



**GELİŞTİRİLMİŞ KATMANLI UZAY  
YERLEŐTİRME YÖNTEMLERİ KULLANILARAK  
HİPERSPEKTRAL GÖRÜNTÜLERİN  
SINIFLANDIRILMASI VE GÖRSELLEŐTİRİLMESİ**

**Mehmet Zahid YILDIRIM**

**2021  
DOKTORA TEZİ  
BİLGİSAYAR MÜHENDİSLİĐİ**

**Tez Danıőmanı  
Dr. Öğr. Üyesi Caner ÖZCAN**

**GELİŞTİRİLMİŞ KATMANLI UZAY YERLEŐTİRME YÖNTEMLERİ  
KULLANILARAK HİPERSPEKTRAL GÖRÜNTÜLERİN  
SINIFLANDIRILMASI VE GÖRSELLEŐTİRİLMESİ**

**Mehmet Zahid YILDIRIM**

**T.C.  
Karabük Üniversitesi  
Lisansüstü Eğitim Enstitüsü  
Bilgisayar Mühendisliği Anabilim Dalında  
Doktora Tezi  
Olarak Hazırlanmıştır**

**Tez Danışmanı  
Dr. Öğr. Üyesi Caner ÖZCAN**

**KARABÜK  
Aralık 2021**

Mehmet Zahid YILDIRIM tarafından hazırlanan “GELİŞTİRİLMİŞ KATMANLI UZAY YERLEŞTİRME YÖNTEMLERİ KULLANILARAK HİPERSPEKTRAL GÖRÜNTÜLERİN SINIFLANDIRILMASI VE GÖRSELLEŞTİRİLMESİ” başlıklı bu tezin Doktora Tezi olarak uygun olduğunu onaylarım.

Dr. Öğr. Üyesi Caner ÖZCAN .....  
Tez Danışmanı, Bilgisayar Mühendisliği Anabilim Dalı

Prof. Dr. Okan ERSOY .....  
Tez Danışmanı, School of Electrical and Computer Engineering, Purdue University

Bu çalışma, jürimiz tarafından oy birliği ile Bilgisayar Mühendisliği Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir. 30/12/2021

Ünvanı, Adı SOYADI (Kurumu) İmzası

Başkan : Prof. Dr. Oğuz FINDIK (KBÜ) .....

Üye : Dr. Öğr. Üyesi Caner ÖZCAN (KBÜ) .....

Üye : Doç. Dr. Okan ERKAYMAZ (MSÜ) .....

Üye : Prof. Dr. İsmail Rakıp KARAŞ (KBÜ) .....

Üye : Doç. Dr. Fatih NAR (AYBÜ) .....

KBÜ Lisansüstü Eğitim Enstitüsü Yönetim Kurulu, bu tez ile, Doktora derecesini onamıştır.

Prof. Dr. Hasan SOLMAZ .....  
Lisansüstü Eğitim Enstitüsü Müdürü

*“Bu tezdeki tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu; ayrıca bu kuralların ve ilkelerin gerektirdiği şekilde, bu çalışmadan kaynaklanmayan bütün atıfları yaptığımı beyan ederim.”*

Mehmet Zahid YILDIRIM

## ÖZET

Doktora Tezi

### GELİŞTİRİLMİŞ KATMANLI UZAY YERLEŞTİRME YÖNTEMLERİ KULLANILARAK HİPERSPEKTRAL GÖRÜNTÜLERİN SINIFLANDIRILMASI VE GÖRSELLEŞTİRİLMESİ

Mehmet Zahid YILDIRIM

Karabük Üniversitesi

Lisansüstü Eğitim Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı:

Dr. Öğr. Üyesi Caner ÖZCAN

Prof. Dr. Okan ERSOY

Aralık 2021, 79 sayfa

Birçok uzaktan algılama uygulaması, yüksek boyutlu bir uzayda bulunan gerçek dünya nesnelerinin anlamlı ve yorumlanabilir kalıplarını ortaya çıkarmaya dayanır. Özellikle görüntüler, farklı malzemelerin özelliklerini yansıtan yüzlerce yüksek çözünürlüklü banttı oluşabilir. Bu işi kolaylaştırmak için, bu tür büyük verilerin karakter özelliklerini koruyarak daha düşük boyutlarda temsil edilmesi gerekmektedir. Boyutsallık problemine bağlı olarak bu tür büyük verilerin analizi ve işlenmesi birçok zorluğu da yanında getirmektedir. Son yıllarda yapılan çalışmalar bu zorlukların üstesinden gelmek için bir boyut indirgeme yaklaşımı olan katmanlı uzay (manifold) öğrenme tekniklerinin oldukça önemli bir çözüm olduğunu göstermektedir. Ancak, bu tekniklerinde karmaşıklığı artan verilerde bazı dezavantajları ve yeterli olmadığı durumlar oluşmaktadır.

Bu tez çalışmasında, manifold öğrenme tekniklerinin bu dezavantajlarından ve çözüm önerilerinden bahsedilmektedir. Bu kapsamda, iki farklı çalışma ile iki yeni geliştirilmiş manifold yerleştirme yöntemi önerilmektedir. İlk yöntem, literatürde bulunan popüler manifold yerleştirme yöntemlerinin hibrit bir yaklaşımla güçlü yönlerinin birlikte kullanılmasını esas almaktadır. İkinci yöntemde ise, optimizasyon tabanlı bir çözüm önerilmektedir. Önerilen yöntemlerin başarısı popüler hiperspektral veri setleri ile test edilmektedir. Her iki çalışmada da mevcut yöntemlere göre sınıflandırma başarısı artırılmakta, kalabalık problemine çözüm getirilmekte ve daha iyi görselleştirmeler sunulmaktadır.

**Anahtar Sözcükler :** Manifold Öğrenme, boyut indirgeme, hiperspektral görüntüleme, çok boyutlu alan yerleştirme, sınıflandırma.

**Bilim Kodu** : 92431

## **ABSTRACT**

**Ph. D. Thesis**

### **CLASSIFICATION AND VISUALIZATION OF HYPERSPECTRAL IMAGE USING ENHANCED MANIFOLD EMBEDDING METHODS**

**Mehmet Zahid YILDIRIM**

**Karabük University  
Institute of Graduate Programs  
Department of Computer Engineering**

**Thesis Advisor:**

**Assist. Prof. Dr. Caner ÖZCAN**

**Prof. Dr. Okan ERSOY**

**December 2021, 79 pages**

Many remote sensing applications rely on uncovering meaningful and interpretable patterns of real-world objects located in a high-dimensional space. In particular, images can consist of hundreds of high-resolution bands that reflect the properties of different materials. To facilitate this work, such big data should be represented in lower dimensions while preserving their character traits. Depending on the dimensionality problem, the analysis and processing of such big data brings with it many difficulties. Recent studies show that manifold learning techniques, which is a dimension reduction approach, are a very important solution to overcome these difficulties. However, as the complexity of the data increases, these techniques have some disadvantages, and their performance are not sufficient.

In this thesis, these disadvantages and solution suggestions of manifold learning techniques are mentioned. In this context, two new improved manifold embedding methods are proposed in two different studies. The first method is based on a hybrid approach that combines the strengths of the popular manifold embedding methods in the literature. In the second method, an optimization-based solution is proposed. The success of the proposed methods is tested with popular hyperspectral datasets. In both studies, classification accuracy is increased, crowding problem is solved and better visualizations are presented according to existing methods.

**Key Word** : Manifold learning, Dimensional Reduction, Hyperspectral Images, Multidimensional Field Embedding, Classification.

**Science Code** : 92431



## TEŐEKKÜR

Bu tez alıőmasının planlanmasında, araőtırılmasında, yürütülmesinde ve oluşumunda ilgi ve desteęini esirgemeyen, engin bilgi ve tecrübelerinden yararlandıęım, yönlendirme ve bilgilendirmeleriyle alıőmamı bilimsel temeller ışığında őekillendiren sayın hocalarım Dr. Öğr. Üyesi Caner ÖZCAN ve Prof. Dr. Okan ERSOY'a sonsuz teşekkürlerimi sunarım.

Manevi destekleriyle beni hiçbir zaman yalnız bırakmayan değerli eşim Öznur YILDIRIM'a ve aileme tüm kalbimle teşekkür ederim.

Bu alıőma TÜBİTAK 1002 Hızlı Destek Projesi olarak 120E404 proje numarasıyla ve Karabük Üniversitesi tarafından Bilimsel Araőtırma Projesi olarak FDT-2020-2315 proje numarasıyla desteklenmiőtir. Projelerde desteklerini esirgemeyen Sinan ASLAN'a da teşekkür ederim.

## İÇİNDEKİLER

	<u>Sayfa</u>
KABUL.....	ii
ÖZET.....	iv
ABSTRACT.....	vi
TEŞEKKÜR.....	viii
İÇİNDEKİLER .....	ix
ŞEKİLLER DİZİNİ.....	xii
ÇİZELGELER DİZİNİ .....	xiv
SİMGELER VE KISALTMALAR DİZİNİ .....	xv
BÖLÜM 1 .....	1
GİRİŞ .....	1
BÖLÜM 2 .....	7
UZAKTAN ALGILAMA VE HİPERSPEKTRAL GÖRÜNTÜLEME.....	7
2.1. UZAKTAN ALGILAMA .....	7
2.1.1. Uzaktan Algılamanın Bileşenleri ve Aşamaları .....	10
2.1.2. Uzaktan Algılamada Algılayıcılardan Elde Edilen Çözünürlük Türleri. 12	
2.2. ELEKTROMANYETİK SPEKTRUM .....	17
2.3. GÖRÜNTÜ TÜRLERİ.....	19
2.4. HİPERSPEKTRAL GÖRÜNTÜLEME.....	22
BÖLÜM 3 .....	24
BOYUT İNDİRGEME VE MANİFOLD ÖĞRENME .....	24
3.1. BANT SEÇİMİ.....	25
3.2. ÖZELLİK ÇIKARIMI .....	26
3.3. MANİFOLD ÖĞRENME .....	27
BÖLÜM 4 .....	29

	<b><u>Sayfa</u></b>
BOYUT İNDİRGEME YÖNTEMLERİ .....	29
4.1. DOĞRUSAL YÖNTEMLER.....	29
4.1.1. Temel Bileşenler Analizi (PCA).....	29
4.1.2. Çok Boyutlu Ölçekleme (MDS) .....	31
4.1.3. Yerel Fisher Diskriminant Analizi (LFDA) .....	31
4.2. DOĞRUSAL OLMAYAN YÖNTEMLER .....	32
4.2.1. Yerel Doğrusal Yerleştirme (LLE).....	32
4.2.2. İzometrik Özellik Haritalama (ISOMAP) .....	33
4.2.3. Laplas Özharitalar (LE) .....	34
4.2.4. Rastlantısal Komşu Yerleştirme (SNE) .....	34
4.2.5. t-Dağılımlı Rastlantısal Komşu Yerleştirme (t-SNE).....	36
4.2.6. Çok Boyutlu Yapay Alan Yerleştirme (MAFE).....	37
BÖLÜM 5 .....	40
BOYUT İNDİRGEME UYGULAMALARINDA OPTİMİZASYON VE ADAPTİF AĞIRLIKLANDIRMA.....	40
5.1. PARÇACIK SÜRÜ OPTİMİZASYONU TABANLI MAFE.....	40
5.2. ADAPTİF AĞIRLIKLANDIRILAN MANİFOLD YERLEŞTİRME .....	43
BÖLÜM 6 .....	47
DENEYSEL ÇALIŞMALAR .....	47
6.1. KULLANILAN VERİ SETLERİ.....	47
6.1.1. Botswana.....	47
6.1.2. Kennedy Uzay Merkezi .....	48
6.1.3. Indiana Pines.....	50
6.1.4. Salinas.....	51
6.1.5. Pavia Üniversitesi .....	52
6.2. HİPERSPEKTRAL GÖRÜNTÜLERİN SINIFLANDIRILMASINDA FARKLI BOYUT İNDİRGEME YÖNTEMLERİNİN KARŞILAŞTIRILMASI 54	
6.3. OPTİMİZASYON TABANLI MANİFOLD YERLEŞTİRME YÖNTEMİ İLE HİPERSPEKTRAL GÖRÜNTÜLERİN SINIFLANDIRILMASI VE GÖRSELLEŞTİRİLMESİ.....	57

6.4. ADAPTİF AĞIRLIKLANDIRILAN MANİFOLD YERLEŐTİRME  
YÖNTEMİ İLE HİPERSPEKTRAL GÖRÜNTÜLERİN SINIFLANDIRMASI 64

BÖLÜM 7 .....	69
SONUÇLAR .....	69
KAYNAKLAR .....	71
ÖZGEÇMİŐ .....	79

## ŞEKİLLER DİZİNİ

### Sayfa

Şekil 2.1. 1840'lı yıllarda kullanılan kameralar [34]. .....	7
Şekil 2.2. İlk hava fotoğrafı örnekleri [35]. .....	8
Şekil 2.3. Fotoğrafçı güvercinler ve çektikleri fotoğraf örnekleri [36]. .....	8
Şekil 2.4. Sputnik-1 uydusu [38]. .....	9
Şekil 2.5. Uzaktan algılama süreci [39]. .....	11
Şekil 2.6. Çok bantlı görüntü örneği. ....	13
Şekil 2.7. Aynı bölgeye ait farklı mekansal çözünürlüğe sahip görüntüler [40]. .....	14
Şekil 2.8. Aynı bölgeye ait farklı radyometrik çözünürlüğe sahip görüntüler [41]. ..	15
Şekil 2.9. Farklı zamanlarda aynı bölgeye ait görüntüler [42]. .....	16
Şekil 2.10. Çeşitli dünya nesnelere ait spektral veri gösterimi [43]. .....	16
Şekil 2.11. Elektromanyetik spektrum [45]. .....	17
Şekil 2.12. Mars yüzeyine ait kaya parçasının spektral yansıma grafiği [46]. .....	18
Şekil 2.13. Pankromatik uydu görüntüsü [47]. .....	19
Şekil 2.14. Multispektral uydu görüntüsü [48]. .....	20
Şekil 2.15. Aynı bölgeye ait uydu görüntüleri [52]. .....	21
Şekil 2.16. Hiperspektral veri küpü [53]. .....	22
Şekil 2.17. Farklı malzemeler için elde edilmiş spektral imzalar [53]. .....	23
Şekil 3.1. Hiperspektral verilerde bant seçimi ve özellik çıkarımı. ....	24
Şekil 3.2. Hiperspektral veri üzerindeki bant seçim işlem adımları [58]. .....	25
Şekil 3.3. Hiperspektral verilerde özellik çıkarımı [60]. .....	26
Şekil 3.4. Çeşitli manifold öğrenme yöntemlerine ait çıktı uzayları [61]. .....	27
Şekil 4.1. PCA algoritması ile boyut indirgeme [64]. .....	30

## **Sayfa**

Şekil 4.2. Veri noktaları arasındaki jeodezik mesafenin gösterimi [10].	33
Şekil 4.3. Gauss ve t-dağılımı grafikleri.	36
Şekil 5.1. PSO algoritmasının temsili çalışma simülasyonu [86].	41
Şekil 6.1. a) Botswana renkli görüntüsü b) doğrulama verisi c) sınıflar [88].	48
Şekil 6.2. a) KSC renkli görüntüsü b) doğrulama verisi c) sınıflar [53].	49
Şekil 6.3. a) IP renkli görüntüsü b) doğrulama verisi c) sınıflar [88].	50
Şekil 6.4. a) Salinas renkli görüntüsü b) doğrulama verisi c) sınıflar [88].	51
Şekil 6.5. a) PU renkli görüntüsü b) doğrulama verisi c) sınıflar [33].	52
Şekil 6.6. Botswana verisi için çalışma zamanı.	55
Şekil 6.7. KSC verisi için çalışma zamanı.	56
Şekil 6.8. PSO parametre optimizasyon grafikleri a) Botswana b) KSC.	58
Şekil 6.9. PSO parametre optimizasyon grafikleri için hata yakınsaması.	59
Şekil 6.10. Botswana verisi için görselleştirme sonuçları a) SNE b) <i>t</i> -SNE c) MAFE d) PSO-MAFE.	62
Şekil 6.11. KSC verisi için görselleştirme sonuçları a) SNE b) <i>t</i> -SNE c) MAFE d) PSO-MAFE.	63
Şekil 6.12. <i>t</i> -SNE, MAFE ve önerilen yöntem için sınıflandırma doğruluğu grafiği.	65
Şekil 6.13. Önerilen yöntemlerin iterasyona bağlı sınıflandırma doğruluğu grafiği.	66
Şekil 6.14. KSC verisi için görselleştirme sonuçları a) <i>t</i> -SNE b) MAFE c) Önerilen yöntem.	67

## ÇİZELGELER DİZİNİ

### Sayfa

Çizelge 2.1. Türkiye'nin sahip olduğu yer gözlem uyduları.....	10
Çizelge 5.1. AWEM algoritması.....	45
Çizelge 6.1. Botswana doğrulama verisi sınıfları ve örnek sayıları.....	48
Çizelge 6.2. KSC doğrulama verisi sınıfları ve örnek sayıları.....	49
Çizelge 6.3. IP doğrulama verisi sınıfları ve örnek sayıları.....	50
Çizelge 6.4. Salinas doğrulama verisi sınıfları ve örnek sayıları.....	52
Çizelge 6.5. PU doğrulama verisi sınıfları ve örnek sayıları.....	53
Çizelge 6.6. Botswana verisi için sınıflandırma sonuçları.....	54
Çizelge 6.7. KSC verisi için sınıflandırma sonuçları.....	55
Çizelge 6.8. Botswana verisi için sınıflandırma sonuçları.....	60
Çizelge 6.9. KSC verisi için sınıflandırma sonuçları.....	60
Çizelge 6.10. IP, PU ve Salinas verileri için sınıflandırma sonuçları.....	61
Çizelge 6.11. KSC verisi için sınıflandırma sonuçları.....	64
Çizelge 6.12. Tüm yöntemler için çalışma zamanları.....	66

## SİMGELER VE KISALTMALAR DİZİNİ

### SİMGELER

- $Y$  : Yüksek boyutlu uzay gözlem verileri  
 $Z$  : Düşük boyutlu uzay yerleştirme koordinatları  
 $\Sigma$  : Kovaryans matrisi  
 $C$  : Merkezleme matrisi  
 $I$  : Birim matrisi  
 $E$  :  $\Sigma$ 'in temel bileşenler matrisi  
 $d_{ij}$  : Veri noktaları arasındaki uzaklık fonksiyonu  
 $W$  :  $N \times N$  boyutlu ağırlık matrisi  
 $w_{ij}$  : Veri noktaları arası ikili ağırlıklar  
 $\mathbb{R}$  : Reel sayılar  
 $\sigma_i$  : Varyans Parametresi  
 $p_{ij}$  : Yüksek Boyutlu Uzayın Yoğunluk Dağılımı  
 $q_{ij}$  : İndirgenmiş Uzayın Yoğunluk Dağılımı  
 $U_r^{ij}$  : İtme Potansiyel Fonksiyonu  
 $U_a^{ij}$  : Çekme Potansiyel Fonksiyonu  
 $\xi_r$  : İtme Kuvveti Büyüklük Parametresi  
 $\xi_a$  : Çekme Kuvveti Büyüklük Parametresi  
 $\nabla_{z_i}$  : Gradyan Operatörü  
 $p$  ve  $q$  : İndirgenmiş Uzay Dağılım Yoğunluk parametreleri  
 $x_i(t)$  :  $i$  parçacığının  $t$  zamanındaki konumu  
 $v_i(t)$  :  $i$  parçacığının  $t$  zamanındaki hızı



## KISALTMALAR

- PCA : Principal Component Analysis (Temel Bileşenler Analizi)
- MDS : Multidimensional Scaling (Çok Boyutlu Ölçekleme)
- LFDA : Local Fisher Discriminant Analysis (Yerel Fisher Diskriminant Analizi)
- LLE : Locally Linear Embedding (Yerel Doğrusal Yerleştirme)
- LE : Laplacian Eigenmaps (Laplas Özharitalar)
- LTSA : Local Tangent Space Alignment (Local Tanjant Uzay Hizalama)
- MVU : Maximum Variance Unfolding (Maksimum Açılma Varyansı)
- SOFM : Self-Organizing Feature Map (Öz-Düzenlemeli Özellik Haritası)
- SNE : Stochastic Neighbor Embedding (Rastlantısal Komşu Yerleştirme)
- t-SNE : t-Distribution Stochastic Neighbor Embedding (t-dağılımlı Rastlantısal Komşu Yerleştirme)
- sSNE : Spherical Stochastic Neighbor Embedding (Küresel Rastlantısal Komşu Yerleştirme)
- MAFE : Multidimensional Artificial Field Embedding (Çok Boyutlu Yapay Alan Yerleştirme)
- KSC : Kennedy Space Center (Kennedy Uzay Merkezi)
- IP : Indiana Pines
- PU : Pavia University
- KS : Kappa Statistic
- KL : Kullback-Leibler Divergence Measure (Kullback-Leibler Sapma Ölçümü)
- AWEM : Adaptive Weighted Embedding Method (Adaptif Ağırlıklandırılmış Gömme Yöntemi)
- 1NN : One Nearest Neighbor (En yakın komşu)

## BÖLÜM 1

### GİRİŞ

Uzaktan algılamada, birçok gerçek dünya nesnesi yüksek boyutlu temsillerle ilişkilendirilebilir. Görüntüler, konuşma sinyalleri, videolar ve metin belgeleri gibi büyük hacimli verilerin analizi ve işlenmesi için araştırmacılar pratik sistemler geliştirmelidir. Büyük veri kümeleri, verilerdeki tüm düzeni ve karakteristiği yakalayan daha düşük boyutlarda bir veri kümesinde özetlenmelidir. Özellikle görüntüler, farklı malzemelerin özelliklerini yansıtan yüzlerce yüksek çözünürlüklü banttandır oluşur. Yüksek boyutlu uzayda çok sayıda farklı bantların bulunması, bu özelliklerin yorumlanmasını zorlaştırmaktadır. Uzaktan algılama verilerinin ön-işlemesi için boyutsallık problemine bağlı olarak çeşitli zorluklar ile karşılaşmaktadır. Bu alanda uzun yıllardır ortaya çıkan araştırma, bunun zor bir problem olduğunu ve tüm sorunlara tek bir çözüm olmadığını ortaya koymaktadır. Son araştırmalar, hiperspektral görüntülerin ön işleminde boyut indirgeme ve manifold öğrenme tekniklerinin çok önemli bir çözüm olduğunu göstermektedir. Bu teknikler, görüntülerde daha iyi sıkıştırma ve görselleştirme olanağı sunarken, yararlı sınıflandırma özelliklerinin çıkarılmasını başarılı bir şekilde sağlamaktadır.

Yüksek boyutlu verilerin yüksek doğrulukla sınıflandırılması, birçok araştırma alanında ortaya çıkan önemli bir sorundur. Yüksek boyutlu bir özellik alanında temsil edilen anlamlı veri kalıplarını bulmak oldukça zordur. Bu zorlukları ele almak için, ilk olarak geleneksel doğrusal temelli boyut indirgeme teknikleri geliştirilmiştir. Bunların en yaygın bilineni temel bileşenler analizi (principal components analysis-PCA) [1-4], yüksek boyutlu uzayda yer alan verileri daha düşük boyutlu bir koordinat sistemine yansıtmak için dikey bir dönüşüm kullanmaktadır. İndirgenmiş uzayda, veri noktaları arasındaki korelasyonun minimuma inmesi hedeflenmektedir. Diğer bir popüler yöntem olan çok boyutlu ölçekleme (multidimensional scaling-MDS) [5], yüksek boyutlu uzaydaki ikili mesafeleri koruyarak daha düşük boyutlu bir koordinat sistemi

üzerine yerleştirir. İkili mesafelerde öklid uzaklığı kullanılmaktadır. Yerel Fisher diskriminant analizi (local Fisher discriminant analysis-LFDA) [6,7], yüksek boyutlu giriş uzayındaki verilerin yerel yapısını koruyan bir doğrusal boyut indirgeme yöntemidir. LFDA, yüksek boyutlu uzaydaki aynı sınıfa ait verilerin varyansını en aza, farklı sınıfların varyansını ise maksimuma çıkararak bu yerel yapıyı korumayı amaçlamaktadır. Yarı denetimli yerel Fisher diskriminant analizi (Semi-supervised Local Fisher discriminant analysis-SELF) [8], farklı sınıflardaki etiketli örnekleri birbirinden ayırmanın yanı sıra, etiketlenmemiş örneklerin global yapısını koruyan yarı denetimli bir boyut azaltma yöntemidir. Doğrusal yöntemlerin tümü yüksek boyutlu girdi uzayının doğrusal bir alt uzay üzerindeki yapısını keşfeder. Bununla birlikte, birçok gerçek dünya verisinin altında yatan yerel yapıları ve doğrusal olmayan geometrileri tespit etme yeteneğine sahip değildirler.

İlerleyen yıllarda özellikle hiperspektral görüntülerde artan veri boyutları ve detaylar, doğrusal yöntemlerin bu verileri indirgemek için çoğu zaman yeterli olmadığını göstermiştir. Bunun sonucunda, doğrusal olmayan yöntemler ortaya çıkmıştır. Yerel doğrusal yerleştirme (locally linear embedding-LLE) [9], doğrusal bir dönüşümle komşuluk ilişkisini temsil eder ve komşu bölgelerden her bir veri noktasının en iyi yeniden yapılandırılmasını sağlar. İzometrik özellik haritalama (isometric feature mapping-ISOMAP) [10], hem verilerin yerel komşuluğunda hem de parçalı kombinasyonlarında doğrusallık varsayılarak küresel doğrusal olmayan yapılar oluşturmak için önerilmiştir. ISOMAP ilk önce LLE’de olduğu gibi yerel bir çözüm üretir ve daha sonra MDS üzerinden doğrusal olmayan veriler için küresel optimal sistemi elde eder. Laplace özharitalar (Laplacian eigenmaps-LE) [11], bir katmanlı uzay üzerindeki Laplas Beltrami operatörü ve ısı denklemi bağlantıları arasındaki ilişkiyi kullanır. LE çözüm alanlarının yanı sıra LLE, lineer transformasyon ve Laplas matrislerinin izleyen öz vektörlerini içerir. Lokal tanjant uzay hizalamada (local tangent space alignment-LTSA) [12], katmanlı uzayın yerel geometrisi, her bir veri noktasında teğet uzayı için bir yaklaşımın oluşturulmasıyla öğrenilir ve küresel katmanlı uzay, daha sonra temel noktaya göre veri noktalarının global koordinatlarını vermek üzere hizalanır. Boyutların azaltılması için etkili bir keşif olan maksimum açılma varyansı (maximum variance unfolding-MVU) [13], yerel mesafelerin korunmasına bağlı varyans yerleştirme haritalarının en üst düzeye çıkarılmasıyla

verilerin düşük boyutlu bir şekilde temsil edilmesini sağlar. Yapay gözetimsiz bir sinir ağı olan öz-düzenlemeli özellik haritası (self-organizing feature map-SOFM) [14], yüksek boyutlu verilerin görselleştirilmesi ve analizi için önerilmektedir. Tüm bu doğrusal olmayan yöntemler genel anlamda, düşük boyutlu uzayda temsil edilen yüksek boyutlu verilerdeki komşuluk ilişkilerini temel alır ve bu komşuluk ilişkilerini doğrusal olmayan yapılar üzerinde korumayı amaçlar.

Bu doğrusal ve doğrusal olmayan yöntemlere ek olarak çeşitli olasılık tabanlı algoritmalar geliştirilmiştir. Bu algoritmalar temel olarak veri noktaları arasında koşullu olasılıkları kullanır. Bu tür ilk yöntem olarak kabul edilebilecek olan rastlantısal komşu yerleştirme (stochastic neighbor embedding-SNE) [15], veri noktaları arasındaki koşullu olasılıklar olarak yüksek boyutlu mesafeleri temsil ederek benzerlikleri tanımlamaktadır. Komşuluk yapılarını koruyarak yüksek boyutlu uzaydan düşük boyutlu uzaya ikili farklılıklar (pairwise dissimilarities) olarak tanımlanan nesnelere yerleştirir. T-dağılımlı rastlantısal komşu yerleştirme (t-distribution stochastic neighbor embedding-t-SNE) [16], SNE'nin daha küçük boyutlu uzayda verilerin yerel yapısını koruyan ve önemli bir küresel yapı ortaya koyan bir varyasyondur. Yüksek boyutlu veri setini ikili bir benzerlik matrisine dönüştürerek oluşturulan benzerlik verilerini kullanır. Bu veriler SNE yönteminden farklı olarak simetriktir. Küresel rastlantısal komşu yerleştirme (spherical stochastic neighbor embedding-sSNE) [17], sabit bir kavis koordinat sistemi göz önünde bulundurularak, öklid olmayan bir katmanlı uzay içine veri katmayı amaçlayan başka bir SNE varyasyondur. Çok boyutlu yapay alan yerleştirme (multidimensional artificial field embedding-MAFE) [18], yüksek boyutlu alandaki komşuluk ilişkilerini korumak için minimum enerji konfigürasyonuna yol açan yeni bir boyut indirgeme çatısıdır. Mekanikteki kuvvet alanı yasasına dayanmaktadır. Optimum enerji çözümünü bulmak için toplam çekme ve itme potansiyeli fonksiyonlarının en aza indirilmesi olarak yorumlanır. Her bir yüksek boyutlu veri noktasının, daha düşük boyutlu bir uzayda bir haritaya sahip olduğu varsayılmaktadır. Optimum gömme haritası, her iki kuvvet alanı tarafından belirlenen grafiğin denge durumuna karşılık gelmektedir. Tüm bu yöntemler, doğrusal olmayan katmanlı uzay verisinde bir koordinat temsili arar ve hiperspektral veriler dahil olmak üzere farklı uygulama alanlarının analizinde önemli bir rol oynar [19-21].

Ayrıca son yıllarda hiperspektral verilerin temsilini iyileştirmek ve yorumlamak için spektral verilere ek olarak uzamsal verileri kullanan derin öğrenme yaklaşımları da kullanılmaktadır [22-24]. Sınıflandırma, hiperspektral görüntüdeki spektral özellikler ile uzamsal özelliklerin birleştirilmesiyle yapılmaktadır [25]. Tüm bu yaklaşımlar etiketlenmiş veriler ve uzun eğitim sürelerine ihtiyaç duymaktadır. Bu çalışmalar incelendiğinde derin öğrenme tabanlı yöntemlerin sınıflandırma doğruluğunun doğrudan iyi bir manifold yerleştirmeye bağlı olduğu görülmektedir [26].

Algoritma verimliliğindeki ve belirli veri uygulamalarındaki farklılıklara rağmen tüm bu boyut indirgeme ve manifold öğrenme teknikleri, sınıflandırıcı girdi özelliklerinin azaltılması, daha iyi sıkıştırma ve daha iyi görselleştirme gibi bazı özellikleri miras almaktadır. Her manifold yerleştirme tekniği, doğrusal olmayan katmanlı uzay verisinde bir koordinat temsili bulmaya çalışmaktadır. Çeşitli doğrusal olmayan manifold öğrenme teknikleri üzerinde yürütülen son karşılaştırmalı çalışmalar [26-28], mevcut yöntemlerin farklı veri setleri üzerindeki kararsız yapılarını ortaya koymaktadır. Ayrıca yöntemler performanslarını etkileyen çok sayıda parametre içermektedir ve parametrelerin seçimi deneysel olarak yapılmaktadır. Bu durum kompleks verilerde büyük bir sorun olarak karşımıza çıkmaktadır. Birçok gerçek dünya nesnesinde bulunan çoklu manifoldları işlemede mevcut yöntemler yetersiz kalmaktadır [29-32]. Ayrıca birçok popüler algoritma, verilerin görselleştirilmesinde kalabalık (crowding) problemi olarak bilinen sorundan etkilenmektedir. Uzaktan algılamada kalabalık problemi, veri noktalarının, yüksek boyutlu uzaydaki verilerin daha düşük boyutlu uzaya yerleştirilmesi sırasında üst üste gelmesi durumudur [16,18]. Bu durum özellikle kompleks verilerde araştırmacıları zorlayan en önemli hususlardan biri olarak karşımıza çıkmaktadır.

Tüm bu manifold öğrenme yöntemlerinin belirli veri uygulamalarına bağlı avantaj ve dezavantajları bulunmaktadır. Boyut indirgeme uygulamalarında en yaygın yöntemlerden olan t-SNE yöntemi hızlı bir algoritmaya sahiptir. Ancak bu algoritma güncel verilerde yeterli sınıflandırma doğruluğu verememektedir [27]. Bunun dışında MAFE yöntemi, sınıflandırma doğruluğu ve verilerin görselleştirilmesinde oldukça başarılıdır. Yine bu yöntemde veri uygulamalarına bağlı olarak parametre belirleme sorunu bir dezavantaj olarak karşımıza çıkmaktadır [18,26,27]. Son yıllarda en yaygın

yaklaşımlardan derin öğrenme tabanlı algoritmalar ise hiperspektral veriler için doğrudan iyi bir ön-işlem gerektirmektedir. Ayrıca yüksek çalışma zamanı ve bellek maliyetleri ortaya çıkarmaktadır [26].

Bu tez çalışması kapsamında, 3 farklı akademik çalışma gerçekleştirilmiştir. Ayrıca kuvvet alanı formülasyonuna dayanan geliştirilmiş 2 farklı yöntem önerilmektedir. İlk çalışmada mevcut yöntemlerin performansları analiz edilmektedir. Popüler hiperspektral veriler kullanılarak literatürde bulunan doğrusal olmayan manifold öğrenme teknikleri karşılaştırmalı olarak incelenmektedir. Diğer çalışmalara da zemin oluşturan bu çalışma ile MAFE algoritmasının güçlü yönleri ortaya koyulmakta ve her iki veri seti için mevcut yöntemlere kıyasla daha yüksek sınıflandırma doğruluğu elde edilmektedir. İkinci çalışmada, hiperspektral görüntülerin sınıflandırılması ve görselleştirilmesi için optimizasyon tabanlı bir manifold öğrenme yöntemi önerilmektedir. Önerilen yöntem MAFE algoritmasının parametre belirleme sorununa çözüm üretmektedir. MAFE algoritmasında yer alan ve deneysel olarak belirlenen parametreler bu çalışmada optimizasyon tabanlı bir yaklaşımla çözülmektedir. Mevcut yöntemlere göre daha yüksek sınıflandırma doğruluğu ve görselleştirme sağlanmaktadır. Kalabalık problemine de başarılı bir çözüm sunulmaktadır. Ayrıca popüler derin öğrenme yöntemleri ile önerilen yöntem karşılaştırmalı olarak analiz edilmektedir. Karşılaştırmalarda literatürdeki en popüler 5 hiperspektral veri seti kullanılmaktadır. Üçüncü ve son çalışmada ise hiperspektral görüntülerin sınıflandırılmasında adaptif ağırlıklandırılan bir manifold yerleştirme yaklaşımı önerilmektedir. Önerilen ikinci yöntem t-SNE ve MAFE algoritmalarının güçlü yönlerini birleştiren hibrit bir yaklaşımdır. Benzer şekilde bu çalışmada da mevcut yöntemlere göre daha yüksek sınıflandırma doğruluğu ve daha iyi görselleştirme sağlanmaktadır. Yapılan tüm çalışmalar için popüler hiperspektral veri setleri üzerinde karşılaştırmalı analizler sunulmaktadır. Çalışmaların tümünde yapılan analizler tablo ve grafiklerle ayrıntılı olarak deneysel çalışmalar bölümünde anlatılmaktadır.

Hazırlanan bu çalışma, 7 ana bölümden oluşmaktadır. Birinci bölümde boyut indirgeme ve manifold öğrenme teknikleri ile ilgili literatür taraması yer almaktadır. En yaygın kullanılan boyut indirgeme yöntemleri ile ilgili temel bilgiler sunulmaktadır. Ayrıca bu tekniklerin avantaj ve dezavantajları kısaca

yorumlanmaktadır. Bunlarla birlikte tez kapsamında yapılan çalışmalar ve önerilen yöntemler hakkında açıklamalar verilmektedir.

İkinci bölümde uzaktan algılama ve hiperspektral görüntüleme hakkında bilgiler verilmektedir. Uzaktan algılama sürecinde verilerin nasıl elde edildiği, bileşenleri ve aşamaları, elektromanyetik spektrum ve uyduların bize sunduğu görüntü türleri bu bölüm içerisinde açıklanmaktadır. Ayrıca hiperspektral görüntüleme ve görüntü analizi hakkında bilgiler yer almaktadır.

Üçüncü bölüm, boyut indirgeme ve manifold öğrenme yöntemleri hakkında bilgiler içermektedir. Bu bölümde boyut indirgeme süreci ve boyut indirgeme sürecinde kullanılan band seçimi ve özellik çıkarımı gibi teknikler ayrıntılı olarak açıklanmaktadır. Ayrıca manifold kavramı ve manifold öğrenme yaklaşımı hakkında bilgiler verilmektedir.

Dördüncü bölüm, literatürde en yaygın bilinen ve deneysel çalışmalarda da kullanılan yöntemler hakkında ayrıntılı bilgiler içermektedir. Doğrusal ve doğrusal olmayan boyut indirgeme yöntemleri formülasyonları ile ayrıntılı olarak bu bölümde sunulmaktadır.

Beşinci bölümde boyut indirgeme çalışmalarında kullanılan optimizasyon tabanlı yaklaşımlar anlatılmaktadır. Bu bölüm, özellikle hiperspektral görüntüler üzerinde uygulanan optimizasyon yöntemleri hakkında bilgiler içermektedir. Ayrıca tez çalışması kapsamında kullanılan parçacık sürü optimizasyonu yöntemi ile ilgili bilgiler de bu bölümde verilmektedir. Bunlara ek olarak, tez çalışması kapsamında tamamlanan iki ayrı çalışmada önerilen algoritmaların detaylı açıklamalarına yer verilmektedir. Önerilen yöntemlerin formülasyonları, algoritmaları ve ayrıntılı açıklamaları bu bölümde yer almaktadır.

Altıncı bölümde deneysel çalışmalar yer almaktadır. Kullanılan veri setleri ile ilgili bilgi ve ön işlem adımları bu bölümde anlatılmaktadır. Ayrıca yapılan üç farklı akademik çalışma ayrı ayrı bu bölümde ele alınmaktadır. Elde edilen tüm bulgular çalışmaların amacına uygun şekilde yorumlanarak sunulmaktadır.

Yedinci ve son bölümde, yapılan çalışmalar ile ilgili tüm analizler ve karşılaştırmalar yorumlanmaktadır. Ayrıca gelecek çalışmalar için bazı öneriler yer almaktadır.



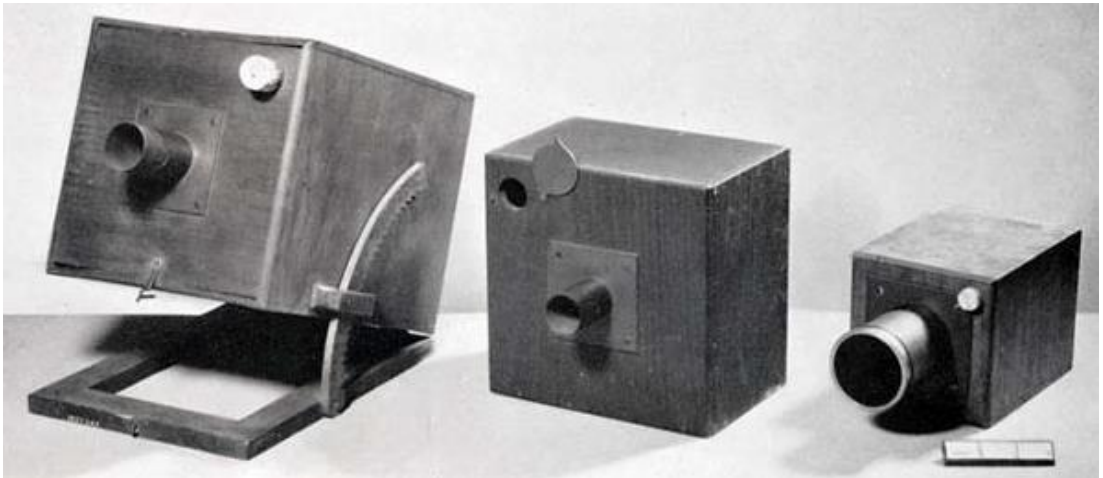
## BÖLÜM 2

### UZAKTAN ALGILAMA VE HİPERSPEKTRAL GÖRÜNTÜLEME

Gelişen uydu ve görüntüleme teknolojileri ile uzaktan algılama ve hiperspektral görüntüleme oldukça önemli bir araştırma konusu olmuştur. Hiperspektral görüntüler günümüzde savunma sanayinde, zirai uygulamalarda, atmosferik olaylarda ve jeodezi biliminde yaygın bir kullanıma sahiptir. Pankromatik hava fotoğrafları ile başlayan uzaktan algılama çalışmalarında artık tüm dünyada olduğu gibi ülkemizde de uydu görüntüleri ve sayısal hava fotoğrafları oldukça yoğun bir şekilde kullanılmaktadır. Bu bölümde Uzaktan algılamanın tarihçesi ve uzaktan algılama süreçleri ile hiperspektral görüntüleme konuları açıklanmaktadır [33].

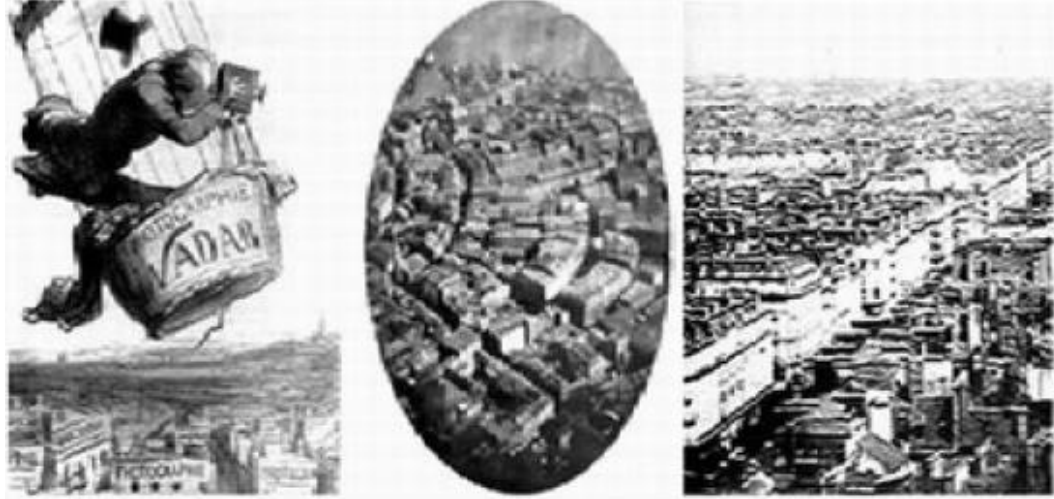
#### 2.1. UZAKTAN ALGILAMA

Uzaktan Algılama, bir temas olmaksızın, algılayıcı sistemleri kullanarak yeryüzü hakkında bilgi edinme bilimidir. Uzaktan algılama kavramı olarak ilk kez 1960 yılında kullanılmıştır. Ancak uzaktan algılamanın ilk örnekleri 19.yüzyıl başlarına dayanmaktadır. Şekil 2.1’de kullanılan ilk kamera örnekleri yer almaktadır.



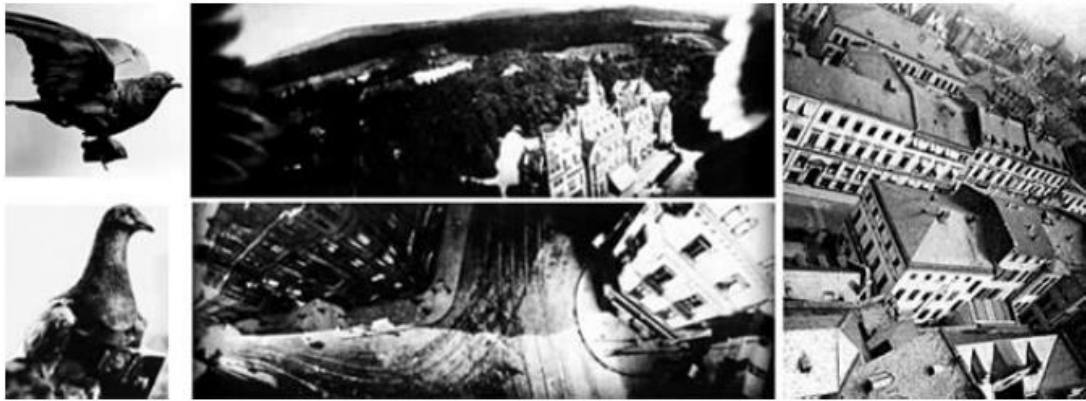
Şekil 2.1. 1840’lı yıllarda kullanılan kameralar [34].

Fotoğrafçılığın 1800'lü yıllarda keşfi ile ilk uzaktan algılama örnekleri elde edilmiştir. 1832'de ilk kez fotoğraf kavramı kullanılmış, ilk renkli fotoğraf ise 1861 yılında üretilmiştir. Sonraki yıllarda gelişen teknoloji ve yöntemler ile günümüz modern fotoğrafçılığı ortaya çıkmıştır.



Şekil 2.2. İlk hava fotoğrafı örnekleri [35].

Fotoğrafın keşfinden sonra güvercinler ve balonlar kullanılarak ilk hava fotoğrafı örnekleri görülmektedir. Bilinen ilk hava fotoğrafı ise Nadar adıyla anılan Fransız bir fotoğrafçı tarafından çekilmiştir. Şekil 2.2'de Nadar'ın bir karikatürü ve çektiği ilk fotoğraf örnekleri görülmektedir.

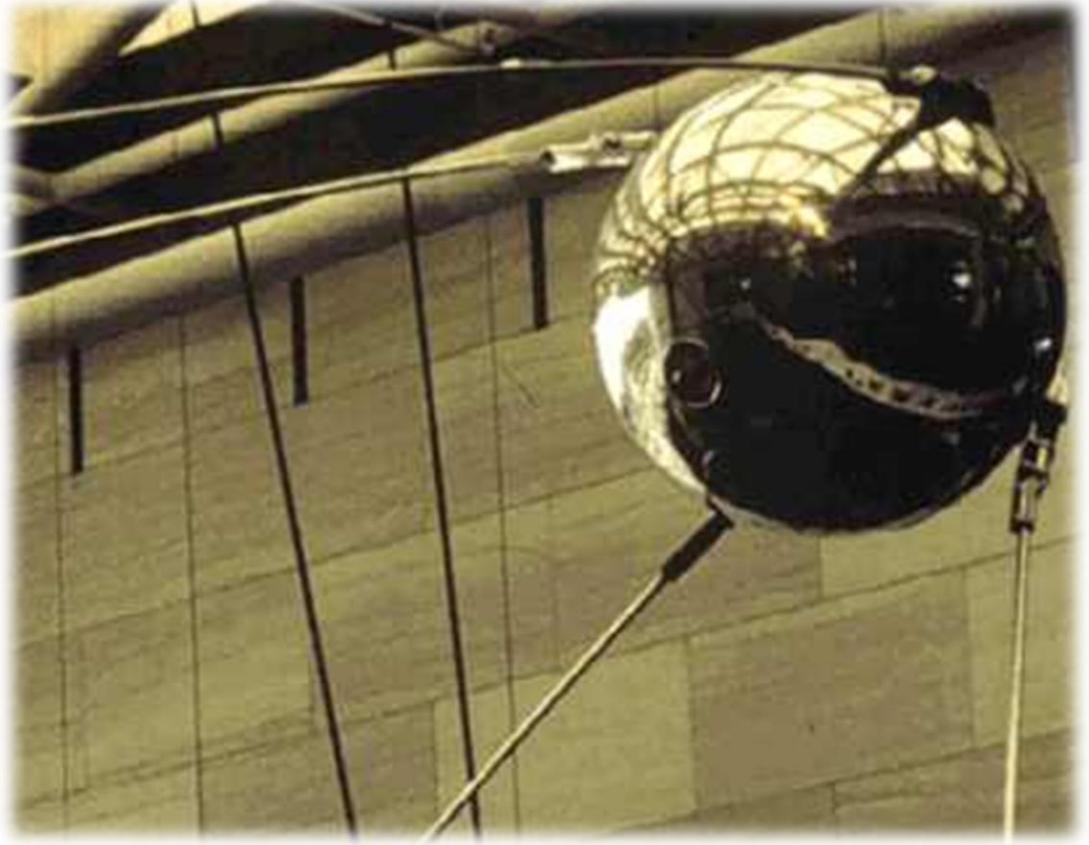


Şekil 2.3. Fotoğrafçı güvercinler ve çektikleri fotoğraf örnekleri [36].

1903 yılında güvercinlerin taşıyabileceği bir otomatik kamera ile havadan fotoğraf örnekleri çekilmiştir. O yıllarda özellikler orduda havadan bilgi edinmek amacıyla bu

güvercinlerin kullanımı görülmektedir. Bu güvercinler ve çektikleri bazı fotoğraf örnekleri Şekil 2.3'te verilmektedir.

Uzaktan algılama çalışmaları 1800'ler de fotoğrafın icadına kadar uzanmakla birlikte uzaktan algılama teknolojisinin daha fazla gelişmesini sağlayan en önemli unsur uzay çağıdır. Uzay çağının 1957'de Sovyet Rusya'nın fırlattığı, Şekil 2.4'te görseli bulunan Sputnik-1 uydusu ile başladığı genel olarak kabul edilmektedir [37]. Sonrasında ise birçok dünya ülkesi tarafından ABD ve Rusya önderliğinde hava gözlem amaçlı çok sayıda uydu fırlatılmıştır.



Şekil 2.4. Sputnik-1 uydusu [38].

Günümüzde uzayda haberleşme, gözlem ve askeri amaçlı otuzdan fazla ülkeye ait yüzlerce uydu yörüngemizde aktif konumdadır. Türkiye'de haberleşme ve gözlem amaçlı uyduları ile 1994 yılından bu yana aktif olarak bu ülkeler arasında yer almaktadır. Çizelge 2.1'de Türkiye'nin uzaya fırlattığı uydulara ait bilgiler verilmektedir.

Çizelge 2.1. Türkiye'nin sahip olduğu yer gözlem uyduları.

Uydu	Amacı	Durumu	Yıl	Hizmet Ömrü (Yıl)
TÜRKSAT 1B	Haberleşme	Pasif	1994	12
TÜRKSAT 1C	Haberleşme	Pasif	1996	14
TÜRKSAT 2A	Haberleşme	Pasif	2001	15
TÜRKSAT 3A	Haberleşme	Aktif	2008	-
RASAT	Yer Gözlem	Aktif	2011	-
Göktürk-2	Yer Gözlem	Aktif	2012	-
TÜRKSAT 3USat	Haberleşme	Aktif	2013	-
TÜRKSAT 4A	Haberleşme	Aktif	2014	-
TÜRKSAT 4B	Haberleşme	Aktif	2015	-
Göktürk-1	Yer Gözlem	Aktif	2016	-
TÜRKSAT 5A	Haberleşme	Aktif	2021	-
IMECE	Yer Gözlem	Proje	2022	-

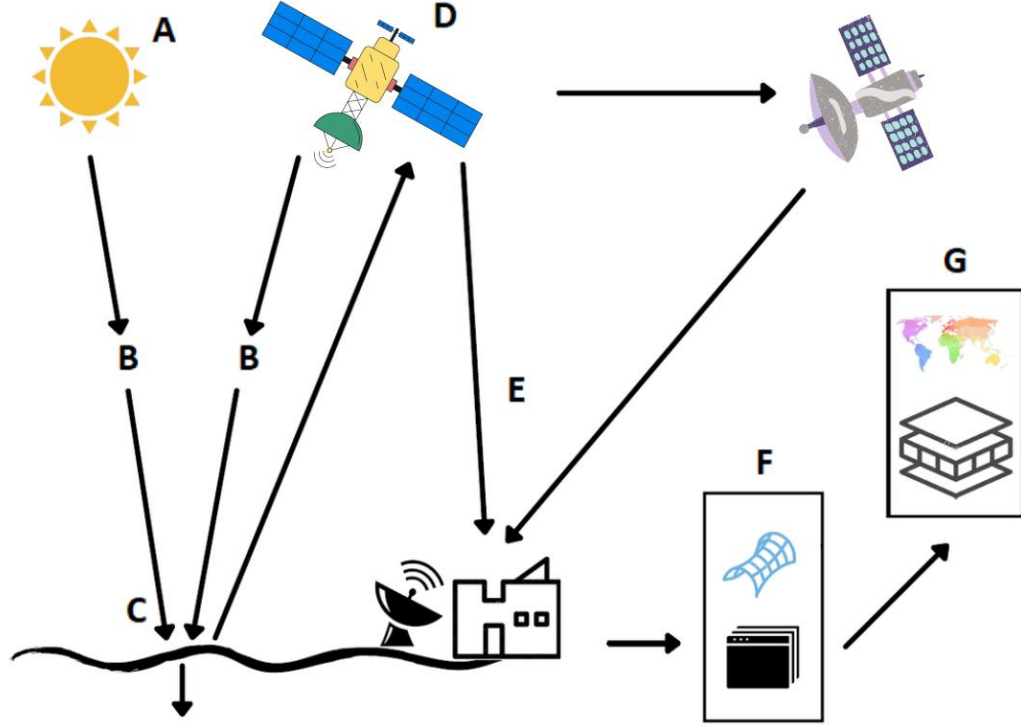
Sonuç olarak uzay çağı ile uzaktan algılama günümüzdeki anlamına daha yakın bir duruma gelmiştir ve artık uzaktan algılama için uydu platformları ve yüksek çözünürlüklü gelişmiş algılayıcılar kullanılır hale gelmiştir. Ayrıca son yıllarda birçok ticari ve akademik çalışmada kameralı drone ve insansız hava aracı kullanımı giderek artmaktadır.

### 2.1.1. Uzaktan Algılamanın Bileşenleri ve Aşamaları

Uzaktan Algılama süreci doğal veya yapay bir enerji kaynağından gelen ve yansıyan enerji dalgalarının algılanması ile başlamaktadır. Enerjinin bir bölümü dünya yüzeyi tarafından yutulmaktadır. Bu enerji dalgaları çeşitli platformlar üzerinden toplanarak veri iletim alanlarında kaydedilirler. Daha sonra bu bilgiler analiz edilir ve yorumlanırlar. Yorumlanan ve anlamlandırılan bu bilgiler sonrasında muhtelif amaçlarla kullanılmaktadır. Bu süreç, uzaktan algılanan nesnelere farklı açılardan görmemizi ve fiziksel olarak bulunmadığımız bir ortamda keşifler yapabilmemizi sağlamaktadır.

Uzaktan algılama sürecine ait bileşenler ve bunların ilişkisi Şekil 2.5'te temsili olarak sunulmaktadır. Uzaktan algılamanın gerçekleşebilmesi için bu yedi bileşeninde bir

araya gelmesi gerekmektedir. Tüm bu bileşenler uzaktan algılama süreci içerisinde birbirlerini tamamlamaktadır.



Şekil 2.5. Uzaktan algılama süreci [39].

*Enerji Kaynağı (A):* Uzaktan algılama sürecinin ilk bileşeni enerji kaynağıdır. Bu kaynak güneş gibi doğal bir kaynak olabileceği gibi radar gibi yapay enerji üreten kaynaklardan nesnelere ışınım gönderen sistemlerde olabilmektedir.

*Işınım ve Atmosfer (B):* Enerji, kaynaktan hedefe doğru yayılırken atmosfer etkisine maruz kalmaktadır. Aynı durum hedeften yansıyan enerji algılayıcıya ulaşırken de gerçekleşmektedir. Atmosfer etkisi bu sebeple enerjinin belirli oranda değişimine yol açar.

*Hedef ve Etkileşim (C):* Atmosferi geçen enerji yüzeye ulaştıktan sonra etkileşime girer. Enerjinin bir kısmı yüzey tarafından soğurulur. Yansıyan enerji ise yüzey veya nesnelere hakkında bilgiler içermektedir. Uzaktan algılama sürecinde önemli olan kısım bilgiyi içeren bu yansıyan enerji dalgasıdır.

*Yansıyan Enerji (D):* Atmosferden geçen enerji hedeften yansıdıktan sonra algılayıcılar tarafından toplanır ve kaydedilir. Bu enerji yüzeyden düzgün veya dağınık yansıyabilmektedir.

*Veri İletimi (E):* Algılayıcıların kaydettiği elektromanyetik enerji verisi bu aşamada işlenecekleri ve görüntü haline getirilecekleri veri istasyonlarına iletilirler.

*Analiz ve Yorumlama (F):* İşlenen görüntüler üzerindeki hedefler ve bu hedeflere ait karakteristik yapılar. Analiz ve yorumlama aşamasında işlenmekte ve anlamlandırılmaktadır.

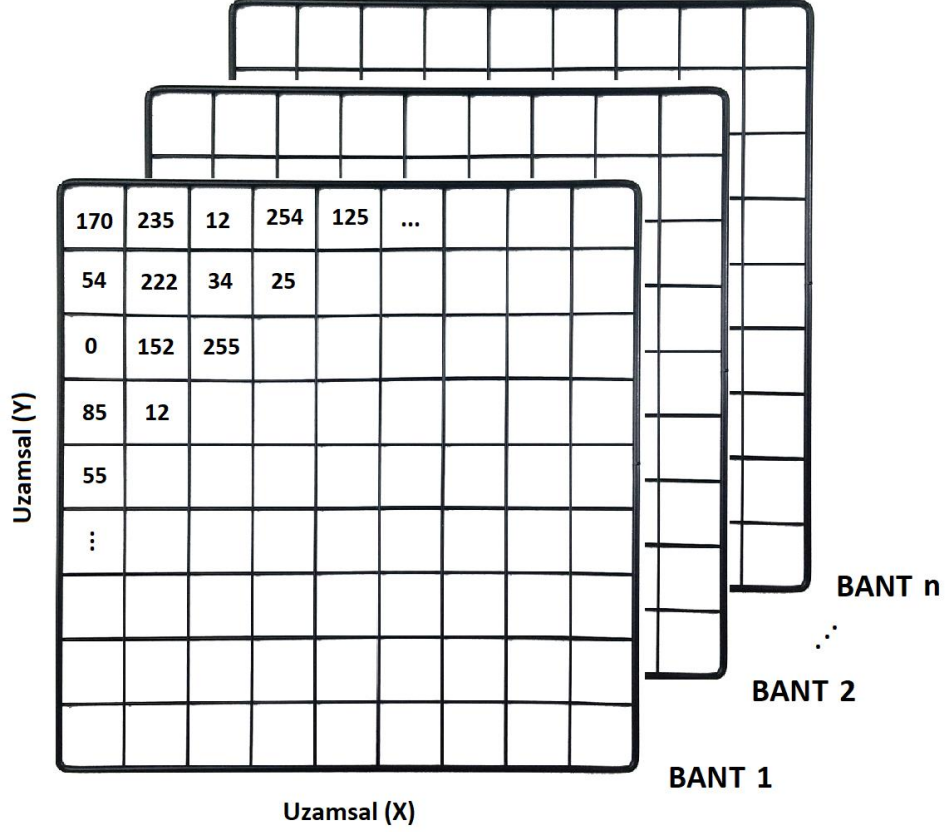
*Uygulama (G):* Uzaktan algılama sürecinin son aşamasında ise elde edilen bilgiler kullanım alanlarına ve amaçlarına bağlı olarak çeşitli uygulamalara entegre edilmektedir.

### **2.1.2. Uzaktan Algılamada Algılayıcılardan Elde Edilen Çözünürlük Türleri**

Uzaktan algılama için görüntü satır ve sütunlardan oluşan ızgara (grid) benzeri bir yapıda elde edilen veridir. Bu ızgaranın büyüklüğü bize çözünürlük kavramını vermektedir. Izgaradaki en küçük veri birimi ise pikseldir. Belirli bir konumdaki yeryüzü nesnelere yansıyan elektromanyetik enerji değeri pikseller içerisindeki sayısal değerlerle temsil edilmektedir. Örnek çok bantlı bir görüntü ızgarası Şekil 2.6'da gösterilmektedir.

Görüntüdeki uzamsal boyutlar veya satır ve sütunlar görüntünün çözünürlüğü ile alakalıdır. Örneğin 1024x768 çözünürlüğe sahip olmak her karede 1024 satır ve 768 sütundan oluşan bir veri matrisi elde edilmesi anlamına gelmektedir. Her veri noktası (piksel) içerisindeki sayısal değerler ise o görüntüye ait verilerdeki derinliği ve hassasiyetin ölçüsünü vermektedir. Örneğin 8 bit ( $2^8$ ) bir görüntü 0-255 aralığında değer üretebilir. Bu değer arttıkça görüntülerdeki her bir piksel değeri arasındaki ayrım daha kolay hale gelecektir. Son olarak görüntüde yer alan bantlar ise her bir veri noktası için farklı dalga boylarındaki verinin saklanabileceğini ifade etmektedir. Bu

da her veri noktası için birden fazla bilginin görüntü matrisinde ifade edilebileceğini göstermektedir.



Şekil 2.6. Çok bantlı görüntü örneği.

Bir görüntüye ait mekansal, radyometrik, zamansal ve spektral bileşenler, yüzeyden yansıyan elektromanyetik enerji ile gerçek dünya nesnelere hakkında analiz ve yorum yapmak için gerekli bilgileri sağlamaktadır. Uzaktan algılamada, algılayıcılar tarafından kaydedilen 4 farklı çözünürlük tipi bulunmaktadır. Bunlar:

- Mekansal çözünürlük,
- Radyometrik çözünürlük,
- Zamansal çözünürlük,
- Spektral çözünürlüktür.

**Mekansal çözünürlük**, konumsal veya geometrik çözünürlük olarak da adlandırılmaktadır. Her bir pikselin temsil ettiği yüzey alanının ölçüsü olarak



tanımlanabilmektedir. Mekansal çözünürlük ne kadar düşük olursa hassasiyet o kadar fazla anlamına gelmektedir. Örneğin 1 metre mekânsal çözünürlük 10 metre mekânsal çözünürlüğe kıyasla çok daha fazla detay vermektedir. Mekansal çözünürlükler kendi içerisinde; düşük çözünürlüklü (100 metre ve üzeri), orta çözünürlüklü (10-100 metre arası) ve yüksek çözünürlüklü (10 metre altı) olmak üzere üç sınıfa ayrılabilirler. Son yıllarda 1 metre altı yüksek mekansal çözünürlüğe sahip uydular giderek artmaktadır.



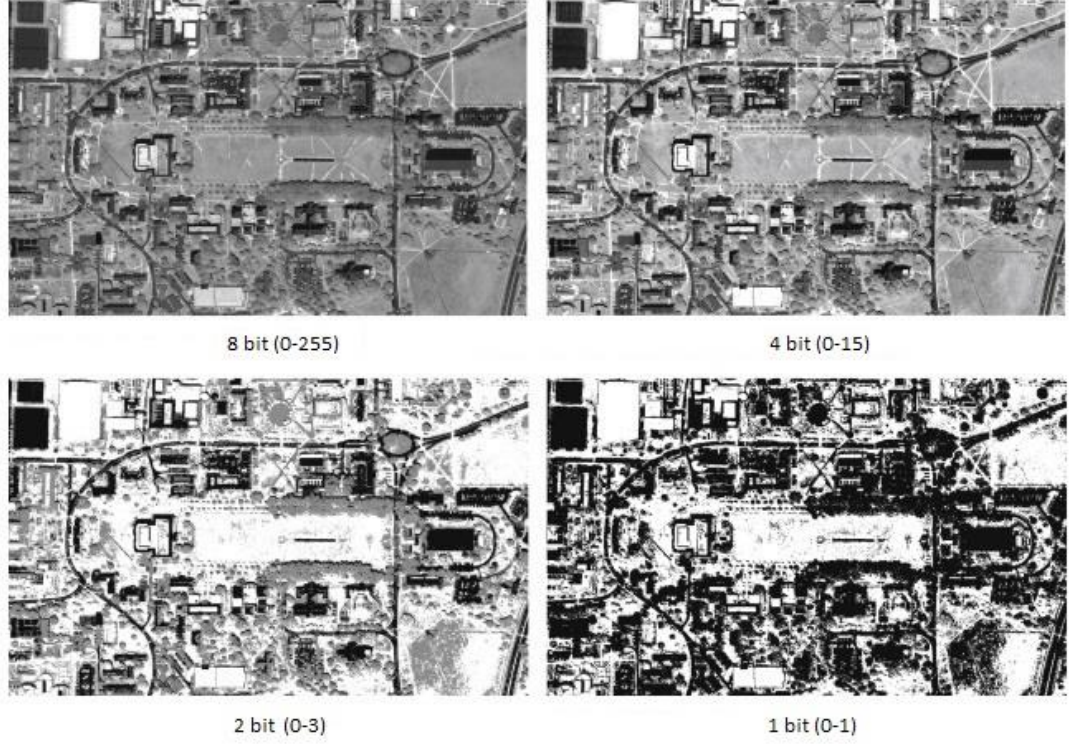
Şekil 2.7. Aynı bölgeye ait farklı mekansal çözünürlüğe sahip görüntüler [40].

Şekil 2.7’de Landsat 7 uydusu ile elde edilmiş farklı mekansal çözünürlüğe sahip görüntüler yer almaktadır. Mekansal çözünürlük değeri arttıkça görüntüdeki bozulma net bir şekilde görülmektedir. Sol üst köşedeki ilk görüntüde her piksel 1 metrekarelik bir alanı temsil ederken, sağ altta yer alan görüntüde 1 piksel 900 metrekarelik bir alanı temsil etmektedir.

**Radyometrik çözünürlük**, elektromanyetik enerji dalgalarının ölçümündeki hassasiyeti temsil etmektedir. Başka bir ifadeyle cisimlerden yansıyan enerjideki farklılıkların ayırt edilebilmesi mümkün olan gri tonu sayısını ifade etmektedir. Radyometrik çözünürlük bit cinsinden tanımlanmaktadır ve  $2^n$  formatında gösterilmektedir. Örneğin; 16 bit radyometrik çözünürlüğe sahip bir algılayıcı,  $2^{16} =$



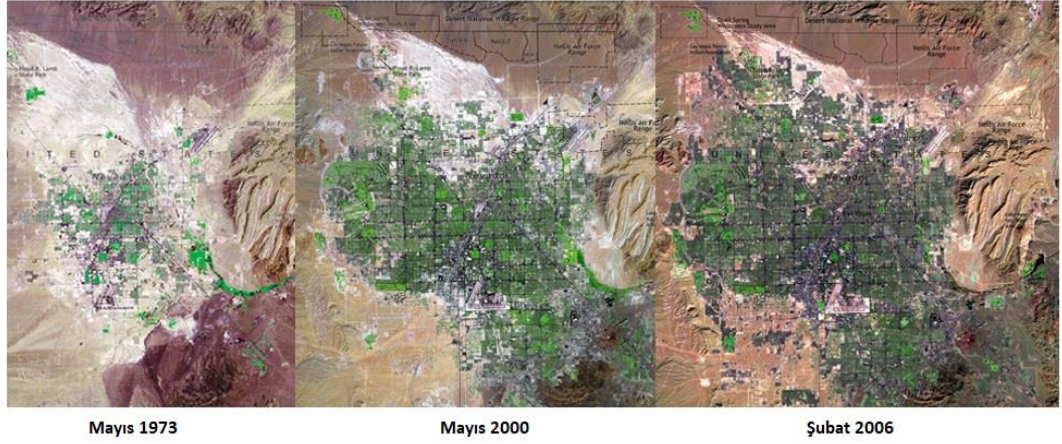
65536 farklı parlaklık değeri veya gri tonu ölçebilmektedir. Radyometrik çözünürlük arttıkça görüntülerin analizi ve yorumlanması da kolaylaşmaktadır.



Şekil 2.8. Aynı bölgeye ait farklı radyometrik çözünürlüğe sahip görüntüler [41].

Şekil 2.8’de aynı bölgeye ait farklı radyometrik çözünürlüklü algılayıcılardan elde edilmiş görüntüler yer almaktadır. Sağ alt kısımda yer alan görüntü 1 bit radyometrik çözünürlüğe sahiptir. Bu sebeple görüntüde sadece 2 farklı parlaklık değeri ölçülebilmektedir. Başka bir ifadeyle görüntü Siyah-Beyaz’dır. Radyometrik çözünürlük arttıkça bölgeye ait görüntüdeki ayrıntılar çok daha net şekilde fark edilmektedir.

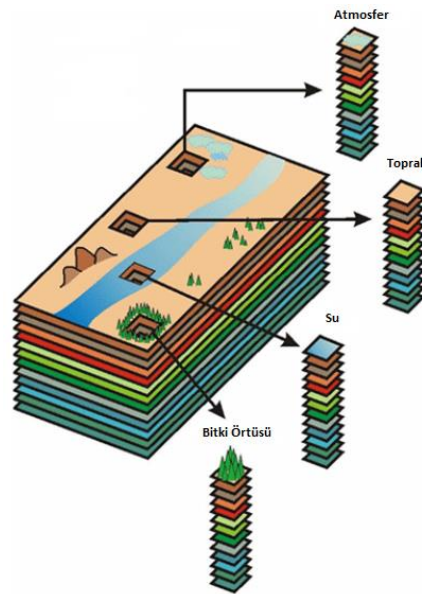
**Zamansal çözünürlük**, uzaktan algılayıcı sistemlerin aynı bölgeyi görüntüleme sıklığı olarak düşünülebilir. Uyduların yüksekli ve yörünge düzenlerine bağlı olarak zamansal çözünürlük değişiklik göstermektedir. Örneğin; Landsat uyduları aynı bölgeyi 16 gün aralıklarla görüntüleyebilmektedir. Şekil 2.9’da Las Vegas bölgesine ait farklı zamanlarda elde edilmiş görüntüler verilmektedir. Görüntüler bölgenin büyümesi, nüfus artış hızı, bitki örtüsündeki değişim ve ulaşım ağı gibi birçok bilgiyi barındırmaktadır.



Şekil 2.9. Farklı zamanlarda aynı bölgeye ait görüntüler [42].

Zamana bağlı elde edilen görüntüler, sayısal arazi modellerinin üretilmesi, doğal afetlerin izlenmesi, arazi kullanımının takibi ve planlanması gibi birçok alanda kullanılmak üzere önemli bilgiler vermektedir.

**Spektral çözünürlük**, algılayıcıların elektromanyetik spektrumun birden fazla bölgesinden aynı anda veri alabilmesini temsil etmektedir. Spektral çözünürlüğün yüksek olması her bir bantın algıladığı dalga boyu aralığının ne kadar küçük olduğu ile ölçülmektedir. Farklı bir ifadeyle, spektrum ne kadar dar ve çok parçaya ayrılırsa spektral çözünürlük o kadar yüksek olur.



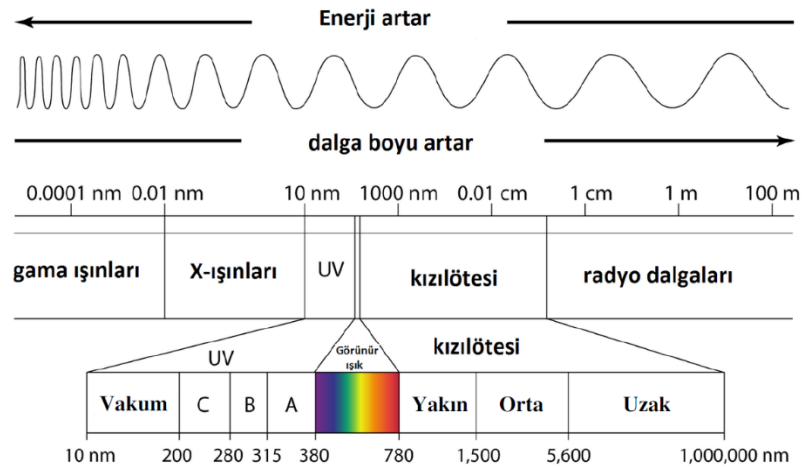
Şekil 2.10. Çeşitli dünya nesnelere ait spektral veri gösterimi [43].

Şekil 2.10'da farklı tür dünya nesnelere ait veri noktalarının spektral veri gösterimi temsili olarak gösterilmektedir. Burada toprak, su veya bitki örtüsüne ait bir veri noktasının birden fazla karakteristik özelliği ölçülmektedir. Her bir özellik elektromanyetik spektrumda bulunan dalga boylarının farklı bölgelerinden elde edilmektedir.

## 2.2. ELEKTROMANYETİK SPEKTRUM

Yeryüzündeki nesnelere enerji kaynağından gelen ışınımın belirli bir kısmını yansıtır, bir bölümünü emer ve bir bölümünü de geçirirler. Yansıtılan enerji elektromanyetik dalgalar şeklinde algılayıcılara nesnelere hakkında bilgiler sağlar [44]. Bu enerji dalgası iki farklı şekilde algılanabilir. Bunlardan ilki pasif algılayıcılar tarafından toplanan verilerdir. Pasif algılayıcılar güneş gibi doğal enerji kaynaklarından gelen ve hedef yüzeylerden yansıyan bilgiyi toplamaktadır. İkinci algılama ise aktif algılayıcılar tarafından elde edilmektedir. Bu tür algılayıcılar kendi enerjisini kendi üreterek bu enerjiden yansıyan bilgileri toplamaktadır. Aktif algılayıcılar harici bir enerji kaynağına ihtiyaç duymazlar.

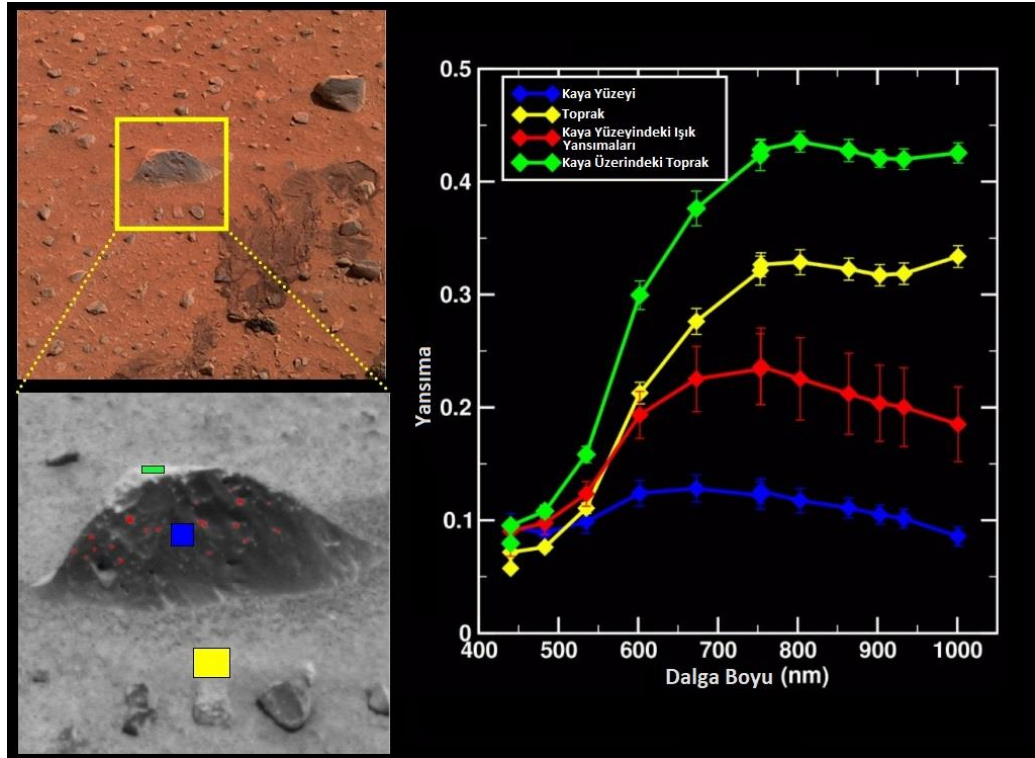
Uzaktan algılama sürecinde veri kaynaklarından çıkan enerji boşlukta dalga şeklinde yol alır. Bu enerji, sahip olduğu dalga boyunun büyüklüğüne göre farklı türlere ayrılmaktadır. Bu farklı enerji yayılımlarının tümüne elektromanyetik spektrum denilmektedir.



Şekil 2.11. Elektromanyetik spektrum [45].

Şekil 2.11’de verilen elektromanyetik spektrumda, en solda kısa dalga boyu ve yüksek enerjiye sahip gama ışınları vardır. En sağda ise uzun dalga boyu ve düşük enerjiye sahip radyo dalgaları mevcuttur. Bu dalga boyları soldan sağa doğru bir atomun boyutlarından bir şehrin büyüklüğüne kadar çok geniş bir uzunlukta yer almaktadır. Elektromanyetik spektrumda, insan gözünün görebildiği ve ilk keşfedilen kısmı ise 0,4 – 0,7 mikrometre arasında kalan küçük bir bölgeyi kapsamaktadır. Bu aralıktaki tüm dalga boyları insan gözü tarafından renk olarak algılanmaktadır.

Her cisim üzerine gelen ışınımı farklı oranlarda yansır, yutar ya da geçirir. Yeryüzündeki cisimlerin tanımlanabilmelerinin en önemli nedeni, spektral özelliklerinin değişiklik göstermesidir. Algılayıcıların tasarımı da bu değişiklikleri fark edecek ve istenilen ayrımları yapabilecek şekilde yapılır. Her spektral bant elektromanyetik spektrumun bir bölümünde duyarlıdır.



Şekil 2.12. Mars yüzeyine ait kaya parçasının spektral yansımaları grafiği [46].

Yeryüzündeki tüm materyaller benzersiz spektral özelliklere sahiptir. Başka bir ifadeyle, tüm nesnelere güneş ışığını veya yapay enerji kaynaklarından gelen elektromanyetik dalgaları farklı şekilde yansıtırlar. Şekil 2.12’de Mars yüzeyinde

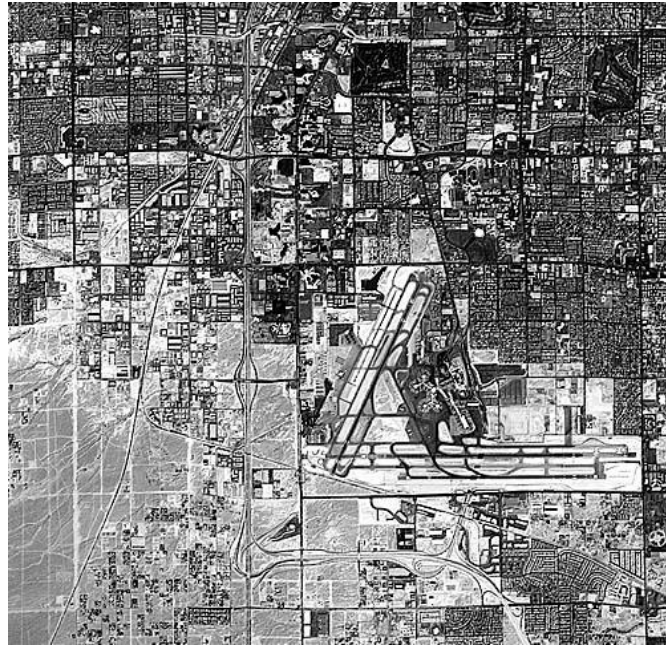


bulunan bir kayanın üzerindeki farklı spektral yansımalar gösterilmektedir. Kaya yüzeyi, kayanın üzerinden yansıyan ışık ve kayanın üzerinde bulunan toprağın hepsi elektromanyetik spektrumun farklı bir bölgesinde ayrılmaktadır. Uygun bölgenin seçilmesi ile birbirine bu kadar yakın materyaller bile spektral bilgileri sayesinde kolaylıkla ayrışabilmektedir.

### 2.3. GÖRÜNTÜ TÜRLERİ

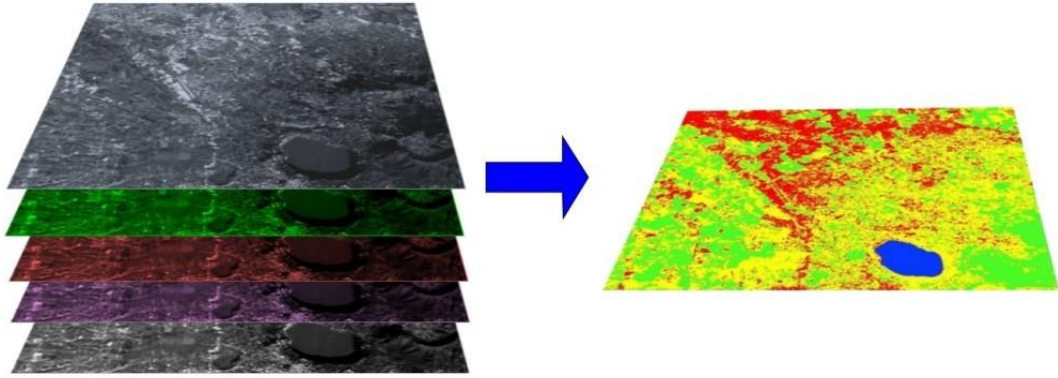
Uzaktan algılamada, elektromanyetik spektrumdaki ölçülen bant aralığına bağlı olarak uyduların bize sunduğu üç farklı görüntü türü bulunmaktadır. Bu görüntü türleri tek banttan oluşan Pankromatik görüntüler, birden fazla banttan oluşan multispektral görüntüler ve yüzlerce spektral bant bilgisi içeren hiperspektral görüntülerdir.

**Pankromatik görüntü**, elektromanyetik spektrumun geniş bir bölümünden yansıyan enerjiyi ölçebilen algılayıcılar tarafından elde edilir. Bu bölümlere genellikle bant adı verilir. Halihazırdaki birçok pankromatik algılayıcılar için bu tek bant genellikle spektrumun görünen ve yakın kızılötesi bölümünü kapsamaktadır. Pankromatik veriler siyah-beyaz ve gri tonlarda görüntü olarak tek bant üzerinden oluşturulurlar. Şekil 2.13'te Las Vegas bölgesine ait bir pankromatik uydu görüntüsü yer almaktadır.



Şekil 2.13. Pankromatik uydu görüntüsü [47].

**Multispektral görüntü**, elektromanyetik spektrumdaki birden fazla bantta ölçülen dijital algılayıcılarla elde edilmektedir. Örneğin; alıcıların bir bölümü görünür kırmızı yansıyan enerjiyi ölçerken, diğer bir grup yakın kızılötesi enerjiyi ölçer. İki ayrı algılayıcı dizilimi aynı dalga boyunun değişik bölümlerindeki enerjiyi ölçebilirler. Bu çoklu yansıma değerleri birleştirilerek renkli multispektral görüntüler Şekil 2.14'te de gösterildiği gibi elde edilmektedir.



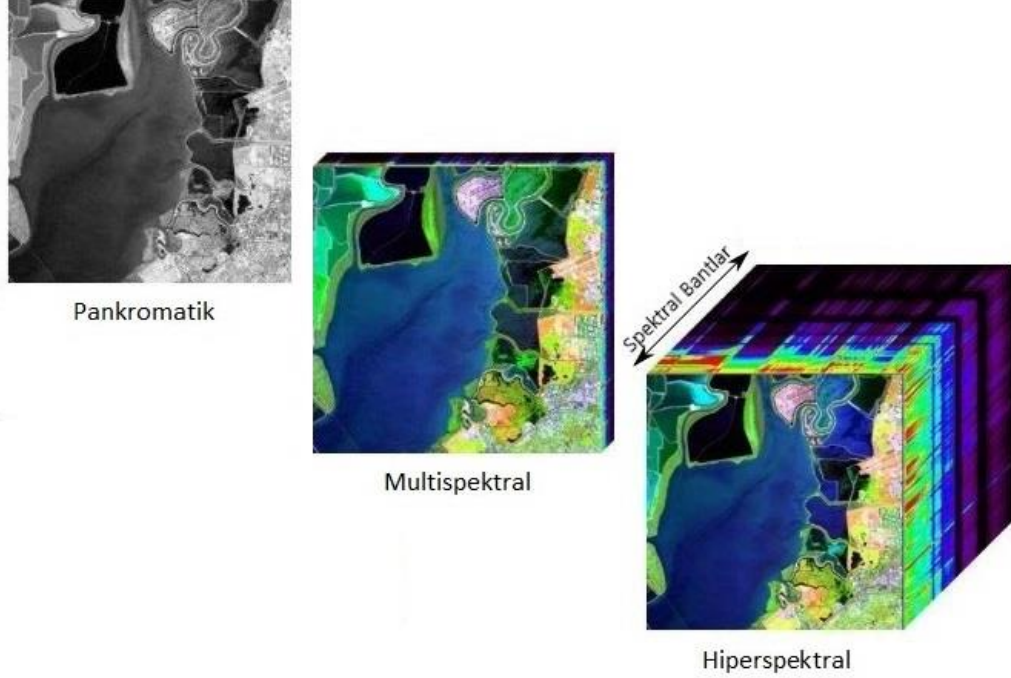
Şekil 2.14. Multispektral uydu görüntüsü [48].

**Hiperspektral görüntüler** de elektromanyetik spektrumun çok dar ve bitişik dalga boyu aralıklarında ve yüzlerce banttan görüntü elde edebilmektedir. Bu sayede görüntülerdeki yüzey özelliklerini çok daha hassas bir şekilde ayırabilmemizi sağlarlar.

İnsan gözü sadece elektromanyetik spektrumdaki görünür bant aralığındaki yansımaları görebilir. Hiperspektral görüntüler ise görünür bant aralığına ek olarak kızılötesi ve yakın kızılötesi bantlardan da ölçüm yapabilmektedir. Hiperspektral görüntüdeki her bir piksel elektromanyetik bantlardaki yansıma verilerinden oluşan vektörel bir ifade ile temsil edilirler. Oluşan bu vektörel ifade her bir piksele ait spektral imzayı vermektedir.

Hiperspektral görüntüler elde etmek için özelleştirilmiş hiperspektral algılayıcı kameralar kullanılır. Bu kameralar elektromanyetik spektrum bileşenlerini dalga boylarına göre ayrıştırma özelliğine sahiptir. Bu kameralar spektroskopi içerirler [49].

Spektroskopi, farklı malzemelerde meydana gelen ışık emisyonu veya yansıması ile ilgili bir bilimdir [50,51].



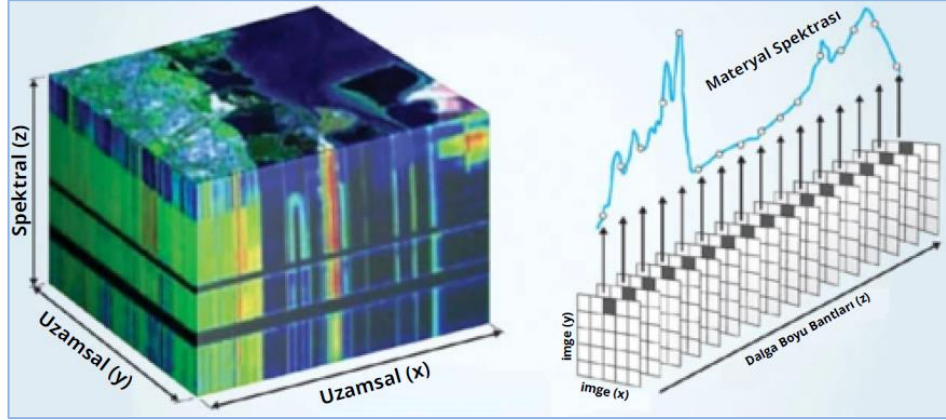
Şekil 2.15. Aynı bölgeye ait uydu görüntüleri [52].

Hiperspektral görüntüleri diğer görüntüleme türlerinden ayıran en önemli özelliği spektral bantlarıdır. Çok düşük dalga boyu aralığında (10-20 nm) yüzlerce bant (200+) üzerinden aldığı veriler ile birbirine çok yakın sınıfları ayırabilmesi ve hedef belirleme uygulamalarındaki başarısı ile önemli bir kaynak haline gelmiştir.

Hiperspektral veriler uzamsal iki boyut ve bu uzamsal boyutlardaki her bir piksel için elektromanyetik spektrumun farklı bantlarından yansıyan yüzlerce bilgi içerirler. Bu 3 boyutlu verilere hiperspektral veri küpü denilmektedir. Her piksel için spektral boyutta oluşan grafik ise Şekil 2.16'da görüldüğü üzere o piksele veya nesneye ait spektral imzayı vermektedir. Bu spektral imza veya materyal spektrası bize nesnelere sınıflandırmada büyük bir kolaylık sağlamaktadır.

Hiperspektral veriler, yeryüzüne ait materyaller hakkında detaylı bilgiler sağlamanın yanında bazı dezavantajları bulunmaktadır. Hiperspektral veri küpleri spektral

düzlemdeki yüksek boyutları sebebiyle görüntülerin işlenmesi aşamasına oldukça büyük hesaplama maliyetleri çıkarmaktadır. Verilerin elde edilmesinde özel, maliyeti yüksek algılayıcılar kullanılmaktadır ve verilerin analizi için güçlü sistemlere ihtiyaç vardır. Ayrıca bu verilerin depolanması da başlı başına çözülmesi gereken bir problemdir.



Şekil 2.16. Hiperspektral veri küpü [53].

Hiperspektral verilerin elde edilmesi için ortaya çıkan bu maliyetler çeşitli boyut indirgeme tekniklerinin gelişimine yol açmıştır. Boyut indirgeme ve kullandığı teknikler sonraki bölümde ayrıntılı olarak açıklanmaktadır.

## 2.4. HİPERSPEKTRAL GÖRÜNTÜLEME

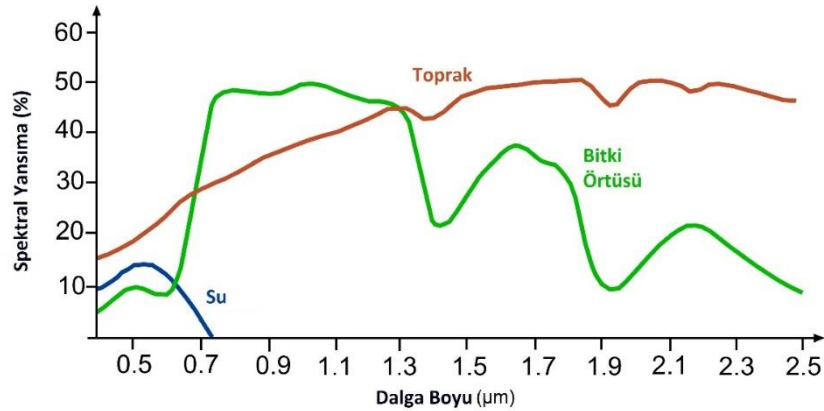
Uzaktan algılamada görüntü analizi üç temel aşamada gerçekleşmektedir. Bunlar sırasıyla görüntü oluşumundaki geometrinin ve radyometrinin anlaşılması, temel görüntü işleme teknikleri ile görüntünün dönüşümü ve istatistiksel örüntü çıkarım algoritmaları ile görüntünün analizi aşamalarıdır [37].

Hiperspektral görüntünün analizinde yine 3 temel uzayda veri küpünü analiz etmek gerekmektedir. Mekansal uzay, görüntü piksellerinin birbirleri ile olan geometrik ilişkilerini barındırmaktadır. Spektral uzay, enerji dalga boyuna ait piksellerin kendi iç dinamiklerine ait değişimleri vermektedir. Son olarak özellik uzayı ise görüntüdeki belirli bölgelere ait piksellerin vektörel gösterimle veri noktalarına dönüşümünü içermektedir [54].



Hiperspektral görüntüler 1970'li yılların başından itibaren kullanılmaya başlanmış olmasına rağmen, yaygın kullanımı geçtiğimiz son 10 yıla dayanmaktadır. Hiperspektral görüntü, spektral anlamda çok sayıda bilgi içeren, malzemeye duyarlı olduğundan dolayı birbirine benzeyen maddelerin ayrıştırılmasında ve maddelerin sınıflandırılmasında kullanılan özel bir görüntü çeşididir.

Çok geniş bir kullanım alanına sahip hiperspektral görüntüler, karakteristik yansıma spektrumlarına sahip çok çeşitli malzemeleri tespit etmek ve haritalamak için kullanılmaktadır. Örneğin, jeologlar tarafından mineral haritalama ve nem, organik içerik veya tuzluluk gibi toprak özelliklerini tespit etmek için kullanılmaktadır. Botanistler, bitki örtüsü türlerini belirlemek, bitki kimyasını incelemek ve bitki örtüsü stresini tespit etmek için hiperspektral görüntüleri başarıyla kullanmaktadır. Askeri alanda da kısmi bitki örtüsü altındaki askeri unsurların tespiti ve diğer birçok askeri hedef tespiti amacıyla hiperspektral görüntüleme kullanılmaktadır [37].



Şekil 2.17. Farklı malzemeler için elde edilmiş spektral imzalar [53].

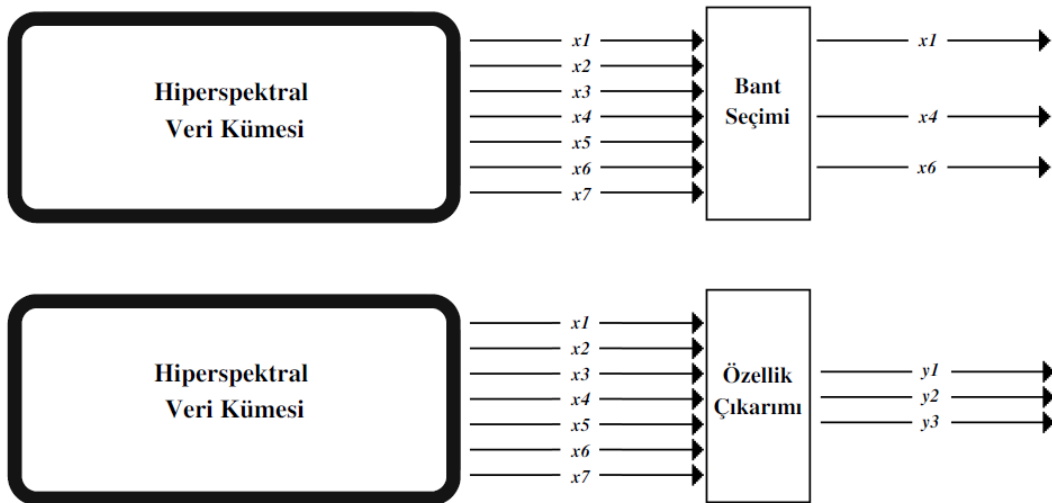
Hiperspektral görüntüleme, Şekil 2.17'deki gibi nesnelerin elektromanyetik spektrum üzerindeki yansıma özelliklerinden faydalanmaktadır. Her nesnenin kendine özgü olan bu yansıma özelliği bize nesnenin spektral imzasını vermektedir. Standart multispektral görüntü işleme tekniklerine kıyasla hiperspektral görüntüleme bize nesnelere hakkında çok daha detaylı bilgiler sağlar.

## BÖLÜM 3

### BOYUT İNDİRGEME VE MANİFOLD ÖĞRENME

Boyut indirgeme, büyük veri kümelerindeki verilerin içerisindeki önemli bilgileri kaybetmeden daha düşük boyutlara indirgenmesidir. Büyük veri kümelerinin analizi ve yorumlanması için bu oldukça önemli bir ön işlemdir. Özellikle yüzlerce yüksek çözünürlüklü banttan oluşan ve farklı nesnelere özelliklerini hassas bir şekilde yansıtan hiperspektral görüntülerin yorumlanabilmesi için boyut indirgeme oldukça önemli bir çalışma alanıdır.

Hiperspektral görüntüler üç boyutlu bir veri küpedir. Ham şekilde bu veri küpünü analiz etmek, depolamak veya görselleştirmek zor bir problemdir. Boyut indirgeme teknikleri ile boyutları azaltılan bu verilerin depolanması daha kolay hale gelmektedir. Ayrıca, iyi bir ön işlem sayesinde indirgenen verilerin mevcut sınıflandırma algoritmaları ile performansları artmaktadır. Bununla beraber, indirgeme işleminden sonra 2 veya 3 boyutta temsil edilebilen verilerin görselleştirilmesi oldukça kolay yapılabilmektedir.

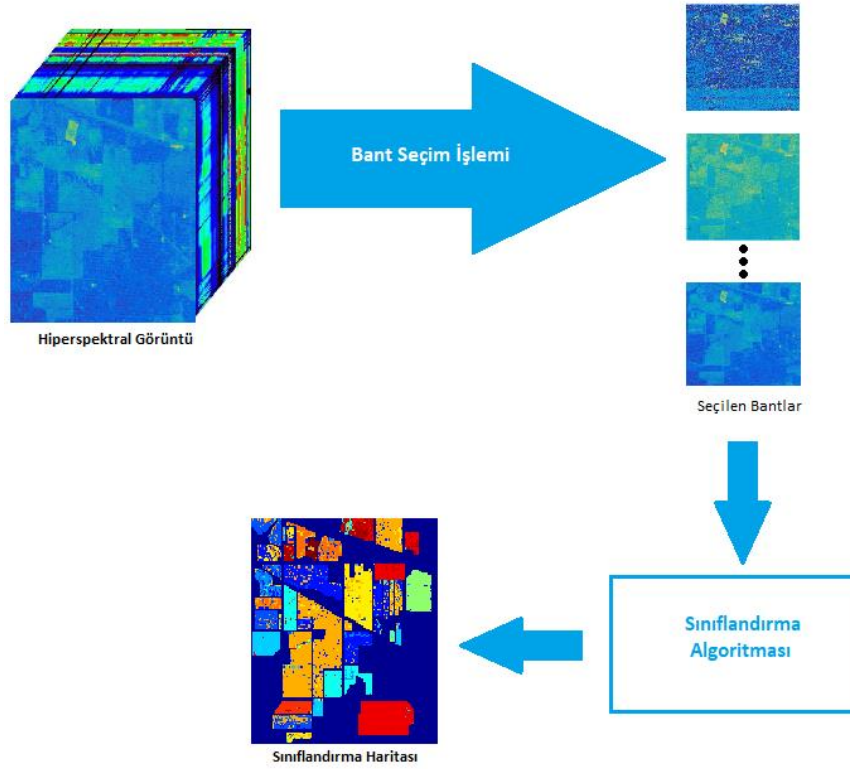


Şekil 3.1. Hiperspektral verilerde bant seçimi ve özellik çıkarımı.

Hiperspektral görüntülerde boyut indirgeme iki farklı yaklaşımla yapılabilmektedir. Şekil 3.1’de gösterilen bu yaklaşımlardan biri spektral düzlemdeki bantların seçilmesini kapsayan bant seçme yaklaşımı diğeri ise görüntülerdeki çeşitli örüntüleri bulmaya çalışan özellik çıkartma yaklaşımıdır.

### 3.1. BANT SEÇİMİ

Özellik seçim yaklaşımları, uygulanacak yöntem veya fonksiyonun gereksinimlerine bağlı olarak veri setinde bulunan gereksiz veya doğrudan sonuca etkisi olmayan özelliklerin kaldırılmasını temel almaktadır [55-57]. Özellik çıkarımı tekniklerinin aksine, özellik seçim yöntemleri orijinal verideki yapıları korumaktadır.



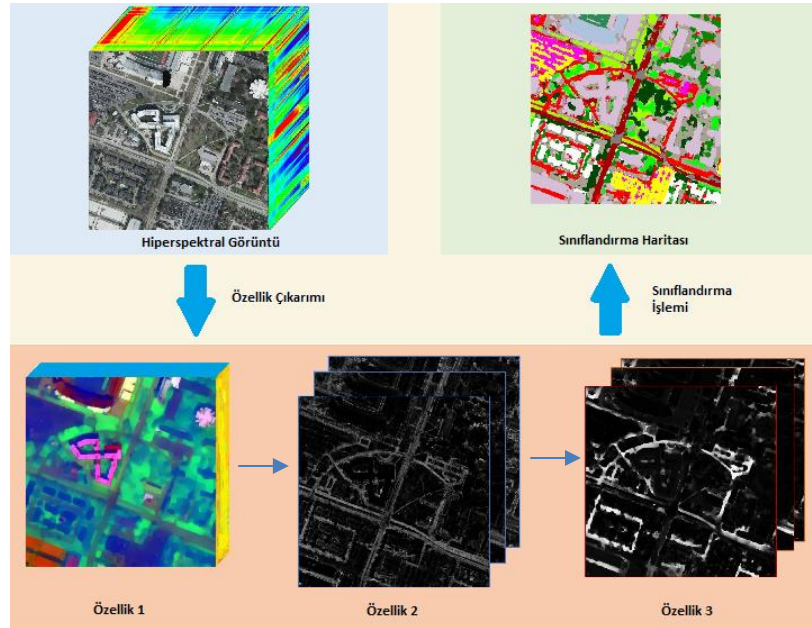
Şekil 3.2. Hiperspektral veri üzerindeki bant seçim işlem adımları [58].

Yüksek boyutlu hiperspektral verilerde geniş dalga boyu aralıklarındaki yüzlerce banttıan oluşan veriler kullanılmaktadır. Bu çok boyutluluğun çözümü için veri kaybını en aza indirecek şekilde doğru bantların seçilerek boyut indirgeme yapılması önemli bir yaklaşımdır.

Şekil 3.2’de gösterimi bulunan bant seçim yöntemlerinde özellik çıkarımındaki gibi dönüşüm işlemleri yapılmaksızın görüntü özelliklerinin birbirlerinden en iyi şekilde ayırt edilebildiği bantların bulunması amaçlanır. Böylelikle orijinal bantlardan önemli olanları seçilerek veri işleme başarısını düşürmeden hesaplama yükünün azaltılması amaçlanmaktadır. Bant seçim yönteminin özellik çıkarımına göre üstünlüğü uzamsal boyuttaki özelliklerin kaybedilmemesidir [59].

### 3.2. ÖZELLİK ÇIKARIMI

Özellik çıkarımı, orijinal büyük veri gruplarını anlamını kaybetmeden daha küçük veri kümeleri olarak temsil etmeyi amaçlamaktadır. Özellik çıkarımı yöntemleri verilerdeki önemli bilgileri kaybetmeden gerekli kaynak sayısını azaltmaktadır. Bu sayede yüksek miktarlarda bellek ve işlem gücü ihtiyacını ortadan kaldırmaktadır [55,56]. Şekil 3.3’te özellik çıkarımı için örnek bir şema verilmektedir.



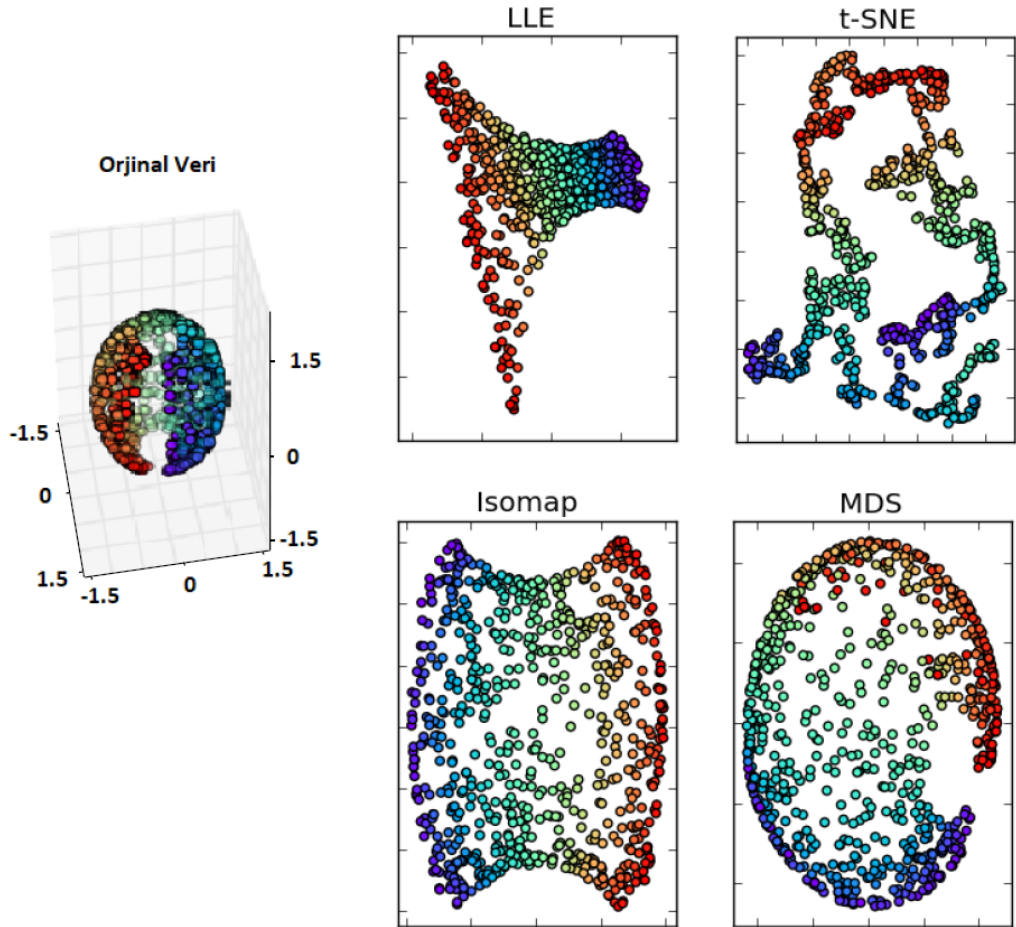
Şekil 3.3. Hiperspektral verilerde özellik çıkarımı [60].

Özellik çıkarım yönteminde orijinal bantlar üzerinden doğrusal ve doğrusal olmayan dönüşümler uygulanmaktadır. Spektral düzlemde indirgeme yapmak yerine uzamsal düzlemdeki verilerin farklı düzlemlere iz düşürülmesi ile boyut azaltma yoluna gidilmektedir [59]. Literatürde en yaygın bilenen özellik çıkarım algoritması PCA’dır.

MDS ve LFDA gibi doğrusal yöntemler diğer bilinen boyut indirgeme yöntemleri olarak karşımıza çıkmaktadır. Doğrusal olmayan yöntemlere geçildiğinde ise karşımıza manifold öğrenme kavramı çıkmaktadır.

### 3.3. MANİFOLD ÖĞRENME

Manifold öğrenme teknikleri yüksek boyutlu giriş uzayında bulunan verilerin daha düşük boyutlu bir uzayda temsilinin bulunduğunu varsaymaktadır. Doğrusal yöntemlerin aksine bu algoritmalar yüksek boyutlu giriş uzayında bulunan doğrusal olmayan karmaşık ilişkileri de başarılı bir şekilde bulabilmektedir. En popüler manifold öğrenme algoritmaları LLE, ISOMAP ve SOFM olarak sıralanabilir. Bu tekniklerle birlikte olasılık tabanlı SNE, t-SEN ve MAFE gibi algoritmalar da bilinen manifold öğrenme yöntemleridir.



Şekil 3.4. Çeşitli manifold öğrenme yöntemlerine ait çıktı uzayları [61].

Manifold öğrenme yöntemleri, en temel anlamda yüksek boyutlu giriş uzayında yer alan verileri daha düşük boyutta bir manifolda yerleştirmektedir. Bu işlemi yaparken veri noktaları arasındaki yerel ve global konum ilişkilerini korumayı amaçlamaktadırlar [62,63]. Şekil 3.4'te bazı popüler manifold öğrenme tekniklerinin aynı veri üzerinde elde ettikleri farklı manifoldlar görülmektedir.

## BÖLÜM 4

### BOYUT İNDİRGEME YÖNTEMLERİ

Hiperspektral veriler gibi büyük veri kümelerinin boyutlarının indirgenmesi için ilk olarak PCA, MDS ve LFDA gibi doğrusal yöntemler önerilmiştir. Bu yöntemler temelde yüksek boyutlu girdi uzayındaki doğrusal yapıları keşfedebilmektedir. Gelişen teknoloji ile görüntüleme sistemleri ve algılayıcılar daha kaliteli ve yüksek boyutlu veriler elde etmeye başlamışlardır. Doğrusal boyut indirgeme yöntemlerinin daha karmaşık yapılar sunan bu veri kümelerinin analizinde ve yorumlanmasında yetersiz kalması ile LLE, ISOMAP ve SOFM gibi doğrusal olmayan yöntemler kullanılmaya başlanmıştır. Bu doğrusal ve doğrusal olmayan yöntemlere ek olarak çeşitli olasılık tabanlı algoritmalarda geliştirilmiştir. Bu algoritmalar temel olarak veri noktaları arasında koşullu olasılıkları kullanmaktadır.

Bu bölümün devamında literatürdeki en yaygın kullanılan boyut indirgeme yöntemleri doğrusal ve doğrusal olmayan yöntemler başlıkları altında iki grup olarak açıklanmaktadır.

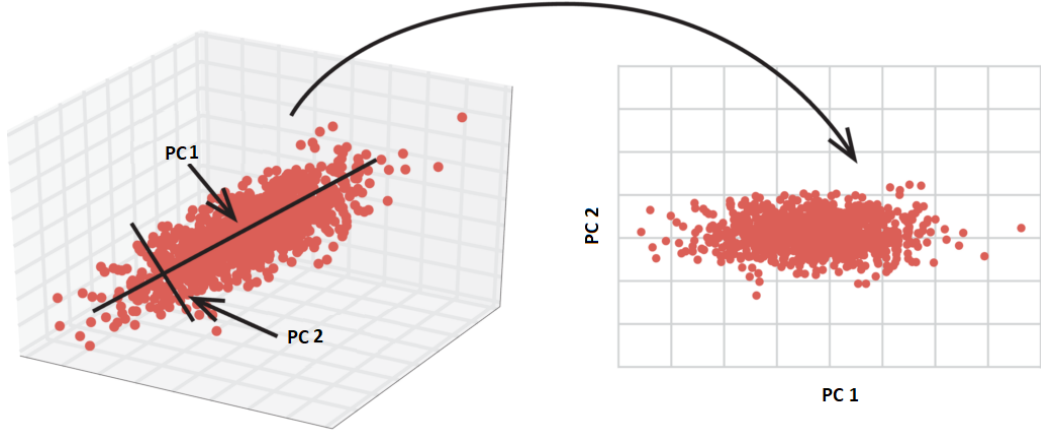
#### 4.1. DOĞRUSAL YÖNTEMLER

Doğrusal boyut indirgeme yöntemleri temelde yüksek boyutlu uzayda yer alan veri noktalarının daha düşük boyutlu doğrusal bir düzlem üzerindeki iz düşümlerini bulmaya çalışır. Bu grupta en yaygın bilinen yöntem PCA algoritmasıdır.

##### 4.1.1. Temel Bileşenler Analizi (PCA)

İlk olarak Pearson (1901) tarafından önerilen ve günümüzde geliştirilmiş birçok varyasyonu ile yaygın olarak kullanılmaya devam eden en temel boyut indirgeme yöntemlerinden biri temel bileşenler analizidir.

PCA, verileri bir koordinat sistemine yansıtmak için dikey bir doğrusal dönüşüm olarak tanımlanır [37]. Verilerin en büyük varyansı birinci koordinat, en büyük ikinci varyansı ikinci koordinat olarak tanımlayarak Şekil 4.1’de gösterildiği gibi ortogonal bir dönüşümle verileri daha düşük boyutta temsil eder.



Şekil 4.1. PCA algoritması ile boyut indirgeme [64].

Giriş uzayından verilen  $\mathbf{Y} = \{y_1, y_2, \dots, y_m\}$  gözlemleri, giriş uzayından daha düşük  $m$  boyutunda doğrusal bir alt uzay arar. Bu alt uzaya yansıtılan haritalar, verilerin kovaryans matrisinin ilk  $m$  özdeğerleri ile ilişkili en büyük varyansa sahiptir. Verileri haritalama işlemi, Eşitlik 4.1’teki merkezleme matrisi  $\mathbf{C}$ ’nin hesaplanması ile bulunmaktadır

$$\mathbf{C} = \mathbf{I}_N - \frac{1}{N} \mathbf{1}\mathbf{1}^T \quad (4.1)$$

Burada  $\mathbf{I}_N$ ,  $N \times N$  boyutlu birim matrisi ve  $\mathbf{1} = [\mathbf{1}, \dots, \mathbf{1}_N]$  ise tüm bir vektör ifadesidir.  $\mathbf{Y}$ ’nin kovaryans matrisi ise Eşitlik 4.2’de ifade edilmektedir.

$$\mathbf{\Sigma} = \mathbf{Y}\mathbf{C}\mathbf{C}^T\mathbf{Y}^T \quad (4.2)$$

Bu formülasyonda, en büyük varyansa sahip alt uzay,  $\mathbf{\Sigma}$ ’in öz değer ayrışmasından elde edilir.  $\mathbf{\Sigma}$ ’in öz değerleri büyükten küçüğe sıralanmış köşegen matrisi ve karşılık gelen  $m$  vektörelere ile  $\mathbf{E} = [e_1, \dots, e_m]$  temel bileşenlerini tanımlanır. İndirgenmiş çıkış



uzayı için elde edilen yeni koordinatlar Eşitlik 4.3'te verilen denklem ile hesaplanmaktadır.

$$\mathbf{Z} = \mathbf{E}^T \mathbf{Y} \mathbf{C} \quad (4.3)$$

#### 4.1.2. Çok Boyutlu Ölçekleme (MDS)

Çok boyutlu ölçekleme yöntemi, yüksek boyutlu uzayda bulunan veri noktaları arasındaki öklid mesafelerini koruyarak bu verileri daha düşük boyutlu bir haritaya yerleştirir [65]. Yüksek boyutlu  $\mathbf{Y}$  uzayındaki veri noktaları arasındaki ikili uzaklıklar Eşitlik 4.4'te verilen formülasyon ile hesaplanır.

$$d_{ij} = \sqrt{\sum_{k=1}^n (y_{ik} - y_{jk})^2} \quad (4.4)$$

Yöntemin temelinde yüksek boyutlu uzayda yer alan uzaklıkların indirgenmiş uzayda da korunması yatmaktadır. Bunu sağlamak için Eşitlik 4.5'te verilen denklem kullanılır. Bu formülasyon ile yüksek boyutlu uzaydan düşük boyutlu uzaya geçiş yapılır.

$$\delta(\mathbf{Z}) = \sum_{ij} (\|y_i - y_j\| - \|z_i - z_j\|)^2 \quad (4.5)$$

Burada  $\|y_i - y_j\|$ , yüksek boyutlu giriş uzayında bulunan  $y_i$  ve  $y_j$  noktaları arasındaki öklid uzaklıkları,  $\|z_i - z_j\|$  ise indirgenmiş çıkış uzayı için  $z_i$  ve  $z_j$  noktaları arasındaki öklid uzaklıklarıdır [66,67].

#### 4.1.3. Yerel Fisher Diskriminant Analizi (LFDA)

Yerel Fisher diskriminant analizi, yüksek boyutlu veri kümesini, aynı sınıfa ait verilerin varyansını en aza indirerek, farklı sınıfların varyansını ise en üst düzeye

çıkarak ve verilerin yerel yapısı koruyarak daha düşük boyutlu bir alana indirgemektedir [68]. LFDA yöntemi, temel olarak Fisher diskriminant analizi [37,69] ve yerelliği koruyan projeksiyon [37,70] yöntemlerinin avantajlarını kullanmaktadır. Bu iki yaklaşımın kombinasyonu ile manifold alanlarında çok modluluğu koruyabilen bir teknik elde edilir. LFDA'nın çözümü, bir dönüşüm matrisi olan  $T_{LFDA}$  optimizasyonunun çözülmesini içermektedir.

$$T_{LFDA} = \operatorname{argmax}_{T \in \mathbb{R}^{d \times m}} \operatorname{tr} \{ (T^T S^{(\omega)} T)^{-1} (T^T S^{(b)} T) \} \quad (4.6)$$

Burada,  $S^{(\omega)}$  yerel sınıf içi dağılım matrisi ve  $S^{(b)}$  yerel sınıflar arası dağılım matrisidir. Genelleştirilmiş özdeğer problemi Eşitlik 4.6 ile çözülerek elde edilmektedir.

## 4.2. DOĞRUSAL OLMAYAN YÖNTEMLER

Doğrusal olmayan boyut indirgeme yöntemleri, karmaşık verilerdeki doğrusal olmayan yapıları ortaya çıkarabilen tekniklerdir. En yaygın bilinen doğrusal olmayan boyut indirgeme algoritmaları bu bölümde ele alınmaktadır.

### 4.2.1. Yerel Doğrusal Yerleştirme (LLE)

LLE algoritması, yüksek boyutlu verilerin yerel komşuluk yapılarını lineer bir kombinasyonla koruyarak yaklaştırılmasını ve bu verilerin daha düşük boyutlu bir temsilini bulmayı amaçlamaktadır [37].

Algoritma yüksek boyutlu uzaydaki komşu noktaların belirlenmesi ile başlar. Bu noktalar K en yakın komşu algoritması ile belirlenir [9]. Veri setindeki tüm noktaların bu belirlenen komşu noktaları ile Öklid mesafeleri hesaplanır. Eşitlik 4.7'de hesaplanan bu mesafelere göre ağırlık matrisi oluşturulur.

$$\delta(Y) = \sum_i |Y_i - \sum_j w_{ij} Y_j|^2 \quad (4.7)$$

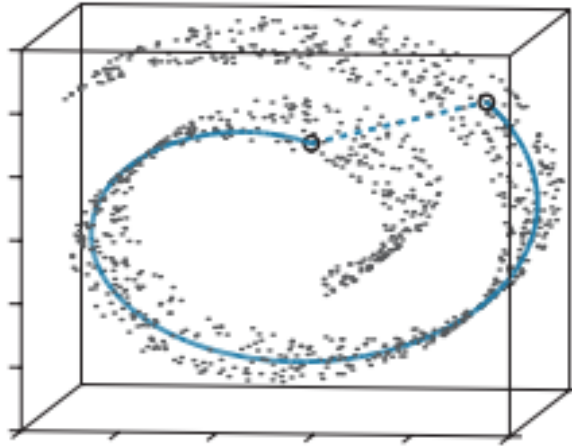
Burada,  $W_{ij}$  bir önceki adımda hesaplanan ağırlıklardır.  $Y_i$  ve  $Y_j$  ise yüksek boyutlu giriş uzayındaki veri noktaları ve komşu noktalarını temsil etmektedir. Belirlenen ağırlıklar ile indirgenmiş çıkış uzayına geçiş için ise Eşitlik 4.8'deki hesaplama kullanılmaktadır.

$$\delta(\mathbf{Z}) = \sum_i \left| \mathbf{z}_i - \sum_j W_{ij} \mathbf{z}_j \right|^2 \quad (4.8)$$

Burada,  $\mathbf{z}_i$  ve  $\mathbf{z}_j$  ise düşük boyutlu indirgenmiş veri uzayındaki noktaları ve komşu noktalarını temsil etmektedir.

#### 4.2.2. İzometrik Özellik Haritalama (ISOMAP)

PCA ve MDS yöntemlerine benzer şekilde bir yüksek boyutlu veri noktaları arasındaki uzaklıkları temel alan ISOMAP yöntemi, veri noktaları arasındaki jeodezik uzaklıkları kullanmaktadır. Jeodezik uzaklık, bir manifold yüzeyi üzerindeki iki nokta arasında bulunan en kısa yol olarak tanımlanabilir [71]. Jeodezik uzaklığın grafiksel gösterimi Şekil 4.2'de verilmektedir.



Şekil 4.2. Veri noktaları arasındaki jeodezik mesafenin gösterimi [10].

ISOMAP, yüksek boyutlu bir girdi uzayını daha düşük boyutlarda indirgenmiş bir uzaya yerleştirme işlemini 3 temel adımda yapmaktadır [66]. Bunlar;

- Komşuluk matrisinin oluşturulması,
- En yakın mesafelerin hesaplanması,
- Düşük boyutlu uzaya yerleşimdir.

Komşuluk matrisinin oluşturulmasında LLE yöntemiyle benzer şekilde Öklid uzaklıkları ile mesafeler belirlenmektedir. Ayrıca K en yakın komşu algoritması ile komşuluklar hesaplanır. İndirgenmiş uzaydaki veri noktaları ise MDS ile aynı formülasyonda elde edilmektedir.

#### 4.2.3. Laplas Özharitalar (LE)

Laplas özharitalar yöntemi, ISOMAP ve LLE yöntemlerine benzer şekilde veri noktaları arasındaki lokal komşuluklarını koruyarak daha düşük boyutlu bir manifolda indirgemeyi amaçlar. LE, veri noktaları arasındaki komşuluk ilişkilerinden elde edilen ağırlıkları diyagonal bir matrise dönüştürür. Laplas matrisi de denilen bu ağırlık matrisi simetriktir. Bir boyutlu bir dönüşüm için LE probleminin çözümü Eşitlik 4.9’da verilen optimizasyon probleminin çözümü ile hesaplanmaktadır.

$$U(\mathbf{Z}) = \sum_{i,j=1,j \neq i} w_{ij}(z_i - z_j)^2 \quad (4.9)$$

Kararlı durum için global optimizasyon problemi ise Eşitlik 4.10’deki şekilde formüle edilmektedir.

$$\mathbf{Z}^* = \operatorname{argmin}_{\mathbf{Z}} U(\mathbf{Z}) \quad (4.10)$$

Burada,  $\mathbf{Z}^*$  indirgenmiş uzayda optimum  $\mathbf{Z}$  olur.

#### 4.2.4. Rastlantısal Komşu Yerleştirme (SNE)

Rastlantısal komşu yerleştirme, yüksek boyutlu bir  $Y = \{y_1, y_2, \dots, y_m\}$  uzayındaki verileri daha düşük boyutlu  $Z = \{z_1, z_2, \dots, z_n\}$  uzayında, birbirlerine olan konumlarını korumayı amaçlayan denetimsiz bir boyut indirgeme ve veri görselleştirme tekniğidir

[15]. Yüksek boyutlu uzayda bulunan her  $i$  verisi ve potansiyel  $j$  komşulukları için asimetrik yoğunluk fonksiyonu Eşitlik 4.11’de verilmektedir. Bu denklem ile asimetrik bir ikili uzaklık matrisi elde edilmektedir.

$$p_{ij} = \frac{\exp\left(-\|y_i - y_j\|^2 / 2\sigma_i^2\right)}{\sum_{k \neq i} \exp\left(-\|y_i - y_k\|^2 / 2\sigma_i^2\right)} \quad (4.11)$$

Burada  $\sigma_i$  deneysel olarak ayarlanan varyans parametresidir. İki veri noktası birbirine yakın ise yoğunluk fonksiyonunun değeri yüksek olmaktadır. Benzer şekilde gauss dağılımı esas alınarak elde edilen düşük boyutlu uzay için çıkış fonksiyonu ise Eşitlik 4.12’de tanımlanmaktadır.

$$q_{ij} = \frac{\exp\left(-\|z_i - z_j\|^2\right)}{\sum_{k \neq i} \exp\left(-\|z_i - z_k\|^2\right)} \quad (4.12)$$

Bu yaklaşım ile yüksek boyutlu uzayda  $p_{ij}$  dağılımı ile düşük boyuta indirgenmiş  $q_{ij}$  dağılımının eşit olması hedeflenmektedir. Bu iki dağılım arasındaki Kullback-Leibler (KL) sapmalarının [37,72] toplamı olan bir maliyet fonksiyonu en aza indirgenerek kolaylıkla elde edilebilmektedir. Bu maliyet fonksiyonu Eşitlik 4.13’te verilmektedir.

$$C = \sum_i KL(P_i || Q_i) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} \quad (4.13)$$

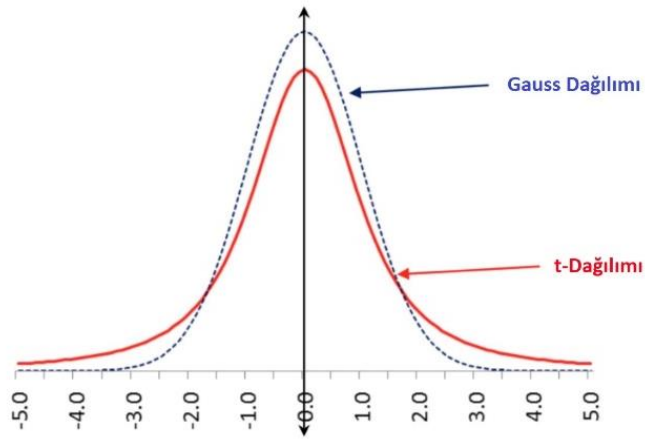
Eşitlik 4.12’deki  $z_k$  teriminin doğrudan  $q_{ij}$ ’yi etkilemesinden dolayı burada  $C$ ’nin ayırt edilebilmesi zahmetli olabilmektedir. Bu dönüşümü kolaylaştırmak için bu fonksiyon gradyan formunda Eşitlik 4.14’teki gibi ifade edilebilir.

$$\frac{\delta C}{\delta z_i} = 2 \sum_j (z_i - z_j)(p_{ij} - q_{ij} + p_{ji} - q_{ji}) \quad (4.14)$$

Burada indirgenmiş uzaydaki yerleştirme koordinatları, rasgele başlatılan gauss örnek haritaları ile bir gradyan iniş algoritması kullanılarak bulunabilmektedir. Maliyet fonksiyonu da bu formülasyon ile elde edilmektedir.

#### 4.2.5. t-Dağılımlı Rastlantısal Komşu Yerleştirme (t-SNE)

Bu yöntem, daha basit bir gradyan hesaplaması kullanarak SNE algoritmasının maliyet fonksiyonunun simetrik olan bir versiyonunu kullanmaktadır. Düşük boyutlu indirgenmiş uzay için gauss dağılımı yerine t-dağılımı (veya student t-dağılımı) kullanılır. Şekil 4.3'te görüldüğü üzere t-dağılımı yapısı gereği uç noktalarda daha hassas bir ayırım yapabilmektedir. Bu da SNE'ye göre sınıflandırma başarısını artırmakta ve kalabalık probleminde de iyileşme sağlamaktadır [16].



Şekil 4.3. Gauss ve t-dağılımı grafikleri.

Giriş uzayı için SNE ile benzer şekilde Eşitlik 4.1'deki formülasyonu kullanan t-SNE yönteminin çıkış uzayı için kullandığı dağılım formülasyonu ise Eşitlik 4.15'te verilmektedir.

$$q_{ij} = \frac{(1 + \|z_i - z_j\|^2)^{-1}}{\sum_{k \neq i} (1 + \|z_i - z_k\|^2)^{-1}} \quad (4.15)$$

Maliyet fonksiyonu SNE'de olduğu gibi yine KL sapmalarının toplamına dayalıdır. Gradyan iniş algoritması ile denklemin indirgenmiş hali Eşitlik 4.16'da ifade edilmektedir.

$$\frac{\delta C}{\delta z_i} = 4 \sum_j (p_{ij} - q_{ij})(z_i - z_j)(1 + \|z_i - z_j\|^2)^{-1} \quad (4.16)$$

#### 4.2.6. Çok Boyutlu Yapay Alan Yerleştirme (MAFE)

MAFE, yüksek boyutlu alandaki komşuluk ilişkilerini korumak için minimum enerji konfigürasyonuna yol açan yeni bir boyut indirgeme çatisıdır. Potansiyel bir alan için mekanikten gelen kuvvet alanı yöntemi, en uygun enerji çözümünü bulmak için toplam çekme ve itme potansiyel fonksiyonlarını en aza indirmek olarak yorumlanır. Daha düşük boyutlu bir uzayda her bir yüksek boyutlu veri noktasının optimum gömme haritası, her iki kuvvet alanının denge durumuna karşılık gelir. Potansiyel enerji fonksiyonu için bir havuz olarak  $i$  ve  $j$  haritaları arasındaki minimum enerji konfigürasyonu için, bir dizi kurala bağlı olarak, Eşitlik 4.17’de verilen itme potansiyeli fonksiyonu kısa mesafelerde baskındır [18].

$$U_r^{ij}(\|z_i - z_j\|) = \xi_r \sigma \exp \left\{ -\frac{\|z_i - z_j\|^q}{\sigma} \right\} \quad (4.17)$$

Burada,  $\xi_r$  itme kuvvetinin büyüklüğü,  $\sigma$  standart sapma ve  $q$  çıktı uzayının dağılımının yoğunluğunu belirler. İtme potansiyeli fonksiyonu düşük boyutlu uzayda  $z_i$  ve  $z_j$  noktalarının çok yakınlaşmasını engellerken, Eşitlik 4.18’de verilen çekme potansiyeli fonksiyonu ise uzun mesafelerde baskındır.

$$U_a^{ij}(\|z_i - z_j\|) = \xi_a w_{ij} \|z_i - z_j\|^p \quad (4.18)$$

Burada, benzer şekilde  $\xi_a$  çekim kuvvetinin büyüklük parametresi,  $p$  çıktı dağılımının yoğunluğunu belirler ve  $w_{ij}$  [73] spektral imza benzerliğini ifade eder. Çekme potansiyel fonksiyonu, itme fonksiyonunun aksine  $z_i$  ve  $z_j$  noktalarının birbirlerinden çok uzaklaşmasını engeller. Bu nedenle, itme potansiyeli fonksiyonu yerel ilişkileri temsil ederken, çekme potansiyeli fonksiyonu uzun mesafelerdeki küresel ilişkileri temsil etmektedir. Kuvvet alanı formülasyonuna göre, çekme ve itme potansiyel fonksiyonları ile her bir harita için hareket dinamiği Eşitlik 4.19’daki gibi yazılmaktadır.

$$\dot{z}_i = - \sum_{j \neq i} \{ \nabla_{z_i} U_a^{i,j}(\|z_i - z_j\|) - \nabla_{z_i} U_r^{i,j}(\|z_i - z_j\|) \} \quad (4.19)$$

Burada,  $U_a^{ij}$  ve  $U_r^{ij}$  sırasıyla çekme ve itme potansiyel fonksiyonlarıdır.  $\nabla_{z_i}$  gradyan operatörüdür.  $(z_i - z_j)$  vektörü, çekme ve itme etkileşim kuvvetlerinin zıt yönlerde hareket ettiği hizalamayı tanımlar. Her bir harita için çekme ve itme potansiyeli Eşitlik 4.20'de verilmektedir.

$$U_i(\mathbf{Z}) = \sum_{j \neq i} \{U_a^{ij}(\|z_i - z_j\|) - U_r^{ij}(\|z_i - z_j\|)\} \quad (4.20)$$

Toplam üst üste gelme potansiyeli ise Eşitlik 4.21'deki gibi ifade edilmektedir.

$$U(\mathbf{Z}) = \sum_{i=1}^N U_i(\mathbf{Z}) \quad (4.21)$$

Burada  $N$ , gömme alanındaki her bir haritanın sayısıdır. Kararlı durum için global optimizasyon problemi ise şu şekilde formüle edilir:

$$\mathbf{Z}^* = \operatorname{argmin}_{\mathbf{Z}} U(\mathbf{Z}) \quad (4.22)$$

Eşitlik 4.22'de,  $\mathbf{Z}^*$  indirgenmiş uzayda optimum  $\mathbf{Z}$  olur. Sonuç olarak itme ve çekme kuvvetleri dengelenmektedir. Maliyet fonksiyonu, KL sapmalarının toplamına dayalı gradyan iniş algoritması ile Eşitlik 4.17, 4.18 ve 4.19 kullanılarak Eşitlik 4.23'teki son halini almaktadır.

$$\begin{aligned} \dot{z}_i = & - \sum_{j \neq i} (z_i - z_j) \xi_a w_{ij} p \|z_i - z_j\|^{p-2} \\ & - (z_i - z_j) \xi_r q \|z_i - z_j\|^{q-2} \exp \left\{ - \frac{\|z_i - z_j\|^q}{\sigma} \right\} \end{aligned} \quad (4.23)$$

Tüm bu doğrusal ve doğrusal olmayan yöntemler temelde yüksek boyutlu uzayda bulunan büyük veri kümelerinin daha düşük boyutlu bir uzayda temsilini bulmaya çalışırlar. Özellikle hiperspektral veriler gibi karmaşık ve büyük boyutlarda veri kümelerinin analizi ve yorumlanmasında yüksek işlem ve bellek maliyetlerinin ortaya çıkması boyut indirgeme işleminin önemini daha çok açığa çıkarmaktadır. Bu tür



büyük veri kümeleri analiz edilmeden önce içerisinde barındırdığı global ve yerel yapıları koruyarak daha düşük boyutlarda temsil edilmelidir. Bu sayede dünya nesnelere ve uzay cisimleri hakkında çok daha ayrıntılı bilgilere temas etmeden erişmemiz mümkün olmaktadır.

## BÖLÜM 5

### BOYUT İNDİRGEME UYGULAMALARINDA OPTİMİZASYON VE ADAPTİF AĞIRLIKLANDIRMA

Bu tez çalışmasında, iki ayrı akademik çalışmada kullanılan iki farklı yöntem önerilmektedir. Yöntemlerin her ikisi de temelde MAFE algoritmasına ve kullandığı kuvvet alanı yasası yaklaşımına dayanmaktadır. İlk yöntemde MAFE yönteminin parametre belirleme sorunu optimizasyon tabanlı bir yaklaşımla çözülmektedir. Bu sayede önerilen yöntem ile veri setinden bağımsız olarak algoritmanın maksimum performansta çalışması sağlanmaktadır. Önerilen ikinci yöntemde ise t-SNE ve MAFE algoritmalarının güçlü yönlerini birlikte kullanmayı amaçlayan hibrit bir yaklaşım sunulmaktadır. Bu yöntemde ise MAFE algoritmasının yüksek sınıflandırma doğruluğu performansı ile t-SNE algoritmasının hızlı çalışma zamanı avantajının birleştirilmesi sağlanmaktadır.

#### 5.1. PARÇACIK SÜRÜ OPTİMİZASYONU TABANLI MAFE

Boyut indirgeme uygulamalarında, büyük veri kümelerindeki bant seçimi, öznitelik çıkarımı, parametre seçimi gibi birçok zorlukla başa çıkmak için optimizasyon algoritmaları yaygın bir şekilde kullanılmaktadır. Literatürde en yaygın bilinen optimizasyon algoritması ise Parçacık sürü optimizasyonu (PSO) yöntemidir.

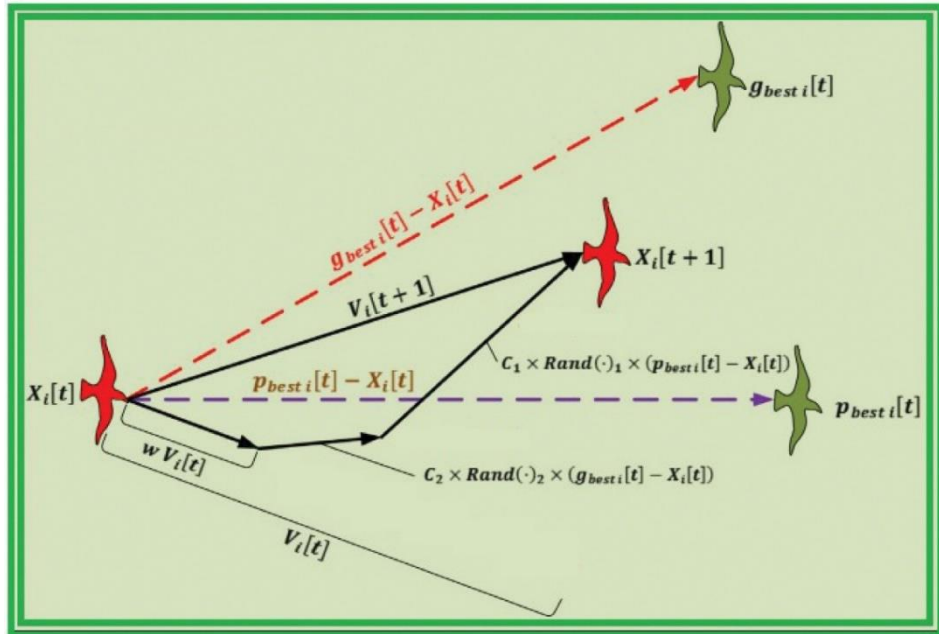
PSO, ilk olarak 1995 yılında Kennedy ve Eberhart tarafından önerilen popülasyon tabanlı bir optimizasyon yöntemidir [74]. Bu yöntem, kuş ve balık sürülerinin davranışlarından esinlenen sezgisel bir yaklaşımdır. Basitliği ve hızlı yakınsama kabiliyeti nedeniyle PSO, son yıllarda birçok alanda araştırma odağı olmuştur [75]. Özellikle boyut indirgeme [76-78], makine öğrenmesi [79], öznitelik çıkarımı [80], görüntü sınıflandırma [81], bant seçimi [82] ve alt-piksel haritalama [83] gibi birçok çalışmada PSO tabanlı yaklaşımlar tercih edilmektedir.

PSO’da problemin çözüm uzayı bir arama uzayı olarak formüle edilmektedir. Arama uzayında yer alan her bir konum, problemin potansiyel bir çözümüdür. Parçacıklar, arama uzayında en iyi konumu bulmak için iş birliği yaparlar. Her bir parçacık kendi hızına göre hareket eder. Her iterasyon için parçacık hareketi ise Eşitlik 5.1 ve 5.2’deki gibi hesaplanmaktadır [84].

$$x_i(t+1) \leftarrow x_i(t) + v_i(t), \quad (5.1)$$

$$v_i(t+1) \leftarrow \omega v_i(t) + c_1 r_1 (pbest_i(t) - x_i(t)) + c_2 r_2 (gbest_i(t) - x_i(t)) \quad (5.2)$$

Burada  $x_i(t)$ ,  $i$  parçacığının  $t$  zamanındaki konumudur.  $v_i(t)$ ,  $i$  parçacığının  $t$  zamanındaki hızını vermektedir.  $pbest_i(t)$ , parçacığın o ana kadarki en iyi konumudur.  $gbest_i(t)$  ise, su ana kadar tüm sürü tarafından bulunan en iyi konumdur.  $\omega$ , önceki zaman adımını ölçeklendiren bir atalet ağırlığı parametresidir.  $c_1$  ve  $c_2$  ise  $pbest_i(t)$  ve  $gbest_i(t)$  fonksiyonlarının etkisini ölçeklendiren hızlanma katsayılarıdır.  $r_1$  ve  $r_2$  parametreleri de deneysel olarak [0-1] aralığında seçilen rasgele değişkenlerdir [84,85]. Şekil 5.1’ de PSO algoritmasının örnek bir çalışma simülasyonu örneği gösterilmektedir.



Şekil 5.1. PSO algoritmasının temsili çalışma simülasyonu [86].

PSO algoritmasının temel çalışma adımları aşağıdaki gibidir.

1. Parçacıkların hız ve konum başlangıç değerlerini belirle.
2.  $pbest$  ve  $gbest$  başlangıç değerlerini hesapla.
3. Eşitlik 5.1 ve 5.2'ye göre konum ve hız değerlerini güncelle.
4.  $pbest$  ve  $gbest$  değerlerini güncelle.
5. Sonlandırma koşulu sağlanmadı ise 3. adıma git.
6.  $gbest$  değerini al ve sonlandır.

MAFE algoritması, mevcut manifold yerleştirme yöntemleri arasında sınıflandırma başarısı en yüksek yöntemdir [27]. Ancak yöntemin Eşitlik 4.23'te verilen maliyet fonksiyonu üzerinde deneysel olarak seçilen çok sayıda parametre bulunmaktadır. Bu parametreler hem birbirlerinden etkilenmekte hem de sınıflandırma sonuçlarını oldukça fazla etkilemektedir. Ayrıca bu parametrelere bağlı olarak, özellikle seçilen karmaşık veri setleri üzerinde, kalabalık problemi görülmektedir. Deneysel çalışmalar, bu parametreler arasında yer alan, indirgenmiş uzayda itme ve çekme kuvvetlerinin yoğunluk parametreleri olan  $p$  ve  $q$  parametrelerinin algoritma performansını en fazla etkileyen parametreler olduğunu göstermektedir. Bu parametreler daha önce yapılan çalışmalarda [0-2] aralığında seçilmiştir [18] ancak deneysel çalışmalarda parametrelere ait optimal değerler kolaylıkla elde edilememektedir. Ayrıca her veri seti için optimal parametreler değişmektedir.

Tüm bu sorunların çözümü için tez kapsamında önerilen ilk yöntem PSO tabanlı MAFE (PSO-MAFE) yöntemidir. Bu yöntem PSO algoritmasını kullanarak  $p$  ve  $q$  parametrelerinin optimal değerlerini elde etmektedir. Bu sayede yöntem uygulandığı verilerden bağımsız olarak her veri seti için yüksek performans sergilemektedir.

PSO-MAFE yönteminde, ilk olarak  $p$  ve  $q$  parametreleri bir  $X$  dizisi olarak ifade edilmektedir. Bu dizi için Eşitlik 5.3'te gösterildiği gibi deneysel olarak minimum ve maksimum değer limitleri belirlenmektedir. Bu limitler çalışılan veri setine bağlı olarak değişkenlik göstermektedir.

$$X_{i,min} \leq X_i \leq X_{i,maks} \quad (5.3)$$

Eşitlik 4.23'te verilen MAFE yönteminin maliyet fonksiyonu ise yeniden düzenlenerek Eşitlik 5.4'teki gibi ifade edilmektedir.

$$\dot{z}_{MAFE} = - \sum_{j \neq i} (\mathbf{z}_i - \mathbf{z}_j) \left\{ \xi_a w_{ij} \|\mathbf{z}_i - \mathbf{z}_j\|^{x_0-2} - \xi_r q \|\mathbf{z}_i - \mathbf{z}_j\|^{x_1-2} \right\} \quad (5.4)$$

Düzenlenen bu formülasyon ile MAFE algoritması PSO içerisinde çalıştırılarak Eşitlik 5.5'te de ifade edildiği şekli ile optimal  $X$  dizisi elde edilmektedir.

$$\operatorname{argmin} X = F_{PSO}(X_{i,\min}, X_{i,\max}, \dot{z}_{MAFE}) \quad (5.5)$$

Burada  $F_{PSO}$  fonksiyonu parameter olarak parametrelerin alabileceği maksimum sınırı, minimum parametre sınırını ve maliyet fonksiyonunu almaktadır. Geriye ise optimum parametre değerlerini içeren  $X$  dizisini döndürmektedir.  $X$  dizisi sırasıyla  $p$  ve  $q$  parametre değerlerini içermektedir. Burada  $p = X_0$  ve  $q = X_1$  olarak ifade edilmektedir.

## 5.2. ADAPTİF AĞIRLIKLANDIRILAN MANİFOLD YERLEŞTİRME

Tez kapsamında önerilen ikinci yöntem olan adaptif ağırlıklandırılan manifold yerleştirme (Adaptive weighted embedding method-AWEM), mevcut manifold öğrenme yöntemlerinin güçlü yönlerini birleştiren bir yaklaşımdır. Buradaki temel amaç mevcut yöntemlerin avantajlarını birleştirmektir. Bu kapsamda, önerilen yöntem t-SNE algoritmasının hızını ve MAFE'nin yüksek sınıflandırma doğruluğu avantajını kullanmaktadır. Ayrıca bu birleşimin neticesinde çıkış uzayları için elde edilen yerleştirme koordinatlarının görselleştirilmesinde de performans artışı sağlanmaktadır.

Önerilen yöntem ilk olarak, Eşitlik 4.15'te verilen t-SNE algoritmasının indirgenmiş uzay için kullandığı dağılım fonksiyonunu Eşitlik 5.6'daki gibi yeniden düzenlemektedir.

$$q_{ij} = \frac{\|\mathbf{z}_i - \mathbf{z}_j\|^{-1}}{\sum_{k \neq i} \|\mathbf{z}_i - \mathbf{z}_k\|^{-1}} \quad (5.6)$$

Yeni dağılım fonksiyonu kullanılarak gradyan iniş algoritması ile maliyet fonksiyonu ise Eşitlik 5.7’de verilmektedir. Eşitlik 4.15’teki orijinal formülasyondaki bazı kısımlar deneysel olarak etkisi göz ardı edilebilir görülmektedir. Bu sebeple denklemdaki bazı bölümler sadeleştirilmiştir.

$$\dot{z}_{tSNE} = 4 \sum_j (p_{ij} - q_{ij})(z_i - z_j) \|z_i - z_j\|^{-1} \quad (5.7)$$

Önerilen yöntemde ayrıca MAFE yönteminin çıkış uzayı formülasyonunda t-SNE ile adaptif olarak uyumlu sonuçlar vermesi için t-dağılımı kullanılmaktadır. Eşitlik 4.23’te verilen MAFE algoritmasının orijinal maliyet fonksiyonu yeni dağılım ile yeniden Eşitlik 5.8’deki gibi ifade edilmektedir. Orijinal maliyet fonksiyonundaki deneysel olarak etkisi göz ardı edilebilecek bazı kısımlarda çıkarılmış veya sadeleştirilmiştir.

$$\dot{z}_{MAFE} = -2 \sum_{j \neq i} (z_i - z_j) \xi_a p_{ij} \|z_i - z_j\|^{-p} - (z_i - z_j) \xi_r q_{ij} \|z_i - z_j\|^{-q} \quad (5.8)$$

Önerilen yöntem, Eşitlik 5.7 ve Eşitlik 5.8’deki maliyet fonksiyonlarını birlikte kullanmaktadır. Bu fonksiyonların ağırlıkları bir dizi kritere bağlı olarak güncellenmektedir. Güncelleme işlemi gerçek zamanlı bir yaklaşımla adaptif olarak gerçekleşmektedir.

$$\dot{z}_i = w_1 \dot{z}_{tSNE} - w_2 \dot{z}_{MAFE} \quad (5.9)$$

Eşitlik 5.9’da verilen  $w_1$  ve  $w_2$  ağırlık parametreleri, algoritmaların maksimum sınıflandırma doğruluğuna yakınsama hızına göre [0-1] aralığında değişmektedir. Başlangıçta, t-SNE algoritmasının ağırlığı yüksek (1’e yakın) seçilir ve ilk yakınsama hızı maksimumda tutulur. Daha sonra MAFE ağırlığı kademeli olarak artırılır ve maksimum doğruluk oranı yukarı doğru çekilebilir. Böylece başlangıçta hızla yükselen sonrasında ise hızı azalmasına karşın yükselmeye devam eden bir sınıflandırma doğruluğu eğrisi elde edilmektedir. Bu sayede her iki yöntemin güçlü yönleri bir arada kullanılmaktadır.

Çizelge 5.1’de AWEM yönteminin algoritması verilmektedir. Burada giriş uzayının ikili benzerlik matrisi ve başlangıç parametreleri hesaplanmaktadır. Sonrasında Çıkış uzayı için rasgele bir örnek uzay oluşturulur. Döngü içerisinde ise indirgenmiş uzay için ikili benzerlik matrisi sürekli güncellenmektedir ve maliyet fonksiyonu değerleri değişmektedir. Ağırlıklar ise her iterasyon sonunda sınıflandırma hatasının hesaplanması ve deneysel olarak belirlenen eşik değerinin üzerinde olması durumunda değiştirilmektedir.

Çizelge 5.1. AWEM algoritması.

---

**Algoritma:** Basit versiyon AWEM

---

**Girdi:** Yüksek boyutlu veri seti  $\mathbf{Y} = \{y_1, y_2, \dots, y_m\}$ ;

**Çıktı:** İndirgenmiş uzay yerleştirme koordinatları  $\mathbf{Z} = \{z_1, z, \dots, z_n\}$ ;

Başlangıç parametrelerini ayarla,  
 İkili benzerlikleri hesapla-  $p_{ij}$  (Eşitlik 4.11 ile)  
 Başlangıç haritasını rastgele değerler ile oluştur-  $\mathbf{Z}^{(1)}$

**while**  $\|\nabla \cup \mathbf{Z}^{(t)}\| > \epsilon$  **do**  
    $t = t + 1$ ;  
   Düşük boyutlu uzaydaki ikili benzerlikleri hesapla-  $q_{ij}$  (Eşitlik 5.6 ile)  
   Gradyan hesapla-  $\frac{\delta C}{\delta z_i}$  (Eşitlik 5.9 ile)  
    $z^{(t)} = z^{(t-1)} + n \frac{\delta C}{\delta y} + \alpha(t)(z^{(t-1)} - z^{(t-2)})$   
    $w_1$  ve  $w_2$ ’yi güncelle  
**end**

---

Bu bölümde önerilen her iki yöntemde temelde kuvvet alanı yasasına dayanan MAFE yönteminin geliştirilmiş birer varyasyonu olarak görülebilir. MAFE hiperspektral verilerin indirgenmesinde ve dolaylı olarak sınıflandırılmasında oldukça başarılı bir yöntemdir [27]. Ancak kullanılan verilerin yapısına ve karmaşıklığına bağlı olarak bazı eksiklikleri mevcuttur. Önerilen bu yaklaşımlar ile bahsedilen dezavantajlar büyük ölçüde giderilmektedir.

Deneysel çalışmalar bölümünde bahsedilen bu çalışmalara ait sonuçlar ayrıntılı olarak analiz edilmektedir. Üç farklı çalışmada, bu bölümde önerilen yöntemler ile sonuçlar alınmaktadır. Elde edilen bulgular önerilen yöntemlerin mevcut yöntemlere olan üstünlüğünü göstermektedir. Beş farklı güncel hiperspektral veri seti üzerinden

gerçekleştirilen deneyler önerilen yöntemlerin hem sınıflandırma doğruluğu hem de verilerin görselleştirilmesinde mevcut yöntemlere göre daha iyi sonuçlar elde ettiğini göstermektedir. Ayrıca hiperspektral veriler gibi büyük ve kompleks veri gruplarının sınıflandırılmasında önemli bir sorun olan kalabalık problemi, önerilen yöntemler ile büyük ölçüde giderilmektedir.



## BÖLÜM 6

### DENEYSEL ÇALIŞMALAR

Deneysel çalışmalarda en yaygın kullanılan hiperspektral veri setlerinden beş tanesi seçilerek çalışmalarda kullanılmaktadır. Yine bu bölüm kapsamında üç farklı çalışmanın detayları verilmektedir.

#### 6.1. KULLANILAN VERİ SETLERİ

Tez kapsamında son yıllardaki güncel boyut indirgeme ve hiperspektral görüntü sınıflandırma çalışmalarında tercih edilen en popüler 5 hiperspektral veri seti seçilmiştir. Çalışmaların içeriğine bağlı olarak bu veri setlerinin tamamı veya bazıları tüm çalışmalarda kullanılmaktadır.

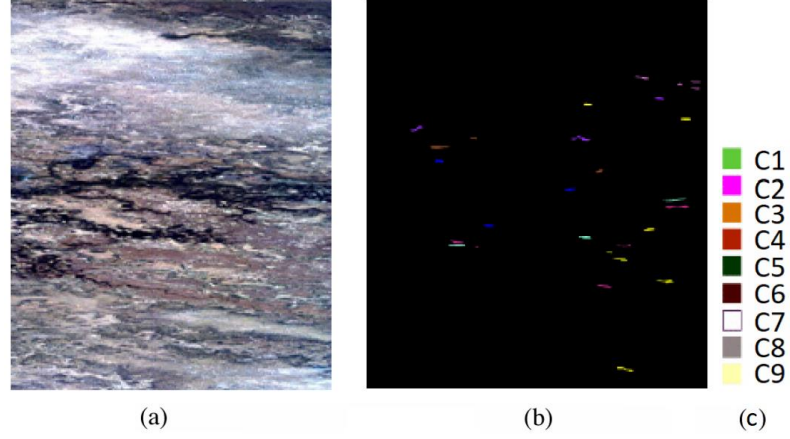
Bunlar sırasıyla;

- Botswana,
- Kennedy Uzay Merkezi,
- Indiana Pines,
- Salinas,
- Pavia Üniversitesidir.

##### 6.1.1. Botswana

Çalışmalar kullanılan ilk ve en eski bilinen veri seti dokuz farklı sınıfa sahip Botswana Hyperion verileridir [18]. Botswana veri seti, NASA EO-1 [87] uydusu ile elde edilmektedir. Yer gözlem uydusu olan EO-1, yörüngede 705 km irtifada bulunmaktadır. Elektromanyetik spektrumun 200'den fazla dalga boyunu ölçebilen bir hiperspektral spektrometreye sahiptir. Botswana verileri, Okavango Deltası, Botswana üzerinde 10 nm pencerelerde (aralıklarla) spektrumun 400-2500 nm aralığında bulunan

bölümde 7,7 km'lik bir şerit üzerinde, 30 m piksel mekansal çözünürlükte verilerden oluşmaktadır. Doğrulama verisi ve sınıf görselleri ise Şekil 6.1'de yer almaktadır.



Şekil 6.1. a) Botswana renkli görüntüsü b) doğrulama verisi c) sınıflar [88].

Çizelge 6.1'de verilen genel sınıf grupları, gürültülü bantların çıkarılmasıyla elde edilen 145 bant içermektedir. Bu bantlarla sınıflandırma çalışmaları yapılmaktadır. Botswana verilerindeki toplam örnek sayısı 1580'dir.

Çizelge 6.1. Botswana doğrulama verisi sınıfları ve örnek sayıları.

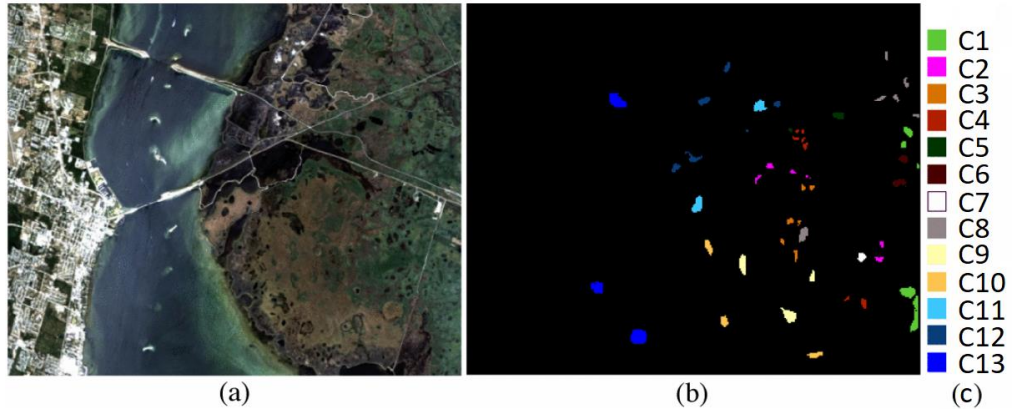
Etiket	Botswana	
	Sınıf	Örnek Sayısı
C1	Su (Water)	158
C2	Taşkın alanı (Floodplain)	228
C3	Dere kenarı (Riparian)	237
C4	Kara yüzey (Firescar)	178
C5	Ada içi (Island interior)	183
C6	Ağaçlık (Woodlands)	199
C7	Çayır (Grasslands)	162
C8	Kısa mopan (Short mopane)	124
C9	Sınıflandırılmamış (Unclassified)	111
<b>Toplam</b>		<b>1580</b>

Botswana verileri, tez kapsamındaki tüm çalışmalarda kullanılmaktadır.

### 6.1.2. Kennedy Uzay Merkezi

Kennedy Uzay Merkezi (KSC) verileri, NASA Airborne Visible / Infrared Imaging Spectrometer (AVIRIS) [89] algılayıcısı ile elde edilmektedir. AVIRIS algılayıcısı,

farklı uçak platformları ile kullanılmış uzaktan algılama alanında elde ettiği veriler ile başarılı bir optik algılayıcıdır. Kullandığı uçaklara bağlı olarak 4 ila 20 km arası irtifalarda ölçüm yapabilmektedir. KSC verileri Florida üzerinden 400-2500 nm dalga boyunda 10 nm genişliğinde 224 banttan elde edilmiş verilerden oluşmaktadır. Yaklaşık 20 km'lik bir rakımdan elde edilen KSC verileri, 18 metrelik bir mekansal çözünürlüğe sahiptir. Doğrulama verisi ve sınıf görselleri ise Şekil 6.2'de yer almaktadır.



Şekil 6.2. a) KSC renkli görüntüsü b) doğrulama verisi c) sınıflar [53].

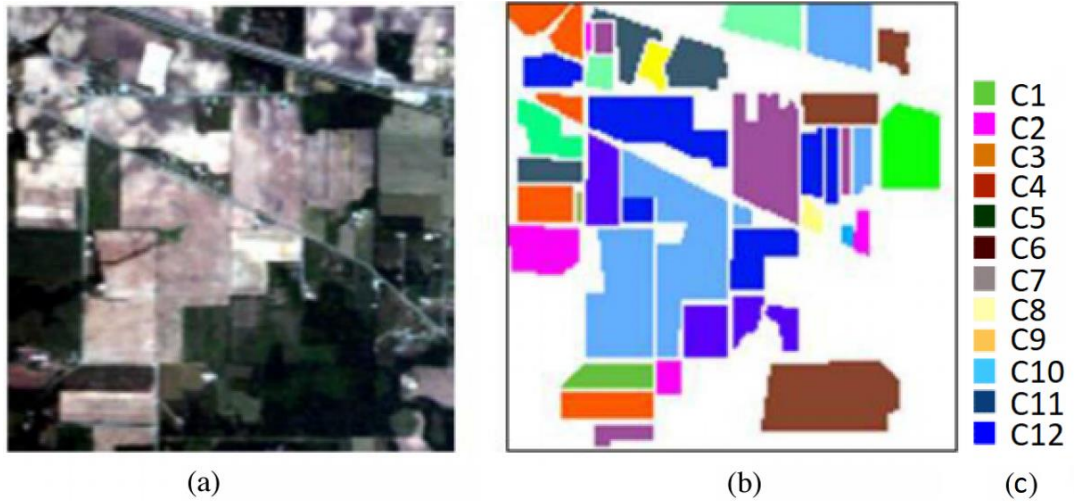
Çizelge 6.2'de verilen genel sınıf grupları, gürültülü bantlar çıkarıldıktan sonra analiz için 176 spektral bant üzerinde kullanılmaktadır. Sınıflandırma çalışmalarında kullanılan veriler 13 farklı sınıftan oluşmakta ve toplam 5211 örnek içermektedir.

Çizelge 6.2. KSC doğrulama verisi sınıfları ve örnek sayıları.

Etiket	KSC	
	Sınıf	Örnek Sayısı
C1	Çalılık (Scrub)	761
C2	Söğüt bataklığı (Willow swamp)	243
C3	CP/Hamak (CP/Hammock)	256
C4	CP/Palmiye/Ham (CP/Palm/Ham)	252
C5	Ağaçlık çam (Slash pine)	161
C6	Meşe/Kuru (Oak/Broad)	229
C7	Meşe bataklığı (Hardwood swamp)	105
C8	Otsu bataklık (Graminoid marsh)	431
C9	Çim bataklık (Spartina marsh)	520
C10	Hasırotu (Cattail marsh)	404
C11	Tuzlu bataklık (Salt marsh)	419
C12	Çamur düzlük (Mud flats)	503
C13	Su (Water)	927
<b>Toplam</b>		<b>5211</b>

### 6.1.3. Indiana Pines

Indiana pines (IP) veri seti, Kuzey-Batı Indiana'daki Pines test alanı üzerinden AVIRIS algılayıcısı tarafından toplanmaktadır. 0,4–2,5 mikrometre arası dalga boyu aralığında 145x145 piksel ve 224 spektral banttandır. Doğrulama verisi ve sınıf görselleri ise Şekil 6.3'te yer almaktadır.



Şekil 6.3. a) IP renkli görüntüsü b) doğrulama verisi c) sınıflar [88].

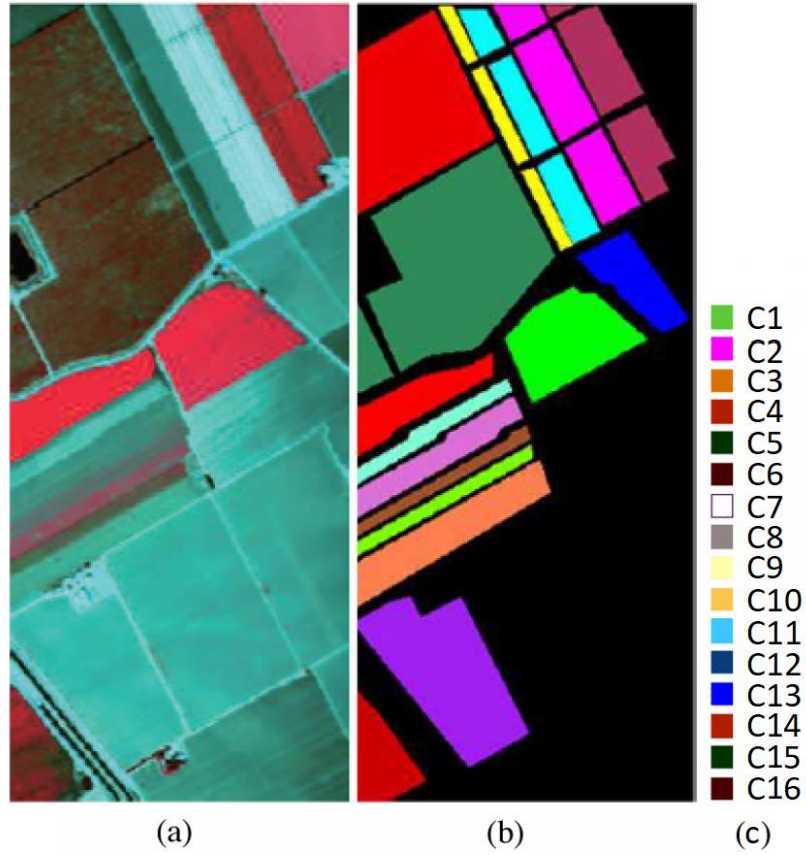
Çizelge 6.3'te gösterilen genel sınıf grupları, gürültülü bantların çıkarılması ile 200 banttandır. Sınıflandırma için kullanılacak veriler 12 sınıf ve 10171 örnekten oluşmaktadır.

Çizelge 6.3. IP doğrulama verisi sınıfları ve örnek sayıları.

Etiket	Indiana Pines	
	Sınıf	Örnek Sayısı
C1	Sürülmemiş mısır (Corn-no till)	1434
C2	Sürmeden ekilmiş mısır (Corn-min till)	834
C3	Mısır (Corn)	234
C4	Otlak (Grass-pasture)	497
C5	Çim Ağacı (Grass-trees)	747
C6	Saman yığını (Hay-windrowed)	489
C7	Sürülmemiş soya fasülyesi (Soybean-no till)	968
C8	Sürmeden ekilmiş soya fasülyesi (Soybean-min till)	2468
C9	Temizlenmiş Soya fasülyesi (Soybean-clean)	614
C10	Buğday (Wheat)	212
C11	Orman (Woods)	1294
C12	Binalar-çim-ağaçlar-sürücüler (Buildings-grass-trees-drives)	380
<b>Toplam</b>		<b>10171</b>

#### 6.1.4. Salinas

Salinas veri seti, AVIRIS algılayıcısı tarafından Salinas Vadisi, Kaliforniya'dan toplanan verilerden oluşmaktadır [90]. Veriler 360-2500 nm dalga boyu aralığında 512x217 piksel ve 224 banttan oluşmaktadır. Ayrıca 3,7 gibi yüksek bir mekansal çözünürlüğe sahiptir. Doğrulama verisi ve sınıf görselleri ise Şekil 6.4'te yer almaktadır.



Şekil 6.4. a) Salinas renkli görüntüsü b) doğrulama verisi c) sınıflar [88].

Çizelge 6.4'te gösterilen genel sınıf grupları, gürültülü bantların çıkarılması ile 204 banttan oluşmaktadır. Sınıflandırma için kullanılacak veriler 16 sınıf ve 54129 örnekten oluşmaktadır.

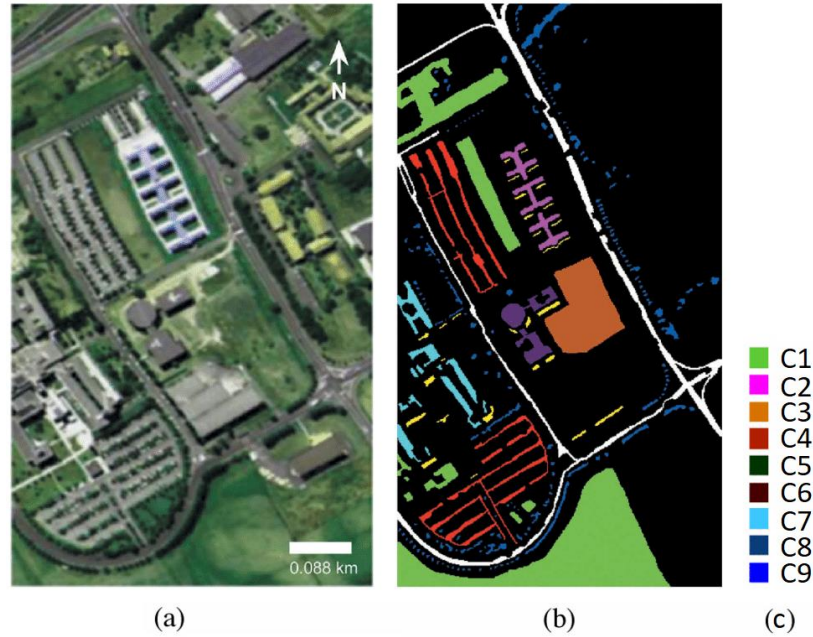
Botswana ve KSC verilerinden farklı olarak Salinas veri seti, Bölüm 6.3'te yer alan çalışmada ayrıntıları bulunan derin öğrenme yöntemleri ile kıyaslama yapılabilmesi için yalnızca bu çalışmada kullanılmaktadır.

Çizelge 6.4. Salinas doğrulama verisi sınıfları ve örnek sayıları.

Etiket	Salinas	
	Sınıf	Örnek Sayısı
C1	Yeşil brokoli otları 1 (Broccoli green weeds 1)	2009
C2	Yeşil brokoli otları 2 (Broccoli green weeds 2)	3726
C3	Nadas (Fallow)	1976
C4	Pürüzlü sürülecek nadas alanı (Fallow rough plow)	1394
C5	Düzgün nadas alanı (Fallow smooth)	2678
C6	Anız (Stubble)	3959
C7	Kereviz (Celery)	3579
C8	Olgunlaşmamış üzüm (Grapes untrained)	11271
C9	Hazırlanmış bağ toprağı (Soil vineyard develop)	6203
C10	Kurumuş mısır otları (Corn senesced green weeds)	3278
C11	Roman marulu 4 hft (Romaine lettuce 4 wk)	1068
C12	Roman marulu 5 hft (Romaine lettuce 5 wk)	1927
C13	Roman marulu 6 hft (Romaine lettuce 6 wk)	916
C14	Roman marulu 7 hft (Romaine lettuce 7 wk)	1070
C15	Olgunlaşmamış üzüm bağı (Vineyard untrained)	7268
C16	Dikey üzüm bağı serası (Vineyard vertical trellis)	1807
<b>Toplam</b>		<b>54129</b>

### 6.1.5. Pavia Üniversitesi

Pavia Üniversitesi (Pavia University - PU) veri seti, ROSIS [91] algılayıcısı tarafından, kuzey İtalya'daki Pavia üniversitesi üzerinden elde edilmektedir [90]. ROSIS, 430 ila 960 nm arası dalga boyunca 4 nm'lik pencereler ile ölçüm yapabilmektedir. Pavia Üniversite'sine ait doğrulama verisi ve sınıf görselleri Şekil 6.5'te yer almaktadır.



Şekil 6.5. a) PU renkli görüntüsü b) doğrulama verisi c) sınıflar [33].



Çizelge 6.5’te gösterilen genel sınıf grupları, gürültülü bantların çıkarılması ile 103 banttan oluşmaktadır. Sınıflandırma için kullanılacak veriler 9 sınıf ve 42776 örnekten oluşmaktadır. Salinas veri seti ile aynı çalışmada kullanılmaktadır.

Çizelge 6.5. PU doğrulama verisi sınıfları ve örnek sayıları.

Etiket	Pavia Üniversitesi	
	Sınıf	Örnek Sayısı
C1	Asfalt (Asphalt)	6631
C2	Çayırlar (Meadows)	18649
C3	Çakıl (Gravel)	2099
C4	Ağaçlar (Trees)	3064
C5	Boyalı metal levhalar (Painted metal sheets)	1345
C6	Çıplak toprak (Bare soil)	5029
C7	Zift (Bitumen)	1330
C8	Tuğlalar (Bricks)	3682
C9	Gölgeler (Shadows)	947
<b>Toplam</b>		<b>42776</b>

Deneysel çalışmalarda sonuçlar her bir sınıf için ayrı olarak sunulmaktadır. Her algoritmanın parametreleri, en iyi sınıflandırma doğruluğunu elde etmek için belirlenmiştir. Bu çalışmada, tanımlanmış nesne sınıflarının veya toprak arazi örtüsünün, alt boyut uzayında ayrılabilir hacimler kullanıp kullanmadığını ve başarılı bir şekilde ayırt edilemeyeceğini belirlemek için en yakın komşu (1NN) sınıflandırma performans doğruluğu kullanılmıştır. 1NN algoritması verimliliği ve sınıf dağılımları hakkında hiçbir varsayımda bulunmaması sebebiyle karşılaştırmalı analizlerde tercih edilmiştir. Ayrıca, algoritma kullanımında hiçbir model parametresine gereksinim duymamaktadır.

Deneysel çalışmalarda farklı yerleştirme uzayları kullanılmış ve gömülü haritalar rastgele eğitim (%70) ve test (%30) kümelerine bölünmüştür. Ayrıca toplam doğruluk oranları için Kappa istatistiği (KS) [92] kullanılarak sınıflandırma doğruluklarının ortalaması alınmıştır. KS formülasyonu Eşitlik 6.1’de ifade edilmektedir.

$$KS = \frac{N \sum_{c=1}^{|C|} t_{cc} - \sum_{c=1}^{|C|} t_{c+} t_{+c}}{N^2 - \sum_{c=1}^{|C|} t_{c+} t_{+c}} \quad (6.1)$$

Burada  $N$  test örneklerinin sayısını,  $t_{cc}$   $C$  sınıfındaki doğru sınıflandırılmış örnek sayısını,  $t_{c+}$   $C$  sınıfı olarak etiketlenmiş test örneklerinin sayısını ve  $t_{+c}$   $C$  sınıfına ait

olduğu tahmin edilen örnek sayısını göstermektedir.  $|C|$ , verilerdeki toplam sınıf sayısını belirtmektedir.

## 6.2. HİPERSPEKTRAL GÖRÜNTÜLERİN SINIFLANDIRILMASINDA FARKLI BOYUT İNDİRGEME YÖNTEMLERİNİN KARŞILAŞTIRILMASI

Bu çalışmada, olasılıksal temelli doğrusal olamayan boyut indirgeme yöntemlerden SNE ve t-SNE yöntemleri ile kuvvet alan formülasyonuna dayanan çok boyutlu bir alan yerleştirme algoritması olan MAFE yöntemi ve MAFE yönteminin geliştirilmiş bir varyasyonu olan Enhanced-MAFE [93] yöntemleri kullanılarak analizler gerçekleştirilmiştir. Bu yöntemler hiperspektral görüntülerden elde edilmiş olan Botswana ve KSC verileri üzerine uygulanmıştır [27].

Çizelge 6.6. Botswana verisi için sınıflandırma sonuçları.

Etiket	Sınıflandırma Doğruluğu (%)			
	SNE	t-SNE	MAFE	Enhanced MAFE
C1	<b>100</b>	<b>100</b>	99,89	<b>100</b>
C2	<b>100</b>	<b>100</b>	<b>100</b>	99,58
C3	76,19	90,91	90,48	<b>97,70</b>
C4	94,74	94,74	79,46	<b>100</b>
C5	94,44	88,89	<b>99,17</b>	96,76
C6	68,42	90	94,19	<b>100</b>
C7	96,18	93,75	93,24	<b>100</b>
C8	91,67	<b>100</b>	<b>100</b>	<b>100</b>
C9	91,67	90,91	90,91	<b>99,13</b>
<b>KS</b>	89,26	93,55	94,27	<b>99,07</b>

Botswana verisi için Çizelge 6.6'da verilen sonuçlar incelendiğinde En iyi KS sonucunu Enhanced-MAFE yönteminin verdiği görülmektedir. Ayrıca 9 sınıfın 7'si için yine Enhanced-MAFE diğer yöntemlerden daha başarılı sonuçlar vermektedir. Diğer iki sınıfta ise kıyaslanan yöntemlere yakın sonuçlar aldığı görülmektedir. Sınıf ayrımlarının en zor olduğu sınıflar C3 ve C6 sınıflarıdır.

Çizelge 6.7'de verilen KSC verileri için KS oranı tüm yöntemlerde yakın olmakla beraber en başarılı yöntem yine MAFE varyasyonu olan Enhanced-MAFE olarak görülmektedir. Enhanced-MAFE 13 sınıfın 8'i için en iyi sınıflandırma doğruluk

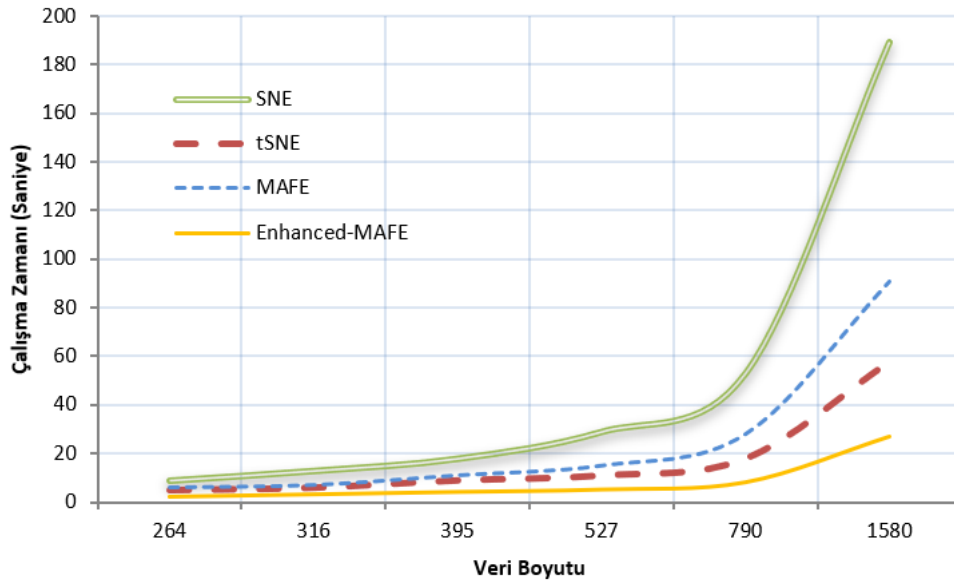


oranlarına sahiptir. Sınıf ayrımlarının en zor olduđu sınıflar C4, C5, C6 ve C8 sınıflarıdır.

Çizelge 6.7. KSC verisi için sınıflandırma sonuçları.

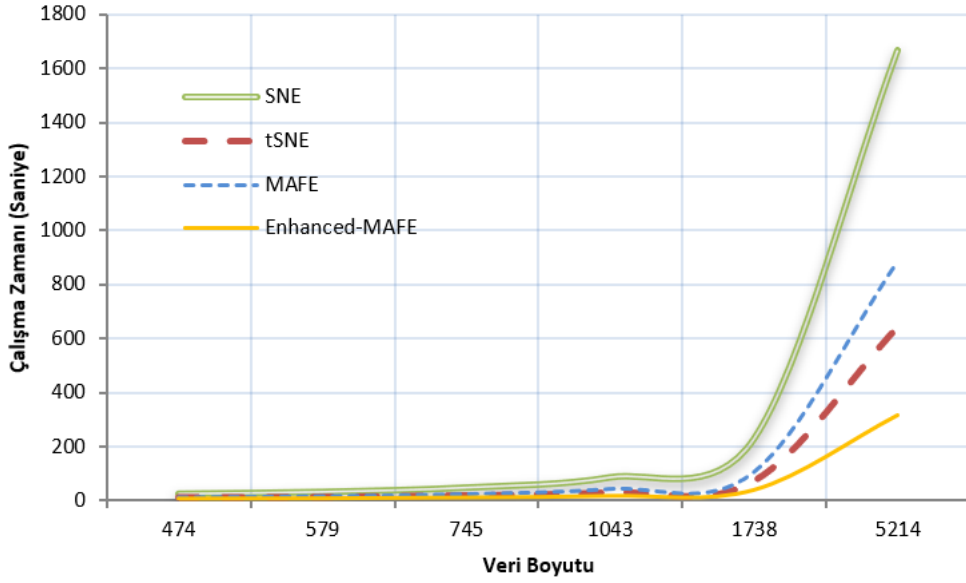
Etiket	Sınıflandırma Doğruluđu (%)			
	SNE	t-SNE	MAFE	Enhanced MAFE
C1	<b>91,20</b>	90,34	65,26	68,54
C2	85,34	<b>87,04</b>	79,03	77,42
C3	80,40	78,33	94,52	<b>98,61</b>
C4	51,90	45,61	<b>68,75</b>	64,62
C5	41,20	41,03	93,33	<b>95,24</b>
C6	39,30	35,71	26,98	<b>44,26</b>
C7	<b>82,77</b>	80,77	35,29	67,65
C8	<b>63,62</b>	60	43,75	42,98
C9	93,56	91,74	<b>99,30</b>	<b>99,30</b>
C10	93,68	93,68	95,08	<b>97,56</b>
C11	93,81	93,81	99,11	<b>100</b>
C12	81,91	80,51	97,60	<b>98,39</b>
C13	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
<b>KS</b>	81,90	80,10	79,92	<b>82,51</b>

Her iki veri seti için sınıflar bağımsız olarak incelendiğinde ağaçlık, dere kenarı ve bataklık bölgelerdeki sınıflandırmanın zor olduđu görülmektedir. Su ise ayırt edilebilmesi en kolay sınıf olarak karşımıza çıkıyor. Ayrıca her iki veri seti içinde Enhanced-MAFE yönteminin sınıflar arasındaki ayrımı oldukça başarılı yapabildiği gözlemlenmektedir.



Şekil 6.6. Botswana verisi için çalışma zamanı.

Şekil 6.6’da verilen veri boyutuna bağlı çalışma zamanı grafiği incelendiğinde, yöntemlerin veri boyutları arttıkça giderek ivmelenen bir çalışma zamanı artış grafiği izledikleri görülmektedir. SNE yönteminin büyük boyutlardaki veri gruplarını analiz etmede diğer yöntemlere kıyasla çalışma zamanı oldukça yüksektir. Enhanced-MAFE diğer yöntemlere göre daha az etkilenen bir grafik sunmaktadır. Grafiklerdeki ayırım veri boyutu arttıkça daha çok ayrılmaktadır.



Şekil 6.7. KSC verisi için çalışma zamanı.

Şekil 6.7’de verilen KSC verileri de Botswana ile benzer şekilde SNE yönteminin çalışma zamanı, veri boyutu arttıkça artan bir ivme ile yükselmektedir. Enhanced-MAFE, KSC verileri içinde en hızlı yöntem olarak karşımıza çıkmaktadır. KSC veri boyutunun Botswana’ya göre daha büyük olmasına rağmen karşılaştırılan yöntemlerin çalışma zamanlarının benzer özellikler gösterdiği görülmektedir. Botswana’da da olduğu gibi yine grafikler verilerin örneklem uzayındaki sayısı arttıkça veri boyutuna bağlı olarak ayrılmaktadır.

Uygulanan algoritmaların çalışma zamanları hesaplanırken en yüksek doğruluk oranını elde etmek için gerekli iterasyon sayıları kullanılmıştır. Bu anlamda incelendiğinde Enhanced-Mafe yöntemi diğer yöntemlere kıyasla çok daha düşük iterasyon sayısı ile maksimum doğruluk oranına ulaşmaktadır. Bu da çalışma zamanına olumlu bir katkı sağlamaktadır.

Verilerin örnekleme doğruluk oranlarını etkileyen en önemli faktörlerden biridir. Örnekleme yapmadan tüm verilerin seçilmesi ise çalışma zamanını oldukça arttırmaktadır. Hem Şekil 6.6 hem de Şekil 6.7’de verilen grafikler incelendiğinde yöntemlerin düşük boyutlu veriler kullanıldığında benzer çalışma zamanı gerektirdikleri, ancak boyutlar arttıkça bu durumun değiştiği görülmektedir. Bunlara ek olarak, çalışma zamanı grafiklerinde her iki veri setinde de oransal olarak benzer sonuçlar elde edilmiştir. Bu sonuçlar bize yöntemlerin çalışma zamanının veriden bağımsız olduğunu da göstermektedir. Sonuç olarak veri boyutu doğrudan çalışma zamanını etkilemektedir. Ayrıca MAFE ve Enhanced-MAFE algoritmaları diğer manifold yerleştirme yöntemlerine kıyasla daha yüksek sınıflandırma doğruluğu elde etmektedir.

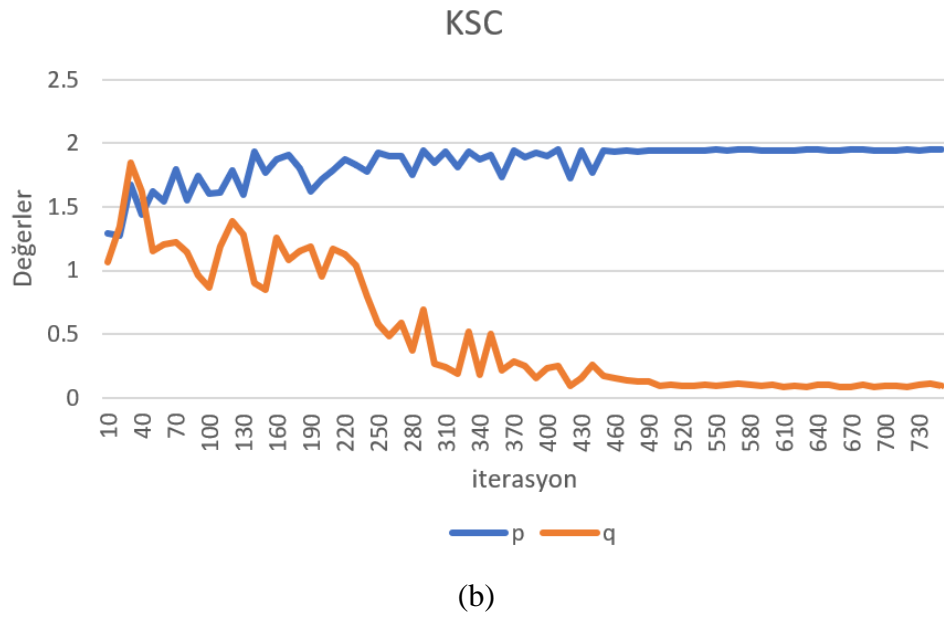
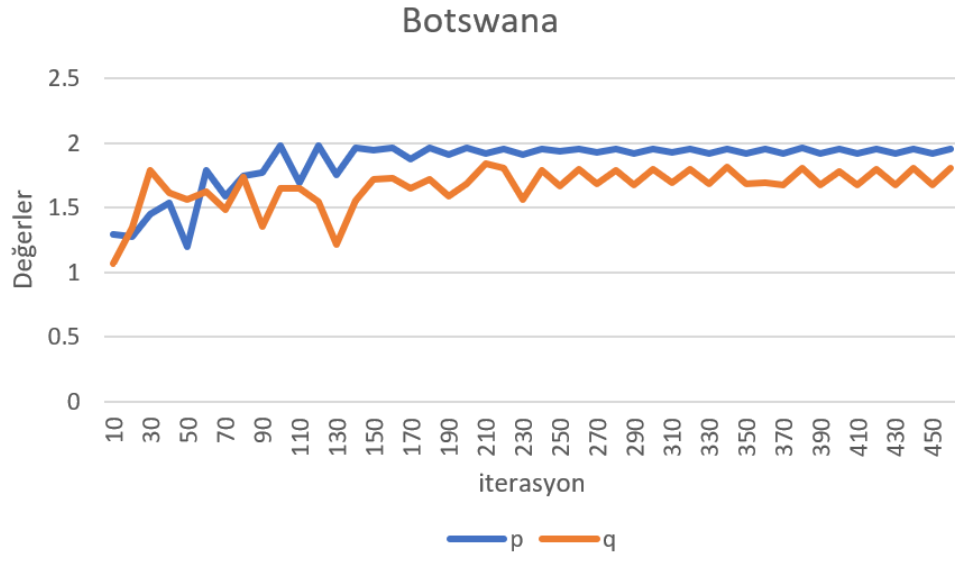
### **6.3. OPTİMİZASYON TABANLI MANİFOLD YERLEŞTİRME YÖNTEMİ İLE HİPERSPEKTRAL GÖRÜNTÜLERİN SINIFLANDIRILMASI VE GÖRSELLEŞTİRİLMESİ**

MAFE yönteminin iki ana eksikliği vardır. Bunlardan ilki, optimum sınıflandırma ve görselleştirme sonuçları elde etmek için özellikle seçilmesi önemli olan iki parametre olmak üzere deneysel parametrelerin belirlenmesindeki zorluktur. İkincisi, karmaşık verilerde daha yaygın olarak ortaya çıkan kalabalık sorunudur. Kalabalık sorunu, yüksek boyutlu uzaydaki verilerin alt boyutlu uzaya gömülmesi sırasında veri noktalarının üst üste gelmesiyle ortaya çıkar [16,18].

Bu çalışmada MAFE yöntemindeki bu problemlerin üstesinden gelmek için literatürde en yaygın optimizasyon yöntemi olan PSO algoritması kullanılmaktadır. MAFE yöntemindeki parametre belirleme problemi PSO tabanlı bir yaklaşımla çözülmektedir. Önerilen yöntem, manifold yerleştirme alanında ilk optimizasyon tabanlı çözümdür. Optimal parametrelerin belirlenmesi ile Botswana, KSC ve diğer popüler hiperspektral veri setleri için doğruluk oranları mevcut yöntemlere göre daha başarılı sonuçlar vermiştir. Ayrıca, kalabalık sorununa da başarılı bir çözüm sunulmaktadır.

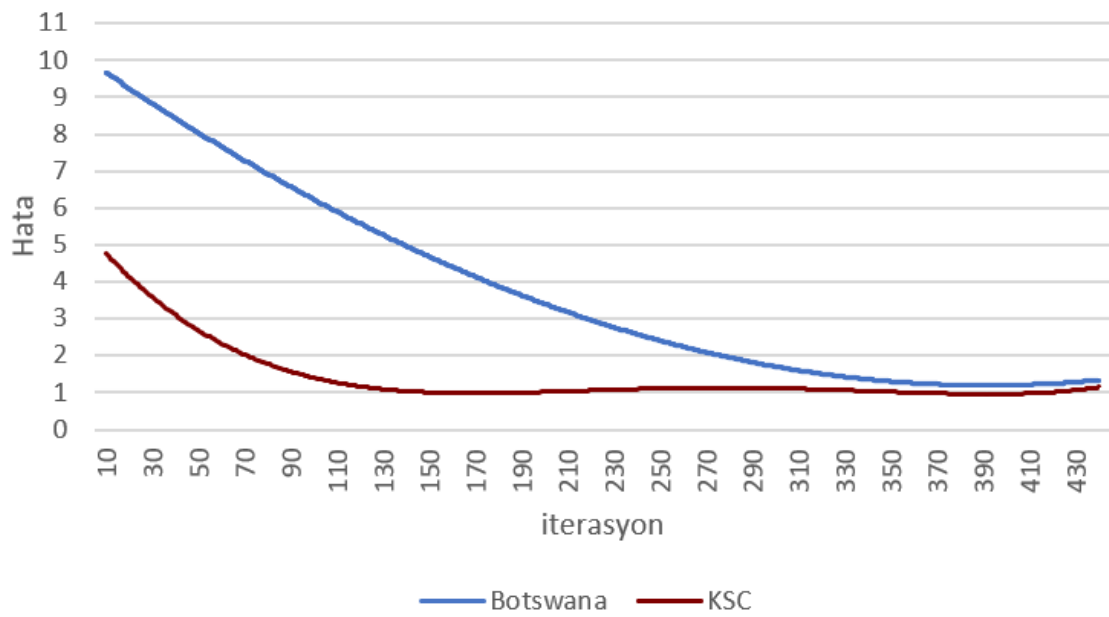
Eşitlik 4.23’te verilen denklemde, maliyet fonksiyonunun birçok değişken parametresi vardır. Bu parametrelerin sonucu büyük ölçüde etkilediği keşfedilmiştir. Tüm

parametreler için yapılan analiz sonucunda sonucu en çok etkileyen parametreler  $p$  ve  $q$  olarak belirlenmiştir. Bu parametreler yapılan çalışmalarda [0-2] aralığında seçilmektedir [18], ancak net bir değer verilmediği için yapılan deneysel çalışmalarda optimum yerleştirme sonuçları elde edilemez. Bu aralıkta, global minimum değerlerini bulmak için PSO tabanlı bir yaklaşım önerilmektedir. Bu yaklaşım parametre optimizasyonu için PSO algoritmasından faydalanarak  $p$  ve  $q$  parametrelerinin optimum değerlerini tespit etmektedir.



Şekil 6.8. PSO parametre optimizasyon grafikleri a) Botswana b) KSC.

PSO algoritması ile optimal  $p$  ve  $q$  yakınsama grafikleri Şekil 6.8’de verilmektedir. Grafiklerde, maksimum sınıflandırma doğruluğu için Botswana verilerinde her iki parametrenin de 2’ye yakınsadığı görülmektedir. KSC verileri için  $p$  yine 2’ye yakınsarken  $q$  parametresi 0’a yakınsamaktadır. Deneysel çalışmaların sonucunda  $p$  ve  $q$  parametrelerinin optimal değerleri, Botswana verileri için sırasıyla 1,89 ve 1,87, KSC için yine sırasıyla 1,94 ve 0,08 olarak bulunmaktadır. PSO algoritmasının çalışması esnasında hata oranınının 1’e yakınsama grafiği ise Şekil 6.9’da sunulmaktadır.



Şekil 6.9. PSO parametre optimizasyon grafikleri için hata yakınsaması.

Şekil 6.9’da verilen parametre optimizasyon grafiğinde, optimal yakınsamanın Botswana verileri için 400 iterasyon ve KSC verileri için 100 iterasyon civarında olduğu görülmektedir.

Botswana verileri için sınıflandırma sonuçları Çizelge 6.8’de verilmektedir. Kalın harflerle yazılan değerler en yüksek sınıflandırma doğruluğunu gösterir. Sunulan tüm yöntemlerin yüksek doğruluk oranlarına sahip olduğu görülmektedir. Önerilen PSO-MAFE yöntemi, Botswana verileri için diğer bilinen yöntemlere kıyasla %99,37 ile en yüksek sınıflandırma doğruluğuna ulaşmaktadır. Enhanced-MAFE de buna yakın bir değer sunarken, diğer yöntemler en yüksek değer olarak %95,89’da kalmaktadır.

Çizelge 6.8. Botswana verisi için sınıflandırma sonuçları.

Etiket	Sınıflandırma Doğruluğu (%)									
	LFDA	LE	LLE	ISOMAP	sSNE	SNE	t-SNE	MAFE	Enhanced MAFE	PSO-MAFE (Önerilen)
c1	<b>100</b>	<b>100</b>	98,79	<b>100</b>	93,41	<b>100</b>	<b>100</b>	98,27	<b>100</b>	<b>100</b>
c2	<b>100</b>	97,06	94,26	<b>100</b>	93,45	98,39	96,74	99,58	99,58	<b>100</b>
c3	91,55	80,28	66,75	91,55	92,10	73,68	88,89	90,68	97,70	<b>98,45</b>
c4	98,11	96,30	96,78	<b>100</b>	92,81	96,85	94,27	84,58	<b>100</b>	99,49
c5	94,55	80	85,30	96,36	90,71	93,44	90,89	98,17	96,76	<b>100</b>
c6	94,92	69,49	67,38	89,83	92,24	67,72	82,64	94,79	<b>100</b>	99,49
c7	97,92	77,08	85,98	97,92	92,23	97,18	96,75	93,84	<b>100</b>	<b>100</b>
c8	97,37	94,59	93,12	<b>100</b>	92,70	92,87	99,17	<b>100</b>	<b>100</b>	<b>100</b>
c9	97,37	97	97,41	<b>100</b>	93,84	92,87	89,97	89,97	<b>99,13</b>	97,91
<b>KS</b>	94,55	88,45	84,45	95,89	91,59	88,21	93,75	94,89	99,07	<b>99,37</b>

Deneysel çalışmalarda Botswana için en düşük sınıflandırma doğruluğu c3 ve c6 ile elde edilmiştir. Bu sınıflar, ayrılması en zor olan Dere kenarı (Riparian) ve Ağaçlık (Woodlands) sınıflarıdır. Bu sınıflar için Enhanced MAFE ve PSO-MAFE, en yüksek sınıflandırma doğruluk değerine sahiptir. Diğer sınıflar için önerilen PSO-MAFE yöntemi yine 9 sınıfın 6'sı için en yüksek oranları vermektedir.

Çizelge 6.9. KSC verisi için sınıflandırma sonuçları.

Etiket	Sınıflandırma Doğruluğu (%)									
	LFDA	LE	LLE	ISOMAP	sSNE	SNE	t-SNE	MAFE	Enhanced MAFE	PSO-MAFE (Önerilen)
c1	92	72	80,38	90,33	94,88	91,49	88,83	65,60	68,54	<b>100</b>
c2	82,61	43,48	51,61	62,22	90	84,78	84,44	79,42	77,42	<b>100</b>
c3	88,46	38,48	39,09	80,77	91,89	79,80	79,27	85,20	98,61	<b>98,99</b>
c4	68	28	18,15	56	85,16	52,17	51,39	68,82	64,62	<b>97,29</b>
c5	52,94	29,41	22,14	47,06	91,14	47,83	41,86	76,40	95,24	<b>98,28</b>
c6	72	28	34,43	40	90,95	44,90	42,31	28,12	44,26	<b>98,13</b>
c7	<b>100</b>	36,36	38,57	90,91	91,18	84,62	80,87	73,33	67,65	<b>100</b>
c8	88,10	35,71	59,11	69,05	91,79	62,26	63,93	44,71	42,98	<b>95,46</b>
c9	96,15	75	67,30	96,15	92,31	92,63	92,23	98,60	99,30	<b>100</b>
c10	<b>100</b>	80,49	68,28	92,68	92,95	88,51	93,54	94,32	97,56	<b>100</b>
c11	97,56	85,37	87,23	97,56	92,30	93,55	95,54	98,11	<b>100</b>	<b>100</b>
c12	98,04	80,39	74,92	92,16	93,75	82,80	78,16	96,87	98,39	<b>100</b>
c13	<b>100</b>	98,39	<b>100</b>	<b>100</b>	93,89	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
<b>KS</b>	89,20	76,71	65,31	86,40	91,71	80,95	79,94	82,42	82,51	<b>98,17</b>

PSO-MAFE benzer şekilde Çizelge 6.9'da KSC verileri için en yüksek sınıflandırma doğruluğunu göstermektedir. Önerilen bu yöntem %98,17 ile en yüksek sınıflandırma

doğruluğuna ulaşırken diğer yöntemler en yüksek %91,71’de kalmaktadır. KSC veri setinde sınıf ayrımlarının en zor olduğu c4, c5, c6 ve c8 sınıfları için yine PSO-MAFE yöntemi en yüksek sınıflandırma başarısını göstermektedir. Sınıf ayrımlarının daha zor ve karmaşık olduğu KSC verilerinde önerilen yöntem 13 sınıfın tümünde en yüksek sınıflandırma doğruluk oranlarını elde etmektedir.

Çizelge 6.10. IP, PU ve Salinas verileri için sınıflandırma sonuçları.

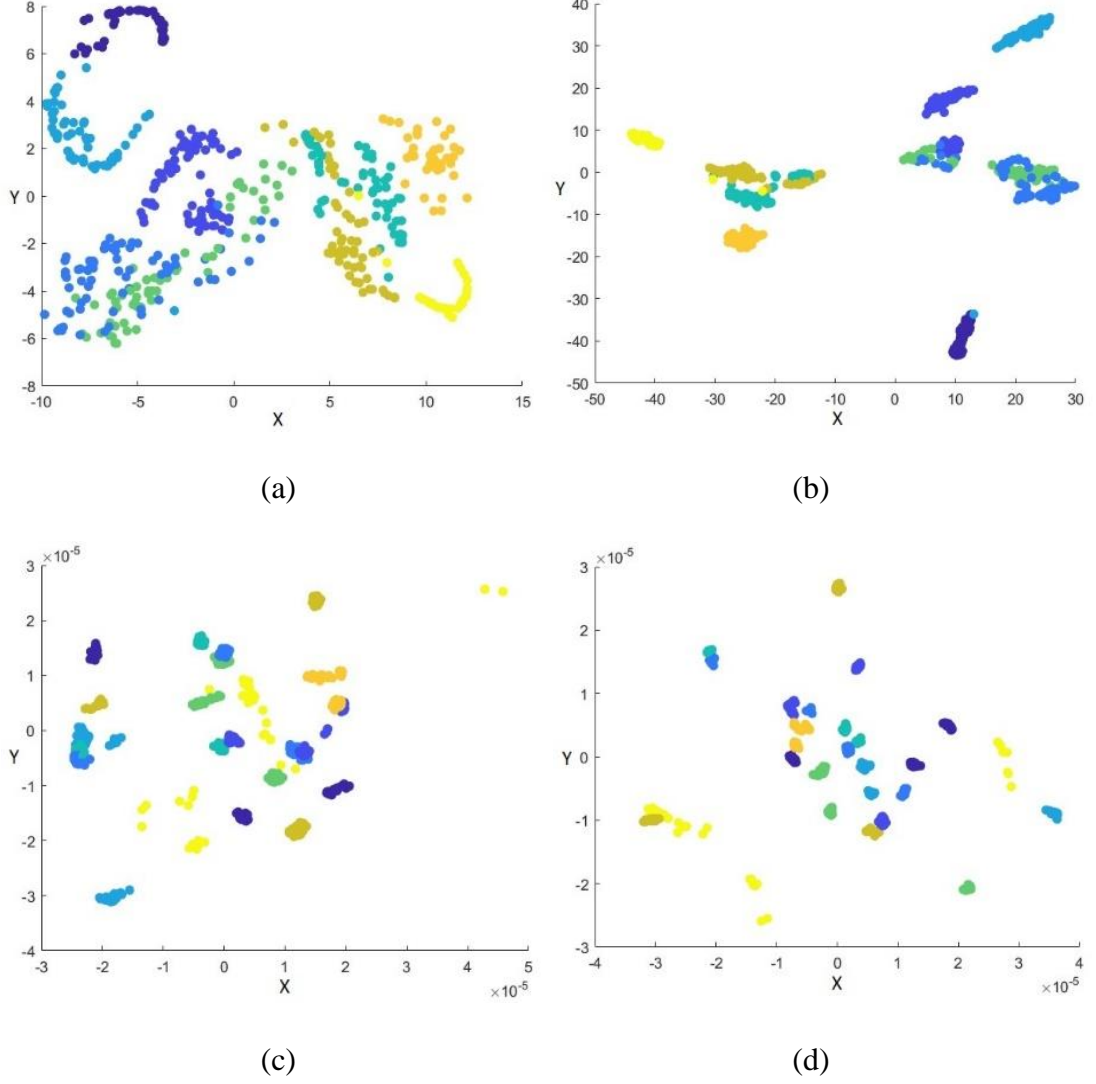
Data Set	Sınıflandırma Doğruluğu (%)										
	LFDA	ISOMAP	sSNE	t-SNE	MAFE	Enhanced MAFE	Resnet 4	DCNN	3D CNN	DR CNN	PSO MAFE (Önerilen)
IP	89,15	86,25	90	75,40	80,75	82,45	95,54	97,56	95,54	<b>98,20</b>	97,70
PU	88,50	86,18	90,96	76,97	82,38	81,98	98,04	96,15	99,03	<b>99,41</b>	98,45
Salinas	89,87	88,50	92,55	80,09	81,45	83,10	98,60	94,87	98,39	98,14	<b>99,30</b>

Çizelge 6.10, literatürdeki diğer popüler hiperspektral veri setleri için elde edilen sonuçları göstermektedir. Önerilen yöntem, tüm veriler için manifold yerleştirme yöntemleri arasında en yüksek sınıflandırma başarısını vermektedir. Bu bize MAFE yönteminin optimal parametreleri seçerek tüm verilerde kabul edilebilir sonuçlar verdiğini göstermektedir.

Çizelge 6.10’da ayrıca Resnet-4 [94], Deep-CNN [95], 3D-CNN [96] ve DR-CNN [97] gibi popüler derin öğrenme yaklaşımlarının sonuçları da verilmektedir. Önerilen yöntem de bu popüler derin öğrenme yöntemlerine benzer veya daha iyi sonuçlar elde etmektedir. Burada iyi bir manifold yerleştirme işleminin, en temel sınıflandırıcılardan biri olan INN algoritması kullanarak, son yılların popüler derin öğrenme yaklaşımları ile rekabet edebileceği görülmektedir.

Botswana hiperspektral verileri için görselleştirme sonuçları Şekil 6.10’da verilmektedir. SNE ve t-SNE yöntemlerinin kalabalık probleminden etkilendiği ve sınıfların örtüştüğü görülmektedir. MAFE sınıf ayrımlarında başarılı olmasına rağmen, ayırt edilmesi zor olan sınıflarda net bir ayırım yapamamaktadır. Önerilen PSO-MAFE yönteminde ise, tüm sınıflar yüksek hassasiyetle ayrılmakta ve neredeyse hiç örtüşme görülmemektedir. KSC verileri, Botswana’ya kıyasla kalabalıklaşma sorunu açısından çok daha zorlayıcıdır [18]. Diğer tüm yöntemler, KSC verilerinde bu sorundan

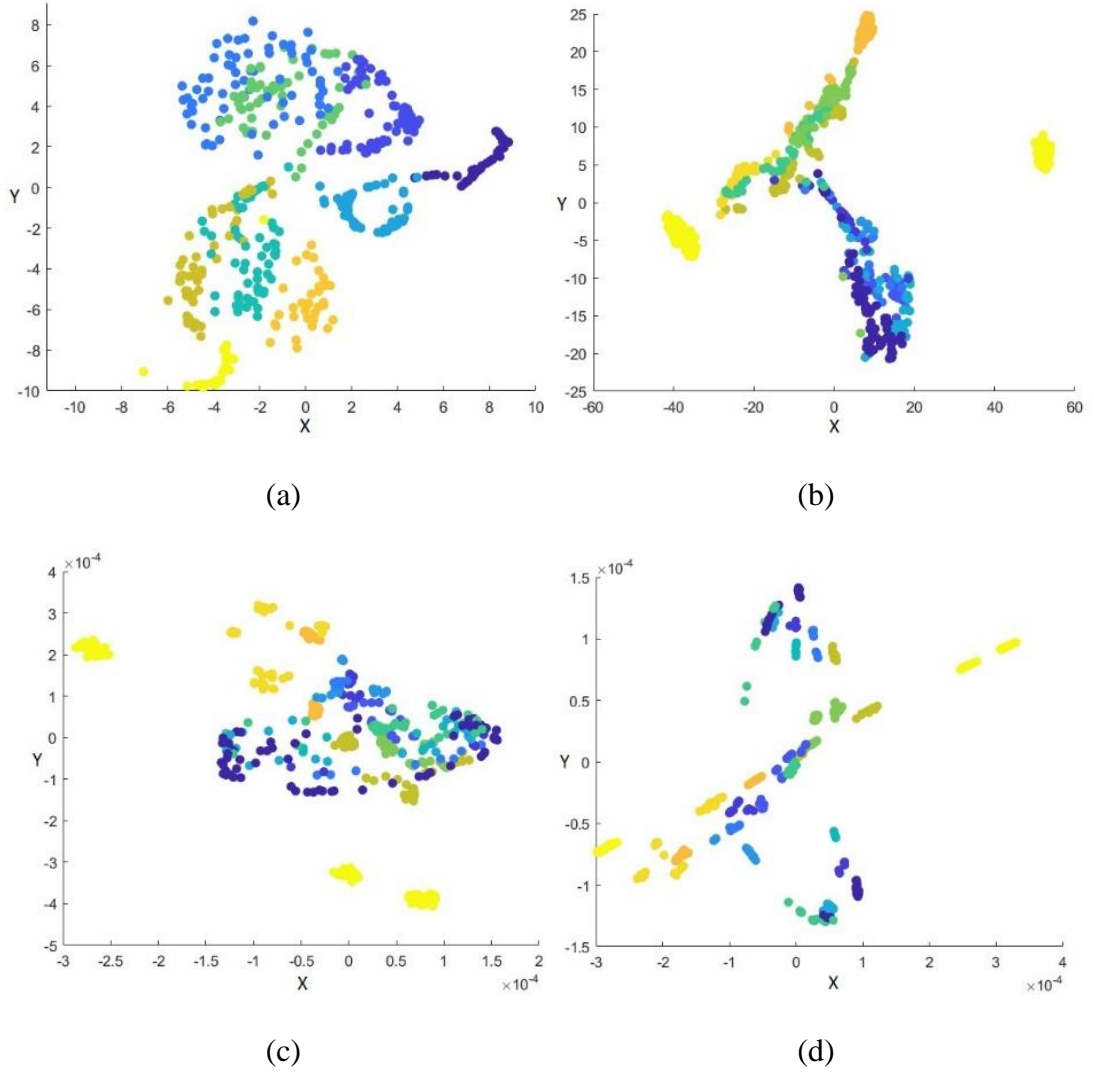
etkilenmektedir. Şekil 6.11'deki görselleştirme sonuçları, önerilen yöntemin bu sorunu KSC verisi içinde büyük ölçüde ortadan kaldırdığını göstermektedir.



Şekil 6.10. Botswana verisi için görselleştirme sonuçları a) SNE b)  $t$ -SNE c) MAFE d) PSO-MAFE.

Burada yine SNE,  $t$ -SNE ve MAFE algoritmaları sınıf ayrımlarında kalabalık problemine takılmış durumdadır. Önerilen yöntem, birkaç küçük temas dışında sınıflandırma verilerinin tümünü hassas bir şekilde KSC verileri içinde gerçekleştirmektedir. Ayrıca her iki veri seti görsellerinin eksen değerlerine bakıldığında MAFE ve önerilen yöntemin bu sınıf ayırma uzayını çok daha küçük bir koordinat aralığında gerçekleştirdiği görülmektedir.





Şekil 6.11. KSC verisi için görselleştirme sonuçları a) SNE b)  $t$ -SNE c) MAFE d) PSO-MAFE.

Sonuç olarak önerilen optimizasyona dayalı çok boyutlu alan gömme yöntemi ile Botswana verileri için %99,37 ve KSC verileri için %98,17 ile diğer tüm manifold yerleştirme yöntemlerine kıyasla en yüksek sınıflandırma doğruluğuna ulaşılmaktadır. Ayrıca MAFE yönteminde parametre belirleme sorunu PSO yaklaşımı kullanılarak ortadan kaldırılmıştır. Önerilen yöntem, özellikle veri karmaşıklığı arttıkça ortaya çıkan kalabalık sorununa da başarılı bir çözüm sunmaktadır. Deneysel sonuçlar, önerilen yaklaşımın bilinen diğer yerleştirme yöntemlerinden daha üstün olduğunu ortaya koymaktadır. Ayrıca son yıllarda popüler olan derin öğrenme yaklaşımlarına karşı da etkili sonuçlar alarak iyi bir manifold yerleştirme ön işleminin sınıflandırma doğruluğuna etkisi ortaya konulmaktadır.

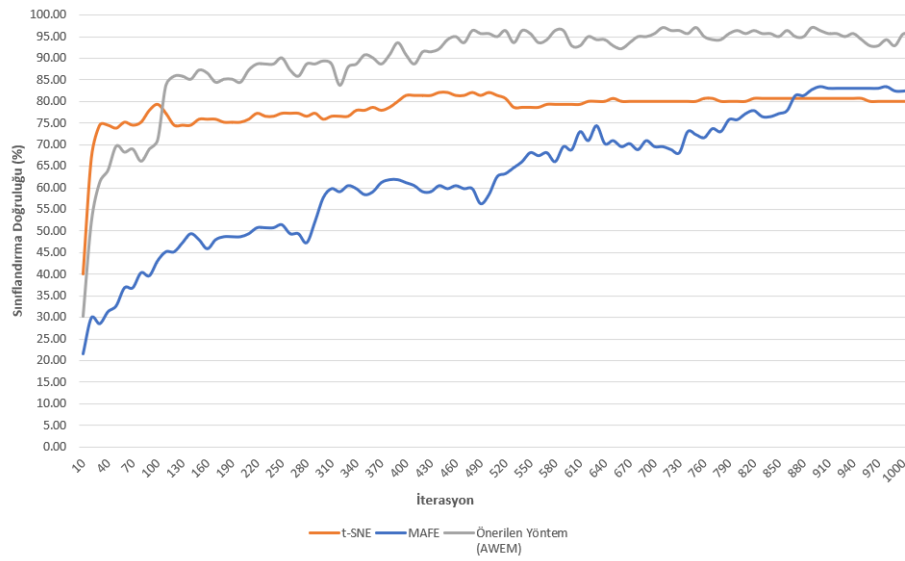
#### 6.4. ADAPTİF AĞIRLIKLANDIRILAN MANİFOLD YERLEŞTİRME YÖNTEMİ İLE HİPERSPEKTRAL GÖRÜNTÜLERİN SINIFLANDIRMASI

Bu çalışmada, mevcut manifold öğrenme yöntemlerinin güçlü yanlarından yararlanan, geliştirilmiş adaptif ağırlıklandırılan bir manifold yerleştirme yöntemi (AWEM) önerilmektedir. Çalışmada ilk olarak, mevcut yöntemlerin çıktı uzaylarında kullanılan maliyet fonksiyonları ve dağıtım fonksiyonları üzerinde bazı iyileştirmeler yapılmıştır. Ardından, iki popüler yöntem olan t-SNE yönteminin hızlı yakınsama yeteneği ile MAFE yönteminin yüksek sınıflandırma doğruluğu yeteneği gibi güçlü özelliklerini kullanabileceğimiz, Eşitlik 5.9’da verilen AWEM yöntemi uygulanmaktadır. Önerilen yöntemin karşılaştırmalı analizi, KSC veri seti ile yapılmaktadır. Deneysel sonuçlar, önerilen yöntemin birçok yönden üstün olduğunu göstermektedir. Bu yöntemle, mevcut yöntemlere göre daha iyi sınıflandırma doğruluğu ve görselleştirme elde edilirken, aynı zamanda kalabalık sorunu da ortadan kalkmaktadır. Ayrıca AWEM yönteminin erken çıkışlı (early-exit) bir versiyonu (AWEM-EE) kullanılarak mevcut yöntemlerle benzer uygulama süreleri elde edilmektedir.

Çizelge 6.11. KSC verisi için sınıflandırma sonuçları.

Etiket	Sınıflandırma Doğruluğu (%)						AWEM (Önerilen)
	LFDA	LLE	ISOMAP	SNE	t-SNE	MAFE	
C1	91,41	78,61	93,33	95,20	89,19	63,64	<b>95,24</b>
C2	80,62	53,15	62,50	87,50	84,62	81,82	<b>100</b>
C3	86,23	38,92	84,62	76,92	81,82	84,62	<b>100</b>
C4	66,40	28,10	53,12	53,85	62,50	71,43	<b>85,71</b>
C5	56,12	23,19	46,26	46,15	44,44	75,00	<b>100</b>
C6	76,20	31,55	40,34	42,86	54,55	37,50	<b>83,33</b>
C7	<b>100</b>	39,70	67,50	82,35	84,62	71,43	<b>100</b>
C8	62,22	61,18	72,12	73,33	63,64	42,86	<b>72,73</b>
C9	95,20	69,10	98,23	93,75	93,33	94,59	<b>100</b>
C10	98,22	68,32	93,33	87,50	95,45	93,33	<b>100</b>
C11	94,16	89,44	94,59	93,33	92,59	94,74	<b>100</b>
C12	97,56	73,98	95,20	93,75	82,61	91,30	<b>100</b>
C13	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
<b>KS</b>	87,25	66,72	84,32	79,59	80,25	82,67	<b>95,63</b>

KSC verileri için sınıflandırma sonuçları Çizelge 6.11’de verilmektedir. Önerilen yöntem %95,63 ile en yüksek sınıflandırma başarısını elde etmektedir. t-SNE ve MAFE yöntemleri ise [%80-83] aralığında çok yakın başarı oranları vermektedir. Önerilen yöntem ayrıca 13 sınıfın tümü için ayrı ayrı en yüksek oranları elde etmektedir. Ek olarak, diğer lineer ve lineer olmayan son teknoloji yöntemlerin sonuçları da Çizelge 6.11’de gösterilmiştir. Doğrusal bir yöntem olan LFDA, %87,25 sınıflandırma doğruluğu ile önerilen yönteme en yakın sınıflandırma doğruluğunu vermektedir.



Şekil 6.12. t-SNE, MAFE ve önerilen yöntem için sınıflandırma doğruluğu grafiği.

KSC verileri için yöntemlerin sınıflandırma doğruluk yakınsama grafiği Şekil 6.12’de verilmektedir. Bu grafikte t-SNE yönteminin 100 iterasyon civarında maksimum doğruluk oranlarına yakınsadığı görülmektedir. Daha sonra yatay bir grafik görüntülenir. MAFE yöntemi ise sürekli artan bir grafik sergilemektedir. Bunlara dayanarak, bu iki yöntemin güçlü yönlerinden yararlanabilecek adaptif ağırlıklandırılan bir manifold yerleştirme yöntemi önerilmektedir.

Önerilen yöntem ilk olarak t-SNE yöntemindeki hızdan yararlanmak için t-SNE’nin ağırlığını maksimize etmektedir. Bu sayede önerilen yöntem hızlı bir şekilde %80 sınıflandırma doğruluğunun üzerine çıkmaktadır. Yakınsama hatası minimuma yaklaştıkça MAFE yönteminin ağırlığı kademeli olarak artırılır ve t-SNE’nin ağırlığı

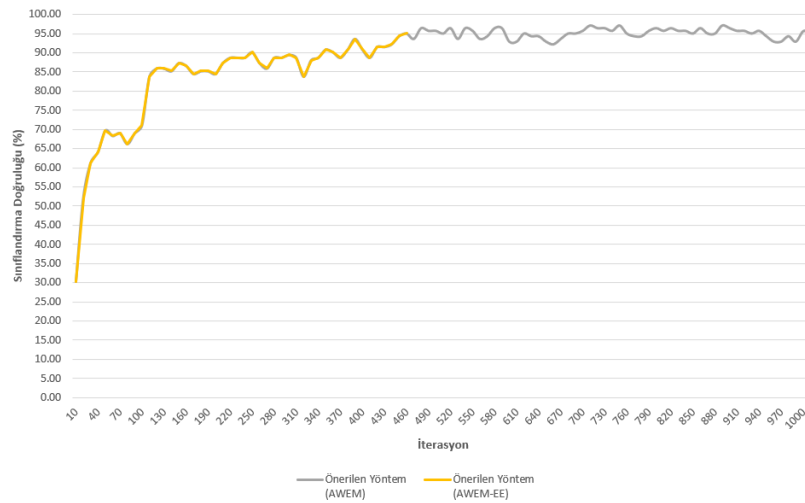
azaltılır. Böylece MAFE yönteminin ağırlığının artması ile sınıflandırma başarısı %95 civarlarına kadar yükselmeye devam etmektedir.

Yöntemler tarafından maksimum sınıflandırma doğruluğuna ulaşılırken harcanan uygulama süresi Çizelge 6.12’de verilmiştir. Tablodan da görüleceği gibi t-SNE yöntemi en hızlı çalışma süresine sahiptir. MAFE daha yüksek bir doğruluk oranına sahip olmasına rağmen, t-SNE’den daha yavaş çalışır. Önerilen yöntem 29,63 sn ile en yüksek çalışma süresine sahiptir.

Çizelge 6.12. Tüm yöntemler için çalışma zamanları.

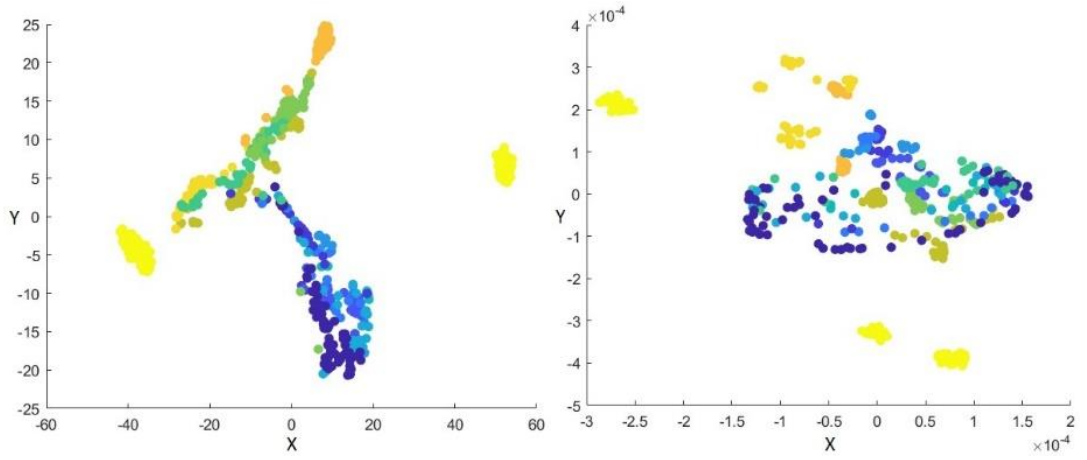
Yöntemler	Sınıflandırma Başarısı (%)	Çalışma Zamanları (sn.)
t-SNE	80,25	12,09
MAFE	82,65	22,55
Önerilen (AWEM)	95,63	29,63
Önerilen (AWEM-EE)	95,47	<b>16,24</b>

Oluşan bu zaman dezavantajını ortadan kaldırmak için önerilen yöntemin erken çıkış versiyonu olan AWEM-EE yöntemi önerilmiştir ve algoritma çalışma süreleri 16,24 saniye ile neredeyse %50 oranında iyileştirilmiştir. Önerilen AWEM-EE yöntemi, Şekil 6.13’te görüldüğü gibi yakınsama hatası minimuma yaklaştığında algoritmayı otomatik olarak durdurur. Bu sayede daha az iterasyon ile benzer başarı oranları elde edilebilir.



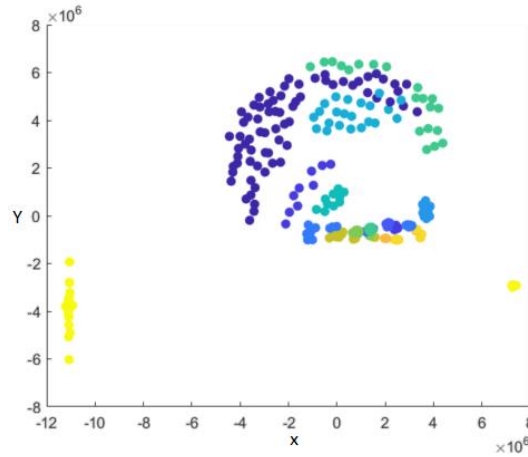
Şekil 6.13. Önerilen yöntemlerin iterasyona bağlı sınıflandırma doğruluğu grafiği.

KSC hiperspektral verileri için görselleştirme sonuçları Şekil 6.14'te verilmektedir. Grafikler, t-SNE yönteminin kalabalık probleminden büyük oranda etkilendiğini göstermektedir. Biri dışında hemen hemen tüm sınıflar çakışmaktadır. MAFE yönteminde de aynı problem t-SNE yöntemine göre daha az olmakla birlikte yine görülmektedir. Önerilen yöntemde ise sınıflar arası ayırım çok daha başarılı bir şekilde yapılmaktadır. Örtüşen sınıfların diğer yöntemlerden çok daha az olduğu grafiklerde açıkça görülmektedir. Sınıf ayırımının zor olduğu sınıflar dışında kalabalık probleminden etkilenen sınıflar oldukça az oranlardadır.



(a)

(b)



(c)

Şekil 6.14. KSC verisi için görselleştirme sonuçları a) t-SNE b) MAPE c) Önerilen yöntem.

Sonuç olarak, bu çalışma KSC hiperspektral veri seti üzerine uygulanmakta ve önerilen yöntem %95,63 ile bilinen diğer yerleştirme yöntemleri arasında en yüksek sınıflandırma doğruluğunu vermektedir. Ayrıca önerilen yöntemin erken çıkış versiyonu ile bu başarı oranı, t-SNE ve MAFE yöntemleri ile benzer bellek ve zaman maliyetleri ile elde edilmektedir. Bunlara ek olarak önerilen yöntem ile karmaşık verilerde ortaya çıkan kalabalık sorununa da etkili bir çözüm ortaya konulmaktadır.

## BÖLÜM 7

### SONUÇLAR

Bu tez çalışmasında, yüksek boyutlu hiperspektral verilerin karakter özelliklerini koruyarak daha düşük boyutlarda temsilini bulabilmek için 2 farklı yöntem önerilmekte ve bu konu kapsamında 3 farklı çalışma yer almaktadır. Çalışmalarda doğrusal olmayan boyut indirgeme yöntemleri içerisinde yer alan manifold öğrenme teknikleri üzerine yoğunlaşmaktadır. Bu yöntemlerin avantaj ve dezavantajları ele alınmakta ve çeşitli çözüm önerileri sunulmaktadır.

İlk çalışmada mevcut manifold yerleştirme yöntemleri karşılaştırmalı olarak analiz edilerek performansları değerlendirilmektedir. Botswana ve KSC verilerinin kullanıldığı bu çalışmada MAFE varyasyonu olan Enhanced-MAFE yönteminin sırasıyla elde ettiği %99,07 ve %82,51 sınıflandırma doğruluğu performansları ile mevcut yöntemlerden üstün olduğunun tespiti yapılmaktadır. Bu çalışma diğer iki çalışmaya da ışık tutması sebebiyle oldukça faydalı bir ön çalışma mahiyetindedir.

İkinci çalışmada, optimizasyon tabanlı bir yaklaşımla MAFE yöntemindeki parametre belirleme sorunu giderilmekte ve Botswana verileri için %99,37, KSC verileri için %98,17 ile mevcut diğer yöntemlere kıyasla en yüksek sınıflandırma doğruluğuna ulaşılmaktadır. Ayrıca bu çalışmada popüler derin öğrenme yöntemleri ve farklı 5 hiperspektral veri seti ile çeşitli analizlerde ele alınmaktadır. Yapılan deneylerde önerilen PSO-MAFE yönteminin derin öğrenme yaklaşımları ile benzer veya daha iyi sonuçlar aldığı görülmektedir. Bu da bize iyi bir yerleştirme işleminin önemini ortaya koymaktadır.

Üçüncü ve son çalışmada ise mevcut yöntemlerin güçlü yönlerini birleştiren hibrit bir yaklaşım önerilmektedir. Bu çalışmada da önerilen AWEM yöntemi sınıf ayrımının zor olduğu bir veri seti olan KSC verileri için %95,63 ile mevcut yöntemlere kıyasla

en iyi sınıflandırma doğruluğunu vermektedir. Ayrıca mevcut veri setindeki kalabalık sorununa başarılı bir çözüm getirilmektedir.

Önerilen her iki algoritma için de mevcut yöntemlere göre üstün sınıflandırma doğruluğu elde edilmektedir. Ayrıca indirgenmiş uzay için elde edilen yerleştirme grafiklerinde daha net ayırım yapılarak daha iyi görseller elde edilmektedir. Tüm yerleştirme yöntemlerinin etkilendiği kalabalık sorunu ise bu çalışmalarda büyük oranlarda ortadan kaldırılmıştır. Gelecek çalışmalarda, son yıllarda popüler olan ve başarılı sonuçlar alan derin öğrenme yaklaşımları için ön işlem aşamasında önerilen yöntemler kullanılabilir. Bu sayede daha hızlı ve verimli algoritmalar tasarlanabilir. Ayrıca önerilen yöntemlerin paralelleştirilmiş versiyonları ile zaman ve bellek maliyetleri düşürülebilir.



## KAYNAKLAR

1. Pearson, K., "On lines and planes of closest fit to systems of points in space", *Philosophical Magazine*, 2 (11): 559-572 (1901).
2. Hotelling, H., "Analysis of a complex of statistical variables into principal components," *Journal of Educational Psychology*, 24: 417-441 (1933).
3. Jolliffe, I. T., "Principal component analysis in regression analysis", *Springer*, 2: 129-155 (1986).
4. Martinez, A. M., and Kak, A. C., "Pca versus lda", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 23 (2): 228-223 (2001).
5. Torgerson, W., "Multidimensional scaling I: theory and method", *Psychometrika*, 17: 401-419 (1952).
6. Sugiyama, M., "Dimensionality reduction of multimodal labeled data by local fisher discriminant analysis", *Journal of Machine Learning Research*, 8: 1027-1061 (2007).
7. Sugiyama, M., "Local fisher discriminant analysis for supervised dimensionality reduction", *Proceedings of 23rd International Conference on Machine Learning*, 905-912 (2006).
8. Sugiyama, M., Ide, T., Nakajima S., and Sese, J., "Semi-supervised local fisher discriminant analysis for dimensionality reduction", *Machine Learning*, 78 (1-2): 35-61 (2010).
9. Roweis, S., and Saul, L. K., "Nonlinear dimensionality reduction by locally linear embedding", *Science*, 290: 2323-2326 (2000).
10. Tenenbaum, J. B., De Silva, V., and Langford, J., "A global geometric framework for nonlinear dimensionality reduction", *Science*, 290: 2319-2323 (2000).
11. Belkin, M., and Niyogi, P., "Laplacian eigenmaps for dimensionality reduction and data representation", *Neural Computation*, 15 (6): 1373-1396 (2003).
12. Zhag, Z., and Zha, H., "Principal manifolds and nonlinear dimensionality reduction via tangent space alignment", *SIAM Journal on Scientific Computing*, 26 (1): 313-338 (2004).
13. Song, L., Smola, A., Borgwardt, K., and Gretton, A., "Colored maximum variance unfolding", *Advances in Neural Information Processing Systems*, vol. 21, pp. 1385-1392, 2008.

14. Kohonen, T., and Honkela, T., “Kohonen network”, *Scholarpedia*, 2 (1): 1568 (2007).
15. Hinton, G., and Roweis, S., “Stochastic neighbor embedding”, *Advances in Neural Information Processing System*, 15: 857-864 (2002).
16. Maaten, L. V. D., and Hinton, G., “Visualizing data using t-sne”, *Journal of Machine Learning Research*, 9: 2579-2605 (2008).
17. Lunga, D., and Ersoy, O., “Spherical stochastic neighbor embedding of hyperspectral data”, *IEEE Transactions on Geoscience and Remote Sensing*, 51 (2): 857-871 (2013).
18. Lunga, D., and Ersoy, O., “Multidimensional artificial field embedding with spatial sensitivity”, *IEEE Transactions on Geoscience and Remote Sensing*, 52 (2): 1518-1532 (2014).
19. Bachmann, C. M., Ainsworth, T. L., and Fusina, R. A., “Exploiting manifold geometry in hyperspectral imagery”, *IEEE Transactions on Geoscience and Remote Sensing*, 43 (3): 441-454 (2005).
20. Huang, H., and Huang, Y., “Improved discriminant sparsity neighborhood preserving embedding for hyperspectral image classification”, *Neurocomputing*, 136 (1): 224-234 (2014).
21. Huang, H., Luo, F., Liu, J., and Yang, Y., “Dimensionality reduction of hyperspectral images based on sparse discriminant manifold embedding”, *ISPRS Journal of Photogrammetry Remote Sensing*, 106: 42-54 (2015).
22. Wang, Z., Du, B., Shi, Q., and Tu, W., “Domain adaptation with discriminative distribution and manifold embedding for hyperspectral image classification”, *IEEE Geoscience and Remote Sensing Letters*, 16 (7): 1155-1159 (2019).
23. Liu, B., Yu, A., Yu, X., Wang, R., Gao, K., and Guo., W., "Deep multiview learning for hyperspectral image classification", *IEEE Transactions on Geoscience and Remote Sensing*, 59 (9) 7758-7772 (2020).
24. Zhang, M., Li, W., and Du, Q., “Diverse region-based CNN for hyperspectral image classification”, *IEEE Transactions on Image Processing*, 27 (6): 2623-2634 (2018).
25. Zhao, W., and Du, S., “Spectral–spatial feature extraction for hyperspectral image classification: a dimension reduction and deep learning approach”, *IEEE Transactions on Geoscience and Remote Sensing*, 54 (8): 4544-4554 (2016).
26. Yıldırım, M. Z., Ozcan, C., and Ersoy, O., “Optimization based manifold embedding for hyperspectral image classification and visualization”, *Remote Sensing Letters*, 12 (11): 1158-1166 (2021).

27. Yıldırım, M. Z., Ozcan, C., ve Ersoy, O., “Hiperspektral görüntülerin sınıflandırılmasında farklı boyut indirgeme yöntemlerinin karşılaştırılması”, *Academic Platform Journal of Engineering and Science*, 9 (1): 159-165 (2021).
28. Hwang, Y., and Ersoy, O., “Comparison of multidimensional swarm embedding techniques by potential fields”, *Swarm Intelligence: Principles, Current Algorithms and Methods*, 1: 143-167 (2018).
29. Maaten, L. V. D., and Postma, E., “Matlab toolbox for dimensionality reduction”. *MICC Report*, Maastricht, 439-440 (2007).
30. Crawford, M. M., Kim, W., and Li, M., “Exploring nonlinear manifold learning for classification of hyperspectral data”, *Optical Remote Sensing*, Berlin, 207-234 (2011).
31. Roy, S. K., Haut, J. M., Paoletti, M. E., Dubey, S. R., and Plaza, A., “Generative adversarial minority oversampling for spectral-spatial hyperspectral image classification”, *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-15 (2021).
32. Di, W., and Crawford, M. M., “Active learning via multi-view and local proximity co-regularization for hyperspectral image classification”, *IEEE Journal of Selected Topics in Signal Processing*, 5 (3): 618-628 (2011).
33. Möller, M., Wittman, T., Bertozzi, A. L., and Burger, M., “A variational approach for sharpening high dimensional images”, *SIAM Journal on Imaging Sciences*, 5 (1): 150-178 (2012).
34. İnternet: Edin Photo, “William Henry Fox Talbot-Cameras 1840 to 1842”, [http://edinphoto.org.uk/1\\_P/1\\_photographers\\_talbot\\_smm\\_cameras\\_1840-1842.htm](http://edinphoto.org.uk/1_P/1_photographers_talbot_smm_cameras_1840-1842.htm) (2021).
35. İnternet: PAPA Intenational, “History of Aerial Photography-Dark Room in a Hot-Air Balloon”, <https://professionalaerialphotographers.com> (2021).
36. İnternet: PAPA Intenational, “History of Aerial Photography- Kites, Pigeons and Rockets”, <https://professionalaerialphotographers.com> (2021).
37. Lunga, W. D., “Multidimensional nonlinear embedding models for image scene dimensionality reduction and visualization”, Ph. D. Thesis, *Purdue University Graduate School*, West Lafayette, Indiana (2020).
38. İnternet: Museum of American Finance, “Launching the Space Race”, <https://www.moaf.org/exhibits/pandemics/asian-flu-1957/sputnik> (2021).

39. Dosapati, A., “Studying SST behaviour along the Italian coast by using Sentinel-3 imagery”, M. Sc. Thesis, *Sapienza University*, Rome (2012).
40. İnternet: NASA Landsat Science, “Landsat Image Gallery”, <https://landsat.visibleearth.nasa.gov/> (2021).
41. Liang, S., and Wang, J., “Advanced remote sensing 2<sup>nd</sup> ed. terrestrial information extraction and applications”, *Academic Press Inc.*, San Diego, 1-57 (2019).
42. İnternet: SEOS-Science Education through Earth Observation for High Schools, “Introduction to Remote Sensing-Resolution”, <https://seos-project.eu/remotesensing/remotesensing-c03-p05.html> (2021).
43. Plaza, J., Plaza, A. J., and Barra, C., “Multi-channel morphological profiles for classification of hyperspectral images using support vector machines”, *Sensors*, 9 (1): 196-218 (2009).
44. Erdin, K., “Elektromanyetik dalgaların oluşumu ve uzaktan algılama”, *Journal of the Faculty of Forestry Istanbul University*, 28 (2): 157-167 (1978).
45. İnternet: Imperial Encyclopedia Traveller, “Electromagnetic Spectrum”, [https://wiki.travellerrpg.com/Electromagnetic\\_Spectrum](https://wiki.travellerrpg.com/Electromagnetic_Spectrum) (2021).
46. İnternet: Selkirk Geospatial Research Centre, “Object Classification using Spectral Signatures”, <https://www.sgrc.selkirk.ca/blog/object-classification-using-spectral-signatures> (2021).
47. İnternet: UCL Earth Science, “The Remote Sensing Tutorial”, <https://www.ucl.ac.uk/EarthSci/people/lidunka/GEOL2014/Geophysics%2010%20-Remote%20sensing/RST%20Part%20II.htm> (2021).
48. İnternet: Semi-Automatic Classification Manual, “Land Cover Classification”, [https://semiautomaticclassificationmanual.readthedocs.io/de/latest/remote\\_sensing.html](https://semiautomaticclassificationmanual.readthedocs.io/de/latest/remote_sensing.html) (2021).
49. Bioucas-Dias, J. M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., and Chanussot, J., “Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches”, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5 (2): 354-379 (2012).
50. Ozdemir, A., and Polat, K., “Deep learning applications for hyperspectral imaging: a systematic review”, *Journal of The Institute of Electronics and Computer*, 2 (1): 39-56 (2020).
51. Mateen, M., Wen, J., Nasrullah, and Akbar, M. A., “The role of hyperspectral imaging: a literature review”, *International Journal of Advanced Computer Science and Applications*, 9 (8): 51-62 (2018).

52. Dhok, S., Bhurane, A., and Kothari, A., "Automated classification of hyperspectral images using spatial-spectral features", *2019 6th International Conference on Signal Processing and Integrated Networks*, 184-189 (2019).
53. Çukur, H., "Sezgisel yöntemlerle hiperspektral görüntülerde boyut indirgeme", Yüksek Lisans Tezi, *Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü*, İstanbul (2015).
54. Demir, B., Ertürk, S., "Improved hyperspectral image classification with noise reduction pre-process", *16th European Signal Processing Conference*, 1-4 (2008).
55. Yağın, F. H., "Gen ifade veri setlerinde boyut indirgeme yöntemlerinin sınıflama performansına etkilerinin karşılaştırılması", Yüksek Lisans Tezi, *İnönü Üniversitesi Sağlık Bilimleri Enstitüsü*, Malatya (2020).
56. Li, G., Li, J., Ju, Z., Sun, Y., and Kong, J., "A novel feature extraction method for machine learning based on surface electromyography from healthy brain", *Neural Computing and Applications*, 31 (12): 9013-9022 (2019).
57. Blum, A. L., and Rivest, R. L., "Training a 3-node neural network is NP-complete", *Neural Networks*, 5 (1): 117-127 (1992).
58. Li, J., Zhao, X., Li, Y., Du, Q., Xi, B., and Hu, J., "Classification of hyperspectral imagery using a new fully convolutional neural network", *IEEE Geoscience and Remote Sensing Letters*, 15 (2): 292-296 (2018).
59. Akgül, B., "Hiperspektral görüntülerde boyut indirgeme", Yüksek Lisans Tezi, *Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü*, İstanbul (2015).
60. Duan, P., Ghamisi, P., Kang, X., Rasti, B., Li, S., and Gloaguen, R., "Fusion of dual spatial information for hyperspectral image classification", *IEEE Transactions on Geoscience and Remote Sensing*, 59 (9): 7726-7738 (2020).
61. İnternet: Scikit Learn, "Comparison of Manifold Learning Methods", [http://lijiancheng0614.github.io/scikitlearn/auto\\_examples/manifold/plot\\_manifold\\_sphere.html](http://lijiancheng0614.github.io/scikitlearn/auto_examples/manifold/plot_manifold_sphere.html) (2021).
62. Taskin, G., "Extension to manifold learning methods via advanced regression methods", *Journal of the Faculty of Engineering and Architecture of Gazi University*, 37 (1): 485-495 (2022).
63. Yan, S., Xu, D., Zhang, B., Zhang, H. J., Yang, Q., and Lin, S., "Graph embedding and extensions: a general framework for dimensionality reduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29 (1): 40-51 (2006).

64. İnternet: Medium, “Dimensionality Reduction with PCA and t-SNE in R”, <https://medium.com/@TheDataGyan/dimensionality-reduction-with-pca-and-t-sne-in-r-2715683819> (2021).
65. Erbayram, T., “Boyut indirgeme tekniklerinin sınıflandırma performanslarının karşılaştırılması”, Yüksek Lisans Tezi, *Necmettin Erbakan Üniversitesi Fen Bilimleri Enstitüsü*, Konya (2020).
66. Akyürek, Ö., “Hiperspektral görüntülerde boyut indirgeme yöntemlerinin karşılaştırmalı analizi”, Yüksek Lisans Tezi, *Kocaeli Üniversitesi Fen Bilimleri Enstitüsü*, Kocaeli (2012).
67. Maaten, L. V. D., Postma, E., and Van den Herik, J., "Dimensionality reduction: a comparative review", *Journal of Machine Learning Research*, 10: 66-71 (2009).
68. Zaatour, R., Bouzidi, S., and Zagrouba, E., “Unsupervised image-adapted local fisher discriminant analysis to reduce hyperspectral images without ground truth”, *IEEE Transactions on Geoscience and Remote Sensing*, 58 (11): 7931-7941 (2020).
69. Fukunaga, K., “Introduction to statistical pattern recognition 2<sup>nd</sup> ed.”, *Academic Press Inc.*, San Diego (2013).
70. He, X., and Niyogi, P., “Locality preserving projections”, *Advances in Neural Information Processing Systems*, 16 (16): 153-160 (2004).
71. İsgüder, G. G., “Manifold learning for image-based gating of intravascular ultrasound(ivas) pullback sequences”, Yüksek Lisans Tezi, *Sabancı Üniversitesi Fen Bilimleri Enstitüsü*, İstanbul (2011).
72. Cover, T. M., and Thomas, J. A., “Information theory and statistics”, *Elements of Information Theory*, 1 (1): 279-335 (1991).
73. Hwang, Y., Ersoy, O., and Tan, Y., “Comparison of multidimensional swarm embedding techniques by potential fields”, *Swarm Intelligence: Principles, Current Algorithms and Methods*, 1: 143-167 (2018).
74. Kennedy, J., and Eberhard, R., “Particle swarm optimization”, *Proceedings of ICNN'95-International Conference on Neural Networks*, 4: 1942-1948, (1995).
75. Elbes, M., Alzubi, S., Kanan, T., Al-Fuqaha, A., and Hawashin, B., “A survey on particle swarm optimization with emphasis on engineering and network applications”, *Evolutionary Intelligence*, 12: 113-129 (2019).
76. Rostami, M., Forouzandeh, S., Berahmand, K., and Soltani, M., “Integration of multi-objective PSO based feature selection and node centrality for medical datasets”, *Genomics*, 112 (6): 4370-4384 (2020).

77. Atyabi, A., Luerssen, M., Fitzgibbon, S., and Powers, D. M., “Dimension reduction in EEG data using particle swarm optimization”, *IEEE Congress on Evolutionary Computation*, 1-8, (2012).
78. Atyabi, A., Luerssen, M., and Powers, D. M., “PSO-based dimension reduction of EEG recordings implications for subject transfer in BCI”, *Neurocomputing*, 119: 319-331 (2013).
79. Tharwat, A., and Hassanien, A. E., “Quantum-behaved particle swarm optimization for parameter optimization of support vector machine”, *Journal of Classification*, 36 (3): 576-598 (2019).
80. Wei, B., Zhang, W., Xia, X., Zhang, Y., Yu, F., and Zhu, Z., “Efficient feature selection algorithm based on particle swarm optimization with learning memory”, *IEEE Access*, 7: 166066-166078 (2019).
81. Gao, Y., and Li, Q., “A segmented particle swarm optimization convolutional neural network for land cover and land use classification of remote sensing images”, *Remote Sensing Letters*, 10 (12): 1182-1191 (2019).
82. Yu, L., Han, Y., and Mu, L., “Improved quantum evolutionary particle swarm optimization for band selection of hyperspectral image” *Remote Sensing Letters*, 11 (9): 866-875 (2020).
83. Li, L., Chen, Y., Xu, T., Huang, C., Liu, R., and Shi, K., “Integration of Bayesian regulation backpropagation neural network and particle swarm optimization for enhancing sub-pixel mapping of flood inundation in river basins”, *Remote Sensing Letters*, 7 (7): 631-640 (2016).
84. Esmín, A. A., Coelho, R. A., and Matwin, S., “A review on particle swarm optimization algorithm and its variants to clustering high-dimensional data”, *Artificial Intelligence Review*, 44 (1): 23-45 (2015).
85. Tsai, C. Y., and Chiu, C. C., “Developing a feature weight self-adjustment mechanism for a k-means clustering algorithm”, *Computational Statistics and Data Analysis*, 52: 4658–4672 (2008).
86. Ghayeb, H. H., Razak, H. A., Sulong, N. H. R., Hanoon, A. N., Abutaha, F., Ibrahim, H. A., Gordan, M., and Alnahhal, M. F., “Predicting the mechanical properties of concrete using intelligent techniques to reduce CO2 emissions”, *Materiales de Construcción*, 69 (334): 190 (2019).
87. Internet: NASA, “Nasa’s Earth Observation System-Earth Observing-1”, <https://eosps.nasa.gov/missions/earth-observing-1> (2021).
88. Wang, K., and Yong, B., “Application of the frequency spectrum to spectral similarity measures”, *Remote Sensing*, 8 (4): 344 (2016).

89. Internet: NASA, “Airborne Visible/Infrared Imaging Spectrometer”, <https://aviris.jpl.nasa.gov> (2021).
90. Davari, A. A., Aptoula, E., and Yanikoglu, B., “On the effect of synthetic morphological feature vectors on hyperspectral image classification performance”, *23rd Signal Processing and Communications Applications Conference*, 653-656 (2015).
91. Xie, F., Lei, C., Li, F., Huang, D., and Yang, J., “Unsupervised hyperspectral feature selection based on fuzzy c-means and grey wolf optimizer”, *International Journal of Remote Sensing*, 40 (9): 3344-3367 (2019).
92. Cohen, J., “A coefficient of agreement for nominal scales”, *Educational and Psychological Measurement*, 20 (1): 37-46 (1960).
93. Ozcan, C., and Ersoy, O., “Enhanced multidimensional field embedding method by potential fields for hyperspectral image classification and visualization”, *Electronics Letters*, 54 (12): 756-758 (2018).
94. Zhong, Z., Li, J., Ma, L., Jiang, H., and Zhao, H., "Deep residual networks for hyperspectral image classification", *2017 IEEE International Geoscience and Remote Sensing Symposium*, 1824-1827 (2017).
95. Hu, W., Huang, Y., Wei, L., Zhang, F., and Li, H., “Deep convolutional neural networks for hyperspectral image classification”, *Journal of Sensors*, 1-12 (2015).
96. Li, Y., Zhang, H., and Shen, Q., “Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network”, *Remote Sensing*, 9 (1): 67 (2017).
97. Zhang, M., Li, W., and Du, Q., “Diverse region-based CNN for hyperspectral image classification”, *IEEE Transactions on Image Processing*, 27 (6): 2623-2634 (2018).



## ÖZGEÇMİŞ

Mehmet Zahid YILDIRIM ilk ve orta öğrenimini Konya’da tamamladı. Fatih Endüstri Meslek Lisesi Bilgisayar Bölümü’nden 2004 yılında mezun oldu. 2006 yılında Selçuk Üniversitesi Teknik Eğitim Fakültesi Bilgisayar Sistemleri Öğretmenliği Bölümü’nde öğrenime başlayıp 2011 yılında mezun oldu. 2012 yılında Karabük Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü Bilgisayar Bilimleri Anabilim Dalı’nda Araştırma Görevlisi olarak göreve ve aynı anabilim dalında yüksek lisans eğitimine başladı. 2013 yılında Mühendislik Tamamlama programı ile Karabük Üniversitesi Bilgisayar Mühendisliği Bölümünden mezun oldu. 2017 yılında yüksek lisanstan mezun olarak Bilgisayar Mühendisliği Anabilim Dalı’nda doktora eğitimine başladı. Doktora eğitimini 2021 yılı Aralık ayında tamamladı. Halen aynı üniversitede görevine devam etmektedir. Evli ve bir kız çocuğu vardır.