

ROAD NETWORK EXTRACTION FROM HIGH-RESOLUTION
MULTI-SPECTRAL SATELLITE IMAGES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

ERSİN KARAMAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF INFORMATION SYSTEMS

DECEMBER 2012

**ROAD EXTRACTION FROM HIGH-RESOLUTION MULTI-SPECTRAL
SATELLITE IMAGES**

Submitted by **Ersin Karaman** in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Information Systems, Middle East Technical University by,

Prof. Dr. Nazife Baykal
Director, **Informatics Institute**

Prof. Dr. Yasemin Yardımcı Çetin
Head of Department, **Information Systems**

Prof. Dr. Yasemin Yardımcı Çetin
Supervisor, **Information Systems, METU**

Examining Committee Members:

Assoc. Prof. Dr. Selim Aksoy
Computer Engineering, **BILKENT**

Prof. Dr. Yasemin Yardımcı Çetin
Information Systems, **METU**

Assist. Prof. Dr. Erhan Eren
Information Systems, **METU**

Assist. Prof. Dr. Banu Günel
Information Systems, **METU**

Assist. Prof. Dr. Alptekin Temizel
Work-based Learning, **METU**

Date:

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Ersin Karaman

Signature : _____

ABSTRACT

ROAD NETWORK EXTRACTION WITH HIGH-RESOLUTION MULTI-SPECTRAL SATELLITE IMAGES

Karaman, Ersin

Ph.D., Department of Information Systems

Supervisor: Prof. Dr. Yasemin Yardımcı Çetin

December 2012, 172 pages

In this thesis, an automatic road extraction algorithm for multi-spectral images is developed. The developed model extracts elongated structures from images by using edge detection, segmentation and clustering techniques. The study also extracts non-road regions like vegetative fields, bare soils and water bodies to obtain more accurate road map. The model is constructed in a modular approach that aims to extract roads with different characteristics. Each module output is combined to create a road score map. The developed algorithm is tested on 8-band WorldView-2 satellite images. It is observed that, the proposed road extraction algorithm yields 47 % precision and 70 % recall. The approach is also tested on the lower spectral resolution images with four-band, RGB and gray level. It is observed that the additional four bands provide an improvement of 12 % for precision and 3 % for recall.

Road type analysis is also in the scope of this study. Roads are classified into asphalt, concrete and unpaved using Gaussian Mixture Models. Other linear objects such as railroads and water canals may also be extracted by this process. An algorithm that classifies drive roads and railroads for very high resolution images is also investigated. It is based on the Fourier descriptors that identify the presence of railroad sleepers. Water canals are also extracted in multi-spectral images by using spectral ratios that employ the near infrared bands. Structural properties are used to distinguish water canals from other water bodies in the image.

Keywords: Road extraction, railroad extraction, water canal extraction, multi-spectral, transportation infrastructure analysis.

ÖZ
YÜKSEK ÇÖZÜNÜRLÜKLÜ ÇOK BANTLI UYDU İMGELERİNDEN
KARAYOLU AĞI ÇIKARIMI

Karaman, Ersin
Doktora, Bilişim Sistemleri
Tez Yöneticisi: Prof. Dr. Yasemin Yardımcı Çetin

Aralık 2012, 172 sayfa

Bu çalışmada, multispektral imgeler için otomatik yol çıkarımı algoritması geliştirilmiştir. Geliştirilen model, kenar çıkarımı, bölütleme ve kümeleme yöntemlerini kullanarak uzun-ince yapıları çıkarmaktadır. Bu çalışma ayrıca daha yüksek performanslı yol çıkarımı sonucu almak için bitki, tarla ve su gibi yol olmayan bölgeleri de tespit etmektedir. Model, farklı özelliklere ait yolları çıkarmak için modüler bir yaklaşımla geliştirilmiştir. Her bir modüle ait sonuçlar bir araya getirilerek puan haritası oluşturulmuştur. Geliştirilen algoritma, WorldView-2 uydu görüntülerinde test edilmiştir. Ortalama %70 geri çağırma ve %47 hassasiyet ile yol bulma başarımı sağladığı gözlenmiştir. Bu yaklaşım daha düşük spektral çözünürlüğü olan 4-bant, RGB ve gri seviye imgelerde de test edilmiştir. Worldview-2 uydusu ile birlikte gelen ek dört bandın % 12 geri çağırma ve % 3 hassasiyet kazandırdığı gözlemlenmiştir.

Bu çalışma kapsamında yol kaplama malzemesi sınıfı belirleme analizi de gerekleřtirilmiřtir. Gauss Karıřım Modeli kullanılarak yol kaplama malzemesi asfalt, toprak ve beton olmak üzere sınıflandırılmıřtır. Demiryolu ve su kanalı gibi diđer dođrusal yapılar da bu yaklařımla ıkarılabilmektedir. Ayrıca, ok yksek öznrlkl imgeler iin demiryolu ve karayolu sınıflandırması yapan algoritma geliřtirilmiřtir. Bu yaklařım demiryolu traverslerin varlıđını belirten Fourier zneliklerini temel almaktadır. Bunun yanında multispektral imgelerde yakın kızıl tesi bantlarını kullanan spektral oranların yardımı ile su blgeleri ıkarılmıřtır. Yapısal zellikler su kanallarının diđer su blgeleri ile ayrımı iin kullanılmıřtır.

Anahtar Kelimeler: Yol ıkarımı, demiryolu ıkarımı, su kanalı ıkarımı, multispektral, ulařım altyapısı analizi.

To my father Fatih Karaman

and

Daughter Elif Naz Zeynep...

ACKNOWLEDGEMENTS

I would like deeply thank my supervisor Prof. Dr. Yasemin Yardımcı Çetin for her guidance, valuable feedback, continuous support and friendly attitude. This thesis would not be complete without her help and generosity in sharing her knowledge and experiences.

I thank my thesis monitoring committee Assoc. Prof. Dr. Selim Aksoy, Assist. Prof. Dr. Erhan Eren for their guidance and feedback. Also I thank my thesis jury members Assist. Prof. Dr. Alptekin Temizel and Assist. Prof. Dr. Banu Günel.

I am grateful to my dearest friends Umut Çınar, Ekin Gedik and Selvi Elif Gök for their invaluable help during my study. They were with me all the time and accompaniments of them were significant in completion of this work.

I also thank my friends Dr. Çağaçan Değer, Nurcan Alkış, Gizem Baybaş, Okan Bilge Özdemir, Bilge Sürün, Aykut Mert Yakut, Ege Saygıner, Önder Gülbeyaz, Mahir Kaya, Adnan Öztürel, Gökçe Oğuz, Kadri Gökhan Yılmaz, Emrullah Kızıltunç and Dr. Serdar Polat for their support. I would also like to credit Informatics Institute members Sibel Gülnar, Yaşar Sayın, Çetin İnci, Sibel Sel, Hakan Güler, Murat Yabancı and Ferda Ercan for their assistance.

I thank The Scientific and Technical Research Council of Turkey (TÜBİTAK), ASELSAN, HAVELSAN and Digital Globe Inc. for their support and cooperation.

Once for all, I would like to thank my family and my in-laws, especially my brother Assoc. Prof. Dr. Selçuk Karaman, for their endless support, love and encouragements. I would like to express my deepest gratitude to my daughter, Elif Naz Zeynep Karaman and my wife Güler Karaman for their love and support and for not losing their patience throughout my study.

TABLE OF CONTENTS

ABSTRACT.....	iv
ÖZ.....	vi
ACKNOWLEDGEMENTS.....	ix
TABLE OF CONTENTS.....	x
LIST OF TABLES.....	xiv
LIST OF FIGURES.....	xv
CHAPTER 1.....	1
1 INTRODUCTION.....	1
1.1 Purpose of the study.....	2
1.2 Justification of the Study.....	2
1.3 Research Questions:.....	3
1.4 Thesis Overview.....	4
CHAPTER 2.....	5
2 LITERATURE REVIEW.....	5
2.1 Road Extraction.....	5
2.1.1 Semi-Automatic Road Extraction Methods.....	5
2.1.2 Automatic Road Extraction Methods.....	12
2.2 Water Canal Extraction.....	19
2.3 Railroad Extraction.....	20
CHAPTER 3.....	22
3 BACKGROUND.....	22
3.1 Canny Edge Detection.....	22
3.2 Mean-Shift Segmentation.....	24

3.3	K-means	26
3.4	Gaussian Mixture Model.....	27
3.5	Segment Analysis.....	30
3.5.1	Calculating Eccentricity	30
3.5.2	Calculating Elongatedness Value.....	30
3.6	Spectral Band Ratios	32
CHAPTER 4.....		34
4	ROAD EXTRACTION ALGORITHM.....	34
4.1	Road Characteristics	34
4.2	Road Extraction Model	35
4.3	Road Region Detection	37
4.3.1	Edge-Based Road Extraction	38
4.3.2	Structure Analysis Based Road Detection	41
4.3.3	Clustering Based Road Detection	45
4.4	Non-Road Region Detection.....	47
4.4.1	Structure Analysis Based Non-Road Region Detection.....	47
4.4.2	Vegetation Detection.....	48
4.4.3	Water Region Detection.....	50
4.5	Score Map Construction.....	53
4.5.1	Score Map Construction Based on Segments	53
4.5.2	Statistical Analysis Based Score Map Construction	54
4.6	Water Canal Extraction.....	55
4.7	Railroad Extraction	58
4.8	Road-Type Classification.....	73

4.9	Verification	75
4.9.1	Precision:.....	76
4.9.2	Recall:	76
4.9.3	F-Measure:	76
4.9.4	Performance Measurement	77
CHAPTER 5.....		79
5	RESULTS AND DISCUSSION	79
5.1	Dataset Used in Experiment.....	79
5.2	Results.....	84
5.3	Road Extraction	85
5.3.1	Edge-Based Road Extraction	85
5.3.2	Structure analysis based road extraction	91
5.3.3	Clustering-Based Road Extraction.....	96
5.3.4	Structure Analysis Based Non-Road Detection	101
5.3.5	Vegetation Detection.....	105
5.3.6	Water Extraction	106
5.3.7	Score map Construction.....	109
5.3.8	Overall Result	111
5.4	Road Material Classification.....	112
5.5	Water Canal Extraction.....	115
5.5.1	WV-2 Image.....	115
5.5.2	Four-band.....	115
5.6	Railroad Extraction	116
5.7	Discussion.....	121

5.7.1	Comparison with Other Methods.....	132
CHAPTER 6.....		138
6	CONCLUSION AND FUTUREWORK.....	138
6.1	Limitations	140
6.2	Future Works	141
REFERENCES		143
APPENDICIES		151
APPENDIX A: ALL ROAD EXTRACTION RESULTS.....		151
APPENDIX B: STRUCTURE BASED ANALYSIS PERFORMANCES OF ALL BAND COMBINATION		162
APPENDIX C: WATER BODY EXTRACTION RESULTS.....		166
APPENDIX D: WATER CANAL EXTRACTION RESULTS		169
APPENDIX E: PARAMETERS USED IN THE STUDY		170
APPENDIX F: THE ROLE OF SPECTRAL BAND OF WV-2.....		171
Curriculum Vitae		172

LIST OF TABLES

Table 1 Sample result for perpendicular axis approach.	62
Table 2 Performance comparison of ring, perpendicular axis and fused method	63
Table 3 Spectral properties of WV-2	80
Table 4 Edge based road extraction module performance for different width values.	87
Table 5 All result of edge based road detection module	90
Table 6 Result of Structure analysis based road extraction module for WV-2 Image.....	92
Table 7 All result of structure analysis based road extraction	96
Table 8 Result of cluster based module for WV-2 image	97
Table 9 Combined results of clustering based road extraction module	100
Table 10 Road extraction results for all bands.....	111
Table 11 Segment analysis decision matrix for area and elongatedness values.	125
Table 12 Comparison of modules	130

LIST OF FIGURES

Figure 1 Eccentricity measure.....	30
Figure 2 Sample segments and basic elongatedness value.	31
Figure 3 Sample segments and elongatedness values	32
Figure 4 General approach to road extraction.....	36
Figure 5 Automatic road extraction algorithm.....	37
Figure 6 Sample RGB and Ground Truth Image	38
Figure 7 Middle Point's Width Calculation.....	39
Figure 8 Edge base road detection module.	39
Figure 9 Edge Based Module Result.....	40
Figure 10 4-Band structure analysis based road extraction result.....	42
Figure 11 Structural Analysis for Road Detection.....	43
Figure 12 Intermediate results for structural analysis module	44
Figure 13 Clustering Based Road Detection.....	46
Figure 14 Cluster based road extraction intermediate results	46
Figure 15 Structural analysis based non-road region detection flow chart	47
Figure 16 Structural analysis based non-road region samples	48
Figure 17 Vegetation detection flow chart.....	49
Figure 18 Vegetation extraction result.....	49
Figure 19 Water region detection.....	51
Figure 20 Water extraction intermediate result.....	52
Figure 21 Segment based score map and thresholded mask.....	54
Figure 22 Automatic threshold computation	55
Figure 23 Water region extraction flowchart	56
Figure 24 Intermediate results of water canal extraction	58
Figure 25 Sample synthetic drive road and railroad images and their Fourier transforms.	59
Figure 26 The Ring approaches result for synthetic data.....	61
Figure 27 Sample railroad synthetic images and Perpendicular approaches elements.....	62

Figure 28 Sample result for combination of perpendicular axis and ring method	63
Figure 29 Adjacent roads and their Fourier transforms	64
Figure 30 Sample image for railroad classification approaches	65
Figure 31 The railroad classification approach flowchart.....	65
Figure 32 Sample Railroad and Highway Images and their Fourier Transforms	66
Figure 33 Histogram of Fourier spectra energy for railroads and drive-roads.....	67
Figure 34 Sample Fourier transform and noise regions.....	68
Figure 35 Histogram of Fourier spectra energy computed by compensating for background energy regions.....	69
Figure 36 Boxplots of Fourier spectra energy for three different images.....	70
Figure 37 Sample Score map	71
Figure 38 Sample thresholded Score map and final mask	71
Figure 39 Labeled given road mask and the algorithm result.....	72
Figure 40 Road Type classification flow chart.....	74
Figure 41 Road type classification sample result.....	75
Figure 42 Mask-Mask Performance measurement.....	77
Figure 43 Mask-Skeleton performance measurement.....	77
Figure 44 Mask-Skeleton Recall Performance	78
Figure 45 WV-2 Image obtained 16 November 2010.....	81
Figure 46 WV-2 Image (Eymir Lake).....	82
Figure 47 WV-2 Image (Yeni Mahalle)	82
Figure 48 General view the image taken from Russia.....	83
Figure 49 Reflectance comparison of some satellite images	84
Figure 50 Sample image for demonstration.....	85
Figure 51 Edge Based Result for WV-2 image.....	86
Figure 52 Edge Based Result for Four-band image.....	88
Figure 53 Edge Based Result for the RGB image.....	89
Figure 54 Edge Based Result for Gray image.....	89
Figure 55 Edge Based Result for Panchromatic image.....	90
Figure 56 Edge based module results	91

Figure 57 Minimum-region parameter of Mean-Shift segmentation.....	92
Figure 58 Structure analysis based module result for WV-2	93
Figure 59 Structure analysis based result for four-band image.....	94
Figure 60 Structure analysis based result for RGB image	94
Figure 61 Structure analysis based for gray image	95
Figure 62 Structure analysis based result for panchromatic image.....	95
Figure 63 Structure analysis based road extraction module results.	96
Figure 64 Clustering based result for WV-2 image	97
Figure 65 Clustering based result for four-band image	98
Figure 66 Clustering based result for RGB image	99
Figure 67 Clustering based result for gray image	99
Figure 68 Clustering based result for panchromatic image	100
Figure 69 All result of clustering based road extraction module.	101
Figure 70 Structure analysis based non-road result	102
Figure 71 Structure analysis based non-road result for four-band.....	103
Figure 72 Structure analysis based non-road result for RGB image.....	103
Figure 73 Structure analysis based non-road result for gray level images.....	104
Figure 74 Structure analysis based non-road result for panchromatic image	104
Figure 75 Vegetation results for WV-2 image.....	105
Figure 76 Vegetation results for four-band image.....	106
Figure 77 Water results for WV-2 image.....	107
Figure 78 Water region detection result for four-band.	108
Figure 79 Road Network result for WV-2 image.....	109
Figure 80 Road Network result for Four-band image.....	109
Figure 81 Road Network result for RGB image.	110
Figure 82 Road Network result for Gray image.....	110
Figure 83 Road Network result for Panchromatic image.....	111
Figure 84 Algorithm Results.....	112
Figure 85 Road Material Classification Result for WV-2 Images	113
Figure 86 Road Material Classification Result for four-band Images	113

Figure 87 Road type classification result on GT and road extraction algorithm output	114
Figure 88 Water results for WV-2 image.....	115
Figure 89 Water results for four-band image.....	116
Figure 90 Railroad result	117
Figure 91 Railroad classification result.	118
Figure 92 Sample railroad classification result.....	119
Figure 93 Sample road parts having periodic structure	120
Figure 94 Sample railroad parts without discernable periodicity	120
Figure 95 Edge linking sample	122
Figure 96 Edge detection result for each band.....	123
Figure 97 Edge registration problem	124
Figure 98 Elongatedness and Area analysis.....	125
Figure 99 Clusters and clustering result for different K values	127
Figure 100 Comparison of EVI index calculated by NIR2 and average of NIR1 and NIR2.....	128
Figure 101 Band ratios segmentation result.....	129
Figure 102 RGB image and the algorithm output.....	130
Figure 103 Comparison of modules.....	131
Figure 104 Sample rural area result and comparison with Bacher & Mayer (2005)	133
Figure 105 Comparison with Senthilnath et.al (2009).....	134
Figure 106 Comparison with Lee et.al (2000)	134
Figure 107 Comparison with Rajeswari et.al (2011).	135
Figure 108 Comparison with Liu et al. (2003).....	136
Figure 109 Comparison with Yuan et al. (2009).	137
Figure 110 All results of Cankaya-1 Image	151
Figure 111 All results of Cankaya-2 Image	153
Figure 112 All results of Yeni Mahalle-1 Image	154
Figure 113 All results of Yeni Mahalle-2 Image	155
Figure 114 All results of METU (1)	156
Figure 115 All results of METU (2)	157
Figure 116 All results of METU (3)	158

Figure 117 All results of METU (4)	159
Figure 118 All results of METU (5)	160
Figure 119 All results of METU (6)	161
Figure 120 Water extraction for Eymir	167
Figure 121 Water Extraction result for METU	167
Figure 122 Water extraction result for east of Ankara.....	168
Figure 123 Water extraction for Yenimahalle	169

CHAPTER 1

1 INTRODUCTION

The importance of satellite systems has constantly been rising in response to the increasing decision needs in several areas. Satellite images provide information for urban planning, traffic management, emergency management, crop estimation, navigation and military operations.

In addition to panchromatic images, contemporary satellites also provide hyper-spectral and multi-spectral images which enable us to analyze the Earth's surface in more detail. With the introduction of these hyper-spectral and multi-spectral images, the spatial and spectral resolutions of satellite images have been continuously improving. In terms of spatial resolution, this means that the objects that were previously covered by one pixel are currently covered by multiple pixels. For instance, previously buildings are represented in one or two pixels in a satellite image; however, they have different sizes in new satellite images. This is also true for spectral resolution.

Object detection from satellite images is also one of the most actively studied areas. Techniques used in this area include image processing, analysis and machine learning algorithms to detect, identify and recognize objects in an image. However, above-mentioned technical improvements also render a continuous necessity for the development of more complex object analysis techniques.

One of the most prominent problems in object detection is the extraction of roads from satellite images. A modular automatic road extraction algorithm based on image processing techniques is developed in this thesis.

1.1 Purpose of the study

Acquiring accurate information is important for faster decision making. To obtain such information for roads, in addition to the construction of an accurate road map, one must also be able to discriminate different types and characteristics of roads. Utilization of multi-spectral and high-spatial resolution images is also critical to realize the needs due to technical improvements. In this context, the purpose of this study is to fulfill the following objectives:

- 1) to develop an automatic road extraction algorithm for high resolution multi-spectral satellite images to get accurate and detailed road map,
- 2) to determine road properties like road type (asphalt, concrete, soil).
- 3) to discriminate roads as motorway, railroad and waterway,
- 4) to determine the contribution of new spectral bands to the accuracy of road extraction

1.2 Justification of the Study

Analysis of high resolution satellite images has been an important research topic for accurate and up-to-date GIS information including roads that are crucial for urban planning (Rajeswari, Gurusurthy, Omkar, Senthilnath, & Reddy, 2011). Roads are very important components of most applications such as; updating existing road maps for car navigation (Zhaoa, Kumagai, Nakagawa, & Shibasaki, 2002), city planning, management, cartography, (Wang, Qin, Du, Chen, & Tao, 2005), data acquisition, GIS systems updates, databases or site models for car navigation and emergency (rescue) system which may need instant map (Vandana, Chandrakant, & Ramachandran, 2002), producing topographic databases for military purposes and disaster management or relief (Bacher & Mayer, 2005), automated mapping (Ravanbakhsh, Heipke, & Pakzad, 2007), trajectory planning for unmanned aerial vehicles (Sırmaçek & Ünsalan, 2010) and topographic mapping (Zhang & Couloigner, 2006),

Developing a road extraction algorithm from multi-spectral images will provide more accurate information to enable effective decision making for military and emergency situations, urban planning, map updating, navigation support, vegetation field analysis and so on.

Extracting roads manually from remotely sensed data is a costly and time consuming process (Mohammadzadeh, Valdan Zoej, & Tavakoli, 2008). Semi-automatic methods, on the other hand, still have their own problems that will be detailed in Chapter 2. It is also stated that the ability of the upcoming sensors will produce better spatial resolution data which motivates researchers to develop automated road extraction approaches (Rajeswari et al., 2011).

Roads may also be composed of different surface materials, and they have different characteristics in urban and rural regions. However, current studies do not acknowledge this point. So, an approach that acknowledges this difference rather than ignoring it is not just necessary but also may improve the performance of road extraction methods.

Ever improving techniques in road extraction methods have also forced the development of more adaptive and flexible approaches. That is, it would be better if a newly developed efficient technique can easily be integrated into the existing ones. One method that satisfies this adaptability is to develop modular approaches. These approaches have also the advantage of being more data set independent.

1.3 Research Questions:

This study aims to extract roads from high spatial resolution and multi-spectral satellite images. In order to achieve this goal, five main research questions will be answered.

R1: How can we develop an algorithm that can extract roads from 8-Band multi-spectral and high spatial resolution satellite images?

R2: How will 8-Band MS image contribute to the accuracy of extracted roads compared to the 4-band, RGB and Gray level satellite image?

R3: How can we develop an algorithm that can be used to classify extracted roads as drive-roads and railroads?

R4: How can we develop an algorithm that can be used to extract water canals?

R5: How can we develop an algorithm that can be used to determine road cover material?

1.4 Thesis Overview

This thesis is organized as seven chapters. The problem is defined in Chapter 1. In Chapter 2, the literature on road extraction from different perspectives is given. This chapter also includes railroad and waterway extraction studies in the literature. Chapter 3 explains the theoretical background of the proposed methods. It also describes the methods that are used in the model. Chapter 4 presents the proposed automatic road extraction algorithm. Each component of the model is detailed in this chapter. Chapter 5 presents data and experiments. In Chapter 6, results are presented separately for each type of data and comparison of the approach with other techniques is also provided. Chapter 7 presents the conclusions. Limitation and future works are also mentioned in this chapter.

CHAPTER 2

2 LITERATURE REVIEW

2.1 Road Extraction

Various road network extractions from satellite images have been proposed in literature. These approaches can be classified with different taxonomies. Mena (2003) proposed a classification system of road extraction studies with a summary of road extraction methods. These classes are determined using on preset objective, the extraction techniques, and sensor type. Moreover, Özkaya (2009) classified road extraction studies into six categories; basic image processing methods, frequency based methods, knowledge base approaches, supervised techniques, segmentation methods and other methods.

In this study, the methods will be classified as automatic and semi-automatic approaches since most of them can be classified into more than one class according to the above mentioned taxonomies. Semi-automatic means that the algorithm requires user feedback before or during the road extraction operation. On the other hand, automatic road extraction methods work without user feedback. The literature review is reported chronologically.

2.1.1 Semi-Automatic Road Extraction Methods

Park and Kim (2001) propose a semi-automatic method for road extraction from IKONOS images. According to the researchers, the distinctive feature of their method is the introduction

of template matching algorithm to be used for road extraction. The method starts with the initial user input that indicates a point on the road center, i.e. road seed, for each segment in the image. Then orientation of the seed is calculated by Burns line extraction algorithm, where the algorithm extracts various lines around the initial point and the orientation of the longest line is taken as the orientation of the seed. In the next stage, by using seed and orientation, a template window and an initial target window (that is formed by shifting the template in the direction of orientation) are defined. Next, by using adaptive least square matching and similarity transformations a true target window that best matches with the template window is formed. This true target window is then used to form another target window and the iterative process continues until two consecutive matches fail. A test for the method is conducted by using a one meter resolution panchromatic IKONOS image with. The results revealed that the method performed fairly well for roads with a few exceptional cases such as the ones with shadows, obstacles, abrupt orientation change, and “complicated structures of exchange”. However, even these cases may be handled by the user’s providing another road seed. One assumption of the method is that there exist small differences between the brightness values of the template window and the target window. Besides, the method is strictly dependent on the validity of the initial road seed input which raises an important limitation, as is also stated by the researchers. They mainly focus on main road extraction.

In the study of Zhao et al. (2002), a semi-automatic approach is proposed. This study uses 4-band (red, green, blue, and NIR) and 1 meter resolution pan-sharpened satellite images obtained from IKONOS. The study consists of three main stages: road mask extraction using commercial remote sensing software, road seed extraction via edge detection and road line extraction using template matching based on the road mask and road seeds. In order to extract the road mask, maximum likelihood method based on training values is used to classify the image. For road seed extraction, Canny edge detection algorithm is used. By tracing edge pixels, sudden and fast change points are determined and edge line is broken. In addition, edge lines which have similar direction and small gaps are merged. Then, starting with the user input template matching algorithm is applied. The user also assigns two directional control points for avoiding wrong direction matching. In terms of performance of the study, an experiment is conducted and it is observed that, the algorithm can return 4005 pixels by using a start and two control points. It is stated that, the final road mask may not include small roads and edges may belong long structures other than roads which may cause a failure of this approach. Even though it is stated

that this approach is valid for main road extraction in urban areas and all roads in rural areas, it actually depends on training data which is used for road mask extraction, edge based road seeds and user feedback.

The method that is proposed in the study of Vandana et al. (2002) integrates cost minimization and path following techniques for semi-automatic road extraction. This method includes a pre-processing stage where Non-Linear Anisotropic Coherence diffusion is used for scale-space routine and Canny edge-detector is used for edge detection. After pre-processing the method starts with two road input points provided by the user. From these points, the orientation and the width of the road are calculated. The results of edge-detection stage are used for the calculation of the width. From the orientation and width, three possible directions are extracted for further propagation. The cost of each direction is estimated by considering variance of intensity, direction, length and width of the road and the direction with the minimal cost is chosen as the next path. The iterative process continues until the cost value reaches the pre-defined values. If the algorithm encounters a junction point (same cost of possible next point), it selects one points to continue and save other points to follow later. The researchers tested the method on both IKONOS and aerial images (1 m and 2 m resolution). The results showed that the method performs well on roads in any orientation with moderate curvature, and on junction points. However, the model has some limitations such as roads cast by shadow may not be extracted, the input roads must have considerable width, and the edges of the roads must be detectable through the pre-processing. Results are presented visually rather than quantitative performance metrics. The performance of this semi-automatic approach depends on selected seed points as well as image quality. It is also important to note that path following operation cannot work if the edge is not extracted by Canny detection.

Peteri and Ranchin (2003) proposed a two-step method for detecting and extracting road networks from high resolution images. In the experiments, 1 meter panchromatic IKONOS satellite images are used. The approach also uses prior knowledge about roads such as models of roads and properties of road network. Algorithm starts with extracting a topologically correct graph. This graph could come from a database or could be extracted automatically from the image by minimizing a cost function which evaluates the homogeneity of the local radiometric variance in several propagation directions. In the next step, roads are reconstructed as surface elements using the road graph obtained in the former step. This step uses specific active contours (snake) technique with multi-resolution analysis (MRA). Reconstruction procedure also consists

of two steps, namely extraction of parallel road sides and their intersections. Road segments with parallel sides are extracted using a new object called Double-Snake which composes two branches where two snakes are evolving jointly. Following the extraction of parallel road sides, intersection regions are identified using simple snakes which are initialized by pairing the extremity points of the Double Snakes. These simple snake objects are called “intersection snakes” in the procedure. A multi-resolution analysis is performed made the original image. Double Snakes is applied onto the several approximated images of different resolution they run on the original image. After Double Snakes minimizes the energy function, Intersection Snakes are initialized and run in same manner. In the results section, only outcomes of the reconstruction phase are given, and the graph is not automatically extracted. It is stated that the algorithm is successful in extracting the contours of the roads.

T. Kim, Park, T. Kim, Jeong and K. O. Kim (2004) suggested a semi-automatic method which is based on template matching. Their method which tracks road centers is tested on 1 m resolution IKONOS images and it is based on least squares template matching. It is stated that the algorithm is successful on extracting road centerlines in different orientations. Proposed method starts with the user input. It is expected in this step that the user selects a point which is on the road centerline. A template window with the selected point at the center is defined and its orientation is calculated. This step is either handled by automatic line extraction or using human interaction once again for direction estimation. An initial target window is then generated by shifting the template window using the orientation information estimated. Least Squares Correlation matching is used after this step to update the position and orientation of the target window. Proposed method works iteratively, estimating the new position and orientation of the target window and updating the corresponding properties of the template window at each step. Tests on IKONOS images showed that using three to five starting points by the user for each road segment as input, most parts of the road network could be extracted. Since the algorithm is based on template matching, it is also vulnerable to false inputs from user. If the user selects a point which is not inside a road region, the algorithm still tries to track regions with similar spectral characteristics. Finally it is stated that the proposed template matching method without global enforcement is a good alternative to energy minimization applications in the feature extraction area.

In the study conducted by Wang et al. (2005), a multi-resolution approach is proposed based on the assumption that roads may show different characteristics at different levels of resolution.

Firstly, the satellite image is divided by using pyramidal decomposition to three different resolution levels: low, mid and high resolution. For the low resolution image, as roads can be approximated by line structures, median filter, Canny edge detection, 8-direction filter are applied sequentially. To get a more accurate road map, a set of operations is applied; breaking edges at acute points, edge tracing, removing small edges and merging close and appropriate edges. For mid-high resolution images, it is assumed that roads are polygonal structures having parallel edges and homogenous spectral information. To extract roads, the image gray value is thresholded to get candidate road masks, and then the morphological operation is applied to remove small components. Edges are extracted by the Canny filter, and Hough transform is applied on subsets of the image to find parallel edge pairs. Then mid-lines of these parallel edges and road regions are extracted. In order to combine the results, all masks are turned into original resolution and overlapped to extract all candidate roads. These candidate road are verified based on central lines and parallel sides. Manual correction is also applied for proper road selection. An IKONOS panchromatic image is used in the experiments, and the results are presented visually rather than quantitatively. It is stated that this semi-automatic method is more suitable for main road extraction.

In the study conducted by Mohammadzadeh et al. (2008), firstly the user selects a few initial road points from IKONOS image. Three bands of the images were used. Then genetic algorithm is applied to calculate mean values of roads for the three bands by using five classes for each band. At the next step, fuzzy segmentation is applied on these mean values. After thresholding the fuzzy matrix, a binary mask is extracted. Morphological operations are applied to thin the binary mask, and ultimate road skeleton is extracted. It is reported that this method provides 77% road extraction accuracy.

In the study conducted by Mohammadzadeh, Valdan Zoj and Tavakoli (2009), a semi-automatic approach is proposed for road extraction from multispectral satellite images. The algorithm requires a seed point generated by the user to create fuzzy classes belonging to the road region in each band based on intensity. Then random particles are generated. For each particle a fuzzy cost function is calculated and best particles that have minimum cost are selected to update the particle population. The best particles that have less cost function value than pre-defined threshold are selected and fuzzy reasoning operation is applied to create the preliminary binary road mask. Mathematical morphological post-processing operations are applied to the road mask to obtain noise removed road centerline. Based on the result, road detection correctness

(i.e. “amount of committed errors”) coefficient is 0.88 while, background detection correctness coefficient is 0.93. This approach is tested on IKONOS images.

Movaghati, Moghaddamjoo, and Tavakoli (2010) introduced a semi-automatic tracking based algorithm which is a combination of Extended Kalman Filtering (EKF) and Particle Filtering (PF) methods to identify road segments in satellite images. Test images are obtained from two different satellites which are Indian Remote Sensing (IRS) with 5.8 m spatial resolution and IKONOS with 0.8 m resolution. The main algorithm starts with an initial point belonging to the road which is supplied to the system by an operator. Using this seed point which inherits the information such as the coordinates of the road center, direction and width estimation, a profile cluster is initialized and Kalman Filtering is performed. Using the initial state and initial profile cluster, EKF module starts to track the road. This step is iterative, so on each step profile clusters are updated and new ones are added. EKF module stops when it comes to an obstacle or a junction where the average profile error exceeds a predefined threshold. Following this step, PF module starts to run. By using a single road branch which is estimated by the former EKF module, N particles are generated. Then those particles will search for a new seed point which EKF module can take as a starting point. If such a seed point cannot be detected after predefined number of steps for any road branch, the algorithm is terminated and the final road network is obtained. In case of an appropriate seed point, EKF module starts to work and continues to trace the road after the obstacle. The proposed method obtained a very high correctness score (0.98) on both satellite images. Completeness scores slightly differ regarding to the satellite type; being 0.92 for IRS and 0.85 for IKONOS. It is stated that the algorithm is generally tested on rural areas and for a better comparison it should be also tested in dense urban areas.

Lin, Zhang, Liu, Shen and Duan (2011) developed a semi-automatic road network extraction algorithm based on tracking the lane markers present on the road for very-high-resolution (VHR) remote sensing images. The proposed approach consists of four steps namely, preprocessing of the image, seed points entry by the user, automatic road tracking, and operator intervention for post-editing. At the preprocessing step, the input image is enhanced and stretched using the dynamic range adjustment method when needed. Following this step, the operator selects three points on the road which define the starting point, direction and width of the road. Using this information, a rectangular window of a road template is initialized and the road markings inside this region are also identified. After the template is initialized, automatic tracking of roads begins. Road tracking is handled by finding the next road axis point by least squares matching

using initial template to obtain the accurate location of the next road center point. After adding this point to the extracted road network, the algorithm checks the condition to see if any stopping measure is fulfilled. When the tracking algorithm stops on such criterion, operator intervention is required. The operator then verifies and modifies the extracted road segments and starts the tracking procedure from another seed point if needed. The algorithm is tested on two different types of images, a QuickBird panchromatic image with 0.6 m resolution and an aerial image acquired using DMC camera with 0.2 m resolution. The quantitative results were given as root mean square error per pixel and it is 1.3 pixels for QuickBird and 1.9 pixels for DMC images. The results are compared with similar matching and tracking methods and the proposed approach outperformed them for the QuickBird image. It is stated in the conclusion that the current method cannot automatically detect the width and direction of the road and update road reference model. Thus, it performs poorly on sharp turns, junctions and shadows over roads.

In their study, Da Silva and Centeno (2012) propose a semi-automatic road extraction method that uses Red band in 3-band (red, green, blue) high resolution aerial images. Their method consists of two stages, i.e. semiautomatic iterative extraction of centerlines and the extraction of the road track. In the study, a road is modeled as a chain of segments. Accordingly, the first stage starts by the first segment defined by the user input and then continues with the extraction of new segments in the search region. This stage uses Radon transform to find candidate segments, and the performance depends to two user defined parameters that determine the search region: the radius (r) and chord (c). The candidate with the largest gray level mean value is regarded as the best solution and the process continues iteratively until the search region is completed or no other candidate can be proposed. In the second stage, road borders are determined by using estimated width of the tracks which must be defined by the user in accordance with the road type. By using three aerial images, the method is evaluated with regard to completeness (i.e. “ability of method to extract roads”), correctness, and RMS index (i.e. “the geometric quality of the detected centerlines”) criteria. Results show that completeness is between 0.60 and 0.94, while correctness is between 0.6-0.94. RMS value varies between 0.7-1.4. Also, after the comparison of their method with the method proposed by Zhang and Couloigner (2006), researchers claim that these two methods are complementary in the sense that their method is better in main roads, whereas the other is better in secondary roads.

2.1.2 Automatic Road Extraction Methods

In Zhang, Shunji and Emmanuel (1999), mathematical morphology was applied on 1-m resolution satellite images to find roads. Roads are assumed as areas rather than lines because of high resolution. It is assumed that road networks are formed by elongated areas. In the study; classification (gray level analysis), segmentation and size distribution analysis, trivial opening, closing hole filling, removing small paths operations have been applied for road extraction. Semi-urban areas obtained from Toronto and Bern are used to demonstrate the algorithm. The approach is also used for aerial images.

In the study conducted by H. Y. Lee, Park and H. K. Lee, 2000, 1 meter resolution IKONOS images have been used. The method consists of two steps. Firstly, road primitives were extracted using an intensity based segmentation approach called hierarchical gradient watershed algorithm. This algorithm was modified to avoid the over segmentation problem. With the assumptions that roads are elongated and large objects with constant intensity and have high contrast with their surroundings, intensity based segmentation is used and elongated and large road segments can be selected by analysis of mean gray value, size (number of pixels in segment) and shape information (major/minor axis). In order to complete missing roads, profile matching methods have been used basing on the primitive road segments direction. Experimental results are provided visually.

Laptev et al. (2000) proposed a fully automatic approach for extracting road network from grayscale aerial images. Their method starts with extraction of lines in the image at coarse scales. In coarse scales, it is easier to disregard disturbances in roads such as spectral changes and artifacts. But it is also easier to complete and verify the road network in finer scales which inherits more information. The proposed algorithm fuses the finer parts of multi-scale information to obtain a more refined road network. After extracting the lines in coarse scale, salient, non-salient roads and crossings are identified in a finer scale using the line information obtained in the former step. Salient roads are defined as complete, easily distinguishable roads where as non-salient ones are disturbed by occlusions from the environment. To extract both types of roads, rubber snake method is employed. False lines are eliminated using the fact that roads generally have uniform widths. By optimizing the rubber snakes, salient roads are found. Non-salient roads are found by first creating an optimal path between two salient road segments and then comparing the optimal path with the extracted one. After extracting the salient and non-

salient roads, junctions are identified with a similar hypothesis step. After the junctions are extracted, these structures form a road network. Although the method is generally developed for rural areas, experiments show that it tends to give successful results in some urban cases. For three 0.5 m aerial images, correctness and completeness measures are given. Proposed algorithm obtained relatively high correctness (>0.95) and completeness (>0.72) scores in those images.

In Peteri et al. (2003), two sequential modules were used to extract roads from IKONOS images; Graph management and Reconstruction. In the graph management module, they aim to obtain a topologically correct graph of road network by selecting the best path for the potential road by minimizing a cost function which evaluates the homogeneity of local radiometric variance for several propagation directions. In the reconstruction module applied on the obtained graph, the snake algorithm and wavelet based multi-resolution approach (MRA) are used. It is stated that this method is not suitable for urban areas, and results have not been quantitatively evaluated.

In the study conducted by Liu, Li and Chapman, 2003, genetic algorithm, fuzzy c-means and clustering analysis have been used to optimize clustering parameters for automatic road extraction from pan-sharpened 61 cm QuickBird satellite image. The experiments have been conducted on 350X214 image frames including vegetation, bare soil and paved roads. Although results look promising, the experiment is conducted on a simple satellite image which is not covering water, railway, buildings and other man-made or natural objects.

The study conducted by Li et al. (2003) proposed an approach based on edge detection method. Firstly, edges are extracted by applying the Laplacian filter. Then, in order to remove non-road edges and complete broken or missing road edges, the snake method is applied. Then, final road lines are vectorized with the help of commercial GIS software. This approach is tested on 2.5 m resolution SPOT-5 images, and experimental results are presented visually.

Shackelford and Davis (2003) used an iterative approach for extracting roads from multispectral high resolution satellite images. The study uses pan-sharpened IKONOS images with 1 meter resolution. The main algorithm starts with masking the pixels in the image which contribute to vegetation areas. Masking is done using NDVI index with a predefined threshold. The length and width features of roads are used for extraction. For each remaining non-vegetation pixel, maximum and minimum length line segments comprising pixels with similar spectral characteristics which pass through the pixel is identified. The maximum length value, angle of

the maximum length line segment, endpoints of the maximum length segment and the minimum length value is stored for each pixel. Then each line segment is iteratively examined and added to the road network if it satisfied the conditions defined by the user. Iterative algorithm starts with the pixel with the highest maximum line segment length. Every pixel inside the line segment is checked if its angle and minimum line segment values are valid for a road part and labeled as road if it is. After this step, line segments are allowed to grow from end points in the general direction of the road segment. This is repeated until no growth is possible. This way, algorithm could track roads around curves and partially occluded areas. Finally, the parts of the obtained road segment that falls inside the buffer area are checked. This buffer defines the area around the extracted segment which potential new line segments will be examined. If the angles of the road sections are not in conflict with the buffer, these parts are also included in the road network set. The algorithm continues until no line segment length values larger than the minimum length set by the user are found. Algorithm is tested on two different data sets with urban and suburban characteristics. It is stated that better results could be obtained with improved ability to track curves and close gaps.

Mena and Malpica (2005) addressed to automatic road network updating in GIS database. The approach uses RGB image and GIS vector information for this aim. As a first step, available road vectors and corresponded areas in the images is selected for creating base information for segmentation. In addition, median filter is applied to the image to eliminate noise. As a subsequent step, segmentation is applied based on the training data. Then segmented image is processed by using edge generalization, skeleton extraction, graph generation, geometrical adjustment and topological adjustment respectively. In the evaluation process different resolution images were used from 1m to 3m resolution. According to the test, this method provides 81% completeness and 87% correctness in average.

An integrated approach is proposed by Long and Zhao (2005). In this approach grey level morphology, mean-shift segmentation, mathematical morphology, edge tracing and convex-hull algorithm are used under the assumption that roads are continuous and narrow regions which are lighter than their neighborhood sides, and have low gray-level variance and approximately fixed widths. Firstly grey-level morphology (cleaning and strengthening) is applied as a pre-processing step to keep edges and smooth the image. Then mean-shift algorithm is applied to this preprocessed image to obtain “simplified image map” including homogenous areas. Then, grey-level thresholding is applied on the image map to get potential binary road mask.

Mathematical morphological opening and closing is applied to the mask to remove small pixel groups and merge small gaps. In order to remove holes, edge tracing is applied. Lastly, convex-hull algorithm is used to detect minimal convex shape of regions. This approach is applied on optical satellite images, and quantitative results are not provided. Even though it is an integrated system and applicable to salient “Manhattan city” regions, still it is sequential approach and the performance depends on previous steps.

An automatic road extraction approach using multi-spectral image is proposed by Bacher and Mayer (2005). This approach focuses on rural and sub urban regions. IKONOS images have been used for demonstration of the approach. First road characteristics are determined including both geometric and functional features. Roads are elongated regions and they have steady grey values. In addition, they form a network which provides navigation from one point to another. In order to extract roads based on these assumptions, firstly Streger line extraction algorithm is applied on all band of the image. To construct training areas of road regions, it is expected that these lines have a parallel pair and the region between parallel lines have constant grey value. Candidate road regions extracted from all bands are fused and used as training areas for classification. Fuzzy classification is applied using grey level variance and mean values and a probability map is constructed. In order to verify road region, fuzzy values belong to road lengths, widths and classification result are used. Minimum value of these three values is used as final weight of road regions. After thresholding the weight map, preliminary road mask is obtained. Then road network is generated by filling gaps. Performance of this approach is measured based on manually digitized reference data. Results show that, this approach provides 0.76 completeness and 0.76 correctness.

In the study of Zhang and Couloigner (2006), road extraction approach was proposed using multispectral satellite images. The approach starts with K-Means image segmentation. K-value was set to “6” based on experimental studies. Then fuzzy logic classifier was applied on the result with the assumption that roads have high reflectance in R, G, B bands and lower reflectance in the NIR band. Since, this spectrally segmented result may include objects having similar spectral characteristics, shape information is used to eliminate noises in the road cluster. Then the road network is extracted. According to the result, average completeness for different images is in the range of 0,45-0,73, while correctness is in 0,30-0,63. Results vary according to the complexity of the image. Since this method’s success essentially depends on the k-means algorithm results, any failure in this step may cause ineffectiveness of the approach.

In the study of Hsu and Yang (2007), a feature extraction method for classification of hyper-spectral and high resolution images is proposed. AVIRIS data set has been used in the study, and it is concluded that wavelet transform is an effective tool for feature extraction. In Hu, Razdan, Femiani, Wonka, and Cui (2007), Fourier and pixel tracking based technique has been used for road extraction with the assumption that road network has low curvature. The technique is based on road tracking which follows lines. In order to find the starting point for road tracking, Fourier shape descriptor of sample road seed (pixel footprint) is used. Validation of method is not based on a quantitative measure. Only sample output images are provided.

In the study of Senthilnath, Rajeshwar and Omkar (2009), Normalized Cut method and texture progressive analysis (TPA) have been used for road extraction from high resolution multispectral satellite images. A commercial street area has been selected as a sample data since the study is intended to extract roads from urban regions. QuickBird satellite images (0,61m panchromatic and 2,4 multispectral resolutions) are used. The approach is separated into three modules; pre-processing, TPA and Normalized Cut method and performance evaluation. The pre-processing module consists of three steps; gray level images thresholding to get binary image, grouping of pixels and masking the vegetation area. In the TPA analysis, firstly training samples including different road textures are selected from the images. Then a window region is defined and its Haralick features are extracted for training samples. Lastly, the image is divided into small windows and their Haralick features are computed for comparison with the Haralick features of training set. Based on the difference of those two feature sets, the image is classified as a road and non-road region. On the other hand, the image is segmented by using normalized cut method. The image is partitioned into classes to form road regions. In order to evaluate performance of the methods (TPA and normalized cut), correctness, completeness and quality measures are calculated in addition to other performance metrics. The result shows that, normalized cut method (correctness 92.7%, completeness 88.4 % and quality 86.1%) is more accurate than TPA (correctness 89.6%, completeness 73.5 % and quality 72.8%).

In Xu, Zhang and Liu (2009), roads are classified based on the number of lanes and widths. Road characteristics are also defined according to their geometric, radiation, topological and context characteristics. In short, roads are elongated structures and their lengths are greater than their widths. They have small curvature, fixed width and clear edges. Gray values are uniform in the area of roads. They have high contrast compared to their surroundings. They form a network which means that they are connected. Based on these assumptions, watershed Dual-Threshold

algorithm has been applied on the image. Then, the deviations from the mean of road pixels and from the 8- neighborhood are weighted to determine the road edges. If edges obtained from water-shed and Multi-weighted methods are the same, morphological operation is applied for noise removal. As a last step, shape index (area/perimeter ratio) is applied to separate road segments from non-road structures. Results are presented visually.

The study provided by Grote, Heipke, Rottensteiner and Meyer (2009) suggests a road network algorithm for CIR images. The approach consists of three main parts; segmentation, road part extraction, sub graph generation. After segmentation of the image by the normalized cut method, segments with similar shapes are merged based on their shapes. Then, each grouped segment is analyzed by using shape (elongatedness, width uniformity) and spectral information (NDVI, color variation) to extract road parts. Then road parts are connected based on their distance, orientation and color similarity. Lastly, by using the contextual information of objects like cars and trees, the results are improved. They are presented in visual format.

Yuan, Wang, Wu, Yan and Li (2009) proposed an approach for automatically extracting road networks from satellite images based on the principle that the roads are relatively thin and homogeneous regions with smooth directional changes. The proposed approach is built on the oscillatory correlation theory and the LEGION (Locally Excitatory Globally Inhibitory Oscillatory Networks) method is used both for segmenting the image and grouping road segments. Proposed method is tested on several input satellite images with different resolutions (4 meters and 30 meters). The algorithm starts with segmenting the gray-level image using the LEGION method. Since the LEGION algorithm is based on leaders which are centered on the center of large homogeneous regions and the roads are relatively thin and elongated regions, road sections are expected to be inside the background segment after the segmentation operation. After obtaining the background segment, the medial axis transform is employed to obtain candidate road regions. Using the Voronoi diagram as the boundary pixels of segments as samples, road regions are obtained with respect to their average distance of medial axis points to the boundary. Following this procedure, resulting segments are mostly road sections. To obtain a more refined representation, resulting segments are regrouped using the LEGION algorithm. Based on the assumption that road segments are aligned collinearly or curvilinearly, the LEGION model with long range horizontal connections to the medial axis points of segments is applied to obtain road segment grouping. In this step, if most of the medial axes of a segment are activated, this segment is regarded as a final road segment. This three step approach is applied

onto both synthetic and real satellite images. It is stated that a prior smoothing for images could result in better outcomes.

Sırmaçek and Ünsalan (2010) proposed a road extraction algorithm from high resolution satellite images. The algorithm consist of five main steps; smoothing the image, edge detection and local feature extraction, spatial matrix creation, analysis of the matrix to find initial point and tracking algorithm. After smoothing the images, Canny edge detection is applied to the data, and gradient and orientation information is extracted. Based on these local features, a spatial voting matrix is constructed. In order to extract the road segment, spatial voting matrix is segmented at different levels. Each segment is analyzed whether it satisfied the elongatedness property of roads by using area and perimeter metrics ratio.

Rajeswari et al. (2011) have developed road extraction algorithm for QuickBird satellite images which have 2,4m MS and 0,6m Pan resolution. The proposed method consists of two main parts; pre-processing and segmentation. For the pre-processing, image is classified into 15 clusters and five of them are considered as road class. Then in order to compensate for the deterioration of road segment due to cars other occluding objects, nearest neighborhood grouping is applied. Then median-filter is applied on the grouped image. Then Normalized-Cut- and Mean-Shift segmentation methods is applied separately and performance result are compared based on the segmentation method. For the panchromatic images, normalized-cut procedure provided 96 % accuracy while mean-shift 80% only. On the other hand, for the MS images (2,4 m resolution) Mean-Shift procedure provides 95 % accuracy while Normalized-Cut provides only 82 % accuracy.

In the study of Rajeswari et al. (2011), level-set and mean-shift methods are compared for road extraction. IKONOS images are used in the experiments. Before applying the methods, the image is clustered into 20 clusters and five of those are selected as the road class. Then for smoothing nearest neighborhood grouping and Median filer are applied. According to the results, mean-shift provide better performance (Completeness, 93%, Correctness. 95%) than level-set (Completeness, 93%, Correctness, 87%) algorithm.

2.2 Water Canal Extraction

Water body extraction is a well-studied problem of the contemporary remote sensing studies domain. Automatic extraction of water bodies from remotely sensed data is important for maintaining an up-to-date geographic database, effective usage of water resources and potential flood estimation. Furthermore, having water map for an image may also enable us to classify other critical objects like ships, bridges, dams.

Most of water extraction studies are developed for low resolution images. Fraizer and Page (2000) summarized and discussed some studies of water body extraction for different low resolution satellite images. Johnston and Barson (1993) proposed a water body extraction method for Landsat TM images. They used images obtained from different seasons in a year to analyze how the water and vegetation fields vary. To do this, classification methods have been used. In the study of Alesheik, Ghorbanali and Nouri (2007), histogram thresholding and band ratio based water extraction (coastline detection) method is proposed. They used Landsat 7, Landsat 5 and Landsat 4 images representing Urmia Lake. Li et al. (2003) proposed a shape-based change detection of water body by using Landsat images.

High resolution satellite images are also used for water extraction and analysis studies. Luo et al. (2007), proposed a water extraction method in IKONOS images to extract bridges. They detect water areas in such images by using Markov Random Fields (MRF) and Support Vector Machines (SVM) based supervised approach. After classification, mathematical morphology operations are used as post-processing step for water extraction. Since the study mainly aims to extract bridges there are no performance values related to water except the error matrix. Chaudhuri and Samal (2008) have also proposed a method for extraction of bridges over water regions. Firstly, the image is classified into eight classes (i.e. ice/snow, shrubs, concrete, water, sand, forest, soil, and rock) by using supervised classification methods. Then classification result divided into three parts as water, concrete and others. IRS-C satellite images have been tested. Another study for bridge extraction is proposed by Gu, Zhu, Shen, Hu and Chang (2011). The approach consists of two main steps. Firstly; the image is filtered according to its spectral features by using iterative thresholding. In the second step, k-means clustering algorithm is employed to use both spatial and textural information to detect water regions.

Nath and Deb (2010) reported that, most of the unsupervised water extraction techniques in the literature use spectral bands and spectral band indexes. These bands and ratios are threshold by pre-defined value to filter image and extract water body. For example, Zhao, Yu and Zhang (2009) proposed a water extraction algorithm for IKONOS images. The method consists of five main parts: NDVI and NDWI calculation, segmentation, feature definitions, classification and feature extraction and post-processing steps.

2.3 Railroad Extraction

Railroads are important for both passenger and freight transportation. Railroads on satellite images are often extracted as road objects due to their common properties with roads such as elongated and curved structures with uniform intensity. In high resolution aerial images, however, the sleepers between railroad tracks depict a distinctive texture from other elongated objects. Since sleeper objects are equally spaced and perpendicular to the railroad tracks, these additional cues can be employed to distinguish railroad objects in aerial images. Even though road extraction has been studied extensively, there are very limited studies on railroad extraction. In the study of Neubert et al. (2008), 5 cm spatial resolution digital ortho-imagery and LIDAR are used to extract railroads. They first attempted to use combined Fourier and Hough transformation, edge detection, object recognition, image segmentation tools of commercial software but were not satisfied with the results. They proposed two algorithms: the first was based on the extraction of straight lines and arcs of railroads using an adapted RANSAC algorithm, and the second was a knowledge-based algorithm that uses railroad properties such as fixed distance between tracks, free of objects, horizontal rail bed.

In the study Hellwitch, Laptev and Mayer (2002), SAR images are used to extract linear structures. In order to test the proposed method, images including pipelines and railroads were used. Firstly, Markov Random Field based line extraction was proposed and employed, then ribbon snake approach was applied on the extracted lines. However, this approach focuses only on the linear structure extraction rather than railroad detection.

There are other periodic structure extraction studies. In Rabatel, Delenne and Deshayes (2008), a method is proposed to detect vine-plots based on Fourier transform and Gabor filters. In Aksoy, Yalniz and Tasdemir (2012), high resolution satellite images have been used to extract orchards with periodicity. These studies aim to extract periodic vegetative field extraction which may have different periodicity distance.

CHAPTER 3

3 BACKGROUND

A summary of methods proposed for road extraction is presented in Chapter 2. In the present chapter, some of the techniques that are utilized in the proposed method are explained in more detail.

3.1 Canny Edge Detection

Edge detection methods are the most commonly used technique for road extraction. There are several edge detection methods including Sobel, Prewitt, Roberts, Laplacian and Canny.

Canny edge detection algorithm is a well-known gradient based edge detection method proposed by Canny (1986). The algorithm simply aims to ensure three criteria: good detection, good localization and single response per edge. The technique is constructed on these requirements in order to ensure robust edge detection. The algorithm is composed of four main steps;

a) Gaussian Smoothing

Raw image noise is a challenging issue for Canny method as it has been built on a gradient based strategy. A 2D Gaussian filtering operator is applied on the input image in order to decrease the

degradation due to the noise. The convolution operation defined in Equation 3.1 is employed for the given input image $f(m, n)$.

$$g(m, n) = G_\sigma(m, n) * f(m, n) \quad (\text{Equation 3.1})$$

where Gaussian kernel is defined as;

$$G_\sigma(m, n) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{m^2+n^2}{2\sigma^2}\right] \quad (\text{Equation 3.2})$$

b) Gradient Filtering

Canny edge detection algorithm aims to locate edge points in all directions, so the derivatives along both x and y axes is computed. By using derivative information, one can compute magnitude the edge gradient and its direction as follows;

$$M(m, n) = \sqrt{G_x^2(m, n) + G_y^2(m, n)} \quad (\text{Equation 3.3})$$

$$\theta(m, n) = \arctan\left(\frac{G_y(m, n)}{G_x(m, n)}\right) \quad (\text{Equation 3.4})$$

where G_x and G_y are the first derivatives along x and y axes, respectively. Gradient direction θ is computed by using Equation 3.4, and the angle values are rounded to one of the eight different directions including diagonals.

c) Non-maximum Suppression

Having reduced the variety of edge directions into eight directions, non-maximum suppression operation is applied in order to thin the edge ridges. The non-maximum suppression operation removes the non-maxima pixels along the edge direction, resulting in thin edges with one pixel width. That is, the edge magnitudes $M(m, n)$ are interpreted such that if the gradient magnitude of the current edge pixel is greater than two of its neighbors along the gradient direction $\theta(m, n)$, then it is marked as an edge; otherwise it is set to null.

d) Hysteresis Thresholding and Edge Linking

It is hard to define a single threshold suitable for determining all of the edge pixels without any false alarms. Canny algorithm uses of hysteresis thresholding technique to extract edges based on their neighborhood. The hysteresis thresholding is defined as;

$$M'(m, n) = \begin{cases} 1, & M(m, n) > T_h \\ 1, & T_h > M(m, n) > T_l \text{ AND } M(m, n) \text{ is neighbour with any pixel } M(m', n') > T_h \\ 0, & M(m, n) < T_l \end{cases} \quad (\text{Equation 3.5})$$

where T_l and T_h are low and high thresholds defined by the user. The edge linking operation begins with a pixel with magnitude greater than high threshold, and the edge tracing continues as long as the pixel values are greater than the low threshold.

3.2 Mean-Shift Segmentation

Segmentation is another method employed in classification problems. There are many segmentation methods in the literature. In our study, mean-shift segmentation is used.

Mean-Shift is a nonparametric density gradient estimation method employing a generalized kernel based mode seeking approach. It was firstly generalized to a simple feature space analysis technique similar to k-nearest neighbor approach in the study of Fukunaga and Hostetler (1975). In a more recent article (Cheng, 1995), the author has further investigated and formulated the mean-shift procedure to reveal its strength for data analysis. Mean-shift algorithm's ability to converge in multimodal feature spaces for discrete data is proved in the study Comaniciu and Meer (2002), and a robust image segmentation algorithm based on Mean-shift is proposed. In the context of feature space analysis, mean-shift considers the local maxima points as modes of the given discrete data sampled from an unknown density function. Therefore, the problem is reduced to a nonparametric function estimation problem. Mean-shift algorithm employs pre-defined kernel functions to iteratively find the weights of nearby points.

In the pattern recognition literature, the most popular kernel density estimation method is the Parzen window technique (Parzen, 1962), which is also used in our implementation. For a given data vector x_i where $i = 1, \dots, n$ defined in the d -dimensional space R^d , let $f(x)$ be the multivariate kernel density estimation function, $K(x)$ is the kernel function and H is the $d \times d$ bandwidth matrix;

$$f(x) = \frac{1}{n} \sum_{i=1}^n K_H(x - x_i) \quad (\text{Equation 3.6})$$

where

$$K_H(x) = |H|^{-\frac{1}{2}} K(H^{-1/2}x) \quad (\text{Equation 3.7})$$

The d-variate kernel $K(x)$ is assumed to satisfy following conditions;

$$\int_{R^d} K(x) dx = 1 \quad (\text{Equation 3.8})$$

$$\int_{R^d} xK(x) dx = 0 \quad (\text{Equation 3.9})$$

$$\lim_{\|x\| \rightarrow \infty} \|x\|^d K(x) = 0 \quad (\text{Equation 3.10})$$

$$\int_{R^d} xx^T K(x) dx = c_k I \quad (\text{Equation 3.11})$$

where c_k is a constant. The kernel function is assumed to be radially symmetric (Comaniciu and Meer, 2002), therefore; we can obtain the final density estimation function as follows,

$$f_{h,K}(x) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \quad (\text{Equation 3.12})$$

In this equation, the size of search window is controlled by bandwidth parameter, h . Based on the definition of density estimation function, the local maxima points in the feature space are the locations of modes having zero gradient, so the mean-shift procedure can be written as;

$$\nabla f_{h,K}(x) = \frac{2c_{k,d}}{nh^{d+2}} \left[\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \right] \left[\frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \right] \quad (\text{Equation 3.13})$$

where $g(x)$ is the kernel profile and assumed to be nonnegative. From Equation 3.13, we can deduce that the density estimate at x computed with the kernel G is proportional to the first term;

$$f_{h,G}(x) = \frac{c_{k,d}}{nh^{d+2}} \left[\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \right] \quad (\text{Equation 3.14})$$

The mean-shift vector is represented by the second term in the Equation 3.15, which is the difference between the weighted means. The final mean-shift vector can be written as;

$$m_{h,G}(x) = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \quad (\text{Equation 3.15})$$

Simplifying these formulas yields to;

$$m_{h,G}(x) = \frac{1}{2} h^2 c \frac{\nabla f_{h,K}(x)}{\nabla f_{h,G}(x)} \quad (\text{Equation 3.16})$$

The proportionality of the mean-shift vector computed with kernel G at location x with the normalized density gradient estimate produced by kernel K is clear in this equation. Therefore, it can be inferred that the mean-shift vector always has the same direction with the maximum intensity increase (Fukunaga and Hostetler, 1975). A stationary point of the estimated density is satisfied due to the fact that the local gradient estimate is aligned with the mean-shift vector. The local maxima points, so called modes, are these stationary points. The convergence of the mean-shift procedure is guaranteed by the successive nature of the algorithm. That is, at each iteration, mean-shift vector $m_{h,G}(x)$ is computed and the kernel $G(x)$ is translated by $m_{h,G}(x)$. Moreover, the Mean-Shift algorithms steps larger when the kernel is aligned with the regions having low density. Likewise, as approaching to the local maxima, the Mean-Shift vector's magnitude begins to decrease. Therefore, an adaptive gradient ascent feature space analysis is accomplished.

3.3 K-means

In order to classify an image into object types, clustering methods are commonly used. K-means is a well-known and popular clustering algorithm commonly used in data mining. K-means aims to partition “n” observations into “k” clusters where the observations are assigned to the nearest mean. Therefore, the data space can be divided into Voronoi cells. The original k-means algorithm was firstly published in Lloyd (1982). For a given set of observations (x_1, x_2, \dots, x_n) where each observation is a d-dimensional vector, the aim is to partition the n observation into k

distinct clusters $C = \{C_1, C_2, \dots, C_k\}$ with the constraint $k \leq n$, so that the average distance of the observations to the cluster centers is minimum, i.e.

$$\operatorname{argmin}_{C_i} \sum_{i=1}^k \sum_{x_j \in C_i} \|x_j - \mu_i\|^2 \quad (\text{Equation 3.17})$$

where μ_i is the mean of the current cluster C_i .

The k-means algorithm requires the number of clusters k and initial cluster centers $m_1^{(1)} \dots m_k^{(1)}$ from the user. In the assignment step, each observation is assigned to its closest cluster center resulting in a different Voronoi diagrams. The assignment operation can be written as;

$$C_i^{(t)} = \{x_p : \|x_p - \mu_i^{(t)}\| \leq \|x_p - \mu_j^{(t)}\| \forall 1 \leq j \leq k\} \quad (\text{Equation 3.18})$$

where x_p represents a single observation. In the update step, new mean values for each cluster are calculated by using;

$$\mu_i^{(t+1)} = \frac{1}{|C_i^{(t)}|} \sum_{x_j \in C_i^{(t)}} x_j \quad (\text{Equation 3.19})$$

The algorithm converges when the cluster assignments no longer change.

In this study, we utilize random partition method to initialize cluster centers which randomly selects the cluster centers from the given observation set. Moreover, as a distance metric, we employ the Euclidean distance.

3.4 Gaussian Mixture Model

Gaussian mixture models are beneficial tools that are able to provide good semi-automatic clustering functionality. That is, sub-clusters within the given set of observations can be identified by estimating the parameters of Gaussian distributions formed by each clusters.

For d -dimension, Gaussian distribution of a set of observations $x = (x^1, x^2, \dots, x^d)^T$ is defined as ;

$$N(x|\mu, \Sigma) = \frac{1}{(2\pi)^{d/2} \sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right) \quad (\text{Equation 3.20})$$

where μ is the mean and Σ is the covariance matrix of the Gaussian.

By using the inferences presented by Bayesian theory, we can write the formulation of a mixture of K Gaussians as follows;

$$P(x) = \sum_{j=1}^K w_j \cdot N(x|\mu_j, \Sigma_j) \quad (\text{Equation 3.21})$$

where Gaussian distribution with j th index has the prior probability of w_j . Thus;

$$\sum_{j=1}^K w_j = 1 \text{ and } 0 \leq w_j \leq 1 \quad (\text{Equation 3.22})$$

The problem that is intended to be solved is the estimation of the parameter set θ of the GMM model fitting the data. For a given set of data $X = \{x_1, x_2, \dots, x_N\}$ belonging to an unknown distribution, we can construct the solution by the help of maximum likelihood formulation as follows;

$$\theta = \operatorname{argmax}_{\theta} p(X|\theta) = \operatorname{argmax}_{\theta} \prod_{i=1}^N p(x_i|\theta) \quad (\text{Equation 3.23})$$

In the context of this study, we employ expectation maximization (EM) algorithm to estimate hidden variables of the given distributions. The EM algorithm has an iterative structure, and at each iteration two major operations take place. The first operation is named as Expectation-Step; the distribution of the hidden variable is estimated for the given data and the current value of the parameters is computed. At the next step, the joint distribution of the observations and the hidden variable is maximized. In fact, the maximization step is dedicated to modify the distribution parameters according to hidden variables in order to increase the likelihood of the data and the hidden variables.

The EM algorithm can be summarized as follows;

$$Q(\theta, \theta^t) = E_Z[\log(p(X, Z|\theta^t))] \quad (\text{Equation 3.24})$$

where Z is the hidden variable. The maximization step involves the solution of ;

$$\theta^{t+1} = \operatorname{argmax}[Q(\theta | \theta^t)] \quad (\text{Equation 3.25})$$

Therefore, the likelihood value of $p(X|\theta)$ is proportional to $Q(\theta, \theta^t)$.

As long as the Z is known, the mixture parameters can be estimated. However, the hidden variables are not always about to be known. Thus, we need to introduce indicator variables in order to write joint likelihood of all X and Z as follows;

For a given definition of the indicator variable z_{ij} ;

$$z_{ij} = \begin{cases} 1 & \text{if } j - \text{th Gaussian emitted } x_j \\ 0 & \text{otherwise} \end{cases} \quad (\text{Equation 3.26})$$

$$L(X, Z, \theta) = p(X, Z | \theta) \quad (\text{Equation 3.27})$$

$$L(X, Z, \theta) = \prod_{i=1}^N \prod_{j=1}^K [p(x_i, j | \theta)]^{z_{ij}} \quad (\text{Equation 3.28})$$

$$L(X, Z, \theta) = \prod_{i=1}^N \prod_{j=1}^K [p(x_i, j | \theta)]^{z_{ij}} [p(j | \theta)]^{z_{ij}} \quad (\text{Equation 3.29})$$

Then taking the logarithm both of the sides yields;

$$\log[p(X, Z)] = \sum_{i=1}^N \sum_{j=1}^K z_{ij} \log[p(x_i, j | \theta)] + z_{ij} \log[p(j | \theta)] \quad (\text{Equation 3.30})$$

The E-step can also be expressed by using the Equation 3.31 as follows;

$$E_z(z_{ij} | \theta^t) = \frac{p(j | \theta^t) p(x_i | j, \theta^t)}{p(x_i | \theta^t)} \quad (\text{Equation 3.31})$$

which also indicates the posterior distribution. Moreover, the M-step is given by the equation;

$$\frac{\partial Q(\theta, \theta^t)}{\partial \theta} = 0 \quad (\text{Equation 3.32})$$

Where

$$\theta = (w_j, \mu_j, \Sigma_j), j = 1, \dots, K, \quad (\text{Equation 3.33, a})$$

$$\sum_{j=1}^K w_j = 1 \text{ and } 0 \leq w_j \leq 1 \quad (\text{Equation 3.33, b})$$

Finally, we can write mean and correlation matrices of the mixtures as follows;

$$\mu_j = \frac{\sum_{i=1}^N p(j | x_i, \theta^t) x_i}{\sum_{i=1}^N p(j | x_i, \theta^t)} \quad (\text{Equation 3.33})$$

$$\Sigma_j = \frac{\sum_{i=1}^N p(j | x_i, \theta^t) [(x_i - \mu_j)(x_i - \mu_j)^T]}{\sum_{i=1}^N p(j | x_i, \theta^t)} \quad (\text{Equation 3.34})$$

3.5 Segment Analysis

After segmentation operation, features are extracted for classification. For road extraction problems, segments are expected to be elongated structures. In order to make this analysis some metrics are used in this study.

3.5.1 Calculating Eccentricity

Eccentricity can be calculated as $E = \sqrt{1 - \frac{b^2}{a^2}}$. This also equals to $\frac{f}{a}$ where “f” is foci point and “a” is the major axis. An eccentricity value close to 1 corresponds to a thin segment, whereas the segment is similar to a circle when the eccentricity is close to zero.

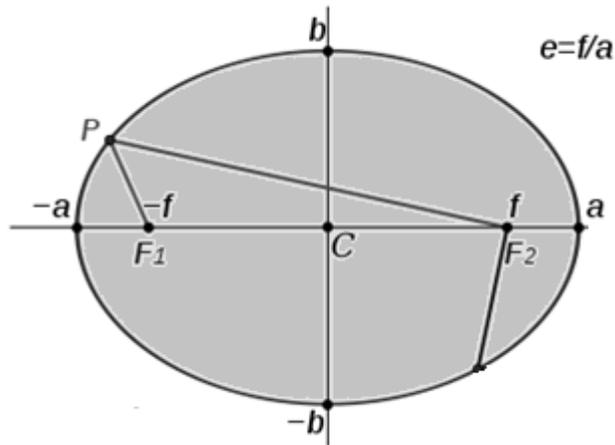
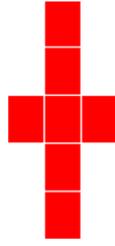


Figure 1 Eccentricity measure

3.5.2 Calculating Elongatedness Value

As stated in the study of Lee et al. (2000), road segments can be selected by analysis of their mean gray values, sizes (number of pixels in the segment) and shape information (major/minor axis). However, a high value of $\frac{a}{b}$ (major/minor axis) ratio is not sufficient to show segment's elongatedness. Below, two simple segments are shown and minor/major values are calculated (Figure 2).



Major/Minor=5/3



Major/Minor=8/1

Figure 2 Sample segments and basic elongatedness value.

Based on the ratio, elongatedness value can be reproduced.

$$\text{Elongatedness} = \frac{a}{b}. \quad (\text{Equation 3.35})$$

For the ideal shape (rectangle) we can say that the area A is given by

$$A = a \cdot b \quad (\text{Equation 3.36})$$

So elongatedness value based on the ratio of minor and major axis values can be calculated as

$$\text{Elongatedness} = \frac{a^2}{A}. \quad (\text{Equation 3.37})$$

Both segments have road characteristics, however, Figure 2.a, may not be detected as a road segment since its elongatedness value is smaller. Therefore, this equation is suitable only if the segment has an ideal shape like rectangle. In order to compensate this effect, Equation 3.38 can be used. In this equation, “extent” represents the ratio of pixels in the segment to number of pixels in the total bounding box.

$$\text{Elongatedness} = \frac{(a \cdot (2 - E))^2}{A} \quad (\text{Equation 3.38})$$

When we consider the structure of segments, for the ideal case of a rectangle (which is presented on Figure 2-b), the extent value is equal to one. However, the extent value for an irregular segment (such as the one that is presented on Figure 3-a) approaches zero. It is worthy to note that the extent value can vary in the range between 0 and 1 depending on the regularity of the

segment. From this equation it can be observed that elongatedness is dependent on the extent value, since (2-Extent) will be the coefficient for an approximation of the ideal major axis length value. For example, in Figure 3, there are three sample segments and respective elongatedness, $\frac{a}{b}$ and extent values are shown



Figure 3 Sample segments and elongatedness values

3.6 Spectral Band Ratios

In remote sensing, spectral band ratios provide essential discrimination for specific such as vegetation, shadow, water and land cover. By the introduction of multispectral imagery, band ratios became popular since they are capable of reducing topographical effects such as shadowing and surface slope directions among the bands of the image. With the help of band ratios, one can construct a robust indicator for the material of interested.

More specifically the band where the material poses high reflectivity and another band where the material poses absorption characteristics can be combined to obtain a good indicator for the material. The band with high reflectivity should exist in the numerator whereas the band with absorption characteristics should be in the denominator. In this study, we employ three different band ratios including Normalized Difference Vegetation Index (NDVI), Soil Adjusted Vegetation Index (SAVI) and Normalized Difference Water Index (NDWI). These band ratios are used in our road extraction model to detect specific materials such as water and vegetation.

NDVI has application areas in the fields of vegetation detection and plant growth monitoring. Its value ranges between -1 and 1. The vegetative regions are bright in the NDVI mask, in other words; vegetation takes larger values in NDVI. The water regions, on the other hand, take values close to 1. NDVI is defined as follows;

$$NDVI = \frac{NIR-Red}{NIR+Red} \quad (\text{Equation 3.39})$$

SAVI is another band ratio useful for discriminating vegetation from the background. In fact, SAVI is an extended version of NDVI exploiting the effects caused by soil brightness. As in the Equation-4.6 The correction factor for the soil brightness is controlled by L parameter.

$$SAVI = \frac{NIR-Red}{NIR+Red+L} * (L + 1) \quad (\text{Equation 3.40})$$

Water regions can be detected by NDWI from a multispectral satellite image, which is defined as;

$$NDWI = \frac{Green-NIR}{Green+NIR} \quad (\text{Equation 3.41})$$

These band ratios will be employed for vegetation and water body detection in the upcoming chapters.

CHAPTER 4

4 ROAD EXTRACTION ALGORITHM

As mentioned in the literature review chapter, many approaches have been proposed in order to extract roads from remotely sensed data. However, the developed algorithms produce promising results only for roads in open landscape and in restricted regions. Furthermore, semi-automatic or supervised approaches which depend on the data characteristics and satellite sensors have been used. Roads may have different characteristic in urban and rural regions. They may also be composed of different surface materials. This implies that it is required to develop a modular approach for road extraction algorithm. Moreover, in order to have a fully automatic road extraction algorithm providing higher road extraction performance for rural, sub-urban area and different surface materials, spatial and spectral features should be used in this modular approach. The proposed modular algorithm of this study focuses on extraction of road and non-road regions of the data. In addition, water-canal, railway detection and road material type classification are also addressed.

4.1 Road Characteristics

Roads are the prominent component of transportation system. Extraction of road from remotely sensed data via image processing techniques bring about some assumptions about road definitions. In the literature, Bacher and Mayer (2005) defines roads as elongated regions and having uniform grey values. They also stated that roads form topological network. It is frequently implied in the literature that roads are elongated structure having steady grey level intensity (Mena et al., 2005; Mohammadzadeh et al., 2009; Zhang and Couloigner, 2006)

In the study conducted by Long et al. (2005), roads are defined as continuous and narrow regions with brighter intensity values lower variance than the surrounding pixels. It is mentioned that roads widths vary slowly (Vandana et.al, 2002; Zhao et al., 2009). Moreover, while roads are linear structures for low resolution images, they are parallelogram regions having homogenous gray values for middle and high resolution images (Wang et.al, 2005).

In this study, in the light of this information roads are assumed to be

- a) Elongated regions
- b) Having fixed width with parallel edges
- c) Form topological network.

The third assumption is not handled in this study since the experiments are conducted on limited size images which may not include roads forming a network.

4.2 Road Extraction Model

The general view of the model is shown in Figure 4 below. The model consists of three main modules; road region detection, non-road region detection and score map construction.

This study aims to develop a modular automatic road extraction model (framework) for high spatial and spectral resolution images. We also aim to improve this model so that it can work independent of the spectral characteristics of satellites and sensors. Therefore, this model is mostly based on common structural properties of roads and spectral differences from surrounding objects.

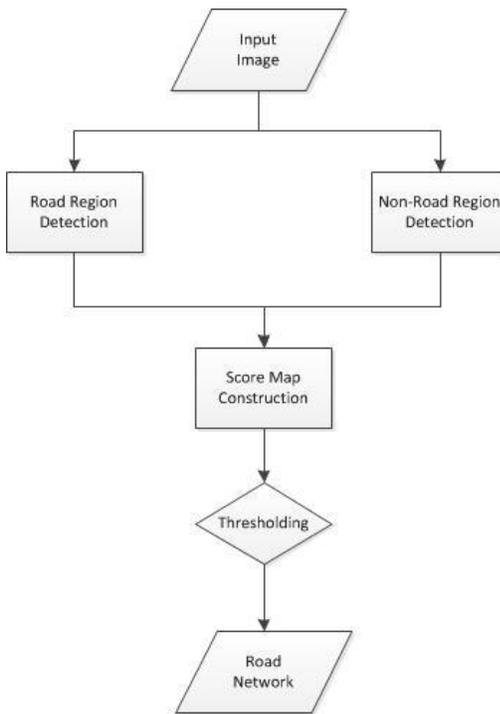


Figure 4 General approach to road extraction

In Figure 5 presented below shows the model in more detail. Both road region and non-road region detection modules can be extended by other approaches. Road region detection module, consist of three sub-modules; edge based, clustering base and structural analysis based road extraction. On the other hand non-road region module includes vegetation detection, structural based non-road region detection and water detection modules. In the score map construction modules, these results are merged and road mask is produced. Each module will be explained in more detail. It is important to note that non-road region detection module is also used as an input to road-region detection modules to reduce the road search region.

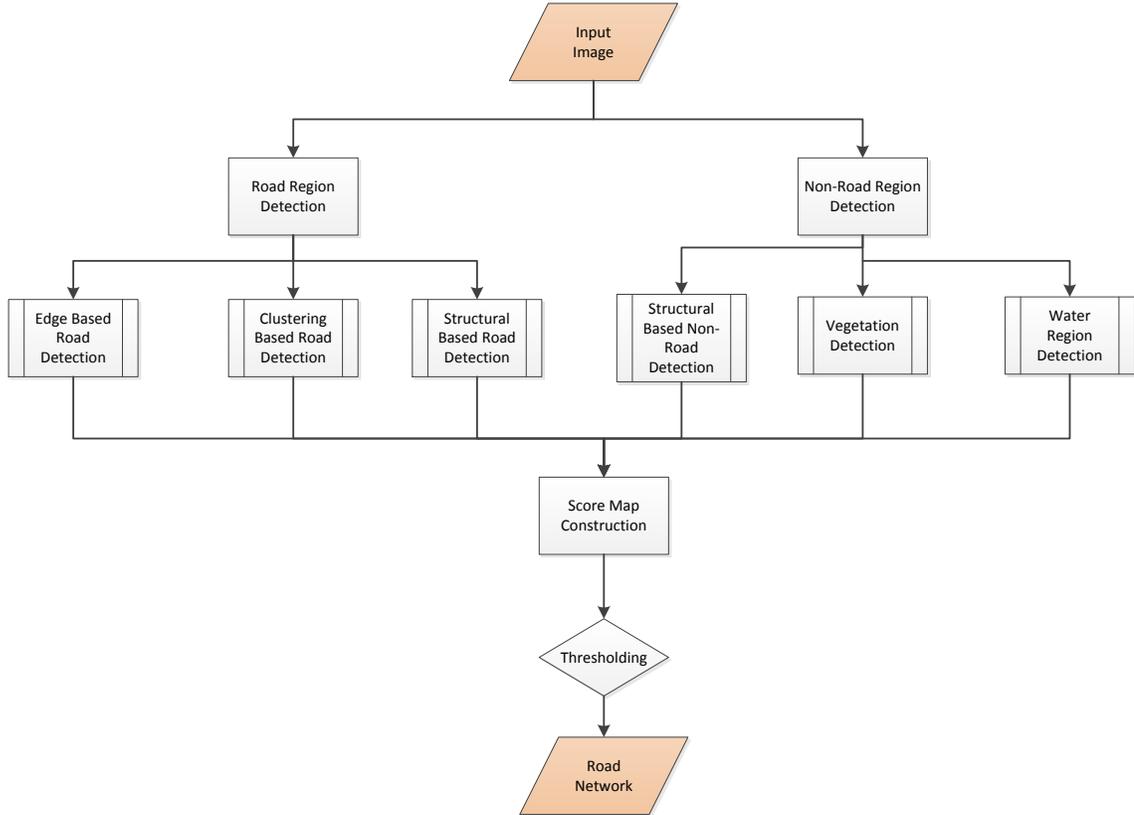


Figure 5 Automatic road extraction algorithm

4.3 Road Region Detection

Based on the assumptions, we attempted to extract roads via tree main modules; edge based road extraction, structural based road extraction and clustering based road extraction. In order to demonstrate the each module's steps; an RGB image and the corresponding ground truth information presented in Figure 6 is used. Performance result will be provided in the Chapter 5.



Figure 6 Sample RGB and Ground Truth Image

4.3.1 Edge-Based Road Extraction

In this part of the study, edge detection based approach will be explained. As a first step of this edge based module, Canny edge detection algorithm is applied on all bands of the data separately. For all resultant edge maps, parallel lines are found to obtain road center lines along with the width information of this line. In order to find parallel line pairs, each edge point pixel is visited and a pair point with in a pre-defined range is searched for. This procedure is applied for both horizontal and vertical directions. If a pair pixel is found in the range, their midpoint is set to as true.

Furthermore, the algorithm continues iteratively for the previously found pair point. That is, for a particular pair point the existence of another pair point in the pre-defined range is also checked, and the mid-point of this new pair is also set to as true.

At the end of this process, vertical and horizontal mid-point masks are obtained. Then, union of these two masks is constructed. Some mid-points are found both from vertical and horizontal procedures. For such cases, middle point width information is calculated as in Equation 4.1

$$1/a^2 + 1/b^2 = 1/w^2 \quad (\text{Equation 4.1})$$

where “a” is the vertical distance, “b” is the horizontal distance and “w” is the distance of the middle point to the edge as shown in the Figure 7.

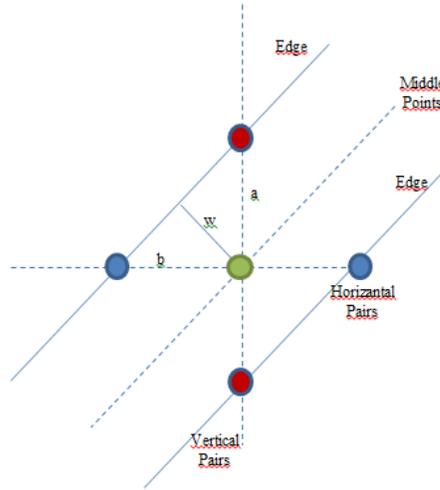


Figure 7 Middle Point's Width Calculation.

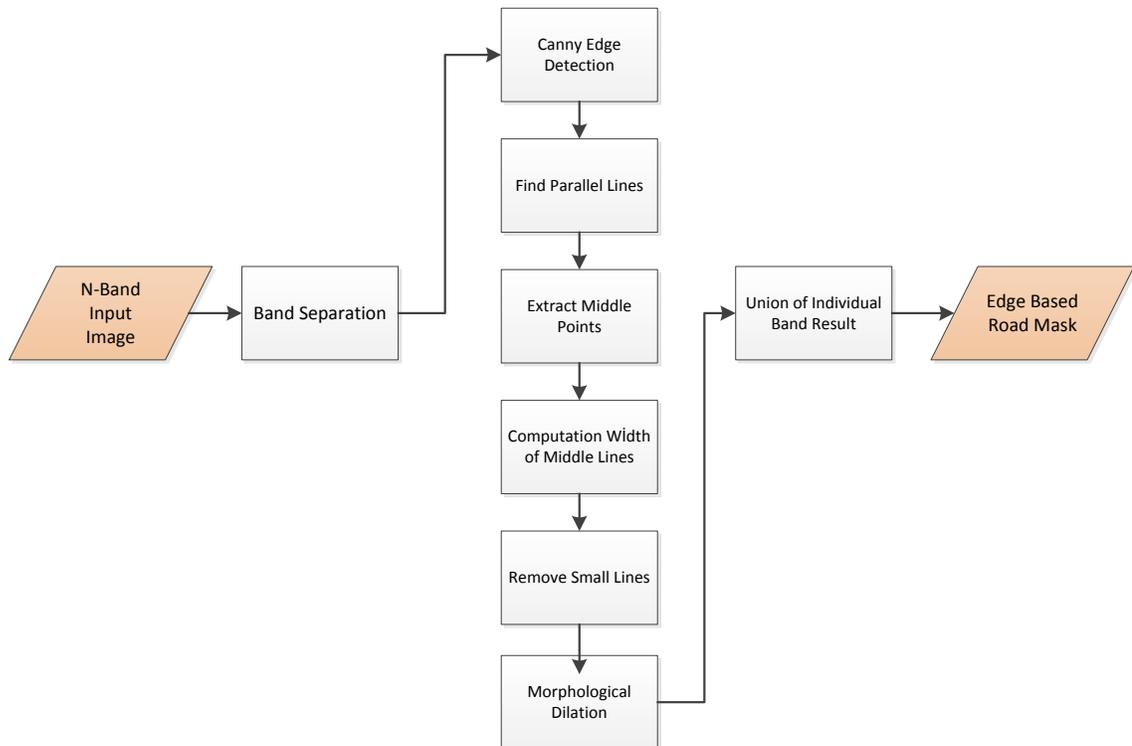


Figure 8 Edge base road detection module.

The middle points from the middle lines of parallel pairs. In order to fill gaps smaller than a certain number of pixels, morphological dilation operation with the disk-structural element with radius “w” is applied. Then, short lines are removed to eliminate parallel edges obtained from non-road structures such as buildings. After the module is applied on all bands individually, the union operation is applied on the results obtained from each band. Edge based road detection module is summarized in Figure 8.

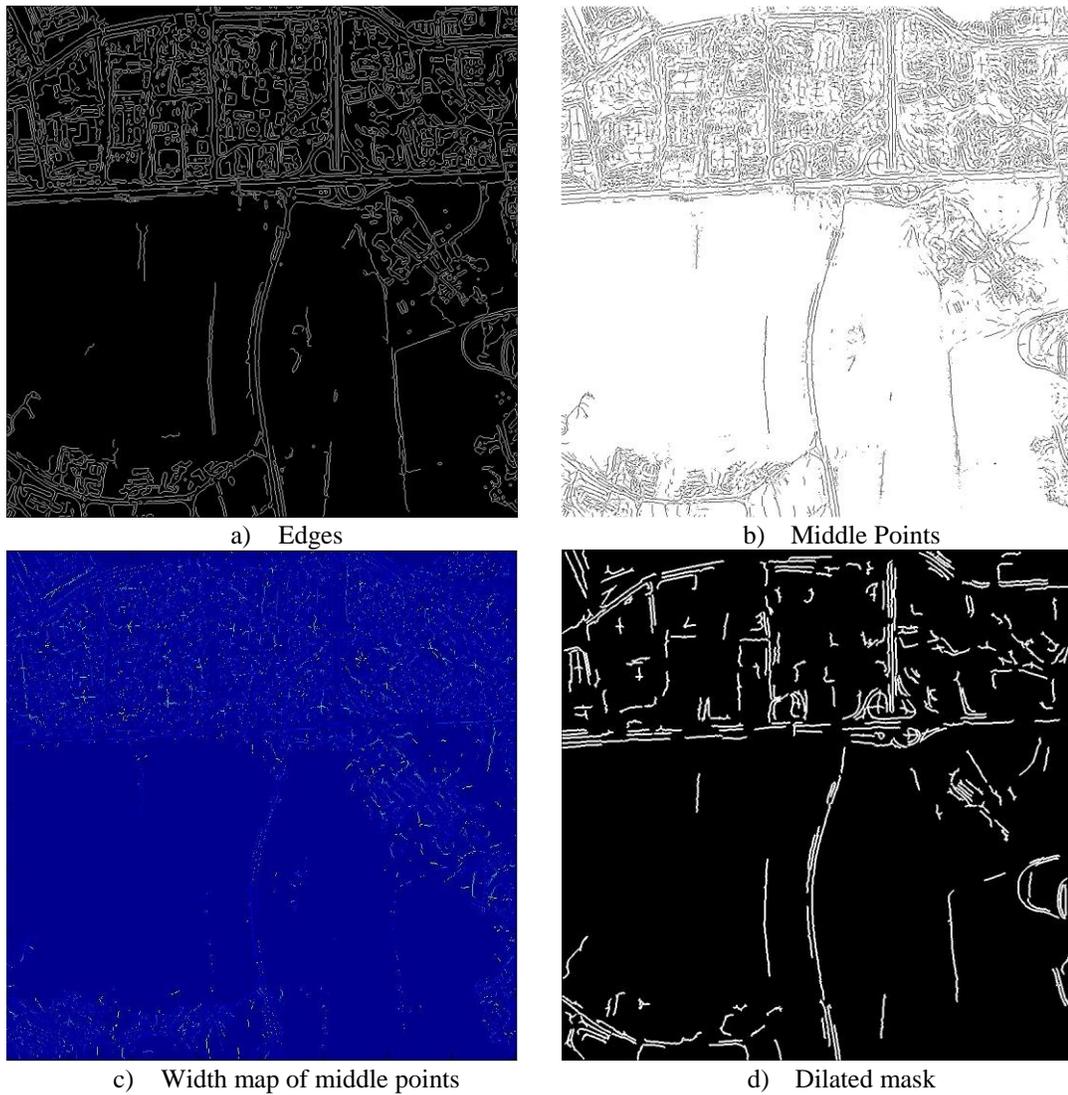


Figure 9 Edge Based Module Result

In Figure 9, sample results for the edge based road extraction are provided. This result corresponds to just one band of a WV-2 image.

4.3.2 Structure Analysis Based Road Detection

In this module, data is segmented by using Mean-Shift algorithm and these segments are analyzed to detect road-like and non-road segments. Regardless of the segmentation method, the segments are classified in to three classes; road segments and non-road segments and unclassified segments. Extracting non-road segments will be described in structure analysis based non-road region extraction module (4.4.1).

In order detect road segments from the segmentation output, each segment is analyzed separately. This stage consists of three steps for which each of the following three criteria are checked

1. Detecting segments having **eccentricity** value greater than a predefined value (defined as 0.99 in this study), or
2. Detecting segments having **elongatedness** greater than a predefined value (defined as 30 in this study),
3. The segment should have large enough **area** (more than 300 pixels in this study).

For the third assumption, each segments area is checked out whether it is greater than pre-defined amount of pixels. This is to avoid small segments that satisfy the conditions which may belong to a building, a tree or another small elongated structure. On the other hand, this may also lead to removal of small road segments. In order to solve this problem, road merging algorithm should be developed as a post-processing module for filling small gaps caused by removing

In order to define discriminative bands, firstly the ground truth of the image is created. Then, all triple and quadruple combination of 8 bands is segmented by using the library developed by EDISON library. (C++ implementation of mean-shift-based image segmentation, a Matlab interface for EDISON) and segmentation results is analyzed based on the verification model stated in Section 4.9. Then structural analysis (Section 3.5) is applied on all segmentation outputs. Precision and recall values are used as the performance indicator (APPENDIX B). For higher precision Coastal Blue, Red Edge, NIR1 and NIR2 bands are used. On the other hand, Blue, Green, Yellow and Red band combination provides higher recall value. Segmentation analysis result is shown in Figure 10. For the WV-2 images, since the all module outputs will be fused into a score map and the unlikely roads will be eliminated, we try to extract road regions

as much as possible at the stage. In order to achieve this goal, the two different four band combination results providing the highest recall and precision are combined by OR operation.

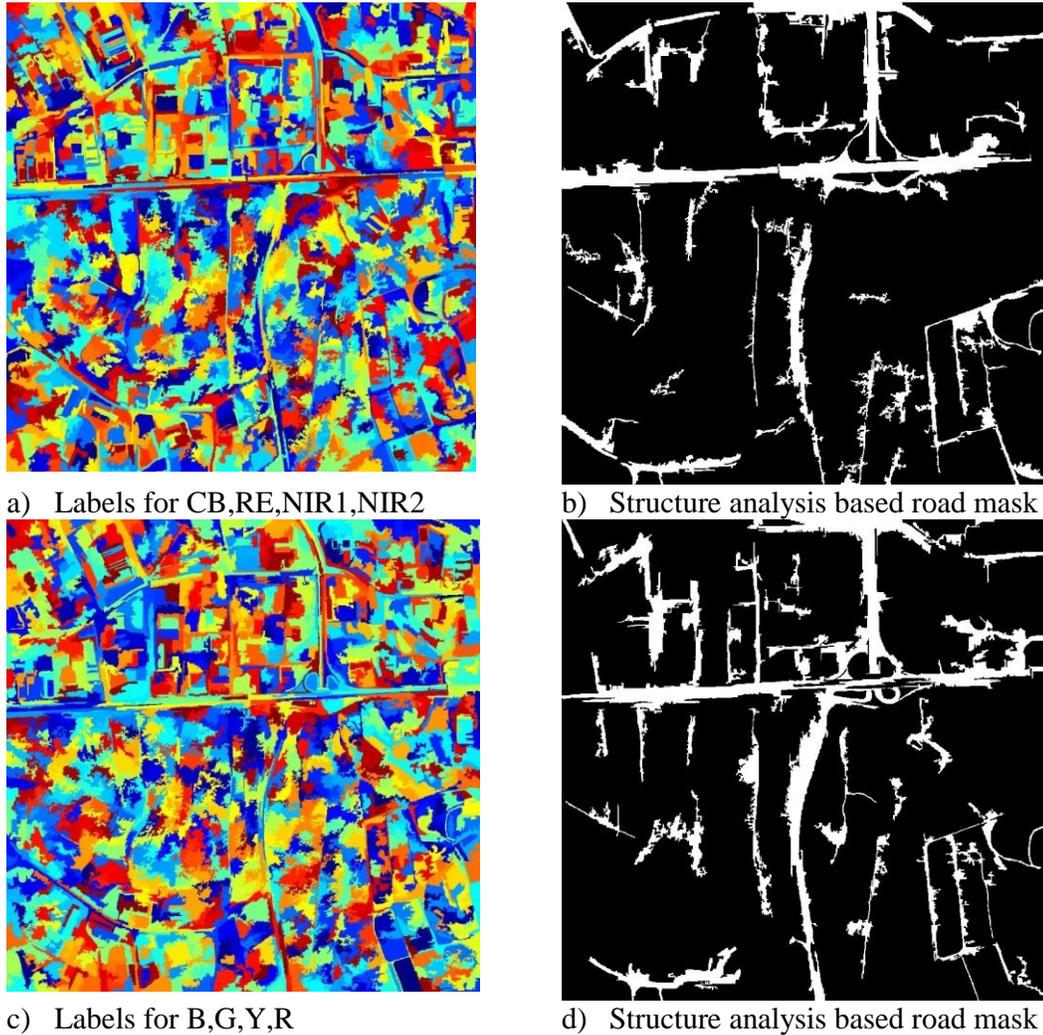


Figure 10 4-Band structure analysis based road extraction result

Briefly, Coastal Blue, Red Edge, NIR1 and NIR2 bands are used for segmentation to achieve higher precision and Blue, Green, Yellow and Red band combination will be used to achieve higher recall.

Structure based analysis road detection module is summarized in the Figure 11

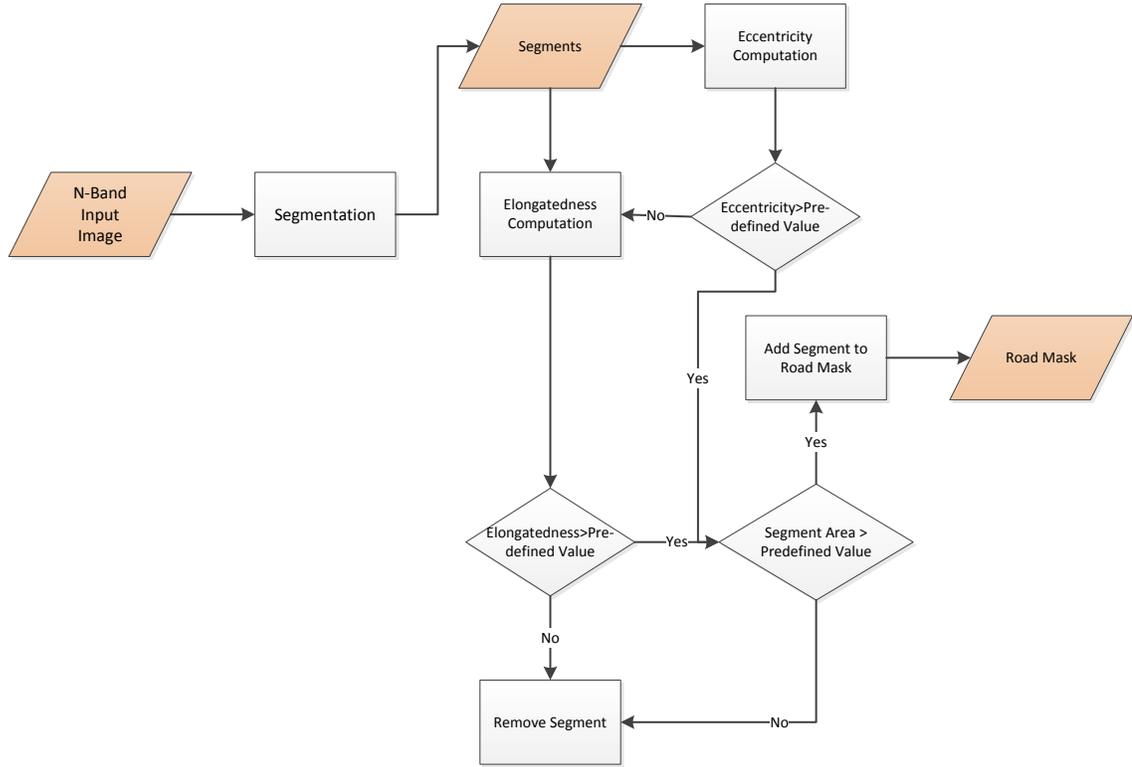
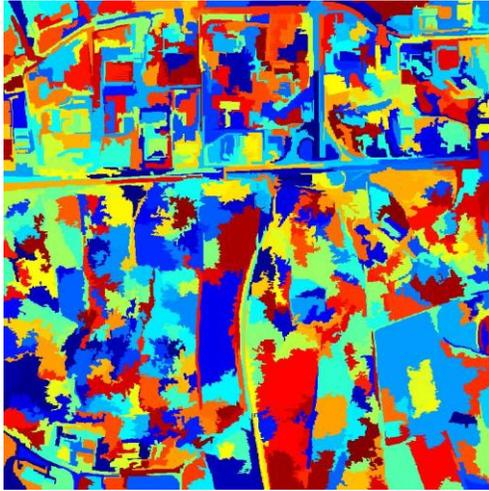


Figure 11 Structural Analysis for Road Detection

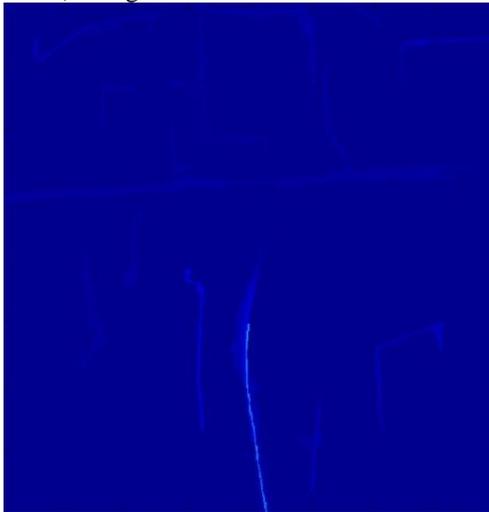
Intermediate results for the structural analysis based road region extraction are provided in Figure 12.



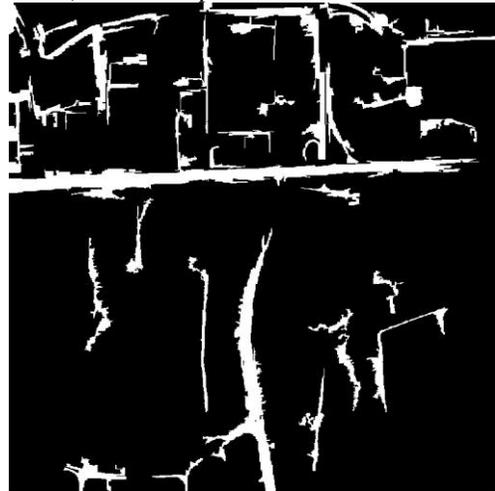
a) Segments



b) Area Map



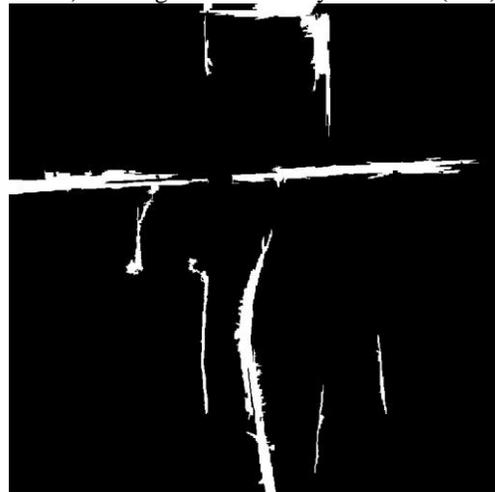
c) Elongatedness Map



d) Elongatedness Analysis Result (>30)



e) Eccentricity Map



f) Eccentricity Analysis Result (>0,99)

Figure 12 Intermediate results for structural analysis module

4.3.3 Clustering Based Road Detection

In this module, K-means clustering algorithm is applied on the image to get road cluster. K-means clustering is used to reduce data into “K” cluster by gathering close mean values as mentioned in the previous chapter. In order to get more accurate cluster result, the vegetation and non-road masks (will be explained in the next section) are excluded from the image. Then, K-means clustering algorithm is applied on the remaining data to obtain the road cluster. For road extraction, in the study of Zhang and Couloigner (2006), K values is suggested to be “six”. After clustering, each cluster is labeled for further analysis. To do this, each pixel of each cluster is labeled based on their connectedness, meaning that connected pixels are grouped under the same label. Similarly, if two structures are identified to be separated by a gap between them, those structures are grouped under distinctive labels. In particular, within this context being connected means being adjacent. Then, each distinct labeled structure in each cluster is checked to define whether it is a road-like component by using the elongatedness assumption technique which is explained in the Structural Analysis Section. Successively, union operation is applied on all results to get the final road mask obtained from the clustering based road detection module. Since, K-means clustering module produces larger blocks or segments than mean-shift segmentation; eccentricity measure is not used in this module. Clustering based road detection module is summarized in Figure 13.

For the some main step of the clustering based road extraction module, intermediate results are provided below in Figure 14.

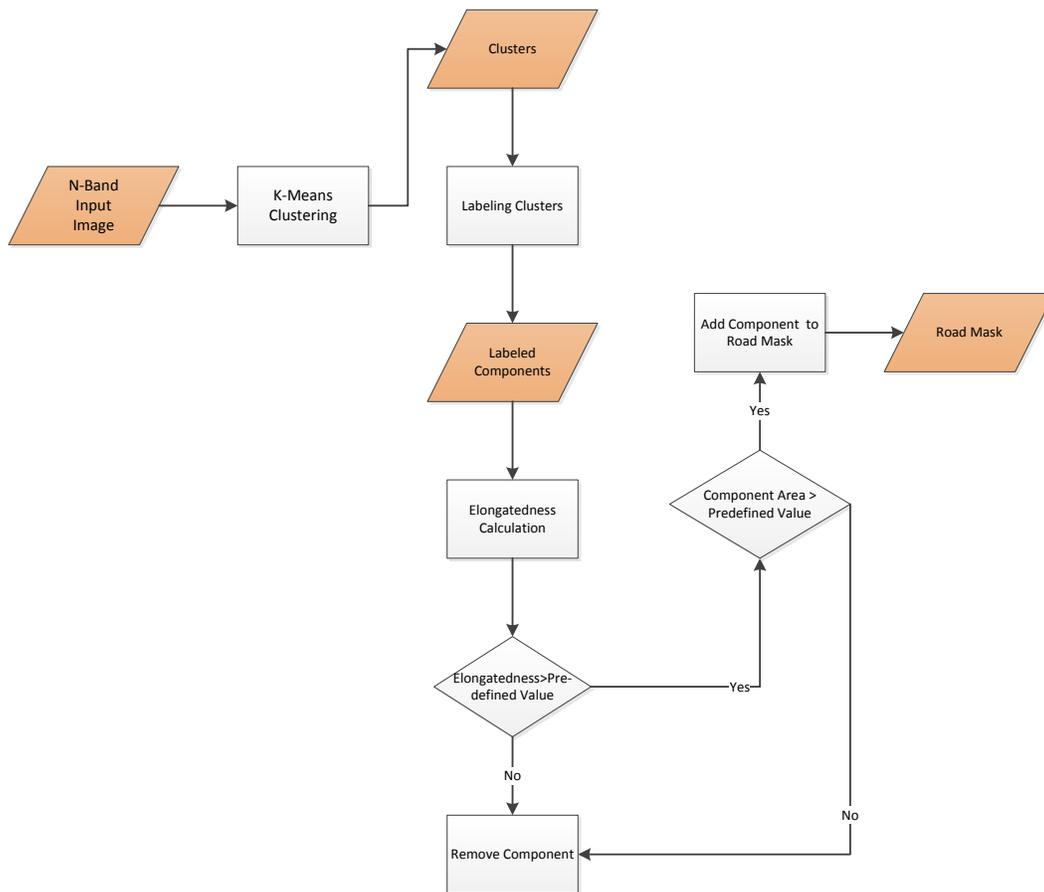
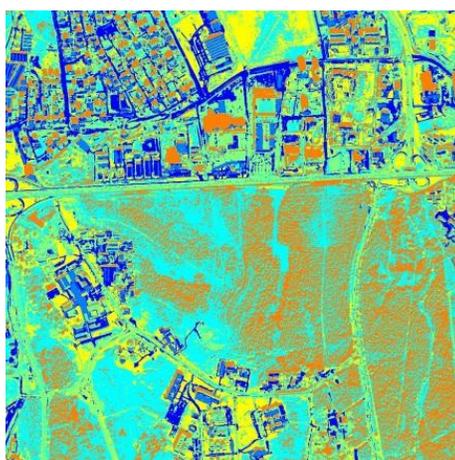
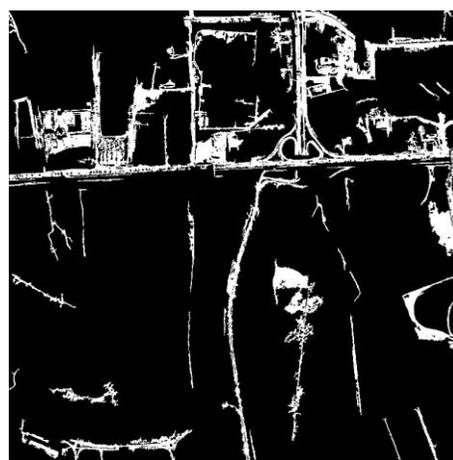


Figure 13 Clustering Based Road Detection



a) Clusters (K=6)



b) Cluster analysis result

Figure 14 Cluster based road extraction intermediate results

4.4 Non-Road Region Detection

In this road extraction model, extraction of non-road region is important not only to verify road extraction result but also to exclude such regions from the image to be processed. In this part of the study, vegetation extraction, water region detection and structural analysis based non-road region detection will be introduced.

4.4.1 Structure Analysis Based Non-Road Region Detection

As it is mentioned in the structure analysis based road detection module, first the data is segmented and then non-road segments are extracted. In order to do this, it is assumed that non-road regions are large and compact. For this purpose, each segment is examined to see whether its “thickness” (t) value and its area (A) are greater than 50 and 600, respectively. The thickness value is computed by using Equation 4.2.

$$t = \frac{A}{a * (2 - Extent)} \quad \text{Equation (4.2)}$$

It is important to note that, while “elongatedness” measure is used for segment flatness calculation, “thickness” measure is used to measure the width of the segment which represents a segment’s fleshiness. Non-road region detection is summarized in Figure 15.

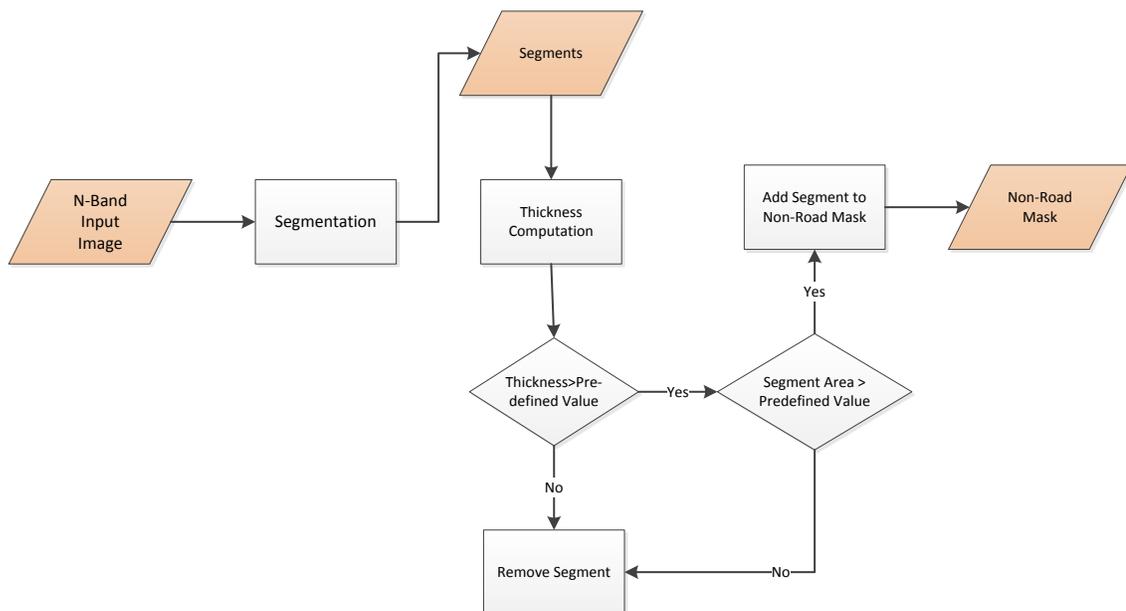


Figure 15 Structural analysis based non-road region detection flow chart

Intermediate result for the non-road region detection is show in Figure 16.

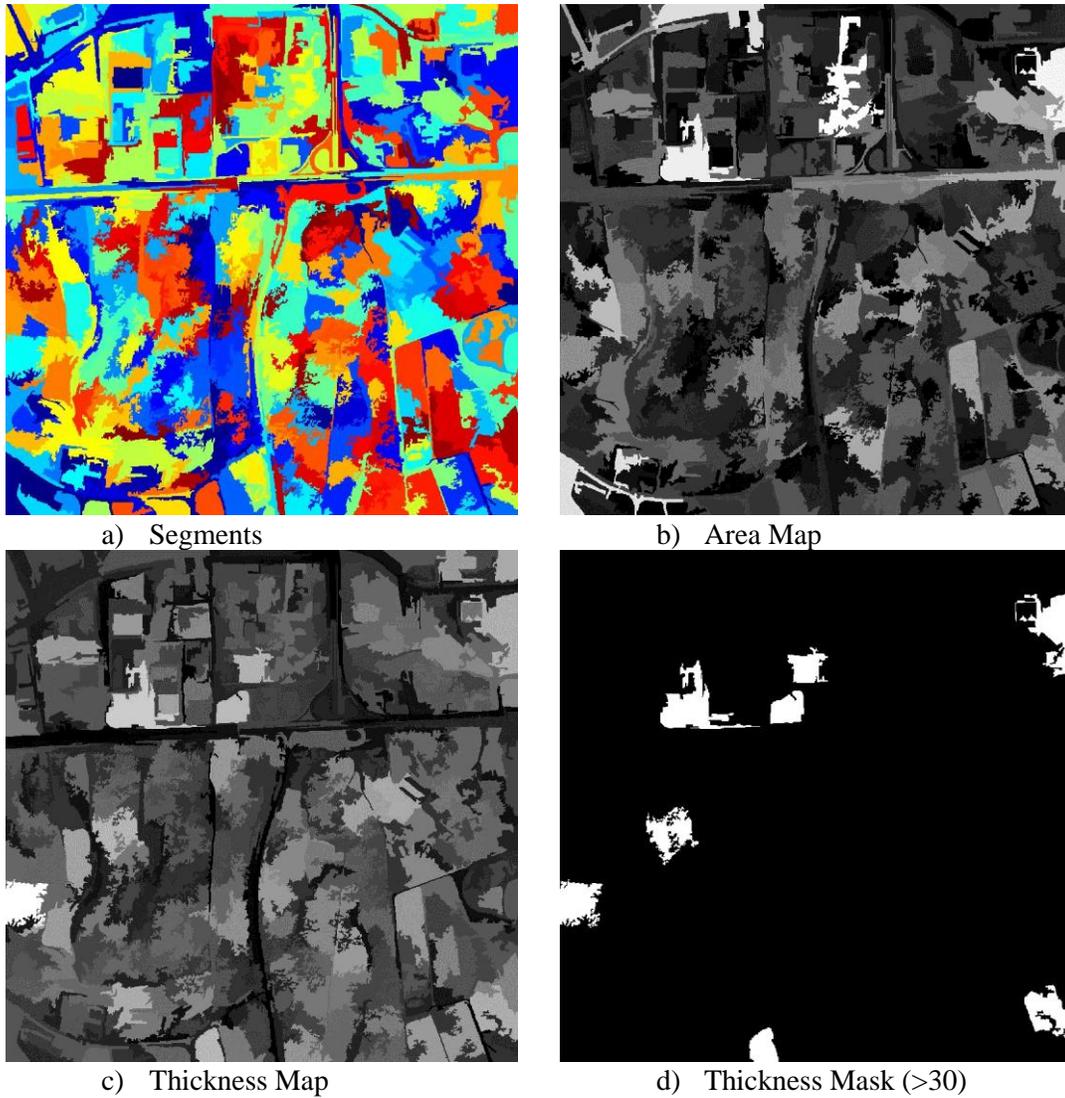


Figure 16 Structural analysis based non-road region samples

4.4.2 Vegetation Detection

Removing vegetative regions from the image may also improve classification results. In road extraction algorithm, it is also important to determine vegetative fields to get more accurate road regions. As mentioned in Chapter 3, commonly used broadband vegetation measures for this purpose are NDVI (Normalized Difference Vegetation Index), SAVI (Soil-Adjusted

Vegetation Index), EVI (Enhanced Vegetation Index), SRI (Simple Ratio Index), ARVI (Atmospherically Resistant Vegetation Index), SGI (Sum Green Index)¹.

Briefly, $NDVI = \frac{NIR-Red}{NIR+Red}$, which ranges between -1.0 and +1.0. The greater the NDVI value of a field, the more probable for it to be vegetative field. The data used in this study, WV-2, provides two different NIR bands: i.e. NIR1 (770-895 nm) and NIR2 (860-1040 nm). Similarly there are two different Red bands i.e. Red (630-690 nm), Red-Edge (705-745 nm). For the vegetation analysis in such images, NDVI index is calculated by $NDVI = \frac{NIR2-Red}{NIR2+Red}$, equation as suggested by Wolf (2010). SAVI or other broadband vegetation detection indices can be used as well. After applying vegetation indices, results are thresholded by the predefined value and the vegetation mask is obtained. The vegetation extraction module is summarized in Figure 17, sample result for WV-2 image is shown in Figure 18.

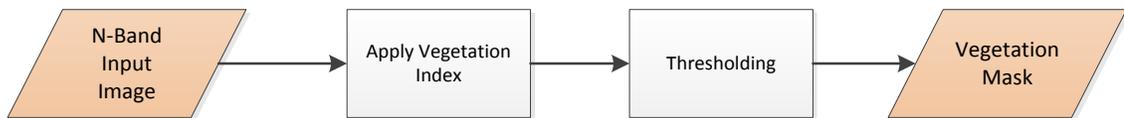


Figure 17 Vegetation detection flow chart

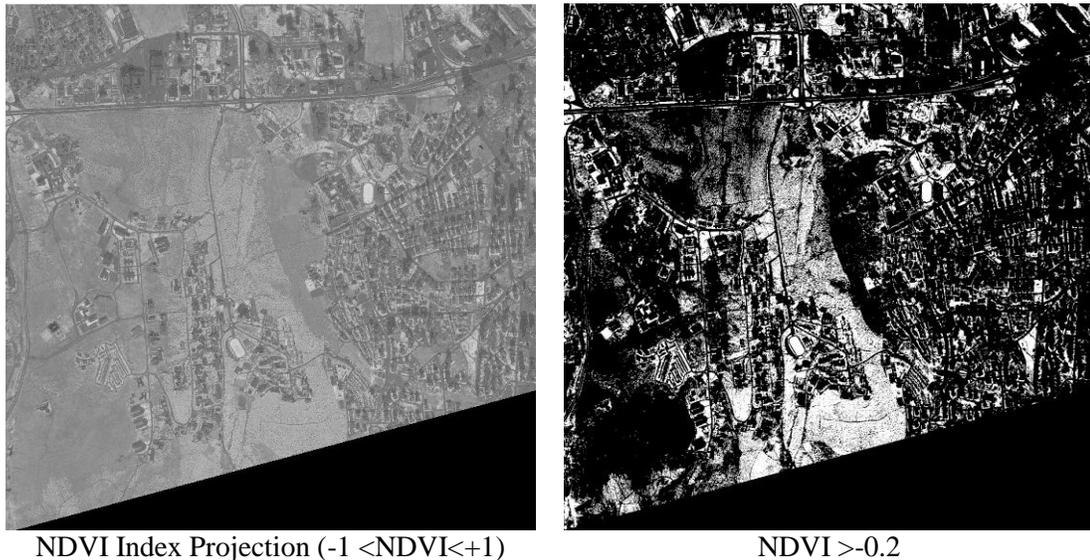


Figure 18 Vegetation extraction result

¹ (http://geol.hu/data/online_help/Vegetation_Indices.html).

4.4.3 Water Region Detection

Water region extraction helps us to define non-road regions. It is known that water regions can be detected by using NIR bands because NIR band's reflection is lower than of visible bands (Zhao et al., 2009). To find water regions, NDWI index is used as well.

$$NDWI = \frac{Green(G)-NIR}{Green(G)+NIR} \quad (\text{Equation 4.3})$$

This formula is used for 4-band images. For the WV-2 images, NDWI can be calculated as follows (Wolf, A., 2010).

$$NDWI = \frac{Coastal(C)-NIR2}{Coastal(C)+NIR2} \quad (\text{Equation 4.4})$$

On the other hand, NDVI value is also used to determine the water regions. It is reported that a smaller NDVI value for the region under investigation will indicate that this region would be a water region rather than a vegetation region (Weier & Herring, 2012). In particular, if NDVI values for a region approaches to -1, then our confidence in classifying that region as water increases. In order to merge these cues for water extraction, the flow chart in Figure 19 is followed.

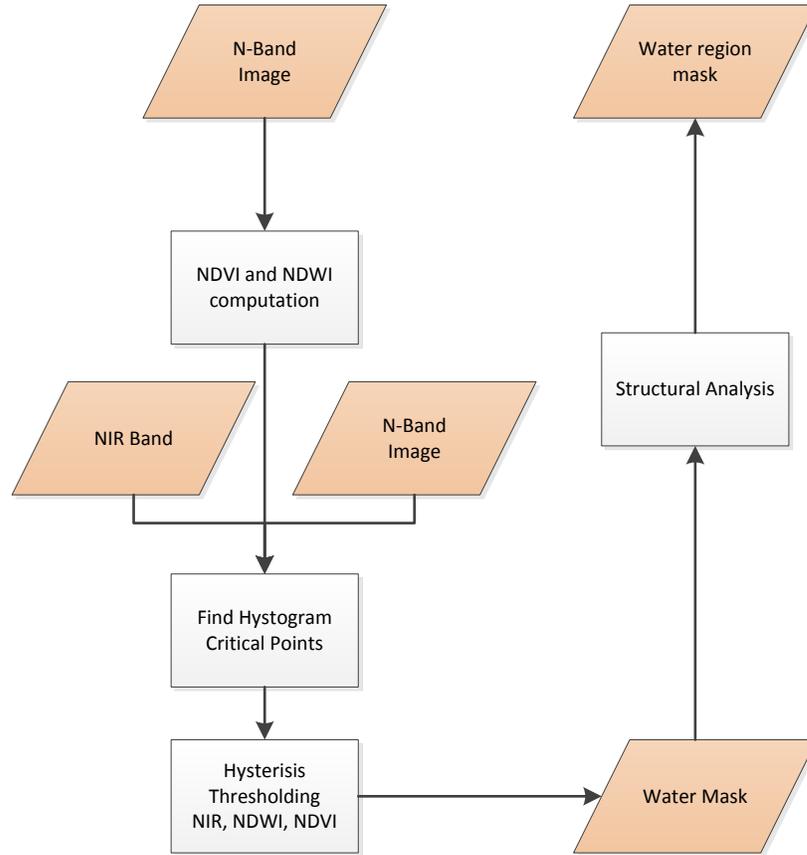


Figure 19 Water region detection

Intermediate thresholded and final results for water extraction are shown in Figure 20. After thresholding the NIR2, NDVI and NDWI values, the union operation is applied to results. Removing connected components having area smaller than a predefined value (500 pixels in this study) provides the final result (Figure 20i).



a) RGB image



b) RGB and Water Region



c) NIR2



d) NIR2 Thresholded



e) NDVI



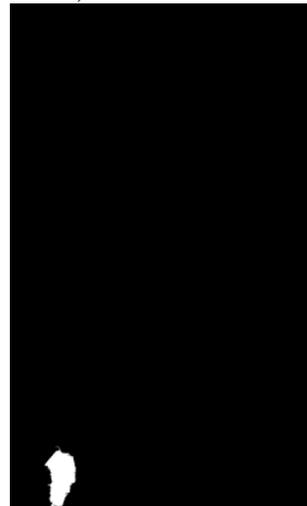
f) NDVI Thresholded



g) NDWI



h) NDVI Thresholded



i) Final Water Mask

Figure 20 Water extraction intermediate result

4.5 Score Map Construction

We developed a score map construction technique in order to fuse the information provided by road and non-road masks. Presently, we have six result masks; edge-based road mask, segment based road mask clustering based road mask, structure analysis based non-road mask, vegetation mask and water mask. Before fusing the masks, unconnected small parts (smaller than 100 pixels) in each mask are removed. The main aim of the score map construction stage is to combine and synthesize the outputs of each module into more convenient form for road extraction. We have two different approaches to create a score map. First one is based on segments' elongatedness values and second approach is based on statistical analysis of contribution of each module on the result.

4.5.1 Score Map Construction Based on Segments

Road masks yielded by the middle point, clustering and structural analysis are re-evaluated in accordance with the segmentation results, while the non-road regions such as vegetation and lands are excluded from the final score map. That is, the road masks are scored based on the elongatedness of the segments that they intersect with. The output label image of the mean-shift procedure is employed to evaluate road masks.

Given a label image obtained by using a segmentation algorithm $L(x, y)$, a set of road masks $R = \{R_1, R_2, \dots, R_n\}$ and non-road masks $N = \{N_1, N_2, \dots, N_m\}$, we can formally express our score map construction method in two steps:

Non-road Mask Exclusion: The pixels which are set to one in non-road mask are simply excluded from the score map P' by adjusting corresponding pixel locations' values to zero.

$$P'(x, y) = \begin{cases} 1 & \text{if } N(x, y) = 0 \\ 0 & \text{otherwise} \end{cases} \quad (\text{Equation 4.5})$$

Road Mask Scoring: Firstly, the road mask is scored by intersecting the road mask with the label mask. That is, each intersected segment's elongatedness value is computed, and the final score map is obtained by taking the maximum pixel value among the road masks. The segments to be scored are extracted by a series of set operators as follows;

$$I_j = ((L = j) \cap R) \quad (\text{Equation 4.6})$$

where I_j is the segment with label number of j , and R is an arbitrary road mask. Then, the extracted segment is scored with the elongatedness function defined before;

$$C(I_j) = \text{elongatedness}(R(I_j)) \quad (\text{Equation 4.7})$$

The final score map is obtained by taking the maximum values from the road masks to the corresponding pixel locations;

$$P(x, y) = \max_{(x,y)}(P'(x, y), C(x, y)) \quad (\text{Equation 4.8})$$

The score map construction technique suggested in this study considers the local structural characteristics of the road masks yielded by different modules. Therefore, the performances of individual road extraction modules are boosted.

Segment based score map and thresholded mask is shown below in Figure 21.

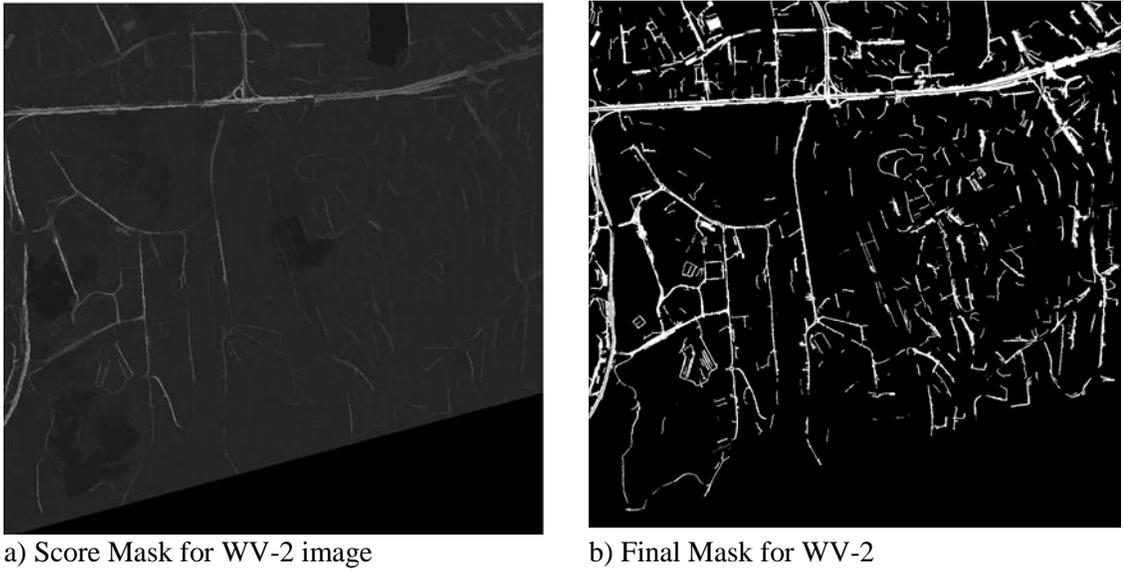


Figure 21 Segment based score map and thresholded mask

4.5.2 Statistical Analysis Based Score Map Construction

In the study of Çinar (2012), a similar score map construction method for a multi-resolution mean-shift segmentation based road extraction algorithm is proposed. The main idea of the method is to find out the optimal parameter sets for each mean-shift segmentation instance

operating on different resolution level. For this purpose, each segment is evaluated using the elongatedness descriptor equation same as our study. From the optimization set, class separation scores are calculated for two classes; road and non-road. The author suggests using a fisher like discrimination equation, which is elaborated in Equation 4.9.

$$d = \frac{\mu_r - \mu_n}{\sigma_r + \sigma_n} \quad (\text{Equation 4.9})$$

Where μ_r and μ_n are means, and σ_r and σ_n are standard deviations of the road and non-road region's distributions respectively.

4.6 Water Canal Extraction

As mentioned in the section 3.1.2.3, water regions are extracted from the image by using NIR, NDWI and NDVI values. While NIR and NDVI values are relatively smaller for water regions, NDWI value are relatively higher. The proposed method aims to detect threshold values automatically. In order to determine threshold values automatically, critical points are extracted by calculating local minima and maxima (Figure 22). For NIR and NDVI values, two local minima and for NDWI two local maxima points are selected to be used as the low and high threshold values for hysteresis thresholding.

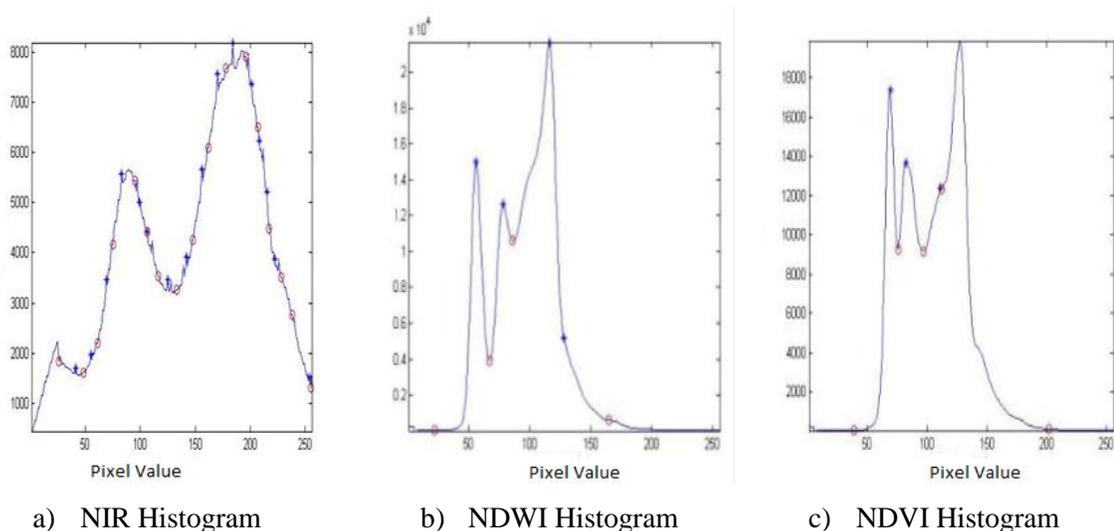


Figure 22 Automatic threshold computation

After thresholding, three different water masks are obtained. These masks are fused by AND operation. This final binary mask is labeled. Each component is checked to define whether it is a road-like component or not by using the elongatedness property explained in the Section 4.3.2. Elongated water regions are assumed to be water canals. As the last step, small components are removed to obtain a clear mask.

Flow chart of the water canal extraction is presented in Figure 23.

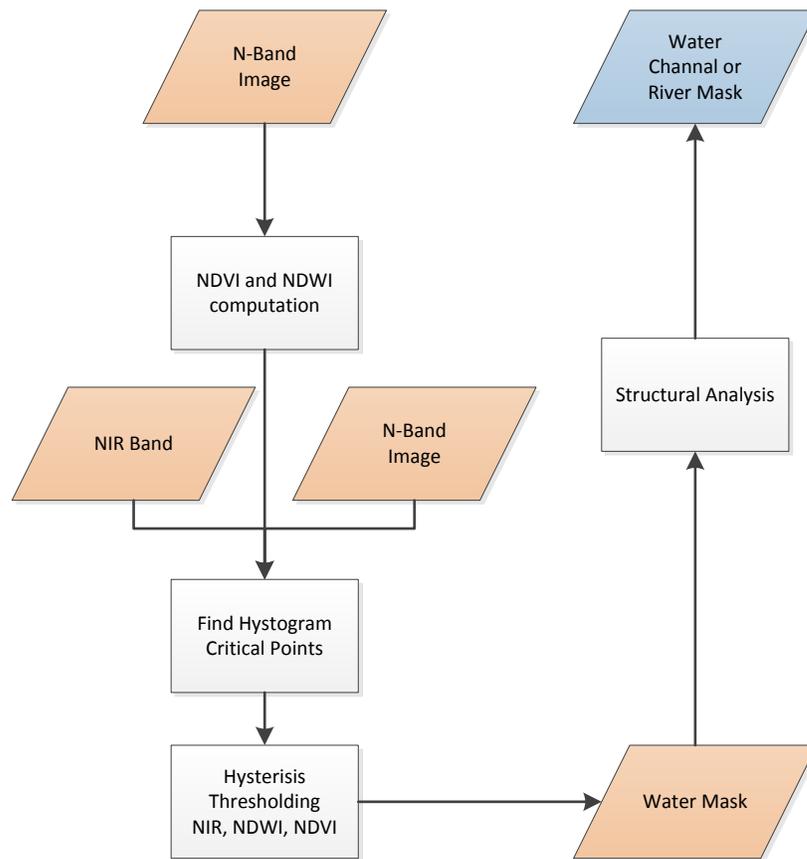
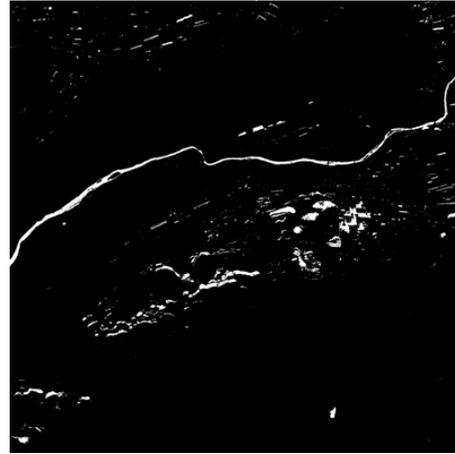


Figure 23 Water region extraction flowchart

Intermediate results for the water canal extraction are shown in Figure 24.



a) NDWI



b) NDWI Water Mask



c) NDVI



d) NDVI Water Mask



e) NIR Band



f) NIR Water Mask

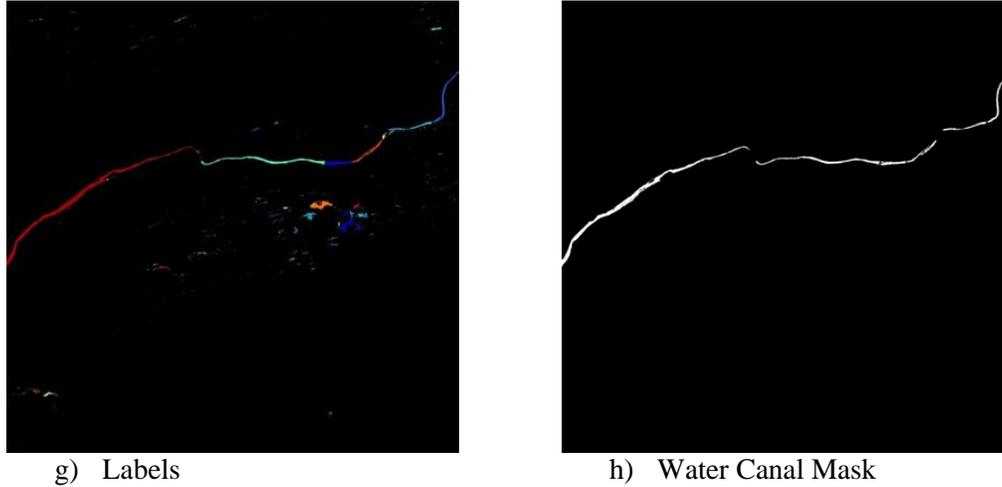


Figure 24 Intermediate results of water canal extraction

4.7 Railroad Extraction

In this part of the study, we treat railroad and road discrimination as a two-class classification problem. For a given road mask including roads and railroads, we strive to distinguish the regions belonging to railroads. The extraction of the initial road mask is achieved via the proposed method. Railroads on satellite images are often extracted as road objects due to their common properties with roads such as elongated and curved structure and uniform intensity. In high resolution aerial images, however, the sleepers between railroad tracks depict a distinctive texture from other elongated objects. Therefore, this additional cue can be employed to distinguish railroad objects in aerial images. The sleepers are perpendicular to the direction of the railroads and they are spatially periodic. The distance between railroad sleepers is standard for a given country. We apply Fourier transform to overlapping windows of the original image and check the presence of spectral peaks corresponding to the spacing between sleepers. These peaks have fixed distance from the origin and are spaced perpendicular to the burst of energy due to the roadsides. Synthetic images are used to analyze the railroad classification method.

Figure 25(a) and (b) depict a synthetic drive road segment and its Fourier Transform with suppressed DC component, respectively. Figure 25 (b) and (d), the burst of energy along the horizontal axes corresponds to the roadsides in the input images (a) and (c). In Figure 25 (c) and (d) depict a portion of the image containing a synthetic railroad portion and the corresponding

Fourier Transform with suppressed DC component, respectively. The vertical series of white dots in Figure 25 (d) corresponds to the sleepers' periodic structure which is perpendicular to the roadsides in Figure 25 (c). As the distance between sleepers, λ_0 , is constant for a given country, the location of the corresponding spectral peaks at the frequency domain can be computed provided $\lambda_0 > 2\lambda_s$ where λ_s is the spatial image resolution. Hence, standard spectral analysis techniques can be used to detect such peaks. For an $N \times N$ window, the $N \times N$ DFT is computed by

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) e^{2\pi i(xu+yv)/N}. \quad (\text{Equation 4.10})$$

and the spectral peaks at λ_0 depict themselves at a distance

$$r = \frac{\lambda_s}{\lambda_0} N \quad (\text{Equation 4.11})$$

from the origin.

Furthermore, the harmonic components of the principal frequency in Equation 4.11 are located at integer multiples of r .

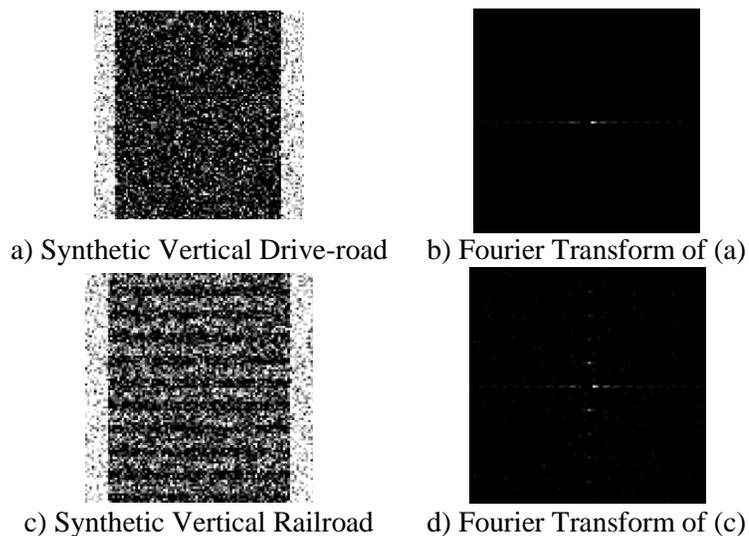


Figure 25 Sample synthetic drive road and railroad images and their Fourier transforms.

The proposed methods rely on detecting this main visual distinction of railroad and drive-road spectra. There are two major cues to detect periodicity in Fourier domain. First, the energy

generated due to sleepers' periodic arrangement in the spatial domain is expected to exist at a fixed distance from the origin as shown in Equation (4.11). Therefore, a circular search region can be employed to seek for energy peaks. Second major cue is the orientation of these peaks with respect to the road direction. This is critical because some of the peaks found in the circular search region may belong to road regions which do not pose periodic behavior in the spatial domain. While "Ring Approach" uses first cue, "Perpendicular Axis Approach" is developed to use the second cue. In the following, these two approaches will be explained.

The Ring Approach:

In this approach the Fourier transform of the image is analyzed with a window including a circular search region as shown in Figure 26(c). The energy peaks in the Fourier domain are investigated by calculating the total energy under the ring (Karaman, Çinar, Gedik, Yardımcı and Halıcı, 2012). The radius of the ring is determined as mentioned in Equation (4.11). For the sake of simplicity, synthetic image will be used as the sample images for demonstrating the approach. The synthetic 100X100 railroad image consists of 10 pixels periodic structures, and 10 pixel road side width. The total energy in the Fourier transform intersecting with the ring(s) is calculated. It is observed that this total energy for railroad spectrum is greater than the one for drive-roads.

In Figure 26, synthetic railroad and drive-roads images and their corresponding steps is shown. For the analysis of harmonic components of the principle frequency, double rings are also used for the region of interest.

Total and average energy for one-ring and double ring region of interest is shown in Figure 26. In the study Karaman et al. (2012), instead of the difference between total and average spectral energy of drive-road and railroad spectra, local peaks extracted by gray level dilation proposed in Heijmans (1991) have been used.

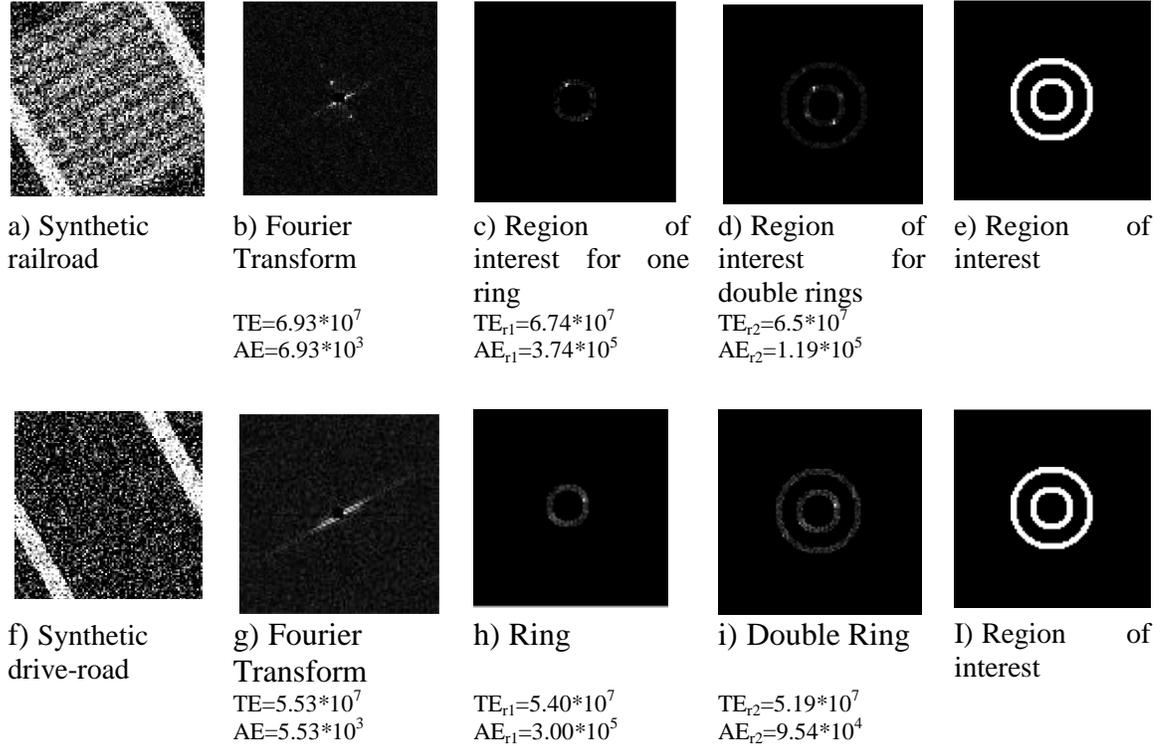


Figure 26 The Ring approaches result for synthetic data

Perpendicular Axis Approach:

In this approach, the orientation of the road should be determined initially. The orientation of the road is determined by the help of the linear structural element in the same window size as show in Figure 27 (b). This linear element is incrementally rotated around the origin and the total energy in Fourier transform intersecting with the line is calculated at each step. It is assumed that the orientation of the maximum energy line corresponds to the roadside. Sample railroad synthetic image and direction of orientation of the maximum energy line is shown in Figure 27.

The next step is determining whether this road is a railroad or drive-road. For this purpose, the total energy along the perpendicular direction of the orientation of the maximum energy line is computed. If the total energy under the perpendicular direction of the maximum energy line is higher than a pre-defined threshold, the region classified as railroad. The sample results for the perpendicular axis approach are shown in Table 1.

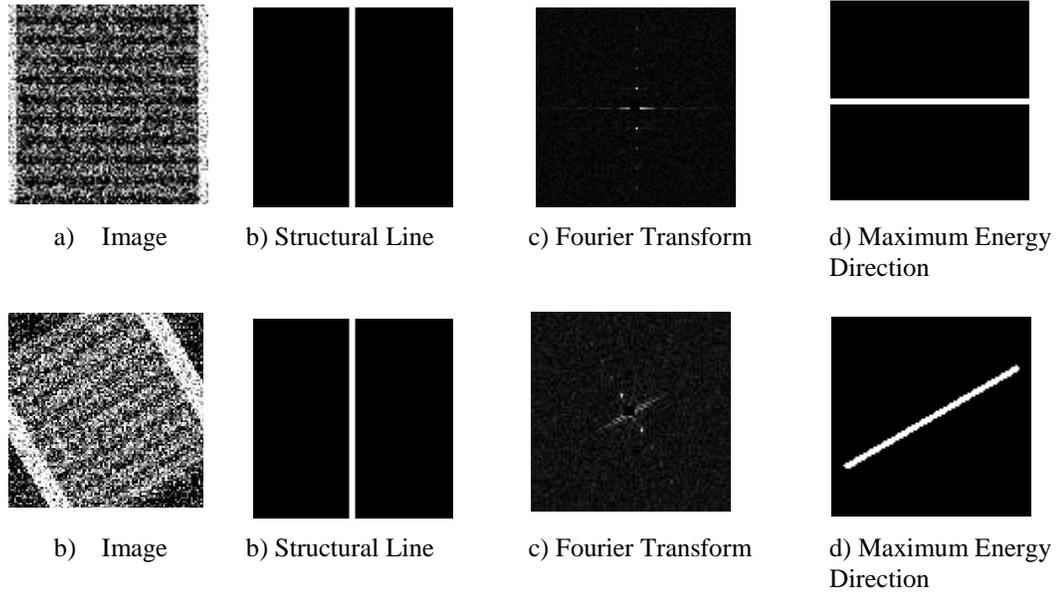
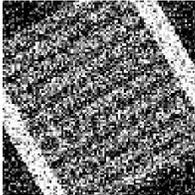
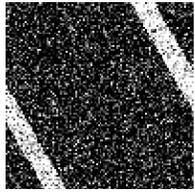


Figure 27 Sample railroad synthetic images and Perpendicular approaches elements

Table 1 Sample result for perpendicular axis approach.

	Image	Fourier Transform	High Energy Regions	Total Energy of Sleepers
Railroad				$2,82 \cdot 10^6$
Drive-road				$1,52 \cdot 10^6$

As seen in Table 1, the road samples having the same orientation and width are used. It is observed that the total energy calculated for the drive road is less than the one for the railroad as proposed by this approach.

Combination of Adaptive Ring and Perpendicular Axis

The two cues, i) energy peaks exist in a fixed distance from the origin, ii) the orientation of these energy peaks is perpendicular to the origin can be used simultaneously. In order to use both of these cues, ring and perpendicular axis approaches are fused. To define region of interest for searching the energy peaks in Fourier domain, the rings used in ring approach and the line used perpendicular axis approach are fused by AND operation after applying small dilation operation to make the region of interest more inclusive. The total energy in Fourier transform intersecting with the region of interest is computed. In Figure 28, synthetic railroad and drive roads and their corresponding steps are shown.

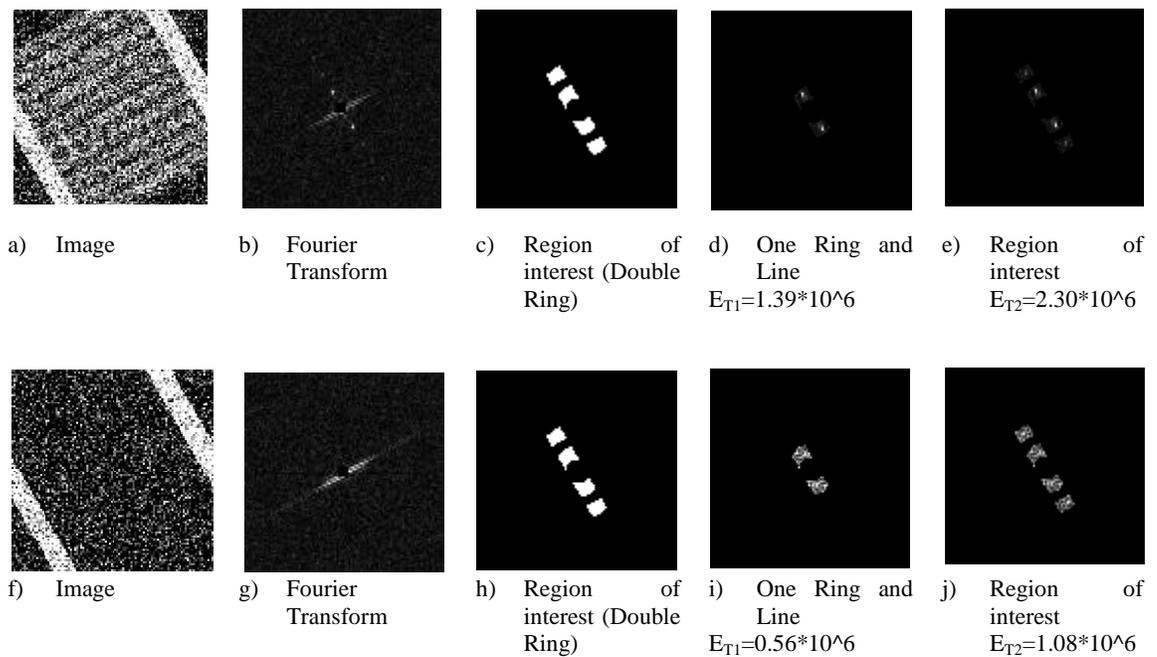


Figure 28 Sample result for combination of perpendicular axis and ring method

In Table 2, comparison of adaptive ring, perpendicular axis and their combination is presented.

Table 2 Performance comparison of ring, perpendicular axis and fused method

	Railroad	Drive-road	Difference	Ratio
Perpendicular Axis	2,82	1,52	1,30	1,86
Single Ring	2,17	1,51	0,66	1,44
Double Rings	4,92	3,73	1,19	1,32
Fused Method (1-Ring)	1,39	0,56	0,83	2,48
Fused Method (2-Rings)	2,30	1,08	1,22	2,13

As shown in Table 2, in terms of difference, perpendicular axis approach provides more distinguishable values. On the other hand, when the fused method is applied with one-ring, it provides more distinguishable result in terms of ratio of the energies of railroad and drive-road. We decided to apply fused method with one ring to real data.

The method is also tested for the multiple synthetic drive roads and railroads. It can be seen that, the energy peaks due to sleepers' periodicity in Fourier spectrum still exist at a fixed distance and orientation from the origin. Sample adjacent roads and their Fourier transforms are shown in Figure 29.

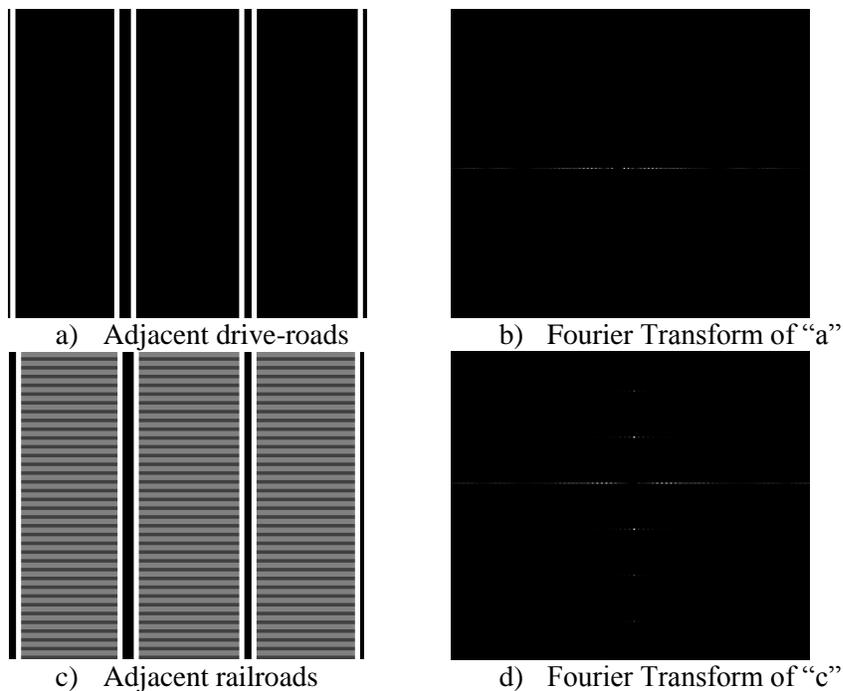


Figure 29 Adjacent roads and their Fourier transforms

Extension of the railroad classification method on real images

The main purpose of this study is to apply these methods on the real remotely sensed image. A set of helicopter-borne aerial images with 20 cm resolution obtained from Google Earth is used as real data. Ground truths including both roads and railroads are prepared manually as shown in Figure 30. White regions represent drive-roads while red pixels depict railroad regions. Union of those white and red masks will be processed to distinguish the railroads in the road network.

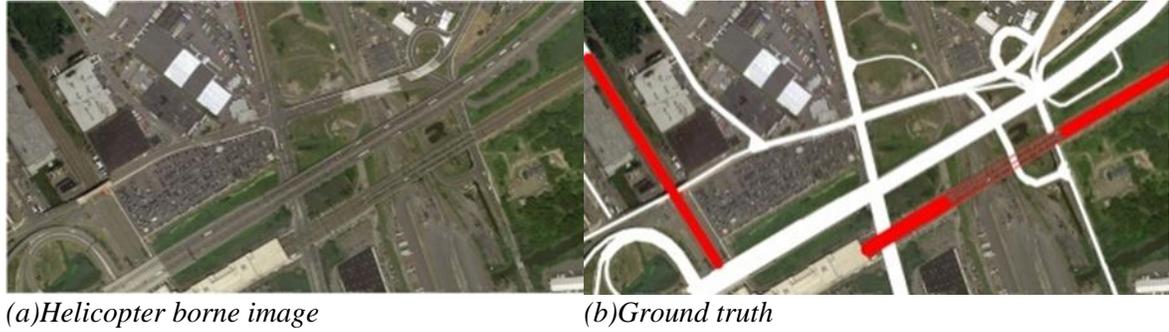


Figure 30 Sample image for railroad classification approaches

The proposed railroad classification method is applied on the real images as presented in Figure 31.

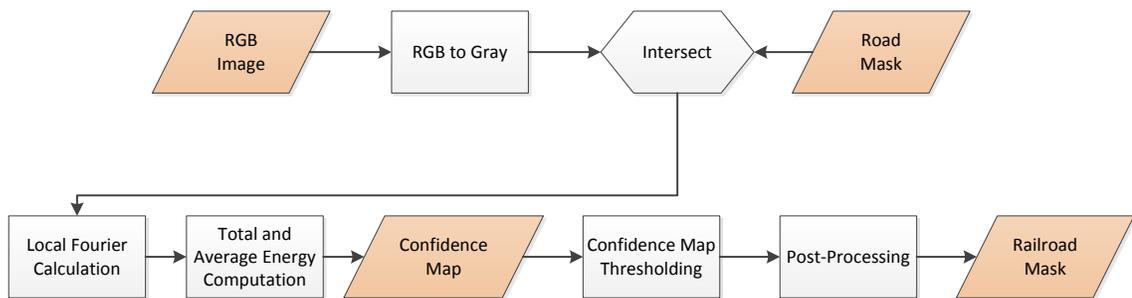


Figure 31 The railroad classification approach flowchart

After RGB image converted in to gray-level image, the road mask to be classified as railroad and drive-road is used to mask the parts that have roads. Then, the fused method of the single ring and perpendicular axis approaches is applied on the masked image. This procedure is employed over the whole image with nonoverlapping NXN windows.

Sample drive-road and railroad window and corresponding Fourier transforms are shown in Figure 32.

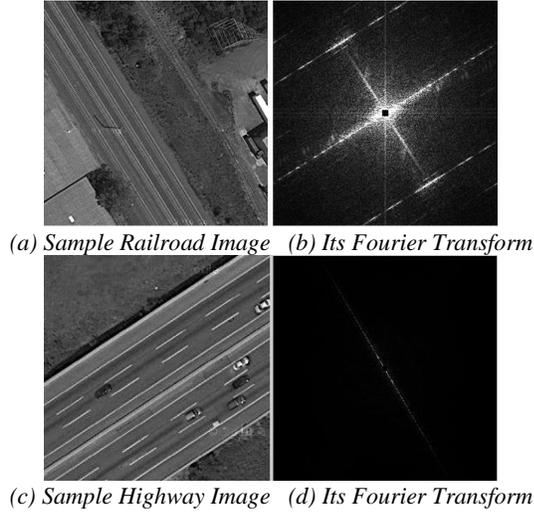
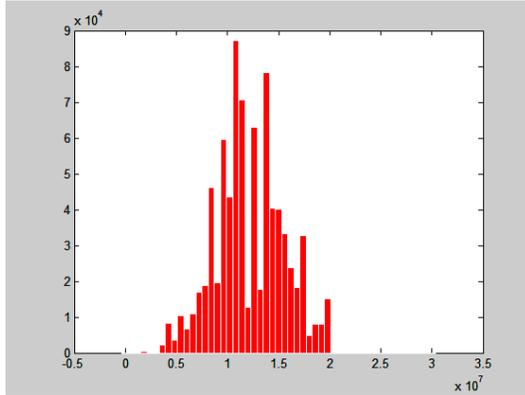


Figure 32 Sample Railroad and Highway Images and their Fourier Transforms

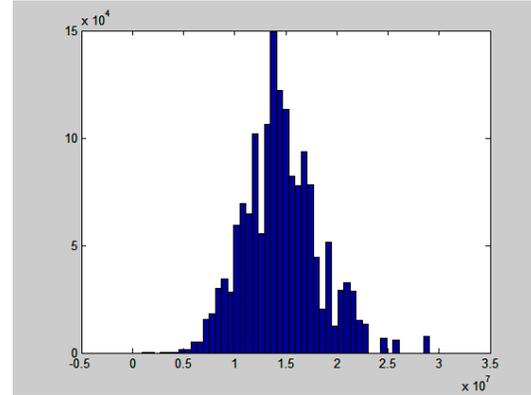
Figure 32 (a) and Figure 32 (b) depict a portion of the image containing a railroad portion and the corresponding Fourier Transform with suppressed DC component. The two parallel lines at North-West and South-East corners of Figure 32 (b) correspond to the sleepers in the Fourier domain. The orientation of these lines is perpendicular to the actual direction of the railway which is represented by the diagonal line running along the South-West and North-East direction through the origin. Figure 32 (c) and (d) depict a highway segment and its Fourier Transform respectively. In Figure 32 (d), the spectral peaks corresponding to the frequency of the sleepers is missing. The periodic structure of road line markings have a frequency that is different that of the sleepers. Furthermore, their direction is identical to those of the roadsides so their contribution will be in the direction perpendicular to that of the sleepers.

Histogram Analysis

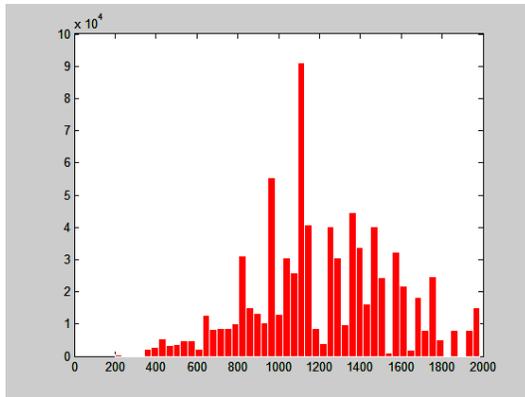
In order to determine the threshold value to discriminate between roads and railroads we need to further analyze the distribution of spectral energies. We have the Fourier energy values for each window. A map is constructed with same size of the real image. Note that only the windows overlapping with the road mask is considered. Fourier spectrum of railroad and drive-roads in the image are analyzed in terms of their total energy, average energy, and average energy in region of interest (ROI) in Figure 33. The average energy under ROI corresponds to the region of interest of the fused method as shown in Figure 28 (i). It is seen that average energy in the ROI histograms for the railroads and roads are significantly different from each other.



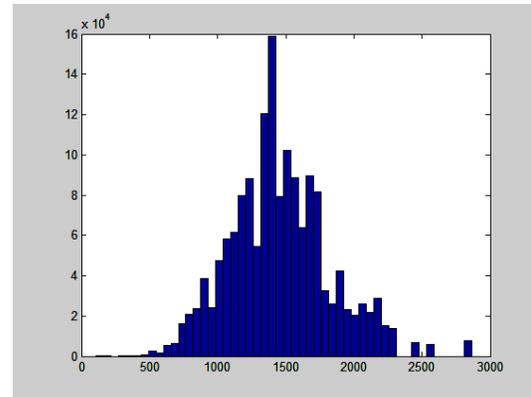
Total energy Railroad



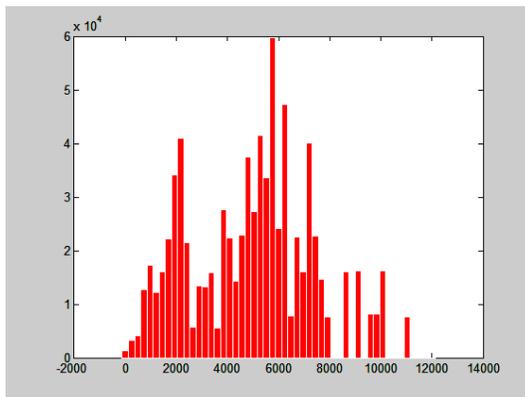
Total energy for drive-road



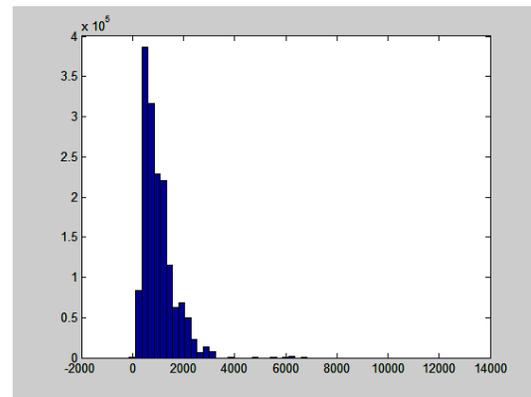
Average energy for Railroad



Average energy for drive-road



Average energy in ROI for Railroad



Average energy in ROI for drive-road

Figure 33 Histogram of Fourier spectra energy for railroads and drive-roads

A total energy in a window is the sum of signal and noise energies. To account for the background noise energy, we propose to subtract its contribution from the total energy. We define background energy regions in two ways i) average energy out of the ROI for sleepers

(BK1) and ii) average energy outside the region consisting of energy location of sleepers, road sides and road direction (BK2). Sample Fourier transform and background noise energy regions are shown Figure 34.

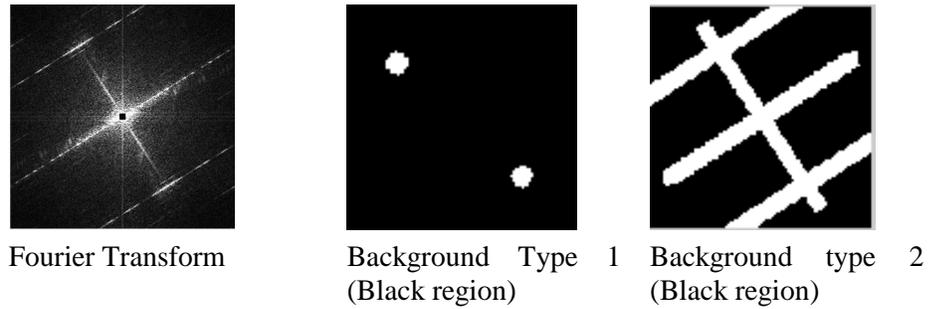
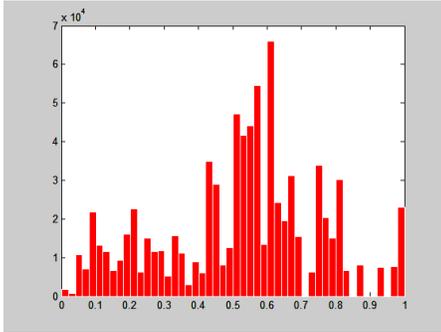


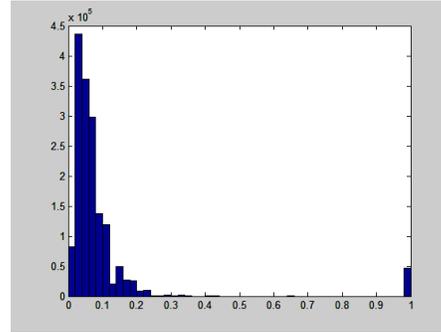
Figure 34 Sample Fourier transform and noise regions.

Histograms of ratio and difference of average energy and background noise energy are shown in in Figure 35.

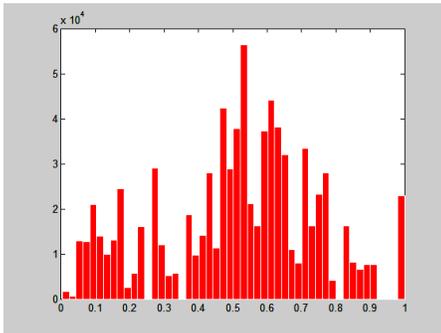
It is observed that, compensating with BK1 or BK2 does not affect the result significantly based on the histograms. In terms of average values ratio Average Energy and BK1 and ratio Average Energy and BK2 are more promising discriminative features. In order to deeply analyze these features, we apply the technique on three different images. In Figure 36, boxplots related to those features are shown.



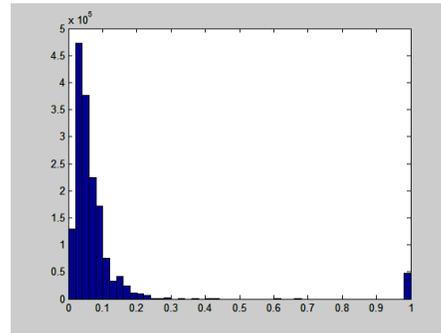
Average Energy/BK1 for Railroad



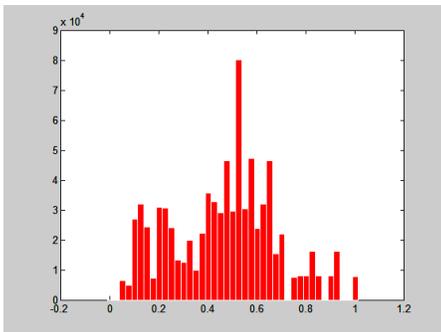
Average Energy/ BK1 for drive-road



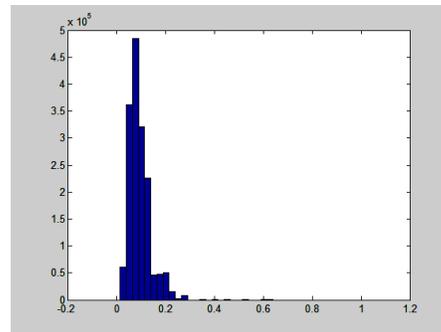
Average Energy/BK2 for Railroad



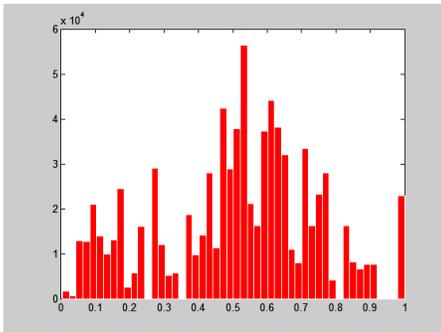
Average Energy/ BK2 for drive-road



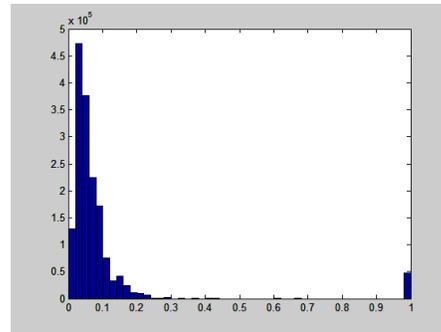
Average energy - BK1 for Railroad



Average energy - BK1 for drive-road

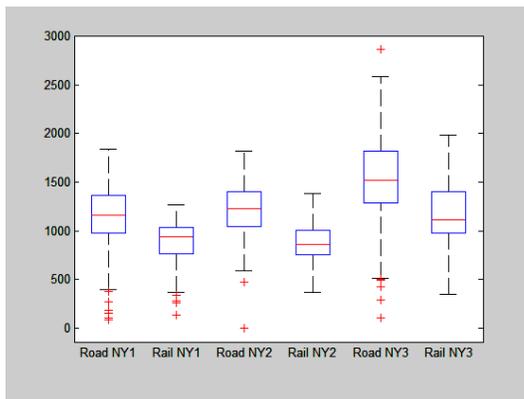


Average energy - BK2 for Railroad

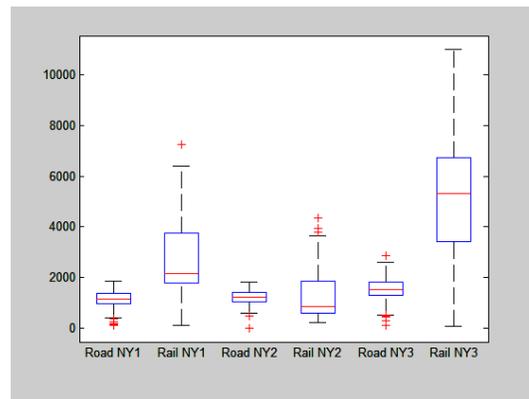


Average energy - BK2 for drive-road

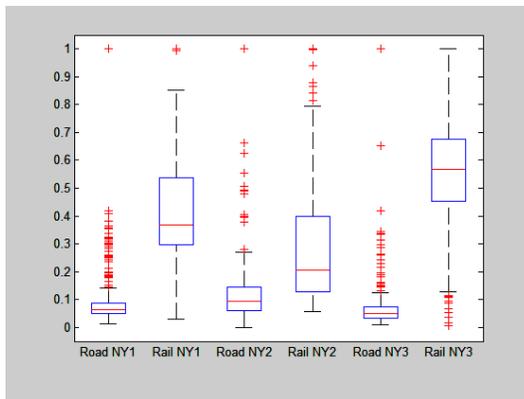
Figure 35 Histogram of Fourier spectra energy computed by compensating for background energy regions



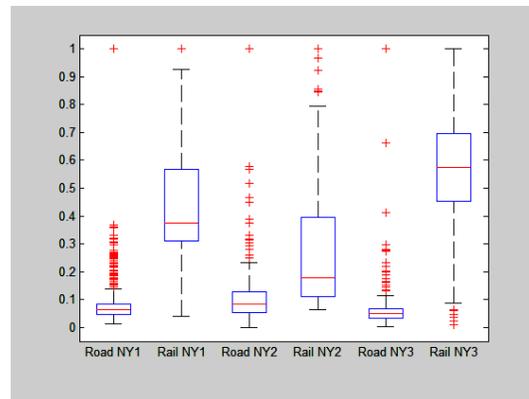
a) Average Energy



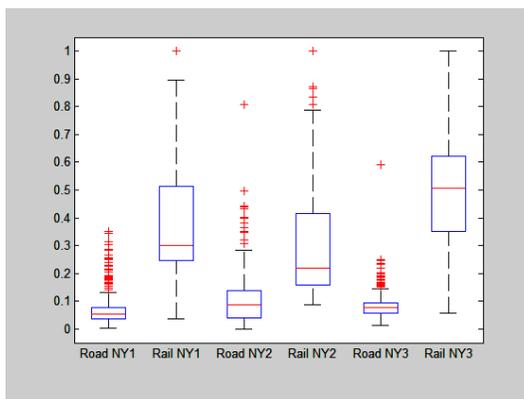
b) Average Energy in ROI



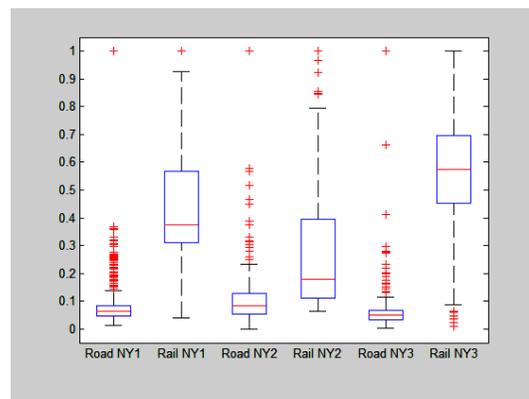
c) Ratio Average Energy and BK1



d) Ratio Average Energy and BK2



e) Average energy - BK1



f) Average energy - BK2

Figure 36 Boxplots of Fourier spectra energy for three different images.

As shown in Figure 36(e) and (f), subtraction of average energy and background value can be used to specify the score map threshold. We preferred using Average Energy-BK1 for score map construction. Sample road mask image and corresponding railroad score map are shown in Figure 37.

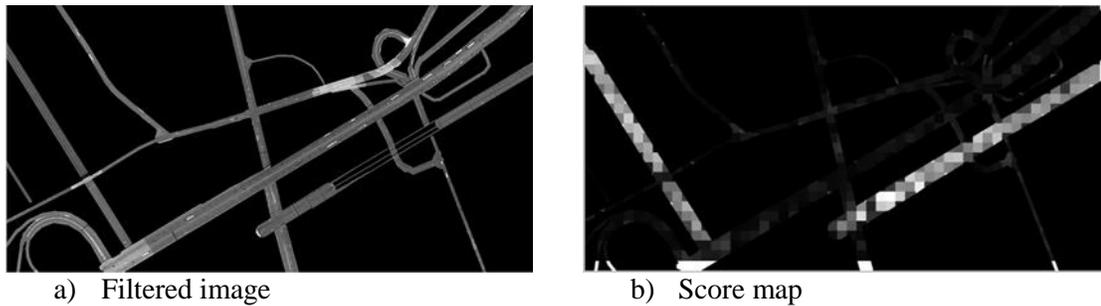


Figure 37 Sample Score map

After score map construction it is normalized between zero and one to use fixed threshold value. Then the map is thresholded by pre-defined value which is 0.18 for this study. Lastly post-processing operations are applied on this mask. The post-processing operation consists of removal of small components (having higher energy due to zebra crossing or other periodic structure), segmenting initial road masks and associating the algorithm result with the original mask. Morphological operations are applied on the railroad mask to eliminate the remaining artifacts in the mask. Sample thresholded score map and railroad mask are shown in Figure 38.

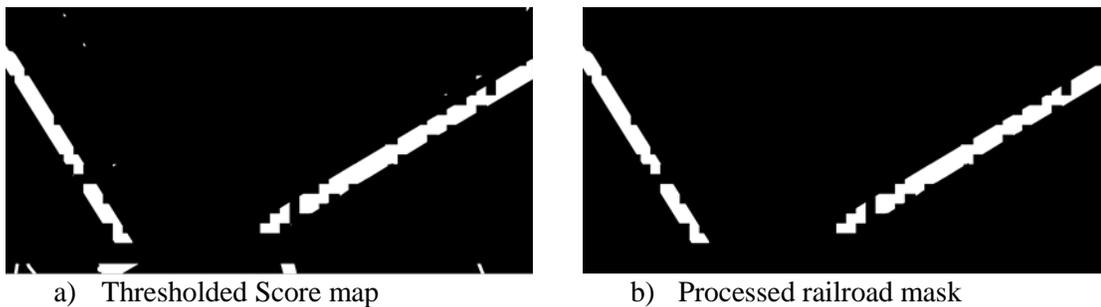
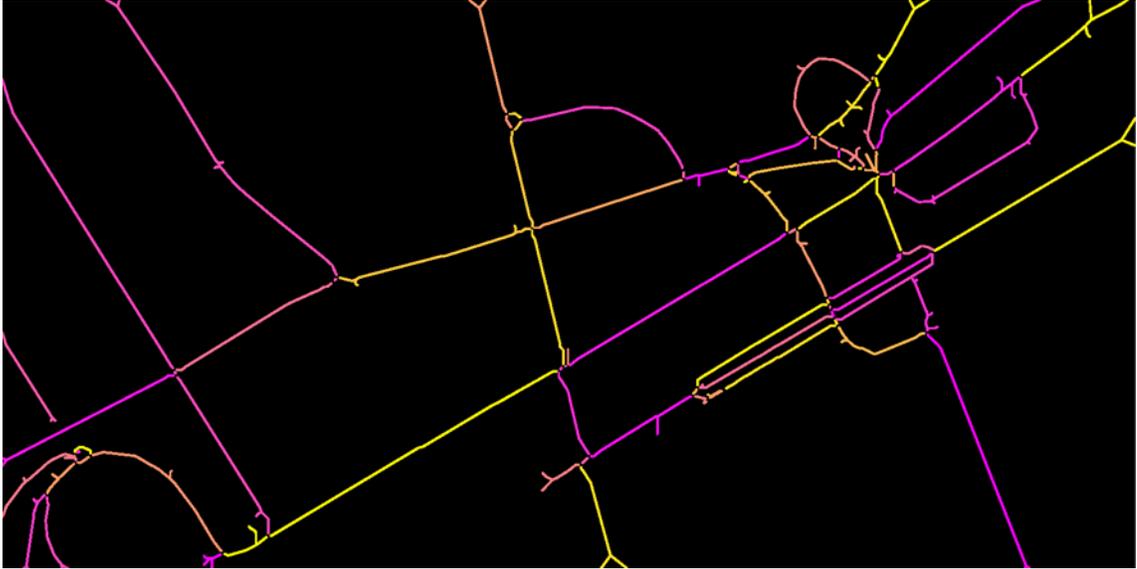


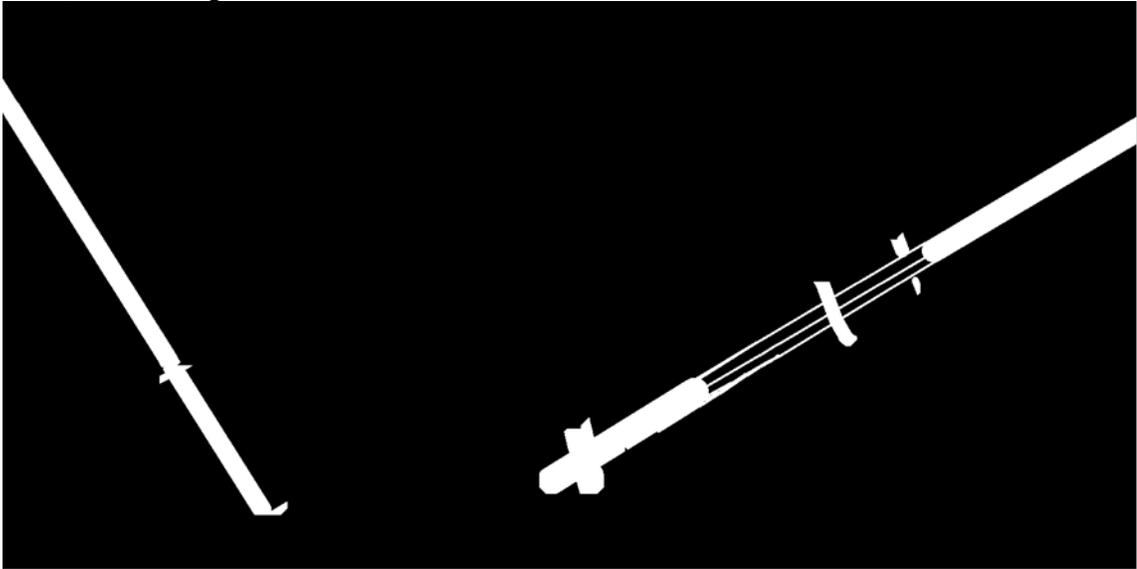
Figure 38 Sample thresholded Score map and final mask

Since we have road network including railroad and drive road already, obtained mask is compared with the given road mask. Firstly, the junction points are found by the method proposed by Çinar, et al (2012). These junction points are excluded from the given road network skeleton and remaining network is labeled. Then, each component in the preliminary railroad mask is compared with the labeled road mask. If a labeled component in the railroad mask has

enough matching number of pixel with the given road mask, this component is classified as railroad. Lastly, morphological dilation is applied on the obtained mask to get clearer railroad mask. Sample labeled road network and algorithm result are shown in Figure 39.



a) Labelled given road mask



b) Algorithm result

Figure 39 Labaled given road mask and the algorithm result

4.8 Road-Type Classification

Road type and surface material information is important for navigation, quality inspection and military applications. In this part of the study, an approach for road type detection is proposed.

After extracting roads using the road detection algorithm, they are classified into three classes: asphalt, concrete and dirt (unpaved). Firstly, a training set is created by the user typifying the three classes; asphalt, concrete and soil. The training set is obtained from the image shown in Figure 41. As surface material such as asphalt could have various states such as dry, wet, old, new and so on, a single Gaussian distribution is not enough to cover all these cases. Gaussian mixture models are useful to model such variations. For asphalt and soil two Gaussian mixtures are created whereas only one Gaussian model is constructed for concrete roads. The mixture probabilities for asphalt and soil are selected as 0.5.

For each point extracted by the road extraction algorithm, Mahalanobis distance to each Gaussian component in Gaussian mixtures in the spectral (color) domain is computed as shown below.

$$D(x) = \sqrt{(x - \mu)^T S^{-1} (x - \mu)} \quad (\text{Equation 4.12})$$

where x represent the corresponding pixel's spectral values of each band $x=(x_{CB}, x_B, x_G, x_Y, x_R, x_{RE}, x_{NIR1}, x_{NIR2})^T$, S represents 8X8 the covariance matrix of each Gaussian distribution and μ represent the 8 dimensional mean vector of Gaussian distribution. The point is classified based on the shortest distance to the individual classes. Once the individual pixels are assigned to the classes, the segmentation result is used to assign each segment into one class. If a segment includes pixels voting for different classes, they are merged into to one class based on the number of pixels voting for each class. The flow chart of road type detection is shown in Figure 40.

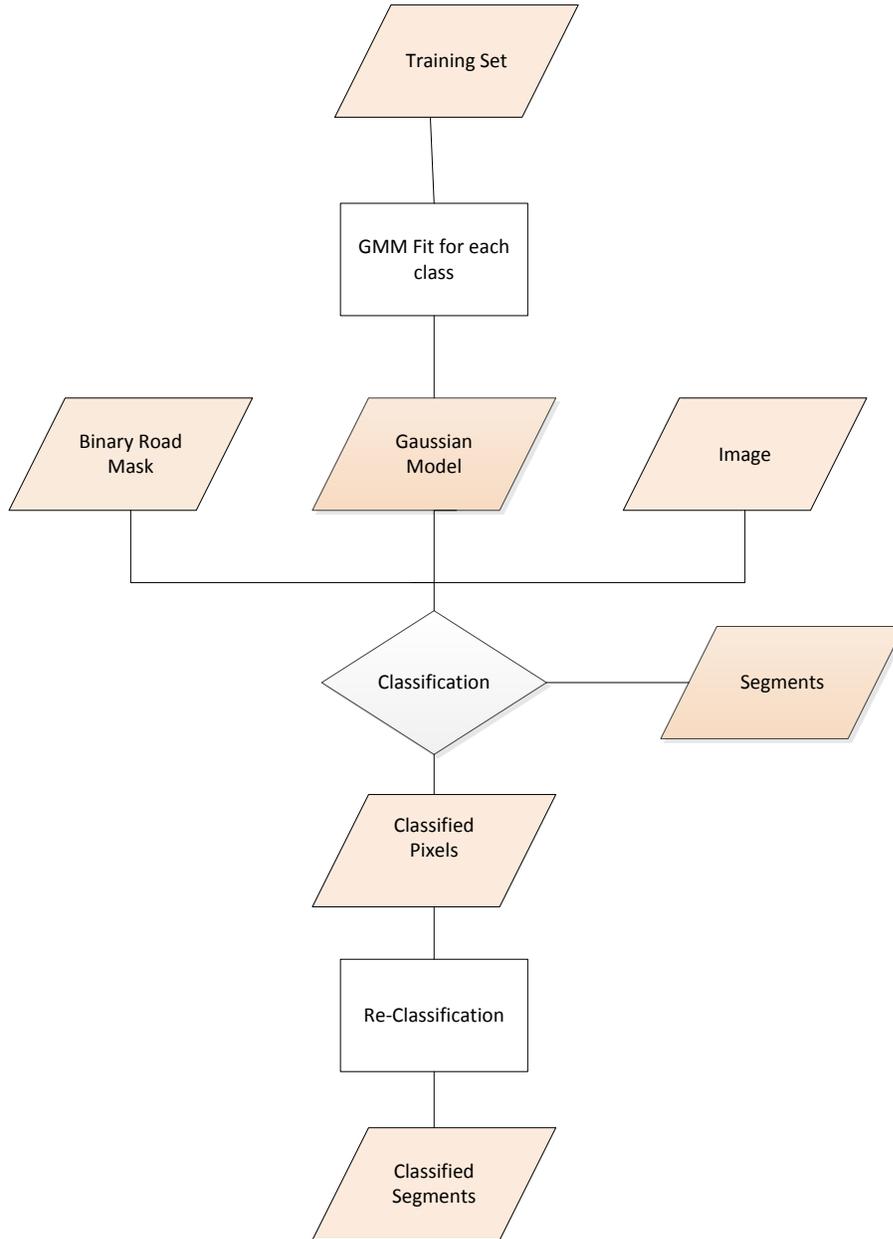


Figure 40 Road Type classification flow chart.

Sample result about road type classification is shown in Figure 41. Blue parts represent asphalt and yellow parts represent unpaved roads. The result of the road material classification method will be presented in Section 5.4.



Figure 41 Road type classification sample result

4.9 Verification

To evaluate effectiveness of the algorithms precision, recall and F-measure metrics have been used. The proposed algorithm produces a road mask and it is compared with the “Ground Truth”. In this part of the study, performance metrics will be defined. We employed two types of performance measurement; “mask to mask” and “mask to skeleton”.

Firstly, variable used in performance measurement is defined below;

True Positives (TP) : Correspond to the number of pixels that belong to the roads in both GT and extracted road mask.

False Positives (FP) : Correspond to the number of pixels that are in the extracted road mask but does not belong to GT.

True Negatives (TN) : Correspond to the number of pixels that are not in GT and road mask.

False Negative (FN) : Correspond to the number of pixels that belong to GT but are not extracted by the algorithm.

4.9.1 Precision:

Precision refers to ratio of true positives over all produced data. It answers the question of “What percentage of the output is indeed what I was looking for?” Equation 4.13 shows the precision computation

$$\text{Precision} = \frac{TP}{TP+FP} \quad (\text{Equation 4.13})$$

4.9.2 Recall:

Recall tries to answer the question “What percentage of all the extracted output is actually correct?” Recall is computed as shown in Equation 4.14.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (\text{Equation 4.14})$$

4.9.3 F-Measure:

There is a tradeoff between precision and recall values so that increasing one typically results in decreasing the other. In order to combine these two metrics, F-measure is used. F-Measure is the harmonic mean of those two metrics. It is given in Equation 4.15 as

$$F_{\beta} = (1+\beta) * \frac{\text{Precision} * \text{Recall}}{(1+\beta) * \text{Precision} + \text{Recall}} \quad (\text{Equation 4.15})$$

“ β ” parameter is used to differentiate the weights of precision and recall on F-Measure results. Generally $F_{0.5}$, F_1 and F_2 are used for F-Measure calculation. The higher and smaller “ β ” value weigh the recall are precision values more, respectively.

4.9.4 Performance Measurement

There are different methods for evaluating the performance of linear object. In the study conducted by Aksoy, et.al. (2010), object-based performance measurement have been used which aims to measure performance based on ground truth and the result objects of the algorithm. However, object based road extraction performance measurement requires pre-processing of ground truth and objectification of the road extraction algorithm result. We preferred to measure the algorithm performance in two ways: mask-mask and mask-skeleton which can be used measure performance of both geometric and thematic object extraction algorithms.

4.9.4.1 Mask – Mask Performance Measurement

In this performance measurement approach, GT and extracted mask is used as in mask format rather than skeleton. In this type of measurement, TPs FPs and FNs are shown in Figure 42.



Figure 42 Mask-Mask Performance measurement.

Blue parts represent FPs, greens part show TPs and red parts represent FN respectively. Precision and recall are calculated based on the area of these parts.

4.9.4.2 Mask – Skeleton Performance Measurement

In this type of the performance measurement TPs and FPs are based on algorithm output skeleton. As shown in Figure 43, yellow parts represent GT mask, while blues are FP and greens are TP pixels.

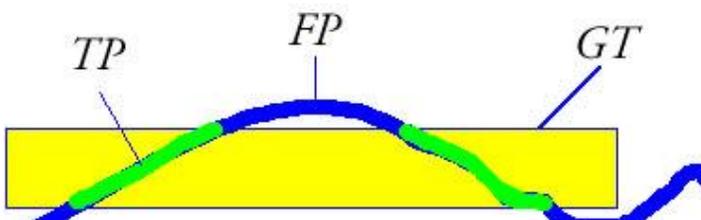


Figure 43 Mask-Skeleton performance measurement.

For FN calculation, the GT is used as skeleton and algorithm result is obtained as mask. As stated below, reds are FN and yellow part is road algorithm result mask in polygon format.

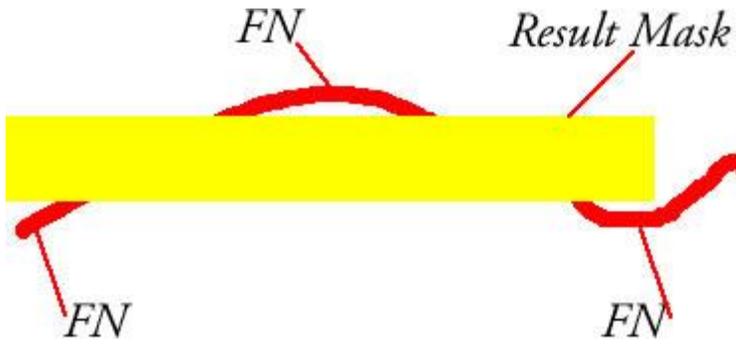


Figure 44 Mask-Skeleton Recall Performance

In short, FP and TP are calculated in the case of GT is a mask while FN is calculated when road output mask is represented as a polygon.

CHAPTER 5

5 RESULTS AND DISCUSSION

5.1 Dataset Used in Experiment

In this thesis, we aimed to develop a road network extraction algorithm for high spatial resolution and multi-spectral satellite images. WorldView-2 (WV-2) images are used for this purpose. WV-2 satellite has been launched 8 October 2009 by DigitalGlobe, Inc. WV-2 provides eight spectral bands including one panchromatic and eight Multi-spectral bands each of which has 40-200 nm bandwidth. It was launched to provide images for spectral analysis, mapping and monitoring, disaster management, defense applications, visualization and simulation of environment. In addition, land-use planning including road network updating is also among expected studies for WV-2 images.

The images have 50 cm and 2.0 m spatial resolution of panchromatic and multi-spectral bands respectively. These bands may contribute to more accurate feature classification and land-cover analysis. Multi spectral bands consist of Coastal Blue, Blue, Green, Yellow, Red, Red-Edge, NIR1 and NIR2 bands.

Spectral properties of WV-2 images is shown in Table 3

Table 3 Spectral properties of WV-2

Band	Spectrum (nm)
Panchromatic	400-1040
Coastal Blue	400-450
Blue	450-510
Green	510-580
Yellow	585-625
Red	630-690
Red Edge	705-745
Near IR 1	770-895
Near IR 2	860-1040

Additional information about WV-2 is provided in Digital Globe official website. Detailed information is provided in APPENDIX F.

In order to test the developed algorithm, four different WV-2 images are used. These images enable us to evaluate the model for different lands and different time. Three of those represent city of Ankara including different land covers. One is obtained on 16 November 2010 and two of them are taken on 26 September 2011.

The first image also covers METU campus and neighborhood area. This enables us to determine the GT information easily. Projection of the images representing city of Ankara is shown in Figure 45.



Figure 45 WV-2 Image obtained 16 November 2010

In addition to this image, two more images for water analysis have been obtained 26 September 2011. The image covering “Eymir lake” is shown in Figure 46.

We have also obtained WV-2 image covering Yeni Mahalle region covering many different land objects. The image is shown in Figure 47.

The fourth one represents a different country, Russia, and it is taken on 7 July 2010. The general view of this image is presented Figure 48.



Figure 46 WV-2 Image (Eymir Lake)



Figure 47 WV-2 Image (Yeni Mahalle)



Figure 48 General view the image taken from Russia.

One of the more important aims of this study is to measure the additional spectral bands' contribution to the performance of the developed algorithm. In order to test this image, we need 4-Band, RGB and Gray level images.

One of the important goals of this study is to measure the contribution of additional spectral bands to the performance of the developed algorithm. In order to test this image, we need 4-Band, RGB and Gray level images.

Some satellites provide 4-band images including RGB and NIR bands. In order to compare the algorithm with 4-bands IKONOS and GEOEYE images can be used. At this point, we do not have IKONOS or GEOEYE (4-Band) images of the same area with the 8-Band WorldView-2 image. Nevertheless, we use corresponding bands of GEOEYE the those of WorldView-2. For this purpose, blue, green, red and NIR1 bands of WV-2 images have been used. Reflectance comparison of some satellite images is shown in Figure 49.

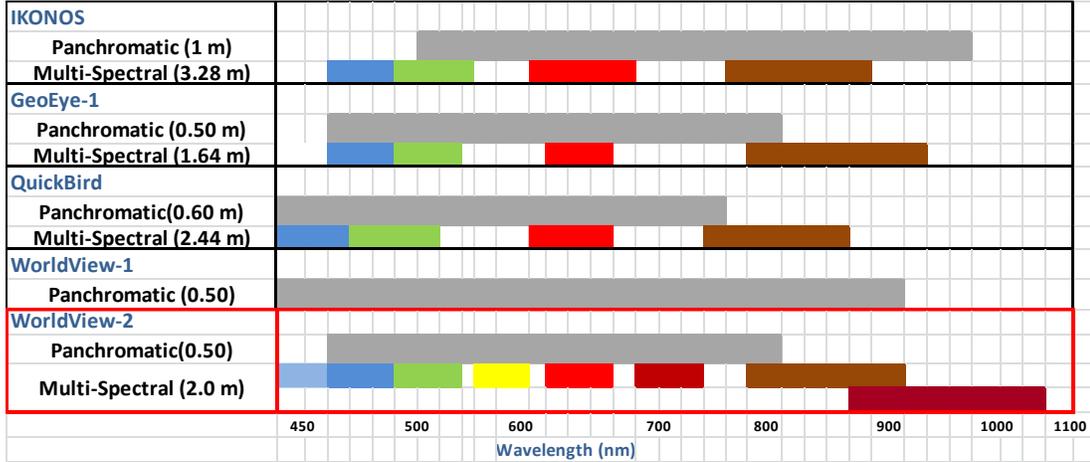


Figure 49 Reflectance comparison of some satellite images

5.2 Results

In this part of the study, algorithm results will be reported for each module and for each image type: WV-2, 4-Band, RGB and Gray level image. Visual and quantitative results of the image shown in Figure 50 are used for demonstration. The results for all data used in this study will be presented in APPENDIX A.



Figure 50 Sample image for demonstration

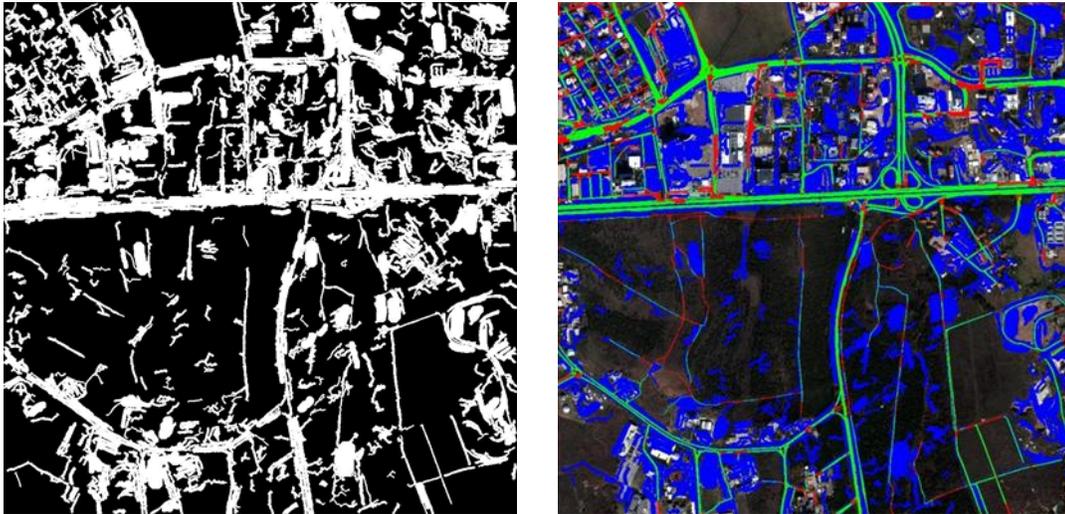
5.3 Road Extraction

5.3.1 Edge-Based Road Extraction

In this part of the study, edge detection based module will be applied on WorldView-2 (WV-2), RGB+NIR four band image, RGB, Grayscale, in addition to panchromatic image to analyze the contribution of the additional bands.

5.3.1.1 WorldView-2 Images

Edge based road detection result for WV-2 images are presented below. In Figure 51, edge based road detection result mask is shown. In Figure 51(b), blue parts corresponds to FP while red parts represent FN and green ones corresponds to the TP.



a) Edge based road mask

b) Edge based road mask (colored)

Figure 51 Edge Based Result for WV-2 image

As mentioned in Chapter 4, we have three parameters for edge based road extraction module. For center line extraction we used three parameters, maximum width, minimum width, and minimum length. These values is set as 4 meters (2 px), 40 meters(20px) and 80 meters (40 px). However, the module is also tested with different parameters. The results are presented in Table 4.

Table 4 Edge based road extraction module performance for different width values.

Max Width	10							
Min Width	1		2		4		6	
	MM	MS	MM	MS	MM	MS	MM	MS
Precision	0,34	0,32	0,41	0,42	0,58	0,65	0,58	0,64
Recall	0,43	0,39	0,42	0,33	0,20	0,15	0,05	0,03
F _{0.5}	0,35	0,33	0,41	0,40	0,42	0,39	0,18	0,12
F ₁	0,38	0,35	0,41	0,37	0,29	0,24	0,09	0,06
F ₂	0,41	0,37	0,41	0,35	0,23	0,17	0,06	0,04

Max Width	20							
Min Width	1		2		4		6	
	MM	MS	MM	MS	MM	MS	MM	MS
Precision	0,29	0,25	0,34	0,32	0,49	0,52	0,55	0,57
Recall	0,61	0,53	0,59	0,46	0,37	0,24	0,22	0,10
F _{0.5}	0,33	0,28	0,38	0,35	0,46	0,42	0,42	0,30
F ₁	0,40	0,34	0,43	0,38	0,42	0,32	0,31	0,18
F ₂	0,50	0,43	0,51	0,42	0,39	0,26	0,25	0,12

Max Width	30							
Min Width	1		2		4		6	
	MM	MS	MM	MS	MM	MS	MM	MS
Precision	0,27	0,24	0,32	0,30	0,41	0,46	0,43	0,49
Recall	0,62	0,55	0,60	0,47	0,39	0,25	0,23	0,11
F _{0.5}	0,31	0,27	0,35	0,32	0,41	0,39	0,37	0,29
F ₁	0,38	0,33	0,41	0,37	0,40	0,32	0,30	0,18
F ₂	0,50	0,43	0,51	0,42	0,39	0,27	0,26	0,13

Max Width	40							
Min Width	1		2		4		6	
	MM	MS	MM	MS	MM	MS	MM	MS
Precision	0,26	0,23	0,30	0,29	0,36	0,42	0,34	0,41
Recall	0,63	0,55	0,61	0,48	0,40	0,25	0,24	0,11
F _{0.5}	0,29	0,26	0,33	0,31	0,37	0,37	0,32	0,27
F ₁	0,37	0,33	0,40	0,36	0,38	0,31	0,29	0,18
F ₂	0,49	0,43	0,50	0,42	0,39	0,27	0,26	0,13

It is also important to note that narrow roads which are mostly soil type roads can be extracted with this module. However, visual results depict that still most part of tiny roads are missing. Increasing minimum width value results in missing more tiny roads as expected.

Since modular approach is employed in this study, we aim to extract higher precision and recall in each module. In order to achieve higher F-measure, we apply union operation for all edge based road masks instead of intersection. For all type of input image, we use the same parameters for this module.

5.3.1.2 Four-Band

As it is mentioned, Blue, Red, Green and NIR1 bands of WV-2 images are combined for testing the proposed approach for Four-band images. Edge based road extraction result for the four-band image is given in Figure 52.

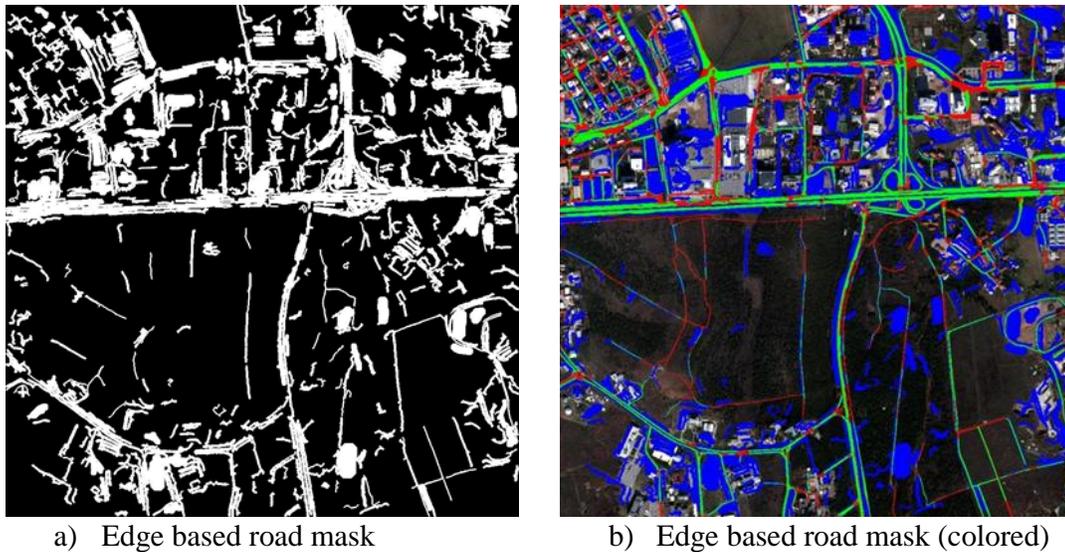


Figure 52 Edge Based Result for Four-band image

5.3.1.3 RGB Image

In Figure 53 Edge Based Result for the RGB image is presented.

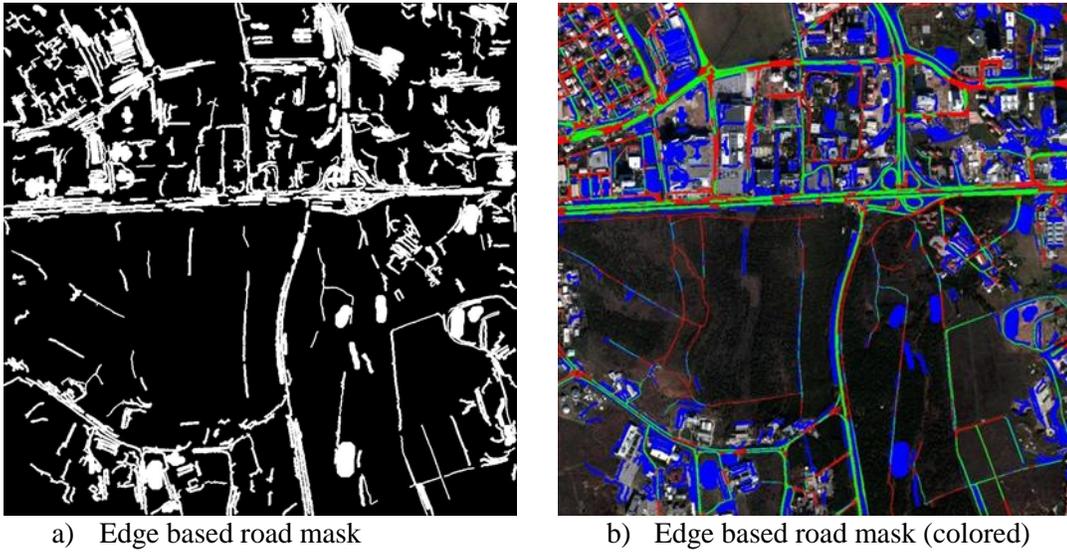


Figure 53 Edge Based Result for the RGB image

5.3.1.4 Gray Level Image

For the gray-level image, edge based road extraction result is presented in Figure 54.

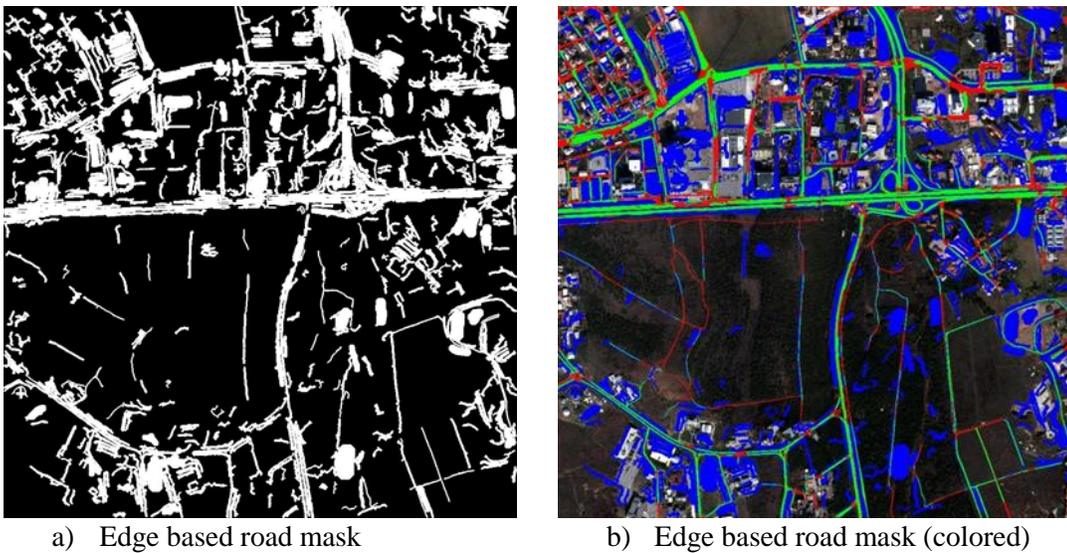


Figure 54 Edge Based Result for Gray image

5.3.1.5 Panchromatic Image

Panchromatic image is also tested in this study. Edge based road extraction result for panchromatic images are presented in Figure 55.

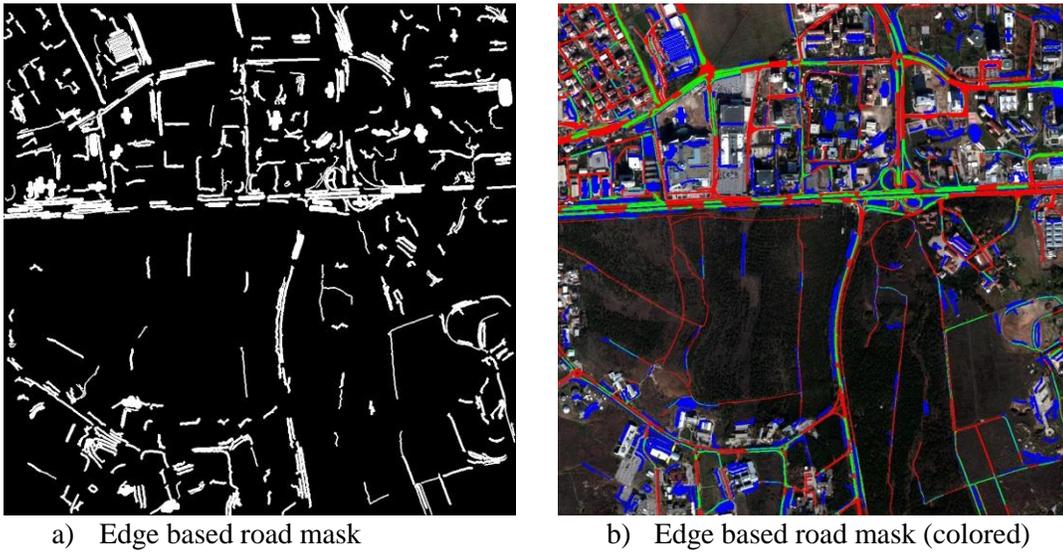


Figure 55 Edge Based Result for Panchromatic image

All result related to this module is summarized in Table 5 and Figure 56.

Table 5 All result of edge based road detection module

	WV-2		Four-band		RGB		Gray		Pan	
	M-M	M-S	M-M	M-S	M-M	M-S	M-M	M-S	M-M	M-S
Precision	0,25	0,27	0,34	0,32	0,28	0,32	0,32	0,34	0,30	0,30
Recall	0,63	0,53	0,59	0,50	0,56	0,46	0,32	0,34	0,35	0,30
F_{0,5}	0,28	0,29	0,38	0,33	0,31	0,35	0,39	0,34	0,31	0,30
F₁	0,36	0,35	0,43	0,38	0,37	0,38	0,34	0,33	0,32	0,30
F₂	0,48	0,44	0,51	0,44	0,46	0,42	0,36	0,33	0,34	0,30

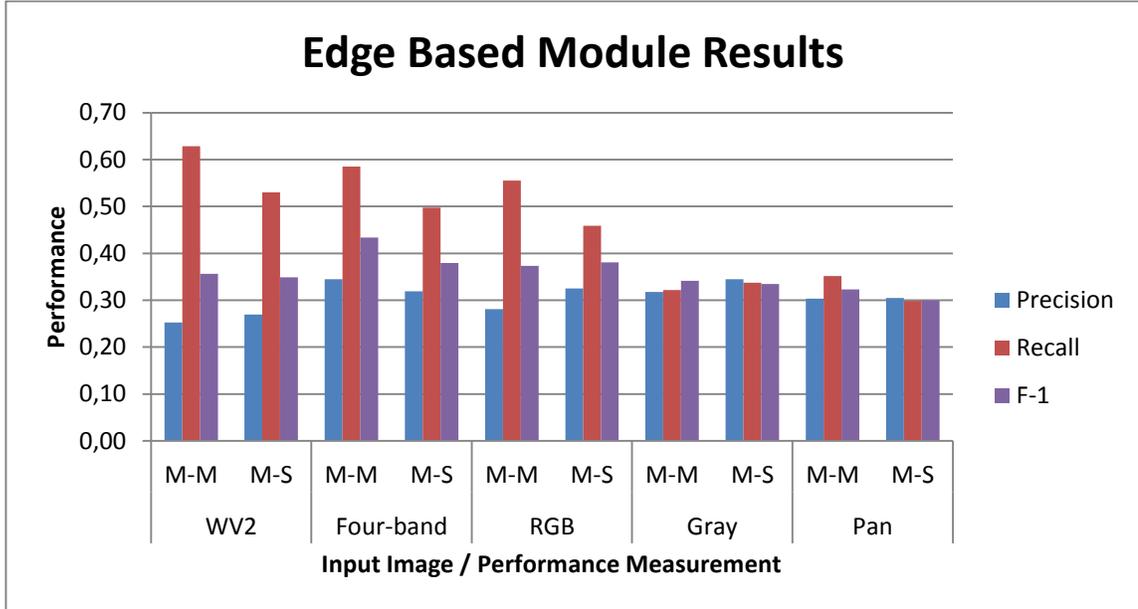


Figure 56 Edge based module results

As shown in Figure 56, the edge based module provide lower precision value for WV-2 images than four-band and RGB images. As WV-2 images, the module may produce more edges, since it uses more spectral bands. Some of these edges may include parallel non-road regions. This causes to achieve higher recall but lower precision.

5.3.2 Structure analysis based road extraction

In this part of the study, structure analysis based road detection module result will be presented for WV-2, RGB+NIR, RGB, Grayscale and Panchromatic images. As mentioned in Chapter 4, Mean-Shift segmentation needs three different parameters; “Minimum Region”, “Range Bandwidth” and “Spatial Bandwidth”. In the study of Çinar (2012), it is suggested to set “range bandwidth” parameter as 8 and “spatial bandwidth” as 4. It is also reported that these bandwidth parameters have a minor effect in the structure analysis based road extraction performance. On the other hand, the “minimum region” parameter is shown to be influential factor for segments shape. Therefore, we conducted experiments for “minimum region” parameter. Different minimum region parameters have been tested to define the parameter of Mean Shift segmentation. As seen in Figure 57, we set the minimum region parameter as 380.

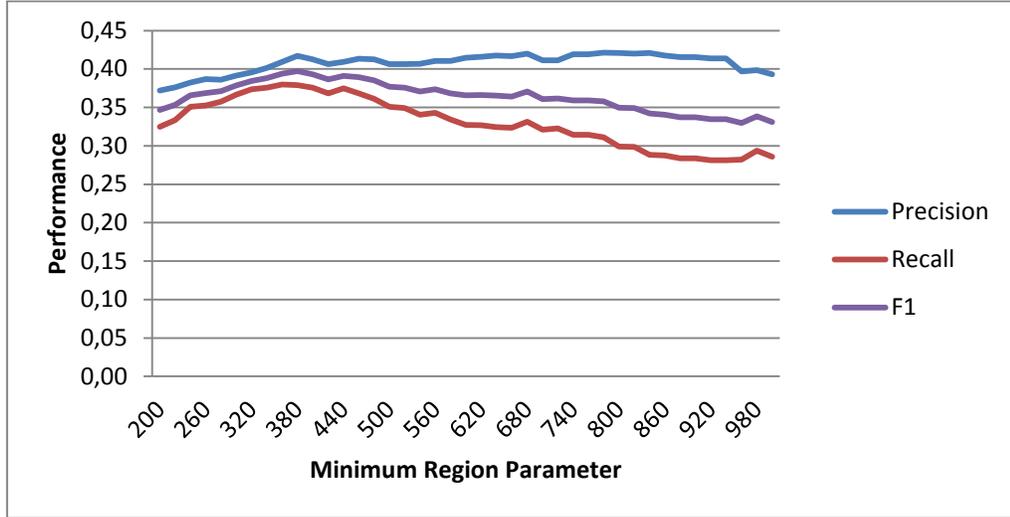


Figure 57 Minimum-region parameter of Mean-Shift segmentation

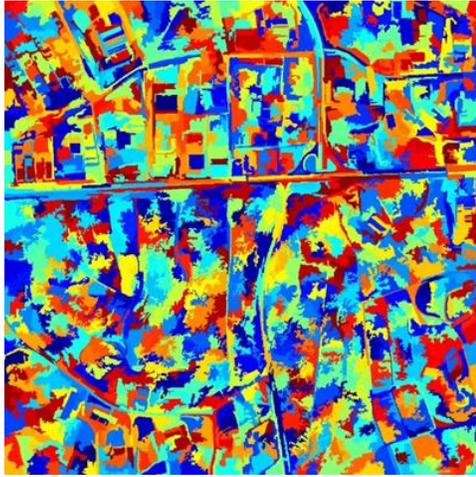
5.3.2.1 WV-2 Images

As mentioned in Chapter 4, structure analysis on segmentation result for WV-2 will be applied on two different four band combinations since one combination provides higher precision while other provides higher recall. Then, these two results are fused by the union operation.

In Table 6, performance result of the demonstration image for structure analysis based road extraction module is presented.

Table 6 Result of Structure analysis based road extraction module for WV-2 Image

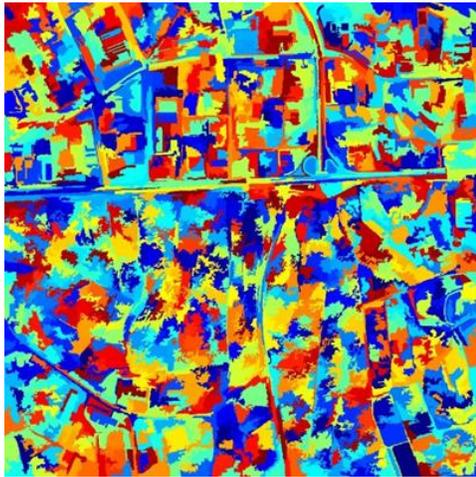
	CB,RE,NIR1,NIR2		B,G,Y,R		Merged for WV-2	
	MM	MS	MM	MS	MM	MS
Precision	0.3801	0.3262	0.3924	0.3472	0.3372	0.2869
Recall	0.5031	0.3276	0.4336	0.3238	0.6039	0.4379
F_{0.5}	0.3992	0.3265	0.3974	0.3423	0.3699	0.3082
F₁	0.4318	0.3269	0.4090	0.3351	0.4328	0.3467
F₂	0.4703	0.3273	0.4214	0.3282	0.5214	0.3962



a) Labels for CB,RE,NIR1,NIR2



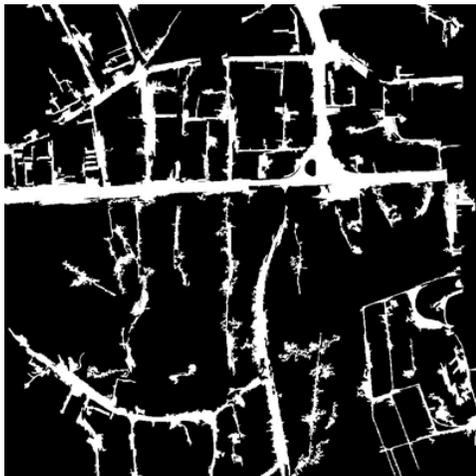
b) Structure analysis based road mask (colored)



c) Labels for B,G,Y,R



d) Structure analysis based road mask (colored)



e) Segmentation Result



f) Segmentation result (Colored)

Figure 58 Structure analysis based module result for WV-2

5.3.2.2 Four-Band image

For four-band image (RGB,NIR-1), visual results of structure analysis based road extraction module are presented in Figure 59.

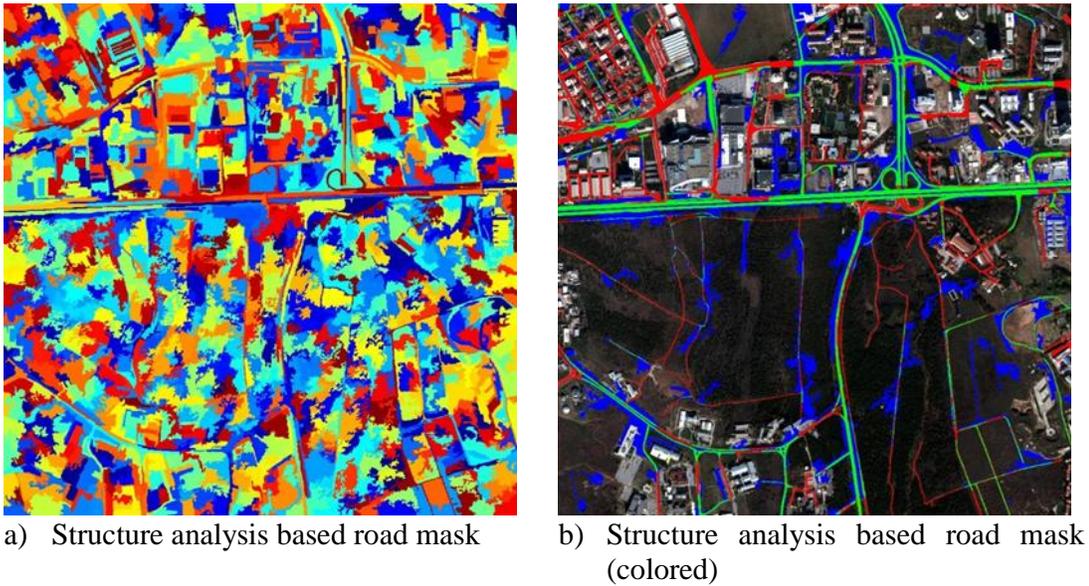


Figure 59 Structure analysis based result for four-band image

5.3.2.3 RGB Image

For RGB image, edge based road extraction result is presented in Figure 60.

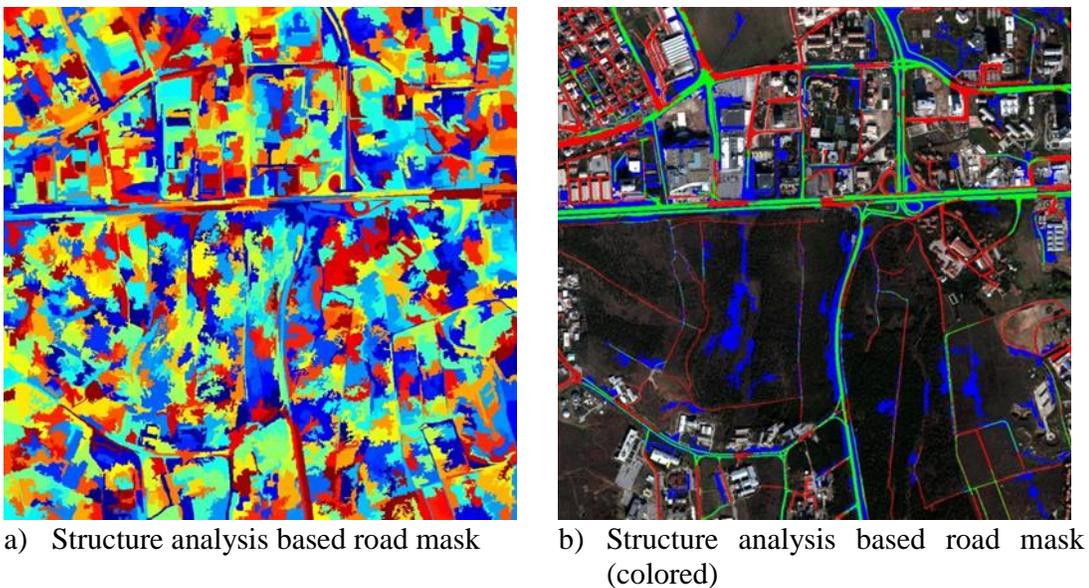


Figure 60 Structure analysis based result for RGB image

5.3.2.4 Gray Level Image

For Gray-level image, structure analysis based road extraction result is presented in Figure 61.

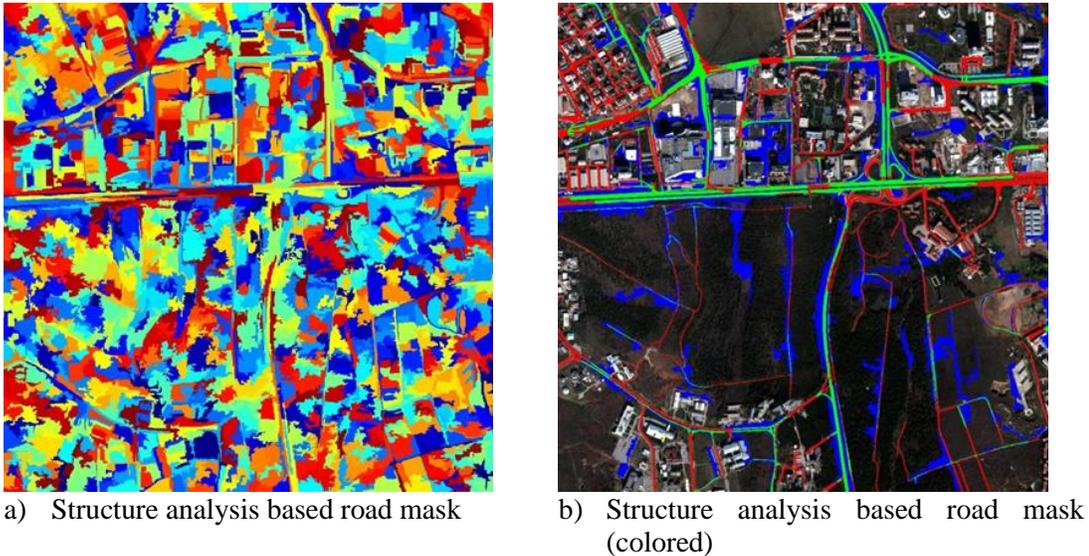


Figure 61 Structure analysis based for gray image

5.3.2.5 Panchromatic Image

Structure analysis based road extraction result for panchromatic images is presented in Figure 62. In

Table 7 All result of structure analysis based road extraction is depicted. The visual comparison of the results of the module is shown in Figure 63.

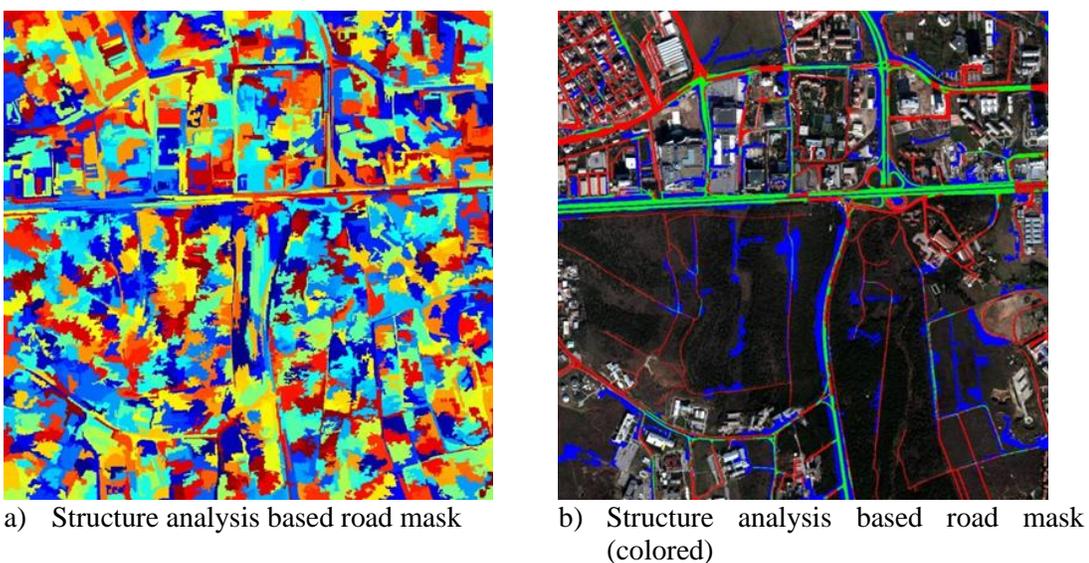


Figure 62 Structure analysis based result for panchromatic image

Table 7 All result of structure analysis based road extraction

	WV-2		Four-band		RGB		Gray		Pan	
	M-M	M-S	M-M	M-S	M-M	M-S	M-M	M-S	M-M	M-S
Precision	0,34	0,29	0,35	0,29	0,38	0,32	0,32	0,27	0,32	0,27
Recall	0,60	0,44	0,50	0,34	0,47	0,31	0,44	0,27	0,39	0,23
F_{0,5}	0,37	0,31	0,37	0,30	0,39	0,32	0,33	0,27	0,34	0,26
F₁	0,43	0,35	0,41	0,32	0,42	0,32	0,37	0,27	0,36	0,25
F₂	0,52	0,40	0,46	0,33	0,45	0,31	0,41	0,27	0,38	0,24

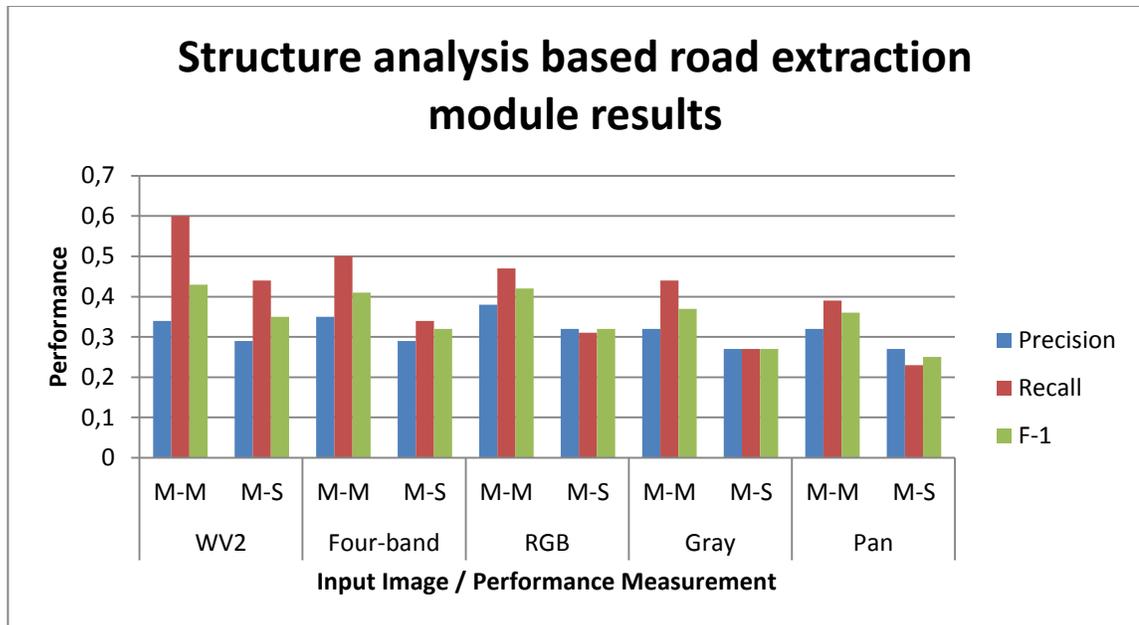


Figure 63 Structure analysis based road extraction module results.

5.3.3 Clustering-Based Road Extraction

Clustering based road detection result will be presented in this part of the study.

5.3.3.1 WV-2 Images

WV-2 images are clustered and analyzed as mentioned in Chapter 4. In Figure 64, road mask and colored mask representing TP, FP and FN values are in different color. As it is mentioned before, we have two different approaches for clustering. One is clustering each band separately and union of each output. Second approach is clustering all data together and analyze clusters.

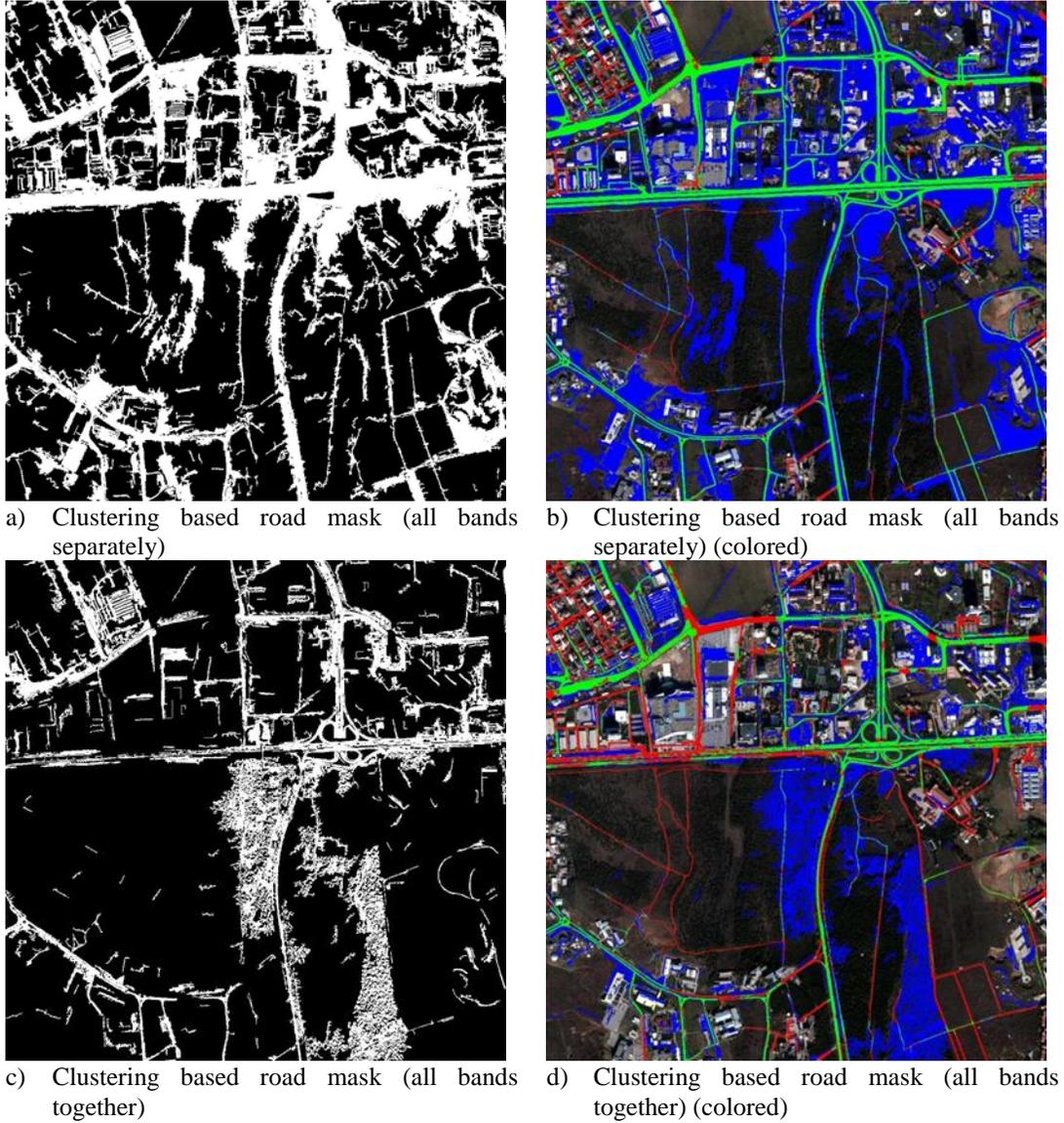


Figure 64 Clustering based result for WV-2 image

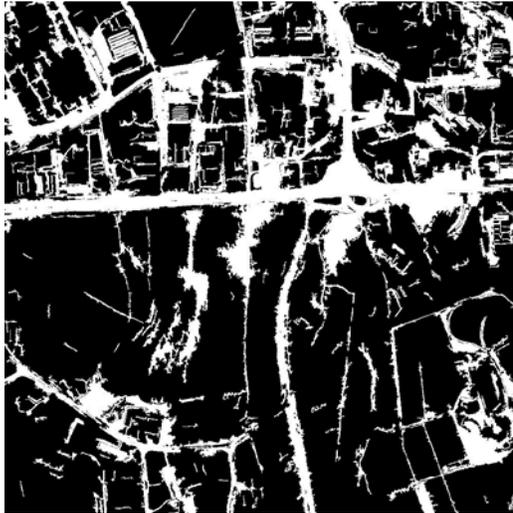
In Table 8, performances of clustering algorithms on the demonstration image for this module are presented.

Table 8 Result of cluster based module for WV-2 image

	Clustering Each Band Separately		Clustering all bands together	
	Mask-Mask	Mask-Skeleton	Mask-Mask	Mask-Skeleton
Precision	0.2316	0.1828	0.2706	0.2303
Recall	0.8105	0.7089	0.4980	0.5182
F_{0.5}	0.2702	0.2147	0.2978	0.2591
F₁	0.3603	0.2907	0.3507	0.3189
F₂	0.5404	0.4499	0.4263	0.4146

5.3.3.2 Four-Band

Combination of Blue, Red, Green and NIR1 bands of WV-2 images are clustered and analyzed. Result of this module is presented below.



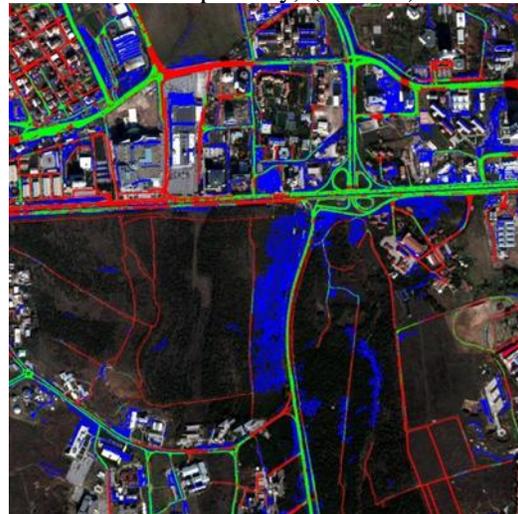
a) Clustering based road mask (all bands separately)



b) Clustering based road mask (all bands separately) (colored)



c) Clustering based road mask (all bands together)



d) Clustering based road mask (all bands together) (colored)

Figure 65 Clustering based result for four-band image

5.3.3.3 RGB Image

For RGB image, clustering based road extraction result is presented in Figure 66.

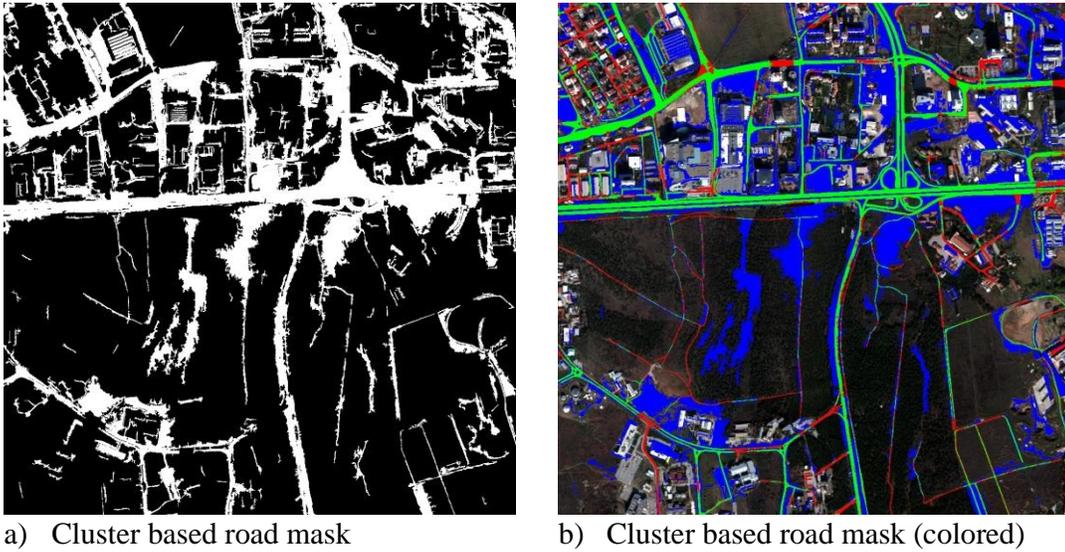


Figure 66 Clustering based result for RGB image

5.3.3.4 Gray Level Image

For Gray-level image, clustering based road extraction result is shown in Figure 67.

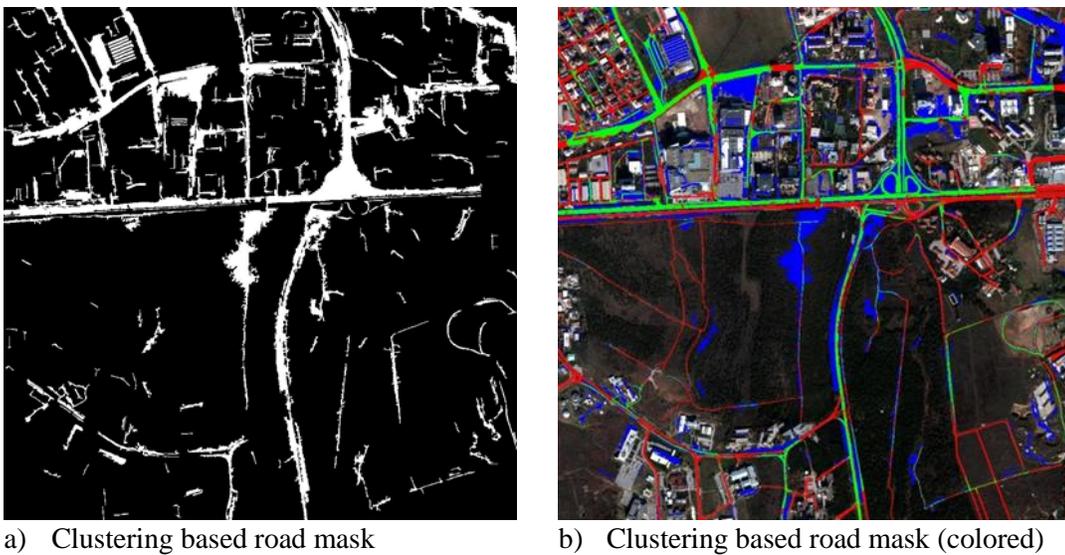


Figure 67 Clustering based result for gray image

5.3.3.5 Panchromatic Image

Clustering based road extraction module output is shown in Figure 68.

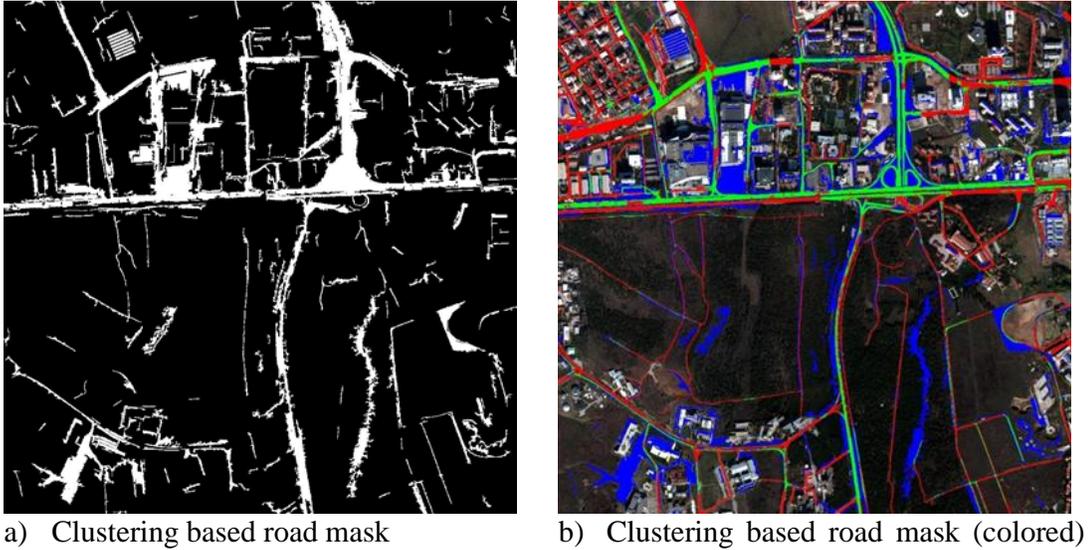


Figure 68 Clustering based result for panchromatic image

When all clustering results are quantitatively compared, it can be seen that, F_1 value is decreased while number of bands is increased. Since the algorithm fuses all band clustering results, the precision values is getting lower since the FPs are increased for each band. Similarly, recall value is increased with each additional band since the TPs are also increasing for each band. However, the increase in recall value is more significant than the precision value which results in lower F_1 value higher number of bands. The results are depicted in Table 9 and Figure 69.

Table 9 Combined results of clustering based road extraction module

	WV-2		Four-band		RGB		Gray		Pan	
	M-M	M-S	M-M	M-S	M-M	M-S	M-M	M-S	M-M	M-S
Precision	0,23	0,18	0,27	0,21	0,29	0,24	0,38	0,31	0,33	0,26
Recall	0,81	0,71	0,74	0,64	0,71	0,60	0,47	0,39	0,42	0,35
$F_{0,5}$	0,27	0,21	0,30	0,25	0,33	0,27	0,40	0,33	0,35	0,28
F_1	0,36	0,29	0,39	0,32	0,41	0,34	0,42	0,35	0,37	0,30
F_2	0,54	0,45	0,54	0,46	0,55	0,46	0,45	0,37	0,40	0,33

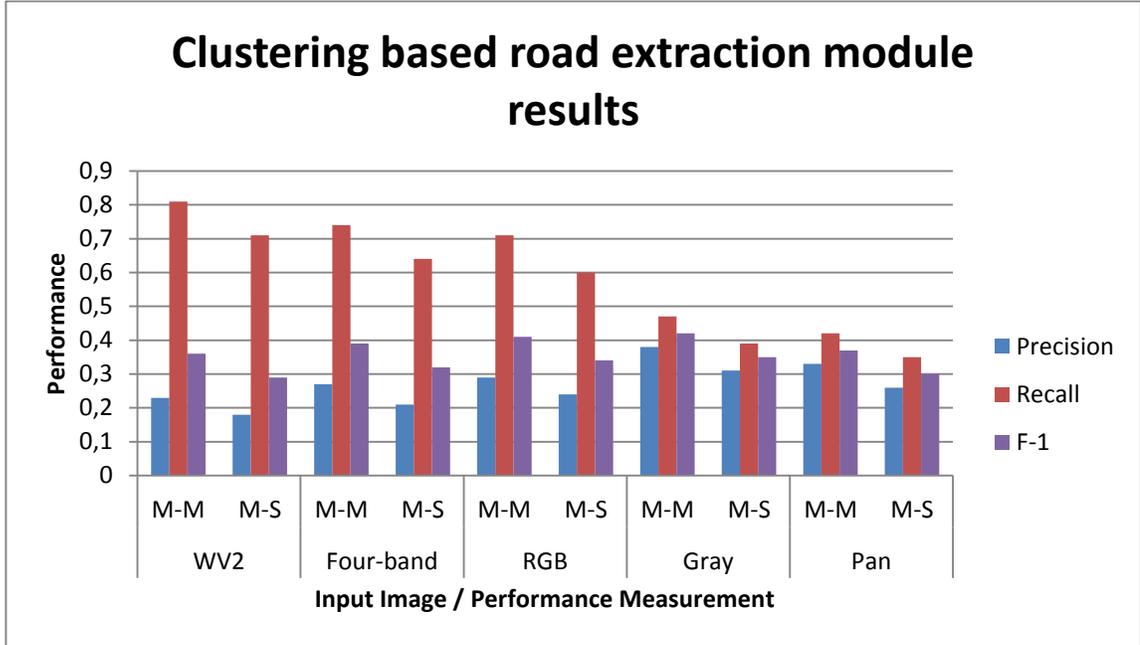


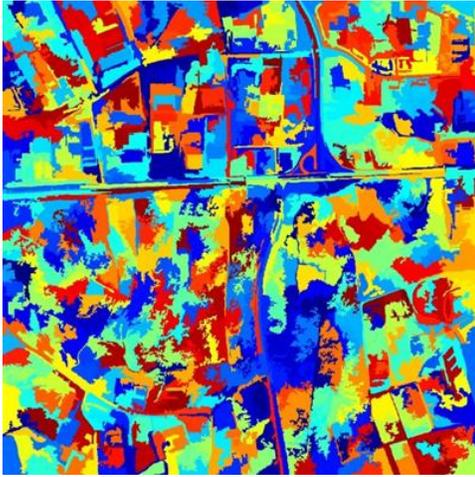
Figure 69 All result of clustering based road extraction module.

5.3.4 Structure Analysis Based Non-Road Detection

In this section, non-road analysis based on structure analysis will be investigated for all types of input images mentioned before. Images are segmented by using mean-shift algorithm. Non-road segments are expected to be thick enough as considered in Chapter 3. It is assumed that these segments are bare soil or uniform large areas that do not have road characteristics.

5.3.4.1 WV-2 Image

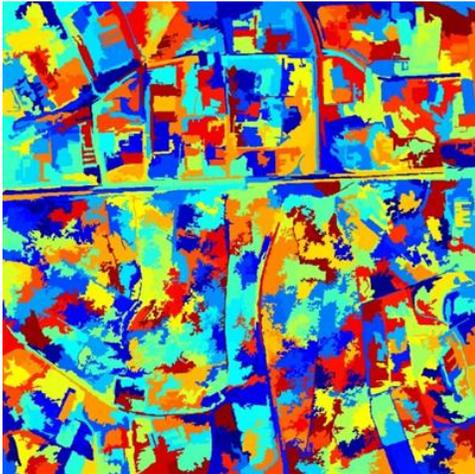
In Figure 70, structure analysis based non-road result is shown. In Figure 70(a), general segments and in Figure 70(b), a non-road mask is shown. For WV-2 images, two different four-band combinations are used for structure analysis for non-road detection. Then these two results are fused by union operation to obtain the final-non-road images.



a) Labels for CB,RE,NIR1,NIR2



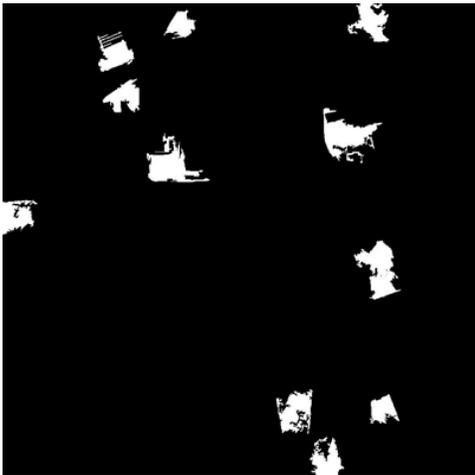
b) Structure analysis based road mask (colored)



c) Labels for B,G, Y,R



d) Structure analysis based road mask (colored)



Not-road structural analysis

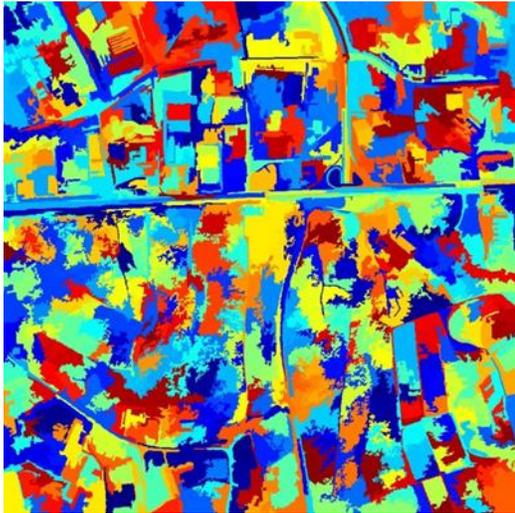


Not-road structural analysis (colored)

Figure 70 Structure analysis based non-road result

5.3.4.2 Four-band image

In Figure 71, structure analysis based non-road result is shown for four-band image which is simulated image of Geoeeye satellite.



a) Labels

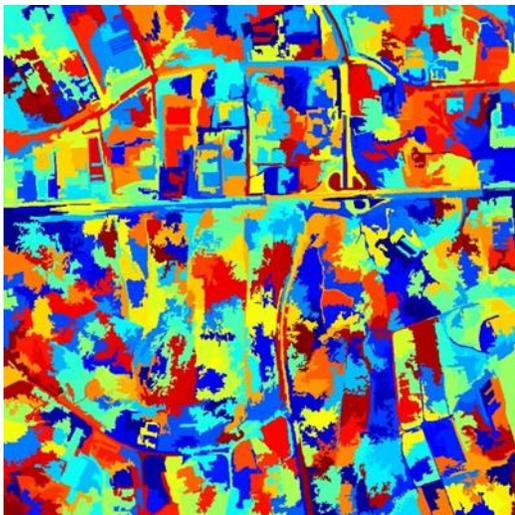


b) Structure analysis based non-road mask (colored)

Figure 71 Structure analysis based non-road result for four-band

5.3.4.3 RGB image

In Figure 72 structure analysis based non-road result is shown for RGB image.



a) Labels



b) Structure analysis based non-road mask (colored)

Figure 72 Structure analysis based non-road result for RGB image

5.3.4.4 Gray-level image

Structure analysis based non-road result for gray level images and segments are shown in Figure 73.

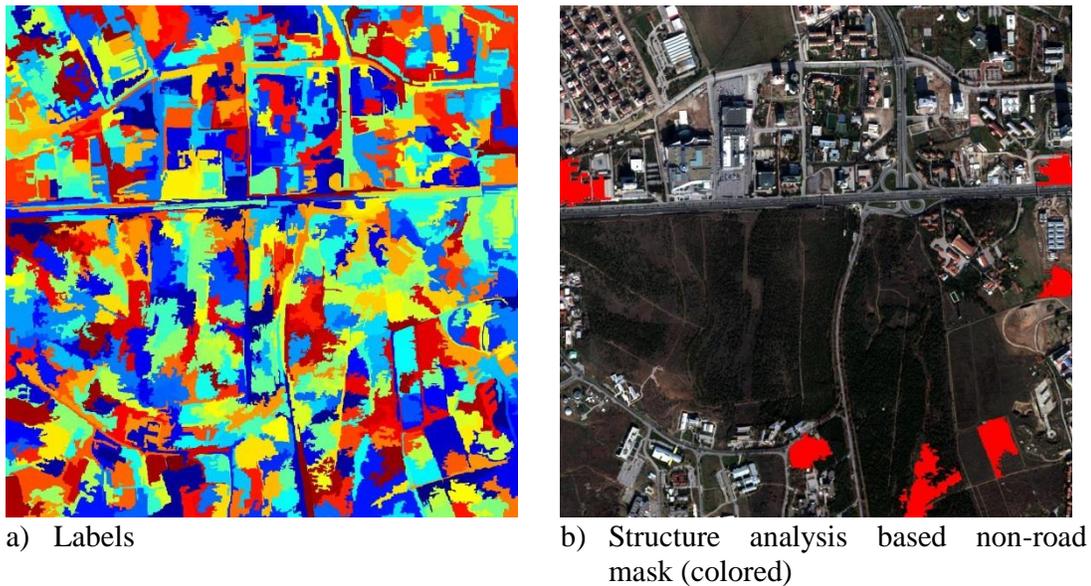


Figure 73 Structure analysis based non-road result for gray level images

5.3.4.5 Panchromatic image

Structure analysis based non-road result for panchromatic image and corresponding segments are shown in Figure 74.

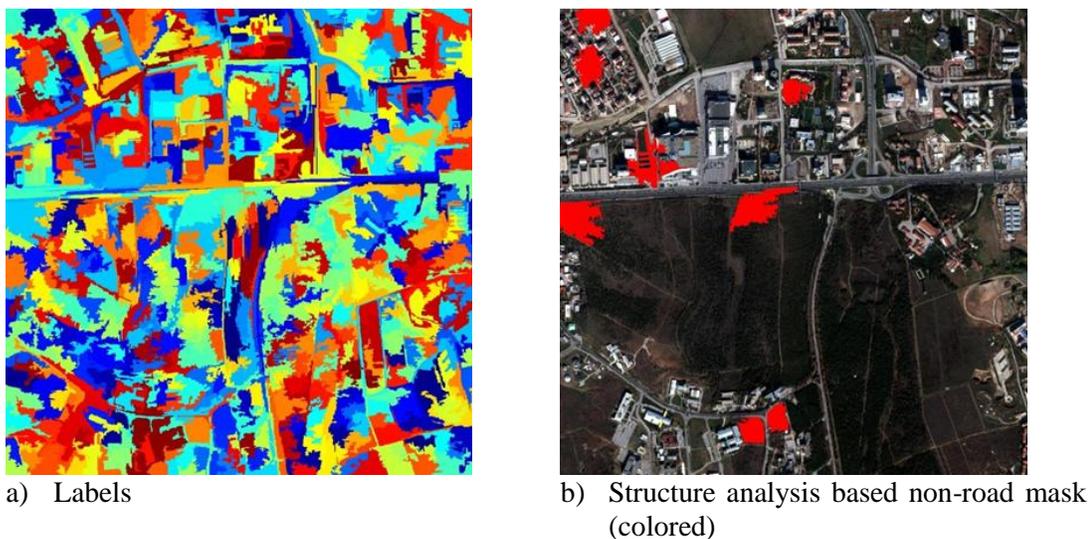


Figure 74 Structure analysis based non-road result for panchromatic image

5.3.5 Vegetation Detection

In this part of the study, result for vegetation analysis of WV-2 and four-band image will be reported.

5.3.5.1 WV-2 Image

NDVI and SAVI results are threshold by the predefined value (0.1) to obtain the vegetation mask. Results are presented in Figure 75.

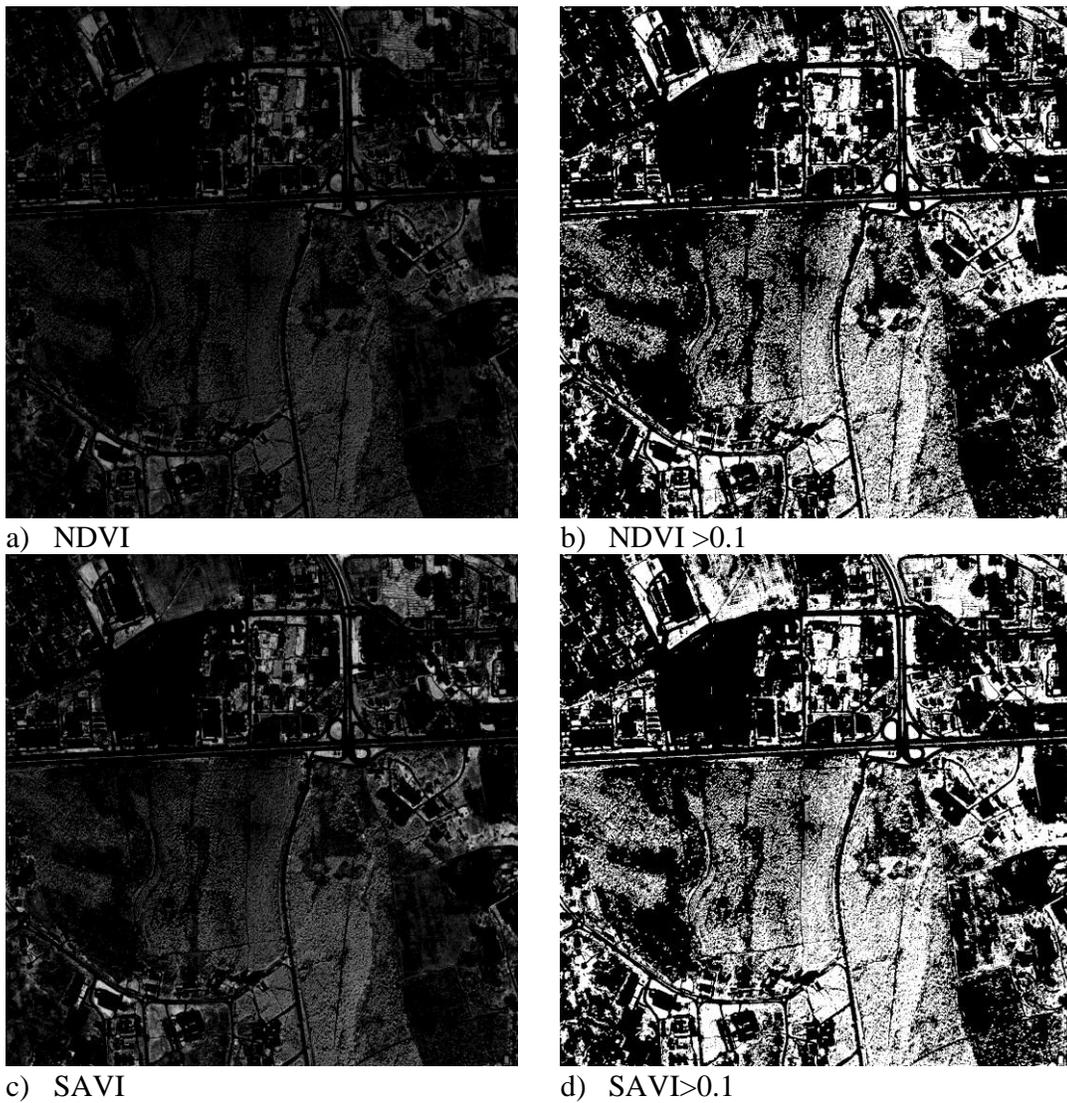


Figure 75 Vegetation results for WV-2 image.

5.3.5.2 Four-band image

Red and NIR1 values are used to calculate the indexes for simulated Geoeye image. Sample vegetation masks are shown in Figure 76.

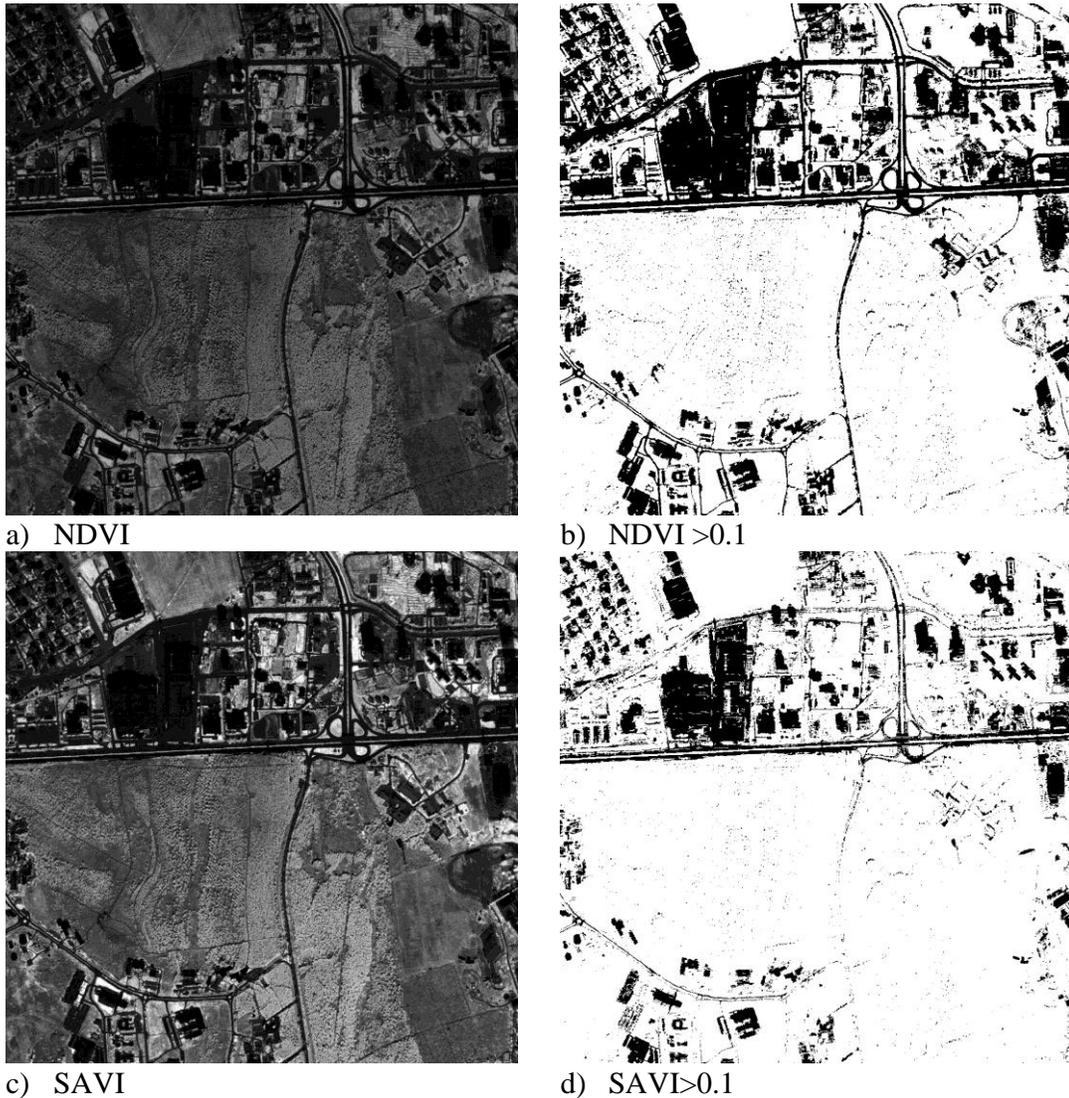


Figure 76 Vegetation results for four-band image.

5.3.6 Water Extraction

In this part of the study, water mask is aimed to extract as a non-road region. Since, there is no water region in the demonstration area, results will be shown other parts of images.

5.3.6.1 WV-2 Image

NDWI, NDVI results and NIR band are employed to detect water regions. In Figure 77 these values and water mask is presented.

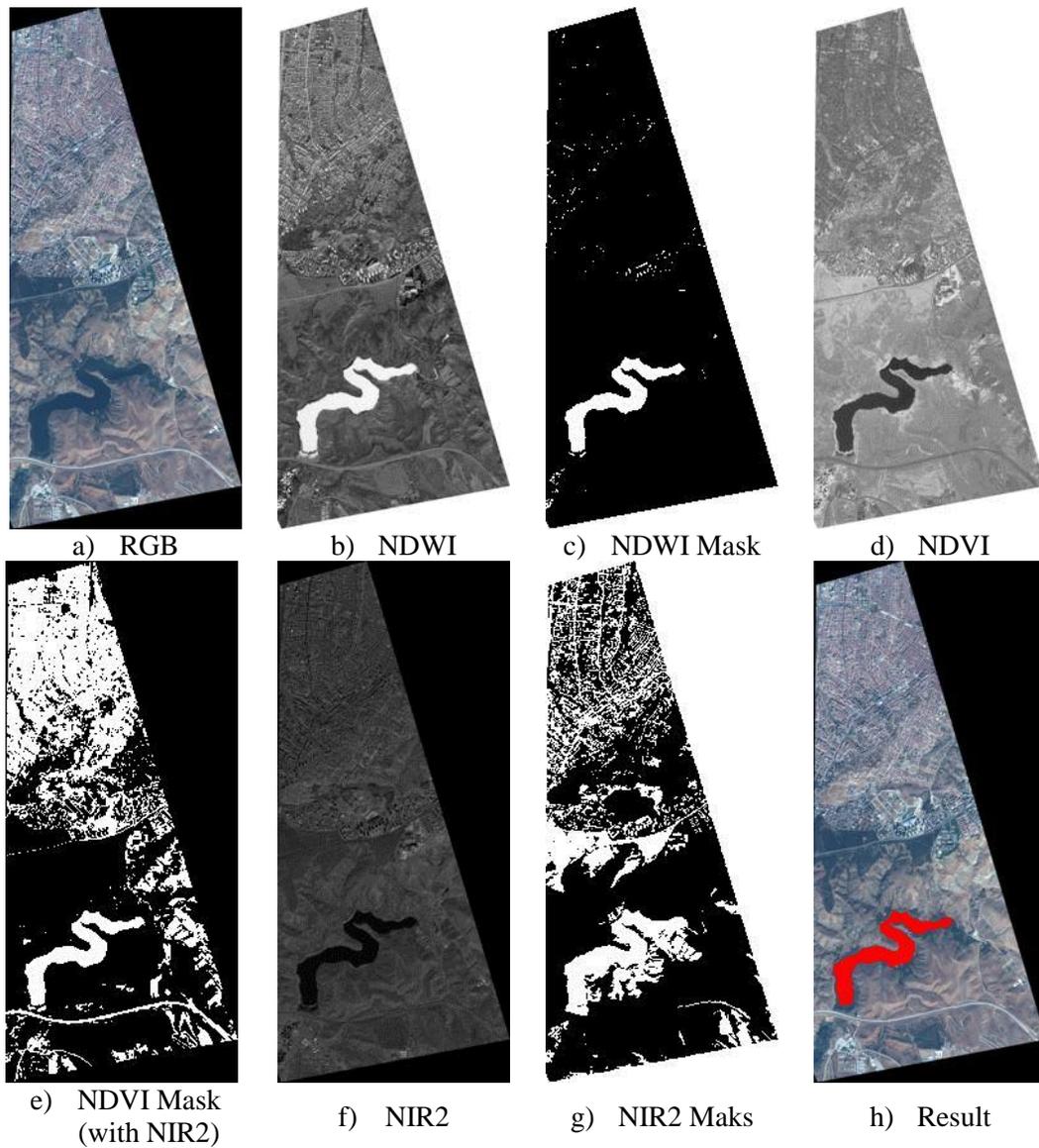


Figure 77 Water results for WV-2 image.

5.3.6.2 Four-band image

Blue, Red, and NIR1 values are used to calculate the indexes for simulated Geoeye image. Results are almost the same with the water region result for WV-2 Image. NDWI, NDVI results

and NIR bands and corresponding automatic hysteresis thresholded masks and final result on RGB image is shown in Figure 78.

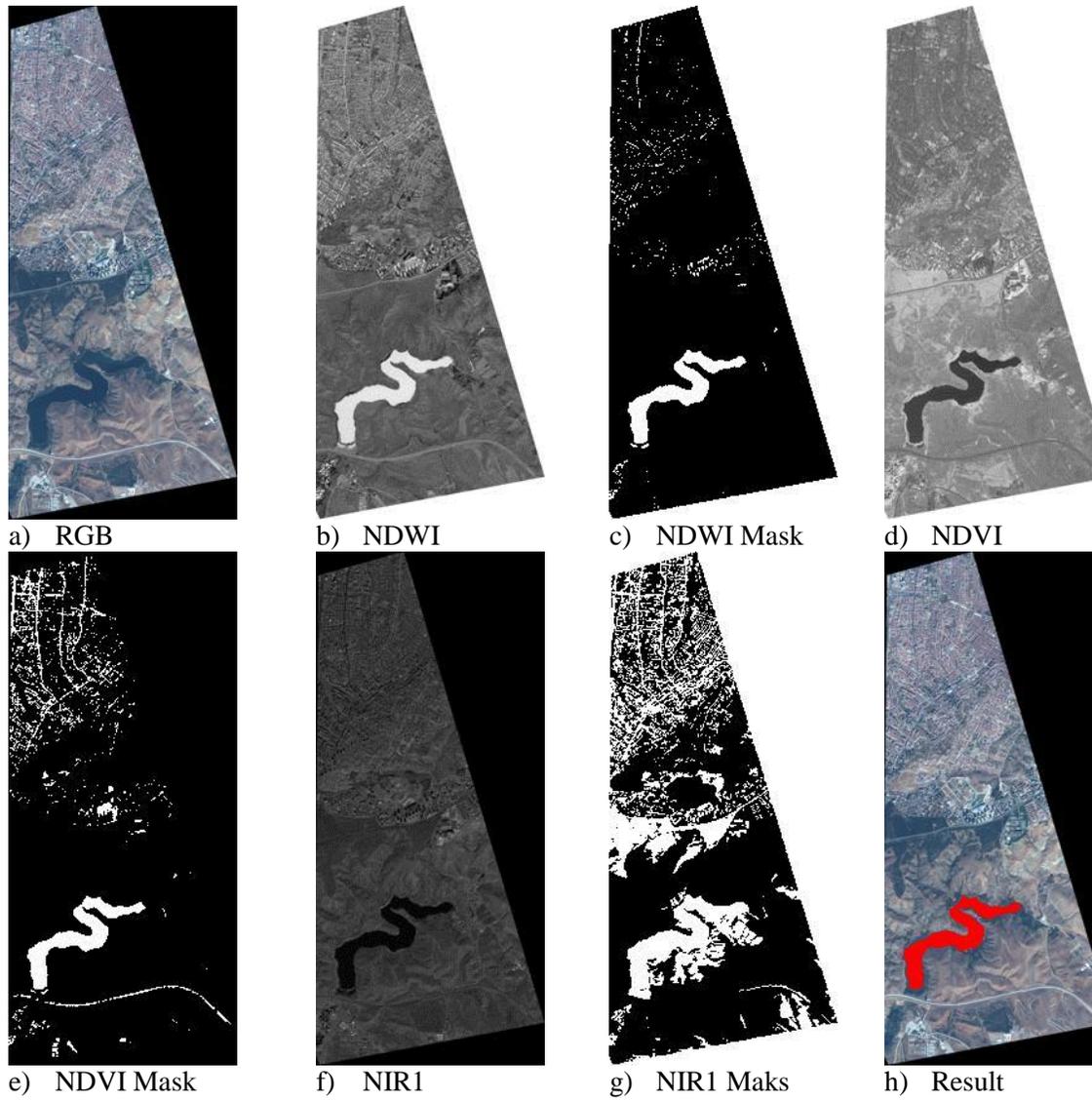


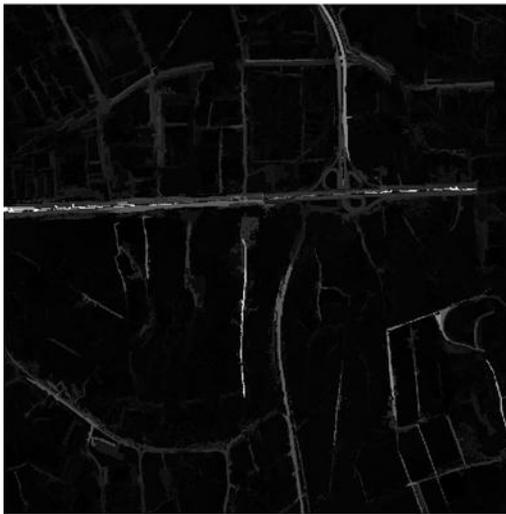
Figure 78 Water region detection result for four-band.

Additional result for water extraction module presented in APPENDIX C.

5.3.7 Score map Construction

5.3.7.1 WV-2 Image

In Figure 79, Score map and thresholded road network is given.



a) Score Map based on segment

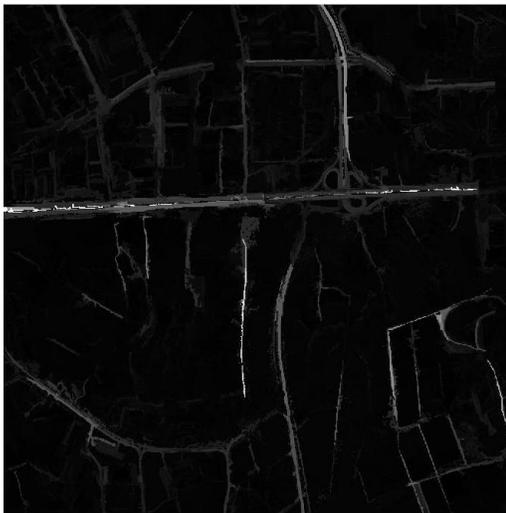


b) Final Result

Figure 79 Road Network result for WV-2 image.

5.3.7.2 Four-Band Image

In Figure 80, Score map and thresholded road network for four-band image is given.



c) Score Map based on segment

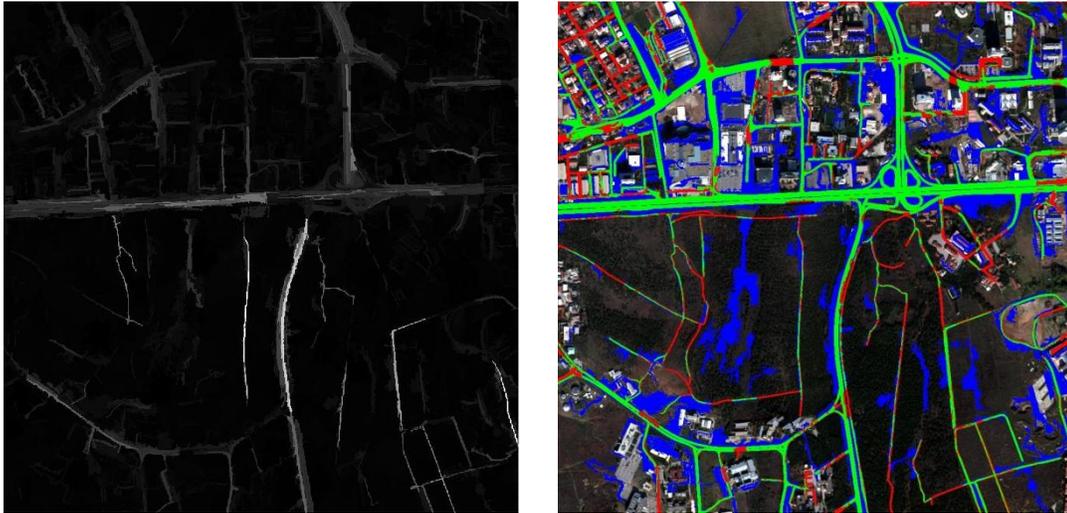


d) Final Result

Figure 80 Road Network result for Four-band image.

5.3.7.3 RGB Image

Score map and thresholded road network of the algorithm for RGB images is presented in Figure 81.



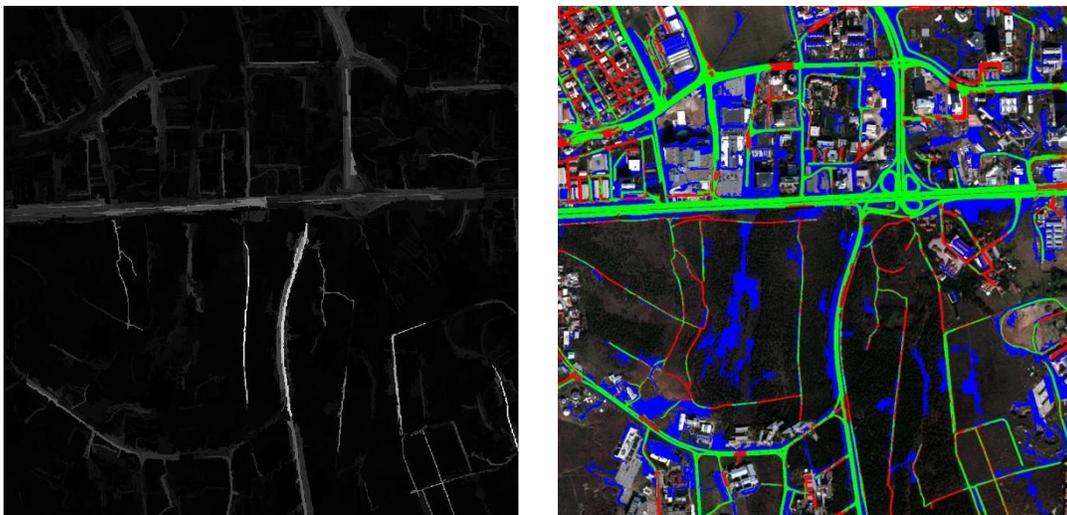
a) Score Map based on segment

b) Final Result

Figure 81 Road Network result for RGB image.

5.3.7.4 Gray Level Image

Score map and thresholded road network of the algorithm for Gray Level images are presented in Figure 82.



c) Score Map based on Segment

d) Final Result

Figure 82 Road Network result for Gray image.

5.3.7.5 Panchromatic Image

Score map and thresholded road network of the algorithm for Gray Level image is presented in Figure 83 .

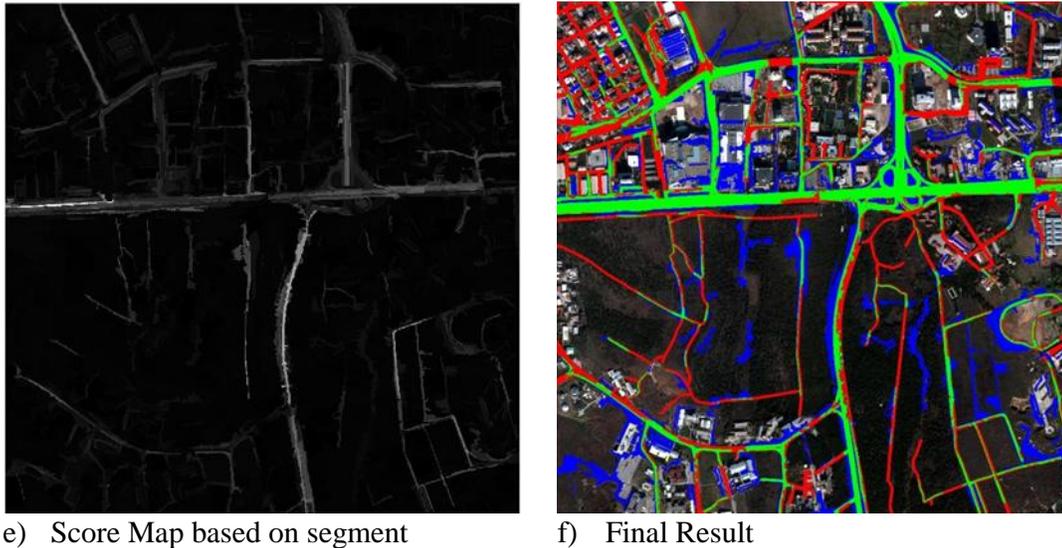


Figure 83 Road Network result for Panchromatic image.

5.3.8 Overall Result

As it is seen in Table 10 and Figure 84, when WV-2 image is used as an input for the proposed algorithm, higher recall is obtained. Even precision is higher for low level images, when we look at F_1 -measure which gives the same importance to recall and precision, it can be inferred that the more number of band is employed, the higher performance is obtained.

Table 10 Road extraction results for all bands

	WV-2		Four-band		RGB		Gray		Pan	
	M-M	M-S	M-M	M-S	M-M	M-S	M-M	M-S	M-M	M-S
Pre	0,61	0,47	0,44	0,35	0,40	0,33	0,41	0,37	0,49	0,43
Rec	0,51	0,70	0,57	0,67	0,49	0,63	0,43	0,59	0,51	0,57
$F_{0.5}$	0,58	0,51	0,46	0,38	0,41	0,37	0,42	0,40	0,49	0,45
F_1	0,55	0,57	0,50	0,46	0,44	0,44	0,42	0,46	0,50	0,49
F_2	0,52	0,64	0,54	0,57	0,47	0,54	0,43	0,53	0,50	0,54

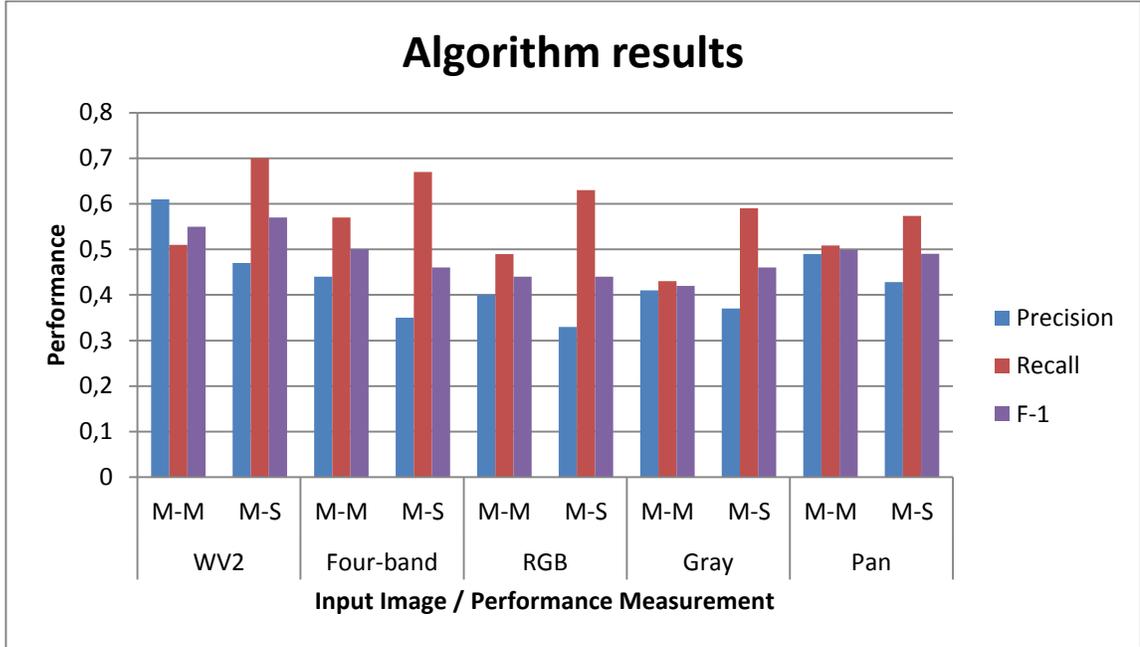


Figure 84 Algorithm Results

5.4 Road Material Classification

Road material classification has been conducted on WV-2 and four-band images. The training set is obtained from the image shown in Figure 41 of Section 4.8. Road Material Classification for WV-2 Images is presented in Figure 85. Yellow parts correspond to unpaved road regions whereas blue parts represent asphalt covered roads.

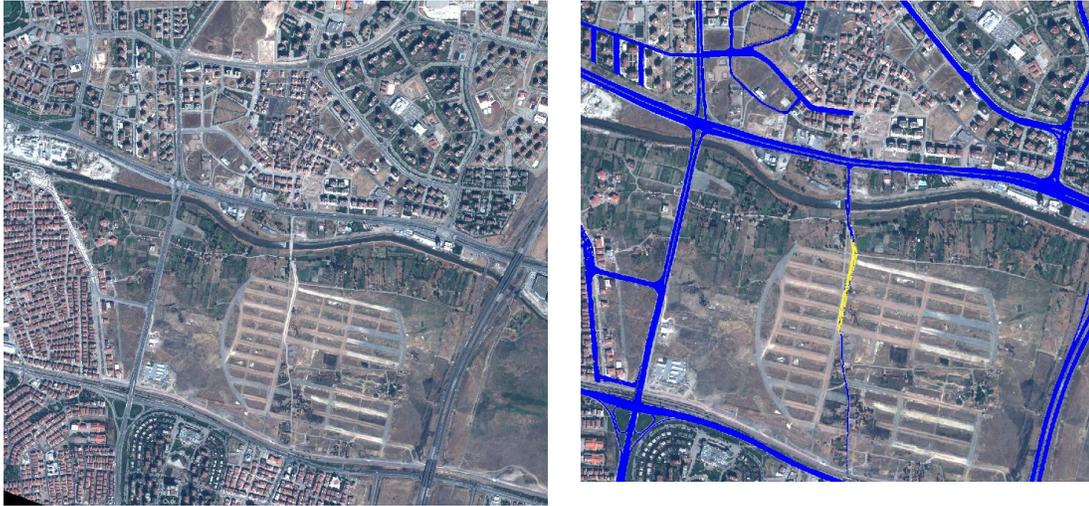


Figure 85 Road Material Classification Result for WV-2 Images

Classification for four-band Images is presented in Figure 86.



Figure 86 Road Material Classification Result for four-band Images

Another result used for road type classification is shown in Figure 87. The algorithm is tested both on ground truth and road extraction algorithm. Yellow parts correspond to unpaved road regions whereas blue parts represent asphalt covered roads.



a) RGB image



b) Ground Truth representing road classes



c) Road type classification result on road GT



d) Road type classification result on the road extraction algorithm output

Figure 87 Road type classification result on GT and road extraction algorithm output

As shown in Figure 87, some regions of asphalted roads are classified as un-paved roads. This case occurs for narrow roads which are either deteriorated or mixed with neighboring soil regions. Some miss classified regions may be due to the fact that road mask may also cover unpaved parts adjacent to the asphalted roads.

5.5 Water Canal Extraction

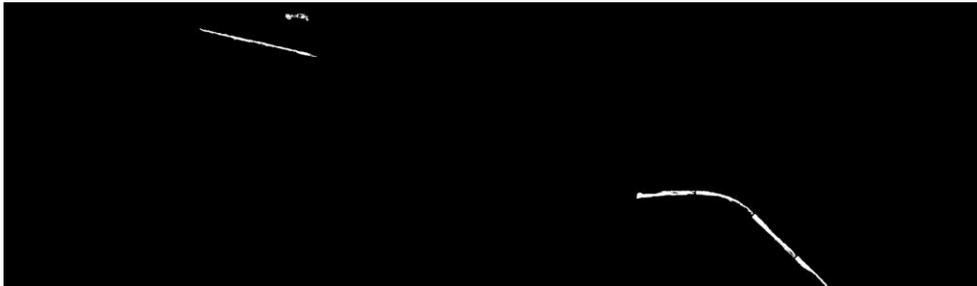
Water regions are found as mentioned Chapter 4. In order to find water canals, elongatedness analysis is applied. All results related to water canal extraction is shown in APPENDIX D.

5.5.1 WV-2 Image

Sample image and water canal detection result for WV-2 images is shown in Figure 88



a) RGB



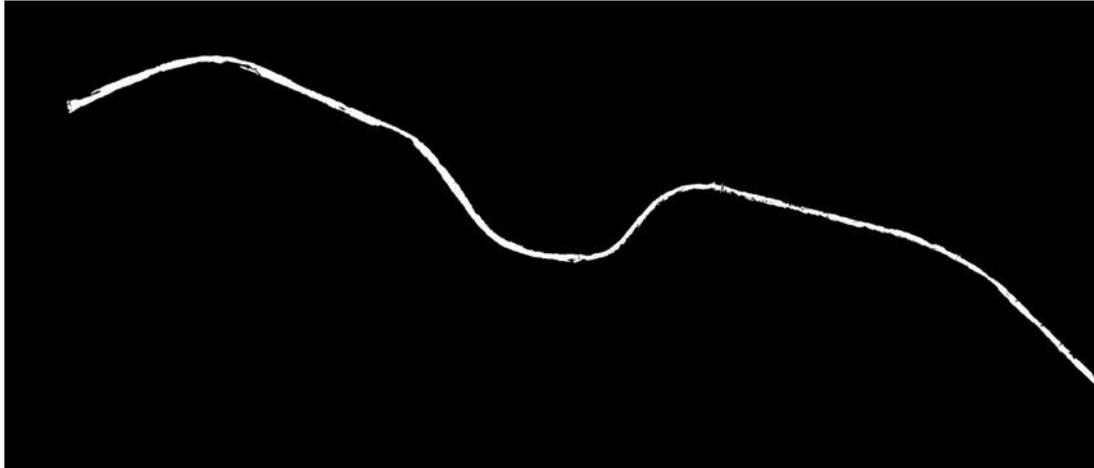
b) Water Mask

Figure 88 Water results for WV-2 image.

When the output is analyzed, it can be seen that, unextracted water canal segment is drier than the extracted parts. This affects the NDWI values and automatic hysteresis thresholding method.

5.5.2 Four-band

Sample image and water canal detection result for four-band image is shown in Figure 89.



a) Water canal mask



b) Water canal mask (colored)

Figure 89 Water results for four-band image.

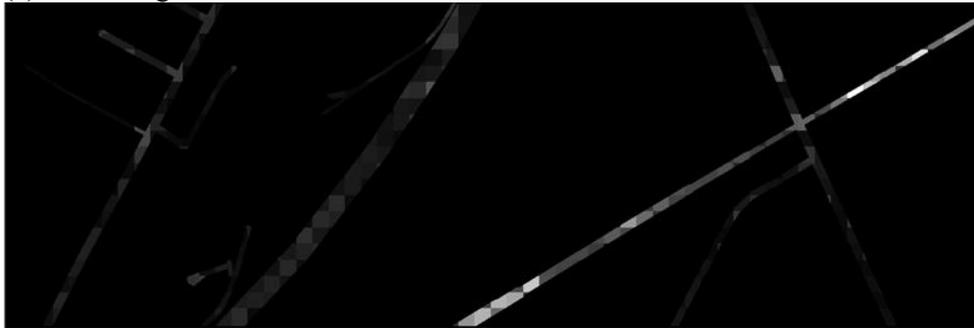
5.6 Railroad Extraction

This approach has been tested on four different 20 cm resolution Google Earth images including both railroad and road structures. Since the proposed method relies on the sleepers' periodicity as a distinctive features of railroads, high resolution images are needed. Some of the sample results can be seen at Figure 90.

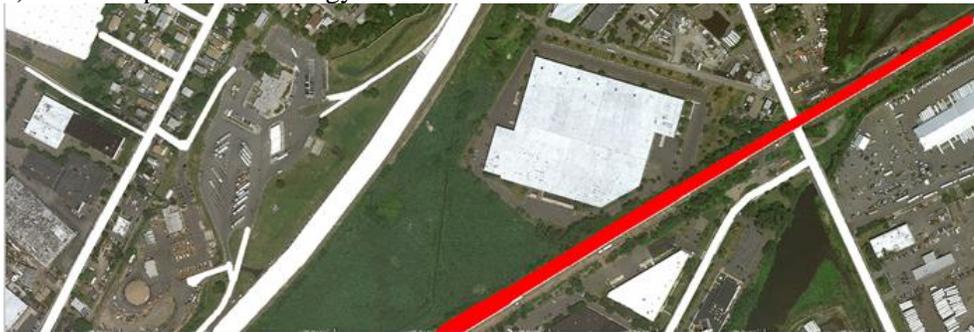
Generally, the algorithm achieves good qualitative results for most circumstances. However in some cases because of the railroad intensity and the deterioration of sleepers' structure, the proposed method may not be successful in distinguishing railroads from other regions.



(a) RGB Image



(c) Score map based on energy



(b) Ground truth. Red parts represent railroads and blue parts represent railroads.



(d) Final Result

Figure 90 Railroad result

As it is seen in Figure 90, some parts cannot be classified. This may be because of bridges. Another factor that affects the result is pillars. Some of those and their shadows occlude some part of railroad. On the other hand, since the approach uses all road map for classification, some zebra crossing on drive roads can be extracted as railroad because of periodicity. Another results of the railway classification is shown in Figure 91 and Figure 92 .



(a) RGB Image



(c) Score map based on average energy



(b) Ground truth. Red parts represent railroads and blue parts represent railroads.

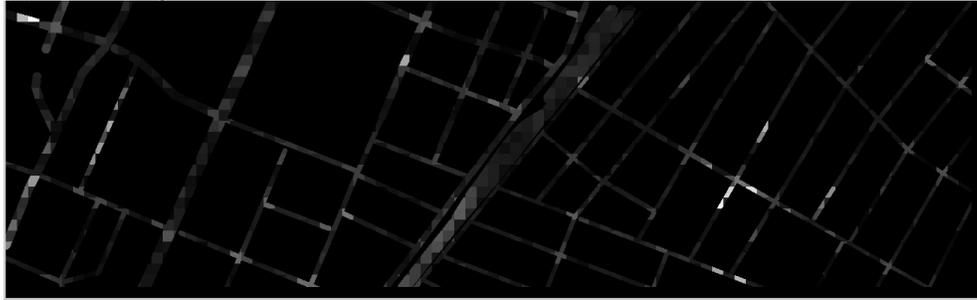


(d) Final Result

Figure 91 Railroad classification result.



(a) RGB Image



(c) Score map based on average energy



(b) Ground truth. Red parts represent railroads and white parts represent railroads.



(d) Final Result

Figure 92 Sample railroad classification result

As seen in the score maps depicted in Figure 90 (c) Figure 91(c) and Figure 92(c), some drive-road parts depict railroad characteristics. This may be because of local periodic structure like zebra crossing, road markings or cars. Another problematic case occurs at junction points for

which it is difficult to detect road direction. Some sample regions are shown in Figure 93. Even though the highest energy is not in the ROI, still there may be higher energy than the background. In Figure 93 (d), parking lines have a periodic structure. Although they do not have the same periodicity with railroad sleepers, the harmonic components of their Fourier spectra can be located at the ROI.

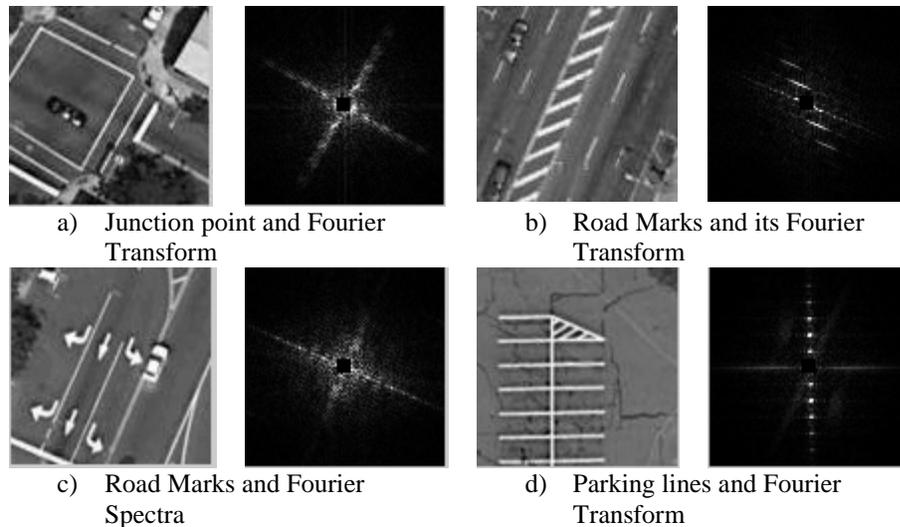


Figure 93 Sample road parts having periodic structure

After thresholding the score map, such cases are classified as railroad. Such disconnected regions are eliminated in post-processing phase by removing small unconnected components as demonstrated in Figure 38 of Section 4.7.

Some railroad parts, on the other hand, may not have enough energy to classify them as railroads. Such railroads may have deteriorated sleepers or they are constructed with a technology that does not incorporate sleepers as shown in Figure 94.

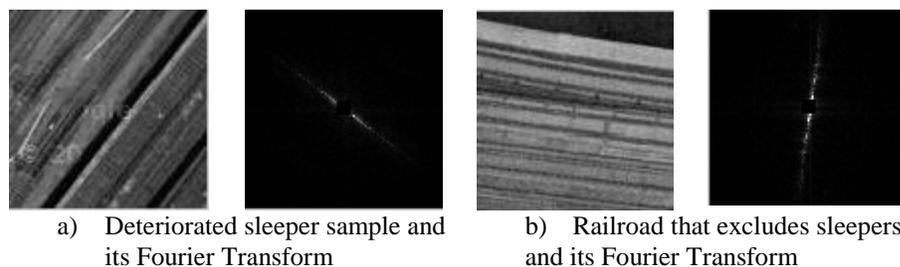


Figure 94 Sample railroad parts without discernable periodicity

5.7 Discussion

In this study we aimed to develop a road network extraction algorithm from remotely sensed multi-spectral data. In order to make the model flexible, a modular approach has been proposed. In this part of the study, each module of the model will be discussed. Each module has some parameters determined according to the road assumptions. The effect of changing the values of these parameters will be investigated in this part of the study. In addition, results will be compared with the existing methods qualitatively and quantitatively.

The first module of the proposed model is edge based road extraction. This module consists of three main steps i.e. edge detection, center line extraction and middle line selection.

Canny edge detector has been used to find contours of the gray level images. This edge detector requires two different parameters. First parameter is the hysteresis threshold (determined by low and high thresholds), to detect edges from gradient magnitudes of each band. Second parameter is sigma that refers to the standard deviation of the Gaussian filter. Setting these parameters to pre-determined fixed values may cause the detection of edges only with a specific gradient value. In order to avoid this problem, these parameters should be set automatically for each image. Automatic thresholds are determined based on the highest value of gradient magnitude of the image and the low threshold equals to 40 % of highest gradient magnitude threshold in this study. Sigma is set as “1” in this study.

For center line extraction we used two parameters, maximum width and minimum width. These values are dependent on the image resolution. Since this study aims to extract roads with width in the range of 4 meters and 80 meters, these values are set as 2 and 40 for the 2 meters resolution multi-spectral images. Middle points are selected as mentioned in Chapter 4. However, there may be some gaps between these middle points. In order to fill these gaps, an approach that checks the orientation of each line and searches for continuing pairs in a pre-defined range is developed.

As it is shown in Figure 95, dashed lines represent the link connecting two middle lines with same orientation and the given gap.

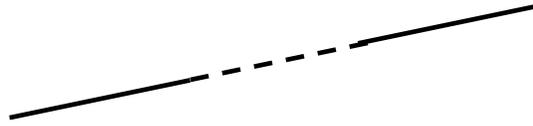
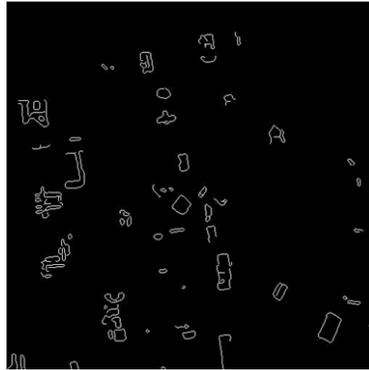


Figure 95 Edge linking sample

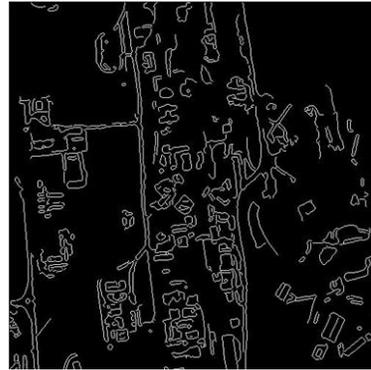
When this approach has been tested on images and it was observed that, this method also merges middle points extracted from buildings in a row. Other approaches such as spline, line fitting, or the one used in Li et al. (2003), have the same problem. We decided to exclude this property from the module in order to reduce FP and increase the precision value.

In order to select appropriate middle lines, small segments are discarded. We aimed to remove short lines obtained from non-road structures such as buildings. Another parameter used in edge based road extraction module is “line length” to be removed. It is assumed in this study that road parts should be greater than 60 meters. This may result in the elimination short road parts obtained by edge based module. On the other hand, if they are not removed many small parallel structures are considered as road segments inducing lower precision.

Also, in the edge based road extraction module each spectral band is processed separately. It is observed that, Canny edge detection result is different for each spectral band (Figure 96). This makes us to consider edge map merging of each band to provide smoother and more complete middle line results.



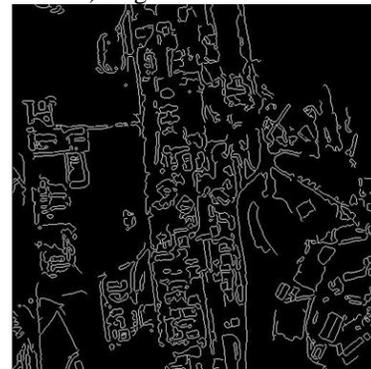
a) Edges coastal blue band



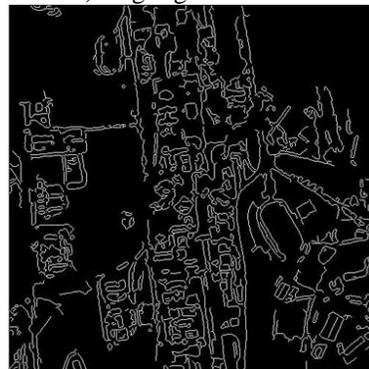
b) Edges of blue band



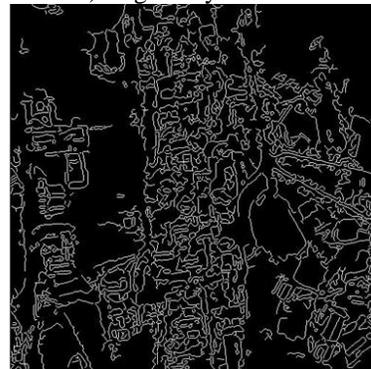
c) Edges green band



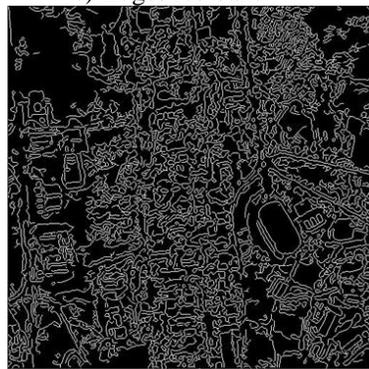
d) Edges of yellow band



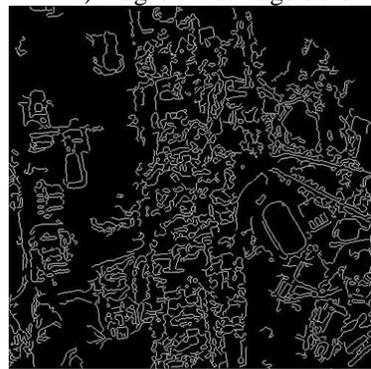
e) Edges red band



f) Edges of red edge band



g) Edges NIR-1 band



h) Edges NIR-2 band

Figure 96 Edge detection result for each band

After merging operation is employed for all edge maps, it is observed that a registration problem occurs. There may be two possible sources of this problem. One is that the gradient may be different in the same region for each band. Second is that the automatic gradient threshold detection may produce different threshold values for different bands. In order to eliminate the second problem, tests with fixed threshold values were conducted and it was observed that the registration problem still could not be solved. This problem also hampers us to employ edge voting method to get all related edges from different spectral characteristics.

Sample result for edge map merging is shown in Figure 97

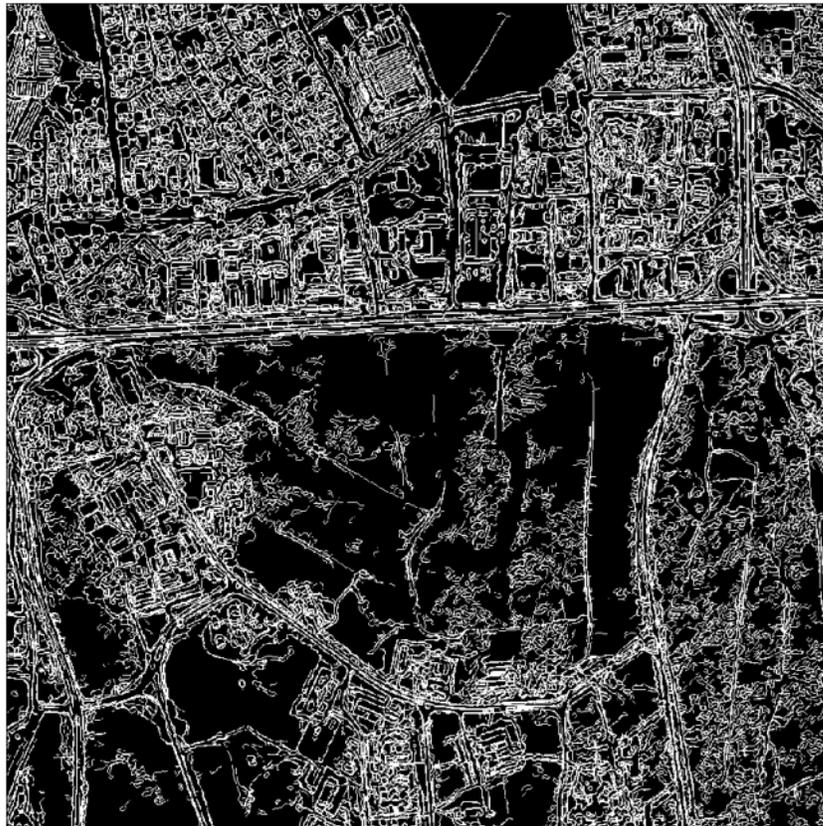


Figure 97 Edge registration problem

In structure analysis based road extraction, mean-shift segmentation is employed. This method needs three parameters: Min Region Area, Spatial bandwidth, Range bandwidth. “Spatial bandwidth” and “Range bandwidth” parameters are optimized in Çinar (2012) as 8, 4 pixels respectively for 1024X1024 IKONOS images. After segmentation is completed, segment analysis is conducted by using elongatedness and eccentricity analysis. Before computing these

values, each segment is checked to ensure its area is large enough. However, in some cases road parts may be eliminated because of segment area threshold. In order to solve this problem, some segments that have lower areas are considered as road segments, if they are elongated enough. Table 11 shows the decision matrix for segment analysis using area and elongatedness threshold values.

Table 11 Segment analysis decision matrix for area and elongatedness values.

Area/Elongatedness	High	Low
High	Road	Not-Road
Low	Probably Road	Not-Road

In Figure 98, mentioned metrics are shown in color.

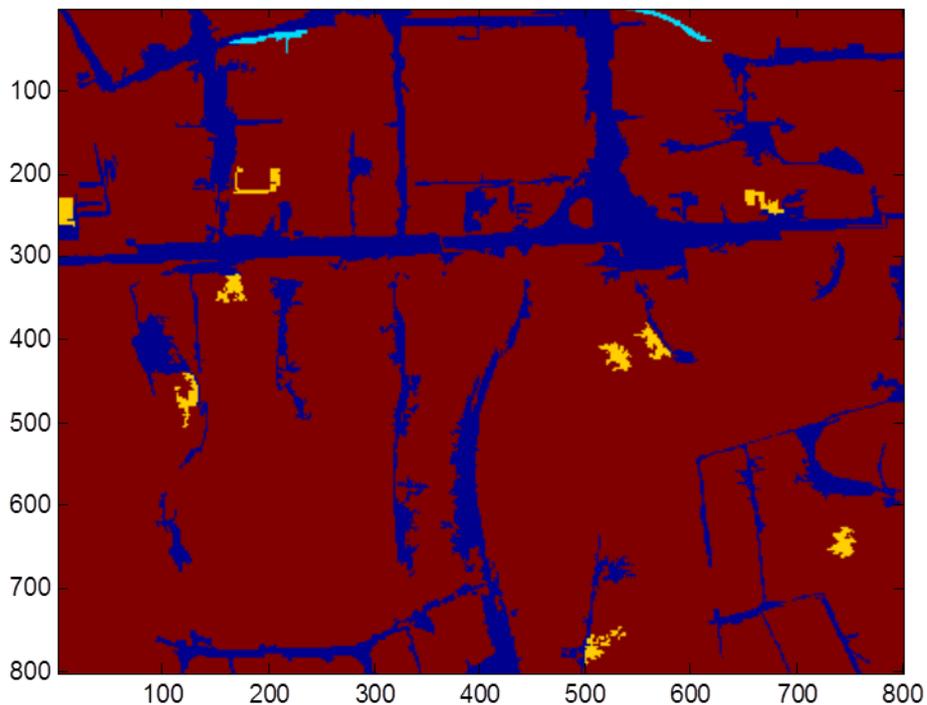


Figure 98 Elongatedness and Area analysis

In Figure 98, yellow parts represent low area and low elongatedness, dark blue parts represent high area and high elongatedness. On the other hand, light blue parts represent high elongatedness and low area, while red regions represent high area and low elongatedness. In the experiments it is observed that, classification of segment elongatedness structure in these classes

is not needed, since pre-defined thresholds already provide the satisfactory solution. For example, structure analysis also produces light blue regions as roads for the above image.

In the clustering part, images are clustered into six clusters as suggested by Zhang and Couloigner (2006). Instead of selecting one cluster for road extraction, each component in each cluster is analyzed to detect roads as mentioned in Chapter 4. It should be noted that more clusters may also be used in our approach since all clusters are analyzed separately. For different K values, clustering and cluster analysis results are shown in Figure 99.

Increasing the number of clusters results in higher recall but lower precision values. Therefore, we set the value of K to “6”.

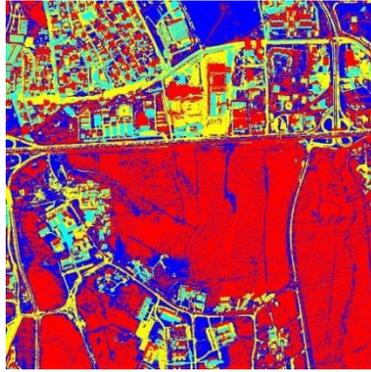
Non-road regions including bare soil, water and vegetation for multi-spectral images are also extracted in this study. These non-road regions can be excluded from the road region detection part of the model. In this manner, structure of the model can be changed and turned into two main consecutive steps i.e. non-road and road region detection parts. This means that, non-road regions can be excluded from all road part detection module result. This also affects the score map construction method. Instead of using non-road result as a reduction factor for the score map, they may not be considered at all as conducted in the study of Shackelford et al. (2003).

However, such a consecutive structure means that overall performance highly depends on the non-road region extraction performance. We conclude that, the modular approach is much more reliable to use.

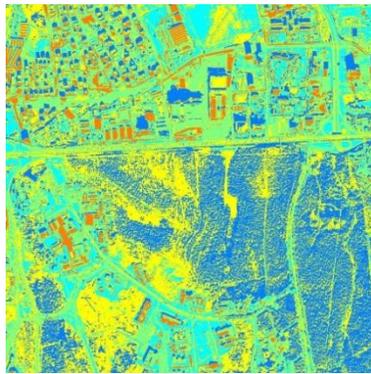
As mentioned in Chapter 4, vegetation, water, and large regions assumed to be bare soil are extracted as non-roads. For vegetation detection different indexes can be employed in this study. Since SAVI requires soil adjusted factor, it depends on image characteristics. EVI index can also be employed. For empirical tests, EVI index is calculated as follows,

$$EVI = G * \frac{NIR-Red}{NIR+c_1*Red-c_2*Blue+L}, \quad (\text{Equation 5.1})$$

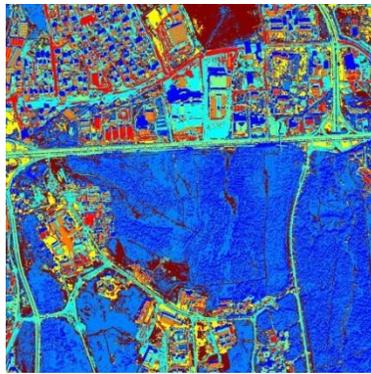
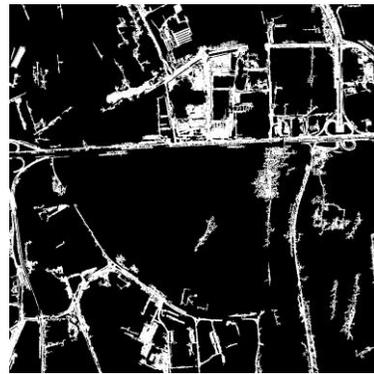
where L is a canopy background adjustment, G is a gain factor and c_1 , and c_2 are the coefficients. In MODIS-EVI algorithms these coefficients are defined as $L=1$, $c_1 = 6$, $c_2 = 7.5$, and $G = 2.5$. However, it is empirically observed that NDVI index is more stable and usable for all type of images and textures.



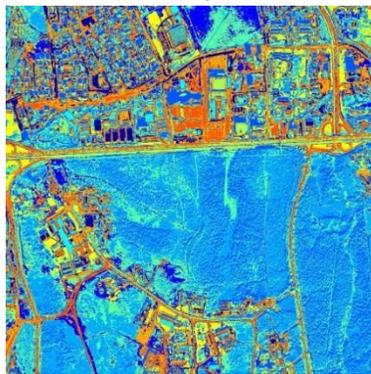
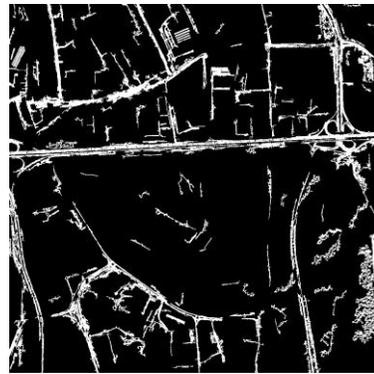
K=4



K=5



K=8



K=10



Figure 99 Clusters and clustering result for different K values

Comparison of NDVI and EVI projections and thresholded sample images are shown in Figure 100.

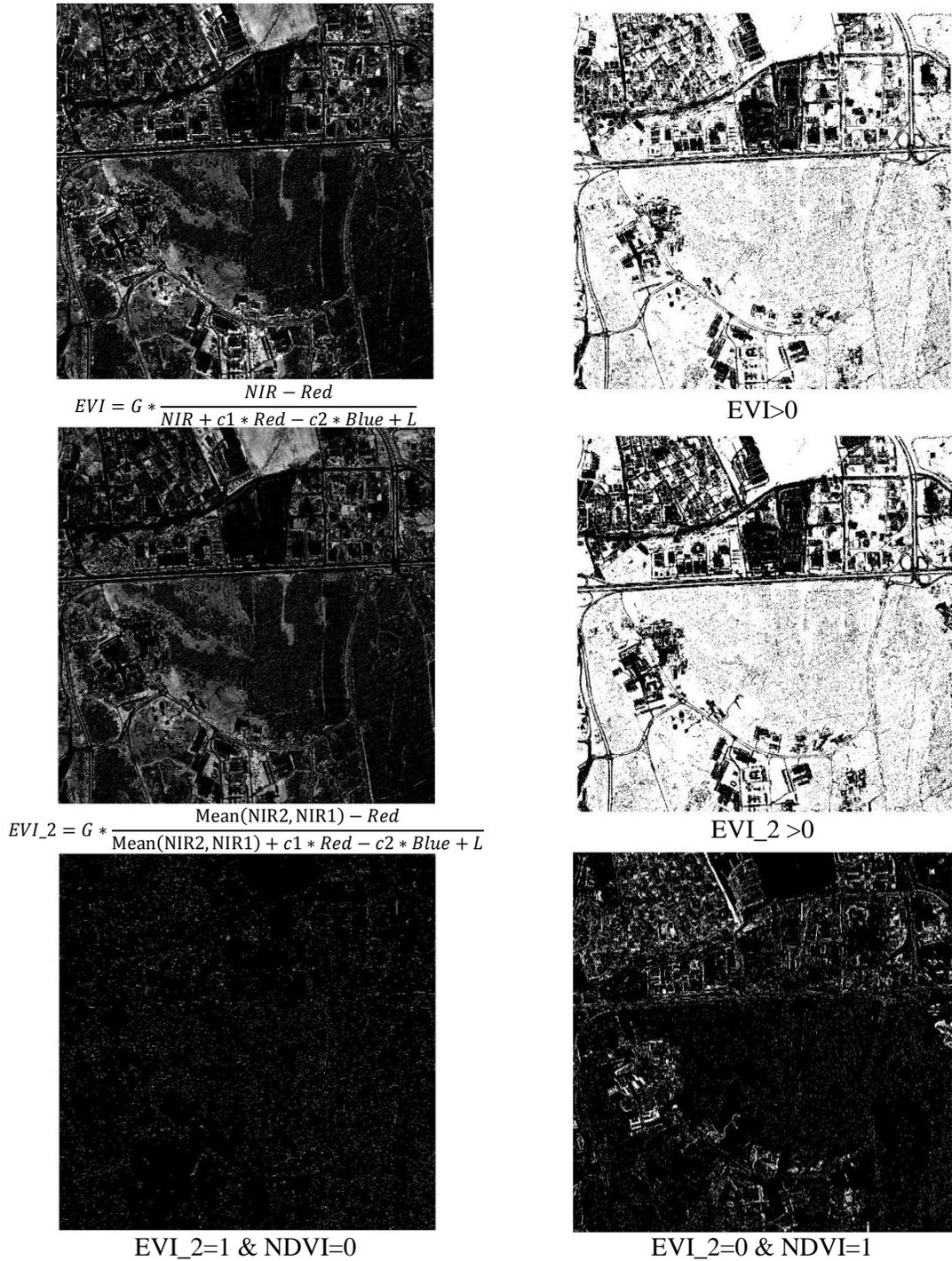


Figure 100 Comparison of EVI index calculated by NIR2 and average of NIR1 and NIR2

Segmentation based non-road region detection module also needs some parameters. One is the minimum region area for the mean-shift algorithm. This parameter is set as 750 for 1024X1024 and 2m resolution images. A higher “minimum region area” threshold may cause the loss of road parts since it forces the merging of segments. Different minimum region area values have been tested. It was qualitatively observed that minimum region area threshold should not be much higher or lower.

Band ratios may be used for segmentation to extract road and non-road regions. In order to test this idea, NDVI index result is segmented by using mean-shift algorithm. In Figure 101, road and non-road analysis based on NDVI segmentation is represented.

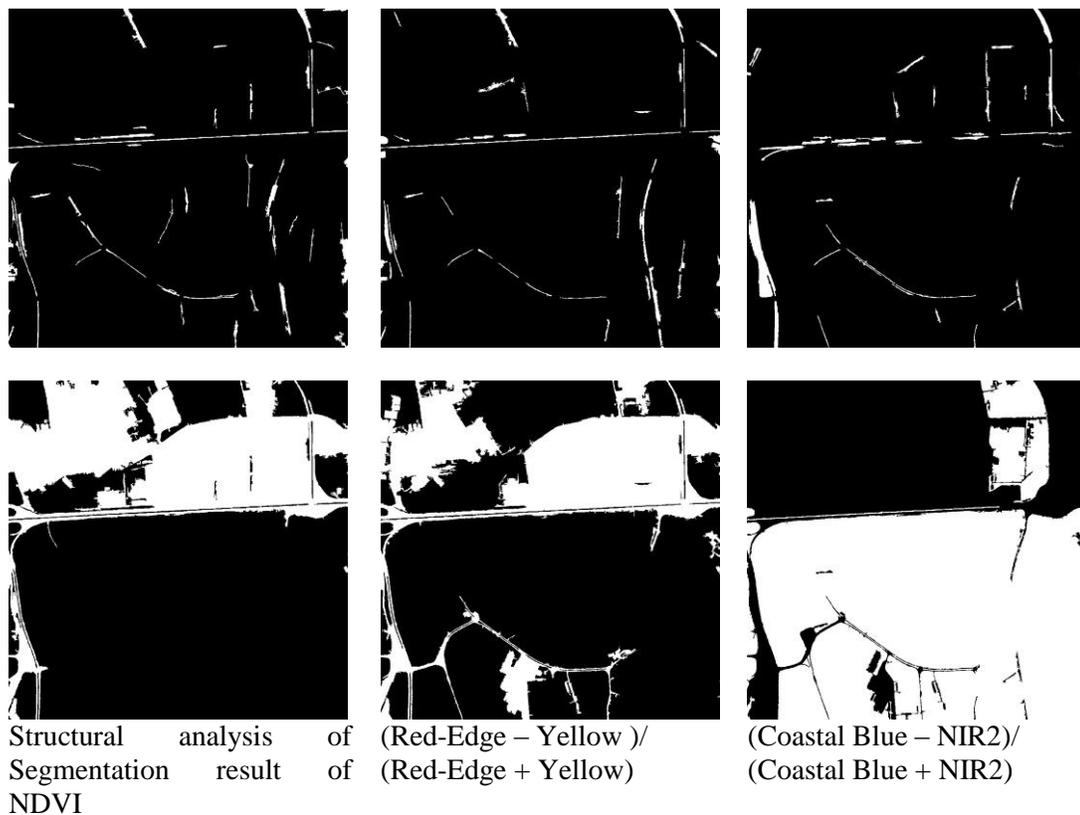


Figure 101 Band ratios segmentation result.

It can be shown that, even though segmentation analysis of spectral information ratios provide better result, they may not be stable as shown for not-road region detection parts.

In the light of this information, the model is finalized as stated in Chapter 4. All results are presented in APPENDIX A. All parameters used in this study are also shown in APPENDIX E.

Each module is investigated in terms of contribution to the final result. In other words, different combination of module fusing is also analyzed. In Table 12, different combination of road extraction modules and corresponding result is shown.

Table 12 Comparison of modules

Edge Based	Segment Based	Clustering Based	Precision	Recall	$F_{0,5}$	F_1	F_2
X			32	45	33,96	37,40	41,62
	X		72	23	50,49	34,86	26,62
		X	70	18	44,37	28,64	21,14
X	X		46	68	49,18	54,88	62,06
X		X	50	50	50,00	50,00	50,00
	X	X	53	59	54,10	55,84	57,69
X	X	X	45	75	48,91	56,25	66,18

We also analyzed the outputs of different module combinations visually. Even though edge-based module provides less contribution to the model performance, we can see that this module extracts tiny roads more frequently than other modules. Furthermore, the extracted roads have smoother boundaries. Although clustering based module provides smaller recall value, the extracted road parts by this module are not fragmented as the outputs of the other modules. The structure analysis based module, on the other hand, provides the highest precision value but tends to extract roads as disconnected parts since it analyzes the segments independently. The input image and the final output is presented in Figure 102 and the contribution of individual modules is depicted in Figure 103.

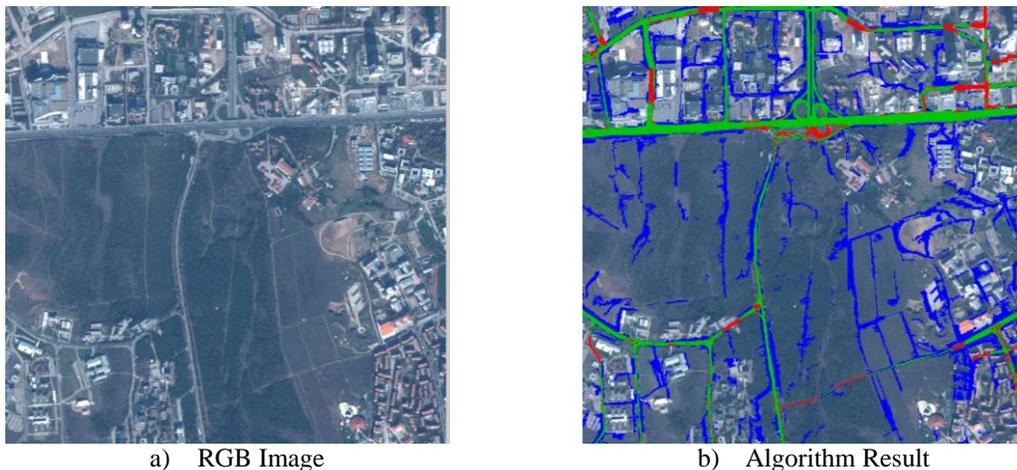
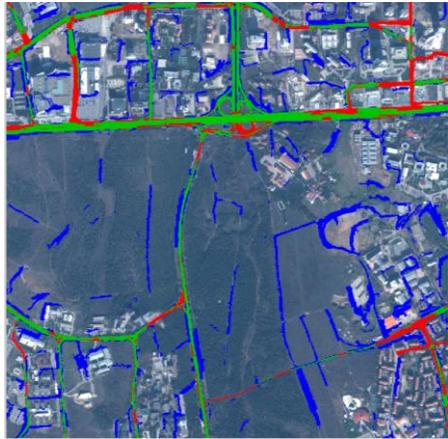
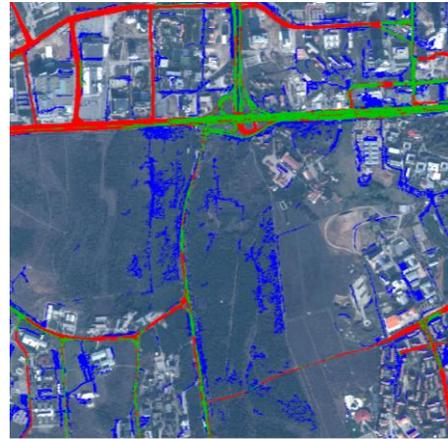


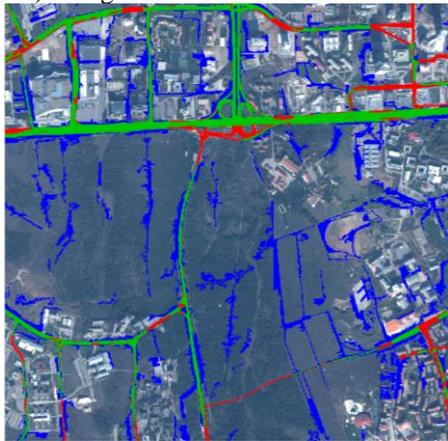
Figure 102 RGB image and the algorithm output



a) Edge based road extraction



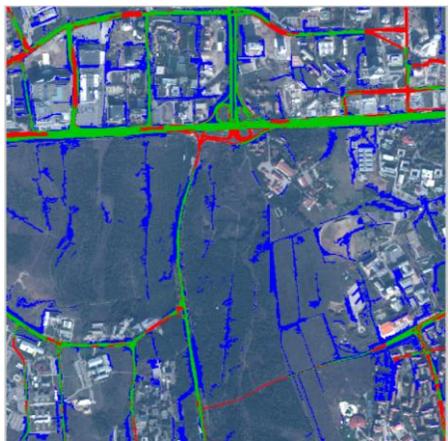
b) Clustering based road extraction



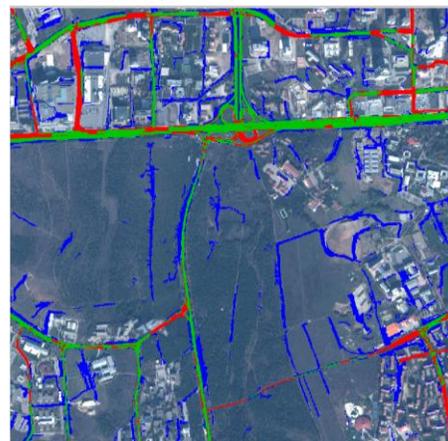
c) Structure analysis based road extraction



d) Structure analysis and edge based road extraction



e) Structure analysis and clustering based road extraction



f) Clustering and edge based road extraction

Figure 103 Comparison of modules

5.7.1 Comparison with Other Methods

Reviewing existing methods showed that most of the studies are semi-automatic. On the other hand, some automatic road extraction algorithms have focused only on specific regions and textures. In this study, we aimed to develop a road network extraction method that is not dependent on any condition like sensor specification, number of bands or specific regions.

In order to compare performances with those of other approaches and determine the contribution of our methods for road extraction, the same data should be employed. However, license agreements have made this impossible and other authors have not agreed to share their codes with us. Hence, the discussion of the proposed method with existing methods will be made based on data characteristics, performance metrics and applied methods.

In the study of Zhao et al. (2002), a semi-automatic approach is proposed. The approach needs user feedback continuously and traces only 4005 pixels in total. This limited number of pixels is not appropriate to give any opinion about the accuracy of the study. Another semi-automatic road extraction algorithm proposed in Vandana et al. (2002) is also based on road tracking approach. Quantitative results are not provided but visual results are promising. However, the test image does not include any occlusions or shadows. Also, the method was not tested on multi-spectral images. In Lin et al. (2011), proposed their approach and tested the approach on 512X512 IKONOS images. In our experience, it is not plausible to generalize results and performance with such small size of images. In Long and Zhao (2005), an integrated approach including segmentation, morphology and thresholding procedures is proposed. The study applies thresholding on simplified images (mean of intensity values of segments) obtained by the mean-shift procedure. This may produce good results; however fixed thresholding on color information may result in removal of the road parts. Fixed color thresholding for pre-processing is also used in Zhang et al. (1999). In our approach, instead of color thresholding of segments', shape information is utilized to extract the road map.

In general, in road tracking and template matching based approaches as in Da Silva et al. (2012), Park et al. (2001), Kim et al. (2004), Movaghati et al. (2010), and Lin et al. (2011) the performances strictly depend on the choice of the initial road seed. In Lin et al. (2011) lane markers are also employed but lane markers are available only for very high resolution images.

In Bacher and Mayer (2005), results are provided quantitatively and look promising but all images are selected from rural areas. They focused on extraction of main roads on such images. In our approach, main roads in rural areas can also be extracted with high performance. Sample result on rural area is shown in Figure 104.

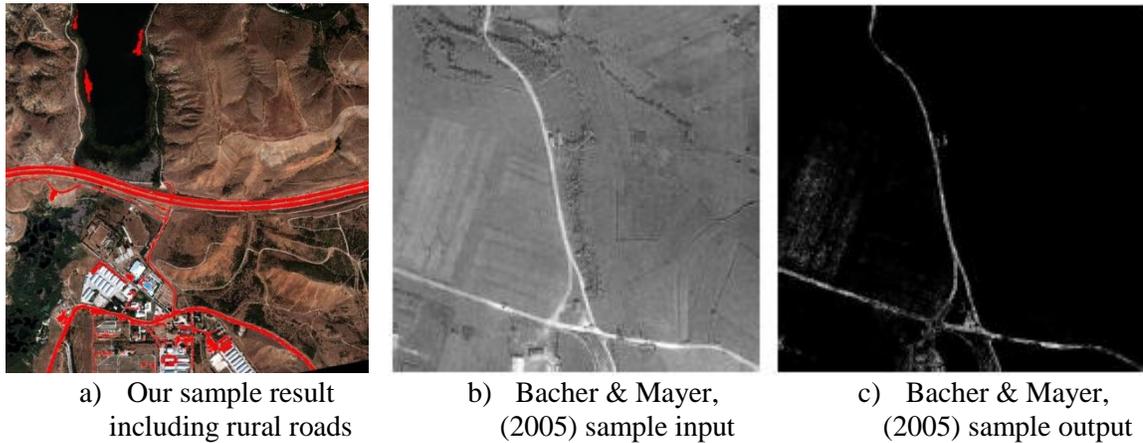


Figure 104 Sample rural area result and comparison with Bacher & Mayer (2005)

In the study conducted by Mohammadzadeh et al. (2008), a semi-automatic approach based on fuzzy cost function calculation according to initial pixel selected by user has been proposed. A small size rural area image has been used in the study. In addition, the approach includes consecutive post-processing steps which may not work well for other regions including other types of land cover.

In the study of Senthilnath et al. (2009), panchromatic QuickBird images have been tested with TPA and Normalized cut methods. In addition to its small size, the image just includes roads, buildings and little vegetation regions. The results are promising, but for the pre-processing method image is converted to binary image by using simple thresholding assuming that roads are darker and buildings are brighter objects. However, this assumption may not hold for images with many land cover objects. Sample result of the study and our algorithm result having similar texture is shown in Figure 105.



a) Sample result of our algorithm



b) Sample result of Senthilnath et. al. (2009)

Figure 105 Comparison with Senthilnath et.al (2009)

In Sırmaçek and Ünsalan (2010), edge map voting based segmentation is applied and each segment is analyzed to check whether it satisfies the elongatedness property of roads by using the ratio of region area and square of perimeter normalized by 4π . This property may not include junction points as mentioned in Chapter 4. This inference is also valid for the study of Xu et al. (2009) which uses ratio of area and perimeter as the shape index.

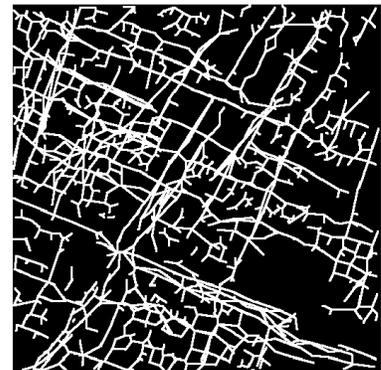
In the study conducted by Lee et al. (2000), road segments are selected by analysis of mean gray value, size (number of pixels in segment) and shape information (major/minor axis). Visual results look promising and supported by quantitative values. However, image characteristics have similar texture, and cover rural areas and main roads. Sample result of Lee et al. (2000) and our algorithm result are shown in Figure 106.



a) Our sample result



b) Lee et.al (2000)
sample input



c) Lee et.al (2000)
sample output

Figure 106 Comparison with Lee et.al (2000)

Color thresholding, shape index and shape size are commonly used features in the literature. However, they should be adaptive. In our experience, fixed color thresholding does not provide road map since it may change according to region and satellite sensors. Also, shape index should handle junction points and grid structure. Hence elongatedness measure is needed for enhancement as proposed in this thesis.

In the study of Rajeswari et al.(2011), the image is clustered into 15 clusters. Five of them are considered to belong to the road class. After smoothing the image to eliminate occlusion effects by filtering, segmentation algorithms are applied. As mentioned above, selecting fixed cluster or clusters as the road cluster may not provide road region. In our approach, each component in each cluster is analyzed to detect roads. Although it is reported that the approach provides over 80 % performance, they focused only on highways in sample images. In addition, pre-processing techniques consist of successive steps which affect overall performance. In Figure 107, a sample result of their study and our algorithm are shown.

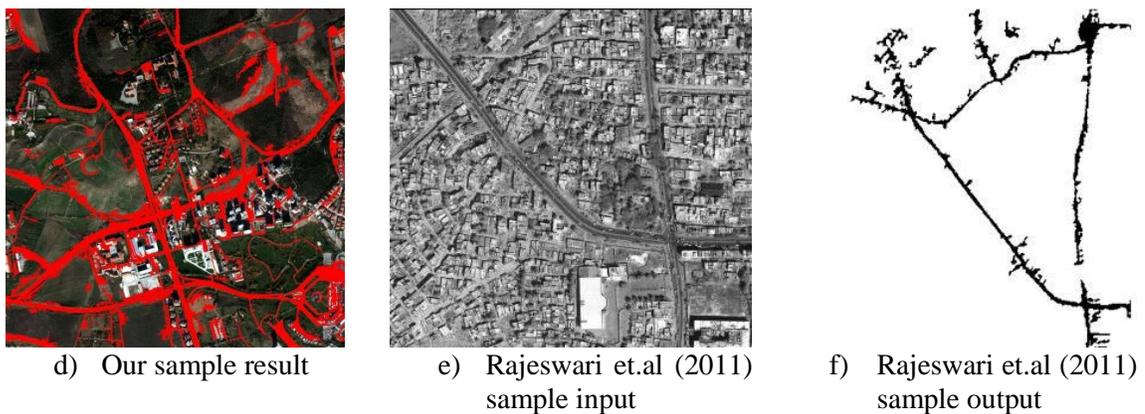


Figure 107 Comparison with Rajeswari et.al (2011).

The study conducted by Mohammadzadeh et al.(2008), proposed fuzzy segmentation based approach. The approach needs initial road segments and based on these initial segments' mean values, membership functions are constructed including five members. Then fuzzy segmentation is applied and the result is thresholded to obtain the road mask. It is obvious that, the performance of the algorithm depends on the initial segment selection.

Mena and Malpica (2005) proposed a method to update existing GIS database. Since they use available road network to enhance the map, it is not possible to compare with our results.

However, it is important to note that, our model would be enhanced by using existing data as well.

In Liu et al. (2003), a genetic algorithm based road extraction method has been proposed. They employed a 350X214 image in the approach consisting of two different objects i.e. road and bare soil. Although the result is very promising; the method should be tested with more images including different type of land covers. Our algorithm result and sample algorithm input and output of the study Liu et al. (2003) are shown in Figure 108.

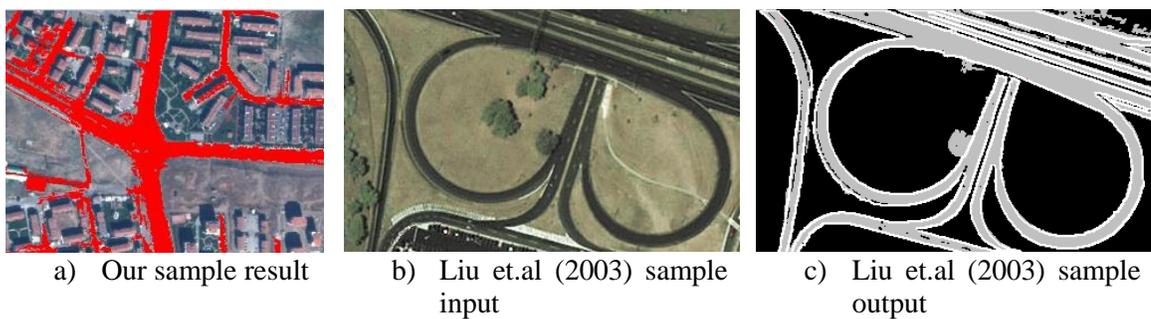


Figure 108 Comparison with Liu et al. (2003).

In the study of Zhang and Couloigner (2006), a clustering base method is proposed. Based on color assumptions, classification is conducted. Results are closer to our approach. However, this method's performance strictly depends on the k-means algorithm results and color assumptions which may vary for different regions which may have different intensity.

In Laptev et al. (2000), it is reported that the proposed algorithm produces 95 % correctness and 72 % completeness. Although the proposed method can be used for improving performance, it can be seen that the images do not include any buildings objects other than road and soil.

In the study of Grote et al., (2009), graph based approach is proposed. After segmentation on 10 cm resolution images, the segments are merged according to their shapes and spectral features. This may cause miss-merging since spectral characteristics of elongated buildings may be similar to roads as discussed in Chapter 6.

In Yuan et al. (2009), a segmentation based method is proposed. It is expected to extract roads by using LEGION segmentation. It is assumed that since roads are thin and elongated, they belong to background segments. Visual results show that this method can be used in rural areas.

Since the method does not employ any post-processing method for roads, it may not differentiate roads with other elongated structures like rivers, railroads, and etc. In Figure 109, sample result of the study and our sample result including rural region roads are represented.

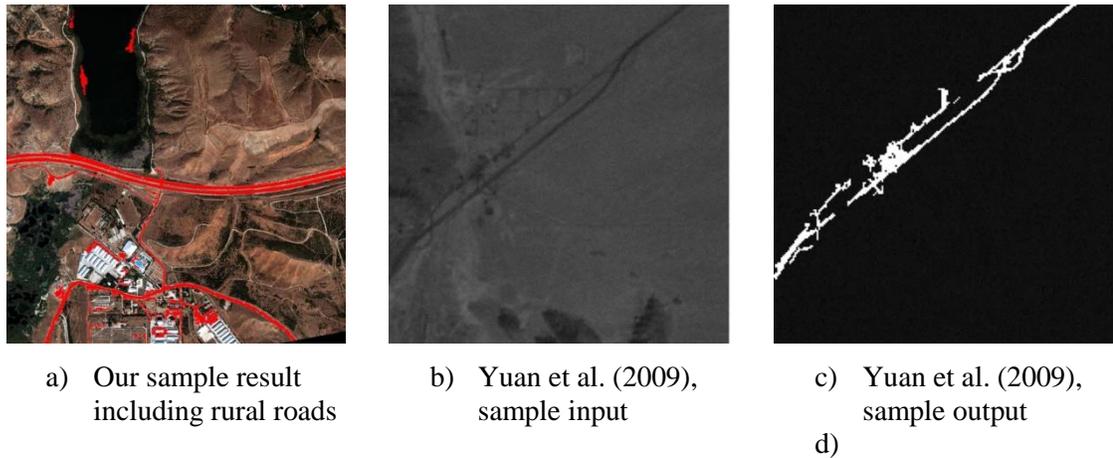


Figure 109 Comparison with Yuan et al. (2009).

In Peteri et al. (2003), a snake based approach has been proposed. Visual result shows that the algorithm is promising for rural areas. However, the algorithm needs the initial road map from GIS database or other resource. Also the edge based road extraction module proposed in this study, provides comparable performance when it is conducted on similar satellite images.

As mentioned before, since we neither have the same data nor do we have the final codes of other authors, we could not compare our results with theirs. On the other hand, we discussed our visual results and quantitative performance values with their reported results. We may say that, our approach is comparable and has demonstrated robust performance over a larger class of roads.

CHAPTER 6

6 CONCLUSION AND FUTUREWORK

In this thesis, transportation infrastructure extraction from multispectral satellite images has been studied. In this context, an algorithm for automatic road extraction from multispectral satellite images is developed. A water canal extraction algorithm and a railroad classification algorithm is also described in this thesis. Moreover, road cover material analysis is investigated.

Road Network Extraction

The road extraction model developed in this thesis does not require user feedback or interaction. In this manner this approach is a fully automatic road network extraction system.

One distinctive feature of the proposed model is its modular structure. A score may be constructed by fusing all the outputs of different modules. This structure enables the model to handle different spatial and material characteristics of roads in urban and rural areas. Also, since this modular approach is organized as a set of parallel modules, it limits the interaction between different modules.

Furthermore, the modular structure increases the adaptability of the model, since it allows us to integrate new modules. Another factor that enhances the adaptability is its applicability to lower and higher spatial resolution images by parameter adjustment. Since each band is processed separately, the model also supports different spectral resolution images including hyper-spectral

images that provide a higher spectral resolution than multi-spectral images. One other property that contributes to the satellite sensor independency of the model is its being an unsupervised-approach.

The algorithm is tested on four different WV-2 images to evaluate the model for different regions and times. Three of those are from the city of Ankara including different land covers. One is obtained on 16 November 2010 and the other two are taken on 26 September 2011. The fourth one is from a different country, Russia, and it is taken on 7 July 2010.

Results show that, this algorithm provides a promising performance on WV-2 satellite images with an average of 45% precision and 70% recall. The algorithm is also employed with four band, RGB and grayscale images. Overall, it was observed that additional bands returned moderately higher performance rates. Each additional band resulted in extraction of more road parts, in particular, for the edge based road extraction module. Moreover, the increasing of the number of bands makes the contribution of clustering based module more significant in terms of recall value.

These performance values are dependent on the road assumptions. If a road path is not elongated enough or it is occluded by some objects, it may not be extracted. These cases specially occur for the urban areas since cars and buildings may cast shadows on roads. Moreover, the proposed method may have some problems if a road segment is recovered by a new or different material than the original, since it violates the uniformity of color or edge gradient. Another problematic case occurs when some trees and branches cover roads, since such parts may be eliminated by the vegetation analysis module. Furthermore, the proposed algorithm unintentionally extracts runways and taxi routes of airports since they are similar to roads in terms of their spatial and spectral characteristics.

Water Canal Extraction

Water canal extraction algorithm is also important for the transportation analysis. In the literature, water regions are extracted only by using band ratio or band thresholding. In our study, NDVI, NDWI and NIR bands are employed together and their thresholds are determined automatically. Structural analysis is also applied on water extraction module to extract only

water canals rather than other type of wetlands like lakes and seas. Results show that the water canal extraction part of the study provides 80% precision and 85% recall.

However, it is important to note that, it is hard to discriminate water and elongated shadow regions because of their spatial and spectral similarities. Moreover, in some cases new asphalted road parts may be labeled as water or vice versa for the same reasons.

Railroad Classification

Railroads are another important component of the transportation network, although there are very few studies on railroad extraction from remotely sensed data. In this study, we aimed to differentiate road and railroad structures. For this purpose, Fourier transform is used since the distance between railroad sleepers is constant and they provide extra spectral peaks in the frequency domain. This approach has been tested on six different 20 cm resolution Google Earth images including both railroad and road structures. The algorithm provides satisfactory performance when the railroad sleepers depict good contrast with the background and they are not deteriorated.

Road Type Classification

Road type and material classification is also studied in this thesis. Roads' cover materials are classified as asphalt, unpaved and concrete by using GMM classification. It is important to note that this part of the study requires training data and performance of this part is highly dependent on the training set. This may cause sensor and region dependency, in the sense that it may limit the applicability of the algorithm to images other than WV-2 images.

6.1 Limitations

Some of the limitations of the proposed method are due to the limited scope of roads definitions. Even though the road extraction model includes some contextual information like “vegetative fields cannot cover roads” and “water regions cannot be on roads”, other contextual information that may improve the performance has not been included.

Another limitation is due to the number of bands used in the segmentation module. In order to improve the adaptability of this model to other types of multispectral satellite images, an N-Band segmentation application is needed, but we used at most 4-bands.

Although the test data set for road extraction model is significantly larger than those used in the literature, it would be better to test the method on data obtained from different geographic regions at different seasons. These tests might help develop more effective and generalizable algorithm, since the spatial and spectral characteristics may change significantly. The test data of water extraction method is also limited due to land cover of test images.

More training data set gathered from different locations is required for more accurate road type classification. Images that cover the wavelengths beyond 1300 nm are not used in this study but they may be useful for classifications.

Lastly, more efficient and reliable evaluation of the model needs more “ground truth” information. However this is a very costly and time consuming process.

6.2 Future Works

The proposed algorithm works both for rural and urban regions but it performs better for rural regions. Accordingly, one possible improvement of the model may be to develop a new module that focuses particularly on extraction of the narrow roads and city roads.

Construction of a better ontology that includes more contextual information can also contribute the proposed model on the problem of developing a line merging method to fill the gaps that are caused by shadows or other objects. Such ontology may additionally enable the removal of non-road objects more efficiently.

New satellite systems have capability to provide stereo images. One possible extension of the study may be the adaptation of the model to use stereo images for road extraction. Additionally, when the high resolution depth information is available, this feature may also be used to improve road extraction performance. These improvements can easily be included to the model due to its modular structure that provides flexibility.

Aside from road network extraction, water canals and railroads are also studied in this thesis. These applications are the first step for the development of new methods for extraction of some important structures, such as bridges and junctions. This thesis lays the foundation to detect such features since it enables the extraction of intersections of different transportation infrastructures.

Based on the proposed model, the possibility of obtaining seed points automatically to be used in the road tracking methods can be investigated. Such an improvement will be also beneficial in the water canal and railroad extraction.

The railroad algorithm developed in this study is specifically for the higher resolution images. However, development of other methods that are applicable to lower resolution images may be valuable, because of the scarcity of the high resolution images.

One important contribution of this study to the current literature is the inclusion of a technique for the discrimination of different road types, since – to our knowledge -- none of the previously proposed methods aim to solve this problem. As a classification problem, the full-fledged solution of this discrimination is beyond the scope of this thesis.

REFERENCES

- Aksoy, S., Akçay, H. G., & Wassenaar, T. (2010). Automatic mapping of linear woody vegetation features in agricultural landscapes using very high resolution imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 48(1), 511-522.
- Aksoy, S., Yalniz, I. Z., & Tasdemir, K. (2012). Automatic Detection and Segmentation of Orchards Using Very High Resolution Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 50(8), 3117-3131.
- Alesheikh, A. A., Ghorbanali, A., & Nouri, N. (2007). Coastline change detection using remote sensing. *International Journal of Environmental Science and Technology*, 4(1), 61-66.
- Amini, J., Lucas, C., Saradjian, M. R., Azizi, A., Sadeghian, S. (2002). Fuzzy Logic System For Road Identification Using IKONOS Images. *Photogrammetric Record*, April 2002.
- Bacher, U., Mayer, H. (2005). Automatic Road Extraction from Multispectral High Resolution Satellite Images. *ISPRS Object Extraction for 3D City Models, Road Databases and Traffic Monitoring - Concepts, Algorithms and Evaluation, CMRT05, Vienna, Austria Volume XXXVI(3)*.
- Beumier, C., & Lacroix, V. (2006). Road extraction for eurosdr contest. *Proceedings of the SPIE, Image and Signal Processing for Remote Sensing XII*, vol. 6365, Stockholm, Sweden.
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679-698.
- Chaudhuri, D., & Samal, A. (2008). An automatic bridge detection technique for multispectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 46(9), 2720-2727.

Cheng, Y. (1995). Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8), 790-799.

Christoudias, C. M., Georgescu, B., & Meer, P. (2002). Synergism in low level vision. *Proceedings of the International Conference on Pattern Recognition, Canada, 4*, 150-155.

Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. , *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5), 603-619.

Çinar, U. (2012). Road extraction from high resolution satellite images using adaptive boosting with multi-resolution analysis. Unpublished master's thesis, Middle East Technical University, Ankara, Turkey

Da Silva, C. R., & Centeno, J. A. S. (2012). Semiautomatic extraction of main road centrelines in aerial images acquired over rural areas. *International Journal of Remote Sensing*, 33(2), 502-516.

Da Silva, D. V. S., Fernando, W. A. C., Kodikaraarachchi, H., Worrall, S. T., & Kondo, A. M. (2010). Adaptive sharpening of depth maps for 3D-TV. *Electronics Letters*, 46(23), 1546-1548.

ENVI Online Help (August 12, 2005). Retrieved from http://geol.hu/data/online_help/Vegetation_Indices.html, December 25, 2012

Frazier, P. S., & Page, K. J. (2000). Water body detection and delineation with Landsat TM data. *Photogrammetric Engineering & Remote Sensing*, 66(12), 1461-1467.

Fukunaga, K., & Hostetler, L. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory*, 21(1), 32-40.

Gardner, M. E., Roberts, D. A., Funk, C. and Noronha, V. (2001). Road Extraction from AVIRIS Using Spectral Mixture and Q-Tree Filter Techniques, Proceedings of the Tenth JPL Airborne Earth Science Workshop; December 2001, pp.145-150.

Gedik, E., Çınar, U., Karaman, E., Yardımcı, Y. and Halıcı, U. and Pakin, K. (2012). Yüksek çözünürlüklü elektro optik imgelerde köprü tespiti için kural tabanlı bir yaklaşım. *Proc. of 20th Signal Processing and Communications Applications Conference (SIU)*, vol., no., pp.1-4, 18-20 April 2012 doi: 10.1109/SIU.2012.6204808.

Géraud, T., & Mouret, J.B. (2004). Fast road network extraction in satellite images using mathematical morphology and Markov random fields. *EURASIP Journal of Applied Signal Processing*, 16, 2503-2514.

Grote, A., Heipke, C., Rottensteiner, F., & Meyer, H. (2009). Road extraction in suburban areas by region-based road subgraph extraction and evaluation. In *Urban Remote Sensing Event, 2009 Joint* (pp. 1-6). IEEE.

Gu, D. Y., Zhu, C. F., Shen, H., Hu, J. Z., & Chang, H. X. (2011). Automatic Bridge Extraction for Optical Images. *Proceedings of Sixth International Conference on Image and Graphics (ICIG)*, 446-451.

Heijmans, H. J. A. M. (1991). Theoretical aspects of gray-level morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6), 568-582.

Hellwich, O., Laptev, I., & Mayer, H. (2002). Extraction of linear objects from interferometric SAR data. *International Journal of Remote Sensing*, 23(3), 461–475.

Hsu, P. H., & Yang, H. H. (2007, July). Hyperspectral image classification using wavelet networks. In *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International* (pp. 1767-1770). IEEE.

Hu, J., Razdan, A., Femiani, J., Wonka, P., & Cui, M. (2007). Fourier Shape Descriptors of Pixel Footprints for Road Extraction from Satellite Images. ICIP.

Johnston, R. & Barson, M. (1993). Remote sensing of Australian wetlands: An evaluation of Landsat TM data for inventory and classification. *Australian Journal of Marine and Freshwater Research*, 44(2), 235-252.

Karaman, E., Çınar, U., Gedik, E., Yardımcı, Y. and Halıcı, U. (2012). A new algorithm for automatic road network extraction in multispectral satellite images. *Proc. of 4th Geographic Object-Based Image Analysis Conference (GEOBIA)*, pp. 298, May 7-9, 2012
URL: <http://mtc-m18.sid.inpe.br/col/sid.inpe.br/mtc-m18/2012/05.15.17.43/doc/122.pdf>.

Karaman, E., Çınar, U., Gedik, E., Yardımcı, Y. and Halıcı, U. (2012). Fourier based feature descriptors for railroad extraction from aerial images. *2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, , vol., no., pp.6013-6015, 22-27 July 2012 doi: 10.1109/IGARSS.2012.6352237.

Karaman, E., Çınar, U., Gedik, E., Yardımcı, Y. and Halıcı, U. (2012). Multi-spektral uydu görüntülerinden otomatik yol haritası çıkarımı. *Proc. of 20th Signal Processing and Communications Applications Conference (SIU)*, vol., no., pp.1-4, 18-20 April 2012 doi: 10.1109/SIU.2012.6204704.

Kim, T., Park, S. R., Kim, T., Jeong, S., & Kim, K.O. (2004). Semi automatic tracking of road centerlines from high resolution remote sensing data. *Journal of The American Society for Photogrammetry and Remote Sensing*, 70, No. 12.

Laptev, I., Mayer, H., Lindeberg, T., Eckstein, W., Steger, C., Baumgartner, A. (2000). Automatic extraction of roads from aerial images based on scale space and snakes. *Machine Vision and Applications* (2000) 12: 23–31.

Lee, H. Y., Park, W., & Lee, H. K. (2000). Automatic road extraction from 1M-resolution satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 55.

Li, J., & Narayanan, R. M. (2003). A shape-based approach to change detection of lakes using time series remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 41(11), 2466-2477.

Li, X., Qiao, Y., Yi, W., Guo, Z. (2003). The Research of Road Extraction for High Resolution Satellite Image. *IEEE International*, Vol. 6, 3949-3951

- Lin, X., Zhang, J., Liu, Z., Shen, J., & Duan, M. (2011). Semi-automatic extraction of road networks by least squares interlaced template matching in urban areas. *International Journal of Remote Sensing*, 32(17), 4943-4959.
- Liu, H., Li, J., & Chapman, M. A. (2003). Automated road extraction from satellite imagery using hybrid genetic algorithms and cluster analysis. *Journal of Environmental Informatics*, 1(2), 40-47.
- Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2), 129-137.
- Long, H., & Zhao, Z. (2005). Urban road extraction from high-resolution optical satellite images. *International Journal of Remote Sensing*, 26(22), 4907-4921.
- Luo, J., Ming, D., Liu, W., Shen Z., Wang, M., & Sheng, H. (2007). Extraction of bridges over water from IKONOS panchromatic data. *International Journal of Remote Sensing*, 28(16), 3633-3648.
- Meer, P., & Georgescu, B. (2001). Edge detection with embedded confidence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(12), 1351-1365.
- Mena, J. B. (2003). State of the art on automatic road extraction for GIS update: a novel classification. *Pattern Recognition Letters*, 24(16), 3037-3058.
- Mena, J. & Malpica, J., (2005). An automatic method for road extraction in rural and semi-urban areas starting from high resolution satellite imagery. *Pattern Recognition Letters* 26, 1201–1220.
- Mohammadzadeh, A., Valadan Zoej, M. J., & Tavakoli, A. (2009). Automatic main road extraction from high resolution satellite imageries by means of particle swarm optimization applied to a fuzzy-based mean calculation approach. *Journal of the Indian Society of Remote Sensing*, 37(2), 173-184.

- Mohammadzadeh, A., Valadan Zoej, M. J., & Tavakoli, A. (2008). Automatic main road extraction from high resolution satellite imageries by means of self-learning Fuzzy-GA algorithm. *Journal of Applied Sciences*, 8(19), 3431-3438.
- Movaghati, S., Moghaddamjoo, A., & Tavakoli, A. (2010). Road extraction from satellite images using particle filtering and extended Kalman filtering. *IEEE Transactions on Geoscience and Remote Sensing*, 48(7), 2807-2817.
- Nath, R. K., & Deb, S. K. (2010). Water-body area extraction from high resolution satellite images - An introduction, review, and comparison. *International Journal of Image Processing*, 3(6), 353-372.
- Neubert, M., Hetch, R., Gedrange, C., Trommler, M., Herold, H., Kruger, T. and Brimmer, F. (2008). "Extraction of Railroad Objects from Very High Resolution Helicopter-Borne Lidar and Ortho-Image Data". *Proc. of GEOBIA 2008*, University of Calgary, Calgary Alberta, Canada, August 05-08.
- Özkaya, M. (2009). Road extraction from high-resolution satellite images. Unpublished master's thesis, Middle East Technical University, Ankara, Turkey
- Park, S. R., & Kim, T. (2001). Semi-automatic road extraction algorithm from IKONOS images using template matching. *Proceedings of the 22nd Asian Conference on Remote Sensing*, Vol. 5, p. 9.
- Parzen, E. (1962). On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3), 1065-1076.
- Peteri, R., & Ranchin, T. (2003). Multiresolution snakes for urban road extraction from IKONOS and QuickBird images. In *23rd EARSeL Annual Symposium Remote Sensing in Transition* (pp. 141-147). Ghent, Belgium.
- Rabatel, G., Delenne, C., & Deshayes, M. (2008). A non-supervised approach using Gabor filters for vine-plot detection in aerial images. *Computers and Electronics in Agriculture*, 62(2), 159-168.

Rajeswari, M.; Gurumurthy, K.S.; Omkar, S.N.; Senthilnath, J.; Reddy, L.P. (2011). Automatic road extraction using high resolution satellite images based on Level Set and Mean Shift methods. IEEE 3rd International Conference on Electronics Computer Technology (ICECT), 2011, vol.2, no., pp.424-428, 8-10 April 2011.

Ravanbakhsh, M., Heipke, C., Pakzad, K. (2007). Knowledge-based road junction extraction from high-resolution aerial images. Urban Remote Sensing Joint Event.

Shackelford, A. K., & Davis, C. H. (2003). Fully automated road network extraction from high-resolution satellite multispectral imagery. In *Geoscience and Remote Sensing Symposium, 2003. IGARSS'03. Proc. of 2003 IEEE International* (Vol. 1, pp. 461-463). IEEE.

Sırmaçek, B., and Ünsalan, C. (2010). Road Network Extraction using Edge Detection and Spatial Voting. 20th International Conference on Pattern Recognition (ICPR) (s. 3313-3316). August 23-26 2010, İstanbul

Tiwari, P. S., Pande, H., & Kumar P. A. (2009). Automatic urban road extraction using airborne laser scanning/altimetry and high resolution satellite data. *Journal of The Indian Society of Remote Sensing*, 37(2), 223-231.

Vandana S., V., Chandrakant, S., & Ramachandran, R. (2002). Semi-automatic road extraction algorithm for high resolution images using path following approach. ICVGIP.

Wang, L., Qin, Q., Du, S, Chen, D., & Tao, J. (2005). Road Extraction from Remote Sensing Image Based on Multi-resolution Analysis. ISPRS, 31. International Symposium on Remote Sensing of Environment, Saint Petersburg.

Wolf, A. (2010). Using WorldView-2 Vis-NIR MSI Imagery to Support Land Mapping and Feature Extraction Using Normalized Difference Index Ratios. *Digital Globe 8-Band Research Challenge*, 1-13

Xu, G., Zhang, D., & Liu, X. (2009). Road extraction in high resolution images from Google Earth. *7th International Conference on Information, Communications and Signal Processing*, 8-10.

Yuan, J., Wang, D. L., Wu, B., Yan, L. & Li, R. (2009). Automatic road extraction from satellite imagery using LEGION networks. *Proceedings of International Joint Conference on Neural Networks*, 3471–3476.

Zhang, C., Shunji, M., & Emmanuel, B. (1999). Road Network Detection by Mathematical Morphology. *ISPRS Workshop "3D Geospatial Data Production: Meeting Application Requirements"*, Paris, France, 185-200.

Zhang, Q. and Couloigner, I. (2006). Automated Road Network Extraction from High Resolution Multi-Spectral Imagery. In: *ASPRS 2006 Annual Conference*, Reno, Nevada.

Zhao, L., Yu, H., Zhang, L. (2009). Water body extraction in urban region from high resolution satellite imagery with Near-Infrared Spectral Analysis. *Proc. of SPIE Vol. 7383* 73833I-1, 1-6

Zhao, H., Kumagai J., Nakagawa, M., & Shibasaki, R. (2002). Semi-Automatic Road Extraction from High-Resolution Satellite Image. *ISPRS Photogrammetric Computer Vision*, Graz, Austria.

APPENDICIES

APPENDIX A: ALL ROAD EXTRACTION RESULTS

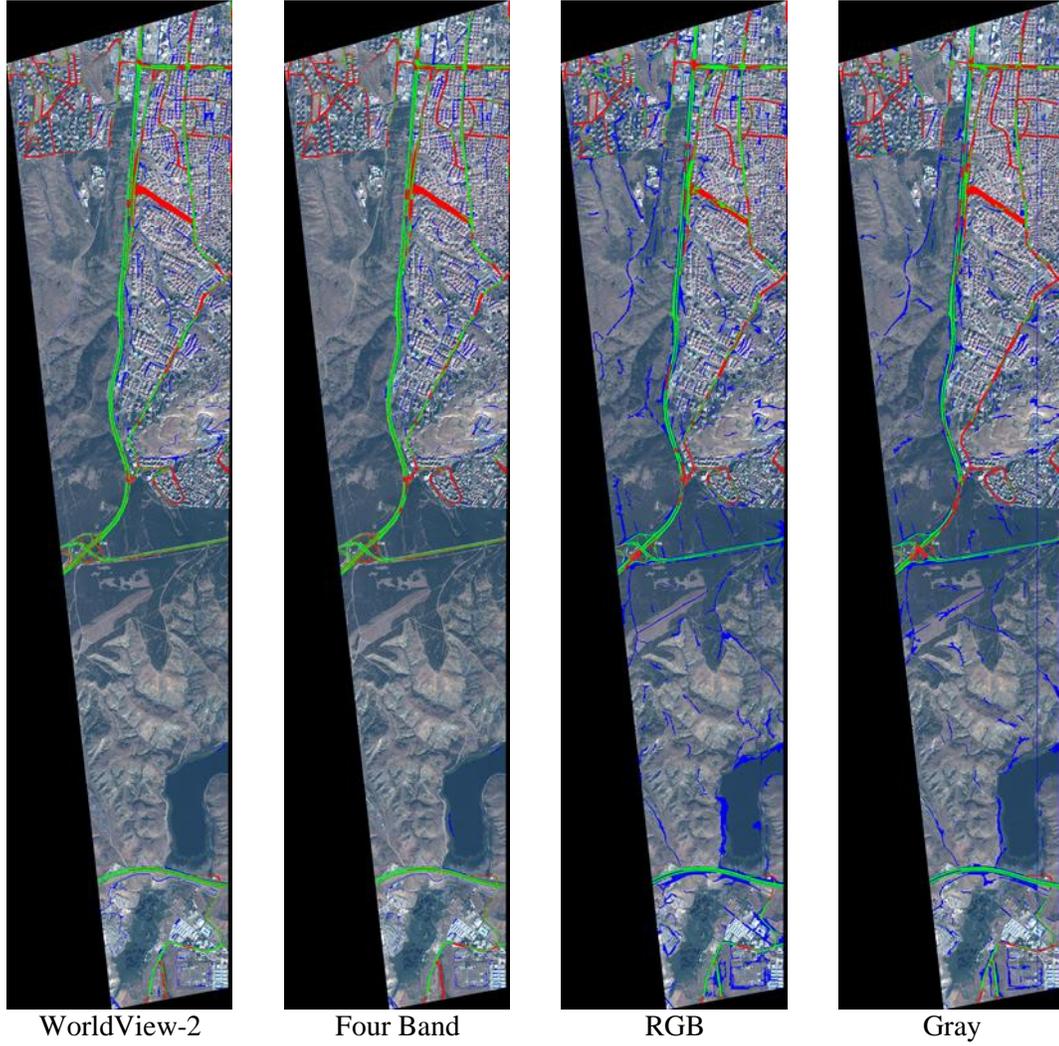
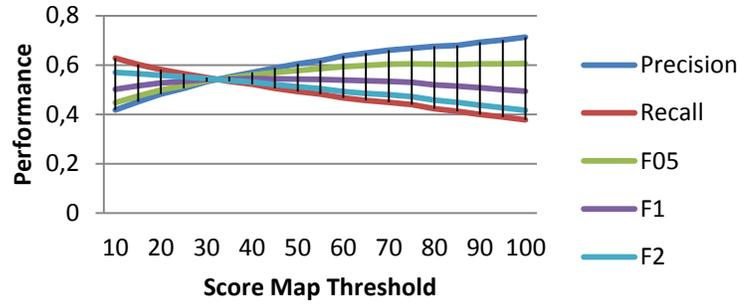
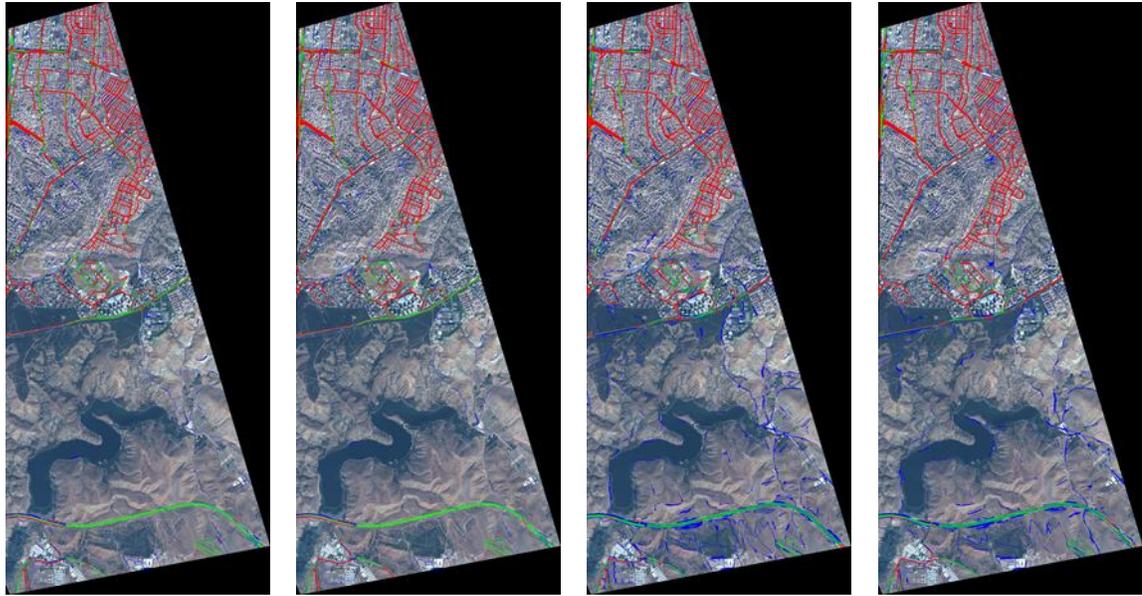


Figure 110 All results of Cankaya-1 Image

Performance Çankaya-1 (WV2)





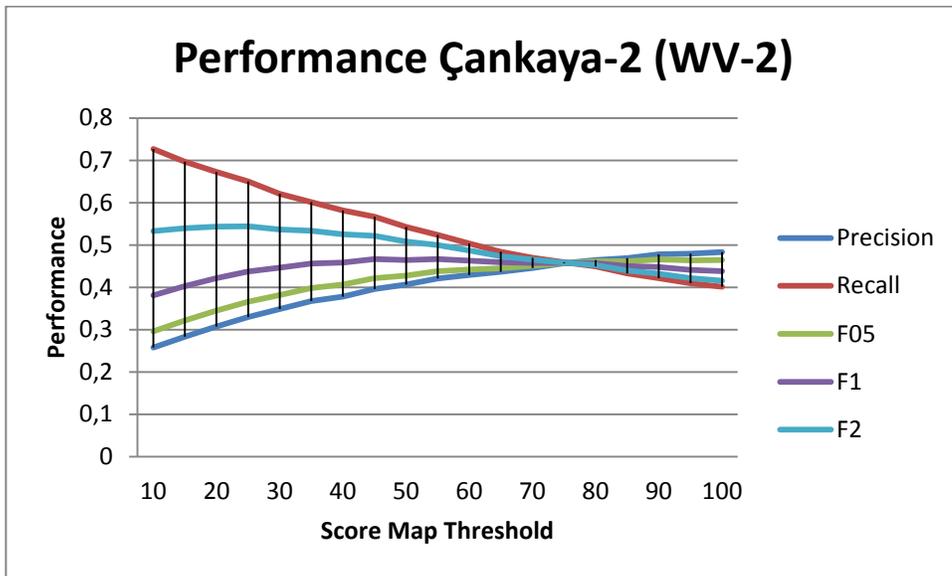
WorldView-2

Four Band

RGB

Gray

Figure 111 All results of Cankaya-2 Image



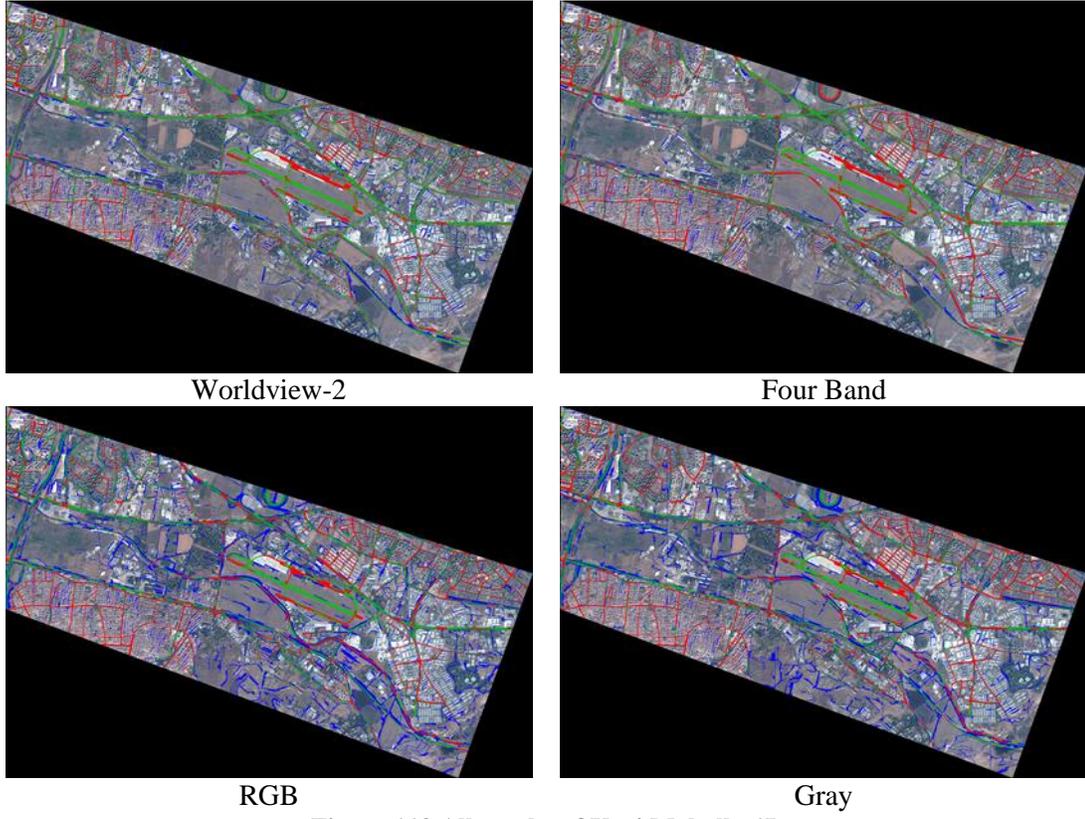
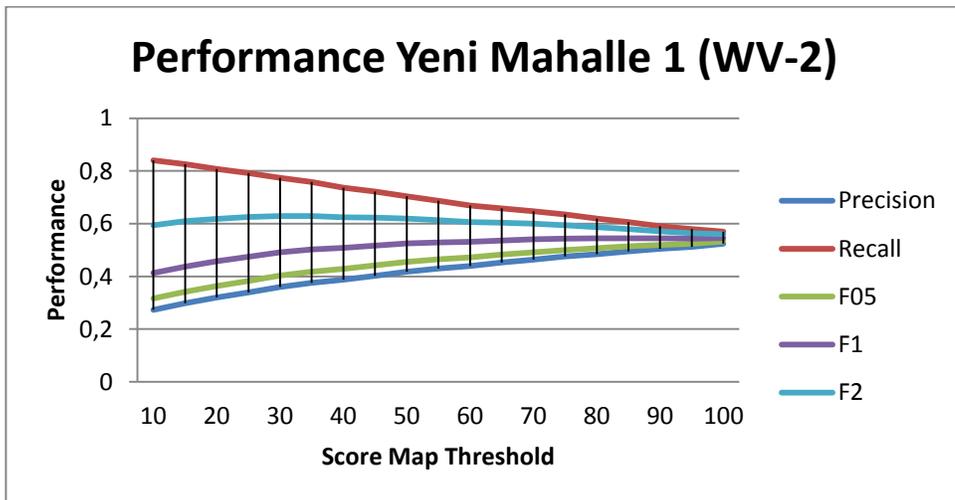


Figure 112 All results of Yeni Mahalle-1Image



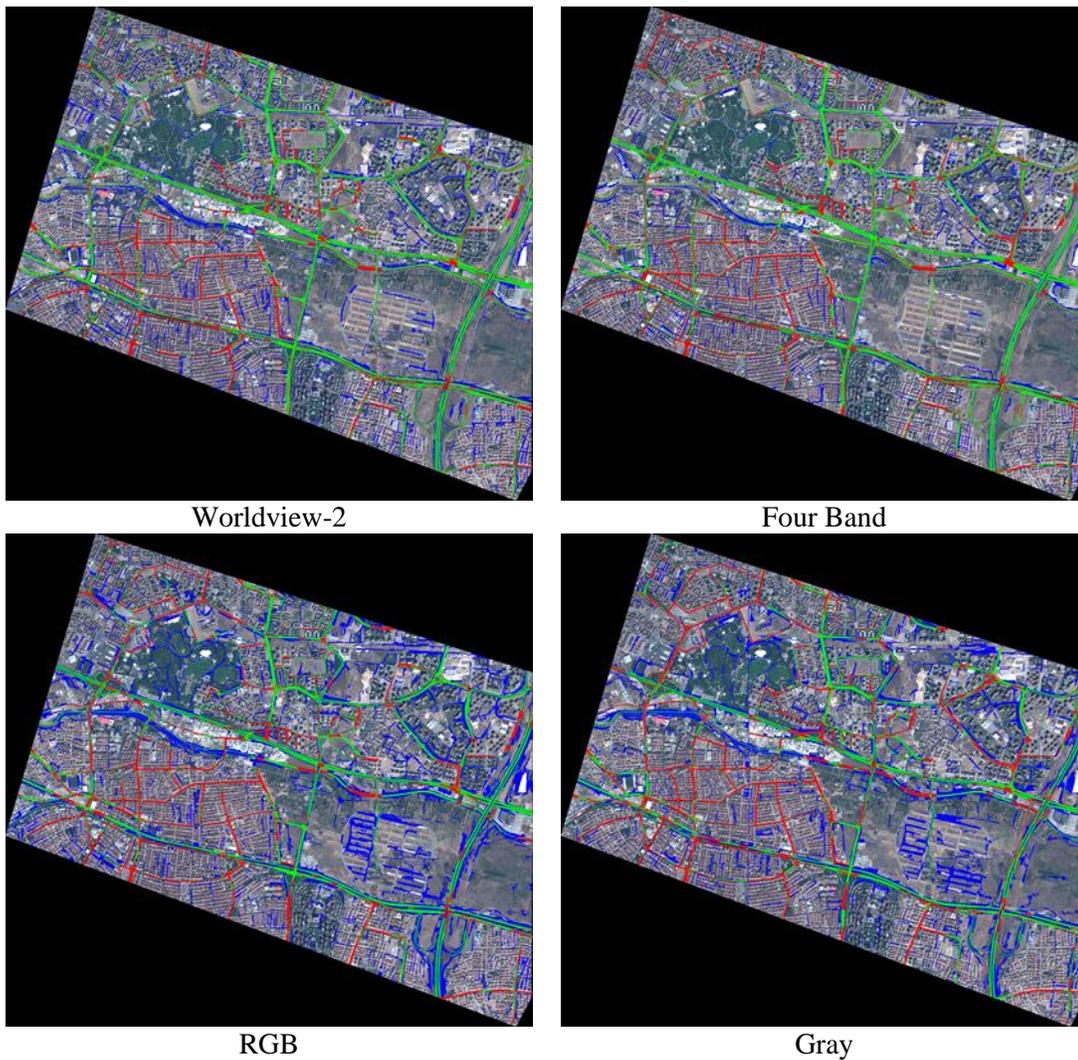
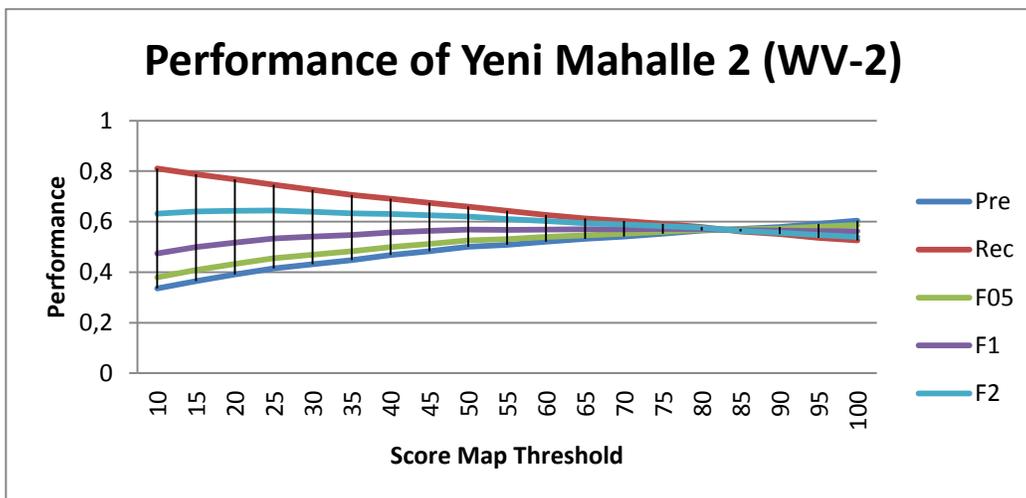
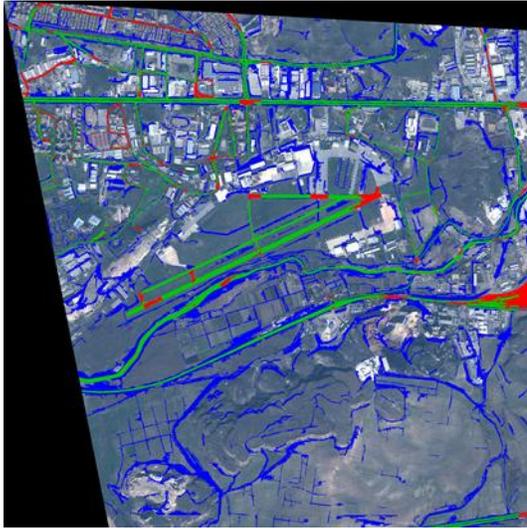
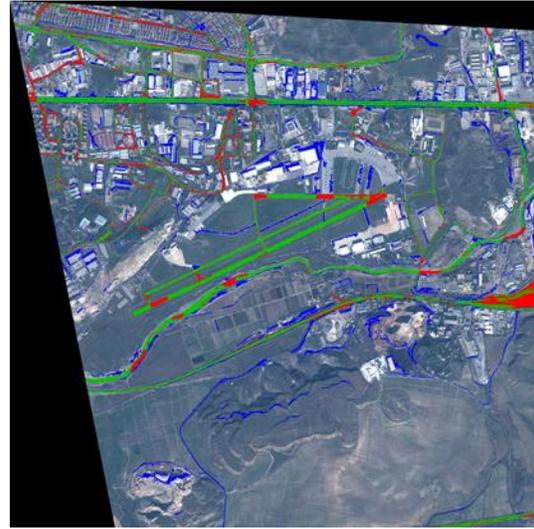


Figure 113 All results of Yeni Mahalle-2 Image

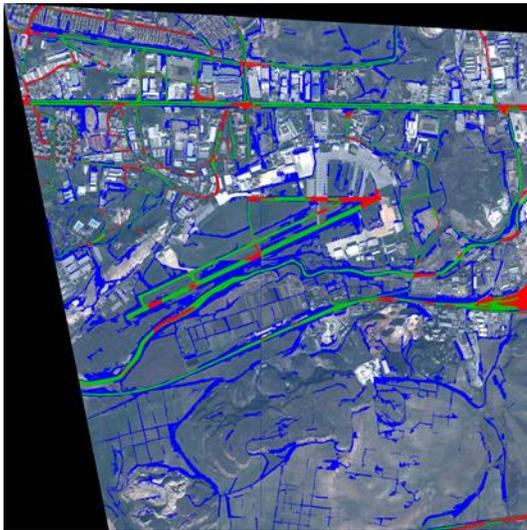




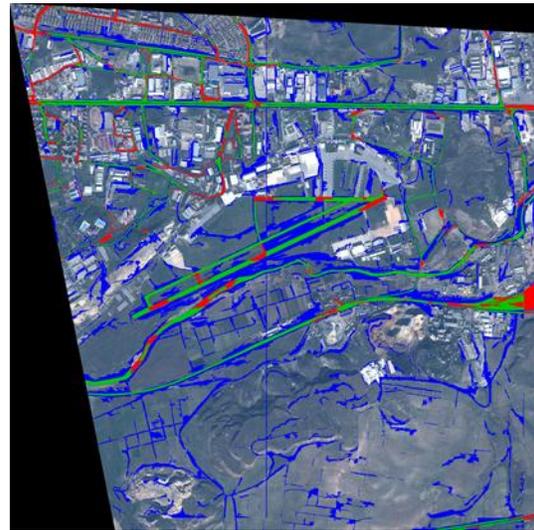
WorldView-2



Four Band

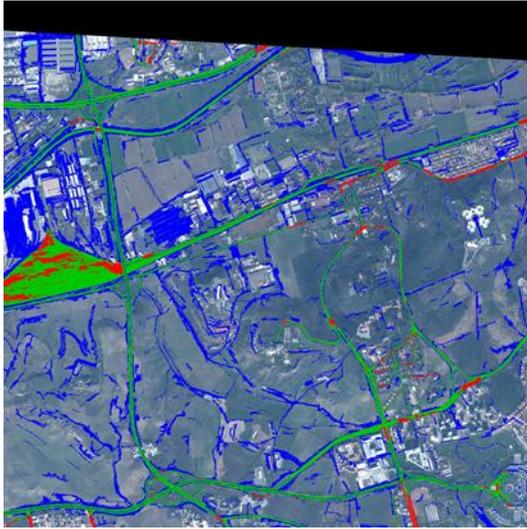


RGB



Gray

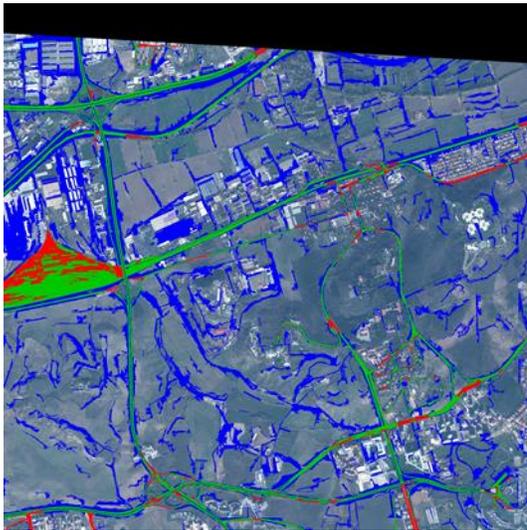
Figure 114 All results of METU (1)



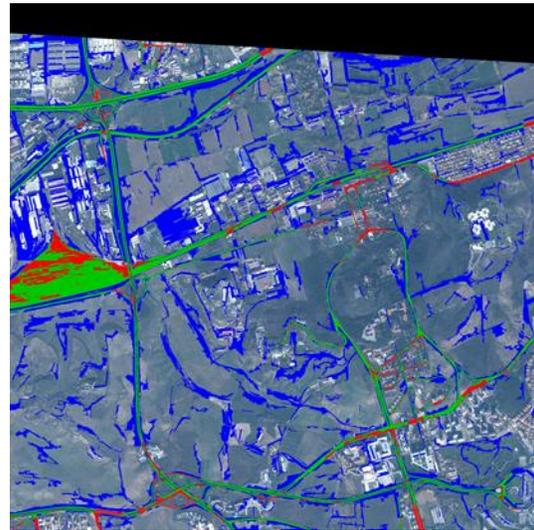
WorldView-2



Four Band

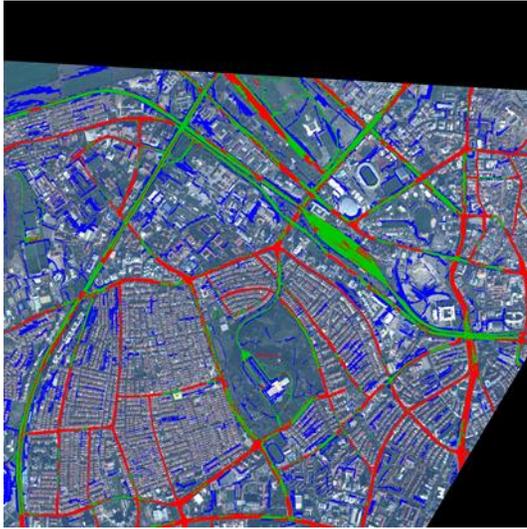


RGB

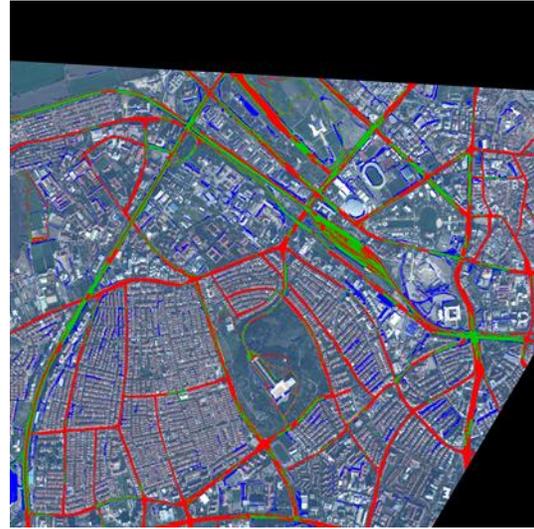


Gray

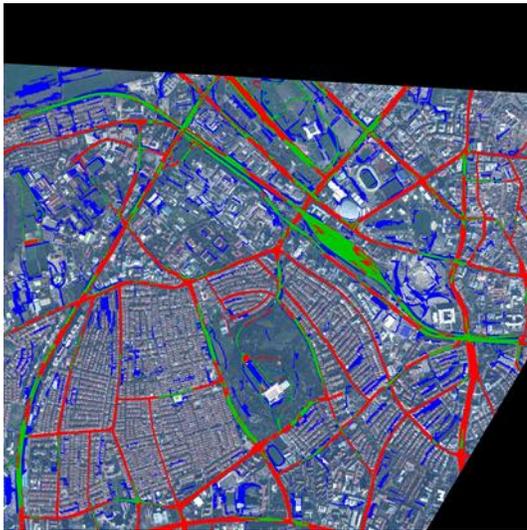
Figure 115 All results of METU (2)



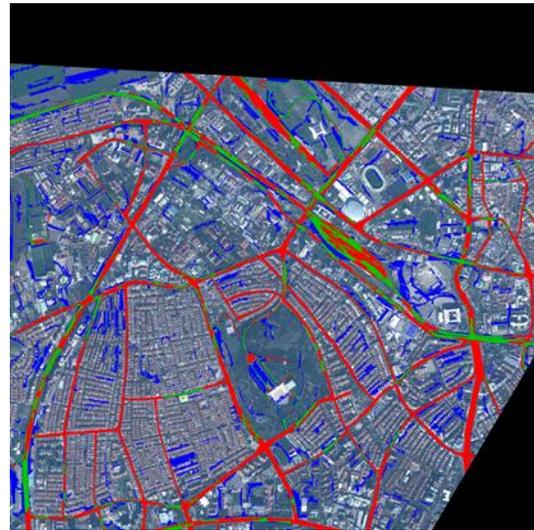
WorldView-2



Four Band

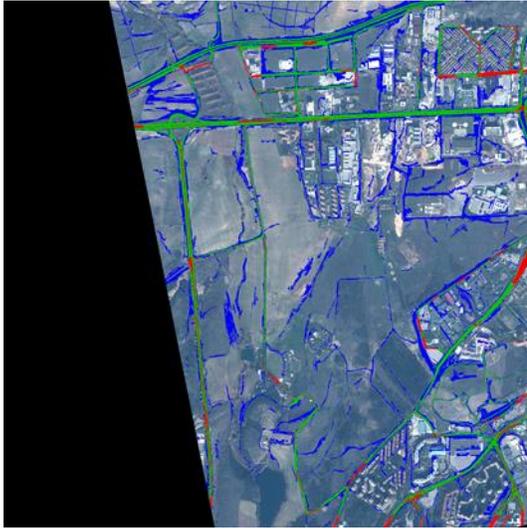


RGB

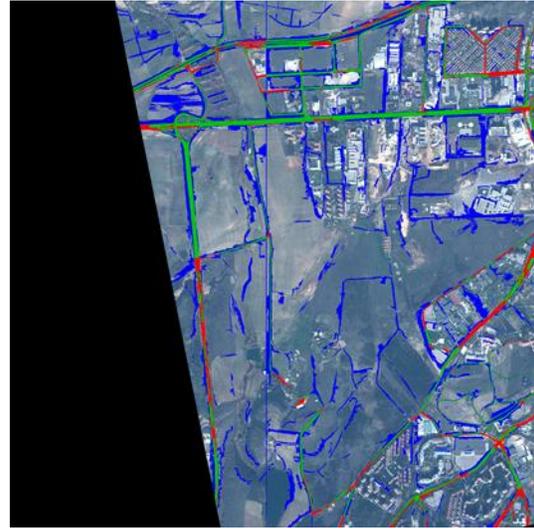


Gray

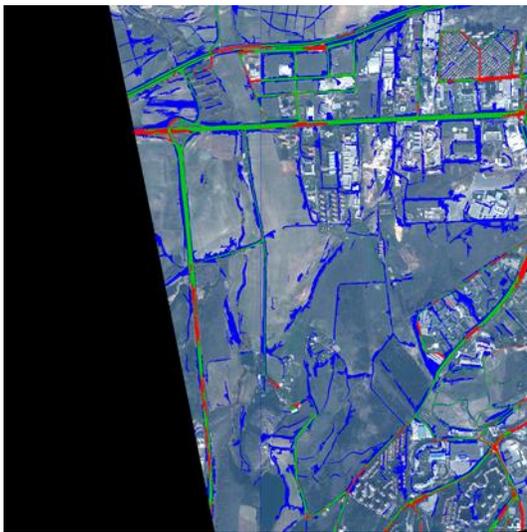
Figure 116 All results of METU (3)



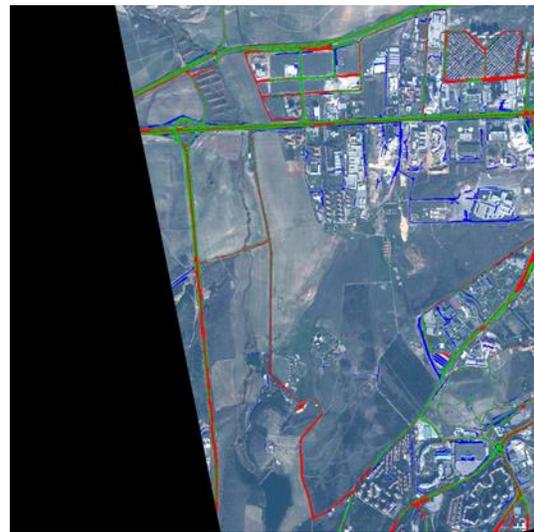
WorldView-2



Four Band

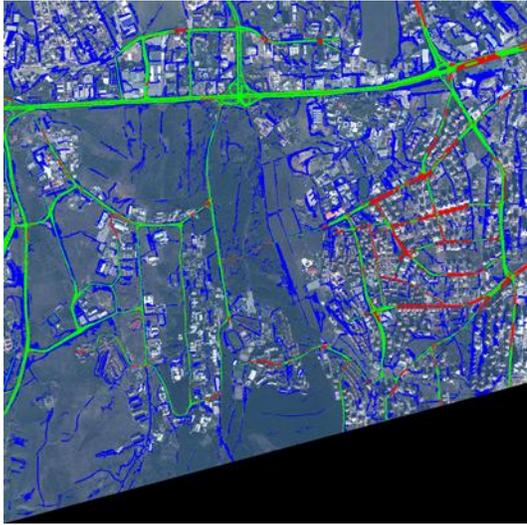


RGB

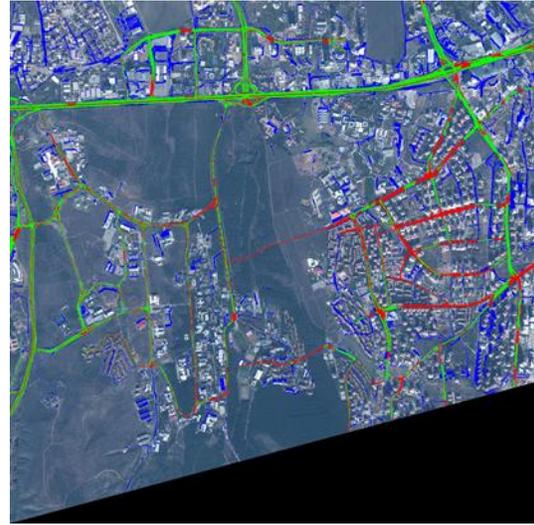


Gray

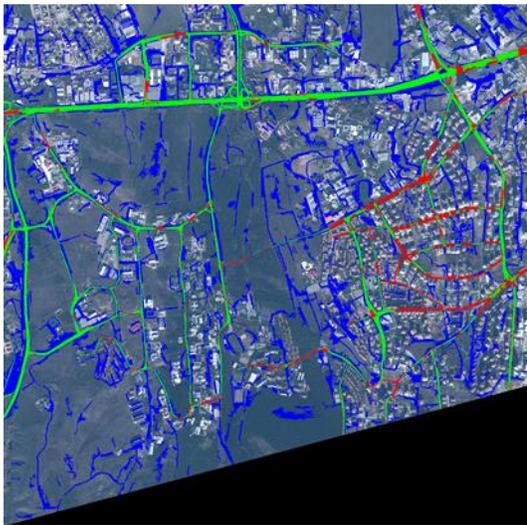
Figure 117 All results of METU (4)



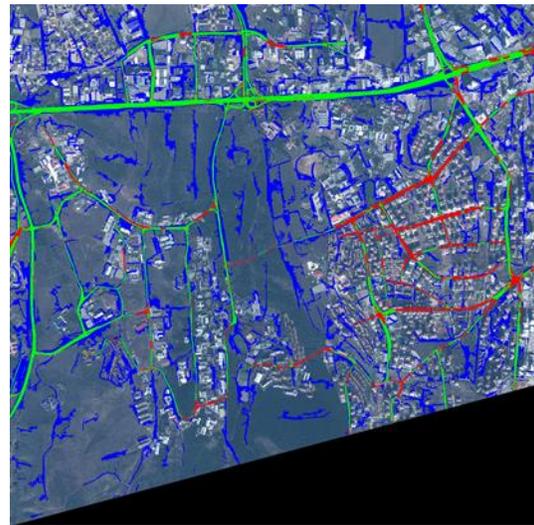
WorldView-2



Four Band



RGB



Gray

Figure 118 All results of METU (5)

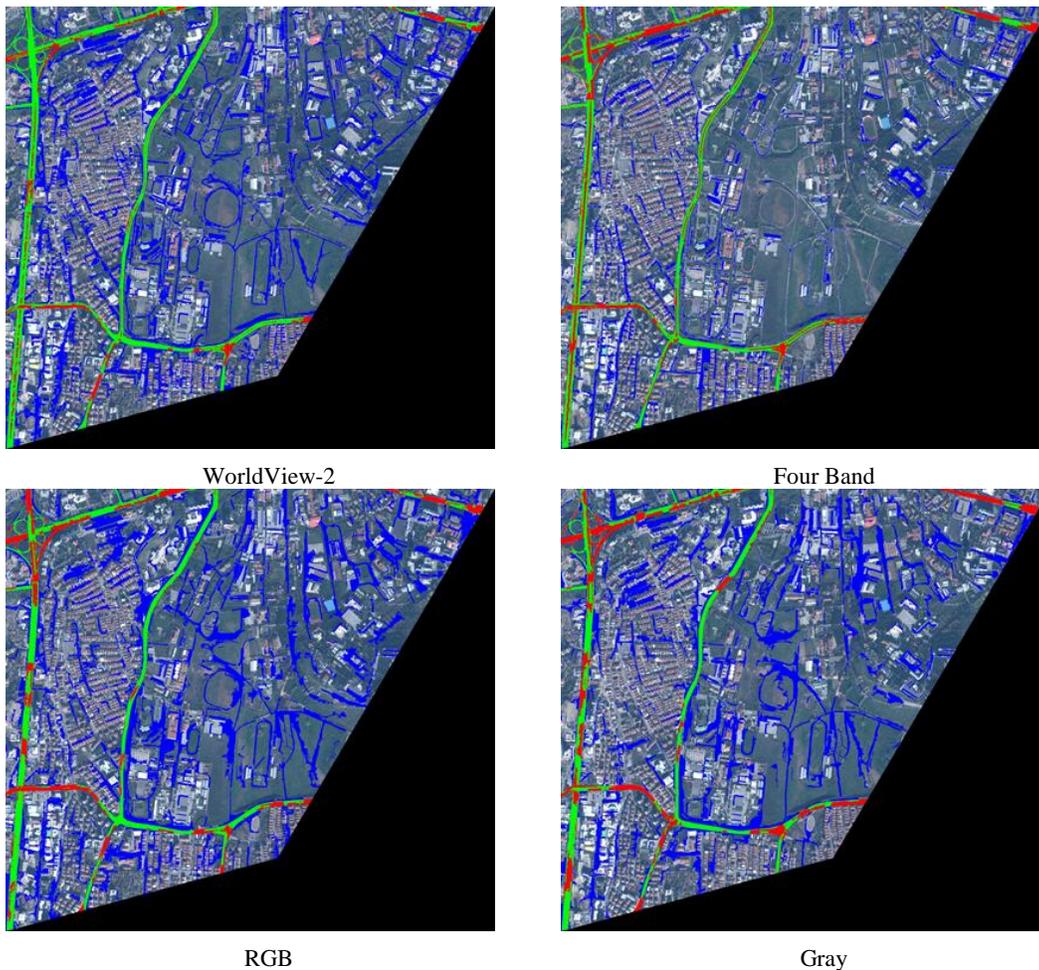
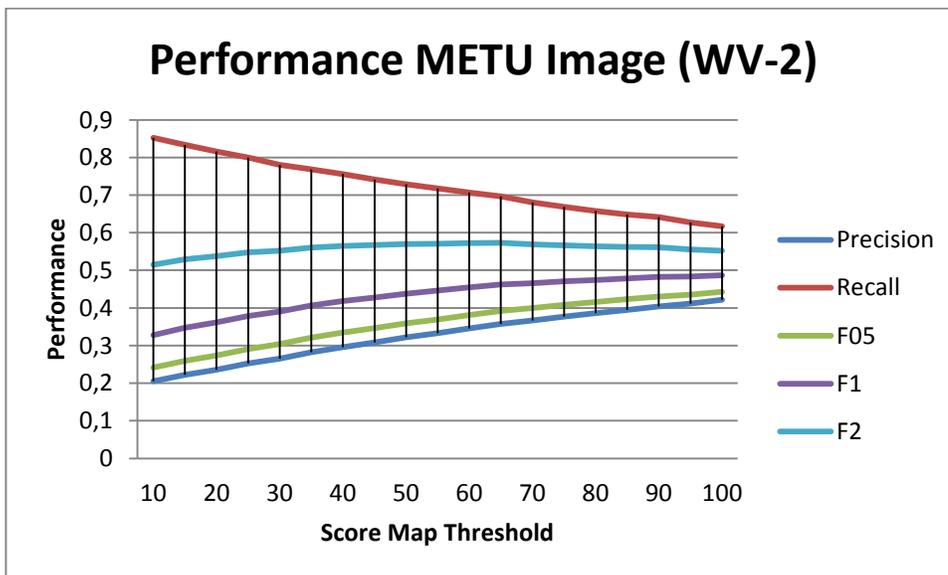


Figure 119 All results of METU (6)



APPENDIX B: STRUCTURE BASED ANALYSIS PERFORMANCES OF ALL BAND COMBINATION

Four Band Combinations

Band1	Band2	Band3	Band4	Recall	Precision	F _{0.5}	F ₁	F ₂
Red	Red Edge	NIR1	NIR2	0,44	0,51	0,49	0,47	0,45
Yellow	Red Edge	NIR1	NIR2	0,47	0,49	0,48	0,48	0,48
Yellow	Red	NIR1	NIR2	0,49	0,51	0,51	0,50	0,50
Yellow	Red	Red Edge	NIR2	0,45	0,47	0,46	0,46	0,45
Yellow	Red	Red Edge	NIR1	0,47	0,46	0,46	0,47	0,47
Green	Red Edge	NIR1	NIR2	0,44	0,46	0,45	0,45	0,45
Green	Red	NIR1	NIR2	0,50	0,46	0,47	0,48	0,49
Green	Red	Red Edge	NIR2	0,51	0,50	0,50	0,50	0,51
Green	Red	Red Edge	NIR1	0,47	0,45	0,45	0,46	0,46
Green	Yellow	NIR1	NIR2	0,42	0,47	0,46	0,44	0,43
Green	Yellow	Red Edge	NIR2	0,47	0,47	0,47	0,47	0,47
Green	Yellow	Red Edge	NIR1	0,46	0,47	0,47	0,47	0,46
Green	Yellow	Red	NIR2	0,46	0,48	0,48	0,47	0,46
Green	Yellow	Red	NIR1	0,46	0,49	0,49	0,48	0,47
Green	Yellow	Red	Red Edge	0,42	0,45	0,44	0,43	0,43
Blue	Red Edge	NIR1	NIR2	0,43	0,50	0,49	0,46	0,44
Blue	Red	NIR1	NIR2	0,45	0,46	0,46	0,45	0,45
Blue	Red	Red Edge	NIR2	0,51	0,47	0,47	0,49	0,50
Blue	Red	Red Edge	NIR1	0,48	0,48	0,48	0,48	0,48
Blue	Yellow	NIR1	NIR2	0,41	0,50	0,48	0,45	0,42
Blue	Yellow	Red Edge	NIR2	0,48	0,50	0,50	0,49	0,49
Blue	Yellow	Red Edge	NIR1	0,46	0,50	0,49	0,48	0,46
Blue	Yellow	Red	NIR2	0,51	0,46	0,47	0,48	0,50
Blue	Yellow	Red	NIR1	0,52	0,50	0,50	0,51	0,51
Blue	Yellow	Red	Red Edge	0,53	0,51	0,51	0,52	0,53
Blue	Green	NIR1	NIR2	0,50	0,49	0,50	0,50	0,50
Blue	Green	Red Edge	NIR2	0,46	0,45	0,45	0,46	0,46
Blue	Green	Red Edge	NIR1	0,45	0,49	0,48	0,47	0,46
Blue	Green	Red	NIR2	0,51	0,47	0,48	0,49	0,50
Blue	Green	Red	NIR1	0,48	0,47	0,47	0,47	0,48

Blue	Green	Red	Red Edge	0,44	0,51	0,49	0,47	0,45
Blue	Green	Yellow	NIR2	0,44	0,47	0,46	0,45	0,45
Blue	Green	Yellow	NIR1	0,47	0,46	0,46	0,46	0,47
Blue	Green	Yellow	Red Edge	0,49	0,50	0,50	0,49	0,49
Blue	Green	Yellow	Red	0,44	0,44	0,44	0,44	0,44
Coastal	Red Edge	NIR1	NIR2	0,46	0,51	0,50	0,48	0,47
Coastal	Red	NIR1	NIR2	0,49	0,52	0,51	0,50	0,50
Coastal	Red	Red Edge	NIR2	0,48	0,47	0,47	0,47	0,48
Coastal	Red	Red Edge	NIR1	0,43	0,47	0,46	0,45	0,44
Coastal	Yellow	NIR1	NIR2	0,43	0,57	0,53	0,49	0,45
Coastal	Yellow	Red Edge	NIR2	0,43	0,48	0,47	0,46	0,44
Coastal	Yellow	Red Edge	NIR1	0,44	0,48	0,47	0,46	0,45
Coastal	Yellow	Red	NIR2	0,46	0,47	0,46	0,46	0,46
Coastal	Yellow	Red	NIR1	0,49	0,51	0,50	0,50	0,49
Coastal	Yellow	Red	Red Edge	0,50	0,47	0,47	0,48	0,49
Coastal	Green	NIR1	NIR2	0,44	0,48	0,47	0,46	0,45
Coastal	Green	Red Edge	NIR2	0,47	0,49	0,49	0,48	0,48
Coastal	Green	Red Edge	NIR1	0,48	0,45	0,45	0,46	0,47
Coastal	Green	Red	NIR2	0,45	0,48	0,47	0,46	0,46
Coastal	Green	Red	NIR1	0,52	0,52	0,52	0,52	0,52
Coastal	Green	Red	Red Edge	0,51	0,47	0,47	0,49	0,50
Coastal	Green	Yellow	NIR2	0,45	0,51	0,50	0,48	0,46
Coastal	Green	Yellow	NIR1	0,47	0,47	0,47	0,47	0,47
Coastal	Green	Yellow	Red Edge	0,49	0,51	0,50	0,50	0,50
Coastal	Green	Yellow	Red	0,46	0,46	0,46	0,46	0,46
Coastal	Blue	NIR1	NIR2	0,48	0,52	0,51	0,50	0,49
Coastal	Blue	Red Edge	NIR2	0,47	0,46	0,46	0,47	0,47
Coastal	Blue	Red Edge	NIR1	0,45	0,48	0,48	0,47	0,46
Coastal	Blue	Red	NIR2	0,53	0,50	0,51	0,51	0,52
Coastal	Blue	Red	NIR1	0,46	0,52	0,51	0,49	0,47
Coastal	Blue	Red	Red Edge	0,42	0,48	0,47	0,45	0,43
Coastal	Blue	Yellow	NIR2	0,44	0,47	0,46	0,46	0,45
Coastal	Blue	Yellow	NIR1	0,46	0,54	0,53	0,50	0,48
Coastal	Blue	Yellow	Red Edge	0,52	0,49	0,50	0,51	0,52
Coastal	Blue	Yellow	Red	0,48	0,50	0,49	0,49	0,48
Coastal	Blue	Green	NIR2	0,49	0,46	0,47	0,47	0,48
Coastal	Blue	Green	NIR1	0,47	0,49	0,48	0,48	0,47

Coastal	Blue	Green	Red Edge	0,50	0,49	0,49	0,49	0,50
Coastal	Blue	Green	Red	0,49	0,49	0,49	0,49	0,49
Coastal	Blue	Green	Yellow	0,44	0,47	0,46	0,45	0,44
Avarege				0,47	0,48	0,48	0,48	0,47

Three Bands

Band1	Band2	Band3	Recall	Precision	F _{0.5}	F ₁	F ₂
Coastal	Blue	Green	0,53	0,47	0,48	0,50	0,52
Coastal	Blue	Yellow	0,52	0,50	0,50	0,51	0,51
Coastal	Blue	Red	0,50	0,45	0,46	0,48	0,49
Coastal	Blue	Red Edge	0,50	0,50	0,50	0,50	0,50
Coastal	Blue	NIR1	0,50	0,48	0,49	0,49	0,50
Coastal	Blue	NIR2	0,50	0,50	0,50	0,50	0,50
Coastal	Green	Yellow	0,50	0,48	0,48	0,49	0,50
Coastal	Green	Red	0,50	0,50	0,50	0,50	0,50
Coastal	Green	Red Edge	0,49	0,47	0,47	0,48	0,49
Coastal	Green	NIR1	0,49	0,56	0,54	0,52	0,50
Coastal	Green	NIR2	0,49	0,51	0,51	0,50	0,49
Coastal	Yellow	Red	0,49	0,45	0,45	0,47	0,48
Coastal	Yellow	Red Edge	0,48	0,51	0,51	0,50	0,49
Coastal	Yellow	NIR1	0,48	0,46	0,47	0,47	0,48
Coastal	Yellow	NIR2	0,48	0,47	0,47	0,48	0,48
Coastal	Red	Red Edge	0,48	0,44	0,44	0,46	0,47
Coastal	Red	NIR1	0,48	0,49	0,48	0,48	0,48
Coastal	Red	NIR2	0,48	0,50	0,49	0,49	0,48
Coastal	Red Edge	NIR1	0,47	0,45	0,46	0,46	0,47
Coastal	Red Edge	NIR2	0,47	0,46	0,46	0,46	0,47
Coastal	NIR1	NIR2	0,47	0,49	0,49	0,48	0,48
Blue	Green	Yellow	0,47	0,47	0,47	0,47	0,47
Blue	Green	Red	0,47	0,47	0,47	0,47	0,47
Blue	Green	Red Edge	0,47	0,51	0,50	0,48	0,47
Blue	Green	NIR1	0,46	0,50	0,49	0,48	0,47
Blue	Green	NIR2	0,46	0,52	0,51	0,49	0,47
Blue	Yellow	Red	0,46	0,49	0,48	0,48	0,47
Blue	Yellow	Red Edge	0,46	0,52	0,50	0,49	0,47
Blue	Yellow	NIR1	0,46	0,50	0,49	0,48	0,47
Blue	Yellow	NIR2	0,46	0,48	0,48	0,47	0,47
Blue	Red	Red Edge	0,46	0,45	0,45	0,45	0,46

Blue	Red	NIR1	0,46	0,46	0,46	0,46	0,46
Blue	Red	NIR2	0,46	0,46	0,46	0,46	0,46
Blue	Red Edge	NIR1	0,46	0,45	0,45	0,45	0,46
Blue	Red Edge	NIR2	0,45	0,48	0,47	0,47	0,46
Blue	NIR1	NIR2	0,45	0,48	0,48	0,47	0,46
Green	NIR1	NIR2	0,44	0,48	0,47	0,46	0,45
Green	Red Edge	NIR2	0,44	0,48	0,48	0,46	0,45
Green	Red Edge	NIR1	0,44	0,51	0,49	0,47	0,45
Green	Red	NIR2	0,44	0,48	0,47	0,46	0,45
Green	Red	NIR1	0,44	0,48	0,47	0,46	0,45
Green	Red	Red Edge	0,44	0,49	0,48	0,47	0,45
Green	Yellow	NIR2	0,44	0,47	0,46	0,45	0,44
Green	Yellow	NIR1	0,44	0,45	0,45	0,45	0,44
Green	Yellow	Red Edge	0,44	0,48	0,47	0,46	0,44
Green	Yellow	Red	0,43	0,44	0,44	0,43	0,43
Yellow	NIR1	NIR2	0,43	0,50	0,49	0,46	0,44
Yellow	Red Edge	NIR2	0,43	0,46	0,45	0,44	0,43
Yellow	Red Edge	NIR1	0,43	0,47	0,46	0,45	0,43
Yellow	Red	NIR2	0,42	0,48	0,47	0,45	0,43
Yellow	Red	NIR1	0,42	0,46	0,45	0,44	0,43
Yellow	Red	Red Edge	0,41	0,41	0,41	0,41	0,41
Red	Red Edge	NIR1	0,41	0,43	0,43	0,42	0,41
Red	Red Edge	NIR2	0,40	0,47	0,46	0,44	0,42
Red	NIR1	NIR2	0,39	0,44	0,43	0,41	0,40
Red Edge	NIR1	NIR2	0,36	0,48	0,45	0,41	0,38
		Average	0,46	0,47	0,47	0,46	0,46

One Band

Band	Recall	Precision	F _{0.5}	F ₁	F ₂
Coastal	0,382596	0,486863	0,461698	0,428478	0,399717
Blue	0,442923	0,425185	0,428618	0,433873	0,439258
Green	0,43242	0,423312	0,425103	0,427818	0,430567
Yellow	0,402127	0,411211	0,409361	0,406618	0,403911
Red	0,410872	0,419455	0,41771	0,415119	0,41256
Red Edge	0,356751	0,491179	0,456757	0,413309	0,37741
NIR1	0,379385	0,458114	0,439858	0,415049	0,392889
NIR2	0,343178	0,474091	0,440485	0,39815	0,363239
Average	0,393782	0,448676	0,434949	0,417302	0,402444

APPENDIX C: WATER BODY EXTRACTION RESULTS



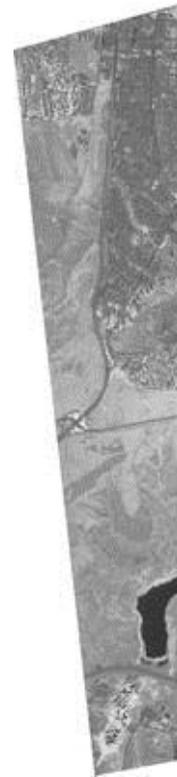
a) RGB



b) NDWI



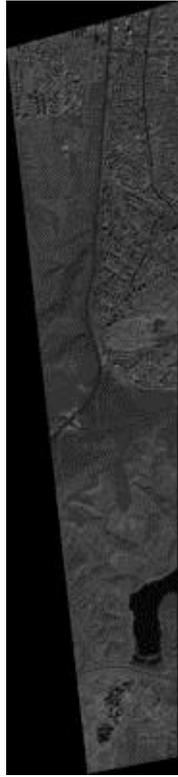
c) NDWI Mask



d) NDVI



e) NDVI Mask (with NIR2)



f) NIR2



g) NIR2 Mask



h) Result

Figure 120 Water extraction for Eymir



a) RGB



b) NDWI



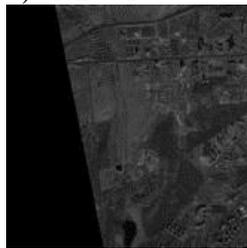
c) NDWI Mask



d) NDVI



e) NDVI Mask (with NIR2)



f) NIR2



g) NIR2 Mask



h) Result

Figure 121 Water Extraction result for METU

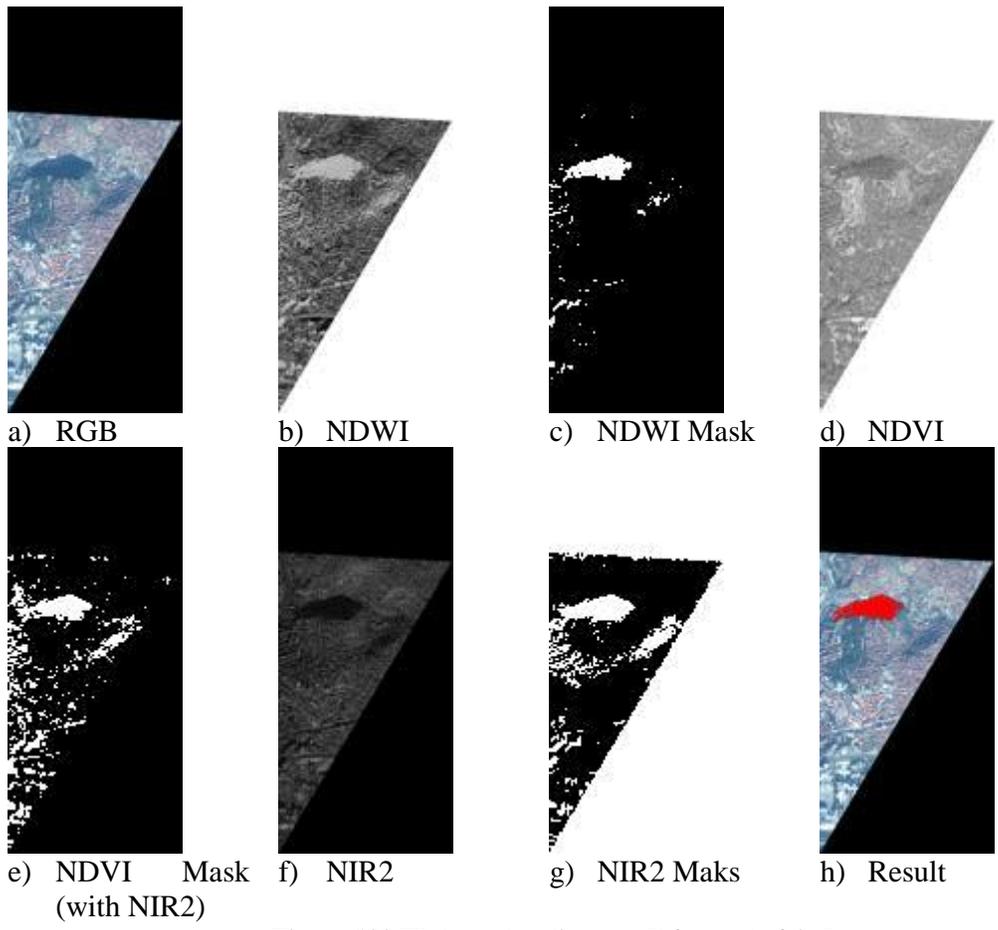


Figure 122 Water extraction result for east of Ankara

APPENDIX D: WATER CANAL EXTRACTION RESULTS

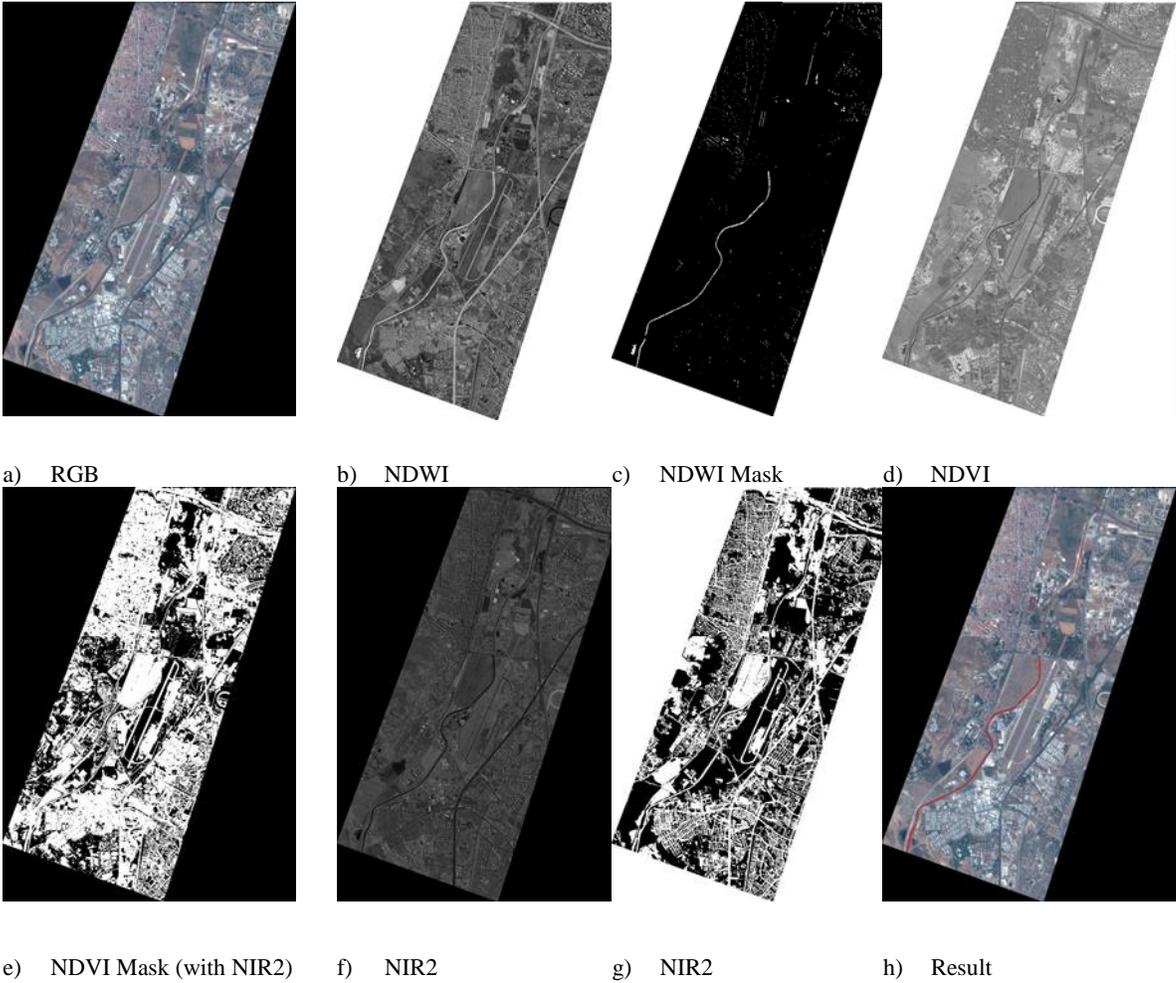


Figure 123 Water extraction for Yenimahalle

APPENDIX E: PARAMETERS USED IN THE STUDY

Module	Parameter	Default Value
Edge Based Road Detection	Maximum Width	40
	Minimum Width	4
	Minimum Length	80
Structure Analysis Based Road Detection	Minimum Region Area	380
	Range Bandwidth	8
	Spatial Bandwidth	4
	Eccentricity	0,99
	Elongatedness	30
	Segment Area	30
Clustering Based Road Detection	K value	6
	Elongatedness	30
	Area	30
Score Map	Maximum Region Area	20
	Score Map Threshold	15
Structure Analysis Based Non-Road	Minimum Region Area	750
	Range Bandwidth	2
	Spatial Bandwidth	2
	Minimum Segment Width	30
	Minimum Segment Area	500
Water Extraction	Min Area	2000
Vegetation	NDVI Threshold	0,1

APPENDIX F: THE ROLE OF SPECTRAL BAND OF WV-2

<p>Coastal Blue (400-450 nm)</p> <ul style="list-style-type: none"> ▪ New band ▪ Absorbed by chlorophyll in healthy plants and aids in conducting vegetative analysis ▪ Least absorbed by water, and will be very useful in bathymetric studies ▪ Substantially influenced by atmospheric scattering and has the potential to improve atmospheric correction techniques 	<p>Red (630-690 nm)</p> <ul style="list-style-type: none"> ▪ Narrower than the red band on QuickBird and shifted to longer wavelengths ▪ Better focused on the absorption of red light by chlorophyll in healthy plant materials ▪ One of the most important bands for vegetation discrimination ▪ Very useful in classifying bare soils, roads, and geological features
<p>Blue (450-510 nm)</p> <ul style="list-style-type: none"> ▪ Identical to QuickBird ▪ Readily absorbed by chlorophyll in plants ▪ Provides good penetration of water ▪ Less affected by atmospheric scattering and absorption compared to the Coastal Blue band 	<p>Red-Edge (705-745 nm)</p> <ul style="list-style-type: none"> ▪ New band ▪ Centered strategically at the onset of the high reflectivity portion of vegetation response ▪ Very valuable in measuring plant health and aiding in the classification of vegetation
<p>Green (510-580 nm)</p> <ul style="list-style-type: none"> ▪ Narrower than the green band on QuickBird ▪ Able to focus more precisely on the peak reflectance of healthy vegetation ▪ Ideal for calculating plant vigor ▪ Very helpful in discriminating between types of plant material when used in conjunction with the Yellow band 	<p>NIR1 (770-895 nm)</p> <ul style="list-style-type: none"> ▪ Narrower than the NIR1 band on QuickBird to provide more separation between it and the Red-Edge sensor ▪ Very effective for the estimation of moisture content and plant biomass ▪ Effectively separates water bodies from vegetation, identifies types of vegetation and also discriminates between soil types
<p>Yellow (585-625 nm)</p> <ul style="list-style-type: none"> ▪ New band ▪ Very important for feature classification ▪ Detects the “yellowness” of particular vegetation, both on land and in the water 	<p>NIR2 (860-1040 nm)</p> <ul style="list-style-type: none"> ▪ New band ▪ Overlaps the NIR1 band but is less affected by atmospheric influence ▪ Enables broader vegetation analysis and biomass studies

Multispectral imagery has provided great value in helping to understand the earth and the impacts of natural processes and man-made activities. In this paper we will look at three major categories of remote sensing applications: feature classification, bathymetric measurements and vegetative analysis and explore how they can be enhanced with the increased spatial resolution, spectral fidelity and collection capacity of WorldView-2.

(http://worldview2.digitalglobe.com/docs/WorldView-2_8-Band_Applications_Whitepaper.pdf)

Curriculum Vitae

ERSİN KARAMAN

METU, Informatics Institute

Inönü Bulvarı. 06800, Çankaya/ANKARA

karamanersin@gmail.com

Education	
2009-Present	METU - Ankara, Turkey <i>PhD, Informatics Institute, Information Systems Program</i>
2007-2009	METU - Ankara, Turkey <i>MSc, Informatics Institute, Information Systems Program</i>
2002-2006	Atatürk University - Erzurum, Turkey <i>B.S., Kazım Karabekir Education Faculty, Department of Computer Education and Instructional Technologies</i>
1998-2001	Erzurum High School - Erzurum, Turkey
Work Experience	
2006-Present	METU-Ankara, Turkey Research Assistant