T.C.
ERCİYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
HARİTA MÜHENDİSLİĞİ ANABİLİM DALI

ÇOK KANALLI OPTİK GÖRÜNTÜLERİN MAKİNE ÖĞRENMESİ TEKNİKLERİ İLE SINIFLANDIRILMASI VE ANALİZİ

(Yüksek Lisans Tezi)

Hazırlayan
Ali Mahir ÇINAR

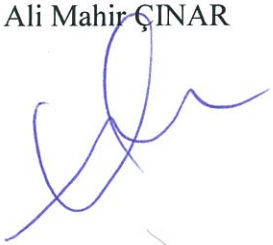
Danışman
Dr. Öğr. Üyesi Ümit Haluk ATASEVER

Temmuz 2019
KAYSERİ

BİLİMSEL ETİĞE UYGUNLUK

Bu çalışmadaki tüm bilgilerin, akademik ve etik kurallara uygun bir şekilde elde edildiğini beyan ederim. Aynı zamanda bu kural ve davranışların gerektirdiği gibi, bu çalışmanın özünde olmayan tüm materyal ve sonuçları tam olarak aktardığımı ve referans gösterdiğimi belirtirim.

Ali Mahir ÇINAR



YÖNERGEYE UYGUNLUK

"Çok Kanallı Optik Görüntülerin Makine Öğrenmesi Teknikleri ile Sınıflandırılması ve Analizi" adlı Yüksek Lisans Tezi, Erciyes Üniversitesi Lisansüstü Tez Önerisi ve Tez Yazma Yönergesi' ne uygun olarak hazırlanmıştır.

Hazırlayan
Ali Mahir ÇINAR

Danışman
Dr. Öğr. Üyesi Ümit Haluk ATASEVER

Harita Mühendisliği ABD Başkanı

Prof. Dr. Abdurrahman EYMEN

KABUL VE ONAY

Dr. Öğr. Üyesi Ümit Haluk ATASEVER danışmanlığında **Ali Mahir ÇINAR** tarafından hazırlanan **"Çok Kanallı Optik Görüntülerin Makine Öğrenmesi Teknikleri ile Sınıflandırılması ve Analizi"** adlı bu çalışma jürimiz tarafından Erciyes Üniversitesi Fen Bilimleri Enstitüsü **Harita Mühendisliği** Anabilim Dalında **Yüksek Lisans** tezi olarak kabul edilmiştir.

01 / 07 / 2019

JÜRİ:

Danışman : Dr. Öğr. Üyesi Ümit Haluk ATASEVER

Üye : Prof. Dr. Erkan BEŞDOK

Üye : Dr. Öğr. Üyesi Ahmet Emin KARKINLI

ONAY:

Bu tezin kabulü Enstitü Yönetim Kurulunun 09/07/2019 tarih ve 2019/40-09 sayılı kararı ile onaylanmıştır.

09 / 07 / 2019

Prof. Dr. Mehmet AKKURT

Enstitü Müdürü

TEŞEKKÜR

Tez süreci boyunca bilgisini ve desteğini hiç esirgemeyen tez danışmanım değerli hocam Sayın Dr. Öğr. Ü. Ümit Haluk ATASEVER' e, jüri olmayı kabul eden saygıdeğer hocalarım Prof. Dr. Erkan BEŞDOK' a ve Dr. Öğr. Üyesi Ahmet Emin KARKINLI' ya ve çalışmalarım boyunca sabırla beni destekleyen aileme en içten teşekkürlerimi sunarım.

Ali Mahir ÇINAR

Temmuz 2019, KAYSERİ

ÇOK KANALLI OPTİK GÖRÜNTÜLERİN MAKİNE ÖĞRENMESİ TEKNİKLERİ İLE SINIFLANDIRILMASI VE ANALİZİ

Ali Mahir ÇINAR

Erciyes Üniversitesi, Fen Bilimleri Enstitüsü
Yüksek Lisans Tezi, Temmuz 2019
Danışman: Dr. Öğr. Üyesi Ümit Haluk ATASEVER

ÖZET

Uzaktan algılama teknolojisindeki son gelişmeler ile yeryüzü hakkında elde ettiğimiz verinin kalitesi önemli bir biçimde artmıştır ve uydu görüntüleri ile yeryüzü hakkında önemli veriler sağlamaktadır. Bu verilerin anlamlandırılabilmesi için tasarlanan sınıflandırma teknikleri gibi metotlar, arazi kullanımı ve değişim analizi gibi uygulamaların hayata geçirilmesine olanak sağlamaktadır.

Sınıflandırma teknikleri ile başarılı sonuçlar elde edebilmek için, doğru algoritma ve parametrelerin ayarlanması önemli bir rol oynamaktadır. Literatürde, yüksek doğruluk değerleri ve etkili kullanımları sebebiyle gelişmiş sınıflandırma teknikleri kullanılmaktadır.

Bu çalışmada çok kanallı optik görüntülerin destek vektör makineleri, komşuluk bileşen analizi, karar ağaçları, derin öğrenme, *k* en yakın komşu, lineer diskriminant gibi gelişmiş makine öğrenmesi teknikleri ile sınıflandırılması hedeflenmiştir ve bu metotların performansları analiz edilmiştir.

Anahtar Kelimeler: Destek Vektör Makineleri, Komşuluk Bileşen Analizi, Karar Ağaçları, Derin Öğrenme, *K* En Yakın Komşu, Lineer Diskriminant, Sınıflandırma

MULTISPECTRAL OPTIC IMAGE CLASSIFICATION AND ANALYSIS WITH MACHINE LEARNING TECHNIQUES

Ali Mahir ÇINAR

Erciyes University, Graduate School of Natural and Applied Sciences
Master Thesis, July 2019
Supervisor: Assist. Prof. Dr. Ümit Haluk ATASEVER

ABSTRACT

Recent developments in remote sensing technologies lead significant improvement in quality of information about the earth and satellite images provide important data about the earth. Method designed for interpretation of such kind of data, such as classification techniques, allows the implementation of land usage and change analysis applications.

In order to obtain successful results in classification, correct algorithm and parameter tuning plays crucial role. Advanced classification techniques are commonly used in the literature due to their high accuracy and effectiveness.

This study aims classification of multispectral optical images with advanced machine learning techniques such as support vector machines, neighborhood component analysis, decision trees, deep learning, *k* nearest neighborhood, linear discriminant and the performance of these methods are analyzed.

Keywords: Support Vector Machine, Neighborhood Component Analysis, Decision Tree, Deep Learning, *K* Nearest Neighborhood, Linear Disciriminant, Classification

İÇİNDEKİLER

ÇOK KANALLI OPTİK GÖRÜNTÜLERİN MAKİNE ÖĞRENMESİ TEKNİKLERİ İLE SINIFLANDIRILMASI VE ANALİZİ

BİLİMSEL ETİĞE UYGUNLUK ii

YÖNERGEYE UYGUNLUK SAYFASI iii

KABUL VE ONAY iv

TEŞEKKÜR v

ÖZET vi

ABSTRACT vii

İÇİNDEKİLER viii

KISALTMALAR xii

TABLOLAR LİSTESİ xiii

ŞEKİLLER LİSTESİ xv

GİRİŞ 1

1. BÖLÜM

GENEL BİLGİLER ve LİTERATÜR ÇALIŞMASI

1.1. Uzaktan Algılama 5

1.2. Uzaktan Algılamanın Temel Bileşenleri 5

1.3. Elektromanyetik Enerji 6

1.4. Uzaktan Algılamada Çözünürlük 7

1.4.1. Mekansal Çözünürlük 7

1.4.2. Spektral Çözünürlük 8

1.4.3. Radyometrik Çözünürlük 9

1.4.4. Zamansal Çözünürlük 10

1.5. Uzaktan Algılamada Uydu Sistemleri 11

1.5.1. Landsat Uydusu 12

1.5.2. Sentinel-2 Uydusu 12

1.6. Uzaktan Algılamada Sınıflandırma 13

1.6.1. Kontrolsüz Sınıflandırma 14

1.6.2. Kontrollü Sınıflandırma 15

1.6.2.1. Paralelkenar Yöntemi 16

1.6.2.2. En Kısa Uzaklık Yöntemi 17

1.6.2.3. Maksimum Olabilirlik Yöntemi 17

2. BÖLÜM

YÖNTEM VE MATERYAL

2.1. Yöntemler 18

2.1.1. Makine Öğrenmesi Teknikleri 18

2.1.2. Destek Vektör Makineleri 18

2.1.2.1. Lineer Destek Vektör Makineleri 20

2.1.2.1.1. Lineer Ayrılabilen Veriler için Destek Vektör Makineleri 20

2.1.2.1.2. Lineer Ayrılamayan Veriler için Destek Vektör Makineleri 22

2.1.2.2. Lineer Olmayan Destek Vektör Makineleri 23

2.1.2.3. Çok Sınıflı Destek Vektör Makineleri 24

2.1.3. Komşuluk Bileşen Analizi 24

2.1.4. Karar Ağaçları 27

2.1.4.1. Karar Ağaçlarının Oluşturulması 28

2.1.4.1.1. Sınıflandırma ve Regresyon Ağaçları 29

2.1.4.1.1.1. Gini İndeksi 29

2.1.4.1.1.2. Twoing Kuralı 30

2.1.4.1.2. Ki-Kare Olasılık 31

2.1.4.2. Karar Ağaçlarının Dallanma Kriterleri 31

2.1.4.2.1. Entropi Bazlı Algoritmalar 32

2.1.4.2.2. Bilgi Kazancı ve Kazanç Oranı 32

2.1.4.3. Karar Ağaçlarının Budanması 33

2.1.4.4. Karar Ağaçlarında Kullanılan Yöntemler 34

2.1.4.4.1. Torbalama Karar Ağaçları Yöntemi 35

2.1.5. Lineer Diskriminant Yöntemi 36

2.1.6. K-En Yakın Komşu Yöntemi 36

2.1.7. Altuzay KNN Yöntemi 37

2.1.8. Derin Öğrenme Yöntemi 37

2.1.8.1. Oto Kodlayıcı 39

2.1.8.2. Softmax 40

2.2. Kullanılan Yazılımlar ve Materyaller 40

3. BÖLÜM

UYGULAMA VE BULGULAR

3.1. Uygulama Aşamaları 42

3.2. Sınıflandırma Sonucu Elde Edilen Değerlerin Matematiksel İfadeleri 43

3.3. Komşuluk Bileşen Analizi ile Bant Ağırlıklarının Bulunması 44

3.4. Anlamlı Sonuç Elde Etmeye Başlanan Bant Sıralamaları 45

3.5. Destek Vektör Makineleri ile Sınıflandırma İşlemleri 46

3.6. Karar Ağaçları ile Sınıflandırma İşlemleri 52

3.6.1. Torbalama Karar Ağaçları ile Sınıflandırma İşlemleri 55

3.7. Lineer Diskriminant ile Sınıflandırma İşlemleri 59

3.8. K-En Yakın Komşu ile Sınıflandırma İşlemleri 62

3.9. Altuzay En Yakın Komşu ile Sınıflandırma İşlemleri 66

3.10. Derin Öğrenme ile Sınıflandırma İşlemleri 69

3.11. Sınıflandırma Sonucu Elde Edilen Değerlerin Yorumu 73

4. BÖLÜM

TARTIŞMA-SONUÇ ve ÖNERİLER

4.1. Tartışma ve Sonuçlar 77

4.2. Öneriler 79

KAYNAKÇA 80

ÖZGEÇMİŞ 85

KISALTMALAR

EMR **:** Elektromanyetik Işınım

ESA **:** Avrupa Uzay Ajansı

KNN **:** K-En Yakın Komşu

NASA **:** National Aeronautical Space Administration

NCA **:** Komşuluk Bileşen Analizi

SVM **:** Destek Vektör Makineleri

TABLOLAR LİSTESİ

Tablo 1.1. Yaygın Kullanılan Uydu Sistemlerinin Özellikleri 11
Tablo 1.2. Landsat 8 Uydusunun Bant Çözünürlükleri 12
Tablo 1.3. Sentinel-2 Uydusunun Tayf Özellikleri ve Uygulama Alanları 13
Tablo 3.1. Akış Şeması 42
Tablo 3.2. Anlamlı Sonuç Elde Etmeye Başlanan Bant Sıralamaları 46
Tablo 3.3. SVM ile Sınıflandırılmış Landsat 8 (Bant:5,1,7) ve Sentinel 2 (Bant: 8,12, 7,5,11) Görüntüleri Doğruluk Değerleri 49
Tablo 3.4. SVM ile Sınıflandırma Doğruluk Değerleri 51
Tablo 3.5. Karar Ağaçları ile Sınıflandırılmış Landsat 8 (Bant:5,1,7) ve Sentinel 2 (Bant:8,12,7,5,11) Görüntüleri Doğruluk Değerleri 53
Tablo 3.6. Karar Ağaçları ile Sınıflandırma Doğruluk Değerleri 55
Tablo 3.7. Torbalama Karar Ağaçları ile Sınıflandırılmış Landsat 8 (Bant:5,1) ve Sentinel 2 (Bant:8,12) Görüntüleri Doğruluk Değerleri 57
Tablo 3.8. Torbalama Karar Ağaçları ile Sınıflandırma Doğruluk Değerleri 58
Tablo 3.9. Lineer Diskriminant ile Sınıflandırılmış Landsat 8 (Bant:5,1,7,6) ve Sentinel 2 (Bant:8,12,7,5,11) Görüntüleri Doğruluk Değerleri 60
Tablo 3.10. Lineer Diskriminant ile Sınıflandırma Doğruluk Değerleri 62
Tablo 3.11. KNN ile Sınıflandırılmış Landsat 8 (Bant:5,1) ve Sentinel 2 (Bant:8,12) Görüntüleri Doğruluk Değerleri 64
Tablo 3.12. KNN ile Sınıflandırma Doğruluk Değerleri 65
Tablo 3.13. Altuzay KNN ile Sınıflandırılmış Landsat 8 (Bant:5,1,7,6) ve Sentinel 2 (Bant:8,12,7,5,11) Görüntüleri Doğruluk Değerleri 67
Tablo 3.14. Altuzay KNN ile Sınıflandırma Doğruluk Değerleri 69
Tablo 3.15. Derin Öğrenme ile Sınıflandırılmış Landsat 8 (Bant:5,1,7) ve Sentinel 2 (Bant:8,12,7,5,11) Görüntüleri Doğruluk Değerleri 71
Tablo 3.16. Derin Öğrenme ile Sınıflandırma Doğruluk Değerleri 72
Tablo 3.17. Landsat 8 Görüntüsünün Yöntemlere Göre Farklı Bant Kombinasyonlarının Doğruluk Değerleri 73

Tablo 3.18. Sentinel 2 Görüntüsünün Yöntemlere Göre Farklı Bant Kombinlerinin Doğruluk Değerleri 74
Tablo 3.19. Landsat 8 Görüntüsü için Yöntemlere Göre Doğruluk Değerleri 75
Tablo 3.20. Sentinel 2 Görüntüsü için Yöntemlere Göre Doğruluk Değerleri 76

ŞEKİLLER LİSTESİ

Şekil 1.1. Uzaktan Algılama Temel Bileşenleri 6
Şekil 1.2. Elektromanyetik Dalga Karakteri 7
Şekil 1.3. Farklı Çözünürlükteki Görüntüler: (a) 1 metre ve (b) 10 metre 8
Şekil 1.4. Spektral Çözünürlük 9
Şekil 1.5. Radyometrik Çözünürlük 10
Şekil 1.6. Sınıflandırma 14
Şekil 1.7. Kontrolsüz Sınıflandırma 15
Şekil 1.8. Kontrollü Sınıflandırma 16
Şekil 1.9. Paralelkenar Yöntemi 17
Şekil 2.1. Hiper Düzlemler 19
Şekil 2.2. Destek Vektörleri 19
Şekil 2.3. Lineer Olarak Ayrılabilen Veriler için Hiper Düzlem 21
Şekil 2.4. Lineer Olarak Ayrılamayan Veriler 22
Şekil 2.5. Lineer Olarak Ayrılamayan Veriler için Hiper Düzlem 23
Şekil 2.6. Lineer Olmayan Verilerin Farklı Bir Uzayda Ayrılması 23
Şekil 2.7. Karar Ağaçları 28
Şekil 2.8. Karar Ağaçlarının Oluşturulması 29
Şekil 2.9. Kararların Birleştirilmesi 35
Şekil 2.10. Torbalama Karar Ağaçları Yöntemi 36
Şekil 2.11. K-En Yakın Komşuluk Yöntemi 37
Şekil 2.12. Derin Öğrenme 38
Şekil 2.13. İlk Derin Öğrenme Mimarisi 38
Şekil 2.14. Kodlayıcı ve Kod Çözücü Yapısı 39
Şekil 2.15. Landsat 8 Uydu Görüntüsü 40
Şekil 2.16. Sentinel 2 Uydu Görüntüsü 41
Şekil 3.1. Sınıflandırmaya Etki Eden Landsat 8 Bantları 44
Şekil 3.2. Sınıflandırmaya Etki Eden Sentinel 2 Bantları 45

Şekil 3.3. SVM ile Sınıflandırılmış Landsat 8 (Bant:5,1,7) Görüntüsü (Kernel: Gaussian) 47
Şekil 3.4. SVM ile Sınıflandırılmış Landsat 8 (Bant:5,1,7) Görüntüsü (Kernel: Lineer) 47
Şekil 3.5. SVM ile Sınıflandırılmış Sentinel 2 (Bant: 8,12,7,5,11) Görüntüsü (Kernel: Gaussian) 48
Şekil 3.6. SVM ile Sınıflandırılmış Sentinel 2 (Bant: 8,12,7,5,11) Görüntüsü (Kernel: Lineer) 48
Şekil 3.7. SVM ile Sınıflandırılmış Landsat 8 Görüntüsü(Kernel: Gaussian) 49
Şekil 3.8. SVM ile Sınıflandırılmış Landsat 8 Görüntüsü(Kernel: Linear) 50
Şekil 3.9. SVM ile Sınıflandırılmış Sentinel 2 Görüntüsü(Kernel: Gaussian) 50
Şekil 3.10. SVM ile Sınıflandırılmış Sentinel 2 Görüntüsü(Kernel: Linear) 51
Şekil 3.11. Karar Ağaçları ile Sınıflandırılmış Landsat 8 (Bant:5,1,7) Görüntüsü 52
Şekil 3.12. Karar Ağaçları ile Sınıflandırılmış Sentinel 2 (Bant:8,12,7,5,11) Görüntüsü 53
Şekil 3.13. Karar Ağaçları ile Sınıflandırılmış Landsat 8 Görüntüsü 54
Şekil 3.14. Karar Ağaçları ile Sınıflandırılmış Sentinel 2 Görüntüsü 54
Şekil 3.15. Torbalama Karar Ağaçları ile Sınıflandırılmış Landsat 8 (Bant:5,1) Görüntüsü 56
Şekil 3.16. Torbalama Karar Ağaçları ile Sınıflandırılmış Sentinel 2 (Bant:8,12) Görüntüsü 56
Şekil 3.17. Torbalama Karar Ağaçları ile Sınıflandırılmış Landsat 8 Görüntüsü 57
Şekil 3.18. Torbalama Karar Ağaçları ile Sınıflandırılmış Sentinel 2 Görüntüsü 58
Şekil 3.19. Lineer Diskriminant ile Sınıflandırılmış Landsat 8 (Bant:5,1,7,6) Görüntüsü 59
Şekil 3.20. Lineer Diskriminant ile Sınıflandırılmış Sentinel 2 (Bant:8,12,7,5,11) Görüntüsü 60
Şekil 3.21. Lineer Diskriminant ile Sınıflandırılmış Landsat 8 Görüntüsü 61
Şekil 3.22. Lineer Diskriminant ile Sınıflandırılmış Sentinel 2 Görüntüsü 61
Şekil 3.23. KNN ile Sınıflandırılmış Landsat 8 (Bant:5,1) Görüntüsü 63

Şekil 3.24. KNN ile Sınıflandırılmış Sentinel 2 (Bant:8,12) Görüntüsü 63
Şekil 3.25. KNN ile Sınıflandırılmış Landsat 8 Görüntüsü 64
Şekil 3.26. KNN ile Sınıflandırılmış Sentinel 2 Görüntüsü 65
Şekil 3.27. Altuzay KNN ile Sınıflandırılmış Landsat 8 (Bant:5,1,7,6) Görüntüsü 66
Şekil 3.28. Altuzay KNN ile Sınıflandırılmış Sentinel 2 (Bant:8,12,7,5,11) Görüntüsü 67
Şekil 3.29. Altuzay KNN ile Sınıflandırılmış Landsat 8 Görüntüsü 68
Şekil 3.30. Altuzay KNN ile Sınıflandırılmış Sentinel 2 Görüntüsü 68
Şekil 3.31. Derin Öğrenme ile Sınıflandırılmış Landsat 8 (Bant:5,1,7) Görüntüsü 70
Şekil 3.32. Derin Öğrenme ile Sınıflandırılmış Sentinel 2 (Bant:8,12,7,5,11) Görüntüsü 70
Şekil 3.33. Derin Öğrenme ile Sınıflandırılmış Landsat 8 Görüntüsü 71
Şekil 3.34. Derin Öğrenme ile Sınıflandırılmış Sentinel 2 Görüntüsü 72

GİRİŞ

Yeryüzü üzerinde meydana gelen değişikliklerin incelenebilmesi, doğal kaynakların etkin kullanımı, yaşam alanlarının gelişim analizi ve daha birçok uygulama için yeryüzüne ait doğruluk değeri yüksek, geniş çapta bilgi sunabilecek hassas ölçekte verilere ihtiyaç duyulmaktadır. Uydu görüntüleri içerdiği bilgi düzeyi ve görüntü işleme aşamalarından sonra elde edilen anlamlı bilgiler açısından sınıflandırma uygulamalarında kullanılan önemli bir altlıktır[1].

Yüzey alanı geniş ve zorlu arazi yapısına sahip alanlarda arazi ile ilgili bilgi toplamak son derece zahmetlidir. Uzaktan algılama ile görüntüleme teknolojisi geliştikten sonraki dönemde arazi yapısı, arazi kullanımı ve ihtiyaç duyulan diğer uygulamalarda istenilen bilgilerin hızlı ve doğru bir biçimde elde edilmesi kolay hale gelmiştir. Ham veri olarak elde edilmiş uydu görüntüleri üzerinden yeryüzüne ait bilgilerin temin edilebilmesi için çeşitli sınıflandırma yöntemleri kullanılmaktadır[1].

Uzaktan algılamada sınıflandırma, yeryüzüne ait optik görüntülerin içerdiği her bir farklı yapının istenen sınıf sayısında özellik gruplarına ayrılmasını içeren işlemler bütündür[2]. Sınıflandırma işleminin temel mantığı yeryüzü üzerindeki benzer yansıtım değerlerine sahip nesnelerin gruplandırılmasıdır. Başka bir deyişle, görüntünün içerisindeki her bir pikselin bant değerleri karşılaştırılarak benzer piksellerin sınıflara ayrılması işlemidir[3]. Sınıflandırma işlemlerinde genel olarak kontrollü sınıflandırma ve kontrolsüz sınıflandırma olmak üzere iki temel yaklaşım kullanılır. Kontrollü sınıflandırma eğitim veri seti kullanılarak sınıflandırma işlemini gerçekleştirdiği için kontrolsüz sınıflandırma işlemine göre genellikle daha yüksek doğruluk değeri elde eder. Sınıflandırma yapılacak bölgenin analizinde doğruluk değerleri, sınıflandırmada uygulanacak yöntemler, sınıf sayısı, uydu görüntüsü kalitesi gibi parametrelerle doğrudan ilişkilidir[4].

Uzaktan algılanmış görüntülerin sınıflandırılmasında ileri sınıflandırma algoritmalarının kullanımı literatürde önemli bir yer kaplamaktadır. Yüksek sınıflandırma yetenekleri ve sınıflandırma sonucu elde edilen doğruluk değerlerinin yanında matematiksel ifadeleri karmaşık olan bu yöntemler parametrik olmayan yapılarıyla sınıflandırma işlemlerinin temelini oluşturmaktadır[2]. Uzaktan algılama sistemleriyle elde edilmiş görüntülerin sınıflandırma işlemlerinde komşuluk bileşen analizi ile bantların sınıflandırmaya etkileri araştırılacaktır ayrıca destek vektör makineleri, karar ağaçları, torbalama karar ağaçları, lineer diskriminant, k-en yakın komşu, altuzay k-en yakın komşu ve derin öğrenme gibi güçlü ve genelleştirme yeteneği yüksek olan sınıflandırma yöntemleri kullanılacaktır.

Destek vektör makineleri istatistiksel öğrenme mantığına dayalı Cortes ve Vapnik tarafından geliştirilmiş bir ileri sınıflandırma yöntemidir[5]. Destek vektör makineleri iki sınıflı lineer verilerin sınıflandırılması problemleri için tasarlanmış, daha sonraki aşamada sınıf sayısı daha fazla olan ve lineer olmayan verilerin sınıflandırılması için kullanılmıştır. Destek vektör makinelerinin temel amacı iki sınıfı birbirinden ayırabilen en uygun aşırı düzlemin tahmin edilebilmesi, iki sınıfı birbirinden optimum şekilde ayırabilen bir düzlemin tanımlanabilmesi esasına dayanır[6].

Destek vektör makinelerini ileri sınıflandırmada kullanılan diğer algoritmalardan farklı kılan en önemli özelliği; matematiksel yapısı gereği sınıflandırma işlemi tamamlandığında sınıflandırma sonucu beklenen riskin üst limitinin minimum olması prensibine dayalı yapısal risk azalımı modelini bünyesinde bulundurmasıdır[7].

Karar ağaçları; uydu görüntülerinin ve işaretlerin sınıflandırılması, karakter tanıma gibi uygulamalarda geniş bir kullanım alanı olan lineer bir sınıflandırma yöntemidir[8]. Karar ağaçları yönteminin bu kadar yaygın kullanılmasının temel sebebi ise doğruluk değeri yüksek sonuçlar verebilmesine karşın kolay anlaşılabilir ve yorumlanabilir olmasıdır[9]. Bu yöntemin temelinde "eğer-ise" şartlarından oluşan kurallar bütünü şeklini alır[10].

Karar ağaçları modelinde bir ağaç yapısı düğüm, dal ve yaprak olmak üzere üç ana kısımdan oluşmaktadır. Her bir öznitelik değeri bu ağaç yapısında bir düğüm tarafından temsil edilir. Ağaç yapısında üst kısım kök, son kısım ise yaprak olarak isimlendirilir[11]. Veri seti tarafından oluşturulan karar iskeletine göre aşağıdan

yukarıya, yukarıdan aşağıya, melez, büyüme - budama yaklaşımları olmak üzere sınıflandırma işlemlerine devam edilir.

Komşuluk bileşen analizi; k en yakın komşu(KNN) yöntemine benzer şekilde çalışan bir yöntemdir. Mahalanobis uzaklığını sınıflandırma sırasında en yüksek performans verecek şekilde öğrenme temelli olarak bulmaya çalışan bir yöntemdir[12]. Komşuluk bileşen analizinin başarılı bir algoritma olmasının sebeplerinden bir tanesi eğitim seti üzerinde sınıflandırma doğruluğunun stokastik varyantını en yükseğe çıkaracak doğrusal bir dönüşüm öğrenmesidir. Bu dönüşüm komşuluk bileşen analizinde sınıflandırma yapılırken uzaklıklar hesaplandığında sınıflandırma hatasını en aza indirgeyen doğrusal dönüşümü temsil eder. Bu yöntem ile yapılacak sınıflandırma çalışması için kullanılan bantların sınıflandırmaya etkilerinin belirlenmesi hedeflenmektedir ve bu sayede görüntüye ait bantların etkileri araştırılmaktadır.

Komşuluk bileşen analizinde bulunan stokastik sınıflandırıcılar sınıflandırma işleminde eğitim verilerinin çoğunluk değerlerine göre sınıflandırmaktadır. Özellikle her bir sorgu için eğitim verilerindeki referans örnekleri, daha uzak örneklerin yakınında bulunan örnekleri destekleyen bir olasılık değerinden alınır[13].

Derin öğrenme çok katmanlı yapay sinir ağlarının özelliklerini kendinde toplayarak sınıflandırma işlemlerini gerçekleştiren bir makine öğrenmesi tekniğidir. Yöntem önceden oluşturulmuş matematiksel modele bağlı öznitelik çıkartıcıya bağlı kalmamaktadır ayrıca bu yönteme ait en önemli avantajlardan biri genelleştirilebilir olmasıdır[14].

Sınıflandırma işlemlerinin doğruluğu, sınıflandırmada kullanılacak yöntemlere ve sınıflandırmaya etki eden diğer faktörlere bağlıdır. Sınıflandırma yöntemleriyle problemin çözümü ve analizi için sınıflandırılacak görüntüye ait eğitim ve test verileri kullanılacak olup, eğitim setindeki her bir örneğin öznitelik bilgileri ve bu bilgilere ait sınıf değerleri belirlenecektir[4].

Bu çalışmada Destek vektör makineleri, Karar ağaçları, Torbalama karar ağaçları, Lineer diskriminant, K-en yakın komşu, Altuzay k-en yakın komşu ve Derin öğrenme yöntemleri kullanılarak sınıflandırma işlemleri yapılacaktır.

Yapılacak sınıflandırma çalışmaları sonucunda elde edilecek bilgilere bağlı olarak; yöntemler arasındaki performans farklarının, yöntemlerin sınıflandırma başarılarının, görüntülerin sahip oldukları bantların sınıflandırmaya katkısının, görüntü çözünürlüklerinin sınıflandırmaya etkisinin belirlenmesi hedeflenmiştir.

1. BÖLÜM

GENEL BİLGİLER ve LİTERATÜR ÇALIŞMASI

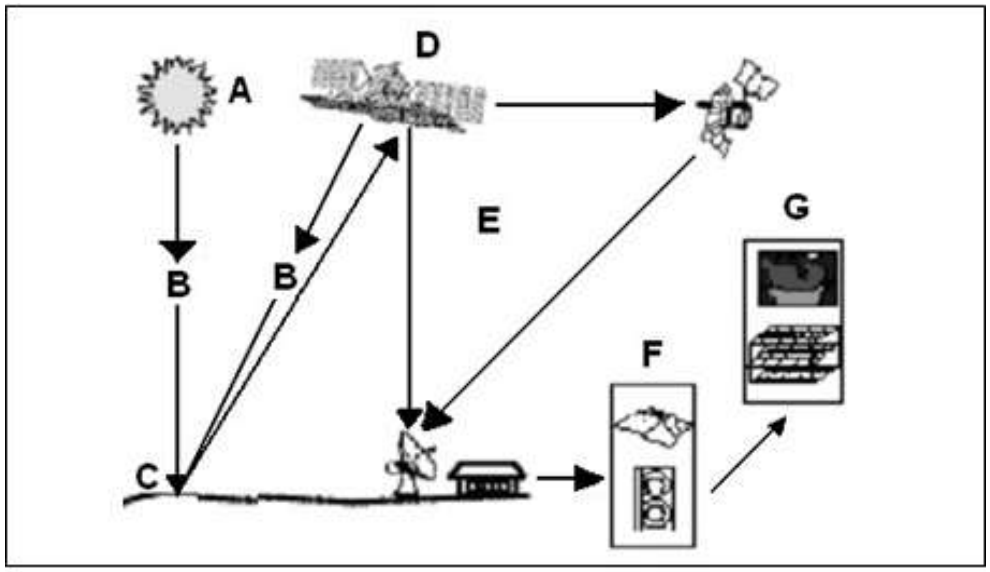
1.1. Uzaktan Algılama

Uzaktan algılama, bir cisim ile arada fiziksel etkileşim olmadan cisimden yansıyan ve yayılan elektromanyetik enerjinin yorumlanması ile bu cismin nicelik ve nitelik özellikleri yönünden incelenerek bilgi elde edilmesi işlemleridir[15]. Çekilen ilk hava fotoğrafları uzaktan algılama çalışmalarının da ilk verileri arasında yer almaktadır. Teknolojinin gelişmesiyle uzaya gönderilen uydulardan elde edilen görüntüler, gelişmiş kamera ve algılayıcılar ile birlikte ise yeryüzü şekillerinin incelenmesi, arazi kullanım faaliyetleri, doğal kaynak tüketiminin incelenmesi, doğal afetlerin meydana getirdiği etkiler gibi önemli verileri gözlemlemek uzaktan algılama çalışmalarının temel amacını oluşturmaktadır[15].

Aktif algılama ve pasif algılama sistemleri olmak üzere uzaktan algılamada iki sistem kullanılmaktadır. Pasif algılama sistemleri güneş gibi bir enerji kaynağı veya yeryüzünün doğal yayılım enerjisine ihtiyaç duyarken aktif algılama sistemleri kendi enerjisini kullanarak hedeften yansıyan enerjiyi gözlemlemektedir. Aktif algılama sistemleri bu özelliğinden dolayı yılın herhangi bir zamanını beklemeksizin görüntü elde edebilir[16].

1.2. Uzaktan Algılamanın Temel Bileşenleri

Uzaktan algılama sürecinin 7 temel bileşeni bulunmaktadır. Bu bileşenler; enerji kaynağı(A), ışınım ve atmosfer(B), yeryüzü ile etkileşimi(C), enerjinin kaydedilmesi(D), verinin iletimi ve işlenmesi(E), veri yorumlaması ve analizi(F) ve uygulama aşaması(G) olarak sıralanmaktadır[17]. Şekil 1.1. 'de uzaktan algılamanın temel bileşenleri gösterilmiştir.



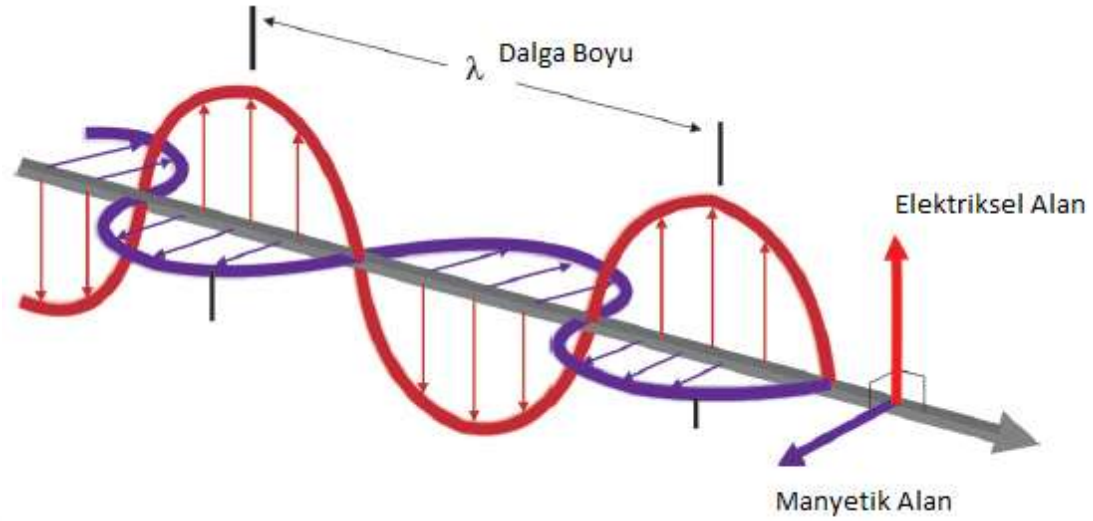
Şekil 1.1. Uzaktan Algılama Temel Bileşenleri

1.3. Elektromanyetik Enerji

Uzaktan algılama sistemleri ile cisimlerin incelenebilmesi için bir enerji kaynağına sahip olmak gerekmektedir. Uzaktan algılama için en önemli enerji kaynağı güneştir. Radyo ve televizyon dalgası, morötesi ve X ışınları, diğer elektromanyetik enerji kaynakları arasında yer almaktadır[18]. Elektromanyetik enerji, uzaktan algılama temel bileşenleri arasındaki bağlantıyı oluşturmaktadır.

Elektromanyetik enerji sahip olduğu dalga ve parçacık karakterleriyle uzaktan algılama için önemli bir yere sahiptir. Elektromanyetik enerjinin dalga karakteriyle elektromanyetik spektrum, parçacık karakteriyle ise enerjinin yeryüzü ve atmosferle etkileşimi açıklanmaktadır[4].

Elektromanyetik enerji, sürekli tekrar eden elektromanyetik dalgalar şeklinde c ışık hızı ile yeryüzüne ulaşmaktadır ve uzaktan algılama sistemlerinde enerjinin ışıma ile taşınması elektromanyetik ışınım(EMR) olarak adlandırılmaktadır[2]. EMR uzaktan algılama sistemlerinin görüntü elde edebilmesinin temelini oluşturur. Algılayıcı sensörler ile süzgeçten geçirilip kaydedilen elektromanyetik ışın atmosferde yol alırken atmosfer ile etkileşim içine girmektedir. Bu etkileşimler; yansıma, saçılma ve yutulma olarak sınıflandırılabilir[4]. Şekil 1.2. 'de elektromanyetik dalga boyunun karakteristiği gösterilmiştir.



Şekil 1.2. Elektromanyetik Dalga Karakteri

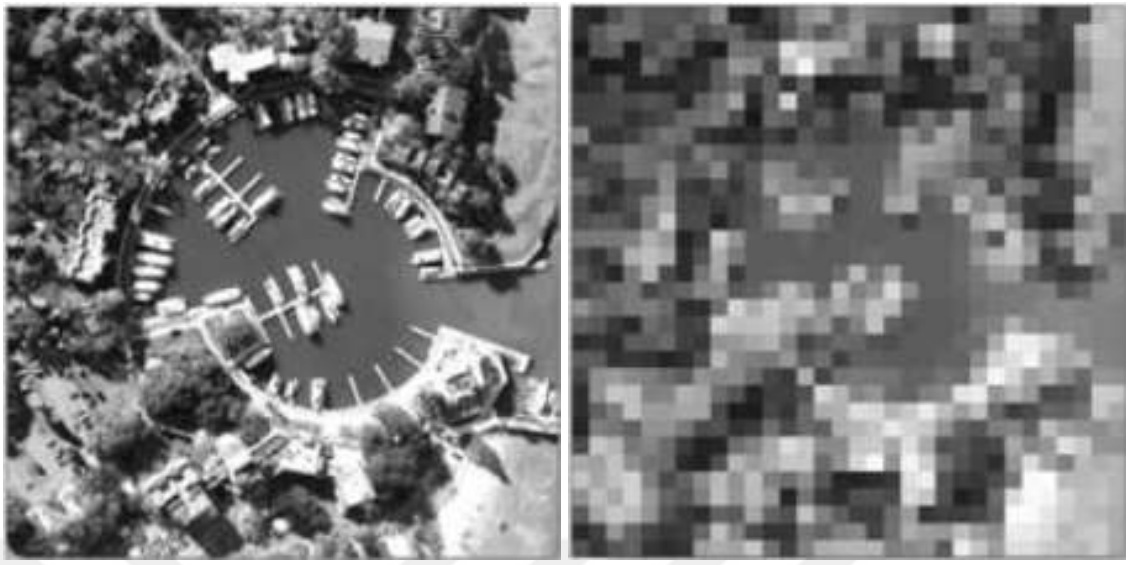
1.4. Uzaktan Algılamada Çözünürlük

Uzaktan algılama çalışmaları için uydular ve diğer görüntü alma sistemleri sayesinde elde edilen görüntülerin ekran çıktısındaki bir piksel değerinin yeryüzünde kapladığı alan mekansal çözünürlük olarak tanımlanmaktadır. Bu pikseller görüntünün en küçük birimini oluşturmaktadır.

Uzaktan algılamada mekansal, spektral, radyometrik ve zamansal çözünürlük olmak üzere dört farklı çözünürlük söz konusudur.

1.4.1. Mekansal Çözünürlük

Elde edilen görüntülerde ayırt edilebilir detay seviyesi mekansal çözünürlük kavramı ile açıklanmaktadır. Bir diğer ifade ile elde edilen görüntüde detaylar mekansal çözünürlüğe bağlıdır ve pikselin yeryüzünde kapladığı alana karşılık gelmektedir[15]. Mekansal çözünürlüğü 1 m olan uydudan alınan görüntü, mekansal çözünürlüğü 10 m olan uydudan alınan görüntüden daha fazla ayırt edilebilir detay seviyesine sahip olmaktadır.



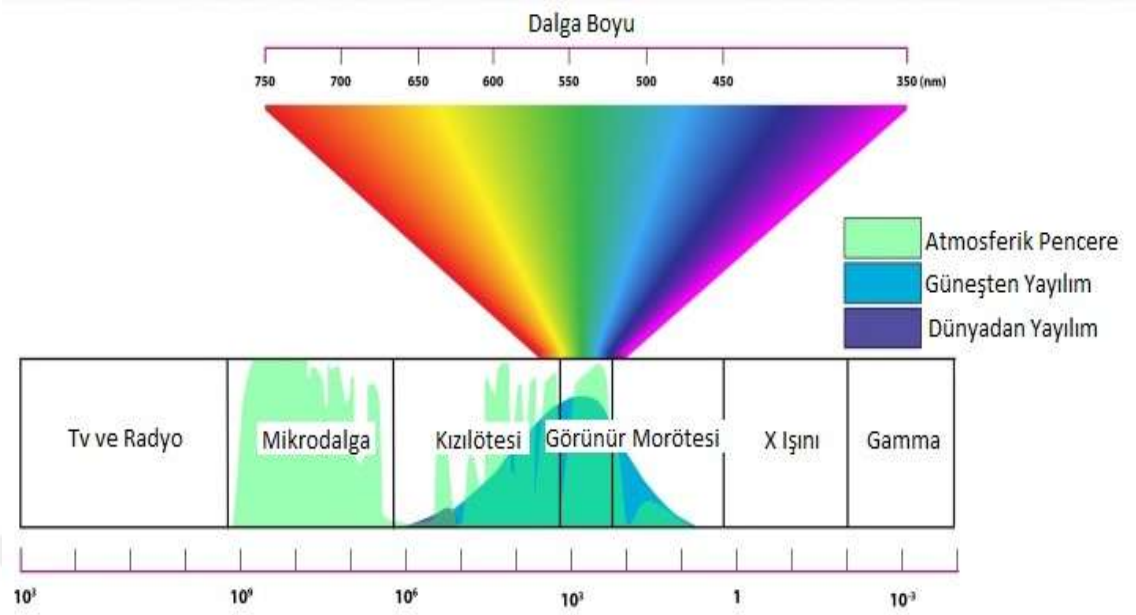
(a) (b)

Şekil 1.3. Farklı Çözünürlükteki Görüntüler: (a) 1 metre ve (b) 10 metre

Şekil 1.3. 'de görüldüğü üzere yüksek çözünürlüğe sahip görüntüler daha yüksek doğruluk değerlerine sahiptir. Bu durum analiz işlemlerinde çıktı ürünün kalitesini doğrudan etkilemektedir.

1.4.2. Spektral Çözünürlük

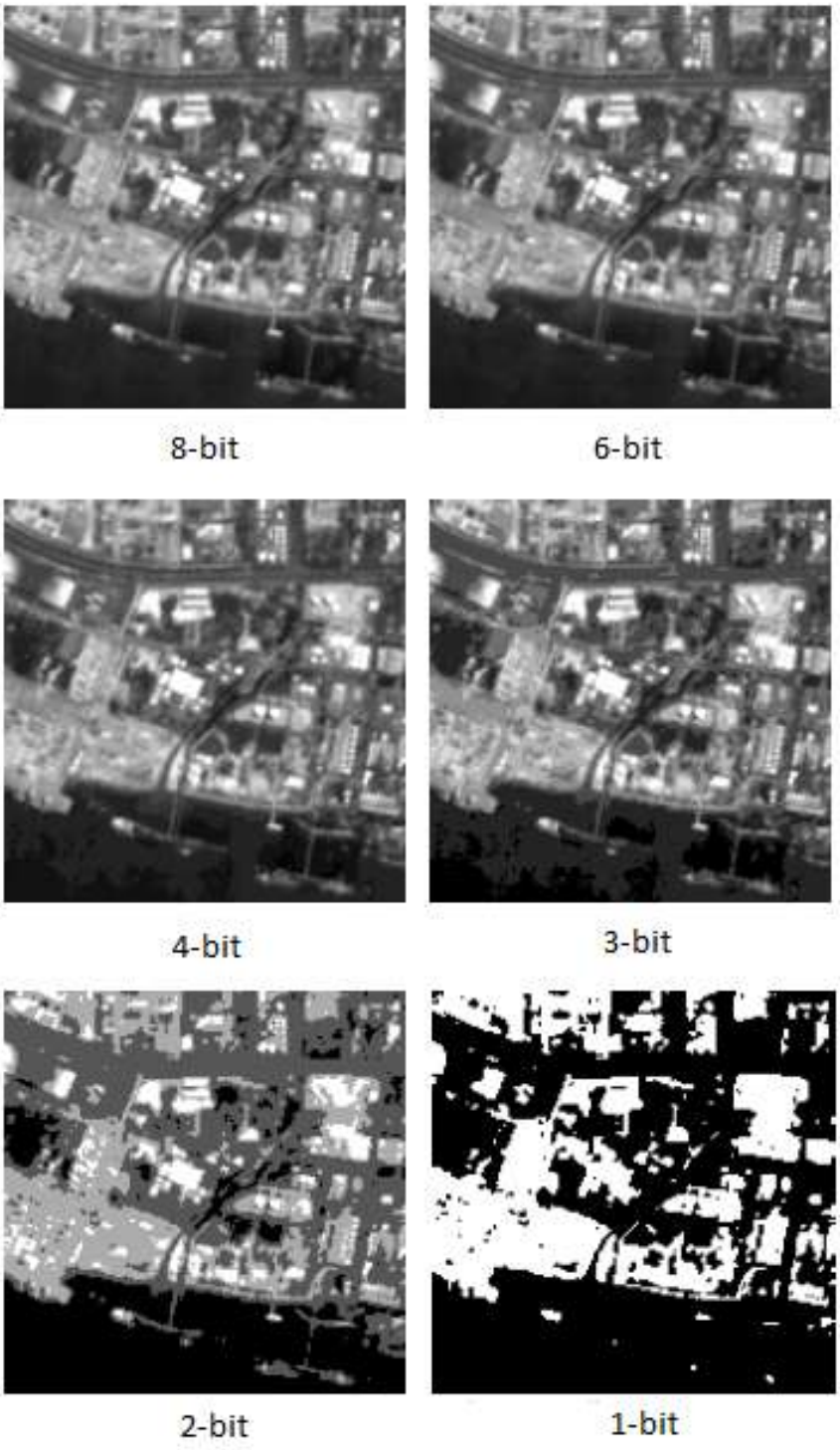
Elektromanyetik spektruma ait farklı bölgelerinde uzaktan algılama sistemleri tarafından algılanan farklı dalga aralıklarını ve yansıtım özelliklerini spektral çözünürlük kavramı açıklamaktadır. Spektral çözünürlük yeryüzündeki ölçümü yapılacak cisim veya arazi yapılarının belirlenebilmesinde önemli bir yere sahiptir[2]. Her cismin sahip olduğu spektral özellikler sayesinde bir görüntüdeki detaylar farklı cisim sınıflarının farklı dalga boyu aralıklarına verdikleri tepkiler sayesinde gözlemlenmektedir. Algılayıcıların tasarımı da bu özellik göz önünde bulundurularak yapılmaktadır. Örnek olarak bir bitki örtüsü ve su içeren bölümler gibi genel arazi örtüsü kısımları yakın kızıl ötesi ve görünür dalga boylarında ayırt edilebilmektedir[15]. Görüntüleme sistemlerine ait spektral çözünürlük uydunun görüntü verebildiği bant sayısına ve görüntü alınan bantların genişliğine bağlıdır[19]. Şekil 1.4. 'de dalga boyuna göre bölgeler gösterilmiştir.



Şekil 1.4. Spektral Çözünürlük

1.4.3. Radyometrik Çözünürlük

Radyometrik çözünürlük, algılayıcı sistemler tarafından elde edilen görüntülerden gelen elektromanyetik ışınımdaki farklılıkları ayırt edebilme özelliğini ifade etmektedir. Algılama sisteminin elektromanyetik ışınımdaki farkları ayırt edebilme hassasiyeti, sistemin radyometrik çözünürlüğünün yüksekliğine bağlıdır[20]. Sistemden sağlanan görüntü üzerindeki her bir piksele karşılık gelen bölümlerden spektral değerine göre farklı bir veri kaydedilmektedir. Bu kaydedilen verilere o algılayıcı sistemin radyometrik çözünürlüğü denilmektedir[2]. Uydu görüntülerinde bu değerler parlaklığın 0 ile 2^n-1 değerleri arasında tanımlanmaktadır. 8bit çözünürlüğünde bir uydu görüntüsünde her bir piksel 0 ile 255 arasında değer almaktadır. Şekil 1.5. 'de radyometrik çözünürlüğü farklı görüntüler verilmiştir.



Şekil 1.5. Radyometrik Çözünürlük

1.4.4. Zamansal Çözünürlük

Dünya çevresinde belirli bir yörüngeye uygun olarak hareket eden uyduların bir bölgeye ait aldığı görüntüdeki zamansal çözünürlük, algılama sisteminin aynı bölge görüntüsünü

ne sıklıkla aldığını belirtmektedir. Uzaktan algılama çalışmalarındaki temel amaçların başında gelen iki farklı zaman dilimi arasında bölgede oluşan farkların incelenmesi zamansal çözünürlüğün önemini ortaya koymaktadır[19].

1.5. Uzaktan Algılamada Uydu Sistemleri

Bir çok farklı disiplinin; yeryüzünün yapısı, arazi kullanımı, arazide meydana gelen değişimler ve doğal kaynaklar hakkında uzaktan algılama sistemleri ile bilgi sahibi olmak istemesi ile birlikte uydu sistemleri gelişmeye başlamıştır. İlk olarak büyük arazileri görüntülemek için kullanılan ve çözünürlükleri kilometre hassasiyetinde olan uydu sistemlerinden faydalanılırken, geliştirilen teknolojiler sonrasında topoğrafik harita üretimi, konum belirleme gibi yüksek çözünürlük ve hassasiyet isteyen uygulamalarda geliştirilmiş uydu sistemleri kullanılmaya başlanmıştır[21]. Tablo 1.1. 'de bazı uyduların özellikleri açıklanmıştır.

Tablo 1.1. Yaygın Kullanılan Uydu Sistemlerinin Özellikleri[14]

Uydu	Mekansal Çözünürlük	Radyometrik Çözünürlük	Spektral Bant Sayısı	Çözünürlük Spektral Bölge	Zamansal Çözünürlük
Landsat-5 TM	30m 30m 120m	8bit	7	Görünür, YKö Kısa Dalga Kö Isıl Kö	16 gün
Landsat-7 ETM+	15m 30m 30m 60m	8bit	8	Pankromatik Görünür, YKö Kısa Dalga Kö Isıl Kö	16 gün
Spot-5 HRG	5m 10m 20m	8bit	5	Pankromatik Görünür, YKö Kısa Dalga Kö	2-3 gün
QuickBird	0.6m 2.4m	11bit	5	Pankromatik Görünür, YKö	1-4 gün
Ikonos	1m 4m	11bit	5	Pankromatik Görünür, YKö	3 gün
GoEye-1	0.5m 1.7m	11bit	5	Pankromatik Görünür, YKö	2-8 gün
Worldview-2	0.5m 1.8m	11bit	9	Pankromatik Görünür, YKö	1-4 gün

*YKö: Yakın Kızıl Ötesi, Kö: Kızıl Ötesi

1.5.1. Landsat Uydusu

Ulusal Havacılık ve Uzay Dairesi(NASA) tarafından yeryüzündeki doğal kaynakları araştırmak amacıyla 1972 yılında Landsat uydularının ilki uzaya fırlatılmıştır. Landsat uydularında dönel ışınlı kamera(RBV), çok bantlı tarayıcı(MSS), tematik harita yapımı(TM) sistemleri olmak üzere üç görüntüleme sistemi kullanılmıştır[21]. Landsat-1,2,3 uydularında RBV ve MSS kullanılmışken, Landsat- 4,5'te ise MSS ve TM algılayıcıları kullanılmaya başlanmıştır. Geliştirilen bu sistemlerle Landsat uydularının ayırma gücü termal bantta 120 metre kapasitesine çıkmıştır. Tablo 1.2. 'de Landsat 8 bant çözünürlükleri gösterilmiştir.

Tablo 1.2. Landsat 8 Uydusunun Bant Çözünürlükleri[21]

Bant Numarası	Landsat 8 Bantları	Konumsal Çözünürlük
1	Kıyı	30
2	Mavi	30
3	Yeşil	30
4	Kırmızı	30
5	Yakın Kızıl Ötesi	30
6	Kısa Dalga Kızıl Ötesi	30
7	Kısa Dalga Kızıl Ötesi	30
8	Pankromatik	15
9	Sirrus	30
10	TIRS 1	100
11	TIRS 2	100

1.5.2. Sentinel-2 Uydusu

Avrupa Uzay Ajansı(ESA) tarafından 2015 yılında uzaya gönderilmiştir. Sentinel-2 uydu görüntüleri, görünür bölge, kısa dalga kızıl ötesi ve yakın kızıl ötesi bölgelerinde farklı mekansal çözünürlükte 13 kanala sahip multispektral verilerdir. Tablo 1.3. 'de Sentinel 2 uydu görüntüsüne özellikler verilmiştir.

Tablo 1.3. Sentinel-2 Uydusunun Tayf Özellikleri ve Uygulama Alanları[23]

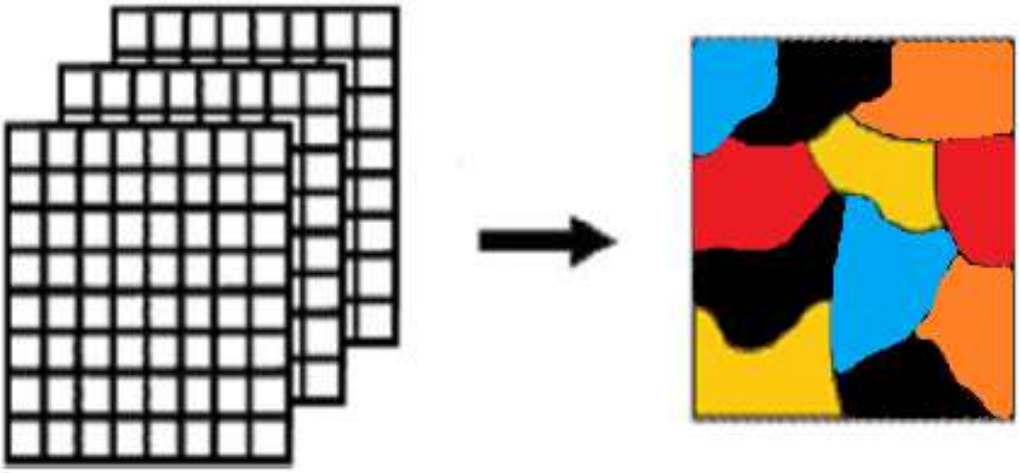
Tayf Kanalları	Uygulama Alanları
B1, B2, B12	Aerosol Giderimi
B8, B8a, B9	Su Buharı Giderimi
B10	Circus Tanımlaması
B2, B3, B4	Yüzey Örtüsü Sınıflandırması
B5, B6, B7	Yaprak Klorofili ve Su Miktarı
B8, B8a, B11, B12	Buz, Kar ve Mineral Tanımlama

1.6. Uzaktan Algılamada Sınıflandırma

Uzaktan algılama sistemleri ile işlenmemiş halde elde edilen, farklı mekansal, spektral, radyometrik ve zamansal bileşenleri mevcut olan görüntüler üzerinden yeryüzüne ait bilgilerin yorumlanabilmesi ve analizi için sınıflandırma işlemlerine ihtiyaç duyulmaktadır. Sınıflandırma işlemlerinde temel amaç, elde edilen görüntü üzerindeki benzer yansıtım özelliklerini taşıyan cisimlerin aynı sınıf içerisinde yer almasını sağlamaktır. Sınıflandırma işlemleri ile görüntüdeki cisimlerin tanımlanabilmesi ve anlaşılabilmesi için kantitatif karar fonksiyonları oluşturulur[15]. Bu fonksiyonların oluşturulmasında görüntü üzerindeki cisimlerin spektral ve mekansal bilgilerinden faydalanılır. Sınıflandırma işlemlerinde temel problem, bir sınıfı diğer sınıftan ayıracak fiziksel karakteristiklerin nasıl belirlenmesi gerektiğidir. Cisimler özelliklerine göre sınıflandırıldıktan sonra ise elde edilen çıktı ürüne tematik harita denilmektedir. Şekil 1.6. 'da verilerin sınıflandırılması gösterilmiştir.

Sınıflandırma işlemlerinde temel anlamda aşağıdaki işlem adımları uygulanmaktadır:

1. Görüntüde yer alan bölge için sınıfların belirlenmesi ve kontrol alanlarının seçilmesi.
2. Yapılacak çalışmanın amacına göre hangi spektral bantlarla çalışılacağının belirlenmesi.
3. İlk adımda belirlenen kontrol alanlarından eğitim ve test verilerinin oluşturulması.
4. Sınıflandırmada kullanılacak yöntemin seçilmesi ve algoritmanın kullanılması.
5. Sınıflandırılmış görüntünün analizi ve yorumlanması[24].



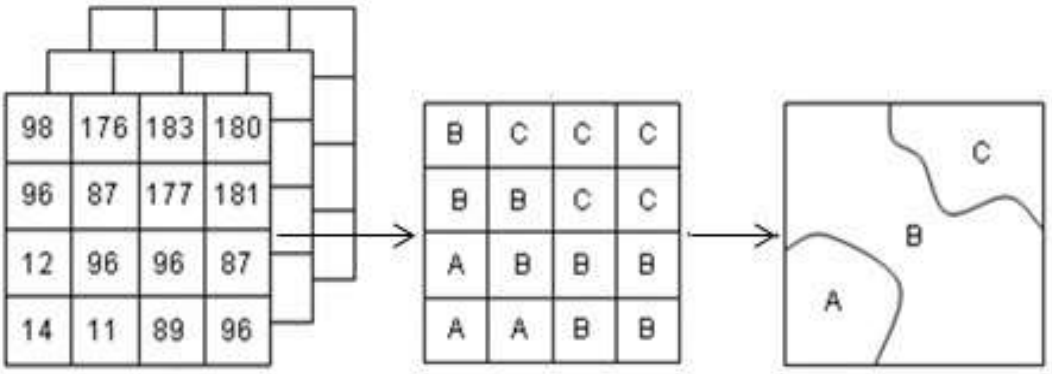
Şekil 1.6. Sınıflandırma

Uzaktan algılama ile sınıflandırma işlemlerinde genel olarak kontrollü sınıflandırma ve kontrolsüz sınıflandırma olmak üzere iki yaklaşım kullanılmaktadır. Kontrollü sınıflandırmada örnek uzay belirlenerek sınıflandırma işlemleri gerçekleştirilirken, kontrolsüz sınıflandırmada kullanılacak yöntemin çalışma prensibine göre eğitim aşaması olmadan gerçekleştirilmektedir.

1.6.1. Kontrolsüz Sınıflandırma

Kontrolsüz sınıflandırma, herhangi bir öncül tematik bilginin kullanılmadığı ve uydu görüntüsünü tercih edilen yöntemin yapısına bağlı olarak (spektral uzaklık vb.) belirli bir kriteri sağlayan grupları ortaya çıkarma işlemi olarak tanımlanabilir. Bir başka deyişle, görüntüdeki benzer özellik taşıyan piksellerin aynı küme etiketine atanmasını sağlayan sınıflandırma sürecinin genel adıdır. Kontrolsüz sınıflandırma yöntemleri tam otomatik sınıflandırma yöntemleridir.

Kontrolsüz sınıflandırma yöntemleri sonucu elde edilen sınıflar spektral sınıflardır. Bu sınıfların ne olduğu daha önceden bilinmemekle beraber o bölgeye ait topografik haritalar ve yardımcı bilgiler ile karşılaştırılarak daha sonradan belirlenmektedir[25]. Şekil 1.7. 'de kontrolsüz sınıflandırmanın akışı gösterilmektedir.



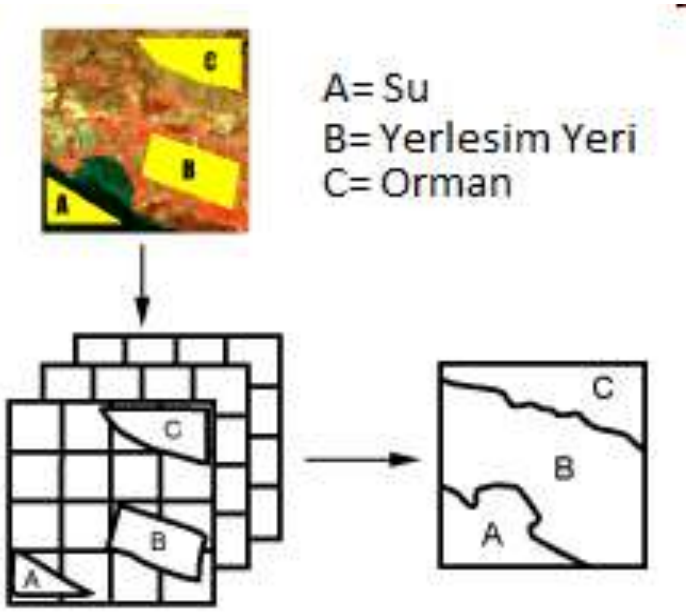
Şekil 1.7. Kontrolsüz Sınıflandırma

Kontrolsüz sınıflandırma uygulamalarında K-Means, Fuzzy C-Means, sıralı kümeleme, tekrarlı veri analizi, yapay zeka optimizasyon algoritmaları gibi yöntemler kullanılmaktadır.

1.6.2. Kontrollü Sınıflandırma

Kontrollü sınıflandırma uygulamalarında istenen sınıfları temsil eden öncül tematik bilgi vardır ve bu bilgilere eğitim verisi denilmektedir. Sınıflandırma yapılacak görüntüdeki cisimlerin sınıf etiketleri bir uzman tarafından önceden belirlenerek sınıflandırma esnasında piksellerin spektral değerlerine göre belirlenen sınıflara atanmasını sağlamaktadır. Ancak kontrollü sınıflandırmada sık karşılaşılan sınıf çakışması problemi işlemin doğruluk değerini etkilemektedir.

İlk aşamada yapılacak olan örnekleme ile arazi örtüsü çeşidi, sınıflandırma yapılacak veriler ve bu bilgiler için gerekli spektral değerler oluşturulur. Bir sonraki aşamada görüntü üzerindeki piksel değerlerinin hangi sınıfa ait olduğu belirlenerek o sınıfa ataması yapılır. Bütün veriler sınıflandırıldıktan sonra çıktı ürün uygulama amacına uygun olarak analiz edilir ve yorumlanır. Şekil 1.8. 'de kontrollü sınıflandırmanın akışı gösterilmektedir.

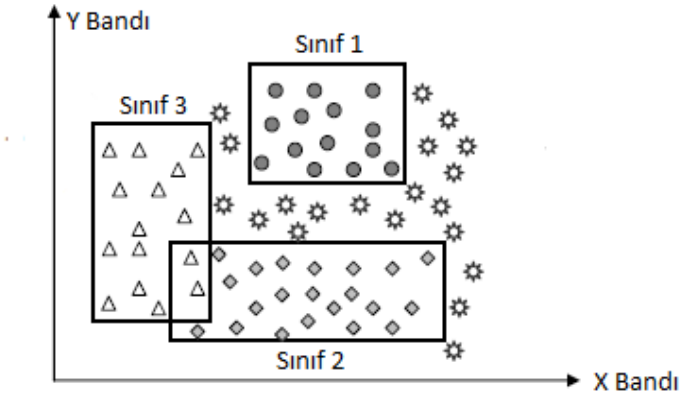


Şekil 1.8. Kontrollü Sınıflandırma

Kontrollü sınıflandırma uygulamalarında yakın zamana kadar en kısa uzaklık, en çok benzerlik, paralelkenar yöntemi gibi yöntemler kullanılmaktayken, günümüzde daha fazla doğruluk değeri sağlayabilen destek vektör makineleri ve hızlandırma algoritmaları, yapay sinir ağları, karar ağaçları, derin öğrenme teknikleri gibi yöntemlerde kullanılmaya başlanmıştır.

1.6.2.1. Paralelkenar Yöntemi

Bu yöntemde bütün sınıflar için eğitim verileri kullanılarak minimum ve maksimum değerler belirlenmektedir ve bu değerlere göre sınıflandırılacak pikseller bulunduğu değer aralığındaki sınıfa yerleştirilmektedir. Herhangi bir aralıkta bulunmayan pikseller ise tanınmayan pikseller olarak adlandırılmaktadır. Parametrik olmayan paralelkenar yöntemi hızlı ve matematiksel ifadesi basit olmasına rağmen, bu yöntemin sakıncası aynı pikselin birden fazla değer aralığına girmesidir. Bu durumda pikselin en yakındaki komşu piksellerin sınıf dağılımlarına göre atanmasına karar verilse de istatistiksel olarak doğruluğu yüksek olmayacaktır[2]. Şekil 1.9. 'da paralelkenar yöntemi belirtilmiştir.



Şekil 1.9. Paralelkenar Yöntemi

1.6.2.2. En Kısa Uzaklık Yöntemi

En kısa uzaklık yöntemi işlemsel olarak basittir ve eğitim verisinin az olduğu uygulamalarda kullanı uygundur. İlk aşama olarak eğitim verilerinden her bir sınıf için her bir banda ait sınıf merkezleri (ortalama değerler) hesaplanır. Daha sonra bilinmeyen her bir pikselin sınıf merkezine olan mesafesi hesaplanır. Bu piksellerin bulundukları uzaklıklara göre en yakın olduğu sınıfa ataması yapılır[15].

Bu yöntem matematiksel olarak basit ve hızlı bir kontrollü sınıflandırma yöntemidir. Ancak görüntüdeki cisimlerin spektral değerlerinin birbirine yakın olması durumunda çıktı ürünün doğruluk değeri yüksek olmayabilir.

1.6.2.3. Maksimum Olabilirlik Yöntemi

Maksimum olabilirlik yönteminde, her bir piksel en yüksek olasılık değerine sahip olduğu sınıfa atanır. Bu yöntemde sınıflandırma yapılacak bölgedeki yeryüzü özelliklerini tanımlayan örnekleme alanları kullanılarak, piksel değerlerinin hesaplanacak olan olasılık değerlerine göre en çok benzediği sınıfa ataması için özellik dosyalarını kaydedilir. Bu sınıflandırma yönteminde tüm sınıflara ait bantların ortalama verileri ve varyansları bulunur. Görüntü üzerinden alınan örneklerin belirlenen sınıflara ait olma ihtimalleri bulunur ve hesaplamalar doğrultusunda alınan örnekler bulunduğu en yakın sınıfa atanır. Bu yöntemde eğitim verisinin normal dağılım gösterdiği kabul edilir. Bu kabul gerçekte doğru olmaması ihtimaline rağmen verilerin modellenmesinde bu yöntemin daha sağlam olduğu görülmektedir[15]. Bu yöntem uzaktan algılama çalışmalarında en yaygın olarak kullanılan ve etkili bir kontrollü sınıflandırma yöntemidir.

2. BÖLÜM

YÖNTEM VE MATERYAL

2.1. Yöntemler

2.1.1. Makine Öğrenmesi Teknikleri

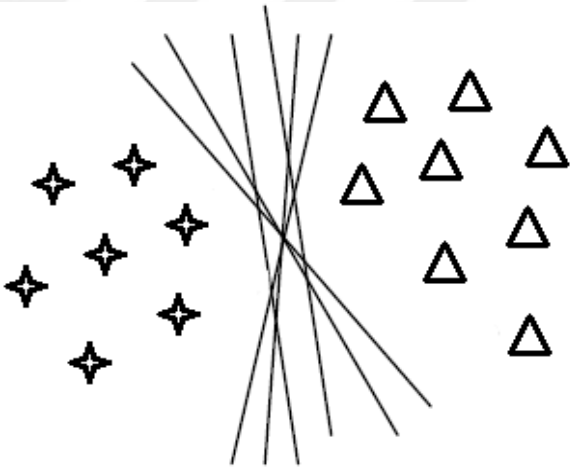
Makine öğrenmesi teknikleri, çözülmesi hedef alınan sınıflandırma işlemleri için çeşitli algılayıcılardan elde edilen görüntüleri kullanarak en uygun sınıflandırma ve regresyon modelini oluşturmayı amaçlamaktadır. Uzaktan algılama ile veri analizi çalışmalarında bir çok makine öğrenmesi teknikleri yaklaşımı ortaya konulmuştur. Bununla birlikte, destek vektör makineleri, yapay sinir ağları, en çok benzerlik yöntemi, karar ağaçları, derin öğrenme gibi yöntemler uygulamalarda en çok kullanılan yöntemler arasında yer almaktadır.

2.1.2. Destek Vektör Makineleri

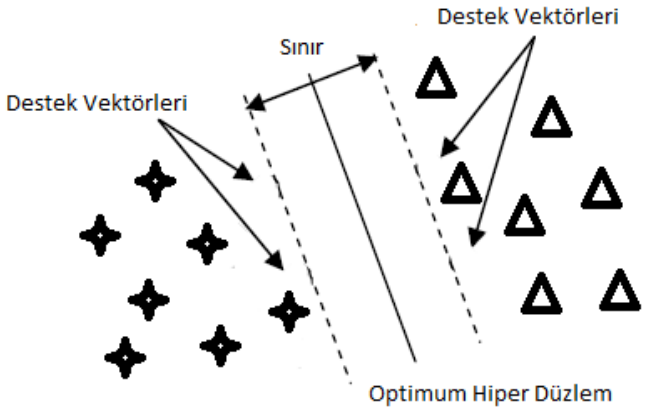
Vladimir Vapnik tarafından ortaya konulan destek vektör makineleri (SVM), istatistiksel öğrenme teorisine dayalı kontrollü bir sınıflandırma yöntemidir. SVM sınıflandırma ve regresyon uygulamalarının analizinde kullanılan yöntemlerin başında gelmektedir. SVM' nin bünyesinde bulundurduğu matematiksel ifadeler lineer verilerin, çok sınıflı verilerin ve lineer olmayan verilerin sınıflandırma işlemleri için geliştirilmiştir[6]. SVM yöntemi karmaşık olarak adlandırılabilecek görüntü verisinin sınıflandırılmasında etkili bir yöntem olarak kullanılmaktadır. Sınıflandırılacak görüntü için elde edilecek az sayıda eğitim verisiyle bile sınıflandırmanın mümkün hale gelmesine olanak sağladıkları için yaygın olarak uygulamalarda kullanılmaktadır[26]. Bu sebeplerden dolayı SVM; yüz tanıma, ses tanıma, duygu analizi, karakter tanımlanması gibi uygulamalarda kullanılmasının yanı sıra uzaktan algılama yöntemleri

ile elde edilmiş görüntülerin sınıflandırılmasında da başarıyla kullanılmaya devam etmektedir[28].

Diğer makine öğrenmesi tekniklerinden farklı SVM yönteminde öğrenme makinesinin riski; eğitim verilerinden elde edilen deneysel gözlemlerin riskinin ve güven aralığının toplamına bağlıdır[2]. Ayrıca SVM yöntemi ile sınıflandırmada öğrenme aşamasını Vapnik - Chervonenkis boyutu ile ele alan temelli yapısal risk minimizasyonunu kullanmaktadır. Bu sayede sınıflar arasında oluşabilecek hata üst sınırının en az olması hedeflenmektedir[27]. SVM, iki sınıf arası mesafeyi maksimize ederek optimize edilmiş hiper düzlemler yardımı ile binary sınıflandırma yapabilen bir yöntemdir. Optimize edilmiş hiper düzlemin en yakın sınıf elemanlarına ise destek vektörleri denir[28]. Şekil 2.1. 'de ve Şekil 2.2. 'de hiper düzlemler ve destek vektörleri gösterilmiştir.



Şekil 2.1. Hiper Düzlemler



Şekil 2.2. Destek Vektörleri

2.1.2.1. Lineer Destek Vektör Makineleri

Destek vektör makineleri kullanılarak yapılan sınıflandırma uygulamalarında mevcut verilerin lineer olarak ayrılabilen veya lineer olarak ayrılamayan durumları ile karşılaşılmaktadır. Lineer olarak ayrılabilen verilerin birbirinden ayrılmasını sağlayan en uygun hiper düzlemin elde edilmesi ve bu verilerin sınıflandırılması basittir. Diğer sınıflandırma uygulamalarında da karşılaşıldığı gibi lineer olarak ayrılamayan verilerin sınıflandırma işlemleri ise daha zor olmaktadır ve verilerin lineer olarak ayrılabilecekleri bir uzaya aktarılması gerekmektedir[29].

2.1.2.1.1. Lineer Ayrılabilen Veriler için Destek Vektör Makineleri

SVM yöntemi ile sınıflandırma işlemlerinde iki sınıfa ait örneklerin, eğitim verileri sayesinde elde edilen karar fonksiyonunun kullanılmasıyla sınıfların birbirlerinden ayrılması hedeflenmektedir. Bu karar fonksiyonu ile eğitim verilerini en uygun biçimde olacak şekilde ayırabilecek hiper düzlemler bulunmaktadır[30].

Lineer ayrılabilen verilerin SVM için eğitimi oldukça basittir. SVM' nin eğitimi için kullanılacak iki sınıflı N sayıda örnekten oluşan lineer ayrılabilen bir eğitim verisinin $\{x_i , y_i\}$, $i = 1,2,.....,N$ şeklinde verildiği kabul edilsin. Bu durumda x özellik uzayını ifade ederken y sınıf etiketi değerlerini ifade etmektedir. Hiper düzlemin merkezden olan uzaklığını temsil eden b eşik değeri ve hiper düzlemin ağırlık vektörü w' nin hesaplanmasıyla birlikte sınıfları birbirinden ayıran hiper düzlem belirlenmektedir[31]. İki sınıfın arasındaki mesafeyi optimum yaparak en uygun hiper düzlem 2.1' deki gibi bulunmaktadır.

$$wx_i + b = 0 \tag{2.1}$$

Hiper düzlemin altında ve üstünde kalan noktalar ise aşağıdaki eşitliklere uymaktadır.

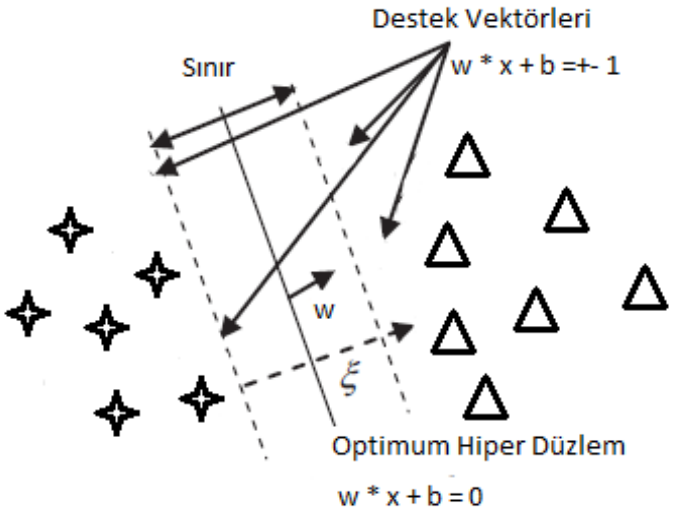
$$wx_i + b \geq +1 \quad için \;\; y_i = +1 \tag{2.2}$$

$$wx_i + b \leq -1 \quad için \;\; y_i = -1 \tag{2.3}$$

Bu iki eşitlikten aşağıdaki eşitlik elde edilir.

$$y_i(wx_i + b) - 1 \geq 0 \tag{2.4}$$

Bu eşitliği sağlayan hiper düzlemin iki sınıfa olan uzaklıkları ise sınır olarak tanımlanmaktadır. Şekil 2.3. 'de lineer ayrılabilen veriler için hiper düzlem gösterilmiştir.



Şekil 2.3. Lineer Olarak Ayrılabilen Veriler için Hiper Düzlem

Hiper düzlem ve bu düzlemlere yakın noktalar arasında *w* ve *b* değerleri ile tanımlanan mesafe 1 / $\|w\|$ 'dir[31]. En uygun hiper düzlem sınırının optimum pozisyona çıkarılması için $\|w\|$ değerinin minimum hale getirilmesi gerekmektedir. Böylece en uygun hiper düzlemin bulunabilmesi için amaç fonksiyonunun çözümlenmesi aşamasına geçilebilmektedir[33]. Bu amaç fonksiyon için koşul aşağıdaki gibidir.

$$Minimize \frac{1}{2}\|w\|^2 \tag{2.5}$$

Amaç fonksiyonunun sınırlamaları ise;

$$y_i(wx_i + b) - 1 \geq 0 \text{ ve } y_i \in \{1,-1\} \tag{2.6}$$

şeklindedir.

Yukarıdaki koşul ve sınırlamalarla ifade edilen problemin çözümü için aşağıda gösterilen Lagrange fonksiyonu kullanılır.

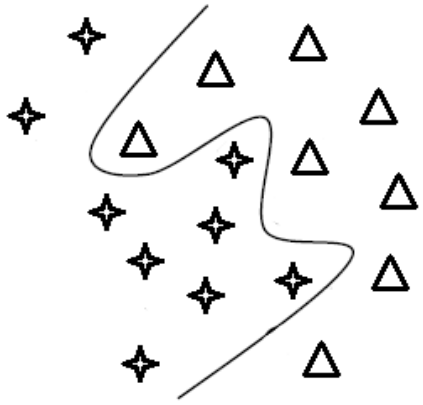
$$L(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \alpha_i y_i(wx_i + b) + \sum_{i=1}^{N} \alpha_i \tag{2.7}$$

Sonuç olarak, lineer ayrılabilen iki sınıflı problemin çözümü için aşağıdaki fonksiyon yazılabilmektedir[30].

$$f(x) = sign(\sum_{i=1}^{N} \lambda_i y_i(xx_i) + b) \tag{2.8}$$

2.1.2.1.2. Lineer Ayrılamayan Veriler için Destek Vektör Makineleri

Uzaktan algılama yöntemleri ile elde edilmiş görüntülerin sınıflandırılmasında diğer uygulamalarda da karşılaşılabileceği gibi verilerin lineer olarak ayrılamama durumu ile karşı karşıya kalınabilmektedir. Bu durum verilerin lineer bir düzlem ile ayrılamadığı anlamına gelmektedir. Şekil 2.4. 'de lineer olarak ayrılamayan veriler belirtilmiştir.



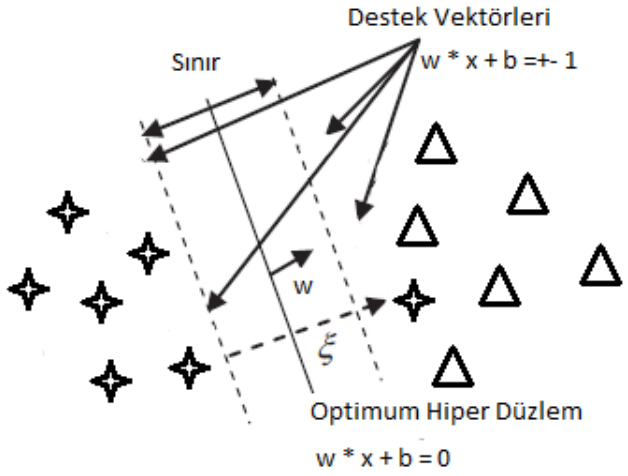
Şekil 2.4. Lineer Olarak Ayrılamayan Veriler

Bu tip veriler içeren problemlerde verilerin bir bölümü hiper düzlemin diğer tarafında kalmaktadır. Bu yüzden çözülmesi amaçlanan problem için sınıflandırma hatalarını temsil eden yapay değişkenin (ξ) tanımlanarak optimizasyon modelinin yeniden çözümlenmesi gerekmektedir[5]. Yapay değişkenin eklendiği ve çözülmesi gerek optimizasyon modeli aşağıdaki gibi olur.

$$y_i(wx_i + b) - 1 + \xi_i \geq 0 \qquad i = 1,2,.....,N \tag{2.9}$$

Kullanılacak bu formülüzasyon ile yöntem sınırın maksimum olmasını sağlamayı ve sınıflandırma hatalarının minimum olmasını amaçlamaktadır. Bu koşulların sağlanması için *C* parametresi tanımlanır. *C* parametresi Lagrange çarpanlarının alabilecekleri maksimum değeri ifade etmektedir. *C* parametresi büyük değerler aldıkça hatalı veri sayısı da azalmaktadır. $C = \infty$ olduğu kabul edildiğinde verilerin lineer olarak ayrılabildiği pozisyon gözlemlenir[33]. Model bu parametrelere bağlı olarak yeniden düzenlendiğinde ise aşağıdaki gibi olmaktadır. Şekil 2.5. 'de lineer olarak ayrılamayan veriler için hiper düzlem gösterilmiştir.

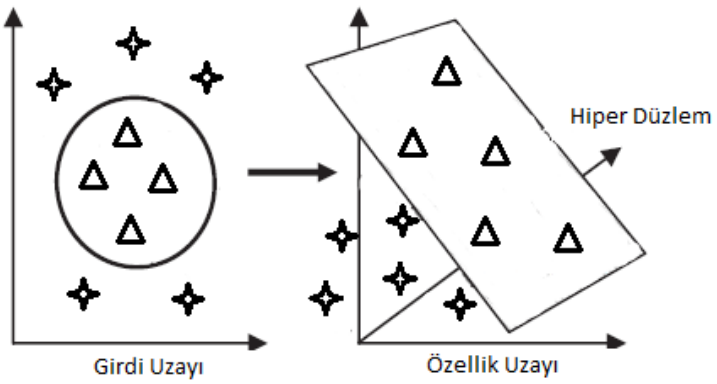
$$L(w, b, \alpha, \xi, \mu) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i - \sum_{i=1}^{N}\alpha_i\{y_i(wx_i + b) - 1 + \xi_i\} - \sum_{i=1}^{N}\mu_i\xi_i \tag{2.10}$$



Şekil 2.5. Lineer Olarak Ayrılamayan Veriler için Hiper Düzlem

2.1.2.2. Lineer Olmayan Destek Vektör Makineleri

Lineer eşitlikler kullanılarak elde edilemeyen hiper düzlemin bulunabilmesinin mümkün olmadığı durumlarda lineer olmayan sınıflandırıcılar kullanılabilmektedir. Problemin çözülebilmesi için verilerin yüksek boyutlu ve lineer bir biçimde ayrılabilecekleri özellik uzayına aktarılması ve uzaylar arası dönüşüm yapabilen fonksiyonlara ihtiyaç duyulmaktadır. Şekil 2.6. 'da lineer olmayan verilerin farklı bir uzayda ayrılması gösterilmiştir.



Şekil 2.6. Lineer Olmayan Verilerin Farklı Bir Uzayda Ayrılması

Destek vektör makineleri Kernel fonksiyonlarını kullanarak lineer olmayan dönüşümler yapabilmektedir. Bu durum verilerin yüksek boyutta lineer olarak ayrımına olanak sağlamaktadır. Kernel fonksiyonu aşağıdaki gibi tanımlanmaktadır[30].

$$K(x_i, x_j) = \Phi(x).\Phi(x_i) \quad (2.11)$$

Destek vektör makineleri ile sınıflandırma işlemlerinde birden çok Kernel fonksiyonunun kullanımı bulunmaktadır. En çok kullanılan fonksiyonlar ise radyal tabanlı fonksiyon, polinom fonksiyon, Pearson VII ve normalleştirilmiş polinom fonksiyonlarıdır.

2.1.2.3. Çok Sınıflı Destek Vektör Makineleri

İlk başlarda iki sınıflı verilerin sınıflandırılması için tasarlanan destek vektör makineleri için, sınıf sayısının ikiden fazla olduğu problemlerin çözümünde yeni çözümler geliştirmiştir. En çok tercih edilen yöntemlerin ise bazıları şunlardır:

- Bire karşı hepsi yaklaşımı
- Bire karşı bir yaklaşımı
- Çok sınıflı amaç fonksiyonu

Bire karşı hepsi yaklaşımında, veri setindeki bir sınıfın bir grubu oluşturması ve diğer sınıfların başka bir grubu oluşturması mantığı bulunmaktadır. Her bir sınıf için bu gruplama işlemi yapılarak iki sınıflı bir sınıflandırma yapılıyormuş gibi destek vektör makineleri eğitilmektedir[2].

Bire karşı bir yaklaşımında, her bir sınıfa ait gruplar diğer gruplar ile ayrı ayrı eğitilir. Test verisindekiler her bir eğitim sonucu bir işaret alır. Eğitim sonucunda en çok işaret aldığı sınıfa ataması yapılır. Bu yöntem simetrik bir yaklaşım olarak karşımıza çıkmaktadır.

Çok sınıflı amaç fonksiyonu yöntemi, Lagrange fonksiyonunun iki sınıftan fazla durumda kullanılacak sınıflandırmalara uygun hale getirilmesini amaçlayan bir yöntemdir ancak bu yöntemin sınıf sayısı arttıkça çıktı üründeki hata miktarının artması, yöntemin tercih edilebilirliğini düşürmektedir[4].

2.1.3. Komşuluk Bileşen Analizi

Komşuluk bileşen analizi bir boyut azaltma, öznitelik seçim tekniğidir. Makine öğrenmesi uygulamalarında özniteliklerin ölçümü çok önemli pozisyondadır. En başarılı

öğrenme algoritmalarından biri olan komşuluk bileşen analizi(NCA) sınıflandırma çalışmaları içinde yaygın bir biçimde kullanılmaktadır[34]. Komşuluk bileşen analizi, en yakın komşu sınıflandırıcısının sınıflandırma doğruluğu ile ilgilenen kriterleri optimize eden vektörlerin projeksiyonunu öğrenerek sınıflandırma işlemlerini gerçekleştirmektedir. Bir başka deyişle, NCA yansıtılan alandaki en yakın komşu sınıflandırıcısının performansını optimize eden bir lineer projeksiyon seçmektedir.

NCA, öngörülen alandaki sınıfları ayırmada etkili olacak projeksiyonu seçerken ilişkili sınıf etiketlerinden oluşan eğitim verilerinden faydalanmaktadır. NCA sınıflandırıcılarını optimize ederken her sınıftaki dağılım hakkında zayıf varsayımlarda bulunmaktadır. Bu durum sınıflardaki dağılımların modellenmesinde Gaussian karışımlarının kullanımına daha yakın bir eşleşme gerçekleştirmektedir[35].

NCA, vektörlerden oluşan bir eğitim seti ve eğitim seti ile ilgili etiket kümesini kullanmaktadır. Daha sonra eğitim setini *z* boyutlu yansıtan, sınıflar arası ayrım yaparken etkili olan bir projeksiyon matrisi kullanmaktadır. Bu projeksiyon matrisi en yakın komşu sınıflandırıcısı tarafından kullanılacak mesafeleri tanımlamaktadır.

NCA yöntemi sınıflandırıcının doğruluğunu en yüksek değere çıkarmayı amaçlamaktadır. Böylece eğitim verisinin üzerinde sınıflandırılan nokta sayısını maksimize etmeye çalışmaktadır[35].

NCA yönteminin içerdiği algoritma sınıflandırma hatasını K Nearest Neighborhood(KNN) sınıflandırmasının stokastik bir varyantından hesaplamaktadır. Algoritma veriler arası mesafe hesaplanırken beklenen sınıflandırma hatasını en aza indiren doğrusal dönüşümü tahmin etmeye çalışır. NCA yönteminde sınıflandırıcılar eğitim setindeki verilerin işaretlerini verileri etiketlemek için kullanır. Her bir sorgu aşaması için eğitim setindeki bilgiler uzaktaki veri yerine yakındaki veriyi tercih etmek için olasılık dağılımından faydalanır. NCA yöntemi diğer algoritmalardan önemli bir açıdan farklıdır. NCA algoritması mesafelerin ölçülmesi bakımından sürekli ve ayırt edilebilir özellikte olmasına rağmen dış bükey bir yapıya sahip değildir[13].

NCA, sınıflandırmada kullanılacak Mahalanobis uzaklığını öğrenmeyi hedefleyen ve bu mesafeyi kullanarak sınıflandırmayı gerçekleştiren bir yöntemdir. NCA yöntemi

Mahalanobis uzaklığını hesaplarken izdüşüm matrisini öğrenir ve matrisin ters çalışmasını engeller.

NCA yöntemi aşağıda tanımlanan formülüzasyon ile Mahalanobis uzaklığını tanımlamayı amaçlar.

$$dist\,(y_i\,,\,y_j)=(A^T y_i - A^T y_j)^T(A^T y_i - A^T y_j) \tag{2.12}$$

Bu eşitlikte A verileri dönüştüren izdüşüm matrisi olarak tanımlanmaktadır. Böylece NCA Mahalanobis uzaklığını öğrenmekten A matrisini öğrenme aşamasına geçmektedir.

Eğitim verisindeki bir veri komşu olarak başka bir veriyi seçerek sınıf etiketini belirler. Dönüştürülmüş olasılık değerleri olan veriler öklid mesafeleri ile softmax regresyonu kullanılarak tanımlanır.

$$p_{ij} = \frac{exp(-\|A^T y_i - A^T y_j\|^2)}{\sum_{k\neq 1} exp(-\|A^T y_i - A^T y_k\|^2)} \;,\; p_{ii} = 0 \tag{2.13}$$

Stokastik seçim kurallarına bağlı olarak verilerin doğru şekilde sınıflandırma olasılığı aşağıdaki gibi hesaplanır ve aynı sınıftaki verilerin kümesi gösterilir.

$$p_i=\sum_{j\in C_i} p_{ij} \tag{2.14}$$

$$C_i=\{j|z_j=z_i\} \tag{2.15}$$

NCA daha sonra A matrisini doğru bir şekilde sınıflandırılmış nokta sayısını maksimize ederek hesaplar.

$$f(A)=\sum_i p_i = \sum_i \sum_{j\in C_i} p_{ij} \tag{2.16}$$

A matrisine göre değişen f değeri aşağıdaki gradyan kuralını doğurur.

$$\frac{\partial f}{\partial A} = 2\sum_i \left(p_i \sum_k p_{ik}(y_i - y_j)(y_i - y_j)^T - \sum_{j\in C_i} p_{ij}\,(y_i - y_j)(y_i - y_j)^T \right)A \tag{2.17}$$

Böylece gradyan tabanlı iyileştirici kullanılmış olunur.

Başka bir deyişle NCA için algoritma şöyle de açıklanabilir.

- Projeksiyon matrisi *A* başlatılır.
- Her adımın ayrıştırılabileceği gradyan tabanlı iyileştirici aşağıdaki adımlarla kullanılır.

  1- Eğitim verisi projeksiyon matrisine uyarlanır.

  2- Veriler arasındaki mesafeler hesaplanır.

  3- Olasılıklar hesaplanır.

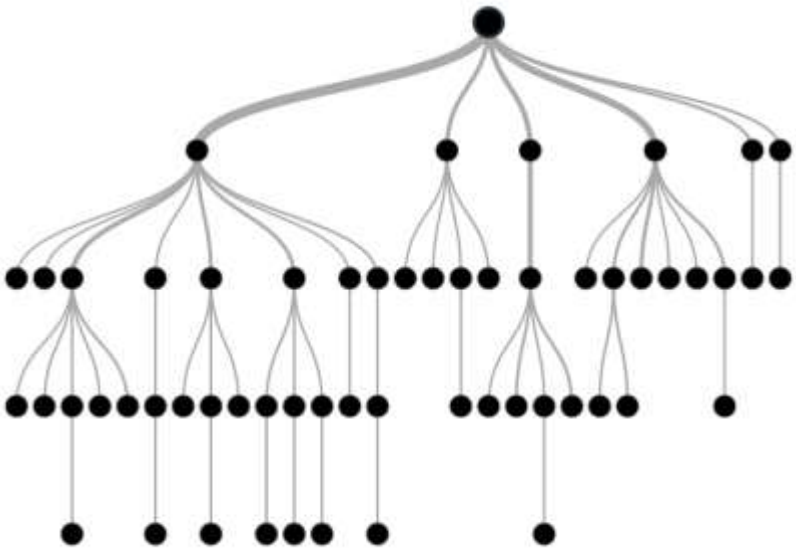
  4- Projeksiyon matrisi *A* hesaplanır ve güncellenir.
- Yukarıdaki işlem adımları tekrarlanır[36].

2.1.4. Karar Ağaçları

Karar ağaçları, ağaç yapıları oluşturulurken kullanılan kuralların basit ve anlaşabilir olmasından dolayı sınıflandırma çalışmalarında sıkça kullanılan yöntemlerden bir tanesi konumundadır[37]. Bu özelliklerinden dolayı makine öğrenmesi teknikleri arasında tercih edilebilirliği fazla olan yöntemdir. Karar ağaçları yöntemi çözülmesi gereken sınıflandırma problemini çözerken veriyi bölümlere ayırarak basitçe karar verme işlemlerini gerçekleştirmektedir.

Karar ağaçlarının içerisinde düğüm, dal ve yaprak olarak isimlendirilen üç ana kısım bulunmaktadır. Karar ağaçlarının bu yapısında her bir düğüm bir öznitelik bilgisini, örnek olarak uydu görüntüsündeki bant değerini, temsil etmektedir. Yöntemde, verilerin öznitelik bilgilerine dair izler belirlenerek karar kuralları oluşturulmaktadır ve eğitim verilerine bağlı test verilerinin hangi sınıfta yer alacağına karar verilmektedir.

Sınıflandırılacak veriler karar ağacı oluşturulurken kullanılacak kural mekanizmasına bağlı olarak sıralı bir biçimde bir yaprağa ulaşana kadar alt bölümlere dağıtılarak sınıflandırılmaktadır. Oluşturulan karar ağacı yapısının genelleme kabiliyetinin anlaşılabilmesi için test verileri ağacın kökünde test edilir ve veriler yaprağa ulaşana kadar tek bir karar kuralına bağlı kalır[2]. Şekil 2.7. 'de karar ağaçlarının yapısı gösterilmiştir.

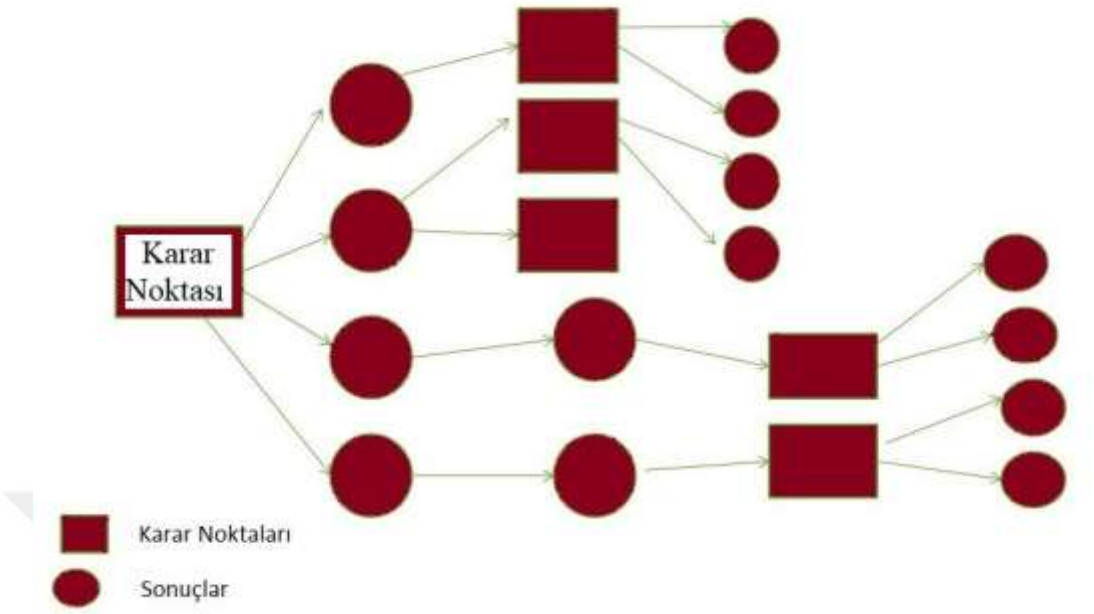


Şekil 2.7. Karar Ağaçları

Değişkenlerin sayısına bağlı olarak karar ağaçlarının yapısı değişmektedir. Bunlar tek değişkenli karar ağacı ve çok değişkenli karar ağacı olarak tanımlanmaktadır. Tek değişkenli karar ağacının çoğunlukla arazi örtüsü sınıflandırması çalışmalarında kullanıldığı görülmektedir. Bu ağaç yapısı ile iç düğümün tek bir öznitelik bilgisi ile test edilerek verinin iki veya daha fazla alt kümeye ayrılması amaçlanmaktadır[38]. Bu alt kümelere ayrılma işlemi yaprak kısmına ulaşana kadar devam etmektedir. Karar ağaçları oluşturulurken kullanılan algoritmanın önemli bir yeri vardır. Bu durumun sebebi kullanılan algoritmaya göre ağaç yapısının şeklinin değişmesidir.

2.1.4.1. Karar Ağaçlarının Oluşturulması

Literatürde yapılan çalışmalara bakıldığında karar ağaçlarının oluşturulması için bir çok algoritma ve yaklaşım geliştirilmiştir. Sınıflandırılması yapılacak veri seti için doğruluk derecesi yüksek olması istenen ağaç yapıları oluşturulurken yüksek doğruluk imkanı sunan algoritmalardan faydalanılmaktadır. Bu algoritmalar genellikle Greedy yaklaşımını esas alarak tasarlanmaktadır. Verilerin dallara ayrılırken hangi gruptan başlanacağı ile ilgili kararlar bu yaklaşım ile alınan kararlar ile mümkün olmaktadır. Geliştirilen bu algoritmalar genel olarak Hunt algoritması bir başka deyişle böl ve elde et yönteminin farklı versiyonlarını kullanmaktadır[39]. Şekil 2.8. 'de karar ağaçlarının oluşturulması belirtilmiştir.



Şekil 2.8. Karar Ağaçlarının Oluşturulması

2.1.4.1.1. Sınıflandırma ve Regresyon Ağaçları

Karar ağaçları bazlı bir yöntem olan regresyon ağaçları(CART) bir çok uygulamada hızlı ve kullanımının kolay olması sebebiyle tercih edilmektedir. İki dallanma yaparak diğer karar ağaçlarından farklılaşan sınıflandırma ağaçları eğitim verisindeki değişkenin kategorik olması durumunda söz konusu olmaktadır. Eğitim verisinde bulunan cevap değişkeninin sürekli olmasında ise regresyon ağacı ismini almaktadır[4].

Sınıflandırma ağaçlarında genellikle iki yöntem kullanılmaktadır.Bunlar;

- Gini indeksi
- Twoing kuralı

olarak isimlendirilmektedir.

2.1.4.1.1.1. Gini İndeksi

Gini indeksi, oluşturulacak sınıfa ait özniteliklerin safsızlığının ölçülmesi için kullanılan bir yöntemdir. Gini indeksi basit ve hızlı hesaplama yapabilen bir algoritmadır. Bu yöntemde kullanılan işlem adımları aşağıdaki gibidir[33]. Bu akışta k sınıf sayısını, d

bir düğümdeki örnek sayısını, *n* sağ taraftaki *i* kategorisindeki örnek sayısını, *m* ise sol taraftaki *i* kategorisindeki örnek sayısını ifade etmektedir.

- Setteki tüm öznitelik değerleri ikiye ayrılır. Oluşturulan sağ ve sol bölünmelere denk gelecek sınıf değerleri gruplandırılır.

- Özniteliklerin sağ ve sol Gini indeks değerleri hesaplanır.

$$G_{sağ} = 1 - \sum_{i=1}^{k} \left(\frac{n_i}{|d_{sağ}|}\right)^2 \tag{2.18}$$

$$G_{sol} = \sum_{i=1}^{k} \left(\frac{m_i}{|d_{sol}|}\right)^2 \tag{2.19}$$

- Her bir öznitelik için aşağıdaki eşitlik hesaplanır. *n* eğitim veri setinde;

$$Gini_i = \frac{1}{n}(|d_{sol}|\ G_{sol} + |d_{sağ}|\ G_{sağ}) \tag{2.20}$$

- Her bir öznitelik için *G* değerleri içerisinden en küçük olanı seçilir ve bölünme işlemi buradan devam ettirilir.

2.1.4.1.1.2. Twoing Kuralı

Twoing kuralı, Gini indeksini de tanımlayan Brieman tarafından 1984 yılında tanımlanmıştır. Bu kural, karar ağaçlarının oluşturulmasında öznitelik seçimi aşamasının farklı bir yaklaşımla belirlenmesini sağlamaktadır. Bu kural aşağıdaki sıralamanın tanımı pozisyonundadır[33].

- Eğitim veri seti öznitelik değerlerine göre iki gruba ayrılır. Bir *T* düğümü sol ve sağ olmak üzere iki dal oluşturur.

- Tüm bölünmeler için $P_{sol,sağ}$ ve $P\ (\ i\ |\ P_{sol,sağ}\ )$ olasılıkları aşağıdaki gibi hesaplanır. $P\ (\ i\ |\ P_{sol,sağ}\ )$ sınıf değerinin sağda veya solda olma olasılığını temsil etmektedir.

$$P_{sol} = \frac{T_{sol}'daki\ öznitelik\ değerlerinin\ ilgili\ sütundaki\ tekrar\ sayısı}{Eğitm\ verisindeki\ kayıtların\ sayısı} \tag{2.21}$$

$$P\ (\ i\ |\ P_{sol}\ ) = \frac{T_{sol}'daki\ kayıtların\ i\ sınıfları\ sayısı}{T_{sol}'daki\ her\ bir\ öznitelik\ sütunundaki\ tekrar\ sayısı} \tag{2.22}$$

$$P_{sağ} = \frac{T_{sağ}'daki\ öznitelik\ değerlerinin\ ilgili\ sütundaki\ tekrar\ sayısı}{Eğitm\ verisindeki\ kayıtların\ sayısı} \quad (2.23)$$

$$P\ (\ i \mid P_{sağ}\ ) = \frac{T_{sağ}'daki\ kayıtların\ i\ sınıfları\ sayısı}{T_{sağ}'daki\ her\ bir\ öznitelik\ sütunundaki\ tekrar\ sayısı} \quad (2.24)$$

- Φ ($s \mid T$), yani uygunluk ölçütü hesaplanır.

$$\Phi\ (\ s \mid T\ ) = 2P_{sağ}P_{sol}\ \sum_{i=1}^{n}\ \left|P\ (\ i \mid P_{sol}\ ) -\ P\ (\ i \mid P_{sağ}\ )\right| \quad (2.25)$$

- Uygunluk ölçütü değerlerinden büyük olanı seçilir ve dallanma işlemi bu değerlere göre yapılır, karar ağacı çizilir.

- İlk adıma dönülerek alt kümelere de aynı işlem tekrarlanır.

2.1.4.1.2. Ki-Kare Olasılık

Bu yöntem karar ağaçlarının oluşturulması sırasında özniteliklerin seçilmesi için kullanılmaktadır ve Mingers tarafından 1989 yılında ortaya atılmıştır. Ki-kare olasılık yönteminde iki değişken arasındaki ilişkinin geleneksel istatistik metoduna bağlı incelenmesi söz konusudur. Sonuç olarak elde edilecek istatistik bilgisinin yani bu fonksiyonun eşitliği aşağıdaki gibidir[2].

$$\chi^2 = \sum\ \sum\ \frac{(x_{ij} - E_{ij})^2}{E_{ij}} \quad (2.26)$$

Bu eşitlikte *E* değeri olasılık tablosundaki her bir hücreye denk gelen olasılık değeridir.

2.1.4.2. Karar Ağaçlarının Dallanma Kriterleri

Karar ağaçları oluşturulurken köklerden yaprağa giden sürecin yani dallanmanın hangi kriterlere göre yapılacağı karar ağacı oluşturma sürecinin önemli bir parçasıdır. Karar ağacı yöntemi yapısı bakımından hızlı ve basit olduğu halde dallanmanın nasıl yapılacağının bilinememesi bu yöntemin problemi konumundadır. Ağaç yapısındaki dallanmanın hangi koşullara göre yapılacağını belirlemek için bazı yaklaşımlar ortaya atılmıştır.

2.1.4.2.1. Entropi Bazlı Algoritmalar

Entropi bilgiye bağlı bir terimdir ve belirsizliğin boyutunu ifade eder. Entropi değeri bire yaklaşırsa belirsizlik artar ve sıfıra yaklaşıldıkça da belirsizlik azalır. C4.5, ID3 ve CART entropi bazlı yöntemler arasında en önemli olanlarıdır.

Veri setini *A* kabul edersek bu veri setinin entropi değeri aşağıdaki gibi olur.

$$E(A) = - \sum_{i=1}^{k} p_i log_2(p_i) \quad (2.27)$$

Bu denklemde *p* olasılık değerini, *k* ise eleman sayısını ifade etmektedir.

Ağaç yapısı oluşturulurken eğitim verisi sayesinde tüm ağaç yapılarının oluşturulması, dallanmanın hangi öznitelik ile oluşturulacağı önemi bir yer kaplamaktadır. Bu aşamaları gerçekleştirmek için ilk adımda bazı değerler hesaplanmaktadır ve ona göre yapılar oluşturulmaktadır[40].

2.1.4.2.2. Bilgi Kazancı ve Kazanç Oranı

Bilgi kazancı, Quinlan tarafından 1993 yılında ortaya atılmıştır. Bu yöntemde ağaç yapısının hangi niteliğine göre dallanma yapılacağının belirlenmesi için entropi kurallarından faydalanılmaktadır. Eğitim verisi setinin entropiye dayalı algoritmalar tarafından işleme alındığında değişkenin durumunun ölçütü bilgi kazancı olarak tanımlanmaktadır.

Alt kümelere ayrılma işlemi bilgi kazancının en yüksek olduğu değişkenden başlanarak gerçekleştirilmektedir. Veri setinin ilk halinin ve alt kümelerinin yani bölündükten sonraki durumun entropi değerlerinin bulunmasıyla bilgi kazancı hesaplanmaya başlanmaktadır. Eşitlik (2.27)' de entropi değerinden bahsedilmiştir.

Veri seti *A* değişkenine göre *k* adet alt kümeye ayrılır ve *E* test seti kullanılarak bilgi kazancı hesaplanmak istenirse aşağıdaki akış takip edilmektedir[2].

$$inf_E(A) = \sum_{i}^{k} \frac{|A_i|}{|A|} x\ inf\ (A_i) \quad (2.28)$$

Bu eşitlik kullanıldıktan sonra *E* test setine göre bilgi kazancı ise;

$$K(E) = inf(A) - inf_E(A) \quad (2.29)$$

şeklinde hesaplanır. Bu eşitlikte *K(E)* değerinin en yüksek değeri alması hedeflenir.

Kazanç kriteri ayırma bilgisi ile normalleştirilir ve bu bilgiler ışığında kazanç oranı belirlenebilir. Ayırma bilgisi aşağıdaki gibi hesaplanır.

$$AyırmaB(E) = - \sum_i^k \frac{|A_i|}{|A|} x \log_2 \frac{|A_i|}{|A|} \quad (2.30)$$

Sınıflandırma işleminde kullanılacak ayırma ile elde edilen bilgi oranını veren eşitlik yani kazanç oranı;

$$KazançO(E) = K(E) - AyırmaB(E) \quad (2.31)$$

ile hesaplanır.

İşlemler her bir yaprak sadece bir sınıfa ait değerleri içerene kadar tekrar eder.

2.1.4.3. Karar Ağaçlarının Budanması

Karar ağacı yöntemi temelde sınıflandırılması istenen verileri tek bir sınıf içerecek şekilde alt kümelere bölme işlemlerini gerçekleştirmektedir. Yapılan bu sınıflandırma sonucunda ortaya çıkan ağaç yapısı kompleks ve kapsamlı olabilmektedir.

Her bir yaprağın tek bir sınıf içermesi verideki gürültü nedeniyle aşırı öğrenme (overfit) problemine yol açmaktadır. Bu gibi problemlerin giderilmesi için ağaç yapısını budamak gerekli olmaktadır[2].

Sınıflandırma işlemine etki etmeyecek bölümlerin sınıflandırma dışına çıkartılarak ağacın bir alt bölümünün yaprak ile sonlandırılması gerekmektedir. Bu işlem adımları karar ağaçlarının budanmasını temsil etmektedir. Böylece budanan ağacın doğruluk değeri artar, kalitesi geliştirilir ve daha anlaşılabilir bir ağaç yapısı elde edilmektedir.

Karar ağaçlarının budanmasında genelde iki yaklaşım vardır. Bunlar; ön budama olarak adlandırılabilen eğitim veri setinin daha fazla bölünüp bölünemeyeceğine karar vermek ve karar ağacı oluşturulduktan sonra bölümlerden bazılarının kaldırılması yaklaşımlarıdır.

İlk yaklaşım durdurma olarak da isimlendirilebilir. Bu yaklaşım bilgi kazancı gibi kuramlar ile verilerin sınıflandırılması için ideal yolu bulmaya çalışmaktadır. İdeal yolu

bulurken veri için eşik değerini geçemediği durumda bölünme durmaktadır ve durulan aşamadaki yaprağın o veri için ideal sınıf olduğuna kanaat getirilmektedir[41].

İkinci yaklaşımda ise ağaç yapısında bulunan tüm yaprakların tek bir sınıfa ait veriyi içermesine kadar devam edilmektedir ve ağaç budama işlemi bu aşamadan sonra gerçekleştirilmektedir. Bu yöntem ile budama yapmak isteyen kullanıcılar sonradan çıkartılacak bölümlerin meydana gelmesinde çok fazla hesaplama yapmak zorunda kalmaktadırlar.

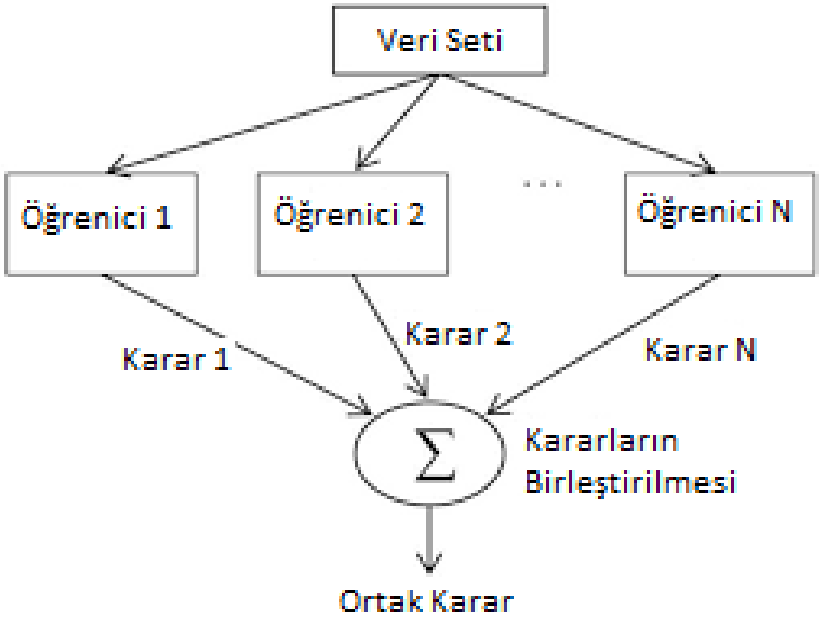
Karar ağaçlarının budanmasında bazı teknikler kullanılmaktadır. Bu teknikler aşağıdaki gibidir[2].

- Karmaşık değerli budama
- Karamsar budama
- Hata azaltmalı budama
- Kritik değerli budama
- Hata tabanlı budama

2.1.4.4. Karar Ağaçlarında Kullanılan Yöntemler

Sınıflandırma çalışmalarında genellikle tek bir sınıflandırıcı kullanılmaktayken günümüzde birden fazla sınıflandırıcının birleştirilerek topluluk adına tek bir sınıflandırıcı ile işlem yapması görüşü ortaya çıkmaktadır. Böylece eğitim verisi yardımıyla sınıflandırıcıların eğitilmesi ve her bir sınıflandırıcının verdiği çıktı ürünün doğruluk değerlerine göre sınıflandırma işlemi gerçekleştirilmektedir. Bu koşulun uygulandığı çalışmalarda sınıflandırıcıların birleştirilerek kullanıldığında verdiği doğruluk oranının, sınıflandırıcıların birer birer kullanıldığında verdiği doğruluk değerlerinden daha fazla olduğu gözlemlenmektedir[42]. Aynı zamanda birleştirilmiş sınıflandırıcıların verdiği hata değerlerinde de azalma meydana gelmektedir.

Sınıflandırma çalışmalarının temel mantığında yüksek doğruluk değerine sahip çıktı ürünü elde etmek olduğundan birleştirme yöntemleri üzerine çalışmaların artmasına sebep olmaktadır. Şekil 2.9. 'da kararların birleştirilmesi gösterilmiştir.



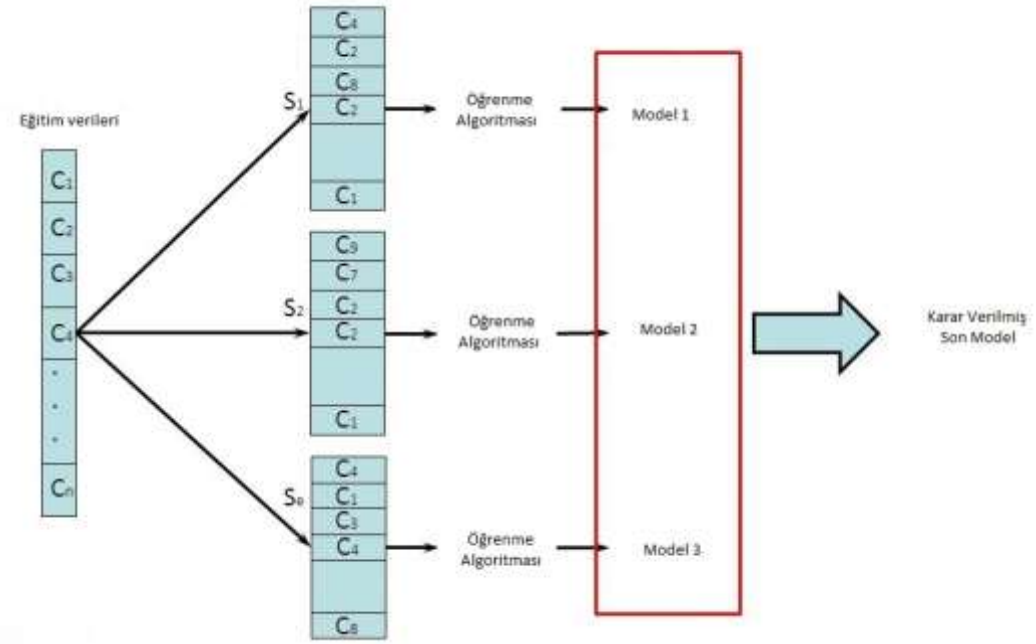
Şekil 2.9. Kararların Birleştirilmesi

2.1.4.4.1. Torbalama Karar Ağaçları Yöntemi

Torbalama yöntemi, sınıflandırıcıların birleştirilmesi için eğitim verileri ile her bir sınıflandırıcının eğitimini tamamlamasıyla birleştirilmesini sağlayan ön yüklemeli bir birleştirme yöntemidir. Bu yöntemde eğitim setinin yeniden örneklenmesinde daha önceki sınıflandırıcıların performans bilgisine ihtiyaç duyulmamaktadır.

Torbalama karar ağaçları yönteminde eğitim veri seti tekrarlı olacak bir biçimde örneklendirildikten sonra oluşturulacak eğitim setinde bazı örneklerin tekrarlandığı gözlemlenirken bazı örneklerin yer almadığı gözlemlenmektedir. Rastgele örnekleme ile eğitim verileri kullanılarak sınıflandırıcının eğitimi tamamlanmaktadır. Farklı koşullar sonucunda sınıflandırıcılar birleştirilerek sonuç sınıflandırıcı oluşturulmaktadır[2]. Sonuç sınıflandırıcı sayesinde en çok oy alan öznitelik sınıfına atanmaktadır.

Sınıflandırıcıların birleştirilmesi aşamasında her bir sınıflandırıcı aynı örnek setini kullanarak işlem yapmadığından, üretilen sonuçlar farklı olmaktadır. Sınıflandırıcıların tahminlerinin oylanmasıyla bu algoritma hatayı azaltmayı hedeflemektedir. Şekil 2.10. 'da torbalama karar ağaçları yöntemi gösterilmiştir.



Şekil 2.10. Torbalama Karar Ağaçları Yöntemi

2.1.5. Lineer Diskriminant Yöntemi

Lineer diskriminant yöntemi, elde edilmiş veri içerisindeki birbirinden farklı sınıflara ait bölümlerin lineer bir biçimde ayrılabilmesini maksimum hale getirerek sınıflandırma işlemini gerçekleştirmektedir. Bu sayede yöntem, ayrılabilmeyi verimli duruma getiren özellik alt uzayını bulur.

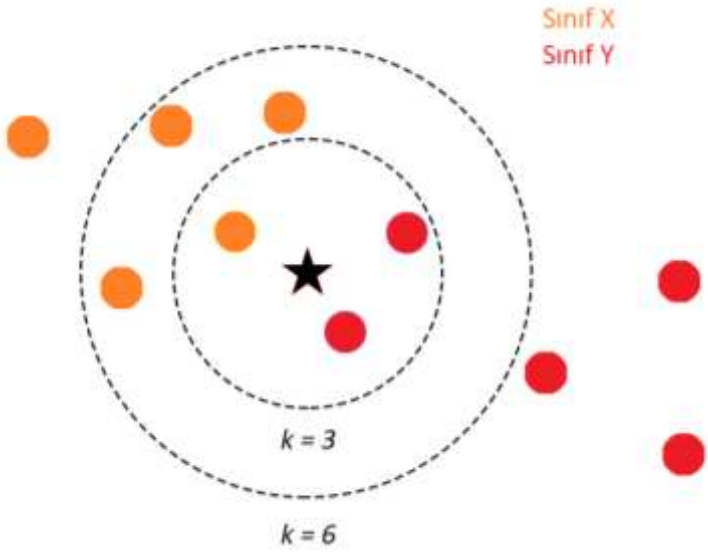
Bu analizde boyut azaltılarak hesaplama verimi artırılarak aşırı öğrenme durumunun azaltılması hedef alınmaktadır. Lineer diskriminant yönteminde verilerin lineer dağıldığı varsayılır[43].

2.1.6. K-En Yakın Komşu Yöntemi

Cover ve Hart tarafından ortaya atılan *k* en yakın komşu algoritması(KNN) basit ve güçlü yapısı sayesinde makine öğrenmesi uygulamalarında ve diğer sınıflandırma uygulamalarında sıkça kullanılan bir yöntemdir. Bu yöntem öznitelik uzayında bulunan eğitim örneklerine yakınlığı baz alarak sınıflandırmayı gerçekleştirmektedir[44]. KNN algoritması *k* değerine göre sınıf atama işlemlerini yapmaktadır.

Yöntemin gerçekleştirilmesi her ne kadar basit ve güçlü olsa da büyük bellek alanına ihtiyaç duyması ve verilerin hacminin artması sonucu işlem yükünün artması bu

yöntemin dezavantajları arasında yer almaktadır. Şekil 2.11. 'de KNN yöntemi gösterilmiştir.



Şekil 2.11. K-En Yakın Komşuluk Yöntemi

KNN yöntemi sınıflandırma sonucuna doğrudan etki eden öklid mesafesi gibi uzaklık değerlerine, *k* komşu sayısına ve ağırlıklandırma parametrelerine bağlı çalışmaktadır.

2.1.7. Altuzay KNN Yöntemi

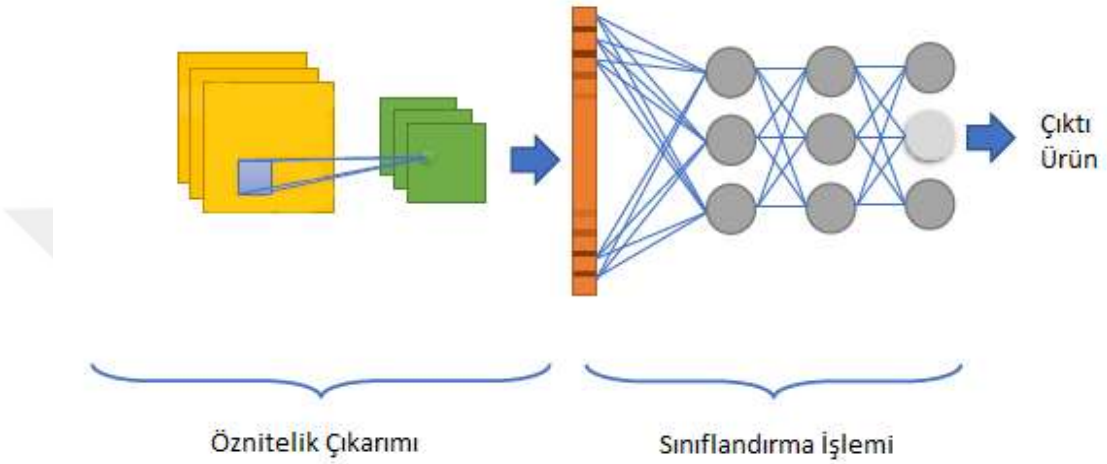
Altuzay KNN yönteminin temel mantığı altuzayların *k* mesafesine bağlı gruplaşmayı sağlamasıdır. Bir alt uzayın alaka durumuna göre kümeleme gerçekleştirilir.

Bu yöntemde öznitelik uzayı üstünde değişiklik yapılarak işlemler gerçekleştirilir. Eğitim seti daha küçük eğitim setlerine bölünerek, her temel öğrenme algoritması ile daha küçük eğitim setleri eğitilir.

2.1.8. Derin Öğrenme Yöntemi

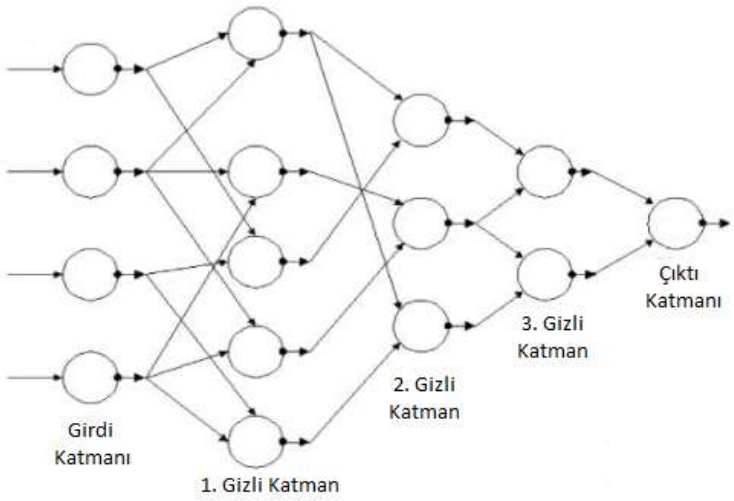
Makine öğrenmesi tekniklerinden biri olan derin öğrenme yöntemi, yapay sinir ağları yapısının çok katmanlı olmasından dolayı veri setinin sınıflandırılmasında kullanılan önemli yöntemler arasında yer almaktadır. Derin öğrenme bu çok katmanlı yapıyı kullanarak temin edilen verinin basitten komplekse doğru çözümlenmesini hedef almaktadır. Derin öğrenme ile ilgili yapılan çalışmalar gelişen işlemci gücüyle beraber 2005 yıllarından sonra ivme kazanmıştır.

Genel olarak makine öğrenmesi tekniklerinde doğruluk değerlerinin yüksek olması özelliklerin nasıl seçildiğiyle ilişkilidir. Derin öğrenme yönteminde verilerin özniteliklerini öğrenmek için çoklu katmanlar bulunmaktadır[45]. Bu katmanlar için girdi verisi kendisinden bir önceki katmanın çıktı verisi olarak tahsis edilmektedir[46]. Şekil 2.12. 'de oto kodlayıcı yapısı verilmiştir.



Şekil 2.12. Derin Öğrenme

Derin öğrenme yöntemi veriler için özellik seviyesinin öğrenilmesini içeren bir yapıya sahiptir. Bu seviye bir görüntü için piksellerin yoğunluk değerleri olabilir. Bu öğrenmeye bağlı sınıflandırma işlemlerini gerçekleştirmek için seçkin algoritmalar kullanılmaktadır. Algoritmalara öncülük eden mimari ise Ivankhnenko'ya aittir[46]. Şekil 2.13. 'de ilk derin öğrenme mimarisi gösterilmiştir.

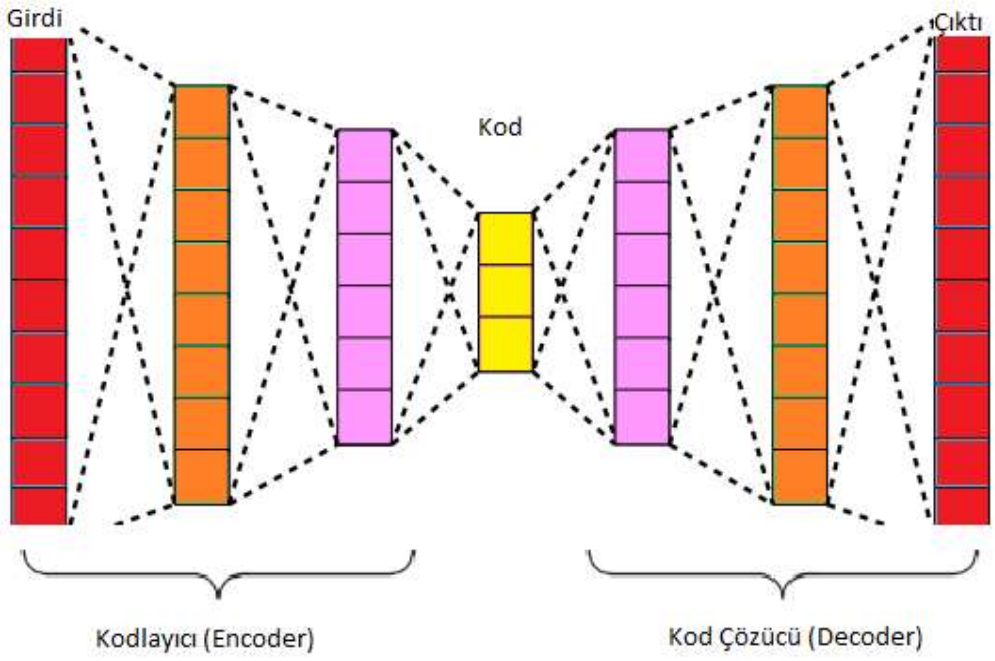


Şekil 2.13. İlk Derin Öğrenme Mimarisi

2.1.8.1. Oto Kodlayıcı

Makine öğrenmesi tekniklerinden biri olan oto kodlayıcı yöntemi yapay sinir ağlarından meydana getirilmiş bir yöntemdir[47]. Oto kodlayıcı temin edilmiş veri seti için boyut küçültme, sınıflandırma gibi işlemler için ideal bir konumdadır. Bir başka deyişle oto kodlayıcı, girdi katmanında bulunan değerleri çıktı katmanına aktaran bir yapay sinir ağıdır.

Oto kodlayıcılar, verilerin eğitimi sırasında üretilen verileri tekrar öğrenerek sınıflandırma işlemini gerçekleştiren yöntemdir. Oto kodlayıcılar, kodlayıcı(enkoder) ve kod çözücü(dekoder) olmak üzere iki ayrı fazdan oluşmaktadır. Kodlayıcı bölümünde işlem yapılırken transfer fonksiyonu kullanılarak veri boyutu farklı bir veriye dönüştürülmektedir. Kod çözücü bölümünde veri yapılan işlemler ile kendisiyle aynı boyutta olan verilere dönüştürülmektedir. Bu aşamalar gizli katmanlar sayesinde gerçekleştirilmektedir[48]. Şekil 2.14. 'de kodlayıcı ve kod çözücü yapısı verilmiştir.



Şekil 2.14. Kodlayıcı ve Kod Çözücü Yapısı

Oto kodlayıcı yönteminde de diğer derin öğrenme yöntemlerinde olduğu gibi aşırı öğrenme problemiyle karşılaşılabilinmektedir. Bu aşırı eğitim durumuyla karşılaşıldığında sınıflandırmanın doğruluk oranında bir azalma gözlemlenmektedir. Problemin aşılması için iterasyon sayısı kısıtlanabilir, var ise gürültü giderilebilir.

2.1.8.2. Softmax

Derin öğrenme ağının son katmanı olan softmax regresyon sınıflandırma işlemlerinin tamamlandığı bölümdür. Softmax, lojistik regresyonun genel hale getirilmiş şekli konumundadır. Lojistik regresyon ile sınıflandırma işlemi için sonuç olasılık değeri gözlemlenmektedir[48]. Oto kodlayıcı yapı olarak softmax regresyonunu sınıflandırma amaçlı kullanmaktadır[49].

2.2. Kullanılan Yazılımlar ve Materyaller

Tez için yapılacak uygulamada kullanılmak üzere Ankara bölgesine 05.06.2018 tarihinde görüntülenmiş Landsat 8 uydu görüntüsü ve 19.09.2018 tarihinde görüntülenmiş Sentinel 2 uydu görüntüsü, https://glovis.usgs.gov ve https://scihub.copernicus.eu sitelerinden temin edilmiştir.

Temin edilen 30 m mekansal çözünürlüğe sahip Landsat 8 uydu görüntüsünün ve 10 m mekansal çözünürlüğe sahip Sentinel 2 uydu görüntüsünün sınıflandırma öncesi görselleri Şekil 2.15. ve Şekil 2.16. 'daki gibidir.



Şekil 2.15. Landsat 8 Uydu Görüntüsü



Şekil 2.16. Sentinel 2 Uydu Görüntüsü

Landsat 8 ve Sentinel 2 uydu görüntülerinin ön işlem adımların gerçekleştirilmesi ve sınıflandırmanın yapılması için ERDAS IMAGINE, ENVI ve MATLAB programları ile gerçekleştirilmiştir.

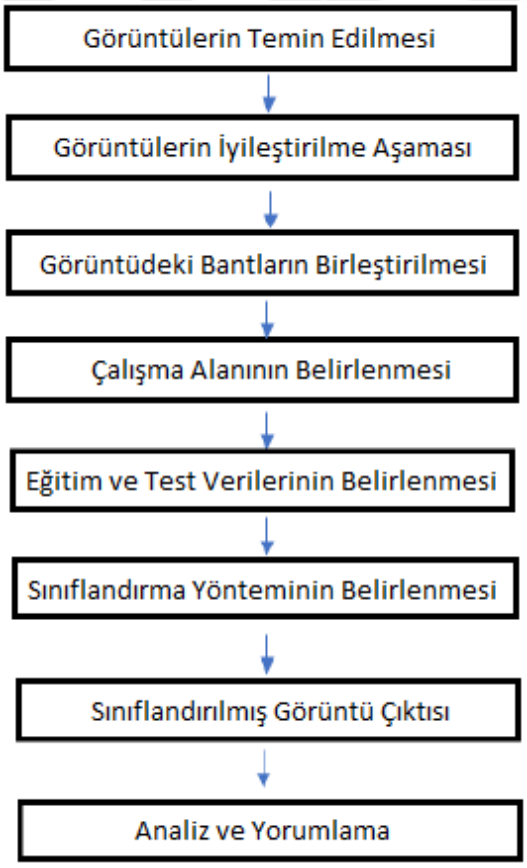
3. BÖLÜM

UYGULAMA VE BULGULAR

3.1. Uygulama Aşamaları

Görüntülerin ön işlemden geçirilmesi ve sınıflandırma işlemleri için aşağıdaki akış şemasına uygun hareket edilmiştir.

Tablo 3.1. Akış Şeması



3.2. Sınıflandırma Sonucu Elde Edilen Değerlerin Matematiksel İfadeleri

Sınıflandırma işlemleri bittikten sonra elde edilen değerler hesaplanırken aşağıdaki değerlere bağlı ifadeler kullanılmıştır.

- *tp* : pozitif sınıftaki örnek sayısı toplamı
- *tn* : negatif sınıftaki örnek sayısı toplamı
- *p* : eğitim örneklerinin bulunduğu matristeki pozitif eleman sayısı
- *n* : eğitim örneklerinin bulunduğu matristeki negatif eleman sayısı
- *N* : eğitim örneklerinin bulunduğu matristeki elemanlar toplamı
- *fp* : yanlış pozitif değerler
- *fn* : yanlış negatif değerler
- *tp_rate* : pozitif sınıftaki örnek sayısı toplamının eleman sayısına oranı
- *tn_rate* : negatif sınıftaki örnek sayısı toplamının eleman sayısına oranı

Bu değerler ışığında doğruluk değeri(accuracy), duyarlılık(sensitivity), özgüllük(specificity), hatırlama(recall), f mesafesi(f-measure) ve g ortalama(g mean) değerleri aşağıdaki gibi hesaplanmıştır.

$$accuracy = (tp + tn) \,/\, N \tag{3.1}$$

$$sensitivity = tp_rate \tag{3.2}$$

$$specificity = tn_rate \tag{3.3}$$

$$precision = tp/(tp + fp) \tag{3.4}$$

$$recall = sensitivity \tag{3.5}$$

$$f_{measure} = 2 * \big((precision * recall)/(precision + recall)\big) \tag{3.6}$$

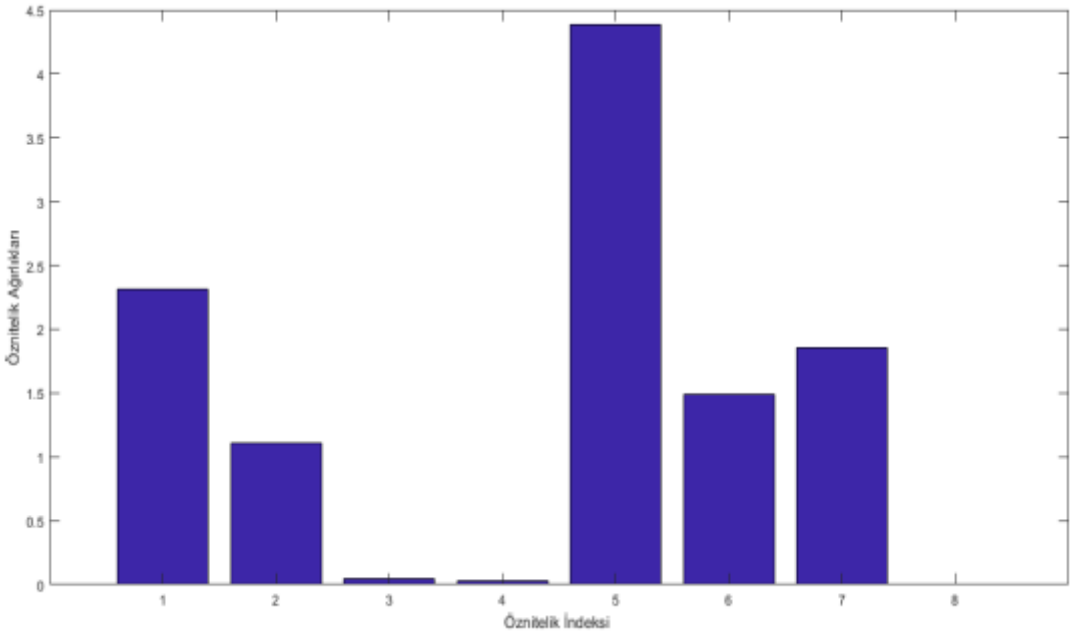
$$g_{mean} = \sqrt{tp_rate * tn_rate} \tag{3.7}$$

3.3. Komşuluk Bileşen Analizi ile Bant Ağırlıklarının Bulunması

Metrik öğrenme algoritmaları arasında başarılı bir pozisyonda bulunan komşuluk bileşen analizi(NCA) öznitelik seçiminde sıkça kullanılmaya başlanmıştır[34]. Bu yöntem ile öznitelik sayısının azaltılarak sınıflandırma işleminin gerçekleştirilmesi amaçlanmaktadır. Kullanılan parametreler aşağıdaki gibidir.

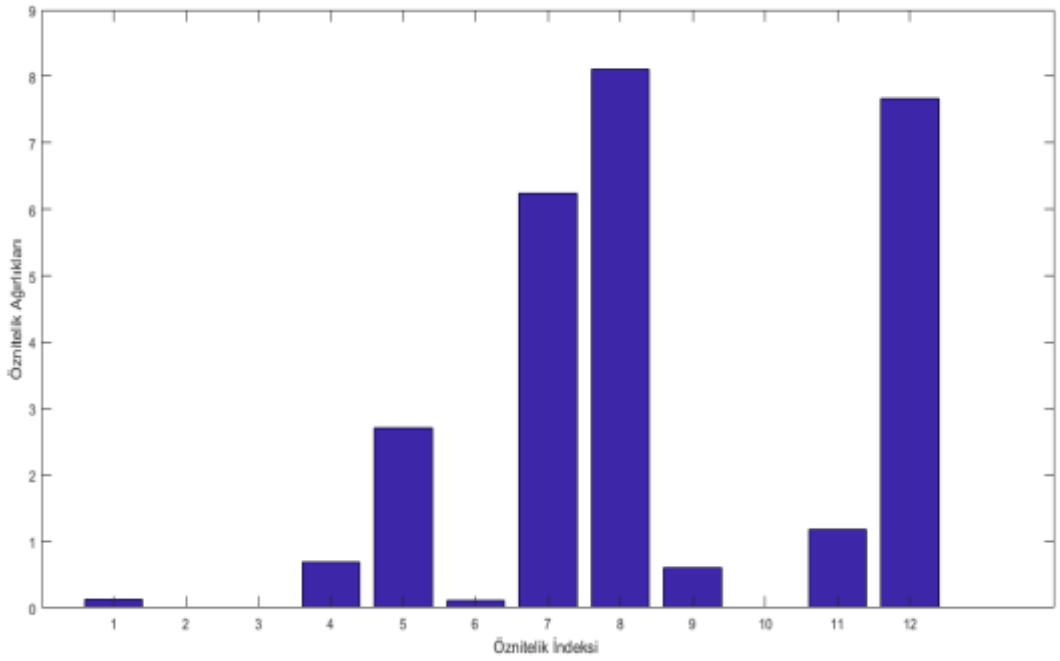
- İterasyon : 1000
- Öğrenme Oranı : 0,5

Sınıflandırma işlemine etki eden bant listeleri her iki görüntü için Şekil 3.1. ve Şekil 3.2. 'deki gibidir.



Şekil 3.1. Sınıflandırmaya Etki Eden Landsat 8 Bantları

Yapılan sınıflandırma işlemine Landsat 8 görüntüsü için en çok etki eden bant 5. bant olarak belirlenmiştir. 5. bant yakın kızılötesi banttır ve toprak, bitki, su yüzeyleri gibi verilerin farklılıklarının belirginleştirilmesinde kullanılmaktadır.



Şekil 3.2. Sınıflandırmaya Etki Eden Sentinel 2 Bantları

Komşuluk bileşen analizi ile elde edilen bant sıralamasında Sentinel 2 görüntüsü için 8. bandın etkisinin diğer bantlara göre daha yüksek olduğu görülmüştür. Sentinel 2 görüntüsünün içerdiği 8. bant Landsat 8 görüntüsünün 5. bandı gibi yakın kızılötesini algılayan bant konumundadır.

İki görüntü tipinden de elde edilen bant sıralaması incelendiğinde sınıflandırma için önemli verileri içeren bantların yakın kızılötesi bölgeyi temsil eden bantlar olduğu ve sınıflandırma işlemleri için önemli bir altlık durumunda olduğu gözlemlenmiştir.

Sınıflandırma işlemleri kullanılan yöntemlerin anlamlı doğruluk değerleri verebilmeye başladığı bantların sınıflandırılması ve tüm bantların sınıflandırılması olarak yapılmıştır. Bu sayede bantların sınıflandırma üzerindeki etkisinin de analizi hedeflenmiştir.

3.4. Anlamlı Sonuç Elde Etmeye Başlanan Bant Sıralamaları

Bu çalışmada tüm bantlarla yapılacak sınıflandırma uygulamasına ek olarak NCA yöntemi ile analiz sonucunda elde ettiğimiz bant sıralamasının sınıflandırmayı hangi sıradan sonra anlamlı hale getireceği de araştırılmıştır. Bu bilgiler ışığında aşağıdaki tabloda Landsat 8 ve Sentinel 2 görüntüleri için hangi yöntemde hangi bantlarla anlamlı

sonuçlar elde edildiği gösterilmiştir ve kullanılan yöntemlerin içerisinde sonuçlar verilmiştir. Tablo 3.2. 'de anlamlı sonuç veren bant sıralamaları gösterilmiştir.

Tablo 3.2. Anlamlı Sonuç Elde Etmeye Başlanan Bant Sıralamaları

Bant Sıralamaları	**Landsat 8**	**Sentinel 2**
Destek Vektör Makineleri	Bant 5,1,7	Bant 8,12,7,5,11
Karar Ağaçları	Bant 5,1,7	Bant 8,12,7,5,11
Torbalama Karar Ağaçları	Bant 5,1	Bant 8,12
Lineer Diskriminant	Bant 5,1,7,6	Bant 8,12,7,5,11
K-En Yakın Komşu	Bant 5,1	Bant 8,12
Altuzay KNN	Bant 5,1,7,6	Bant 8,12,7,5,11
Derin Öğrenme	Bant 5,1,7	Bant 8,12,7,5,11

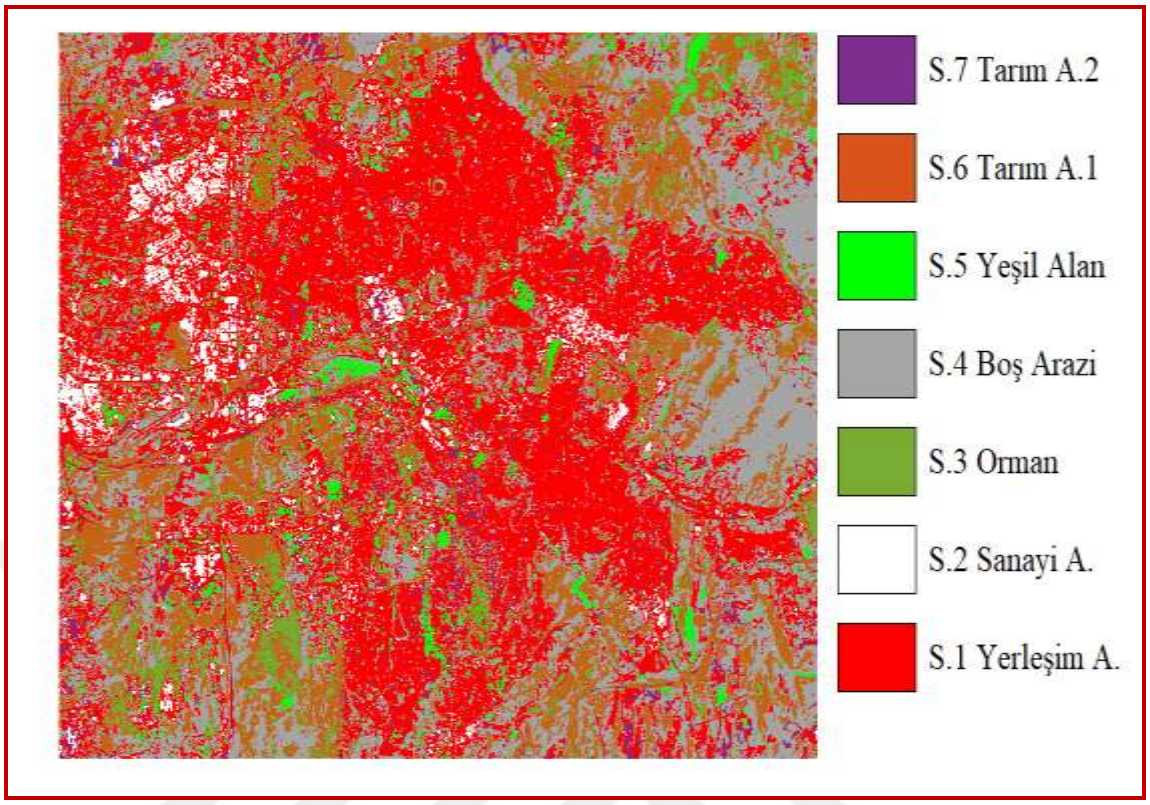
3.5. Destek Vektör Makineleri ile Sınıflandırma İşlemleri

Destek vektör makineleri, parametrik olmayan ve istatistiksel öğrenme temelli sınıflandırma algoritmasıdır. Çok sınıflı veri setlerinin sınıflandırılmasında yaygın bir kullanıma sahip olan destek vektör makineleri için hangi kernel fonksiyonunun kullanılacağı görüntünün sınıflandırma sonrası doğruluk değerine doğrudan etki eden bir konumdadır.

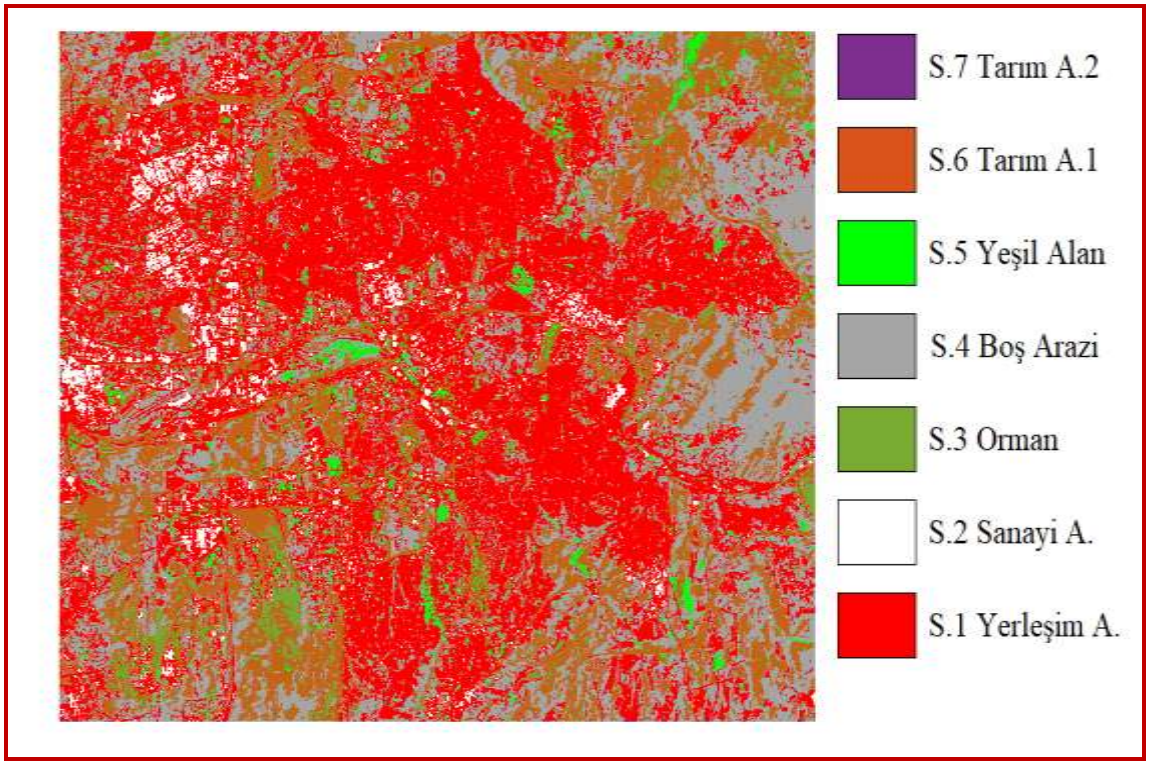
Tezin bu kısmında destek vektör makineleri ile sınıflandırma işlemleri için aşağıdaki parametreler kullanılmıştır.

- Kernel Fonksiyonu : Gaussian, Linear
- Kernel Ölçeği : 5
- Sınıflandırma : Bire Karşı Bir

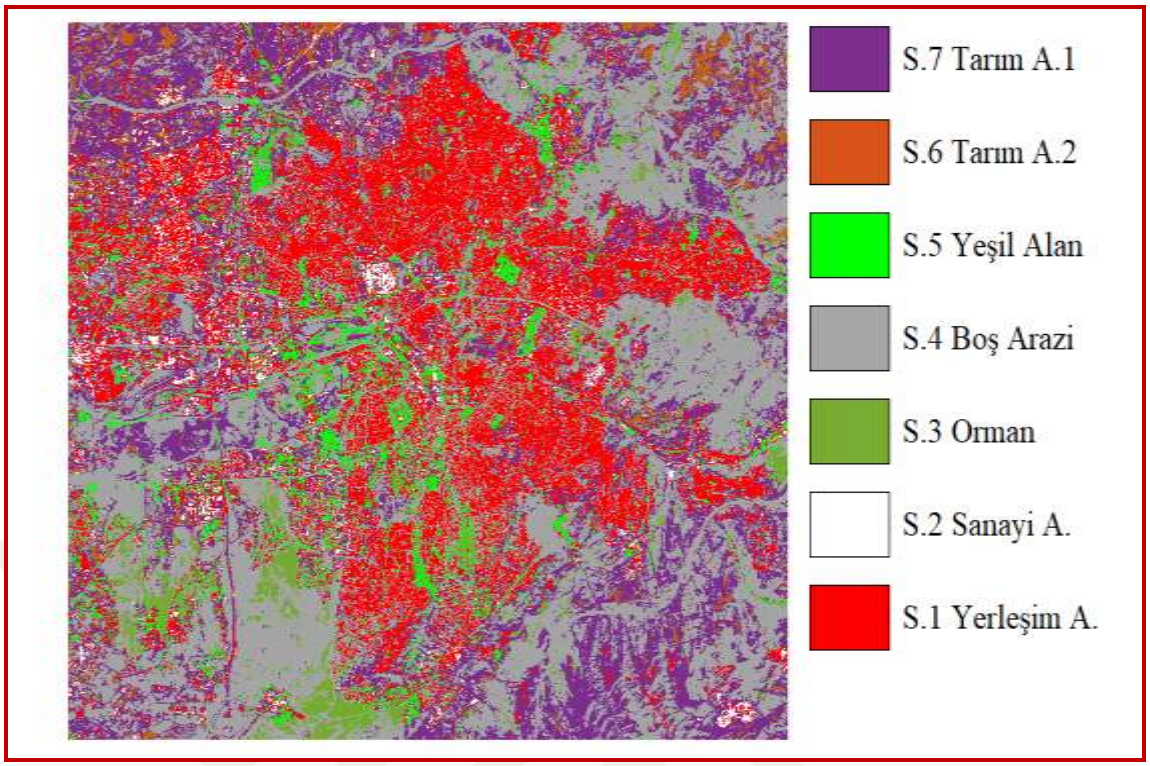
Sınıflandırma işlemine NCA ile bulduğumuz sınıflandırmaya etki eden bantlardan Landsat 8 görüntüsünün 5., 1., 7., bantları ve Sentinel 2 görüntüsünün 8., 12., 7., 5., 11., bantları SVM ile sınıflandırılarak başlanmıştır. Şekil 3.3. ve Şekil 3.4.' de Landsat 8 görüntüsünün Gaussian ve Lineer kerneli kullanılarak sınıflandırılmış çıktı ürünleri ve Şekil 3.5. ve Şekil 3.6.' da ise Sentinel 2 görüntüsünün Gaussian ve Lineer kerneli kullanılarak sınıflandırılmış çıktı ürünleri gösterilmiştir.



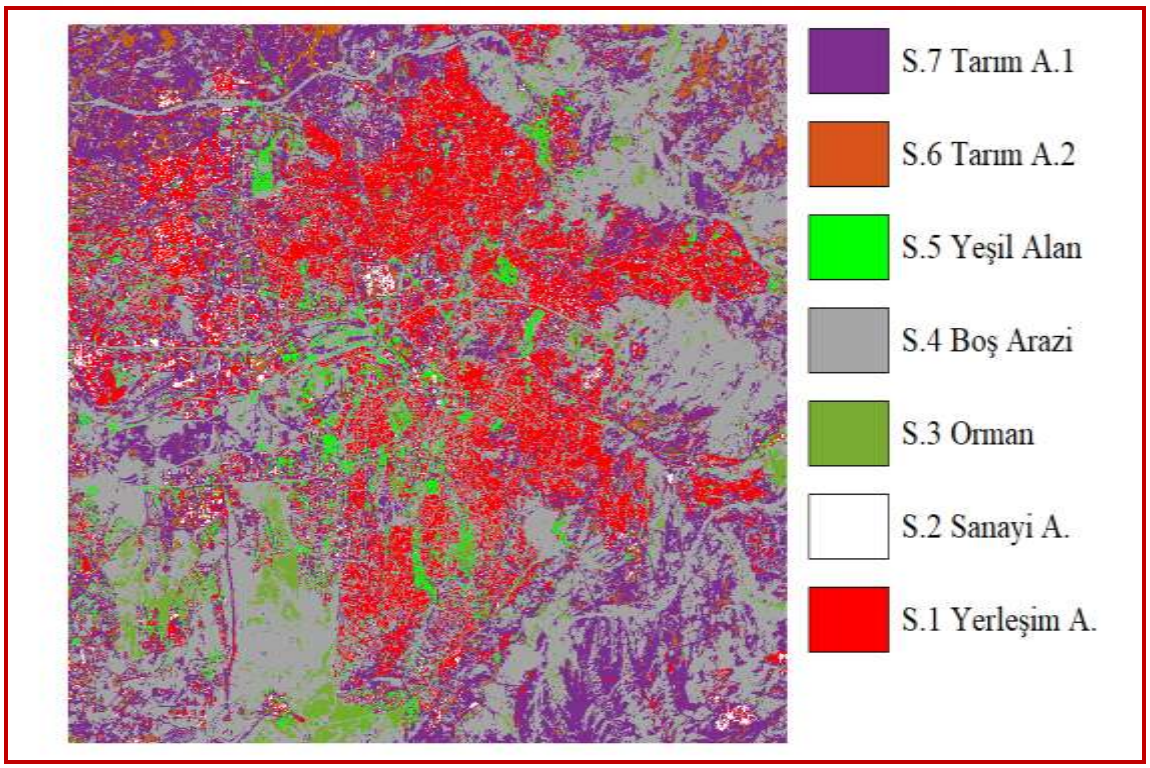
Şekil 3.3. SVM ile Sınıflandırılmış Landsat 8 (Bant:5,1,7) Görüntüsü (Kernel: Gaussian)



Şekil 3.4. SVM ile Sınıflandırılmış Landsat 8 (Bant:5,1,7) Görüntüsü (Kernel: Lineer)



Şekil 3.5. SVM ile Sınıflandırılmış Sentinel 2 (Bant: 8,12,7,5,11) Görüntüsü (Kernel: Gaussian)



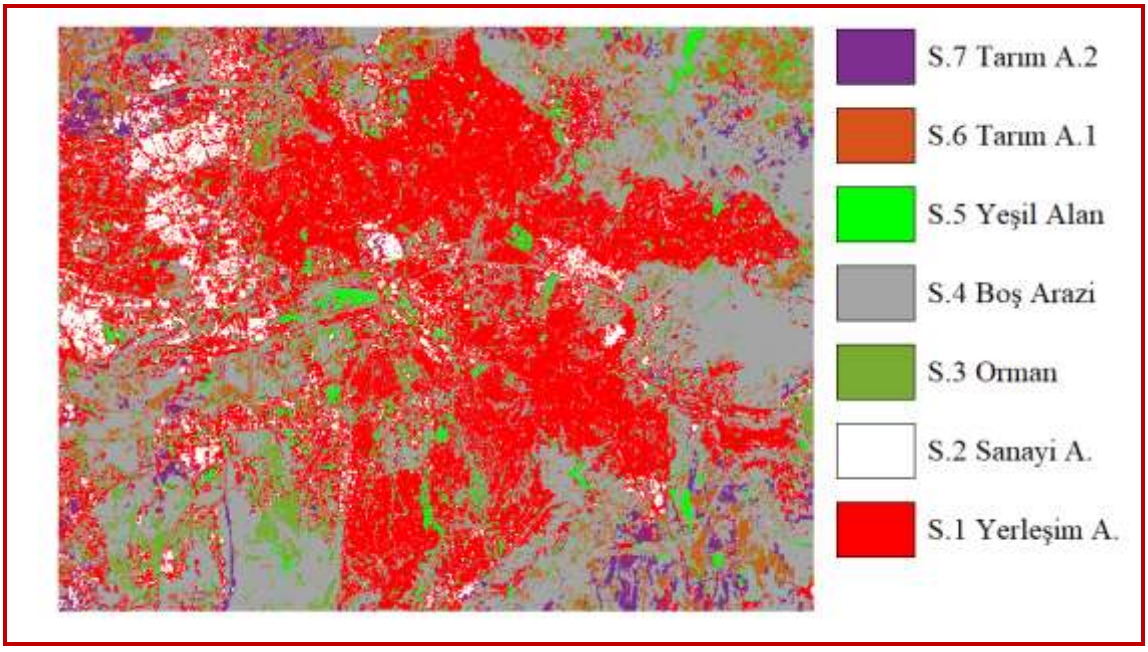
Şekil 3.6. SVM ile Sınıflandırılmış Sentinel 2 (Bant: 8,12,7,5,11) Görüntüsü (Kernel: Lineer)

Tablo 3.3. SVM ile Sınıflandırılmış Landsat 8 (Bant:5,1,7) ve Sentinel 2 (Bant: 8,12, 7,5,11) Görüntüleri Doğruluk Değerleri

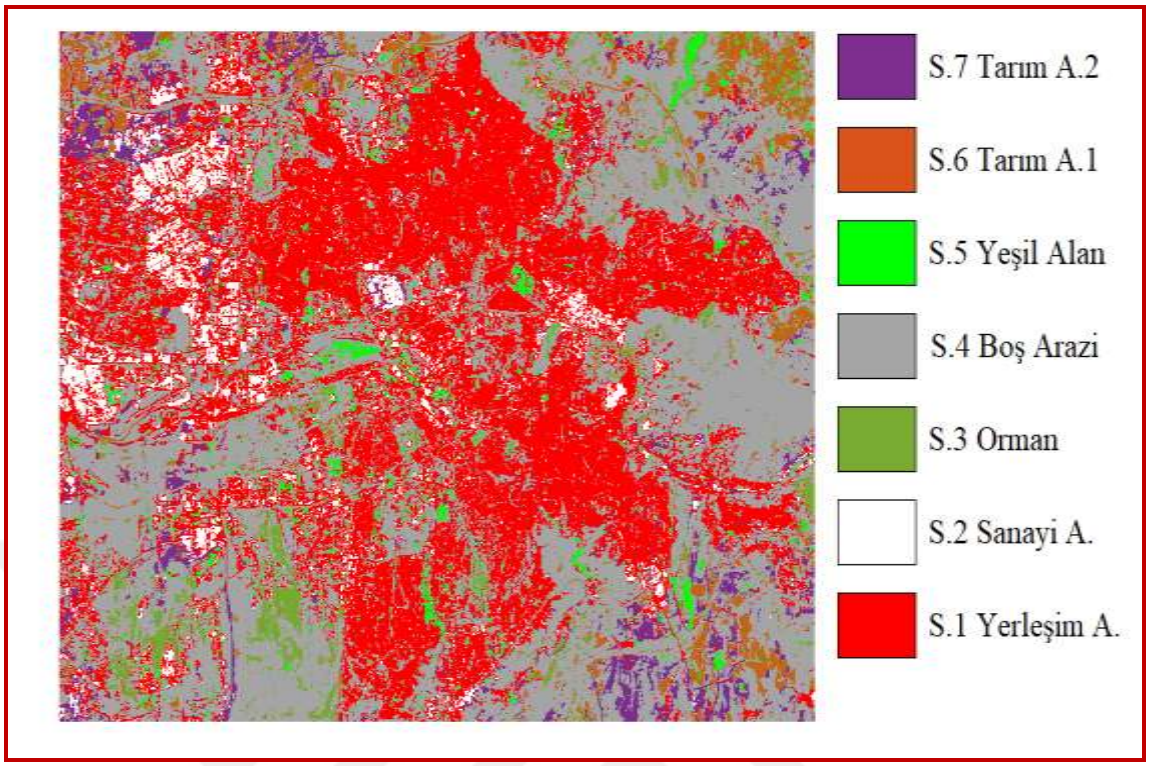
	Eğitim Doğruluğu	Test Doğruluğu	Duyarlılık (Sensitivity)	Özgüllük (Specificity)	Hassaslık (Presicion)	Hatırlama (Recall)	F-measure	G-mean
Landsat 8	0,8630 0,8384	0,8644 0,8292	0,9480 0,9664	0,8394 0,7883	0,6379 0,5766	0,9480 0,9664	0,7626 0,7223	0,8921 0,8728
Sentinel 2	0,9111 0,8908	0,9098 0,8913	0,9330 0,9152	0,9068 0,8883	0,5573 0,5074	0,9330 0,9152	0,6978 0,6529	0,9198 0,9017

Landsat 8 ve Sentinel 2 görüntüleri için sınıflandırmaya etki eden ilk 3 bant ile yapılan sınıflandırma sonucu elde edilen değerler analiz edildiğinde destek vektör makinelerinin yapısının Sentinel 2 görüntüsüyle daha yüksek sonuçlar verdiği görülmüştür. Ancak diğer bantların içerdiği bilgileri sınıflandırma işlemine eklediğimiz zaman SVM yöntemi Sentinel 2 için daha iyi sonuçlar vermektedir. Bu duruma rağmen Landsat 8 görüntüsünün hassaslık değeri ise daha yüksektir.

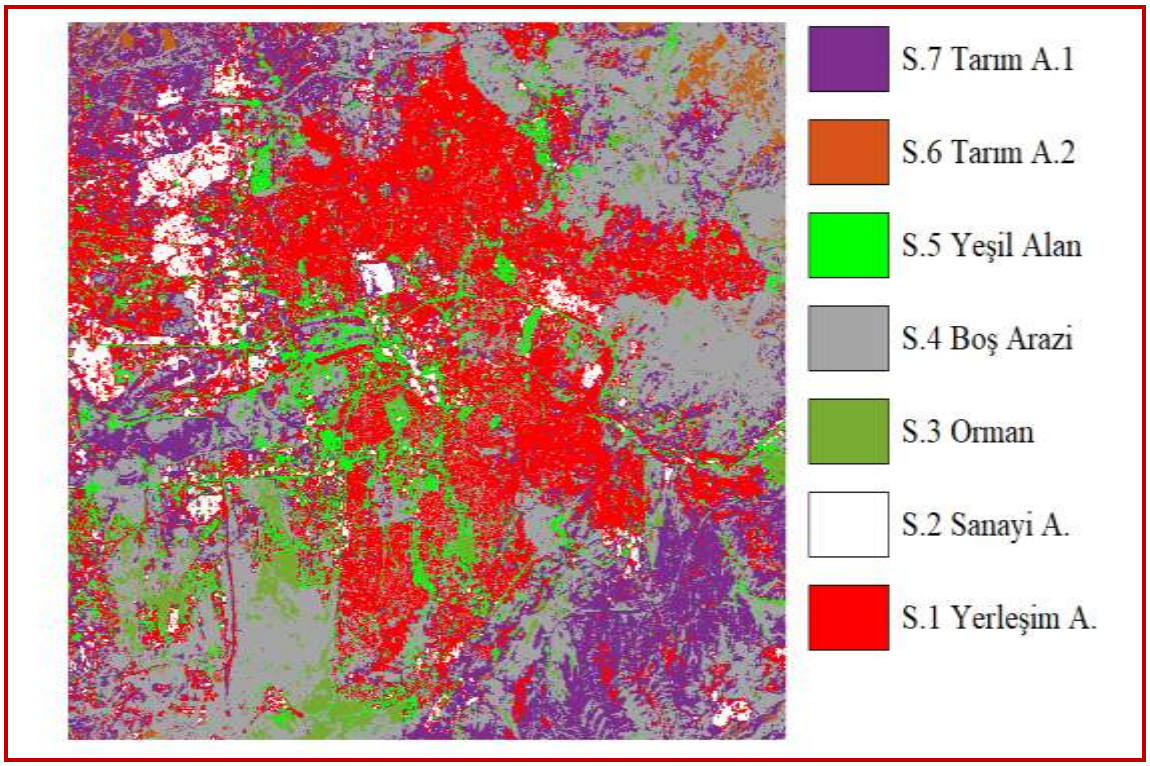
Şekil 3.7., Şekil 3.8., Şekil 3.9. ve Şekil 3.10. 'da ise Landsat 8 ve Sentinel 2 görüntülerine ait tüm bantların sınıflandırma işlemine dahil edilerek elde edilen çıktı ürünler gösterilmiştir.



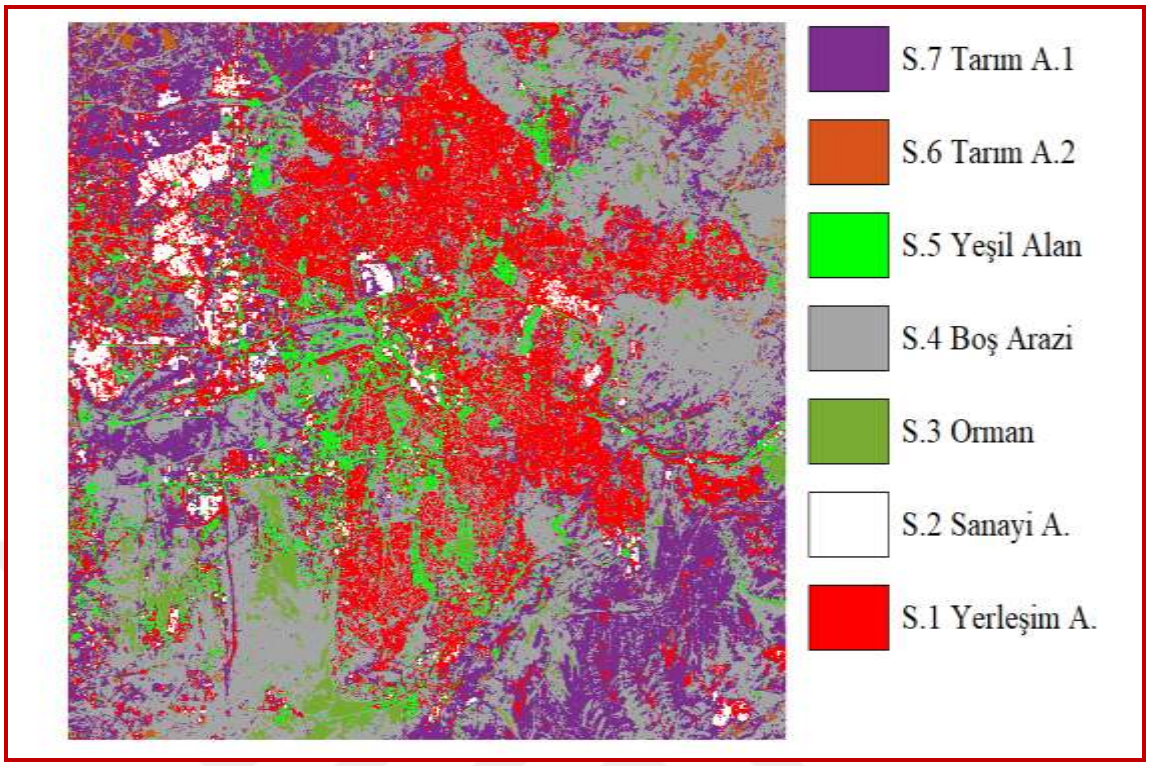
Şekil 3.7. SVM ile Sınıflandırılmış Landsat 8 Görüntüsü(Kernel: Gaussian)



Şekil 3.8. SVM ile Sınıflandırılmış Landsat 8 Görüntüsü(Kernel: Linear)



Şekil 3.9. SVM ile Sınıflandırılmış Sentinel 2 Görüntüsü(Kernel: Gaussian)



Şekil 3.10. SVM ile Sınıflandırılmış Sentinel 2 Görüntüsü(Kernel: Linear)

Görüntülere ait tüm bantların sınıflandırma sonucunda verdiği oğruluk değerleri Tablo 3.4.' de verilmiştir.

Tablo 3.4. SVM ile Sınıflandırma Doğruluk Değerleri

Tüm Bantlar	Eğitim Doğruluğu	Test Doğruluğu	Duyarlılık (Sensitivity)	Özgüllük (Specificity)	Hassaslık (Presicion)	Hatırlama (Recall)	F-measure	G-mean
Landsat 8	0,9567	0,9529	0,9840	0,9442	0,8320	0,9840	0,9016	0,9639
	0,9244	0,9227	0,9872	0,9046	0,7440	0,9872	0,8485	0,9450
Sentinel 2	0,9741	0,9681	0,9643	0,9686	0,7941	0,9643	0,8710	0,9664
	0,9632	0,9591	0,9420	0,9613	0,7536	0,9420	0,8373	0,9516

Yukarıdaki tabloda 1. satır Gaussian kerneli ile yapılan işlem sonucu elde edilen değerleri gösterirken 2. satır Lineer kerneli ile elde edilen değerleri göstermektedir.

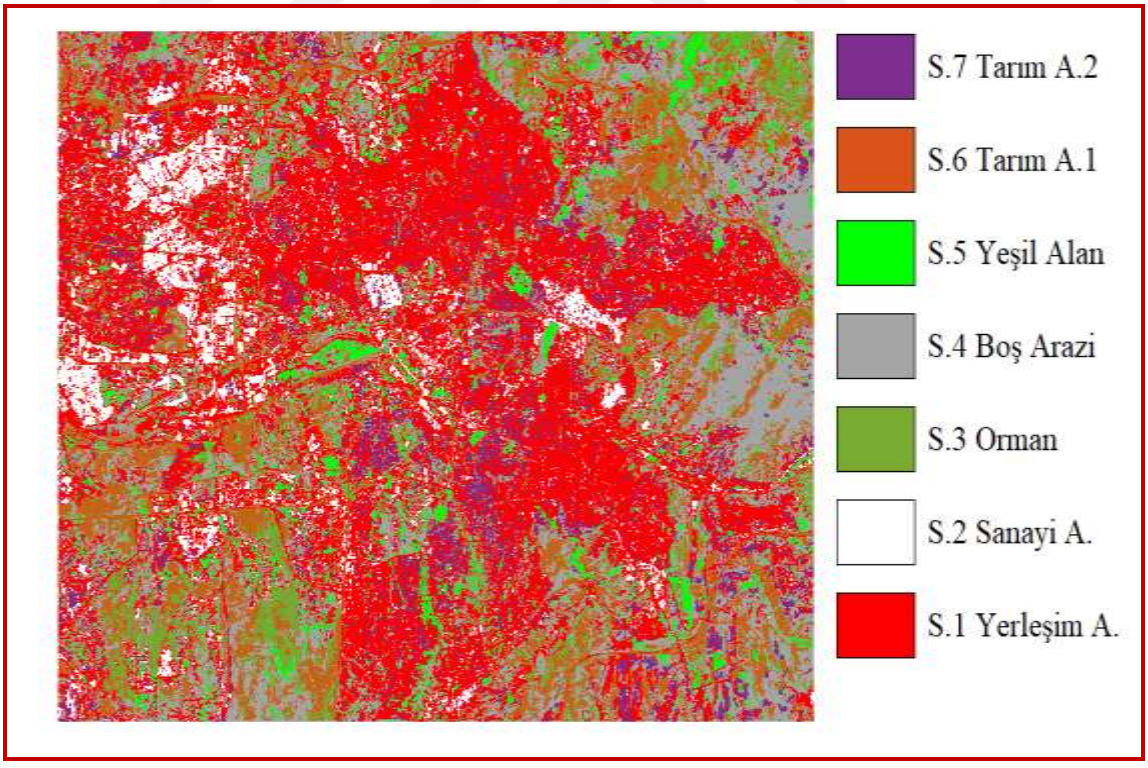
Tablodaki değerlere bakıldığında Gaussian kernelinin Lineer kerneline göre daha başarılı sonuçlar verdiği görülmüştür. Sentinel 2 görüntüsünün SVM ile yapılan sınıflandırmalarda daha iyi sonuçlar verdiği belirlenmiştir. Tüm bantlar ile yapılan sınıflandırmada daha yüksek değerler alınmıştır.

3.6. Karar Ağaçları ile Sınıflandırma İşlemleri

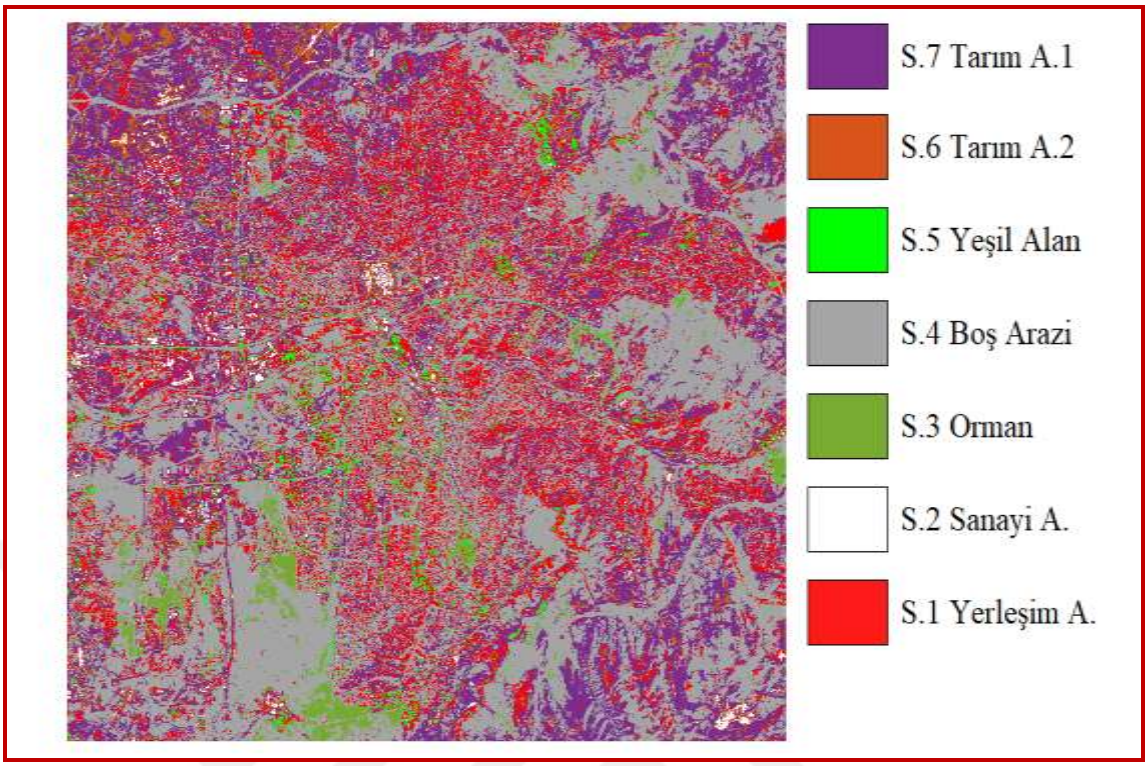
Karar ağaçları yöntemi karmaşık haldeki sınıflandırma problemlerini yapısı gereği aşamalara ayırarak karar verme sürecini basitleştirmesinden dolayı sınıflandırma uygulamalarının tercih edilen yöntemlerinden biridir. Karar ağaçları ile yapılan sınıflandırmada kullanılan parametreler aşağıda belirtilmiştir.

- Bölünme Kriteri : Gini İndeksi
- Maximum Bölünme Sayısı : 20

Landsat 8 ve Sentinel 2 görüntüleri için sınıflandırmaya etki eden bantların sonuçları anlamlı sonuç elde edecek şekilde sınıflandırma işlemlerine dahil edilerek elde edilen çıktı ürünler ve sonuçlar Şekil 3.11., Şekil 3.12., Şekil 3.13., ve Şekil 3.14. 'deki gibidir.



Şekil 3.11. Karar Ağaçları ile Sınıflandırılmış Landsat 8 (Bant:5,1,7) Görüntüsü



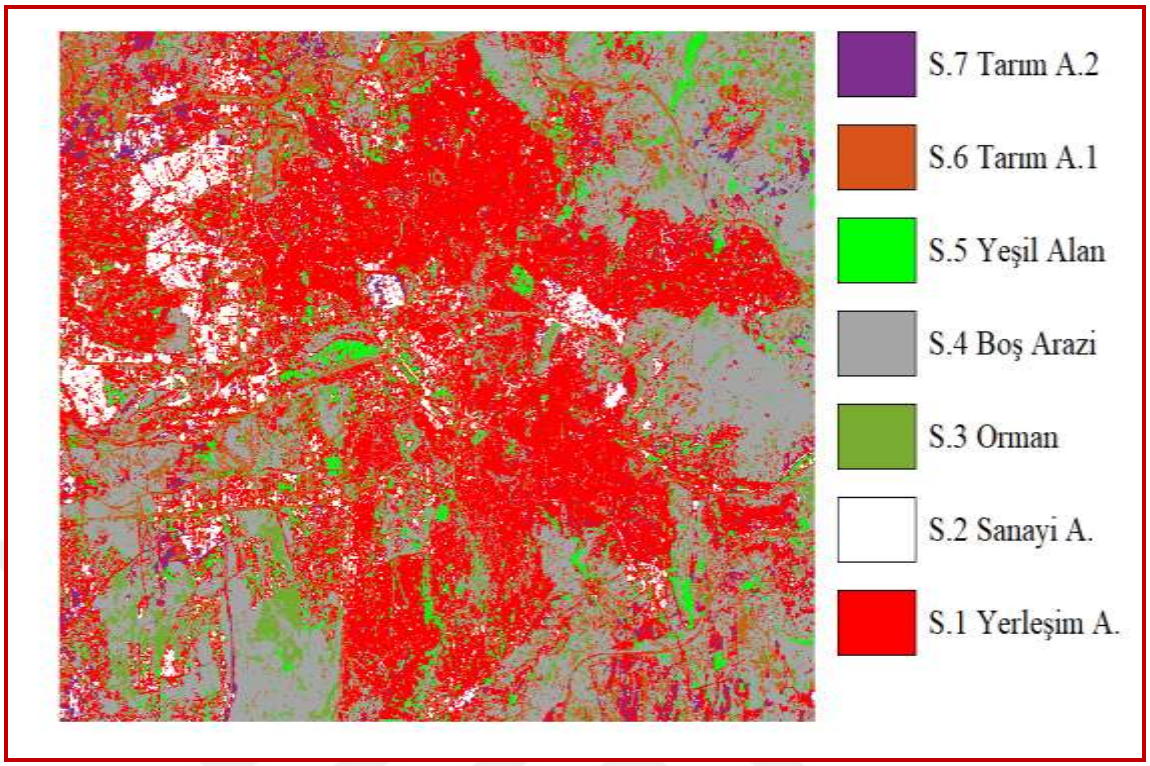
Şekil 3.12. Karar Ağaçları ile Sınıflandırılmış Sentinel 2 (Bant:8,12,7,5,11) Görüntüsü

Tablo 3.5. 'de anlamlı sonuçlar elde edilen bantlar ile yapılan sınıflandırma sonuçları yer almaktadır.

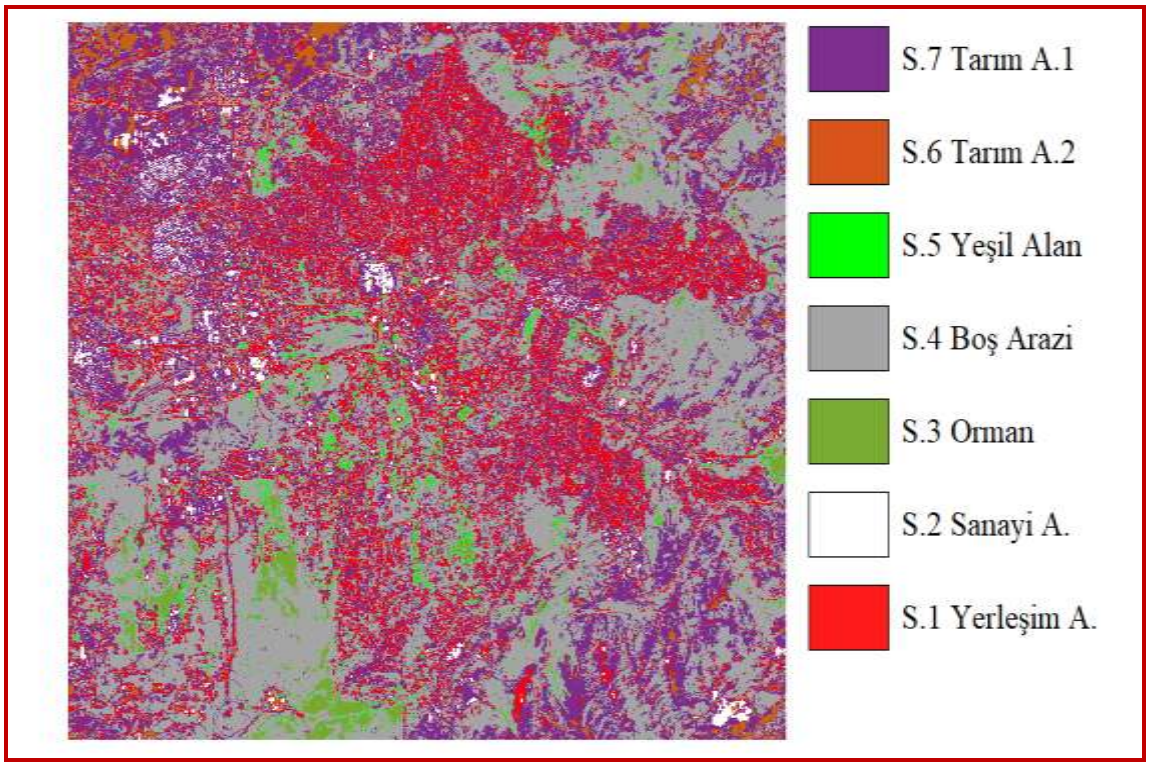
Tablo 3.5. Karar Ağaçları ile Sınıflandırılmış Landsat 8 (Bant:5,1,7) ve Sentinel 2 (Bant:8,12,7,5,11) Görüntüleri Doğruluk Değerleri

	Eğitim Doğruluğu	**Test Doğruluğu**	**Duyarlılık (Sensitivity)**	**Özgüllük (Specificity)**	**Hassaslık (Presicion)**	**Hatırlama (Recall)**	**F-measure**	**G-mean**
Landsat 8	0,8654	0,8391	0,7982	0,8513	0,6156	0,7982	0,6951	0,8243
Sentinel 2	0,8483	0,8340	0,6429	0,8580	0,3627	0,6429	0,4638	0,7427

Seçilen bantlarla yapılan sınıflandırma sonucunda hassaslık değerinin yüksek olması karar ağaçları yönteminin doğruluk değerine etki etmektedir.



Şekil 3.13. Karar Ağaçları ile Sınıflandırılmış Landsat 8 Görüntüsü



Şekil 3.14. Karar Ağaçları ile Sınıflandırılmış Sentinel 2 Görüntüsü

Tüm bantlar kullanılarak elde edilen doğruluk değerleri Tablo 3.6. 'daki gibidir.

Tablo 3.6. Karar Ağaçları ile Sınıflandırma Doğruluk Değerleri

Tüm Bantlar	Eğitim Doğruluğu	Test Doğruluğu	Duyarlılık (Sensitivity)	Özgüllük (Specificity)	Hassaslık (Presicion)	Hatırlama (Recall)	F-measure	G-mean
Landsat 8	0,9025	0,8856	0,9006	0,8776	0,6739	0,9006	0,7709	0,8890
Sentinel 2	0,8650	0,8674	0,6607	0,8934	0,4379	0,6607	0,5267	0,7683

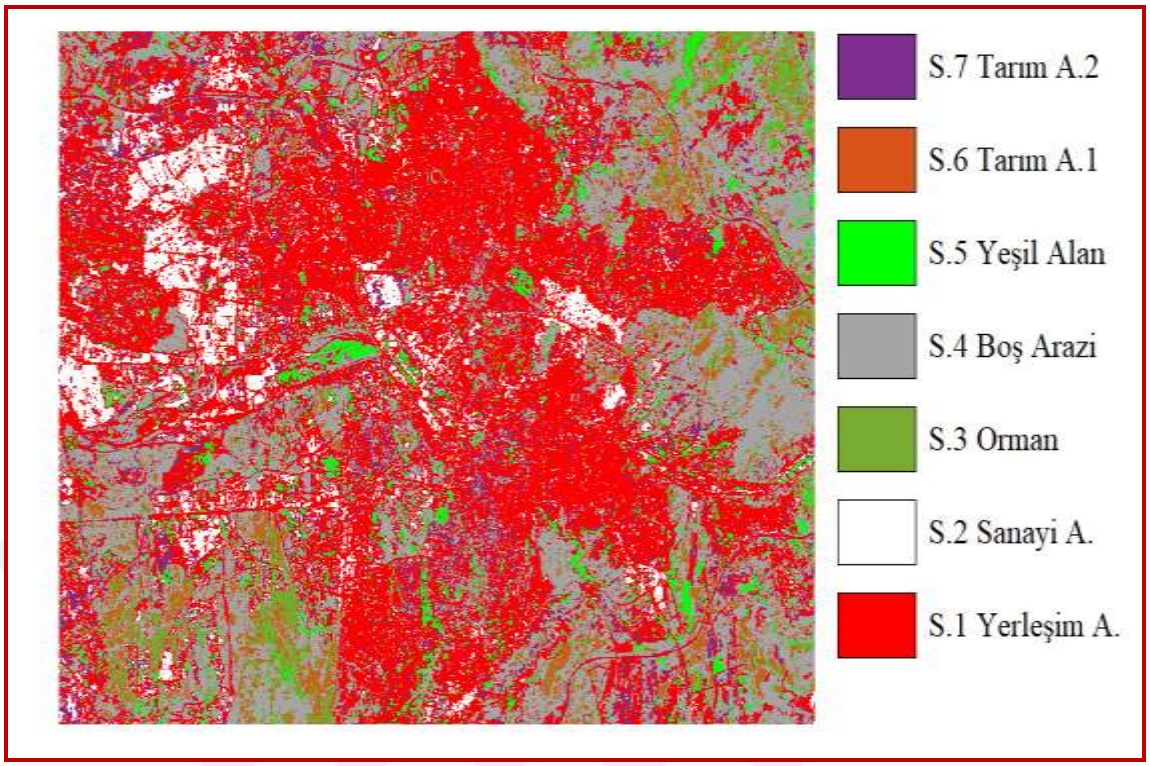
Tabloya göre Landsat 8 verisinin daha yüksek doğruluk değerlerinde çıktı ürün verdiği ve duyarlılık değerinin Sentinel 2 görüntüsüne göre oldukça yüksek olduğu belirlenmiştir. Görüntüler için tüm bantlar kullanılarak elde edilmiş sınıflandırma sonucunun daha yüksek olduğu gözlemlenmiştir.

3.6.1. Torbalama Karar Ağaçları ile Sınıflandırma İşlemleri

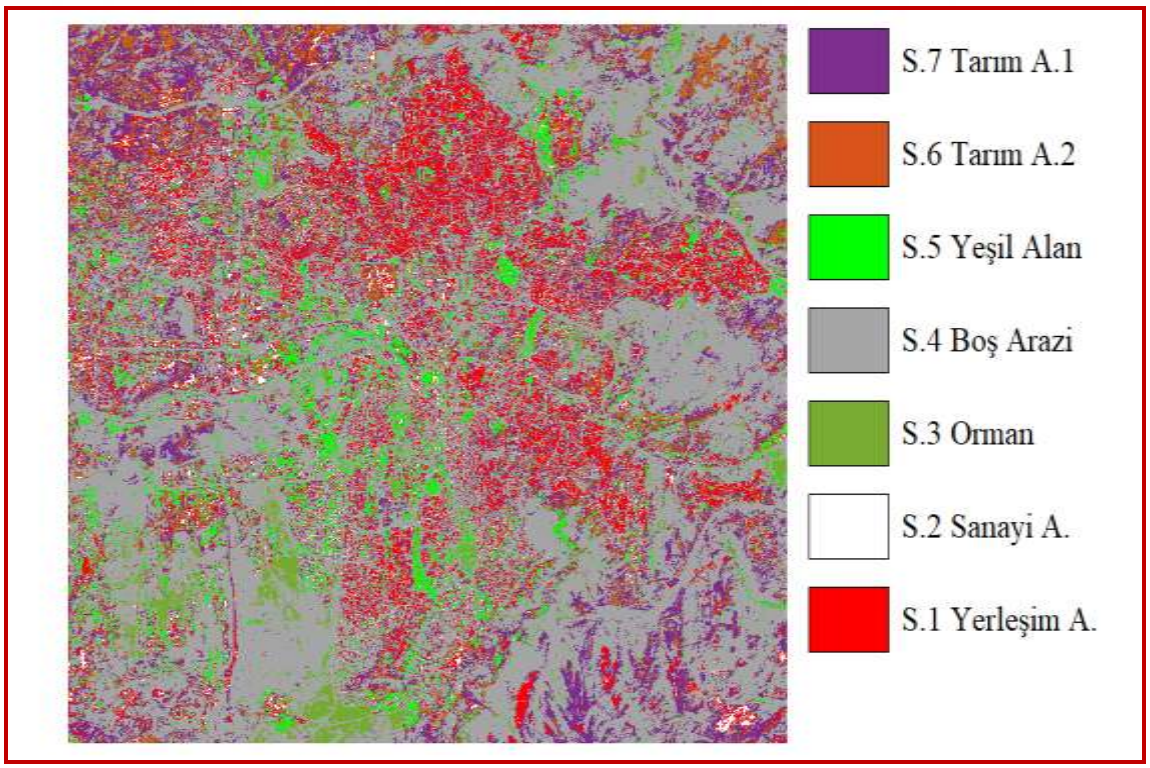
Eğitim setinin belirli bir düzen dışında rastgele olacak bir biçimde yeniden meydana getirilmesi ile sınıflandırıcının eğitiminin tamamlanmasını hedef alan bu yöntem torba boyutlarına ve iterasyon sayısına göre sınıflandırma işlemlerini gerçekleştirmektedir. Torbalama karar ağaçları ile yapılan sınıflandırmada kullanılan parametreler aşağıdaki şekildedir.

- Öğrenme Döngü Sayısı : 30
- Maximum Bölünme Sayısı : 100

Bant sıralamasına göre tatmin edici doğruluk veren bantlar ile yapılan sınıflandırma sonuçları Şekil 3.15. ve Şekil 3.16.' da gösterilmiştir.



Şekil 3.15. Torbalama Karar Ağaçları ile Sınıflandırılmış Landsat 8 (Bant:5,1) Görüntüsü



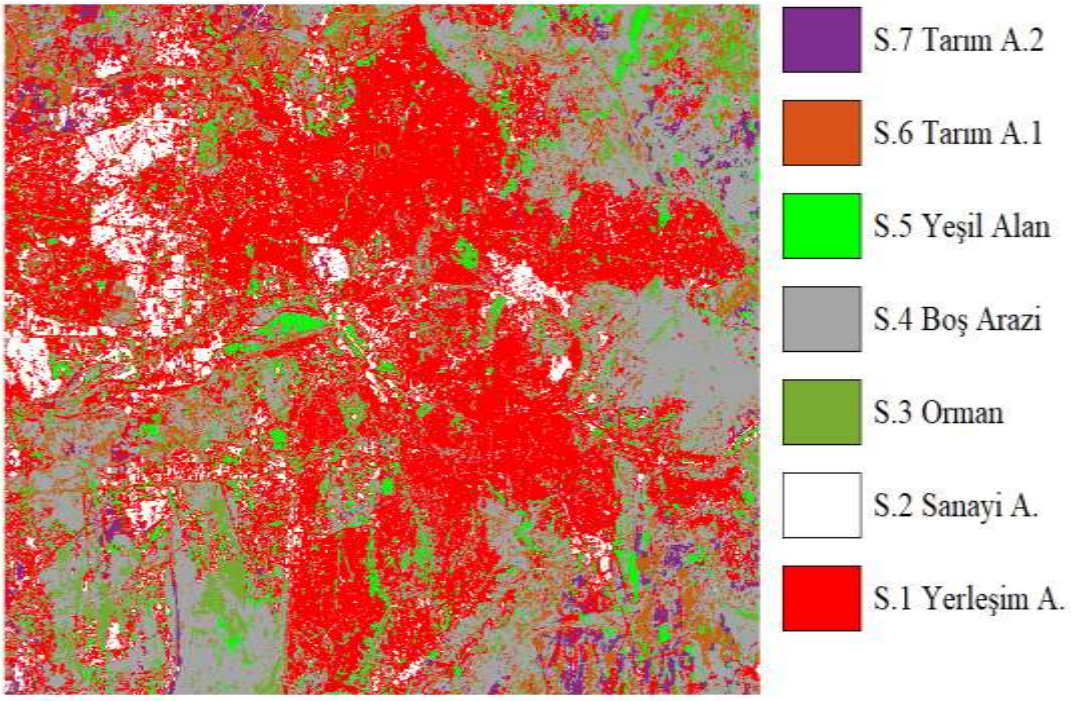
Şekil 3.16.. Torbalama Karar Ağaçları ile Sınıflandırılmış Sentinel 2 (Bant:8,12) Görüntüsü

Tablo 3.7. 'de torbalama karar ağaçları ile anlamlı sonuç elde eden bantların sınıflandırma sonuçları yer almaktadır.

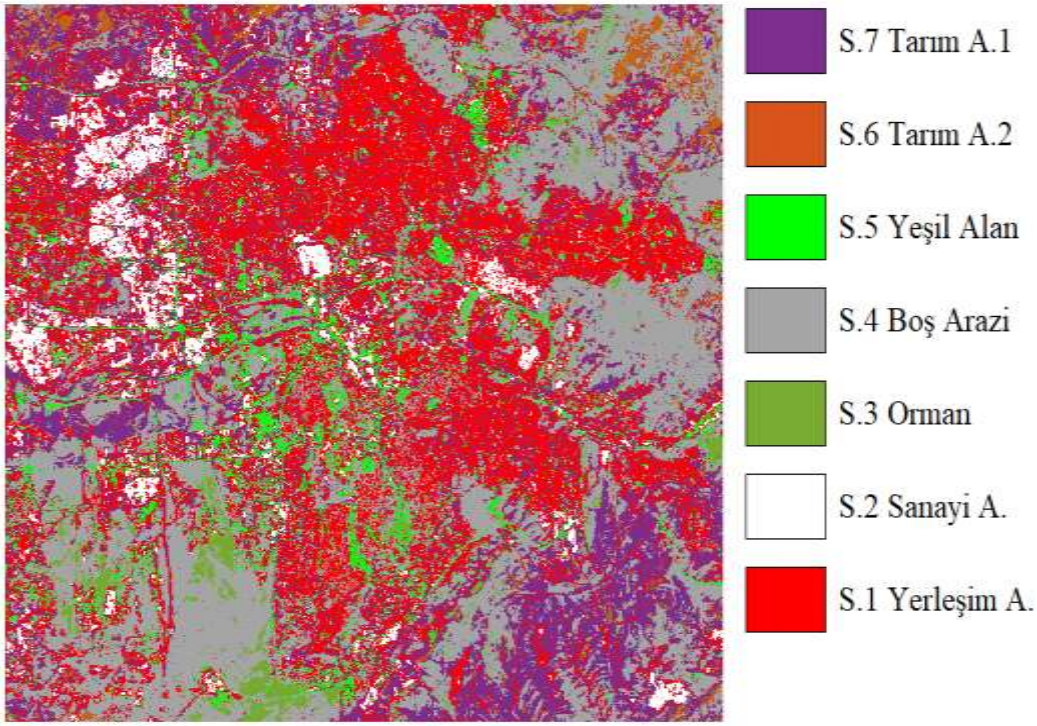
Tablo 3.7. Torbalama Karar Ağaçları ile Sınıflandırılmış Landsat 8 (Bant:5,1) ve Sentinel 2 (Bant:8,12) Görüntüleri Doğruluk Değerleri

	Eğitim Doğruluğu	**Test Doğruluğu**	**Duyarlılık (Sensitivity)**	**Özgüllük (Specificity)**	**Hassaslık (Presicion)**	**Hatırlama (Recall)**	**F-measure**	**G-mean**
Landsat 8	0,8685	0,8124	0,8899	0,7892	0,5575	0,8899	0,6855	0,8381
Sentinel 2	0,8752	0,8425	0,6830	0,8625	0,3844	0,6830	0,4920	0,7675

Torbalama karar ağaçları ile seçilen bantların sınıflandırılmasında 2 bant ile bile yüksek doğruluk değerleri elde edilmiştir. Yöntemin iki görüntü içinde benzer sonuçlar verdiği görülmüştür. Şekil 3.17. ve Şekil 3.18. 'de tüm bantlar ile yapılan sınıflandırma sonuçları gösterilmiştir.



Şekil 3.17. Torbalama Karar Ağaçları ile Sınıflandırılmış Landsat 8 Görüntüsü



Şekil 3.18. Torbalama Karar Ağaçları ile Sınıflandırılmış Sentinel 2 Görüntüsü

Bu yöntemle tüm bantlar kullanılarak yapılan sınıflandırma sonucu doğruluk değerleri Tablo 3.8.' deki gibidir.

Tablo 3.8. Torbalama Karar Ağaçları ile Sınıflandırma Doğruluk Değerleri

Tüm Bantlar	Eğitim Doğruluğu	Test Doğruluğu	Duyarlılık (Sensitivity)	Özgüllük (Specificity)	Hassaslık (Presicion)	Hatırlama (Recall)	F-measure	G-mean
Landsat 8	0,9895	0,9592	0,9840	0,9523	0,8528	0,9840	0,9137	0,9680
Sentinel 2	0,9840	0,9636	0,9509	0,9652	0,7745	0,9509	0,8537	0,9580

Performans olarak karar ağaçlarını geride bırakan bu yöntem ile görüntüler memnun edici doğruluk değerlerine ulaşarak sınıflandırılmıştır. Her bir bant kullanılarak yapılmış sınıflandırma işlemlerinde her yeni eklenen bandın doğruluk değerlerine olumlu yönde katkısı olduğu gözlemlenmiştir.

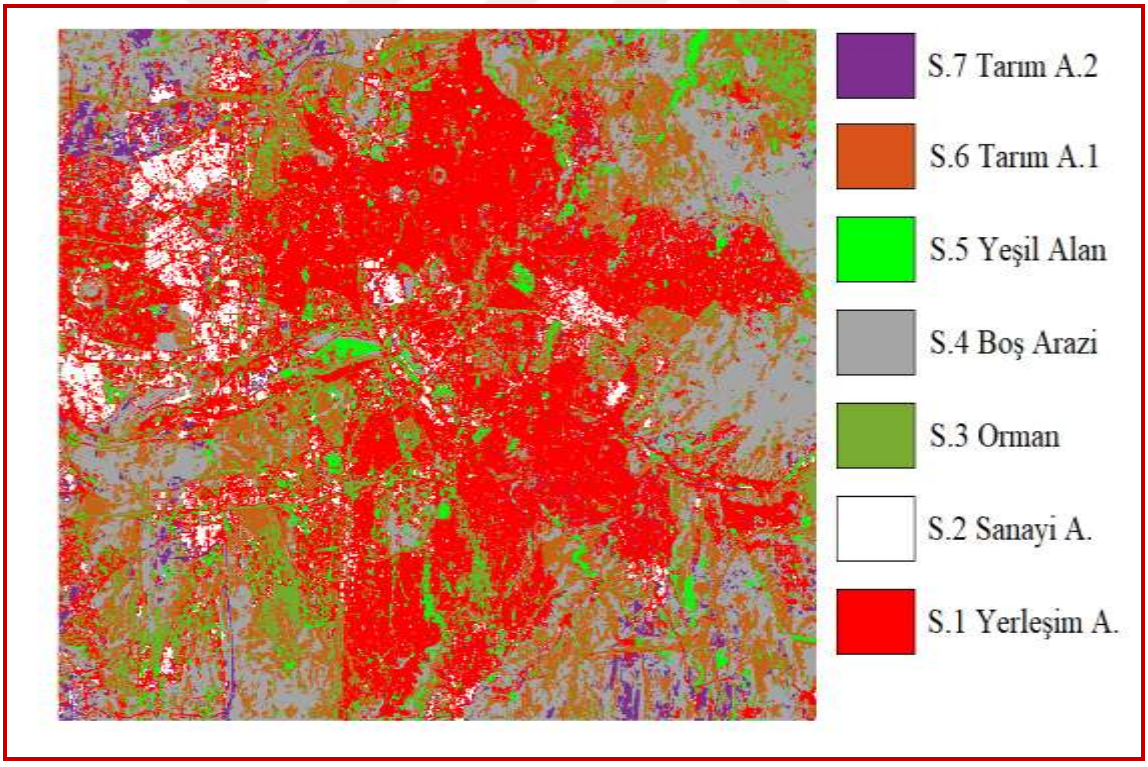
3.7. Lineer Diskriminant ile Sınıflandırma İşlemleri

1936 yılında Fisher tarafından geliştirilen yöntem, karmaşık problemlerde iyi sonuçlar üretebilen bir sınıflandırma yöntemidir. Bu yöntem, sınıfları en iyi bir biçimde ayırabilen değişkenlerin lineer formunu bulmaya dayanmaktadır. Bu yöntem ile yapılan sınıflandırmada kullanılan parametreler;

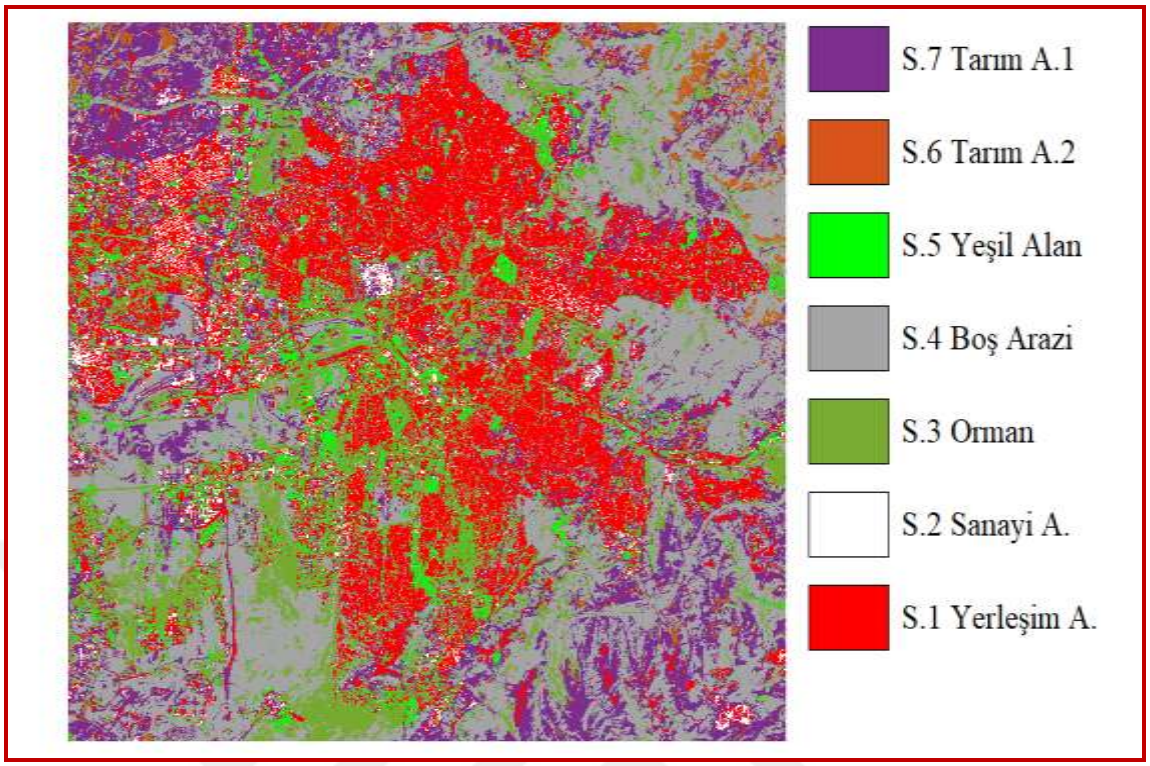
- Sınıflandırma Tipi : Lineer
- Gama Katsayısı : 0

şeklindedir.

Lineer diskriminant ile anlamlı çıktı ürün verecek şekilde bantlar kullanılarak sınıflandırılmış görüntülerin çıktıları Şekil 3.19. ve Şekil 3.20. 'deki gibidir.



Şekil 3.19. Lineer Diskriminant ile Sınıflandırılmış Landsat 8 (Bant:5,1,7,6) Görüntüsü



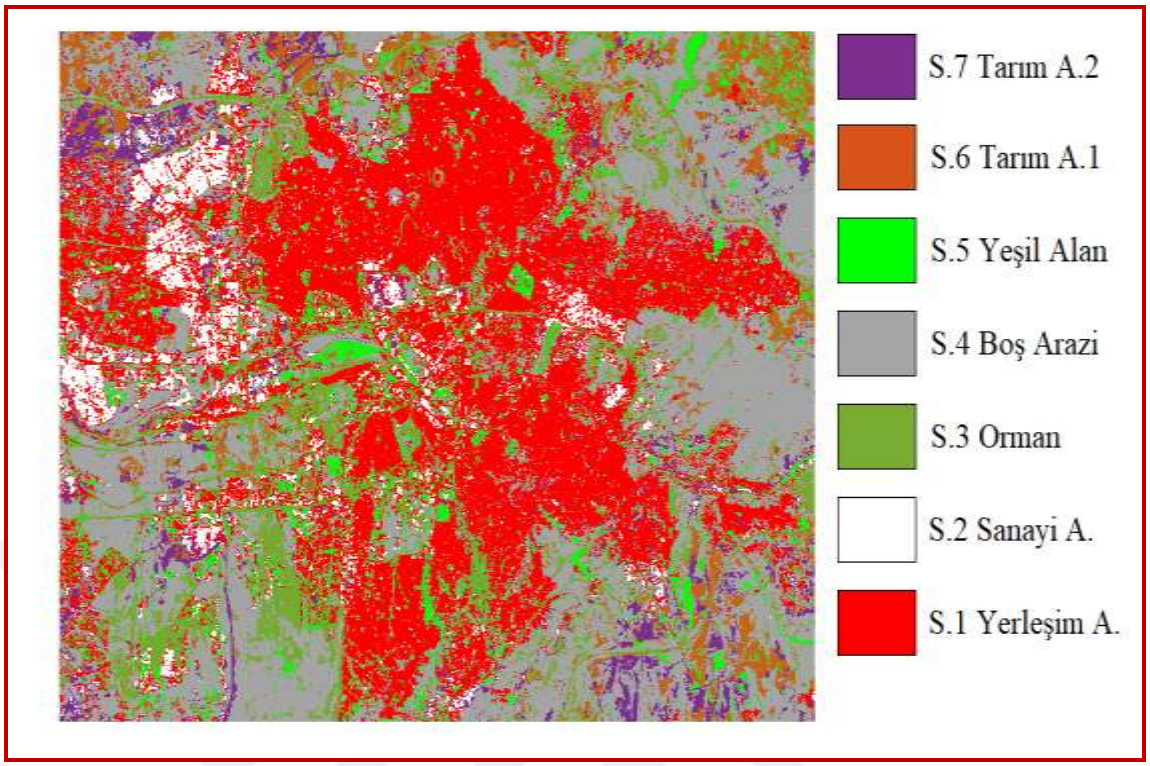
Şekil 3.20. Lineer Diskriminant ile Sınıflandırılmış Sentinel 2 (Bant:8,12,7,5,11) Görüntüsü

Tablo 3.9. 'da lineer diskriminant ile yapılan sınıflandırma sonucu doğruluk değerleri verilmiştir.

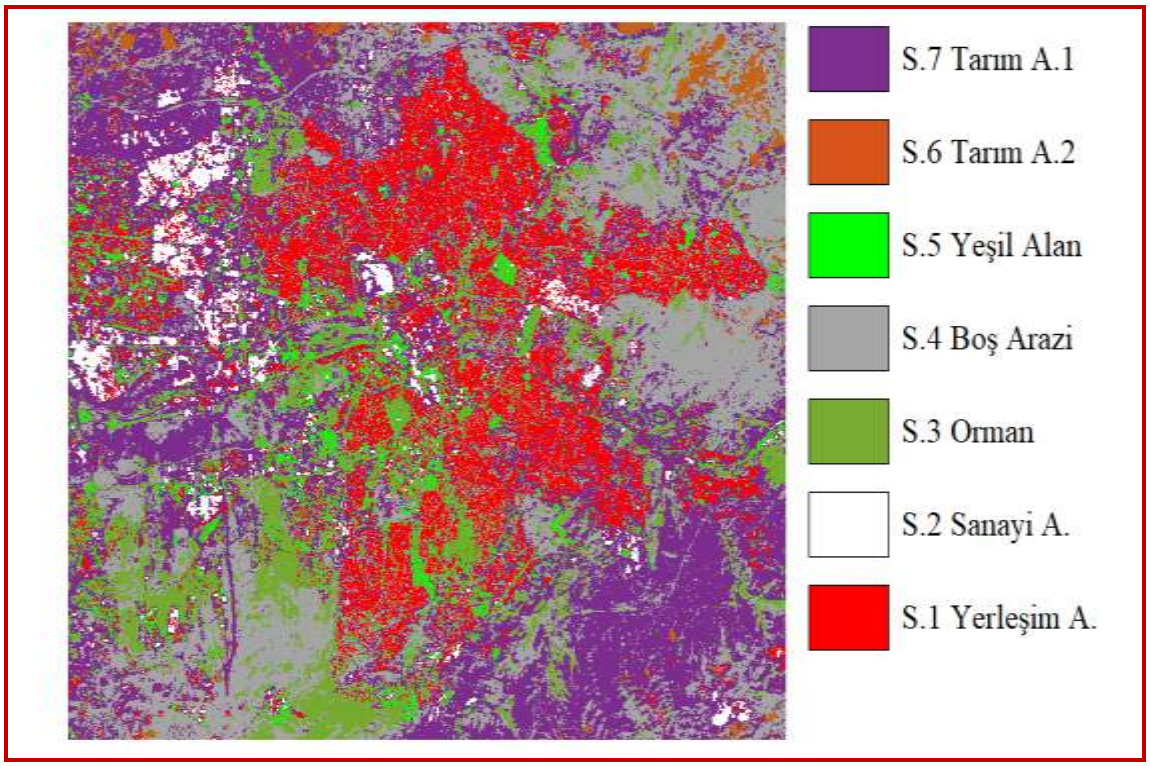
Tablo 3.9. Lineer Diskriminant ile Sınıflandırılmış Landsat 8 (Bant:5,1,7,6) ve Sentinel 2 (Bant:8,12,7,5,11) Görüntüleri Doğruluk Değerleri

	Eğitim Doğruluğu	**Test Doğruluğu**	**Duyarlılık (Sensitivity)**	**Özgüllük (Specificity)**	**Hassaslık (Presicion)**	**Hatırlama (Recall)**	**F-measure**	**G-mean**
Landsat 8	0,8657	0,8665	0,9878	0,8303	0,6346	0,9878	0,7727	0,9056
Sentinel 2	0,8673	0,8584	0,9464	0,8474	0,4380	0,9464	0,5989	0,8955

Sınıflandırmaya en fazla etki eden bantların çıktı değerlerini analiz edebilecek seviyede sonuç istenmesiyle birlikte yapılmış sınıflandırma uygulamasında Sentinel 2 görüntüsü için daha yüksek doğruluk değerleri elde edilmesine rağmen Landsat 8 görüntüsünün duyarlılık ve hassasiyet değerlerinin daha yüksek olduğu belirlenmiştir. Bu durumun görüntüler arasındaki mekansal çözünürlükten kaynaklandığı düşünülmektedir. Şekil 3.21. ve Şekil 3.22. 'de tüm bantlar ile yapılmış sınıflandırma sonuçları gösterilmiştir.



Şekil 3.21. Lineer Diskriminant ile Sınıflandırılmış Landsat 8 Görüntüsü



Şekil 3.22. Lineer Diskriminant ile Sınıflandırılmış Sentinel 2 Görüntüsü

Yöntemin doğruluk değerleri Tablo 3.10. 'daki gibidir.

Tablo 3.10. Lineer Diskriminant ile Sınıflandırma Doğruluk Değerleri

Tüm Bantlar	Eğitim Doğruluğu	Test Doğruluğu	Duyarlılık (Sensitivity)	Özgüllük (Specificity)	Hassaslık (Presicion)	Hatırlama (Recall)	F-measure	G-mean
Landsat 8	0,9175	0,9129	0,9840	0,8929	0,7207	0,9840	0,8320	0,9373
Sentinel 2	0,9382	0,9342	0,9375	0,9338	0,6402	0,9375	0,7609	0,9356

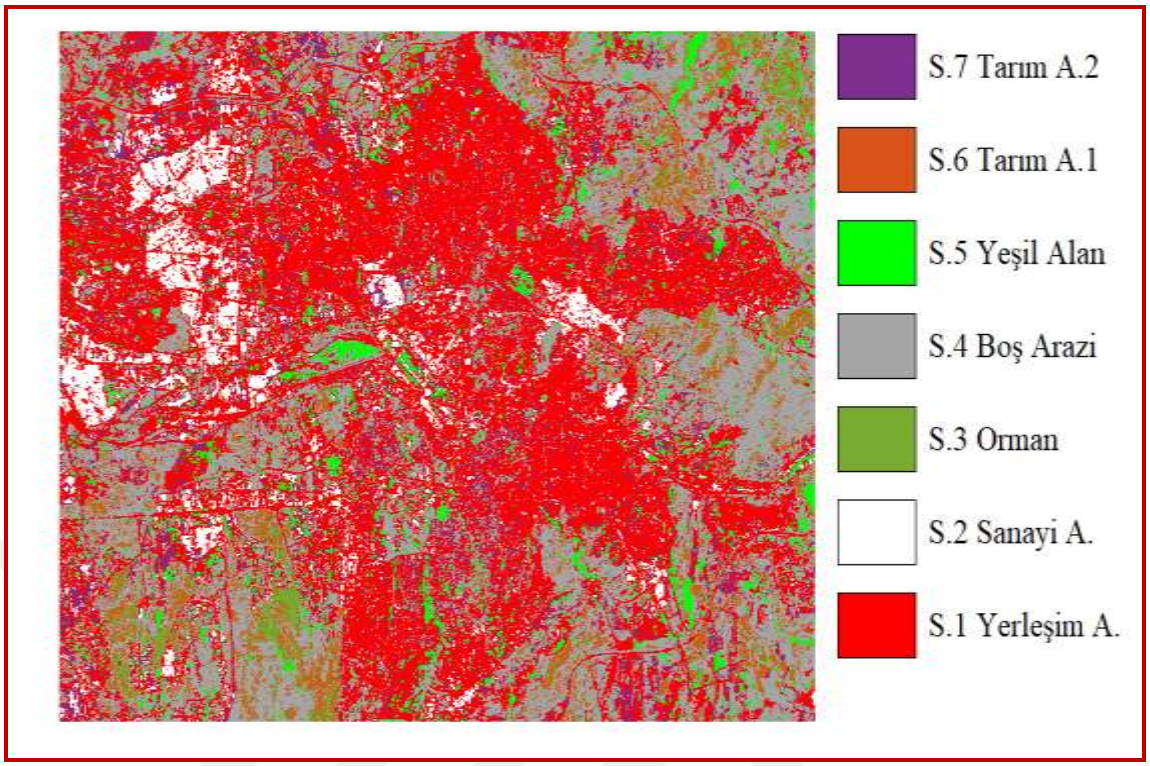
Lineer diskriminant yöntemi ile yapılan sınıflandırmada eğitim ve test verilerinin yüksek duyarlılıkta işlendiği ve bu durumun çıktı ürüne doğrudan etki ettiği görülmüştür.

3.8. K-En Yakın Komşu ile Sınıflandırma İşlemleri

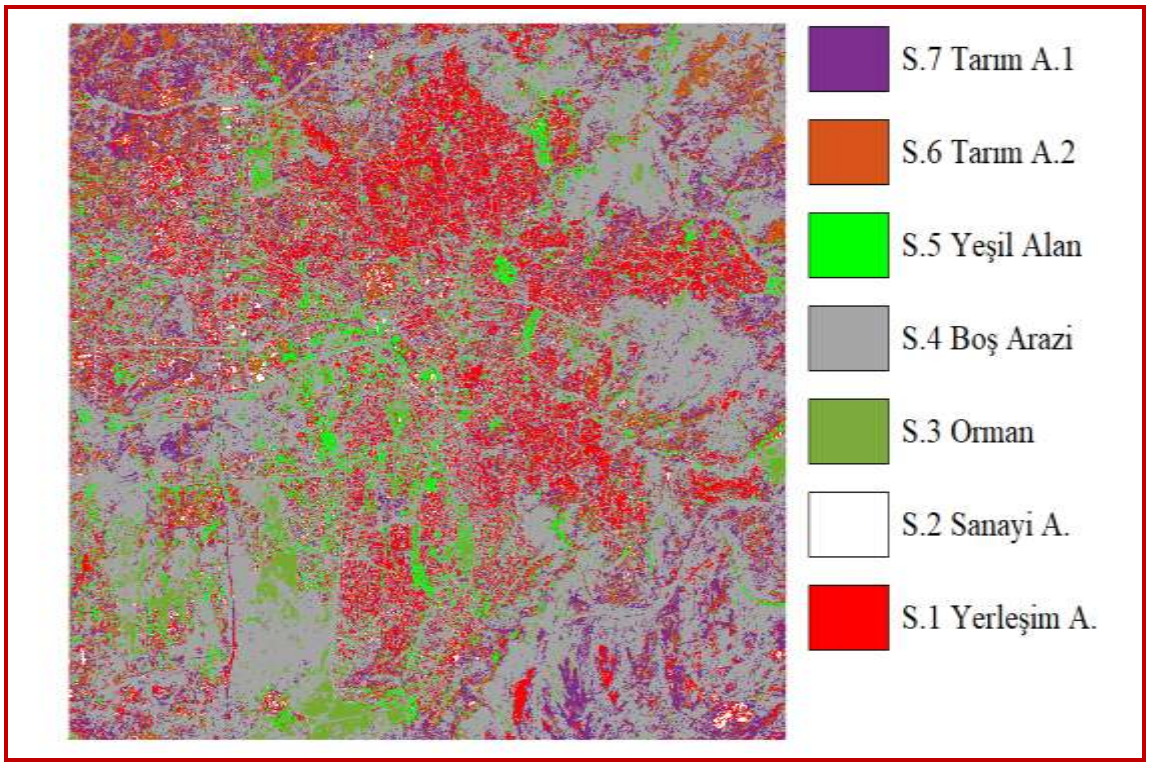
K-en yakın komşu algoritması denetimli öğrenmeyi hedef alan sınıflandırma ve regresyon işlemleri için kullanılan bir algoritmadır. Bu yöntemin mantığında en yakın noktalar aranır ve *k;* bilinmeyen noktaların en yakın komşularının sayısını temsil eder. KNN ile yapılan sınıflandırmada aşağıdaki parametreler kullanılmıştır.

- Mesafe Kriteri: Öklid Mesafesi
- NumNeighbors: 5

Yöntemin her iki görüntü tipi içinde kullandığı 2 adet bant sayesinde sınıflandırdığı görüntüler için çıktı ürünler ise Şekil 3.24. ve Şekil 3.25. 'deki gibidir.



Şekil 3.23. KNN ile Sınıflandırılmış Landsat 8 (Bant:5,1) Görüntüsü



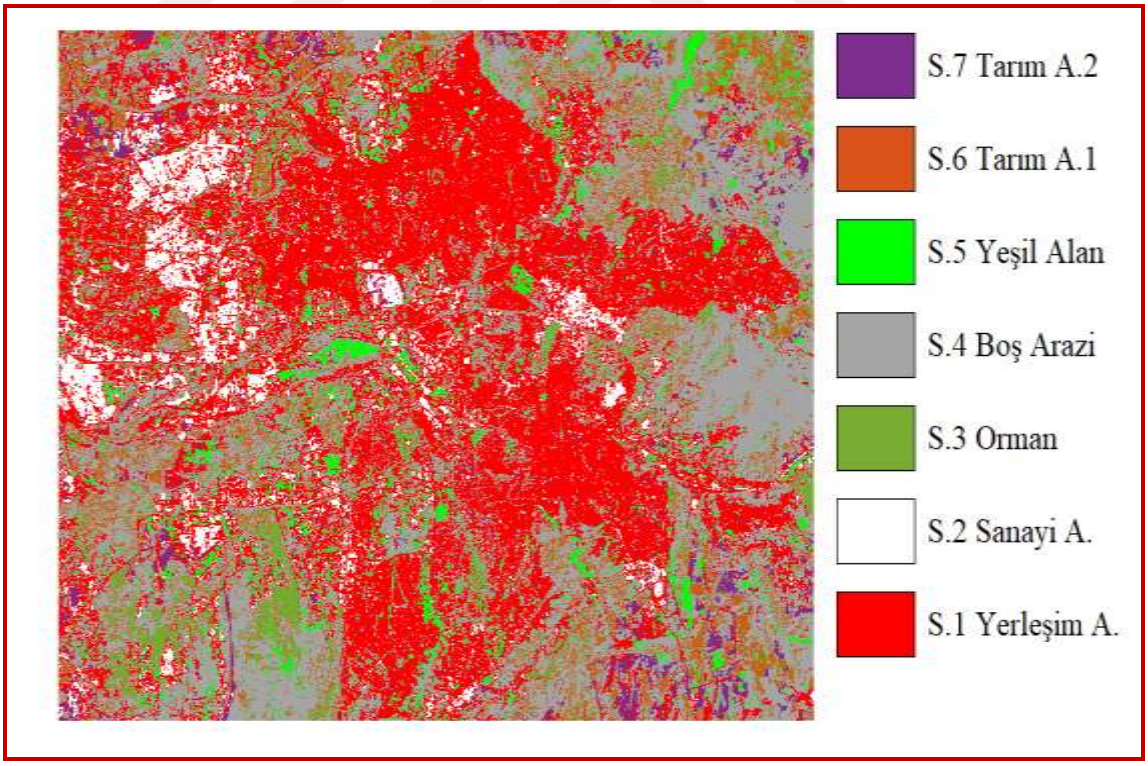
Şekil 3.24. KNN ile Sınıflandırılmış Sentinel 2 (Bant:8,12) Görüntüsü

Tablo 3.11. 'de KNN ile sınıflandırılmış görüntülerin doğruluk değerleri yer almaktadır.

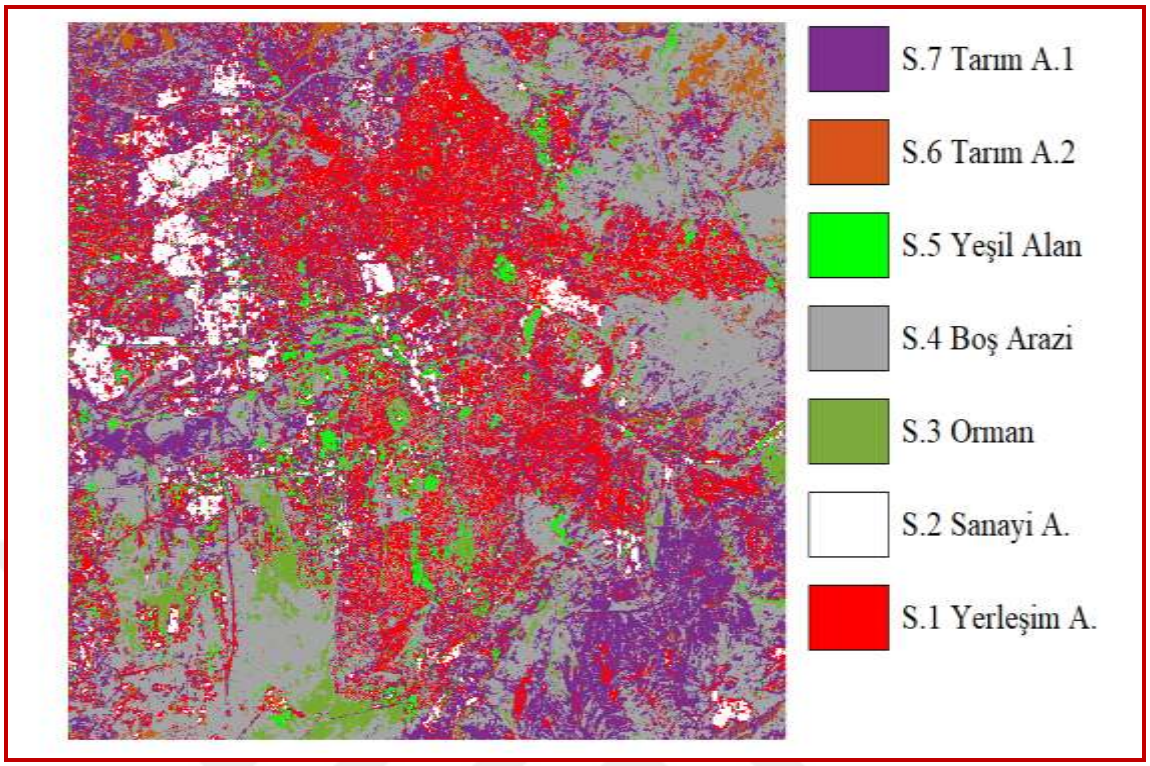
Tablo 3.11. KNN ile Sınıflandırılmış Landsat 8 (Bant:5,1) ve Sentinel 2 (Bant:8,12) Görüntüleri Doğruluk Değerleri

	Eğitim Doğruluğu	**Test Doğruluğu**	**Duyarlılık (Sensitivity)**	**Özgüllük (Specificity)**	**Hassaslık (Presicion)**	**Hatırlama (Recall)**	**F-measure**	**G-mean**
Landsat 8	0,8576	0,8067	0,8746	0,7865	0,5500	0,8746	0,6753	0,8294
Sentinel 2	0,8744	0,8375	0,7054	0,8541	0,3780	0,7054	0,4922	0,7762

KNN yöntemi az bant sayısına rağmen sonuçlarıyla bazı yöntemlerin önüne geçmiştir. Her iki görüntü tipi içinde yüksek doğruluk değerleri sağlamıştır. Şekil 3.25. ve Şekil 3.26. 'da tüm bantlar ile yapılan sınıflandırma çıktı ürünleri gösterilmiştir.



Şekil 3.25. KNN ile Sınıflandırılmış Landsat 8 Görüntüsü



Şekil 3.26. KNN ile Sınıflandırılmış Sentinel 2 Görüntüsü

KNN yöntemi ile tüm bantlar kullanılarak yapılan sınıflandırma sonucu elde edilen doğruluk değerleri Tablo 3.12.'de verilmiştir.

Tablo 3.12. KNN ile Sınıflandırma Doğruluk Değerleri

Tüm Bantlar	Eğitim Doğruluğu	Test Doğruluğu	Duyarlılık (Sensitivity)	Özgüllük (Specificity)	Hassaslık (Presicion)	Hatırlama (Recall)	F-measure	G-mean
Landsat 8	0,9561	0,9262	0,9583	0,9172	0,7647	0,9583	0,8506	0,9375
Sentinel 2	0,9786	0,9711	0,9286	0,9764	0,8320	0,9286	0,8776	0,9522

Sentinel 2 görüntüsün sınıflandırılması sonucu Landsat 8 görüntüsüne göre daha yüksek doğruluk değerlerine ulaşmıştır. Sınıflandırmaya etkisi fazla olan banttan sonra eklenen her bir bant için doğruluk değerleri gelişmiştir. Bantlardaki veriyi işlemekte yöntemin başarısı gözlemlenmiştir.

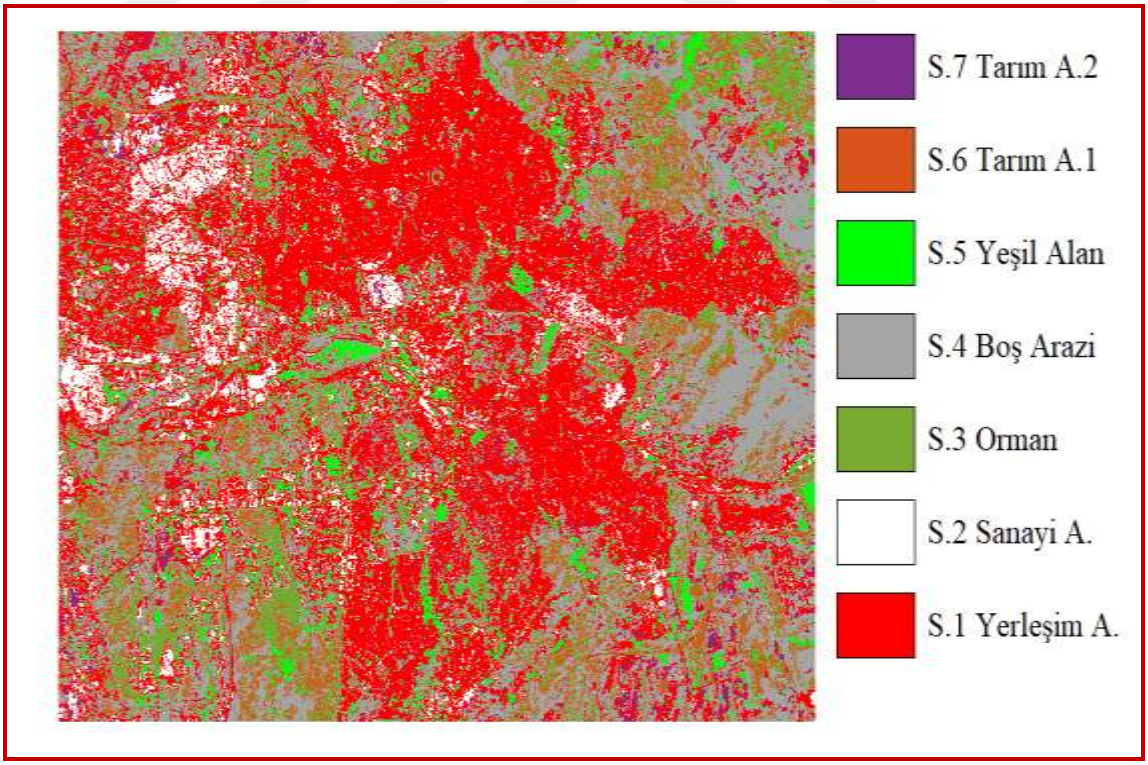
3.9. Altuzay En Yakın Komşu ile Sınıflandırma İşlemleri

Altuzay *k* en yakın komşu yönteminin amacı, belirli bir sorgunun yakın komşu aramasıyla ilgili daha önce bilinmeyen alt uzaylarını algılaması ve her bir ilgili altuzayın yakın komşu sınıfını belirlemesidir. Bu yöntemde sınıflandırmaya etki eden parametreler;

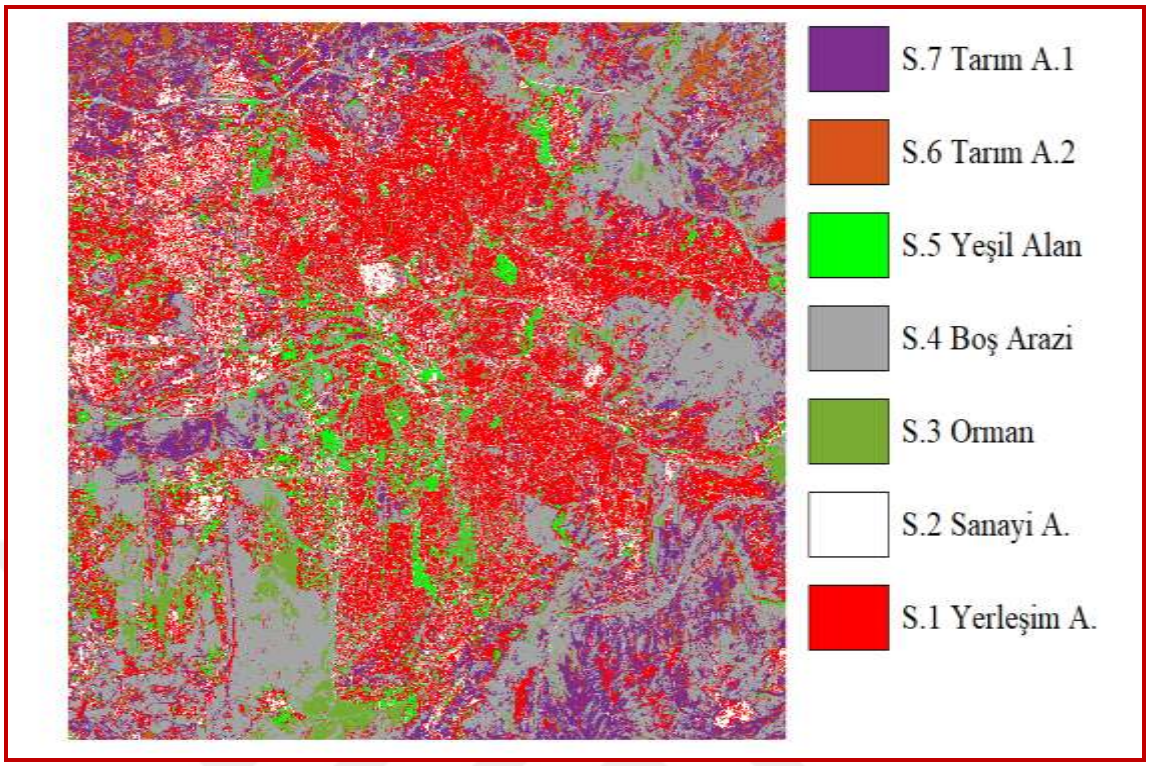
- Öğrenme Döngü Sayısı : 30
- Altuzay Sayısı : Öznitelik / 2

şeklindedir.

Seçilen bantların bu yöntemle sınıflandırılması ve görüntülerin çıktısı Şekil 3.27.'de ve Şekil 3.28.'de verilmiştir.



Şekil 3.27. Altuzay KNN ile Sınıflandırılmış Landsat 8 (Bant:5,1,7,6) Görüntüsü



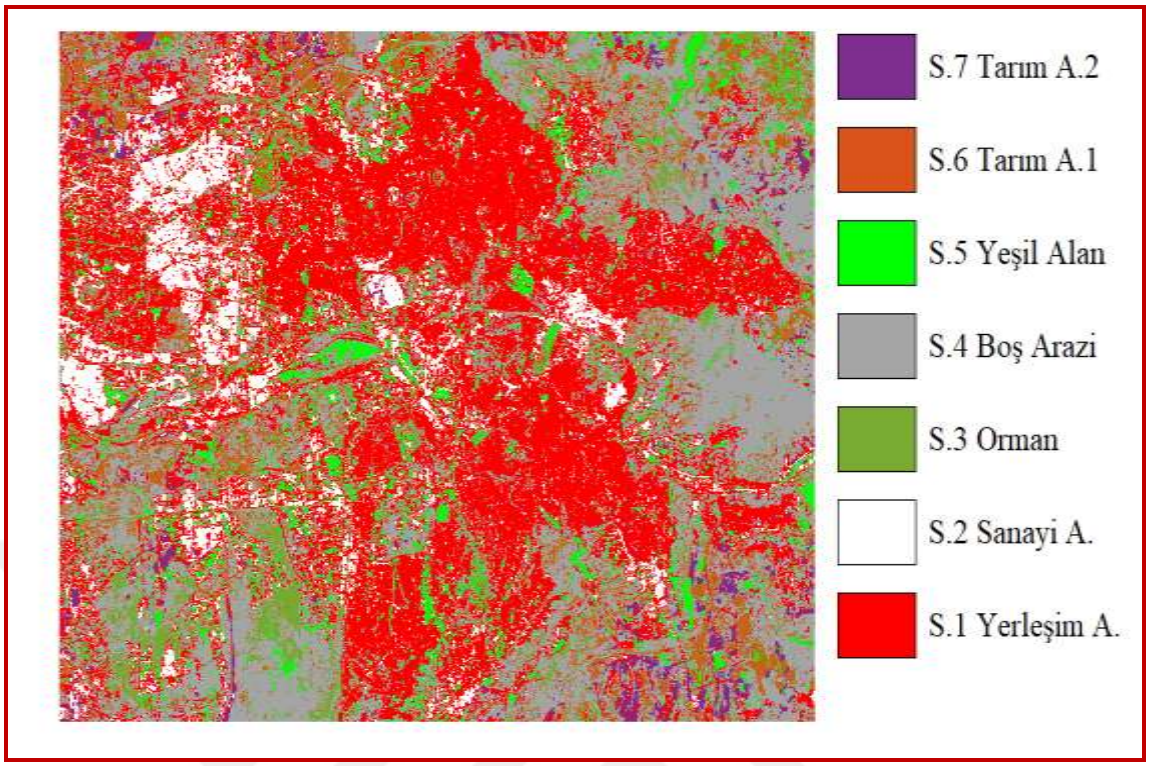
Şekil 3.28. Altuzay KNN ile Sınıflandırılmış Sentinel 2 (Bant:8,12,7,5,11) Görüntüsü

Tablo 3.13. 'de altuzay KNN ile belirli bantlar kullanılarak yapılmış sınıflandırmaların doğruluk değerleri verilmiştir.

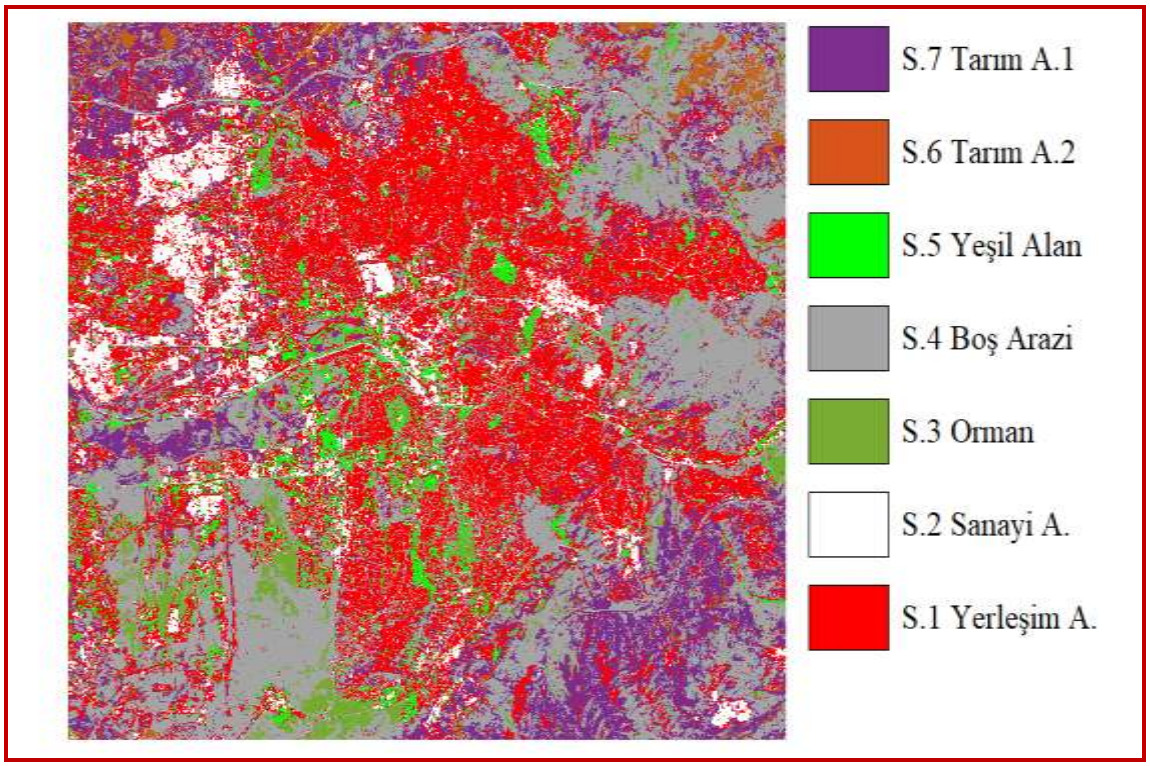
Tablo 3.13. Altuzay KNN ile Sınıflandırılmış Landsat 8 (Bant:5,1,7,6) ve Sentinel 2 (Bant:8,12,7,5,11) Görüntüleri Doğruluk Değerleri

	Eğitim Doğruluğu	Test Doğruluğu	Duyarlılık (Sensitivity)	Özgüllük (Specificity)	Hassaslık (Presicion)	Hatırlama (Recall)	F-measure	G-mean
Landsat 8	1,0000	0,8595	0,9052	0,8458	0,6366	0,9052	0,7475	0,8750
Sentinel 2	1,0000	0,9501	0,9464	0,9506	0,7067	0,9464	0,8092	0,9485

Sınıflandırma işlemine etki eden bantların yüksek doğruluk verdiği aşama seçilerek yapılan uygulamada yüksek doğruluk değerleri elde edildiği görülmüştür. Şekil 3.29. ve Şekil 3.30. 'da tüm bantlar ile yapılan sınıflandırma sonucu gösterilmiştir.



Şekil 3.29. Altuzay KNN ile Sınıflandırılmış Landsat 8 Görüntüsü



Şekil 3.30. Altuzay KNN ile Sınıflandırılmış Sentinel 2 Görüntüsü

Sınıflandırılmış görüntülerin doğruluk değerleri Tablo 3.14. 'deki gibidir.

Tablo 3.14. Altuzay KNN ile Sınıflandırma Doğruluk Değerleri

Tüm Bantlar	Eğitim Doğruluğu	Test Doğruluğu	Duyarlılık (Sensitivity)	Özgüllük (Specificity)	Hassaslık (Presicion)	Hatırlama (Recall)	F-measure	G-mean
Landsat 8	1,0000	0,9599	0,9817	0,9535	0,8629	0,9817	0,9185	0,9675
Sentinel 2	1,0000	0,9776	0,9554	0,9804	0,8594	0,9554	0,9049	0,9678

Altuzay KNN ile yapılan sınıflandırmada Sentinel 2 görüntüsü için daha yüksek sonuçlar elde edilmiştir. Görüntüler için duyarlılık değerlerinin doğruluğa etki ettiği belirlenmiştir.

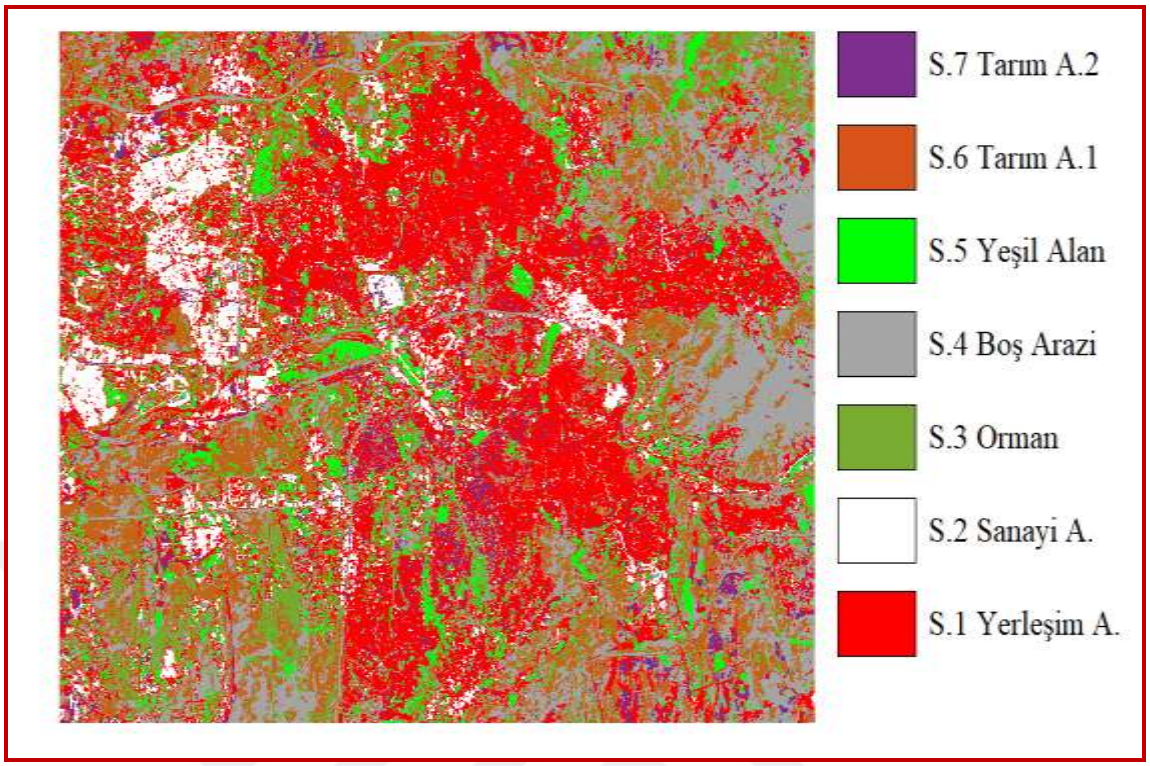
3.10. Derin Öğrenme ile Sınıflandırma İşlemleri

Derin öğrenme yönteminde sınıflandırma aşamasının gerçekleştirildiği bölüm softmax bölümüdür. Sınıflandırma sonucunda doğruluk oranlarının çok yüksek olmasından dolayı literatürde yapılan son çalışmalarda derin öğrenme tekniklerinin kullanımı oldukça artmıştır. Sınıflandırma işleminde oto kodlayıcı kullanılmıştır. Etki eden parametreler aşağıda gösterilmiştir.

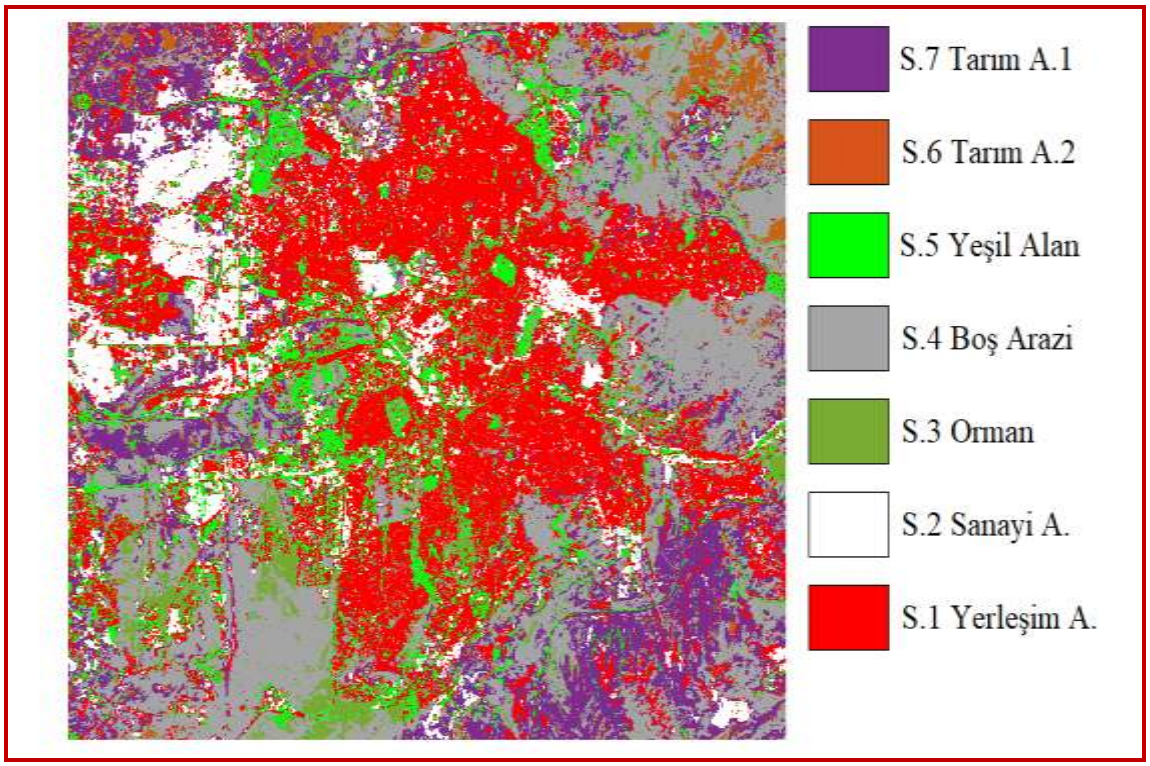
- Kodlayıcı Fonksiyon : Purelin
- Kod Çözücü Fonksiyon : Logsig
- Epok : 2000

Oto kodlayıcı ile yapılan sınıflandırma sonucu oluşan görüntüler aşağıdaki gibidir.

Sınıflandırmaya etki eden bantların sınıflandırma işlemlerine teker teker eklenerek verdiği anlamlı verilerden elde edilen anlamlı ilk görüntüler Şekil 3.31. ve Şekil 3.32. 'deki gibidir.



Şekil 3.31. Derin Öğrenme ile Sınıflandırılmış Landsat 8 (Bant:5,1,7) Görüntüsü



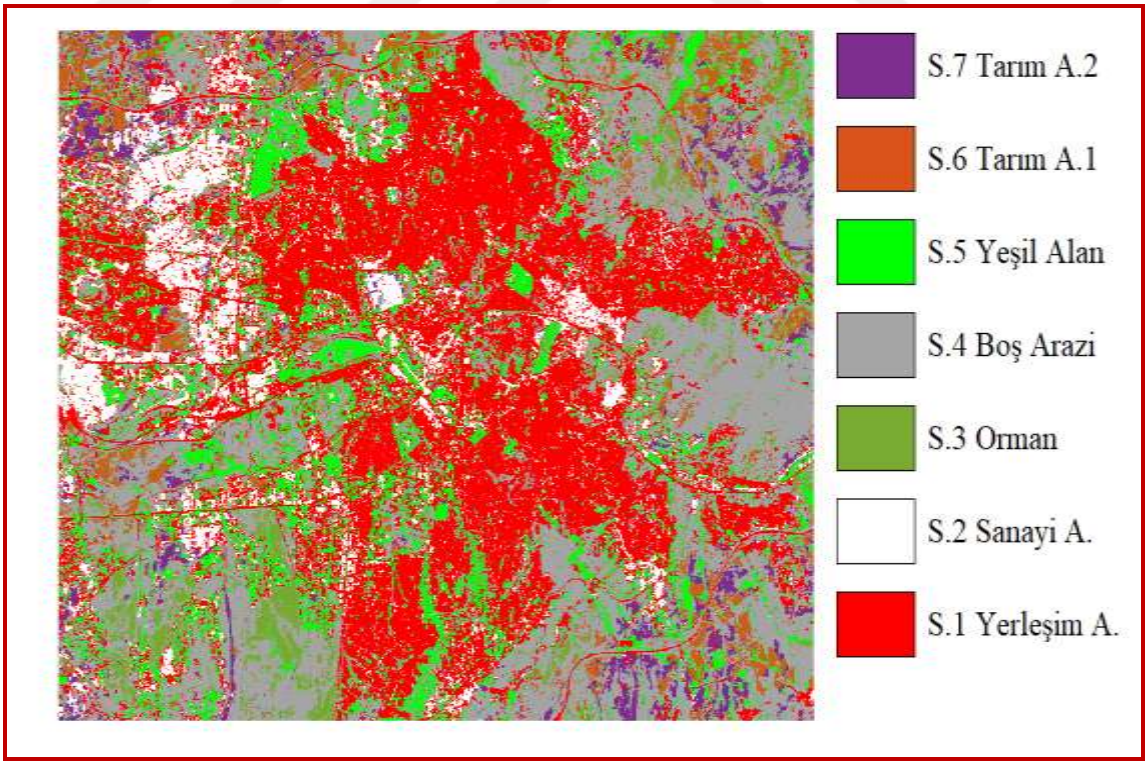
Şekil 3.32. Derin Öğrenme ile Sınıflandırılmış Sentinel 2 (Bant:8,12,7,5,11) Görüntüsü

Doğruluk değerleri Tablo 3.15. 'deki gibi elde edilmiştir.

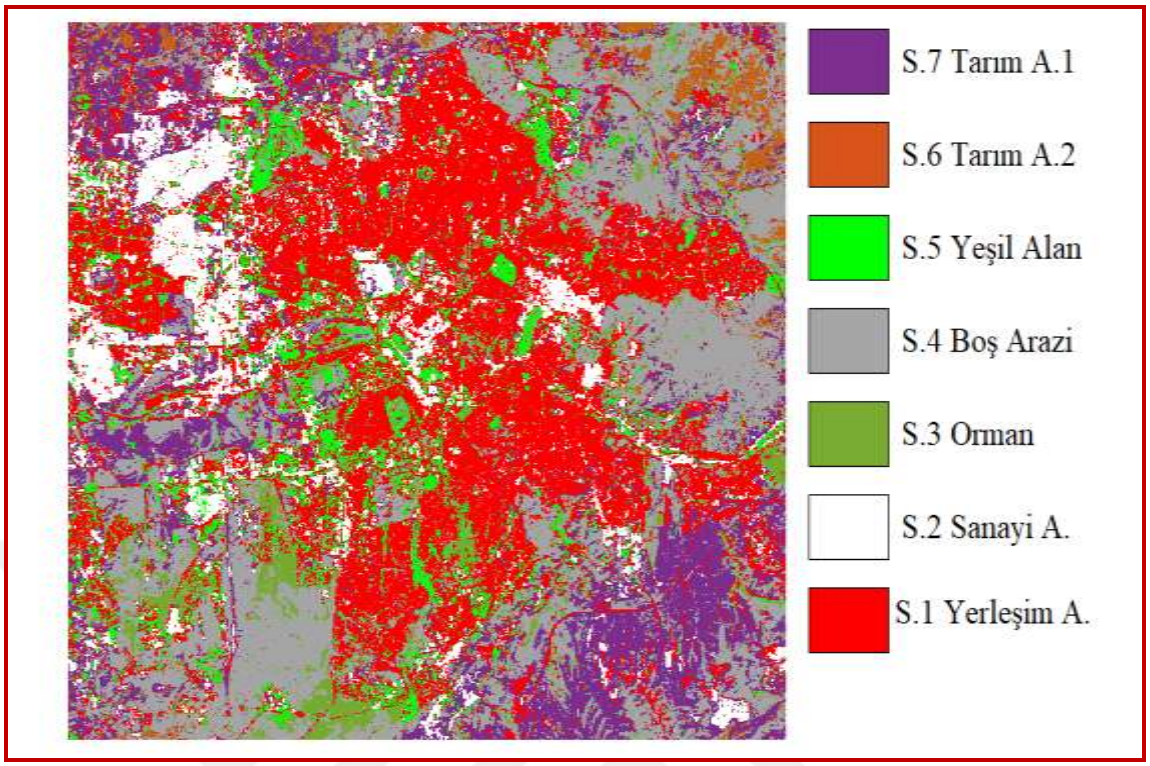
Tablo 3.15. Derin Öğrenme ile Sınıflandırılmış Landsat 8 (Bant:5,1,7) ve Sentinel 2 (Bant:8,12,7,5,11) Görüntüleri Doğruluk Değerleri

	Eğitim Doğruluğu	Test Doğruluğu	Duyarlılık (Sensitivity)	Özgüllük (Specificity)	Hassaslık (Presicion)	Hatırlama (Recall)	F-measure	G-mean
Landsat 8	1,0000	0,8833	0,8746	0,8859	0,6959	0,8746	0,7751	0,8803
Sentinel 2	1,0000	0,9342	0,9464	0,9327	0,6386	0,9464	0,7626	0,9395

Kullanılan bant sayısının az olmasına rağmen verdiği doğruluk değeriyle derin öğrenme yönteminin yeni nesil uygulamalarda önemli bir yeri olması beklenmektedir. Tüm bantlar kullanılarak yapılan sınıflandırma sonucu elde edilen görüntüler Şekil 3.33. ve Şekil 3.34. 'deki gibi olmuştur.



Şekil 3.33. Derin Öğrenme ile Sınıflandırılmış Landsat 8 Görüntüsü



Şekil 3.34. Derin Öğrenme ile Sınıflandırılmış Sentinel 2 Görüntüsü

Tüm bantlar kullanılarak gerçekleştirilen sınıflandırma sonucu oluşan doğruluk değerleri Tablo 3.16. 'daki gibidir.

Tablo 3.16. Derin Öğrenme ile Sınıflandırma Doğruluk Değerleri

Tüm Bantlar	Eğitim Doğruluğu	Test Doğruluğu	Duyarlılık (Sensitivity)	Özgüllük (Specificity)	Hassaslık (Presicion)	Hatırlama (Recall)	F-measure	G-mean
Landsat 8	1,0000	0,9698	0,9847	0,9653	0,8944	0,9847	0,9374	0,9750
Sentinel 2	1,0000	0,9811	0,9777	0,9815	0,8690	0,9777	0,9202	0,9796

Epok sayısına bağlı olarak yaptığı işlemlerin kalitesini artıran bu yöntem yüksek duyarlılık oranına sahip olarak sınıflandırma sonucu elde edilen çıktı ürünün kalitesini artırmıştır.

3.11. Sınıflandırma Sonucu Elde Edilen Değerlerin Yorumu

Landsat 8 görüntüsüne ait bantlarının yöntemlere göre farklı kombinleri ile yapılan sınıflandırma sonucunda elde edilen değerler ve her bir yöntem için kullanılan bantlar aşağıdaki gibidir.

- Destek Vektör Makineleri: Bant 5, 1, 7
- Karar Ağaçları: Bant 5, 1, 7
- Torbalama Karar Ağaçları: Bant 5,1
- Lineer Diskriminant: Bant 5, 1, 7, 6
- K-En Yakın Komşu: Bant 5, 1
- Altuzay K-En Yakın Komşu: Bant 5, 1, 7, 6
- Derin Öğrenme: Bant 5, 1, 7

Tablo 3.17. 'de Landsat 8 görüntüsü için yöntemlerin farklı bantlar için verdiği doğruluk değerleri gösterilmiştir.

Tablo 3.17. Landsat 8 Görüntüsünün Yöntemlere Göre Farklı Bant Kombinasyonlarının Doğruluk Değerleri

	Eğitim Doğruluğu	**Test Doğruluğu**	**Duyarlılık (Sensitivity)**	**Özgüllük (Specificity)**	**Hassaslık (Presicion)**	**Hatırlama (Recall)**	**F-measure**	**G-mean**
SVM	0,8630 0,8384	0,8644 0,8292	0,9480 0,9664	0,8394 0,7883	0,6379 0,5766	0,9480 0,9664	0,7626 0,7223	0,8921 0,8728
Karar Ağaçları	0,8654	0,8391	0,7982	0,8513	0,6156	0,7982	0,6951	0,8243
T. Karar Ağaçları	0,8685	0,8124	0,8899	0,7892	0,5575	0,8899	0,6855	0,8381
Lineer D.	0,8657	0,8665	0,9878	0,8303	0,6346	0,9878	0,7727	0,9056
KNN	0,8576	0,8067	0,8746	0,7865	0,5500	0,8746	0,6753	0,8294
Altuzay KNN	1,0000	0,8595	0,9052	0,8458	0,6366	0,9052	0,7475	0,8750
Derin Öğrenme	1,0000	0,8833	0,8746	0,8859	0,6959	0,8746	0,7751	0,8803

Sınıflandırmaya en fazla katkıyı sağlayan bantlar kullanılarak yapılan sınıflandırma çalışmasında Landsat 8 görüntüsü için en yüksek doğruluk değerini derin öğrenme yöntemi, altuzay KNN yöntemi, destek vektör makineleri(kernel:gaussian) ve lineer

diskriminant yöntemi vermiştir. Karar ağaçları yöntemi ve KNN yöntemi ise kullanılan bant sayısından dolayı tatmin edici sonuçlarına rağmen en düşük değerleri veren yöntem olarak belirlenmiştir.

Sentinel 2 görüntüsünün farklı bantları bir araya getirilerek yöntemlerin verdiği cevaplara göre yapılan sınıflandırma sonrasında elde edilen doğruluk değerleri ve her bir yöntem için kullanılan bantlar aşağıda verilmiştir.

- Destek Vektör Makineleri: Bant 8, 12, 7, 5, 11
- Karar Ağaçları: Bant 8, 12, 7, 5, 11
- Torbalama Karar Ağaçları: Bant 8, 12
- Lineer Diskriminant: Bant 8, 12, 7, 5, 11
- K-En Yakın Komşu: Bant 8, 12
- Altuzay K-En Yakın Komşu: Bant 8, 12, 7, 5, 11
- Derin Öğrenme: Bant 8, 12, 7, 5, 11

Tablo 3.18. 'de Sentinel 2 görüntüsü için yöntemlerin farklı bantlar için verdiği doğruluk değerleri gösterilmiştir.

Tablo 3.18. Sentinel 2 Görüntüsünün Yöntemlere Göre Farklı Bant Kombinlerinin Doğruluk Değerleri

	Eğitim Doğruluğu	**Test Doğruluğu**	**Duyarlılık (Sensitivity)**	**Özgüllük (Specificity)**	**Hassaslık (Presicion)**	**Hatırlama (Recall)**	**F-measure**	**G-mean**
SVM	0,9111 0,8908	0,9098 0,8913	0,9330 0,9152	0,9068 0,8883	0,5573 0,5074	0,9330 0,9152	0,6978 0,6529	0,9198 0,9017
Karar Ağaçları	0,8483	0,8340	0,6429	0,8580	0,3627	0,6429	0,4638	0,7427
T. Karar Ağaçları	0,8752	0,8425	0,6830	0,8625	0,3844	0,6830	0,4920	0,7675
Lineer D.	0,8673	0,8584	0,9464	0,8474	0,4380	0,9464	0,5989	0,8955
KNN	0,8744	0,8375	0,7054	0,8541	0,3780	0,7054	0,4922	0,7762
Altuzay KNN	1,0000	0,9501	0,9464	0,9506	0,7067	0,9464	0,8092	0,9485
Derin Öğrenme	1,0000	0,9342	0,9464	0,9327	0,6386	0,9464	0,7626	0,9395

Belirlenen bantlar kullanılarak elde edilen sınıflandırma sonuçlarına bakıldığında Sentinel 2 görüntüsü için en yüksek doğruluk veren yöntem altuzay KNN yöntemi ve derin öğrenme yöntemidir. Destek vektör makineleri kullanılan bant sayısına rağmen oldukça yüksek değerler vermektedir. Karar ağaçları yöntemi ise Landsat 8 görüntüsünde olduğu gibi en düşük doğruluk değerini veren yöntemlerden biridir.

Tablo 3.19.'da Landsat 8 görüntüsü için farklı yöntemlerle tüm bantlar kullanılarak yapılan sınıflandırma sonucu doğruluk değerleri yer almaktadır.

Tablo 3.19. Landsat 8 Görüntüsü için Yöntemlere Göre Doğruluk Değerleri

Tüm Bantlar	**Eğitim Doğruluğu**	**Test Doğruluğu**	**Duyarlılık (Sensitivity)**	**Özgüllük (Specificity)**	**Hassaslık (Presicion)**	**Hatırlama (Recall)**	**F-measure**	**G-mean**
SVM	0,9567 0,9244	0,9529 0,9227	0,9840 0,9872	0,9442 0,9046	0,8320 0,7440	0,9840 0,9872	0,9016 0,8485	0,9639 0,9450
Karar Ağaçları	0,9025	0,8856	0,9006	0,8776	0,6739	0,9006	0,7709	0,8890
T. Karar Ağaçları	0,9895	0,9592	0,9840	0,9523	0,8528	0,9840	0,9137	0,9680
Lineer D.	0,9175	0,9129	0,9840	0,8929	0,7207	0,9840	0,8320	0,9373
KNN	0,9561	0,9262	0,9583	0,9172	0,7647	0,9583	0,8506	0,9375
Altuzay KNN	1,0000	0,9599	0,9817	0,9535	0,8629	0,9817	0,9185	0,9675
Derin Öğrenme	1,0000	0,9698	0,9847	0,9653	0,8944	0,9847	0,9374	0,9750

Bu tabloya göre Landsat 8 görüntüsü için en yüksek doğruluk değerini veren yöntemler derin öğrenme ve Altuzay KNN yöntemleri olurken, en düşük doğruluk değerini veren yöntem ise karar ağaçları yöntemidir.

Tablo 3.20.' de ise Sentinel 2 uydu görüntüsü için tüm bantlar dahil edilerek kullanılan tekniklerin verdiği doğruluk oranları sergilenmiştir.

Tablo 3.20. Sentinel 2 Görüntüsü için Yöntemlere Göre Doğruluk Değerleri

Tüm Bantlar	Eğitim Doğruluğu	Test Doğruluğu	Duyarlılık (Sensitivity)	Özgüllük (Specificity)	Hassaslık (Presicion)	Hatırlama (Recall)	F-measure	G-mean
SVM	0,9741 0,9632	0,9681 0,9591	0,9643 0,9420	0,9686 0,9613	0,7941 0,7536	0,9643 0,9420	0,8710 0,8373	0,9664 0,9516
Karar Ağaçları	0,8650	0,8674	0,6607	0,8934	0,4379	0,6607	0,5267	0,7683
T. Karar Ağaçları	0,9840	0,9636	0,9509	0,9652	0,7745	0,9509	0,8537	0,9580
Lineer D.	0,9382	0,9342	0,9375	0,9338	0,6402	0,9375	0,7609	0,9356
KNN	0,9786	0,9711	0,9286	0,9764	0,8320	0,9286	0,8776	0,9522
Altuzay KNN	1,0000	0,9776	0,9554	0,9804	0,8594	0,9554	0,9049	0,9678
Derin Öğrenme	1,0000	0,9811	0,9777	0,9815	0,8690	0,9777	0,9202	0,9796

Tablo 3.20.' ye göre Sentinel 2 görüntüsü için doğruluk değeri en yüksek yöntemler derin öğrenme ve Altuzay KNN yöntemleridir. Doğruluk değeri en düşük olan yöntem Karar Ağaçları yöntemidir.

4. BÖLÜM

TARTIŞMA-SONUÇ ve ÖNERİLER

4.1. Tartışma ve Sonuçlar

Geliştirilen teknolojiler doğrultusunda uzaktan algılama çalışmaları gibi birçok bilim dalına hizmet eden ve yapısı gereği geniş bölgeler için bilgi içeren görüntülerin temini günümüzde yüksek çözünürlükte gerçekleştirilebilinmektedir. Bu imkanlar doğrultusunda elde edilen görüntüler, amaca yönelik uygulamalar için altlık oluşturmakta ve çıktı ürünün kalite değerini artırmaktadır.

Görüntü üzerinden bilgi çıkarımı sınıflandırma çalışmalarıyla mümkün olmaktadır. Çıktı ürünün doğruluk değerlerinin ise sınıflandırma işlemlerinde kullanılacak yöntemler ve bu yöntemlerin içerdiği parametrelerle doğrudan ilişkili olduğu gözlemlenmiştir. Bu sebeplerden dolayı yöntem seçiminin sınıflandırma sonrası yapılacak çıkarımlar için önemli bir konumda olduğu görülmüştür.

Tezimizde makine öğrenmesi teknikleri ile Landsat 8 ve Sentinel 2 uydu görüntülerinin sınıflandırılması hedeflenmiş ve gerçekleştirilmiştir. Her bir teknik ile doğruluk değerleri ve sınıflandırılmış görüntü elde edilmiştir.

Yapılan sınıflandırmalar sonucunda gözlemlenen detaylar ise şu şekilde sıralanabilir;

- Komşuluk bileşen analizi ile sınıflandırmaya etki eden bantlar belirlendiğinde Landsat 8 görüntüsü için 5.bandın, Sentinel 2 görüntüsü için 8.bandın sınıflandırmaya en fazla etki eden bant olduğu belirlenmiştir. Her iki bandında yakın kızılötesi bölge için olduğu bilindiğinden sınıflandırma işlemi için bu bilgilerin önemli olduğu ortaya koyulmuştur.
- Sınıflandırmaya en çok etki eden bandın tek başına sınıflandırılmasıyla elde edilen doğruluk değerleri düşük olmakla beraber diğer bantların içerdiği

bilgilere ihtiyaç duyduğu belirlenmiştir ve diğer bantlar eklendiğinde kademeli olarak değerlerin arttığı görülmüştür.

- Eğitim ve test verilerinin kalitesine göre sınıflandırma işlemleri sonucu elde edilen ürünün kalitesi değişmektedir.
- Yakın kızılötesi bölgede elde edilen verilerin sınıflandırma için önemli bir altlık olduğu belirlenmiştir.
- Destek vektör makinelerinin sınıflandırma performansına seçilecek kernel fonksiyonunun etki ettiği görülmüştür. Gaussian kerneli lineer kerneline göre daha başarılı sonuç vermiştir.
- Karar ağaçları yönteminin diğer yöntemlere göre daha düşük doğruluk değerlerine sahip olduğu görülmüştür.
- Torbalama karar ağaçları yöntemi verdiği doğruluk değerleri ile karar ağaçları yöntemini geride bırakmıştır.
- Torbalama karar ağaçları ve *k* en yakın komşu yöntemleri az sayıda bant kullanarak yapılan sınıflandırma işlemlerinde de yüksek doğruluk değerleri vererek bantların içerdiği bilgileri işlemek konusunda başarılı olduklarını göstermişlerdir.
- Yapılan sınıflandırma çalışması ışığına en yüksek doğruluk değerini veren yöntem derin öğrenme(oto kodlayıcı) yöntemidir.
- Gelişen GPU gücüne bağlı olarak derin öğrenme ve komşuluk bileşen analizinin verdiği sonuç ürünler tatmin edici olup, literatüre kazandırılacak yeni çalışmalar için önemli birer yöntem konumunda olacağı tahmin edilmektedir. Ayrıca işlemci gücüne göre yöntemin tepki süresi de değişmektedir ve doğruluk değerlerini azda olsa değiştirdiği belirlenmiştir.

Görüntülerden elde edilmiş sınıflandırma sonuçlarının doğru bir şekilde yorumlanabilmesi, ilerleyen zamanlarda verilecek kararları etkilemektedir. Tezde kullanılan yöntemlerin doğruluk değerlerinden anlaşıldığı üzere geliştirilen her teknolojiyle beraber gelen yöntemlerin doğruluk değerlerinin artırıldığı gözlemlenmiştir ve yakın gelecekte daha da iyi sonuçlar veren yöntemlerin geliştirileceği düşünülmektedir.

4.2. Öneriler

Sınıflandırma uygulamalarında kullanılacak yöntemlerin başarılı olması; kullanılacak görüntünün kalitesine, görüntüyü sınıflandırmakta kullanılacak eğitim ve test verisinin doğru seçimine ve sınıflandırmayı gerçekleştirecek kullanıcıya bağlıdır. Bu sebepler doğrultusunda yapılan her yeni uygulamada değerlerin değişebileceği sonucuna varılmıştır. Makine öğrenmesi tekniklerinin sınıflandırma performansları incelenirken de bu kriterler göz önünde bulundurulmalıdır ve aynı veri seti kullanılarak sınıflandırma işlemi tamamlanmalıdır.

KAYNAKÇA

1. Kavzoğlu, T., Çölkesen, İ., 2010. Karar ağaçları ile uydu görüntülerinin sınıflandırılması: Kocaeli örneği. **Harita Teknolojileri Elektronik Dergisi, 2**(1): 36-45.

2. Çölkesen, İ., 2009. Uzaktan Algılamada İleri Sınıflandırma Tekniklerinin Karşılaştırılması ve Analizi. Gebze Yüksek Teknolojisi Enstitüsü, Yüksek Lisans Tezi, Gebze, 138 s.

3. Campbell, J.B., 1996. Introduction to remote sensing. Guilford Press, New York, 621 s.

4. Atasever, Ü.H., 2011. Uydu Görüntülerinin Sınıflandırılmasında Hızlandırma (Boosting), Destek Vektör Makineleri, Rastgele Orman (Random Forest) ve Regresyon Ağaçları Yöntemlerinin Kullanılması. Erciyes Üniversitesi, Yüksek Lisans Tezi, Kayseri, 65 s.

5. Cortes, C. Vapnik, V.N., 1995. Support vector network. **Machine Learning, 20** (3): 273–297.

6. Vapnik, V.N., 2013. The Nature of Statistical Learning Theory. Springer Science & Business Media, 299 s.

7. Kechman, V., 2001. Learning and Soft Computing. MIT Press, 536 s.

8. Atasever, Ü.H., Özkan, C., 2012. Arazi örtüsünün belirlenmesinde torbalama karar ağaçları yönteminin kullanılması. UZAL - CBS, Zonguldak.

9. Safavian S.R., Landgrebe D., 1991. A survey of decision tree classifier methodology. **IEEE Transactions On Systems Man And Cybernetics, 21** (3): 660-674.

10. Akman, M., 2010. Veri Madenciliğine Genel Bakış ve Random Forests Yönteminin İncelenmesi: Sağlık Alanında Bir Uygulama. Ankara Üniversitesi, Yüksek Lisans Tezi, Ankara, 90 s.

11. Quinlan, J. R., 1993. C4.5: Programs for machine learning. **Morgan Kaufmann Publishers, 16** (3): 235-240.

12. Goldberger, J., Roweis, S., Hinton, G., Salakhutdinov, R., 2005. Neighbourhood components analysis. **In Advances in Neural Information Processing Systems, 17**: 513-520.

13. Weinberger, K.Q., Saul, L.K., 2009. Distance metric learning for large margin nearest neighbor classification. **Journal of Machine Learning Research, 10**: 207-244.

14. Karahan, Ş., Akgül, Y., 2016. Eye detection by using deep learning. 24th Signal Processing and Communication Application Conference, Zonguldak.

15. Sunar, F., Özkan, C., Osmanoğlu, B., 2011. Uzaktan Algılama. Anadolu Üniversitesi Yayınları, Eskişehir. 210 s.

16. Okur, A.Y., 2017. Uzaktan Algılama ve Coğrafi Bilgi Sistemi Kullanımı ile Potansiyel Termal Su Alanlarının Araştırılması: Kaz Dağları ve Çevresi Örneği. İstanbul Teknik Üniversitesi, Yüksek Lisans Tezi, İstanbul, 88s.

17. Özkan, C., 2013. Uzaktan Algılama Ders Notları (Yayınlanmamış), Lisans Programı, Kayseri.

18. Örmeci, C., 1987. Uzaktan algılama:Temel esaslar ve algılama sistemleri. **İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, 1**(6): 1345-1346.

19. Gibson, P.J., 2001. Introductory remote sensing: Principles and concepts. **Routledge publishers, 138**(2):219-232.

20. Yılmaz, M., 2016. Taşkın Riski Ön Değerlendirmesinde Uzaktan Algılama Sistemlerinin Kullanılabilirliği: Meriç Nehrinde Uygulanması. T.C. Orman ve Su İşleri Bakanlığı, Uzmanlık Tezi, Ankara, 84 s.

21. Özbalmumcu, M., Erdoğan, M., 2001. Uzaktan algılama amaçlı uydu görüntüleme sistemleri. **Harita Dergisi, 125**(6): 59-82 s.

22. Şenyiğit, S., 2019. Coğrafi Bilgi Sistemleri ve Uzaktan Algılama Yöntemleri ile Arazi Örtüsü/Alan Kullanım Değişimlerinin Belirlenmesi: Bingöl Kent Merkezi Örneği. Atatürk Üniversitesi, Yüksek Lisans Tezi, Erzurum, 97s.

23. Erbay, A., 2014. Farklı Özellikleriyle Yeni Nesil Dünya Gözlem Uyduları ve Kullanım Alanları. Uzaktan Algılama-CBS Sempozyomu, İstanbul.

24. Kavzoglu, T., 2001. An Investigation Of The Design And Use Of Feed-Forwad Artificial Neural Networks İn The Classification Of Remotely Sensed Images. Nottingham University, Phd Thesis, UK, 308 s.

25. Özkan, C., 1998. Uzaktan Algılama Verileriyle Orman Yangını Analizi. İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, İstanbul, 100 s.

26. Gu, Y., Liu T., Jia, X., Benediktsson, J. A., ve Chanussot, J., 2016. Nonlinear multiple kernel learning with multiple-structure-element extended morphological profiles for hyperspectral image classification. **IEEE Transactions on Geoscience and Remote Sensing, 54**(6): 3235–3247 s.

27. Schölkopf, B., Simard, P., Smola, AJ.,Vapnik, V., 1998. Prior knowledge in support vector kernels. **Advances in Neural Information Processing Systems, 10:**640-646.

28. Erdal, H., 2015. Contribution of machine learning methods to the construction industry: Prediction of compressive strength. **Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 21**(3):109-114.

29. Demirci, D., 2007. Destek Vektör Makineleri İle Karakter Tanıma. Yıldız Teknik Üniversitesi, Yüksek Lisans Tezi, İstanbul, 53 s.

30. Kavzoğlu, T., Çölkesen, İ., 2010. Destek vektör makineleri ile uydu görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi. **Harita Dergisi, 144** (7): 73-82 s.

31. Osuna, E.E., Freund, R., Girosi, F., 1997. Support vector machines: training and applications, **Massachusetts Institute of Technology and Artificial Intelligence Laboratory, 144:**42.

32. Foody, G.M. and Mathur A., 2004. A relative evaluation of multiclass image classification by support vector machines. **IEEE Transactions on Geoscience and Remote Sensing, 42**(6): 1335–1343.

33. Özkan, Y., 2008. Veri Madenciliği Yöntemleri. Papatya Yayıncılık, İstanbul, 216 s.

34. Yang, W., Wang K., Zuo, W., 2012. Fast neighborhood component analysis. **Neurocomputing, 83**:31-37 s.

35. Singh-Miller, N., Collins, M., Hazen, T.J, 2007. Dimensionality reduction for speech recognition using neighborhood components analysis. **Eighth Annual Conference of the International Speech Communication Association**: 1158-1161.

36. Liu, J., Chen, S, 2006. Discriminant common vectors versus neighbourhood components analysis and Laplacianfaces: A comparative study in small sample size problem. **Image and Vision Computing, 24**(3): 249-262.

37. Kavzoğlu, T., Çölkesen, İ., 2010. Karar ağaçları ile uydu görüntülerinin sınıflandırılması: Kocaeli örneği, **Harita Teknolojileri Dergisi, 2**(1): 36-45.

38. Swain P.H., Hauska H., 1977. Decision tree classifier - design and potential, **IEEE Transactions on Geoscience and Remote Sensing, 15**(3): 142-147 s.

39. Tan, P., Steinbach, M., Kumar, V., 2005. Introduction to data mining. **Addison Wesley, 321321367**:169.

40. Kantardzic, M., 2003. Data mining: concepts, methods and Algorithms. **Technometrics,45**(3):277.

41. Pal, M., 2002. Factors Influencing the Accuracy of Remote Sensing Classifications: a Comparative Study. University of Nottingham, Phd thesis, 279 s.

42. Opitz, D., Maclin, R., 1999. Popular ensemble methods: An empirical study. **Journal of Artificial Intelligence Research, 11**: 169–198.

43. Ye, J., Janardan, R., Li, Q., 2005. Two diensional linear discriminant analysis. Advenced in Neural Information Processing Systems, Kanada.

44. Alpaydin, E., 2004. Introduction to Machine Learning. MIT Press, 407 s.

45. LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. **Nature, 521**: 436–444.

46. Deng, L., Yu, D., 2014. Deep learning: Methods and applications. **Foundations and Trends Signal Process, 7**(3-4): 197–387.

47. Liou, C.Y., Cheng, W.C., Liou, J.W., Liou, D.R., 2014. Autoencoder for words. **Neurocomputing, 139**: 84–96.

48. Güven, M.A., Atasever, Ü.H., Beşdok, E., 2018. Hiperspektral görüntülerin oto-kodlayıcı derin öğrenme ile sınıflandırılması. VII. UZAL - CBS, Eskişehir.

49. Yuksel, M.E., Basturk, N.S., Badem, H., Caliskan, A., Basturk, A., 2018. Classification of high resolution hyperspectral remote sensing data using deep neural networks. **Journal of Intelligent & Fuzzy Systems, 34**(4): 2273-2285.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı: Ali Mahir ÇINAR
Uyruğu: Türkiye (T.C)
Doğum Tarihi ve Yeri: 17.06.1992 - Ankara
Medeni Durum: Bekar
e-mail: alimahircinar@gmail.com

EĞİTİM

Derece	Kurum	Mezuniyet Tarihi
Yüksek Lisans	Erciyes Üniversitesi, Fen Bilimleri Enstitüsü	2019
Lisans	Erciyes Üniversitesi, Harita Mühendisliği	2016
Lise	Bekir Gökdağ Lisesi, ANKARA	2010

İŞ DENEYİMLERİ

Yıl	Kurum	Görev
2019-	ALPKON İNŞAAT	Harita Mühendisi
2017-2018	MSK ENERJİ	Harita Mühendisi
2016-2017	ALPKON İNŞAAT	Harita Mühendisi

YABANCI DİL

İngilizce