



EIT DIGITAL DOCTORAL SCHOOL
BUSINESS DEVELOPMENT EXPERIENCE

**Deep learning for counting cars in
satellite images - an AI application
to estimate the impact of COVID-19**

TELEFÓNICA - CHIEF DATA OFFICE

AUTHOR: Roberto Valle Fernández
SUPERVISORS: Richard Benjamins
Pedro Antonio de Alarcón

Madrid, 2020

Acknowledgments

- I would first like to thank my supervisors, Richard Benjamins and Pedro de Alarcón, for their valuable guidance during these six months. They provided me with the tools that I needed to choose the right direction and successfully complete the Business Development Experience (BDExp). Their insightful feedback pushed me to sharpen my thinking and brought my work to improve the conditions of our world.
- Particularly, I wish to express my deepest gratitude to Richard Benjamins for placing his trust and confidence in my abilities.
- Finally, I would also like to acknowledge the funding support received by the EIT Digital Doctoral School and its valuable courses in innovation and entrepreneurship (*i.e.*, “Raising Awareness”, “Opportunity Recognition”, “Business Development”, “Business Modelling” and “Growth and Harvest”).

Thank you with all of my heart.

Abstract

In this work, we deal with the problem of analyzing data acquired from high-resolution satellite imagery to provide accurate insights, trends and predictive signals for humanitarian purposes. Motivated by the fact that Coronavirus Disease (COVID-19) pandemic has caused worldwide turmoil in a short period of time since December 2019, we estimate the negative impact of COVID-19 lockdown in the capital of Spain, Madrid, using commercial satellite imagery courtesy of *Maxar*[©]. We observed that satellite imagery is a rich source of information, thus we design a product to deal with the aforementioned problem using advanced computer vision and machine learning algorithms to generate robust, predictive and consumable information for policy makers.

Automated analytics involving detection, tracking and counting vehicles from satellite imagery are still relevant and open problems in computer vision. During this BDExp, we focus on the development of a car-counting solution to monitor the presence of visible cars within these high-resolution images. In this way, we process different “regions of interest” (*i.e.*, hospital, shopping centre, industry, to name a few) before and during the COVID-19 lockdown. We hypothesize that COVID-19 outbreak would have negative impacts on the global economy as a consequence of the urban mobility restrictions. For this reason, it is essential to estimate the reduction of vehicles during the confinement.

In the experiments we also measure the performance of the our proposal, and compute car-counting statistics to quantify the dramatic drop in the number of vehicles during the lockdown. As a result, we corroborate these statistics using additional indicators such as telco, traffic and economic data respectively. We reach the conclusion that these insights correlate with official statistics on activity, thus car-counting statistics can complement traditional measures of economic activity in helping policy makers tailor their responses to flatten the recession curve.

Contents

1	Introduction	1
1.1	Motivation	1
1.1.1	Economic activity	1
1.1.2	Natural disasters	3
1.1.3	World environment	3
1.1.4	Artificial Intelligence in Telefónica	4
1.2	Objectives	5
2	Remote sensing data	7
2.1	Satellite viewers	8
2.2	Public labelled data sets	11
2.3	High-resolution data set over Madrid	13
3	Monitoring vehicles	19
3.1	Business plan	19
3.2	Market study	20
3.3	Technology survey	21
3.3.1	Counting by regression	23
3.3.2	Counting by detection	23
4	Results	25
4.1	Implementation	26
4.2	Additional telco and traffic data	26
4.3	Vehicle presence analysis	27
5	Conclusions	32
5.1	Future work	32
A	Satellites	34

Introduction

The ambition of EIT Digital is to educate a new generation of leaders in future technologies, as an answer to the evolving needs of the European economy. The EIT Digital Doctoral School (<https://doctoralschool.eitdigital.eu/>) innovates with a new kind of doctoral programme based on offering a deep expertise in key digital areas together with a strong background in Innovation and Entrepreneurship (I&E). In this document, we present the last phase of this doctoral education also denoted as Business Development Experience (BDExp). It involves an internship period of six months at a large company to assure that PhD candidates attain I&E skills.

During my doctoral research, we deal with the problem of accurately estimating the pose of human faces in images. This involves the location in the image of a set of facial fiducial points or “landmarks” that represents the combination of the rigid and non-rigid components of face pose [20, 1, 21, 14, 19, 22]. Analogously, during the BDExp, we also consider the problem of locating visible vehicles by processing aerial images using similar computer vision algorithms.

In this context, *Telefónica* provided me with the opportunity to work with one of the largest telephone operators and mobile network providers in the world. They believe big data and artificial intelligence can be of great use to society’s development (something they call “Big Data for Social Good” within their big data business, LUCA). Therefore, they work in various social, humanitarian and environmental areas to give the world back the value of data and contribute to the UN Sustainable Development Goals for 2030. For more information, visit <https://luca-d3.com/data-for-good/data-social>

1.1 Motivation

There is a growing interest in the analysis of satellite imagery for multiple domains:

1.1.1 Economic activity

When governments, international institutions or NGOs want to monitor population displacements (see Fig. 1.1) or identify the economic growth in cities, they can also act fast and efficiently thanks to satellite imagery. In a normal year, millions of Muslims flock to Mecca. However with COVID-19 only a few thousand have been allowed to make the journey this year, under strict rules.

The expansion of COVID-19 is causing a healthcare crisis with an impact on economic sectors. The authorities in Spain are adopting all necessary measures to contain the spread of the virus and mitigate its impact on the national economy. Economic activities impact social behaviours, which leave signatures in satellite images that can be automatically detected and classified.

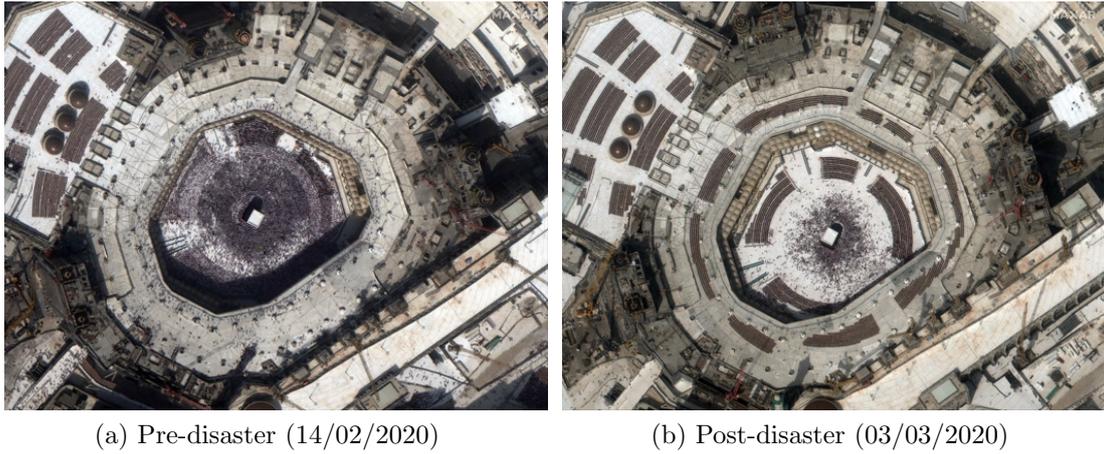


Figure 1.1: Satellite images courtesy of *Maxar*[©] illustrate a smaller crowd surrounding the Kaaba at the Grand Mosque in the holy city of Mecca, Saudi Arabia, during COVID-19.

Additionally, recent literature estimates the poverty in different regions, which is the first “Sustainable Development Goal” by using night-time light satellite images. However, non-commercial sensors from public satellites, *e.g.*, Day/Night Band sensor (DNB) from Suomi National Polar-orbiting Partnership (SNPP) Visible Infrared Imaging Radiometer Suite (VIIRS), only provide images at a resolution of about 750m.

As a result, commercial night-time images are required to infer economic activities, *e.g.*, quantifying the lights emissions of each city [8]. The Jilin-1 satellite constellation are Chinese remote sensing satellite sensors, which are owned by *Chang Guang Satellite Technology Company*[©]. Jilin-1 satellites have the capability to capture high definition images (1m), video sequences and hyperspectral imagery for various mapping applications including environmental monitoring, forest management, energy, mining, land planning and more (see Fig. 1.2).

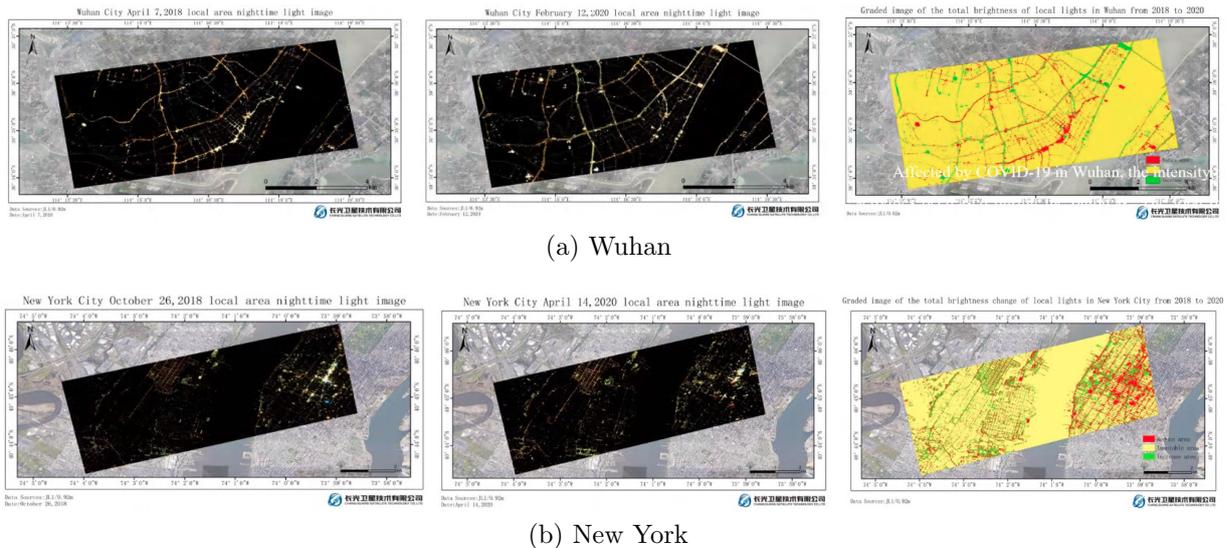


Figure 1.2: Comparison of night-time satellite imagery pre/post COVID-19 lockdown in two different cities. Samples provided by *Veritas Imagery Services*[©].

1.1. Motivation

1.1.2 Natural disasters

This is likely the line of action that most directly contributes to saving human lives. We can help in several ways by locating vulnerable population centres and identifying pre and post disaster patterns such as floods, storms, earthquakes, eruptions, wildfires, shoreline erosion, oil spills and burning gas fields prevention and damage assessment [5] (see Fig. 1.3). When countries are affected by major floods events (see Fig. 1.4), efficient response is almost unthinkable without the analysis of satellite imagery to gather detailed information about water extent and affected population.

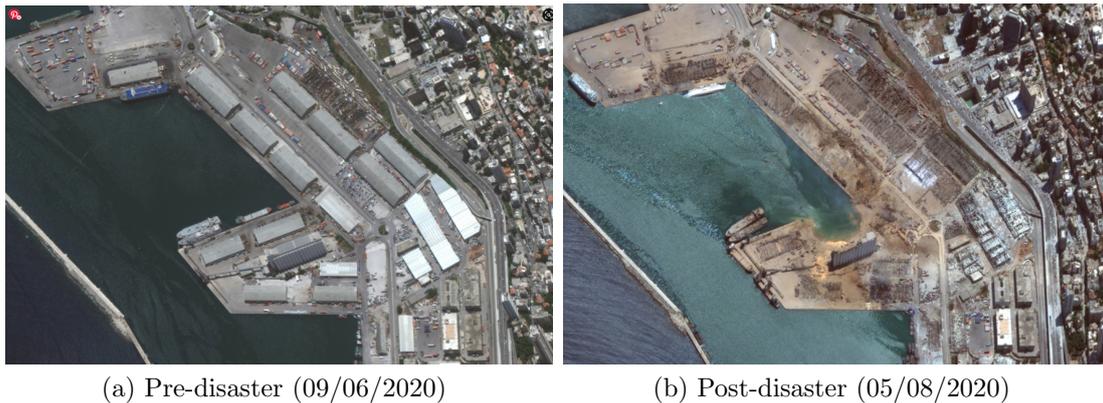


Figure 1.3: Satellite images courtesy of *Maxar*® show the scale of the damage caused by Beirut explosion of 4th August 2020, with buildings miles from the port lying in ruin.

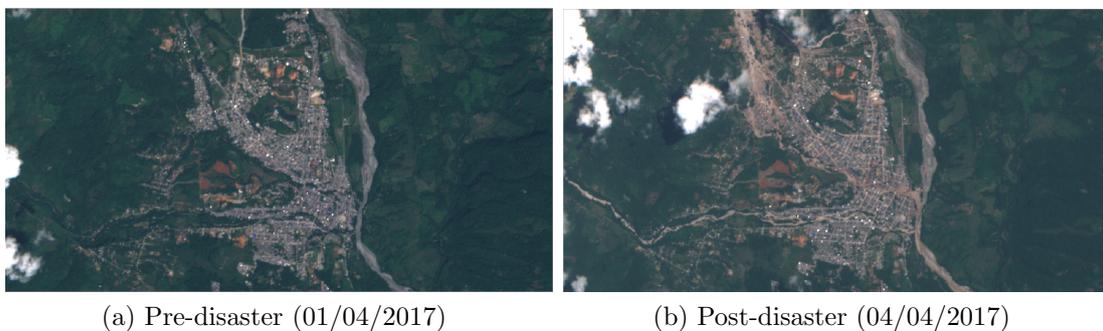


Figure 1.4: The Mocoa landslide was a natural disaster that took place in 2017. During the pre-dawn hours of 1st April, locally heavy rain triggered flash flooding and landslides in the city of Mocoa.

1.1.3 World environment

When governments or financial institutions want to assess the evolution of dam/bridge construction work in remote areas (urban change analysis), they can do it using satellite imagery, without having to send somebody to the field too often (see Fig. 1.5). Moreover, it is crucial to evaluate the achievement of the “Sustainable Development Goals” in countries with low financial resources, *e.g.*, quantifying the lack of water resources, the reduction of forests, mapping to establish ownership of land and buildings, monitoring the size and health of their crops (agriculture), and growing threats due to climate change.



(a) Pre-disaster (28/04/2017)

(b) Post-disaster (22/02/2020)

Figure 1.5: Satellite images courtesy of *Maxar*[©] show the hospital being built to handle COVID-19 pandemic at the Chinese city of Wuhan.

At the same time, instruments onboard environmental satellites are showing the impact of the COVID-19 on the environment. As factories halt operations and businesses close, air quality is improving in hard-hit regions. Recent studies [15] prove that the “Tropospheric Monitoring Instrument” sensor onboard Sentinel-5P have detected reductions in nitrogen dioxide, an atmospheric pollutant created by fuel combustion. Nitrogen dioxide concentrations fell sharply from March 2019 to March 2020 near major cities including New York, Paris, Madrid and Milan, according to the European Space Agency.

1.1.4 Artificial Intelligence in Telefónica

Telefónica is also a member of the Spanish Observatory for Big Data, Artificial Intelligence and Data Analytics (BIDA). This group consists of around 30 large enterprises that study the possibility of a B2G (Business to Government) data sharing initiative to provide future insights on climate change to policy makers (https://aeca.es/observatorio_bida/). BIDA believes that this would be the first time business data would be shared for the common good on such a large scale. Applying Artificial Intelligence (AI) to this unique combination of public and privately held data has the potential to uncover so-far unknown insights about the relation between economic activities and potential measures to reduce climate change.

However, when the COVID-19 pandemic erupted, companies had to change. It has become a key concern with a major impact on businesses across sectors. Consequently, all decision-makers of the BIDA initiative (*i.e.*, *Telefónica*[©], *BBVA*[©], *Mapfre*[©], *Red Eléctrica de España*[©], *Instituto Nacional de Estadística*, *Banco de España*, *Instituto de Estudios Fiscales*) decided to share statistical data correctly anonymized in order to infer economic indicators for a post COVID-19 economic recovery.

It is also worthwhile to notice an increasing role of other “computer vision” products in *Telefónica* (<https://www.telefonica.com/en/web/press-office/-/telefonica-designs-the-most-advanced-system-to-monitor-the-number-of-people-and-occupancy-on-the-beaches>). Recently, they develop the most advanced system on the market to monitor the number of people and occupancy on the beaches. In the context of COVID-19 restrictions, the protocol for the opening of beaches recommends the authorities (town halls, coasts, autonomous communities) to determine the capacity of bathers and set the maximum number of people who may come in groups to the facilities. Through a video surveillance

1.2. Objectives

system, multiple cameras detect which areas are being occupied and which are free and calculate the percentage of occupation in real time. This capacity and occupancy control system has more than 99% reliability thanks to the cameras equipped with powerful software using deep learning techniques through a neural network capable of discriminating against other elements and discarding them from the counting metrics as it only counts people (see Fig. 1.6).



Figure 1.6: With this *Telefónica*[©] solution, town halls can control, simply and effectively, the capacity of its beaches through cameras that capture frames from the area.

1.2 Objectives

The European Commission has issued a coordinated economic response to the COVID-19 outbreak. This includes an analysis of the impact that COVID-19 pandemic has in the private equity sector and proposed measures to mitigate harm. *Telefónica* observed that satellite imagery is a rich source of information, yet it is less investigated and sustainable results can be converted into a business.

Right now, organisations providing services based on remote sensing data do not know how to automatically extract all the information implicitly included. My role will be to design the guidelines to standardize a software to deal with the aforementioned problem, and how to carry out an intensive analysis of the satellite information that allows other institutions or authorities to evaluate the economic impact of COVID-19 policies. In this way, statisticians and managers would improve their tools and skills for working with such satellite data, by integrating this AI software in their standard statistical business processes.

Nowadays, there is an increasing number of organizations providing services based on satellite imagery (see Section 3.2), and the vast majority have reportedly used techniques like car-counting, tracking oil inventories or watching corn fields to make profitable forecasts of equity and commodity markets, to name a few. We also believe that car-counting estimated through AI techniques would provide useful insights to enable evidence-based policy making for solving societal problems. Recent studies reveal up to a 90% increase when comparing cars traffic between fall of 2018 and 2019 in several hospitals from Wuhan (see Fig. 1.7). We hypothesize that the amount of vehicles, which decreases drastically during the COVID-19 lockdown, could be correlated with the economic activity. Thus, we further investigate on how to accurately detect these vehicles using computer vision techniques.



Figure 1.7: Satellite images courtesy of *RS Metrics*[©] suggest that COVID-19 may have been present and spreading through China before the outbreak was first reported to the world. A dramatic increase in hospital traffic reveals that an infection was growing in the community and people have to see a doctor.

Finally, we would like to confirm the results obtained monitoring vehicles from satellite imagery comparing them with additional information such as: telco, traffic and economic data. Further analysis of each decision is also interesting to learn what “things” can be useful to maintain during the “new normal” status such as teleworking, etc.

Remote sensing data

In the GIT context (Geospatial Information Technology), when people are referring to remote sensing they are generally talking about Earth data captured by devices mounted on satellites, airplanes or drones. But strictly speaking, human eyes or a dolphin's sonar are also types of remote sensing systems. In this chapter, we categorize satellite remote sensing systems according to three different criteria:

1. The spatial resolution specifies the pixel size of satellite images covering the Earth surface (low spatial resolution: 30m - 1000m; high spatial resolution: 0.3m - 4m). Sensors with resolution lower than 0.3m are restricted to military services, so public images with low resolution are obtained using a drone, plane or globe (see Fig. 2.1).

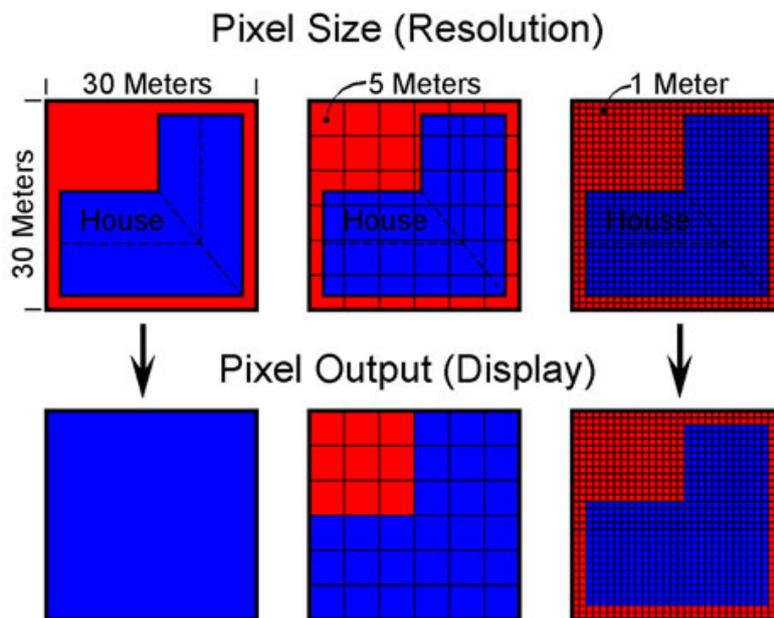


Figure 2.1: Spatial Resolution.

In the case of an image that has 10m spatial resolution, features that are smaller than 100m² will not be correctly detected. Satellite imagery has a maximum spatial resolution of about 30 centimetres, thus it is not possible to detect individual persons using either commercial or free satellites.

2. The temporal resolution specifies the revisiting frequency of a satellite sensor for a specific location (low temporal resolution: > 16 days; high temporal resolution: < 1 day - 3 days). Geostationary satellites always fly over the same area and can take several shots of the same location per day. On the contrary, most other satellites

(called “sun-synchronous”) are in constant movement, and can acquire images from a specific zone only every few days.

- Satellite sensors can provide complex types of information, including information that cannot be directly perceived by human eyes. In the first instance, a sensor’s spectral resolution specifies the number of spectral bands in which the sensor can collect reflected radiance (see Fig. 2.2). But the number of bands is not the only important aspect of spectral resolution. The wavelength of bands in the electromagnetic spectrum is also important (low spectral resolution: 3 bands; medium spectral resolution: 3 - 15 bands; high spectral resolution: 220 bands).

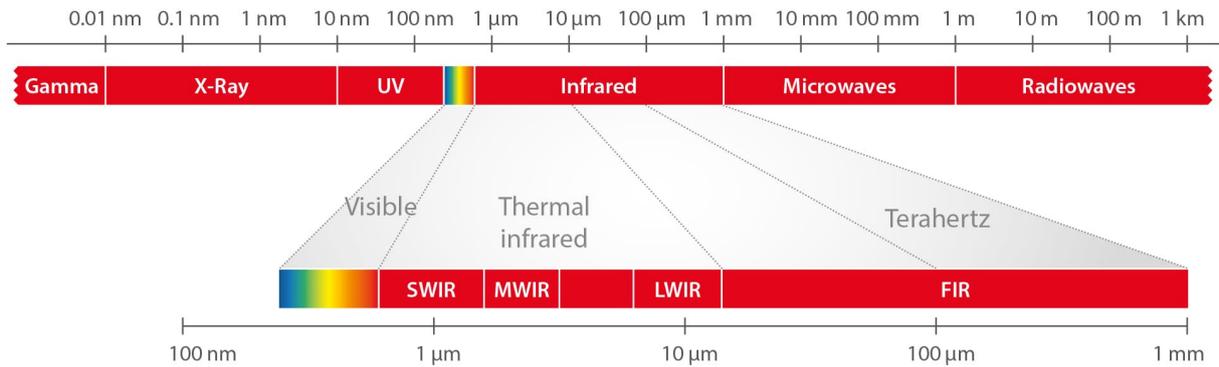


Figure 2.2: Spectral Resolution.

Unfortunately, because of technical constraints, satellite remote sensing systems can only offer the following relationship between spatial and spectral resolution: a high spatial resolution is associated with a low spectral resolution and vice versa.

Most satellites are built and maintained by **private** enterprises (see Table A.1). *DigitalGlobe*® and *MDA Holdings*® merged to become *Maxar*® on 2017, the most important commercial vendor of space imagery and geospatial content (<https://www.digitalglobe.com/products/satellite-imagery>). It is also worthwhile to mention other big and small companies, e.g., *Airbus*® (<https://www.airbus.com/space.html>), *Planet*® (<https://www.planet.com/>), *Earth-i*® (<https://earth-i.space/>), *CG-Satellite*® (<https://www.cgsatellite.com/>), *Satellologic*® (<https://satellologic.com/>), to name a few.

Satellite imaging of the Earth surface is of sufficient public utility that many countries also maintain remote sensing programs. USA has led the way in making this data **public** for scientific use, e.g., National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA), European Space Agency (ESA), etc.

2.1 Satellite viewers

Ordering high-resolution satellite imagery can be a real hassle. It requires finding a reputable satellite imagery distributor, knowledge on what type of image data is needed for the job, running a search through several image search and discover services, obtaining a price quote, negotiating terms, sending a purchase order, waiting for the order to process, and finally downloading very large data files.

2.1. Satellite viewers

There are multiple online subscription services available, but we only mention the most relevant:

- **SecureWatch** delivers the +100 petabyte imagery library from *Maxar*[©], the first satellite imagery provider, which includes all commercial daily collections from the WorldView-2, WorldView-3, WorldView-4 and GeoEye-1 series, for online viewing or on-demand archive retrieval (<https://securewatch.digitalglobe.com/>). For example, it has been utilized by *Google Earth Pro*[©].

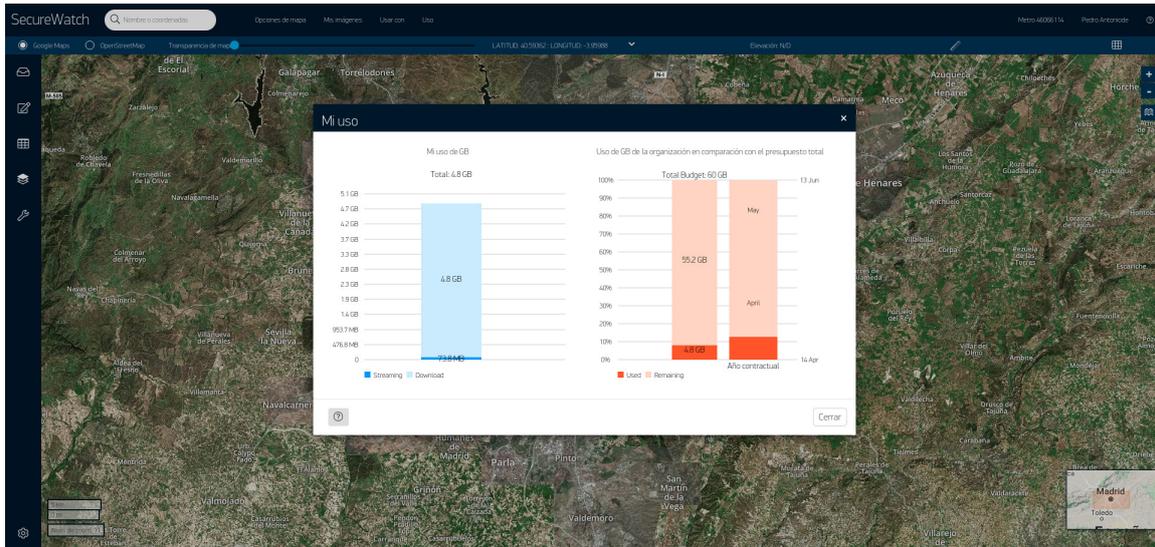


Figure 2.3: SecureWatch trial accounts allow us to download up to 60GB data.

We obtain a trial license for two months from 14th April to 13rd June (see Fig. 2.3). Our account provides full access to the SecureWatch catalogue, including FirstLook, Metro and Vivid modes (see Fig. 2.4). *FirstLook* is everything customers need for emergency management providing access to pre/post event imagery. *Metro* supplies up-to-date high-resolution imagery (0.3m) of more than 6000 world capitals. Images are mosaicked together to provide a consistent view of the urban landscape. *Vivid* also provides a high-resolution mosaic (0.5m) of a whole country.

We have downloaded multiple high-resolution images from their catalogue to create an extensive data set of Madrid (see Section 2.3). However, it is worth mentioning two main disadvantages of the SecureWatch platform: 1) the WorldView series, GeoEye, QuickBird, KompSat and IKONOS sensors always capture images at 11:00am approximately. Thus, it is impossible to acquire images from night-time to analyze the illumination [8], or earlier at 08:30am to deal with traffic jams. 2) the amount of images at most cities is insufficient. In the most important districts, such as Madrid or Barcelona, we approximately find one image per month, but, in other cities, there are only one or two images per year. Additionally, we have to discard certain images due to some appearance problems, *e.g.*, noise, presence of clouds, etc.

- **OneAtlas** is a unique collaborative environment to easily access premium imagery from Pleiades, SPOT, and open source data such as Sentinel-2. We perform large scale image processing, since we benefit from *Airbus*[©] assets (<https://oneatlas.airbus.com/>).

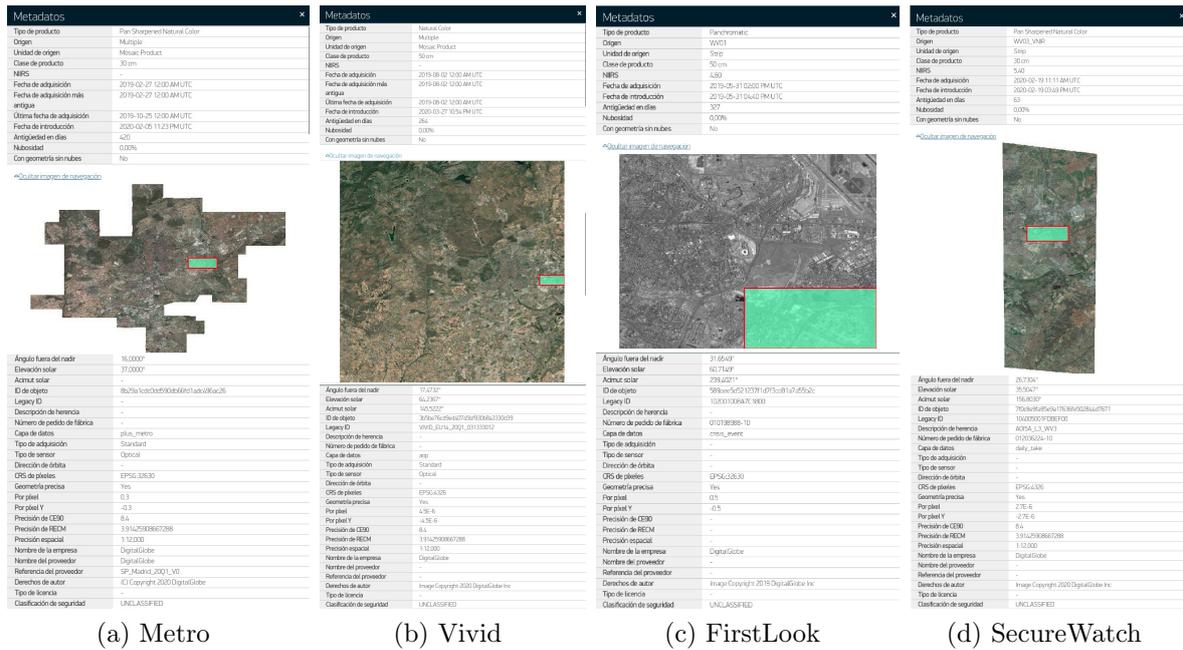


Figure 2.4: Metadata from samples acquired in Metro, Vivid, FirstLook and SecureWatch modes around “El Cañaveral”.

- **Planet Explorer** offers its own API integrations with direct integration to ArcGIS and the open source geospatial platform, QGIS. We also consider *Planet*[®]'s biggest advantage as the frequency of image collections with its signature huge constellation of PlanetScope satellites (<https://www.planet.com/explorer/>).
- **ESRI** (Environmental Systems Research Institute) is an international supplier of geographic information system (GIS) software. ESRI uses the name ArcGIS to refer to its suite of GIS software products (<http://maps.arcgis.com/apps/MapSeries/index.html?appid=9766dba97c954fcaa175da83b72ccf06>). It provides access to Landsat (<https://livingatlas2.arcgis.com/landsatexplorer>) or Sentinel data (<https://sentinel2explorer.esri.com>).
- **NASA's GIBS** explores the past and present of our planet through NASA's Global Imagery Browse Services. GIBS provides quick access to over 900 satellite imagery products available within 3-5 hours of being observed. View GIBS in action by using Worldview (<https://worldview.earthdata.nasa.gov/>) and Earthdata Search (<https://search.earthdata.nasa.gov/>). Additionally, NASA-ARSET provides online seminars to train people to visualize, interpret and apply remote sensing data (<https://arset.gsfc.nasa.gov/all/upcoming-arset-trainings>).

Download free satellite data from Amazon Web Services (AWS) (<https://registry.opendata.aws/tag/satellite-imagery/>). Example using Landsat-8 data:

```
aws s3 cp --recursive
s3://landsat-pds/c1/L8/201/032/LC08_L1TP_201032_20190904_20190917_01_T1/ .
```

However, anonymous access to most satellites stored in AWS is not allowed (CBERS-4, Sentinel-2, Sentinel-1 or Terra/MODIS). The request authentication enables Amazon S3 to identify and charge the requester for their use of the bucket.

2.2. Public labelled data sets

Additionally, Sentinel satellite data is also available via Copernicus (<https://scihub.copernicus.eu/>) and 5 Data and Information Access Services (DIAS) through the following platforms: CreoDIAS (<https://creodias.eu/>), Mundi Web Services (<https://mundiwebservices.com/>), ONDA DIAS (<https://www.onda-dias.eu/cms/>), WEkEO (<https://www.wekeo.eu/>) and Sobloo (<https://sobloo.eu/>).

We have downloaded via Copernicus multiple images from the same location at different spatial resolutions to visually compare their quality (see Fig. 2.5).

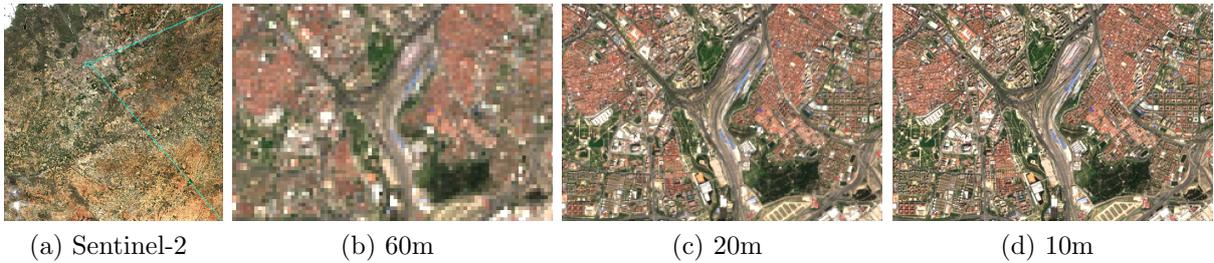


Figure 2.5: TCI (True Colour Image) images acquired from the Sentinel-2 satellite using different spatial resolutions. Best viewed in color and zoomed in.

As a result, we reach the conclusion that public satellite images of 10m resolution does not contain enough information to accurately detect vehicles in the scene (see Fig. 2.6).

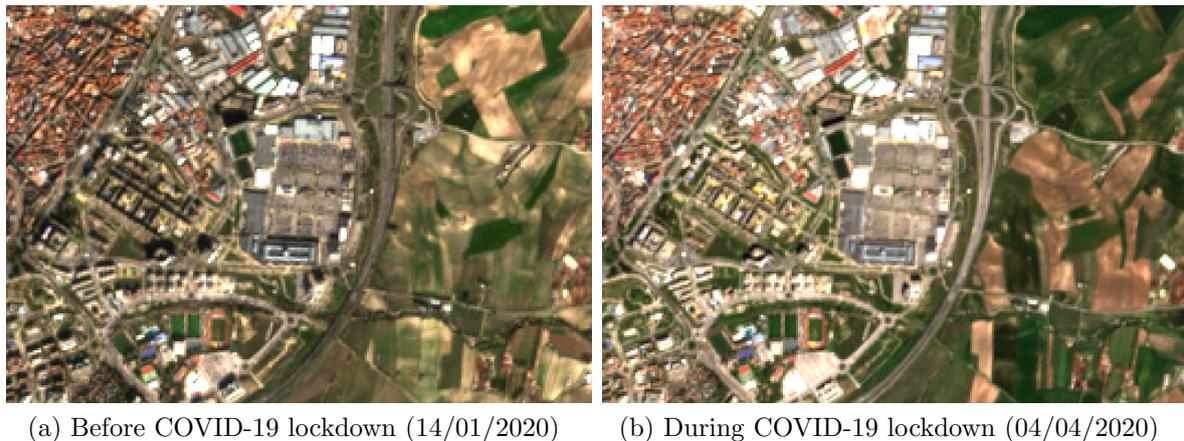


Figure 2.6: Differences between two Sentinel-2 satellite images of the same location before and during COVID-19 pandemic (shopping centre “Plaza Norte 2”).

Moreover, the satellite imagery downloaded is not labelled, *e.g.*, annotated with the location of each visible vehicle. Consequently, other annotated data sets are required to train sophisticated computer vision models that learn how to extract relevant information from these images.

2.2 Public labelled data sets

- DOTA [23] (<https://captain-whu.github.io/DOTA/>) is a large-scale data set for **object detection** in aerial images. The fully annotated DOTA images contains 188282 instances, each of which is labelled by an arbitrary (8 dof) quadrilateral (see Fig. 2.7).

- COWC [16] (<https://gdo152.llnl.gov/cowc/>). The Cars Overhead With Context data set is a large set with annotated cars from overhead (0.15m). It provides 32716 annotated cars that are useful for training models such as deep neural networks to learn how to **detect and count cars**. COWC-M extends the original data set by adding labels for the classes of cars “sedan”, “pickup”, “other” or “unknown”.
- CARPK [7] (<https://lafi.github.io/LPN/>). The Car Parking Lot data set contains nearly 90000 cars from 4 different parking lots collected by means of a drone at approximate 40 meters height. The image set is annotated by one bounding box per car. All labelled bounding boxes have been well recorded with the top-left points and the bottom-right points required to **detect and count cars**.
- VEDAI [17] (<https://downloads.greyc.fr/vedai/>). VEDAI is a data set for Vehicle Detection in Aerial Imagery, provided as a tool to **detect and count cars** in unconstrained environments (0.125m). The vehicles contained in the database, in addition of being small, exhibit different variabilities such as multiple orientations, lighting/shadowing changes or occlusions.
- DLR-MVDA [12] (<https://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-12760>). The DLR Munich Vehicle data set was collected over the Munich city, Germany to **detect vehicles**. The images were captured from an airplane by a Canon Eos 1Ds Mark III camera. The optical image is taken at a height of 1000m above ground, and the ground sampling distance is approximately 0.13m.
- EAGLE [2] (<https://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-12760>) is a large scale data base for **multi-class vehicle detection**. It consists of 215986 instances manually annotated with oriented bounding boxes by airborne imagery experts.

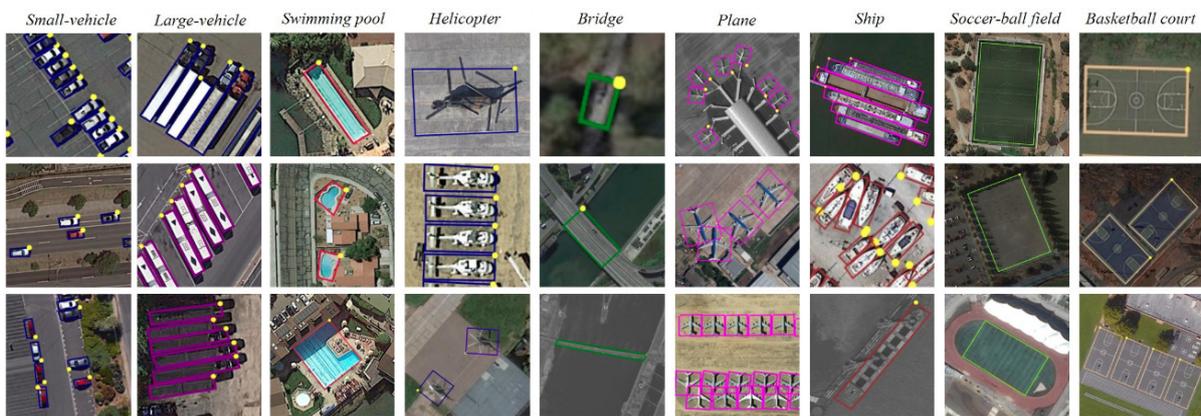


Figure 2.7: Multi-class object detection.

- fMoW [3] (<https://www.iarpa.gov/challenges/fmow.html>) consists of over one million images from over 200 countries and includes images and metadata released in the benchmark. For each image, we provide at least one bounding box annotation containing one of 63 categories. In this second **classification problem**, we learn to recognize objects such as “airport”, “hospital”, “military facility”, “prison”, “zoo”, “stadium”, etc.

2.3. High-resolution data set over Madrid

- xView [9] (<https://xviewdataset.org>) holds images from complex scenes around the world, annotated using bounding boxes with 60 classes at 0.3m spatial resolution. In such a **classification problem**, we learn to recognize objects such as “aircraft”, “barge”, “building”, “bus”, “tower”, “ferry”, etc.
- xView2 [5] (<https://xview2.org>) provides high-resolution satellite imagery for **building damage assessment**. It uses imagery from the *Maxar*®’s Open Data Program. It has annotated polygons and damage scores for each building, giving particular attention to on-the-ground changes between pre-disaster and post-disaster imagery (see Fig. 2.8).

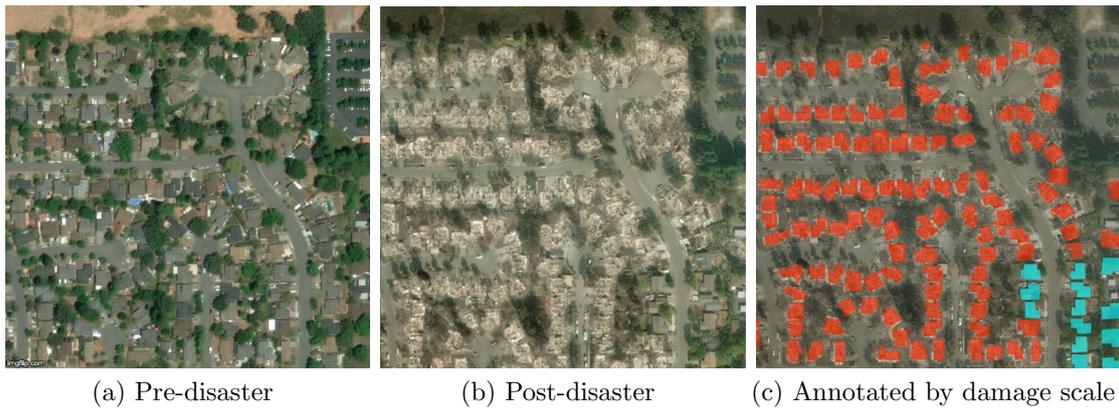


Figure 2.8: Building damage assessment.

- iSAID [25] (<https://captain-whu.github.io/iSAID/>) is a large-scale data base for **instance segmentation** in aerial images. It contains 2806 high-resolution images densely annotated with 655451 object instances and 15 categories.
- SpaceNet [4] (<https://spacenetchallenge.github.io/>) also focuses on the use of off-nadir imagery for building footprint extraction **segmentation**. The data set includes building labels for 6 different areas of interest (Rio, Vegas, Paris, Shanghai, Khartoum and Atlanta).

2.3 High-resolution data set over Madrid

In this section we have collected multiple images over Madrid using the SecureWatch platform. High-resolution imagery, as provided by commercial satellites that daily collect images with a ground sample distance of 0.3m around the globe, can be a valuable asset to estimate the impacts of COVID-19 in society. We first consider the option to download mosaicked images over Madrid using the SecureWatch’s *Metro* catalogue (see Section 2.1). However, it only provides 1 mosaic per year covering up to 2432 km^2 which represents a 30% of the total area that the surface of the whole autonomous community occupies (see Fig. 2.9).

Alternatively, we use the standard *Secure Watch* catalogue to download/stream images acquired from 22 different hot spots in Madrid. They approximately delimit 163.5 km^2 of land area, which represents a 2% of the whole autonomous community. For each region of interest, we obtain all samples available in the SecureWatch platform (via WorldView-4

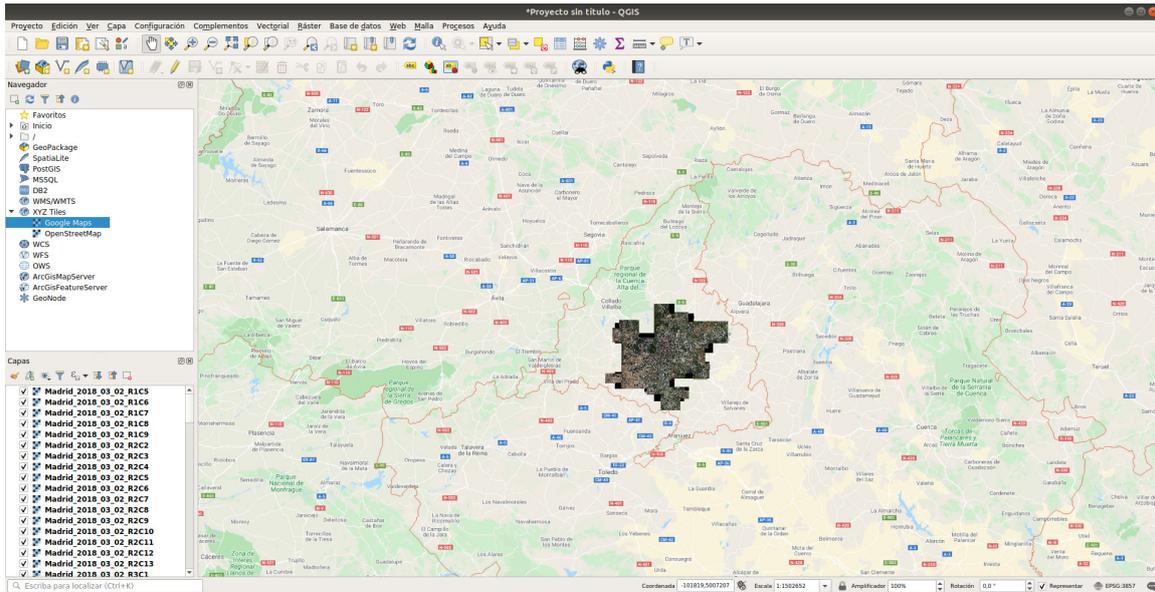


Figure 2.9: Sample mosaic obtained using the *Metro* catalogue from SecureWatch. With committed annual refresh, the Metro image layer provides a predictable and reliable, current view of the ever-changing urban landscape, however it is not useful to evaluate the economic impact of COVID-19.

satellite), which on average represents 1 image per month. As a result, we have collected 153 images saved in a private repository in AWS ([bobetocalo-maxar](#)).

Table 2.1 denotes the area and acquisition dates related to each region of interest. The selection of these hot spots in Madrid would allow us to handle different action lines (see Section 1.1). For instance, “Vicente Calderón” to assess the evolution of building constructions; “Navacerrada” to analyze the effect of climate change; “Morata de Tajuña” to evaluate agriculture activity; “Barajas Airport”/“EMT Carabanchel” to measure economic activity detecting transports such as planes/buses, etc.

In fact, as our goal is to evaluate the economic impact of COVID-19 lockdown through the presence of vehicles, we also define different subcategories such as shopping centres (e.g., “Parquesur”); hospitals (e.g., “Doce de Octubre”); industrial areas (e.g., “Polígono Cobo Calleja”); traffic congestion (e.g., “Manoteras”); parks (e.g., “Retiro”); universities (e.g., “Ciudad universitaria”), etc.

2.3. High-resolution data set over Madrid

Place	Area (km^2)	Date
Barajas Airport	30.52	(24/10/19) (19/02/20) (20/02/20) (04/03/20) (26/03/20)
Vicente Calderón	7.64	(19/12/18) (20/01/19) (25/03/19) (25/05/19) (21/07/19) (28/08/19) (27/09/19) (05/10/19) (16/10/19) (24/10/19) (06/12/19) (13/01/20) (20/02/20) (29/03/20) (24/05/20) (06/06/20)
El cañaveral	27.50	(11/11/15) (27/04/16) (09/03/17) (22/02/18) (23/09/18) (27/02/19) (24/10/19) (19/02/20)
Ciudad universitaria	7.63	(05/10/19) (24/10/19) (06/12/19) (13/01/20) (20/02/20) (29/03/20) (24/05/20) (06/06/20)
Polígono Cobo Calleja	30.63	(27/02/19) (24/10/19) (20/02/20) (29/03/20)
Mercamadrid	7.64	(27/02/19) (24/10/19) (04/03/20) (29/03/20) (24/05/20)
EMT Carabanchel	0.47	(27/02/19) (24/10/19) (20/02/20) (29/03/20) (24/05/20)
Doce de Octubre	1.91	(24/10/19) (20/02/20) (29/03/20) (24/05/20)
Fuenlabrada	1.91	(02/10/18) (27/02/19) (24/10/19) (04/04/20)
Alcalá de Henares	1.91	(27/02/19) (01/06/19) (16/09/19) (28/09/19) (24/10/19) (13/01/20) (19/02/20) (04/03/20)
Hospital San Carlos	1.90	(01/01/19) (07/01/19) (14/01/19) (20/01/19) (27/02/19) (25/05/19) (21/07/19) (28/08/19) (27/09/19) (05/10/19) (24/10/19) (06/12/19) (13/01/20) (20/02/20) (29/03/20) (24/05/20) (06/06/20)
Manoteras	7.63	(05/10/19) (06/12/19) (13/01/20) (20/02/20) (29/03/20) (24/05/20) (06/06/20)
Morata de Tajuña	7.66	(07/11/15) (27/04/16) (23/09/18) (02/10/18) (24/10/19) (24/05/20)
Navacerrada	7.59	(28/03/19) (02/10/19)
Parque de Atracciones	1.91	(13/01/19) (25/05/19) (09/07/19) (21/07/19) (28/08/19) (05/10/19) (24/10/19) (06/12/19) (20/02/20) (29/03/20) (24/05/20) (06/06/20)
Retiro	7.64	(06/12/19) (13/01/20) (20/02/20) (29/03/20) (24/05/20)
Puerta del Sol	1.91	(06/12/19) (13/01/20) (20/02/20) (29/03/20) (24/05/20)
Distrito Telefónica	1.90	(28/08/19) (24/10/19) (04/03/20) (24/05/20)
Vallecas	1.91	(01/01/19) (27/02/19) (25/05/19) (09/07/19) (21/07/19) (28/08/19) (27/09/19) (05/10/19) (24/10/19) (06/12/19) (13/01/20) (20/02/20) (29/03/20) (24/05/20) (06/06/20)
La Gavia	1.91	(27/02/19) (24/10/19) (19/02/20) (04/03/20) (29/03/20)
Las Rozas	1.90	(23/09/18) (27/02/19) (02/10/19) (04/04/20)
Parquesur	1.91	(24/10/19) (28/02/20) (29/03/20) (24/05/20)

Table 2.1: Data set collection with 22 different regions of interest over Madrid.



(a) Before COVID-19 lockdown (27/02/2019)



(b) During COVID-19 lockdown (04/04/2020)

Figure 2.10: Parking located at “The Style Outlets” shopping centre, the first outlet centre open in Spain.



(a) Before COVID-19 lockdown (20/02/2020)



(b) During COVID-19 lockdown (29/03/2020)

Figure 2.11: The “Polígono Cobo Calleja” is a huge economic area overwhelmingly dedicated to the wholesale distribution of Chinese imports.

2.3. High-resolution data set over Madrid



(a) Before COVID-19 lockdown (20/02/2020)



(b) During COVID-19 lockdown (26/03/2020)

Figure 2.12: “Barajas” is one of the world’s largest airport terminals in terms of area. The image includes the main building, T4 and a satellite building, T4S which is approximately 2.5km apart.



(a) Before COVID-19 lockdown (20/02/2020)



(b) During COVID-19 lockdown (29/03/2020)

Figure 2.13: The “12 de Octubre” is a public University Hospital for the southern zone of Madrid.

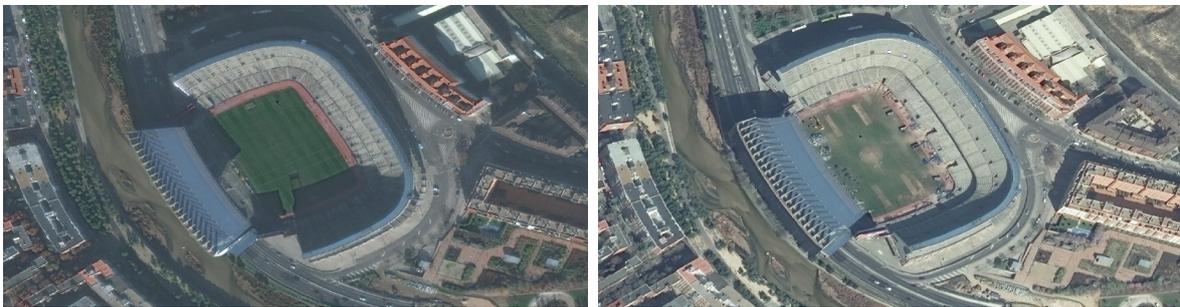


(a) Before COVID-19 lockdown (04/03/2020)



(b) During COVID-19 lockdown (24/05/2020)

Figure 2.14: The “Distrito Telefónica” is the headquarters of the Spanish telecom company Telefónica in Las Tablas, Madrid.



(a) 19/12/2018

(b) 25/03/2019



(c) 25/05/2019

(d) 20/02/2020

Figure 2.15: The “Vicente Calderón” stadium was the home stadium of Atlético Madrid since its completion in 1966 to 2017, with a seating capacity of 54907 and located on the banks of the Manzanares, in the Arganzuela district of Madrid, Spain.

Monitoring vehicles

Our main objective is the **location of cars** within each satellite image. In this way, by counting the number of cars in a certain region pre-post coronavirus confinement, we would measure its relationship with the economic growth.

First, it is crucial to annotate the aforementioned images with the position of each vehicle of interest in the scene. We use publicly available software as LabelImg (<https://github.com/tzutalin/labelImg>), which stores as XML files the coordinates manually annotated. It is worth mentioning that labelling such an image with up to 5000 cars would take 5 days approximately (see Fig. 3.1). Consequently, as we do not have enough time to annotate all images in our data set (see Table 2.1), we have just selected 10 images in order to evaluate the performance of our vehicle detector.

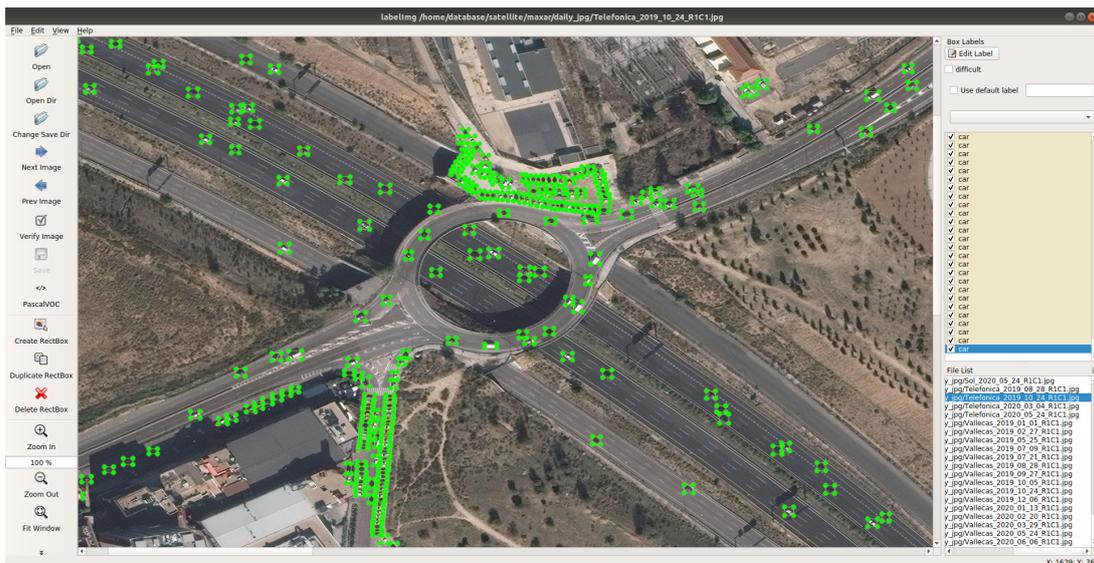


Figure 3.1: LabelImg is a graphical image annotation tool. It is written in Python and uses Qt for its graphical interface.

3.1 Business plan

The detection of vehicles in satellite images is important for various applications *e.g.*, traffic management, parking lots utilization, urban planning, etc. Collecting traffic and parking data from airborne platforms such as drones, airplanes, satellites, etc. gives fast coverage over a larger area. Getting the same coverage by terrestrial sensors would need the deployment of more sensors, more manual work, thus higher costs. In fact, investors who can afford to hire satellites to extract information are finding a significant edge in the stock market.

For example, a real application to the localization of buildings is to know exactly where to deliver assistance (high resolution population maps <https://ai.facebook.com/blog/mapping-the-world-to-help-aid-workers-with-weakly-semi-supervised-learning>). Similarly, it is in full alignment with relevant EU initiatives (Green Deal, Data Strategy, B2G data sharing report). The European Commission requested the sharing of satellite data related to the COVID-19 for research purposes (<https://www.euspaceimaging.com/eu-commission-asks-eo-community-for-help-with-covid-19/>).

However, to unleash the power in satellite images, we need automated AI-based computer algorithms to extract these kinds of information from them. The data space will be fed with a combination of privately-held and publicly available data. As our main contribution, we present a framework that recognizes specific elements in strategic locations to compute such economic insights. 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. Intellectual property developing the product is fully owned by *Telefónica*.

3.2 Market study

Satellite data value continues moving downstream towards big data analytics, as newer data sets and applications come online. Industry incumbents with operational capacities, *e.g.*, *Maxar*[©], *Airbus*[©], *Planet*[©], *Earth-i*[©], and so on, continue to partner and evolve their businesses towards offering data applications as part of their services. For example: Geospatial Big Data platform (GBDX) enables extraction of meaningful insights such as objects or materials, with access to *Maxar*[©] high-resolution satellite imagery (<https://www.maxar.com/products/gbdx>); *Earth-i*[©] launched AI-enabled SPECTRUM geospatial intelligence platform (<https://earthi.space/spectrum/>).

Alternatively, multiple venture capital firms invests in high-tech companies for the sole purpose of extracting data from satellite imagery to process custom imagery data feeds, and create vertical apps, trends, alerts and predictive signals. There are a few that we think are positioned for success: *RSMetrics*[©] (<https://rsmetrics.com/>) is the veteran providing fundamental insights, trends and predictive signals for businesses and investors in metals, industrials, retail and commercial real estate; *Orbital insight*[©] (<https://orbitalinsight.com/>) has attracted investment on agriculture, retail car counting, and oil inventory; *Descartes Labs*[©] (<https://www.descarteslabs.com/>) became known for correctly predicting a drop in domestic corn production based on its analysis of the change in color of plants over time; *Overstory*[©] (<https://www.overstory.ai/>) helps to prevent forest fires, deforestation and power outages with the mission to create a more sustainable environment; *Starlab*[©] (<https://www.starlab.es/>), based in Barcelona, adopts a different business strategy, focusing on direct contracts with ESA under support of the Catalan and Spanish Governments.

In terms of technology, we have also found different projects: *CosmiQ Works*[©] (<https://www.cosmiqworks.org/>) focuses on the development of cutting-edge technologies exploring the rapid advances delivered by computer vision; *Azavea*[©] provides an open source Raster Vision project making it easy to build computer vision models to understand and analyze geospatial imagery (<https://rastervision.io/>).

3.3 Technology survey

We develop a software to easily integrate technology able to detect objects of interest (*e.g.*, buildings, vehicles, roads, etc.), delimit their location in the scene, recognize their class (*e.g.*, truck, small car, tennis court, swimming pool, etc.) or estimate damage scores post disaster such as floods, storms, earthquakes, eruptions, wildfires, etc. As a result, it provides vector data used to display and work on geographic information systems (GIS). Vector data (.geojson) represents real world features using points and lines within the GIS environment. Raster data (.geotiff) stores the image of the world divided into tiles.

Object detection in remote sensing images is full of challenges. Present approaches detecting objects [12, 16, 17, 7, 23, 9, 2] typically fail or lose precision due to the relatively small size of the target objects and the vast amount of data to be processed in the presence of multiple “in-the-wild” factors, such as, different cities/countries, viewpoint changes, occlusions, illuminations, blurriness, and so on (see Fig. 3.2).



Figure 3.2: Challenging appearance variability due to different factors including viewpoint changes (nadir angle), shadows, daylight changes marked by weather and seasons, etc.

Additionally, the large scenes of interest processed, and the difficulty of dense object detection on traffic jams or parking lots are the main problems needed to solve. When the spatial resolution gets worse, the algorithm has less information to work with, and it gets harder to distinguish individual cars (see Fig. 3.3). Zooming into one part of the parking lot below causes quite the arguments. Are there five cars or six? Some say they can see the shadow of something between the white car and dark car on the right side of the image, so a car must be there. But others say the parking space is clearly empty. The assessment of whether or not a car is there is highly subjective. A human must make these decisions and this is using the highest resolution satellite imagery commercially available.

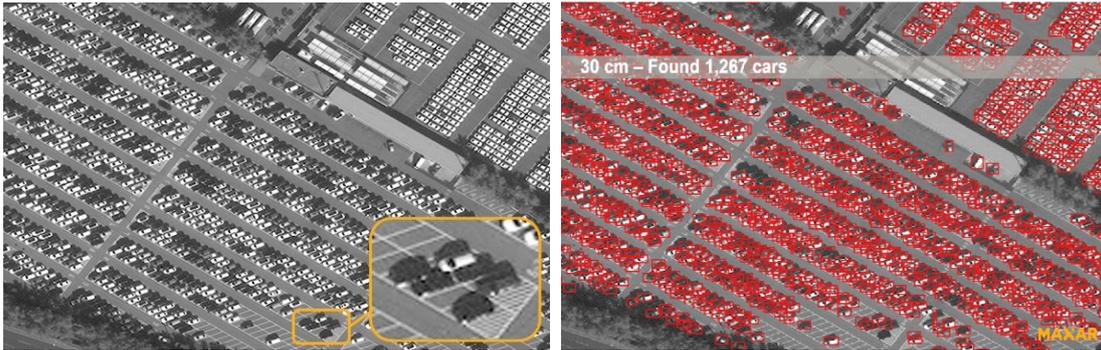
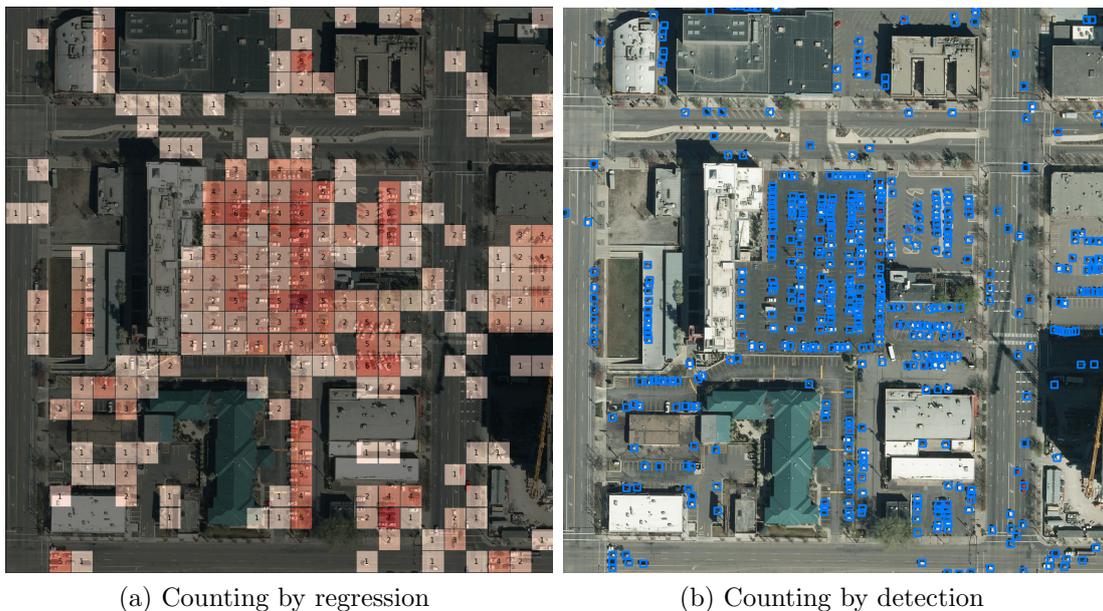


Figure 3.3: Monitoring vehicles around the Hyundai manufacturing plant in Ulsan, South Korea. The completed cars are stored in surface lots outside and parked closely together.

The counting of vehicles plays an important role in measuring the behaviour patterns of traffic flow in cities, as streets and avenues can get crowded easily. We categorize recent approaches into two groups according to whether they estimate the number of cars evading or not the hard task of learning to detect individual instances (see Fig. 3.4).



(a) Counting by regression

(b) Counting by detection

Figure 3.4: Both supervised approaches require a set of training images with annotations. (a) defines the presence of a vehicle by putting a single dot on each instance in the image; (b) specifies the vehicle position by setting the bounding box coordinates on each instance.

3.3.1 Counting by regression

These methods avoid solving the hard detection problem [10, 16, 13]. Instead, a direct mapping from the whole image, or from large image patches, to the number of vehicles is learned. Counting by density algorithms rely on regressors trained to estimate the objects density per unit area so that the total number can be obtained by integration, without explicit detection being required. They tend not to require many training samples, but are usually constrained to the same scene on which it was trained.

3.3.2 Counting by detection

Detection counters work in the more intuitive fashion of localizing each car uniquely and then counting the localizations [12, 17, 7, 11, 24]. In this work, we follow a deep learning approach based on SCRDet [24] that combines strategic location sampling and an ensemble of lightweight convolutional neural networks (CNNs) to quickly count small vehicles in satellite images that could be used to compute economic indicators based on it, automatically.

In Fig. 3.5 we show the most popular object detection architecture regarded as Faster R-CNN [18]. This network consist of: a) feature generation stage to obtain features of the objects, usually using a CNN (*e.g.*, ResNet-101 [6]); b) Region Proposal Network (RPN) to generate bounding boxes or locations of possible objects in the image; c) classification layer to predict which class each object belongs to; and d) regression layer to make the coordinates of the object bounding box more precise.

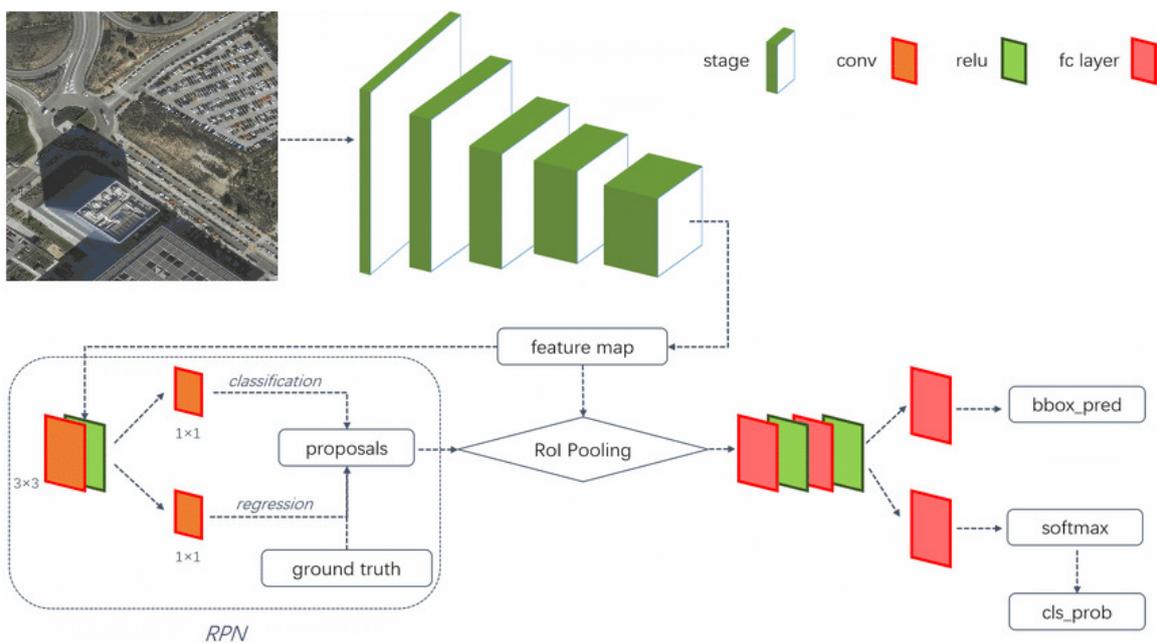


Figure 3.5: The Faster R-CNN architecture [18] consists of the RPN as a region proposal algorithm and the well-known Fast R-CNN as detector baseline.

We crop each satellite image into multiple squared tiles passed through the backbone CNN (*e.g.*, ResNet-101 [6]) to get the feature maps. For every point in the feature maps, the RPN has to learn whether an object is present in the input image at its corresponding location and estimate its size. This is done by placing in the input image a set of “anchors” for each location given by the output feature map from the backbone CNN. These anchors

indicate possible objects in various sizes and aspect ratios at this location (see Fig. 3.6b). Then, the bounding box proposals from the RPN are used to pool features from previous feature maps. After passing them through two fully connected layers, the features are fed into the sibling classification and regression branches. The loss is the weighted sum of the Softmax loss (classification problem) and the regression L_2 loss (localization bounding box coordinates). Finally, we apply non-maximum suppression (NMS) to discard overlapped bounding boxes due to repeated vehicles in consecutive tiles.

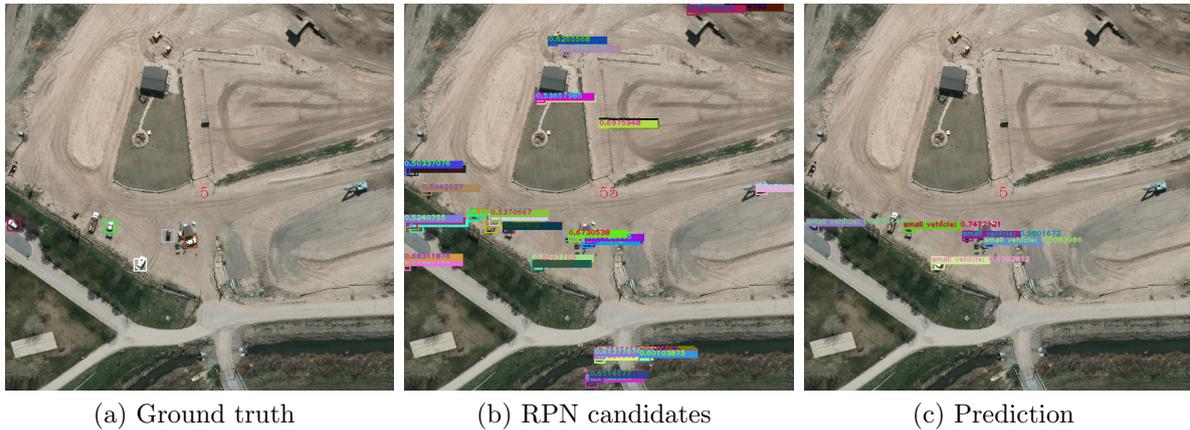


Figure 3.6: The learning process involves the generation of candidates categorized by the predicted class probability (`cls_prob`) and the coordinates position (`bbox_pred`). So, we count occurrences to set an overall number of small vehicles in the image. We show the results in the image center (best viewed in color and zoomed in).

In addition, as we only need “small vehicles” to infer economic insights, we ignore the remaining classes annotated (*e.g.*, bus, plane, ship, building, bridge, hospital, swimming pool, etc.) to save computational time, and we reduce the original anchors size to properly fit bounding box proposals to the standard car dimension.

Results

The vehicle detection is a challenging problem due to the small size of the vehicles and the complex background of man-made objects, which appear visually similar to the cars. However, recent literature reports noticeable results in the vehicle detection task [9, 24, 11, 2]. A critical question here is whether the models trained and evaluated in the same labelled data set, generalize or not to the situations present in real-life operation. In this way, we demonstrate that Mundhenk *et al.* [16] techniques, which correctly localize cars 90% of the time, and achieve an error rate on car-counting of 5% using 0.15m resolution images, do not work properly in satellite images with a lower or higher spatial resolution (see Fig. 4.1).

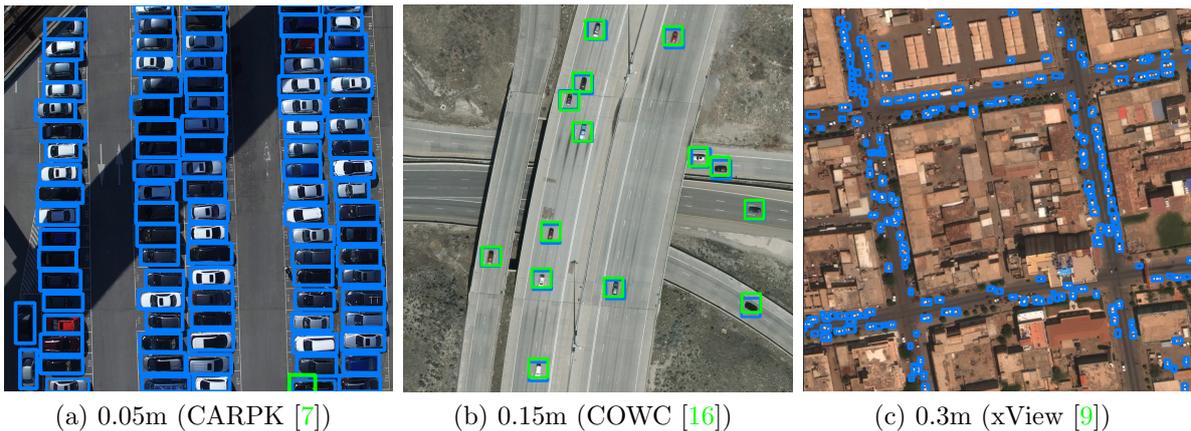


Figure 4.1: Car detection model following [16] algorithm trained on COWC (0.15m) and evaluated using different resolutions. Blue and green bounding boxes represent the ground truth and predictions respectively.

The experimental results of the aforementioned articles, according to the overall detection rates of the “small vehicle” class in the most challenging data sets, leave much to be desired. For instance, the average precision (AP) metric, which computes the average precision value for recall value over 0 to 1, is around 0.7-0.8 in the DOTA benchmark [24, 11], and 0.4-0.5 in the xView [9] and EAGLE [2] competitions respectively.

In addition, Azimi *et al.* [2] also demonstrate the lower generalization capabilities of models trained on a single data base. For instance, a model trained on DOTA imagery achieves a 0.38 AP on EAGLE, which represents a 27% AP reduction. In this chapter, we alleviate the problem of detecting vehicles using multi-scale training samples. As we have just annotated 10 images in our data set over Madrid (see Section 2.3), we train our model using also different vehicle annotations at different spatial resolutions from COWC [16] (32 images) and DOTA [23] (1869 images).

4.1 Implementation

All experiments have been carried out with the settings described in this section. We crop each satellite image into multiple tiles of 800×800 pixels as input passed through a generic object detector based on Faster R-CNN [18]. In this way, we follow SCRDet [24] public code to initialize our model parameters, but we train a single model using multiple vehicle annotations at different spatial resolutions from different data bases (*i.e.*, COWC + DOTA + Madrid subset), which improves the generalization capabilities of regressors trained on a single data set.

In our framework, we also generate different samples in each training epoch by applying random in plane rotations between $\pm 45^\circ$, and randomly mirroring images horizontally to increase the amount of variability in the data bases. In addition, we apply random scale changes w.r.t a spatial resolution of 0.3m, resizing each image according to its Ground Sampling Distance (GSD), which denotes the distance between two consecutive pixels measured on the ground. A GSD of 0.3m means that one pixel in the image represents linearly 30cm on the ground ($30 \times 30 = 900\text{cm}^2 = 0.09\text{m}^2$). As far as we know, this is the first time that an algorithm successfully combines images at different resolutions to deal with the lack of satellite data properly annotated.

Training the CNN takes 278 hours using a NVidia Tesla K80 (12GB) GPU provided by *Telefónica*, using a batch size of 1 image. The total number of iterations is 268840, which corresponds to 20 epochs \times 13442 image tiles (see Fig. 4.2). At runtime our approach processes test images on average at a rate of 0.025 FPS, being a surface area between 1.91 and 30.63 km^2 (see Table 2.1) processed in approximately 30 to 200 seconds respectively.

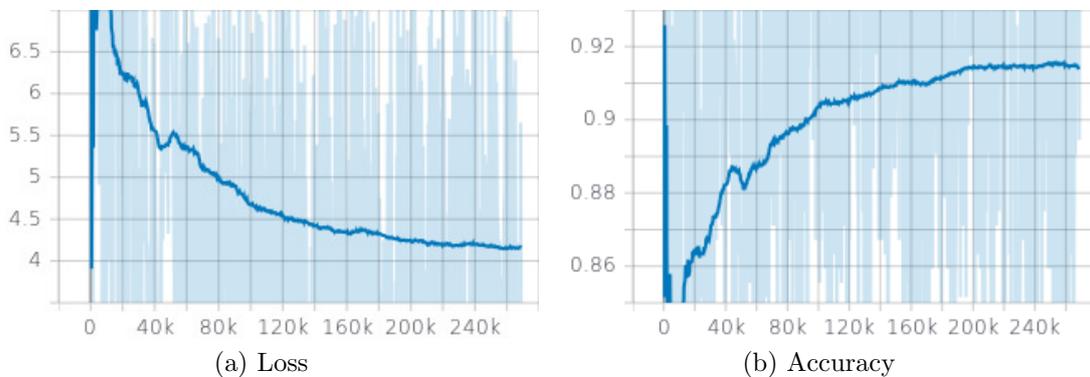


Figure 4.2: Visualizing the training loss and accuracy of the whole model over the total number of epochs. The bold curves denote the smoothed curves for visualization, and the faint curves denote the original curves.

4.2 Additional telco and traffic data

We use additional data sources in order to support the economic activity indicators inferred by monitoring vehicles. First, we analyze whether it is possible to include information from the remaining BIDA members (see Section 1.1.4).

- *Telefónica Movistar*[©] offers insights from the mobile phone network about the global trends of groups of people, using the Smart Steps platform (<https://luca-d3.com/>)

4.3. Vehicle presence analysis

[es/productos-servicios/business-insights/crowd-analytics](#)), which studies antenna activities and divides the city on a grid, while complies with existing regulations and data anonymization laws. It reports how many people is connected to each antenna, and every time a mobile changes antenna (handover) this is registered in the network, billions of times due to the fact that Telefónica have millions of customers.

The remaining BIDA members such as *BBVA*[©], *Mapfre*[©], *Banco de España*, *Instituto de Estudios Fiscales*, etc. communicate data infrequently (*e.g.*, one-year lag) and/or with a coarse-granularity level (*e.g.*, the whole autonomous community of Madrid). Thus, we explore other public data sources to provide further insights that correlate with the vehicle activity.

- *City Council of Madrid* provides an open data portal where preprocessed data from traffic sensors located on strategic points in the roads and streets of Madrid city, can be downloaded (<https://datos.madrid.es/portal/site/egob>). It provides data not only on the vehicle count but also on its speed and geographical location.
- *Axesor Rating*[©] estimates Spanish economic activity through 300 enterprises, which represent a 1.5% of the gross domestic product (PIB) approximately (<https://www.axesor.com/es-ES/es/covid-19/>).

4.3 Vehicle presence analysis

First, we analyze the performance of our vehicle detection technology in unrestricted satellite scenarios. Similar to [2], which reports a 0.38 AP in EAGLE by training a CNN model with DOTA [23], we achieve a 0.25 AP, which is also a poor detection score by processing a subset of 10 labelled images in our data set (see Table 2.1). Thus, we consider that the vehicle detection problem is still far from being completely solved using satellite images. However, it is worth mentioning that our main objective is to count vehicles. In this context, we obtain on average a 70% accuracy in the car-counting task, with a $\pm 15\%$ standard deviation from the mean. Sometimes it is difficult to extract the contour of a car for accurate detection, which decreases the vehicle detection accuracy, whereas it does not affect the vehicle count. In Fig. 4.3 we show a representative image where we get a 82.11% car-counting accuracy. We reach the conclusion that our software fails under most well-known “in-the-wild” conditions (*e.g.*, densely crowded parking lots, non-distinguishable objects, etc.), which are common situations in the vast majority of images.

We would like to measure the negative impact of COVID-19 lockdown in Madrid by automatically monitoring the presence of small vehicles through satellite imagery. The COVID-19 pandemic in Spain was first confirmed to have spread on 31st January 2020. On 15th March the national confinement became effective, and it ended on 21st June. All residents were mandated to remain in their normal residences except to purchase food and medicines, work or attend emergencies. In Fig. 4.4 we visually perceive a reduction in the total number of cars before and during the COVID-19 pandemic. Thus, it seems reasonable to study the overall impact of the lockdown on the traffic volume.

In this way, we consider as 100% the samples with the highest number of localized cars for each region of interest in Madrid, and consequently, we analyze how this percentage (*i.e.*, proportion of visible vehicles within the scene) increments/decrements its value along

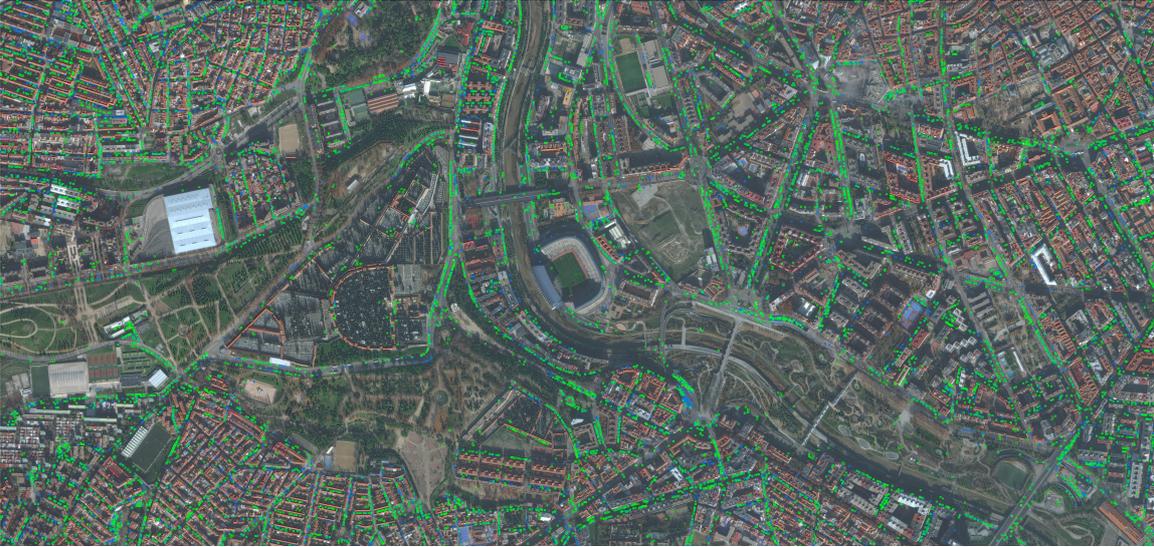


Figure 4.3: Counting metrics comparison between manually annotated and automatically detected vehicles using a “Vicente Calderón” sample (7.64 km^2). We show in the image centre 9998 and 8210 occurrences respectively (best viewed in color and zoomed in). Blue and green colours represent ground truth and predictions respectively.



(a) 2450 small vehicles before COVID-19 lockdown



(b) 365 small vehicles during COVID-19 lockdown

Figure 4.4: Representative satellite images from “Distrito Telefónica”. It can be noticed a dramatic reduction in the presence of vehicles automatically detected during COVID-19 lockdown.

4.3. Vehicle presence analysis

time. First, we show in Fig. 4.5a the evolution of the number of visible cars detected in the whole autonomous community of Madrid, obtained through the combination of data extracted from the 22 hot spots (see Table 2.1). In Fig. 4.5b, we also study the results acquired from a single instance, such as “Manoteras” in the north of the capital. Finally, in Fig. 4.5c and Fig. 4.5d, we analyze the statistics acquired from two relevant groups, such as shopping centers (“La Gavia”, “Las Rozas” and “Parquesur”) and hospitals (“Doce de Octubre”, “Fuenlabrada”, “Alcalá de Henares” and “Hospital San Carlos”) respectively.

We measure the impact of COVID-19 lockdown in Madrid by analyzing these curves. In Fig. 4.5, we notice this effect in the amount of detected vehicles, which on average decreases up to 55% from March to April. Surprisingly, we report a significant mean reduction of 75% in *shopping centres* probably because non-essential shops closed due to COVID-19 restrictions. We achieve on average a 40% reduction in traffic occupancy around *hospitals*, suggesting that people refused to go, unless they experienced severe symptoms. Finally, in the *Manoteras highway crossings*, we set a reduction of only 25%, because in such a large area (7.63 km²) we also process data around its neighbourhoods (*i.e.*, Hortaleza, Chamartín, and so on), and our software does not allow the distinction of moving and stationary traffic (see Section 5.1).

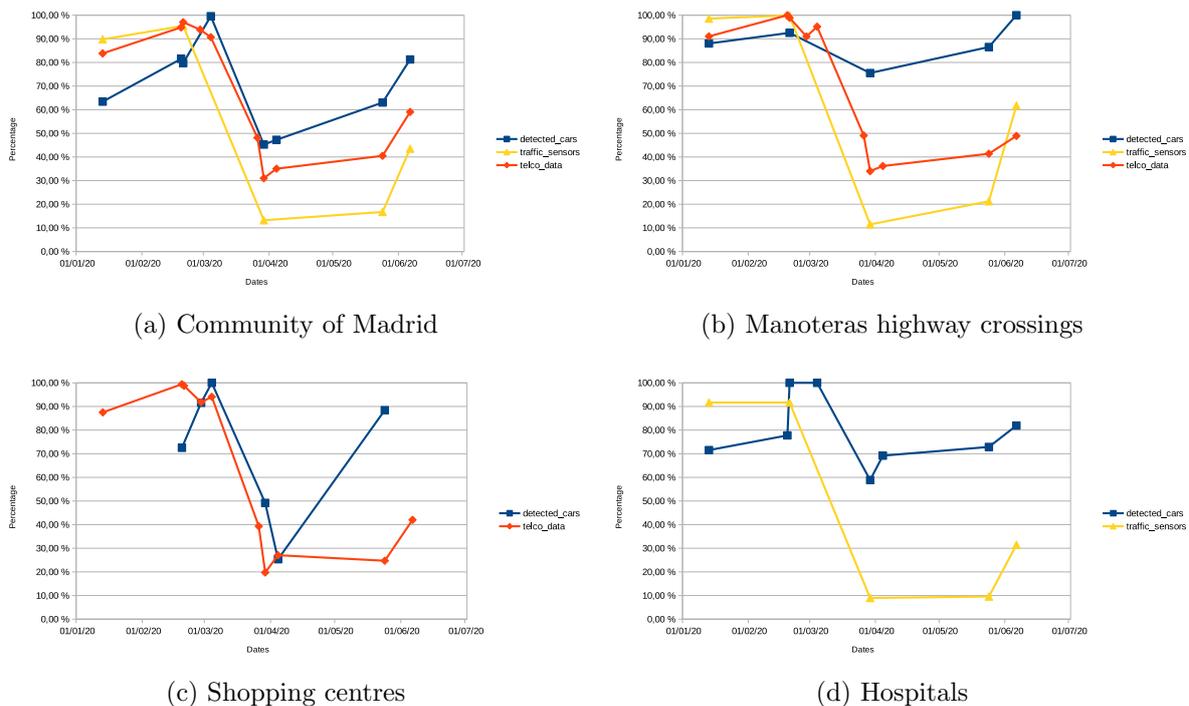


Figure 4.5: Timeline curves of how the COVID-19 outbreak is evolving in Madrid since 2020. Red, yellow and blue colours compare different curves obtained using anonymized and aggregated telco data from *Telefónica Movistar*[©] antennas, traffic statistics acquired from *City Council of Madrid* sensors, and by estimating presence of visible cars with our satellite technology respectively.

As we mentioned before (see Section 4.2), we also use additional data sources. First, we include anonymized and aggregated telco data acquired from *Telefónica Movistar*[©] mobile phone devices connecting to antennas within each region of interest. At this point, we also get on average a reduction of 60% in the autonomous community of Madrid from March to April in 2020, similar to the 55% obtained using the car-counting measures. In

Fig. 4.5 we plot these results in red colour. Although these statistics are consistent, it is worth noticing that telco data is generated using a time period of 10 minutes, whereas car-counting is computed at a limited frame (a fixed time value). In addition, those antennas located near the image boundary would receive data from *Movistar*[®] devices out of the region of interest. Moreover, we are also restricted because we cannot obtain telco data before 2020 and/or insights from healthcare areas due to the code of conduct¹.

Additionally, we include traffic statistics publicly available from the *City Council of Madrid* website. In this case, we show a mean reduction of 80% in the traffic around the city centre from March to April, which represents a 25% difference with the amount of vehicles detected (our technology does not allow the distinction of moving and stationary traffic). In Fig. 4.5 we plot these results in yellow colour. Anyway, we conclude that traffic reduction rates holds when compared with previous car-counting and telco indicators. It is also disappointing that traffic information is generated using a time period of 15 minutes, and traffic sensors are only located at the city centre, thus we cannot download data from outlying regions in Madrid (“Las Rozas”, “Fuenlabrada”, “Alcalá de Henares”, etc.).

In Fig. 4.6 we report traffic sensors and car-counting statistics since December 2018. Here, we appreciate in both curves a notable reduction of around 30%-55% during summer vacations from July to August, and the winter breaks in December. This decrement is reasonable since Madrid citizens rather spend their holidays in cooler regions in the north of Spain, or south/east to the Mediterranean beaches. During COVID-19 lockdown, which was somewhat unexpected, these statistics even fall to a 55%-80% reduction compared to previous March 2019, which relates to the obedience to a stay-at-home order.

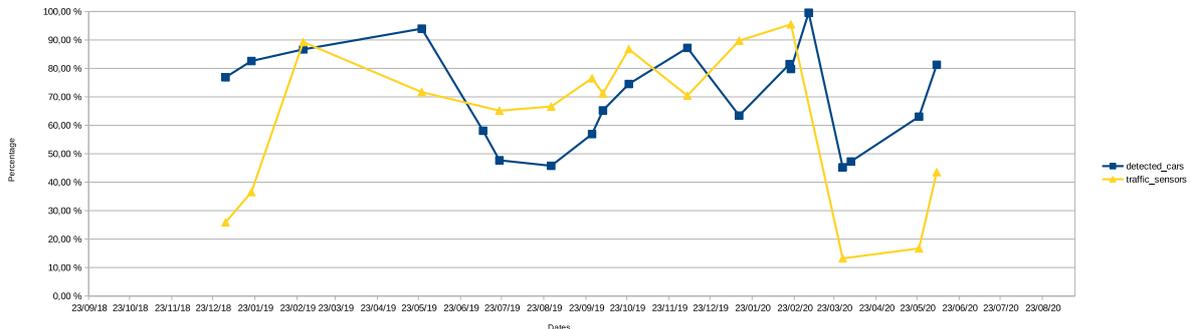


Figure 4.6: Comparison between traffic statistics acquired from sensors and car-counting estimations over the whole autonomous community of Madrid (see Table 2.1). Note the differences from March to May 2020 compared to the same period last year.

These indicators serve not only to verify stay-at-home compliance, but also to infer the economic impact of COVID-19 in a chain of companies such as shopping facilities, hotels, airlines, petrol filling stations, etc. 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. In Fig. 4.7 we prove that economic activity is also correlated with the aforementioned indicators. *Axesora Rating*[®] data, extracted from 300 enterprises in Spain, confirm an impressive 36.55% and 34.49% reduction in the number of sales during COVID-19 outbreak, similar to the 55% obtained using the car-counting measures in April-May.

We reach the conclusion that these insights correlate with official statistics, which experienced a 34% fall in economic activity during the first two weeks of the COVID-19 lock-

¹We gratefully acknowledge Javier Carro Calabor and Arturo Canales González this telco data.

4.3. Vehicle presence analysis

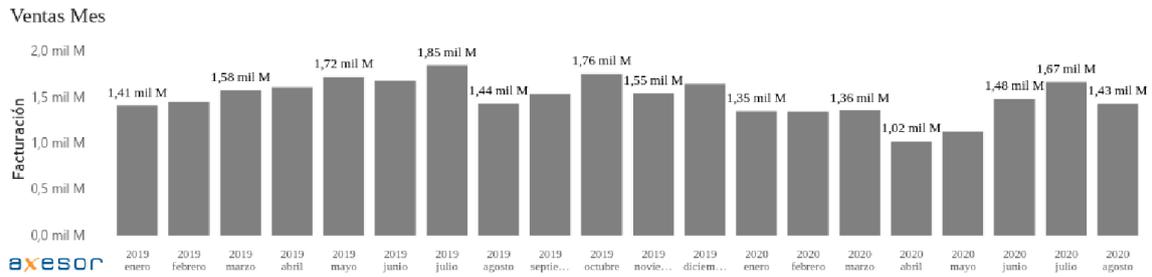


Figure 4.7: Accurate and timely economic data extracted from small, medium and large enterprises that represent a 1.5% of the gross domestic product in Spain.

down (https://english.elpais.com/economy_and_business/2020-06-18/spains-economy-fell-3.html). Consequently, one could imagine trying to improve worldwide measures from the observed insights through satellite imagery at this spatial resolution. In fact, using more high-resolution sensors and improved algorithms, these tasks could be done continuously and at low cost.

Conclusions

The activities done at *Telefónica*[©] helped me to foster my professional skills within an industrial environment, putting into practice the knowledge acquired during my doctoral studies. In this report, we deal with the problem of analyzing information acquired from commercial satellite imagery to provide robust insights during the COVID-19 pandemic in Madrid. The COVID-19 outbreak forced governments worldwide to impose lockdowns and quarantines to prevent virus transmission, which leave signatures in satellite images that can be automatically detected and classified using advanced computer vision algorithms. As our main contribution, we present a framework that counts small vehicles to compute such indicators automatically.

The availability of annotated data sets has also encouraged research in this area [9, 24, 11, 2]. However, a careful analysis of commercial software based on current algorithms, reveals the existence of a knowledge gap in how to train a regressor using satellite images from multiple data bases at different spatial resolutions. In this way, we propose a novel strategy to combine labelled data, training a vehicle detector that processes WorldView-4 imagery courtesy of *Maxar*[©] (see Table A.1) on average at 40 seconds per image. It is also worth noticing that repurposing the framework for other economic activities is simple. The objects to monitor, *i.e.*, roads, planes, ships, etc., can be changed easily.

As a result, we download a set of 153 high-resolution satellite images extracted from 22 hot spots around the community of Madrid (see Table 2.1). It includes images before, during and after COVID-19 lockdown. First, we measure the car-counting precision using 10 manually labelled images, which leads on average to a 70% precision in the car-counting task. Secondly, we extract car-counting statistics using all 153 images, where we notice a 55% mean drop in the presence of visible vehicles during the COVID-19 lockdown. Finally, we also prove the correlation of this indicator using additional data such as, a) telco data related to *Movistar*[©] mobile phone devices, b) traffic statistics publicly available from the *City Council of Madrid*, and c) economic activity from SMEs acquired from *Axesor*[©].

5.1 Future work

Commercial satellite imagery is currently in a sweet spot, powerful enough to see a car, but not enough to tell the make and model; collected frequently enough for a farmer to keep tabs on crops health, but not so often that people could track the comings and goings of a neighbour. The last few years have seen a number of interesting developments involving satellites. When future *Maxar*[©] satellites legion start launching in the first half of 2021, they will provide up to 0.15m high-resolution images of a location in a single day. And that is just the beginning when it comes to the usability of satellite imagery. *Planet*[©] and other new satellite companies are working to deploy higher resolution cameras, videos and radar sensors that can see through clouds to gather even more detail about what is happening on Earth.

5.1. Future work

An analyst does not have time in one shift to count the cars manually in each of those satellite frames. But, they have plenty of time to run all images through computer vision software, as the one provided in this work, to notice patterns and trends about when the numbers of cars change and investigate the cause, creating mission-critical information for influencing decisions, etc.

Although, we consider that the goal of the BDExp has been accomplished, there are plenty of improvements that could be done in order to achieve better results. We do not have time-series resources to reason about the moving or static condition of a vehicle, but we would combine these raster images with vector data that exploits common geometric properties of roads in maps for filtering those parked vehicles, which are not relevant to measure economic activities.

Satellites

Name	Orbit type	Orbit (km)	Days	Start	End	Organisation	Resolution (m)
Landsat 1	Sun Synchronous	907	18	1972	1978	NASA - USA	80
Landsat 2	Sun Synchronous	908	18	1975	1982	NASA - USA	80
Landsat 3	Sun Synchronous	915	18	1978	1983	NASA - USA	30
Landsat 4	Sun Synchronous	705	16	1982	2001	NASA - USA	30
Landsat 5	Sun Synchronous	705	16	1984	2013	NASA - USA	30
RADARSAT-1	Sun Synchronous	798	24	1995	2013	CSA - Canadian Space Agency	
Landsat 7	Sun Synchronous	705	16	1999		NASA - USA	15
Terra	Sun Synchronous	705	16	1999		NASA - USA	15
EO-1	Sun Synchronous	705	16	2000	2017	NASA - USA	
Aqua	Sun Synchronous	705		2002		NASA - USA	
ENVISAT	Sun Synchronous	783	35	2002	2012	ESA	
SPOT 5	Sun Synchronous	832	26	2002	2015	CNES - France	2.5-5
Metop-A	Sun Synchronous	817	29	2006		EUMETSAT	
RADARSAT-2	Sun Synchronous	798	24	2007		MDA - Canada	
GeoEye-1	Other	681		2008		DigitalGlobe	0.46
RapidEye	Sun Synchronous	630	5.5	2008	2020	Planet - USA	5
WorldView-2	Sun Synchronous	770	1.1	2009		DigitalGlobe	0.46
Pleiades 1A	Sun Synchronous	694	26	2011		Airbus Defence and Space	0.5
Metop-B	Sun Synchronous	817	29	2012		EUMETSAT	
Pleiades 1B	Sun Synchronous	694	26	2012		Airbus Defence and Space	0.5
SPOT 6	Sun Synchronous	694	26	2012		Airbus Defence and Space	1.5
CBERS-3	Sun Synchronous	778	26	2013	2013	INPE - Brazil	5
Landsat 8	Sun Synchronous	705	16	2013		NASA - USA	15
PROBA-V	Sun Synchronous	820	1	2013		ESA	
SkySat-1	Sun Synchronous	578		2013		Planet - USA	0.8
CBERS-4	Sun Synchronous	778	26	2014		INPE - Brazil	5
Deimos-2	Sun Synchronous	620	2	2014		DEIMOS - Spain	0.75
Sentinel-1A	Sun Synchronous	693	6	2014		ESA	
SPOT 7	Sun Synchronous	694	26	2014		Airbus Defence and Space	1.5
WorldView-3	Sun Synchronous	617	1	2014		DigitalGlobe	0.31
Planet - Dove	Sun Synchronous	400		2015		Planet - USA	3
Sentinel-2A	Sun Synchronous	786	5	2015		ESA	10
GOES-16	Geostationary	35786		2016		NOAA	
Sentinel-1B	Sun Synchronous	693	6	2016		ESA	
Sentinel-3A	Sun Synchronous	814	27	2016		ESA	
WorldView-4	Sun Synchronous	617	1	2016	2019	DigitalGlobe	0.31
Sentinel-2B	Sun Synchronous	786	5	2017		ESA	
Sentinel-5P	Sun Synchronous	824	17	2017		ESA	
GOES-17	Geostationary	35786		2018		NOAA	
Metop-C	Sun Synchronous	817	29	2018		EUMETSAT	
Sentinel-3B	Sun Synchronous	814		2018		ESA	
VividX2	Sun Synchronous	505		2018		earth-i	

Table A.1: General information about Earth Observation satellites.

Name	Description	Bands	Mission
AATSR	Advanced Along Track Scanning Radiometer	7	ENVISAT
ABI	Advanced Baseline Imager	16	GOES-16 (GOES-EAST), GOES-17 (GOES-WEST), GOES-T, GOES-U
AIRS	Atmospheric Infrared Sounder	7	Aqua
ALI	Advanced Land Imager	9	EO-1
AMSR-E	Advanced Microwave Scanning Radiometer for EOS	6	Aqua
AMSU-A	Advanced Microwave Sounding	15	Aqua, Metop-A, Metop-B, Metop-C, NOAA-15, NOAA-16, NOAA-17, NOAA-18, NOAA-19
ARGOS-3 (A/DCS)	Advanced Data Collection System	6	Metop-A, Metop-B, Metop-C, NOAA-19, SARAL/AltiKa
ASAR	Advanced Synthetic Aperture Radar	1	ENVISAT
ASCAT	Advanced scatterometer	1	Metop-A, Metop-B, Metop-C
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer	14	Terra
AVHRR/3	Advanced Very High Resolution Radiometer	6	Metop-A, Metop-B, Metop-C, NOAA-15, NOAA-16, NOAA-17, NOAA-18, NOAA-19
CERES	Clouds and Earth's Radiant Energy System	3	Aqua, NOAA-20 (JPSS-1), Terra, TRMM
DORIS	Doppler Orbitography and Radiopositioning Integrated by Satellite	2	ENVISAT, Jason-1, Jason-2, SARAL/AltiKa, SPOT 2, SPOT 3, SPOT 4, SPOT 5
ETM	Enhanced Thematic Mapper	8	Landsat 7
EXIS	Extreme Ultra Violet and X-ray Irradiance Sensor	2	GOES-16 (GOES-EAST), GOES-17 (GOES-WEST), GOES-T, GOES-U
GIS-MS	Visible and NIR	4	GeoEye-1
GIS-PAN	Panchromatic	1	GeoEye-1
GLM	Geostationary Lightning Mapper	1	GOES-16 (GOES-EAST), GOES-17 (GOES-WEST), GOES-T, GOES-U
GOME-2	Global Ozone Monitoring Experiment-2	4	Metop-A, Metop-B, Metop-C
GOMOS	Global Ozone Monitoring by Occultation of Stars	3	ENVISAT
HiRAIS	High Resolution Advanced Imaging System	5	Deimos-2
HiRI	High Resolution Optical Imager	5	Pleiades 1A, Pleiades 1B
HIRS/4	High Resolution Infra Red Radiation Sounder	20	Metop-A, Metop-B, Metop-C, NOAA-18, NOAA-19
HSB	Humidity Sounder for Brazil	4	Aqua
Hyperion	High resolution hyperspectral imager with 220 spectral bands	220	EO-1
IASI	Infrared Atmospheric Sounding Interferometer	3	Metop-A, Metop-B, Metop-C
IRMSS-2	Infrared Multispectral Scanner-2	4	CBERS-3, CBERS-4
LAC	LEISA Atmospheric Corrector	256	EO-1
MERIS	Medium Resolution Imaging Spectrometer Instrument	15	ENVISAT
MHS	Microwave Humidity Sounder	5	Metop-A, Metop-B, Metop-C, NOAA-18, NOAA-19
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding	5	ENVISAT
MISR	Multi-angle Imaging SpectroRadiometer	4	Terra
MODIS	Moderate Resolution Imaging Spectroradiometer (PFM on Terra, FM1 on Aqua)	36	Aqua, Terra
MOPITT	Measurement of carbon monoxide in the troposphere	64	Terra
MSI	Multispectral Imager (Sentinel-2)	13	Sentinel-2A, Sentinel-2B, Sentinel-2C, Sentinel-2D
MSS (LS 1-3)	Multispectral Scanner - Landsat 1,2,3	5	Landsat 1, Landsat 2, Landsat 3
MSS (LS 4-5)	Multispectral Scanner - Landsat 4,5	4	Landsat 4, Landsat 5
MUXCam	Multispectral Camera	4	CBERS-3, CBERS-4
MWR	Microwave Radiometer	2	ENVISAT, ERS-1, ERS-2, Sentinel-3A, Sentinel-3B, Sentinel-3C, Sentinel-3D
NAOMI	New AstroSat Optical Modular Instrument	5	AlSat-2A, AlSat-2B, SPOT 6, SPOT 7, VNREDSat-1A
OLCI	Ocean and Land Color Instrument	21	Sentinel-3A, Sentinel-3B, Sentinel-3C, Sentinel-3D
OLI	Operational Land Imager	9	Landsat 8
PanMUX	Panchromatic and Multispectral Camera	4	CBERS-3, CBERS-4
PlanetScope	Planet.com, Dove satellites	4	Planet - Dove
RA-2	Radio Altimeter	3	ENVISAT
RADARSAT 2	Radar	1	RADARSAT-2
RBV	Return Beam Vidicon Camera	4	Landsat 1
SAR-C Radarsat1	Synthetic Aperture Radar on RADARSAT-1	1	RADARSAT-1
SAR-C Sentinel1	C-band SAR on Sentinel-1A/Sentinel-1B	1	Sentinel-1A, Sentinel-1B, Sentinel-1C, Sentinel-1D
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Chartography	8	ENVISAT
SLSTR	Sea and Land Surface Temperature Radiometer	11	Sentinel-3A, Sentinel-3B, Sentinel-3C, Sentinel-3D
SpaceView 110 Imaging System	Formerly GIS-2, GeoEye Imaging System-2	5	WorldView-4
SRAL	SAR Radar Altimeter	2	Sentinel-3A, Sentinel-3B, Sentinel-3C, Sentinel-3D
SUVI	Solar Ultra Violet Imager	6	GOES-16 (GOES-EAST), GOES-17 (GOES-WEST), GOES-T, GOES-U
TIRS	Thermal Infrared Sensor	2	Landsat 8
TM	Thematic Mapper	7	Landsat 4, Landsat 5
Tropomi	Tropospheric Monitoring Instrument (O3, NO2, SO2, HCHO, CO, CH4)	8	Sentinel-5P
UHD Camera	Ultra High Definition camera (video up to 25fps and stills)	3	VividX2
WFI-2	Wide Field Imager - 2	4	CBERS-3, CBERS-4
WV-3 CAVIS	WorldView-3: Cloud, Aerosol, water Vapor, Ice, Snow	12	WorldView-3
WV-3 MSS	WorldView-3: Multi spectral sensor	8	WorldView-3
WV-3 PAN	WorldView-3: Panchromatic sensor	1	WorldView-3
WV-3 SWIR	WorldView-3: Shortwave Infrared sensor	8	WorldView-3
WV110	Standard 4 colors + New 4 colors	8	WorldView-2
WV60	PAN band for WorldView -2	1	WorldView-2

Table A.2: General information about Earth Observation sensors.

Bibliography

- [1] E. Amador, R. Valle, J. M. Buenaposada, and L. Baumela. Benchmarking head pose estimation in-the-wild. In *Proc. Iberoamerican Congress on Pattern Recognition*, pages 45–52, 2017. [1](#)
- [2] S. M. Azimi, R. Bahmanyar, C. Henry, and F. Kurz. EAGLE: Large-scale vehicle detection dataset in real-world scenarios using aerial imagery. *CoRR*, abs/2007.06124, 2020. [12](#), [21](#), [25](#), [27](#), [32](#)
- [3] G. Christie, N. Fendley, J. Wilson, and R. Mukherjee. Functional map of the world. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 6172–6180, 2018. [12](#)
- [4] A. V. Etten, D. Lindenbaum, and T. M. Bacastow. SpaceNet: A remote sensing dataset and challenge series. *CoRR*, abs/1807.01232, 2018. [13](#)
- [5] R. Gupta, B. Goodman, N. Patel, R. Hosfelt, S. Sajeew, E. Heim, J. Doshi, K. Lucas, H. Choset, and M. E. Gaston. Creating xBD: A dataset for assessing building damage from satellite imagery. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 10–17, 2019. [3](#), [13](#)
- [6] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016. [23](#)
- [7] M. Hsieh, Y. Lin, and W. H. Hsu. Drone-based object counting by spatially regularized regional proposal network. In *Proc. International Conference on Computer Vision*, pages 4165–4173, 2017. [12](#), [21](#), [23](#), [25](#)
- [8] N. Jean, M. Burke, S. M. Xie, W. M. Davis, D. B. Lobell, and S. Ermon. Combining satellite imagery and machine learning to predict poverty. *Science*, 353:790–794, 2016. [2](#), [9](#)
- [9] 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*, abs/1802.07856, 2018. [13](#), [21](#), [25](#), [32](#)
- [10] V. S. Lempitsky and A. Zisserman. Learning to count objects in images. In *Proc. Annual Conference on Neural Information Processing Systems*, pages 1324–1332, 2010. [23](#)
- [11] C. Li, C. Xu, Z. Cui, D. Wang, Z. Jie, T. Zhang, and J. Yang. Learning object-wise semantic representation for detection in remote sensing imagery. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 20–27, 2019. [23](#), [25](#), [32](#)
- [12] K. Liu and G. Mátyus. Fast multiclass vehicle detection on aerial images. *IEEE Geoscience and Remote Sensing Letters*, 12(9):1938–1942, 2015. [12](#), [21](#), [23](#)
- [13] W. Liu, M. Salzmann, and P. Fua. Context-aware crowd counting. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 5099–5108, 2019. [23](#)
- [14] P. D. López, R. Valle, and L. Baumela. Facial landmarks detection using a cascade of recombinator networks. In *Proc. Iberoamerican Congress on Pattern Recognition*, pages 575–583, 2018. [1](#)

Bibliography

- [15] A. Lorente, K. Boersma, H. Eskes, J. P. Veeffkind, J. van Geffen, M. B. de Zeeuw, H. A. D. van der Gon, S. Beirle, and M. I. Król. Quantification of nitrogen oxides emissions from build-up of pollution over Paris with TROPOMI. *Scientific Reports*, 9, 2019. [4](#)
- [16] T. N. Mundhenk, G. Konjevod, W. A. Sakla, and K. Boakye. A large contextual dataset for classification, detection and counting of cars with deep learning. In *Proc. European Conference on Computer Vision*, pages 785–800, 2016. [12](#), [21](#), [23](#), [25](#)
- [17] S. Razakarivony and F. Jurie. Vehicle detection in aerial imagery : A small target detection benchmark. *Journal of Visual Communication and Image Representation*, 34:187–203, 2016. [12](#), [21](#), [23](#)
- [18] S. Ren, K. He, R. B. Girshick, and J. Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 39(6):1137–1149, 2017. [23](#), [26](#)
- [19] R. Valle, J. M. Buenaposada, and L. Baumela. Cascade of encoder-decoder CNNs with learned coordinates regressor for robust facial landmarks detection. *Pattern Recognition Letters*, 136:326–332, 2020. [1](#)
- [20] R. Valle, J. M. Buenaposada, A. Valdés, and L. Baumela. Head-pose estimation in-the-wild using a random forest. In *Proc. Articulated Motion and Deformable Objects*, pages 24–33, 2016. [1](#)
- [21] R. Valle, J. M. Buenaposada, A. Valdés, and L. Baumela. A deeply-initialized coarse-to-fine ensemble of regression trees for face alignment. In *Proc. European Conference on Computer Vision*, pages 609–624, 2018. [1](#)
- [22] R. Valle, J. M. Buenaposada, A. Valdés, and L. Baumela. Face alignment using a 3D deeply-initialized ensemble of regression trees. *Computer Vision and Image Understanding*, 189:102846, 2019. [1](#)
- [23] G. Xia, X. Bai, J. Ding, Z. Zhu, S. J. Belongie, J. Luo, M. Datcu, M. Pelillo, and L. Zhang. DOTA: A large-scale dataset for object detection in aerial images. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 3974–3983, 2018. [11](#), [21](#), [25](#), [27](#)
- [24] X. Yang, J. Yang, J. Yan, Y. Zhang, T. Zhang, Z. Guo, X. Sun, and K. Fu. SCRDet: Towards more robust detection for small, cluttered and rotated objects. In *Proc. International Conference on Computer Vision*, pages 8231–8240, 2019. [23](#), [25](#), [26](#), [32](#)
- [25] S. W. Zamir, A. Arora, A. Gupta, S. H. Khan, G. Sun, F. S. Khan, F. Zhu, L. Shao, G. Xia, and X. Bai. iSAID: A large-scale dataset for instance segmentation in aerial images. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 28–37, 2019. [13](#)