A Low-Cost Video-Based Solution for City-Wide Bicycle Counting in Starter Cities

Eduardo Peixoto[1], João Moutinho[2], and Rui José[1](✉)

[1] Centro Algoritmi, University of Minho, Braga, Portugal
eduardopeixotop@gmail.com, rui@dsi.uminho.pt
[2] Centro de Computação Gráfica (CCG), Guimarães, Portugal
Joao.Moutinho@ccg.pt

Abstract. Cycling is increasingly popular as a sustainable urban mobility mode. Data can play a key role in the success of cycling promotion initiatives, by helping to understand cycling demand and assess the real impact of investments. However, creating a global view of the cycling activity of a city can be a major challenge, as there are no obvious sources from which to obtain the necessary cycling data. While there are now many bike counters in the market, their hardware and deployment costs severely limit the number of sensors that can be deployed and consequently the spatial coverage of the city. In this work, we explore the viability of a large-scale bicycle counting infrastructure for city-wide cycling analytics, which explores the trade-off between costs and spatial/temporal coverage. The proposed solution uses a set of temporary low-cost video-based counters, which can flexibility be rotated among multiple counting locations. To understand the viability of the approach we developed a prototype counter, where we tested two video processing techniques: OpenCV and Yolo. Results suggest that the overall approach could indeed support a low-cost and universal bike counting functionality, as long as delayed access to data is acceptable. Even though this is not envisioned as a general bike counting solution, it may provide a smart way to approach the complex issue of universal city-wide bike counting.

Keywords: Smart mobility · Bike counters · Cycling data · OpenCV · Yolo

1 Introduction

Cycling and other soft mobility modes are increasingly seen as crucial for sustainable urban mobility [1]. Many cities across the world are developing cycling programmes, which may include the development of dedicated infrastructures, bike sharing services or public awareness activities. A key element for the success of those initiatives is the ability to understand their impact by measuring cycling activity [2]. The lack of systematic and consistent monitoring processes has been identified as one of the obstacles towards more sustainable development of cycling mobility policies [3]. This lack of data about cycling activity affects all elements of the ecosystem, but is particularly problematic for municipalities and other local entities who very much need concrete and actionable data

H. Santos et al. (Eds.): SmartCity 360 2019, LNICST 323, pp. 139–150, 2020.
https://doi.org/10.1007/978-3-030-51005-3_14

to justify investments in cycling infrastructures, to make more informed decisions and to assess the real impact of their policies. Likewise, private companies in the cycling sector also need this type of data to create solid business plans, make informed business decisions and optimize operations management.

Data is therefore critical for the ability of cycling mobility initiatives to create real, profound, incremental and measurable impact. However, creating a global view of the cycling activity of a city can be a major challenge, as there are no obvious sources from which to obtain the necessary cycling data. This is a very generic problem, which affects any city, regardless of its dimension or its level of cycling readiness. In this research, we are mainly concerned with the case of starter cities, where the first initiatives are emerging, but infrastructures and cycling adoption are still very incipient. For these cities, data can help promoting the type of societal and political transformation needed to bring cycling mobility to the centre of mobility policies. Basic data, such as the number of cycling trips at new infrastructures or the level of city-wide cycling activity are particularly crucial for sustaining a strong Return on Investment argument. Without this data, the reality of cycling mobility is largely invisible to urban planners and to society at large.

Recent years have seen the emergence of a wide range of bike counters, which can be deployed on streets to count passing cyclists. Their key advantage is the ability to continuously count the universe of cyclists at those locations, without depending on any action from cyclists themselves. Their key problem is cost, which includes not only the cost of the counting device itself, but also the often very high costs associated with deployment in public space. This severely limits the number of sensors that can be deployed and consequently the spatial coverage of the city. This is particularly severe for cities where cycling investment is still at its early stages.

In this work, we explore the viability of a large-scale bicycle counting infrastructure for city-wide cycling analytics. Rather than improving the technical performance of any particular technology, we are mainly concerned with the development of a bike counting service in the most economically efficient way and using whatever combination of technologies may be suitable for this purpose. This should allow any starter medium-size city to deploy a few hundreds of counting points, which would constitute a paradigm change for city-wide cycle sensing. Our early solution is centred on video analysis. We have developed a low-cost video-based counting device based on the Raspberry Pi platform. In this study, we analyse the ability of this low-cost platform to serve as the backbone for a city-wide bicycle counting service. The approach is based on a rotating scheme, where sensors are regularly moved to pre-defined locations and possibly combined with partial data sources to provide a global view of the city cycling activity. To understand the viability of the approach we compare the reliability and deployment flexibility of two video processing techniques. Despite using only well-known and widely available video processing techniques, this approach seems to offer enough precision to be seen as a valuable solution, in which some loss in precision may be counterbalanced by the wide spatial coverage of the counting process.

2 Related Work

A bicycle counting system can be based on manual or automated techniques, or even in a combination of both. Manual techniques involve having people manually counting passing bicycles, either physically at counting gates or by observation of videos recorded at those locations. They can offer a valuable approach for more occasional measurements or when a more thorough characterisation of cyclists is needed [4]. Still, a manual counting is a monotonous and, especially, a time-consuming task, which can only provide occasional snapshots of the cycling reality. It does not scale when a wide spatial/temporal coverage is needed.

An automated process involves placing a sensing infrastructure that is able to autonomously perform the count. A bicycle counter is normally composed by a sensor, which collects the data used for detecting the passage of bicycles, and a central apparatus, which supports any data processing needs, as well as sharing or storing the collected information. Recent years have seen the emergence of a wide range of automated bike counters, and there are now several market solutions. They explore a very broad range of bicycle detection techniques, with very different properties and performances, such as Pneumatic tubes, Inductive loop detectors, Piezoelectric Strips, Pressure or acoustic pads, Active infrared, Pyroelectric, Laser scanning, Radio Wave, Video image processing, Magnetometers or Bicycle Barometers [4]. Some of these technologies were initially developed for counting motor vehicles, but they have since been repurposed for detecting pedestrians and cyclists.

The current state of the art in commercial solutions may be exemplified by CITIX-3D by Eco-Counter [5]. This is a wide-range counter, with the capability to automatically count and classify, not just cyclists, but also pedestrians and vehicles. The technology developers claim that it offers a greater precision than traditional video analysis, requires zero calibration and can be reliably operated day or night, rain or shine. However, the very high cost of the solution and the sophistication of the data produced, make it more suitable to generate broader usage profiles of public spaces. For example, they can support the detailed study of movement patterns around specific crossings, helping planners to understand how space is used and consequently, identify the need for specific interventions. Bike counters are also often used in combination with public displays, showing daily, monthly or annual bike counts in real-time. These displays serve mainly as a public celebration of all those who already cycle and as a medium to improve the public visibility to the city's efforts to promote bicycle traffic.

While some of these techniques can be highly sophisticated and accurate, they do not seem to offer a compelling solution for a city-wide counting system. The first reason is that they typically assume that their key performance indicator is precision. This naturally favours more sophisticated and consequently more costly technologies. While precision is an obviously central criterium, the added value brought by a more complex solution may not always counterbalance the added costs, especially when considering the existence of a large number of counters. The second reason is that the viability of their deployment can be highly dependent on environmental elements, such as the possibility to physically install sensors or other hardware, the profiles of the cycling paths or even the level of cycling/pedestrian traffic. Consequently, when we consider the huge diversity of urban settings that compose any city, it becomes very difficult to identify a technique that could be the best solution for the whole city. Moreover, many of these deployment approaches may also represent significant additional costs. This severely limits the number of sensors that can be deployed and consequently the spatial coverage of the city. Large communities can sometimes afford to install two or three permanent counters at key locations, and very large communities, may reach upwards of ten permanent counters throughout the city [6]. With just a few counting devices, mobility analysis is not complete and the monitoring data is not rich enough. As a consequence, these solutions fail to provide a large scale and cost-effective solution for the problem of measuring urban cycle traffic. They are mainly deployed in very small numbers and to cover high profile locations, such as new flagship infrastructures.

3 A Video-Based Solution for Scalable Urban Bike Counting

Our exploration of this design space is driven by the need to optimize three key variables of the solution, cost, spatial coverage and temporal coverage, while also guaranteeing viable precision levels. Assuming continuous spatial and temporal coverage would always imply very high costs because a large number of devices would have to be permanently used at a very large number of locations. However, reducing costs by cutting the number of counters will have a negative impact on spatial and temporal coverage. Therefore, the fundamental design issue for this system is the optimisation of the inherent trade-off between coverage and cost.

The first and most obvious approach to address this optimisation is to simply reduce the cost of the counting devices. A solution based on low-cost counters would be more prone to be economically viable for large scale deployment across the city. For that, the cost of the counting devices should be an order of magnitude lower than current solutions. A key step to reduce costs is to focus only on concrete data needs, which in our case is information about bike passages at multiple reference locations around the city, so that we can build a global account of bike trips in that city. A clear focus on these concrete requirements should allow to reduce or eliminate any development, deployment and maintenance costs not directly related with those needs. For example, assuming that no live data is needed, the sensing device could operate without any network connectivity. It could simply store counting data, which could then be retrieved directly from the device itself, whenever there was a maintenance operation or when the device was moved to another location. For most urban planning processes, this delay in data gathering should not be a problem.

The second way to optimize the cost/coverage equation is by relaxing the requirements for continuous temporal and spatial coverage. Since we want to create a city-wide system, our approach is designed to favour spatial coverage, while relaxing time coverage. The goal is to reduce the number of bike counters that are needed, and consequently the cost of the whole infrastructure. However, since we still need to obtain data at a potentially large and representative set of collection points, we consider that the system would be based on a network of temporary bike counters to be regularly rotated across different locations. In this case, a lower number of sensors should be able to cover many counting points without compromising the essence of the data produced over time by the system.

A key requirement emerging from this approach is that a temporary counter should be very simple to deploy at multiple locations and it should not require any external energy source or cable connections. A very flexible and low-cost deployment process would be essential to reduce the cost of the overall solution, as the costs involved in the public deployment of technology can be substantially higher than the costs of the devices themselves. Flexibility is also crucial to support the deployment of bike counters into the very diverse urban spaces where they may need to be deployed.

In regard to time coverage, basic information about the number of cyclists passing by a particular counter should easily accommodate only occasional measurements. A partial temporal coverage should not be a limitation as, for most cases, bike passage numbers are essentially about trends and medium or long-term evolution, and not so much about live data. The average values are not likely to face dramatic changes, and the total number of passages over a month or a year should be easily estimated based only on samples obtained from regular, but not continuous measurements. With the proper corrections for seasonal or weather factors, it should be reasonably simple to make adequate estimations.

This process could also be improved even further by combining these video-based counters with other, more accessible, but less accurate data sources. Many of them may provide simple and reliable data about cycling activity, but they fail to provide universal counts, as they are can only count a subset of cyclists. For example, only a few cyclists will have detectable Bluetooth devices and only a few will use any particular cycling application. The complementarity between technologies may help to mitigate some of their limitations and create a view of cycling mobility that is bigger than the sum of the parts.

From this perspective, the video-based counter described in this work is mainly seen as providing the baseline data about the universe of cyclists in a particular location. When deployed alongside other sensors, they can be used to determine the percentage of passing cyclists which can be detected through other basic sensors. For example, one can estimate the percentage of cyclists using a specific cycling application from which public data can be obtained. This can then be compared against the data generated from those other sources to provide an estimate of their representativeness in regard to the universe of cyclists. With these data, it would then be possible to support dynamic and more accurate estimations of cycling counts based only on data from those partial sources. This would also reduce the need for a very complete temporal coverage of the bike counters. The process can be regularly recalibrated through the scheduled redeployment of video sensors at those locations.

3.1 A Portable Video-Based Bike Counter

We decided to focus on video analysis as the core sensing technology for this bike counting system. This decision was based on its low cost, its high reliability and the wide availability of cameras and video processing tools [7], but especially on the flexibility with which it can be deployed across many different types of city locations. Using video as a core technology does not exclude other types of sensing technology, it just assumes that video will play a central role as the source of universal counting data for direct counts and for adjusting estimation parameters.

The counting device is based on a Raspberry Pi v2 with a 900 MHz quad-core ARM Cortex-A7 CPU and 1 GB RAM. The device also includes a Raspberry Pi Camera Module 8MP V2 with auto-focus and a 10,4 mAh Power Bank. These are all simple hardware components, which are widely available on the market and should cost a total of about 100€, making it significantly less expensive than common bike counting solutions. Despite its simplicity and low performance, we have managed to use this device to run the software needed to capture and process images. Its small size and energy autonomy are also important to make it a very suitable solution for simple deployments and for regular rotation between multiple counting locations. Ideally, battery capacity should be aligned with transfer cycles, so that their charging or replacement could be made as part of that process.

The essence of our bike counting device is the ability to detect the presence of specific objects of interest, in this case, bicycles in the images captured by the camera. We explored the use of two video processing techniques which may be seen as representative of two major approaches for video processing: Background Subtraction with OpenCV and Deep Learning with YOLO.

The reason why we tested two different techniques was not to make any generic comparison between them, but to understand their implications in regard to our specific requirements of flexible usage across many different deployment locations. The software was developed using python and the respective libraries for those methods.

The OpenCV approach was based on background subtraction [8]. This method is suitable for videos, and compares each frame with the previous frames, thus allowing to distinguish background from moving objects of interest, as shown in Fig. 1.

Fig. 1. Application of background subtraction.

To extract the information of interest contained in the videos, we used multiple OpenCV library functions, including GaussianBlur, createBackgroundSubtractorMOG2, morphologyEx and connectedComponentsWithStats. This method creates bounding boxes around the white components which are attached in the mask. Bounding boxes whose area is below or above a certain threshold are peremptorily rejected.

Once the algorithm detects the object of interest, it proceeds to the bounding box design, which delimits the object in question for the study (see Fig. 2). We used a basic classification scheme where squares were classified as pedestrians and rectangles as classified as bicycles.

Fig. 2. Determining bounding box around the object of interest.

YOLO (You Only Look Once) [9] uses neural networks to detect objects in images (not taking video into account). The popularity of YOLO comes from the fact that only a neural network can describe the bounding box and classes of the objects, all at once. We used an implementation of this neural network, called DarkFlow, which implements YOLOv2 using the TensorFlow package in Python. This neural network has already been trained with bicycle and pedestrian images. For the purpose of this work, we rejected objects whose probability was less than 0.8.

4 Methodology

To evaluate and fully explore the range of design and deployment alternatives, we analysed the different possibilities associated with the application of the two presented methods (Background Subtraction and Deep Learning). For each of them, we tried to identify the operational conditions that could optimise their precision and the extent to which this could affect their viability as flexible and easy to deploy cycling sensor. The key independent variable in our study was the camera position.

The goal is to represent the diversity of scenarios that may arise when using video cameras in different urban settings. This was from the beginning an important requirement, as we wanted to guarantee that the cameras could easily be deployed across many locations in a variety of situations, which could require very diverse camera mounts. With that in mind, we defined 3 different positions for the counter device (*c.f.* Fig. 3), which we believe may be representative of most common settings:

- View 1: A vertical line pointing directly from above to the cyclist (Vertical);
- View 2: A lateral view pointing to the side of the cyclist (Lateral);
- View 3: An obliquus view with ~45° angle on passing cyclists (Obliquus).

The perspective obtained with each camera view is shown in the 3 video frames which compose Fig. 4.

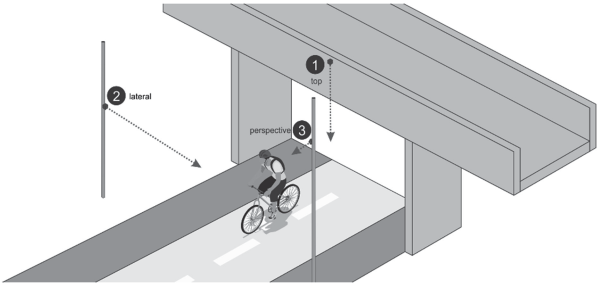

Fig. 3. Evaluation deployment with video being captured from three references angles

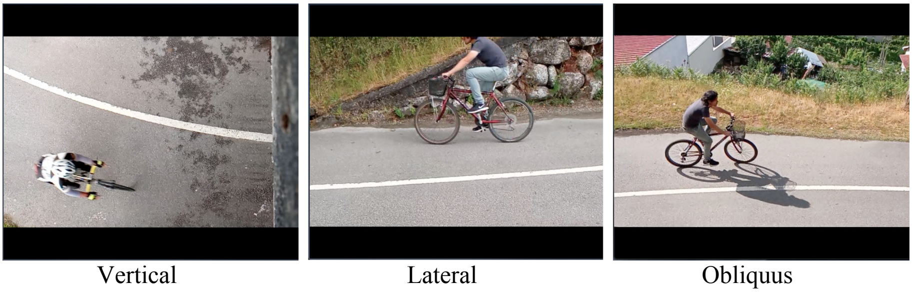

Fig. 4. Three video capture views used for evaluation

We collected videos at each of the three angle views and the same videos were used as input to our two video processing methods. Even though this was a cycle path, there was also pedestrian traffic and there were cyclists passing by in small groups and at various speeds. The evaluation addresses two complementary perspectives. The first was the general efficacy of the counting process, measured in terms of false positives, false negatives and the ratio between manual and automated counts, for each of the technique/angle combinations. The second was the deployment implications, assessed by the impact of different operational conditions on the efficacy of the techniques. The efficacy of the counting process is largely dependent on the deployment conditions and they both affect the final cost of the process.

5 Results

To support the evaluation process, we started by manually annotating the video streams with the events corresponding to the passage of cyclists. Each annotation corresponded to a passage event with the times of the first and last frame where the bicycle is detected. This was used as the ground truth for assessing the results of each experimental condition.

We then applied the OpenCV and Yolo techniques to automatically count the bicycles in each of the three video streams (vertical, lateral and obliquus). This was made by automatically generating a similar type of annotation where a bicycle passage event was associated with two timestamps: the time when it was first detected and the time where detection ends.

These values were compared against the begin and end times produced by manual annotations. We consider that an object had been correctly identified, whenever the absolute difference was less than two seconds. If a passage annotation is not identified by the algorithms, this was counted as a false positive. If the algorithms produced any count that did not had a corresponding passage in the video, this was counted as a false positive. Table 1 shows the results generated for the six experimental conditions.

Table 1. Results of bike counting for each experimental condition

Method	View	Man count	Auto count	False negative	False positive	Diff	%Diff
OpenCV	Vertical	177	209	3	35	32	18,1%
OpenCV	Lateral	77	82	0	5	5	6,5%
OpenCV	Obliquus	73	61	21	9	−12	−16,4%
Yolo	Vertical	177	189	6	18	12	6,8%
Yolo	Lateral	77	79	0	2	2	2,6%
Yolo	Obliquus	73	71	2	0	−2	−2,7%

A general analysis of these results shows that in general YOLO (28 false counts) seems to perform better than OpenCV (73 false counts). Also, for both techniques, the above view seems to be worst performing approach, as can more easily be perceived from Fig. 5 and Fig. 6.

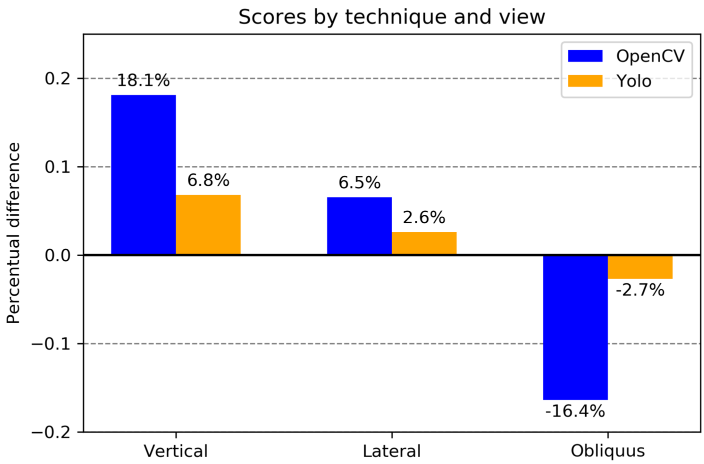

Fig. 5. Comparison between techniques for each of the camera positions

The key observation is that Yolo seems to clearly outperform OpenCV in the ability to correctly detect bicycle passages. In particular, its Deep Learning approach seems more suitable for dealing with abrupt changes in the scene caused by light variations or camera vibration.

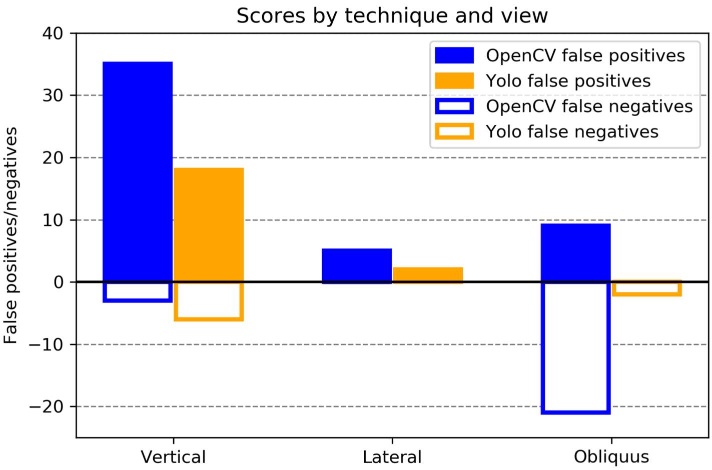

Fig. 6. Comparing false results for each of the camera positions

However, a key part of this analysis is also to identify some of the causes behind the situations where the system performed worst, mainly to understand if they were merely circumstantial and potentially easy to solve by using some different set of configurations. OpenCV scored particularly bad in Vertical or Obliquus views. This seems to have been mainly caused by the high sensitivity of this technique to variations caused by shadows and their change throughout the day. The camera focus may also have interfered, and seems the likely cause of many of the false positives that were detected with OpenCV. The key implication of these results is that OpenCV could only be considered a viable approach for side views, which would break our requirement of achieving maximum flexibility in regard to camera deployment options. However, OpenCV has a much better computational performance, which could be essential whenever there is the need for near real time data.

With YOLO, there might also be some space for improvement, particularly in regard to the less positive score in the vertical view. The YOLO package used for this work had been previously trained for bicycle recognition using mainly videos with lateral and obliquus views. This may justify the better performance of the algorithm in these cases. However, this also opens a potentially simple path for improving performance from the above view by adding new training cases using videos capture from vertical points of view. This vertical view is considered to be the most effective to properly segment bicycles when multiple cyclists are riding together in close proximity and generating occlusion situations.

The key limitation with YOLO, however, was computational performance. The Raspberry Pi device used in the experiment had very low processing power. While enough for the OpenCV approach, these devices were not able to cope with the requirements of YOLO processing. More specifically, they were not able to perform real time processing in which image capture and image processing would both be performed concurrently. A first solution would be to off-load the images to more powerful servers, which could be much more efficient in completing the task. However, this would bring connectivity requirements and increase energy demand. Also, from a privacy perspective, there is a strong argument to discard approaches where the video data leaves the counting device.

An alternative is to separate video capturing and recording from video analysis. The system could be capturing data for a certain period of time and then stop data collection to process the images. From our experience with the processing algorithms this seems to be viable and something which could potentially be aligned with day/night cycles.

6 Conclusions and Future Developments

In this study, we have analysed the conception of a low-cost, video-based and movable bike counting device for supporting large-scale counting. The approach explores the optimization of the inherent trade-off between cost and spatial/temporal coverage.

Results suggest that the overall approach could support a form of low-cost and universal bike counting functionality, providing a path to approach the complex issue of how to enable a city to understand the reality of its emerging cycling mobility. The flexibility and simple deployment that characterises this system seems well suited to the concept of intermittent counting, where sensors would be deployed for relatively short periods, i.e. 1–4 weeks, and then moved to another location, possibly returning to the original location every few months. This would provide the core data for city-wide cycling analytics, which could then be combined with other data sources to estimate a more dynamic and accurate perspective of the cycling reality.

Future work will explore the use of Bluetooth radio identification to create origin/destination matrixes. As many cyclists use Bluetooth equipment (Smartphones, wearables or speedometers), to identify their radio device while passing through a counter, this may provide partial, but very relevant, information to support the decision process in cycling mobility in municipalities. Our preliminary tests revealed that this technology is a possibility, especially considering the ever-growing number of devices that use Bluetooth. The occasional deployment of the video-based approach described in this work, would provide the data needed to estimate bicycle counts based only the percentage of users that carry detectable Bluetooth devices.

We also plan to investigate the system properties of this counting system. This may involve determining the minimum number of bike counters needed to estimate global cycling activity within particular levels of confidence. It may also involve the optimization of measurement cycles and their duration.

Acknowledgements. This work has been supported by national funds through FCT, Fundação para a Ciência e Tecnologia, within the Project Scope: UID/CEC/00319/2019, and also by the European Structural and Investment Funds in the FEDER component, through the Operational Competitiveness and Internationalization Programme (COMPETE 2020) [Project nº 039334; Funding Reference: POCI-01-0247-FEDER-039334].

References

1. Bulc, V.: Cycling: green and efficient transport for the future. In: European Commission (2016)
2. Lee, J., Leem, Y.T., Lee, S.H.: Categorizing U-Bike service and assessing its adoptability under it. In: 12th World Conference on Transport Research, pp. 1–10 (2010)
3. Ricci, M.: Bike sharing: a review of evidence on impacts and processes of implementation and operation. Res. Transp. Bus. Manag. **15**, 28–38 (2015)
4. Johnstone, D., Nordback, K., Lowry, M.: Collecting Network-Wide Bicycle and Pedestrian Data : A Guidebook for When and Where to Count (2017)
5. Eco-Counter: CITIX 3D (2019). https://www.eco-compteur.com/en/produits/citix-range/citix-3d-2/. Accessed 7 July 2019
6. Lowry, M.B.: Spatial Analysis of Bicycle Count Data: Final Project Report (2017)
7. Parekh, H., Thakore, D., Jaliya, U.: A survey on object detection and tracking methods. Int. J. Innov. Res. Comput. Commun. Eng. **2**, 2970–2979 (2014)
8. KaewTraKulPong, P., Bowden, R.: An improved adaptive background mixture model for real-time tracking with shadow detection. In: Remagnino, P., Jones, G.A., Paragios, N., Regazzoni, C.S. (eds.) Video-Based Surveillance Systems. Springer, Boston (2011)
9. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, December 2016, pp. 779–788 (2016)