

**ISTANBUL TECHNICAL UNIVERSITY ★ EARTHQUAKE ENGINEERING
AND DISASTER MANAGEMENT INSTITUTE**

**EARTHQUAKE DAMAGE IDENTIFICATION
FROM VERY HIGH RESOLUTION POST EVENT IMAGE
USING ADVANCED METHODS IN MATHEMATICAL MORPHOLOGY**

M.Sc. THESIS

Enes Oğuzhan ALATAŞ

Earthquake Engineering and Disaster Management Institute

Earthquake Engineering Programme

DECEMBER 2019

**ISTANBUL TECHNICAL UNIVERSITY ★ EARTHQUAKE ENGINEERING
AND DISASTER MANAGEMENT INSTITUTE**

**EARTHQUAKE DAMAGE IDENTIFICATION
FROM VERY HIGH RESOLUTION POST EVENT IMAGE
USING ADVANCED METHODS IN MATHEMATICAL MORPHOLOGY**

M.Sc. THESIS

**Enes Oğuzhan ALATAŞ
(802171203)**

Earthquake Engineering and Disaster Management Institute

Earthquake Engineering Programme

Thesis Advisor: Asst. Prof. Gülşen TAŞKIN KAYA

DECEMBER 2019

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ DEPREM MÜHENDİSLİĞİ
VE AFET YÖNETİMİ ENSTİTÜSÜ

**DEPREM HASARININ YÜKSEK ÇÖZÜNÜRLÜKLÜ
AFET SONRASI GÖRÜNTÜLERDEN İLERİ MATEMATİKSEL
MORFOLOJİ YÖNTEMLERİ KULLANILARAK BELİRLENMESİ**

YÜKSEK LİSANS TEZİ

**Enes Oğuzhan ALATAŞ
(802171203)**

Deprem Mühendisliği ve Afet Yönetimi Enstitüsü

Deprem Mühendisliği Programı

Tez Danışmanı: Dr. Öğr. Üyesi Gülşen TAŞKIN KAYA

Aralık, 2019

Enes Oğuzhan ALATAŞ, a M.Sc. student of ITU Earthquake Engineering Engineering and Technology 802171203 successfully defended the thesis entitled “EARTHQUAKE DAMAGE IDENTIFICATION FROM VERY HIGH RESOLUTION POST EVENT IMAGE USING ADVANCED METHODS IN MATHEMATICAL MORPHOLOGY”, which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

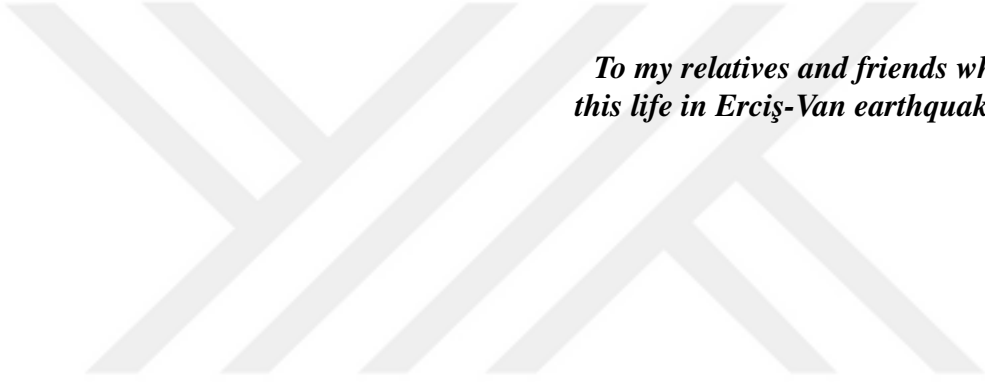
Thesis Advisor : **Asst. Prof. Gülşen TAŞKIN KAYA**
Istanbul Technical University

Jury Members : **Asst. Prof. Tuğrul TURAN**
Istanbul Technical University

Assoc. Prof. Koray KAYABOL
Gebze Technical University

Date of Submission : **15 November 2019**
Date of Defense : **13 December 2019**





*To my relatives and friends who departed
this life in Erciř-Van earthquake of 2011...*



FOREWORD

First of all, I would like to express my deepest appreciation to my advisor Asst. Prof. Gülşen TAŞKIN KAYA for her assistance throughout my master study. She was always a mentor for me, encouraged me to study, answered my all questions in a inspirational way, and did not refrain to be sympathetic.

I would also like to thank Ezgi MENEKAY who has always been with me in all the good and the bad times, without making any discrimination. She was the one who helped me many times that I felt exhausted.

Laslty, I have to thank my family for their endless support, love and trust. I always feel warmth of their love.

DECEMBER 2019

Enes Oğuzhan ALATAŞ
Civil Engineer



TABLE OF CONTENTS

	<u>Page</u>
FOREWORD	ix
TABLE OF CONTENTS	xi
ABBREVIATIONS	xiii
LIST OF TABLES	xv
LIST OF FIGURES	xvii
SUMMARY	xix
ÖZET	xxi
1. INTRODUCTION	1
1.1 Literature Review	2
1.2 Thesis Outline.....	7
2. THEORETICAL BACKGROUND	9
2.1 Remote Sensing Principals	9
2.2 Fundamentals of Machine Learning	10
2.2.1 Classification methods.....	11
2.2.1.1 k Nearest neighbours	11
2.2.1.2 Support vector machines.....	12
2.2.2 Feature selection.....	15
2.2.2.1 Minimum redundancy maximum relevance	15
2.3 Mathematical Morphology	16
2.3.1 Morphological profiles	17
2.3.2 Morphological attribute profiles	19
3. EXPERIMENTS AND RESULTS	21
3.1 An Evaluation of The Proposed Approach.....	22
3.2 Earthquake Damage Assessment: Description of Datasets.....	25
3.3 Morphological Feature Extractions	27
3.4 Classification Schemes	29
3.5 Feature Selection and Subset Decision.....	31
3.6 Results	35
3.6.1 Results: Bam dataset	35
3.6.2 Results: Haiti dataset.....	37
3.7 Damage Maps.....	40
4. CONCLUSIONS AND DISCUSSIONS	45
REFERENCES	47
CURRICULUM VITAE	53



ABBREVIATIONS

RS	: Remote Sensing
ML	: Machine Learning
FS	: Feature Selection
FE	: Feature Extraction
MM	: Mathematical Morphology
MPs	: Morphological Profiles
APs	: Attribute Profiles
kNN	: k-Nearest Neighbour
SVM	: Support Vector Machines
RBF	: Radial Basis Factor
DR	: Dimensionality Reduction
EMS	: European Macro-seismic Scale
KKT	: Karush-Kuhn-Tucker
mRMR	: Minimum Redundancy Maximum Relevance
MI	: Mutual Information
SE	: Structuring Element
VHR	: Very High Resolution
PSI	: Pan-Sharpned Image
OA	: Overall Accuracy
AA	: Average Accuracy



LIST OF TABLES

	<u>Page</u>
Table 3.1 : The created profiles and the expected subsets for each toy data.....	23
Table 3.2 : The data and the corresponding selection results.....	24
Table 3.3 : The toy data and the corresponding selected subsets.....	25
Table 3.4 : Features in each attribute profile sets and their specifications.	28
Table 3.5 : Features in each morphological profile sets and their specifications.	28
Table 3.6 : The number of training and test samples for each class used in the experiments on the Bam study area.	29
Table 3.7 : The number of training and test samples for each class used in the experiments on the Haiti study area.....	29



LIST OF FIGURES

	<u>Page</u>
Figure 2.1 : Examples of bandwidths in each optical imaging system.	10
Figure 2.2 : Effects of neighbourhood number in kNN.	12
Figure 2.3 : SVM’s separating hyperplane and the margin.....	13
Figure 2.4 : An object has different surfaces with relatively different heights and its corresponding pixel representation.	17
Figure 2.5 : SEs with some of the possible shapes and sizes.....	18
Figure 2.6 : Demonstration of the fundamental operators.	18
Figure 2.7 : The difference between operations with reconstruction and without reconstruction.	18
Figure 2.8 : Connected components in an image (a), and their tree representation (b).....	19
Figure 3.1 : Flowchart of the proposed approach.....	21
Figure 3.2 : The created toy data (a) highlights area, (b) highlights moment of inertia, (c) highlights standard deviation.	22
Figure 3.3 : Assigned labels for toy data (a) highlights area, (b) highlights moment of inertia,(c) highlights standard deviation.	23
Figure 3.4 : True color and panchromatic images of the Bam training area.....	26
Figure 3.5 : True color and panchromatic images of the Haiti training area.	26
Figure 3.6 : Representative images from each attribute profile set of the Bam study area.	27
Figure 3.7 : Representative images from morphological profile sets of the Bam study area created by opening and closing operators without reconstruction.....	28
Figure 3.8 : Ground truths created by a visual comparison of pre- and post-event images.....	29
Figure 3.9 : Evaluation of the labeled pixels in the Bam study area.....	30
Figure 3.10 : Evaluation of the labeled pixels in the Haiti study area.	31
Figure 3.11 : Top 40 features ranked with respect to the feature selection for (a) Bam study area and (b) Haiti study area.....	31
Figure 3.12 : PCA based learning curve.	32
Figure 3.13 : Feature frequencies in the mRMR rankings of the Bam extended dataset and their corresponding thresholds.	33
Figure 3.14 : Feature frequencies in the mRMR rankings of the extended dataset from Haiti study area and their corresponding thresholds.....	34
Figure 3.15 : The features selected from all the considered top rankings in each cases, individually and commonly.	35

Figure 3.16: Thematic maps of the Bam study area obtained with k-NN classifier using (a) pixel densities, (b) Subset-1, (c) Subset-2, (d) Subset-3.....	36
Figure 3.17: Confusion matrices of the Bam study area obtained with kNN classifier using (a) pixel densities, (b) Subset-1, (c) Subset-2, (d) Subset-3.....	36
Figure 3.18: Thematic maps of the Bam study area obtained with SVM classifier using (a) pixel densities,(b) Subset-1, (c) Subset-2, (d) Subset-3.....	37
Figure 3.19: Confusion matrices of the Bam study area obtained with SVM classifier using (a) pixel densities, (b) Subset-1, (c) Subset-2, (d) Subset-3.....	37
Figure 3.20: Thematic maps of the Haiti study area obtained with kNN classifier using (a) pixel densities, (b) selected feature subsets from top 10, (c) 20, (d) 40 features.....	38
Figure 3.21: Confusion matrices of the Haiti study area obtained with kNN classifier using (a) pixel densities, (b) selected feature subsets from top 10, (c) 20, (d) 40 features.....	38
Figure 3.22: Thematic maps of the Haiti study area obtained with SVM classifier using (a) Pixel densities, (b) selected feature subsets from top 10, (c) 20, (d) 40 features.....	38
Figure 3.23: Confusion matrices of the Haiti study area obtained with SVM classifier using (a) Pixel densities, (b) selected feature subsets from top 10, (c) 20, (d) 40 features.....	38
Figure 3.24: Selected feature sets for the Bam study area.	39
Figure 3.25: Selected feature sets for the Haiti study area.....	40
Figure 3.26: Damage maps of the entire Bam image obtained using (a) Subset-1, (b) Subset-2.....	42
Figure 3.27: Damage maps of the entire Haiti image obtained using (a) Subset-1, (b) Subset-2.....	43
Figure 3.28: The reference damage map of the Bam dataset.....	44
Figure 3.29: The reference damage map of the Haiti dataset.	44

**EARTHQUAKE DAMAGE IDENTIFICATION
FROM VERY HIGH RESOLUTION POST EVENT IMAGE
USING ADVANCED METHODS IN MATHEMATICAL MORPHOLOGY**

SUMMARY

Our planet is subjected to earthquakes, which is a devastating natural disaster that affects human life, wild life, and their habitats. Earthquake engineering discipline is trying to manage risks and to protect all living creatures from destructive effects of the earthquakes for over a hundred years. One of the most important issues in this discipline is to detect damaged or totally collapsed buildings right after the earthquake. Because the damage information provided in a short time of period could help the decision makers to build a fast emergency plan, and to guide research and rescue teams. Even though the field based surveys can provide a detailed illustration of the earthquake-induced damage, they are not sufficient to rapidly serve critical information. Therefore, remote sensing data sources are becoming a popular direction in the earthquake-induced damage assessment. Light detection and ranging (LIDAR), synthetic aperture radar (SAR), and optical images are three main types of the remote sensing data that are used for earthquake-induced damage detection tasks. Each data type has its own advantages and disadvantages for such tasks. For example, LIDAR images can be useful to detect the damaged in detail, while SAR is not affected by severe weather conditions, or optical images are the most accessible and interpretable type of the data. Thanks to recent technological improvements, optical imaging sensors are able to acquire very high resolution images. Therefore, in this study, a very high resolution post-earthquake image was considered to detect earthquake-induced damage.

For an accurate earthquake damage assessment from very high resolution (VHR) images, contextual relations between pixels need to be included in conjunction with spectral information during the classification. To utilize the spatial information in an efficient way, specific patterns representing the earthquake-induced damage should properly be modelled. Morphological Profiles (MPs) and Attribute Profiles (APs) provide a multi-dimensional representation of an image with a successive implementation of different attribute filters, and they are able to generate the complicated features for a specific pattern. In this study, the APs and the MPs were used to extract the additional contextual features for two different very high resolution post-event satellite images, acquired from City of Bam in Iran and Porto-Prince. These contextual features were then analyzed by means of a feature selection algorithm to find the optimal features, contributing the damage the most, in those profiles. A feature selection method, called Minimum Redundancy Maximum Relevance (mRMR) was used to analyze performance of the morphological and attribute features. A final subset of selected features was analyzed also using two different classifiers, that are k-Nearest Neighbours (kNN) and Support Vector Machines (SVM). The results showed that use of a proper configuration of those profiles can significantly improve the classification accuracy and the quality of the thematic map.



DEPREM HASARININ YÜKSEK ÇÖZÜNÜRLÜKLÜ AFET SONRASI GÖRÜNTÜLERDEN İLERİ MATEMATİKSEL MORFOLOJİ YÖNTEMLERİ KULLANILARAK BELİRLENMESİ

ÖZET

Dünya üzerinde her yıl insan hayatına, çevreye ve ekonomiye ciddi zararlar veren depremler meydana gelmektedir. Mühendisler ve bilim insanları mevcut teknolojiyle bu doğal afetlerin önceden tahmin edilebileceği bir yöntemi henüz geliştirememişlerdir. Bu nedenle depremler insan hayatını bir çok yönden ciddi anlamda etkilemeye devam etmektedir. Deprem Mühendisliği, bu etkileri en aza indirmeyi amaçlayan çalışmaların yapıldığı bir disiplindir. Bu etkilerin en az indirilmesi söz konusu olduğunda çoğu zaman ilk akla gelen yapısal güvenilirlik olmaktadır. Ancak yapısal güvenilirliğin sağlanmaması durumunda da Deprem Mühendisliği disiplini mevcut risklerin yönetimi konusunu da ele almaktadır. Bu risklerin doğru yönetilmesini sağlayan bileşenlerden birisi de deprem sonrası ortaya çıkan yapısal hasarları hızlı bir şekilde belirlenebilmesidir. Yapısal hasarların hızlı belirlenebilmesi, kurtarma çalışmalarının planlanmasıyla daha çok ve daha hızlı hayat kurtarabilmenin yanında, hasarın ekonomik boyutu ve afet sonrası kalkınma planlarının yapılması gibi faydalar da sağlamaktadır. Bu amaç doğrultusunda, afet sonrası hasarın belirlenmesinde uzaktan algılama sistemlerinden yaygın bir şekilde faydalanılmaktadır. Uzaktan algılama verilerinden, mevcut bilimsel birikimin ve teknolojinin de olumlu etkisiyle, görüntü işleme ve makine öğrenmesi teknikleri kullanılarak afet durum değerlendirmesinin yapılması öne çıkan bir araştırma konusudur.

Bu çalışmada deprem nedeniyle meydana gelen yapısal hasarın belirlenerek acil durum planlanmasına kaynak oluşturacak hasar haritalarının belirlenebilmesinde çok yüksek çözünürlüklü optik uydu fotoğraflarının kullanılması konusu ele alınmıştır. Literatürde söz konusu verilerin bu amaçla kullanıldığı bir çok çalışma yapılmış olsa da henüz istenilen hız ve doğruluğa ulaşabilecek genel bir yaklaşıma erişilememiştir. Bu nedenle görüntü işleme ve yapay öğrenme alanlarındaki mevcut gelişmeler de göz önüne alınarak ihtiyaç duyulan hız ve doğrulukta sonuç verebilecek bir yaklaşım oluşturmak amaçlanmıştır. Önerilen yaklaşımın şekillenmesinde ve değerlendirilmesinde vaka çalışması olarak; 12 Ocak 2010'da 7.0 Mw büyüklüğünde meydana gelen Haiti ve 26 Aralık 2003'te 6.6 Mw büyüklüğünde meydana gelen Bam depremlerine ait çok yüksek çözünürlüklü veri setleri kullanılmıştır. Söz konusu depremlerin meydana getirdiği yapısal hasarlar sadece deprem sonrası görüntülerin görüntü işleme teknikleriyle detaylandırılmış ve bu detaylı veriler yapay öğrenme yöntemleri kullanılarak mevcut hasarın belirlenmesinde kullanılmıştır.

İncelenen veri setlerinden Bam şehrine ait olanı depremden sadece sekiz gün sonra Quickbird isimli ticari uydu aracılığıyla çekilmiştir. Bam verisinin oluşturulmasında kullanılan sensör kırmızı, yeşil, mavi ve yakın kızıl ötesi dalga boylarında elde edilen dört kanallı bir multi-spektral veriye ek olarak daha geniş bir dalga boyu aralığında tek kanallı, 0.6 metrelik uzamsal çözünürlükte pankromatik bir veri sunabilmektedir. Haiti şehrinden elde edilen veri setinin alındığı Worldview-1 isimli uydu ise kırmızı, yeşil ve mavi olmak üzere üç kanallı multi-spektral, 0.5 metrelik uzamsal çözünürlükte de pankromatik veri sunmaktadır. Sensörlerin sunduğu multispektral veri parçaları pankromatik verilerle karşılaştırıldığında spektral çözünürlük bakımından daha zengin olsa da, uzamsal çözünürlük açısından çok daha zayıftır. Matematiksel olarak değerlendirildiğinde multispektral kanallar çekilen alandaki malzeme kimyasıyla ilgili daha çok bilgi verirken, pankromatik veriler nesnelere ayırt etmek konusunda daha faydalı bilgi vermektedir. Zira, uzamsal çözünürlük olarak belirtilen büyüklükler veri içerisinde bir piksel içerisinde sığdırılan alanın gerçekte kaç metrekairelik bir alan olduğunu temsil etmektedir. Dolayısıyla bu iki veriyi bir arada kullanmak analizleri daha doğru sonuçlara yönlendirecektir. Bu doğrultuda “pan-sharpening” adı verilen bir işlem uygulanarak iki veri türü birleştirilerek analizlerde kullanılmaktadır.

Söz konusu uzaktan algılama verileri makine öğrenmesi veya başka bir deyişle yapay öğrenme kullanılarak bir çok amaç için sınıflandırılabilir. Örneğin tarım arazileri üzerinde bu sınıflandırma yaklaşımı kullanılarak arazi üzerindeki bitki örtüsü kolaylıkla belirlenebilir. Bu çalışmada sınıflandırma yaklaşımını deprem sonrası hasarlı yapıların tespit edilmesi için kullanılmıştır. Sınıflandırılacak veri her ne kadar yüksek çözünürlükte olsa da böyle zor bir problemin çözülmesi ancak ileri düzey görüntü işleme metodlarının kullanılmasıyla mümkün olacaktır. Bu nedenle çalışmada başarısı benzer analizlerde başarısının literatürdeki bir çok çalışmayla doğrulanmış bir görüntü işleme dalı olan matematiksel morfolojinin ileri yöntemleri kullanılarak mevcut görüntülerden ek özellikler hesaplanmıştır. Bu ileri yöntemler morfolojik profiller ve morfolojik nitelik profilleridir. Bu yöntemlerle sınıflandırma sonucunu iyileştirmek mümkün olsa da, profillerin oluşturulmasında kullanılacak parametreler bölgeye has geometrik özelliklere bağlı olduğundan bir süpervizyon gerektirmektedir. Bu gereklilik de yöntemleri otomatik bir hasar tespit yaklaşımında kullanmayı zorlaştırmaktadır. Fakat başka bir yapay öğrenme dalı olan öznitelik seçme yöntemleri bu konuda bir çözüm üretmek için kullanılabilir. Bu konuda bir çözüm üretmek için kullanılabilir.

Bu çalışmada, incelenen iki farklı bölgeye ait deprem sonrası uydu görüntüsünün içerisinde seçilen, görece olarak küçük alanlara ait, çalışma alanları üzerinden yukarıda bahsedilen morfolojik yöntemler kullanılarak çok fazla sayıda öznitelik oluşturulmuş ve bu öznitelikler öznitelik seçme yöntemlerinden maksimum alaka minimum gereksizlik (ing: maximum relevance minimum redundancy (mRMR)) isimli yöntem uygulanarak öznitelikler bir önem sırasına göre sıralanmışlardır. İlgili deprem görüntüsündeki bütün bölgelerde bu sıralamalar üzerinden elde edilen en önemli özniteliklerle aynı parametrelere sahip öznitelikler hesaplanarak bütün bölgeler sınıflandırılmış ve hasar haritaları elde edilmiştir. Ulaşılan sonuçlar önerilen yöntemin iki ayrı veri üzerinde de sınıflandırma doğruluklarını ve tematik harita kalitelerini önemli ölçüde iyileştirdiğini göstermiştir.

Bu tezde önerilen yaklaşım ve bu yaklaşımda kullanılan bütün yöntemler detaylı bir şekilde açıklanmıştır. Dört bölümden oluşan bu tezin, ilk bölümünde problemin tanımı anlatılarak literatürde daha önce yapılan benzer çalışmalar incelenmiştir. İkinci

bölümde ise kullanılan yöntemlerin teorik altyapıları anlatılmıştır. Üçüncü bölümde ise ele alınan deprem sonrası çok yüksek çözünürlüklü görüntüler için yapılan analizler ve sonuçları ayrıntılı olarak sunulmuştur. Son bölümde ise çalışmanın sonuçları değerlendirilmiş, olası ileri çalışmalar tartışılmıştır.





1. INTRODUCTION

Earthquakes, as the one of the most destructive natural disasters, affect a lot of people in every year. Detecting earthquake-induced building damage, first and foremost, is a vital task since it provides information for search and rescue teams to respond to the emergency. Due to this, the damage assessment should be as accurate and as fast as possible. Remote sensing technologies offer valuable sources to respond to these necessities. A damage map can be quickly generated by using remote sensing imagery, and this damage map can be used for managing the help and rescue operations as well as for evaluating the natural, social and economical impacts and for planning the possible further restorations.

There have been several studies in the literature on the earthquake damage identification by using remote sensing images [1–5]. The damage assessment problem is still an open problem for researchers as it involves many challenges such as the need for a satisfactory accuracy and a quick response time. Therefore, this thesis aims to propose an accurate and a fast earthquake damage identification approach using the state-of-the-art tools along with the image processing and machine learning.

In order to build a fast and an accurate damage assessment, very high resolution (VHR) optical post-earthquake satellite images were considered since they are the most accessible and the most interpretable type of the remote sensing data. The information in the original data was spatially enhanced by a proper implementation of mathematical morphology tools, which are morphological profiles (MPs) and attribute profiles (APs). However, for such a proper implementation of MPs and APs the corresponding parameters can not be selected in the first step since both of the methods can give different output images based on different datasets with the same parameters. Because the outputs depend not only on the parameters but also on the spatial content of the datasets. Therefore, a data-driven implementation should be designed. This design can be possible by selecting the most relevant features from an extensive feature set generated by using the morphological tools. In this sense, the thesis proposes

an approach based on three steps: i) an extensive morphological feature set from a selected study area on the original post-earthquake image is generated. ii) a feature selection by using maximum relevant minimum redundancy (mRMR) method on this extensive feature set is conducted to select the relevant features for the entire post-event image. iii) with the joint use of those created features and the spectral features of the image, the post-event image is classified into some land cover classes, including damage, to generate a damage map showing the regions, that have light, moderate, and severe damages is created based on the damage intensities in the classified data. In the evaluation of the results, two different classification methods k-Nearest Neighbours (kNN) and support vector machines (SVM) were used to remove the effects of the classifier from the results. In the implementations, two different post-event images, acquired from City of Bam in Iran and Port-au-prince in Haiti, were considered as case studies. The performance of the proposed approach were evaluated in terms of the classification accuracy, and the quality of the thematic maps in particular areas of interest was also interpreted.

To sum up, this thesis proposes to utilize the APs and MPs to assess the earthquake damage from a post-event VHR satellite image. To find the most proper feature configuration describing the damage patterns, a high dimensional data-set was created with extensive configuration intervals (e.g: criterion values, size and shape of structuring elements.), then a supervised feature selection procedure was implemented on this high dimensional dataset. The outcomes of the case studies conducted on were examined if there are essential common feature configurations independent of the characteristics of the dataset, such as the structural characteristics and constructional details of the building. According to the extensive literature search conducted in this thesis, there has been no study reported in the literature extensively exploring the effects of the APs and MPs together on VHR image for an earthquake damage assessment.

1.1 Literature Review

Remote sensing technologies have been used as primary source to response the need for an overall post-earthquake damage assessment for decades [2]. In the vast majority of the related studies, three types of remote sensing data have commonly been used; light

detection and ranging (LIDAR), synthetic aperture radar (SAR), and optical images. Therefore the related studies can be categorized into three main group according to the used data type. For instance, Gamba et al. [6] discussed the performance of optical remote sensing images in earthquake damage assessment by utilizing Geographic Information Systems (GIS). Ito et al. [7] proposed an earthquake damage extraction method for using SAR data by using artificial neural networks. Schweier and Markus [8], used LIDAR data to assess building damage and losses caused by the earthquake. Dong and Shan [1] comprehensively reviewed the studies related to the post-earthquake building damage assessment by using remote sensing in 2013. The review showed that use of different types of the data comes with different advantages and disadvantages. For instance, LIDAR images can be useful to detect the pancake type of building collapse by exploring the changes occurred in the building heights. SAR images are not affected by the cloudy weather conditions, whereas optical images are very useful data for damage assessment due to their easier visual interpretation.

The existing studies are not only dissociate from each other by the data types, but also dissociated by the methods which can be split into two general approaches: mono- and multi-temporal. Mono-temporal approaches identify the damage by interpreting the features of the objects coming from a single post-earthquake image, whereas multi-temporal approaches detect the damage by comparing pre- and post-earthquake images. In 2002, Ishii et al. [9] firstly examined both approaches on an aerial photograph image pair of Kobe city taken before and after the Hyogoken-Nanbu earthquake. They discussed the potential success and shortcomings of those approaches specific to aerial image. The study yielded two important results: in the mono-temporal analysis using only the pixel intensity levels to determine damaged buildings can cause errors since the damaged areas may represent different intensities depending upon its source material, whereas in the multi-temporal analysis shadows lead to several confusions as it can cause high redundant differences between the pre- and post-earthquake images.

When both pre- and post-event data are available, a multi-temporal analysis can be performed for detecting the damaged areas by focusing the changes related to earthquake. Turker and San [10], calibrated the post- and pre-event optical satellite images from Izmit earthquake of 1999 and used the difference between near-infrared bands to detect the damaged areas. The reason that they used the near-infrared

bands is because of its insensitivity to atmospheric variations. Kusogi et al. [11] proposed a nonlinear change detection method and tested it on an aerial image pair of Hanshin-Awaji earthquake. Turker and Cetinkaya [12] detected the collapsed buildings using the idea that the building heights change if the building is collapsed. They used digital elevation models (DEMs) to extract the building heights and detected the damaged buildings according to their height differences between pre- and post-event images based on a threshold value. On the contrary of these data types, SAR data can not be able to be used for such an analysis except for very limited case studies, where suitable pre-event SAR data is available. In 2014, Plank [13] reviewed SAR based multi-temporal damage assessment studies and showed the limitations of the data type for such problem. For instance, at least three SAR images are required for the damage assessment, or after a disaster SAR amplitude does not only increase, the amplitude can also decrease and this fluctuation causes several misdetections.

Multi-temporal approaches might not be very feasible in terms of time efficiency since they require the pre-processing steps, including registration and correction of radiometric and atmospheric effects to make both pre- and post-images comparable [1]. Therefore, use of a single post-earthquake image in earthquake damage assessment is an effective alternative way to a image-to-image comparison. Vu et al. [14], in 2007, succeed to generate a damage map that is consistent with the field-based investigations by using only post-event data. Trianni and Gamba [15] tested the joint use of post-event and ancillary data, such as GIS, and showed that use of extra spectral and spatial information are required for an accurate damage assessment. However, in this case, a very high resolution optical image should be used for an accurate damage identification, which makes the damage assessment task even more difficult due to the large variations occurred in the urban areas. Especially the damage, the open ground, and the building classes might be very challenging classes to be detected using only a post-event image because of the possible geometric irregularities, different reflections, or strong similarities between the spectral values of these classes. To address such difficulties, post-event image are mostly used by incorporating spatial context into per-pixel spectral classification. This spatial context can be extracted based on textural or geometrical characteristics of the objects.

There have been several studies in the literature focusing on the textural patterns, which express spatial distributions of variations in pixel densities [16–18]. Rathje et al. [19] used the textural features combined with the spectral features on the VHR post-event image acquired from Bam earthquake and showed that the damaged areas are better identified with additional textural features compared to pixel-based classification. Sun et al. [20] used the different types of textural features based on gray level histograms, gray level co-occurrence matrices (GLCMs), local binary patterns, gaussian markov random fields, and gabor filters, for determining the collapsed buildings from the post-event images. After a timely expensive processing, a comprehensive experiments conducted with 122 features yields a conclusion that no specific textural features exist representing the all types of damage and building patterns in this feature set.

To extract geometrical information from the VHR earthquake images, mathematical morphology has commonly been used as an image processing tool [21–23]. Chini et al. [24] proposed a damage detection approach based on the fusion of SAR and VHR images using a mathematical morphology tool, called morphological profiles (MPs) [25], which implements a sequence of the opening and closing operators to extract additional features. They created a building mask with MPs from the VHR image as to use a reference data for change detection and presented the potential benefits of using the MPs in the damage detection tasks. Dell’Acqua et al. [26] also showed that MPs are very effective tools in earthquake damage identification when using a single post-event image. Wang et al. [27] also proved the contribution of the Mathematical Morphology to the damage assessment. They proposed a method that includes different morphological operations, such as closing, erase etc., to extract the damaged parts of the roads after an earthquake.

Besides, the geometrical and textural information can be complementary to each other when using them together in the damaged building extraction. Ma and Quin [28], developed a method that uses both textural and morphological information together in mono-temporal damage assessment process. Li et al. [29] used MPs and textural features together on a post-earthquake image with 4 different window sizes and the study achieved a reasonable improvement in the collapsed building detection from the post-event images. However, their method also resulted apparent errors and they stated

that a future study, that includes more features and seeks the most relevant features through adopting a feature reduction method, may be valuable for the literature.

In the literature of the mathematical morphology, Dalla Mura et al. [30] have recently used attribute profiles (APs) and multi-attribute profiles (MAPs) to extract geometrical and textural information from VHR images. They created APs using attribute filters sequentially and compared the performance of the APs to the conventional MPs resulting in that the APs outperforms MPs in terms of both classification accuracy and computational time. Although both methods have proved their efficiency in image understanding, the parameter selection for constructing elements is a difficult task to consider. A solution constructing the most possible profiles with very high number of parameter between very small intervals might be useful, but this approach dramatically increases the number of dimensions and also brings a very high computational time, as well as need for more training samples [31]. As a result of this, there can be a large number of redundant features in the dataset. The dimensionality reduction is a way to deal with such difficulties. In 2013, Pedernana et al. [32] used unsupervised manifold embedding and supervised feature selection methods to reduce this high dimensionality, and they showed the potential success of the dimensionality reduction techniques in solving such problems. However, manifold embedding methods requires all the extracted features of the entire area of interest to be classified, meaning that they are not very convenient to be used in classification tasks. Therefore, feature selection techniques are more feasible in such cases since they can be used to select a representative subset from those high dimensional feature profiles. This selection can be conducted both supervised and unsupervised way. In 2008, Bhardwaj and Patra [33] proposed an unsupervised feature selection method to select the most relevant features in the high dimensional morphological profiles. The unsupervised feature selection method is very efficient in terms of computational time, supervised methods are more credible in terms of classification accuracy for such the problems.

The damage maps, also called as thematic maps, provide information about pixels or segments whether they are collapsed or not. However, determining the intensity of earthquake-induced building damage is another problem to solve. Earthquake engineers developed many different intensity scales to express the damage levels in different areas of the subjected cities [34]. Remote sensing based damage detection

applications can be ended up with the damage maps created with the implementation of the intensity scales. In the studies creating a damage map by using remote sensing datasets, European Macroseismic Scale 1998 (EMS-98) [35] is one of the mostly used intensity scales for mapping the impact of earthquakes. Yamazaki et. al. [36] used EMS-98 to interpret the building damage in five different levels and created a damage map of Bam Earthquake based on these classified damage information by a multi-temporal analysis. In 2013, Rastiveis et. al. [37] proposed a fuzzy decision making system considering the building shape differences between pre- and post-event images and created a five level damage map. In some cases, interpreting the damage with five different scales may not be possible. In such cases, the damage can be classified into less number of levels due to some limitations of remote sensing data, which for example remote sensing can not capture the pancake type building collapse. Romaniello et. al. [38] conducted a case study on 2011 Haiti Earthquake and created a damage map with three different classes, that are heavy, moderate, and light damages. They also reported the accuracy loss in comparison to detection of only collapsed buildings. In the mono-temporal approaches, it is much more challenging to identify the damage levels since the approach has not another source to decide the intensity of damage.

1.2 Thesis Outline

This thesis consists of four chapters. The first chapter introduces the definition of the problem and motivation of the thesis followed by giving a literature review to support the motivation. The second chapter explains the details of the proposed approach and provides a brief introduction to the machine learning (ML) techniques considered in this study followed by describing the image processing tools that are used to extract the features for a better representation of damage patterns. An overview of feature selection (FS) methods are also given in this chapter. The third chapter focuses on the experimental studies conducted in this thesis by using the proposed approach. The last chapter presents the discussion of experimental results, the conclusions, and the problems addressed for possible future researchers.



2. THEORETICAL BACKGROUND

For making an accurate earthquake damage assessment possible from the VHR images, contextual relations among the pixels should be used hand in hand. This contextual relations can be derived from modelling specific patterns, representing the earthquake damage. Morphological Profiles (MPs) and Attribute Profiles (APs) are able to generate complicated features for those specific patterns. It can be possible to create a great variety of features with different configurations of the APs and the MPs, such as different attributes, different criterion values, or different shapes and size of structuring element (SE). In that variety, an optimal subset, that has the most separability power, should be determined. To determine such a subset, a feature selection process can be used.

To be brief, the thesis proposes a damage detection approach exploiting the theories of Remote Sensing, Mathematical Morphology, and Machine Learning. Therefore, this chapter focuses on those topics especially by explaining briefly the methods considered in this thesis.

2.1 Remote Sensing Principals

In the most general sense, remote sensing is a process that measures the physical characteristics of an area by remotely measuring reflected and absorbed radiation coming from the objects. The sensors mounted on the satellites take images of the Earth's surface that offer a deep understanding about the objects on the ground. The different imaging systems can be developed by using different types of sensor technologies; hence providing the different types of the data. For instance, optical imaging systems use passive sensors that measure the reflected solar radiation at different wavelengths, whereas SAR imaging systems use active sensors that transmit microwave pulses to the target object and create an image based on time difference between transmitted and back-scattered pulses.

Optical remote sensing systems are generally classified into four different types including panchromatic imaging, multi-spectral imaging, super-spectral imaging, and hyper-spectral imaging. The panchromatic imaging systems use a sensor sensitive to radiation within a wide wavelength range. The measured physical quantity in panchromatic systems is the brightness of the target objects. In multi-spectral remote sensing systems, the sensor collects a multi-channel data in a certain wavelength range. The spectral channels are sensitive to radiation within different narrow wavelength bands. The resulting image contains both the brightness and spectral information of the targets being observed. If the sensors, which have more than ten channels in narrower wavelength bands are used, then the imaging system is called super-spectral. As for hyper-spectral imaging systems, the sensors have capability to capture reflectance in more than 100 adjacent spectral bands. The bandwidths in images from all of those optical imaging systems were illustrated in Figure 2.1.

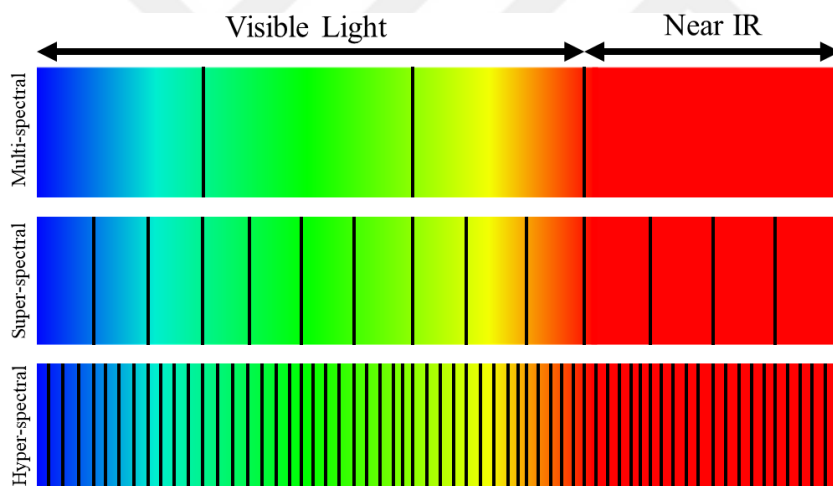


Figure 2.1 : Examples of bandwidths in each optical imaging system.

2.2 Fundamentals of Machine Learning

The analysis of remote sensing data reveal a lot of information about the object on the Earth, therefore machine learning is a very effective tool for knowledge extraction for remote sensing images. Machine learning optimizes a specific criterion value subject to some constraints to build a learning model that extracts descriptive or inferential knowledge from the past experience [39]. This model can be a boundary to classify instances in a population, a set of cluster centres to cluster a group, or a function which estimates the outcome of the new observations.

The machine learning methods that are used in this study can be split into two categories as classification and feature selection methods which have been explained in the following subsections.

2.2.1 Classification methods

Classification is a supervised ML application that requires a labelled dataset to be used during the learning process. The classification methods basically split the data into several classes by providing class boundaries, so called decision boundary, the features of the given data. In this study, k-Nearest Neighbours (kNN) and Support Vector Machines (SVM) are mostly used classification methods in the analysis of the remote sensing data. These two classification methods are explained in the following subsections.

2.2.1.1 k Nearest neighbours

k-Nearest Neighbours method classifies an instances with respect to the the most frequent label of the number of k nearest labelled samples based on the closeness of the neighbours which is defined by a distance function [40]. Those distance functions could be depended on the numerical or categorical features of the data. The distance function can be based on different distance measures such as Euclidean, Hamming, Manhattan, or Chebychev. The number of neighbours, so called k, is a parameter that needs to be optimized in kNN. In order to find an optimum k, 2-fold cross validation data is used with different k numbers, then the optimum k is selected based on the highest average accuracy. A visual example to demonstrate the effects of the k is shown in Figure 2.2 for a synthetic dataset. The sample represented by a black solid square is classified with different number of k nearest neighbours as k = 1,2 and 3, respectively. In the synthetic dataset the new sample is classified into the blue class for the 1- and 2-NN cases, whereas the 3-NN model classifies the sample into the red class.

In terms of computational complexity, kNN is much more simple than other classification algorithms. However, the less complexity does not mean the less accuracy [41]. It should be noted that performance of a classifier always depends

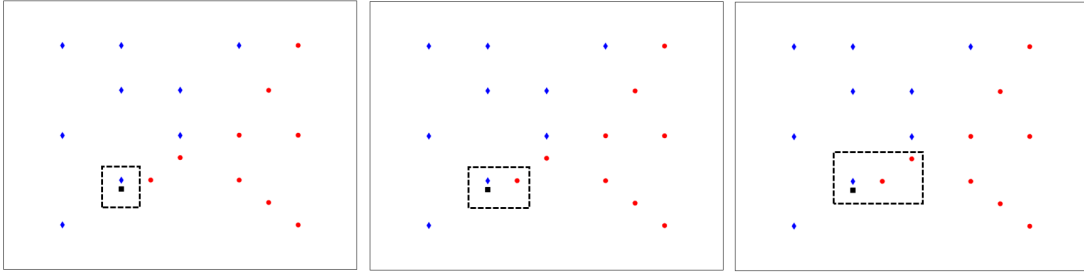


Figure 2.2 : Effects of neighbourhood number in kNN.

on the data to be classified, and there is not possible to create a model that fits all types of data [42].

2.2.1.2 Support vector machines

Support vector machines (SVM) is a classification method based on the idea that: estimation of the actual class densities is not required for separating two classes, but an estimation of the class boundaries is sufficient to separate the classes [43, 44]. In this sense, the method simply classifies the instances by creating an optimal hyperplane by maximizing the margin between the classes [45]. The concept of SVM is explained as follows through an example for a linearly separable data with two classes in two dimensional space.

$$X = \{(x_1, c_1), \dots, (x_m, c_m), x_i \in \mathbb{R}^2, c_i \in \{-1, +1\}\} \quad (2.1)$$

where X is a set of training samples consisting of x_i which corresponds to location vector of the i^{th} sample. c_i corresponds to class label of associated i^{th} sample. To separate this data into two classes, a discrimination function, also called a decision boundary, $f(x)$ is defined as follows:

$$f(x) = w^T x_i + w_0 \quad (2.2)$$

The discriminant function should also satisfy the following constraint related to the class labels such that

$$r_i(w^T x_i + w_0) \geq 1 \quad (2.3)$$

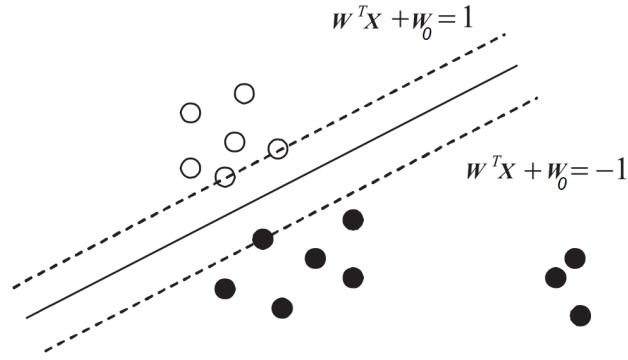


Figure 2.3 : SVM's separating hyperplane and the margin.

It should be noted that the constraint in Equation 2.3 drives the hyperplane to some distance from the classes, not to the exact boundary of any class. This distance, called margin, is the distance between the hyperplane and the instances that are the closest ones to the hyperplane, for both classes. SVM classifier tries to maximize this distance to separate classes with the best generalization capacity. There are an infinite number of possible hyperplanes in the considered space \mathbb{R} . However, there is one unique hyperplane with the maximum margin. Thus, the margin term has to be maximised in order to find the optimal hyperplane. As can be seen in Fig. 2.3, the margin is in an inverse proportion of the weight vector w . Therefore, the optimization problem can be expressed as:

$$\min_{w, w_0} \quad \frac{1}{2} \|w\|^2 \quad (2.4a)$$

$$\text{subject to} \quad r_i(w^T x_i + w_0) \geq 1 \quad (2.4b)$$

The constrained optimization problem can be solved by implementing Lagrange multipliers α_i . Then, the problem in the unconstrained form can be expressed as in Equation 2.5.

$$L(w, w_0, \alpha) = \frac{1}{2} \|w\|^2 + \sum_{i=1}^m \alpha_i \{r_i(w^T x_i + w_0) - 1\} \quad (2.5)$$

The problem needs to be more manipulated by expressing the w and w_0 in terms of the Lagrange multipliers with the use of Karush-Kuhn-Tucker (KKT) conditions.

After the implementation of KKT conditions, the problem becomes a dual optimization problem yielding as follows:

$$L(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j r_i r_j (x_i^T x_j) \quad (2.6a)$$

$$\text{subject to} \quad \sum_{i=1}^n \alpha_i r_i = 0, \quad \alpha_j \geq 0 \quad (2.6b)$$

Equation 2.6 can be solved by a quadratic programming approach. The samples having nonzero α_i s are called support vectors. Finally, this solution makes the discriminant function able to analytically separate instances. The decision boundary is a linear combination of the support vectors. All the expressions given in above are based on the assumption that the data is linearly separable. However, most of the real world problems the is not sufficient to provide a satisfactory accuracy especially for the nonlinearly separable data. By utilizing the nonlinear kernel functions, data can be mapped into a higher dimensional space in which the data can be classified more accurate. There are three types of the kernel functions generally used in this context: polynomial, hyperbolic tangent function, and radial basis factor (RBF). In this study, RBF kernel was used due to its confirmed suitability to similar problems [46,47]. The RBF kernel for two samples is determined by the following equation:

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2} \quad (2.7)$$

where γ is called as kernel width parameter.

In the implementation of the nonlinear SVM, the parameter γ in the kernel function and a penalty parameter for the second term of Equation 2.6 called C has to be tuned to maximize the performance of classifier and its generalization capacity. This selection, also called model selection in the literature, can be done by using a k-fold cross validation. The k-fold cross validation is a simple concept that folds the training data into number of k folds and conducts pre-executions using the folded parts of the data as training and validation parts. Based on this idea, a grid search can be implemented to select the most relevant parameters for the problem [48].

2.2.2 Feature selection

In machine learning applications data is handled with respect to variation of instances in its features. It can be possible to analyse the data sufficiently with the less number of features than it has. Furthermore, it is also possible to reach a better success level with the less number of features. Dimensionality reduction framework tries to find or to compute those valuable features. The methods in that frame work are generally categorized as feature extraction (FE) and feature selection (FS). FS methods seek for a feature subset that have more ability to represent underlying structure of the data than others, while FE methods compute new features based on the originals to decrease the dimensionality. As can be clearly understood from the definitions, FS has an advantage over FE since that the physical meaning of the data is preserved.

In this study, a great number of features were extracted by using image processing tools called MPs and APs and a feature selection method was used in order to not only decrease such a high dimensionality and but also to look for a subset that contains optimal synthetic features. In this sense, a feature selection method, named Minimum Redundancy Maximum Relevance (mRMR), was used due to its proved success on similar analyses reported in the literature [49–51].

2.2.2.1 Minimum redundancy maximum relevance

Minimum Redundancy Maximum Relevance (mRMR) [52] is a feature selection method that measures feature-feature and feature-label correlations according to mutual information (MI), expressed in Equation 2.8.

$$MI(x,y) = \sum_{x \in O_x, y \in O_y} p(x,y) \log \left(\frac{p(x,y)}{p(x)p(y)} \right) \quad (2.8)$$

where O_x and O_y correspond to the instance spaces for x and y , respectively. While $p(x,y)$ is the joint probability of x and y . $p(x)$ and $p(y)$ are the marginal probabilities.

The feature-feature and feature-label correlations are called redundancy and relevance, respectively. The method solves a dual optimization problem iteratively to sort the features from the most relevant to the least. The objective of the optimization problem introduced in Equation 2.9 is maximizing the relevance V , while minimizing the redundancy W .

$$\min_W \quad W = \frac{1}{|S|^2} \sum_{i,j \in S} MI(i, j) \quad (2.9a)$$

$$\max_V \quad V = \frac{1}{|S|} \sum_{i \in S} MI(h, i) \quad (2.9b)$$

where S refers to set of features, and $MI(i, j)$ refers to mutual information between i -th and j -th features. $MI(h, i)$ corresponds to mutual information between target class h and i -th feature.

The mRMR feature selection method combines the relevance and the redundancy terms in two different ways, that are shown in Equation 2.10.

$$\max_{V-W} \quad V - W = MI(h, i) - \frac{1}{|S|} \sum_{i,j \in S} MI(i, j) \quad (2.10a)$$

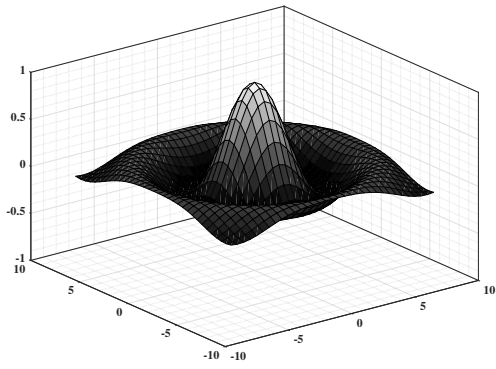
$$\max_{V/W} \quad V/W = MI(h, i) / \left[\frac{1}{|S|} \sum_{i \in S} MI(i, j) \right] \quad (2.10b)$$

The alternatives are called mutual information difference (MID) and mutual information quotient (MIQ), respectively. Based on the data characteristics, MID and MIQ approaches can lead different or similar selection results. In this study, both of the approaches are considered and the MIQ is decided to use since it gives a bit more consistent selection results in comparison to the MID.

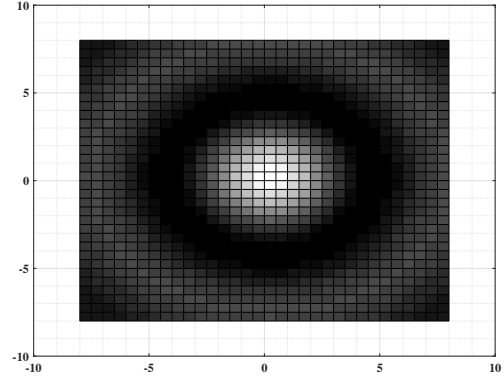
2.3 Mathematical Morphology

Mathematical morphology is a widely used framework in the field of image processing, that is constructed based on simple mathematical concepts from set theory. The components of the framework are used in a wide range of of problems in image analysis such as edge detection, noise reduction, image restoration, or segmentation.

In a gray-scale image, the pixel intensities represent the relative height of a surface in three-dimensional Euclidean space. Morphological operations probe those surfaces according to pixel intensities that represent their heights. Figure 2.4 illustrates the relative heights represented by pixel intensities.



(a) Surfaces in 3D Euclidean space.



(b) Representative pixel intensities.

Figure 2.4 : An object has different surfaces with relatively different heights and its corresponding pixel representation.

2.3.1 Morphological profiles

Morphological Profiles (MPs) are very effective image processing tools when considering to detect heterogeneous structures in an image [25]. The MPs are defined as a cumulative function that combines varied morphological opening and closing operators with different scales of the structuring element as in the following equation:

$$\text{MP}(f) = \{\phi_{\lambda_i}(f), \dots, \phi_{\lambda_1}(f), \gamma_{\lambda_1}(f), \dots, \gamma_{\lambda_i}(f)\} \quad (2.11)$$

where $\phi_{\lambda_i}(f)$ and $\gamma_{\lambda_i}(f)$ are the images created by implementing closing and opening operators, respectively, on an image f with a structuring element λ_i .

To explain the concept of the MPs more clear, the fundamental terms of morphological image processing, such as structuring element, erosion and dilation, should be clarified. The structuring element (SE) is a binary template matrix that situates in all possible locations in the considered image. The SE can be created with different shapes and sizes. SEs with some possible shapes and sizes can be shown in Fig. 2.5. The erosion operator slides the pre-defined SE in the image, and set the minimum value of all pixels in the neighbourhood to pixel that corresponds to the origin of the SE. The dilation analogously set the maximum value to the origin. Sequential use of erosion and dilation, generates morphological opening and closing operations. The opening operator contains a dilation of an eroded image with the same SE. In a similar manner, the closing of an image is an erosion after an dilation of this image. These operations are demonstrated on a binary image illustrated in Fig. 2.6.

The opening and closing operations can be performed with morphological reconstruction to eliminate effects of those operations on the image, such as totally

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (a) \ 3 \times 3 \text{ square SE}$$

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \quad (b) \ 4 \times 4 \text{ disk SE}$$

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad (c) \ 3 \times 3 \text{ line SE}$$

Figure 2.5 : SEs with some of the possible shapes and sizes.

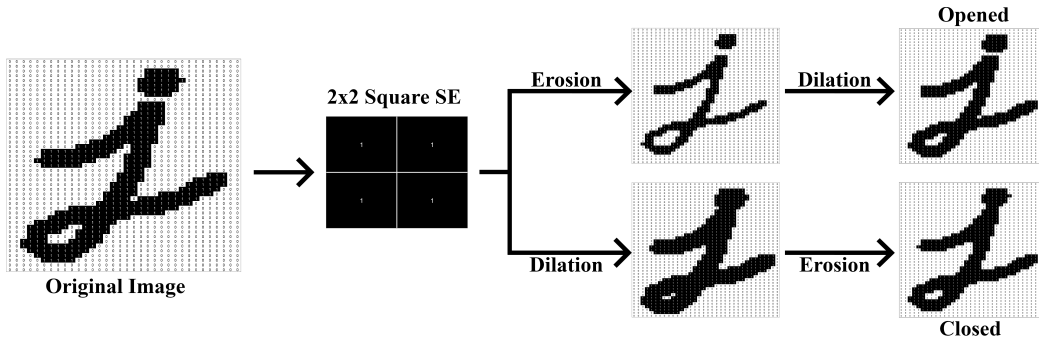


Figure 2.6 : Demonstration of the fundamental operators.

deforming the original shape, caused by the large size of the SE compared to the objects in the image. These effects of the operation and the elimination by the reconstruction can be shown in the Fig. 2.7.

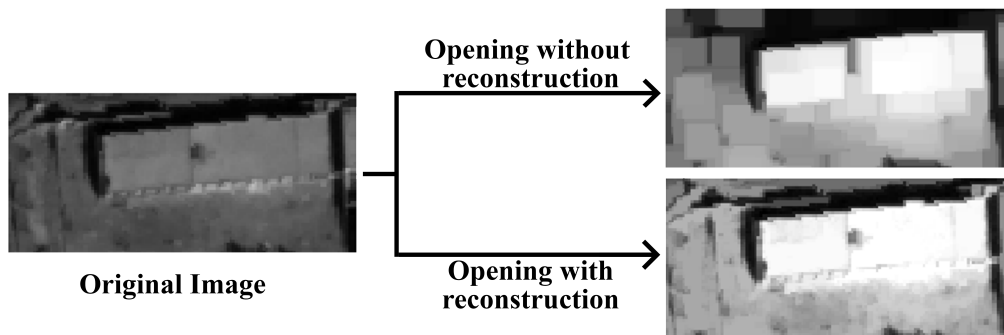


Figure 2.7 : The difference between operations with reconstruction and without reconstruction.

2.3.2 Morphological attribute profiles

Attribute Profiles (APs) has been suggested as using MPs' cumulative concept in conjunction with the attribute filters that can extract different types of the structural patterns from an image as described in the Eq. 2.12. To retain simplicity, this thesis explains fundamental principles of attribute filtering. Therefore, the reader is advised to read [53, 54] for a complete understanding of the details.

$$AP(f) = \{\phi^{AT_i}(f), \dots, \phi^{AT_1}(f), f, \gamma^{AT_1}(f), \dots, \gamma^{AT_i}(f)\} \quad (2.12)$$

Where $\phi^{AT_i}(f)$ and $\gamma^{AT_i}(f)$ refer to the images, generated by using the thickening and thinning operators, respectively. The operators compare the attributes of the objects in the image to the criterion T_i which corresponds to the type of the attribute to be considered. The different types of attributes, such as area, standard deviation, or moment of inertia, can be used for modelling the different geometrical or textural information of the objects [55].

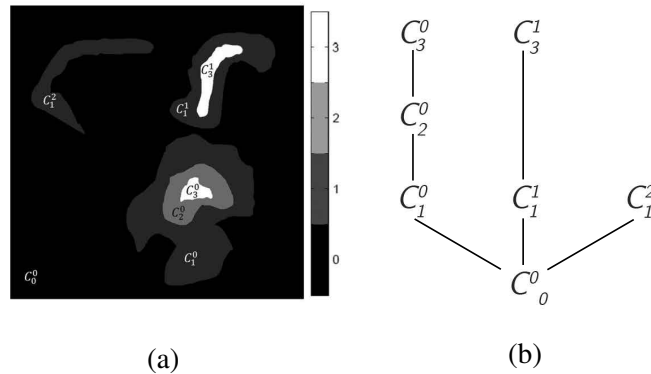


Figure 2.8 : Connected components in an image (a), and their tree representation (b).

Attribute filters probe connected components in an image. The connected components structure in an image illustrated in Figure 2.8. The operations look for a criterion whether is satisfied or not by the components. An attribute opening removes the component that is not satisfies an increasing criterion. If the criterion is not increasing, the operator is not an opening anymore. The operator based on a non-increasing criterion is called as attribute thinning. The increasingness of a criterion can be explained by a simple principle; when the all of supersets a connected component set can satisfy the criterion that is also satisfied by the main set, the criterion is defined

as increasing. Differently from opening and thinning, attribute closing and thickening operators compounds the different connected components in the image that satisfies the criterion.

For the images subjected to the considered, a joint use of the different attributes can offer valuable information [56]. To use the information from the different attributes together, a MAP can be created as in Eq. 2.13.

$$\text{MAP}(f) = \{\text{AP}(A_1), \dots, \text{AP}(A_n)\} \quad (2.13)$$

where A_n is the n -th AP added into $\text{MAP}(f)$. In this study, MAPs are actually created but the features in this MAP are not used since the study aims to achieve an improved performance with the less number of features. In other words, the APs is evaluated separately in this study.

3. EXPERIMENTS AND RESULTS

The aim of this study is to determine the most representative morphological features detecting earthquake damage. Only VHR post-event optical satellite images were considered and hundreds of different morphological features were created based on the original images. In this extended dataset a feature selection procedure was conducted. To make a robust subset selection, feature selection operation was repeated based on 100 different realizations of the training data and the final subset was statistically determined from these 100 different selection results. Figure 3.1 shows the flowchart of the proposed approach.

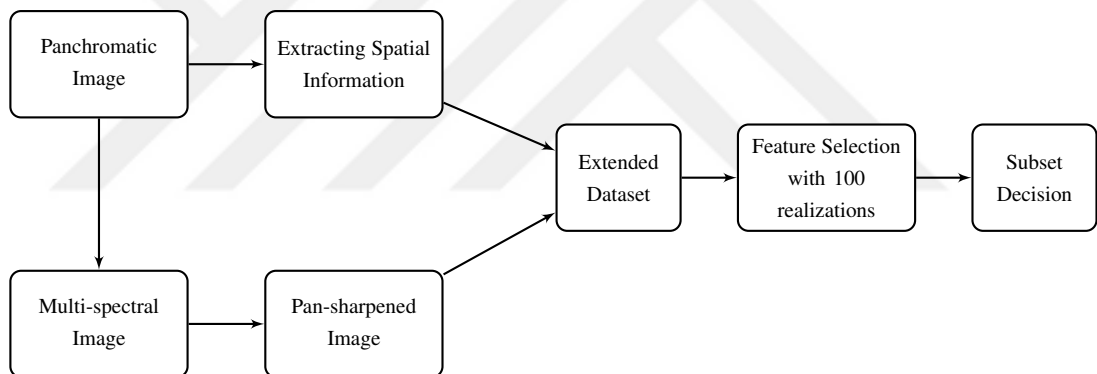


Figure 3.1 : Flowchart of the proposed approach.

After the subset selection procedure on the training area, the finalized features were also generated for all the entire image. Then the entire area was classified with the selected morphological features and spectral bands and a thematic map of the area was obtained. Based on the thematic map, a damage map was determined by a decision making approach that considers the proportion of pixels classified as damage.

The performances of the finally selected features were evaluated in terms of classification accuracy and the quality of thematic map. All the experiments were conducted on two different test images selected from two different earthquake datasets in order to evaluate the generalization capacity of the proposed approach.

Before using earthquake images, an evaluation of the proposed subset decision approach was completed on a synthetic dataset. The synthetic dataset contain three different images which were created to make them separable by the different attributes. This chapter firstly focuses on the evaluation of the proposed selection approach on the synthetic data. It continues by presenting the description of the datasets considered and the selecting training areas, classifications schemes on the training areas, the results of the feature selection procedures, and finally the transformation of the classification results into damage maps.

3.1 An Evaluation of The Proposed Approach

The proposed approach creates a feature subset that includes different morphological features which have different geometrical meanings. Those geometrical meanings can be evaluated visually. However, in a complex dataset, such as VHR post-earthquake images, the evaluation might be deceptive. For this reason, the proposed process should be verified through a properly created toy data.

To verify the proposed method, three toy datasets were created. As can be seen in Fig. 3.2, the toy data contain three different type of geometrical shapes. These shapes was manipulated to highlight particular geometrical attributes, such as area, moment of inertia, and standard deviation.

The different geometrical objects in each of the toy data were labeled as illustrated in Fig. 3.3. For each toy data, three extended attribute profiles that include area, moment of inertia, and standard deviation attribute profiles with the same criterion values were created. Then the selection procedure was implemented on the each extended toy dataset. As shown in Table 3.1, we expect to find the subsets that represent the

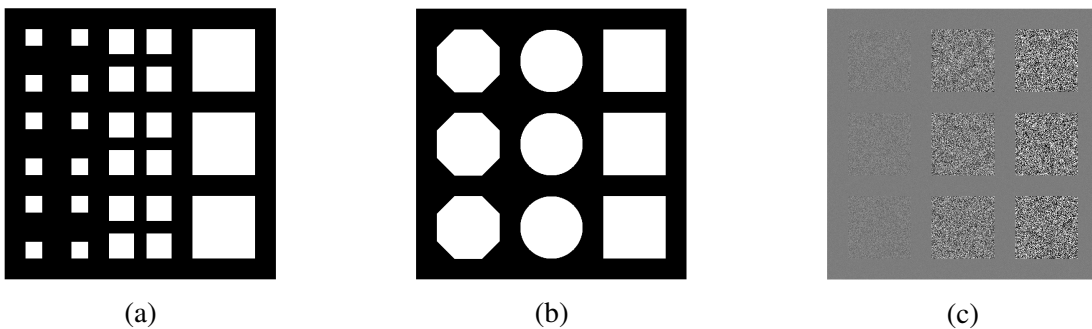


Figure 3.2 : The created toy data (a) highlights area, (b) highlights moment of inertia, (c) highlights standard deviation.

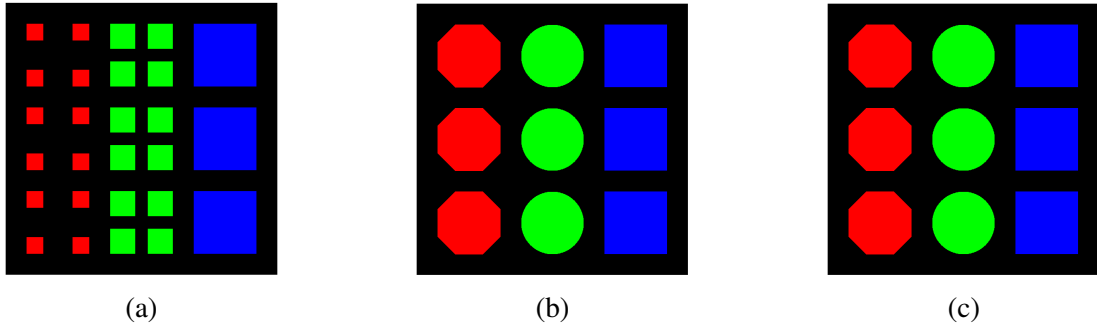
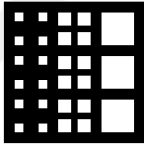
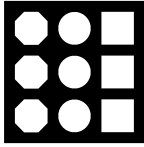
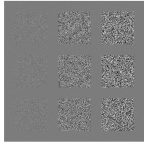


Figure 3.3 : Assigned labels for toy data (a) highlights area, (b) highlights moment of inertia,(c) highlights standard deviation.

geometric differences of each toy data. For instance, the area attribute profile should be clearly distinguishable in the top rankings of the selection results of the toy data which includes shapes with different areas, whereas the other attribute profiles should take place in the behind.

Table 3.1 : The created profiles and the **expected subsets** for each toy data.

Data	Created Profiles
	<ul style="list-style-type: none"> • Area: {100, 500, 2000, 2500, 4000, 6000, 25000} • Inertia: {0.05, 0.1, 0.1593, 0.1595, 0.2, 0.4, 0.8} • Std: {10, 50, 100, 150, 200, 250, 300}
	<ul style="list-style-type: none"> • Area: {100, 500, 2000, 2500, 4000, 6000, 25000} • Inertia: {0.05, 0.1, 0.1593, 0.1595, 0.2, 0.4, 0.8} • Std: {10, 50, 100, 150, 200, 250, 300}
	<ul style="list-style-type: none"> • Area: {100, 500, 2000, 2500, 4000, 6000, 25000} • Inertia: {0.05, 0.1, 0.1593, 0.1595, 0.2, 0.4, 0.8} • Std: {10, 50, 100, 150, 200, 250, 300}

To find a representative subset in the extended profiles, feature selection experiment was conducted with 100 different realizations with 40 pixels for each classes. In these experiments maximum relevance and minimum redundancy (mRMR) feature selection algorithm was used. The results of the realizations, which were visualized in the second column of Table 3.2, can give an insight about the feature relevancy. As can

be seen from the selection results, the expected profiles appear at the first places of the feature-axis, meaning that those are the most important ones among the rest profiles. However, subset from the ranking variations, including the most important features, should also be determined. To determine this subset, the number of top features in ranking should be decided. The final subsets were determined by considering feature frequencies in the decided number of top ranking. The standard deviation of all the feature frequencies in the decided ranking range was used as a decision boundary. The features having a frequency above the decision boundary were included in the final subset.

For the toy data, considering the first 5 features in the ranking as top features rankings is sufficient. The determined subsets and the relevant features in these subsets were shown in Table 3.3. The selected features satisfy the expectations for the proposed method.

Table 3.2 : The data and the corresponding selection results.

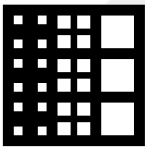
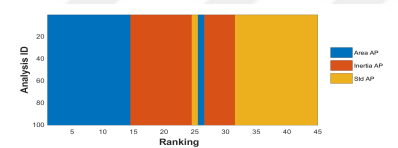
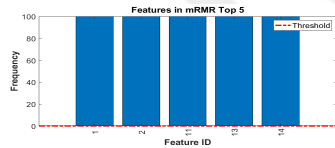
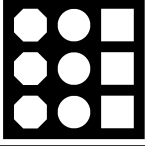
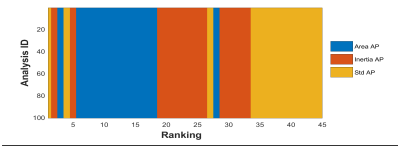
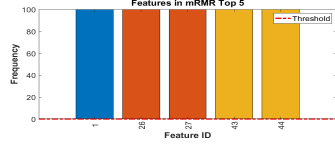
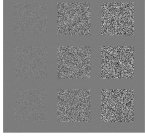
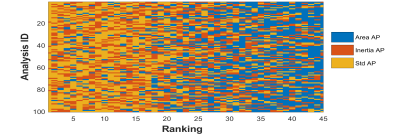
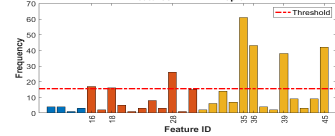
Data	Selection Result	Subset Decision
		
		
		

Table 3.3 : The toy data and the corresponding selected subsets.

Data	Selected Features					

3.2 Earthquake Damage Assessment: Description of Datasets

As the first dataset, the image of City of Bam (Iran) acquired by high-resolution commercial satellite QuickBird with a multi-spectral and a panchromatic image pairs was used in the experiments. The spatial resolutions of multi-spectral and the panchromatic images are 2.4 m. and 0.6 m., respectively. A region with the area of 45 hectares from the entire image was selected as a study area to find the best features by the proposed approach. A true color representation from the multi-spectral pair and the panchromatic image of the study area in Bam can be seen in Figure 3.4.

The second dataset is the image acquired after the Haiti Earthquake of 2011 by high-resolution commercial satellite WorldView-1. The original dataset contains a pan-sharpened multi-spectral and a panchromatic images both with the 0.5 m. spatial resolution. The study region in this dataset was selected with the area of 40 hectares to test the proposed approach and to train the classification model. In Figure 3.5, the true color representation and the panchromatic images of the Haiti study area can be seen.

In the analysis of such datasets, to use both of the spatial and spectral content of the original data, a pan-sharpened image (PSI) is created [57]. Pan-sharpening, also

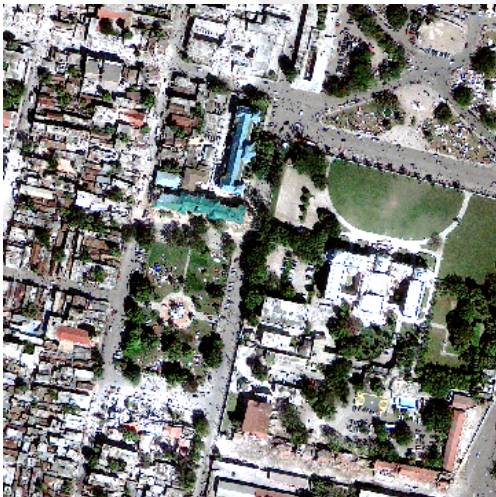


(a) A true color representation of the Bam training area.



(b) Panchromatic image of the Bam training area.

Figure 3.4 : True color and panchromatic images of the Bam training area.



(a) A true color representation of the Haiti study area.



(b) Panchromatic image of the Haiti study area.

Figure 3.5 : True color and panchromatic images of the Haiti training area.

called panchromatic sharpening, is using a panchromatic image to enhance the spatial resolution of multi-spectral images.

3.3 Morphological Feature Extractions

The panchromatic images were used to extract the spatial information due to its high spatial resolution. In the implementation of the APs for damage assessment, four attributes, including the area (AP_a), moment of inertia (AP_i), standard deviation (AP_s), and length of the diagonal (AP_d) were considered due to the fact that they were reported as the most representative attributes for detecting the similar problems [30].

On the side of the MPs, eight different profiles with four shape of the SE, including disk, square, diamond, and line were created with and without geodesic reconstruction. Some of the images obtained by the defined profiles for each attribute and morphological operation are shown in Fig. 3.6 and 3.7. The criterion value used in the attribute filters, and the size of the SEs used in the morphological profiles were defined in an increasing order as shown in the Table 3.4 and 3.5. Finally, the Bam and Haiti datasets was reached 557 and 556 dimensions, respectively, with the extracted features and the original panchromatic along with the pan-sharpened image of the selected study areas.

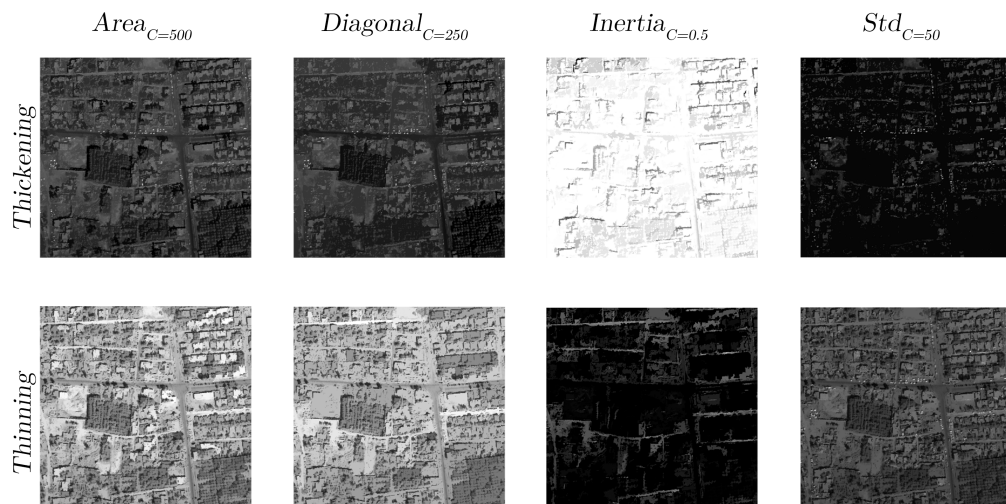


Figure 3.6 : Representative images from each attribute profile set of the Bam study area.

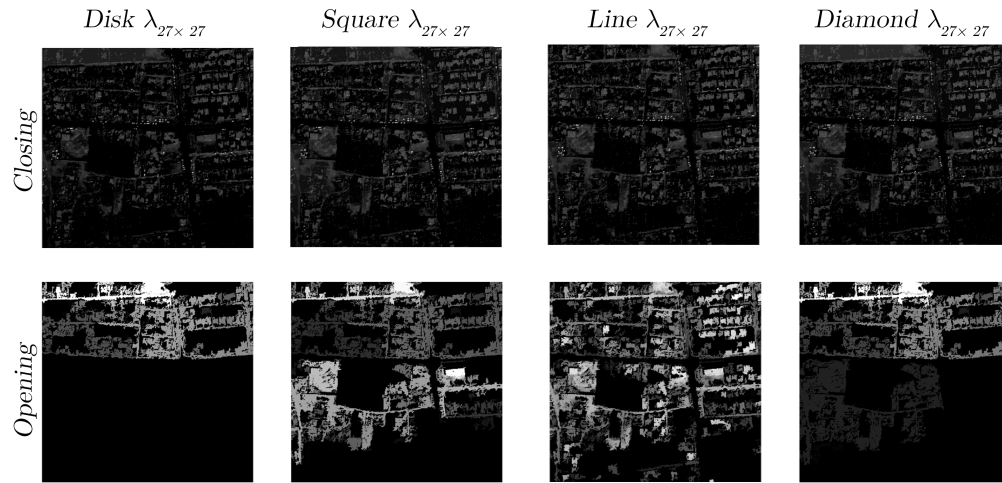


Figure 3.7 : Representative images from morphological profile sets of the Bam study area created by opening and closing operators without reconstruction.

Table 3.4 : Features in each attribute profile sets and their specifications.

Feature Set	Assigned Colors	Description	Feature IDs	Method	Attribute	Criteria (C)
1		AP_a	{1, 2, ..., 20}	Thickening	Area	{1000, 950, ..., 50}
1		AP_a	{21, 22, ..., 40}	Thinning	Area	{50, 100, ..., 1000}
2		AP_d	{41, 42, ..., 59}	Thickening	Length of diagonal	{500, 475, ..., 50}
2		AP_d	{60, 61, ..., 78}	Thinning	Length of diagonal	{50, 75, ..., 500}
3		AP_i	{79, 80, ..., 87}	Thickening	Moment of inertia	{0.9, 0.8, ..., 0.1}
3		AP_i	{88, 89, ..., 96}	Thinning	Moment of inertia	{0.1, 0.2, ..., 0.9}
4		AP_s	{97, 98, ..., 116}	Thickening	Standard deviation	{100, 95, ..., 5}
4		AP_s	{117, 118, ..., 136}	Thinning	Standard deviation	{5, 10, ..., 100}

Table 3.5 : Features in each morphological profile sets and their specifications.

Feature Set	Assigned Colors	Description	Feature IDs	Method	Shape of SE	Size of SE
5		MP_{D1}	{137, 138, ..., 162}	Closing with reconstruction	Disk	{51x51, 49x49, ..., 1x1}
5		MP_{D1}	{163, 164, ..., 188}	Opening with reconstruction	Disk	{1x1, 3x3, ..., 51x51}
6		MP_{D0}	{189, 190, ..., 214}	Closing	Disk	{51x51, 49x49, ..., 1x1}
6		MP_{D0}	{215, 216, ..., 240}	Opening	Disk	{1x1, 3x3, ..., 51x51}
7		MP_{S1}	{241, 242, ..., 266}	Closing with reconstruction	Square	{51x51, 49x49, ..., 1x1}
7		MP_{S1}	{267, 268, ..., 292}	Opening with reconstruction	Square	{1x1, 3x3, ..., 51x51}
8		MP_{S0}	{293, 294, ..., 318}	Closing	Square	{51x51, 49x49, ..., 1x1}
8		MP_{S0}	{319, 320, ..., 344}	Opening	Square	{1x1, 3x3, ..., 51x51}
9		MP_{L1}	{345, 346, ..., 370}	Closing with reconstruction	Line	{51x51, 49x49, ..., 1x1}
9		MP_{L1}	{371, 372, ..., 396}	Opening with reconstruction	Line	{1x1, 3x3, ..., 51x51}
10		MP_{L0}	{397, 398, ..., 422}	Closing	Line	{51x51, 49x49, ..., 1x1}
10		MP_{L0}	{423, 424, ..., 448}	Opening	Line	{1x1, 3x3, ..., 51x51}
11		MP_{DA1}	{449, 450, ..., 474}	Closing with reconstruction	Diamond	{51x51, 49x49, ..., 1x1}
11		MP_{DA1}	{475, 476, ..., 500}	Opening with reconstruction	Diamond	{1x1, 3x3, ..., 51x51}
12		MP_{DA0}	{501, 502, ..., 526}	Closing	Diamond	{51x51, 49x49, ..., 1x1}
12		MP_{DA0}	{527, 528, ..., 552}	Opening	Diamond	{1x1, 3x3, ..., 51x51}

3.4 Classification Schemes

For both of the considered datasets, the number of six land-cover classes, including damage, building, road, vegetation, shadow, and open-ground, were determined for the classification, and the ground truth data were visually created by comparing the pre- and post-event images of the selected study areas. The ground truth data for both of the study areas can be seen in Figure 3.8.

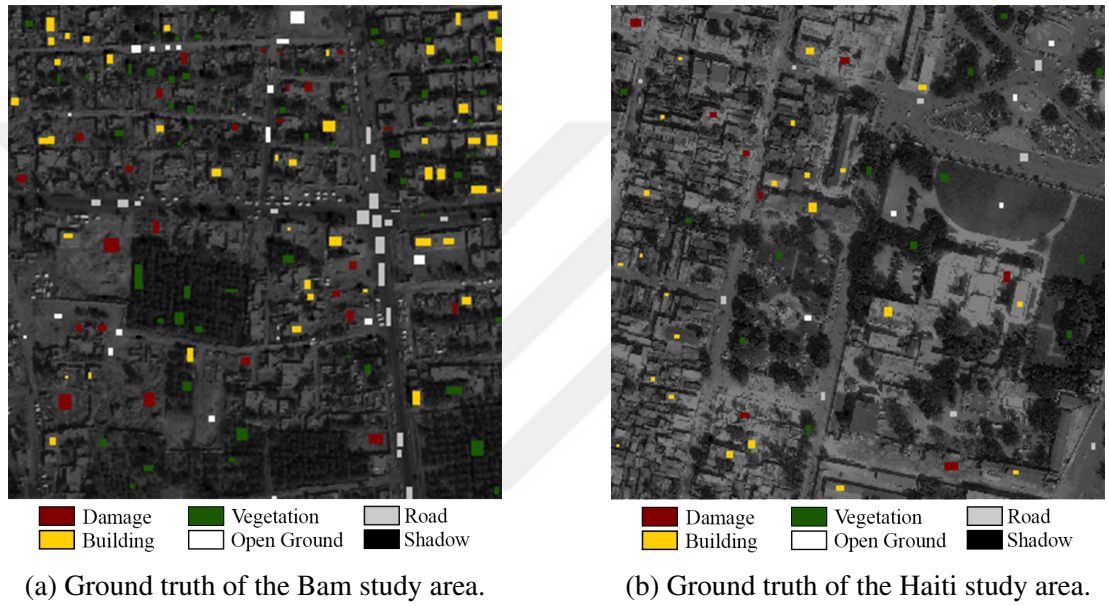


Figure 3.8 : Ground truths created by a visual comparison of pre- and post-event images.

Table 3.6 : The number of training and test samples for each class used in the experiments on the Bam study area.

Class	Damage	Building	Shadow	Vegetation	Road	Open Ground	Total
Training	150	150	150	150	150	150	900
Test	2539	5849	1017	3446	2559	1028	16438
Total	2689	5999	1167	3596	2709	1178	17338

Table 3.7 : The number of training and test samples for each class used in the experiments on the Haiti study area.

Class	Damage	Building	Shadow	Vegetation	Road	Open Ground	Total
Training	150	150	150	150	150	150	900
Test	919	1297	227	1088	719	292	4542
Total	1069	1447	357	1238	869	442	5442

The author's previous work [58] based on those ground truth data showed that, even if a high classification accuracy is achieved the quality of thematic map might be very

poor. Thus, the ground truth data need to be evaluated and analyzed to measure the variations in the original multi-spectral image of the all labeled pixels for all classes. The variations of each classes in original bands were illustrated in Figure 3.9 and Figure 3.10. As can be seen in the variation plots, some labeled pixels causes high variance in the ground truth. To eliminate these effects, the samples that are not coherent with the normal distribution were removed from the ground truth.

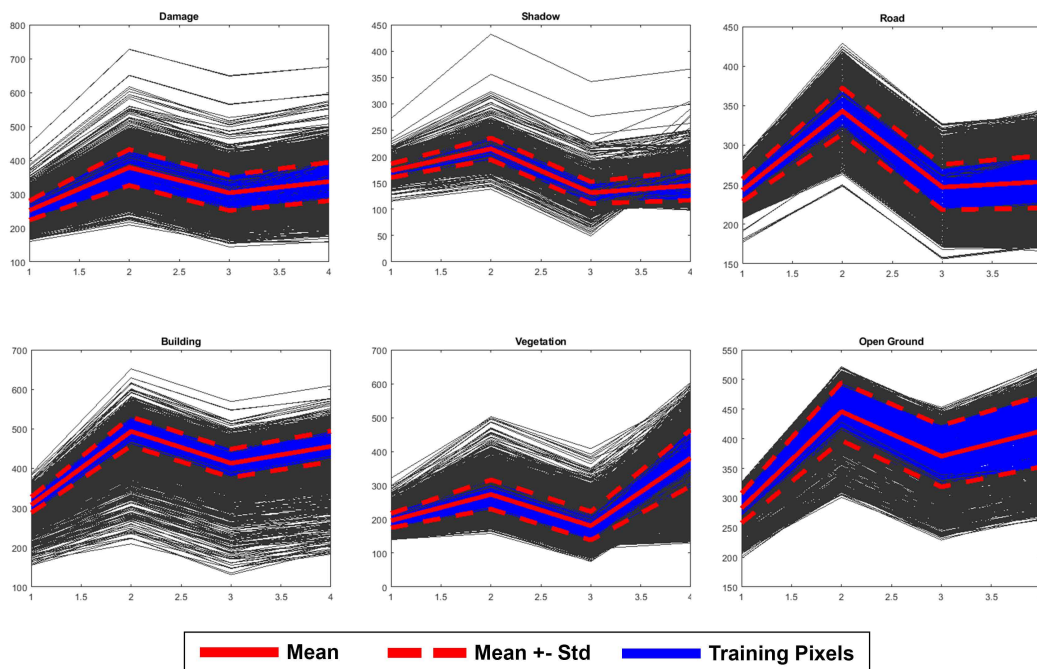


Figure 3.9 : Evaluation of the labeled pixels in the Bam study area.

To perform a reliable comparison, all the classification experiments were repeated ten times based on randomly selected training and test samples from the aforementioned ground truth data, that the number of samples for each class was given in Table 3.6 and Table 3.7 for Bam and Haiti datasets, respectively.

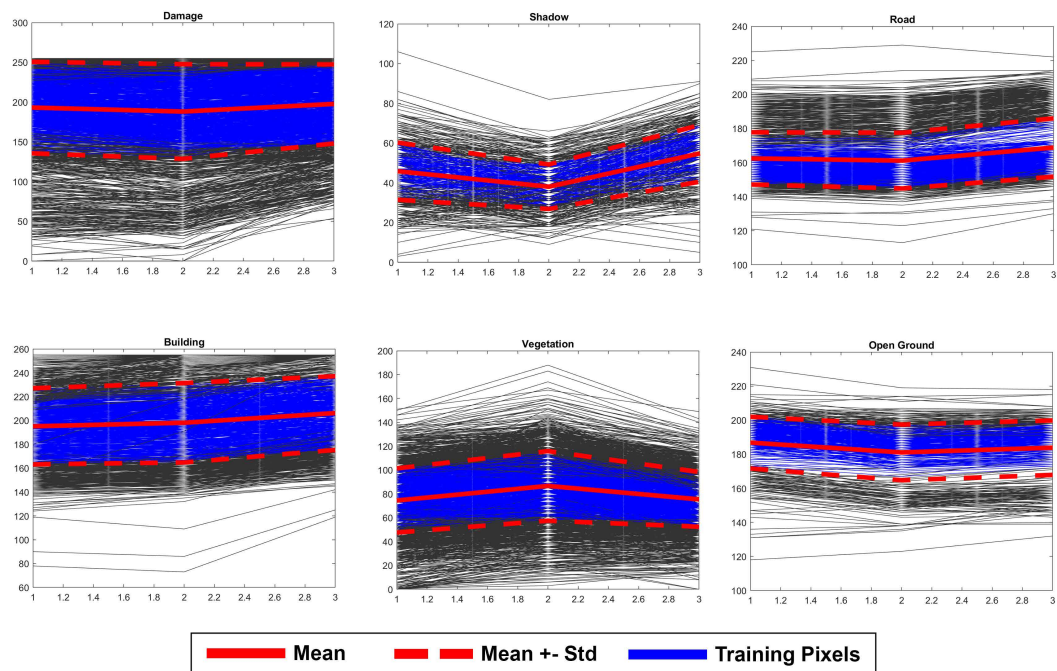


Figure 3.10 : Evaluation of the labeled pixels in the Haiti study area.

3.5 Feature Selection and Subset Decision

To find the most representative features in the each extended dataset, feature selection experiments were conducted with 100 different realizations of training samples. The outputs of selection process for the top 40 features were visualized as the heat maps as in Fig. 3.11. The colors in the heat maps represent the features given in Table 3.5 and 3.4.

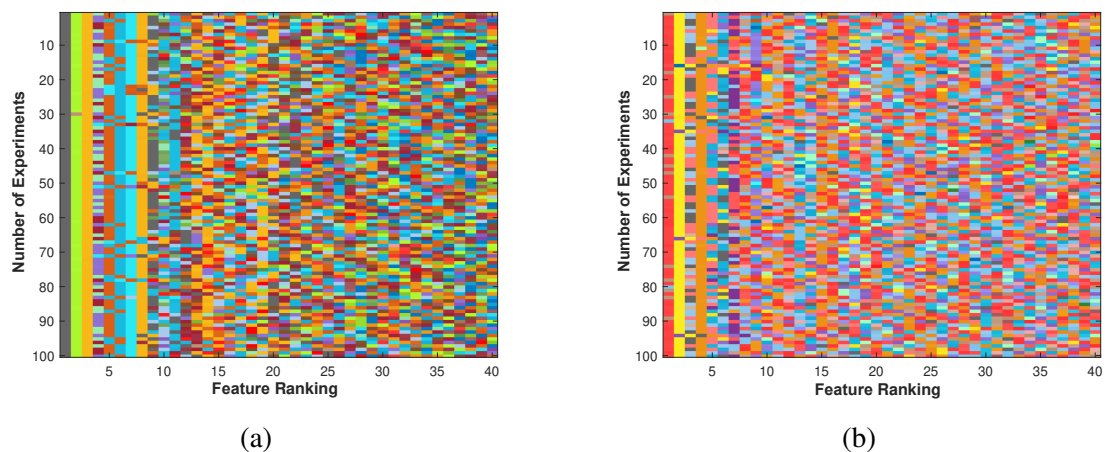


Figure 3.11 : Top 40 features ranked with respect to the feature selection for (a) Bam study area and (b) Haiti study area.

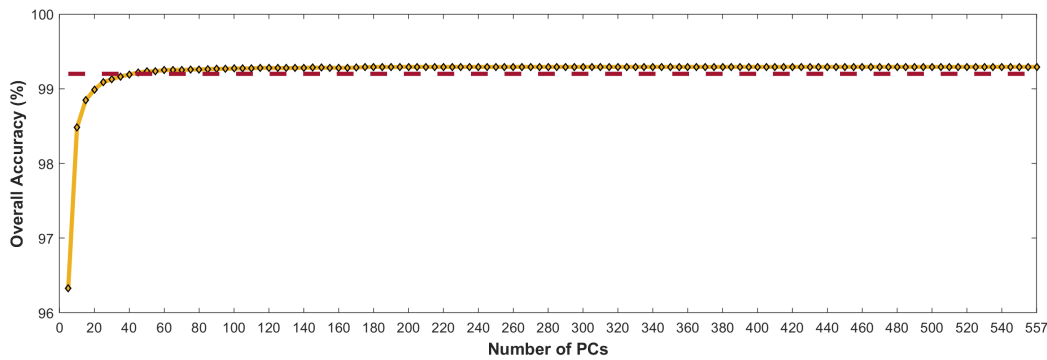


Figure 3.12 : PCA based learning curve.

The different realizations of the feature selection operations brings different rankings as those use different pixels for training, meaning that the training data affect performance of the feature selection algorithm. Despite of this, a feature pattern might be more dominant among the others even when using different training set, if this feature is really effective in damage assessment. Therefore, the more dominant features out of hundred experiments should be determined. To determine this subset, the one should decide how many of those features are the most important ones. For this aim, a learning curve based on principal component analysis was created to find the number of features which affects the classification performance the most. The curve created by classifying the data using different size of the first principal components in an increasing order with the step size of five, and shows the corresponding overall accuracies from ten different realizations of the training data. According to the result of this analysis, illustrated in Fig. 3.12, the top 10, 20, 40 feature rankings were decided as the ranges for selecting the optimal features as they have a significant effect on classification accuracy.

The final subsets from a number of different feature selection experiments can be determined by three simple approaches, such as keeping an absolute number of the most frequent features in the experiments, or keeping a percentage of the most frequent features, or keeping features that exceeds a pre-defined importance threshold. Keeping an absolute number of features or percentage can pass over features that have too close importance. Therefore, we used a threshold to determine the final feature subsets. The threshold should not have a large bias, hence the standard deviation of feature frequencies in related part of the selection result was used as the threshold and the features above the threshold are selected as the most important features. This process

is illustrated with frequency plots as shown in Fig. 3.13 and 3.14. It should be noted that, the features which are not ranked in the top 10, 20 and 40 are not included into the related frequency plots.

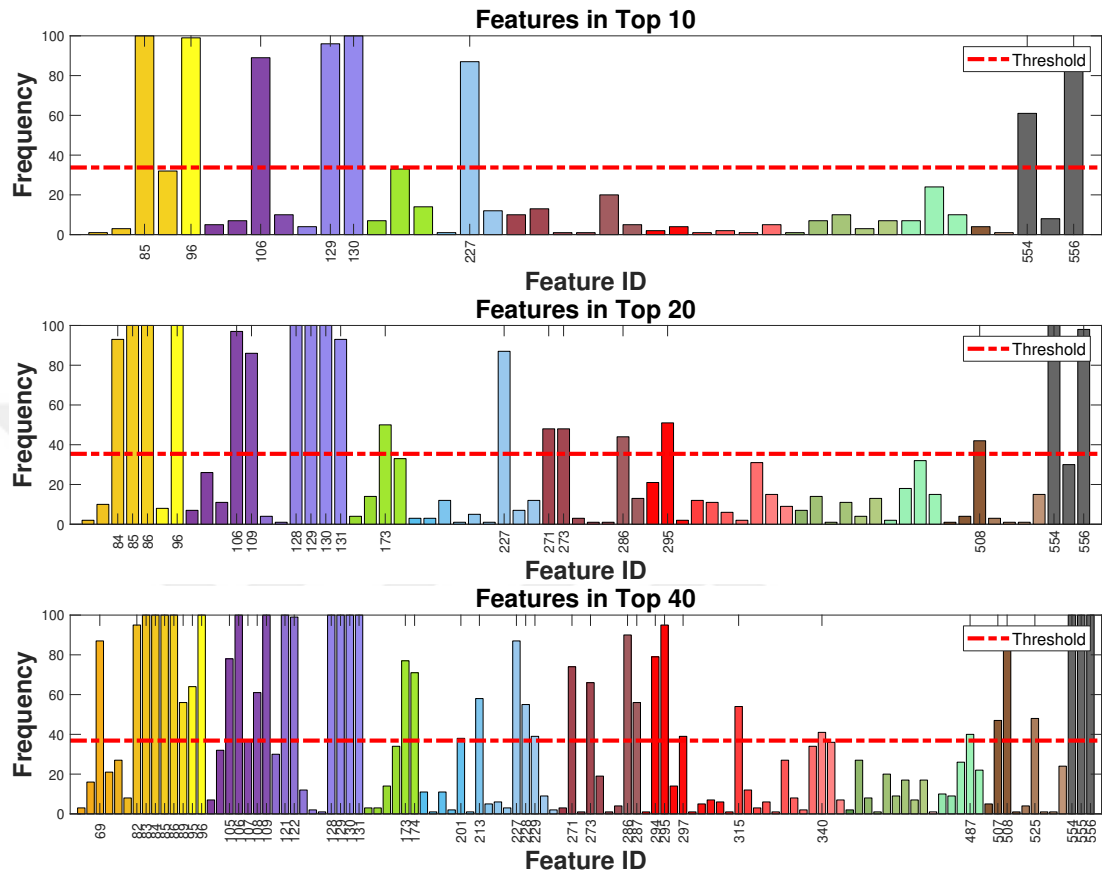


Figure 3.13 : Feature frequencies in the mRMR rankings of the Bam extended dataset and their corresponding thresholds.

It can be clearly seen in both set of frequency plots, selecting the most frequent features in the first 10, 20 and 40 rankings directly could cause to miss important features or to select irrelevant features. For instance, in Figure 3.13 eight features are selected within the features ranked in the first top ten features for the Bam dataset, but if the most frequent ten features were directly selected two irrelevant features would be selected.

Regarding the results of the Bam dataset, shown in Fig. 3.13, a number of features from the moment of inertia and the standard deviation APs and red, blue, and green spectral bands come to the fore in all the top rankings. The features from MPs created using disk, square, and diamond shaped SEs are also observed several times in the top 20 and 40 rankings. On the contrary, the area and the length of diagonal APs, and MPs created using line shaped SEs are nearly not observed in any top ranking.

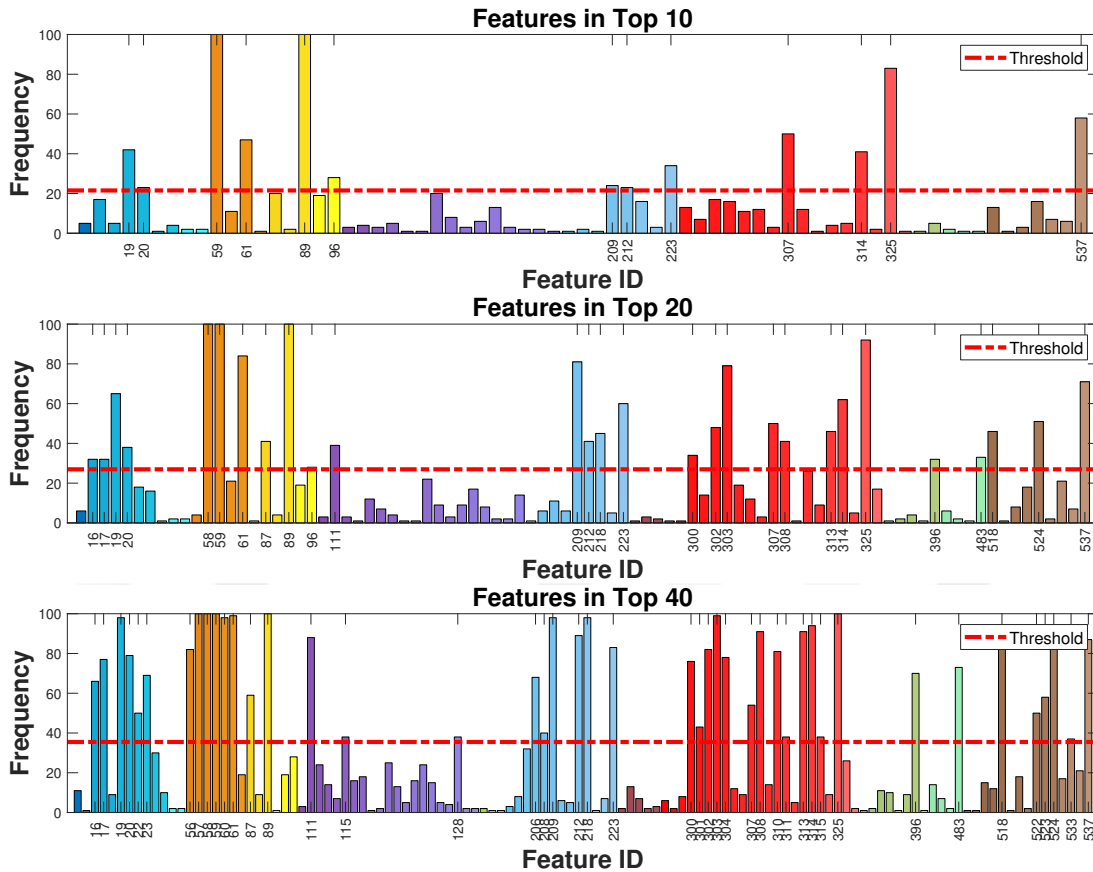


Figure 3.14 : Feature frequencies in the mRMR rankings of the extended dataset from Haiti study area and their corresponding thresholds.

The results from the feature analysis of the Haiti extended dataset were shown in Figure 3.14. In all the selected top rankings of this dataset, a variety of different feature types distinguish themselves including the area and the length of diagonal APs and the MPs based on the disk, square, and diamond shaped SEs.

The selection results from the both of the extended datasets should be also analyzed together. Each dataset has its own challenges and the selected features are expected to overcome these challenges. On the one hand, the Haiti dataset includes buildings with sharp edges and the damaged buildings were extracted with the use of area or length of the diagonal attributes. On the other hand, the buildings in the Bam dataset can not be easily separated by the edges as they have similar textures and colours with the open ground class, hence in the Haiti extended dataset the features from standart deviation and moment of inertia APs led to the most significant improvements on the assessment. With regard to the MPs, it can be possible to select many different features in such wide implementation since they directly related to the shape of SE, not the texture. Consequently, the selected features from all the top rankings of Bam and

Haiti extended datasets are visualized in Figure 3.15 showing that a number of the MPs obtained by using the SEs with disk, square, and diamond shapes, and a few features from the inertia and the standard deviation APs are commonly observed in the both analysis. However, there is not any attribute type nor a specific SE type that can be prominently noticed in both of the results.

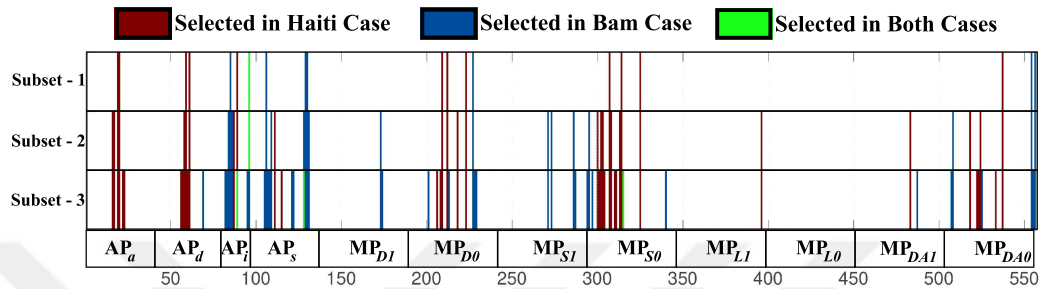


Figure 3.15 : The features selected from all the considered top rankings in each cases, individually and commonly.

3.6 Results

Based on the aforementioned threshold approach, three feature subsets each of which consists of features selected in top 10, 20, and 40 rankings were determined for each study area extracted from both Bam and Haiti dataset, and these areas were classified with k-NN and SVM classifiers based on the feature subsets. The results for each dataset were discussed separately in the following subsection. The performance of the proposed approach was evaluated in terms of both classification accuracy and the quality of thematic maps. Therefore, the confusion matrices were also given to provide per class accuracy. All the confusion matrices were obtained by averaging 10 trials of different random realizations from the training data.

3.6.1 Results: Bam dataset

In the study area of Bam dataset, the hardest classes to separate are the open ground, the building, and the damage classes since they have similar textures and spectral signatures. For this reason, it is expected to observe a significant improvement on the classification performance especially for those classes.

The classification accuracies obtained using k-NN and SVM classifiers were given by the confusion matrices as shown in Figure 3.17 and 3.19 respectively. The confusion matrices indicate that the use of new feature subsets in the classification does not only improved the overall and average accuracy, but also improve the accuracy for the classes open ground, building, and damage classes compared to the case obtained by using only the pixel densities of the original data. The thematic maps, given in 3.16 and 3.18, also showed the improvement for those classes by providing less noisy representation of those classes. Both the accuracies and quality of the thematic maps in terms of extracting the most difficult classes are found to be consistent for both SVM and k-NN. According to the classification accuracies obtained by the both classifiers, the most significant improvement is resulted from the use of selected feature subsets in top 10 features. However, the feature subsets from the top 20 and 40 features have also an apparent positive effect on the classification accuracy and the quality of the thematic map. As a result, the subset from the top 40 features achieves the highest classification accuracy, while the subset from the top 10 features makes the most significant improvement of the classification accuracy.

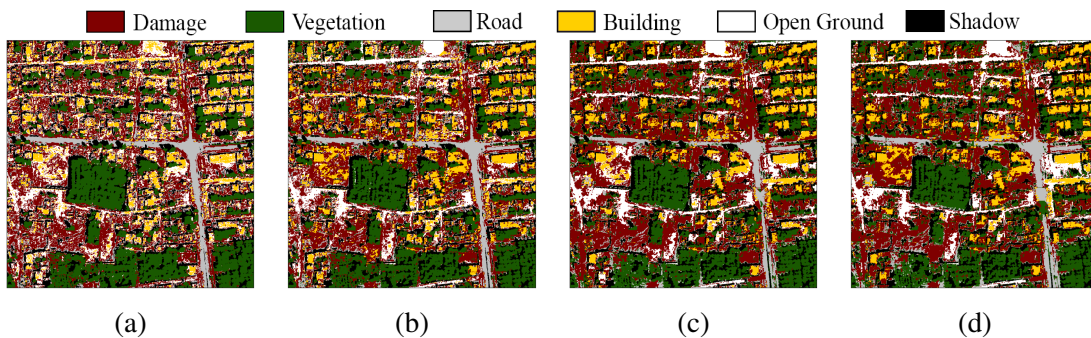


Figure 3.16 : Thematic maps of the Bam study area obtained with k-NN classifier using (a) pixel densities, (b) Subset-1, (c) Subset-2, (d) Subset-3.

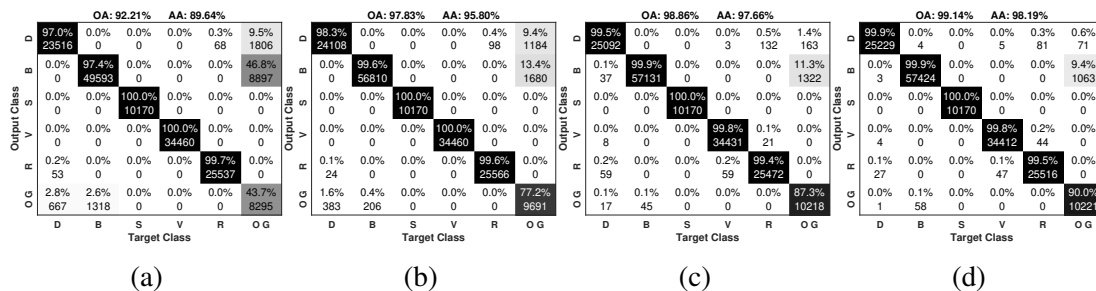


Figure 3.17 : Confusion matrices of the Bam study area obtained with kNN classifier using (a) pixel densities, (b) Subset-1, (c) Subset-2, (d) Subset-3.

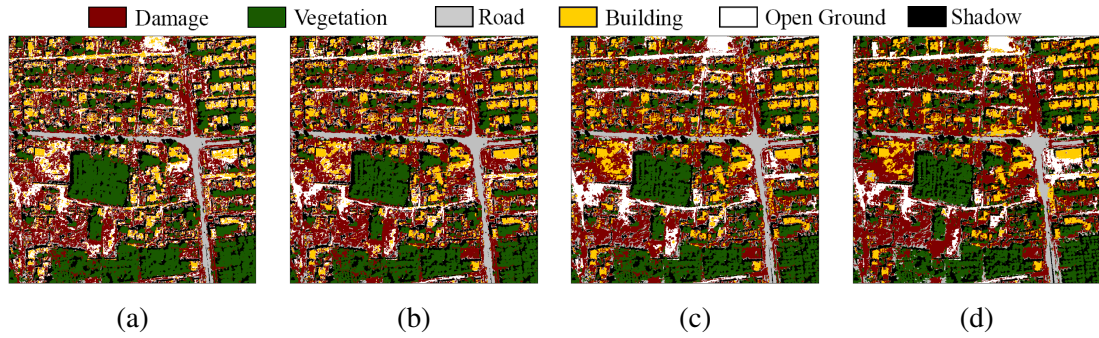


Figure 3.18 : Thematic maps of the Bam study area obtained with SVM classifier using (a) pixel densities,(b) Subset-1, (c) Subset-2, (d) Subset-3.

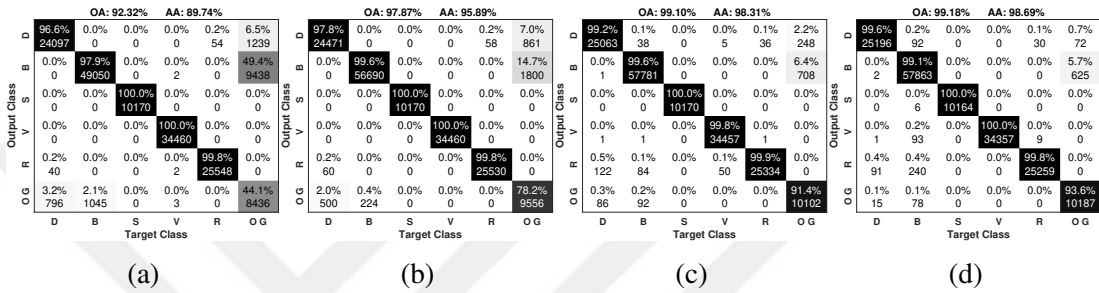


Figure 3.19 : Confusion matrices of the Bam study area obtained with SVM classifier using (a) pixel densities, (b) Subset-1, (c) Subset-2, (d) Subset-3.

3.6.2 Results: Haiti dataset

The challenge in the Haiti study area is to separate the building and the damage classes as in the case of the Bam dataset. The study area includes the buildings with many different types of shapes, roofs, colours and sizes. Therefore, the extracted feature subsets are expected to help for handling this difficulty.

The confusion matrices shown in Figure 3.21 and 3.23 clearly express the effect of the extracted subsets in the classification accuracy compared to the case obtained with only using pixel densities of the original feature set. The accuracy for the most confused classes were significantly improved with the use of the new features. As be seen in Figure 3.20 and 3.22, the quality of the thematic maps in terms of classification accuracy were also improved by using additional features, which are consistent with the confusion matrices as well. In the thematic maps obtained by using only the original features, the damaged areas can not be distinguished well and some of the damage patterns were classified as undamaged buildings, while the thematic maps obtained by using the extracted features provides better representations for the damage class.

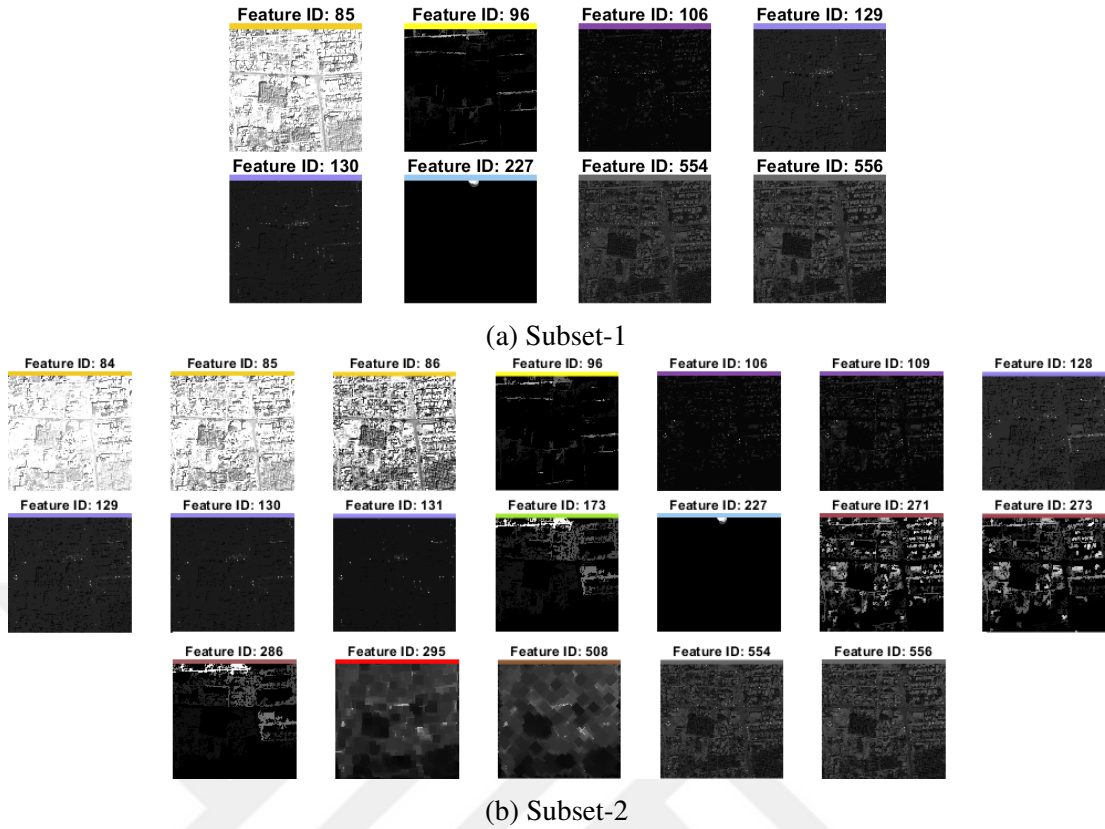
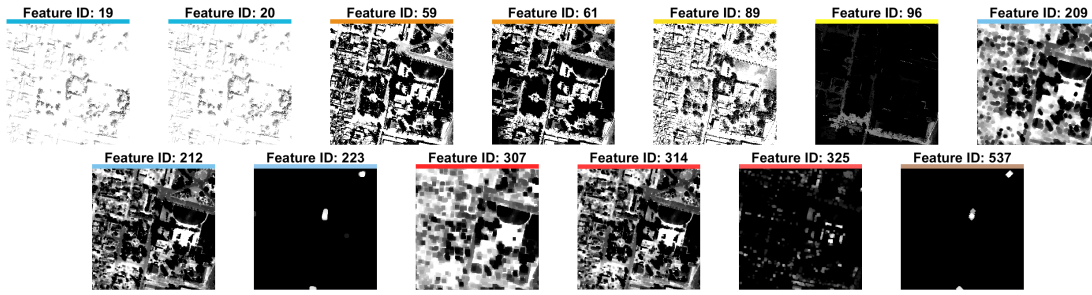
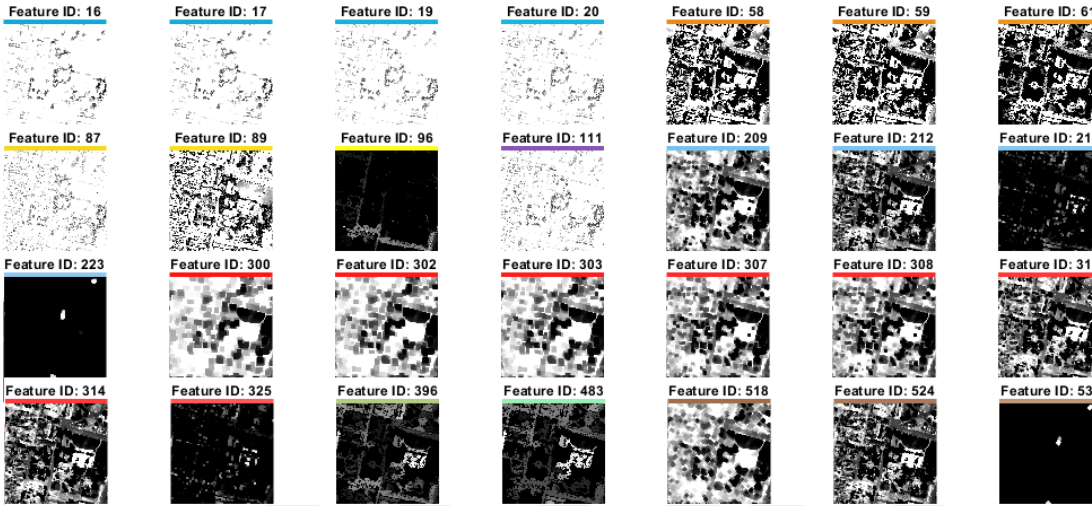


Figure 3.24 : Selected feature sets for the Bam study area.

To sum up, both of the case studies conducted with Bam and Haiti datasets show the efficiency of the selected feature subsets in classification accuracy for both of the classification methods compared to the pixel-based classification results. All the feature subsets from each number of the top features improve the classification accuracy. However, the number of features to be generated is also important in terms of the computational time for an analysis on the entire images. Therefore, only the features selected from the top ten and the top twenty rankings were used in the analysis of the entire areas in the considered images, since they have the most significant effect on the results. The selected feature subsets for Bam and Haiti datasets are shown in Figure 3.24 and 3.25, respectively.



(a) Subset-1



(b) Subset-2

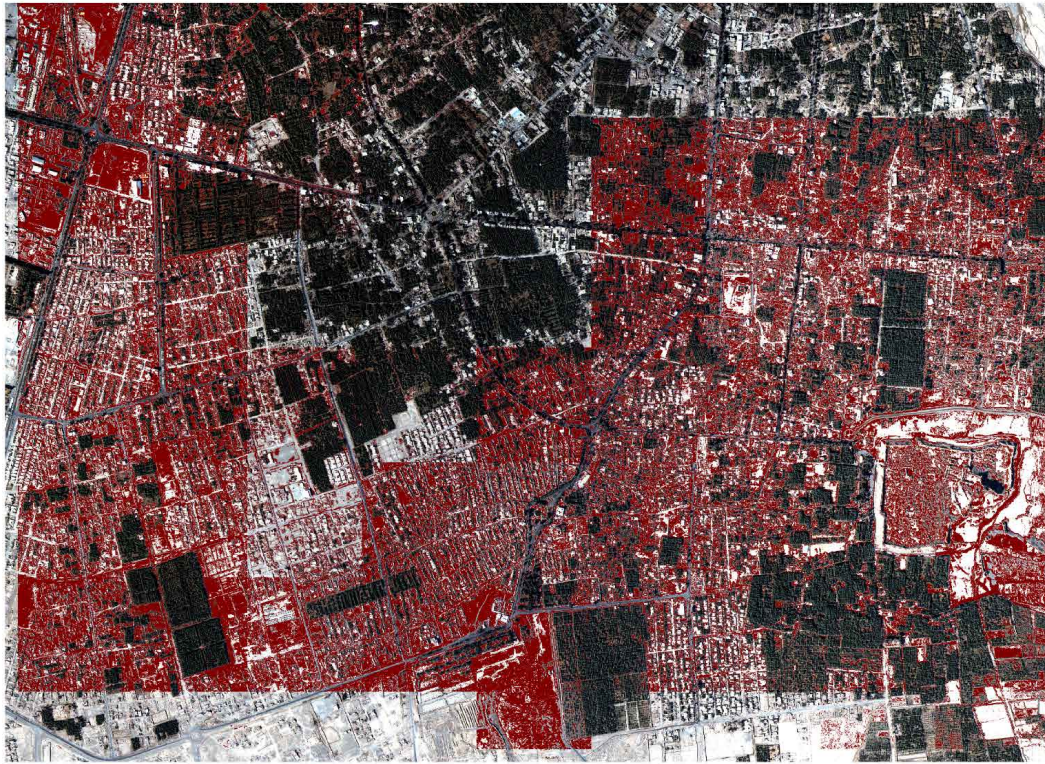
Figure 3.25 : Selected feature sets for the Haiti study area.

3.7 Damage Maps

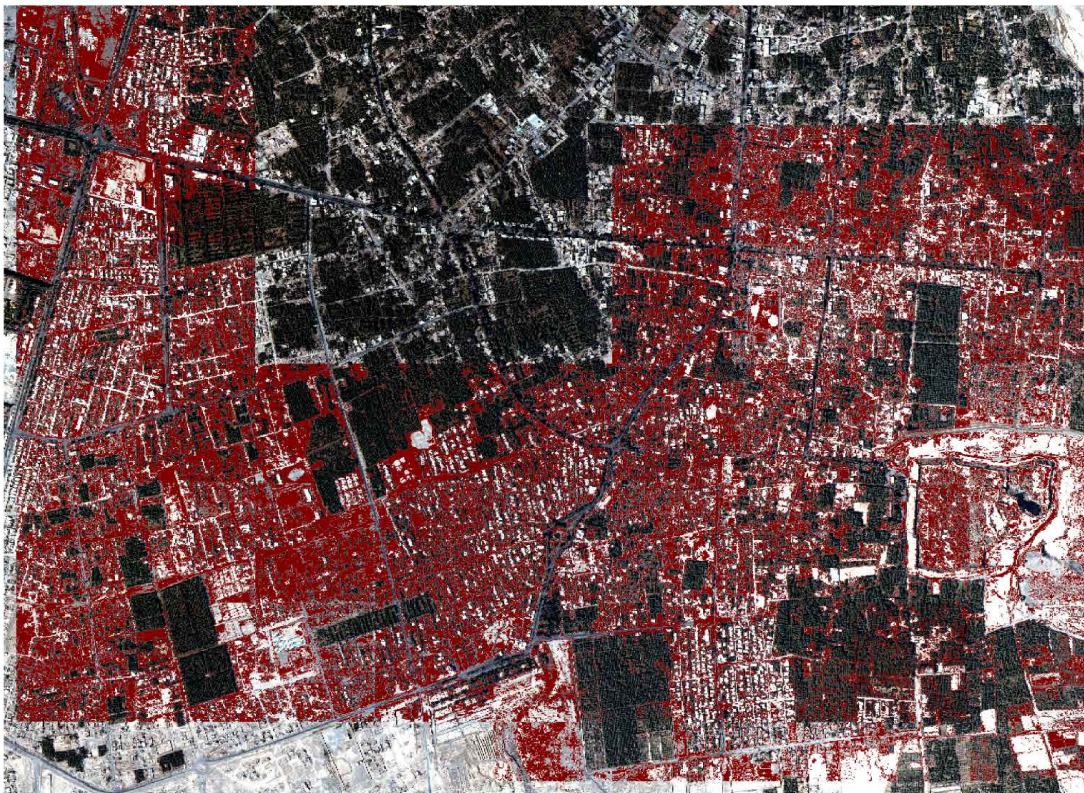
In this study, an automatic earthquake-induced damage detection approach in the way that is as fast and as accurate as possible was proposed. Therefore, any ancillary data or pre-earthquake data were not used. The use of such data needs time consuming pre-processing steps such as image registration and atmospheric correction. Due to this, creating a damage map without using any pre-earthquake information is another challenge for the proposed earthquake-induced damage detection approach. In order to build the damage map, the entire area in the post-earthquake image is classified and a thematic map is obtained to represent all the land cover classes. However, to classify the entire area at only one execution is not a practical approach due to its need for high memory. The entire area is generally classified in small pieces and the general thematic map is obtained with the synthesis of those small pieces. In this study, both of the Bam and Haiti datasets were subdivided into small parts consisting of 900×900

pixels, and then they were classified using the training data associated with the study areas by using selected features from the top ten rankings of FS analysis.

The thematic maps include all the land cover classes by definition. But the damage maps should only provide the level of the earthquake-induced damage. Therefore, the thematic maps are not directly used to interpret the damage distributions. In this study, the true color representation of the images are masked red colored pixels that represent the damage labels from corresponding classification results. The classes different from damage are eliminated in the damage maps. Finally, the obtained damage maps of entire areas for each dataset, can be seen in Figure 3.26 and 3.27. The damage maps generated by the proposed approach are evaluated in comparison to the reference maps obtained by European Space Agency (ESA), which are shown in Figure 3.28 and 3.29. The comparison shows that the generated damage maps are consistent with the reference works in broad strokes.



(a)



(b)

Figure 3.26 : Damage maps of the entire Bam image obtained using (a) Subset-1, (b) Subset-2.



(a)



(b)

Figure 3.27 : Damage maps of the entire Haiti image obtained using (a) Subset-1, (b) Subset-2.

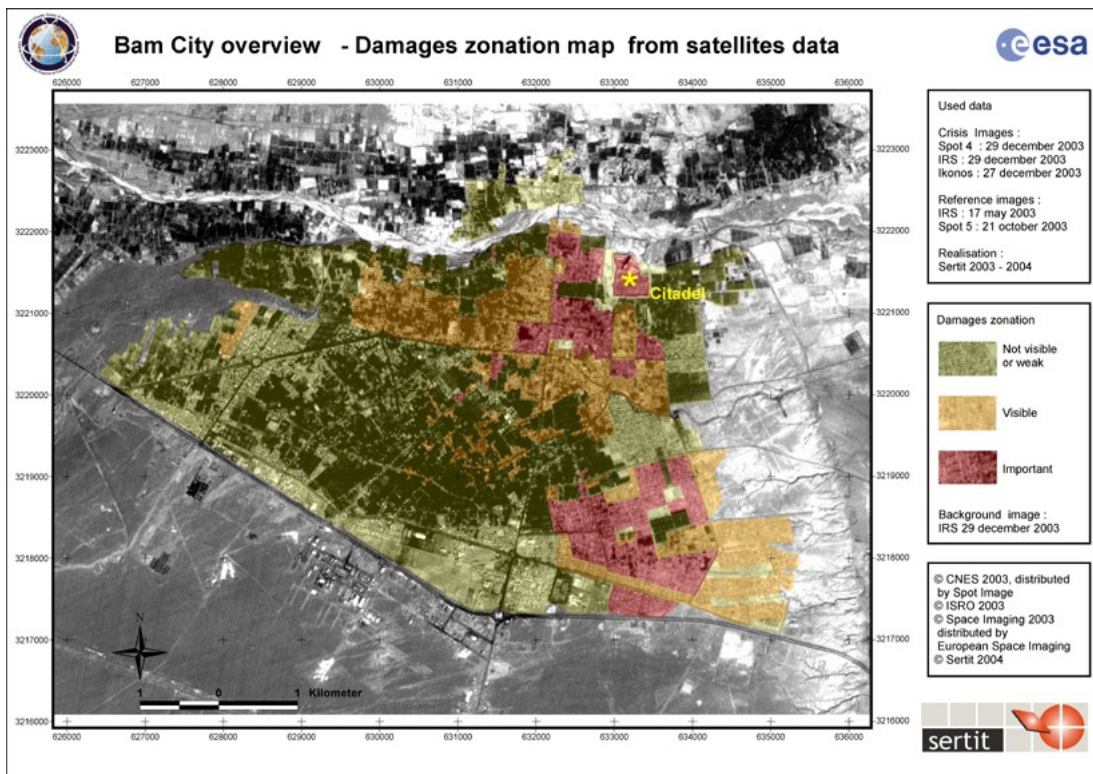


Figure 3.28 : The reference damage map of the Bam dataset.

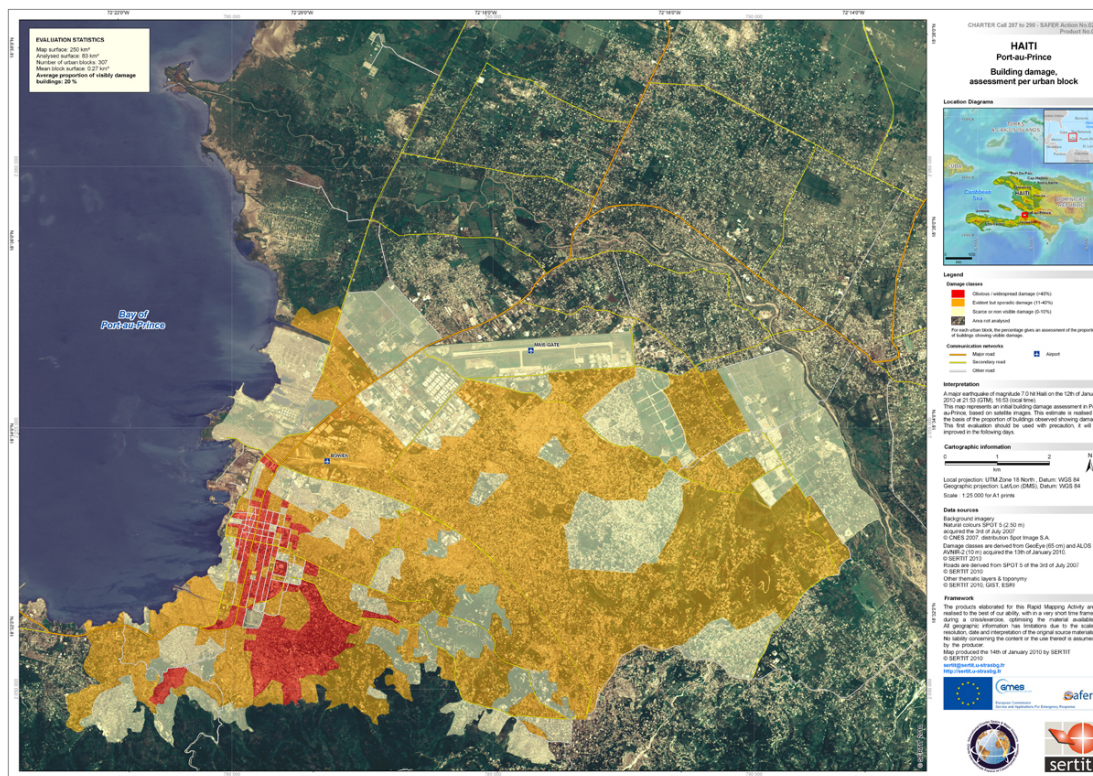


Figure 3.29 : The reference damage map of the Haiti dataset.

4. CONCLUSIONS AND DISCUSSIONS

A fast and accurate earthquake damage identification from VHR post-earthquake image is very challenging task, requiring additional contextual information. This study aims to implement the recently proposed spatial feature extraction methods, that are MPs and APs, for the earthquake damage assessment. There can be a great variety of extracted features by the use of different parameters for these methods. When the data size is considered, to generate that variety of features is not a practical approach as such process would be expensive in terms of both the computational time and the memory usage. Furthermore, it can be possible to generate irrelevant features in such variety. Therefore, in this study, wide feature sets were created by using MPs and APs and a feature selection procedure using mRMR was implemented to find the most relevant features in those feature sets. The proposed approach was firstly tested with a synthetic dataset set and the results of this synthetic analysis reached the expected improvements compared to conventional pixel-based classification.

In order to evaluate the proposed approach in an unbiased manner, two post-event VHR images were selected as case studies, Bam and Haiti. In those case studies, two training areas were selected to be used for the FS process and also to use as training area in classification of the entire areas. Three different feature subsets for each case studies, that contain the most frequent features in top ten, twenty and forty rankings, were determined. The determined subsets were evaluated in terms of the classification accuracies and the quality of the thematic maps obtained based on the training areas. The evaluation showed that all the determined subsets improved classification accuracies and the quality of the thematic maps, severely. However, the most severe improvement were observed on the results from the feature subsets selected from top ten features. Furthermore, the visual examination of those selected subsets, as shown in Fig. 3.24 and 3.25, proves their capability to separate different patterns in the selected study areas of the considered images. Therefore, the entire damage maps of the considered post-earthquake images, which were shown in Figure

3.26 and 3.27, were computed by using only the features selected from top ten, and top twenty rankings.

The evaluation also yielded two important results: only one solution for these problem is not possible as the generated features may extract similar information from the image, and the features created by using the same methods and the same parameters may provide different information for the different datasets. For instance, the feature #96 was selected in both of the case studies but in the Bam case it highlights the open ground class, whereas it highlights the damage class in the Haiti case. As a summary, even the selected feature sets significantly improve the performance of the earthquake-induced damage assessment, it is not possible to specify a morphological feature type nor a parameter range for the morphological operations for such a multi-class problem. That might be possible when the damage assessment task is addressed with a two class approach.

The damage map generating process also revealed an important challenge. As mentioned before, the damage maps were generated by the classification results of the entire areas, and these areas were classified within subdivided small parts. In other words, the features in the selected feature subsets were created for those small areas, individually. However, it should be noted that the MPs and APs are extracted based on the geometrical properties in the images and dividing the entire images into small parts also means dividing the geometries in the entire image. It can be defined as a generalization problem of the such image processing tools.

To conclude, advanced methods in mathematical morphology are very efficient source for a fast and an accurate earthquake-induced damage identification from VHR post-event image. The methods provides the needed spatial details that can not be extracted directly from the source image. However, their usage may be improved further by handling the generalization problems, and/or by transforming the classification scheme into two class scheme. The author's future study will be focused on these two challenges.

REFERENCES

- [1] **Dong, L. and Shan, J.** (2013). A comprehensive review of earthquake-induced building damage detection with remote sensing techniques.
- [2] **Voigt, S., Kemper, T., Riedlinger, T., Kiefl, R., Scholte, K. and Mehl, H.** (2007). Satellite Image Analysis for Disaster and Crisis-Management Support, *IEEE Transactions on Geoscience and Remote Sensing*, 45(6), 1520–1528.
- [3] **Stramondo, S., Bignami, C., Chini, M., Pierdicca, N. and Tertulliani, A.** (2006). Satellite radar and optical remote sensing for earthquake damage detection: results from different case studies, *International Journal of Remote Sensing*, 27(20), 4433–4447.
- [4] **Brunner, D., Lemoine, G. and Bruzzone, L.** (2010). Earthquake Damage Assessment of Buildings Using VHR Optical and SAR Imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 48(5), 2403–2420.
- [5] **Dell’Acqua, F. and Gamba, P.** (2012). Remote Sensing and Earthquake Damage Assessment: Experiences, Limits, and Perspectives, *Proceedings of The IEEE*, 100(10, SI), 2876–2890.
- [6] **Gamba, P. and Casciati, F.** (1998). GIs and Image Understanding for Near-Real-Time Earthquake Damage Assessment, *Photogrammetric Engineering & Remote Sensing*, 64(10), 987–994.
- [7] **Ito, Y., Hosokawa, M., Lee, H. and Liu, J.G.** Extraction of damaged regions using SAR data and neural networks.
- [8] **Schweier, C. and Markus, M.** (2006). Classification of Collapsed Buildings for Fast Damage and Loss Assessment, *Bulletin of Earthquake Engineering*, 4(2), 177–192, <https://doi.org/10.1007/s10518-006-9005-2>.
- [9] **ISHII, M.** (2002). Detection of Earthquake Damage Area from Aerial Photographs by Using Color and Edge Information, *ACCV2002*, 27–32, <https://ci.nii.ac.jp/naid/10011073649/en/>.
- [10] **Turker, M. and San, B.T.** (2003). SPOT HRV data analysis for detecting earthquake-induced changes in Izmit, Turkey, *International Journal of Remote Sensing*, 24(12), 2439–2450, <https://doi.org/10.1080/0143116031000070427>, <https://doi.org/10.1080/0143116031000070427>.
- [11] **Kosugi, Y., Sakamoto, M., Fukunishi, M., Doihara, T. and Kakumoto, S.** (2004). Urban change detection related to earthquakes using an adaptive

nonlinear mapping of high-resolution images, *IEEE Geoscience and Remote Sensing Letters*, 1(3), 152–156.

- [12] **Turker, M. and Cetinkaya, B.** (2005). Automatic detection of earthquake-damaged buildings using DEMs created from pre- and post-earthquake stereo aerial photographs, *International Journal of Remote Sensing*, 26(4), 823–832, <https://doi.org/10.1080/01431160512331316810>.
- [13] **Plank, S.** (2014). Rapid Damage Assessment by Means of Multi-Temporal SAR — A Comprehensive Review and Outlook to Sentinel-1, *Remote Sensing*, 6(6), 4870–4906, <http://www.mdpi.com/2072-4292/6/6/4870>.
- [14] **Vu, T.T., Matsuoka, M. and Yamazaki, F.** (2005). Detection and Animation of Damage Using Very High-Resolution Satellite Data Following the 2003 Bam, Iran, Earthquake, *Earthquake Spectra*, 21(S1), 319–327, <https://doi.org/10.1193/1.2101127>, <https://doi.org/10.1193/1.2101127>.
- [15] **Trianni, G. and Gamba, P.** (2008). Damage detection from SAR imagery: Application to the 2003 Algeria and 2007 Peru earthquakes, *International Journal of Navigation and Observation*, 2008.
- [16] **Mitomi, H., Matsuoka, M. and Yamazaki, F.** (2002). Application of automated damage detection of buildings due to earthquakes by panchromatic television images, *The 7th US National Conference on Earthquake Engineering, CD-ROM*, p. 10.
- [17] **Samadzadegan, F. and Rastiveisi, H.** Automatic detection and classification of damaged buildings, using high resolution satellite imagery and vector data.
- [18] **Dubois, D. and Lepage, R.** (2014). Fast and Efficient Evaluation of Building Damage From Very High Resolution Optical Satellite Images, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(10), 4167–4176.
- [19] **Rathje, E.M., Crawford, M., Woo, K. and Neuenschwander, A.** (2005). Damage Patterns from Satellite Images of the 2003 Bam, Iran, Earthquake, *Earthquake Spectra*, 21(S1), 295–307, <https://doi.org/10.1193/1.2101047>, <https://doi.org/10.1193/1.2101047>.
- [20] **Sun, W., Shi, L., Yang, J. and Li, P.** (2016). Building Collapse Assessment in Urban Areas Using Texture Information From Postevent SAR Data, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(8), 3792–3808.
- [21] **Pesaresi, M. and Benediktsson, J.A.** (2001). A new approach for the morphological segmentation of high-resolution satellite imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 39(2), 309–320.

- [22] **Benediktsson, J.A., Pesaresi, M. and Amason, K.** (2003). Classification and feature extraction for remote sensing images from urban areas based on morphological transformations, *IEEE Transactions on Geoscience and Remote Sensing*, 41(9), 1940–1949.
- [23] **Huang, X. and Zhang, L.** (2012). Morphological Building/Shadow Index for Building Extraction From High-Resolution Imagery Over Urban Areas, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(1), 161–172.
- [24] **Chini, M., Pierdicca, N. and Emery, W.J.** (2009). Exploiting SAR and VHR Optical Images to Quantify Damage Caused by the 2003 Bam Earthquake, *IEEE Transactions on Geoscience and Remote Sensing*, 47(1), 145–152.
- [25] **Pesaresi, M. and Benediktsson, J.A.** (2001). A new approach for the morphological segmentation of high-resolution satellite imagery, *IEEE transactions on Geoscience and Remote Sensing*, 39(2), 309–320.
- [26] **Dell’Acqua, F., Bignami, C., Chini, M., Lisini, G., Polli, D.A. and Stramondo, S.** (2011). Earthquake Damages Rapid Mapping by Satellite Remote Sensing Data: L’Aquila April 6th, 2009 Event, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(4), 935–943.
- [27] **Wang, J., Qin, Q., Zhao, J., Ye, X., Feng, X., Qin, X. and Yang, X.** (2015). Knowledge-Based Detection and Assessment of Damaged Roads Using Post-Disaster High-Resolution Remote Sensing Image, *Remote Sensing*, 7(4), 4948–4967, <http://www.mdpi.com/2072-4292/7/4/4948>.
- [28] **Ma, J. and Qin, S.** (2012). Automatic depicting algorithm of earthquake collapsed buildings with airborne high resolution image, *2012 IEEE International Geoscience and Remote Sensing Symposium*, pp.939–942.
- [29] **Li, L., Li, Z., Zhang, R., Ma, J. and Lei, L.** (2010). Collapsed buildings extraction using morphological profiles and texture statistics - A case study in the 5.12 wenchuan earthquake, *2010 IEEE International Geoscience and Remote Sensing Symposium*, pp.2000–2002.
- [30] **Mura, M.D., Benediktsson, J. and Bruzzone, L.** (2009). Modeling structural information for building extraction with morphological attribute filters, *Image and Signal Processing for Remote Sensing XV*, volume 7477, International Society for Optics and Photonics, p.747703.
- [31] **Fukunaga, K.** (1990). Chapter 10 - Feature extraction and linear mapping for classification, **K. Fukunaga**, editor, *Introduction to Statistical Pattern Recognition (Second Edition)*, Academic Press, Boston, second edition edition, pp.441 – 507, <http://www.sciencedirect.com/science/article/pii/B9780080478654500168>.
- [32] **Pedergnana, M., Marpu, P.R., Dalla Mura, M., Benediktsson, J.A. and Bruzzone, L.** (2013). A novel technique for optimal feature selection in attribute profiles based on genetic algorithms, *IEEE Transactions on Geoscience and Remote Sensing*, 51(6), 3514–3528.

- [33] **Bhardwaj, K. and Patra, S.** (2018). An unsupervised technique for optimal feature selection in attribute profiles for spectral-spatial classification of hyperspectral images, *ISPRS Journal of Photogrammetry and Remote Sensing*, 138, 139 – 150, <http://www.sciencedirect.com/science/article/pii/S0924271618300340>.
- [34] **Musson, R.M. and Cekić, I.** (2012). Intensity and intensity scales, *New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, 1–41.
- [35] **Grünthal, G.** (1998). European macroseismic scale 1998, **Technical Report**, European Seismological Commission (ESC).
- [36] **Yamazaki, F., Yano, Y. and Matsuoka, M.** (2005). Visual Damage Interpretation of Buildings in Bam City Using QuickBird Images Following the 2003 Bam, Iran, Earthquake, *Earthquake Spectra*, 21(S1), 329–336, <https://doi.org/10.1193/1.2101807>, <https://doi.org/10.1193/1.2101807>.
- [37] **Rastiveis, H., Samadzadegan, F. and Reinartz, P.** (2013). A fuzzy decision making system for building damage map creation using high resolution satellite imagery, *Natural Hazards and Earth System Sciences*, 13(2), 455.
- [38] **Romaniello, V., Piscini, A., Bignami, C., Anniballe, R. and Stramondo, S.** (2017). Earthquake damage mapping by using remotely sensed data: the Haiti case study, *Journal of Applied Remote Sensing*, 11(1), 016042.
- [39] **Alpaydin, E.** (2014). *Introduction to machine learning*, MIT press.
- [40] **Cover, T. and Hart, P.** (1967). Nearest neighbor pattern classification, *IEEE transactions on information theory*, 13(1), 21–27.
- [41] **Ran, Q., Li, W., Du, Q. and Yang, C.** (2015). Hyperspectral image classification for mapping agricultural tillage practices, *Journal of Applied Remote Sensing*, 9(1), 097298.
- [42] **Ho, Y. and Pepyne, D.** (2002). Simple Explanation of the No-Free-Lunch Theorem and Its Implications, *Journal of Optimization Theory and Applications*, 115(3), 549–570, <https://doi.org/10.1023/A:1021251113462>.
- [43] **Vapnik, V. and Vapnik, V.** (1998), *Statistical learning theory*.
- [44] **Cortes, C. and Vapnik, V.** (1995). Support-vector networks, *Machine learning*, 20(3), 273–297.
- [45] **Melgani, F. and Bruzzone, L.** (2004). Classification of hyperspectral remote sensing images with support vector machines, *IEEE Transactions on geoscience and remote sensing*, 42(8), 1778–1790.
- [46] **Taskin, G., Ersoy, O.K. and Kamasak, M.E.** (2015). Earthquake-induced damage classification from postearthquake satellite image using spectral and spatial features with support vector selection and adaptation, *Journal of Applied Remote Sensing*, 9, 9 – 9 – 20, <https://doi.org/10.1117/1.JRS.9.096017>.

- [47] **Mountrakis, G., Im, J. and Ogole, C.** (2011). Support vector machines in remote sensing: A review, *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 247–259.
- [48] **Chang, C.C. and Lin, C.J.** (2011). LIBSVM: A library for support vector machines, *ACM transactions on intelligent systems and technology (TIST)*, 2(3), 27.
- [49] **Wang, M., Wan, Y., Ye, Z., Gao, X. and Lai, X.** (2018). A band selection method for airborne hyperspectral image based on chaotic binary coded gravitational search algorithm, *Neurocomputing*, 273, 57–67.
- [50] **Çukur, H., Binol, H., Uslu, F.S., Kalaycı, Y. and Bal, A.** (2015). Cross correlation based clustering for feature selection in hyperspectral imagery, *2015 9th International Conference on Electrical and Electronics Engineering (ELECO)*, IEEE, pp.232–236.
- [51] **Leignel, C., Caelen, O., Debeir, O., Hanson, E., Leloup, T., Simler, C., Beumier, C., Bontempi, G. and Wolff, E.** (2010). Detecting man-made structure changes to assist geographic data producers in planning their update strategy, *ISPRS Archive*, 38(Part 4), 8–2.
- [52] **Hanchuan Peng, Fuhui Long and Ding, C.** (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8), 1226–1238.
- [53] **Breen, E. and Jones, R.** (1996). Attribute openings, thinnings, and granulometries, *Computer Vision and Image Understanding*, 64(3), 377–389.
- [54] **Dalla Mura, M., Benediktsson, J.A., Chanussot, J. and Bruzzone, L.,** (2011). The Evolution of the Morphological Profile: from Panchromatic to Hyperspectral Images, **Prasad, S and Bruce, LM and Chanussot, J,** editor, *Optical Remote Sensing: Advances in Signal Processing and Exploitation Techniques*, volume 3 of *Augmented Vision and Reality*, pp.123–146.
- [55] **Dalla Mura, M., Benediktsson et al.** (2010). Morphological attribute profiles for the analysis of very high resolution images, *IEEE Transactions on Geoscience and Remote Sensing*, 48(10), 3747–3762.
- [56] **Mura, M.D., Benediktsson, J.A. et al.** (2010). Extended profiles with morphological attribute filters for the analysis of hyperspectral data, *International Journal of Remote Sensing*, 31(22), 5975–5991, <https://doi.org/10.1080/01431161.2010.512425>, <https://doi.org/10.1080/01431161.2010.512425>.
- [57] **Ghassemian, H.** (2016). A review of remote sensing image fusion methods, *Information Fusion*, 32, 75–89.
- [58] **Alataş, E.O. and Taşkın, G.** (2019). Attribute Profiles in Earthquake Damage Identification from Very High Resolution Post Event Image, *IGARSS 2019*

- 2019 IEEE International Geoscience and Remote Sensing Symposium,
pp.9299–9302.



CURRICULUM VITAE



Name Surname: Enes Oğuzhan Alataş

Place and Date of Birth: Erciş 07.09.1993

E-Mail: oguzenesalatas@gmail.com

EDUCATION:

- **B.Sc.:** 2017, Yıldız Technical University, Civil Engineering
- **M.Sc.:** 2020, Istanbul Technical University, Earthquake Engineering

PROFESSIONAL EXPERIENCE AND REWARDS:

- 08/2018-12/2019 - Research Scholar in TUBITAK Project ID: 217E032

PUBLICATION LIST:

- **E. O. Alataş** and G. Taşkın, "Attribute Profiles in Earthquake Damage Identification from Very High Resolution Post Event Image," IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019, pp. 9299-9302.
- **E. O. Alataş**, G. Taşkın, A. Ertürk, K. Kayabol, "Post-Earthquake Image Classification with The Joint Use of Attribute Profiles and Simple Linear Iterative Clustering " M2GARSS 2020 - 2020 Mediterranean and Middle-East Geoscience and Remote Sensing Symposium, Tunis, Tunisia, 2020.