International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 4 Issue 06, June-2015

Vehicle Detection and Classification from High Resolution Aerial Views using Morphological Operations

Anu Anto
M-Tech: Dept of Electronics and Communication
Marian Engineering College
Thiruvananthapuram, India

Minnu Jayan. C
Asst. professor: Dept of Electronics and Communication
Marian Engineering College
Thiruvananthapuram, India

***Abstract*—Traffic monitoring by aerial images is an important issue. Here, a novel method for vehicle detection from high resolution aerial images and videos is put forward. The system exploits the background of the aerial view for better performance. It works by identification of structural element followed by morphological operations. The morphological operations include opening and top-hat transformation for light background components and closing followed by bot - hat transformation for dark background components. Then big objects are sieved using an area threshold which is larger than a typical vehicle followed by morphological opening transformation to remove targets whose width is narrower than the diameter of structural element utilized in the morphological operations. Finally the results are overlaid in order to amalgamate those vehicles detected by both cases. The experimental result of this method on highway aerial image of 0.15×0.15 m spatial resolution and a video footage shows that the method is robust and efficient. In addition to this a method for classification of vehicles is also proposed based on morphological operations and component parameters of the images which classify the detected vehicles into big and small vehicles. Direction of vehicles is also found.**

Keywords—Morphological operation, structural element, vehicle classification, vehicle detection.

I. INTRODUCTION

In this current century, the number of vehicles on roads is increasing as the population. This has resulted in traffic jam and with our life style of moving with time where even a second is important traffic congestion is a serious issue and should be controlled. For controlling traffic, traffic monitoring was using induction loops, ground sensors, stationary cameras, etc. for collecting data. But these could acquire only partial information of the vehicles passing through the areas where they are placed. But inorder to get a complete solution we need to monitor those roads connected to these congested roads like highways and inorder to acquire area wide information of highways aerial images were used.

Traffic monitoring by using aerial images require high resolution images since then only vehicles could be detected efficiently. In [1] 0.15m resolution aerial image is used where cars are identifiable and gave promising result.

For vehicle detection from aerial images the existing techniques had classified the vehicles as implicit or explicit model before detection. After the model is decided then only the detection is done. The explicit model classify vehicle as a box or uses a wide frame representation and then use top-down or bottom-up approach for classification where the implicit model uses a training set to extract features of vehicles and detect them.

In [2] satellite image is used for vehicle detection based on region of interest and classification of vehicles is also done.

H. Zheng and Li Li in [3] present an immunological vehicle detection method using an aerial image of 0.6m resolution. This method had an advantage of direct detection of vehicles from single satellite image since till then multi-temporal image were used but its overall performance is low.

R. Mishra in [4] introduced automatic vehicle detection using MS-1 and MS-2 images. This method is a low cost method to extract moving vehicle information which can be used for traffic control but it is time consuming.

R. Ruskone et.al in [5] used a hierarchical method for vehicle detection. This is a noise resistant method and can be used for different resolution without changing learning bases but arises a question whether it's operational or not.

All the above mentioned works have limitation that decreases the performance of vehicle detection. The main aim of vehicle detection is to detect vehicles with maximum efficiency (approximately 100%). So a novel method is proposed in this paper where the detection is based on background where if light background is encountered then opening and top-hat transformation are performed and if dark background then closing and bot-hat transformation are done.

After detection classification of the detected vehicles based on its size and direction is also done.

II. PROPOSED METHOD

For vehicle detection we propose a novel method whose flowchart is given in Fig. 1. First of all structural element is identified based on input. Then based on background of the image morphological operations are done. If it is a light background image/frame then we do opening followed by grayscale top-hat transformation and Otsu partitioning [11]. For dark background we make use of closing followed by grayscale bot-hat transformation and Otsu partitioning. Then we do opening to both cases in order to filter out the vehicles from the image.

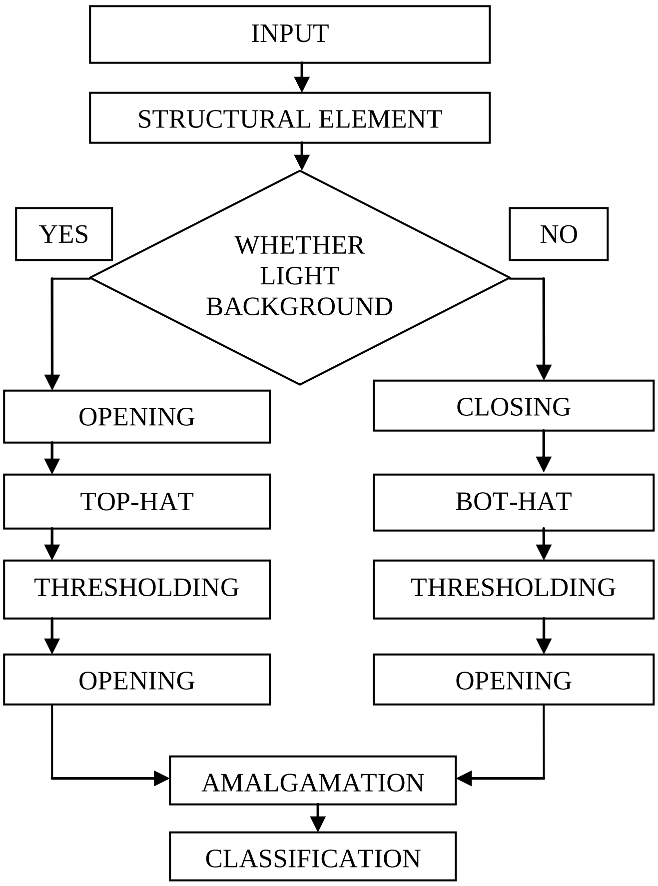


Fig. 1. Flowchart of proposed system

A. *Identification of Structural Element*

Fig. 2. A highway image with a spatial resolution of 0.15 × 0.15m

Fig. 3. Frame of a highway footage

A GIS road vector map is taken as the input in Fig. 2 in case of image as input and 1st frame is taken as input in the case of video input to constrain the vehicle detection to highway networks. By using 0.15 × 0.15m ground resolution aerial image (Fig. 2) even small cars are visible. A car is symmetrical in width and its width dimension is 14-16 pixels. So we select a disc with radius 5 pixels as the structural element.

B. Morphological Operation

Opening generally smoothes the contour of an object, breaks narrow isthmuses and eliminates thin protrusions. Closing also tends to smooth sections of contours but as opposed to opening, it generally fuses narrow breaks and long thin gulfs, eliminates small holes and fills gaps in the contour. Both are defined as:

$$Opening: f \circ b = (f \ominus b) \oplus b \qquad (1)$$
$$Closing: f \bullet b = (f \oplus b) \ominus b \qquad (2)$$

where f presents the original highway image, b is the structure element; '○' presents the grayscale opening transformation, and '●' presents the grayscale closing transformation and '⊖' and '⊕' are erosion operator and dilation operators respectively. The grayscale top-hat transformation and grayscale bot-hat transformation are defined as:

$$Top - hat: T = f - f \circ b \qquad (3)$$
$$Bot - hat: B = f \bullet b - f \qquad (4)$$

where '*T*' presents the image after the grayscale top-hat transformation and '*B'* is the image after the grayscale bot-hat transformation.

After opening transformation bright pixels representing dashed white lines for lane division and concrete road dividers with bright surface background as well as a little target of bright color are filtered out as noise. After closing pixels with low intensity like new pavement with tar or a target with a dark or black color are filtered out as noise. So both opening and closing transformation can be considered as background estimation techniques.

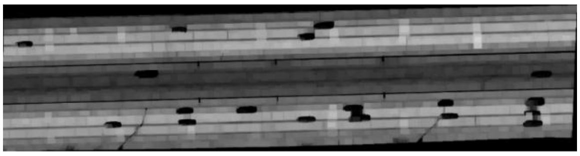


Fig. 4. A grayscale opening transformed image

Fig. 5. A grayscale opening transformed frame

Fig. 6. A grayscale top-hat transformed image

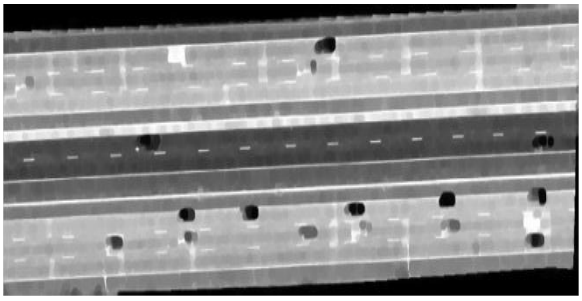


Fig. 7. A grayscale closing transformed image

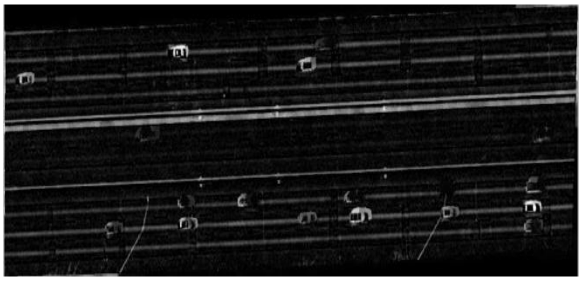


Fig. 8. A grayscale bot-hat transformed image

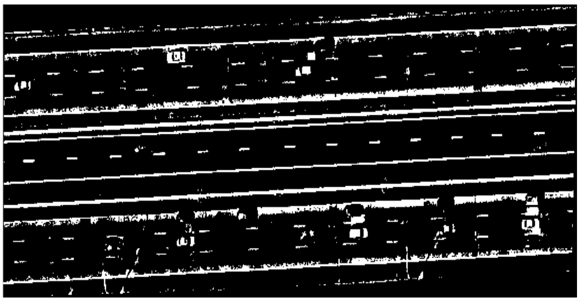


Fig. 9. A grayscale light background image after thresholding using the Otsu's method

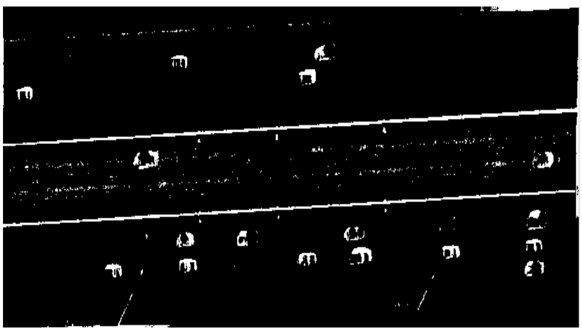


Fig. 10. A grayscale dark background image after thresholding using the Otsu's method

Fig. 11. A grayscale top-hat transformed frame

Fig. 12. A grayscale closing transformed frame

Fig. 13. A grayscale bot-hat transformed frame

Fig. 14. A grayscale light background frame after thresholding using the Otsu's method

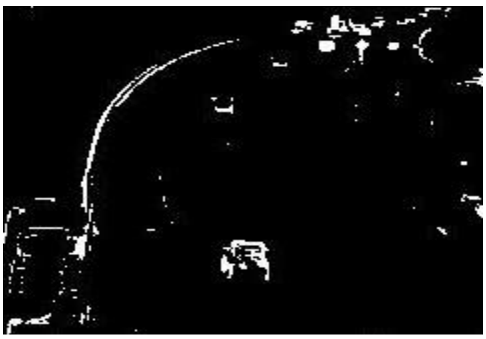


Fig. 15. A grayscale dark background frame after thresholding using the Otsu's method

C.Vehicle Detection

By applying Otsu's threshold method to grayscale top-hat and bot-hat transformation results black and white or binary image could be obtained. Targets with area larger than threshold pixels are sieved in the detection process in the case of top-hat. It is not possible in the case of bot-hat since vehicles of dark background will be sieved together with background. In addition little targets whose width are less than 10 pixels can be filtered out using morphological opening transformation with a disc shaped structural element with a radius of 4 pixels. Thus vehicles of dark background and light background are detected.

Fig. 16. Vehicle detection results using grayscale morphological operation on light background image

Fig. 17. Vehicle detection result using grayscale morphological operation on dark background image

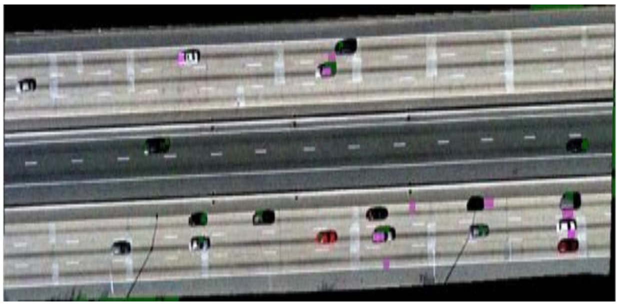


Fig. 18. Vehicle detection results from two cases of image

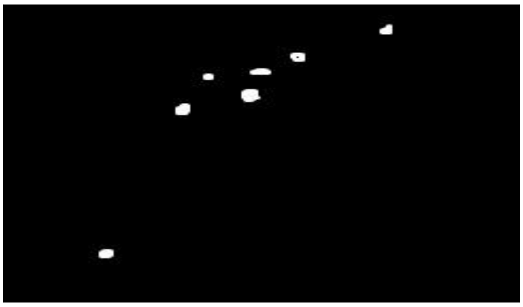


Fig. 19. Vehicle detection result using grayscale morphological operation on light background frame

Fig. 20. Vehicle detection result using grayscale morphological operation on dark background frame.

Fig. 21. Vehicle detection results from two cases of frame

D. Classification

(a) Based on Size

The algorithm for classifying the detected vehicles into cars and trucks is given below:

i. Dilate the image.
ii. Select the structural element SE as

$$SE = [1\ 1\ 1;\ 1\ 1\ 1;1\ 1\ 1]$$

iii. Use connected component labeling with

$$\text{Connectivity} = 4$$

iv. Find area, major axis length and minor axis length of each connected components.
v. If area > threshold value and minor axis/major axis length < 0.5, increment truck count by 1 else increment car count by 1.

The detected result (amalgamation of vehicles of both light and dark background) is taken and is dilated inorder to combine those vehicles which may have broken into pieces during segmentation process. It also removes irregular shapes occurring after segmentation which may reduce the detection performance. The structural element is selected as above since the vehicles are detected as rectangle/square. If it is of other shapes then other structural element could be used.

After the algorithm is performed display the result in a message box as shown in Fig. 22 and Fig. 23 by converting the numeric to string value for car and truck counts respectively.

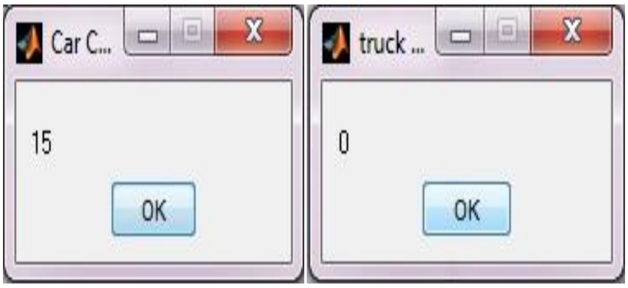


Fig. 22. Car count and Truck count of aerial image shown in Fig. 18

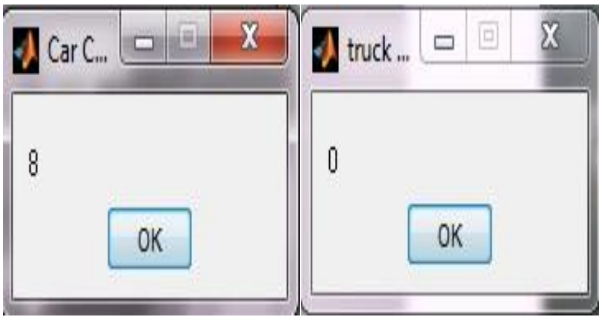


Fig. 23. Car count and Truck count of aerial frame shown in Fig. 21

(b) Based on Direction

The algorithm for finding the direction of flow of vehicle is given below:

i. From input select region of interest.
ii. Perform opening transformation
iii. Find centroid of the vehicle.
iv. The centroid of components (x1,y1) and (x2,y2) where if x2>x1 then the vehicle is moving right (leaving highway) and if x1>x2 then vehicle is moving left(entering highway) and if both are equal then stationary.

Fig. 2 and Fig. 3 are the inputs. In case of images the marked vehicles in Fig. 24 are taken as region of interest. Both their centroids are found and are found that only the x axis values vary whereas the y axis values are almost constant. Hence if the centroid value corresponding to x-bar varies more than 5 then are considered as moving left or right. The y-bar value varies 2 to 3 units only and hence is not considered.

In the case of videos the 1st and 10th frame are taken and the marked car in Fig. 25 is taken as region of interest. Since the car is moving forward the 2nd frame x–bar value is less compared to 1st frame.

After the direction is found using the algorithm display the result in a message box as string value as moving left right or whether vehicle is stationary.

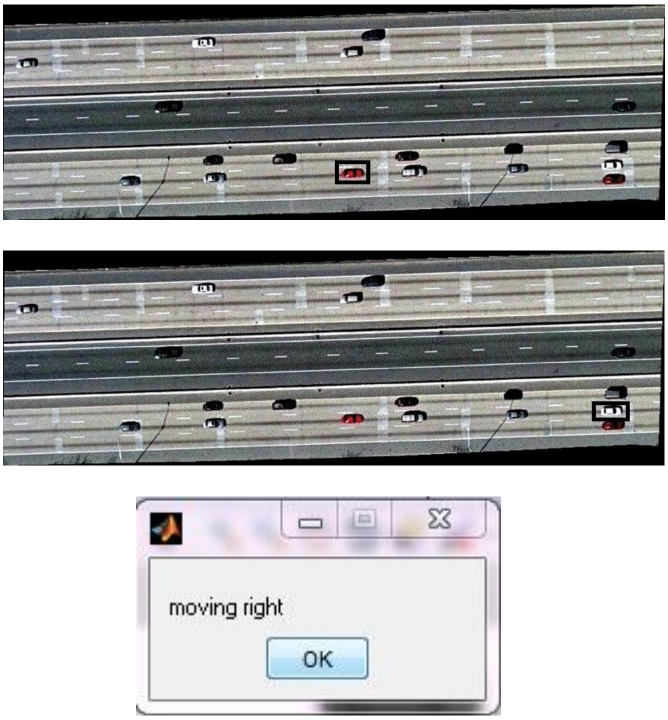


Fig. 24. (a) First region of interest of aerial image (b) Second region of interest of aerial image (c) Direction of flow of vehicle

Fig. 25. (a) First frame showing region of interest (b) Tenth frame showing region of interest (c) Direction of flow of vehicle

III. EVALUATION OF THE RESULT

A. *Datasets*

In this paper we used a GIS road vector map covering streets and highways of the city of Norfolk, Virginia, USA which was downloaded from the Hampton Roads Transportation Planning Organization (http://hrtpo.org/) as the aerial image input. The space resolution of the image was 0.15 × 0.15m. Six highway scenes were selected for evaluation of proposed method using aerial images. (Fig. 2 was numbered as scene 2).

Also we used a aerial footage as the input in the case of detection using aerial video clip.

B. *Vehicle Detection Results*

(i) *Image*

By using the proposed method vehicles were detected from both cases using scene 2. Then both cases were overlaid. The green dots denote those vehicles detected by opening and pink dots denote those vehicles denoted by closing. From Fig. 18 we could see that some cars were detected by both cases. So by amalgamating both the cases we were able to count these as a single car. Similarly vehicle detection results of five typical scenes are shown in Fig. 26.

For evaluation of vehicles detection results we used a numerical accuracy assessment by comparing the input image Fig. 2 and the output image Fig. 18. According to Wiedemann et. al [13] three categories of extraction results are defined as follows:

True Positive (TP): the number of correctly extracted true vehicles.

False Positive (FP): the number of incorrectly extracted false vehicles.

False negative (FN): the number of omitted vehicles.

Based on these definitions, three statistical measures are used in our study:

$$Correctness\ (Cr) = \frac{TP}{TP+FP} \times 100\% \qquad (5)$$

$$Completeness\ (Cm) = \frac{TP}{TP+FN} \times 100\% \qquad (6)$$

$$Quality\ (Q) = \frac{TP}{TP+FP+FN} \times 100\% \qquad (7)$$

The correctness is a measure ranging between 0 and 1 that indicates the detection accuracy rate relative to ground truth. The correctness and completeness are the converse of commission and omission errors, respectively. The two measures are complementary and need to be interpreted simultaneously. The quality shows the overall accuracy of the extraction method, and the quality value can never be higher than either the completeness or correctness value. Table I reports the three calculated measures for the six highway scenes.

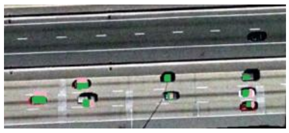
TABLE I Vehicle Detection Results Of The 6 Highway Scenes

Highway	TP	FP	FN	Cr%	Cm%	Q%
Scene 1	6	0	0	100	100	100
Scene 2	15	2	2	88.24	88.24	78.95
Scene 3	8	0	1	100	88.88	88.88
Scene 4	17	2	1	89.47	94.44	85
Scene 5	16	2	0	88.88	100	88.88
Scene 6	35	1	2	97.22	94.59	92.1
Total	97	7	6	93.96	94.36	88.97

TABLE II Vehicle Detection Results Of The 6 Highway Scenes After Classification Based On Size

Highway	TP	FP	FN	Cr%	Cm%	Q%
Scene 1	6	0	0	100	100	100
Scene 2	15	0	2	100	88.24	88.24
Scene 3	8	0	1	100	88.88	88.88
Scene 4	17	0	1	100	94.44	94.44
Scene 5	16	0	0	100	100	100
Scene 6	35	0	2	100	94.59	94.59
Total	97	0	6	100	94.36	94.36


(a)


(b)


(c)


(d)


(e)

Fig. 26. Vehicle detection results of other five typical scenes. (a) Scene 1 (b) Scene 3 (c) Scene 4 (d) Scene 5 (e) Scene 6

(ii) Video

The video frames were taken and same evaluation was conducted. The result is shown in Fig. 27 and Table III and IV.

(a)Frame 10

(b)Frame 20

(c)Frame 30

(d) Frame 40

(e) Frame 53

Fig. 27. Vehicle Detection result of other selected frames

TABLE III Vehicle Detection Results Of The 6 Highway Frames

Highway	TP	FP	FN	Cr%	Cm%	Q%
Scene 1	8	0	2	100	80	80
Scene 2	5	1	1	83.3	83.33	71.42
Scene 3	5	0	1	100	83.33	83.33
Scene 4	3	2	2	60	60	42.85
Scene 5	4	2	1	66.66	80	57.14
Scene 6	3	2	1	60	75	50
Total	28	7	8	78.32	76.94	64.12

TABLE IV Vehicle Detection Results Of The 6 Highway Frames After Classification Based On Size

Highway	TP	FP	FN	Cr%	Cm%	Q%
Scene 1	8	0	2	100	80	80
Scene 2	5	0	1	100	83.33	83.33
Scene 3	5	0	1	100	83.33	83.33
Scene 4	3	0	2	100	60	60
Scene 5	4	0	1	100	80	80
Scene 6	3	0	1	100	75	75
Total	28	0	8	100	76.94	76.94

C. *Discussion*

From Fig. 18 fifteen vehicles were detected and two noisy areas were detected as cars and two cars were not detected. No false vehicles were detected from light background but two were detected from dark background, since the vehicle targets of dark background adhered to the background together, false vehicles with big area could not be sieved using area threshold. In scene 2 the correctness and completeness were up to 88.24% respectively and the quality came up to 78.95%.

From TABLE 1, the correctness, completeness and quality of each scene could be found out. In summary of all the six scenes our proposed method for vehicle detection resulted in a correctness of 93.96%, completeness of 94.36% and quality of 88.97%. Thus vehicle detection results of our proposed method are satisfactory.

After application of classification based on size from TABLE II, we could see that the correctness increased to 100% since those false vehicles detected where removed using multiple thresholding and the quality increased to 94.36%. The quality or the performance depends upon the threshold value and so it should be chosen with utmost care.

As in the case of video from TABLE III and TABLE IV we can see that the false vehicles detected were eliminated but the quality increased to 76.94%. So we could say that the proposed method is an efficient one.

As in the case of finding the direction of flow of vehicle, by analyzing Fig. 24 and Fig. 25 we could say that the algorithm is efficient.

IV. CONCLUSION AND FUTURE WORK

A novel method for vehicle detection and classification from aerial views has been proposed which used morphological operations namely opening, closing, gray scale bot-hat and gray scale top-hat transformations. A GIS road vector map of 0.15×0.15 m ground resolution was used as image input and footage as video input. The method was tested on six highway scenes and got correctness, completeness and quality as 100%, 94.36% and 94.36% respectively for image input and 100%, 76.94% and 76.94% respectively for video input. So we could conclude that this method is efficient. Also the flow of vehicles has also been found out efficiently.

For future work, the system can be improved and extended with the following aspects:

- Increasing the ground resolution of the image from 0.15 to 1m by taking the input image of 1m resolution and adjusting the radius of disc used as structural element.

ACKNOWLEDGMENT

I would like to express my gratitude to my institution for giving me an opportunity to do this work. I would also like to express my sincere gratitude to my colleagues and staff of my department for their advice and encouragement throughout this work.

REFERENCES

[1]. Z. Zheng, G. Zhou, Y. Wang, Y. Liu, X. Li,X. Wang, and L.Jiang, "A Novel Vehicle Detection Method With High Resolution Highway Aerial Image" IEEE Journal of Selected topics in Applied Earth Observations and Remote Sensing, vol. 6, no. 6, pp. 2338- 2343 ,2013

[2]. Lizy Abraham and M.Sasikumar, "Vehicle detection and classification from high resolution satellite images", ISPRS Annals of the Photogrametry, Remote Sensingand Spatial Information Sciences, Volume II-1, 2014,pp. 1-8.

[3]. S. Hinz, J. Leitloff, and U. Stilla, "Context-supported vehicle detection in optical satellite images of urban areas," in Proc. IEEE IGARSS, 2005, vol. 4, pp. 2937–2941

[4]. H. Zheng and L. Li, "An artificial immune approach for vehicle detection from high resolution space imagery," International Journal of Computer Science and Network Security, vol. 7, no. 2, pp. 67–72, 2007

[5]. R. K. Mishra, "Automatic moving vehicle's information extraction from one-pass worldview-2 satellite imagery," in International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences, 2012, vol. XXXIX-B7, XXII ISPRS Congress, pp. 323–328

[6]. R. Ruskone, L. Guigues, S. Airault, and O. Jamet, "Vehicle detection on aerial images: A structural approach," in Proc. Int. Pattern Recognition, 1996, pp. 900–904

[7]. Papageorgiou and T. Poggio, "A trainable system for object detection," Int. J. Computer Vision, vol. 38, no. 1, pp. 15–33, 2000

[8]. S. Kluckner, G. Pacher, H. Grabner, H. Bischof, and J. Bauer, "A 3-D teacher for car detection in aerial images," in Proc. IEEE International Conference Computer Vision, 2007, pp. 1–8

[9]. T. Nguyen, H. Grabner, B. Gruber, and H. Bischof, "On-line boosting for car detection from aerial images," in Proc. IEEE Int. Conf. Research, Innovation and Vision for the Future (RIVF07), 2007, pp. 87–95

[10]. H. Grabner, T. Nguyen, B. Gruber, and H. Bischof, "On-line boosting based car detection from aerial images," ISPRS Journal for Photogrammetry and Remote Sensing, vol. 63, no. 3, pp. 382–396, 2008

[11]. N. Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. Syst., Man Cybern., vol. 9, no. 1, pp. 62–66, 1979.

[12]. C. Wiedemann, C. Heipke, H. Mayer, and O. Jamet, "Empirical evaluation of automatically extracted road axes," in Empirical Evaluation Methods in Computer Vision, K. Bowyer and P. Phillips, Eds. New York, NY, USA: IEEE Comput. Soc. Press, 1998, pp. 172–187.