

# Measuring Human and Economic Activity from Satellite Imagery to Support City-Scale Decision-Making during COVID-19 Pandemic

Rodrigo Minetto<sup>1</sup>, Mauricio Pamplona Segundo<sup>2</sup>, Gilbert Rotich<sup>2</sup> and Sudeep Sarkar<sup>2</sup>

<sup>1</sup> The Federal University of Technology - Paraná (UTFPR)

<sup>2</sup> Institute for Artificial Intelligence (AI+X), University of South Florida (USF)

**Abstract**—The COVID-19 outbreak forced governments worldwide to impose lockdowns and quarantines over their population to prevent virus transmission. As a consequence, there are disruptions in human and economic activities all over the globe. The recovery process is also expected to be rough. Economic activities impact social behaviors, which leave signatures in satellite images that can be automatically detected and classified. Satellite imagery can support the decision-making of analysts and policymakers by providing a different kind of visibility into the unfolding economic changes. Such information can be useful both during the crisis and also as we recover from it.

In this work, we use a deep learning approach that combines strategic location sampling and an ensemble of lightweight convolutional neural networks (CNNs) to recognize specific elements in satellite images and compute economic indicators based on it, automatically. This CNN ensemble framework ranked third place in the US Department of Defense xView challenge, the most advanced benchmark for object detection in satellite images. We show the potential of our framework for temporal analysis using the US IARPA Function Map of the World (fMoW) dataset. We also show results on real examples of different sites before and after the COVID-19 outbreak to demonstrate possibilities.

Among the future work is the possibility that with a satellite image dataset that samples a region at a weekly (or biweekly) frequency, we can generate more informative temporal signatures that can predict future economic states. Our code is being made available at <https://github.com/maups/covid19-satellite-analysis>

## I. INTRODUCTION

The COVID-19 outbreak is changing the world as never seen before. The lockdowns and quarantines implemented worldwide can be noticed even from space. Spatial agencies such as the US National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA) observed a significant decrease in nitrogen dioxide emissions over major metropolitan areas around the world as a consequence of the economic slowdown. However, the potential use of remote sensing data goes far beyond. As an example, the European Union Commission requested the sharing of any satellite imagery related to the pandemic for research purposes<sup>1</sup>. Such images will support decisions concerning:

- traffic issues, to ensure citizens' mobility but at the same time to avoid traffic jams that block the exchange of essential supplies;

- medical infrastructure, to have knowledge about any temporary medical facility construction around Europe and to gain awareness on the impacts and actions taken in the face of the outbreak;
- facilities activity, to safely and economically maximize resources; and
- social distancing, to appraise if people are following orders during a quarantine.

High-resolution imagery, as provided by sophisticated satellites like WorldView-3 [1] that daily collect panchromatic images with a ground sample distance (GSD) of 0.3 meters around the globe, can be a valuable asset to estimate the impacts of COVID-19 in society.

As shown in Figure 1, temporal analyses of strategic sites can provide critical indicators of human and economic activities. They provide information on traffic or distancing issues through detecting and classifying vehicles. They can keep track of new medical infrastructure being built by identifying construction elements such as bulldozers, excavators, trucks, and tents. And they can measure economic activity by detecting commercial transports such as planes, ships, and locomotives. However, to unleash the power in satellite images, we need automated AI-based computer algorithms to extract these kinds of information from them, without requiring extensive manual labor, so local decision-makers all over the world can use them without time lag.



(a) Airport before and after the COVID-19



(b) Car rental parking lot before and after the COVID-19

<sup>1</sup><https://www.euspaceimaging.com/eu-commission-asks-eo-community-for-help-with-covid-19/>

In this work, we present a framework that recognizes specific elements in strategic locations to compute such indicators automatically. As part of this work, we describe an ensemble of convolutional neural networks (CNN) for simultaneous detection and classification of objects in high-resolution aerial images. This approach is state of the art and ranked third place in the US Department of Defense conducted xView challenge [2]. The xView dataset is very relevant to the COVID-19 problem because it used WorldView-3 as a source for more than 1,100 high-resolution images spanning about 800,000 aerial objects around the world, and covering a total area of 1,400 square kilometers. The organizers provided annotations for 60 classes of objects, with many of them being particularly relevant to the task of this work. We employ a combination of strategic location sampling and a lightweight CNN architecture to perform satellite image processing and analysis within an acceptable time frame. With that, we hope to support regular economic assessment and decision making processes.

We present our framework within the context of stay-at-home order enforcement (Section III) and discuss later how to adapt it to other scenarios (Section V). In our experiments, we first evaluate the detection performance on the xView dataset (Section IV-A) and then show its potential for temporal analysis using the US IARPA Functional Map of the World (fMoW) dataset [3] (Section IV-B). Finally, we show our framework in action on real examples of world scenes before and after the COVID-19 outbreak (Section V).

## II. RELATED WORK

The use of satellite imagery is of paramount importance to support the management of natural disasters, humanitarian assistance, and environmental conservation policies. In recent years, the unprecedented amount of data captured by sophisticated satellites has profoundly impacted the information quality and the demand for techniques to extract knowledge from it.

When concerning the infrastructure for the quick development of algorithms, the number of large-scale challenge problems devised to foster breakthroughs in this field is remarkable. The SpaceNet challenge [4] focused on the automated building footprint extraction and road network detection. The organizers used the WorldView-2 satellite to collect high-resolution images to cover more than 683,000 buildings and 8,676 road stretches of five metropolitan areas. As observed by them, these mappings are of particular interest in natural disasters. By using a satellite with a daily revisit time, it would be possible to quickly identify damaged buildings and blocked/destroyed roads and prepare the logistics for humanitarian assistance accordingly. This topic was later embraced by the xView2 challenge [5], which focused on assessing building damage after a natural disaster. The organizers released pre- and post-disaster images from 850,000 buildings around the world, depicting the effects of earthquakes, tsunamis, floodings, volcanic eruptions, wild-fires, tornados, and hurricanes. The US IARPA Functional Map of the World (fMoW) challenge [3] encouraged the

design of automated solutions for land use classification in satellite images. Its dataset comprised more than one million excerpts of multispectral images from 63 categories, including satellite metadata and temporal views. Among these categories are hospitals, educational institutions, airports, prisons, parks, electric and fire substations, places that are worth monitoring during a pandemic.

Remote sensing also plays a vital role in the study of human diseases, with many works stipulating associations between terrain characteristics and disease incidence. Rogers [6] observed a correlation between African trypanosomias causing sleeping sickness and indices of temperature, rainfall, and vegetation obtained from satellite imagery. Rogers *et al.* [7] later perceived that sensing seasonal climate could help to predict mosquito vectors that are responsible for malaria transmission. Dister *et al.* [8] investigated the relationship between Lyme disease and measurements of vegetation structure, wetness, and abundance. Cyranoski [9] mapped wetlands to study the spreading of avian influenza. Ford *et al.* [10] showed how sea surface temperature, sea surface height, and chlorophyll A levels can be used to predict outbreaks of cholera. Garni *et al.* [11] used land cover and topography information to map the risk of occurrence of cutaneous leishmaniasis.

In the field of economics, satellite images have helped to estimate different indicators. There are plenty of works based on satellite-recorded nighttime lights, as they provide a reasonable valuation of economic activity. Regression of gross domestic product (GDP) [12], [13], poverty levels [14], [15] and development indices [16] based on this information were deemed plausible in the literature. Recently, the analysis of high-resolution daytime satellite images improved such measurements [17], [18] thanks to the advances brought by deep learning [19].

The bio-inspired CNN [20], a popular deep learning choice nowadays, is composed of multiple layers of artificial neurons and is used to learn representations with various levels of abstraction. Its ability to discover intricate patterns in massive data [21] has made it a perfect tool for remote sensing. It currently supports a myriad of applications in the literature, such as semantic segmentation [22], target localization [23], region classification [24], super-resolution [25], regression models for environmental knowledge extraction [26], understanding of temporal and spatial variations [27], study of semantic relationships between aerial targets [28], 3D reconstruction [29], and hyperspectral image generation [30]. As detailed in the next section, we also use deep learning as a tool to extract knowledge from satellite imagery. The main difference to other works is that we do not regress indicators directly from the images, but from information obtained from them, such as the number of vehicles, trucks, buildings, and so on. This strategy allows us to create indicators that are informative, understandable, and supportive in decision-making.

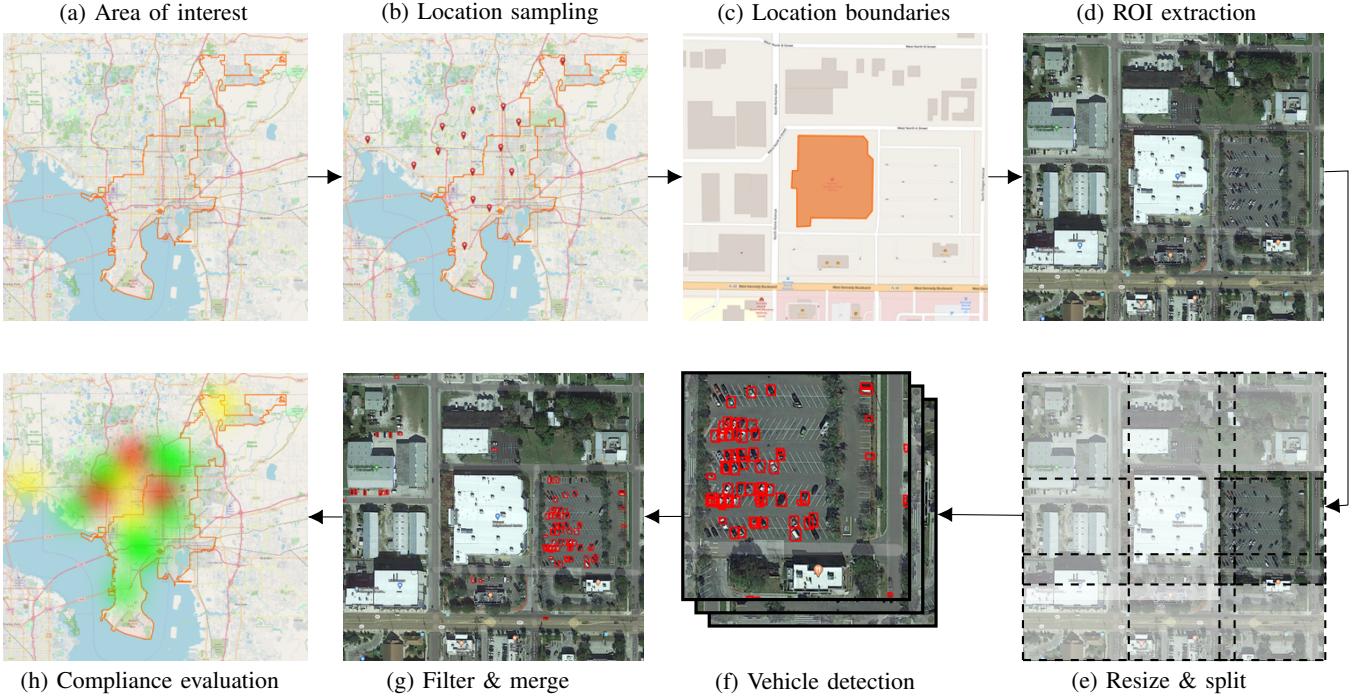


Fig. 2. We present a workflow to analyze the pattern of vehicles over time to monitor compliance with stay-at-home order. Similar workflows are possible for other aspects of the economy, such as supply-chain disruptions. Other than the first step, the rest is fully automatic. (a) A human demarcates an area of interest. (b) The algorithm samples strategic locations. (c) The algorithm then looks for their boundaries in open knowledge sources to delimit regions to arrive at (d) regions of interest (ROI) in satellite images. (e) Each ROI is then automatically resized and split into several small parts to be (f) processed by a vehicle detector. (g) And the results are filtered and merged into a single outcome. (h) Finally, and this part is still conceptual, an algorithm will analyze the history of vehicle occupancy in each sampled location to help identify non-compliant zones. Maps, landmarks, and boundaries were obtained from OpenStreetMap<sup>\*</sup>. Satellite images were obtained from Google Maps<sup>\*\*</sup>.

\* © OpenStreetMap contributors (<https://www.openstreetmap.org/copyright>)

\*\* Image © 2020 Google, Maxar Technologies (<https://www.google.com/permissions/geoguidelines/>)

### III. PROPOSED FRAMEWORK

The idea of analyzing the flow of vehicles under a stay-at-home order in a large area, such as a city or a county, using satellite images is hard to execute due to the vast amount of data to be processed. Thus, data sampling is necessary to reduce the computational cost so that it is possible to generate content to aid hazard assessment and decision making by authorities within an acceptable time frame. The sampling strategy, however, has to take the relevance of the selected regions to the problem into account. This because traditional sampling methods, such as random and grid sampling, tend to pick too many meaningless regions, which could lead authorities to incorrect assumptions. In a stay-at-home scenario to avoid disease proliferation, like the ones occurring due to the recent COVID-19 outbreak, the surrounding of places that gather crowds, like airports, schools, hospitals, churches, malls, and supermarkets, should be prioritized over underpopulated areas.

In this work, we present a complete framework to map increases and decreases in the flow of vehicles over an area of interest by combining open knowledge sources, such as OpenStreetMaps<sup>2</sup>, satellite images, and a machine learning-based vehicle detector. Figure 2 illustrates the sequence of

stages that compose our proposed framework. These stages are detailed in the following sections.

#### A. Location sampling and region of interest (ROI) extraction

The first stage in our framework consists of a strategic sampling of locations within an area of interest. To do so, first, we need to define what is a strategic location. In this work, it can be any place with a high circulation of people that may contribute to the proliferation of pathogens. More specifically, we look for items with the following tags in the OpenStreetMap database: ‘shop=supermarket’, ‘aeroway=aerodrome’, ‘amenity=hospital’, ‘amenity=university’, ‘amenity=school’, ‘shop=mall’, and ‘amenity=place\_of\_worship’. This tag list can be easily extended if necessary, or even redesigned for other applications.

Among the recovered items within the area of interest (see Figure 2(b)), we select the ones that contain annotations for the boundary contour (see Figure 2(c)) in the form of a sequence of latitude and longitude coordinates. We find the smallest enclosing bounding box for the coordinates of each item, which is then expanded  $m$  meters in all directions to delimit the item’s ROI in a satellite image (see Figure 2(d)). Smarter ROI extraction strategies can be used, such as parking lot detection near strategic locations,

<sup>2</sup><https://www.openstreetmap.org>

upon the availability of reliable techniques and resources to support them.

### B. Vehicle detection

The input for this stage is a group of ROI images extracted in the previous stage, and the output for the  $i$ -th ROI is a set of  $n_i$  detected regions  $\mathcal{R}^i = \{r_1^i, r_2^i, \dots, r_{n_i}^i\}$ . Each region  $r_j$  is defined by an axis-aligned rectangular box  $b(r_j) = (x_1, y_1, x_2, y_2)$  where  $(x_1, y_1)$  and  $(x_2, y_2)$  represent the upper left and bottom right corners, respectively. A score  $w(r_j)$  expresses the confidence of the detection within the interval  $[0, 1]$ .

In this work, we create an ensemble of Single Shot Multibox Detectors (SSD) [31] for vehicle detection. To do so, we combine two models released as baselines for the xView dataset [2], *Vanilla* and *Multires*, by using different parameters for image resizing, splitting, and output merging.

Even though ROIs are small parts of satellite images, they may still be too large for carrying out vehicle detection directly. Both baseline models receive  $300 \times 300$  images as input. So we split our ROI images into blocks of  $300 \times 300$  pixels without overlap whenever possible for the *Vanilla* model, as illustrated in Figure 2(e), or with an overlap of 100 pixels for the *Multires* model. Adding an overlap helps to detect vehicles that lay at the edge of two or more adjacent blocks. Besides, a second copy of the *Vanilla* model uses ROI images scaled by a factor of 1.3 to increase the detection accuracy of smaller vehicles. Table I summarizes this arrangement.

Each detector in the ensemble runs separately on each block of its input image (see Figure 2(f)), and we end up with multiple results per block that must be merged into a single outcome. Before that, we eliminate regions whose confidence value is below a threshold  $t$ . Table I indicates the value of  $t$  for each detector. Non-discarded regions for all blocks are mapped back to the ROI coordinate space. Many vehicles may be detected multiple times, either by being detected by different models of the ensemble or by appearing in overlapped block regions. The Non-Maximum Suppression (NMS) [32] algorithm with minor adaptations is used to discard duplicate regions belong to the same object. Consider  $\bar{\mathcal{R}} = \{\bar{r}_1, \bar{r}_2, \dots\}$  the set of regions not yet filtered by NMS. NMS selects the region  $\bar{r}_i \in \bar{\mathcal{R}}$  with the highest confidence score and loops through  $\bar{\mathcal{R}}$  looking for other regions  $\bar{r}_j$  that have an Intersection over Union (IoU) greater than a given threshold  $\sigma$ :

$$\text{IoU}(b(\bar{r}_i), b(\bar{r}_j)) = \frac{\text{area}(b(\bar{r}_i) \cap b(\bar{r}_j))}{\text{area}(b(\bar{r}_i) \cup b(\bar{r}_j))} > \sigma \quad (1)$$

The IoU metric takes into account the total area of both analyzed regions, which is particularly interesting for satellite imagery, where in many cases a significant intersection of objects does not mean that they should be merged (see Figure 3(a)). The region  $\bar{r}_i$  and all other regions that satisfy Equation 1 will form a subset  $\bar{\mathcal{R}}_k \subseteq \bar{\mathcal{R}}$  that will be merged

into a single region  $r_i$ :

$$r_i(b) = \frac{\sum_{\bar{r} \in \bar{\mathcal{R}}_k} w(\bar{r}) \times b(\bar{r})}{\sum_{\bar{r} \in \bar{\mathcal{R}}_k} w(\bar{r})} \quad (2)$$

with  $r_i$  being part of the final result  $\mathcal{R}$ . As may be seen, instead of discarding overlapped regions with lower confidence score, we combine all regions in  $\bar{\mathcal{R}}_k$  through a weighted average. This avoids noise in the final bounding box coordinates, as shown in Figures 3(b) and 3(c).  $\bar{\mathcal{R}}_k$  is then removed from  $\bar{\mathcal{R}}$  and the process is repeated until  $\bar{\mathcal{R}}$  is empty.

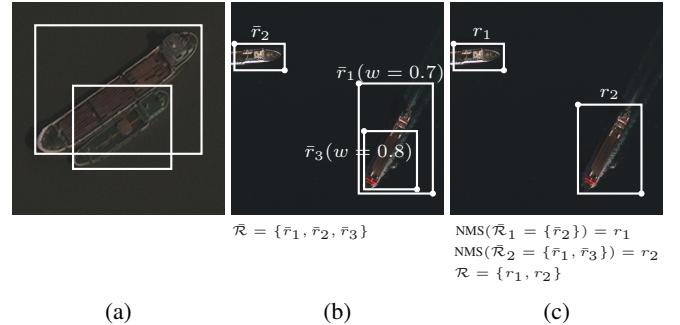


Fig. 3. Example of (a) two overlapped regions from the same category that must not be merged, and of (b) three detected regions in which (c) two of them were merged using their confidence score to define the new bounding box dimensions.

An example of the final output for a ROI is illustrated in Figure 2(g).

### C. Temporal analysis

Given enough time and acquisition frequency, it is possible to apply time series analysis techniques [33] to learn the standard behavior-patterns in each ROI and then identify trends in these areas. It is vital to cope with typical variations in different scales, such as seasonal variations along the year, monthly variations, weekly variations, or even daily variations. Recent approaches based on concept drift [34] can help to identify a change in behavior while avoiding outlier data. However, this aspect has not been implemented yet. We need access to appropriate data to study these variations.

## IV. EXPERIMENTS

### A. Detector evaluation

To evaluate the detection stage, we used the xView dataset detailed by Darius *et al.* [2]. It contains a training set with 847 high-resolution images and about one million annotations for 60 classes of objects. It also contains an evaluation subset with 282 images to which no annotations were provided, and a sequestered testing subset with 284 images. Image sizes range from  $2564 \times 2576$  to  $3187 \times 4994$  pixels. The interpolated mean average precision (mAP), detailed by Henderson and Ferrari [35], can be computed for the training set through its object annotations. A mAP value for the entire evaluation set could be obtained in an online submission system while

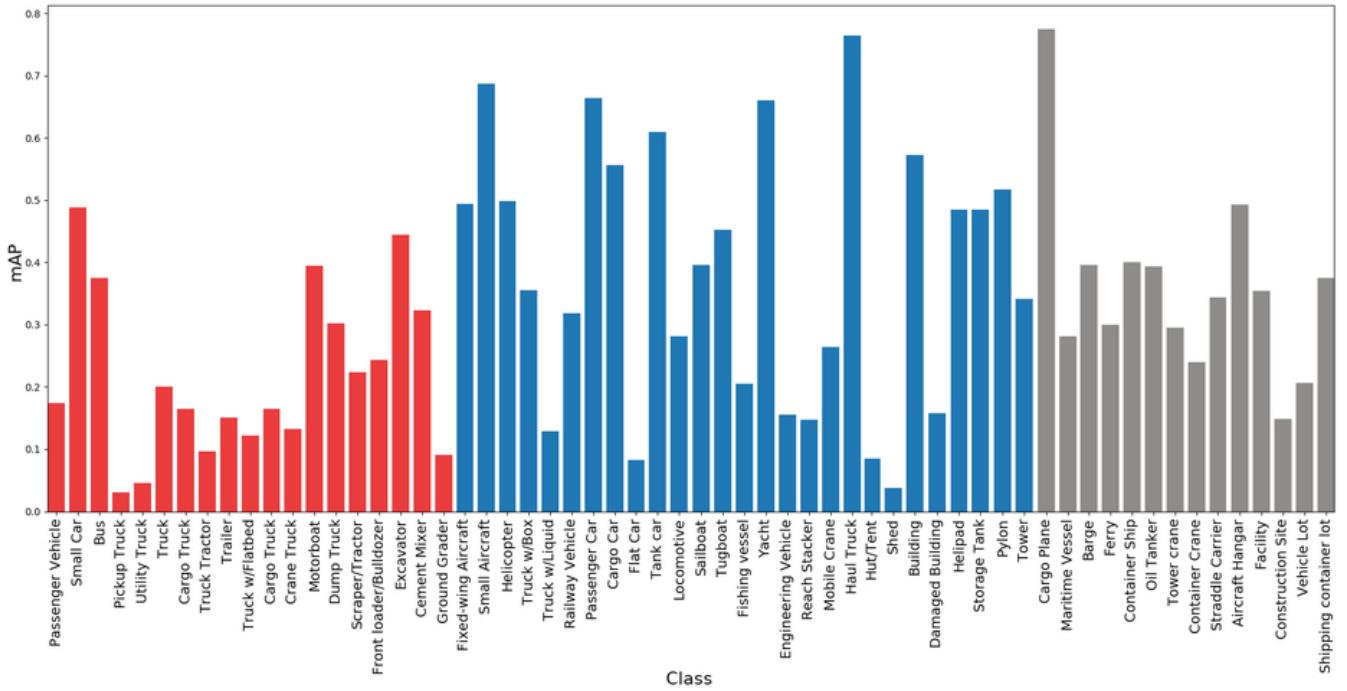


Fig. 4. Mean average precision per xView class. Red, blue and gray bars represent small, medium and large targets, respectively.

the xView competition was running. The precision for the sequestered testing set could only be computed by the xView organizers.

The xView classes were divided into three groups according to the object size: small, medium and large. A complete list of classes per group is available in the competition website<sup>3</sup>. The ensemble configuration described in Section III-B is used for the small group only, which includes the class ‘*small\_car*’ that we use for vehicle detection. A complete description of our ensemble is shown in Table I, including additional detectors and their parameters for objects of medium and large sizes as well. As other classes are useful for future analyses (discussed in Section V), we report detection results for them as well.

TABLE I  
PARAMETERS OF OUR ENSEMBLE OF SSDS FOR ALL CLASSES IN THE XVIEW DATASET. DETECTORS #1-#3 ARE USED FOR SMALL OBJECTS (e.g., ‘*passenger-vehicle*’, ‘*small-car*’, ‘*pickup-truck*’); DETECTORS #1-#4 FOR MEDIUM OBJECTS (e.g., ‘*small-aircraft*’, ‘*helicopter*’, ‘*sailboat*’); AND DETECTORS #3-#5 FOR LARGE OBJECTS (e.g., ‘*passenger-cargo-plane*’, ‘*container-ship*’, ‘*vehicle-lot*’).

	Scale	Overlap	Thr.	Model	Size group
Det. #1	1.0	0px	0.15	Vanilla	Small & medium
Det. #2	1.3	0px	0.06	Vanilla	Small & medium
Det. #3	1.0	100px	0.06	Multires	All
Det. #4	0.7	100px	0.5	Multires	Medium & large
Det. #5	0.6	0px	0.06	Multires	Large

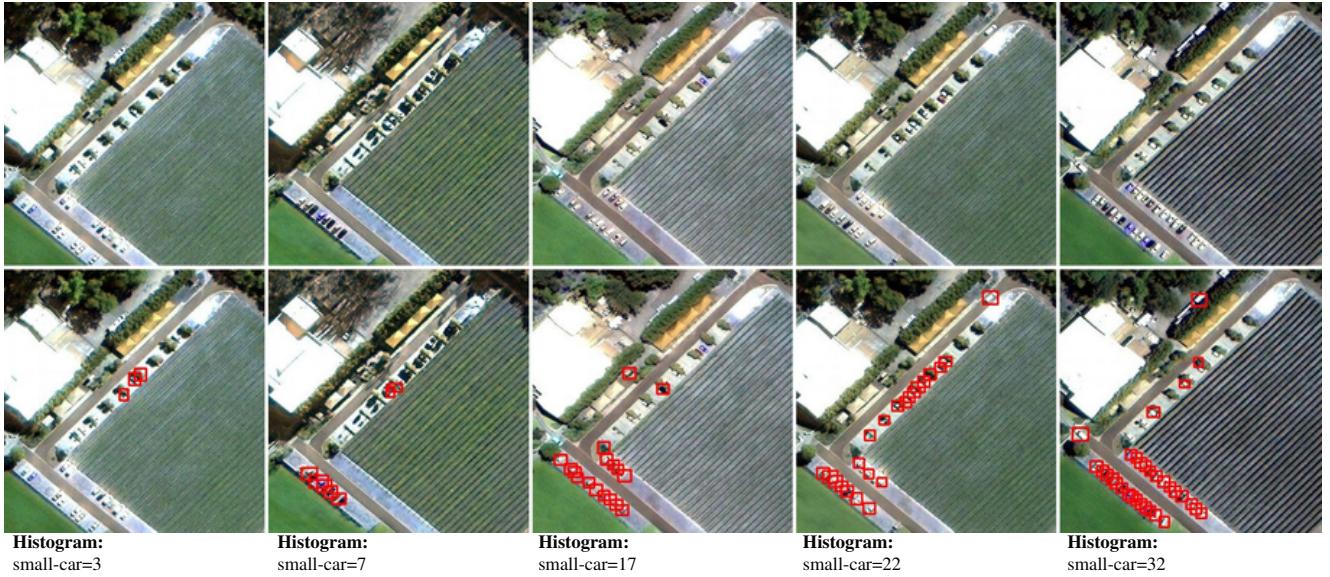
In Figure 4 we show the mAP per class of our ensemble for the training set, which is the only set with annotations that allowed us to do so. Even though the baseline detectors used in our ensemble were trained with this set, the problem is hard enough to prevent detectors from reaching perfect accuracy. Still, this figure gives a good idea of which classes are more accurate than others. The class ‘*small\_car*’, for instance, reaches nearly 0.5 mAP.

This ensemble was submitted to the xView competition and achieved a mAP of 29.88 in the evaluation set, while the baselines *Vanilla* and *Multires* achieved 20.87 and 18.14, respectively. It ranked third over all contestants in the sequestered testing set (see Table II), evidencing the potential of the approach for detecting targets in satellite images.

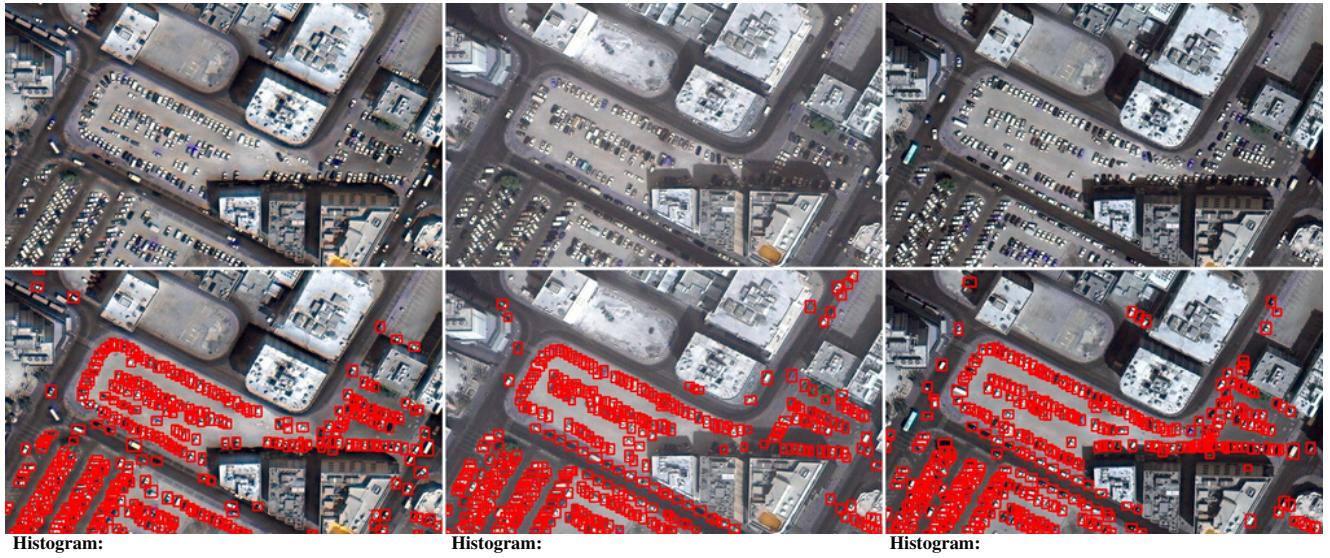
TABLE II  
FINAL XVIEW LEADERBOARD: MAP PER SIZE GROUP AND OVERALL.

Rank	Country	mAP			
		Small	Medium	Large	Score
1	Russia	0.1965	0.3371	0.3400	0.2932
2	Australia	0.1632	0.3595	0.2536	0.2727
3 (proposed)	USA	0.1733	0.3261	0.3039	0.2726
4	Italy	0.1680	0.3339	0.2821	0.2693
5	USA	0.1587	0.2657	0.2511	0.2284

<sup>3</sup><https://challenge.xviewdataset.org/challenge-evaluation>



(a) The difference in vehicle count between the first and last images is 29.



(b) The difference in vehicle count between the first and last images is 12.

Fig. 5. Framework detection results (bottom row) for multiple temporal views of a same region (top row) for a small variation in the flow of vehicles. The samples are sorted from left to right in a non-decreasing order of the number of detected vehicles.

### B. Temporal evaluation

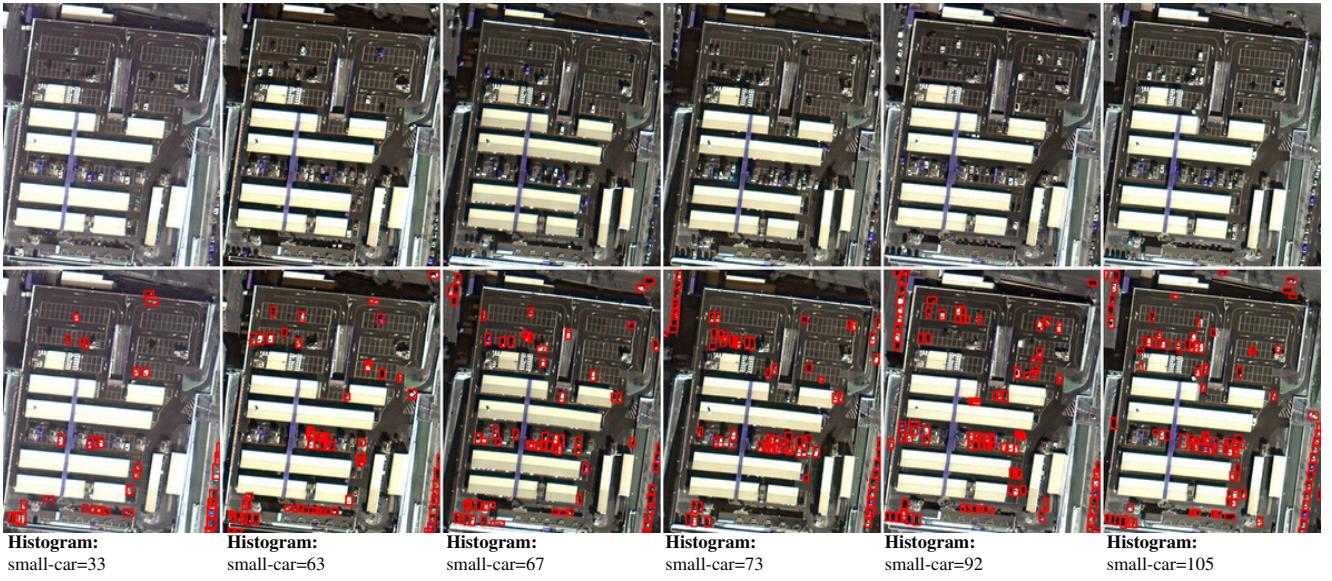
The US IARPA fMoW dataset [3] contains more than one million excerpts of satellite images split into training, evaluation, and testing subsets. Even though it provides high-resolution pan-sharpened images [36], most of them do not have a GSD as low as the ones in the xView dataset. This because this dataset was created for land use classification, not for small object detection. However, it provides temporal views of the same region, which is very interesting for this experiment. Temporality brings variations in shadows, viewpoints, whether, and vehicles in the scene, the last being our primary focus.

Each region in the dataset represents one of the 62 classes or false detection. Among those classes, we are particularly

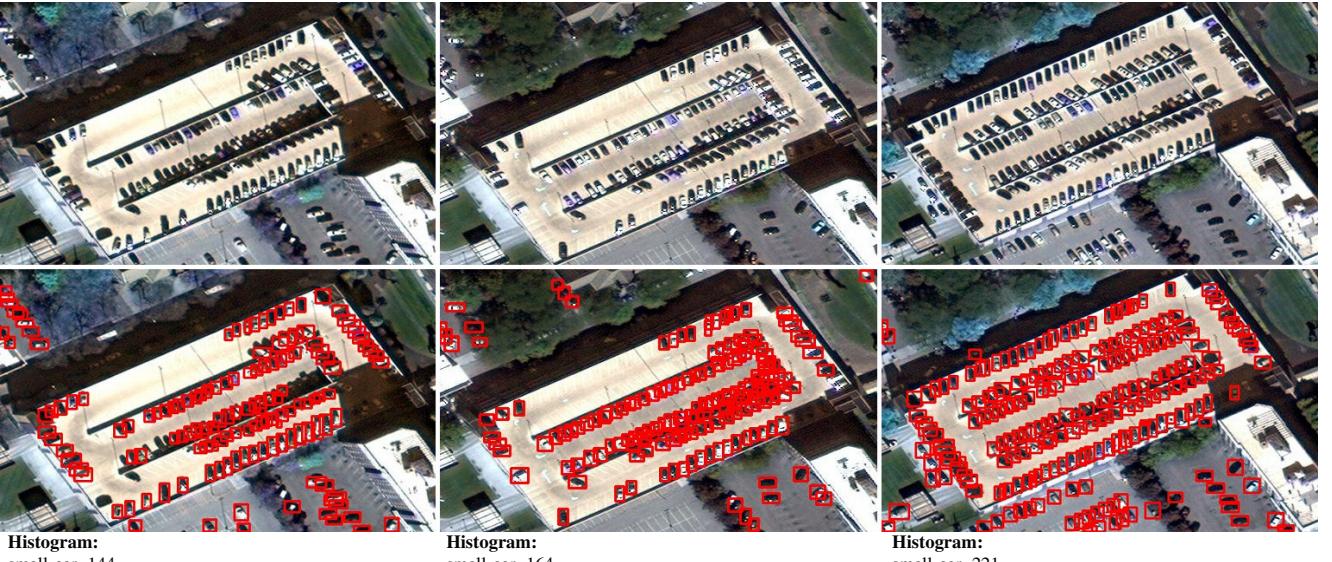
interested in one: ‘*parking\_lot\_or\_garage*’. We looked for regions of this class that:

- have three or more samples with GSD smaller than 0.4 and dimensions greater than  $300 \times 300$  pixels (one block in Section III-B);
- show an open-air parking lot (as ‘*parking\_lot\_or\_garage*’ includes closed garage buildings).

Following these criteria, we were able to obtain images from nine regions. We ran our vehicle detector using each of them as our ROI, and even though there are false positives and negatives in nearly all regions, the count is consistent enough for automated analysis. We sorted samples of the same region in a non-decreasing order of the number of detected vehicles to illustrate the potential of the framework



(a) The difference in vehicle count between the first and last images is 72.



(b) The difference in vehicle count between the first and last images is 77.

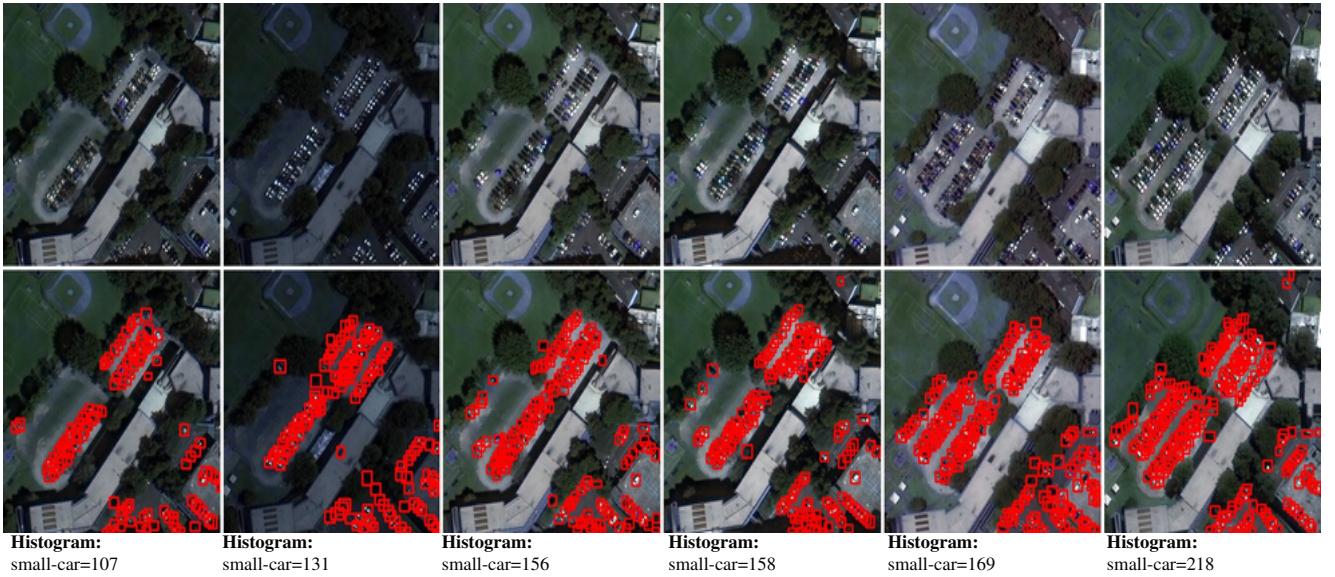
Fig. 6. Framework detection results (bottom row) for multiple temporal views of a same region (top row) for a medium variation in the flow of vehicles. The samples are sorted from left to right in a non-decreasing order of the number of detected vehicles.

to perceive gradual changes in the flow of vehicles. We also sorted these regions in non-decreasing order of the absolute difference in the number of detected vehicles between the samples with the highest and lowest vehicle count. Regions with small, medium and high variation in the number of vehicles are respectively shown in Figures 5, 6 and 7. One can argue if the sorting is correct for some of the samples, but the overall quality of the process is evident.

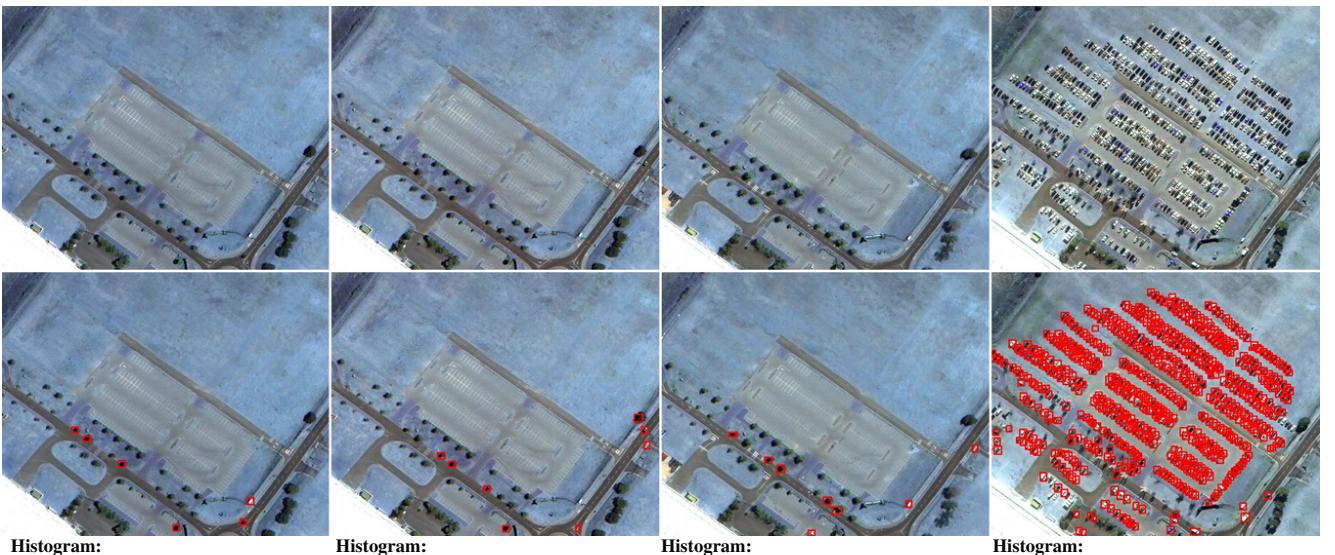
Figure 5 shows the regions with the lowest variation in the flow of vehicles. As can be observed, we can identify regions with a stable count regardless of the volume of cars. This ability is essential for stay-at-home enforcement in situations where a decrease is expected but is not confirmed, requiring further action from the authorities (*e.g.*, suspension

of activities in the public, social and private sectors).

Figures 6 and 7 respectively show regions with medium and high variation. Recognizing such changes is important in both directions, either increasing or decreasing. For instance, a decreasing vehicle count in hospitals can point out a reduction in the outbreak, and in airports can indicate a reduction in the economic activity, which relates to the obedience to a stay-at-home order. Meanwhile, a similar trend in residential areas can suggest both an outbreak reduction or stay-at-home disobedience, depending on the context. Finally, an increasing vehicle count can reveal critical regions that require more attention from authorities. Sudden increases in supermarkets can detect panic buying, and in convention centers the occurrence of large unauthorized events (see the



(a) The difference in vehicle count between the first and last images is 111.



(b) The difference in vehicle count between the first and last images is 648.

Fig. 7. Framework detection results (bottom row) for multiple temporal views of a same region (top row) for a large variation in the flow of vehicles. The samples are sorted from left to right in a non-decreasing order of the number of detected vehicles.

bottom row of Figure 7).

As we can measure variations in the vehicle count, we can also estimate hazard levels for each ROI given an expected behavior (*e.g.* stable count in supermarkets, decrease in schools, increase in rental car facilities). When ROIs are associated to geographic coordinates, it is possible to interpolate these levels to neighboring areas to produce a risk map as the one illustrated in Figure 2(h).

## V. DISCUSSION

We can automatically detect different types of vehicle from satellite imagery, which can be indicators for underlying economic activity. We describe our workflow in the context of enforcing a stay-at-home order (Section III), but it can be used for different purposes, as we will discuss

later. By using two publicly available datasets, xView and fMoW, we estimated how well our detector worked for 60 distinct classes of objects (Section IV-A) and exemplified how well we can perceive the flow of these objects over time (Section IV-B). These results show the potential of the proposed framework to address the intended problem, but miss the real thing – the COVID-19 outbreak.

To evaluate our framework in real-world situations, we collected satellite imagery released to the press to illustrate the impact of the COVID-19 over the globe. It is worth noting that these are not always raw, high-resolution images, such as those provided by xView and fMoW. Many of these images had artifacts that impair automated analysis, like watermarks and low-resolution. We improved the detection

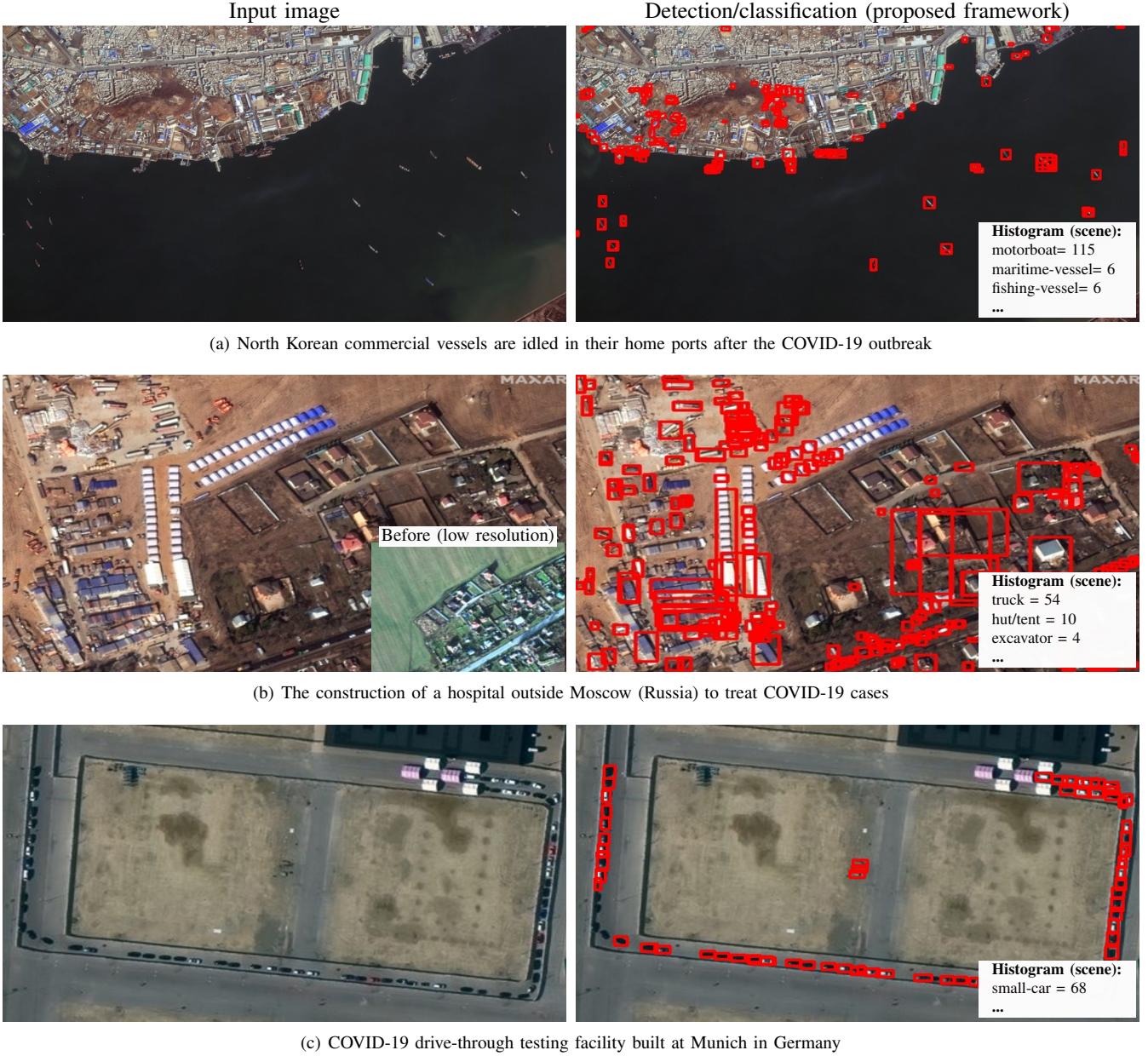


Fig. 8. Satellite imagery from world scenes related to the COVID-19 pandemic and statistics about vehicles/infrastructure available. In order to avoid a large number of detections we are showing the results for some selected categories for a classification confidence threshold of 0.10 and 0.15. **Photo credit:** SATELLITE IMAGE 2020 MAXAR TECHNOLOGIES.

outcome in some cases by upsampling the images or adjusting the confidence threshold. We believe these modifications would not be necessary if we had access to the original files.

In Figure 8(c), we show a drive-through COVID-19 testing site built in Munich, Germany. As can be seen, our automated count is very close to the real number and could help authorities to measure attendance over time in health facilities.

Figures 9(a) and 9(b) show a car rental parking lot in the Phoenix Airport, Arizona, before and after the COVID-19 outbreak, respectively. As fewer people are traveling, the volume of parked vehicles increases substantially. This information serves not only to verify stay-at-home compliance but also to estimate the economic impact in a chain of companies

such as car rental companies, airlines, insurance, etc.

Finally, Figure 9(e) shows car lines in a tollbooth at Wuhan, China, before the COVID-19 outbreak, and Figure 9(f) shows the same tollbooth empty after. This example shows how a simple redefinition of the list of strategic locations (e.g., in this case, to tollbooths, highways, and border crossings) allows our framework to detect traffic jams and migration flows.

Repurposing the framework for other economic activities is simple. The places to monitor, i.e. roads, ports, etc., can be changed easily. The objects of interest can be changed, and our detector can handle many different options. See Figure 4 for a complete list. For instance, Figure 8(a) shows an image

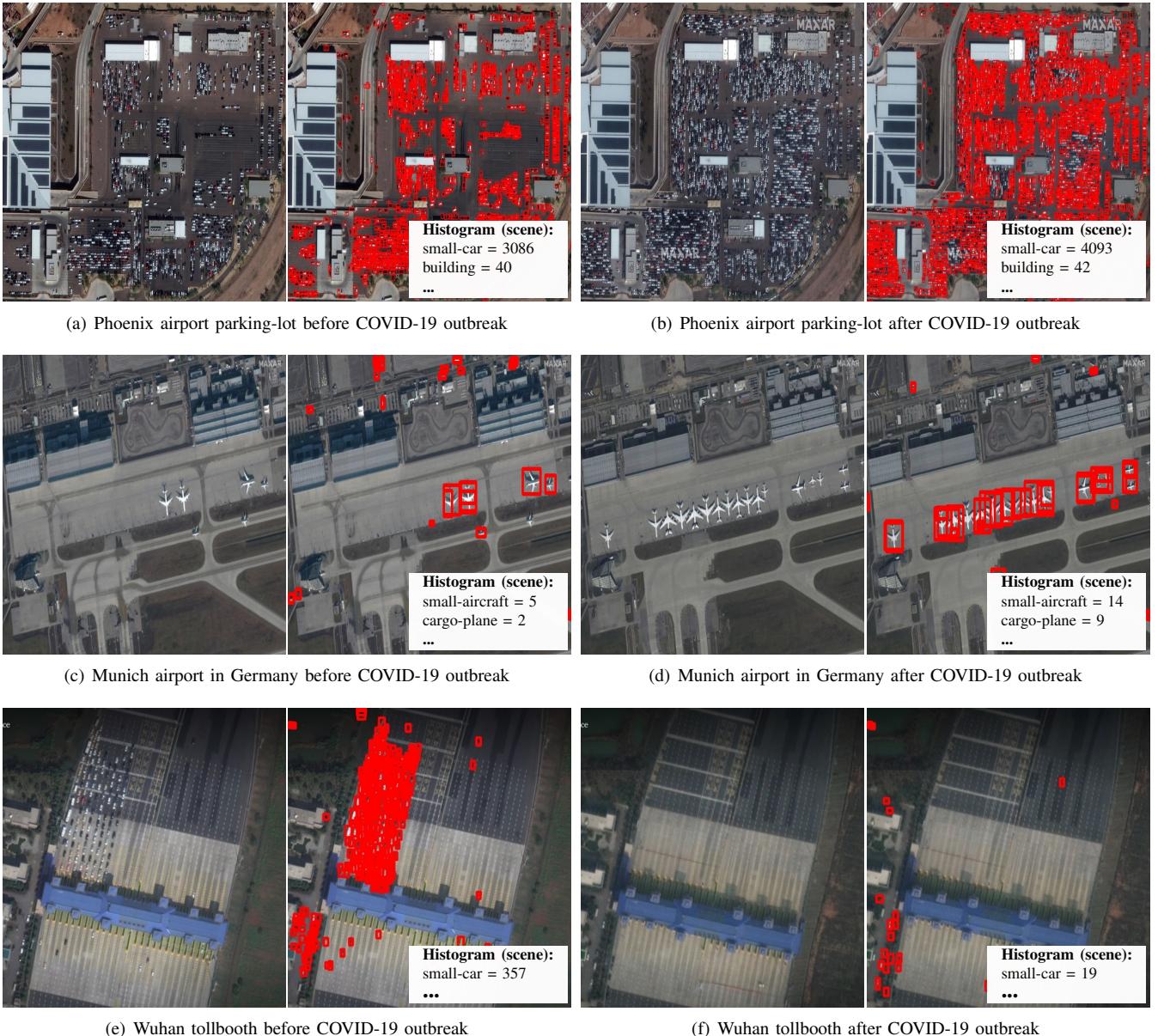


Fig. 9. Satellite imagery from world scenes before and after the COVID-19 outbreak and statistics about vehicles/infrastructure available. In order to avoid a large number of detections we are showing the results for some selected categories for a classification confidence threshold of 0.1. **Photo credit:** SATELLITE IMAGE 2020 MAXAR TECHNOLOGIES.

from North Korean commercial vessels used to transport coal and other commodities, according to Christoph Koettl from the New York Times<sup>4</sup>, stopped in their home ports as an attempt to prevent the virus transmission. We did not find an image with a reasonable resolution before the outbreak. Still, according to other satellite data, this concentration of ships is not typical. Anyway, detecting maritime vessels in commercial ports can be an excellent indicator of economic activity, as maritime transport carries out more than 90% of the world's trade<sup>5</sup>.

Figures 9(c) and 9(d) present another example of an

<sup>4</sup><https://www.nytimes.com/2020/03/26/video/coronavirus-north-korea.html>

<sup>5</sup><https://business.un.org/en/entities/13>

economic indicator. They respectively show the number of grounded planes at the Munich International Airport in Germany before and after the COVID-19 outbreak, which, according to Jonathan O'Callaghan from Forbes<sup>6</sup>, has dramatically increased. Our framework can automatically produce the same observation.

As our final visual example, in Figure 8(b), we show a campaign hospital being built in a field 31 miles outside of Moscow, Russia, to treat COVID-19 patients, as reported by Dave Mosher from Business Insider<sup>7</sup>. We were able

<sup>6</sup><https://www.forbes.com/sites/jonathanocallaghan/2020/03/26/these-erie-photos-reveal-the-impact-of-the-coronavirus-pandemic-from-space/>

<sup>7</sup><https://www.businessinsider.com/coronavirus-covid-19-russian-hospital-field-near-moscow-satellite-photos-2020-3>

to detect different construction-related equipment, such as trucks, tents, and excavators. Although, in this case, the location must be determined beforehand, our framework can keep track of the construction site proportions, which in turn can indicate the magnitude of the outbreak in that location.

Adaptations of the proposed framework could measure many other indicators. For instance, counting trucks on roads and highways, trains on railways and stations, or containers in dry ports are all possibilities of economic indicators that can also point out problems in the supply chain. Observing tractors in rural areas can help to evaluate agriculture activity, the same way the volume of vehicles in factories can assess the level of industrial activity. We are currently looking for high-resolution collection of satellite images over time to evaluate our framework under real, continual circumstances. To this end, we are seeking partnerships with the industry and government agencies to gain access to such data. With that, hopefully, we will be able to provide a flexible tool that can be explored by authorities during this COVID-19 outbreak or in future events demanding a similar awareness over human activities.

## VI. ACKNOWLEDGMENTS

Part of the equipment used in this project are supported by a grant (CNS-1513126) from the USA National Science Foundation. Internal funding support from USF Research and Innovation to the Institute for Artificial Intelligence (AI+X) is also acknowledged. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research. The authors would like to thank also the research Brazilian agencies CNPq, CAPES and FAPESP.

## REFERENCES

- [1] S. Asadzadeh and C. R. de Souza Filho, "Investigating the capability of worldview-3 superspectral data for direct hydrocarbon detection," *Remote Sensing of Environment*, vol. 173, pp. 162 – 173, 2016.
- [2] D. Lam, R. Kuzma, K. McGee, S. Dooley, M. Laielli, M. Klaric, Y. Bulatov, and B. McCord, "xView: Objects in Context in Overhead Imagery," *CoRR*, 2018.
- [3] G. Christie, N. Fendley, J. Wilson, and R. Mukherjee, "Functional Map of the World," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 6172–6180.
- [4] A. V. Etten, D. Lindenbaum, and T. M. Bacastow, "Spacenet: A remote sensing dataset and challenge series," *CoRR*, vol. abs/1807.01232, 2018. [Online]. Available: <http://arxiv.org/abs/1807.01232>
- [5] R. Gupta, B. Goodman, N. Patel, R. Hosfelt, S. Sajeev, E. Heim, J. Doshi, K. Luscas, H. Choset, and M. Gaston, "Creating xBD: A dataset for assessing building damage from satellite imagery," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 10–17.
- [6] D. J. Rogers, "Satellite imagery, tsetse and trypanosomiasis in Africa," *Preventive Veterinary Medicine*, vol. 11, no. 3, pp. 201 – 220, 1991.
- [7] D. J. Rogers, S. E. Randolph, R. W. Snow, and S. I. Hay, "Satellite imagery in the study and forecast of malaria," *Nature*, vol. 415, no. 6872, pp. 710 – 715, 2002.
- [8] S. W. Dister, D. Fish, S. M. Bros, D. H. Frank, and B. L. Wood, "Landscape characterization of peridomestic risk for lyme disease using satellite imagery," *The American Journal of Tropical Medicine and Hygiene*, vol. 57, no. 6, pp. 687–692, 1997.
- [9] D. Cyranoski, "Putting china's wetlands on the map," *Nature*, vol. 458, no. 134, 2009.
- [10] T. E. Ford, R. R. Colwell, J. B. Rose, S. S. Morse, D. J. Rogers, and T. L. Yates, "Using satellite images of environmental changes to predict infectious disease outbreaks," *Emerging infectious diseases*, vol. 15, no. 9, pp. 1341–1346, 2009.
- [11] R. Garni, A. Tran, H. Guis, T. Baldet, K. Benallal, S. Boubidi, and Z. Harrat, "Remote sensing, land cover changes, and vector-borne diseases: Use of high spatial resolution satellite imagery to map the risk of occurrence of cutaneous leishmaniasis in ghardaia, algeria," *Infection, Genetics and Evolution*, vol. 28, pp. 725 – 734, 2014.
- [12] G. Tilottama, R. Powell, C. Elvidge, K. Baugh, and P. Sutton, "Shedding light on the global distribution of economic activity," *The Open Geography Journal*, vol. 3, 01 2010.
- [13] Y. Hu and J. Yao, "Illuminating economic growth," *IMF Working Papers*, vol. 19, p. 1, 01 2019.
- [14] C. D. Elvidge, P. C. Sutton, T. Ghosh, B. T. Tuttle, K. E. Baugh, B. Bhaduri, and E. Bright, "A global poverty map derived from satellite data," *Computers and Geosciences*, vol. 35, no. 8, pp. 1652–1660, 2009.
- [15] A. M. Noor, V. A. Alegana, P. W. Gething, A. J. Tatem, and R. W. Snow, "Using remotely sensed night-time light as a proxy for poverty in africa," *Population Health Metrics*, vol. 6, no. 5, 2008.
- [16] G. Tilottama, C. Elvidge, P. Sutton, K. Baugh, and D. Ziskin, "Estimating the information and technology development index (idi) using nighttime satellite imagery," *Proceedings of the Asia-Pacific Advanced Network*, vol. 30, 2010.
- [17] N. Jean, M. Burke, M. Xie, W. M. Davis, D. B. Lobell, and S. Ermon, "Combining satellite imagery and machine learning to predict poverty," *Science*, vol. 353, no. 6301, pp. 790–794, 2016.
- [18] P. K. Suraj, A. Gupta, M. Sharma, S. B. Paul, and S. Banerjee, "On monitoring development indicators using high resolution satellite images," 2017.
- [19] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 5 2015.
- [20] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36, pp. 193–202, 1980.
- [21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.
- [22] M. Kampffmeyer, A.-B. Salberg, and R. Jenssen, "Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2016, pp. 1–9.
- [23] E. Guirado, S. Tabik, M. Rivas, A. Domingo, and F. Herrera, "Whale counting in satellite and aerial images with deep learning," *Scientific Reports*, vol. 9, no. 1, 2019.
- [24] R. Minetto, M. Pamplona Segundo, and S. Sarkar, "Hydra: An ensemble of convolutional neural networks for geospatial land classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6530–6541, 2019.
- [25] S. Lei, Z. Shi, and Z. Zou, "Super-resolution for remote sensing images via local-global combined network," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 8, pp. 1243–1247, 2017.
- [26] S. A. Kulp and B. H. Strauss, "New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding," *Nature communications*, vol. 10, no. 1, 2019.
- [27] Q. Zhang, Q. Yuan, C. Zeng, X. Li, and Y. Wei, "Missing data reconstruction in remote sensing image with a unified spatial-temporal-spectral deep convolutional neural network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4274–4288, 2018.
- [28] G. Rotich, S. Aakur, R. Minetto, M. P. Segundo, and S. Sarkar, "Using semantic relationships among objects for geospatial land use classification," in *IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, 2018, pp. 1–7.
- [29] M. J. Leotta, C. Long, B. Jacquet, M. Zins, D. Lipsa, J. Shan, B. Xu, Z. Li, X. Zhang, S.-F. Chang, M. Purri, J. Xue, and K. Dana, "Urban semantic 3d reconstruction from multiview satellite imagery," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2019, pp. 1–10.
- [30] Y. Zhan, D. Hu, Y. Wang, and X. Yu, "Semisupervised hyperspectral image classification based on generative adversarial networks," *IEEE*

*Geoscience and Remote Sensing Letters*, vol. 15, no. 2, pp. 212–216, 2018.

- [31] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, “SSD: Single Shot Multibox Detector,” in *European Conference on Computer Vision (ECCV)*. Springer, 2016, pp. 21–37.
- [32] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, “Object detection with discriminatively trained part-based models,” *IEEE Transactions Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 32, no. 9, pp. 1627–1645, Sep. 2010.
- [33] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications*. Berlin, Heidelberg: Springer-Verlag, 2005.
- [34] J. a. Gama, I. Žliobaitūn defined, A. Bifet, M. Pechenizkiy, and A. Bouchachia, “A survey on concept drift adaptation,” *ACM Comput. Surv.*, vol. 46, no. 4, Mar. 2014. [Online]. Available: <https://doi.org/10.1145/2523813>
- [35] P. Henderson and V. Ferrari, “End-to-end training of object class detectors for mean average precision,” *CoRR*, vol. abs/1607.03476, 2016. [Online]. Available: <http://arxiv.org/abs/1607.03476>
- [36] H. Li, L. Jing, Y. Tang, Q. Liu, H. Ding, Z. Sun, and Y. Chen, “Assessment of pan-sharpening methods applied to worldview-2 image fusion,” in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 3302–3305.