THE APPLICATION OF NEURAL NETWORKS, IMAGE PROCESSING AND CAD-BASED ENVIRONMENTS FACILITIES IN AUTOMATIC ROAD EXTRACTION AND VECTORIZATION FROM HIGH RESOLUTION SATELLITE IMAGES

F. Farnood Ahmadi[a,] *, M. J. Valadan Zoej[a], H. Ebadi[a], M. Mokhtarzade[a]

[a] K.N.Toosi University of Technology, Faculty of Geodesy and Geomatics Engineering, Tehran, Iran - farshid_farnood@yahoo.com, -valadanzouj@kntu.ac.ir,- ebad)@kntu.ac.ir, m_mokhtarzade@yahoo.com

Commission III

KEY WORDS: Automatic, CAD, Image processing, High resolution images, Neural networks, Road extraction, Vectorization

ABSTRACT:

In this article a new procedure that was designed to extract road centerline from high resolution satellite images, is presented. The results (road Networks) are fully structured in vector formed in Computer Aided Design (CAD) based system that could be used in Geographical Information System (GIS) with minimum edit. The designed procedure is the combination of image processing algorithms and exploiting CAD-based facilities. In the first step, artificial neural networks are used to discriminate between road and non-road pixels. Then road centerlines are extracted using image processing algorithms such as morphological operators, and a road raster map is produced. Some cleaning algorithms were designed to reduce the existing noises and improve the obtained results. Finally, the edited raster map was vectorized using the CAD-based facilities. Obtained results showed that the structured vector based road centerlines are confirming when compared with road network in the reference map.

1. INTRODUCTION

Satellite and aerial images are the most important available data sources for map generation and updating of available maps. They provide accurate, easily accessible and reliable spatial information for Geographical Information Systems (GIS). The traditional manual methods of data capture from these images are expensive, laborious and time consuming and do not let full exploitation of available data in image archives. Nowadays when satellite images have highly improved in terms of spatial, spectral and temporal resolutions and Geomatics communities are overwhelmed by the sheer volume of collected images, the necessity of automation in feature extraction and map updating seems urgent.

Roads as one of the most important man made objects are in high concern to be extracted (semi)automatically and many researches have been carried out in this area. Geometrically constrained template matching (Gruen et al., 1995; Vosselman and Knecht, 1995), active contours or snakes (Neuenschewander et al., 1995; Trinder and Li, 1995; Gruen and Li, 1997) and fuzzy set and morphological operators (Mohammadzadeh et al., 2006) are some of the semi-automatic methods for road extraction.

Road extraction could be defined as the process of road identification and accurate localization in the image so that when the image to ground systems transformation is performed, the road network is truly presented in the object space. Automatic road extraction concentrates on automating all or some parts of this process to facilitate and expedite the road extraction task.

In high resolution satellite images, roads could be regarded as elongated homogeneous regions that contrast from background with distinct spectral behavior. Based on this model, automatic road extraction from this kind of images can be categorized in three steps as road detection, road thinning and centerline extraction and finally vectorization of extracted road skeleton.
Road detection is defined as the process of assigning a value to each pixel that can be used as a criterion to extinguish between road and background pixels. This process classifies the entire image into two different classes and has a major influence on the success of next stages. The segmented image, usually containing some unwanted and missed road pixels, is then introduced to some noise removal and other image processing algorithms to extract road centerline. Finally, the extracted road centerline is vectorized and transformed into CAD-based environments to be ready for GIS applications.

In this research, a back propagation neural network with its different input parameters is proposed for road detection, which is described in section 2. This is followed by morphological thinning and other image processing algorithms for road centerline extraction accompanied by noise reduction and quality improvement techniques, as well as automatic vectorization, which is outlined in section 3. Conclusions and recommendations for further studies are presented in section 4.

2. ROAD DETECTION USING ARTIFICIAL NEURAL NETWORKS

Neural Networks are computational systems made up of simple processing units called neurons which are usually organized into layers with fully or partially connections. The main task associated with a neuron is to receive the activation values from its neighbors (the output of other neurons), compute an output based on its weighted input parameters and send that output to its neighbours.

* Corresponding author.

Learning (training) a network is the process of adapting or modifying the connection weights between neurons so that the network can fulfil a specific task. Back Propagation is the most common learning algorithm which is an iterative gradient decent algorithm designed to minimize error function expressed in equation 1:

$$E = \frac{1}{2}\sum_{j=1}^{L}\left(d_j - o_j^M\right)^2 \quad \text{(Eq.1)}$$

Where d_j and o_j represent desired output and current response of the neuron "j" in the output layer respectively and "L" is the number of neurons in the output layer. In an iterative method, corrections to weight parameters are computed and added to previous values as below:

$$\begin{cases} \Delta w_{i,j} = -\eta \dfrac{\partial E}{\partial w_{i,j}} \\ \Delta w_{i,j}(t+1) = \Delta w_{i,j} + \alpha \Delta w_{i,j}(t) \end{cases} \quad \text{(Eq2)}$$

Where $w_{i,j}$ is weight parameter between neuron i and j, η is a positive constant that controls the amount of adjustment and is called learning rate, α is a momentum factor that can take values between 0 and 1 and "t" denotes the iteration number. The parameter α smoothes rapid changes between the weights. Learning rate and momentum parameter have a major influence on the success of learning process.

2.1 Network Structure

In order to use neural networks in road detection, input layer is consisted of neurons the same number as input parameters and output layer is made up of just one neuron that shows whether the input parameters can represent a road pixel or not. Usually, one hidden layer is sufficient, although the number of neurons in the hidden layer is often not readily determined (Richard, 1993).

In this research, a back propagation neural network with one hidden layer is implemented that uses 500 road and 500 background pixels as its training set.

An adaptive strategy is used to avoid trail and error learning rate and momentum assignment. In this method, both parameters are adjusted downwards as half after some training intervals if the overall training error has increased and upward 1.2 times if the overall error has decreased (Heerman and Khazeinie, 1992).

One of the most important factors in employing ANNs is to decide what type of information should be extracted from input image to be fed through the network as its input parameters. The discrimination ability of the network is highly affected by chosen input parameters. In continue different input vectors are evaluated to find the optimum network design.

2.2 Network Input Parameters Design

As a case study, a part of a pan-sharpened true colour (RGB) Ikonos image with the size of 550*550 pixels from Kish Island in Iran was chosen. Figure 1 shows the original image and its manually produced reference map which is used for accuracy assessment.

In the first try, the spectral information for each pixel, after normalizing the RGB values between 0 and 1, is simply entered to the network as its input parameters.

Thus, three neurons are designed in input layer in charge of receiving spectral values for each pixel in entire image. Figure 2 shows the network structure and obtained results are presented in Table 1.

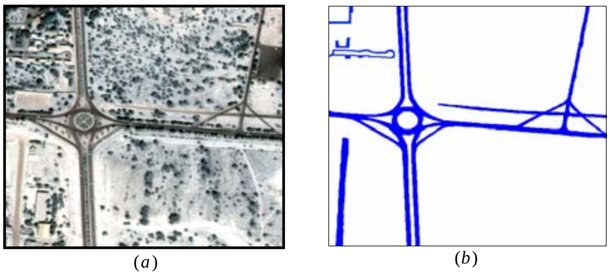


Figure 1; a) RGB Ikonos Image from Kish Island in Iran. b) Manually produced reference map.

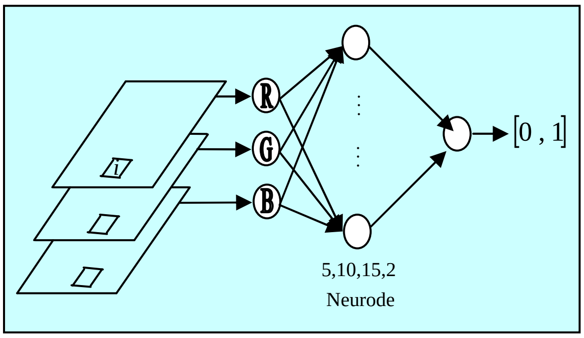


Figure 2: Network Structure when spectral values form input parameters.

Hi.N	Best Iteration	RCC	BCC	RMSE	Kappa Coeff.	Over all Acc.
5	15000	77.66	88.96	0.2459	67.92	93.81
10	5000	73.29	88.85	0.2240	68.34	94.66
15	5000	73.60	89.97	0.2204	69.46	94.75
20	10000	74.79	90.85	0.2235	69.30	94.47

Table 1: Three spectral values as input parameters

In this table, RCC and BCC are the road and background correctness coefficients that show the percentage of true functionality of the network about road and background pixels respectively. The RMSE value is computed by comparing of network responses in the output neuron and desired value from manually produced reference map (1 for road and 0 for background pixels). Kappa coefficient and overall accuracy parameters are obtained the same way as conventional classification methods where the network response about each pixel is classified into two classes as road and background pixels using an appropriate threshold value.

In order to improve network's functionality and considering the fact that roads are presented as homogeneous areas in high resolution images, neighboring pixels in a 3*3 window were participated in input vector generation. Furthermore, the normalized distances of all 9 pixels in the mentioned window to the mean vector of road pixels are added to input parameters. It was intended to use more spectral information to the network in order to increase the discrimination ability of the network between road and background pixels and also to reduce the iteration times in learning stage. The proposed distance parameter witch is calculated for each pixel as below:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix}_i , \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{mean}$$

$$d_i = \frac{1}{441.673}\left[(R_m - R_i)^2 + (G_m - G_i)^2 + (B_m - B_i)^2\right]^{1/2} \quad \text{(Eq. 3)}$$

While, $[R_m \quad G_m \quad B_m]_i^T$ is the mean of road pixels in training set. The maximum distance in the spectral space, is 441.673.

Thus the input layer is consisted of 9 red, 9 green, 9 blue and finally 9 normalized distances to the road mean vector which means 36 neurons are designed in this layer. Figure 3 shows network's structure and Table 2 presents obtained results.

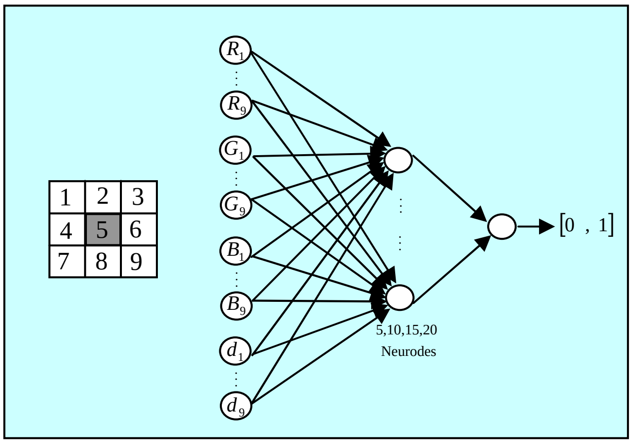


Figure 3: Network structure when neighbor pixels with their normalized distances form input parameters

Hi.N	Best Iteration	RCC	BCC	RMSE	Kappa Coeff.	Overall Acc.
5	1000	75.60	93.85	0.2011	71.31	95.11
10	1000	75.20	94.60	0.2002	71.60	95.15
15	1500	75.51	95.57	0.2014	71.86	95.15
20	1500	76.60	95.24	0.2008	72.18	95.17

Table 2: Spatial information and normalized distance as input parameters

Comparison between Table 1 and 2 shows that designed input parameters could improve network's ability in both road and background detection. Although input layer enlargement can make training stage more time consuming, but this problem is compensated to some extent by decrease in requested hidden layer size and iteration time.

Finally, as a comparison with statistical methods, obtained results from Maximum-Likelihood classification method and best network from Tables 1 and 2 are shown together in Figure 4 with their accuracy assessment parameters. A detailed explanation of the road detection strategy used in this research can be found in (Mokhtarzade and Valadan Zoej, 2007).

3. AUTOMATIC ROAD VECTORIZATION

In the image shown in Figure 4, which is obtained from improved BNN, all pixels on the road surface were detected. But what we need for GIS is the vectorized road centerline that should be extracted from road raster map.

An algorithm was devised to extract road centerline and vectorize it so that it is ready to be entered into GIS as a road vector layer as explained in the following sections.

3.1 Road Thinning and Centerline Extraction

In order to find road centerline from the available road raster map, thinning algorithms could be used to find road skeleton, which is a one pixel width road representation. Morphologic operators and an iterative boundary erosion algorithm were implemented to find the road centerline and omit other road pixels from the binary image. The obtained result is shown in Figure 5.
It can be seen that the obtained image shows the main skeleton of the roads, but there are two kinds of unwanted pixels that should be removed before vectorization:
Single elements distributed all over the image which are not in the road class
Spur pixels in the road network that are caused by thinning algorithm.
In the following sections the strategies for removing these pixels are explained.

3.1.1 Removing Single Elements

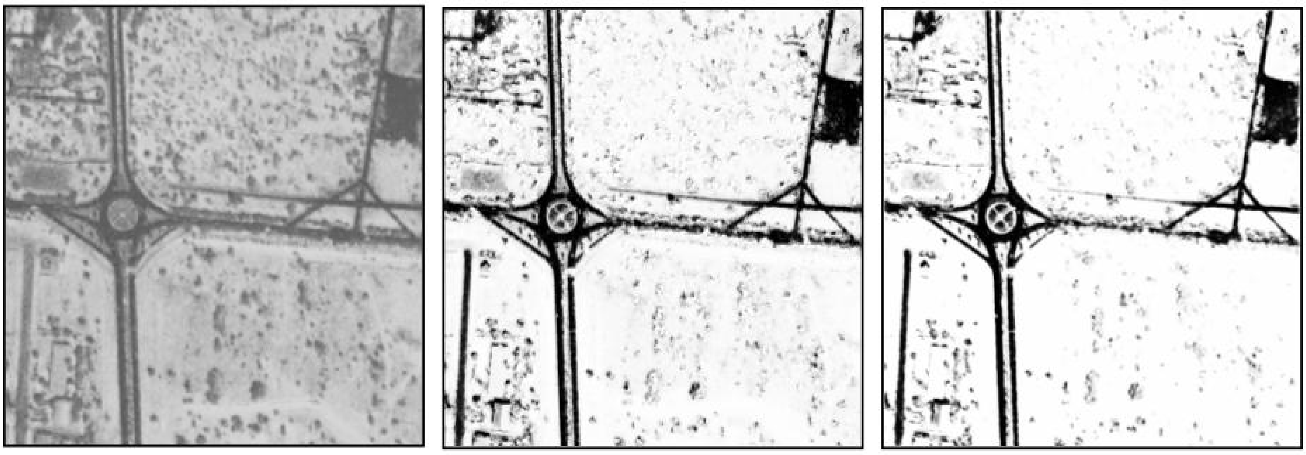
Single elements that do not belong to road class are distributed all over the image as small sets of isolated pixels. The following algorithm has been used to remove them. At fist a run length encoding algorithm is implemented (Haralick et al., 1992) and then contiguous runs are combined to form a single run. Finally, consistent runs are encoded as separate objects. Those objects that are comprised from less than a predetermined pixel number are regarded as unwanted isolated pixels and are removed in the next step.

It should be noted that the success of this method is directly influenced by the defined threshold that is a problem dependent value. In this research, the optimum threshold value was found in an trial-and-error as 20 pixels. The obtained result is shown in Figure 6.

3.1.2 Removing Spur Pixels

The term "Spur Pixels" is assigned to non-road small linear elements that are connected to the main road skeleton at one end and the other ends is free. The length of these elements is negligible in comparison with the main road network. They are not omitted in the previous section as there is a link between them and main road pixels. In order to remove them, the following algorithm which can be classified as morphological operators is suggested.

For each road pixel, neighboring road pixels are counted in a 3*3 window. If the number of road pixels is equal or less than one, the center pixel is omitted from the road raster map, otherwise it is kept unchanged. The main problem with this algorithm is that, in each run one pixel from the ends of main road skeleton is also removed. Figure 7 depicts the obtained result after successive iterations when all spur elements are removed.



Maximum Likelihood Method
RCC: 55.91
BCC: 70.12
RMSE: 0.3349
Kappa Coefficient: 63.10
Overall Acc.: 91.20

Hidden Neurons: 10
Iteration: 5000
RCC: 73.29
BCC: 88.85
RMSE: 0.2240
Kappa Coefficient: 68.34
Overall Acc.: 94.63

Hidden Neurons: 15
Iteration: 1500
RCC: 75.51
BCC: 95.57
RMSE: 0.2014
Kappa Coefficient: 72.18
Overall Acc.: 95.15

Figure 4: Left image: Maximum-Likelihood result, Middle image: Simple BNN with 10 neurons in hidden layer from Table1, Right image: Improved BNN with 15 neurons in hidden layer from Table 2

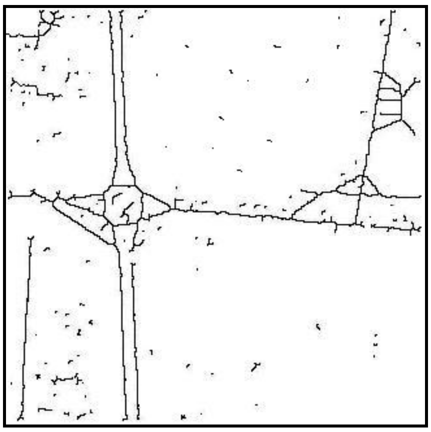


Figure 5: Road skeleton after morphological thinning operator

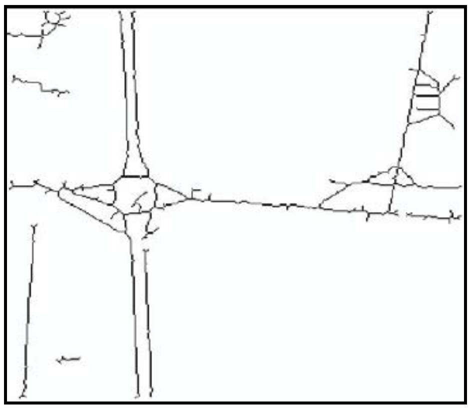


Figure 6: The result of removing single elements

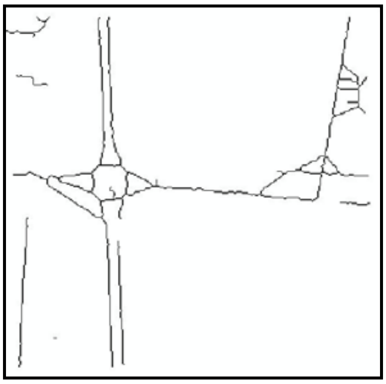


Figure 7: The result of removing spur pixels

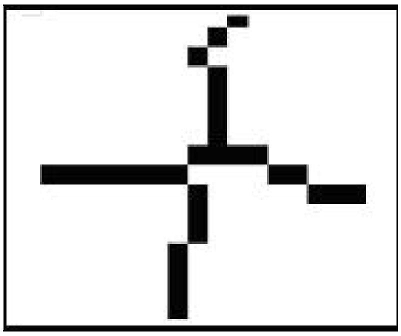


Figure 8: Cross pixel structure

3.2 Road Segmentation

In automatic road extraction, each road section should be defined separately in a way that belonging pixels are grouped together representing a unique section. This task can be fulfilled in two stages:

- Isolation of different road sections
- Grouping road pixels into their relevant road sections

3.2.1 Isolation of Different Road Sections

Road section isolation can be performed in cross pixels where two or more road sections intersect. Considering the fact that road thinning has reduced the road width into one pixel, those pixels with more than three 8-neighbouring road pixels could be regarded as cross pixels (Figure 8).

When cross pixels are found, they are removed to separate road sections into isolated segments.

3.2.2 Grouping Pixels into Their Relevant Road Sections

After road network separation, it would be possible to group pixels into their relevant pixels. It could be performed by sweeping the whole image and using run length encoding scheme. In the second sweep, neighboring runs are connected to each other to form a unique run and a run code is assigned. The strategy in run code assignment is to use successive positive integer numbers for road runs where non-road pixels are represented by zero values. . When neighboring runs are connected, coding is modified so that the maximum code is equal to the number of available isolated segments. Figure 9 represent the coded run segments in different colors.

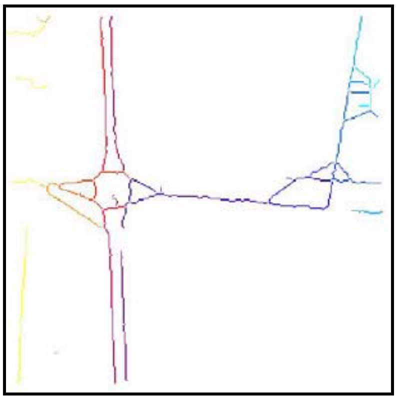


Figure 9: Isolated run codes presented in different colors

3.3 Automatic Vectorization Using CAD Environment Facilities

In this research, vectorization of segmented road pixels was performed with the aid of MicroStation^{TM} software.

In general, vectorization process of segmented road sections in this software can be expressed as:

- Coordinate assignment to all pixels belonging to each road segment.
- Coordinate transformation to MicroStation^{TM} software (to establish a relationship between road raster map and MicroStation^{TM} environment).
- Connecting successive points so that the segmented road would be represented as a multi line.
-

Since road pixels belonging to each segment were categorized and coded in the previous section, each segment could be transformed independently into the MicroStation^{TM}. This task was performed by a designed software port to establish the relationship between programming and MicroStation^{TM} environments. This port provides the system with all the facilities available in the MicroStation^{TM} for vector-based data editing and data control processes.

In digitizing a sample road segment, the first step is to find its beginning point. For this reason, the road raster map is swept top-down and left-right to find the first pixel belonging to the interest road segment. Then the pixels located in its direct 8-neighbourhood are followed in both directions until the starting point, having only one 8-neighbour, is found. Line following procedure begins from this point and road pixels are transformed into the MicroStation as^{TM} separated points. Consecutive points are connected to each other to form a structured line string for each road segment. This vectorized line string contains all the geometric information of the represented road segment. This procedure is repeated for all road segments available in the road raster map. Figure 10 shows the vectorized road segments in the MicroStation^{TM} environment.

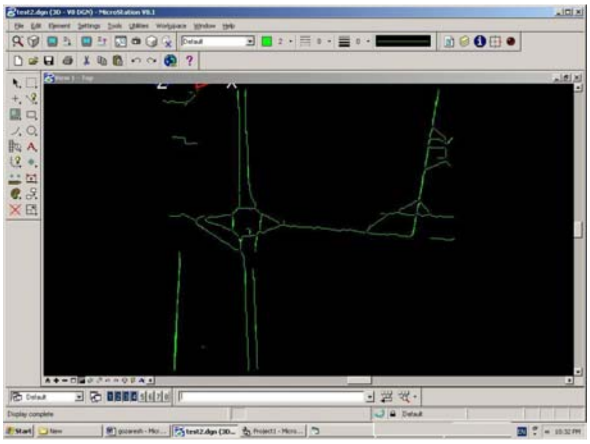


Figure 10: The extracted road network presented in CAD-based environment

In order to evaluate the obtained results, the road network available in the test area was manually extracted to produce a road network reference map, which is shown in Figure 11.

The comparison between Figures 10 and 11 showed that wherever a road pixel is extracted by the automatic procedure, its geometric position is exactly the same as manually produced road network, but the completeness of extracted road network is with doubt. Figure 12 shows those areas where the designed automatic procedure failed to detect road pixels, which are presented in boxes.

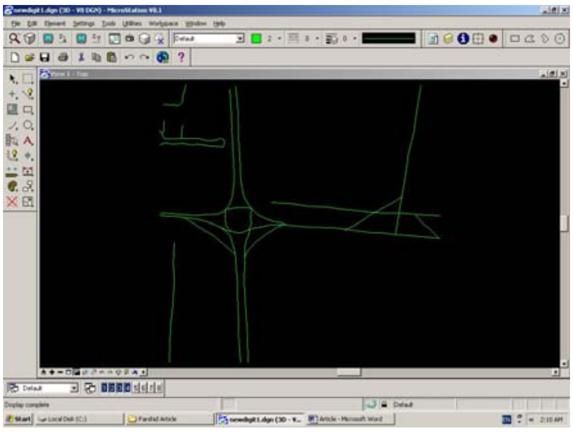


Figure 11: Manually produced road network

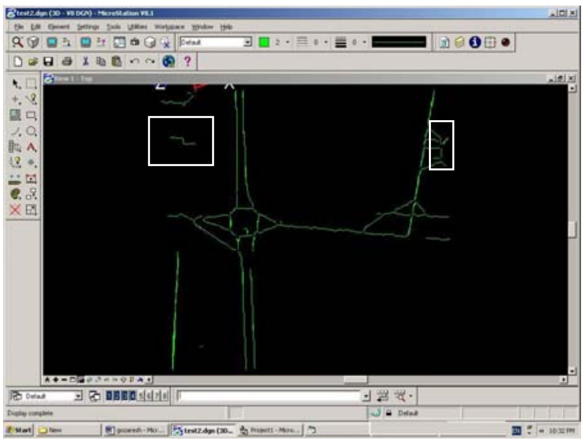


Figure 12: The differences between automatic extracted road centerline and its manually produced road network

4. CONCLUSIONS AND RECOMMENDATIONS

Image processing techniques such as pattern recognition and object detection algorithms do not provide fully structured data for GIS environments. This is mainly due to the fact that the results of image processing algorithms are raster maps while GIS environments need structured vector based data as their input entities. Furthermore, it should be noticed that CAD based facilities could not be performed directly on source imaged because of the complicated nature of them, especially when dealing with high resolution ones.

Therefore, the combination of image processing techniques and CAD facilities can be a good alternative to automate the process of data preparation and entering into GIS.

A new approach was designed and implemented in this research, and a fully automatic procedure was designed and implemented to extract road centerlines from high resolution satellite images. Neural networks and digital image processing algorithms such as morphological operators were used to extract road centerline from satellite images and present it as an edited road raster map. Then obtained raster map was prepared and vectorized using the facilities provided by CAD-based environments. The final result is the structured vector based road network presented in CAD environment where it can be easily transformed to GIS. The comparison between obtained results and road network reference map proved acceptable geometric accuracy of designed procedure.

The designed and implemented method in this research for automatic road extraction was mainly based on spectral and geometric properties of road networks. Knowledge-based road extraction methods, which are based on an appropriate road model, could be an efficient method for further study. In these methods, even when some parts of the road in input source image, road raster map or thinned road center line are missed, they could be recovered based on the road knowledge injected into the system via knowledge-based road model.

REFERENCES

GRUEN, A., AGOURIS, P., LI, H., 1995. Linear Feature Extraction with Dynamic Programming and Globally Enforced Least Squares Matching. In: Gruen, A., Kuebler, O., Aqouris, P. (Eds.), Automatic Extraction of Man-Made Objects from Aerial and Space Images. Brikhauser, Basel. 83-94.

GRUEN, A., LI, H., 1997. Semi-Automatic Linear Feature Extraction by Dynamic Programming and LSB-Snakes. Photogrammetric Engineering and Remote Sensing. 63(8): 985-995.

HARALICK, ROBERT M., AND LINDA G. SHAPIRO. 1992. Computer and Robot Vision. Volume I. Addison-Wesley, pp. 28-48.

HEERMAN, P.D. AND KHAZENIE, N., 1992. Classification of Multi-spectral Remote Sensing Data Using a Back Propagation Neural Network. IEEE Transactions on Geoscience and remote sensing. 30(1): 81-88.

MOHAMMADZADEH, A., TAVAKOLI, A., VALADAN ZOEJ, M. J., 2006. Road Extraction Based on Fuzzy Logic and

Mathematical Morphology from Pan-sharpened IKONOS Images. The Photogrammetric Record, Vol. 21 (113):44-60.

MOKHTARZADEH, M., VALADAN ZOEJ, M.J., 2007. Road Detection from High Resolution Satellite Images Using Artificial Neural Networks. International Journal of Applied Earth Observation and Geoinformation, 9(1):32-40.

NEUENSCHWANDER, W., FUA, P., SZEKELY G., KUBLER, O., 1995. From Ziplock Snakes to Velcro Surfaces. In: Gruen, A., Kuebler, O., Aqouris, P. (Eds.), Automatic Extraction of Man-Made Objects from Aerial and Space Images. Brikhauser, Basel. 105-114.

RICHARD, J.A., 1993. Remote Sensing Digital Image Analysis: Introduction. Second Edition, Springer, ISBN 0-387-5480-8, New York.

TRINDER, J., LI, H., 1995. Semi-Automatic Feature Extraction by Snakes. In: Gruen, A., Kuebler, O., Aqouris, P. (Eds.), Automatic Extraction of Man-Made Objects from Aerial and Space Images. Brikhauser, Basel. 105-114.

VOSSELMAN, G., KNECHT, J., 1995. Road Tracing by Profile Matching and Kalman Filtering. In: Gruen, A., Kuebler, O., Aqouris, P. (Eds.), Automatic Extraction of Man-Made Objects from Aerial and Space Images. Brikhauser, Basel. 165-274.