Freight Volume Modeling on Major Highway Links

August 2021

A Research Report from the Pacific Southwest Region University Transportation Center

Chrysovalantis Anastasiou, Ph.D. Student, University of Southern California

Seon Ho Kim, Associate Director, IMSC, University of Southern California

Genevieve Giuliano, Co-Principal Investigator, University of Southern California

Luciano Nocera, Co-Principal Investigator, University of Southern California

Cyrus Shahabi, Principal Investigator, University of Southern California

Pacific Southwest Region UTC

USC University of Southern California

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. PSR-20-02	**2. Government Accession No.** N/A	**3. Recipient's Catalog No.** N/A
4. Title and Subtitle Freight Volume Modeling on Major Highway Links		**5. Report Date** 08/02/2021 **6. Performing Organization Code** N/A
7. Author(s) Anastasiou: https://orcid.org/0000-0003-4966-5129 Kim: https://orcid.org/0000-0002-8410-0839 Giuliano: https://orcid.org/0000-0002-9257-8269 Nocera: https://orcid.org/0000-0002-8100-4286 Shahabi: https://orcid.org/0000-0001-9118-0681		**8. Performing Organization Report No.** TBD
9. Performing Organization Name and Address METRANS Transportation Center University of Southern California University Park Campus, RGL 216 Los Angeles, CA 90089-0626		**10. Work Unit No.** N/A **11. Contract or Grant No.** USDOT Grant 69A3551747109
12. Sponsoring Agency Name and Address U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology 1200 New Jersey Avenue, SE, Washington, DC 20590		**13. Type of Report and Period Covered** Final report (08/16/2020 - 08/16/2021) **14. Sponsoring Agency Code** USDOT OST-R
15. Supplementary Notes https://imsc.usc.edu:8103 (Freight Dashboard)		

16. Abstract

With this research we would like to validate the feasibility of freight volume estimation on major highways links from accurate but sparse sensor data, by studying a restricted control area covering approximately 12 square miles around the Ports of Los Angeles and Long Beach where freight volume is most relevant. We researched creating a real-world dataset of real-world data in this region, however, due to the COVID 19 pandemic we had to limit the collection to publicly available Caltrans CCTV video footage and leveraged our ADMS traffic database to generate synthetic data. We have implemented state-of-the-art computer vision algorithms, which were used to classify trucks and define truck categories optimized for best performance. These results show that it is feasible to use CCTV cameras to detect and classify trucks and that the process can be fully automated. In parallel, we created a truck simulator to generate realistic truck trajectories between predefined locations, and developed algorithms to estimate freight on links with different heuristics. Our approach provides the best results by relating compatible observations across sensors using travel-time information estimated on current traffic conditions. Our preliminary results show that freight volume estimation on major highways links is feasible.

17. Key Words Truck Traffic estimation, OD-Matrix, image truck classification		**18. Distribution Statement** No restrictions.	
19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 24	**22. Price** N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Contents

About the Pacific Southwest Region University Transportation Center

The Pacific Southwest Region University Transportation Center (UTC) is the Region 9 University Transportation Center funded under the US Department of Transportation's University Transportation Centers Program. Established in 2016, the Pacific Southwest Region UTC (PSR) is led by the University of Southern California and includes seven partners: Long Beach State University; University of California, Davis; University of California, Irvine; University of California, Los Angeles; University of Hawaii; Northern Arizona University; Pima Community College.

The Pacific Southwest Region UTC conducts an integrated, multidisciplinary program of research, education and technology transfer aimed at *improving the mobility of people and goods throughout the region.* Our program is organized around four themes: 1) technology to address transportation problems and improve mobility; 2) improving mobility for vulnerable populations; 3) Improving resilience and protecting the environment; and 4) managing mobility in high growth areas.

U.S. Department of Transportation (USDOT) Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Disclosure

Shahabi (PI), Giuliano (co-PI), Nocera (co-PI), and Anastasiou (PhD student), conducted this research titled, "Freight Volume Modeling on Major Highway Links" at METRANS and Integrated Media Systems Center, USC. The research took place from August 16, 2020, to August 16, 2021 and was funded by a grant from METRANS in the amount of $99,980. The research was conducted as part of the Pacific Southwest Region University Transportation Center research program.

Acknowledgements

We want to thank Mustafa Alturki, Yingzhe Liu , Luke Nelson and Yiqing Fan for their contributions to the project.

Abstract

With this research we would like to validate the feasibility of freight volume estimation on major highways links from accurate but sparse sensor data, by studying a restricted control area covering approximately 12 square miles around the Ports of Los Angeles and Long Beach where freight volume is most relevant. We researched creating a real-world dataset of real-world data in this region, however, due to the COVID 19 pandemic we had to limit the collection to publicly available Caltrans CCTV video footage and leveraged our ADMS traffic database to generate synthetic data. We have implemented state-of-the-art computer vision algorithms, which were used to classify trucks and define truck categories optimized for best performance. These results show that it is feasible to use CCTV cameras to detect and classify trucks and that the process can be fully automated. In parallel, we created a truck simulator to generate realistic truck trajectories between predefined locations, and developed algorithms to estimate freight on links with different heuristics. Our approach provides the best results by relating compatible observations across sensors using travel-time information estimated on current traffic conditions. Our preliminary results show that freight volume estimation on major highways links is feasible.

Research Report

Executive Summary

Freight traffic in urban areas has a huge impact on traffic, health, and the infrastructure. The COVID-19 pandemic has significantly affected shopping behavior, which in turn has impacted freight traffic, specifically with online shopping and the increased demand for home delivery of goods. This trend is building on the continuing growth in e-commerce sales: according to the United States Quarterly E-Commerce Report, e-commerce sales in the United States have increased at double-digit rates for the past two decades. Therefore, monitoring freight traffic in metropolitan areas is becoming increasingly important for state and local policy making, to alleviate congestion, limit emissions (e.g., Port of Los Angeles Clean Truck Program) and plan to maintain roadways. One of the most challenging problems in urban transportation planning is the lack of fine grain data on freight movements. Cities and regions do not know how many trucks operate in the region and have only limited information on freight flows. A particularly important information problem is the absence of a consistent and current source for freight volume and origin-destination data. Without such information, it is difficult to manage or plan for freight in metropolitan areas. This research is aimed at developing a method for generating freight (truck) volume and origin-destination estimations at the traffic analysis zone level from streamed data so that estimations can be constantly updated. Specifically, we seek to learn if it is possible to automate the estimates of time-dependent flow of trucks and the OD matrix (number of trucks traveling from each origin to each destination) and find out what is the most accurate way to do so. The lack of fine grain truck flow data limits the research that can be performed and leads to the proliferation of one-off studies. Examples of existing data in Southern California include special surveys to generate information on port related truck trips, or to document truck traffic associated with warehouse clusters, RFID sensors at the ports of LA and LB, WIMS stations, UCI TAMS [5, 6] sensors using existing inductive loop sensors. Currently, precise freight information can only be achieved in the vicinity of the sources of information over the highway network where it is most useful, and therefore will remain a restricted source of information only locally actionable. One strategy for solving the freight data problem is to obtain more data, which we propose to do by applying modern image learning algorithms on currently deployed CCTV cameras to detect and classify truck class of interest. To extend freight volume estimates to an entire highway system away from where the truck sensors are located, we propose to develop algorithms that take into consideration the traffic conditions to estimate travel time and reconcile observations for compatible truck types. We developed and tested our freight volume estimation models with synthetic data from a truck simulation software we have built, that we hope can also be useful to Caltrans to utilize our research results and as a planning tool. The outcomes of this research can provide a valuable tool for transportation planning, a low cost, always current best estimate of freight flows on the highway and arterial system which can reveal spatial patterns that can help gain valuable insights into the impact of freight in the entire Los Angeles Metropolitan Area.

Introduction

The economic and social cost induced by freight traffic remains a major concern of the policymakers and all citizens in Los Angeles with the port of Los Angeles ranking number one container port in the United States each year since 2000, and air cargo from LAX, the fourth-busiest airport in the world and second busiest in the United States, and 10th in the world in air cargo tonnage processed, with more than 2.4 million tons of air cargo in 2018[1]. While vehicle traffic has been largely reduced due to the pandemic in 2020, Los Angeles is still ranked as the most congested city in the U.S. according to TomTom Traffic Index Ranking[2]. With the COVID-19 pandemic we have also witnessed an increase in e-shopping triggering a corresponding increase in the transport of goods. One potential tool to help understand and mitigate the effects of freight is a data-driven freight volume tracking system. Because truck information is sporadic and limited, accurate freight volume models require (i) more data sources and (ii) extending the truck observations at the sensors to the entire road network.

Today, we use surveys and region demographics to estimate OD Matrix leading to raw spatial and temporal freight information. [13] have addressed the lack of resolution, by proposing to estimate a 24/7 OD Matrix, casting the problem as a minimization that accounts for traffic volume in addition to survey and region demographics information. Still in this work truck volume is inferred rather than observed and there is no differentiation between types of trucks as the traffic information provided from traditional induction loop sensors is not able to differentiate between vehicle types. Instead in this research we want to analyze truck data obtained directly from sensors to estimate the flow and OD matrix at a finer temporal and spatial resolution. For this we leverage existing truck sensors such as TAMS, WIMS and RFID that can provide very accurate information and research how CCTV sensors used for monitoring can be leveraged to augment the existing sources of information.

To validate the feasibility and automation of freight volume estimation on links from accurate but sparse data sources (e.g., WIMS, CCTV), we focus on a restricted control area in the Los Angeles metropolitan area where freight volume is most relevant. We considered an area covering approximately 12 square miles around the Ports of Los Angeles and Long Beach for which Caltrans and UCI can provide sensor data. Due to the COVID pandemic collecting a data has proven challenging, therefore we modified our research strategy to be able to proceed with limited data by (1) generating synthetic datasets on the area of interest that can be used to train and verify different freight volume estimation models, and (2) leverage existing Caltrans CCTV video footage from public Caltrans feeds to investigate cameras can be used to provide truck observations (i.e., timestamp, truck type and truck type confidence). One advantage of this approach is that we will be in a better position to understand what data to collect and how to use the data to validate our work with future studies.

Specifically, the project includes the following five tasks:

---

[1] https://www.lawa.org/News%20Releases/2019/News%20Release%2018

[2] https://www.tomtom.com/en_gb/traffic-index/ranking/?country=US

- Task 1: Truck classification from CCTV video
- Task 2: Traffic-aware freight simulation
- Task 3: Freight modeling from sensor observations
- Task 4: Freight volume visualization
- Task 5: Final report

Truck classification from CCTV video

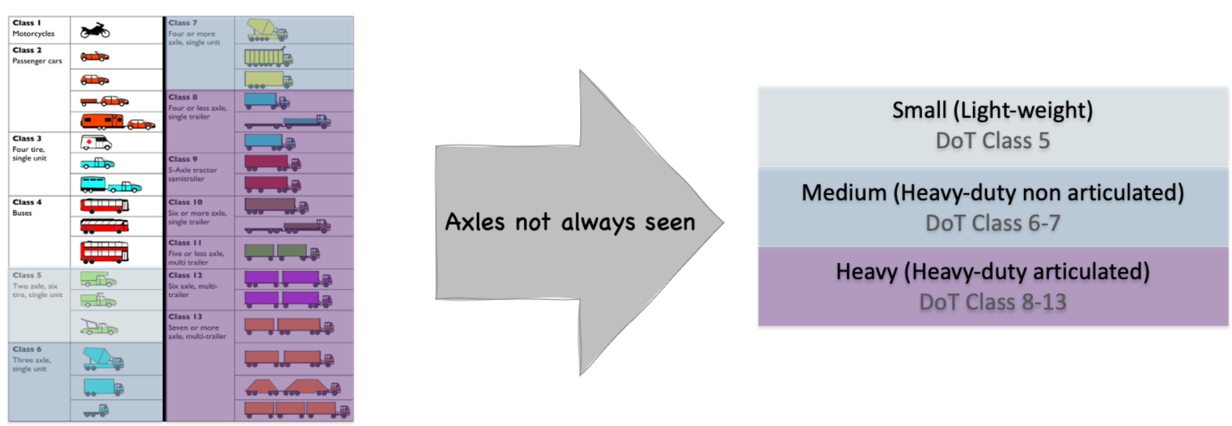
The problem we are trying to solve within this task is to detect and classify trucks given an image that is manually extracted from the CCTV video footage. We are working off still images as this is a prospective study to investigate the feasibility of using these sensors to provide truck observations for our freight volume modeling algorithms, with the idea that if successful we will extend the methods to using videos and select for each observation the best view to use for the classification task.

Method

We initially started on this task by defining a simplified set of initial truck classes of interest to Caltrans that we modeled off the US DoT vehicle types as shown in Figure 1. These classes were defined on trucks and if articulated or not, while DoT vehicle classes are based on the number of axles, which is not usually visible in Caltrans videos.

We based our initial choice of detection and classification on state-of-the-art YOLOv3/Darknet-53 algorithms. While some work in classifying pickup trucks and vans exists, it does not support the truck classes we would like to cover. Also, in previous work YOLOv3/Darknet-53 are trained with publicly available databases that contain footage with different angles than the typical Caltrans videos and that do not label trucks with the needed number of classes. In fact, our truck classification problem is different from existing studies found in the literature, in that (1) CCTV cameras capture footage in a variety of angles, (2) class of truck usually depends on the number of axles but is not always seen and (3) no available labeled datasets.

Figure 1. Simplified truck classes



US DoT vehicle types*
* https://www.fhwa.dot.gov/policyinformation/tmguide/tmg_2013/vehicle-types.cfm

We therefore have built a new dataset by capturing CCTV footage from Caltrans CCTV Map [7] which we used to create a set of labeled image frames for training and a separate set for testing the algorithms. The initial set of labels was set to three classes, *small* (lightweight, e.g., vans, pickups), medium (heavy-duty non articulated, e.g., delivery vans) and *heavy* (heavy-duty articulated trucks). After testing and examination of the model performance and misclassified trucks we have iteratively improved the performance by fine-tuning the parameters of the algorithms and optimizing the classes, so they are better adapted to visually discriminate the trucks based on the image features.

Experiments and Results

We have iterated through 4 versions of the datasets, updating the class definitions to improve classification performance as summarized in Table 1.

Table 1. Labeled datasets v2-4

	Dataset v2	Dataset v3	Dataset v4
# of images	1253	1253	1253
# of background images	300	300	300
Total # of images	1553	1553	1553
articulated_truck count	1029	1029	
articulated_truck_front count			634
articulated_truck_rear count			395
single_unit_truck count	922	922	
single_unit_truck_front count			501
single_unit_truck_rear count			421
lightweight count	866		
van count		225	
van_front count			128
van_rear count			97
pickup count		641	
pickup_front count			253
pickup_rear count			388

Performance results were obtained by implementing TruckCV, a YOLOv3 model initialized from pre-trained Darknet53 weights, using 10% of the dataset for testing and 5% of the dataset for validation. Figures 2-4 present the performance metrics for datasets v2-4 in terms of: precision, recall, F1-score, average precision (AP), and mean average precision (mAP). With each dataset iteration we have introduced new labels to improve classification performance. For example, we added front and rear labels as trucks appear visually different depending on the highway direction they travel. Appendix A presents for datasets v2-4, image frames annotated with classification results showing how the trucks in the images were labeled.

Figure 2. Dataset v-2 performance results

```
confidence_threshold = 0.25
 - articulated_truck, AP = 86.02%   (TP = 84, FP = 21)
 - single_unit_truck, AP = 84.20%   (TP = 76, FP = 33)
 - lightweight,       AP = 63.00%   (TP = 56, FP = 44)

Precision = 0.69 / Recall = 0.77 / F1-score = 0.73
TP = 216, FP = 98, FN = 65, Average IoU = 55.63 %

confidence_threshold = 0.50
 - articulated_truck, AP = 86.02%   (TP = 75, FP = 11)
 - single_unit_truck, AP = 84.20%   (TP = 74, FP = 18)
 - lightweight,       AP = 63.00%   (TP = 51, FP = 22)

Precision = 0.80 / Recall = 0.71 / F1-score = 0.75
TP = 200, FP = 51, FN = 81, Average IoU = 64.57 %
```

```
mAP@0.50 = 0.777391, or 77.74 %
mAP@0.60 = 0.731565, or 73.16 %
mAP@0.70 = 0.625322, or 62.53 %
mAP@0.80 = 0.357104, or 35.71 %
mAP@0.90 = 0.025788, or  2.58 %
```

Figure 3. Dataset v-3 performance results

```
confidence_threshold = 0.25
 - articulated_truck, AP = 86.13%   (TP = 88, FP = 27)
 - single_unit_truck, AP = 84.87%   (TP = 76, FP = 31)
 - van,               AP = 49.90%   (TP = 13, FP =  8)
 - pickup,            AP = 76.33%   (TP = 46, FP = 30)

Precision = 0.70 / Recall = 0.79 / F1-score = 0.74
TP = 223, FP = 96, FN = 58, average IoU = 56.40 %


confidence_threshold = 0.50
 - articulated_truck, AP = 86.13%   (TP = 83, FP = 18)
 - single_unit_truck, AP = 84.87%   (TP = 72, FP = 22)
 - van,               AP = 49.90%   (TP = 12, FP =  6)
 - pickup,            AP = 76.33%   (TP = 41, FP = 16)

Precision = 0.77 / Recall = 0.74 / F1-score = 0.75
TP = 208, FP = 62, FN = 73, average IoU = 62.45 %
```

```
mAP@0.50 = 0.743076, or 74.31 %
mAP@0.60 = 0.705901, or 70.59 %
mAP@0.70 = 0.614002, or 61.40 %
mAP@0.80 = 0.352933, or 35.29 %
mAP@0.90 = 0.035619, or  3.56 %
```

Figure 4. Dataset v-4 performance results

```
confidence_threshold = 0.25
 - articulated_truck_front, AP = 90.93% (TP = 61, FP = 23)
 - articulated_truck_rear,  AP = 84.24% (TP = 29, FP = 12)
 - single_unit_truck_front, AP = 78.12% (TP = 37, FP = 13)
 - single_unit_truck_rear,  AP = 84.53% (TP = 36, FP = 16)
 - van_front,               AP = 30.44% (TP =  7, FP = 10)
 - van_rear,                AP = 59.75% (TP =  3, FP =  4)
 - pickup_front,            AP = 66.82% (TP = 16, FP = 12)
 - pickup_rear,             AP = 69.27% (TP = 25, FP =  7)

Precision = 0.69 / Recall = 0.76 / F1-score = 0.72
TP = 214, FP = 97, FN = 67, average IoU = 56.09 %

confidence_threshold = 0.50
 - articulated_truck_front, AP = 90.93% (TP = 60, FP = 17)
 - articulated_truck_rear,  AP = 84.24% (TP = 26, FP =  5)
 - single_unit_truck_front, AP = 78.12% (TP = 33, FP = 11)
 - single_unit_truck_rear,  AP = 84.53% (TP = 34, FP = 11)
 - van_front,               AP = 30.44% (TP =  6, FP =  6)
 - van_rear,                AP = 59.75% (TP =  3, FP =  4)
 - pickup_front,            AP = 66.82% (TP = 14, FP =  5)
 - pickup_rear,             AP = 69.27% (TP = 23, FP =  6)

Precision = 0.75 / Recall = 0.71 / F1-score = 0.73
TP = 199, FP = 65, FN = 82, average IoU = 61.76 %
```

```
mAP@0.50 = 0.705120, or 70.51 %
mAP@0.60 = 0.675986, or 67.60 %
mAP@0.70 = 0.571626, or 57.16 %
mAP@0.80 = 0.306439, or 30.64 %
mAP@0.90 = 0.032269, or  3.23 %
```

Conclusion

As seen in the performance results, the addition of the front and rear sub-classed generally improve performance (e.g., *articulated_truck_front* class of dataset v3, achieves 90.9% average precision, while *articulated_truck* of dataset v1-2 only achieves 86.1%). We are in the process of further revising the dataset to add images/examples and revise labels. Adding labeled images is needed for some classes to improve performance as we have very few examples (e.g., vans and pickups), and to deal with rare cases where trucks and vehicles are recognized as single trucks (some examples of which can be seen in the classified images of Appendix A). Also, labels correspond to trucks that are not fully in the image frame or appear very small need not to be considered. With our future work we have started to explore strategies to go from images to video, by implementing continuous detection and classification where we track/deduplicate objects and implementing an adaptive frame selection mechanism to choose the best frame to use to classify a given truck. Finally, we are looking at selectively enable/disable classifier labels based on the direction of travel, either by using the tracker to infer direction or by matching to a known label corresponding to the side of the highway the truck is traveling, e.g., if the truck is moving away from the camera we will use *_rear* labels and if the truck is moving towards the camera we will use *_front* labels to classify. We expect these developments to resolve misclassification issues such as those seen in Appendix A, where the classification fails because the truck is too small in the image and/or is detected together with another object or is detected as rear when traveling towards the camera and vice versa.

Traffic-aware freight simulation

We describe here after the main features of the freight simulation software we have developed. The purpose of the simulator is to generate synthetic data to help us develop and test our truck flow estimation algorithms.

Unlike existing traffic flow simulators that simulate the traffic, our freight simulation simulates the trajectories of trucks using our real-world traffic data that we provide, and sensors locations and truck observations at each sensor including a time, a type of truck and a confidence for that truck type.

Method

We have developed a parameterized truck simulation taking in input real-world historical traffic information to generate truck trajectories assuming departure times and locations of the trucks and destinations and assuming trucks travel according to traffic speeds. Using real-world traffic information will allow us to test the algorithms under different traffic conditions. The simulation can also take as input a set of sensor locations to produce truck observations based on the truck trajectories. Both the truck destinations and the sensor locations can be specified with configuration files or generated using an algorithm that randomly puts them on the map within a highway link while considering the direction of travel. For traffic information we leverage our ADMS database to (see Freight volume visualization for details). The simulator was developed as a command line tool in Java, with configuration files in JSON format (spatial data is encoded in GeoJSON a standard geospatial format for mapping).

Experiments and Results

Using the freight simulator, we have generated a number of datasets, e.g., we generated synthetic datasets that were used to test our freight volume estimation algorithms, by varying the number of sensors with values in the set {100, 150, 200, 250, 300}, and number of trucks with values in the set {250, 500, 750, 1000, 5000} and considering one origin at the Port of Los Angeles (simplified test case). Simulation configuration data and synthetic datasets can be visualized using a freight visualization dashboard we have developed and presented in detail in a later section.

Conclusion

We expect to expand the capabilities of our freight simulator and test our predictive algorithms on a variety of conditions (e.g., different traffic conditions, sensor layout, algorithms parameters) to gain insights into improving volume estimation performance and help Caltrans understand our research and its implications, e.g., as a planning tool to understand how many sensors are needed to achieve a certain level of estimation accuracy or what is the estimation accuracy that can be achieved with existing layout of sensors.

Freight modeling from sensor observations

In this section we present our approach to model freight volume from truck observation provided by sensors on the road network.

Method

The purpose of developing predictive freight models is to examine the feasibility and estimation accuracy that can be achieved. Specifically, we would like to learn if it is possible to accurately estimate the time-dependent flow of trucks in a road network and if it is possible to accurately estimate the OD matrix (number of trucks traveling from each origin to each destination). In addition we would like to learn what is the best/most accurate way to achieve these objectives. Related work includes truck flow estimation works such as the Heavy-Duty Truck (HDT) Model developed as part of SCAG [8, 9], and air quality works from Perugo et al. [10, 11, 12]. Available sensors in the region of interest (12 sq. miles around the Ports of Los Angeles and Long Beach) includes the following sensors.

RFID sensors: very accurate truck classification. Typically available at port exits. Provides information for <location, timestamp, truck_type> + <truck_id>, for specific truck ids.
Weigh-in-motion (WIM) sensors: very accurate truck classification. Sparse but provides checkpoints with information for <location, timestamp, truck type> + <truck id>.
TAMS sensors: very accurate truck classification. Sparse and probabilistic, provides information for <location, timestamp, truck_type, truck_type_probability>.
CCTV cameras: variable accuracy truck classification. Sparse and probabilistic, provides information for <location, timestamp, truck_type, truck_type_probability>.
Inductive Loop Detectors: traffic information only (speed and occupancy).
Double Loop Detectors: e.g., Caltrans vehicle counting system.

We formalize the flow estimation problem by considering a number of sensors, each providing truck observations in a tuple form:

<location, timestamp, truck_type, truck_type_probability>

For example, in the case of an RFID sensor all tuple fields are precisely known with a truck type probability of 1. Given a number of such observations we seek to predict truck flow for all links in a given road network. We have developed three approaches:

Baseline: only estimates at edges where data is sensed.

FlowPath: extrapolates flow based on logic.

Reachability-based: extrapolates flow based on travel time, i.e., computing if observations can be reached taking into account existing traffic conditions.

Experiments and Results

Evaluation is conducted for the three approaches on the synthetic datasets generated using our freight simulator. We report on performance using the following metrics: percentage of graph edges in estimation that exist in ground truth (Precision), percentage of graph edges in ground truth that are in estimation (recall), Mean Absolute Error of flow estimation (MAE), Mean Absolute Percentage Error (MAPE). We have experimented by varying the number of sensors and the number of trucks Figures 6 and 7 shows the maps of estimated link flows obtained for 200 sensors randomly distributed and 1000 trucks, and 300 sensors randomly distributed and 1000 trucks respectively. Both datasets consider one origin located at the Port of Los Angeles and the same set of 30 destinations. Performance statistics are printed below each figure. Our results indicate that the Reachability-based approach provides the most accurate and most complete estimations. While Flow Path and Reachability-based perform equally well in terms of precision of the estimate (97-98% and 94-95% respectively) and the recall rates are greater for the Reachability-based algorithm (25-37% and 63-75% respectively).

Figure 6. Experiments with 200 sensors and 1000 trucks

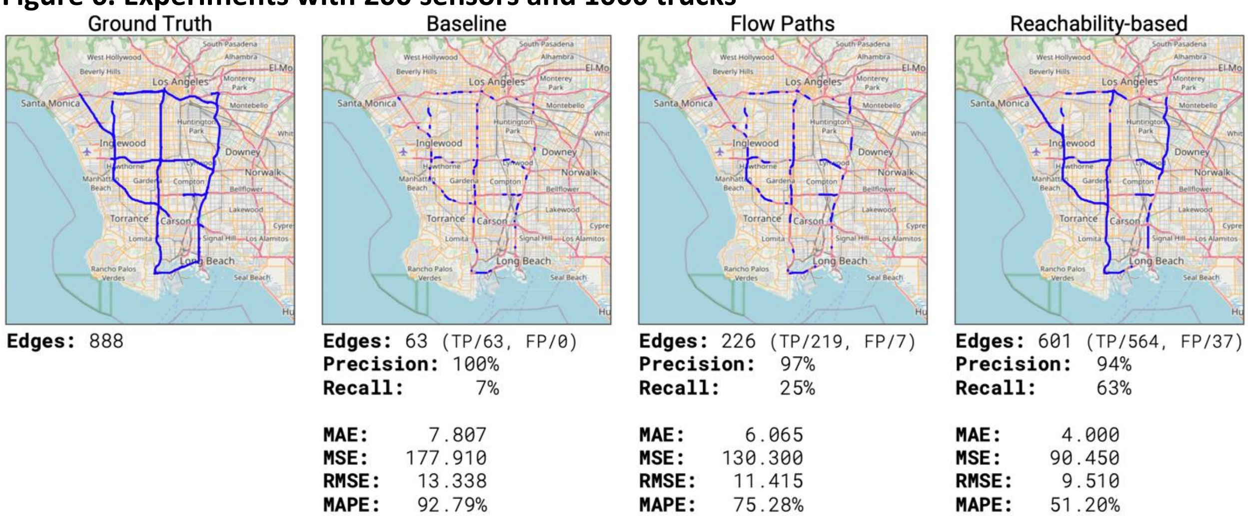


Figure 7. Experiments with 300 sensors and 1000 trucks

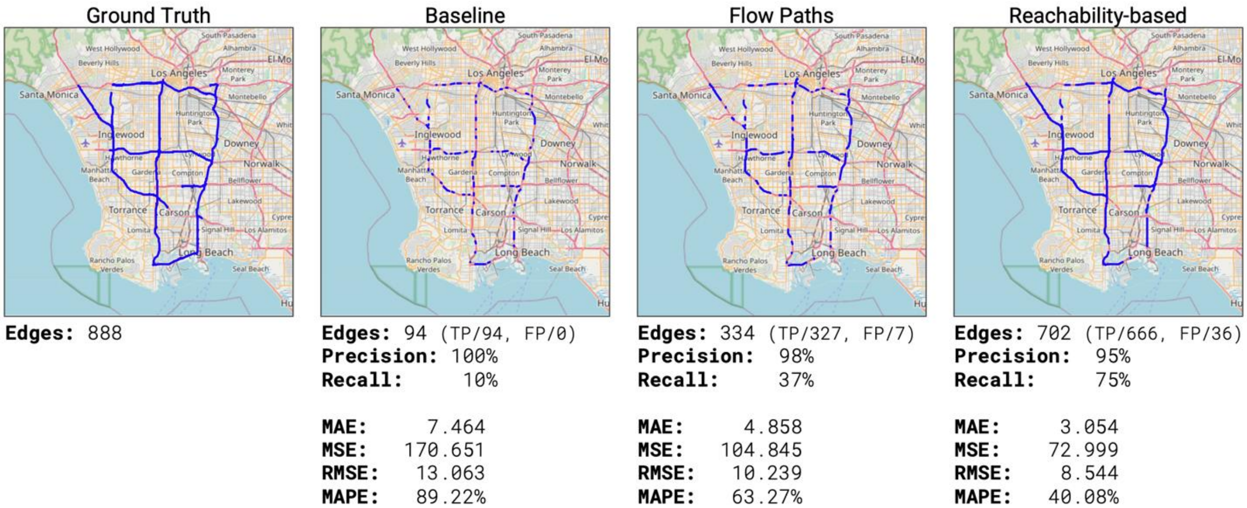
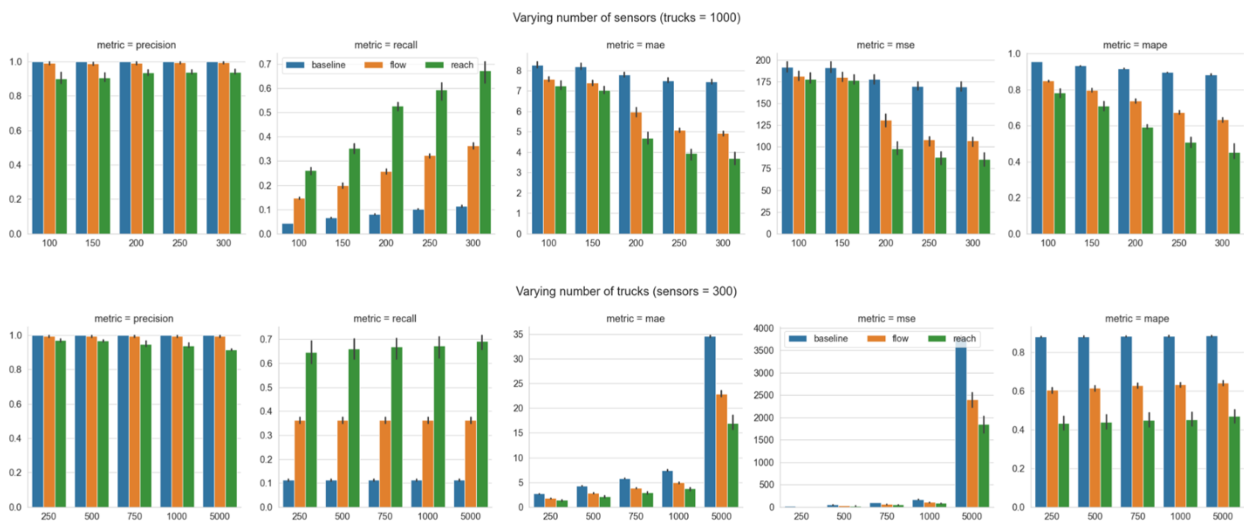


Figure 8 presents bar charts for experiments with varying the number of sensors and trucks respectively. We observe similar patterns where the Reachability-based outperforms the other approaches with the Flow Path as the second-best performer across all metrics.

Figure 8. Experiments with varying number of trucks (top) and sensors (bottom)



Conclusion

We have shown that with these initial experiments that a relatively large number of sensors allows us to automate the estimation of truck flow on highway links. These estimates can produce hourly truck counts on each link of the highway continuously, provided the sensor observations are available. With our future work we will relax the hypothesis made (one truck origin, truck class probability set to 1.0) and explore failure modes. For example, when no sensors are available on the segments leading to two separate destinations, we have adopted a greedy approach to decide which segment a truck needs to travel; this leads to errors. Instead, we will explore how a probabilistic approach can lead to better prediction performance. Another possible solution that we will explore is to place sensors strategically, so they maximize the predictions precision and recall. This last approach can be of interest to Caltrans to optimize sensor placement in an eventual deployment. Finally, with this experimental set-up we can investigate what sensors are most practical and useful to use.

Freight volume visualization

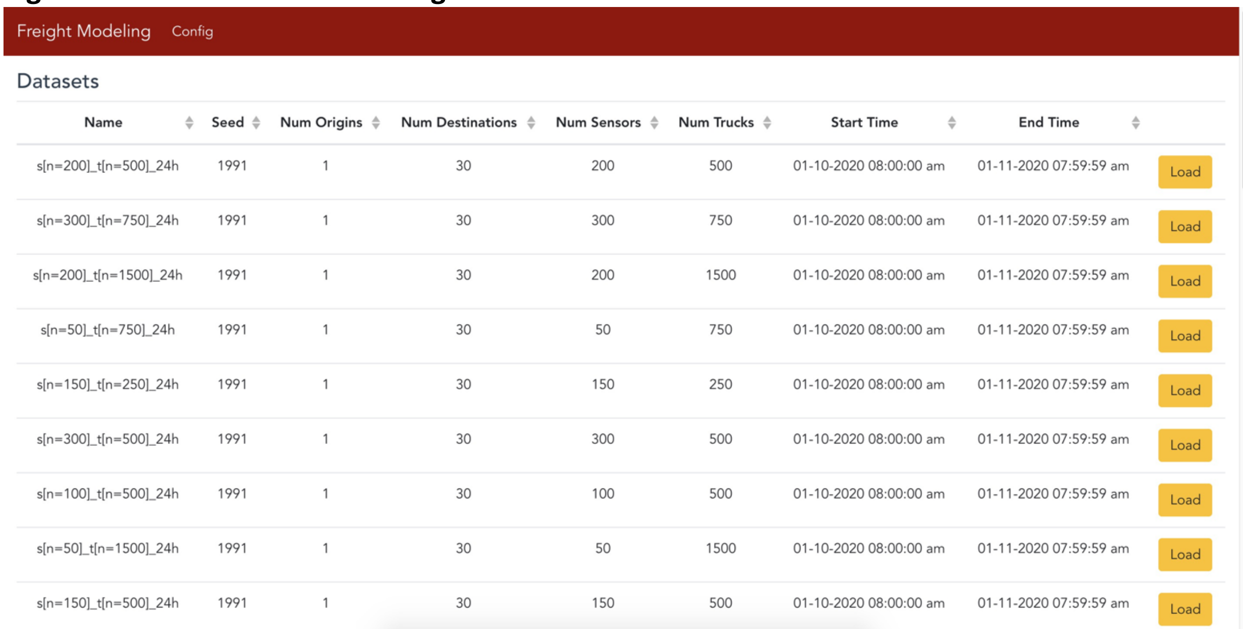
We develop both the front-end dashboard and back-end services to understand the simulated data and results of the algorithms. We hope this tool can be useful to Caltrans to understand the value of our research and as a planning tool.

The tool interfaces with the ADMS Database to obtain the traffic data to visualize.

ADMS Database. With our partnership with the Los Angeles Metropolitan Transportation Authority, we developed a large transportation data warehouse called Archived Traffic Data Management System (ADMS). The ADMS Database stores the bus and traffic flow information. This dataset includes both sensor metadata and real-time data for freeway and arterial traffic sensors (approximately 16,000 loop-detectors) covering 4,300 miles and 2,000 bus and train automatic vehicle locations with an update rate as high as 30 seconds. We have been continuously collecting and archiving the datasets mentioned above for the past 12 years. These datasets are used to train and test our traffic forecasting and Bus ETA models.

Web-based UI. To display the synthetic data and the results of the predictive models we built a comprehensive dashboard. The dashboard includes a screen that displays available simulation parameter files listing their main parameters (i.e., seed used to randomize, number of origins, and destinations, number of sensors, number of trucks, departure times, start and end time) as shown in Figure 9. Once a simulation is loaded the user can examine all the parameters using a map and a table. For example, clicking on truck will display the simulated truck trajectory in the map with the corresponding sensor observations. Figure 10 shows this display screen where the user can control the information displayed on the map using an interactive legend. Finally, the picture of Figure 11 shows a screen dedicated to display and compare volume predictions in a side-by-side fashion, e.g., to allow the user to visually examine the differences between the ground truth (the actual simulated data) and the prediction (produced by freight volume modeling algorithms). We also support displaying model performance metrics (e.g., confusion matrix) and examining discrepancies with the ground truth for individual trucks. Finally, a dedicated view allows comparison of ground truth and predicted OD matrix.

Figure 9. Truck simulations listing screen

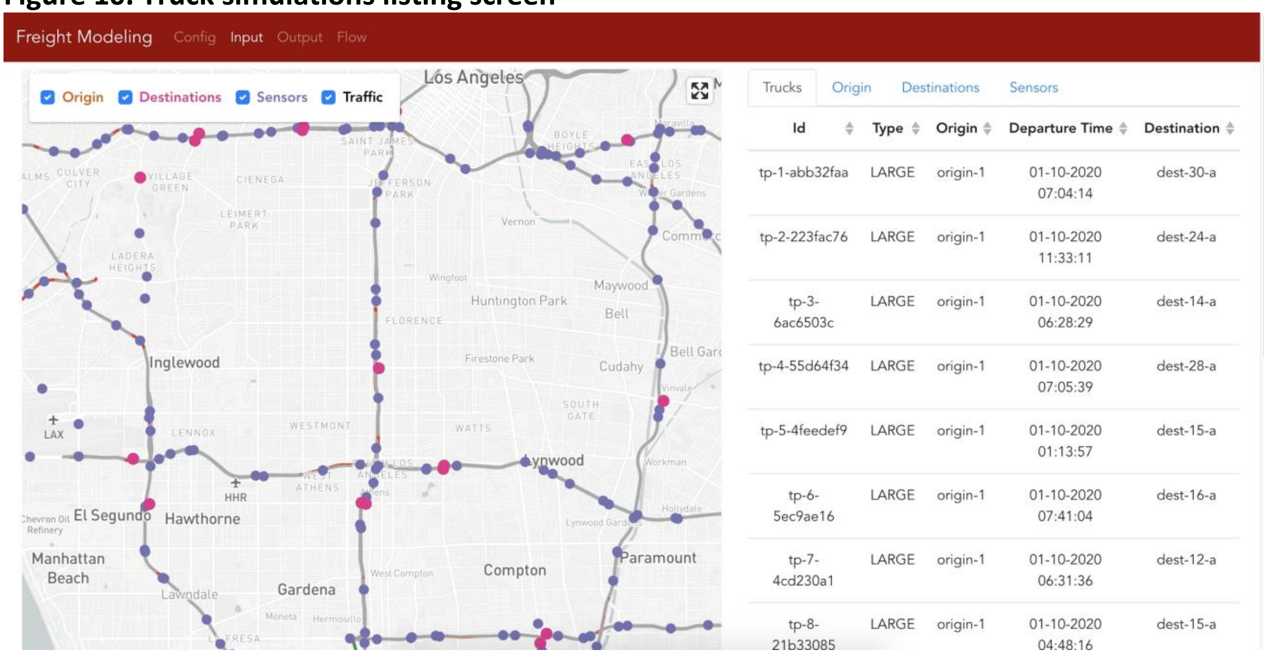


Freight Modeling Config

Datasets

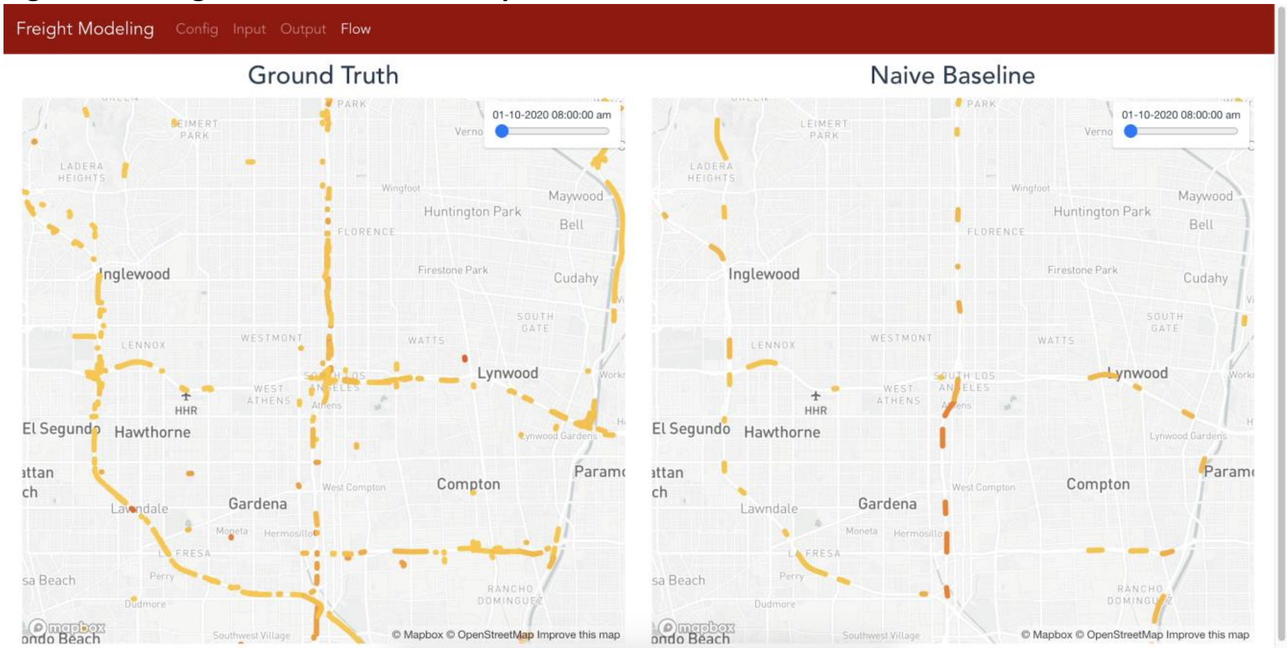
Name	Seed	Num Origins	Num Destinations	Num Sensors	Num Trucks	Start Time	End Time	
s[n=200]_t[n=500]_24h	1991	1	30	200	500	01-10-2020 08:00:00 am	01-11-2020 07:59:59 am	Load
s[n=300]_t[n=750]_24h	1991	1	30	300	750	01-10-2020 08:00:00 am	01-11-2020 07:59:59 am	Load
s[n=200]_t[n=1500]_24h	1991	1	30	200	1500	01-10-2020 08:00:00 am	01-11-2020 07:59:59 am	Load
s[n=50]_t[n=750]_24h	1991	1	30	50	750	01-10-2020 08:00:00 am	01-11-2020 07:59:59 am	Load
s[n=150]_t[n=250]_24h	1991	1	30	150	250	01-10-2020 08:00:00 am	01-11-2020 07:59:59 am	Load
s[n=300]_t[n=500]_24h	1991	1	30	300	500	01-10-2020 08:00:00 am	01-11-2020 07:59:59 am	Load
s[n=100]_t[n=500]_24h	1991	1	30	100	500	01-10-2020 08:00:00 am	01-11-2020 07:59:59 am	Load
s[n=50]_t[n=1500]_24h	1991	1	30	50	1500	01-10-2020 08:00:00 am	01-11-2020 07:59:59 am	Load
s[n=150]_t[n=500]_24h	1991	1	30	150	500	01-10-2020 08:00:00 am	01-11-2020 07:59:59 am	Load

Figure 10. Truck simulations listing screen



Id	Type	Origin	Departure Time	Destination
tp-1-abb32faa	LARGE	origin-1	01-10-2020 07:04:14	dest-30-a
tp-2-223fac76	LARGE	origin-1	01-10-2020 11:33:11	dest-24-a
tp-3-6ac6503c	LARGE	origin-1	01-10-2020 06:28:29	dest-14-a
tp-4-55d64f34	LARGE	origin-1	01-10-2020 07:05:39	dest-28-a
tp-5-4feedef9	LARGE	origin-1	01-10-2020 01:13:57	dest-15-a
tp-6-5ec9ae16	LARGE	origin-1	01-10-2020 07:41:04	dest-16-a
tp-7-4cd230a1	LARGE	origin-1	01-10-2020 06:31:36	dest-12-a
tp-8-21b33085	LARGE	origin-1	01-10-2020 04:48:16	dest-15-a

Figure 11. Freight volume results comparison screen



Conclusion

In summary, we developed an approach for freight volume modeling and for classifying trucks from CCTV cameras. Because of the lack of real-world data, we have developed a freight simulator that creates synthetic datasets under different scenarios of number of trucks and sensors based on historical traffic data. Our freight volume modeling experiments were conducted on a restricted control area covering approximately 12 square miles around the Ports of Los Angeles and Long Beach where freight volume is most relevant. The approach is capable of handling large road networks and can provide the most accurate predictions when compared to baseline approaches. The CCTV truck classification results show that it is possible to repurpose currently deployed Caltrans monitoring cameras for counting and classifying trucks.

References

[1] Oh, C., Tok, A. and Ritchie, S.G., 2005. Real-time freeway level of service using inductive-signature-based vehicle reidentification system. *IEEE Transactions on intelligent transportation systems*, *6*(2), pp.138-146.
[2] Tok, A., Hernandez, S.V. and Ritchie, S.G., 2009, January. Accurate individual vehicle speeds from single inductive loop signatures. In *Proceedings of 88th Annual Meeting of the Transportation Research Board, National Research Council, Washington, DC, USA*.
[3] Lamas-Seco, José J., Paula M. Castro, Adriana Dapena, and Francisco J. Vazquez-Araujo. "Multi-loop inductive sensor model for vehicle traffic applications." *Sensors and Actuators A: Physical* 263 (2017): 580-592.
[4] Marszalek, Z., Zeglen, T., Sroka, R. and Gajda, J., 2018. Inductive loop axle detector based on resistance and reactance vehicle magnetic profiles. *Sensors*, *18*(7), p.2376.
[5] Tok, A. (2016) An overview of the California Statewide Freight Forecasting Model. UC Irvine: Institute of Transportation Studies:
http://www.scag.ca.gov/committees/CommitteeDocLibrary/mtf012716_CSFFM.pdf
[6] Truck Activity Monitoring System online at: http://freight.its.uci.edu/tams
[7] Caltrans CCTV Map online at: http://cwwp2.dot.ca.gov/vm/iframemap.htm
[8] SCAG REGIONAL TRAVEL DEMAND MODEL AND 2012 MODEL VALIDATION available online at: https://www.princeton.edu/~alaink/Orf467F16/SCAG_RTDM_2012ModelValidation.pdf
[9] Yoon, S., & Kim, J. H. 4-Step truck model: SCAG case study available online at: https://silo.tips/download/4-step-truck-model-scag-case-study
[10] Perugu, H., Wei, H., & Yao, Z. (2014). *Modeling truck activity using short-term traffic counts for reliable estimation of heavy-duty truck emissions in urban areas,* Transportation Research Board. (link)
[11] Perugu, H., Wei, H., & Yao, Z. (2016). *Integrated data-driven modeling to estimate PM2. 5 pollution from heavy-duty truck transportation activity over metropolitan area*. *Transportation Research Part D: Transport and Environment*, *46*, 114-127. (link)
[12] Perugu, H., Wei, H., & Yao, Z. (2017). *Developing high-resolution urban scale heavy-duty truck emission inventory using the data-driven truck activity model output*. Atmospheric environment, 155, 210-230. (link)
[13] Ma, Wei, and Zhen Sean Qian. "Estimating multi-year 24/7 origin-destination demand using high-granular multi-source traffic data." *Transportation Research Part C: Emerging Technologies* 96 (2018): 96-121.

Data Management Plan

Products of Research

We have assembled a dataset of labeled truck images of Caltrans online CCTV cameras that can be used to train truck detection and classification algorithms.

We created algorithms to estimate truck flow from sensor observations.

We created a truck simulator that uses real-world traffic data to generate synthetic truck path and sensor observations.

Data Format and Content

We will describe the format for the data that is part of this research as we release the datasets publish the algorithms. The truck simulator will be made available on request.

Data Access and Sharing

We have published the finalized truck image dataset online for researchers to access at https://bit.ly/3IMfbYh. The dataset contains training, validation, and test images. Readme files provides details about the dataset.

Algorithms for truck flow prediction and truck simulation will be publish online in workshops or conference papers.

Reuse and Redistribution

The image dataset will be provided as is under GPL license terms.

Appendix

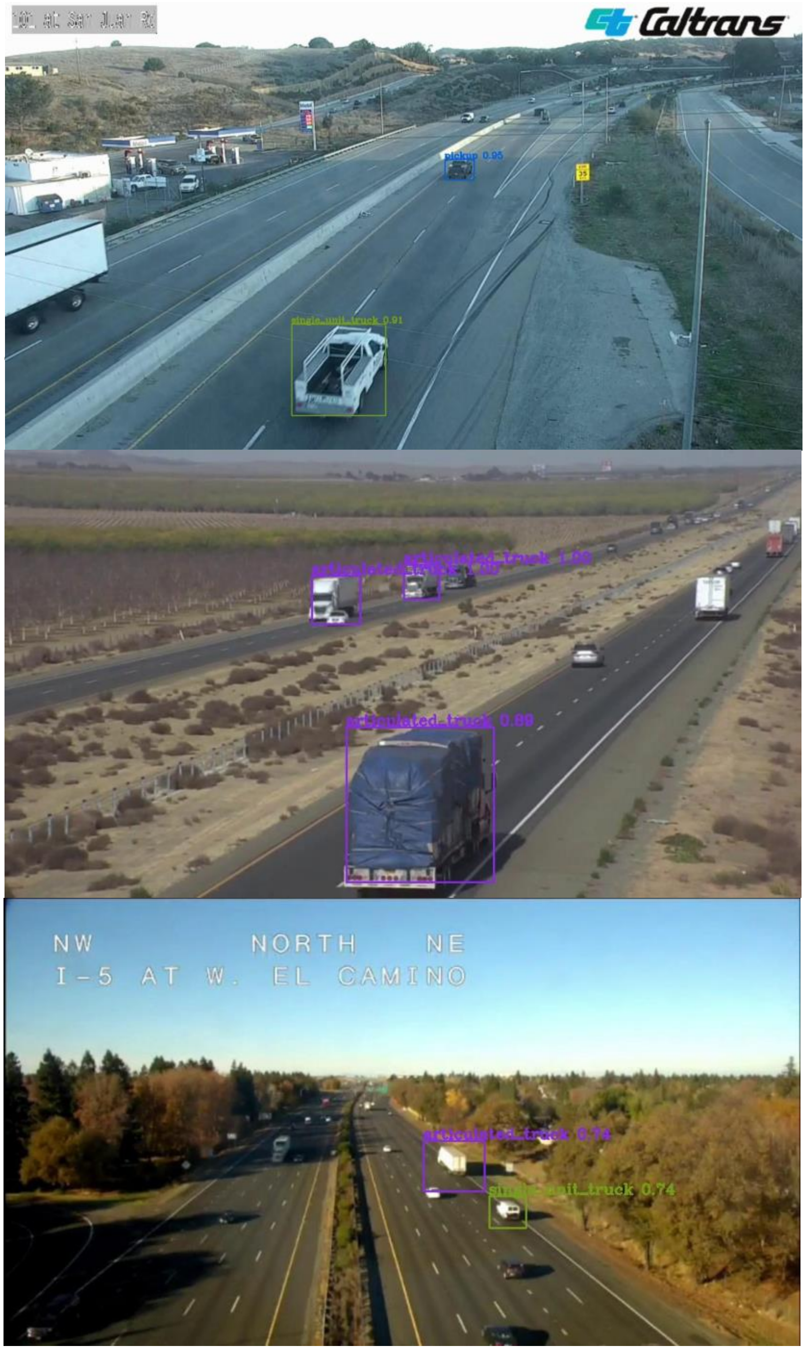
Examples of classification results obtained on single images extracted from CCTV footage for the datasets v2-4.

Dataset v2

Classes annotations: lightweight single_unit_truck articulated_truck

Dataset v3

Classes annotations: pickup single_unit_truck articulated_truck



Dataset v4

Classes annotations: pickup_rear articulated_truck_front articulated_truck_rear