An Interactive Data Visualization and Analytics Tool to Evaluate Mobility and Sociability Trends During COVID-19

Fan Zuo
New York University
New York City, NY, USA
fan.zuo@nyu.edu

Jingxing Wang
University of Washington
Seattle, WA, USA
wangjx@uw.edu

Jingqin Gao
New York University
New York City, NY, USA
jingqin.gao@nyu.edu

Kaan Ozbay
New York University
New York City, NY, USA
Kaan.Ozbay@nyu.edu

Xuegang Jeff Ban
University of Washington
Seattle, WA, USA
banx@uw.edu

Yubin Shen
New York University
New York City, NY, USA
Ys77@nyu.edu

Hong Yang
Old Dominion University
Norfolk, VA, USA
hyang@odu.edu

Shri Iyer
New York University
New York City, NY, USA
shri.iyer@nyu.edu

ABSTRACT

The COVID-19 outbreak has dramatically changed travel behavior in affected cities. The C2SMART research team has been investigating the impact of COVID-19 on mobility and sociability. New York City (NYC) and Seattle, two of the cities most affected by COVID-19 in the U.S. were included in our initial study. An all-in-one dashboard with data mining and cloud computing capabilities was developed for interactive data analytics and visualization to facilitate the understanding of the impact of the outbreak and corresponding policies such as social distancing on transportation systems. This platform is updated regularly and continues to evolve with the addition of new data, impact metrics, and visualizations to assist public and decision-makers to make informed decisions. This paper presents the architecture of the COVID related mobility data dashboard and preliminary mobility and sociability metrics for NYC and Seattle.

CCS CONCEPTS

• **Human-centered computing** → **Visualization**; • **Software and its engineering** → **Software creation and management**.

KEYWORDS

COVID-19, interactive data visualization, human mobility pattern, cloud computing, data mining, social distancing, object detection

ACM Reference Format:
Fan Zuo, Jingxing Wang, Jingqin Gao, Kaan Ozbay, Xuegang Jeff Ban, Yubin Shen, Hong Yang, and Shri Iyer. 2020. An Interactive Data Visualization and Analytics Tool to Evaluate Mobility and Sociability Trends During COVID-19. In *San Diego '20: The 9th SIGKDD International Workshop for Urban Computing, August 24, 2020, San Diego, CA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The novel Coronavirus COVID-19 spreading rapidly throughout the world was recognized by the World Health Organization (WHO) as a pandemic on March 11, 2020 [15]. As of May 17, confirmed coronavirus cases in the U.S. surpassed 1,400,000, and New York has the highest number of confirmed cases (348,232 confirmed cases as of May 15) in the country [13]. The COVID-19 pandemic and ensuing social distancing orders have had dramatic impacts on the use of every mode of transportation. With little information on similar historic pandemics, collection of perishable mobility, safety, and behavior data related to COVID-19 and learning from the collected data becomes the key for decision-makers.

C2SMART researchers have developed an interactive data dashboard (http://c2smart.engineering.nyu.edu/covid-19-dashboard/) that consolidates multiple public data sources listed in Table 1, to investigate changes in mobility and sociability patterns.

The main objective of this paper is to introduce the developed platform based on the analysis of data obtained in NYC and Seattle. Time lag between different cities may offer some insights into how travel patterns evolve in the context of COVID-19.

2 PLATFORM ARCHITECTURE

Figure 1 illustrates the architecture of the tool. The core of the platform resides on cloud computing that provides a reliable infrastructure for solid performance on large scale data processing and analysis. The ingestion instances inside the platform consume the raw data stream and evaluate it based on data accuracy, timeliness, validity and granularity. Once the evaluation is completed, the outcome data will be dispatched to the internal data warehouse for further analysis. The data dashboard is capable of performing real-time data analytics and visualization.

Table 1: Data Collection List

| Name	Region
No. of COVID-19 case	NYC and Seattle
MTA subway/bridges and tunnels volume	NYC
Weight-in-motion stations	NYC
NYCDOR rea-time traffic speed map	NYC
Travel time (INRIX and Virtual sensors)	NYC and NJ
NYPD crash reports	NYC
MTA bus time	NYC
Parking and camera violations	NYC
311 service requests	NYC
Citibike count	NYC
NYCDOT real-time cameras	NYC
Travel time (Google API)	Seattle
Traffic volume (WSDOT detectors)	Seattle
SDOT bike/pedestrian counts	Seattle
Paid parking occupancy	Seattle
Public transit demand	Seattle
SDOT real-time cameras	Seattle
Taxi ridership	Chicago
Subway ridership	Six cities in China
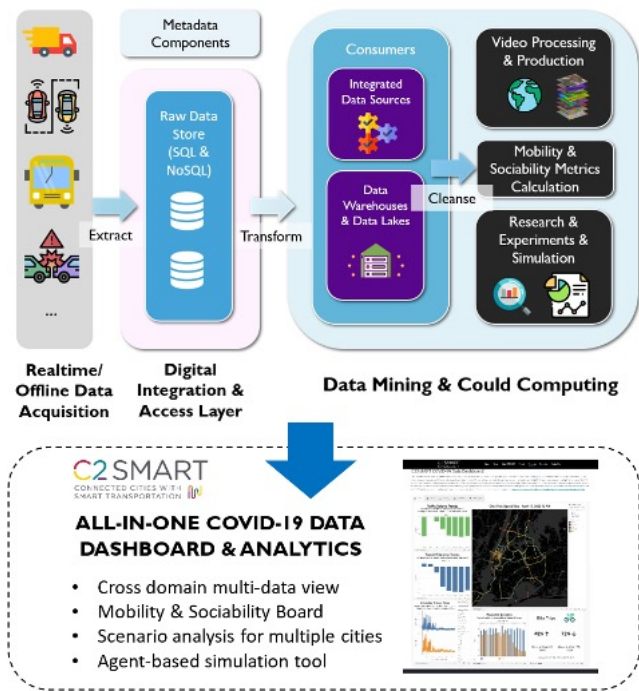
Figure 1: Data Dashboard Architecture

The dashboard contains two main sub-boards: (a) mobility board and (b) sociability board. The mobility board presents multi-data views in terms of vehicular traffic volume, corridor travel time, transit ridership, freight traffic by gross vehicle weight (GVW), as well as risk indicators in terms of reported crashes, pedestrian and cyclist fatalities and speeding tickets. Scenario analysis can be performed directly to these metrics in time series or with spatiotemporal aggregations. The sociability board presents average and maximum intersection crowd density, bike and pedestrian counts related to open streets and pedestrian social distancing compliance rates based on information extracted from real-time traffic cameras. The tool also connects to C2SMART's MATSim agent-based simulation virtual testbed [4] that provides network performance and emission evaluation for reopening phases.

3 MOBILITY BOARD

The section illustrates some results offered by the mobility board.

3.1 New York City

The closing of essential businesses and stay-at-home policies caused an immediate direct impact on transportation. Data shows steep declines in both transit ridership and vehicular traffic. Travel time on corridor 495 that connects Long Island and New Jersey (NJ) via Queens and Manhattan in NYC was analyzed in Figure 2 [12] [8]. A flatter travel time pattern of near free-flow speeds was observed following the stay-at-home orders, instead of typical spikes of commuter peaks. Increased traffic volume and travel time were observed in the week of May 4, indicating the start of the recovery period for NYC.
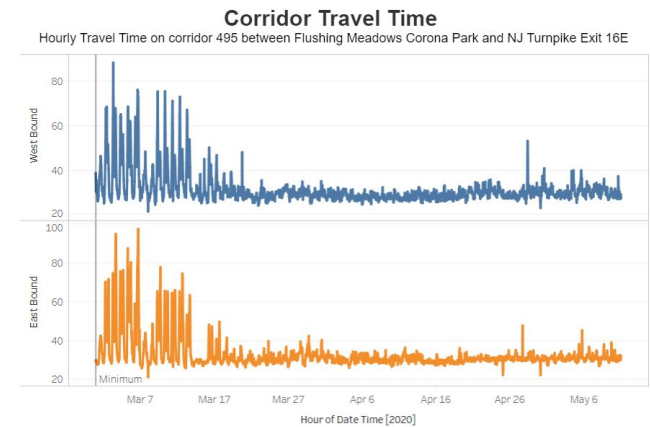

Figure 2: Hourly Travel Time on 495 Corridor (Flushing Meadows Corona Park to NJ Turnpike Exit 16E)

An increase in traffic speed and school zone speeding tickets was observed with some city streets operating at or near free-flow speed during the pandemic. For example, on April 15, during the traditional morning peak hour, 8.3% of the 145 local road segments where data is available, had an average traffic speed over the city speed limit of 25 mph in morning peak hour. 28% of them were over 20mph, while only 3.4% had an average traffic speed over 20mph in February prior to the pandemic.

More mobility metrics for NYC are summarized in Table 2 and notable changes have been seen. For example, crashes remain low but the severity of crashes measured by fatalities is up. This may possibly be due to higher vehicular speeds, but more data is needed

for further investigation. Bikeshare ridership remains down with the exception of the first 12 days of March. Ridership patterns did change, with 15% fewer Friday and Saturday trips and a 20% increase in average trip duration in March 2020 compared to the same month last year. If not specified, the percentage shown in the table compares the data in 2020 with 2019.

3.2 Seattle

The greater Seattle area is the first epicenter of the COVID-19 outbreak in U.S., as it reported the first confirmed case in late January. Washington state, in particular the greater Seattle area, therefore responded very early to set up guidelines and executive orders of social distancing to slow down the spread of the novel coronavirus.

Travel time data for a segment on Interstate-5 were collected via Google Directions API [6] and analyzed to investigate the trends on corridor travel time. Travel time started to decrease since early March and reached the lowest level (near free-flow travel time) since early April. It remained low for the entire April and began to increase in early May.

The critical date analysis in Figure 3 demonstrated that the typical spikes of commute peaks disappeared since the release of stay-at-home order on March 23, 2020, indicating most people reacted immediately to the order. Due to less traffic, the reliability of travel time became higher: the standard deviation of daytime travel time of the corridor reduced to 0.67 minutes in late April compared with 6.43 minutes in late February.
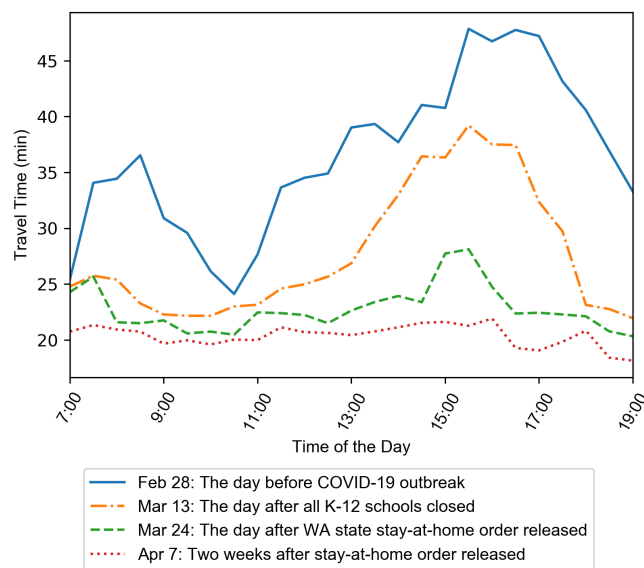

Figure 3: Critical Day Demonstration for I-5 Travel Time

Traffic volumes at three freeways locations (I-5 downtown Seattle near Freeway Park, I-5 NE of Green Lake Park, and SR-520 toll bridge) showed a consistent drop in traffic volume in March, followed by a gradual slow increase in April. Using the same week of previous year as a baseline, traffic volume at the I-5 Downtown Seattle station dropped to its lowest level -46.91% in the week of

March 30, and gradually increased to -41.95%, -35.50%, and -26.31% in the weeks of April 13, 27, and May 11, respectively. Similar trends have been observed for the other two stations. The continuous increasing trend of traffic volume indicates that the impact of traffic in Seattle is slowly returning to pre-pandemic despite stay-at-home restrictions remaining in place. On the other hand, public transit demand [2] remained low since late March and has not started to recover yet. Mobility metrics for Seattle are shown in Table 3.

4 SOCIABILITY BOARD

4.1 New York City

Although social distancing orders are announced, it remains unclear how people are responding to these policies. Understanding the actual reduction in social contact and actual compliance rate is important to measuring the effectiveness of the policy. Therefore, 311 complaints and real-time camera videos are used to examine the compliance of social distancing policies. An increase in social distancing complaints has made the Non-emergency police matter category the 2nd most common complaint among NYC311 reports. In addition, the real-time social distancing behavior has been exampled with the analysis of video data. Figure 4 presents the video processing and data analysis process. Applying object detection to real-time traffic camera videos [16] [17] with a frame rate of 30 seconds at multiple key locations in different zones within the city provides additional information about crowd density [9]. One of the state-of-the-art algorithms, RetinaNet [10], and ResNet-50 [7] are used as the backbone network architecture. The model was pre-trained using the COCO data set [11]. The open-source platforms TensorFlow [1] and Keras [5] are utilized as the machine learning library's support, along with OpenCV [3] for basic video processing.

To calculate the social distancing compliance rate, the centroids of detected pedestrians' bounding box are identified and the distance between the centroids are calculated [14]. Next, the ratio of real height and pixel height (R-P ratio) are computed to project the real distance by assuming every person has the same heights (1.70 meters/5.58 feet is used in this study). Traffic-related objects (person, car, truck, bicycle, bus) and the total number of violated social distancing pairs are reported for each frame. Figure 5 presents an example of the video processing output. The program was run on an instance which was configured with Intel Xeon processors up to 3.1GHz and 16GiB of memory. The GPU is Nvidia Tesla M60. The running time is around 0.94 sec/frame without real-time visualization and 1.06 second/frame with visualization. The sociability metrics for NYC is summarized in Table 4.

4.2 Seattle

Biking and walking for recreational purposes become more popular than ever in Seattle during the pandemic. Data from SDOT bike and pedestrian counters show that these activities dramatically changed during the COVID-19 outbreak period. At Fremont Bridge, a critical commuting connector between North Seattle and downtown Seattle, higher bike counts (+3,500) were observed on weekdays while fewer counts (1,000) on weekends in February. Bike counts reduced to the lowest level in mid-March, but grew back to nearly 2,500, without a

Table 2: COVID-19 Mobility Metrics (NYC)

Week	Mar 9	Mar 16	Apr 13	May 4
Traffic Volume*	-16%	-42%	-67%	-31%
Transit Ridersship	-20%	-72%	-92%	-89%
Corridor Travel Time	-14%	-32%	-36%	-35%
Citibike Trips	+13%	-38%	-	-
Crashes	-22%	-49%	-81%	-
Citywide speeds	Average speeds on Midtown Avenues: 108% increase at 8AM-6PM (Apr vs. Feb 2020)			
School Zone Speeding Tickets	Increased by 72% during Mar 13 to Apr 19 vs. Jan 1 to Mar 12 in 2020			
Freight Traffic	The number of very heavy trucks (GVW > 100 kips) was down for 30% for QB and 44% for SIB traffic during Mar 13 to Apr 12 compared with Feb 3 to Mar 13			
Fatality Rate	Increased from 1.4 to 1.9 fatalities/1000 crashes in the first three weeks for Apr compared to the same period of Feb 2020			

*Via inter-borough crossing

Table 3: COVID-19 Mobility Metrics (Seattle)

Week	Mar 30	Apr 13	Apr 27	May 11
Traffic Volume	-46.91%	-41.95%	-35.50%	-26.31%
Peak Hour Travel Time*	-30.27%	-41.75%	-42.09%	-36.59%
Public Transit Demand	-78%	-81%	-81%	-81%
Travel Time Reliability	Travel time standard deviation for daytime (7AM-7PM) decreased from 6.43 (week of Feb 24) to 0.67 min (week of Apr 30)			

*Travel times are compared with the week of Feb 24, 2020
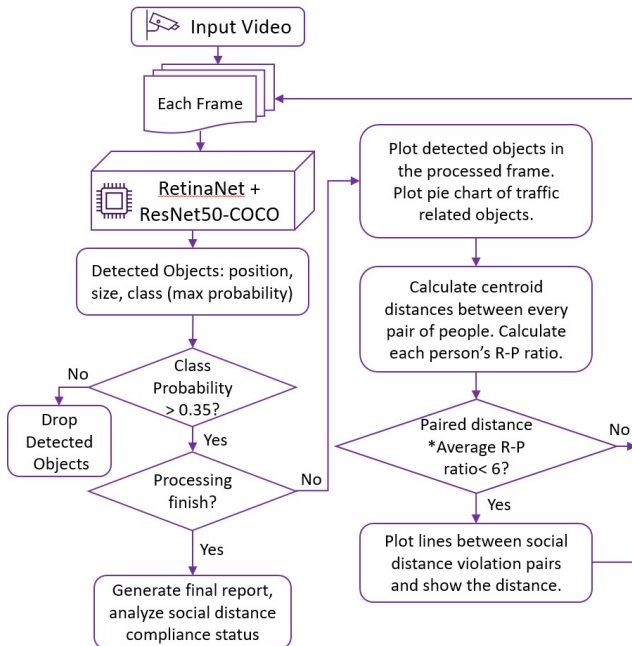

Figure 4: Video Processing and Data Analysis Process

clear weekday/weekend pattern in April, possibly due to an increase in the use of bikes for recreational purposes.
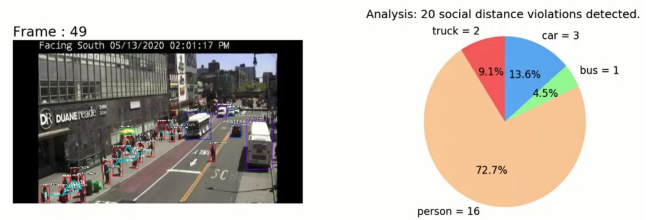

Figure 5: Example of Video Processing Output

Table 4: COVID-19 Sociability Metrics (NYC)

	Apr 2	May 13
Average Peds Density (#/frame)	2.6	2.8
Max Peds Density (#/frame)	20	24
Peds Social Distancing Compliance	91%	87%
311 Social Distancing Complaint Ranking	Top 2 (since Mar 30)	

The daily pedestrian pattern of Burke Gilman Trail, a 27-mile multi-use trail was also investigated. From mid-March, the daily pedestrian counts almost doubled compared with the period of pre-pandemic and such immediate shift reflected the impact of social distancing rules announced in mid-March. Since small businesses, recreational and commercial facilities were closed, people switched to public or community trails for recreational purposes. The statistics for bike/pedestrian counts are summarized in Table 5. The

pedestrian counts increased over 50% in April compared with the same period in 2019.

Similar to NYC, the social distancing violation detection algorithm was implemented for a local street intersection (Broadway E Pike St EW) in Seattle with a frame rate of 30 seconds. Results on May 18 show an average pedestrian density of 3.2 people/frame, a maximum pedestrian density of 12 people/frame, and a pedestrian social distancing compliance rate of 89%. More camera data will be collected and enlarge the coverage of social distancing compliance detection.

Table 5: COVID-19 Sociability Metrics (Seattle)

Week	Mar 30	Apr 6	Apr 13	Apr 20
Bike Counts	-45.33%	+8.80%	-11.38%	-40.02%
Peds Counts	+61.31%	+106.87%	+64.85%	+53.83%

5 CONCLUSIONS AND FUTURE WORK

This paper introduces an interactive data dashboard developed by C2SMART researchers for COVID-19 analysis. An integrated database that fuses information from multiple data sources including traditional traffic detectors, crowdsourcing applications, probe vehicles, real-time traffic cameras, and police and hotline reports for NYC and Seattle was build. The platform is based on cloud computing and data mining techniques for data acquisition and processing and supports interactive data visualization and data analytics for quantifying multiple mobility and sociability metrics. It serves as an all-in-one data fusion tool for open data that is not easy to find in a single place and is highly scalable that can be extended to other cities.

As of the first week of May, data from both NYC and Seattle shows that auto traffic volumes began to rise even when the stay-at-home restrictions were still in place. Moreover, while vehicular traffic volume has started to increase, transit usage remains low indicating a potential shift in mode choice and increased preference for non-public modes of travel, at least in the near future. The crowd density and social distancing compliance for pedestrians based on data obtained from publicly available DOT cameras using state-of-the-art video processing techniques can provide useful insights into daily pedestrian and cyclist demands and their behavior in terms of maintaining social distance.

Through this interactive data dashboard, we are hoping to provide researchers, transportation authorities, and the general public with information they can use to track the impact of the outbreak on our transportation systems and thus to support them to make more effective data-driven decisions. The current dashboard contains data from NYC, Seattle, Chicago and six cities from China and will eventually expand to cover more cities worldwide. As the world continues to adjust to the new reality of COVID-19, C2SMART researchers are continuing to collect data, including perishable mobility, safety, and behavior data, and will continue to monitor these trends and regularly update findings for different reopening stages. In the future, the integrated database will be further utilized for predictive analysis to assist developing effective actionable strategies to plan for potential future scenarios.

REFERENCES

[1] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek Gordon Murray, Benoit Steiner, Paul A. Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zhang. 2016. TensorFlow: A system for large-scale machine learning. *CoRR* abs/1605.08695 (2016). arXiv:1605.08695 http://arxiv.org/abs/1605.08695

[2] Transit App. 2020. *How coronavirus is disrupting public transit.* https://transitapp.com/coronavirus

[3] Gary Bradski. 2000. The opencv library. *Dr Dobb's J. Software Tools* 25 (2000), 120–125.

[4] C2SMART. [n.d.]. *Development and Tech Transfer of Multi-agent Virtual Simulation Testbed Ecosystem.* http://c2smart.engineering.nyu.edu/tech-transfer-simulation-test-bed/#1556735549504-aa267ce9-8698

[5] Francois Chollet et al. 2015. *Keras.* https://keras.io/

[6] Google. [n.d.]. *Google Directions API.* https://developers.google.com/maps/documentation/directions/intro

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2015. Deep Residual Learning for Image Recognition. *CoRR* abs/1512.03385 (2015). arXiv:1512.03385 http://arxiv.org/abs/1512.03385

[8] Abdullah Kurkcu, Ender Faruk Morgul, and Kaan Ozbay. 2015. Extended implementation method for virtual sensors: Web-based real-time transportation data collection and analysis for incident management. *Transportation Research Record* 2528, 1 (2015), 27–37.

[9] Abdullah Kurkcu and Kaan Ozbay. 2017. Estimating pedestrian densities, wait times, and flows with wi-fi and bluetooth sensors. *Transportation Research Record* 2644, 1 (2017), 72–82.

[10] Tsung-Yi Lin, Priya Goyal, Ross B. Girshick, Kaiming He, and Piotr Dollár. 2017. Focal Loss for Dense Object Detection. *CoRR* abs/1708.02002 (2017). arXiv:1708.02002 http://arxiv.org/abs/1708.02002

[11] Tsung-Yi Lin, Michael Maire, Serge J. Belongie, Lubomir D. Bourdev, Ross B. Girshick, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. 2014. Microsoft COCO: Common Objects in Context. *CoRR* abs/1405.0312 (2014). arXiv:1405.0312 http://arxiv.org/abs/1405.0312

[12] Ender Faruk Morgul, Hong Yang, Abdullah Kurkcu, Kaan Ozbay, Bekir Bartin, Camille Kamga, and Richard Salloum. 2014. Virtual sensors: Web-based real-time data collection methodology for transportation operation performance analysis. *Transportation Research Record* 2442, 1 (2014), 106–116.

[13] The Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) 2020. *COVID-19 Dashboard.* The Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). https://coronavirus.jhu.edu/map.html.

[14] Ding Wang, Fan Zuo, Jingqin Gao, Yueshuai He, Zilin Bian, Suzana Duran Bernardes, Chaekuk Na, Jingxing Wang, John Petinos, Kaan Ozbay, Joseph Y.J. Chow, Shri Iyer, Hani Nassif, and Xuegang Jeff Ban. 2020. *White Paper: Agent-based Simulation Model and Deep Learning Techniques to Evaluate and Predict Transportation Trends around COVID-19.* C2SMART center. http://c2smart.engineering.nyu.edu/wp-content/uploads/2020/06/C2SMART-COVID-19-Whitepaper-Issue3-1.pdf

[15] World Health Organization 2020. *WHO Director-General's opening remarks at the media briefing on COVID-19.* World Health Organization. https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020.

[16] Kun Xie, Chenge Li, Kaan Ozbay, Gregory Dobler, Hong Yang, An-Ti Chiang, and Masoud Ghandehari. 2016. Development of a comprehensive framework for video-based safety assessment. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC).* IEEE, 2638–2643.

[17] Kun Xie, Kaan Ozbay, Hong Yang, and Chenge Li. 2019. Mining automatically extracted vehicle trajectory data for proactive safety analytics. *Transportation research part C: emerging technologies* 106 (2019), 61–72.