Autonomous Vehicle Detection and Classification in High Resolution Satellite Imagery

Conference Paper · March 2019

DOI: 10.1109/ACIT.2018.8672712

3 authors:

Ali J Ghandour
National Council for Scientific Research

Houssam Krayem
Islamic University of Lebanon

Abdul Karim Gizzini
École Nationale Supérieure de l'Electronique et de ses Applications

Autonomous Vehicle Detection and Classification in High Resolution Satellite Imagery

Ali J. Ghandour
National Council for Scientific Research (CNRS)
Beirut, Lebanon
aghandour@cnrs.edu.lb

Houssam A. Krayem
Lebanese University
Beirut, Lebanon
krayem.hossam@gmail.com

Abedelkarim A. Jezzini
Islamic University of Lebanon
Beirut, Lebanon
abedelkarim.j@gmail.com

Abstract

High resolution remote sensing data can provide worldwide images rapidly contrasted with conventional strategies for information accumulation. Therefore tiny objects like cars can be easily detected. Automatic vehicles enumeration research domain plays an important role in various applications including traffic monitoring and management. In this paper, we propose autonomous vehicle detection and classification approach in highway environment. Proposed approach consists mainly from three stages: (*i*) first, preprocessing operations are applied in order to eliminate noisy objects including soil, vegetation, water. (*ii*) Then, built-up area index is utilized to detect and delineate road networks. (*iii*) Finally, Multi-thresholding segmentation is implemented, resulting in vehicle detection and classification, where detected vehicles are classified into cars and trucks. Quality percentage assessment is carried over different study areas, illustrating the great efficiency of the proposed approach especially in highway environment.

***Keywords*-Remote Sensing, Object Identification, Road extraction, Vehicle enumeration and Classification.**

I. Introduction

Traffic flow data analysis is an important task in urban planning, especially for road networks design, and noise pollution estimation.Vehicle enumeration and monitoring the continuously growing vehicles fleet over time is an important issue required for the strategic planning of transport infrastructure.

Cameras and motion sensors can be considered as a traditional techniques for traffic flow monitoring. These techniques are accurate, however are spatially-limited. In developing countries, vehicle fleet monitoring and management necessary to identify alternative solutions for better transportation facilities, traffic management, and air pollution estimation, can't be applied by field based equipments.

Relying on remote sensing techniques to detect traffic flow from satellite images is an efficient, affordable, sustainable and reliable solution for the aforementioned problems, resulting in automatic traffic flow management.

In this manuscript we propose an autonomous vehicle detection and classification approach from high resolution satellite images. The aim behind our work is to: (*i*) address the power of using remote sensing in vehicle detection and enumeration research area, (*ii*) design and implement a robust autonomous approach for vehicles detection and classification from high-resolution satellite images and (*iii*) finally carry an accuracy assessment analysis over the proposed method where results for different study areas shows the high quality percentage of the proposed approach that reaches 82.25%.

The rest of this paper is organized as follows: Section 2 presents a literature survey of existing research research in the area of vehicular detection from satellite imagery. Section 3 discusses our proposed method made of three main stages for vehicular detection and classification. Performance and accuracy metrics of the proposed method are shown in Section 4 over three different study areas. And finally, conclusion and future work appear in Section 5.

II. Literature Review

Airborne imagery or aerial photography can be used to provide more global view of land scape objects like roads, traffic trajectories as described in [1] [2]. Moreover, very high resolution satellite imagery provides a new opportunity for traffic monitoring as proposed in [3]. Synthetic Aperture Radar (SAR), is a form of radar that is used to create two and three dimensional representation of objects, such as landscapes. It is based on the motion of the radar antenna over a target region to generate finer spatial resolution images. Optical data includes the visible wavebands and therefore can produce color band images, which is similar to how the human eye sees the world. Advantages and limitations of Both sources are described in [4].

978-1-7281-0385-3/18/$31.00 2018 IEEE

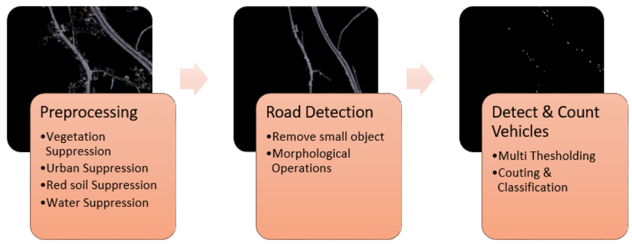


Figure 1: Methodology Flow Chart.

For instance, day/night and weather conditions affects optical data not SAR. On the other hand, optical sensors have a horizontal view compared to SAR which is oblique.This issue causes distortion and hides some image features [5]. Vehicle detection from high resolution satellite imagery has not been well investigated in the literature. Some existing research works has proposed different schemes in the literature for car detection using aerial images.

The authors of [6] propose a 3D model approach that detect and count cars in aerial images. In [7], authors uses the power of shape features for car detection and classification operation. In [8], cars are modeled to have a small rectangular form. Edge detection is implemented to extract the four sides of the rectangular boundary. A statistical framework is proposed in [9] by which cars samples and their relative features are recorded in vectors. The detection algorithm is based on computing the closeness value of each car candidate against recorded statistics.

III. Methodology

In this paper, we propose a hybrid approach for vehicle detection in highway environment using four band high resolution satellite imagery. These bands are: red, green, blue, and near infrared ($\boldsymbol{NIR}$). In this study we use four bands multi-spectral GeoEye-1 images to compute the indices needed to automate the vehicular detection process.

First of all, preprocessing operations are carried over the area under study. These operations consist of vegetation and water regions elimination using Normalized Difference Vegetation Index ($\boldsymbol{NDVI}$) and the Water Suppression Index ($\boldsymbol{WSI}$), respectively. Second, road network is masked out using the Build Up Area index $\boldsymbol{BAI}$, besides applying morphological operations for noise and unwanted objects elimination. Finally, multi-level Otsu thresholding is applied for vehicle detection, where segmented vehicles are classified into cars and trucks according to hight, width, and area parameters. Figure **??** presents a flow chart of our proposed methodology.

A. Preprocessing

1) Vegetation and Red soil Suppression: The presence of vegetation areas within satellite images can be considered as noisy regions that needs to be eliminated.

Normalized Difference Vegetation Index ($\boldsymbol{NDVI}$) defined in Equation 1, is used to generate vegetation area mask, that will be used to mask out all green regions within image [18]. It depends on measuring chlorophyll percentages of green vegetation that absorbs red light and reflects near-infrared wavelengths [11].

Extracting trees (and other vegetation areas) is done by computing $\boldsymbol{NDVI}$ and then a thresholding operation is applied on the NDVI image, resulting in masking out all the areas with vegetation. Figure 2 show the $\boldsymbol{NDVI}$ layer of Study Area 3, where we can notice that vegetation has the highest intensity, and road has the lowest intensity. Study Areas used for the evaluation of this paper will be discussed in the coming sections.

$$NDVI = \frac{NIR - R}{NIR + R} \tag{1}$$

2) Water Suppression: Water bodies are characterized by high reflectance values in $\boldsymbol{G}$ band, with low ones in $\boldsymbol{R}$ band. Then, in order to mask out water regions, we use the Water Suppression Index $\boldsymbol{WSI}$ defined in Equation 2 [12].

$$WSI = \frac{G - R}{G + R} \tag{2}$$

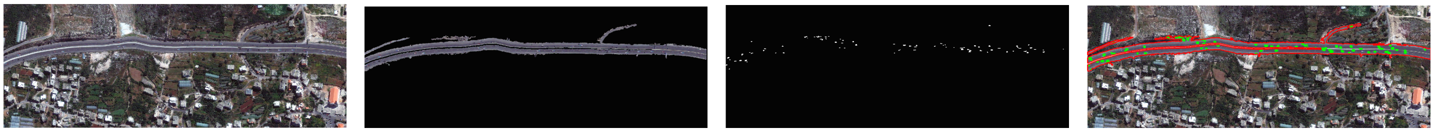
3) Urban Suppression: After eliminating vegetation and water regions, once can only notice the presence of buildings in addition to the road network. In order to obtain accurate results in the road detection, all bright pixels must be removed since road network can be identified by its low intensity. Bright pixels are pixels that have high intensity in each channel in the RGB image. Figure 4 shows the preprocessing resultant image of Study Area 3 after that the urban suppression step has been applied.

B. Road Detection

Generally, highway road networks in satellite images have gray color with a straight elongated geometric shape. There exist a numerous set of features that can be used in order to define the road network. Generally, roads in urban areas are presented as the signs of $\boldsymbol{T}$ or $\boldsymbol{U}$ shape, also known as geometric features.

Texture features can also define the road network where we depend on the homogeneity of the road regions in the

Figure 2: NDVI layer of Study Area 3



(a) (b) (c) (d)

Figure 3: Figures 3a, 3b, 3c and 3d show the original image, besides applying road detection, vehicle count, and the results fusion, respectively, for Study Area 1.

extraction process. Moreover, photometric features, which we implemented in our paper, rely on the reflectance values of roads color within the satellite image in order to mask out the road networks. A detailed description and analysis about road networks is presented in [19].

1) Built-Up Area Index: Similarly as $\boldsymbol{NDVI}$, remote sensing provides another robust and reliable index to detect asphalt and concrete surfaces in satellite images [16] [17].

$\boldsymbol{BAI}$ defined in Equation 3, is used in our algorithm for road detection operation.

$$BAI = \frac{B - NIR}{B + NIR} \tag{3}$$

2) Bright Vehicles and Morphological Operations: Applying the last preprocessing operation for urban areas suppression has a side effect on some of the vehicles existing in the area under study. In fact, the urban suppression operation leads to the problem that bright vehicles has been removed and hence isolated pixels appeared in the road network.

To resolve this issues, we transform the preprocessed image to a binary image, in order to remove isolated pixels, fill holes, and re-mask the new binary image with original image.

3) Small Objects Removal: The final challenge in the road detection algorithm is the existence of unwanted small objects. Since the street is the largest object in the image, we remove all objects having a size smaller than the size of the smallest road.

And then, finally the road network is extracted as shown in Figures 3b, 5b, and 6b which depict the road detection resultant output of Study Area 1, 2, and 3 respectively.

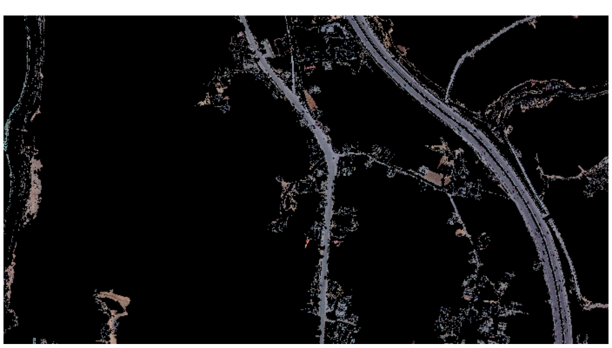


Figure 4: Study Area 3 after Preprocessing

C. Detect & Count Vehicles

1) Multi-Thresholding Segmentation: Image segmentation techniques is used to group pixels into different image regions, belonging different objects. Segmentation is utilized in several applications including object recognition, image compression, and image editing.

k-means clustering, k-nn segmentation, maximum likelihood are popular segmentation techniques. All these methods characterized by its implementation complexity.

In our approach, we use multi-thresholding technique as an image segmentation operation as a segmentation technique, which is more efficient with less implementation complexity.

The Otsu threshold [13] is the bedrock of multi-thresholding segmentation. An optimal threshold is selected by the discriminant criterion to maximize the separability of the resultant classes by utilizing only the zeroth-and the first-order cumulative moments of the test image histogram. Using this technique, bright vehicles will be detected, in addition to some false objects like highway lane markers and road dividers. Eliminating false detected misclassification, is applied through implementing a sliding neighborhood filter to the test image.

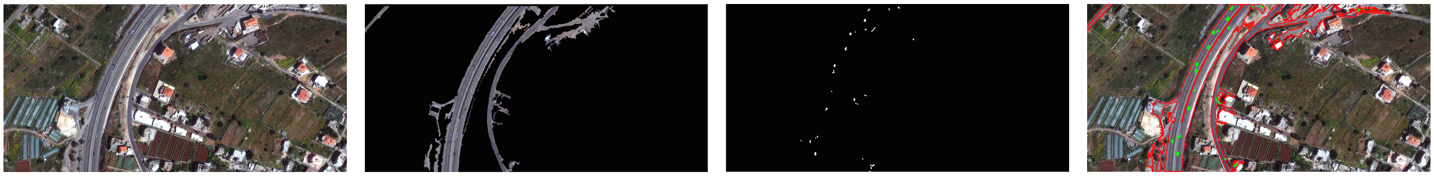
This filter is defined in Equation 4, where $\boldsymbol{F(x,y)}$ represents the new image generated after applying the max filter by using the $a \times b$ sliding window $\boldsymbol{W(s,t)}$ on the original image $\boldsymbol{I}$. The behavior of this filter assign the maximum neighborhood intensity value for each pixel within the image.

$$F(x,y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} \max(W(s,t).I(x+s, y+t)) \tag{4}$$

Sliding neighborhood filter will brighter vehicles pixels, while maintaining pixels brightness corresponding to irrelevant objects such as lane markers. As a result, vehicles will be highlighted, while lane markers will be eliminated.

Finally, we apply a filter to remove all objects with a size smaller than and greater than the vehicle size to obtain the vehicles output image as shown in Figures 3c, 5c, and 6c that depict vehicle detection resultant output of Study Area 1, 2, and 3 respectively.

2) Classification: Width, height and area of the detected vehicles are considered as classification parameters distinguishing vehicles into cars and trucks. The threshold for each parameter is set to be the average value of all vehicles candidates. Vehicles with width, height, and area greater than the average values is classified as a truck, otherwise it is a car.



(a) (b) (c) (d)

Figure 5: Figures 5a, 5b, 5c and 5d show the original image, besides applying road detection, vehicle count, and the results fusion, respectively, for Study Area 2.

IV. Evaluation Results

In this section we will discuss the implementation results of the proposed vehicle detection algorithm using an accuracy assessment model.

A. Study Areas

We choose three Study Areas to assess our proposed algorithm. In the following, we elaborate on the details of each Study Area:

- Study Area 1: a wide image for a straight section of the Saida-Beirut HighWay. This Study Area contains buildings and vegetation areas.
- Study Area 2: is characterized by the same environmental conditions as Study Area 1 but the main difference is that it contain an additional road branch close to the highway with a ramp structure. This study area is useful to assess the robustness of the road detection stage in our proposed method.
- Study Area 3: besides buildings and vegetation areas, this Study Area is characterized by the presence of red soil areas and sea water.

B. Performance Metrics

We will analyze vehicles detection performance using Lin and Nevatias formulas proposed in [10]. They define the Miss Factor ($\boldsymbol{MF}$), Branching Factor ($\boldsymbol{BF}$), and the Quality Percentage ($\boldsymbol{QP}$) for reporting algorithms performances. Therefore, all objects in the image will be categorized as follows:

- True Positive (TP): Both manual and automated methods label the vehicle.
- True Negative (TN): Both manual and automated methods label the object belonging to the background.
- False Positive (FP): The automated method incorrectly labels the object as a vehicle.
- False Negative (FN): The automated method does not correctly label truly vehicles objects.

Our approach performance will be evaluated using the following matrices:

$$MissFactor(MF) = \frac{FN}{TP} \tag{5}$$

$$BranchingFactor(BF) = \frac{FP}{TP} \tag{6}$$

$$Qualitypercentage(QP) = 100 \times \frac{TP}{TP + FP + FN} \tag{7}$$

The miss factor represents the error where our approach incorrectly labels vehicles as background. The branching factor measures the error where our approach incorrectly labels background as vehicle. Quality percentage measures the overall performance of vehicle detection operation.

Our proposed method was tested on the Study Areas listed above. Ground truth data for each study area is compared with the obtained algorithm results to estimate the overall quality percentage. Figures 3, 5, and 6 shows the resultant images of applying our vehicle detection approach on Study Area 1, 2, and 3 respectively.

Table I: Accuracy Assessment Metrics for cars detection.

Study Area	Ground Truth	TP	FP	FN	BF	MF	QP
1	40	38	4	2	0.1	0.05	86.36%
2	12	10	2	2	0.2	0.2	71.4%
3	38	35	5	3	0.14	0.08	81.39%

Accuracy assessment records for different study areas is shown in table I. The main advantage witnessed is that $\boldsymbol{TP}$ value is very high in all tested scenarios which means that our proposed vehicle detection algorithm have succeeded in the detection operation in most cases. We noticed that the performance of the proposed algorithm obviously varies according to certain conditions, like the presence of bushes and trees. Moreover, in heavy traffic cases, our approach would count cars very close to each other as one object, thus affecting the final quality percentage.

Trucks detection accuracy assessment analysis is shown in Table II. The quality percentage results recorded for Study Areas 1, 2, and 3 respectively reveals the high performance of the classification process done by our introduced approach. We can notice that the $\boldsymbol{FP}$ values in truck detection is less than $\boldsymbol{FP}$ values for car detection. This is due to the separability parameters used in the classification phase where unwanted objects are removed so the probability to obtain false trucks objects is low. Moreover, some real trucks are not detected

(a) (b) (c) (d)

Figure 6: Figures 6a, 6b, 6c and 6d show the original image, besides applying road detection, vehicle count, and the results fusion respectively, on Study Area 3.

since it is close to the road boundary or close to another truck, so as a result our algorithms count them as one truck.

Table II: Accuracy Assessment Metrics for trucks detection.

Study Area	Ground Truth	TP	FP	FN	BF	MF	QP
1	15	13	3	2	0.2	0.15	72.22%
2	6	5	0	1	0	0.2	83.33%
3	19	15	1	4	0.05	0.26	75%

As an overall analysis, we can say that our hybrid approach developed for vehicle detection is capable to detect and classify vehicles in highway road networks. Further improvements will be applied later to decrease the false detected vehicles, thus improving the overall quality percentage of the suggested approach.

V. Conclusion and Future Work

In this manuscript, we propose a vehicle detection approach based mainly on three phases: *(i)* Preprocessing, *(ii)* Road Detection, and *(iii)* Vehicle detection and classification. In the first phase, vegetation regions, soil, and water areas is eliminated using the reflectance values in different image bands. The Second phase consist mainly of road detection where $\boldsymbol{BAI}$ is used, besides a series of morphological operations in order to detect roads in the test image under study. After that, multi-thresholding segmentation is used to detect vehicle objects. Finally, width, height, and area parameters of the detected vehicles is used to classify vehicles as car or trucks.

As future work, machine learning will be utilized in thresholds estimation [14]. Moreover, we will use more vehicles features such as color and shape, in addition to implementing a probabilistic model for road detection [15],improving the overall quality percentage of the proposed approach.

Acknowledgment

This project has been funded with support from National Council for Scientific Research in Lebanon.

References

[1] Chengcui Zhang, Chen, S.C. Shyu,and M.L.Peeta, “Adaptive Background Learning for Vehicle Detection and Spatio-Temporal Tracking. Proceedings of the Fourth IEEE Pacific-Ri Conference On Multimedia,” Singapore, pp: 1-5, December, 2003.

[2] S. Hinz, “Detection and Counting of Cars in Aerial Images,” Proceedings of the International Conference on Image Processing,Barcelona, vol: 3, pp:997-1000, September, 2003.

[3] Rocio Alba Flores, “Evaluation of the use of high resolution satellite imagery in transportation applications. Final Report, Department of Electrical and Computer Engineering, University of Minnesota Duluth 271 MWAH,” Duluth, 2004.

[4] U. Stilla, E. Michaelsen, U. Soergel, S. Hinz, and J. Ender, “Airborne monitoring of vehicle activity in urban areas. Proceedings of the XXth ISPRS Congress,” Istanbul, pp: 973, July, 2004.

[5] P. Lohmann, K. Jacobsen, K. Pakzad, and A. Koch, “Comparative information extraction from SAR and optical imagery. Proceedings of the XXth ISPRS Congress,” Istanbul, pp: 535, July, 2004.

[6] Schlosser, C. Reitberger, and S. Hinz, “Automatic Car Detection in High Resolution Urban Scenes Based on an Adaptive 3D-Model,” Proceedings of the IEEE/ISPRS joint Workshop on Remote Sensing and Data Fusion over Urban Areas, Berlin, pp:167-171, 2003.

[7] T. Zhao, and R. Nevatia, “Car Detection in Low Resolution Aerial Image,” Proceeding of the 8th IEEE International Conference on Computer Vision, Vancouver, Canada, Vol:1, pp: 710-717, July, 2001.

[8] H. Moon, R. Chellappa, and A. Rosenfeld, “Performance Analysis or a Simple Vehicle Detection Algorithm,” Proceedings of the 3th ARL Federal Laboratory Symposium, February, 1999.

[9] H. Schneiderman, and T. Kanade, “A Statistical Method for 3D Object Detection Applied to Faces and Cars,” Proceedings of the Computer Vision and Pattern Recognition Hilton Head, vol:1, pp. 746-751, June, 2000.

[10] C. Lin and R. Nevatia, “Building detection and description from a single intensity image,” Computer Vision and Image Understanding,vol. 72, no. 2, pp. 101-121, November 1998.

[11] J.G. Lyon, D. Yuan, “A Change Detection Experiment using Vegetation Indices,” Photogramm.Eng. Remote Sensing 64, pp: 143-150, 1998.

[12] X. Jin, C.H. Davis, Information Fusion 6, pp: 257273, 2005.

[13] Otsu, “A threshold selection method from gray-level histograms,” IEEE transactions on systems, man, and cybernetics. vol: SMC9, no:1, pp.62-66. January, 1979.

[14] L. Shapiro, and G. Stockman, “Compute Vision,” Prentice Hall, p: 114, 2001.

[15] Cem nsalan and Beril Sirmacek, “IEEE transactions on geoscience and remote sensing,” Vol: 50, no: 11, November, 2012.

[16] B. Guindon, Y. Zhang, and Dillabaugh, “Landsat urban mapping based on a combined spectral-spatial methodology,” Remote Sensing Environ, pp: 218-232, 2004.

[17] M. Wulder, J. White, G. Masek, and J. Dwyer, “D.P. Continuity of landsat observations: Short term considerations,” Remote Sensing Environ, pp: 747-751, 2011.

[18] C. Small, and Lu, “J.W.T. Estimation and vicarious validation of urban vegetation abundance by spectral mixture analysis,” Remote Sens. Environ, pp: 441-456, 2006.

[19] Weixing Wang, Nan Yang, Yi Zhang, Fengping Wang, Ting Cao, and Patrik Eklund, “A review of road extraction from remote sensing images,” Journal of Traffic and Transportation Engineering, pp: 271-282, 2016.