Note

SatTrafikk

Project results 2008

Note no	SAMBA/50/08
Authors	Siri Øyen Larsen, Rune Solberg
Date	December 2, 2008
Contract	JOP.12.08.2 (Norwegian Space Centre)

Norsk Regnesentral

Norsk Regnesentral (Norwegian Computing Center, NR) is a private, independent, non-profit foundation established in 1952. NR carries out contract research and development projects in the areas of information and communication technology and applied statistical modeling. The clients are a broad range of industrial, commercial and public service organizations in the national as well as the international market. Our scientific and technical capabilities are further developed in co-operation with The Research Council of Norway and key customers. The results of our projects may take the form of reports, software, prototypes, and short courses. A proof of the confidence and appreciation our clients have for us is given by the fact that most of our new contracts are signed with previous customers.

Norsk Romsenter

Norsk Romsenter (Norwegian Space Centre, NSC) is a government agency under the Ministry of Trade and Industry. NSC promotes the development, co-ordination and evaluation of national space activities as well as supports Norwegian interests in the European Space Agency (ESA). Earth observation involves all activities related to collection of information on the Earth's surface or atmosphere from instruments on board satellites. The Norwegian Space Centre's application programme supports users, research communities and businesses in testing the potential of Earth observation from satellites. Priority is given to the development of applications having public benefit.

Statens Vegvesen

Statens Vegvesen (The Norwegian Public Roads Administration (NPRA)) is responsible for the planning, construction and operation of the national and county road networks, vehicle inspection and requirements, driver training and licensing. It is also authorized to grant subsidies for ferry operations. The objective of the NPRA is to develop and maintain a safe, ecofriendly and efficient transport system. This is being done on a sound, professional basis by interacting with politicians, users and other interested parties.

Norsk Regnesentral
Norwegian Computing Center
Postboks 114, Blindern
NO-0314 Oslo, Norway

Besøksadresse
Office address
Gaustadalléen 23
NO-0373 Oslo, Norway

Telefon · telephone
(+47) 22 85 25 00
Telefaks · telefax
(+47) 22 69 76 60

Internett · internet
www.nr.no
E-post · e-mail
nr@nr.no

Title	**SatTrafikk Project results 2008**
Authors	**Siri Øyen Larsen, Rune Solberg**
Date	December 2
Year	2008
Publication number	SAMBA/50/08

Abstract

Information about traffic density (vehicle counts) is necessary for planning, construction, and maintenance of the road network. Very high resolution (VHR) satellite imagery may provide supplimentary information to the traditional ground-based sensors used for traffic monitoring. As a part of the SatTrafikk project, we present an approach for vehicle detection from QuickBird images based on image segmentation and pattern recognition methods. The proposed strategy is an attempt to solve some of the difficulties related to analysis under the conditions that are present at Norwegian roads. We also present the most recent experimental results, where the methods have been tested on images of Kristiansund and Sollihøgda acquired in July and August of 2008. The results are discussed in light of earlier experiments with images from the same areas. The major challenge seems to relate to segmentation, especially since many vehicles have poor contrast to the local background, while certain road marks have higher contrast in comparison. Finally, we discuss how we plan to approach some of the deficiencies of our algorithm in the future. The SatTrafikk project is funded by the Norwegian Space Centre and the Norwegian Public Roads Administration.

Keywords	Remote sensing, pattern recognition, vehicle detection, road traffic statistics, very high resolution satellite images, QuickBird
Target group	Road traffic authorities
Availability	Open
Project number	220 339
Research field	Earth Observation
Number of pages	4

Contents

1 Introduction

1.1 Background

Planning, construction, and maintenance of the road network are important tasks for several public entities, including the Norwegian Public Roads Administration (Statens Vegvesen Vegdirektoratet). The quality of this work relies heavily on the availability of updated traffic statistics. Today's primary source of traffic monitoring data is ground-based sensors in the road. These sensors count the number of vehicles that pass a given point over time, and also measure vehicle speed. The equipment is expensive to purchase and operate, and the methodology has evident shortcomings due to the very limited geographical coverage of the system. A number of statistical measures might be derived from these data. The most important information is the Annual Average Daily Traffic (AADT), which represents the traffic on an average day of the year.

Satellite imagery may provide a supplement to the traditional ground-based sensors. Vehicles are evident in very high resolution (VHR) satellite images, such as 0.6-m resolution Quickbird images. The main focus of the SatTrafikk project is to develop methodology for automatic detection of vehicles in satellite images of Norwegian roads. The approach is based on image segmentation and pattern recognition methods. It is an attempt to solve the difficulties related to analysis under the conditions present for Norwegian roads.

The spatial coverage offered by satellite images is much greater than can be achieved from ground-based sensors. A difficulty is that VHR satellite sensors do not allow frequent temporal acquisitions. Thus, we must ask ourselves whether it is possible to infer traffic statistics from satellite data with satisfactory accuracy. This question is also adressed by the SatTrafikk project, and documented in a separate report [1].

In the current report we will focus on the analysis of the satellite images and present vehicle detection results from the second year of the SatTrafikk project. The problem was previously adressed by NR in the project "Road Traffic Snapshot" in collaboration with the SVV and the Institute of Transport Economics (Transportøkonomisk institutt, TØI) in 2006-2007. The project was funded by the European Space Agency (ESA), and its main objective was to develop and test the necessary methodology for vehicle detection in VHR satellite imagery, and to discuss possibilities for future evolution. The Road Traffic Snapshot project gave positive outcome, and SVV and NR made an initiative to continue. The Norwegian Space Centre (NSC) supported the idea, and contributed – together with SVV – to the funding of the SatTrafikk project. The outcome of the SatTrafikk project during its first year (2007) was presented in [2].

1.2 About this report

There has not been any new methodological development concerning vehicle detection in this second year of the SatTrafikk project. However, we will give a brief description of the previously developed methods here. This is done in chapter 2. The main purpose of the work described in this report has been to further validate the methods on new data. In chapter 3 we describe the new image data. We also present and discuss the corresponding experimental results. Finally, in chapter 4, we summarize our work and discoveries, and suggest an approach for future work.

2 Methods

In this chapter we present a summary of the methods that have been developed for vehicle detection in VHR satellite images. The interested reader should refer to chapter 2 in [2] for details.

2.1 Preprocessing

The only preprocessing that is done to the images is foreground masking. During the main processing, we restrict our attention to the roads in the image. These roads are outlined in a road mask. The foreground mask is constructed using both a road mask and a vegetation mask. The purpose of the vegetation mask is to subtract from the foreground mask any parts of the image representing vegetation in the road.

2.2 Segmentation

The segmentation routine is applied to the masked panchromatic image, and finds segments that are either brighter or darker than the surroundings on the road. We apply Otsu's method for threshold selection [3], which groups pixels into classes depending on their graytone intensity value. The result will yield pixels with small intra-class intensity variation, and as large separation as possible between the classes. The Otsu threshold is found from the intensity histogram of the image.

We apply the method in two stages; to the lower and upper parts of the image histogram, where it finds thresholds for dark and bright segments, resepectively. On each stage, we use two overlapping, but slightly different, sections of the histogram – one yields a strict threshold, the other yields a more relaxed threshold. This approach is meant to make the procedure less dependent on one threshold. If the threshold is too strict, the segments will be poorly defined. If the threshold is too relaxed, many unwanted segments may be included. Therefore, the image is thresholded twice, and we combine the two results to construct the final segmentation.

2.3 Tree shadows

Norwegian roads are often narrow and often close to forest on one or both sides. A frequently encountered problem is that much of the road is hided by tree shadows. The problem is impaired at large view angles from the satellite. As a consequence, the segmentation produces many unwanted dark segments that represent tree shadows. We handle this problem making the simple assumption that no real vehicle segments should be located on the outer edge of the road. We use dilation to compute a road-edge mask from the road mask, such that the resulting road-edge mask is only one pixel wide. Any segment overlapping the road-edge mask is then regarded a non-vehicle segment, and discarded.

2.4 Feature extraction

The segmentation procedure finds not only vehicle segments, but also other segments that differ in graytone intensity from the local background. We want to extract segment features with power to discrimate between the different classes of objects. The features should describe both spatial and spectral characteristics of the segments. It is a difficult challenge to find the best possible combination of features. The selected features in our application are

- the intensity mean of the segment,
- the gradient mean of the segment,
- the standard deviation of the intensity within the segment,

- the length of the segment's bounding box,
- the 1st Hu moment of the segment, and
- the spatial spread of the segment.

These features are used for the main classification.

In addition, we perform an initial, rule-based classification in order to discard obvious non-vehicles before the main classification. The pre-classification is based on the area and elongation of the segment. It also discards some segments based on mean intensity and gradient values.

Furthermore, we compute a feature called *distance to nearest vehicle shadow*, which is used for post-classification. Based on information about the sun direction, we first calculate an image mask that represents the expected shadow zones of the vehicles found in the image. The expected shadow zones mask is then compared with a segmented image of dark objects that lie close to a bright object. Whereever there is overlap between these two images, the dark object is assumed to be a shadow. The result of this process is a shadow mask. This mask is used to estimate the distance to nearest shadow, which is a feature that is calculated for each object resulting from segmentation. Ideally, the shadow distance is zero for objects that are true vehicle shadows, positive for other dark objects, small for bright vehicles, and (relatively) large for road marks and other bright segments that may be confused with vehicles.

2.5 Classification

We define six object classes:

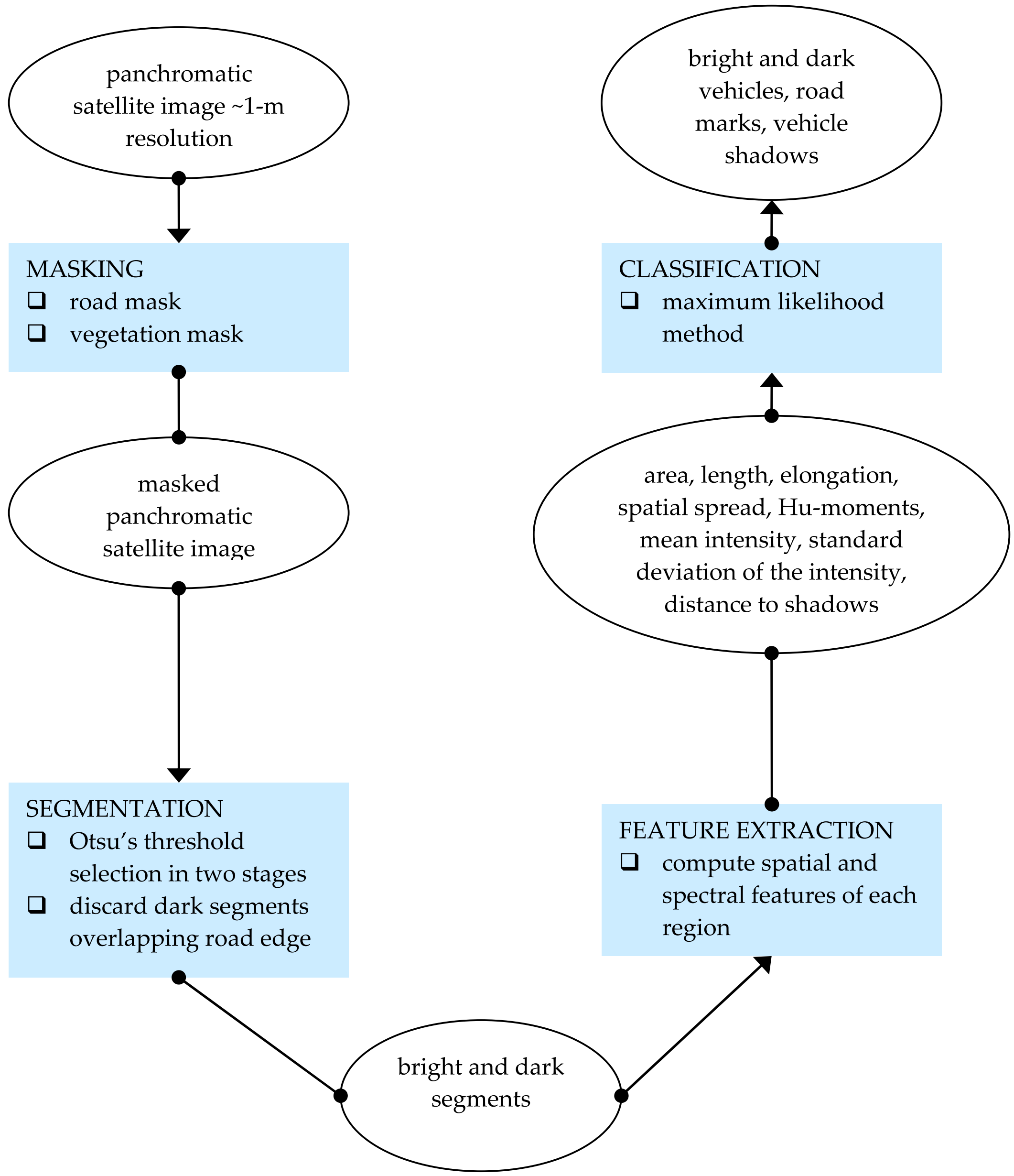
1. Bright car (light vehicle)
2. Dark car (light vehicle)
3. Bright truck (heavy vehicle)
4. Bright vehicle fragment
5. Vehicle shadow
6. Road mark - arrow

We assume that the feature vectors are normally distributed within each class, and the covariance matrices are arbitrary, i.e., we use a standard maximum-likelihood classifier with quadratic discriminant functions.

After main classification, we do a post-classification based on the distance to shadow feature, as mentioned above. More specifically, the class label of a dark car or vehicle fragment is changed to shadow if its shadow distance is zero. Furthermore, the classification of a road mark is changed to bright vehicle fragment if its distance to shadow is below a predefined (small) value.

2.6 System flow chart

The diagram below gives a methodological overview and describes the flow of data through the system.



3 Validation of methods on new data

3.1 Description of the data

The image data consist of two QuickBird scenes of 0.6-m resolution in the panchromatic band. Both acquisitions were ordered by the project and made in the summer of 2008. The locations, Sollihøgda and Kristiansund, were also used in a previous year, as reported in [2]. Table 1 gives an overview of the image data.

Location	Roads	Date	Time (UTC)	Upper left latitude	Mean sun elevation angle	Mean off nadir view angle
Sollihøgda	EV16	Aug 21 2008	10:48	60	41,7	10,0
Kristiansund	RV70	Jul 08 2008	10:57	63	49,1	23,2

Table 1 Image data.

Traffic counts from ground-based count sites are provided by the Norwegian Public Road Administration. We study two count sites[1] in each image. The traffic data contain the number of vehicles that pass a given count site each hour, as well as the average speed in each direction. We want to compare these counts to the number of vehicles found in the image. Therefore, at a given count site, we consider the longest strecth of road in the image that is unbroken by junctions to ensure that the same number of vehicles would pass any given point in the segment. Assuming there is free flow of traffic, we can then estimate the number of vehicles that may be expected to appear within the road segment in the image. Since the time of image acquisition (for both images) is late within the time interval from 10 to 11 hours UTC, we make an equivalent estimation for the subsequent hour interval as well. Information about the count sites is provided in Table 2, along with the expected number of vehicles in the image.

Location	Count site	Length of road segment (m)	Road dire-ction	10 - 11 UTC			11 - 12 UTC		
				Average speed (km/h)	Predicted number of vehicles in image		Average speed (km/h)	Predicted number of vehicles in image	
Sollihøgda	Sollihøgda	7819	1	67	40	73	66	41	72
			2	68	33		70	31	
	Rasteplass EV16	6139	1	58	36	66	60	35	61
			2	57	30		61	26	
Kristiansund	Atlanten	1055	1	47	14	29	48	14	32
			2	43	15		44	18	
	Rensvik	5775	1	66	18	37	66	17	38
			2	67	19		67	21	

Table 2 Count data.

[1] These are the same count sites that were studied in the 2002 and 2004 images of Sollihøgda and Kristiansund, respectively. In the 2007 report [2] these count sites were denoted "Kristiansund #1", "Kristiansund #2", "Sollihøgda #1", and "Sollihøgda #2". In this report, we will instead refer to the names of the actual locations, i.e., "Atlanten", "Rensvik", "Sollihøgda", and "Rasteplass EV16", respectively.

3.2 Validation experiments and results

The 2008 images were both delivered in UTM projection zone 32 North. As a first attempt, we use the road masks created for "old" image data, i.e., the 2004 Kristiansund image and the 2002 Sollihøgda image, as described in [2]. These images were delivered in UTM projection zone 33 North, thus we have to do a map conversion of the road masks. We use the built-in ENVI function to perform the conversion of projections. However, the resulting road masks do not match the image adequately. Therefore, the final road masks are manually constructed for the new image data.

We apply the vehicle detection algorithm to sub-images that represent the road segments around the count sites, as described above. We use the same class description database (training data) as described in [2], i.e., the training data is collected from the images Bodø (July 2003), Kristiansund (June 2004), Eiker (June 2002), and Sollihøgda (May 2002). The number of vehicles that are classified as vehicles are then compared to: 1) the predicted number of vehicles in the image from ground-based traffic data, and 2) the manually counted number of vehicles in the image. We have not considered the overall classification performance on these data, i.e., the classification of non-vehicle objects.

Each of the objects classified as vehicles are manually inspected and categorized as 1) a correct detection, 2) a false detection (not a vehicle), or 3) an extra wagon on a vehicle (trailer) that has already been registered as a correct detection. Furthermore, all vehicles omitted by the algorithm were registered during visual inspection.

Validation results are presented in Table 3. The last column represents the number of vehicles that were found by manual inspection of the image, and is the sum of correctly detected vehicles and omitted vehicles. Hence, it is here interesting to compare the numbers in the columns "Manual count" and "Correct detection" to the two columns "Predicted number of vehicles in image" (representing the ground-based number-of-vehicles estimate) in Table 2 above.

Location	Objects classified as vehicles				Omitted vehicles	Manual count
	Correct detection	False detection	Extra wagon	Sum		
Sollihøgda	42	6	2	50	35	77
Rasteplass EV16	40	13	0	53	23	63
Atlanten	37	4	0	41	9	46
Rensvik	21	5	0	26	18	39

Table 3 Validation results.

Count site	n_a	n_m	n_g	relative error between n_a and n_m (%)	relative error between n_g and n_m (%)
Sollihøgda	48	77	73	-37,66	-5,19
Rasteplass EV16	53	63	66	-15,87	4,76
Atlanten	41	46	29	-10,87	-36,96
Rensvik	26	39	37	-33,33	-5,13

Table 4. Relative errors between different vehicle counts. Here n_a denotes the number of vehicles found by the automatic algorithm, n_m the number of vehicles counted manually, and n_g the number of vehicles prediced based on ground-sensor data in the hour interval 10-11 UTC.

As can be seen in Table 4, the number of vehicles found by the algorithm is in general lower than for manual count. The relative error between the automatic vehicle count, n_a, and the manual vehicle count, n_m, is

$$\frac{n_a - n_m}{n_m} \cdot 100\% ,$$

thus, a negative relative error means that the automatic vehicle count is less than the manual. It should be noted that n_a here includes the false detections. The absolute relative error is between 11% and 38%. If we count only the *correctly detected* vehicles, the relative error is between negative 20% and 46%, or, correspondingly, the algorithm correctly detects between 55% and 80% of the manually counted cars (Table 5).

Location	Manual count	Correct detections	Correct detection rate (%)
Sollihøgda	77	42	54,55
Rasteplass EV16	63	40	63,49
Atlanten	46	37	80,43
Rensvik	39	21	53,85

Table 5. Comparison of manual vehicle count and the number of correctly detected vehicles. The correct detection rate is given in percent of manual count, which is considered to be the "true" number of vehicles in the image.

The manual count agrees fairly well with the ground-based measurements at three out of four locations. At Atlanten, the absolute relative error is 37%. A likely explanation to this is that the assumptions made when predicting n_g do not hold, i.e., that there is free traffic flow (no queues) and traffic density is uniformly distributed in the given time interval. Hence, we should rather trust the manual count when evaluating the performance of our algorithm.

In Table 6 we present the number of false detections and compare it to the total number of detections. We see that roughly 10-20% of the objects classified as vehicles are false detections, i.e., various types of non-vehicle objects.

Location	False detections	All detections	False detection rate (%)
Sollihøgda	6	50	12,00
Rasteplass EV16	13	53	24,53
Atlanten	4	41	9,76
Rensvik	5	26	19,23

Table 6. The number of falsely detected vehicles relative to the automatic vehicle count (all detections).

3.3 Discussion

The number of falsely detected vehicles is relatively low, compared to previous experiments.[1] Nevertheless, the false detections are not neglicible. The clearly dominant type of false detections is road marks (Figure 1). A few other objects, such as road bridges or signs, are also registered as detections.

The perhaps most serious problem is that a fair amount of vehicles are not detected at all. As noted above, the true detection rate was only 54-80% on the most recent experimental data.

Figure 2 shows examples of vehicles that were not detected. These vehicles typically have low contrast to the local background, or they are partially hidden by vegetation near the road. On some occasions, the algorithm correctly detects the vehicle shadow, without detecting the actual vehicle (Figure 3).

Figure 4 illustrates another recurrent problem; the sharp reflection of sunlight from vehicles. Such reflections may produce bright segments during segmentation. A dark segment in the shadow zone of a bright segment will be classified as vehicle shadow, hence, some dark vehicles are wrongly classified due to the sunlight reflections.

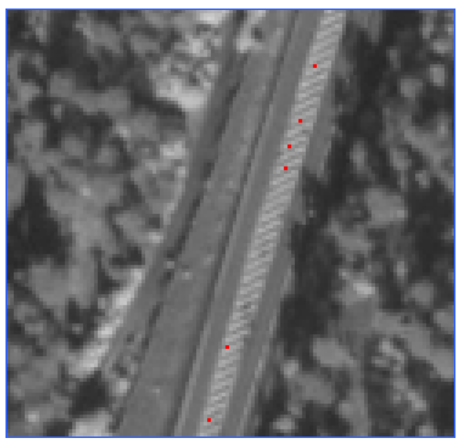


Figure 1. Road marks falsely detected as vehicles (indicated by red points).

[1] In the 2004 image of Sollihøgda, the detected number of vehicles at the "Rasteplass EV16" count site was actually overestimated compared with the manual count.

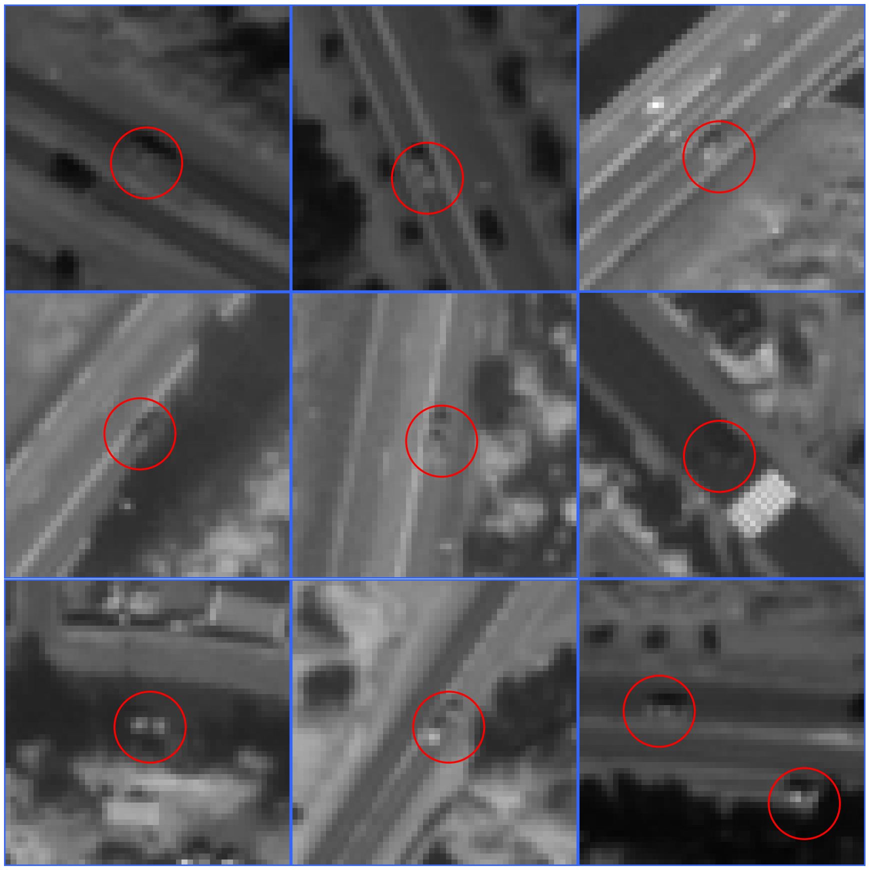


Figure 2. Examples of vehicles that are challenging to find to due low contrast, trees by the road, or nearby road marks. Each of the circled vehicles was found by manual inspection, and either not found or mis-classified by the automatic algorithm.

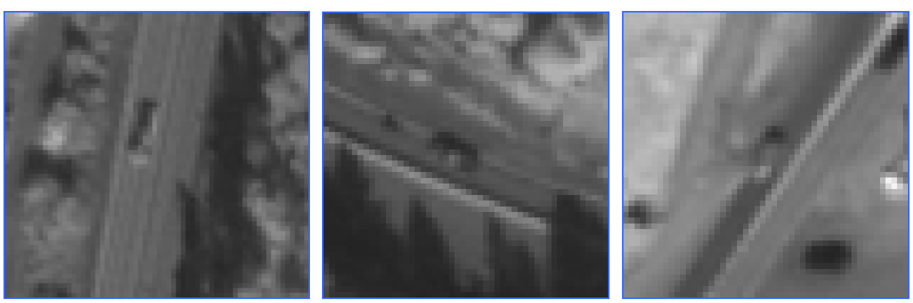


Figure 3. In the above images, the vehicle shadows were correctly classified, but the vehicles themselves were not found.

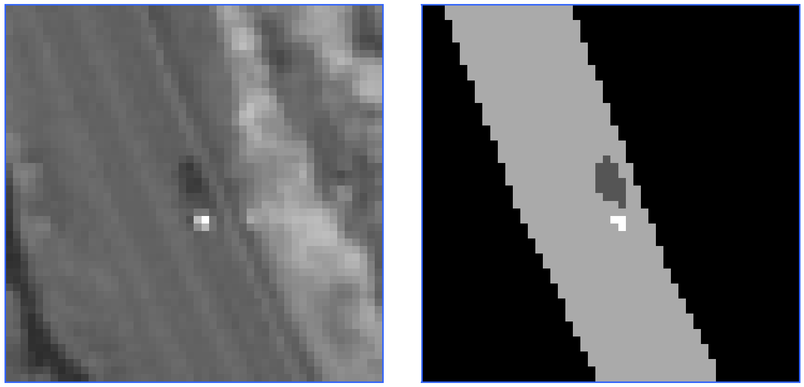


Figure 4. Panchromatic image (left) and segmentation result (right). According to manual interpretation, this is a dark car and a sharp reflection of sunlight. The algorithm classifies the dark segment as vehicle shadow, while the small bright segment is discarded during preclassification.

4 Conclusions

4.1 Recommendations for further work

We have seen many situations in which the presented methods fails to capture true vehicle objects in the image. Although some errors are made during classification, most of the difficulties lie in the segmentation. Vehicles with low contrast to the background, or vehicles composed of several parts with different colours, form the greatest challenge. In the future, we should improve the segmentation approach to be more robust than the current approach. The local segmentation conditions, i.e., both image lighting conditions and road surface conditions, change from place to place within a scene, but also within the subimages we use as input to the algorithm. We could try to adapt the thresholds for segmentation to even more local conditions than the whole subimage scene. Another idea is to apply a preprocessing step so that the contrast becomes more even over the entire image. One way to do this is to apply a contrast enhancement filter before segmentation. A concrete suggestion is to follow Alba-Flores' example in [4], where it is suggested to apply a filter that assigns to each pixel the maximum or minimum intensity value of a 3-by-3 pixels neighbourhood, before using Otsu's threshold selection for bright or dark objects, respectively.

We have shown that vehicle shadows can provide valuable contextual information. However, it can be discussed whether our model is too simple; it throws away the information about the objects that *cast* shadows, and only keeps the information about the actual shadows, i.e., a shadow mask. An example where this goes wrong is seen in Figure 4 above. An alternative approach is to model the vehicle and the expected shadow as one composited object. Since local radiometric disturbances often make vehicle detection more difficult, the algorithm may become more robust if structural and geometric features are emphasized. Hinz [5] use an explicit vehicle model to detect cars in high-resolution aerial imagery. Their model consists mainly of geometric features and also some radiometric properties. The car is modeled as a 3D object by a wire-frame representation, and an accurate computation of the car's shadow projection is derived from date, daytime, and image orientation parameters, and added to the model. The expected saliency of particular features, such as the edge between a car's hood and windshield, is also - depending on vehicle color, orientation, view point and sun direction – included in the model. Of course, this strategy is not suitable in our application, where much less details are visible in the image. However, we could make an attempt to generalize the idea.

The disadvantage of the detailed description is that a large number of models is needed to cover all types of vehicles. Hinz suggest using a tree-like model hierarchy for this.

In further work we should also seek to reduce the number of false vehicle detections. An improved segmentation approach may help by reducing the number of road marks that disturbs the segmentation result. At the same time, we will unlikely liberate ourselves of all the road marks during segmentation, as some road marks have even higher contrast to the background than certain vehicles. The road mark segments consitute a very heteregeneous class of objects, especially geometrically. Therefore, our classifier was trained on arrow-shaped road marks only. Since finding geometric features *common* to most road marks has proven to be very difficult, we should instead focus on finding features that *discriminate* road marks from vehicles.

4.2 Summary

We have presented an approach for vehicle detection in VHR satellite images of Norwegian roads. The algorithm was previously described in the first SatTrafikk report [2], and only briefly restated here. The main focus of this report has been to describe the latest validation experiments and discuss the results. We have also outlined ideas for future work in Section 4.1.

Validation was conducted on two QuickBird images; Kristiansund (July 2008) and Sollihøgda (August 2008). Both locations were also used in the previous validation experiments with image data from earlier years. New road masks were manually constructed. The classification was based on training data from earlier experiments.

The experimental results show that the proposed algorithm *underestimates* the number of vehicles in the images. Vehicles are omitted due to local radiometric disturbances, low local contrast, occlusion by trees, or other factors. These concerns are similar to what was found in previous validation experiments. Compared to experiments in the past, falsely detected vehicles were less influental. However, the amount of false detections is still too large to ignore.

Acknowledgements

The work presented in this report was funded by the Norwegian Public Roads Administration (Statens Vegvesen, Vegdirektoratet) and the Norwegian Space Centre (Norsk Romsenter).

References

[1] Larsen, S. Ø., M. Aldrin, and O. Haug (2008). Estimating Annual Average Daily Traffic (AADT) based on extremely sparse traffic counts – A study of the feasibility of using satellit data for AADT estimation, NR Note, Norwegian Computing Center.

[2] Larsen, S. Ø., H. Koren, and R. Solberg (2008). SatTrafikk – Vehicle Detection in Satellite Images for Development of Traffic Statistics, NR Note no SAMBA/18/08, Norwegian Computing Center.

[3] Otsu, N. (1979). A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man and Cybernetics, 9(1), pp. 62-66.

[4] Alba-Flores, R. (2005). Evaluation of the use of high-resolution satellite imagery in transportation applications, Final report, CTS 05-11, Department of Electrical and Computer Engineering, University of Minnesota Duluth, USA.

[5] Hinz, S. (2005). Detection of vehicles and vehicle queues for road monitoring using high resolution aerial images, Proceedings of 9th World Multiconference on Systemics, Cybernetics and Informatics, 10-13 July 2005, Orlando, Florida, unpaginated CD-ROM.