# Turning Existing Traffic Cameras into Smart Sensors Using Artificial Intelligence and Edge Computing

**Proposed by the River Hawk Team:**

Zubin Bhuyan
Jennifer Capistano
Qilei Chen
Shania Ruiz
Xinzi Sun

**Supervised by:**
Benyuan Liu, Ph.D.
Yuanchang Xie, Ph.D.

**In Collaboration with:**
MassDOT

June 13, 2021

# Table of Contents

# List of Figures

# List of Tables

# 1. Overview

This report is prepared by the River Hawk Team from the University of Massachusetts (UMass) Lowell for the 2021 Transportation Technology Tournament (TTT) organized by the National Operations Center of Excellence (NOCoE) and the U.S. Department of Transportation (DOT) Intelligent Transportation System (ITS) Joint Program Office (JPO) Professional Capacity Building (PCB) program. This document establishes the Concept of Operations (ConOps) for a proposed ITS solution, which aims to turn existing traffic surveillance cameras owned and operated by the Massachusetts DOT (MassDOT) into smart sensors leveraging the recent developments in computer vision and Artificial Intelligence (AI) technologies. This report starts with providing some background information, followed by a description of the proposed solution and its implementation strategy. It then provides an analysis of the proposed solution. Following that, the feasibility of the solution and the proposed implementation schedule are discussed. Lastly, the various benefits of the solution are identified and presented.

## *1.1 Background*

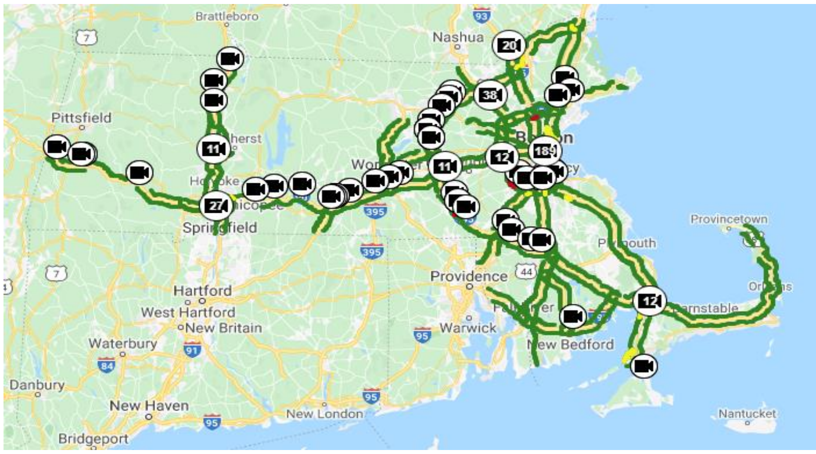


**Figure 1.1 Traffic Cameras in Massachusetts**

MassDOT has approximately 370 traffic cameras installed along its major roadways (See Figure 1.1). The videos captured by these traffic cameras can be streamed to the MassDOT Highway Operations Center (HOC) and displayed on the MassDOT's traveler information system (TIS) website. The TIS website provides real-time traffic, traffic camera feed, roadwork, and incident status for the public to access. Currently, these cameras are used only for incident verification and traffic conditions monitoring, which are conducted manually. Although the captured traffic videos contain useful information, they are not further analyzed and often are not archived. Such practices are prevalent in other states as well. It would be very interesting to equip these cameras with automated data collection and incident detection features, setting up an excellent example of low-cost solutions to retrofit old ITS equipment and maximize their service life and benefits.

## *1.2 System Requirements*

Figure 1.2 illustrates how the proposed solution works, and more details can be found in Section 2. An inexpensive and low power consumption edge computing device will be added to each camera to process traffic videos and generate valuable traffic parameters. The proposed system should meet the following:

1. *Significant Benefits*: The system can automatically detect lane changes and generate lane-wise traffic classification counts, time headway, and speed. Such data are essential for TIS

and Transportation Systems Management and Operations (TSMO) and can be further used to detect stop-and-go traffic, incidents (e.g., significantly reduce incident detection time), speeding and other unsafe activities, etc.

2. *High Accuracy*: It should be highly accurate and reliable to be useful in practice.
3. *Low Cost*: It is based on existing traffic cameras and requires minimum investment.
4. *Easy Maintenance*: The added hardware can be maintained together with existing cameras. The developed algorithm can be downloaded and updated remotely on the added edge computing device. This allows new features (e.g., detecting stalled vehicles in the emergency lane) to be easily implemented (i.e., no field trips needed).
5. *High Scalability and Flexibility*: The system can be applied to both existing cameras and new cameras MassDOT may want to install in the future.

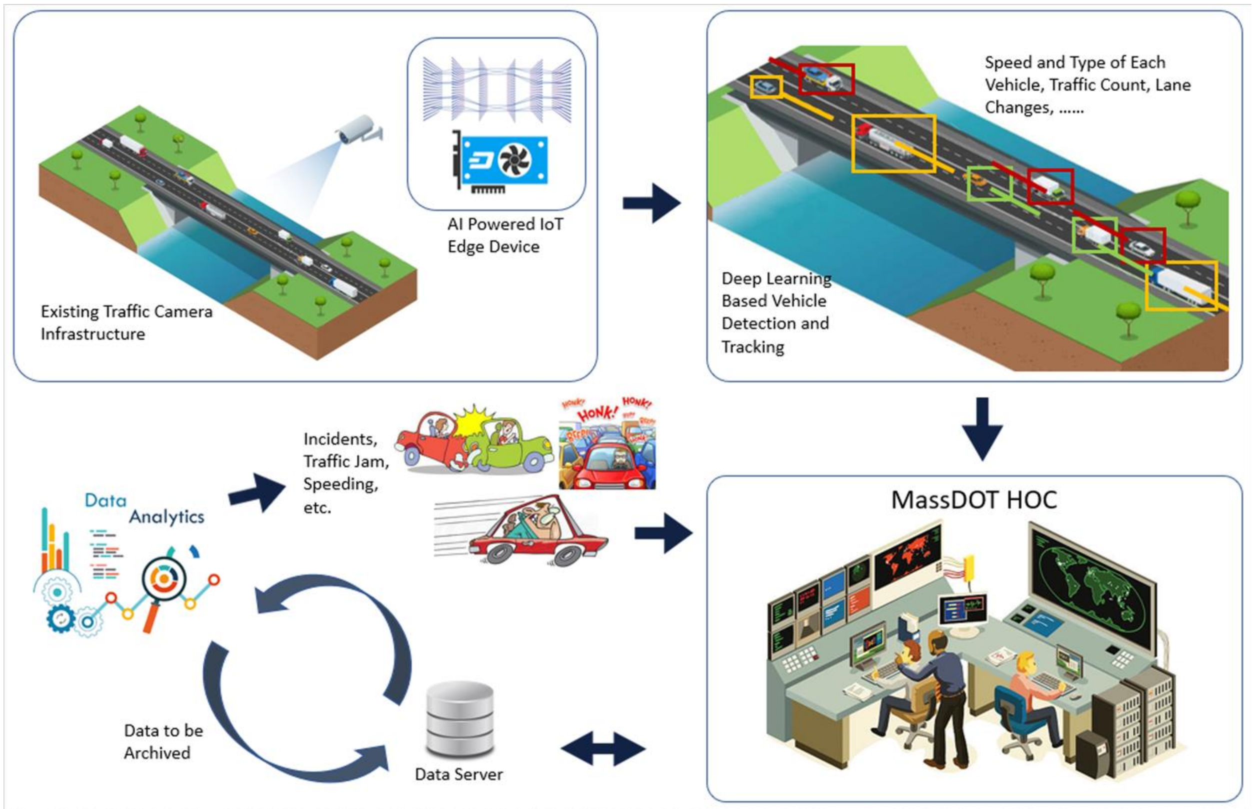


**Figure 1.2 Overview of the Proposed Solution**

## *1.3 Stakeholders*

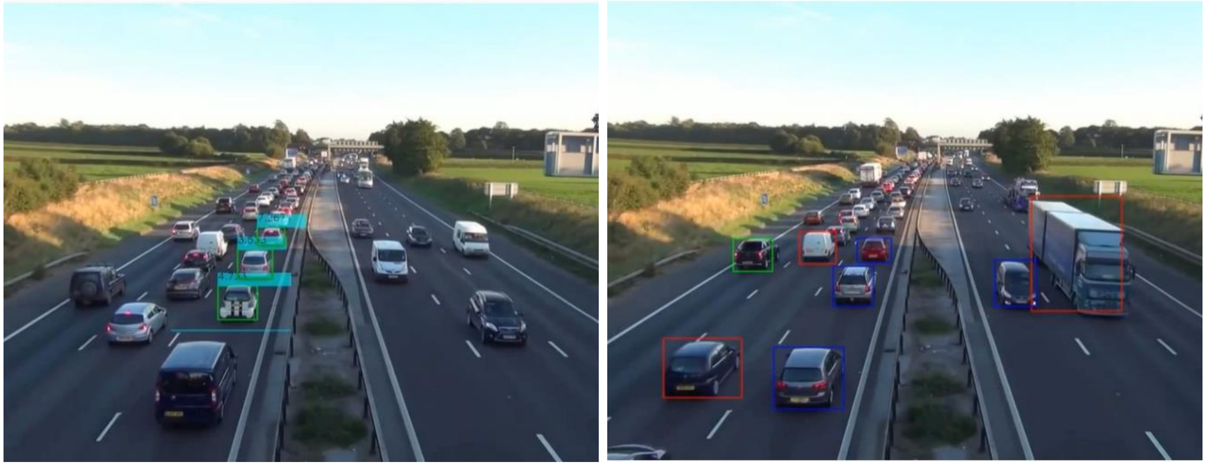
Any parties that may benefit from or be affected by the proposed solution are considered stakeholders. The stakeholders identified for this project include, but not limited to, the following:

- MassDOT ITS, HOC, Safety, Planning, and Performance Management divisions (e.g., TIS, incident detection, travel time reliability, and crash risk vs. real-time traffic patterns)
- City/Town Traffic Department (e.g., traffic monitoring, incident detection)
- Public Safety Agencies such as Massachusetts State Police Department (e.g., targeted law enforcement based on spatiotemporal distributions of risky traffic patterns)
- Travelers who benefit from the TIS
- MassDOT Information Technology division (e.g., data security and system integration)

# 2. Proposed Solution Description

## *2.1 Overview of the Solution*

The main objective of this project is to make the best use of existing ITS infrastructure. The solution wishes to turn existing MassDOT traffic surveillance cameras into smart sensors using artificial intelligence (AI) technologies. The existing cameras are mainly used to verify incidents and monitor traffic conditions, such as stop-and-go traffic. The videos are reviewed manually as needed and often not archived, although MassDOT has the capability to stream and record the traffic videos. Different from loop detectors, the existing cameras are not currently being used for data collection purposes. On that note, the ITS technology being proposed will allow MassDOT to collect several key traffic parameters automatically, including speed, time headway, number of lane changes, and traffic count by vehicle class based on existing infrastructure. The generated data will enable automatic incident detection in real time or enhance existing incident detection methods with additional critical information. The proposed solution will be based on the existing cameras. The algorithm can be deployed at the edge, and there is no need to stream the entire traffic video to the Highway Operations Center (HOC) (unless mandated by MassDOT). If the algorithm detects an incident, it will immediately trigger an alert and start to stream video to the HOC. An HOC operator can then review the video and confirm if the incident really occurred. Besides incident detection, the algorithm can act like a data collector and stream the processed data (instead of raw videos) to the HOC. This will help MassDOT significantly cut data communication costs.



*(a) Time Headway Estimation* *(b) Lane-Wise Vehicle Detection*

**Figure 2.1 Sample Model Outputs**

Figures 2.1 shows screenshots illustrating lane-wise traffic counting and time headway estimation results. For lane-wise traffic counting (see Figure 2.1(b)), distinct colors are used for each lane to draw bounding boxes around detected vehicles. For time headways, the algorithm keeps track of when each vehicle crosses a predefined position on a lane, as shown by the straight line in Figure 2.1(a), and calculates time headways. The estimated time headways are shown in seconds above the bounding box.

The proposed solution is a typical Internet of Thing (IoT) application. These smart edge sensors powered by AI technologies will be the cornerstone of future smart transportation systems and cities. In a July 2020 FHWA report (Vasudevan et al., 2020), AI-powered traffic cameras are considered a very promising solution for detecting traffic incidents. In the future, the proposed solution can also provide critical traffic data to support connected and automated vehicles for route and lane choice decisions.

### *2.2 Technical Details*

This study utilizes YOLO (You Only Look Once) and Kalman filter to detect and track vehicles, respectively. YOLO is a very popular AI method that uses a single convolutional neural network (CNN) for computer vision-based object detection and recognition. It divides an image into smaller regions and classifies them into various classes marked by bounding boxes. YOLO has different versions, and YOLOv3 is chosen for this study due to its lightweight and outstanding performance (Redmon and Farhadi, 2018). Figure 2.2 (Xu et al., 2020) shows the architecture of YOLOv3.

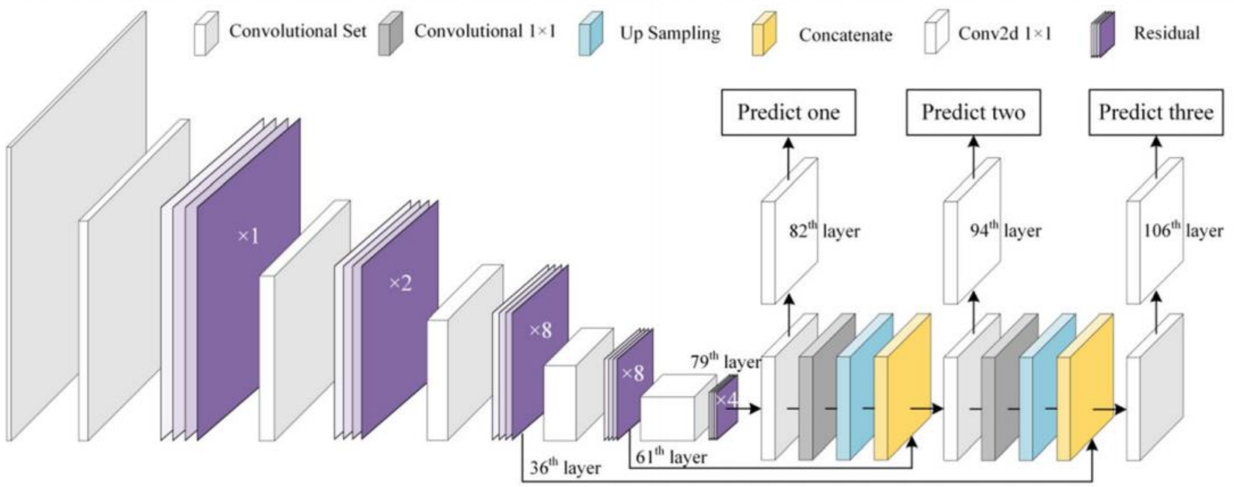


**Figure 2.2 Model Architecture of YOLOv3**

A significant challenge in developing deep learning models is to compile the training data. This study utilizes the Common Objects in Context (COCO) (Lin et al., 2014) dataset (comprising 23,300 annotated images belonging to '*car*', '*bus*', and '*truck*' categories), the STREETS dataset (Snyder and Do, 2019) (containing about 33,000 segmented vehicles in 3000 traffic images captured from 100 cameras in Chicago suburbs), and traffic videos provided by MassDOT (510 frames were extracted from traffic videos and manually annotated by the team).

**Table 2.1 Precision, recall, and F1 score of the best-performing model**

| | IoU = 0.5 | IoU = 0.3 |
|---|---|---|
| Precision | 0.82 | 0.92 |
| Recall | 0.87 | 0.89 |
| F1 | 0.84 | 0.90 |

Several trials with different combinations of training datasets and the YOLO with various settings were performed. For each trial, the data was split into training and testing sets with a 90:10 ratio. Data augmentation techniques were also used to expand the training set further and improve the trained model's generalization ability and performance. The performance was measured by *precision*, *recall*, and *F1 score* at two Intersection over Union (IoU) levels. The definitions of these performance measures and IoU can be found in Tan (2020). Higher performance measures mean better models. The best-performing model results are shown in Table 2.1, and some sample output videos are available at this *link*.

All computational experiments were performed on a system with the NVIDIA TITAN RTX graphics card with 24 GB of memory, which can process 20 video frames per second. A nice feature of YOLOv3 is its computation efficiency, making the trained algorithm suitable for deployment on edge devices such as the NVIDIA Jetson series (NVIDIA, 2021) in Figure 2.3. These small edge devices are of the size of a credit card. They have impressive computational capacity (Stein et al., 2019) and consume little power (5-15 Watt). This makes it possible to deploy the proposed solution at remote locations that use solar panels for power supply and 4G for data communications.

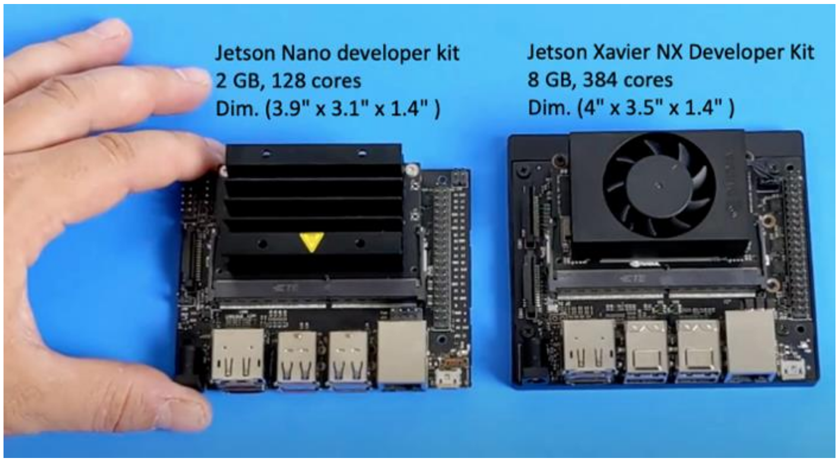


**Figure 2.3 NVIDIA Jetson Nano (left) and Jetson Xavier NX (right)**

### *2.3 Implementation Strategy*

Since MassDOT already has an extensive network of traffic cameras, we propose to deploy the developed YOLOv3 model on one of the small NVIDIA Jetson devices (Figure 2.3) at existing camera sites so that there is no need to worry about the cost of camera installation, power supply, and data communications. To further reduce deployment cost and increase the feasibility of the proposed solution, the system can be installed during scheduled maintenance trips of existing cameras. Additionally, the YOLOv3 model will be updated (i.e., retrained) as we accumulate more training data, and new features may be added such as detecting stalled vehicles in the emergency lane. Finally, the system design will allow the updated YOLOv3 model to be remotely downloaded to the chosen Jetson edge devices via the existing communication network. In other words, no field trips will be needed to upgrade the system when a new YOLO model is available.

## 3. ConOps for Solution

### *3.1 High-Level Functional Architecture*

The architecture functional objects from the National ITS Reference (USDOT, 2007; 2021) are used to analyze the functional elements relevant to this study and their interactions. Table 3.1 shows the physical objects applicable to this project.

**Table 3.1 Functional Architecture for the Proposed Solution**

| Physical Object | Architecture Functional Objects (National ITS Reference) | Description |
|---|---|---|
| **Traffic Management Center (TMC)** | TMC Basic Surveillance | *'TMC Basic Surveillance'* remotely monitors and controls traffic sensor systems and surveillance (e.g., CCTV) equipment, and collects, processes, and stores the collected traffic data. Current traffic information and other real-time transportation information is also collected from other centers. The collected information is provided to traffic operations personnel and made available to other centers. |
| | TM01: Infrastructure-based Traffic Surveillance | *'TMC Passive Surveillance'* collects time stamped vehicle identities from different detection zones, correlates the identities, and calculates link travel times and derives other traffic measures. |
| | PS11: Early Warning System | The bidirectional interface with this terminator allows potential threat information that is collected by ITS systems to be provided to the alerting and advisory systems to improve their ability to identify threats and provide useful and timely information. |
| | TM08: Traffic Incident Management System | *'TMC Incident Detection'* identifies and reports incidents to Traffic Operations Personnel. It remotely monitors and controls traffic sensor and surveillance systems that support incident detection and verification. |
| | TM09: Integrated Decision Support and Demand Management | It collects traffic data from sensors and surveillance equipment and uses this information to measure traffic network performance. The planned control strategies can be passed back to the transportation information center so that the intended strategies can be reflected in future route planning. |
| **ITS Roadway Equipment** | DM01: ITS Data Warehouse | *'Roadway Data Collection'* collects traffic, road, and environmental conditions information for use in transportation planning, research, and other off-line applications where data quality and completeness take precedence over real-time performance. It includes the sensors, supporting roadside infrastructure, and communications equipment that collects and transfers information to a center for archival. |

## 3.2 *High-Level Physical Architecture*

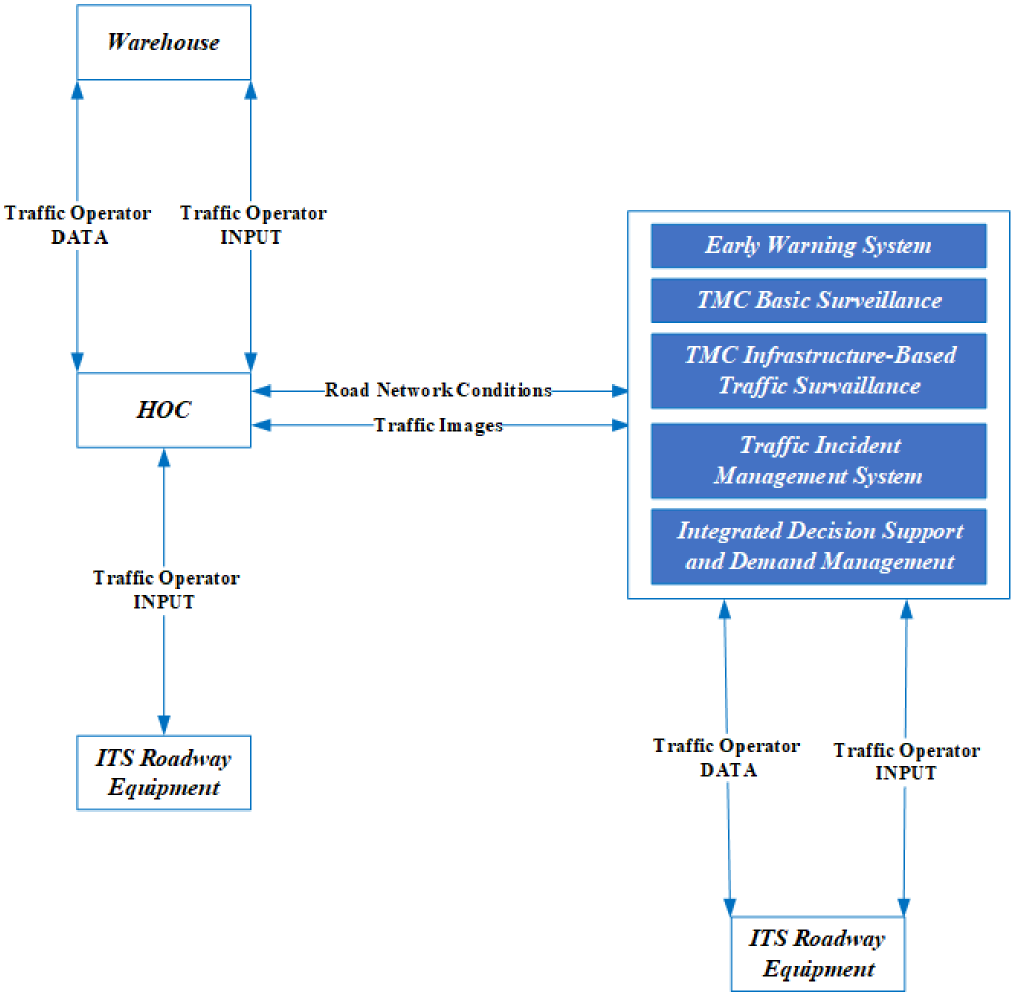


**Figure 3.1 Physical Architecture of the Proposed Solution**

The High-Level Physical Architecture provided in Figure 3.1 breaks the solution system down to show how these physical objects interact with one another to achieve the desired outcome.

## 3.3 *High-Level Enterprise Architecture*

A high-level enterprise architecture shown in Figure 3.2 below provides an overview of the proposed system's people, processes, data, applications, technologies, performance, and relationships among them.

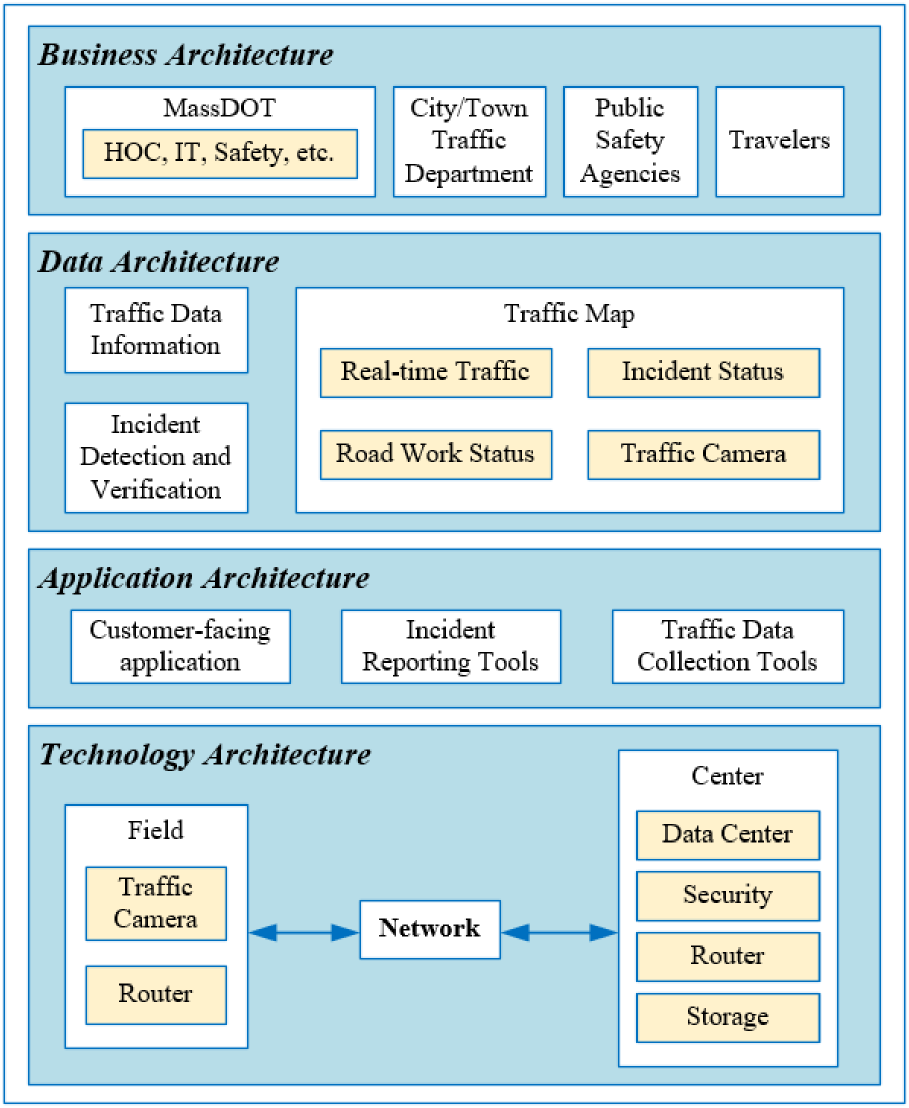


**Figure 3.2 Enterprise Architecture of the Proposed Solution**

# 4. Implementation Feasibility and Schedule

**Table 4.1 Cost Estimates for Upgrading One Camera**

| Item | Unit Price |
|---|---|
| YOLOv3 model and future upgrades | $0 |
| Jetson device (see links for data sources) | $59 for Jetson Nano and $479 for Jetson Xavier NX |
| 3D printed case for the Jetson device | $20 |
| Installation (assuming installed during a camera maintenance trip) | $100 |
| Data communication costs (assuming these cameras are already connected to HOC) | $0 (We will use the existing communication infrastructure owned by MassDOT) |
| Operation and maintenance costs | $200 for Nano and $400 for Xavier NX (both labor and parts) over a **10-year** period |
| **Total** | $379 for Nano and $999 for Xavier NX |

Table 4.1 provides some cost estimates for upgrading one camera using the proposed solution. Since the team will provide the developed YOLOv3 model and future upgrades to MassDOT for free, the cost for algorithm or software is $0. The proposed solution does not require any new hardware or software at the MassDOT HOC. It has excellent scalability and adding it to another camera will not affect any already equipped cameras (e.g., slowing them down), as all the computation will be performed on the edge (not at the HOC). Jetson Nano is a more economical solution. However, choosing Xavier NX in the long term might be a better option, especially if MassDOT has a plan to upgrade existing cameras to high-definition ones, which will require more powerful edge computing devices.

Table 4.2 is a suggested timeline for upgrading existing MassDOT cameras with the YOLOv3 solution. The first month will be used to engage stakeholders (Task 1). Six months after that are allocated for trial operations and evaluation to cover different operating conditions such as snow and extreme temperatures (Task 2). The system will be used for two main applications: data collection and incident detection. Its performance will be evaluated using metrics such as accuracy (e.g., traffic count, speed, time headway), incident detection rate, false alarm rate, and mean time to detect. If both the hardware and software systems work correctly during the trial operations, nine months will be allocated for MassDOT HOC and Information Technology (IT) divisions to modify their existing traffic monitoring platform to take advantage of the new data (Task 3). In the meantime, MassDOT can start to upgrade the remaining cameras during their regular camera maintenance trips (Task 4). This task may take a long time, depending on MassDOT's camera maintenance schedule. Upon the completion of Task 3, two months will be allocated to train traffic operators and other stakeholders and show them how to use the system (Task 5).

**Table 4.2 Timeline for Implementing the Proposed Solution**

| Task | Month | | | | |
|---|---|---|---|---|---|
| | 1 | 2-7 | 8-16 | 17-18 | 19 - Future |
| 1. Stakeholder meetings | ■ | | | | |
| 2. Trial operations and evaluation | | ■ | | | |
| 3. Integration with HOC system | | | ■ | | |
| 4. Large scale deployment | | | ■ | ■ | ■ |
| 5. Employee training | | | | ■ | |

## 5. Anticipated Impacts

The proposed solution aligns well with the MassDOT requirements described in Section 1.2. It also offers many benefits as discussed below.

### *5.1 Operational Benefits*

Upgrading existing traffic cameras into smart sensors provides a cost-effective solution to expand state DOTs' data collection capability. The generated data can be used for automatic incident detection, and human efforts are only needed in verifying incidents. Compared to manually reviewing traffic videos, this solution can significantly reduce the incident detection time and lead to a speedy recovery after an incident occurs. In addition, the collected data can be used for many applications, such as Traveler Information Systems, to improve traffic system performance and travel time reliability, which are critical aspects of TSMO.

### *5.2 Safety Benefits*

Besides traffic incidents, turning traffic cameras into smart sensors will enable the timely detection and verification of other unsafe traffic conditions such as stop-and-go traffic. Early detection of traffic incidents can save precious time for quickly deploying first responders to incident sites and reducing the risk of secondary incidents (Karlaftis et al., 1999). In addition, based on the spatiotemporal distributions of stop-and-go traffic and speeding, police officers can develop risk-targeted patrol plans to improve highway safety.

### *5.3 Mobility Benefits*

It is estimated that over 50% of freeway congestion is caused by traffic incidents (Wang et al., 2008). Non-recurring congestion (e.g., due to traffic incidents) also significantly increases travel time uncertainty. With the proposed solution, response teams can be deployed more quickly to clear the incidents, restore traffic movement, and improve travel time reliability. Additionally, the risk of secondary accidents, which also significantly contribute to congestion and delay, can be reduced with the timely deployment of response teams.

### *5.4 Environmental Benefits*

Non-recurring congestion is also a major contributor to carbon emissions, fuel consumption, and even noise pollution (Qiao et al., 2007). The proposed solution can bring environmental benefits by reducing incident detection time. Additionally, little to no construction is needed for the proposed solution to be implemented. Therefore, the harmful environmental effects of construction and the congestion it may bring about will be mitigated.

### *5.5 Economic Benefits*

The proposed solution will be implemented based on an existing traffic camera network. This will save the costs of purchasing and installing new traffic cameras and significantly enhance the capability of existing ITS equipment with little investment. Also, the proposed solution will process information continuously and independently, requiring minimal human intervention. As a result, labor can be minimized or directed to other tasks. More importantly, congestion costs the U.S. $166 Billion a year (TTI, 2019). Therefore, by improving incident detection and reducing response time, the economic benefits of this solution should never be underestimated.

## 6. Conclusions

The proposed solution applies advanced Artificial Intelligence (AI) and low-cost edge computing technologies to existing traffic cameras and turns them into smart sensors. It will enhance old ITS infrastructure to maximize its reach and expand the data collection and incident detection capacity for the client (i.e., MassDOT). The ConOps, cost, feasibility, and impacts analyses suggest that the solution has great potential to be successful. Also, its system design is very flexible and upgraded detection algorithms can be easily deployed to edge devices without field trips.

## 7. Acknowledgments

The team would like to sincerely thank Neil E. Boudreau, Corey O’Connor, Carrie McInerney, and Gregory M. Khirallah from MassDOT for their kind support and valuable comments. Note that this proposed solution is only for MassDOT to consider, and MassDOT is under no obligation to implement anything described in this report.

## 8. References

USDOT. Systems Engineering for Intelligent Transportation Systems – An Introduction for Transportation Professionals, 2007. Available online at https://ops.fhwa.dot.gov/publications/seitsguide/, Accessed on June 3, 2021.

USDOT. *ARC-IT Version 9.0. The National ITS Reference Architecture - Architecture Reference for Cooperative and Intelligent Transportation*, May 2021. Available online at https://local.iteris.com/arc-it/index.html, Accessed on May 3, 2021.

Karlaftis, M.G., Latoski, S.P., Richards, N.J. and Sinha, K.C., 1999. ITS impacts on safety and traffic management: an investigation of secondary crash causes. *Journal of Intelligent Transportation Systems*, *5*(1), pp.39-52.

Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P. and Zitnick, C.L., 2014, September. Microsoft coco: Common objects in context. In *European conference on computer vision* (pp. 740-755). Springer, Cham.

NIVIDIA, *Jetson Nano: Deep Learning Inference Benchmarks*, Available online at https://developer.nvidia.com/embedded/jetson-nano-dl-inference-benchmarks, Accessed on 5/30/2021.

Qiao, F., Yu, L. and Li, L., 2007. *Estimating Impact of Nonrecurring Congestion on Vehicle Emissions* (No. 07-0427). In Proceedings of the 86th Transportation Research Board Annual Meeting, Washington, D.C., on January 21-25, **2007**

Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

Snyder, C. and Do, M., 2019. *STREETS: A Novel Camera Network Dataset for Traffic Flow* (No. SAND2019-11477C). Sandia National Lab. (SNL-NM), Albuquerque, NM (United States).

Stein, E., Liu, S. and Sun, J., 2019. *Real-Time Object Detection on an Edge Device* (CS230 Final Report). Department of Electrical Engineering, Stanford University, CA. Available online at http://cs230.stanford.edu/projects_fall_2019/reports/26261995.pdf, Accessed on 5/30/2021.

Tan, I., 2020. *Measuring Labelling Quality with IOU and F1 Score*. Available online at https://medium.com/supahands-techblog/measuring-labelling-quality-with-iou-and-f1-score-1717e29e492f, Accessed on 5/30/2021.

Texas A&M Transportation Institute (TTI), 2019. *Nationwide Gridlock Costs $166 Billion Per Year*. Available online at https://today.tamu.edu/2019/08/22/tti-report-nationwide-gridlock-costs-166-billion-per-year/, Accessed on 5/30/2021.

Vasudevan, M., Townsend, H., Schweikert, E., Wunderlich, K.E., Burnier, C., Hammit, B.E., Gettman, D. and Ozbay, K., 2020. *Identifying Real-World Transportation Applications Using Artificial Intelligence (AI)-Real-World AI Scenarios in Transportation for Possible Deployment* (No. FHWA-JPO-20-810). United States. Department of Transportation. Intelligent Transportation Systems Joint Program Office.

Wang, Y., Hallenbeck, M.E., Cheevarunothai, P. and Northwest, T., 2008. *Quantifying incident-induced travel delays on freeways using traffic sensor data* (No. WA-RD 700.1). Seattle: Transportation Northwest, University of Washington.

Xu, Z.F., Jia, R.S., Sun, H.M., Liu, Q.M. and Cui, Z., 2020. Light-YOLOv3: fast method for detecting green mangoes in complex scenes using picking robots. *Applied Intelligence, 50*(12), pp.4670-4687.