Smart Intersections – IoT Insights using Video Analytics and Artificial Intelligence

Background

Like many other countries around the world, Australia is seeing significant population growth, particularly in our larger cities and built up urban environs. At the same time, there is a digital transformation underway, with Mobility-as-a-Service, smart phone ride-hailing and ride-sharing services, micro-mobility and other new forms of personal transport becoming increasingly available and adopted, altering the use of our core transport corridors and roads.

These two globally applicable mega trends are putting the existing transportation networks under increasing pressure, with the resultant traffic congestion outpacing the ability to invest in new infrastructure. At the same time, there is a worrying trend whereby vulnerable road users are increasingly at risk with a disproportionate increase in frequency of incidents for these road users.

Introduction

As road trauma rates increase amongst all road users including the likes of pedestrians and cyclists and congestion impacts our daily commute – new technologies are emerging to create an accurate picture of the road environment in real-time. Thus, enabling intersection and corridor risk profiles to be created which can be used to intervene and mitigate future incidents from occurring or understanding traffic flow across multiple modes.

In this trial, IoT, Video Analytics, Deep Learning (DL) and Artificial Intelligence (AI), for the purpose of traffic flow assessment and insights into road user behaviour, were evaluated at an intersection at the AIMES testbed in Melbourne[1] in partnership with: University of Melbourne, Department of Transport (DOT), IAG and Cisco.

Two use cases were conducted to evaluate the capability and accuracy of the technology:

Use Case 1 – Traffic Count
Count the number of road users traversing the intersection over an extended period of time.

Use Case 2 – Traffic Type Classification
Determine the type of traffic (car, truck, bus, vulnerable road users (VRU)) over a period of time.

1. https://eng.unimelb.edu.au/industry/aimes

Challenge

- While nearly 50% of all crashes within metropolitan area occur at intersections, lack of visibility into road environment makes it impossible to take proactive measures.
- There is a complete lack of data about road user interaction and near miss experiences.

Solution

- A first step towards being able to capture profile of an intersection and subsequently enable proactive intervention is through enabling connectivity at the intersections. Video Analytics and AI provide the connectivity mechanism between road users and the data network.
- Insights into traffic volumes and types delivered through Video Analytics, AI and Edge computing residing at roadside infrastructure.

Outcome

- New insights and visibility into volume and flow of various vehicle and pedestrian types across a prolonged period, thus providing a foundation for mitigation of incidents, better planning and traffic routing decisions.
- New insights into road user behaviour and traffic patterns connected with other systems and applications enable broader range of safety and road efficiency solutions.

Our major findings from this trial found that the technology provided:

- 90-95% accuracy in road user count in both day light and low light conditions
- 95% accuracy in classifying cars, trucks, buses and VRUs (pedestrians and cyclists combined)
- Detailed real-time directional traffic information that otherwise doesn't exist today. For example, it was observed that a typical traffic mix of the intersection includes high volume of VRUs. Observation showed on average a ratio of 78% cars, 12% VRUs, 8% trucks and 2% buses.

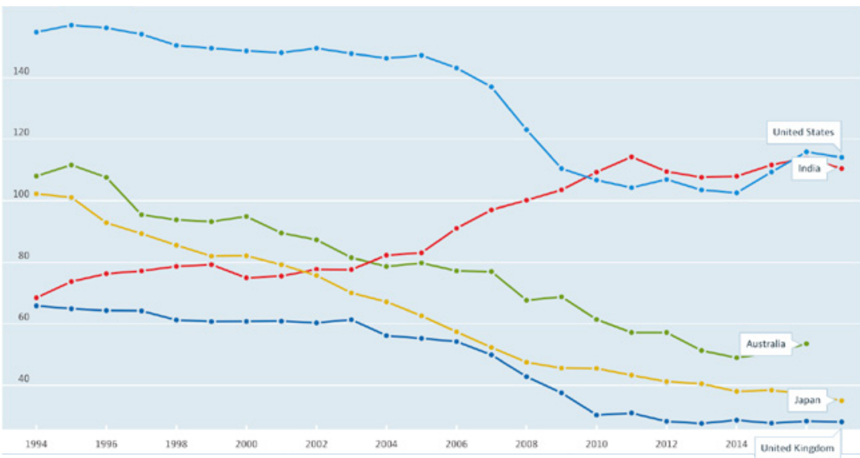
Future work will be undertaken to tune the AI/DL models to identify a broader range of road user types: for example, trams, types of trucks (semi-trailer; utility), cyclists/motorcyclist as well as introduce Deep Fusion Reasoning Engine (DFRE) for event and behavior recognition.

The rising cost of road trauma

Every year, more than 1 million lives are lost to traffic accidents around the world. With a relatively constant rate of road fatalities to population (~180/million), we are observing a steady increase in the number of lost lives year after year.

Australia has been among the countries with a better-than-average rate of fatalities to population and has been significantly improving from a rate above 100/million in 1990s to below 60/million in 2010s. However, after decades of improving, we are currently observing a worsening in trends. The number of road deaths recorded in Australia and the fatality rate to population has begun to rise since 2014.

Figure 1: Road Fatality Rates per Million Residents (Source: www.oecd.org)



Study of data between 2013 and 2019 by the University of Melbourne showed that 50% of all crashes within metropolitan areas occurred at intersections.

According to the most recent data from BITRE, there are around 100 road deaths per month in Australia and a third of these are pedestrians, motorcyclists and cyclists. Furthermore, for every fatal accident there are 30 others that require hospitalisation and nearly 6 times more non-hospitalised injuries.[2] Between 2001 and 2016, there has been an increase of 29% in hospitalised injuries from road crashes.[3]

Table 1: BITRE fatal accident data for the 12 months ending July 2019[4]

	2015	2016	2017	2018	2019
Driver	555	623	567	525	571
Passenger	251	208	255	205	223
Pedestrian	161	182	161	176	176
Motorcyclist	203	249	211	191	208
Pedal Cyclist	31	29	39	34	31
Total	1,204	1,293	1,223	1,137	1,214

Aside from the obvious and personal impact of road trauma on individuals, families and the society, there is also an economic cost estimated at about 30 billion[5] per year associated with roadway trauma. It is imperative to understand the roots of this recent change in trends and to identify possible solutions to help reverse it.

2. https://www.aaa.asn.au/wp-content/uploads/2018/03/AAA-ECON_Cost-of-road-trauma-summary-report_Sep-2017.pdf
3. Hospitalised Injury, BITRE. https://www.bitre.gov.au/publications/ongoing/hospitalised-injury.aspx
4. Australian Road Deaths Database, BITRE. https://www.bitre.gov.au/statistics/safety/fatal_road_crash_database.aspx
5. These are estimates of the total social costs for 2015 and 2016 consisting property damage costs, fatality costs and injury cost.

The value of risk profiling our roads

To explain and predict intersection crash probabilities, researchers have studied different measures for characterisation of intersection traffic conflict patterns.[6] Research indicates that intersection crash probabilities and severities can be predicted using surrogate models based on the intersection conflict patterns.[7] For car dominated intersections, the conflict patterns can be directly derived from the geometric design of the facility or simulated using traffic microsimulation. However, for multimodal traffic flow, which is becoming a dominant trend in our roadway space usage, the realisation of potential conflicts are significantly more difficult because:

1. As compared to car movements, pedestrians and bicycles display a more heterogenous and flexible use of roadway space. For example, a pedestrian jaywalking or a cyclist riding in the wrong direction or lane are more frequently observed than a car running the red light or driving in the wrong direction. As such the multimodal conflict patterns are difficult to predict from the geometric design of the intersection.

2. The existing traffic sensor infrastructure are geared towards counting cars and heavy vehicles at stop line and not over a distance, and as a result, the flow of lighter and more vulnerable road users are unknown and the multimodal conflict patterns are never directly recorded for them.

3. Microsimulation is another technique to characterise the intersection conflict patterns, However, the lack of real-world multimodal flow data (item 2) has impeded multimodal microsimulation.

New sensor technology such as Video Analytics and the broader Internet of Things (IoT) can therefore play a pivotal role in understanding the conflict patterns of multimodal traffic at intersections and the benefits of risk profiling of intersection to enable proactive intervention.

Understanding road user behaviour

With the increasing popularity of new transport modes, including ride-hailing and ride-sharing services, electric and shared bike and scooter systems, urban transport is bearing a paradigm shift in roadway space utilisation. This is a positive transition to multimodal and sustainable use of transport resources; however, the shared use of roadway space introduces inevitable and new conflicts among modes and significant safety risks, traffic jams and delays to all road users.

6. Simulation of safety: A review of the state of the art in road safety simulation modelling. W Young, A Sobhani, MG Lenné, M Sarvi. Accident Analysis & Prevention 66, 89-103.
7. Gettman, D. and Head, L., 2003. Surrogate safety measures from traffic simulation models. Transportation Research Record, 1840(1), pp.104-115.

The transport infrastructure of the future should be capable of accommodating the multimodal travel demands efficiently and at the centre of it lies the smart intersection.

Every intersection, depending on geometric design, neighbourhood land use and position in the network, may display different pattern of usage and variability of demand within and across days. Some intersections are dominantly used by private vehicles and some accommodate multimodal travellers. Depending on time of day and location, some intersections may also bear a heavy weight from freight trucks and heavy vehicles. Depending on patterns of users and demand volumes, urban intersections may expose their users to different levels and types of safety risks and efficiency issues. As such, understanding traffic volumes and types, dominant delays, potential conflicts, safety risks and near miss experiences carried out for individual intersections can shed light on what types of interventions are essential and what pieces of technology can serve those needs most effectively.

Connecting the unconnected

Living in a super connected and digitally rich world today, it is difficult to imagine that there is anything left to connect. But the reality is that we are only scratching the surface of possibilities. Across all industries and our personal lives, we are seeing physical objects, big and small connecting to computer networks for the purpose of extracting new insights, making these objects smarter, interactive and self-aware. Sensors attached to a bridge can tell us about its structural health. Water level sensors can provide early warning for road flooding. RFID tags can help track equipment in a hospital. All these are examples of Internet of Things (IoT) in action.

As part of the trial, Video Analytics and AI were used to provide the connectivity mechanism between road users and the data network. In the context of mobility and transportation, connecting the unconnected – intersections, bus stops, trains, traffic lights - represents a huge opportunity to (a) improve safety and efficiency; (b) create new services and opportunities through data insights and decisions.

Table 2: Smart Intersection Use Cases

Adjust level crossing timing to allow enough time for large group of pedestrians to cross
Prioritise heavy vehicles, to take them away from pedestrian streets near by
Evaluate intersection risk profile based on traffic mix observed
Add cycling lanes, based on volume of cyclists observed
Finetune traffic signals based on volume of traffic
Identify near misses between road users to improve intersection's safety

Making sense of the data with Artificial Intelligence, Machine Learning and Deep Learning

Road intersections can become more intelligent and self-reliant through the use of technologies such as IoT, Video Analytics, AI/DL and Edge computing. Combining Artificial Intelligence (AI) with IP based video analytics makes it possible to leverage visual analysis in a very targeted manner. Instead of constant indiscriminate video recording, it makes it possible to focus on specific objects and events of interest.

Machine Learning inference models can be trained to identify different events such as jaywalking, near-misses between road users, person down, crashes, objects on road and many others. Benefits of video analytics technology are twofold:

1. Reduce amount of data collected by focusing on data points of interest;
2. Reduce concern for personal privacy, as there is no constant video streaming recording.

At the trial intersection, this technology was used to differentiate and count different road users, such as cars, trucks, buses and vulnerable road users (VRUs).

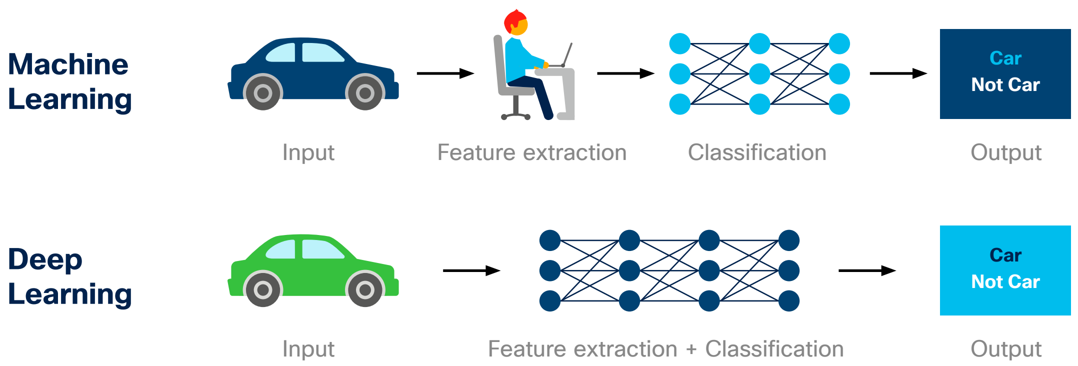
Edge computing played a major role in the project. It brings real-time processing, analyses and decisions closer to the data source to reduce time latency for sensitive and safety critical applications. Running video analytics at the edge, for example at an intersection, means that all the data processing takes place locally and only the data insights – for example aggregated traffic classification data, are sent to the Cloud.

It also makes it possible to take immediate action to events unfolding at an intersection. For example, extend pedestrian crossing time when a large group of pedestrians detected at an intersection.

Artificial Intelligence (AI) is driving powerful transformations across a broad spectrum of industries. AI is a system of solving complex problems and taking actions without human intervention. Machine Learning (ML) is the ability to "statistically learn" from data without explicit programming. AI and ML enable us to learn from data, identify patterns and make smarter decisions that augment human capabilities. Deep Learning (DL) is the process of self-learning and making decisions with complex data. As millions of physical objects get connected, by 2025 data volumes expected to reach 79ZB. By 2020, more than 1 billion cameras will be deployed by cities, generating almost 9 trillion hours of video recorded each year. AI, ML and DL become indispensable tools in making sense of all this data.

Machine Learning inference models can be trained to recognise certain objects or events, with a high degree of accuracy. For example, identifying objects such as buses, trucks, people, cyclists, wheelchairs etc or events such as jaywalking, man down, vehicles driving erratically or at high speed etc. Typically, Machine Learning models are built from previously classified data and require human guidance. If, for example, ML returns inaccurate prediction, a programmer would need to fix it. The key advantage of Deep Learning is that it is able to perform this learning on its own by using its own computational logic and meta algorithm. In an environment where cameras might move and would require calibration, Deep Learning simplifies operations by handling this autonomously and forms the basis of the AI component of the trial.

Figure 2: Autonomous learning enabled by combining reasoning and AI



Applying Artificial Intelligence and Deep Learning to Smart Intersections

This project belongs to a series of collaborative transport technology trials conducted within the Australian Integrated Multimodal Ecosystem (AIMES) which is an academia-industry-government partnership program founded by The University of Melbourne in 2016. The AIMES testbed is an urban ecosystem supporting the implementation, validation and testing of transport technologies and associated operational scenarios. The testbed provides a large scale, complex and multimodal urban environment for live transport technology and operations experiments. This area includes 100km of roadways bound by Alexandra Parade to the north, Victoria Street on the south, Hoddle Street on the east and Lygon Street on the west, plus the Eastern Freeway and Eastlink.

The Smart Intersection trial was focussed on proving the efficacy and accuracy of Cisco Intelligent Edge Video Analytics (CIEVA) for traffic flow assessment at one intersection within the AIMES testbed. This intersection (Nicholson St./Johnson) was chosen based on its layout and the mix of pedestrian, bicycle and vehicles at a high enough traffic density.

The main goal was to use Video Analytics, Artificial Intelligence (AI), Deep Learning (DL) and Edge computing to create a digitised understanding of traffic volume of all modes and deliver valuable insights into traffic behaviour and therefore risk.

Two use cases were designed for this trial:

Use Case 1 - Traffic Count

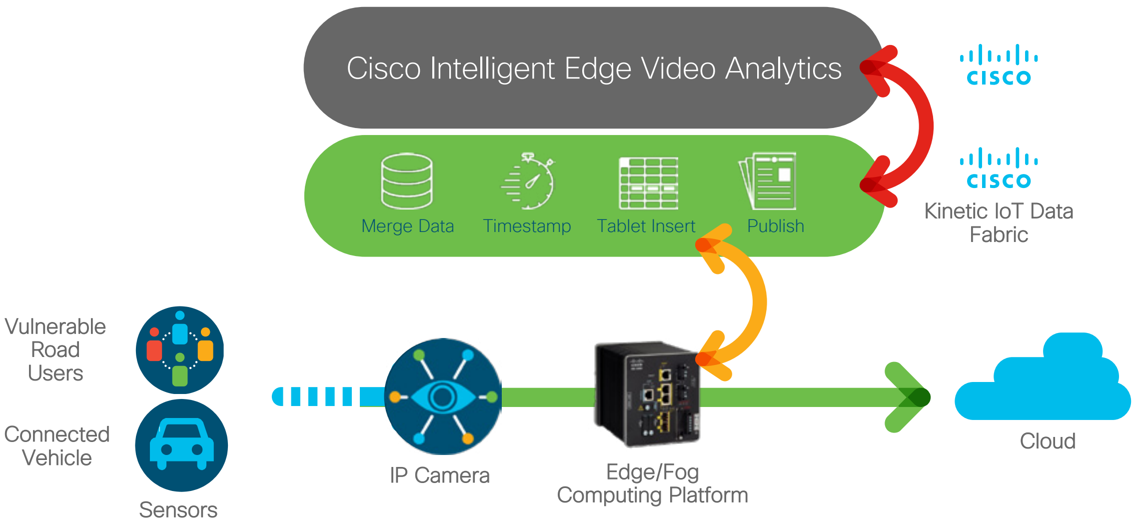
- In this use case, the number of road users and direction of travel through the intersection was measured over an extended period of time.

Use Case 2 - Traffic Type Classification

- In this use case, Cisco's IEVA artificial intelligence software was used to determine the type of traffic (car, truck, bus, pedestrian, bike, motorbike) traversing the intersection over a period of time.

The technology solution trialled in this project includes the Cisco Intelligent Edge Video Analytics (CIEVA) platform, Cisco IP Cameras, IC3000 with on-board edge compute capability as well as a dedicated fog compute node. The physical equipment was mounted on poles overhead or in roadside cabinets. Local network connectivity was provided using ethernet cables which also supply power using Power over Ethernet (POE). Finally, the Wide Area Network (WAN) connectivity is provided via mobile wireless 4G/LTE services.

Figure 3: High level Solution overview



The Cisco Integrated Edge Video Analytics (CIEVA) platform is a software application that provides edge compute optimised Deep Learning models that can detect and count people, vehicles, objects, and can take action on detection of events via an alert notification sub-system. During this project, existing CIEVA models were used to identify and count trucks, cars, motorcycles, buses and vulnerable road users as they travelled through the intersections. Two IP cameras were used to provide full coverage of the intersection.

Figure 4 : Intersection layout

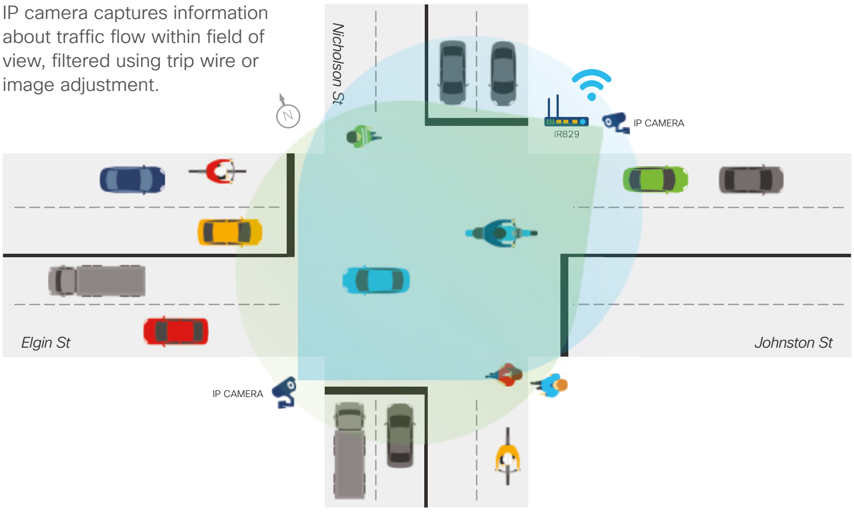


Figure 5: Applying AI to a single video frame

To capture traffic pattern and road user behaviour at an intersection using AI/DL we defined a number of key parameters such as geo-zone, virtual tripwires, objects of interest and directions of travel. Geo-zone (blue square in figure 5) defines an area of interest and makes it possible to keep count of objects within it at any given time. Tripwires (orange line in Figure 5) provide a digital point of reference for identifying when an object is crossing the intersection. The combination of these identifiers enables directional traffic analysis.

As road users traverse the intersection, video frames are taken every second and analysed in real time locally at the intersection by leveraging Edge computing to pick up specific information that we are interested in. Trucks, buses, trams, vulnerable road users, motorcycles were identified and counted. Only the metadata related to the actual capture (e.g road user count, direction of travel) is sent to the Cloud for storage, further analysis and visualisation. No video recording is stored at either the intersection or the Cloud, thus providing privacy protection.

Results

The project concluded the following accuracy and capability levels when comparing results with a manual observation of the road traffic environment:

Use Case 1 - Traffic Count

Count the number of road users traversing the intersection over an extended period of time.

- 90-95% accuracy in road user count in both day light and low light conditions.

Use Case 2 - Traffic Type Classification

Determine the type of traffic (car, truck, bus, vulnerable road users (VRU)) over a period of time.

- 95% accuracy in classifying cars, trucks, buses and VRUs (pedestrians and cyclists combined).
- Detailed real-time directional traffic information that otherwise doesn't exist today. For example, it was observed that a typical traffic mix of the intersection includes high volume of VRUs. Observations on average showed a ratio of 78% cars, 8% trucks, 2% buses and 12% VRUs.

From the results it can be seen that the project was successful with a high level of accuracy obtained in counting the number of road users. In addition, the ability to classify the road user types also proved successful.

Beyond IoT insights

The technology verified as part of this trial lays a foundation for future solutions focusing on safety and efficiency of our roads using AI based Video Analytics. The current setup extended across a network of intersections and corridors can expand the depth and breadth of insights collected.

Additionally, further capabilities using AI technology can extend the solution beyond identifying and counting object to also capturing traffic queue length and high-risk events such as jaywalking and speeding. Monitoring road user (pedestrians, trams, vehicles etc) interactions can provide insights into near misses and improve the overall accuracy of the intersection risk profile.

Figure 6: Network of smart intersections and road corridors improve overall journeys for all road users

The insights and/or the individual data sets captured at an intersection can be integrated with other systems and applications thus enabling broader set of use cases targeting safety and efficiency of our roads. Future work in this area will be around the use of Artificial Intelligence to understand the interaction and behavior between multiple road user objects.

For example:

- Detect emergency vehicles, or heavy freight vehicles and hold light sequences to allow them priority right of way
- Detect vehicles driving in a wrong direction and alert authorities
- Detect vehicles parked illegally or broken down
- Detect erratic driving and alert authorities
- Extend pedestrian crossing time when large queues are detected
- Assess risk and take immediate proactive steps leveraging logic that sits at edge
- Cycle lane planning based on volume of users observed and rate of near misses

Conclusions

No doubt, rising road trauma and congestion are the two biggest issues facing transport agencies today. However, the first step in arresting these issues is understanding road user behaviour and then using the information to create a formidable set of mitigation actions.

Through this trial we have proven that technology today can play a major role in understanding road user behaviour and can be applied across intersections and corridors with minimal infrastructure, setup and training. Achieving a 95% accuracy level of traffic counts and the ability to identify vehicle types demonstrates that the maturity of the technology is at a level where it can be deployed and generate immediate benefits.

The trial concluded that the combination of IoT, Video Analytics, Artificial Intelligence, Deep Learning and Edge computing can successfully be used to achieve accurate data on the types and numbers of road users within a defined area. Furthermore, these set of capabilities can be extended to support and understand more complex behaviour including understanding and predicting road user interaction.

In collaboration with

Cisco would like to acknowledge the contribution of the following people:

Majid Sarvi (University of Melbourne), Neema Nassir (University of Melbourne), Daniel Hoyne (DoT), Russell Brett (IAG), John Judge (Innovation Central), Peter Papaioannou (Innovation Central), Shwetha Srinivasan (Innovation Central), Stu Breslin (Innovation Central), Simon Young (Cisco), Rohan Locke (Cisco), Helen Airiyan (Cisco), Fabrizio Guarnieri (Cisco).