Highway traffic monitoring on medium resolution satellite images

Sébastien Drouyer, Carlo de Franchis

Highway traffic monitoring on medium resolution satellite images

Sebastien Drouyer
CMLA
ENS Paris-Saclay
Cachan, France
sebastien@drouyer.com

Carlo de Franchis
CMLA
ENS Paris-Saclay
Cachan, France
carlo.de-franchis@ens-cachan.fr

***Abstract*—These last years, earth observation imagery has significantly improved. Public satellites such as WorldView-3 can now produce images with a Ground Sample Distance of 31cm, reaching an equivalent resolution than aerial images. Perhaps more importantly, the revisit frequency has also been greatly enhanced: providers such as Planet can now acquire images of an area on a daily basis. These major improvements are fueled by an increasing demand for frequent objects detection. An application generating a particular interest is vehicle detection. Indeed, vehicle detection can give to public and private actors valuable data such as traffic monitoring and parking occupancy rate estimations. Several datasets, such as DOTA or VehSat, already exist, allowing researchers to train machine learning algorithms to detect vehicles. However, these datasets focus on relatively high definition and expensive aerial and satellite images. In this paper, we will present a method for detecting vehicles on medium resolution satellite images, with a GSD comprised between 1 and 5 meters. This approach can notably be used on Planet images, allowing to monitor traffic of an area on a daily basis.**

***Index Terms*—vehicle detection, satellite image, Planet**

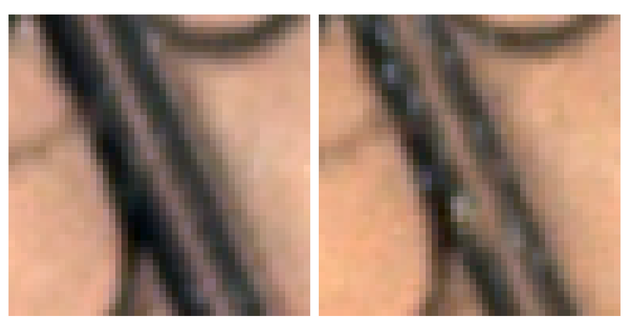
I. Introduction

Automatic traffic monitoring can allow public and private institutions to take better informed decisions. In addition to being an indicator of economic activity [1], it is also correlated to oil consumption and pollution. In order to monitor traffic, several solutions have been proposed. Vehicule detection can notably be done using video cameras [2] [3], aerial images [4] [5], or satellite images [6]. However, using cameras or planes to acquire images is an expensive solution if a full scale daily traffic monitoring is needed. Using satellite images can be cheaper, but the resolution of satellites that are able to revisit an area daily is generally much lower than 1 meter.

In such low resolutions, detecting vehicles is challenging, as a vehicle is represented by a few pixels. However, on highways, they are visible and appear as dots on 3m resolution Planet images. This is especially true in Saudi Arabia, where the majority of vehicles are white and roads are dark. See figures 1 and 2.

In this paper, we present a method able that automatically monitors traffic density on a series of Planet images. It takes as input a road mask and a series of Planet images and returns as output an estimation of the number of vehicles for each image. As the conditions of acquisition can vary, the method

Fig. 1: Planet images of the studied highway in Saudi Arabia on 09/19/2017.



(a) 09/18/2017 (b) 09/19/2017

Fig. 2: Crops of the studied highways in Saudi Arabia at different dates. Contrast has been enhanced using normalization.

has been developped so that it is relatively robust to changes in illuminations, contrast, and blurriness.

We will start by describing the methods developped, and then we validate the method using simulated and real images.

II. Methodology

As described in the introduction, the method takes as input a road mask and a series of Planet images of the same area on different dates. The road mask is an binary image indicating where roads are located. It can either be conceived manually or automatically using semantic segmentation or tools such as OpenStreetMap. The method returns as outputs an estimation of the number of vehicles for each image.

A major assumption of our approach is that vehicles are lighter than the road at least in one RGB channel. The detector will therefore detect white, silver, blue, red, brown or green vehicles, but won't detect black or grey vehicles. According to PPG Industries and DuPont manufacturers [7], black and grey vehicles account for 35% of the total number of vehicles around the world. On most scenes, our detector will only detect around 65% of all vehicles present. This isn't ideal, but provided a sufficiently large sample, this detection rate is enough for estimating traffic density.

The method developped consists of two steps: a preprocessing step and a detection step.

A. Image serie preprocessing

First, we detect if clouds are present in each image (a process similar to [8] can be used): if the highway is partly occluded by clouds, then the traffic estimation is likely to be unreliable so the image is removed from the series and won't be processed.

Secondly, as raw images are generally not perfectly aligned, they are aligned using a phase correlation method as described in [9].

B. Vehicle detection

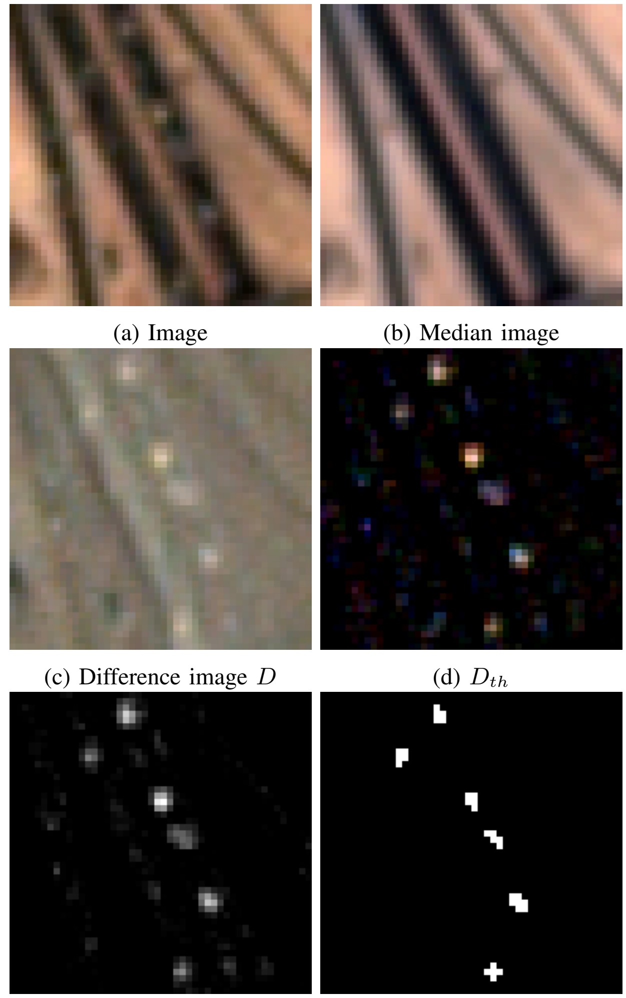
Vehicle detection is done in several steps.

The first step consists of comparing each image in the series to an image of the same area but emptied of all vehicles. Such an image can be obtained by computing the median image of the series: as there is generally a significant space between vehicles on highways, they do not cover the majority of the space. Provided a sufficiently large sample, computing a median image on the series removes all vehicles. For each image, we compute the difference image D such as:

$$D = I - M$$

I being the image in the series being processed, M being the median image. See figures 3a, 3b and 3c.

In order to detect white spots, we compute D_{th}, the minimum of all directional top-hat transform (size of structuring element: 7px) on D. See Figure 3d. We used this approach because a simple top-hat doesn't process correctly borders of roads as they can appear significantly different from an image to another. See Figure 4.



(a) Image (b) Median image (c) Difference image D (d) D_{th} (e) D_{thm} (f) Final detections

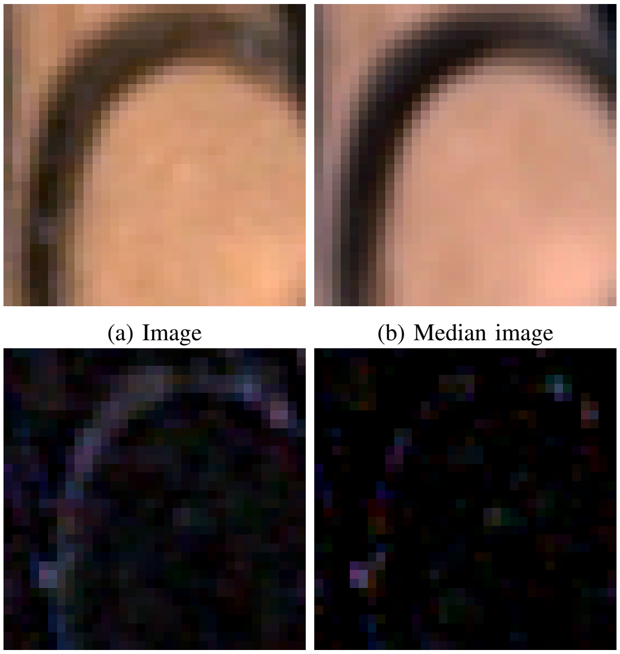
Fig. 3: Overall process for vehicule detection.

We then compute D_{thm}, a gray level version of D_{th} where for each pixel the maximum of all channels is retained, and where all pixels that are not in the road mask are set to 0. See Figure 3e.

The last step consists of detecting regions in D_{thm} that are sufficiently bright and large. For doing that, we first compute $S = \max(D_{thm})$. We then compute the threshold image $T = D_{thm} >= S$. We remove in T connected components with an area smaller than $L = 5$ pixels. Each remaining component is marked as an object. We multiply S by R (with R set to 0.9) and repeat the process several time until $S < M$, where M has been set to 100. In each step we detect connected components as vehicles only if they have not been already detected in the previous step. See Figure 3f.

III. Experimental validation

We have extracted a series of 318 images of an area in Saudi Arabia containing an highway. The acquisition date range from 09/13/2017 to 09/11/2018.



(a) Image

(b) Median image

(c) Simple top-hat on D

(d) MD top-hat on D

Fig. 4: Comparison between a simple top-hat to the minimum of all direcitonal top-hat (named MD top-hat). The change in illumination observed in the road appears when applying a simple top-hat, disrupting the car detection process at a later stage. The MD top-hat is robust to such changes.

From this data, we will validate our approach in two steps. As a ground truth is not available, we will first generate simulation images where vehicles are artificially added on the median images, and compare the number of vehicles added during the simulation to the number of vehicles detected. It will allow us to check the accuracy of our method and its robustness to noise and buriness. In the second step, we apply our algorithm to real images, and check if observed trends are consistent with reality.

A. Validation on simulation images

We first compute the median image M from the aligned series as described in Section II-B. This image will serve as a basis for the simulation, as the highways are emptied of their car.

For each image of the series, we also estimate their blur kernel using the method described in [10]. The resulting list L_{blur} of blur kernel will prove useful later during the post-processing stage of the simulation.

From this median image M, we generate 200 simulated images. For each simulated image:

1) We add a random number of vehicules: between 0 and 100. As we don't know exactly how vehicles look like on Planet images, we first enlarge the image so that its GSD is 0.8m, similar to SkySat images. At this resolution, we know that vehicules can be represented by blurry 6×2px rectangles. See Figure 5. Those rectangles are added on the highways (using the road mask) with a direction consistent with the highway they are on. The image is then reduced to its original size.
2) At this stage, the image is much less noisy and blurry as what we could expect on real acquisitions, as it is derived from the median image of the whole sample. We must therefore add noise and blur.
3) We apply a gaussian noise with $\mu = 0$ and $\sigma = 200$ on each channel separately.
4) We apply the same gaussian noise with $\mu = 0$ and $\sigma = 300$ on all channels.
5) We choose a random blur kernel from the list L_{blur}, and apply it on the image.

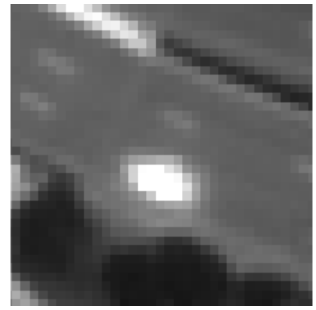


Fig. 5: SkySat image of a vehicle on a road.

Fig. 6: Simulated image.

An example of generated image can be seen in Figure 6.

The relationship between the number of detected vehicles and the number of simulated vehicles (serving as ground truth) is shown in Figure 8, from which several observations can be made.

A first observation is that, in a few cases, the number of

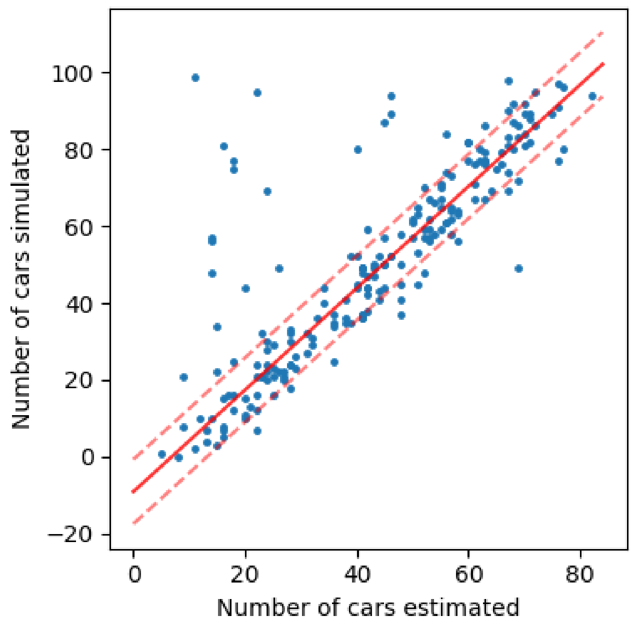


Fig. 7: Number of vehicles simulated vs detected on simulation images. $y = 1.323x - 9.077 \pm 8.377$, x being the number of detected vehicles and y being the number of simulated vehicles in the image. Linear regression obtained using RANSAC.

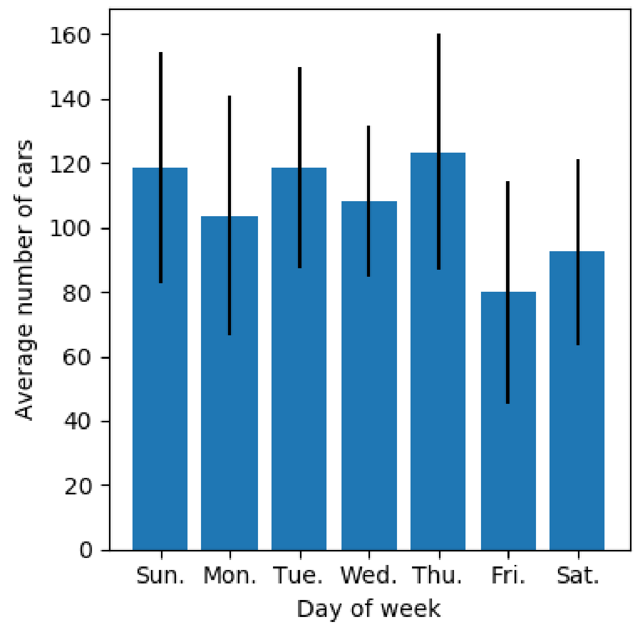


Fig. 8: Average of vehicle detected per weekday.

vehicles is significantly underestimated. After investigation, we have found that this happens when images are particularly blurry, destroying the majority of the information in the image. This is an indication that images that are too blurry should be removed from the series.

Secondly, the derived linear regression suggests that the detector under-estimates the number of vehicles by about 30%, but with a positive offset of about $9.077/1.323 \approx 6.86$ detections, probably due du noise.

Finally, overall, the relationship is linear, and the observed error seems to be independent from the number of vehicles in the images.

B. Validation on real images

The detector has been applied on the series of real images. For validating our approach, we plotted the average number of vehicle detected per weekday in Figure 8. There is clear drop of about 25% in the number of vehicles detected on weekends (from Friday to Saturday in Saudi Arabia).

IV. Conclusion

We presented in this paper a dot detector able to count vehicles on a series of Planet images of highways. We measured its precision on simulated images, and checked that its estimations are consistent with reality on real images of an highway in Saudi Arabia.

As this study has been limited to a single area in Saudi Arabia, we plan to check how our approach generalize on other areas.

The presented detector is general and could be applied to a range of different detection tasks.

References

[1] Floris J van Ruth. Traffic intensity as indicator of regional economic activity. Technical report, Citeseer.

[2] Jie Zhou, Dashan Gao, and David Zhang. Moving vehicle detection for automatic traffic monitoring. *IEEE transactions on vehicular technology*, 56(1):51–59, 2007.

[3] Zhigang Zhu, Guangyou Xu, Bo Yang, Dingji Shi, and Xueyin Lin. Visatram: A real-time vision system for automatic traffic monitoring. *Image and Vision Computing*, 18(10):781–794, 2000.

[4] Gui-Song Xia, Xiang Bai, Jian Ding, Zhen Zhu, Serge Belongie, Jiebo Luo, Mihai Datcu, Marcello Pelillo, and Liangpei Zhang. Dota: A large-scale dataset for object detection in aerial images. In *CVPR*, 2018.

[5] T Nathan Mundhenk, Goran Konjevod, Wesam A Sakla, and Kofi Boakye. A large contextual dataset for classification, detection and counting of cars with deep learning. In *European Conference on Computer Vision*, pages 785–800. Springer, 2016.

[6] S. Drouyer. Vehsat: A large-scale dataset for vehicle detection in satellite images. Manuscript submitted for publication.

[7] Wikipedia contributors. Car colour popularity — Wikipedia, the free encyclopedia, 2018. [Online; accessed 6-January-2019].

[8] T. Dagobert, J.-M. Morel, and C. de Franchis. Visibility detection in time series of planetscope images. In *International Geoscience and Remote Sensing Symposium*, 2019.

[9] Martin Rais. *Fast and accurate image registration. Applications to on-board satellite imaging.* PhD thesis, Université Paris-Saclay, 2016.

[10] J. Anger, C. de Franchis, and G. Facciolo. Assessing the sharpness of satellite images: Study of the planetscope constellation. In *International Geoscience and Remote Sensing Symposium*, 2019.