Measuring Human and Economic Activity From Satellite Imagery to Support City-Scale Decision-Making During COVID-19 Pandemic

Rodrigo Minetto, *Member, IEEE*, Maurício Pamplona Segundo, *Member, IEEE*, Gilbert Rotich, and Sudeep Sarkar, *Fellow, IEEE*

Abstract—The COVID-19 outbreak forced governments worldwide to impose lockdowns and quarantines 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. In this article, 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 that could be used to 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 illustrate different measurable indicators. Our code and annotated high-resolution aerial scenes before and after the outbreak are available on GitHub.[1]

Index Terms—Remote sensing, CNN-based object detection, human and economic activity assessment, COVID-19 pandemic

---

1 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.[2] Such images will support decisions concerning: (1) traffic issues, to ensure citizens' mobility but at the same time to avoid traffic jams that block the exchange of essential supplies; (2) 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; (3) facilities activity, to safely and economically maximize resources; and (4) 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 collect panchromatic images daily 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. Fig. 1 presents two scenarios in which the analysis of strategic sites can provide critical indicators of human and economic activities over time. Fig. 1a shows parked aircraft before and after the COVID-19 outbreak, while Fig. 1b shows a car rental parking lot within a similar time frame. These examples illustrate the decrease in traveling caused by this pandemic and its consequential impact on aviation and car rental businesses. Other examples include obtaining information on traffic or distancing issues through detecting and classifying vehicles; keeping track of new medical infrastructure being built by identifying construction elements such as bulldozers, excavators, trucks, and tents; and measuring economic activity by detecting commercial transports such as planes, ships, and locomotives.

Although many indicators computed through remote sensing are also measurable from other data sources, the former is advantageous for its versatility. Monitoring systems based on satellite images can support new indicators and new areas of interest with little effort, as all of them share the same database and input format. Besides, scaling these systems up to a global level is a matter of satellite coverage and computational power, eliminating complications associated

---

1. https://github.com/maups/covid19-satellite-analysis
2. https://www.euspaceimaging.com/eu-commission-asks-eo-community-for-help-with-covid-19/

- *Rodrigo Minetto is with the Universidade Tecnológica Federal do Paraná (UTFPR), Curitiba 80230-901, Brazil. E-mail: rodrigo.minetto@gmail.com.*
- *Maurício P. Segundo, Gilbert Rotich, and Sudeep Sarkar are with the Department of Computer Science and Engineering, University of South Florida, Tampa, FL 33620 USA. E-mail: {mauriciop, grotich, sarkar}@usf.edu.*

(a) Airport before and after the COVID-19

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

Fig. 1. COVID-19 impacts on human and economic activities. *Photo credit*: Satellite image 2020 Maxar Technologies.

with data collection from heterogeneous sources at this scale. These characteristics favor the adaptation and application of such systems when a fast response is critical.

Nevertheless, to unleash the full potential use of 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. As our main contribution, 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 World-View-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. Furthermore, we manually annotated nearly 16,500 objects from high-resolution aerial scenes before and after the COVID-19 outbreak. We made them publicly available in our github repository hoping that they will be useful to other researchers addressing in the same problem.

We present our framework within the context of stay-at-home order enforcement (Section 3) and discuss later how to adapt it to other scenarios (Section 4.3). In our experiments, we first evaluate the detection performance on the xView dataset (Section 4.1) and then show its potential for temporal analysis using the US IARPA Functional Map of the World (fMoW) dataset [3] (Section 4.2). Finally, we show our framework in action on real examples of world scenes before and after the COVID-19 outbreak (Section 4.3).

2 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.

The rise in the number of large-scale challenge datasets that has recently become available to foster breakthroughs in this field has been 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, wildfires, 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 trypanosomiases 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]. Other relevant efforts in this

(a) Area of interest   (b) Location sampling   (c) Location boundaries   (d) ROI extraction

(h) Compliance evaluation   (g) Filter & merge   (f) Vehicle detection   (e) Resize & split
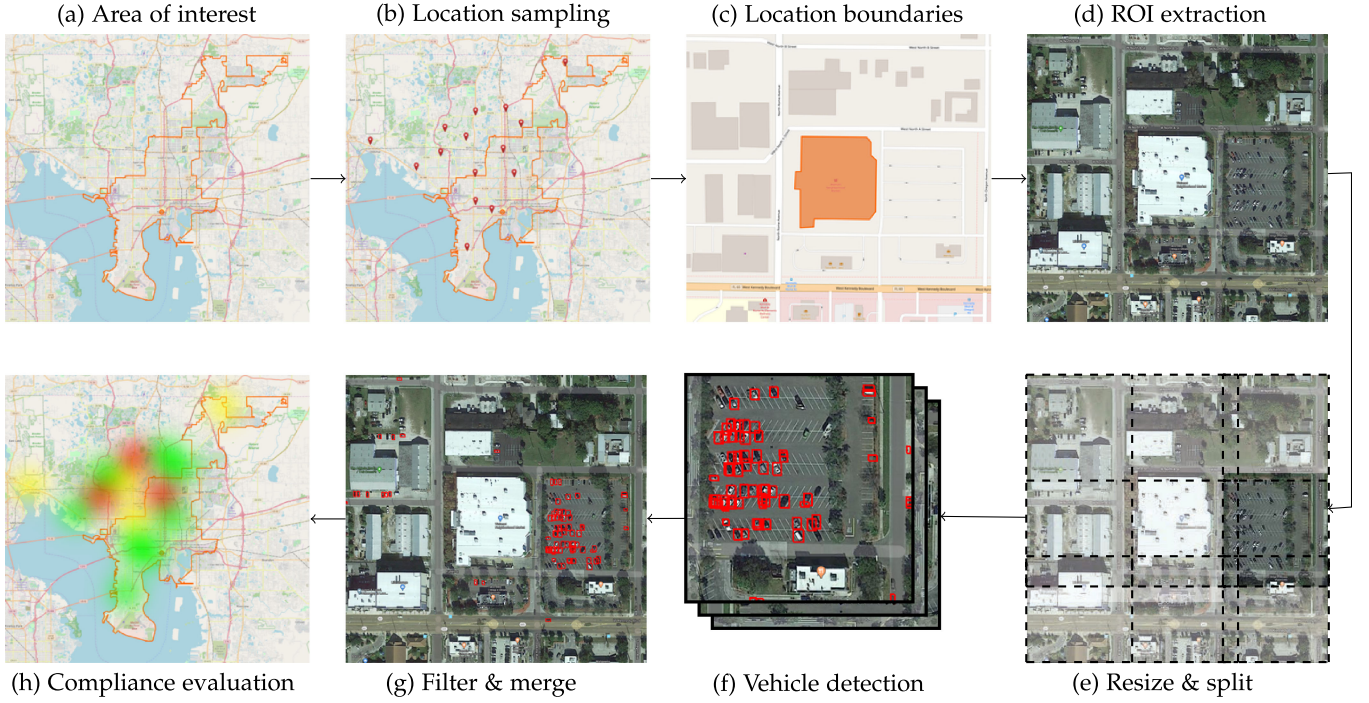
Fig. 2. We present a workflow to analyze the pattern of vehicles over time to monitor compliance with stay-at-home order. Similar workflows are possible for other aspects of the economy, such as supply-chain disruptions. Other than the first step, the rest is fully automatic. (a) A human demarcates an 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[*]. Satellite images were obtained from Google Maps[**].

[*]OpenStreetMap contributors   [**] Image © 2020 Google, Maxar Technologies

field include estimating asset wealth across thousands of African villages from publicly-available multispectral satellite imagery [20] and predicting key food security metrics such as z-scores of stunting or wasting [21].

The bio-inspired CNN [22], 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 [23] has made it a perfect tool for remote sensing. It currently supports a myriad of applications in the literature, such as semantic segmentation [24], target localization [25], region classification [26], [27], image retrieval [28], super-resolution [29], regression models for environmental knowledge extraction [30], understanding of temporal and spatial variations [31], study of semantic relationships between aerial targets [32], 3D reconstruction [33], and hyperspectral image generation [34]. 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.

3 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,[3] satellite images, and a machine learning-based vehicle detector. Fig. 2 illustrates the sequence of stages that compose our proposed framework. These stages are detailed in the following sections.

3.1 Location Sampling and Region of Interest Extraction

The first stage in our frameworks 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

3. https://www.openstreetmap.org

TABLE 1
Parameters of Our Ensemble of SSDs for All
Classes in the xView Dataset

	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

Detectors #1-#3 are used for small objects; detectors #1-#4 for medium objects; and detectors #3-#5 for large objects.
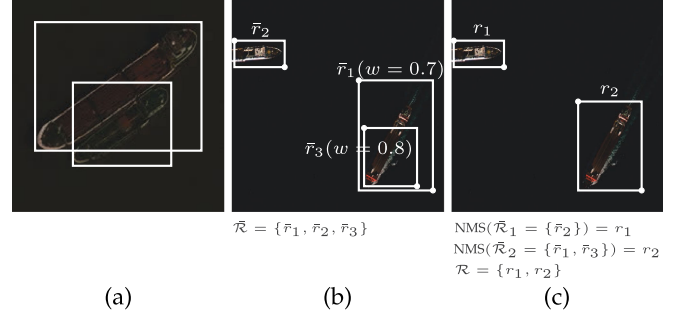

Fig. 3. Example of (a) two overlapped regions from the same category that must not be merged, and of (b) three detected regions in which (c) two of them were merged using their confidence score to define the new bounding box dimensions.

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 Fig. 2b), we select the ones that contain annotations for the boundary contour (see Fig. 2c) 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 region of interest (ROI) in a satellite image (see Fig. 2d). Smarter ROI extraction strategies can be used, such as parking lot detection near strategic locations, upon the availability of reliable techniques and resources to support them.

3.2 Vehicle Detection

The input for this stage is a group of ROI images extracted in the previous stage, and the output for the ith ROI is a set of n_i detected regions $\mathcal{R}^i = \{r_1^i, r_2^i, \ldots, r_{n_i}^i\}$. 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) [35] 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 300×300 images as input. So we split our ROI images into blocks of 300×300 pixels without overlap whenever possible for the *Vanilla* model, as illustrated in Fig. 2e, 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 1 summarizes this arrangement.

Each detector in the ensemble runs separately on each block of its input image (see Fig. 2f), 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 1 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) [36] algorithm with minor adaptations is used to discard duplicate regions belong to the same object. Consider $\bar{\mathcal{R}} = \{\bar{r}_1, \bar{r}_2, \ldots\}$ 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 σ:

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

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 Fig. 3a). 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:

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

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 Figs. 3b and 3c. $\bar{\mathcal{R}}_k$ is then removed from $\bar{\mathcal{R}}$ and the process is repeated until $\bar{\mathcal{R}}$ is empty.

An example of the final output for a ROI is illustrated in Fig. 2g. This module of our framework can be updated in the future to incorporate recent state-of-the-art solutions that address some recurring challenges in satellite data, such as class unbalancing [37], scale variations [38], and unrealistic false alarms [39], as a way to increase the detection accuracy or reduce the ensemble size.

3.3 Temporal Analysis

Given enough time and acquisition frequency, it is possible to apply time series analysis techniques [40] to learn the
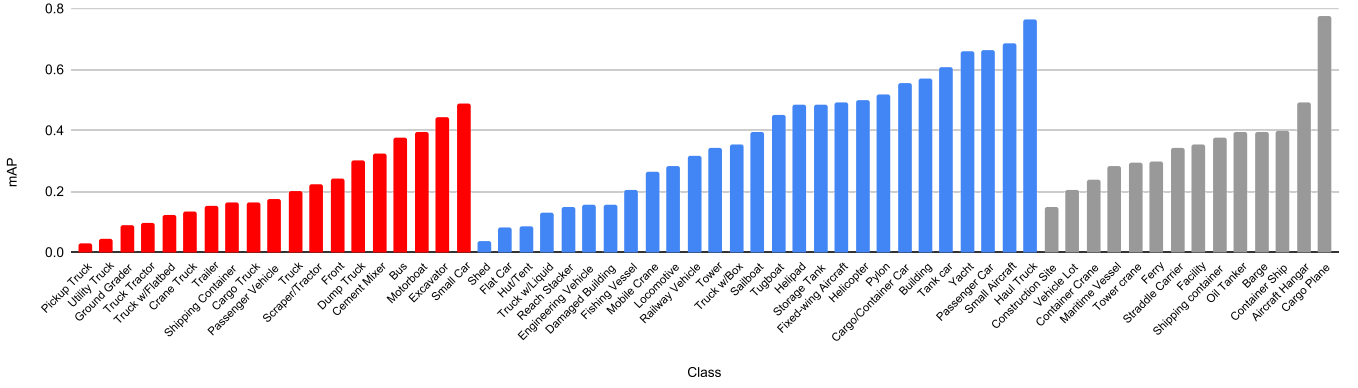
Fig. 4. Mean average precision per xView class. Red, blue, and gray bars represent small, medium and large targets, respectively.

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 [41] can help to identify a change in behavior while avoiding outlier data. Depending on the amount of data available for training, one can also explore the use of recurrent neural networks for time-series forecasting [42], [43] followed by the detection of abnormal behaviors. However, this module has not been implemented yet. We need access to appropriate data to study these variations.

4 EXPERIMENTS

4.1 Detector Evaluation

To evaluate the detection stage, we used the xView dataset [2]. It contains a training set with 847 high-resolution images (0.3 GSD) 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 2564×2576 to 3187×4994 pixels. The interpolated mean average precision (mAP), detailed by Henderson and Ferrari [44], 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 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 Fig. 4. The ensemble configuration described in Section 3.2 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 1, 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 5), we report detection results for them as well.

In Fig. 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 2), evidencing the potential of the approach for detecting targets in satellite images. Our framework can process up to ten image blocks per second on a modern GPU, which allows updating the object count of thousands of ROIs per hour. This pace is more than enough to handle city-scale applications, even if image acquisition occurs on a daily basis.

4.2 Temporal Evaluation

The 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 [45], 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, weather, and vehicles in the scene, the last being our primary focus.

Each region in the dataset represents one of the 63 categories, including a false detection category that aggregates different types of regions that do not fit into the other 62 defined categories. Among those classes, we are particularly interested in one: 'parking_lot_or_garage'. We looked for regions of this class that:

TABLE 2
Final xView Leaderboard: mAP Per Size Group and Overall

| Rank | mAP | | | |
Country	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

- have three or more samples with GSD smaller than 0.4 and dimensions greater than 300×300 pixels (one block in Section 3.2); and
- show an open-air parking lot (as '*parking_lot_or_garage*' includes closed garage buildings).

Following these criteria, we ended up with nine regions with three to nine images each. To improve the detection outcome in the regions with lower resolution, we manually upsampled them so that their objects' sizes looked closer to how they were supposed to look on xView images (i.e., GSD ≈ 0.3). We also increased the confidence threshold to 0.25 to eliminate false positives caused by the upsampling operation (e.g., edge blurring and noise amplification). We ran our vehicle detector using each of these regions as our ROI, and even though there are false positives and negatives in nearly all regions, the count is consistent enough for further automated analyses. 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 to perceive gradual changes in the flow of vehicles. Three regions with small, medium and high variation in the number of vehicles are respectively shown in Figs. 5a, 5b, and 5c. One can argue if the sorting is correct or not for some of the samples, but the overall quality of the process is evident.

Fig. 5a shows the region with the lowest variation in the flow of vehicles. As can be observed, we can identify regions with a stable count even when there is a high volume of cars. This ability is essential for stay-at-home enforcement when a decrease is expected but is not confirmed, requiring further action from the authorities (e.g., suspension of activities in public, social, and private sectors). Figs. 5b and 5c 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. 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 Fig. 5c).

We manually annotated more than 2,000 cars in the images shown in Fig. 5 using an open-source tool[4] to compare our detection results with the ground truth. Each annotation is an axis-aligned rectangle delimited by its top-left and bottom-right corners and categorized as one of the 60 xView object classes. Our detector achieved a 0.59 mAP for small cars in these images, which is on par with the accuracy on xView. The Mean Absolute Percentage Error (MAPE) in car counting for images with more than 100 annotations is approximately 15 percent, which indicates that the number of detections and annotations are relatively close to each other. Although these count values have a larger deviation in some regions (see Fig. 5a), the error tends to be similar in images of the same region. As a result, we can estimate the amount of change in the number of vehicles

accurately. With this information, we can devise indicators for decision-makers using different ROI groups and their expected behavior (e.g., a stable count in supermarkets, a decrease in schools, an increase in rental car facilities). Besides that, with proper temporal sequences of images from local businesses' parking lots, we could quantify the impact of COVID-19 on their earnings [46]. Finally, when ROIs have geographic coordinates, it is possible to interpolate these estimates to neighboring areas and produce heatmaps, as illustrated in Fig. 2h.

4.3 COVID-19 Case Studies

The results presented so far 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. Many of these images had artifacts that impair automated analysis, like watermarks and low-resolution. We followed the same procedure employed for fMoW images, and manually upsampled images so that the GSD is approximately 0.3 and increased the confidence threshold to 0.25.

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 Fig. 4 for a complete list. For instance, Fig. 6a shows an image from North Korean commercial vessels used to transport coal and other commodities. According to Christoph Koettl from the New York Times,[5] they 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 percent of the world's trade.[6]

In Fig. 6b, 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.[7] We were able 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. In Fig. 6c, 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 (over 60 vehicles[8]) and could help authorities to measure attendance over time in health facilities.

---

4. https://github.com/tzutalin/labelImg

5. https://www.nytimes.com/2020/03/26/video/coronavirus-north-korea.html
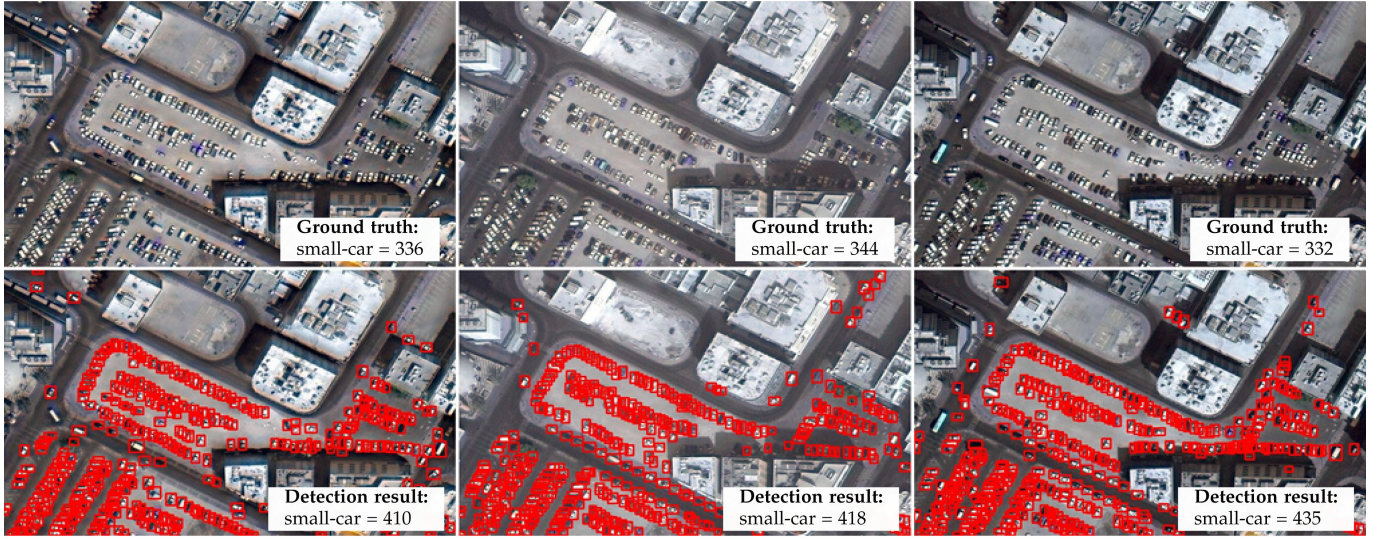
6. https://business.un.org/en/entities/13

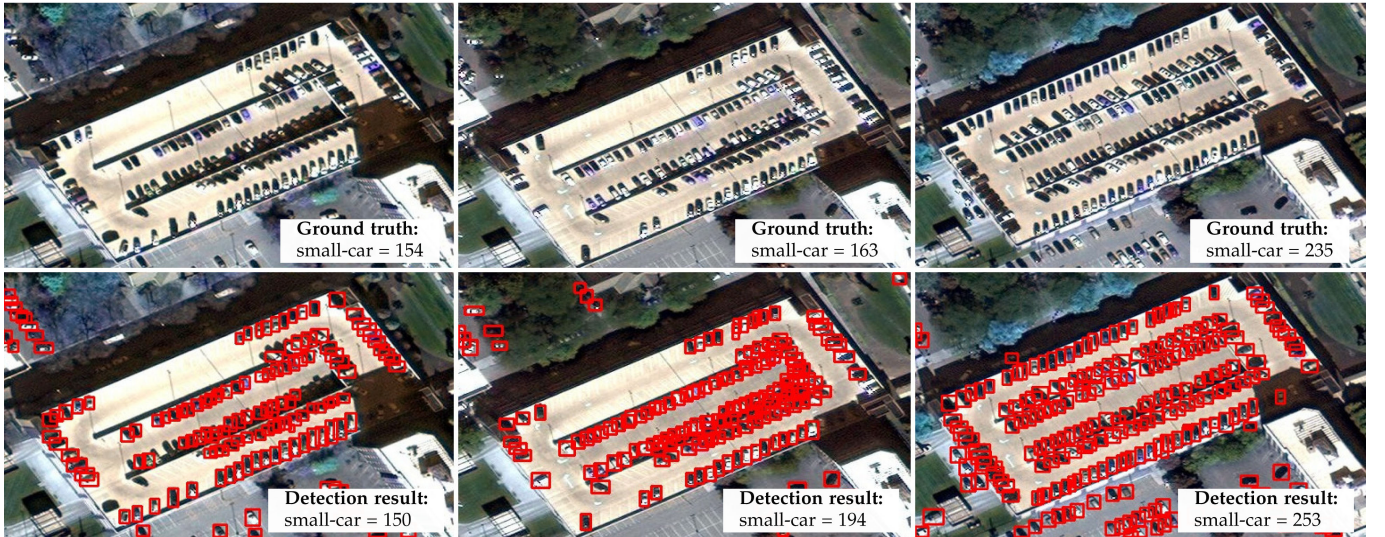7. https://www.businessinsider.com/coronavirus-covid-19-russian-hospital-field-near-moscow-satellite-photos-2020-3

8. https://www.gim-international.com/content/news/satellite-imagery-covid-19-testing-facilities-in-munich-germany
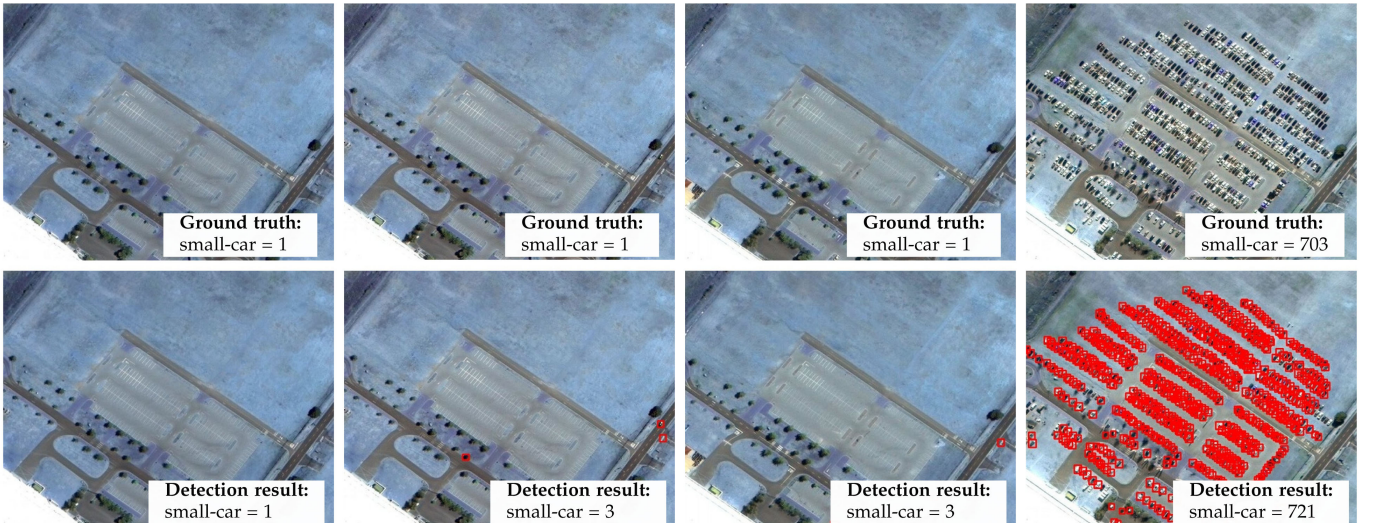
(a) The difference between the lowest and the highest vehicle count is 12 in the ground truth and 25 in our count

(b) The difference between the lowest and the highest vehicle count is 81 in the ground truth and 103 in our count

(c) The difference between the lowest and the highest vehicle count is 702 in the ground truth and 720 in our count

Fig. 5. Detection results for multiple samples of a same region with a (a) small, (b) medium, and (c) 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.
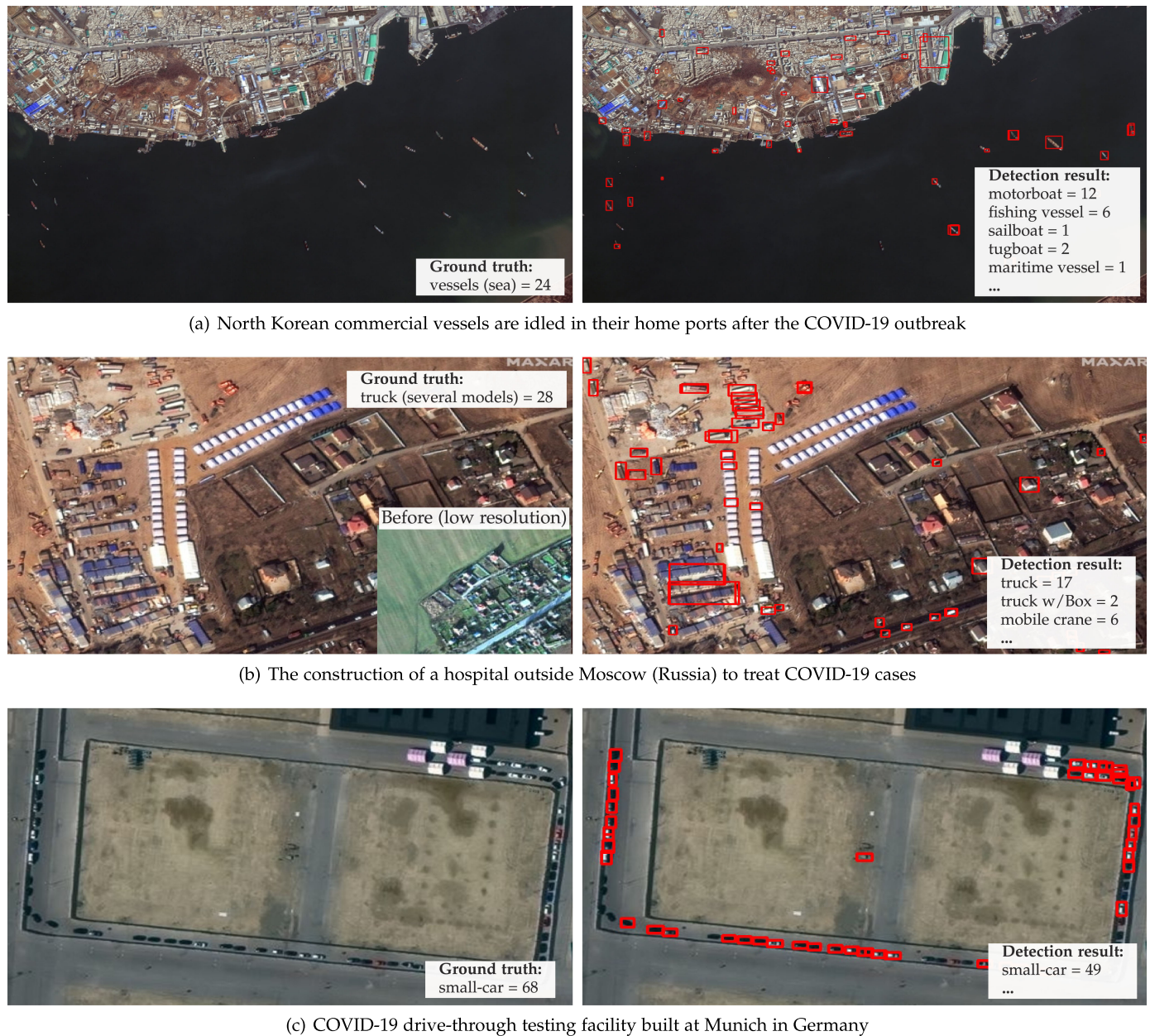
(a) North Korean commercial vessels are idled in their home ports after the COVID-19 outbreak

(b) The construction of a hospital outside Moscow (Russia) to treat COVID-19 cases

(c) COVID-19 drive-through testing facility built at Munich in Germany

Fig. 6. Satellite imagery from world scenes related to the COVID-19 pandemic and statistics about vehicles/infrastructure available. *Photo credit:* SATELLITE IMAGE 2020 MAXAR TECHNOLOGIES.

For some rare scenes, we were able to find the images before and after the pandemic. Thus, it was possible to use the temporal information to illustrate the economic effect of the outbreak. For example, Fig. 7 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.

Fig. 8 presents another example of a plausible economic indicator. It shows the number of planes in activity at the Salt Lake City International Airport (USA) before and after the COVID-19 outbreak, indicating that the pandemic has dramatically impacted the airplane

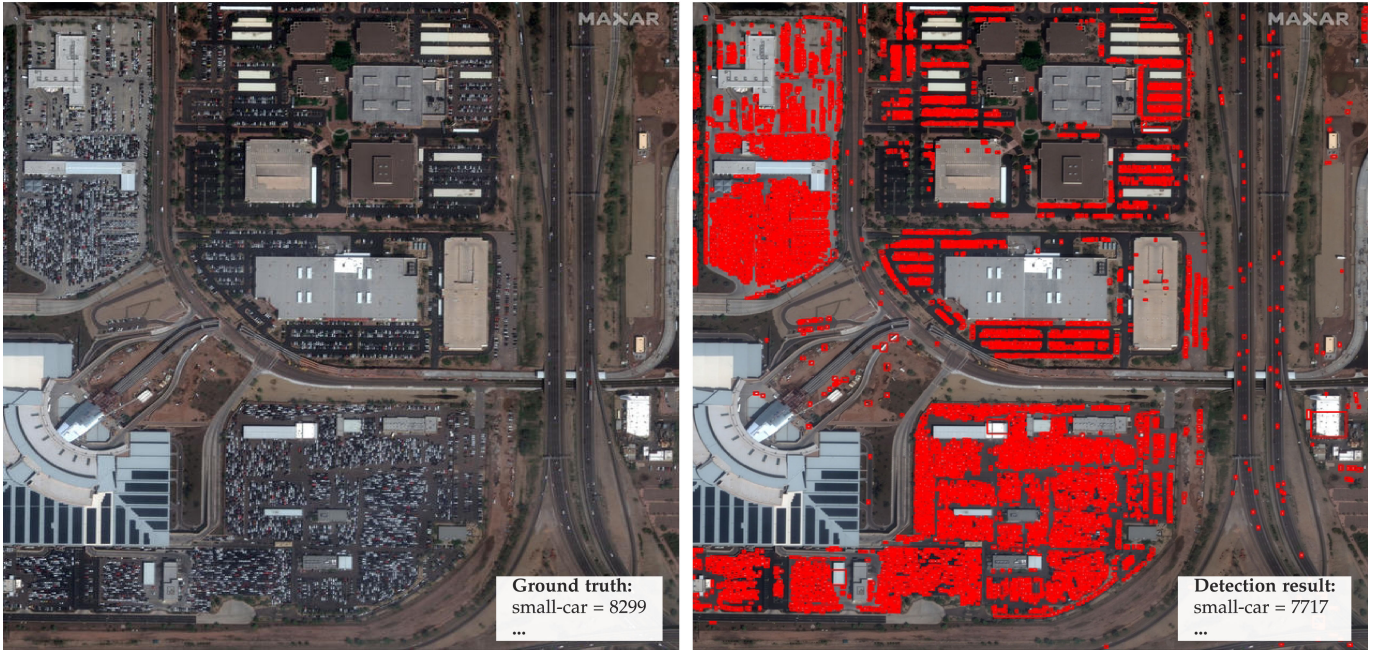travel segment. Our framework can automatically detect this decrease.

Fig. 9 shows an interesting aspect. Before the pandemic it is possible to see a golf course (left) and a supermarket (top right) and their respective parking lots. The same location after the pandemic shows that the golf course parking lot is almost empty while the supermarket parking lot had a much smaller change in the number of cars. This information could be used to indicate the compliance to stay-at-home orders and the occurrence of panic buying.

Finally, Fig. 10 shows car lines in a tollbooth at Wuhan, China, before the COVID-19 outbreak, and 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.

Ground truth:
small-car = 5336
...

Detection result:
small-car = 5989
...
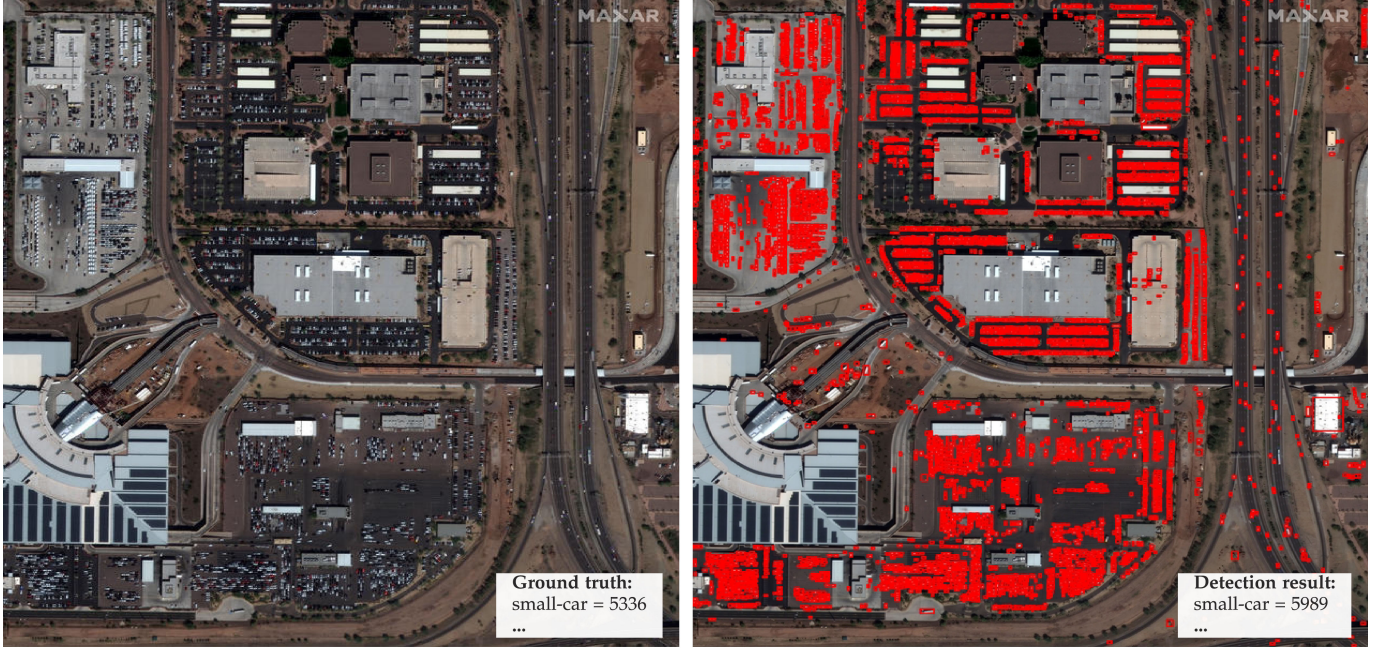
(a) Phoenix airport (USA) parking-lot before COVID-19 outbreak

Ground truth:
small-car = 8299
...

Detection result:
small-car = 7717
...

(b) Phoenix airport (USA) parking-lot after COVID-19 outbreak

Fig. 7. Satellite imagery from world scenes before and after the COVID-19 outbreak and statistics about vehicles/infrastructure available. *Photo credit*: Satellite image 2020 Maxar Technologies.

We annotated more than 16,000 vehicles on the images shown in Figs. 7, 8, 9, and 10 for performance analysis purposes. The detector achieved a 0.94 mAP for airplanes, 0.66 mAP for small cars, and a 10 percent MAPE in object counting for images with more than 100 annotations. These results were slightly better than the ones obtained for the fMoW images in Section 4.2. This outcome, allied to the fact that we were able to automatically perceive the impact of COVID-19 in real scenes (e.g., decrease in aircraft on airport gates, increase in vehicles available for rental, small changes in parking occupancy of grocery

stores), reinforces the proposed framework's aptness to the intended application.

5   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 3), but it can be used for different purposes, as we discussed in Section 4.3. By using two publicly available datasets, xView and fMoW,

(a) Salt Lake City International Airport (USA) before COVID-19 outbreak

(b) Salt Lake City International Airport (USA) after COVID-19 outbreak
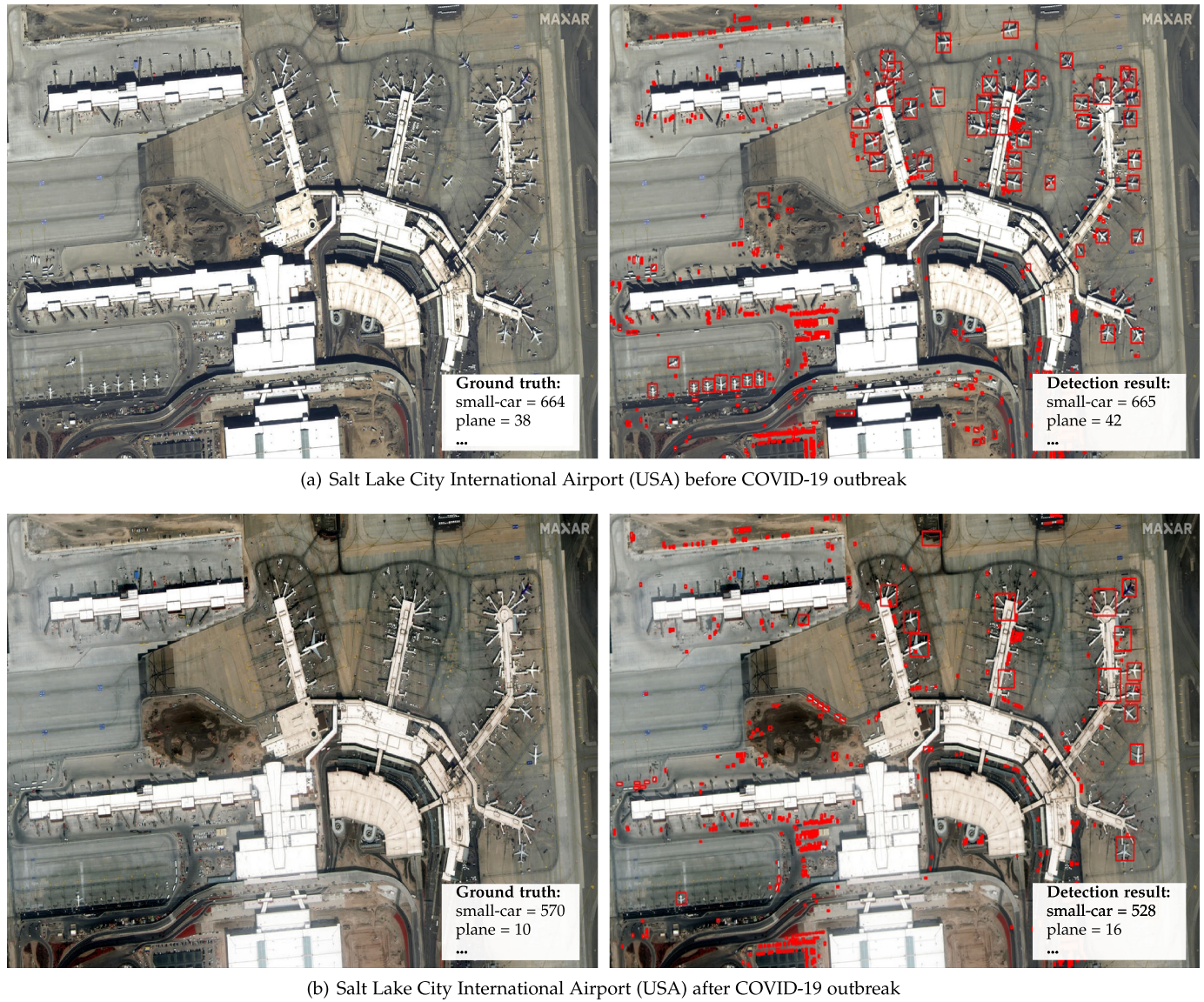
Fig. 8. Satellite imagery from world scenes before and after the COVID-19 outbreak and statistics about vehicles/infrastructure available. *Photo credit*: SATELLITE IMAGE 2020 MAXAR TECHNOLOGIES.

we estimated how well our detector worked for 60 distinct classes of objects (Section 4.1) and exemplified how well we can perceive the flow of these objects over time (Section 4.2).

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 evaluate agriculture activity, the same way the volume of vehicles in factories can assess industrial activity. The pool of indicators varies according to the local interests of each city. At this point, we can identify changes in the volume of objects using a small set of images and possibly classify these changes in different orders of magnitude. This information by itself could compose a dashboard to help decision-makers in spotting situations that demand immediate attention.

While human activities can be categorized in qualitative terms (e.g., low, normal, high), economic activities are better represented as continuous values. However,

any attempt to model and forecast changes in consolidated indicators (e.g., tourism revenue, local businesses' earnings, real estate vacancy rate) requires longer satellite image sequences with a high revisit rate and a history of indicator values. We are currently looking for a 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.

It is worth noting that a framework based on satellite images has different limitations. The first one relates to the high cost of the data, especially when high-resolution images and high-frequency acquisition are required. Second, satellites hardly capture object dynamics within a day and cannot detect objects at night. Moreover, object visibility is affected by different factors, such as weather
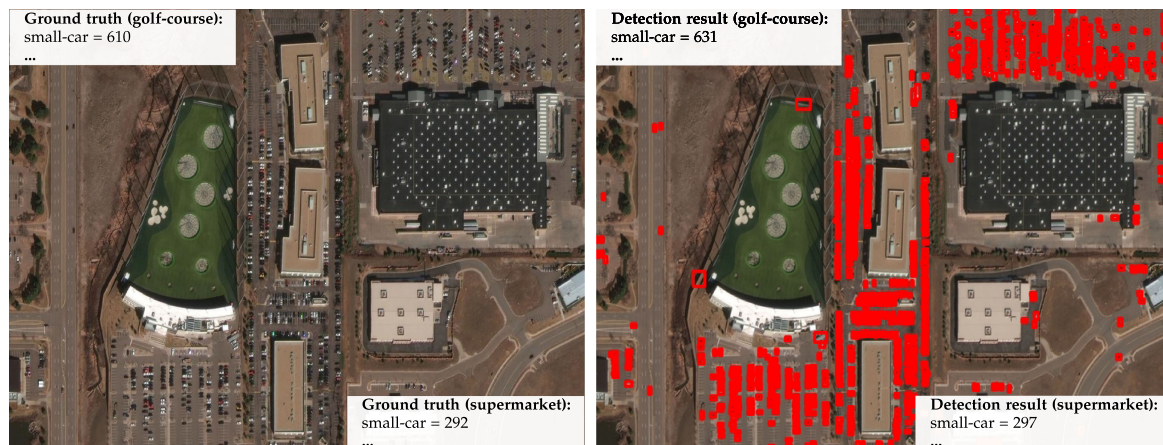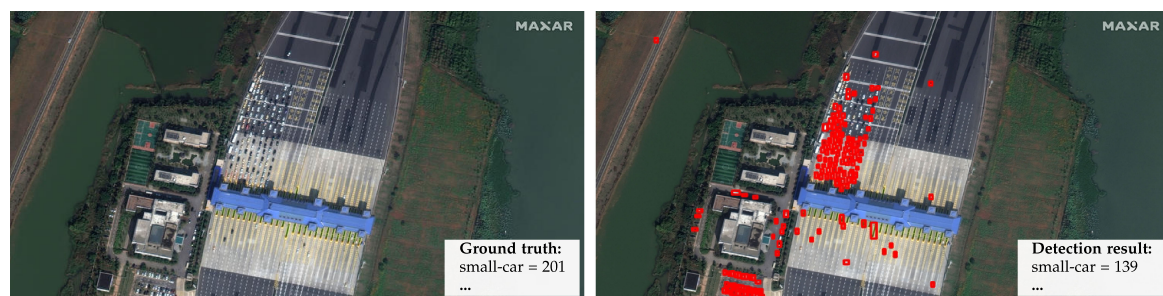
(a) Golf course and a supermarket in Colorado (USA) before COVID-19 outbreak
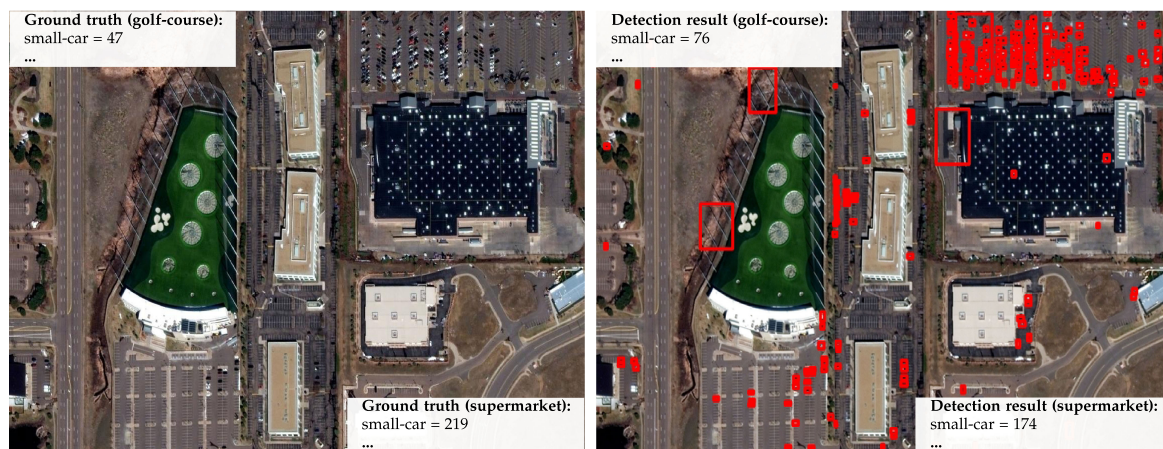
(b) Golf course and a supermarket in Colorado (USA) after COVID-19 outbreak

Fig. 9. Satellite imagery from world scenes before and after the COVID-19 outbreak and statistics about vehicles/infrastructure available. *Photo credit*: Satellite image 2020 Maxar Technologies.

(a) Wuhan (China) toll plaza before COVID-19 outbreak

(b) Wuhan (China) toll plaza after COVID-19 outbreak

Fig. 10. Satellite imagery from world scenes before and after the COVID-19 outbreak and statistics about vehicles/infrastructure available. *Photo credit*: Satellite image 2020 Maxar Technologies.

conditions (cloud cover) and satellite point of view (occlusions caused by trees and buildings). Finally, vehicle-based indicators cannot use ROIs with enclosed parking garages. Depending on the application's accuracy requirements, it may be necessary to combine satellite data with other data sources, such as mobile phone data, GPS signals, tollbooth records, etc. On the other hand, these alternative sources also have their limitations. For instance, smartphone data may not cover all models, operating systems, and carriers; GPS signals are affected by weather and buildings; most ROIs may not have a tollbooth nearby. Additionally, many developing nations may not have the infrastructure for aggregating these sources. Thus, in scenarios where a fast response is critical, satellite data stands out for its portability and coverage.

ACKNOWLEDGMENTS

## REFERENCES

[1] S. Asadzadeh and C. R. S. Filho, "Investigating the capability of worldview-3 superspectral data for direct hydrocarbon detection," *Remote Sens. Environ.*, vol. 173, pp. 162–173, 2016.

[2] D. Lam *et al.*, "xView: Objects in context in overhead imagery," vol. 1, pp. 1–16, 2018.

[3] G. Christie, N. Fendley, J. Wilson, and R. Mukherjee, "Functional map of the world," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 6172–6180.

[4] A. V. Etten, D. Lindenbaum, and T. M. Bacastow, "Spacenet: A remote sensing dataset and challenge series," 2018, *arXiv:1807.01232*.

[5] R. Gupta *et al.*, "Creating xBD: A dataset for assessing building damage from satellite imagery," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 10–17.

[6] D. J. Rogers, "Satellite imagery, tsetse and trypanosomiasis in Africa," *Preventive Veterinary Med.*, 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," *Amer. J. Tropical Med. 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," *Emerg. Infectious Diseases*, vol. 15, no. 9, pp. 1341–1346, 2009.

[11] R. Garni *et al.*, "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 Evol.*, 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," *Open Geography J.*, vol. 3, pp. 147–160, 2010.

[13] Y. Hu and J. Yao, "Illuminating economic growth," *IMF Working Papers*, vol. 19, 2019, Art. no. 1.

[14] C. D. Elvidge *et al.*, "A global poverty map derived from satellite data," *Comput. Geoscie.*, 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, 2008, Art. no. 5.

[16] G. Tilottama, C. Elvidge, P. Sutton, K. Baugh, and D. Ziskin, "Estimating the information and technology development index (IDI) using nighttime satellite imagery," *Proc. Asia-Pacific Advanced Netw.*, vol. 30, 2010, pp. 143–171.

[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," *CoRR*, vol. abs/1712.02282, pp. 1–36, 2017.

[19] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[20] C. Yeh *et al.*, "Using publicly available satellite imagery and deep learning to understand economic well-being in africa," *Nat. Commun.*, vol. 11, 2020, Art. no. 2583.

[21] S. Ganguli, J. Dunnmon, and D. Hau, "Predicting food security outcomes using convolutional neural networks (CNNs) for satellite tasking," 2019, *arXiv:1902.05433*.

[22] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biol. Cybern.*, vol. 36, pp. 193–202, 1980.

[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Advances Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[24] 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 *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2016, pp. 1–9.

[25] E. Guirado, S. Tabik, M. Rivas, A. Domingo, and F. Herrera, "Whale counting in satellite and aerial images with deep learning," *Sci. Rep.*, vol. 9, 2019, Art. no. 14259.

[26] R. Minetto, M. P. Segundo, and S. Sarkar, "Hydra: An ensemble of convolutional neural networks for geospatial land classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6530–6541, Sep. 2019.

[27] F. Hu, G. Xia, W. Yang, and L. Zhang, "Mining deep semantic representations for scene classification of high-resolution remote sensing imagery," *IEEE Trans. Big Data*, vol. 6, no. 3, pp. 522–536, Sep. 2020.

[28] X. Tong, G. Xia, F. Hu, Y. Zhong, M. Datcu, and L. Zhang, "Exploiting deep features for remote sensing image retrieval: A systematic investigation," *IEEE Trans. Big Data*, vol. 6, no. 3, pp. 507–521, Sep. 2020.

[29] S. Lei, Z. Shi, and Z. Zou, "Super-resolution for remote sensing images via local–global combined network," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 8, pp. 1243–1247, Aug. 2017.

[30] S. A. Kulp and B. H. Strauss, "New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding," *Nat. Commun.*, vol. 10, 2019, Art. no. 4844.

[31] 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 Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4274–4288, Aug. 2018.

[32] G. Rotich, S. Aakur, R. Minetto, M. P. Segundo, and S. Sarkar, "Using semantic relationships among objects for geospatial land use classification," in *Proc. IEEE Appl. Imagery Pattern Recognit. Workshop*, 2018, pp. 1–7.

[33] M. J. Leotta *et al.*, "Urban semantic 3D reconstruction from multi-view satellite imagery," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2019, pp. 1–10.

[34] Y. Zhan, D. Hu, Y. Wang, and X. Yu, "Semisupervised hyperspectral image classification based on generative adversarial networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 2, pp. 212–216, Feb. 2018.

[35] W. Liu *et al.*, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 21–37.

[36] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, Sep. 2010.

[37] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 2, pp. 318–327, Feb. 2020.

[38] P. Zhou, B. Ni, C. Geng, J. Hu, and Y. Xu, "Scale-transferrable object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 528–537.

[39] H. Hu, J. Gu, Z. Zhang, J. Dai, and Y. Wei, "Relation networks for object detection," in *IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 3588–3597.

[40] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications*. Berlin, Germany: Springer, 2005.

[41] J. a. Gama, I. Žliobaitundefined, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A survey on concept drift adaptation," *ACM Comput. Surv.*, vol. 46, no. 4, 2014, Art. no. 44.

[42] G. Lai, W.-C. Chang, Y. Yang, and H. Liu, "Modeling long- and short-term temporal patterns with deep neural networks," in *Proc. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2018, pp. 95–104.

[43] Y. Qin, D. Song, H. Cheng, W. Cheng, G. Jiang, and G. W. Cottrell, "A dual-stage attention-based recurrent neural network for time series prediction," in *Proc. 26th Int. Joint Conf. Artif. Intell.*, 2017, pp. 2627–2633.

[44] P. Henderson and V. Ferrari, "End-to-end training of object class detectors for mean average precision," *CoRR*, vol. abs/1607.03476, pp. 1–15, 2016.

[45] H. Li *et al.*, "Assessment of pan-sharpening methods applied to worldview-2 image fusion," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2015, pp. 3302–3305.

[46] Z. Katona, M. Painter, P. N. Patatoukas, and J. Zeng, "On the capital market consequences of alternative data: Evidence from outer space," in *Proc. 9th Miami Behavioral Finance Conf.*, 2018, pp. 1–63.

Rodrigo Minetto received the PhD degree in computer science from the University of Campinas (UNICAMP), Brazil and Université Pierre et Marie Curie, France (UPMC), in 2012. He is currently an assistant professor with the Federal University of Technology - Paraná (UTFPR), Brazil. His research interests include image processing, computer vision and machine learning. Currently he is a visiting scholar with the University of South Florida (USF).
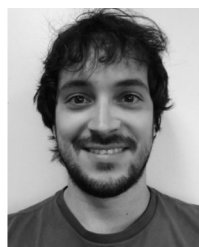
Maurício Pamplona Segundo received the BSc, MSc, and DSc degrees in computer science from the Federal University of Paraná (UFPR). He is currently a postdoctoral researcher with the Institute for Artificial Intelligence (AI+X), University of South Florida. His research interests include computer vision and pattern recognition, and his research interests include biometrics, remote sensing, 3D reconstruction, accessibility tools, and vision-based automation.

Gilbert Rotich received the BSc degree in computer engineering from Bethune-Cookman University. He is currently working toward the PhD degree at the University of South Florida. His research interests include computer vision, machine learning, deep learning, and remote sensing.

Sudeep Sarkar received the MS and PhD degrees in electrical engineering with University Presidential Fellowship, from The Ohio State University. He is currently a professor of computer science and engineering and an associate vice president for Research & Innovation with the University of South Florida, Tampa. He is the recipient of the National Science Foundation CAREER Award in 1994, the USF Teaching Incentive Program Award for Undergraduate Teaching Excellence in 1997, the Outstanding Undergraduate Teaching Award in 1998, and the Theodore and Venette Askounes-Ashford Distinguished Scholar Award, in 2004. He is a Fellow of the American Association for the Advancement of Science (AAAS), Institute of Electrical and Electronics Engineers (IEEE), American Institute for Medical and Biological Engineering (AIMBE), and International Association for Pattern Recognition (IAPR); and a charter member and member of the Board of Directors of the National Academy of Inventors (NAI). He has 25 year expertise in computer vision and pattern recognition algorithms and systems, holds three U.S. patents and has published high-impact journal and conference papers.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/csdl.