



**T.C.
İSTANBUL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**



YÜKSEK LİSANS TEZİ

**OBJE TABANLI SINIFLANDIRMA YÖNTEMLERİNİN
KARŞILAŞTIRILMASI**

Hüseyin DOĞAN

Orman Mühendisliği Anabilim Dalı

Orman Mühendisliği Programı

DANIŞMAN

Yrd. Doç. Dr. Muhittin İNAN

Aralık, 2017

İSTANBUL

Bu çalışma 7.12.2017 tarihinde aşağıdaki jüri tarafından Orman Mühendisliği Anabilim Dalı, Orman Mühendisliği Programında Yüksek Lisans tezi olarak kabul edilmiştir.

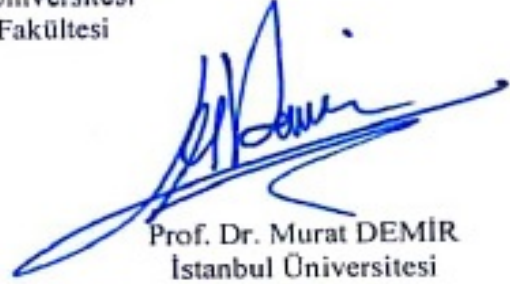
Tez Jürisi



Yrd. Doç. Dr. Muhittin İNAN (Danışman)
İstanbul Üniversitesi
Orman Fakültesi



Prof. Dr. Hakan YENER
İstanbul Üniversitesi
Orman Fakültesi



Prof. Dr. Murat DEMİR
İstanbul Üniversitesi
Orman Fakültesi



Doç. Dr. Esra ERTEN
İstanbul Teknik Üniversitesi
İnşaat Fakültesi



Yrd. Doç. Dr. Emrah ÖZDEMİR
İstanbul Üniversitesi
Orman Fakültesi



20.04.2016 tarihli resmi gazetede yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince; Bu Lisansüstü teze, İstanbul Üniversitesi'nin abonesi olduğu intihal yazılım programı kullanılarak Fen Bilimleri Enstitüsü'nün belirlemiş olduğu ölçütlere uygun rapor alınmıştır.

ÖNSÖZ

Yüksek Lisans eğitimim süresince ve tez konumun belirlenmesinden başlayarak sonuçlanana kadar bilgi, birikim ve zamanını benimle paylaşan tez danışmanım ve değerli hocam Yrd. Doç. Dr. Muhittin İNAN'a,

Tezimin geliştirilmesi ve açıklarının belirlenmesi için sorduğum tüm soruları sabırla cevaplayan ve görüşlerini bildiren değerli hocalarım Yrd. Doç. Dr. Hüseyin YURTSEVEN, Doç. Dr. O. Yalçın YILMAZ, Prof. Dr. Hakan YENER ve Prof. Dr. Ayhan KOÇ'a ve çalışma arkadaşım Arş. Gör. Burcu KURTOĞLU'na,

Çalışma konumun belirlenmesinde yardımları bulunan Universitat für Bodenkultur Wien'deki Erasmus+ tez danışmanım Prof. Dr. Clement ATZBERGER'a,

Eğitimim süresince sabırla yanımda olan ve desteklerini hiçbir zaman esirgemeyen annem, babam, ablalarım ve arkadaşlarıma en içten teşekkürlerimi sunarım.

Aralık 2017

Hüseyin DOĞAN

İÇİNDEKİLER

Sayfa No

ÖNSÖZ.....	iv
İÇİNDEKİLER	v
ŞEKİL LİSTESİ.....	vii
TABLO LİSTESİ	x
SİMGE VE KISALTMA LİSTESİ	xi
ÖZET.....	xiii
SUMMARY	xv
1. GİRİŞ.....	1
2. GENEL KISIMLAR	6
2.1. GÖRÜNTÜ SINIFLANDIRMA TEKNİKLERİ	6
2.1.1. Piksel Tabanlı Sınıflandırma Yöntemi.....	7
2.1.2. Obje Tabanlı Sınıflandırma Yöntemi	10
2.1.2.1. Parametrik ve Parametrik Olmayan Obje Tabanlı Sınıflandırma Algoritmaları	10
2.1.2.2. Karar Ağaçları ve Random Forest	12
2.1.2.3. Makine Öğrenimi ve Nearest Neighbor Sınıflandırma Yöntemi	17
3. MALZEME VE YÖNTEM	21
3.1. MALZEME	21
3.1.1. Çalışma Alanı ve Özellikleri.....	21
3.1.1.1. Coğrafi Konum.....	22
3.1.1.2. İklim ve Topografik Yapı.....	23
3.1.1.3. Bitki Örtüsü	24
3.1.2. Çalışmada Kullanılan Veriler ve Yazılımlar.....	24
3.1.2.1. Uydu Görüntüleri	25
3.1.3. Amenajman Planları.....	36
3.2. YÖNTEM	36
3.2.1. Görüntü Ön İşleme.....	38
3.2.2. Segmentasyon	43
3.2.3. Obje Tabanlı Sınıflandırma.....	45
3.2.3.1. Random Forest Sınıflandırma Yöntemi	46

3.2.3.2. <i>Nearest Neighbor Sınıflandırma Yöntemi</i>	49
3.2.3.3. <i>Threshold Sınıflandırma Yöntemi</i>	49
3.2.4. Doğruluk Analizi ve Sınıflandırma Doğrulukları	51
4. BULGULAR	55
4.1. NEAREST NEIGHBOR SINIFLANDIRMA SONUÇLARI	58
4.2. RANDOM FOREST SINIFLANDIRMA SONUÇLARI	58
4.3. THRESHOLD SINIFLANDIRMA SONUÇLARI	60
5. TARTIŞMA VE SONUÇ	65
KAYNAKLAR	70
EKLER	78
EK-1 Çalışmada Kullanılan R-Studio Makroları.....	78
ÖZGEÇMİŞ	79

ŞEKİL LİSTESİ

Sayfa No

Şekil 1.1: Farklı parametrelerle segmentasyon işlemi gerçekleştirilmiş uydu görüntüleri; a) istenilen kriterlere sahip segmentasyon b) orijinal görüntü c) aşırı segmente edilmiş görüntü.	4
Şekil 2.1: Kontrollü Sınıflandırma yöntemi.	8
Şekil 2.2: Tuz ve Biber etkisine örnek. (a) yüksek çözünürlüklü bir uydu görüntüsünün piksel tabanlı sınıflandırma yöntemi ile sınıflandırılmış hali. (b) aynı görüntünün obje tabanlı sınıflandırma yöntemi ile sınıflandırılmış hali (Maggie ve diğ. 2011).	9
Şekil 2.3: Tablo 2.1 ‘deki eğitim sınıflarından oluşturulan veri kümeleri.....	13
Şekil 2.4: Şekil 2.3 ‘deki veri kümelerinden oluşturulmuş bir sınıflandırma kuralı (Karar Ağacı). (Çölkesen, 2009) (Değiştirilerek kullanılmıştır).	14
Şekil 2.5: T sayıdaki örnek alanın bir kısmını <i>In-Bag</i> , geri kalan kısmını ise <i>Out-Of-Bag</i> (OOB) olarak test aşamasında kullanılan ve sınıflandırıcılar arasından en zayıf olanların ağırlığı doğrultusunda karar sınıfı oluşturan bir “ensemble learning” algoritması olan <i>Random Forest</i> (Vuolo ve Ng, 2016).	15
Şekil 2.6: Nearest Neighbor Sınıflandırıcı Prensibi.	17
Şekil 2.7: <i>NN</i> ve <i>k-NN</i> sınıflandırıcılar arasındaki fark. İki boyutlu boşluk uzayında verilerin dağılımı(1). 3 farklı sınıf kullanılarak(yeşil-kırmızı-mavi) oluşturulan karar sınırları arasındaki farkı görebiliriz. (2) <i>NN</i> sınıflandırıcısında karar sınırları içerisinde bulunan kırmızı ve yeşil adacıklar yanlış sınıflandırılma sonucudur, bu durumda beklenen kırmızı ve yeşil adacıkların mavi sınıfa dahil edilmesidir. (3) <i>k-NN</i> sınıflandırıcıda <i>genelleştirmenin</i> yani karar sınırları içerisinde bulunan noktaların aynı sınıflara atanma işleminin daha düzgün yapıldığı görülebilir. (http://cs231n.github.io/classification kaynağından değiştirilerek alınmıştır, 2017)	19
Şekil 2.8: Derrac, Garcia ve Herrera tarafından yazılan 2013 yılı çıkışlı “Bulanık en yakın komşu algoritmaları: taksonomi ve deneysel analizler” araştırmasında önerilen <i>Bulanık En Yakın Komşu</i> sınıflandırıcıları taksonomisi.	20
Şekil 3.1: İstanbul İli ve Çalışma Alanı sınırları.	22
Şekil 3.2: Sentinel-2a uydusu için bantların çözünürlüğü ve elektromanyetik spektrumdaki yerleri. (ESA, 2016). (Değiştirilerek kullanılmıştır).....	27
Şekil 3.3: 24 Temmuz 2016 tarihli 13 bantlı Sentinel-2a görüntüsü.	28

Şekil 3.4: Sentinel 2 uydusuyla zirai alanların gözlemlenmesi için belirlenen bant kombinasyonu (Kırmızı Kanal: 11, Yeşil Kanal: 8a ve Mavi Kanal: 2).....	30
Şekil 3.5: Sentinel 2 uydusuyla atmosferik geçirgenlik sağlaması için kullanılan bant kombinasyonu. (Kırmızı Kanal: 12, Yeşil Kanal: 11, Mavi Kanal: 4)	30
Şekil 3.6: Sentinel 2 uydusuyla sağlıklı vejetasyonun belirlenmesi için kullanılan bant kombinasyonu. (Kırmızı Kanal: 8a, Yeşil Kanal: 11, Mavi Kanal: 2)	32
Şekil 3.7: Sentinel 2 uydusuyla vejetasyon analizleri için kullanılan bant kombinasyonu.(Kırmızı Kanal: 11, Yeşil Kanal: 8a, Mavi Kanal: 4)	32
Şekil 3.8: Sentinel 2 uydusu ile elde edilen standart Kızılötesi bant kombinasyonu (Kırmızı Kanal: 8, Yeşil Kanal: 4, Mavi Kanal: 3).....	33
Şekil 3.9: Sentinel 2 uydusunun 4 ve 5. bantları ile oluşturulan NDI45 vejetasyon indeksi ve spektral değerleri. (Bant Formülü: $(NIR - R) / (NIR + R)$)	34
Şekil 3.10: Sentinel 2 uydusunun 4 ve 7. Bantları ile oluşturulan NDVI vejetasyon indeksi ve spektral değerleri. (Bant Formülü: $(NIR - R) / (NIR + R)$).....	35
Şekil 3.11: Sentinel 2 uydusunun 4, 5, 6 ve 7. Bantları ile oluşturulan IRECI vejetasyon indeksi ve spektral değerleri. (Bant Formülü: $(NIR - R) / (RE1 / RE2)$)	37
Şekil 3.12: Sentinel 2 uydusunun 7. bandının 4. banda oranlanmasıyla elde edilen PSSR vejetasyon indeksi ve spektral değerleri. (Bant Formülü: (NIR / R))....	37
Şekil 3.13: Sınıflandırma doğruluklarının belirlenmesi için kullanılacak yer gerçeklerinin elde edileceği işletme şeflikleri sınırları ve bölmecikler.	38
Şekil 3.14: Random Forest sınıflandırma yöntemi için iş akış diyagramı.....	39
Şekil 3.15: Nearest Neighbor sınıflandırma yöntemi için iş akış diyagramı.....	39
Şekil 3.16: Threshold sınıflandırma yöntemi için iş akış diyagramı.....	40
Şekil 3.17: Sen2Cor yöntemi ile atmosferik düzeltmenin gerçekleştirilmesi. İşlenmemiş Sentine-2a Level-1C görüntüsü ve bant değerleri (üstte). Atmosferik düzeltmesi gerçekleştirilmiş Level-2A nitelikli uydu görüntüsü ve bant değerleri (altta).	42
Şekil 3.18: Multiresolution Segmentation algoritması ile gerçekleştirilen segmentasyon işlemi ve kullanılan parametreler (Ölçek Parametresi: 20, Pürüzsüzlük: 0.5, Yoğunluk: 0.5).	44
Şekil 3.19: Large Scale Meanshift Segmentation algoritması ile gerçekleştirilen segmentasyon işlemi ve kullanılan parametreler (Alansal Mesafe: 10, Spektral Mesafe: 10, Min. Alan Boyutu: 10).	45
Şekil 3.20: QGIS içerisinde eğitim alanlarının seçilmesi.....	46

Şekil 3.21: eCognition ile eğitim alanlarının seçilmesi.....	50
Şekil 3.22: Belirlenen bir eşik değeri ile NDVI üzerinde vejetasyonun görüntülenmesi (Mavi alanlar 0.5 üzerindeki yansıma değerlerini belirtmektedir).....	51
Şekil 3.23: Doğruluk analizi için kullanılan uydu görüntüsü üzerine bindirilmiş meşçere ve arazi kullanım sınıfları.	53
Şekil 3.24: ERDAS Imagine ile rastgele nokta atılması (Sarı noktalar doğrulukları belirlenmemiş, kırmızılar ise belirlenmiş olanlardır).....	54
Şekil 4.1: Hiyerarşik sınıflandırma ile yerleşim ve su yüzeylerinden ayrılan vejetasyon alanları.....	56
Şekil 4.2: Hiyerarşik sınıflandırma ile vejetasyon ve su yüzeylerinden ayrılan yerleşim alanları, toprak ve yapay yüzeyler.....	56
Şekil 4.3: Hiyerarşik sınıflandırma ile vejetasyon ve yerleşim alanlarından ayrılmış su yüzeyleri.	57
Şekil 4.4: Vejetasyon sınıfına ait alanların daha sonra kendi içerisinde objelere ayrılması.....	58
Şekil 4.5: Nearest Neighbor algoritması ile sınıflandırılmış Sentinel-2a görüntüsü....	58
Şekil 4.6: RF algoritması ile sınıflandırılmış Sentinel-2a uydu görüntüsü.	60
Şekil 4.7: Threshold Sınıflandırma ile görüntünün 3 ana sınıfa ayrılması.	62

TABLO LİSTESİ

Sayfa No

Tablo 2.1: Sadeleştirilmiş bir eğitim sınıfları tablosu.....	13
Tablo 3.1: R Studio “matrixStats” paketi ile bantların spektral değerlerinin görüntülenmesi.....	27
Tablo 3.2: RF için sınıf hiyerarşisi ve örnek alan sayıları	47
Tablo 3.3: NN için sınıf hiyerarşisi ve örnek alan sayıları	48
Tablo 4.1: Hiyerarşik sınıflandırma ile 3 ana sınıfa ayrılan görüntünün hata matrisi.....	52
Tablo 4.2: NN Sınıflandırma doğrulukları.....	59
Tablo 4.3: Random Forest Sınıflandırma Doğrulukları	61
Tablo 4.4: THC Sınıflandırma doğrulukları.....	63
Tablo 4.5: Threshold, Random Forest ve Nearest Neighbor Sınıflandırıcılarının 3 ana sınıfta karşılaştırılması.....	65
Tablo 4.6: Random Forest ve Nearest Neighbor Sınıflandırıcılarından 9 sınıfta elde edilen sonuçlar ve sınıfların alanlara göre dağılımı.....	66

SİMGE VE KISALTMA LİSTESİ

Kısaltmalar	Açıklama
RF	Random Forest
NN	Nearest Neighbor (En Yakın Komşu)
SVM	Support Vector Machines (Destek Vektör Makineleri)
ML	Maximum Likelihood (En Yüksek Benzerlik)
THC	Threshold Classification (Eşik Değer Sınıflandırıcısı)
DN	Digital Number (Sayısal Değer)
LULC	Land Use Land Cover (Arazi Kullanımı ve Örtüsü)
OBIA	Object Based Image Analysis (Objekt Tabanlı Görüntü Analizi)
UAV-İHA	Unmanned Aerial Vehicle (İnsansız Hava Aracı)
DEM	Digital Elevation Model (Sayısal Yükseklik Modeli)
ESA	European Space Agency (Avrupa Uzay Ajansı)
CART	Classification and Regression Trees
ELM	Ensemble Learning Methods (Birleşik Öğrenme Metodları)
OOB/IB	Out of Bag / In Bag
VI	Vegetation Indices (Vejetasyon İndeksleri)
LAI	Leaf Area Index (Yaprak Yüzey Alanı)
LCC	Leaf Chlorophyll Concentration (Yaprak Klorofil Yoğunluğu)
S2	Sentinel 2
NDVI	Normalized Difference Vegetation Index
IRECI	Inverted Red Edge Chlorophyll Index
SR	Simple Ratio
PSSR	Pigment Specific Simple Ratio
AOT	Aerosol Optical Thickness
TOA	Top-of-Atmosphere
BOA	Bottom-of-Atmosphere
MT	Multi-temporal (Çok zamanlı)
MS	Multi-spectral (Çok Bantlı)
NIR	Near Infrared (Yakın Kızılötesi)
SWIR	Short Wave Infrared (Kısa Dalga Kızılötesi)

ÖZET

YÜKSEK LİSANS TEZİ

OBJE TABANLI SINIFLANDIRMA YÖNTEMLERİNİN KARŞILAŞTIRILMASI

Hüseyin DOĞAN

İstanbul Üniversitesi

Fen Bilimleri Enstitüsü

Orman Mühendisliği Anabilim Dalı

Danışman : Yrd. Doç. Dr. Muhittin İNAN

Son yıllarda dijital görüntü işleme tekniklerinin önemli gelişmeler göstermesi ve görüntülere olan erişim imkânlarının kolaylaşmasının sonucu olarak, arazi örtü ve kullanım haritalarının uydu görüntüleri veya hava fotoğraflarından elde edilmesi amacıyla farklı metotlar geliştirilmiştir. Bunlardan öne çıkanlar Piksel ve Obje tabanlı sınıflandırma yöntemleri olarak kabul edilmektedir. Bu çalışma ile amaçlanan son yıllarda oldukça sık kullanılan görüntü sınıflandırma yöntemleri olan obje tabanlı sınıflandırma yöntemlerinden 3 tanesinin karşılaştırılması ve kullanılan veriler doğrultusunda performanslarının değerlendirilmesidir. Ayrıca uzaktan algılama verilerinin özellikleri, elde edilme yöntemleri ve sınıflandırma aşamasında kullanılmadan önce yapılması gereken ön işlemlerin neler olduğuna yer verilmiştir.

Çalışma kapsamında görüntü zenginleştirme tekniklerinden bazılarına; ayrıca vejetasyon indeksleri ve görüntü yorumlama için kullanılan bant kombinasyonlarına da yer verilmiştir. Bu sayede özellikle Threshold Sınıflandırıcı için gereken Vejetasyon İndeksleri hakkında bilgilendirmelerde bulunulmuş, ayrıca 13 bantlı Sentinel 2a uydusu ile elde edilebilecek vejetasyon indeksleri üzerinde durulmuştur. Bu çalışmanın bir diğer özelliği ise orta çözünürlükteki uydulara göre oldukça yeni sayılabilecek bir uydu sistemi olan Sentinel 2a verileriyle çalışılmış olmasıdır; bu verilerin özelliği uydu tarafından görüntülediği tarihten itibaren erişime açık ve ücretsiz olmasıdır. Böylece, obje sınıflandırma yöntemleri gibi aynı zamanda Sentinel 2a verilerinin performansının değerlendirilmesi ile gelecekte bu uydu verileri ile yapılacak çalışmalar için faydalı bir kaynak görevi görmesi amaçlanmıştır.

Obje tabanlı sınıflandırıcılar arasında bu çalışma içerisinde kullanılacak olan Random Forest ve Nearest Neighbor algoritmalarının çalışma prensipleri ve günümüze kadar nasıl bir gelişim geçirdikleri üzerinde durulmuştur. Bu yöntemleri anlayabilmek için gerekli olan terminoloji de detaylı bir şekilde yöntem başlıkları altında verilmiştir. Bu sınıflandırma yöntemlerinin tanımlanması, kullanım alanlarının ortaya konulması ve gerçekleştirilmesi için kullanılan yazılımların özellikleri de bu tez kapsamında üzerinde durulan bir diğer konudur. Birçoğu açık kaynak ya da ücretsiz yazılımlar olan QGIS, R Studio ve SNAP gibi yazılımların verilerin işlenmesini nasıl gerçekleştirdiği ve kullanılan yöntemlerin neler olduğu ilgili başlıklar altında yer almaktadır.

Sonuç olarak verilerin elde edilmesi, işlenmesi ve farklı algoritmalar ile sınıflandırma işlemlerinin gerçekleştirilmesi ile sınıflandırılmış görüntüler oluşturulmuş ve elde edilen doğruluk değerleri üzerinden bu sınıflandırma yöntemlerinin Sentinel 2 gibi orta çözünürlüklü uydu sistemlerinde kullanımlarının ne kadar verimli kullanılabildiği belirlenmiştir. 9 sınıf üzerinde yapılan sınıflandırma sonuçlarında Nearest Neighbor sınıflandırıcı %77.41 toplam doğruluğa sahipken Random Forest sınıflandırıcıda bu oran %87.77 olmuştur. Threshold sınıflandırıcıyı 9 farklı sınıfta kullanmak mümkün olmadığından 3 ana sınıf üzerinde (vejetasyon, su yüzeyleri, vejetasyon dışı alanlar) yapılan sınıflandırma ile sınıflandırma doğruluğu %94.44 olarak bulunmuştur. Random Forest ve Nearest Neighbor sınıflandırıcılarının da Threshold Sınıflandırma ile karşılaştırılmak için 3 ana sınıf üzerinde kullanılması ile ulaşılan doğruluk değerleri sırasıyla %98.76 ve %96.24 olmuştur.

Aralık 2017, 94 sayfa.

Anahtar kelimeler: Görüntü Sınıflandırma Yöntemleri, Sentinel 2, Vejetasyon İndeksleri, Karar Ağaçları, En Yakın Komşu Sınıflandırma

SUMMARY

M.Sc. THESIS

COMPARISON OF OBJECT BASED CLASSIFICATION METHODS

Hüseyin DOĞAN

Istanbul University

Institute of Graduate Studies in Science and Engineering

Department of Forest Engineering

Supervisor : Asst. Prof. Dr. Muhittin İNAN

In recent years, different methods have been developed to obtain land cover and usage maps from satellite images or aerial photographs as a result of significant improvements in digital image processing techniques and facilitation of access to images. The prominent ones are considered as pixel and object based classification methods. The aim of this study is to compare three of object-based classification methods which are frequently used image classification methods in recent years and to evaluate their performance in the direction of the input data. Additionally, the features of remote sensing data, how to obtain them, and the precautions to be taken before using them in the classification phase are given in this research.

Within the scope of the study, some of the image enhancement techniques; as well as vegetation indexes and band combinations used for image interpretation are mentioned. In this issue, we have been informed especially about the vegetation indexes required for the Threshold Classifier and also the vegetation indices which can be obtained with the Sentinel 2a conformity with 13 band. Another feature of this study is that it is based on Sentinel 2a, a satellite system that is quite new to the medium/high-resolution satellites; the feature of this data is that access is open and free from the date the satellite is displayed. Thus, as with object classification methods, it is also intended to serve as a useful resource for evaluating the performance of Sentinel 2a data and future work with these satellite data.

As a result, classification images were obtained by processing, processing and classification with different algorithms, and the accuracy of the obtained images was used to determine how efficiently these classification methods could be used in medium/high resolution satellite systems such as Sentinel 2. Nearest Neighbor classifier has a total accuracy of 77.41% in classification results made on 9 classes, whereas it is 87.77% in Random Forest classifier. Since the Threshold classifier can not be used on 9 classes, the classification accuracy was found to be 94.44% on the 3 main classes (vegetation, water surfaces, non-vegetation areas). The accuracy of Random Forest and Nearest Neighbor classifiers on the 3 main classes to compare with Threshold Classification was 98.76% and 96.24% respectively.

December 2017, 94 pages.

Keywords: Image Classification Methods, Sentinel 2, Vegetation Indices, Decision Trees, Nearest Neighbor Classification

1. GİRİŞ

Görüntü sınıflandırma teknikleri uzun süredir uzaktan algılama uygulayıcıları için vazgeçilmez birer araç olmuştur. Özellikle ormancılıkta, ormanların haritalanması ve gözlemlenmesi için büyük kolaylıklar sağlayan bu araçların sağladığı avantajlar artarak devam etmektedir. Bu tekniklerin klasik arazi ölçüm uygulamalarına oranla uygun maliyetli olduğu kabul edilmekte, gelişen teknoloji ile birlikte doğrulukların artması da bu tekniklerinin kullanımına olan yönelimin artmasını sağlamaktadır.

Orman niteliklerinin anlaşılmasında kullanılan sınıflandırma yöntemleri önceleri stratejik amaçlı (orman alanı tespiti, kadastral sınırlar vb.) uygulamalarda kullanılıyorken, son yıllarda bu sınıflandırma yöntemlerinin kullanım alanları daha detaylı (operasyonel) hale gelmiştir: Bu sınıflandırma yöntemleri meşcere tür tespitinden, bireysel ağaç hacim hesaplamalarına kadar kullanılabilir (Desheng ve Fang, 2010; Lefebre ve diğ. 2016). Bu noktada yapılacak sınıflandırmaların amacının belirlenmesi ve kullanılacak materyallerin bu doğrultuda seçilmesi büyük önem taşımaktadır. Çalışmanın içerisinde bu sınıflandırma yöntemlerinin en uygun kullanım alanları, verilerin seçilmesi ve işlenmesi gibi noktalara değinilmiştir.

McRoberts ve Tomppo (2007) ‘nın Avrupa’nın ulusal orman varlıkları üzerine yapmış olduğu çalışma, sınıflandırılmış görüntülerin bizlere hız, kesinlik ve verimlilik kazandırmanın yanında yukarıda bahsedilen orman envanterinin belirlenmesinde ve haritalanmasında daha önceden mümkün olmayacak doğruluklarda sonuçlar elde etme olanağı sağladığını göstermiştir.

Görüntü sınıflandırma teknikleri kapsamında ele alınan iki temel yöntem vardır: piksel ve obje tabanlı sınıflandırma. Arazi kullanımı ve arazi örtüsü belirlemek için yapılan sınıflandırma işlemlerinde kullanılan ilk yöntemler piksel tabanlı sınıflandırma metotlarıydı. Bu yöntem kendi içinde ikiye ayrılmakta (kontrollü ve kontrolsüz); ayrıca bu yöntemlerin kombinasyonlarıyla da doğruluk değerleri arttırılabilmektedir (Enderle ve Weih, 2005).

Araştırmacılar tamamen otomatik gerçekleşen bir sınıflandırma tekniği oluşturmak için yaptığı çalışmalar sonucu piksel tabanlı sınıflandırmayı temel alarak obje tabanlı sınıflandırma yöntemini teoriye dökmüşlerdir (Blaschke ve diğ. 2000; Csatho ve diğ. 1999). Bu sınıflandırma yöntemindeki temel mantık; benzer spektral değerlere sahip komşu pikselleri bir araya toplayıp, meydana gelen piksel kümelerine tek bir objeymiş gibi davranarak farklı algoritmalarla sınıflandırmaya dahil edilmesini sağlamaktır (Dronova ve diğ. 2011). Bu çalışmada obje tabanlı sınıflandırmadaki obje kavramı ve farklı algoritmalarda bu objelerin davranış şekilleri incelenmiş, ayrıca obje tabanlı sınıflandırmada kullanılan yazılımlar hakkında bilgilendirmelere yer verilmiştir.

Bu çalışma ile amaçlanan, bahsedilen bu sınıflandırma yöntemlerinden obje tabanlı yaklaşımların aynı görüntü üzerindeki doğruluk değerlerinin karşılaştırılmasıdır. Bu yöntemlerin kullanımının ormancılık açısından sağlayacağı önemli kazanımlar McRobberts ve Tomppo'nun 2007 yılında yaptıkları çalışmada şu şekilde aktarılmıştır: Ormancılıkta arazi örtü haritasının uygun ölçeklerle ve düzenli aralıklarla oluşturulması esastır, obje tabanlı yaklaşımlar ile bu haritaların oluşturulması hem maliyeti hem de doğrulukları bakımından oldukça tatmin edicidir.

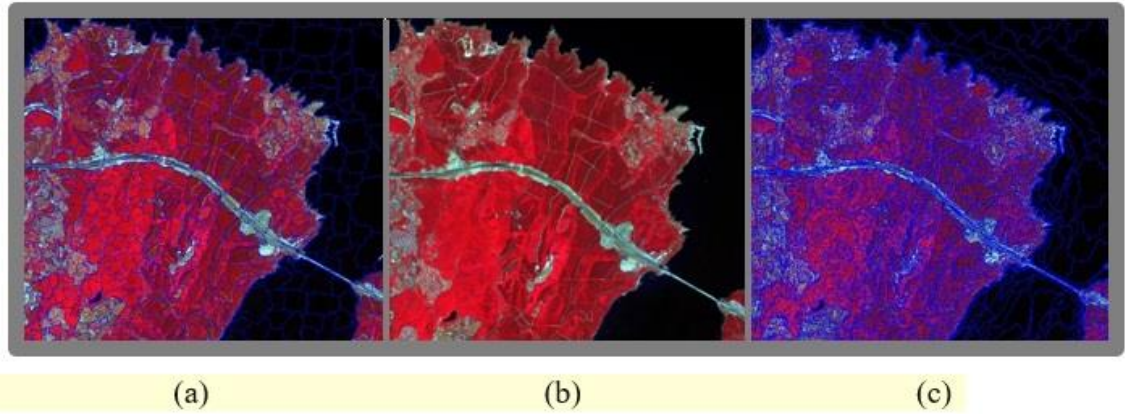
Ancak, obje tabanlı yaklaşımlar her ne kadar son teknoloji sınıflandırma yöntemlerinden biri olarak kabul edilse de zayıf kaldığı noktalar da vardır. Desheng ve Fang'ın 2010 tarihli araştırmalarında bu yöntemlerin piksel tabanlı yöntemlerden ileride olduğu ancak hala geliştirilmesi gereken yönlerinin bulunduğu üzerinde durulmuştur.

Bu yöntemlerin öne çıkan dezavantajları, segmentasyon işlemi için uygun değerlerin belirlenmesi ve kullanılan verilerin çözünürlüğünün yüksek olması gerekliliğidir. Çözünürlüğü yüksek olan uydu görüntüleri (SPOT, IKONOS, Quickbird, WorldWiev-2 ile yapılan araştırmalarda bu yöntemin piksel tabanlı yöntemlere kıyasla daha üstün sonuçlar verdiği görülmektedir (Cleave ve diğ. 2008; Kalkan ve Maktav, 2010; Colkesen ve Kavzoglu, 2015). Ancak bu verilere ulaşımın yüksek maliyetli olması ve işlenmesi için gereken sürelerin orta çözünürlükteki görüntülerden çok daha yüksek olması;

Sentinel-2a gibi arşiv verilerinin son kullanıcılar için ücretsiz olduğu ve işlenme süresi bakımından çok hızlı olan uydu görüntülerinin, obje tabanlı yaklaşımlar için uygunluğunun değerlendirilmesinde önemli bir konu olduğunu göstermiştir.

Bu çalışmada, obje tabanlı sınıflandırma yöntemlerinin verimlilikleri orta çözünürlüklü uydu verileri üzerinde karşılaştırılıp, yöntemlerin kabiliyetleri ve eksik yönlerine değinilmiştir. Elde edilen sonuçların ne ifade ettiği ise tartışma ve sonuç bölümünde ele alınmıştır.

Çalışma kapsamında karşılaşılan problemler, temelde kullanılan uydu verilerinin mekansal çözünürlükleriyle alakalıdır. Obje tabanlı sınıflandırma yöntemlerinin doğruluğu değerleri sınıflandırma öncesi yapılan segmentasyon işlemiyle alakalıdır (Pal ve Pal, 1993) ve segmentasyon işleminde en önemli kriter kullanılan verilerin mekansal çözünürlüğüdür. “Malzeme ve Yöntem” bölümünde üzerinde durulacak bir konu olan segmentasyon işlemi temel olarak birbirine benzeyen homojen desenlerin (piksellerin) bir araya toplanması olarak ifade edilebilir. Bu işlem için kullanılan değerler çoğu zaman kullanıcının ilgi alanı doğrultusunda, kısmen deneme yanılma yöntemiyle (eyeball estimation) bulunur, bu nedenle kullanıcı hatasına açık ve sınıflandırma doğruluklarına doğrudan etki eden bir aşamadır (Martin ve diğ, 2001; N. Clinton ve diğ. 2010). Segmentasyon ile ilgili yapılan araştırmalar, bu yöntem her ne kadar tatmin edici sonuçlar verse de geliştirilmesi gerektiğine işaret etmektedir (Rajvardhan, 2016).



Şekil 1.1: Farklı parametrelerle segmentasyon işlemi gerçekleştirilmiş uydu görüntüleri:

- a) istenilen kriterlere sahip segmentasyon b) orijinal görüntü c) aşırı segmente edilmiş görüntü.

Şekil 1.1'de obje boyutları arasındaki farkı görebiliriz; Şekil 1.1(b)'de orijinal görüntüyü, Şekil 1.1(a)'da homojen desenleri bir araya almış, alan gerçekliğini tam olarak yansıtabilecek yapılar oluşturan segmentlerin, Şekil 1.1(c)'de boyutlarının çok küçüldüğü, insan gözü açısından homojen sayılamayacak bir yapıda olduğu görülmektedir.

Ayrıca, çalışma alanı içerisinde birbirine çok yakın yansıma değerleri olan farklı arazi örtü sınıfları bulunmaktadır. Sınıflandırma aşamasında bahsedilen bu alanlardaki sınıflandırma doğruluklarının düşük olması, genel doğruluğu ciddi oranlarda düşürmüştür (örneğin: kumullar, kazı alanları ve maden sahaları birbiri içerisinde karışıyor iken; beyaz çatılı sanayi bölgesi ve seralar aralarında karışmaktadır). Yansıma değerlerinin benzerlik göstermesinin temel nedeni kullanılan uydu verisinin (Sentinel-2a) yeterince yüksek spektral ve mekansal çözünürlüğe sahip olmamasından ileri gelir.

Daha önce bahsedilen araştırmaların da sonucunda belirtildiği gibi, uygun maliyetli ve yüksek doğruluk içeren arazi kullanım haritaları üretimi ormancılık ve ormancılığın alt dalları için oldukça önemli bir konuyu oluşturmaktadır. Ülkemiz ormancılığında kullanılan altlıkların birçoğunun hala yoğun yersel ölçümler gerektirdiğini göz önünde bulundurursak (Yurtseven, 2008), son teknoloji ürünü orta/yüksek çözünürlüklü uydulardan: amenajman planları (özellikle ulaşımı güç bölgeler için), böcek/yangın zararı, orman tahribatı haritaları ve bunlar gibi birçok ormancılık operasyonu için yardımcı altlıkların oluşturulması önemli bir girişim olacaktır.

Bu çalışmanın çıkış noktalarından bir tanesi de, kullanılacak olan Sentinel-2a uydu görüntülerinin son kullanıcı tarafından ücretsiz olarak temin edilebiliyor olmasıdır. Böylece, sadece uygun yazılım ve görüntü sınıflandırma uzmanları yardımıyla bahsedilen bu tematik haritalar çok uygun maliyetlerle elde edilebilecek ve arazi çalışması için ayrılan harcamaların miktarı azalacaktır. 2015 yılında ESA tarafından fırlatılan Sentinel-2a uydusu üzerine yapılan araştırmalar gün geçtikçe artmakta (Atzberger ve diğ., 2012; Richter ve diğ., 2011; Immitzer ve diğ., 2016) ve uzaktan algılama uzmanları tarafından Landsat uydularından sağlanan faydanın benzerinin Sentinel uydu sistemlerinden beklenmektedir.

Uydu görüntülerinin ücretsiz temin edilmesi, ayrıca geometrik ve atmosferik düzeltmelerin SNAP isimli Sentinel-2a görüntü işleyici platformunun içerisinde kolaylıkla gerçekleştirilmesi bu uydunun kullanımının gelecek yıllarda artacağını habercisidir. Bu çalışma içerisinde Sentinel-2a uydu görüntülerinin obje tabanlı sınıflandırma yöntemleri için verimliliğinin incelenmesinin yanında sınıflandırılan görüntülerin potansiyel kullanım alanlarının belirlenmesi de amaçlanmıştır. Ayrıca bu çalışmanın, Sentinel uydu sistemleri ile yapılacak çalışmalarda yararlanılacak bir kaynak niteliği taşıması hedeflenmiştir.



2. GENEL KISIMLAR

Obje tabanlı sınıflandırmanın anlaşılabilmesi için öncelikle; görüntü sınıflandırma teknikleri, bilgi sınıfları, spektral sınıflar, arazi kullanım haritaları, segmentasyon, yansıma değerleri, karar ağaçları, makine öğrenimi gibi kavramların biliniyor olması önemlidir. Bu bölümde obje tabanlı sınıflandırmanın temellerini oluşturan bileşenler ve bunların bağlı olduğu kavramlara yer verilecektir.

Bunların yanında, sınıflandırma işlemine doğrudan etki eden bir diğer faktör, uzaktan algılama verilerinin seçimidir. Bu aşamada çalışma alanına ait verilerin uydu görüntülerinden mi yoksa hava fotoğraflarından mı temin edileceği, yapılacak araştırmaya göre verilerin toplandığı tarihin belirlenmesi vb. gibi detaylara dikkat edilmesi gerekir. Kullanılacak uzaktan algılama verilerinin belirlenmesine yönelik birçok araştırma mevcuttur. Bunlara Barnsley(1999) ve Lefsky ve Cohen(2003) 'in farklı uzaktan algılama verilerinin spektral, radyometrik, mekânsal ve zamansal çözünürlüklerinin; ayrıca polarizasyon ve açısallık gibi özelliklerinin karşılaştırıldığı araştırmaları örnek gösterilebilir.

2.1. GÖRÜNTÜ SINIFLANDIRMA TEKNİKLERİ

Görüntü sınıflandırma teknikleri birçok açıdan karmaşık bir işlemler bütünüdür. Uzaktan Algılama araştırmalarının bu alana odaklanmış çalışmaları, sınıflandırma sonuçlarının çevresel ve sosyoekonomik çalışmaların sonucunu doğrudan etkilemesi nedeniyle, yıllardır uzaktan algılama topluluklarının ilgisini çekmiştir (D. Lu ve Q. Weng 2006). Bilim adamları ve uygulamacılar bu alanda büyük emekler harcamış ve sınıflandırma doğruluklarını arttırmak amacıyla birçok teknik ve yöntem geliştirmişlerdir (Franklin ve diğ. 2002; Pal ve Mather 2003; Gallego 2004). Buna rağmen, uzaktan algılanmış bir görüntünün tematik bir haritaya dönüştürülmesi günümüzde bile kolay bir işlem değildir, buna arazi yapısının karmaşıklığı, uzaktan algılama verileri, görüntü işleme ve sınıflandırma yöntemlerinin seçilmesi vb. gibi nedenler gösterilebilir (Tso ve Mather 2001; Landgrebe 2003).

Bu araştırmada en sık kullanılan iki sınıflandırma yönteminden biri olan obje tabanlı sınıflandırma yöntemi üzerinde durulacaktır. Nitekim bu yöntemin daha iyi anlaşılması

için bu yöntemle temel olan piksel tabanlı sınıflandırmanın da bilinmesi önemlidir. Bu sınıflandırma metodu genel hatları ile incelenip daha sonra obje tabanlı sınıflandırma ile elde edilen gelişmelerin neler olduğu ilerleyen bölümlerde verilmiştir.

2.1.1. Piksel Tabanlı Sınıflandırma Yöntemi

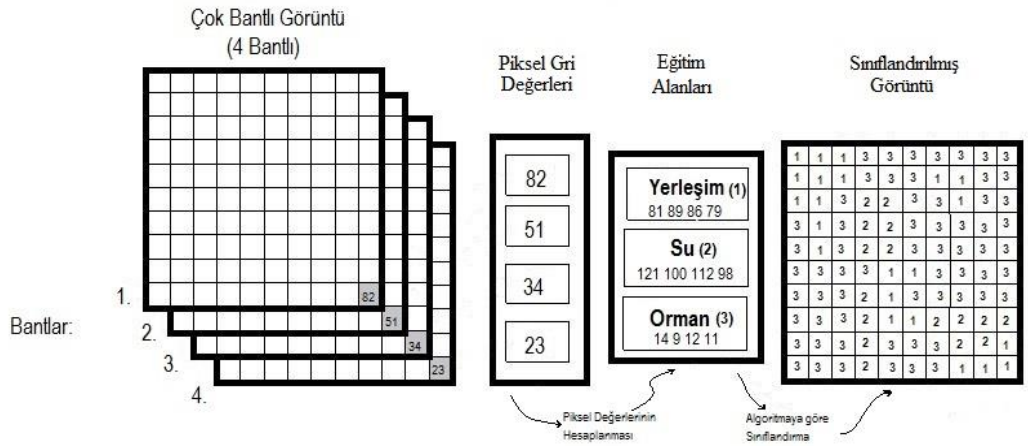
Uzaktan algılanmış görüntüler sınıflandırma işleminde kullanılacak bilgiler içerir, bunlar; spektral, mekânsal, radyometrik ve zamansal olarak kendi içerisinde ayrılır. Bu bilgiler aynı zamanda *çözünürlük* kavramıyla açıklanır.

Spektral çözünürlük, bir algılayıcının elektromanyetik spektrumda kaydedebildiği belirli dalga boyu aralığıdır; aralığın daraltılması spektral çözünürlüğü artırırken spektral aralığın arttırılması çözünürlüğü düşürür. Mekansal çözünürlük, algılayıcı tarafından algılanan bir pikselin temsil ettiği alanın veya ayırt edilebilen en küçük objenin boyutudur. Çözünürlük arttıkça bu obje ya da alanın boyutu küçülür. Radyometrik çözünürlük, her bir banttaki verilerin numaralandırılma aralığıdır; örneğin 8bit veriler 0 ile 255 arasında kaydedilirken, 7bit veriler 0 ile 127 aralığında değerler alır. Zamansal çözünürlük ise algılayıcının aynı alandan hangi sıklıkla veri aldığı ile ilgilidir; arazi kullanımındaki değişimlerin gözlemlenmesi için oldukça önemli bir yere sahiptir (Yener, Koç ve Çoban, 2006).

Sınıflandırma işleminde temel amaç doğal ve yapay objelere ilişkin spektral özellikleri temsil eden görüntü piksellerinin kullanıcı tarafından belirlenen sınırlı sayıda sınıfa atanmasıdır. Bu açıdan ele alındığında uzaktan algılamada sınıflandırma işlemi farklı spektral yansıma değerleri esasına dayanarak orijinal görüntüdeki her bir pikselin arazi örtüsü sınıflarından birisine otomatik olarak etiketlenmesidir (Lillesand ve diğ, 2007).

Uzaktan algılanmış görüntülerden arazi örtüsü belirlemeye yönelik çıkarımlarda bulunabilmek için birçok sınıflandırma metodu geliştirilmiştir. Bu görüntülerin piksellerin bilgilerini içeren satır ve sütunlardan oluşuyor olması nedeniyle kontrollü ve kontrolsüz sınıflandırma yöntemleri arazi örtüsü belirlenmesi için uzun yıllardır kullanılan en temel piksel tabanlı sınıflandırma yöntemlerinden ikisidir (Dean ve Smith 2003).

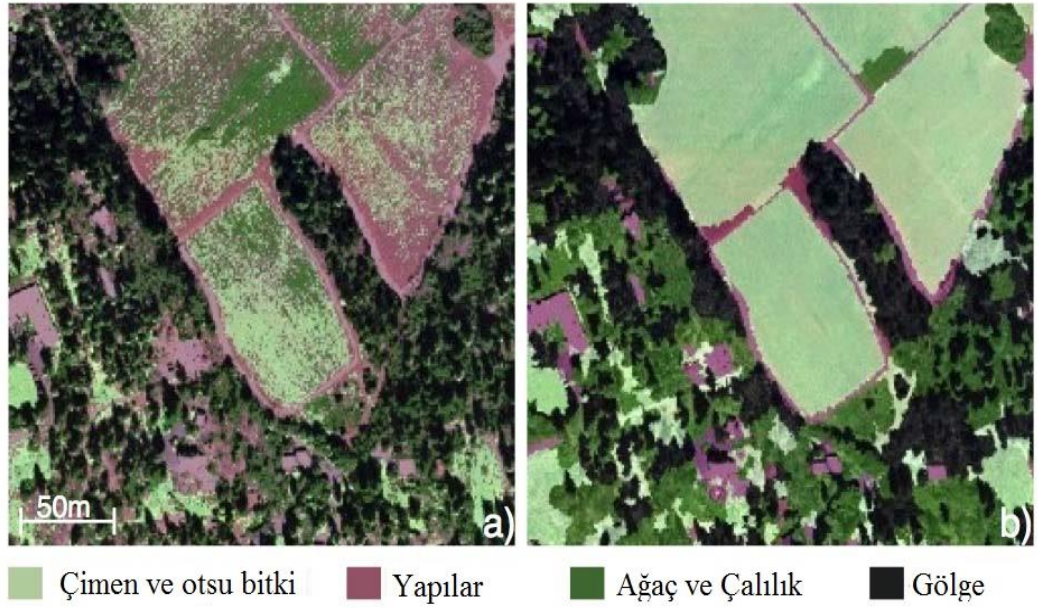
Piksel tabanlı sınıflandırma başlığı altında incelenen bu metotlardaki en belirgin fark; kontrolsüz sınıflandırmada sınıflar algoritma tarafından belirlenirken, kontrollü sınıflandırmada sınıflar kullanıcı tarafından belirlenir. Her iki yöntemde de elde edilecek sınıf sayısı kullanıcı tarafından belirlenir. Ancak kontrolsüz sınıflandırmada elimizdeki spektral değerler, belirlenen sınıf sayısına (n) bölünerek en ayrılabilir n tane sınıf üretilirken, kontrollü sınıflandırmada sınıflar kullanıcı tarafından eğitim alanı atamasıyla gerçekleşir. Kontrollü sınıflandırmanın eğitim alanı atama aşamasında, belirlenen eğitim alanlarındaki spektral değerlerin ortalaması alınır ve görüntünün tamamında bu değere yakın spektral değerlere sahip pikseller, belirlediğimiz sınıfa dahil edilir. Belirlenen sınıfların spektral ayrılabilirlikleri ne kadar yüksek ise sınıflandırmanın o kadar başarılı olacağı öngörülür (Gong ve diğ. 1992).



Şekil 2.1: Kontrollü sınıflandırma yöntemi.

Piksel ve Obje tabanlı sınıflandırma yöntemlerinin karşılaştırıldığı birçok dokümana ulaşmak mümkündür (Cleave ve diğ. 2008; Robertson ve King, 2011), her iki yönteminde birbirleri karşısında kullanım alanlarına göre avantaj ve dezavantajları bulunmaktadır. Ancak günümüzde yüksek çözünürlüklü uzaktan algılama verilerinde obje tabanlı sınıflandırma yöntemlerinin çok daha verimli olduğu bilinmektedir. Bunun nedenlerinden bir tanesi piksel tabanlı sınıflandırmada oluşan

tuz ve biber etkisi (salt & pepper effect)'dir. Bu etkiyi şu şekilde tanımlayabiliriz; sınıflandırma aşamasında örnek alan olarak verilen piksellerin, görüntünün içindeki bazı münferit piksellerin spektral aralığını yakalayamaması nedeniyle görüntüde sınıflandırılmamış piksellere rastlanır. Bu piksellerin oluşma olasılığı görüntünün çözünürlüğü yani piksel sayısı arttıkça artar. Böylece görüntü üzerinde münferit olarak hatalı sınıflandırma kabul edilen piksellere rastlanır. Obje tabanlı sınıflandırmada ise piksellerin komşuluk ilişkileri de işin içine girmektedir, bu durumda her bir obje belirlenen parametrelere göre birçok pikseli bir arada bulundurduğundan sınıflandırılmamış piksellere rastlanılmaz. Tüm bunların ötesinde, piksel tabanlı sınıflandırma günümüzde hala verimlilikle kullanılabilen bir sınıflandırma metodu olarak kabul edilir. Ancak sınıflandırma yapılacak görüntülerin düşük/orta çözünürlüklü olması durumunda obje tabanlı sınıflandırmaya tercih edilmelidir (D. Lu ve Q. Weng 2007).



Şekil 2.2: Tuz ve Biber etkisine örnek. (a) yüksek çözünürlüklü bir uydu görüntüsünün piksel tabanlı sınıflandırma yöntemi ile sınıflandırılmış hali. (b) aynı görüntünün obje tabanlı sınıflandırma yöntemi ile sınıflandırılmış hali (Maggie ve diğ. 2011).

2.1.2. Obje Tabanlı Sınıflandırma Yöntemi

Piksel tabanlı sınıflandırma metodundan farklı olarak, bu sınıflandırma metodunda piksellerin direkt olarak sınıflandırmaya katılması yerine; öncelikle spektral olarak homojen yapıda olan pikseller segmentasyon işlemi ile bir araya getirilip, obje dediğimiz yapıları oluştururlar ve sınıflandırmaya bu şekilde dahil edilirler. Yani, temel olarak bakıldığında bu sınıflandırma yöntemi piksel tabanlı sınıflandırma yönteminden iki yönüyle ayrılır: sınıflandırmaya giren birimler ve sınıflandırmada kullanılan özellikler (Desheng ve Fan 2010). Obje tabanlı sınıflandırma yönteminde oluşturulan objeler kendisini oluşturan piksellerin spektral değerlerinin ortalamasını aldığı için görüntü bazındaki spektral değer farklılıklarını azaltır ve bu sayede daha önce de bahsedilen Tuz & Biber etkisini ortadan kaldırır. Bir diğer yönüyle obje tabanlı sınıflandırmada, objelerin içerisinde bulunan mekânsal/dokusal ve içeriksel özellikler, spektral değerlere ek olarak, sınıflandırma doğruluklarını arttırmak üzere kullanılabilirler (Guo ve diğ. 2007).

2.1.2.1. Parametrik ve Parametrik Olmayan Obje Tabanlı Sınıflandırma Algoritmaları

Sınıflandırma algoritmaları ve aynı zamanda obje tabanlı sınıflandırma algoritmaları parametrik ve parametrik olmayan olarak ikiye ayrılmaktadır (Boiman ve diğ. 2008).

Parametrik algoritmalar, Eğitim-tabanlı sınıflandırma algoritmaları olarak da literatürde geçmektedir (Boiman ve diğ. 2008). Parametrik, yani eğitim-tabanlı algoritmalar, sınıflandırıcının yoğun bir eğitim/örneklendirme süreci istediği algoritmalarlardır (Boch, Zisserman ve Munoz, 2007). Bu sınıflandırıcı algoritmalara örnek olarak Support Vector Machine (Grauman ve Darrell, 2005), Boosting (Opelt ve diğ. 2004), parametrik üretici modeller (Fei-Fei ve diğ. 2004), karar ağaçları (Hamilton, Megown ve Mellin, 2007) ve yapay sinir ağları (Neural Networks) verilebilir. Günümüzde, doğruluk değerleri bakımından en güvenilir olan ve bu nedenle en çok kullanılan sınıflandırma algoritmaları eğitim tabanlı sınıflandırma algoritmalarıdır (Boch, Zisserman ve Munoz, 2007).

Her iki sınıflandırma yönteminde de bir eğitim süreci bulunur, ancak parametrik yöntemlerin sınıflandırma aşamasında karar kuralları (decision rules) oluşturulması için

saf ve tüm görüntüyü temsil edebilecek sayıda eğitim alanının kullanılması esastır. Eğitim alanlarının sayısı ve özellikleri hakkında bilgiler “obje tabanlı sınıflandırma” başlığında verilmiştir.

Parametrik olmayan sınıflandırma algoritmalarında ise sınıflandırma kararları eğitim verileri üzerinden ve eğitim alanlarının istatistiki değerleri ile gerçekleştirilir (Mallinis ve diğ. 2007). Parametrik olmayan sınıflandırma algoritmaları temelinde Maximum Likelihood (ML/En Yüksek Olabilirlik), Nearest Neighbor(NN/En Yakın Komşu), Minimum Distance (MD/En Kısa Mesafe) ve Parallelepiped (Paralel Kenar) gibi yaklaşımlara dayanırlar. Bu sınıflandırıcılar arasında ise NN-Görünü sınıflandırıcısı çok sık kullanılan bir yöntem olsa da doğruluk değerleri parametrik sınıflandırıcılara göre çok daha düşüktür (Varma ve Ray, 2007). Ancak parametrik olmayan sınıflandırıcıların da parametrik sınıflandırıcılara göre önemli avantajları bulunmaktadır.

Parametrik olmayan sınıflandırma metotları ile çok büyük sayıda sınıflar kolaylıkla üretilebilir; bunun yanında bu yöntem ile parametrik sınıflandırmada önemli bir sorun olan algoritmaların aşırı uyum gösterme (overfitting) problemi ortadan kaldırılmış olur (Boiman ve diğ. 2008). Bu sınıflandırma yönteminde eğitim/örnekleme aşaması çok hızlıdır; bu aşama her ne kadar bir kerelik ön işleme süreci olarak görülse de, sınıfların ve eğitim alanlarının değiştirilmesi gerektiğinde büyük veri miktarlarıyla çalışılan projelerde günler alabilmektedir. Sınıfları değiştirmek, eklemek veya çıkartmak parametrik olmayan sınıflandırma yöntemlerinde anlık uygulanabilen işlemlerdir (Hamilton, Megown ve Mellin 2007; (Mallinis ve diğ. 2007).

Ancak bu avantajlara rağmen parametrik olmayan NN görüntü sınıflandırıcıları ve son model parametrik, eğitim bazlı yaklaşımlar arasında doğruluk performansı bakımından uçurumlar bulunmaktadır. Bu açıdan bakıldığında özellikle NN görüntü sınıflandırmaları eğitim bazlı(parametrik) sınıflandırmalar yanında çok güçsüz görünebilir, ancak parametrik olmayan sınıflandırmalara getirilebilecek yeni modifikasyonlar bu açığın kapanmasına yardımcı olabilir, hatta sahip olduğu avantajlar çerçevesinde tercih edilme sebebi dahi olabilir (Boch, Zisserman ve Munoz, 2007).

2.1.2.2. Karar Ağaçları ve Random Forest

Karar ağaçlarının temelinde “veri madenciliği” yöntemleri yatmaktadır. Veri madenciliği, büyük boyutlardaki veriyi değişik perspektiflerden ele alarak analiz etme ve veriyi işe yarar bilgiye dönüştürme işlevi olarak tanımlanabilir (Silahtaroğlu, 2008). Bu yaklaşım iki yöntem ile ele alınabilir: tanımlayıcı ve tahmin edici modeller. Tahmin edici modellerde, sonuçları bilinen veriler ile yola çıkılarak bir model oluşturmak ve oluşturulan bu modelden sonuçları bilinmeyen verilerin sonuç ve değerlerinin tahmini amaçlanır. Tanımlayıcı modellerde ise eldeki verilerden strateji geliştirme ve karar verme süreçlerinde kullanılmak üzere bilgiler sağlanmaktadır. Tanımlayıcı modeller daha çok veriler arasındaki ilişkiyi ortaya çıkartmak için kullanılır (Akpınar H., 2000).

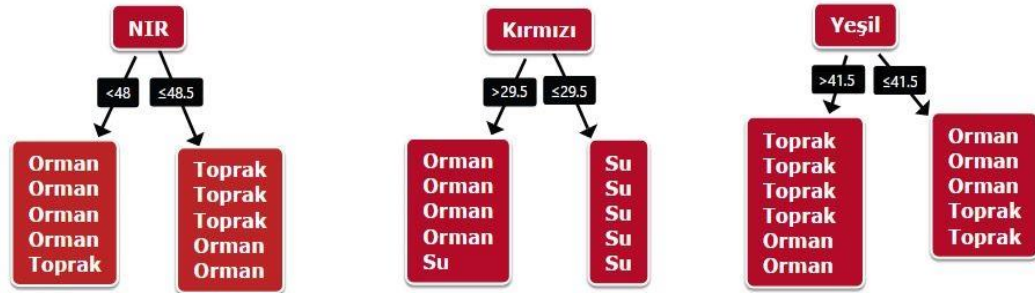
Sınıflandırma genel anlamda, geçmişte toplanan verilerin hangi sınıfa ait olduğu bilindiğinde, yeni gelen verinin hangi sınıfa ait olduğunu bulma işlemidir ve tahmini bir modeldir. Sınıfı bilinen nesneler ile (eğitim alanları) bir model kurulur; kurulan model eğitim kümesinde yer almayan nesneler (test verileri) ile test edilerek performansı ölçülür (Akman ve diğ. 2010).

Karar ağaçları temelinde Kök, Karar ve Yaprak düğümlerinden oluşan ve iş akış şemaları şeklinde dizilim gösteren yapılardır (Şekil 2.4). Buradaki en üst düğüm kök, kökten sonra dallanma gösteren her bir düğüm karar, en diptekiler yaprak, bunların arasındaki bağlantılar ise dal olarak adlandırılır. Karar ağaçlarında dallanma modeli olarak veriler hiyerarşik olarak dallandırılır. Hiyerarşik yapının kullanılmasındaki amaç ise farklı seviyelerdeki aynı detayların objeler arasındaki ilişkiyi nasıl etkilediğini gözlemlemektir. Bu yapıyı içerisinde objelerin özellikleri ile tanımlandığı bir evren olarak hayal edebiliriz. Her bir özellik, objenin önemli bir özelliğini ölçmekte ve birbirinden küçük farklılıklar gösteren ancak tamamen benzersiz olan sonuçlar çıkartmaktadır. Bu yapıyı örneklendirecek olursak; yapıdaki her bir obje, özellikleri bakımından birbirinden ayrı olan sınıflara aittir. Hesaba katılacak en önemli girdiler, hangi sınıflara ait olduğu bilinen örnek alanlar olacaktır.

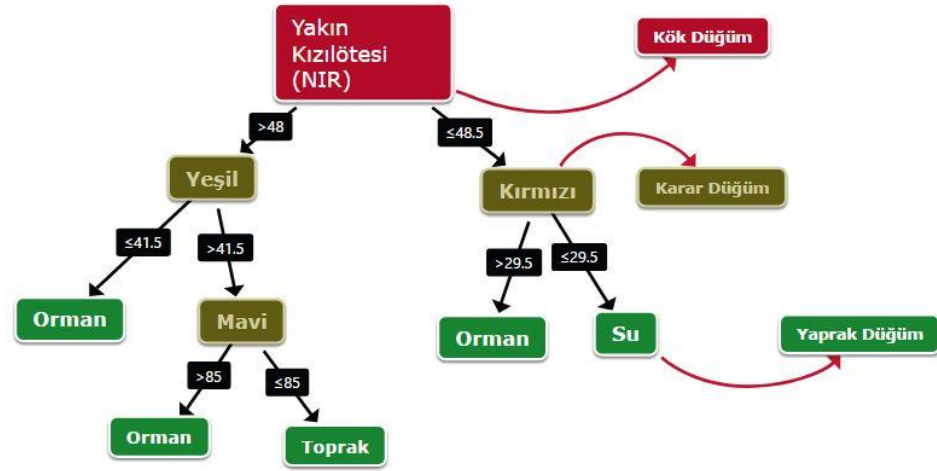
Buradaki amaç bir sınıflandırma kuralı oluşturmak ve bu kural aracılığıyla objelerin özellik değerlerine bakarak onların hangi sınıfa ait olduğunu belirlemektir. İşte bu bahsettiğimiz sınıflandırma kuralı “Karar ağacı” olarak adlandırılır. Bu sınıflandırma kuralının en altında bulunan yaprak düğümleri ise sınıfları belirtmektedir (Silahtaroglu, 2008).

Tablo 2.1: Sadeleştirilmiş bir eğitim sınıfları tablosu.

Numara	Özellikler				Sınıf
	NIR	Kırmızı	Yeşil	Mavi	
1	≤ 48.5	> 40.5	> 41.5	≤ 85	Toprak
2	≤ 48.5	≤ 40.5	> 41.5	> 85	Orman
3	< 48	≤ 40.5	≤ 41.5	> 85	Orman
4	≤ 48.5	≤ 40.5	> 41.5	> 85	Orman
5	< 48	> 40.5	≤ 41.5	> 85	Orman
6	< 48	> 40.5	> 41.5	≤ 85	Su
7	< 48	≤ 40.5	≤ 41.5	≤ 85	Toprak
8	≤ 48.5	> 40.5	> 41.5	≤ 85	Orman
9	≤ 48.5	> 40.5	> 41.5	> 85	Orman
10	< 48	≤ 40.5	≤ 41.5	> 85	Toprak



Şekil 2.3: Tablo 2.1 ‘deki eğitim sınıflarından oluşturulan veri kümeleri.



Şekil 2.4: Şekil 2.3 'deki veri kümelerinden oluşturulmuş bir sınıflandırma kuralı (Karar Ağ -acı). (Çölkesen, 2009) (Değiştirilerek kullanılmıştır.)

Karar ağaçları son yıllarda temel olarak kabul edilip birçok algoritma geliştirilmesinde yardımcı olmuştur. Bu yöntemler tek bir karar ağacı değil, birçok karar ağacının bir araya gelmesi ile oluşturulmuştur (*Ensemble Learning Methods*). Bu karar ağaçlarının bir araya getirilmesiyle oluşturulan yöntemlerden en çok öne çıkanlardan bir tanesi Breiman (2001) tarafından ortaya konulan *Random Forest* yöntemidir. Bu yöntem, *CART* (Classification and Regression Trees/ Sınıflandırma ve regresyon ağaçları) ve rastgele oluşturulan alt uzaylardan (Random Subspace) meydana gelir (Akman ve diğ. 2010).

RF en çok kullanılan Birleşik Öğrenme metotlarından birisi olup (Breiman 2001), çoklu değişken ve sınıfları, karmaşık model ve parametrelere ihtiyaç duymadan sınıflandırabilen bir sınıflandırıcıdır (Akar, Ö. 2017). Aynı zamanda RF birçok karar ağacı temeli olan algoritmalara göre, gürültüye karşı duyarlı olmaması ve model uyumu gerektirmemesi nedeniyle daha üstün kabul edilir (Pal 2003; Gislason ve diğ. 2004; Waske ve diğ. 2007).

Random Forest çok sayıda karar ağacından oluşan, sınıflandırma ve ya regresyon ağaçları oluşturmak için kullanılan ve aynı zamanda kümeleme yapabilen bir sınıflayıcıdır. Sınıfların elde edildiği veri setinden elde edilen değerler kategorik ise sınıflama, sürekli ise karar ağaçları oluşturulur.

Karar ağaçlarının oluşturulmasında kullanılan algoritmalar Classification and Regression Trees (CART) çatısı altında toplanmıştır. Bu algoritmardan bir tanesi olan RF ile oluşturulan ağaçlar *budanmaz* (Şekil 2.3 & Şekil 2.4), ağacın hangi düğümden başlayarak dallara ayrılacağına denk geline her bir değişkendeki *bilgi kazanımına* göre karar verilir. RF algoritmasında her bir düğümden gerçekleşen bu dallanmalar GINI(ismini Corradini'den alır) indeksiyle belirlenir (Breiman ve diğ. 1984; Kononenko and Hong, 1997; Pal, 2002).

RF algoritmasındaki karar ağaçlarını oluşturmak için, veri kümelerinin hangi kademedeki birleşeceklerine karar vermemize yarayacak değer olan GINI indeksi; ilk defa Breiman ve arkadaşları tarafından 1984 yılındaki araştırmalarında bu amaç doğrultusunda kullanılmıştır.

Bu araştırmaya göre GINI fonksiyonu sınıflara ilişkin özniteliklerin safsızlığını ölçer.

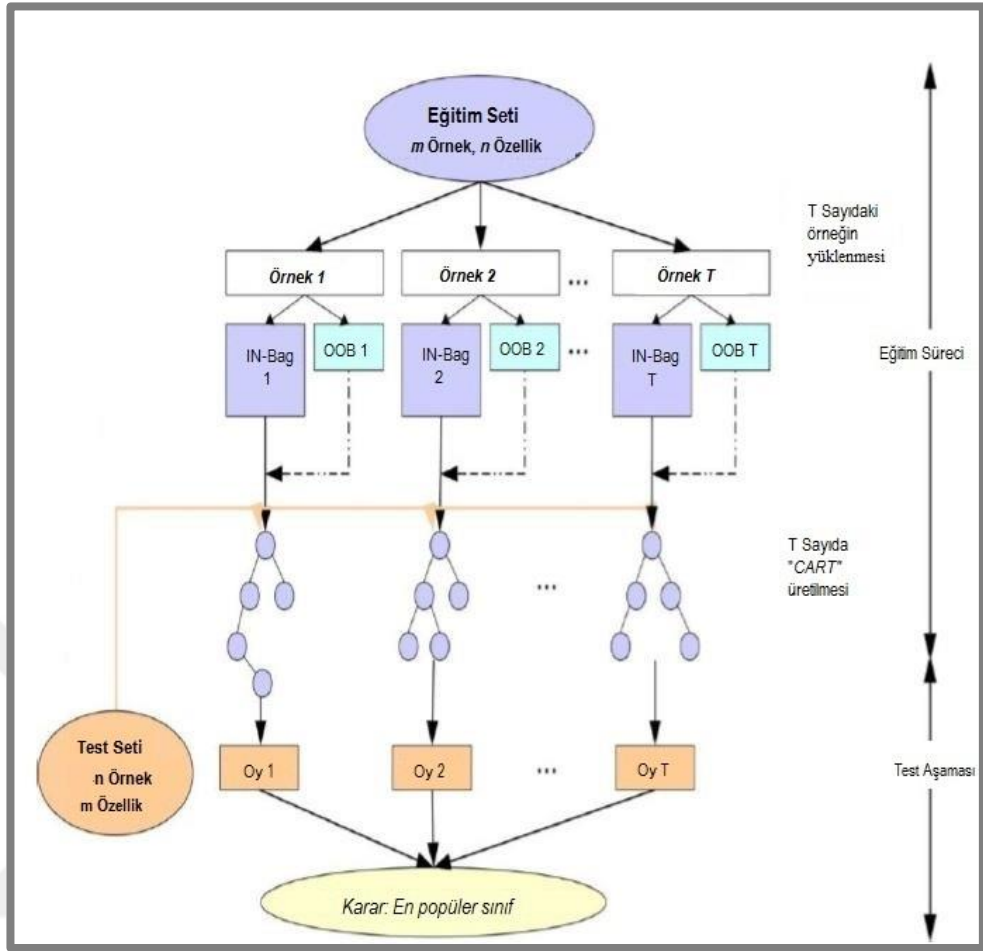
C_i gibi bazı sınıflara sahip T şeklinde verilen bir eğitim setinin;

$$f(C_i, T) / |T| \quad (2.1)$$

şeklinde bir olasılığa sahip olduğu söylenebilir. Genel Gini fonksiyonu veya safsızlığın ölçümü için;

$$\sum \sum_{j \neq i} (f(C_i, T) / |T|)(f(C_j, T) / |T|) \quad (2.2)$$

formülü kullanılır. Gini indeksi basittir ve hızlı bir şekilde hesaplanabilir. Bu indeks rastgele seçilen bir objenin bir düğümden $f(C_i, T) / |T|$ olasılığı ile i sınıfına atanması olarak özetlenebilir (Çölkesen, 2008).

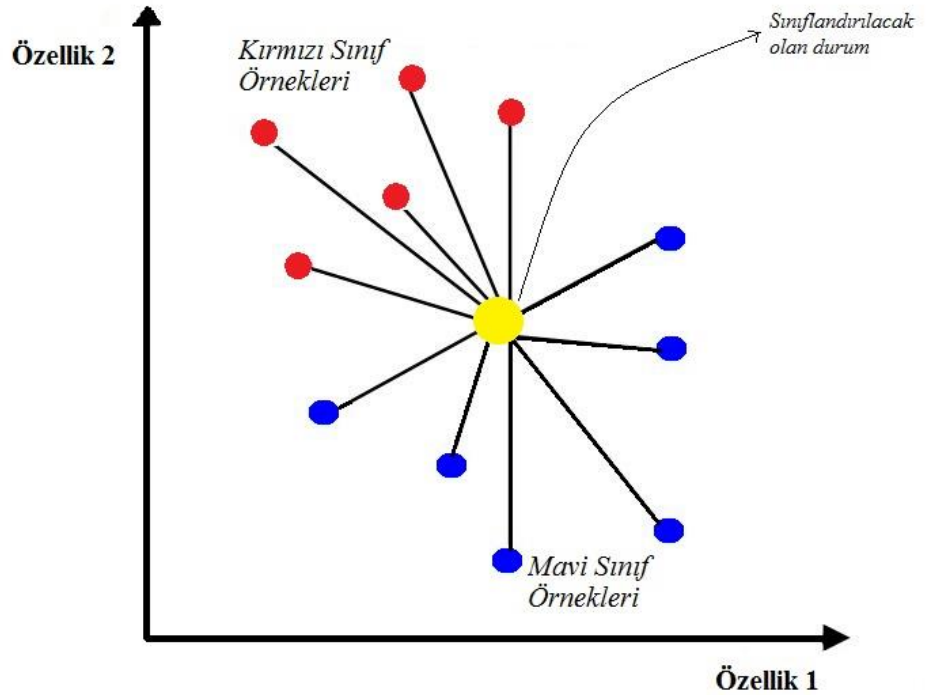


Şekil 2.5: T sayıdaki örnek alanın bir kısmını *In-Bag*, geri kalan kısmını ise *Out-Of-Bag*(OOB) olarak test aşamasında kullanılan ve sınıflandırıcılar arasından en zayıf olanların ağırlığı doğrultusunda karar sınıfı oluşturan bir “ensemble learning” algoritması olan *Random Forest* (Vuolo ve Ng, 2016).

RF sınıflandırıcısının çalışma mantığını anlatırken ele almamız gereken bir diğer önemli konu karar ağaçlarının doğruluk analizlerinin nasıl yapıldığıdır. Bu aşamada görüntüden elde edilen ham verilerle n sayıda düğüm örnekleme oluşturulduğunu varsayarsak, bu örnekleme oranının $2/3$ 'ü karar ağaçlarında öğrenme verisi olarak kullanılır. Öğrenme verisi, yani ağaçları oluşturmak için kullanılan bu düğüm örnekleme oranı *In-Bag* olarak isimlendirilir. Geriye kalan $1/3$ oranındaki veri ise her bir ağacı kendi içinde test etmek için kullanılır. Bu test verileri ise *Out-Of-Bag* olarak adlandırılır. Sonuç olarak her bir ağaç için elde edilen OOB değerleri tüm ormanın hata oranını hesaplamak için bir araya getirilir (Şekil 2.5).

2.1.2.3. Makine Öğrenimi ve Nearest Neighbor Sınıflandırma Yöntemi

Parametrik ve parametrik olmayan algoritmaları anlatırken bu algoritmaların çalışma mantığına değinmiştik. Bu çalışma kapsamında parametrik olmayan sınıflandırma yöntemleri üzerinde durulduğundan; *Makine Öğrenimi* terimini de kısaca tanımlamamız gerekir. Daha önce de bahsettiğimiz gibi parametrik olmayan bu algoritmalar makine öğrenme metotları olarak da bilinir. Bu konu hakkındaki öncü araştırmacılardan biri olan Leo Breiman'ın 2001 çıkışlı "*Random Forests*" yayınında: "Parametrik yöntemlerin aksine makine öğrenme yöntemleri verinin dağılımı ile ilgili herhangi bir varsayıma dayanmadığı için parametrik değildir. Bunlar veri temelli yöntemlerdir, tahminler ve buna karşılık gelen cevaplar arasındaki ilişkileri öğrenirler" tanımlamasında bulunmuştur.



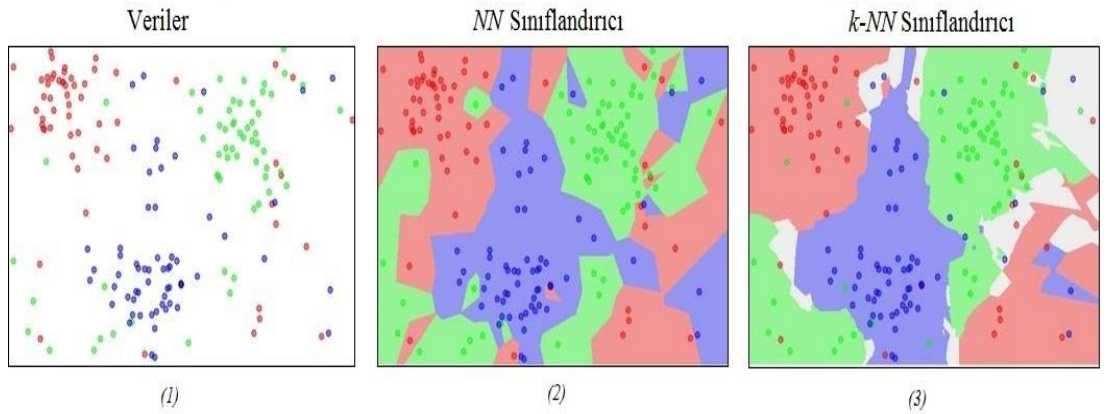
Şekil 2.6: Nearest Neighbor sınıflandırıcı prensibi.

Makine Öğrenimi başlığı altında yer alan ve en temelde bulunan konulardan bir tanesi *Kontrollü Öğrenme (Supervised Learning)* olarak geçer. Bu yöntemde eğitim ve test örneklemeleri mevcuttur. Eğitim örnekleri (x,y) şeklinde tanımlanmış bir veridir; burada x durum, y ise etiketi yani kullanıcı tarafından verilen isimi belirtir.

A olarak hesaba katacağımız test verileri ise etiketleri bilinmeyen x durumlarından oluşur. Buradaki amaç test örneklemeleri için etiketlerin belirlenmesi olacaktır. Kontrollü Öğrenmede kullanılan “*Kontrollü*” kelimesi, eğitim verilerinde her x durumu için bilinen y etiketlerinin bir gözlemci ya da öğretici tarafından tanımlanmasının gerekmesindendir (Kibriya A. ve diğ. 2007).

NN, Fix ve Hodges tarafından 1951 yılında sunulan, parametrik olmayan ve durumlar üzerinden hareket eden bir doku sınıflandırıcısıdır. Daha sonra Cover ve Hart’ın 1967 yılında bu konu üzerine yaptıkları detaylı araştırma ve geliştirmeler sonucu bu yöntem doku sınıflandırıcıları ve makine öğrenimi ile uğraşan araştırmacılar için bir kilometre taşı haline gelmiştir. Veri madenciliği algoritmaları arasında ise ilk on algoritma arasında yer almaktadır (Derrac, Garcia, Herrera, 2013).

NN yöntemi her ne kadar farklı alandaki araştırmalar için temel oluştursa da, en başta gelen kullanım alanı kontrollü sınıflandırma yöntemidir. NN ‘in kontrollü sınıflandırma yöntemlerinde kullanımı k-Nearest Neighbor(KNN); yani yukarıda bahsedilen test alanlarındaki her bir dokunun (örneğin; piksel) kendi etrafında bulunan “ k ” sayıdaki en yakın komşularının arasından en çok benzerlik gösterdiği sınıfa atandığı yöntemdir (Grother, Candela, Blue, 1997) (Şekil 2.6).

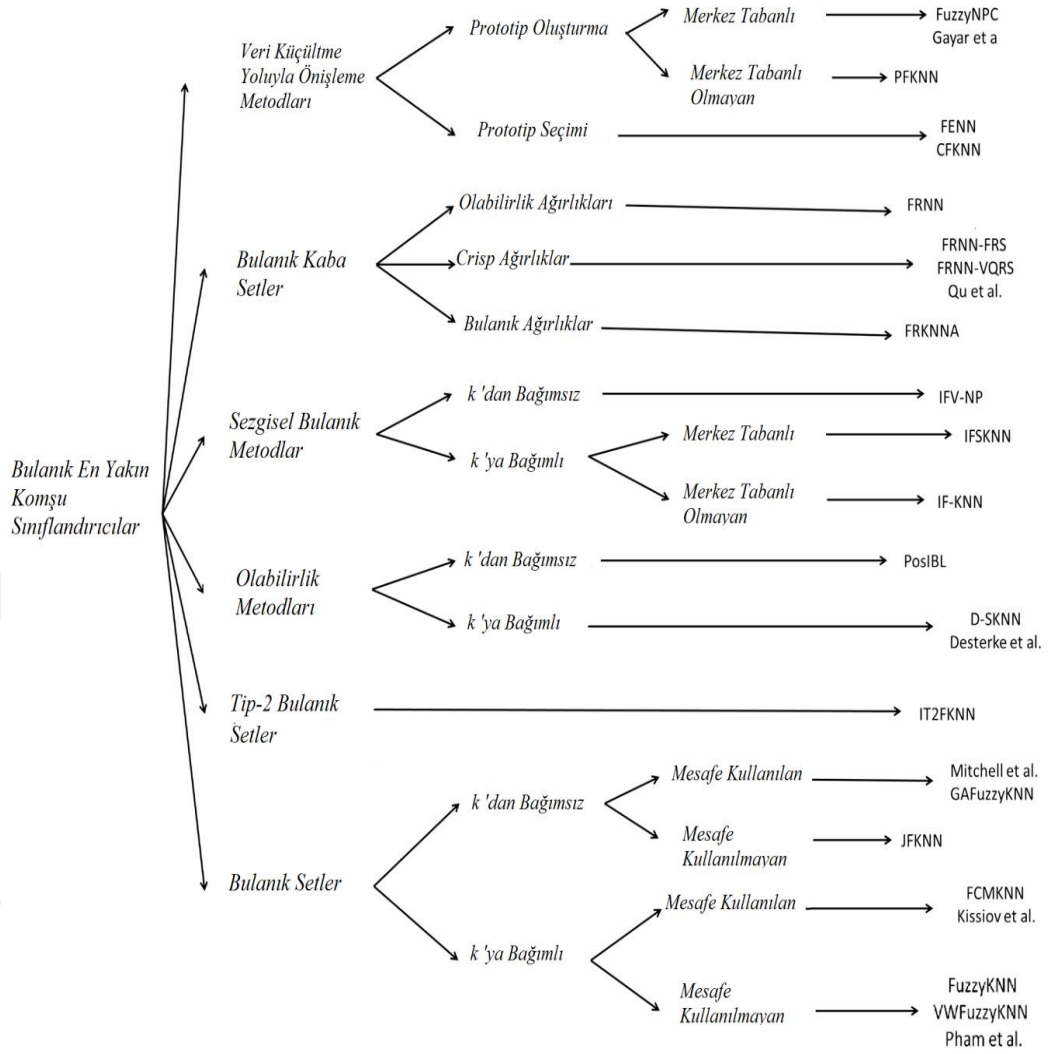


Şekil 2.7: NN ve k -NN sınıflandırıcılar arasındaki fark. (<http://cs231n.github.io/classification> kaynağından değiştirilerek alınmıştır, 2017)

Şekil 2.7'de (1) İki boytlı boşluk uzayında verilerin dağılımını ve 3 farklı sınıf kullanılarak (yeşil - kırmızı -mavi) oluşturulan karar sınırları arasındaki farkı görebiliriz. (2) NN sınıflandırıcısında karar sınırları içerisinde bulunan kırmızı ve yeşil adacıklar yanlış sınıflandırılma sonucudur, bu durumda beklenen kırmızı ve yeşil adacıkların mavi sınıfa dahil edilmeleridir. (3) k-NN sınıflandırıcıda genelleştirmenin yani karar sınıfları içerisinde bulunan noktaların aynı sınıflara atanma işleminin daha düzgün yapıldığı görülebilir.

Son yıllarda ön planda olan ve sıklıkla kullanılan obje tabanlı sınıflandırma yazılımı eCognition, sınıflandırma yöntemi olarak bulanık mantık (fuzzy logic) ile ilişkilendirilmiş k-NN algoritmasını kullanmaktadır. Literatürde bu yöntem Fuzzy Nearest Neighbor Classification (bulanık en yakın komşu sınıflandırma) olarak geçmektedir. Bu yöntemi klasik k-NN sınıflandırıcılarından ayıran en önemli özelliklerinden biri k değerinin otomatik olarak veriler üzerinden belirlenmesidir. (Derrac, Garcia, Herrera, 2013).

Bulanık En Yakın Komşu sınıflandırıcıları da kendi arasında belirli parametrelere sahip olma ve olmama durumuna göre birbirinden ayrılmaktadır, Şekil 2.8 'de bu sınıflandırıcıların taksonomisi verilmiştir.



Şekil 2.8: Derrac, Garcia ve Herrera tarafından yazılan 2013 yılı çıkışlı “Bulanık en yakın komşu algoritmaları: taksonomi ve deneysel analizler” araştırmasında önerilen Bulanık En Yakın Komşu sınıflandırıcıları taksonomisi.

3. MALZEME VE YÖNTEM

Çalışmada kullanılan uydu görüntüleri ve haritalar; 2016 tarihli *Sentinel-2a* ve 2011 tarihli *Worldview-2* uydu görüntüleri, meşçere tipi haritaları ve *Google Earth Pro* yazılımından elde edilmiştir. Doğruluk analizleri için gereken yer ölçümlerinde 2012 tarihli amenajman planından yararlanılmıştır.

2016 tarihli *Sentinel-2a* görüntüleri *ESA (European Space Agency)*'nin resmi sayfasından temin edilmiş ve yöntem içerisinde anlatılacak olan işlemlerden; örnekleme, vejetasyon indeksleri ve görüntü ön işleme süreçleri *SNAP (Sentinel Application Platform)* içerisinde gerçekleştirilmiştir. Görüntülerin zenginleştirilme, bant birleşimleri, doğruluk tespitleri, sınıflandırma vb. uygulamaları için ise Quantum GIS, ArcGIS, Erdas Imagine, Envi ve eCognition yazılımları kullanılmıştır.

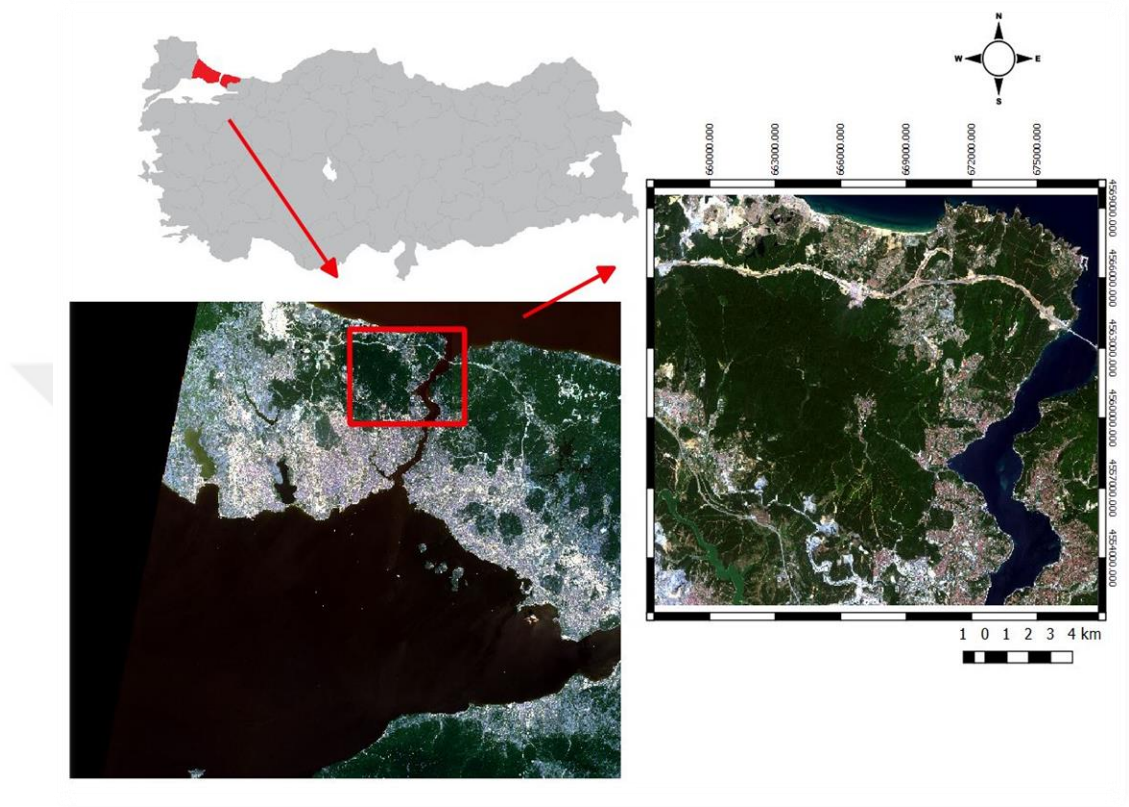
Araştırmada kullanılan asıl uydu görüntüleri *Sentinel-2a* olmasının yanında, doğruluk analizleri, örnekleme doğrulukları ve segmentasyon aşamasındaki parametrelerin belirlenmesi için gereken yüksek çözünürlüklü uydu görüntüsü olarak *Worldview-2*'den faydalanılmıştır. Ayrıca çalışma kapsamındaki en önemli noktalardan birisi olarak kabul edilen *Random Forest* algoritmaları, alansal ve spektral istatistiklerin hesaplanması için gereken istatistiksel yazılım için *R Studio* kullanılmıştır.

3.1. MALZEME

3.1.1. Çalışma Alanı ve Özellikleri

Çalışma alanı olarak seçilen bölge İstanbul il sınırları içerisinde ve değişimin kolayca gözlemlenebileceği bir bölge olan kuzey ormanlarıdır. Bu alanın seçilme nedenlerden bir diğeri ise arazi çalışmalarının yapılmayacak olması ve gözlemlenen bölgenin yersel verilerine kolayca erişilebilecek olmasıdır. Bu nedenle Belgrad ormanları araştırmada, değişime neden olan faktörlerin de bilinmesi nedeniyle, kullanılması kararlaştırılmış bir alandır. Çalışma alanı 6 farklı işletme şefliğinin sınırları içerisinde bulunmaktadır,

bunlar: Atatürk Arboretumu, Bentler, Kemerburgaz, Kurtkemer, Beykoz ve Sarıyer Orman İşletme Şeflikleridir.



Şekil 3.1: İstanbul ili ve çalışma alanı sınırları.

3.1.1.1. Coğrafi Konum

Belgrad Ormanı, İstanbul ilinin Avrupa Yakası'nda, yerleşim alanlarının kuzeyinde, $28^{\circ} 53' 25''$ - $29^{\circ} 00' 55''$ doğu boylamı ile $41^{\circ} 09' 44''$ - $41^{\circ} 14' 40''$ kuzey enlemleri içerisinde yer almaktadır. Mülki sınırlarına bakıldığında Sarıyer ve Eyüp ilçeleri arasında yayılım gösterdiği görülmektedir. İdari olarak İstanbul Orman Bölge Müdürlüğü'nün yetki alanı içinde kalmaktadır. Orman sınırları içindeki mesire yerleri ise Bahçeköy Orman İşletme Müdürlüğü'nün yetki alanı içerisinde (Bıçak G., 2007).

3.1.1.2. İklim ve Topografik Yapı

Çalışma alanı Marmara Bölgesi'nde bulunmaktadır ve iklim olarak bölgenin karakteristiklerini taşımaktadır. Yazları sıcak ve kurak kışları ılıman ve yağışlıdır. Deniz seviyesine göre bölgenin en alçak yeri 30 metre ve en yüksek noktası 219 metredir.

Boğaz platosunun kuvvetli bir şekilde yarılmış olması, Boğaz oluşunun hakim rüzgar yönünde açılmış bulunması ve muhtelif kısımlar arasındaki yükselti ve bakı farkları İstanbul Boğazı çevresinde oldukça kuvvetli özelliklerle beliren yerel iklim tiplerinin oluşumuna yol açan esas sebeptir. Bu topoğrafya ve bakı nedeniyle boğaz kıyıları arasında önemli ısı, nem ve rüzgar şiddeti farklılıkları görülmektedir. İstanbul Boğazı ve çevresi Akdeniz ikliminin etkisi altındadır. Fakat bu yörede kuraklık batı ve güney kıyılarımızda olduğu kadar şiddetli olmadığı gibi, süresi de kısadır. Yörede Akdeniz makro ikliminin, don tehlikesinin de mevcut olduğu kışı şiddetli ve genellikle daha nemli bölgesel bir tipi olarak karşımıza çıkar. Akdeniz iklimi ile Karadeniz iklimi arasında bir geçiş iklimine sahiptir (Anonim / OGM, 2012, syf. 46).

İstanbul yağış özellikleri ile kuzey ve güney arasında bir geçiş alanı meydana getirir. Yıllık yağış güneyden kuzeye genel bir artış gösterir. Marmara kıyılarında 650mm'ye kadar olan yıllık yağış, kuzey kıyılarda 750mm dolayındadır. Yörenin 200metreden daha yüksek kısımlarında ise bu miktar 1000mm'ye yaklaşır veya geçer(Bahçeköy 1074,4mm). Jeokompleksin sınırları içinde yağış miktarının dağılışında yükselti, bakı, konum ve rölyef başlıca rol oynar (örneğin derin yarılmış vadilerin tabanlarında yağış daha azdır). Doğu ve batı yarıları arasında rölyef bakımından mevcut olan asimetri dolayısıyla, kuzeye doğru ilerledikçe meydana gelen yağış artışı batıda tedrici, doğuda ise çok daha kısa mesafelerde değişebilir. Geçiş karakterinin diğer bir belirtisi, yörede yaz kuraklığının kuzeye doğru hafiflemesi ve kısılmasıdır. Kurak ve yarı kurak ayların sayısı Trakya'nın Marmara kıyılarında 5 ay olduğu halde, yörenin kuzeyinde 2 aya kadar düşer. İstanbul'un güney kesimleri kuzeye oranla daha fazla ısınmakta ve vejetasyon devresi nispeten daha sıcak geçmektedir (Anonim / OGM, 2012, syf. 47).

3.1.1.3. Bitki Örtüsü

Belgrad Ormanının hakim ağaç türü meşelerdir (*Quercus robur*, *Q. fainetto*, *Q. petraea*, *Q. cerris*). Ayrıca, Doğu Kayını (*Fagus orientalis*), Adi Gürgen (*Carpinus betulus*), Anadolu Kestanesi (*Castanea sativa*), Titrek Kavak (*Populus tremula*), Adi Fındık (*Corylus avellana*), Akçaağaç türleri (*Acer campestre*, *A. trautvetteri*), Ova Karaağacı (*Ulmus minor*), Gümüş İhlamur (*Tilia argentea*), Kızılağaç (*Alnus glutinosa*), Söğüt türleri (*Salix caprea*, *S. cinerea*), gibi nem isteği yüksek olan türler plan ünitesinde ön plana çıkmataırlar. Doğal meşcereler yukarıda belirtilen meşe, gürgen, kayın, kestane, kızağaç, ıhlamur saf ya da karışık meşcereleri ile kendi aralarında karışım oluşturduğu meşcerelerden oluşturmaktadırlar. Diğer türler meşcere tiplerine karışım oranının altında serpili olarak katılmaktadırlar (Anonim / OGM, 2012, syf. 55).

Plan ünitesinde yukarıda verilen doğal ağaç, ağaççık türleri dışında yapay yolla oluşturulan meşcereler de yer almaktadır. Bu meşcereler ise; *Pinus brutia* (Kızılçam), *Pinus nigra* (Karaçam), *Pinus sylvestris* (Sarıçam), *Abies bornmuelleriana* (Uludağ Göknarı), *Pinus pinea* (Fıstıkçamı), *Pinus maritima* (Sahilçamı) ve *Robinia pseudoaccacia* (Yalancı akasya) türlerinden oluşmaktadır (Anonim / OGM, 2012, syf. 56).

Çalışma kapsamında uydu görüntülerinden elde edilmesi hedeflenen bitki örtüsü sınıfları ibreli, yapraklı ve zayıf vejetasyon olmak üzere 3 sınıfa ayrılacaktır. Bu nedenle ilgi alanı doğrultusunda belirtilen asli odunsu türler yapraklılar için Meşe, Kayın, Gürgen, Kestane; ibreli türler için Kızılçam, Karaçam, Sarıçam, Göknar, Fıstıkçamı ve Sahilçamı olarak belirlenmiştir.

3.1.2. Çalışmada Kullanılan Veriler ve Yazılımlar

Çalışmada kullanılan veriler Sentinel2a, WorldView2 uydu görüntüleri ve 2012 yılına ait amenajman planlarıdır. Görüntülerin işlenmesi ve sınıflandırma aşamasında QGIS, eCognition ve RStudio yazılımları kullanılmıştır, ayrıca sınıflandırma doğruluklarının belirlenmesi için Erdas Imagine yazılımından yararlanılmıştır.

Bu veriler hakkında detaylı bilgiler ve yazılımların hangi aşamada, nasıl kullanıldıkları ilgili başlıkların içeriğinde verilecektir.

3.1.2.1. Uydu Görüntüleri

Uydular Dünya'nın haritalanma, ölçme ve gözlemlenmesi açısından sahip olduğumuz en önemli teknolojilerden birisini oluşturmaktadır (Twonshend ve diğ. 2008). Bu uydular ile küresel anlamda vejetasyon varlığı ortaya koyulabilirken, daha lokal ve spesifik çalışmalar gerçekleştirilebilir (Loveland ve diğ. 1999). Bu durumda kullanılacak olan uydunun özellikleri düşünülmelidir. Günümüzde hangi uydular bulunmakta; araştırmacılar için erişim imkanları; hangi algılayıcılarla görüntüleme sağladıkları; mekânsal çözünürlüklerinin neler olduğu ve hangi kurum tarafından geliştirilmiş olduğu gibi sorular ilk akla gelenlerdir (Cihlar, 2000).

Kamusal, ticari ve askeri kullanıma tabi yüzlerce uydu bulunmaktadır. Kamusal sistemler devletler tarafından finanse edilmiş olup herkesin kullanımına açıktır; ticari uydular ise devletler tarafından yetkilendirilip şirketler tarafından finanse edilir ve bu uydular tarafından elde edilen veriler ticari malzeme olarak kar elde etmek için kullanılır. Askeri uydular ise devletler tarafından finanse edilir ancak kamusal kullanıma çok kısıtlı olarak açık olup istihbarat sağlamak için kullanılırlar (Baker ve diğ. 2001).

Uydu görüntüleyicileri teknoloji ile birlikte yıllar boyunca değişimlere uğramıştır. Birinci nesil uydu görüntüleyicileri ilk olarak askeri istihbarat sağlamak için geliştirilmiş ve algılayıcı olarak geleneksel fotoğraf filmlerini kullanmıştır (örneğin *Corona* ilk istihbarat uydularından olarak kabul edilir). Daha sonraki dönemlerde bu geleneksel yöntem yerini dijital görüntüleme metotlarına bırakmıştır (Davis, 2007).

Günümüzde Avrupa Birliği, ABD, Kanada, Japonya, Çin, Rusya ve Hindistan gibi gelişmiş ülkelerin uzay ajansları ve ayrıca ticari şirketler tarafından finanse edilmiş birçok uydu sistemleri mevcuttur. Bu çalışma kapsamında kullanılması amaçlanan uydu verilerinin en önemli özellikleri gerçekleştirilecek obje tabanlı sınıflandırma metotlarının ideal hız ve doğrulukta gerçekleşmesini sağlayacak mekânsal çözünürlük-

-lülere, ayrıca erişim bakımından en az maliyete sahip olmasıdır. Uyduların sahip olması gereken bu özellikler göz önünde bulundurulduğunda *Avrupa Uzay Ajansı* tarafından 2015 yılında faaliyete giren *Sentinel-2a* uydusu bu çalışma için uygun görülmüştür.

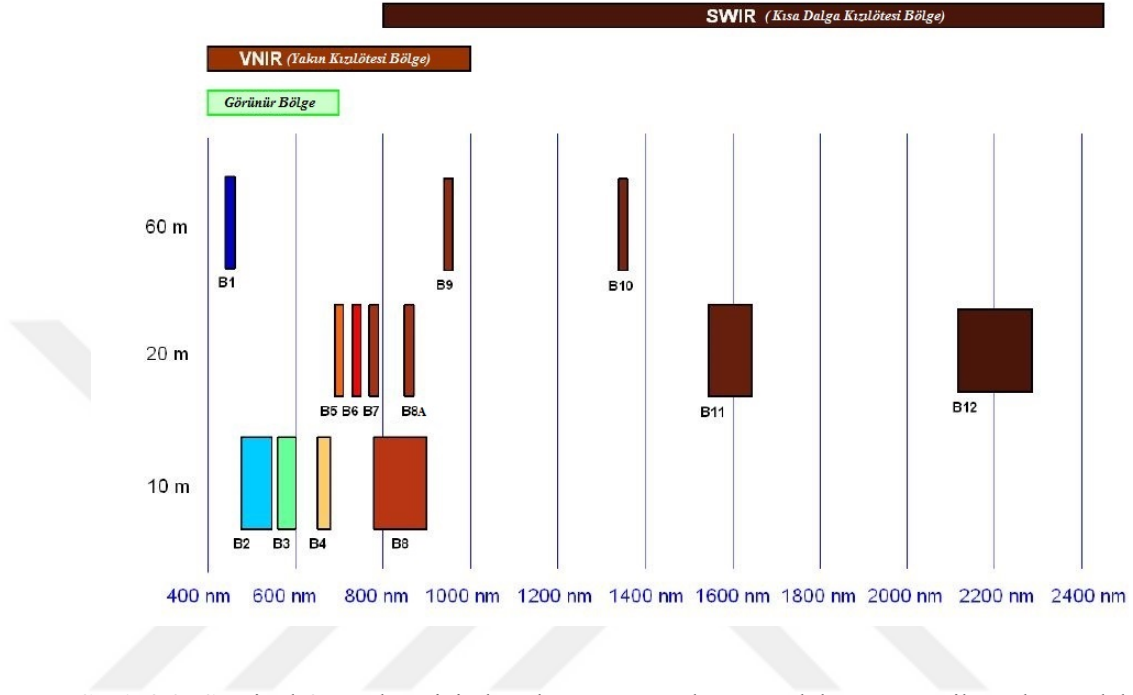
Ayrıca sınıflandırma doğruluklarının kontrol edilmesi için gereken yüksek çözünürlüklü uydu verileri için WorldView-2 (0.46 metre) ticari uydusunun çalışma alanına ait 2011 tarihli görüntüleri çalışma kapsamında kullanılmıştır.

Sentinel Uydu Görüntüleri

Avrupa Uzay Ajansı (ESA) tarafından finanse edilen *Copernicus* programı dahilinde geliştirilen *Sentinel* uyduları, endüstriyel ve çevresel olayların gözlemlenmesi için gereken sürdürülebilir ve kontrollü veri kaynakları sağlanması amacıyla geliştirilmiştir. Dünya'nın gözlenmesine yönelik başlatılan Copernicus programı dahilinde bulunan Sentinel uyduları 6 farklı uydu ve amaç doğrultusundaki ekipmanlar ile geliştirilmiştir. Her bir Sentinel görevi ise kendi içerisinde 2 farklı uydu takımı ile uyduların tekrarlama(dünyanın etrafında bir dönüşü tamamlama) ve kapsama alanlarının zenginleştirilmesi amacıyla tasarlanmıştır (örneğin Sentinel 3a ve Sentinel 3b).

Sentinel 1a ve 1b uyduları gece-gündüz radar görüntüleme sağlayan ve sahip olduğu cihazlar ile yeryüzü ve okyanus yüzeylerinin incelenmesi amacıyla geliştirilmiştir. Sentinel 2a ve 2b uyduları ise vejetasyon, toprak, su ve kıyısız alan gözlemlerinde kullanılmak amacıyla multispektral yüksek çözünürlüklü görüntü elde etmek üzere geliştirilmiş, aynı zamanda doğal afet yönetimi ve risk analizleri için altlık oluşturması planlanan uydu takımıdır. Sentinel 3a ve 3b uyduları deniz ve yeryüzü topografyası, sıcaklık ve renk değişimlerinin gözlemlenmesi için tasarlanmış uydulardır. Sentinel 4 ve 5 uyduları atmosferdeki değişimlerin görüntülenmesi için geliştirilen uydulardır; Sentinel 5 güneş senkronizasyonunda iken Sentinel 4'ün yersel istasyonlar ile senkronize edilmesi planlanmıştır. Son olarak Sentinel 6 içeri-

-indeki *Radar Altimetre* ile deniz ve yeryüzü yükseltilerinin hesaplanması ve iklimsel değişimlerin bunlar üzerindeki etkilerinin gözlemlenmesi hedeflenmiştir (ESA, 2016).

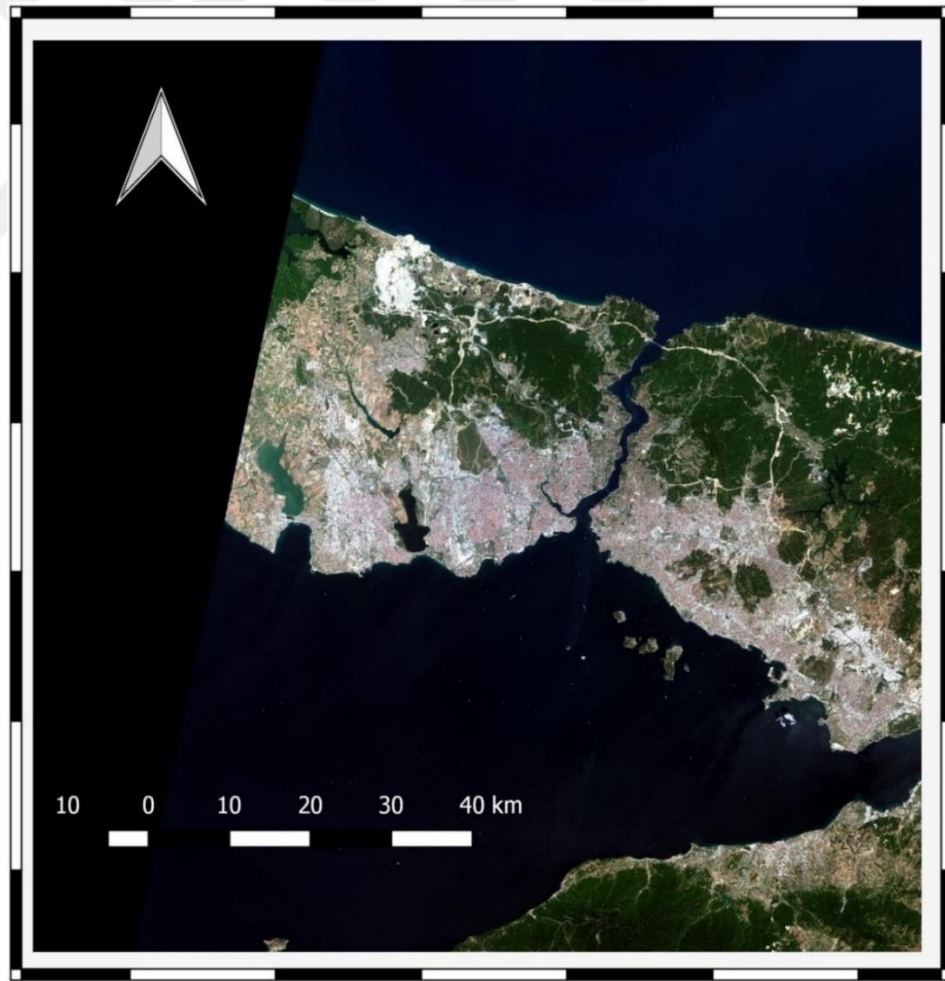


Şekil 3.2: Sentinel-2a uydusu için bantların çözünürlüğü ve elektromanyetik spektrumdaki yerleri. (ESA, 2016). (Değiştirilerek kullanılmıştır)

Tablo 3.1: Sentinel 2a uydusu bant genişlikleri. (ESA, 2016) (Değiştirilerek Kullanılmıştır)

Mekansal Çözünürlük (m)	Bant Numarası	Merkez Dalgaboyu (nm)	Bant Genişliği (nm)
10	2	490	65
	3	560	35
	4	665	30
	8	842	115
20	5	705	15
	6	740	15
	7	783	20
	8a	865	20
	11	1 610	90
	12	2 190	180
60	1	443	20
	9	945	20
	10	1 380	30

Bu çalışma kapsamında kullanılan asıl veri, Sentinel 2a uydusundan elde edilen 24 Temmuz 2016 tarihli ve 35TPF sıra numaralı multispektral, orta / yüksek çözünürlüklü uydu görüntüsüdür (Şekil 3.3). 2015 yılında görüntü alımına başlanan Sentinel-2 programı diğer Sentinel görevleri gibi 2 senkronize uydudan oluşmaktadır; buradaki amaç uyduların yüzeydeki tarama alanını arttırmak ve başlanılan noktaya dönmesi için gereken süreyi azaltmaktır; Landsat-8 için 16 gün olan bu süre Sentinel-2 programında 5 güne düşürülmüştür. Sentinel 2a / 2b uydularındaki bantlar spektrum'un görünür kısmından başlayarak kısa dalga kızılötesine (SWIR) kadar, 10 metreden 60 metreye kadar mekansal çözünürlüklere sahip (Şekil 3.2) görüntüler almak için tasarlanmıştır (Lefebvre, A. ve diğ. 2016).



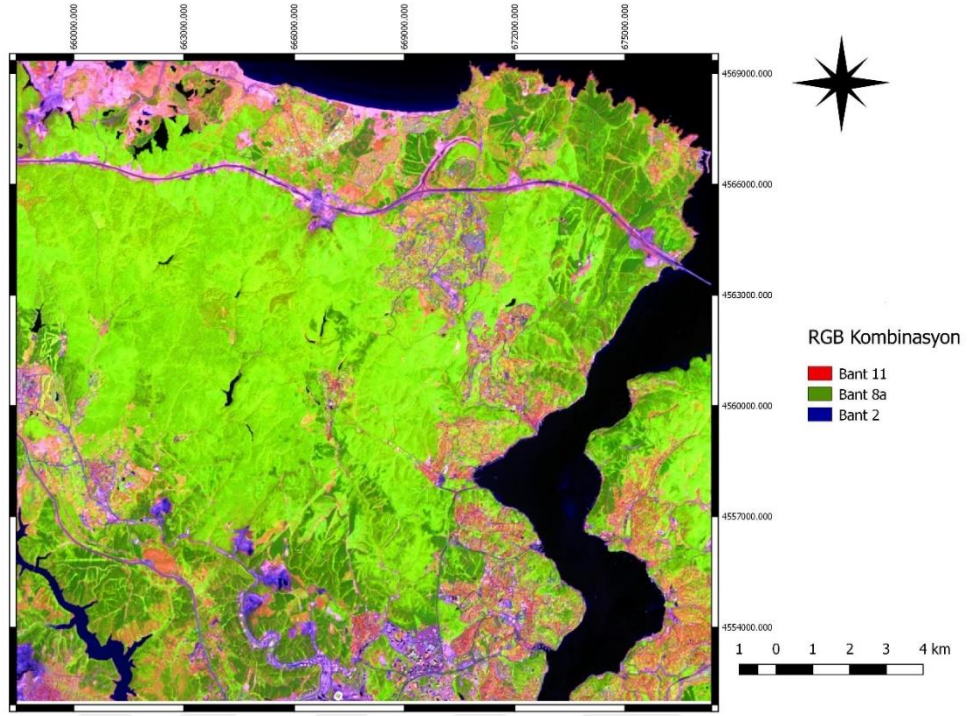
Şekil 3.3: 24 Temmuz 2016 tarihli 13 bantlı Sentinel-2a görüntüsü.

Bant Kombinasyonları ve Vejetasyon İndeksleri

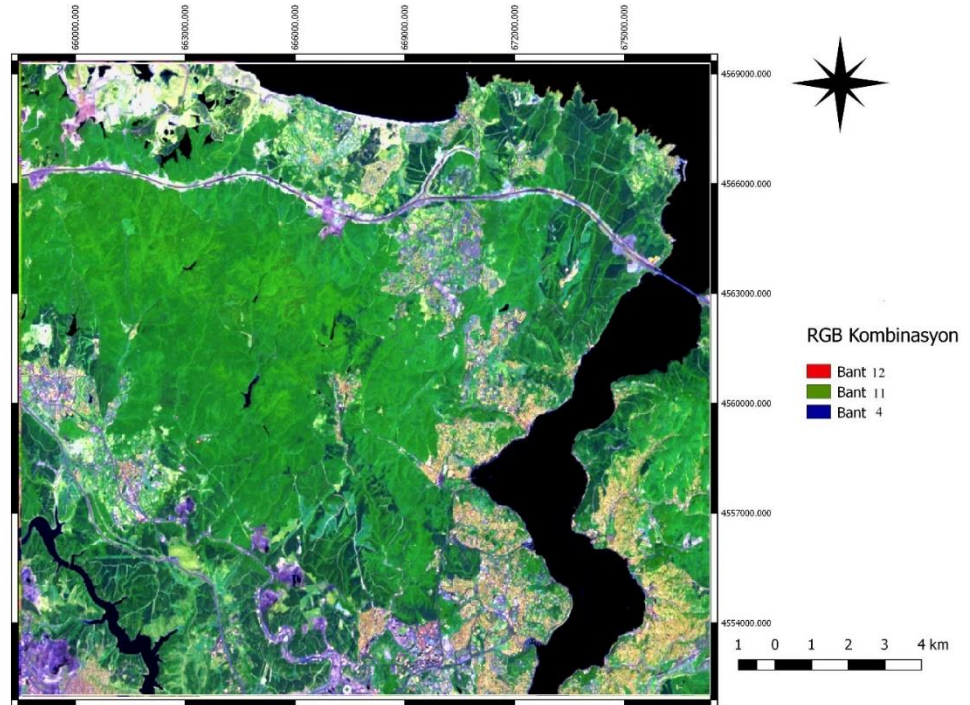
Yüksek mekansal çözünürlüğe sahip uydu görüntüleri yardımıyla yeryüzü nesnelerin konumlarına ve dağılımlarına ilişkin detaylı bilgilerin elde edilmesi mümkündür. Buna karşın yüksek çözünürlüklü görüntüler heterojen yapıda ve benzer yansıma değerlerine sahip yoğun görüntü piksellerinden oluşmaktadır. Karmaşık yapıdaki pikseller arasındaki spektral ayrımın artırılması amacıyla uydu görüntüleri ile birlikte yardımcı verilerin bir arada kullanılması gereklidir. Entegrasyon sonucunda elde edilen veri seti yüksek boyutlu ve korelasyonlu bantları (özellikleri) içermektedir. Bu nedenle yüksek boyutlu veri seti içerisinde sınıflandırma probleminin çözümüne katkı sağlayacak bantların belirlenmesi önemli bir araştırma konusudur (Çölkesen ve diğ. 2015).

Bant kombinasyonlarının kullanımı, çıplak gözle fark edilemeyecek detayların insan gözü tarafından algılanabilecek renk ayrımları elde ederek görüntüler üzerinde yorumlama yapabilmemize olanak verir. Burada renk ayrımlarından kasıt farklı yansıma değerlerine sahip olmasına rağmen insan gözüne aynı renkte görünen cisimlerin elektromanyetik spektrumdaki farklı kanallar ile yeniden renklendirilmesidir. Bu yöntemin ormancılık, ziraat ve şehircilik çalışmalarında kullanımı oldukça yaygındır, Sentinel 2 görüntüleri için oluşturulan bant kombinasyonları aşağıda yer almaktadır.

Şekil 3.4’de gösterilen bant kombinasyonu zirai ürünlerin gözlemlenmesi için kullanışlıdır. Görüntüdeki açık yeşil alanlar vejetasyonun güçlü ve sağlıklı olduğu bölgeleri ve geniş yapraklı ağaçları, koyu yeşil alanlar ibreli türleri belirtir. Vejetasyonun zayıf olduğu alanlar ve açık toprak kahverengi ile leylak rengi arasında yer alır (ESA, 2016).



Şekil 3.4: Sentinel 2 uydusuyla zirai alanların gözlemlenmesi için belirlenen bant kombinasyonu (Kırmızı Filtre: 11, Yeşil Filtre: 8a ve Mavi Filtre: 2)

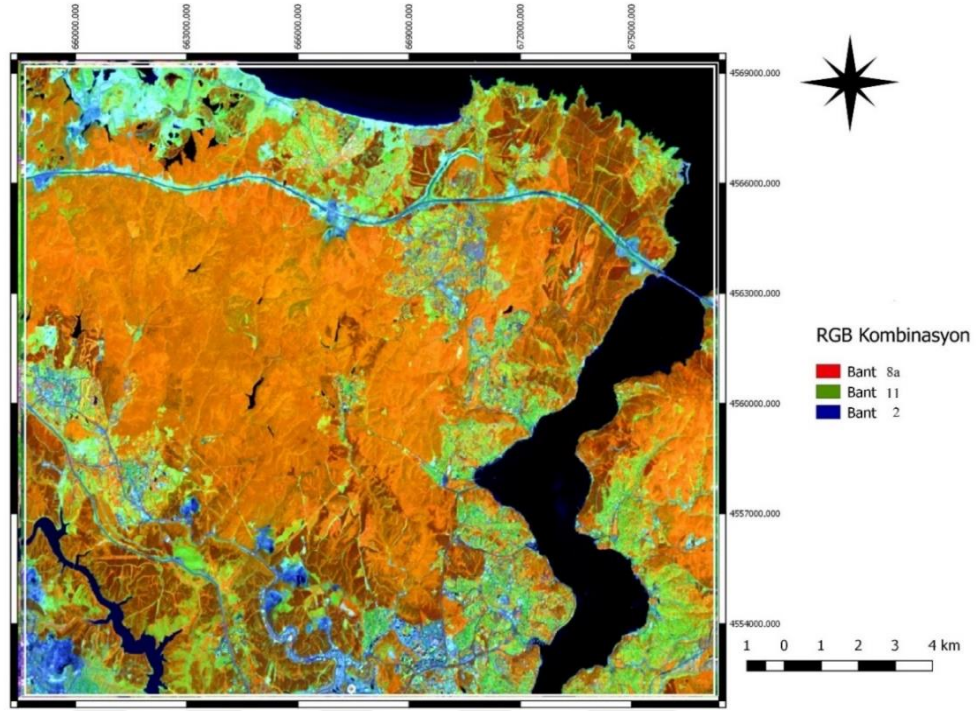


Şekil 3.5: Sentinel 2 uydusuyla atmosferik geçirgenlik sağlaması için kullanılan bant kombinasyonu. (Kırmızı Filtre: 12, Yeşil Filtre: 11, Mavi Filtre: 4)

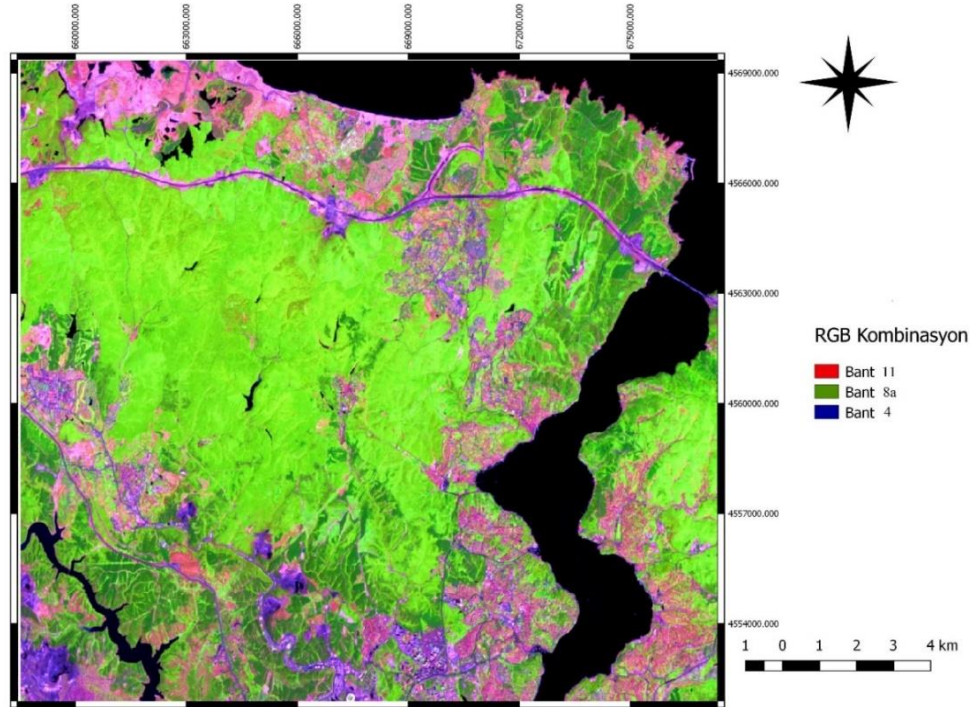
Bu bant kombinasyonu ile atmosfer partiküllerinin etkisi azaltılıp, sis ve duman gibi etmenlerin aşılması sağlanırken “doğal” renklere yakın renkler elde edilir. Vejetasyon süresi boyunca bitki örtüsü yeşil ve yeşilin tonları olarak görünürken şehirler beyaz, gri, siyan ya da mordur. Suyun yansımalarının neredeyse tamamen absorbe edildiği bu bantlarda su diğer cisimlerden kesin bir şekilde ayrılır. Bu kombinasyonda kar ve buz koyu mavi, sığ sular koyu mavi görünürken derin su siyahtır. Orman yangınları ya da volkan kraterleri gibi ısının yüksek olduğu bölgeler ise sarı ve kırmızının tonları olarak görülebilir. Bu kombinasyon özellikle orman yangınlarının gözetlenmesi için kullanılmaktadır. Ayrıca sel felaketine uğramış bölgeler koyu mavi ya da siyah olarak göze çarpar (Şekil 3.5 / ESA, 2016).

Sağlıklı vejetasyonunun belirlenmesi için kullanılan bu bant kombinasyonunda bitki örtüsü sarı, turuncu, kahverengi ve kırmızının tonları olarak göze çarpar. Toprak yeşil veya kahverengidir. Şehir ve yapılaşma beyaz, siyan ve gri, parlak mavi alanlar tıraşlama kesimleri, kırmızımsı renkler ise genç vejetasyonu ya da mera alanlarını gösterir. Temiz ve derin sular koyu mavi, sığ ya da sediment içeren sular ise mavinin açık tonları olarak göze çarpar. Vejetasyon üzerine yapılacak çalışmalarda, orta-kızılötesi bandın eklenmiş olması bitkinin büyüme evrelerini ve ya stresini gözetleme şansı sunar. Bu bant kombinasyonu şehirler ve yapılaşmanın olduğu bölgeler için iyi bir seçim değildir (Şekil 3.6 / ESA, 2016).

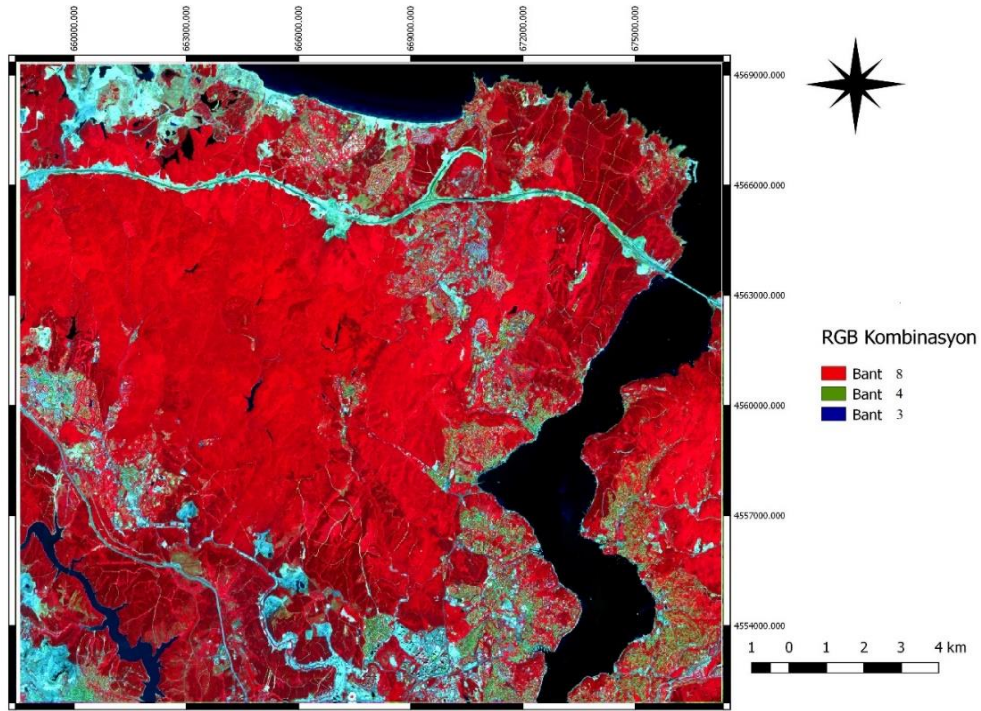
Vejetasyon analizleri için sıklıkla kullanılan 11, 8a, 4 kombinasyonu kullanıcıya büyük bir yelpazede yansıma ve renk karşıtlığı değerleri sağlar. Bu kombinasyonda sağlıklı vejetasyon parlak yeşil ve toprak leylak rengidir. Vejetasyon analizleri arasında en çok kullanıldığı alanlar kereste üretimi ve böcek zararları ile mücadeledir (Şekil 3.7 /ESA, 2016).



Şekil 3.6: Sentinel 2 uydusuyla sağlıklı vejetasyonun belirlenmesi için kullanılan bant kombinasyonu. (Kırmızı Filtre: 8a, Yeşil Filtre: 11, Mavi Filtre: 2)



Şekil 3.7: Sentinel 2 uydusuyla vejetasyon analizleri için kullanılan bant kombinasyonu. (Kırmızı Filtre: 11, Yeşil Filtre: 8a, Mavi Filtre: 4)



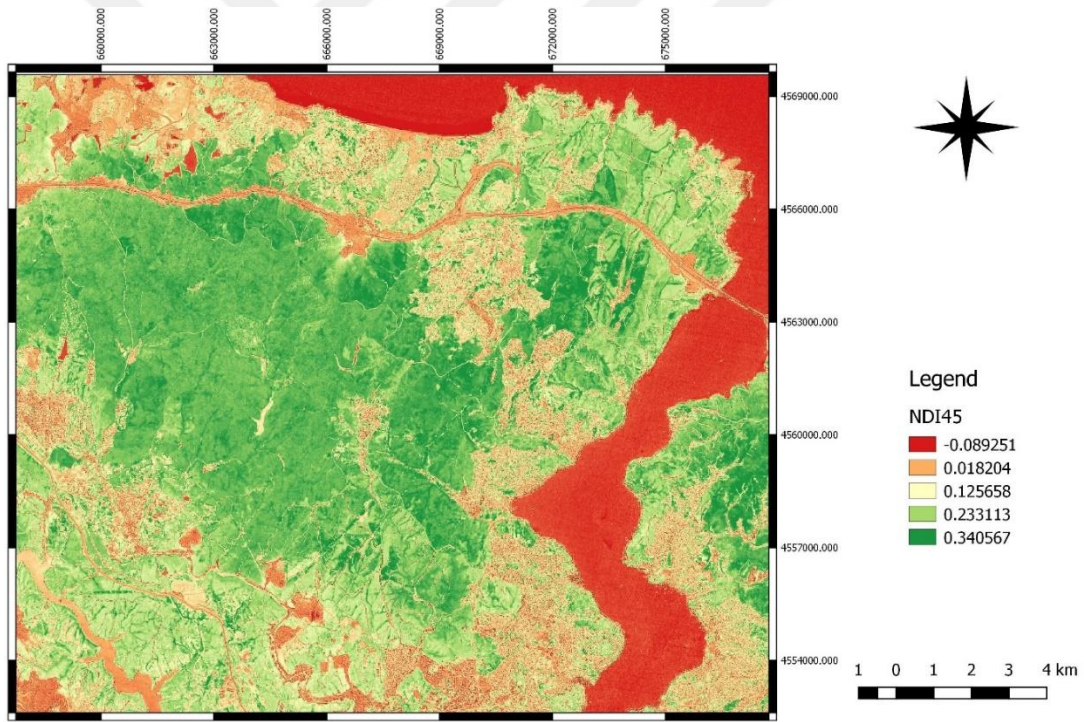
Şekil 3.8: Sentinel 2 uydusu ile elde edilen standart Kızılötesi bant kombinasyonu (Kırmızı Filtre: 8, Yeşil Filtre: 4, Mavi Filtre: 3).

Uzaktan algılamada vejetasyon görüntüleme ve su yüzeylerinin belirlenmesinde en sık kullanılan bant kombinasyonu olan standart Kızılötesi renklendirme, Sentinel 2 görüntülerinde 8, 4, 3 bantları ile gerçekleştirilir. Vejetasyon kırmızısının tonları, yerleşim alanları siyan mavisi ve toprak açık/koyu kahverengi tonları arasında değişim gösterir. Buz, kar ve bulutlar beyaz ve açık siyan renkleri arasındadır. İbrelili türler öteki odunsu türlere göre daha koyu bir kırmızıdır. Bu kombinasyon yukarıda bahsedildiği gibi vejetasyon çalışmaları, drenaj ve toprak yapısı görüntüleme, ayrıca zirai ürünlerin farklı büyüme evrelerindeki değişimi gözlemlemek için kullanılır. Sentinel 2 uydusundan elde edilen bu kombinasyon geleneksel kızılötesi hava fotoğrafçılığıyla elde edilen kompozitlerle benzer özelliklere sahiptir (Şekil 3.8 / ESA, 2016).

Yukarıda bahsedilen bant kombinasyonları dışında farklı ilgi alanları doğrultusunda (örneğin şehir planlama, risk analizi, dağ ormancılığı vb.) çeşitli bant kombinasyonları da mevcuttur. Bununla birlikte uzaktan algılamada kullanılan bir başka görüntüleme ve bilgi elde etme yöntemi *Vejetasyon İndeksleri* oluşturmaktır.

VI'lerine dayanan yöntemler genellikle basit hesaplanıp, kolay uygulanabilir olmaları nedeniyle sıklıkla tercih edilmektedir. Temelde VI ile yapılan uygulama; vejetasyonun spektral yansıma özelliklerini karşılaştırarak bitki örtüsünün biyofiziksel parametrelerini niceliksel olarak elde etmektir. Uydu görüntülerinden VI'lerinin elde edilmesi, verilerin geniş multispektral kapsama alanlarını değiştirmeden vejetasyonun biyofiziksel parametrelerinin elde edildiği en iyi yöntemlerden bir tanesi olarak kabul edilir. Bu nedenle vejetasyon indekslerinin kullanımı ve geliştirilmesi uzaktan algılama uygulamaları için büyük önem taşır (Frampton W. J. ve diğ. 2013).

Burada elde edilen çıktı tek bir banttır ve içerisinde belirli formüllerle birleştirilen bantların istenilen özelliklerini barındırır.



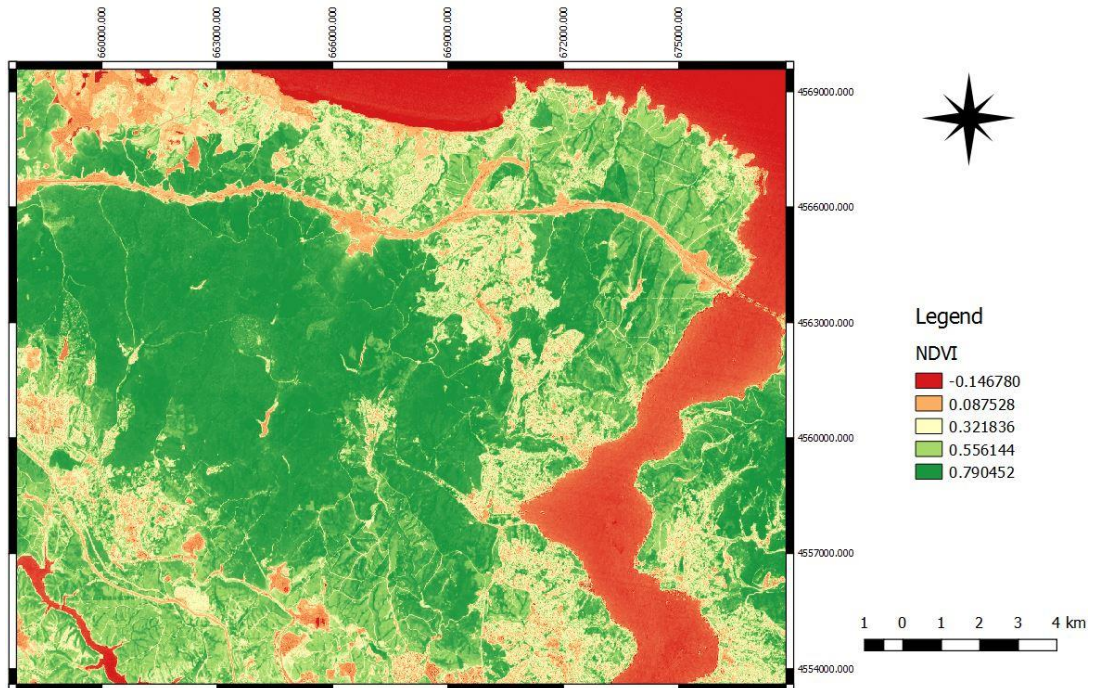
Şekil 3.9: Sentinel 2 uydusunun 4 ve 5. bantları ile oluşturulan NDVI45 vejetasyon indeksi ve spektral değerleri. (Bant Formülü: $(RE1 - R) / (RE1 + R)$)

Willian James Frampton ve arkadaşlarının 2013 yılında Sentinel 2a uydusundan elde edilen *Yaprak Yüzey Alanı (LAI)* ve *Yaprak Klorofil Konsantrasyonu (LCC)* gibi vejetasyon analizleri için önem arz eden vejetasyon değerlerine yönelik yaptıkları

çalışmada öne çıkan 4 tane vejetasyon indeksi olmuştur, bunlar: NDVI, NDI45, PSSR ve IRECI'dir.

Bu noktada LAI ve LCC üzerinde durulması gerekir. LCC değeri meşçerenin sağlığı ve potansiyel büyüme verimliliğini ortaya koymada (Gitelson ve diğ. 2006), LAI ise meşçerenin fonksiyonu ve yapısını belirlemede kullanılır (Wilhelm ve diğ. 2000). Yapılan çalışma doğrultusunda vejetasyon analizleri ön planda olduğu için kullanılacak vejetasyon indeksleri LAI ve LCC'nin en yüksek doğrulukta elde edildiği bu dört VI'ı olacaktır.

NDI45 vejetasyon indeksi (Şekil 3.9) Delegido J. ve arkadaşları tarafından 2011 yılında LCC değerinin yüksek çözünürlüklü uydu görüntülerinden temini için oluşturulmuştur. Bu indeks'te NDVI için kullanılan kırmızı bant 7. bant olan NIR yerine 5. bant RE1 kullanılmıştır, kırmızı bant yine 4. bant olarak sabit kalmıştır. Bu bantın özelliği, yüksek yansıma değerlerindeki "doygunluğun" NDVI'a göre daha yüksek olmasıdır.



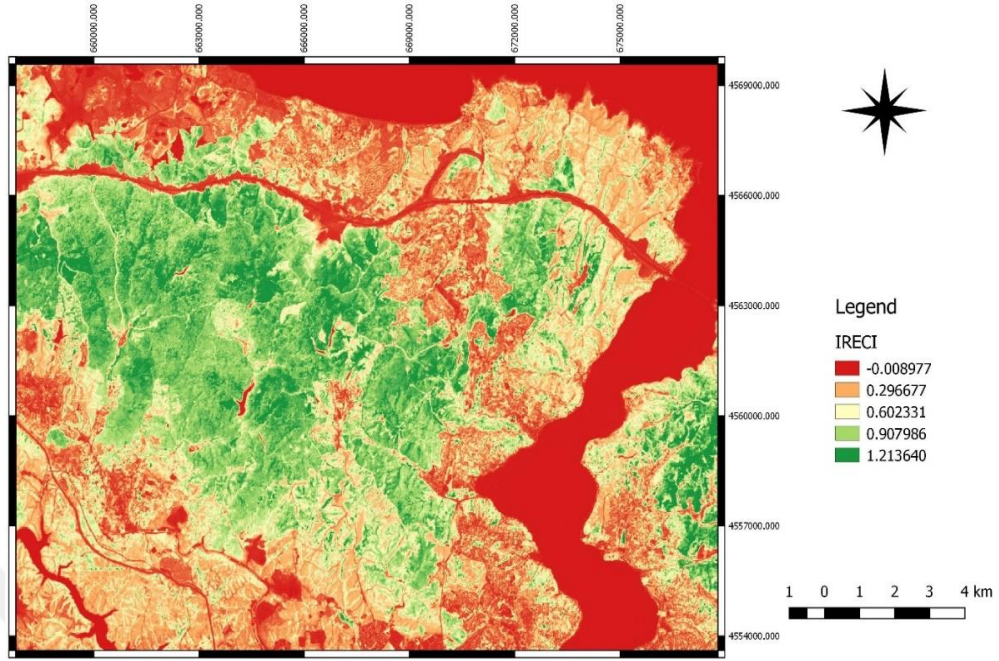
Şekil 3.10: Sentinel 2 uydusunun 4 ve 7. Bantları ile oluşturulan NDVI vejetasyon indeksi ve spektral değerleri. (**Bant Formülü:** $(NIR(Bant\ 7) - R(Bant\ 4)) / (NIR(Bant\ 7) + R(Bant\ 4))$)

NDVI (Normalized Difference Vegetation Index, Şekil 3.10) günümüzde vejetasyon analizleri ve diğer birçok araştırma konusu için en sık kullanılan VI'lerinin başında gelir ve ilk olarak Rouse ve arkadaşları tarafından 1973 yılında kullanılmıştır. Vejetasyon analizleri sırasında kullanılan LAI ve LCC'nin elde edilmesi dışında, yersel ölçümlerle birleştirilerek bitki verimliliği, parçalı meşçere yapısı, kapalılık, su toplama kapasitesi gibi ormancılık açısından kritik olan değerlerin elde edilmesinde de kullanılmaktadır (Myneni R. B. Ve diğ. 1995).

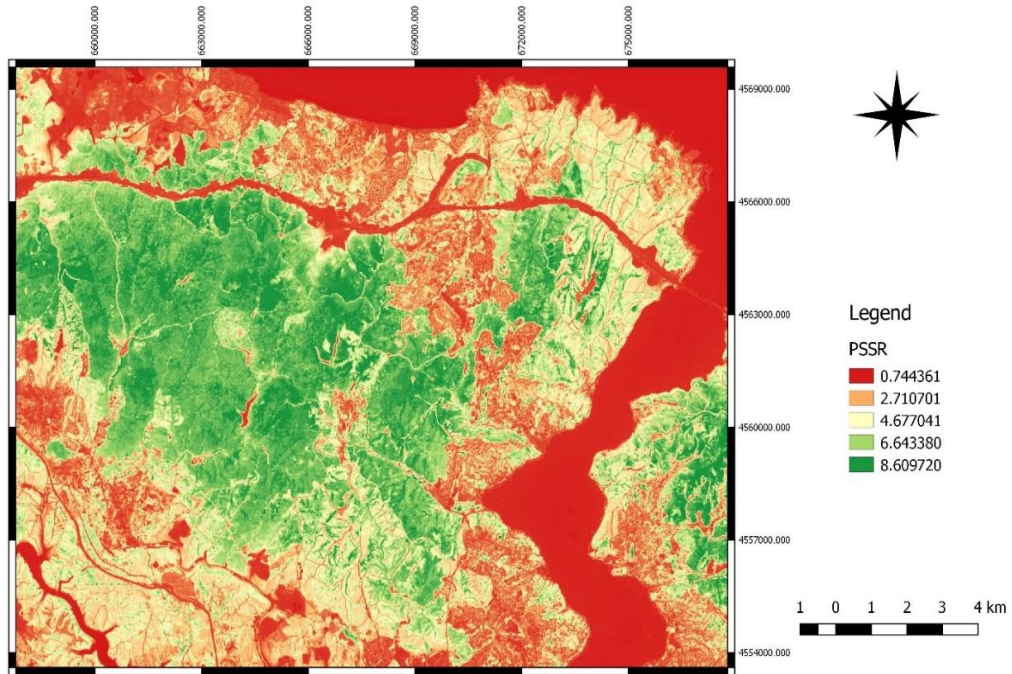
IRECI (Inverted Red Edge Chlorophyll Index, Şekil 3.11) indeksi de NDI45 ve NDVI ile birlikte *spektrometre* ile elde edilen data setleriyle birleştirilip LCC ve LAI değerlerinin elde edilmesi için kullanılan ve bu indekslerden LCC ile en yüksek korelasyon gösteren S2 vejetasyon indeksidir (William J. F. ve diğ. 2013).

Jordan ve arkadaşlarının 1969 yılında yaptıkları araştırma kapsamında ortaya konulan SR (Simple Ratio) kavramı LAI ve LCC değerlerinin elde edilmesi için Yakın Kızılötesi (NIR) bandın Kırmızı (R) bandına oranı ile elde edilen değeri belirtir. Bu değer Rouse ve arkadaşlarının 1973 senesinde yaptıkları çalışmada NDVI'nin temeli olarak belirtilmektedir. Blackburn 1998 yılında yaptığı çalışmada SR için optimum olan dalga boylarını bulmaya çalışmış ve sonuç olarak PSSR (Pigment Specific Simple Ratio) adını verdiği vejetasyon indeksini ortaya çıkarmıştır (Şekil 3.12).

Çalışma kapsamında VI'lerinin kullanımı, Sentinel-2a uydusunun sahip olduğu 10 ve 20 metrelik bantlara “görüntü ön işleme” aşamasında eklenerek 14 bantlı bir multispektral görüntü elde etme yoluyla gerçekleşmiştir. VI'lerini orijinal bantlardan ayıran özellik, içerdiği spektral değerlerin kullanım alanı doğrultusunda (toprak, su, vejetasyon vb. gibi) ayrılabilirliği çok iyi sağlaması ve bu yönüyle sınıflandırmanın doğruluğunu arttırmasıdır.

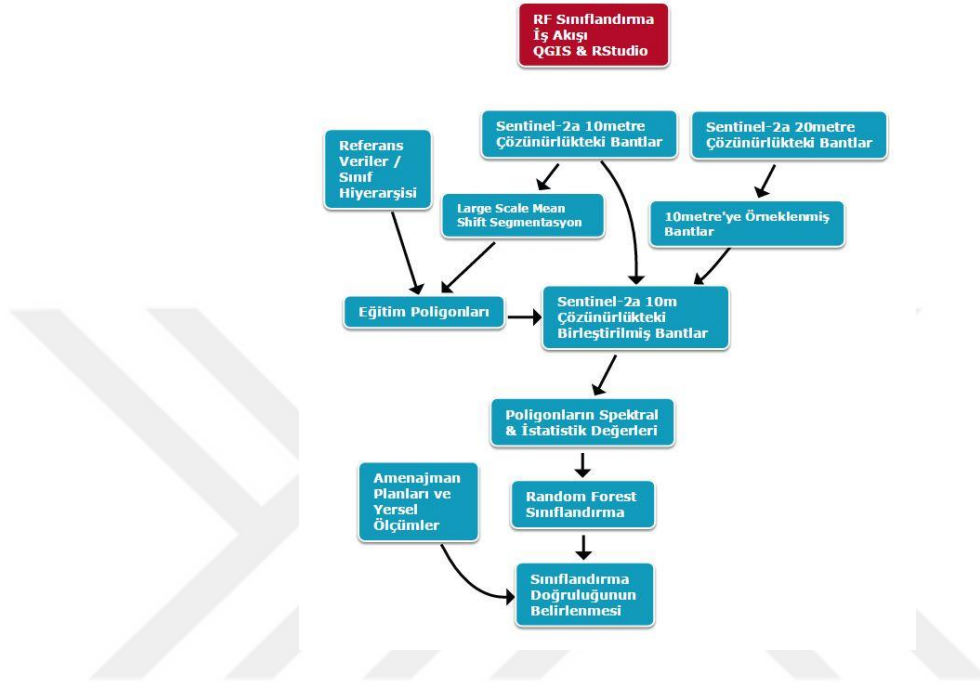


Şekil 3.11: Sentinel 2 uydusunun 4, 5, 6 ve 7. Bantları ile oluşturulan IRECI vejetasyon indeksi ve spektral değerleri. (Bant Formülü: $\text{NIR}(\text{Bant } 7) - \text{R}(\text{Bant } 4) / (\text{RE1} / \text{RE2})$)

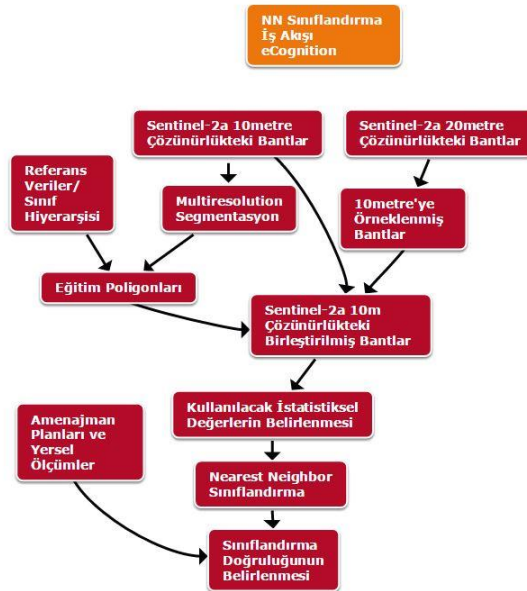


Şekil 3.12: Sentinel 2 uydusunun 7. bandının 4. banda oranlanmasıyla elde edilen PSSR vejetasyon indeksi ve spektral değerleri. (Bant Formülü: $\text{NIR}(\text{Bant } 7) / \text{R}(\text{Bant } 4)$)

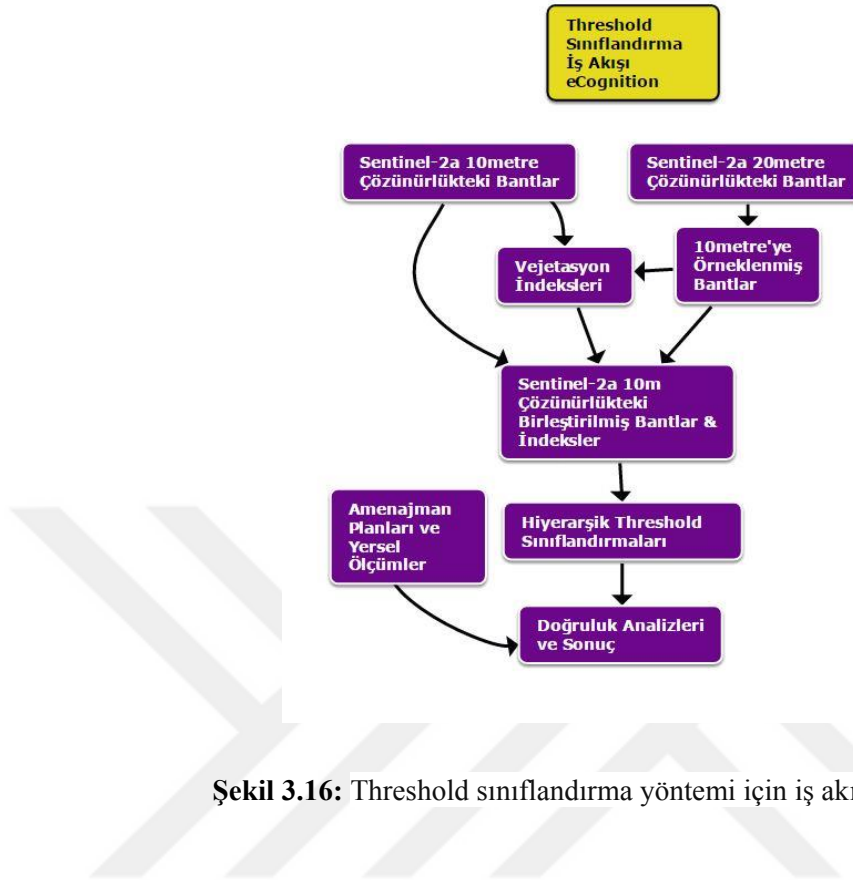
hangi amaçla gerçekleştirildiği konusunda bilgiler de ilgili başlıklar altında yer almaktadır. RStudio yazılımı içerisinde kullanılan makrolar (script) ise "Ek" başlığında verilmiştir.



Şekil 3.14: Random Forest sınıflandırma yöntemi için iş akış diyagramı.



Şekil 3.15: Nearest Neighbor sınıflandırma yöntemi için iş akış diyagramı.



Şekil 3.16: Threshold sınıflandırma yöntemi için iş akış diyagramı.

3.2.1. Görüntü Ön İşleme

Görüntülerin elde edilmesinden kullanımına kadar takip edilmesi gereken bir dizi işlemler vardır. Terminolojide bunlar atmosferik, topoğrafik düzeltme, bulut tespiti ve yok edilmesi vb. olarak adlandırılır. Bu işlemler kullanılan görüntü ve alanın topoğrafyasına göre aşamalı olarak gerçekleştirilir. Sınıflandırmanın doğru sonuçlar vermesi bu işlemlerin yapılmasıyla doğrudan ilişkilidir. Bu çalışma kapsamında yapılacak olan işlem, çalışma alanı topoğrafyasının nispeten düz olması ve yükselti farklarının kayda değer olmaması nedeniyle yalnızca atmosferik düzeltme olacaktır.

Optik cihazlarla ve uzaydan gerçekleştirilen uzaktan algılama işlemlerinde optik düzlemin yansıma aralıkları yeryüzünün incelenmesi ve değişikliklerin gözlemlenmesi için oldukça güçlü bir araçtır. Ancak, bu işlemde optik cihazların işlevleri atmosfer tarafından büyük oranlarda etkilenmektedir; bulutlar, gaz molekülleri ve aerosoller güneşten gelen ve ya dünyadan yansıyan ışınların hem kırılmasına hem de absorbe edil-

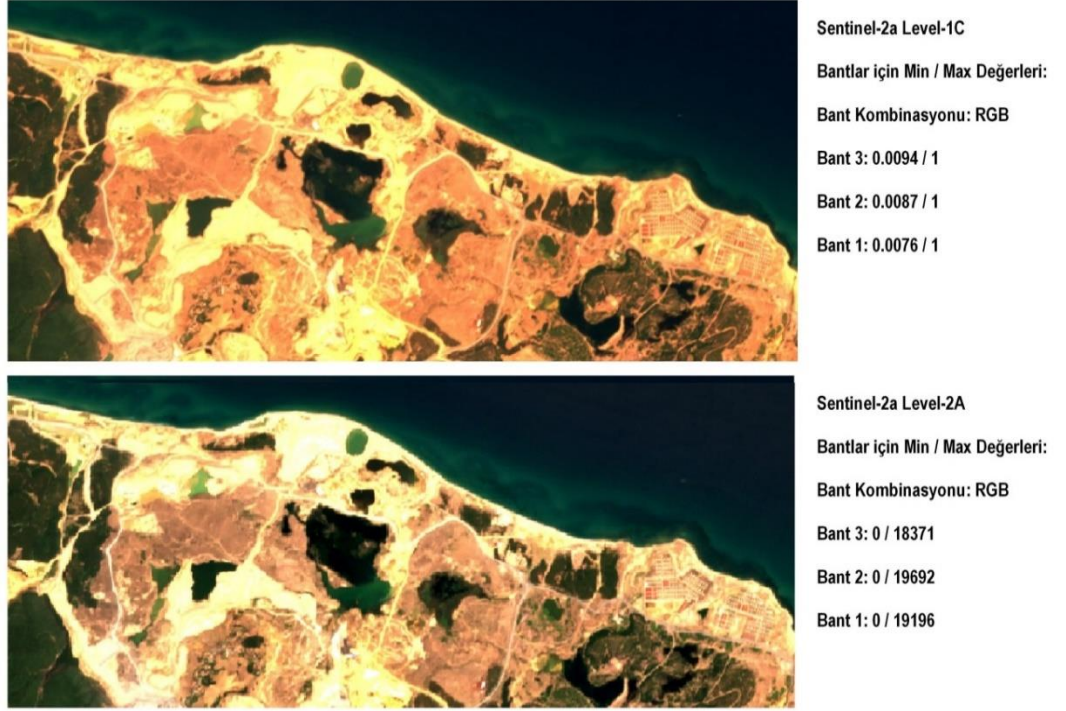
-mesine neden olur. Bu nedenle uydular tarafından elde edilen ham görüntülerin kullanılmadan önce bu gibi etkilerin giderilmesi gerekmektedir (Hagolle ve diğ. 2014).

Atmosferik düzeltmelerin yapılması aşamasında bilinmesi gereken bazı değerler bulunmaktadır, bunlar; AOT (Aerosol Optical Thickness / Aerosol Tabakası Kalınlığı), TOA (Top-of-Atmosphere / Atmosfer Üstü) yansımadır. Burada AOT değerleri aerosol'ün etkilerini ve kalınlığını belirtir, bu değerlerin *ışınımsal iletim kodları* ile işlenmesi ile uyduların TOA (atmosfer üstü) yansıma değerleri BOA (*yüzey yansıma değerleri*)'ya dönüştürülür.

Bu noktada bilinmesi, tahmin edilmesi gereken en önemli değer AOT değeridir. Bunun birçok yöntemi bulunmasına karşın Sentinel 2 uydularıyla bu değer multi-temporal (MT / çok zamanlı) ve multi-spektral (MS) metotlarıyla elde edilebilir. Sentinel 2 uydularının yeryüzünü aynı açı ile görüntülemesi ve 6 gün gibi kısa bir süre içerisinde aynı noktaya tekrar gelmesi aynı görüntüyü 6 gün farkla tekrar elde etmesini sağlar. Bu şekilde aynı alanın 2 farklı görüntüsündeki spektral değerlerin değişimleri hesaplanarak aerosol tabakasının kalınlığı belirlenebilir, bu yöntem multi-temporal yöntem denilir (Hagolle ve diğ. 2008). Atmosferik düzeltme işlemleri için kullanılan bir diğer metot ise MS metotlar olarak geçer. Bu yöntemde AOT değeri farklı bantların spektral değerleri arasındaki ampirik ilişkiden elde edilmektedir; kullanılan bantlar genellikle kısa dalga kıızılötesi (SWIR) ve yakın kıızılötesi (NIR) bantlardır (Guanter ve diğ. 2008).

Sentinel 2 uydularının atmosferik düzeltmeleri ESA'nın geliştirdiği Sen2Cor yöntemiyle gerçekleştirilmektedir. Burada işlenmemiş ham veriler Level-1C, işlenmiş görüntüler ise Level-2A olarak isimlendirilir. Burada Level-1C ürünler orto-rektifiye edilmiş TOA yansıma değerine sahip çok bantlı ve birden çok tarih damgalı verileri belirtirken: Level-2A ürünleri TOA değerleri BOA değerlerine çevrilmiş, ek olarak AOT haritası, Su buharı haritası çıktılarına sahip verileri belirtir. Sen2Cor işlemi sırasında ayrıca 60 ve 20 metrelik bantlar isteğe bağlı olarak 10 metrelik 2. 3. 4. ve 8. bantlara örneklenir (ESA, 2016).

Bu çalışmada atmosferik düzeltme için Sen2Cor görüntü işleyicisi kullanılmış, aynı zamanda 20 metrelik bantlar 10 metreye örneklenip birleştirilerek 10 bantlı uydu görüntüsü elde edilmiştir. 60 metre çözünürlüklü bantlar sınıflandırma aşamasında kullanışsız olduklarından işleme alınmamıştır (Bu bantlar termal ve atmosferik parametrelerin elde edilmesi için gereklidirler).



Şekil 3.17: Sen2Cor yöntemi ile atmosferik düzeltmenin gerçekleştirilmesi. İşlenmemiş Sentine 2a Level-1C görüntüsü ve bant değerleri (üstte). Atmosferik düzeltmesi gerçekleştirilmiş Level-2A nitelikli uydu görüntüsü ve bant değerleri (altta).

Görüntülerin atmosferik düzeltmesi ve örneklenme işlemi Sen2Cor görüntü işleyicisi tarafından verilen parametreler doğrultusunda otomatik olarak gerçekleşmektedir (görüntü hakkında gereken bilgiler; görüntünün alındığı tarih, güneşin açısı vb. görüntünün *metadata*'sından temin edilir). Kullanılacak olan görüntünün 10 metre çözünürlükte olan bantlardan oluşması gerektiğinden 13 bantlı görüntüdeki 20 metre çözünürlüğündeki bantların tamamının (5, 6, 7, 8a, 11, 12. bantlar) 10 metreye örneklenmesi gerekmektedir. Örnekleme işlemi belirli bir piksel boyutuna sahip

görüntünün piksel boyutu büyütülüp (downsample) ya da küçültülüp (upsample) görüntünün kabalaştırılması ya da keskinleştirilmesi yöntemidir.

Görüntü örnekleme işleminde birçok farklı metot vardır, bunlardan en yaygın olanları; En yakın komşu (Nearest Neighbor), Bilineer Enterpolasyon (Bilinear Interpolation), Kübik Katlama (Cubic Convolution), Kübik Eğrisi (Cubic Spline), Kübik B-Eğrisi (Cubic B-Spline) vs. olarak isimlendirilir. Sentinel-2a uydusunun görüntü işleme yazılımı olan SNAP bu yöntemlerden kübik eğrisini (Cubic Spline) kullanmaktadır (ESA, 2016).

Bu işlem ile birlikte tüm bantların atmosferik düzeltmeleri yapılmış ve 20 metre çözünürlüğe sahip bantlar 10 metreye örneklenmiştir. Bu aşamadan sonra görüntü kullanıma hazırdır, çalışma alanı doğrultusunda görüntü kesilip (Şekil 3.13) sınıflandırma için gereken segmentasyon işlemi gerçekleştirilebilir.

3.2.2. Segmentasyon

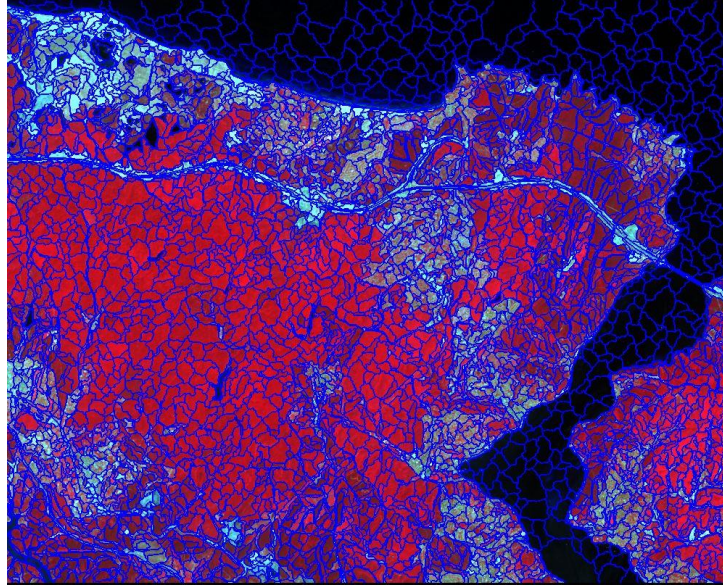
Daha önceki kısımlarda bahsedildiği üzere obje tabanlı sınıflandırma yöntemlerinde, piksel tabanlı sınıflandırma yöntemlerinde kullanılan temel birim olan pikselin yerine objeler kullanılmaktadır. Obje, belirlenen parametreler doğrultusunda (yansıma değerleri, tekstür, mekansal değerler vb.) benzer pikselleri bir araya getirip onlardan oluşturulan kümelere verilen isimdir. İyi bir segmentasyon işlemi sonuç olarak farklı tekstüre sahip homojen objelerden oluşan bir yapı oluşturmalıdır. Segmentasyon işlemi için geliştirilmiş birçok algoritma bulunmaktadır; bu çalışma içeriğinde kullanılan QGIS ve eCognition yazılımları segmentasyon işlemleri için Large Scale Meanshift Segmentation (LSMS) ve Multiresolution Segmentation (MRS) algoritmalarını kullanmaktadır.

MRS algoritmasında kullanılan temel yöntem alan birleştiricisi olarak tanımlanabilir. Bu yöntemde objeler piksellerin birbirleri ile birleşmesi ile oluşmaya başlar. Her bir birleşme aşamasında piksellerden oluşan objeler, alansal homojenlik kriterini sağlayana kadar birbirleri ile birleştirilmeye devam eder. Bir pikselin komşu objeye katılması için gereken koşullar “uygun” ya da “uygun değil” olarak gerçekleşmez. Burada geçerli olan

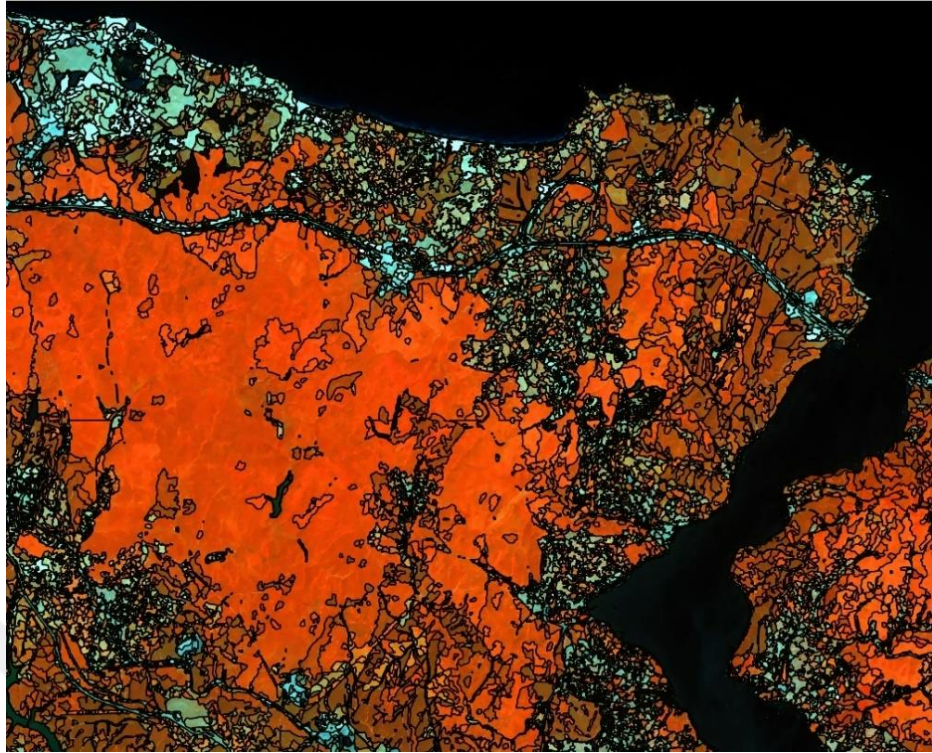
kural “birleşme koşulu” denilen bir derecelendirme sistemidir. Her bir olası birleşme için “birleşme kuralı” kendisini değiştirir. Birleşme kuralının oluşması için segmentasyon öncesi verilen parametreler kullanılır; bu parametreler doğrultusunda oluşan birleşme kuralının (eşiğin) altına düşüldüğünde segmentasyon işlemi tamamlanmış olur. Bu noktadan sonra birleştirilebilecek objeler kalmamıştır (Baatz ve Schape, 2015).

Birleşme kuralının düşük değerlere sahip olması (girilen parametre doğrultusunda) oluşan objelerin boyutunun küçük olmasına neden olur, yani birleşme kuralı ile obje boyutu arasında doğru orantı bulunmaktadır. Birleşme kuralının oluşması için kullanılan parametre ölçek parametresidir (Scale Parameter). Ayrıca homojenliği sağlamak için kullanılan renk kombinasyonları için spektral değerler ve ayrıca obje şekli için pürüzsüzlük (smoothness) ile yoğunluk (compactness) kullanılır.

LSMS algoritmasında da benzer bir yöntem kullanılmaktadır. Bu algoritmada istenilen parametreler “alansal mesafe (spatial radius)”, “spektral mesafe (range radius)” ve “minimum obje boyutudur (minimum object size)”. Burada minimum obje boyutu birleştirme koşulunu; alansal ve spektral mesafeler ise homojenliği sağlamak için kullanılır.



Şekil 3.18: Multiresolution Segmentation algoritması ile gerçekleştirilen segmentasyon işlemi ve kullanılan parametreler (Ölçek Parametresi: 20, Pürüzsüzlük: 0.5, Yoğunluk: 0.5).



Şekil 3.19: Large Scale Meanshift Segmentation algoritması ile gerçekleştirilen segmentasyon işlemi ve kullanılan parametreler (Alansal Mesafe: 10, Spektral Mesafe: 10, Min. Alan Boyutu: 10).

Segmentasyon işlemi sonucunda elde edilen sonuçlar incelendiğinde her iki segmentasyonun da objeleri homojen olarak birbirinden ayırabildiği ancak LSMS ile elde edilen obje boyutlarının minimum obje boyutunun (birleştirme kuralı) üzerinde çıktığı görülmüştür. 50 piksel olması beklenen obje boyutlarının bunun çok üzerinde çıktığı görülmüş, homojen olmasına rağmen büyük objelerin sınıflandırma aşamasında problemler çıkarabileceği düşünüldükçe her iki sınıflandırma yöntemi için de eCognition ile üretilen MRS kullanılmıştır (Şekil 3.18). Çalışma alanı içerisinde elde edilen toplam obje (segment) sayısı 12533'dür.

3.2.3. Obje Tabanlı Sınıflandırma

Çalışmanın bir sonraki aşamasındaki işlem, daha önce belirlenen obje tabanlı sınıflandırma yöntemlerinin her bir segmentin içerisinde bulunan özelliklere göre kullanılmasıdır. Belirlenen üç farklı sınıflandırma yöntemi ile görüntülerin sınıflandırma işlemi sırasıyla; R Studio ve QGIS yazılımları kullanılarak Random Forest

ve eCognition yazılımı kullanılarak Nearest Neighbor ile Threshold Sınıflandırma yöntemleri olacaktır.

3.2.3.1. Random Forest Sınıflandırma Yöntemi

Görüntü sınıflandırma işlemi, sayısallaştırılmış görüntüdeki sayısal değerlerin (DN) arazi kullanım bilgileri barındırması için işlenmesi olarak tanımlanır. Bir diğer deyişle, görüntü sınıflandırma işlemiyle her bir piksel içerdiği spektral değer ya da mekânsal özelliklerine göre önceden tanımlanmış sınıflardan birisine atanır. Spektral düzlemde gerçekleşen sınıflandırma işleminde karar kuralları kullanılmaktadır, bu noktada piksellerin değerleri; mekânsal özelliklerle gerçekleşen sınıflandırma da ise yine karar kuralları geçerlidir ancak buradaki değerler komşuluk ilişkilerine bağlı olarak gerçekleşir (Şekil, tekstür ve doku vb. gibi) (Gao, 2009).

Objeye tabanlı yaklaşımların piksel tabanlı yaklaşımlarla olan benzerliklerinden daha önceki başlıklarda bahsetmiştik. Temelinde bütün objeye tabanlı sınıflandırma yöntemleri piksel tabanlı yaklaşımları kullanmaktadır, sonucuna buradan ulaşabiliriz.

Birleşik Öğrenme Metotları (ELM) tanımından genel bilgiler başlığında söz edilmişti, bu kavram RF algoritmasının temelini oluşturmaktadır. ELM sınıflandırıcılar tek bir sınıflandırıcı oluşturmak yerine birçok sınıflandırıcı üretip her bir sınıflandırıcının derecelendirmesine göre sınıf atamalarını gerçekleştirir. Bu sınıflandırıcılar da kendi içinde 3'e ayrılır; Bagging, Boosting ve Random Forest (Akar, 2013). Bagging ve Boosting bu araştırmanın konusu değildir.

RF algoritmasının çalışması için gereken iki parametre bulunmaktadır: $\langle\langle N \rangle\rangle$ ve $\langle\langle m \rangle\rangle$. Bunlar oluşturulacak ağaç sayısı ve ağaçların her bir karar düğümü noktasında yapacağı ayrılma sayısını belirtmektedir. Bu çalışmada RF için kullanılacak parametreler için Immitzer, Vuolo ve Atzberger'in 2016 tarihli çalışmaları dikkate alınarak $N = 500$ olarak belirlenmiş, m 'in ise toplam örnek alan sayısının karekökü olarak alınmasına karar verilmiştir.

Görüntü ön işleme, örnekleme ve segmentasyon işlemleri sonrasında elde edilen 10 bantlı görüntüden verilerin elde edilmesi ve bu bantların aynı ağırlıkta değere sahip olması için bütün bantların yansıma değerleri 0 ile 255 arasında esnetilmesi (stretch) gerekmektedir. Bu işlem için R Studio yazılımındaki “raster” paketi kullanılmıştır. Bu paket ile çok büyük alanlara sahip çok bantlı görüntülerin coğrafi ve spektral özellikleri hızlı bir şekilde işlenebilir (CRAN, 2016).

Her bir bandın aynı ağırlıkta öneme sahip olduğu noktada bu bantlardan sınıflandırma aşamasında kullanılacak istatistiki değerlerin elde edilmesi işlemine geçilir. Bu çalışma içerisinde görüntülerden elde edilmesi kararlaştırılan değerler; ortalama (mean), standart sapma (standart deviation /sd) ve minimum/maksimum ‘dur. Bu aşamada segmentasyon işlemi ile elde edilen her bir objenin içerisindeki bu istatistiki değerlerin ortalaması elde edilir, yani her bir obje bir pikselmiş gibi işleme girer.

Segmentasyon ile oluşturulan 12533 objenin istatistiki değerleri her bir bant için farklıdır, bu durumda elde edilen değerler her bir bant için 4 sütundan oluşur; böylece toplamda 40 sütundan oluşan bir tablo elde edilmiş olur. Bu tabloların düzenlenmesi ve görüntülenebilmesi için R Studio yazılımının “matrixStats” paketi kullanılmıştır (Tablo 3.2).

Tablo 3.2: R Studio “matrixStats” paketi ile bantların spektral değerlerinin görüntülenmesi.

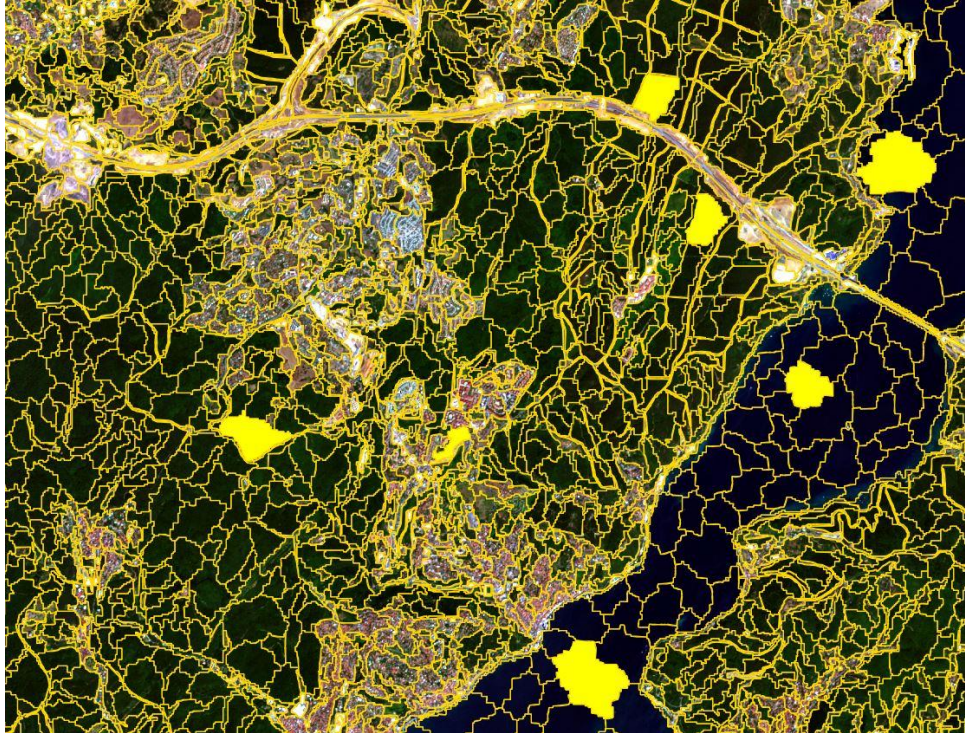
	zone	mean.red	mean.green	mean.blue	mean.nir	mean.ndvi	mean.ireci	mean.pssr	mean.ndi45	sd.red	sd.green	sd.blue	sd.nir
1	0	12.537449	22.153569	20.3669157	1.038789e+02	194.40365	63.375582	85.548428	150.34584	6.9551330	7.3645396	10.1206068	19.18549694
2	1	34.390658	47.046238	49.7433425	1.110171e+02	159.35021	42.383544	52.859570	126.01367	16.3410200	16.5548907	21.5368742	24.59331137
3	2	9.167148	17.995701	13.0573727	1.318483e+02	218.77348	104.150210	124.574942	160.79759	3.5315766	4.7763919	5.7815477	21.36192391
4	3	48.542114	58.210611	67.9022987	1.097737e+02	141.89651	35.892761	41.786180	104.06383	17.4552892	18.0514599	21.4321251	7.45438696
5	4	9.768548	20.496232	16.7146498	8.760438e+01	190.47165	51.125509	77.686446	156.33441	2.1323090	2.5321962	2.9171045	5.21620718
6	5	10.895530	20.929118	14.7433746	1.320711e+02	217.07289	97.645583	119.491178	165.35132	4.9955866	8.9316579	6.7041219	23.90331295
7	6	14.685334	24.176273	22.3160866	1.032937e+02	193.02879	61.577071	82.642734	148.36046	8.6035260	10.6387930	12.0111338	19.03571325
8	7	19.465857	26.567028	29.1985740	6.241732e+01	140.89435	28.509119	43.166427	108.07005	18.8152390	23.7002466	26.5716497	50.97241754
9	8	44.268288	63.480648	61.0755412	1.010297e+02	144.29264	31.927643	42.704860	126.33686	19.0146043	19.0308858	22.7466903	17.28713549
10	9	13.551666	24.000302	20.1255858	1.006382e+02	194.36795	60.867567	83.715053	154.81511	5.0785455	5.8050458	6.4152487	16.00913534
11	10	37.455579	50.482740	55.6755890	1.089395e+02	154.63970	39.374189	50.078257	127.60680	17.4419026	17.4794416	23.803237	20.54469932
12	11	11.506331	19.651779	15.3952645	1.198276e+02	211.84404	86.367266	108.765200	155.77663	3.3277239	4.2261144	4.7596705	15.26905824
13	12	22.856556	35.763700	33.8697220	1.267002e+02	189.37959	70.864564	82.602006	145.88498	14.4744307	18.4112009	21.7911452	37.65093361
14	13	28.565656	42.447609	45.2682597	1.280287e+02	178.40520	62.385608	72.237460	144.58958	18.0584989	22.2001434	29.6822939	22.24669310
15	14	68.076242	103.139563	120.8568567	9.383424e+01	83.88887	12.230271	17.846411	83.00036	22.8091070	24.8249667	32.9039975	36.31207117
16	15	35.910517	55.358634	55.8520373	8.841916e+01	138.93730	29.175625	40.192826	125.57215	10.4885224	14.0722883	18.6428001	23.83454189
17	16	109.720820	154.329604	184.2488115	1.651519e+02	94.42131	19.717978	20.673361	74.24596	18.2516888	20.9527334	28.0148912	11.73191247
18	17	80.193669	122.759450	162.6513078	1.500756e+02	100.58684	21.269862	22.796627	87.19083	16.9483106	22.9749386	32.2689965	19.40780009

Her bir segmentasyonun *Tablo 3.2*'de gösterildiği gibi spektral değerlerinin elde edilmesinden sonra QGIS yazılımı ile bu poligonlar içinden örnek alanlar belirlenerek (su, yerleşim, ibreli vb.) RF algoritması için karar kurallarının oluşturulması aşamasına geçilebilir (*Şekil 3.20*). Bu aşamada her bir sınıf için belirlenen eğitim alanı sayısı 25 ile 40 arasında ve *Tablo 3.3*'de gösterildiği gibidir.

Tablo 3.3: RF için sınıf hiyerarşisi ve eğitim alanı sayıları.

1. Su	30
2. İbreli Türler	30
3. Yapraklı Türler	29
4. Zayıf / Bozuk Vejetasyon	31
5. Yerleşim 1	28
6. Yerleşim 2	27
7. Açık Toprak	33
8. Kumul	29
9. Yerleşim 3 (Sanayi)	33

Eğitim alanlarının toplam sayısı 270'dir; RF algoritması bu poligonlardan elde edilen istatistiki değerlerin (ortalama, min/maks, standart sapma) 2/3 sini karar kurallarını oluşturmada 1/3 ini ise bu karar kurallarının doğruluğunu ölçmede kullanacaktır.



Şekil 3.20: QGIS içerisinde eğitim alanlarının seçilmesi.

Son olarak eğitim alanlarından oluşturulan karar kuralları ile objelerin her biri değerlendirilip R Studio yazılımında bulunan “RandomForest” paketinin “predict (tahmin)” komutuyla tüm poligonların belirlenen bu 9 sınıf içerisine atanması gerçekleştirilir. Sonuç olarak elde edilen sınıflandırılmış görüntü ve doğruluk değerleri “bulgular” başlığında verilecektir.

3.2.3.2. Nearest Neighbor Sınıflandırma Yöntemi

Bu sınıflandırma yöntemi eCognition yazılımına özel olarak geliştirilmiş ve En Yakın Komşu algoritması temeline dayanmaktadır. Bu yazılım bir “black box software / kara kutu yazılım” olarak değerlendirilebilir, öyle ki yazılım içerisinde oluşturulan *Kural Dizinlerinin* (Rule Sets) hangi ağırlıklarla ve algoritmalarla birleştirilip sınıflandırma sonucunu ortaya çıkardıkları saklı tutulmaktadır (Beizer, 2001).

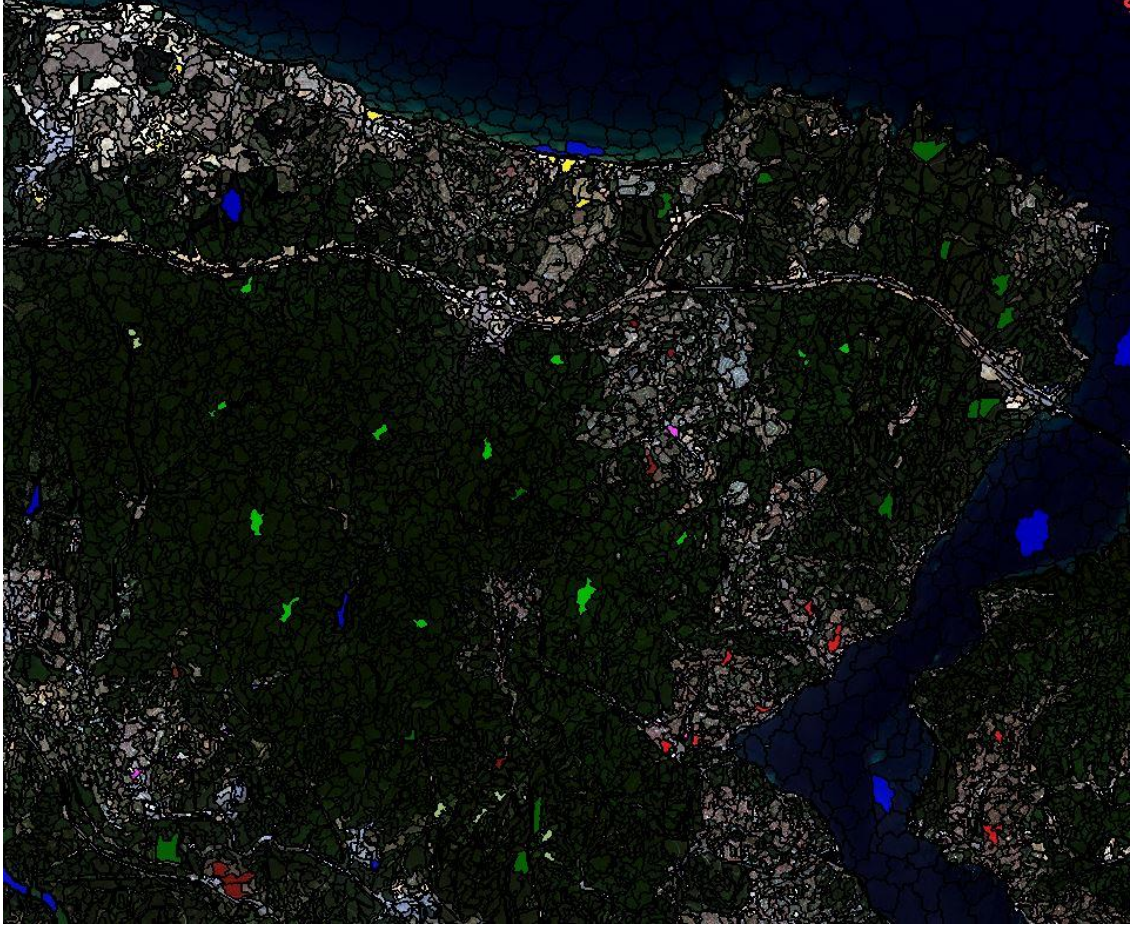
Bu sınıflandırma yönteminde de RF’da olduğu gibi segmentasyon işlemi sonucu oluşan objelerin (12533 adet) spektral değerleri dikkate alınmaktadır. Kullanılan 10 bandın her birinde objelerin sahip olduğu ortalama, standart sapma ve minimum/maksimum değerleri sınıflandırma aşamasında kullanılmıştır, elde edilen değerler RF için kullanılan değerler ile aynıdır (Tablo 3.2).

NN sınıflandırıcısı daha önce bahsedildiği üzere bir kontrollü sınıflandırma yöntemidir, çalışma alanının tamamına ait istatistiki değerler elde edildikten sonra bu objeler içinden hangi sınıfa ait olduğu bilinen eğitim alanlarının seçilmesi gerekmektedir (Şekil 3.21).

Tablo 3.4: NN için sınıf hiyerarşisi ve örnek alan sayıları .

1. Su	30
2. İbrelili Türler	30
3. Yapraklı Türler	29
4. Zayıf / Bozuk Vejetasyon	31
5. Yerleşim 1	28
6. Yerleşim 2	27
7. Açık Toprak	33
8. Kumul	29
9. Yerleşim 3 (Sanayi)	33

Seçilen örnek alanlar her bir sınıf için 25 ile 40 arasında ve belirlenen sınıf hiyerarşisi Tablo 3.4’deki gibidir.



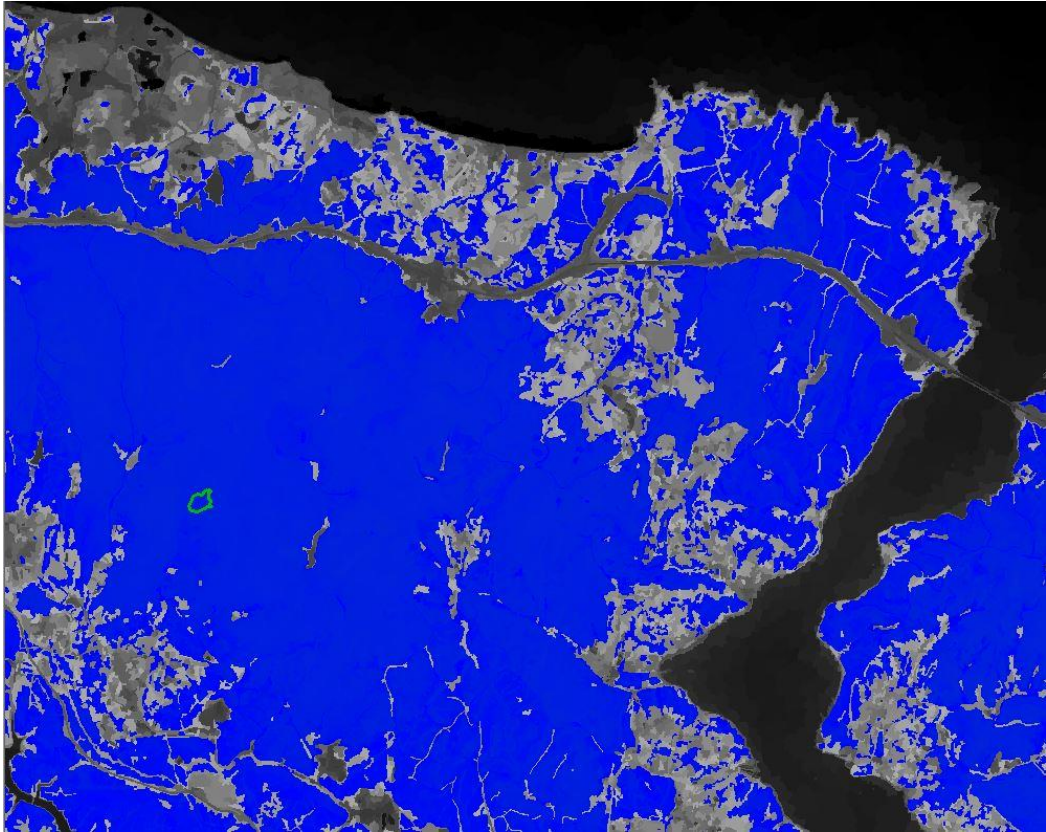
Şekil 3.21: eCognition ile eğitim alanlarının seçilmesi.

Sınıflandırma sonucu ve doğruluk analizi “bulgular” başlığında verilecektir.

3.2.3.3. Threshold Sınıflandırma Yöntemi

Bu sınıflandırma yöntemi de NN sınıflandırıcısı gibi eCognition yazılımıyla gerçekleştirilmektedir. Yöntemin isminden de anlaşılacağı üzere bu yöntemde sınıflandırma, belirlenen bir ya da birden fazla eşik değerlerin bir araya gelip, bantların sahip oldukları spektral değerlerin bu eşik değerlere göre birbirinden ayrılması şeklinde gerçekleşir. Belirlenen bu eşik değerler ile birbirinden ayrılan her bir spektral grup kendi içerisinde sınıfları oluşturur. Örnek ile açıklamak gerekirse, bir vejetasyon indeksi

olan NDVI'nin (Şekil 3.10) spektral yansıma değerleri -1 ile 1 arasında yer almaktadır ve değer -1'e yaklaştıkça sadece suyun bulunduğu yüzeyler görünür hale gelir, değer 1'e yaklaştıkça vejetasyon görünür duruma gelir. Yapraklı türlerin yansıma değeri 1'e yakın (örneğin 0.85) değerler alırken aynı alandaki ibreli türlerin yansıma değerleri 0.5'e yakındır. Bu durumda 3 farklı eşik değeri ile alandaki su, ibreli ve yapraklı türlerin sınıflandırılması sağlanabilmektedir.



Şekil 3.22: Belirlenen bir eşik değeri ile NDVI üzerinde vejetasyonun görüntülenmesi (Mavi alanlar 0.5 üzerindeki yansıma değerlerini belirtmektedir).

Bu sınıflandırma yönteminde belirli bir algoritma kullanılmadığı için eşik değerlerin belirlenmesi kullanıcıya bırakılmıştır. Yüksek doğrulukların elde edilmesi oluşturulan *Kural Dizinlerinin* doğruluğuyla ilişkilidir ve bu kurallar oluşturulduktan sonra son kullanıcının yalnızca eşik değerleri girmesi gerekir.

Threshold Sınıflandırma yönteminde eşik değerler tek bir bant üzerinden verilir, bu nedenle spektral ayrılabilirliğin yüksek olduğu vejetasyon indekslerinin kullanılması

(vejetasyon ve su yüzeyleri için NDVI, zirai alanlar için IRECI ve yerleşim alanları için NDBI vb. / ESA 2016) ve belirlenen kural dizinleri ile ağırlıklarının belirlenip birleştirilmesi gerekmektedir. Bu yöntem ile elde edilen sınıflandırılmış görüntü “bulgular” başlığında verilecektir.

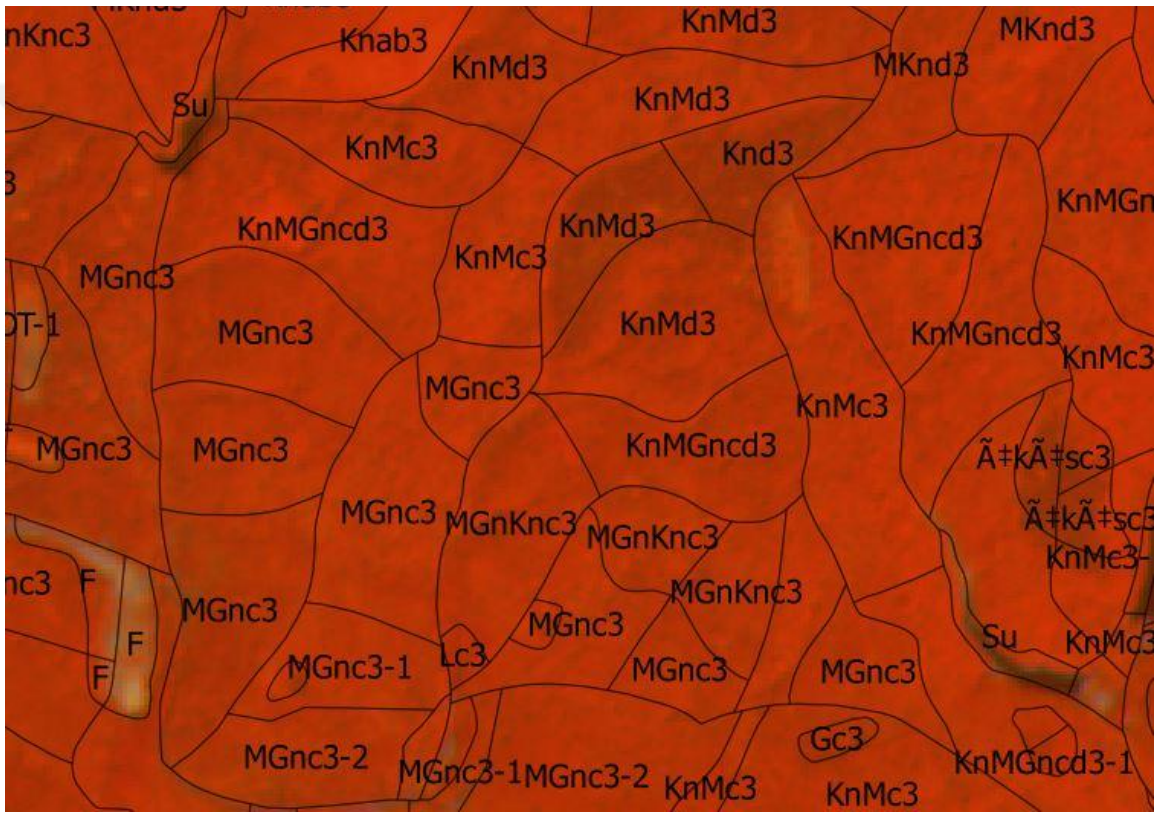
3.2.4. Doğruluk Analizi ve Sınıflandırma Doğrulukları

Gerçekleştirilen sınıflandırma işlemlerinden sonra elde edilen tematik haritaların pratikte ne kadar güvenilir olduğunun test edilmesi gerekmektedir, ancak bu sayede elde edilen haritalar üzerinden gerçekleştirilecek uygulamaların hangi hassasiyette sonuç vereceği öngörülebilir. Doğruluk analizi, uzaktan algılamanın temel konularından bir tanesidir, bunun nedeni yapılan ölçümlerde kaynağı ne olursa olsun her zaman bir hata payının olabileceğini düşünmemiz gerektiğidir.

Doğruluk değerlendirmeleri sınıflandırma aşamasından sonra gerçekleştirilir ancak obje tabanlı sınıflandırma yönteminde doğruluğu doğrudan etkileyen ve sınıflandırma aşamasından önce üzerinde durulması gereken en önemli nokta segmentasyon evresidir. Segmentasyon aşamasında kullanılan parametrelerin her bir görüntü için değişmesi ve kararın kullanıcı tarafından verilmesi bu sistemin bir kusuru olarak görülebilir. Bunun nedeni aynı algılayıcı tarafından alınıyor bile olsa bir görüntünün tekrar alınamayacak olmasından öte gelir; burada güneşin açısı, iklim, hava koşulları, sensörün açısı vb. gibi etkenler sebep olmaktadır. Bu nedenle aynı sınırlar görüntüleniyor bile olsa spektral değerler farklılık gösterecektir ve obje sınırları bu doğrultuda değişecektir. Segmentasyonun kullanıcı tarafından olabildiğince “doğru” gerçekleştirildiği ve sınıflandırmanın tamamlandığı noktada bu bölümün konusu olan “doğruluk” hakkında değerlendirmelerde bulunabiliriz.

Doğruluk analizinin yapılabilmesi için çalışılan alanın yer gerçekliklerine ihtiyaç vardır. Bu sayede oluşturulan harita ve yer gerçekliklerinin olduğu mekânsal veri üst üste bindirilerek ne derece doğru bir harita elde edildiği hesaplanabilir. ERDAS Imagine yazılımı bahsedilen bu iki girdi (tamamlanmış tematik harita ve yer gerçekleri) arasında örnek alanlar yoluyla bir bağlantı kurarak doğruluk analizini gerçekleştirir. Bu aşamada yazılımdan, sınıflandırmada kullanılan her bir sınıf için 50 adet rastgele nokta (referans

noktaları) belirlemesi istenir (Toplamda 400 adet). Bu sayının doğruluk belirlemek için minimum değeri olduğu belirtilmiştir (Yener H., Koç A. ve Çoban O., 2006). Ancak bu çalışma içerisinde kullanılan sınıf sayısının 9 adet olması nedeniyle referans noktaların her bir sınıf için 30'ar tane alınmasına karar verilmiştir. Doğruluk analizi için kullanılan yer gerçekleri OGM'nin 2013 yılı amenajman planlarından elde edilmiştir, WV2 uydu görüntüsü ve Google Earth haritaları da bu verileri doğrulamak için eş zamanlı olarak kullanılmıştır.



Şekil 3.23: Doğruluk analizi için kullanılan uydu görüntüsü üzerine bindirilmiş meşçere tipi hariri taşı ve arazi kullanım sınıfları.

Bu çalışmada verilecek olan doğruluk değerleri Kullanıcı Doğruluğu (User's Accuracy), Üretici Doğruluğu (Producer's Accuracy), Toplam Doğruluk (Total Accuracy) ve Kappadır.



Şekil 3.24: ERDAS Imagine ile doğruluk analizinde kullanılmak üzere rastgele nokta atılması.

Her bir sınıflandırma yöntemi için doğruluk değerleri ve sınıflandırılmış görüntüler bulgular başlığında verilecektir.

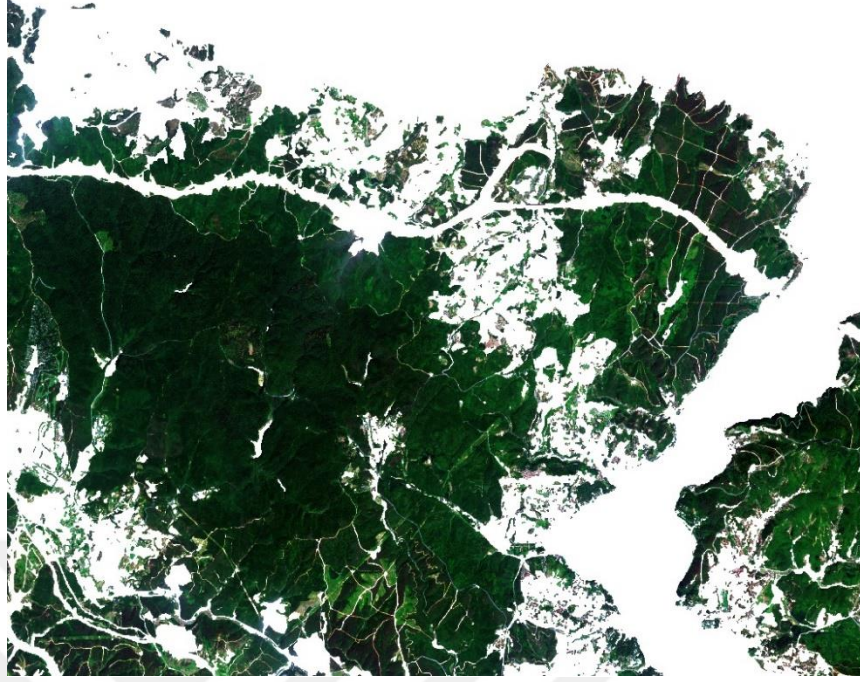
4. BULGULAR

Çalışmanın sonucunda elde edilen sınıflandırılmış görüntüler ve bu görüntülerin doğruluk analizleri bu başlık altında verilecektir. Bu araştırmanın ana amacı olan “obje tabanlı sınıflandırma yöntemlerinin karşılaştırılması”, sınıflandırma algoritmalarının verimliliğini ölçmeye yöneliktir. Ancak sınıflandırma doğruluklarının arttırılması yalnızca algoritmaların yetenekleri ile sınırlı değildir. Daha önceki başlıklarda anlatıldığı üzere algoritmaların sınıflandırma aşamasından önce üzerinde durulması gereken atmosferik düzeltme, segmentasyon, örnek alanların seçilimi, çalışma alanının spektral ayrılabilirliği, vejetasyon indekslerinin kullanılan bantlara dahil edilmesi; gibi birçok etmen doğruluk değerlerine doğrudan etki etmektedir.

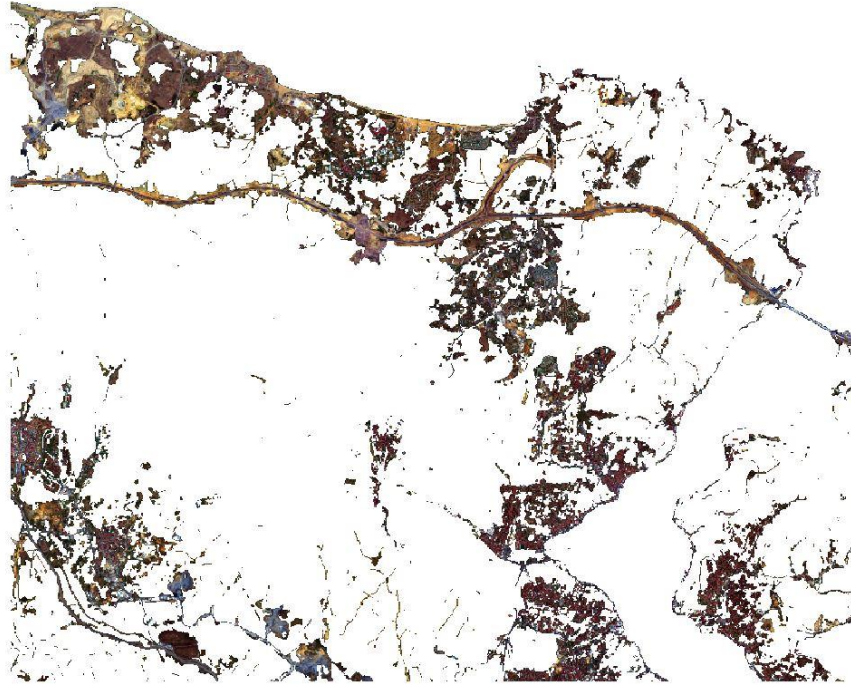
Sınıflandırma için kullanılan görüntülerin her bir sınıflandırma yöntemi için (RF, NN, THC) aynı olması nedeniyle yukarıda bahsedilen etmenlerin doğruluk üzerindeki etkilerinin incelenmesi gerçekleştirilmemiştir, ancak bu başka bir araştırma konusudur.

Sınıflandırma sonucunda elde edilen değerlere bakıldığında NN ile %77.41 ve RF ile %87.77 doğruluklarına ulaşıldığı görülür (Genel Doğruluk). Bu doğruluk değerleri her ne kadar tatmin edici görünse de çalışma alanının heterojen bir yapıda olması ve seçilen sınıfların (özellikle yerleşim alanları için) spektral ayrılabilirliklerinin çok düşük olması bu doğruluk değerlerinin daha yüksek çıkmasını engellemiştir. Bu kanıya aynı görüntüye *hiyerarşik sınıflandırma* uygulanarak yalnızca su yüzeyleri, yerleşim ve vejetasyonun birbiriyle ayrılması sonucu ulaşılmıştır.

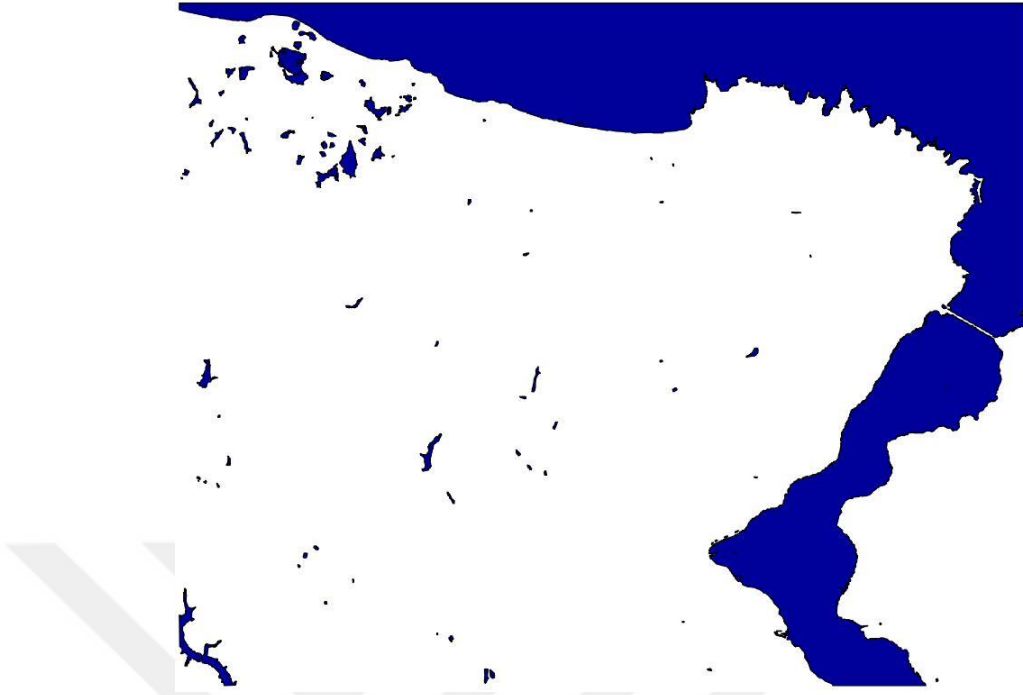
Şekil 4.1, 4.2 ve 4.3’de belirlenen 3 ana sınıfın birbirinden ayrılabilirliği gözlemlenmiş ve elde edilen hata oranı %2.66, yani genel doğruluk değeri %97.34 olmuştur (Tablo 4.1).



Şekil 4.1: Hiyerarşik sınıflandırma ile yerleşim ve su yüzeylerinden ayrılan vejetasyon alanları.



Şekil 4.2: Hiyerarşik sınıflandırma ile vejetasyon ve su yüzeylerinden ayrılan yerleşim alanları, toprak ve yapay yüzeyler.



Şekil 4.3: Hiyerarşik sınıflandırma ile vejetasyon ve yerleşim alanlarından ayrılmış su yüzeyleri.

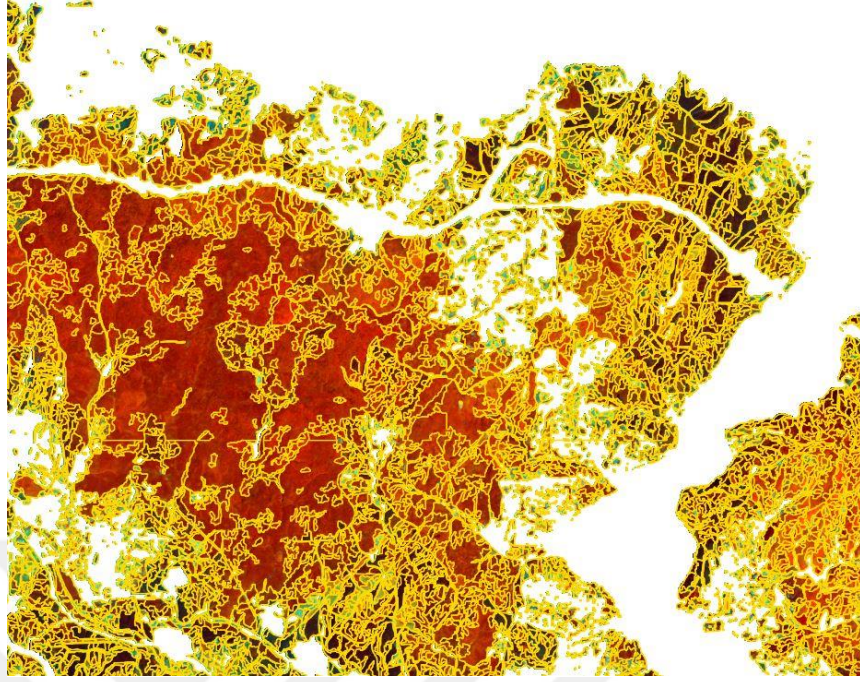
Tablo 4.1: Hiyerarşik sınıflandırma ile 3 ana sınıfa ayrılan görüntünün hata matrisi.

OOB tahmini hata oranı: 2.66%

Hata Matrisi:

	Su	Vejetasyon	Yerleşim Alanı	Sınıf Hatası
Su	46	4	0	0.08888888
Vejetasyon	0	50	0	0.00000000
Yerleşim Alanı	0	0	50	0.00000000

Buradan ulaşılabilecek sonuç, sınıf sayısının artırılması ve spektral ayrılabilirliği birbirine benzerlik gösteren sınıfların sınıflandırmada kullanılması, genel doğruluğu etkilemektedir. Bu nedenle sınıflandırmanın amacı doğrultusunda ana sınıflar belirlenip daha sonra bu alanlar üzerinde tekrar sınıflandırma yapılması yüksek doğruluk oranlarına ulaşmamıza yarayacaktır.



Şekil 4.4: Vejetasyon sınıfına ait alanların daha sonra kendi içerisinde objelere ayrılması.

Söz konusu bulgular sınıflandırma doğruluklarının arttırılmasına yöneliktir, ancak araştırma konusunun sınıflandırma algoritmalarının karşılaştırılması olması nedeniyle görüntüler hiyerarşik sınıflandırma gerçekleştirilmeden sınıflandırmaya alınmıştır. Ancak yüksek doğruluk değerlerine ulaşılması planlanıyorsa hiyerarşik sınıflandırmanın kullanılması önemli bir araç olacaktır (Akar, Ö. 2017).

4.1. NEAREST NEIGHBOR SINIFLANDIRMA SONUÇLARI

eCognition yazılımının NN modülü ile yapılan sınıflandırmada belirlenen 9 sınıfın doğruluk değerleri ve sınıflandırılmış görüntü Tablo 4.2 ve Şekil 4.5’de verilmiştir.

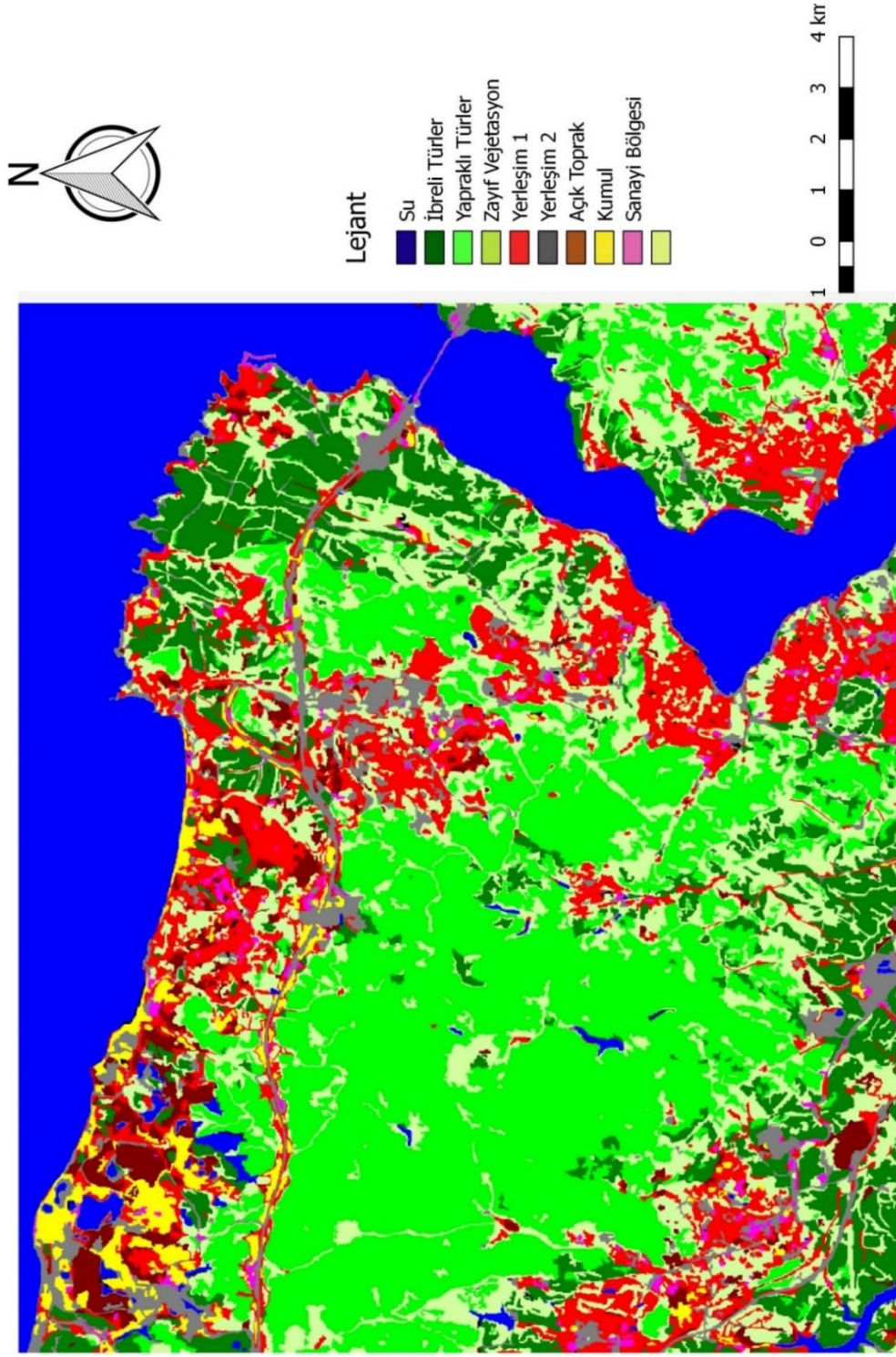
Sınıflandırılmış görüntü ve doğruluk değerlerine bakıldığında su yüzeyleri ve vejetasyon bakımından oldukça verimli sonuçlar (%77 ile %93 aralığında) alındığı söylenebilir. Yapay alanlarda ise bu sonuç oldukça düşük görünmektedir (%56 ile %70 aralığında). Bunun nedeni daha önce bahsedildiği üzere bu alanların spektral ayrılabilirliğinin oldukça düşük olmasından ileri gelmektedir.

Tablo 4.2: NN Sınıflandırma doğrulukları.

Hata Matrisi	Su Yüzeyleri	İbrelili Türler	Yapraklı Türler	Zayıf Vejetasyon	Yerleşim 1	Yerleşim 2	Açık Alanlar	Kumullar	Endüstriyel Alanlar
Su Yüzeyleri	25	3	0	0	0	0	2	0	0
İbrelili Türler	2	26	0	0	0	0	2	0	0
Yapraklı Türler	0	0	25	5	0	0	0	0	0
Zayıf Vejetasyon	0	0	0	24	4	0	2	0	0
Yerleşim 1	0	0	0	0	24	4	1	0	1
Yerleşim 2	0	0	0	2	0	17	3	3	5
Açık Alanlar	0	0	0	0	0	4	23	3	0
Kumullar	0	0	0	0	0	0	4	21	5
Endüstriyel Alanlar	0	0	0	0	0	3	0	3	24
Toplam	27	29	25	31	28	28	37	30	35

SINIF İSİMLERİ	Referans Toplamları	Toplam Sınıflandırılmış	Doğru Sınıflandırılmış	Üretici Doğruluğu	Kullanıcı Doğruluğu
Su Yüzeyleri	27	30	25	92.59%	83.33%
İbrelili Türler	29	30	26	89.66%	86.67%
Yapraklı Türler	25	30	25	100.00%	83.33%
Zayıf Vejetasyon	31	30	24	77.42%	80.00%
Yerleşim 1	28	30	24	85.71%	80.00%
Yerleşim 2	28	30	17	60.71%	56.67%
Açık Alanlar	37	30	23	62.16%	76.67%
Kumullar	30	30	21	70.00%	70.00%
Endüstriyel Alanlar	35	30	24	68.57%	80.00%
Toplam	270	270	209		
Toplam Sınıflandırma Doğruluğu:					77.41%

Kappa İstatistikleri	Sonuç Kappa Değeri:
Her Sınıfın Kendi İçerisindeki Kappa Değerleri	0.7458
Su Yüzeyleri	0.8148
İbrelili Türler	0.8506
Yapraklı Türler	0.8163
Zayıf Vejetasyon	0.7741
Yerleşim 1	0.7769
Yerleşim 2	0.5165
Açık Alanlar	0.7296
Kumullar	0.6625
Endüstriyel Alanlar	0.7702



Şekil 4.5: Nearest Neighbor algoritması ile sınıflandırılmış Sentinel-2a görüntüsü.

4.2. RANDOM FOREST SINIFLANDIRMA SONUÇLARI

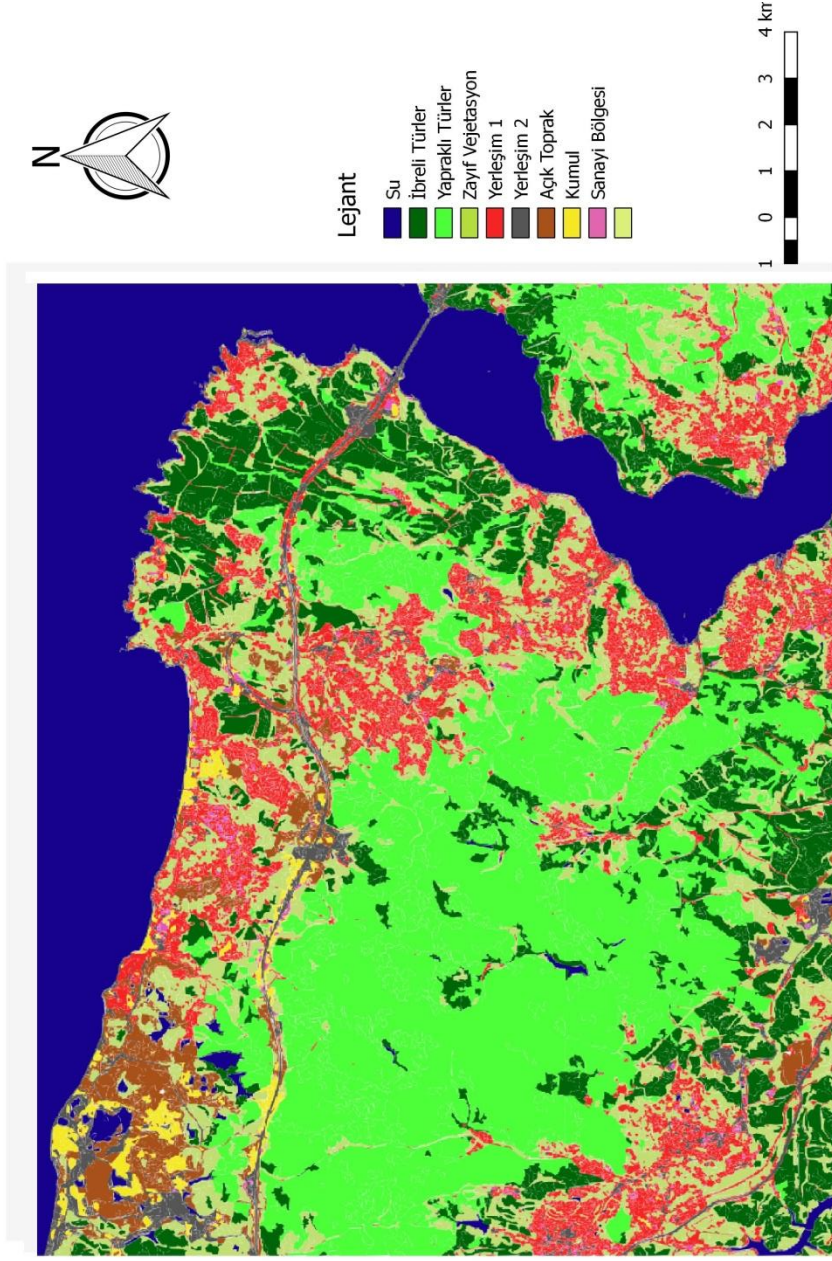
QGIS ve R Studio yazılımları kullanılarak gerçekleştirilen RF sınıflandırmasında belirlenen 9 sınıfın doğrulukları ve sınıflandırılmış görüntüsü Tablo 4.3 ve Şekil 4.6’da verilmiştir.

Tablo 4.3: Random Forest Sınıflandırma Doğrulukları.

Hata Matrisi	Su Yüzeyleri	İbrelili Türler	Yapraklı Türler	Zayıf Vejetasyon	Yerleşim 1	Yerleşim 2	Açık Alanlar	Kumullar	Endüstriyel Alanlar
SINIFLANDIRILMIŞ VERİLER									
Su Yüzeyleri	30	0	0	0	0	0	0	0	0
İbrelili Türler	0	29	0	1	0	0	0	0	0
Yapraklı Türler	0	0	28	1	0	0	0	0	0
Zayıf Vejetasyon	0	1	1	28	0	0	0	0	0
Yerleşim 1	0	0	0	0	24	2	2	0	1
Yerleşim 2	0	0	0	0	1	23	3	1	2
Açık Alanlar	0	0	0	0	1	1	25	2	1
Kumullar	0	0	0	0	1	0	3	24	3
Endüstriyel Alanlar	0	0	0	0	1	1	0	2	26
Toplam	30	30	29	31	28	27	33	29	33

SINIF ISIMLARI	Referans Toplamları	Toplam Sınıflandırılmış	Doğru Sınıflandırılmış	Üretici Doğruluğu	Kullanıcı Doğruluğu
Su Yüzeyleri	30	30	30	100%	100%
İbrelili Türler	30	30	29	96.66%	96.66%
Yapraklı Türler	29	30	28	96.55%	96.66%
Zayıf Vejetasyon	31	30	28	90.32%	93.33%
Yerleşim 1	28	30	24	85.71%	80%
Yerleşim 2	27	30	23	85.18%	76.66%
Açık Alanlar	33	30	25	75.75%	83.33%
Kumullar	29	30	24	82.75%	80%
Endüstriyel Alanlar	33	30	26	78.78%	86.66%
Toplam	270	270	237		
Toplam Sınıflandırma Doğruluğu: 87.77%					

Kappa İstatistikleri	Sonuç Kappa Değeri:	Her Sınıfın Kendi İçerisindeki Kappa Değerleri
		1
		0.9666
		0.9333
		0.9032
		0.8571
		0.8518
		0.7575
		0.8275
		0.7878



Şekil 4.6: RF algoritması ile sınıflandırılmış Sentinel-2a uydu görüntüsü.

Sınıflandırma sonuçlarına bakıldığında su yüzeylerinin hatasız olarak sınıflandırılmış olması göze çarpmaktadır. Bununla birlikte vejetasyonun sınıflandırılma doğruluklarının %90 üzerinde olduğu görülür. Sınıflandırmanın doğruluğunu NN sınıflandırıcısında da azaltan sınıflar olan yapay yüzeyler (yerleşim, sanayi bölgesi, maden sahaları vb.) RF sınıflandırıcısında en düşük değerlere sahiptir.

4.3. THRESHOLD SINIFLANDIRMA SONUÇLARI

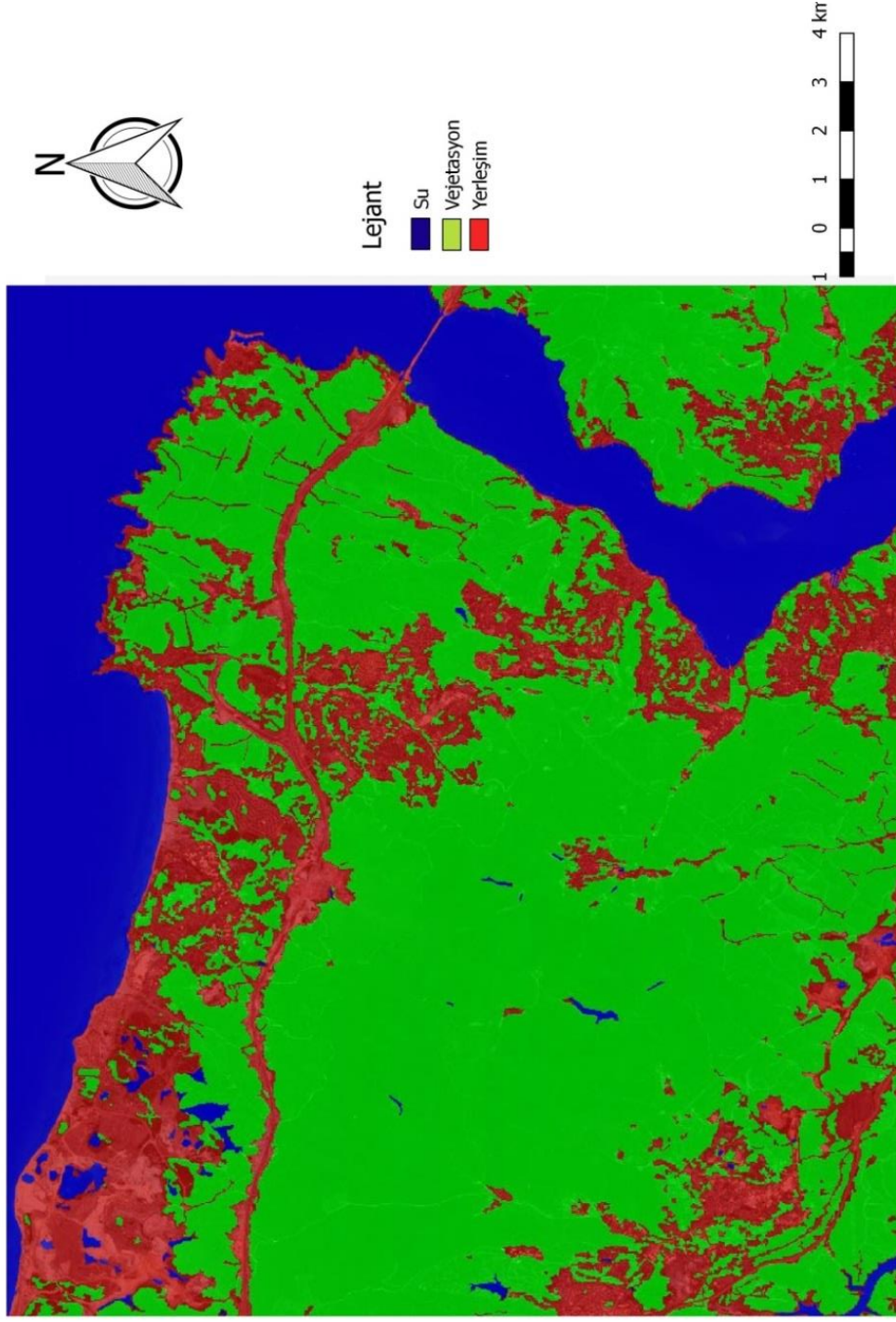
Belirlenen 3 farklı eşik değeri ve sınıfların bu eşik değerler arasında yer alma durumuna göre sınıflandırmanın gerçekleştirilmesi, bu sınıflandırmanın temelini oluşturur. Basit bir iş akışına sahip olan bu sınıflandırma, farklı bantların spektral aralıklarının, ortalama, standart sapma ve kullanılabilecek diğer istatistiksel (min/max, median, percentile vb.) değerlerinin tek bir banda indirgenmesiyle (vejetasyon indeksleri gibi) birçok sınıfın ayrılabilirliği sağlanabilecektir.

Bu yöntem ile belirlenen 3 ana sınıfın (su yüzeyleri, vejetasyon ve vejetasyon dışı alanlar) birbirinden ayrılması oldukça başarılı bir şekilde gerçekleştirilebiliyorken, ana sınıfların daha sonra kendi içinde sınıflandırılması için bir yöntem belirlenememiştir. Bu nedenle bu sınıflandırma yönteminin diğer sınıflandırma yöntemleri ile belirlenen 9 sınıfta karşılaştırılması mümkün değildir. Bu durum göz önünde bulundurularak diğer iki sınıflandırma yöntemleriyle elde edilen 9 sınıf, Threshold sınıflandırıcı ile elde edilebilen 3 ana sınıfa (su yüzeyleri, vejetasyon ve vejetasyon dışı alanlar) daraltılıp doğruluklarını karşılaştırmak mümkün olmuştur.

Tablo 4.4: THC Sınıflandırma doğrulukları

Hata Matrisi					
SINIFLANDIRILMIŞ VERİLER	Su Yüzeyleri	Vejetasyon	Yerleşim		
Su Yüzeyleri	28	2	0		
Vejetasyon	1	29	0		
Yerleşim	0	2	28		
Toplam	29	0	0		

SINIF İSİMLERİ	Referans Toplamları	Toplam Sınıflandırılmış	Doğru Sınıflandırılmış	Üretici Doğruluğu	Kullanıcı Doğruluğu
Su Yüzeyleri	30	30	28	96.55%	100.00%
İbrelili Türler	30	30	29	87.87%	100.00%
Yapraklı Türler	30	30	28	93.33%	100.00%
Yapraklı Türler	90	90	85		
Toplam Sınıflandırma Doğruluğu:			94.44%		



Şekil 4.7: Threshold Sınıflandırma ile görüntünün 3 ana sınıfa ayrılması.

Tablo 4.5: Threshold, Random Forest ve Nearest Neighbor Sınıflandırıcılarının 3 ana sınıfta karşılaştırılması.

	Random Forest Sınıflandırma		Alanların Dağılımı	Nearest Neighbor Sınıflandırma		Alanların Dağılımı	Threshold Sınıflandırma		Alanların Dağılımı
	Üretici Doğruluğu(%)	Kullanıcı Doğruluğu (%)		Üretici Doğruluğu(%)	Kullanıcı Doğruluğu (%)		Üretici Doğruluğu(%)	Kullanıcı Doğruluğu (%)	
Su Yüzeyleri	100	100	6999.79	99.6	100	6971.79	96.55	100	6758.29
Vejetasyon	98.7	100	18555.86	97.2	99.8	19170.55	87.87	100	17519.79
Vej. Dışı Alanlar	96.2	97.6	6507.87	94.5	100	5921.181	93.33	100	7785.44
Toplam (% / ha)	98.76		32063.52	96.24		32063.52	94.44		32063.52

Tablo 4.6: Random Forest ve Nearest Neighbor Sınıflandırıcılarından 9 sınıfta elde edilen sonuçlar ve sınıfların alanlara göre dağılımı.

	Random Forest Sınıflandırma			Alanların Dağılımı (ha)	Nearest Neighbor Sınıflandırma			Alanların Dağılımı (ha)
	Üretici Doğruluğu (%)	Kullanıcı Doğruluğu (%)			Üretici Doğruluğu (%)	Kullanıcı Doğruluğu (%)		
Su Yüzeyleri	100	100		6999.79	92.59	83.33		7053.46
İbrelili Türler	96.66	96.66		4576.55	89.66	86.67		3503.54
Yapraklı Türler	90.55	96.66		9144.98	100	83.33		8413.87
Zayıf Vejetasyon	90.32	93.33		5078.74	77.42	80		5916.96
Yerleşim 1	85.71	80		3581.48	85.71	80		3732.32
Yerleşim 2	85.18	76.66		1026.6	60.71	56.67		1665.14
Açık Toprak	75.75	83.33		1120.17	62.16	76.67		813.01
Kumul	85.75	80		385.3	70	70		634.55
Sanayi Bölgesi	78.78	86.66		152.25	68.57	80		291.55
Toplam (% / ha)	87.77			32063.52	77.41			32063.52

5. TARTIŞMA VE SONUÇ

Sınıflandırma işlemi ile görüntülerin hangi arazi örtü ya da kullanım sınıflarına sahip olduğu; bu sınıfların kapladığı alanlar, sınırları ve görüntünün kullanım alanına göre birçok dijital bilgi içermesi sağlanabilir. Bu bilgilerin kullanım alanı ise ormancılık ve ziraat açısından bakıldığında, iş yükünü hafifletmesi bakımından alternatifi olmayan bir yöntemdir. Orman yollarının belirlenmesi ve yeni güzergahların oluşturulması, orman yangınları sonucu oluşan tahribatın hesaplanması, böcek zararına uğramış alanların tespiti, ağaç türlerinin tespiti (yüksek çözünürlüklü ve hiperspektral görüntüler ile), su yönetimi, ormancılık politikalarının saptanması (makro veya mikro ölçekli), kadastral sınırların belirlenmesi ve kaçak yapılaşmanın tespiti gibi bir çok kullanım alanı türetilir.

Görüntü sınıflandırmada kaydedilen ilerlemeler ile son teknoloji sınıflandırma yöntemlerinden biri olan obje tabanlı sınıflandırma yöntemine yönelim hızla artmaktadır. Bu yöntemlerin kullanımı için gereken verilerin seçilmesi çalışmanın sonucunu etkileyen en önemli faktörlerden birisidir. Bu noktada yeni ve ücretsiz olan ESA'nın geliştirdiği Sentinel uyduvarının bu yöntemlerin değerlendirilmesinde kullanılması hem sınıflandırma algoritmalarının orta/yüksek çözünürlüklü uydu verilerindeki performansını ölçmek hem de Sentinel 2a uydu verilerinin obje tabanlı sınıflandırma yöntemleri için ne kadar verimli olduğunu saptamak için tercih edilmiştir.

Sınıflandırma aşamasında Sentinel uydu verilerinin kullanımı ile ilgili yapılan araştırmalarda farklı yöntemlere başvurulmuş (zaman serileri veya radar verileriyle birlikte kullanılması gibi) ve bunların sonuçlara olumlu etkileri görülmüştür (Clerici ve diğ. 2017, Xiong ve diğ. 2017). Ancak bu çalışma kapsamında uydunun tek tarihli görüntüsünün spektral özellikleri kullanılmıştır, dolayısıyla çalışmanın sonuçlarını bahsedilen bu yöntemler ile geliştirmek mümkündür. Sentinel-2 uydusunun spektral

görüntüleme potansiyeli, su yüzeylerinin belirlenmesi (Du ve diğ. 2016), zirai alanların haritalanması (Immitzer ve diğ. 2016); yaprak nitrojen konsantrasyonu tahmini (Ramoelo ve diğ. 2015), meşçere klorofil içeriği, yaprak yüzey alanı (LAI) ve yaprak klorofil yoğunluğu (LCC) gibi parametrelerin tahmini (Frampton ve diğ. 2013), yangın zararı tespiti (Fernandez-Manso ve diğ. 2016) gibi farklı konular üzerinde ortaya konulmuştur. Bu araştırmalar ile ulaşılan sonuçlardan biri; sınıflandırma doğruluklarının arttırılması amacıyla bantlar arasındaki korelasyonun etkili bir biçimde kullanılması; dolayısıyla vejetasyon ve zirai alanlar hakkında bilgi içeren bilgi içeren bantların ve indekslerin sınıflandırmaya katılması olmuştur. Bu çalışma kapsamında kullanılan bantlar ve vejetasyon indekslerinin seçilmesinde dikkate alınan başlıca kaynaklar bunlar olmuştur.

Yine yukarıda bahsedilen bu araştırmalar kapsamında kullanılan sınıflandırma algoritmalarından Random Forest ve NN sınıflandırıcısı çalışma alanının homojen dağılım gösterdiği yerlerde sınıflandırmayı yüksek doğruluklarla sağlamıştır. Kullanılan yöntemlerin benzer homojenliğe sahip farklı veri setleri üzerinde denenmesi ile neredeyse aynı sonuçlara varılması, sınıflandırma sonuçlarının veriden bağımsız olduğunu göstermektedir. Aynı zamanda sınıflandırma doğruluklarının arttırılması için gereken en önemli etmenlerin eğitim alanlarının saf piksellerden elde edilmesi ve seçilen sınıfların spektral ayrılabilirliklerinin olabildiğince yüksek olması gerektiği daha önce yapılan bu araştırmalar sonucunda da ortaya konulmuştur.

Bu çalışma içerisinde karşılaştırılan obje tabanlı sınıflandırma yöntemleri olan NN ve RF 'ın verimlilikleri, “bulgular” başlığındaki doğruluk değerlerinin karşılaştırılması ile mümkün olur. Sınıflandırma aşamasında her iki yöntem için Sentinel-2a uydusundan elde edilen aynı 14 bantlı görüntü (10 orijinal bant ve 4 adet vejetasyon indeksi) ve bu görüntülerden elde edilen aynı istatistiksel veriler (ortalama, standart sapma, minimum ve maksimum değerler) kullanılmıştır. Segmentasyon sonucu elde edilen 12533 adet objenin sınıflandırma doğrulukları bulgular başlığında verilmiştir. Bu sonuçlardan yola çıkarak RF sınıflandırıcısının NN sınıflandırıcısına olan üstünlüğü görülebilir. eCognition içerisinde gerçekleştirilen NN sınıflandırma sürecinde objelerin spektral

değerlerine ek olarak obje özelliklerinin (tekstür) de hesaba katılıp (örneğin: yollar uzun ve düz bir çizgi, binalar dörtgen, ağaçlar ise yüksek çözünürlüklü görüntülerde dairemsi yapılar oluştururlar) sınıflandırma doğruluklarının artırılması mümkündür. Ancak objelerin spektral değerleri dışında oluşturduğu biçimlerin hesaba katılması yoğun bir iş gücü gerektirmektedir.

Bu sonuçların makro ölçekli zirai ve ormancılık uygulamalar için yeterli olduğunu söyleyebiliriz. Ancak RF ve NN sınıflandırıcılarının Sentinel-2 uydu verileri üzerinde daha etkin bir şekilde kullanımının mümkün olduğu daha önce bu uydular üzerinde yapılan araştırmalarda ortaya konulmuştur. Sınıflandırmanın daha detaylı ve ilgi alanı doğrultusunda daha fazla bilgi içermesi (toprak yapısı, tür çeşitliliği, arazi kullanım biçimi, yangın/böcek zararı vb.), çalışmanın amacı doğrultusunda belirlenen sınıfların spektral ayrılabilirliklerinin tam anlamıyla hesaba katılmasıyla; ayrıca Sentinel-1 Radar verileri ve Sentinel-2 uydu çiftinden elde edilen zaman serilerinin de kullanımıyla gerçekleştirilip, operasyonel tarım ve ormancılık uygulamalarında kullanılacak seviyeye getirilebilir.

Bu iki yaygın sınıflandırma yönteminden ayrı olarak THC sınıflandırıcısı kontrolsüz bir sınıflandırma yöntemidir. Ancak eCognition içerisinde oluşturulacak farklı argümanlarla (sınıflandırma kuralları) başarılı sonuçlar verebilecek potansiyele sahip bir sınıflandırma yöntemi olduğu görülmüş: 3 ana sınıfta gerçekleştirilebilen sınıflandırma sonuçlarına bakıldığında su yüzeyleri, vejetasyon ve vejetasyon dışı alanları diğer obje tabanlı yöntemler kadar yüksek doğruluklarla (%94.44) sınıflandırılabilirdiği ortaya konulmuştur.

KAYNAKLAR

- A. Fernández-Manso, O. Fernández-Manso, C. Quintano. SENTINEL-2A red-edge spectral indices suitability for discriminating burn severity. *Int. J. Appl. Earth Obs. Geoinf.*, 50 (2016), pp. 170-175.
- Akpınar H. [Knowledge discovery from databases and data mining]. *İÜ İşletme Fakültesi Dergisi* 2000 ; 29(1):1.
- Alan S. Belward, Jon O. Skøien, 2014. Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. *ISPRS Journal of Photogrammetry and Remote Sensing* 103 (2015) 115–128
- Antoine Lefebvre, Christophe Sannier ve Thomas Corpetti, 2016. Monitoring Urban Area with Sentinel-2A Data: Application to the Update of the Copernicus High Resolution Layer Imperviousness Degree. *Remote Sensing* 2016, 8, 606; doi:10.3390/rs8070606.
- Andrej Karpathy, 2017. *Convolutional Neural Networks for Visual Recognition*. <http://cs231n.github.io/object-detection/>.
- A. Opelt, M. Fussenegger, A. Pinz, and P. Auer. Weak hypotheses and boosting for generic object detection and recognition. *In ECCV, 2004*.
- A. Ramoelo, M. Cho, R. Mathieu, A.K. Skidmore. Potential of Sentinel-2 spectral configuration to assess rangeland quality. *J. Appl. Remote. Sens.*, 9 (2015)
- Baker, J.C., O'Connell, K.M., Williamson, R. (Eds.), 2001. *Commercial observation satellites*. Rand Corporation ISmpvEApA 0-1aa0-29sing, pe. 6uI.
- Barnsley, M.J., 1999, Digital remote sensing data and their characteristics. In P. Longley, Longley, M. Goodchild, D.J. Maguire and D.W. Rhind (Eds), *Geographical Information Systems: Principles, techniques, applications, and management*, 2nd edn, pp. 451–466 (New York: John Wiley and Sons).
- Beizer, B., *Black Box Testing*. New York: John Wiley & Sons, Inc., 1995. [3] A. Bertoli "Chapter 5: Software Testing," in *IEEE SWEBOK Trial Version 1.00*, May 2001.
- Blackburn, G.A., 1998. Quantifying chlorophylls and carotenoids at leaf and canopy scales: an evaluation of some hyperspectral approaches. *Remote Sens. Environ.* 66, 273–285.

- Breiman L. 2001. *Random forests*. *Machine Learning* 45(1): p.5–32
- Bosch, A. Zisserman, and X. Munoz. *Image Classification using random forests and ferns*. In ICCV, 2007.
- C. Atzberger, K. Richter, 2012. Spatially constrained inversion of radiative transfer models for improved LAI mapping from future Sentinel-2 imagery. *Remote Sensing of Environment; Volume 120*, 15 May 2012, Pages 208–218.
- Casey Cleave, Maggi Kelly, Faith R. Kearns, Max Moritz, 2008. Classification of Wildland/Urban Interface: A Comparison of Pixel and Object-Based Classifications Using High Resolution Aerial Photography. *Computers, Environment and Urban Systems; Volume 32*, Issue 4, July 2008, Pages 317–326.
- C. E. Shannon, 1948. A Mathematical Theory of Communication. *The Bell System Technical Journal*, Vol. 27, pp. 379–423, 623–656, July, October, 1948.
- Cihlar, J., 2000. Land cover mapping of large areas from satellites: status and research priorities. *Int. J. Remote Sens.* 21 (6–7), 1093–1114.
- Davis, G., 2007. History of the NOAA satellite program. *J. Appl. Remote Sens.* 1 (1). <http://dx.doi.org/10.1117/1.2642347>, pp. 012504–012504.
- Delegido, J., Vergara, C., Verreist, J., Gandia, S., Moreno, J., 2011a. *Remote estimation of crop chlorophyll content by means of high-spectral-resolution reflectance techniques*. *Agron. J.* 103, 1834–1842.
- Desheng Liu and Fan Xia, 2010. Assessing Object-Based Classification: Advantages and Limitations, *Remote Sensing Letters*, 1:4, 187-194.
- Dronova I., Gong P., Wang L., (2011), Object-based analysis and change detection of major wetland cover types and their classification uncertainty during the low water period at Poyang Lake, China, *Remote Sensing of Environment*, 115(12), 3220-3236.
- Enderle, D, and R.C. Weih, Jr., 2005. Integrating supervised and unsupervised classification methods to develop a more accurate land cover classification, *Journal of the Arkansas Academy of Science*, 59:65-73.
- Orman Genel Müdürlüğü, *Fonksiyonel Orman Amenajman Planı*, sayfa 44-48, 2012.
- Franklin, S.E., Peddle, D.R., Dechka, J.A. and Stenhouse, G.B., 2002, Evidential reasoning with Landsat TM, DEM and GIS data for land cover classification in support of grizzly bear habitat mapping. *International Journal of Remote Sensing*, 23, pp. 4633–4652.

- Gallego, F.J., 2004, Remote sensing and land cover area estimation. *International Journal of Remote Sensing*, 25, pp. 3019–3047.
- Gao J., 2009, *Digital analysis of remotely sensed imagery*, The Mc Graw-Hill Companies, USA.
- Gao, Y. and J.F. Mas, 2008. *A comparison of the performance of pixel-based and object-based classifications over images with various spatial resolutions*, GEOBIA 2008–Pixels, Objects, Intelligence. GEOgraphic Object Based Image Analysis for the 21st Century, (G.J.Hay, T. Blaschke and D. Marceau, editors). University of Calgary, Alberta, Canada, ISPRS Vol. No. XXXVIII-4/C1. Archives ISSN No.:1682-1777,
- Georgios Mallinis, Nikos Koutsias, Maria Tsakiri-Strati, Michael Karteris, 2007. Object-Based Classification Using Quickbird Imagery for Delineating Forest Vegetation Polygons in a Mediterranean Test Site. *ISPRS Journal of Photogrammetry & Remote Sensing* 63 (2008) 237–250.
- Gislason PO, Benediktsson JA, Sveinsson JR. 2004. Random Forest classification of multisource remote sensing and geographic data. In: *IEEE International Geoscience and Remote Sensing Symposium Proceedings (IGARSS 2004)*; September 20–24; Anchorage (AK), vol. II, Hoboken (NJ): IEEE International.
- Gitelson, A.A., Viña, A., Verma, S.B., Runquist, D.C., Arkebauer, T.J., Keydan, G., Leavitt, B., Ciganda, V., Burba, G.G., Suyker, A.E., 2006. *Relationship between gross primary production and chlorophyll content in crops: Implications for the synoptic monitoring of vegetation productivity*. J. Geophys. Res. 111, D08S11.
- Gong, P., D. Marceau, and P. J. Howarth. 1992. A comparison of spatial feature extraction algorithms for land use classification with SPOT HRV data. *Remote Sensing of Environment* 40, 137–151.
- Guanter, L.; Gómez-Chova, L.; Moreno, J. Coupled retrieval of aerosol optical thickness, columnar water vapor and surface reflectance maps from ENVISAT/MERIS data over land. *Remote Sens. Environ.* 2008, 112, 2898–2913.
- Guo, Q., Kelly, M., Gong, P. and Liu, D., 2007, *An object-based classification approach in mapping tree mortality using high spatial resolution imagery*. *GIScience & Remote Sensing*, 44, pp. 24–47.
- Hagolle, O.; Dedieu, G.; Mougnot, B.; Debaecker, V.; Duchemin, B.; Meygret, A. Correction of aerosol effects on multi-temporal images acquired with constant viewing angles: Application to Formosat-2 images. *Remote Sens. Environ.* 2008, 112, 1689–1701.

- Hamilton, R.; Megown, K.; Mellin, T.; Fox, I. 2007. *Guide to automated stand delineation using image segmentation*. Rep. No. RSAC-0094-RPT1. Salt Lake City, UT: U.S. Department of Agriculture Forest Service, Remote Sensing Applications Center. 17p.
- Hakan Yener, Ayhan Koc, H.Oguz Coban, 2006. Uzaktan Algılama Verilerinde Sınıflandırma Doğruluğunun Belirlenmesi Yöntemleri. *İstanbul Üniversitesi Orman Fakültesi Dergisi*, Seri B, Cilt 56, Sayı 2, 2006.
- Hakan Yener, Ayhan Koç ve Oğuz Çoban, 2006. Uzaktan Algılama Verileri ve Teknik Özellikleri. *İstanbul Üniversitesi Orman Fakültesi Dergisi*, Seri B; Cilt 56; Sayı 1; 2006.
- Hay, G. J., and G. Castilla. 2006. Object-based image analysis: Strengths, Weaknesses, Opportunities and Threats (SWOT). Paper presented at *1st International Conference on Object-based Image Analysis (OBIA 2006)*.
- Hüseyin Yurtseven, 2008. *Using of High Resolution Satellite Images In Object Based Image Analysis*, Istanbul University Department of Forest Engineering, Thesis (Phd), 5-15 pages.
- Ismail Colkesen, 2009. *Uzaktan Algılamada İleri Sınıflandırma Tekniklerinin Karşılaştırılması ve Analizi*. Thesis (Msc.). Gebze Technical University.
- Ismail Colkesen, Taskin Kavzoglu, Tahsin Yomralioglu, 2015. The use of support vector machine for optimal band selection for remotely sensed images. *Türkiye Ulusal Fotogrametri ve Uzaktan Algılama Birliği VIII. Sempozyumu (TUFUAB2015)*, 21–23 Mayıs 2015.
- Joaquín Derrac, Salvador García, Francisco Herrera, 2013. Fuzzy nearest neighbor algorithms: Taxonomy, experimental analysis and prospects. *Information Sciences* 260 (2014) 98–119
- Jordan, C.F., 1969. *Derivation of leaf-area index from quality of light on the Forest floor*. Ecology 50, 663–666.
- Jing Qian, Qiming Zhou, Quan Hou, 2006, Comparison Of Pixel-Based And Object-Oriented Classification Methods For Extracting Build-up Areas In Aridzone, *ISPRS Workshop on Updating Geo-spatial Databases with Imagery & The 5th ISPRS Workshop on DMGISs*.
- Kaan Kalkan ve Derya Maktav, 2010. Nesne Tabanlı ve Piksel Tabanlı Yöntemlerin Karşılaştırılması (IKONOS örneği). III. *Uzaktan Algılama ve Coğrafi Bilgi Sistemleri Sempozyumu*.
- Katja Richter, Clement Atzberger, Francesco Vuolo, Guido D'Urso, 2010. Evaluation of Sentinel-2 Spectral Sampling for Radiative Transfer Model Based LAI Estimation of Wheat, Sugar Beet, and Maize. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (Volume: 4, Issue: 2, June 2011)

- Katja Richter, Clement Atzberger, Francesco Vuolo, Guido D'Urso, 2010. Evaluation of Sentinel-2 Spectral Sampling for Radiative Transfer Model Based LAI Estimation of Wheat, Sugar Beet, and Maize. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (Volume: 4, Issue: 2, June 2011)
- K. Grauman and T. Darrell. *The pyramid match kernel: Discriminative classification with sets of image features*. In ICCV, 2005.
- Kibriya, A. M. and Frank, E. (2007). An empirical comparison of exact nearest neighbour algorithms. In *Proceedings of the 11th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD'07)*, volume 4702 of Lecture Notes in Computer Science, pages 140–151. Springer.
- Kononenko, I. and Hong, S.J., 1997. *Attribute selection for modelling*. *Future Generation Computer Systems*, 13, 181-195.
- Landgrebe, D.A., 2003, *Signal Theory Methods in Multispectral Remote Sensing* (Hoboken, NJ: John Wiley and Sons).
- Laura Dingle Robertson and Douglas J. King, 2011. Comparison of Pixel and Object Based Classification in Land Cover Change Mapping. *International Journal of Remote Sensing*, Volume 32 – 2011, Issue 6, pages 1505 – 1529.
- Lefsky, M.A. and Cohen, W.B., 2003, Selection of remotely sensed data. In M.A. Wulder and S.E. Franklin (Eds), *Remote Sensing of Forest Environments: Concepts and case studies*, pp. 13–46 (Boston: Kluwer Academic Publishers).
- Lillesand, T.M., Kiefer, R.W. ve Chipman, J.W. (2007). *Remote Sensing and Image Interpretation*. 6th Edition, John Wiley & Sons., New York, USA.
- Lowry, J. H. Jr., R.D. Ramsey, K. Boykin, D. Bradford, P. Comer, S. Falzarano, W. Kepner, J. Kirby, L. Langs, J. Prior-Magee, G. Manis, L. O'Brien, K. Pohs, W. Rieth, T. Sajwaj. 2007. Mapping meso-scale land cover over very large geographic areas within a collaborative framework: A case study of the Southwest Regional Gap Analysis Project (SWReGAP). *Remote Sensing of Environment* 108:59-73.
- Loveland, T.R., Zhu, Z., Ohlen, D.O., Brown, J.F., Reed, B.C., Yang, L., 1999. An analysis of the IGBP global land-cover characterization process. *Photogrammetric Eng. Remote Sens.* 65 (9), 1021–1032.
- Lu, D. and Weng, Q., 2004, Spectral mixture analysis of the urban landscapes in Indianapolis with Landsat ETM+ imagery. *Photogrammetric Engineering and Remote Sensing*, 70, pp. 1053–1062.
- Maggi Kelly, Samuel D. Blanchard, Kevin Koy and Ellen Kersten 2011. Terrestrial Remotely Sensed Imagery in Support of Public Health: New Avenues of Research Using Object-Based Image.. *Remote Sens.* 2011, 3, 2321-2345.

- Marpa, P.R., I. Niemeyer, and R. Gloaguen, 2006. A procedure for automatic object-based classification, *Proceedings from 1st International Conference on Object-based Image Analysis*, 4-5 July, Salzburg, Austria.
- Martin BAATZ und Arno SCHÄPE, 2016. *Multiresolution Segmentation: an optimization approach for high quality multi-scale image Segmentation*.
- Markus Immitzer, Francesco Vuolo and Clement Atzberger, 2016. First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe, *Remote Sens.* 2016, 8, 166; doi:10.3390/rs8030166
- Martin, D., C. Fowlkes, D. Tal, and J. Malik, 2001. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics, *Proceedings of the Eighth International Conference on Computer Vision (ICCV'01), Volume 2*
- McRoberts, R.E. and E.O. Tomppo, 2007. Remote sensing support for national forest inventories, *Remote Sensing of Environment*, 110:412-419.
- Muhammet AKMAN, Yasemin GENÇ, Handan ANKARALI, 2010. *Random Forests Yöntemi ve Sağlık Alanında Bir Uygulama*. Türkiye Klinikleri J Biostat 2011;3(1)
- M. Varma and D. Ray. *Learning the discriminative power invariance trade-off*. In *ICCV*, 2007.
- Myneni, R. B., F. G. Hall, P.J. Sellers, and A.L. Marshak (1995) 'The interpretation of spectral vegetation indexes', *IEEE Transactions on Geoscience and Remote Sensing*, 33, 481-486.
- N. Clerici, Cesar Augusto Valbuena Calderón & Juan Manuel Posada (2017). *Fusion of Sentinel-1A and Sentinel-2A data for land cover mapping: a case study in the lower Magdalena region, Colombia*, *Journal of Maps*. 13:2, 718-726, DOI: 10.1080/17445647.2017.1372316 .
- N. Clinton, A. Holt, J. Scarborough, L. Yan, P. Gong Accuracy assessment measure for object-based image segmentation goodness. *Photogramm. Eng. Remote Sen.*, 76 (3) (2010), pp. 289–299
- Oren Boiman, Eli Shechtman and Michal Irani, 2008. *In Defense of Nearest-Neighbour Based Image Classification*. pp. 9-21.
- Olivier Hagolle, Mireille Huc, David Villa Pascual ve Gerard Dedieu, 2014. A Multi-Temporal and Multi-Spectral Method to Estimate Aerosol Optical Thickness over Land, for the Atmospheric Correction of FormoSat-2, LandSat, VEN_S and Sentinel- 2 Images. *Remote Sens.* 2015, 7, 2668-2691; doi:10.3390/rs70302668.

- Özlem AKAR, 2017. The Rotation Forest algorithm and object-based classification method for land use mapping through UAV images, *Geocarto International*, DOI: 10.1080/10106049.2016.1277273.
- Özlem Akar, Oğuz Güngör, 2013. Classification of multispectral images using Random Forest algorithm. *Journal of Geodesy and Geoinformation Vol.1, Iss.2*, pp.105-112, November 2012, Journal No.106.
- Pal M. 2003. Random Forest for land cover classification. *Geoscience and Remote Sensing Symposium*, IGARSS '03. Proceedings. 2003 *IEEE International*. 6:3510–3512.
- Pal, M. and Mather, P.M., 2003, An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment*, 86, pp. 554–582
- Pal, N. R., and S. K. Pal. 1993. A Review on Image Segmentation Techniques. *Pattern Recognition* 26:9, 1277-1294.
- P. Grother, G.T. Candela, J.L. Blue, 1997. Fast implementations of nearest-neighbor classifiers, *Pattern Recognition* 30 (1997) 459–465.
- Rajvardhan Oak, 2016. A Study of Digital Image Segmentation Techniques. *International Journal of Engineering And Computer Science* ISSN: 2319-7249, Volume 5 Issue 12 Dec. 2016, Page No. 19770-19783
- R. Fei-Fei, L. and Fergus and P. Perona. Learning generative visual models from few training examples: an incremental bayesian approach tested on 101 object categories. In *CVPR Workshop on Generative-Model Based Vision*, 2004.
- Robert J. Hijmans, Jacob van Etten, Joe Cheng, Matteo Mattiuzzi, Michael Sumner, Jonathan A. Greenberg, Oscar Perpinan Lamigueiro, Andrew Bevan, Etienne B. Racine, Ashton Shortridge, 2016. *Geographic Data Analysis and Modeling (Package "raster")*.
- Santos, T., Tenedório, J., A., Encarnação, S., Rocha, J., 2006. Comparing pixel vs. Object based classifiers for land cover mapping with Envisat-MERIS data. *26th EARSeL Symposium*, Maio, Varsóvia.
- Silahtaroğlu G. [Data mining models]. *Kavram ve Algoritmalarıyla Temel Veri Madenciliği*. 1. Baskı. İstanbul: Papatya Yayıncılık; 2008. s.29-30.
- Tso, B. and Mather, P.M., 2001, *Classification Methods for Remotely Sensed Data (New York: Taylor and Francis Inc)*.

- Townshend, J.R., Latham, J., Arino, O., Balstad, R., Belward, A., Conant, R., Elvidge, C., Feuquay, J., El Hadani, D., Herold, M., Janetos, A., Justice, C.O., Liu Jiyuan, Loveland, T., Nachtergaele, F., Ojima, D., Maiden, M., Palazzo, F., Schmillius, C., Sessa, R., Singh, A., Tschirley, J., Yamamoto, H., 2008. Integrated Global Observation of the Land: an IGOS-P Theme, *IGOL Report No. 8*, GTOS 54 (FAO, Rome) pp. 74.
- U. Mueller-Wilm, 2016. *Sen2Cor Configuration and User Manual*. S2-PDGS-MPC-L2A-SUM-V2.3.
- Watts J. D., Powell S. L., Lawrence R. L., Hilker T., (2011), Improved classification of conservation tillage adoption using high temporal and synthetic satellite imagery, *Remote Sensing of Environment* 115, 66–75.
- Waske B, Heinzel V, Braun M, Menz G. 2007. Random Forests for classifying multi-temporal sar data. Proc. ‘*Envisat Symposium 2007*’; 2007 April 23–27; Montreux, Switzerland.
- Wilhelm, W.W., Ruwe, K., Schlemmer, M.R., 2000. Comparisons of three leaf area index meters in a corn canopy. *Crop Science* 40, 1179–1183.
- William James Frampton, Jadunandan Dash, Gary Watmough, Edward James Milton, 2013. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing* 82 (2013) 83–92.
- Yavuz H. S, Çevikalp H., (2008), A new distance measure for hierarchical clustering, *IEEE 16th Signal Processing, Communication and Applications Conference*, 20-22 April 2008, Aydin, pp.1-4.
- Y. Du, Y. Zhang, F. Ling, Q. Wang, W. Li, X. Li. Water bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band. *Remote Sens*, 8 (2016), p.354.

EKLER

EK 1. Çalışmada Kullanılan RStudio Makroları

```

##### Gerekli Paketler #####
#
# Paketlerin indirilmesi
install.packages("raster", dependencies = TRUE)
install.packages("foreign", dependencies = TRUE)
install.packages("matrixStats", dependencies = TRUE)
install.packages("randomForest", dependencies = TRUE)
#
# Paketlerin yüklenmesi (eğer indirildilerse)
library(raster)
library(foreign)
library(matrixStats)
library(randomForest)
#
##### Çalışma Klasörünün Belirlenmesi #####
# ("D:/your_folder/...")
setwd("D:/GIS workplace/Thesis/Sentinel_2a/LC81800312015249LGN00/StackedLayers/R_Results_LSMS2")
#
##### Adım 1: Raster görüntünün belirlenmesi #####
#
# Görüntü ya da Fotoğrafın Yüklenmesi ve RStudio ile görüntülenmesi (raster)

raster <- brick("Ortho_15cm.tif")

raster
plot(raster)
plotRGB(raster, 4,1,2, stretch='lin')

crop_255_resample <- aggregate(crop_255, fact=5, fun=mean, expand=TRUE, na.rm=TRUE)

crop_255_resample
plotRGB(crop_255_resample, 4,1,2)

writeRaster(crop_255_resample,filename=".../Belgradormanj_255_resample.tif",format="Griff",datatype="FLT4S",overwrite=TRUE)

ras <- brick("D:/GIS workplace/Thesis/Landsat/LC81800312015249LGN00/StackedLayers/R_Results_LSMS2/LS7_10bands_PS_255.tif")
seg <- raster("D:/GIS workplace/Thesis/Landsat/LC81800312015249LGN00/StackedLayers/R_Results_LSMS2/LS7_LSMS_Rasterized.tif")
shape <- shapefile("D:/GIS workplace/Thesis/Landsat/LC81800312015249LGN00/StackedLayers/R_Results_LSMS2/LS7_LSMS_Segmenated.dbf")

names(ras) <- c("blue","green","red","nir","swir1","swir2","savi","ndvi","msavi","evi")

stat_mean <- zonal(ras, seg, mean)
colnames(stat_mean) <- c("zone","mean.blue","mean.green","mean.red","mean.nir","mean.swir1","mean.swir2","mean.savi","mean.ndvi","mean.msavi","mean.evi")
stat_sd <- zonal(ras, seg, sd)
colnames(stat_sd) <- c("zone","sd.blue","sd.green","sd.red","mean.nir","sd.swir1","sd.swir2","sd.savi","sd.ndvi","sd.msavi","sd.evi")
colnames(stat_min) <- zonal(ras, seg, min)
colnames(stat_min) <- c("zone","min.blue","min.green","min.red","min.nir","min.swir1","min.swir2","min.savi","min.ndvi","min.msavi","min.evi")
stat_max <- zonal(ras, seg, max)
colnames(stat_max) <- c("zone","max.blue","max.green","max.red","max.nir","max.swir1","max.swir2","max.savi","max.ndvi","max.msavi","max.evi")

head(stat_mean)

extr_stat <- cbind(stat_mean[,-1], stat_sd[,-1], stat_min[,-1], stat_max[,-1])

head(extr_stat)

save(extr_stat, file="D:/GIS workplace/Thesis/Landsat/LC81800312015249LGN00/StackedLayers/R_Results_LSMS2/LS7_LSMS_stat_seg.Rdata")

require(foreign)

out_dir_main <- "D:/GIS workplace/Thesis/Landsat/LC81800312015249LGN00/StackedLayers/R_Results_LSMS2/"
out_dir_models <- "D:/GIS workplace/Thesis/Landsat/LC81800312015249LGN00/StackedLayers/R_Results_LSMS2/Models/"

load(file="D:/GIS workplace/Thesis/Landsat/LC81800312015249LGN00/StackedLayers/R_Results_LSMS2/LS7_LSMS_stat_seg.Rdata")

extr_stat <- extr_stat

ref.file.dbf.name <- "D:/GIS workplace/Thesis/Landsat/LC81800312015249LGN00/StackedLayers/R_Results_LSMS2/LS7_LSMS_Segmenated.dbf"

ref.file.dbf <- read.dbf(file=ref.file.dbf.name)

extr_stat <- cbind(ref.file.dbf,extr_stat)

summary(extr_stat[,~c(1:3)])

```

ÖZGEÇMİŞ

Kişisel Bilgiler	
Adı Soyadı	Hüseyin DOĞAN
Doğum Yeri	Ankara
Doğum Tarihi	27.02.1989
Uyruğu	☒ T.C. ☐ Diğer:
Telefon	+90 505 749 59 58
E-Posta Adresi	ozgurhsyndgn@gmail.com
Web Adresi	



Eğitim Bilgileri	
Lisans	
Üniversite	Bartın Üniversitesi
Fakülte	Orman Fakültesi
Bölümü	Orman Mühendisliği
Mezuniyet Yılı	2012

Yüksek Lisans	
Üniversite	İstanbul Üniversitesi
Enstitü Adı	Fen Bilimleri Enstitüsü
Anabilim Dalı	Orman Mühendisliği
Programı	Orman Mühendisliği
Mezuniyet Tarihi	2017