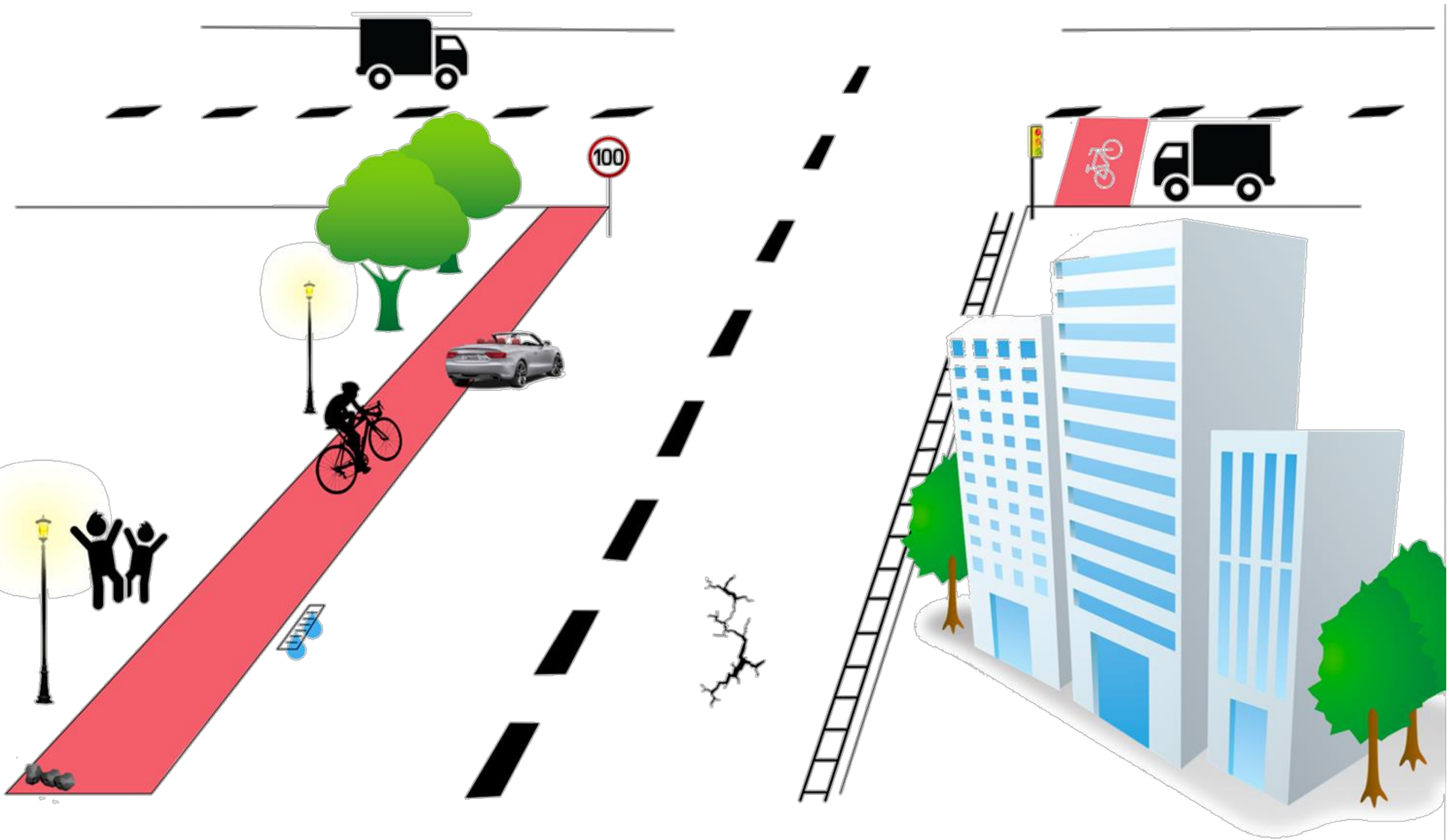
Imperial College
London

Using Deep Learning to Identify Cyclists' Risk Factors in London

MRes Biomedical Research (Data Science) Standard Project

Luís Rita, Ricky Nathvani & Majid Ezzati

20 August 2020



I certify that this thesis, and the research to which it refers, are the product of my own work, conducted during the current year of the MRes in Biomedical Research at Imperial College London. Any ideas or quotations from the work of other people, published or otherwise, or from my own previous work are fully acknowledged in accordance with the standard referencing practices of the discipline.

Luís Rita

Abstract

Cycling encompasses many societal benefits. It influences a community's safety, economy, environment, equity and health. The number of cyclists on the roads is highly influenced by their perception of safety. To determine road safety, it is fundamental to have a common metric, so that risk factors can be determined and compared. Using Google Street View (GSV) imagery is a cost-effective approach to analyse urban environments. Due to the high number of images needed to extract accurate results, models to automatically detect objects and structures are used.

The aim of this project was to use object detection and image segmentation models to extract cyclists' road risk factors from GSV images of London. This involved compiling road safety indicators and risk factors; analysing a GSV dataset, before using two state-of-the-art tools, YOLOv5 and PSPNet101, to detect objects and segment images, respectively, and further analysing their results; determining the limitations of YOLOv5, PSPNet101 and suggesting ways of making cyclists' safety assessment more accurate.

Approximately 2 million objects were identified, and 200 billion pixels labelled in the 500 000 images available in the imagery dataset. On average, there were 108 images per Lower Layer Super Output Area (LSOA). Using YOLOv5, the distribution of the following risk factors was (in)directly identified at an LSOA level: high vehicle speed, tram/train rails, truck circulation, parked cars and pedestrian presence. Statistically significant negative correlations between cars x buses, cars x cyclists and cars x people (strongest) were found. And positive correlations between people x bus and people x bicycles (strongest). Long-tail distribution on the number of heavy-vehicles was observed. Using PSPNet101, building (27%), sky (22%) and road (22%) pixels

were the most common. Thus, objects in any of these areas can be equally detected. All results and implementations were made available in the project's repository.

Future developments include increasing the availability and resolution of GSV images. Training YOLOv5 and PSPNet101 with bigger datasets containing more categories. Defining a safety metric to account simultaneously for detected objects and segmented structures.

Deep Learning | **Cycling** | **Risk Factors** | **London**

Acknowledgments

Deeply grateful to my supervisors Majid Ezzati and Ricky Nathvani. And all others from Imperial's School of Public Health, including Barbara Metzler, Emily Muller, Esra Suel and Theo Rashid. Kavi Bhalla (University of Chicago), Jill Baumgartner (McGill University) and Michael Brauer (University of British Columbia) have importantly contributed.

☆ Dimitris ☆ Jordi

1. Introduction

Cycling comprises many benefits for societies. It influences our safety, economy, environment, equity and health [1]. Although there are countries where most of the daily commuting is done by bicycle, in others, it is rarely used [2].

In the actual Covid-19 pandemic context, governments are promoting cycling as an alternative to driving or, even, crowded public transports. Not only it is known more polluted areas to pose an additional infection risk to their inhabitants, but also public transports where it may be hard to assure social distancing. For these reasons, UK government is boosting this sustainable way of transportation with a 2-billion-pound package [3].

The number of cyclists on the roads is highly influenced by their perception of safety [4]. To determine road safety, it is fundamental to have a clear common metric so that risk factors can be determined and compared.

Google Street View (GSV) imagery is publicly available, plus, it covers most of the developed countries (Figure 1) [5]. For this reason, this is a cost-effective approach to analyse city environments.

Due to the high number of images needed to extract accurate results, datasets and models trained on them are used to automatically detect objects and other structures present. Object detection (OD) [6] and image segmentation (IS) [7] tools are used with this end.

The goal of this project was to compile a list of the most relevant risk factors for cyclists in London, based on specific safety metrics such as accident, injury and fatality rates. Using a GSV image dataset, to extract the identified risk factors from Greater London using OD and IS models. To study, at a Lower Layer Super Output Area (LSOA) level, how the different safety factors are distributed across London. To identify correlations among the most detected objects. To detect the most common misclassifications done by both algorithms and suggest ways to mitigate them, after individually analysing images from all LSOAs. Finally, to provide new guidelines on how OD and IS models can detect additional road safety risk factors based on the experience of this project.

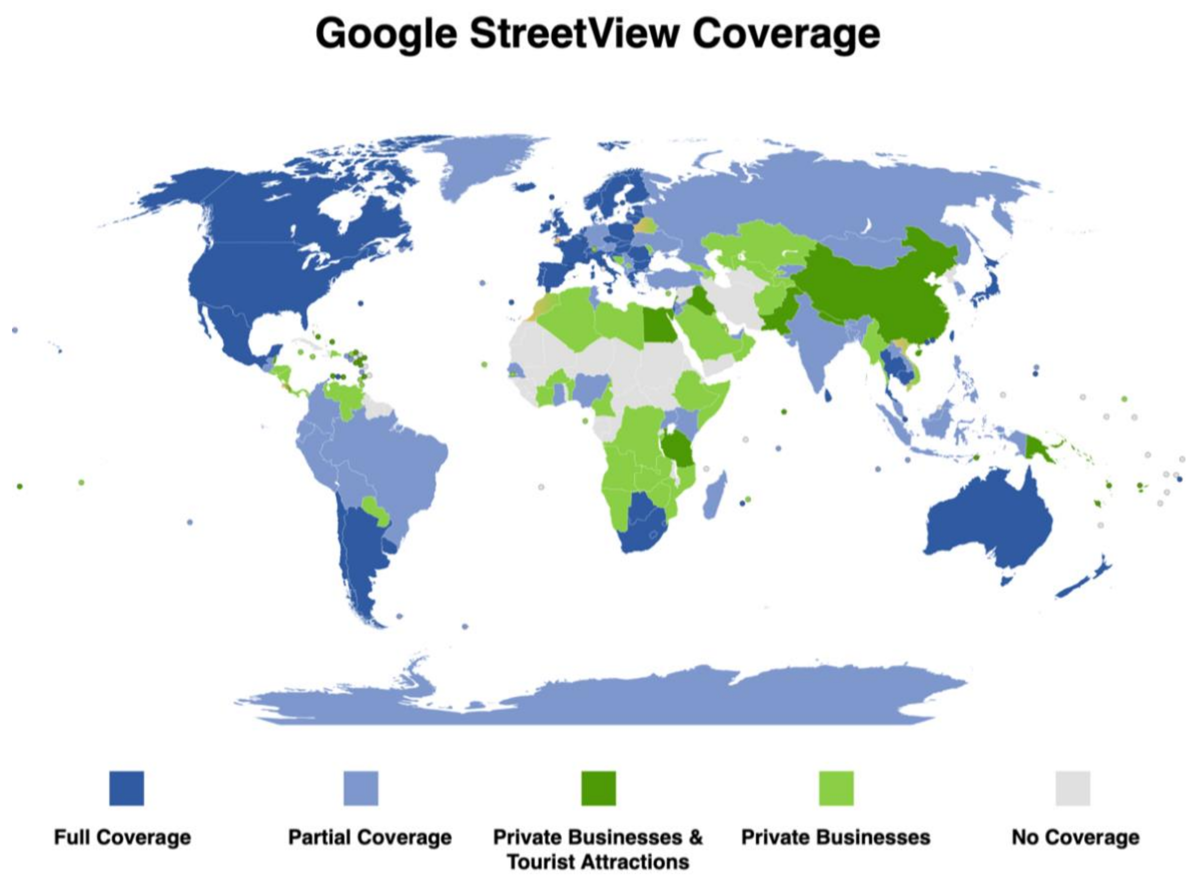


Figure 1 GSV world coverage. Adapted from [5].

In the next section, concepts crucial for the understanding of this project’s goals and results are introduced.

2. Background

Cycling benefits are first introduced. Then, road safety indicators are explored, before identifying cyclist's risk factors in London and defining how they will be captured using OD and IS methods. YOLOv5 (OD) and PSPNet101 (IS) models along with their respective datasets used to train them are detailed.

2.1. Cycling Benefits

This section includes an overview on the most important societal benefits associated to cycling. They can be categorized in five main domains: safety, economy, environment, equity and health [1].

The main cause of death in the USA in youngers is traffic accidents. Accounting for 41% of the total number of deaths in the age group 15 to 24 (CDC). In the EU, in the past 10 years, deaths among cyclists remained constant, while for car drivers and passengers had decreased 24%. Among pedestrians fell by 19% [8]. UK was one of the only 3 EU nations which fatality rate among pedestrians increased, 1.3% a year. On average, it fell 2.6% in EU per year. For cyclists, in the UK, the number of fatalities decreased 1.3%, with the 13th-best average annual drop. Given that 99% of the pedestrians killed were struck by motor vehicles and 1% by bikes, it is evident the necessity of promoting cyclists' safety and increase their number in the streets. In parallel, this promotes safety in numbers: cyclists are safer if their number increases. The awareness drivers develop by contacting more frequently with cyclists is the root cause.

Many economic benefits for individuals, companies and communities are known from promoting walking and cycling as alternative ways of transportation [9]. According to 2015 Urban Mobility Scorecard, the cost of congestion for the US in 2014 was 160 billion dollars. For an individual was estimated to be 960 dollars each year. These values account for time and fuel expenses. While the cost of having a car in the USA in 2018 was calculated to be 8849 dollars, for a bike was 308 dollars and walking was considered free. 2018 Benchmarking Report adds that bike tourism has a positive economic impact in multiple regions worldwide. Protected Bike Lanes Mean Business Report shows the positive impact cycling may have in business. It was found that workers cycling in their commutes, on average, spent more time and travelled more often to their companies, then car drivers. Moreover, in a 2011 study, the Political Economy Research Institute found that 11.41 jobs were on average created when investing 1 million pounds in bicycle-only projects, comparing to 7.75 jobs while investing the same amount in road-only projects.

Reducing the dependency on non-renewable sources is one of the positive aspects of cycling. It was estimated by the United States Environmental Protection Agency that the transportation sector was responsible by the largest share of greenhouse gas emissions – 28% (in the USA). From those, passenger vehicles and light-duty trucks account for most of the overall transport sector – 60%. Moreover, it is known structures such as roads and parking lots increase significantly the probability of urban flooding, stormwater runoff and urban heat island effect (due to the lack of shadows and exposed land in the cities, they often register higher temperatures than the surroundings, resultant from the low levels of air humidity). Promoting cycling will reduce the need of the previous infrastructures and mitigate some of the negative consequences.

Promoting cycling in a country promotes equity. Due to the high cost of car ownership, when a city prioritizes road infrastructure for these vehicles, it puts in higher risk low-income families that cannot afford it. This is particularly important in low-income communities where a brief from Bridging the Gap estimated that only 50% of the roads have sidewalks, comparing to 90% in high-income homologous. This results in higher threat for pedestrians and cyclists. The New Majority: Pedalling Towards Equity reported that 26% of people of colour would like to cycle but do not do it due to safety concerns, comparing to 19% in white respondents.

Physical activity, such as cycling and walking, has numerous benefits to physical and mental health [9]. Centre for Disease Control and Prevention reported that 1 in 10 premature deaths, 1 in 8 breast cancers, 1 in 8 cases of colorectal cancer, 1 in 12 people suffering from diabetes and 1 in 15 cases of heart disease could be prevented if citizens became more active. It can also reduce the risk for coronary heart disease, stroke and many respiratory chronic diseases, which are intimately related to air quality [9]. 2018 State of the Air report states over 133.9 million Americans live in counties with unhealthy levels of ozone and/or particle pollution. A factor that is highly influenced by the transportation patterns inside a community. Vehicles are one of the main contributors accordingly to the United States Environmental Protection Agency. Due to all these benefits, it becomes evident the importance of promoting a less sedentary lifestyle among the population.

By promoting cyclists' safety, the whole society benefits. Cyclists and pedestrians will be safer. Drivers will reduce their commute times and translate that into society gains such as lower levels of pollution and economic losses.

2.2. Road Safety Indicators

Road safety indicators are essential for policy making. According to the European Road Safety Charter, they allow us to assess the current situation of the roads, observe the impact on accident rates after an intervention, monitor its progress over time and predict further evolutions.

To be useful, road safety indicators should comply with several criteria:

1. Relate to some aspect of road safety, such as the causes or consequences of a road accident;
2. Be measurable in a reliable way;
3. Be monitorable over time;
4. Allow road safety engineers or public health experts to set targets;
5. Be useful for establishing comparisons and benchmarking different safety performances.

There are six dimensions common to all indicators: geographical scope, time span, numerical format, representation/visualization, reliability, accuracy, representativeness and a specific "level" of road safety. The first encompasses where the measurement takes place: organisation, city, region, country, Europe or global. The second relates to the time frame comprehensive to the analysis: day, week, month, quarter, year, decade or longer. The units of the measurement are represented by the third feature. They can be a proportion, a percentage or some other well-defined ratio. Representation described the way in which data is presented, for example, in the form of a map, graph or table. Reliability, accuracy and representativeness are linked to the design and implementation of the measurement system. Finally, the "level" of the indicator differs on whether it considers one of the following: impact of the crashes, post-crash response, crash outcomes, crash causes and predictors, road safety policy and measures, or safety culture and safety systems.

Crash outcomes, including indicators such as mortality, severely/slightly injured and accident rates, were the ones considered while selecting the risk factors presented in the next section.

Road safety indicators were introduced so that a clear ranking of the most relevant cyclist's risk factors could be established

2.3. Risk Factors

To list and order the most relevant risk factors for cyclists in London, accident, injury and fatality rates were considered. In London, the fatality rate for cyclists is relatively low (Figure 2), consequently, priority was given to the other two more discriminative ones. Accident and injury rates were used to order all risk factors when designing Figure 21 diagram. Note there is also a strong qualitative and experience-based component inherent to these rankings, once a common safety metric to all risk factors was not found.

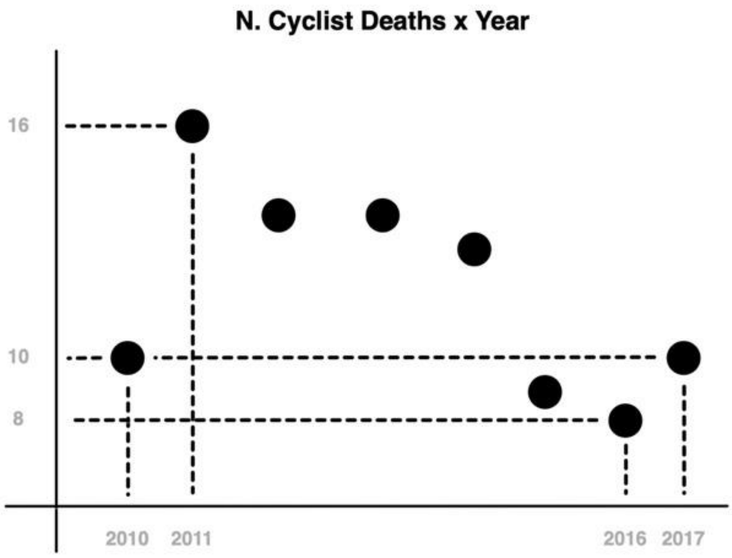


Figure 2 Number of cyclists' deaths in Greater London between 2010 and 2017. Data from [10].

The top 3 most relevant factors identified that influence cyclists' safety was the presence of a cycle lane, road speed limits and road lane width. Next, statistical data that supports the rankings defined in Figure 21 are provided. Note that only a small number of accidents involving cyclists are reported [11], consequently, the statistics presented in the next paragraphs may not totally reflect a real-life scenario.

2.3.1. Cycle Lane and Parked Cars

Cycle lanes can be physically separated or located on the road.

Physically separated lanes reduce the probability of crash when a car tries to overtake a cyclist or, in the case of fall, to be hit. One of the main causes of injury among cyclists are falls caused by bad pavement quality [12] [13] [14]. With no cars parked in the surroundings, these lanes reduce the risk of injury among cyclists by half. [15] For all these reasons, risks associated to high speed road limits, narrow bicycle lane widths, road pavement quality and parked cars were not considered.

In the case of an on-road cycle lane, vehicular speeds tend to be lower and there are less interactions between these and the cyclists, when comparing to no lane [16].

This makes the first scenario the safest, followed by on-road and no cycle lane.

The presence of a cycle lane was considered the most decisive factor by preventing many of the previously identified risks. It was considered the number one in the rankings of risk factors.

2.3.2. Vehicle Speed

Speed was found to be one of the major factors involved in around 10% of all accidents and 30% of the fatal ones. Speed of vehicles involved in a crash is the single most important factor in determining the severity of injuries. [17]

There are two distinct factors when considering speed. Not only higher speeds are known to be responsible for an increased rate of accidents, injuries and deaths, but also large speed differences. Roads with high speed variance are more unpredictable, once they favour the number of encounters and an increased number of overtaking manoeuvres. Consequently, reducing speed limits sometimes may only result in the decrease of the vehicles' average speed and not its variance. [18]

In the core of the danger posed by vehicles' high speeds are the increase in the braking distance and kinetic energy that is transferred from the vehicle to the cyclist. Once both increase with the square of the velocity, the possibility of avoiding or surviving a crash decreases quadratically. [18]

In a biological perspective, it is known the human body can only resist the transference of a limited amount of kinetic energy in a crash. [19] This amount varies for different body parts, age groups and gender. Considering the best-designed car, if the vehicle exceeds 30 km/h, this limit can be exceeded. [20] Studies also show if a car travels at a speed lower than 30 km/h, the probability for a pedestrian to survive a crash is higher than 90%. When hit by a car at 45 km/h, the chance of surviving decreases to 50%. [21] Or, as the speed of a car rises from 30 km/h to 50 km/h, the probability of surviving a crash decreases by a factor of 8. [22]

This was considered the second most relevant factor. In the case of an on-road cycle lane next to a low-speed limit road, the risk factors related to parallel traffic were considered negligible, regardless of the width of the lane.

2.3.3. Lane Width

In the United Kingdom, recommended cycle lane width is 2 meters. Minimum required is 1.5 meters. All values below 1.5 m are considered too narrow, allowing little room to manoeuvre around obstacles, such as: debris, potholes and water drainers. It is frequently referred it is safer not having a cycle lane, than one that is too narrow, once motorists tend to drive right up to the line and cyclists too close to the kerb.

After the road speed limits, cycle lane width was considered the following most important factor. Whenever it is considered wide, traffic risk factors were not considered. Regardless of the width of an on-road cycle lane, low-speed limits were enough to discard all traffic related risk factors.

2.3.4. Streetlight

After the top 3, streetlight was considered the most relevant criterion in determining road safety. It is known to affect drivers and cyclists' reaction time and make cyclists unnoticeable particularly when not using any reflective or luminous gear. Moreover, in the cyclists' perspective, they are less aware of other road risks associated for example to the quality of the pavement. It is not expected to encounter many significantly under illuminated roads in London, consequently, it was placed in fourth place of the ranking.

2.3.5. Pavement Quality, Tram/Train Rails and Water Drainers

As frequently referred in literature, pavement quality is a crucial factor to consider when evaluating safety. [12] [13] [14] Pavement quality refers to the quality of the road when there is no cycle lane, or to the cycle lane itself, when it is present. Along with the presence of water drainers and trails, these were the following most important risk factors. This was placed under streetlight once with enough luminosity and circulating at a moderate velocity, it should not pose a significant threat.

2.3.6. Number Intersections and Intersections Visibility

The majority of bike and car crashes occur in intersections. In [23], the reported percentage was 60% over the total number of crashes. Additionally, as part of the same study, intersections where streets do not meet at right angles, posed an additional danger to cyclists. Crashes at these areas were 31% more likely to cause serious injury to the cyclist. The main reason is the decreased intersection visibility.

2.3.7. Lorries and other Large Vehicles

In the last years, economic development and consumer demand have been increasing, and so as the number of trucks in the cities. [24] [25] While cycling has following the same trend, the number of encounters among them has significantly increased. As an example, in New York City, 15% of bicycle networks overlap with 11% of truck networks. [26] The increased number of encounters has contributed to a higher accident and mortality rates involving trucks. Truck-bicycle accidents have usually more severe consequences than any other type of accidents. [27] [28] [29] [30] In some EU countries, 30% of all cycling fatalities are associated to trucks. [31] Studies in the past 2 decades have identified trucks as the most common vehicle category involved in cyclist deaths, in London. [28] [32] [33]

2.3.8. Advanced Stop Line

These lines present in several European countries such as Belgium, Denmark and United Kingdom, allow a head start to certain types of vehicles (namely, bicycles) when the traffic signal changes from red to green. This has several advantages. First, drivers behind the line can clearly realize the presence of cyclists around them and take the right precautions to avoid danger manoeuvres. Second, it becomes safer for a cyclist to turn to the left avoiding a crash with the cars that are behind. In terms of statistical data on accident, injury or mortality rate, it was not found any.

2.3.9. Bend Visibility

Several sources identify bends as a risk factor. Bends and intersections are often jointly considered as posing similar risks to the cyclist. Low visibility, in the cyclist’s perspective, make risky situations that usually are not – sudden presence of pedestrians or intrusive vegetation. In the driver’s perspective, it can make cyclists unnoticeable and, consequently, a vulnerable element. [12] There is no clear statistical data showing how bends affect cyclists' accident, injury or fatality rates.

2.3.10. Pedestrians

Among all age-groups, pedestrian fatalities most often occur in children younger than 14 years old, when comparing with adults aged between 15 and 64 or 65 or more. In terms of gender, men are at a greater risk than women. [34] For these reasons, locations with higher concentration of people satisfying these criteria (e.g. – school areas) are at additional risk. Nevertheless, in car-free zones, accidents between pedestrians and cyclists are extremely rare and almost never serious. [35] This was considered the least important of the risk factors.

Considering **Error! Reference source not found.**, from the left to the right of the diagram, the number of risk factors decreases. The most unsafe situation was the absence of cycle lane. Secondly, high speed limits with a narrow lane in an on-road scenario. Thirdly, high-speed limits but with a wider lane. Fourthly, low-speed limits regardless of the presence of a narrow or wide on-road lane. Finally, a physically separated lane was considered the safest scenario.

In the next two sections are described the tools used to capture objects and structures from imagery.

2.4. Object Detection

OD is a computer technology that closely relates to image processing and computer vision fields. It is used in tasks such as image annotation, activity recognition, face detection, face recognition, video object co-segmentation and object tracking [36].

Every object has a set of special features that helps classifying it as belonging to a specific class. For example, when looking for circles or rectangles, objects that are curve or perpendicular on the corners are searched [36].

Methods for OD fall in one of these two categories: machine-learning or deep-learning based approaches. In the case of the first, a list of the most relevant features to look is defined *a priori*. A support vector machine is one possible example. Contrarily, in a deep-learning approach, end-to-end OD can be performed without specifying the relevant features (YOLOv5). Convolution neural networks are typical examples [36].

2.4.1. YOLOv5

YOLOv5 is the most recent version of YOLO which was originally developed by Joseph Redmon. First version runs in a framework called Darknet which was purposely built to execute YOLO [37].

Version 5 is the second model that was not developed by the original author (after version 4), and the first running in a state-of-the-art machine learning framework – PyTorch [6].

YOLOv5 GitHub repository contains a pre-trained model in the MS Coco dataset. Plus, benchmark tests (Figure 3) on the same dataset and detailed documentation on how to execute or retrain it using different data.

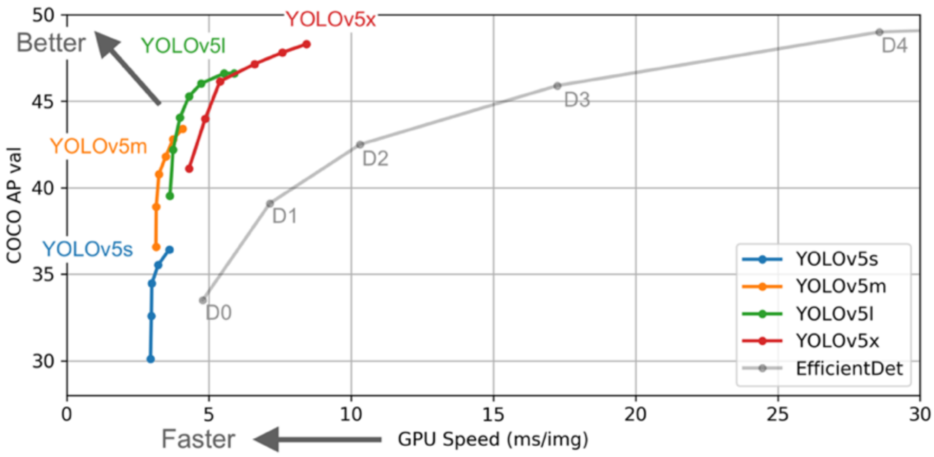


Figure 3 The most up to date YOLO model is the version 5 (July 2020). It was released with 4 different sets of weights varying in accuracy and storage requirements. The presence of *EfficientDet* (the most accurate OD model) highlights the speed of detection of YOLOv5, while keeping the same high accuracy. [6]

In Table 1, there is a comparison in terms of precision, speed and storage requirements for each YOLOv5 set of available weights.

Table 1 Specifications for all sets of weights released with YOLOv5. Generally, as average precision increases, more processing power is required from the GPU to be executed. Adapted from [6].

Model	AP^{val}	AP^{test}	AP_{50}	$Speed_{GPU}$	FPS_{GPU}	Params	FLOPS	Weights Size (MB)
YOLOv5s	36.6	36.6	55.8	2.1	476	7.5M	13.2B	14
YOLOv5m	43.4	43.4	62.4	3.0	333	21.8M	39.4B	42
YOLOv5l	46.6	46.7	65.4	3.9	256	47.8M	88.1B	92
YOLOv5x	48.4	48.4	66.9	6.1	164	89.0M	166.4B	170

The architecture of YOLOv5 consists in three important parts, as in any single-stage object detector: model backbone, neck and head. The first is used to extract the main features of a given input image. In version 5, Cross Stage Partial Networks are used. These have shown significant improvements in processing time with deeper networks. Model neck PANet was used to obtain feature pyramids and helps generalizing the model on object scaling. The final detection part is performed by the head of the model (same as in YOLOv3 and YOLOv4). It applies anchor boxes on features and generates output vectors including class probabilities and bounding boxes.

Each potential detection has an associated confidence score. This indicates how certain is the model about the presence of an object inside the bounding box and, at the same time, whether the box is capturing it correctly.

Due to the high accuracy and speed of YOLOv5x, this model was chosen to process the GSV dataset.

2.5. Image Segmentation

In computer vision and image processing, IS refers to the process of partitioning a digital image into multiple segments or sets of pixels. The goal is to simplify image representation to the point that multiple structures can be easily retrieved. More precisely, IS is the process of assigning one label to every pixel in an image and the ones with the same label share some characteristic. Consequently, this method provides information on the presence of certain structures, their shape and location in the image [38].

2.5.1. PSPNet101

Pyramid Scene Parsing Network (PSPNet) is one of the most accurate IS models. It won ImageNet Scene Parsing Challenge 2016, PASCAL VOC 2012 benchmark and Cityscapes benchmark. It achieved a mIoU accuracy of 85.4% on PASCAL VOC 2012 and 80.2% on Cityscapes [7]. Since that time, segmentation models' accuracy has reached a plateau in the 2 previous years (Figure 4).

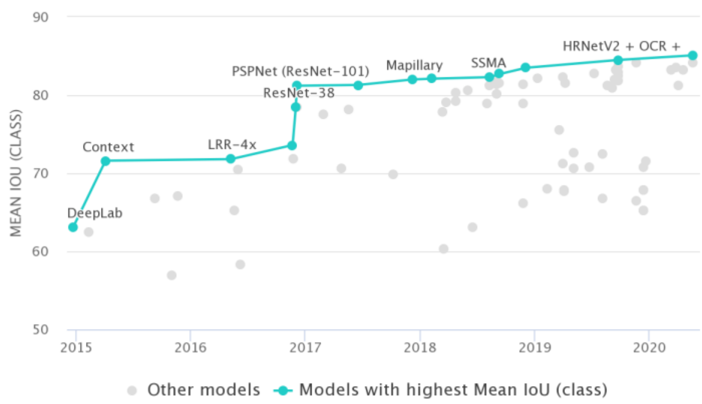


Figure 4 Evolution of IS models over time. Since 2017, the increase in the mean IoU has been very small. [39]

After receiving an input image, PSPNet executes a Convolutional Neural Network (CNN) to extract a feature map from the last convolutional layer. Then, a pyramid parsing module is used to harvest different sub-region representations, followed by up sampling and concatenation layers to create the final feature representation. This carries both local and global context information. In the last step, the representation is fed into a convolution layer and the final per-pixel prediction obtained.

Due to the long execution times of IS models, it was chosen the fastest, with the highest precision and best documentation.

2.6. Training Datasets

There are multiple available datasets with labelled objects and segmented images. For OD model training and benchmarking, two of the most used and containing the highest number of road categories are MS Coco [40] and Open Images V6 [41]. Similarly, for IS, Cityscapes [42] and ADE20K [43] are the current state-of-the-art datasets. In Table 2, all the 4 are represented together and compared in terms of their relevant categories to assess cyclists' road safety.

Table 2 Comparison between 4 of the biggest OD and IS datasets, with relevant data to assess road safety.

	Object Detection		Image Segmentation	
Risk Factor	MS Coco	Open Images V6	Cityscapes	ADE20K
Cycle Lane	-	-	Sidewalk*	-
Streetlight	-	Streetlight*	-	Streetlight* | Street Lamp*
Pedestrians (e.g.: children near schools)	People	Girl | Man | Person	Person	Person | Individual | Someone | Somebody | Mortal
Water Drainers	-	-	-	-

Tram/Train Rails	Train	Train	-	-
Number Intersections	-	-	Sidewalk* | Road	Sidewalk* | Pavement*
Intersections Visibility	-	-	Sidewalk* | Road	Sidewalk* | Pavement*
Bends Visibility	-	-	-	-
Vehicle Speed	Stop Sign* | Traffic Light*	Stop Sign* | Traffic Light* | Traffic Sign*	Traffic Light* | Traffic Sign*	Traffic Light* | Traffic Signal* | Stoplight*
Parked Cars	Car | Parking Meter*	Car | Taxi | Vehicle	Parking*	Car | Auto | Automobile | Machine | Motorcar
Lorries and other Large Vehicles	Bus | Train | Truck	Bus | Train | Van	Bus | Truck | On Rails | Caravan	Truck | Motortruck | Van
Road Width	-	-	Road*	Road* | Route*
Pavement Quality (Pits, Trenches, Tree Root Encroachments)	-	-	-	-
Advanced Stop Line	-	-	-	-

Some of the road safety objects can be directly extracted, others indirectly. The same applies for the segmented structures. Features in Table 2 with no asterisk represent their dynamic behaviour. Consequently, they are expected to vary in number over time within the same location.

Within OD category, the following cyclists risk factors were identified: car and parking meter were used to account for parked cars. Person for pedestrians. Truck and bus for truck circulation. Bicycles for number cyclists. Traffic light and stop sign for vehicle speed, once they are traffic calming factors. Finally, train for the presence of tram/train rails due to the close association. For IS, road, sidewalk and streetlights were used to retrieve road and sidewalk width, and streetlight, respectively.

2.6.1. MS Coco

Microsoft Coco is one of the biggest and most popular datasets used for OD, segmentation, and captioning. It contains 330K images, from which 200K are labelled. 1.5M labelled objects are present and distributed across 80 categories. Figure 5 contains some examples of labelled objects. This dataset contains several classes of everyday objects ranging from home appliances, until the most commonly seen on the roads [40].

Figure 5 (Top) Example of annotated images in the MS Coco dataset. (Bottom) Objects present in each image [40].

2.6.2. Cityscapes

Cityscapes dataset focuses on the semantic understanding of urban street scenes. It contains 20 000 coarse annotated images and 5 000 fine annotations from 50 German, French and Swiss cities. These were captured over several months (summer, spring and fall), good/medium weather conditions and during daytime. 30 classes of structures are present (Figure 6) [42].

Figure 6 Example of three segmented images available in Cityscapes [42].

MS Coco and Cityscapes datasets are the most complementary and with the highest number of relevant objects and road structures. YOLOv5 and PSPNet101 models pre-trained in these two datasets were used to detect objects and segment images, respectively.

3. Methodology

In this section it will be detailed how GSV imagery dataset was storage and processed using YOLOv5 and PSPNet101 models. The choice of the respective parameters. Plus, the software and hardware that was used to execute them.

3.1. GSV Imagery Dataset

Instead of considering a road safety analysis of London, it was chosen to perform it at an LSOA level. Greater London can be divided in multiple smaller areas: Output Area, Middle Layer Super Output Area and LSOA. Each one of these subdivisions of London differs at a geographical scale. A zip folder containing multiple shapefiles for each of these divisions was obtained from London Datastore.

GSV imagery dataset was obtained using Google API before the start of this project by Emily Muller. A separate file associating each image identification to a given London LSOA was also available.

Due to the high memory requirements, all images used in this project were storage in the Imperial servers. Images in the Research Data Store were accessed using *Globus* platform.

3.2. YOLOv5

To execute YOLOv5, Imperial's High-Performance Computing cluster was used. It was remotely accessed using the VPN connection tool – *Tunnelblick*.

Due to the speed of execution of YOLOv5, it was used one single P1000 GPU. The most accurate set of weights was chosen- YOLOv5x.

It was defined a minimum confidence of 0.5 for every detection. Higher than the standard value of 0.4. Only text files containing the detected objects and respective locations were saved. Each line includes a numerical designation for each object and space-separated are the coordinates of the centre of the detection box, along with two values for the width and height of the rectangle.

Using *Matplotlib* and *Seaborn* Python visualization frameworks, it was plotted the correlation matrix for the top 15 most detected objects. Pearson correlation factors and p-values were obtained using *SciPy pearsonr* function.

Pandas DataFrame was used to plot the relative and absolute distribution histograms for the top 15 most detected objects.

Misclassifications, limitations and future directions analysis focused on the objects identified before as relevant for road safety. Moreover, all observations resulted from the individual assessment of one image from all London LSOAs, plus the overall project experience.

3.3. PSPNet101

Although it was already available a preliminary version of the implementation executing PSPNet101, provided by Esra Suel, modifications were made to overcome incompatibilities with the new version of TensorFlow.

IS methods are, generally, slower than OD. Originally it was used Python *multiprocessing* tool to accelerate execution. At the end, the original GSV dataset was split in 13 batches and executed parallelly in 13 P100 GPUs. This way, 13 jobs were submitted to the HPC. P100 was the chosen GPU due to its high processing power for numerical analysis when compared to the others available (Table 3).

Table 3 GPU types available on the Imperial High-Performance Computing cluster. P100 was used in this project.

GPU Type	Single Precision (TFLOPS)	Double Precision (TFLOPS)	Memory (GB)	Memory Bandwidth (GB/s)
P1000	1.8	<<1	4	80
K80	5.6	2.9	12	240
P100	8.0	4.0	16	730
RTX6000	16.3	<1	24	670

All images in the GSV dataset were segmented. After, two Python functions were implemented. One that generates a dictionary linking each RGB colour to a given object class. Another, that receives as input the full dataset of segmented images and outputs the total number of labelled pixels for each category. Relative and absolute distribution of all labels were analysed and represented using *Pandas DataFrame* Python library.

Misclassifications, limitations and future directions analysis focused on the structures identified before as relevant for road safety. Moreover, all observations resulted from the individual assessment of one image from all London LSOAs, plus the overall project experience.

4. Results & Discussion

In this section, results are presented along with their discussion. First, it is provided an overview of the GSV imagery dataset across all London LSOAs. Then, OD and IS outputs are analysed.

4.1. GSV Dataset

GSV dataset contains 518 350 images spread across Greater London. There are 512 812 images identified with a London LSOA. From those, 478 724 are unique (Table 4).

Table 4 Not all images in the GSV dataset are LSOA identified. For this reason, a smaller set was used in the analysis – 478 724.

N. Images GSV Dataset	N. LSOA Identified Images	N. Non-Repeated Identified Images	N. LSOAs with Images
518 350	512 812	478 724	4832

For each datapoint there are 4 images available, ranging from 0 to 360 degrees. There are 119 681 unique LSOA identified points, each with four 90 degrees images (Figure 7).

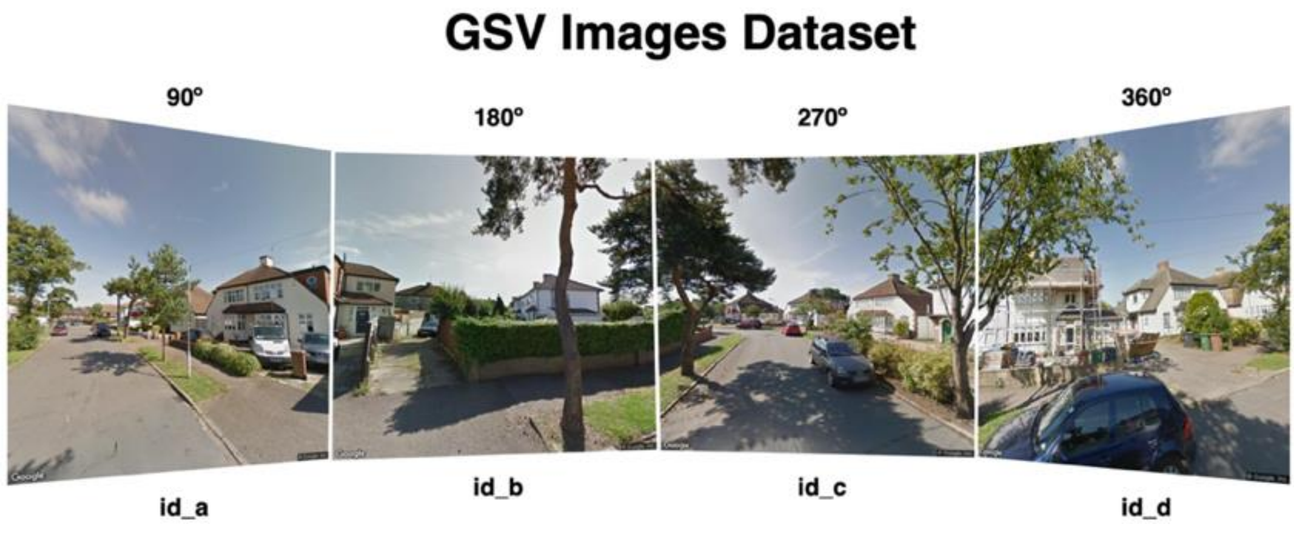


Figure 7 For each datapoint in Figure 8, there are 4 images associated. Each capturing a 90 degrees angle of the surroundings.

There are more images available near Central London, decreasing to the periphery. In Figure 8, it is represented an LSOA atlas, along with the respective geographical distribution of all images.

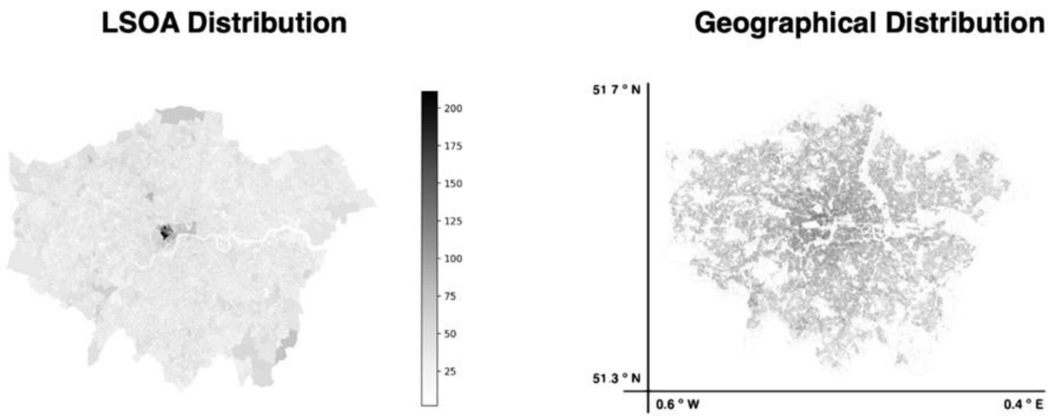


Figure 8 (Left) LSOAs coloured accordingly to the number of available images. (Right) Geographical distribution (latitude and longitude) of the same set of images.

The LSOA with the highest number of datapoints, 211, is in Central London. On average, there are 27 datapoints available per LSOA. There is one LSOA in the dataset with only 1 (Table 5). The wide distribution on the number of images compromises the accuracy of the estimation of the number of objects and segmented structures in less represented LSOAs.

Table 5 Availability of GSV points in the dataset, across all London LSOAs.

Minimum	Maximum	Mean	Standard Deviation	Mode	Median
1	211	27	24	25	11

4.2. Object Detection | YOLOv5

An example of an image after running YOLOv5 is provided in Figure 9. All cars, trucks and people in the image were accurately detected with high confidence values.

Figure 9 Example of a GSV image after executing YOLOv5.

Next, the relative and absolute distributions of all objects are presented at a dataset level.

4.2.1. Dataset Object Distribution

Relative distribution of the top 15 most detected categories of objects was represented in Figure 10.

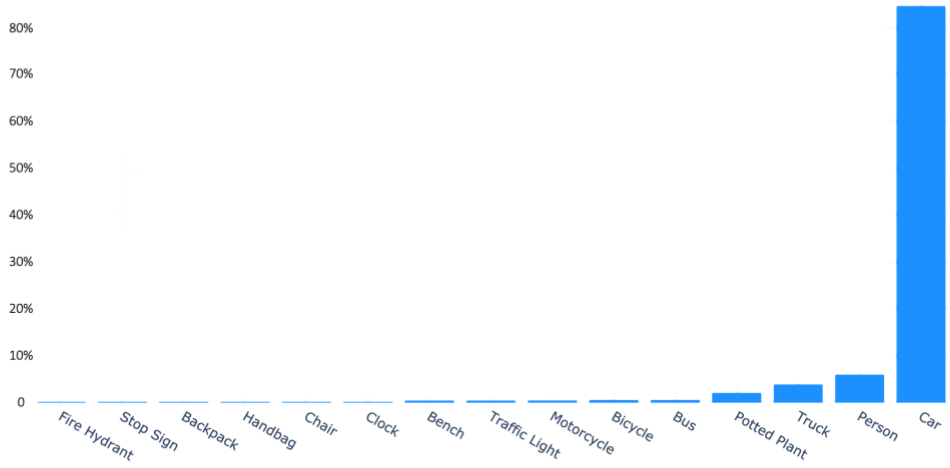


Figure 10 Relative distribution of top 15 detected objects across all LSOAs.

Once the dataset exclusively contains street view images, it was expected to detect a big fraction of cars. London is a city with a high populational density and where public transports are common. This may justify the high number of pedestrians and buses. Potted plant detections are closely related with the number of parks and green areas in London.

Table 6 contains the absolute numbers of the top 15 most common objects detected in the GSV dataset. Highlighted with green are the ones that were identified as contributing positively to cyclists' road safety. Red ones are the negative factors.

In addition to the coloured risk factor objects in Table 6, trains (657) and parking meters (968) were also considered to extract the risk factors in Figure 21.

Table 6 Absolute counting for the top 15 most commonly detected objects. Coloured are the objects that were used to extract cyclists' road safety factors.

Object	Number Detections	Object	Number Detections	Object	Number Detections
Car	1.51M	Bicycle	10.9K	Chair	2.19K
Person	107K	Motorcycle	8.97K	Handbag	2.09K
Truck	70.1K	Traffic Light	6.31K	Backpack	1.94K
Potted Plant	37.9K	Bench	5.01K	Stop Sign	1.28K
Bus	11.5K	Clock	2.75K	Fire Hydrant	1.17K

4.2.2. LSOA Object Distribution

Figure 11 contains the LSOA distribution of the objects identified as relevant to capture cyclists' risk factors.

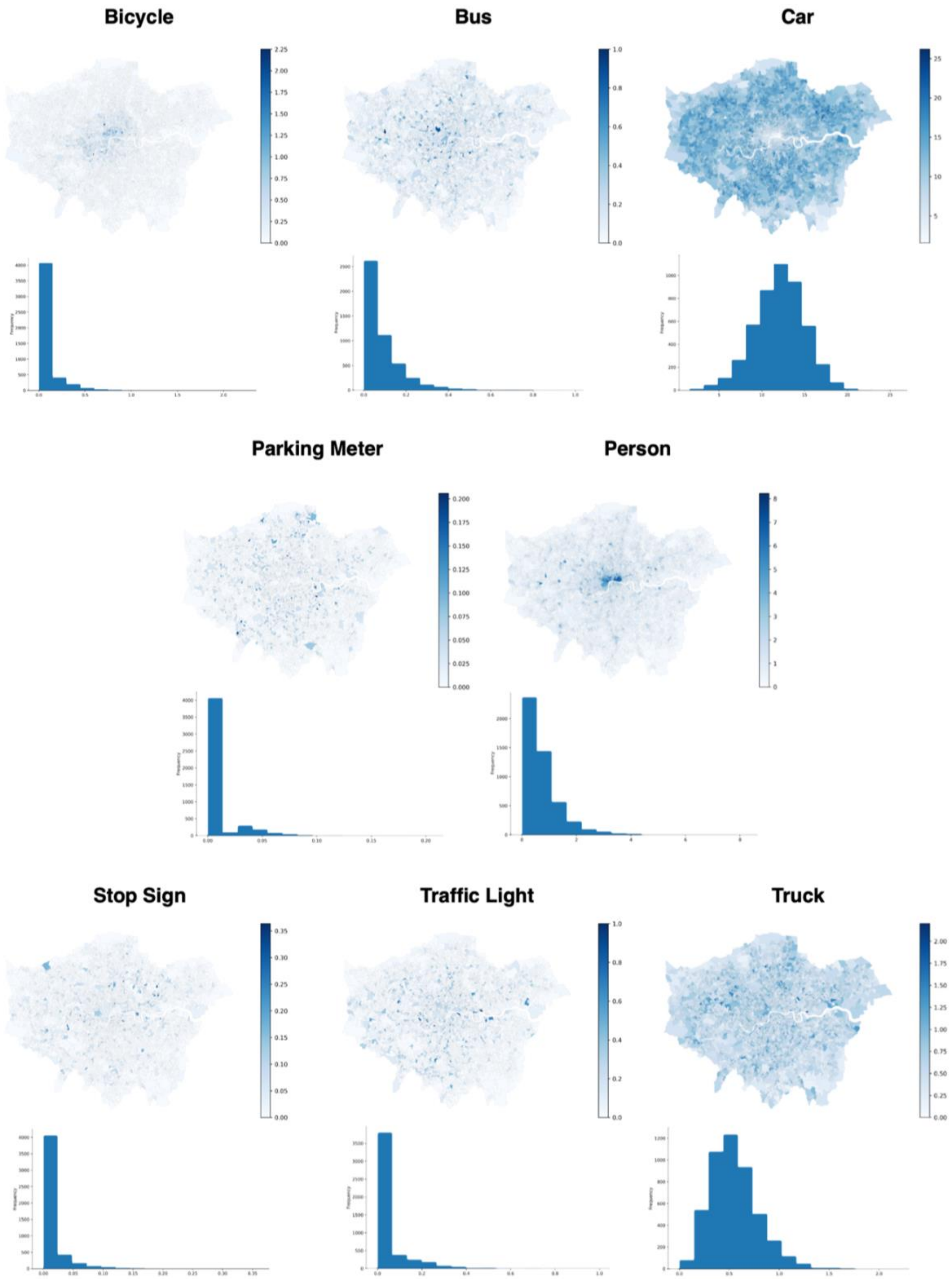


Figure 11 Detected objects' distribution across all London LSOAs. Plus, respective distribution histograms on the bottom of each atlas.

Most bicycles were detected in Central London. Once this is a type of transport for short distances, they were essentially found in this area. Buses are slightly more concentrated in Central London than in the periphery.

Trucks were in higher number away from the city centre. For many roads in the centre, their circulation is forbidden according to the London Lorry Control Scheme. Both bus and truck distributions are long-tailed distributed, increasing the chances of unexpected occurrences, posing cyclists at higher risk. Car distribution is significantly denser outside of Central London. Roads have a higher number of lanes far from the centre, this may imply higher chance of detecting these objects. Fewer parking meters in the centre positively correlates with the number of cars in this area. There is a significant higher number of people detections in the centre. Particularly in the City of London where it is located its historic centre and central business district. Stop signs and traffic lights were essentially detected outside of Central London. Both have a very similar histogram distribution. Although, the number of detected traffic lights is almost five times bigger.

4.2.3. Correlation Matrix

In the context of road safety, the strongest positive correlations were: Person x Bicycle (0.52), Person x Bus (0.48), Bus x Bicycle (0.25) and Bus x Truck (0.20). The strongest negative correlations were: Person x Car (-0.23) and Bicycle x Car (-0.20) (Figure 17).

In terms of positive correlations, Person x Bicycle high value suggests pedestrians and cyclists feel safe occupying the same space. Once bus is a public transport, it is expected high number of people in the surroundings. Consequently, the Pearson coefficient for Person x Bus is relatively high. There is a significant difference between the correlation values of Bicycle x Bus and Bicycle x Truck. This might suggest cyclists feel safer next to buses than trucks. This should not be surprising once bus drivers are more experienced driving close to vulnerable pedestrians. Relative high correlation between buses and trucks, and the fact of being structurally similar suggests one might have been wrongly classified as the other sometimes.

For negative correlations, the statistically significant Pearson coefficients for Person x Car and Bicycle x Car suggest areas with high concentration of cars are dissuasive for cycling and walking.

One factor that cannot be discarded is that bigger objects occluding smaller ones might contribute for the latter being less detected. Consequently, resulting in a negative correlation. Due to the height from the ground GSV images were captured, this is unlikely to be a recurrent phenomenon.

4.2.4. Holistic Cyclist Safety Metric

It was not found a precise metric to estimate cyclist's road safety based on the detected objects in the roads of London. Although, based on the objects' distributions in Figure 11, one positive and another negative combinative measure of safety were formulated.

One of the features that influences cyclist's safety is the number of other cyclists in the surroundings. This happens because drivers become more aware of their presence when in large numbers. Moreover, the vast majority of serious injurious are caused by crashes between vehicles and cyclists. It was found there is a statistically significant negative correlation between the presence of cars and people. This way, the higher the presence of pedestrians, the lower the number of cars. Consequently, less risky for cyclists to get injured. *Bicycle* and *Person* LSOAs were combined into one, after summing the average number of each of these objects per image.

If cars are the main contributors for injury rates, heavy vehicles are particularly relevant when analysing the fatality rates of a certain area. A second LSOA map was created joining the average number per image of the following objects: *Bus*, *Car* and *Truck*.

It is important to highlight we cannot extract a holistic metric of safety from these 2 generated LSOAs (Figure 12). A simple example is a road where cyclists are physically isolated from the traffic is not necessarily unsafe for cycling.

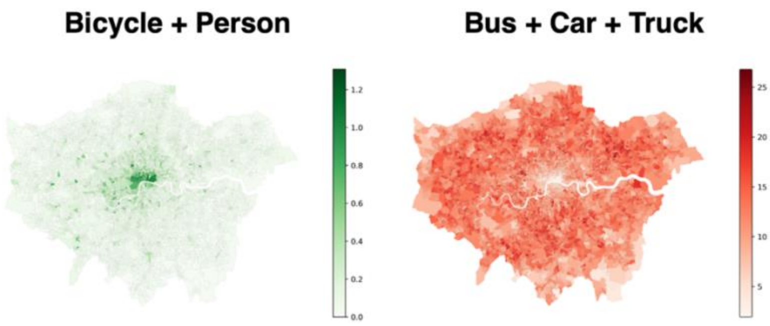


Figure 12 (Left) *Bicycle* and *Person* LSOA distributions were combined into a combinative metric reflecting a positive score for cyclists' safety. (Right) *Bus*, *Car* and *Truck* distributions combined into a final atlas showing the traffic in London. This is inversely correlated with cyclists' safety.

4.2.5. Limitations and Misclassifications

High level of confidence was found for all the detected road safety objects (Figure 11). YOLOv5 can detect a wide range of object sizes, even when partially occluded (Figure 18). Moreover, low contrast between objects and background do not appear to have caused a high number of non-detections. An example of this was when the algorithm detected a car reflected in a window on the roads of London (Figure 9). Nevertheless, this is considered a misclassification – misclassification of sorts.

It was compiled a set of ten random object detected images, so that the reader can verify by themselves the accuracy of YOLOv5 (Figure 18).

4.3. Image Segmentation | PSPNet101

An example of a segmented image is provided in Figure 13. It was identified with the respective labels of the detected structures.

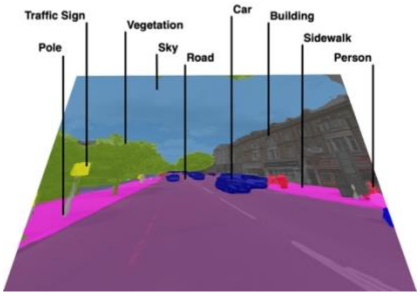


Figure 13 GSV image after segmentation using PSPNet101. Pixel labels were included.

In the next section are detailed the absolute and relative label distributions across the GSV dataset.

4.3.1. Dataset Pixel Labels Distribution

Buildings (27%), sky (22%), road (22%), vegetation (18%) and cars (6%) cover 95% of the total area of all images (Figure 14). In the case of the first four structures, this is explained by their intrinsic dimensions. Once this dataset exclusively contains images obtained from the roads of London, it was expected to find a higher number of cars than any other objects. This pixel frequency number is not only associated with their size, but their frequency in the images. This way, both OD and IS seemed to be in accordance.

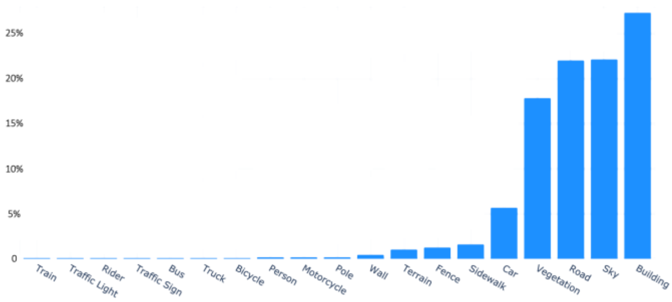


Figure 14 Relative distribution of the labelled pixels after executing PSPNet101 in GSV dataset.

The relative distribution of segmented pixels across the different categories also suggests that not only objects present on the roads can be detected, but also in the surroundings, including in buildings and in the sky. Although for some of the following objects it was not found a clear correlation with road safety, they show the power of using GSV images to detect structures in a wide range of locations. Embedded in the buildings landscape, 2 750 clocks were wrongly classified instead of satellite dishes. In the same area, 37 917 potted plants were identified. In the sky, 234 airplanes were detected. Having detected a significant number of pixels labelled as sidewalk suggests objects regularly present are likely to be captured (107 266 people were found, along with 5 013 benches and 1 168 fire hydrants).

In Table 7 are the absolute counting for all pixel labels distributed by 19 segmented categories.

Table 7 Absolute number of labelled pixels detected across all imagery dataset.

Label	Number Pixels	Label	Number Pixels	Label	Number Pixels	Label	Number Pixels
Building	47.4B	Sidewalk	2.77B	Motorcycle	299M	Traffic Sign	58.1M
Sky	38.4B	Fence	2.18B	Person	232M	Rider	13.9M
Road	38.2B	Terrain	1.79B	Bicycle	95.5M	Traffic Light	12.5M
Vegetation	31.0B	Wall	766M	Truck	91.3M	Train	6.84M
Car	9.83B	Pole	303M	Bus	81.5M		

Again, from an IS point of view, GSV dataset appears to be useful to estimate the area of roads and sidewalks due to the relative high number of pixels detected and consistent shapes. The same applies to streetlights. In spite only 303M pixels were identified, the dimensions of this object suggest that a significant number of those should have been detected.

After individually analysing one segmented image per LSOA for the complete dataset, it appears that both the area and shape of the roads and sidewalks can be accurately retrieved.

As it was identified in the Introduction, these last properties are relevant in a road safety context because they allow to calculate road and sidewalk width. And the presence of streetlights or poles, as they are called in Cityscapes, it is a proxy to assess road visibility.

In Figure 20, these concepts are exemplified along with an illustrative image on the side.

4.3.2. Limitations and Misclassifications

Due to the dimensions of certain structures, PSPNet101 was not able to accurately capture their shape. One example are the poles which are thin and, consequently, their detection is highly influenced by the resolution of the images in the dataset. This is not particularly problematic once the most important about these structures are their detection and not capturing their shape.

In terms of the roads and sidewalks, sometimes occlusion appears to be an issue. Nevertheless, by accounting for the objects that are usually present in any of these areas, considering their overlapped areas simultaneously seems to be an effective workaround. Specifically, this was observed for cars on the roads and people on the sidewalks. This way, it should still be possible to extract information on the shape and size of these structures.

Extracting absolute dimensions of these structures on the roads of London can be a hard task. Criteria can widely vary accordingly to the angle the images were taken. One way to overcome this would be to focus on the relative dimensions across the objects (Figure 19).

4.4. Future Directions

For OD, one step forward in assessing cyclist's road safety would be to find a common road safety indicator allowing us to compare, for example, in terms of accident or injury rate, how objects influence safety differently. Training YOLOv5 in a bigger dataset and with a higher number of object categories than MS Coco. This would involve annotating additional images. Documentation on how to train YOLOv5 is very detailed and easy to follow. Static images cannot capture several features. Using video recordings including pedestrian, cyclist and vehicle movement would allow to capture variables that are highly influenced by the time images are obtained.

In terms of IS, these models can be exclusively used for detecting new objects. One structure that was identified as determinant for cyclist's road safety was the presence of cycle lanes. It was not found any dataset including those. Creating one including this category is of great importance to assess cyclists' road safety.

Although we did not find evidence whether sidewalks correlate with cyclist's safety, in the future, proven there is an association, PSPNet101 pre-trained in Cityscapes will be able to capture it.

Figure 20 summarizes all road safety information (identified in this project) that can be extracted using PSPNet101 in the GSV dataset.

5. Conclusion

The goal of this project was to use OD and IS to extract cyclists' road risk factors from a dataset of GSV images of Greater London. This included the study of image distribution across all LSOAs. Identifying relevant road safety indicators to determine and rank cyclists risk factors. Using YOLOv5 and PSPNet101 to detect objects and segment images. Analysing object distributions and correlations among the different categories. Finally, identifying common misclassifications and limitations of both methods, and propose new ways of advancing road safety assessment.

Approximately 2 million objects were identified, and 200 billion pixels labelled in the 500 000 images available in the dataset. On average, there were 108 images per LSOA. Using YOLOv5, the distribution of the following risk factors was (in)directly identified at an LSOA level: high vehicle speed, tram/train rails, truck circulation, parked cars and pedestrians. Car (84.5% of all objects), person (6.01%), truck (3.92%), bus (0.60%), traffic light (0.4%), stop sign (0.07%), parking meter (0.05%) and train (0.04%) counting were used to identify the previous risk factors. London road traffic was found to be higher outside of Central London. In contrast, the presence of cyclists and pedestrians was higher inside this area. Former was defined as a negative combinative measure of safety and, the second, positive. It was found a statistically significant negative correlations between cars x buses, cars x cyclists and cars x people. Positive correlations between people x bicycles and people x buses. Long-tail distributions on the number of heavy-vehicles (buses and trucks) was observed. Using PSPNet101, building (27%), sky (22%) and road (22%) pixels were the most common. Thus, objects in any of these areas can be equally detected. All results and implementations were made available in the project's repository.

Future directions include increasing the availability and resolution of GSV images. Train YOLOv5 and PSPNet101 with datasets containing a higher number of categories relevant for road safety. Define a safety metric to weight and combine (at a road level) detected objects or segmented structures. Finally, to process street view images or video in real-time would allow to better capture the dynamics of road safety (video).

References

[1] "pedbikeinfo," Pedestrian and Bicycle Information Center, [Online]. Available: http://www.pedbikeinfo.org/. [Acedido em 19 April 2020].

[2] C. Allan, "Cycling UK's Cycling Statistics," November 2019. [Online]. Available: https://www.cyclinguk.org/statistics. [Acedido em 19 August 2020].

[3] "£2 billion package to create new era for cycling and walking," GOV.UK, 9 May 2020. [Online]. Available: https://www.gov.uk/government/news/2-billion-package-to-create-new-era-for-cycling-and-walking. [Acedido em 19 August 2020].

[4] S. M. Balogh, "Perceived safety of cyclists- The role of road attributes," KTH Royal Institute of Technology, Stockholm, Sweden, 2017.

[5] "Coverage of Google Street View," Wikipedia, 14 August 2020. [Online]. Available: https://en.wikipedia.org/wiki/Coverage_of_Google_Street_View. [Acedido em 19 August 2020].

[6] G. Jocher, "YOLOv5," Ultralytics, 23 July 2020. [Online]. Available: https://github.com/ultralytics/yolov5. [Acedido em 23 July 2020].

[7] J. S. X. Q. X. W. J. J. Hengshuang Zhao, "Pyramid Scene Parsing Network," em *Computer Vision and Pattern Recognition*, 2017.

[8] E. R. S. Observatory, "Traffic Safety Basic Facts 2018," European Commission, 2018.

[9] P. A. a. O. CDC (Division of Nutrition, "Physical Activity," 4 February 2019. [Online]. Available: https://www.cdc.gov/physicalactivity/about-physical-activity/pdfs/healthy-strong-america-201902_508.pdf. [Acedido em 19 April 2020].

[10] T. f. London, "Travel in London, Report 11," 2018.

[11] SafetyNet, "Pedestrians & Cyclists," 2009. [Online]. Available: https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/specialist/knowledge/pdf/pedestrians.pdf. [Acedido em 13 May 2020].

[12] "Bicycle Facilities," iRAP, [Online]. Available: http://toolkit.irap.org/default.asp?page=treatment&id=1. [Acedido em 8 April 2020].

[13] "Road Safety Commission," Government of Western Australia, [Online]. Available: https://www.rsc.wa.gov.au/RSC/media/Documents/Resources/Cyclists-INFO-SHEET.pdf. [Acedido em 8 April 2020].

[14] "Bicycle Safety," National Highway Traffic Safety Administration, [Online]. Available: https://www.nhtsa.gov/road-safety/bicycle-safety. [Acedido em 8 April 2020].

[15] K. T. e. al, "Route Infrastructure and the Risk of Injuries to Bicyclists: A Case-Crossover Study," *American Journal of Public Health,* vol. 102, nº 12, pp. 2336-2343, 2012.

[16] L. Chen, C. Chen, R. Srinivasan, C. E. McKnight, R. Ewing e M. Roe, "Evaluating the Safety Effects of Bicycle Lanes in New York City," *American Journal of Public Health,* vol. 102, nº 6, p. 1120–1127, 2012.

[17] "Cycling- preventing injury," Victoria State Government, November 2013. [Online]. Available: https://www.betterhealth.vic.gov.au/health/HealthyLiving/cycling-preventing-injury. [Acedido em 7 May 2020].

[18] "Speed and accident risk," European Commission, 8 May 2020. [Online]. Available: https://ec.europa.eu/transport/road_safety/specialist/knowledge/speed/speed_is_a_central_issue_in_road_safety/speed_and_accident_risk_en. [Acedido em 8 May 2020].

[19] "What forces can be tolerated the human body?," European Commission, Mobility and Transport, 9 May 2020. [Online]. Available: https://ec.europa.eu/transport/road_safety/specialist/knowledge/vehicle/key_issues_for_vehicle_safety_design/what_forces_can_be_tolerated_the_human_body_en. [Acedido em 9 May 2020].

[20] C. a. H. N. Tingvall, "Vision Zero- An ethical approach to safety and mobility," em *6th ITE International Conference Road Safety & Traffic Enforcement: Beyond 2000*, Melbourne, 1999.

[21] E. Pasanen, "Driving speeds and pedestrian safety," 1991.

[22] S. M. G. Ashton, "Benefits from changes in vehicle exterior design," em *Proceedings of the Society of Automotive Engineers*, Detroit, MI, 1983.

[23] V. S. M. R. e. a. Asgarzadeh M, "The role of intersection and street design on severity of bicycle-motor vehicle crashes," *Injury Prevention,* vol. 23, pp. 179-185, 2017.

[24] L. Dablanc, "Goods transport in large European cities: Difficult to organize, difficult to modernize," *Transportation Research Part A: Policy and Practice,* vol. 41, nº 3, p. 280–285, 2007.

[25] M. Jaller, J. Holguín-Veras e S. Hodge, "Parking in the city: Challenges for freight traffic," *Transportation Research Record: Journal of Transportation Research Board,* vol. 2379, p. 46–56, 2013.

[26] A. Conway, N. Tavernier, V. Leal-Tavares, N. Gharamani, L. Chauvet, M. Chiu e X. Bing Yeap, "Freight in a bicycle-friendly city," *Transportation Research Record: Journal of Transportation Research Board,* vol. 2547, p. 91–101, 2016.

[27] J.-K. Kim, S. Kim, G. F. Ulfarsson e L. Porrello, "Bicyclist injury severities in bicycle-motor vehicle accidents," *Accident Analysis & Prevention,* vol. 39, nº 2, p. 238–251, 2007.

[28] J. Manson, S. Cooper, A. West, E. Foster, E. Cole e N. R. M. Tai, "Major trauma and urban cyclists: physiological status and injury profile," *Emergency Medicine Journal,* vol. 30, nº 1, p. 32–37, 2012.

[29] S. Kaplan, K. Vavatsoulas e C. G. Prato, "Aggravating and mitigating factors associated with cyclist injury severity in Denmark," *Journal of Safety Research,* vol. 50, p. 75–82, 2014.

[30] P. Chen e Q. Shen, "Built environment effects on cyclist injury severity in automobile-involved bicycle crashes," *Accident Analysis & Prevention,* vol. 86, p. 239–246, 2016.

[31] P. e. al., "Transportation Research Procedia," vol. 25, p. 999–1007, 2017.

[32] M. McCarthy e K. Gilbert, "Cyclist road deaths in London 1985–1992: Drivers, vehicles, manoeuvres and injuries," *Accident Analysis & Prevention,* vol. 28, nº 2, p. 275–279, 1996.

[33] A. S. Morgan, H. B. Dale, W. E. Lee e P. J. Edwards, "Deaths of cyclists in London: Trends from 1992 to 2006," *BMC Public Health,* vol. 10, nº 669, p. 1–5, 2010.

[34] P. Tuckela, W. Milczarskib e R. Maiselc, "Pedestrian injuries due to collisions with bicycles in New York and California," *Journal of Safety Research,* vol. 51, pp. 7-13, 2014.

[35] D. Dufour, "European Commission - Intelligent Energy Europe," February 2010. [Online]. Available: https://ec.europa.eu/energy/intelligent/projects/sites/iee-projects/files/projects/documents/presto_fact_sheet_cyclists_and_pedestrians_en.pdf. [Acedido em 11 May 2020].

[36] "Object detection," Wikipedia, 24 July 2020. [Online]. Available: https://en.wikipedia.org/wiki/Object_detection. [Acedido em 19 August 2020].

[37] J. Redmon, S. Divvala, R. Girshick e A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," em *Computer Vision and Pattern Recognition*, 2016.

[38] "Image segmentation," Wikipedia, 6 July 2020. [Online]. Available: https://en.wikipedia.org/wiki/Image_segmentation. [Acedido em 19 August 2020].

[39] "Semantic Segmentation on Cityscapes test," Papers With Code, [Online]. Available: https://paperswithcode.com/sota/semantic-segmentation-on-cityscapes. [Acedido em 18 August 2020].

[40] M. M. S. B. L. B. R. G. J. H. P. P. D. R. C. L. Z. P. D. Tsung-Yi Lin, "Microsoft COCO: Common Objects in Context," em *Computer Vision and Pattern Recognition*, 2014.

[41] H. R. N. A. J. U. I. K. J. P.-T. S. K. S. P. M. M. A. K. T. D. V. F. A. Kuznetsova, "The Open Images Dataset V4: Unified image classification, object detection, and visual relationship detection at scale," *Resultados da procura Resultado da Web com ligações entre sites International Journal of Computer Vision,* 2020.

[42] M. O. S. R. T. R. M. E. R. B. U. F. S. R. B. S. M. Cordts, "The Cityscapes Dataset for Semantic Urban Scene Understanding," em *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[43] H. Z. X. P. S. F. A. B. A. T. Bolei Zhou, "Scene Parsing through ADE20K Dataset," em *Computer Vision and Pattern Recognition*, 2017.

[44] "Factsheet: Cyclist safety," European Commission, [Online]. Available: https://ec.europa.eu/transport/sites/transport/files/themes/urban/doc/cyclist_safety_onepager.pdf. [Acedido em 8 April 2020].

[45] "Bike rider safety," Victorian Transport Resources, [Online]. Available: https://www.vicroads.vic.gov.au/safety-and-road-rules/cyclist-safety/bike-rider-safety. [Acedido em 8 April 2020].

[46] "Driving & cycling safety," Transport for London, [Online]. Available: https://tfl.gov.uk/travel-information/safety/road-safety-advice/driving-and-cycling-safety. [Acedido em 8 April 2020].

Appendices

YOLOv5 | Detected Objects

One of the main goals of this project was to show the potential of Street View imagery. Given a large dataset, there are plenty of information that can be extracted.

While analysing all the processed LSOAs atlas, it was found two that illustrate the potential of this technique: *Airplane* and *Potted Plant* categories (Figure 15).

In the case of the first, it was detected a higher density of planes per image in the areas next to the airports of Heathrow and City of London. Moreover, all detected planes are on the right of each of these structures. This phenomenon is explained by the wind direction West-> East, which makes the planes preferably landing from East-> West, and the fact landing takes significantly longer than take-off. Thus, only images taken on the right contain them. Finally, it is also clear the difference on the number of detections next to each of these airports. Due to the increased air traffic of Heathrow, most of them are in its proximities.

Potted plants were also frequently detected. These were mainly present in images closer to the biggest parks of London. This category includes all vegetation inserted in any type of pot. Given vegetation was the second most labelled type of pixels across the GSV dataset after executing PSPNet101, it is not surprising the high levels of captured potted plants (fourth most detected object).

Figure 15 (Left) Density of planes present in images taken next to the closest London airports agrees with expectations. (Right) Identically, the biggest density of potted plants was observed closer to the biggest parks.

YOLOv5 | Limitations & Misclassifications

For the objects we defined as relevant to cyclist's road safety, the number of misclassifications was very small. This was achieved because it was defined a high threshold of 0.5 to count as a detection and in MS Coco training dataset the most common objects are the ones we are interested.

Although, there were objects consistently misclassified. The most common were satellite dishes being detected as clocks. Depending on the angle, arm dishes can easily resemble a clock pointer. It was detected 2750 clocks in the complete GSV imagery dataset (Figure 16). Other less represented objects were also wrongly identified. Sometimes due to their shape, others because of their texture. An example of the former was the detection of boats instead of construction containers or, for the latter, benches instead of fences.

Figure 16 The most common misclassification identified after executing YOLOv5 was the detection of clocks instead of satellite dishes.

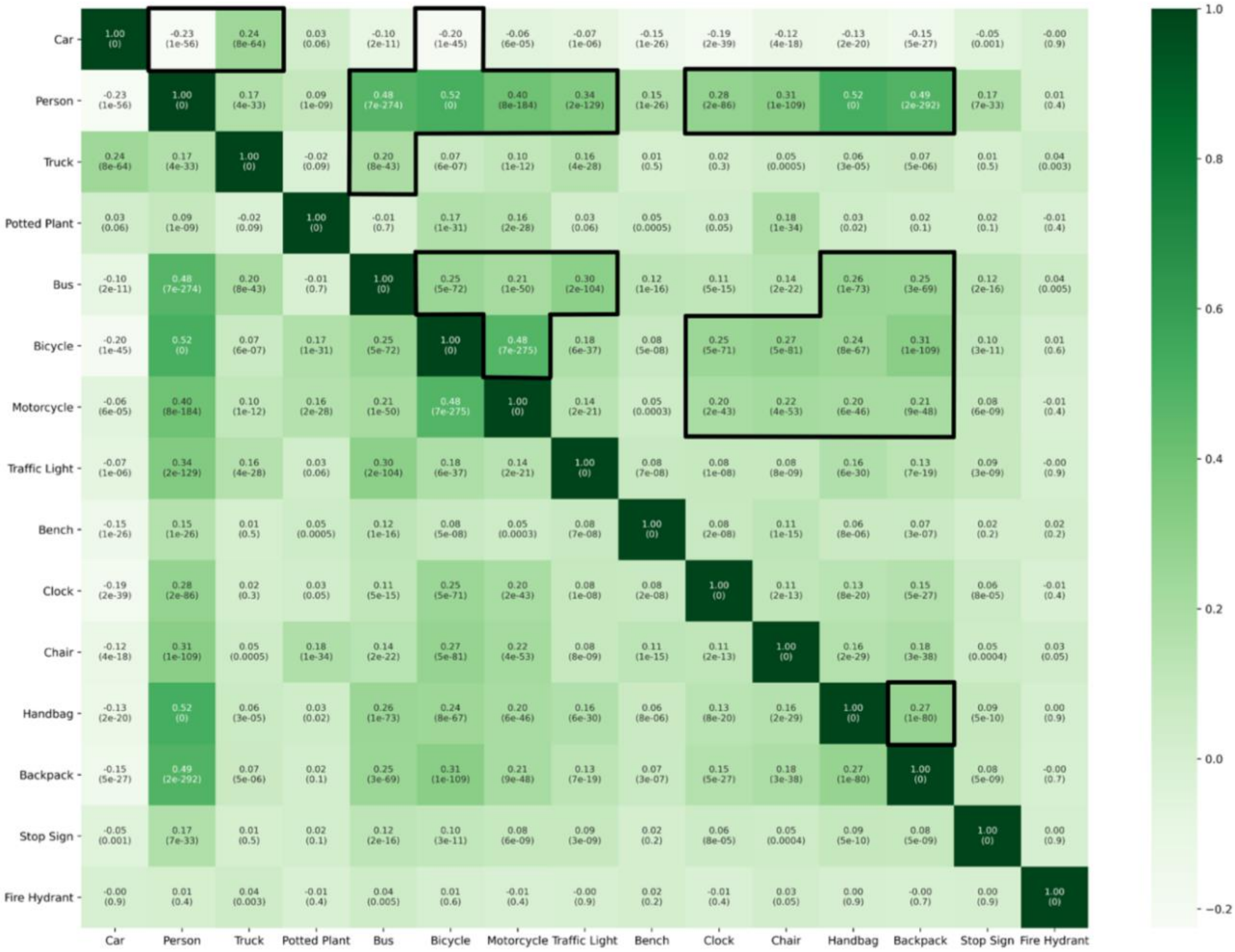


Figure 17 Top 15 detected objects correlation matrix. Each cell contains the Pearson correlation coefficient (top) and the associated p-value (bottom).

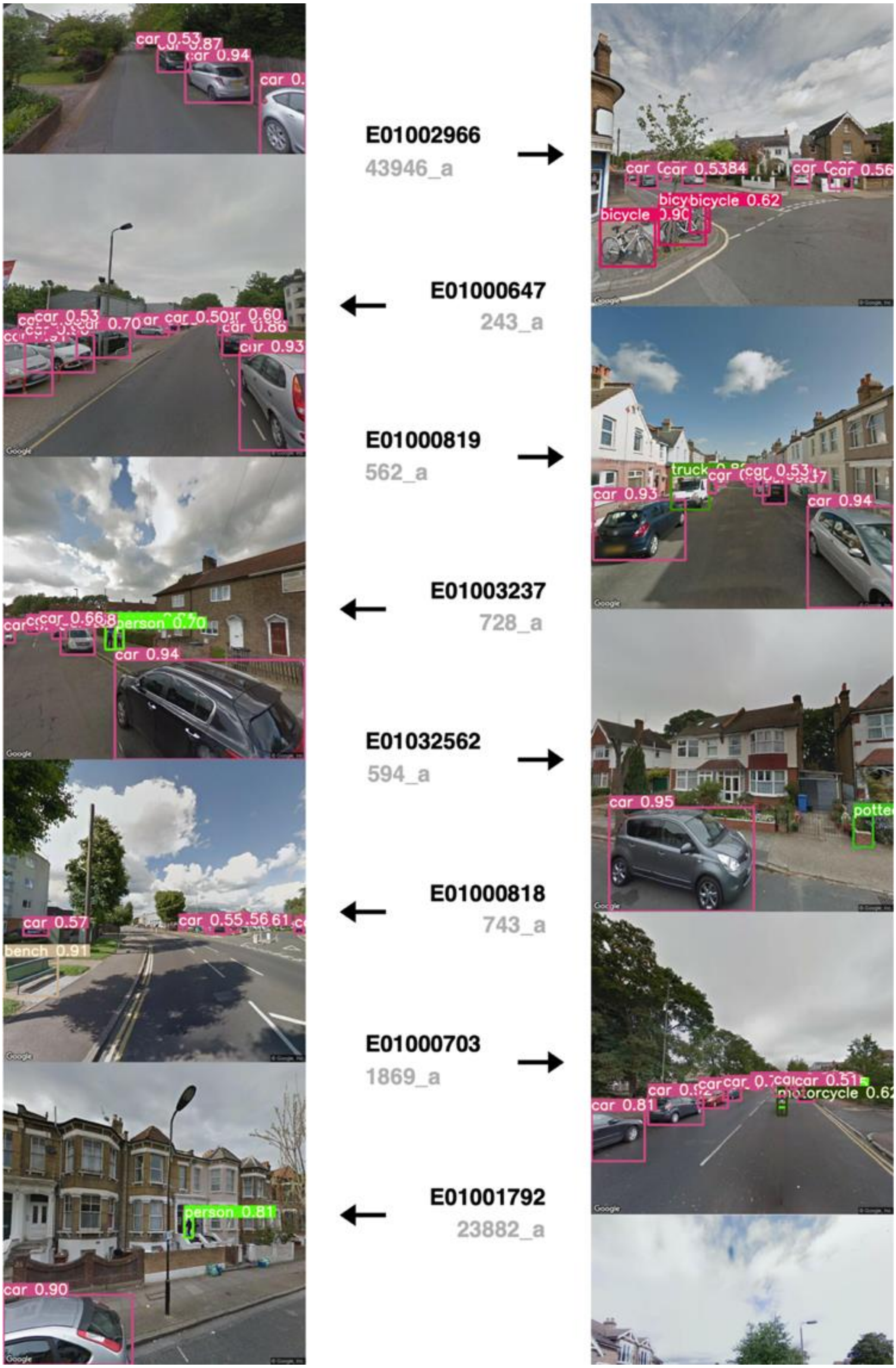


Figure 18 Ten randomly chosen object detected images from different LSOAs show a high accuracy of detection among MS Coco categories.

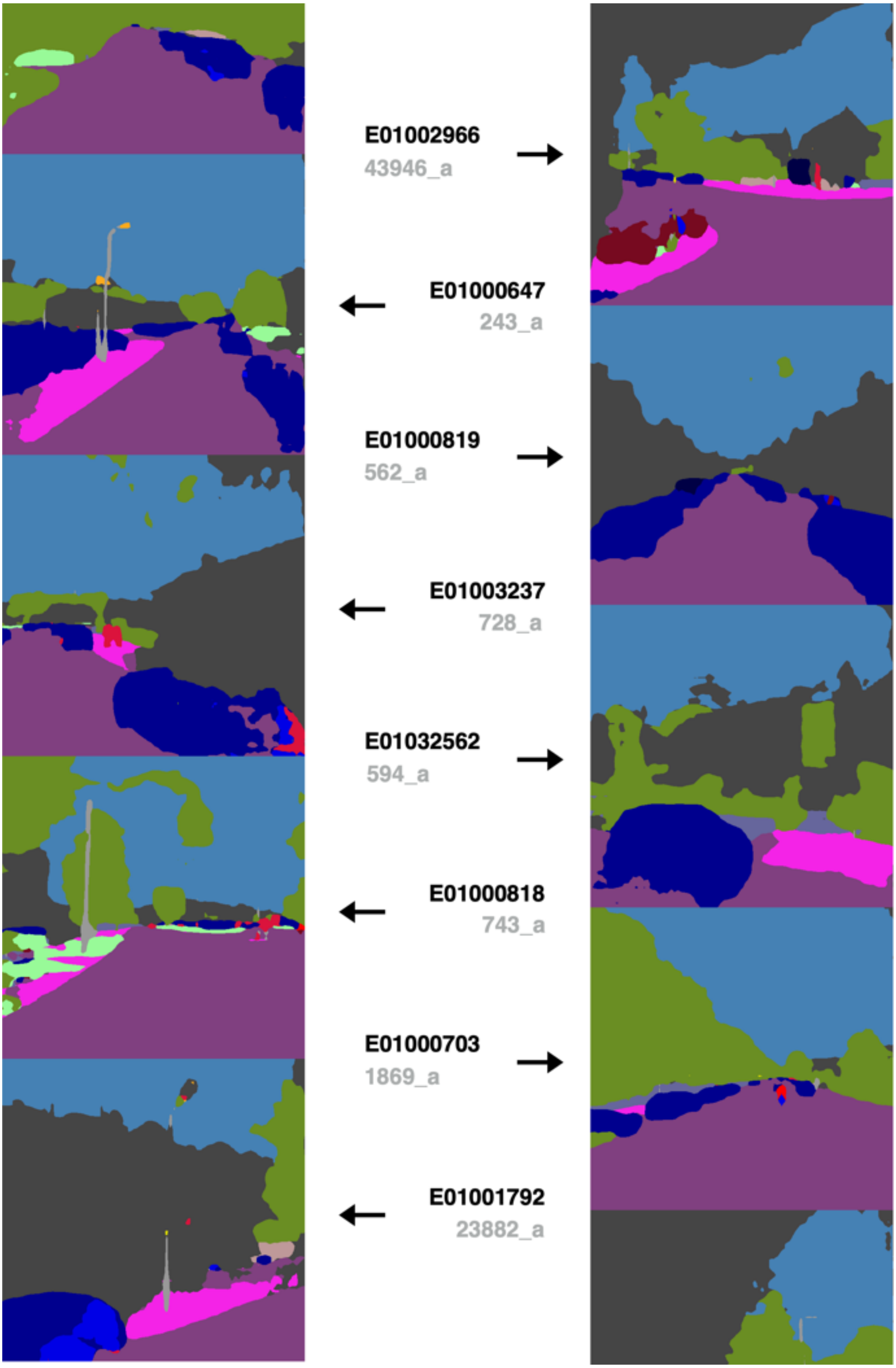


Figure 19 A small sample of randomly segmented images from different LSOAs shows the importance of accounting for structure occlusion while capturing sizes and shapes.

Image Segmentation

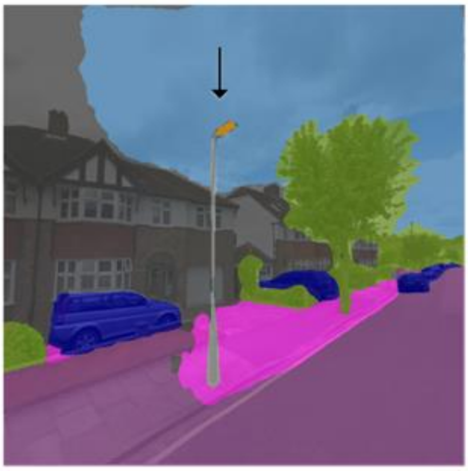
PSPNet101

Road

Road width is determinant for cyclists' safety.
In the presence of a shared cycling lane, this is crucial. Keeping a safe lateral distance between vehicles and cyclists (generally, legally enforced >1.5 m) decreases fatality rate of the latter.

Pole

Light conditions influence drivers and cyclists reaction time. In high speed roads, it tends to be very short. Cyclists are more aware on the presence of pedestrians during the night when there are streetlights.



Sidewalk

Cityscapes incudes in this same category walking paths and physically separate cycling lanes.
The width of cycling lanes is one main factor contributing to cyclists safety. Allowing them to keep their distance from other vehicles and avoiding holes or other obstacles on the floor.

Figure 20 Infographic illustrating the potential of IS to extract road and sidewalk width, and streetlights.

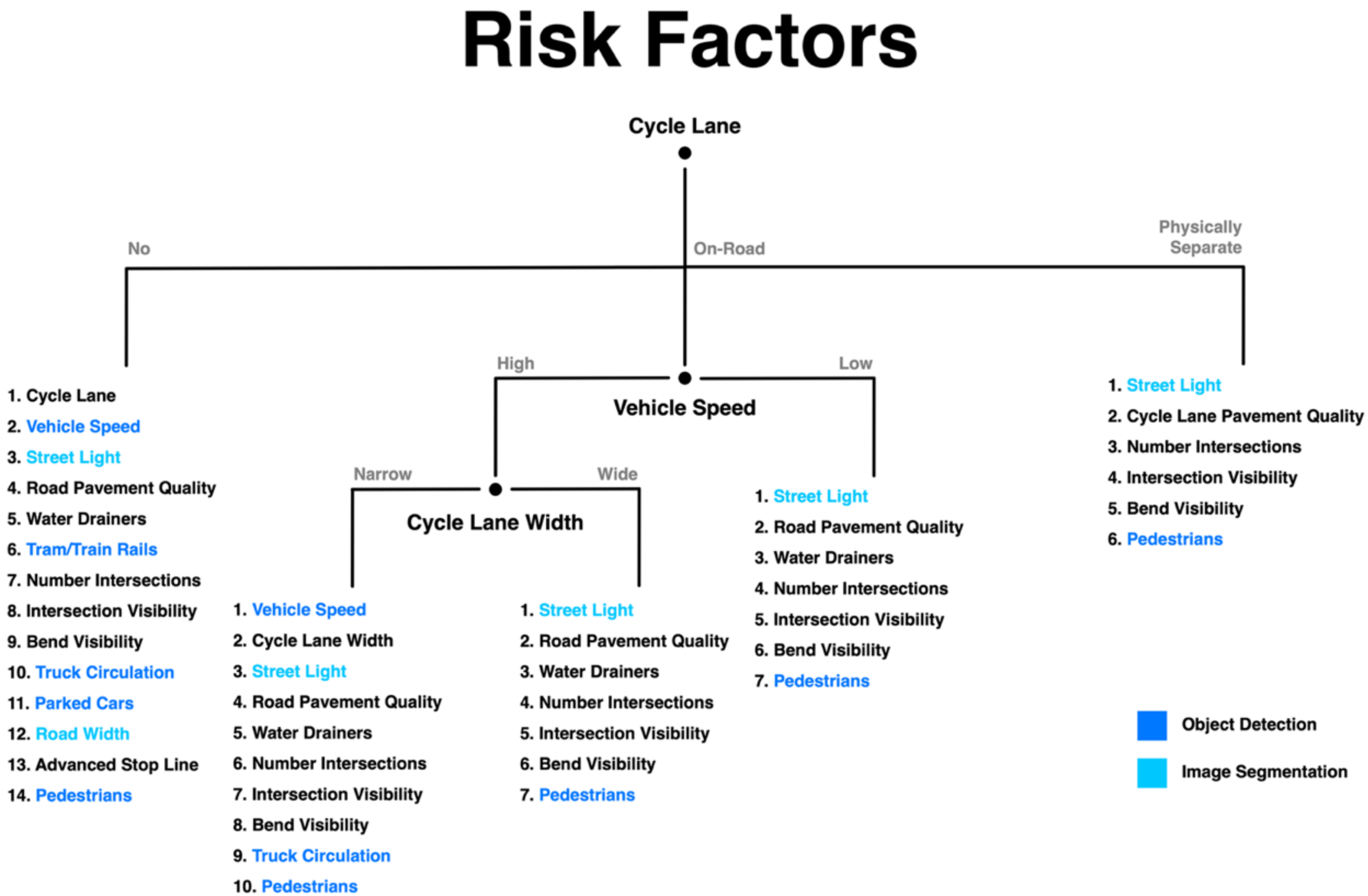


Figure 21 Identified cyclists' risk factors in 5 scenarios on the roads of London.

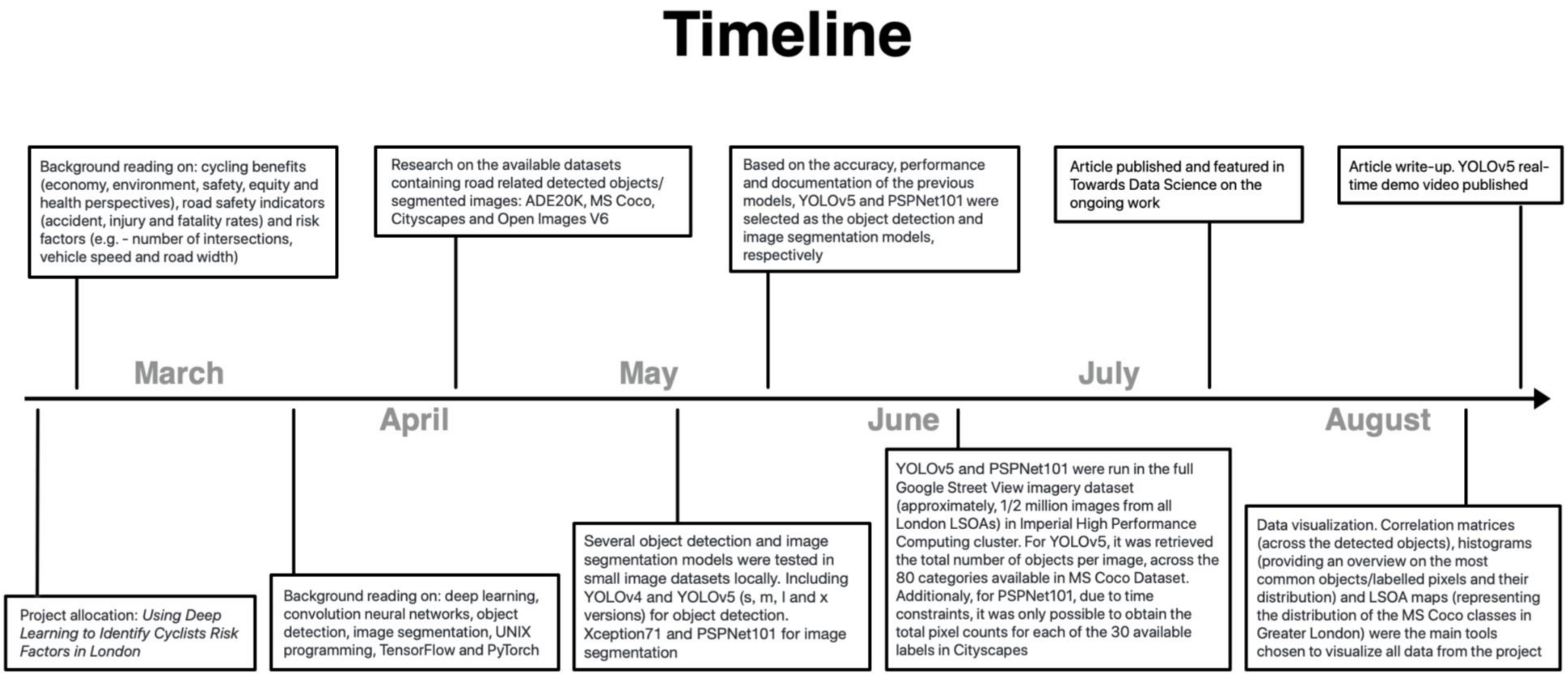


Figure 22 Project's roadmap since it started in March, until the submission month, August.