.

5.1.1 Vehicle detection performance

The performance of the different models to detect and recognize different types of vehicles (i.e., car, bus, and truck) was evaluated based on a subset of the collected dataset. Table 8 provides a summary of the data used for this purpose

Table 8: Details of the test set used to evaluate model performance for vehicle detection.

Date	Time	Camera	# Cars (Total)	# Buses (Total)	# Trucks (Total)
17-Jul-2023	0550	All four cameras	625	154	0
	0715	All four cameras	526	0	97
	1610	All four cameras	3526	0	177
26-Jul-2023	2220	All four cameras	1821	59	0
20-Sep-2023	1744	All but Thermal_ingress	492	41	0

Table 9 presents the performance of the different models trained on images from the PTZ and thermal cameras. The PTZ camera model performed excellently in detecting and recognizing various types of vehicles, especially cars, achieving an AP value of 0.96 and an F1-score of 0.92. It also showed comparable performance for buses and trucks.

In contrast, the model trained on thermal camera images showed inferior performance. Although it performed satisfactorily for cars, with an AP of 0.86 and an F1-score of 0.82, its performance for buses and trucks was poor.

Table 9: Results of vehicle detection from PTZ and thermal camera images. The higher the AP and F1-score values, the better the performance.

Model	Car		Bus		Truck	
	AP	F1-score	AP	F1-score	AP	F1-score
PTZ	0.96	0.92	0.90	0.84	0.91	0.85
Thermal	0.86	0.82	0.60	0.63	0.70	0.58

A possible reason for the thermal camera model's poor performance could be the limited horizontal field of view compared to the PTZ cameras. Large vehicles like buses, especially bendy buses, and trucks traveling in the near lane would occupy the entire horizontal field of view for the thermal camera, making it difficult for the model to recognize and localize these vehicles. Figure 13 illustrates two example frames from the Thermal_egress camera with detection results. The vehicle traveling in the far lane in the left image was detected, while the one in the near lane was missed. However, further research is warranted to determine the exact cause of the thermal camera model's poor performance on large vehicles.

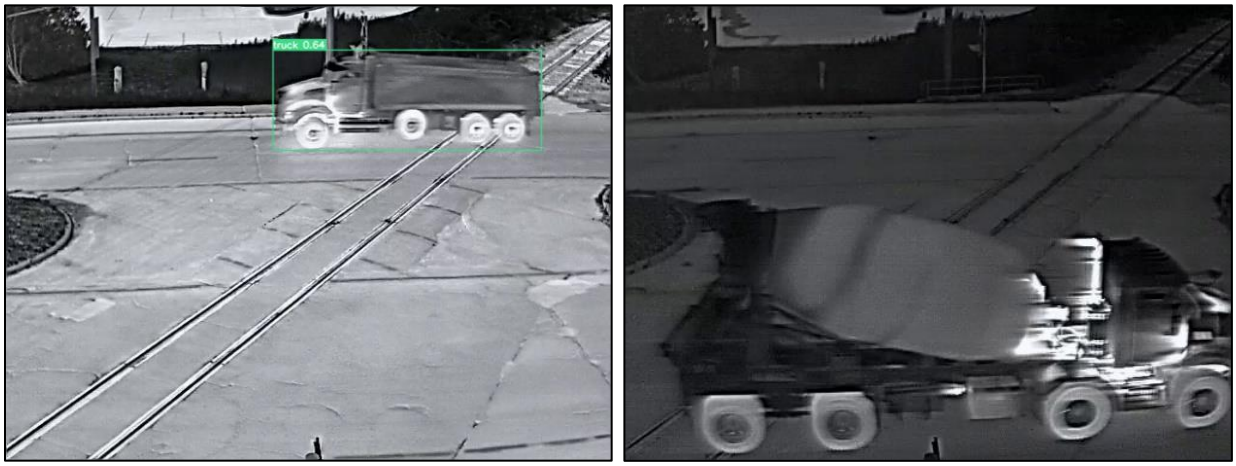

Figure 13: Two example images captured by the Thermal_egress camera. The model was able to detect the truck travelling along the far lane (left image), while it missed the one travelling along the near lane (right image).

5.1.2 Performance under varying lighting conditions

To evaluate how the vehicle/VRU detection performance varied depending on lighting conditions across the different sensors, two separate splits of the test data were created – one that included only day-time traffic scenes, while the other included only night-time traffic scenes. Table 10 shows the details of the test splits.

Table 10: Details of the test splits showing day vs. night time traffic.

Split	Date	Time	Camera	# Cars (Total)	# Persons (Total)	# Cyclists (Total)
Day	17-Jul-2023	1100	PTZ_egress	115	0	0
			Thermal_egress	95	0	0
		1240	Both PTZ cameras	525	185	42
			Both thermal cameras	415	155	32
	20-Sep-2023	1727	PTZ_egress	120	10	60
			Thermal_egress	45		25
Night	26-Jul-2023	2250	Both PTZ cameras	155	35	0
			Both thermal cameras	98	7	0
	20-Sep-2023	1953	Both PTZ cameras	253	90	95
			Both thermal cameras	220	17	53
		1957	Both PTZ cameras	264	52	0
			Both thermal cameras	188	100	0

Since the dataset was collected from a live grade crossing and not from any controlled testbed, the scene dynamics (e.g., occlusion level, traffic volume per min, vehicle/VRU proximity to the crossing) varied widely during different times of the day. Therefore, to allow for a fair comparison, care was taken to select only those video segments that exhibited similar scene dynamics across the two splits. Moreover, from each video min, only those frames were selected that depicted the vehicles/VRUs in close proximity to the crossing.

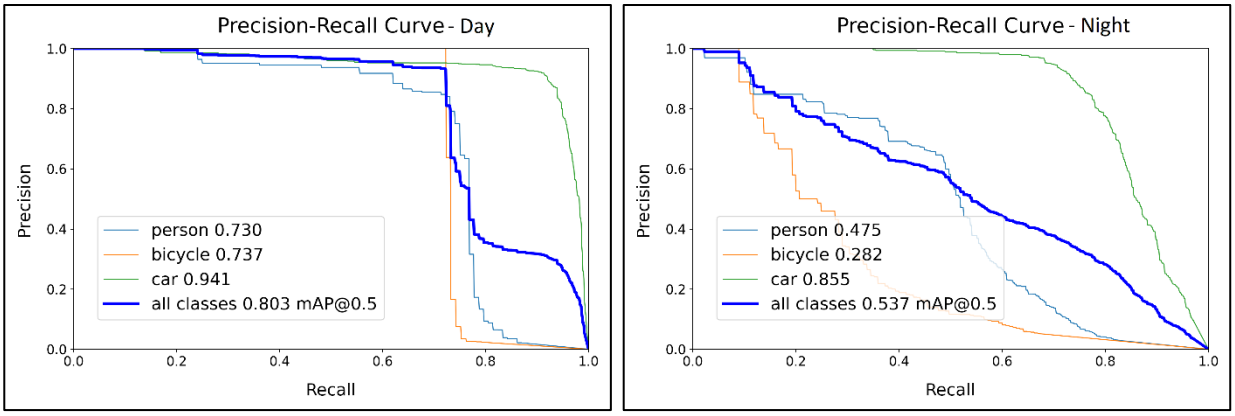

Figure 14: Precision-Recall curve for day time (left image) and night time traffic (right image) based on detection from the PTZ cameras. The higher the AP value, the better the performance.

Figure 14 shows the results of a precision-recall analysis based on PTZ camera images. While the detection of cars was observed to be satisfactory both during day and night time, the performance during night time saw a decline of approximately 9% in the AP value. However, VRU detection was severely affected during night-time for PTZ camera images, with detection of persons and cyclists seeing an AP value drop of approximately 26% and 46%, respectively.

The reduced performance during night time could be mainly attributed to the lack of visual features during night time, especially for VRUs which are generally much smaller in size than any vehicles. Another possible reason for the significantly reduced performance during night time could be the poor quality of the night time images from the PTZ cameras. Figure 15 shows two example images captured by the PTZ_ingress camera. As can be seen, the night time image (right) was blurry and pixelated, essentially diminishing visual features for the cyclist, thus causing a missed detection.

Figure 15: Example detection on day time image (left) and night-time image (right) captured by the PTZ_ingress camera. The cyclist on the left image captured during the day was accurately detected by the model, while another cyclist in the same position (denoted by the white arrow) captured by the same camera during the night was missed.

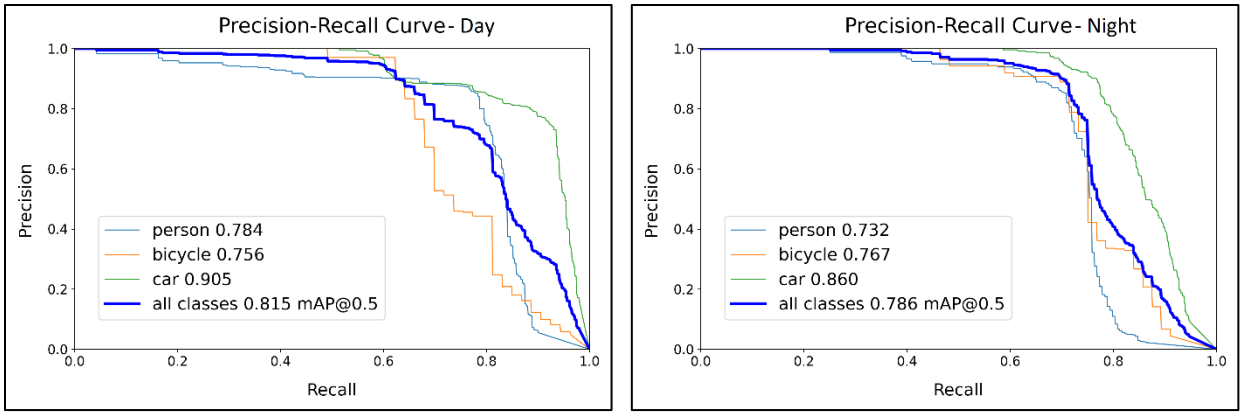

Figure 16: Precision-Recall curve for day time (left) and night time traffic (right) based on detection from the thermal camera images. The higher the AP value, the better the performance.

Figure 16 shows the results for the same exercise based on thermal camera images. As the precision-recall plots reveal, vehicle and VRU detection performance was comparable across day and night time. This is inline with the expectation since images captured by thermal camera sensors were not affected by lighting variation. Figure 17 shows detection performance on two images captured by the PTZ_ingress and Thermal_ingress cameras, respectively. While the VRU blended into the background in the dark in the captured PTZ camera image, the clear thermal signature of the VRU helped it appear bright in the thermal camera image, thereby allowing the model to accurately detect the VRU.

Figure 17: VRU detection during night time on PTZ (left) and thermal (right) camera images. While the model trained on PTZ camera images failed to detect the VRU (denoted by the white arrow on the right image), the model trained on thermal camera images was able to detect it in the corresponding thermal camera image.

5.1.3 Performance under varying traffic conditions

In order to evaluate the performance of vehicle detection under different traffic conditions, two different splits of the test data were created – one depicting a high volume of traffic and the other depicting a low traffic condition. These splits were motivated by the analysis of traffic patterns as illustrated in Figure 4. To be specific, a 1 min length of video data from 0830 in the first batch was selected for the high-traffic condition, as this minute saw 78 separate vehicles passing through the crossing. On the other hand, for

the low-traffic split, a 1 min video at 1100 was selected, during which 18 separate vehicles passed through the crossing. Table 11 provides details of the test splits.

Table 11: Details of the test splits showing low traffic and high traffic conditions.

Camera	# Cars (Low traffic)		# Cars (High Traffic)	
	Count	Total	Count	Total
PTZ_ingress	18	201	78	929
PTZ_egress		171		801
Thermal_ingress		153		405
Thermal_egress		85		376

Figure 18 shows the results of the analysis. The detection of cars was satisfactory in both low and high traffic conditions, for both the PTZ and thermal camera images. However, the models showed slightly reduced performance in high traffic conditions on both cameras. This may be due to factors like the increased likelihood of occluded scenes during heavy traffic compared to low traffic situations. However, the performance drop was more significantly pronounced for the thermal camera images (~8% drop in the AP value) than those from the PTZ cameras (~2% drop in the AP value). This could be explained by the superior ability of the model trained on PTZ camera images in distinguishing cars in occluded scenarios than that trained on thermal camera images, thanks to the rich and fine details captured in PTZ camera images.

An example situation based on two images from the high traffic split is illustrated in Figure 19.

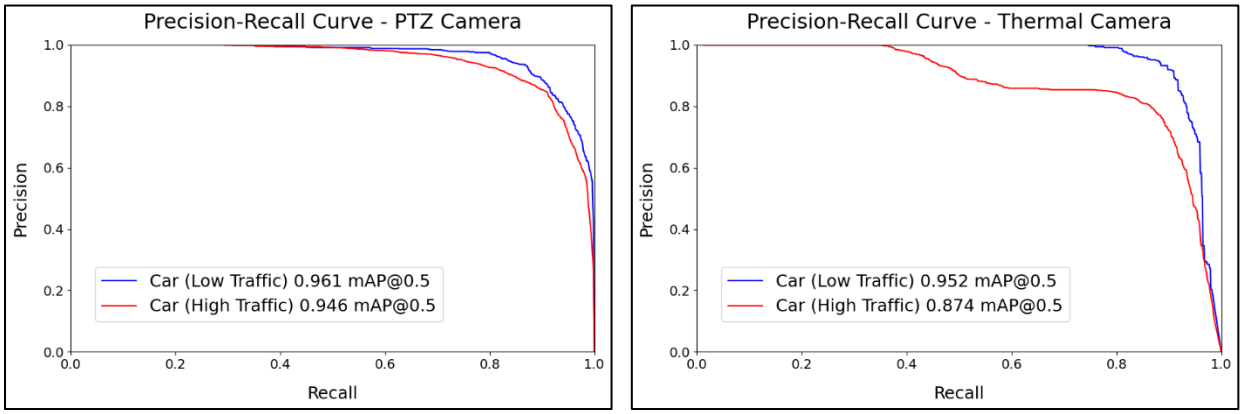

Figure 18: Precision-Recall curve for the low-traffic and high-traffic splits of the test data. (left) PTZ camera, (right) thermal camera. The higher the AP value, the better the performance.

Figure 19: Two occluded cars captured by the PTZ_egress (left image) and Thermal_egress (right image) cameras. The model trained on the PTZ camera images was successful in detecting both cars (one in the foreground and a second in the background) as shown in the left image, while the model trained on the thermal camera images failed to detect the occluded car as revealed in the right image.

5.1.4 Effect of video frame rate

Video frame rate is an important factor that dictates the logistics required to collect, process and analyze data from multi-modal vision sensors in a real-time setting. While higher frame rate offers fine-grained temporal dynamics of the scene captured in the camera, it incurs increased computational and storage demands on the processing and analysis fronts due to the sheer volume of data generated per unit of time, especially in a multi-modal vision sensor environment. Since increased computational demand would typically incur increased power consumption, capturing videos at a higher frame rate would be very challenging on a road-side infrastructure deployment scenario due to the limitations of logistics and resources such deployment environments are typically subjected to. Therefore, it is crucial to find the optimal video frame rate required for the application at hand. To this end, this section investigates if an optimal frame rate could be established to detect vehicle and VRUs that would ultimately allow for the detection of unsafe behaviours involving them at or near a grade crossing.

As illustrated in Figure 8, the detected vehicles/VRUs needed to be tracked over time so that their trajectories could be studied to recognize their behaviour as belonging to one of the rail unsafe scenarios mentioned in Table 6. While detection works on individual frames and is thus irrelevant to video frame rate, the tracking of the detected vehicles/VRUs is influenced by video frame rate. However, among the different potentially unsafe events, only U-turn and jaywalking/intrusion events involved vehicles and/or VRUs in motion, while the other events (i.e., queuing, stopped on the crossing surface, and stopped elsewhere) involved stopped vehicles, thus making them much less relevant to video frame rate. Therefore, it is sufficient to study the impact of video frame rate with relation to U-turn and jaywalking/intrusion events.

Table 12 shows the effect of varying video frame rate on the tracking performance for the U-turn event. There was a total of 36 separate vehicles (thus 36 separate trajectories) observed during the video duration selected for the analysis, with one vehicle involving in the U-turn event. The original frame rate of the video was 15 FPS. When the video was sampled at 10 FPS, the trajectories of all 36 vehicles were

successfully identified by the tracking algorithm, as opposed to only 11 when the video was sampled at 5 FPS. Since the cameras were installed near the east side vehicle lane, vehicles travelling north along that lane were observed for a shorter duration in the camera frames due to the limited field of view compared to those travelling south along the west lane. As a result, most of the missing trajectories came from vehicles travelling along the east side vehicle lane.

Table 12: Analysis of video frame rate on vehicle tracking performance for a U-turn event.

Date	Time	Camera	Frame rate (FPS)	# Vehicle trajectories	# Vehicle trajectories identified	# U-turn events	# Vehicle trajectories identified for the U-turn event
21-Sep-2023	1756	PTZ_ingress	10	36	36	1	1
			5		13		1
			1		0		0

However, it is noteworthy to mention that the vehicle performing the U-turn was successfully tracked both at 10 FPS and 5 FPS due to the reduced speed of the vehicle while performing the event. This finding is crucial as the optimal frame rate for recognizing unsafe events involving vehicles could be effectively reduced to 5 FPS, thereby significantly reducing the processing demands. However, any other events that would involve tracking a vehicle at a greater speed (e.g., speeding vehicle) would require a frame rate higher than 5 FPS. None of the vehicle trajectories including the one performing the U-turn event could be identified when the video was sampled at 1 FPS as shown in the table.

A similar exercise was carried out for the intrusion/jaywalking events, but the results have not been presented in a tabular form. Since VRUs moved at a much slower speed in the videos gathered in the dataset, 5 FPS was observed to be adequate to track the VRUs for jaywalking/intrusion events. However, it remains to be investigated if 5 FPS would be adequate to track a running VRU, as there were no such VRU instances recorded in the collected dataset.

Figure 20 illustrates the impact of video frame rate on tracking performance. As shown in the top row, the trajectory of the vehicle (denoted by a white arrow) travelling along the near lane was not identified when the video was sampled at 5 FPS. However, the vehicle performing the U-turn event was successfully tracked at both 5 FPS and 10 FPS.

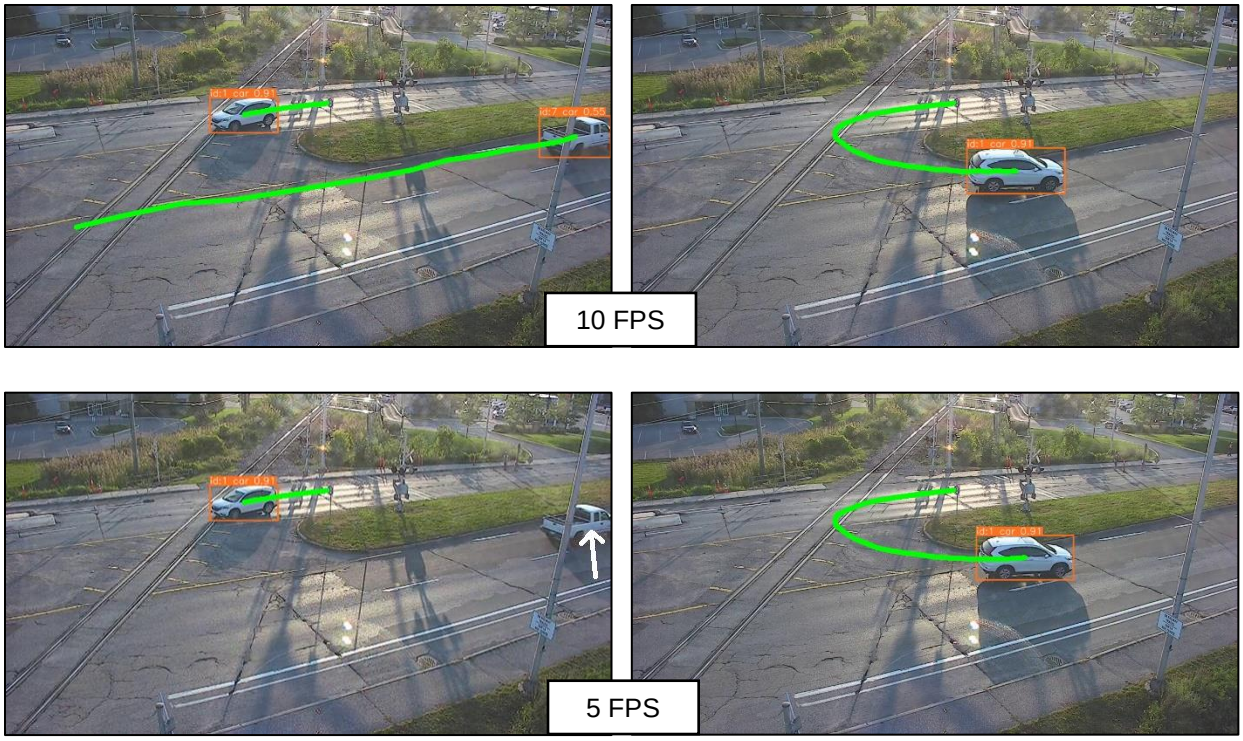

Figure 20: Illustration of the impact of video frame rate on vehicle tracking performance. Top-row shows the results when tracking was performed at 10 FPS, while bottom row shows results for 5 FPS. Trajectory of each vehicle, as identified by the tracking algorithm, is marked in green. While the trajectory of the vehicle at the near lane (denoted by white arrow in the bottom-left image) could not be identified at 5 FPS, the U-turn vehicle's trajectory was successfully identified both at 5 FPS and 10 FPS.

5.1.5 Sensor-specific image artifacts and their impact on detection performance

While the preceding sections offered a quantitative analysis of the performance of vehicle/VRU detection and tracking across the PTZ and thermal camera sensors, this section attempts to pinpoint the different peculiarities and artifacts manifested in the images captured by the different sensors. Additionally, it presents a qualitative analysis of how these artifacts impacted the detection performance across the sensors.

Apart from the blurry and pixelated image issue observed in the PTZ camera images captured during the night time (refer to section 5.1.2 and Figure 15 (right)), there were three other main artifacts noticed in the PTZ camera images. These included the presence of strong shadow, glare, and lens flare which manifested in the form of a reflection of the camera lens itself in the captured images. These artifacts were primarily observed in the third batch of data, likely because the third batch was collected in September when the sun's angle was different than it was in July, when the other two batches were collected.

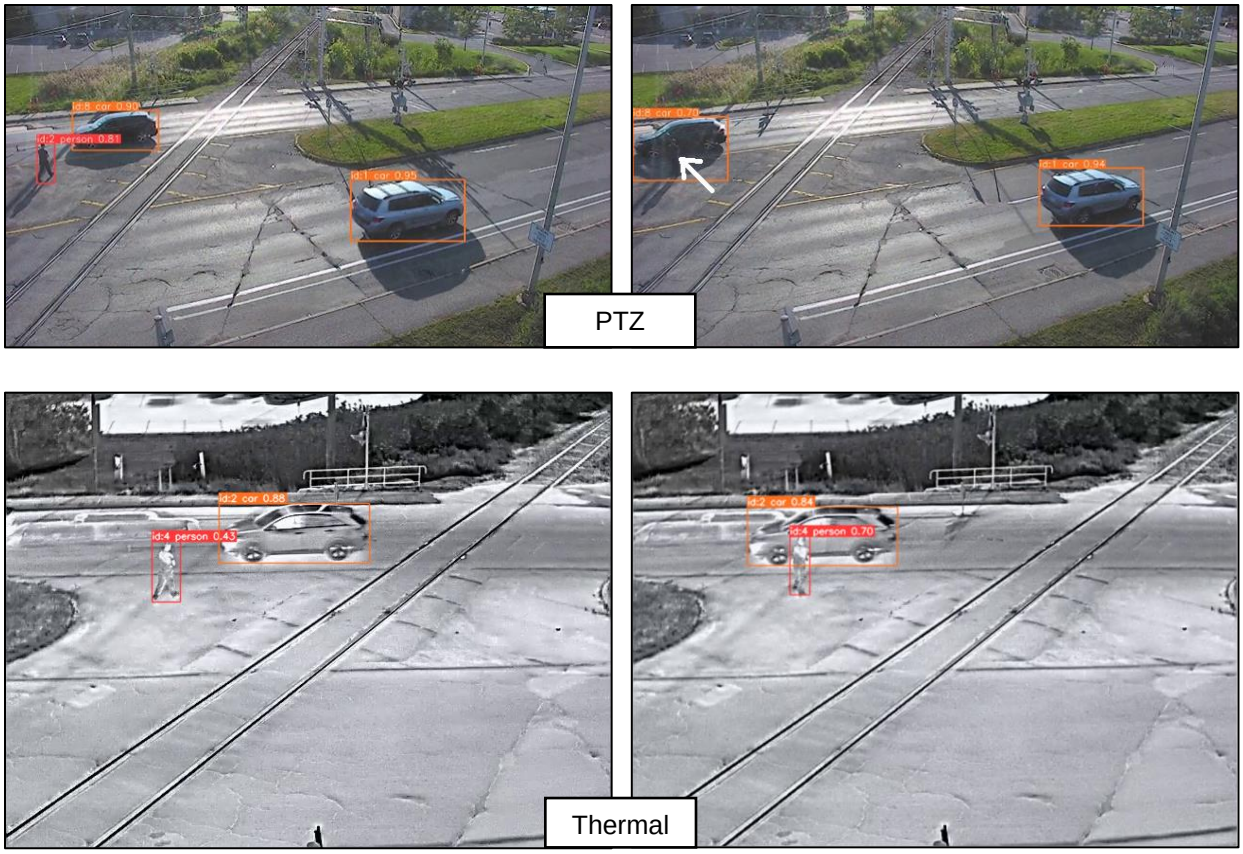

Figure 21: Top row shows detection results on two example images captured by the PTZ_ingress camera under strong shadow, while the bottom row shows the detection results for the corresponding images captured by the Thermal_egress camera. In the top-left image, the person was detected accurately until a subsequent frame when the person was completely covered by strong shadow. The missed person due to the passing-car strong shadow that occurred in the next frame sequence (top-right image) is shown by a white arrow. Detection on the thermal camera images remained unaffected by shadow.

The detection performance was observed to decline due to the presence of strong shadows, resulting in both false negatives as well as poor localization of objects. Figure 21 shows examples of detections on images from both the PTZ and thermal cameras under strong shadow conditions. As illustrated in the figure, the VRU in the top-right image went undetected because it was completely covered by strong shadow from a passing car, even though it was detected accurately in the preceding frame sequences. Furthermore, the strong shadow from the car caused the model to incorrectly extend the car's bounding box to include the shadow. In contrast, detections on the corresponding thermal camera images were unaffected.

Figure 22: Example of detections on PTZ_ingress camera images under strong glare conditions; (left image) the white car on the top-right corner of the image travelling along the far lane was accurately detected before entering the glare zone; (middle image) the car was missed in the next frame sequence as it passed through the strong glare zone (denoted by the white arrow); (right image) the car was detected again as it left past the glare zone. (Best viewed when zoomed in.)

Strong glare was another prominent artifact that posed challenges for the model to detect objects from the PTZ camera images, eventually causing missed detections. Figure 22 presents some visualizations of some example detections under strong glare conditions. As the figure illustrates, the glare zone caused the white car travelling on the far lane to go undetected as it passed through that zone (middle image, marked by the white arrow), though it was accurately detected in the preceding and following frame sequences (left and right image, respectively).

Furthermore, the lens flare phenomenon, which occurs when strong light directly hits the camera lens, was frequently observed in the third batch of PTZ camera images. Images affected by lens flare would typically display a bright, often circular or semicircular flare or artifact that originates from the camera lens itself. The result is usually a series of bright spots or a hazy, washed-out area in the image. Detection performance was observed to decline due to the lens flare effect in the PTZ camera images. Figure 23 shows some example detections on one PTZ camera image that was affected by lens flare along with the detections on the corresponding thermal camera image.

Figure 23: Lens flare causing the pedestrian (encircled in white on the left image) to become completely obscured in the PTZ_ingress camera image. The corresponding thermal image (right) allowed the model to detect the pedestrian accurately.

One image artifact issue that was found to degrade the thermal camera images was thermal washout. A thermal washout image loses clarity and contrast because of excessive heat or extended exposure to high temperatures. This issue typically arises when the thermal sensor or camera overheats, resulting in less accurate thermal readings and an image that looks overly bright or "washed out". Consequently, important details become obscured, making it difficult to identify temperature differences, hot spots, or thermal leaks. This essentially caused the detection model trained on thermal camera images to fail to detect any objects in the thermal washout regions.

Table 13 shows nine separate video minutes in the collected dataset that were affected by thermal washout. It also reports the extent to which the images were impacted by this phenomenon. It appears that only the Thermal_ingress camera exhibited this phenomenon in the collected dataset. Figure 24 shows two cases where detection failed due to thermal washout in the images. The top left shows a case where a pedestrian was not detected because it was completely obscured by a washed-out region. The top-right shows the pedestrian was once again successfully detected after leaving the washed-out region. Similarly, the vehicle in the bottom-left image was accurately detected up until the point it entered into the washed-out region as shown in the bottom-right image.

Table 13: Details of the video segments in the collected dataset that were affected by thermal washout.

Camera	Date	Time	Washout severity and extents
Thermal_ingress	17-Jul-2023	0900	Light washout on the top-left corner of the images
		1245	Light washout on the top-left corner of the images
		1250	Light washout on the top-left corner of the images
		1635	Heavy washout on the left half of the images
		1640	Heavy washout on the left half of the images
		1845	Light washout on the top-left corner of the images
		1850	Light washout on the top-left corner of the images
		1855	Light washout on the top-left corner of the images
	26-Jul-2023	2000	Light washout on the top-left corner of the images

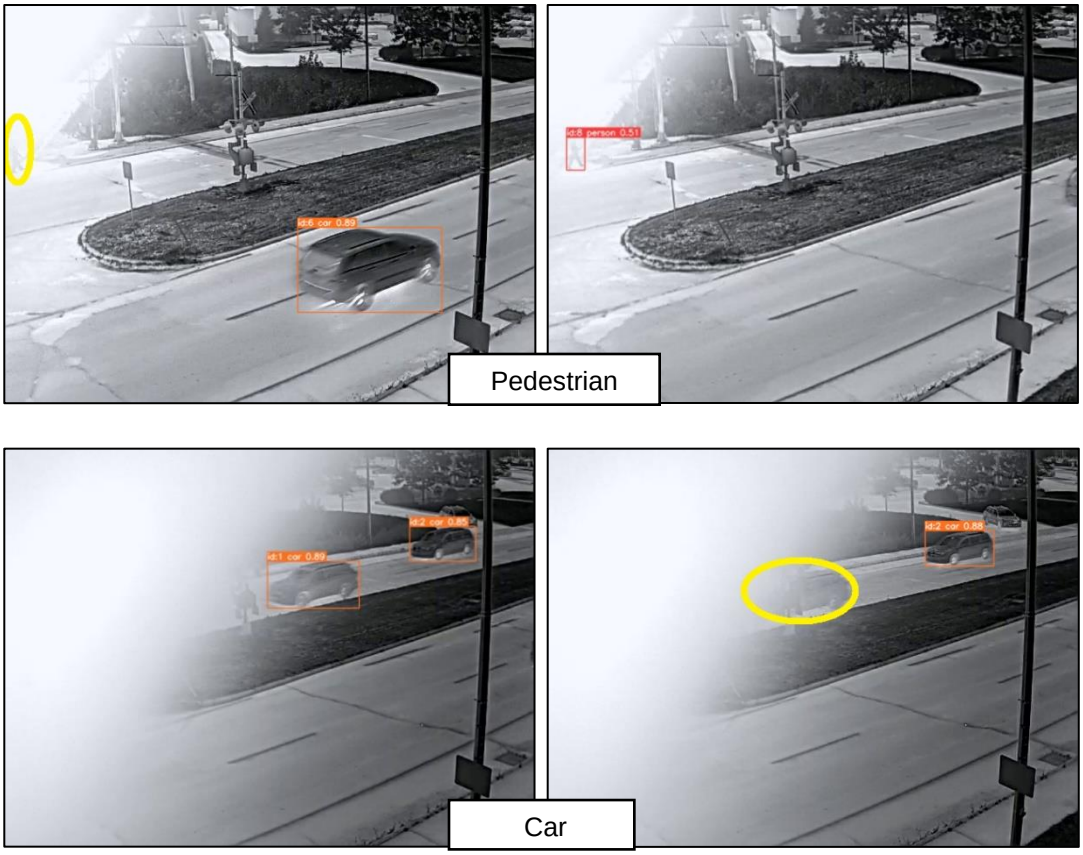

Figure 24: Example images captured by the PTZ_ingress camera exhibiting a thermal washout phenomenon. In the top-left image, the pedestrian (encircled in yellow) was buried under washout was missed by the model, but was detected as soon as it appeared out of the washout region (top-right). In the bottom pair of images, the car was detected successfully (bottom-left image) until it entered into the washout region (marked in yellow in the bottom-right image).

5.2 Detection and recognition of unsafe events

This section presents the results of the event detection algorithms described in section 4.3.

5.2.1 U-turn event

The ground truth (GT) in the test set consisted of 10 illegal U-turn events. Table 14 shows the algorithm detected all 10 U-turns with no false positive or false negative results. This represented a 100% success rate of the algorithm for this test set.

Table 14: Confusion matrices for the U-turn algorithm applied to 4 camera views.

Camera view	PTZ_ingress	PTZ_egress	Thermal_ingress	Thermal_egress
Number of events	22	5	14	3
GT (actual U-turn)	5	2	2	1
TP	5	2	2	1
FP	0	0	0	0
TN	17	3	12	2
FN	0	0	0	0

Figure 25 shows an illegal U-turn event captured by both the PTZ ingress and PTZ egress cameras.

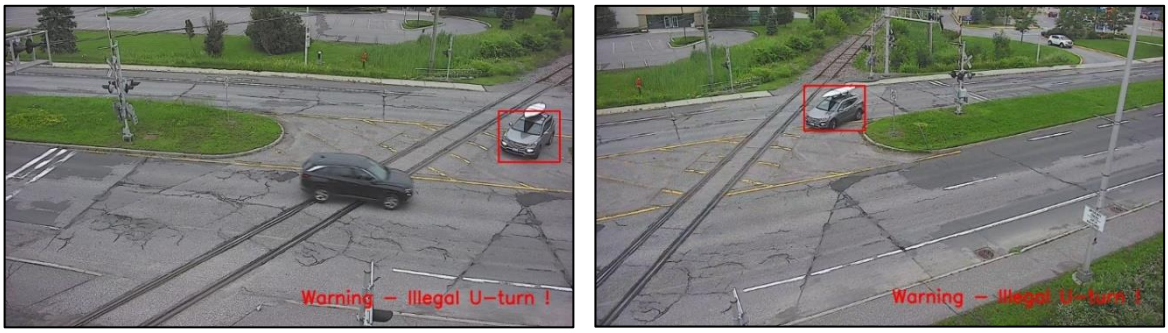

Figure 25: Illegal U-turn event detected from PTZ_ingress (right) and PTZ_egress (left) views.

Figure 26 shows an illegal U-turn event captured by the thermal ingress and thermal egress views. The detection of the U-turn event was confirmed by combining the decision results obtained from several camera views.

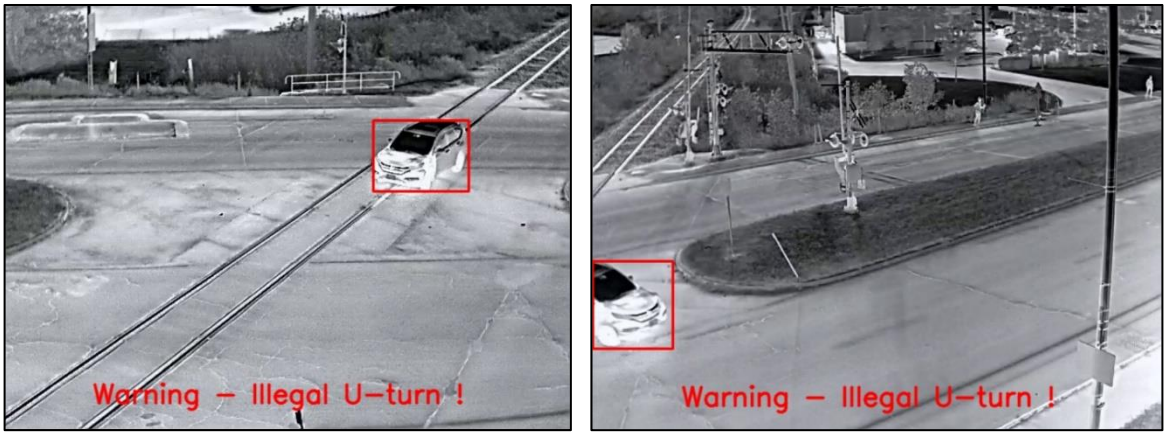

Figure 26: Illegal U-turn event detected from thermal ingress (right) and thermal egress (left) views.

5.2.2 Stopped on the crossing event

The test set included two stopped on the crossing event. Table 15 shows that, out of all 45 vehicle events, the algorithm detected 2 stopped on the crossing event with no false positive or false negative results. The success rate of the algorithm for this test set was thus 100%.

Table 15: Confusion matrices for the stopped-on-crossing algorithm.

Camera view	PTZ_ingress	PTZ_egress	Thermal_ingress	Thermal_egress
Number of events	22	5	14	4
GT (stopped on crossing event)	1	0	0	1
TP	1	0	0	1
FP	0	0	0	0
TN	21	5	14	3
FN	0	0	0	0

Figure 27 shows the algorithm result when detecting a vehicle stopped on the crossing (railway track).

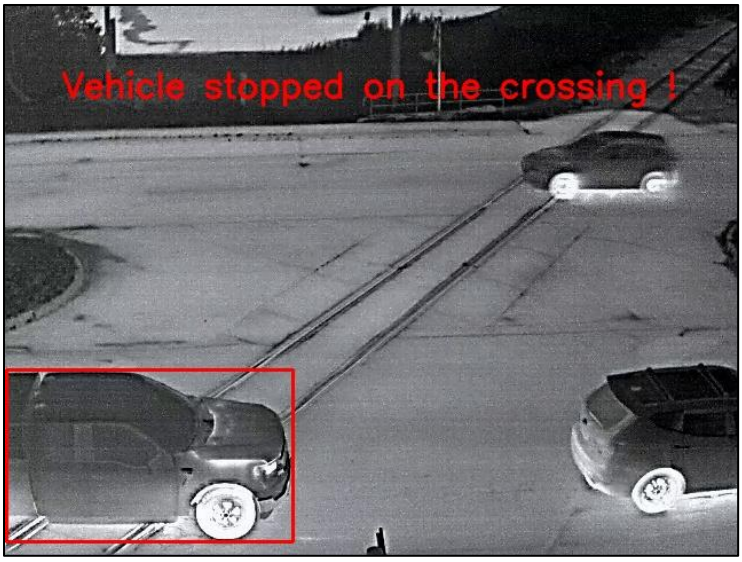

Figure 27: Vehicle (with red rectangle) stopped on the crossing detected by the algorithm.

5.2.3 Stopped elsewhere event

The results of the algorithm for detection of vehicles stopped elsewhere are shown in Table 16. The dataset included 9 stopped elsewhere events. The algorithm correctly detected all stopped elsewhere events with no false positives or false negatives. The success rate of the algorithm for this test set was 100%.

Table 16: Confusion matrices for the stopped elsewhere algorithm.

Camera view	PTZ_ingress	PTZ_egress	Thermal_ingress	Thermal_egress
Number of events	22	5	14	3
GT (Stopped elsewhere)	2	2	4	1
TP	2	2	4	1
FP	0	0	0	0
TN	20	3	10	2
FN	0	0	0	0

Figure 28 shows a stopped paramedic vehicle detected as stopped elsewhere by the algorithm.

Figure 28: Vehicle (with a red rectangle) detected by the algorithm as stopped elsewhere.

5.2.4 Queuing event

Table 17 presents the results of the queuing detection algorithm. The dataset included 24 queuing events. The algorithm detected all queuing events with no false positives or false negatives.

Table 17: Confusion matrices for the queuing algorithm.

Camera view	PTZ_ingress	PTZ_egress	Thermal_ingress	Thermal_egress
Number of events	22	5	14	3
GT (queuing event)	14	1	8	1
TP	14	1	8	1
FP	0	0	0	0
TN	8	4	6	2
FN	0	0	0	0

Figure 29 shows the detection results of a queuing event involving four vehicles located in traffic lane 3 (see Figure 11 in section 4.3.3 for more information on the traffic lanes). The top row in Figure 29 shows that as soon as two vehicles in succession are stopped, the algorithm recognizes that both vehicles (with a red rectangle) are queuing. Even though the stopped vehicles are past the railway tracks themselves, there is a risk that other vehicles may come along later and stop in a place where they block the tracks themselves.

The bottom row shows that other vehicles (each with a red rectangle) joining the queue are also recognized as part of the same queue event.

Figure 29: Queuing event involving four vehicles (each with a red rectangle) detected by the algorithm on lane 3.

Figure 30 shows the detection results of a queuing event involving three vehicles located on the fourth traffic lane. In the top row, two stopped vehicles (each with a red rectangle) following one another are identified as queuing. In the bottom row, a third vehicle (with a red rectangle) slowed down and then stopped, and so was recognized as queuing.

Figure 30: Queuing event involving three vehicles (each with a red rectangle) detected by the algorithm in lane 4.

5.2.5 Jaywalking event

Table 18 presents the results of the jaywalking detection algorithm. The dataset included 29 jaywalking events and the algorithm correctly detected 27 events. The two missing events were not real false negatives because two pedestrians who were walking side by side were detected as a single pedestrian by the detection algorithm. It was expected that the detection algorithm would only recognize the pedestrian facing the camera because the second one would have been hidden by him (see Figure 31). Thus, the jaywalking algorithm only received the trajectory of one pedestrian. Figure 32 shows the detection result of a jaywalking event.

Table 18: Confusion matrices for the jaywalking algorithm.

Camera view	PTZ_ingress	PTZ_egress	Thermal_ingress	Thermal_egress
Number of events	7	14	2	17
GT (jaywalking event)	5	10	2	12
TP	5	10	2	10
FP	0	0	0	0
TN	2	4	0	5
FN	0	0	0	2

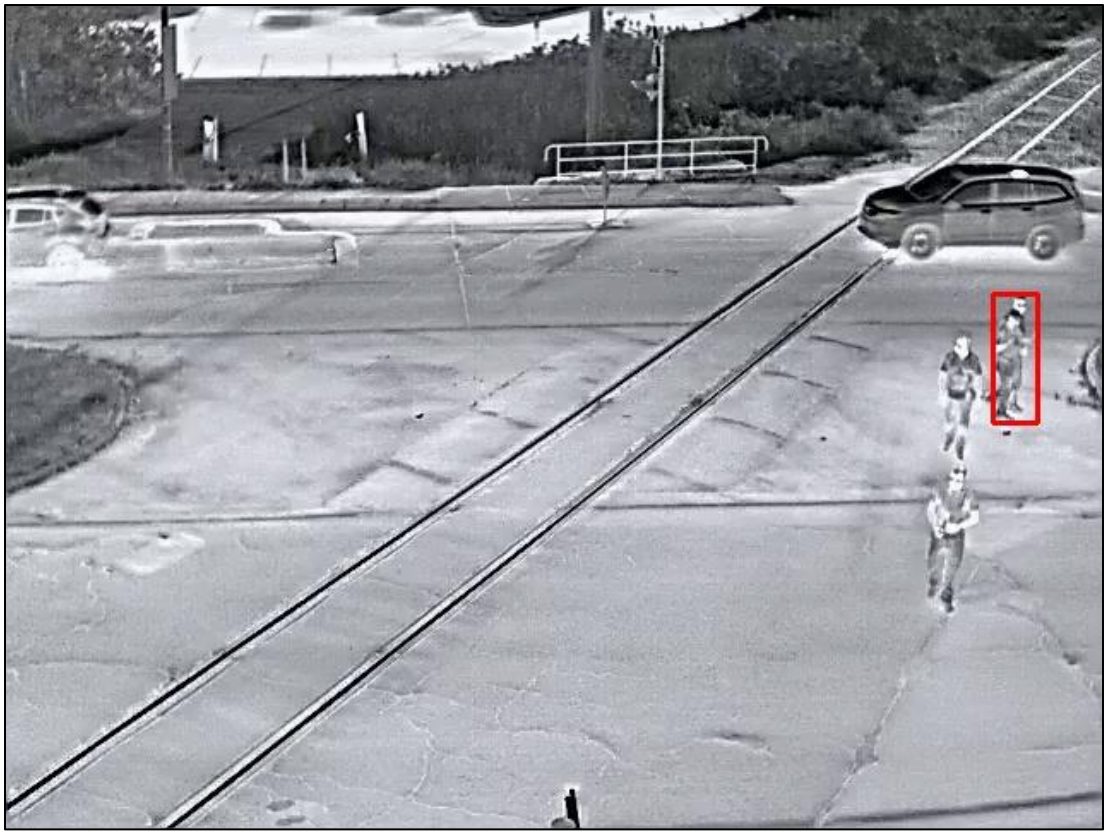

Figure 31: Two pedestrians (with the red rectangle), walking side by side, were detected as one person.

Figure 32: Detection of a jaywalking event.

5.2.6 Intrusion event

Table 19 presents the results of the intrusion detection algorithm. A pedestrian crossing the railway track was recognized by the algorithm as an intrusion event in the danger zone. The dataset included 11 intrusion events and the algorithm correctly detected 10 events. The missing event was not a real false negative because it was related to the same scenario seen in section 5.2.5. Two pedestrians who were walking side by side were detected as one person by the detection algorithm due to the occlusion issue (see Figure 33 for the intrusion example and Figure 31 for the similar jaywalking example). Thus, the intrusion algorithm only received the trajectory of one pedestrian in that event. Figure 34 shows the detection result of an intrusion event.

Table 19: Confusion matrices for the intrusion algorithm.

Camera view	PTZ_ingress	PTZ_egress	Thermal_ingress	Thermal_egress
Number of events	7	14	2	17
GT (intrusion event)	2	4	0	5
TP	2	4	0	4
FP	0	0	0	0
TN	5	10	2	12
FN	0	0	0	1

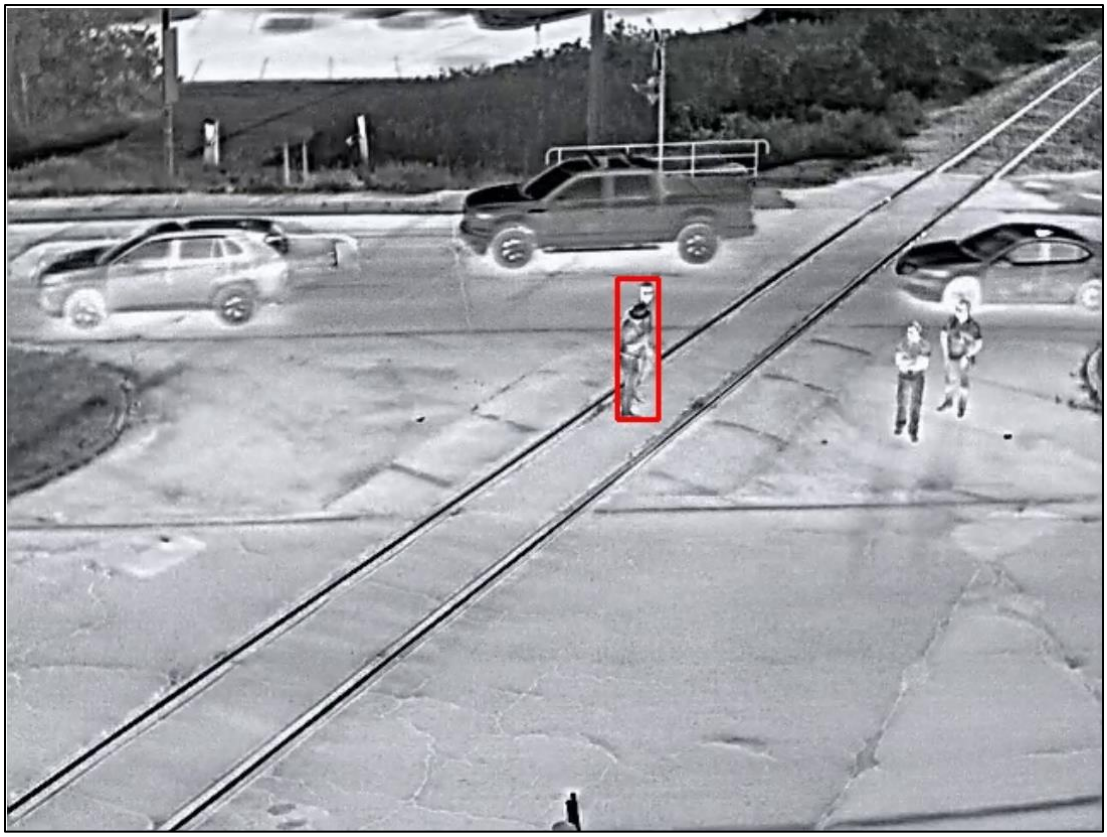

Figure 33: Two pedestrians (with a red rectangle), walking side by side, were detected as one person due to the occlusion issue.

Figure 34: Detection of an intrusion event, a pedestrian crossing the railway track.

5.3 Summary

Section 5 presents the results of vehicle and VRU detection and tracking based on multi-modal vision sensors including PTZ and thermal cameras. Additionally, the results of the algorithms developed to recognize the potentially unsafe events involving and vehicles and VRUs at or near a grade crossing are presented.

6 Key findings and lessons learned

This section presents a consolidated summary of the key findings and lessons learned in this project.

6.1 Key findings

The key findings that emerged from the literature review, the collection, processing and analysis of data, as well the development of deep learning algorithms for vehicle/VRU detection are summarized as follows:

- Potential opportunities to reduce the likelihood of grade crossing collisions using emerging machine vision technologies were identified through a comprehensive review of published reports on crossing collisions and crossing inspections.
- The analysis of real-world data collected from a public grade crossing validated the rail safety use case scenarios studied in this project.
- Vehicle queuing on either end of the crossing was observed to be particularly concerning in the collected dataset due to their prolonged duration which could lead to potentially dangerous situations causing vehicles to become trapped on the crossing surface – thereby increasing the chances of collision with a train.
- The performance of vehicle and VRU detection using advanced AI and deep learning algorithms based on RGB and thermal camera images was found to be satisfactory.
- Use of both RGB and thermal cameras to detect vehicles and VRUs at or near a grade crossing were deemed necessary due to the complementary information available from the different sensors.
- RGB cameras demonstrated excellent performance for vehicle detection during the day, while thermal cameras performed better at night.
- Thermal cameras were particularly well suited for VRU detection, both during the day and at night.
- Detection of vehicles under occlusion was observed to be better dealt with using RGB camera images.
- An optimal video frame rate of 5 FPS was found to be adequate to capture the trajectories of vehicles and VRUs involved in the potentially unsafe events captured in the collected dataset.
- Both types of cameras, especially the RGB cameras, were subjected to several different image artifacts causing a decline in performance for vehicle/VRU detection.

Additionally, the development of the event detection algorithms generated the following key results:

- Event detection algorithms mainly deal with the analysis and modeling of vehicle/VRU trajectories and therefore rely on the results obtained from the detection and tracking algorithms.
- The frame rate of the camera is an important parameter to quantify the vehicle's motion in pixels between two successive frames.
- The event algorithms have the potential to share the number and position of vehicles/VRUs involved in the event through a V2X communication system. The vehicle positions need to be

converted from camera coordinates to real world coordinates. Cameras should be calibrated to be able to perform this conversion.

- Vehicles and VRUs can be associated with several events. For example, a vehicle may be identified as queuing and stopped on the crossing at the same time.
- The detection of an event could be confirmed by combining the decision results obtained from different camera models (i.e., PTZ and thermal) and different views (i.e., ingress and egress).
- The development of deep learning based solutions for the detection and recognition of events is possible in the case of a large dataset.
- The automatic detection and recognition of the intersection configuration (e.g., danger zone, jaywalking area, etc.) is an important task when dealing with several grade crossings.

6.2 Lessons learned

The lessons learned through the execution of this project included the following:

- The visual artifacts observed in the thermal camera images (i.e., thermal washout) could be prevented by deploying anti-washout techniques (e.g., cooling, sun-protective enclosures etc.).
- The visual artifacts observed in the RGB camera images, especially lens flare, were primarily caused by the angle of sun to the camera sensors. Therefore, positioning of the cameras, particularly the azimuth angle, could play a vital role in preventing such artifacts from appearing in the first place.
- Detection of large vehicles (e.g., bendy buses, large trailer trucks, etc.) could be improved by using thermal cameras with a larger horizontal field of view.

Acronyms and Abbreviations

AI	artificial intelligence
CN	Canadian National Railway
CP	Canadian Pacific Railway
EDT	eastern daylight time
EST	eastern standard time
FPS	frames per second
GCR	Grade Crossing Regulations
GCWD	grade crossing warning device
ID	identification
IoU	Intersection-over-union
LiDAR	light detection and ranging
MST	mountain standard time
PDT	pacific standard time
PTZ	pan, tilt, zoom
RADAR	radio detection and ranging
RGB	red, green, blue
ROI	region of interest
TC	Transport Canada
TSB	Transportation Safety Board of Canada
V2P	vehicle-to-pedestrian
V2X	vehicle-to-everything
VRU	vulnerable road user

References

[1] "Rail transportation occurrences in 2022," Transportation Safety Board of Canada, [Online]. Available: https://www.tsb.gc.ca/eng/stats/rail/2022/sser-ssro-2022.html. [Accessed November 2023].

[2] Deloitte, "Implications of Connected and Autonomous Vehicles in Ontario: Insurance and data access & security," 2022. [Online]. Available: https://www2.deloitte.com/ca/en/pages/consulting/articles/connectedvehiculesontario.html. [Accessed June 2023].

[3] Transportation Safety Board of Canada (TSB), "Rail transportation safety investigation report R21H0087," 6 April 2022. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2021/r21h0087/r21h0087.html. [Accessed June 2023].

[4] J.K. Caird, J.I. Creaser, C.J. Edwards, and R.E. Dewar, "A Human Factors Analysis OF Highway-Railway Grade Crossing Accidents In Canada," Cognitive Ergonomics Research Laboratory, Department of Psychology, University of Calgary, Alberta, 2002.

[5] Transportation Safety Board of Canada (TSB), "Rail transportation safety investigation report R20D0013," 1 March 2021. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2020/r20d0013/r20d0013.html. [Accessed June 2023].

[6] Qualcomm, "Cellular -V2X Technology Overview," [Online]. Available: https://www.qualcomm.com/content/dam/qcomm-martech/dm-assets/documents/c-v2x_technology.pdf. [Accessed June 2023].

[7] Transportation Safety Board of Canada (TSB), "Rail transportation safety investigation report R19T0191," 9 February 2023. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2019/r19t0191/r19t0191.html. [Accessed June 2023].

[8] Transportation Safety Board of Canada (TSB), "Rail transportation safety investigation report R18V0127," 23 July 2019. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2018/R18V0127/R18V0127.html. [Accessed June 2023].

[9] Transportation Safety Board of Canada (TSB), "Rail transportation safety investigation report R18T0006," 17 July 2019. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2018/R18T0006/R18T0006.html. [Accessed June 2023].

[10] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R17H0015," 23 February 2018. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2017/r17h0015/r17h0015.html. [Accessed June 2023].

[11] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R16M0026," 15 February 2018. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2016/r16m0026/r16m0026.html. [Accessed June 2023].

[12] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R13T0192," 2 December 2015. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2013/r13t0192/r13t0192.html. [Accessed June 2023].

[13] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R13E0015," 12 August 2014. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2013/R13E0015/R13E0015.html. [Accessed June 2023].

[14] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R13D0001," 15 January 2014. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2013/r13d0001/r13d0001.html. [Accessed June 2023].

[15] Transportation Safety Board of Canada (TSB), "Rail Investigation Report R11T0175," 4 October 2012. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2011/r11t0175/r11t0175.html. [Accessed June 2023].

[16] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R09V0219," 21 July 2010. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2009/r09v0219/r09v0219.html. [Accessed June 2023].

[17] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R08T0158," 24 August 2009. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2008/r08t0158/r08t0158.html. [Accessed June 2023].

[18] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R08M0002," 17 September 2008. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2008/r08m0002/r08m0002.html. [Accessed June 2023].

[19] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R07D0111," 26 February 2009. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2007/r07d0111/r07d0111.html. [Accessed June 2023].

[20] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R05E0008," 9 January 2007. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2005/r05e0008/r05e0008.html. [Accessed June 2023].

[21] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R04C0110," 7 March 2006. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2004/r04c0110/r04c0110.html. [Accessed June 2023].

[22] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R04H0009," 10 January 2006. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2004/r04h0009/r04h0009.html. [Accessed June 2023].

[23] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R02T0149," 7 January 2003. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2002/r02t0149/r02t0149.html. [Accessed June 2023].

[24] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R00C0159," 4 February 2002. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/2000/r00c0159/r00c0159.html. [Accessed June 2023].

[25] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R99T0298," 11 September 2001. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/1999/r99t0298/r99t0298.html. [Accessed June 2023].

[26] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R99S0100," 17 December 2002. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/1999/r99s0100/r99s0100.html. [Accessed June 2023].

[27] Transportation Safety Board of Canada (TSB), "Railway Investigation Report R99H0009," 16 January 2002. [Online]. Available: https://www.tsb.gc.ca/eng/rapports-reports/rail/1999/r99h0009/r99h0009.html. [Accessed June 2023].

[28] "Notices and Orders," Transport Canada, [Online]. Available: https://tc.canada.ca/en/rail-transportation/enforcement-action-measures-mitigate-threats-rail-safety/notices-orders?region=ppoqa. [Accessed June 2023].

[29] ITS America, "The Future of V2X: 30 MHz Application Map," 2021. [Online]. Available: https://itsa.org/wp-content/uploads/2021/01/ITS-America-30-MHz-Application-Map-1-27-21.pdf. [Accessed June 2023].

[30] Government of Ontario, "Highway Traffic Act, R.S.O. 1990, c. H.8," [Online]. Available: https://www.ontario.ca/laws/statute/90h08#BK245. [Accessed June 2023].

[31] Goverment of Canada, "Grade Crossings Regulations (SOR/2014-275)," [Online]. Available: https://laws-lois.justice.gc.ca/eng/regulations/SOR-2014-275/page-4.html#h-808059. [Accessed June 2023].

[32] Transport Canada, "Grade Crossing Inventory," [Online]. Available: https://open.canada.ca/data/en/dataset/d0f54727-6c0b-4e5a-aa04-ea1463cf9f4c. [Accessed May 2023].

[33] DarkLabel, 2021. [Online]. Available: https://github.com/darkpgmr/DarkLabel.

[34] G. Singh, S. Akrigg, M. D. Maio, V. Fontana, R. J. Alitappeh, S. Khan, S. Saha, K. Jeddisaravi, F. Yousefi, J. Culley, T. Nicholson, J. Omokeowa, S. Grazioso, A. Bradley, G. D. Gironimo and F. Cuzzolin, "ROAD: The Road Event Awareness Dataset for Autonomous Driving," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 45, no. 1, pp. 1036-1054, 2023.

[35] M. A. Rahman and R. Laganière, "Spatio-Temporal Activity Detection via Joint Optimization of Spatial and Temporal Localization," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) Workshops*, 2024.

[36] G. Jocher, A. Chaurasia and J. Qiu, *Ultralytics YOLOv8,* 2023.

[37] Y. Zhang, P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, P. Luo, W. Liu and X. Wang, "ByteTrack: Multi-Object Tracking by Associating Every Detection Box," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022.

[38] M. Braun, S. Krebs, F. B. Flohr and D. M. Gavrila, "EuroCity Persons: A Novel Benchmark for Person Detection in Traffic Scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2019.

[39] Y. P. Gong, J. Cao, Y. Li, J. Xie, H. Sun and Jinfeng, "TJU-DHD: A Diverse High-Resolution Dataset for Object Detection," in *IEEE Transactions on Image Processing*, 2020.

[40] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick and P. Dollár, "Microsoft COCO: Common Objects in Context," *CoRR,* vol. abs/1405.0312, 2014.

[41] Qualcomm, "Cellular-V2X Technology Overview," [Online]. Available: https://www.qualcomm.com/content/dam/qcomm-martech/dm-assets/documents/c-v2x_technology.pdf. [Accessed 23 May 2023].

This page intentionally left blank

NRC.CANADA.CA • INFO@NRC-CNRC.GC.CA • 877-672-2672 •