

COMPARATIVE ANALYSIS OF  
MULTIPLE MACHINE LEARNING ALGORITHMS  
FOR POST-EARTHQUAKE BUILDING DAMAGE ASSESSMENT  
IN HATAY CITY FOLLOWING THE 2023 EARTHQUAKE

A THESIS SUBMITTED TO  
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES  
OF  
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ÖMER KAYA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR  
THE DEGREE OF MASTER OF SCIENCE  
IN  
GEODETIC AND GEOGRAPHIC INFORMATION TECHNOLOGIES

APRIL 2024



Approval of the thesis:

**COMPARATIVE ANALYSIS OF  
MULTIPLE MACHINE LEARNING ALGORITHMS  
FOR POST-EARTHQUAKE BUILDING DAMAGE ASSESSMENT  
IN HATAY CITY FOLLOWING THE 2023 EARTHQUAKE**

submitted by **ÖMER KAYA** in partial fulfillment of the requirements for the degree of **Master of Science in Geodetic and Geographic Information Technologies, Middle East Technical University** by,

Prof. Dr. Naci Emre Altun  
Dean, **Graduate School of Natural and Applied Sciences** \_\_\_\_\_

Prof. Dr. Sevda Zuhall Akyürek  
Head of the Department, **Geodetic and Geographic Information Technologies** \_\_\_\_\_

Prof. Dr. Sevda Zuhall Akyürek  
Supervisor, **Civil Engineering, METU** \_\_\_\_\_

Assoc. Prof. Dr. Semih Kuter  
Co-Supervisor, **Forest Engineering, Çankırı Karatekin University** \_\_\_\_\_

**Examining Committee Members:**

Prof. Dr. Mehmet Lütfi Süzen  
Geological Eng., METU \_\_\_\_\_

Prof. Dr. Sevda Zuhall Akyürek  
Civil Eng., METU \_\_\_\_\_

Assoc. Prof. Dr. Meltem Şenol Balaban  
Department of City and Regional Planning, METU \_\_\_\_\_

Assoc. Prof. Dr. Semih Kuter  
Forest Eng., Çankırı Karatekin Uni. \_\_\_\_\_

Assoc. Prof. Dr. Berk Anbarođlu  
Geomatics Eng., Hacettepe Uni. \_\_\_\_\_

Date: 26.04.2024

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

Name Last name : Ömer Kaya

Signature :

## **ABSTRACT**

### **COMPARATIVE ANALYSIS OF MULTIPLE MACHINE LEARNING ALGORITHMS FOR POST-EARTHQUAKE BUILDING DAMAGE ASSESSMENT IN HATAY CITY FOLLOWING THE 2023 EARTHQUAKE**

Kaya, Ömer

Master of Science, Geodetic and Geographic Information Technologies

Supervisor: Prof. Dr. Sevda Zuhul Akyürek

Co-Supervisor: Assoc. Prof. Dr. Semih Kuter

April 2024, 161 pages

This thesis presents earthquake damage assessment methodologies, focusing on integrating satellite imagery and machine learning (ML) algorithms. The study aims to enhance the accuracy and efficiency of identifying damaged buildings using only post-earthquake satellite images by combining remote sensing and ML techniques. ML algorithms, like Multivariate Adaptive Regression Splines (MARS), Artificial Neural Networks (ANN), Support Vector Machines (SVM), Random Forest (RF), and ensemble learning methods, are evaluated for their effectiveness in automatic feature extraction from satellite imagery. Principal Component Analysis (PCA) is employed with the derived texture image bands, augmenting the discriminative power of the models. A ten-fold cross-validation process is used in the study.

The study also tackles challenges in classifying building footprints from non-nadir imagery, where intricate building shapes and densities impede shadow detection, requiring a customized approach integrating vector data refinement. Examining building footprint areas across folds uncovers considerable variations in size

distribution, particularly evident in Fold 5. Evaluating the highest accuracy fold with the Damage Proximity Map emphasizes the criticality of aligning ML classification outcomes with ground truth data.

The results of the study highlight several key findings. SVM with a linear kernel emerges as the top-performing algorithm, mainly exhibiting superior accuracy, achieving 57% to 64.82% accuracy. Additionally, MARS demonstrates stable performance across folds, maintaining accuracies around 60% to 62%. In contrast, SVM with 2nd-degree polynomial kernel and ensemble methods exhibit different inaccuracies across folds.

**Keywords:** Multispectral Imagery, Post-Earthquake Building Damage Assessment, Machine Learning.

## ÖZ

### HATAY İLİNDE 2023 DEPREMİ SONRASI BİNA HASAR TESPİTİ İÇİN ÇOKLU MAKİNE ÖĞRENME ALGORİTMALARININ KARŞILAŞTIRMALI ANALİZİ

Kaya, Ömer  
Yüksek Lisans, Jeodezi ve Coğrafi Bilgi Teknolojileri  
Tez Yöneticisi: Prof. Dr. Sevda Zuhul Akyürek  
Ortak Tez Yöneticisi: Doç. Dr. Semih Kuter

Nisan 2024, 161 sayfa

Bu tez, uydu görüntüleme ve makine öğrenme algoritmalarının entegrasyonuna odaklanarak deprem hasar değerlendirme metodolojilerine yer vermektedir. Çalışmanın amacı, uzaktan algılama ve makine öğrenme tekniklerini kullanarak sadece deprem sonrası uydu görüntüleri kullanılarak hasarlı binaların tespit edilmesinin doğruluğunu ve verimliliğini artırmaktır. Makine öğrenme algoritmaları, Çok Değişkenli Uyarlanabilir Regresyon Eğrileri (MARS), Yapay Sinir Ağları (ANN), Destek Vektör Makineleri (SVM), Rastgele Orman (RF) ve toplu öğrenme yöntemleri gibi, uydu görüntülerinden otomatik özellik çıkarma etkinliği değerlendirilmiştir. Temel Bileşenler Analizi (PCA), doku görüntü bantlarını kullanarak modellerin ayırt edici gücünü artırmak amacı ile uygulanmıştır. Çalışmada 10-sıfırlı çapraz doğrulama yöntemi kullanılmıştır.

Çalışma ayrıca, gölge algılamayı engelleyen karmaşık bina şekilleri ve yoğunlukları gibi sorunlarla uğraşırken, özel bir yaklaşım gerektiren vektör veri iyileştirmesini entegre eden özel bir yaklaşıma ihtiyaç duyan nadir olmayan görüntülerden bina izlerini sınıflandırmadaki zorlukları ele almaktadır. Katmanlar

arasında bina izi alanlarını incelemek, özellikle Katman 5'te belirgin bir şekilde görülen boyut dağılımındaki önemli deęişkenlikleri açığa çıkarmaktadır. En yüksek doğruluk katmanının Hasar Yakınlık Haritası ile deęerlendirilmesi, makine öğrenme sınıflandırma sonuçlarının bina sınırı verileriyle hizalanmasının önemini vurgulamaktadır.

Çalışmanın sonuçları birkaç önemli bulguyu ortaya koymaktadır. Doğrusal kernel SVM, özellikle üstün doğruluk sergileyen en iyi performans gösteren algoritma olarak ortaya çıkmaktadır ve %57 ila %64,82 arasında doğruluk sağlamaktadır. Ayrıca, MARS, katmanlar arasında stabil performans sergileyerek %60 ila %62 arasında doğruluk sağlamaktadır. Bununla birlikte, ikinci dereceden polinom kernel SVM ve toplu öğrenme yöntemleri katmanlar arasında farklı hatalar sergilemektedir.

**Anahtar kelimeler:** Çok Spektralı Görüntüler, Deprem Sonrası Bina Hasar Deęerlendirmesi, Makine Öğrenimi.



*To my loving wife and daughter*

## ACKNOWLEDGMENTS

The author extends his deepest gratitude to his esteemed advisor, Prof. Dr. Zuhâl Akyürek, and his dedicated co-advisor, Assoc. Prof. Dr. Semih Kuter, for their unwavering support and understanding during his challenges. He feels genuinely fortunate to have benefited from their invaluable guidance and contributions to this study.

The author also expresses indebtedness to the examining committee members, Prof. Dr. Mehmet Lütfi Süzen, Assoc. Prof. Dr. Meltem Şenol Balaban, and Assoc. Prof. Dr. Berk Anbarođlu, for their insightful advice and constructive feedback.

Heartfelt appreciation is extended to the author's three most significant pillars of support: his mother, Menşure Kaya; his father, Enver Kaya; and his wife, Nazlı Dilay Çakır Kaya. The author acknowledges that their unwavering belief in them sustained them through this journey, and they are profoundly grateful for their love and encouragement.

Special thanks are offered to Yeşeren Çakır for her support since the birth of their child and her unwavering presence, which was a constant motivation throughout this process.

Additionally, the author acknowledges the Directorate General of Geographic Information Systems for providing the vector data in Hatay City with ground truth, which was instrumental in the research conducted for this thesis.

Lastly, the author extends their most profound appreciation to Esmâ Hışıl and Dr. Hüseyin Hışıl for their boundless encouragement. They acknowledge that their unwavering support has been their strength through every step of this thesis, and for that, the author is eternally grateful.

## TABLE OF CONTENTS

ABSTRACT.....	v
ÖZ .....	vii
ACKNOWLEDGMENTS .....	x
TABLE OF CONTENTS.....	xi
LIST OF TABLES .....	xiii
LIST OF FIGURES .....	xiv
LIST OF ABBREVIATIONS .....	xix
1 INTRODUCTION .....	1
2 LITERATURE REVIEW .....	5
2.1 Background on Earthquake Damage Assessment.....	5
2.2 Application of earthquake Damage Assessment with Remote Sensing Imagery and Artificial Intelligence. ....	7
2.2.1 Limitations and Proposed Solutions on Earthquake Damage Assessment in Literature .....	10
3 MATERIALS AND METHODS.....	13
3.1 Study Area.....	13
3.2 The Dataset.....	15
3.2.1 Satellite Data.....	15
3.2.2 Vector Data .....	19
3.2.3 Ancillary Data .....	27
3.3 Methodology .....	28
3.3.1 Satellite Image Band Reflectance Calculations .....	31
3.3.3 Input Clusters of the Dataset.....	35

3.3.4	Building Classification .....	37
3.4	Machine Learning Algorithms.....	40
3.4.1	Support Vector Machine (SVM) .....	40
3.4.2	Artificial Neural Networks (ANN).....	41
3.4.3	Random Forest (RF) .....	42
3.4.4	Multivariate Adaptive Regression Splines (MARS) .....	43
3.4.5	Ensemble Learning .....	45
4	RESULTS AND DISCUSSIONS .....	49
4.1	Discussions .....	85
4.1.1	Consistency Evaluation of Machine Learning Algorithms .....	86
4.1.2	Challenges and Considerations in Building Footprint Classification from Non-Nadir Imagery.....	87
4.1.3	Challenges in Shadow Detection for Building Footprints Due to Complex Building Shapes and Density .....	88
4.1.4	Examining Variability in Building Footprint Areas Across Folds ....	90
4.1.5	Computational Efficiency Analysis of Machine Learning Models ...	92
4.1.6	Comparison with Existing Literature .....	94
4.1.7	Comparison of Highest Accuracy Fold with Damage Proximity Map 96	
4.1.8	Assessing the Ground Truth Vector Data.....	101
4.1.9	Exploring Factors Influencing Misclassification.....	102
5	CONCLUSIONS AND RECOMMENDATIONS .....	109
	APPENDICES .....	118
A.	Histogram Charts of Texture Images Produced from Reflectance Image Bands 118	

## LIST OF TABLES

### TABLES

Table 3.1. Time, date of acquisition, image location, sensor, spatial resolution, and band composition of the satellite imagery. ....	16
Table 3.2 Spectral bands of Pleiades – 1B (Laurent Coeurdevey and Karine Fernandez, 2012).....	17
Table 3.3. Building condition percentages of the dataset. ....	22
Table 3.4. Main input combinations. ....	36
Table 4.1. Training data error matrix for RF ML algorithm.....	67
Table 4.2. Test data error matrix for RF ML algorithm.....	67
Table 4.3. Prediction accuracy for the building classes.....	67
Table 4.4. User accuracy for the building classes.....	68
Table 4.5. Minimum Correlation of the Directional Texture Image Bands (Direct_PCA_minCor) and added reflectance image bands (Ref_Direct_PCA_minCor). ....	75
Table 4.6. ML algorithms accuracy percentage for Direct_PCA_minCor dataset .	82
Table 4.7. ML algorithms accuracy percentage for Ref_Direct_PCA_minCor dataset .....	82
Table 4.8. Confusion matrix for the 5 <sup>th</sup> fold of SVM linear kernel ML algorithm building classification with respect to vector data provided by GD of GIS (2023). ....	100
Table 4.9. Confusion matrix for the 5 <sup>th</sup> fold of SVM linear kernel ML algorithm building classification with respect to DPM. ....	100

## LIST OF FIGURES

### FIGURES

Figure 3.1. The location of the study area. ....	14
Figure 3.2. Image footprints provided by Pleiades and UHUZAM (UHUZAM, 2023).....	16
Figure 3.3. Study area with pan-sharpened 50 cm Pleiades 1B image (Satellite Image Acquisition Date: 08.02.2023). ....	18
Figure 3.4. Vector data acquired from the Directorate of Geographical Information Systems, ATLAS Project with classified building footprints (GD of GIS, 2023). .	20
Figure 3.5. Landmark checkpoints. ....	23
Figure 3.6. Pre-edit vector data where the vector data is not aligned with the buildings from the satellite images.....	24
Figure 3.7. Post-edited vector data where the satellite image and vector data building footprints match. ....	25
Figure 3.8. Part of the study area shows the density of buildings where shadows cannot be detected individually. ....	27
Figure 3.9. NASA/ARIA program DPM for the study area (NASA/JPL-Caltech/USGS/JAXA/Earth Observatory of Singapore – Remote Sensing Lab, 2023).....	29
Figure 3.10. Flowchart of the methodology. ....	30
Figure 3.11. Building classification with EMS98 (Grünthal, 1998). ....	38
Figure 3.12. The proposed building damage scale is tailored to RS vertical imagery (Cotrufo <i>et al.</i> , 2018). ....	39
Figure 3.13. The study used two building classes compared to the EMS98 building scale. ....	40
Figure 3.14. Symmetric boolean functions of MARS (Hastie <i>et al.</i> , 2009).....	44
Figure 4.1. Native (DN) image band one (Red) histogram .....	50
Figure 4.2. Reflectance (TOA) image band one (Red) histogram .....	50
Figure 4.3. Native (DN) image band two (Green) histogram.....	51
Figure 4.4. Reflectance (TOA) image band two (Green) histogram.....	51

Figure 4.5. Native (DN) image band three (Blue) histogram .....	52
Figure 4.6. Reflectance (TOA) image band three (Blue) histogram.....	52
Figure 4.7. Native (DN) image band four (NIR) histogram .....	53
Figure 4.8. Reflectance (TOA) image band four (NIR) histogram.....	53
Figure 4.9. Sample images of produced texture image bands overlapping with building footprints.....	54
Figure 4.10. The mean texture histogram for Band one of the heavily damaged class.....	55
Figure 4.11. The mean texture histogram for Band one of the heavily no-damage class.....	56
Figure 4.12. The mean texture histogram for Band one of the collapsed class. ....	56
Figure 4.13. The mean texture histogram for Band one of the urgent demolish class.....	57
Figure 4.14. The mean values of the image bands with standard deviation error bars.....	58
Figure 4.15. The mean values of the contrast texture image bands with standard deviation error bars. ....	58
Figure 4.16. The mean values of the correlation texture image bands with standard deviation error bars. ....	59
Figure 4.17. The mean values of the dissimilarity texture image bands with standard deviation error bars.....	59
Figure 4.18. The mean values of the entropy texture image bands with standard deviation error bars. ....	60
Figure 4.19. The mean values of the homogeneity texture image bands with standard deviation error bars.....	60
Figure 4.20. The mean values of the angular 2 <sup>nd</sup> moment texture image bands with standard deviation error bars.....	61
Figure 4.21. Heavily damaged building class correlation matrix. ....	63
Figure 4.22. No-damage building class correlation matrix.....	64
Figure 4.23. Collapsed building class correlation matrix. ....	65

Figure 4.24. Urgent demolish building class correlation matrix. ....	66
Figure 4.25. Correlation matrix for the PCA analysis for band one_texture .....	69
Figure 4.26. Correlation matrix for the PCA analysis for band two_texture. ....	70
Figure 4.27. The correlation matrix for the PCA analysis for band three_texture..	71
Figure 4.28. Correlation matrix for the PCA analysis for band four_texture.....	72
Figure 4.29. Correlation matrix for the PCA Bands 2 and 4 with directional texture image applied to No-Damage and Damaged building class.....	73
Figure 4.30. Minimum correlation matrix for directional image bands used for ML models.....	74
Figure 4.31. Mean value extraction for all buildings in the study area. ....	76
Figure 4.32. Distribution of Folds (Partitions) of the study area.....	78
Figure 4.33. Training data accuracy for Direct_PCA_minCor dataset with standard deviation error bars (11 Image Bands). ....	79
Figure 4.34. Training data accuracy for Ref_Direct_PCA_minCor dataset with standard deviation error bars (15 Image Bands).....	79
Figure 4.35. Test data accuracy for Direct_PCA_minCor dataset with standard deviation error bars (11 Image Bands). ....	80
Figure 4.36. Test data accuracy for Ref_Direct_PCA_minCor dataset with standard deviation error bars (15 Image Bands). ....	80
Figure 4.37. The distribution of the buildings in Fold number 5 in the study area.	84
Figure 4.38. Area distribution (MIN) with standard deviation error bars across folds. ....	91
Figure 4.39. Area distribution (MAX) with standard deviation error bars across folds. ....	91
Figure 4.40. Standard deviation of the area distribution with standard deviation error bars across folds.....	91
Figure 4.41. CPU runtimes for ML algorithms. ....	93
Figure 4.42. The comparison of fold number five to DPM.....	98
Figure 4.43. True negative and true positive distribution of the fold five building classification using SVM linear kernel.....	103



Figure 4.44. False negative and false positive distribution of the fold five building classification using SVM linear kernel. ....	104
Figure 4.45. True positive (green) and true negative (red) examples of uniform building reflectance. ....	105
Figure 4.46. False positive (yellow) and false negative (blue) tilted building samples. ....	106
Figure 4.47. Large area building with non-uniform building footprint reflectance. ....	107
Figure 4.48. Satellite dishes reflect the sunlight, causing changes in the reflectance values of the building footprint. ....	107

## LIST OF ABBREVIATIONS

### ABBREVIATIONS

AFAD	Disaster and Emergency Management Presidency
AI	Artificial Intelligence
ARIA	Advanced Rapid Imaging and Analysis
ANN	Artificial Neural Networks
ALOS-2	Advanced Land Observing Satellite-2
CNN	Convolutional Neural Networks
GCN	Graph Convolutional Networks
DL	Deep Learning
DN	Digital Number
GIS	Geographic Information Systems
OBIA	Object-Based Image Analysis
MARS	Multivariate Adaptive Regression Splines
ML	Machine Learning
MLP	Multi-Layer Perceptron
UAV	Unmanned Aerial Vehicles
UTC	Coordinated Universal Time
RF	Random Forest
RNN	Recurrent Neural Networks
RS	Remote Sensing

SAR	Synthetic Aperture Radar
SVM	Support Vector Machine
TRT	Türkiye Radio and Television Corporation
TOA	Top of Atmosphere
YOLO	You Only Look Once Object Detection Algorithm
xBD	Large Satellite Imagery Dataset for Disaster Management



# CHAPTER 1

## INTRODUCTION

Natural disasters, particularly earthquakes, pose significant challenges to urban landscapes, necessitating advanced methodologies for efficient and accurate damage assessment. The aftermath of earthquakes often demands rapid and precise evaluations of building-damaged properties to facilitate targeted recovery efforts.

In Turkey, we encountered two earthquakes on Monday, February 6th, 2023, with local times recorded at 04:17 (UTC + 3) and 10:24 (UTC + 3). These earthquakes, measuring 7.8 and 7.7 in magnitude, respectively, profoundly impacted thirteen cities. The affected cities include Adana, Adıyaman, Batman, Diyarbakır, Elazığ, Gaziantep, Hatay, Kahramanmaraş, Kilis, Malatya, Mardin, Osmaniye, and Şanlıurfa. This seismic activity was a significant event with widespread implications for the region and the whole country (TRT, 2023).

After the earthquake, our observations unveiled a significant challenge confronting national rescue missions coordinated by the relevant governmental agency, Disaster and Emergency Management Presidency (AFAD). The enormity of the seismic event, affecting 13 cities, highlighted the necessity for a more nuanced and efficient approach to guide the allocation of essential aid to buildings across these urban centers. Given the scale of devastation and the imperative to reach all collapsed and damaged structures promptly, traditional rescue missions encountered constraints in deploying the requisite resources, including excavators and human personnel with specialized expertise. This scenario underscores the demand for alternative methods to determine the urgency of a building's damage status, transitioning from damaged to no damage.

In recent years, the integration of remote sensing (RS) and Geographic Information Systems (GIS) have emerged as pivotal tools for disaster management and

response, particularly in the assessment of post-earthquake damage. Earthquakes, as natural phenomena, often leave widespread destruction in their wake, causing significant damage to infrastructure and posing imminent threats to human lives. Conventional damage assessment techniques, dependent on ground surveys and manual inspection, consume considerable time and encounter logistical hurdles, particularly in inaccessible or difficult zones. However, the advent of RS technologies, coupled with GIS tools, has revolutionized this process, offering a rapid and comprehensive means of assessing the extent and severity of the damage. Using satellite and aerial platforms, incorporating high-resolution imagery acquired by multispectral sensors alongside sophisticated spatial analysis methodologies, it becomes feasible to identify, quantify, and map the effects of earthquakes with unparalleled precision and efficacy. In this context, this thesis delves into the application of such technologies, explicitly focusing on detecting post-earthquake damage using RS data from Planet Pleiades and employing various ML algorithms for comprehensive and comparative assessment.

An early study by Yamazaki *et al.* (2003) emphasizes the main strides in RS technologies and their applications in capturing the distribution of damaged urban areas. These advancements have proven instrumental in detecting and guiding emergency management and recovery efforts toward the necessary sites. The data captured is particularly pertinent for analysis, with the optimal timeframe being immediately post-earthquake. Areas requiring urgent attention are susceptible to significant contamination from a satellite imaging perspective. The presence of heavy machinery engaged in emergency response activities can impact reflectance values. Additionally, buildings that have collapsed onto each other undergo alterations in their building footprints. Moreover, individuals displaced from their homes often establish temporary living spaces, such as tents, introducing changes to the reflective properties of the affected structures.

Recent advancements in ML techniques offer promising solutions to improve the accuracy and efficiency of damaged building detection from RS data. Convolutional Neural Networks (CNNs) have emerged as particularly effective

tools for image classification tasks, including identifying damaged buildings in high-resolution satellite imagery (Yang *et al.*, 2019). By automatically extracting relevant features from RS data, CNNs can overcome the limitations of traditional classification methods and achieve higher accuracy rates in detecting damaged structures (Zhu *et al.*, 2020).

However, despite the potential of ML-based approaches, several challenges persist in damaged building detection using RS data. One major challenge is the variability in environmental conditions and damage patterns, which can hinder the generalization capability of ML models trained on specific datasets (Chen *et al.*, 2021). Additionally, the availability of labeled training data poses a significant constraint, as manually annotating large-scale RS datasets for damaged building detection is labor-intensive and costly (Kerle *et al.*, 2020).

The study focuses on the urgent need for a sophisticated and data-driven approach to earthquake damage assessment, particularly in densely populated urban areas. Integrating advanced methodologies such as RS and ML aims to improve the accuracy and efficiency of identifying and assessing building damage following earthquakes. It is aimed to present a comparative analysis of multiple ML algorithms for post-earthquake building damage assessment. Through an interdisciplinary approach, the goal is to develop a comprehensive methodology that can swiftly identify damaged buildings, providing crucial information to emergency response teams.

In dense urban environments, managing emergency responses becomes increasingly challenging, making it crucial to provide response teams with an instant understanding of the affected areas. The thesis emphasizes the importance of directing emergency response teams specifically to damaged buildings rather than inundating them with unnecessary information, such as open roads and potential traffic jams. To achieve this goal, the proposal suggests utilizing satellite imagery and vector map data, combined with ML algorithms, to produce and compare optimal image band compositions for accurate damage detection.

The approach stands out for its focus on compiling and comparing multiple ML algorithms with post-event satellite images and current vector map data. By systematically evaluating various algorithms and data source combinations, the methodology aims to identify the most effective techniques for rapidly and accurately identifying damaged buildings. This targeted approach has the potential to streamline emergency response efforts and improve overall disaster management in densely populated urban areas.

The thesis unfolds in a structured manner, beginning with Chapter 1, which introduces the study's goals and provides an overview of the organizational framework of the thesis. Chapter 2 thoroughly reviews recent literature relevant to the subject matter. In Chapter 3, detailed descriptions of the materials and methods utilized in the research are presented, including discussions on selected study areas, input data, the methodologies employed for data acquisition, and the resulting outcomes. Chapter 4 is dedicated to presenting and discussing the obtained results in depth. Finally, Chapter 5 serves as the conclusion of the thesis, summarizing the work conducted, highlighting its significance, and outlining potential avenues for future research endeavors.



## CHAPTER 2

### LITERATURE REVIEW

#### 2.1 Background on Earthquake Damage Assessment

The field of earthquake damage assessment has been actively explored since 1988, notably employing Object-Based Image Analysis (OBIA) as a pivotal methodology (Gamba and Casciati, 1998). OBIA is a two-step process involving image segmentation and subsequent classification, where each segment is assigned, specific classes using spectral, textural, spatial, and topological properties. This automated approach has proven effective in delineating damaged buildings, particularly showcased in its successful application during the aftermath of the 2004 Bam, Iran earthquake (Janalipour and Mohammadzadeh, 2016).

The field of earthquake damage assessment has witnessed significant advancements, particularly with the integration of RS technologies and artificial intelligence. Researchers have explored various methodologies to enhance the efficiency and accuracy of damage evaluation processes, addressing the critical need for rapid and precise assessments of the outcome of seismic events.

A groundbreaking investigation utilized high-resolution satellite imagery (Yamazaki *et al.*, 2003), notably QuickBird, to evaluate building damage in Zemmouri, Algeria, after the 2003 earthquake. The authors performed visual inspections, classifying buildings into distinct damage levels based on pre-event and post-event images. This study highlighted the severe potential of high-resolution satellite imagery for post-disaster damage assessment.

Kouchi *et al.* (2004) focused on damage detection for the 2003 Boumerdes, Algeria earthquake, employing QuickBird satellite images. Like Yamazaki *et al.* (2003), the authors utilized visual interpretation techniques to classify buildings into

different damage levels, emphasizing the importance of high-resolution satellite imagery in earthquake damage assessment.

The integration of unmanned aerial vehicles (UAVs) and deep learning for post-earthquake damage detection is presented by Takhtkeshha *et al.* (2023), focusing on a case study in Sarpol-e Zahab, Iran. This work introduces a rapid, self-supervised deep learning method using UAV data, providing an efficient and timely approach to assess earthquake-induced damage.

The application of CNNs in building damage detection from satellite imagery is also explored by Ma, Liu, Ren, Wang, *et al.* (2020). CNNs, a deep learning algorithm, are employed to automatically detect building damage, showcasing the effectiveness of ML techniques in image analysis.

Transfer learning is investigated for automatic hurricane damage detection using satellite images by Kaur *et al.* (2022). This study explores the advantages and outcomes of employing transfer learning techniques, which leverage pre-existing knowledge from one task to improve the performance of another in the context of hurricane damage detection.

An unsupervised learning framework for region-based damage assessment on an extensive satellite image database (i.e., xBD) is introduced in Mittal and Bafna (2023). This framework aims to assess damage in large-scale satellite imagery datasets without needing labeled training data.

The study of Ma, Liu, Ren, and Yu (2020) focuses on collapsed building detection using various techniques, including texture and CNN features, multi-temporal optical satellite imagery, synthetic aperture radar (SAR), and improved algorithms like “you only look once” (YOLO) version 3. These works contribute to developing robust methodologies for identifying and classifying collapsed buildings after earthquakes.

## **2.2 Application of earthquake Damage Assessment with Remote Sensing Imagery and Artificial Intelligence.**

The field of earthquake damage assessment has witnessed significant advancements with the integration of RS imagery and artificial intelligence (AI) techniques. Pioneering studies, such as the work by Yamazaki *et al.* (2003), marked a turning point by exploring the use of high-resolution satellite images, specifically QuickBird, to assess building damage in Zemmouri, Algeria, after the earthquake. The authors conducted visual inspections and classified buildings into different damage levels, highlighting the potential of high-resolution satellite imagery for post-disaster damage assessment. Similarly, Kouchi *et al.* (2004) extended this approach to the 2003 Boumerdes, Algeria earthquake, emphasizing the importance of high-resolution satellite imagery in earthquake damage detection.

In recent years, the application of AI, particularly ML and deep learning, has revolutionized earthquake damage assessment. A noteworthy instance is the research conducted by scholars in Iran, where they presented a rapid, self-supervised deep-learning technique for detecting post-earthquake damage using UAV data. This approach, exemplified in the case study of Sarpol-e Zahab, Iran, showcases the efficiency and timeliness of using AI for post-earthquake damage assessment (Takhtkeshha *et al.*, 2023). Moreover, studies focusing on using CNNs for building damage detection in satellite imagery have demonstrated the effectiveness of deep learning techniques in automatically identifying and classifying damaged structures.

In addition to these advancements, traditional ML algorithms have contributed significantly to earthquake damage assessment. Li *et al.* (2017) demonstrated the effectiveness of SVM in classifying post-earthquake damage using features extracted from satellite imagery. Similarly, Rao *et al.* (2023) employed RF classifiers to identify damaged buildings after earthquakes, showcasing the utility of ensemble learning methods in this context. These approaches often involve

feature engineering to extract relevant information from RS data, followed by model training and validation to achieve accurate damage classification results.

As explored in the literature, transfer learning offers another dimension to earthquake damage assessment. Researchers have demonstrated the advantages of transfer learning techniques by leveraging pre-trained models for automatic hurricane damage detection using satellite images. This approach capitalizes on knowledge gained from one domain (e.g., hurricanes) to enhance the performance of models in earthquake damage detection, showcasing the adaptability and versatility of AI methods (Kaur *et al.*, 2022).

The work of Turker and San (2004) on collapsed buildings caused by the 1999 earthquake in Izmit, Turkey, demonstrates the utility of digital analysis of aerial photographs in identifying and understanding the extent of structural damage following an earthquake. This approach enables researchers to assess the damage to buildings and infrastructure quickly and accurately, facilitating more effective disaster response and recovery efforts.

Similarly, Menderes *et al.* (2015) focus on automatically detecting damaged buildings using RS and information technologies. By combining data from various sources such as satellite imagery, GIS, and ML algorithms, the researchers develop automated methods for identifying and mapping damaged buildings after an earthquake. This approach streamlines the assessment process and provides valuable information for emergency responders and urban planners.

Furthermore, Rathje *et al.* (2005) utilize multi-temporal high-resolution optical satellite imagery to identify and assess earthquake damage. By comparing images taken before and after an earthquake, researchers can detect changes in the landscape and infrastructure, allowing for a more thorough analysis of the impact of the seismic event.

More into large-scale damage detection using satellite imagery, the literature hints at broader applications of RS and AI techniques. While specific findings may vary

across studies, the overall trend indicates a transformative impact on earthquake damage assessment methodologies. This literature review explores the evolution, procedures, and outcomes in earthquake damage assessment, leveraging the synergies between RS imagery and artificial intelligence.

However, achieving high accuracy in earthquake damage assessment requires careful consideration of several factors, including image quality, feature selection, and model validation. Low-quality imagery or incomplete coverage can compromise the accuracy of damage assessments, leading to inaccurate conclusions about the extent of destruction. As seen in the studies done by Menderes *et al.* (2015) and Turker and San (2004), the study area consists of a very low building count for the coverage location, denoted as the density of buildings. Although generating building footprints from shadows and conducting building damage classification yields highly accurate results, limitations arise concerning the requirement for building shapes to adhere to rectangular forms, the absence of proximity to vegetation, and shadows not overlapping between buildings. It is also seen in pre- and post-earthquake images or image-derived products in the study of Menderes *et al.* (2015), whereas in this research, only post-earthquake satellite images are used. As within the research of Turker and San (2004), damaged building detection using shadow edges results showed up to 96.15% overall accuracy where the digital number (DN) of the pixel is found to be 52 and 50% threshold is used for the percentage for the edge pixels of shadow-casting edges to successfully separate collapsed from uncollapsed building and detected damaged buildings from shadows (Turker and San, 2004). Also, with damage detection going further into automation using multi-temporal techniques, there is a drop in the overall accuracy of the measure to 82% (Menderes *et al.*, 2015). The primary understanding is the number of buildings in the calculation area and coverage. When earthquakes affect thirteen cities with buildings of varying shapes, automatic building damage classification becomes less accurate. This is especially true in densely populated city centers where using building shadows for automation isn't

feasible. Additionally, dealing with tens of thousands of buildings adds to the complexity of the process.

Small ground sampling distance (GSD) was used by Naito *et al.* (2020). Using 20 cm GSD orthophoto images and defining four level damage classes (level 1, 2, 3, 4) from level 1 buildings, there was no significant damage to the level 4 collapsed building structure. The study achieved 56.6% accuracy using visual interpretation and bag-of-visual-words with SVM with Radial Basis Function (RBF) model and 88.4% overall accuracy using visual interpretation and CNN model. For the two models presented in the paper, the CNN model has a higher overall accuracy. In contrast to the related research, this study with the SVM model achieved a 64.82% accuracy with the use of PCA and directional texture image bands where the GSD for the image is 50 cm; therefore, the processing power required for the model is relatively less where the results in this study given in a matter of seconds instead of hours.

Furthermore, selecting relevant features for damage classification is crucial, as it directly impacts the performance of ML models. Additionally, rigorous validation procedures, such as cross-validation and independent testing, are essential to ensure the reliability of damage assessment models and validate their accuracy across different datasets and scenarios. Overall, prioritizing accuracy in earthquake damage assessment methodologies is necessary for informing decision-making processes and facilitating effective disaster response and recovery efforts.

### **2.2.1 Limitations and Proposed Solutions on Earthquake Damage Assessment in Literature**

While advancements in earthquake damage assessment using RS and AI are promising, several limitations persist, prompting researchers to propose innovative solutions to address these challenges.

One standard limitation revolves around the issue of temporal resolution. Most satellite imagery is constrained by revisit intervals, hindering the ability to capture real-time or near-real-time data during and immediately after seismic events. Proposed solutions include the integration of high-temporal-resolution satellite constellations or the fusion of satellite data with UAV imagery for more timely assessments (Takhtkeshha *et al.*, 2023).

Spatial resolution is another challenge, especially in large urban areas where detailed damage assessment requires fine-scale imagery. Researchers propose using higher-resolution sensors or combining multi-resolution datasets to enhance spatial clarity in damage detection (Rathje *et al.*, 2005). However, this produces time constraints and difficulty in achieving high-resolution image data. Airborne imagery with ortho-rectification, as utilized by researchers (Naito *et al.*, 2020), reveals that acquiring and post-production such images entails significant waiting times in data analysis. Consequently, with recent technological advancements, efforts are made to achieve comparable accuracy using satellite images with lower Ground Sample Distance (GSD). However, the processing time for image production and preparation of information is significantly reduced despite the broader coverage of satellite imagery.

Addressing the challenge of post-disaster information overload, where the sheer volume of data can overwhelm response teams, proposed solutions involve integrating automated damage assessment tools with GIS to provide decision-makers with actionable insights in a more accessible format (Gamba and Casciati, 1998).

While significant strides have been made in earthquake damage assessment through RS and AI, persistent limitations demand ongoing research and innovation. The proposed solutions discussed here underscore the interdisciplinary nature of this field, combining advancements in sensor technology, ML methodologies, and data management strategies to enhance the effectiveness of post-earthquake damage assessments (Liu, 2014).

Cloud cover poses a significant obstacle, as it can obscure critical information during image acquisition. Researchers suggest the development of advanced cloud-penetrating sensors or using SAR, which is less affected by atmospheric conditions, as potential solutions to mitigate this limitation (Ge *et al.*, 2019).

A notable challenge is the interpretability of ML models' intense learning algorithms like CNNs. Proposals for more explainable models or the integration of interpretability techniques aim to enhance the transparency and trustworthiness of these algorithms in damage assessment applications (Ma, Liu, Ren, Wang, *et al.*, 2020).

Labeling and annotating large datasets for training ML models can be time-consuming and resource-intensive. Researchers advocate for developing transfer learning techniques, where pre-trained models are adapted to new datasets with limited labeled samples, streamlining the training process and overcoming data scarcity challenges (Kaur *et al.*, 2022).

The efficacy of damage detection methodologies in structural assessment for buildings has been a subject of scholarly scrutiny, particularly concerning the application of ML, DL, and traditional classification techniques. Recent research underscores the superior performance of DL methodologies, notably CNNs, in accurately discerning structural damage patterns with accuracies ranging between 85% and 95% (Yang *et al.*, 2021). While trailing DL in sophistication, ML paradigms demonstrate commendable accuracies within the 70% to 90% range, contingent upon nuanced factors such as feature engineering and model selection (Naito *et al.*, 2020). In contrast, traditional classification approaches, such as SVM and decision trees, tend to exhibit relatively diminished accuracies, typically hovering between 60% and 80%, owing to their dependence on manually engineered features and less nuanced algorithmic frameworks. It's important to emphasize that these percentages can vary based on dataset size, quality, and preprocessing methods. This prompts ongoing exploration to refine damage detection techniques in structural assessment.



## CHAPTER 3

### MATERIALS AND METHODS

#### 3.1 Study Area

The city of Hatay has been selected for this study, focusing specifically on the Antakya district center (Figure 3.1). Antakya presents a unique mix of modern and culturally significant buildings, making it an ideal location for studying the seismic resilience of diverse structures. The study encompasses an area of 7973 m<sup>2</sup>, encompassing 8298 buildings. The General Directorate of Geographical Information Systems classified these buildings into four categories with the help of the site emergency response team AFAD (GD of GIS, 2023): No Damage, Heavily Damaged, Urgent Demolish, and Collapsed. The choice of Antakya district center as the study area ensures a comprehensive examination of the seismic impact on a wide range of structures, providing valuable insights for earthquake damage assessment and strategic urban planning in Hatay.

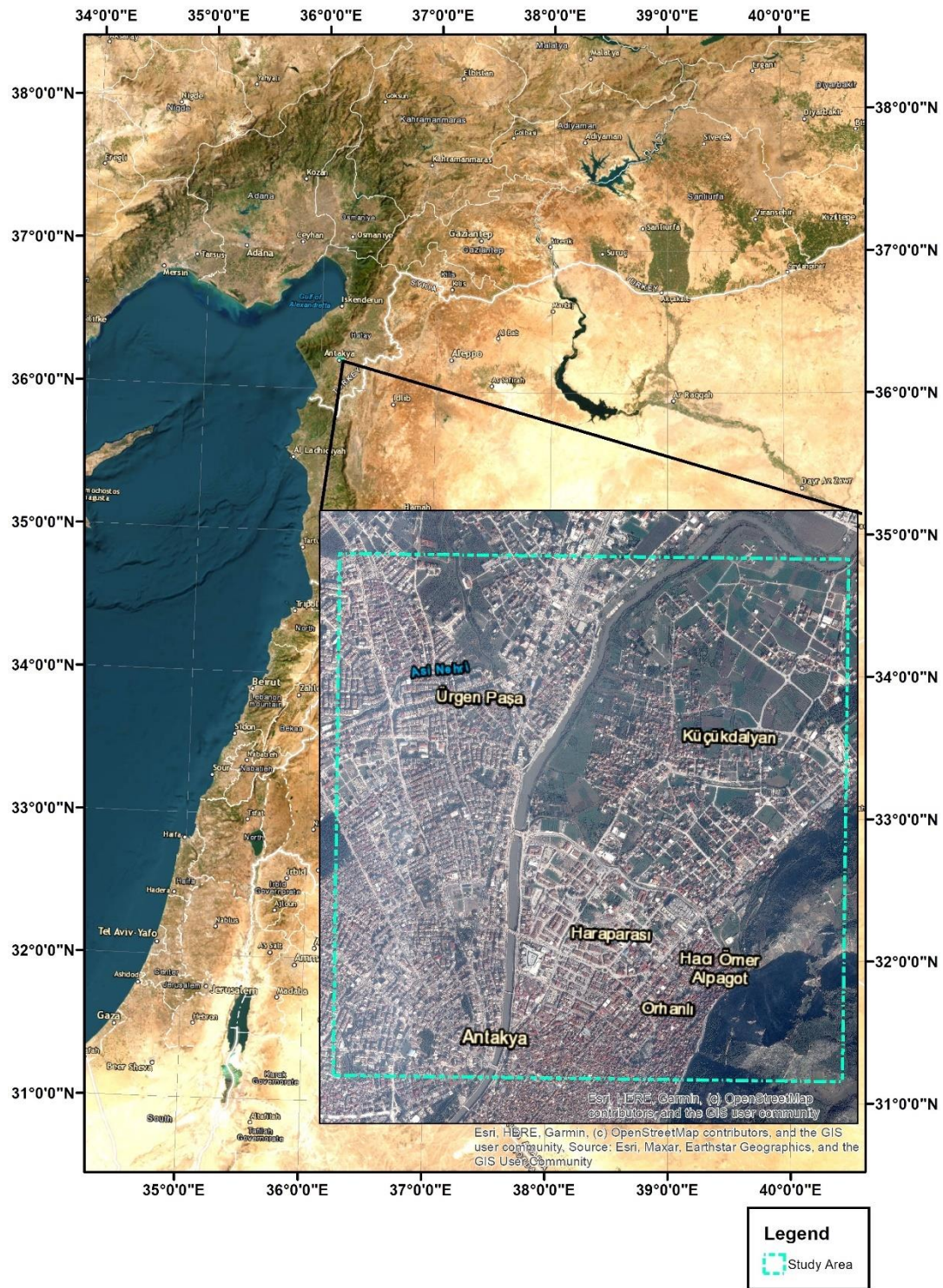


Figure 3.1. The location of the study area.

## **3.2 The Dataset**

### **3.2.1 Satellite Data**

The satellite data for this study is the Pleiades PHR1B 50 cm Imagery, generously provided by Airbus Pleiades (Airbus, 2023), and further processed by Implementation and Research Center for Satellite Communication and Remote Sensing (UHUZAM) (UHUZAM, 2023).

Notably, the satellite imagery procured from UHUZAM is not confined to Hatay/Antakya alone; rather, it encompasses every city affected by the earthquake. This expansive coverage is a deliberate effort to create a comprehensive dataset for academic use in scientific studies related to earthquake impact and building damage assessment.

All images and their corresponding footprints have been meticulously cataloged and organized. Figure 3.2 visually represents the images and their geographic footprints provided by Pleiades and UHUZAM (UHUZAM, 2023)

The acquired satellite data have been systematically cataloged, detailing the coverage areas and corresponding spatial resolutions. Table 3.1 provides a comprehensive overview of this crucial information, offering insights into the geographic extent covered by the satellite imagery and the spatial resolution achieved. In Table 3.2, the details of the image dataset (i.e., wavelength and native resolution properties, as well as separate image band compositions) are provided.

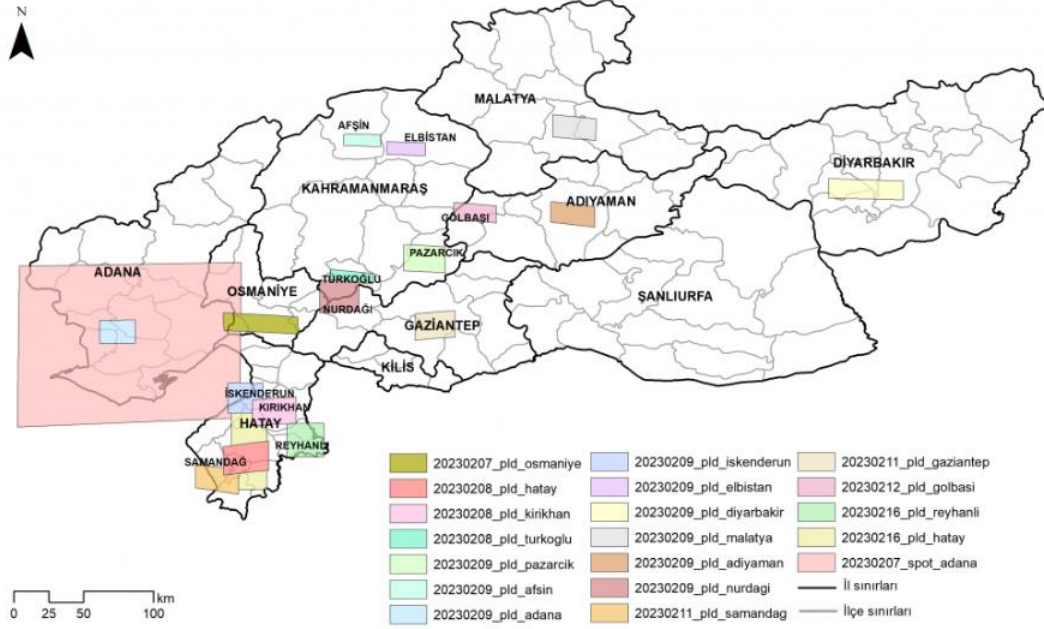


Figure 3.2. Image footprints provided by Pleiades and UHUZAM (UHUZAM, 2023)

Table 3.1. Time, date of acquisition, image location, sensor, spatial resolution, and band composition of the satellite imagery.

Date of Acquisition	Image Location	Sensor	Spatial Resolution	Band Composite
16.02.2023	Hatay	Pleiades-1B	0.5 Meter	RGB+NIR
16.02.2023	Hatay	Pleiades-1A	0.5 Meter	RGB+NIR
16.02.2023	Hatay	Pleiades-1A	0.5 Meter	RGB+NIR
08.02.2023	Hatay	Pleiades-1A	0.5 Meter	RGB+NIR
09.02.2023	Hatay	Pleiades-1B	0.5 Meter	RGB+NIR

Table 3.2 Spectral bands of Pleiades – 1B (Laurent Coeurdevey and Karine Fernandez, 2012)

Pleiades – 1B Bands	Native Resolution (m)	Wavelength Interval ( $\mu\text{m}$ )
Panchromatic Band (1)	0.5	0.47-0.83
Blue Band (2)	2	0.43-0.55
Green Band (3)	2	0.50-0.62
Red Band (4)	2	0.59-0.71
NIR Band (5)	2	0.74-0.94

Two satellite images acquired on 8 Feb 2023 were used for the study area. These satellite images, which consist of four-band images, use red, green, and blue bands and NIR (Near Infrared) channels. The image can be seen in Figure 3.3.

In this study, PCA alongside directional texture analysis is employed to enhance the detection of damaged building features within satellite or aerial imagery (Angulo-Saucedo *et al.*, 2022). PCA is a dimensionality reduction technique that preserves essential information while reducing redundancy. PCA image bands are derived by transforming the original spectral bands of the imagery into a new set of orthogonal variables known as principal components. Each principal component represents a linear combination of the original bands, ordered by their importance in explaining the variance within the dataset. These PCA image bands effectively condense the information contained in the original spectral bands, allowing for more efficient analysis and interpretation of the imagery. Integrating PCA image bands into the analysis aims to capture the underlying patterns and structures in the imagery related to damaged building features, thereby improving the accuracy and effectiveness of the ML algorithms.

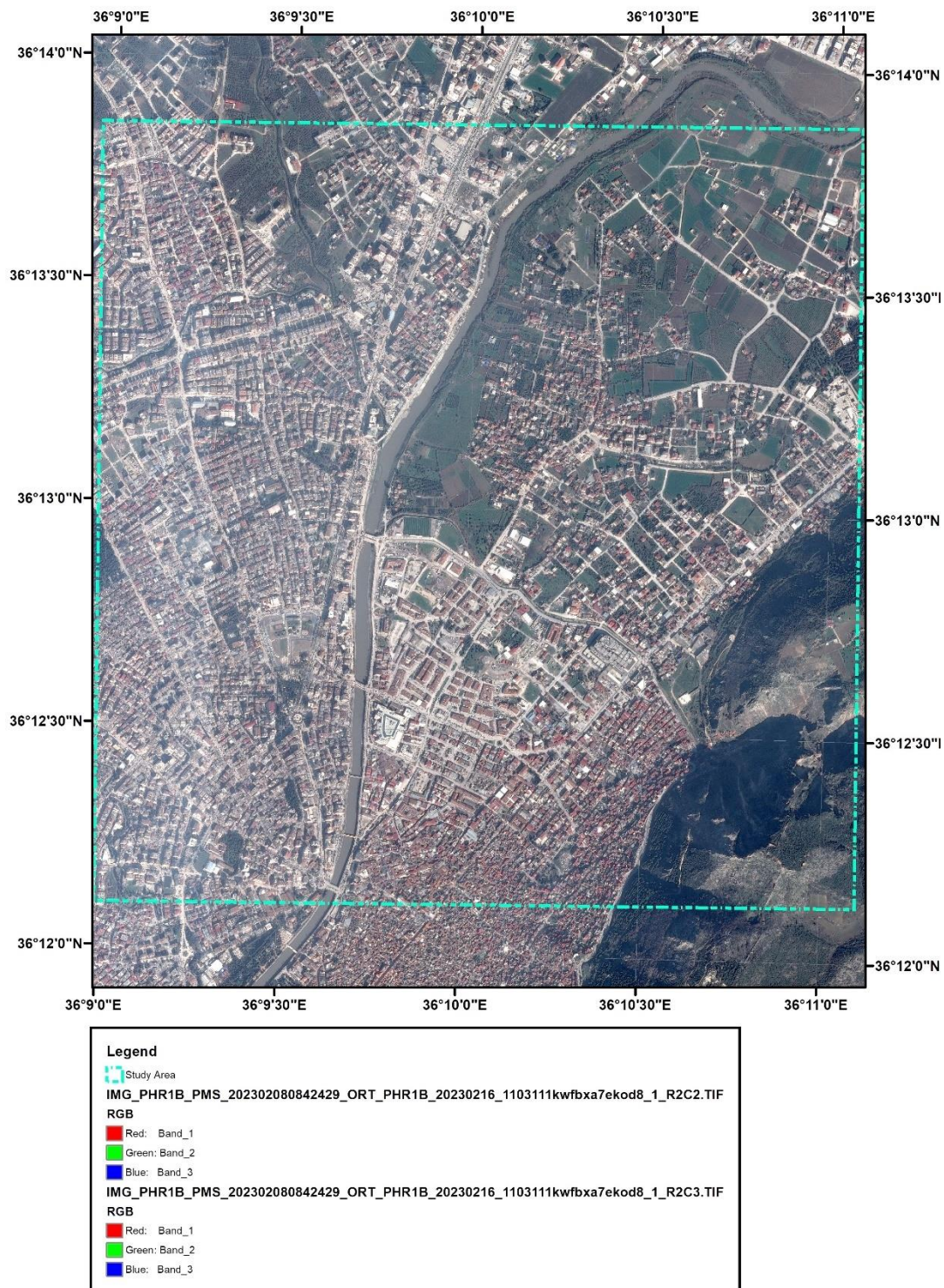


Figure 3.3. Study area with pan-sharpened 50 cm Pleiades 1B image (Satellite Image Acquisition Date: 08.02.2023).

### **3.2.2 Vector Data**

In support of the comprehensive analysis undertaken in this study, diverse vector datasets have been instrumental in enhancing the geospatial understanding of the chosen study area. These datasets encompass a variety of spatial information, notably building footprints integrated with the metadata of ground truth (GD of GIS, 2023). Sourced meticulously from reputable authorities, including the Republic of Türkiye Ministry of Environment, Urbanization, and Climate Change Directorate General of Geographic Information Systems and the United States Geological Survey (USGS), these datasets ensure the utmost reliability and accuracy in representing the study area.

To ensure broad accessibility and compatibility with commonly used GIS software, the vector datasets have been formatted in the standard Shapefile (\*.shp) and adhere to the World Geodetic System (WGS 84). Researchers keen on utilizing these datasets for further analysis are encouraged to consult the accompanying documentation for detailed insights into attributes, data sources, and any preprocessing steps that have been undertaken. The visual representation of the project area and building footprints are depicted in Figure 3.4, offering a tangible illustration of the valuable ground truth data obtained through these collaborative initiatives.

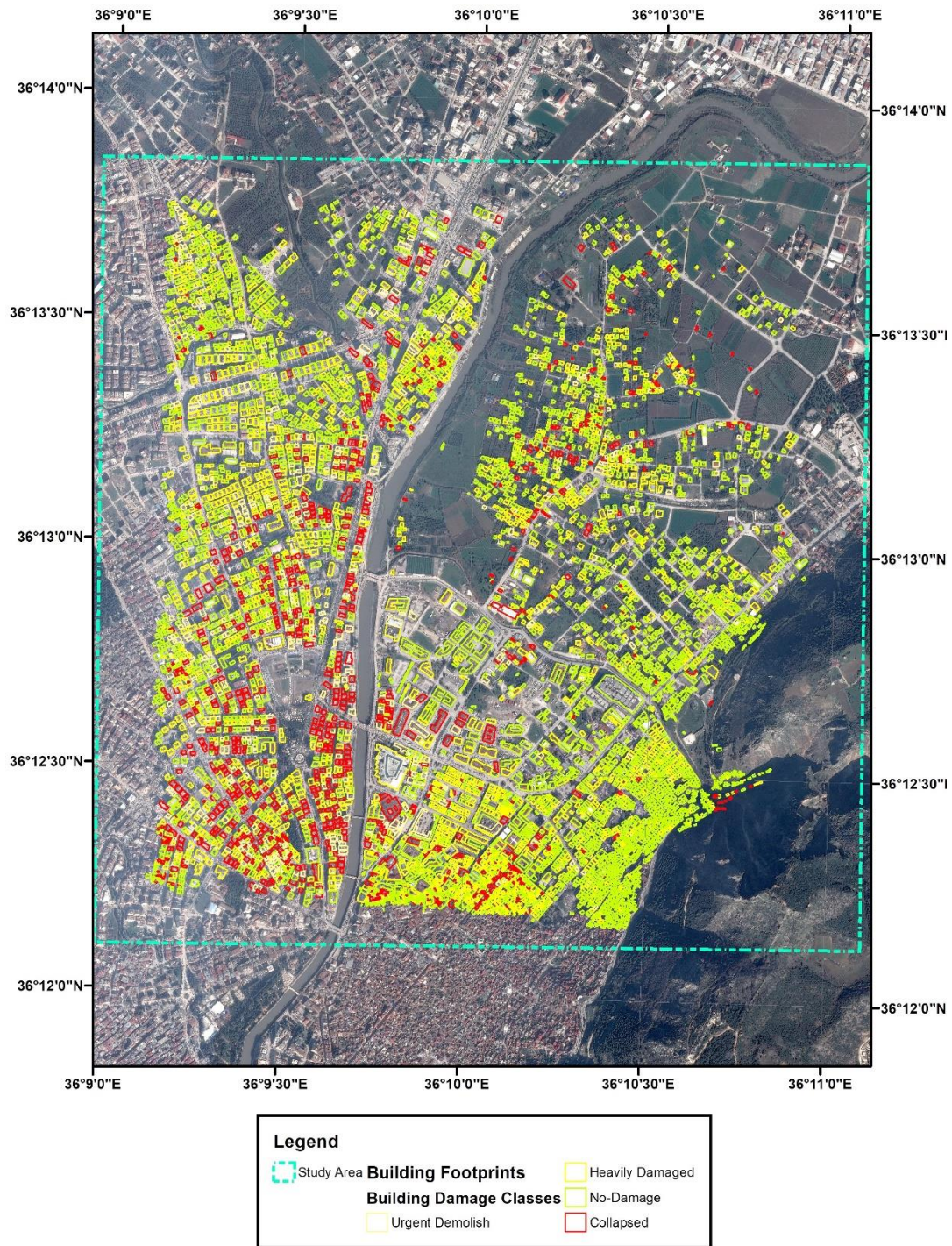


Figure 3.4. Vector data acquired from the Directorate of Geographical Information Systems, ATLAS Project with classified building footprints (GD of GIS, 2023).



İL BANK produces the data presented in Figure 3.4 in compliance with the Regulation on the Production of Large-Scale Maps and Map Information (Regulation on the Production of Large-Scale Maps and Map Information, 2005). Produced through photogrammetric methods and workflows, the vector data bears the production date of 29.03.2021 (GD of GIS, 2023).

The metadata accompanying the building footprint data contains vital information such as the current condition of each building post-earthquake, acquisition date of data, owner's address, detailed cadaster, and generic land ownership details. This extensive metadata forms a robust basis for comprehending the status of each building, serving as the cornerstone for the building classification in this study. Leveraging current building conditions enhances the relevance and accuracy of the building classifications, thereby enriching the seismic impact assessment with detailed insights.

Within the study area, 8,298 building footprints have been identified after the earthquake and classified by the General Directorate of Geographical Information Systems with the help of the site emergency response team AFAD (GD of GIS, 2023). The distribution of these classifications is as follows: 589 buildings are labeled as "Urgent Demolish" 2,135 are categorized as "Heavily Damaged" 4,484 as "No Damage" and 1,090 are marked as "Collapsed". The data is presented in Table 3.3, showcasing the respective percentages of each class.

Post-earthquake satellite images obtained from UHUZAM (UHUZAM, 2023) and the building footprints obtained from the General Directorate of Geographical Information Systems (GD of GIS, 2023) are the fundamental inputs in this study.

Table 3.3. Building condition percentages of the dataset.

Class	Object Count	Percentage
Total Data	8298	100.00%
Urgent Demolish	589	7.10%
Heavily Damaged	2135	25.73%
No Damage	4484	54.04%
Collapsed	1090	13.14%

Determining Ground Sample Distance (GSD) in satellite imagery for a given study area entails a meticulous process reliant on identifiable objects and building outlines. This procedure involves assessing the pixel size, termed pixel size, within the satellite imagery. Additionally, distances to recognized landmarks are measured and cross-referenced with ground truth data to validate accuracy. The landmarks used are shown in Figure 3.5. This verification step is a crucial quality assurance measure to ensure that the selected satellite images meet the requisite resolution and fidelity standards for subsequent analysis.

References to authoritative sources underpin this process, notably drawing from the Pleiades user manual authored by Laurent Coeurdevey and Karine Fernandez in 2012 (Laurent Coeurdevey and Karine Fernandez, 2012). This manual furnishes essential metadata, including GSD information, crucial for satellite image analysis. Furthermore, supplementary data from UHUZAM, an entity presumed to be associated with satellite imagery processing, enhances the validation process.

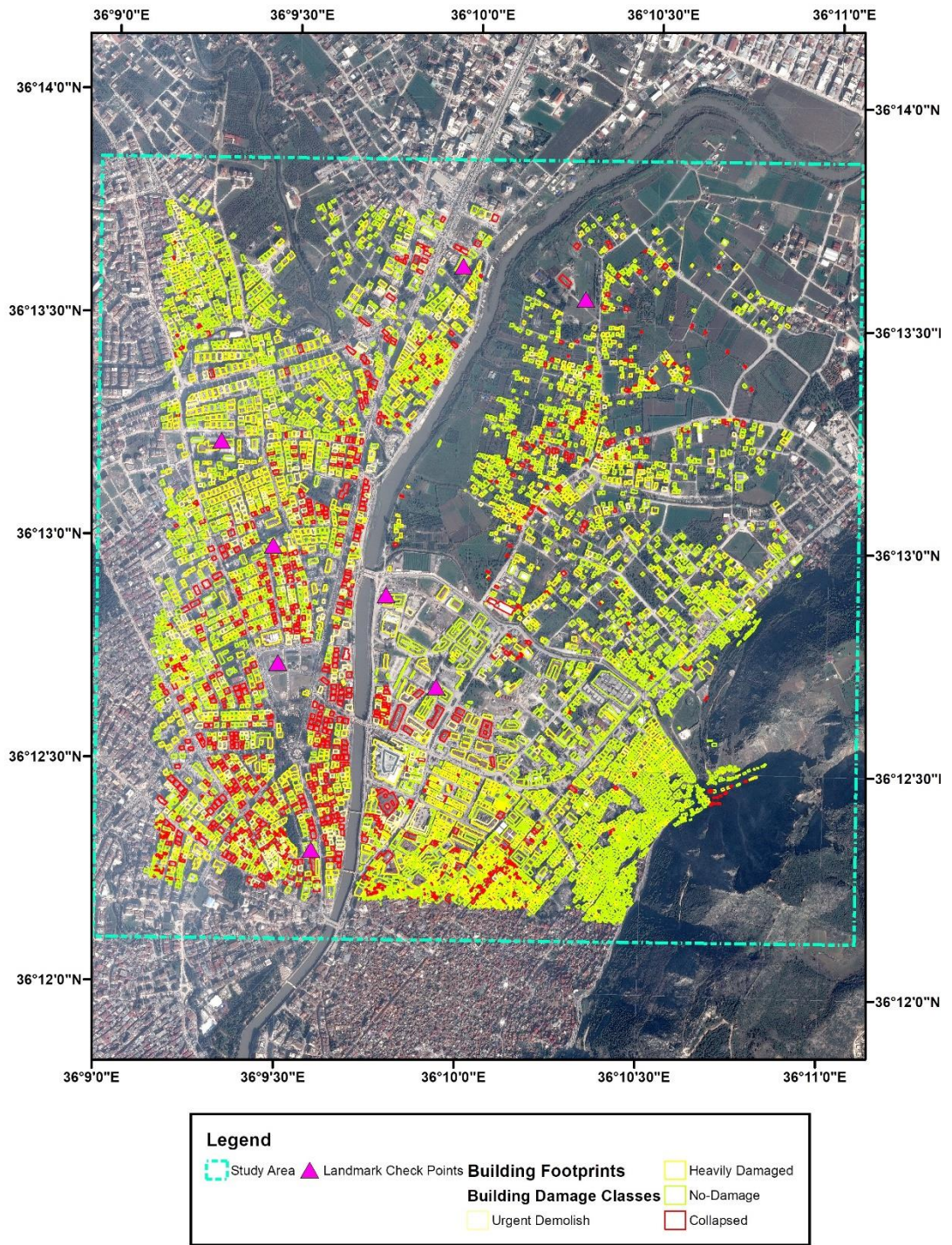


Figure 3.5. Landmark checkpoints.

Since the Pleiades image is off-nadir, the building footprints do not match the buildings in the image; the vector data manually to be aligned with the buildings in the satellite image. Figure 3.6 and Figure 3.7 show the pre and post-edited vector data, respectively.



Figure 3.6. Pre-edit vector data where the vector data is not aligned with the buildings from the satellite images.



Figure 3.7. Post-edited vector data where the satellite image and vector data building footprints match.

This process was performed to ensure the building footprints from the vector data directly covering the area of the building to be able to represent every building's footprint for the statistics analysis.

Editing vector data involves meticulous manual adjustments within each neighborhood, ensuring a uniform and consistent approach to alterations. Specifically, editors utilize the facade of undamaged buildings within the neighborhood as reference points to guide the editing process. By carefully comparing and aligning alterations with the features of undamaged building facades, editors ensure that changes are applied uniformly across all building components. This method maintains spatial consistency within the neighborhood and minimizes edited data discrepancies throughout the study area.

Different studies have shown that acquiring the building's footprint or cover area can be derived from the building's shadows (Dikmen, 2014; Turker and San, 2004).

The primary objective of the research is to leverage the intricate relationship between the shadows cast by damaged buildings and their surroundings, particularly in the context of multi-story structures. Specifically, the study aims to discern distinctive patterns and cues within building shadows, exploiting them as vital indicators for identifying and delineating damaged polygons. Notably, collapsed buildings are anticipated to exhibit an absence of shadow due to their structural failure, contrasting with intact structures that maintain discernible shadow features. This approach underscores the significance of shadow analysis as a pivotal tool in detecting and characterizing damaged building footprints. As highlighted by previous research by (Turker and San, 2004), the resulting polygonal representations of buildings, derived from shadow cues and stored within Geographic Information Systems (GIS) as vector polygons, serve as invaluable assets for comprehensive disaster management and urban planning endeavors. Unfortunately, in this case, it can be seen that while some areas have a low density in means of buildings where the shadows can be extracted, most of the study areas for this research do not meet the feature criteria for the building features. Therefore, extracting the building's footprint vectors from the shadows is not feasible, as mentioned in the cited articles. This situation is presented in Figure 3.8.



Figure 3.8. Part of the study area shows the density of buildings where shadows cannot be detected individually.

After editing the vector data, top-of-atmosphere (TOA) Reflectance calculations were performed on the satellite images, guided by the Pléiades Satellite Image Manual (Laurent Coeurdevey and Karine Fernandez, 2012). By converting DN to reflectance, we can correct for variations in factors such as solar angle, atmospheric conditions, and sensor calibration, affecting the DN values captured by the satellite sensor.

### 3.2.3 Ancillary Data

The collaboration between the Earth Observatory of Singapore's RS Lab, NASA's Jet Propulsion Laboratory, and Caltech's Advanced Rapid Imaging and Analysis (ARIA) team underpins the derivation of this map. The data originates from Japan's Advanced Land Observing Satellite-2 (ALOS-2) satellite, collected on February 8, 2023. Equipped with a synthetic aperture radar, this satellite emits pulses of

microwaves toward the Earth's surface, recording reflections to map the landscape, including buildings. By comparing the February 8 data with observations predating the earthquake on April 7, 2021, and April 6, 2022, scientists meticulously tracked changes, discerning and identifying areas likely affected by the seismic event (NASA/JPL-Caltech/USGS/JAXA/Earth Observatory of Singapore – Remote Sensing Lab, 2023). The data is represented in Figure 3.9.

The damage proxy map is found to be a rough input to be considered as an input for the ML algorithms; therefore, it is used for a comparison of the resulting damage classification maps. The comparison is to check the data classified as corresponding to the damage proxies provided. This comparison is presented in section 4.1.7.

The color variations from yellow to red indicate an increasingly more significant change in the area covered by the pixel; each pixel in the damage proximity map has a corresponding dimension of about 30m (Yun *et al.*, 2015).

### **3.3 Methodology**

The methodology section of this thesis presents a systematic framework that guides the investigation into earthquake damage assessment using RS imagery and ML techniques. This section is a comprehensive guide detailing the research design, data collection methods, data analysis procedures, and validation techniques employed to achieve the study's objectives. The employed methodology is illustrated in Figure 3.10.



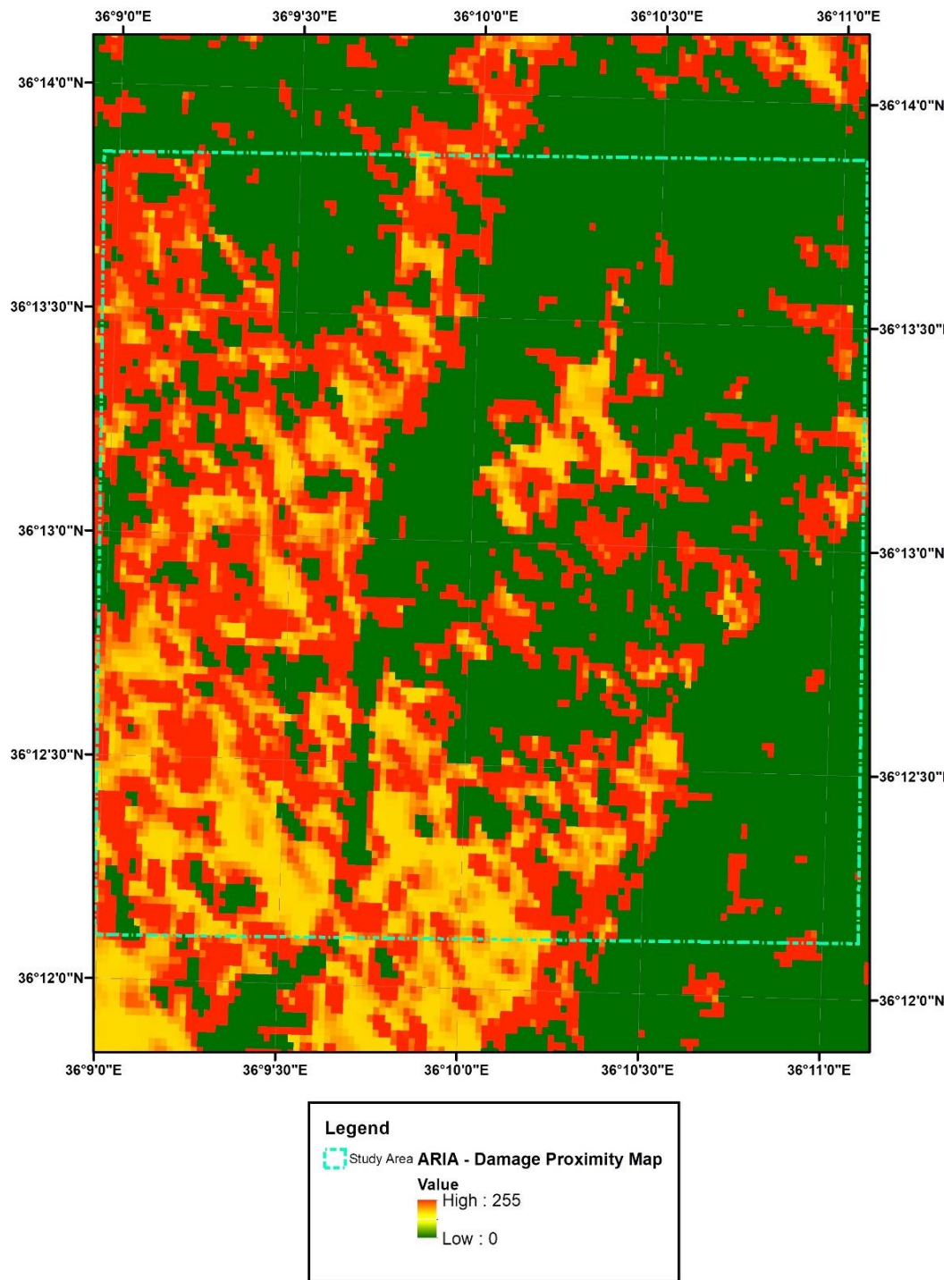


Figure 3.9. NASA/ARIA program DPM for the study area (NASA/JPL-Caltech/USGS/JAXA/Earth Observatory of Singapore – Remote Sensing Lab, 2023)

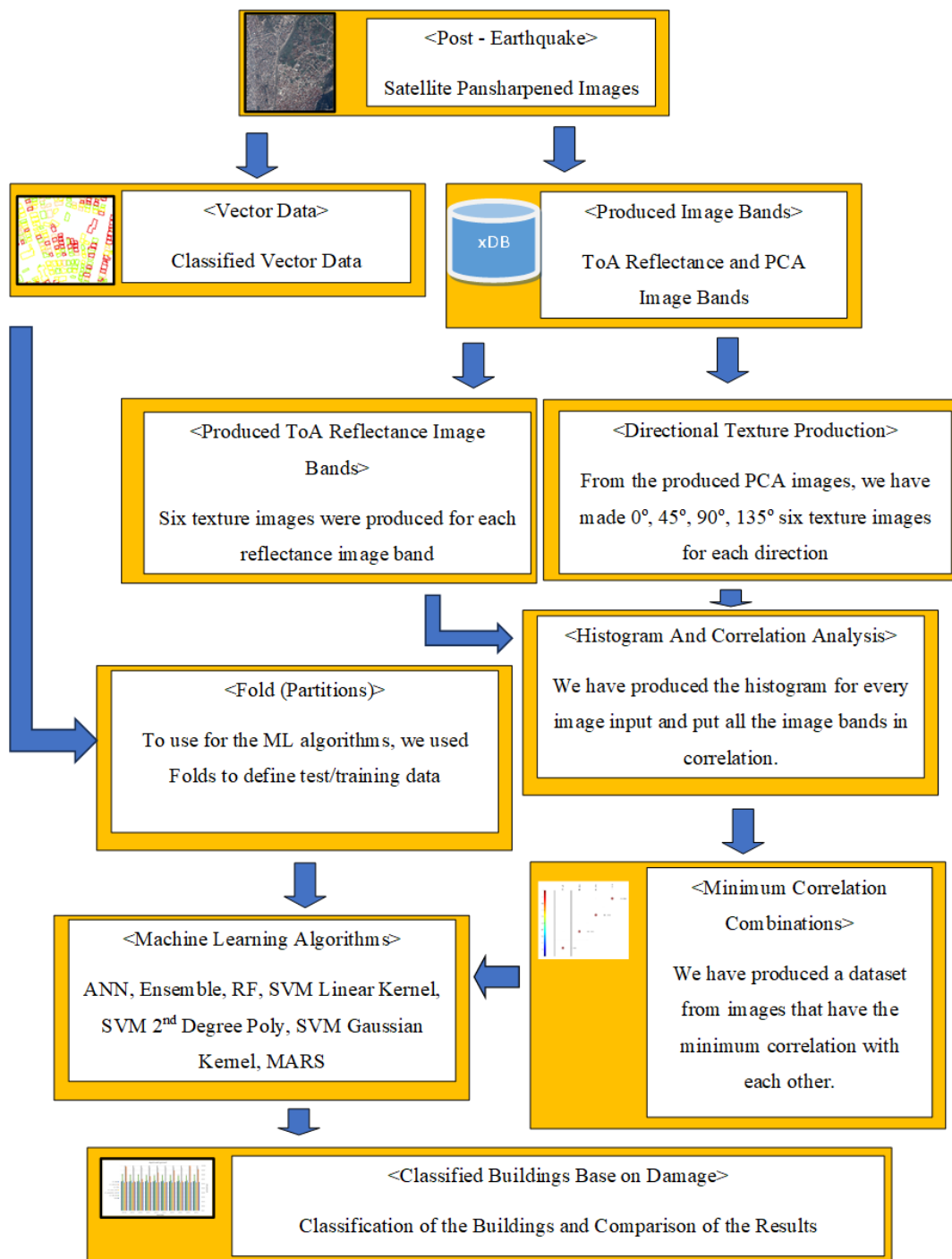


Figure 3.10. Flowchart of the methodology.

### 3.3.1 Satellite Image Band Reflectance Calculations

The Pleiades 1B satellite provides imagery with a spatial resolution of 50 centimeters, allowing for detailed and precise geospatial analysis (Laurent Coeurdevey and Karine Fernandez, 2012). The instrument has red, green, and blue (visible bands) and near-infrared (NIR) bands, enabling a multi-spectral analysis that captures diverse information about the Earth's surface.

Each spectral band of the imagery is transformed from raw DNs to TOA reflectance. This process involves normalizing pixel values to account for atmospheric variations and sensor characteristics, ensuring that the resulting data accurately represents surface properties. Reflectance values provide a standardized measure of how much incident light the Earth's surface reflects, enabling more consistent and comparable analysis across different scenes and conditions (Chander *et al.*, 2009).

Airbus's Pleiades user manual (Laurent Coeurdevey and Karine Fernandez, 2012) supplies two distinct equations instrumental in this conversion process, addressing specific sensor characteristics and atmospheric conditions. As outlined in the user manual, this conversion was achieved using Equation 3.1, where  $L_b(p)$  represents the TOA Radiance pixel value,  $X_{(p)}$  represents the DN of each pixel, and  $GAIN(b) - BIAS(b)$  is the absolute radiometric coefficients resulting from sensor calibration performed during the satellite's lifecycle (Laurent Coeurdevey and Karine Fernandez, 2012). The equation ensures an accurate representation of radiance at the top of the Earth's atmosphere for each band of the Pleiades image dataset.

$$L_b(p) = \frac{X_{(p)}}{GAIN(b)} + BIAS(b) \quad (\text{Eq 3.1})$$

Then, those TOA radiances were converted to TOA reflectance values using Equation 3.2:

$$\rho_b(p) = \frac{\pi \cdot L_b(p)}{E_0(b) \cdot d \cdot \cos(\theta_s)} \quad (\text{Eq 3.2})$$

In Equation 3.2,  $\rho_b(p)$  denotes the TOA Reflectance pixel value. In contrast,  $L_b(p)$  represents the previously derived TOA Radiance value. Additionally, the equation introduces several essential parameters:  $E_0(b)$ , the solar spectral irradiance in the considered band;  $\theta_s$ , the Sun zenith angle;  $d$ , a correction coefficient concerning the mean Earth-Sun distance; and  $\pi$ , the mathematical constant Pi.

A detailed analysis of reflectance histograms has been conducted to gain a comprehensive understanding of the spectral characteristics of the Pleiades 1B satellite imagery concerning different building label classes. This step is instrumental in providing a preliminary insight into the distribution of pixel reflectance values across various image bands for each building category, namely "Urgent Demolish" "Heavily Damaged" "No Damage" and "Collapsed"

A histogram table has been computed for every image band, encapsulating the frequency distribution of pixel reflectance values within the specified building label classes. The histograms serve as a valuable tool for elucidating the general trends and variations in reflectance exhibited by the diverse building footprints. This preliminary analysis is essential for identifying potential patterns or discrepancies in the spectral response of different building types.

Furthermore, a mean pixel reflectance value has been calculated for each building footprint within the designated areas. This mean value serves as a representative measure of the overall reflectance characteristics associated with each building label class. By aggregating this information, histogram charts have been meticulously generated for every building category, offering a graphical representation that facilitates a more intuitive and accessible understanding of the reflectance patterns.

The reflective analysis of Pleiades 1B satellite imagery through histogram charts is a precursor to the subsequent steps in the methodology. This information

contributes to the comprehensive feature set utilized in the ML models, enhancing their ability to discern and classify building conditions accurately. The graphical representation of reflectance histograms not only aids in the initial exploration of spectral patterns but also forms an integral part of the broader geospatial analysis framework employed in earthquake damage assessment.

After aligning the footprints, obtaining initial insights from the data on a building-by-building basis is paramount. Comprehensive statistical measures are essential for each building. Utilizing the principal component analysis, bands derived from the reflectance satellite imagery were subsequently transformed into texture bands. Minimum, maximum, mean, and standard deviation values were calculated within each building's footprint area.

### **3.3.2 Satellite Image Texture Dataset Production**

Developing an exhaustive satellite image texture dataset is pivotal in advancing earthquake damage identification and building detection methodologies. The methodology integrates texture features (Haralick, 1979), which is crucial for earthquake damage assessment and building detection. The dataset comprises six distinct texture layers: contrast, correlation, dissimilarity, entropy, homogeneity, and second moment. Contrast and dissimilarity layers capture spatial variations in pixel intensities, aiding in discerning changes in building structures post-earthquake. Entropy, a measure of randomness in pixel values, characterizes texture patterns related to earthquake-induced damages. Homogeneity and second-moment layers provide insights into the uniformity and orderliness of pixel distributions, aiding in the differentiation of damaged and undamaged areas. The correlation layer captures relationships between pixel values, facilitating nuanced texture representation for damage detection.

The methodology integrates texture features crucial for earthquake damage assessment and building detection. The dataset comprises eight distinct texture layers:

The contrast and dissimilarity layers are vital for capturing spatial variations in pixel intensities, which prove instrumental in distinguishing changes in building structures post-earthquake. The contrast Equation 3.3 computes the absolute differences between pixel values, emphasizing the range of intensity variations. On the other hand, dissimilarity also considers absolute differences but accentuates changes between neighboring pixels. Both equations utilize  $P(i, j)$  to represent the probability of occurrence of pixel intensity values  $i$  and  $j$ , where  $N$  and  $M$  represent the image's dimensions.

$$\sum_{i=1}^N \sum_{j=1}^M |i - j| P(i, j) \quad (\text{Eq 3.3})$$

Entropy Equation 3.4 measures randomness in pixel values, providing insights into the texture patterns related to earthquake-induced damages. It is calculated by summing the products of pixel probabilities and their natural logarithms. The log denotes the natural logarithm, and  $P(i, j)$  represents the probability of pixel intensity values  $i$  and  $j$ .

$$\sum_{i=1}^N \sum_{j=1}^M P(i, j) \log(P(i, j)) \quad (\text{Eq 3.4})$$

Homogeneity Equation 3.5 and second moment Equation 3.6 layers offer insights into the uniformity and orderliness of pixel distributions, aiding in the differentiation of damaged and undamaged areas. Homogeneity calculates the inverse of the absolute differences between pixel values, emphasizing uniform regions, while the second moment captures the overall orderliness of pixel distributions.

$$\sum_{i=1}^N \sum_{j=1}^M \frac{1}{1 + |i - j|} P(i, j) \quad (\text{Eq 3.5})$$

$$\sum_{i=1}^N \sum_{j=1}^M P(i, j)^2 \quad (\text{Eq 3.6})$$

The correlation layer Equation 3.7 captures relationships between pixel values, facilitating nuanced texture representation for damage detection. It considers the covariance of pixel values normalized by the product of their standard deviations.  $\mu_i$  and  $\sigma_i$  represent the mean and standard deviation of pixel intensities in the  $i$ -th row.

$$\frac{\sum_{i=1}^N \sum_{j=1}^M (i - \mu_i)(j - \mu_j)P(i, j)}{\sigma_i \sigma_j} \quad (\text{Eq 3.7})$$

Structures impacted by seismic activity are classified into No Damage and Damaged. The classification framework provided by the Republic of Turkey Ministry of Environment, Urbanization, and Climate Change's Directorate General of Geographic Information Systems (GD of GIS, 2023) establishes the ground truth. This authoritative classification forms the basis for the assessments, ensuring a reliable reference for evaluating the structural aftermath of seismic events.

### 3.3.3 Input Clusters of the Dataset

For this study, 60 different statistical variables were generated to serve as inputs for ML algorithms. These variables were derived from 15 image bands, collectively called the "Ref\_Direct\_PCA\_minCor" dataset. Each statistical variable represents a specific aspect of building footprints, including minimum, maximum, standard deviation, and mean values obtained from the vector data footprints.

To organize the input data effectively, various combinations of image bands, each serving a distinct purpose, are employed:

- The combinations of reflectance bands from the Pleiades images (bands 1-4) are labeled "ref\_band."
- PCA bands, generated from the reflectance image bands 1 to 4, are denoted as "PCA" with corresponding band numbers.

- Texture image bands, derived from reflectance images, encompassing contrast, correlation, dissimilarity, entropy, homogeneity, and the angular second moment, are designated "ref\_tex\_band."
- Similarly, texture image bands derived from PCA images follow the same parameters and are labeled "pca\_tex\_band."

Table 3.4 provides detailed definitions for each composition mentioned above, clarifying the components utilized in the input data preparation process. Furthermore, the subsequent sections of this chapter delve into the methods employed to acquire these specific inputs. It is important to note that all spectral bands utilized in the input combinations have undergone rigorous atmospheric and topographic corrections to ensure data accuracy and consistency.

Table 3.4. Main input combinations.

Combination Name	<i>Content of the Combination</i>
ref_band	Top of Atmosphere Reflectance of Native Image Bands 1-4
PCA	First principal component from PCA derived from ref_band 1-4
ref_tex_band	Six texture images produced from ref_band 1-4
pca_tex_band	Six texture images produced from PCA band 1-4

The initial combinations are assessed using histogram and correlation analysis from the histogram and correlation analyses. Different combinations of image bands were obtained, resulting in 44 to 60 different minimum, maximum, standard deviation, and mean values for 8,298 building footprints from the vector data. Directional texture bands were produced for the PCA image combination of 0, 45, 90, and 135 degrees. The additional combinations are detailed in Table 3.6.



Table 3.6. Additional input combinations.

Combination Name	<i>Content of the Combination</i>
Direct_PCA	Directionally produced texture image bands from first PCA combination from band 1- 4 with the directions of $0^0$ , $45^0$ , $90^0$ and $135^0$ for six texture metrics.
Direct_PCA_minCor	After conducting a correlation analysis of the "Direct_PCA" dataset, the dataset with minimal correlation was found.
Ref_Direct_PCA_minCor	Four band Reflectance image bands were added to the "Direct_PCA_minCor"

### 3.3.4 Building Classification

Building classification following a seismic event traditionally relies on the EMS98 scale (Grünthal, 1998), which categorizes buildings based on their structural integrity after hazardous natural disasters. However, obtaining ground truth classes from authorities can be challenging. The study aims to identify damaged buildings directly from post-earthquake imagery, recognizing the discrepancies between the observed ground truth and the nadir view determination of buildings. The traditional building damage classes can be seen in Figure 3.11.


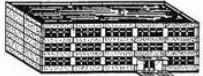



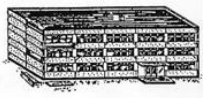




Classification of damage to masonry buildings		Classification of damage to buildings of reinforced concrete	
	<b>Grade 1: Negligible to slight damage</b> (no structural damage, slight non-structural damage) Hair-line cracks in very few walls. Fall of small pieces of plaster only. Fall of loose stones from upper parts of buildings in very few cases.		<b>Grade 1: Negligible to slight damage</b> (no structural damage, slight non-structural damage) Fine cracks in plaster over frame members or in walls at the base. Fine cracks in partitions and infills.
	<b>Grade 2: Moderate damage</b> (slight structural damage, moderate non-structural damage) Cracks in many walls. Fall of fairly large pieces of plaster. Partial collapse of chimneys.		<b>Grade 2: Moderate damage</b> (slight structural damage, moderate non-structural damage) Cracks in columns and beams of frames and in structural walls. Cracks in partition and infill walls; fall of brittle cladding and plaster. Falling mortar from the joints of wall panels.
	<b>Grade 3: Substantial to heavy damage</b> (moderate structural damage, heavy non-structural damage) Large and extensive cracks in most walls. Roof tiles detach. Chimneys fracture at the roof line; failure of individual non-structural elements (partitions, gable walls).		<b>Grade 3: Substantial to heavy damage</b> (moderate structural damage, heavy non-structural damage) Cracks in columns and beam column joints of frames at the base and at joints of coupled walls. Spalling of concrete cover, buckling of reinforced rods. Large cracks in partition and infill walls, failure of individual infill panels.
	<b>Grade 4: Very heavy damage</b> (heavy structural damage, very heavy non-structural damage) Serious failure of walls; partial structural failure of roofs and floors.		<b>Grade 4: Very heavy damage</b> (heavy structural damage, very heavy non-structural damage) Large cracks in structural elements with compression failure of concrete and fracture of rebars; bond failure of beam reinforced bars; tilting of columns. Collapse of a few columns or of a single upper floor.
	<b>Grade 5: Destruction</b> (very heavy structural damage) Total or near total collapse.		<b>Grade 5: Destruction</b> (very heavy structural damage) Collapse of ground floor or parts (e. g. wings) of buildings.

Figure 3.11. Building classification with EMS98 (Grünthal, 1998).

The research attempted to classify damaged buildings into distinct categories based on varying damage severity levels, utilizing imagery with a ground sampling distance (GSD) of 50 cm. However, the efforts encountered challenges in effectively discerning specific categories, explicitly distinguishing between EMS98 G1 from G2 and G3 and G4 from G5.

Despite employing higher-resolution imagery, the differences in reflectance properties between these categories needed more significant differences to facilitate reliable classification. Studies show the need for an up-to-date classification process using RS, where the image's ground sampling distance plays a higher role in the classification process (Cotrufo *et al.*, 2018). Classification for building damage for the sole purpose of RS can be seen in Figure 3.12.























Damage Grade	Description of the Related Damage	Post-Event Imagery and Brief Interpretation Guidelines			
		Pre-Event Aerial (0.1m < GSD < 0.3m)	Post-Event Satellite (0.3m < GSD < 0.5m)	Post-Event Aerial (0.1m < GSD < 0.3m)	Post-Event UAV (GSD < 0.1m)
<b>Destroyed</b>	Total Collapse: collapse of part of the building (>50%); Building structure not distinguishable (The Walls have been destroyed or collapsed).				
	<p><b>Major:</b> Partial collapse of the roof; severe failure of the walls.</p> <p><b>Damaged</b> "Major" and "Minor" shall be used only when Aerial / UAV Imagery (GSD &lt; 0.3m) is available—only Damaged Grade when using Satellite Imagery.</p>		<p>Only use the "Damaged" grade when using satellite imagery.</p> 		
<b>Damaged</b>	<p><b>Minor:</b> The roof remains intact mainly but presents partial damage.</p>		<p>Black Spots on the rooftop suggest the collapse of part of the roof.</p> 		
			<p>The Structure appears intact—large debris deposited on the ground.</p> 		
<b>Possibly Damaged</b>	Uncertain interpretation due to image quality. Building surrounded by damaged/destroyed buildings.				
	<b>No Visible Damage</b>	The Structure appears to have complete structural integrity: The walls remain standing, and the roof is visually undamaged.			

Figure 3.12. The proposed building damage scale is tailored to RS vertical imagery (Cotrufo *et al.*, 2018).

Consequently, to mitigate these interpretational ambiguities and streamline the classification process, a decision was made to simplify the approach by consolidating the classification into two broader damaged building classes. This pragmatic adjustment underscores the necessity of critically assessing the suitability of data resolution and classification methodologies in RS analyses while adapting research strategies to accommodate inherent data limitations and achieve meaningful outcomes. The two-class approach compared to the EMS98 scale can be seen in Figure 3.13.

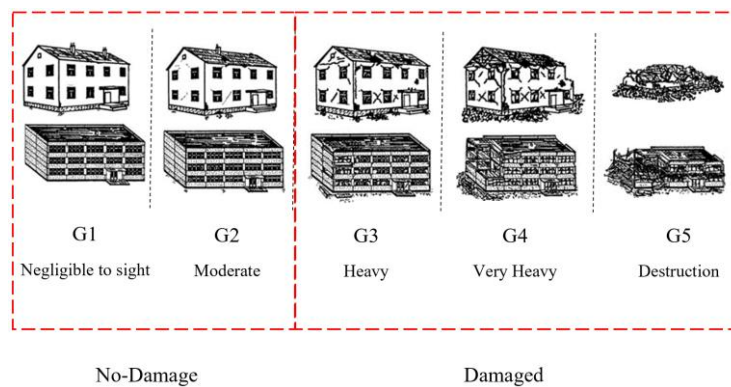


Figure 3.13. The study used two building classes compared to the EMS98 building scale.

### 3.4 Machine Learning Algorithms

#### 3.4.1 Support Vector Machine (SVM)

SVM has emerged as a powerful tool in ML, particularly for binary classification tasks. At its core, SVM operates by finding the optimal hyperplane that best separates the data points belonging to different classes while maximizing the margin, i.e., the distance between the hyperplane and the nearest data points from each class. This unique characteristic of SVM makes it robust and effective, especially in scenarios where the data is not linearly separable (Vapnik *et al.*, 1996).

One of the key strengths of SVM lies in its ability to utilize different kernel functions, which enable it to handle non-linear decision boundaries and capture complex patterns within the data. The linear kernel is the simplest among the commonly used kernels, operating in the original feature space and drawing linear decision boundaries. However, when the data is not linearly separable, more complex kernels come into play.

The second-order polynomial kernel is one example that maps the input features into a higher-dimensional space, allowing SVM to capture non-linear relationships through quadratic decision boundaries. This kernel is particularly useful when dealing with moderately complex datasets where linear separation is insufficient.

On the other hand, the Radial Basis Function (RBF) kernel is arguably the most widely used kernel in SVM. It transforms the input features into an infinite-dimensional space, where the decision boundary becomes non-linear and highly flexible. The RBF kernel is adept at capturing intricate patterns and can effectively handle datasets with complex structures and overlapping classes.

In summary, SVM with different kernels offers a flexible framework for binary classification, capable of handling a wide range of datasets with varying degrees of complexity. By intelligently selecting the appropriate kernel function based on the characteristics of the data, SVM can deliver superior performance in capturing intricate decision boundaries and accurately classifying data points.

### **3.4.2 Artificial Neural Networks (ANN)**

The Multi-Layer Perceptron (MLP) ANN stands as a cornerstone in the field of ML, revered for its ability to tackle complex tasks ranging from pattern recognition to classification. At its core, the MLP comprises multiple layers of interconnected neurons, each layer contributing to the network's ability to learn and make predictions (Basheer and Hajmeer, 2000). The architecture of an MLP typically consists of an input layer, one or more hidden layers, and an output layer.

The feedforward nature of the MLP denotes that data flows through the network in a unidirectional manner, starting from the input layer, passing through the hidden layers, and culminating in the output layer. During this process, each neuron receives inputs from the neurons in the preceding layer, computes a weighted sum of these inputs, and applies an activation function to produce an output. This sequential propagation of data allows the MLP to model complex relationships between inputs and outputs, making it a versatile tool for a myriad of tasks.

Central to the functioning of the MLP is the backpropagation algorithm, which serves as the cornerstone of training the network. Through a process of forward and backward passes, backpropagation enables the network to adjust the weights of its connections iteratively, minimizing the discrepancy between the predicted outputs and the actual targets. This iterative optimization process is driven by the gradient descent algorithm, wherein the network computes the gradient of a chosen loss function with respect to its parameters and adjusts these parameters in the direction that minimizes the loss.

The versatility and power of MLPs lie in their ability to learn complex, non-linear mappings from input to output. By virtue of their layered architecture and the iterative optimization afforded by backpropagation, MLPs excel in tasks requiring the extraction of intricate patterns and relationships from data. From image recognition to natural language processing, MLPs have found widespread application across various domains, cementing their status as a fundamental building block in the realm of artificial intelligence and ML.

### **3.4.3 Random Forest (RF)**

RF is a versatile and powerful machine-learning algorithm renowned for its robustness and accuracy, particularly in classification tasks. It belongs to the ensemble learning family, wherein multiple base models are combined to form a more robust and accurate predictor. The essence of RF lies in its ability to construct

a multitude of decision trees, each trained on a subset of the data and employing a random selection of features (Breiman, 2001).

At its core, RF generates an ensemble of decision trees, where each tree is constructed using a bootstrapped sample of the training data. This process introduces diversity among the trees, ensuring they capture different aspects of the underlying data distribution. Additionally, a random subset of features is considered for splitting at each node of the decision tree, further enhancing the diversity and reducing the correlation between individual trees.

During the prediction phase, the output of each decision tree is aggregated to yield the final prediction. This aggregation typically takes the form of a majority or weighted voting scheme for classification tasks, where the class with the most votes across all trees is assigned as the predicted class. This ensemble-based approach imbues RF with robustness to noise and outliers while mitigating overfitting, thereby improving generalization performance on unseen data.

One of the key strengths of RF lies in its ability to efficiently handle high-dimensional data, making it suitable for tasks involving large numbers of features. Additionally, its inherent parallelizability lends itself well to efficient training on modern computational architectures, enabling rapid experimentation and deployment.

#### **3.4.4 Multivariate Adaptive Regression Splines (MARS)**

The MARS algorithm is another method to detect damaged buildings from undamaged ones. In this subsection, an exploration of the MARS algorithm, introduced by (Friedman, 1991; Hastie *et al.*, 2009), will be undertaken.

MARS, which stands for Multivariate Adaptive Regression Splines, serves as a regression algorithm employed for modeling relationships between a response variable (in this case, the identification of damaged buildings) and predictor variables (the inputs or features represented by the Ref\_Direct\_PCA\_minCor

dataset). Unlike parametric methods, MARS does not assume a linear relationship between classes; instead, it models the relationship between the response and predictors using piecewise linear functions. Rather than assuming a single linear relationship across all the data, MARS divides the data into segments (knots) and fits a separate linear function to each segment. This approach allows for a divide-and-conquer strategy between the response and predictors.

MARS can effectively capture nonlinear changes between the response and predictors by utilizing piecewise linear functions and permitting different slopes between segments without abrupt changes. These piecewise linear functions are described in Equation 3.10 (Friedman, 1991).

$$\begin{aligned} [x - t]_+ &= \begin{cases} x - t, & \text{if } x > t, \\ 0, & \text{otherwise,} \end{cases} \\ [t - x]_+ &= \begin{cases} t - x, & \text{if } x < t, \\ 0, & \text{otherwise,} \end{cases} \end{aligned} \quad (\text{Eq. 3.10})$$

Within the equation,  $t$  can be seen as the location of a knot point ( $x, t \in \mathbb{R}$ ),  $[x - t]_+$  and  $[t - x]_+$  are the hinge functions of MARS. The demotion of “+” means only the positive values for the function are considered; this can be seen as represented in Figure 3.14.

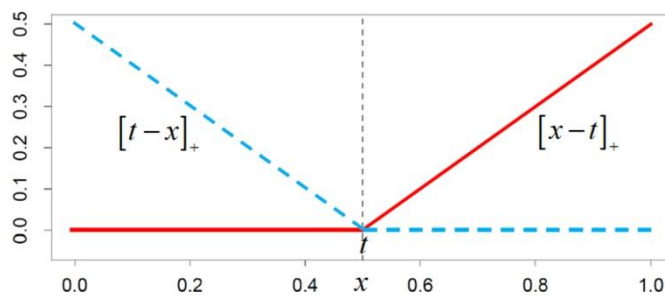


Figure 3.14. Symmetric boolean functions of MARS (Hastie *et al.*, 2009).

A nonparametric regression model that defines the relation between input and output variables is defined mathematically as;



$$Y = f(X) + \varepsilon \quad (\text{Eq. 3.11})$$

In equation 3.11,  $Y$  denotes the output, and  $X = (x_1, x_2, \dots, x_p)$ .  $T$  denotes input variables called predictors, and small epsilon is the observation error.

Hastie et al. (1994) and Hastie et al. (2009) state that the MARS ML algorithm can be modified to handle classification problems.

### 3.4.5 Ensemble Learning

Ensemble learning represents a paradigm shift in the field of ML, where the focus shifts from relying on a single model to leveraging the collective wisdom of multiple models. This approach has gained widespread popularity due to its ability to improve prediction accuracy, robustness, and generalization performance across a variety of classification tasks (Zhang and Ma, 2012).

At its core, ensemble learning involves combining the predictions of multiple base models, known as learners or weak learners, to form a more robust and accurate predictor, known as the ensemble. The fundamental principle behind ensemble learning is that by aggregating the predictions of diverse models, the ensemble can effectively mitigate individual model biases and errors, leading to improved overall performance.

One of the most popular ensemble learning techniques is bagging (Bootstrap Aggregating), exemplified by algorithms such as RF. Bagging involves training multiple instances of a base model on different subsets of the training data, typically sampled with replacement. This diversity in training data ensures that each base model learns various aspects of the underlying data distribution, thereby reducing variance and improving generalization performance.

Another widely used ensemble method is boosting, which aims to sequentially train a series of weak learners, with each subsequent learner focusing on the samples that were misclassified by the previous ones. Gradient Boosting Machines (GBMs),

including algorithms like AdaBoost and Gradient Boosting Trees (GBTs), are prominent examples of boosting algorithms. GBMs iteratively improve the ensemble's performance by fitting new models to the residuals of the previous ones, gradually reducing both bias and variance.

Moreover, there are ensemble techniques that focus on combining the predictions of base models rather than training them sequentially or in parallel. Examples include Stacking and Voting classifiers. Stacking involves training multiple base models and using a meta-model, often a simple linear model, to combine their predictions. Voting classifiers, on the other hand, aggregate predictions by majority voting or weighted averaging, where each base model's prediction carries a certain weight.

Ensemble learning encompasses a diverse array of techniques, each with its strengths and weaknesses, but united by the common goal of harnessing the collective knowledge of multiple models to improve predictive performance. By embracing the principles of diversity, randomness, and aggregation, ensemble learning has emerged as a cornerstone of modern ML, offering unparalleled flexibility and performance across a wide range of classification tasks.

Various ML algorithms were used in MATLAB to tackle the binary classification problem of detecting building damage. An Artificial Neural Network (ANN) model was trained with the “fitcnet” function, capturing complex data patterns. Additionally, Support Vector Machine (SVM) models with different kernels, including linear, polynomial (2nd degree), and Gaussian (i.e., RBF), were trained using the “fitsvm” function, which helps define decision boundaries in feature space effectively. RF models were trained using the “treebagger” function, adjusting parameters like the number of trees and minimum leaf size for optimal performance. Ensemble models were created using the “fitcensemble” function to harness the collective intelligence of diverse models. Through hyperparameter optimization, each algorithm's parameters were fine-tuned to enhance the models'

accuracy in distinguishing between damaged and undamaged buildings, laying a solid groundwork for practical predictive analytics.

In extending the exploration of ML algorithms for detecting building damage, MARS was integrated using the “earth” module within the R environment (Milborrow, 2011). With a focus on comprehensibility and interpretability, MARS offers a flexible framework for capturing complex relationships between predictors and the binary response variable denoting building damage. A grid search methodology was employed to systematically explore the parameter space and identify optimal settings, ensuring robustness and reliability in model fitting. Leveraging the earth module, model construction was initiated with specified tuning parameters, notably setting the maximum number of basis functions in the forward stage to 100 and the degree of interaction to 3. These parameters govern the complexity of the model, balancing the trade-off between model interpretability and predictive performance.



## CHAPTER 4

### RESULTS AND DISCUSSIONS

This study offers the examination of the outcomes derived from the implemented methodologies in post-earthquake building damage assessment. Through the aspect of integrated satellite imagery and ML algorithms, this chapter unfolds the empirical findings and embarks on insightful analysis, striving to unravel the implications and significance of the research outcomes.

This chapter explores the efficacy of various ML models in automatically detecting damaged buildings from post-earthquake satellite images. The performance metrics of each approach, shedding light on their strengths, weaknesses, and overall effectiveness in damaged building classification using ML algorithms, are presented. Through rigorous experimentation and analysis, providing an understanding of the capabilities and limitations of these methodologies is aimed, thereby empowering decision-makers with valuable insights for optimized disaster management strategies.

By situating the findings within existing literature and theoretical frameworks, their relevance and potential impact on shaping future disaster response protocols and urban planning initiatives are elucidated. Furthermore, we navigate through the practical implications of this research, considering its ramifications for emergency responders, policymakers, and stakeholders vested in bolstering disaster resilience and community well-being.

Firstly, the histograms of the satellite data for each band are obtained. The histogram charts for the given DN and reflectance image bands can be seen in Figures 4.1 - 4.8.

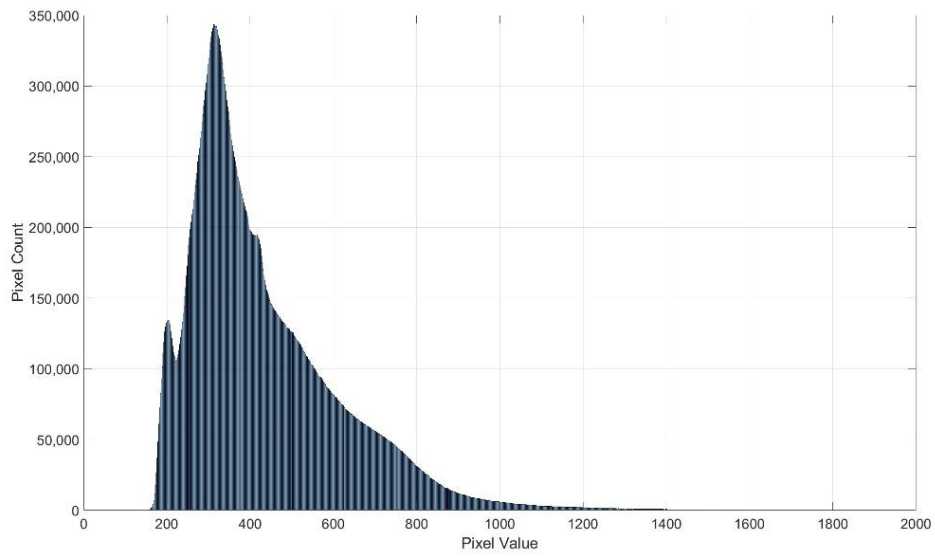


Figure 4.1. Native (DN) image band one (Red) histogram

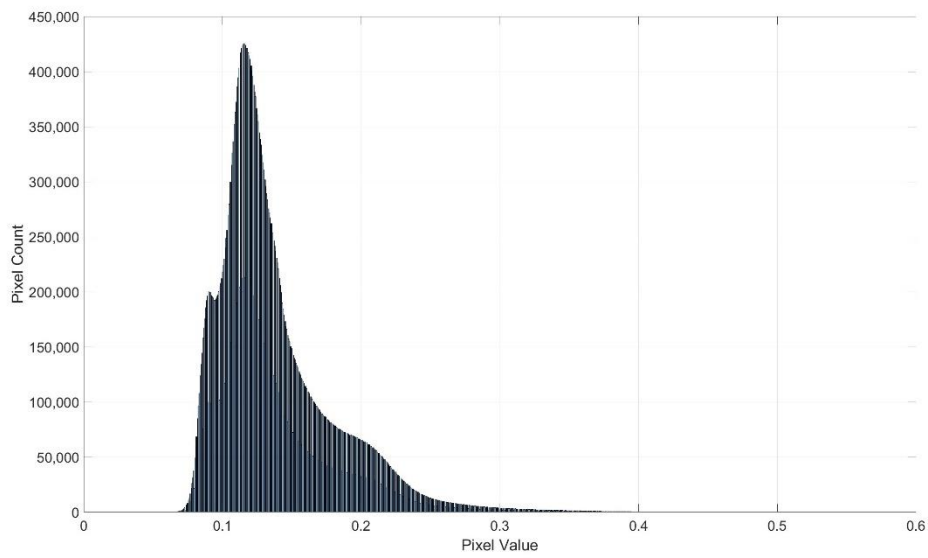


Figure 4.2. Reflectance (TOA) image band one (Red) histogram

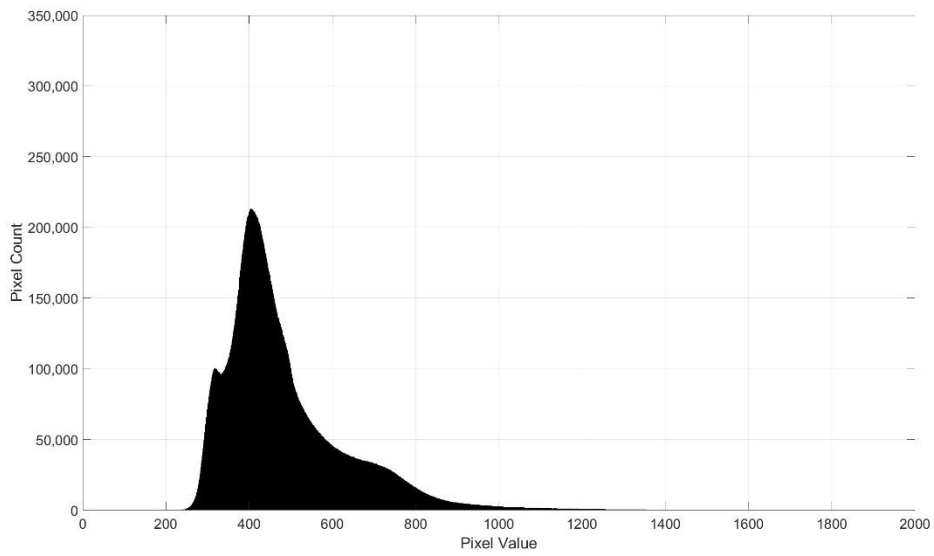


Figure 4.3. Native (DN) image band two (Green) histogram

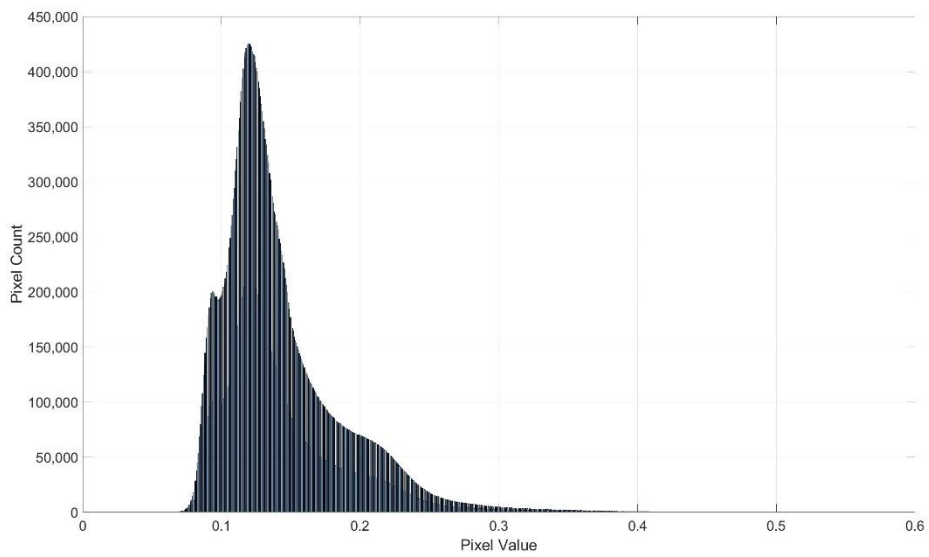


Figure 4.4. Reflectance (TOA) image band two (Green) histogram

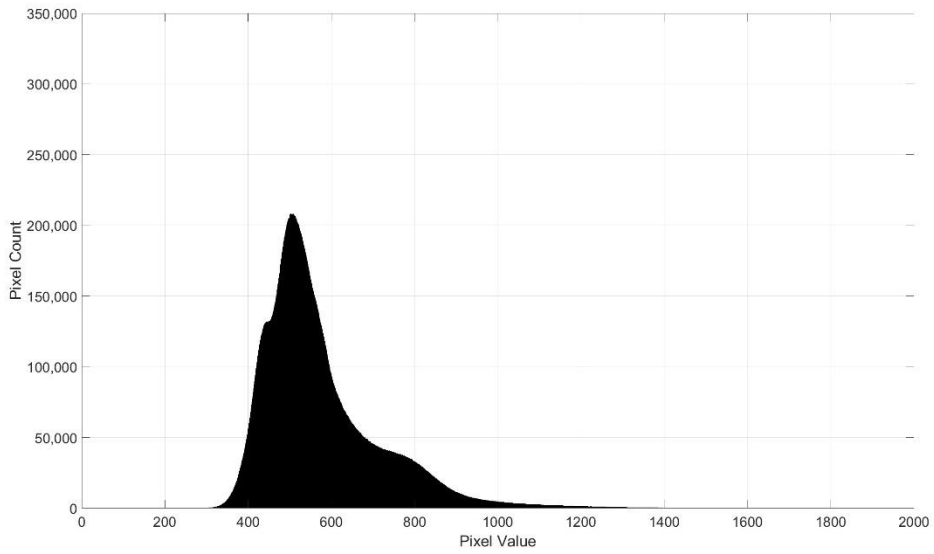


Figure 4.5. Native (DN) image band three (Blue) histogram

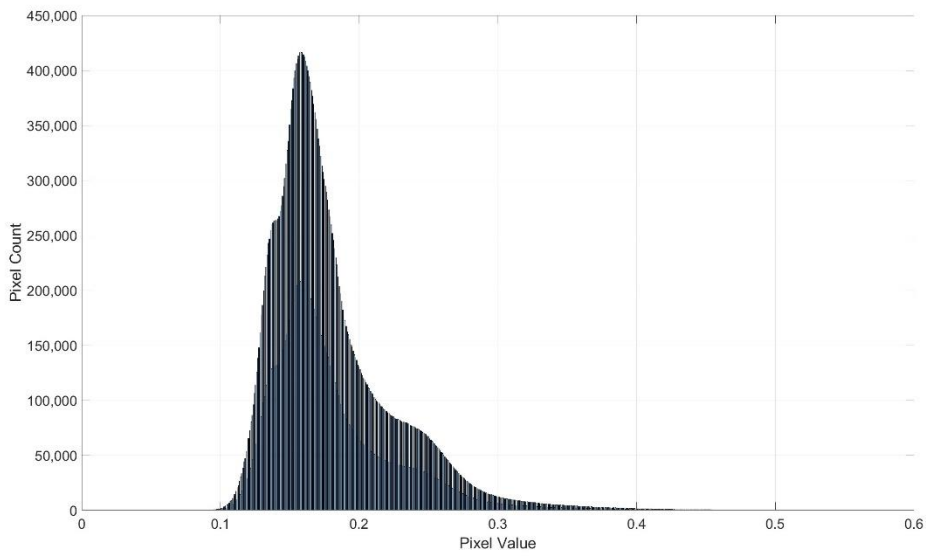


Figure 4.6. Reflectance (TOA) image band three (Blue) histogram



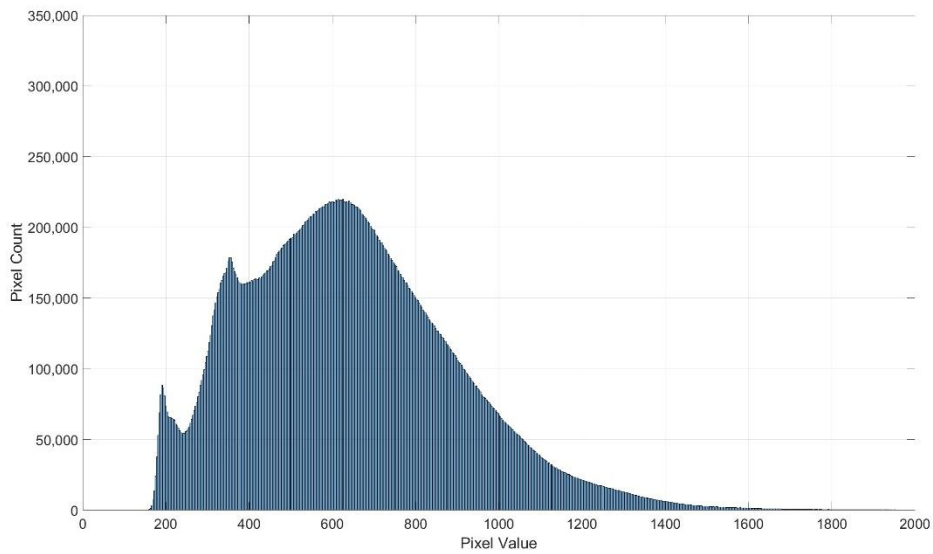


Figure 4.7. Native (DN) image band four (NIR) histogram

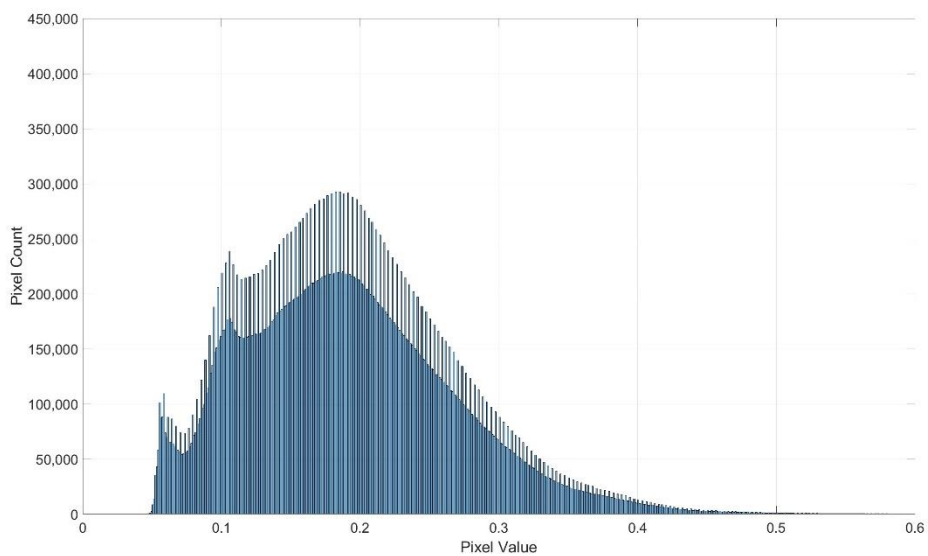


Figure 4.8. Reflectance (TOA) image band four (NIR) histogram

For every native image band of the satellite image provided by UHUZAM, the bands were separated as R, G, B, and NIR image bands, and the reflectance images from them were also calculated. The texture images for every image band separated as R, G, B, and NIR were obtained with the reflectance images produced.

Concerning the study (Rathje *et al.*, 2005), the contrast, dissimilarity, entropy, homogeneity, angular second moment, and correlation texture images were produced and presented in Figure 4.9.

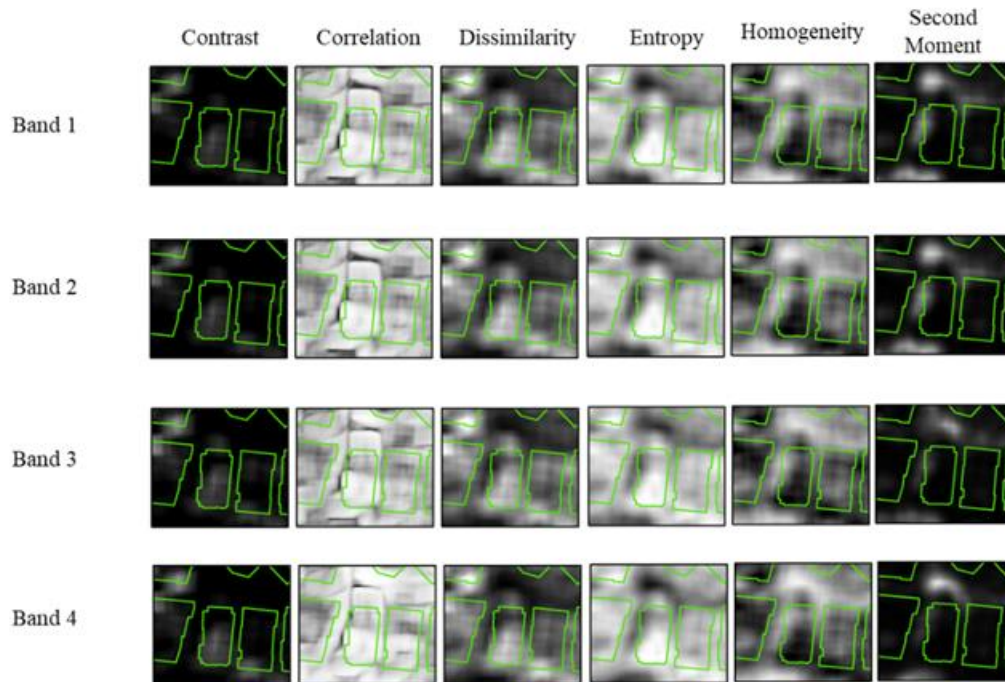


Figure 4.9. Sample images of produced texture image bands overlapping with building footprints.

A thorough histogram analysis was conducted following the production of necessary image and texture bands for each reflectance image band. This analysis aimed to delineate reflectance values and identify any distinctive reflectance values indicative of damaged buildings. Subsequently, eight different texture images were generated for each reflectance band. These texture images were then subjected to histogram analysis, specifically focusing on four native building classes obtained from the Directorate General of Geographical Information Systems of Turkey (GD of GIS, 2023).

The comprehensive analysis aimed to generate texture image histograms corresponding to four native building classes. However, upon meticulous scrutiny, it became evident that the texture data histograms exhibited discernible similarities

across various building damage classes. Consequently, the efficacy of distinguishing between the diverse building damage classes based solely on these texture data histograms was limited. Upon examination of the texture histograms (sample of the histograms produced illustrated in Figures 4.10 - 4.13), it was observed that the texture images displayed negligible discrepancies among the four classes within the constraints of texture images produced from reflectance image bands. This observation suggests that the texture features derived from reflectance images fail to discriminate adequately between the different building damage classes. For a comprehensive overview, the complete ensemble of 128 histogram figures illustrating the texture images derived from the four-class reflectance data are given in Appendix A.

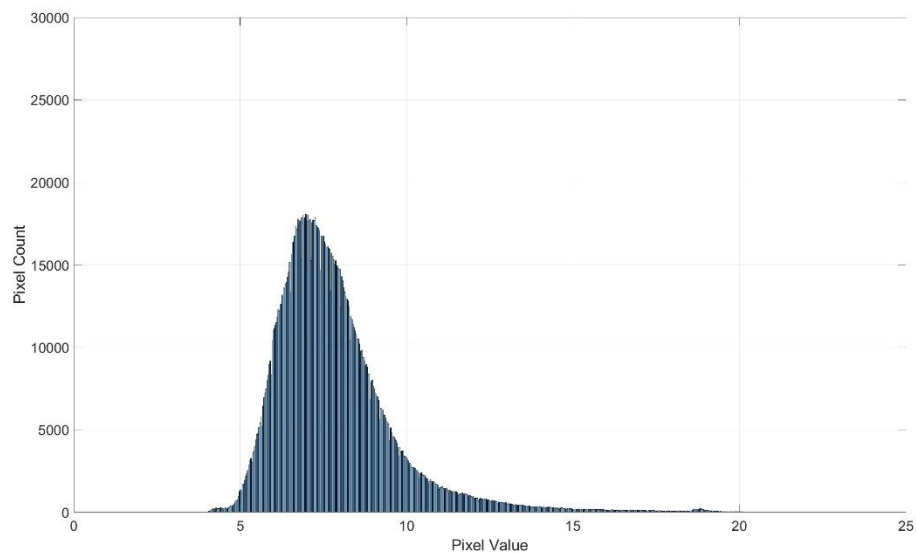


Figure 4.10. The mean texture histogram for Band one of the heavily damaged class.

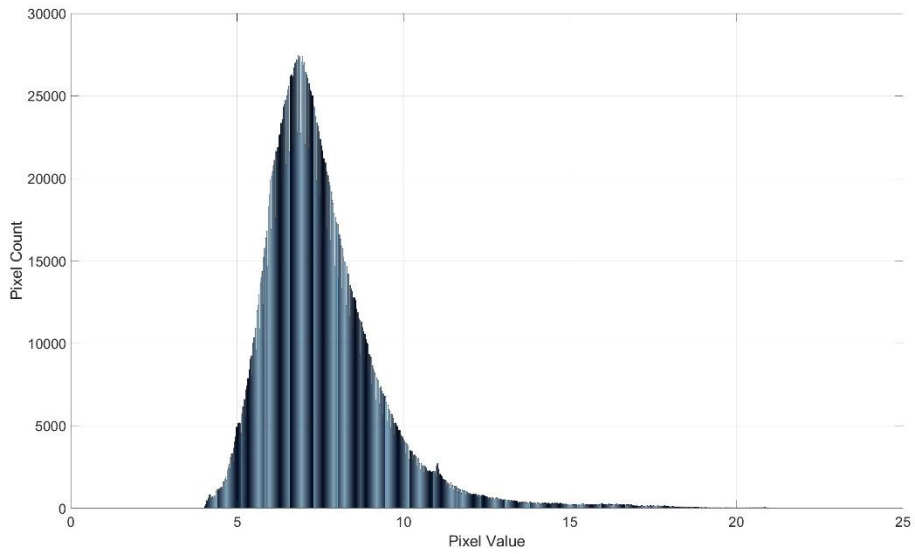


Figure 4.11. The mean texture histogram for Band one of the heavily no-damage class.

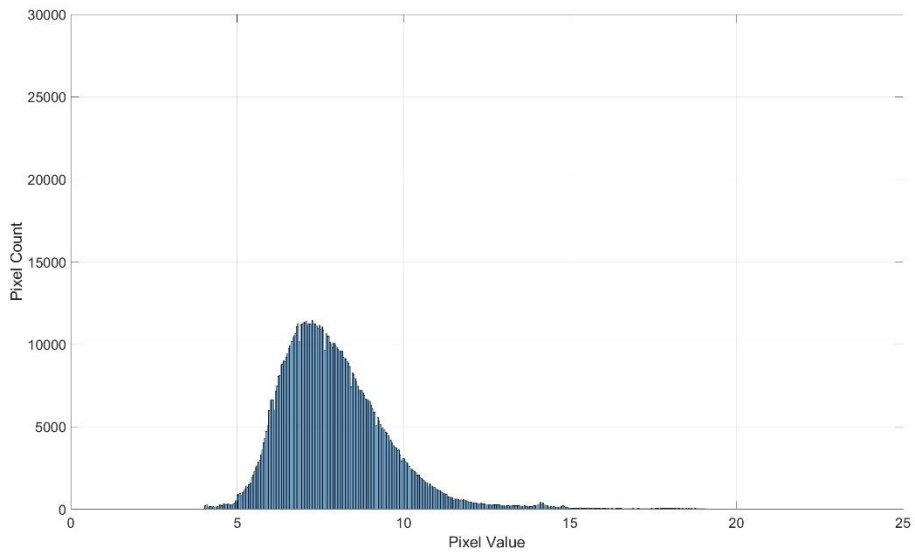


Figure 4.12. The mean texture histogram for Band one of the collapsed class.

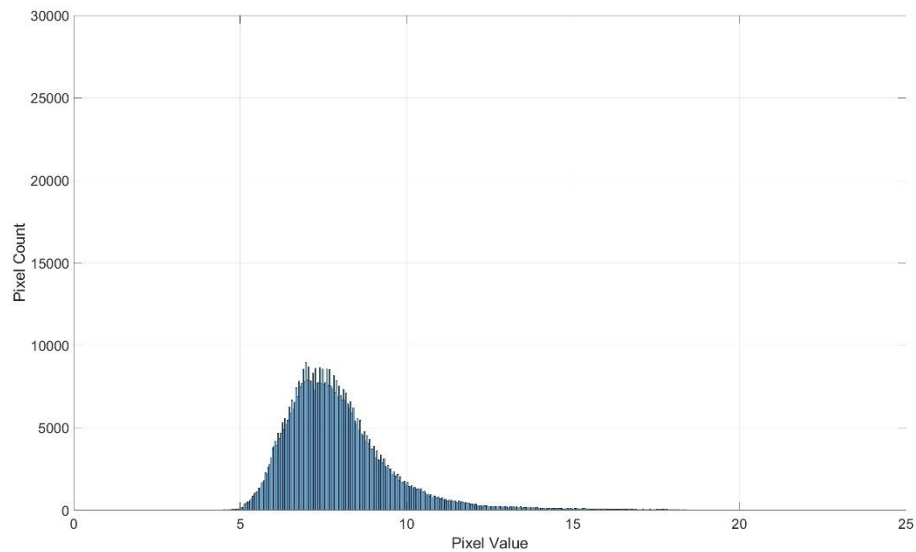


Figure 4.13. The mean texture histogram for Band one of the urgent demolish class.

The research methodology entailed the initial utilization of histogram analysis to assess the distribution characteristics of the dataset across building classes. However, given the complexity inherent in the data, this approach might not comprehensively elucidate all relevant nuances. To understand the dataset in detail, mean reflectance charts were generated separately for each building class. Concurrently, texture images were also analyzed to discern spatial characteristics within the dataset. The absence of discernible differentiation between classes in reflectance and texture images suggests a potential inadequacy of these features in class discrimination. In response, alternative features or data sources warrant exploration. These charts can be seen in Figure 4.14 - 4.20.

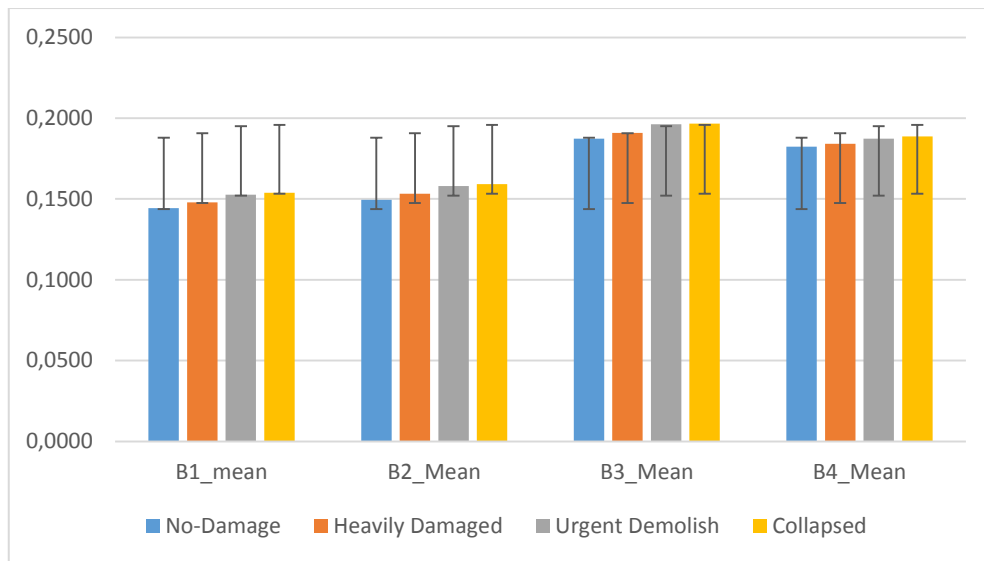


Figure 4.14. The mean values of the image bands with standard deviation error bars.

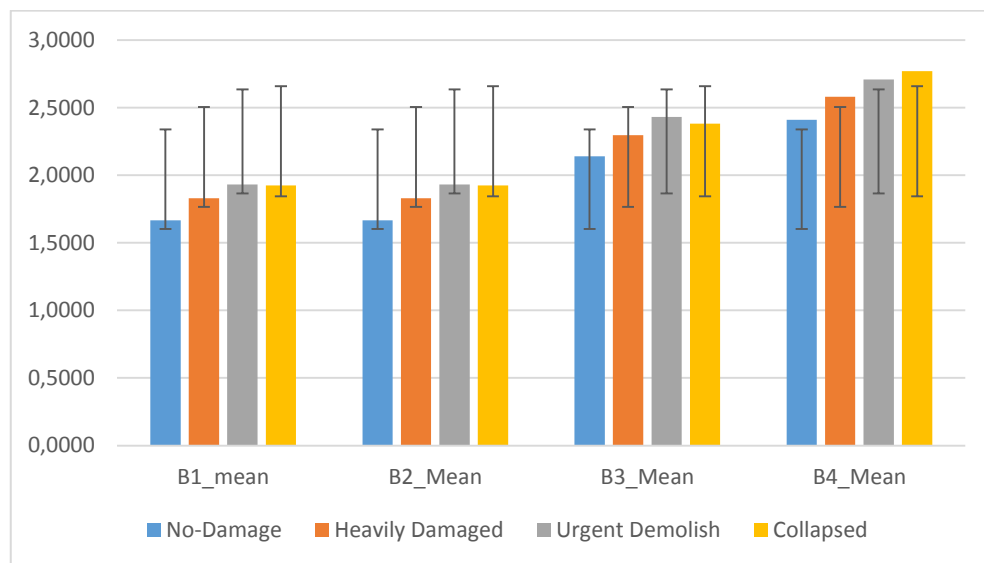


Figure 4.15. The mean values of the contrast texture image bands with standard deviation error bars.

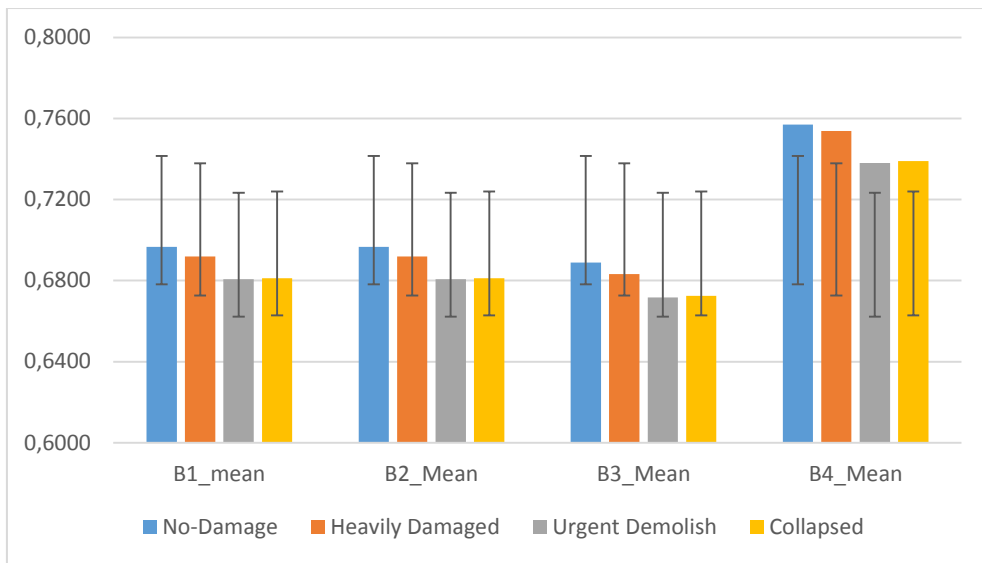


Figure 4.16. The mean values of the correlation texture image bands with standard deviation error bars.

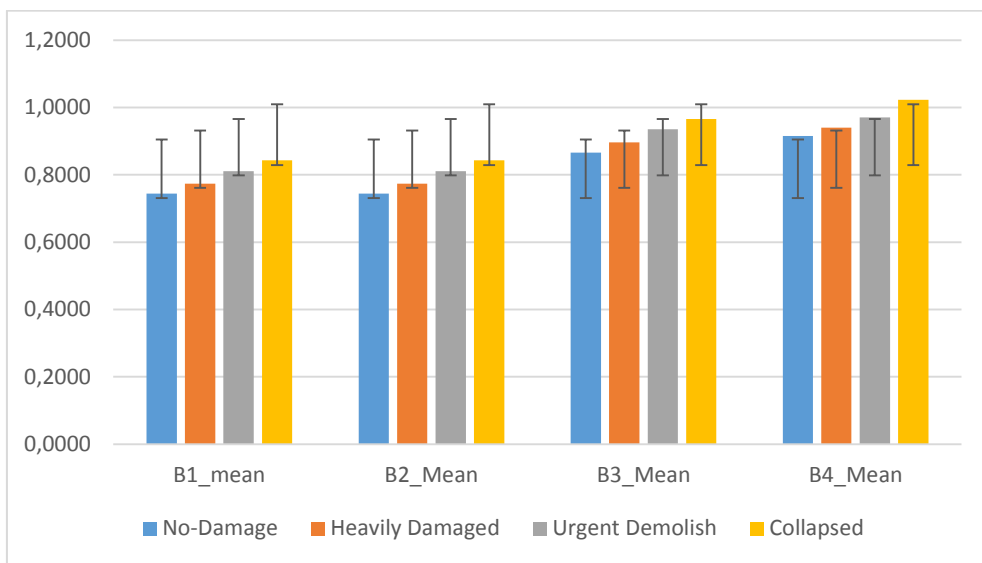


Figure 4.17. The mean values of the dissimilarity texture image bands with standard deviation error bars.

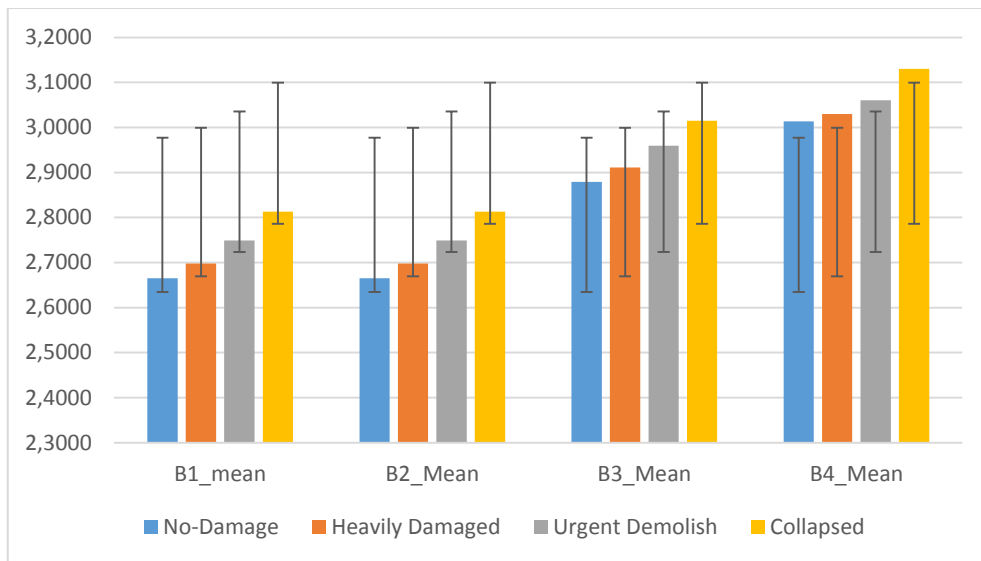


Figure 4.18. The mean values of the entropy texture image bands with standard deviation error bars.

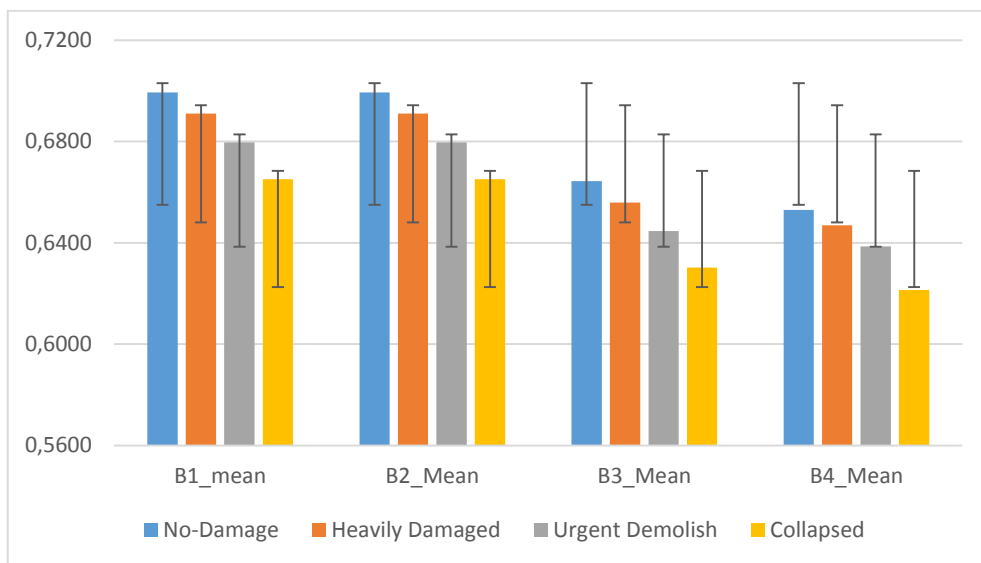


Figure 4.19. The mean values of the homogeneity texture image bands with standard deviation error bars.



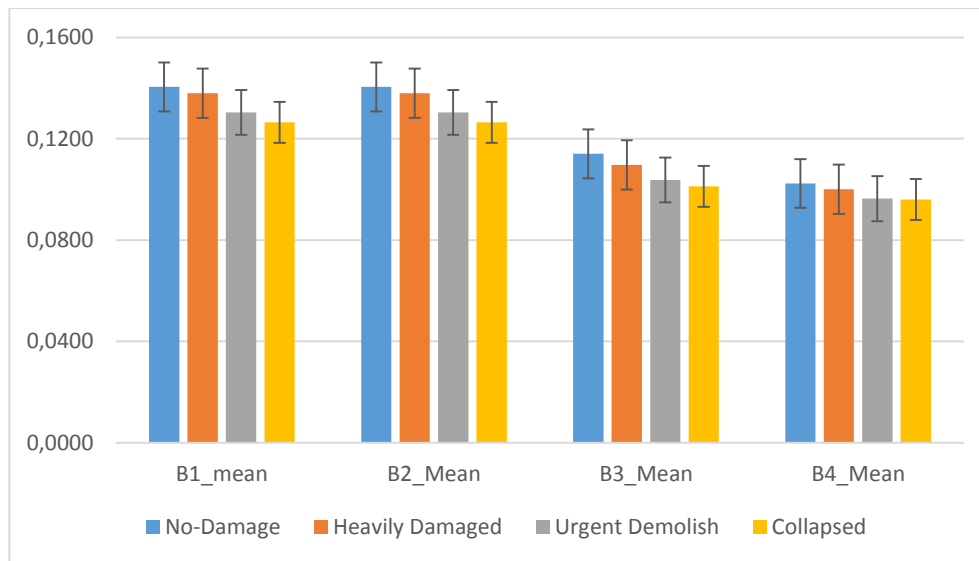


Figure 4.20. The mean values of the angular 2<sup>nd</sup> moment texture image bands with standard deviation error bars.

Despite the extensive analysis conducted on the texture data histograms derived from the reflectance images, the observed similarities among the histograms posed challenges in distinguishing between different building damage classes. This finding underscores the complexity of accurately characterizing building damage using texture analysis alone, highlighting the need for further investigation and refinement of analytical methodologies in this domain.

By applying PCA to the imagery dataset, a series of orthogonal principal components were obtained, and each represents distinct spatial patterns or features inherent in the data. Three visible image bands and the near-infrared (NIR) image band were used to produce four PCA image bands.

Furthermore, the directional texture analysis was conducted on the PCA-derived images, considering 0, 45, 90, and 135-degree orientations. This analytical approach allowed us to identify and emphasize spatial patterns or textures aligned with specific orientations, enhancing the ability to discern damaged building features.

The interpretation of each directional component is as follows:

- Direction 0 degrees denotes a horizontal orientation, capturing spatial patterns predominantly aligned along the horizontal axis. Such features may include roads, coastlines, or other linear structures parallel to the image's horizontal axis.
- Direction 45 degrees corresponds to a diagonal orientation, capturing spatial patterns aligned at a 45-degree angle relative to the horizontal axis. These patterns may highlight sloping terrain, vegetation edges, or other diagonal structures within the imagery.
- Direction 90 degrees signifies a vertical orientation, emphasizing spatial patterns predominantly aligned in a vertical orientation. This orientation may reveal tall buildings, trees, or other vertical structures perpendicular to the horizontal axis.
- Direction 135 degrees reflects an opposite diagonal orientation, capturing spatial patterns aligned at a 135-degree angle relative to the horizontal axis. Features like those captured at 45 degrees are emphasized, albeit oppositely oriented.

Following the inconclusive results from the histogram analysis, the investigation shifted towards producing correlations for each of the four native building damage classes individually. The objective was to determine if the textures derived from the reflectance data could effectively distinguish between these classes. Subsequently, correlation analysis was conducted for each class separately to identify relationships within the data. All texture images were employed as input data for each building class. Subsequently, total correlation matrices were computed to ascertain the absence of correlations between the texture images and the respective building classes. These correlation analysis outcomes are presented in Figure 4.21 - 4.24.

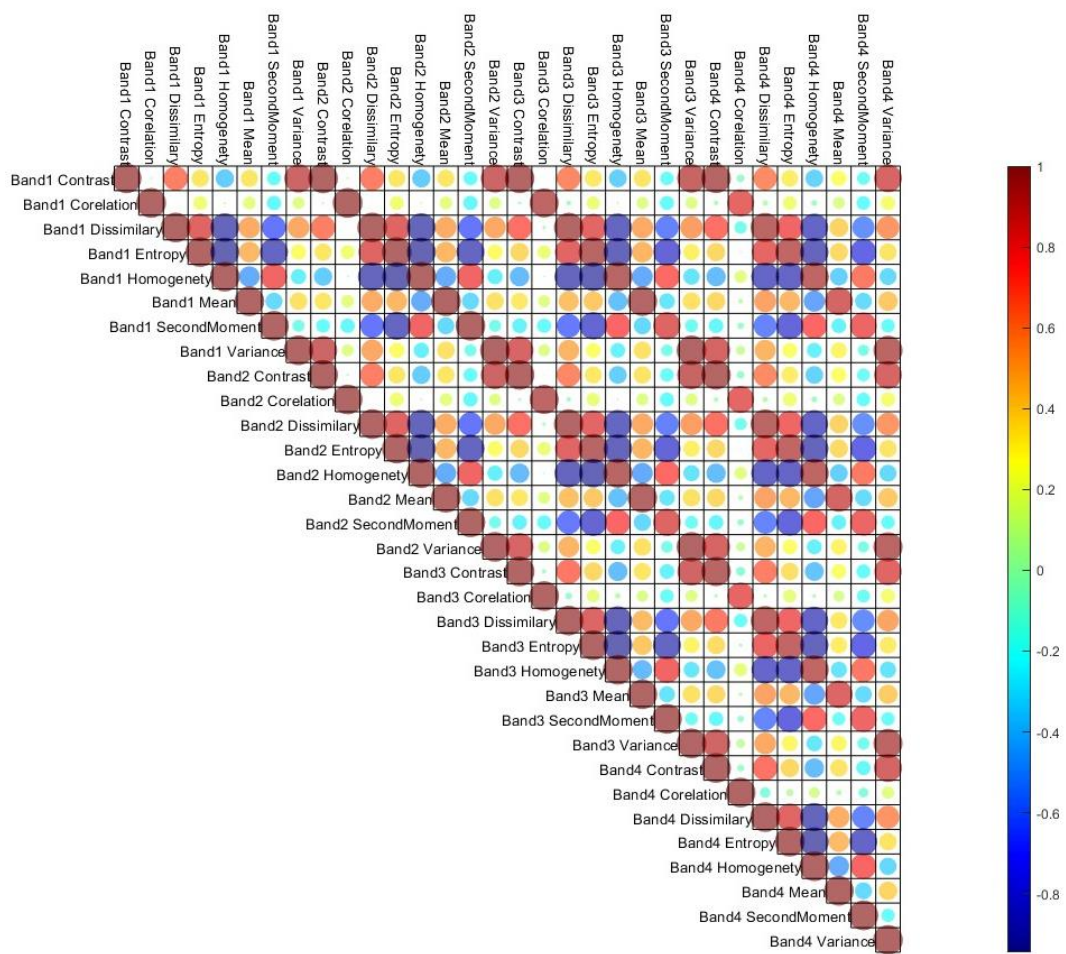


Figure 4.21. Heavily damaged building class correlation matrix.

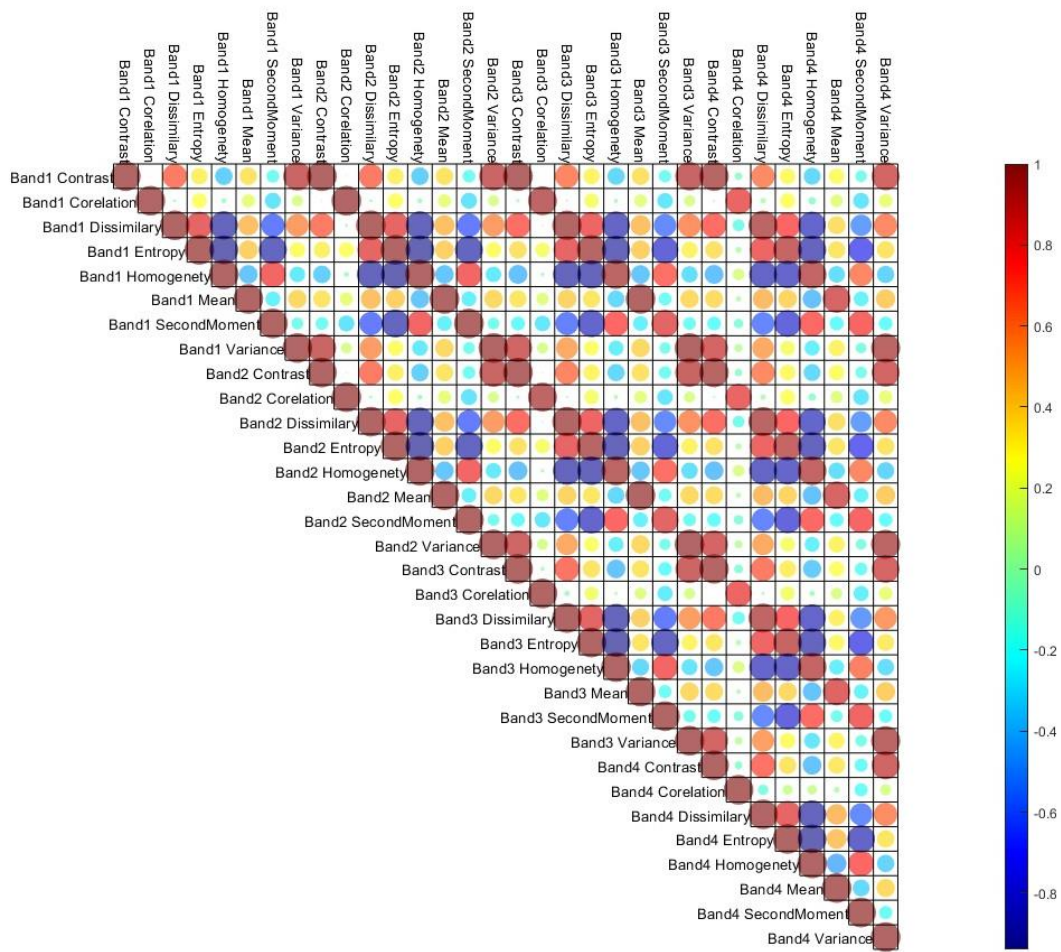


Figure 4.22. No-damage building class correlation matrix.

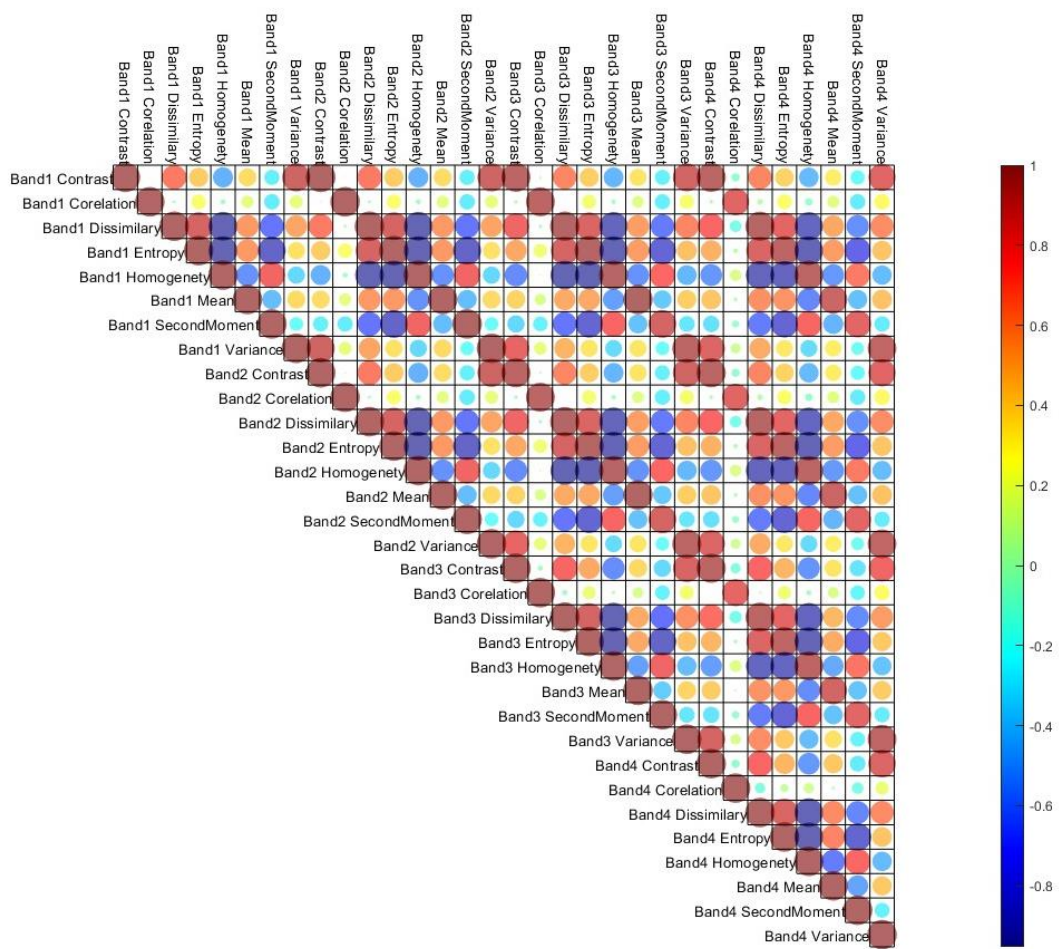


Figure 4.23. Collapsed building class correlation matrix.

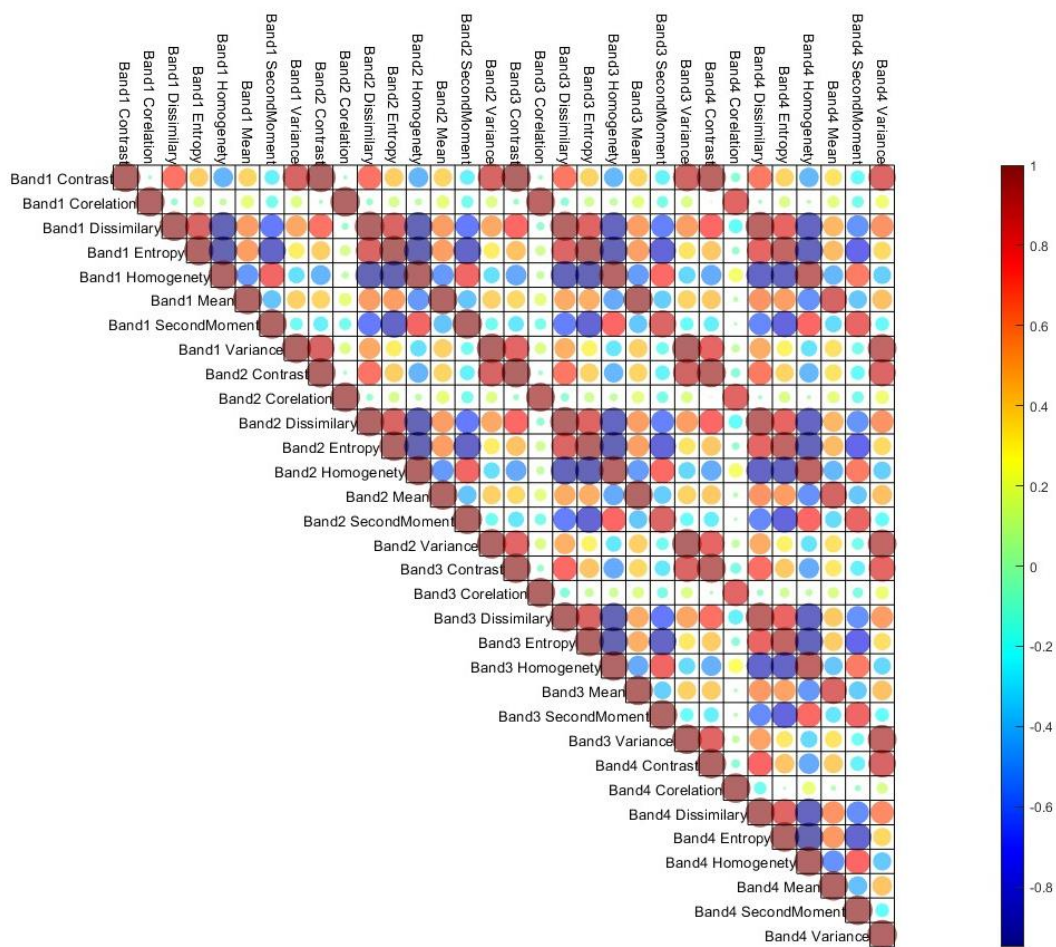


Figure 4.24. Urgent demolish building class correlation matrix.

The correlation analysis revealed both negative and positive correlations among the data. However, despite these correlations, the analysis did not yield sufficiently discernible patterns to enable the RF ML algorithm to achieve an overall accuracy rate exceeding 51%. Error matrices for the RF algorithm detection can be inspected for training data in Table 4.1 and for test data in Table 4.2. Additionally, prediction and user accuracy metrics can be seen in Table 4.3 and Table 4.4.

Table 4.1. Training data error matrix for RF ML algorithm.

	No-Damage	Heavily Damaged	Urgent Demolish	Collapsed	Row Total
No-Damage	2037	0	0	0	2037
Heavily Damaged	0	1051	0	0	1051
Urgent Demolish	0	0	369	0	369
Collapsed	0	0	0	533	533
Column Total	2037	1051	369	533	3990

Table 4.2. Test data error matrix for RF ML algorithm.

	No-Damage	Heavily Damaged	Urgent Demolish	Collapsed	Row Total
No-Damage	754	360	99	143	1356
Heavily Damaged	79	74	37	30	220
Urgent Demolish	5	0	3	2	10
Collapsed	35	17	19	53	124
Column Total	873	451	158	228	1710

Table 4.3. Prediction accuracy for the building classes

	No-Damage	Heavily Damaged	Urgent Demolish	Collapsed
Training	100.00	100.00	100.00	100.00
Test	86.37	16.41	1.90	23.25

Table 4.4. User accuracy for the building classes

	No-Damage	Heavily Damaged	Urgent Demolish	Collapsed
Training	100.00	100.00	100.00	100.00
Test	55.60	33.64	30.00	42.74

This outcome suggests that while some degree of correlation between texture data and building damage classes may exist, the correlations are not robust enough to facilitate accurate classification using ML algorithms alone. Therefore, additional factors or features may need to be considered in conjunction with texture data to enhance the classification accuracy.

The original goal of the thesis was to classify damaged buildings into four distinct classes. However, upon further consideration, it was recognized that prioritizing the accuracy of detected buildings is paramount. Therefore, a strategic decision was made to reclassify the building classes into two categories: "No-damage" and "Damaged." The rationale behind this reclassification and detailed explanations of the classification methodology are provided in Chapter 3.3.3 of the thesis.

This reclassification allows for a more streamlined and focused approach to building damage assessment, simplifying the classification process while prioritizing the accuracy of identifying damaged buildings. By consolidating the classes into two broad categories, the analysis can better serve the overarching goal of effectively identifying and assessing building damage in satellite or aerial imagery.

This strategic shift in the classification approach reflects a commitment to optimizing the accuracy and efficiency of the building damage assessment process, thereby enhancing the utility and practicality of the research outcomes for disaster response efforts. With the building classes reduced to two classes and the directional texture for the PCA image bands in hand, we first produced the



correlation matrices for each building class. These images can be seen in Figures 4.25 - 4.28.

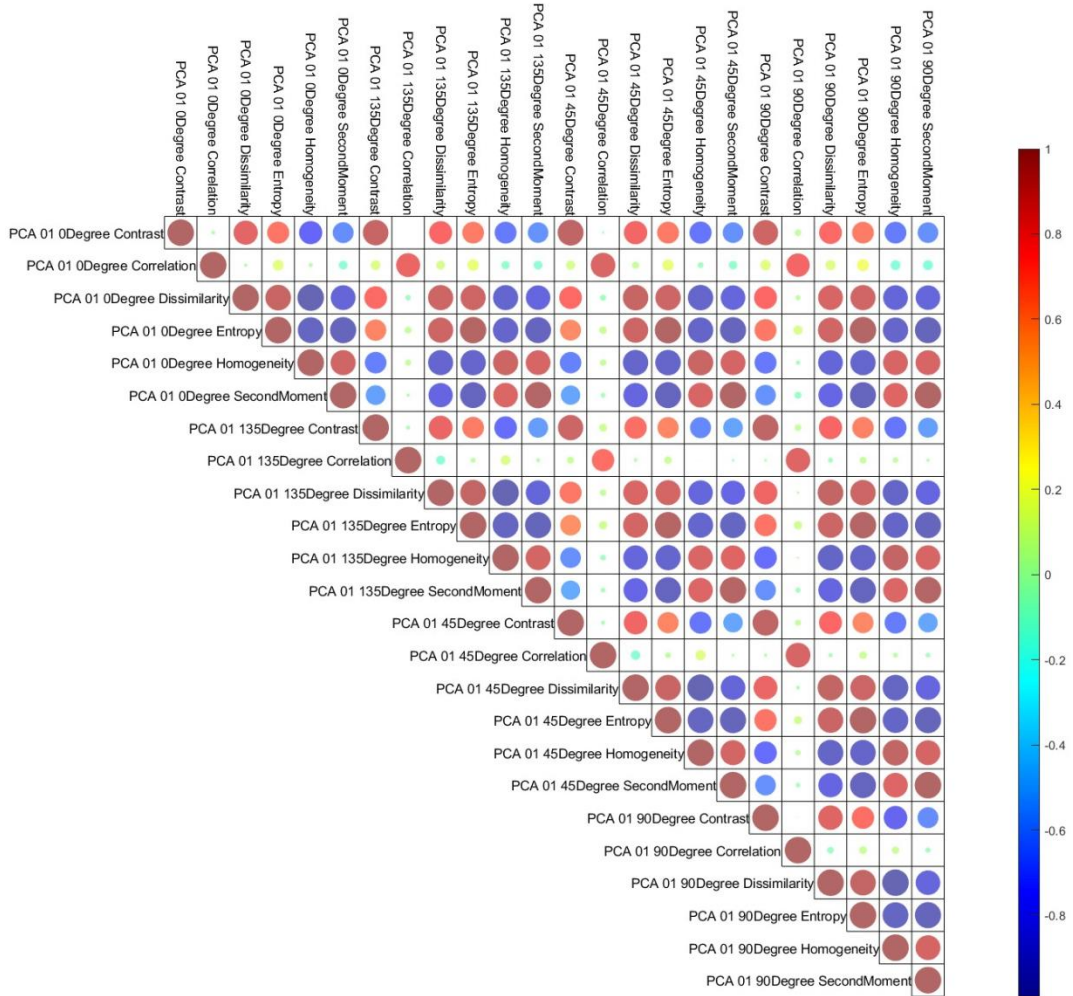


Figure 4.25. Correlation matrix for the PCA analysis for band one\_texture

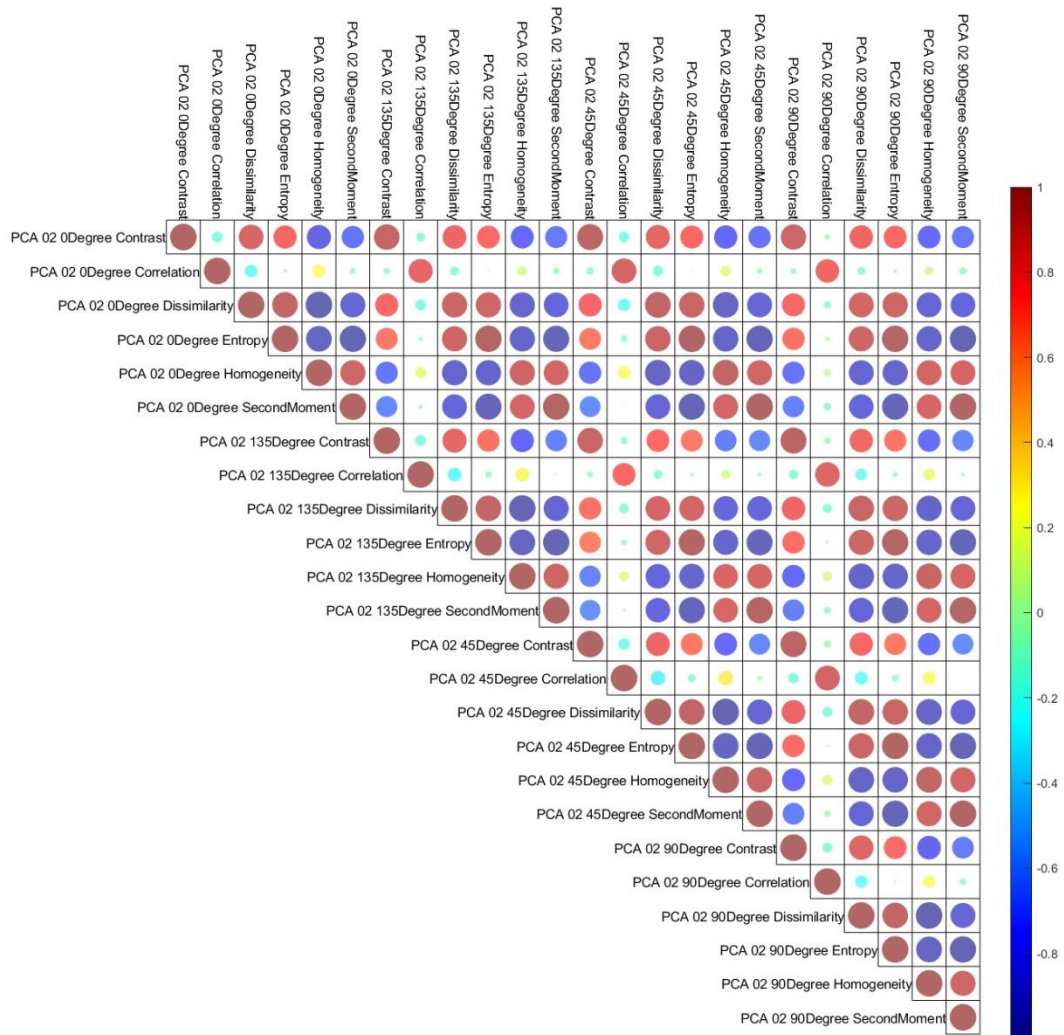


Figure 4.26. Correlation matrix for the PCA analysis for band two\_texture.

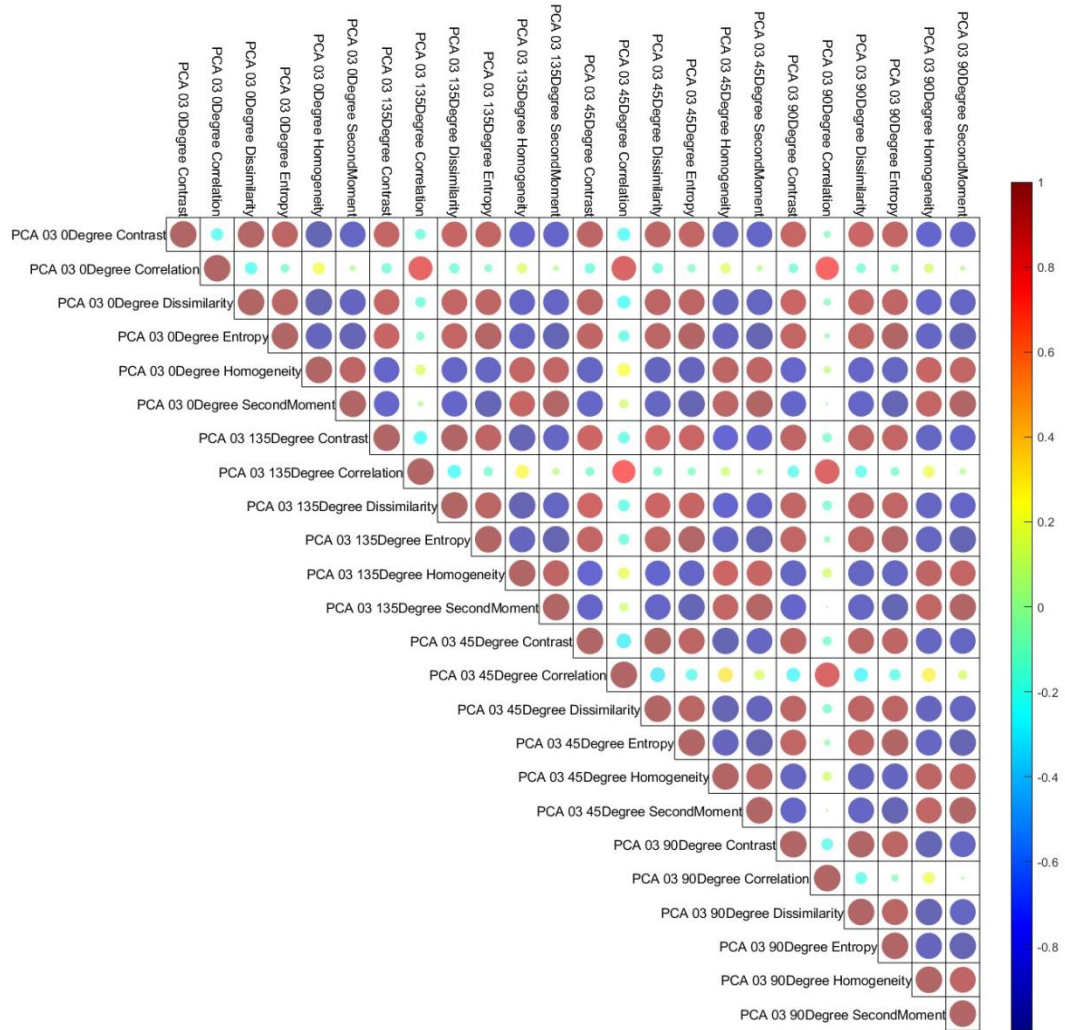


Figure 4.27. The correlation matrix for the PCA analysis for band three\_texture

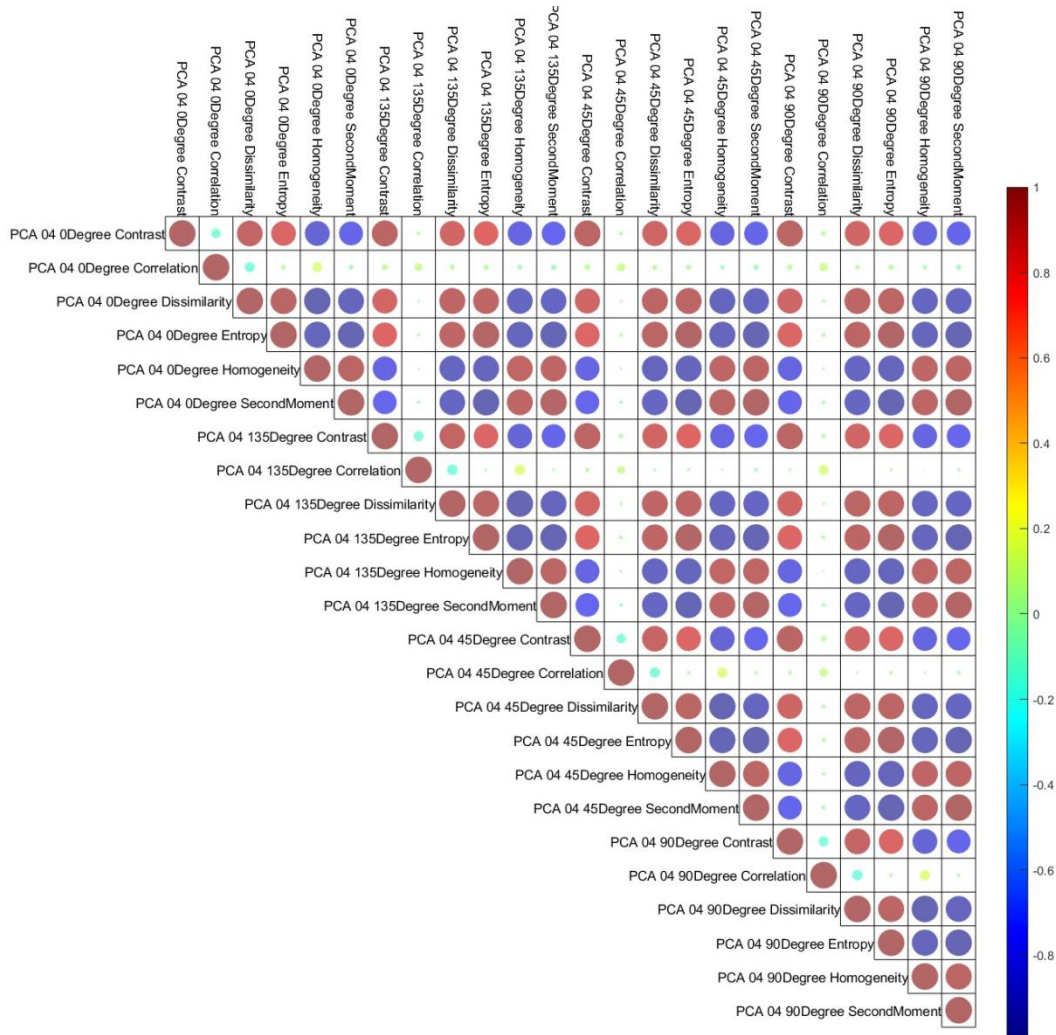


Figure 4.28. Correlation matrix for the PCA analysis for band four\_texture.

The correlation charts show that specific image band and texture values exhibit strong correlations close to +1 and -1, indicating significant relationships within the dataset. Conversely, specific PCA image band textures demonstrate correlations nearing 0, suggesting a lack of pronounced association with the building classes essential for achieving distinctive separation. We prioritize discerning a clear distinction between the "No-Damage" and "Damaged" classes in the investigation. To this end, the most discriminative texture image bands are amalgamated, and distinct building footprints are utilized to optimize class differentiation. Specifically, PCA image bands 2 and 4, which exhibit the most apparent separation

between classes, were leveraged, and a correlation matrix based on this refined selection was generated. The resulting matrix is depicted in Figure 4.29.

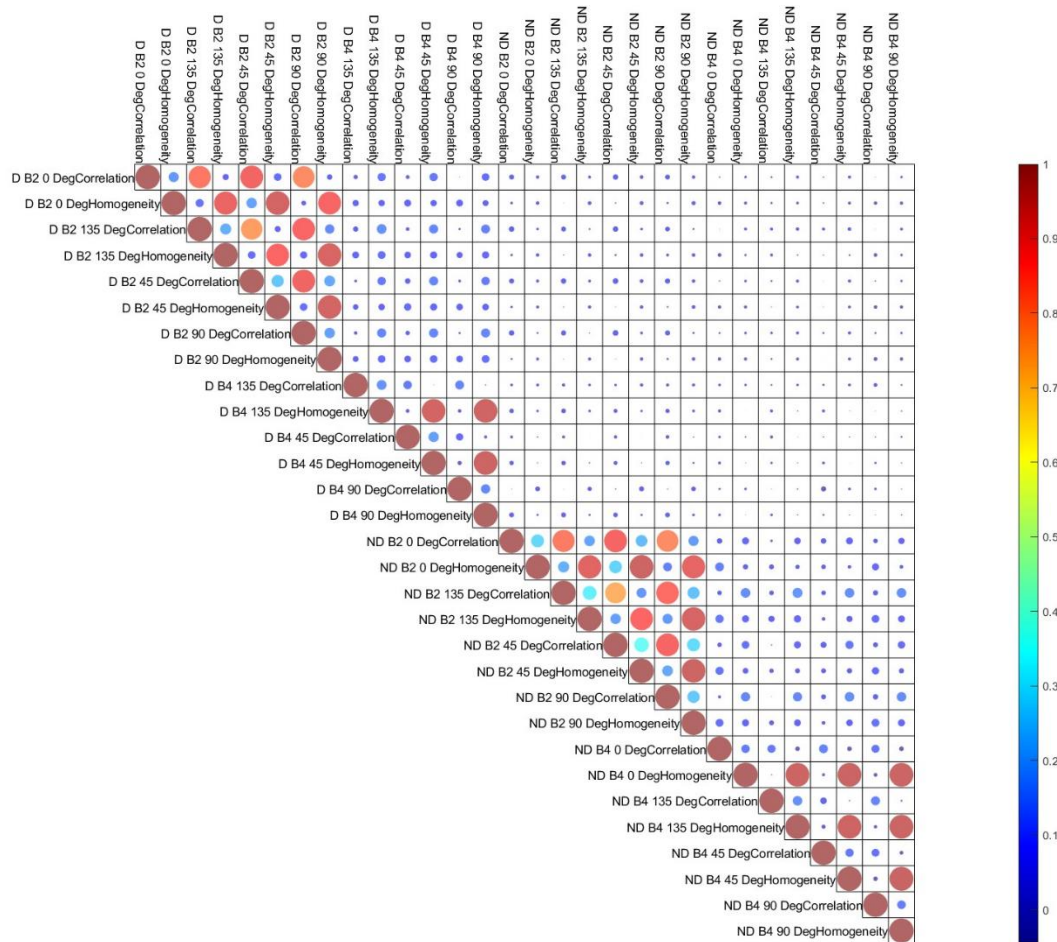


Figure 4.29. Correlation matrix for the PCA Bands 2 and 4 with directional texture image applied to No-Damage and Damaged building class.

A correlation matrix was generated using the insights derived from Figure 4.29 to minimize correlations across all image bands and directional texture images. This meticulous approach enabled the identification of the most suitable image texture for integration into the machine-learning model. The resulting correlation matrix is visually represented in Figure 4.30.

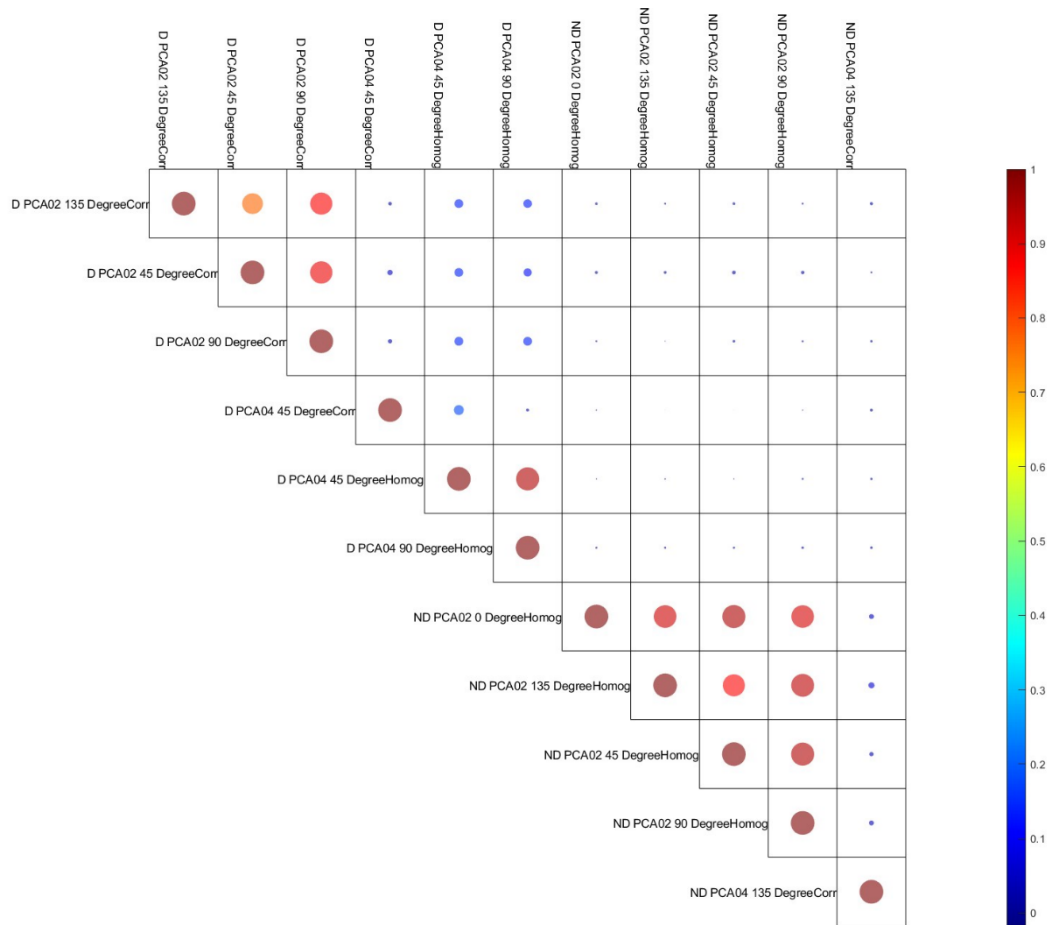


Figure 4.30. Minimum correlation matrix for directional image bands used for ML models.

The correlation analysis has identified eleven directional image bands that exhibit optimal suitability for integration into ML algorithms. These selected bands demonstrate the lowest correlations between the "Damaged" and "No-Damage" classes and among the images. The specific combinations of these image bands are presented comprehensively in Table 4.5. In addition to the selected eleven directional texture images, four reflectance image bands were included to see the differences in accuracy. In contrast, the four reflectance images were produced from the native image bands (Ref\_Direct\_PCA\_minCor).

Table 4.5. Minimum Correlation of the Directional Texture Image Bands (Direct\_PCA\_minCor) and added reflectance image bands (Ref\_Direct\_PCA\_minCor).

<b>Produced Image</b>	<b>Image Band</b>	<b>Texture</b>	<b>Direction</b>
Reflectance	1 <sup>st</sup> Band	-	-
Reflectance	2 <sup>nd</sup> Band	-	-
Reflectance	3 <sup>rd</sup> Band	-	-
Reflectance	4 <sup>th</sup> Band	-	-
PCA	2 <sup>nd</sup> Band	Homogeneity	0 Degree
PCA	2 <sup>nd</sup> Band	Correlation	45 Degree
PCA	2 <sup>nd</sup> Band	Homogeneity	45 Degree
PCA	2 <sup>nd</sup> Band	Correlation	90 Degree
PCA	2 <sup>nd</sup> Band	Homogeneity	90 Degree
PCA	2 <sup>nd</sup> Band	Correlation	135 Degree
PCA	2 <sup>nd</sup> Band	Homogeneity	135 Degree
PCA	4 <sup>th</sup> Band	Correlation	45 Degree
PCA	4 <sup>th</sup> Band	Homogeneity	45 Degree
PCA	4 <sup>th</sup> Band	Homogeneity	90 Degree
PCA	4 <sup>th</sup> Band	Correlation	135 Degree

These two datasets are produced to give the ML algorithms the best possible distinctions between the two building classes. Given the nature of the study, the most distinct image and texture bands were determined to be used for the ML algorithm for future studies as well. The vector data was used to determine the training and test datasets.

Within each dataset, we have identified the specific image that exhibits the lowest correlation between the "Damaged" and "No-Damage" classes. This determination was facilitated by utilizing vector-building footprints. The mean value for the corresponding image was calculated for each building footprint. Each unique object

ID within the dataset corresponds to a distinct building footprint, providing a dataset-specific mean value corresponding to each object ID across all dataset images. The workflow can be seen in Figure 4.31.

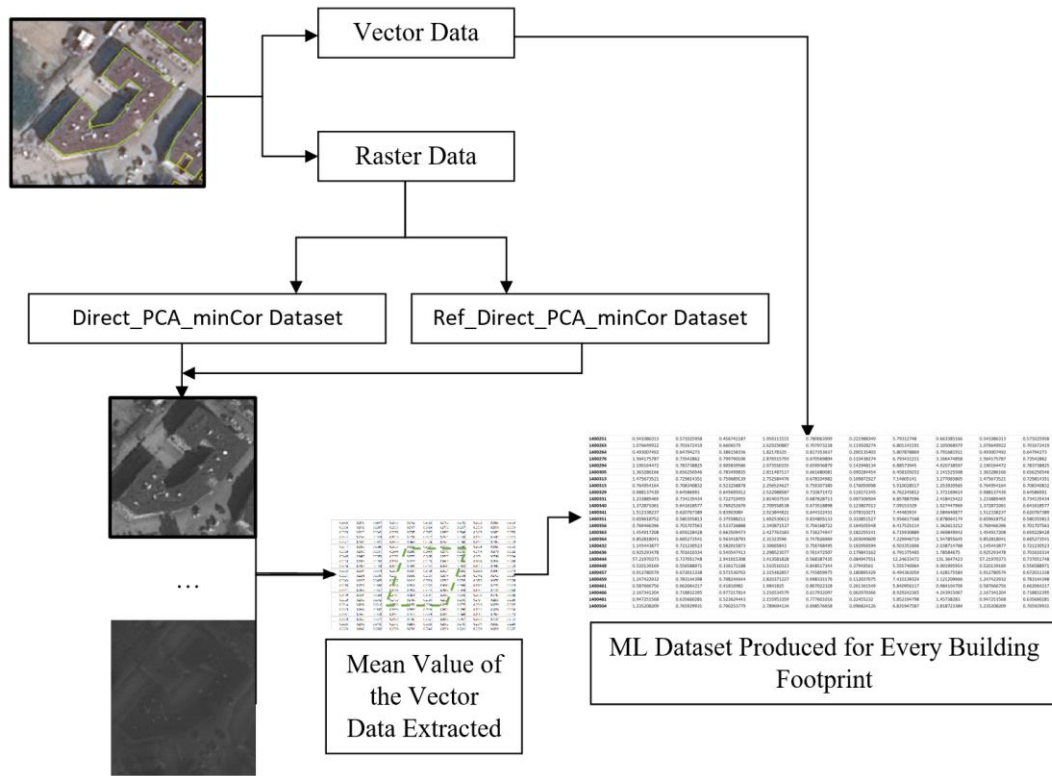


Figure 4.31. Mean value extraction for all buildings in the study area.

In this study, a ten-fold cross-validation process for the data was used. Ten-fold cross-validation is a prevalent technique in ML for assessing the robustness and generalization capabilities of predictive models. This process entails systematically partitioning the dataset into ten mutually exclusive subsets called "folds." Each fold comprises an approximately equal proportion of the dataset, ensuring that data distribution remains consistent across all iterations. Subsequently, the model undergoes ten iterations, during which nine folds are utilized for training, while the remaining fold serves as a validation set. This procedure is repeated ten times, with each fold being used exactly once as the validation set. The performance metrics obtained from each iteration are then averaged to estimate the model's predictive accuracy and generalization ability. Ten-fold cross-validation is favored for its



capacity to mitigate biases associated with single train-test splits, thereby fostering a comprehensive assessment of its robustness and generalization to unseen data. The partitions (i.e., folds) distribution of the study area can be seen in Figure 4.32.

This map delineating every building's footprint alongside corresponding fold numbers is a pivotal visual aid in comprehending ML experimentation distribution and allocation of data subsets. Such a visualization offers a holistic depiction of how the dataset is partitioned across distinct folds, elucidating spatial patterns and potential correlations between geographic locations and fold assignments. Moreover, overlaying fold numbers onto building footprints gives a nuanced understanding of how individual instances are distributed across different folds, thereby facilitating insights into model performance variations, data stratification strategies, and potential sources of bias or variance.

When interpreting the results of model accuracy, it is crucial to consider the performance on both training and test datasets. High accuracy on the training data suggests that the model has effectively learned the underlying patterns and relationships in the training examples. However, the ultimate measure of a model's efficacy lies in its performance on the test data, which represents unseen instances. If the accuracy of the test data closely mirrors that of the training data, it indicates that the model has successfully generalized its learned patterns to new, unseen instances. Conversely, a significant disparity between training and test accuracies may signal overfitting, wherein the model has memorized the training data rather than learning meaningful patterns, resulting in poor performance on unseen data. Therefore, a thorough interpretation of accuracy metrics necessitates an assessment of both training and test performances to gauge the model's reliability and generalization capability across diverse datasets.

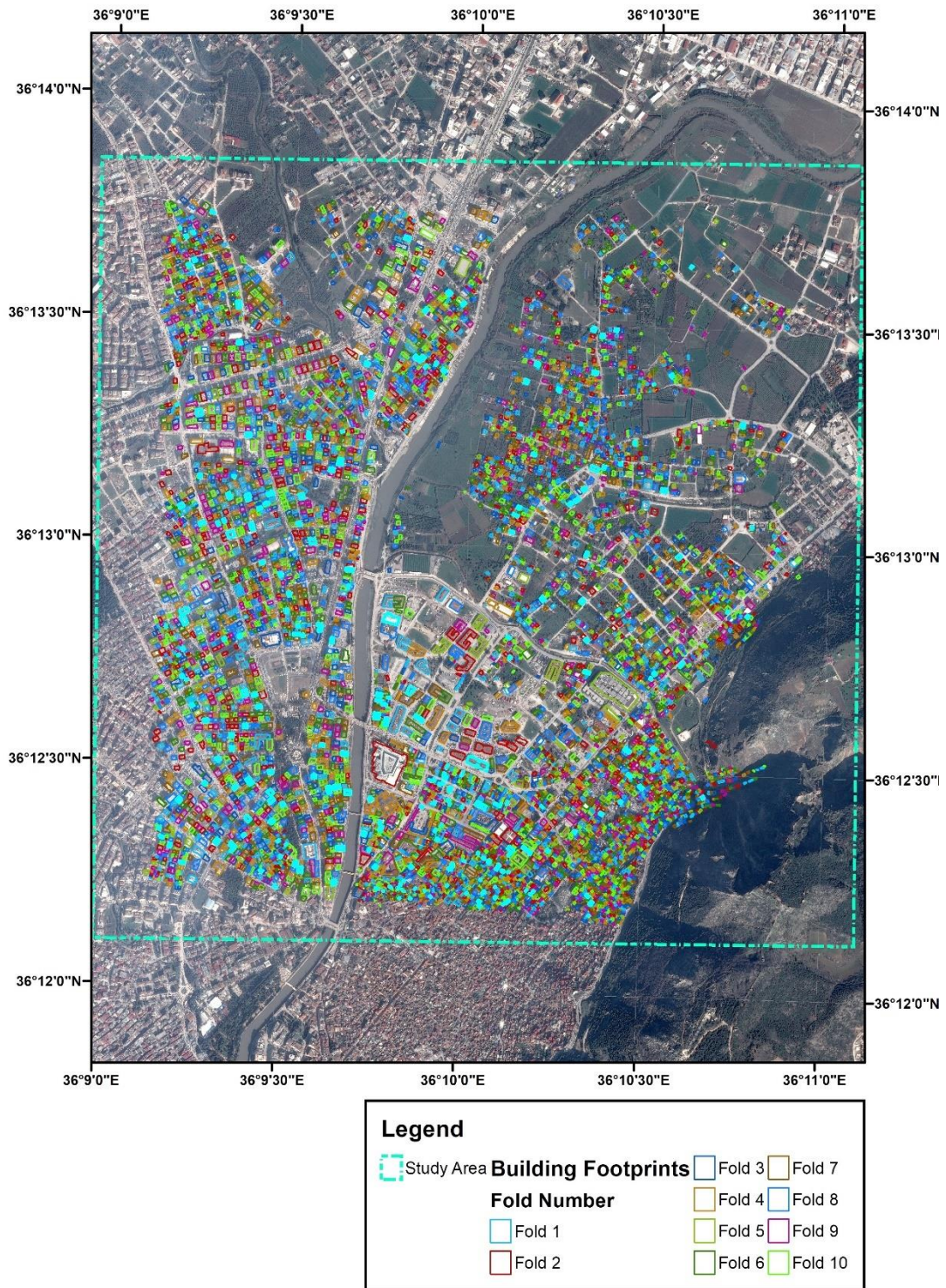


Figure 4.32. Distribution of Folds (Partitions) of the study area.

With the ML algorithms described in section 3.4, the test and training data accuracies for inspection were produced and compared. Accuracy results for the training data for every fold for both datasets Direct\_PCA\_minCor and Ref\_Direct\_PCA\_minCor are as follows in Figure 4.33 and Figure 4.34.

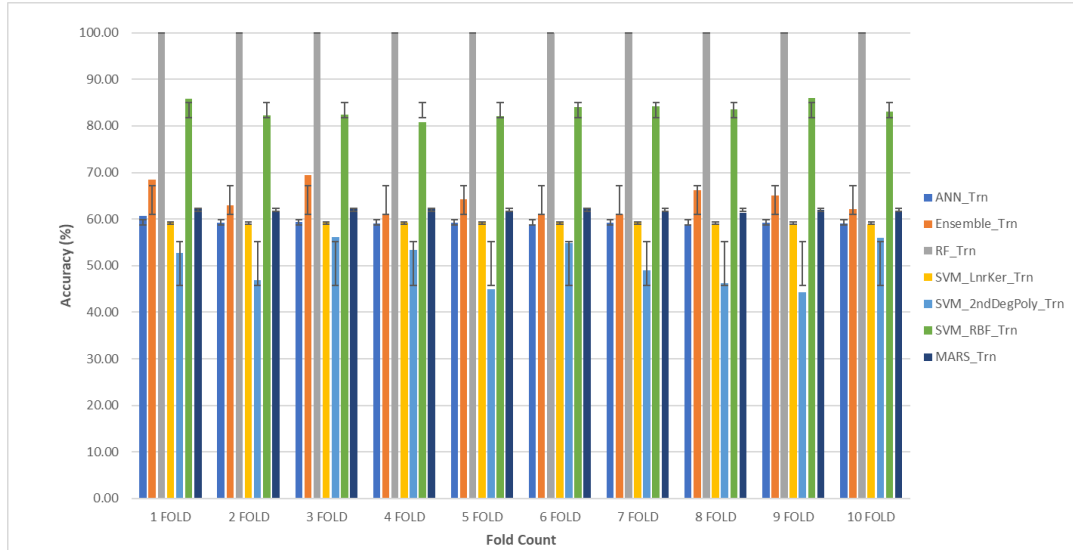


Figure 4.33. Training data accuracy for Direct\_PCA\_minCor dataset with standard deviation error bars (11 Image Bands).

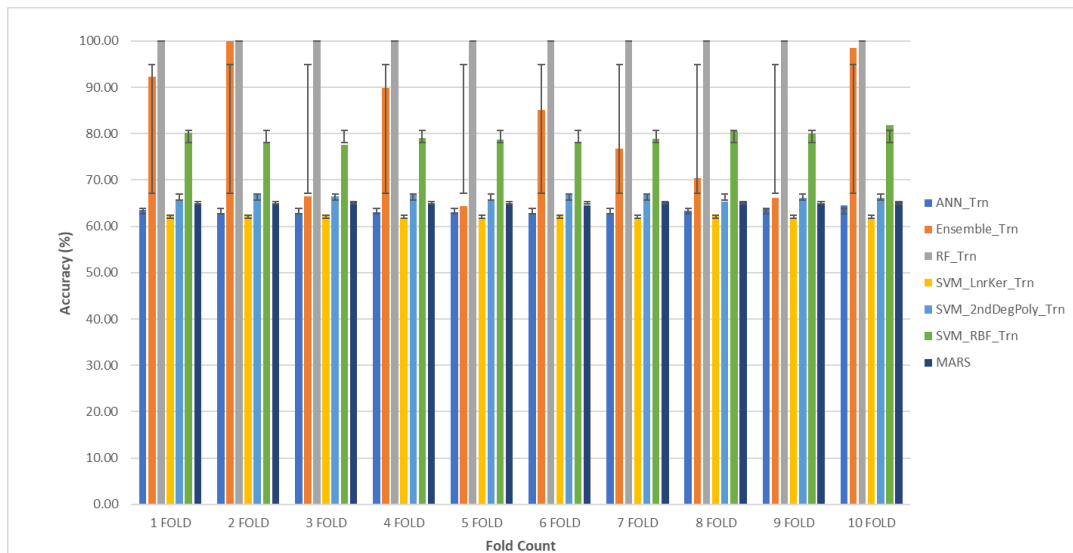


Figure 4.34. Training data accuracy for Ref\_Direct\_PCA\_minCor dataset with standard deviation error bars (15 Image Bands).

Accuracy results for the test data for every fold for both datasets Direct\_PCA\_minCor and Ref\_Direct\_PCA\_minCor are as follows in Figure 4.35 and Figure 4.36.

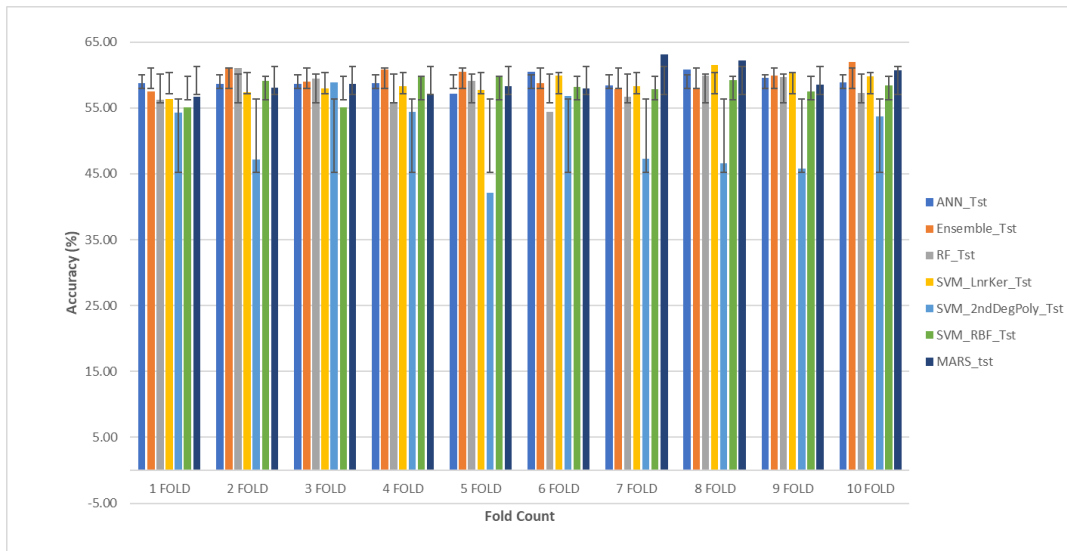


Figure 4.35. Test data accuracy for Direct\_PCA\_minCor dataset with standard deviation error bars (11 Image Bands).

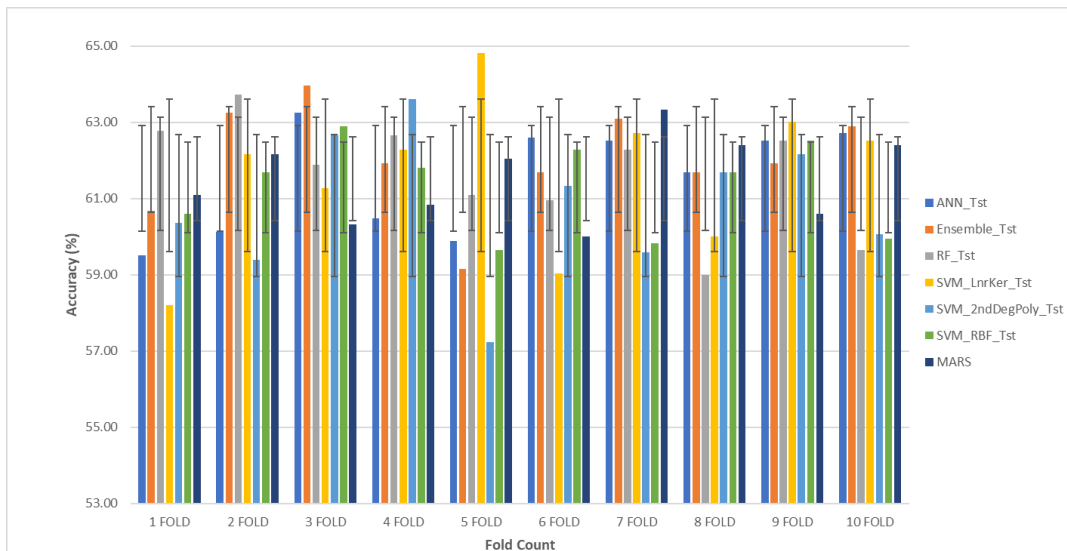


Figure 4.36. Test data accuracy for Ref\_Direct\_PCA\_minCor dataset with standard deviation error bars (15 Image Bands).

Figure 4.35 and Figure 4.36 clearly show that with the Ref\_Direct\_PCA\_minCor dataset, the SVM linear kernel ML algorithm produces the highest accuracy of 64.82% in the 5<sup>th</sup> Fold. While the accuracy for both datasets is observed to be around 60%, the lowest accuracy we have seen for any given dataset and ML algorithm with the maximum accuracy is with the Direct\_PCA\_minCor dataset, and the ML algorithm used is SVM 2<sup>nd</sup> Degree Polynomial with the accuracy of 58.92%. The total accuracy for the separated datasets is given in Table 4.3 and Table 4.4.

The attainment of the highest accuracy using the SVM linear kernel within Fold number 5's training partition underscores the efficacy of this specific model configuration. Such results carry significant implications for the robustness and discriminative power of the SVM linear kernel, particularly within the context of geospatial analysis. The selection of Fold number 5 as the optimal training partition further suggests that this subset of the data encapsulates salient characteristics that are effectively captured and utilized by the SVM model. This finding may prompt further investigation into the underlying features and spatial distributions within Fold number 5.

Moreover, identifying Fold number 5 as the optimal training partition necessitates a deeper exploration of the spatial distribution of data points within this fold. The forthcoming depiction in Figure 4.37 offers a visual representation of how instances within Fold number 5 are geographically dispersed, shedding light on potential spatial clusters, outliers, or patterns that may correlate with the observed high accuracy of the SVM linear kernel.

The tables presented the accuracy results for various ML algorithms across eleven directional PCA image bands and instance images. The accuracy values, expressed as percentages, reflect the performance of each algorithm when applied to the corresponding image bands.

It is evident that the choice of algorithm and image band greatly influences model performance. Specific algorithms demonstrate consistent performance across multiple image bands, while others exhibit more variability.

Table 4.6. ML algorithms accuracy percentage for Direct\_PCA\_minCor dataset

FOLD_ID				SVM	SVM 2 <sup>nd</sup>	SVM RBF	
	ANN	ENSEMBLE	RF	Linear Kernel	Degree Polynomial	(Gaussian Kernel)	MARS
FOLD1	58.80	57.47	56.27	56.39	54.28	55.06	56.75
FOLD2	58.67	61.20	61.04	57.35	47.23	59.16	58.07
FOLD3	58.67	58.99	59.40	57.95	58.92	55.06	58.62
FOLD4	58.80	60.84	55.78	58.31	54.46	59.88	57.11
FOLD5	57.11	60.48	59.11	57.71	42.17	59.76	58.31
FOLD6	60.48	58.80	54.46	59.88	56.87	58.19	57.95
FOLD7	58.38	57.83	56.75	58.26	47.29	57.90	63.09
FOLD8	60.84	58.07	59.88	61.57	46.63	59.28	62.17
FOLD9	59.52	59.88	59.64	60.24	45.78	57.47	58.55
FOLD10	58.87	62.00	57.23	59.83	53.73	58.38	60.72

Table 4.7. ML algorithms accuracy percentage for Ref\_Direct\_PCA\_minCor dataset

FOLD_ID				SVM	SVM 2 <sup>nd</sup>	SVM RBF	
	ANN	ENSEMBLE	RF	Linear Kernel	Degree Polynomial	(Gaussian Kernel)	MARS
FOLD1	59.52	60.68	62.77	58.19	60.36	60.60	61.08
FOLD2	60.12	63.25	63.73	62.17	59.40	61.69	62.17
FOLD3	63.25	63.98	61.88	61.28	62.65	62.89	60.31
FOLD4	60.48	61.93	62.65	62.29	63.61	61.81	60.84
FOLD5	59.88	59.16	61.08	64.82	57.23	59.64	62.05
FOLD6	62.61	61.69	60.96	59.04	61.33	62.29	60.00

---

FOLD7	62.53	63.09	62.29	62.73	59.59	59.83	63.33
FOLD8	61.69	61.69	58.99	60.00	61.69	61.69	62.41
FOLD9	62.53	61.93	62.53	63.01	62.17	62.53	60.60
FOLD10	62.73	62.89	59.64	62.53	60.07	59.95	62.41

---

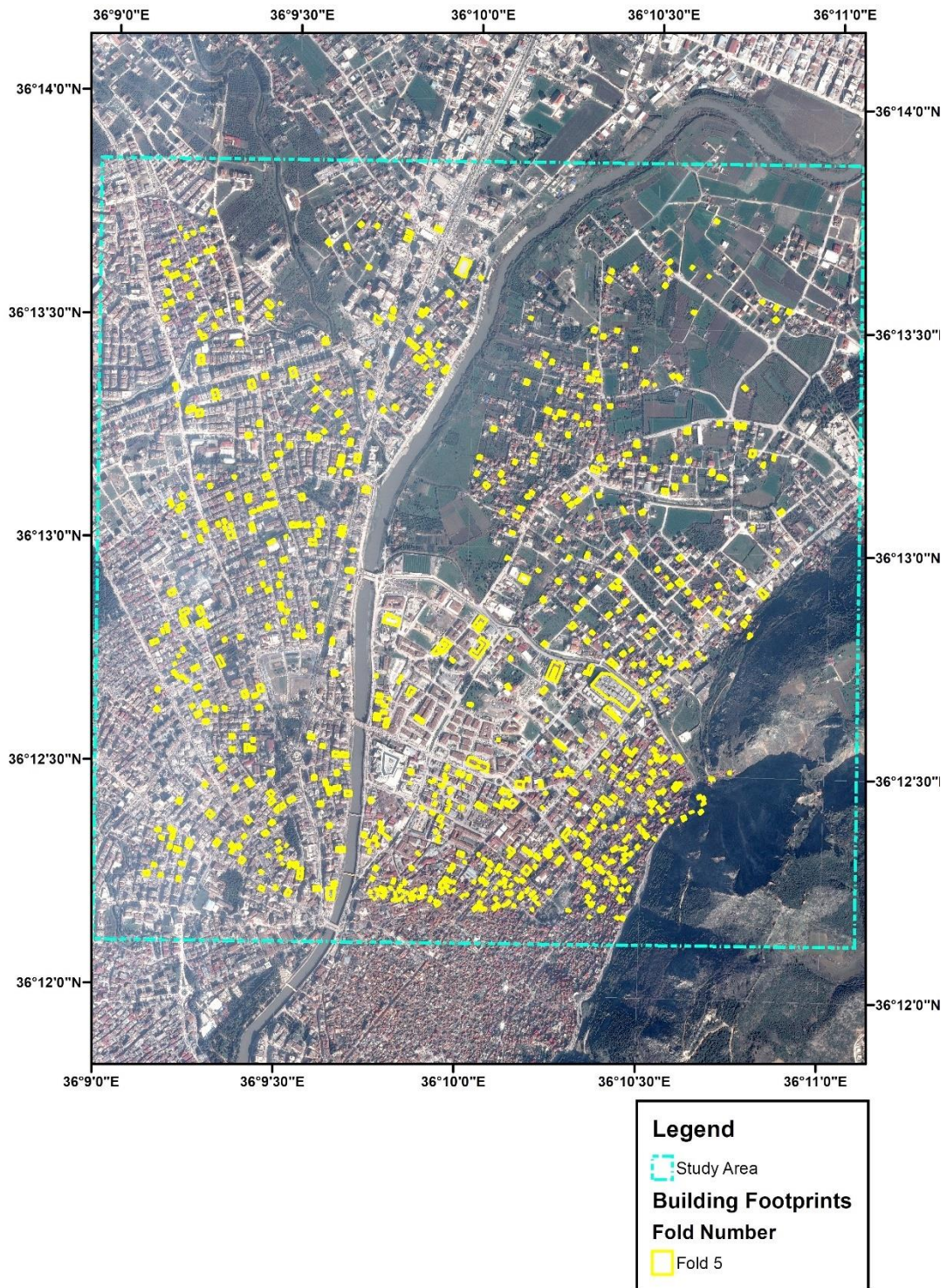


Figure 4.37. The distribution of the buildings in Fold number 5 in the study area.



## 4.1 Discussions

The findings presented in this study offer valuable insights into the application of ML algorithms for building damage assessment using satellite imagery. A robust framework for automated building damage classification in post-disaster scenarios is developed through rigorous data preprocessing, feature selection, and model evaluation.

One of the key observations from the analysis is the importance of feature selection in optimizing model performance. Identifying and prioritizing texture image bands with minimal correlations between damaged and undamaged building classes aims to enhance the discriminative power of the machine-learning models. The results demonstrate that leveraging a subset of texture image bands derived from PCA alongside reflectance image bands yields promising accuracy and predictive capability outcomes.

Furthermore, using ten-fold cross-validation enabled a comprehensive assessment of model generalization across diverse datasets. This approach mitigates biases associated with single train-test splits and provides valuable insights into the robustness and stability of the models across different data partitions. Identifying Fold number 5 as the optimal training partition highlights the spatial heterogeneity within the dataset and underscores the importance of considering geographic variability in model training.

The choice of ML algorithm also emerged as a critical factor influencing model performance. The results indicate that SVM with a linear kernel exhibits superior accuracy compared to other algorithms, mainly when applied to the Ref\_Direct\_PCA\_minCor dataset. This finding suggests that the linear separation boundaries imposed by the SVM algorithm align well with the underlying patterns in the dataset, resulting in the more effective classification of building damage.

#### 4.1.1 Consistency Evaluation of Machine Learning Algorithms

The performance evaluation of ML algorithms is crucial for understanding their reliability and effectiveness in classifying building footprints based on satellite imagery. This section presents the consistency of various algorithms across different folds of the dataset. The aim is to uncover insights into their stability and generalization capabilities.

Firstly, the training and test accuracies attained by a diverse set of ML algorithms are analyzed across ten folds of the dataset. Notably, algorithms like SVM with linear kernel and MARS exhibit relatively stable performance across folds, maintaining consistent accuracies around 60% to 62%. In contrast, algorithms such as SVM with 2nd-degree polynomial kernel and ensemble methods show more variability in their accuracies across folds.

Comparing the accuracies achieved on the test dataset reveals further insights into algorithmic performance. Again, SVM with linear kernel and MARS demonstrate consistent performance across folds, indicating their ability to generalize learned patterns effectively, with accuracies ranging from 57% to 63%. Conversely, SVM with 2nd-degree polynomial kernel and ensemble methods display more pronounced variability in their test accuracies, ranging from 56% to 64%.

The observed consistencies and variabilities provide valuable insights into the robustness and generalization capabilities of the algorithms. While some algorithms maintain stable accuracies across folds, others exhibit sensitivity to dataset fluctuations or task complexities. These observations underscore the importance of thoroughly evaluating and fine-tuning algorithms for geospatial analysis tasks.

Expanding the analysis to include the Ref\_Direct\_PCA\_minCor dataset reinforces these findings. Despite a broader feature space, algorithms still face challenges in generalization, as evidenced by fluctuations in test accuracies. Notably, the RF

algorithm demonstrates perfect training accuracy but struggles with variability in test performance, indicating potential overfitting issues.

#### **4.1.2 Challenges and Considerations in Building Footprint Classification from Non-Nadir Imagery**

The fact that the images are not orthorectified, meaning they are not captured from a nadir viewpoint, introduces additional challenges and considerations in the classification process. When images are captured from non-nadir viewpoints, distortions such as perspective and relief displacement can occur, impacting objects' appearance and spatial relationships in the imagery.

In building footprint classification, non-nadir images can lead to distortions in building shapes and sizes and variations in shadow lengths and directions. These distortions can introduce inconsistencies in feature extraction and classification, potentially affecting the accuracy and reliability of ML algorithms.

Furthermore, occlusions and overlapping objects in non-nadir images can complicate the delineation of building footprints, as parts of buildings may be obscured or hidden. This can result in incomplete or inaccurate feature representations, hindering classification.

Intricacies stem from various factors, the foremost being the rectification process in areas ravaged by earthquakes, where debris and rubble clutter thoroughfares. Obtaining viable rectification points with ground elevation poses a formidable challenge. Moreover, seismic events often induce ground shifts, necessitating meticulous computations to ascertain their extent, directly impacting the image's coordinate systems. Regrettably, an official comprehensive calculation to determine the absolute shift in the study area was unavailable during this study.

Furthermore, earthquakes can result in the tilting or displacement of buildings, as observed in satellite imagery, thereby deviating from their intended footprints. Addressing these shifts and deviations necessitates precise adjustments to the

vector data, aligning them with the corresponding satellite images. Such adjustments were executed based on a minimum neighborhood scale to ensure accuracy and fidelity to the ground truth.

In summary, the rectification process and vector data editing in earthquake-affected regions are multifaceted, influenced by the challenges of debris-laden terrain, ground shifts, and structural distortions. These complexities underscore the criticality of meticulous data processing techniques to maintain the integrity and accuracy of geospatial analyses in disaster-stricken areas.

#### **4.1.3 Challenges in Shadow Detection for Building Footprints Due to Complex Building Shapes and Density**

This study encountered notable challenges regarding detecting and utilizing shadows for building footprint extraction. This issue arises primarily due to the intricate shapes and high density of buildings within the study area. In densely populated urban environments characterized by closely situated buildings, shadows cast by neighboring structures often intersect or overlap, making it challenging to delineate individual building footprints based on shadow information alone accurately.

The proximity of buildings to one another exacerbates the complexity of shadow detection. When buildings are positioned closely, their shadows can blend or merge, resulting in ambiguous shadow patterns that are challenging to interpret accurately. In such cases, it becomes increasingly more difficult to discern the boundaries of individual buildings solely by analyzing their shadows, leading to inaccuracies in the building footprint extraction process.

Furthermore, the shape variability of buildings further complicates shadow detection efforts. Buildings with irregular shapes, such as L-shaped or T-shaped structures, may cast shadows that deviate from typical geometric patterns, making them less suitable for automated shadow detection algorithms. As a result, relying

solely on shadow information to extract building footprints becomes impractical and less effective in environments characterized by diverse building shapes and configurations.

Given the limitations posed by shadow detection in this study, a comprehensive approach incorporating vector data provided by the General Directorate of Geographical Information Systems is followed. The vector data were modified to align with the actual building configurations observed in satellite imagery to address the challenges associated with complex building shapes and dense urban landscapes. This process involved refining and adjusting the vector data to accurately represent the footprints of individual buildings, considering their irregular shapes and proximity to neighboring structures. This tailored approach allowed us to mitigate the limitations of shadow detection and achieve more precise results in delineating building footprints within the study area.

The approach to addressing the challenges posed by shadow detection in this study aligns closely with findings from prior research, underscoring its resonance with established methodologies in the field. This tailored strategy, which involved modifying vector data to represent building footprints accurately, resonates with the recommendations put forth by (Angulo-Saucedo *et al.*, 2022), (Dikmen, 2014), and (Turker and San, 2004). In their respective studies, similar adjustments to vector data were made to accommodate the complexities of urban landscapes and ensure the fidelity of building footprint extraction processes. By adhering to these established practices and refining vector data to better align with observed building configurations, we could navigate the limitations of shadow detection and enhance the methodology's accuracy, consistent with the broader trajectory of research in the field.

#### 4.1.4 Examining Variability in Building Footprint Areas Across Folds

In extending the analysis of building footprint areas across folds, additional statistical measures were incorporated to include the standard deviation alongside the minimum and maximum values to better understand the spatial distribution and variability of building sizes among the folds.

Upon examination of the standard deviation values, significant disparities in the degree of dispersion or spread of building footprint areas around the mean within each fold were observed. Notably, Fold 2 has the highest standard deviation of 626.10 m<sup>2</sup>, indicating considerable variability in building sizes across this fold. Conversely, Fold Nine exhibits the lowest standard deviation of 186.61 m<sup>2</sup>, suggesting a more uniform distribution of building sizes within this fold.

Of particular interest is Fold 5, previously noted for its diverse building sizes. Consistently, it maintains its significance when considering standard deviation, with a notable value of 556.74 m<sup>2</sup>. This underscores the pronounced variability in building footprint areas within Fold 5, highlighting its unique characteristics within the dataset. The distributions for the minimum, maximum, and standard deviation of the study area where the area of each building footprint is presented can be seen in Figure 4.38 to Figure 4.40.

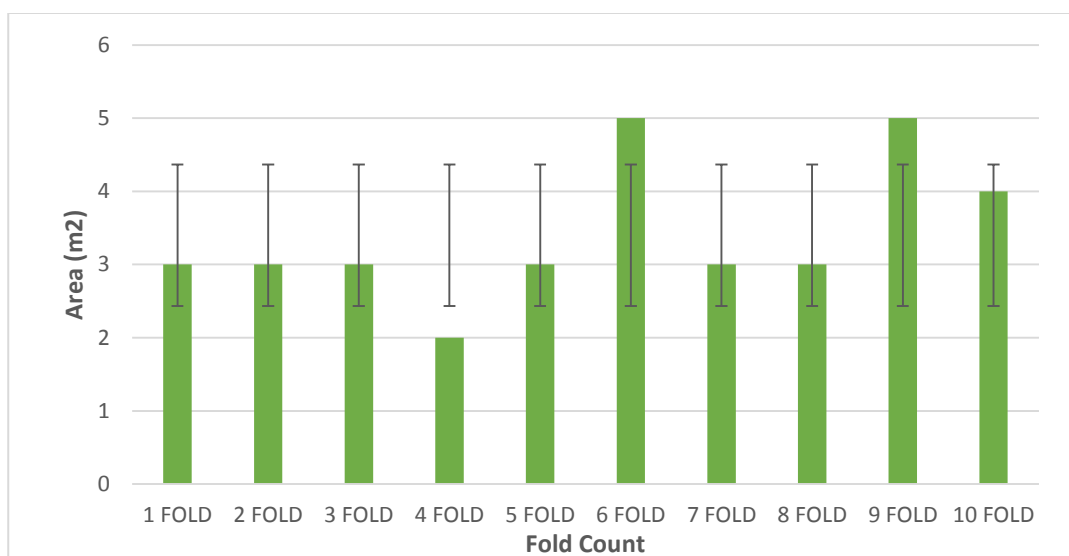


Figure 4.38. Area distribution (MIN) with standard deviation error bars across folds.

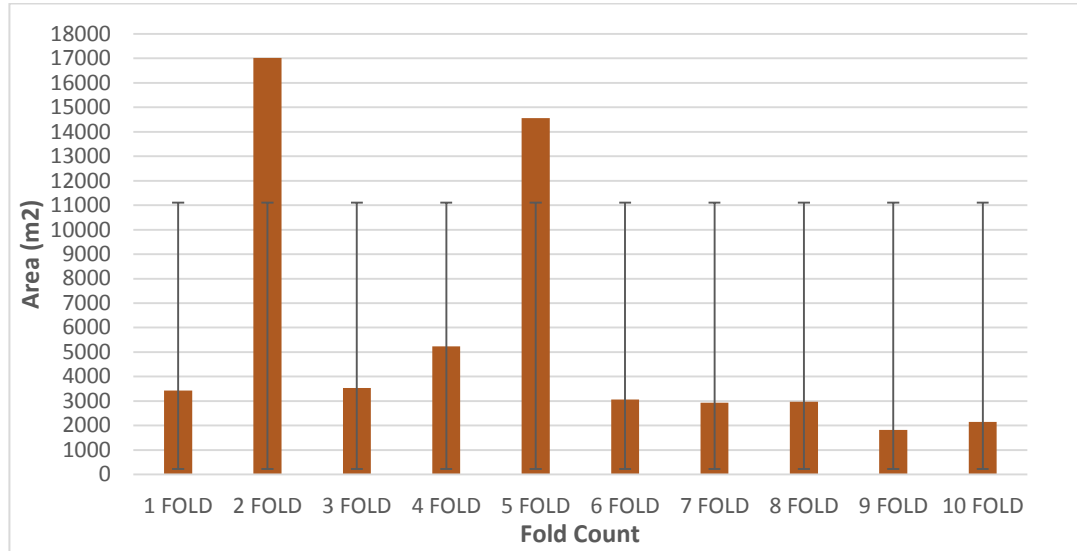


Figure 4.39. Area distribution (MAX) with standard deviation error bars across folds.

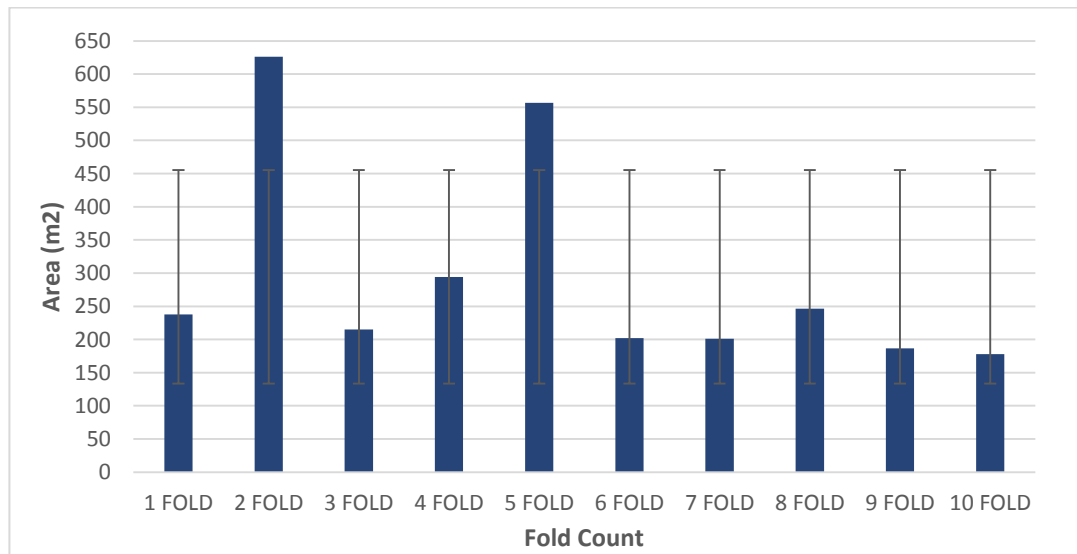


Figure 4.40. Standard deviation of the area distribution with standard deviation error bars across folds.

Furthermore, comparing standard deviation values across folds provides insights into the spatial heterogeneity of building sizes within the study area. Folds with

higher standard deviations, such as Folds 2 and 5, likely encompass regions with more pronounced variations in building sizes, which may pose challenges for accurate classification.

Conversely, folds with lower standard deviations, such as Fold 9, may indicate more homogeneous building size distributions, potentially simplifying the classification task. However, the nuanced relationship between standard deviation and classification accuracy warrants further investigation to comprehensively elucidate spatial variability's impact on ML model performance.

In summary, incorporating standard deviation alongside minimum and maximum values enhances understanding of the distribution and variability of building footprint areas across folds.

#### **4.1.5 Computational Efficiency Analysis of Machine Learning Models**

The computational efficiency of ML models plays a crucial role in determining their practical applicability, particularly in scenarios involving large-scale datasets and complex algorithms. This section explores the CPU times associated with various ML models applied to two datasets: Direct\_PCA\_minCor and Ref\_Direct\_PCA\_minCor.

Analyzing the CPU times across different models reveals notable variations in computational efficiency, influenced by both the model architecture and the dataset's characteristics.

Across the board, models with fewer predictors generally exhibit lower CPU times compared to those with a higher number of predictors. For instance, SVM with linear kernels and the ensemble model for the Direct\_PCA\_minCor dataset demonstrate relatively low CPU times, ranging from 0.77 to 2.94 seconds. These findings align with expectations, as linear models typically involve more straightforward computations and are computationally less intensive than models with polynomial or radial basis function (RBF) kernels. CPU times for the given



dataset and ML algorithm can be seen in Figure 4.41, which compares the two used datasets. The computational power harnessed in this study was anchored by a computer with an AMD Ryzen 7 5800H processor clocked at 3.20 GHz, coupled with 16 GB 3200 MHz of RAM, running on a Windows 10 64-bit environment.

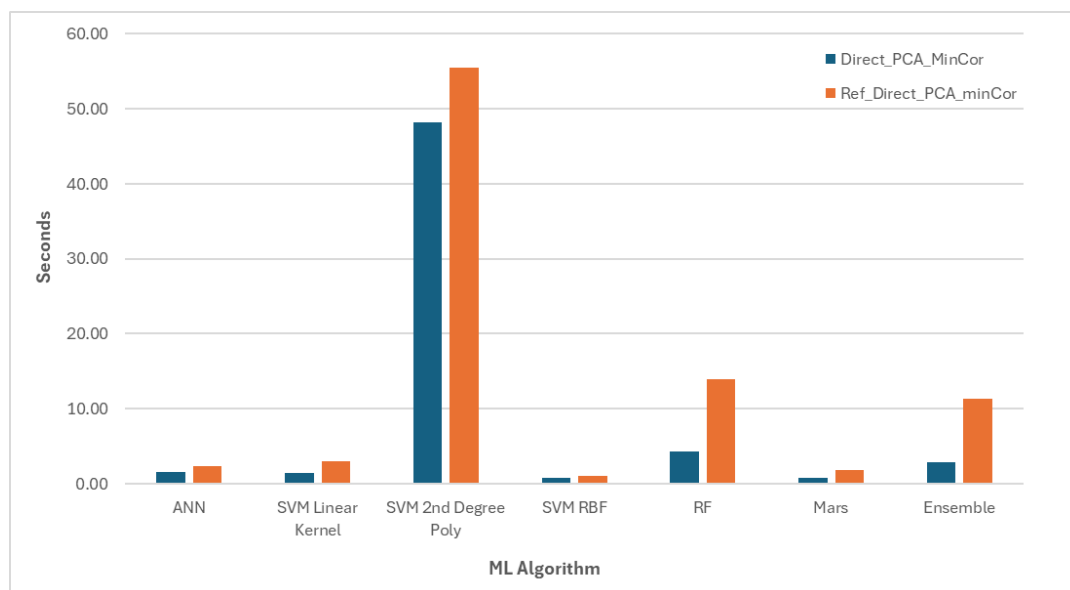


Figure 4.41. CPU runtimes for ML algorithms.

For the Direct\_PCA\_minCor dataset, the Support Vector Machine (SVM) with a linear kernel is the most computationally efficient model, with a CPU time of 1.41 seconds. This indicates that SVM with a linear kernel balances computational speed and predictive performance on this dataset.

However, when considering more complex models, such as SVM with a 2nd-degree polynomial kernel, the computational time increases significantly to 48.21 seconds. This substantial increase in CPU time suggests that the model's complexity introduces computational overhead, impacting its efficiency.

In contrast, RF and ensemble models exhibit moderate CPU times of 4.26 seconds and 0.77 seconds, respectively, indicating reasonable efficiency for models leveraging ensemble learning techniques.

Upon transitioning to the Ref\_Direct\_PCA\_minCor dataset, which includes additional information from reflectance images, a general increase in CPU times across all models compared to the Direct\_PCA\_minCor dataset is observed. For instance, the CPU time for SVM with a linear kernel rises to 2.94 seconds, highlighting the impact of dataset complexity on computational efficiency. Similarly, SVM with a 2nd-degree polynomial kernel experiences a substantial increase in CPU time to 55.51 seconds, underscoring the computational demands of higher-dimensional datasets.

Interestingly, despite the dataset's increased complexity, the models' relative performance in terms of computational efficiency remains consistent. SVM with a linear kernel remains one of the most efficient models. In contrast, SVM with a 2nd-degree polynomial kernel exhibits the highest CPU time, emphasizing the trade-off between model complexity and computational efficiency.

Overall, the analysis of CPU times provides valuable insights into the computational demands of ML models across different datasets. Understanding these computational characteristics is essential for selecting appropriate models based on computational resources and application requirements.

#### **4.1.6 Comparison with Existing Literature**

The study significantly contributes to the field by evaluating various ML algorithms and employing a unique methodology tailored for post-earthquake building damage assessment. While existing literature often focuses on individual algorithms or specific data sources, the research adopts an interdisciplinary approach, integrating RS data from Planet Pleiades with advanced ML techniques.

In contrast to traditional methods that rely solely on ground surveys or manual interpretation of satellite imagery, the methodology leverages the high-resolution capabilities of Pleiades imagery to provide detailed and accurate assessments of building damage (Airbus, 2023). By transforming raw DN values into reflectance

and subsequently calculating TOA spectral radiance and reflectance values (Chander *et al.*, 2009), the reliability and consistency of the data for analysis were ensured.

Furthermore, the study incorporates a comprehensive feature set derived from satellite imagery, including texture layers generated using PCA and directional texture bands (Haralick *et al.*, 1973). This multi-dimensional approach enhances the discriminative power of the analysis, enabling the detection of subtle variations in building conditions post-earthquake. Unlike previous studies that may focus on a limited set of features or neglect the importance of texture information, the methodology captures a broader range of spatial and spectral characteristics, thereby improving the accuracy and reliability of damage detection.

The variability inaccuracies observed in studies utilizing ML algorithms for automatic classification, such as the 55% accuracy achieved for discriminating between undamaged buildings by Naito *et al.* (2020), highlight the multifaceted nature of building footprint classification. Factors like dataset diversity, methodological intricacies in feature extraction and model training, and the inherent difficulty distinguishing between different building classes contribute to this variance. Underscoring the dataset with 20 cm aerial photos compared to 50 cm GSD satellite images, the classification results with a lower quality image of 64.82% is significantly higher.

While the study demonstrates promising accuracies achieved through the utilization of ML algorithms such as ANN, ensemble, RF, SVM with linear 2<sup>nd</sup> Degree Polygon and RBF kernels, and MARS, it is essential to acknowledge the inherent challenges in accurately separating damaged buildings from undamaged ones solely based on the inputs utilized in this research. Despite the advancements in supervised learning techniques, including the notable accuracies showcased, the limitations of the input data must be considered. While valuable, RS imagery and ancillary data may only capture some pertinent features necessary for precise classification. Factors like varying environmental conditions, alterations in building

materials, and the intricate nature of damage patterns contribute to the complexity of achieving complete separation. Moreover, the dynamic urban environment adds further complexity, with temporary structures, debris, and other elements potentially obscuring accurate damage assessment.

#### **4.1.7 Comparison of Highest Accuracy Fold with Damage Proximity Map**

Figure 4.37 depicts the SVM's highest accuracy, especially when the classification of Fold 5 is conducted using Fold 1 through Fold 4 and Fold 6 through Fold 10 as training data. Leveraging the ARIA Damage Proximity Map (DPM), the highest accuracy fold was compared to explore potential consistency in fold selection with high damage areas identified by the DPM. This comparison is illustrated in Figure 4.42.

Additionally, as described in Section 3.2.3, a damage proximity map shows potentially damaged areas with a color code from yellow to red, indicating the potential damage's severity (Yun *et al.*, 2015). This study has classified the four building damage classes into two, described in Section 3.3.4. As the damage proximity map color scheme from 0-255 provides the damaged areas, the assumption was made to classify the damage proximity map into two classes, which are the no-damage class, denoted by a pixel value of zero, and the damaged class, denoted by pixel values higher than zero. This assumption is made due to the pixel values other than zero holding even the smallest possible damaged area information, which in this study corresponds to the classification method used for buildings.

Figure 4.42 shows that significant building footprints coincide with the DPM's valid data zone, while smaller areas fall within the no-data zone. This observation provides valuable insights: buildings located within regions of valid DPM values are more likely to have incurred damage, whereas those within no-data areas lack

valid information about damage, thus contributing to a more comprehensive classification assessment.

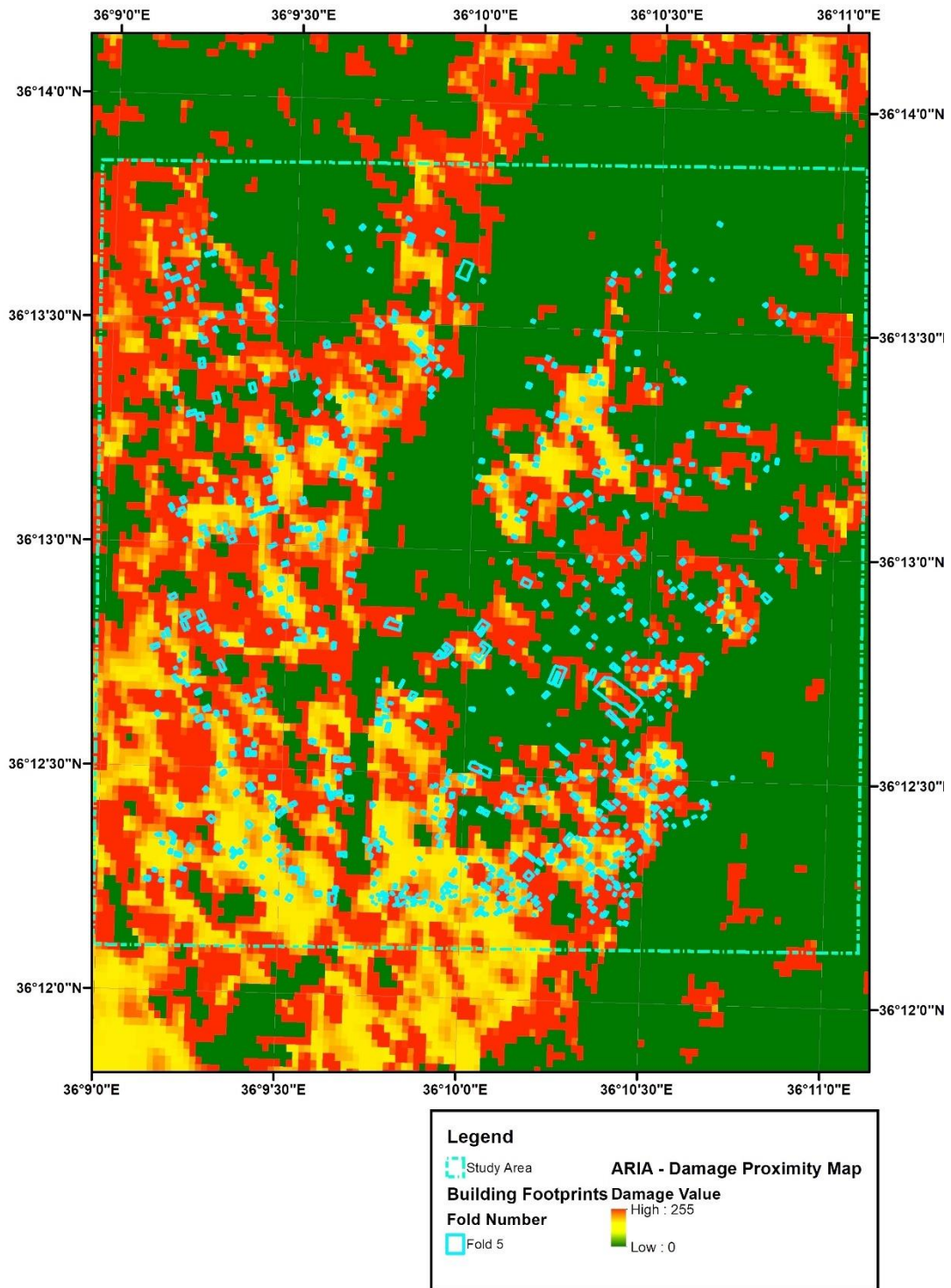


Figure 4.42. The comparison of fold number five to DPM

ML classification using the SVM linear kernel algorithm with the fifth fold, which has the highest accuracy with the Ref\_Direct\_minCor dataset, 294 out of 830 buildings were classified in accordance with the vector data acquired from the GD of GIS (2023). Comparing these results to the DPM, 246 buildings were classified in accordance with both the vector data acquired from the GD of GIS (2023) and DPM, where the DPM value shows zero.

The ML classification utilizing the SVM linear kernel algorithm, aligned with vector data obtained from GD of GIS (2023), identified 126 buildings as damaged, with the fifth fold demonstrating correct alignment. Compared with the Damage Proxy Map (DPM), only 29 out of these 126 buildings fall within the proximity of damage outlined by the DPM, where the DPM value exceeds zero.

On the other hand, 245 buildings out of 830 were wrongly classified, where the ML algorithm classified them as having no damage, and the vector data acquired from GD of GIS (2023) states that they are damaged. Among these 245 misclassified buildings, 200 correctly align with the DPM proximity data, where the DPM value is zero. Forty-five of these buildings' DPM values are more significant than zero, which aligns with the acquired vector data.

With a similar approach, 165 out of 830 buildings were classified as damaged, where the vector data acquired from GD of GIS (2023) states that they have no damage. 29 of these 165 buildings align with the DPM value where the DPM value for the building footprints is more than zero. One hundred thirty-six of these buildings align with the vector data acquired from the GD of GIS (2023), where their DPM value shows a mean value in their footprint area as zero. The confusion matrix for the SVM linear kernel ML algorithm with the 5<sup>th</sup> fold is presented in Table 4.8, and the confusion matrix for the DPM regarding the 5<sup>th</sup> fold is presented in Table 4.9.

Table 4.8. Confusion matrix for the 5<sup>th</sup> fold of SVM linear kernel ML algorithm building classification with respect to vector data provided by GD of GIS (2023).

	<b>Predicted Negative</b>	<b>Predicted Positive</b>
<b>Actual Negative</b>	294	165
<b>Actual Positive</b>	245	126

Table 4.9. Confusion matrix for the 5<sup>th</sup> fold of SVM linear kernel ML algorithm building classification with respect to DPM.

	<b>Predicted Negative</b>	<b>Predicted Positive</b>
<b>Actual Negative</b>	446	233
<b>Actual Positive</b>	93	58

The table provided presents a confusion matrix, a fundamental tool for evaluating the performance of classification models. The table titles categorize the model's predictions and the actual ground truth into four quadrants. "Predicted Negative" and "Predicted Positive" denote the instances predicted by the model, while "Actual Negative" and "Actual Positive" represent the true status of the instances according to the ground truth in this study; in this case, the ground truth is the data acquired from the GD of GIS (2023).

Evaluating the classification model's performance provides valuable insights into its effectiveness in accurately predicting instances within the dataset. This discussion aims to compare and analyze the metrics obtained from two sets of experiments, highlighting the observed differences and implications for the model's practical applicability. The metrics, including accuracy, precision, and recall, were



derived from a confusion matrix, a tabular representation of the model's predictions compared to the ground truth labels. The confusion matrix organizes the predictions into four categories: true positives (correctly predicted positive instances), true negatives (correctly predicted negative instances), false positives (incorrectly predicted positive instances), and false negatives (incorrectly predicted negative instances). Examining these metrics makes it possible to understand the model's strengths and weaknesses in correctly classifying instances and avoiding misclassifications. This information is crucial for assessing the model's reliability and suitability for real-world applications.

In the first set of experiments, the model achieved an accuracy of 0.506, a precision of 0.432, and a recall of 0.339. This initial evaluation revealed a balance between precision and recall, indicating a moderate ability to identify positive instances while minimizing false positives correctly.

In contrast, the second set of experiments improved overall accuracy, reaching 0.607. However, this gain was accompanied by a significant trade-off between precision and recall. Precision decreased substantially to 0.199, suggesting a higher rate of false positives, while recall increased slightly to 0.384, indicating a modest improvement in capturing positive instances. This discrepancy suggests a tendency for the DPM to overclassify, as it inaccurately predicts a more significant number of negative instances as positive.

#### **4.1.8 Assessing the Ground Truth Vector Data**

In discussing the ground truth methodology employed in this research, it is essential to recognize its strengths and limitations. The ground truth data utilized in this study comprises a subset of information sourced from Hatay city, generated by GD of GIS (2023) between the 7th of February 2023 and the 11th of February 2023, using a UAV. This approach has its drawbacks. While it offers a rapid means of assessing building damage levels, the process may inadvertently introduce errors

and inaccuracies into the classification. One notable concern arises from the subjective nature of damage classification, which relies on the expertise and judgment of specialists conducting the assessments. The variability introduced by different specialists underscores the potential for inconsistencies and discrepancies in the classification data.

Moreover, the angle of observation by the UAV presents another challenge in accurately capturing building damage. Depending on the viewing angle, buildings may exhibit varying degrees of damage, complicating the classification process. Additionally, occlusions caused by structural elements or debris further impede the assessment of damage levels, particularly on specific sides of buildings. These factors contribute to the complexity and time constraints of generating accurate damage classification data. Furthermore, the distance between the UAV camera and the building being assessed during the data collection period introduces another source of potential error, as it can affect the resolution and clarity of the captured imagery. As such, careful consideration of these factors is paramount in interpreting and contextualizing the ground truth data utilized in this research.

#### **4.1.9 Exploring Factors Influencing Misclassification**

In the investigation of building damage detection accuracy, a notable finding emerged regarding the misclassification rate, particularly the 35.18% discrepancy observed in the results. This raises significant questions about the underlying factors contributing to this discrepancy and prompts a deeper examination of potential influences. The true negative and true positive classified buildings' distribution in the study area is plotted to see if there is any clustering in the study area regarding building classification. The true negative and positive distribution can be seen in Figure 4.43; similarly, false negative and positive can be seen in Figure 4.44.

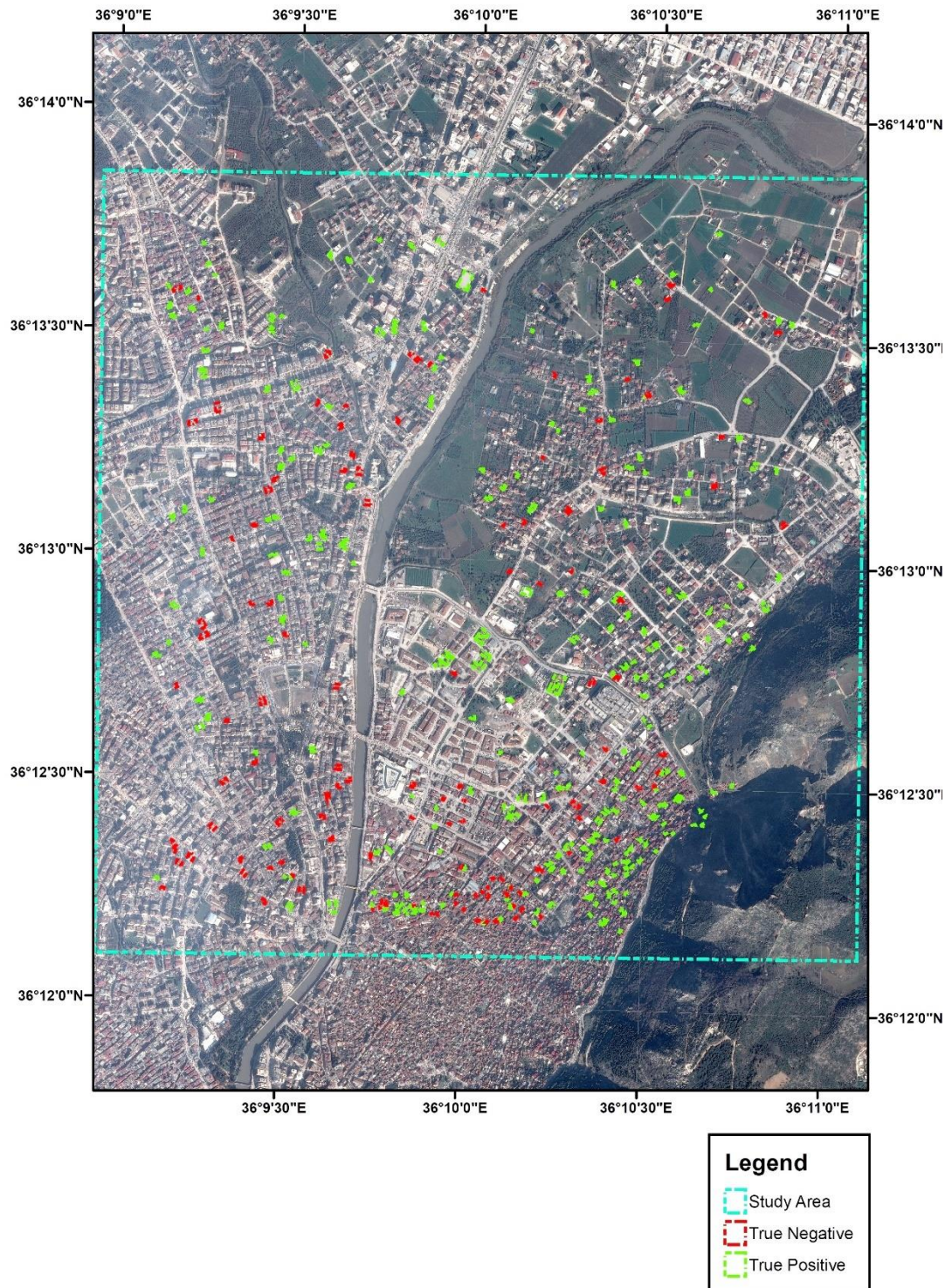


Figure 4.43. True negative and true positive distribution of the fold five building classification using SVM linear kernel.

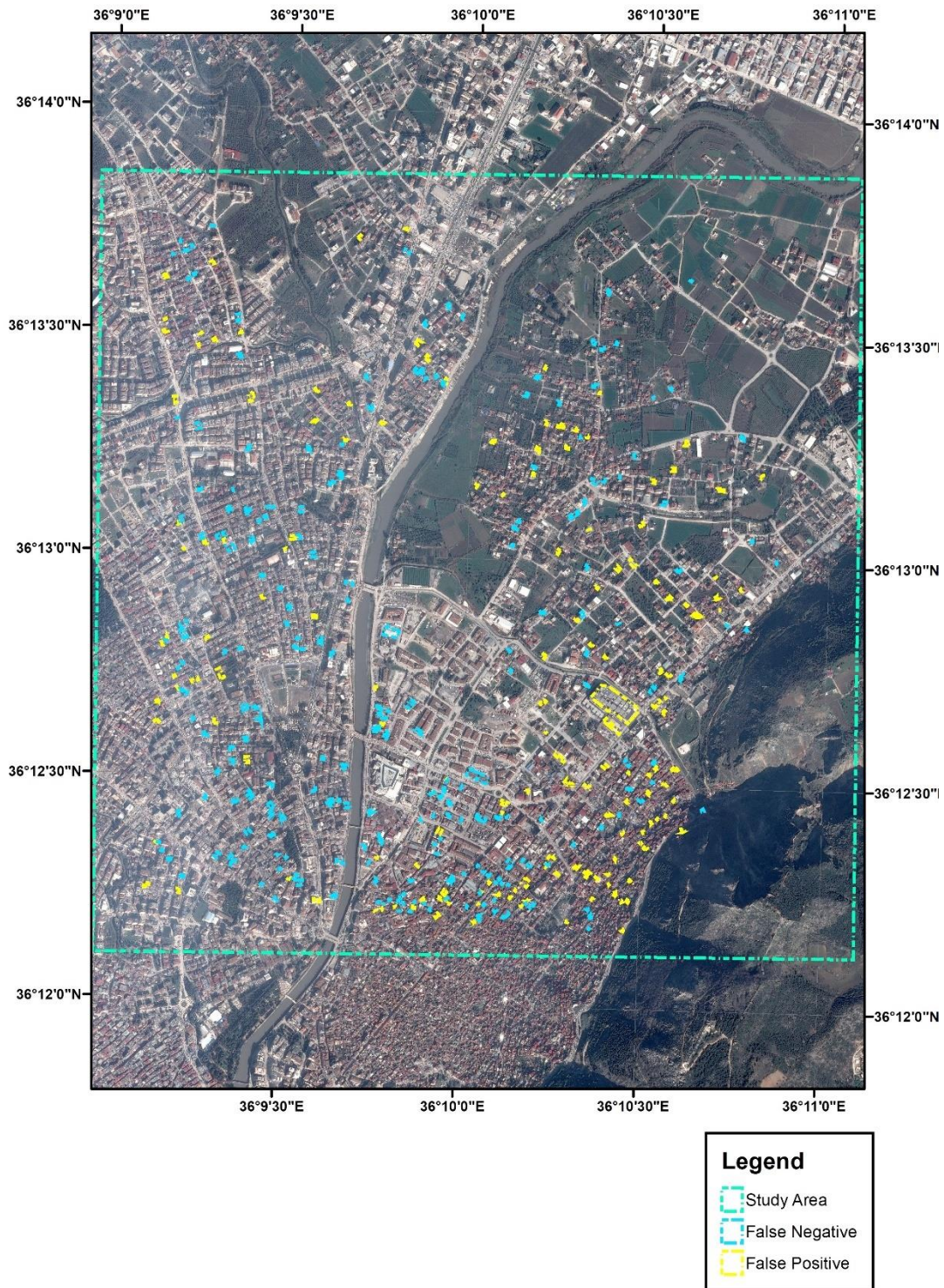


Figure 4.44. False negative and false positive distribution of the fold five building classification using SVM linear kernel.

Upon examining Figure 4.43 and Figure 4.44, it becomes evident that no clear correlation exists between building density and classification accuracy. The distribution of both true and false classifications appears reasonably even across the entire image. However, upon closer inspection of Figure 4.43, a notable clustering of correctly classified large buildings is observed in the southeast corner, warranting further investigation.

The study area's southeastern region comprises buildings close to each other, and shadows from nearby hills add complexity to the analysis. It becomes apparent that the footprint of each building aligns directly with the roof, which is visible in satellite imagery. This trend is consistent throughout the study area. The uniformity in the reflectance of building footprints significantly influences the classification results of the ML algorithm. Figure 4.45 provides visual examples of both true positive and true negative classifications.



Figure 4.45. True positive (green) and true negative (red) examples of uniform building reflectance.

Several generalizations emerge from the misclassifications observed in the dataset, particularly concerning false positive and false negative classifications. Misclassification due to tilted buildings presents a significant factor. Tilted buildings contribute to misclassifications in two main ways. Firstly, in false positive cases, tilted buildings that remain standing may exhibit a relatively uniform reflection within their building footprints, leading to misclassification. Secondly, in cases where a building has collapsed on one side but retains the reflective properties of an intact building, misclassification can occur. An example can be seen in Figure 4.46.

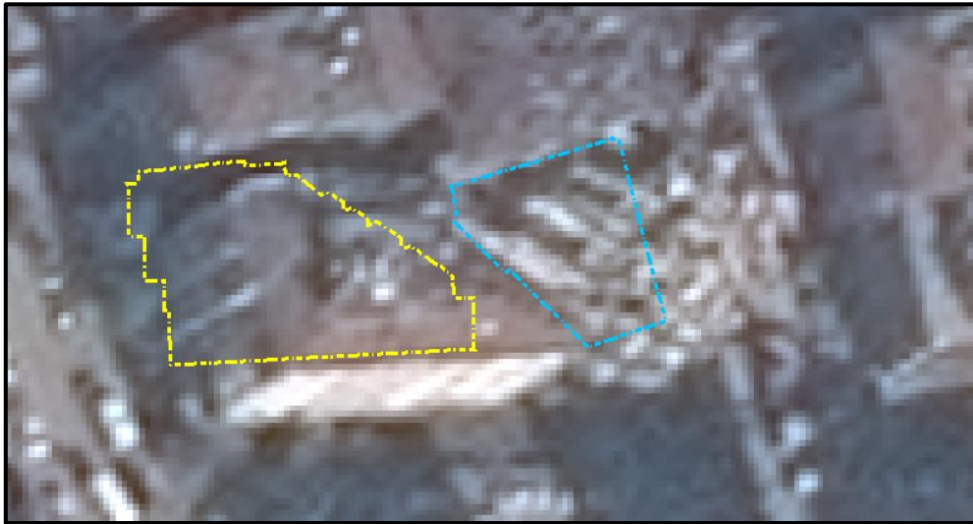


Figure 4.46. False positive (yellow) and false negative (blue) tilted building samples.

As previously discussed, achieving optimal classification of buildings relies on the uniformity of a building's rooftop reflectance values. Any object that disrupts this uniformity, such as vegetation or satellite dishes, can lead to false classifications by the ML algorithm. Additionally, while the area of a building's footprint does not directly impact the classification problem, architectural trends favoring larger building areas can introduce variations in rooftop uniformity, thus complicating the ML classification process. Sample images can be seen in Figure 4.47 and Figure 4.48.



Figure 4.47. Large area building with non-uniform building footprint reflectance.

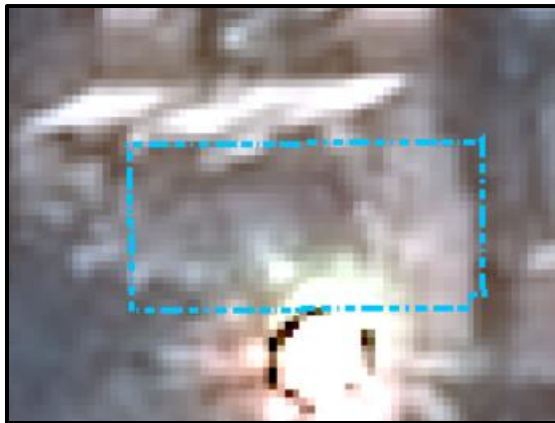


Figure 4.48. Satellite dishes reflect the sunlight, causing changes in the reflectance values of the building footprint.





## CHAPTER 5

### CONCLUSIONS AND RECOMMENDATIONS

In this study, earthquake damage assessment methodologies, particularly in densely populated urban regions, are studied. The devastating earthquakes that struck Turkey on February 6th highlighted emergency responders' critical challenges in swiftly evaluating and addressing the extensive damage inflicted upon buildings and infrastructure. Traditional methods, reliant on manual surveys and inspections, often struggle to cope with the scale and urgency of such disasters, prompting the exploration of alternative approaches.

Integrating RS technologies and GIS has emerged as a transformative response disaster management and response tool. By harnessing data from satellites and aerial platforms, coupled with advanced spatial analysis techniques, these technologies offer a rapid and comprehensive means of assessing the extent and severity of damage remotely. This enables timely decision-making for emergency responders and lays the foundation for effective disaster mitigation and urban planning strategies.

Moreover, recent advancements in ML present promising avenues for improving the accuracy and efficiency of damaged building detection from RS data. These ML algorithms can automatically extract relevant features from imagery, overcoming the limitations of traditional classification methods and achieving higher accuracy rates in identifying damaged structures.

This study contributes to the related literature by undertaking a comparative analysis of various machine-learning algorithms for post-earthquake building damage assessment utilizing post-earthquake satellite imagery exclusively. However, it is imperative to acknowledge the inherent challenges associated with this endeavor. One major challenge lies in acquiring or generating vector data for

building footprints. Notably, even within a relatively small area encompassing 8,298 buildings, the considerable time required to edit building footprints underscores this task's complexity. Additionally, tilted buildings within the scenery pose another significant challenge, necessitating corrective measures to represent building footprints accurately. Since pixel values within the building footprint are utilized to reconstruct various aspects, meticulous correction of the building footprint polygons becomes essential. In the current scenario, it is evident that RS images, even with centimeter-level precision, need to be revised to accurately identify damaged buildings solely from RGB datasets, as studies have shown results with only 60% accuracy. However, incorporating PCA and texture data significantly improves the results. It's crucial to highlight that achieving the desired accuracy in obtaining the four damage classes has proven challenging, with a maximum accuracy of only 64.82% achieved when classified as damaged and not damaged.

Furthermore, satellite image acquisition date and time are crucial from a raster data standpoint. Timeliness is paramount, especially following an earthquake, as the landscape changes rapidly. It is possible to observe instances where multiple buildings are removed from the scenery, rubble accumulates near tilted buildings, and other dynamic alterations occur.

Furthermore, the findings illuminate a crucial limitation: relying solely on RS data may not be adequate for binary detection of damaged and undamaged buildings. While RS data provides valuable insights, augmenting it with additional data sources and employing refined methodologies is imperative to enhance the accuracy of damage detection. For instance, pre-earthquake satellite images can serve as supplementary inputs for machine-learning algorithms alongside change detection outputs, enriching the analysis and improving detection accuracy.

The study highlights the critical role of sophisticated, data-driven methodologies in earthquake damage assessment. Integrating RS, GIS, and ML techniques aims to provide valuable insights and methods to expedite the detection of damaged

buildings. Additionally, this research offers guidance on decision-making during emergencies, facilitating more efficient response efforts.

Based on the analysis conducted in this study, several recommendations can be made for future research in building damage classification using satellite imagery and ML algorithms.

Firstly, future research could explore multi-class classification approaches to capture a broader spectrum of damage severity levels. While this study focused on the binary classification of building footprints, extending the analysis to include multiple classes could provide more nuanced insights into the extent and severity of damage following natural disasters. Emergency responders and policymakers can better prioritize resources and aid efforts in affected areas by categorizing building footprints into different damage levels.

Moreover, the generalization of the findings of this study to different geographic regions and disaster types remains an open question. The performance of ML models is contingent upon the availability and quality of training data, which may vary across different contexts. Therefore, further validation and refinement of the approach using diverse datasets are essential to ensure its applicability in real-world disaster response scenarios. Researchers should seek to collect and incorporate data from a wide range of geographic regions and disaster types to enhance ML models' robustness and generalization capabilities for building footprint classification. These datasets serve as the foundation for training and evaluating ML models. Researchers can more accurately assess their models' performance and generalization capabilities using labeled datasets for training and testing. This approach helps to identify potential biases, limitations, and areas for improvement, ultimately leading to more reliable and effective classification algorithms.

Additionally, future research could investigate integrating additional data sources and features to improve classification accuracy and reliability. Incorporating complementary datasets such as elevation data, land cover maps, or historical

imagery could provide valuable context and enhance the discriminatory power of ML models. Furthermore, exploring advanced preprocessing techniques and feature engineering approaches could help address challenges related to data quality, noise, and variability in building characteristics. Therefore, while the methodologies offer valuable insights and advancements, future research endeavors should explore incorporating additional data sources and more comprehensive feature extraction techniques to enhance the accuracy and robustness of damage detection algorithms.

Despite the promising results obtained in this study, several limitations and areas for future research should be acknowledged. Firstly, the reliance on texture image bands derived from PCA may introduce biases or artifacts into the analysis, warranting further investigation into alternative feature extraction techniques. Additionally, the classification of building damage into binary categories (damaged vs. undamaged) needs to be more accurate in the complexity of post-disaster scenarios, and nuanced variations in building conditions may be overlooked.

## REFERENCES

- Airbus. (2023). SIA01-WS. Defence and Space.
- Angulo-Saucedo, G. A., Leon-Medina, J. X., Pineda-Muñoz, W. A., Torres-Arredondo, M. A., and Tibaduiza, D. A. (2022). Damage Classification Using Supervised Self-Organizing Maps in Structural Health Monitoring. *22*(4), 1484.
- Basheer, I. A. and Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*, *43*(1), 3-31. doi:[https://doi.org/10.1016/S0167-7012\(00\)00201-3](https://doi.org/10.1016/S0167-7012(00)00201-3)
- Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5-32. doi:10.1023/A:1010933404324
- Chander, G., Markham, B. L., and Helder, D. L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, *113*(5), 893-903. doi:<https://doi.org/10.1016/j.rse.2009.01.007>
- Chen, X., Chen, Z., Wang, J., and Chen, C. W. (2021). A review of deep learning methods for remote sensing image classification %J Remote Sensing. *13*(1), 142.
- Cotrufo, S., Sandu, C., Giulio Tonolo, F., and Boccardo, P. (2018). Building damage assessment scale tailored to remote sensing vertical imagery. *European Journal of Remote Sensing*, *51*(1), 991-1005. doi:10.1080/22797254.2018.1527662
- Dikmen, M. (2014). A Shadow based trainable method for building detection in satellite images.
- Friedman, J. H. (1991). Multivariate Adaptive Regression Splines. *19 %J The Annals of Statistics*(1), 1-67, 67.
- Gamba, P. and Casciati, F. (1998). GIS and image understanding for near-real-time earthquake damage assessment. *Photogrammetric engineering remote Sensing Letters*, *64*, 987-994.
- GD of GIS. (2023). Hatay Vector. Retrieved from [https://basic.atlas.gov.tr/?\\_appToken=&metadataId=](https://basic.atlas.gov.tr/?_appToken=&metadataId=)

- Ge, P., Gokon, H., and Meguro, K. (2019). Building Damage Assessment Using Intensity SAR Data with Different Incidence Angles and Longtime Interval. *Journal of Disaster Research*, 14(3), 456-465. doi:10.20965/jdr.2019.p0456
- Grünthal, G. (1998). Conseil de l'Europe, Cahiers du Centre Europeen de Geodynamique et de Seismologie. *European Macroseismic Scale, Lüksemburg*.
- Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5), 786-804. doi:10.1109/PROC.1979.11328
- Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-3*(6), 610-621. doi:10.1109/TSMC.1973.4309314
- Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2): Springer.
- Janalipour, M. and Mohammadzadeh, A. (2016). Building Damage Detection Using Object-Based Image Analysis and ANFIS From High-Resolution Image (Case Study: BAM Earthquake, Iran). *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(5), 1937-1945. doi:10.1109/JSTARS.2015.2458582
- Kaur, S., Gupta, S., Singh, S., Hoang, V. T., Almakdi, S., Alelyani, T., and Shaikh, A. (2022). Transfer Learning-Based Automatic Hurricane Damage Detection Using Satellite Images. *11*(9), 1448.
- Kerle, N., Voigt, S., and Nex, F. (2020). Editorial: Machine learning for earth observation data: Models, methods, and big data management %J *Frontiers in Earth Science*. 8, 1-3.
- Kouchi, K. i., Yamazaki, F., Kohiyama, M., Matsuoka, M., and Muraoka, N. (2004). *Damage detection from QuickBird high-resolution satellite images for the 2003 Boumerdes, Algeria earthquake*. Paper presented at the Proc. 1st Asia Conference on Earthquake Engineering.
- Laurent Coeurdevey, I. Q. M. and Karine Fernandez, P. P. M. (2012). *Pléiades Imagery – User Guide* (USRPHR-DT-125-SPOT-2.0). Retrieved from France: <https://www.intelligence-airbusds.com/en/8718-user-guides>

- Liu, S. B. (2014). Crisis Crowdsourcing Framework: Designing Strategic Configurations of Crowdsourcing for the Emergency Management Domain. *Computer Supported Cooperative Work (CSCW)*, 23(4), 389-443. doi:10.1007/s10606-014-9204-3
- Ma, H., Liu, Y., Ren, Y., Wang, D., Yu, L., and Yu, J. (2020). Improved CNN Classification Method for Groups of Buildings Damaged by Earthquake, Based on High Resolution Remote Sensing Images. *12(2)*, 260.
- Ma, H., Liu, Y., Ren, Y., and Yu, J. (2020). Detection of Collapsed Buildings in Post-Earthquake Remote Sensing Images Based on the Improved YOLOv3. *12(1)*, 44.
- Menderes, A., Erener, A., and Sarp, G. (2015). Automatic Detection of Damaged Buildings after Earthquake Hazard by Using Remote Sensing and Information Technologies. *Procedia Earth and Planetary Science*, 15, 257-262. doi:<https://doi.org/10.1016/j.proeps.2015.08.063>
- Milborrow, S. (2011). Multivariate adaptive regression spline models. In: Retrieved from CRAN Project: <http://cran.rproject.org/web/packages/earth> ....
- Mittal, P. V. and Bafna, R. (2023). Unsupervised learning framework for region based damage assessment on xBD, a large satellite imagery.
- Naito, S., Tomozawa, H., Mori, Y., Nagata, T., Monma, N., Nakamura, H., . . . Shoji, G. (2020). Building-damage detection method based on machine learning utilizing aerial photographs of the Kumamoto earthquake. *36(3)*, 1166-1187. doi:10.1177/8755293019901309
- NASA/JPL-Caltech/USGS/JAXA/Earth Observatory of Singapore – Remote Sensing Lab. (2023). Satellites Assess Earthquake Damage in Turkey. Retrieved from <https://www.jpl.nasa.gov/images/pia25564-satellites-assess-earthquake-damage-in-turkey>
- Rao, A., Jung, J., Silva, V., Molinario, G., and Yun, S. H. (2023). Earthquake building damage detection based on synthetic-aperture-radar imagery and machine learning. *Nat. Hazards Earth Syst. Sci.*, 23(2), 789-807. doi:10.5194/nhess-23-789-2023
- Rathje, E. M., Kyu-Seok, W., Crawford, M., and Neuenschwander, A. (2005, 29-29 July 2005). *Earthquake damage identification using multi-temporal high-resolution optical satellite imagery*. Paper presented at the Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05.

- Regulation on the Production of Large-Scale Maps and Map Information*. (2005). Resmi Gazete: Türkiye Cumhuriyeti Retrieved from <https://www.resmigazete.gov.tr/eskiler/2005/07/20050715-5.htm>
- Takhtkeshha, N., Mohammadzadeh, A., and Salehi, B. (2023). A Rapid Self-Supervised Deep-Learning-Based Method for Post-Earthquake Damage Detection Using UAV Data (Case Study: Sarpol-e Zahab, Iran). *15*(1), 123.
- TRT. (2023). Türkiye'nin unutmayaacağı en acı tarih: 6 Şubat 2023. Retrieved from <https://www.trthaber.com/haber/gundem/turkiyenin-unutmayacagi-en-aci-tarih-6-subat-2023-786582.html>
- Turker, M. and San, B. T. (2004). Detection of collapsed buildings caused by the 1999 Izmit, Turkey earthquake through digital analysis of post-event aerial photographs. *International Journal of Remote Sensing*, *25*(21), 4701-4714. doi:10.1080/01431160410001709976
- UHUZAM. (2023). Kahramanmaraş Depremleri sonrasında bilimsel araştırmalara destek için İTÜ-UHUZAM tarafından araştırmacılarla paylaşılan uydu görüntüleri. Retrieved from <https://web.cscrs.itu.edu.tr/kahramanmaras-depremleri/>
- Vapnik, V., Golowich, S., and Smola, A. (1996). Support vector method for function approximation, regression estimation and signal processing. *Advances in neural information processing systems*, *9*.
- Yamazaki, F., Kouchi, K.-i., Kohiyama, M., Estrada, M., and Matsuoka, M. (2003). Remote Sensing Technology for Earthquake Damage Detection.
- Yang, J., Zhang, X., Zhang, D., and Zhang, X. (2019). A CNN-based method for damaged building detection in high-resolution remote sensing imagery %J IEEE Geoscience and Remote Sensing Letters. *16*(12), 1893-1897.
- Yang, W., Zhang, X., and Luo, P. (2021). Transferability of Convolutional Neural Network Models for Identifying Damaged Buildings Due to Earthquake. *13*(3), 504.
- Yun, S.-H., Hudnut, K., Owen, S., Webb, F., Simons, M., Sacco, P., . . . Coletta, A. (2015). Rapid Damage Mapping for the 2015 Mw 7.8 Gorkha Earthquake Using Synthetic Aperture Radar Data from COSMO–SkyMed and ALOS-2 Satellites. *Seismological Research Letters*, *86*(6), 1549-1556. doi:10.1785/0220150152 %J Seismological Research Letters
- Zhang, C. and Ma, Y. (2012). *Ensemble machine learning: methods and applications*: Springer.



Zhu, X., Tuia, D., Mou, L., Xia, G. S., Zhang, L., Xu, F., and Fraundorfer, F. (2020). Deep learning in remote sensing: A comprehensive review and list of resources %J IEEE Geoscience and Remote Sensing Magazine. 8(4), 8-36.

## APPENDICES

### **A. Histogram Charts of Texture Images Produced from Reflectance Image Bands**

The histograms demonstrate no direct separation between building damage classes using texture images produced from reflectance images. The reader is encouraged to inspect the histograms within each building damage class of the same texture image across all image bands. It is seen that after the inspection, the four building damage classes show the same trend, with the peak points and the curvatures being the same. This comparison allows the inspector to discern variations between different image bands featuring the same texture image. Notably, no distinct differentiation is observed among damage classes within these histograms.

