

Segmentation based traversing-agent approach for road width extraction from satellite images using volunteered geographic information

Extraction
from satellite
images

131

Received 3 May 2018
Revised 22 June 2018
Accepted 23 July 2018

Prajowal Manandhar and Prashanth Reddy Marpu

Department of Electrical and Computer Engineering,

Khalifa University of Science and Technology, Abu Dhabi, United Arab Emirates, and

Zeyar Aung

*Department of Computer Science, Khalifa University of Science and Technology,
Abu Dhabi, United Arab Emirates*

Abstract

We make use of the Volunteered Geographic Information (VGI) data to extract the total extent of the roads using remote sensing images. VGI data is often provided only as vector data represented by lines and not as full extent. Also, high geolocation accuracy is not guaranteed and it is common to observe misalignment with the target road segments by several pixels on the images. In this work, we use the prior information provided by the VGI and extract the full road extent even if there is significant mis-registration between the VGI and the image. The method consists of image segmentation and traversal of multiple agents along available VGI information. First, we perform image segmentation, and then we traverse through the fragmented road segments using autonomous agents to obtain a complete road map in a semi-automatic way once the seed-points are defined. The road center-line in the VGI guides the process and allows us to discover and extract the full extent of the road network based on the image data. The results demonstrate the validity and good performance of the proposed method for road extraction that reflects the actual road width despite the presence of disturbances such as shadows, cars and trees which shows the efficiency of the fusion of the VGI and satellite images.

Keywords Road extraction, Segmentation, Agent, GIS, VGI

Paper type Original Article

© Prajowal Manandhar, Prashanth Reddy Marpu and Zeyar Aung. Published in *Applied Computing and Informatics*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) license. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this license may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

Publishers note: The publisher wishes to inform readers that the article “Segmentation based traversing-agent approach for road width extraction from satellite images using volunteered geographic information” was originally published by the previous publisher of *Applied Computing and Informatics* and the pagination of this article has been subsequently changed. There has been no change to the content of the article. This change was necessary for the journal to transition from the previous publisher to the new one. The publisher sincerely apologises for any inconvenience caused. To access and cite this article, please use Manandhar, P., Reddy Marpu, P., Aung, Z. (2021), “Segmentation based traversing-agent approach for road width extraction from satellite images using volunteered geographic information”, *Applied Computing and Informatics*. Vol. 17 No. 1, pp. 131-152. The original publication date for this paper was 26/07/2018.



Applied Computing and
Informatics
Vol. 17 No. 1, 2021
pp. 131-152
Emerald Publishing Limited
e-ISSN: 2210-8327
p-ISSN: 2634-1964
DOI 10.1016/j.aci.2018.07.004

1. Introduction

The work on extraction of topographic objects like roads from satellite imagery started from late 1970s [43,37]. Yet, road extraction is a challenging research topic in the field of remote sensing especially with the advent of high spatial resolution satellite images. Since the past few decades, remote sensing has been an important data source for geographic information. Nowadays, with the availability of multiple earth observation satellite imagery for the same site/region, the problem of mapping registration between those available imagery and/or even between Volunteered Geographic Information (VGI) and the image arises. Thus, there is a need to represent these imagery data from multiple sources accurately.

Moreover, having an accurate and latest up-to-date road infrastructure database is important for many applications such as topographic mapping and map updating, along with disaster monitoring and safety analysis [65]. Practical urban applications such as automated road navigation [32], updating geographic information systems [7,38], and geometric correction of urban remote sensing images [3] require accurate and up-to-date road information. But the high complexity of remote sensing information of the road and other limitations such as lagging-behind of modern computer automation in catching up with the level of the corresponding demand, road extraction has still remain an active research topic in remote sensing.

The general characteristics of roads include relatively a slight width difference and changes in its direction, along with relatively uniform color but contrasting with its adjacent area being interlinked to form a road network [51]. Because of the variations of materials used in road and road widths that usually occur in an image and because of the presence of disturbance such as nearby buildings, trees and shadows cast by them, extraction of the actual shape of the road is an inherently difficult task. The presence of vehicles and lane marks on the road creates an arbitrary spatial and spectral texture that hinders the road extraction process. Furthermore, composition of objects which are similar to road pavement (such as roadside rooftops and parking areas) add unwanted noise to the extraction process. These additional disturbances make separation of the road from its surrounding objects even more difficult [12]. The information of road width is important in several applications. For example, a fully extracted road network can be used as one of the inputs to perform accurate geometric correction between images acquired at different times. As another example, in the context of urban micro-climate studies, it is very important to get the exact extent of the road as the thermal properties of the roads significantly affect the urban heat island effect. An accurate representation of the urban land cover which mainly consists of roads and buildings is very important for urban thermal modeling [1].

Previous works such as Jin and Davis [22,13,20,56] proposed the use of various modern techniques and methods to deal with the problems in road extraction. But these existing methods still could not satisfactorily address the problem of separability stated above. Because of the presence of many unavoidable disturbances in the road extraction process, a accurate extraction becomes difficult, especially when the relation between roads and other background objects (such as vehicles, trees, buildings, or shadows) is ignored. In fact, in aerial imagery these background context objects often tend to have powerful influences on the occurrence of roads in aerial imagery as tall objects tend to cast shadows on the road which may cause to be a disturbance. Therefore, additional data like context information (e.g., noise) is used in dealing with complex scenarios to steer the extraction process. For example, Strat [54] and Bordes, Giraudon, and Jamet [8] made use of neighboring objects to achieve better extraction.

With the introduction of geo-referenced digital road map, the accuracy and precision of road extraction can be improved significantly. Nowadays, road network is generally available in vector format with vertices and its corresponding edges connecting them. With the rise of the web technology, publicly available information such as Volunteered

Geographic Information (VGI) has undergone fast development [33]. VGI is a crowd-sourcing tool which allows members of the general public to create and contribute towards a global database that contains geo-referenced facts about the Earth's surface and its near-surface. An example of VGI is OpenStreetMap (OSM) [44], which provides spatial, geometrical, and attribute information about road networks. Most of the previous works use supervised learning approaches which require large amounts of processing time and storage space to train the system along with the demand for huge computation power. Thus, use of VGI helps to extract such road width to the full extent in a cost effective manner especially when the remote sensing data is growing day by day with even higher resolution.

1.1 Our contributions

The overview of our approach can be seen in the Figure 1. We use a segmentation approach where we make autonomous agents traverse through segments in a known road direction provided by VGI, and try to approximate the shape of the segment by its area to extract the road width accurately. As the width of the roads is mostly fixed, the road segments appear as rectangular objects, so the approximate width of the segment can be obtained by the corresponding area of the image segment and its corresponding length is obtained from VGI line.

In our proposed work, the segmentation process is performed prior to the traversal by the agent. Then, our autonomous agent performs a local operation to extract road width by

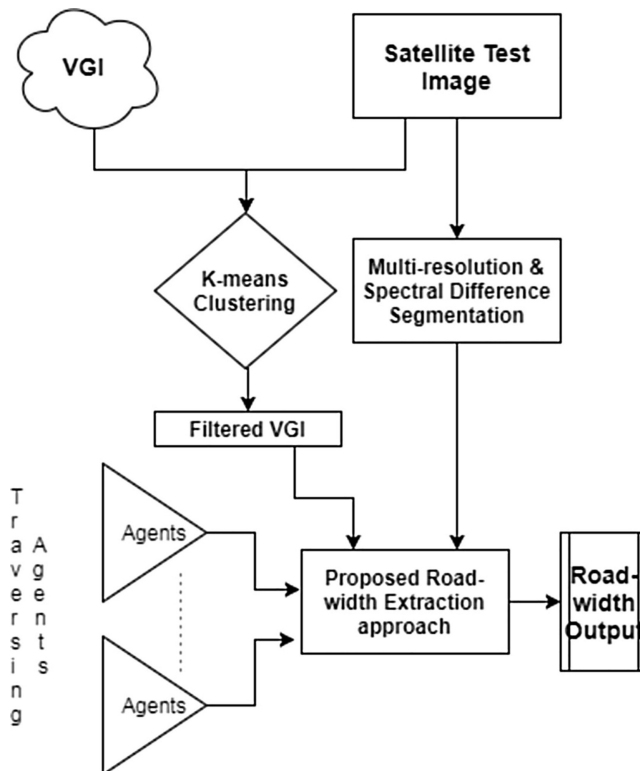


Figure 1.
Overview of our
approach.

getting inputs in the form of VGI and segmentation results of the satellite image in the direction guided by VGI. Apart from these inputs, agents tend to traverse with the knowledge of the width of previously traversed segments and also the spectral features i.e. color to verify the existing road. Also, if there is an issue with the image registration, it becomes very difficult to utilize the VGI, which is the reason why we used the agents approach, where we traverse the VGI and the images in parallel. Thus, the method is also adapted to account for problems associated with geographic mis-registration between the VGI and the image where VGI may not necessarily overlap with the available satellite images as shown in [Figures 2 and 10](#). Out of many possible mis-alignments between VGI and the satellite image, our proposed approach helps to deal with two major mis-alignment problems i.e translation and rotation. For the case of translation, if we know the starting position from the VGI, then we would only need to locate its neighbourhood point in the underlying image. Once we map the starting position of the VGI in the underlying image, we are able to traverse through the route in the direction of the VGI to extract the road network. While in case of rotation, we would need two reference points in order to determine the angle of rotation and then, we can proceed in the similar manner by mapping the starting position of VGI onto the underlying image, we are able to traverse to detect the road. The process also detects the cases where the road appears in the VGI and not in the satellite images, and thereby ignores such segments.

2. Related works

The fusion of digital image data has become essential in remote sensing for many applications such as topographic mapping and map updating, as well as disaster monitoring with the availability of multi-sensor, multi-temporal, multi-resolution and multi-frequency data from different Earth observation satellites [48].

Satellite images have been mostly used for the purpose of classification of land use/cover [49]. The presence of in-depth information content of the multi-spectral data, along with the inputs from optical sensors are being used for various classification purpose. Thus, most of the research work found uses the spectral aspect [2,58,9,63,13], while the temporal and spatial aspect has not been utilized much [49]. Also, image texture provides the spatial aspect of

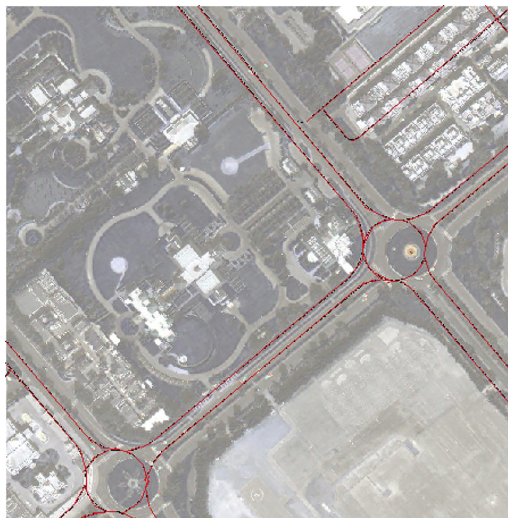


Figure 2.
The case of translation,
where satellite image
(Image 3) not
overlapping with VGI
information shown in
red dotted line.

remote sensing images, which can increase accuracy of classification based on information of local spatial structure and variation among land cover categories [17].

A number of works have been carried out in the area of road extraction from satellite images. Few examples include road width extraction, road centre-line extraction and road network detection. Road extraction can be categorized in two ways [2]. First, road extraction methodologies can be categorized based on either road-area extraction or road-centerline extraction. Road extraction largely depends on image classification and segmentation [41,46,58,31]. Road-centerline extraction methods relies on detecting road-skeletons [33,55,52,39,10]. Secondly, road extraction method can be automatic or semi-automatic. In semi-automatic, some prior information such as user inputs such as seed points or prior geographical information are required [33,11]; while in automatic approaches, no such prior information is required [42,36,59,21,41]. Also, additional inputs to the classifiers have proven to be important factors for increasing the classification accuracy [25,30,60,61]. Most of the existing approaches use supervised learning, which requires large amounts of training time and computational resources. Hence, there is a need for effective system to estimate the extent of the road width.

With the availability of the VGI, it has been used in number of applications including to map land-use patterns [23] and disaster management [19]. Even though the credibility of the VGI has been questioned [16] because it is voluntarily collected from public and it is freely available, it is usually reliable due to the engagement of a lot of volunteers, especially when it comes from sources or data warehouses such as OpenStreetMap [14,28].

Zhang [64] initially used road database along with various other data sources to develop automated system to verify whether the road exists or not. Baltsavias [6] also provided an overview about the use of knowledge and existing geodata to develop object identification system. The work by Liu et al. [33] integrates Mathematical Morphology (e.g., path opening/closing) approach by adapting VGI captured in the OSM database along with shape features (e.g., elongation and compactness) as prior knowledge to extract main road-networks from satellite images.

Yuan and Cheriyyadat [63] presented a method guided by OSM data that precisely segments road regions and uses a factorization-based segmentation method in order to localize boundaries for both texture and non-texture regions. Mattyus et al. [34] also combined OSM data with aerial images to estimate road width at global scale using a sophisticated random field probabilistic model.

[62] proposed an automatic method to align raster image to Google Maps by finding out local translation within each image tile. The process is then followed by a method of 'thin-plate-spline' warping which uses tile control point pairs along with confidence values. [53] proposed a method to define binary road mask by using spatial length-width contextual measure. This process was used in finding intersection and termination points. Thus, obtained matched point pairs were used to perform rubber-sheeting transform. Then modified snake algorithm is used to form the final binary image by taking into consideration intermediate vector road points.

Mnih and Hinton [41] used neural network trained with a labeled datasets which were derived from rasterized road map in order to detect road pixels. Their approach was able to detect roads with moderate occlusions but their method seemed to have failed when the occlusions are large. In other works, generally the road features obtained from the satellite image, such as edge and parallel boundaries, are not always prominent enough to be important features in all situations. Moreover, methods like the one proposed by Chen, Sun, and Vodacek [12] needed to perform separate map registration to align road vectors to its corresponding image road centerlines prior to extracting road segments. Recently, Kaiser et al. [26] used deep convolutional neural networks (CNN) on online maps to detect roads and buildings. Also, Landsiedel and Wollherr [29] used semantic knowledge from 3D metric

environment to infer street geometry information. However, our proposed approach is more effective than the above mentioned approaches as it is less resource demanding and we tend to determine the total road width and also, provide the way to deal with map registration problem through the traversal of agents.

3. Proposed method

This section covers the procedures used in our approach. Before applying our approach, we first need prior information about the road centre-line, which can be obtained from VGI.

3.1 Preprocessing

We use WorldView-2 satellite [15] images, which consist of 8 bands. In the first step, we extract the Normalized Difference Vegetation Index (NDVI) [18] feature to differentiate vegetation and also to improve the contrast of other objects in the satellite image. The NDVI feature is normalized to range of [0, 1000] to be able to use in segmentation process. In this work, we perform Principal Component Analysis (PCA) [24] on the 8 bands of the satellite data to reduce the dimensionality of the data. We choose the set of the first three components which account for 99% of the variance of the original satellite data. The PCA components are also normalized between 1 and 1000. The normalized PCA components are stacked with the normalized NDVI layer to form the basis of our segmentation process. The main idea is to improve the contrast between the objects and also reduce dimensionality to efficiently use the segmentation algorithms which are mostly based on Euclidean distance between objects in the spectral feature space.

3.2 Segmentation

Segmentation is a process of performing partition of an image space into non-overlapping homogeneous segments which has similar spectral features. We perform segmentation at two levels using eCognition software [57]. The first level uses multi-resolution segmentation and the second is based on spectral difference segmentation ([5]).

3.2.1 Multi-resolution segmentation. Multi-resolution segmentation [4] is based on scale, color and shape of an object for dealing specially with a very high resolution imagery, as it consists of very fine details. As objects of interest typically appear at different scales in an image in a very high resolution image, it allows the extraction of segments consisting of different size of the image objects. This process starts with each pixel which leads to form one image object or region. Then, a pair of image objects gets merged iteratively to form a larger object, where merging is performed based on the criteria of local homogeneity, which depends on similar spectral values of adjacent objects.

The criteria for creating image objects having relatively homogeneous pixels using the general segmentation function is

$$S_{\hat{m}} = w_{clr} \cdot h_{clr} + (1 - w_{clr}) \cdot h_{shp} \quad (1)$$

where w (user-defined weight) for spectral color (clr) versus shape is $0 \leq w_{clr} \leq 1$.

Spectral heterogeneity (h) of an object is defined as the sum of the standard deviations of spectral values of each band (σ_k) multiplied by the weights for each band (w_k):

$$h = \sum_{k=1}^m w_k \cdot \sigma_k \quad (2)$$

Difference between adjacent objects: When two image objects are similar and closer to each other in the feature space, the degree of fitting d_f for d-dimensional in this feature space is defined as:

$$d_f = \sqrt{\sum_d (f_{1d} - f_{2d})^2} \quad (3)$$

Here, the distance can be generalized in each dimension by its standard deviation over all segments of the feature:

$$d_f = \sqrt{\sum_d \left(\frac{f_{1d} - f_{2d}}{\sigma_{fd}} \right)^2} \quad (4)$$

The degree of fitting of two adjacent image objects can be defined by describing the change of heterogeneity h_{diff} in a virtual merge.

$$h_{diff} = h_m - \frac{h_1 + h_2}{2} \quad (5)$$

The generalized equation to an arbitrary number of channels c each consisting n different objects and each having weight w_c :

$$h_{diff} = \sum_c w_c (n_1 (h_{mc} - h_{1c}) + n_2 (h_{mc} - h_{2c})) \quad (6)$$

3.2.2 Spectral difference segmentation. On the results of multi-resolution segmentation, we perform spectral difference segmentation. This is a merging procedure, where neighbouring objects with a spectral mean below the defined threshold are merged to produce the final object segments. This reduces the total number of objects and also improves the results of multi-resolution segmentation. Figure 4 shows a sample segmentation output of a satellite image in Figure 6(a).

3.3 Road extraction

We consider the known road centre-line information (available from VGI) as our starting point to extract the road region in the satellite image. We utilize a traversing multi-agent approach, where multiple agents traverse the network and perform local operations simultaneously in multiple directions across the segmented images along the VGI. The agents traverse each segment utilizing the information of the segment such as segment length, width, and the angle of the line segment. We use a semi-automatic process where we manually need to specify the starting point on which our agent gets initiated both on the satellite image and the VGI. We then extract the whole segment in which the starting point lies. Once we extract the road segment, we obtain the end points of that segment after the thinning process [27]. We then traverse through each end point and try to see if there is an adjacent segment that can be connected in the direction of known road provided VGI meets the existence of road criteria. This criteria is based on mean color variation of the traversing segment such that threshold of the variation difference is defined at 1.5 of the previous segment. We choose this threshold based on experiments using sample images which were manually examined. Apart from this, we also account for exception cases of occlusions like shadow and trees, where NDVI was found to be less than 0 for shadows and more than 0.3 for trees based on visual observation on sample images. With this, we try to see if the pixel in neighbouring segment along the direction of known road can be treated as road segment or not. If it is a road segment then, at each end point, a sub-agent gets created and then it traverses through each detected road segment, thereby determining the entire road network. This process is described in Algorithm 1. An example is shown in Figure 3 (from (b) to (d), . . . , up until (e) and (f)).

We illustrate the demonstration of our approach in Figure 5. With reference to Figure 5 (i,ii,iii), we demonstrate how our agent traverses to segment 'Q' from 'P'. Let us consider that we have three segments X,Y & Z as shown in (i). But due to the noise, during segmentation we

Algorithm 1 Road_Extraction() algorithm.

Require Segmented image, existing road centre-line, starting point

Ensure Road segments

- 1: Make the agent traverse through point $P(i,j)$
- 2: Extract segment containing of point $P(i,j)$ in segmented image
- 3: Perform the thinning process in extracted segment
- 4: Find end points $E(i,j)$ of the VGI line segment corresponding to the traversed segment
- 5: **for** each end points **do**
- 6: Consider the neighbouring pixels $E(i+k,j+l)$ of $E(i,j)$, where $k,l = \pm 1, 2$
- 7: Among these neighbouring pixels, find only those pixels $E(i+k,j+l)$ such that these pixels belong to the different segment other than the current traversing segment
- 8: Validate the neighbouring segment as road either based on mean color deviation, where threshold is 1.5 times of previous segment or, by considering it as shadow ($NDVI < 0$) or trees ($NDVI > 0.3$)
- 9: **if** neighbouring segment is validated as existing road:
 then
- 10: Consider neighbouring pixel $E(i+k,j+l)$ as road if it is in direction of known existing road
- 11: Extract segment of neighbouring pixel $E(i+k,j+l)$
- 12: REPEAT the entire process for the obtained segment
- 13: Combine all extracted road segment

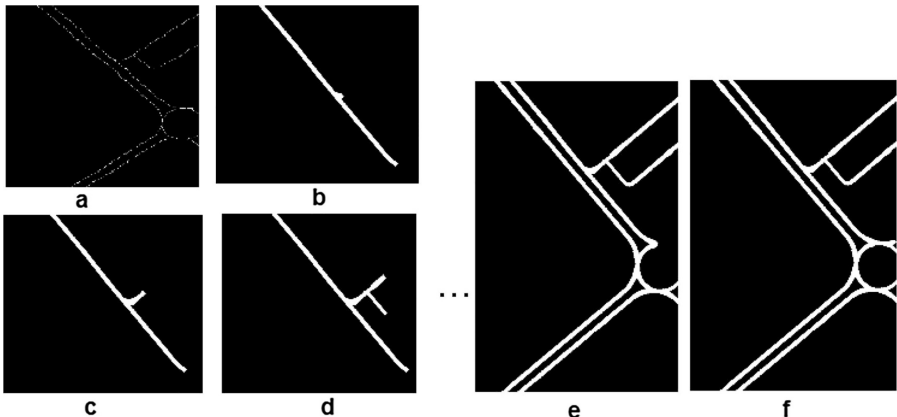


Figure 3. Illustration of our approach. Starting from the road centre-line information from VGI in subfigure (a), autonomous multiple sub-agents discover the road segments incrementally in subfigures (b), (c), (d), . . . , (e), and (f).

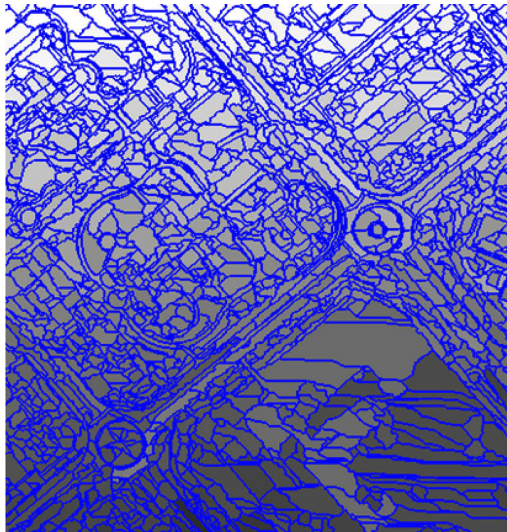


Figure 4.
Example of image
segmentation.

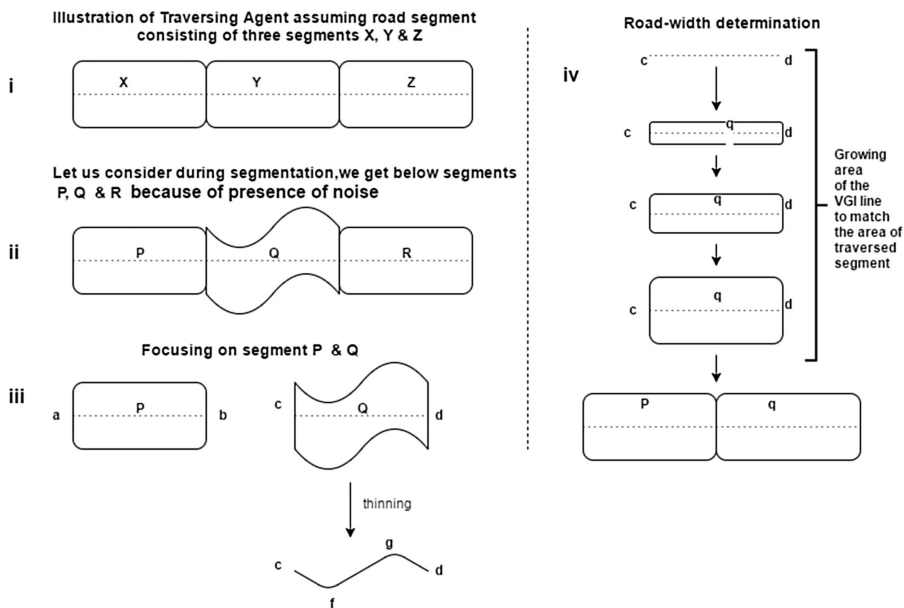


Figure 5.
Illustration of
Traversing Agent and
Road-width
Determination.

yield segments such as P, Q & R as shown in (ii). We also consider its corresponding VGI centre-line segment as shown by dotted line in the figure. Suppose segment 'P' has end points 'a' and 'b'; segment 'Q' has end points 'c' and 'd'. Then our agent starts with segment 'P' and traverses through points 'a' and 'b'. Now, the agent searches for the neighbouring pixels of 'b' belonging to the different segment, which happens to be 'c' as it belongs to different segment 'Q'. Then, agent extracts the segment 'Q', keep record of its segment area and performs the

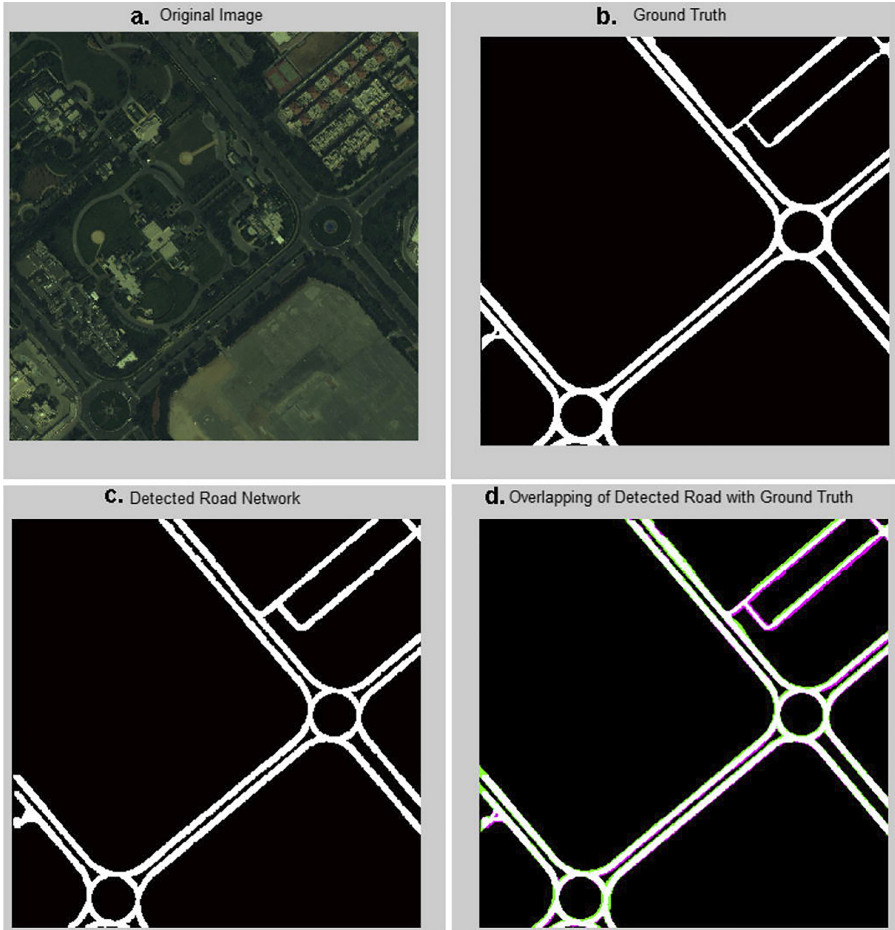


Figure 6. Results for Image 3. (a) Original RGB image, (b) Ground truth, (c) Detected road network by our method, and (d) Overlapping of detected road with ground truth. Pink indicates over detection (false positives) and green under detection (false negatives).

thinning process on segment 'Q' and determines end points ('c', 'f', 'g' & 'd'), which includes joint points ('f' & 'g') as well apart from end points ('c' & 'd'). Then we determine the angle of the segment by considering the median of the angles between two points lying in that particular segment. This allows us to approximate the angle of that segment. In [Figure 5 \(iv\)](#), we illustrate our road-width determination process. Then, we take its corresponding VGI centre-line and grow its area in the direction of the determined angle. The resultant segment outcome will be driven by the shape of VGI. Then, we start to grow the size of VGI centre-line in segment 'Q' such that the grown segment 'q', as shown in (iv) has an area pertaining to that of segment 'Q', provided it meets the criteria of the width-ratio; which we will describe in next section. Thus, our agent traverses to segment 'R' from segment 'Q' in the similar manner.

3.3.1 Determination of road width. The width of the road tends to be more or less constant over the region defined by the extracted road segment. The width tends to change from segment to segment, but consistency of the width is maintained somewhat in the region of the change where the segment gets changed. Thus, by getting the underlying segmented area directed by VGI, we try to approximate the width of that corresponding VGI segment with its

corresponding area obtained from segmentation. But the segmentation process may not always provide a perfect segment that overlays with existing roads, as there may be noises (such as vehicles, trees, street paints, etc.) which could prove to be hindrances in this process. In order to address this, we keep track of the width of the road by defining the width ratio based on the area of that particular road segment to the length of road within that segment.

For each detected segment, we start with growing the region of the known road centre-line such that it covers the maximum region of the road segment. Here, the centre road line of the VGI segment is grown in a rectangular way from its initial skeleton until the area of the rectangular region matches the area of the segment detected. We take this region area into consideration and define the width of the road. This process is illustrated in [Algorithm 2](#). Since the segmented area is vulnerable to the outliers (because of the presence of noises), we maintain the width of the road for each traversed segment and perform analysis based on its previous segment width before defining the current width of the road. We keep the ratio of width to length of the road constant, and this allows us to deal with the noises present in the direction of the road. For example, a few trees can be a part of a road and hence can lead to a greater extension of the segment when the areas of trees expand outside the road region. Thus, this could lead to defining more width of the road than the actual one. Hence, by keeping track of the ratio, we would be able to deal with the noises while defining the width of the road. In order to make this road-width ratio more consistent, apart from initial 5 traversed segments; we use each time median value of the determined ratio from the last 5 traversed segments. Thus, making our road-width ratio more robust to width changes.

Algorithm 2 RoadWidth_Determination() algorithm.

Require Segmented image, existing road centre-line

Ensure Width of the road corresponding to the extracted segment

- 1: Find the common region between the road centre-line and segmented region
 - 2: Perform the thinning process in extracted segment to find end points
 - 3: **if** Segment contains only two end points **then**
 - 4: Find the angle between the two end points of the extracted segment with reference to X-axis
 - 5: **else if** Segment contains more than two end points **then**
 - 6: Find the median angle between all the possible combination of the end points
 - 7: Find the average angle between the median angle of the current segment and angle of previously traversed segment
 - 8: Grow the obtained common region in the direction of the angle detected until the area of the region is equal to the area of the extracted segment, and provided it meets the width ratio
 - 9: Grown region provides the road-width for the corresponding segment
-

4. Results and discussion

4.1 Dataset

We have used two datasets in this experiment which are Abu Dhabi Dataset and Massachusetts Dataset.

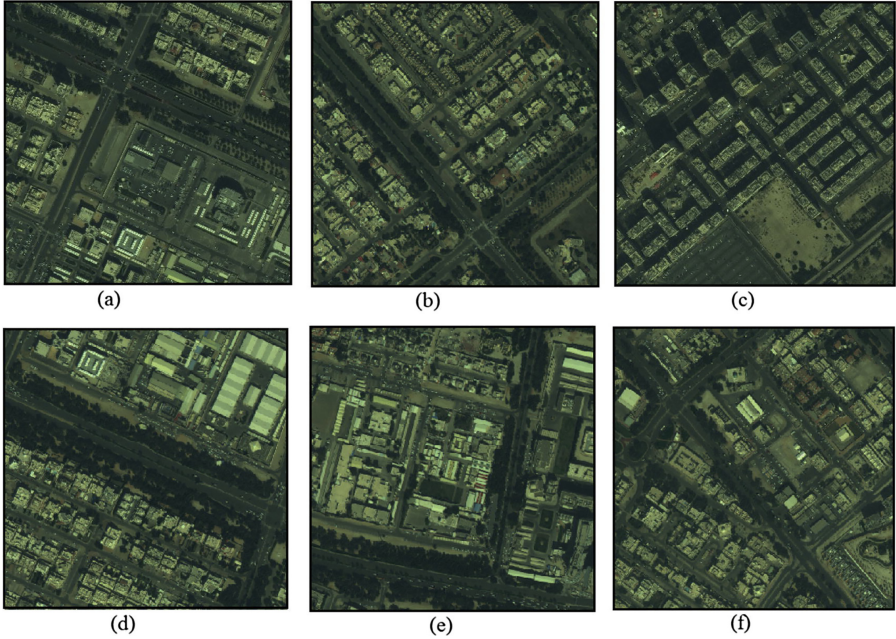


Figure 7. RGB images for (a) Image 1, (b) Image 2, (c) Image 4, (d) Image 5, (e) Image 6, (f) Image 7. Each test images are of size 1200 by 1200 pixels.

4.1.1 Abu Dhabi Dataset. The test data includes images acquired in 2014 over Abu Dhabi, United Arab Emirates by the WorldView-2 satellite with 0.46 meters ground sampling distance for panchromatic and 1.84 meters for multispectral images. All the images are composed of eight multispectral bands (Coastal, Blue, Green, Yellow, Red, Red Edge, NIR 1, and NIR 2) with a radiometric resolution of 11 bits per pixel. All the multi-spectral images are pan-sharpened with Gram-Schmidt approach [35] using Panchromatic images to create higher resolution images which helps to enhance the shape of the objects in color images. The VGI data is manually created by four volunteers by drawing vector lines overlaid on Google Earth [47]. This forms a nice case to validate our method on incorrectly registered data, as we observed significant misregistration between Google Earth images and the WorldView-2 data used in this study. The assessment of our proposed approach is performed over 7 test images (each image has a size of 1200×1200 pixels) which differ in the width of the road and the noises present on the road. The ground truth of the actual road network in each test image is manually traced by the authors in a very careful manner before applying our method to the image in order to avoid any bias.

4.1.2 Massachusetts dataset. This data set consist of 1171 aerial images of the Massachusetts state which was made public by Mnih [40]. Each image has a size of 1500×1500 pixels which covers an area of 2.25 square kilometers consisting of urban, suburban and rural areas. We have randomly selected over 50 contrasting images to perform comparison analysis. And the VGI data is developed in a similar way to Abu Dhabi dataset.

4.2 Detection performance

Pixel-based performance measures such as accuracy, precision, recall, and F1-score are determined to assess the quality of the results of the road extraction. A false positive (FP) is defined as an “over detection” where an actual non-road pixel is detected as a road pixel by

our method. A false negative (FN) is defined as an “under detection” where an actual road pixel is left out as a non-road one by our method. A true positive (TP) (true negative (TN) respectively) means that an actual road (non-road respectively) pixel is correctly identified as such.

$$\textit{Precision (Correctness)} = \frac{TP}{TP + FP} \tag{7}$$

$$\textit{Recall (Completeness)} = \frac{TP}{TP + FN} \tag{8}$$

$$\textit{Quality} = \frac{TP}{TP + FP + FN} \tag{9}$$

$$\textit{F1 - score} = \frac{2TP}{2TP + FP + FN} \tag{10}$$

We can visualize the results in [Figure 8](#) and [Table 1](#). The results demonstrate that our proposed framework is effective for the extraction of existing roads. We are able to achieve

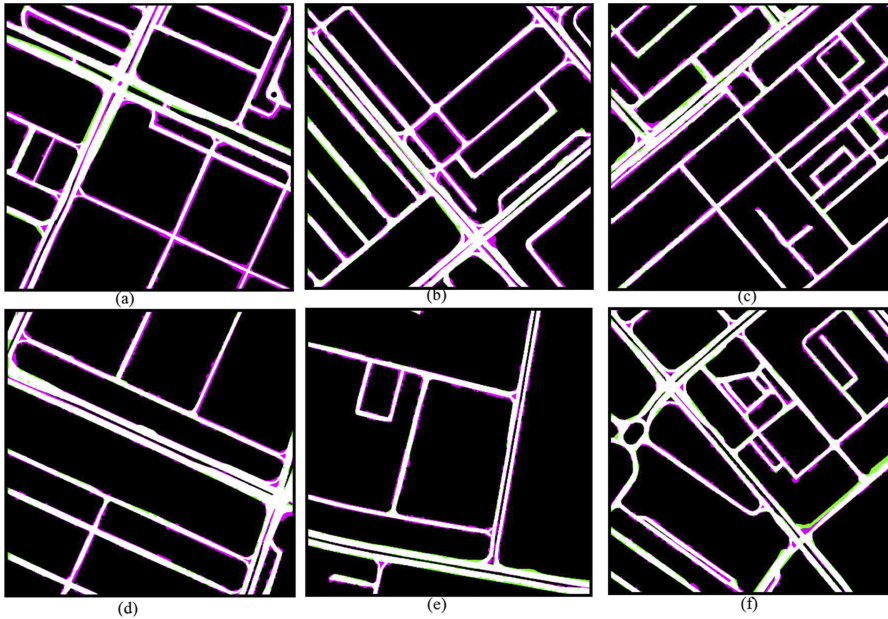


Figure 8. Superimpose of detected road with ground truth for (a) Image 1, (b) Image 2, (c) Image 4, (d) Image 5, (e) Image 6 and (f) Image 7. Pink indicates over detection (false positives) and green under detection (false negatives).

Measure	Img1	Img2	Img3	Img4	Img5	Img6	Img7
Precision	0.76	0.82	0.74	0.86	0.76	0.81	0.82
Recall	0.94	0.91	0.98	0.98	0.94	0.96	0.91
Quality	0.79	0.81	0.79	0.78	0.81	0.82	0.82
F1-score	0.82	0.86	0.83	0.80	0.84	0.88	0.86

Table 1. Performance evaluation in different images (for tile size 300×300).

average accuracy of above 93% in all the seven test images with the maximum being 96% in Image 6 and minimum being 88% in Image 2. The pixel-based overall computed mean precision is about 80% which shows correctness, mean recall about 94% which shows completeness of the road-width detection, and mean F1-score around 83% for the 7 test images.

In Figure 8, the white color portion of road shows the overlap of detected road regions with ground truth, pink color shows the overly detected road regions (false positives), and green shows the undetected road regions (false negatives). The results show that most of the times, we extract more width than the ground truth as indicated by pink color. Since we tend to approximate the areas of traversed segment, we tend to overestimate the width of the road resulting in the lower mean precision value. This is also because a clear distinction cannot be made between road and non-road regions while identifying the ground truth because of presence of disturbances such as shadows, trees, and parking vehicles, etc. Our approach sometimes tends to miss when the actual width of the road is much wider than usual. This is because we take into consideration the initial width traversed and then slowly build upon that width ratio of the road which helps to deal with the noises such as trees, vehicles and shadows. We also observe that by growing the region, we are able to connect the disjointed roads which are incurred because of the problem with disjoint center lines if any.

4.3 Computational complexity

From Table 1 and Figure 9, we can analyze the execution time of our proposed road extraction method. (We use a standard workstation using Microsoft Windows 7 with Intel Core i7-3770 with 3.40 GHz, 16 GB DDR3 RAM, and 1 TB SATA HDD with 7,200 RPM.) Since the proposed method can be easily parallelized, we experimented with tiling to handle large datasets. For each image, by increasing the tile size from 150×150 pixels to 300×300 pixels tile size, the execution time decreases. After that, the execution time tends to increase with the increase of the tile size. Thus, we can say that the execution time for the running of the road network extraction generally decreases with the decrease of tile size. However, if the tile size becomes too small, our method tends to divide the object segments into multiple tiles, and then the execution time is observed to have increased (as in case of 150×150 pixels tile size).

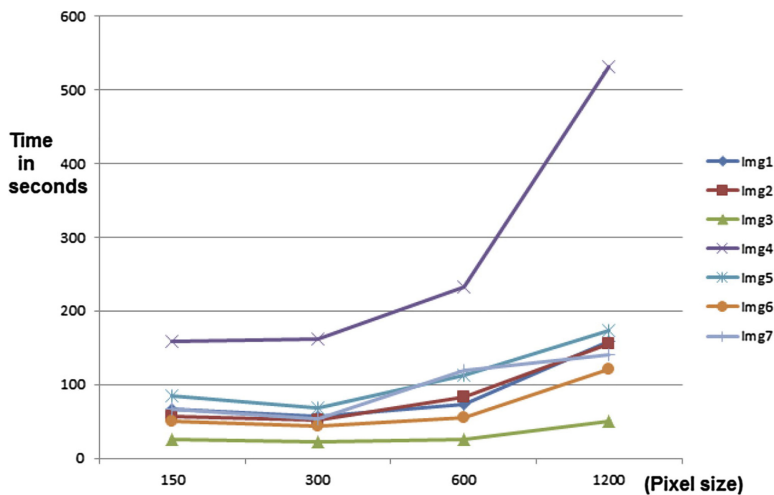


Figure 9. Comparison of execution time in seconds for different tile sizes (pixels).

So, we found 300×300 pixels tile size to be the most effective size for road extraction in a parallel scheme. Figure 9 shows that the image 4 (shown in Figure 7 (c)) has higher execution time in comparison to the other images. This is because the image 4 consists of the shadow which falls across the entire road width and beyond. Therefore, much finer segmentation of the image is performed. The traversal of one big segment takes lot lesser time than the traversal of smaller segments within the same big segment because of repetitive traversal function calls and other checks that need to be performed. Thus, the traversing agent took more time in image 4 than the other images.

In terms of the four performance metrics (accuracy, precision, recall, and F1-score), the change in tile size only affects them by a very small order of magnitude resulting in very minor changes like 0.01, which are statistically negligible.

4.4 Non-overlapping VGI

For the cases where the road information from VGI does not overlap with the underlying satellite image, by using the starting point and the direction of the existing roads, our approach has proven robust to traverse through the route in the direction of the known road to extract the complete road network with same performance measures as with overlapping case. A sample case is demonstrated in the image shown in Figure 2. Figure 10 shows case of rotation, where we manually rotated the image. With the help of two reference points, we determine the angle of rotation. Then, we use initial mapping reference seed in order to apply

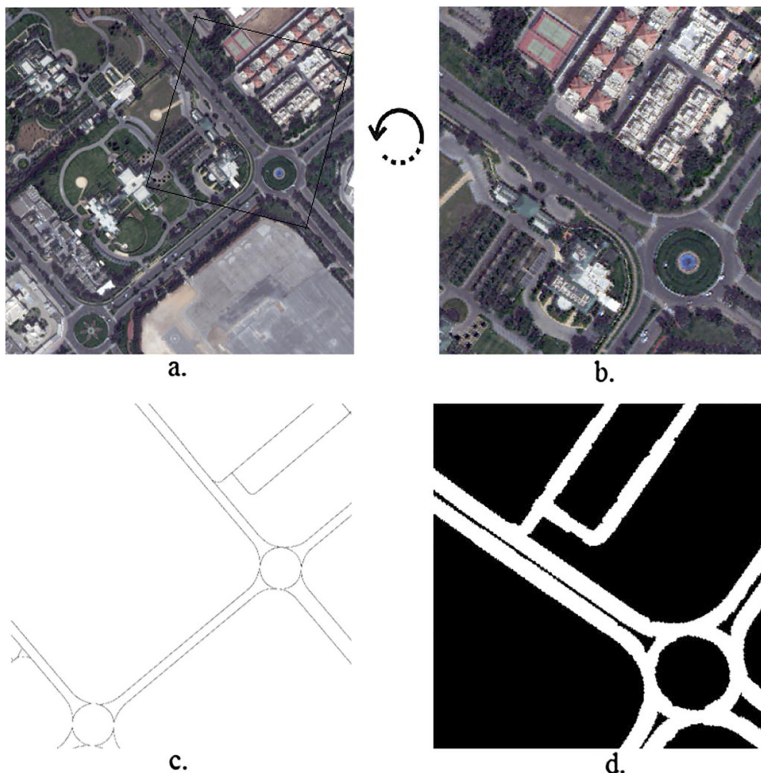


Figure 10.
A case of rotation: (a)
Test image in RGB, (b)
A part of test image,
which is rotated, (c)
Corresponding VGI
road centre-line for test
image, and (d) Detected
road for rotated image.

our algorithm in the direction of angle of rotation. The result (i.e., detected road) of the process is shown in [Figure 10\(d\)](#).

4.5 Comparison with existing methods

In ZY-3 dataset (as shown in [Figures 11 and 12](#)), our approach tends to provide smoother road detection with better handling of noises as compared to output obtained from the approach of [\[33\]](#); which also makes use of VGI. [Table 3](#) shows the comparison of our results with other different approaches mentioned in [Alshehhi and Marpu \[2\]](#), values listed are average values over 50 contrasting images taken from Massachusetts dataset. Also other approaches like Saito, Yamashita, and Aoki [\[50\]](#) and Panboonyuen et al. [\[45\]](#) performed experiments on Massachusetts dataset, but are not the exact images from this study. [\[50\]](#) claimed to obtain their best recall measure of 91.18% whereas we were able to obtain recall of 92.9%. Meanwhile Panboonyuen et al. [\[45\]](#) used a highly resource demanding approach of the deep CNN along with landscape metrics and conditional random fields to get completeness of 89.4%, correctness of 85.8% and F1-score of 87.6%. From the results as shown in [Figure 13](#), by using our approach, we can justify that the use of VGI information is able to aid in extracting the width of the road most of the times. Moreover, some algorithms which do not use VGI tends to detect even railway tracks as the roads [\[2\]](#). However, the failure to detect newly constructed road still remains the drawback of our method, but regularly updated VGI will help to deal even with these situations. [Figure 13](#) shows that our approach tend to overestimate the width which can be explained with the over-segmented regions lying underneath the VGI indicating road because of similar texture of neighbourhoods regions next to roads.

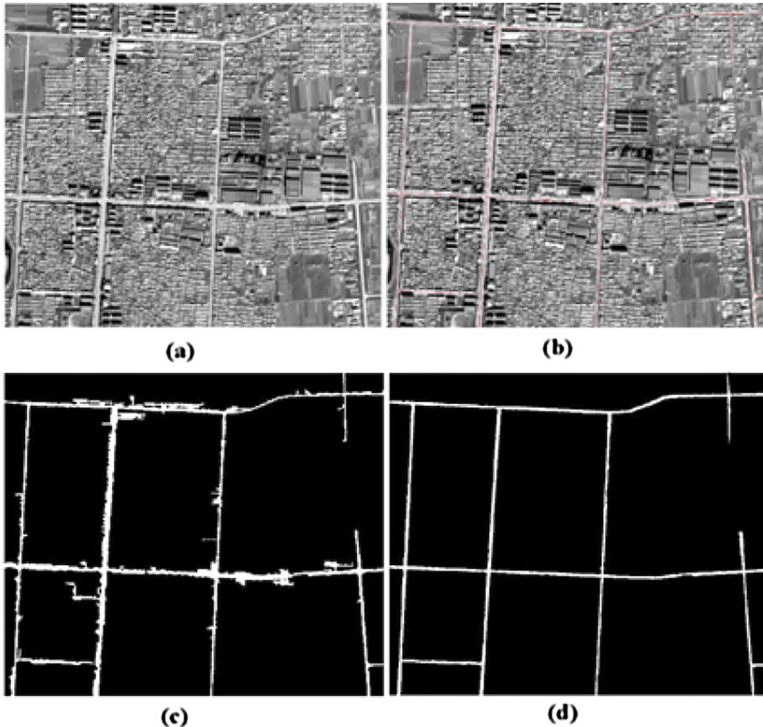


Figure 11.
Examples on ZY3 dataset: (a) Test image 8, (b) VGI (red line) on test image 1, (c) Output of [\[33\]](#) (d) Output of the proposed approach.

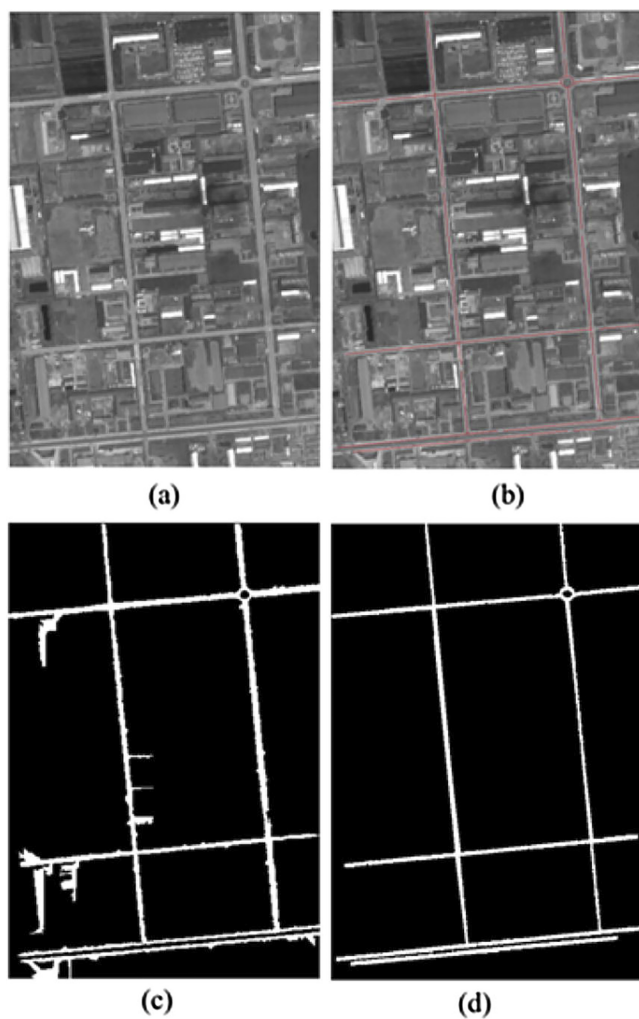


Figure 12.
Examples on ZY3
dataset: (a) Test image
9, (b) VGI (red line) on
test image 2, (c) Output
of [33] (d) Output of the
proposed approach.

Tile size	Img1	Img2	Img3	Img4	Img5	Img6	Img7
150×150	66	57	26	159	85	50	67
300×300	57	52	22	162	69	43	54
600×600	74	83	25	233	113	55	120
1200×1200	158	155	50	531	174	121	141

Table 2.
Comparison of running
time in seconds for
different images for
different tile sizes
(pixels).

5. Conclusion and future works

In this paper, the extraction of the total width of a road network from a satellite image is presented with the help of a traversing agents in the direction guided by VGI over the segmentation results. Assessments performed on 7 test instances of Abu Dhabi selected from WorldView-2 satellite images, 2 test images from ZY-3 dataset and over 50 images from Massachusetts dataset demonstrate that our approach provides the ability to extract smooth road network with generalization of the width of the road based on area of the segment traversed using the multi-agent-based road network extraction process. Our approach is able to deal with noises such as trees and shadows of the building casted on to the road, thus providing a smooth road network.

The results from this work form an important input to our future work to extract the new developments in the road network which allows us to update the global road database. For updating the global road network, which involves adding new road regions and amending

Table 3.

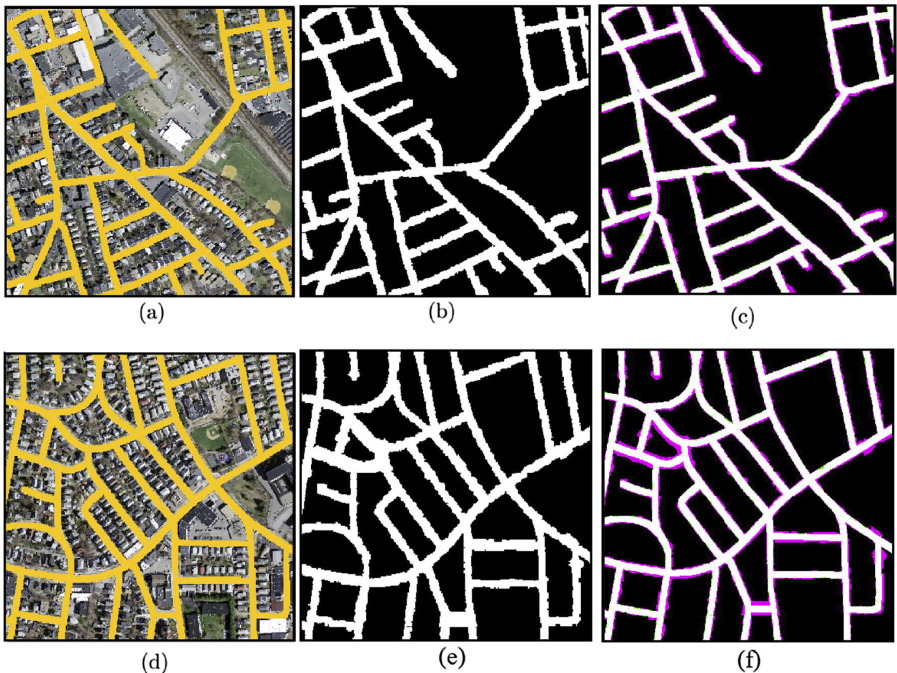
Comparison with other approaches.

(*Experiment might have been performed in different images of the same dataset, N/A means data was not reported and all values are reported in ‘%).

Methods	Completeness	Correctness	Quality	F1-score
Maurya et al. [36]	82.3 ± 4.7	70.5 ± 4.3	61.2 ± 6.1	N/A
Sujatha and Selvathi [55]	83.5 ± 4.3	76.6 ± 4.5	66.5 ± 6.4	N/A
Alshehhi and Marpu [2]	92.5 ± 3.2	91.0 ± 3.0	84.7 ± 5.4	N/A
**Panboonyuen et al. [45]	89.4	85.8	N/A	87.6
**Saito et al. [50]	92.9	90.1	N/A	N/A
Proposed-Method	93.4 ± 4.1	85.9 ± 2.9	82.2 ± 3.4	90.2 ± 3.1

Figure 13.

Examples on Massachusetts dataset: (a) Ground Truth superimposed over test image 10, (b) Results obtained by proposed method on test image 10, (c) Superimposition of detected road with ground truth for Massachusetts test image 10, (d) Ground truth superimposed over test image 11, (e) Results obtained by proposed method on test image 11, and (f) Superimposition of detected road with ground truth for Massachusetts test image 11 (pink color indicates over detection; green means under detection).



the regions of the existing ones, we plan to make use of a machine learning approach to extract those additional segments.

References

- [1] Afshin Afshari, A new model of urban cooling demand and heat island- application to vertical greenery systems (VGS), *Energy Build.* 157 (2017) 204–217.
- [2] Alshehhi Rasha, Prashanth Reddy Marpu, Hierarchical graph-based segmentation for extracting road networks from high-resolution satellite images, *ISPRS J. Photogramm. Remote Sens.* 126 (2017) 245–260.
- [3] Auclair Fortier Marie-Flavie, Djemel Ziou, Costas Armenakis, Shengrui Wang, Automated correction and updating of roads from aerial ortho-images, *Int Arch Photogramm Remote Sens XXXIII* (2000) 90–97.
- [4] Martin Baatz, Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation, *Angewandte Geographische Informations-Verarbeitung XII* (2000) 12–23.
- [5] Martin Baatz, Ursula Benz, Seyed Dehghani, M. Heynen, A. Höltje, P. Hofmann, I. Lingenfelder, et al., *eCognition user guide*, Definiens Imaging GmbH, 2001.
- [6] E.P. Baltasvias, Object extraction and revision by image analysis using existing geodata and knowledge: Current status and steps towards operational systems, *ISPRS J. Photogramm. Remote Sens.* 58 (2004) 129–151.
- [7] R. Bonnefon, Pierre Dhéréte, Jacky Desachy, Geographic information system updating using remote sensing images, *Pattern Recogn. Lett.* 23 (2002) 1073– 1083.
- [8] Ghislaine Bordes, Gerard Giraudon, Olivier Jamet, Road modeling based on a cartographic database for aerial image interpretation, in: *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps*, Birkhäuser Verlag, 1997, 123–139.
- [9] Lorenzo Bruzzone, Mingmin Chi, Mattia Marconcini, A novel transductive SVM for semisupervised classification of remote-sensing images, *IEEE Trans. Geosci. Remote Sens.* 44 (2006) 3363–3373.
- [10] Chuqing Cao, Ying Sun, Automatic road centerline extraction from imagery using road GPS data, *Remote Sens.* 6 (2014) 9014–9033.
- [11] D. Chaudhuri, N.K. Kushwaha, A. Samal, Semi-automated road detection from high resolution satellite images by directional morphological enhancement and segmentation techniques, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 5 (2012) 1538–1544.
- [12] Bin Chen, Weihua Sun, Anthony Vodacek, Improving image-based characterization of road junctions, widths, and connectivity by leveraging OpenStreetMap vector map, in: *Proceedings of the 2014 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2014, pp. 4958–4961.
- [13] Ching-Chien Chen, Craig A. Knoblock, Cyrus Shahabi, Automatically conflating road vector data with orthoimagery, *GeoInformatica* 10 (2006) 495–530.
- [14] Alexis Comber, Linda See, Steffen Fritz, Marijn Van der Velde, Christoph Perger, Giles Foody, Using control data to determine the reliability of volunteered geographic information about land cover, *Int. J. Appl. Earth Obs. Geoinf.* 23 (2013) 37–48.
- [15] DigitalGlobe, WorldView-2, 2017. <http://www.satimagingcorp.com/satellitesensors/worldview-2/>.
- [16] Andrew J. Flanagan, Miriam J. Metzger, The credibility of volunteered geographic information, *GeoJournal* 72 (2008) 137–148.
- [17] B. Ghimire, J. Rogan, J. Miller, Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic, *Remote Sens. Lett.* 1 (2010) 45–54.

-
- [18] Samuel N. Goward, Brian Markham, Dennis G. Dye, Wayne Dulaney, Jingli Yang, Normalized difference vegetation index measurements from the Advanced Very High Resolution Radiometer, *Remote Sens. Environ.* 35 (1991) 257–277.
- [19] Flávio Eduardo Aoki Horita, L'via Castro Degrossi, Luiz Fernando Gomes de Assis, Alexander Zipf, João Porto de Albuquerque, The use of volunteered geographic information (VGI) and crowdsourcing in disaster management: a systematic literature review, in: *Proceedings of the 19th Americas Conference on Information Systems*, 2013, pp. 1–10.
- [20] Jiuxiang Hu, Anshuman Razdan, John C. Femiani, Ming Cui, Peter Wonka, Road network extraction and intersection detection from aerial images by tracking road footprints, *IEEE Trans. Geosci. Remote Sens.* 45 (2007) 4144–4157.
- [21] Xin Huang, Liangpei Zhang, Road centreline extraction from high-resolution imagery based on multiscale structural features and support vector machines, *Int. J. Remote Sens.* 30 (2009) 1977–1987.
- [22] Xiaoying Jin, Curt H. Davis, An integrated system for automatic road mapping from high-resolution multi-spectral satellite imagery by information fusion, *Inf. Fusion* 6 (2005) 257–273.
- [23] Jokar Arsanjani, Marco Helbich Jamal, Mohamed Bakillah, Julian Hagenauer, Alexander Zipf, Toward mapping land-use patterns from volunteered geographic information, *Int. J. Geograph. Inf. Sci.* 27 (2013) 2264–2278.
- [24] Ian Jolliffe, *Principal Component Analysis*, John Wiley and Sons, 2002.
- [25] Martin Jung, Kathrin Henkel, Martin Herold, Galina Churkina, Exploiting synergies of global land cover products for carbon cycle modeling, *Remote Sens. Environ.* 101 (2006) 534–553.
- [26] Pascal Kaiser, Jan Dirk Wegner, Aurélien Lucchi, Martin Jaggi, Thomas Hofmann, Konrad Schindler, Learning aerial image segmentation from online maps, *IEEE Trans. Geosci. Remote Sens.* 55 (2017) 6054–6068.
- [27] Kamaljeet Kaur, Mukesh Sharma, A method for binary image thinning using gradient and watershed algorithm, *Int. J. Adv. Res. Comput. Sci. Software Eng.* 3 (2013) 287–290.
- [28] Thomas Koukoletsos, Mordechai Haklay, Claire Ellul, Assessing data completeness of VGI through an automated matching procedure for linear data, *Trans. GIS* 16 (2012) 477–498.
- [29] Christian Landsiedel, Dirk Wollherr, Road geometry estimation for urban semantic maps using open data, *Adv. Rob.* 31 (5) (2017) 282–290.
- [30] R.M. Lark, J.V. Stafford, Classification as a first step in the interpretation of temporal and spatial variation of crop yield, *Ann. Appl. Biol.* 130 (1997) 111–121.
- [31] Junyang Li, Lizuo Jin, Shumin Fei, Junyong Ma, Robust urban road image segmentation, In *Proceedings of the 2014 11th IEEE World Congress on Intelligent Control and Automation (WCICA)* (2014) 2923–2928.
- [32] Qingquan Li, Long Chen, Ming Li, Shih-Lung Shaw, Andreas Nuchter, A sensorfusion drivable-region and lane-detection system for autonomous vehicle navigation in challenging road scenarios, *IEEE Trans. Veh. Technol.* 63 (2014) 540–555.
- [33] Bo Liu, Wu. Huayi, Yandong Wang, Wenming Liu, Main road extraction from ZY-3 grayscale imagery based on directional mathematical morphology and VGI prior knowledge in urban areas, *PLoS One* 10 (2015) e0138071.
- [34] Gellert Mattyus, Shenlong Wang, Sanja Fidler, Raquel Urtasun, Enhancing road maps by parsing aerial images around the world, *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (2015) 1689–1697.
- [35] T. Maurer, How to pan-sharpen images using the Gram-Schmidt pan-sharpen method: a recipe, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* 1 (2013) W1.
- [36] Rohit Maurya, P.R. Gupta, Ajay Shankar Shukla, Road extraction using k-means clustering and morphological operations, in: *Proceedings of the 2011 IEEE International Conference on Image Information Processing (ICIIP)*, 2011, pp. 1–6.

-
- [37] David M. McKeown, Jerry L. Denlinger, Cooperative methods for road tracking in aerial imagery, In Proceedings of the 1988 Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) (1988) 662–672.
- [38] Juan B. Mena, State of the art on automatic road extraction for GIS update: a novel classification, *Pattern Recogn. Lett.* 24 (2003) 3037–3058.
- [39] Zelang Miao, Wenzhong Shi, Hu.a. Zhang, Xinxin Wang, Road centerline extraction from high-resolution imagery based on shape features and multivariate adaptive regression splines, *IEEE Geosci. Remote Sens. Lett.* 10 (2013) 583–587.
- [40] Volodymyr Mnih, *Machine Learning for Aerial Image Labeling*, University of Toronto, Canada, 2013 (Ph.D. diss.).
- [41] Volodymyr Mnih, Geoffrey E. Hinton, Learning to detect roads in high-resolution aerial images, *Proceedings of the 2010 European Conference on Computer Vision (ECCV)* (2010) 210–223.
- [42] Mehdi Mokhtarzade, M.J. Valadan Zoj, Road detection from high-resolution satellite images using artificial neural networks, *Int. J. Appl. Earth Obs. Geoinf.* 9 (2007) 32–40.
- [43] Makoto Nagao, Takashi Matsuyama, Yoshio Ikeda, Region extraction and shape analysis in aerial photographs, *Comput. Graph. Image Process.* 10 (1979) 195–223.
- [44] OpenStreetMap Volunteers, OpenStreetMap project, 2016. <https://www.openstreetmap.org/>.
- [45] Teerapong Panboonyuen, Kulsawasd Jitkajornwanich, Siam Lawawirojwong, Panu Srestasathiern, Peerapon Vateekul, Road segmentation of remotely sensed images using deep convolutional neural networks with landscape metrics and conditional random fields, *Remote Sens.* 9 (2017) 680.
- [46] Bo Peng, Lei Zhang, David Zhang, Automatic image segmentation by dynamic region merging, *IEEE Trans. Image Process.* 20 (2011) 3592–3605.
- [47] Muo Ping-hao, A method of making satellite photo map using Google Earth, *Electr. Power Survey Des.* 2 (2008) 30–31.
- [48] Cle Pohl, John L. Van Genderen, Review article multisensor image fusion in remote sensing: concepts, methods and applications, *Int. J. Remote Sens.* 19 (1998) 823–854.
- [49] John Rogan, DongMei Chen, Remote sensing technology for mapping and monitoring land-cover and land-use change, *Prog. Plann.* 61 (2004) 301–325.
- [50] Shunta Saito, Takayoshi Yamashita, Yoshimitsu Aoki, Multiple object extraction from aerial imagery with convolutional neural networks, *Electron. Imaging* 2016 (2016) 1–9.
- [51] Q. Sheng, Feida Zhu, S. Chen, Huifang Wang, Hanzhen Xiao, Automatic road extraction from remote sensing images based on fuzzy connectedness, in: *Proceedings of the 2013 5th International Conference on Geo-Information Technologies for Natural Disaster Management (GIT4NDM)*, 2013, pp. 143–146.
- [52] Wenzhong Shi, Zelang Miao, Johan Debayle, An integrated method for urban main-road centerline extraction from optical remotely sensed imagery, *IEEE Trans. Geosci. Remote Sens.* 52 (2014) 3359–3372.
- [53] Wenbo Song, James M. Keller, Timothy L. Haithcoat, Curt H. Davis, Automated geospatial conflation of vector road maps to high resolution imagery, *IEEE Trans. Image Process.* 18 (2009) 388–400.
- [54] Thomas M. Strat, Using context to control computer vision algorithms, in: *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Springer, 1995, pp. 3–12.
- [55] Chinnathevar Sujatha, Dharmar Selvathi, Connected component-based technique for automatic extraction of road centerline in high resolution satellite images, *EURASIP J. Image Video Process.* 2015 (2015) 8.
- [56] Weihua Sun, David W. Messinger, Knowledge-based automated road network extraction system using multispectral images, *Opt. Eng.* 52 (2013) 047203.

-
- [57] Trimble Navigation Ltd, eCognition Software, 2016. <http://www.ecognition.com/>.
- [58] Cem Unsalan, Beril Sirmacek, Road network detection using probabilistic and graph theoretical methods, *IEEE Trans. Geosci. Remote Sens.* 50 (2012) 4441–4453.
- [59] Jan D. Wegner, Javier A. Montoya-Zegarra, Konrad Schindler, A higher-order CRF model for road network extraction, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 1698–1705.
- [60] Peter T Wolter, David J. Mladenoff, George E. Host, Thomas R Crow, Improved forest classification in the Northern Lake States using multi-temporal Landsat imagery, *Photogramm. Eng. Remote Sens.* 61 (1995) 1129–1144.
- [61] C.E. Woodcock, S.A. Macomber, M. Pax-Lenney, W.B. Cohen, Large area monitoring of temperate forest change using Landsat data: generalization across sensors, time and space, *Remote Sens. Environ.* 78 (2001) 194–203.
- [62] Xiaqing Wu, Rodrigo Carceroni, Hui Fang, Steve Zelinka, Andrew Kirmse, Automatic alignment of large-scale aerial rasters to road-maps, in: *Proceedings of the 15th Annual ACM International Symposium on Advances in Geographic Information Systems (GIS)*, 2007, pp. 17:1–17:8.
- [63] Jiangye Yuan, Anil M. Cheryadat, Road segmentation in aerial images by exploiting road vector data, in: *Proceedings of the 2013 Fourth International Conference on Computing for Geospatial Research and Application (COMGeo)*, 2013, pp. 16–23.
- [64] Chunsun Zhang, Towards an operational system for automated updating of road databases by integration of imagery and geodata, *ISPRS J. Photogramm. Remote Sens.* 58 (2004) 166–186.
- [65] Jiaping Zhao, Suya You, Jing Huang, Rapid extraction and updating of road network from airborne LiDAR data, in: *Proceedings of the 2011 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, 2011, pp. 1–7.

Appendix A

Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.aci.2018.07.004>.

Corresponding author

Prajowal Manandhar can be contacted at: prajowal.manandhar@ku.ac.ae