

T.C.  
YILDIZ TEKNİK ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

MAKİNE ÖĞRENME ALGORİTMALARI KULLANILARAK LAZER  
VERİLERİNDEN AĞAÇ TÜRLERİNİN SINIFLANDIRILMASI  
OLANAKLARININ ARAŞTIRILMASI

Zehra ÇETİN

DOKTORA TEZİ

Harita Mühendisliği Anabilim Dalı  
Uzaktan Algılama ve CBS Programı

Danışman  
Prof. Dr. Naci YASTIKLI

Haziran, 2022

T.C.  
YILDIZ TEKNİK ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

**MAKİNE ÖĞRENME ALGORİTMALARI KULLANILARAK LAZER  
VERİLERİNDEN AĞAÇ TÜRLERİNİN SINIFLANDIRILMASI  
OLANAKLARININ ARAŞTIRILMASI**

Zehra ÇETİN tarafından hazırlanan tez çalışması 23.06.2022 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Harita Mühendisliği Anabilim Dalı, Uzaktan Algılama ve CBS Programı **DOKTORA TEZİ** olarak kabul edilmiştir.

Prof. Dr. Naci YASTIKLI  
Yıldız Teknik Üniversitesi  
Danışman

**Jüri Üyeleri**

Prof. Dr. Naci YASTIKLI, Danışman  
Yıldız Teknik Üniversitesi

Prof. Dr. Füsun BALIK ŞANLI, Üye  
Yıldız Teknik Üniversitesi

Prof. Dr. Dursun Zafer ŞEKER, Üye  
İstanbul Teknik Üniversitesi

Dr. Öğr. Üyesi Uğur ACAR, Üye  
Yıldız Teknik Üniversitesi

Prof. Dr. Umut Güneş SEFERCİK, Üye  
Gebze Teknik Üniversitesi

Danışmanım Prof. Dr. Naci YASTIKLI sorumluluğunda tarafımda hazırlanan Makine Öğrenme Algoritmaları Kullanılarak Lazer Verilerinden Ağaç Türlerinin Sınıflandırılması Olanaklarının Araştırılması başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Zehra ÇETİN

İmza



*Sevgili eřim*  
*ve*  
*biricik kızıma*

## TEŞEKKÜR

---

Doktora tez danışmanlığımı üstlenerek çalışmalarım süresince yardımlarını esirgemeyen, bilgi ve tecrübesi ile bana her zaman yol gösteren, beni cesaretlendiren çok değerli hocam Prof. Dr. Naci YASTIKLI'ya teşekkürlerimi arz ederim.

Doktora tez çalışmamın izlenmesi sırasında bilimsel desteklerini esirgemeyen, pozitif düşünceleri ile beni motive eden sayın hocalarım Prof. Dr. Fusün Balık ŞANLI ve Prof. Dr. Dursun Zafer ŞEKER'e teşekkürlerimi sunarım.

Tez çalışmam süresince her zaman yanımda olan, beni daima motive eden ve kendime inanmamı sağlayan oda arkadaşım Arş. Gör. Güldane OKU TOPAL'a çok teşekkür ederim. Ayrıca, bu süreçte bana yardım ve hoşgörülerini hiçbir zaman eksik etmeyen başta arkadaşım Arş. Gör. Özge GÜNEŞ ve sayın hocam Dr. Öğr. Üyesi Deniz ÖZ DEMİR olmak üzere tüm hocalarım ve arkadaşlarıma teşekkür ederim.

Bu çalışmada kullanılan veri setini sağladıkları için İstanbul Büyükşehir Belediyesi'ne teşekkürlerimi bildirmeyi borç bilirim.

Bugünlere gelmemde büyük emek ve fedakarlık gösteren başta dile getirmesem de hep yanımda hissettiğim canım babam Yılmaz ERİŞİR ve canım annem Fatma ERİŞİR olmak üzere çok kıymetli ailem ve eşimin değerli ailesine en içten sevgi ve sonsuz teşekkürlerimi sunarım.

Bu tez çalışmasının tamamlanmasında çok büyük emeği olan, hayattaki en büyük şansım ve destekçim, yol arkadaşım, sevgili eşim Erkut ÇETİN'e göstermiş olduğu sabır ve anlayıştan dolayı sonsuz teşekkür ve minnetlerimi sunarım. Ayrıca doktora çalışmam süresince, her zaman en büyük moral ve motivasyon kaynağım olan, hayatıma renk katan biricik kızım İpek ÇETİN'e çok teşekkür ederim.

Zehra ÇETİN

# İÇİNDEKİLER

<b>SİMGE LİSTESİ</b>	<b>vii</b>
<b>KISALTMA LİSTESİ</b>	<b>viii</b>
<b>ŞEKİL LİSTESİ</b>	<b>x</b>
<b>TABLO LİSTESİ</b>	<b>xv</b>
<b>ÖZET</b>	<b>xviii</b>
<b>ABSTRACT</b>	<b>xx</b>
<b>1 GİRİŞ</b>	<b>1</b>
1.1 Literatür Özeti.....	6
1.2 Tezin Amacı .....	21
1.3 Hipotez .....	23
<b>2 SINIFLANDIRMA</b>	<b>24</b>
2.1 Sınıflandırma Yöntemleri .....	24
2.2 LiDAR Nokta Bulutunun Sınıflandırılması .....	25
2.2.1 Nokta Bulutu Segmentasyonu.....	27
2.2.2 Nokta Bulutu Sınıflandırması.....	28
2.3 LiDAR Nokta Bulutu Verilerinden Ağaç Türlerinin Nokta Tabanlı Sınıflandırılması.....	32
2.4 Sınıflandırmanın Doğruluk Analizi.....	33
<b>3 MAKİNE ÖĞRENMESİ</b>	<b>36</b>
3.1 Makine Öğrenmesinin Temelleri .....	36
3.2 Makine Öğrenmesi Yöntemleri .....	38
3.2.1 Denetimsiz Öğrenme.....	38
3.2.2 Denetimli Öğrenme.....	39
3.2.3 Pekiştirmeli Öğrenme.....	48
3.3 K-Katlamalı Çapraz Doğrulama (K-Fold Cross Validation).....	48
<b>4 AĞAÇ TÜRLERİNİN SINIFLANDIRILMASI</b>	<b>50</b>
4.1 Çalışma Alanı .....	51
4.2 LiDAR Verisi ve Referans Veri .....	54
4.2.1 LiDAR Verisi.....	54
4.2.2 Referans Veri.....	58
4.3 Genel İşlem Adımları.....	63
4.4 Sınıflandırma ile Yüksek Bitki Örtüsü Sınıfının Elde Edilmesi .....	64

4.5 Bireysel Ağaç Tacı Segmentasyonu.....	71
4.5.1 “MeanShift” Kümeleme ile Bireysel Ağaç Taçlarının Elde Edilmesi.72	
4.5.2 “K-means” Kümeleme ile Bireysel Ağaç Taçlarının Elde Edilmesi...78	
4.5.3 “DBSCAN” Kümeleme ile Bireysel Ağaç Taçlarının Elde Edilmesi...84	
4.5.4 Kentsel Ağaç Türlerinin Sınıflandırılmasında Kullanılacak Segmentasyonun Belirlenmesi.....90	
4.6 Özellik Çıkarımı .....	92
4.7 Ağaç Türlerinin Makine Öğrenme Algoritmaları ile Sınıflandırılması .....	98
4.7.1 Destek Vektör Makineleri Sınıflandırması.....101	
4.7.2 Rastgele Orman Sınıflandırması.....112	
4.7.3 Çok Katmanlı Algılayıcı Sınıflandırması.....128	
4.7.4 Makine Öğrenme Tabanlı Sınıflandırma Algoritmalarının Performanslarının Karşılaştırılması.....139	
4.7.5 Sınıflandırma Özelliklerinin Önemlerinin Belirlenmesi.....152	
4.8 Tekil Cadde Ağaçlarının Makine Öğrenme Tabanlı Segmentasyon Yöntemleri ile Otomatik Belirlenmesi.....	161
4.8.1 Test Alanı ve Genel İşlem Adımları.....162	
4.8.2 Yüksek Bitki Örtüsünün Nokta Tabanlı Sınıflandırılması.....163	
4.8.3 Tekil Cadde Ağaçlarının Nokta Tabanlı Segmentasyonu ve Doğruluk Analizi.....164	
<b>5 SONUÇ VE ÖNERİLER</b>	<b>169</b>
<b>KAYNAKÇA</b>	<b>179</b>
<b>TEZDEN ÜRETİLMİŞ YAYINLAR</b>	<b>194</b>

## SİMGE LİSTESİ

---

$w_1, w_2, w_3$	Ağırlıklar
$y$	Çıktı
$k$	Eğitim Verisi Sayısı
$x_1, x_2, x_3$	Girdiler
$\mathcal{H}$	Hiperdüzlem
$X, Y, Z$	Kartezyen Koordinatlar
$R_i$	Sağ Taraftaki $i$ Sınıfında Bulunan Örnek Sayısı
$D_{sağ}$	Sağda Belirlenen Eşik Değerden Büyük Değere Sahip Örneklerin Sayısı
$b$	Sapma
$n$	Sınıf Sayısı
$L_i$	Sol Taraftaki $i$ Sınıfında Bulunan Örnek Sayısı
$D_{sol}$	Solda Belirlenen Eşik Değerden Küçük Değere Sahip Örneklerin Sayısı

## KISALTMA LİSTESİ

---

AGFLVQ	Adaptive Gaussian Fuzzy Learning Vector Quantization
AI	Artificial Intelligence
ALS	Airborne Laser Scanning, Airborne LiDAR System
ANN	Artificial Neural Network
API	Application Programming Interface
ASCII	Amerikan Standard Code for Information Interchange
ASPRS	American Society for Photogrammetry and Remote Sensing
AVIRIS	Airborne Visible/Infrared Imaging Spectrometer
CAD	Computer Aided Desing
CHM	Canopy Height Model
ÇKA	Çok Katmanlı Algılayıcı
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DCM	Digital Canopy Model
DSM	Digital Surface Model
DTM	Digital Terrain Model
DVM	Destek Vektör Makineleri
EM	Expectation Maximum
EVI	Enhance Vegetation Index
FN	False Negative
FP	False Positive
GD	Gap Distribution
ISODATA	Iterative Self-Organizing Data Analysis Technique
İBB	İstanbul Büyükşehir Belediyesi
k-MSN	k-most similar neighbor
k-NN	k-Nearest Neighbor
LAS	Log ASCII Standard
LDA	Lineer Diskriminant Analizi
LiDAR	Light Detection and Ranging
MCC	Multi Class Classifier
MDA	Mean Decrease in Accuracy
MDG	Mean Decrease in Gini

ML	Machine Learning
MLC	Maximum Likelihood Classification
MNF	Minimum Noise Fraction Transformation
NDVI	Normalised Difference Vegetation Index
nDSM	Normalised Digital Surface Model
OBIA	Object Based Image Analysis
OOB	Out-of-Bag
PCA	Principal Component Analysis
RCS	Relative Clustering Scale
RBF	Radial Basis Function
RF	Random Forest
RO	Rastgele Orman
SVM	Support Vector Machine
SYM	Sayısal Yükseklik Modeli
TCO	Tree Crown Object
TN	True Negative
TP	True Positive
YM	Sayısal Yüzey Modeli
2B	2 Boyutlu
3B	3 Boyutlu

## ŞEKİL LİSTESİ

Şekil 2.1	Değişken dağılımındaki simetri derecesi (eğrilik katsayısı-skewness) (a) ve bir dağılımın zirve yaptığı ölçü (sivrilik katsayısı-kurtosis)(b).....33
Şekil 3.1	Yapay zekâ, makine öğrenmesi ve derin öğrenme ilişkisi [117]..38
Şekil 3.2	Denetimli öğrenme (a) ve denetimsiz öğrenme(b) [122].....40
Şekil 3.3	Destek vektör makineleri [140] (a) ve en uygun hiper düzlemin belirlenmesi [138] (b).....41
Şekil 3.4	Bazı çekirdek eşitlikleri [139].....42
Şekil 3.5	RO algoritması [146] (a) ve bir karar ağacı örneği [147] (b).....42
Şekil 3.6	Rastgele orman algoritması için veri setlerinin oluşturulması [151].....43
Şekil 3.7	Rastgele orman yönetiminde OBB işleyiş süreci [152].....44
Şekil 3.8	Üç katmanlı ileri beslemeli bir sinir ağı modeli [165].....47
Şekil 3.9	Yapay bir nöron modeli [161].....47
Şekil 3.10	10-katlamalı çapraz doğrulama.....49
Şekil 4.1	Yıldız Teknik Üniversitesi Davutpaşa Kampüsü (Google Earth)...51
Şekil 4.2	Test alanı 1 (a) ve Test alanı 2 (b) (Hava fotoğrafı – 2013).....52
Şekil 4.3	Test alanlarında yer alan kent ağaçlarının konumları (Google Earth – 2013) (a) ve bu ağaçlara ilişkin fotoğraflar (b).....53
Şekil 4.4	Test alanı 1'in yüksekliğe göre renklendirilmiş nokta bulutu verisi (a) ve nokta bulutunun perspektif görünümü (b).....56
Şekil 4.5	Test alanı 2'nin yüksekliğe göre renklendirilmiş nokta bulutu verisi (a) ve nokta bulutunun perspektif görünümü (b).....57
Şekil 4.6	Geliştirilen Şehir Haritası tabanlı uygulamanın veri toplama ekranı.....58
Şekil 4.7	Geliştirilen Şehir Haritası tabanlı uygulamada ekran üzerinden işaretlenen her bir ağaç için tür bilgisinin seçildiği veri toplama penceresi.....59
Şekil 4.8	Renkli solar harita üzerinde toplanan ağaç verisinin görünümü (698 ağaç) (a) ve renkli solar harita üzerinde toplanan ağaç verisinin yakın görünümü (sarı dikdörtgen) (b).....60
Şekil 4.9	Test alanı 1'in referans verisi (a) ve Test alanı 2'nin referans verisi (b).....61
Şekil 4.10	Çalışma bölgesindeki geniş yapraklı (a) ve iğne yapraklı (b) ağaç örnekleri.....62

<b>Şekil 4.11</b>	Geniş yapraklı ağaç (a), geniş yapraklı ağacın LiDAR verisinde tepeden görünümü (b) ve yandan görünümü (c); iğne yapraklı ağaç (d), iğne yapraklı ağacın LiDAR verisinde tepeden görünümü (e) ve yandan görünümü (f).....62
<b>Şekil 4.12</b>	Önerilen makine öğrenme tabanlı ağaç türleri sınıflandırması yaklaşımının genel işlem adımları.....64
<b>Şekil 4.13</b>	Yüksek bitki örtüsü sınıfının elde edilmesi amacıyla önerilen yaklaşımın genel işlem adımları [72].....65
<b>Şekil 4.14</b>	Test alanı 1'in hiyerarşik kural seti ile sınıflandırma sonucu (a) ve sınıflandırma sonucunun perspektif görünümü (b).....67
<b>Şekil 4.15</b>	Test alanı 2'nin hiyerarşik kural seti ile sınıflandırma sonucu (a) ve sınıflandırma sonucunun perspektif görünümü (b).....68
<b>Şekil 4.16</b>	Test alanı 1'de elde edilen yüksek bitki örtüsü noktaları (a) ve Test alanı 2'de elde edilen yüksek bitki örtüsü noktaları (b).....70
<b>Şekil 4.17</b>	Hiyerarşik kural tabanlı sınıflandırma ile edilen yüksek bitki örtüsü noktaları (a), 1 m'lik tampon bölge ile elde edilen hatalı bina noktaları (b) ve hatalı bina noktalarının çıkarıldığı son yüksek bitki örtüsü noktaları (c).....72
<b>Şekil 4.18</b>	Tüm çalışma alanında "MeanShift" kümeleme sonucunun grafiksel gösterimi (küme merkezlerini içeren).....73
<b>Şekil 4.19</b>	Tüm çalışma alanında "MeanShift" kümeleme sonucu: tepeden görünüm (a), perspektif görünüm (b) ve yakın görünüm (sarı dikdörtgen) (c).....74
<b>Şekil 4.20</b>	"MeanShift" kümelemede her bir kümedeki en yüksek nokta (a) ve en yüksek nokta ve tüm yüksek bitki örtüsü sınıfı noktalarının birlikte görünümü (b).....75
<b>Şekil 4.21</b>	"MeanShift" kümeleme sonucunun referans veri ile karşılaştırılması.....76
<b>Şekil 4.22</b>	Sıralı ağaçlarda "MeanShift" kümeleme sonucu örnekleri.....77
<b>Şekil 4.23</b>	Tek ağaçlarda "MeanShift" kümeleme sonucu örnekleri.....77
<b>Şekil 4.24</b>	Karışık taç yapılı ağaçlarda "MeanShift" kümeleme sonucu örnekleri.....78
<b>Şekil 4.25</b>	Tüm çalışma alanında "K-means" kümeleme sonucunun grafiksel gösterimi (küme merkezlerini içeren).....79
<b>Şekil 4.26</b>	Tüm çalışma alanında "K-means" kümeleme sonucu: tepeden görünüm (a), perspektif görünüm (b) ve yakın görünüm (sarı dikdörtgen) (c).....80
<b>Şekil 4.27</b>	"K-means" kümelemede her bir kümedeki en yüksek nokta (a) ve en yüksek nokta ve tüm yüksek bitki örtüsü sınıfı noktalarının birlikte görünümü (b).....81
<b>Şekil 4.28</b>	"K-means" kümeleme sonucunun referans veri ile karşılaştırılması..82

Şekil 4.29	Sıralı ağaçlarda “K-means” kümeleme sonucu örnekleri.....	83
Şekil 4.30	Tek ağaçlarda “K-means” kümeleme sonucu örnekleri.....	83
Şekil 4.31	Karışık taç yapılı ağaçlarda “K-means” kümeleme sonucu örnekleri.....	84
Şekil 4.32	Tüm çalışma alanında “DBSCAN” kümeleme sonucunun grafiksel gösterimi (küme merkezlerini içeren).....	85
Şekil 4.33	Tüm çalışma alanında “DBSCAN” kümeleme sonucu: tepeden görünüm (a), perspektif görünüm (b) ve yakın görünüm (sarı dikdörtgen) (c).....	86
Şekil 4.34	“DBSCAN” kümelemede her bir kümedeki en yüksek nokta (a) ve en yüksek nokta ve tüm yüksek bitki örtüsü sınıfı noktalarının birlikte görünümü (b).....	87
Şekil 4.35	“DBSCAN” kümeleme sonucunun referans veri ile karşılaştırılması..	88
Şekil 4.36	Sıralı ağaçlarda “DBSCAN” kümeleme sonucu örnekleri.....	89
Şekil 4.37	Tek ağaçlarda “DBSCAN” kümeleme sonucu örnekleri.....	89
Şekil 4.38	Karışık taç yapılı ağaçlarda “DBSCAN” kümeleme sonucu örnekleri.....	90
Şekil 4.39	Test alanı 1’de yer alan doğru segmente edilmiş bireysel ağaç taçları (difüzyon harita ile karşılaştırılmış).....	91
Şekil 4.40	Test alanı 2’de yer alan doğru segmente edilmiş bireysel ağaç taçları (difüzyon harita ile karşılaştırılmış).....	92
Şekil 4.41	Test alanı 1’de rastgele belirlenen eğitim (a) ve test (b) verileri.....	99
Şekil 4.42	Test alanı 2’de rastgele belirlenen eğitim (a) ve test (b) verileri.....	100
Şekil 4.43	Test alanı 1’in 8 özellik kullanılarak DVM sınıflandırması.....	102
Şekil 4.44	Test alanı 1’in 25 özellik kullanılarak DVM sınıflandırması.....	105
Şekil 4.45	Test alanı 2’nin 8 özellik kullanılarak DVM sınıflandırması.....	107
Şekil 4.46	Test alanı 2’nin 25 özellik kullanılarak DVM sınıflandırması.....	110
Şekil 4.47	Test alanı 1’in 8 özelliğin kullanıldığı RO’da 1. karar ağacı (a) ve 68. karar ağacı (b).....	114
Şekil 4.48	Test alanı 1’in 8 özellik kullanılarak RO sınıflandırması.....	115
Şekil 4.49	Test alanı 1’de 25 özelliğin kullanıldığı RO’da 83. karar ağacı (a) ve 103. karar ağacı (b).....	118
Şekil 4.50	Test alanı 1’in 25 özellik kullanılarak RO sınıflandırması.....	119
Şekil 4.51	Test alanı 2’de 8 özelliğin kullanıldığı RO’da 127. karar ağacı...	121
Şekil 4.52	Test alanı 2’nin 8 özellik kullanılarak RO sınıflandırması.....	122
Şekil 4.53	Test alanı 2’de 25 özelliğin kullanıldığı RO’da 72. karar ağacı...	125

<b>Şekil 4.54</b>	Test alanı 2'nin 25 özellik kullanılarak RO sınıflandırması.....	126
<b>Şekil 4.55</b>	ÇKA ağ yapısı (8 sınıflandırma özelliği kullanılan).....	129
<b>Şekil 4.56</b>	Test alanı 1'in 8 özellik kullanılarak ÇKA sınıflandırması.....	130
<b>Şekil 4.57</b>	ÇKA ağ yapısı (25 sınıflandırma özelliği kullanılan).....	132
<b>Şekil 4.58</b>	Test alanı 1'in 25 özellik kullanılarak ÇKA sınıflandırması.....	132
<b>Şekil 4.59</b>	Test alanı 2'nin 8 özellik kullanılarak ÇKA sınıflandırması.....	135
<b>Şekil 4.60</b>	Test alanı 2'nin 25 özellik kullanılarak ÇKA sınıflandırması.....	137
<b>Şekil 4.61</b>	Test alanı 1'in 8 özellik kullanılarak gerçekleştirilen DVM (a), RO (b) ve ÇKA (c) sınıflandırmaları.....	141
<b>Şekil 4.62</b>	Test alanı 1'in 25 özellik kullanılarak gerçekleştirilen DVM (a), RO (b) ve ÇKA (c) sınıflandırmaları.....	141
<b>Şekil 4.63</b>	Test alanı 1'de her bir geniş ve iğne yapraklı ağaç için 8 özellik kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırıcılarının hassasiyet, kesinlik ve F1-skor değerleri.....	142
<b>Şekil 4.64</b>	Test alanı 1'de her bir geniş ve iğne yapraklı ağaç için 25 özellik kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırıcılarının hassasiyet, kesinlik ve F1-skor değerleri.....	144
<b>Şekil 4.65</b>	Test alanı 2'nin 8 özellik kullanılarak gerçekleştirilen DVM (a), RO (b) ve ÇKA (c) sınıflandırmaları.....	147
<b>Şekil 4.66</b>	Test alanı 2'in 25 özellik kullanılarak gerçekleştirilen DVM (a), RO (b) ve ÇKA (c) sınıflandırmaları.....	148
<b>Şekil 4.67</b>	Test alanı 2'de her bir geniş ve iğne yapraklı ağaç için 8 özellik kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırıcılarının hassasiyet, kesinlik ve F1-skor değerleri.....	150
<b>Şekil 4.68</b>	Test alanı 2'de her bir geniş ve iğne yapraklı ağaç için 25 özellik kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırıcılarının hassasiyet, kesinlik ve F1-skor değerleri.....	151
<b>Şekil 4.69</b>	Test alanı 1'de RO sınıflandırmasındaki ortalama Gini'deki azalmaya göre 8 sınıflandırma özelliğinin önemleri (feature importance scores).....	153
<b>Şekil 4.70</b>	Test alanı 1'de RO sınıflandırmasındaki ortalama Gini'deki azalmaya göre 25 sınıflandırma özelliğinin önemleri (feature importance scores).....	154
<b>Şekil 4.71</b>	Test alanı 1'de 8 sınıflandırma özelliğinin Pearson korelasyon matrisi.....	155
<b>Şekil 4.72</b>	Test alanı 1'de 25 sınıflandırma özelliğinin Pearson korelasyon matrisi.....	156
<b>Şekil 4.73</b>	Test alanı 2'de RO sınıflandırmasındaki ortalama Gini'deki azalmaya göre 8 sınıflandırma özelliğinin önemleri (feature importance scores).....	157

<b>Şekil 4.74</b>	Test alanı 2’de RO sınıflandırmasındaki ortalama Gini’deki azalmaya göre 25 sınıflandırma özelliğinin önemleri (feature importance scores).....158
<b>Şekil 4.75</b>	Test alanı 2’de 8 sınıflandırma özelliğinin Pearson korelasyon matrisi.....159
<b>Şekil 4.76</b>	Test alanı 2’de 25 sınıflandırma özelliğinin Pearson korelasyon matrisi.....160
<b>Şekil 4.77</b>	Test alanı A: Şehir haritası (2013) (a), cadde ağaçları (b) ve yoğunluğa göre renklendirilmiş ham LiDAR nokta bulutu (c)...162
<b>Şekil 4.78</b>	Test alanı B: Şehir haritası (2013) (a), cadde ağaçları (b) ve yoğunluğa göre renklendirilmiş ham LiDAR nokta bulutu (c)...162
<b>Şekil 4.79</b>	Tekil cadde ağaçlarının elde edilebilmesi için belirlenen iş akış şeması.....163
<b>Şekil 4.80</b>	Test alanı A’nın nokta tabanlı sınıflandırılması (a) ve Test alanı A’da bulunan yüksek bitki örtüsü noktaları (b).....163
<b>Şekil 4.81</b>	Test alanı B’nin nokta tabanlı sınıflandırılması (a) ve Test alanı B’de bulunan yüksek bitki örtüsü noktaları (b).....164
<b>Şekil 4.82</b>	Test alanı A’nın yüksekliğe göre renklendirilmiş ham nokta bulutu (a), “MeanShift” kümeleme ile segmente edilmiş cadde ağaçları (b), “MeanShift” kümeleme ile segmente edilmiş cadde ağaçlarının gri tonlu DSM ile üst üste bindirilmesi (c), “DBSCAN” kümeleme ile segmente edilmiş cadde ağaçları (d), “DBSCAN” kümeleme ile segmente edilmiş cadde ağaçlarının gri tonlu DSM ile üst üste bindirilmesi (e).....165
<b>Şekil 4.83</b>	Test alanı B’nin yüksekliğe göre renklendirilmiş ham nokta bulutu (a), “MeanShift” kümeleme ile segmente edilmiş cadde ağaçları (b), “MeanShift” kümeleme ile segmente edilmiş cadde ağaçlarının gri tonlu DSM ile üst üste bindirilmesi (c), “DBSCAN” kümeleme ile segmente edilmiş cadde ağaçları (d), “DBSCAN” kümeleme ile segmente edilmiş cadde ağaçlarının gri tonlu DSM ile üst üste bindirilmesi (e).....166
<b>Şekil 4.84</b>	Test alanı A’da referans ağaçlar (a) “MeanShift” kümeleme ile elde edilmiş TP, FP, FN cadde ağaçları (b) ve “DBSCAN” kümeleme ile elde edilmiş TP, FP, FN cadde ağaçları (c).....167
<b>Şekil 4.85</b>	Test alanı B’de referans ağaçlar (a) “MeanShift” kümeleme ile elde edilmiş TP, FP, FN cadde ağaçları (b) ve “DBSCAN” kümeleme ile elde edilmiş TP, FP, FN cadde ağaçları (c).....167

## TABLO LİSTESİ

<b>Tablo 2.1</b>	ASPRS standart nokta sınıfları [81].....	26
<b>Tablo 2.2</b>	İkili değişkenli hata matrisi [112].....	34
<b>Tablo 4.1</b>	Test alanlarında bulunan geniş ve iğne yapraklı ağaç türleri.....	54
<b>Tablo 4.2</b>	“RIEGL LMS-Q680i” lazer tarayıcı özellikleri [72].....	55
<b>Tablo 4.3</b>	Yüksek bitki örtüsünün sınıflandırılması için oluşturulan kural seti ve oluşun sınıflar.....	66
<b>Tablo 4.4</b>	8 elemanlı özellik seti.....	93
<b>Tablo 4.5</b>	25 elemanlı özellik seti.....	93
<b>Tablo 4.6</b>	Test alanı 1’de yer alan her bir bireysel ağaç tacı segmentlerinin hesaplanan 8 özelliği.....	95
<b>Tablo 4.7</b>	Test alanı 1’de yer alan her bir bireysel ağaç tacı segmentlerinin hesaplanan 25 özelliği.....	96
<b>Tablo 4.8</b>	Test alanı 2’de her bir bireysel ağaç tacı segmentlerinin hesaplanan 8 özelliği.....	97
<b>Tablo 4.9</b>	Test alanı 2’de yer alan her bir bireysel ağaç tacı segmentlerinin hesaplanan 25 özelliği.....	97
<b>Tablo 4.10</b>	Test alanlarının eğitim ve test verileri içerisinde bulunan ağaç türlerinin dağılımı.....	101
<b>Tablo 4.11</b>	DVM için “Grid search” yaklaşımında kullanılan parametreler....	102
<b>Tablo 4.12</b>	Test alanı 1’in 8 özellik kullanılarak DVM sınıflandırmasına ilişkin hata matrisi.....	103
<b>Tablo 4.13</b>	Test alanı 1’in 8 özellik kullanılarak DVM sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları.....	104
<b>Tablo 4.14</b>	Test alanı 1’in 25 özellik kullanılarak DVM sınıflandırmasına ilişkin hata matrisi.....	105
<b>Tablo 4.15</b>	Test alanı 1’in 25 özellik kullanılarak DVM sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları.....	106
<b>Tablo 4.16</b>	Test alanı 2’nin 8 özellik kullanılarak DVM sınıflandırmasına ilişkin hata matrisi.....	108
<b>Tablo 4.17</b>	Test alanı 2’nin 8 özellik kullanılarak DVM sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları.....	109
<b>Tablo 4.18</b>	Test alanı 2’nin 25 özellik kullanılarak DVM sınıflandırmasına ilişkin hata matrisi.....	110
<b>Tablo 4.19</b>	Test alanı 2’nin 25 özellik kullanılarak DVM sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları.....	112

<b>Tablo 4.20</b>	RO için “Grid search” yaklaşımında kullanılan parametreler.....	113
<b>Tablo 4.21</b>	Test alanı 1’in 8 özellik kullanılarak RO sınıflandırmasına ilişkin hata matrisi.....	115
<b>Tablo 4.22</b>	Test alanı 1’in 8 özellik kullanılarak RO sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları.....	116
<b>Tablo 4.23</b>	Test alanı 1’in 25 özellik kullanılarak RO sınıflandırmasına ilişkin hata matrisi.....	119
<b>Tablo 4.24</b>	Test alanı 1’in 25 özellik kullanılarak RO sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları.....	120
<b>Tablo 4.25</b>	Test alanı 2’nin 8 özellik kullanılarak RO sınıflandırmasına ilişkin hata matrisi.....	122
<b>Tablo 4.26</b>	Test alanı 2’nin 8 özellik kullanılarak RO sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları.....	124
<b>Tablo 4.27</b>	Test alanı 2’nin 25 özellik kullanılarak RO sınıflandırmasına ilişkin hata matrisi.....	126
<b>Tablo 4.28</b>	Test alanı 2’nin 25 özellik kullanılarak RO sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları.....	128
<b>Tablo 4.29</b>	ÇKA için “Grid search” yaklaşımında kullanılan parametreler.....	129
<b>Tablo 4.30</b>	Test alanı 1’in 8 özellik kullanılarak ÇKA sınıflandırmasına ilişkin hata matrisi.....	130
<b>Tablo 4.31</b>	Test alanı 1’in 8 özellik kullanılarak ÇKA sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları.....	131
<b>Tablo 4.32</b>	Test alanı 1’in 25 özellik kullanılarak ÇKA sınıflandırmasına ilişkin hata matrisi.....	132
<b>Tablo 4.33</b>	Test alanı 1’in 25 özellik kullanılarak ÇKA sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları.....	134
<b>Tablo 4.34</b>	Test alanı 2’nin 8 özellik kullanılarak ÇKA sınıflandırmasına ilişkin hata matrisi.....	135
<b>Tablo 4.35</b>	Test alanı 2’nin 8 özellik kullanılarak ÇKA sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları.....	137
<b>Tablo 4.36</b>	Test alanı 2’nin 25 özellik kullanılarak ÇKA sınıflandırmasına ilişkin hata matrisi.....	138
<b>Tablo 4.37</b>	Test alanı 2’nin 25 özellik kullanılarak ÇKA sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları.....	139
<b>Tablo 4.38</b>	Test alanı 1’de 8 ve 25 sınıflandırma özelliği kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırma sonuçları.....	141
<b>Tablo 4.39</b>	Test alanı 1’de 8 ve 25 sınıflandırma özelliği kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırmalarının 10 katlamalı çapraz doğrulamaya göre ortalama doğrulukları.....	145

<b>Tablo 4.40</b>	Test alanı 2’de 8 ve 25 sınıflandırma özelliği kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırma sonuçları.....	149
<b>Tablo 4.41</b>	Test alanı 2’de 8 ve 25 sınıflandırma özelliği kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırmalarının 10 katlamalı çapraz doğrulamaya göre ortalama doğrulukları.....	152
<b>Tablo 4.42</b>	Test alanı A ve test alanı B’de gerçekleştirilen “MeanShift” ve “DBSCAN” kümeleme algoritmalarının doğruluk değerleri.....	168



## Makine Öğrenme Algoritmaları Kullanılarak Lazer Verilerinden Ağaç Türlerinin Sınıflandırılması Olanaklarının Araştırılması

Zehra ÇETİN

Harita Mühendisliği Anabilim Dalı

Doktora Tezi

Danışman: Prof. Dr. Naci YASTIKLI

Kentsel çevrenin en önemli bileşenlerinden biri yeşil alanlardır ve yeşil alanların en baskın elemanı olan ağaçlar kent ekolojisinde önemli bir yere sahiptir. Ağaç türleri hakkında detaylı ve doğru bilgi, afet yönetimi, şehir planlaması, çevre koruma veya kentsel kalkınma politikası üretimi gibi çalışmalar için yerel yönetimler açısından büyük önem taşımaktadır. Bu çalışmanın amacı, makine öğrenme algoritmaları ile kentsel çalışma alanı olarak seçilen Yıldız Teknik Üniversitesi, Davutpaşa Kampüsü'ne ait 3B ham LiDAR nokta bulutu verileri kullanılarak geniş ve iğne yapraklı ağaç türlerinin nokta tabanlı yöntemlerle sınıflandırılması olanaklarının araştırılmasıdır. İlk olarak ham LiDAR verisinden hiyerarşik kural tabanlı sınıflandırma yaklaşımı ile belirlenen Test alanı 1 ve Test alanı 2'de yüksek bitki örtüsü sınıfı elde edilmiştir. Bireysel ağaç taçlarının elde edilmesi için üç farklı kümeleme algoritması ("MeanShift", "K-means" ve "Density-Based Spatial Clustering of Applications with Noise - DBSCAN") kullanılarak nokta tabanlı segmentasyon gerçekleştirilmiştir. Geometri ve yoğunluk (intensity) tabanlı özelliklerden oluşan iki farklı sınıflandırma özellik seti (8 ve 25 adet)

kullanılarak Destek Vektör Makineleri (DVM), Rastgele Orman (RO) ve Çok Katmanlı Algılayıcı (ÇKA) algoritmaları ile ağaç tacı segmentleri geniş ve iğne yapraklı ağaç türleri olarak sınıflandırılmıştır. Test alanı 1’de 70-30 bölümlenmede (eğitim-test) gerçekleştirilen makine öğrenme tabanlı DVM, RO ve ÇKA için genel sınıflandırma doğrulukları sırasıyla 8 özellik kullanılarak %76.25, %80.00 ve %67.50; 25 özellik kullanılarak %80, %83.75 ve %73.75 olarak elde edilmiştir. Test alanı 2’de aynı 70-30 bölümlenmede DVM, RO ve ÇKA için sınıflandırma doğrulukları sırasıyla 8 özellik kullanılarak %71.70, %69.81 ve %72.33; 25 özellik kullanılarak %72.33, %72.33 ve %46.54 olarak belirlenmiştir. Sınıflandırma doğruluklarına göre, kentsel çalışma alanında genel olarak ağaç türlerinin sınıflandırılması için DVM ve RO sınıflandırıcılarının ÇKA’ya göre daha başarılı oldukları görülmüştür. Kentsel ağaç türlerinin sınıflandırılmasının ardından RO ile ortalama Gini’deki azalmaya (Mean Decrease in Gini - MDG) göre sınıflandırmada kullanılan özellikler analiz edilmiş ve geometri tabanlı özelliklerin yoğunluk tabanlı özelliklerden önemli olduğu görülmüştür.

**Anahtar Kelimeler:** Sınıflandırma, makine öğrenmesi, destek vektör makineleri, rastgele orman, çok katmanlı algılayıcı

# Investigation of Tree Species Classification Possibilities From Laser Data Using Machine Learning Algorithms

Zehra CETIN

Department of Geomatics Engineering

Doctor of Philosophy Thesis

Supervisor: Prof. Dr. Naci YASTIKLI

One of the most important components of the urban environment is the green areas and, the trees, which are most dominant element of urban environment, have an important place in urban ecology. Detailed and accurate information about tree species is of great importance for all local governments for many studies such as disaster management, city planning, environmental protection or production of urban development policy. The aim of this study is to investigate the possibilities for deciduous and coniferous tree species classification with machine learning algorithms in the study area by point-based methods from 3D raw LiDAR point cloud data of Davutpaşa Campus of Yıldız Technical University, which was chosen as the urban study area. First, high vegetation class was obtained from raw LiDAR data in Test area 1 and Test area 2 using the hierarchical rule-based classification approach. To obtain individual tree crowns, point-based segmentation was performed using three different clustering algorithms (MeanShift, K-means, Density-Based Spatial Clustering of Applications with Noise - DBSCAN). Tree crown segments were classified into deciduous and coniferous tree species in urban area with Support Vector Machines (SVM), Random Forest

(RF), and Multi-Layer Perceptron (MLP) algorithms by using two different classification feature sets (8 and 25 features) consisting spatial- and intensity-based features. The overall classification accuracies for machine learning-based SVM, RF, and MLP algorithms performed in 70-30 partitioning (training-test) in Test area 1 were obtained as 76.25%, 80.00%, and 67.50% respectively using 8 features, and obtained as 80%, 83.75%, and 73.75% respectively using 25 features. In the same 70-30 partitioning in Test area 2, the classification accuracies for the SVM, RF, and MLP algorithms were obtained as 71.70%, 69.81%, and 72.33% respectively using 8 features, and obtained as 72.33%, 72.33%, and 46.54% using 25 features. According to the classification accuracies, it has been observed that SVM and RF classifiers are more successful in the urban study area than the MLP classifier in general for the classification of tree species. After the classification of urban tree species, the features used in the classification were analyzed according to the Mean Decrease in Gini (MDG) in RF, and it was seen that the spatial-based features were important than the intensity-based features.

**Keywords:** Classification, machine learning, support vector machine, random forest, multi-layer perceptron

Kentsel alanlar insanlığın ana yaşam alanlarından biridir ve artan kentleşme eğilimi ile birlikte, kentsel alanlarda yaşayan insan nüfusu da gün geçtikçe artış göstermektedir. Buna paralel olarak, kentsel çevreyi anlama ve yönetme ihtiyacı da insanlığın yaşamının doğası gereği çok önemli hale gelmiştir [1-2]. Kentsel çevrenin en önemli bileşenlerinden biri yeşil alanlardır ve bu alanlar sürdürülebilir kentlerin planlanmasında kilit unsurlardır [3]. Kentlerin yoğun ve yüksek yapıları arasında yer alabilen yeşil alanlar ve onların en baskın elamanı olan ağaçlar, kent ekolojisinde çok önemli bir yere sahiptir.

Bir kentsel alan içindeki tüm ağaçları ifade eden kentsel ormanlar; terleme, gölgeleme ve ısı emme yoluyla hava, su ve arazi kalitesini iyileştirme; iklim kontrolü; karbon birikimini azaltma; gürültü, toz, gaz ve rüzgar zararlarını önleme; toprak ve su dengesini koruma; enerji tasarrufu sağlamada önemli rol oynamaktadırlar [4-9]. Ayrıca, kent ağaçları sosyo-kültürel (anıt ağaç, endemik türler vb.), estetik, peyzaj, psikolojik ve maddi olmak üzere çok yönlü başka işlevlere de sahiptirler. Kent ağaçları, kentsel peyzajın temel varlıklarıdır ve kentlerdeki ağaç peyzajı tasarımında temel amaç; ağacın fonksiyonel özelliklerinden yararlanarak, olumsuz çevre koşullarını iyileştirmek ve ağacın estetik özelliklerini kullanarak yapısal kullanımını güçlendirmektir. Son on yıllarda yapılan araştırmalar, yağmur suyu akışını azaltma, temiz hava, daha iyi zihinsel ve fiziksel sağlık için destek, suçları azaltma, konut enerjisi maliyetlerini azaltma, yaşama ve çalışma için daha cazip yerler gibi kent ağaçlarının sağladığı birçok farklı faydayı ortaya koyma üzerine yoğunlaşmıştır [10].

Kentsel ağaçların sayısız olumlu etkilerinin yanında az sayıda olumsuz etkileri de bulunmaktadır. Bazı ağaç türleri, polenler sebebiyle alerjik reaksiyonlara yol açma [2], çevre kirliliği yaratma, fazla kökleşmeden dolayı kentsel yapılara ve tarihi dokuya zarar verme, fazla dallanma ve büyümeden dolayı kentlerin silüetini bozma gibi olumsuzluklara yol açabilmektedirler.

Kent ağalarının tm bu olumlu ve olumsuz etkilerinden dolayı, kentlerdeki ağa kltr yerel ynetimlerin bařlıca faaliyet konularından birini oluřturmaktadır. Ağa trleri hakkında detaylı ve doęru bir bilgi, afet ynetimi, Őehir planlaması, evre koruma veya kentsel kalkınma politikası retimi iin tm yerel ynetimler aısından byk nem tařımaktadır [11]. Ağalardan beklenen iřlevlerin yerine getirilebilmesi iin kent ii ve evresinde yapılmıř veya yapılacak olan ağalar ile ilgili alıřmaların teknięine uygun bir Őekilde gerekleřtirilmesi gerekmektedir. Bu nedenle, ağalandırma alıřmalarının planlanması ve ynetimi, ağaların korunması iin daha iyi bir kontroln saęlanması amacıyla ncelikle kent ağalarının mevcut durumunun iyi bilinmesi gerekmektedir [7].

Detaylı  boyutlu (3B) kent modeli retimi ve coęrafi bilgi sistemi uygulamaları gibi eřitli alıřmalar da doęru ve gncel ağa tr bilgisine ihtiya duymaktadır. Bu uygulamalar iin de zamanında ve hassas bir Őekilde kentsel alanlardaki ağa trlerinin elde edilmesi kritik neme sahiptir.

Kentlerdeki ağa tr sınıflandırmasına ynelik yapılan arařtırmalarda, arazi lmleri ve hava fotoęrafı yorumlanması kentsel ağa trleri hakkında bilgi edinmenin geleneksel iki yoludur [4, 12-13]. Fakat ağaların sınıflandırılmasına ynelik arazi alıřmaları zaman alıcı, pahalı ve genellikle geniř alanları tamamen kapsamazken, hava fotoęraflarının yorumlanması da ağa trlerinin byk lekli haritalanması iin yavař ve pahalı kalmaktadır [4]. Ayrıca, hava fotoęraflarının yorumlanmasında yorumlayıcının deneyimine baęlı olarak farklı yorumcular tarafından yapılan fotoęraf yorumları arasında byk tutarsızlıklar oluřabilmektedir [14]. Bu nedenle, ağa tr belirleme alıřmaları iin, aynı anda geniř bir alanda veri elde edebilme avantajına sahip uydu uzaktan algılama grntlerini kullanmak alternatif bir yol olarak ortaya ıkmıřtır [7]. Bununla birlikte uydu grntlerinin kullanımı genel olarak bitki rts sınıflandırmasında bařarılı olsa da, ağa rtsnn yzey yansıtma deęerleri zerine eřitli spektral ve meknsal faktrlerin etkileri (atmosfer kořulları, aydınlatma geometrisi ve ağaların optik zellikleri nedeniyle) ve grnt sensrlerinin spektral ve meknsal konfigrasyonlarındaki sınırlamalar nedeniyle bireysel ağa tr sınıflandırmasında doęruluęu dřk kalmaktadır [8, 15-16]. Daha aık bir ifadeyle, ağaların dikey yapısı, karıřık ta yapıları ve uydu grntlerinin detay

seviyesindeki otomatik algılama, arařtırmacıların uydu görüntüleri kullanarak ağaç türleri belirlemede karşılařtığı temel zorluklar olarak ifade edilebilir [17]. Geleneksel uydu görüntülerinde farklı ağaç türleri benzer spektral özelliklere sahip olabilirler [14].

Ağaç türü sınıflandırmasında uzaktan algılanmış görüntülerin kendi özelliklerinin yanı sıra kentsel çevreden dolayı oluşan sınırlamalar da mevcuttur. Ağaç taçlarının yoğun olarak dağıldığı ve çevrelerinin nispeten homojen olduğu ormanlık alanlarla karşılaştırıldığında kentsel alanlar, karmaşık ve heterojen arazi örtüsüne sahiptir ve bu nedenle ağaç türleri sınıflandırmasında belirli zorluklarla karşılaşabilmektedir [13]. Örneğin; yüksek katlı nesnelere tarafından oluşan gölgeler, gölgelenen ağaç türlerinin spektral bilgilerinin azalması veya tamamen kaybolmasına yol açar, böylece ağaçların sınıflandırılmaları ve yorumlanması zorlaşır veya ağaç taçlarının altındaki, toprak, asfalt ya da çimento gibi çeşitli arka plan materyalleri, yüksek çözünürlüklü sensörlerle gözlemlenen taçların spektral bilgilerini etkileyebilir [13].

Aktif bir uzaktan algılama lazer teknolojisi olan LiDAR (Light Detection and Ranging) ise, kentsel ağaçların yapısının ayrıntılı üç boyutlu bilgisinin sunulmasında arazi ölçümleri ve diğer birçok uzaktan algılama teknolojisine göre büyük bir avantaj sağlamaktadır [18]. Hava LiDAR, ağaç tacına nüfuz edebilir (ağaçların bazı iç yapılarını gösterebilir), geometrik bilgileri sağlayabilir ve ayrıca LiDAR'ın nokta yoğunluğu artırılarak bu yeteneklerinin daha da artması sağlanabilir [17]. Hava LiDAR'ın üç boyutlu bilgileri doğrudan sunması, bitki örtüsünden birden çok dönüş sinyali alabilme kabiliyetinin yüksek olması ve yoğunluk (intensity) verisi toplama özelliği LiDAR verilerini hava fotoğraflarının önemli bir alternatifi haline getirmiştir [17]. Havadan lazer tarama, ağaç taçlarının yapısı ve şeklini ölçtüğü için fotoğraflarda benzer spektral yansıma veren ağaç türleri (ladin ve çam gibi) arasında ayırım yapmak için efektif bir şekilde kullanılabilir [19].

Farklı ağaç türleri, yaprak dağılımı ve dallanma şekillerinde farklılık gösterirler [16]. Sonuç olarak, taç içinde birçok LiDAR veri noktası kullanılarak karakterize edilen bir ağacın yapısal özellikleri, bir yorumlayıcının ağaç türlerini doğru bir

şekilde tanımlama başarısını arttırabilir. Buradaki zorluk, çok sayıda LiDAR verilerinden gelen sayısız türler arasında önemli tanı özelliklerini çıkarmak ve seçmektir [16]. Her bir ağacı ayrı ayrı açıklayan önemli özellikleri bulmak, ağaç türü sınıflamasında kritik öneme sahiptir [20]. Yüksek yoğunluklu hava LiDAR verileri kullanılarak otomatik ağaç türleri sınıflandırması daha hassas kentsel ağaçların belirlenmesini sağlayacaktır, ancak buna ilişkin yöntemleri geliştirmek için daha fazla araştırma yapılması gerekmektedir [16].

LiDAR verileri, bireysel sinyal geri dönüşlerine (discrete return) veya tam dalga formuna (full-waveform) dayanan bir dizi ağaç yapısal değişkeni üretmeye olanak sağlar [8]. LiDAR verileri kullanılarak ağaç türlerinin sınıflandırılması, LiDAR verilerinden türetilen ağaçlara ilişkin geometrik ve radyometrik özellikler (features) sayesinde gerçekleştirilir [21-22]. Burada ifade edilen özellikler, nokta bulutundan doğrudan ya da daha ileri işlemler (radyometrik kalibrasyon) sonucu elde edilen tek bir ağacın nicelik bilgilerini (yükseklik, taç genişliği, genliği, oranı (crown length-width-ratio) vb.) veya geri yansıyan lazer sinyalinin bir gözlemine (sinyal genişliği (echo width), geri yansıma kesiti (backscatter cross-section) vb.) tanımlamaktadır [22]. Daha açık bir ifadeyle geometrik özellikler temel olarak LiDAR nokta dağılımına, radyometrik özellikler ise lazer tarafından yayılan ve geri yansıyan sinyalin miktarı, dağılımı ve yansıtma özelliklerine bağlıdır.

LiDAR nokta bulutunun sınıflandırılması için kullanılan yaygın sınıflandırıcılar karar ağaçları (desicion tree), rastgele orman (Random Forest - RF), yapay sinir ağları (Artificial Neural Network - ANN) ve destek vektör makineleri (Support Vector Machine - SVM) olarak ifade edilebilirler [18]. LiDAR verileri kullanılarak ağaç türleri sınıflandırması genellikle birçok alt adımı barındıran üç temel aşamaya ayrılır; ilk olarak ağaç taçları, örtü yükseklik modelinden (Canopy Height Model – CHM) segmentasyon ile elde edilir, daha sonra bireysel ağaçların karakteristik özellikleri çıkarılır ve son olarak da çıkarılan özelliklere dayalı olarak, ağaç türleri uygun bir sınıflandırıcı kullanılarak sınıflandırılır [18-20, 23].

Öğrenme tabanlı yaklaşımlar gibi çoğu sınıflandırma yöntemi, 3B nokta bulutu verilerini düzenli 2B yapıya dönüştürerek, nokta sınıflandırması için 2 boyutlu bir algoritma kullanır fakat 3 boyuttan 2 boyuta dönüşüm işlemi bilgi kayıplarına

sebepe olur [24-25]. Nokta bulutlarının doğrudan 3B noktalar kullanılarak sınıflandırılması ile genel sınıflandırma doğruluğunun iyileştirilebileceği Miller vd. [26] tarafından (yaklaşık %10) ve Ekhtari vd. [27] tarafından (yaklaşık %4) ifade edilmiştir.

Bu tez çalışmasında, makine öğrenme algoritmaları ile LiDAR nokta bulutu verileri kullanılarak kentsel ağaç türlerinin sınıflandırılmasına yönelik bir yaklaşım geliştirilmesi amaçlanmıştır. Girdi olarak tüm LiDAR noktalarının ham 3B koordinat (X, Y, Z) ve yoğunluk verilerini doğrudan kullanan, önerilen makine öğrenme tabanlı sınıflandırma yaklaşımı ile çalışma alanı olarak seçilen Yıldız Teknik Üniversitesi, Davutpaşa Kampüsü'nde yer alan geniş ve iğne yapraklı kentsel ağaçlar otomatik olarak LiDAR nokta bulutu verilerinden elde edilmiştir. Çalışma alanındaki belirlenen Test alanlarında (Test alanı 1 ve Test alanı 2) kullanılan destek vektör makineleri (DVM), rastgele orman (RO) ve çok katmanlı algılayıcı (ÇKA) makine öğrenme algoritmaları için doğruluk analizleri gerçekleştirilmiş ve en iyi sınıflandırma doğruluğuna sahip algoritma belirlenmiştir. Kentsel ağaç türlerinin elde edilmesi amacıyla gerçekleştirilen nokta tabanlı sınıflandırma için en önemli özellikler RO sınıflandırması ile ortalama Gini'deki azalmaya (Mean Decrease Gini - MDG) göre belirlenmiştir. Son olarak belirlenen kentsel çalışma alanında ağaç türlerinin sınıflandırılmasının ardından, burada geliştirilen hiyerarşik kural tabanlı sınıflandırma ve nokta tabanlı segmentasyon yaklaşımları ("MeanShift" ve "DBSCAN") kullanılarak, iki farklı test alanında (Test alanı A ve Test alanı B) kentsel cadde ağaçlarının elde edilebilirliği test edilmiştir.

Bu tez çalışması kapsamında geliştirilen yaklaşım, kentsel alanlardaki geniş ve iğne yapraklı ağaçların LiDAR nokta bulutu verilerinden makine öğrenme algoritmaları ile otomatik olarak sınıflandırılmasına olanak sağlayacaktır. Pahalı, uzun zaman gerektiren, emek yoğun ve deneyime bağlı tutarsızlıklara sebep olabilen geleneksel yöntemler yerine bu çalışmada önerilen otomatik yaklaşım ile büyük alanlarda daha düşük maliyetli, daha hızlı ve otomatik olarak ağaç türleri elde edilebilecektir. Elde edilen kentsel ağaç türleri yerel yönetimler ya da ilgili diğer kamu ve özel sektör kuruluşları tarafından afet yönetimi, şehir planlaması, çevre koruma veya kentsel kalkınma politikası üretimi gibi çalışmalarda rahatlıkla

kullanılabilecektir. Kentsel alanlarda tür belirleme çalışmaları belirli periyotlarda kolayca tekrarlanabilecek ve endemik türler veya anıt ağaç gibi değerli ağaç türlerinin belirlenmesi çalışmaları da hızlı ve sürekli bir şekilde gerçekleştirilebilecektir. Ayrıca, 3B kent modeli, coğrafi bilgi sistemi üretimi vb. çalışmalar için de detaylı kentsel geniş ve iğne yapraklı ağaç türü verisi elde edilmiş olacaktır.

Bu tez çalışması 5 bölümden oluşmaktadır. Giriş bölümü olan birinci bölüm, literatür özeti, tezin amacı ve hipotez alt başlıklarını içermektedir. Sınıflandırma, segmentasyon yöntemleri ve sınıflandırmanın doğruluk performansına ilişkin bilgiler ikinci bölüm olan sınıflandırma bölümünde yer almaktadır. Üçüncü bölümde makine öğrenmesi algoritmalarına ilişkin temel bilgilere yer verilmiştir. Dördüncü bölümde makine öğrenme algoritmaları ile ham LiDAR nokta bulutu verileri kullanılarak kentsel geniş ve iğne yapraklı ağaç türlerinin sınıflandırılması olanakları incelenmiştir. Bu bölümde ham LiDAR verileri kullanılarak hiyerarşik kural tabanlı sınıflandırma yaklaşımı ile yüksek bitki örtüsü sınıfının elde edilmesi, yüksek bitki örtüsü sınıfı noktalarının nokta tabanlı segmentasyonu sonucu bireysel ağaç taçlarının elde edilmesi, sınıflandırma için farklı özellik setlerinin oluşturulması ve bu özellikler kullanılarak üç farklı makine öğrenme algoritması (DVM, RO ve ÇKA) ile kentsel çalışma alanında yer geniş ve iğne yapraklı ağaçların sınıflandırılması işlem adımları ayrıntılı bir şekilde verilmiştir. Ayrıca nokta tabanlı sınıflandırmaların doğruluk performanslarının değerlendirilmesi ve en önemli sınıflandırma özelliklerinin belirlenmesi aşamalarına ilişkin detaylar da dördüncü bölümde yer almaktadır. Çalışma sonuçları ve ileriki çalışmalar için öneriler ise çalışmanın son bölümü olan beşinci bölümde verilmiştir.

## 1.1 Literatür Özeti

Geçmişten günümüze LiDAR verileri kullanılarak gerçekleştirilen ağaç türleri sınıflandırma uygulamalarında, çalışma alanları, kullanılan veri setleri, sınıflandırma yöntemleri ve sınıflandırılan ağaç türü sayıları çok çeşitlilik göstermektedir. Ağaç türleri sınıflandırma uygulamaları, sürdürülebilir orman yönetimi, ekolojik ve çevresel koruma, kereste kalite özelliklerinin, biyokütle ve tür çeşitliliğinin hassas haritalanması, bireysel ağaç büyümesinin modellenmesi

gibi etkenlerden dolayı bugüne kadar yaygın olarak ormanlık alanlarda gerçekleştirilse de; terleme, gölgeleme ve ısı emme yoluyla hava, su ve arazi kalitesini iyileştirme, iklim kontrolü, karbon birikimini azaltma, gürültü, toz, gaz ve rüzgar zararlarını önleme, toprak ve su dengesini koruma, enerji tasarrufu sağlama gibi önemli nedenlerden dolayı günümüzde kentsel alanlarda da sıklıkla gerçekleştirilen bir çalışma haline gelmiştir [28-30]. Bu uygulamalarda, kesikli dönüşlü (discrete-return) LiDAR verileri kullanıldığı gibi, daha yeni bir teknoloji olan tam dalga formu kaydeden (full-waveform) LiDAR verilerinden de faydalanılmaktadır [31-36]. Ağaç türleri sınıflandırma uygulamaları için LiDAR verileri tek başına kullanabileceği gibi, uydu görüntüleri, hava fotoğrafları ya da hiperspektral görüntüler gibi farklı birkaç veri türü bütünleşik olarak da sınıflandırma uygulamalarında kullanılabilir [37-45]. Ayrıca LiDAR verileri, ağaç türleri sınıflandırması için ham olarak kullanılmasının yanında geleneksel görüntü işleme algoritmalarının doğrudan uygulanabilmesi nedeniyle düzenli grid yapıda da yaygın olarak temsil edilebilmektedir [46-49]. Gerçekleştirilen ağaç türü sınıflandırma uygulamalarının bazılarında sadece iğne yapraklı ve yapraklarını döken, geniş yapraklı ağaç türleri ayrılırken, bazılarında da çok sayıda farklı türün sınıflandırılması yapılmaktadır [50-52]. Özetle çalışmaların amacına ve çalışma alanlarına göre, farklı uygulama aşamalarında farklı veri türü ve sınıflandırma yöntemlerinin kombinasyonları ile son on yıllarda başarılı ağaç türleri sınıflandırmaları sıklıkla gerçekleştirilmektedir. Fassnacht vd. [53] gerçekleştirdikleri inceleme çalışmalarında, günümüze kadar uzaktan algılama verileri kullanılarak yapılan ağaç türü sınıflandırmalarına ilişkin bahsedilen tüm bu bilgileri de içeren kapsamlı derlemelere yer vermişlerdir.

Holmgren ve Persson [19] çalışmalarında İskoç çamı (*Pinus sylvestris*) ve Norveç ladini (*Picea abies*) arasındaki ayırım için lazer verilerindeki özellikleri bulmayı ve farklı orman türlerinde bu sınıflandırmayı doğrulamayı amaçlamışlardır. İlk olarak, tek tek ağaçların segmentasyonu, lazer verilerinden üretilen sayısal örtü modeli kullanılarak gerçekleştirilmiştir. İkinci olarak, ağaç yükseklikleri ve taç alanları türetilmiştir. Daha sonra taç şeklinin ve yapısının tipik özellikleri bulunarak, çam ve ladin ağaç segmentlerini ayırmak için lazer verilerinden bir dizi değişken çıkarılmıştır. Çalışmada, tek tek ağaçları tanımlama ve bu ağaçların

yükseklik ve taç çapını tahmin etme yöntemi altı adımdan oluşmuştur. Bu adımlar; bir sayısal yüzey modeli üretme (Digital Surface Model - DSM) (1), bir sayısal arazi modeli üretme (Digital Terrain Model - DTM) (2), ağaçların örtüsünün modellenmesi ve bir sayısal örtü modeli (Digital Canopy Model - DCM) üretme (3), sayısal örtü modelinin farklı ölçeklerle düzeltilmesi (4), görüntünün farklı bölümleri için hangi ölçeğin seçileceğini belirlemek için yükseklik verilerine parabolik bir yüzey yerleştirilmesi (5), tanımlanan ağaçlar için yükseklik ve taç çaplarının hesaplanması (6) olarak ifade edilebilir. Ağaç türü sınıflandırmasında kullanılacak çıkarılan değişkenler de farklı türlerin lazer dönüşlerinin oranı ve yükseklik dağılımı, geometri ve yoğunluk ölçümlerine dayandırılmıştır. Lazer verilerinden çıkarılan değişkenler, aralarındaki korelasyona göre gruplandırılmıştır. Aralarında mutlak korelasyon katsayısı  $r \geq 0.70$  olan değişkenler bir grup oluşturmuş ve "Student's t" testi iki ağaç türü grup ortalamasındaki farkı test etmek için tüm değişkenler için gerçekleştirilmiştir. Her grupta en yüksek mutlak t değerine sahip değişken, sınıflandırmada kullanılmak üzere seçilmiştir. Seçilen değişkenler kullanılarak sınıflandırma hem klasik lineer hem de kuadratik ayırt edici (quadratic discriminant) fonksiyon kullanılarak yapılmıştır. Bu çalışma ile İskoç çamı ve Norveç ladin sınıflandırması için genel sınıflandırma doğruluğu %95 ile umut verici olarak belirlenmiştir.

Reitberger vd. [23] tam dalga formlu LiDAR verilerinin bir dalga formu ayrıştırma tekniği ile türetilen özelliklerine dayalı bir ağaç türü sınıflandırma yaklaşımı belirlemek için bir çalışma gerçekleştirmişlerdir. Genel amaç, zaman kazandıran ve pahalı orman envanter yöntemlerini, tam dalga tarayıcılarındaki LiDAR verilerini kullanan yeni tekniklerle değiştirmektir. Öncelikle, 3B noktalar ve özellikleri, geleneksel ilk ve son darbe tekniklerine kıyasla çok daha fazla sayıda nokta üreten dalga formlarından çıkarılmıştır. Ardından çıkarılan 3B noktalardan hesaplanan ağaçların karakteristik özellikleri belirlenmiş ve son olarak bu özellikler kullanılarak iki aşamalı bir prosedür olan ağaç türlerinin kümelenmesi ve ardından gelen Bayes sınıflandırması ile ağaç türleri belirlenmiştir. Çalışma sonucunda verileri iki küme halinde gruplandıran (yaprak döken, iğne yapraklı) sınıflandırma ile yaprakların olduğu bir zamanda (leaf-on) %80'lik bir genel doğruluk elde edilmiştir.

Reitberger vd. [20] küçük ayak izi tam dalga formu LiDAR verilerinden türetilen özellikleri kullanarak ağaç türü sınıflandırması için bir yöntem geliştirmişlerdir. İlk olarak, lazer ışını yansımalarının, yoğunluğunun ve darbe genişliğinin 3 boyutlu koordinatları, dalga formuna bir dizi Gauss sinyaline uyan bir dalga formu ayrıştırma tekniği ile çıkarılmıştır. Daha sonra, ağaç taçları “watershed” algoritması kullanılarak örtü yükseklik modelinden (CHM) türetilmiştir. Segmentasyon işleminin ardından da yansımaların 3 boyutlu koordinatlarından hesaplanan ağaç özellikleri, yoğunluk ve darbe genişliği, iğne yapraklı ve yaprak döken ağaçları kontrolsüz bir sınıflandırma ile tespit etmek için kullanılmıştır. Ağaç türleri, çıkarılan özellikler kullanılarak bir “K-means” ve bir “Expectation Maximum (EM)” sınıflandırıcısı ile sınıflandırılmıştır. Son olarak, sonuçlar doğrulanmış ve gerçek doğruluk verileri ile karşılaştırılmıştır. Verileri iki küme (iğne yapraklı, yaprak döken) halinde sınıflandıran sınıflandırma, yaprakların olduğu durumda %85, yaprakların olmadığı durumda %96'lık genel bir doğruluğa ulaşmıştır (en iyi durumda).

Hollaus vd. [54]'nin çalışmalarının amacı karışık ormanlık alanın ağaç türlerinin sınıflandırılması için tam dalga formu kaydeden hava LiDAR'ın potansiyelini analiz etmektir. Farklı ağaç türlerinin (ladin, karaçam, kırmızı kayın) yaprakların olmadığı zamanda saçılma mekanizmalarının karakterize edilmesine odaklanmışlardır. Bu amaçla, kullanılan nicelikler sinyal genişliği, geri yansıma kesiti ve sinyalin dikey yöndeki dağılımıdır. Ağaç taç segmentlerini elde etmek için bir kenar tabanlı segmentasyon yaklaşımı kullanılmıştır. Ardından araştırma verileri analizinin bulgularına dayanarak, ladin, karaçam ve kırmızı kayın ağaç türlerini sınıflandırmak için bir karar ağacı ile birlikte bulanık bir kural belirlenmiştir. İlk aşamada, yapraklarını döken ve iğne yapraklı ağaçlar, olasılıkların (p) 0.05 ile 0.95 arasında değiştiği doğrusal olasılık fonksiyonlarına dayanarak birbirinden ayrılmıştır. Bu sınıflandırmada elde edilen toplam doğruluk %83 olarak belirlenmiştir. Ardından, ladin ve karaçamı birbirinden ayırmak için sınıflandırılmış iğne yapraklı segmentler bir eşik değeri kullanılarak alt sınıflara ayrılmıştır. Sınıflandırılan ladin, karaçam ve kırmızı kayın ağaçları için ise toplam doğruluk %75 olarak belirlenmiştir. Çalışma sonucunda elde edilen sonuçlar, tam dalga formundaki havadan lazer tarama verilerinin geometrik bilgileri ve geri

yansıtma özelliklerini birleştirmenin ağaç türü sınıflamasında yüksek bir potansiyele sahip olduğunu göstermektedir.

Kim vd. [12]'nin Washington Park Arboretum, Seattle, Washington, USA'da gerçekleştirdikleri çalışmalarının amacı LiDAR yoğunluk (intensity) verisinin ağaç türlerini ve tür gruplarını ayırmada kullanılabilirliğini test etmek ve türler arasındaki farklı yoğunluk özellikleriyle hangi faktörlerin ilişkili olacağını araştırmaktır. Yaprakların olduğu ve olmadığı zamanlarda LiDAR verisi toplanmış ve aynı zamanda arazi çalışmaları, sekiz geniş yapraklı ve yedi adet iğne yapraklı tür için ağaç konumlarını, ağaç türlerini ve yüksekliklerini, taç taban yüksekliklerini ve taç çaplarını ölçmek için yürütülmüştür. Yaprakların olmadığı zamanda elde edilen LiDAR veri seti, çalışma alanı için bir sayısal arazi modeli oluşturmak için kullanılmıştır. İlk olarak bireysel ağaçlar, arazide ölçülen gövde konumu yardımıyla yaklaşık ağaç konumuna yakın LiDAR geri dönüş verilerini gösteren "FUSION/LDV" yazılımı kullanılarak saptanmıştır. Ağaç merkezi türlere bağlı olarak farklı bir şekilde; geniş yapraklı türler için taç içindeki x ve y koordinatlarının ortalama bir değeri, iğne yapraklı türler için ağacın en yüksek noktası olarak kullanılmıştır. Yine "FUSION/LDV" yazılımı kullanılarak örtü yüzey modeli oluşturulduktan sonra taç yüzeyini temsil eden lazer noktaları çıkarılmıştır. Ardından her taç içindeki lazer noktaları kullanılarak, her ağaç için yoğunluk verilerini analiz etmek amacıyla değişkenler hesaplanmıştır. Tüm değişkenler taç taban yüksekliğinin üzerinde bulunan lazer geri dönüşleri kullanılarak türetilmiştir. Ortalama yoğunluk değerleri, hem yaprakların olduğu hem de yaprakların olmadığı veri setlerindeki her bir ağaç tacı, üst taç ve taç yüzeyini temsil eden lazer dönüşleri kullanılarak hesaplanmıştır. Temel bileşenler analizi (Principal Component Analysis - PCA), bu değişkenlerin tamamında mevcut olan hemen hemen tüm bilgileri içeren orijinal değişkenlerin bir alt kümesini belirlemek için gerçekleştirilmiş ve geniş yapraklı ve iğne yapraklı türler için basit bir ağaç türü sınıflandırması, klasik lineer ayırıcı fonksiyon (linear discriminant function) kullanılarak orijinal değişkenlerin seçilen alt kümesi üzerinde yapılmıştır. Gerçekleştirilen bu çalışma genel olarak LiDAR yoğunluk verilerinin, geniş yapraklı türleri iğne yapraklılardan ve bu geniş gruplar içindeki çeşitli ağaç türlerini birbirinden ayırt etmek için kullanılabileceğini göstermiştir.

Çapraz doğrulama ile lineer ayırıcı fonksiyon kullanılarak gerçekleştirilen geniş yapraklı ve iğne yapraklı türler için sınıflandırma sonuçlarına göre yaprakların olmadığı zamandaki sınıflandırma oranı (%83.4) yaprakların olduğu durumdan (%73.1) yüksek, iki verinin birlikte kullanımı ise en yüksek (%90.6) olarak göze çarpmaktadır.

Ørka vd. [31]'nin amacı, iğne yapraklı ve geniş yapraklı ağaç türlerini ayırmak için hava lazer tarayıcı (ALS) verilerinden türetilen aday özellikleri belirlemek ve bu özelliklerin sınıflandırma performanslarını ölçmektir. Hem geometrik hem de yoğunluk tabanlı özellikleri dikkate aldıkları bu çalışmayı Norveç'te bir orman rezervinde yer alan 197 Norveç ladini ve 180 huş ağacı (şartlara göre yapraklar) üzerinde gerçekleştirmişlerdir. Tek aday özellikler için sınıflandırma işlemi gerçekleştirilerek genel doğruluğu düşük olan özellikler, aday özellikler kümesinden çıkarılarak sınıflandırma kullanılacak özellikler belirlenmiştir. Tür sınıflandırması için Lineer Diskriminant Analizi (LDA) kullanılarak belirlenen geometrik ve yoğunluk tabanlı özellikler için sırasıyla %77 ve %73'lük genel sınıflandırma doğruluğu elde etmişler ve bu özelliklerin birleştirilmesi ile ise %88'lik bir genel sınıflandırma doğruluğuna ulaşmışlardır.

Korpela vd. [55] gerçekleştirdikleri araştırma uygulamasında Güney Finlandiya'daki 13000'den fazla orman ağacında iki LiDAR sensörünün tek ağaç seviyesinde tepkisini incelemişlerdir. Çalışmanın amaçları ilgili LiDAR özelliklerini keşfetmek ve gövde ve ağaç değişkenlerine bağımlılıklarını, iki sensörü ve bunların füzyonunu, yoğunluk normalizasyonundan kazanımı, eğitim setinin boyutunun önemini ve yaş ve yerinde doğurganlığın etkilerini, ekonomik olarak daha az önemli türlerdeki LiDAR sinyalinin incelemektir. Çalışmanın amaçları doğrultusunda, kullanılmak üzere LiDAR değişkenleri belirlenmiş ve bu değişkenlerden çeşitli özellikler hesaplanmıştır. Bu uygulamada sınıflandırma çalışmalarında doğrusal diskriminant analizi (linear discriminant analysis - LDA), k-en yakın komşu (k-Nearest Neighbor - k-NN), k-en benzer komşu (k-most similar neighbor - k-MSN) ve rasgele orman (random forest - RF) algoritmaları kullanılmıştır. Kesikli dönüşlü LiDAR yoğunluk verilerindeki bireysel çam, ladin ve huş ağaçlarının tür sınıflandırmasında %88-90 doğruluk elde edilmiştir. Test edilen sınıflandırıcılar arasında RO, en yüksek sınıflandırma performansına sahip

olmuştur. Çalışmanın sonuçları, sadece yansımadan etkilenmeyen yaprak boyutu, yönlendirme, yeşillik yoğunluğunun LiDAR yoğunluk verisini etkilediğini göstermiştir. Sensörlerin karşılaştırılması, tür sınıflandırması için yoğunluk verisinin performansının, belirsiz kalan nedenlerden dolayı sensörler arasında değiştiğini ortaya koymuştur. Hem mesafe (range) hem de kazanç alıcı normalizasyonun (gain receiver normalization) tür sınıflandırmasını; yoğunluk değerlerinin ağırlıklandırılmasının, iki LiDAR veri kümesinin birleştirilmesini geliştirdiği belirlenmiştir.

Li vd. [15] araştırmalarında, ağaç yapısının özelliklerinden tam olarak yararlanmak için yüksek yoğunluklu hava LiDAR verilerini (23 nokta/m<sup>2</sup>) kullanarak, iğne yapraklı ağaçlar veya yaprak döken ağaçlardaki türlerin sınıflandırılmasını gerçekleştirmeyi amaçlamışlardır. Bu amaçla LiDAR verilerinden yapısal değişken türetme yöntemleri, yüksek yoğunluklu veriler için geliştirilmiş ve yeni yapısal değişkenler oluşturulmuştur. Farklı türlerin taç şekli ve yaprak dağılımı farklı olduğundan bunları karakterize etmek için, her bir ağaç noktası segmentinden iki özellik grubu çıkarılmıştır. İlk önce her segmentten 79 değişken çıkarılmış ve temel bileşenler analizi (Principal Component Analysis - PCA) yönteminin uygulanmasından sonra, değişken sayısı 25'e düşürülmüş ve daha sonra da korelasyon matrisi temel alınarak 15 yüksek korelasyonlu değişken daha kaldırılmıştır. Çalışmada LiDAR nokta bulutlarından elde edilen 276 ayrı ağaç segmentinden yarısı karar ağacı sınıflandırıcısının eğitimi için, diğerleri ise doğrulama için kullanılmıştır. Çalışma bölgesinde yer alan beş dominant türün (jack çamı (*Pinus banksiana*), beyaz çam (*Pinus strobus*), beyaz huş (*Betula papyrifera*), kavak (*Populus tremuloides*) ve şeker akçaağaç (*Acer saccharum*) genel sınıflandırma doğruluğu % 81.3 olarak elde edilmiştir. Çalışma, Kuzey Amerika'nın kıta ormanlarında türlerin sınıflandırılması için yüksek yoğunluklu hava LiDAR verilerinin yararlı olduğunu göstermiştir. Ayrıca, deneysel sonuçlar, karar ağacı algoritmasının tür sınıflaması için etkin olduğunu da ortaya çıkarmıştır.

Heinzel ve Koch [56] çalışmalarında, tam dalga formlu LiDAR verileri ağaç türü sınıflandırmasında daha büyük miktarda çıkarılabilir bilgi sunduğundan, tam dalga formlu LiDAR verilerinin büyük bir kısmını ele almak ve kapsamlı bir dizi

değişkenin kombinasyonunu çıkarmak için bir algoritma geliştirmişlerdir. Bu çalışmanın ilk aşaması, yüksek miktarda veriyi (test yeri için yaklaşık 15 GB ASCII formatında LiDAR verileri) yönetilebilir bir biçime getirmektir. Bu nedenle, farklı nokta parametrelerini, istatistikleri ve diğer özellikleri birleştirerek nokta bulutunu düzenli yapıya (raster) dönüştüren bir program geliştirilmiştir. Algoritma, TreesVis kitaplığından optimize edilmiş LiDAR işleme algoritmalarını kullanarak C++'da uygulanmıştır. Tam dalga formu parametrelerinden genlik, genişlik, yoğunluk ve bir sinyal içindeki toplam hedef sayısı 10 km<sup>2</sup>'lik çalışma alanını kapsayan düzenli yapıya dönüştürülmüştür. Bu aşamadan sonra lineer ayırım analizi (linear discriminant analysis - LDA) kullanılarak ağaç türü sınıflandırılması için daha fazla sayıda değişken elde edilmesi sağlanmıştır. İleri adımlı LDA, her yineleme sırasında bir sonraki en yüksek ayırmacı güce sahip değişkeni belirleyen iteratif bir işlemidir ve bu nedenle çok sayıda değişkeni sınırlandırmak için toplam 10 üst düzey değişken için 10 LDA uygulanmıştır ve bunlardan da en uygun değişkenler belirlenmiştir. Yoğunluk, genişlik ve toplam hedef sayısına dayanan en önemli üç değişkeni belirlemek mümkün olmuş ve altı ağaç türüne kadar toplam %57 hassasiyetle sınıflandırma gerçekleştirilmiştir. Dört ana tür doğruluğu %78'e yükselmiş ve sadece iğne yapraklı ağaçlar ve geniş yapraklı ağaçlar için %91 oranında doğru sınıflandırma yapılabilmektedir.

Dalponte vd. [57]'nin çalışmalarının ana amacı, gelecekte Alplerde ağaç türü haritalamasında kullanılmak üzere farklı yoğunluklarda multispektral, hiperspektral ve LiDAR verilerine dayalı farklı veri gruplarını analiz etmektir. LiDAR verileri ile birleştirilmiş yüksek mekânsal çözünürlüklü hiperspektral görüntüler ve LiDAR verileriyle birleştirilmiş yüksek mekânsal çözünürlüklü uydu multispektral görüntüler olmak üzere iki çoklu sensör kurulumunu analiz etmişlerdir. Araştırmacılar çalışmanın özelleşmiş amaçlarını ise hem geniş yapraklı hem de iğne yapraklı birçok türe sahip bir dağlık alan üzerinde yüksek spektral ve mekânsal çözünürlüklü çoklu sensör kurulumunun sınıflandırma performansını anlamak (çok yüksek mekânsal ve spektral çözünürlüğe sahip hiperspektral görüntüler ile yüksek nokta yoğunluklu LiDAR verilerinin entegrasyonu) (1), optimum kurulumla göre türlerin ayrılabilirliği üzerine veri karakteristiklerinin düşürülmesinin sonuçlarını anlamak (spektral görüntülerin

spektral çözünürlüğü ve LiDAR verilerinin nokta yoğunluğuna göre) (2) ve yüksek ve alçak nokta yoğunluklu LiDAR veri toplamının ağaç türleri sınıflandırılmasındaki farklarını anlamak (3) olarak açıklamışlardır. Her bir veri kurulumu için ham verilerin ön işlenmesinden sınıflandırmaya kadar belirli bir akış zinciri benimsenmiştir. Bu çalışmada, doğrusal olmayan bir destek vektör makinesi (Support Vector Machine - SVM) sınıflandırıcısı ve rastgele orman (Random Forest - RF) sınıflandırıcısı olmak üzere iki kontrollü sınıflandırıcı kullanılarak ağaç türlerinin belirlenmesi işlemi gerçekleştirilmiştir. Analiz edilen türlerin farklı ağaç yapılarını karakterize eden özellikleri, yüksek yoğunluklu LiDAR verilerinden çıkarılmıştır. Deneysel sonuçlar, hiperspektral verilere dayanan kurulumun, genel makro sınıflar, orman türleri ve tek türlerle etkili olduğunu ve yüksek kappa doğruluklarına (sırasıyla %93.2, %82.1 ve %76.5) ulaştığını göstermiştir. Tek ağaç türleri için keskin, orman türleri için hala yüksek olmak üzere multispektral verilerin kullanılması sınıflandırma doğruluğunda bir azalmaya neden olmuştur. Genel makro sınıflar düşünüldüğünde ise, çoklu spektral kurulumun hala çok yüksek (%85.8) olduğu tespit edilmiştir. LiDAR verileriyle ilgili deneysel analiz de, yüksek yoğunluklu LiDAR verilerinin, hiperspektral veya multispektral verilerle kombine edildiğinde, düşük yoğunluklu verilere göre ağaç türü sınıflandırması için daha fazla bilgi sağladığını göstermiştir.

Zhang ve Qiu [58] çalışmalarında LiDAR ve hiperspektral görüntülerden kentsel ağaç türlerini tanımlamak için sinir ağı tabanlı (neural network based) bir yaklaşım geliştirmişlerdir. Tek ağaçlar ilk olarak LiDAR nokta bulutu verilerinden tespit edilmiş ve daha sonra ağaç türleri ilişkili hiperspektral bilgiden tanımlanmıştır. LiDAR verisinden tek ağaçların belirlenmesi iki aşamada gerçekleştirilmiştir. İlk olarak her bir LiDAR noktasının arazi topoğrafyasından bağımsız yüksekliklerinin belirlenmesi için bir sayısal arazi modeli üretilerek orijinal LiDAR dönüşlerinden çıkarılmıştır. İkinci aşama olan tek ağaçların belirlenmesinde öncelikle, vektör tabanlı yaklaşım ve hiperspektral görüntüden üretilen normalize edilmiş fark bitki örtüsünden (Normalised Difference Vegetation Index - NDVI) de yararlanılarak bireysel ağaç segmentasyonu gerçekleştirilmiş, ardından bir ağaç tepe noktası belirleme algoritması ile de her

bir ağacın en yüksek noktası; bir ağaç tacı türetme algoritması ile de ağaç taçları elde edilmiştir. Son olarak da ağaç türleri, hiperspektral görüntüden geliştirilen, hem bir sinir ağı hem de bir bulanık sistemin temel bileşenleri içeren “Adaptive Gaussian Fuzzy Learning Vector Quantization (AGFLVQ)” yaklaşımı ile elde edilmiştir. Gerçekleştirilen bu çalışma ile LiDAR verisi ve hiperspektral görüntülerin birlikte kullanımı ile sadece bireysel ağaçları tespit etme ve bu ağaçların metriklerini tahmin etme yanında geliştirilen algoritma ile ağaç türlerinin belirlenebileceği ortaya konmuş ve geleneksel arazi ölçümleri yerine bu iki verinin entegrasyonun büyük potansiyele sahip olduğu belirlenmiştir.

Li vd. [16]'nin çalışmalarının ana amacı, ağaçlar arasındaki türlerin sistematik olarak farklılık gösteren ayrıntılı yaprak ve dallanma şekillerini yakalamak üzere tasarlanmış, yeni ağaç taç yapısal özelliklerine dayalı orman ağacı türlerini doğru sınıflandırmak için küçük ayak izi LiDAR verilerinin kullanımını geliştirmektir. İkincil amaç ise, LiDAR nokta yoğunluğunun sınıflandırma doğruluğunu etkileyip etkilemediğini, etkiliyorsa nasıl etkilediğini araştırmaktır. Tam dalga formu verileri ön işlenmiş ve sağlanan tam veri seti ayrık LiDAR noktalarını içerecek şekilde üretilmiştir ve bunlar daha sonra TerraScan yazılımı kullanılarak zemine ve zemine ait olmayan noktalara sınıflandırılmıştır. Her dalga formundan beşe kadar geri dönüş elde edilmiş ve zemin noktaları, 0.15 m x 0.15 m mekânsal çözünürlükte bir sayısal arazi modeli (Digital Terrain Model - DTM) üretmek için kullanılmıştır. Bu aşamanın ardından üretilen örtü yükseklik modelinden yararlanılarak bir bireysel ağaç tacı üretme yöntemi kullanılarak ağaç tacı segmentasyonu gerçekleştirilmiştir. Ağaç türlerinin sınıflandırılması için, bir ağacın 3B dokusu, yapraklanma kümeleşmesinin göreceli derecesi, ağaç ögesi kümelenmesinin maksimum hale getirildiği nispi ölçek ve ağaç tacındaki boşluk dağılımını tanımlayan dört tip LiDAR özelliği tasarlanmış ve sonrasında bu dört özellik grubu 3 boyutlu doku (3D texture), göreceli kümeleme derecesi (relative clustering groups -RCD), göreceli kümeleme ölçeği (relative clustering scale - RCS) ve boşluk dağılımı (gap distribution - GD) özellik grupları olarak ifade edilmiştir. Özellikler bir genetik algoritma kullanılarak seçilmiş ve daha sonra ağaç türleri, seçilen özelliklere dayalı doğrusal ayırım analizi (Linear Discriminant Analysis - LDA) kullanılarak sınıflandırılmıştır. Dört tür olan şeker akçaağaç (Acer

saccharum Marsh.), titreyen kavak (*Populus Tremuloides Michx.*), jack çam (*Pinus banksiana Lamb.*) ve dođu çam (*Pinus strobus L.*) için sınıflandırma sonucunda toplam dođruluk %77.5 ve Kappa katsayısı 0.7 olarak elde edilmiştir. Çalışmanın sonuçları, ağaç türlerinin sınıflandırılmasına yardımcı olan türetilmiş yapısal özelliklerin önemini ortaya koymuştur. Ayrıca bu araştırma, LiDAR nokta yoğunluğu ile tür sınıflandırması dođruluđu arasında pozitif dođrusal bir korelasyon ( $R^2 = 0.88$ ) olduğunu da göstermiştir.

Alonzo vd. [59]'nin çalışmalarının amacı, "AVIRIS" görüntüsü ve yüksek yoğunluklu LiDAR verisinin taç nesne düzeyinde birleştirilmesi yoluyla, biyolojik çeşitlilik gösteren California, Santa Barbara şehrindeki ağaç türlerinin haritalanma dođruluđunu iyileştirmektir. 29 ortak ağaç türünü haritalamak için yüksek mekânsal çözünürlüklü (3.7 m) hiperspektral görüntü ve 22 nokta/m<sup>2</sup> LiDAR verisi birleştirilmiştir. Ön işleme, segmentasyon, özellik çıkarımı, birleştirme ve sınıflandırma olmak üzere dört temel aşamadan oluşan çalışmada ağaç taçları standart ayırım analizi (discriminant analysis) ile sınıflandırılmıştır. Tür düzeyinde ve yaprak tipi seviyesinde haritalar sırasıyla %83.4 (kappa = 82.6) ve %93.5 genel dođrulukları ile üretilmiştir. Çalışmanın sonucunda, ağaç türlerinin sınıflandırılması için uygulamaya LiDAR verilerinin eklenmesinin, tek başına spektral verilere kıyasla sınıflandırma dođruluđunda %4.2'lik bir artış sağladığı anlaşılmıştır.

Verlič vd. [60] uygulamalarında literatürde bulunan "Worldview-2" görüntüsü ile LiDAR verilerinin birlikte kullanıma ilişkin boşluğu kapatmak amacıyla, "Worldview-2" görüntüsü ve LiDAR verilerinin bir kombinasyonu ile Slovenya Ljubljana'daki dođal, karışık, heterojen kent ormanının egemen tabakasındaki beş farklı ağaç türünün bireysel taçlarının basit bir nesne tabanlı görüntü analizi (Object Based Image Analysis - OBIA) metodu kullanılarak dođrudan başarılı bir şekilde sınıflandırılıp sınıflandırılmadığını belirlemeye çalışmışlardır. Model eğitim verileri ormandaki sistematik bir arsa çizelgesinden elde edilmiştir. Çalışma alanında 304 iğne yapraklı (Norveç ladin ve İskoç çamı) ve 270 yaprak döken (Avrupa kayını, Sessile ve Pedunculate meşe (kombine) ve Tatlı kestane) ağaçları tespit edilmiştir. Çalışmada ağaç türü sınıflandırması, destek vektör makinesi (Support Vector Machine - SVM) modelinin uygulanmasıyla gerçekleştirilmiştir.

Sınıflandırmanın doğruluğunun bağımsız test verileriyle kıyaslayarak değerlendirilmesi için, bir hata (confusion) matrisi hesaplanarak doğruluk değerlendirmesi gerçekleştirilmiştir. Sınıflandırmanın genel doğruluğu %58 ve kappa katsayısı 0.431 olarak elde edilmiştir. Araştırmacılar bu çalışma ile kullanılan nesne tabanlı görüntü analizi yönteminin, ağaç taçlarının sık sık iç içe geçmiş yoğun, ağaç türü açısından zengin, oldukça heterojen doğal kentsel ormanlık alanlar için birleştirilmiş “WorldView-2” ve lazer tarama verilerinin umut verici bir şekilde uygulanabilirliği sonucuna varmışlardır. Bununla birlikte kent orman alanını oluşturan ağaç türlerinin doğruluğunun diğer çalışmalara göre daha düşük, yaprak döken ve iğne yapraklı ağaç türleri arasındaki ayrımın ise daha güvenilir olduğu sonucu da ortaya çıkmıştır.

Cao vd. [18]'nin çalışmalarının genel amacı, sub-tropikal ormanlarda (tipik güneydoğu Çin'de) ağaç türü sınıflandırması için küçük ayak izi tam dalga formu verisinin kapasitesini keşfetmektir. Voksel tabanlı bir kompozit dalga biçimi yaklaşımı kullanarak 6 sub-tropikal ağaç türünün (3 sınıflandırma düzeyinde) ayırım sonuçlarını değerlendirmek ve ağaç türünün sınıflandırılması için en önemli metrikleri ve kompozit dalga modellemesinde en uygun voksel boyutunu araştırmak çalışmanın özelleşmiş amaçlarıdır. Bu amaçlarla ilk olarak “Gauss” ayırıştırma teknikleri kullanılarak önceden işlenmiş dalga formlarından çıkarılan filtrelenmiş nokta bulutlarına dayanan bir sayısal arazi modeli (Digital Terrain Model - DTM) oluşturulmuştur. Ardından ağaçların belirlenmesi için, örtü yükseklik modeline (Canopy Height Model - CHM) dayanan değişken büyüklükteki bir dairesel pencere ile yüksek değerleri (ağaç tepeleri olması muhtemel) belirten bir lokal maksimum algoritması uygulanmıştır. Üçüncü olarak, belirli bir yere ait dikey profili içeren tek tek ağaç tabanlı bileşik dalga biçimleri belirlenen taçlar tarafından çıkarılmış ve metrikleri hesaplamak için bir veri modeli olarak kullanılmıştır. Son olarak da, bu metrikler, altı sub-tropikal ağaç türünün üç ayırım seviyesinde birbirinden ayrılması için bir rastgele orman (Random Forests - RF) sınıflandırıcıda kullanılmıştır. Çalışmanın sonuçları, tüm ağaç türlerinin nispeten yüksek doğrulukla (altı sınıf için %68.6, dört ana tür için %75.8 ve iğne yapraklı ağaçlar ve geniş yapraklı ağaçlar için %86.2) sınıflandırıldığını göstermektedir. Tam dalga formu metrikleri (medyan enerjinin

yüksekliğine, dalga formu mesafesine ve dalga formu tepe sayısına dayanılarak) yüksek sınıflandırma önemi göstermiş ve çeşitli voksel boyutları arasında istikrarlı kalmıştır. Sonuçlar ayrıca, voksel temelli yaklaşımın geniş tarama açılarıyla ilgili bazı sorunları hafifletilebileceğini önermektedir. Özetle bu çalışma ile elde edilen deneysel sonuçlara göre tam dalga formu LiDAR verilerinin sub-tropikal ormandaki ağaç türü sınıflandırması için önemli bir potansiyele sahip olduğu belirlenmiştir.

Dian vd. [8] Anyang, Henan, Çin'de 8 ortak ağaç türünü sınıflandırmak için yüksek mekânsal çözünürlüklü (0.5 m) hiperspektral görüntüler ile LiDAR verilerini ağaç tacı nesne düzeyinde birleştirilmişlerdir. İlk olarak LiDAR nokta bulutundan dikey kentsel orman özelliklerini ifade eden örtü yükseklik modeli (Canopy Height Model – CHM) çıkarılmış ve bu adımı takiben ağaç tacı nesne (Tree crown object – TCO) bilgileri “MeanShift” algoritması kullanılarak örtü yükseklik modelinden elde edilmiştir. Daha sonra örtü yükseklik modeli hiperspektral görüntülerden elde edilen minimum gürültü fraksiyon dönüşümü (minimum noise fraction transformation – MNF) ve iyileştirilmiş bitki indeksi (enhanced vegetation index – EVI) ile birleştirilmiştir. Bu kombine özellikler, farklı ağaç türleri için kaba bir sınıflandırma şeması oluşturmak için destek vektör makinaları (Support Vector Machine - SVM) sınıflandırıcısına girdi olarak kullanılmıştır. Son olarak da, son ağaç türü haritasını üretmek için TCO'ya çoğunluk oy kullanma yöntemi uygulanmıştır. Çalışmada hiperspektral görüntüler ve LiDAR verileri kullanılarak gerçekleştirilen ağaç türü sınıflandırma uygulamasına ilişkin iş akışı verilmiştir. Gerçekleştirilen bu deneysel çalışma sonucunda örtü yükseklik modeli-mekânsal-spektral özelliklerin birlikte kullanımının piksel tabanlı sınıflandırmada sadece özelliklerinin kullanımına kıyasla daha yüksek doğruluğa ulaştığı ortaya konulmuştur. Bununla birlikte, CHM ve EVI özelliklerinin kendi sınırlamalarının da mevcut olduğu, bunların büyük oranda farklı ağaç türlerinin farklı özelliklerine bağlı olduğu bu çalışma ile gözler önüne serilen başka bir sonuçtur.

Koenig ve Höfle [61] çalışmalarında, ağaç türleri sınıflandırması için, elde edilen tam dalga formu özelliklerine özel önem gösteren tam dalga formu kaydeden lazer tarama ile gerçekleştirilen uygulamaları derlemişlerdir. Tam dalga formu verisi ve tek ağaç sınıflandırmasını, ağaç türlerinin sınıflandırılması için türetilen

özellikleri, ağaç türleri sınıflandırması için kullanılan özellik setlerini detaylı bir şekilde incelemiştir. Ayrıca özellik karakteristiği ve ağaç sınıflandırmasını etkileyen ve sınırlayan faktörleri, ağaç türlerine özgü özellik karakteristiklerini de mevcut çalışmalardan yola çıkarak analiz etmişlerdir. Çalışmada genellikle ağaç türü sınıflandırması için öncelikle tek ağaçların belirlendiği, belirlenen bu tek ağaç türü objelerine dayanarak farklı özelliklerin türetildiği ve kombine edilerek ağaç türü sınıflandırmasında kullanıldığı vurgulanmıştır. Ağaç türü sınıflandırmasına ilişkin özellikler geometrik, radyometrik ve dalga formu özellikleri olmak üzere üç türe ayrılmıştır. Bu üç türe göre belirlenen, ağaç nesnesinin belirli bir uzamsal alt birimi içindeki bir özelliği sınıflandıran sınıf alt-kümesi özellikleri ile tüm nesneye ilişkin olarak bir özelliği kapsayan sınıf nesnesi özellikleri tablolar halinde verilmiştir. Tablo halinde verilen bir diğer önemli bilgi ise büyük ağaç türlerinin tür özelliklerine ait özelliklerinin özetidir.

Koma vd. [22] Viyana'nın kent merkezinde (Avusturya) kaydedilen tam dalga formu havadan lazer tarama verileri kullanılarak taksonomik aile seviyesinde kent ağaçlarının nesne tabanlı sınıflandırılmasını incelemiştir. Kentsel bir çevrede ağaç ailesi sınıflandırması için sıklıkla kullanılan geometrik (taç uzunluğu, genişliği, oranı vb.) ve radyometrik (geri yansıma kesiti vb.) özellik kümelerini ele almışlardır. Bu çalışmada ormanlık alanlarda ağaçların sınıflandırılmasında başarıyla uygulanan çeşitli özellik türleri kullanılmıştır. Çalışmada iş akışı, veri ön işleme, özelliklerin araştırılması ve türetilmesi, bunu takiben sınıflandırma, özellik seçimi ve doğrulamayı içermektedir. Ön işleme aşamasının ardından geometrik ve radyometrik özelliklerden yola çıkılarak, tek ağaç segmentlerinden yedi set (F1-F7) özellik türetilmiştir. Çalışma alanına ait verilerin sınıflandırılması F1-F7 özellik setleri kullanılarak rastgele orman (Random Forest - RF) sınıflandırıcısı ile gerçekleştirilmiştir. Çalışmanın sonucunda en güvenilir sınıflandırmanın, geometrik ve radyometrik özelliklerin kombinasyonu kullanılarak %87.5'lik toplam doğrulukla elde edildiği ortaya konmuştur.

Zhang vd. [62] Washington'da güneydoğu Seattle ve Park Arboretum olmak üzere belirledikleri iki çalışma alanında hava hiperspektral görüntüler ve LiDAR verileri kullanılarak kent ormanlarının ağaç türlerinin sınıflandırılması etkinliğini araştırmışlardır. Çalışmadaki sınıflandırma prosedürü; LiDAR'dan türetilmiş örtü

yükseklik modellerinden (Canopy Height Model - CHM) nesne tabanlı segmentasyon ve görsel yorumlama (1), belirlenmiş ağaç taçlarından spektral bilgi çıkarımı ve “Minimum Noise Fraction (MNF)” dönüşümü ile spektral bant indirgeme (2), rastgele orman (Random Forest - RF) ve çoklu sınıf sınıflandırıcısı (Multi Class Classifier - MCC) kullanarak ağaç türlerinin tanımlanması (3) ve farklı algoritmalar ve veri setleri kullanılarak sınıflandırmaların doğruluk değerlendirmesi (4) şeklinde ifade edilmiştir. Gerçekleştirilen uygulama sonucunda, yedi ağaç türünün 118 bandın tümü kullanılarak genel sınıflandırma doğruluğu RO için %46.3, MCC için %79.6 olarak elde edilmiştir. MNF dönüşümü ile çıkarılan yalnızca en iyi 20 bant kullanılarak ise, hem RO hem de MCC sınıflandırmaları sonucu sırasıyla %87.0 ve %88.9'a kadar genel doğrulukta bir artış elde edilmiştir. Sınıflandırma doğruluklarından hareketle MNF band seçimi sürecinin, hiperspektral verileri kullanırken ağaç türü sınıflandırması için tercih edilen bir yaklaşım olduğu, küçük örneklem boyutlarına sahip yüksek boyutlu veri kümelerini kullanırken MCC sınıflandırıcısı daha sağlam iken, RO sınıflandırıcısının hiperspektral verilerde mevcut olan yüksek boyutluluk ve gürültüden dolayı oldukça dezavantajlı olduğu sonuçlarına varılmıştır. Çalışmanın genel sonuçları da, kentsel ormanlardaki bireysel ağaç türlerinin belirlenmesi için, ağaç taçlarının LiDAR verilerinden nesne tabanlı segmentasyonu ve hiperspektral karakterizasyonun birlikte kullanımı ile başarılı bir şekilde gerçekleştirilebileceğini göstermiştir.

Shen ve Cao [63] çalışmalarında güneydoğu Çin'in subtropikal ormanlarındaki ağaç türlerini sınıflandırmak için eşzamanlı olarak elde edilen hiperspektral görüntüler ve LiDAR verilerini kullanarak sınıflandırma için en önemli değişkenleri tespit etmeyi amaçlamış ve iki farklı veri setinin birleştirilmiş kullanımının katkısını değerlendirmek istemişlerdir. Bu amaçla, ilk olarak, bir nokta bulutu segmentasyon algoritması ile LiDAR verilerinden her bir ağaç tacını elde etmişlerdir. Ardından temel bileşen analizi (Principal Component Analysis - PCA) ve rastgele orman algoritmasına dayanan Gini indeksindeki ortalama azalmaya (MDG) göre farklı hiperspektral ve LiDAR özelliklerini sınıflandırma için seçmişlerdir. Son olarak, rastgele orman sınıflandırıcısı ile hem hiperspektral hem de LiDAR özelliklerini kullanarak beş ağaç türünü sınıflandırmışlardır. Burada

önerilen sınıflandırma yaklaşımı, çalışma alanındaki beş ağaç türünü sınıflandırmak için nispeten yüksek bir doğruluk (genel doğruluk > %85.4) sağlamıştır. Hem hiperspektral hem de LiDAR özelliklerini kullanan sınıflandırmanın, yalnızca hiperspektral özelliklerden daha yüksek doğrulukla sonuçlandığı vurgulanmıştır (genel doğruluklardaki iyileşme = %0.4 - 5.6). Sonuçlar ayrıca, ağaç türlerini etkili bir şekilde sınıflandırmak için en önemli özelliklerin daha azının da kullanılabileceğini ortaya koymuştur (genel doğruluk = %85.8 – 91).

Shi vd. [64] Almanya'da doğal bir ormanda altı farklı ağaç türünü ayırt etmek için yaprakların olduğu ve yaprakların bulunmadığı koşullarda elde edilen 37 sık kullanılan LiDAR metriğinin sınıflandırma performanslarını değerlendirmeyi amaçlamışlardır. Bu amaçla, ilk olarak LiDAR metrikleri arasındaki korelasyonu ve rastgele orman algoritması aracılığı ile belirlenen önemlerini dikkate alarak en önemli LiDAR metriklerini belirlemişlerdir. Elde edilen sonuçlar, LiDAR metriklerinin yaklaşık %60'ının birbiriyle yüksek düzeyde ilişkili olduğunu göstermiştir ( $|r| > 0.7$ ). Yaprakların olduğu ve yaprakların olmadığı durumdaki LiDAR metriklerinin kullanımı arasında ağaç türlerinin sınıflandırma doğruluğunda istatistiksel olarak anlamlı bir fark olmadığını belirlemişlerdir. Bununla birlikte, yaprakların olduğu ve yaprakların olmadığı durumdaki LiDAR metriklerini birleştirmenin, genel doğruluğu %58.2 (yaprakların olduğu) ve %62'den (yaprakların olmadığı) %66.5'e ve kapa katsayısını da 0.47 (yaprakların olduğu) ve 0.51'den (yaprakların olmadığı) 0.58'e önemli ölçüde arttırdığını vurgulamışlardır. Genel olarak radyometrik özelliklerin (özellikle yoğunlukla ilgili metrikler), ağaç türlerinin ayırımında geometrik özelliklerden daha tutarlı olduğu ve önemli katkılar sağladığını ortaya koymuş ve çalışmalarındaki sonuçlara göre hava LiDAR verilerinden türetilen metriklerin karışık ılıman bir ormandaki ağaç türlerinin ayırt edilmesine yardımcı olabileceğini kanıtlamışlardır.

## 1.2 Tezin Amacı

Bu tez çalışmasının amacı, LiDAR ile elde edilmiş ham 3B nokta bulutu verilerinden kentsel alanlardaki ağaç türlerinin otomatik olarak sınıflandırılması için makine öğrenme tabanlı bir yaklaşım geliştirmektir. Arazi ölçümleri ya da

hava fotoğraflarının yorumlanması gibi zahmetli, zaman alıcı ve pahalı yöntemler yerine LiDAR sistemi verileri kullanılarak kentsel ağaç türlerinin elde edilmesi özellikle büyük alanlarda daha hızlı ve uygun maliyetlidir. Bu çalışmada ham LiDAR verilerinin herhangi bir ön işlem gerçekleştirilmeden doğrudan nokta tabanlı yöntemlerle sınıflandırılması hedeflenmiştir. Bu amaçla, kentsel çalışma alanı olarak belirlenen İstanbul ili Esenler ve Güngören ilçelerinde yer alan Yıldız Teknik Üniversitesi, Davutpaşa Kampüsü'ne ait ham nokta bulutu verileri kullanılarak makine öğrenme algoritmaları ile geniş yapraklı (yaprak döken) ve iğne yapraklı ağaçların sınıflandırılması için bir yaklaşım geliştirilmiştir. Öncelikle belirlenen kentsel test alanlarında (Test alanı 1 ve Test alanı 2) hiyerarşik kural tabanlı sınıflandırma yaklaşımı ile yüksek bitki örtüsü sınıfı noktalarının ve yüksek bitki örtüsü noktalarının nokta tabanlı segmentasyonu sonucu bireysel ağaç taçlarının elde edilmesi hedeflenmiştir. Ardından, makine öğrenme tabanlı ağaç türleri sınıflandırması için her bir LiDAR noktası kullanılarak elde edilen iki farklı sınıflandırma özellik setinin (8 özellik ve 25 özellik) belirlenmesi ve "Grid search" yaklaşımı ile makine öğrenme algoritmalarının en uygun parametrelerinin elde edilmesi istenmiştir. En iyi sınıflandırma doğruluğunu veren parametreler kullanılarak DVM, RO ve ÇKA sınıflandırıcıları ile Test alanı 1 ve Test alanı 2'de geniş yapraklı ve iğne yapraklı kentsel ağaçları belirlemek amaçlanmıştır. Sınıflandırma yöntemlerinin doğruluğunu test etmek ve kentsel ağaç türlerinin sınıflandırılması için hangi makine öğrenme algoritmasının daha iyi sonuç verdiğini belirlemek için istatistiksel ölçümlerle performans değerlendirmesinin gerçekleştirilmesi yine bu tez çalışmasının amaçları arasında yer almaktadır. Makine öğrenme algoritmaları ile sınıflandırmanın gerçekleştirilebilmesi için türetilen geometri ve yoğunluk tabanlı özelliklerin RO sınıflandırması ile ortalama Gini'deki azalmaya (Mean Decrease in Gini - MDG) göre sınıflandırmadaki önemini belirlemek çalışmanın diğer bir hedefidir. Belirlenen kentsel çalışma alanında ağaç türlerinin sınıflandırılmasının ardından, burada geliştirilen hiyerarşik kural tabanlı sınıflandırma ve nokta tabanlı segmentasyon yaklaşımları kullanılarak, kentsel cadde ağaçlarının elde edilebilirliğinin test edilmesi çalışmanın son amacıdır.

### 1.3 Hipotez

Makine öğrenme algoritmaları ile sıklıkla ormanlık alanlarda gerçekleştirilen iğne yapraklı ve geniş yapraklı ağaç türleri sınıflandırması çalışmaları kentsel alanlarda da hızlı ve uygun maliyetli olarak LiDAR verileri kullanılarak otomatik bir şekilde gerçekleştirilebilir.

LiDAR verilerinden ağaç türlerinin sınıflandırılması için bireysel ağaç taçlarının elde edilmesi amacıyla segmentasyon aşamasında normalize edilmiş yüzey modeli (normalised Digital Surface Model - nDSM) ya da kanopi yüksekliği modeli (Canopy Height Model – CHM) kullanılması yerine doğrudan nokta tabanlı yöntemler ile elde edilmiş yüksek bitki örtüsü noktaları kullanılabilir. Yani doğrudan nokta tabanlı segmentasyon ile bireysel ağaç taçları elde edilebilir. Buna paralel olarak rastgele dağılmış 3B noktaları düzenli grid yapıya enterpole etmek yerine doğrudan 3B noktalar kullanılarak makine öğrenme tabanlı sınıflandırma ile kentsel alanlardaki ağaç türleri başarılı bir şekilde elde edilebilir.

Makine öğrenme algoritmaları ile ağaç türleri sınıflandırılmasının gerçekleştirilebilmesi için, yaygın olarak kullanılan LiDAR verilerinin dalga formu analizi sonucu elde edilen özellikleri yerine sadece geometri ve yoğunluk tabanlı sınıflandırma özellikleri kullanılarak kentsel alanlardaki geniş ve iğne yapraklı ağaçlar sınıflandırılabilir.

Sınıflandırma, en basit tanımıyla ortak özelliklere göre gruplara ayırma olarak ifade edilir. Bilinmeyen desenleri özelliklerine göre ait oldukları sınıfa atama anlamına gelen sınıflandırma, günümüzde tematik bilgi çıkarımında yaygın olarak kullanılan yöntemlerden biridir. Bu bölümde sınıflandırma yöntemleri kısaca açıklanacak, nokta bulutunun sınıflandırılmasına ve ağaç türlerinin nokta tabanlı sınıflandırılması ilişkin bilgiler verilecek, son olarak sınıflandırmanın doğruluk analizi irdelenecektir.

### 2.1 Sınıflandırma Yöntemleri

Sınıflandırma yöntemleri, piksel tabanlı sınıflandırma, nesne tabanlı sınıflandırma ve nokta tabanlı sınıflandırma olmak üzere üç gruba ayrılır [65].

Piksel tabanlı sınıflandırma yöntemlerinde, piksellerin spektral yansıma değerleri dikkate alınarak sınıflandırma gerçekleştirilir. Bu yöntemler, piksellerle ilgili mekânsal veya bağlamsal bilgileri hesaba katmadan, ilgi alanındaki her bir pikselin spektral özelliklerini analiz ederler [66]. Klasik görüntü analiz çalışmaları genellikle piksel tabanlı sınıflandırma yöntemleri kullanılarak gerçekleştirilir. En çok benzerlik sınıflandırması (Maximum Likelihood Classification - MLC) ve “ISODATA (Iterative Self-Organizing Data Analysis Technique)” en çok tercih edilen iki piksel tabanlı sınıflandırma yöntemidir [67]. Teorik olarak, piksel tabanlı görüntü sınıflandırma tekniği, kontrolsüz (unsupervised) ve kontrollü (supervised) sınıflandırma olarak ikiye ayrılır [68]. Kullanıcı müdahalesi başka bir ifade ile herhangi bir örneklem (kontrol) veri seti gerekmeden piksellerin otomatik olarak gruplandırılması kontrolsüz sınıflandırmayı ifade ederken, hedeflenen sınıflar için yeterli sayıda örneklem alanı kullanılarak gerçekleştirilen sınıflandırma kontrollü sınıflandırma olarak isimlendirilir.

Nesne tabanlı sınıflandırma yöntemlerinde, tekil pikseller yerine nesnelere sınıflandırılır [69]. Piksel tabanlı sınıflandırmadan farklı olarak, nesne tabanlı sınıflandırma yaklaşımları, görüntü segmentasyonu yoluyla anlamlı bir şekilde bir

araya getirilmiş birçok homojen pikselden oluşan nesnelere üzerinde çalışır [70]. Pksel tabanlı sınıflandırma yöntemlerinde kullanılan spektral bilgilere ek olarak, nesne tabanlı sınıflandırmada şekil ve doku bilgileri de dikkate alınır. Nesne tabanlı sınıflandırma yöntemi, birçok uygulamada piksel tabanlı sınıflandırma yöntemlerine göre daha üstün olmuştur [67, 71-72]. Bu yaklaşımın olumsuz yanı, sınıflandırmanın doğruluğunun segmentasyonun doğruluğuna bağlı olmasıdır [73].

Nokta tabanlı sınıflandırma yöntemlerinde, tekil pikseller ya da nesnelere yerine noktalar sınıflandırılır. Pksel tabanlı yaklaşımdaki piksellere benzer şekilde, sınıflandırma aşamasında en küçük birim noktalardır ve tüm noktaların ayrı ayrı bireysel özelliklerine göre sınıflandırma gerçekleştirilir [72]. Nokta bulutu sınıflandırma yaklaşımları, lazer tarama sistemleri (hava ya da yersel sistemler) ile elde edilen nokta bulutu verileri ya da fotoğraflardan çeşitli yöntemler (yoğun görüntü eşleme ya da hareket ile nesne oluşturma) kullanılarak elde edilen nokta bulutu verilerinin sınıflandırılmasında kullanılabilir.

Bu çalışma, makine öğrenme algoritmaları ile LiDAR nokta bulutunun nokta tabanlı olarak sınıflandırılmasına dayandığından, nokta tabanlı sınıflandırmaya ilişkin bilgiler bir sonraki “2.2 LiDAR Nokta Bulutunun Sınıflandırılması” bölümü içerisinde verilmiştir.

## **2.2 LiDAR Nokta Bulutunun Sınıflandırılması**

Günümüzde en önemli veri toplama tekniklerinden biri olan LiDAR, yer yüzeyine ait coğrafi referanslı nokta bulutlarını elde etmek için etkili bir araçtır ve büyük ölçekli, yüksek mekansal çözünürlüklü nokta bulutu verilerini hızlı bir şekilde elde etme avantajına sahiptir [74-75]. LiDAR sistemlerinde lazer sinyali yeryüzeyinde bulunan hedeflere çarpar ve geri yansıyan enerji kaydedilerek 3B nokta bulutu elde edilir [76]. Buna göre, noktalar farklı nesnelere ait olsalar da sistem tarafından aynı veri türü (konum, yoğunluk vb.) olarak kaydedilirler [77]. Yeryüzeyi, binalar, bitki örtüsü ve diğer doğal ya da insan yapımı nesnelere hepsi LiDAR nokta bulutu verisinde aynı sınıfta yer almaktadır [72]. Hava LiDAR verilerini kategorik nesne örneklerine sınıflandırmak, daha fazla veri işleme ve model yeniden yapılandırması için ilk ve en kritik adımdır [78]. LiDAR verilerinin

sınıflandırması, coğrafi bilgileri çıkarmak için gereklidir [79]. Nokta bulutu sınıflandırması, topografik haritalama, şehir planlama, orman izleme, enerji hattı tespiti, yol tespiti ve planlama, sayısal arazi modeli üretimi, değişim belirleme, obje çıkarımı, 3B binaların yeniden oluşturulması gibi çok çeşitli hava LiDAR uygulamalarında önemli bir rol oynar [25]. Otomatik nokta bulutu sınıflandırması ihtiyacı, hava LiDAR'ın yaygın olarak kullanılmasıyla önemli ölçüde artmıştır [78]. Ancak, 3B nokta bulutlarının ham doğası, yüksek sensör gürültüsü, eşit olmayan yoğunluk dağılımı, yoğun kentsel alanlarda güvenilir nokta etiketlemenin zorluğu, karmaşık nesne yapısı ve kentsel çevrelerde bulunan nesne sınıflarının (örneğin binalar, yollar, ağaçlar ve düşük bitki örtüsü) yüksek değişkenliği nedeniyle nokta bulutunun sınıflandırılması oldukça zorlu bir iştir [25, 80]. Bu nedenle son yıllarda, LiDAR nokta bulutunu yeryüzeyini ifade eden anlamlı sınıflara ayırmak popüler araştırma alanlarından biri olmuştur. Amerikan Fotogrametri ve Uzaktan Algılama Topluluğu (American Society for Photogrammetry and Remote Sensing - ASPRS) tarafından önerilen temel standart LiDAR sınıfları Tablo 2.1'de listelenmiştir.

**Tablo 2.1** ASPRS standart nokta sınıfları [81]

Sınıflandırma Değeri	Sınıflar
0	Oluşturulan (Created, Never Classified)
1	Sınıflandırılmamış (Unclassified)
2	Zemin (Ground)
3	Alçak bitki örtüsü (Low vegetation)
4	Orta bitki örtüsü (Medium vegetation)
5	Yüksek bitki örtüsü (High vegetation)
6	Bina (Building)
7	Alçak nokta (Low point) (Gürültü)
8	Model anahtar noktası (Model key-Point) (Mass point)
9	Su (Water)
10	ASPRS tanımı için ayrılmıştır
11	ASPRS tanımı için ayrılmıştır
12	Üst üste binen noktalar (Overlap points)
13-31	ASPRS tanımı için ayrılmıştır

LiDAR nokta bulutu verileri doğrudan sınıflandırılabilirdiği gibi sınıflandırma işlemi segmentasyon ve sınıflandırma olmak üzere iki adımda da gerçekleştirilebilir.

LiDAR verilerinin segmentasyonu ve sınıflandırılmasına sırasıyla “2.2.1 Nokta Bulutu Segmentasyonu” ve “2.2.2 Nokta Bulutu Sınıflandırması” başlıkları altında yer verilmiştir.

### **2.2.1 Nokta Bulutu Segmentasyonu**

Bir segment, aynı sınıftaki noktaların bir alanıdır [82]. Belirli homojenlik kriterlerine göre gruplama işlemi segmentasyon olarak isimlendirilir. LiDAR nokta bulutunun segmentasyonu, mimarinin 3B yeniden yapılandırılması için oldukça önemli ancak bir o kadar da zor bir adımdır. LiDAR nokta bulutlarının segmentasyonu için yıllar içinde çok sayıda algoritma geliştirilmiştir. Birçok araştırmacı, kenar tabanlı, yüzey tabanlı ve küme tabanlı segmentasyon vb. dahil olmak üzere çeşitli segmentasyon yöntemleri geliştirmeye çalışmışlardır [83]. Lazer tarayıcı verilerine özgü mevcut segmentasyon yaklaşımları, görüntü tabanlı (image-based) yaklaşımlar ve nokta bulutu tabanlı (point cloud-based) olarak gruplandırılabilir [84].

LiDAR verilerinden üretilen raster sayısal Yüzel Modellerini (YM) kullanan görüntü tabanlı algoritmalar, daha hızlı bir segmentasyon için mevcut görüntü işleme tekniklerine izin vermesine rağmen 3B nokta bulutunun raster (grid) yapıya dönüştürülmesi sırasında gerçekleşen veri kaybından etkilenirler [85].

Nokta bulutu tabanlı segmentasyon daha basit bir ifade ile nokta tabanlı segmentasyon yaklaşımlarında segmentasyon, bir nokta bulutundaki her bir bireysel noktayı etiketleme işlemidir, böylece aynı yüzeye veya bölgeye ait tüm noktalara aynı etiket verilir [83]. Nokta tabanlı segmentasyon yaklaşımları, verileri öncelikle 2B ya da 3B geometrik özellikleri kullanarak bölümlere ayırır [84]. Bir LiDAR nokta bulutunu segmentlere ayırmak, noktaları yükseklik, yoğunluk ve normallik gibi homojen özelliklere sahip farklı gruplara ayırmaktır [86]. LiDAR verilerinin segmentasyonu genellikle yakınlık ve/veya nokta dağılımının tutarlılığı kriterlerine dayalı olarak elde edilir [87-89]. İyi bir segmentasyonda segmentlerin gerçek nesnelere uyumlu olması beklenir ve noktaların yerel komşuluğu, ölçeği ve bir noktanın yerel komşuluğunun özellikleri bu tür bir uyumun sağlanması için temel öneme sahiptir [90]. Segmentasyon sonuçları doğru bölümlenme, yetersiz bölümlenme (under-segmentation), aşırı

bölümleme (over-segmentation), kaçan noktalar ve gürültüler olmak üzere kategorize edilebilir [91].

Nokta tabanlı segmentasyon için genel olarak geometrik yerleştirme (geometric fitting), bölge büyütme (region growing), kümeleme (clustering) olmak üzere üç önemli strateji benimsenmiştir [84, 92]. Bu stratejilerden yaygın olarak kullanılan kümelemenin en büyük avantajlarından biri, bölge büyütmenin aksine gruplandırmayı başlatmak için tohum (seed) noktalarına veya bölgelerine ihtiyaç duymaması; diğer önemli avantajı ise, yerleştirme temelli yöntemlerden farklı olarak, kümelemenin çok düzlemlili ve 3B cepheleri bölümlere ayırabilmesidir [92].

“Octree-based” kümeleme, “K-means” kümeleme, bulanık kümeleme ve “MeanShift” kümeleme dahil olmak üzere nokta kümelemeye dayalı yöntemler, noktalar arasındaki mesafe ve nokta yoğunluğu gibi benzerlik ölçümlerini kullanarak nokta bulutunu nokta gruplarına ayırır. Bu yöntemler kararlı sonuçlar üretebilir, ancak yanlış kümeleme algoritması kurulumu (örneğin, çekirdek genişliği parametreleri ve “MeanShift” bölümlemesinde geçerli bir bölgenin minimum nokta sayısı) nedeniyle aşırı bölümleme (over-segmentation) veya yetersiz bölümlemelere (under-segmentation) yol açabilirler [86].

Bu çalışmada ağaç tacı segmentasyonu için kullanılan “MeanShift”, “K-means” ve “DBSCAN” kümeleme algoritmaları detaylı bir şekilde üçüncü bölüm olan “Makine Öğrenmesi” bölümünde verilmiştir.

### **2.2.2 Nokta Bulutu Sınıflandırması**

LiDAR nokta bulutunun sınıflandırılması genel olarak; arazi ve araziye ait olmayan noktalar şeklinde sınıflandırma (filtreleme) ve yollar, binalar, yeşil alanlar vb. farklı objelere ayırarak sınıflandırma şeklinde oluşacak sınıflara göre ikiye ayrılabilir.

Filtreleme, arazi ve araziye ait olmayan noktalar şeklinde iki sınıfın elde edildiği ve araziye ait olmayan noktaların temizlendiği sınıflandırma yaklaşımı olarak bilinir [72, 93-94]. Filtreleme, aynı zamanda LiDAR verisinden çıplak yeryüzünü ifade eden Sayısal Yükselik Modeli (SYM) elde etmek için kullanılan bir yaklaşımdır. Morfolojik filtreler, artan filtreler, yüzey tabanlı filtreler ve segmentasyon tabanlı filtreler LiDAR nokta bulutu verilerinin filtrelenmesi için

geliştirilen çok sayıda algoritmadan bazılarıdır [95-97]. Veri yapısı, ilk gönderilen ve yansıyan verinin kullanılması, test komşuluğu ve tek seferde filtrelenen nokta yapısı, tek adım ya da iterasyon, devamsızlıkların ölçümü, değiştirme ve silme bu filtreleme algoritmalarının dayandığı temel özelliklerdir [72, 95].

Filtrelemeden farklı olarak LiDAR nokta bulutu verilerinin sınıflandırılması sonucu bina, yol, çıplak yeryüzü, ağaç vb. gibi çok sayıda farklı sınıf elde edilebilmektedir. Nokta bulutları sınıflandırmaya tabi olduğunda nokta bulutlarının raster yapıya (gride örneklenmiş) dönüştürülüp sınıflandırılması ya da doğrudan nokta tabanlı sınıflandırılması olmak üzere iki yöntem mevcuttur [27, 98]. LiDAR verilerinin grid tabanlı ve nokta tabanlı sınıflandırılması sırasıyla “2.2.2.1 Grid Tabanlı Sınıflandırma” ve “2.2.2.2 Nokta Tabanlı Sınıflandırma” alt başlıklarında incelenmiştir.

#### **2.2.2.1 Grid Tabanlı Sınıflandırma**

Nokta bulutlarının sırasız ve yapılandırılmamış (unstructured) doğası nedeniyle, ilk uygulamalar esas olarak 3B nokta bulutlarını düzenli 2B yapıya dönüştürmeye çalışmışlardır. Grid tabanlı sınıflandırma yaklaşımları, 2B yapıya enterpole edilmiş LiDAR verilerini kullanır ve bu verilerde her bir piksel X, Y, Z koordinat bilgisinin yanında, yoğunluk, çoklu dönüş sayısı, yansıyan sinyalin genişliği, geri yansıma katsayısı gibi tanımlayıcı bilgiler içerir [98]. LiDAR verisi ile multispektral ya da hiperspektral görüntülerin birlikte kullanıldığı çalışmalarda grid tabanlı sınıflandırma yaklaşımları kullanılır [99].

Gride örneklenmiş LiDAR verilerinin kullanıldığı grid tabanlı sınıflandırma yaklaşımlarının en büyük avantajı görüntü işleme algoritmalarının doğrudan kullanıma olanak sağlamasıdır [72]. Böylece büyük verilerle bile hızlı bir şekilde sınıflandırma uygulamalarının gerçekleştirilmesi mümkün olmaktadır. Fakat, özellikle yüksek nokta yoğunluklu LiDAR verilerinde bir pikselin içine birden çok noktanın düşme olasılığına rağmen bir piksel bu noktaların özelliklerinin ortalama değerlerini ifade eder yani o alanda sadece bir nokta için grid tabanlı sınıflandırmada işlem yapılır. Benzer şekilde, bir pikselin içinde düşey olarak farklı sınıflara ait çoklu yansımalar bulunmasına rağmen, grid tabanlı sınıflandırmada bir piksel sadece bir sınıfı temsil eder [72, 98]. Bu nedenlerden dolayı ham LiDAR

nokta bulutunun grid yapıya enterpole edilmesi önemli derecede veri kaybına yol açabilir [100].

Bu çalışmada, nokta tabanlı algoritmalar kullanılarak LiDAR verilerinin sınıflandırılması gerçekleştirildiği için grid tabanlı yaklaşıma çok değinilmemiştir. Görüntü işleme algoritmalarının doğrudan kullanıldığı grid tabanlı sınıflandırmaya ilişkin detaylı bilgiler Uzar Dinlemek [65]'de yer almaktadır.

### **2.2.2.2 Nokta Tabanlı Sınıflandırma**

Grid tabanlı sınıflandırma yaklaşımından farklı olarak LiDAR nokta bulutu için nokta tabanlı sınıflandırma yöntemi, her bir tekil lazer noktasını ait olduğu nesne sınıfına atamayı amaçlar. Nokta tabanlı sınıflandırma yaklaşımları, ham (unstructured) 3B LiDAR noktalarının herhangi bir ön işleme ihtiyaç duymadan ve buna paralel olarak herhangi bir veri kaybına uğramadan doğrudan sınıflandırılmalarını sağlar. LiDAR nokta bulutlarının veri boyutuna bağlı olarak hesaplama maliyetinin pahalı ve işlem süresinin uzun olmasına rağmen yayılan her lazer sinyalinin çoklu dönüşlerinin, sinyallerin çarptığı materyale göre farklı sınıflara atanabilmesi nokta tabanlı sınıflandırma yaklaşımının en büyük yararlarıdır.

LiDAR nokta bulutu verilerinin sınıflandırılması için doğrudan noktaların kullanıldığı (nokta tabanlı) bir çok farklı kural tabanlı ve makine öğrenme tabanlı sınıflandırıcılar (DVM, RO, ÇKA vb.) mevcuttur [101]. Nokta tabanlı bu sınıflandırıcılar, 3B LiDAR nokta bulutundaki her bir tekil noktadan hesaplanan farklı özellikleri kullanarak sınıflandırma işlemini gerçekleştirirler [98, 102-104]. Nokta tabanlı sınıflandırma için, öncelikle hedeflenen sınıflara göre kullanılacak özellikler belirlenir ve genellikle eğitim aşamaları sırasında sınıflandırmada kullanılacak özelliklerin en uygun parametreleri belirlenerek sınıflandırma işlemi gerçekleştirilir [72]. Sınıflandırma algoritmalarının temelini oluşturan bu özellikler; geometrik özellikler (spatial-based features), dönüş tabanlı özellikler (echo-based features) ve dalga formuna dayanan özellikler (waveform-based features) olmak üzere üç gruba ayrılabilir [104].

Geometrik özellikler, sınıflandırma algoritmalarında yaygın olarak kullanılan özelliklerdendir. Bu özellikler, hem geleneksel kesikli-dönüslü hem de daha yeni teknoloji olan tam dalga boyu kaydeden (full-waveform) LiDAR sistemleri ile

kaydedilen verilerden hesaplanabilirler. Yerel çevre (the local environment), yükseklik (height features), zemin seviyesinden yükseklik (hight from ground level), özdeğerler (eigenvalues), yerel düzlem özellikleri (local plane features), yüzey tabanlı özellikler (surface-based features), dış bükey kabuk tabanlı özellikler (convex hull-based features) ve düşey profil özellikleri (vertical profile features) en önemli geometrik özelliklerdendir [98, 102-104].

Dönüş tabanlı özellikler, LiDAR sistemlerinin çoklu dönüşleri kullanılarak belirlenir. İlk dönen (first return), ortada dönen (intermediate return), son dönen (last return) ve tekli dönen (single return) LiDAR noktalarının farklı birleşimleri kullanılarak elde edilen zemin dönüşü, bina dönüşü, bitki örtüsü dönüşü ve elektrik hattı dönüşü dönüş tabanlı özelliklerdendir [103]. Bir yüzeyden yansıyan LiDAR nokta sayısına bağlı olarak belirlenen nokta yoğunluğu (point density) ve yoğunluk oranı (density ratio) diğer dönüş tabanlı özelliklerdendir [103].

Dalga formuna dayanan özellikler, sadece tam dalga boyu kaydeden LiDAR sistemleri ile elde verilerden hesaplabilen özelliklerdir. Dönüş genliği (echo amplitude), dönüş genişliği (echo width), dönüş şekli (echo shape) ve dönüş enine kesiti (echo cross-section) dalga formuna dayanan sınıflandırma özelliklerindendir [102].

LiDAR nokta bulutlarının nokta tabanlı sınıflandırılmasında kullanılan sınıflandırma özellikleri ve geometrik ve dönüş tabanlı özelliklerin kullanıldığı hiyerarşik kural tabanlı sınıflandırmaya ilişkin detaylı bilgiler Erişir [72]'de yer almaktadır. LiDAR nokta bulutu tam dalga formu analiz yaklaşımları ve dalga formuna dayanan özelliklerin kullanıldığı LiDAR nokta bulutunun sınıflandırılmasına ilişkin detaylı bilgiler ise Kurtak ve Yastıklı [105] ve Kurtak [106]'da bulunmaktadır.

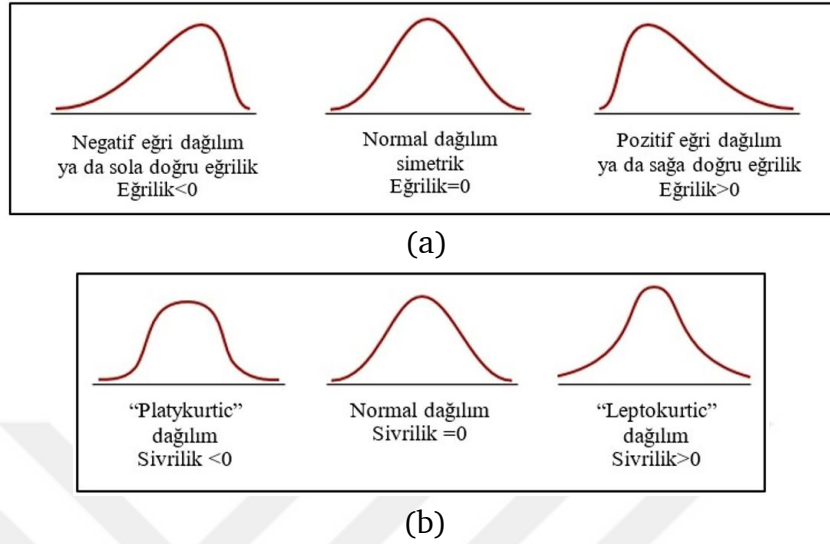
Bu çalışmada LiDAR verilerinden elde edilen bireysel ağaç taçlarının ağaç türlerine sınıflandırılması aşamasında kullanılan nokta tabanlı destek vektör makineleri (Support Vector Machine – SVM), rastgele orman (Random Forest – RF) ve çok katmanlı algılayıcı (Multi-Layer Perceptron – MLP) makine öğrenme algoritmalarına ilişkin bilgiler ise bir sonraki bölüm olan “3. Makine Öğrenmesi” bölümünde verilmiştir.

## 2.3 LiDAR Nokta Bulutu Verilerinden Ağaç Türlerinin Nokta Tabanlı Sınıflandırılması

LiDAR nokta bulutu verileri kullanılarak ağaç türlerinin sınıflandırılması bir önceki bölümde bahsedilen nokta tabanlı sınıflandırma algoritmaları kullanılarak gerçekleştirilebilir. Ağaç türleri sınıflandırmasında kilit nokta, her bir ağacı bireysel olarak tanımlayan önemli özelliklerin (features) bulunmasıdır [23]. Taç içerisinde birçok LiDAR veri noktası kullanılarak karakterize edilen bir ağacın yapısal özellikleri, ağaç türlerini doğru bir şekilde tanımlama yeteneğini artırabilir fakat buradaki zorluk, büyük LiDAR verilerinden çok sayıda tür arasında temel tanımlama özelliklerini çıkarmak ve seçmektir [16]. İstatistiksel analizler ile LiDAR nokta bulutunun geometrik, dönüş ya da dalga formu verilerine dayalı olarak hesaplanan çeşitli özellikleri kullanılarak sınıflandırma gerçekleştirilir. LiDAR verileri kullanılarak gerçekleştirilen ağaç türleri sınıflandırmasında yaygın olarak kullanılan bazı özellikler aşağıdaki gibidir [8, 15-16, 18-20, 23, 56, 59, 62]:

- Ağaçların taç özellikleri (taç uzunluğu, taç genişliği, taç uzunluk-genişlik oranı, ağaç tepesinin sivriliği, ağaç tepesinin simetrisi, taç hacmi, normalize edilmiş hacim vb.),
- Nokta yoğunluğu ve nokta dağılımları (ortalama nokta yoğunluğu, ortalama nokta sıklığı vb.),
- Yoğunluk (intensity) özellikleri (ortalama yoğunluk değeri, yoğunluğun standart sapma değeri, maksimum yoğunluk, yoğunluğun simetri derecesi (eğrilik katsayısı - skewness), yoğunluğun zirve derecesi (sivrilik katsayısı - kurtosis), belli bir dalga boyundaki yoğunluk değeri vb.),
- Noktaların yüksekliklerinden türetilen özellikler (maksimum yükseklik, ortalama yükseklik, yükseklik değişimi, yüksekliğin standart sapması, bağlı ortanca yükseklik, yüksekliğin simetri derecesi, yüksekliğin zirve derecesi vb.),
- Lazer sinyal özellikleri (sinyal genliği, sinyal genişliği, ortalama sinyal genliği, ortalama sinyal genişliği, maksimum genlik, dalga formu mesafesi, dönen dalga formu enerjisi vb.),
- Çoklu dönüş özellikleri (ilk dönen, ortada dönen, son dönen oranları, dalga formu başına düşen dönüş sayısı vb.).

Şekil 2.1’de değişken dağılımındaki simetri derecesi (eğrilik katsayısı - skewness) ve bir dağılımın zirve yaptığı ölçü (sivrilik katsayısı - kurtosis) değişkenleri görülmektedir.



**Şekil 2.1** Değişken dağılımındaki simetri derecesi (eğrilik katsayısı - skewness) (a) ve bir dağılımın zirve yaptığı ölçü (sivrilik katsayısı - kurtosis) (b)

#### 2.4 Sınıflandırmanın Doğruluk Analizi

Doğruluk analizi (accuracy analysis) sınıflandırma gibi bilgi çıkarımı uygulamalarının önemli bir parçasıdır ve sonuç ürünlerin kalitesini belirler. Dokuların gerçekte ait oldukları sınıfa atanması sınıflandırmada doğruluğu, yanlış sınıflara atanması ise sınıflandırma hatasını ifade eder [72]. Sınıflandırma uygulamalarında sonuçların kalitesini değerlendirmek için kullanılan en yaygın yaklaşım doğruluğun istatistiksel ölçümüdür [107]. Tüm doğruluk değerlendirmeleri doğruluk analizi örneklerinin tanımlanması, bu örnekler için yer gerçekliği verilerinin toplanması ve sonuçların analizi olmak üzere üç temel adımı içerir [108]. Özellikle multispektral görüntülerin sınıflandırma doğruluklarının değerlendirilmesi için kullanılan bazı yöntemler hata matrisi (error (confusion) matrix), içerme hatası (error of commission), dâhil etmeme hatası (error of omission), üretici doğruluğu (producer’s accuracy), kullanıcı doğruluğu (consumer’s (user’s) accuracy), genel doğruluk (overall accuracy) ve Kappa katsayısı (Kappa coefficient) olarak ifade edilebilir [109]. Bunlara ek olarak, nesne ile arka planı arasındaki farkı belirlemek için bütünlük (completeness), doğruluk (correctness) ve kalite (quality) gibi doğruluk

değerlendiricileri de mevcuttur [110]. Bütünlük ve doğruluk değerlendiricileri sırasıyla üretici ve kullanıcı doğruluğu olarak da ifade edilebilir [72, 110].

Hata matrisi, öngörülen ve gerçek sınıf etiketlerinin basit bir çapraz tablosudur [111]. Başka bir ifade ile satır ve sütunlar halinde düzenlenen hata matrisi, sınıflandırma sonucu elde edilen veriler ile yer gerçeği (ground-truth) verilerini birlikte içerir [72]. Hata matrisinde sütunlar genellikle referans veriyi, satırlar ise sınıflandırma sonuçlarını ifade eder [108]. Genel doğruluk, üretici doğruluğu ve kullanıcı doğruluğu gibi daha önce ifade edilen doğruluk değerlendiricilerinin tümü hata matrisi kullanılarak hesaplanır. İkili değişkenli (binary) bir hata matrisi örneği Tablo 2.2’de verilmiştir.

**Tablo 2.2** İkili değişkenli hata matrisi [112]

		Yer gerçeği verisi	
		Doğru (True)	Yanlış (False)
Sınıflandırma sonucu	Doğru (True)	TP	FP
	Yanlış (False)	FN	TN

Tablo 2.2’deki TP, TN, FP ve FN sırasıyla “True Positive”, “True Negative”, “False Positive” ve “False Negative” örneklemleri ifade etmektedir. Buradaki TP, yer gerçeği verilerine göre doğru olarak sınıflandırılan varlıkları ifade ederken TN, sınıflandırmada elde edilemeyen ve aynı zamanda yer gerçeği verilerinde de bulunmayan varlıkları tanımlar. FP, sınıflandırma ile elde edilen ancak yer gerçeği verilerinde bulunmayan ve FN ise sınıflandırmada elde edilmeyen yer gerçeği verilerinde doğru olarak tanımlanan varlıkları ifade eder [101].

İçerme hatası (error of commission) “False positive” lerin bir ölçüsü iken, dâhil etmeme hatası (error of omission) ise “False negative” lerin bir ölçüsüdür. Köşegen değerleri hariç, içerme hataları hata matrisinin satırlarında gösterilir. Dahil etmeme hataları ise köşegen değerleri hariç hata matrisinin sütunlarında gösterilir.

Doğruluk (accuracy), hassasiyet (recall), kesinlik (precision) ve F1-skor da TP, TN, FP ve FN kullanılarak hesaplanan ve özellikle makine öğrenme algoritmaları ile sınıflandırma uygulamalarında yaygın olarak kullanılan doğruluk

değerlendiricileridir [101]. Buradaki hassasiyet sıklıkla kullanıcı doğruluğu, kesinlik ise üretici doğruluğu olarak adlandırılır [113]. Doğruluk, doğru sınıflandırılmış örneklerin tüm örneklere oranını gösterir (Eşitlik 2.1). Hassasiyet, sınıflandırma algoritması tarafından doğru olarak tahmin edilen pozitif örneklerin, doğru olarak tanınması gereken toplam örnek sayısına oranını değerlendirmek için kullanılır (Eşitlik 2.2). Kesinlik, doğru sınıflandırılmış pozitif örneklerin toplam tahmin edilen pozitif örneklere oranını belirlemek için kullanılır (Eşitlik 2.3). F1-skor, hassasiyet ve kesinlik değerlerinin harmonik ortalaması alınarak hesaplanır (Eşitlik 2.4). Doğruluk, hassasiyet, kesinlik ve F1-skor'un potansiyel değerleri 0 ile 1 arasındadır ve 1'e yakın değerler daha iyi bir sınıflandırma performansını tanımlarken ve 0'a yakın değerler daha düşük sınıflandırma sonuçlarını ifade eder [114]. Doğruluk, sınıflandırma sonuçlarının toplam doğruluğunu ifade ederken her bir sınıfın ayrı ayrı sınıflandırma performansı hakkında bilgi vermez. Hassasiyet, kesinlik ve F1-skor değerleri ise her bir sınıf için ayrı ayrı hesaplanan doğruluk değerlendiricileridir.

$$\text{Doğruluk (Accuracy)} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2.1)$$

$$\text{Hassasiyet (Recall)} = \frac{TP}{TP+FN} \quad (2.2)$$

$$\text{Kesinlik (Precision)} = \frac{TP}{TP+FP} \quad (2.3)$$

$$F1 - \text{skor} = 2 \times \frac{\text{Hassasiyet} \times \text{Kesinlik}}{\text{Hassasiyet} + \text{Kesinlik}} \quad (2.4)$$

Bu bölümde ifade edilen doğruluk analiz değerlendiricilerinin hiçbiri nokta bulutu verisinin nokta tabanlı sınıflandırma performansını değerlendirmek için tasarlanmamıştır [72]. Fakat, hata matrisi elemanları olarak pikseller yerine noktalar kullanılarak istatistiksel yöntemler ile nokta tabanlı sınıflandırmaların performansları hakkında bilgi sahibi olunabilir [115].

Verileri akıllı eyleme dönüştürmek için bilgisayar algoritmalarının geliştirilmesiyle ilgilenen çalışma alanı makine öğrenmesi (Machine Learning – ML) olarak bilinir [116]. Bu bölümde makine öğrenmesinin temellerinden bahsedilecek, makine öğrenmesi türlerine ve bunun alt başlığında kullanılan makine öğrenmesi algoritmalarına detaylı bir şekilde değinilecektir. Bölümün sonunda ise bir model değerlendirme yöntemi olan çapraz doğrulama yaklaşımına yer verilecektir.

### 3.1 Makine Öğrenmesinin Temelleri

Öğrenme, makinelerin de önemli bir yönü haline getirilmiş doğal bir insan davranışıdır [117]. Makine öğrenmesi (machine learning), insanların yeteneklerini makineler tarafından taklit etmeyi amaçlayan önemli bir yapay zekâ (artificial intelligence – AI) dalıdır [118-119]. Makine öğrenimi, verilerden öğrenebilen, deneyimlerden bilgi edinebilen ve zaman içinde öğrenme davranışlarını iyileştirebilen algoritmaların tasarımı ve geliştirilmesi için zekâyı anlamakla ilgilenir [120].

Geleneksel bir makine öğrenmesi işlem adımları sırasıyla; problemi tanımlamak, eğitim ve test için veri toplanması, verileri tanımlayan özelliklerin belirlenmesi, analizler sonucu en uygun parametreleri belirleyerek modeli eğitmek ve eğitilmiş modelin performansını değerlendirmek için görünmeyen test verileri kullanılarak modelin test edilmesi olarak ifade edilebilir [121].

Makine öğrenmesi, maliyet ve süre temel kaygı olduğunda veya problemin tam olarak incelenemeyecek kadar karmaşık görüldüğü durumlarda geleneksel mühendislik akışına verimli bir alternatif sunabilir [122]. Bir makine öğrenmesi yönteminin performansı, üzerinde eğitildiği etiketlenmiş verilere bağlıdır. Dünyadaki veri miktarı hızla artarken, bu verileri etiketlemek ise genellikle maliyetlidir [123].

Makine öğrenmesi arařtırmaları son birkaç yıldır son derece yaygındır ve buna baęlı olarak bir uygulayıcı için kullanımı oldukça kolay olan çok sayıda doęru ve verimli algoritma ortaya çıkmıřtır [119]. Günümüzde makine öğrenimi en hızlı büyüyen teknik alandır ve akıllı saęlık, akıllı fabrika (Endüstri 4.0) vb. gibi birçok kullanım alanına sahiptir. Objeler tanıma, konuşma tanıma, görüntü (yüz) tanıma, öneri sistemleri, obje çıkarımı, otonom sürüş, semantik analiz, borsa analizi, kredi puanlama, bilgi güvenlięi vb. makine öğrenmenin kullanıldığı yaygın uygulamalardandır [120, 124-125].

Makine öğrenimi, karmařık içerięi sınıflandırmak için kullanılacak güçlü bir matematiksel araçtır. Makine öğrenimi algoritmalarında, önceden tanımlanmış keyfi kurallar ve güçlü varsayımlara dayalı parametreler olarak tanımlanmak yerine, sınıflandırma kuralları eğitim verileri kullanılarak otomatik olarak öğrenilir [78]. Makine öğrenimi ve otomatik özellik seçimi sayesinde, geleneksel sınıflandırma metodolojisinin doęasında bulunan zahmetli tasarım ve programlama çalışmalarının çoęundan kaçınılır. Bu nedenle, birden çok nesne türünden oluşan karmařık arazi yapısında, geleneksel sınıflandırma yöntemlerinden daha etkili olmaları nedeniyle makine öğrenme yöntemleri daha uygundur [78].

Makine öğrenme algoritmaları geliřtirmek için Scikit-learn, TensorFlow, Theano, Keras, Caffe, Torch, Deep Learning 4j, MxNet, CNTK, Lasagne, BigDL ve benzeri birçok çerçeve kullanılmaktadır [126].

Makine öğrenmesi yapay zekânın alt alanı iken, derin öğrenme ise makine öğrenmesinin bir alt kümesidir [117, 126]. Yapay zekâ, makine öğrenmesi ve derin öğrenme iliřkisi Şekil 3.1'de verilmiştir. Derin öğrenme, makine öğrenme arařtırmalarında yeni bir alandır. Motivasyonu, insan beyninin analiz ve öğrenme sinir aęını kurmak ve simüle etmektir. Derin öğrenme görüntü, ses ve metin gibi verileri yorumlamak için insan beynini taklit eder [118]. Derin öğrenme çok sayıda katman ve parametre içeren bir sinir aęıdır [117].



**Şekil 3.1** Yapay zekâ, makine öğrenmesi ve derin öğrenme ilişkisi [117]

### 3.2 Makine Öğrenmesi Yöntemleri

Makine öğrenmesi yöntemleri temel olarak üç ana kategoriye ayrılabilir [122, 125-126]:

- Denetimsiz öğrenme (unsupervised learning),
- Denetimli öğrenme (supervised learning),
- Pekiştirmeli öğrenme (reinforcement learning).

#### 3.2.1 Denetimsiz Öğrenme

Denetimsiz öğrenme, genellikle gizli düzenlilikleri (örneğin kümeler) ortaya çıkarmaya veya verilerdeki anormallikleri (örneğin bazı olağandışı makine işlevleri veya bir ağa izinsiz giriş) tespit etmeye çalışır [119]. Denetimsiz öğrenme, etiketlenmemiş, sınıflandırılmamış veya kategorize edilmemiş test verilerinden öğrenme görevidir [126]. Denetimsiz öğrenmede, eğitim seti etiketlenmemiş girdilerden, yani herhangi bir atanmış istenen çıktısı olmayan girdilerden oluşur [122]. Denetimsiz öğrenmenin ana uygulamaları küme analizi (clustering analysis), temel bileşen analizi (principal component analysis), vektör niceleme (vector quantization) ve kendi kendine organizasyon (self-organization) alanlarındadır [126].

Bu çalışmada gerçekleştirilen segmentasyon aşamasında kullanılan kümeleme algoritmaları (“MeanShift”, “K-means”, “DBSCAN”) bu bölümde açıklanmıştır.

##### 3.2.1.1 “MeanShift”

1975 yılında Fukunaga ve Hostetler [127] tarafından önerilen “MeanShift”, her bir veriyi yerel maksimum yoğunluk fonksiyonuna göre kaydıran parametrik olmayan, yinelemeli bir prosedürdür [128]. Yöntem, ilk olarak küme merkezi olarak veri setinden rastgele bir nokta seçer [129]. “MeanShift” belirli bir

bölgedeki noktaların ortalaması olacak şekilde, merkezler için adayları güncelleyerek çalışan, merkez tabanlı bir algoritmadır. Algoritma, aranacak bölgenin boyutunu belirleyen bir parametre olan bant genişliğine (bandwidth) göre kümelerin sayısını otomatik olarak ayarlar. Bant genişliği manuel olarak ayarlanabilir, ayarlanmadığında öngörülen bant genişliği fonksiyonu kullanılarak tahmin edilebilir.

### **3.2.1.2 “K-means”**

“K-means” kümeleme [130] genellikle bir veri kümesini otomatik olarak bölümlenmek için uygulama kolaylığı, basitlik, verimlilik ve deneysel başarı açısından en yaygın hiyerarşik algoritmadır [131-132]. “K-means” algoritması, k sayısındaki ilk küme merkezlerini seçerek her bir veri noktasını en yakın kümeye tahsis eder ve her bir kümenin ortalama değerini hesaplayarak küme merkezlerini günceller [133]. Algoritma tekrarlı bir şekilde veri noktalarının merkeze uzaklıklarını inceleyerek değişim olmayıncaya dek işlemi sürdürür. Bu algoritma veriye ilişkin belirlenecek küme sayısını bilmeyi gerektirir.

### **3.2.1.3 “DBSCAN”**

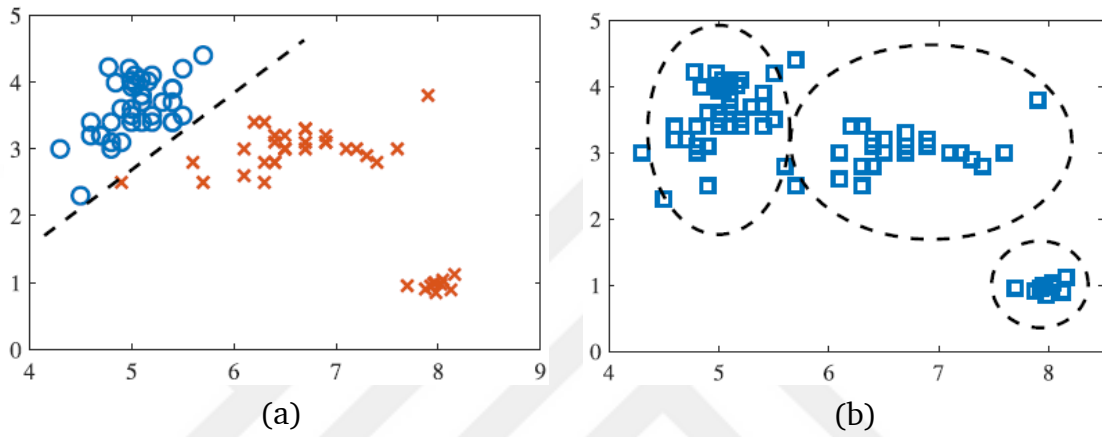
“DBSCAN (Density-Based Spatial Clustering of Applications with Noise)”, parametre olarak küme sayısı gerektirmeden keyfi olarak şekillendirilmiş kümeleri ve gürültüleri belirleyebilen kümeleme algoritmalarından biridir (“K-means” gibi diğer kümeleme algoritmalarının aksine) [134-135]. “DBSCAN” algoritması, kümeleri üç eleman ile ifade eder; ilki kümenin çekirdekleri (cores) veya tohum (seed) üyeleri, ikincisi kümenin sınırları ve üçüncüsü ise kümede hiçbir etkisi olmayan gürültüden veya aykırı değerlerden oluşur [136]. “DBSCAN” algoritmasının ana fikri, her bir çekirdek noktasının belirli bir yarıçapta belirli sayıda komşuya sahip olması gerekliliğine dayanmaktadır [137]. Algoritma rastgele bir nokta ile başlar, bu noktanın belirli komşuluğundaki noktalar alınır ve komşulukta yeterli yoğunlukta nokta varsa, bir küme başlatılır; aksi halde, nokta gürültü olarak etiketlenir. İşlem yinelemeli olarak devam eder [137].

## **3.2.2 Denetimli Öğrenme**

Denetimli öğrenmede her örnekle ilişkili bir etiket vardır. Örnekle ilgili bir sorunun cevabı olması gerektiği varsayılır [119]. Denetimli öğrenmede, eğitim seti

girdi ve istenen çıktı çiftlerinden oluşur ve amaç girdi ve çıktı uzayları arasında bir eşlemeyi öğrenmektir [122, 126]. Denetimsiz ve denetimli öğrenme farkı Şekil 3.2'de verilmiştir.

Bu çalışmada gerçekleştirilen sınıflandırma aşamasında kullanılan algoritmalar (destek vektör makineleri, rastgele orman, çok katmanlı algılayıcı) bu bölümde detaylı bir şekilde açıklanmıştır.



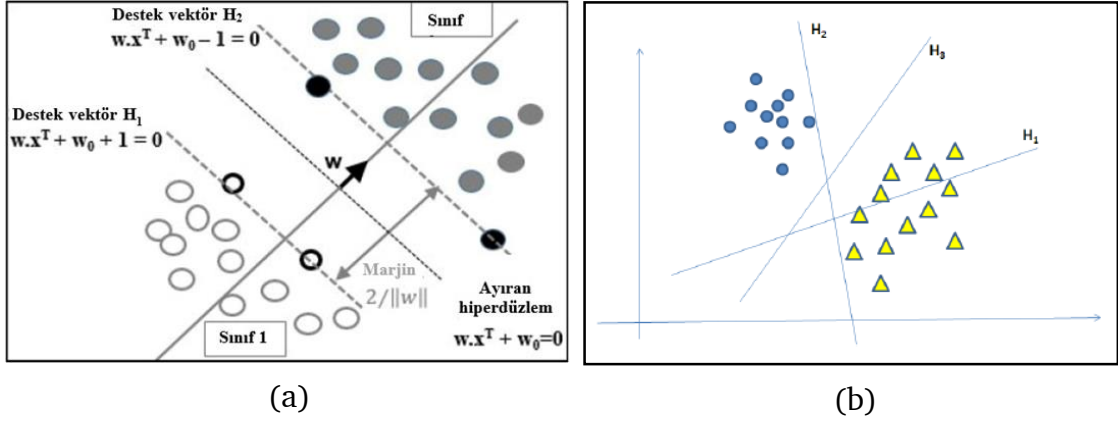
Şekil 3.2 Denetimli öğrenme (a) ve denetimsiz öğrenme (b) [122]

### 3.2.2.1 Destek Vektör Makineleri (DVM)

Destek vektör makineleri (Support Vector Machines - SVMs) sınıflandırma, regresyon ve aykırı değerlerin tespiti için kullanılan bir dizi denetimli öğrenme yöntemidir [138-139]. Destek vektörü makinesi, gruplar veya gözlem sınıfları arasındaki optimum ayırıştırma hiper düzlemini (hyperplane) hesaplar [140]. Hiper düzlem eğitim yoluyla oluşturulan modeli temsil eden denklemdir (Eşitlik 3.1). Destek vektörleri, hiper düzleme en yakın olan veri noktalarıdır. Bu noktalar, kenar boşluklarını hesaplayarak ayırma çizgisini daha iyi tanımlar. Ayırım mesafesi (margin), en yakın sınıf noktalarındaki iki çizgi arasındaki boşluktur. Bu, vektörleri veya en yakın noktaları desteklemek için çizgiden dik mesafe olarak hesaplanır (Eşitlik 3.2). Şekil 3.3'te destek vektör makineleri algoritması ve en uygun hiper düzlem belirleme örnekleri görülmektedir.

$$\text{Hyperplane } \mathcal{H} = \{x | \langle \omega, x \rangle + b = 0\} \quad (3.1)$$

$$\gamma_S(\omega, b) = \min_{x_i \in S} d_{x_i}(\omega, b) \quad (3.2)$$

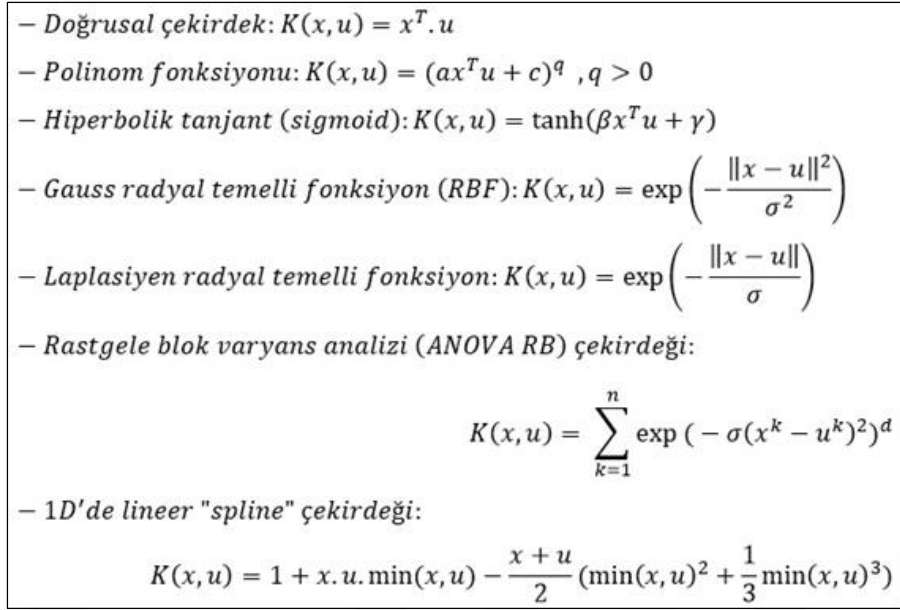


**Şekil 3.3** Destek vektör makineleri [140] (a) ve en uygun hiper düzlemin belirlenmesi [138] (b)

Bir sorun girdi alanında doğrusal olarak ayrılmazsa, yumuşak kenarlı (soft-margin) DVM, yanlış sınıflandırılmış veri noktalarının sayısını en aza indiren ve iyi genelleyen sağlam bir ayırma hiper düzlemi bulamaz. Bunun için, verilerin doğrusal olarak ayrılabilceği çekirdek alan olarak adlandırılan daha yüksek boyutlu bir alana dönüştürmek için bir çekirdek (kernel) kullanılabilir [139]. Sıklıkla kullanılan çekirdek fonksiyonları aşağıdaki şekilde tanımlanabilir [141-144]:

- Lineer çekirdek (Linear kernel),
- Radyal temelli fonksiyon çekirdeği (Radial basis function (rbf) kernel),
- Polinomal çekirdek (Polynomial kernel),
- Sigmoid çekirdek (Sigmoid kernel),
- Normalize edilmiş kernel (Normalization kernel),
- Laplasiyen radyal temelli fonksiyonu (Laplacian radial basis function),
- Rastgele blok varyans analizi çekirdeği (Randomized blocks analysis of variance (ANOVA RB) kernel)

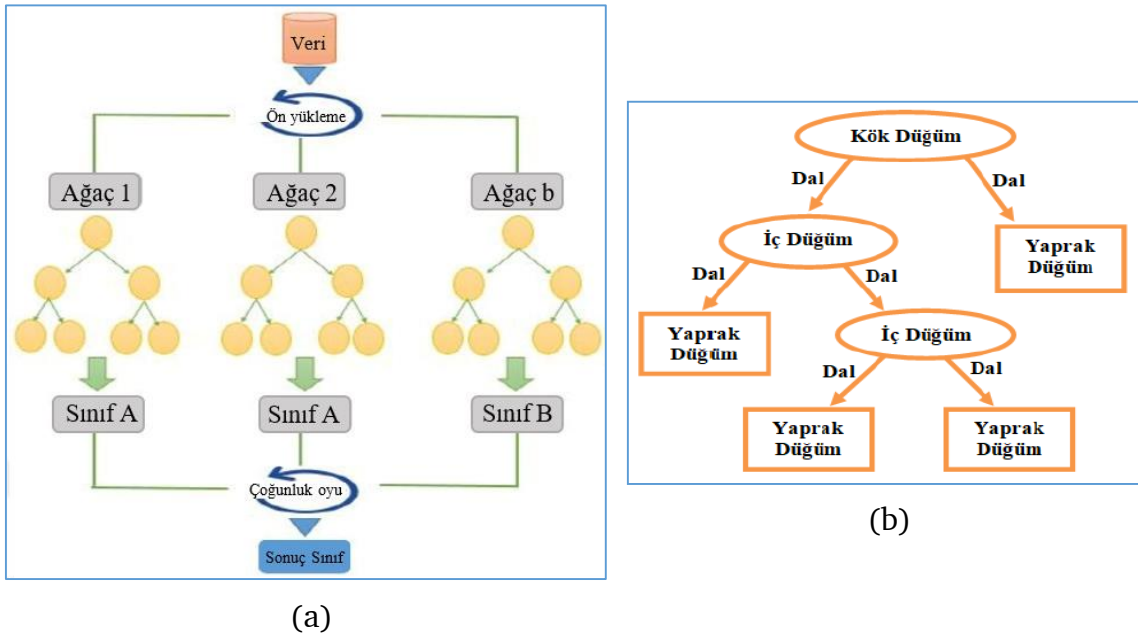
Daha açık bir ifade ile çekirdek numarası (kernel trick) ayrıştırılmayan bir problemi, daha fazla boyut ekleyerek ayrılabilir problemlere dönüştürür. Çekirdek numarası doğrusal olmayan ayırma probleminde en faydalı olan yöntemdir ve daha doğru bir sınıflandırıcı oluşturmaya yardımcı olur. Bazı çekirdek eşitlikleri Şekil 3.4'de yer almaktadır.



Şekil 3.4 Bazı çekirdek eşitlikleri [139]

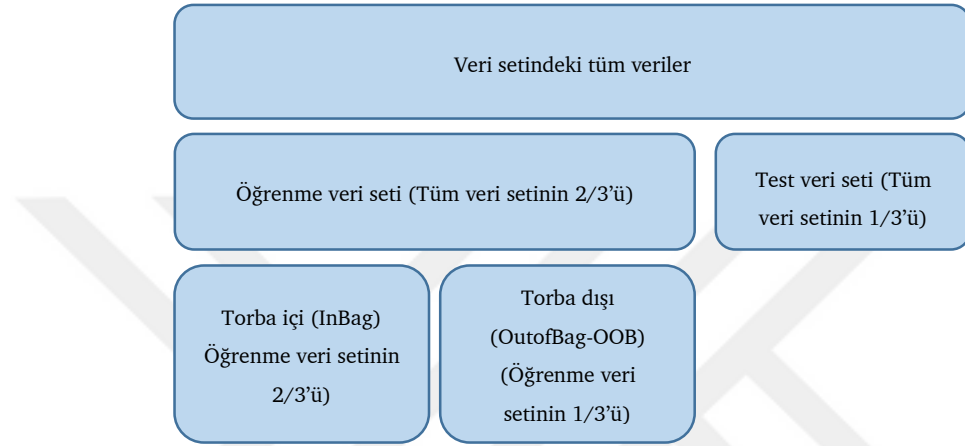
### 3.2.2.2 Rastgele Orman (RO)

Rastgele orman [145], karar ağacı sınıflandırmasına dayanan parametrik olmayan bir makine öğrenme algoritmasıdır. Karar ormanının oluşturulması esnasında kök düğüm, iç düğüm, dal ve yaprak düğüm gibi yapılardan oluşan bireysel ağaçların geliştirilmesinde rastgele vektörler kullanılır. Şekil 3.5'de RO algoritması ve bir karar ağacı örneği verilmiştir.

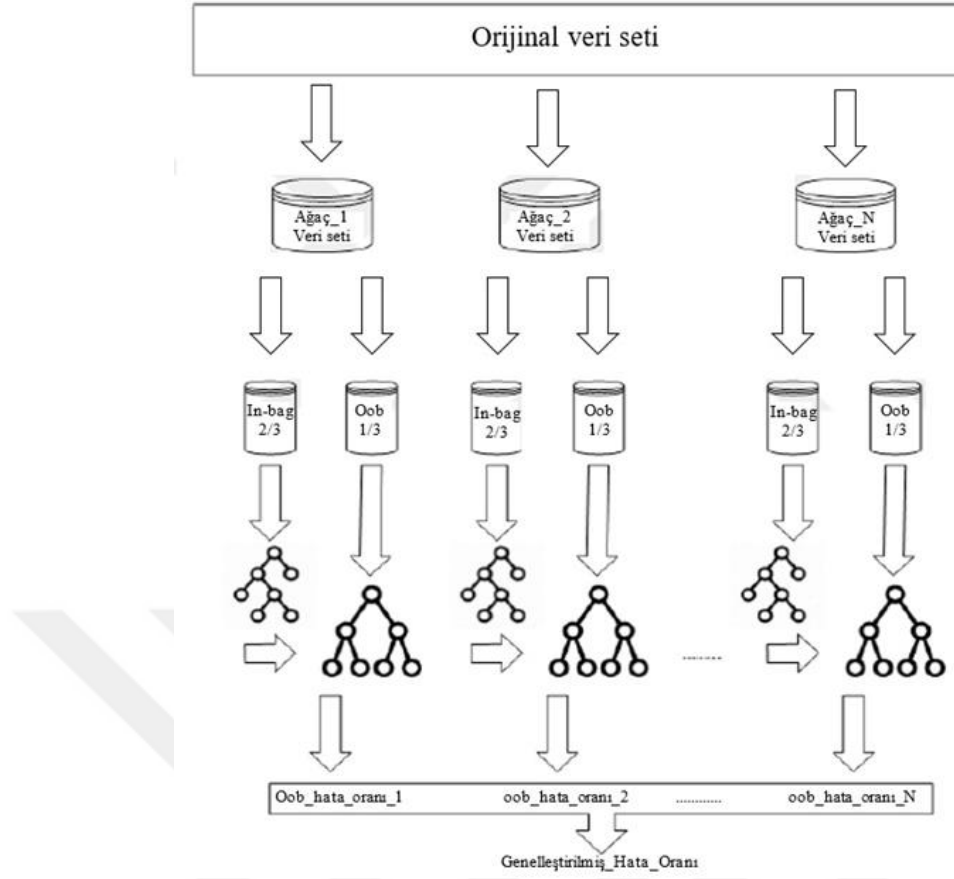


Şekil 3.5 RO algoritması [146] (a) ve bir karar ağacı örneği [147] (b)

Her karar ağacının inşası için, orijinal veri setinden ayrı bir rastgele önyükleme (bootstrap) örneği alınır [148]. Birden çok karar ağacı oluşturmak için örneklerin belirlenmesinde torbalama (bagging) yaklaşımı kullanılır [149]. Torba dışı (Out-of-Bag - OOB) örnekler olarak bilinen örneklerin üçte biri doğrulama amaçları için saklanırken, kalan torba içi (in-of-bag) örnekler karar ağacını oluşturmak için kullanılır [150] (Şekil 3.6). Rastgele orman yönteminde OOB işleyiş süreci Şekil 3.7'de yer almaktadır.



**Şekil 3.6** Rastgele orman algoritması için veri setlerinin oluşturulması [151]



**Şekil 3.7** Rastgele orman yönetiminde OOB işleyiş süreci [152]

Ağaçlar, regresyon ve sınıflandırma problemleri için kullanılan özyinelemeli bir bölünme yöntemi olan “CART (Classification and Regression Tree)” metodolojisi kullanarak büyütülür ve budanmaz [145]. Her bir düğümdeki bölünme genellikle “CART” algoritmasında kullanılan “Gini” indeksi kriterine dayanır ve standart karar ağaçlarında, düğümler en iyi bölünmeyi veren değişkene göre (Gini'deki en yüksek düşüşe göre) bölünür [148]. Bölünme yeni bir dal oluşturulmasına gerek kalmayana dek devam eder. Bu algoritma, en basit haliyle iki parametrenin tanımlanmasını gerektirir [153]:

- Ormanı oluşturacak ağaç sayısı ( $n$ ),
- Ağaçlardaki her düğüm için dikkate alınan özellik / öngörücü sayısı ( $m$ ).

Sınıflandırma için, rastgele ormanlardaki her ağaç, her girdi örneğindeki en popüler sınıf için bir birim oyu verir ve girdi etiketleri ağaçların oy çokluğu ile belirlenir [102].

Rastgele orman, sınıflandırma sürecinde torba dışı (OOB) verileri kullanarak doğrudan değişken önemine sahip ölçümler sağlayan bir sınıflandırma algoritmasıdır [102, 153]. Rastgele orman, girdi değişkenlerinin önemini belirlemek için ortalama doğrulukta azalma (Mean Decrease in Accuracy - MDA) ve ortalama Gini'de azalma (Mean Decrease in Gini - MDG) olmak üzere iki ölçüm sağlar [148]. Yüksek boyutlu veri kümelerinde özellik seçimi ve en önemli özelliklerin belirlenmesi model yorumlanabilirliğini geliştirebilir ve örnek eğitim sürecini hızlandırabilir [154].

### ***Gini İndeksi***

Gini indeksi, sınıflara ilişkin özniteliklerin safsızlığının ölçüldüğü bir yöntemdir ve bu algorithmada sağ ve sol olmak üzere ikili dallanma söz konusudur [155]. Eşitlik 3.3 ve Eşitlik 3.4'de sırasıyla sağ ve sol dalın Gini değeri hesabı için kullanılan eşitlik, Eşitlik 3.5'de ise bir düğümün Gini değeri hesabı için kullanılan eşitlik görülmektedir.

$$G_{sağ} = 1 - \sum_{i=1}^n \left( \frac{R_i}{|D_{sağ}|} \right)^2 \quad (3.3)$$

$$G_{sol} = 1 - \sum_{i=1}^n \left( \frac{L_i}{|D_{sol}|} \right)^2 \quad (3.4)$$

$$Gini_i = \frac{1}{k} (|D_{sol}|G_{sol} + |D_{sağ}|G_{sağ}) \quad (3.5)$$

Bu eşitliklerdeki  $k$  eğitim verisi sayısını,  $n$  sınıf sayısını,  $D_{sağ}$  sağ taraftaki belirlenen eşik değerden büyük değere sahip örneklerin sayısını,  $D_{sol}$  sol taraftaki belirlenen eşik değerden küçük değere sahip örneklerin sayısını,  $R_i$  sağ taraftaki  $i$  sınıfında bulunan örnek sayısını,  $L_i$  sol taraftaki  $i$  sınıfında bulunan örnek sayısını ifade etmektedir.

Gini indeksi 0 ile 1 arasında bir sayıdır ve küçük bir değer, bir düğümün ağırlıklı olarak tek bir sınıfa ait gözlemleri içerdiğini gösterir [156]. Gini indeksi büyüdükçe düğüm heterojenliği artarken, Gini indeksi azaldıkça düğüm homojenliği artar. Bir alt düğümün Gini indeksi bir üst düğümün Gini indeksinden daha az olduğunda o dal başarılıdır [157]. İdeal dal bölümlenme Gini değeri sıfıra ulaştığında, yani her bir yaprak düğümünde yalnızca bir sınıf kaldığında gerçekleşir [158].

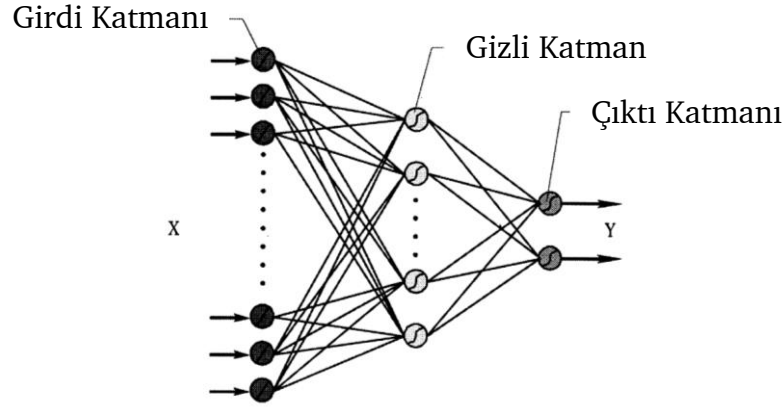
### 3.2.2.3 Sinir Ağı (Neural Network) Modelleri

Sinir ağları (neural networks), deneyimlerden öğrenme yeteneğini taklit eden beynin sinir yapısına dayanan bir tekniktir [159]. Sinir ağları, bir veri kümesindeki temel ilişkileri tanımlayan bir dizi algoritmadan oluşur. Bir sinir ağının temel işlem ögesi bir nörondur [160]. Bir sinir ağının davranışı, nöronlarının aktarım fonksiyonları, öğrenme kuralı ve mimarinin kendisi tarafından belirlenir [161]. Farklı uygulamalar için uygun olan, farklı özelliklere sahip çeşitli sinir ağı mimarileri mevcuttur [159, 162-164] ve bunlardan bazıları aşağıda verilmiştir:

- Çok katmanlı algılayıcılar (Multi-Layer Perceptrons),
- Konvolüsyonel sinir ağları (Convolution Neural Network),
- Tekrarlayan sinir ağları (Recurrent Neural Networks),
- Radyal temel fonksiyonlu sinir ağları (Radial Basis Function Neural Networks),
- Derin inanç ağı (Deep Belief Network),
- Bayes sinir ağları (Bayesian neural networks),
- Olasılıksal sinir ağı (Probabilistic Neural Network).

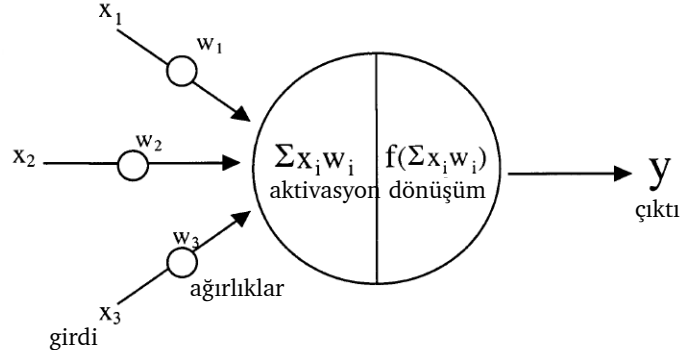
#### ***Çok Katmanlı Algılayıcı (ÇKA)***

Çok katmanlı algılayıcı (Multi-Layer perceptron - MLP), bir ağ biçiminde bağlanmış birden çok nöronun birleşiminden oluşan en popüler yapay sinir ağlarından biridir [164]. En basit ÇKA, en az üç düğüm katmanından oluşur: bir girdi katmanı (input layer), bir gizli katman (hidden layer) ve bir çıktı katmanı (output layer) [163]. ÇKA, doğrusal olmayan öğelerin (nöronların) ardışık katmanlar halinde düzenlendiği ve bilginin giriş katmanından çıktı katmanına, gizli katman boyunca tek yönlü olarak aktığı bir ileri beslemeli sinir ağıdır [165]. Şekil 3.8'de üç katmanlı ileri beslemeli bir sinir ağı modeli verilmiştir.



**Şekil 3.8** Üç katmanlı ileri beslemeli bir sinir ağı modeli [165]

ÇKA'da her bir nöron, bir sonraki katmandaki nörona ağırlık adı verilen bağlantılar aracılığıyla bağlanır [163]. Yapay bir nöron modeli Şekil 3.9'de yer almaktadır. Giriş katmanı, girdileri almaktan sorumludur ve giriş katmanındaki nöron sayısı, giriş veri kümesindeki özelliklerin (modeldeki bağımsız değişkenlerin) sayısına eşittir [166]. Gizli katmanlar, çok katmanlı algılayıcının ana hesaplama çekirdeğidir ve çıktı katmanı, girdinin kararını veya tahminini verir [167].



**Şekil 3.9** Yapay bir nöron modeli [161]

ÇKA, gizli katmanların, nöronların ve yinelemelerin sayısı gibi bir dizi hiperparametrenin ayarlanmasını gerektirir. Çok katmanlı algılayıcı sınıflandırıcısındaki önemli değişkenler aşağıdaki gibi ifade edilebilir:

- Gizli katman boyutu (hidden\_layer\_sizes); sinir ağı sınıflandırıcısında istenilen katman ve nöron sayısını belirlemeyi sağlar,
- Maksimum iterasyon (max\_iter); iterasyon sayısını gösterir,
- Aktivasyon (activation); gizli katmanlar için aktivasyon işlevini gerçekleştirir,

- Çözücü (solver); düğümler arasında ağırlık optimizasyonu için ilgili algoritmayı belirtir,
- Rastgele durum (random\_state); aynı sonuçları yeniden üretmek için bir tohum ayarlamaya izin verir.

### 3.2.3 Pekiştirmeli Öğrenme

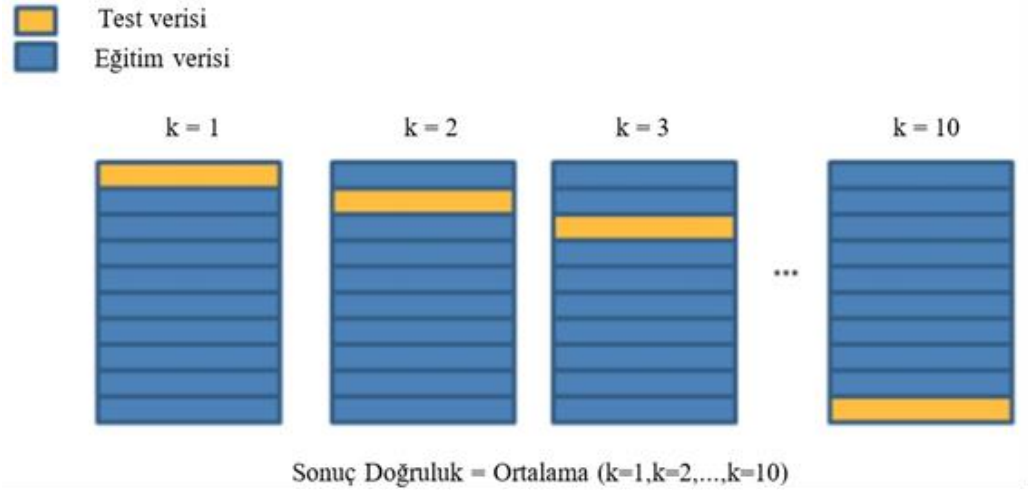
Pekiştirmeli öğrenme, bazı birikimli ödül kavramlarını en üst düzeye çıkarmak için nasıl harekete geçileceği ile ilgilidir [126]. Pekiştirmeli öğrenme, gerçekleştirilecek eylemleri öğrenmek için robotik, yatırım kararları ve envanter yönetiminde kullanılır [126].

Pekiştirmeli öğrenme, bir anlamda, denetimli ve denetimsiz öğrenme arasında yer alır. Denetimsiz öğrenmenin aksine, bir tür denetim vardır, ancak bu, verilerdeki her girdi için istenen bir çıktının belirtilmesi biçiminde gelmez. Bunun yerine, bir pekiştirmeli öğrenme algoritması, yalnızca belirli bir girdi veya gözlem için bir çıktı seçtikten sonra çevreden geri bildirim alır. Geri bildirim, pekiştirmeli öğrenmede eylem olarak bilinen çıktının öğrencinin hedeflerini ne ölçüde yerine getirdiğini gösterir [122].

Bu çalışmada, yalnızca denetimsiz ve denetimli öğrenme algoritmaları kullanıldığından pekiştirmeli öğrenme ayrıntılarına değinilmemiştir.

### 3.3 K-Katlamalı Çapraz Doğrulama (K-Fold Cross Validation)

Bu yaklaşım, gözlem kümesini yaklaşık olarak eşit büyüklükte k gruplarına ayırarak, makine öğrenme model tahminlerinin değerlendirilmesi için kullanılan rastgele k-kat çapraz doğrulama içerir [156]. Döngünün her bir adımında o adıma ait k-1 kat, model gelişimi için eğitim verisi olarak, kalan 1 kat ise ortaya çıkan model geçerliliği için test verisi olarak kullanılır [168-169]. Her bir adımda kullanılacak test verileri farklı bir kattan alınmaktadır. k sayısı kadar tekrarlayan döngü sonucunda, her bir adımda elde edilen sınıflandırıcı başarılarının ortalaması modelin başarısı olarak değerlendirilir. Uygulamada, genellikle k=5 ya da k=10 alınarak k-katlamalı çapraz doğrulama gerçekleştirilir [156, 170]. Şekil 3.10'da 10-katlamalı bir çapraz doğrulamanın görselleştirilmesi verilmiştir.



Şekil 3.10 10-katlamalı çapraz doğrulama



Bu çalışmada makine öğrenme algoritmaları kullanılarak LiDAR nokta bulutu verileri ile ağaç türlerinin otomatik olarak sınıflandırılması olanakları araştırılmıştır. Yaygın olarak grid yapıya dönüştürülmüş LiDAR nokta bulutu verileri kullanılarak gerçekleştirilen klasik görüntü sınıflandırma yaklaşımları (piksel ya da nesne tabanlı) yerine doğrudan ham LiDAR nokta bulutu verileri kullanılarak nokta tabanlı bir sınıflandırma yaklaşımı benimsenmiştir. Gride enterpolasyon sırasında yaşanabilecek veri kayıplarını en aza indirmek amacıyla belirlenen yaklaşımda, ilk olarak hiyerarşik kural tabanlı sınıflandırma yöntemi ile kentsel çalışma alanındaki yüksek bitki örtüsü sınıfı noktaları elde edilmiştir. Daha sonra makine öğrenme algoritmaları kullanılarak bireysel ağaç tacı segmentasyonu ve geniş ve iğne yapraklı ağaç türleri sınıflandırılması gerçekleştirilmiştir.

Bu bölümde, ilk olarak çalışma alanı ve çalışma alanı içerisinde yer alan biri küçük ve biri büyük olmak üzere belirlenen iki test alanı tanımlanmıştır. İkinci kısımda kullanılan LiDAR nokta bulutu verileri ve referans (ground-truth) verilere ilişkin ayrıntılar verilmiştir. Makine öğrenme tabanlı kentsel ağaç türleri sınıflandırması için gerçekleştirilen çalışmanın genel iş akışı üçüncü kısımda yer almaktadır. Hiyerarşik kural tabanlı tabanlı sınıflandırma yaklaşımı ile yüksek bitki örtüsü sınıflandırma işlemi bir sonraki kısımda verilmiştir. Beşinci kısımda denetimsiz makine öğrenme algoritmaları ile gerçekleştirilen bireysel ağaç tacı segmentasyonu yer almaktadır. Makine öğrenme algoritmaları kullanılarak gerçekleştirilecek sınıflandırma uygulamalarında kullanılmak üzere elde edilen iki farklı set halindeki sınıflandırma özellikleri altıncı kısımda gösterilmiştir. Üç farklı sınıflandırıcı kullanılarak gerçekleştirilen geniş ve iğne yapraklı ağaç türleri sınıflandırma uygulamaları ve bu sınıflandırmaların istatistiksel doğruluk analizleri, üç sınıflandırıcının doğruluk performanslarının karşılaştırılması ve sınıflandırmadaki en önemli özelliklerinin belirlenmesi yedinci kısımda

verilmiştir. Ağaç türlerinin sınıflandırılması için geliştirilen hiyerarşik kural tabanlı sınıflandırma ve nokta tabanlı segmentasyon yaklaşımı kullanılarak kent alanlarında tekil cadde ağaçlarının belirlenmesinin test edildiği sekizinci kısım ile bu bölüm sonlandırılmıştır.

#### 4.1 Çalışma Alanı

Bu çalışma kapsamında, İstanbul ili Avrupa yakası Esenler ve Güngören ilçeleri sınırları içerisinde yer alan Yıldız Teknik Üniversitesi Davutpaşa Kampüsü kentsel çalışma alanı olarak seçilmiştir. Yaklaşık 125 hektar büyüklüğündeki Yıldız Teknik Üniversitesi Davutpaşa Kampüsü, farklı şekil ve yükseklikteki binalar, farklı tür ve boyutlardaki zengin bitki örtüsü, asfalt, patika ve yaya yolları, yol işaretleri ve otoparklar gibi çeşitli kentsel unsurları bir arada bulundurmaktadır (Şekil 4.1).



**Şekil 4.1** Yıldız Teknik Üniversitesi Davutpaşa Kampüsü (Google Earth)

Bu çalışmada kullanılmak üzere seçilen kentsel çalışma alanında belirlenen Test alanı 1'in büyüklüğü yaklaşık 4.7 hektar (Şekil 4.2a) ve Test alanı 2'nin büyüklüğü ise yaklaşık yaklaşık 8.9 hektardır (Şekil 4.2b). Şekil 4.3'de test alanında yer alan örnek kent ağaçlarının konumları ve bu ağaçlara ilişkin fotoğraflar verilmiştir. Test

alanlarında çok sayıda geniş yapraklı ve iğne yapraklı ağaç türü mevcuttur (Tablo 4.1). Tablo 4.1’de görüldüğü üzere test alanlarında geniş yapraklı ağaçların tür çeşitliliği iğne yapraklılardan oldukça fazladır.



(a)



(b)

**Şekil 4.2** Test alanı 1 (a) ve Test alanı 2 (b) (Hava fotoğrafı – 2013)



(a)



(b)

**Şekil 4.3** Test alanlarında yer alan kent ağaçlarının konumları (Google Earth – 2013) (a) ve bu ağaçlara ilişkin fotoğraflar (b)

**Tablo 4.1** Test alanlarında bulunan geniş ve iğne yapraklı ağaç türleri

Geniş yapraklı ağaçlar				İğne yapraklı ağaçlar
Katalpa	Erik	Nar	Mahlep	Mavi ladin
Erguvan	Mürdüm eriği	Kiraz	Ayva	Avrupa ladini
Ak dut	Elma	Oya ağacı	Ova akçaağacı	Kızılçam
Japon kurtbağı	Yalancı akasya	Sofora	Süs kirazı	Sahil çamı
Türk fındığı	Doğu çınarı	Defne	Pirnal meşe	Fıstık çamı
Adi dişbudak	Karayemiş	Kobuşu manolyası	Büyük çiçekli manolya	Anadolu karaçamı
Amerikan sığla ağacı	Amerikan glaçdiyası	Gümüşi ıhlamur	Kafkas ıhlamuru	
Sivri meyveli dişbudak	Büyük yapraklı ıhlamur	Dişbudak yapraklı akçaağaç		

## 4.2 LiDAR Verisi ve Referans Veri

Bu bölümde, bu çalışmada kullanılan LiDAR verisinin özellikleri ayrıntılı bir şekilde verilmiştir. Referans olarak kullanılan, test alanlarında yer alan geniş ve iğne yapraklı ağaçları içeren referans veri toplama aşamaları yine bu bölümde irdelenmiştir.

### 4.2.1 LiDAR Verisi

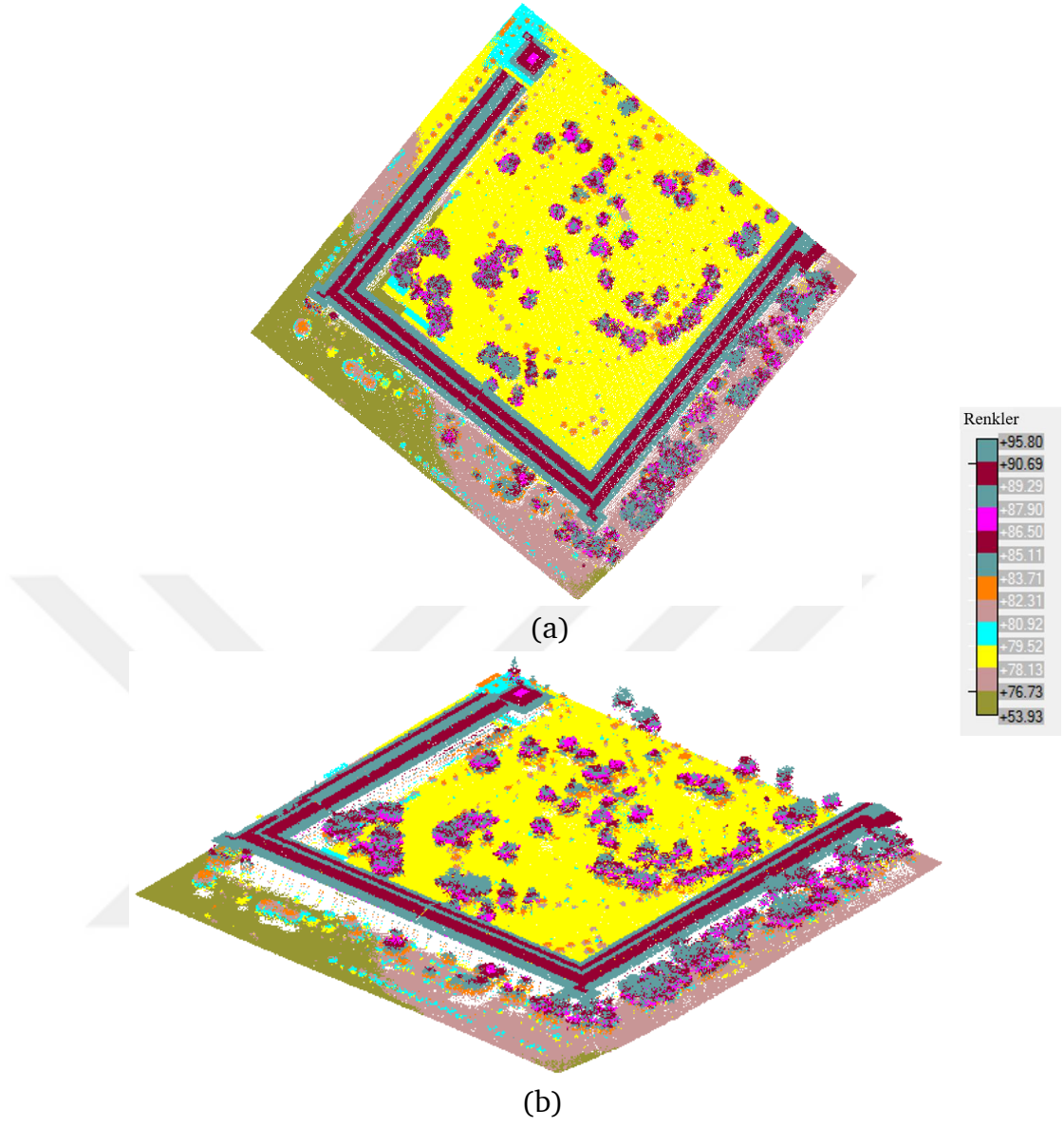
Bu çalışmada kullanılan Yıldız Teknik Üniversitesi Davutpaşa Kampüsü'ne ait LiDAR nokta bulutu verileri "Hava LiDAR Teknolojisiyle Sayısal Yüzey Modeli ve 3 Boyutlu Kent Modeli Projesi" kapsamında toplanmıştır. Eylül, 2013'te BİMTAŞ tarafından uçuşu gerçekleştirilen veriler İstanbul Büyükşehir Belediyesi'nden temin edilmiştir. "Eurocopter AS350" helikoptere monte edilmiş tam dalga boyu kaydeden (full-waveform) "Riegl LSM-Q680i" lazer tarayıcı ile elde edilen LAS (Log ASCII Standard) formatındaki verilerin yoğunluğu 16 nokta/m<sup>2</sup> dir. Bütünleşik sistemde LiDAR verileri ile birlikte eş zamanlı olarak "IGI DigiCam" kamera ve "IGI Aerocontrol" georeferans sistemleri ile de veri toplanmıştır. Helikopterin hızı 148 km/h ve uçuş yüksekliği yaklaşık 600 metredir. Tablo 4.2'de

“Riegl LSM-Q680i” hava lazer tarayıcısının Reigl firması tarafından belirlenen teknik özellikleri yer almaktadır. İlgili proje kapsamında verilerin tüm kalibrasyon ve kolon dengelemesi işlemleri BİMTAŞ tarafından gerçekleştirilmiştir.

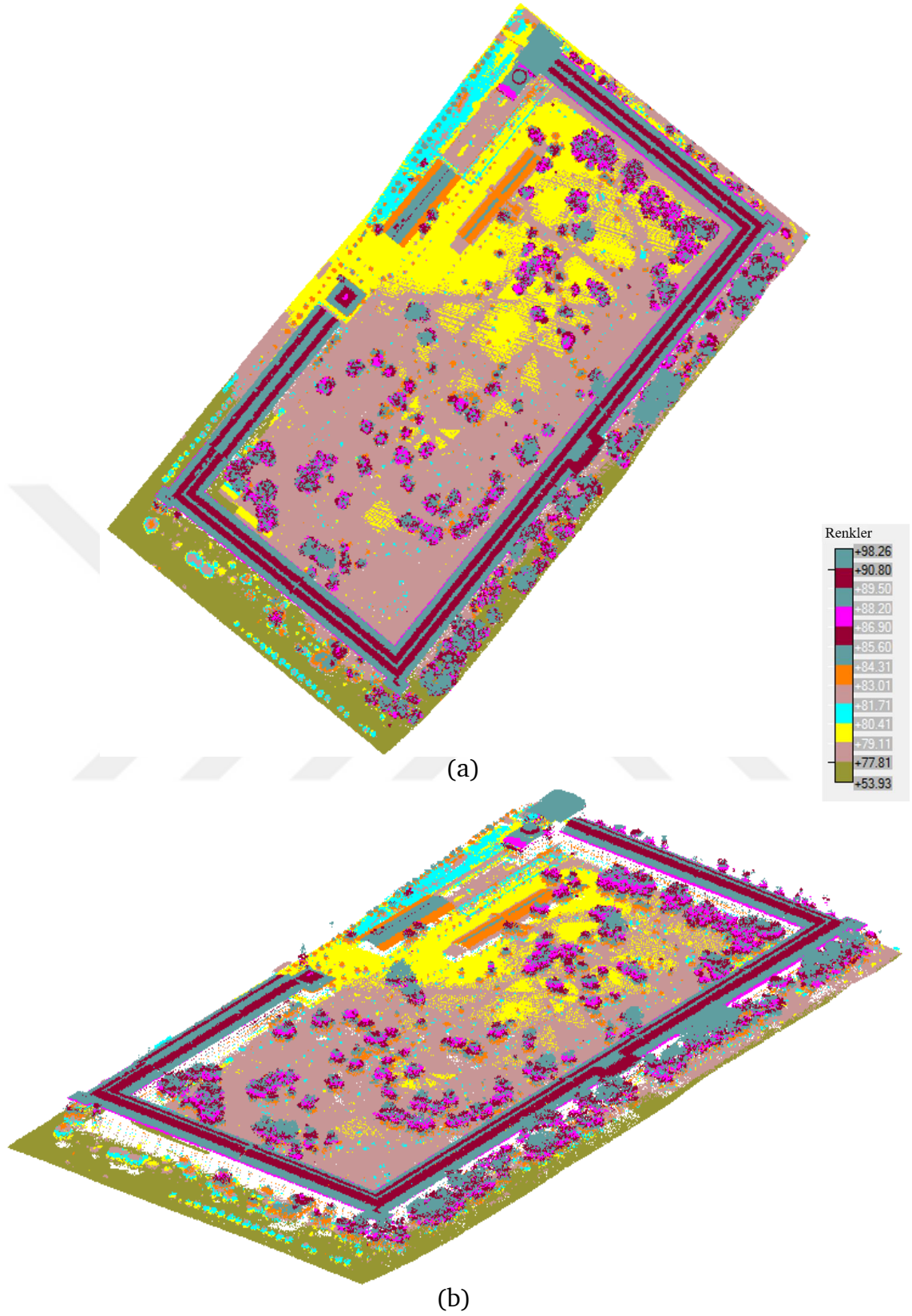
**Tablo 4.2** “RIEGL LMS-Q680i” lazer tarayıcı özellikleri [72]

Eğik (Range) Ölçüm Performansı	Minimum mesafe	30 m
	Doğruluk (accuracy)	20 mm
	Hassasiyet (precision)	20 mm
	Lazer sinyali tekrarlama oranı	400 000 Hz (e varan)
	Etkin ölçüm oranı	266 Hz (e varan) (60° tarama açısında)
	Paralel tarama çizgileri	8956 x 6708 piksel
	Lazer dalga boyu	Yakın kızılötesi
	Lazer ışını iraksaması	≤ 0.5 mrad
	Sinyal başına hedef sayısı	Sayısallaştırılmış dalga işleme: Sınırsız Veri çıkışı gözleme: İlk dönen
Tarayıcı Performansı	Tarama mekanizması	Döner ayna
	Tarama deseni	Paralel tarama çizgileri
	Tarama açısı aralığı	+30°/-30° (toplam 60°)
	Tarama hızı	10 - 200 çizgi/sn
	Açı ölçüm çözünürlüğü	0.001°
Yoğunluk (intensity) ölçümü		Her dönüş sinyali için 16 bit yoğunluk verisi
Genel Özellikleri	Ağırlık	17.5 kg
	Maksimum uçuş yüksekliği	5000 m

Test alanı 1’e ait LiDAR verisi 553 086, Test alanı 2’ye ait LiDAR verisi ise 1 216 449 noktadan oluşmaktadır. Şekil 4.4’te Test alanı 1’e ait nokta bulutu verileri ve Şekil 4.5’de Test alanı 2’ye ait nokta bulutu verileri yer almaktadır.



**Şekil 4.4** Test alanı 1'in yüksekliğe göre renklendirilmiş nokta bulutu verisi (a) ve nokta bulutunun perspektif görünümü (b)

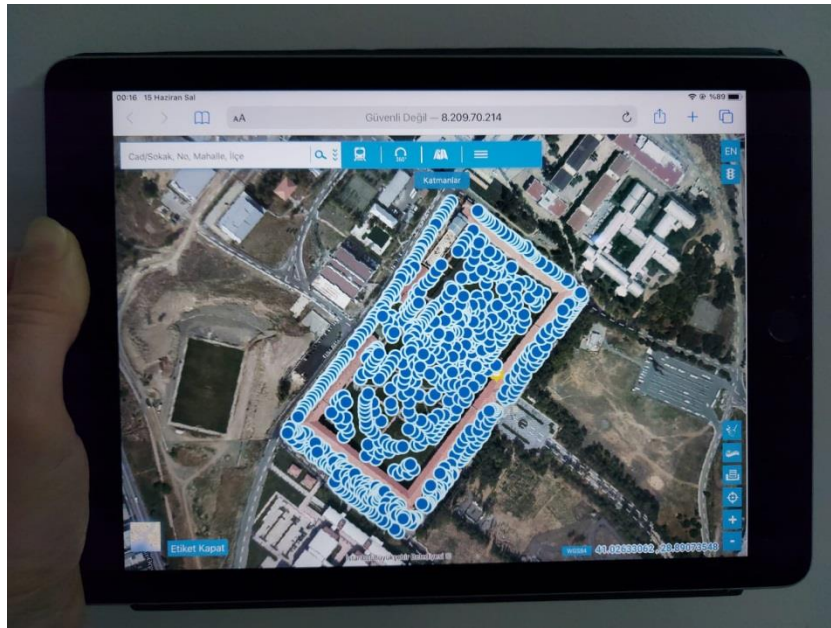


Şekil 4.5 Test alanı 2'nin yüksekliğe göre renklendirilmiş nokta bulutu verisi (a) ve nokta bulutunun perspektif görünümü (b)

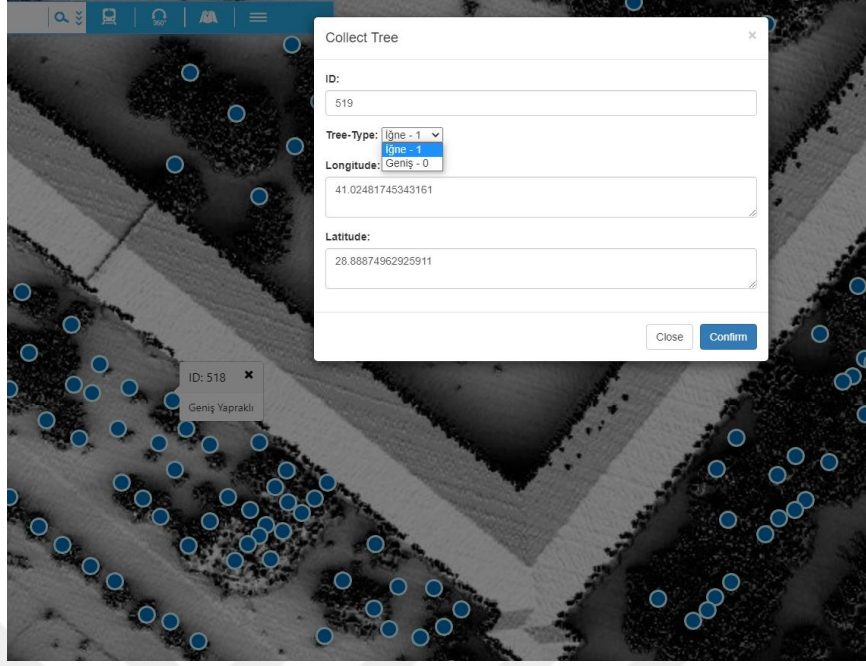
#### 4.2.2 Referans Veri

Bu çalışmada, Yıldız Teknik Üniversitesi Davutpaşa Kampüsü'nde belirlenen test alanlarında gerçekleştirilmesi planlanan ağaç türlerinin makine öğrenme algoritmaları kullanılarak sınıflandırması için referans veri olarak kullanılmak üzere test alanlarında bulunan geniş ve iğne yapraklı ağaç türlerinin bilinmesine ihtiyaç vardır. Detaylı arazi çalışması ile ağaç türleri verisi elde edilmiştir.

Test alanlarında bulunan geniş ve iğne yapraklı ağaç türlerinin arazide elde edilmesi aşamasında saha çalışmasını kolaylaştırmak amacıyla İBB Şehir Haritası'nın tüm geliştiricilere sunduğu API kütüphanesi ile bir veri toplama uygulaması geliştirilmiştir. Ön yüzde "Şehir Haritası API", arka yüzde "Django" ve "NGINX" web uygulama hazırlama platformları kullanılarak sahada ağaç türü verisinin toplanması için bir uygulama hazırlanmış ve internet üzerinden erişilebilecek şekilde bir sunucu aracılığıyla yayınlanmıştır. "Tree Collector" olarak isimlendirilen İBB Şehir Haritası tabanlı uygulama kullanılarak test alanlarında bulunan ağaç tür bilgileri (geniş ya da iğne olarak) arazide yapılan incelemeler sonucunda bir tablet aracılığıyla ekran üzerinden işaretlenen her bir ağaç için belirlenmiştir (Şekil 4.6 ve Şekil 4.7). Böylece saha çalışması sırasında kritik öneme sahip referans veri sayısal olarak doğrudan elde edilmiş ve bu aşamada oluşabilecek kaba hataların önüne geçilmiştir.

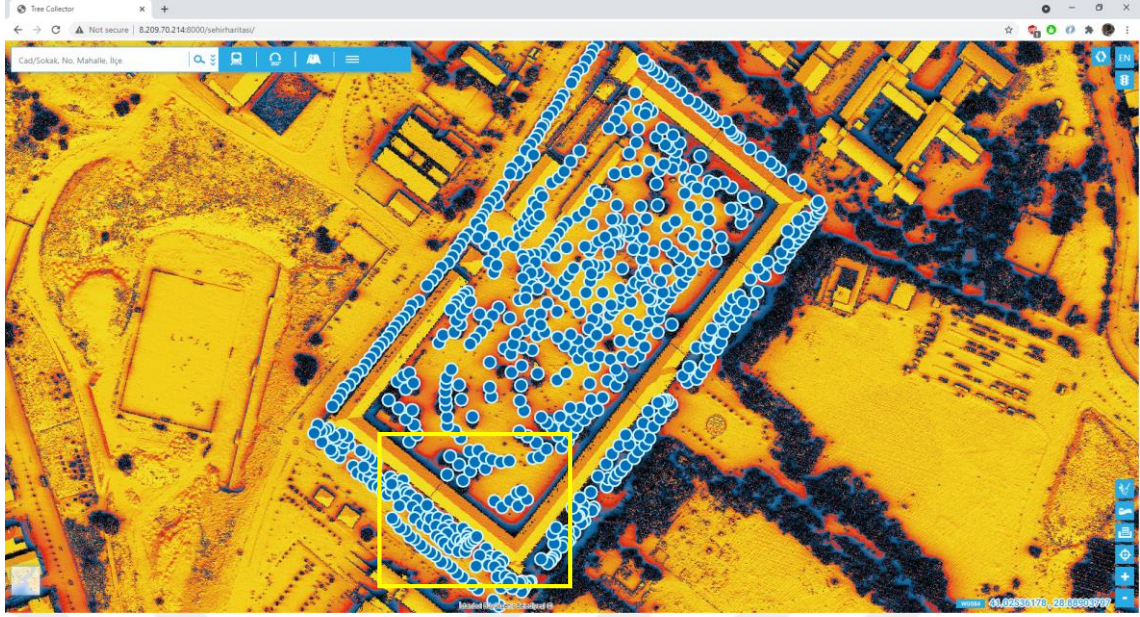


**Şekil 4.6** Geliştirilen Şehir Haritası tabanlı uygulamanın veri toplama ekranı

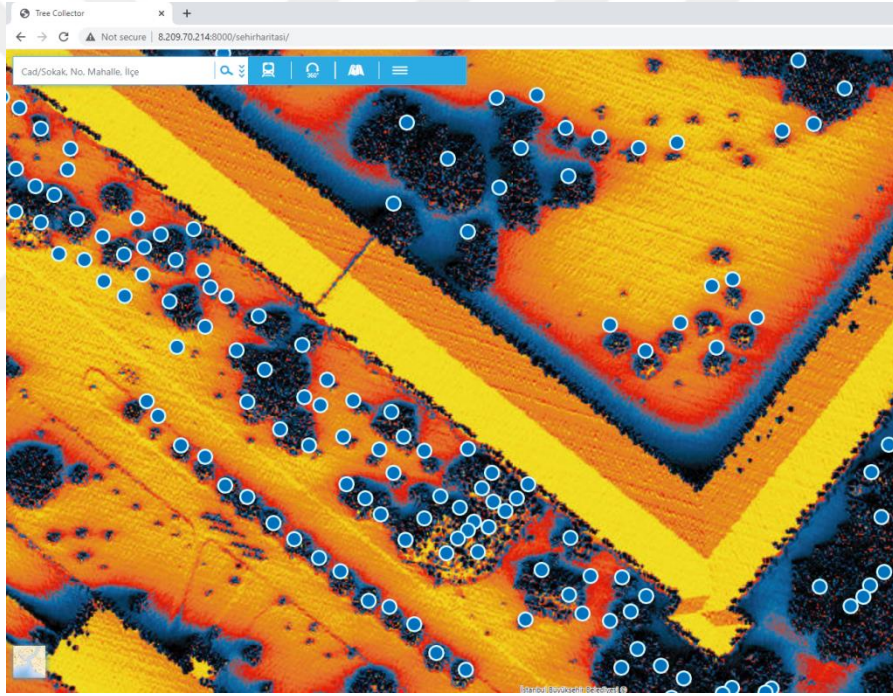


**Şekil 4.7** Geliştirilen Şehir Haritası tabanlı uygulamada ekran üzerinden işaretlenen her bir ağaç için tür bilgisinin seçildiği veri toplama penceresi

Referans veri toplama aşamasında ekran üzerinden ağaçların işaretleme işleminin daha rahat bir şekilde yapılabilmesi için altlık olarak Şehir Haritası'nın 2013 ortofoto, 2014 uydu görüntüsü, siyah beyaz solar harita, renkli solar harita ya da difüzyon haritası verilerinden birisinin seçilerek kullanımı hazırlanan uygulama ile sağlanmıştır. Toplanan ağaç verisinin renkli solar harita üzerinde görünümü Şekil 4.8'de verilmiştir.



(a)

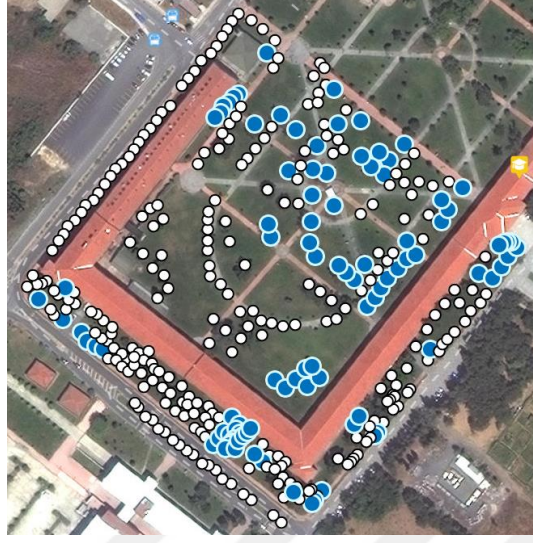


(b)

**Şekil 4.8** Renkli solar harita üzerinde toplanan ağaç verisinin görünümü (698 ağaç) (a) ve renkli solar harita üzerinde toplanan ağaç verisinin yakın görünümü (sarı dikdörtgen) (b)

Test alanı 1’de 269 adet geniş yapraklı ve 97 adet iğne yapraklı ağaç olmak üzere toplam 366 adet ağaç bulunmaktadır (Şekil 4.9a). Test alanı 2’de ise 548 adet geniş yapraklı ve 150 adet iğne yapraklı olmak üzere toplamda 698 adet ağaç mevcuttur (Şekil 4.9b). Çalışma bölgesinde yer alan geniş ve iğne yapraklı ağaç

örnekleri Şekil 4.10'da, geniş ve iğne yapraklı ağaçların LiDAR verisinde görünüm örnekleri Şekil 4.11'de verilmiştir.



(a)

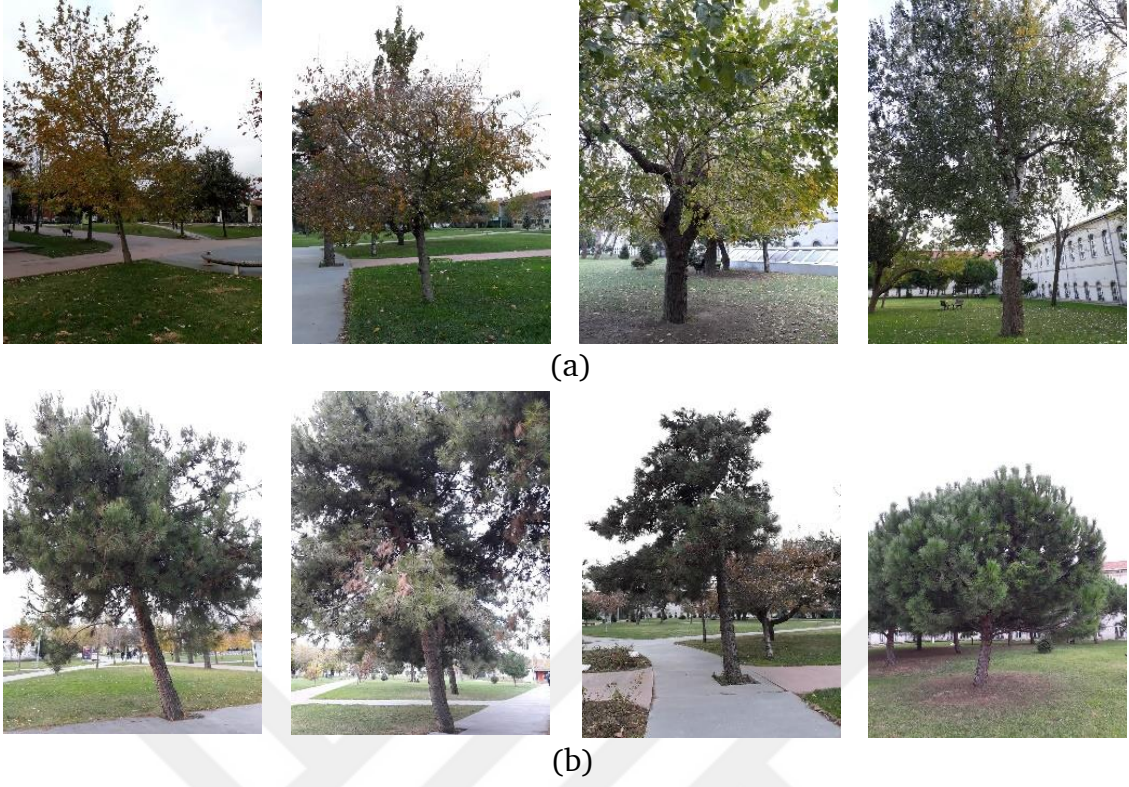


○ Geniş yapraklı

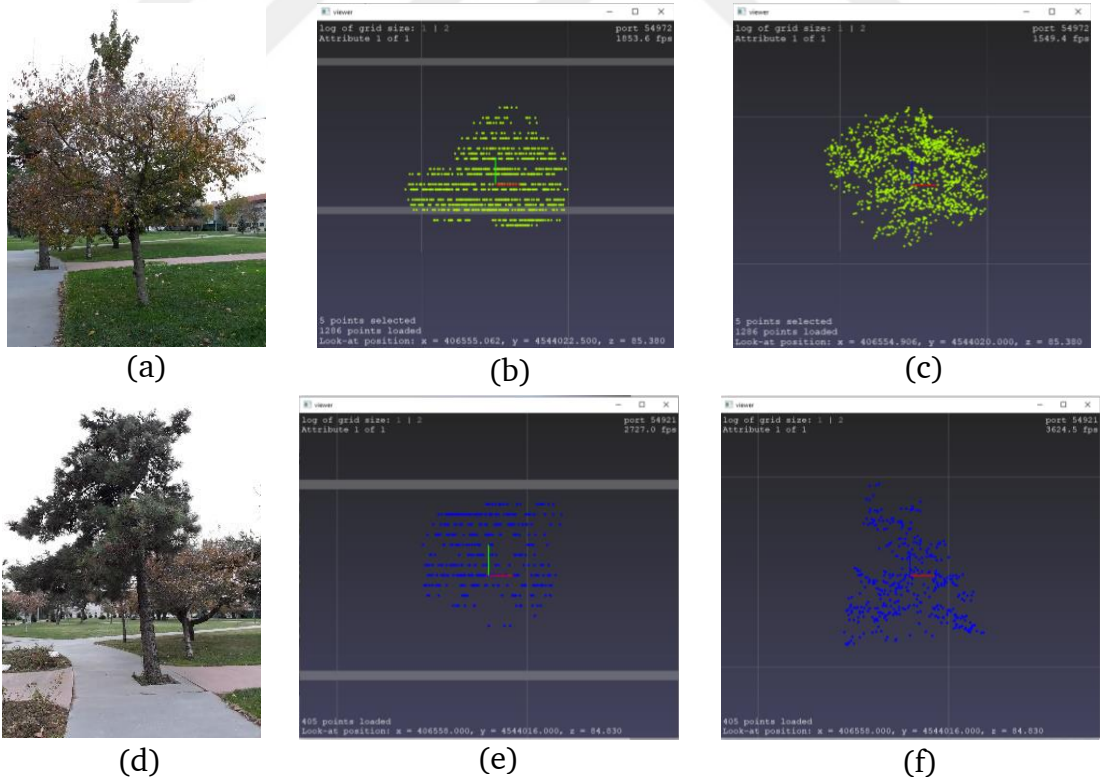
● İğne yapraklı

(b)

Şekil 4.9 Test alanı 1'in referans verisi (a) ve Test alanı 2'nin referans verisi (b)



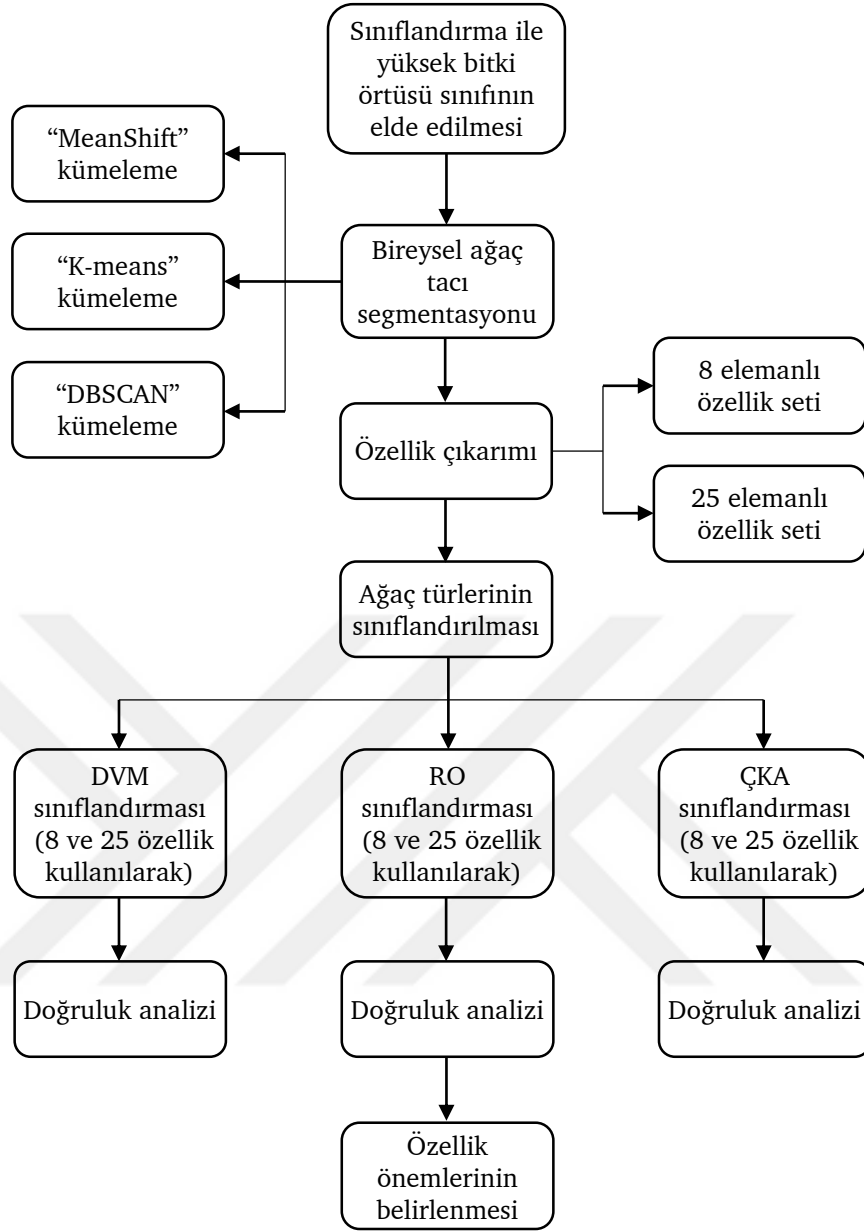
**Şekil 4.10** Çalışma bölgesindeki geniş yapraklı (a) ve iğne yapraklı (b) ağaç örnekleri



**Şekil 4.11** Geniş yapraklı ağaç (a), geniş yapraklı ağacın LiDAR verisinde tepeden görünümü (b) ve yandan görünümü (c); iğne yapraklı ağaç (d), iğne yapraklı ağacın LiDAR verisinde tepeden görünümü (e) ve yandan görünümü (f)

### 4.3 Genel İşlem Adımları

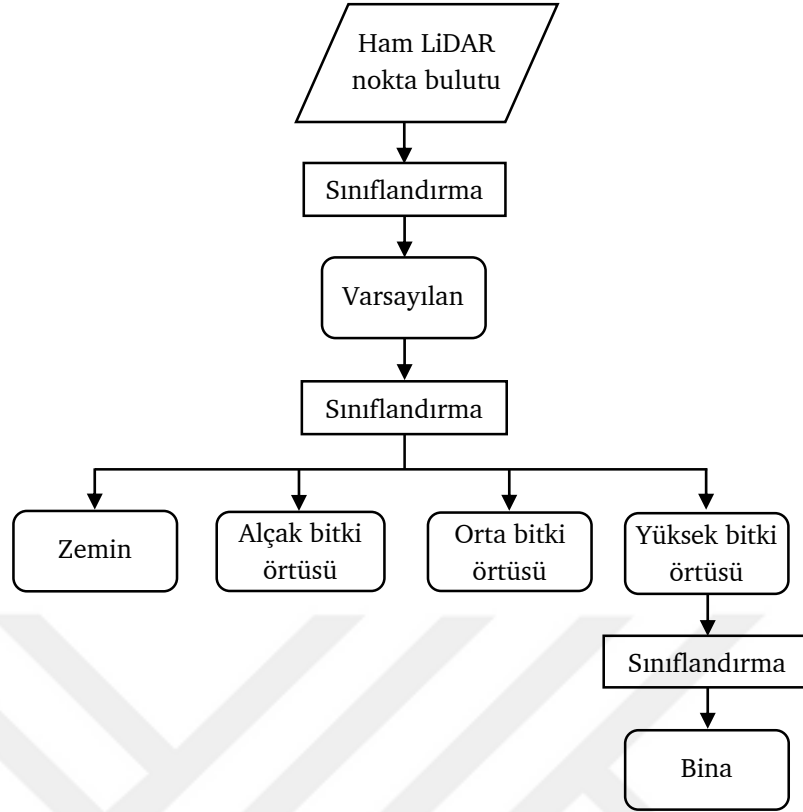
Bu çalışmada, özellikle büyük ölçekli uygulamalarda kentsel ağaç türlerini hızlı ve maliyet açısından yüksek verimlilikle belirleyebilmek için 3 boyutlu ham LiDAR verileri kullanılarak makine öğrenme algoritmalarına dayalı ağaç türleri sınıflandırma yöntemi geliştirilmiştir. Makine öğrenme algoritmalarına dayalı ağaç türleri sınıflandırma yöntemi genel olarak beş temel adımı içermektedir. İlk adım, bir hiyerarşik kural tabanlı sınıflandırma yöntemi ile ham LiDAR nokta bulutu verisinden zemin, bina, alçak bitki örtüsü, orta bitki örtüsü, yüksek bitki örtüsü, alçak nokta, hava nokta ve varsayılan sınıflarının elde edilmesidir. Daha sonraki adım, bireysel ağaç taçlarının, yüksek bitki örtüsü noktalarından makine öğrenme tabanlı kümeleme algoritmaları kullanılarak segmente edilmesidir. Her bir ağaç için sınıflandırma özelliklerinin elde edilmesi ve bu özellikler kullanılarak kentsel çalışma alanındaki geniş yapraklı ve iğne yapraklı ağaç türlerinin DVM, RO ve ÇKA makine öğrenme algoritmaları ile sınıflandırılması takip eden işlem adımıdır. Son işlem adımı ise, sınıflandırılmış ağaç türleri için doğruluk analizinin gerçekleştirilmesidir. “TerraScan” yazılımı (sadece sınıflandırma ile yüksek bitki örtüsü sınıfının elde edilmesi aşamasında) ve “Jupyter Notebook (açık kaynak kodlu bir web uygulaması)” ortamında “Python programlama dili (Python 3.6.4)” kullanılarak gerçekleştirilen bu çalışmanın genel işlem adımlarını içeren akış şeması Şekil 4.12’de yer almaktadır.



**Şekil 4.12** Önerilen makine öğrenme tabanlı ağaç türleri sınıflandırması yaklaşımının genel işlem adımları

#### 4.4 Sınıflandırma ile Yüksek Bitki Örtüsü Sınıfının Elde Edilmesi

Bu çalışmada kentsel ağaç türlerinin sınıflandırılması için önerilen yaklaşım, ham LiDAR nokta bulutu verilerinden yüksek bitki örtüsü noktalarının sınıflandırılması ile başlamaktadır. Ham LiDAR nokta bulutundan yüksek bitki örtüsü sınıfının elde edilmesi amacıyla hiyerarşik kural tabanlı bir sınıflandırma yaklaşımı kullanılmıştır. Her tekil lazer noktası için bir nesne sınıfı belirlemeyi amaçlayan nokta tabanlı bu sınıflandırma yaklaşımının genel işlem adımları Şekil 4.13'de verilmiştir.



**Şekil 4.13** Yüksek bitki örtüsü sınıfının elde edilmesi amacıyla önerilen yaklaşımın genel işlem adımları [72]

Yüksek bitki örtüsü sınıfının elde edilmesi amacıyla oluşturulan hiyerarşik nokta tabanlı kural seti, ilk olarak LiDAR verisinde yer alan tüm noktaların varsayılan sınıfına atanması için basit olarak tüm noktaları bir sınıftan değerine geçirmeyi sağlayan “Sınıfa göre” kuralını içermektedir. Daha sonra belirlenen yarıçap değerine göre her bireysel noktanın komşuluğundaki diğer noktalarla yüksekliğini karşılaştıran “Alçak noktalar” kuralı, zemin sınıfına ait olamayacak kadar alçakta bulunan gürültü noktalarının alçak nokta sınıfına atanması amacıyla kural setine eklenmiştir. Zemin noktalarının sınıflandırılması amacıyla iteratif olarak bir üçgen yüzey modeli oluşturan [100] “Zemin” kuralı kullanılmıştır. Oluşan zemin yüzeyinin altında kalan hatalı noktaların ayıklanması amacıyla bir sınıftaki komşu noktalardan alçakta olan noktaları sınıflandıran “Yüzeyin altı” kuralı bu kural setine dâhil edilmiştir. Zemin sınıfı noktalarının elde edilmesinin ardından ham nokta bulutunda kalan yüksek noktalar için belirli bir yarıçap değeri dikkate alınarak havadaki gürültü noktalarını temizleyen “Hava noktaları” kuralı eklenmiştir. Havadaki noktaların temizlenmesinin ardından ise, zemin yüzeyine göre belirli yükseklik aralığındaki noktaların zemin sınıfına eklenmesi ve alçak

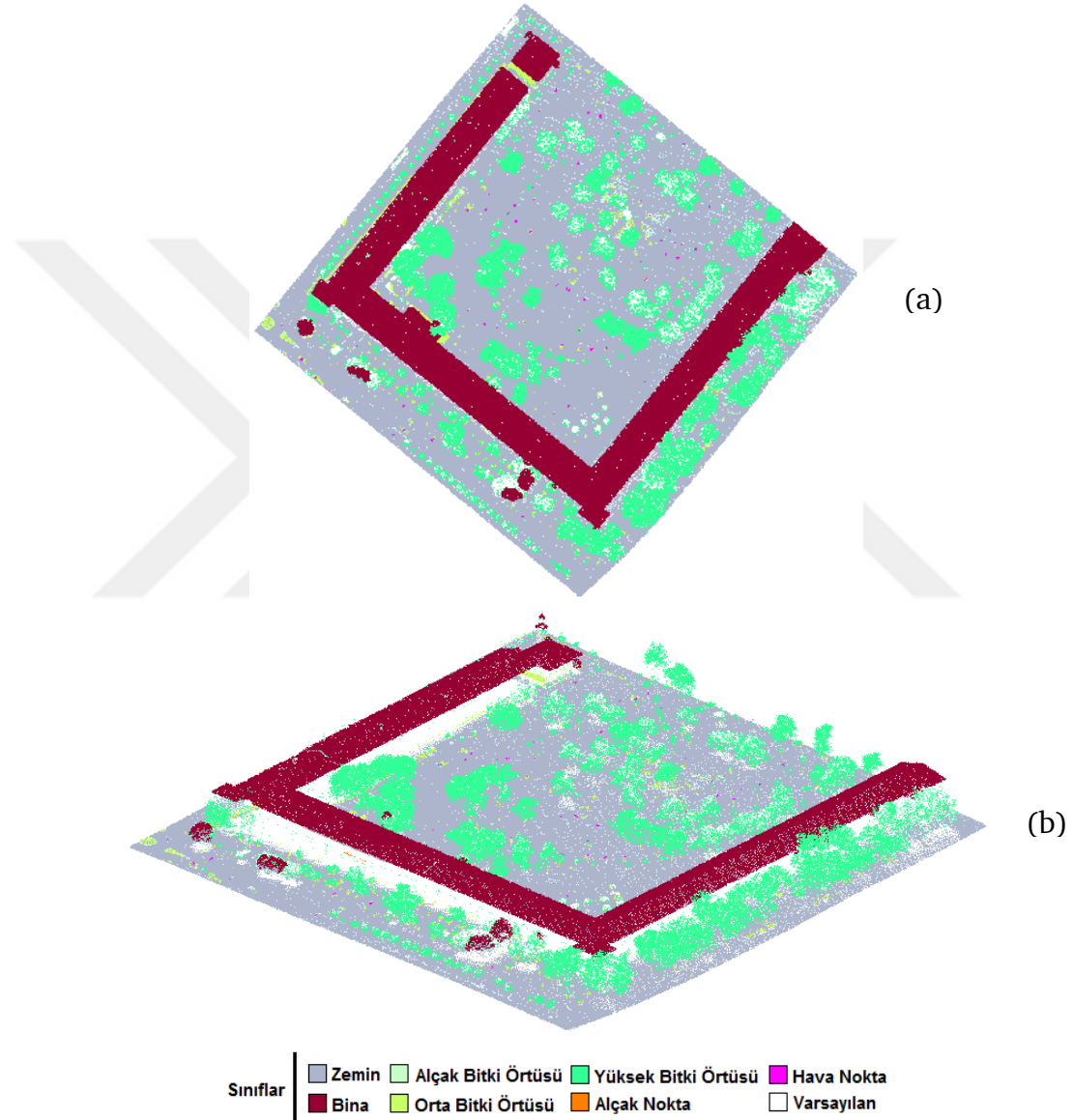
bitki örtüsü, orta bitki örtüsü ve yüksek bitki örtüsü sınıflarının elde edilmesi için “Yerden yüksekliğe göre 1”, “Yerden yüksekliğe göre 2”, “Yerden yüksekliğe göre 3”, “Yerden yüksekliğe göre 4” kuralları kullanılmıştır. Yüksek bitki örtüsü sınıfı noktaları arasından bina sınıfı noktalarının ayrılması amacıyla özdeğerler ya da yerel düzlemin özellikleri kullanılarak bina çatılarındaki düzlemsel yüzeyleri oluşturan “Bina” sınıflandırma kuralı hiyerarşik kural setine ilave edilmiştir. Son olarak yüksek bitki örtüsü noktaları arasında kalan hatalı bina noktalarını elemine etmek için noktaları dönüş bilgilerine dayanarak sınıflandıran “Çoklu dönüşe göre” kuralı ile kural seti tamamlanmıştır. Oluşturulan hiyerarşik nokta tabanlı kural seti Tablo 4.3’de verilmiştir.

**Tablo 4.3** Yüksek bitki örtüsünün sınıflandırılması için oluşturulan kural seti ve oluşan sınıflar

<b>Nokta Tabanlı Sınıflandırma</b>	
<b>Kural</b>	<b>Oluşan Sınıf</b>
Sınıfa göre	Varsayılan
Alçak noktalar	Alçak nokta
Zemin	Zemin
Yüzeyin altı	Alçak nokta
Hava noktaları	Hava nokta
Yerden yüksekliğe göre 1	Zemin
Yerden yüksekliğe göre 2	Alçak bitki örtüsü
Yerden yüksekliğe göre 3	Orta bitki örtüsü
Yerden yüksekliğe göre 4	Yüksek bitki örtüsü
Bina	Bina
Çoklu dönüşe göre	Varsayılan

Nokta tabanlı sınıflandırma yaklaşımında her bir bireysel LiDAR noktası oluşturulan hiyerarşik kural setine göre ait olduğu sınıfa otomatik olarak atanmaktadır. Bu nedenle, doğru bir sınıflandırma işlemi için kullanılan her bir kural içindeki parametrelerin hassas bir şekilde belirlenmesi gerekmektedir. Özellikle zemin, bitki örtüsü ve bina sınıfları için detaylı parametre analizleri gerçekleştirilmiştir.

Detaylı parametre analizlerinin ardından belirlenen parametrelerin kullanıldığı hiyerarşik kural seti ile Test alanı 1 ve Test alanı 2'ye ait LiDAR verileri ASPRS standartlarına uygun olarak nokta tabanlı sınıflandırılmıştır. Sınıflandırma sonucunda zemin, bina, alçak bitki örtüsü, orta bitki örtüsü, yüksek bitki örtüsü, alçak nokta, hava nokta ve varsayılan olmak üzere sekiz sınıf elde edilmiştir (Şekil 4.14 ve Şekil 4.15).

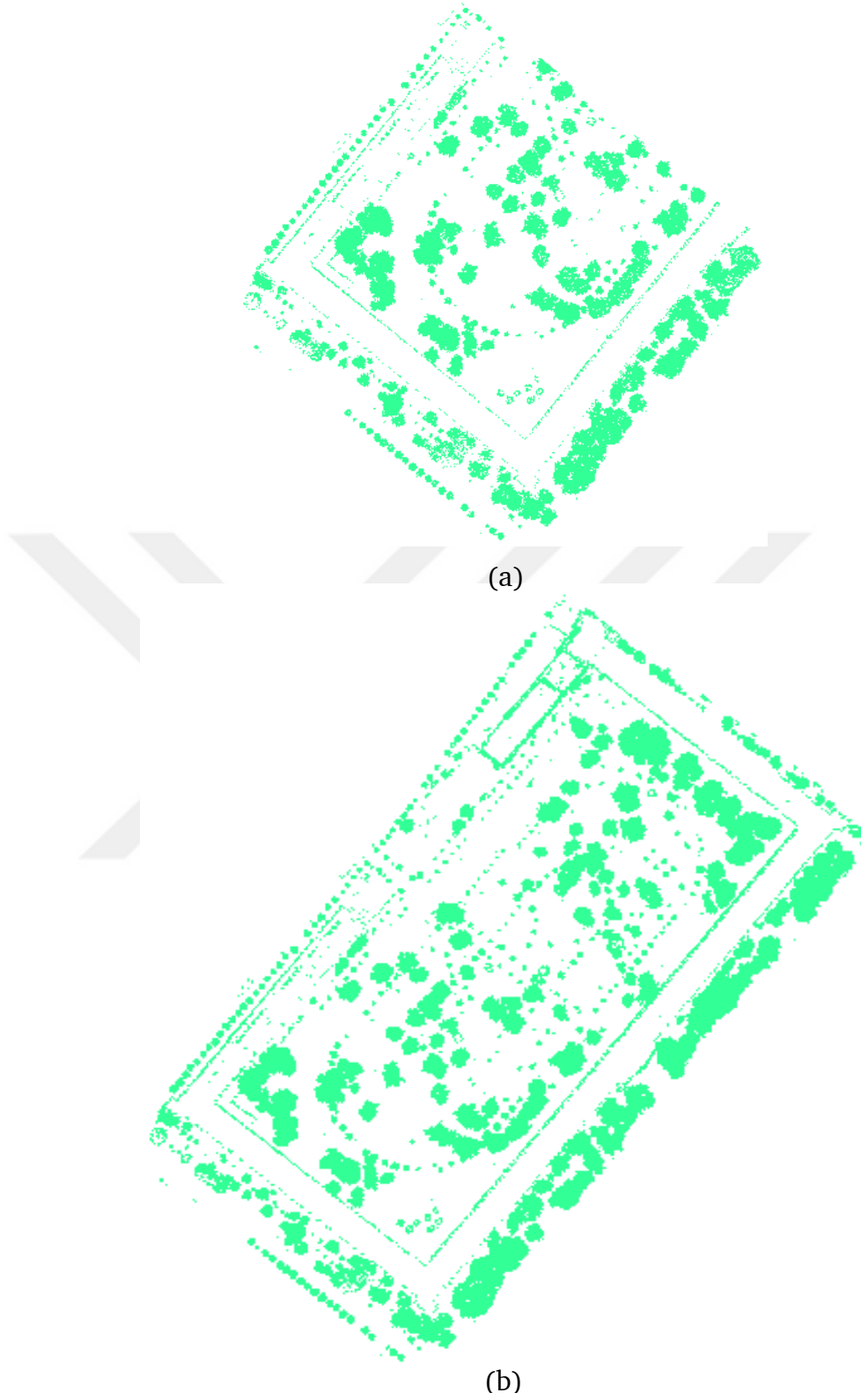


**Şekil 4.14** Test alanı 1'in hiyerarşik kural seti ile sınıflandırma sonucu (a) ve sınıflandırma sonucunun perspektif görünümü (b)



Test alanlarında nokta tabanlı sınıflandırma sonucu elde edilen yüksek bitki örtüsü sınıfı ağaç tacı segmentasyonunda kullanılmak üzere diğer veriden ayrılmıştır (Şekil 4.16). Nokta tabanlı sınıflandırma sonuçları irdelendiğinde, yüksek bitki örtüsü sınıfına ait noktaların büyük oranda sınıflandırıldığı ve bu bağlı olarak çalışma alanındaki neredeyse tüm ağaç noktalarının elde edildiği görülmektedir.

Bu çalışmadaki sınıflandırma işlemi, “TerraSolid” yazılım ailesinin ana modülü olan “TerraScan” ile yukarıda detayları açıklanan, Tablo 4.3’te verilen kural setleri kullanılarak gerçekleştirilmiştir. Nokta bulutu yönetim, işleme ve görselleştirme, noktalardan 3B vektör veri oluşturma, otomatik 3B bina modeli üretme ve elle modifiye gibi özellikleri ile ön plana çıkan “TerraScan” bir CAD ortamı olan “MicroStation” ortamında bir uygulama olarak çalışmaktadır. Bu çalışmada, hiyerarşik kural seti ile gerçekleştirilen nokta tabanlı sınıflandırma uygulaması ve bu uygulamanın doğruluğuna ilişkin detaylı bilgiler Erişir [72]’de yer almaktadır.

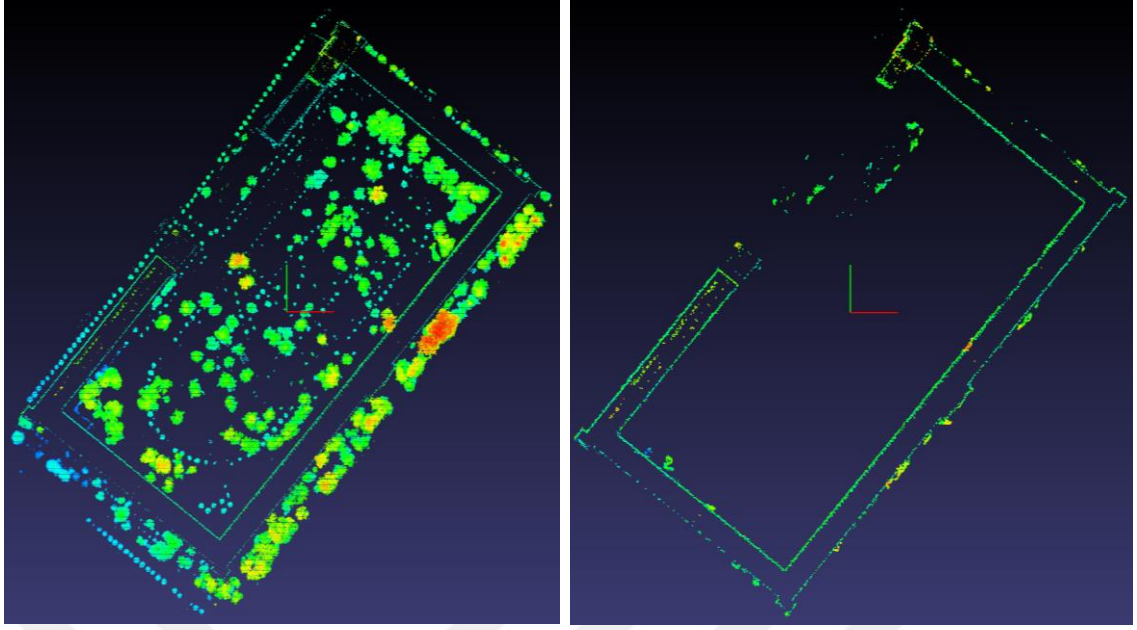


**Şekil 4.16** Test alanı 1’de elde edilen yüksek bitki örtüsü noktaları (a) ve Test alanı 2’de elde edilen yüksek bitki örtüsü noktaları (b)

#### 4.5 Bireysel Ağaç Tacı Segmentasyonu

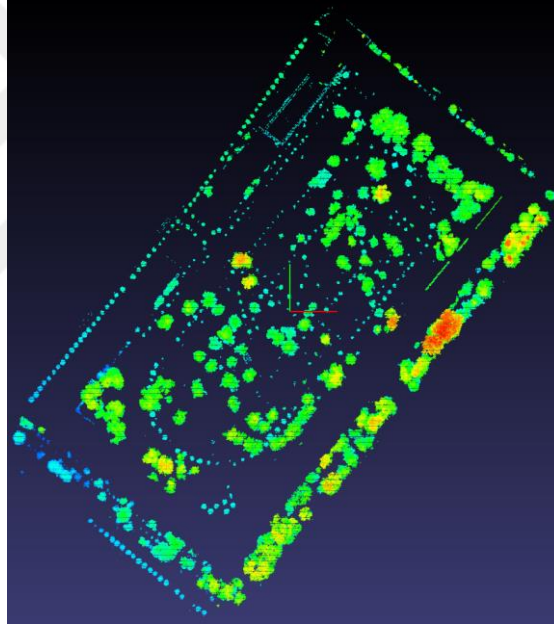
Bu çalışmada, hiyerarşik kural tabanlı sınıflandırma ile elde edilen yüksek bitki örtüsü noktalarının denetimsiz makine öğrenme algoritmaları kullanılarak segmente edilmesi sonucu test alanlarında yer alan bireysel ağaç taçları elde edilmiştir. Öncelikle bireysel ağaç taçlarının daha hassas bir şekilde elde edilmesi için daha önce sınıflandırılan yüksek bitki örtüsü noktaları arasında bulunan (özellikle bina kenarlarına ait noktalar) hatalı bina sınıfına ait noktaların temizlenmesi hedeflenmiştir. Bu amaçla “Python programla dili” kullanılarak daha önce sınıflandırılmış her bir bina noktasının 1 metrelik tampon bölgesi içerisinde bulunan yüksek bitki örtüsü noktaları tespit edilmiştir. Buradaki amaç bina çatı kenarı ya da bina duvarına ait noktaların binaya yakın yüksek bitki örtüsü sınıfına ait noktalar ile karışmasının engellenmesidir. Her iki test alanını da içeren Test alanı 2’de yer alan tüm bina noktalarının tampon bölgesine düşen hatalı yüksek bitki örtüsü noktaları tespit edilip mevcut yüksek bitki örtüsü noktalarından çıkarılmıştır. Bu işlemin sonucunda daha doğru, mümkün olduğunca az bina noktası içeren yüksek bitki örtüsü noktaları elde edilmiştir (Şekil 4.17).

Yüksek bitki örtüsü noktaları arasında yer alan hatalı bina noktalarının ayıklanmasının ardından bireysel ağaç taçlarını elde etmek için her ham LiDAR noktasının x ve y Kartezyen koordinat çiftine dayalı olarak 2 boyutlu nokta tabanlı segmentasyonu gerçekleştirilmiştir. “MeanShift”, “K-means” ve “DBSCAN” kümeleme algoritmaları kullanılarak her iki test alanını da içeren Test alanı 2’de bulunan yüksek bitki örtüsü noktaları bireysel ağaç taçlarını elde etmek için segmente edilmiştir.



(a)

(b)



(c)

**Şekil 4.17** Hiyerarşik kural tabanlı sınıflandırma ile edilen yüksek bitki örtüsü noktaları (a), 1 m'lik tampon bölge ile elde edilen hatalı bina noktaları (b) ve hatalı bina noktalarının çıkarıldığı son yüksek bitki örtüsü noktaları (c)

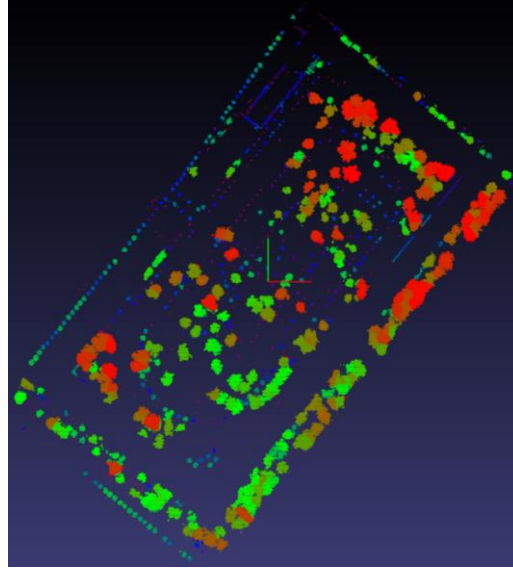
#### 4.5.1 “MeanShift” Kümeleme ile Bireysel Ağaç Taçlarının Elde Edilmesi

Bireysel ağaç taçlarının elde edilmesi amacıyla “MeanShift” kümeleme işlemi için gerçekleştirilen parametre analizleri sonucunda bant genişliği değeri 3 ve komşular araması için çalıştırılacak paralel işlerin sayısı ( $n\_jobs$ ) ise -1 (tüm işlemcileri kullanarak) olarak belirlenmiştir. Yüksek bitki örtüsü noktalarının

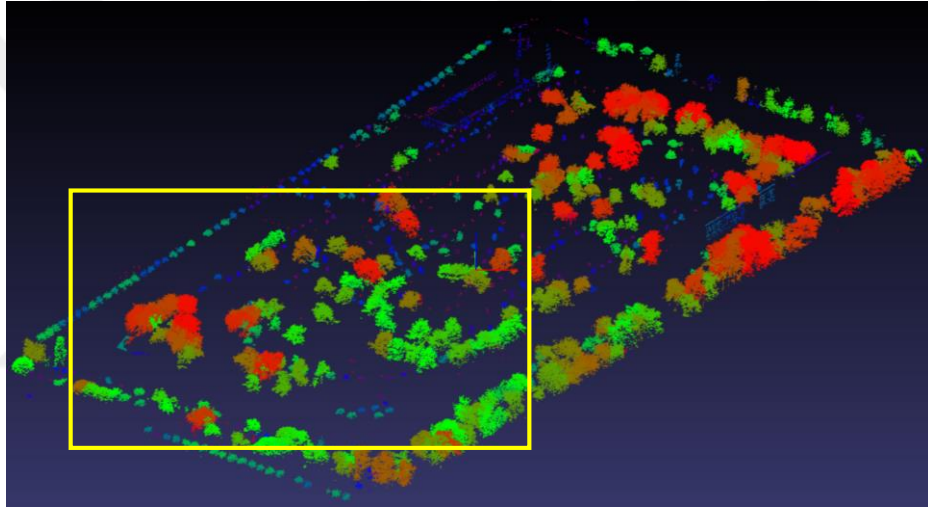
“MeanShift” segmentasyonu sonucu tüm çalışma alanında 705 bireysel ağaç tacı elde edilmiştir (Şekil 4.18). Yüksek bitki örtüsü noktalarının “MeanShift” kümeleme sonucu Şekil 4.19’de yer almaktadır. Kümelerde yer alan en yüksek noktanın görünümü ve en yüksek nokta ile tüm noktaların bir arada görünümü ise Şekil 4.20’de verilmiştir.



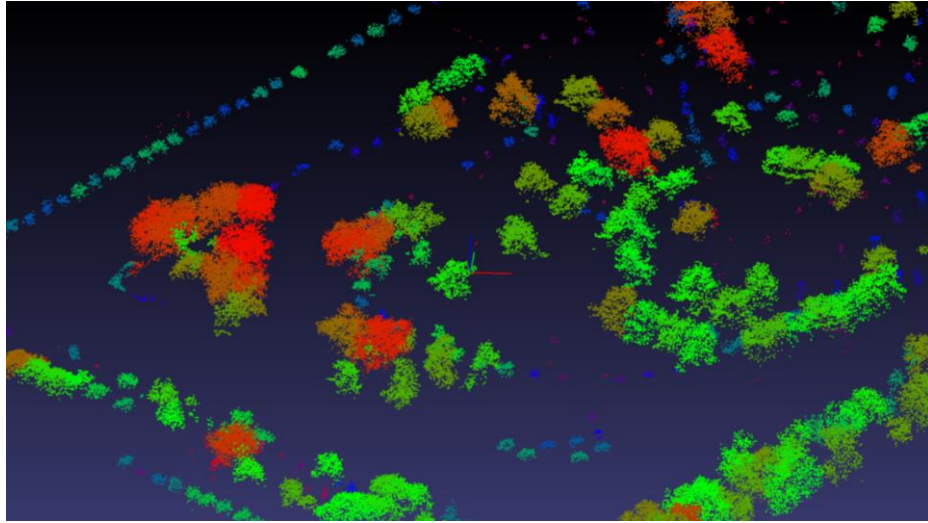
Şekil 4.18 Tüm çalışma alanında “MeanShift” kümeleme sonucunun grafiksel gösterimi (küme merkezlerini içeren)



(a)

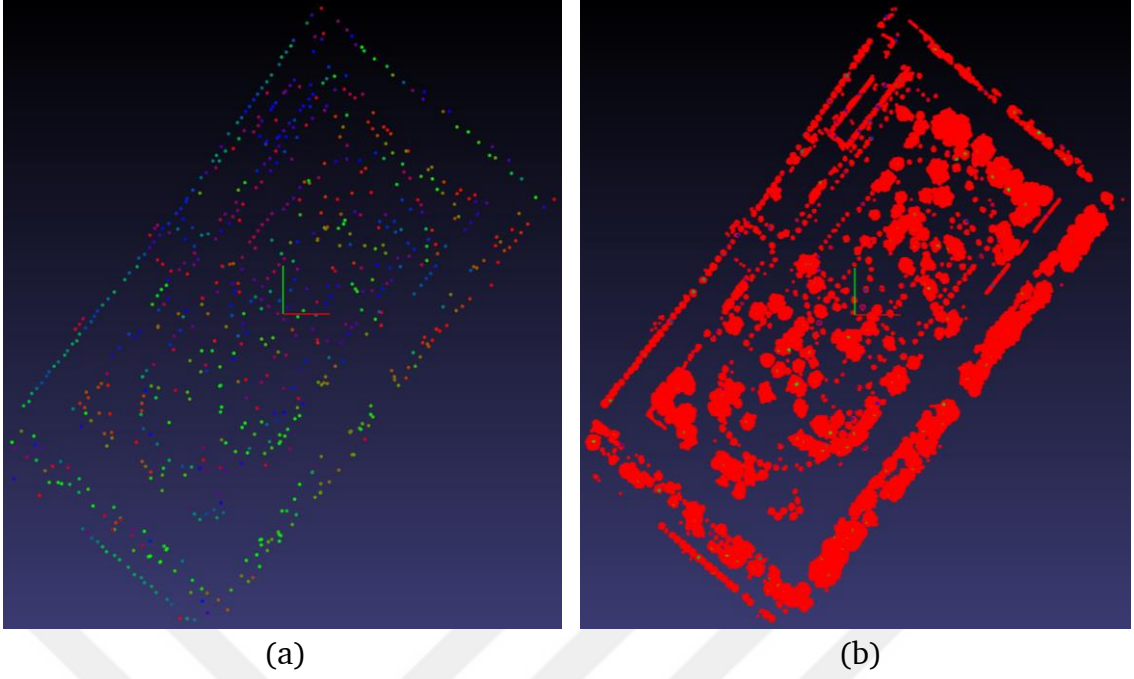


(b)



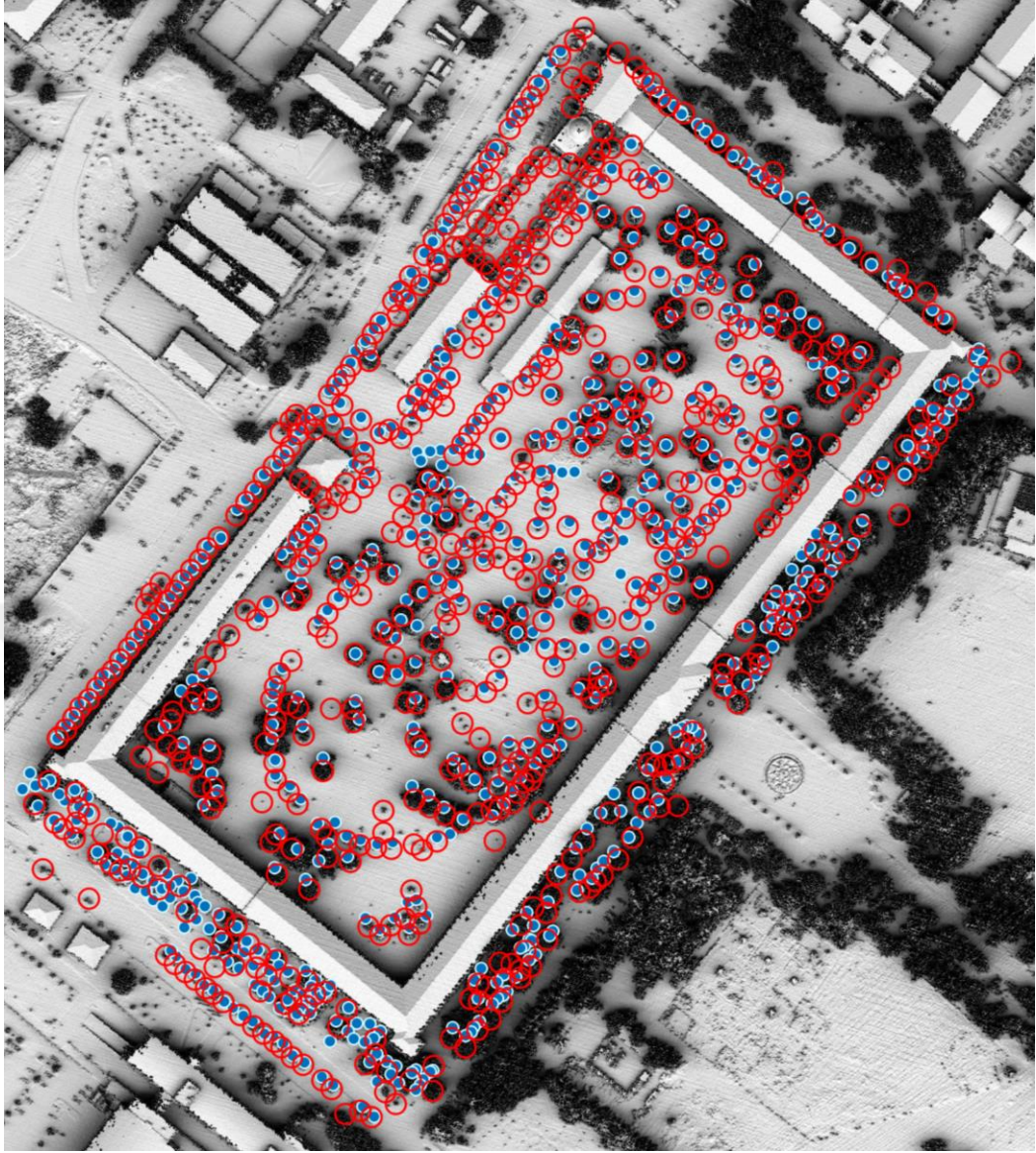
(c)

**Şekil 4.19** Tüm çalışma alanında “MeanShift” kümeleme sonucu: tepeden görünüm (a), perspektif görünüm (b) ve yakın görünüm (sarı dikdörtgen) (c)



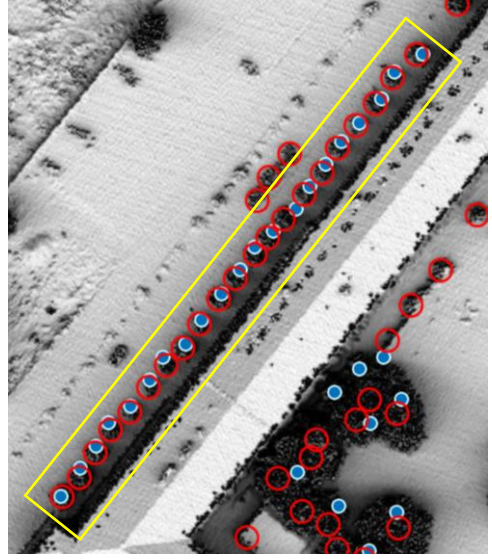
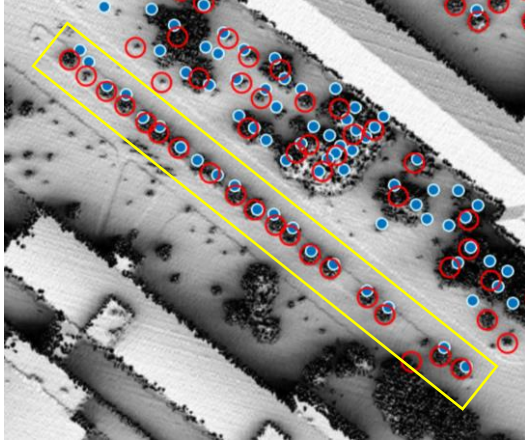
**Şekil 4.20** “MeanShift” kümelemede her bir kümedeki en yüksek nokta (a) ve en yüksek nokta ve tüm yüksek bitki örtüsü sınıfı noktalarının birlikte görünümü (b)

“MeanShift” kümeleme yaklaşımı ile segmente edilen bireysel ağaç taçlarının doğruluğu detaylı bir şekilde analiz edilmiştir. Analiz işleminin gerçekleştirilebilmesi için her bir ağaç tacını ifade eden “MeanShift” kümeleme sonuçları ve referans veri “Şehir Haritası Siyah Beyaz Solar Harita” altlığı üzerinde gösterilmiştir (Şekil 4.21). Sıralı ağaçlarda “MeanShift” kümeleme sonucu örnekleri, tek ağaçlarda “MeanShift” kümeleme sonucu örnekleri ve karışık taç yapısına sahip ağaçlarda “MeanShift” kümeleme sonucu örnekleri sırasıyla Şekil 4.22, Şekil 4.23 ve Şekil 4.24’de verilmiştir. Kümeleme sonuçları incelendiğinde “MeanShift” yöntemi ile gerçekleştirilen segmentasyon sonucunun özellikle sıralı ve tek ağaçlarda oldukça başarılı olduğu görülmüştür. Birbirine karışan karışık taç yapılı ağaçlarda doğru segmentasyonun yanında yetersiz segmentasyon söz konusudur. Bazı ağaç taçları “MeanShift” kümeleme yöntemi ile elde edilememiştir. Buna göre sıralı ve tek ağaçlara göre karışık taç yapılı ağaçlarda yöntemin nispeten daha az başarılı olduğu sonucuna varılmıştır.



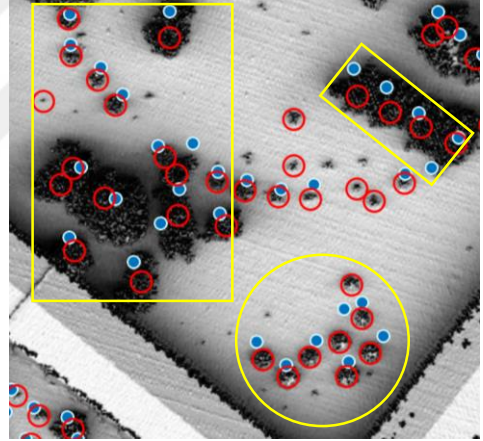
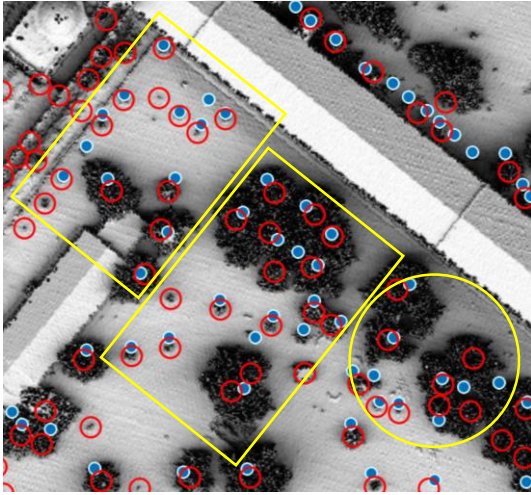
● Referans      ● Kümeleme

Şekil 4.21 “MeanShift” kümeleme sonucunun referans veri ile karşılaştırılması



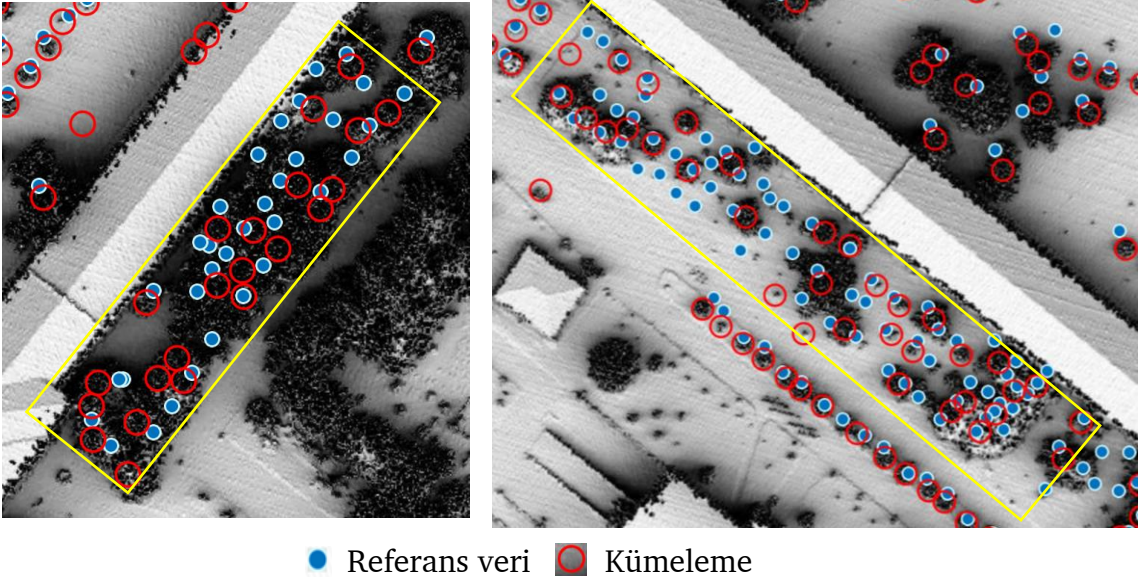
● Referans veri ● Kümeleme

Şekil 4.22 Sıralı ağaçlarda “MeanShift” kümeleme sonucu örnekleri



● Referans veri ● Kümeleme

Şekil 4.23 Tek ağaçlarda “MeanShift” kümeleme sonucu örnekleri

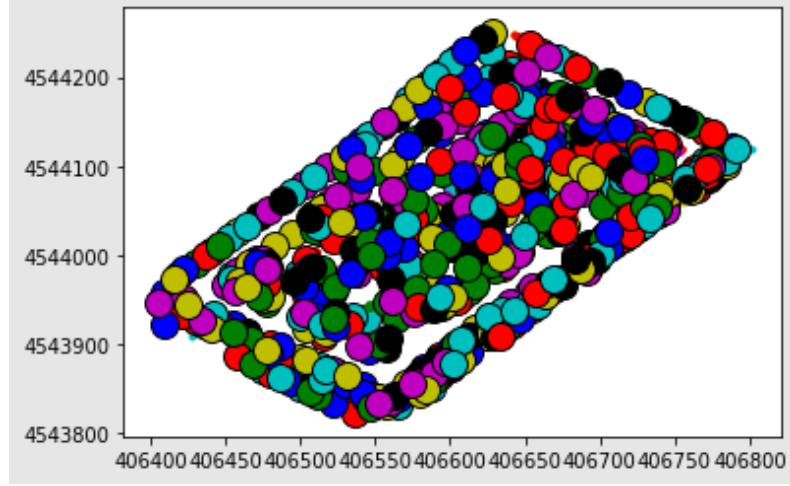


Şekil 4.24 Karışık taç yapılı ağaçlarda “MeanShift” kümeleme sonucu örnekleri

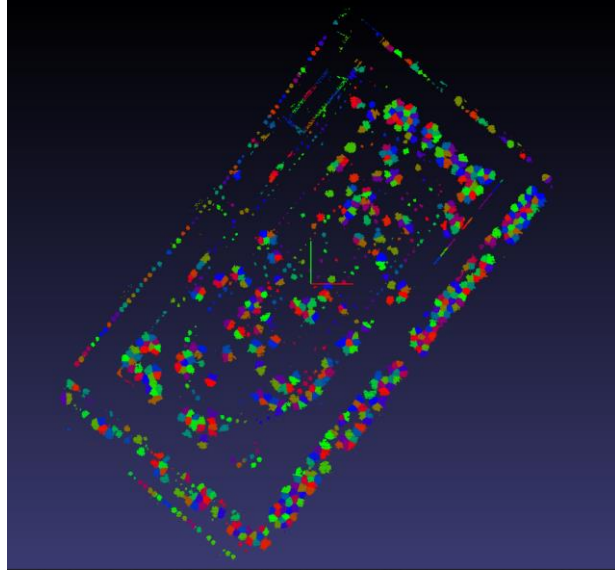
#### 4.5.2 “K-means” Kümeleme ile Bireysel Ağaç Taçlarının Elde Edilmesi

Bireysel ağaç taçlarının elde edilmesi amacıyla kullanılan “K-means” kümeleme algoritmasında küme sayısı bilinmesi gereken bir değişken olduğundan tüm çalışma alanında yer alan 698 ağaç küme sayısı parametresi olarak kullanılmıştır. Parametre analizleri sonucunda “K-means” kümeleme algoritmasında kullanılan, merkez başlatma için rastgele sayı üretimini belirleyen “random state” değeri 0 olarak belirlenmiştir. Yüksek bitki örtüsü noktalarının “K-means” segmentasyonu sonucu tüm çalışma alanında elde edilen 698 bireysel ağaç taçlarının küme merkezlerini içeren grafiksel görünümü Şekil 4.25’de verilmiştir. Yüksek bitki örtüsü noktalarının “K-means” kümeleme sonucu Şekil 4.26’da yer almaktadır. Kümelerde yer alan en yüksek nokta ve en yüksek nokta ile tüm noktaların bir arada görünümü ise Şekil 4.27’de verilmiştir.

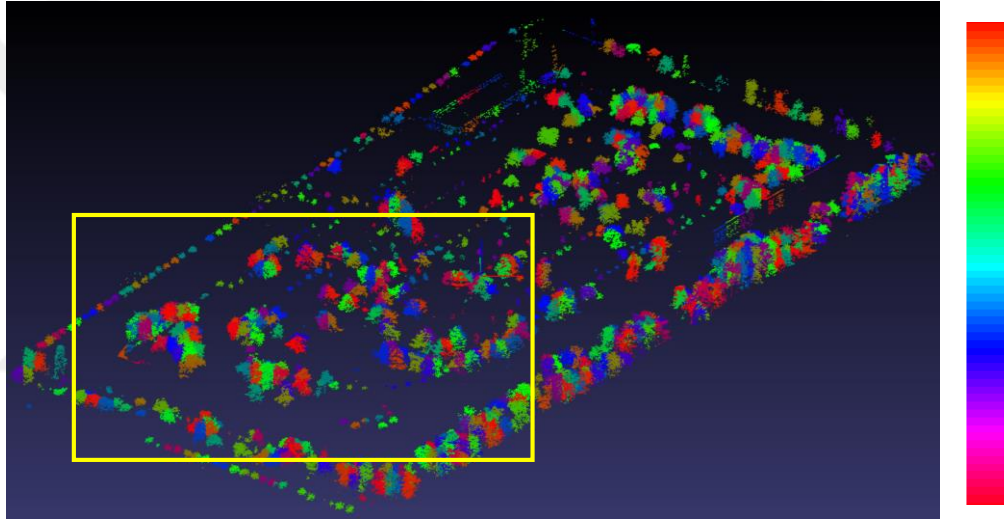
Belirlenen küme sayısı: 698



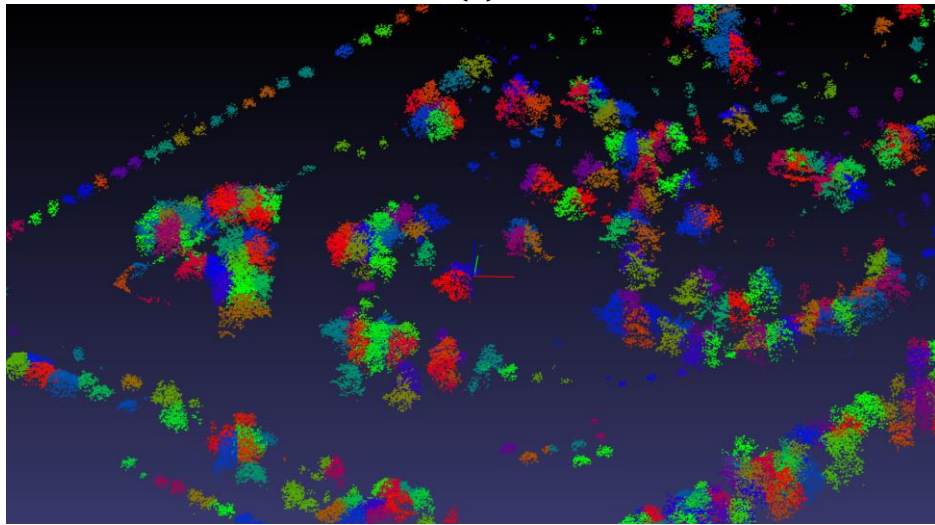
Şekil 4.25 Tüm çalışma alanında “K-means” kümeleme sonucunun grafiksel gösterimi (küme merkezlerini içeren)



(a)

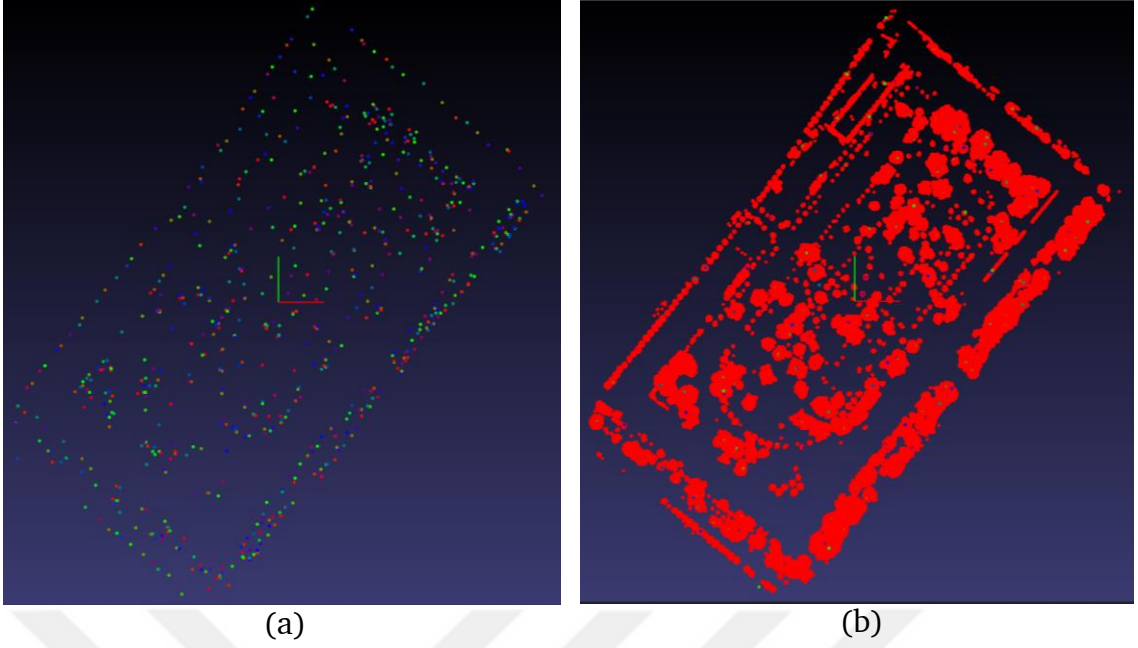


(b)



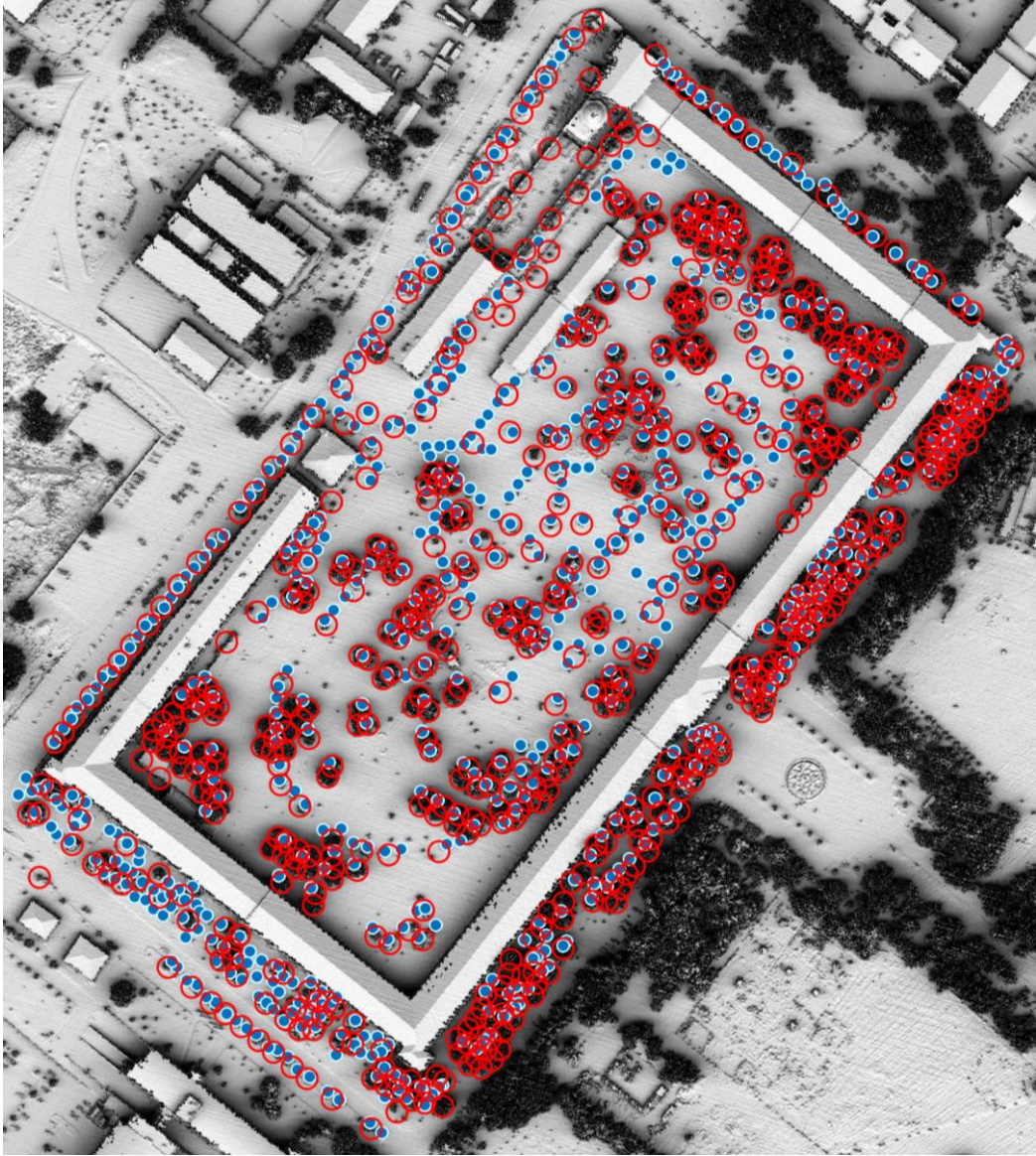
(c)

**Şekil 4.26** Tüm çalışma alanında “K-means” kümeleme sonucu: tepeden görünüm (a), perspektif görünüm (b) ve yakın görünüm (sarı dikdörtgen) (c)



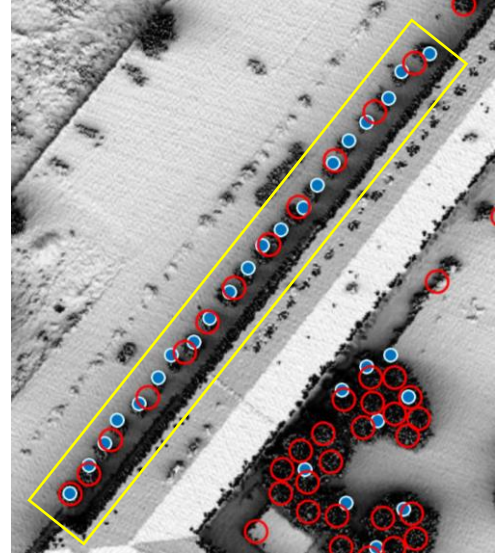
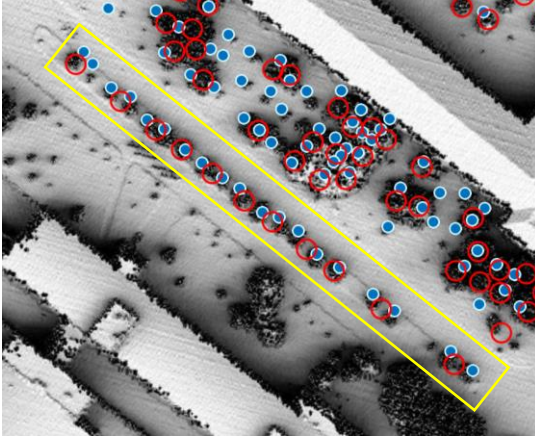
**Şekil 4.27** “K-means” kümelemede her bir kümedeki en yüksek nokta (a) ve en yüksek nokta ve tüm yüksek bitki örtüsü sınıfı noktalarının birlikte görünümü (b)

“K-means” kümeleme yaklaşımı ile segmente edilen bireysel ağaç taçlarının doğruluğu da detaylı bir şekilde analiz edilmiştir. Segmente edilen her bir ağaç tacının yer aldığı “K-means” kümeleme sonuçları ve referans veri analiz işlemi için “Şehir Haritası Siyah Beyaz Solar Harita” altlığı üzerinde gösterilmiştir (Şekil 4.28). Sıralı ağaçlarda “K-means” kümeleme sonucu örnekleri, tek ağaçlarda “K-means” kümeleme sonucu örnekleri ve karışık taç yapısına sahip ağaçlarda “K-means” kümeleme sonucu örnekleri sırasıyla Şekil 4.29, Şekil 4.30 ve Şekil 4.31’de verilmiştir. Kümeleme sonuçları incelendiğinde “K-means” yöntemi ile gerçekleştirilen segmentasyon sonucuna göre sıralı ağaçlarda doğru segmentlerin yanı sıra yetersiz segmentasyon göze çarpmaktadır. Tek ve karışık taç yapılı ağaçlarda ise doğru segmentasyona ek olarak yer yer yetersiz segmentasyon, yer yer de aşırı segmentasyon söz konusudur. Buna göre çalışma alanında “K-means” kümeleme yönteminin yeterli doğruluğa ulaşamadığı sonucuna varılmıştır. Özellikle “MeanShift” kümeleme yaklaşımı ile karşılaştırıldığı “K-means” ile gerçekleştirilen ağaç tacı segmentasyonun daha az başarılı olduğu görülmüştür.



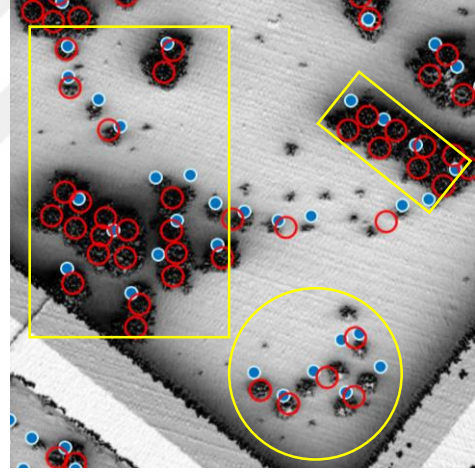
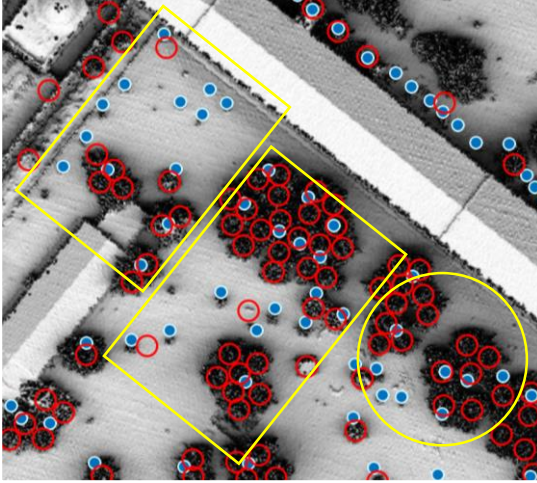
● Referans veri ● Kümeleme

Şekil 4.28 “K-means” kümeleme sonucunun referans veri ile karşılaştırılması



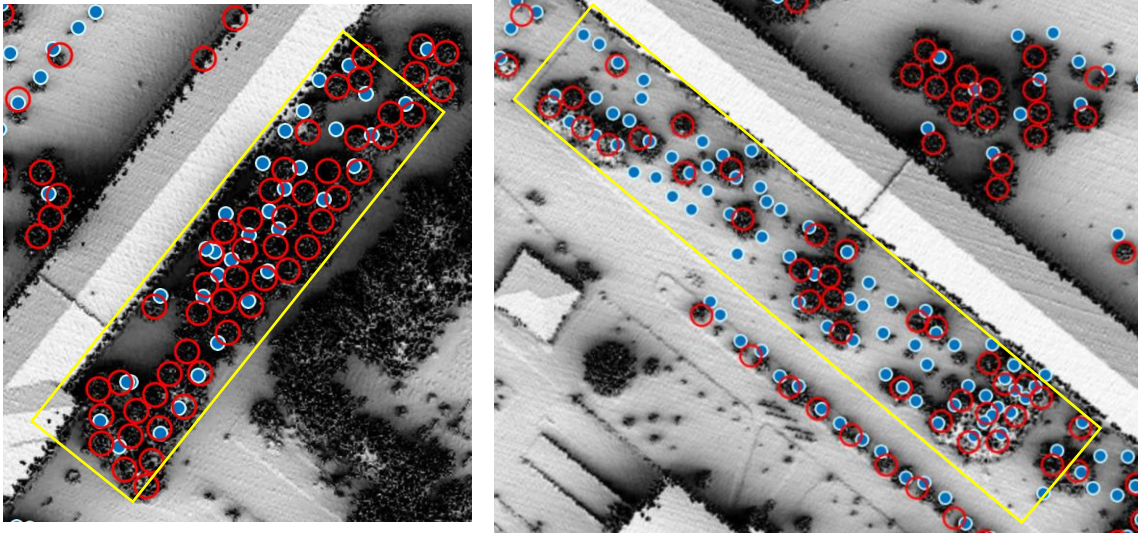
● Referans veri ● Kümeleme

Şekil 4.29 Sıralı ağaçlarda “K-means” kümeleme sonucu örnekleri



● Referans veri ● Kümeleme

Şekil 4.30 Tek ağaçlarda “K-means” kümeleme sonucu örnekleri



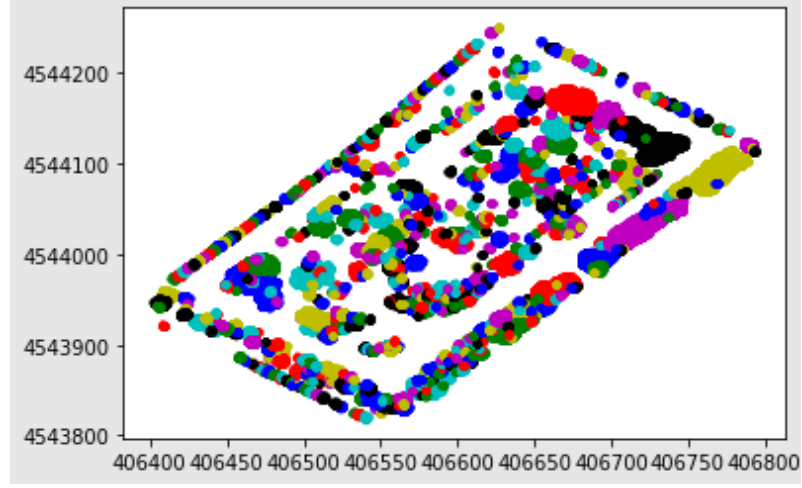
● Referans veri    ◻ Kümeleme

**Şekil 4.31** Karışık taç yapılı ağaçlarda “K-means” kümeleme sonucu örnekleri

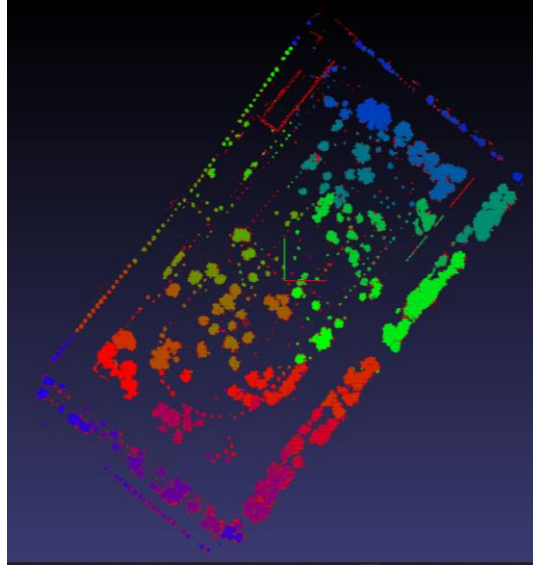
#### 4.5.3 “DBSCAN” Kümeleme ile Bireysel Ağaç Taçlarının Elde Edilmesi

Bireysel ağaç taçlarının segmentasyonu amacıyla “DBSCAN” kümeleme işlemi için gerçekleştirilen parametre analizleri sonucunda, kümeleme algoritmasındaki “eps (biri diğerinin komşuluğunda iki örnek arasındaki maksimum mesafe)” değeri 0.566, “min\_samples (bir komşulukta çekirdek nokta olarak kabul edilecek nokta için örnek sayısı veya toplam ağırlık)” değeri ise 10 olarak belirlenmiştir. Yüksek bitki örtüsü noktalarının “DBSCAN” segmentasyonu sonucu tüm çalışma alanında 699 bireysel ağaç tacı elde edilmiştir (Şekil 4.32). “DBSCAN” kümeleme sonucu Şekil 4.33’de yer almaktadır. Her bir kümede yer alan en yüksek noktaların görünümü ve en yüksek noktalar ile tüm noktaların bir arada görünümü ise Şekil 4.34’de verilmiştir.

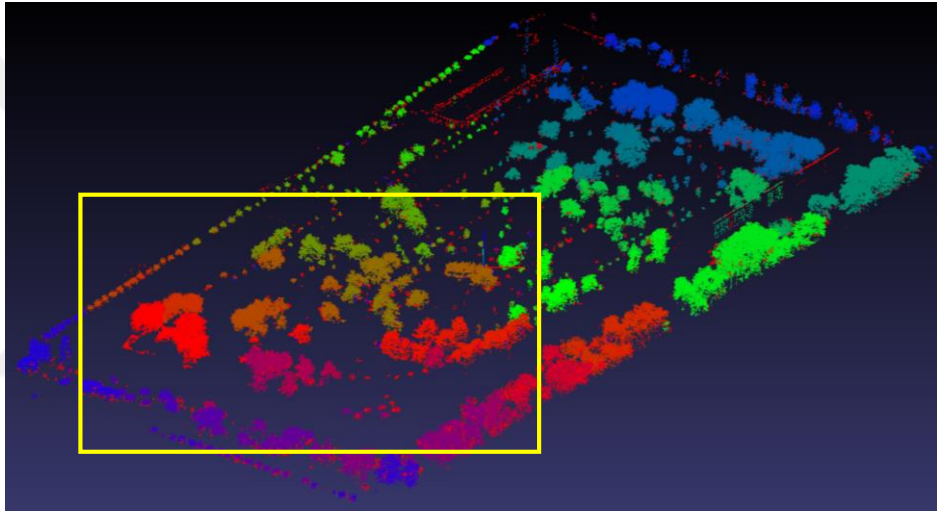
Belirlenen küme sayısı:



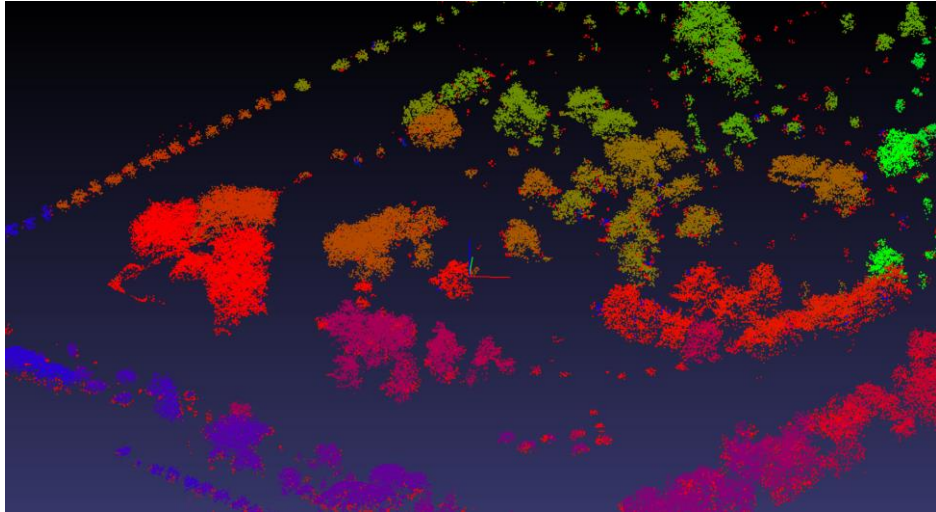
Şekil 4.32 Tüm çalışma alanında “DBSCAN” kümeleme sonucunun grafiksel gösterimi (küme merkezlerini içeren)



(a)

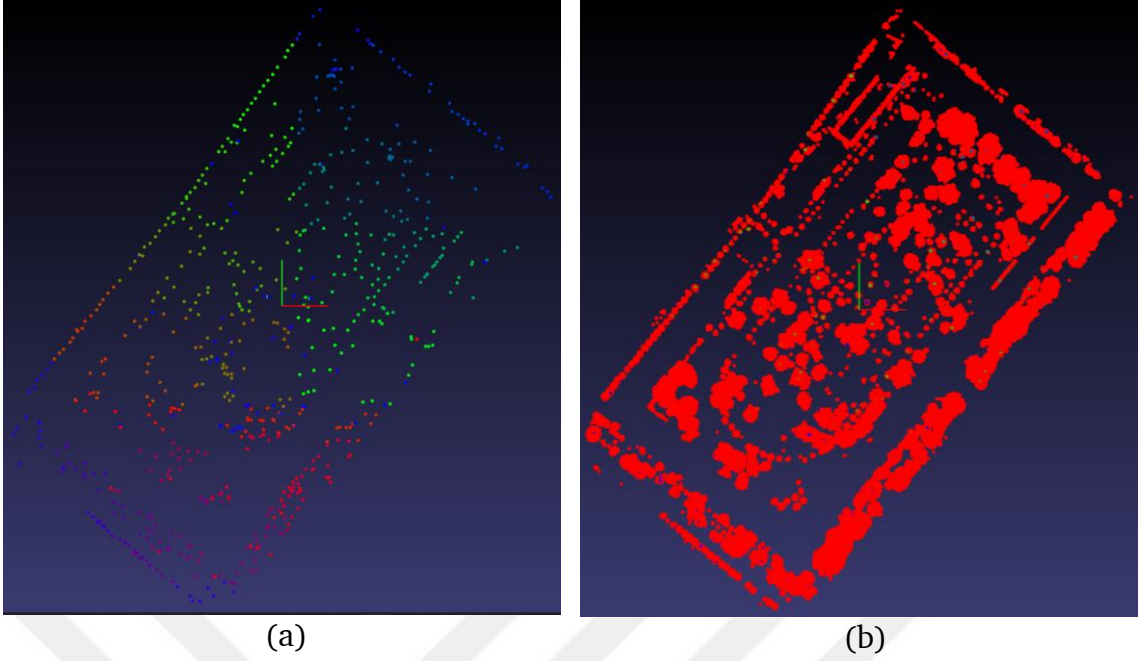


(b)



(c)

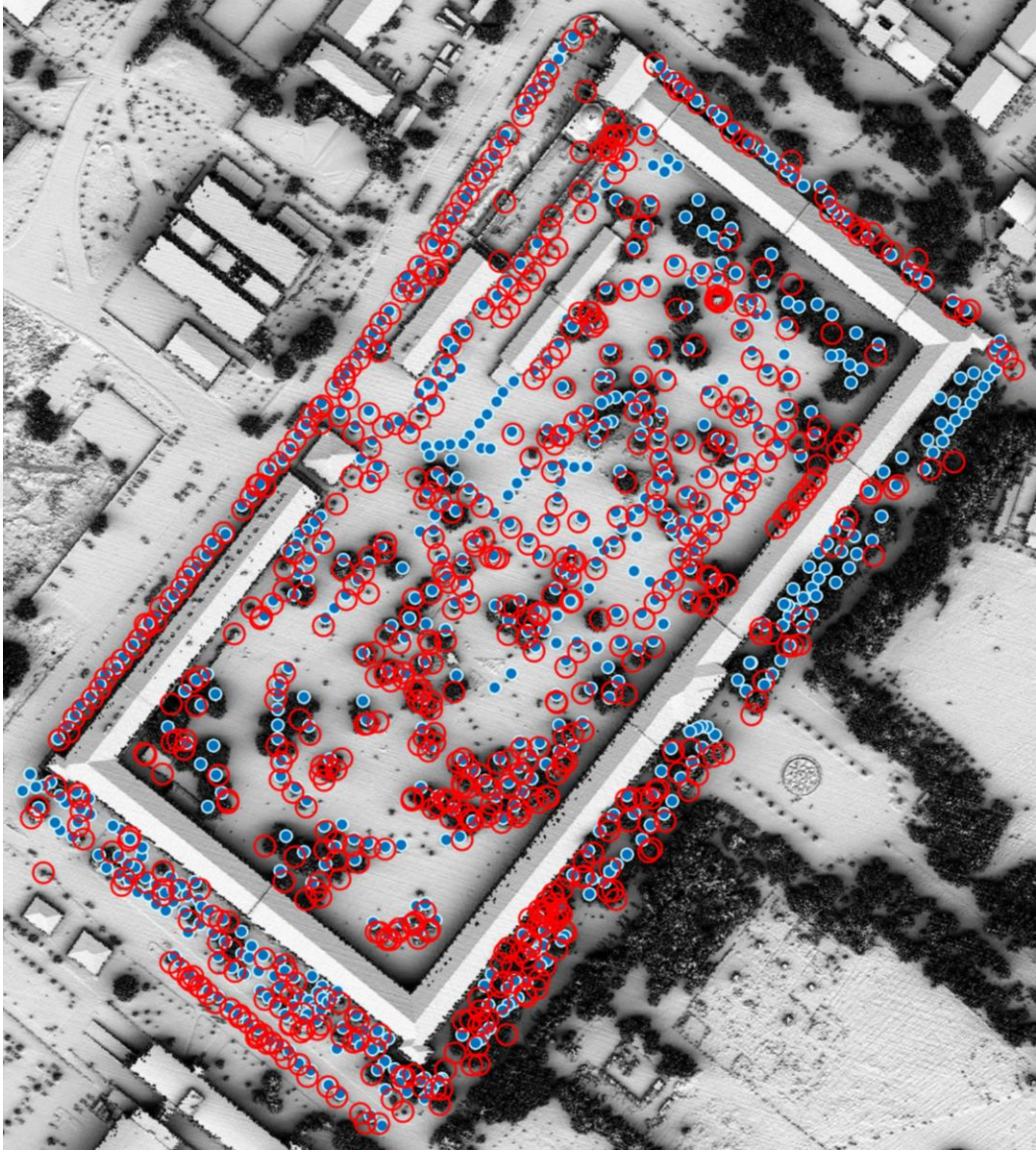
**Şekil 4.33** Tüm çalışma alanında “DBSCAN” kümeleme sonucu: tepeden görünüm (a), perspektif görünüm (b) ve yakın görünüm (sarı dikdörtgen) (c)



**Şekil 4.34** “DBSCAN” kümelemede her bir kümedeki en yüksek nokta (a) ve en yüksek nokta ve tüm yüksek bitki örtüsü sınıfı noktalarının birlikte görünümü (b)

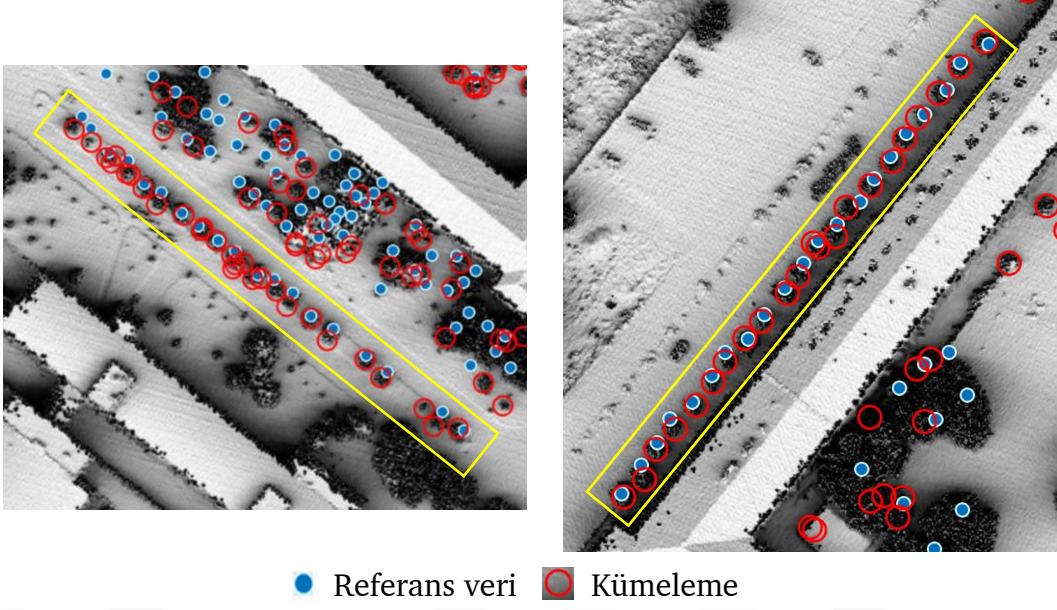
Çalışma bölgesinde “DBSCAN” kümeleme yaklaşımı ile segmentasyon sonucu elde edilen bireysel ağaç taçlarının doğruluğu da detaylı bir şekilde analiz edilmiştir. “DBSCAN” kümeleme ile elde edilen her bir ağaç tacının yer aldığı segmentasyon sonuçları ve referans veri analiz işlemi için “Şehir Haritası Siyah Beyaz Solar Harita” altlığı üzerinde gösterilmiştir (Şekil 4.35). Çalışma alanında bulunan sıralı ağaçlarda “DBSCAN” kümeleme sonucu örnekleri, tek ağaçlarda “DBSCAN” kümeleme sonucu örnekleri ve karışık taç yapısına sahip ağaçlarda “DBSCAN” kümeleme sonucu örnekleri sırasıyla Şekil 4.36, Şekil 4.37 ve Şekil 4.38’de verilmiştir. Kümeleme sonuçları incelendiğinde “DBSCAN” yöntemi ile gerçekleştirilen segmentasyon sonucuna göre sıralı ağaçlarda doğru segmentlerin yanı sıra “K-means” yönteminin tersine aşırı segmentasyon söz konusudur. Tek ve karışık taç yapılı ağaçlarda ise “K-means” yöntemine benzer olarak doğru segmentasyona ek yer yer yetersiz segmentasyon, yer yer de aşırı segmentasyon mevcuttur. Buna göre bireysel ağaç taçlarının elde edilmesi için “DBSCAN” kümeleme yönteminin de yeterli doğruluğa ulaşamadığı sonucuna varılmıştır. “K-means” yöntemi gibi “DBSCAN” yöntemi ile gerçekleştirilen ağaç tacı

segmentasyonun da “MeanShift” ile gerçekleştirilen segmentasyona göre daha az başarılı olduğu görülmektedir.

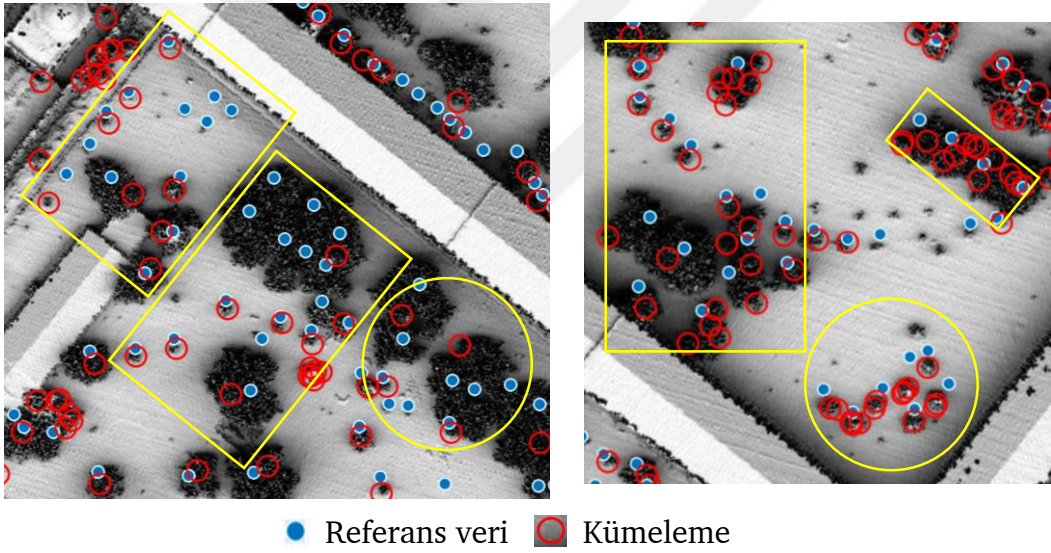


● Referans veri ● Kümeleme

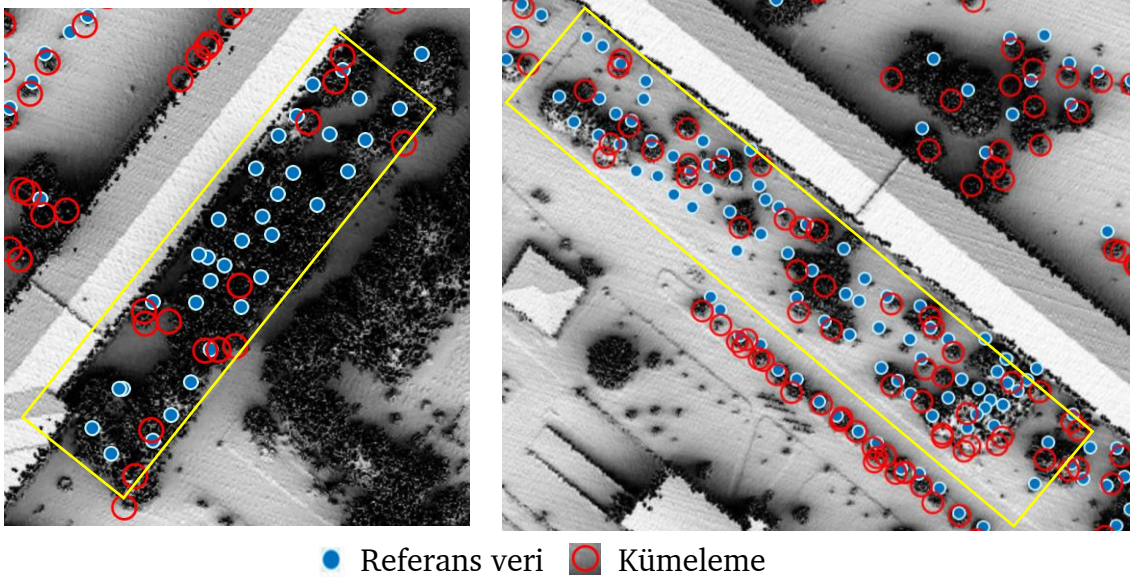
Şekil 4.35 “DBSCAN” kümeleme sonucunun referans veri ile karşılaştırılması



Şekil 4.36 Sıralı ağaçlarda “DBSCAN” kümeleme sonucu örnekleri



Şekil 4.37 Tek ağaçlarda “DBSCAN” kümeleme sonucu örnekleri



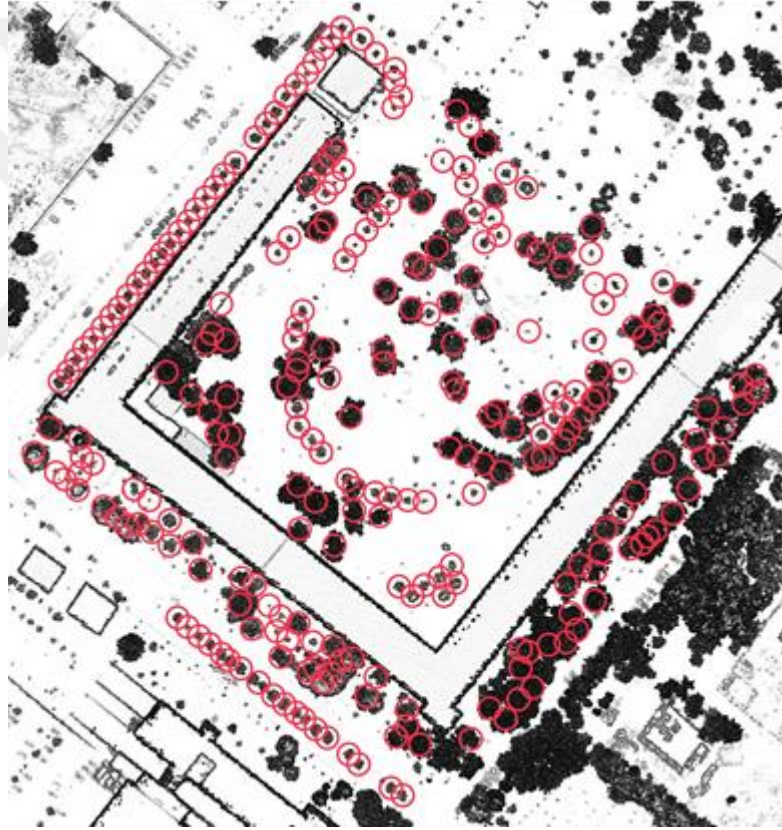
**Şekil 4.38** Karışık taç yapılı ağaçlarda “DBSCAN” kümeleme sonucu örnekleri

#### 4.5.4 Kentsel Ağaç Türlerinin Sınıflandırılmasında Kullanılacak Segmentasyonun Belirlenmesi

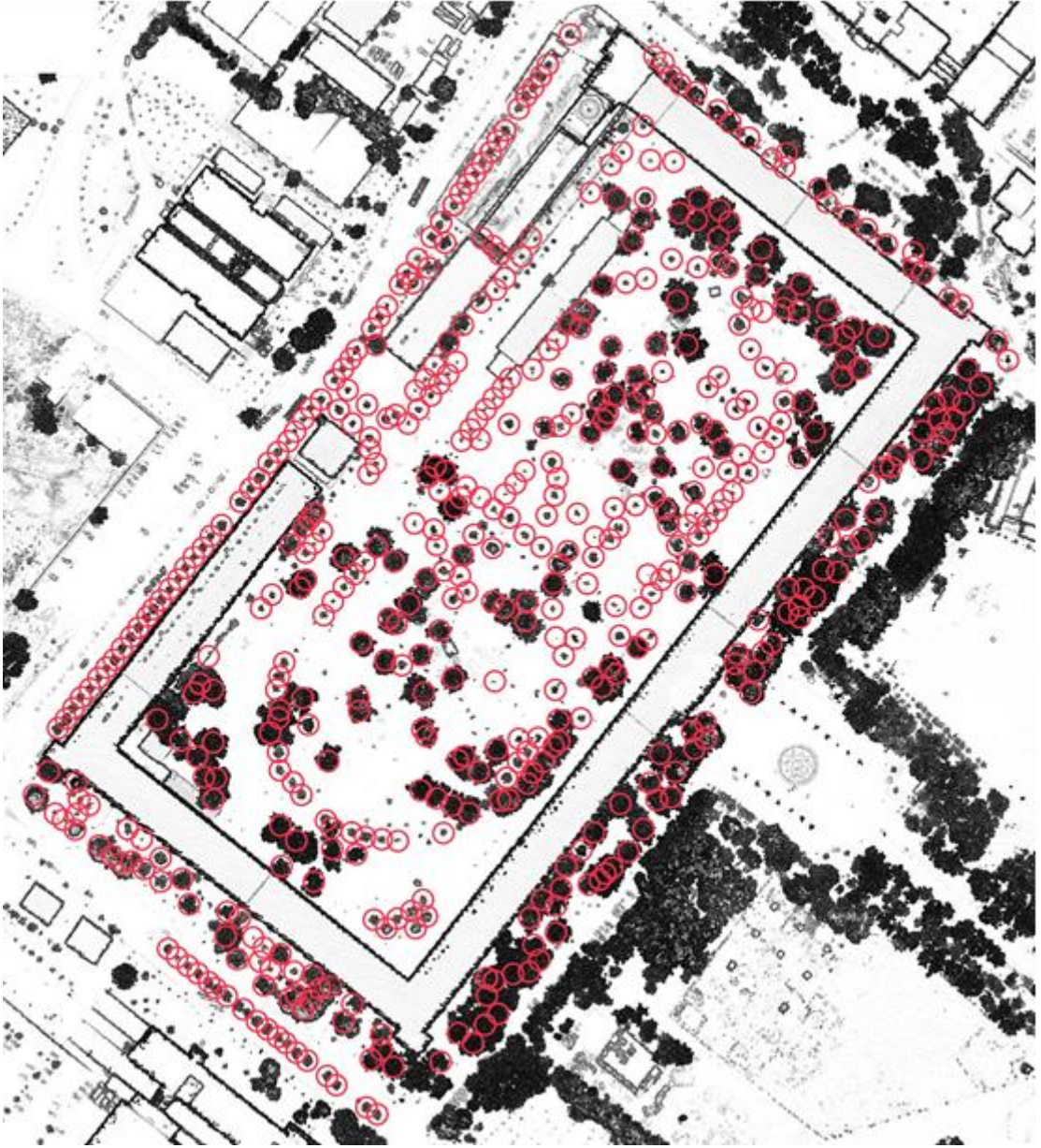
Bu çalışmada, ağaç türleri sınıflandırılmasında kullanılmak üzere gerçekleştirilen 2B nokta tabanlı segmentasyon uygulamalarından en iyi sonucu veren segmentasyon algoritması seçilip elde edilen ağaç taçları makine öğrenme tabanlı sınıflandırma yöntemleri ile sınıflandırılmıştır. Gerçekleştirilen segmentasyon sonuçlarının Şehir Haritası ile karşılaştırılması sonuçları görsel olarak yorumladığında, “MeanShift” segmentasyonun tüm çalışma bölgesini içeren Test alanı 2’de yer alan bireysel ağaç taçlarını kümelemede en iyi yöntem olduğu sonucuna varılmıştır. “K-means” ve “DBSCAN” kümeleme algoritmalarına kıyasla “MeanShift” ile elde edilen segmentasyon sonuçları çok daha doğru olmasına rağmen yine de hatalı segmentleri barındırmaktadır. Doğru tespit edilen ağaçların yanı sıra, yetersiz segmente edilen ağaçlar (birden fazla ağaç tacının tek ağaç olarak belirlenmesi), aşırı segmente edilen ağaçlar (tek ağaç tacının birden fazla ağaç olarak belirlenmesi), kaçırılan ağaçlar ve gürültüler bulunmaktadır.

Bundan sonraki aşamada gerçekleştirilecek kentsel ağaç taçlarının makine öğrenme algoritmaları ile sınıflandırılma performansının doğru bir şekilde test edilebilmesi için doğru segmente edilmiş ağaç taçlarının eğitim ve test verileri olarak kullanılması gerekmektedir. Bu nedenle çalışma alanında doğru segmente

edilen ağa lar korunarak,  zellikle yoęun ve ağa  ta larının birbirine karışmış olduęu b lgelerde elde edilen yetersiz ve aşırı segmentler “MeanShift” k meleme sonu larından  ıkarılmıştır. İlk sonu lara g re 346 bireysel ağa  tacının elde edildięi Test alanı 1’de, hatalı segmentlerin ayrılması sonucu 265 doęru segmente edilmiş bireysel ağa  tacı bulunmaktadır (Şekil 4.39). Buna g re Test alanı 1’de “MeanShift” k meleme ile ger ekleřtirilen segmentasyonun doęruluęu yaklaşık %77 olarak belirlenmiştir. 705 bireysel ağa  tacının elde edildięi Test alanı 2’de ise, hatalı segmentlerin ayrılması sonucu 527 doęru segmente edilmiş bireysel ağa  tacı elde edilmiştir (Şekil 4.40). Test alanı 2’de “MeanShift” k meleme ile ger ekleřtirilen segmentasyonun doęruluęu ise yaklaşık %75 olarak belirlenmiştir.



**Şekil 4.39** Test alanı 1’de yer alan doęru segmente edilmiş bireysel ağa  ta ları (difüzyon harita ile  akıştırılmış)



**Şekil 4.40** Test alanı 2’de yer alan doğru segmente edilmiş bireysel ağaç taçları (difüzyon harita ile çıkarılmış)

#### 4.6 Özellik Çıkarımı

Çalışma bölgesinde bulunan kentsel ağaç türlerinin sınıflandırılması için makine öğrenme algoritmalarının eğitilmesi amacıyla sınıflandırma özelliklerine ihtiyaç vardır. Bu çalışmada, LiDAR nokta bulutundaki her nokta için geometri ve yoğunluk bilgileri kullanılarak hesaplanan sınıflandırma özellikleri belirlenmiştir. Geometri ve yoğunluk tabanlı 8 özellik ve 25 özelliğin yer aldığı iki farklı özellik seti oluşturulmuştur. 8 elemanlı özellik seti ve 25 elemanlı özellik setinde bulunan özellikler sırasıyla Tablo 4.4 ve Tablo 4.5 verilmiştir. Tablo 4.4 ve Tablo 4.5’de yer

alan “Z” her bir LiDAR noktasının ulusal koordinat sistemindeki Z koordinatı ve yoğunluk ise LiDAR noktalarının kalibre edilmemiş yoğunluk değerleridir.

**Tablo 4.4** 8 elemanlı özellik seti

<b>Sınıflandırma özellik seti (8 özellik)</b>	
<b>Geometri tabanlı özellikler</b>	<b>Yoğunluk (intensity) tabanlı özellikler</b>
Nokta sayısı	Yoğunluğun standart sapması
Maksimum Z	Ortalama yoğunluk
Ortalama Z	-
Z'nin “Skewness” değeri	-
Z'nin “Kurtosis” değeri	-
Z aralığı	-

**Tablo 4.5** 25 elemanlı özellik seti

<b>Sınıflandırma özellik seti (25 özellik)</b>	
<b>Geometri tabanlı özellikler</b>	<b>Yoğunluk (intensity) tabanlı özellikler</b>
Nokta sayısı	-
Maksimum Z	Maksimum yoğunluk
Minimum Z	Minimum yoğunluk
Z'nin standart sapması	Yoğunluğun standart sapması
Ortalama Z	Ortalama yoğunluk
Z'nin “Skewness” değeri	Yoğunluğun “Skewness” değeri
Z'nin “Kurtosis” değeri	Yoğunluğun “Kurtosis” değeri
Z aralığı	Yoğunluk aralığı
Z aralığının %5'inde yer alan nokta sayısı	Yoğunluk aralığının %5'inde yer alan nokta sayısı
Z aralığının %25'inde yer alan nokta sayısı	Yoğunluk aralığının %25'inde yer alan nokta sayısı
Z aralığının %50'sinde yer alan nokta sayısı	Yoğunluk aralığının %50'sinde yer alan nokta sayısı
Z aralığının %75'inde yer alan nokta sayısı	Yoğunluk aralığının %75'inde yer alan nokta sayısı
Z aralığının %90'ında yer alan nokta sayısı	Yoğunluk aralığının %90'ında yer alan nokta sayısı

Kentsel ağaç türlerinin sınıflandırılması için belirlenen geometri tabanlı sınıflandırma özelliklerinden “nokta sayısı” her bir kümede yer alan LiDAR nokta sayılarını ifade etmektedir. “Maksimum Z” ve “Minimum Z” sırasıyla her bir kümede yer alan LiDAR noktalarının en büyük Z koordinat değerini ve en küçük Z koordinat değerini göstermektedir. Her bir ağaç segmentinde yer alan noktaların Z koordinatlarının standart sapma değeri “Z'nin standart sapması” özelliğidir. “Ortalama Z” değeri her bir kümede bulunan noktaların Z koordinatlarının ortalama değerini ifade etmektedir. Benzer şekilde “Z'nin skewness değeri” ve “Z'nin kurtosis değeri” de sırasıyla her bir kümede yer alan lazer noktalarının Z koordinatlarının istatistiksel olarak hesaplanan “Skewness” ve “Kurtosis” değerleridir. “Z aralığı” her bir ağaç tacında bulunan en büyük Z koordinatı ile en küçük Z koordinatı arasındaki farkı belirtmektedir. “Z aralığının %5'inde yer alan nokta sayısı”, “Z aralığının %25'inde yer alan nokta sayısı”, “Z aralığının %50'sinde yer alan nokta sayısı”, “Z aralığının %75'inde yer alan nokta sayısı” ve “Z aralığının %90'ında yer alan nokta sayısı” özellikleri sırasıyla tüm çalışma alanında en büyük Z koordinatı ile en küçük Z koordinatı arasındaki farkın %5'inin altındaki aralıkta her bir küme içerisinde bulunan nokta sayısını, %25'inin altındaki aralıkta her bir küme içerisinde bulunan nokta sayısını, %50'sinin altındaki aralıkta her bir küme içerisinde bulunan nokta sayısını, %75'inin altındaki aralıkta her bir küme içerisinde bulunan nokta sayısını ve %90'ının altındaki aralıkta her bir küme içerisinde bulunan nokta sayısını ifade etmektedir.

Kentsel ağaç türlerinin sınıflandırılması için belirlenen yoğunluk tabanlı sınıflandırma özelliklerinden “Maksimum yoğunluk” ve “Minimum yoğunluk” sırasıyla her bir kümede yer alan LiDAR noktalarının en büyük yoğunluk değerini ve en küçük yoğunluk değerini ifade etmektedir. Her bir ağaç segmentinde yer alan noktaların yoğunluk değerlerinden hesaplanan standart sapma değeri “yoğunluğun standart sapması”dır. “Ortalama yoğunluk” değeri her bir kümede bulunan noktaların yoğunluk değerlerinin ortalama değerini göstermektedir. “Yoğunluğun skewness değeri” ve “yoğunluğun kurtosis değeri” de sırasıyla her bir kümede yer alan lazer noktalarının yoğunluk değerlerinin istatistiksel olarak hesaplanan “Skewness” ve “Kurtosis” değerleridir. “Yoğunluk aralığı” her bir ağaç tacında bulunan noktaların en büyük yoğunluk değeri ile en küçük yoğunluk

değeri arasındaki farkı belirtmektedir. “Yoğunluk aralığının %5’inde yer alan nokta sayısı”, “yoğunluk aralığının %25’inde yer alan nokta sayısı”, “yoğunluk aralığının %50’sinde yer alan nokta sayısı”, “yoğunluk aralığının %75’inde yer alan nokta sayısı” ve “yoğunluk aralığının %90’ında yer alan nokta sayısı” özellikleri sırasıyla tüm çalışma alanında en büyük yoğunluk ile en küçük yoğunluk arasındaki farkın %5’inin altındaki aralıkta her bir küme içerisinde bulunan nokta sayısını, %25’inin altındaki aralıkta her bir küme içerisinde bulunan nokta sayısını, %50’sinin altındaki aralıkta her bir küme içerisinde bulunan nokta sayısını, %75’inin altındaki aralıkta her bir küme içerisinde bulunan nokta sayısını ve %90’ının altındaki aralıkta her bir küme içerisinde bulunan nokta sayısını göstermektedir.

Kentsel ağaç türlerinin belirlenmesi için her bir ağaç tacına ilişkin belirlenen sınıflandırma özelliklerinin hesaplanması gerekmektedir. Tüm LiDAR noktalarının bireysel özellikleri kullanılarak her bir ağaç tacının özellik değerleri hesaplanmıştır. Tablo 4.6 ve Tablo 4.7’de sırasıyla Test alanı 1’deki ağaç segmentlerine ilişkin hesaplanan 8 ve 25 özellik değerleri görülmektedir. Tablo 4.8 ve Tablo 4.9’da ise sırasıyla Test alanı 2’deki ağaç segmentlerine ilişkin hesaplanan 8 ve 25 özellik değerleri görülmektedir.

**Tablo 4.6** Test alanı 1’de yer alan her bir bireysel ağaç tacı segmentlerinin hesaplanan 8 özelliği

<b>Sınıflandırma Özelliklerinin Hesabı</b>						
		<b>Segment numarası (Kümeler)</b>				
		<b>0</b>	<b>1</b>	<b>2</b>	<b>...</b>	<b>264</b>
<b>Özellikler</b>	<b>Nokta sayısı</b>	2349	1783	2034	...	4
	<b>Max_Z (m)</b>	92.15	92.06	90.82	...	78.69
	<b>Ortalama_Z (m)</b>	87.75673	87.15456	86.01412	...	78.4575
	<b>Skew_Z</b>	-0.45965	-0.19981	-0.0748	...	0.134795
	<b>Kurtosis_Z</b>	-0.00952	-0.25729	-0.90935	...	-3.1273
	<b>Z aralığı (m)</b>	11.12	10.89	9.63	...	0.45
	<b>Std_yoğunluk</b>	13348.71	13233.46	13078.92	...	13380.41
	<b>Ortalama_yoğunluk</b>	19879.99	17711.55	22252.83	...	12990.5

Tablo 4.7 Test alanı 1’de yer alan her bir bireysel ağaç tacı segmentlerinin hesaplanan 25 özelliği

Sınıflandırma Özelliklerinin Hesabı						
		Segment numarası (Kümeler)				
		0	1	2	...	264
Özellikler	Nokta sayısı	2349	1783	2034	...	4
	Max_Z (m)	92.15	92.06	90.82	...	78.69
	Min_Z (m)	81.03	81.17	81.19	...	78.24
	Std_Z (m)	2.03888	2.02896	2.123756	...	0.20468
	Ortalama_Z (m)	87.7567	87.1546	86.01412	...	78.4575
	Skew_Z	-0.4597	-0.19981	-0.0748	...	0.13480
	Kurtosis_Z	-0.0095	-0.25729	-0.90935	...	-3.1273
	Z aralığı (m)	11.12	10.89	9.63	...	0.45
	Z_5_nokta sayısı	0	0	0	...	0
	Z_25_nokta sayısı	1	0	0	...	4
	Z_50_nokta sayısı	398	476	960	...	4
	Z_75_nokta sayısı	2226	1731	2033	...	4
	Z_90_nokta sayısı	2349	1783	2034	...	4
	Max_yoğunluk	33298	33262	33265	...	33059
	Min_yoğunluk	5993	5987	5999	...	6059
	Std_yoğunluk	13348.7	13233.5	13078.92	...	13380.4
	Ortalama_yoğunluk	19880.0	17711.6	22252.83	...	12990.5
	Skew_yoğunluk	-0.0538	0.27290	-0.412	...	1.99873
	Kurtosis_yoğunluk	-1.9985	-1.92738	-1.83171	...	3.99578
	Yoğunluk aralığı	27305	27275	27266	...	27000
	Yoğunluk_5_nokta sayısı	1143	1012	812	...	3
	Yoğunluk_25_nokta sayısı	1143	1012	812	...	3
	Yoğunluk_50_nokta sayısı	1143	1012	812	...	3
	Yoğunluk_75_nokta sayısı	1143	1012	812	...	3
	Yoğunluk_90_nokta sayısı	1143	1012	812	...	3

**Tablo 4.8** Test alanı 2’de her bir bireysel ağaç tacı segmentlerinin hesaplanan 8 özelliği

Sınıflandırma Özelliklerinin Hesabı						
		Segment numarası (Kümeler)				
		0	1	2	...	526
Özellikler	Nokta sayısı	2349	1783	2034	...	6
	Max_Z (m)	92.15	92.06	90.82	...	82.69
	Ortalama_Z (m)	87.75673	87.15456	86.01412	...	82.10667
	Skew_Z	11.12	10.89	9.63	...	0.87
	Kurtosis_Z	-0.45965	-0.19981	-0.0748	...	1.246172
	Z_aralığı (m)	-0.00952	-0.25729	-0.90935	...	0.391101
	Std_yoğunluk	19879.99	17711.55	22252.83	...	28433
	Ortalama_yoğunluk	13348.71	13233.46	13078.92	...	10806.83

**Tablo 4.9** Test alanı 2’de yer alan her bir bireysel ağaç tacı segmentlerinin hesaplanan 25 özelliği

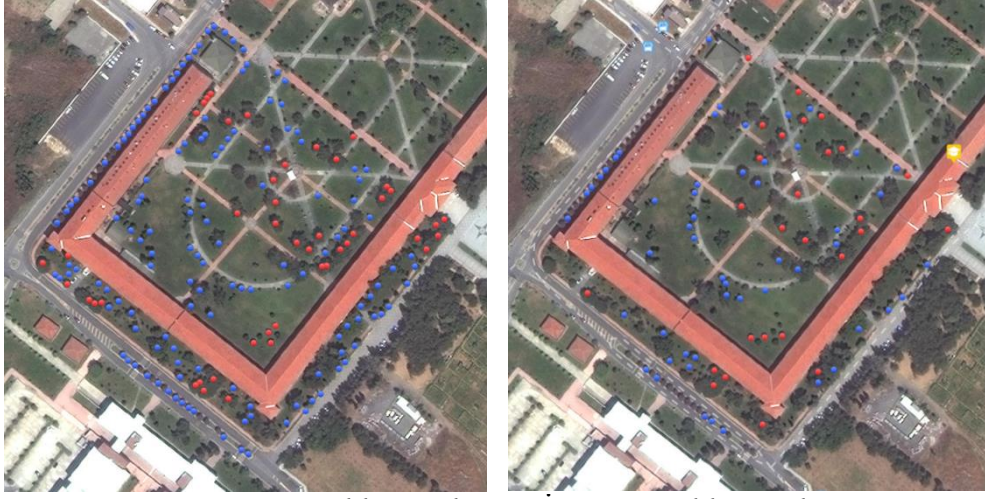
Sınıflandırma Özelliklerinin Hesabı						
		Segment numarası (Kümeler)				
		0	1	2	...	526
Özellikler	Nokta sayısı	2349	1783	2034	...	6
	Max_Z (m)	92.15	92.06	90.82	...	82.69
	Min_Z (m)	81.03	81.17	81.19	...	81.82
	Std_Z (m)	2.03888	2.02896	2.12376	...	0.34273
	Ortalama_Z (m)	87.7567	87.1546	86.0141	...	82.1067
	Skew_Z	-0.4597	-0.19981	-0.0748	...	1.24617
	Kurtosis_Z	-0.0095	-0.25729	-0.90935	...	0.39110
	Z_aralığı (m)	11.12	10.89	9.63	...	0.87
	Z_5_nokta sayısı	0	0	0	...	0
	Z_25_nokta sayısı	14	6	14	...	0
	Z_50_nokta sayısı	868	892	1322	...	6
	Z_75_nokta sayısı	2349	1783	2034	...	6
	Z_90_nokta sayısı	2349	1783	2034	...	6
	Max_yoğunluk	33298	33262	33265	...	32960
	Min_yoğunluk	5993	5987	5999	...	6374
	Std_yoğunluk	13348.7	13233.4	13078.9	...	10806.8
	Ortalama_yoğunluk	19880.0	17711.6	22252.8	...	28433
	Skew_yoğunluk	-0.0538	0.27290	-0.412	...	-2.44932
	Kurtosis_yoğunluk	-1.9985	-1.92738	-1.83171	...	5.99935
	Yoğunluk_aralığı	27305	27275	27266	...	26586
	Yoğunluk_5_nokta sayısı	1143	1012	812	...	1
	Yoğunluk_25_nokta sayısı	1143	1012	812	...	1
	Yoğunluk_50_nokta sayısı	1143	1012	812	...	1
	Yoğunluk_75_nokta sayısı	1143	1012	812	...	1
	Yoğunluk_90_nokta sayısı	1143	1012	812	...	1

#### 4.7 Ağaç Türlerinin Makine Öğrenme Algoritmaları ile Sınıflandırılması

Bu bölümde, Test alanı 1 ve Test alanı 2'de yüksek bitki örtüsü noktalarının segmentasyonu sonucu elde edilen bireysel ağaç taçlarının geniş ve iğne yapraklı ağaçlar olmak üzere üç farklı makine öğrenme algoritması (DVM, RO ve ÇKA) kullanılarak sınıflandırılması ayrıntılı bir şekilde açıklanacaktır.

Her iki test alanında yer alan kentsel ağaç türlerinin belirlenen geometrik ve yoğunluk tabanlı özelliklerden oluşan iki farklı özellik seti kullanılarak sınıflandırılabilmesi için öncelikle eğitim ve test verileri belirlenmiştir. Makine öğrenimi tabanlı sınıflandırma modellerinin eğitimi ve test edilmesi için gerekli olan referans veri, elde edilen doğru bireysel ağaç segmentlerinin görsel yorumlama ve saha araştırmasına göre geniş veya iğne yapraklı ağaçlar olarak manuel etiketlenmesiyle oluşturulmuştur. Test alanı 1'de 3B nokta tabanlı sınıflandırma için 193 geniş ve 72 iğne yapraklı ağaç olmak üzere toplam 265 ağaç tacı bulunmaktadır. Bu ağaç taçlarının %70'i olan 185 ağaç DVM, RO ve ÇKA sınıflandırıcılarının eğitilmesi için ve kalan %30'u olan 80 ağaç ise sınıflandırıcıların performanslarının belirlenmesi amacıyla test verisi olarak kullanılmıştır (Şekil 4.41). Test alanı 2'de ise 414 geniş ve 113 iğne yapraklı olmak üzere toplam 527 ağaç tacı bulunmaktadır. Benzer şekilde bu ağaç taçlarının %70'i olan 368 ağaç DVM, RO ve ÇKA sınıflandırıcılarının eğitilmesi için ve kalan %30'u olan 159 ağaç ise sınıflandırıcıların performanslarının belirlenmesi amacıyla test verisi olarak kullanılmıştır (Şekil 4.42). Test alanı 1 ve Test alanı 2'de eğitim ve test verilerinde yer alan geniş ve iğne yapraklı ağaçların dağılımı Tablo 4.10'da görülmektedir.

Test alanı 1 ve Test alanı 2'de kentsel ağaç türlerinin elde edilmesi için tüm makine öğrenmesi algoritmalarında (DVM, RO ve ÇKA) aynı rastgele belirlenmiş ağaç tacı verisi eğitim için, aynı rastgele belirlenmiş ağaç tacı verisi test için kullanılmıştır. Böylelikle sınıflandırma işleminin ardından makine öğrenme algoritmalarının performanslarının karşılaştırılması olanaklı olmuştur.



● Geniş yapraklı ağaçlar ● İğne yapraklı ağaçlar  
(a) (b)

Şekil 4.41 Test alanı 1’de rastgele belirlenen eğitim (a) ve test (b) verileri



(a)



- Geniş yapraklı ağalar • İęne yapraklı ağalar  
(b)

Şekil 4.42 Test alanı 2'de rastgele belirlenen eğitim (a) ve test (b) verileri

**Tablo 4.10** Test alanlarının eğitim ve test verileri içerisinde bulunan ağaç türlerinin dağılımı

		Eğitim %70	Test %30	Toplam
Test alanı 1	Geniş yapraklı ağaçlar	140	53	193
	İğne yapraklı ağaçlar	45	27	72
Test alanı 2	Geniş yapraklı ağaçlar	297	117	414
	İğne yapraklı ağaçlar	71	42	113

#### 4.7.1 Destek Vektör Makineleri Sınıflandırması

Test alanı 1 ve Test alanı 2'nin destek vektör makineleri (DVM) algoritması kullanılarak kentsel ağaç türlerinin 3 boyutlu nokta tabanlı sınıflandırılması için öncelikle parametre analizi, sonrasında en iyi sınıflandırma sonucunu veren parametreler kullanılarak iki farklı sınıflandırma özelliği (8 özellik ve 25 özellik) ile sınıflandırma işlemi gerçekleştirilmiştir. Ayrıca DVM sınıflandırma yönteminin başarısı çapraz doğrulama (cross-validation) kullanılarak da test edilmiştir. Test alanı 1 ve Test alanı 2'de her iki özellik setinin kullanıldığı sınıflandırmalar için 10 katlamalı çapraz doğrulama gerçekleştirilmiştir. Parametre analizi, DVM sınıflandırması, sınıflandırmanın doğruluğu ve 10 katlamalı çapraz doğrulama sonuçları bu bölümde irdelenecektir.

##### 4.7.1.1 DVM Sınıflandırması İçin Parametre Analizi

Kentsel ağaç türlerinin makine öğrenme algoritmaları ile doğru bir şekilde sınıflandırılabilmesi için özellik çıkarım işleminin ardından sınıflandırma algoritması için en uygun parametrelerin belirlenmesi gerekmektedir. Bu nedenle burada DVM algoritması için "grid search" yöntemi ile en uygun sınıflandırma parametrelerinin otomatik olarak belirlenmesi amaçlanmıştır. DVM algoritmasında en önemli parametreler olan düzenleme (regularization) parametresi C ve gama ( $\gamma$ ) parametresi için belirlenen parametre uzayı (Tablo 4.11) kullanılarak en iyi sonuç veren parametreler analiz edilmiştir. Parametre analizleri, hızlı olması ve belirlenen özellik uzayını test edebilmesi için Test alanı 1'de, daha güvenli olabilmesi için daha fazla özelliğin bulunduğu 25 sınıflandırma özelliği ile gerçekleştirilmiştir. Parametre analizleri sonucunda C için en uygun

değer 100 000 000,  $\gamma$  için en uygun değer ise  $5.092462164188164 \times 10^{-10}$  olarak belirlenmiştir.

**Tablo 4.11** DVM için “Grid search” yaklaşımında kullanılan parametreler

Parametre uzayı	
C	0.00001, 0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000, 10000, 100000, 1000000, 10000000, 100000000, 1000000000
$\gamma$	'auto', 'scale', 0.1, 0.001, 0.0001, 0.0000000005, 5.092462164188164e-10, 0.0000000006, 0.000000000005

#### 4.7.1.2 Test Alanı 1’in Belirlenen Parametreler ile DVM Sınıflandırması ve Doğruluk Analizi

Test alanı 1, parametre analizi sonucu belirlenen parametreler kullanılarak geniş ve iğne yapraklı ağaçların elde edilebilmesi için DVM algoritması ile nokta tabanlı olarak sınıflandırılmıştır. Sınıflandırma işleminde DVM algoritmasında radyal tabanlı fonksiyon (RBF) çekirdeği kullanılmıştır. En önemli parametreler olan C ve gama ( $\gamma$ ) için sırasıyla belirlenen 100 000 000 ve  $5.092462164188164 \times 10^{-10}$  değerleri kullanılmıştır. Test alanı 1’de makine öğrenimi tabanlı eğitim aşamasının ardından test verilerinin 8 özellik kullanılarak gerçekleştirilen DVM sınıflandırma sonuçları Şekil 4.43 ve bu sınıflandırmaya ilişkin hata (confusion) matrisi Tablo 4.12’de yer almaktadır.



- Geniş yapraklı ağaçlar
- İğne yapraklı ağaçlar

**Şekil 4.43** Test alanı 1’in 8 özellik kullanılarak DVM sınıflandırması

**Tablo 4.12** Test alanı 1'in 8 özellik kullanılarak DVM sınıflandırmasına ilişkin hata matrisi

		Referans veri		
		Geniş	İğne	Toplam
Tahmin verisi	Geniş	42	8	50
	İğne	11	19	30
	Toplam	53	27	80

Test alanı 1'de eğitim aşamalarının ardından 8 özellik kullanılarak gerçekleştirilen DVM sınıflandırma sonuçlarına göre doğru olarak elde edilen geniş ve iğne yaprakları ağaç türlerinin yanında 11 geniş yapraklı ağaç hatalı bir şekilde iğne yapraklı ağaç olarak, 8 iğne yapraklı ağaç ise hatalı bir şekilde geniş yapraklı olarak sınıflandırılmıştır. Test alanındaki toplam ağaç sayılarına göre geniş yapraklı ağaçların iğne yapraklı ağaçlarla karışma oranı, iğne yapraklı ağaçların geniş yapraklı ağaçlarla karışma oranından daha düşüktür. Makine öğrenme tabanlı sınıflandırma sonuçları, test veri setinin (%30) doğruluk (accuracy), hassasiyet (recall), kesinlik (precision) ve F1-skoru ile yorumlanmıştır. Test alanı 1'in 8 özellik kullanılarak gerçekleştirilen DVM sınıflandırma sonuçları aşağıdaki şekildedir:

$$\text{— Toplam doğruluk (overall accuracy)} = (42+19)/(42+19+8+11) = 0.7625 = \%76$$

$$\text{— Hassasiyet (recall) (geniş yapraklı ağaçlar)} = 42/(42+11) = 0.7925$$

$$\text{— Hassasiyet (recall) (iğne yapraklı ağaçlar)} = 19/(19+8) = 0.7037$$

$$\text{— Kesinlik (precision) (geniş yapraklı ağaçlar)} = 42/(42+8) = 0.84$$

$$\text{— Kesinlik (precision) (iğne yapraklı ağaçlar)} = 19/(19+11) = 0.6333$$

$$\text{— F1-skor (geniş yapraklı ağaçlar)} = 2x(0.7925x0.84)/(0.7925+0.84) = 0.8156$$

$$\text{— F1-skor (iğne yapraklı ağaçlar)} = 2x(0.7037x0.6333)/(0.7037+0.6333) = 0.6667$$

Test alanı 1'de yer alan geniş ve iğne yapraklı ağaçların elde edilebilmesi için daha önceden belirlenen 8 sınıflandırma özelliği kullanılarak gerçekleştirilen sınıflandırmanın genel doğruluğu %76 olarak elde edilmiştir. Geniş yapraklı

ağaçlar için hassasiyet, kesinlik ve F1-skor değerleri sırasıyla 0.7925, 0.84 ve 0.8156 olarak hesaplanmıştır. İğne yapraklar ağaçlar ise için ise hassasiyet, kesinlik ve F1 skor değerleri sırasıyla 0.7037, 0.6333 ve 0.6667 olarak hesaplanmıştır. Sınıflandırma sonuçlarına göre DVM algoritmasının 8 özellik kullanılarak ağaç türlerini sınıflandırmada başarılı olduğu ancak sınıflandırma doğruluğunun daha da iyileştirilebileceği sonucuna varılmıştır. Buradaki sonuçlara göre DVM algoritması 8 özellik kullanılarak geniş yapraklı ağaçları sınıflandırmada iğne yapraklı ağaçlara göre daha başarılıdır.

Bu çalışmada sınıflandırma modellerinin başarıları çapraz doğrulama kullanılarak da test edilmiştir. Tüm sınıflandırıcılar için sınıflandırma sonuçlarının güvenilirliğini test etmek ve kararlılığını değerlendirmek, fazla uyumu önlemek için 10 katlamalı çapraz doğrulama gerçekleştirilmiştir. 8 özellikli DVM algoritmasının 10 katlamalı çapraz doğrulamaya (Tablo 4.13) göre ortalama doğruluğu %76 olarak belirlenmiştir. 10 katlamalı çapraz doğrulama sonuçları, DVM algoritmasının %70 eğitim, %30 test bölümlenmesinde tutarlı ve güvenilir sonuçlar ürettiğini göstermektedir.

**Tablo 4.13** Test alanı 1'in 8 özellik kullanılarak DVM sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları

10 Katlamalı Çapraz Doğrulama		
1. katlama doğruluğu	0.8519	Ortalama doğruluk: 0.7624 (%76)
2. katlama doğruluğu	0.7037	
3. katlama doğruluğu	0.7778	
4. katlama doğruluğu	0.8148	
5. katlama doğruluğu	0.6296	
6. katlama doğruluğu	0.8462	
7. katlama doğruluğu	0.6923	
8. katlama doğruluğu	0.7308	
9. katlama doğruluğu	0.7308	
10. katlama doğruluğu	0.8462	

Test alanı 1'de test verilerinin 25 özellik kullanılarak gerçekleştirilen DVM sınıflandırma sonuçları Şekil 4.44 ve bu sınıflandırmaya ilişkin hata matrisi de Tablo 4.14'de verilmiştir.



- Geniş yapraklı ağaçlar ● İğne yapraklı ağaçlar

Şekil 4.44 Test alanı 1'in 25 özellik kullanılarak DVM sınıflandırması

Tablo 4.14 Test alanı 1'in 25 özellik kullanılarak DVM sınıflandırmasına ilişkin hata matrisi

		Referans veri		
		Geniş	İğne	Toplam
Tahmin verisi	Geniş	47	10	57
	İğne	6	17	23
	Toplam	53	27	80

Test alanı 1'de DVM ile 25 özellik kullanılarak gerçekleştirilen sınıflandırmaya göre test alanındaki 47 geniş yapraklı ağaç ve 17 iğne yapraklı ağaç doğru bir şekilde sınıflandırılmıştır. Sadece 6 geniş yapraklı ağaç yanlış bir şekilde iğne yapraklı olarak, 10 iğne yapraklı ağaç ise yanlış bir şekilde geniş yapraklı ağaç olarak sınıflandırılmıştır. 8 özelliğin kullanıldığı DVM sınıflandırmasına benzer şekilde burada da geniş yapraklı ağaçların iğne yapraklı ağaçlarla karışma oranı, iğne yapraklı ağaçların geniş yapraklı ağaçlarla karışma oranından daha düşüktür. Test alanı 1'in 25 özellik kullanılarak gerçekleştirilen DVM sınıflandırma sonuçları aşağıdaki şekildedir:

- Toplam doğruluk (overall accuracy) =  $(47+17)/(47+17+10+6) = 0.80$   
= %80
- Hassasiyet (recall) (geniş yapraklı ağaçlar) =  $47/(47+6) = 0.8868$
- Hassasiyet (recall) (iğne yapraklı ağaçlar) =  $17/(17+10) = 0.6296$

- Kesinlik (precision) (geniş yapraklı ağaçlar) =  $47/(47+10) = 0.8246$
- Kesinlik (precision) (iğne yapraklı ağaçlar) =  $17/(17+6) = 0.7391$
- F1-skor (geniş yapraklı ağaçlar) =  $2x(0.8868x0.8246)/(0.8868+0.8246)$   
= 0.8545
- F1-skor (iğne yapraklı ağaçlar) =  $2x(0.6296x0.7391)/(0.6296+0.7391)$   
= 0.68

Test alanı 1’de 25 sınıflandırma özelliği kullanılarak gerçekleştirilen DVM sınıflandırmasının genel doğruluğu %80’dir. Hassasiyet, kesinlik ve F1-skor değerleri geniş yapraklı ağaçlar için sırasıyla 0.8868, 0.8246 ve 0.8545 olarak elde edilmiştir. Hassasiyet, kesinlik ve F1-skor değerleri, iğne yapraklar ağaçlar için ise sırasıyla 0.6296, 0.7391 ve 0.68’dir. Sınıflandırma sonuçlarına göre DVM algoritmasının 25 özellik kullanılarak ağaç türlerini sınıflandırmada 8 özelliğe göre daha başarılı olduğu görülmüştür. Buradaki sonuçlara göre DVM algoritması, 8 özelliğe olduğu gibi 25 özellik kullanılarak geniş yapraklı ağaçları sınıflandırmada iğne yapraklı ağaçlara göre daha başarılıdır. Test alanı 1’de 25 özellik kullanılarak gerçekleştirilen DVM algoritmasının 10 katlamalı çapraz doğrulamaya (Tablo 4.15) göre ortalama doğruluğu %81 olarak belirlenmiştir. 25 özellik kullanılarak %70 eğitim, %30 test bölümlenmesinde gerçekleştirilen DVM sınıflandırması sonucu 10 katlamalı çapraz doğrulama sonuçları ile benzerdir.

**Tablo 4.15** Test alanı 1’in 25 özellik kullanılarak DVM sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları

10 Katlamalı Çapraz Doğrulama		
1. katlama doğruluğu	0.8148	<b>Ortalama doğruluk:</b> 0.8110 (%81)
2. katlama doğruluğu	0.6296	
3. katlama doğruluğu	0.9259	
4. katlama doğruluğu	0.9630	
5. katlama doğruluğu	0.8148	
6. katlama doğruluğu	0.8462	
7. katlama doğruluğu	0.8077	
8. katlama doğruluğu	0.7308	
9. katlama doğruluğu	0.7692	
10. katlama doğruluğu	0.8077	

#### 4.7.1.3 Test Alanı 2'nin Belirlenen Parametreler ile DVM Sınıflandırması ve Doğruluk Analizi

Test alanı 2, Test alanı 1'e benzer şekilde parametre analizi sonucu en iyi sonucu veren parametreler kullanılarak geniş ve iğne yapraklı kentsel ağaçların elde edilebilmesi için DVM algoritması ile 3B nokta tabanlı olarak sınıflandırılmıştır. Burada da sınıflandırma işleminde DVM algoritmasında radyal tabanlı fonksiyon (RBF) çekirdeği kullanılmış ve C ve gama ( $\gamma$ ) parametreleri ise sırasıyla 100 000 000 ve  $5.092462164188164 \times 10^{-10}$  olarak alınmıştır. Şekil 4.45'de Test alanı 2'de test verilerinin 8 özellik kullanılarak gerçekleştirilen DVM sınıflandırma sonuçları ve Tablo 4.16'da bu sınıflandırma sonuçlarına ilişkin hata matrisi verilmiştir.



● Geniş yapraklı ağaçlar ● İğne yapraklı ağaçlar

Şekil 4.45 Test alanı 2'nin 8 özellik kullanılarak DVM sınıflandırması

**Tablo 4.16** Test alanı 2'nin 8 özellik kullanılarak DVM sınıflandırmasına ilişkin hata matrisi

		Referans veri		
		Geniş	İğne	Toplam
Tahmin verisi	Geniş	106	34	140
	İğne	11	8	19
	Toplam	117	42	159

Test alanı 2'de DVM algoritması ile 8 özellik kullanılarak gerçekleştirilen sınıflandırma ile geniş yapraklı ağaçlar büyük oranda doğru sınıflandırılmıştır (106 geniş yapraklı ağaç). Geniş yapraklı ağaçların tersine sadece 8 iğne yapraklı ağaç doğru sınıflandırılmıştır. Test alanında yalnızca 11 geniş yapraklı ağaç iğne yapraklı ağaç olarak sınıflandırılırken, 34 iğne yapraklı ağaç geniş yapraklı olarak sınıflandırılmıştır. Test alanı 2'de iğne yapraklı ağaçların sınıflandırılma oranı geniş yapraklı ağaçlara göre düşüktür. Test alanı 2'nin 8 özellik kullanılarak gerçekleştirilen DVM sınıflandırma sonuçları aşağıdaki şekildedir:

— Toplam doğruluk (overall accuracy) =  $(106+8)/(106+8+34+11) = 0.7170 = \%72$

— Hassasiyet (recall) (geniş yapraklı ağaçlar) =  $106/(106+11) = 0.9060$

— Hassasiyet (recall) (iğne yapraklı ağaçlar) =  $8/(8+34) = 0.1905$

— Kesinlik (precision) (geniş yapraklı ağaçlar) =  $106/(106+34) = 0.7571$

— Kesinlik (precision) (iğne yapraklı ağaçlar) =  $8/(8+11) = 0.4211$

— F1-skor (geniş yapraklı ağaçlar) =  $2x(0.9060x0.7571)/(0.9060+0.7571) = 0.8249$

— F1-skor (iğne yapraklı ağaçlar) =  $2x(0.1905x0.4211)/(0.1905+0.4211) = 0.2623$

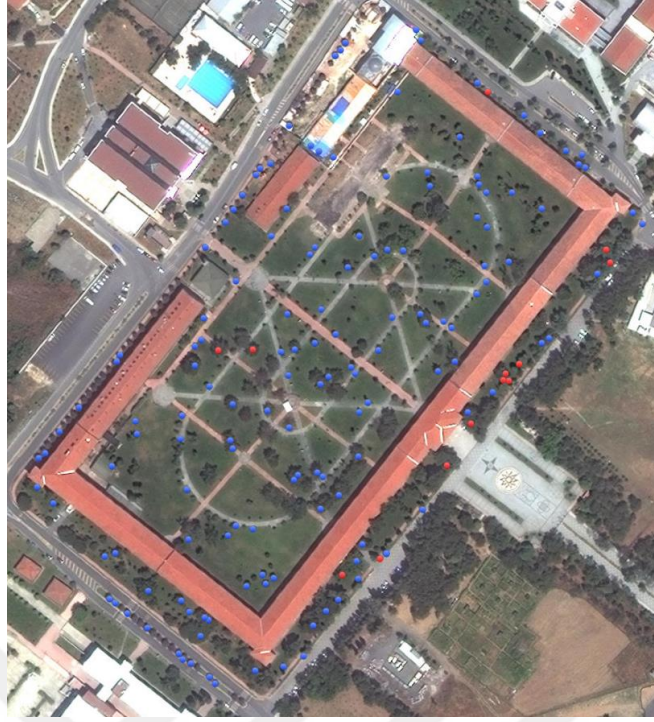
Test alanı 2'nin ağaç türlerinin elde edilmesi amacıyla 8 özellik kullanılarak gerçekleştirilen DVM sınıflandırmasının genel doğruluğu %72 olarak elde edilmiştir. Test alanı 2'de yer alan geniş yapraklı ağaçlar için hassasiyet, kesinlik ve F1-skor değerleri sırasıyla 0.9060, 0.7571 ve 0.8249'dur. İğne yapraklı ağaçlar için ise hassasiyet, kesinlik ve F1-skor değerleri sırasıyla 0.1905, 0.4211 ve 0.2623'tür. DVM algoritması Test alanı 2'de 8 özellik kullanılarak Test alanı 1'e

göre daha düşük doğrulukla kentsel ağaç türlerini belirlemiştir. Test alanı 2’de DVM algoritması ile 8 özellik kullanılarak geniş yapraklı ağaçlar başarılı bir şekilde sınıflandırılrsa da iğne yapraklı ağaçlarda aynı başarı söz konusu olmamıştır. Bu durum, Test alanı 2’de iğne yapraklı ağaçların azlığı ve iğne yapraklı ağaç dağılımının homojen olmamasından kaynaklanmaktadır. Test alanı 2’nin 8 özellik kullanılarak DVM sınıflandırmasının 10 katlamalı çapraz doğrulamaya (Tablo 4.17) göre ortalama doğruluğu %77 olarak belirlenmiştir. Elde edilen sonuçlar 8 özellik kullanılarak DVM sınıflandırmasının Test alanı 2’de de yine tutarlı sonuçlar ürettiğini göstermektedir.

**Tablo 4.17** Test alanı 2’nin 8 özellik kullanılarak DVM sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları

10 Katlamalı Çapraz Doğrulama		
1. katlama doğruluğu	0.7925	<b>Ortalama doğruluk:</b> 0.7780 (%77)
2. katlama doğruluğu	0.8113	
3. katlama doğruluğu	0.6792	
4. katlama doğruluğu	0.7925	
5. katlama doğruluğu	0.7547	
6. katlama doğruluğu	0.8491	
7. katlama doğruluğu	0.7736	
8. katlama doğruluğu	0.75	
9. katlama doğruluğu	0.7692	
10. katlama doğruluğu	0.8077	

Test alanı 2 verileriyle 25 özellik kullanılarak gerçekleştirilen DVM sınıflandırma sonuçları Şekil 4.46’da ve 25 özelliğin kullanıldığı sınıflandırma sonuçlarına ilişkin hata (confusion) matrisi ise Tablo 4.18’de görülmektedir.



● Geniş yapraklı ağaçlar ● İğne yapraklı ağaçlar

**Şekil 4.46** Test alanı 2'nin 25 özellik kullanılarak DVM sınıflandırması

**Tablo 4.18** Test alanı 2'nin 25 özellik kullanılarak DVM sınıflandırmasına ilişkin hata matrisi

		Referans veri		
		Geniş	İğne	Toplam
Tahmin verisi	Geniş	109	36	145
	İğne	8	6	14
	Toplam	117	42	159

Test alanı 2'de 25 özellik kullanılarak gerçekleştirilen DVM ile 109 geniş yapraklı ağaç doğru olarak sınıflandırılırken, yalnızca 6 iğne yapraklı ağaç doğru olarak elde edilmiştir. Geniş yapraklı ağaçların sadece 8 tanesi iğne yapraklı olarak hatalı sınıflandırılmıştır. 36 iğne yapraklı ağaç ise geniş yapraklı olarak hatalı bir şekilde elde edilmiştir. Buradaki geniş ve iğne yapraklı ağaçların karışım oranı ise Test alanı 2'de 8 özelliğin kullanıldığı DVM sınıflandırması ile birbirine yakındır. Test alanı 2'nin 25 özellik kullanılarak gerçekleştirilen DVM sınıflandırma sonuçları aşağıdaki şekildedir:

- Toplam doğruluk (overall accuracy) =  $(109+6)/(109+6+36+8) = 0.7233 = \%72$
- Hassasiyet (recall) (geniş yapraklı ağaçlar) =  $109/(109+8) = 0.9316$
- Hassasiyet (recall) (iğne yapraklı ağaçlar) =  $6/(6+36) = 0.1429$
- Kesinlik (precision) (geniş yapraklı ağaçlar) =  $109/(109+36) = 0.7517$
- Kesinlik (precision) (iğne yapraklı ağaçlar) =  $6/(6+8) = 0.4286$
- F1-skor (geniş yapraklı ağaçlar) =  $2x(0.9316x0.7517)/(0.9316+0.7517) = 0.8320$
- F1-skor (iğne yapraklı ağaçlar) =  $2x(0.1429x0.4286)/(0.1429+0.4286) = 0.2143$

Test alanı 2'nin 25 özellik kullanılarak gerçekleştirilen DVM sınıflandırması sonucunda genel doğruluk %72 olarak belirlenmiştir. Geniş yapraklı kentsel ağaçlar için 25 özellik kullanılarak gerçekleştirilen DVM sınıflandırması sonucu elde edilen hassasiyet, kesinlik ve F1-skor değerleri sırasıyla 0.9316, 0.7517 ve 0.8320'dir. İğne yapraklı ağaçlar için ise hassasiyet, kesinlik ve F1-skor değerleri sırasıyla 0.1429, 0.4286 ve 0.2143'tür. Test alanı 2'nin 25 özellik kullanılarak gerçekleştirilen DVM sınıflandırma sonuçları aynı alanın 8 özellik kullanılarak gerçekleştirilen sınıflandırma sonuçlarına oldukça yakındır. DVM sınıflandırma yaklaşımı ile 25 özellik kullanılarak geniş yapraklı ağaçlar oldukça başarılı sınıflandırılmış fakat iğne yapraklı ağaçlarda aynı başarı elde edilememiştir. Daha önce 8 özelliğin sınıflandırma sonuçlarında da ifade edildiği gibi test alanında iğne yapraklı ağaç sayısı ve bu ağaçların dağılımının yetersiz olmasının başarı oranının düşük olmasında etkindir. 25 özelliğin kullanıldığı DVM sınıflandırmasının 10 katlamalı çapraz doğrulamasına (Tablo 4.19) göre ortalama doğruluk ise Test alanı 2'de %79 olarak elde edilmiştir.

**Tablo 4.19** Test alanı 2'nin 25 özellik kullanılarak DVM sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları

10 Katlamalı Çapraz Doğrulama		
1. katlama doğruluğu	0.7736	Ortalama doğruluk: 0.7858 (%79)
2. katlama doğruluğu	0.7925	
3. katlama doğruluğu	0.6604	
4. katlama doğruluğu	0.7736	
5. katlama doğruluğu	0.7736	
6. katlama doğruluğu	0.8113	
7. katlama doğruluğu	0.7736	
8. katlama doğruluğu	0.8269	
9. katlama doğruluğu	0.8462	
10. katlama doğruluğu	0.8269	

#### 4.7.2 Rastgele Orman Sınıflandırması

Test alanı 1 ve Test alanı 2'nin rastgele orman (RO) algoritması kullanılarak kentsel ağaç türlerinden geniş ve iğne yapraklı ağaç olarak 3 boyutlu nokta tabanlı sınıflandırılması gerçekleştirilmiştir. İlk olarak makine öğrenmesi tabanlı RO algoritması için parametre analizi, ardından en iyi sınıflandırma sonucunu veren parametreler kullanılarak 8 sınıflandırma özelliği ve 25 sınıflandırma özelliği kullanılarak Test alanı 1 ve Test alanı 2'de sınıflandırma işlemi gerçekleştirilmiştir. RO yönteminin performansını değerlendirmek amacıyla her iki test alanında, her iki özellik setinin kullanıldığı sınıflandırmalar için ayrı ayrı 10 katlamalı çapraz doğrulama da gerçekleştirilmiştir. Parametre analizi, RO sınıflandırması, sınıflandırmanın doğruluğu ve çapraz doğrulama sonuçları bu bölümde ayrıntılı bir biçimde verilecektir.

##### 4.7.2.1 RO Sınıflandırması İçin Parametre Analizi

Bu çalışmada kullanılacak RO algoritması için de “grid search” yöntemi ile en uygun sınıflandırma parametrelerinin otomatik olarak belirlenmesi hedeflenmiştir. RO algoritması için en önemli parametreler olan ormandaki ağaç sayısı ( $n\_estimators$ ) ve en iyi bölünmeyi ararken göz önünde bulundurulması gereken özelliklerin sayısı ( $max\_features$ ) için belirlenen özellik uzayı (Tablo 4.20) kullanılarak en iyi sonuç veren sınıflandırma parametreleri analiz edilmiştir. Buradaki parametre analizleri de test alanı 1'de 25 sınıflandırma özelliği ile

gerçekleştirilmiştir. “Grid search” yöntemi ile RO’da en uygun parametreler olarak; ağaç sayısı 150 ve en iyi bölünmeyi ararken göz önünde bulundurulması gereken özelliklerin sayısı 10 olarak belirlenmiştir.

**Tablo 4.20** RO için “Grid search” yaklaşımında kullanılan parametreler

Parametre uzayı	
Ağaç sayısı	1, 5, 10, 20, 50, 75, 100, 150, 200, 250, 500, 1000
En iyi bölünme için maksimum özellik sayısı	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25

#### 4.7.2.2 Test Alanı 1’in Belirlenen Parametreler ile RO Sınıflandırması ve Doğruluk Analizi

Detaylı parametre analizlerinin ardından Test alanı 1, en iyi sonucu veren parametreler kullanılarak geniş ve iğne yapraklı ağaç türlerini elde etmek amacıyla makine öğrenme tabanlı RO algoritması ile sınıflandırılmıştır. Nokta tabanlı sınıflandırmada 8 özelliğin kullanıldığı durumda RO’da toplamda 150 ağaç ( $n\_estimators$ ) ve her bölmede maksimum 8 özellik ( $max\_features$ ) kullanılan parametreler olmuştur. RO için gerçekleştirilen parametre analizinde en iyi bölünmeyi ararken göz önünde bulundurulması gereken özelliklerin sayısı ( $max\_features$ ) 10 olarak belirlenmesine rağmen buradaki sınıflandırma özellik seti 8 elemanlı olduğu için her bölmede maksimum 8 özellik kullanılmıştır. 8 özelliğin kullanıldığı RO sınıflandırmasında farklı bölümlenmeye sahip 2 karar ağacı örneği Şekil 4.47’de görülmektedir. Test alanı 1’de eğitim işleminin ardından 8 özellik kullanılarak gerçekleştirilen RO sınıflandırma sonuçları Şekil 4.48 ve bu sınıflandırmaya ilişkin hata (confusion) matrisi Tablo 4.21’de verilmiştir.





- Geniş yapraklı ağaçlar
- İğne yapraklı ağaçlar

**Şekil 4.48** Test alanı 1'in 8 özellik kullanılarak RO sınıflandırması

**Tablo 4.21** Test alanı 1'in 8 özellik kullanılarak RO sınıflandırmasına ilişkin hata matrisi

		Referans veri		
		Geniş	İğne	Toplam
Tahmin verisi	Geniş	49	12	61
	İğne	4	15	19
	Toplam	53	27	80

Test alanı 1'de gerçekleştirilen eğitim işleminin ardından 8 özellik kullanılarak gerçekleştirilen RO sınıflandırmasına göre 49 geniş yapraklı ağaç ve 15 iğne yapraklı ağaç doğru bir şekilde sınıflandırılmıştır. Yalnızca 4 adet geniş yapraklı ağaç iğne olarak sınıflandırılırken bu sayı iğne yapraklı ağaçlarda daha fazladır (12 iğne yapraklı ağaç geniş yapraklı olarak sınıflandırılmış). Buna göre Test alanı 1'de 8 özellikli RO sınıflandırması ile iğne yapraklı ağaçların geniş yapraklı olarak sınıflandırılması geniş yapraklı ağaçların iğne yapraklı ağaç olarak sınıflandırılmasından daha olasıdır. Test alanı 1'in 8 özellik kullanılarak gerçekleştirilen RO sınıflandırma sonuçları aşağıdaki gibidir:

- Toplam doğruluk (overall accuracy) =  $(49+15)/(49+15+12+4) = 0.80$   
= %80
- Hassasiyet (recall) (geniş yapraklı ağaçlar) =  $49/(49+4) = 0.9245$
- Hassasiyet (recall) (iğne yapraklı ağaçlar) =  $15/(15+12) = 0.5556$

- Kesinlik (precision) (geniş yapraklı ağaçlar) =  $49/(49+12) = 0.8033$
- Kesinlik (precision) (iğne yapraklı ağaçlar) =  $15/(15+4) = 0.7895$
- F1-skor (geniş yapraklı ağaçlar) =  $2x(0.9245x0.8033)/(0.9245+0.8033)$   
= 0.8596
- F1-skor (iğne yapraklı ağaçlar) =  $2x(0.5556x0.7895)/(0.5556+0.7895)$   
= 0.6522

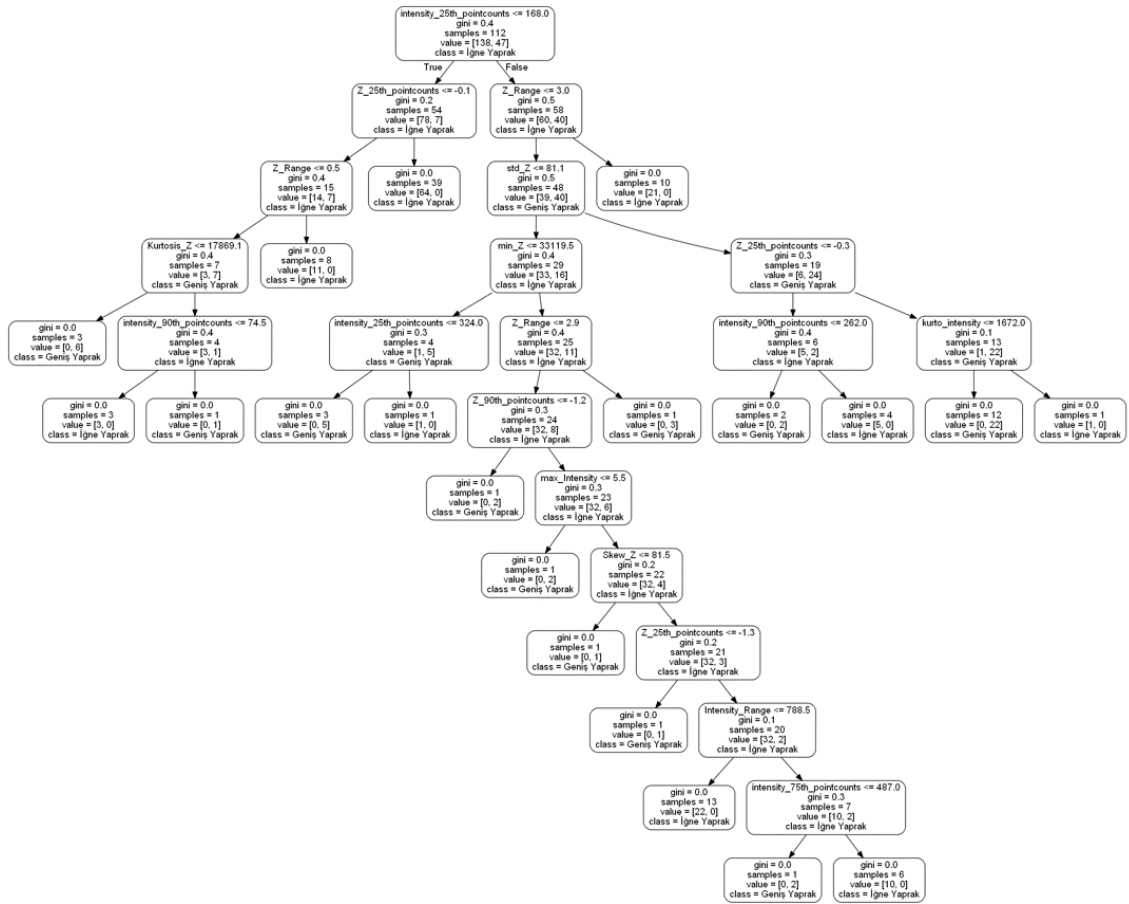
RO algoritması ile 8 özellik kullanılarak Test alanı 1’de geniş ve iğne yapraklı ağaçların elde edilmesi için gerçekleştirilen sınıflandırmanın genel doğruluğu %80’dir. Geniş yapraklı ağaçlar için hassasiyet değeri 0.9245, kesinlik değeri 0.8033 ve F1-skor değeri 0.8596 olarak hesaplanmıştır. İğne yapraklı ağaçlar için hassasiyet değeri 0.5556, kesinlik değeri 0.7895 ve F1-skor değeri 0.6522 olarak belirlenmiştir. Sınıflandırma sonuçlarına göre 8 özelliğin kullanıldığı RO algoritmasının ağaç türlerini sınıflandırmada başarılı olduğu ifade edilebilir. Hassasiyet, kesinlik ve F1-skor değerleri incelendiğinde RO sınıflandırıcısının Test alanı 1’de 8 özellik kullanılarak geniş yapraklı ağaçları iğne yapraklı ağaçlardan daha iyi sınıflandırdığı görülmektedir. 8 özelliğin kullanıldığı RO sınıflandırmasının Test alanı 1’de 10 katlamalı çapraz doğrulamaya (Tablo 4.22) göre %78 olan ortalama doğruluğu sınıflandırmanın tutarlılığını ortaya koymaktadır.

**Tablo 4.22** Test alanı 1’in 8 özellik kullanılarak RO sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları

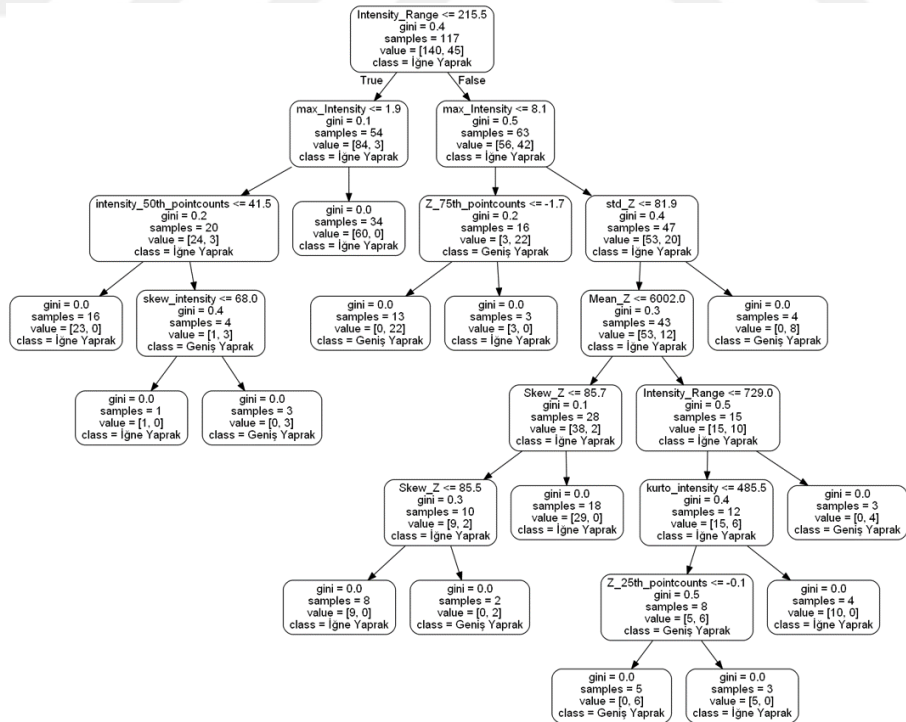
10 Katlamalı Çapraz Doğrulama		
1. katlama doğruluğu	0.7778	Ortalama doğruluk: 0.7812 (%78)
2. katlama doğruluğu	0.6296	
3. katlama doğruluğu	0.8889	
4. katlama doğruluğu	0.9259	
5. katlama doğruluğu	0.6667	
6. katlama doğruluğu	0.8462	
7. katlama doğruluğu	0.6923	
8. katlama doğruluğu	0.7308	
9. katlama doğruluğu	0.8077	
10. katlama doğruluğu	0.8462	

Test alanı 1’de 25 özellik kullanılarak gerçekleştirilen RO sınıflandırmasında toplamda 150 ağaç (n\_estimators) ve her bölmede maksimum 8 özellik (max\_features) yerine 10 özellik en iyi sonucu veren parametrelerdir. Test alanı 1’de 25 özelliğin kullanıldığı RO sınıflandırmasında farklı bölümlenmeye sahip 2 karar ağacı örneği Şekil 4.49’da yer almaktadır. Test alanı 1’in 25 özellik ile gerçekleştirilen RO sınıflandırma sonuçları Şekil 4.50 ve sınıflandırmanın hata matrisi Tablo 4.23’de verilmiştir.





(a)



(b)

Şekil 4.49 Test alanı 1'de 25 özelliğin kullanıldığı RO'da 83. karar ağacı (a) ve 103. karar ağacı (b)



- Geniş yapraklı ağaçlar
- İğne yapraklı ağaçlar

Şekil 4.50 Test alanı 1'in 25 özellik kullanılarak RO sınıflandırması

Tablo 4.23 Test alanı 1'in 25 özellik kullanılarak RO sınıflandırmasına ilişkin hata matrisi

		Referans veri		
		Geniş	İğne	Toplam
Tahmin verisi	Geniş	50	10	60
	İğne	3	17	20
	Toplam	53	27	80

Test alanı 1'de 25 özellik kullanılarak gerçekleştirilen RO sınıflandırması ile doğru olarak sınıflandırılan geniş yapraklı ağaç sayısı oldukça yüksek iken (50 doğru sınıflandırılan geniş yapraklı ağaç), doğru olarak sınıflandırılan iğne yapraklı ağaçlar için aynı durum söz konusu değildir. 27 ağaçtan 17 iğne yapraklı ağaç doğru olarak sınıflandırılmıştır. Buna göre 10 iğne yapraklı ağaç geniş yapraklı olarak sınıflandırılırken yalnızca 3 geniş yapraklı ağaç iğne yapraklı olarak sınıflandırılmıştır. Test alanı 1'in 25 özellik kullanılarak gerçekleştirilen RO sınıflandırma sonuçları aşağıdaki görülmektedir:

- Toplam doğruluk (overall accuracy) =  $(50+17)/(50+17+10+3) = 0.8375 = \%84$
- Hassasiyet (recall) (geniş yapraklı ağaçlar) =  $50/(50+3) = 0.9434$
- Hassasiyet (recall) (iğne yapraklı ağaçlar) =  $17/(17+10) = 0.6296$
- Kesinlik (precision) (geniş yapraklı ağaçlar) =  $50/(50+10) = 0.8333$

- Kesinlik (precision) (iğne yapraklı ağaçlar) =  $17/(17+3) = 0.85$
- F1-skor (geniş yapraklı ağaçlar) =  $2x(0.9434x0.8333)/(0.9434+0.8333)$   
= 0.8849
- F1-skor (iğne yapraklı ağaçlar) =  $2x(0.6296x0.85)/(0.6296+0.85) =$   
0.7234

25 sınıflandırma özelliği kullanılarak Test alanı 1’de gerçekleştirilen RO sınıflandırmasının toplam doğruluğu %84 olarak elde edilmiştir. Hassasiyet, kesinlik ve F1-skor değerleri geniş yapraklı ağaçlar için sırasıyla 0.9434, 0.8333 ve 0.8849 iken iğne yapraklı ağaçlar için sırasıyla 0.6296, 0.85 ve 0.7234’dır. Test alanı 1’de 25 özelliğin kullanıldığı RO sınıflandırması 8 özelliğin kullanıldığı RO sınıflandırmasından daha başarılıdır. Geniş ve iğne yapraklı ağaç sınıflandırmaları da 8 özelliğin kullanıldığı RO sınıflandırmasından nispeten daha başarılı sonuçlar üretmiştir. Test alanı 1’de 25 özelliğin kullanıldığı sınıflandırmanın 10 katlamalı çapraz doğrulamasının (Tablo 4.24) ortalama doğruluğu da %82’dir.

**Tablo 4.24** Test alanı 1’in 25 özellik kullanılarak RO sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları

10 Katlamalı Çapraz Doğrulama		
1. katlama doğruluğu	0.8148	Ortalama doğruluk: 0.8154 (%82)
2. katlama doğruluğu	0.6296	
3. katlama doğruluğu	0.8889	
4. katlama doğruluğu	0.9630	
5. katlama doğruluğu	0.7037	
6. katlama doğruluğu	0.8846	
7. katlama doğruluğu	0.8077	
8. katlama doğruluğu	0.8077	
9. katlama doğruluğu	0.8077	
10. katlama doğruluğu	0.8462	

#### 4.7.2.3 Test Alanı 2’nin Belirlenen Parametreler ile RO Sınıflandırması ve Doğruluk Analizi

Test alanı 2, parametre analizinin ardından geniş ve iğne yapraklı kentsel ağaçların belirlenmesi için RO algoritması ile 3B nokta tabanlı olarak sınıflandırılmıştır. 8 özelliğin kullanıldığı bu ilk sınıflandırmada Test alanı 1’deki duruma benzer şekilde RO’da toplamda 150 ağaç (n\_estimators) ve her bölmede





● Geniş yapraklı ağaçlar ● İğne yapraklı ağaçlar

Şekil 4.52 Test alanı 2'nin 8 özellik kullanılarak RO sınıflandırması

Tablo 4.25 Test alanı 2'nin 8 özellik kullanılarak RO sınıflandırmasına ilişkin hata matrisi

		Referans veri		
		Geniş	İğne	Toplam
Tahmin verisi	Geniş	100	31	131
	İğne	17	11	28
	Toplam	117	42	159

Test alanı 2'de 8 sınıflandırma özelliği kullanılarak gerçekleştirilen RO sınıflandırması ile 100 geniş yapraklı ağaç ve 11 iğne yapraklı ağaç doğru sınıflandırılmıştır. 17 geniş yapraklı ağaç hatalı olarak iğne yapraklı ağaç şeklinde sınıflandırılırken bu oran iğne yapraklı ağaç da oldukça yüksektir. 42 ağaçtan 31'i yanlış bir şekilde geniş yapraklı olarak elde edilmiştir. Burada da iğne yapraklı ağaçların geniş yapraklı ağaçlarla karışıp, geniş olarak sınıflandırılması oldukça olasıdır. Test alanı 2'nin 8 özellik kullanılarak gerçekleştirilen RO sınıflandırma sonuçları aşağıdaki şekildedir:

- Toplam doğruluk (overall accuracy) =  $(100+11)/(100+11+31+17) = 0.6981 = \%70$
- Hassasiyet (recall) (geniş yapraklı ağaçlar) =  $100/(100+17) = 0.8547$
- Hassasiyet (recall) (iğne yapraklı ağaçlar) =  $11/(11+31) = 0.2619$
- Kesinlik (precision) (geniş yapraklı ağaçlar) =  $100/(100+31) = 0.7634$
- Kesinlik (precision) (iğne yapraklı ağaçlar) =  $11/(11+17) = 0.3929$
- F1-skor (geniş yapraklı ağaçlar) =  $2x(0.8547x0.7634)/(0.8547+0.7634) = 0.8065$
- F1-skor (iğne yapraklı ağaçlar) =  $2x(0.2619x0.3929)/(0.2619+0.3929) = 0.3143$

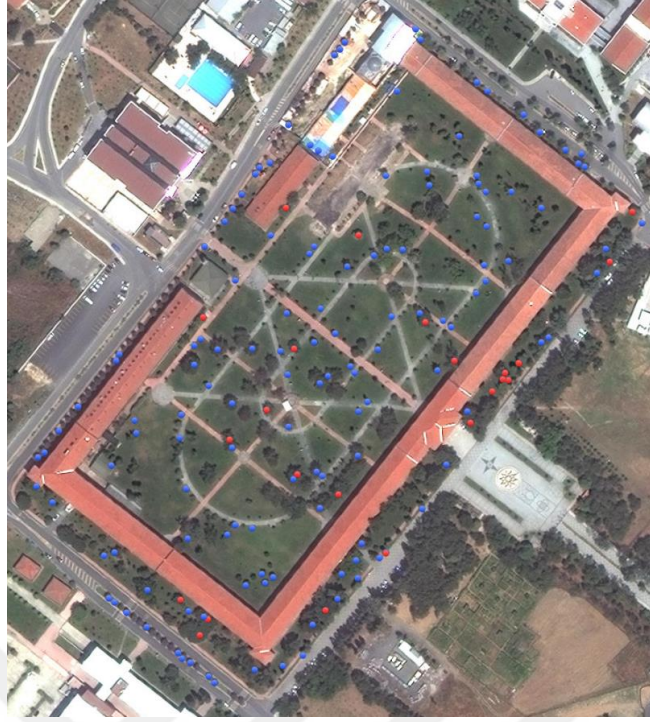
Test alanı 2’de kentsel ağaç türlerinin elde edilmesi için 8 özellik kullanılarak gerçekleştirilen RO sınıflandırmasının toplam doğruluğu %70 olarak belirlenmiştir. RO sınıflandırması sonucu Test alanı 2’de yer alan geniş yapraklı ağaçlar için hassasiyet, kesinlik ve F1-skor değerleri sırasıyla 0.8547, 0.7634 ve 0.8065; iğne yapraklı ağaçlar için ise sırasıyla 0.2619, 0.3929 ve 0.3143’tür. Test alanı 2’de 8 özellik kullanılarak gerçekleştirilen RO algoritması ile geniş ve iğne yapraklı kent ağaçları Test alanı 1’e göre daha düşük doğrulukla belirlenmiştir. Hassasiyet, kesinlik ve F1-skor değerleri incelendiğinde geniş yapraklı ağaçlar Test alanı 2’de RO ile başarılı bir şekilde sınıflandırılrsa da iğne yapraklı ağaçların sınıflandırılma doğruluğu oldukça düşüktür. DVM sınıflandırıcısına benzer şekilde düşük doğruluklu iğne yapraklı ağaç sınıflandırmasının nedeni test alanında yer alan iğne yapraklı ağaçların az sayıda olması ve dağılımıdır. 8 özellik kullanılarak gerçekleştirilen RO sınıflandırmasının 10 katlamalı çapraz doğrulamaya (Tablo 4.26) göre ortalama doğruluğu %76’dır.

**Tablo 4.26** Test alanı 2'nin 8 özellik kullanılarak RO sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları

<b>10 Katlamalı Çapraz Doğrulama</b>		
1. katlama doğruluğu	0.7358	<b>Ortalama doğruluk:</b> 0.7554 (%76)
2. katlama doğruluğu	0.8113	
3. katlama doğruluğu	0.6604	
4. katlama doğruluğu	0.7736	
5. katlama doğruluğu	0.6415	
6. katlama doğruluğu	0.8113	
7. katlama doğruluğu	0.7358	
8. katlama doğruluğu	0.7885	
9. katlama doğruluğu	0.7115	
10. katlama doğruluğu	0.8846	

Test alanı 2'de 25 özelliğin kullanıldığı RO sınıflandırmasında toplamda 150 ağaç (n\_estimators) ve her bölmede maksimum 10 özellik (max\_features) kullanılmıştır. Test alanı 2'de gerçekleştirilen 25 özelliğin kullanıldığı RO sınıflandırmasında oluşan bir karar ağacı örneği Şekil 4.53'de görülmektedir. Şekil 4.54'de Test alanı 2'de 25 özellik ile gerçekleştirilen RO sınıflandırma sonuçları yer alırken sınıflandırmanın hata matrisi Tablo 4.27'de verilmiştir.





● Geniş yapraklı ağaçlar ● İğne yapraklı ağaçlar

**Şekil 4.54** Test alanı 2'nin 25 özellik kullanılarak RO sınıflandırması

**Tablo 4.27** Test alanı 2'nin 25 özellik kullanılarak RO sınıflandırmasına ilişkin hata matrisi

		Referans veri		
		Geniş	İğne	Toplam
Tahmin verisi	Geniş	104	31	135
	İğne	13	11	24
	Toplam	117	42	159

Test alanı 2'de 25 özellik kullanılarak gerçekleştirilen RO sınıflandırmasına göre doğru sınıflandırılan geniş yapraklı ağaçlar 8 özelliğin kullanıldığı duruma göre artış gösterirken doğru sınıflandırılan iğne yapraklı ağaçların sayısı değişmemiştir (104 doğru sınıflandırılan geniş yapraklı ağaç, 11 doğru sınıflandırılan iğne yapraklı ağaç). Burada 25 özelliğin kullanıldığı durumda 13 geniş yapraklı ağaç hatalı olarak iğne yapraklı şeklinde sınıflandırılmıştır; yine 31 iğne yapraklı ağaç hatalı olarak geniş yapraklı şeklinde sınıflandırılmıştır. Buna göre ağaçların birbirleri ile karışma oranları 8 özelliğin kullanıldığı durum ile çok benzerdir. Test alanı 2'nin 25 özellik kullanılarak gerçekleştirilen RO sınıflandırma sonuçları aşağıda verilmiştir:

- Toplam doğruluk (overall accuracy) =  $(104+11)/(104+11+31+13) = 0.7233 = \%72$
- Hassasiyet (recall) (geniş yapraklı ağaçlar) =  $104/(104+13) = 0.8889$
- Hassasiyet (recall) (iğne yapraklı ağaçlar) =  $11/(11+31) = 0.2619$
- Kesinlik (precision) (geniş yapraklı ağaçlar) =  $104/(104+31) = 0.7704$
- Kesinlik (precision) (iğne yapraklı ağaçlar) =  $11/(11+13) = 0.4583$
- F1-skor (geniş yapraklı ağaçlar) =  $2x(0.8889x0.7704)/(0.8889+0.7704)=0.8254$
- F1-skor (iğne yapraklı ağaçlar) =  $2x(0.2619x0.4583)/(0.2619+0.4583) = 0.3333$

Test alanı 2'nin 25 özellik kullanılarak gerçekleştirilen RO sınıflandırması sonucunda genel doğruluk %72 olarak elde edilmiştir. Geniş yapraklı ağaçlar için 25 özellik kullanılarak gerçekleştirilen RO sınıflandırması sonucu elde edilen hassasiyet, kesinlik ve F1-skor değerleri sırasıyla 0.8889, 0.7704 ve 0.8254'dir. İğne yapraklı ağaçlar için ise hassasiyet, kesinlik ve F1-skor değerleri sırasıyla 0.2619, 0.4583 ve 0.3333'tür. Test alanı 2'nin 25 özellik kullanılarak gerçekleştirilen sınıflandırma sonucu aynı test alanının 8 özellik kullanılarak gerçekleştirilen sınıflandırma sonucu ile aynıdır. 8 özelliğin kullanıldığı RO sınıflandırmasında olduğu gibi buradaki 25 özelliğin kullanıldığı RO sınıflandırmasında da geniş yapraklı ağaçların aksine iğne yapraklı ağaçların sınıflandırma doğruluğu aynı nedenlerden dolayı oldukça düşüktür. 25 özelliğin kullanıldığı RO sınıflandırmasının 10 katlamalı çapraz doğrulamasına (Tablo 4.28) göre ortalama doğruluğu ise Test alanı 2'de %80 olarak elde edilmiştir.

**Tablo 4.28** Test alanı 2'nin 25 özellik kullanılarak RO sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları

10 Katlamalı Çapraz Doğrulama		
1. katlama doğruluğu	0.7358	Ortalama doğruluk: 0.8030 (%80)
2. katlama doğruluğu	0.8113	
3. katlama doğruluğu	0.6981	
4. katlama doğruluğu	0.8302	
5. katlama doğruluğu	0.7925	
6. katlama doğruluğu	0.8113	
7. katlama doğruluğu	0.7547	
8. katlama doğruluğu	0.8462	
9. katlama doğruluğu	0.8269	
10. katlama doğruluğu	0.9231	

#### 4.7.3 Çok Katmanlı Algılayıcı Sınıflandırması

DVM ve RO sınıflandırmalarının ardından Test alanı 1 ve Test alanı 2'de çok katmanlı algılayıcı (ÇKA) algoritması kullanılarak geniş ve iğne yapraklı kent ağaçlarının 3 boyutlu nokta tabanlı sınıflandırılması gerçekleştirilmiştir. ÇKA algoritması için parametre analizi ve ardından en uygun parametreler ile her bir test alanı için sırasıyla 8 ve 25 özellik kullanılarak nokta tabanlı sınıflandırma uygulaması gerçekleştirilmiştir. ÇKA sınıflandırma yönteminin performansını değerlendirmek amacıyla Test alanı 1 ve Test alanı 2'de, 8 ve 25 özellik setinin kullanıldığı sınıflandırmalar için ayrı ayrı 10 katlamalı çapraz doğrulama da gerçekleştirilmiştir. Gerçekleştirilen parametre analizi, ÇKA sınıflandırması, sınıflandırmanın doğruluğu ve çapraz doğrulama sonuçları bu bölümde açıklanmıştır.

##### 4.7.3.1 ÇKA Sınıflandırması İçin Parametre Analizi

ÇKA sınıflandırma algoritması için de "grid search" yöntemi ile en uygun sınıflandırma parametrelerinin otomatik olarak belirlenmesi hedeflenmiştir. ÇKA algoritması için en önemli parametreler olan gizli katmandaki nöronların sayısı (hidden\_layer\_sizes) ve maksimum iterasyonların sayısı (max\_iter) için Tablo 4.29'da verilen parametre uzayı kullanılarak en iyi sonuç veren sınıflandırma parametreleri analiz edilmiştir. Buradaki parametre analizleri de Test alanı 1'de 25 sınıflandırma özelliği ile gerçekleştirilmiştir. Gerçekleştirilen parametre analizi

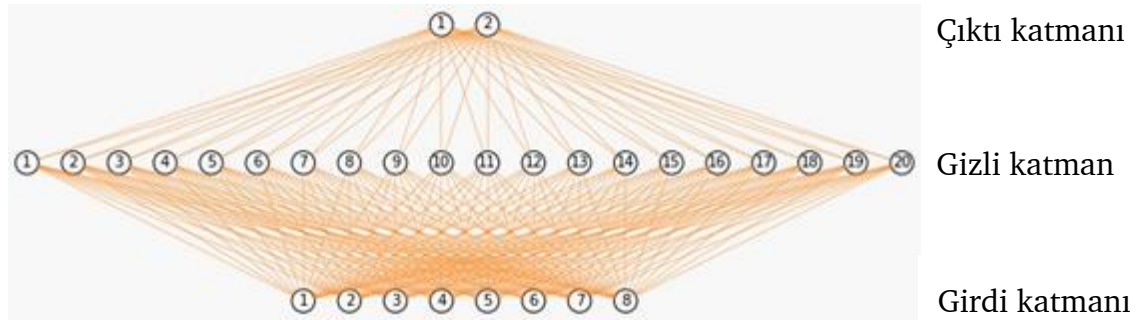
sonrasında gizli katmandaki nöronların sayısı 20 ve maksimum iterasyonların sayısı 2000 olarak belirlenmiştir.

**Tablo 4.29** ÇKA için “Grid search” yaklaşımında kullanılan parametreler

Parametre uzayı	
Gizli katmandaki nöronların sayısı	1, 5, 10, 20, 50, 75, 100, 150, 200, 250, 500, 1000
Maksimum iterasyonların sayısı	(10), (20), (30), (50), (10.), (100.), (100, 100.) (50), (30, 50.), (50, 50), (250.), (500), (2000)

#### 4.7.3.2 Test Alanı 1’in Belirlenen Parametreler ile ÇKA Sınıflandırması ve Doğruluk Analizi

Test alanı 1, geniş ve iğne yapraklı kent ağaçlarının elde edilmesi amacıyla detaylı parametre analizlerinin ardından belirlenen parametreler kullanılarak makine öğrenme tabanlı ÇKA algoritması ile sınıflandırılmıştır. 8 özelliğin kullanıldığı ÇKA sınıflandırmasında girdi katmanı 8 nörondan, gizli katman 20 nörondan ve çıktı katmanı geniş ve iğne yapraklı ağaçlar olmak üzere 2 nörondan oluşmaktadır (Şekil 4.55). ÇKA’da maksimum iterasyonların sayısı ise 2000 olarak kullanılmıştır. Test alanı 1’de eğitim işleminin ardından 8 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırma sonuçları Şekil 4.56’da ve ÇKA sınıflandırmasına ilişkin hata (confusion) matrisi de Tablo 4.30’da verilmiştir.



**Şekil 4.55** ÇKA ağ yapısı (8 sınıflandırma özelliği kullanılan)



- Geniş yapraklı ağaçlar
- İğne yapraklı ağaçlar

**Şekil 4.56** Test alanı 1'in 8 özellik kullanılarak ÇKA sınıflandırması

**Tablo 4.30** Test alanı 1'in 8 özellik kullanılarak ÇKA sınıflandırmasına ilişkin hata matrisi

		Referans veri		
		Geniş	İğne	Toplam
Tahmin verisi	Geniş	52	25	77
	İğne	1	2	3
	Toplam	53	27	80

Test alanı 1'de eğitimin ardından 8 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırmasına göre 53 geniş yapraklı ağaçtan 52'si doğru olarak sınıflandırılırken, 27 iğne yapraklı ağaçtan sadece 2'si doğru olarak sınıflandırılmıştır. Yani sadece 1 geniş yapraklı ağaç iğne yapraklı ağaç ile karışırken tam 25 iğne yapraklı ağaç geniş yapraklı ağaçlarla karışmıştır. Buna göre ÇKA sınıflandırması ile Test alanı 1'de 8 özelliğin kullanıldığı durumda iğne yapraklı ağaçların geniş yapraklı olarak sınıflandırılması oldukça olasıdır. Test alanı 1'in 8 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırma sonuçları aşağıdaki gibidir:

- Toplam doğruluk (overall accuracy) =  $(52+2)/(52+2+25+1) = 0.675$   
= %68
- Hassasiyet (recall) (geniş yapraklı ağaçlar) =  $52/(52+1) = 0.9811$
- Hassasiyet (recall) (iğne yapraklı ağaçlar) =  $2/(2+25) = 0.0741$

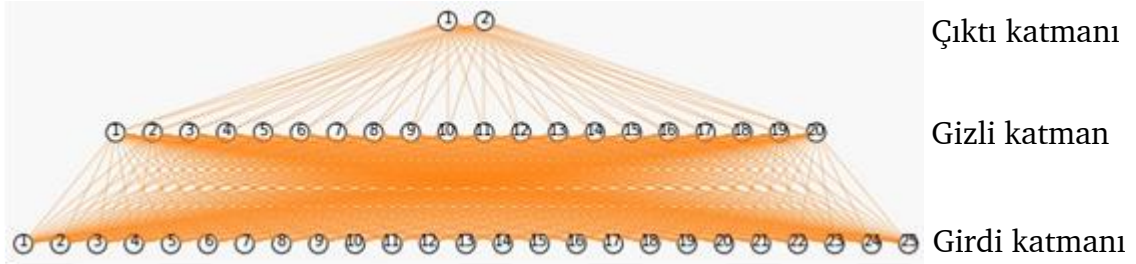
- Kesinlik (precision) (geniş yapraklı ağaçlar) =  $52/(52+25) = 0.6753$
- Kesinlik (precision) (iğne yapraklı ağaçlar) =  $2/(2+1) = 0.6667$
- F1-skor (geniş yapraklı ağaçlar) =  $2x(0.9811x0.6753)/(0.9811+0.6753)$   
= 0.80
- F1-skor (iğne yapraklı ağaçlar) =  $2x(0.0741x0.6667)/(0.0741+0.6667)$   
= 0.1334

Test alanı 1’de 8 özellik kullanılarak geniş ve iğne yapraklı kent ağaçlarının belirlenmesi amacıyla gerçekleştirilen ÇKA sınıflandırmasının toplam doğruluğu %68’dir. Geniş yapraklı ağaçlar için hassasiyet değeri 0.9811, kesinlik değeri 0.6753 ve F1-skor değeri 0.80 olarak hesaplanırken iğne yapraklı ağaçlar için bu değerler sırasıyla 0.0741, 0.6667, 0.1334 olarak hesaplanmıştır. Sınıflandırma sonuçlarına göre 8 özelliğin kullanıldığı ÇKA algoritmasının ağaç türlerini sınıflandırmada diğer yöntemlerin aksine yeterince başarılı olamadığı söylenebilir. Hassasiyet, kesinlik ve F1-skor değerleri incelendiğinde ise ÇKA sınıflandırıcısının diğer yöntemlere benzer şekilde Test alanı 1’de 8 özellik kullanılarak geniş yapraklı ağaçları iğne yapraklı ağaçlardan çok daha iyi sınıflandırdığı görülmektedir. Test alanı 1’de 8 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırmasının 10 katlamalı çapraz doğrulamaya (Tablo 4.31) göre ortalama doğruluğu %49 olarak oldukça düşük elde edilmiştir.

**Tablo 4.31** Test alanı 1’in 8 özellik kullanılarak ÇKA sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları

10 Katlamalı Çapraz Doğrulama		
1. katlama doğruluğu	0.2963	Ortalama doğruluk: 0.4895 (%49)
2. katlama doğruluğu	0.6296	
3. katlama doğruluğu	0.2962	
4. katlama doğruluğu	0.8889	
5. katlama doğruluğu	0.6296	
6. katlama doğruluğu	0.6923	
7. katlama doğruluğu	0.3846	
8. katlama doğruluğu	0.3462	
9. katlama doğruluğu	0.3846	
10. katlama doğruluğu	0.3462	

Test alanı 1'in 25 özelliğın kullanıldıđı ÇKA sınıflandırmasında girdi katmanı 25 nörondan, gizli katman 20 nörondan ve çıktı katmanı geniş ve iğne yapraklı ağaçlar olmak üzere 2 nörondan oluşmaktadır (Şekil 4.57) ve maksimum iterasyonların sayısı ise 2000 olarak kullanılmıştır. Test alanı 1'in 25 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırma sonuçları Şekil 4.58'de ve bu ÇKA sınıflandırmasına ilişkin hata matrisi Tablo 4.32'de verilmiştir.



Şekil 4.57 ÇKA ağ yapısı (25 sınıflandırma özelliđi kullanılan)



- Geniş yapraklı ağaçlar
- İğne yapraklı ağaçlar

Şekil 4.58 Test alanı 1'in 25 özellik kullanılarak ÇKA sınıflandırması

Tablo 4.32 Test alanı 1'in 25 özellik kullanılarak ÇKA sınıflandırmasına ilişkin hata matrisi

		Referans veri		
		Geniş	İğne	Toplam
Tahmin verisi	Geniş	49	17	66
	İğne	4	10	14
	Toplam	53	27	80

Test alanı 1’de 25 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırmasına göre 49 geniş yapraklı ağaç ve 10 iğne yapraklı ağaç doğru bir şekilde sınıflandırılmıştır. 4 iğne yapraklı ağaç hatalı olarak geniş yapraklı ağaç şeklinde sınıflandırılmıştır. Geniş yapraklı ağaçlara göre karışma oranı yine daha yüksek olsa da 8 özelliğin kullanıldığı duruma göre daha iyi bir şekilde 17 tane iğne yapraklı ağaç geniş yapraklı ağaçlar ile karışmıştır. Test alanı 1’in 25 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırma sonuçları aşağıda verilmiştir:

- Toplam doğruluk (overall accuracy) =  $(49+10)/(49+10+17+4) = 0.7375 = \%74$
- Hassasiyet (recall) (geniş yapraklı ağaçlar) =  $49/(49+4) = 0.9245$
- Hassasiyet (recall) (iğne yapraklı ağaçlar) =  $10/(10+17) = 0.3704$
- Kesinlik (precision) (geniş yapraklı ağaçlar) =  $49/(49+17) = 0.7424$
- Kesinlik (precision) (iğne yapraklı ağaçlar) =  $10/(10+4) = 0.7143$
- F1-skor (geniş yapraklı ağaçlar) =  $2x(0.9245x0.7424)/(0.9245+0.7424) = 0.8235$
- F1-skor (iğne yapraklı ağaçlar) =  $2x(0.3704x0.7143)/(0.3704+0.7143) = 0.4878$

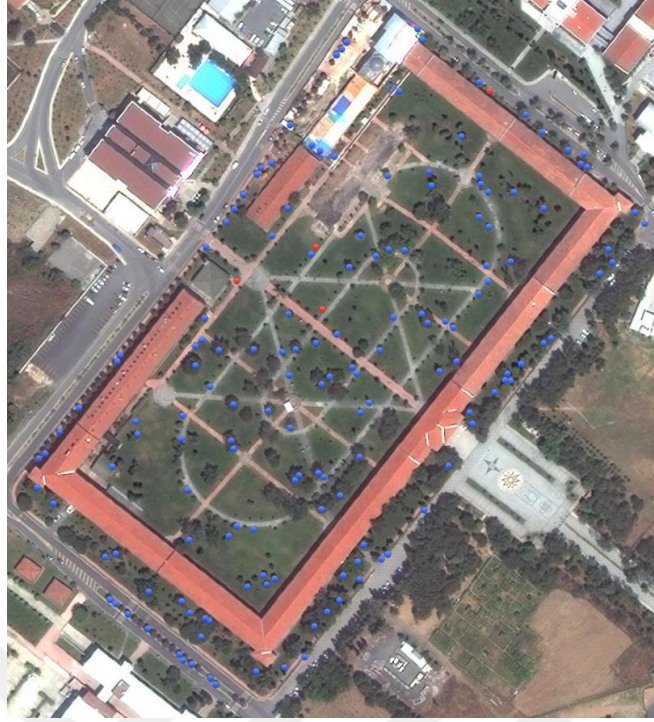
Test alanı 1’de 25 özellik kullanılarak geniş ve iğne yapraklı ağaç türlerinin elde edilmesi amacıyla gerçekleştirilen ÇKA sınıflandırmasının genel doğruluğu %74’tür. ÇKA sınıflandırma sonuçlarına göre geniş yapraklı ağaçlar için hassasiyet değeri 0.9245, kesinlik değeri 0.7424 ve F1-skor değeri 0.8235 olarak elde edilmiştir. İğne yapraklı ağaçlar için hassasiyet değeri 0.3704, kesinlik değeri 0.7143 ve F1-skor değeri 0.4878 olarak hesaplanmıştır. Test alanı 1’in 25 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırması, 8 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırmasından daha başarılı sonuçlar üretmiştir. Test alanı 1’de 25 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırmasında iğne yapraklı ağaçların sınıflandırılma başarısı 8 özellik ile gerçekleştirilen sınıflandırmadan daha iyi olsa da geniş yapraklı ağaçlardan düşüktür. Ayrıca, 25 özellik kullanılarak Test alanı 1’de gerçekleştirilen ÇKA sınıflandırmasının 10 katlamalı çapraz doğrulamaya (Tablo 4.33) göre ortalama doğruluğu %64 olarak elde edilmiştir.

**Tablo 4.33** Test alanı 1'in 25 özellik kullanılarak ÇKA sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları

<b>10 Katlamalı Çapraz Doğrulama</b>		
1. katlama doğruluğu	0.5185	<b>Ortalama doğruluk:</b> 0.6372 (%64)
2. katlama doğruluğu	0.6667	
3. katlama doğruluğu	0.6296	
4. katlama doğruluğu	0.8889	
5. katlama doğruluğu	0.6296	
6. katlama doğruluğu	0.7308	
7. katlama doğruluğu	0.6923	
8. katlama doğruluğu	0.3846	
9. katlama doğruluğu	0.6154	
10. katlama doğruluğu	0.6154	

#### 4.7.3.3 Test Alanı 2'nin Belirlenen Parametreler ile ÇKA Sınıflandırması ve Doğruluk Analizi

Test alanı 2'ye ait nokta bulutu verileri de parametre analizi ile belirlenen parametreler kullanılarak kentsel ağaç türlerinin belirlenmesi için 8 özelliğin kullanıldığı ÇKA sınıflandırıcısı ile 3B nokta tabanlı olarak sınıflandırılmıştır. Test alanı 1'de 8 özelliğin kullanıldığı sınıflandırmada olduğu gibi buradaki ÇKA'da 8 nöronlu bir girdi katmanı, 20 nöronlu bir gizli katman ve 2 nöronlu bir çıktı katmanı bulunmaktadır. Test alanı 2'de kullanılan ÇKA'da da maksimum iterasyonların sayısı 2000 olarak kullanılmıştır. Test alanı 2'de geniş ve iğne yapraklı ağaçların yer aldığı 8 özellikli ÇKA sınıflandırma sonucu Şekil 4.59'da verilmiştir. Tablo 4.34'de ise ÇKA sınıflandırma sonuçlarına ilişkin hata matrisi yer almaktadır.



● Geniş yapraklı ağaçlar ● İğne yapraklı ağaçlar

Şekil 4.59 Test alanı 2'nin 8 özellik kullanılarak ÇKA sınıflandırması

Tablo 4.34 Test alanı 2'nin 8 özellik kullanılarak ÇKA sınıflandırmasına ilişkin hata matrisi

		Referans veri		
		Geniş	İğne	Toplam
Tahmin verisi	Geniş	114	41	155
	İğne	3	1	4
	Toplam	117	42	159

Test alanı 2'de ÇKA algoritması ile gerçekleştirilen 8 özellikli sınıflandırmaya göre test alanında 114 geniş yapraklı ağaç doğru olarak sınıflandırılırken, yalnızca 1 iğne yapraklı ağaç doğru olarak sınıflandırılmıştır. Çok iyi bir oranla yalnız 3 geniş yapraklı ağaç iğne yapraklı olarak sınıflandırılmıştır. İğne yapraklı ağaçların ise 41 tanesi geniş yapraklı olarak sınıflandırılmıştır. Buradaki geniş ve iğne yapraklı ağaçların karışımları Test alanı 1'deki 8 özelliğin kullanıldığı durum ile oldukça benzerdir. Test alanı 2'nin 8 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırma sonuçları aşağıdaki şekildedir:

- Toplam doğruluk (overall accuracy) =  $(114+1)/(114+1+41+3) = 0.7233 = \%72$
- Hassasiyet (recall) (geniş yapraklı ağaçlar) =  $114/(114+3) = 0.9744$
- Hassasiyet (recall) (iğne yapraklı ağaçlar) =  $1/(1+41) = 0.0238$
- Kesinlik (precision) (geniş yapraklı ağaçlar) =  $114/(114+41) = 0.7355$
- Kesinlik (precision) (iğne yapraklı ağaçlar) =  $1/(1+3) = 0.25$
- F1-skor (geniş yapraklı ağaçlar) =  $2 \times (0.9744 \times 0.7355) / (0.9744 + 0.7355) = 0.8383$
- F1-skor (geniş yapraklı ağaçlar) =  $2 \times (0.9744 \times 0.7355) / (0.9744 + 0.7355) = 0.8383$

Test alanı 2'de geniş ve iğne yapraklı ağaçların elde edilebilmesi amacıyla gerçekleştirilen 8 özelliğin kullanıldığı ÇKA sınıflandırmasının toplam doğruluğu %72'dir. Geniş yapraklı ağaçlar için hassasiyet, kesinlik ve F1-skor değerleri sırasıyla 0.9744, 0.7355 ve 0.8383 olarak belirlenmiştir. İğne yapraklı ağaçlar için hassasiyet, kesinlik ve F1-skor değerleri ise sırasıyla 0.0238, 0.25 ve 0.3143 olarak hesaplanmıştır. Test alanı 2'de 8 özellik kullanılarak gerçekleştirilen ÇKA algoritması ile geniş ve iğne yapraklı kent ağaçları Test alanı 1'e göre daha yüksek doğrulukla belirlenmesine rağmen, sınıflandırma doğruluğunun iyileştirilmesine ihtiyaç vardır. Test alanı 2'de 8 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırma sonuçlarına göre geniş yapraklı ağaçlar yüksek doğrulukla tahmin edilse de iğne yapraklı ağaçlar için tam tersi durum söz konusudur. Test alanı 2'de iğne yapraklı ağaçların azlığı ve heterojen dağılımı ÇKA sınıflandırmasında başarı oranının beklenen seviyede olmamasında etkili olmuştur. Test alanı 2'de 8 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırmasının 10 katlamalı çapraz doğrulamaya (Tablo 4.35) göre ortalama doğruluğu ise %67 ile genel doğrulukla yaklaşık olarak tutarlıdır.

**Tablo 4.35** Test alanı 2'nin 8 özellik kullanılarak ÇKA sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları

10 Katlamalı Çapraz Doğrulama		
1. katlama doğruluğu	0.7170	Ortalama doğruluk: 0.6716 (%67)
2. katlama doğruluğu	0.7358	
3. katlama doğruluğu	0.6792	
4. katlama doğruluğu	0.3019	
5. katlama doğruluğu	0.7547	
6. katlama doğruluğu	0.8113	
7. katlama doğruluğu	0.7547	
8. katlama doğruluğu	0.2692	
9. katlama doğruluğu	0.7885	
10. katlama doğruluğu	0.9038	

Test alanı 2'nin 25 özelliğın kullanıldığı ÇKA sınıflandırmasında 25 nörondan oluşan bir girdi katmanı, 20 nörondan oluşan bir gizli katman ve 2 nörondan oluşan bir çıktı katmanı mevcuttur ve maksimum iterasyonların sayısı 2000 olarak kullanılmıştır. Test alanı 2'nin 25 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırma sonuçları Şekil 4.60 ve bu ÇKA sınıflandırmasına ilişkin hata matrisi Tablo 4.36'de yer almaktadır.



● Geniş yapraklı ağaçlar ● İğne yapraklı ağaçlar

**Şekil 4.60** Test alanı 2'nin 25 özellik kullanılarak ÇKA sınıflandırması

**Tablo 4.36** Test alanı 2'nin 25 özellik kullanılarak ÇKA sınıflandırmasına ilişkin hata matrisi

		Referans veri		
		Geniş	İğne	Toplam
Tahmin verisi	Geniş	37	5	42
	İğne	80	37	117
	Toplam	117	42	159

Test alanı 2'de 25 özelliğin kullanıldığı ÇKA sınıflandırmasında, diğer sınıflandırmalara göre doğru olarak sınıflandırılan geniş yapraklı ağaçların sayısında oldukça büyük bir düşüş söz konusu olmuştur (117 ağaçtan 37 doğru sınıflandırılan geniş yapraklı ağaçlar). Tam tersi olarak doğru olarak sınıflandırılan iğne yapraklı ağaçların sayısında da büyük bir yükseliş söz konusu olmuştur. 80 geniş yapraklı ağaç iğne yapraklı ağaç ile karışırken, yalnızca 5 iğne yapraklı ağaç geniş yapraklı ağaç ile karışmıştır. Buradaki sınıflandırmada ağaçların karışım oranları diğer tüm sınıflandırmalara göre tersine dönmüştür. Test alanı 2'nin 25 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırma doğrulukları aşağıdaki şekildedir:

- Toplam doğruluk (overall accuracy) =  $(37+37)/(37+37+5+80) = 0.4654 = \%47$
- Hassasiyet (recall) (geniş yapraklı ağaçlar) =  $37/(37+80) = 0.3162$
- Hassasiyet (recall) (iğne yapraklı ağaçlar) =  $37/(37+5) = 0.8810$
- Kesinlik (precision) (geniş yapraklı ağaçlar) =  $37/(37+5) = 0.8810$
- Kesinlik (precision) (iğne yapraklı ağaçlar) =  $37/(37+80) = 0.3162$
- F1-skor (geniş yapraklı ağaçlar) =  $2x(0.3162x0.8810)/(0.3162+0.8810) = 0.4654$
- F1-skor (iğne yapraklı ağaçlar) =  $2x(0.8810x0.3162)/(0.8810+0.3162) = 0.4654$

25 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırması sonucu Test alanı 2'de geniş ve iğne yapraklı ağaçların genel sınıflandırma doğruluğu %47'dir. Geniş yapraklı ağaçların 25 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırması sonucu elde edilen hassasiyet değeri 0.3162, kesinlik değeri 0.8810 ve F1-skor

değeri 0.4654 'dir. İğne yapraklı ağaçlar için ise hassasiyet değeri 0.8810, kesinlik değeri 0.3162 ve F1-skor değeri 0.4654'tür. Test alanı 2'nin 25 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırma sonuçları aynı alanın 8 özellik kullanılarak gerçekleştirilen sınıflandırma sonuçlarından oldukça düşüktür. 25 özelliğin kullanıldığı ÇKA, Test alanı 2'de başarılı sonuçlar üretememiştir. ÇKA algoritması ile 25 özellik kullanılarak gerçekleştirilen sınıflandırmada diğer tüm yöntemlerin aksine iğne yapraklı ağaçların elde edilme oranı geniş yapraklı ağaçlara göre çok daha başarılıdır. 25 özelliğin kullanıldığı ÇKA sınıflandırmasının 10 katlamalı çapraz doğrulamaya (Tablo 4.37) göre ortalama doğruluğu ise Test alanı 2'de %69 olarak elde edilmiştir. Genel doğruluk ve 10-katlama doğruluğundan da anlaşılacağı üzere ÇKA algoritma sınıflandırmasının sonuçları çok güvenilir değildir.

**Tablo 4.37** Test alanı 2'nin 25 özellik kullanılarak ÇKA sınıflandırmasının 10 katlamalı çapraz doğrulama sonuçları

10 Katlamalı Çapraz Doğrulama		
1. katlama doğruluğu	0.7736	Ortalama doğruluk: 0.6940 (%69)
2. katlama doğruluğu	0.7170	
3. katlama doğruluğu	0.5283	
4. katlama doğruluğu	0.7547	
5. katlama doğruluğu	0.7170	
6. katlama doğruluğu	0.8113	
7. katlama doğruluğu	0.8302	
8. katlama doğruluğu	0.75	
9. katlama doğruluğu	0.75	
10. katlama doğruluğu	0.3077	

#### 4.7.4 Makine Öğrenme Tabanlı Sınıflandırma Algoritmalarının Performanslarının Karşılaştırılması

Bu bölümde Test alanı 1 ve Test alanı 2 olmak üzere her iki test alanında da geniş ve iğne yapraklı ağaç türlerinin sınıflandırılmasında kullanılan üç farklı makine öğrenme algoritmalarının performanslarının karşılaştırılması gerçekleştirilecektir.

##### 4.7.4.1 Test Alanı 1'de DVM, RO ve ÇKA Algoritmalarının Karşılaştırılması

Segmentlere ayrılmış bireysel ağaçlar belirlenen parametreler kullanılarak destek vektör makinesi, rastgele orman ve çok katmanlı algılayıcı makine öğrenimi

sınıflandırma yöntemleriyle LiDAR verilerinin geometrik ve yoğunluk tabanlı özelliklerine dayalı olarak geniş yapraklı veya iğne yapraklı ağaç türleri olarak sınıflandırılmıştır. Test alanı 1’de 8 özellik kullanılarak gerçekleştirilen nokta tabanlı DVM, RO ve ÇKA sınıflandırma sonuçları Şekil 4.61’de ve 25 özellik kullanılarak gerçekleştirilen sınıflandırma sonuçları Şekil 4.62’de verilmiştir. Test alanı 1’de tüm yaklaşımlar (DVM, RO ve ÇKA) kullanılarak elde edilen sınıflandırmanın genel doğrulukları ise Tablo 4.38’de yer almaktadır. Sınıflandırma sonuçları incelendiğinde DVM, RO ve ÇKA sınıflandırıcılarının kentsel ağaç türlerinden geniş ve iğne yapraklı ağaçları ayırmada kullanılabileceği sonucuna varılabilir. Test alanı 1’de 8 sınıflandırma özelliğinin kullanıldığı DVM ile RO algoritmalarının birbirlerine oldukça yakın, ÇKA’nın ise bu iki yöntemden farklı sonuçlar ürettiği ifade edilebilir. Geniş ve iğne yapraklı ağaçların sınıflandırılmasında %80 genel doğruluk ile en iyi sınıflandırma sonucunu veren rastgele orman sınıflandırıcısıdır. DVM sınıflandırıcısı ise %76.25 genel doğruluk ile RO’dan sonra en iyi sonucu veren algoritmadır. 8 özelliğın kullanıldığı Test alanı 1’de geniş ve iğne yapraklı ağaçların sınıflandırılmasında % 67.5 genel doğruluk üreten ÇKA algoritması ise 3 sınıflandırma yöntemi arasındaki en düşük genel doğruluk sonucunu vermiştir.

Test alanı 1’de 25 sınıflandırma özelliğinin kullanıldığı sınıflandırmalarda DVM ile RO algoritmalarının sınıflandırma sonuçları 8 özelliğın kullanıldığı durumda olduğu gibi birbirlerine oldukça yakın, ÇKA ile gerçekleştirilen sınıflandırma sonuçlarından ise nispeten farklıdır. Test alanı 1’de 25 özelliğın kullanıldığı sınıflandırma uygulamalarında, 8 özelliğın kullanıldığı duruma benzer şekilde %83.75 genel doğruluk ile RO en iyi sonucunu veren makine öğrenimi tabanlı sınıflandırıcı olmuştur. %80 genel doğruluk ile DVM algoritması RO’ı takip etmektedir. 25 özelliğın kullanıldığı sınıflandırmada %73.75 genel doğruluk ile ÇKA sınıflandırıcısı diğer algoritmalarından daha kötü sonuç üretmiştir. Genel olarak ise Test alanı 1’de 25 özelliğın kullanıldığı tüm DVM, RO ve ÇKA sınıflandırma modelleri, 8 özelliğın kullanıldığı sınıflandırma modellerinden daha doğru sınıflandırma gerçekleştirmişlerdir. 8 özelliğın kullanıldığı sınıflandırmalara göre 25 özellik kullanılarak DVM ve RO %3.75, ÇKA algılayıcı ise %6.25 daha yüksek sınıflandırma doğruluğuna ulaşmıştır.



● Geniş yapraklı ağaçlar ● İğne yapraklı ağaçlar  
(a) (b) (c)

Şekil 4.61 Test alanı 1'in 8 özellik kullanılarak gerçekleştirilen DVM (a), RO (b) ve ÇKA (c) sınıflandırmaları



● Geniş yapraklı ağaçlar ● İğne yapraklı ağaçlar  
(a) (b) (c)

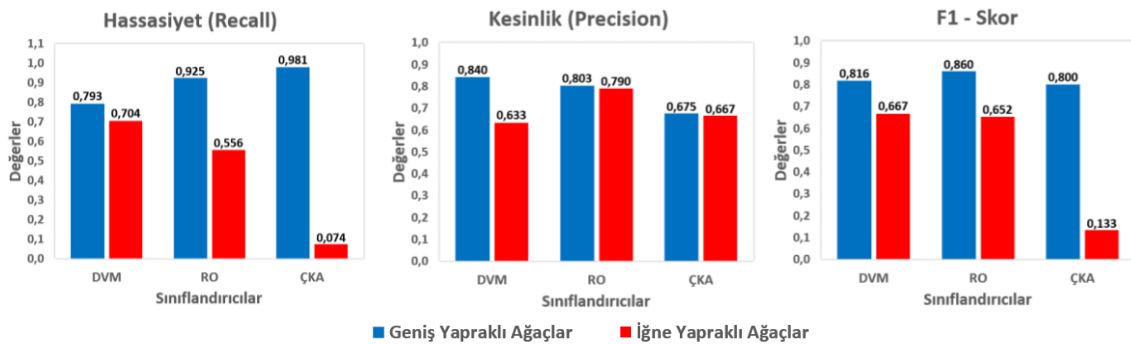
Şekil 4.62 Test alanı 1'in 25 özellik kullanılarak gerçekleştirilen DVM (a), RO (b) ve ÇKA (c) sınıflandırmaları

Tablo 4.38 Test alanı 1'de 8 ve 25 sınıflandırma özelliği kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırma sonuçları

Test Alanı 1		
Sınıflandırma yöntemi	Genel doğruluk (8 özellik kullanılarak)	Genel doğruluk (25 özellik kullanılarak)
DVM	% 76.25	% 80.00
RO	% 80.00	% 83.75
ÇKA	% 67.50	% 73.75

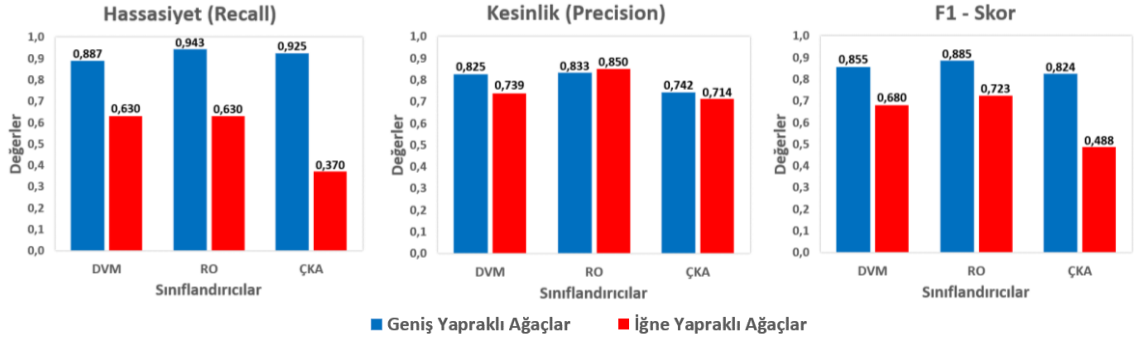
Test alanı 1'de 8 özellik kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırma sonucunda her bir ağaç için ayrı ayrı elde edilen hassasiyet (recall), kesinlik (precision) ve F1-skor değerleri Şekil 4.63'de gösterilmiştir. 8 özelliğin

kullanıldığı sınıflandırmalarda 0.981 olan hassasiyet değerine göre ÇKA geniş yapraklı ağaçlar için en iyi sınıflandırma metodu olarak göze çarpmaktadır. Geniş yapraklı ağaçlar için hassasiyet değeri RO algoritmasında ÇKA'a benzer şekilde yüksektir (0.925), DVM'de ise nispeten düşüktür (0.793). Geniş yapraklı ağaçlar ile kıyaslandığında iğne yapraklı ağaçlar için tüm sınıflandırıcıların hassasiyet değerleri daha düşük elde edilmiştir. İğne yapraklı ağaçların hassasiyet değeri DVM sınıflandırıcısı ile en yüksek (0.704), ÇKA sınıflandırıcısı ile 0.074 değeri ile en düşük olarak elde edilmiştir. Çok düşük hassasiyet değerinden dolayı geniş yapraklı ağaçların tersine ÇKA sınıflandırıcısı iğne yapraklı ağaçların elde edilmesi için uygun bir yöntem değildir. Geniş yapraklı ağaçlar için DVM ve RO ile birbirine yakın kesinlik değerleri (0.840 ve 0.803) elde edilirken ÇKA ile en düşük kesinlik değeri (0.675) elde edilmiştir. Kesinlik değerine göre ÇKA geniş yapraklı ağaçlar için en başarısız sınıflandırma yöntemidir. İğne yapraklı ağaçlar için RO 0.790 kesinlik değeri ile en iyi sınıflandırma yöntemidir. 0.633 ve 0.667 sırasıyla DVM ve ÇKA yaklaşımları için daha düşük kesinlik değerleridir. F1-skor değerine göre, 8 özelliğin kullanıldığı tüm sınıflandırma yaklaşımları ile geniş yapraklı ağaçlar iğne yapraklı ağaçlardan daha iyi sınıflandırılmıştır. 0.860 F1-skor değerine göre RO geniş yapraklı ağaçlar için en başarılı yöntemdir. İğne yapraklı ağaçlar için ÇKA sınıflandırıcısı, 0.133 F1-skor değeri ile en başarısız yöntemdir. DVM (0.667 F1-skor) ve RO (0.652 F1-skor) yöntemleri ile iğne yapraklı ağaçlar ÇKA'ya göre daha başarılı sınıflandırılmışlardır. Genel olarak, 8 özelliğin kullanıldığı tüm sınıflandırma yaklaşımları geniş yapraklı ağaçları iğne yapraklı ağaçlara göre çok daha iyi sınıflandırmıştır.



**Şekil 4.63** Test alanı 1'de her bir geniş ve iğne yapraklı ağaç için 8 özellik kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırıcılarının hassasiyet, kesinlik ve F1-skor değerleri

Test alanı 1'de 25 özellik kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırma sonucunda her bir ağaç için ayrı ayrı elde edilen hassasiyet (recall), kesinlik (precision) ve F1-skor değerleri Şekil 4.64'de verilmiştir. 25 özelliğin kullanıldığı sınıflandırmalarda hassasiyet değerine göre geniş yapraklı ağaçlar başarılı bir şekilde sınıflandırılmışlardır. RO (0.943) geniş yapraklı ağaçlar için en iyi sınıflandırma yaklaşımı olmuştur. Sırasıyla ÇKA (0.925) ve DVM sonraki başarılı yaklaşımlardır. Geniş yapraklı ağaçlar ile kıyaslandığında iğne yapraklı ağaçlar için daha düşük hassasiyet değerleri elde edilmiştir. DVM ve RO'nun hassasiyet değeri birbirinin aynı iken (0.630), ÇKA en düşük hassasiyet değerine sahip (0.370) algoritma olmuştur. Bu nedenle, ÇKA'nın iğne yapraklı ağaçların sınıflandırılması için yeterli bir yaklaşım olmadığı ifade edilebilir. 25 özelliğin kullanıldığı sınıflandırmalarda geniş yapraklı ağaçlar için DVM ve RO ile benzer kesinlik değerleri elde edilmiştir (sırasıyla 0.825 ve 0.833). Kesinlik değerine göre ÇKA (0.742), üç yaklaşım arasında en başarısız yöntem olmuştur. İğne yapraklı ağaçlar için RO (0.850) kesinlik değerine göre iyi sınıflandırma sonucunu veren yaklaşım olmuştur. 0.739 ve 0.714 değerleri sırasıyla DVM ve ÇKA yaklaşımları için daha düşük kesinlik değerleri olmuştur. 8 sınıflandırma özelliğinin kullanıldığı sınıflandırmalarda olduğu gibi, 25 özelliğin kullanıldığı tüm sınıflandırmalarda da F1-skor değerine göre geniş yapraklı ağaçlar, iğne yapraklı ağaçlardan daha başarılı elde edilmişlerdir. Geniş yapraklı ağaçlar için ÇKA algılayıcı 0.488 F1-skor değeri ile en başarısız sınıflandırma yaklaşımı olarak belirlenmiştir. F1-skor değerine göre RO (0.73) en başarılı sınıflandırma metodu, DVM (0.680) ise benzer şekilde iyi bir sınıflandırma yaklaşımıdır. Test alanı 1'de, genel olarak 25 özelliğin kullanıldığı sınıflandırma algoritmalarında geniş yapraklı ağaçlar için hassasiyet, kesinlik ve F1-skor değerleri, iğne yapraklı ağaçlardan daha yüksektir. Sonuç olarak, Test alanı 1'de 25 özellik kullanılarak gerçekleştirilen makine öğrenimi tabanlı sınıflandırmalarda DVM, RO ve ÇKA algoritmaları iğne yapraklı ağaçlar ile kıyaslandığında geniş yapraklı ağaçlar için daha başarılı sonuçlar üretmişlerdir.



**Şekil 4.64** Test alanı 1’de her bir geniş ve iğne yapraklı ağaç için 25 özellik kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırıcılarının hassasiyet, kesinlik ve F1-skor değerleri

Test alanı 1’de 8 ve 25 sınıflandırma özelliği kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırmalarının 10 katlamalı çapraz doğrulamaya göre ortalama doğrulukları Tablo 4.39’da verilmiştir. 8 özelliğin kullanıldığı sınıflandırmalarda RO algoritması ile gerçekleştirilen sınıflandırmanın 10 kat çapraz doğrulamaya göre ortalama doğruluğu DVM’nin ortalama doğruluğundan %1.88, ÇKA’nın ortalama doğruluğundan %29.17 daha yüksek olarak en yüksek doğruluktur (%78.12). ÇKA’nın 10 kat çapraz doğrulamaya göre ortalama doğruluğu en düşük doğruluktur (%48.95). Kentsel ağaç türlerinin 10 kat çapraz doğrulamaya göre ortalama sınıflandırma doğruluğu DVM ve RO algoritmaları için birbirine oldukça yakinken, ÇKA algoritması için DVM ve RO’ya kıyasla oldukça düşüktür. 25 özelliğin kullanıldığı DVM, RO ve ÇKA sınıflandırmalarında da algoritmaların birbirlerine göre performansları 8 özelliğin kullanıldığı sınıflandırmalar ile aynıdır fakat 25 özelliğin kullanıldığı sınıflandırmalarda 10 kat çapraz doğrulamaya göre ortalama sınıflandırma doğrulukları 8 özellikli sınıflandırmalardan daha yüksektir. Buna göre 25 özelliğin kullanıldığı RO sınıflandırmasında, ortalama sınıflandırma doğruluğu %81.54 olarak DVM’den %0.44, ÇKA’dan %17.82 daha yüksek bir şekilde elde edilmiştir. 25 özelliğin kullanıldığı sınıflandırmada da ÇKA en kötü ortalama sınıflandırma doğruluğuna sahip algoritma olmuştur. Genel olarak 10 katlamalı çapraz doğrulama sonuçları incelendiğinde 25 özelliğin kullanıldığı sınıflandırmaların 8 özelliğe göre daha başarılı sonuçlar ürettiği görülmektedir (Tablo 4.39). Tüm sınıflandırıcıların (DVM, RO ve ÇKA) 10 kat çapraz doğrulamaya göre ortalama doğruluklarının, genel doğrulukları değerleri

ile orantılı olduğu ifade edilebilir. Buradan hareketle gerçekleştirilen tüm sınıflandırma uygulamalarının güvenilir ve tutarlı sonuçlar ürettiği vurgulanabilir.

**Tablo 4.39** Test alanı 1’de 8 ve 25 sınıflandırma özelliği kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırmalarının 10 katlamalı çapraz doğrulamaya göre ortalama doğrulukları

Test Alanı 1		
Sınıflandırma yöntemi	10 katlamalı çapraz doğrulama ortalama doğruluk (8 özellik kullanılarak)	10 katlamalı çapraz doğrulama ortalama doğruluk (25 özellik kullanılarak)
DVM	% 76.24	% 81.10
RO	% 78.12	% 81.54
ÇKA	% 48.95	% 63.72

#### 4.7.4.2 Test Alanı 2’de DVM, RO ve ÇKA Algoritmalarının Karşılaştırılması

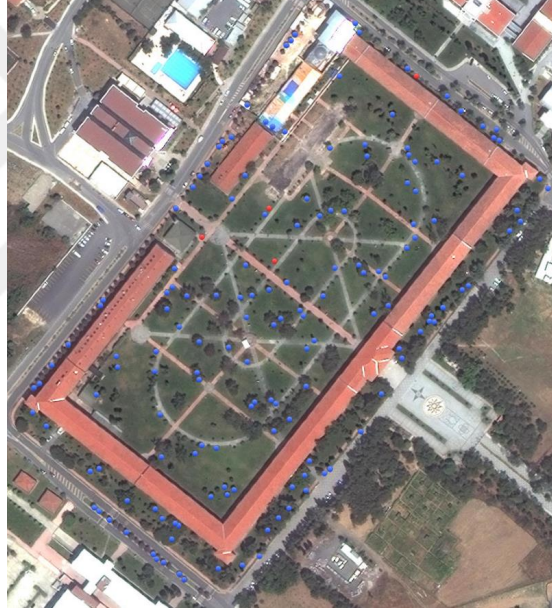
Test alanı 2’de de Test alanı 1’de olduğu gibi segmentasyon sonucu elde edilen bireysel ağaç taçları, LiDAR verilerinin geometrik ve yoğunluk tabanlı özellikleri kullanılarak elde edilen 2 farklı özellik seti kullanılarak DVM, RO ve ÇKA sınıflandırıcıları ile kentsel ağaç türleri elde edilmiştir. Test alanı 2’de 8 özellik kullanılarak gerçekleştirilen makine öğrenimi tabanlı DVM, RO ve ÇKA sınıflandırma sonuçları Şekil 4.65’de ve 25 özellik kullanılarak gerçekleştirilen sınıflandırma sonuçları ise Şekil 4.66’da verilmiştir. DVM, RO ve ÇKA kullanılarak elde edilen 3B nokta tabanlı sınıflandırmanın genel doğrulukları ise Tablo 4.40’de yer almaktadır. Sınıflandırma sonuçları incelendiğinde Test alanı 2’de de DVM, RO ve ÇKA sınıflandırıcılarının geniş ve iğne yapraklı ağaç türlerini ayırmada kullanılabileceği ifade edilebilir. Test alanı 2’de eğitim aşamalarının ardından gerçekleştirilen test verilerinin (%30) 8 özelliğin kullanıldığı makine öğrenme tabanlı sınıflandırmalarında geniş ve iğne yapraklı ağaçların ayırt edilebilmesi için kullanılan tüm algoritmaların (DVM, RO ve ÇKA) sınıflandırma sonuçları birbirine oldukça yakındır. En iyi sınıflandırma çok katmanlı algılayıcı (ÇKA) ile %72.33 olarak elde edilmiştir. DVM algoritması ÇKA’ya oldukça yakın sınıflandırma toplam doğruluğu (%71.70) sahiptir. Test alanı 1’in tersine Test alanı 2’de %69.8 genel doğrulukla RO en düşük sınıflandırma sonucunu vermiştir.

Test alanı 2’de 25 özelliğin kullanıldığı sınıflandırma sonuçları iki yaklaşım için (DVM ve RO) birbirinin aynısı, diğer yaklaşımın (ÇKA) doğruluğu ise bu yaklaşımlara göre oldukça düşüktür. Test alanı 2’de 25 özelliğin kullanıldığı geniş ve iğne yapraklı ağaçları sınıflandırma uygulamalarında %72.33 ile DVM ve RO aynı genel doğruluğa sahip, en iyi sonucu veren iki algoritma olmuştur. Test alanı 2’de kentsel ağaç türlerinin sınıflandırılması için 25 özelliğin kullanıldığı durumda 8 özelliğin aksine ÇKA sınıflandırıcısı en kötü sınıflandırma sonucuna sahip (% 46.17) algoritma olmuştur. Genel olarak Test alanı 2’de 25 özelliğin kullanıldığı DVM ve RO sınıflandırma modelleri 8 özelliğin kullanıldığı sınıflandırma modellerinden daha başarılı sınıflandırma gerçekleştirmişlerdir. Fakat bu iki yöntemin aksine Test alanı 2’de 8 özelliğin kullanıldığı ÇKA sınıflandırma modeli 25 özelliğin kullanıldığı sınıflandırmadan daha başarılı sonuç üretmiştir. 8 özelliğin kullanıldığı sınıflandırmalara göre 25 özellik kullanılarak DVM ve RO algoritmaları sırasıyla %0.63 ve %2.53 daha yüksek sınıflandırma doğruluğuna ulaşmış, ÇKA algılayıcı ise büyük farkla %26.16 daha düşük sınıflandırma doğruluğuna sahip olmuştur.



(a)

(b)



● Geniş yapraklı ağa lar ● İğne yapraklı ağa lar  
(c)

 ekil 4.65 Test alanı 2'nin 8  zellik kullanılarak ger ekleřtirilen DVM (a), RO (b) ve  KA (c) sınıflandırmaları



(a)



(b)



● Geniş yapraklı ağaçlar ● İğne yapraklı ağaçlar  
(c)

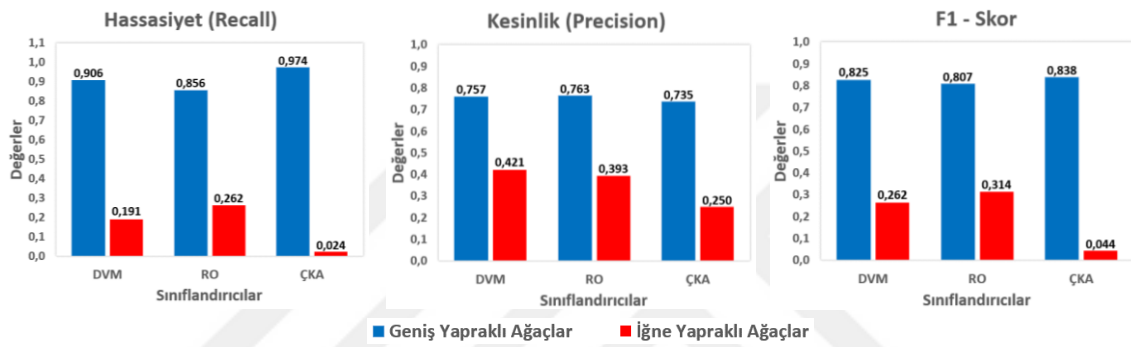
**Şekil 4.66** Test alanı 2'nin 25 özellik kullanılarak gerçekleştirilen DVM (a), RO (b) ve ÇKA (c) sınıflandırmaları

**Tablo 4.40** Test alanı 2’de 8 ve 25 sınıflandırma özelliği kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırma sonuçları

Test Alanı 2		
Sınıflandırma yöntemi	Genel doğruluk (8 özellik kullanılarak)	Genel doğruluk (25 özellik kullanılarak)
DVM	% 71.70	% 72.33
RO	% 69.81	% 72.33
ÇKA	% 72.33	% 46.54

Test alanı 2’de 8 özellik kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırma sonucunda her bir ağaç için ayrı ayrı elde edilen hassasiyet (recall), kesinlik (precision) ve F1-skor değerleri Şekil 4.67’de gösterilmiştir. 8 özelliğin kullanıldığı sınıflandırmalarda Test alanı 1’de olduğu gibi Test alanı 2’de de hassasiyet değerine göre (0.974) ÇKA geniş yapraklı ağaçlar için en iyi sınıflandırma yöntemi olmuştur. Geniş yapraklı ağaçlar için hassasiyet değeri DVM ve RO algoritmalarında da oldukça yüksektir (sırasıyla 0.906 ve 0.856). Test alanı 2’de 8 özelliğin kullanıldığı sınıflandırmalarda tüm sınıflandırıcılar ile elde edilen iğne yapraklı ağaçların hassasiyet değeri geniş yapraklı ağaçlara göre oldukça düşüktür. ÇKA sınıflandırıcısı geniş yapraklı ağaçların aksine 0.024 hassasiyet değeri ile iğne yapraklı ağaçlar için en düşük sınıflandırma doğruluğuna sahiptir. Hassasiyet değerleri dikkate alındığında DVM, RO ve ÇKA sınıflandırıcılarının hiçbirinin iğne yapraklı ağaçlar için uygun sınıflandırıcılar olmadığı ifade edilebilir. 8 özelliğin kullanıldığı sınıflandırmalarda geniş yapraklı ağaçlar için tüm sınıflandırma modellerinin (DVM, RO ve ÇKA) kesinlik değerleri birbirine yakındır. Kesinlik değerine göre 0.763 ile RO en yüksek doğruluğa sahip, 0.735 değeri ile ÇKA en düşük doğruluğa sahiptir. Hassasiyet değerine benzer şekilde 8 sınıflandırma özelliğinin kullanıldığı tüm DVM, RO ve ÇKA sınıflandırma modellerinin kesinlik değerleri iğne yapraklı ağaçlar için yetersizdir. İğne yapraklı ağaçların sınıflandırılmasında 0.421 kesinlik değeri ile DVM en yüksek sınıflandırma doğruluğuna sahip, 0.250 kesinlik değeri ile de ÇKA en düşük sınıflandırma doğruluğuna sahiptir. Test alanı 2’de 8 özelliğin kullanıldığı sınıflandırma uygulamalarında hassasiyet ve kesinlik için mevcut olan durum F1-skor değeri için de geçerlidir. F1-skor değeri geniş yapraklı ağaçlarda tüm

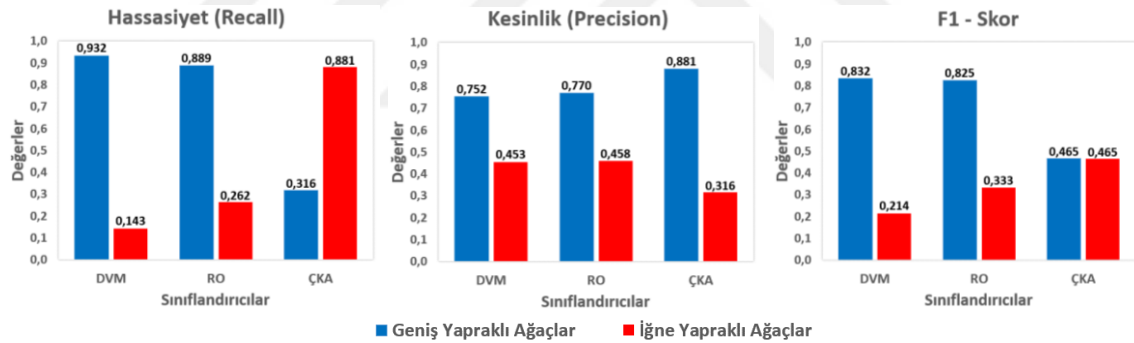
sınıflandırıcılar için birbirine yakın (0.825, 0.807 ve 0.838) olarak elde edilmiştir. Bu sınıflandırılardan ÇKA 0.838 değeri ile geniş yapraklı ağaçlar için en uygun yöntem olarak ön plana çıkmaktadır. İğne yapraklı ağaçlar için ise F1-skor değerleri diğer doğruluk değerlendiricilerine benzer şekilde oldukça düşüktür. 8 özellikli nokta tabanlı sınıflandırmalarda en yüksek F1-skor değeri RO algoritması ile (0.314) elde edilmiştir, en düşük değer ise ÇKA ile (0.044) elde edilmiştir. Genel olarak, Test alanı 2’de 3B nokta tabanlı sınıflandırmalarda 8 özelliğin kullanıldığı tüm sınıflandırma yaklaşımları geniş yapraklı ağaçları iğne yapraklı ağaçlara göre çok daha iyi sınıflandırmıştır.



**Şekil 4.67** Test alanı 2’de her bir geniş ve iğne yapraklı ağaç için 8 özellik kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırıcılarının hassasiyet, kesinlik ve F1-skor değerleri

Test alanı 2’de 25 özellik kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırma sonucunda her bir ağaç için ayrı ayrı elde edilen hassasiyet (recall), kesinlik (precision) ve F1-skor değerleri Şekil 4.68’de yer almaktadır. Test alanı 2’de 25 özelliğin kullanıldığı sınıflandırmalarda hassasiyet değerine göre DVM (0.932) geniş yapraklı ağaçlar için en başarılı sınıflandırıcı olmuştur. Geniş yapraklı ağaçlar için hassasiyet değerine göre RO (0.889) algoritması da başarılı bir yöntem iken 0.316 hassasiyet değeri ile ÇKA yeterince başarılı bir yaklaşım değildir. Bunun aksine iğne yapraklı ağaçlar için 0.881 hassasiyet değeri ile ÇKA en iyi sınıflandırma sonucunu veren yaklaşım olmuştur. Sırasıyla 0.143 ve 0.262 hassasiyet değerleri ile DVM ve RO’nun iğne yapraklı ağaçların nokta tabanlı sınıflandırma ile elde edilmesi için uygun yaklaşımlar olmadığı ortaya konmuştur. Test alanı 2’de 25 özellikli sınıflandırmaların kesinlik değerleri incelendiğinde geniş yapraklı ağaçların tüm sınıflandırma modelleri ile iğne yapraklı ağaçlara göre daha hassas sınıflandırıldığı görülmektedir. ÇKA, 0.881 kesinlik değeri ile

geniş yapraklı ağaçları en doğru sınıflandıran yaklaşımdır. ÇKA'yı sırasıyla RO (0.770) ve DVM (0.752) sınıflandırıcıları takip etmektedir. İğne yapraklı ağaçlar için kesinlik değerleri analiz edildiğinde tüm sınıflandırıcıların (DVM, RO ve ÇKA) sınıflandırma doğruluklarının ortalamasının altında kaldığı ifade edilebilir. Kesinlik değerine göre RO (0.458) ve DVM (0.453) en başarılı yöntemler iken ÇKA (0.316) en yetersiz sınıflandırma yöntemidir. Test alanı 2'de 25 özelliğin kullanıldığı sınıflandırmalarda F1-skor değeri DVM (0.832) ve RO (0.825) sınıflandırma modelleri için yüksektir. F1-skor değeri ÇKA algoritması ile ise 0.465 olarak oldukça düşük elde edilmiştir. 25 özelliğin kullanıldığı sınıflandırmalarda iğne yapraklı ağaçların elde edilebilmesi için tüm yaklaşımların F1-skor değerleri ortalamasının altında kalmıştır. Özellikle 0.214 F1-skor değeri ile DVM algoritması iğne yapraklı ağaçlar için en başarısız yöntem olarak belirlenmiştir. Sonuç olarak, çoğunlukla Test alanı 2'de 25 özellik kullanılarak gerçekleştirilen sınıflandırmalarda her üç sınıflandırma algoritması da geniş yapraklı ağaçlar için daha başarılı olmuştur.



**Şekil 4.68** Test alanı 2'de her bir geniş ve iğne yapraklı ağaç için 25 özellik kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırıcılarının hassasiyet, kesinlik ve F1-skor değerleri

Test alanı 2'de 8 ve 25 sınıflandırma özelliği kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırmalarının 10 katlamalı çapraz doğrulamaya göre ortalama doğrulukları Tablo 4.41'de verilmiştir. 8 özelliğin kullanıldığı sınıflandırmalarda DVM ile gerçekleştirilen sınıflandırmanın 10 kat çapraz doğrulamaya göre ortalama doğruluğu %77.80 ile en yüksek doğruluktur. DVM, RO'dan %2.26, ÇKA'dan ise %10.64 daha yüksek ortalama doğrulukla nokta tabanlı sınıflandırma gerçekleştirmiştir. Test alanı 1'de olduğu gibi 8 özelliğin kullanıldığı ÇKA sınıflandırmasının 10 kat çapraz doğrulamaya göre ortalama doğruluğu en düşük

doğrudur (%67.16). Test alanı 2’de hem 8 ve hem de 25 özelliğin kullanıldığı sınıflandırmalarda 10 kat çapraz doğrulamaya göre ortalama sınıflandırma doğrulukları DVM ve RO algoritmaları için birbirine oldukça yakınken, ÇKA algoritması için nispeten DVM ve RO’dan düşüktür. 25 özelliğin kullanıldığı sınıflandırmalarda 8 özellikten farklı olarak RO algoritması ortalama %80.30 ile en doğru sonucu veren sınıflandırma algoritması olmuştur. ÇKA ise %69.40 ortalama doğruluk ile en düşük doğrulukla kentsel ağaç türlerini sınıflandırılan algoritma olmuştur. Genel olarak 10 katlamalı çapraz doğrulama sonuçları analiz edildiğinde Test alanı 2’de de Test alanı 1’de olduğu gibi 25 özelliğin kullanıldığı sınıflandırmaların 8 özelliğe göre daha başarılı sonuçlar ürettiği ifade edilebilir (Tablo 4.41). Tüm sınıflandırıcıların (DVM, RO ve ÇKA) 10 kat çapraz doğrulamaya göre ortalama doğrulukları, genel doğruluklarından daha yüksektir. Test alanı 1’e göre, Test alanı 2’de elde edilen sınıflandırma doğrulukları biraz daha tutarsızdır. Daha açık bir ifade ile DVM, RO ve ÇKA algoritmaları Test alanı 1’de daha başarılı ve daha güvenilir sınıflandırma sonuçları üretmiştir.

**Tablo 4.41** Test alanı 2’de 8 ve 25 sınıflandırma özelliği kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırmalarının 10 katlamalı çapraz doğrulamaya göre ortalama doğrulukları

Test Alanı 2		
Sınıflandırma yöntemi	10 katlamalı çapraz doğrulama ortalama doğruluk (8 özellik kullanılarak)	10 katlamalı çapraz doğrulama ortalama doğruluk (25 özellik kullanılarak)
DVM	% 77.80	% 78.58
RO	% 75.54	% 80.30
ÇKA	% 67.16	% 69.40

#### 4.7.5 Sınıflandırma Özelliklerinin Önemlerinin Belirlenmesi

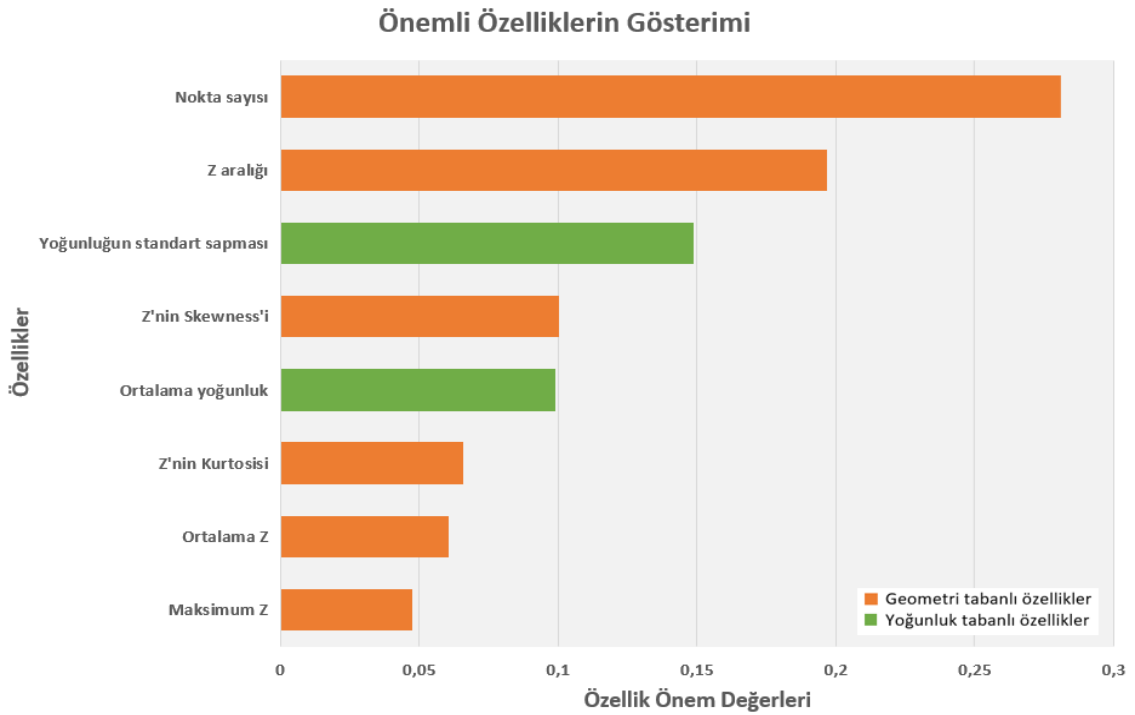
Bu çalışmada, kentsel ağaç türlerinin 3B nokta tabanlı sınıflandırılmasında kullanılan iki farklı özellik setindeki (8 ve 25 adet özellikten oluşan) geometrik ve yoğunluk tabanlı özelliklerin önemlerinin (feature importance) belirlenmesi için bir özellik analiz çalışması gerçekleştirilmiştir. RO sınıflandırması ile ortalama Gini’deki azalmaya (Mean Decrease in Gini - MDG) göre sınıflandırma özelliklerinin önemleri belirlenmiştir. Özellik önemlerinin belirlenmesinin

ardından, iki farklı özellik seti için (8 ve 25 adet özellik için) özelliklerin birbirleriyle ilişkilerini gösteren “Pearson” korelasyon matrisleri (correlation coefficient) de analiz edilmiştir.

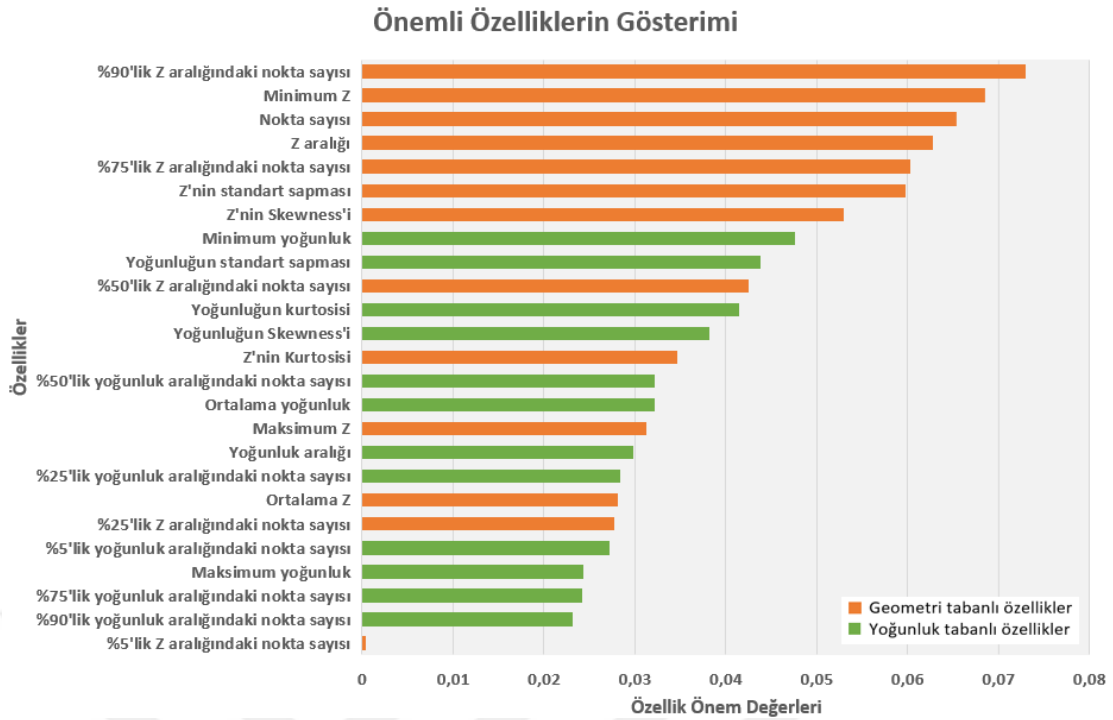
Test alanı 1 ve Test alanı 2’de gerçekleştirilen özellik önem analizleri ve 8 ve 25 özellik için korelasyon matrisleri bu bölümde gösterilmiştir.

#### 4.7.5.1 Test Alanı 1’de Önemli Özelliklerin Belirlenmesi

Bir test alanında ortalama Gini’deki azalmaya (Mean Decrease in Gini - MDG) göre belirlenen sınıflandırma özellik önem değerlerinin toplamı 1’e eşittir. Özellik önem değeri yüksek olan özellik sınıflandırma uygulaması için daha önemli iken özellik önemi düşük özellik ise sınıflandırma uygulaması için daha az önemlidir. Test alanı 1’de RO sınıflandırması ile ortalama Gini’deki azalmaya göre belirlenen 8 sınıflandırma özelliğinin önemleri Şekil 4.69 ve 25 sınıflandırma özelliğinin önemleri ise Şekil 4.70’de verilmiştir.



**Şekil 4.69** Test alanı 1’de RO sınıflandırmasındaki ortalama Gini’deki azalmaya göre 8 sınıflandırma özelliğinin önemleri (feature importance scores)



**Şekil 4.70** Test alanı 1’de RO sınıflandırmasındaki ortalama Gini’deki azalmaya göre 25 sınıflandırma özelliğinin önemleri (feature importance scores)

Tüm özelliklerin özellik önem değerleri toplamı 1’e eşittir. Şekil 4.69’daki özellik önemleri incelendiğinde 8 özellik arasında en önemli sınıflandırma özelliğinin “Nokta sayısı”, en az önemli sınıflandırma özelliğinin ise “Maksimum Z” olduğu görülmektedir. “Nokta sayısı” ndan sonra “Z aralığı”, “Yoğunluğun standart sapması” ve “Z’nin skewness’i” en önemli 3 sınıflandırma özelliği olarak göze çarpmaktadır. Test alanı 1’de RO sınıflandırmasındaki ortalama Gini’deki azalmaya göre, 8 özellikten özellikle geometri tabanlı iki özelliğin (“Nokta sayısı” ve “Z aralığı”) yoğunluktan tabanlı iki özellikten (“Yoğunluğun standart sapması” ve “Ortalama yoğunluk”) daha önemli olduğu sonucuna varılabilir.

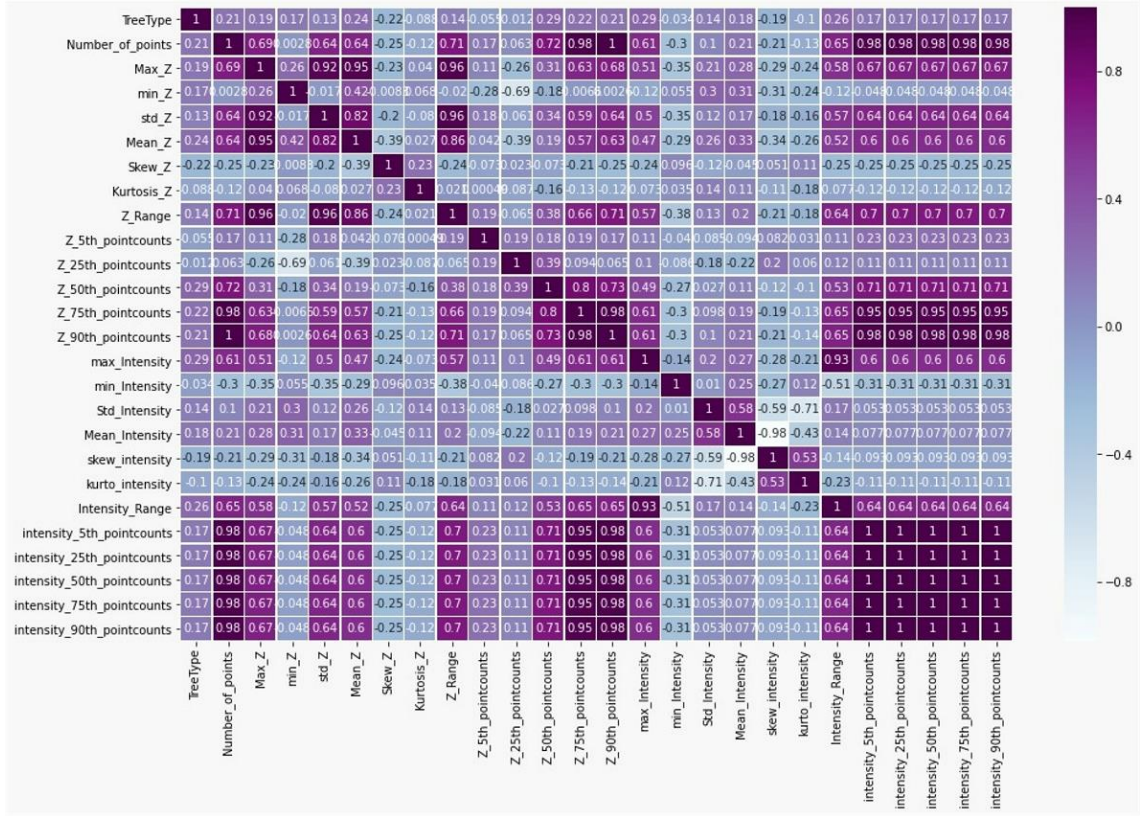
Şekil 4.70’deki özellik önemleri incelendiğinde 25 özellik arasında en önemli sınıflandırma özelliğinin “%90’lık Z aralığındaki nokta sayısı”, en az önemli olan sınıflandırma özelliğinin ise “%5’lik Z aralığındaki nokta sayısı” olduğu açıkça görülmektedir. “Minimum Z”, “Nokta sayısı”, “Z aralığı”, “%75’lik Z aralığındaki nokta sayısı” ve “Z’nin standart sapması”, “%90’lık Z aralığındaki nokta sayısı” dan sonraki en önemli beş özellik olmuştur. Genel olarak, 25 özelliğin kullanıldığı RO sınıflandırmasında ortalama Gini’deki azalmaya göre önerilen 3B nokta tabanlı

sınıflandırma için geometrik özelliklerin yoğunluk tabanlı özelliklerden daha yüksek bir öneme sahip olduğu ifade edilebilir.

Bir test alanında gerçekleştirilen sınıflandırma özelliklerinin birbiriyle ilişkileri -1 ile 1 arasında değer alabilir. Buna göre mutlak değer olarak aralarındaki korelasyon 1'e yakın olan özellikler yüksek korelasyonlu, 0'a yakın özellikler ise düşük korelasyonlu olarak ifade edilebilir. Özellik önemlerine ek olarak özelliklerin birbirleriyle ilişkilerinin analiz edildiği "Pearson" korelasyon matrisleri de Test alanı 1'de 8 sınıflandırma özelliği ve 25 sınıflandırma özelliği için sırasıyla Şekil 4.71 ve Şekil 4.72'de yer almaktadır.



Şekil 4.71 Test alanı 1'de 8 sınıflandırma özelliğinin Pearson korelasyon matrisi



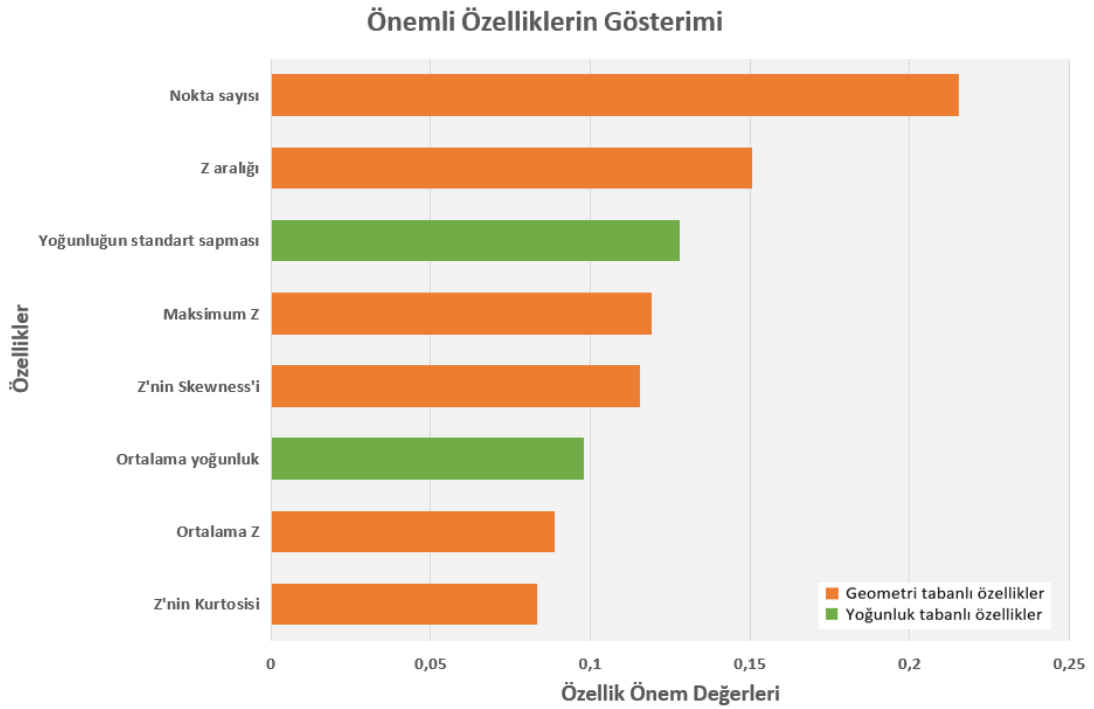
**Şekil 4.72** Test alanı 1’de 25 sınıflandırma özelliğinin Pearson korelasyon matrisi  
 Test alanı 1’de 8 sınıflandırma özelliğinin birbirlerine göre ilişkileri incelendiğinde tüm sınıflandırma özelliklerinin hedef sınıf olan ağaç türü ile korelasyonu birbirine yakın olarak elde edilmiştir ve bu korelasyonlar çok yüksek değildir (0.24 ile -0.22 arasında değişen korelasyonlar). 8 sınıflandırma özelliğinin korelasyon matrisine göre geometri tabanlı “Nokta sayısı”, “Maximum Z”, “Ortalama Z” ve “Z aralığı” özelliklerinin birbirleriyle yüksek korelasyonlu olduğu göze çarpmaktadır (korelasyon 0.64 ve üzeri). Yoğunluk tabanlı özelliklerin (“Ortalama yoğunluk” ve “Yoğunluğun standart sapması) de birbirleriyle korelasyonu 0.58 olarak elde edilmiştir. Test alanı 1’de 8 özelliğin birbirlerine göre ilişkileri incelendiğinde genel olarak geometri tabanlı ve yoğunluk tabanlı özelliklerin birbirleri ile düşük korelasyonlu olduğu sonucuna varılabilir.

Test alanı 1’de 25 sınıflandırma özelliğinin birbirlerine göre ilişkileri incelendiğinde 8 sınıflandırma özelliğine benzer şekilde tüm sınıflandırma özelliklerinin hedef sınıf olan ağaç türü ile korelasyonu birbirine yakın olarak elde edilmiş ve bu korelasyonlar burada da çok yüksek değildir (0.29 ile -0.22 arasında değişen korelasyonlar). Geometri tabanlı özellikler kendi içlerinde hem yüksek

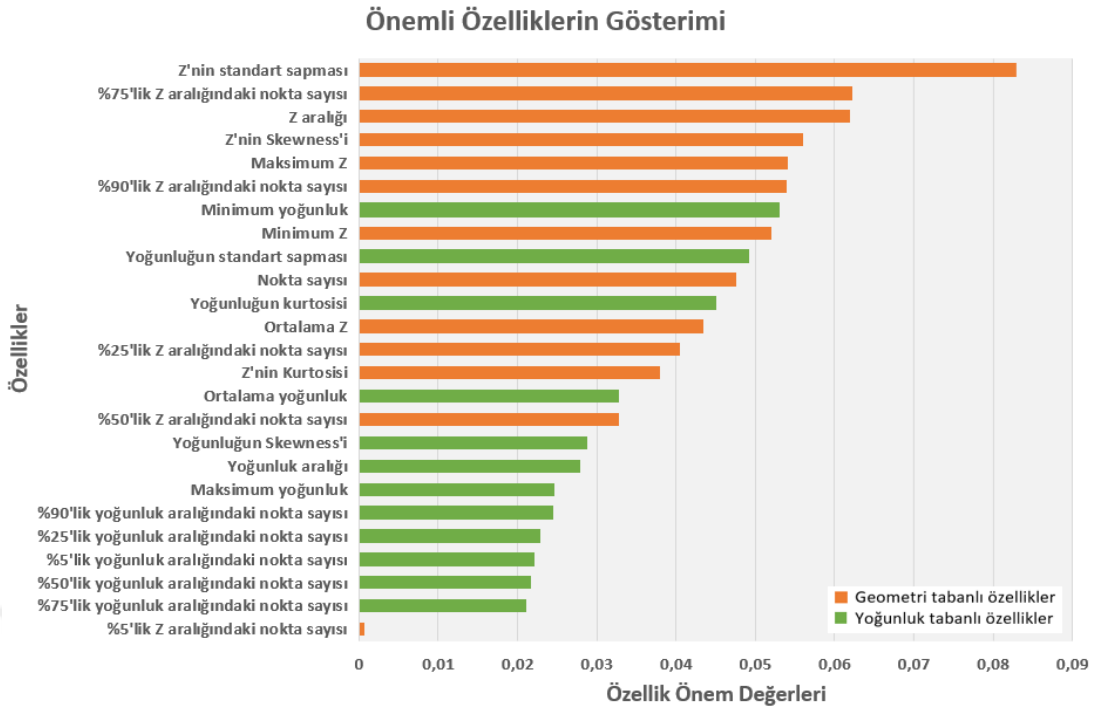
korelasyonlu hem de düşük korelasyonlu özellikleri barındırmaktadır ve yoğunluk tabanlı özelliklerin kendi içinde de benzer durum söz konusudur. Test alanı 1’de 8 özelliğin aralarındaki korelasyondan farklı olarak 25 özelliğe bulunan bazı geometri tabanlı özellikler ile yoğunluk tabanlı özellikler birbirleri ile yüksek korelasyonlu olarak belirlenmiştir (örneğin “%50’lik Z aralığındaki nokta sayısı”, “%75’lik Z aralığındaki nokta sayısı” ve “%90’lık Z aralığındaki nokta sayısı” ile “Yoğunluk aralığı”, “%5’lik yoğunluk aralığındaki nokta sayısı”, “%25’lik yoğunluk aralığındaki nokta sayısı”, “%50’lik yoğunluk aralığındaki nokta sayısı”, “%75’lik yoğunluk aralığındaki nokta sayısı”, “%90’lık yoğunluk aralığındaki nokta sayısı” ve “Maksimum yoğunluk” yüksek korelasyonlu özelliklerdir).

#### 4.7.5.2 Test Alanı 2’de Önemli Özelliklerin Belirlenmesi

Test alanı 2’de RO sınıflandırması ile ortalama Gini’deki azalmaya (Mean Decrease in Gini - MDG) göre belirlenen 8 sınıflandırma özelliğinin önemleri Şekil 4.73 ve 25 sınıflandırma özelliğinin önemleri ise Şekil 4.74’de verilmiştir.



**Şekil 4.73** Test alanı 2’de RO sınıflandırmasındaki ortalama Gini’deki azalmaya göre 8 sınıflandırma özelliğinin önemleri (feature importance scores)



**Şekil 4.74** Test alanı 2’de RO sınıflandırmasındaki ortalama Gini’deki azalmaya göre 25 sınıflandırma özelliğinin önemleri (feature importance scores)

Şekil 4.73’deki özellik önemleri incelendiğinde Test alanı 2’de 8 özellik arasında yer alan en önemli 3 sınıflandırma özelliği sırasıyla Test alanı 1’deki en önemli 3 özellekle birebir aynı aynıdır. Yani Test alanı 2’de de 8 özellekten en önemli sınıflandırma özelliği de “Nokta sayısı”dır. “Nokta sayısı”nı “Z aralığı”, “Yoğunluğun standart sapması” ve “Maksimum Z” özellikleri takip etmektedir. Test alanı 2’de 8 özellekten en az önemli sınıflandırma özelliğinin ise “Z’nin kurtosisi” olduğu görülmektedir. Test alanı 2’de, Test alanı 1’deki duruma benzer şekilde RO sınıflandırmasındaki ortalama Gini’deki azalmaya göre, 8 özellekten geometri tabanlı özelliklerin yoğunluktan tabanlı özelliklerden daha önemli olduğu ifade edilebilir.

Şekil 4.74’de yer alan Test alanı 2’deki özellik önemleri incelendiğinde, 25 özellik arasında en önemli sınıflandırma özelliğinin “Z’nin standart sapması”, en az önemli olan sınıflandırma özelliğinin ise Test alanı 1’de olduğu gibi “%5’lik Z aralığındaki nokta sayısı” olduğu görülmektedir. “%75’lik Z aralığındaki nokta sayısı”, “Z aralığı”, “Z’nin Skewness’i”, “Maksimum Z” ve “%90’lık Z aralığındaki nokta sayısı”, “Z’nin standart sapması” dan sonraki en önemli beş özellik olarak belirlenmiştir. Test alanı 2’de 25 özelliğın önemleri özellekle en iyi özelliklerde

Test alanı 1’dekenden farklı olsa da genel olarak RO sınıflandırmasında ortalama Gini’deki azalmaya göre önerilen 3B nokta tabanlı sınıflandırma için geometrik özelliklerin yoğunluk tabanlı özelliklerden çok daha yüksek bir öneme sahip olduğu burada da açıkça görülmektedir.

Özellik önemlerine ek olarak özelliklerin birbirleriyle ilişkilerinin analiz edildiği “Pearson” korelasyon matrisleri de Test alanı 2’de 8 sınıflandırma özelliği ve 25 sınıflandırma özelliği için sırasıyla Şekil 4.75 ve Şekil 4.76’da yer almaktadır.



Şekil 4.75 Test alanı 2’de 8 sınıflandırma özelliğinin Pearson korelasyon matrisi



Şekil 4.76 Test alanı 2’de 25 sınıflandırma özelliğinin Pearson korelasyon matrisi

Test alanı 2’de 8 sınıflandırma özelliğinin birbirlerine göre ilişkileri incelendiğinde Test alanı 1’deki duruma benzer şekilde tüm sınıflandırma özelliklerinin hedef sınıf olan ağaç türü ile korelasyonu birbirine yakın olarak elde edilmiştir ve bu korelasyonlar çok yüksek değildir (0.29 ile -0.04 arasında değişen korelasyonlar). Test alanı 2’de de 8 sınıflandırma özelliğinin korelasyon matrisine göre geometri tabanlı “Nokta sayısı”, “Maximum Z”, “Ortalama Z” ve “Z aralığı” özelliklerinin birbirleriyle yüksek korelasyonlu olduğu göze çarpmaktadır (korelasyon 0.65 ve üzeri). Yoğunluk tabanlı özelliklerin (“Ortalama yoğunluk” ve “Yoğunluğun standart sapması) de birbirleriyle korelasyonu 0.65 olarak belirlenmiştir. Test alanı 2’de 8 özelliğinin birbirlerine göre korelasyonu dikkate alındığında genel olarak geometri tabanlı ve yoğunluk tabanlı özelliklerin birbirleri ile düşük korelasyonlu olduğu sonucuna varılabilir.

Test alanı 2’de 25 özelliğinin birbirleri ile ilişkileri incelendiğinde tüm sınıflandırma özelliklerinin hedef sınıf olan ağaç türü ile korelasyonu birbirine yakın olarak belirlenmiş ve bu korelasyonların burada da çok yüksek olmadığı söylenebilir

(0.31 ile -0.12 arasında deęişen korelasyonlar). Geometri tabanlı özellikler ve yoğunluk tabanlı özelliklerin kendi aralarında tutarlı bir korelasyon söz konusu değildir. Geometri tabanlı özellikler kendi içlerinde hem yüksek yoğunluklu hem de düşük korelasyonlu özellikleri barındırmaktadır ve yoğunluk tabanlı özelliklerin de kendi içinde benzer durum söz konusudur. Test alanı 2’de 25 özellikte bulunan bazı geometri tabanlı özellikler ile yoğunluk tabanlı özellikler birbirleri ile yüksek korelasyonlu olarak elde edilmiştir (örneğin “%50’lik Z aralığındaki nokta sayısı”, “%75’lik Z aralığındaki nokta sayısı” ve “%90’lık Z aralığındaki nokta sayısı” ile “%5’lik yoğunluk aralığındaki nokta sayısı”, “%25’lik yoğunluk aralığındaki nokta sayısı”, “%50’lik yoğunluk aralığındaki nokta sayısı”, “%75’lik yoğunluk aralığındaki nokta sayısı” ve “%90’lık yoğunluk aralığındaki nokta sayısı” korelasyonludur).

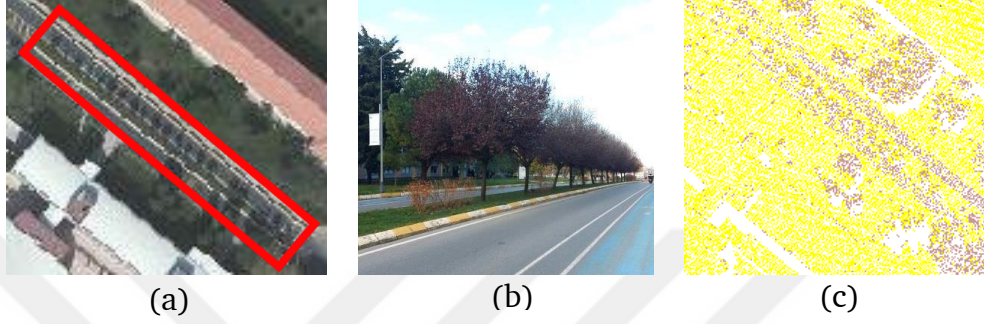
Sonuçta özellik önemlerine bakılarak, makine öğrenme tabanlı algoritmalar ile kentsel ağaç türlerinin sınıflandırılması uygulamalarında farklı test alanlarında da aynı algoritma ve aynı parametreler kullanıldığı durumda önemli özelliklerin birbirine yakın olduğu çıkarımı yapılabilir. Benzer şekilde farklı test alanlarında da aynı algoritma ve aynı parametreler kullanıldığı durumda tüm özellikler arasındaki korelasyonların birbirine çok benzer olarak elde edildiği çıkarımı yapılabilir.

#### **4.8 Tekil Cadde Ağaçlarının Makine Öğrenme Tabanlı Segmentasyon Yöntemleri ile Otomatik Belirlenmesi**

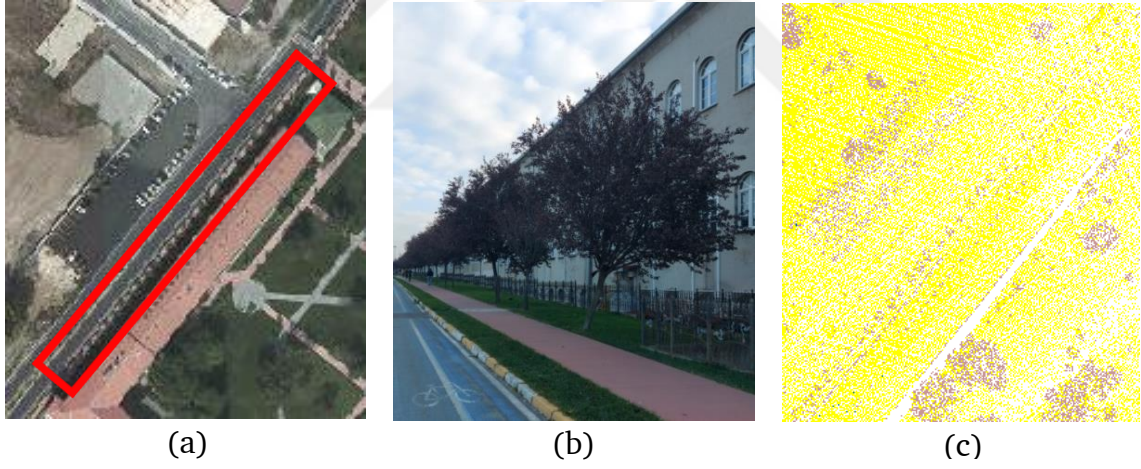
Bu çalışmada, makine öğrenme algoritmaları kullanılarak LiDAR verilerinden 3B nokta tabanlı sınıflandırma ile kentsel geniş ve iğne yapraklı ağaç türlerinin elde edilmesi uygulamasına ek olarak tekil cadde ağaçlarının kentsel ağaç taçlarının elde edilmesi için önerilen makine öğrenme tabanlı segmentasyon yöntemleri ile elde edilmesi de araştırılmıştır. Bu bölümde tekil cadde ağaçlarının elde edilebilmesi için belirlenen çalışma alanı, genel işlem adımları, nokta tabanlı sınıflandırma ve segmentasyon aşamaları ve doğruluk analizi çalışmalarına yer verilecektir.

#### 4.8.1 Test Alanı ve Genel İşlem Adımları

LiDAR verilerinden tekil cadde ağaçlarının otomatik olarak elde edilebilmesi için ağaç türlerinin sınıflandırmasına benzer olarak çalışma alanında iki farklı test alanı belirlenmiştir. Test alanı A ve Test alanı B olarak isimlendirilen cadde ağaçlarının yer aldığı test alanları sırasıyla Şekil 4.77 ve Şekil 4.78'de yer almaktadır.

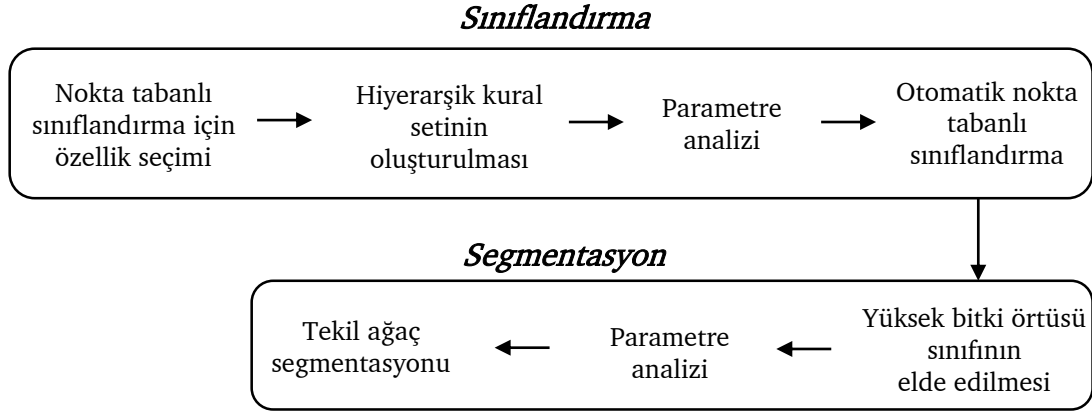


**Şekil 4.77** Test alanı A: Şehir haritası (2013) (a), cadde ağaçları (b) ve yoğunluğa göre renklendirilmiş ham LiDAR nokta bulutu (c)



**Şekil 4.78** Test alanı B: Şehir haritası (2013) (a), cadde ağaçları (b) ve yoğunluğa göre renklendirilmiş ham LiDAR nokta bulutu (c)

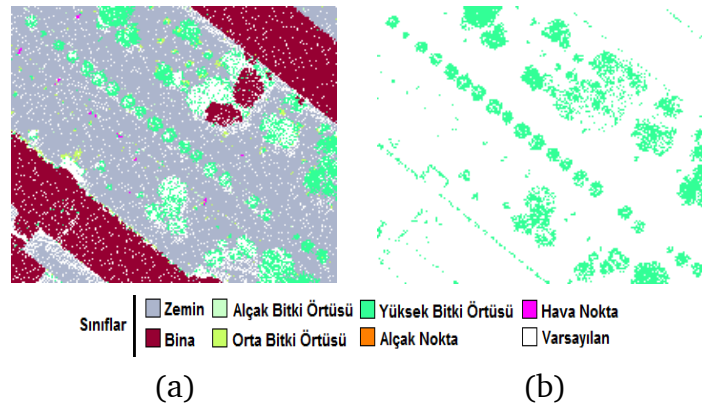
Belirlenen test alanlarında tekil cadde ağaçlarının belirlenmesi için nokta tabanlı işlem adımları sınıflandırma ve segmentasyon olmak üzere iki bölüme ayrılır (Şekil 4.79).



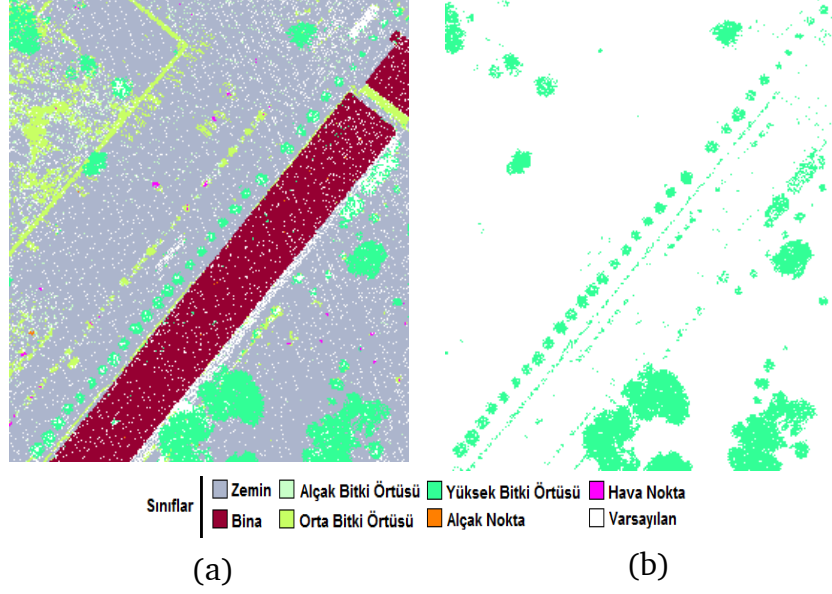
**Şekil 4.79** Tekil cadde ağaçlarının elde edilebilmesi için belirlenen iş akış şeması

#### 4.8.2 Yüksek Bitki Örtüsünün Nokta Tabanlı Sınıflandırılması

Bu çalışmada, test alanlarının LiDAR nokta bulutu verilerinin daha önce de yüksek bitki örtüsü sınıfını elde etmek için kullanılan otomatik hiyerarşik kural tabanlı sınıflandırma yöntemi ile sınıflandırılması gerçekleştirilmiştir. Nokta tabanlı sınıflandırma için kentsel ağaç türlerinin elde edilmesi aşamasında belirlenen hiyerarşik kural seti (Tablo 4.3) (en sondaki çoklu dönüşe göre kuralı hariç) ve en uygun parametreler kullanılarak 3B nokta tabanlı sınıflandırma ile Test alanı A ve Test alanı B’de zemin, bina, alçak bitki örtüsü, orta bitki örtüsü, yüksek bitki örtüsü, alçak nokta, hava nokta ve varsayılan sınıfları elde edilmiştir (Şekil 4.80a ve Şekil 4.81a). Nokta tabanlı sınıflandırma sonucunda elde edilen tüm yüksek bitki örtüsü noktaları, diğer arazi sınıflarının noktalarından ayrılmıştır (Şekil 4.80b ve Şekil 4.81b). LiDAR nokta bulutunun sınıflandırma işlemi, burada da “Terrasolid” yazılımının “TerraScan” modülü kullanılarak gerçekleştirilmiştir.



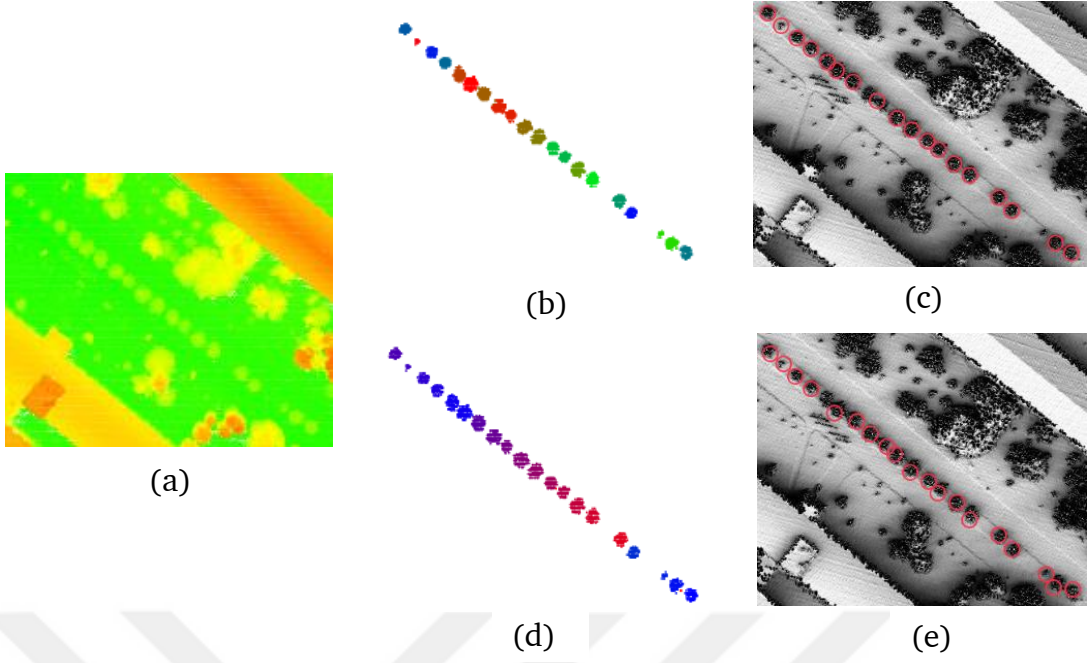
**Şekil 4.80** Test alanı A’nın nokta tabanlı sınıflandırılması (a) ve Test alanı A’da bulunan yüksek bitki örtüsü noktaları (b)



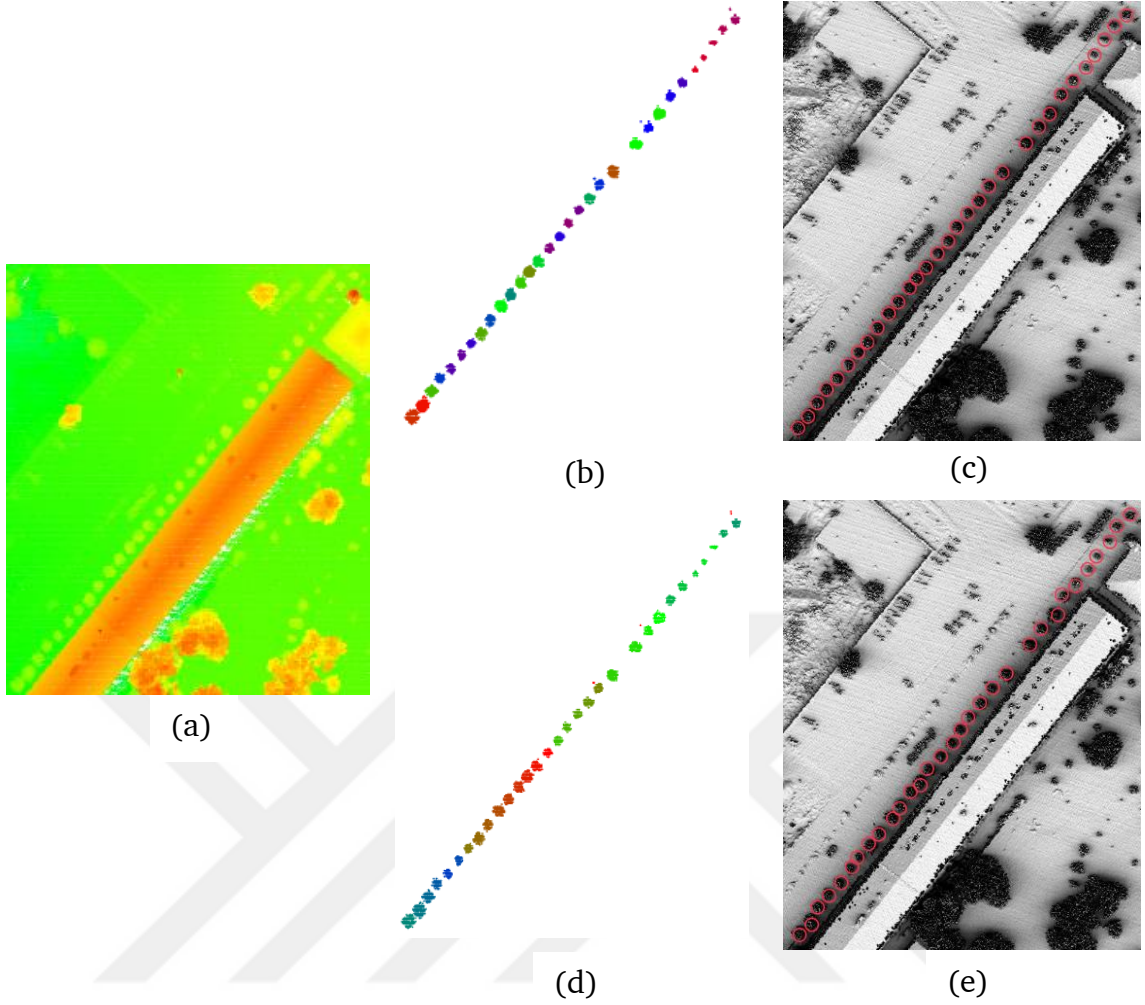
**Şekil 4.81** Test alanı B'nin nokta tabanlı sınıflandırılması (a) ve Test alanı B'de bulunan yüksek bitki örtüsü noktaları (b)

#### 4.8.3 Tekil Cadde Ağaçlarının Nokta Tabanlı Segmentasyonu ve Doğruluk Analizi

Test alanı A ve Test alanı B'deki yüksek bitki örtüsü sınıfındaki cadde ağaçlarının segmentasyonu için de kentsel ağaç taçlarının elde edilme aşamasında olduğu gibi makine öğrenme tabanlı “MeanShift” ve “DBSCAN” kümeleme algoritmaları kullanılmıştır. Daha önce kentsel ağaç taçlarının belirlenmesi için gerçekleştirilen parametre analizlerinde elde edilen parametreler (Bölüm 4.5.1 ve Bölüm 4.5.3) kullanılarak yine “Jupyter Notebook (açık kaynak kodlu bir web uygulaması)” ortamında “Python programlama dili (Python 3.6.4)” kullanılarak 2B ağaç segmentasyon işlemi gerçekleştirilmiştir. “MeanShift” ve “DBSCAN” kümeleme ile Test alanı A ve Test alanı B'deki 2B tek ağaç segmentasyonu sonuçları sırasıyla Şekil 4.82 ve Şekil 4.83'de yer almaktadır. Segmentasyon sonuçları analiz edildiğinde, hem “MeanShift” hem de “DBSCAN” kümeleme algoritmaları kullanılarak nokta tabanlı segmentasyon ile tekil sokak ağaçlarının başarılı bir şekilde belirlendiği açıkça görülmektedir. “MeanShift” ve “DBSCAN” kümeleme algoritmaları ile gerçekleştirilen segmentasyon sonuçları, Test alanı A'da birbirine oldukça yakın ve Test alanı B'de ise tamamen aynıdır.



**Şekil 4.82** Test alanı A'nın yüksekliğe göre renklendirilmiş ham nokta bulutu (a), "MeanShift" kümeleme ile segmente edilmiş cadde ağaçları (b), "MeanShift" kümeleme ile segmente edilmiş cadde ağaçlarının gri tonlu yüzey modeli ile üst üste bindirilmesi (c), "DBSCAN" kümeleme ile segmente edilmiş cadde ağaçları (d), "DBSCAN" kümeleme ile segmente edilmiş cadde ağaçlarının gri tonlu yüzey modeli ile üst üste bindirilmesi (e)

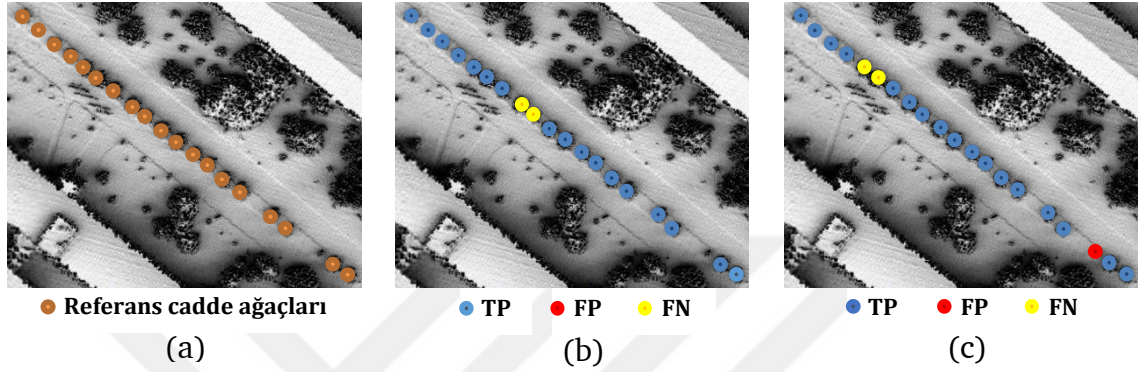


**Şekil 4.83** Test alanı B'nin yüksekliğe göre renklendirilmiş ham nokta bulutu (a), “MeanShift” kümeleme ile segmente edilmiş cadde ağaçları (b), “MeanShift” kümeleme ile segmente edilmiş cadde ağaçlarının gri tonlu yüzey modeli ile üst üste bindirilmesi (c), “DBSCAN” kümeleme ile segmente edilmiş cadde ağaçları (d), “DBSCAN” kümeleme ile segmente edilmiş cadde ağaçlarının gri tonlu yüzey modeli ile üst üste bindirilmesi (e)

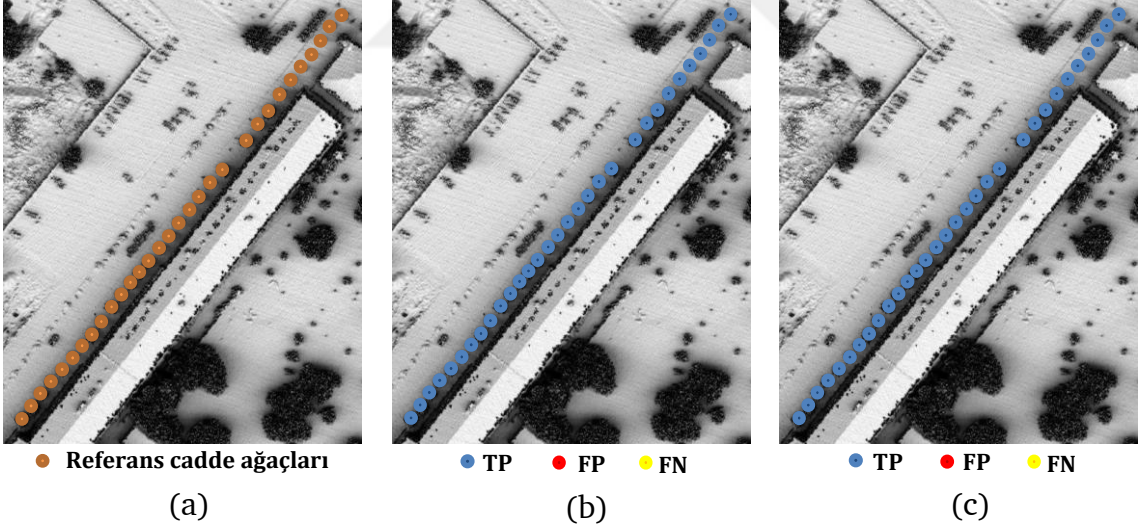
Makine öğrenme tabanlı “MeanShift” ve “DBSCAN” kümeleme ile 2B segmentasyon sonucu elde edilen tekil cadde ağaçları için doğruluk analizi, belirlenme oranına (detection rate) göre bütünlük ve doğruluk analizleri kullanılarak gerçekleştirilmiştir. Test alanı A'da, hem “MeanShift” hem de “DBSCAN” kümeleme algoritmaları kullanılarak 17 küme tekil cadde ağaçları olarak doğru (TP) bir şekilde belirlenmiştir (Şekil 4.84). “MeanShift” kümeleme ile yanlış kümelenmiş hiçbir cadde ağacı bulunmazken, “DBSCAN” kümeleme ile sadece 1 küme tekil cadde ağacı (FP) olarak yanlış elde edilmiştir.

Ayrıca 2 cadde ağacının noktaları da, hem “MeanShift” hem de “DBSCAN” kümeleme ile ağaç kümeleri (FN) olarak segmente edilememiştir.

Test alanı B’de, 31 kümenin tümü, hem “MeanShift” hem de “DBSCAN” kümeleme algoritmaları kullanılarak doğru bir şekilde (TP) tekil cadde ağaçları olarak belirlenmiştir (Şekil 4.85). Test alanı B’de yanlış belirlenmiş (FP) veya segmente edilemeyen (FN) hiçbir cadde ağacı bulunmamaktadır.



**Şekil 4.84** Test alanı A’da referans ağaçlar (a) “MeanShift” kümeleme ile elde edilmiş TP, FP, FN cadde ağaçları (b) ve “DBSCAN” kümeleme ile elde edilmiş TP, FP, FN cadde ağaçları (c)



**Şekil 4.85** Test alanı B’de referans ağaçlar (a) “MeanShift” kümeleme ile elde edilmiş TP, FP, FN cadde ağaçları (b) ve “DBSCAN” kümeleme ile elde edilmiş TP, FP, FN cadde ağaçları (c)

Test alanı A’daki “MeanShift” kümeleme algoritması için bütünlük ve doğruluk değerleri sırasıyla %89.47 ve %100, “DBSCAN” kümeleme algoritması için bütünlük ve doğruluk değerleri de sırasıyla %89.47 ve %94.44 olarak elde edilmiştir (Tablo 4.42). Test alanı B’de ise, hem “MeanShift” hem de “DBSCAN”

kümeleme algoritmaları için tüm bütünlük ve doğruluk değerleri %100 olarak elde edilmiştir (Tablo 4.42). 2B nokta tabanlı segmentasyon sonuçları, hava LiDAR verilerini kullanarak tekil cadde ağaçlarının otomatik olarak elde edilmesi için önerilen “MeanShift” ve “DBSCAN” kümeleme yöntemlerinin başarısını açıkça ortaya koymaktadır. Otomatik segmentasyon sonuçları her iki kümeleme algoritmaları için de oldukça tatmin edicidir.

**Tablo 4.42** Test alanı A ve Test alanı B’de gerçekleştirilen “MeanShift” ve “DBSCAN” kümeleme algoritmalarının doğruluk değerleri

	“MeanShift” Kümeleme		“DBSCAN” kümeleme	
	Bütünlük	Doğruluk	Bütünlük	Doğruluk
<b>Test alanı A</b>	%89.47	%100	%89.47	%94.44
<b>Test alanı B</b>	%100	%100	%100	%100

# 5

## SONUÇ VE ÖNERİLER

---

Bu tez çalışmasında, afet yönetimi, şehir planlaması, çevre koruma veya kentsel kalkınma politikası üretimi gibi bir farklı çalışmalarda kullanılabilmesi için makine öğrenme algoritmaları ile lazer tarayıcı verilerinden kentsel ağaç türlerinin sınıflandırılması olanakları araştırılmıştır. Destek vektör makineleri (DVM), rastgele orman (RO) ve çok katmanlı algılayıcı (ÇKA) sınıflandırma algoritmaları kullanılarak ham 3B LiDAR verilerinden kentsel çalışma alanında geniş ve iğne yapraklı ağaç türlerinin elde edilmesi amaçlanmıştır.

Bu amaçla, kentsel çalışma alanı olarak seçilen Yıldız Teknik Üniversitesi, Davutpaşa Kampüsü'nde küçük bir alanı içeren Test alanı 1 ve daha büyük bir alanı içeren Test alanı 2 (Test alanı 1'in tamamını da kapsayan) ağaç türlerinin sınıflandırılması için seçilmiştir. Öncelikle ham LiDAR verisinden hiyerarşik kural tabanlı sınıflandırma yaklaşımı ile Test alanı 1 ve Test alanı 2'de yeryüzeyini ifade eden zemin, bina, alçak bitki örtüsü, orta bitki örtüsü, yüksek bitki örtüsü, alçak nokta, hava nokta ve varsayılan sınıfları elde edilmiştir. Ağaç türlerinin sınıflandırılmasında kullanılmak üzere yüksek bitki örtüsü sınıfı noktaları diğer noktalardan ayrılmış ve bireysel ağaç taçlarının elde edilmesi için makine öğrenme tabanlı üç farklı kümeleme algoritması ("MeanShift", "K-means", "Density-Based Spatial Clustering of Applications with Noise - DBSCAN") kullanılarak 2B nokta tabanlı segmentasyon işlemi gerçekleştirilmiştir. En iyi segmentasyon sonucunu veren "MeanShift" kümeleme yaklaşımı ile elde edilen bireysel ağaç taçları, makine öğrenme tabanlı sınıflandırma algoritmalarının eğitim ve test aşamalarında kullanılmak üzere referans veri kullanılarak etiketlenmiştir. Ağaç türlerinin sınıflandırmalarında kullanılmak üzere geometri ve yoğunluk (intensity) tabanlı özelliklerden oluşan iki farklı sınıflandırma özellik seti (8 özellik ve 25 özellik) oluşturulmuş ve DVM, RO ve ÇKA sınıflandırma yöntemleri için "Grid search" yaklaşımı ile en uygun parametreler belirlenmiştir. Eğitim aşamalarının ardından (%70 eğitim, %30 test verisi) iki farklı sınıflandırma özellik

seti kullanılarak destek vektör makineleri, rastgele orman ve çok katmanlı algılayıcı algoritmaları ile test alanlarındaki (Test alanı 1 ve Test alanı 2) bireysel ağaç tacı segmentleri geniş yapraklı ve iğne yapraklı ağaç türlerine sınıflandırılmıştır. Takip eden paragraflarda, öncelikle DVM, RO ve ÇKA sınıflandırma yöntemleri ile Test alanı 1’de 8 özellik ve 25 özellik kullanılarak elde edilen sınıflandırma sonuçları ve sonrasında Test alanı 2’de 8 özellik ve 25 özellik kullanılarak elde edilen sınıflandırma sonuçları analiz edilmiştir.

Test alanı 1’de 8 sınıflandırma özelliği kullanılarak gerçekleştirilen geniş ve iğne yapraklı ağaçların Destek Vektör Makinesi (DVM) sınıflandırmasına göre genel doğruluk %76 olarak elde edilmiştir. Geniş yapraklı ağaçlar için hassasiyet, kesinlik ve F1-skor değerleri sırasıyla 0.7925, 0.84 ve 0.8156; iğne yapraklı ağaçlar için ise bu değerler sırasıyla 0.7037, 0.6333 ve 0.6667 olarak hesaplanmıştır. Test alanı 1’in 25 sınıflandırma özelliği kullanılarak gerçekleştirilen DVM sınıflandırmasına göre ise genel doğruluk %80 olarak belirlenmiştir. Hassasiyet, kesinlik ve F1-skor değerleri geniş yapraklı ağaçlar için sırasıyla 0.8868, 0.8246 ve 0.8545; iğne yapraklı ağaçlar için ise sırasıyla 0.6296, 0.7391 ve 0.68’dir. Buradan hareketle makine öğrenme tabanlı DVM algoritmasının genel olarak Test alanı 1’de geniş ve iğne yapraklı ağaçları başarılı bir şekilde sınıflandırdığı değerlendirilmiştir. 8 ve 25 özelliğin kullanıldığı nokta tabanlı sınıflandırmalar ayrı ayrı incelendiğinde ise, 25 sınıflandırma özelliğinin kullanıldığı durumda sınıflandırma doğruluklarının daha yüksek olduğu görülmektedir. Geniş ve iğne yapraklı ağaç türlerinin bireysel sınıflandırma doğrulukları incelendiğinde de 25 özelliğin kullanıldığı durumda özellikle geniş yapraklı ağaç türünün sınıflandırılma doğruluğunun 8 özelliğin kullanıldığı durumdan daha yüksek olduğu ifade edilebilir. Geniş ve iğne yapraklı ağaç türlerinin birbirlerine göre sınıflandırma doğrulukları karşılaştırıldığında ise genel olarak her iki test alanında da (Test alanı 1 ve Test alanı 2) geniş yapraklı ağaçlar iğne yapraklı ağaçlardan daha başarılı bir şekilde sınıflandırılmıştır.

Rastgele Orman (RO) algoritması ile Test alanı 1’de 8 sınıflandırma özelliği kullanıldığında genel sınıflandırma doğruluğu %80 olarak elde edilmiştir. Geniş yapraklı ağaçlar için hassasiyet değeri 0.9245, kesinlik değeri 0.8033 ve F1-skor değeri 0.8596 iken iğne yapraklı ağaçlar için bu değerler sırasıyla 0.5556, 0.7895

ve 0.6522 olarak belirlenmiştir. Test alanı 1'de 25 sınıflandırma özelliğinin kullanıldığı RO sınıflandırmasının genel doğruluğu %84 olarak elde edilmiştir. Hassasiyet, kesinlik ve F1-skor değerleri geniş yapraklı ağaçlar için sırasıyla 0.9434, 0.8333 ve 0.8849; iğne yapraklı ağaçlar için sırasıyla 0.6296, 0.85 ve 0.7234 olarak hesaplanmıştır. Buna göre RO makine öğrenme algoritmasının kentsel Test alanı 1'de geniş ve iğne yapraklı ağaçları başarılı bir şekilde türlerine ayırabildiği söylenebilir. 25 özellikli RO algoritmasının sınıflandırma sonuçları, 8 özellikli RO algoritmasından %4 daha başarılı sonuç üretmiştir. Geniş ve iğne yapraklı ağaçların ayrı ayrı sınıflandırma doğruluklarına göre de yine 25 özelliğin kullanıldığı RO algoritmasının her iki ağaç türünü 8 özelliğin kullanıldığı sınıflandırmadan daha iyi sınıflandırdığı görülmüştür. Test alanı 1'de gerçekleştirilen 8 özellik ve 25 özelliğin kullanıldığı her iki sınıflandırma sonucunda da geniş yapraklı ağaçlar genel olarak RO algoritması ile iğne yapraklı ağaçlardan çok daha başarılı bir şekilde sınıflandırılmışlar.

Test alanı 1'de 8 sınıflandırma özelliğinin kullanıldığı Çok Katmanlı Algılayıcı (ÇKA) algoritması ile gerçekleştirilen sınıflandırma sonucu %68 olarak elde edilmiştir. Geniş yapraklı ağaçlar için hassasiyet değeri 0.9811, kesinlik değeri 0.6753 ve F1-skor değeri 0.80 olarak hesaplanırken, iğne yapraklı ağaçlar için bu değerler sırasıyla 0.0741, 0.6667, 0.1334 olarak hesaplanmıştır. ÇKA sınıflandırıcısı ile Test alanı 1'de 25 özelliğin kullanıldığı geniş ve iğne yapraklı ağaçların genel sınıflandırma doğruluğu %74'tür. Geniş yapraklı ağaçlar için hassasiyet değeri 0.9245, kesinlik değeri 0.7424 ve F1-skor değeri 0.8235; iğne yapraklı ağaçlar için bu değerler sırasıyla 0.3704, 0.7143 ve 0.4878 olarak hesaplanmıştır. Bu sonuçlara göre, ÇKA sınıflandırmasının Test alanı 1'de geniş ve iğne yapraklı ağaçların sınıflandırılmasında çok yüksek doğruluğa ulaşamadığı görülmüştür. Özellikle 8 sınıflandırma özelliğinin kullanıldığı ÇKA sınıflandırıcısının sınıflandırma doğruluğu geniş ve iğne yapraklı ağaçların sınıflandırılması için tatmin edici değildir. 25 özelliğin kullanıldığı ÇKA sınıflandırma sonucu 8 özelliğe göre daha yüksektir. Geniş ve iğne yapraklı ağaç türlerinin ayrı ayrı sınıflandırma doğrulukları incelendiğinde, 25 sınıflandırma özelliğinin kullanıldığı durumda her iki ağaç türü de 8 özelliğin kullanıldığı durumdan daha başarılı sınıflandırılmışlardır. Geniş ve iğne yapraklı ağaçların

birbirlerine göre sınıflandırma doğruluklarına bakılacak olursa her iki özellik setinin (8 özellik ve 25 özellik) de kullanıldığı ÇKA sınıflandırmalarında geniş yapraklı ağaçlar iğne yapraklı ağaçlardan çok daha başarılı bir şekilde elde edilmiştir. ÇKA sınıflandırma sonucunda özellikle hassasiyet ve F1-skor değerlerine göre iğne yapraklı ağaçların elde edilme başarısı oldukça düşüktür.

Test alanı 2'de 8 özelliğin kullanıldığı DVM sınıflandırmasının genel doğruluğu %72 olarak elde edilmiştir. 8 özellik ile gerçekleştirilen sınıflandırmada geniş yapraklı ağaçlar için hassasiyet, kesinlik ve F-skor değerleri sırasıyla 0.9060, 0.7571 ve 0.8249; iğne yapraklı ağaçlar için ise bu değerler sırasıyla 0.1905, 0.4211 ve 0.2623'tür. Test alanı 2'de 25 özelliğin kullanıldığı DVM sınıflandırmasının genel doğruluğu da %72 olarak belirlenmiştir. Hassasiyet, kesinlik ve F1-skor değerleri 25 özelliğin kullanıldığı durumda sırasıyla 0.9316, 0.7517 ve 0.8320; iğne yapraklı ağaçlar için ise sırasıyla 0.1429, 0.4286 ve 0.2143'tür. Bu sonuçlara göre DVM sınıflandırıcısının Test alanı 2'de geniş ve iğne yapraklı ağaçları ortalama bir doğrulukla sınıflandırdığı ifade edilebilir. Sınıflandırma özellik sayısının değişmesi Test alanı 2'de DVM algoritmasının sınıflandırma sonuçlarını değiştirmemiştir. Ağaç türlerinin bireysel sınıflandırma doğrulukları incelendiğinde, geniş ve iğne yapraklı ağaçların her iki özellik setinde de (8 özellik ve 25 özellik) sınıflandırma doğruluklarının birbirine oldukça yakın olduğu ifade edilebilir. İki ağaç türünün birbirlerine göre sınıflandırma doğrulukları analiz edildiğinde ise geniş yapraklı ağaçlar hem 8 özellikli hem de 25 özellikli DVM sınıflandırmasında iğne yapraklı ağaçlardan çok daha iyi bir şekilde elde edilmiştir.

Test alanı 2'de 8 sınıflandırma özelliği kullanılarak gerçekleştirilen makine öğrenme tabanlı RO algoritması ile toplam sınıflandırma doğruluğu %70 olarak belirlenmiştir. Bu sınıflandırmada geniş yapraklı ağaçlar için hassasiyet, kesinlik ve F1-skor değerleri sırasıyla 0.8547, 0.7634 ve 0.8065; iğne yapraklı ağaçlar için ise sırasıyla 0.2619, 0.3929 ve 0.3143'tür. 25 özelliğin kullanıldığı RO sınıflandırıcısı ile Test alanı 2'de ağaç türlerinin sınıflandırılma doğruluğu %72'dir. RO sınıflandırıcı ile 25 özellik kullanıldığında hassasiyet, kesinlik ve F1-skor değerleri sırasıyla 0.8889, 0.7704 ve 0.8254; iğne yapraklı ağaçlar için ise sırasıyla 0.2619, 0.4583 ve 0.3333 olarak hesaplanmıştır. Test alanı 2'de nokta

tabanlı RO sınıflandırma sonuçlarına göre geniş ve iğne yapraklı ağaçlar ortalama bir doğruluk ile sınıflandırılmıştır. 25 özelliğin kullanıldığı RO sınıflandırma sonuçları 8 özelliğin kullanıldığı sınıflandırmaya göre %2 daha başarılıdır. Geniş ve iğne yapraklı ağaçların ayrı ayrı sınıflandırılmasında yine 25 özellikli RO algoritması, 8 özelliğin kullanıldığı RO sınıflandırmasına göre her iki ağaç türünü de biraz daha yüksek doğrulukla elde etmiştir. Test alanı 2'de ağaç türlerinin birbirlerine göre sınıflandırma doğrulukları karşılaştırıldığında, geniş yapraklı ağaçların her iki özellik seti (8 özellik ve 25 özellik) kullanılarak gerçekleştirilen sınıflandırmada iğne yapraklı ağaçlardan çok daha iyi bir şekilde sınıflandırıldığı görülmüştür.

Makine öğrenme tabanlı ÇKA sınıflandırıcısı ile 8 özellik kullanılarak Test alanı 2'de gerçekleştirilen sınıflandırma sonucu %72 olarak belirlenmiştir. 8 özellikli ÇKA ile geniş yapraklı ağaçlar için elde edilen hassasiyet, kesinlik ve F1-skor değerleri sırasıyla 0.9744, 0.7355 ve 0.8383; iğne yapraklı ağaçlar için elde edilen hassasiyet, kesinlik ve F1-skor değerleri ise sırasıyla 0.0238, 0.25 ve 0.3143 olarak hesaplanmıştır. Test alanı 2'de 25 özellik kullanılarak gerçekleştirilen ÇKA sınıflandırmasının genel doğruluğu %47'dir. Bu sınıflandırma da geniş yapraklı ağaçlar için hassasiyet değeri 0.3162, kesinlik değeri 0.8810 ve F1-skor değeri 0.4654 iken iğne yapraklı ağaçlar için bu değerler sırasıyla 0.8810, 0.3162 ve 0.4654'tür. Buradaki sonuçlara göre, 8 özelliğin kullanıldığı durumda ÇKA ile geniş ve iğne yapraklı ağaçlar ortalama bir doğrulukla elde edilirken 25 özelliğin kullanıldığı ÇKA ile geniş ve iğne yapraklı ağaçlar ise oldukça düşük bir doğrulukla sınıflandırılmışlardır. Test alanı 2'de geniş ve iğne yapraklı ağaçların ayrı ayrı sınıflandırma doğrulukları incelendiğinde 25 özelliğin kullanıldığı durumda geniş yapraklı ağaçlar 8 özelliğin kullanıldığı durumdan daha yüksek bir şekilde sınıflandırılmışlardır. Bunun tersi olarak 8 özelliğin kullanıldığı durumda iğne yapraklı ağaçlar 25 özelliğin kullanıldığı duruma göre daha iyi sınıflandırılmışlardır. Geniş ve iğne yapraklı ağaçların bireysel sınıflandırma doğrulukları karşılaştırıldığında 8 özelliğin kullanıldığı durumda geniş yapraklı ağaçların iğne yapraklı ağaçlardan çok daha iyi bir şekilde sınıflandırıldığı görülmektedir. 25 özelliğin kullanıldığı durumda ise her iki ağaç türünün sınıflandırılma doğruluğu birbirinin aynıdır.

Test alanı 1’de 8 sınıflandırma özelliğinin kullanıldığı makine öğrenme tabanlı DVM, RO ve ÇKA algoritmalarının nokta tabanlı olarak kentsel ağaç türlerini sınıflandırma doğrulukları analiz edildiğinde DVM ile RO algoritmalarının birbirlerine yakın (sırasıyla %76.25 ve %80), ÇKA’nın ise bu iki yöntemden daha düşük sınıflandırma doğruluğu (%67.50) ürettiği görülmüştür. Buna göre Test alanı 1’de 8 özelliğın kullanıldığı durumda RO algoritması en iyi sınıflandırma doğruluğunu veren, ÇKA ise en düşük sınıflandırma doğruluğuna veren sınıflandırma yöntemidir. Test alanı 1’de 25 sınıflandırma özelliğini kullanan DVM, RO ve ÇKA sınıflandırmalarının genel doğrulukları karşılaştırıldığında sırasıyla %80 ve %83.75 ile DVM ve RO algoritmaları, %73.75 ile ÇKA’dan daha yüksek doğrulukla sınıflandırma işlemi gerçekleştirmişlerdir. 25 özelliğın kullanıldığı sınıflandırmada da RO en iyi sonucu veren makine öğrenme tabanlı sınıflandırma algoritması iken ÇKA en kötü sınıflandırma sonucu veren makine öğrenme tabanlı sınıflandırma algoritmasıdır. Genel olarak Test alanı 1’de 25 özelliğın kullanıldığı tüm DVM, RO ve ÇKA sınıflandırma yöntemleri, 8 özelliğın kullanıldığı sınıflandırma yöntemlerinden daha doğru başarılı sınıflandırma gerçekleştirmişlerdir ve kentsel ağaç türlerinin elde edilebilmesi için RO algoritması en başarılı sınıflandırıcı, ÇKA ise en az başarılı sınıflandırıcı olmuştur.

Test alanı 2’de 8 sınıflandırma özelliğı kullanılarak gerçekleştirilen nokta tabanlı DVM, RO ve ÇKA sınıflandırma sonuçları analiz edildiğinde üç sınıflandırma yönteminin genel doğruluklarının (sırasıyla %71.70, %69.81 ve %72.33) birbirlerine oldukça yakın olduđu görülmüştür. Test alanı 1’deki durumun tersine, Test alanı 2’de 8 özelliğın kullanıldığı durumda ÇKA en başarılı sınıflandırıcı olmuşken, RO en başarısız sınıflandırıcı olmuştur. Test alanı 2’de 25 sınıflandırma özelliğı kullanılarak gerçekleştirilen DVM, RO ve ÇKA sınıflandırma sonuçları analiz edildiğinde ise DVM ve RO’ın genel sınıflandırma doğrulukları aynı (%72.33) ve ÇKA sınıflandırıcısının genel sınıflandırma doğruluğı (% 46.54) ise diğeri yaklaşımlara göre oldukça düşüktür. Buna göre 8 özelliğın kullanıldığı durumun tersine Test alanı 2’de 25 özelliğın kullanıldığı durumda DVM ve RO en başarılı sınıflandırıcılar iken, ÇKA en başarısız sınıflandırıcı olmuştur. Genel olarak Test alanı 2’de 25 özelliğın kullanıldığı DVM, RO ve ÇKA sınıflandırmalarında DVM ve RO algoritmaları 8 sınıflandırma özelliğının kullanıldığı durumdan daha

yüksek sonuçlar üretirken ÇKA sınıflandırıcısı için bu durum tamamen tersi olmuştur. 8 özellikli ÇKA sınıflandırma sonuçları 25 özellikli ÇKA sınıflandırma sonuçlarından daha başarılı olmuştur.

Test alanı 1’de RO sınıflandırmasındaki ortalama Gini’deki azalmaya (Mean Decrease in Gini - MDG) göre sınıflandırma özelliklerinin önemi analiz edildiğinde 8 özellik arasından en önemli sınıflandırma özelliğinin “Nokta sayısı”, en az önemli sınıflandırma özelliğinin ise “Maksimum Z” olduğu görülmüştür. Test alanı 2’de ise 25 özellik arasından en önemli sınıflandırma özelliği “%90’lık Z aralığındaki nokta sayısı”, en az önemli olan sınıflandırma özelliği ise “%5’lik Z aralığındaki nokta sayısı”dır. Test alanı 1’de her iki özellik seti için (8 özellik ve 25 özellik) genel olarak ortalama Gini’deki azalmaya göre önerilen 3B nokta tabanlı sınıflandırma için geometrik özelliklerin yoğunluk tabanlı özelliklerden daha yüksek bir öneme sahip olduğu ifade edilebilir.

Test alanı 2’de RO sınıflandırması ile ortalama Gini’deki azalmaya göre belirlenen 8 sınıflandırma özelliklerinin önemleri incelendiğinde en önemli sınıflandırma özelliği Test alanı 1’de olduğu gibi “Nokta sayısı” ve en az önemli sınıflandırma özelliği ise “Z’nin kurtosisi” olarak belirlenmiştir. Test alanı 2’de 25 özellik arasında ise en önemli sınıflandırma özelliğinin “Z’nin standart sapması”, en az önemli olan sınıflandırma özelliğinin ise “%5’lik Z aralığındaki nokta sayısı” olduğu görülmüştür. Buna göre, Test alanı 1’deki duruma benzer şekilde tüm özelliklerin önemleri dikkate alındığında RO sınıflandırmasında ortalama Gini’deki azalmaya göre önerilen 3B nokta tabanlı sınıflandırma için geometrik özelliklerin yoğunluk tabanlı özelliklerden çok daha yüksek bir öneme sahip olduğu belirlenmiştir.

Test alanı 1 ve Test alanı 2’de RO sınıflandırmasında ortalama Gini’deki azalmaya göre gerçekleştirilen özellik önemleri birbirleri ile karşılaştırıldığında genel olarak her iki alanda da iki farklı set sınıflandırma özelliği arasından (8 özellik ve 25 özellik) en önemli özellikler ve en az önemli özelliklerin birbirleri ile tutarlı olduğu görülmüştür. Ayrıca her iki alanda her iki sınıflandırma özellik setinin önemleri incelendiğinde de genel olarak tüm sınıflandırmalarda geometri tabanlı

özelliklerin yoğunluk tabanlı özelliklerden çok daha önemli olduğu sonucu çıkmıştır.

Kentsel ağaç türlerinin sınıflandırılması için makine öğrenme tabanlı sınıflandırma algoritmalarının performanslarının detaylı analizlerinin ardından genel olarak, Test alanı 1'deki DVM, RO ve ÇKA sınıflandırma doğruluklarının Test alanı 2'den daha yüksek olduğu sonucu çıkarılabilir. Bunun en büyük nedeni olarak Test alanı 1'e göre, Test alanı 2'deki geniş ve iğne yapraklı ağaç türlerinin dağılımının homojenliğinin azalması gösterilebilir. Buna ek olarak her iki test alanı ve her iki özellik setinin kullanımında da geniş yapraklı ağaçların iğne yapraklı ağaçlardan çok daha hassas bir şekilde sınıflandırılabilirdiği, başka bir ifade ile geniş yapraklı ağaçların iğne yapraklı ağaçlarla karışım oranının düşük olduğu vurgulanabilir. Bu durum da benzer şekilde kentsel çalışma alanında yer alan geniş yapraklı ağaç örneklemelerinin iğne yapraklı ağaç örneklerinden sayıca daha fazla olması ile açıklanabilir. Ayrıca genel olarak tüm sınıflandırmalarda (Test alanı 2'de 8 özelliğin kullanıldığı durum hariç) DVM ve özellikle RO sınıflandırıcısının benzer özellikler gösteren geniş ve iğne yapraklı ağaçlar için bile güvenilir, yüksek doğruluklu sonuçlar ürettiği, buna rağmen ÇKA'nın nispeten daha düşük sonuçlar ürettiği ifade edilebilir. RO sınıflandırıcısının DVM ve ÇKA sınıflandırıcılarına göre başarısı, RO sınıflandırıcısının dengesiz örneklemelerin işlenmesi için uygun olması, büyüyen ağaçlar sırasında sınıflandırma modeline ek rastgelelik eklemesi ve ayrıca her bir düğümün bölme işleminde rastgele bir özellik alt kümesi arasından yalnızca en iyi özellikleri araması gibi özellikleri ile açıklanabilir. DVM ve RO sınıflandırma modelleri ile karşılaştırıldığında, ÇKA modellerinin her yinelemede optimizasyon için çok sayıda ağırlığa sahip olması ve buna bağlı olarak ÇKA'nın eğitim aşamasında daha fazla eğitim verisine ve daha fazla parametre analizine ihtiyacının olması da ÇKA'nın daha düşük sınıflandırma başarısına sahip olmasının nedenleri olarak açıklanabilir.

Kentsel çalışma bölgesinde ağaç türlerinin sınıflandırılmasının ardından iki farklı test alanında (Test alanı A ve Test alanı B) bireysel cadde ağaçlarının otomatik olarak elde edilebilmesi için daha önce önerilen makine öğrenme tabanlı 2B segmentasyon yaklaşımı ("MeanShift" ve "DBSCAN" kümeleme algoritmaları ile) kullanılarak pilot bir uygulama da gerçekleştirilmiştir. Kentsel ağaç türlerinin

sınıflandırılmasına benzer iş adımları ile gerçekleştirilen uygulamada test alanı A'da "MeanShift" kümeleme algoritması için bütünlük (completeness) ve doğruluk (correctness) değerleri sırasıyla %89.47 ve %100, "DBSCAN" kümeleme algoritması için ise bu değerler sırasıyla %89.47 ve %94.44 olarak elde edilmiştir. Test alanı B'de ise hem "MeanShift" hem de "DBSCAN" kümeleme algoritmaları için tüm bütünlük ve doğruluk değerleri %100 olarak elde edilmiştir. Genel olarak kentsel alanda 2B nokta tabanlı segmentasyon sonucu tekil cadde ağaçlarının elde edilmesi için "MeanShift" kümeleme algoritması "DBSCAN" kümeleme algoritmasından daha yüksek doğruluğa sahip sonuçlar üretmesine rağmen, "DBSCAN" algoritması ile elde edilen tekil cadde ağaçlarının doğruluklarının oldukça yüksek olduğu görülmüştür. Test alanı A ve Test alanı B'de elde edilen sonuçlar esas alındığında, "MeanShift" ve "DBSCAN" kümeleme algoritmaları ile nokta tabanlı olarak kentsel alanlarda tekil cadde ağaçları başarılı bir şekilde elde edilebilir.

Literatürdeki ağaç türleri sınıflandırma çalışmalarının çoğu ormanlık alanlarda homojen dağılımlara sahip olan ağaçlar için gerçekleştirilirken bu çalışmada ise kentsel alanlarda heterojen dağılıma sahip geniş ve iğne yapraklı ağaç türleri hızlı ve uygun maliyetli olarak LiDAR verilerinden elde edilmiştir. LiDAR verilerinden ağaç türlerinin belirlenmesi, geleneksel çalışmaların çoğunda normalize edilmiş yüzey modeli (normalised Digital Surface Model - nDSM) ya da kanopi yüksekliği modeli (Canopy Height Model – CHM) kullanılarak segmentasyon ve sınıflandırma işlemleri sonrasında elde edilirken, bu çalışmada 3B noktalar düzenli grid yapıya enterpole edilmeden doğrudan kullanılarak nokta tabanlı yaklaşımlar ile segmentasyon ve sınıflandırma işlemi sonrasında elde edilmiştir. Böylece gride enterpolasyon sırasında yaşanabilecek veri kayıplarının önüne geçilmiştir. Bunlara ek olarak bu çalışmada makine öğrenme algoritmaları ile ağaç türleri sınıflandırılmasının gerçekleştirilebilmesi için, yaygın olarak kullanılan LiDAR verilerinin dalga formu analizi sonucu elde edilen sınıflandırma özellikleri yerine sadece geometri ve yoğunluk tabanlı sınıflandırma özellikleri kullanılarak kentsel alanlardaki geniş ve iğne yapraklı ağaçlar başarılı bir şekilde sınıflandırılmıştır. Bu tez çalışması kapsamında önerilen makine öğrenme tabanlı sınıflandırma yaklaşımları ile çalışma alanının büyüklüğü ve küçüklüğünden bağımsız olarak

aynı işlem adımları gerçekleştirilmek suretiyle ağaç türleri sınıflandırılabilir ve özellikle büyük alanlarda hız ve maliyet açısından çok daha büyük avantaj sağlanabilir.

Bu çalışmada, özellikle ÇKA algoritması ile gerçekleştirilen sınıflandırma sonuçlarının iyileştirilebilmesi için daha homojen dağılmış örneklemelerin yer aldığı, daha büyük eğitim alanları eğitim için kullanılabilir. Daha başarılı sınıflandırma sonuçlarının elde edilebilmesi için LiDAR nokta bulutu verilerinden türetilen geometrik ve yoğunluk tabanlı özelliklere ek olarak tam dalga formulu LiDAR verilerinden elde edilen çok sayıda farklı sınıflandırma özelliği kullanılabilir. Ayrıca, özellikle sayıca fazla olan sınıflandırma özelliklerinin önemlerinin analizinin ardından, özellik seçim işlemi gerçekleştirilerek yalnızca daha fazla etkiye sahip özellikler kullanılarak sınıflandırma sonuçlarında bir iyileşme sağlanabilir.

Bu tez çalışması, gelecekteki büyük ölçekli kentsel uygulamalarda makine öğrenmesi algoritmaları kullanılarak gerçekleştirilecek geniş yapraklı ve iğne yapraklı ağaç türleri sınıflandırmaları ve kent alanlarında tekil cadde ağaçlarının segmentasyonu için pilot bir çalışma niteliğindedir. Bu çalışmada önerilen makine öğrenme algoritmaları ile ağaç türlerinin 3B nokta tabanlı sınıflandırma yaklaşımı kullanılarak elde edilen kentsel geniş ve iğne yapraklı ağaç türleri, yine makine öğrenme algoritmaları ile 2B nokta tabanlı segmentasyon yaklaşımı ile elde edilen tekil cadde ağaçları afet yönetimi, şehir planlaması, çevre koruma veya kentsel kalkınma politikası üretimi gibi birçok farklı çalışmada kolaylıkla kullanılabilir.

- [1] R. R. Jensen, P. J. Hardin ve A. J. Hardin, "Classification of urban tree species using hyperspectral data," *Geocarto International*, vol. 27, pp. 443–458, 2012.
- [2] J. Xu, Z. Cai, T. Wang, G. Liu, P. Tang ve X. Ye, "Exploring Spatial Distribution of Pollen Allergenic Risk Zones in Urban China," *Sustainability*, vol. 8, no. 10, pp. 978, 2016.
- [3] F. Chianucci, N. Puletti, E. Giacomello, A. Cutini ve P. Corona, "Estimation of leaf area index in isolated trees with digital photography and its application to urban forestry," *Urban For Urban Greening*, vol. 14, pp. 377–382, 2015.
- [4] R. Pu ve S. Landry, "A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species," *Remote Sensing of Environment*, vol.124, pp. 516–533, 2012.
- [5] J.Tigges, T. Lakes ve P. Hostert, "Urban vegetation classification: Benefits of multitemporal RapidEye satellite data," *Remote Sensing of Environment*, vol. 136, pp. 66-75, 2013.
- [6] D. Li, Y. Ke, H. Gong, B. Chen ve L. Zhu, "Tree species classification based on WorldView-2 imagery in complex urban environment," 2014 Third International Workshop on Earth Observation and Remote Sensing Applications (EORSA), 2014, pp. 326-330.
- [7] Y. T. Mustafa, H. N. Habeebb, A. Steinc ve F. Y. Sulaimanb, "Identification and Mapping of Tree Species in Urban Areas Using WORLDVIEW-2 Imagery," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 175-181, 2015.
- [8] Y. Dian, Y. Pang, Y. Dong ve Z. Li, "Urban tree species mapping using airborne LiDAR and hyperspectral data," *Journal of the Indian Society of Remote Sensing*, pp. 1–9, 2016.
- [9] R. Shojanoori, H. Z. M. Shafri, S. Mansor ve M. H. Ismail, "The Use of WorldView-2 Satellite Data in Urban Tree Species Mapping by Object-Based Image Analysis Technique," *Sains Malaysiana*, vol.45, no.7, pp. 1025–1034, 2016.
- [10] J. L. Strunk, J. R. Mills, P. Ries, H. Temesgen ve L. Jeroue, "An urban forest-inventory-and-analysis investigation in Oregon and Washington," *Urban Forestry and Urban Greening*, vol. 18, pp. 100–109, 2016.
- [11] C. Iovan, D. Boldo ve M. Cord, "Detection, Characterization, and modeling vegetation in urban areas from high-resolution aerial imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 1, no. 3, pp. 206-213, 2008.

- [12] S. Kim, R. J. McGaughey, H. E. Andersen ve G. Schreuder, "Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data," *Remote Sensing of Environment*, vol. 113, pp. 1575–1586, 2009.
- [13] D. Li, Y. Ke, H. Gong ve X. Li, " Object-Based Urban Tree Species Classification Using Bi-Temporal WorldView-2 and WorldView-3 Images," *Remote Sensing*, vol. 7, no. 12, pp. 16917-16937, 2015.
- [14] R. Pu, "Broadleaf species recognition with in situ hyperspectral data," *International Journal of Remote Sensing*, vol. 30, no. 11, pp. 2759–2779, 2009.
- [15] J. Li, B. Hu, G. Sohn ve L. Jing, "Individual tree species classification using structure features from high density airborne lidar data," 2010 IEEE International Geoscience and Remote Sensing Symposium, 2010, pp. 2099-2102.
- [16] J. Li, B. Hu ve T. L. Noland, "Classification of tree species based on structural features derived from high density LiDAR data," *Agricultural and Forest Meteorology*, vol. 171–172, pp. 104–114. 2013.
- [17] A. Moradi, M. Satari ve M. Momeni, "Individual Tree of Urban Forest Extraction From Very High Density LiDAR Data," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLI, no. B3, pp. 337-343, 2016.
- [18] L. Cao, N. Coops, J. L. Innes, J. Dai, H. Ruan ve G. She, "Tree species classification in subtropical forests using small-footprint full-waveform LiDAR data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 49, pp. 39-51, 2016.
- [19] J. Holmgren ve Å. Persson, "Identifying species of individual trees using airborne laser scanning," *Remote Sensing of Environment*, vol. 90, pp. 415-423, 2004.
- [20] J. Reitberger, P. Krzystek ve U. Stilla, "Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees," *International Journal of Remote Sensing*, vol. 29, no.5, pp. 1407-1431, 2008.
- [21] B. Höfle, M. Hollaus ve J. Hagenauer, "Urban vegetation detection using radiometrically calibrated small-footprint full-waveform airborne LiDAR data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 67, pp. 134–147, 2012.
- [22] Z. Koma, K. Koenig ve B. Höfle, "Urban Tree Classification Using Full-Waveform Airborne Laser Scanning," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. III-3, pp. 185–192, 2016.

- [23] J. Reitberger, P. Krzystek ve U. Stilla, "Analysis of full waveform lidar data for tree species classification," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 36, pp. 228–233, 2006.
- [24] Z. Zhongyang, C. Yinglei, S. Xiaosong, Q. Xianxiang ve S. Li, "Classification of LiDAR Point Cloud based on Multiscale Features and PointNet," *2018 Eighth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 2018, pp. 1-7.
- [25] C. Wen, L. Yang, X. Li, L. Peng ve T. Chi, "Directionally constrained fully convolutional neural network for airborne lidar point cloud classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 162, pp. 50-62, 2020.
- [26] C. I. Miller, J. J. Thomas, A. M. Kim, J. P. Metcalf ve R. C. Olsen, "Application of image classification techniques to multispectral lidar point cloud data," *Proceedings SPIE Laser Radar Technology and Applications*, vol. 9832, 2016.
- [27] N. Ekhtari, C. Glennie ve J. C. Fernandez-Diaz, "Classification of Airborne Multispectral Lidar Point Clouds for Land Cover Mapping," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 6, pp. 2068-2078, 2018.
- [28] M. Voss ve R. Sugumaran, "Seasonal effect on tree species classification in an urban environment using hyperspectral data, LiDAR, and an object-oriented approach," *Sensors*, vol. 8, pp. 3020–3036, 2008.
- [29] M. A. A. Khameneh, "Tree Detection and Species Identification using LiDAR Data," M.S. thesis, School of Architecture and the Built Environment Royal Institute of Technology (KTH), Stockholm, Sweden, 2013.
- [30] Z. Zhang, X. Liu ve W. Wright, "Statistical analysis of LiDAR-derived structure and intensity variables for tree species identification," In: *2013 GIS and Remote Sensing Research Conference (SIRC NZ 2013)*, 2013, pp. 29–30.
- [31] H. O. Ørka, E. Næsset ve O. M. Bollandsås, "Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data," *Remote Sensing Environment*, vol. 113, pp. 1163–1174, 2009.
- [32] N. R. Vaughn, L. M. Moskal ve E. C. Turnblom, "Fourier transformation of waveform LiDAR for species recognition," *Remote Sensing Letters*, vol.2, pp. 347–356, 2011.
- [33] C. Ko, T. K. Rimmel ve G. Sohn, "Mapping tree genera using discrete LiDAR and geometric tree metrics," *Bosque*, vol. 33, no. 3, pp. 313–319, 2012.

- [34] N. R. Vaughn, L. M. Moskal ve E. C. Turnblom, “Tree species detection accuracies using discrete point lidar and airborne waveform lidar,” *Remote Sensing*, vol. 4, no. 2, pp. 377–403, 2012.
- [35] E. Lindberg, L. Eysn, M. Hollaus, J. Holmgren ve N. Pfeifer, “Delineation of tree crowns and tree species classification from full-waveform airborne laser scanning data using 3-D ellipsoidal clustering,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, pp. 3174–3181, 2014.
- [36] X. Yu, P. Litkey, J. Hyypä, M. Holopainen ve M. Vastaranta, “Assessment of low density full-waveform airborne laser scanning for individual tree detection and tree species classification,” *Forest*, vol. 5, pp. 1011–1031, 2014.
- [37] T. Brandtberg, “Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 61, pp.325–340, 2007.
- [38] J. N. Heinzl, H. Weinacker ve B. Koch, “Full automatic detection of tree species based on delineated single tree crowns – a data fusion approach for airborne laser scanning data and aerial photographs,” In 8th international conference on LiDAR applications in forest assessment and inventory, 2008, pp. 76–85.
- [39] J. Holmgren, A. Persson ve U. Söderman, “Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images,” *International Journal of Remote Sensing*, vol.29, no. 5, pp. 1537–1552, 2008.
- [40] T. G. Jones, N. C. Coops ve T. Sharma, “Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada,” *Remote Sensing of Environment*, vol. 114, pp. 2841–2852, 2010.
- [41] S. Kim, T. Hinckley ve D. Briggs, “Classifying individual tree genera using stepwise cluster analysis based on height and intensity metrics derived from airborne laser scanner data,” *Remote Sensing of Environment*, vol. 115, no. 12, pp. 3329–3342, 2011.
- [42] J. Li, “Individual Tree Delineation And Species Identification In Deciduous And Mixed Canadian Forests Using High Spatial Resolution Airborne Lidar And Image Data,” PhD Thesis, Graduate Program In Earth And Space Science York University Toronto, Ontario, Canada, 2013.
- [43] A. Ghosh, F. E. Fassnacht, P. K. Joshi ve K. Barbara, “A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 26, pp. 49–63, 2014.

- [44] C. Sommer, S. Holzwarth, U. Heiden, M. Heurich, J. Mueller ve W. Mauser, “Feature-based tree species classification using airborne hyperspectral and lidar data in the Bavarian Forest National Park,” *EARSeL eProceedings*, vol. 14, no. 2, pp. 49-70, 2015.
- [45] M. Alonzo, J. P. McFadden, D. J. Nowak ve D. A. Roberts, “Mapping urban forest structure and function using hyperspectral imagery and lidar data,” *Urban Forestry and Urban Greening*, vol. 17, pp. 135–147, 2016.
- [46] T. Sasaki, J. Imanishi, K. Ioki, Y. Morimoto ve K. Kitada, “Object-based classification of land cover and tree species by integrating airborne LiDAR and high spatial resolution imagery data,” *Landscape and Ecological Engineering*, vol. 8, pp. 157–171, 2012.
- [47] Z. Zhang ve X. Liu. “Support Vector Machines for Tree Species Identification Using Lidar Derived Structure and Intensity Variables,” *Geocarto International*, vol. 28, no. 4, pp. 364–378, 2013.
- [48] J. Schumacher ve T. Nord-Larsen, “Wall-to-wall tree type classification using airborne lidar data and CIR images,” *International Journal of Remote Sensing*, vol. 35, no. 9, pp. 3057-3073, 2014.
- [49] A. Hovi, L. Korhonen, J. Vauhkonen ve I. Korpela, “LiDAR waveform features for tree species classification and their sensitivity to tree- and acquisition related parameters,” *Remote Sensing of Environment*, vol. 173, pp. 224–237, 2016.
- [50] S. Kim, G. Schreuder, R. J. McGaughey ve H. -E. Andersen, “Individual tree species identification using LIDAR intensity data,” In *ASPRS 2008 Annual Conference*, 2008, pp. 382–393.
- [51] M. A. Cho, R. Mathieu, G. P. Asner, L. Naidoo, J. A. Van Aardt, A. Ramoelo, P. Debba, K. Wessels, R. Main, I. P. Smit ve B. F. Erasmus, “Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system,” *Remote Sensing of Environment*, vol. 125, pp. 214–226, 2012.
- [52] J. Marrs ve W. Ni-Meister, “Machine Learning Techniques for Tree Species Classification Using Co-Registered LiDAR and Hyperspectral Data,” *Remote Sensing*, vol. 11, pp. 819, 2019.
- [53] F.E. Fassnacht, H. Latifi, K. Sterenczak, A. Modzelewska, M. Lefsky, L.T. Waser, C. Straub ve A. Ghosh, “Review of studies on tree species classification from remotely sensed data,” *Remote Sensing of Environment*, vol. 186, pp. 64–87, 2016.
- [54] M. Hollaus, W. Mücke, B. Höfle, W. Dorigo, N. Pfeifer, W. Wagner, C. Bauerhansl ve B. Regner, “Tree species classification based on full-waveform airborne laser scanning data,” In *9th International SilviLaser Conference*, 2009, pp. 54–62.

- [55] I. Korpela, , H. O. Ørka, M. Maltamo, T. Tokola ve J. Hyypä, “Tree species classification using airborne LiDAR — Effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type,” *Silva Fennica*, vol. 44, no. 2, pp. 319–339, 2010.
- [56] J. Heinzl ve B. Koch, “Exploring full-waveform LiDAR parameters for tree species classification,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 13, pp. 152–160, 2011.
- [57] M. Dalponte, L. Bruzzone ve D. Gianelle, “Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data,” *Remote Sensing of Environment*, vol. 123, pp. 258–270, 2012.
- [58] C. Zhang ve F. Qiu, “Mapping individual tree species in an urban forest using airborne lidar data and hyperspectral imagery,” *Photogrammetric Engineering and Remote Sensing*, vol. 78, no. 10, pp. 1079–1087, 2012.
- [59] M. Alonzo, B. Bookhagen ve D. A. Roberts, “Urban tree species mapping using hyperspectral and lidar data fusion,” *Remote Sensing of Environment*, vol. 148, pp.70–83, 2014.
- [60] A. Verlič, N. Đurić, Ž. Kokalj, A. Marsetič, P. Simončič ve K. Oštir, “Tree Species Classification Using WorldView-2 Satellite Images and Laser Scanning Data in a Natural Urban Forest,” *Šumarski list*, vol. 138, no. 9–10, pp. 477–488, 2014.
- [61] K. Koenig ve B. Höfle, “Full-Waveform airborne laser scanning in vegetation studies - A review of point cloud and waveform features for tree species classification,” *Forests*, vol. 7, 2016.
- [62] Z. Zhang, A. Kazakova, L. M. Moskal ve D. M. Styers, “Object-Based tree species classification in urban ecosystems using LiDAR and hyperspectral data,” *Forests*, vol. 7, no. 6, pp. 122, 2016.
- [63] X. Shen, L. Cao, “Tree-Species Classification in Subtropical Forests Using Airborne Hyperspectral and LiDAR Data,” *Remote Sensing*, vol. 9, no. 11, pp. 1180, 2017.
- [64] Y. Shi, A. Skidmore ve M. Heurich, “Important LiDAR metrics for discriminating forest tree species in Central Europe,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 137, pp. 163–174, 2018.
- [65] M. Uzar Dinlemek, “Otomatik Bina Çıkarımı Uygulamalarında Çoklu Algılama Sistemi Kullanım Olanaklarının Analizi,” *Doktora Tezi, YTÜ Fen Bilimleri Enstitüsü, İstanbul*, 2012.
- [66] R. C. Weih ve N. D. Riggan, “Object-Based Classification vs. Pixel-Based Classification Comparative Importance of Multi-Resolution Imagery,” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVIII- 4/C7, 2010.

- [67] Y. Brik, N. Zerrouki ve D. Bouchaffra, "Combining Pixel- and Object-Based Approaches for Multispectral Image Classification Using Dempster-Shafer Theory," In 2013 International Conference on Signal-Image Technology and Internet-Based Systems, 2013, pp. 448-453.
- [68] I. Bello, N. Chigbu ve G. Agbaje, "Large scale mapping: an empirical comparison of pixel-based and object-based classifications of remotely sensed data," South African Journal of Geomatics, vol. 6, pp. 277 – 294, 2017.
- [69] Y. Gao, J. - F. Mas, B. H. P. Maathuis, X. Zhang ve P. M. Van Dijk, "Comparison of pixel- based and object- oriented image classification approaches-a case study in a coal fire area, Wuda, Inner Mongolia, China," International Journal of Remote Sensing, vol.27, no. 18, pp. 4039-4055, 2006.
- [70] N. Zerrouki ve D. Bouchaffra, "Pixel-based or Object-based: Which approach is more appropriate for remote sensing image classification?," In 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2014, pp. 864-869.
- [71] T. Blaschke, "Object Based Image Analysis For Remote Sensing", ISPRS Journal of Photogrammetry and Remote Sensing, vol. 65, pp. 2–16, 2010.
- [72] Z. Erişir, "Nokta Tabanlı Sınıflandırma Yöntemleri ile LiDAR Verilerinin Sınıflandırılması," Yüksek Lisans Tezi, YTÜ Fen Bilimleri Enstitüsü, İstanbul, 2015.
- [73] A. Zhang ve P. Tang, "Fusion algorithm of pixel-based and object-based classifier for remote sensing image classification," In 2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS, 2013, pp. 2740-2743.
- [74] Y. Sun, C. Wang, J. Li, Z. Zhang, D. Zai, P. Huang ve C. Wen, "Automated segmentation of LiDAR point clouds for building rooftop extraction," In 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2016, pp. 1472-1475.
- [75] C. Wen, X. Li, X. Yao, L. Peng ve T. Chi, "Airborne LiDAR point cloud classification with global-local graph attention convolution neural network," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 173, pp. 181-194, 2021.
- [76] G. Forlani ve C. Nardinocchi, "Adaptive Filtering of Aerial Laser Scanning Data," International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 36(Part 3/W52), pp. 130–135, 2007.
- [77] D. Gajski, T. Fiedler ve A. Krtalić, "Classification and Filtering of Airborne Topographic Lidar Data," International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 34(6/W11), pp. 100-104, 2003.

- [78] B. Guo, X. Huang, F. Zhang ve G. Sohn, "Classification of airborne laser scanning data using JointBoost," *ISPRS Journal of Photogrammetry Remote Sensing*, vol. 100, pp. 71–83, 2015.
- [79] C. Zhao, H. Guo, J. Lu, D. Yu, D. Li ve X. Chen, "ALS Point Cloud Classification With Small Training Data Set Based on Transfer Learning," *In IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 8, pp. 1406-1410, 2020.
- [80] J. Niemeyer, F. Rottensteiner ve U. Soergel, "Conditional random fields for lidar point cloud classification in complex urban areas," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 1, no. 3, pp. 263-268, 2012.
- [81] LAS Specification 1.4–R15; The American Society for Photogrammetry and Remote Sensing: Bethesda, ML, USA, 2019.
- [82] P. Colling, M. Rottmann, L. Roese-Koerner ve H. Gottschalk, "False Positive Detection and Prediction Quality Estimation for LiDAR Point Cloud Segmentation," *In 2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI)*, 2021, pp. 18-25.
- [83] Q. Zhan ve L. Yu, "Segmentation of LiDAR point cloud based on similarity measures in multi-dimension euclidean space," *In Advances in Computer Science and Engineering*, 2012, pp. 349-357.
- [84] E. Che, M. J. Olsen, C. E. Parrish ve J. Jung, "Pavement marking retro-reflectivity estimation and evaluation using mobile lidar data," *Photogrammetric Engineering and Remote Sensing*, vol. 85, no. 8, pp. 573-583, 2018.
- [85] M. Awrangjeb ve C. S. Fraser, "Rule-based segmentation of LIDAR point cloud for automatic extraction of building roof planes," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 2(3/W1), pp. 1-6, 2013.
- [86] T. Wu, X. Hu ve L. Ye, "Fast and Accurate Plane Segmentation of Airborne LiDAR Point Cloud Using Cross-Line Elements," *Remote Sensing*, vol. 8, no. 5, pp. 383, 2016.
- [87] G. Vosselman, B. Gorte ve G. Sithole, "Change Detection For Updating Medium Scale Maps Using Laser Altimetry," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 35(Part B3), pp. 207–212, 2004.
- [88] T. Melzer, "Non-parametric segmentation of ALS point clouds using mean shift," *Journal of Applied Geodesy*, vol. 1, no. 3, 159-170, 2007.
- [89] M. Wang ve Y. H. Tseng, "Incremental segmentation of lidar point clouds with an octree- structured voxel space," *The Photogrammetric Record*, vol. 26, no. 133, pp. 32-57, 2011.
- [90] S. Ural ve J. Shan, "Min-cut based segmentation of airborne lidar point clouds," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 39, 2012.

- [91] M. Le Louarn, P. Clergeau, E. Briche ve M. Deschamps-Cottin, “Kill Two Birds with One Stone: Urban Tree Species Classification Using Bi-Temporal Pléiades Images to Study Nesting Preferences of an Invasive Bird.” *Remote Sensing*, vol. 9, pp. 916, 2017.
- [92] S. M. I. Zolanvari, D. F. Laefer ve A. S. Natanzi, “Three-dimensional building façade segmentation and opening area detection from point clouds,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 143, pp. 134-149, 2018.
- [93] R. Passini ve K. Jacobsen, “Filtering of Digital Elevation Model,” *Proceedings of the ASPRS Annual Convention*, Washington DC, 2002.
- [94] X. Hu ve L. Ye, “A Fast and Simple Method of Building Detection From Lidar Data Based On Scan Line Analysis”, In *The ISPRS Workshop on 3D Virtual City Modeling*, Regina, Canada, 2013.
- [95] G. Sithole ve G. Vosselman, “Experimental Comparison of Filter Algorithms For Bare-Earth Extraction From Airborne Laser Scanning Point Clouds”, *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 59, no. (1-2), pp. 85-101, 2004.
- [96] D. Tóvári ve N. Pfeifer, “Segmentation Based Robust Interpolation-A New Approach To Laser Data Filtering”, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 36, no. 3, 2005.
- [97] N. Pfeifer ve G. Mandlbürger, “LiDAR Data Filtering and DTM Generation, Topographic Laser Ranging And Scanning: Principles and Processing,” *CRC Press, Taylor & Francis Group*, New York, pp. 308-331, 2008.
- [98] H. B. Kim ve G. Sohn, “Point-based classification of power line corridor scene using random forests”, *Photogrammetric Engineering and Remote Sensing*, vol. 79, pp. 821-33, 2013.
- [99] Y. Yang ve Y. Lin, “A Novel Deformable Model For Urban Vegetation Detection Using LiDAR Data, In Image and Signal Processing,” In *2nd International Congress on Image and Signal Processing*, 2009, pp. 1-5.
- [100] P. Axelsson, “DEM Generation From Laser Scanner Data Using Adaptive TIN Models”, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 33, 2000.
- [101] Z. Cetin ve N. Yastikli, “The Use of Machine Learning Algorithms in Urban Tree Species Classification,” *ISPRS International Journal of Geo-Information*, vol. 11, no. 4, pp. 226, 2022.
- [102] N. Chehata, L. Guo ve C. Mallet, “Airborne LiDAR feature selection for urban classification using random forests,” *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 39 (Part 3/W8), pp. 207-212, 2009.

- [103] H. B. Kim, ve G. Sohn, "3D classification of power-line scene from airborne laser scanning data using random forests," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Saint-Mandé, France, vol. 38(3A), 2010, pp.126-132.
- [104] C. Mallet, F. Bretar, M. Roux, U. Soergel ve C. Heipke, "Relevance assessment of full-waveform lidar data for urban area classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 6, pp. 71–84, 2011.
- [105] A. B. Kurtak ve N. Yastıklı, "Tam Dalga Boyu LiDAR Verileri ile Bitki Örtüsü Sınıflandırma Olanaklarının Araştırılması," In VII. Uzaktan Algılama ve CBS Sempozyumu (UZAL-CBS 2018), Eskişehir, 2018, pp. 1-10.
- [106] A. B. Kurtak, "Tam Dalga LiDAR Verileri ile Bitki Örtüsü Sınıflandırma Olanaklarının Araştırılması," Yüksek Lisans Tezi, YTÜ Fen Bilimleri Enstitüsü, İstanbul, 2019.
- [107] J. O'Neil-Dunne, S. MacFaden ve A. Royar, "A Versatile, Production Oriented Approach To High-Resolution Tree-Canopy Mapping In Urban And Suburban Landscapes Using GEOBIA and Data Fusion," *Remote Sensing*, vol. 6, pp. 12837-12865, 2014.
- [108] R. G. Congalton ve K. Green, "Assessing the Accuracy of Remotely Sensed Data: Principles and Practises," Second Edition, CRC Press Taylor & Francis Group, New York, 2009.
- [109] K. Navulur, "Multispectral Image Analysis Using The Object-Oriented Paradigm," CRC Press, Taylor and Frances Group, New York, 2007.
- [110] M. Rutzinger, F. Rottensteiner ve N. Pfeifer, "A Comparison of Evaluation Techniques For Building Extraction From Airborne Laser Scanning," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 2, no. 1, pp. 11–20, 2009.
- [111] G. M. Foody, "Thematic map comparison: evaluating the statistical significance of differences in classification accuracy," *Photogrammetric Engineering and Remote Sensing*, vol. 70, no. 5, pp. 627-633, 2004.
- [112] A. E. Maxwell, M. Sharma, J. S. Kite, K. A Donaldson, S. M. Maynard ve C. M. Malay, "Assessing the Generalization of Machine Learning-Based Slope Failure Prediction to New Geographic Extents," *ISPRS International Journal of Geo-Information*, vol. 10, pp. 293, 2021.
- [113] J. Weaver, B. Moore, A. Reith, J. McKee ve D. Lunga, "A Comparison of Machine Learning Techniques to Extract Human Settlements from High Resolution Imagery," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018, pp. 6412-6415.

- [114] M. Yancho, N. Coops, P. Tompalski, T. Goodbody ve A. Plowright, "Fine-Scale Spatial and Spectral Clustering of UAV-Acquired Digital Aerial Photogrammetric (DAP) Point Clouds for Individual Tree Crown Detection and Segmentation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, pp. 4131–4148, 2019.
- [115] N. Yastikli ve Z. Cetin, "Classification of raw LiDAR point cloud using point-based methods with spatial features for 3D building reconstruction," *Arabian Journal of Geosciences*, vol. 14, pp. 146, 2021.
- [116] C. Malhotra, V. Kotwal ve S. Dalal, (2018). "Ethical Framework for Mahine Learning," 2018 ITU Kaleidoscope: Machine Learning for a 5G Future (ITU K), 2018, pp. 1-8.
- [117] P. P. Shinde ve S. Shah, "A Review of Machine Learning and Deep Learning Applications," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1-6.
- [118] M. Chai, Y. Lin ve Y. Li, "Machine Learning and Modern Education," In *International Conference on E-Learning, E-Education, and Online Training*, 2018, pp. 41-46.
- [119] G. Ratsch, "A Brief Introduction into Machine Learning," Friedrich Miescher Laboratory of the Max Planck Society, 2004.
- [120] A. Holzinger, "Introduction to Machine Learning and Knowledge Extraction (MAKE)," *Machine Learning and Knowledge Extraction.*, vol. 1, no. 1, pp. 1-20, 2017.
- [121] O. Güldali, "Deep Feature Transfer From Deep Learning Models Into Machine Learning Algorithms to Classfy Covid-19 From Chest X-Ray Images," Master Thesis, Istanbul Technical University, Graduate School, Istanbul, Turkey, 2021.
- [122] O. Simeone, "A Very Brief Introduction to Machine Learning With Applications to Communication Systems," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 4, pp. 648-664, 2018.
- [123] A. Tegen, P. Davidsson ve J.A. Persson, "A Taxonomy of Interactive Online Machine Learning Strategies," In: Hutter, F., Kersting, K., Lijffijt, J., Valera, I. (eds) *Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2020. Lecture Notes in Computer Science*, vol. 12458. pp. 137-153, 2021.
- [124] M. Stamp, "Introduction to machine learning with applications in information security," Chapman and Hall/CRC, 2017.
- [125] M. Stamp, "Introduction to machine learning with applications in information security," Chapman and Hall/CRC, 2017.

- [126] K. Jung, "A Study on Machine Learning for Steganalysis," In 3rd International Conference on Machine Learning and Soft Computing, 2019, pp. 12-15.
- [127] K. Fukunaga ve L. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," IEEE Transactions on Information Theory, vol. 21, pp. 32–40, 1975.
- [128] Z. Q. Wen ve Z. X. Cai, "Mean shift algorithm and its application in tracking of objects," In 5th International Conference on Machine Learning and Cybernetics, Dalian, 2006, pp. 4024- 4028.
- [129] Y. Du, B. Sun, R. Lu, C. Zhang ve H. Wu, "A method for detecting high-frequency oscillations using semi-supervised K-means and mean shift clustering," Neurocomputing, vol. 350, pp. 102-107, 2019.
- [130] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," In the fifth Berkeley symposium on mathematical statistics and probability, vol. 1, 1967, pp. 281–297.
- [131] T. Kanungo, D.M. Mount, N. Netanyahu, C. Piatko, R. Silverman ve A.Y. WuAn, "Efficient K-means clustering algorithm: analysis and implementation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 881–892, 2002.
- [132] A. K. Jain,"Data clustering: 50 years beyond K-means," Pattern Recogn Letters, vol. 31, pp. 651–66, 2010.
- [133] K. Wagstaff, C. Cardie, S. Rogers ve S. Schroedl, "Constrained K-means clustering with background knowledge," In 18th International Conference on Machine Learning, 2001, pp. 577-584.
- [134] T. Boonchoo, X. Ao, Y. Liu, W. Zhao ve Q. He, "Grid-based DBSCAN: indexing and inference," Pattern Recogn, vol. 90, pp. 271–284, 2019.
- [135] P. Paral, A. Chatterjee ve A. Rakshit, "Vision Sensor-Based Shoe Detection for Human Tracking in a Human–Robot Coexisting Environment: A Photometric Invariant Approach Using DBSCAN Algorithm," IEEE Sensors Journal, vol. 19, no. 12, pp. 4549-4559, 2019.
- [136] M. Khatoun ve W. A. Banu, "An efficient method to detect communities in social networks using DBSCAN algorithm," Social Network Analysis and Mining, vol. 9, no. 9, 2019.
- [137] R. Nasiboglu, B. T. Tezel ve E. Nasibov, " Learning the stress function pattern of ordered weighted average aggregation using DBSCAN clustering," International Journal of Intelligent Systems, vol. 34, pp. 477- 492, 2019.
- [138] A. C. G. Thomé, "SVM Classifiers – Concepts and Applications to Character Recognition, Advances in Character Recognition," IntechOpen, 2012.

- [139] M. Awad ve R. Khanna, "Efficient learning machines: Theories, concepts, and applications for engineers and system Designers," Apress Open, New York, 2015.
- [140] P. R. Nicolas, "Scala for Machine Learning," Packt Publishing Ltd, 2014.
- [141] V. N. Vapnik, "The Nature of Statistical Learning Theory," Springer, New York, 1995.
- [142] K.-R. Müller, S. Mika, G. Ratsch, K. Tsuda ve B. Schölkopf, "An introduction to kernel-based learning algorithms," IEEE Transactions on Neural Networks, vol. 12, pp. 181–201, 2001.
- [143] B. Schölkopf ve A. J. Smola, "Learning with Kernels," MIT Press, 2002.
- [144] D. Huson, "Algorithms in Bioinformatics," WS'06, ZBIT, 2006.
- [145] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
- [146] L. R. Hazım, "Four Classification Methods Naïve Bayesian, Support Vector Machine, K-Nearest Neighbors and Random Forest Are Tested For Credit Card Fraud Detection," Master Thesis, Master of Information Technology, Altınbas University, Istanbul, 2018.
- [147] D. Sabancı, "Rastgele Orman Yaklaşımı Kullanılarak Çok Değişkenli Uyumlu Regresyon Şeritlerinde Model Seçimi," Doktora Tezi, Ondokuz Mayıs Üniversitesi, Fen Bilimleri Enstitüsü, Samsun, 2019.
- [148] M. Immitzer, C. Atzberger ve T. Koukal, "Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data," Remote Sensing, vol. 4, pp. 2661-2693, 2012.
- [149] Y. Chen, S. Zhao, Z. Xie, D. Lu ve E. Chen, "Mapping multiple tree species classes using a hierarchical procedure with optimized node variables and thresholds based on high spatial resolution satellite data," GIScience and Remote Sensing, vol. 57, no. 4, pp. 526-542, 2020.
- [150] A. Hennessy, K. Clarke ve M. Lewis, "Hyperspectral Classification of Plants: A Review of Waveband Selection Generalisability," Remote Sensing, vol. 12, pp. 113, 2020.
- [151] B. Bayraktar, "Rastgele Ormanlar ve Aşırı Öğrenme Makineleri Teknikleri ile Oltalama Saldırılarının Tespiti," Yüksek Lisans Tezi, İstanbul Üniversitesi, Cerrahpaşa Lisansüstü Eğitim Enstitüsü, İstanbul, 2019.
- [152] M. Ercire, "Kısa Süreli Güç Kalitesi Bozulmalarının Dalgacık Analizi ve Rastgele Orman Yönetimi ile Sınıflandırılması," Yüksek Lisans Tezi, Kütahya Dumlupınar Üniversitesi, Fen Bilimleri Enstitüsü, Kütahya, 2019.

- [153] C. Sothe, C. M. De Almeida, M. B. Schimalski, L. E. C. La Rosa, J. D. B. Castro, R. Q. Feitosa, M. Dalponte, C. L. Lima, V. Liesenberg, G. T. Miyoshi ve A. M. G. Tommaselli, "Comparative performance of convolutional neural network, weighted and conventional support vector machine and random forest for classifying tree species using hyperspectral and photogrammetric data," *GIScience and Remote Sensing*, vol. 57, no. 3, pp. 369-394, 2020.
- [154] Y. Wu ve X. Zhang, "Object-Based Tree Species Classification Using Airborne Hyperspectral Images and LiDAR Data," *Forests*, vol. 11, no. 1, pp. 32, 2020.
- [155] Ü. H. Atasever, "Uydu Görüntülerinin Sınıflandırılmasında Hızlandırma (Boosting), Destek Vektör Makineleri, Rastgele Orman (Random Forest) ve Regresyon Ağaçları Yöntemlerinin Kullanılması," Yüksek Lisans Tezi, Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Kayseri, 2011.
- [156] G. James, D. Witten, T. Hastie ve R. Tibshirani, "An Introduction to Statistical Learning," Springer, 2013.
- [157] Ö. Akar, "Rastgele Orman Sınıflandırıcısına Doku Özellikleri Entegre Edilerek Benzer Spektral Özellikteki Tarımsal Ürünlerin Sınıflandırılması," Doktora Tezi, Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Trabzon, 2013.
- [158] J. Watts ve R. Lawrence, "Merging random forest classification with an object-oriented approach for analysis of agricultural lands," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. vol. 37, pp. 579-582, 2008.
- [159] P. Jeatrakul ve K. W. Wong, "Comparing the performance of different neural networks for binary classification problems," In *Eighth International Symposium on Natural Language Processing*, 2009, pp. 111-115.
- [160] D. Anderson ve G. McNeill, "Artificial Neural Networks Technology," A DACS State of-the-Art Report, Contract No. F30602-89-C-0082, ELIN: A011, Rome Laboratory RL/C3C, Griffiss Air Force Base, New York, 1992.
- [161] S. Agatonovic-Kustrin ve R. Beresford, "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 22, no. 5, pp. 717-727, 2000.
- [162] G. Dorffner, "Neural Networks for Time Series Processing," *Neur Ntwk Wor*, vol. 6, no. 4, pp. 447-468, 1996.
- [163] M. Shoaib, A. Y. Shamseldin, B. W. Melville, "Comparative study of different wavelet based neural network models for rainfall-runoff modeling," *Journal of Hydrology*, vol. 515, pp. 47-58, 2014.

- [164] A. P. Piotrowski, M. J. Napiorkowski, J. J. Napiorkowski ve M. Osuch, "Comparing various artificial neural network types for water temperature prediction in rivers," *Journal of Hydrology*, vol. 529, pp. 302-315, 2015.
- [165] S. Lek ve J-F. Guégan, "Artificial neural networks as a tool in ecological modelling: an introduction," *Ecological Modelling*, vol. 120, no. 2-3, pp. 65-73, 1999.
- [166] Y.-S. Park ve S. Lek, "Chapter 7 - Artificial Neural Networks: Multilayer Perceptron for Ecological Modeling," in *Developments in Environmental Modelling*, Sven Erik Jørgensen, Eds, Elsevier, vol. 28, 2016, pp. 123-140.
- [167] S. Nezami, E. Khoramshahi, O. Nevalainen, I. Pölönen ve E. Honkavaara, "Tree Species Classification of Drone Hyperspectral and RGB Imagery with Deep Learning Convolutional Neural Networks," *Remote Sensing*, vol. 12, pp.1070, 2020.
- [168] M. Stone, "Cross-validatory choice and assessment of statistical predictions," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 36, pp. 111-147, 1974.
- [169] Y. Y. Ziggah, H. Youjian, A. R. Tierra ve P. B. Laari, "Coordinate transformation between global and local datums based on artificial neural network with K-Fold cross-validation: a case study, Ghana," *Earth Sciences Research Journal*, vol. 23, pp. 67-77, 2019.
- [170] S. Gholinejad, A. A. Naeini ve A. Amiri-Simkooei, "Robust particle swarm optimization of RFMs for high-resolution satellite images based on K-fold cross-validation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 8, pp. 2594-2599, 2019.

## TEZDEN ÜRETİLMİŞ YAYINLAR

---

### Konferans Bildirileri

1. N. Yastikli ve Z. Cetin, "Detection of Individual Trees in Urban Areas Using Raw LiDAR Data," International Symposium on Applied Geoinformatics (ISAG2019), Istanbul, Turkey, 7 - 09 November 2019.
2. Z. Cetin ve N. Yastikli, "Automatic Detection of Single Street Trees From Airborne LiDAR Data Based on Point Segmentation Methods," International Symposium on Applied Geoinformatics (ISAG2021), Riga, Latvia, 2-3 December 2021. <https://doi.org/10.15659/isag2021.12489>

### Makaleler

1. Z. Cetin ve N. Yastikli, "The Use of Machine Learning Algorithms in Urban Tree Species Classification," ISPRS International Journal of Geo-Information, vol. 11, no. 4, pp. 226, 2022. <https://doi.org/10.3390/ijgi11040226>