

**ANKARA ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**YÜKSEK LİSANS TEZİ**

**HİPERSPEKTRAL VERİ SINIFLANDIRMA İÇİN BOYUT İNDİRGEME  
YÖNTEMLERİNİN BAŞARIMLARININ KARŞILAŞTIRILMASI**

**Ali Ömer KOZAL**

**ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI**

**ANKARA  
2014**

**Her hakkı saklıdır**

## TEZ ONAYI

Ali Ömer KOZAL tarafından hazırlanan “**Hiperspektral Veri Sınıflandırma İçin Boyut İndirgeme Yöntemlerinin Başarımlarının Karşılaştırılması**” adlı tez çalışması 16/04/2014 tarihinde aşağıdaki jüri tarafından oy birliği ile Ankara Üniversitesi Fen Bilimleri Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı’nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

**Danışman :** Yrd. Doç. Dr. Hakkı Alparslan ILGIN

Ankara Üniversitesi Elektrik-Elektronik Mühendisliği Anabilim Dalı

### Jüri Üyeleri:

**Başkan** : Doç. Dr. Süleyman TOSUN

Ankara Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı

**Üye** : Yrd. Doç. Dr. Selçuk TAŞCIOĞLU

Ankara Üniversitesi Elektrik-Elektronik Mühendisliği Anabilim Dalı

**Üye** : Yrd. Doç. Dr. Hakkı Alparslan ILGIN

Ankara Üniversitesi Elektrik-Elektronik Mühendisliği Anabilim Dalı

**Yukarıdaki sonucu onaylarım.**

**Prof. Dr. İbrahim DEMİR**

**Enstitü Müdürü**

## **ETİK**

Ankara Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım bu tez içindeki bütün bilgilerin doğru ve tam olduğunu, bilgilerin üretilmesi aşamasında bilimsel etiğe uygun davrandığımı, yararlandığım bütün kaynakları atıf yaparak belirttiğimi beyan ederim.

16.04.2014

Ali Ömer KOZAL

## ÖZET

Yüksek Lisans Tezi

### HİPERSPEKTRAL VERİ SINIFLANDIRMA İÇİN BOYUT İNDİRGEME YÖNTEMLERİNİN BAŞARIMLARININ KARŞILAŞTIRILMASI

Ali Ömer KOZAL

Ankara Üniversitesi  
Fen Bilimleri Enstitüsü  
Elektrik - Elektronik Mühendisliği Anabilim Dalı

Danışman: Yrd. Doç. Dr. Hakkı Alparslan ILGIN

Hiperspektral algılayıcılar, devamlı ve düşük dalga boyu aralıklarında yüzlerce spektral banda sahip yüksek boyutlu görüntüler üretmektedir. Bu bantlar genel olarak elektromanyetik spektrumun görünür bölgesi ve kızılötesi aralığını içermektedir. Alınan bu yüksek boyutlu veri sayesinde bir bölgenin hiperspektral görüntü verisi kullanılarak, diğer görüntü verilerine göre daha yüksek sınıflandırma başarımı sağlanabilmektedir. Ancak çok sayıda banda sahip bu verilerin sınıflandırılması için yüksek işlem gücü ve işlem süresi gerekmektedir. Bu yüzden, hiperspektral görüntülerin sınıflandırılması aşamasında, artıklıkları azaltmak ve daha az boyutta veri kullanılarak sınıflandırma performansını artırabilmek için boyut indirgeme yöntemleri kullanılması gereklidir.

Bu tez çalışmasında, Hiperspektral görüntülerin sınıflandırılması sürecinde yapılacak boyut indirgeme aşaması için farklı boyut indirgeme yöntemlerinin sınıflandırma açısından başarımları ve hesaplama performansları karşılaştırılmıştır. Sınıflandırıcı olarak Destek Vektör Makineleri ve En Yakın Komşuluk Sınıflandırıcı kullanılmıştır.

Çalışmanın ilk aşamasında hiperspektral verilere alt örnekleme yaparak sınıflandırma başarımı ve işlem süresi açısından sınıflandırma performansları incelenmiştir. Boyut indirgeme tekniklerinin kullandığı veri azaltılarak işlem süresi ve sınıflandırma başarımlarındaki performans değişiklikleri ortaya konmuştur. İkinci aşama olarak, iki farklı çoğunluk oylaması yöntemi kullanarak sınıflandırma başarımının artışı denenmiştir. Son olarak boyut laneti olarak da bilinen; yüksek boyutlu verilerin sınıflandırılmasında, eğitim verisinin yetersiz kalması ile sınıflandırma performansındaki düşüş etkisi incelenmiş; farklı boyut indirgeme tekniklerinin bu olgunun etkisini azaltabildiği ortaya konmuştur.

**Nisan 2014, 104 sayfa**

**Anahtar Kelimeler:** Hiperspektral görüntüleme, boyut indirgeme, uzaktan algılama, hiperspektral görüntülerde sınıflandırma, uydu görüntüleme, görüntü sınıflandırma

## **ABSTRACT**

Master Thesis

### **COMPARING PERFORMANCE OF DIMENSION REDUCTION TECHNIQUES FOR HYPERSPECTRAL DATA CLASSIFICATION**

Ali Ömer KOZAL

Ankara University  
Graduate School of Natural and Applied Sciences  
Department of Electrical and Electronics Engineering

Supervisor: Asst. Prof. Dr. Hakkı Alparslan ILGIN

Hyperspectral sensors generate images in narrow bands in continuous manner with hundreds of spectral bands. In general, these bands contain the data in visible and infrared part of the electromagnetic spectrum. With the use of this high dimensional image data, a scene is classified with higher accuracy than other types of image data of the same scene. However, to classify the data with large number of spectral bands requires more processing power and processing time. Therefore, for classifying hyperspectral images, it is required to apply dimension reduction techniques in order to eliminates redundant information and increase classifying efficiency with less number of bands.

In this thesis, variety of dimension reduction methods is compared in classification accuracy and calculation time during classification process. Support Vector Machines and Nearest Neighbor Classifier are used for classification.

In the first phase of the study, the effect of sub-sampling hyperspectral images for dimension reduction techniques is explored and compared in classification performance and calculation time. Reducing the amount of data used by dimension reduction methods, change in processing time and classification accuracy is presented. For the second phase, increase in classification accuracy is experienced by utilizing two different majority voting methods. Finally, phenomena known as curse of dimensionality, that is the decrease in the accuracy when classifying high dimensional data with few training sample is analyzed and it is shown that some dimension reduction methods can cope with this effect.

**April 2014, 104 pages**

**Key Words:** Hyperspectral imaging, dimension reduction, remote sensing, hyperspectral image classification, satellite imaging, image classification

## TEŞEKKÜR

Uzun süren yüksek lisans eğitim sürecinde bana çok yardımcı olan ve rehberlik eden danışmanım Yrd. Doç. Dr. Hakkı Alparslan ILGIN'a (Ankara Üniversitesi Elektrik-Elektronik Mühendisliği Anabilim Dalı) öncelikle teşekkür etmek isterim.

Ardından beni hiperspektral görüntü işleme konusunun içine iten ve her konuda yol gösteren çalışma arkadaşım Mustafa TEKE'ye,

Özelikle tez yazdığım sürede bana hep destek veren eşim Bahar, kuzularım Zeynep ve Zehra'ya,

Her konuştuğumuzda ne zaman bitireceğimi sorgulayan, benden çok dert eden anneme ve babama

sonsuz teşekkürler...

Ali Ömer KOZAL

Ankara, Nisan 2014

## İÇİNDEKİLER

TEZ ONAY SAYFASI	
ETİK.....	i
ÖZET.....	ii
ABSTRACT.....	iii
TEŞEKKÜR .....	iv
KISALTMALAR DİZİNİ .....	vii
ŞEKİLLER DİZİNİ .....	viii
ÇİZELGELER DİZİNİ .....	x
1. GİRİŞ.....	1
2. UZAKTAN ALGILAMA .....	4
2.1 Giriş .....	4
2.2 Görüntüleme .....	6
2.2.1 Pankromatik görüntüleme .....	6
2.2.2 Multispektral görüntüleme.....	7
2.2.3 Hiperspektral görüntüleme .....	9
2.3 Çözünürlük .....	15
2.3.1 Uzaysal çözünürlük .....	16
2.3.2 Spektral çözünürlük.....	17
2.3.3 Radyometrik çözünürlük.....	18
2.4 Hiperspektral Özet .....	19
3. BOYUT İNDİRGEME.....	20
3.1 Giriş .....	20
3.1.1 Öznitelik çıkarımı.....	20
3.1.2 Öznitelik seçimi .....	21
3.2 Çalışmada Kullanılan Boyut İndirgeme Yöntemleri .....	21
3.2.1 Doğrusal yöntemler .....	22
3.2.2 Doğrusal olmayan yöntemler .....	25
3.3 Boyut seçimi .....	28
4. SINIFLANDIRMA .....	30
4.1 Kullanılan Veri .....	30
4.1.1 AVIRIS - Indian Pines .....	31
4.1.2 ROSIS – Pavia University.....	33
4.2 Destek Vektör Makineleri .....	36
4.3 K - En Yakın Komşuluk Sınıflandırıcı.....	37
5. DENEYSEL ÇALIŞMALAR .....	39
5.1 Giriş .....	39
5.1.1 Destek Vektör Makineleri parametre seçimi.....	39
5.1.2 Sınıflandırıcı eğitimi.....	40
5.1.3 Boyut seçimi .....	40
5.1.4 Örneklem dışı yayma .....	41
5.1.5 Tekrarlı çoğunluk oylaması.....	41
5.1.6 Uzaysal çoğunluk oylaması.....	43
5.2 Test 1: Alt Örneklemenin Hiperspektral Boyut İndirgemeye Etkisi .....	44
5.2.1 Test akışı.....	45
5.2.2 Test adımları .....	45

5.2.3 Test sonuçları.....	49
5.3 Test 2: Çoğunluk Oylamasıyla Sınıflandırma Başarımının Artırımı.....	57
5.3.1 Test akışı.....	58
5.3.2 Test adımları.....	59
5.3.3 Test sonuçları.....	63
5.4 Test 3: Hughes Olgusunun Yöntemler Üzerindeki Etkisi .....	74
5.4.1 Test akışı.....	75
5.4.2 Test adımları.....	77
5.4.3 Test sonuçları.....	79
6. DEĞERLENDİRME.....	81
7. SONUÇ.....	98
KAYNAKLAR .....	100
ÖZGEÇMİŞ.....	104



## KISALTMALAR DİZİNİ

AE	Oto-Kodlayıcı (Auto Encoder)
Ar-Ge	Araştırma – Geliştirme
AVIRIS	Airborne Visible InfraRed Imaging Spectrometer
DIF	Difüzyon Haritaları (Diffusion Maps)
FA	Faktör Analizi (Factor Analysis)
FIR	Uzak kızılötesi (Far Infrared)
FMVU	Hızlı Maksimum Değişinti Çözücü (Fast Maximum Variance Unfolding)
ISO	Eş Ölçekli Öznitelik Haritalama (Isometric Feature Mapping)
JPL	Jet Propulsion Laboratory, NASA
KNN	K-En Yakın Komşu (K-Nearest Neighbor)
LDA	Doğrusal Ayraç Analizi (Linear Discriminant Analysis)
LPP	Bölgeselliği Koruyan İzdüşüm (Locality Preserving Projections)
LWIR	Uzun dalga boylu kızılötesi (Long-wavelength infrared)
MANC	Çok Katmanlı Haritalama (Manifold Charting)
MCML	En Fazla Daralan Metrik Öğrenme (Maximally Collapsing Metric Learning)
MWIR	Orta dalga boylu kızılötesi (Mid-wavelength infrared)
NASA	National Aeronautics and Space Administration
NCA	Komşuluk Bileşen Analizi (Neighbourhood Components Analysis)
NIR	Yakın kızılötesi (Near Infrared)
NPE	Komşuluk Korumalı İç Katma (Neighborhood Preserving Embedding)
PCA	Temel Bileşen Analizi (Principal Component Analysis)
PPCA	Olasılıksal PCA (Probabilistic Principal Component Analysis)
RGB	Kırmızı, Yeşil, Mavi (Red, Green, Blue)
ROSIS	Reflective Optics System Imaging Spectrometer
SPE	Olasılıksal Yakınlık İç Katma (Stochastic Proximity Embedding)
SVM	Destek Vektör Makineleri (Support Vector Machines)
SWIR	Kısa dalga boylu kızılötesi (Short-wavelength infrared)

## ŞEKİLLER DİZİNİ

Şekil 2.1 Elektromanyetik spektrum .....	5
Şekil 2.2 Görünür ve Kızılötesi Bantlar .....	6
Şekil 2.3 Pankromatik, Multispektral ve Hiperspektral Bantlar .....	6
Şekil 2.4 Pankromatik görüntü.....	7
Şekil 2.5 Aynı bölgenin farklı çoklu bant görüntüsü .....	8
Şekil 2.6 Landsat 7 Bant Karşılaştırması .....	9
Şekil 2.7 Multispektral ve Hiperspektral görüntü özellikleri.....	10
Şekil 2.8 Ormanlık bir arazinin spektral imzası.....	12
Şekil 2.9 Sağlıklı ve hastalıklı şeker pancarının spektral imzası .....	13
Şekil 2.10 Hiperspektral Veri Küpü .....	14
Şekil 2.11 Hiperspektral görüntülerin elde edilmesi (Shaw ve Burke 2003).....	15
Şekil 2.12 Spektral ve uzaysal çözünürlükler (Schowengerdt 2006).....	17
Şekil 2.13 Aynı görüntünün farklı radyometrik çözünürlükte ifadesi .....	18
Şekil 4.1 Indian Pines görüntüsü gösterimi .....	31
Şekil 4.2 Indian Pines görüntüsünün Sınıflandırma Haritası .....	32
Şekil 4.3 Pavia University çok bantlı görüntüsü.....	34
Şekil 4.4 Pavia University görüntüsünün Sınıflandırma Haritası .....	34
Şekil 4.5 Pavia University altküme verisi .....	35
Şekil 5.1 Tekrarlı çoğunluk oylaması .....	42
Şekil 5.2 Uzaysal çoğunluk oylaması .....	43
Şekil 5.3 Test akışı, Test 1 .....	46
Şekil 5.4 Veri filtresi, Test 1 .....	47
Şekil 5.5 Alt örnekleme.....	47
Şekil 5.6 Tam Veri Seti Boyut İndirgeme, Test 1 .....	47
Şekil 5.7 Alt Örneklenmiş Veri Seti Boyut İndirgeme, Test 1 .....	48
Şekil 5.8 Örnekleme dışı yayma .....	48
Şekil 5.9 SVM Eğitim, Test 1 .....	49
Şekil 5.10 Sınıflandırma ve karşılaştırma, Test 1 .....	49
Şekil 5.11 Tam veri seti, ilk 3 bandının RGB gösterimi.....	50
Şekil 5.12 Alt örneklenmiş veri seti, ilk 3 bandının RGB gösterimi .....	51
Şekil 5.13 AE sınıflandırma sonuçları .....	52
Şekil 5.14 FA sınıflandırma sonuçları .....	52
Şekil 5.15 FMVU sınıflandırma sonuçları.....	52
Şekil 5.16 ISO sınıflandırma sonuçları .....	53
Şekil 5.17 LDA sınıflandırma sonuçları .....	53
Şekil 5.18 LPP sınıflandırma sonuçları.....	53
Şekil 5.19 MANC sınıflandırma sonuçları .....	54
Şekil 5.20 MCML sınıflandırma sonuçları .....	54
Şekil 5.21 NCA sınıflandırma sonuçları .....	54
Şekil 5.22 NPE sınıflandırma sonuçları .....	55
Şekil 5.23 PCA sınıflandırma sonuçları.....	55
Şekil 5.24 PPCA sınıflandırma sonuçları .....	55
Şekil 5.25 Test akışı, Test 2 .....	59
Şekil 5.26 Sınıflandırma Haritasına veri filtresi uygulaması .....	60
Şekil 5.27 Boyut İndirgeme, Test 2 .....	60

Şekil 5.28 SVM Eğitim, Test 2 .....	61
Şekil 5.29 Sınıflandırma, Test 2.....	61
Şekil 5.30 KNN Eğitim, Test 2 .....	61
Şekil 5.31 Çoğunluk oylaması uygulamaları .....	62
Şekil 5.32 AE sınıflandırma sonuçları .....	64
Şekil 5.33 DIF sınıflandırma sonuçları .....	65
Şekil 5.34 FA sınıflandırma sonuçları .....	65
Şekil 5.35 FMVU sınıflandırma sonuçları .....	66
Şekil 5.36 ISO sınıflandırma sonuçları .....	66
Şekil 5.37 LDA sınıflandırma sonuçları .....	67
Şekil 5.38 LPP sınıflandırma sonuçları.....	67
Şekil 5.39 MANC sınıflandırma sonuçları .....	68
Şekil 5.40 NPE sınıflandırma sonuçları .....	68
Şekil 5.41 PCA sınıflandırma sonuçları.....	69
Şekil 5.42 PPCA sınıflandırma sonuçları .....	69
Şekil 5.43 SPE sınıflandırma sonuçları.....	70
Şekil 5.44 Faktör Analizi için sınıflandırma sonuçları .....	71
Şekil 5.45 Test akışı, Test 3 .....	76
Şekil 5.46 Sınıflandırma Haritasına veri filtresi uygulaması Test 3 .....	77
Şekil 5.47 Boyut İndirgeme, Test 3 .....	77
Şekil 5.48 Rastgele veri kümesi oluşturma .....	78
Şekil 5.49 Sınıflandırıcı Eğitimi, boyutu indirgenmiş veri .....	78
Şekil 5.50 Sınıflandırıcı Eğitimi, boyutu indirgenmemiş veri .....	79
Şekil 6.1 Yöntemlerin sınıflandırma doğruluğunun karşılaştırılması, Test 1 .....	82
Şekil 6.2 Yöntemlerin sürelerinin karşılaştırılması, Test 1 .....	83
Şekil 6.3 SVM ile yapılan sınıflandırma sonuçları .....	87
Şekil 6.4 KNN ile yapılan sınıflandırma sonuçları .....	87
Şekil 6.5 Indian Pines, ilk aşama çoğunluk oylaması yöntemleri.....	90
Şekil 6.6 Pavia University, ilk aşama çoğunluk oylaması yöntemleri.....	91
Şekil 6.7 Indian Pines, ikinci aşama çoğunluk oylaması yöntemleri.....	93
Şekil 6.8 Pavia University, ikinci aşama çoğunluk oylaması yöntemleri.....	93
Şekil 6.9 Test 2 SVM, boyut indirgeme yöntem sürelerinin karşılaştırılması .....	95
Şekil 6.10 Indian Pines, Test 3 sonuç karşılaştırması .....	96
Şekil 6.11 Pavia University, Test 3 sonuç karşılaştırması .....	96

## ÇİZELGELER DİZİNİ

Çizelge 2.1 Hiperspektral görüntüleyiciler .....	11
Çizelge 4.1 Indian Pines görüntüsünde kullanılan sınıflar ve örnek sayıları .....	33
Çizelge 4.2 Pavia University görüntüsünde kullanılan sınıflar ve örnek sayıları .....	35
Çizelge 5.1 Sınıflandırma sonucu, Test 1 .....	56
Çizelge 5.2 Faktör Analizi için sınıflandırma doğrulukları .....	70
Çizelge 5.3 Indian Pines verisi için SVM çoğunluk oylaması test sonuçları .....	72
Çizelge 5.4 Indian Pines verisi için KNN çoğunluk oylaması test sonuçları .....	72
Çizelge 5.5 Pavia University verisi için SVM çoğunluk oylaması test sonuçları.....	73
Çizelge 5.6 Pavia University verisi için KNN çoğunluk oylaması test sonuçları.....	73
Çizelge 5.7 Boyut indirgeme yöntem sürelerinin karşılaştırılması.....	74
Çizelge 5.8 Indian Pines verisi için KNN çoğunluk oylaması test sonuçları .....	79
Çizelge 6.1 Alt örneklenmiş veri seti ile tam veri seti sonuç karşılaştırması .....	84
Çizelge 6.2 Farklı veri ve sınıflandırıcı sonuçlarının karşılaştırılması .....	85
Çizelge 6.3 Indian Pines verisi SVM çoğunluk oylaması karşılaştırması .....	88
Çizelge 6.4 Pavia University verisi SVM çoğunluk oylaması karşılaştırması .....	89
Çizelge 6.5 İlk aşama çoğunluk oylamanın sınıflandırma sonucuna etkisi .....	90
Çizelge 6.6 İkinci aşama çoğunluk oylamanın sınıflandırma sonucuna etkisi .....	92
Çizelge 6.7 Boyut indirgeme yöntem sürelerinin karşılaştırılması Test 2 .....	94

## 1. GİRİŞ

Hiperspektral görüntüler; genellikle elektromanyetik spektrumun görünür, yakın kızılötesi ve kısa dalga kızılötesi bölümünde devamlı ve dar aralıklı dalga boyu aralıklarında elde edilen verilerdir. Hiperspektral bir görüntü 50-250 arasında banttan oluşabilmektedir. Görüntü bantları birbirine yakın olduğu için artıklık içerebilmekte ve sınıflandırma için kullanılacak başka bir öznelik uzayında, daha düşük boyutlu veride, istenilen sınıflandırma performansına ulaşmak mümkün olmaktadır (Kozal vd. 2013a).

Hiperspektral görüntüler, görünür bölge görüntülerinden daha fazla bant içerdiğinden insan gözünün algılayamadığı bilgileri ortaya koyar. Bu nedenle tarım, madencilik, astronomi, kimya gibi birçok alanda kullanılır. Spektrumları benzer görüntülerin sınıflandırılması ile nesne veya maddelerin ayırt edilmesi hiperspektral görüntülerin temel amacı iken en büyük dezavantajı içerdiği yüksek sayıda bant nedeniyle sınıflandırmadaki işlem yükünün fazla olmasıdır. Bunun yanında yüksek boyutundan dolayı sınıflandırma yöntemlerinin performansı düşmektedir (Landgrebe 2003). Bu dezavantajların önüne geçmek için hiperspektral görüntülerin işlenmesi öncesi boyut indirgeme uygulanır.

Yüksek boyutlu veriyi, daha düşük boyutlu bir öznelik bir uzayında temsil etmek için yapılan işlemler, boyut indirgeme olarak tanımlanmaktadır. Boyut indirgeme ile yüksek boyutlu veride bulunan bilgi içeriği daha az sayıda öznelik ile daha az boyutlu temsil edilmekle birlikte sınıflandırma işlemi için de daha yüksek performans alınacak bir uzaya veri dönüşümü yapılması amaçlanmaktadır.

Hiperspektral görüntülerde boyut indirgeme yöntemi olarak genelde Temel Bileşen Analizi (Principal Component Analysis, PCA) kullanılmaktadır (Rodarmel ve Shan 2002). Bunun yanında birçok farklı yöntem bulunmaktadır.

Bu çalışmanın amacı, son yıllarda ihtiyacı artan ve başta ülkemizde olmak üzere projelendirilmeye başlamış olan hiperspektral görüntülerin işlenmesini de içerecek operasyonel bir hiperspektral görüntü işleme yazılımı için, farklı boyut indirgeme yöntemlerinin sınıflandırma başarımına etkisi açısından yol göstermektir.

Projelere örnek olarak, Haziran 2013 tarihinde başlayan, müşterisi Milli Savunma Bakanlığı olan Kalkınma Bakanlığı tarafından desteklenen Yüksek Çözünürlüklü Yer Gözlem Uydularına yönelik alt sistemlerin ülkemizde geliştirilmesi ve bunun için gereken Ar-Ge, entegrasyon ve test alt yapılarının oluşturulmasını hedefleyen İMECE projesi (<http://www.uzay.tubitak.gov.tr> 2013) verilebilir. Proje kapsamında Hiperspektral Uydu Kamerası Alt Yapısı geliştirilecektir.

Bir diğer örnek, müşterisinin Savunma Sanayii Müsteşarlığı olduğu, ülkemizdeki görüntü istihbaratına dayalı gelişmelere paralel olarak yerli imkanlarla multispektral ve hiperspektral görüntüleme teknolojilerinin geliştirilmesi ve bu teknolojilerle ilgili altyapıların ülkemize kazandırılmasını hedefleyen TUYGUN (İleri Görüntüleme Teknolojileri) projesidir (<http://www.ssm.gov.tr> 2014).

Bu tez çalışmasında hiperspektral görüntüler; doğrusal olan ve olmayan, öğreticili ve öğreticisiz gibi birçok farklı yöntem ile boyutu indirgenmiştir. Boyutu indirgenen bu görüntüler sınıflandırmaya tabi tutularak, yöntemler süre ve sınıflandırma doğruluğu açısından karşılaştırılmıştır.

Çalışma üç test aşamasından oluşmaktadır. Testler öncesinde hiperspektral görüntülere boyut indirgeme öncesi herhangi bir ön işlem uygulanmamıştır. Testlerde farklı olmakla birlikte 12şer adet boyut indirgeme yöntemi kullanılmıştır.

İlk testte (Bölüm 5.2; Test 1: Alt Örneklemenin Hiperspektral Boyut İndirgemeye Etkisi), sınıflandırmanın daha kısa sürede daha az işlem yüküyle yapılması açısından tüm görüntünün değil, görüntünün bir kısmının boyutu indirgenmiş ve bu işlem tüm görüntüye uygulanmıştır. Bu işlem, görüntüdeki sınıflardan örnek alarak bir alt örneklenmiş küme oluşturarak yapılmıştır. Tüm veri yerine oluşturulan alt örneklenmiş kümenin boyutunu indirgenmiş ve elde edilen dönüşüm fonksiyonu tüm kümeye uygulanmıştır. Bu yöntemle elde edilen görüntü Destek Vektör Makineleri ile sınıflandırılmış ve sınıflandırma başarımı geleneksel yöntem ile karşılaştırılmıştır.

İkinci testte (Bölüm 5.3; Test 2: Çoğunluk Oylamasıyla Sınıflandırma Başarımının Artırımı) yine boyut indirgeme ve sonrasında sınıflandırma yapılmış, sınıflandırıcı

olarak Destek Vektör Makinelerinin yanında K – En Yakın Komşuluk Sınıflandırıcı kullanılmıştır. Elde edilen sınıflandırma sonuçları iki farklı çoğunluk oylamasının kombinasyonları işlenmiş ve bu işlemlerin başarıma etkisi incelenmiştir.

Üçüncü testte (Bölüm 5.4; Test 3: Hughes Olgusunun Yöntemler Üzerindeki Etkisi) K – En Yakın Komşuluk Sınıflandırıcı kullanarak boyutu indirgenmiş ve indirgenmemiş veriler karşılaştırılmıştır.

Çalışma MATLAB ortamında yapılmış, çalışmalar sırasında standart MATLAB kütüphanelerinin yanında libSVM kütüphanesi (Chang ve Lin 2011) ve boyut indirgeme (dimension reduction) kütüphanesi (Van der Maaten 2007) kullanılmıştır.

## 2. UZAKTAN ALGILAMA

### 2.1 Giriş

Uzaktan algılama, yeryüzünde bulunan nesnenin özelliklerinin uzaktan çeşitli alıcılar yardımıyla bilgi elde etme bilimi olarak tanımlanır (Schowengerdt 2006). Dolayısı ile bu özellikler, yerinde değil belirli bir mesafeden ölçülür. Nesne ile doğrudan bir temas olmadığından; optik, akustik ya da mikrodalga gibi yayılan sinyaller ölçülür.

Uzaktan algılamada alıcılar genel olarak atmosferde ya da uzayda bulunan platformlara yerleştirilirler. Elde edilen bilgiler meteoroloji, tarım, şehir bölge planlama, ormancılık gibi alanlarda kullanılmaktadır.

Uzaktan algılamada yaygın olarak alıcılar uçaklara ya da uydulara yerleştirilmekte ve kilometrelerce uzaktan alınmaktadır. Örneğin hâlihazırda ülkemizin sahip olduğu RASAT ve GÖKTÜRK-2 adlı yer gözlem uyduları sırasıyla 689 ve 686 km yörüngeden görüntü almaktadır. Alınan bu görüntüler indirildikten sonra yerde değerlendirilmektedir (<http://www.uzay.tubitak.gov.tr> 2013).

Uzaktan algılama sistemleri, yeryüzünden yansıyan farklı dalga boylarındaki enerjinin sensörler tarafından algılanıp veriye dönüştürülmesi prensibi ile çalışır. Bunlardan çoğunluğu pasif sistemlerdir, yani güneş ışınlarının ya da başka kaynaklardan gelen ışınların yansımaları ile oluşan enerjiyi algılayıp veriye çevirirler. Aktif sistemler ise ürettikleri enerjiyi yeryüzüne gönderir ve yansıyan ışınları kaydeder. Uçaklarda kullanılan radar ve yapay açıklıklı radar görüntüleyiciler bunlara örnek verilebilir.

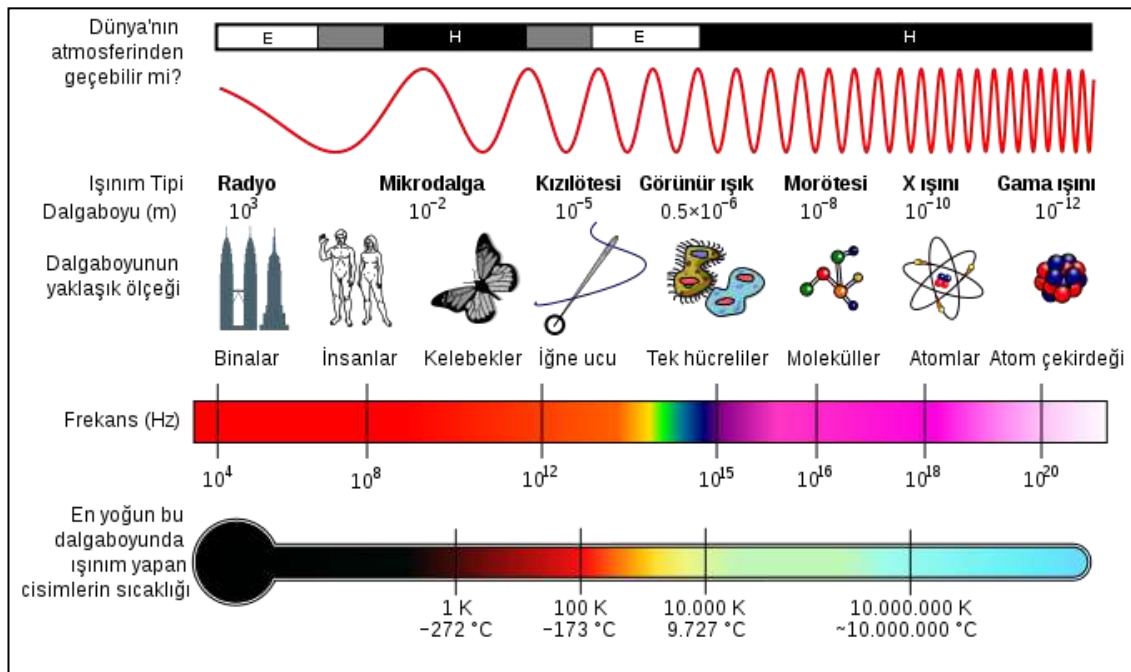
Elektromanyetik dalgaların tümünün oluşturduğu diziye elektromanyetik spektrum denir. Elektromanyetik spektrum şu aralıklara ayrılmıştır:

Enerjisi en yüksek kısımda gamma ışınları ve X ışınları vardır. 1 nm ile 0,6 µm arasını morötesi kısım oluşturur. Görünür ışığın dalga boyu 0,4 ila 0,7 µm'dir. Kızılötesi kısım 0,7 ile 1000 µm arasında yer alır. Daha kısa dalga boylarında yani 0,7 µm yakınlarında kızılötesi ışınları özel filmlerle algılanabilirken, daha uzun dalga boyları sıcaklık olarak



hissedilebilir. Daha uzun dalga boylu aralıklar, milimetre ve metre ile ölçülür. Mikrodalga bölgesi 1 mm ile 1 m arasında değişir ve radar sistemlerini de içerir. Son olarak en uzun dalga boylu ve frekansı en düşük kısımda radyo dalgaları yer almaktadır (<http://earthobservatory.nasa.gov> 2013).

Bu aralıkların isimlendirilmesi, atmosferin geçirgenliği, frekansları ve bu dalga boyunda en yoğun ışımayı yapan cisimlerin sıcaklığı şekil 2.1 ile verilmiştir (<http://myasadata.larc.nasa.gov> 2013).



Şekil 2.1 Elektromanyetik spektrum

Uzaktan algılama uygulamalarında genel olarak elektromanyetik spektrumun 0,4 – 15  $\mu\text{m}$  dalga boylu enerjiye sahip bölgesi kullanılır. Bu, görünür bant (0,4 – 0,75  $\mu\text{m}$ ) ve kızılötesi kısımdan oluşur.

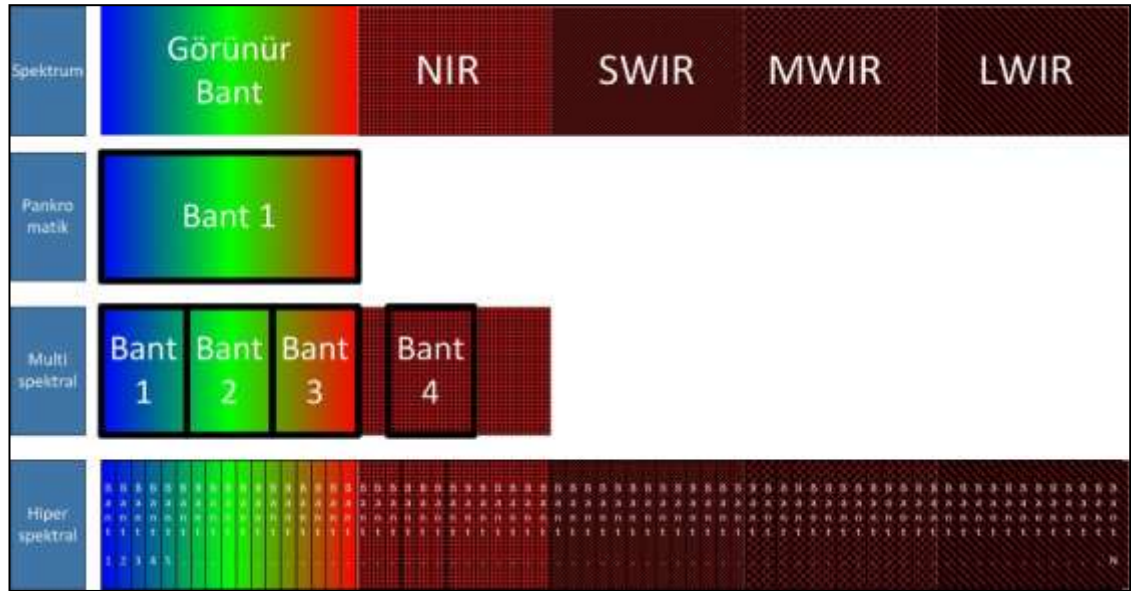
Kızılötesi; yakın kızılötesi (NIR, 0,75 – 1,4  $\mu\text{m}$ ), kısa dalga boylu kızılötesi (SWIR, 1,4 – 3  $\mu\text{m}$ ), orta dalga boylu kızılötesi (MWIR, 3 – 8  $\mu\text{m}$ ), uzun dalga boylu kızılötesi (LWIR, 8 – 15  $\mu\text{m}$ ) ve uzaktan algılamada çok kullanılmayan uzak kızılötesi (FIR, 15 – 1000  $\mu\text{m}$ ) olarak beşe ayrılır (Şekil 2.2). NIR ve SWIR yansıyan kızılötesi, MWIR ve LWIR de termal kızılötesi olarak da adlandırılabilir (Byrnes 2009).



Şekil 2.2 Görünür ve Kızılötesi Bantlar

## 2.2 Görüntüleme

Yaygın olarak kullanılan görüntüleyiciler insan gözünün algıladığı görünür dalga boyunda (0,4 – 0,75 µm) görüntü alan optik pasif görüntüleyicilerdir. Bu dalga boyları için Pankromatik (siyah-beyaz) ve Multispektral (çok bantlı) sensörler kullanılmaktadır.



Şekil 2.3 Pankromatik, Multispektral ve Hiperspektral Bantlar

### 2.2.1 Pankromatik görüntüleme

Pankromatik algılayıcılar, tek bantta genellikle elektromanyetik spektrumun görünür bölgesinin tamamını algılar, şekil 2.3. Örneğin RASAT uydusunda yer alan pankromatik algılayıcının spektral aralığı 0,42 – 0,73 µm'dir ve insan gözünün algılayabildiği spektrumun tamamı ile örtüşmektedir (<http://www.uzay.tubitak.gov.tr> 2014).



Şekil 2.4 Pankromatik görüntü

Pankromatik görüntülerde tüm görünür bölge tek bir bantla ifade edildiğinden ortaya şekil 2.4 ile gösterildiği gibi bir gri-seviye görüntü çıkar.

### 2.2.2 Multispektral görüntüleme

Multispektral görüntüleyiciler, çoğunlukla görünür dalga boyunun birden fazla algılayıcı ile görüntülenmesi prensibi ile çalışır. Genel olarak kırmızı, yeşil ve mavi renklerden oluşan üç spektral banttandır. Bu bantların birliktedir, spektral bant sayısı daha da artabilir (Şekil 2.3). İnsan gözünün de elektromanyetik spektrumda yer alan 3 rengi ve bunların karışımını algıladığı düşünüldüğünde çok geniş kullanım alanları bulunmaktadır. Günümüzde fotoğraf makinalarının görüntü alma sistemleri, bilgisayar monitörlerinin görüntü gösterme sistemleri gibi birçok alanda 3 bantlı multispektral görüntüleme kullanılmaktadır.

Uzaktan algılamada, multispektral görüntülerin farklı bantları birleştirilerek renkli görüntü (Şekil 2.5.a) ya da yalancı renk (false color) görüntü (Şekil 2.5.b) gibi uygulamalarda kullanılır.



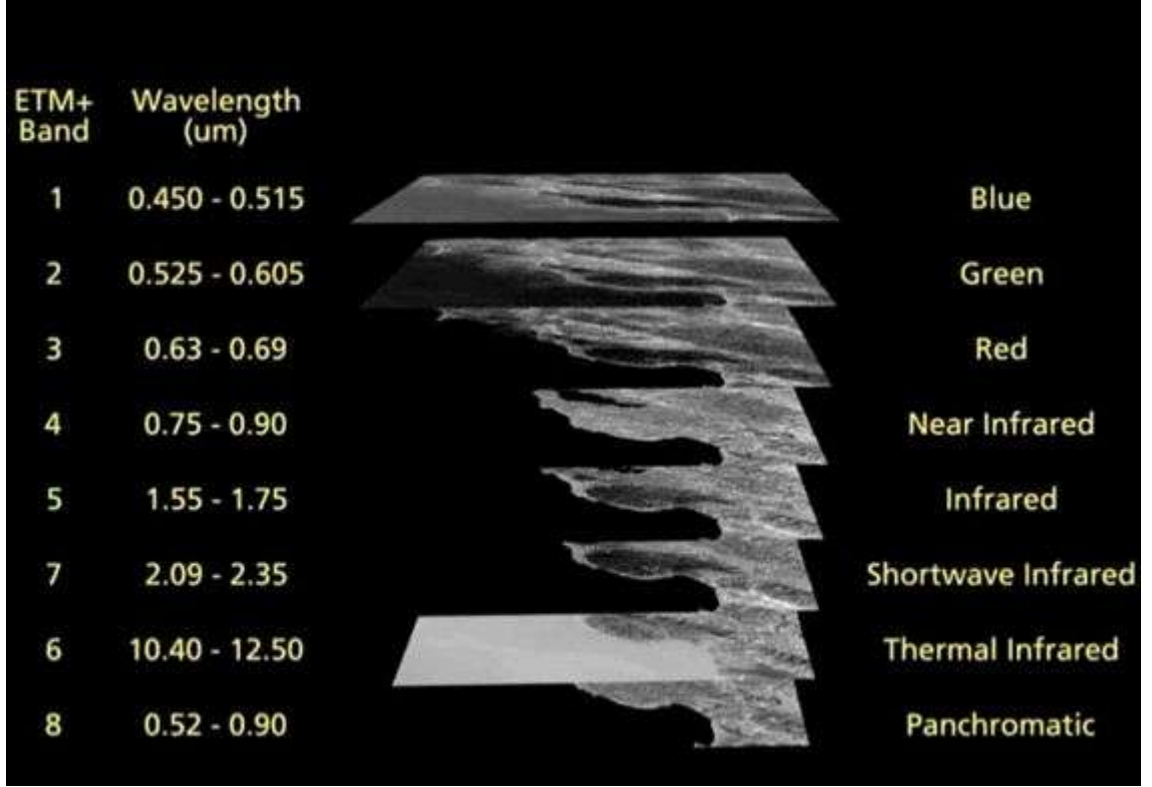
Şekil 2.5 Aynı bölgenin farklı çoklu bant görüntüsü

a. RGB, b. Yalancı renk

Multispektral görüntüleme bantlar devamlı, birinin bittiği dalga boyunda diğeri başlıyor olabileceği gibi ayırık da olabilir. Bu sensörler, görünür dalga boylarını taradıkları gibi, spektrumun NIR, SWIR, MWIR ve LWIR kısımlarını taradığı uygulamalar görülmektedir.

Bu teknoloji, uzaktan algılama uydularında da kullanılmakta olup, ülkemiz uydularından RASAT uydusu için bu sensörlerin spektral bantları 1. Bant: 0,42 – 0,55  $\mu\text{m}$  (Mavi), 2. Bant: 0,55 – 0,58  $\mu\text{m}$  (Yeşil), 3. Bant: 0,58 – 0,73  $\mu\text{m}$  (Kırmızı) bantlarından oluşmaktadır. Göktürk-2 uydusunda ise Multispektral görüntüleyici 4 spektral banttan oluşup; kırmızı, mavi ve yeşil yanında yakın kızılötesi banttan görüntü alan dördüncü bir sensör bulunmaktadır (<http://www.uzay.tubitak.gov.tr> 2013).

Uzaktan algılama uydularında genel olarak bu algılayıcılardan birden fazlası bir arada kullanılabilir. Ülkemizin sahip olduğu RASAT ve Göktürk-2 uyduların, pankromatik algılayıcı ile Multispektral algılayıcılar bir arada kullanılmıştır. Bunun yanında çok daha geniş bantları bir arada algılayan farklı uydular da mevcuttur.

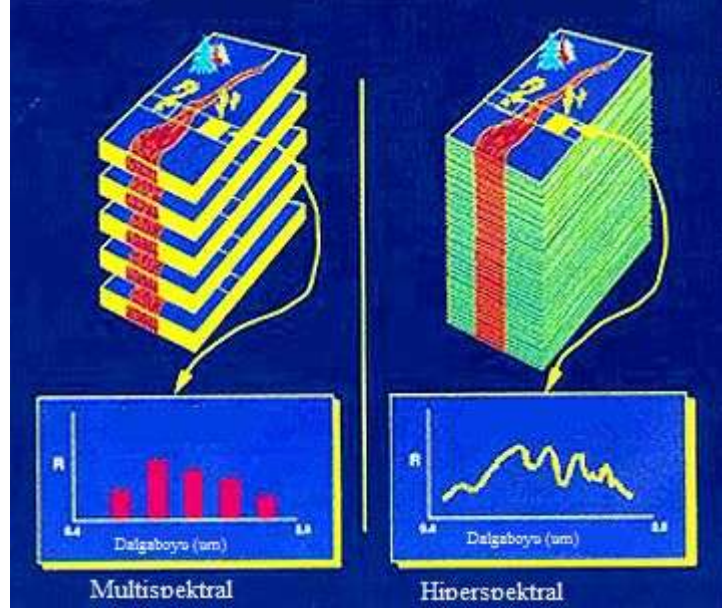


Şekil 2.6 Landsat 7 Bant Karşılaştırması

1999’da NASA tarafından üretilip yörüngeye oturtulan Landsat 7 uydusunun bantları şekil 2.6 ile gösterilmiştir (<http://landsat.gsfc.nasa.gov> 2013).

### 2.2.3 Hiperspektral görüntüleme

Hiperspektral görüntüleyiciler, multispektral görüntüleyicilere göre elektromanyetik spektrumun çok daha geniş bölgesini tarayan, sayıca çok daha fazla banttan oluşan ve çok daha dar ve devamlı dalga boyu aralıklarında sensörleri barındıran görüntüleyicilerdir. Şekil 2.7 ile (<http://rst.gsfc.nasa.gov> 2013) gösterildiği gibi, multispektral görüntülerde bantlar arasında aralık vardır ve bantlar daha geniştir, hiperspektral görüntülerin dalga boylarına bakıldığında ise dar ve devamlı dalga boyları görülür. Multispektral görüntüleyiciler gibi hiperspektral görüntüleyiciler de, görünür ışığın yanında spektrumun NIR, SWIR, MWIR ve LWIR kısımlarından görüntü alabilirler.



Şekil 2.7 Multispektral ve Hiperspektral görüntü özellikleri

Hiperspektral görüntüleyiciler, pankromatik ve multispektral görüntüleyiciler gibi pasif algılama ile çalışmakta olup, güneşten yansıyan ışınların farklı dalga boylarındaki enerjilerini kullanır.

Hiperspektral görüntüler; genellikle elektromanyetik spektrumun görünür, yakın kızılötesi ve kısa dalga kızılötesi bölümünde devamlı ve dar dalga boyu aralıklarında elde edilen verilerdir. Bir hiperspektral görüntü 50-250 arasında spektral banttandır. Her bant bir görüntü olduğundan, bir hiperspektral görüntü, bant sayısınınca elemanı olan bir görüntü kümesi şeklinde düşünülebilir.

Hiperspektral görüntüleyiciler çoğunlukla aynı görüntüleyicide görünür bant, NIR, SWIR bantlarından görüntü alacak şekilde tasarlanmakla birlikte; MWIR ve LWIR bantlarından veri alan hiperspektral görüntüleyiciler de vardır. Örneğin, hiperspektral görüntü işleme camiasında en çok kullanılan görüntüleyicilerden AVIRIS (Airborne Visible InfraRed Imaging Spectrometer) ve Hyperion 400 – 2500 nm aralığında ve 10 nm aralıklarla, ROSIS (Reflective Optics System Imaging Spectrometer) 430 – 860 nm aralığında ve 4 nm aralıklarla görüntü almaktadır.

### 2.2.3.1 Hiperspektral görüntüleyiciler

Yakın zamanda kullanılmış ve günümüzde halen kullanılmakta olan hiperspektral görüntüleyicilerin bazıları çizelge 2.1 ile verilmiştir (Shippert 2003). Çizelgede en alt satırda yer alan PRISMA uydusu henüz fırlatılmamış, yakın zamanda fırlatılması beklenmektedir.

Çizelge 2.1 Hiperspektral görüntüleyiciler

Algılayıcı	Üretici	Platform	Bant	Spektral Çözünürlük <sup>a</sup>	Spektral Aralık <sup>b</sup>
Hyperion	NASA USGS	Uydu (EO-1)	220	10	0,4 – 2,5
CHRIS	ESA	Uydu (Proba-1)	63	1,25	0,4 – 1,05
AVIRIS	NASA JPL	Uçak	224	10	0,4 – 2,5
RODIS	DLR	Uçak	115	4	0,43 – 0,86
CASI 1500	ITRES Research Limited	Uçak	288	2,5	0,36 – 1,05
SASI 600	ITRES Research Limited	Uçak	100	15	0,95 – 2,45
MASI 600	ITRES Research Limited	Uçak	64	32	3 – 5
TASI 600	ITRES Research Limited	Uçak	32	250	8 – 11,5
HYDICE	Naval Research Lab	Uçak	210	10	0,4 – 2,5
PROBE-1	Earth Search Sciences Inc.	Uçak	128	12	0,44 - 2,54
HyMap	Integrated Spectronics	Uçak	128	15	0,45 – 2,5
EPS-H	GER Corporation	Uçak	148	9	0,43 – 12,5
EPS-A	GER Corporation	Uçak	31	25	0,43 – 12,5
AVNIR	OKSI	Uçak	60	10	0,43 – 1,03
PRISMA	ASI	Uydu (PRISMA)	237	12	0,4 – 2,5

<sup>a</sup> nm, <sup>b</sup> µm

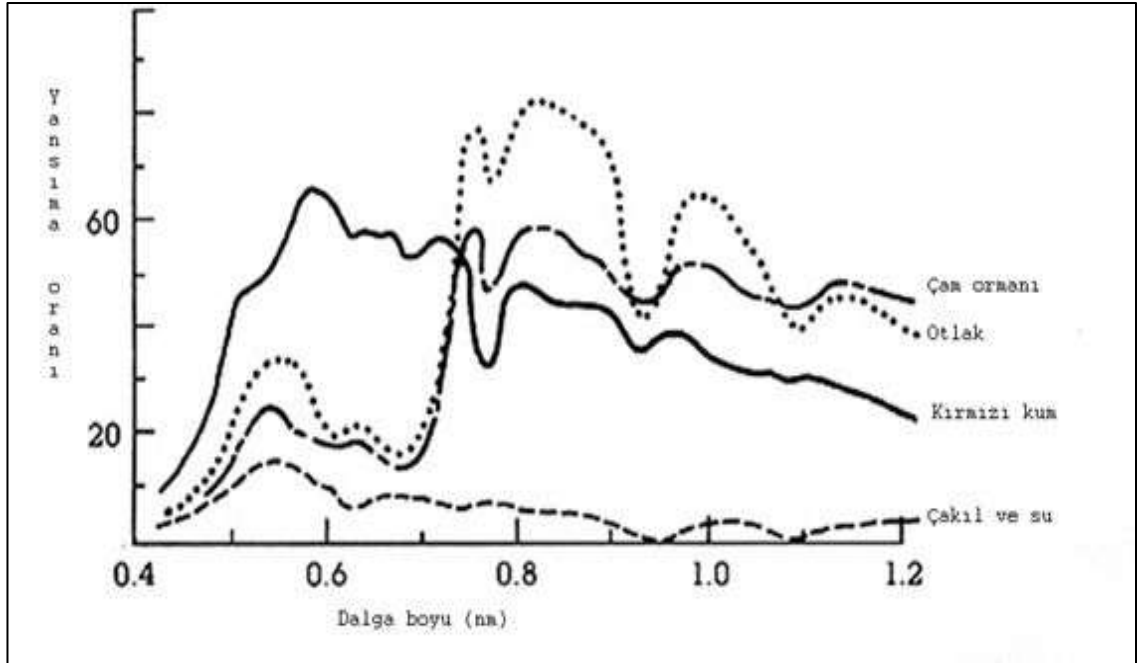
### 2.2.3.2 Hiperspektral görüntülerin kullanım alanları

Bir cismin, elektromanyetik spektrumun farklı bölgelerinde yaydığı, yansıttığı ve soğurduğu radyasyon birleşimine o maddenin spektral imzası denir. Bu spektral imzalar kullanılarak nesnelere birbirinden ayrılabilir. Hiperspektral görüntüler, sahip olduğu geniş spektrum ve bant sayısı ile farklı spektral imzalara sahip nesnelere ayırmada kullanılırlar. Bu ayırım şu alanlarda kullanılabilir:

- Bitki türlerinin tespiti

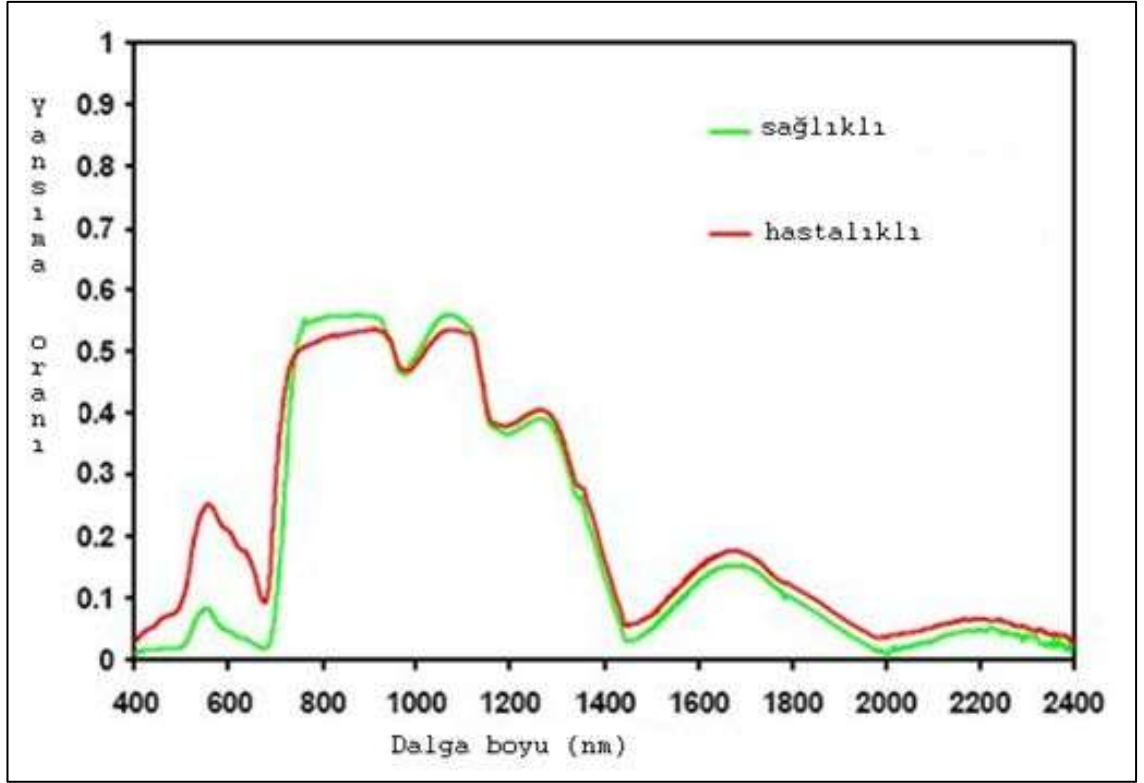
- Bitki hastalıklarının tespiti
- Atmosferik olayların tespiti
- Arazi şekillerinin tespiti
- Jeolojik haritalama
- Maden ve mineral tespiti
- Petrol yataklarının tespiti
- Denizde çözülmüş organik maddelerin tespiti
- Yangın tespiti
- Hedef tespiti
- Su ve su buharı tespiti

Örneğin, çam ormanı, otlak, kırmızı kum ve çakıl üzerindeki suyun spektral imzaları şekil 2.8 ile gösterilmiştir. Sağlıklı ve hastalıklı şeker pancarının 400 – 2400 nm aralığındaki yansıtma oranı şekil 2.9 ile gösterilmiştir ve buradan hangi dalga boyu kullanılarak sağlıklı ve hastalıklı şeker pancarını birbirinden ayrılabilceği gözükmektedir.



Şekil 2.8 Ormanlık bir arazinin spektral imzası



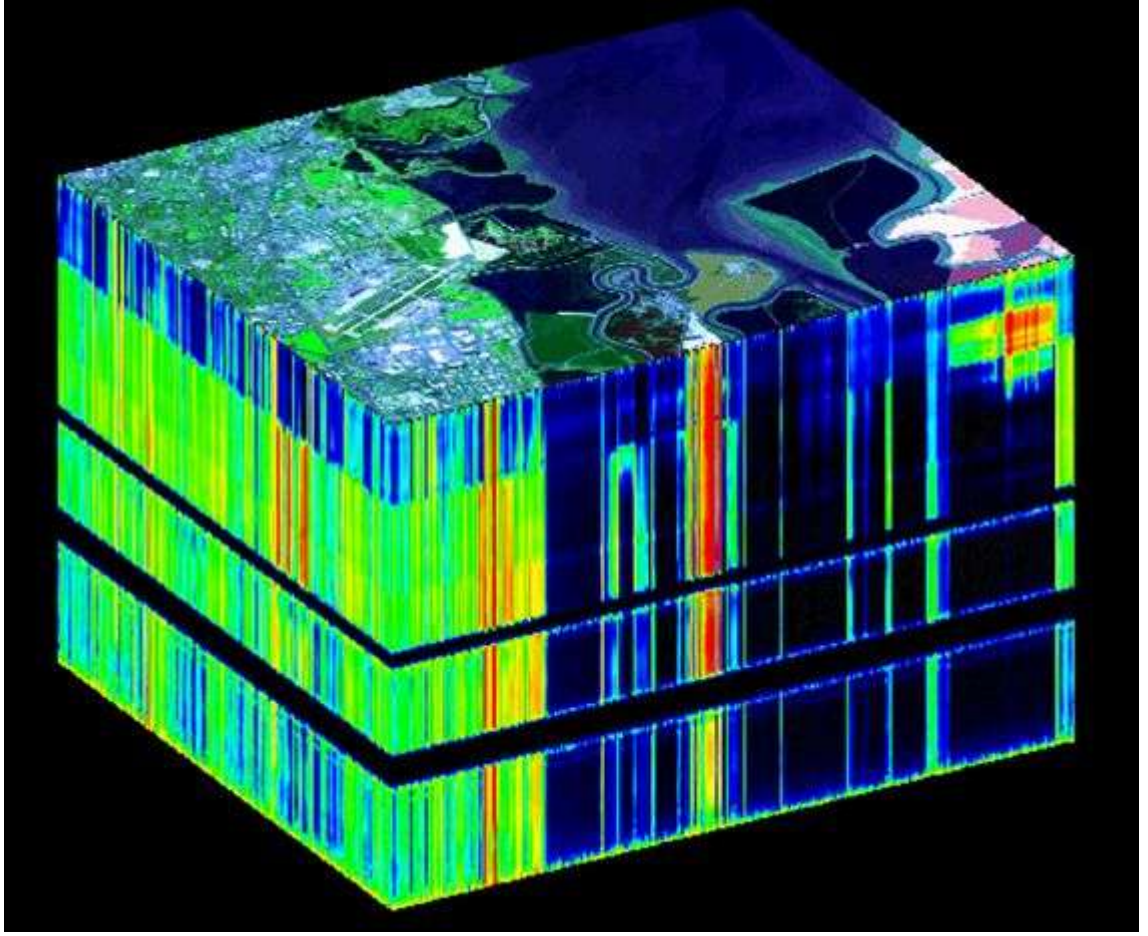


Şekil 2.9 Sağlıklı ve hastalıklı şeker pancarının spektral imzası

### 2.2.3.3 Hiperspektral verilerin gösterimi

Hiperspektral görüntüler, çok fazla sayıda dar dalga boyu aralıklı ve devamlı bantları içerdiğinden genel olarak görüntü verisi küpü şeklinde alınır ve saklanır.

Örnek olarak, hiperspektral görüntü işleme alanında en çok kullanılan görüntüleyici AVIRIS'in 20 Ağustos 1992'de NASA ER-2 uçağı ile 20000 metre yükseklikten San Francisco körfezinin güney ucundaki Moffett Field California'dan aldığı hiperspektral görüntü, veri küpü olarak şekil 2.10 ile gösterilmiştir (<http://aviris.jpl.nasa.gov> 2013).

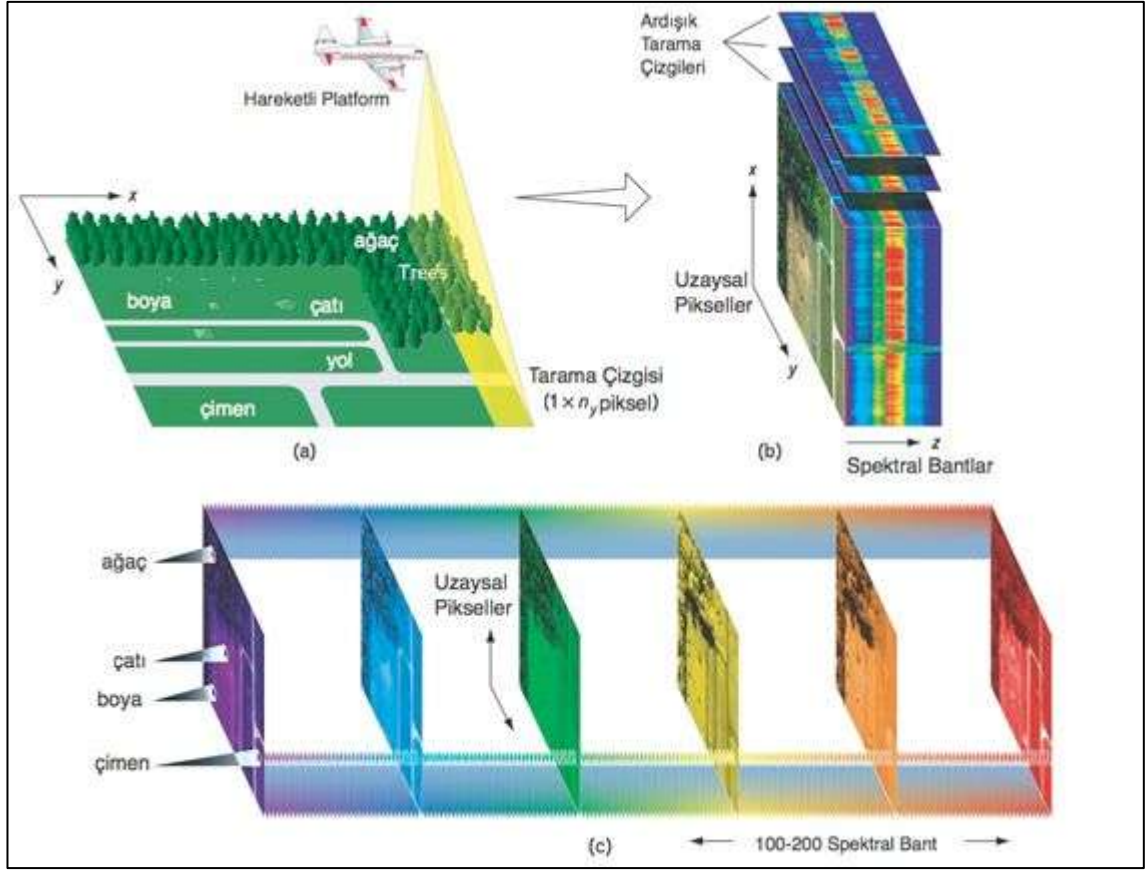


Şekil 2.10 Hiperspektral Veri Küpü

Bu veri küpünde, AVIRIS'teki 224 bant, en küçük dalga boyundan (400 nm) en büyüğe doğru (2500 nm) üstten alta sıralanmıştır.

#### 2.2.3.4 Hiperspektral görüntülerin elde edilmesi

Hiperspektral görüntüler, diğer uzaktan algılama görüntülerine benzer şekilde elde edilir. Uydu ya da uçak gibi hareketli bir platforma yerleştirilen hiperspektral algılayıcı, altındaki sahne hareket ettikçe, bir çizgiyi tarayarak görüntülemeye devam eder (Şekil 2.11.a). Platform ilerledikçe şekil 2.11.b'de gösterildiği gibi ardışık tarama çizgileri elde edilir ve hiperspektral veri küpü oluşturulur (Shaw ve Burke 2003).



Şekil 2.11 Hiperspektral görüntülerin elde edilmesi (Shaw ve Burke 2003)

a. Görüntünün alınması, b. Veri küpü oluşturulması, c. Bantların gösterimi

Görüntülemenin tamamlanmasının ardından, aynı dalga boyundaki sensör tarafından alınan tarama çizgilerindeki veriler, görüntünün bir bandı olarak bir arada tutulur, farklı dalga boyundaki veriler farklı bant olarak ifade edilir (Şekil 2.11.c).

### 2.3 Çözünürlük

Bir görüntünün çözünürlüğü, o görüntünün sahip olduğu ayrıntı miktarı olarak tanımlanabilir. Çözünürlüğü daha yüksek olan bir görüntü, daha fazla bilgi içerir. Genel olarak, görüntüde yer alan piksel sayısı olarak bilinse de; piksel sayısının (uzamsal çözünürlük) yanında, o pikselin kaç adet bant bilgisine sahip olduğu (spektral çözünürlük) ve her bant bilgisinin kaç bitlik veri olarak tutulduğu (radyometrik çözünürlük) bilgilerini de içerir.

Farklı çözünürlük kavramları, sıradaki bölümlerde anlatılmaktadır.

### 2.3.1 Uzaysal çözünürlük

Yer Örnekleme Mesafesi olarak da bilinen uzaysal (uzamsal) çözünürlük, yerde bir pikselin temsil ettiği alandır. Uzaktan algılama ile elde edilmiş bir görüntüde yan yana yer alan iki piksel arasındaki mesafenin yerdeki iz düşümü olarak bilinir.

Genel olarak yatay ve dikeydeki uzaysal çözünürlükler, bir biri ile eşitir. Uzaysal çözünürlük arttıkça, bir pikselin yerde temsil ettiği alan azalır. Örneğin, RASAT uydusu Pankromatik kamerasının uzaysal çözünürlüğü 7,5 metre, daha yüksek çözünürlüklü Göktürk-2 uydusunun Pankromatik kamerasının uzaysal çözünürlüğü 2,5 metredir (<http://www.uzay.tubitak.gov.tr> 2013).

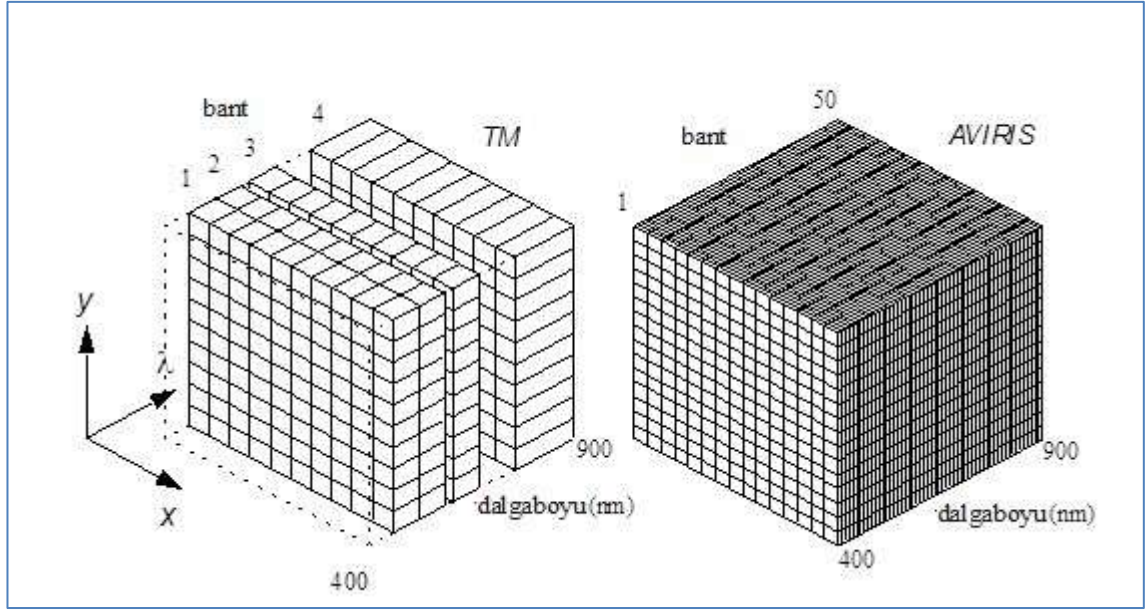
Uzaysal çözünürlük, görüntüleyicide bulunan piksel sayısı ile doğru, kamerayı taşıyan platformun yere uzaklığı ile ters orantılı olarak artar. Ayrıca görüntüleyicinin görüntüleme açısı ya da yerde görüntülediği alan ile de bağlantılıdır.

Örneğin, üzerinde yatayda 1000 piksele sahip ve bulunduğu yükseklikte 15 km görüntü genişliği olan bir görüntüleyicinin çözünürlüğü 15 metre olarak bulunur.

Şekil 2.11.a'da gösterildiği gibi, hareketli platform üzerinde çizgisel görüntüleme yapan görüntüleyiciler için, dikey eksenindeki görüntüleyicinin çözünürlüğü, görüntüleyicinin görüntüleme frekansına ve platformun yer hızına bağlıdır. Platform yavaşladıkça ya da görüntüleyicinin görüntüleme sıklığı arttıkça uzaysal çözünürlük artar.

Şekil 2.11.b'de  $x$  ve  $y$  ekseninde yer alan piksellerin görüntülediği alan uzaysal çözünürlüğü ifade eder.

Yer örnekleme mesafesi 30 metre olan LANDSAT TM ile 20 metre olan AVIRIS görüntülerinin karşılaştırılması şekil 2.12 ile verilmiştir. Uzaysal çözünürlüğün yer aldığı  $x$  ve  $y$  eksenlerinde aynı karede çözünürlüğü daha yüksek olan AVIRIS görüntüsünde daha fazla piksel bulunmaktadır (Schowengerdt 2006).



Şekil 2.12 Spektrel ve uzaysal çözünürlükler (Schowengerdt 2006)

### 2.3.2 Spektrel çözünürlük

Spektrel çözünürlük, çok bantlı bir görüntüde ardışık iki bant arasında yer alan elektromanyetik spektrumun genişliği olarak tanımlanır.

Bantlar arası spektrel çözünürlüğü sabit olması zorunluluğu yoksa bile çoğu görüntüleyicide spektrel çözünürlük belli bir toleransla sabit olarak tutulmuştur. Spektrel çözünürlük arttıkça, iki bant arasında algılanan bölgenin boyu azalır.

Örneğin; EO-1 uydusu üzerindeki Hyperion hiperspektrel görüntüleyici, 400 – 2500 nm aralığında devamlı 220 bant ile görüntü almaktadır (Folkman vd. 2001). Dolayısı ile 2100 nm'lik aralıkta bulunan 220 bandın spektrel çözünürlüğü 9,5 olarak bulunur.

Şekil 2.11.c'de yer alan görüntülerin arasındaki spektrel mesafe, spektrel çözünürlüğü ifade eder.

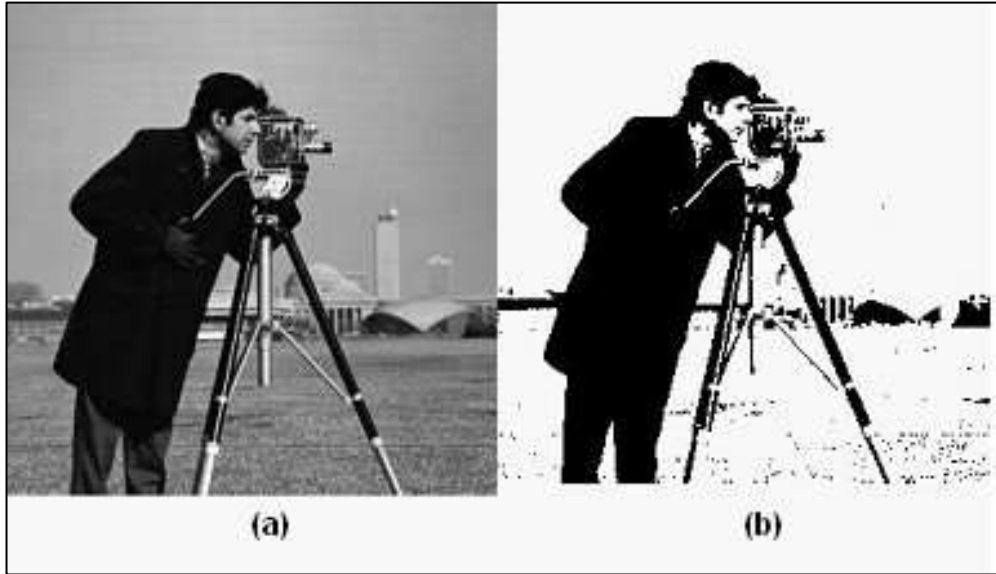
Multispektrel bir görüntüleyici olan Landsat TM ile hiperspektrel görüntüleyici AVIRIS 400-900 nm dalga boylarındaki görüntülerinin karşılaştırılması şekil 2.12 ile verilmiştir. Dalga boyunu gösteren  $\lambda$  ekseninde LANDSAT TM görüntüleyicisinde spektrel

çözünürlüğü değişen ve aralıklı 4 bant varken, AVIRIS'te ardışık olarak devamlı olan 50 bant bulunmaktadır (Schowengerdt 2006).

### 2.3.3 Radyometrik çözünürlük

Radyometrik çözünürlük, algılayıcının elektromanyetik enerjinin büyüklüğüne karşı olan algılama hassasiyetini ifade eder. Diğer bir ifade ile radyometrik çözünürlük, algılayıcının kaydettiği enerjide ne kadar ayırt edebilme kabiliyetini olduğunu gösterir. Bir görüntü üzerinde ise radyometrik çözünürlük, bir pikselin bir bandının kaç bitle saklandığı ile anlaşılır.

Radyometrik çözünürlük arttıkça, görüntüde algılanan enerjideki daha küçük farklılıklar algılanabilir.



Şekil 2.13 Aynı görüntünün farklı radyometrik çözünürlükte ifadesi

a. 8bit görüntü, b. 1bit görüntü

Radyometrik çözünürlük genel olarak 2'nin kuvvetleri olarak ifade edilir. Mesela 8 bit veri kaydetme hassasiyetine sahip olan bir algılayıcıda veya her pikselin 8 bit ile bir bandının ifade edildiği görüntüde gelen enerji  $2^8$  yani 256 farklı değer ile temsil edilir.

Şekil 2.13 ile aynı sahnenin farklı radyometrik çözünürlüklerle olan gösterimi verilmiştir. Şekillerden a ve b görüntüsünün ikisinin de spektral çözünürlüğü tek bant ve uzaysal çözünürlükleri yani piksel sayıları eşit olmakla birlikte radyometrik çözünürlüğü daha yüksek olan a şeklinde daha fazla detay vardır.

## **2.4 Hiperspektral Özet**

Elektromanyetik spektrumun çok daha geniş bölgesini taradığı ve dar ve devamlı dalga boyu aralıklarında veri alındığı için, hiperspektral görüntüler, sayıca fazla banttan oluşur. Sahip olduğu geniş spektrum ve bant sayısı sayesinde çok fazla kullanım alanı vardır.

Ancak yüksek boyut, dolayısı ile yüksek verinin getirdiği hesap yükü bu görüntülerin işlenmesi ve sınıflandırılmasını zorlaştırmaktadır. Ayrıca yüksek boyut, sınıflandırma doğruluğunda düşüşe neden olmaktadır. Bu nedenlerle, sınıflandırma öncesi hiperspektral görüntülerin boyutunun indirgenmesi gerekir.

Bu çalışmada kullanılan boyut indirgeme teknikleri, Bölüm 3 ile açıklanmıştır.

### **3. BOYUT İNDİRGEME**

#### **3.1 Giriş**

Boyut indirgeme; çok boyutlu bir uzaydan, az boyutlu uzaya yapılan dönüşüm olarak tanımlanabilir. Bu işlem, bir verinin gösterim kolaylığının yanında, veri işlenmesi sırasında oluşan işlem yükünü hafifletmek ve dolayısı ile işlem süresini kısaltmak için kullanılır. Bunun yanında, hiperspektral görüntülerin bir takım sınıflandırma zorluklarının da önüne geçilmiş olur. Hughes olgusu olarak adlandırılan etki, az sayıdaki eğitim verisi ile çok sayıdaki özniteliğin sınıflandırması ve sonuç olarak da sınırlı sayıdaki veri ile boyut arttıkça sınıflandırma doğruluğunun düşmesidir (Landgrebe 2003). Hughes olgusu, “boyut laneti (curse of dimensionality)” olarak da isimlendirilmektedir.

Boyut indirgeme; hiperspektral görüntülerde, yüksek boyutundan kaynaklanan problemlerin aşılması için, görüntülerin işlenmesinde etkili ve gerekli bir ön işleme fazıdır. Amaç, daha az veri ve daha az hesap yükü ile yüksek doğruluk yüzdesi elde etmektir.

Boyut indirgeme işlemi genel olarak iki farklı yolla yapılabilir. Bunlar, öznitelik çıkarımı (feature extraction) ve öznitelik seçimidir (feature selection) ve sıradaki bölümlerde anlatılmıştır.

#### **3.1.1 Öznitelik çıkarımı**

Öznitelik çıkarımı işleminde, eldeki veri farklı yöntemlere özel bir uzayda dönüşüme tabi tutulup farklı uzaylarda temsil edilir ve o uzayda öznitelik çıkarımı işlemi yapılır. Dönüşüm sonrası elde edilen kümedeki öznitelikler, orijinal verideki özniteliklerin bir kombinasyonudur.

Hiperspektral görüntülerin boyutunun indirgenmesinde öznitelik çıkarımı için farklı yöntemler vardır. Bu tez çalışmasında, öznitelik çıkarımı yöntemi kullanarak boyut indirgeyen teknikler kullanılmıştır ve bu teknikler Bölüm 3.2’de anlatılmıştır. İlgili



yöntemler, hiperspektral görüntüye dönüşümler uygular ve hedeflenen boyut kadar özneliğe indirgeyerek yeni bir veri elde eder.

Yöntemler; öğreticili ve öğreticisiz, doğrusal olan ve olmayan olarak gruplandırılabilir.

### **3.1.2 Öznitelik seçimi**

Öznitelik seçimi yöntemi aynı zamanda alt küme yöntemi olarak da adlandırılır. Verideki öznelik kümesinin bir altkümesini seçerek boyut indirgeme yapılır, dolayısı ile dönüşüm sonrasındaki her öznelik, orijinal verideki özneliklerden biridir.

Hiperspektral görüntülerde öznelik seçimi ile boyut indirgemek demek, görüntüye herhangi bir dönüşüm uygulamadan görüntüde bulunan bantların uygulamaya yönelik önemli olanlarını seçip diğerlerini atarak yapılır. Hiperspektral görüntülerde her bant bir boyutu temsil ettiği için öznelik seçimi, bant seçimi olarak da adlandırılır. Herhangi bir dönüşüm yapmadan; görüntüdeki bantların en fazla bilgi taşıyanları, en fazla ayırt edici olanları, ya da birbiri arasında ilintisi en az olanları seçilerek bant sayısının yani veri boyutun azaltılması işlemi yapılır. Sınıf bilgisinin kullanımına göre öğreticili ya da öğreticisiz olarak yapılabilir.

## **3.2 Çalışmada Kullanılan Boyut İndirgeme Yöntemleri**

Boyut indirgeme teknikleri genel olarak doğrusal (linear) ve doğrusal olmayan (non-linear) olarak ikiye ayrılır.

Doğrusal teknikler, yüksek boyutlu uzaydaki verinin doğrusal yakın ya da doğrusal bir alt uzayda temsil edilebileceğini kabul eder. Diğer yandan doğrusal olmayan yöntemler ise bu varsayımı kabul etmez ve yüksek boyutlu uzaydaki çok daha karmaşık verileri düşük uzayda tanımlayabilir (Van der Maaten vd. 2009).

### **3.2.1 Doğrusal yöntemler**

#### **3.2.1.1 Temel bileşen analizi**

Temel Bileşen Analizi (Principal Component Analysis, PCA), verileri deęişintinin en yüksek oldukları deęerlere göre birbirleri ile ilişkisiz (bağımsız) uzaylara izdüşüm yapan doğrusal ve öğreticisiz boyut indirgeme yöntemidir.

Değişkenleri; daha az sayıda, deęişintisi en çok olacak temel bileşenlerle ifade ederek boyut indirgemeye çalışır. Dönüşüm sonrası oluşan ilk temel bileşen en büyük deęişinti deęerine sahip olur. İkinci bileşen de ilk bileşene ortogonal olan ve ikinci büyük deęişinti deęerine sahip olan bileşendir. Sıralama bu şekilde devam eder (Theodoridis vd. 2010). Dönüşüm sonrası, orijinal verideki deęişkenler kadar temel bileşen bulunur. En büyük deęişinti deęerine sahip temel bileşenden itibaren, hedeflenen boyut sayısı kadar temel bileşen alınır ve böylece boyut indirgenmiş olur.

Farklı alanlarda Hotelling dönüşümü, amprik ortogonal fonksiyonu, Tekil Deęer Ayırımı olarak da adlandırılır (Fodor 2002).

PCA, hiperspektral görüntülerde boyut indirgeme alanında en fazla kullanılan ve referans olarak alınan yöntemdir.

#### **3.2.1.2 Faktör analizi**

İlk olarak psikoloji alanında kullanılan istatistiksel bir yöntem olan faktör analizi (Factor Analysis, FA); verinin bilinmeyen ve genelde ölçülemeyen ortak faktörlere dayalı olduğunu kabul eder. Verideki korelasyonu tanımlayan bu ortak faktörleri arayarak veriyi daha az deęişkenle tanımlamaya çalışır. Böylece, birbirleriyle ilintili deęişkenler yerine daha az sayıda ve birbirinden bağımsız deęişken kümeleri elde edilir (Fodor 2002).

Faktör analiz de PCA gibi, oluşturduğu faktörler için ortogonal yapılar kurar ancak PCA'den farklı olarak, faktör modelleri verinin ölçeğine baęlı deęildir.

Faktörleri hesaplamak için sıkça kullanılan iki yöntem vardır. Bunlar, Temel Faktör Yöntemi ve En Büyük Olabilirlik (Maximum Likelihood) Yöntemi'dir (Johnson ve Wichern 2002). Bu çalışmada faktörler, En Büyük Olabilirlik (Maximum Likelihood) Yöntemi ile hesaplanmıştır (Rubin ve Thayer 1982).

### **3.2.1.3 Bölgeselliği koruyan izdüşüm**

Bölgeselliği Koruyan İzdüşüm (Locality Preserving Projections; LPP), Laplace işleci öz-haritalarında kullanılan ceza fonksiyonunu optimize etmeye çalışan öğreticisiz, doğrusal boyut indirgeme yöntemidir (He ve Niyogi 2004). Yöntem, yerel doğrusal olmayan yöntemlerden türetilmiştir. LPP, doğrusal olan ve doğrusal olmayan yöntemlerin avantajlarını bir arada barındırmaya çalışır (Van der Maaten 2007). Laplace Özharitalarının ceza fonksiyonunu minimize etmek için doğrusal bir dönüşüm bulmayı amaçlar.

Doğrusal bir yöntem olmasına karşın, çok katmanlı verilerin doğrusal olmayan yapısını ifade etmede kullanılır.

Bölgesellik, verideki noktaların komşu sayısı ile tanımlanır. Çalışmalar sırasında farklı komşu sayısı ile denemeler yapılmış, en iyi sonuç  $k=12$  ile alınmış ve bu şekilde kullanılmıştır.

### **3.2.1.4 Komşuluk korumalı içe katma**

Komşuluk Korumalı İçe Katma (Neighborhood Preserving Embedding, NPE), Doğrusallığı Koruyan İzdüşüm gibi doğrusal olmayan lokal tekniklerdeki ceza fonksiyonunu optimize etmeye çalışır ve doğrusal olmayan tekniklerden türetilmiştir. Yerel Doğrusal İçe Katma yönteminin doğrusal olarak yaklaşıklama ile çözmeye çalışan sürümüdür (He vd. 2005).

Bölgesellik veri noktalarının korunumunu amaçladığı için, evrensel dönüşümlerdeki tepe noktaların bozunumlarından etkilenmez.

Bölgesellik, verideki noktaların komşu sayısı ile tanımlanır. Çalışmalar sırasında farklı komşu sayısı ile denemeler yapılmış, en iyi sonuç  $k=1$  ile alınmış ve bu şekilde kullanılmıştır.

### 3.2.1.5 Olasılıksal PCA

Olasılıksal PCA (Probabilistic Principal Component Analysis, PPCA), PCA'in veri vektör asal eksenlerini Beklenti Enbüyüklenme (Expectation-Maximization; EM) kullanarak bulmaya çalışan öğreticisiz ve doğrusal bir yöntemdir (Tipping ve Bishop 1999).

FA gibi, veriyi bilinmeyen ortak faktörlere dayandırır.

PPCA; boyut sayısı binlerin üzerinde olan ve veri sayısı çok fazla iken az sayıda özdeğer (eigenvalue) ve özvektör (eigenvector) hesaplama konusunda PCA'e göre daha yüksek başarımlar gösterir. Bunun dışında, veri setinde eksik olduğu durumlarda EM kestirimci sayesinde en büyük olabilir veri noktalarını hesaplayarak çözüm sağlar (Roweis 1998).

### 3.2.1.6 Doğrusal ayraç analizi

Doğrusal Ayraç Analizi (Linear Discriminant Analysis, LDA), boyut indirgendikten sonra bile sınıf ayrımını korumaya çalışan bir yöntemdir. Öğreticili bir yöntem olup, eğitim için verilen sınıflarda sınıf içi değişinti ( $S_W$ ) ve sınıflar arası değişinti ( $S_B$ ) değerlerini göz önüne alarak (3.1) ile verilen dönüşüm matrisi ( $W$ ) hesaplanır (Theodoridis vd. 2010):

$$W = \operatorname{argmax}_W \frac{W^T S_B W}{W^T S_W W} \quad (3.1)$$

PCA maksimum değişintiye göre boyut indirgeme yapıp, sınıflar arası değişintileri göz önünde bulundurmadığı için, yapılan işlem her zaman için sınıflandırma için uygun olmayabilir. LDA ise sınıflar arası ayrımın daha iyi yapılabilmesine olanak sağlayacak

şekilde izdüşüm yapılacak eksenleri belirler.

Bir kısıt olarak, LDA ile yapılan boyut indirgeme çalışmalarında, veri kümesinde yer alan  $N$  sınıf için, maksimum  $N-1$  boyuta indirgenebilmektedir. Bu çalışmada LDA için diğer yöntemlerden farklı bir hedef boyut seçilmesinin sebebi, bu kısıttır.

### 3.2.1.7 Komşuluk bileşen analizi

Komşuluk Bileşen Analizi (Neighbourhood Components Analysis, NCA), en yakın komşuluk yöntemine benzer çalışan, yakın komşuluk metriği olarak Mahalanobis metriğini kullanan öğrenme temelli yöntemdir (Goldberger vd. 2004). Öğreticili bir yöntem olduğu için, sınıf bilgilerine ihtiyaç duyar.

Mahalanobis uzaklığı ( $D(x,y)$ ),  $x$  ve  $y$  noktaları için,  $A$  bir yarı kesin pozitif matris olmak üzere:

$$D(x,y) = (x - y)' A (x - y) \quad (3.2)$$

şeklinde hesaplanır

### 3.2.1.8 En fazla daralan metrik öğrenme

En Fazla Daralan Metrik Öğrenme (Maximally Collapsing Metric Learning, MCML) Mahalanobis uzaklığını kullanarak sınıflandırma yapan, aynı sınıftaki verilerin metrik olarak birbirine en yakın olacağını ve bu sebeple tek bir noktaya daraltılabileceğini kabul eden doğrusal ve öğreticili metrik öğrenme yöntemidir (Globerson ve Roweis 2005).

## 3.2.2 Doğrusal olmayan yöntemler

### 3.2.2.1 Oto-kodlayıcı

Oto-Kodlayıcı (Auto Encoder, AE), çok katmanlı ileri beslemeli yapay sinir ağ tabanlı

bir boyut indirgeme yöntemidir. Yöntem, kodlanmış verileri sıkıştırılmış bir gösterimde bulmaya çalışır.

Oto-kodlayıcı; PCA'in, doğrusal olmayan uyarlanabilir çok katmanlı bir kodlayıcı ve bir çözücünden oluşan genelleştirmesidir. Kodlayıcı, çok boyutlu veriyi düşük boyutlu koda çevirirken; çözücü de kodu veriye çevirir (Hinton ve Salakhutdinov 2006).

Genel olarak çok fazla bağlantıya sahip oldukları için geri yayılım algoritması çok yavaş yakınsar ve yerel minimumlarda takılabilir. Bu dezavantajların önüne geçmek için Kısıtlı Boltzmann Makineleri ile bir ön eğitim yapılmaktadır (Hinton vd. 2006).

### **3.2.2.2 Eş ölçekli öznitelik haritalama**

Çok boyutlu veri, eğimli çok katmanlı gibi bir yapıda olduğunda; Öklid uzaklığı ile ölçülen uzaklık, eğimli yüzeydeki gerçek uzaklıktan farklı olacağından veri noktalarının yakınlığını hesaplamada yanıltıcı olabilir. Eş Ölçekli Öznitelik Haritalama (Isometric Feature Mapping, Isomap, ISO); bu problemi, eğrisel uzaklığı koruyarak çözmeye çalışan bir yöntemdir (Tenenbaum 1998).

Bir topoloji üzerindeki tüm ikili mesafeleri hesaplayıp tablolarda tutan Isomap, kısa devre denilen hatalardan dolayı düşük performans gösterebilmektedir (Van der Maaten 2007).

### **3.2.2.3 Hızlı maksimum değişinti çözücü**

ISO gibi, Hızlı Maksimum Değişinti Çözücü (Fast Maximum Variance Unfolding, FMVU) de bir komşuluk grafiği tanımlar ve tüm uzaklıkları burada tutar. FMVU'nun ISO'dan farkı, doğrudan çok katmanlı verileri açmaya çalışır. Bunu, veri noktaları arasındaki Öklid uzaklıklarını komşuluk grafiğindeki uzaklıkları değiştirmeden maksimize ederek yapmaya çalışır. Böylece çok katmanlının açılması sonrası, birbirine yakın noktalar yine yakın, uzak noktalar da uzak kalır ve çok katmanlının bölgesel geometrisi değişmez (Weinberger ve Saul 2006). Çok katmanlının açılmasından sonra problem bir yarı kesin programlama (Semidefinite Programming) ile çözülür (Van der

Maaten 2007).

#### 3.2.2.4 Çok katmanlı haritalama

Haritalama, yüksek boyuttaki veri noktalarının düşük boyuttaki koordinat sistemine atanmasıdır.

Çok Katmanlı Haritalama (Manifold Charting, MANC) İlk adımda EM algoritması kullanılarak belirlenen sayıda bölgesel faktör çözümleyicisi oluşturur (Ghahramani vd. 1996). Bu çalışmada, bölgesel faktör çözümleyicisi sayısı, çeşitli denemeler sonucu 2 olarak seçilmiştir. Böylelikle veri düşük boyutta bölgesel faktör çözümleyicilerin karışımı ile temsil edilmiş olur. Ardından yöntem, evrensel koordinatlardaki veri noktalarında yer alan doğrusal modeller arasındaki farkı ölçen ceza fonksiyonunu minimize etmeye çalışır (Brand 2002).

#### 3.2.2.5 Difüzyon haritaları

Difüzyon Haritaları (Diffusion Maps, DIF) orijinal olarak dinamik sistemlerde kullanılmaktadır. Veri grafiğinde Markov rastgele yürüyüş tanımlamaya dayanır (Lafon ve Lee 2006). Tanımlanan bu rastgele yürüyüşler ile veri noktalarının yakınlığı bulunmuş olur. Çalışmada rastgele yürüyüş sayısı 1 olarak seçilmiştir.

Yöntem, öncelikle bir veri grafiği oluşturur. Grafikteki kenarların ağırlıkları Gauss çekirdek fonksiyonu kullanılarak bulunur.

$$w_{ij} = e^{-\frac{x_i - x_j}{2\sigma^2}} \quad (3.3)$$

Burada  $\sigma$ , Gauss çekirdek fonksiyonunun değışintisidir ve 1 olarak seçilmiştir. Ardından normalizasyon yapılır ve sonucunda  $P$  matrisi bulunur.

$$p_{ij}^{(1)} = \frac{w_{ij}}{\sum_k w_{ik}} \quad (3.4)$$

Burada  $P$ , ileri geçiş matrisidir. Farklı rastgele yürüyüşler  $t$  ile ifade edildiğinde difüzyon uzaklıkları

$$D^t(x_i, x_j) = \frac{(p_{ik}^{(t)} - p_{jk}^{(t)})^2}{k \psi(x_k^{(0)})} \quad (3.5)$$

olarak bulunur.

### 3.2.2.6 Olasılıksal yakınlık içe katma

Olasılıksal Yakınlık İçe Katma (Stochastic Proximity Embedding, SPE), çok boyutlu ölçeklendirmenin ham gerilim fonksiyonunu minimize eden özyinelemeli (iterative) bir yöntemdir. Ham gerilim fonksiyonu şu şekilde verilir:

$$\phi(Y) = \sum_{ij} (d_{ij} - r_{ij})^2 \quad (3.6)$$

Burada  $r_{ij}$ , yüksek boyuttaki veri noktaları  $x_i$  ve  $x_j$  arasındaki yakınlığı,  $d_{ij}$  de bu noktaların düşük boyuttaki izdüşümleri  $y_i$  ve  $y_j$  arasındaki Öklid uzaklığını temsil eder (Van der Maaten 2007).

Yöntem bir başlangıç konfigürasyonu ile başlar,  $y_i$  ve  $y_j$  noktalarının yerini rastgele seçer. Sonrasında rastgele iki nokta seçerek ve bunların düşük boyuttaki koordinatları  $y_i$  ve  $y_j$ 'yi değiştirerek  $d_{ij}$ 'yi  $r_{ij}$  ile eşitlemeye çalışır (Agrafiotis 2003).

## 3.3 Boyut Seçimi

Yüksek boyut veri analiz camiasında, yüksek boyutlu verilerde farklı yöntemlerin işe yaramasının sebebinin, verinin gerçekten yüksek boyutlu olmadığı, daha ziyade yüksek boyuta yayılmış düşük boyutta ifade edilebilecek veri olduğu kabul görmüştür (Levina ve Bickel 2004). Bu sebeple; gerçek hayatta karşılaşılan çoğu yüksek boyutlu veride, fazla bilgi kaybı olmadan boyutun indirilebileceği kabul edilmektedir.

Elimizdeki verinin özelliklerini yansıtmak için gerekli minimum parametre sayısı



“Etkin boyut” olarak tanımlanmıştır. İdealde, boyut indirgeme sonrası hedeflenen boyutun etkin boyut olması gerekir (Van der Maaten 2007).

Etkin boyutu hesaplamak için farklı yöntemler vardır. Bunlar yerel ve evrensel yöntemler olarak ikiye ayrılır. Bu çalışmada yerel yöntemlerden “En Büyük Olabilirlik Kestirimi” (Maximum Likelihood Estimation, MLE) yöntemi kullanılmıştır (Levina ve Bickel 2004). Bu yöntem artan  $r$  yarıçaplı bir hiperkürenin kapsadığı veri noktası sayısını kestirir ve bu hiperküre içindeki veri noktalarını Poisson şeklinde modeller.

Etkin boyut; herhangi bir boyut indirgeme yönteminden bağımsız olarak hesaplanmakta olup, direkt veri üzerinde yapılan bir analize dayanmaktadır. Tez kapsamında yapılan boyut hesaplama detayları Bölüm 5.1.3’te anlatılmıştır.

Boyut indirgeme sonrası yapılacak sınıflandırma işlemi ve sınıflandırmada kullanılan veri ile ilgili detaylı bilgiler Bölüm 4’te verilmiştir.

#### **4. SINIFLANDIRMA**

Görüntü işlemede; sınıflandırma, piksellerin hangi sınıfa ait olduğunu belirleme işidir. Genel olarak her piksel, spektral bantlardaki değerlerden oluşan bir birim olarak ele alınır. Sınıfı bilinen diğer piksellerle karşılaştırılarak hangi sınıfa ait olduğu bulunur, böylelikle sınıflar oluşturulur. Görüntü üzerindeki bu sınıflar da belirli bölgeleri oluşturur. Sonunda her dijital görüntü bölümlerden oluşan ve farklı renk ya da sayılarla gösterilen bir haritaya dönüşür (Campbell 2002). Örnek olarak hiperspektral görüntülerdeki Sınıflandırma Haritaları verilebilir (Şekil 4.2, 4.4).

Teoride bu sınıflar homojendir; aynı sınıfa ait pikseller farklı sınıflardaki piksellere göre birbirine daha benzerdir. Ancak pratikte her sınıfın içinde farklılıklar olabilir (Campbell 2002).

1970'lerin başında Landsat uydusundan ilk görüntüler alınmaya başladığından beri, uzaktan algılama alanında birçok sınıflandırıcı kullanılmıştır.

##### **4.1 Kullanılan Veri**

Hiperspektral görüntülerin sınıflandırılması ve sınıflandırma doğruluğunun ölçülmesi çalışmalarında sonucunun test edilmesi amacıyla yapılan çalışmalarda belli başlı bir kaç görüntü kullanılmaktadır. Bunun sebebi deneyin başarımını sınıflandırma doğruluğu olarak ölçerken, uzaktan algılama ile görüntülenen sahnedeki yerin sınıf bilgisine duyulan ihtiyaçtır. Sınıflandırma Haritası olarak adlandırılan bu bilgi, her pikselde hangi sınıfın yer aldığını bildirir. Literatürde kullanılan görüntülerin kısıtlı olmasının sebebi, açık kaynak olarak sınıflandırma haritası yayımlanmış hiperspektral görüntüsünün sayıca az olmasıdır.

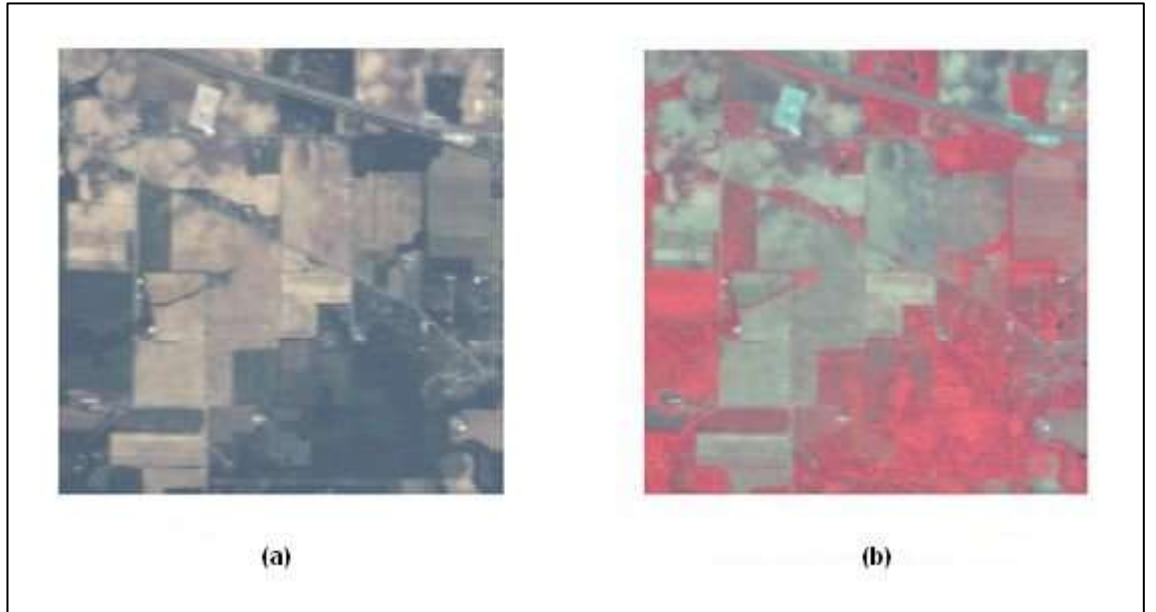
Görüntüler, sınıflandırma haritasının yanında gürültü içeren bantları elenmiş şekilde de yayımlanabilmektedir.

#### 4.1.1 AVIRIS - Indian Pines

AVIRIS Indian Pines, hiperspektral görüntü işleme alanında en fazla kullanılan görüntülerden biridir. Sınıflandırma haritasına sahiptir.

Görüntü, JPL tarafından üretilen AVIRIS görüntüleyici ile 12 Haziran 1992 tarihinde Amerika'da İndiana eyaleti bölgesinden West Lafayette'in 10 km batısından alınmıştır. Görüntü, Purdue Üniversitesi'nin internet sitesinden temin edilmiştir (<ftp://ftp.ecn.purdue.edu> 2013).

Görüntünün alındığı bölge bir görüntü test bölgesi olup, bu bölgenin sınıflandırma haritası çıkartılmıştır. Görüntü 145x145 pikselden oluşmakta, yerde 3,2 km x 3,2 km'lik bir bölgeyi kapsamaktadır. Spektral olarak 220 banttan oluşan görüntü 400-2500 nm aralığında alınmıştır. Görüntü, yayımlanmadan önce kamera üreticisi JPL tarafından verilen bilgilerle kalibre edilmiş, kameranın kullanmadığı 4 bant çıkartılıp 224 banttan 220'ye indirilmiştir. 220 bantlı kalibrasyon verileri görüntünün alındığı internet sitesinde yer almaktadır.



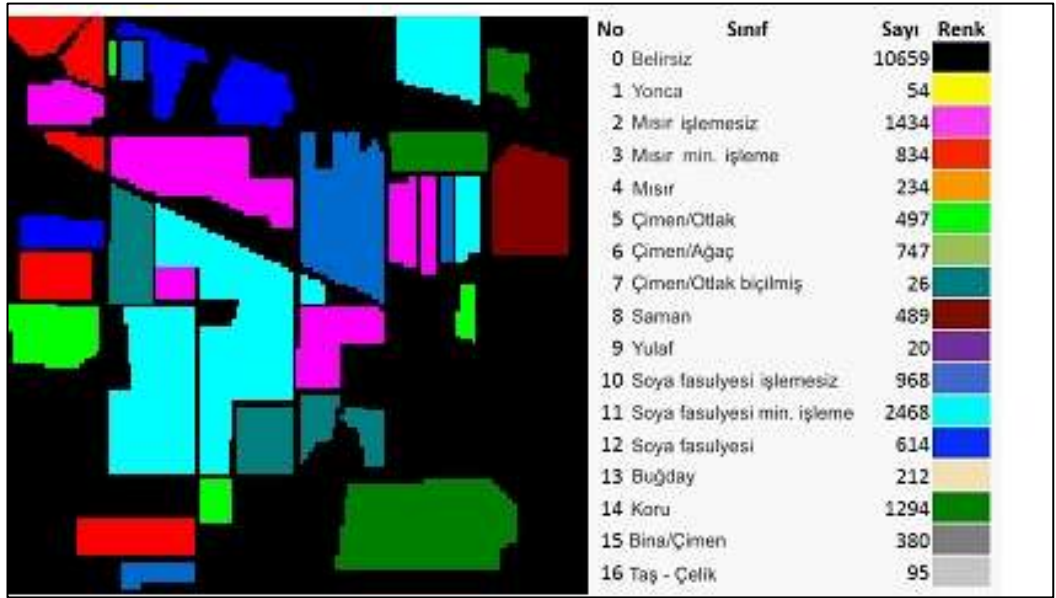
Şekil 4.1 Indian Pines görüntüsü gösterimi

a. RGB, b. Yalancı renk

Görüntü bantlarından kırmızı, yeşil ve maviye karşılık gelen bantlarından alınmış birer görüntü ve yalancı renk görüntü şekil 4.1 ile gösterilmiştir.

Görüntü 16 sınıftan oluşmaktadır. Bu sınıflar genelde ekin sınıfları olup, her sınıfın piksel sayısı, Sınıflandırma Haritasında temsil edilen renkler ve Sınıflandırma Haritası şekil 4.2’de yer almaktadır.

Çalışmada görüntü yayımlandığı şekliyle 220 bant ile değil; literatürde kullanıldığı şekilde, yani su buharı tarafından soğurulmuş bantlar ve gürültüden oluşan bantlar çıkardıktan sonra (104–108, 150–163 ve 220. bantlar) 200 bantlı haliyle kullanılmıştır.



Şekil 4.2 Indian Pines görüntüsünün Sınıflandırma Haritası

Ayrıca yine literatürde kullanılan şekilde, 400’den az eleman sayısı olan sınıfların çıkarılması sonucunda elde edilen 9 sınıflı sınıflandırma haritası kullanılmıştır. Ekinlerden oluşan bu sınıflar çizelge 4.1 ile verilmiştir.

Çizelge 4.1 Indian Pines görüntüsünde kullanılan sınıflar ve örnek sayıları

No	Sınıf	Örnek Sayısı
2	Mısır işlemez	1434
3	Mısır min işleme	834
5	Çimen/Otlak	479
6	Çimen/Ağaç	747
8	Saman	489
10	Soya fasulyesi işlemez	968
11	Soya fasulyesi min işleme	2468
12	Soya fasulyesi	614
14	Koru	1294

#### 4.1.2 ROSIS – Pavia University

Görüntü ROSIS (Reflective Optics System Imaging Spectrometer) tarafından İtalya'nın kuzeyindeki Pavia kendinden alınmıştır.

Görüntü 610x340 pikselden oluşmakta olup, görüntüde Pavia Üniversitesi'nin kampüs alanı yer almaktadır. Görüntünün uzaysal çözünürlüğü 1,3 m'dir. Spektral olarak 115 banttandır oluşan görüntü 430-860 nm aralığında ve 4 nm'lik spektral çözünürlükle alınmıştır. 115 bandın veri taşımayan 12 bandı çıkartılmış ve kalibrasyon yapılarak yayımlanmıştır ( <http://www.ehu.es> 2013).

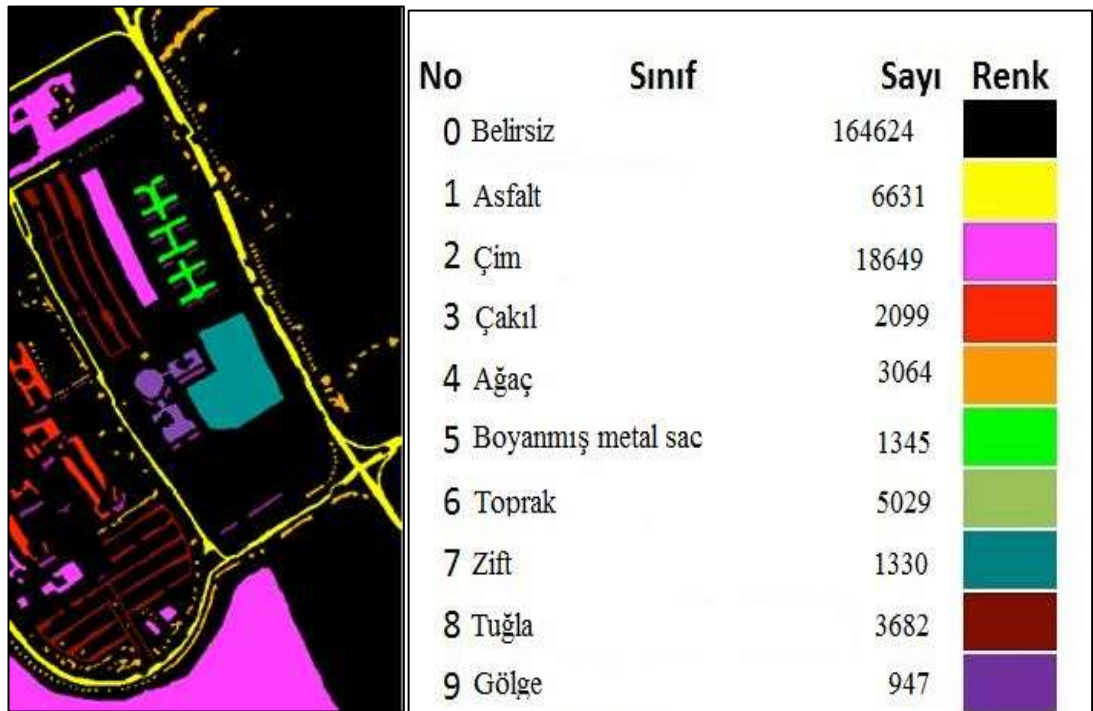
Görüntü bantlarından kırmızı, yeşil ve maviye karşılık gelen bantlarından alınıp birleştirilmiş RGB görüntü ve yalancı renk görüntü şekil 4.3 ile gösterilmiştir.

Pavia University görüntüsünde 9 arazi örtüsü sınıfı belirlenmiştir. Bu sınıflar, her sınıfın piksel sayısı, Sınıflandırma Haritasında temsil edilen renkler ve Sınıflandırma Haritası şekil 4.4 ile gösterilmiştir. Sınıflandırma haritasında sınıfı belirtilmiş toplam 42776 piksel bulunmaktadır.



Şekil 4.3 Pavia University çok bantlı görüntüsü

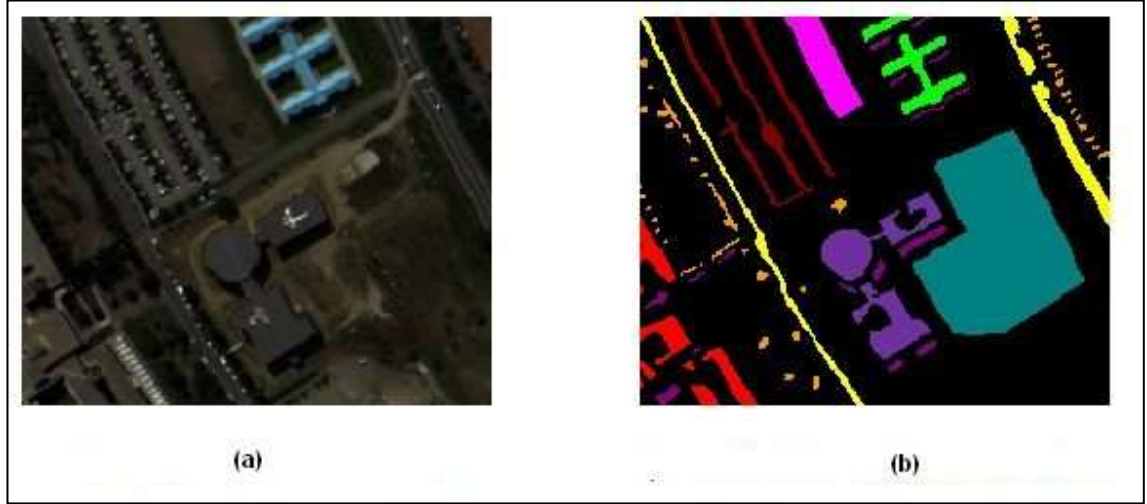
a. RGB, b. Yalancı Renk



Şekil 4.4 Pavia University görüntüsünün Sınıflandırma Haritası

Çalışmada, Pavia University görüntüsünün 190x230'luk bir alt kümesi kullanılmıştır. Böylelikle etkin piksel sayısı 15504 olmuştur.

Altküme görüntüsü ve altküme sınıflandırma haritası şekil 4.5 ile gösterilmiştir.



Şekil 4.5 Pavia University altküme verisi

a. RGB görüntü, b. Sınıflandırma haritası

Altküme görüntüsündeki sınıflar çizelge 4.2 ile verilmiştir

Çizelge 4.2 Pavia University görüntüsünde kullanılan sınıflar ve örnek sayıları

No	Sınıf	Örnek Sayısı
1	Asfalt	1102
2	Çim	779
3	Çakıl	818
4	Ağaç	472
5	Boyanmış metal sac	1345
6	Toprak	5029
7	Zift	1330
8	Tuğla	3682
9	Gölge	947

## 4.2 Destek Vektör Makineleri

Destek Vektör Makineleri (Support Vector Machines, SVM) sınıflandırma, öğrenme ve regresyon problemlerinde sıkça kullanılan bir yöntemdir. Son yıllarda hiperspektral görüntülerin sınıflandırılmasında en yaygın olarak kullanılan sınıflandırıcıdır (Bazi ve Melgani 2006).

Multispektral görüntüler gibi, daha az banda sahip görüntülerde SVM yanında çok katmanlı algılayıcılar, radyal tabanlı sinir ağları da başarıyla kullanılsa da yüksek boyut sayısından dolayı hiperspektral görüntülerin sınıflandırılmasında bu yöntemler kullanışlı değildir (Demir 2010).

SVM'nin hiperspektral görüntülerdeki başarımının sebepleri, boyut lanetinden daha az etkilenmesi yani çok boyutlu veriyi daha iyi kullanabilmesidir. Bunun sebebi, dayandığı en büyük sınır prensibidir. Diğer önemli özelliği, karar fonksiyonunun sağladığı seyrek çözümlerdir (Bazi ve Melgani 2006).

SVM, orijinal olarak ikili bir sınıflandırıcı olarak geliştirilmiştir (Cortes ve Vapnik 1995). Bu sınıflandırıcı sadece iki sınıfı birbirinden ayırır. MATLAB'da bulunan SVM fonksiyonu da sadece iki sınıfı ayırmak için kullanılabilir. Bu dezavantajı aşabilmek için birkaç yöntem geliştirilmiştir. Bunlardan bazıları, çoklu sınıflandırmayı ikili sınıflandırma tabanlı yaparken, diğerleri tüm veri seti için tek bir optimizasyon problemi kullanırlar (Hsu ve Lin 2002). İkili sınıflandırma tabanlı çoklu sınıflandırma yöntemine bir örnek,  $N$  sınıflı problemi çözmek için her ikili sınıf için bir destek vektör makinesi yaratmaktır. Yani  $N$  sınıf için toplam  $N(N-1)/2$  makine olur. Buna "bire karşı bir" yöntemi denir. Bir diğer yöntem; eğitim yapılırken her sınıf için, o sınıfa 1, geri kalan sınıfların hepsine 0 vermektir. Bu işlem, tüm sınıflar için tekrarlanır. Bu yöntem "bire karşı tamamı" yöntemi denir ve toplam  $N$  tane makine olur.

Bu çalışmada, ikiden fazla sınıf kullanıldığı için MATLAB ortamında gerçekleşmiş ve farklı çoklu SVM tekniğini barındıran libSVM paketinin 3.14 sürümü kullanılmıştır (Chang ve Lin 2011).



SVM; Landsat-5 Thematic Mapper görüntüleyicisinin aldığı, arazi örtüsü içeren 6 sınıflı ve 6 bantlı Maryland veri seti üzerinden yapılan çalışmada, En Büyük Olabilirlik Sınıflandırıcı (Maximum Likelihood Classifier), sinir ağı sınıflandırıcıları (neural network classifiers) ve karar ağacı sınıflandırıcıları (decision tree classifier) ile karşılaştırılmıştır (Huang vd. 2002). Çalışma sonucunda multispektral görüntüde en iyi sonuca sinir ağı sınıflandırıcıları ve SVM ile ulaşıldığı bulunmuştur.

SVM'nin çekirdek fonksiyonu olarak; doğrusal, polinom ya da radyal tabanlı bir fonksiyon seçilebilir. Hiperspektral görüntülerin sınıflandırılmasında doğrusal çekirdekli SVM, radyal tabanlı SVM, k-komşuluk sınıflandırıcısı ve radyal tabanlı sınıflandırıcının karşılaştırılması Melagni ve Bruzzone tarafından AVIRIS Indian Pines veri seti ile 9 mahsul sınıfı üzerinden yapılmış, en iyi sonuçlara radyal taban fonksiyon çekirdekli SVM ile ulaşıldığı görülmüştür (Melgani ve Bruzzone 2002).

2004'te yapılan başka bir çalışmada ise bu kez 128 bantlı HyMAP görüntüleri kullanılmış; 128, 6, 3 ve 2 boyuta indirgeme sonrası sınıflandırıcılar karşılaştırılmıştır (Camps-Valls vd. 2004). Performansları karşılaştırılan sınıflandırıcılar doğrusal SVM, polinom SVM, radyal tabanlı SVM, çok katmanlı algılayıcılar, radyal tabanlı sinir ağlarıdır. Çalışma sonucunda SVM sınıflandırıcının en iyi sonuçları verdiği, yüksek boyutlarda polinom çekirdekli SVM'nin, düşük boyutlarda ise radyal taban fonksiyonlu SVM'nin en iyi sonuçları verdiği görülmüştür.

### **4.3 K - En Yakın Komşuluk Sınıflandırıcı**

K - En Yakın Komşuluk Sınıflandırıcı (*k*-Nearest Neighbors, KNN) sınıflandırma ve regresyon problemlerinde kullanılan, komşuluk sayısı hariç parametrik olmayan bir yöntemdir (Cover ve Hart 1967). Parametrik olmaması ve basit bir sınıflandırıcı olması sebebiyle sınıflandırma alanında referans olarak kullanılan yöntemdir.

Sınıflandırılacak veri kümesinden alınan bir nokta, eldeki eğitim verisinde bulunan tüm noktalarla yakınlık açısından karşılaştırılır. Noktalar arası yakınlığı ölçmek için en çok kullanılan metrik Öklid uzaklığı olsa da Hamming uzaklığı, Mahalanobis uzaklığı gibi farklı yöntemler kullanılabilir. Burada *k*; metriğin ölçülmesi sonucu bulunacak

komşuluk sayısını belirtmektedir. Komşuluk sayısı 1 alındığında; en basit yöntemle veri kümesindeki nokta, eğitim kümesinde en yakın olduğu nokta ile aynı sınıfa koyulur, bu En Yakın Komşuluk Sınıflandırıcı olarak adlandırılır. Komşuluk sayısı arttığında; sınıflandırma sonucu,  $k$  sayısınca çıkan sonuçlara çoğunluk oylaması yapılarak bulunur.

Testler için MATLAB yazılımında hâlihazırda bulunan fonksiyon kullanılmış,  $k$  komşuluk sayısı olarak 1 seçilmiştir. Yani, En Yakın Komşuluk sınıflandırıcı olarak kullanılmıştır. Eğitim verisi olarak, doğru şekilde karşılaştırma yapabilmek amacı ile SVM’de seçilen veri ile aynı veri kullanılmıştır.

Yapılan çalışmalar boyut indirgeme aşaması ve sınıflandırma aşaması olarak iki kısma ayrılabilir. Bölüm 3 ile kullanılan boyut indirgeme teknikleri, Bölüm 4 ile de sınıflandırma yöntemleri ve sınıflandırılacak görüntüler verilmiş olup, bir sonraki bölümde yapılan deneysel çalışmalar anlatılmaktadır.

## 5. DENEYSEL ÇALIŞMALAR

### 5.1 Giriş

Yapılan çalışmalar üç ana test çatısında toplanmıştır. Bu testler için kullanılacak boyut indirgeme yöntemleri Bölüm 3.2’de, sınıflandırıcılar Bölüm 4.2 ve 4.3’te, kullanılan veri Bölüm 4.1’de anlatılmıştır. Ayrıca SVM parametre seçimi Bölüm 5.1.1’de, SVM ve KNN eğitimi Bölüm 5.1.2’de, hedeflenen boyutun belirlenmesi Bölüm 5.1.3’de, örneklem dışı yayma fonksiyonu da Bölüm 5.1.4’de anlatılmıştır.

Sınıflandırma başarımının artırılması için kullanılan çoğunluk oylaması (majority voting) yöntemleri Bölüm 5.1.5 - 5.1.6’da verilmiştir.

Bölüm 5.2’de bir hiperspektral veri seti ve bu veri setinin bir alt kümesi için 12 farklı boyut indirgeme yönteminin SVM sınıflandırıcı ile sınıflandırılarak karşılaştırması anlatılmıştır.

Bölüm 5.3’te iki farklı hiperspektral görüntü için 12 farklı boyut indirgeme yönteminin SVM ve KNN sınıflandırıcı ile iki farklı çoğunluk oylaması yöntemi uygulandığında elde edilen sınıflandırma başarımındaki artışı verilmiştir.

#### 5.1.1 Destek vektör makineleri parametre seçimi

Testlerde SVM'nin çekirdek fonksiyonu olarak (5.1) ile gösterilen radyal tabanlı çekirdek fonksiyonu kullanılmıştır. Çekirdek fonksiyonu, sınıflar arasındaki sınırın belirlenmesinde büyük rol oynamaktadır.

$$K(x_1, x_1) = e^{-\gamma(x_1-x_2)^2} \quad (5.1)$$

SVM’de kullanılacak parametreler olan  $C$  ceza parametresi ve radyal taban fonksiyonu  $\gamma$  parametresi için her yöntem için ayrıca testler yapılmış, bu testlerde iki farklı parametre tarama (grid search) yöntemi ile denenerek en iyi sonucu veren parametreler bulunmuştur. Yapılan testlerde her boyut indirgeme yöntemi için  $C$  için  $2^{-10}$  ile  $2^{30}$

arasında  $\gamma$  için de  $2^{-60}$  ile  $2^{15}$  arasında tarama yapılarak parametre değerleri aranmıştır. Sonuç olarak her boyut indirgeme testi için 3000 sınıflandırma testi yapılmış ve sınıflandırma doğruluğu için evrensel maksimum değerler bulunmuştur.

Bulunan parametrelerin başarımını doğrulamak adına, k-en yakın komşu sınıflandırıcısı  $k=1$  için test yapılmış ve bulunan SVM sınıflandırma sonucu ile karşılaştırılmıştır.

### **5.1.2 Sınıflandırıcı eğitimi**

Testlerde sınıflandırıcı olarak Destek Vektör Makineleri ve En Yakın Komşu Sınıflandırıcı kullanılmıştır. Boyutu, Bölüm 5.1.3'de verilen hedef boyuta indirilmiş ve sınıflandırmaya hazır verilerin tüm sınıflardan rastgele seçilen 100'er veri sınıflandırıcıların eğitimi için kullanılmıştır.

Alınan 100'er veri, AVIRIS Indian Pines için tüm kümenin % 9,6'sı, ROSIS Pavia University için tüm kümenin % 5,8'idir.

Sınıflandırıcının sınıfları ezberlememesi için veri sayısı sınırlı ve belli bir limit değerinin altında olan sınıflar çalışmaya dâhil edilmemiş, literatürdeki diğer çalışmalarda kullanıldığı üzere bu limit değer 400 olarak belirlenmiştir. Çalışmaya dâhil edilen sınıflar 4.1.1 ve 4.1.2 bölümlerinde anlatılmıştır.

### **5.1.3 Boyut seçimi**

Testlerde kullanılan veriler için boyut indirgeme yöntemleri ile indirilecek hedef boyut, Etkin Boyut (Intrinsic/Effective Dimensionality) hesaplanarak yapılmıştır.

#### **5.1.3.1 Etkin boyut**

Yüksek boyutlu veriler işlenmek için çok büyük işlem gücüne ihtiyaç duyduğundan, bu verilerde belli bir yapıya sahip olduğu ve bu yapının belirli sayıda parametre tarafından belirlendiği kabul edilmektedir (Fukunaga 1990). Verinin gözlemlenebilen özelliklerini yansıtmak için gerekli minimum parametre sayısı Etkin boyut olarak tanımlanmıştır

(Van der Maaten vd. 2009).

Geleneksel olarak etkin boyutu hesaplamak için kovaryans matrisin özdeğer ve özvektörleri hesaplanıp, baskın olan özdeğerlerin sayısı alınır ancak bu teknik doğrusal dönüşüme dayandığından, doğrusal olmayan çok katmanlılardaki (manifold) başarısı yüksek değildir. Bu tip çok katmanlı veriler için Etkin Boyut genel olarak En Büyük Olabilirlik Kestirimi yöntemi ile hesaplanır (Levina ve Bickel 2004).

Çalışmada kullanılan görüntü verileri “Indian Pines” ve “Pavia University” için En Büyük Olabilirlik Kestirimi yöntemi ile etkin boyut hesaplanmış, her iki veri seti için etkin boyut sayısı 9 olarak bulunmuştur.

#### **5.1.4 Örneklem dışı yayma**

Örneklem dışı yayma (Out of sample extension), belli veri noktaları için geliştirilmiş bir dönüşümün, geliştirme sırasında kullanılmayan noktalara uygulanması ve o veri noktalarının da dönüşüme katılması olarak tanımlanabilir. Parametrik dönüşümler için çıkan parametrelerle, doğrusal dönüşümler için doğrusal eşleme uygulanarak bu yeni noktalar dönüşüme katılabilir (Van der Maaten vd. 2009).

Bu çalışmada, veriyi yüksek boyuttan düşük boyuta indirgerken tüm verideki tüm noktalar yerine her sınıftan belli sayıda nokta alınarak oluşturulan altkümeyle boyut indirgeme yapılmış ve sonrasında örneklem dışı yayma yöntemiyle verinin geri kalanına uygulanmıştır. Altküme ile yapılan testin başarımı sınıflandırma doğruluğu ve süre olarak ölçülmüştür.

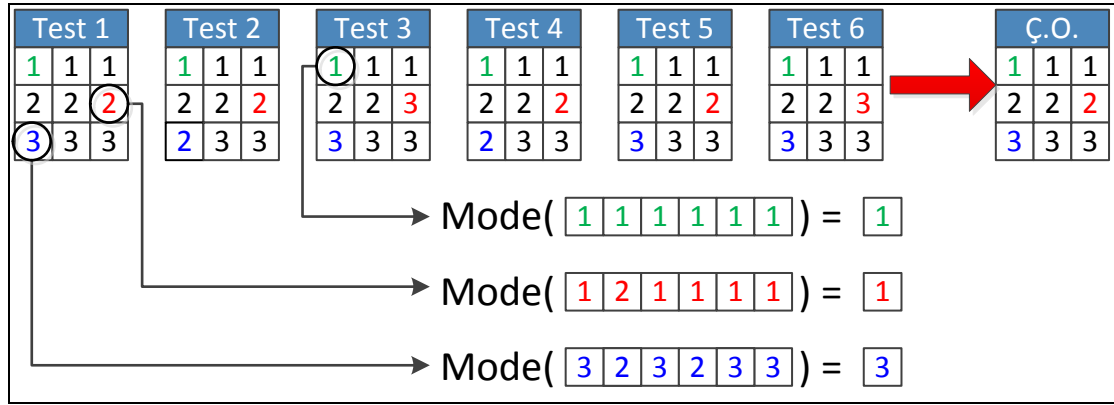
#### **5.1.5 Tekrarlı çoğunluk oylaması**

Hiperspektral görüntülerin boyut indirgenmesi sırasında, sınıflandırıcının eğitimi için sınıflardan rastgele örnek alınır. Bu örneklerin seçimi, sınıflandırma sonucuna etki eder (Foody vd. 1995). Bu sebeple yapılan testler tekrar edilmiş, her seferinde rastgele seçilen eğitim verisi sayesinde farklı sonuçlar alınmıştır. Test tekrar sayısı 9 olarak belirlenmiş, bu testler sonucu çıkan sınıflandırma doğruluğu yüzdesi, 9 testin ortalaması

alınarak verilmiştir.

Tekrarlı çoğunluk oylamasında; testlerin sonucunda çıkan sınıflandırma doğruluğu değerlerinin ortalamasını almak yerine, tekrar edilen testlerin her birinin sonucunu ayrı ayrı ele alıp, sınıflandırılmış her pikselde tekrar sayısınca çıkan görüntüdeki değerler karşılaştırılmış ve en çok bulunan değer, pikselin sınıflandırma sonucu olarak kaydedilmiştir.

Bu işlemi yapmak için MATLAB’da bulunan “mode” fonksiyonu kullanılır. Bu fonksiyonun görevi, bir dizideki en sık bulunan elemanı bulmaktır.



Şekil 5.1 Tekrarlı çoğunluk oylaması

Şekil 5.1 ile verilen örnekte bir test 6 kez tekrarlanmış ve 6 sonuç bulunmuştur. Test sonuçları farklı çıkan pikseller renklerle işaretlenmiştir. (1,1)’de yer alan piksel için çoğunluk oylaması yapıldığında, tüm veriler 1 olduğunda oylama sonucu 1 olur (yeşil pikseller). (3,1)’de yer alan piksel için 4 tane 3 sonucuna karşı, 2 tane 2 sonucu vardır ve çoğunluk oylaması sonucu 3 olur (mavi pikseller). Çoğunluk oylaması sonucu oluşan pikseller Ç.O. sonucu ile belirtilmiştir.

145x145x200’lük “Indian Pines” verisi için 9 test tekrarı ile 9 adet 145x145’lik sınıflandırma sonucu bulunur ve bu 9 sonuç, 145x145x9’luk 3 boyutlu bir matris haline getirilir.

145x145x9’luk veri ile 3.boyut üzerinden “mode” fonksiyonu çağırılır ve böylece

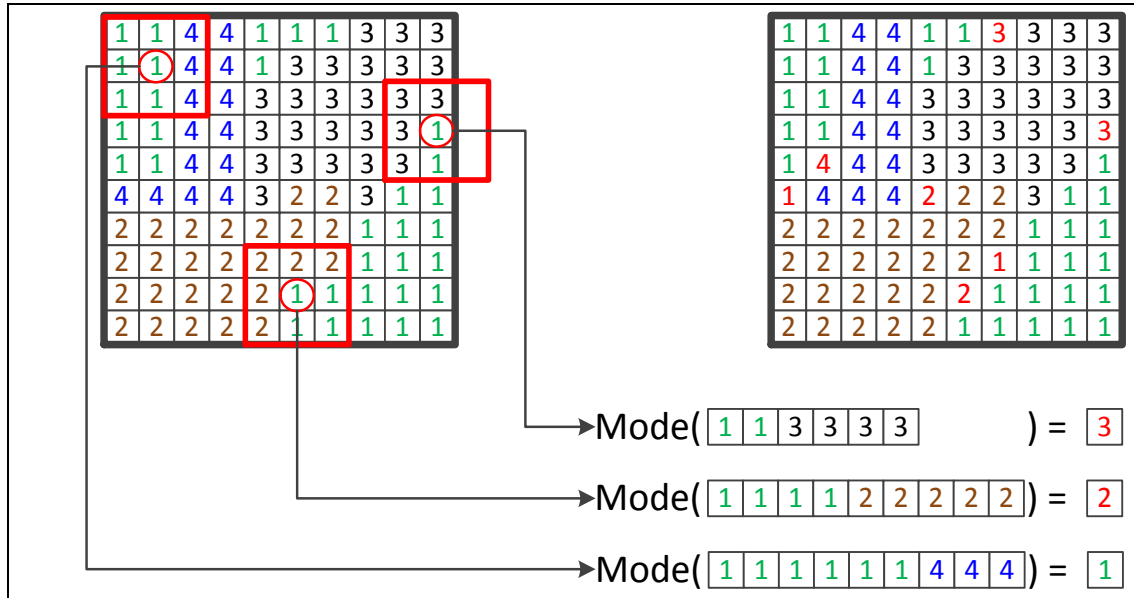
21025 pikselin her biri için en sık bulunan sınıflandırma sonucundan oluşan 145x145x1'lik yeni bir sınıflandırma sonucu ortaya çıkmış olur.

### 5.1.6 Uzaysal çoğunluk oylaması

Uzaysal çoğunluk oylaması, görüntü üzerinde yer alan bir filtreleme gibi düşünülebilir. Filtrenin uygulanması esnasında; her piksel için, ilgili pikselin komşularından yararlanır. Örneğin, 3x3'lük medyan filtrenin piksel ve komşuluğundaki 8 pikselin ortanca elemanını bulması gibi, uzaysal çoğunluk oylamasında filtre boyutu içindeki elemanların en sık yer alanı bulunur.

Bu işlemi gerçekleştirmek için, yine tekrarlı çoğunluk oylamasında yapıldığı gibi "mode" fonksiyonu kullanılır. Seçilen piksel ve komşuları bir dizi haline getirilerek mode fonksiyonu çağırılır ve fonksiyon çıktısı ilgili piksel değeri olarak kaydedilir.

Eğer en sık olan birden fazla eleman varsa, bunların küçük olanını döndürür.



Şekil 5.2 Uzaysal çoğunluk oylaması

Şekil 5.2 ile örnek bir uzaysal çoğunluk oylaması örneği verilmiştir. Bu oylama (2,2), (9,6) ve (4,10) pikselleri için ayrıntılı olarak incelenmiştir. Tüm sınıflandırma sonucuna

uygulandığında deęeri deęişen pikseller kırmızı ile gösterilmiştir.

## **5.2 Test 1: Alt Örneklemenin Hiperspektral Boyut İndirmeye Etkisi**

Bu testin amacı; boyut indirgeme esnasında tüm verinin boyutunu indirgemek yerine verinin belli bir kısmının boyutunu indirgeyip, çıkan dönüşüm modelini tüm veriye uygulayarak, bunun zaman ve sınıflandırma doğruluęu üzerindeki etkisini ölçmektir. Bu etkiyi ölçmek için; alt örnekleme (sub-sampling) yapmadan aynı işlemler yapılmış ve sonuçlar karşılaştırılmıştır. Böylelikle, operasyonel bir hiperspektral görüntü sınıflandırıcı gerçekleştirildięi zaman, daha önceden boyut indirgemesi yapılan bir alanın tekrardan görüntüsü alındığında ya da aynı sınıfları içeren benzer bir görüntü alındığında daha önce yapılan boyut indirgeme işleminde elde edilen dönüşüm fonksiyonunun yeni görüntüde kullanılması ile ilgili sonuçlar elde edilecektir.

Test için “Indian Pines” verisi kullanılmıştır. Çalışma kapsamında kullanılan 12 adet farklı boyut indirgeme yöntemi şunlardır:

- AE
- FA
- FMVU
- ISO
- LDA
- LPP
- MANC
- MCML
- NCA
- NPE
- PCA
- PPCA

Tekniklerin seçimi esnasında, alt örnekleme sonrası örnekleme dışı yayma fonksiyonu çağırıldığından, bu fonksiyonu destekleyen teknikler seçilmiştir.



### 5.2.1 Test akışı

Testin genel akışı şekil 5.3 ile verilmiştir.

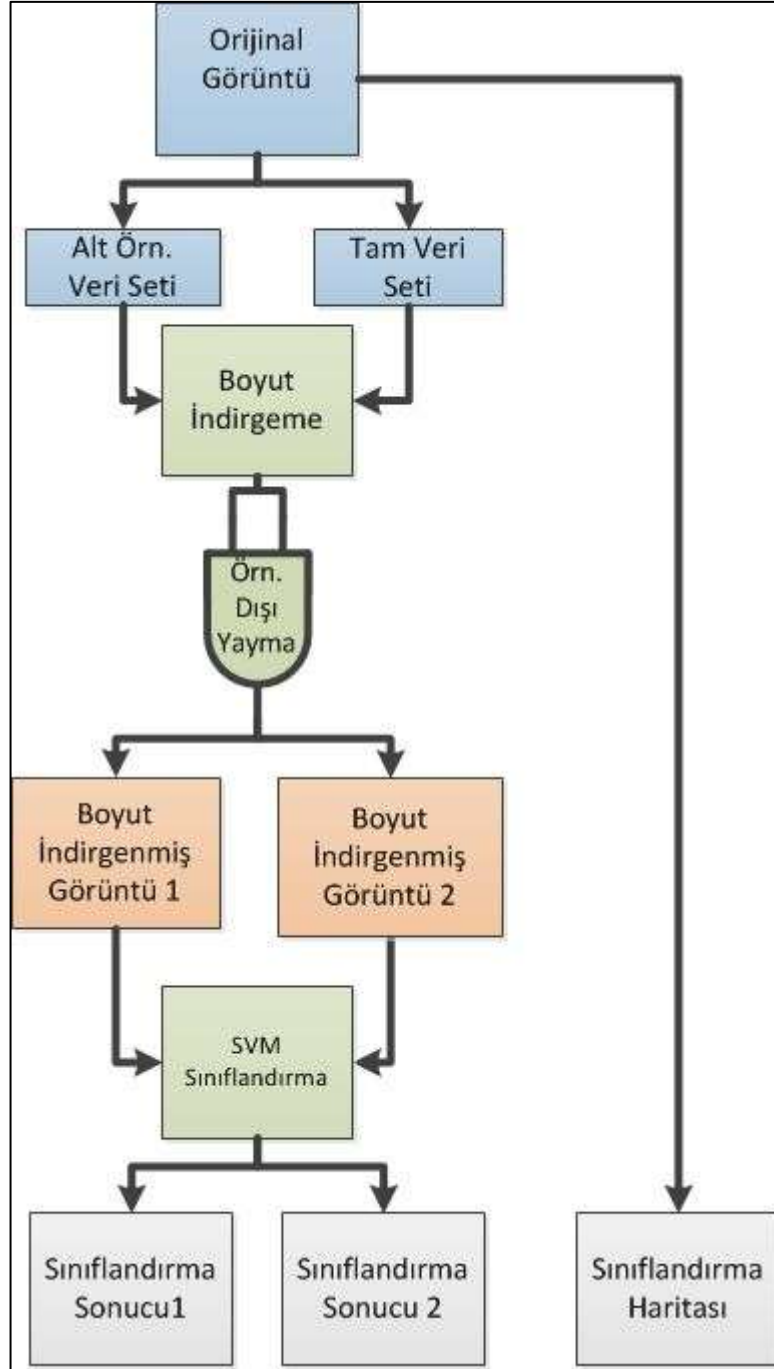
Orijinal görüntü, iki farklı veri kümesine çevrilir. Bu veri kümelerine boyut indirgeme yöntemleri uygulanır. Sonra boyutu indirgenmiş veri kümeleri örneklem dışı yayma ile boyutu indirgenmiş görüntüye dönüştürülür. Boyutu indirgenmiş görüntüler SVM ile sınıflandırılır ve sonuçlar Sınıflandırma Haritası ile karşılaştırılır.

### 5.2.2 Test adımları

Testin adımları her boyut indirgeme yöntemi için ayrı ayrı uygulanır:

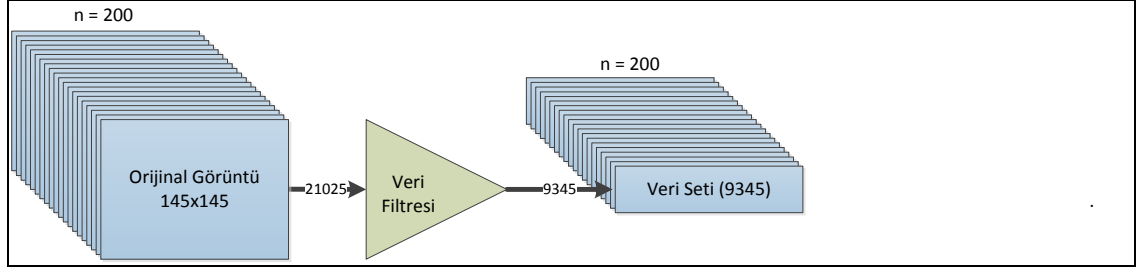
Şekillerde “ $n$ ” ile bant/boyut sayısı ifade edilmektedir. Görüntüler ve veri kümeleri, 200 boyutlu veri için mavi, 9 boyutlu veri için turuncu renkle ifade edilmiştir. Yine veri setlerinin üzerinde piksel/veri noktası sayısı rakam ile belirtilmiştir.

1. Sınıflandırma Haritasında yer alan sınıf bilgilerine göre görüntü, veri kümesi haline getirilir. Sınıflandırma Haritasında, sınıflandırılmış ve 400’den fazla örneği olan sınıfların elemanı olan veriler alınır, küçük sınıfların verileri ve Sınıflandırma Haritasında belirsiz olarak belirtilip sınıflandırılmamış bölgeler çıkartılır.



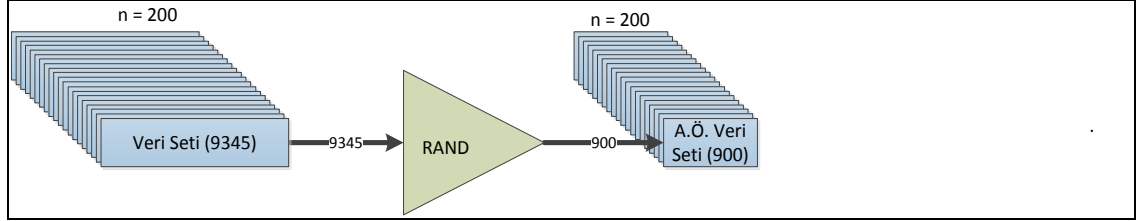
Şekil 5.3 Test akışı, Test 1

Bu işlemin sonunda, “Indian Pines” için her birinin 200 bandı olan 21025 (145x145) piksellik görüntüden, her birinin boyut sayısı 200 olan 9345 elemanlı veri kümesine geçilmiş olur. Bu veri kümesi, “Tam Veri Seti” olarak isimlendirilir. Tüm piksellerin yer aldığı bu veri seti, şekillerde “Veri Seti (9345)” olarak gösterilmiştir (Şekil 5.4).



Şekil 5.4 Veri filtresi, Test 1

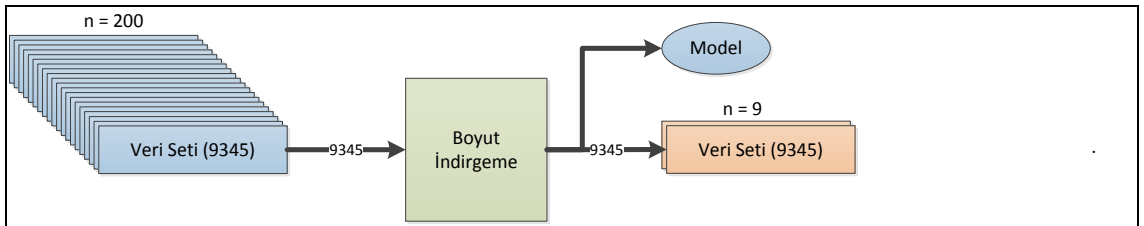
2. Veri setine alt örnekleme yapmak için, Tam Veri Seti alınarak, her sınıftan rastgele 100'er eleman seçen bir fonksiyona verilir. Bu fonksiyon, şekillerde “RAND” olarak belirtilmiştir (Şekil 5.5).



Şekil 5.5 Alt örnekleme

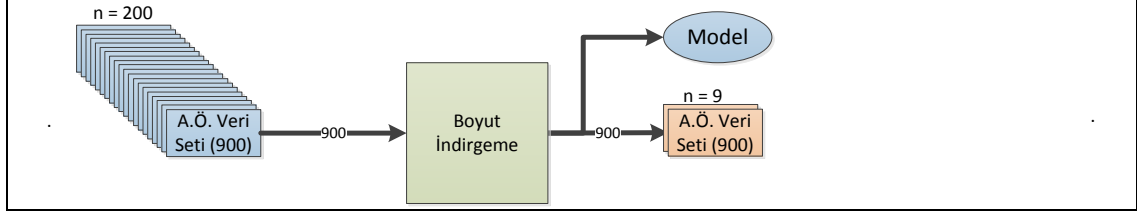
Bu işlem sonucunda 9345 piksellik 200 boyutlu Tam Veri Setinden 900 piksellik 200 boyutlu bir veri kümesine geçilmiş olur. Bu veri kümesi, “Alt Örneklenmiş Veri Seti” olarak isimlendirilir ve şekillerde “A.Ö. Veri Seti (900)” olarak gösterilmiştir.

3. Her iki veri kümesi için boyut indirgeme yöntemleri uygulanır. Yöntem sonucunda, Tam Veri Seti için şekil 5.6 ile gösterildiği gibi 9345 elemanlı ve 9 boyutlu bir veri seti ve ilgili dönüşüm modeli, Alt Örneklenmiş Veri Seti için şekil 5.7 ile belirtildiği gibi 900 elemanlı ve 9 boyutlu veri seti ve dönüşüm modeli elde edilir.



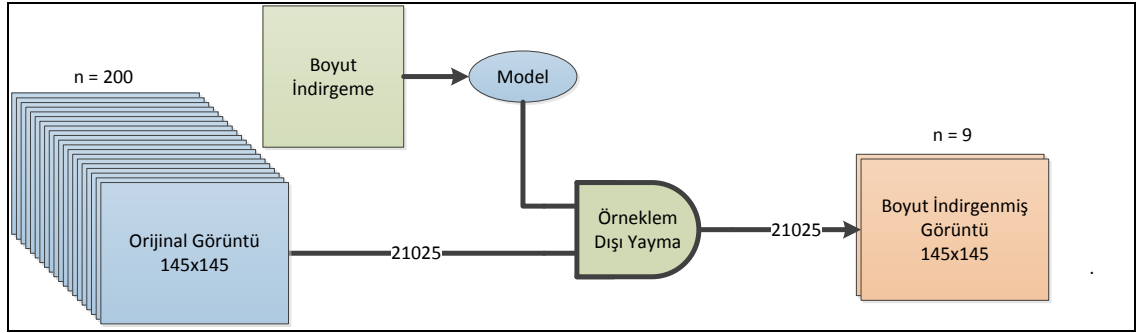
Şekil 5.6 Tam Veri Seti Boyut İndirgeme, Test 1

Örneklem dışı yayma yapılacağı için, elde edilen bu veri setleri testin geri kalanında kullanılmaz. Burada amaç, dönüşüm modellerini elde etmektir.



Şekil 5.7 Alt Örneklenmiş Veri Seti Boyut İndirgeme, Test 1

4. Örneklem Dışı Yayma için, elde edilen dönüşüm modelleri kullanılır. Bunun için Tam Veri Seti ve Alt Örneklenmiş Veri Seti için elde edilen dönüşüm modelleri, orijinal görüntüye uygulanır (Şekil 5.8).

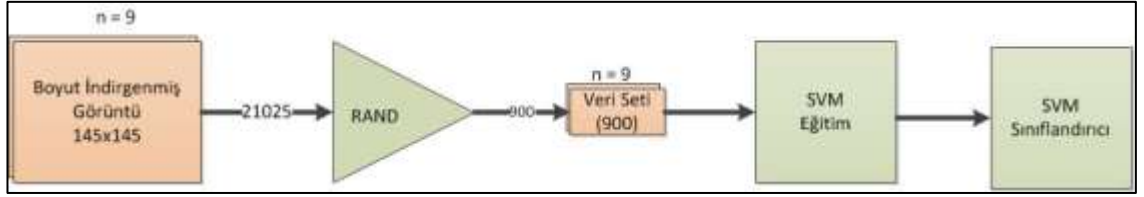


Şekil 5.8 Örneklem dışı yayma

İşlem sonucunda 145x145 piksellik 200 boyutlu görüntüden 145x145 piksellik 9 boyutlu görüntü elde edilmiş olur. Böylelikle testin boyut indirgeme kısmı tamamlanır ve bir boyut indirgenme yöntemi için iki farklı boyutu indirgenmiş görüntü elde edilir.

Boyut indirgeme için 3. adımda geçen süre ve örneklem dışı yayma için 4. adımda geçen süre toplanarak boyut indirgeme süresi olarak not edilir.

5. Sınıflandırma aşamasında, boyutu indirgenmiş görüntülerden, 9 sınıftan rastgele 100'er örnek seçilir ve bu örnekler ile SVM eğitilerek sınıflandırıcı elde edilir (Şekil 5.9).



Şekil 5.9 SVM Eğitim, Test 1

9 sınıftan 100'er örnek, Sınıflandırma Haritasında karşılaştırma yapılacak 9345 örneğin %9,6'sını oluşturmaktadır.

6. Eğitilen sınıflandırıcı ile boyutu indirgenmiş görüntü sınıflandırılıp koordinatlara oturtularak sınıflandırma sonucu oluşturulur (Şekil 5.10).



Şekil 5.10 Sınıflandırma ve karşılaştırma, Test 1

7. Sınıflandırma sonucu ile Sınıflandırma Haritası karşılaştırılır ve sınıflandırma doğruluğu hesaplanır.

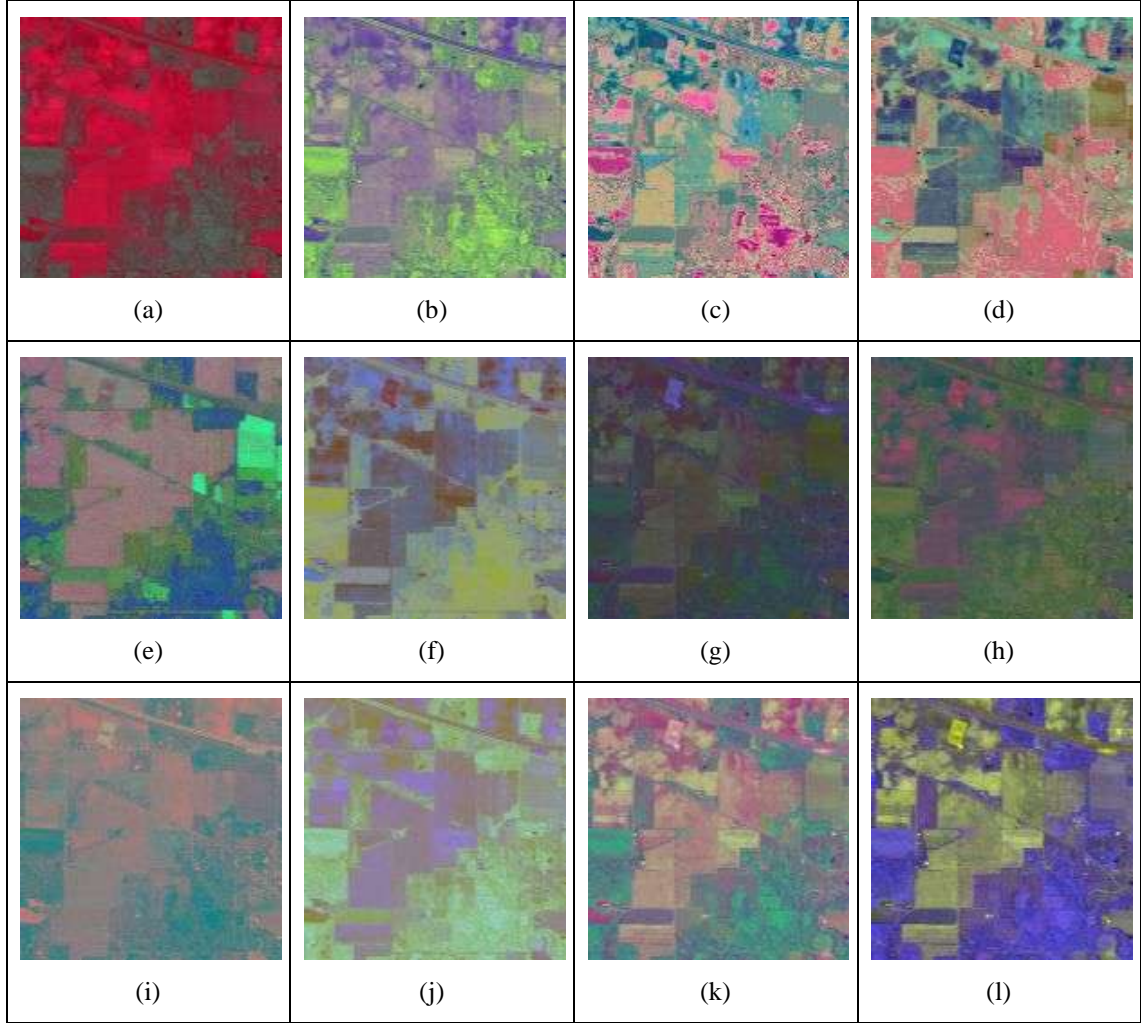
Bu işlemler 12 adet boyut indirgeme yöntemi için, her birinde 2 farklı veri seti olmak üzere yapılır ve test tamamlanır.

### 5.2.3 Test sonuçları

Yapılan testlerin tümü; hem alt örnekleme yapılırken hem de sınıflandırıcı eğitilirken yapılan rastgele örnek seçiminin sonuca etkisini en aza indirmek amacıyla 9 kez tekrar edilmiştir. Sınıflandırma doğruluğu bulunurken, yapılan bu 9 testin ortalaması alınmıştır. Yani 12 yöntem ve 2 veri seti için 9 tekrardan oluşan toplam 216 test yapılmıştır.

Denenen yöntemlerde, 200 boyuttan 9 boyuta indirgeme sonrası, ortaya çıkan

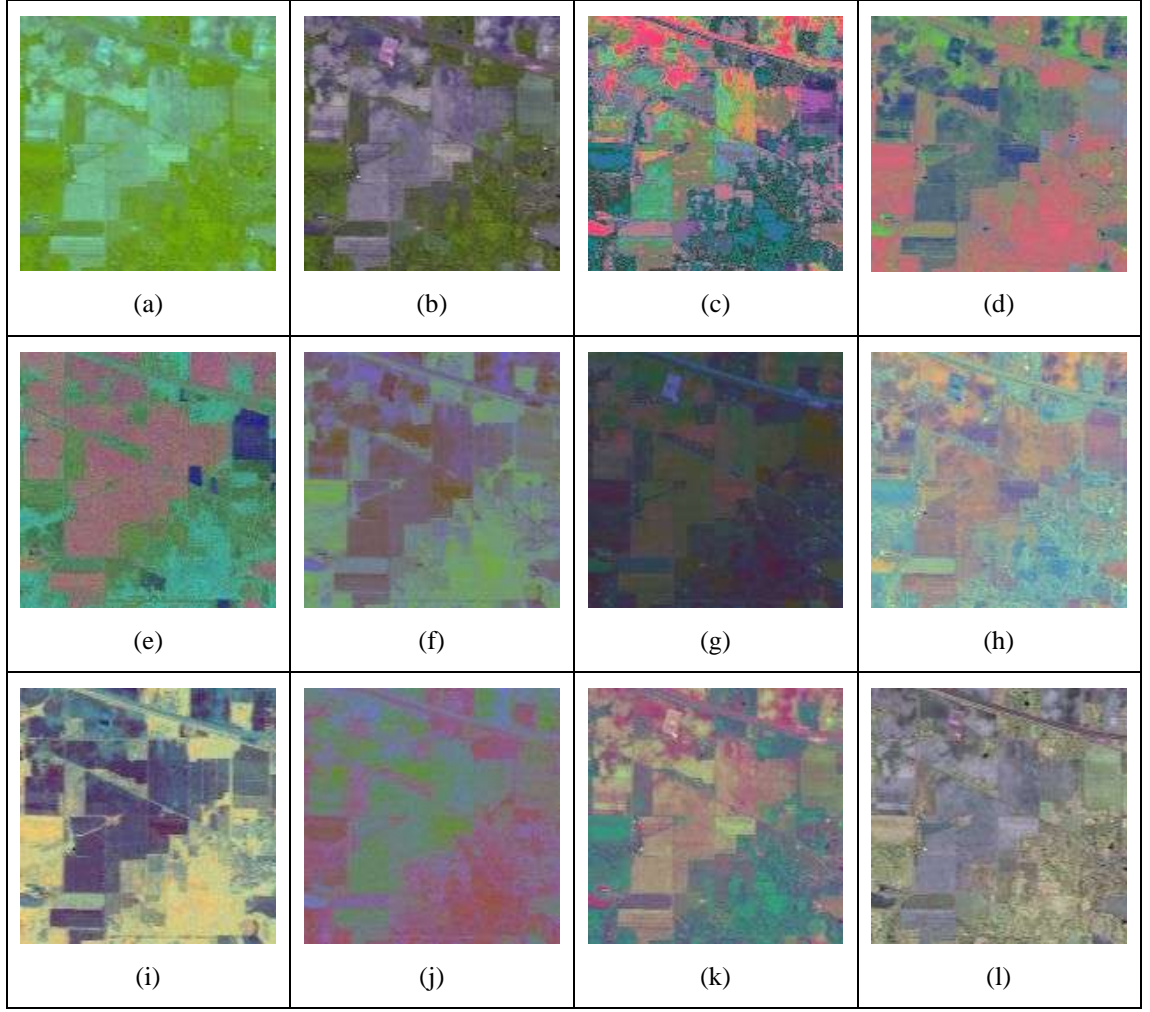
bantlardan ilk 3 tanesi alınarak sırasıyla kırmızı, yeşil ve mavi renge atanıp RGB olarak gösterimi tam veri seti için şekil 5.11 ile alt örneklenmiş veri seti için şekil 5.12 ile verilmiştir. Harf sırasına göre şekillerde aşağıda belirtilen boyut indirgeme yöntemleri gösterilmektedir



Şekil 5.11 Tam veri seti, ilk 3 bandının RGB gösterimi

a. AE, b. FA, c. FMVU, d. ISO, e. LDA, f. LPP, g. MANC, h. MCML, i. NCA, j. NPE, k. PCA, l. PPCA

Şekil 5.11-5.12 bilgi amacıyla verilmiş olup, boyut indirgemenin sınıflandırmaya etkisi ile ilgili kesin bir bilgi vermeyebilir çünkü 9 boyutun sadece 3 tanesinin gösterimi yapılabilmektedir. Şekillerden de anlaşılacağı gibi, tam veri seti ve alt örneklenmiş veri seti ile yapılan boyut indirgeme ve örneklem dışı yayma sonrası oluşan görüntüler birbirinden farklıdır.



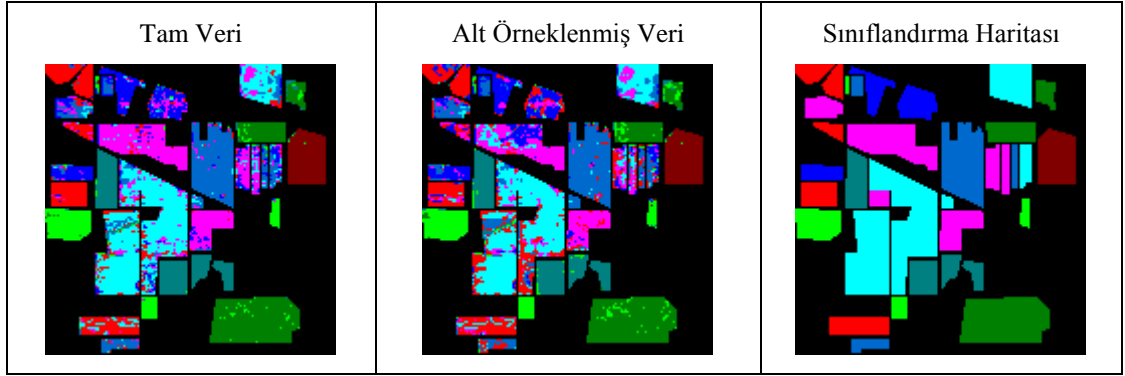
Şekil 5.12 Alt örneklenmiş veri seti, ilk 3 bandının RGB gösterimi

a. AE, b. FA, c. FMVU, d. ISO, e. LDA, f. LPP, g. MANC, h. MCML, i. NCA, j. NPE, k. PCA, l. PPCA

Testler sonrası her yöntem için tam veri seti sınıflandırma sonucu , alt örneklenmiş veri seti sınıflandırma sonucu ve sınıflandırma haritası harf sırasına göre şu şekillerde gösterilmiştir:

- Şekil 5.13 → AE
- Şekil 5.14 → FA
- Şekil 5.15 → FMVU
- Şekil 5.16 → ISO
- Şekil 5.17 → LDA
- Şekil 5.18 → LPP
- Şekil 5.19 → MANC

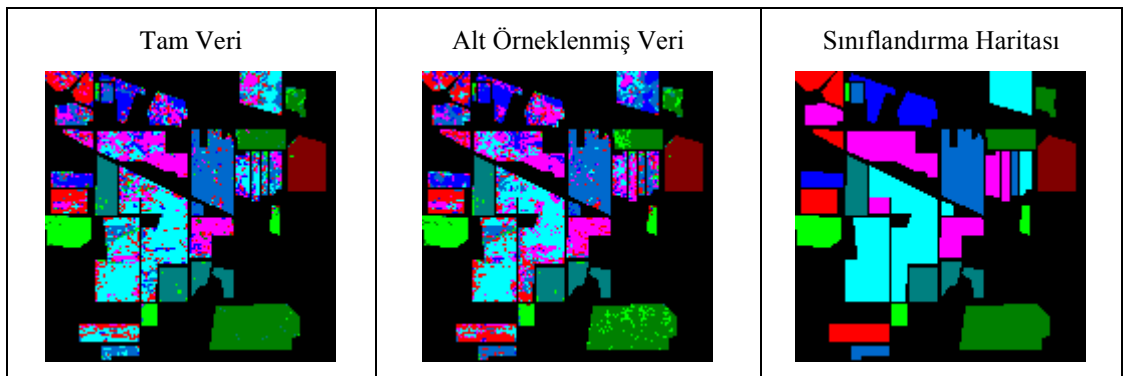
- Şekil 5.20 → MCML
- Şekil 5.21 → NCA
- Şekil 5.22 → NPE
- Şekil 5.23 → PCA
- Şekil 5.24 → PPCA



Şekil 5.13 AE sınıflandırma sonuçları

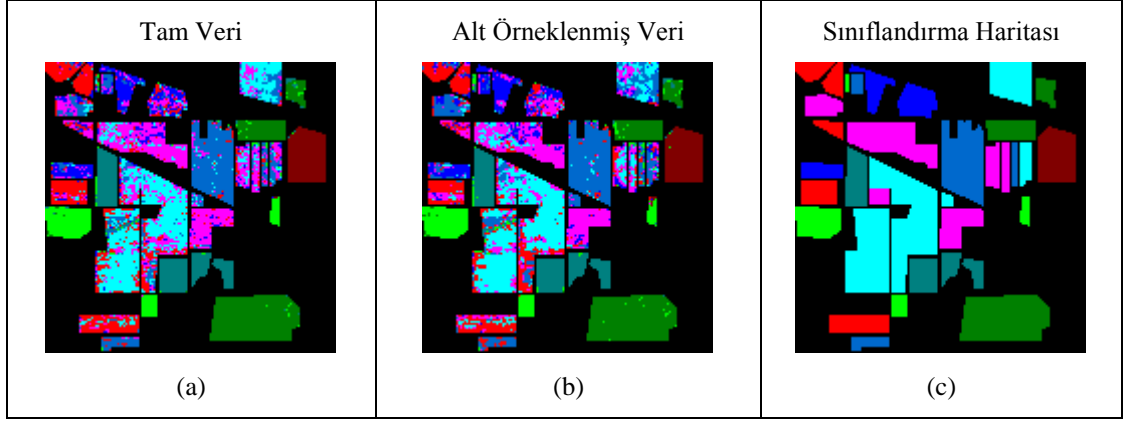


Şekil 5.14 FA sınıflandırma sonuçları

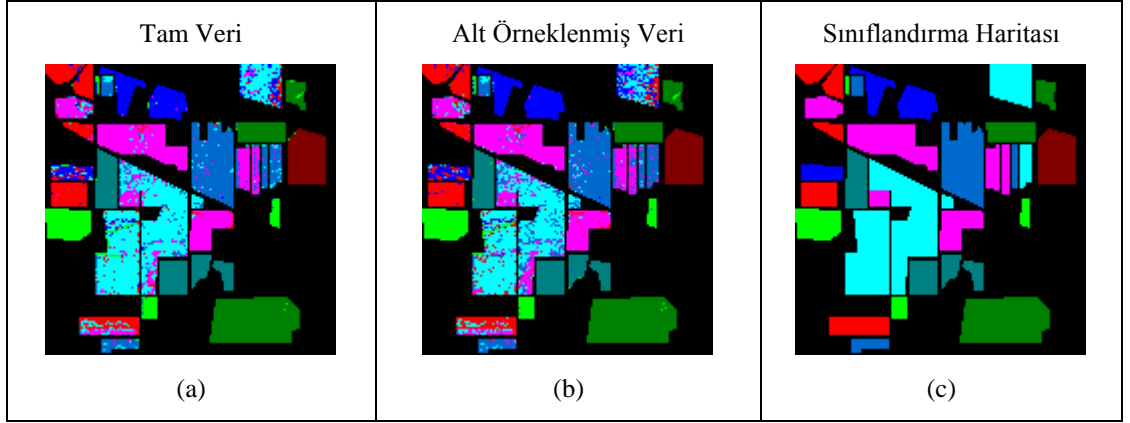


Şekil 5.15 FMVU sınıflandırma sonuçları

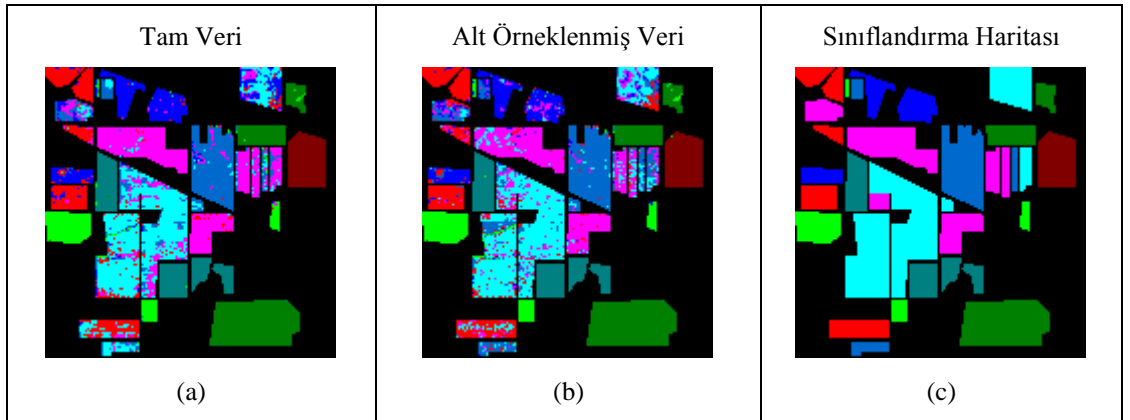




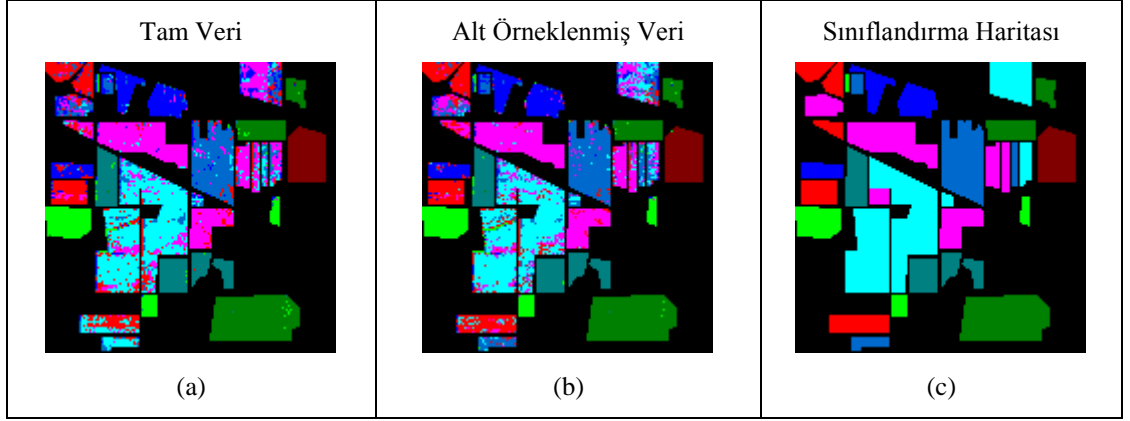
Şekil 5.16 ISO sınıflandırma sonuçları



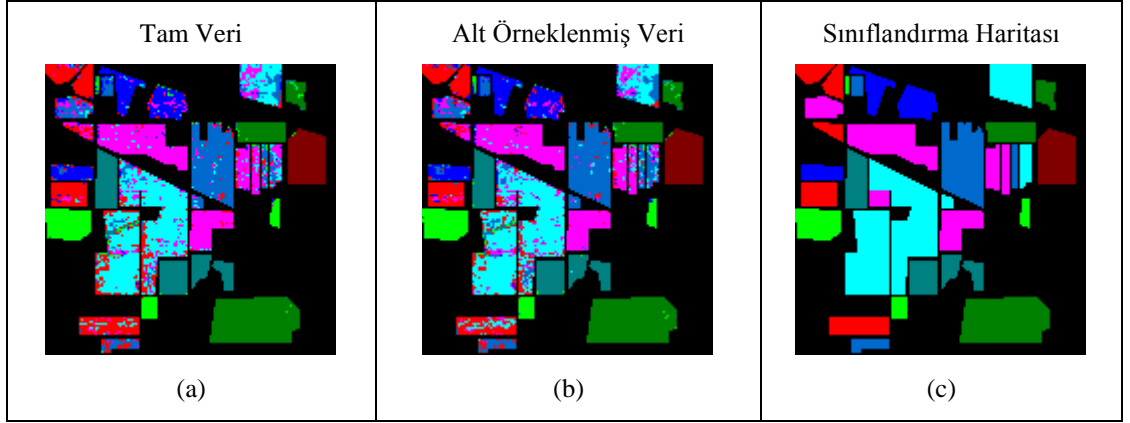
Şekil 5.17 LDA sınıflandırma sonuçları



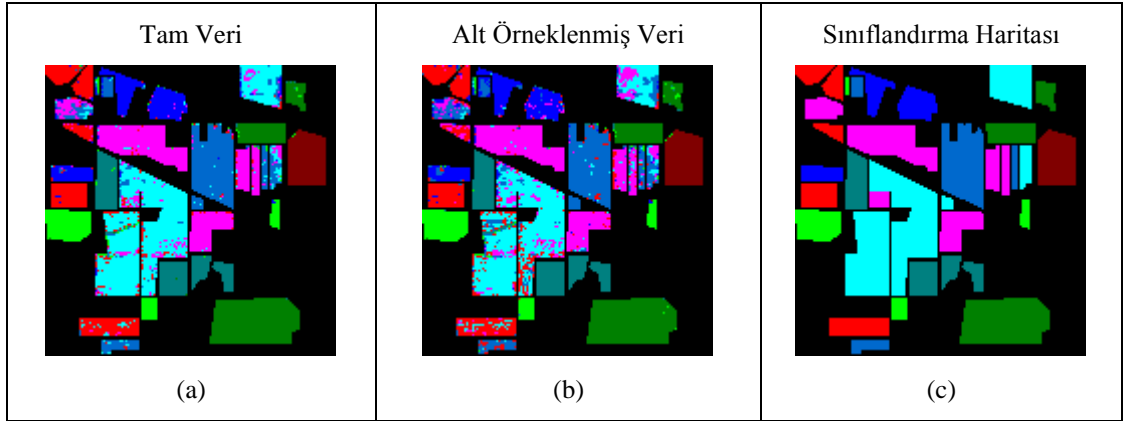
Şekil 5.18 LPP sınıflandırma sonuçları



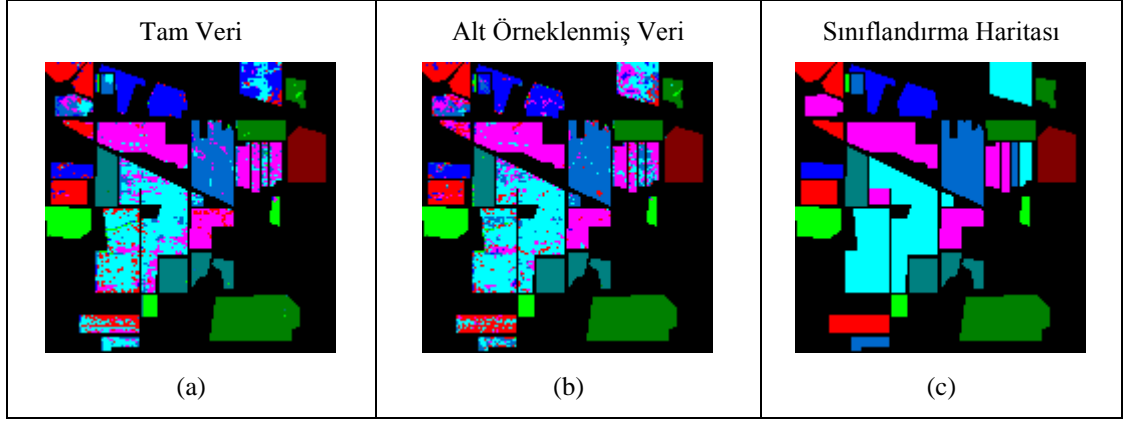
Şekil 5.19 MANC sınıflandırma sonuçları



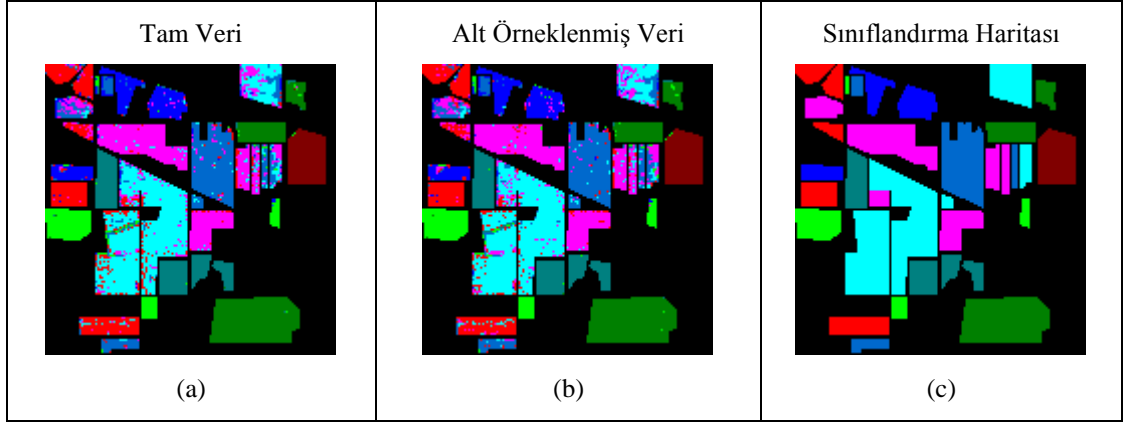
Şekil 5.20 MCML sınıflandırma sonuçları



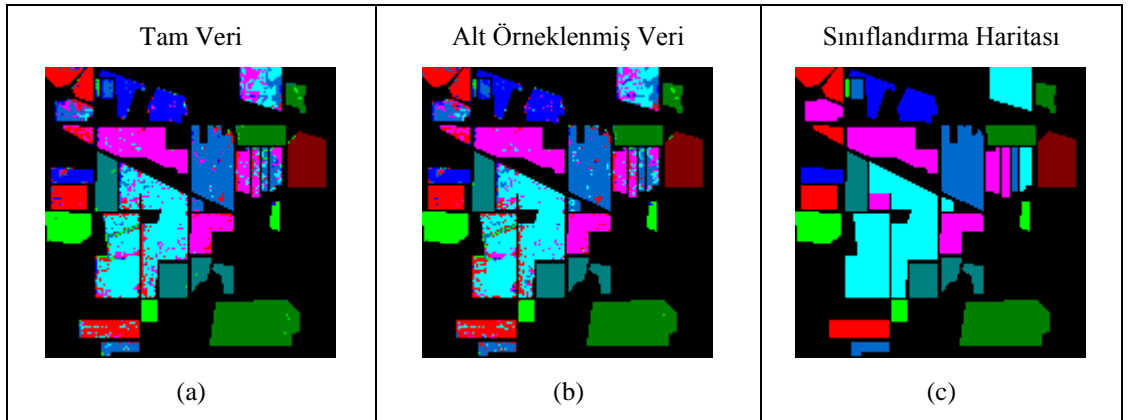
Şekil 5.21 NCA sınıflandırma sonuçları



Şekil 5.22 NPE sınıflandırma sonuçları



Şekil 5.23 PCA sınıflandırma sonuçları



Şekil 5.24 PPCA sınıflandırma sonuçları

Testte yer alan yöntemlerin doğrusal olma durumu ve eğitim şekline göre sıralanarak alt örneklenmiş ve tam veri ile verdikleri sınıflandırma doğruluğu ve test süresi çizelge 5.1

ile verilmiştir.

Çizelge 5.1 Sınıflandırma sonucu, Test 1

Doğrusal Olma Durumu	Eğitim Şekli	Yöntem Adı	Sınıflandırma Doğruluğu a		Süre b	
			A.Ö.	Tam	A.Ö.	Tam
Doğrusal	Öğreticisiz	FA	84,76	85,31	1,55	13,10
		LPP	74,27	81,77	0,11	11,68
		NPE	74,09	80,30	0,15	11,50
		PCA	82,96	82,44	0,07	0,28
		PPCA	82,45	81,98	4,36	43,58
	Öğreticili	LDA	77,26	86,90	0,04	0,09
		MCML	73,87	74,24	130,03	18796,24
		NCA	77,12	84,59	323,23	38052,02
Doğrusal olmayan	Öğreticisiz	AE	70,57	76,31	5,50	57,28
		FMVU	60,85	65,45	107,55	576,00
		ISO	66,39	69,99	34,29	1396,43
		MANC	77,44	75,54	1,41	14,83

a %, b s

Testlerde sınıflandırma performansı yönünden doğrusal yöntemler daha iyi sonuçlar vermiştir. Süre yönünden de Mahalanobis uzaklığını kullanmayan doğrusal yöntemlerin daha iyi sonuçlar verdiği görülmüştür (Kozal vd. 2013a).

Alt örneklenmiş veri kümesi ile yapılan testlerde, beklendiği gibi boyut indirgeme süreleri yönünden tüm yöntemler için daha iyi sonuçlar vermiştir. Sınıflandırma doğruluğu olarak bakıldığında, doğrusal eğiticiisiz yöntemler olan LPP, NPE gibi yöntemler için %5 civarı bir düşüş olmuş, PCA, FA, PPCA ise fazla etkilenmemiştir. Öğreticili yöntemler ise, eğitim verisi sayısı %10 civarına düşmesinden dolayı performans düşüşü göstermişlerdir. Doğrusal olmayan yöntemler ise MANC hariç doğrusal yöntemlerden daha fazla etkilenmişlerdir. MANC ise PCA ve PPCA ile birlikte, diğer yöntemlerin tamamından farklı bir şekilde daha az veri ile performansını artırmıştır (Kozal vd. 2013b).

### 5.3 Test 2: Çoğunluk Oylamasıyla Sınıflandırma Başarımının Artırımı

Bu çalışmanın amacı; hiperspektral görüntülerin boyut indirgenerek Destek Vektör Makineleri tarafından sınıflandırılması ile elde edilen sonuçların başarımını çoğunluk oylaması yöntemlerini kullanarak arttırmak, çıkan sonuçları En Yakın Komşuluk sınıflandırıcı ile karşılaştırmaktır.

Bu testte, 5.2 bölümünde anlatılan testten farklı olarak:

- Sınıflandırmada kullanılmayacak bölgeler için, boyut indirgeme sonrası görüntü örneklem dışı yayma fonksiyonu ile tekrar oluşturulmamış, bunun yerine sınıflandırma haritası da veri kümesi haline getirilmiş ve görüntüler değil veri kümeleri karşılaştırılmıştır. Bu sayede örneklem dışı yayma fonksiyonu olmayan boyut indirgeme yöntemleri de kullanılabilmiştir.
- Sınıflandırıcı olarak SVM yanında KNN de kullanılmıştır.
- Sınıflandırıcılar eğitilirken, rastgele seçim adımında seçilen 100'er örnekte aynı örnek veri noktasının alınması engellenmiştir.
- Testler hem "Indian Pines" hem de "Pavia University" görüntüleri için yapılmıştır.

İlk çalışmada, alt örnekleme aşamasının gerektirdiği bir zorunluluk olarak, sadece örneklem dışı yayma yapılabilen boyut indirgeme yöntemleri kullanılabilirken, bu çalışmada böyle bir kısıt bulunmadığından birçok yöntem denenmiş, sınıflandırma başarımı yüksek yöntemler seçilerek kullanılmıştır. Bu, daha fazla doğrusal olmayan yöntemin kullanımına olanak sağlamıştır. Ayrıca süre olarak kötü sonuçlar vermiş olan yöntemler elenmiştir.

Çalışma kapsamında kullanılan 12 adet farklı boyut indirgeme yöntemi şunlardır:

- AE
- DIF

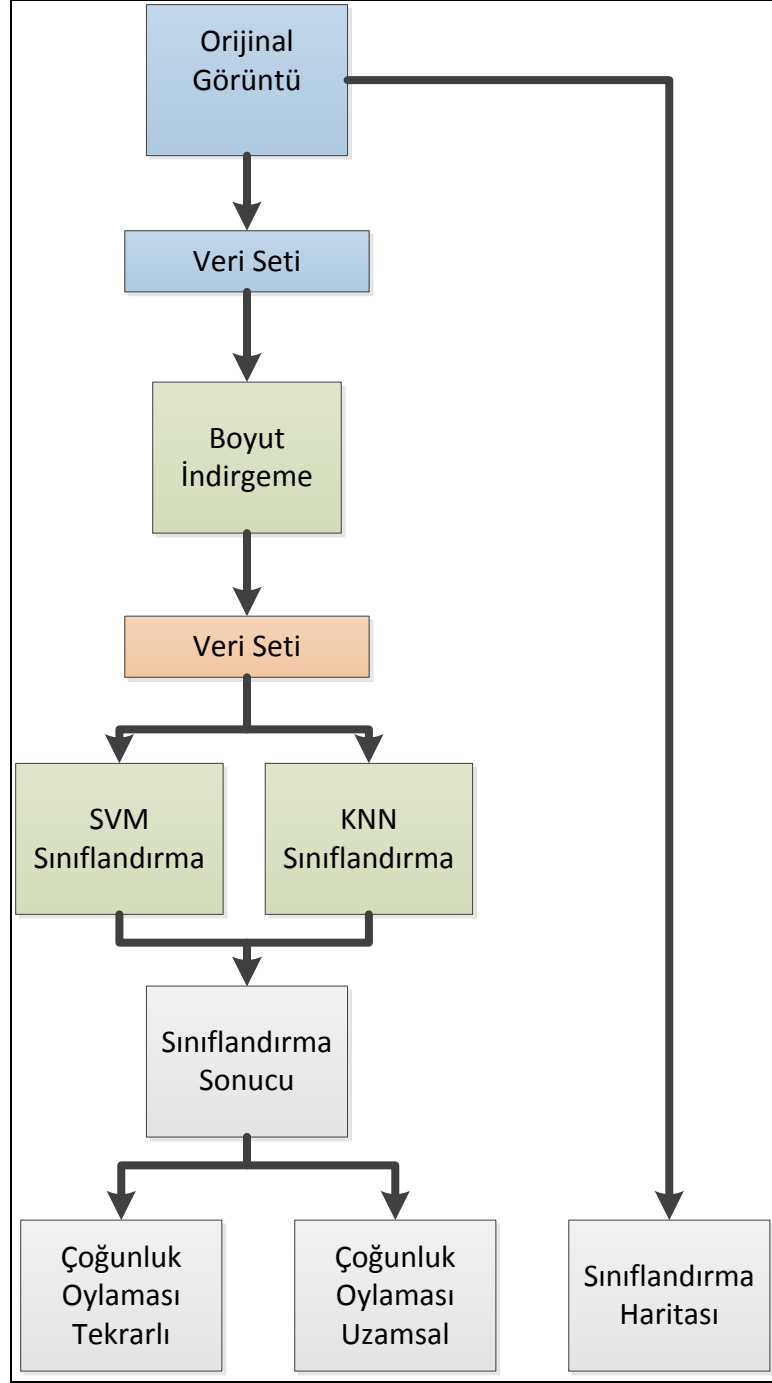
- FA
- FMVU
- ISO
- LDA
- LPP
- MANC
- NPE
- PCA
- PPCA
- SPE

Boyut indirgeme sonrası örneklem dışı yayma fonksiyonu kullanmama kararı alınsa da LDA için ilgili fonksiyon kullanılmıştır çünkü kullanılan diğer yöntemlerden farklı olarak LDA öğreticili bir yöntemdir. Diğer yöntemlerde tüm veri girdi olarak verilirken, LDA'de hem tüm veri, hem de tüm sınıf bilgileri veri olarak girmesi gerekir. Bunun yerine LDA için sınıflandırıcı eğitimi sırasında kullanılan verinin aynısı kullanılmıştır. Böylece ek bir veri girişine gerek kalmamıştır.

### **5.3.1 Test akışı**

Testin genel akışı şekil 5.25'te verilmiştir.

Orijinal görüntüden Sınıflandırma Haritası ele alınarak veri seti çıkartılır. Bu veri seti, boyut indirgeme yöntemi ile görüntünün etkin boyutuna indirgenir. Boyutu indirgenen veri seti, KNN ve SVM sınıflandırıcı ile sınıflandırılır. Elde edilen sınıflandırma sonucu, tekrarlı çoğunluk oylaması ve uzaysal çoğunluk oylaması aşamalarından geçirilerek bunların sınıflandırma doğruluğuna etkisi gözlemlenir.



Şekil 5.25 Test akışı, Test 2

### 5.3.2 Test adımları

Testin detaylı adımları ise ilerleyen maddelerde verilmiştir. Her boyut indirgeme yöntemi için adımlar ayrı ayrı uygulanır:

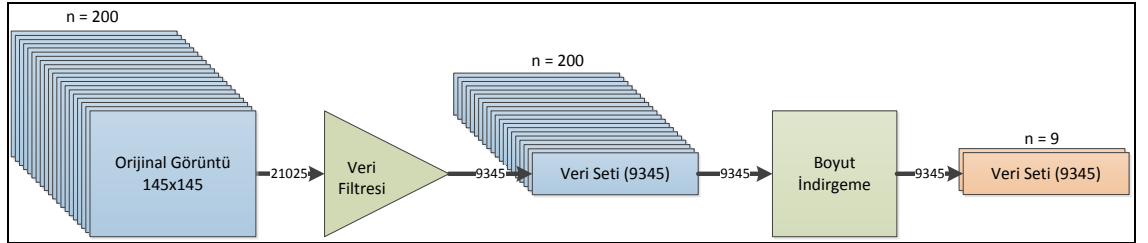
Örnekleme dışı yama yapılmayacağından, verilerin 145x145'lik görüntüler haline getirilmesine gerek yoktur. Bu yüzden hem Sınıflandırma Haritası, hem de test akışında veriler hep veri kümesi halinde kullanılmıştır.

1. Sınıflandırma haritasına veri filtresi uygulanır ve 400'den fazla örneği olan sınıfların elemanı olan veriler alınır, küçük sınıfların verileri ile belirsiz olarak belirtilip sınıflandırılmamış bölgeler çıkartılır. Sınıflandırma haritası bir veri kümesi haline getirilir (Şekil 5.26).



Şekil 5.26 Sınıflandırma Haritasına veri filtresi uygulaması

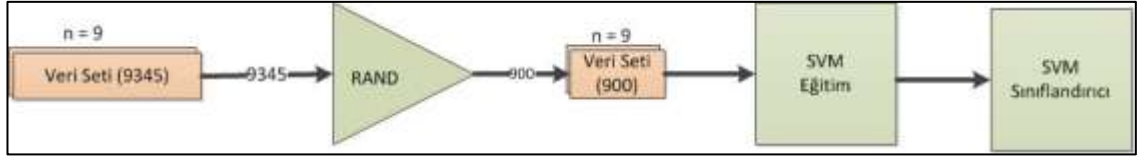
2. Orijinal görüntüye de veri filtresi uygulanır ve görüntü veri seti haline getirilir. Ardından bu veri setine boyut indirgeme yöntemleri uygulanır. Yöntem sonucunda, şekil 5.27 ile gösterildiği gibi 9345 elemanlı ve 9 boyutlu bir veri seti elde edilir.



Şekil 5.27 Boyut İndirgeme, Test 2

3. Sınıflandırma aşamasında, boyutu indirgenmiş görüntülerden 9 sınıftan rastgele 100'er örnek seçilir ve bu örnekler ile Destek Vektörü Makineleri eğitilerek sınıflandırıcı elde edilir (Şekil 5.28).

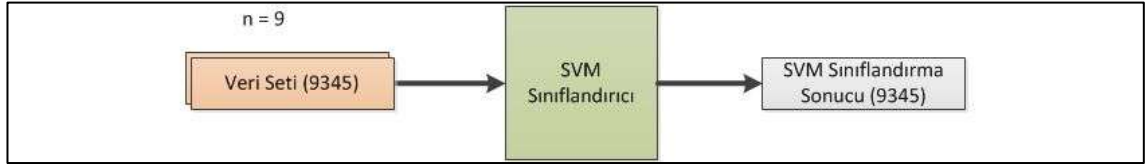




Şekil 5.28 SVM Eğitim, Test 2

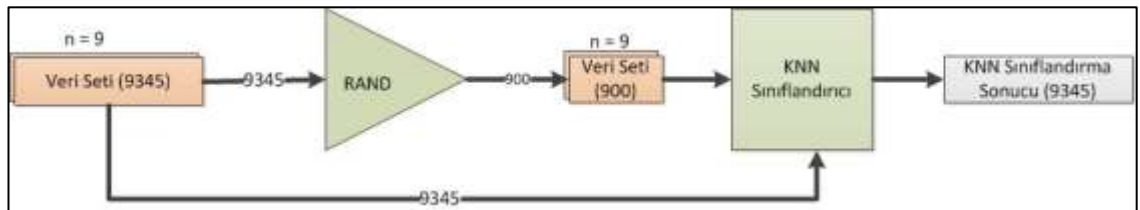
9 sınıftan 100 örnek, Sınıflandırma Haritasında Indian Pines için karşılaştırma yapılacak 9345 örneğin %9,6'sını oluşturmaktadır. Pavia University için bu oran % 5,8'dir.

4. Eğitilen sınıflandırıcı ile veri kümesi sınıflandırılır, sınıflandırma sonucu elde edilir (Şekil 5.29).



Şekil 5.29 Sınıflandırma, Test 2

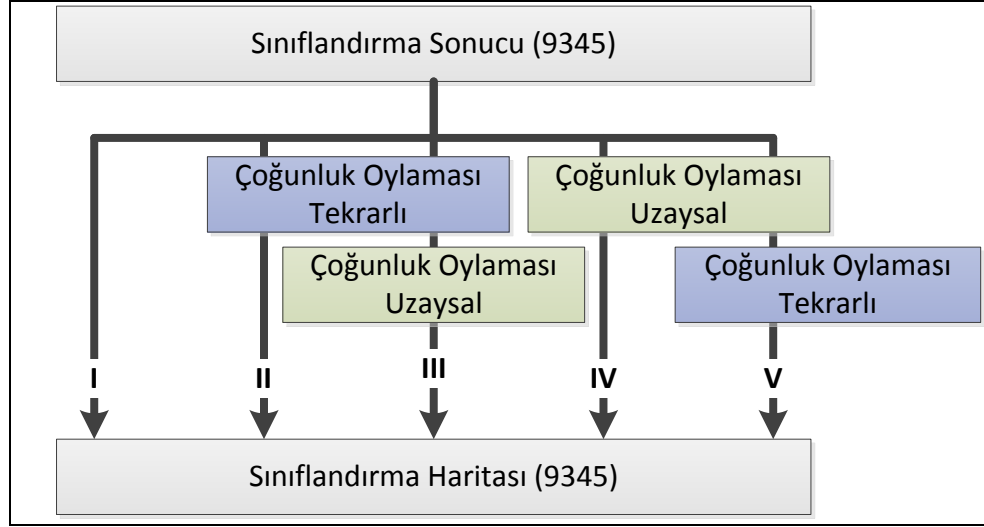
5. KNN için sınıflandırma aşamasında, boyutu indirgenmiş görüntülerden 9 sınıftan rastgele 100'er örnek seçilir ve bu örnekler KNN sınıflandırıcıya eğitim verisi olarak verilir. Ardından tüm veri sınıflandırıcıya sokularak eğitim kümesindeki en yakın komşu bulunur ve veri sınıflandırılır (Şekil 5.30).



Şekil 5.30 KNN Eğitim, Test 2

3, 4 ve 5. adımlar 9 kez tekrar edilerek, RAND işleminin etkisi azaltılır. Ayrıca tekrarlı çoğunluk oylaması için birden fazla sonuca ihtiyaç duyulduğundan, bu şekilde gerekli veri seti elde edilir.

5. Testin ayırt edici kısmı olan çoğunluk oylaması bu adımda yapılır. Çoğunluk oylaması kısmı birden fazla aşamada yapılır. Bu aşamalar şekil 5.31 ile gösterilmiştir.



Şekil 5.31 Çoğunluk oylaması uygulamaları

Her boyut indirgeme yöntemine ait sınıflandırma sonucu için:

- I. Tekrar edilen 9 testin sınıflandırma sonuçları alınır.
- II. 9 sınıflandırma sonucuna tekrarlı çoğunluk oylaması yapılır. (ilk aşama)
- III. II’de elde edilen sonuca uzaysal çoğunluk oylamasına yapılır. (ikinci aşama)
- IV. 9 sınıflandırma sonucuna uzaysal çoğunluk oylaması yapılır. (ilk aşama)
- V. IV’de elde edilen sonuca tekrarlı çoğunluk oylaması yapılır (ikinci aşama)

I. adımında sınıflandırma sonucu elde edildikten sonra II. ve IV. adımda yapılan işlemler direkt olarak I. adımdan sonra olduğu için ilk aşama işlemler olarak adlandırılır. İlk aşama işlemlere III. ve V. adımda yapılan ikini işlemler de ikinci aşama işlemler olarak adlandırılır.

Bu testler sonunda; tekrarlı çoğunluk oylamasının yapısı gereği, elde edilen test tekrar sayısınınca sonuç tek bir sonuca dönüşür. Dolayısı ile tekrarlı çoğunluk oylaması yapılan adımlarda (II. III. V.) tek bir sınıflandırma sonucu elde edilirken, diğer adımlarda (I. IV.) test tekrarı kadar yani 9 adet sonuç elde edilir. Böylelikle her boyut indirgeme yöntemi için 21 tane sınıflandırma sonucu elde edilmiş olur.

6. Sınıflandırma sonuçları ile Sınıflandırma Haritası karşılaştırılır ve sınıflandırma doğruluğu hesaplanır.

Bu işlemler 12 adet boyut indirgeme yöntemi için tekrarlanır. Boyut indirgeme kısmında herhangi bir rastgele seçim fonksiyonu olmadığı için, bu adımlar birer kez yapılırken, sınıflandırma kısmı 9 kez tekrar edilir.

### 5.3.3 Test sonuçları

Yapılan testler; tekrarlı çoğunluk oylaması için veri oluşturmak amacıyla 9 kez tekrar edilmiştir. Tekrarlı çoğunluk oylaması içermeyen adımlarda sınıflandırma doğruluğu yapılan 9 testin ortalaması olarak verilmiştir. 12 yöntem ve 2 farklı hiperspektral görüntü için 9 tekrardan oluşan toplam 216 test yapılmıştır.

- I. Tekrar edilen 9 testin ortalamasının sınıflandırma sonuçlarını
- II. 9 sınıflandırma sonucuna yapılan ilk aşama tekrarlı çoğunluk oylamasını
- III. II. adımda elde edilen sonuca yapılan ikinci aşama uzaysal çoğunluk oylamasını
- IV. 9 sınıflandırma sonucuna yapılan ilk aşama uzaysal çoğunluk oylamasını
- V. IV. adımda elde edilen sonuca yapılan ikinci aşama tekrarlı çoğunluk oylamasını temsil eder.

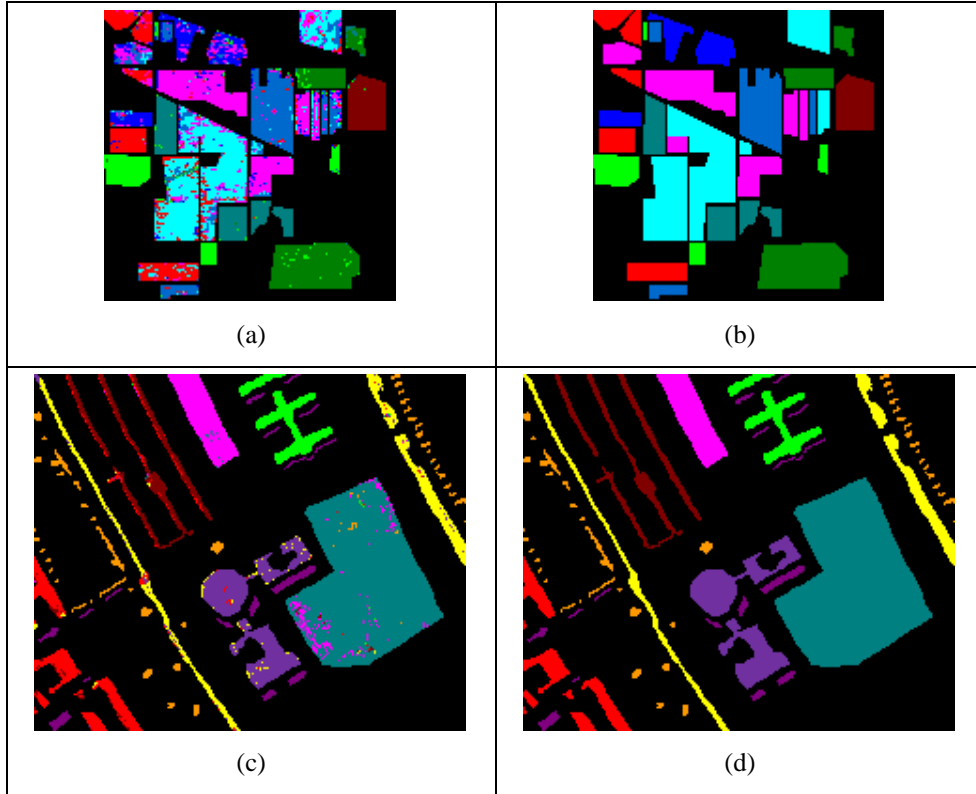
Testler sonrası her yöntem için SVM ile sınıflandırıldıktan sonra tekrarlı çoğunluk oylaması uygulanmış sınıflandırma sonuçları (II) ve Sınıflandırma Haritası

- (a) Indian Pines tekrarlı çoğunluk oylaması sınıflandırma sonucu
- (b) Indian Pines Sınıflandırma Haritası
- (c) Pavia University tekrarlı çoğunluk oylaması sınıflandırma sonucu
- (d) Pavia University Sınıflandırma Haritası

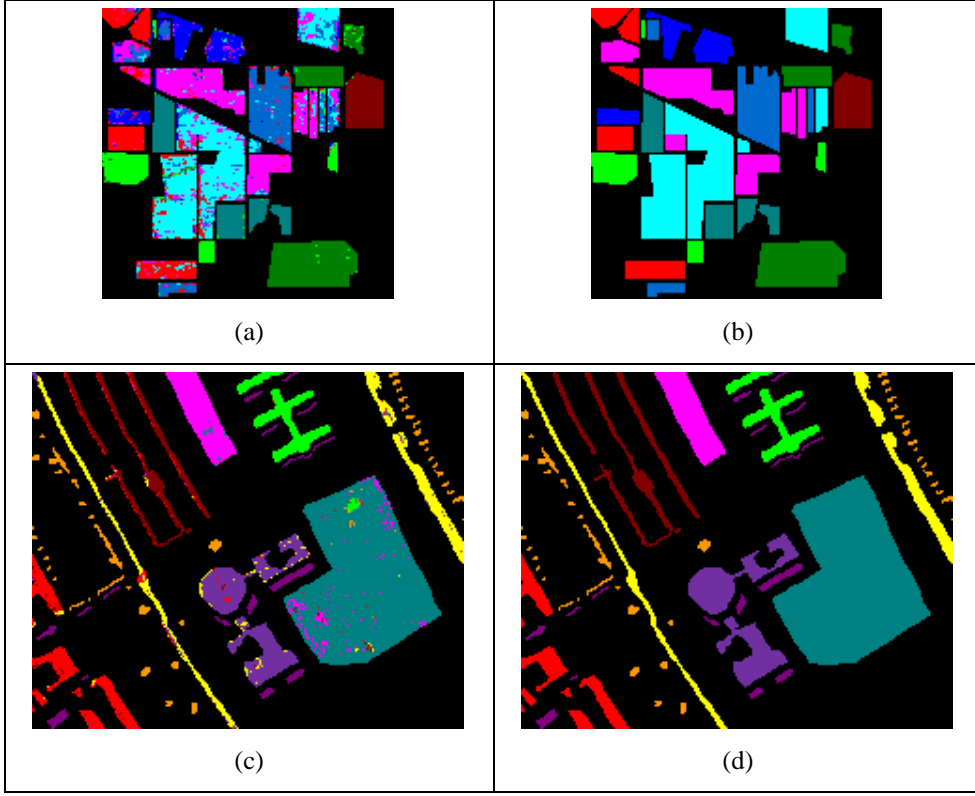
olmak üzere alfabetik sıraya göre şu şekillerde gösterilmiştir:

- Şekil 5.32 → AE
- Şekil 5.33 → DIF
- Şekil 5.34 → FA
- Şekil 5.35 → FMVU

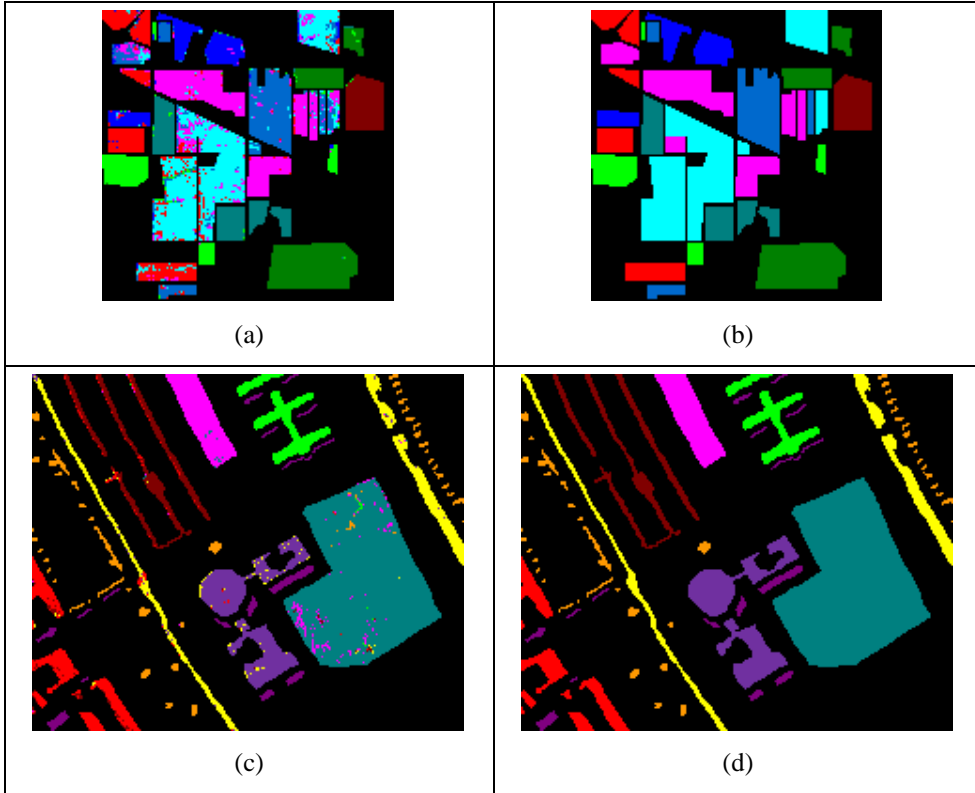
- Şekil 5.36 → ISO
- Şekil 5.37 → LDA
- Şekil 5.38 → LPP
- Şekil 5.39 → MANC
- Şekil 5.40 → NPE
- Şekil 5.41 → PCA
- Şekil 5.42 → PPCA
- Şekil 5.43 → SPE



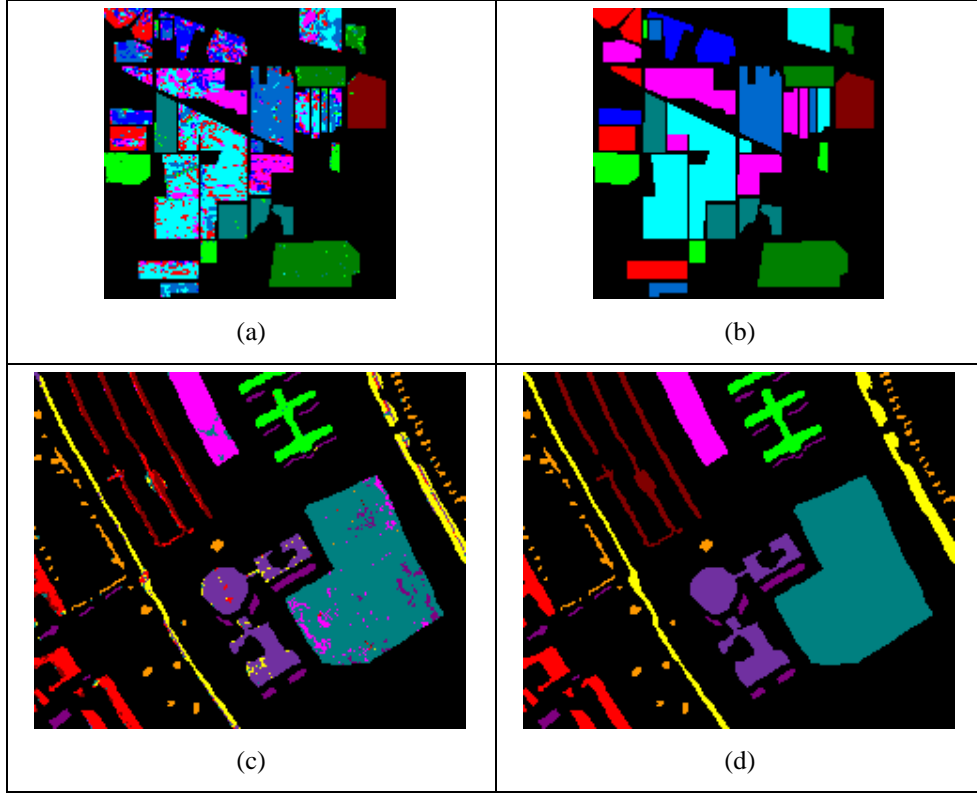
Şekil 5.32 AE sınıflandırma sonuçları



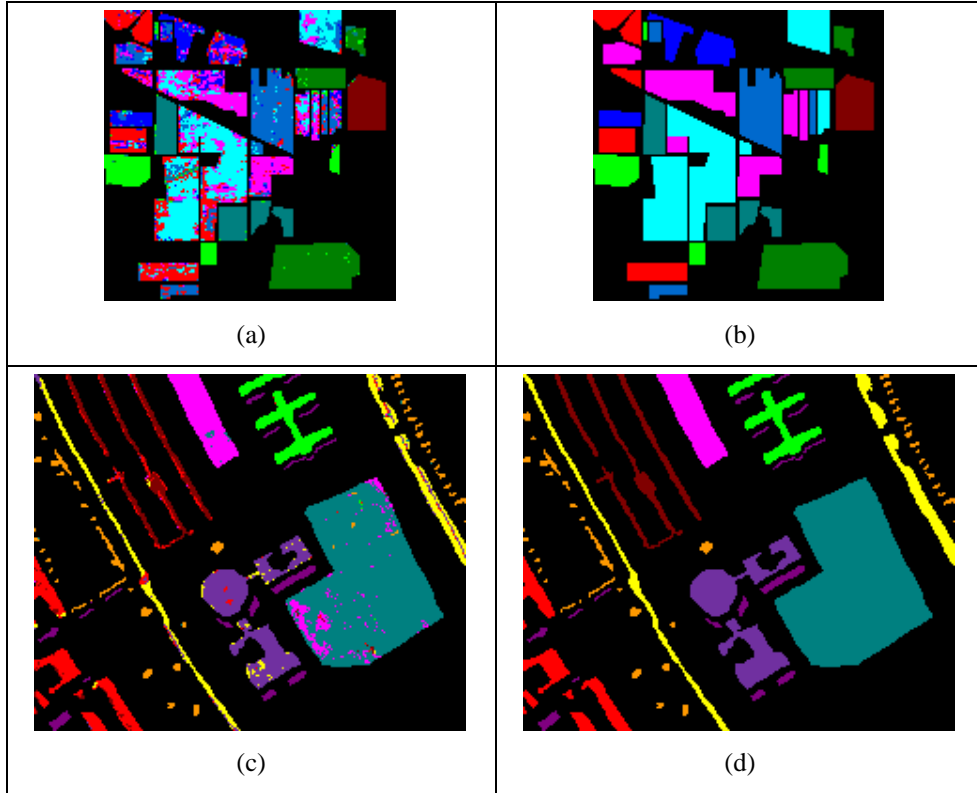
Şekil 5.33 DIF sınıflandırma sonuçları



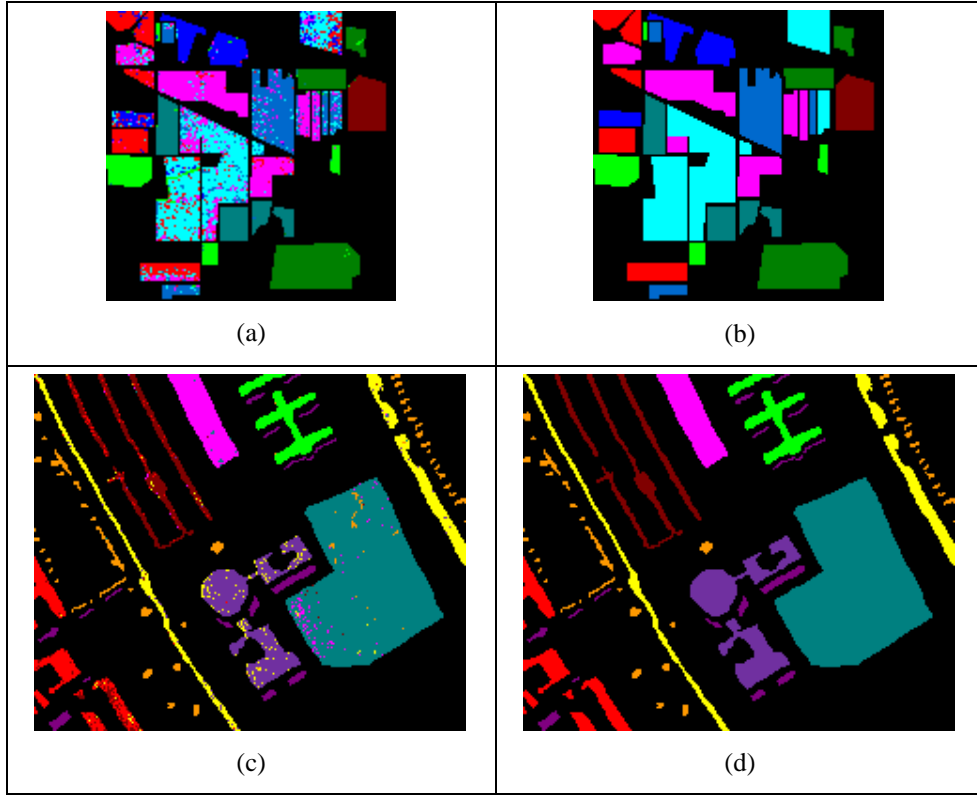
Şekil 5.34 FA sınıflandırma sonuçları



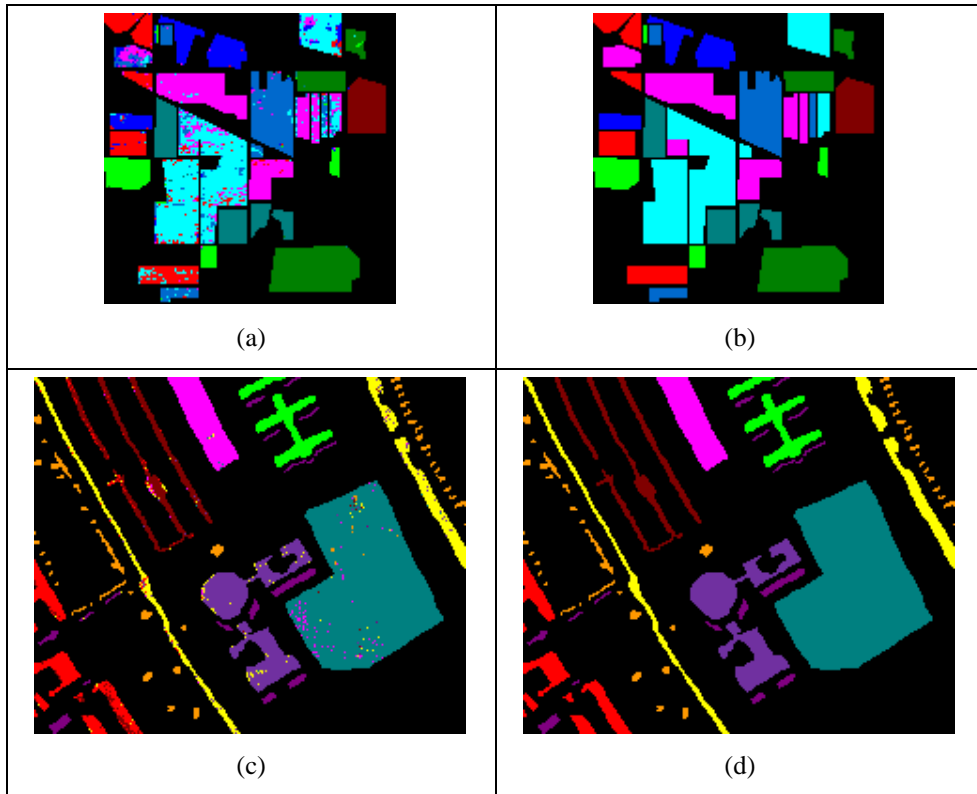
Şekil 5.35 FMVU sınıflandırma sonuçları



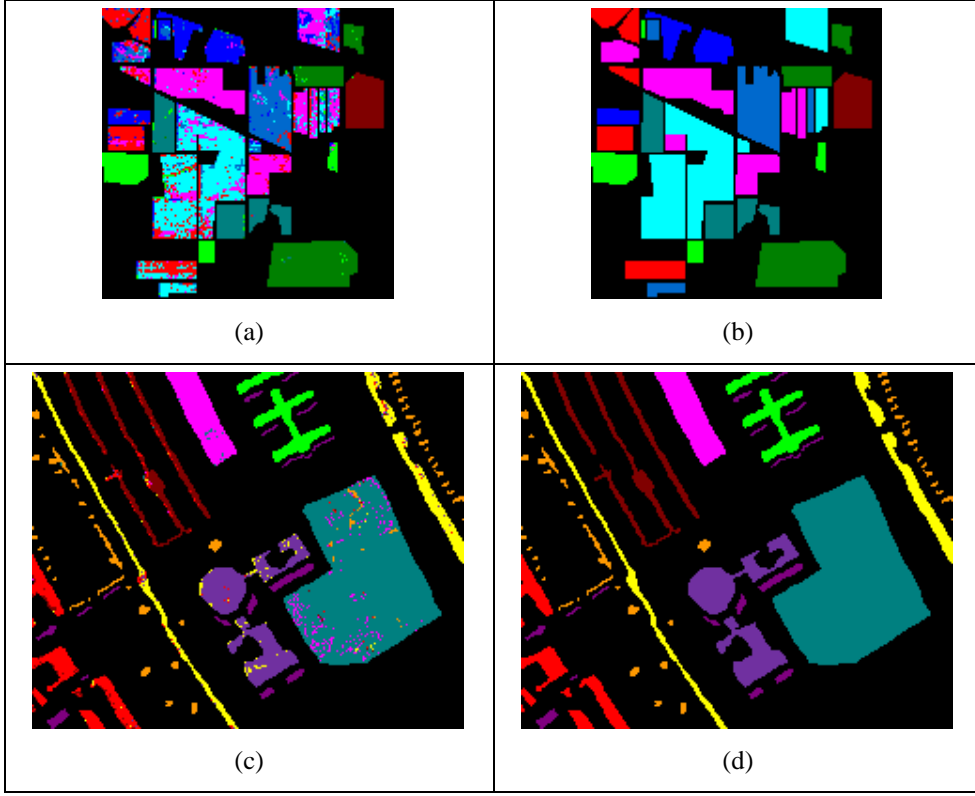
Şekil 5.36 ISO sınıflandırma sonuçları



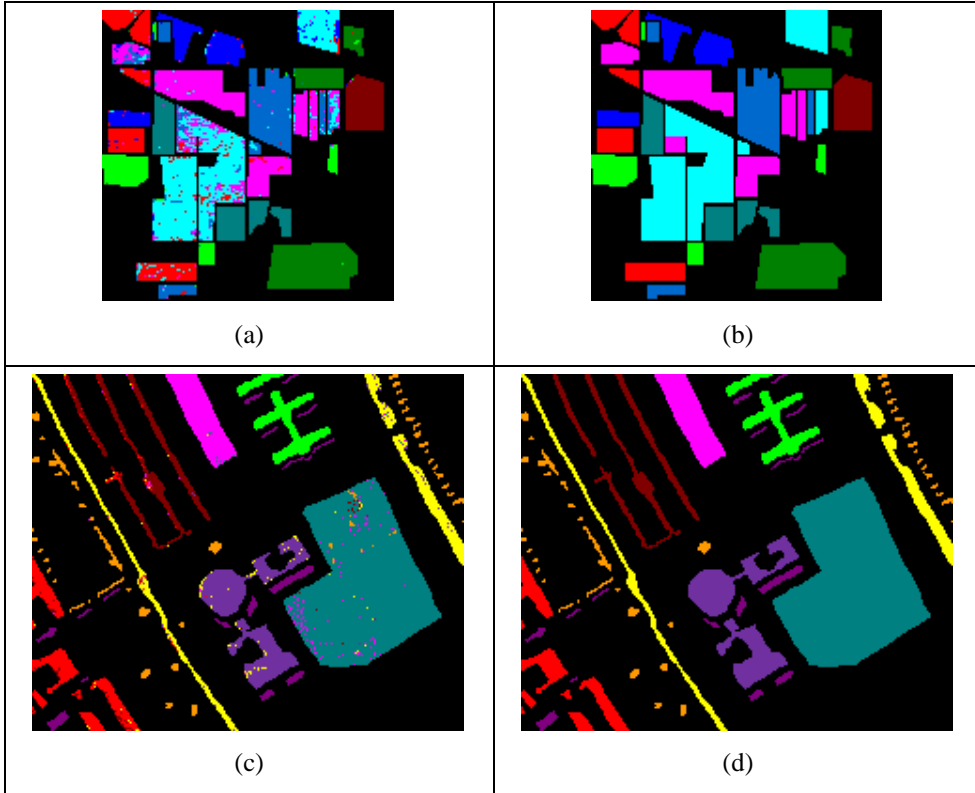
Şekil 5.37 LDA sınıflandırma sonuçları



Şekil 5.38 LPP sınıflandırma sonuçları

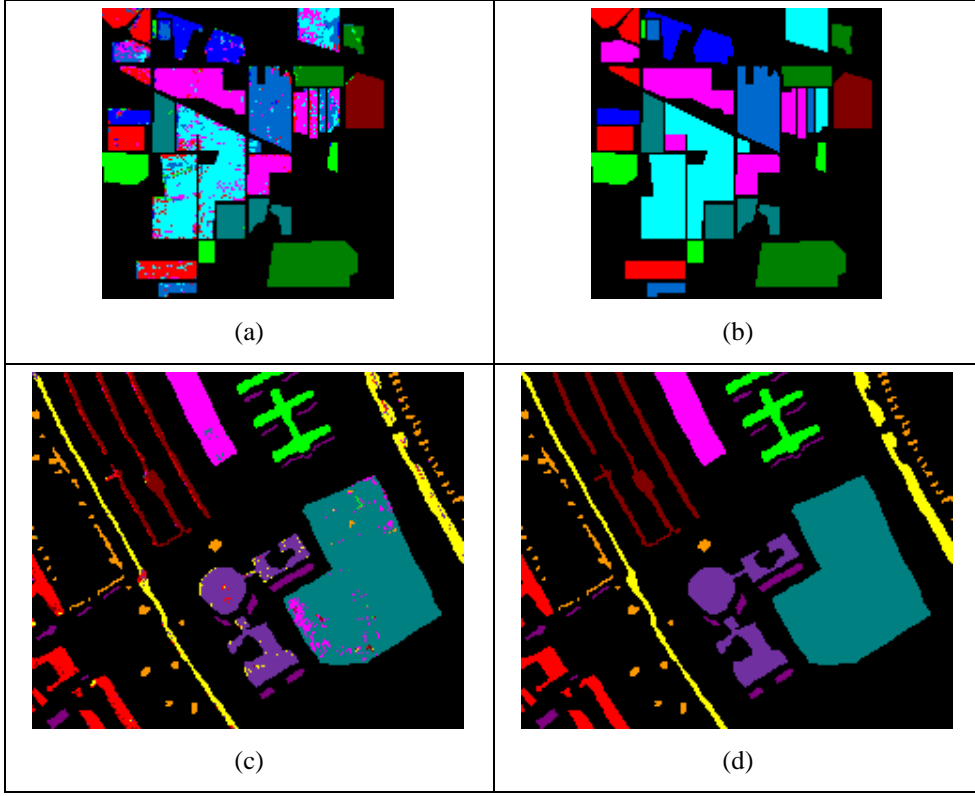


Şekil 5.39 MANC sınıflandırma sonuçları

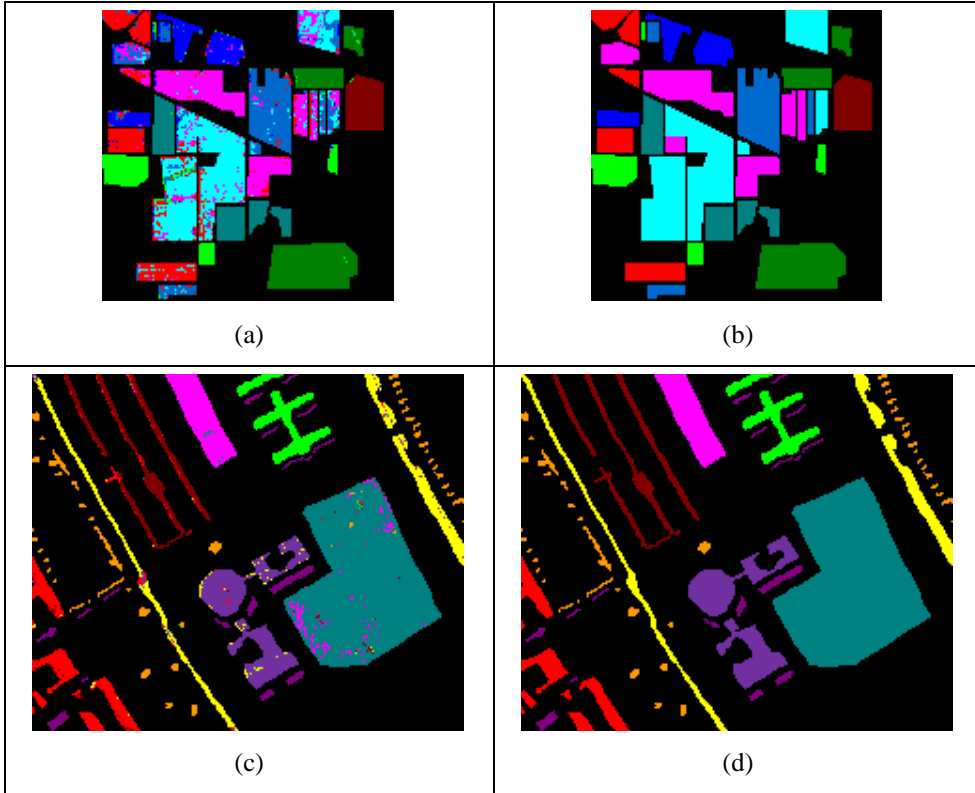


Şekil 5.40 NPE sınıflandırma sonuçları

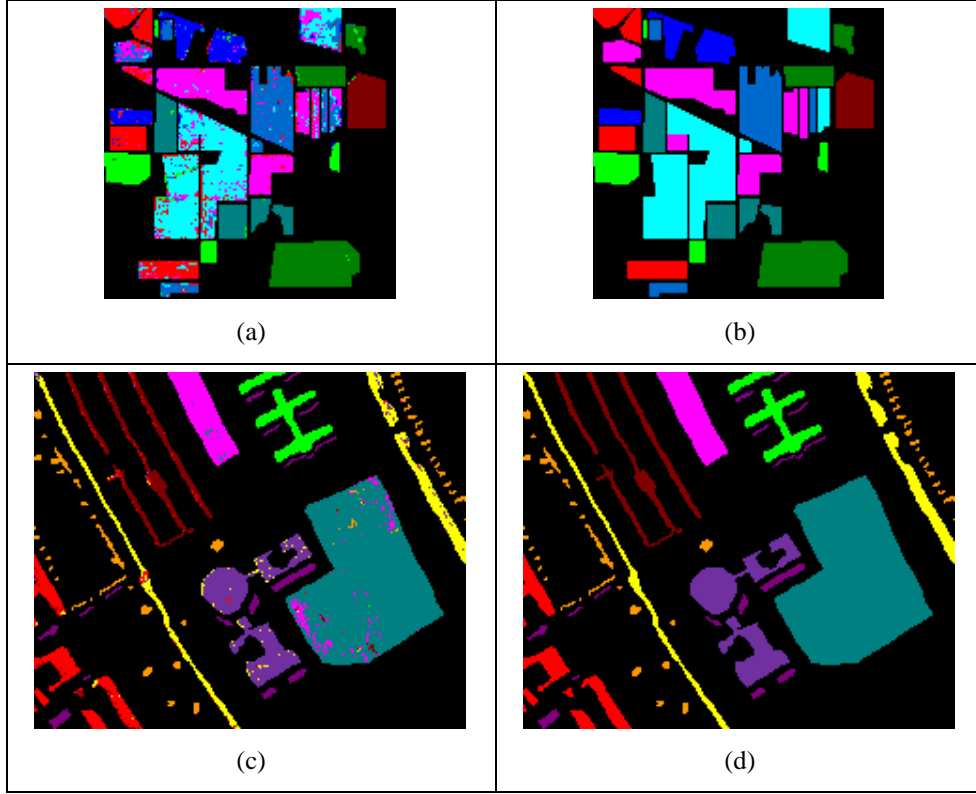




Şekil 5.41 PCA sınıflandırma sonuçları



Şekil 5.42 PPCA sınıflandırma sonuçları



Şekil 5.43 SPE sınıflandırma sonuçları

Yöntemlerden bir tanesi olan Faktör Analizi için Indian Pines verisi kullanılarak, çoğunluk oylamaları ile elde edilen değerler örnek olarak çizelge 5.2 ile verilmiştir.

Çizelge 5.2 Faktör Analizi için sınıflandırma doğrulukları

	Sınıflandırma Doğruluğu <sup>a</sup>				
	I	II	III	IV	V
1.Test	86,41	-	-	91,74	-
2.Test	84,69	-	-	89,80	-
3.Test	85,18	-	-	91,28	-
4.Test	85,26	-	-	91,26	-
5.Test	85,78	-	-	91,71	-
6.Test	84,97	-	-	91,08	-
7.Test	84,03	-	-	89,36	-
8.Test	84,22	-	-	89,87	-
9.Test	84,99	-	-	91,11	-
Sonuç	85,06	88,03	92,78	90,80	93,68

<sup>a</sup> %

Çizelge, testlerin sınıflandırma doğruluğundan oluşmaktadır. I. ve IV. adımdaki yöntemler için 9 adet testin sonucu ve en alt satırda da bu değerlerin ortalaması verilirken II. III. ve V. adımdaki yöntemlerde tekrarlı çoğunluk oylaması kullanıldığından tek bir sonuç vardır ve en alt satırda sonuç değerleri verilmiştir.

Çizelge 5.2 ile doğruluğu verilen sınıflandırma sonuçları şekil 5.44 ile gösterilmiştir. Şekilde I ve IV yöntemler için 9 testten Test 1'in sınıflandırma sonuçları verilmiştir.



Şekil 5.44 Faktör Analizi için sınıflandırma sonuçları

a. I 1.Test, b. II, c. III, d. IV 1.Test, e. V, f. Sınıflandırma haritası

Indian Pines verisi ile tüm boyut indirgeme yöntemleri ve tüm çoğunluk oylaması yöntemleri kullanılarak SVM sınıflandırıcı ile elde edilen test sonuçları çizelge 5.3'te KNN sınıflandırıcı ile elde edilen test sonuçları çizelge 5.4'te verilmiştir. Benzer şekilde Pavia University verisi için SVM sınıflandırıcı kullanılarak alınan test sonuçları çizelge

5.5, KNN sınıflandırıcı kullanılarak alınan sonuçlar çizelge 5.6 ile verilmiştir.

Çizelge 5.3 Indian Pines verisi için SVM çoğunluk oylaması test sonuçları

	Sınıflandırma Doğruluğu <sup>a</sup>				
	I	II	III	IV	V
AE	74,15	81,97	88,61	82,37	88,87
DIF	80,24	84,98	91,14	87,9	91,57
FA	85,06	88,03	92,78	90,8	93,68
FMVU	66,17	70,04	75,69	72,94	75,98
ISO	71,51	75,06	80,28	78,61	81,48
LDA	76,05	84,67	87,27	91,81	92,45
LPP	86,97	89,60	94,03	92,14	93,94
MANC	75,46	77,56	83,03	81,59	83,56
NPE	87,62	90,90	94,98	93,08	95,41
PCA	82,38	86,33	92,28	89,51	92,57
PPCA	82,26	84,99	90,13	88,21	90,12
SPE	81,30	85,96	91,54	88,73	92,15

a %

Çizelge 5.4 Indian Pines verisi için KNN çoğunluk oylaması test sonuçları

	Sınıflandırma Doğruluğu <sup>a</sup>				
	I	II	III	IV	V
AE	62,32	69,89	78,90	73,69	80,76
DIF	73,17	79,96	87,33	83,17	88,85
FA	68,81	76,44	84,19	80,49	85,89
FMVU	62,76	69,13	78,13	74,34	79,98
ISO	69,66	75,49	83,15	79,8	84,32
LDA	75,14	84,37	86,85	91,67	92,08
LPP	84,41	89,34	94,56	91,85	95,25
MANC	69,40	79,37	86,9	81,85	87,68
NPE	85,28	89,81	95,16	92,20	95,52
PCA	74,31	81,67	88,62	85,08	90,18
PPCA	75,61	82,33	88,97	86,12	90,60
SPE	75,23	82,75	89,55	85,56	90,49
Yüksek Boyut	75,29				

a %

Çizelge 5.5 Pavia University verisi için SVM çoğunluk oylaması test sonuçları

	Sınıflandırma Doğruluğu <sup>a</sup>				
	I	II	III	IV	V
AE	91,28	93,61	97,47	96,45	97,59
DIF	90,96	93,51	97,53	96,30	97,36
FA	93,31	95,19	98,50	97,64	98,69
FMVU	88,61	89,41	94,69	94,42	94,72
ISO	89,18	90,90	95,20	94,40	95,20
LDA	90,46	94,46	97,28	98,60	98,60
LPP	94,24	95,68	98,85	98,27	99,03
MANC	90,79	93,63	98,40	97,11	98,52
NPE	94,29	95,75	99,16	98,48	99,21
PCA	92,70	93,58	97,58	96,93	97,59
PPCA	92,85	94,59	97,99	97,37	98,23
SPE	92,47	94,11	97,71	97,10	97,79

<sup>a</sup> %

Çizelge 5.6 Pavia University verisi için KNN çoğunluk oylaması test sonuçları

	Sınıflandırma Doğruluğu <sup>a</sup>				
	I	II	III	IV	V
AE	87,38	90,18	95,02	94,03	95,48
DIF	86,11	88,85	94,16	93,16	94,77
FA	87,94	90,66	95,04	94,14	95,67
FMVU	85,97	89,39	94,88	93,94	95,39
ISO	88,52	90,25	94,94	94,63	95,53
LDA	88,49	93,45	96,39	97,61	98,15
LPP	88,13	92,02	97,97	96,24	98,28
MANC	78,71	85,71	94,08	89,93	94,71
NPE	85,70	89,86	96,58	94,79	96,81
PCA	89,43	92,12	95,32	95,98	96,23
PPCA	89,46	91,77	95,20	94,82	95,63
SPE	89,97	92,16	95,98	95,53	96,31
Yüksek Boyut	89,99				

<sup>a</sup> %

Çizelge 5.4, 5.6 ile verilen KNN sınıflandırma sonuçları için, en alt satırda ‘Yüksek Boyut’ ile boyut indirgeme yapmadan yapılan sınıflandırma sonuçları verilmiştir.

Çizelge 5.7 ile 12 yöntem için her iki görüntü verisini kullanarak ve SVM ile sınıflandırarak yapılan testlerin süreleri verilmiştir. Test 1’den farklı olarak boyut indirgeme süresinin yanında, sınıflandırma süreleri de toplam süreye dâhil edilmiştir. Bu süre içerisinde testlerin 9 kez tekrar edilmesi için geçen zaman da dâhildir.

Çizelge 5.7 Boyut indirgeme yöntem sürelerinin karşılaştırılması

	Test Süresi <sup>a</sup>	
	Indian Pines	Pavia University
AE	791,87	469,21
DIF	3269,11	6434,92
FA	138,35	66,40
FMVU	3025,62	2448,82
ISO	9904,13	1489,73
LDA	3,08	6,46
LPP	115,00	117,74
MANC	99,56	73,64
NPE	104,29	117,79
PCA	10,66	8,00
PPCA	386,87	422,02
SPE	20223,29	36131,14

a s

Çizelgede doğrusal boyut indirgeme teknikleri, doğrusal olmayanlara göre çok daha kısa sürede tamamlanabildiği gözlemlenmektedir. Doğrusal olmayan tekniklerden yalnız MANC süre olarak doğrusal tekniklere yakın başarımlar gösterebilmiştir.

#### 5.4 Test 3: Hughes Olgusunun Yöntemler Üzerindeki Etkisi

Bu çalışmanın amacı; Hughes olgusunun yöntemler üzerindeki etkisini araştırmaktır. Bölüm 3.1’de anlatıldığı gibi Hughes olgusu, az sayıdaki eğitim verisi ile çok sayıdaki özniteliğin sınıflandırması ve sonuç olarak da sınırlı sayıdaki veri ile boyut arttıkça sınıflandırma doğruluğunun düşmesidir (Landgrebe 2003).

Yüksek boyutun sebep olacağı sınıflandırma başarımlarındaki düşüşün, farklı boyut indirgeme yöntemleri ile veriyi azaltarak önüne geçebilmek, çok daha az veri ile daha

yüksek sınıflandırma başarımı elde etmek amaçlanmıştır. Sınıflandırıcı olarak KNN sınıflandırıcı kullanılmıştır. Bunun sebebi, KNN'in parametrik olmayan bir yöntem olması, veriye göre ayar gerektiren parametreleri bulunmamasından kaynaklanmaktadır.

Test sonunda; farklı yöntemlerle boyutu indirgenen verilerin bir kısmının, boyutu indirgenmemiş veri ile elde edilen sınıflandırma doğruluğundan daha yüksek sınıflandırma doğruluğuna ulaşılması beklenmektedir.

Çalışmada kullanılan verilerde Bölüm 4.1'de anlatıldığı gibi Indian Pines verisi için 200, Pavia University için 103 bant bulunmaktadır. Verilerin boyutu 9'a indirildiğinde ise boyutu indirgenmemiş veriye göre Indian Pines için %4,5, Pavia University için %8,7'si kullanılmış olacaktır.

Bu testte, Bölüm 5.3'te anlatılan testten farklı olarak:

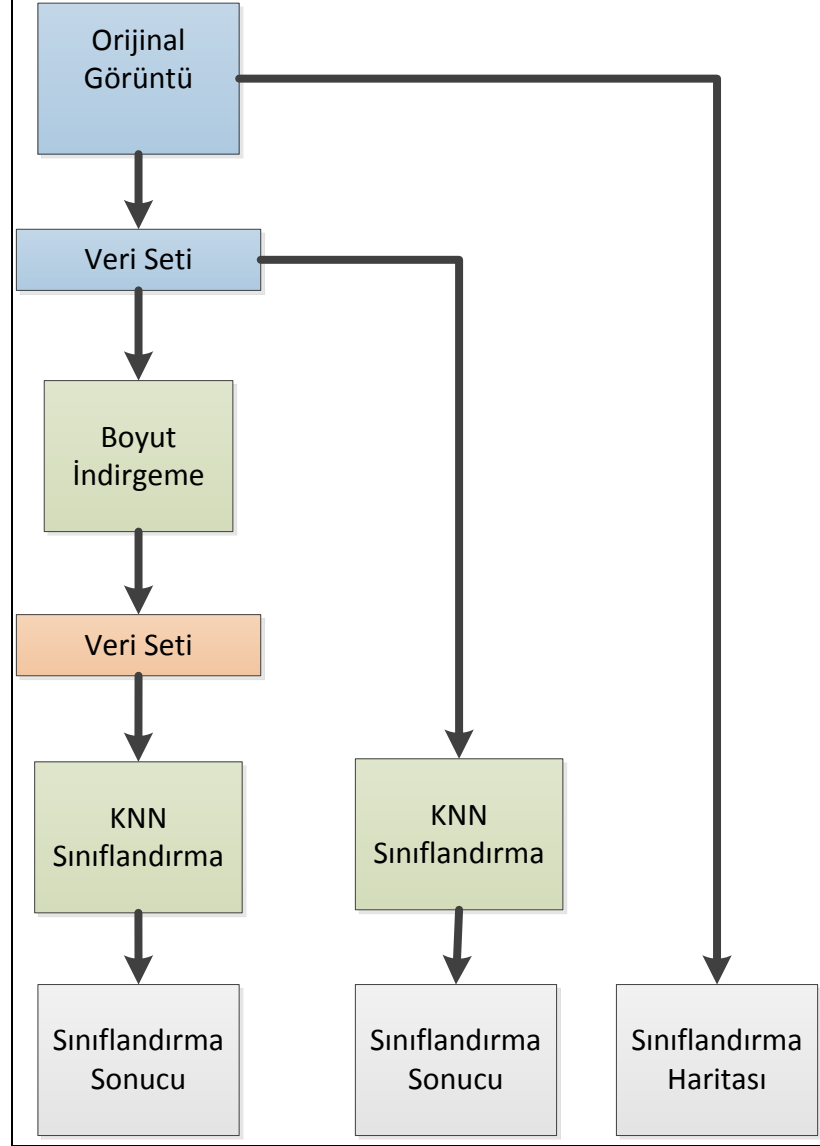
- Her test sırasında rastgele veri seçmek yerine en başta 10 farklı rastgele veri kümesi oluşturulmuş ve tüm yöntemler için aynı veriler kullanılmıştır.
- Sınıflandırıcı olarak sadece KNN kullanılmıştır.
- Öğreticili yöntemler kullanılmamıştır.

Testin ayırıcı kısmı olarak, 103 ve 200 öznitelikli (boyutlu) veriyi sınıflandırırken sınıf başına 100'er veri noktasının yanında 50'şer ve 10'ar örnek de eğitim için kullanılacak ve örnek sayısı düştükçe daha fazla yöntemin boyut indirgemenin yapılan veriye göre yüksek performans vermesi beklenecektir.

#### **5.4.1 Test akışı**

Test akışı olarak, Bölüm 5.2'de anlatılan testin KNN'le yapılan kısmının birebir aynısı kullanılacak, farklı olarak test sonunda herhangi bir çoğunluk oylaması işlemi yapılmayacaktır. Ayrıca boyut indirgeme yöntemlerinin yanında boyut indirgemesiz bir test de yapılacak ve elde edilen sonuçlar karşılaştırılacaktır. Testler 100'er, 50'şer ve 10'ar örnekle KNN eğitimi için tekrarlanacaktır. Test akışı şekil 5.45 ile gösterilmiştir. Diğer şekillerde olduğu gibi mavi renk fazla öznitelikli veriyi (200 boyut, 103 boyut),

turuncu renk düşük öznelikli veriyi (9 boyut), beyaz renk de sınıflandırma sonucunu temsil etmektedir.



Şekil 5.45 Test akışı, Test 3

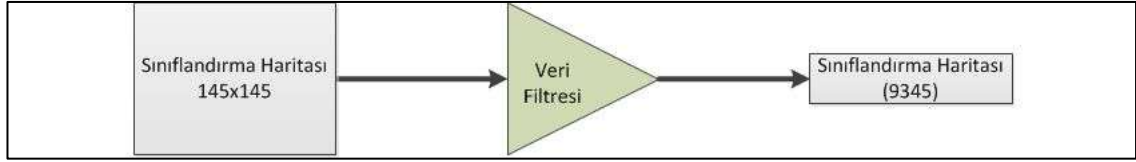
Test, her boyut indirgeme yöntemi için, 10 boyut indirgeme tekrarı ve 10 farklı sınıflandırıcı eğitimi yapacak rastgele veri olmak üzere 100 kez tekrar edilmiş ve bir veri sayısı için sonuç elde edilmiştir. 3 farklı veri sayısı için testler ayrıca tekrar edilmiş, boyut indirgmeden yapılan sınıflandırma için de yine 10 rastgele veri seti kullanılmıştır.



## 5.4.2 Test adımları

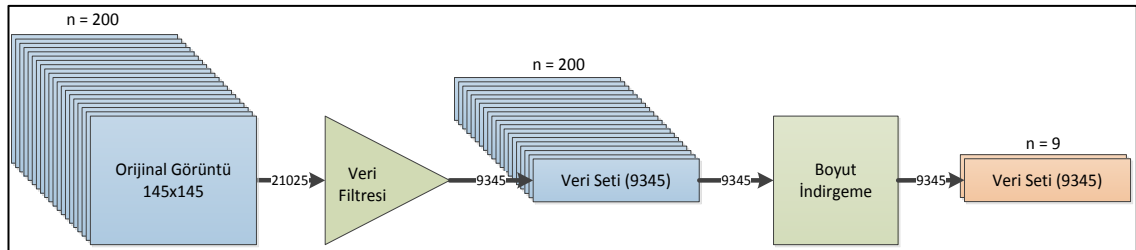
Testin detaylı adımları ise ilerleyen maddelerde verilmiştir. Her boyut indirgeme yöntemi için adımlar ayrı ayrı uygulanır:

1. Sınıflandırma haritasına veri filtresi uygulanır ve 400'den fazla örneği olan sınıfların elemanı olan veriler alınır, küçük sınıfların verileri ile belirsiz olarak belirtilip sınıflandırılmamış bölgeler çıkartılır. Sınıflandırma haritası bir veri kümesi haline getirilir (Şekil 5.46).



Şekil 5.46 Sınıflandırma Haritasına veri filtresi uygulaması Test 3

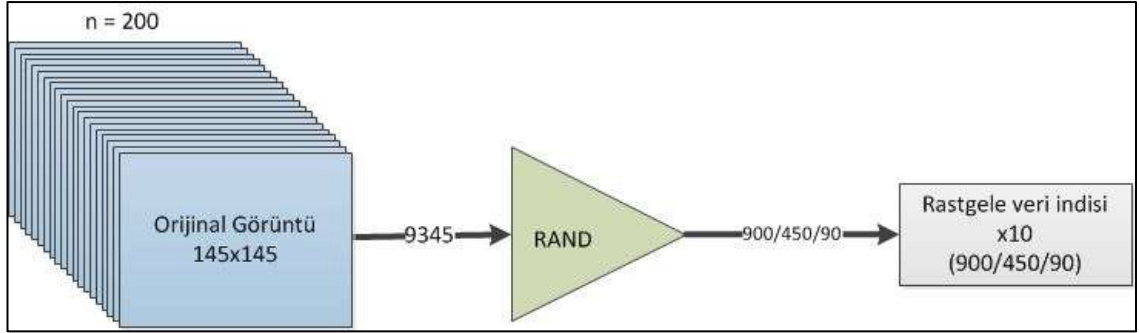
2. Orijinal görüntüye de veri filtresi uygulanır ve görüntü veri seti haline getirilir. Ardından bu veri setine boyut indirgeme yöntemleri uygulanır. Yöntem sonucunda, şekil 5.47 ile gösterildiği gibi 9345 elemanlı ve 9 boyutlu bir veri seti elde edilir. Bu işlem her yöntem için 10 kez tekrar edilir. Yüksek boyutlu veri için ise boyut indirgeme aşaması uygulanmaz.



Şekil 5.47 Boyut İndirgeme, Test 3

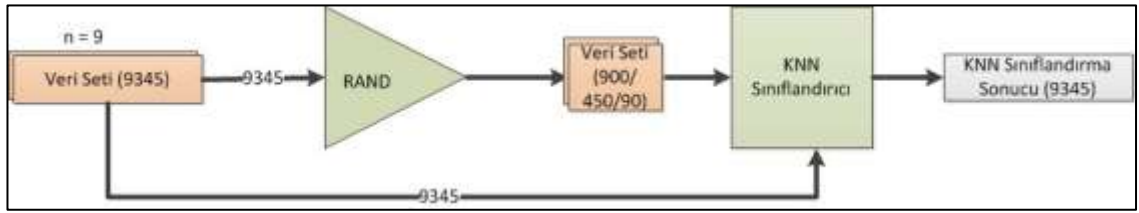
3. Sınıflandırma aşamasından önce her sınıftan 100'er, 50'şer ve 10'ar örnekli veri indisi 10kez seçilir ve tüm sınıflandırmalar için bu veri kümesi kullanılır (Şekil 5.48). Burada verilerin kendilerinin değil indislerinin saklanması sebebi, her boyut indirgeme yöntemi için ilgili pikselin farklı öznelikleri olması, boyut indirgemesiz

durumda da 9 değil daha yüksek boyut içermesidir.



Şekil 5.48 Rastgele veri kümesi oluşturma

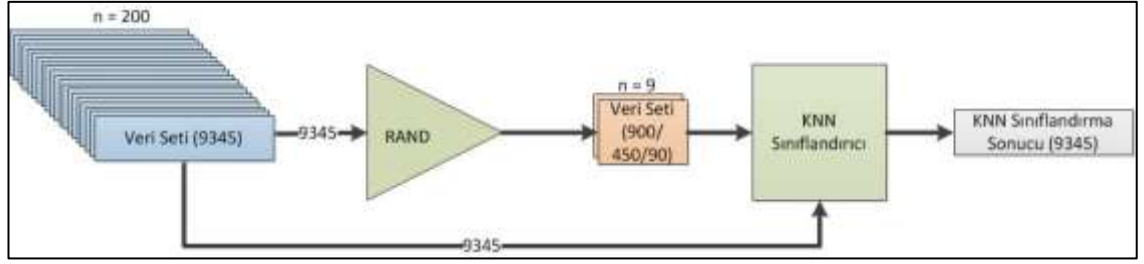
4. KNN için sınıflandırma aşamasında, boyutu indirgenmiş görüntülerden 9 sınıftan rastgele seçilen indisler için veriler alınarak sınıflandırma yapılır, (Şekil 5.49).



Şekil 5.49 Sınıflandırıcı Eğitimi, boyutu indirgenmiş veri

Hughes olgusuna göre, yöntemlerin hiperspektral veriler üzerindeki tesirini araştırmak için 4. Adım boyut indirgenmemiş veri üzerinden 10, 50 ve 100'er örnek için tekrarlanır, (Şekil 5.50).

5. Sınıflandırma sonuçları ile Sınıflandırma Haritası karşılaştırılır ve sınıflandırma doğruluğu hesaplanır.



Şekil 5.50 Sınıflandırıcı Eğitimi, boyutu indirgenmemiş veri

### 5.4.3 Test sonuçları

Yapılan testlerde her bir boyut indirgeme için, boyut indirgeme yöntemi 10 kez uygulanmış ve bunların her biri 10 farklı rastgele veri kümesi ile denenmiştir. Böylece her bir test 100 kez tekrar edilmiş olur. Ayrıca farklı eğitici nokta sayısı (100, 50, 10) için testler 3 kez yinelenmiştir.

Çizelge 5.8 Indian Pines verisi için KNN çoğunluk oylaması test sonuçları

	Sınıflandırma Doğruluğu <sup>a</sup>					
	Indian Pines			Pavia University		
	100	50	10	100	50	10
AE	61,44	56,92	47,51	88,15	86,28	79,83
DIF	72,80	68,28	<b>58,33</b>	86,42	84,87	79,12
FA	69,02	63,33	50,54	88,62	86,84	80,95
FMVU	62,42	57,57	45,73	86,37	82,91	72,24
ISO	70,17	65,79	55,81	88,52	86,53	<b>82,28</b>
LPP	<b>84,59</b>	<b>81,31</b>	<b>70,92</b>	88,05	85,91	78,47
MANC	69,55	65,41	<b>56,49</b>	79,52	75,38	67,37
NPE	<b>85,49</b>	<b>82,04</b>	<b>71,34</b>	86,10	83,40	75,47
PCA	74,75	69,32	55,89	89,68	87,72	81,18
PPCA	75,26	<b>70,20</b>	<b>58,54</b>	89,81	<b>88,04</b>	<b>82,15</b>
SPE	75,51	69,97	<b>56,46</b>	90,00	87,95	<b>81,41</b>
Boyut İndirgemesiz	75,76	70,05	56,32	90,02	87,98	81,35

<sup>a</sup> %

Çizelge 5.8’de test sonuçları verilmiştir. Her test yüz kez yinlendiğinden, sonuçlar ortalama olarak gösterilmiştir. En alt satırda boyut indirgeme yapılmadan sınıflandırılan veriler için sonuç verilmiştir. Boyut indirgeme yapılmadan alınan sonuçlardan daha

yüksek olan sonuçlar kalın yazı tipi ile belirtilmiştir.

Buna göre, piksel başına daha fazla öznitelik içeren Indian Pines verisinde Hughes etkisi daha açıkça görülmüş, daha fazla yöntem boyut indirgmeden yapılan testten başarılı olmuştur. Yine beklendiği gibi öznitelik başına düşen örnek sayısı 100'den 50 ve 10'a azaltıldıkça Hughes etkisi daha fazla yöntem için kendisini göstermiştir.

## 6. DEĞERLENDİRME

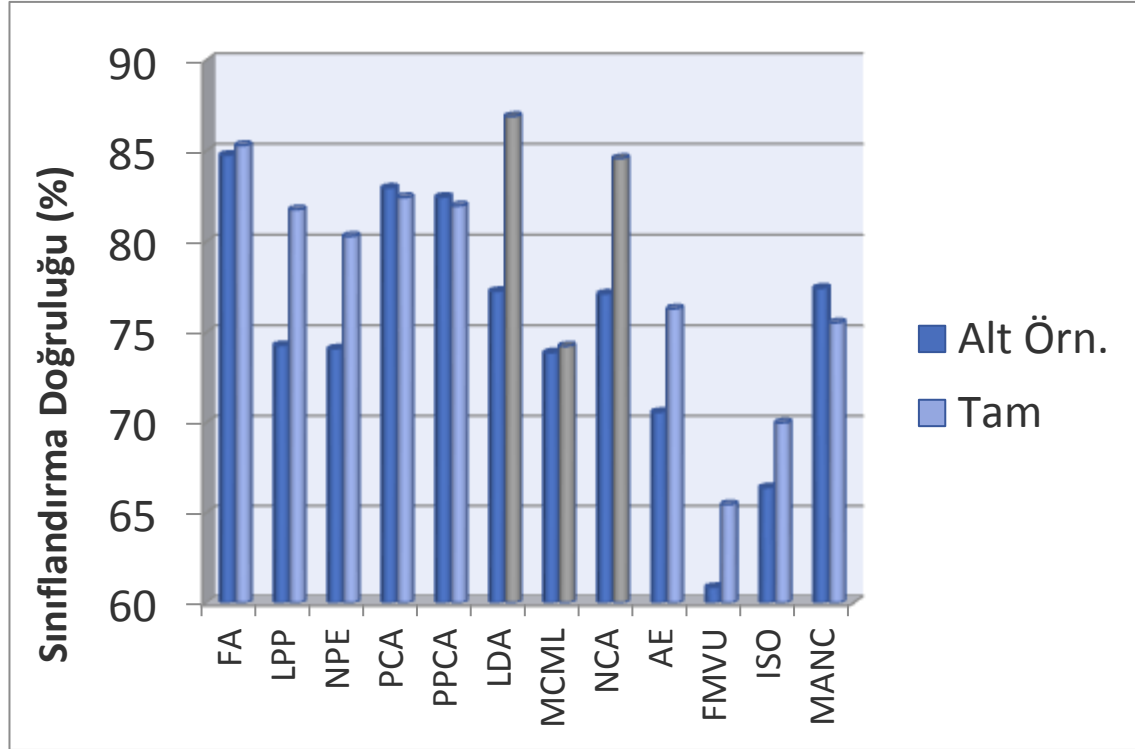
Bu tezde hiperspektral verilerin sınıflandırılmasında boyut indirgeme yöntemlerinin etkisi araştırılmıştır. Hiperspektral verilerin yüksek boyutlu olmalarından dolayı kısıtlı sayıdaki örnek ile sınıflandırıldığında Hughes olgusu ile tanımlanan sınıflandırma sorununa sebep olmaktadır. Ayrıca yüksek boyut hesaplama zamanlarını arttırmaktadır.

Bir hiperspektral görüntüleme sisteminde, örnek veriler kullanılarak elde edilen uzayların, tüm veriye uygulanmasının etkisi uzay örnekleme ile test edilmiştir. Diğer bir ifade ile tüm veriden seçilen alt örneklenmiş küme ile boyut indirgeme dönüşüm fonksiyonu elde edilmiş ve bu tüm veriye uygulanmıştır. Gerçek bir görüntüleme sisteminde örnek test verileri kısıtlı olacağı için yapılan çalışmada böyle bir durumda hangi yöntemlerin sınıflandırma sonucuna daha az etki yapacağı araştırılmıştır. Bu sayede benzer sınıflara ait görüntüler alındığında ya da bir yerin farklı zamanlarda görüntüleri alındığında, boyut indirgeme aşaması bir kez uygulanarak elde edilen dönüşüm fonksiyonunun diğer görüntülere uygulanabilirliği incelenmiştir.

Alt örnek ve tam veri seti kullanımının karşılaştırılması sonuçları çizelge 5.1 ile verilmektedir. Öğreticisiz ve doğrusal Faktör Analizi yöntemi hem tam veri seti hem de alt örneklenmiş veri seti ile yüksek sınıflandırma performansını elde etmiştir. Yine öğreticisiz doğrusal yöntemler olan PCA ve PPCA yöntemleri yüksek sonuçlar elde etmiştir. Sınıflandırma doğruluğu sonuçları grafiksel olarak şekil 6.1 ile de gösterilmiştir.

Öğreticili yöntemler için ayrı bir parantez açmak gerekir. Öğreticili yöntemlerde alt örneklenmiş veri seti için bir deney yapmak uygundur. Ancak tam veri seti ile yapılan deneyde tüm noktalar için sınıf bilgilerinin en başta bilinmesi ve boyut indirgeme yöntemine uygulanması, gerçek hayatta karşılaşılabilecek mümkün olmayan bir durumdur. Bu yöntemleri kullanabilmek için bütün pikseller için sınıflandırma sonucunun elde olması gerekir ki bu aşamadan sonra tekrar sınıflandırmaya ihtiyaç duyulmayacağı için tüm veri kümesi kullanarak öğreticili yöntemler ile test yapmak sınıflandırma açısından yanıltıcı sonuçlar verebilir. Zaten elimizde tüm noktalarının sınıfları bilinen bir veri varsa, bu sınıfları verip boyut indirgeme uygulamak ve sonrasında sınıflandırma

doğruluğu ölçmek, diğer yöntemlerle karşılaştırma açısından uygun olmamaktadır. Öğreticili yöntemler için Test 1, alt örnekleme için Test 2, tam veri kümesi için Test 3 olarak adlandırılmıştır. Öğreticisiz yöntemler için Test 1, alt örnekleme için Test 2, tam veri kümesi için Test 3 olarak adlandırılmıştır. Bu yöntemler sadece alt örnekleme için Test 1 ile karşılaştırılabilirler. Bu yüzden şekil 6.1 ile yapılan gösterimde öğreticili yöntemlerin Tam veri kümesi ile olan sonuçları gri ile belirtilmiştir.

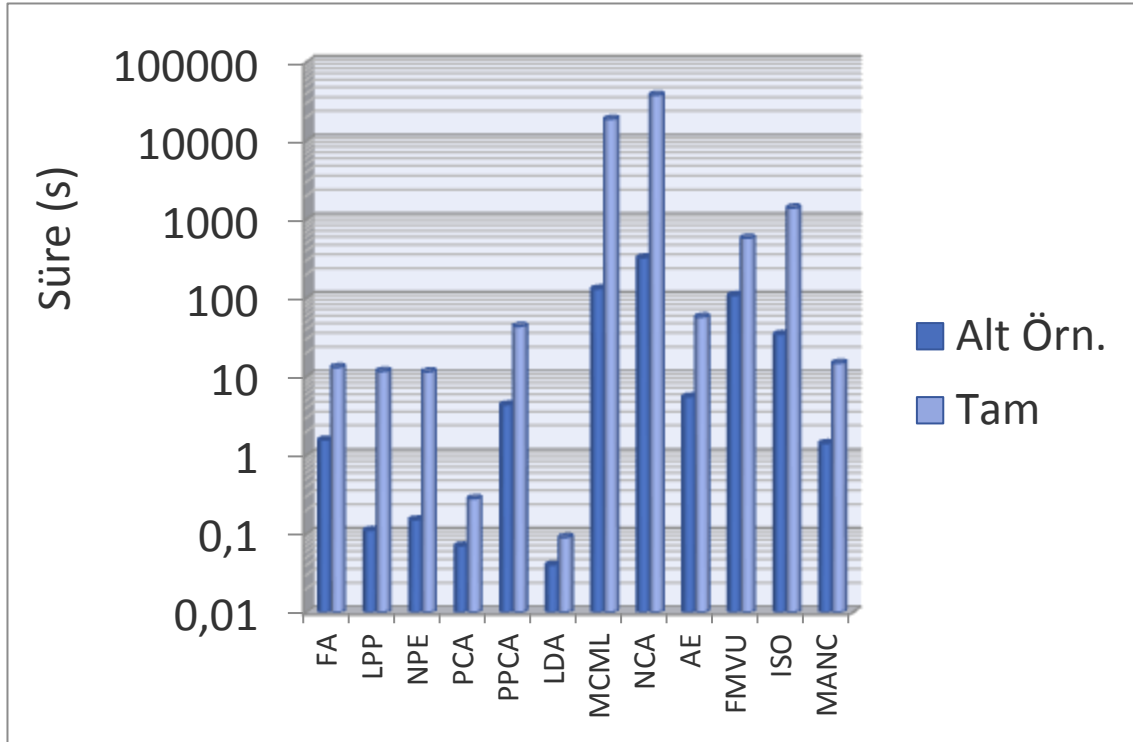


Şekil 6.1 Yöntemlerin sınıflandırma doğruluğunun karşılaştırılması, Test 1

Öğreticili yöntemlerden LDA, sahip olduğu kısıtlardan dolayı 9 yerine 8 boyuta indirmiştir ve süre olarak en iyi değerlere ulaşmıştır.

Yöntemlerin süre açısından performansları, şekil 6.2 ile grafiksel olarak gösterilmiştir. Görüldüğü üzere doğrusal ve öğreticisiz yöntemler daha kısa süre içerisinde gerekli hesaplamaları yapabilmektedir. Özellikle Faktör Analizi yöntemi alt örnekleme hesaplama süresi olarak 10 kata kadar hız artışı yakalamakla birlikte tam veri seti ile benzer sınıflandırma performansına sahip olabilmektedir. Öğreticili yöntem olan LDA en iyi süre performansını verirken diğer öğreticili yöntemler MCML ve NCA Mahalanobis uzaklığı hesapladığından süre açısından en kötü performansı vermektedir.

Bu yöntemler, alt örneklenmiş veri ile kullanıldığında süre olarak 100 kattan daha iyi sonuçlar vermektedir. PPCA, türetildiği PCA ile benzer sınıflandırma doğruluk sonuçlarını verse de süre olarak PCA'den daha kötü performans vermiştir.



Şekil 6.2 Yöntemlerin sürelerinin karşılaştırılması, Test 1

Doğrusal olmayan yöntemlerin Mahalanobis uzaklığı hesaplayan yöntemler hariç daha kötü süre sonuçları vermesi, daha karmaşık algoritmaların ve örneklem dışı yayma fonksiyonunun getirdiği hesap yükünden kaynaklanmaktadır. Bu yöntemlerin sınıflandırma doğruluğu olarak daha düşük performans vermesi ile hiperspektral verilerin yapısının doğrusal yöntemler için daha uygun olduğu çıkarımı yapılabilir.

Test 1 sonucundaki değerlendirmelerde, çizelge 5.1'e göre alt örnekleme süre yönünden çok büyük katkısı olduğu, bunun yanında sınıflandırma doğruluğundaki düşüşün yonteme göre değiştiği gözlemlenmektedir. FMVU, ISO, MCML, NCA gibi çok uzun süren yöntemlerin alt örnekleme ile kullanılabilir hale geldiği, bu yöntemler için alt örnekleme kullanmanın bir zorunluluk olduğu söylenebilir.

Çizelge 6.1 ile alt örneklenmiş veri seti sonuçlarının tam veri seti ile karşılaştırması

verilmiştir. Sınıflandırma doğruluğundaki fark yüzde olarak verilmiş, süre sütununda ise kaç kat kazanım yapıldığı gösterilmiştir.

Çizelge 6.1 Alt örneklenmiş veri seti ile tam veri seti sonuç karşılaştırması

Sınıflandırma Yöntemi	Sınıflandırma Doğruluğu <sup>a</sup>	Süre <sup>b</sup>
FA	-0,55	8,45
LPP	-7,5	106,18
NPE	-6,21	76,67
PCA	0,52	4,00
PPCA	0,47	10,00
LDA	-9,64	2,25
MCML	-0,37	144,55
NCA	-7,47	117,72
AE	-5,74	10,41
FMVU	-4,6	5,36
ISO	-3,6	40,72
MANC	1,9	10,52

a % b kat

Sınıflandırma doğruluğu olarak bakıldığında doğrusal ve öğreticisiz yöntemlerin daha iyi sonuçlar verdiği görülmüştür. Bu yöntemlerin süre olarak performanslarının da iyi olduğu düşünüldüğünde, alt örnekleme ile kazanılacak süre ile kaybedilecek sınıflandırma doğruluğu arasında bir seçim yapmak gereklidir.

Örneğin FA için %0,55'lik bir sınıflandırma doğruluğu kaybı varken süre açısından 8,5 katlık bir kazanım vardır. Tam veri seti ile yapılan bir boyut indirgeme için geçen sürenin 13 saniye olduğu dikkate alındığında %0,55'lik kayıp ile kazanılacak süre, kullanım uygulamasına göre değerlendirilmelidir.

Ancak PCA, PPCA, MANC gibi yöntemler alt örnekleme ve sonrasında yapılan örnekleme dışı yayma işleminde, tüm veriye göre daha iyi sonuçlar vermişlerdir. Bu yöntemlerde alt örneklenmiş veri ile çalışmanın daha uygun olduğu söylenebilir.

Test 2'nin sonuçları Indian Pines için çizelge 5.3 - 5.4, Pavia University için çizelge 5.5 - 5.6 ile verilmiştir.



Pavia University ile yapılan testlerde Indian Pines verisine göre çok daha yüksek başarımlar elde edilmiştir. Bunun en büyük sebebi, Indian Pines verisinde sınıflandırmaya tâbi tutulan verilerin hepsinin tarım ürünü olması yani sınıfların birbirine benzemesidir. 9 sınıfın 3 tanesi soya fasulyesi türleri, 2 tanesi mısır türleri, 2 tanesi de çimen içermektedir. Ayrıca tarım verilerinde, bitkiler farklı olsa bile bitki sınıflarının olduğu kısımlarda bitkinin yanında toprak da görüntülediği için aynı toprak bilgisi tüm sınıfların görüntüsüne karışmıştır. Bu sınıfların spektral imzalarının birbirine benzediği gibi öğreticisiz boyut indirgeme yöntemlerinin bu sınıfların farklılığını bilmeden boyut indirgediğinden, boyut indirgeme sonrası kalan verideki benzerliğin daha fazla olma ihtimali vardır. Pavia University sınıfları ise genel olarak insan yapımı, spektral imzaları birbirinden farklı maddelerden olduğundan daha yüksek sınıflandırma performansı elde edilmiştir. Birbirinden farklı sınıflar içeren görüntüler seçilmesinin sebebi, çoğunluk oylamasının tarım arazisi ya da şehir merkezi gibi farklı alanlardaki etkisini ölçmektir.

Çizelge 6.2’de yöntemler için bulunan sonuçların ortalaması alınmış ve ortalamanın üzerindeki sonuçlar kalın yazı tipi ile belirtilmiştir.

Çizelge 6.2 Farklı veri ve sınıflandırıcı sonuçlarının karşılaştırılması

Görüntü Verisi	Sınıflandırma Doğruluğu <sup>a</sup>			
	Indian Pines		Pavia University	
Sınıflandırma Yöntemi	SVM	KNN	SVM	KNN
AE	74,15	62,32	91,28	<b>87,38</b>
DIF	<b>80,24</b>	<b>73,17</b>	90,96	86,11
FA	<b>85,06</b>	68,81	<b>93,31</b>	<b>87,94</b>
FMVU	66,17	62,76	88,61	85,97
ISO	71,51	69,66	89,18	<b>88,52</b>
LDA	76,05	<b>75,14</b>	90,46	<b>88,49</b>
LPP	<b>86,97</b>	<b>84,41</b>	<b>94,24</b>	<b>88,13</b>
MANC	75,46	69,40	90,79	78,71
NPE	<b>87,62</b>	<b>85,28</b>	<b>94,29</b>	85,70
PCA	<b>82,38</b>	<b>74,31</b>	<b>92,60</b>	<b>89,43</b>
PPCA	<b>82,26</b>	<b>75,61</b>	<b>92,85</b>	<b>89,46</b>
SPE	<b>81,30</b>	<b>75,23</b>	<b>92,47</b>	<b>89,97</b>
Ortalama	79,10	73,01	91,76	87,15

<sup>a</sup> %

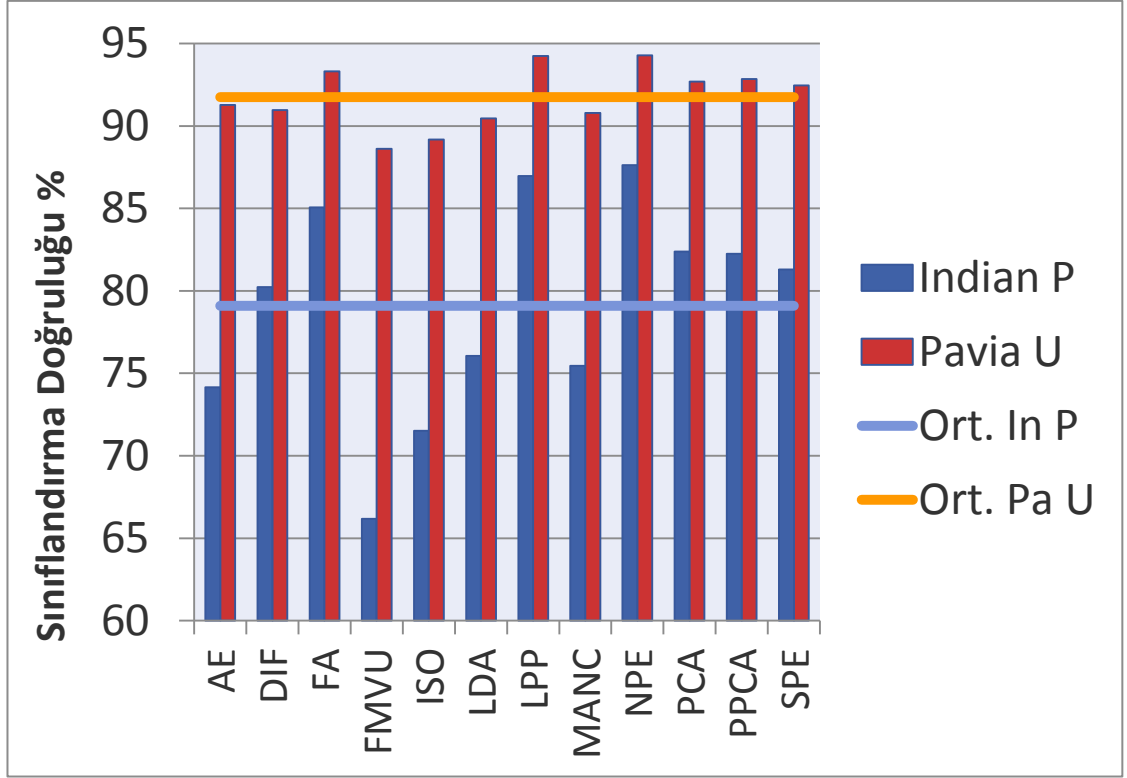
Sonuçlar incelendiğinde, tüm boyut indirgeme yöntemleri için SVM sınıflandırıcı ile KNN'e göre daha yüksek sonuçlar elde edildiği görülmüştür. Böylece, en basit sınıflandırıcılardan olan  $k = 1$  KNN sınıflandırıcıdan daha yüksek başarımlar elde edilmiş ve SVM parametrelerini doğrulanmıştır. Buradan SVM'nin  $k = 1$  KNN'e göre hiperspektral verilerde daha iyi bir sınıflandırıcı olduğu söylenebilir.

Sonuçlar karşılaştırıldığında; LPP, PCA, PPCA ve SPE'nin her zaman ortalamanın üzerinde sonuçlar verdiği, FMVU ve MANC'in ise her zaman ortalamadan düşük sonuçlar verdiği görülmüştür.

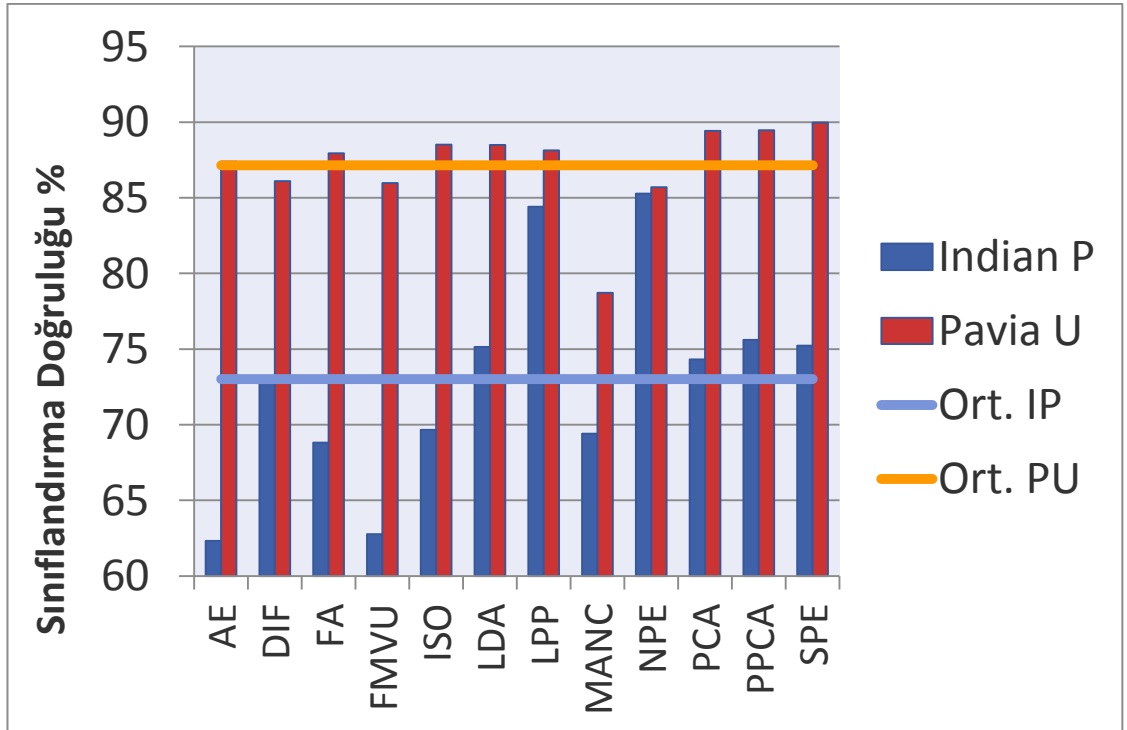
Sınıflandırma doğruluğu yöntemler açısından veriye ve kullanılan sınıflandırma yöntemine göre de farklılık göstermektedir. SVM ile sınıflandırmada PCA, PPCA'den daha üstün sonuç verirken, KNN'de ise PPCA daha üstün sonuç vermiştir.

Daha yüksek sınıflandırma doğruluğu veren SVM sınıflandırıcısı ile yapılan testlere bakıldığında, her iki görüntü için de sonuçların yüksekten düşüğe NPE, LPP, FA, PCA, PPCA, SPE diye sıralandığı gözlemlenmiştir. İncelenen 11 yöntemin 5 tanesinin doğrusal olduğu ve bu yöntemlerin sınıflandırma doğruluğu bakımından ilk 5'i oluşturduğu, herhangi bir doğrusal olmayan yöntemin doğrusal bir yöntemden daha iyi sonuç vermediği görülmüştür.

Çizelge 6.2 ile verilen sonuçlar grafiksel olarak şekil 6.3 - 6.4 ile gösterilmiştir. Grafiklerde sınıflandırma sonuçlarının ortalaması "Ort." olarak, etiketlenmiştir.



Şekil 6.3 SVM ile yapılan sınıflandırma sonuçları



Şekil 6.4 KNN ile yapılan sınıflandırma sonuçları

SVM ile sınıflandırılan ve sonrasında çoğunluk oylamasına tabi tutulan sonuçlar Indian Pines verisi için çizelge 6.3, Pavia University verisi için çizelge 6.4’de verilmiştir. Son satırda sonuçların ortalaması göstermiştir.

Çizelge 6.3 Indian Pines verisi SVM çoğunluk oylaması karşılaştırması

	Sınıflandırma Doğruluğu <sup>a</sup>				
	I	II	III	IV	V
AE	74,15	81,97	88,61	82,37	88,87
DIF	80,24	84,98	91,14	87,9	91,57
FA	85,06	88,03	92,78	90,8	93,68
FMVU	66,17	70,04	75,69	72,94	75,98
ISO	71,51	75,06	80,28	78,61	81,48
LDA	76,05	84,67	87,27	91,81	92,45
LPP	86,97	89,60	94,03	92,14	93,94
MANC	75,46	77,56	83,03	81,59	83,56
NPE	87,62	90,90	94,98	93,08	95,41
PCA	82,38	86,33	92,28	89,51	92,57
PPCA	82,26	84,99	90,13	88,21	90,12
SPE	81,30	85,96	91,54	88,73	92,15
Ortalama	79,10	83,34	88,48	86,47	89,32

<sup>a</sup> %

Çizelge 6.3’de görüldüğü gibi, Indian Pines verisi için sınıflandırma sonucu I. adım sonrası yapılan çoğunluk oylamaları II. ve IV. adımlar, sonucu pozitif olarak etkilemiştir. Tekrarlı çoğunluk oylaması ile sınıflandırma doğruluğu %79,10’dan %83,34’e çıkarken uzaysal çoğunluk oylamasında %86,47’ye yükselmiştir.

Çizelge 6.4 incelendiğinde, sınıflandırma doğruluğunun Pavia University verisinde tekrarlı çoğunluk oylaması ile %91,76’dan %93,70’e, uzaysal çoğunluk oylaması ile %96,92’ye yükseldiği görülmüştür.

Çizelge 6.4 Pavia University verisi SVM çoğunluk oylaması karşılaştırması

	Sınıflandırma Doğruluğu <sup>a</sup>				
	I	II	III	IV	V
AE	91,28	93,61	97,47	96,45	97,59
DIF	90,96	93,51	97,53	96,3	97,36
FA	93,31	95,19	98,5	97,64	98,69
FMVU	88,61	89,41	94,69	94,42	94,72
ISO	89,18	90,90	95,2	94,4	95,20
LDA	90,46	94,46	97,28	98,60	98,60
LPP	94,24	95,68	98,85	98,27	99,03
MANC	90,79	93,63	98,4	97,11	98,52
NPE	94,29	95,75	99,16	98,48	99,21
PCA	92,70	93,58	97,58	96,93	97,59
PPCA	92,85	94,59	97,99	97,37	98,23
SPE	92,47	94,11	97,71	97,1	97,79
Ortalama	91,76	93,70	97,53	96,92	97,71

a %

Uzaysal çoğunluk oylaması yöntemi sınıflandırma sonuçlarını iyileştirmede yani hatalı sınıflandırma sonuçlarını düzeltmede tekrarlı çoğunluk oylamasına göre daha başarılıdır. Bunun en büyük sebebi, birbirine yakın noktalarda olan piksellerin aynı sınıfta olmasından kaynaklanmaktadır.

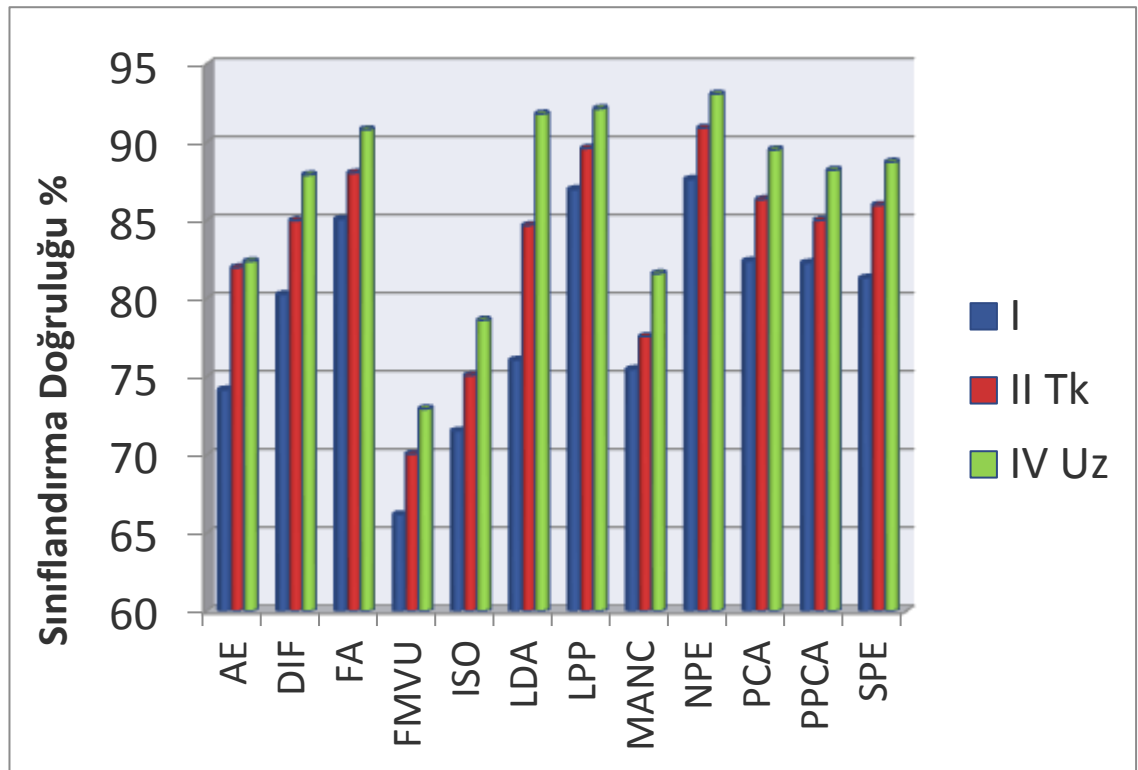
Çizelge 6.5 ile her iki görüntü verisi için, SVM ile yapılan sınıflandırma sonrasında, ilk aşamada yapılan tekrarlı çoğunluk oylamasının etkisi (II-I) ve uzaysal çoğunluk oylamasının etkisi (IV-I) bu yöntemler uygulanarak bulunan sınıflandırma doğruluğundan çoğunluk uygulaması olmadan bulunan sınıflandırma doğruluğu çıkartılarak gösterilmiştir.

Çizelge 6.5'in açıklaması olarak SVM ile ilk aşamada yapılan çoğunluk oylamasının etkileri Indian Pines görüntüsü için şekil 6.5 ile Pavia University için şekil 6.6 ile grafiksel olarak ifade edilmiştir.

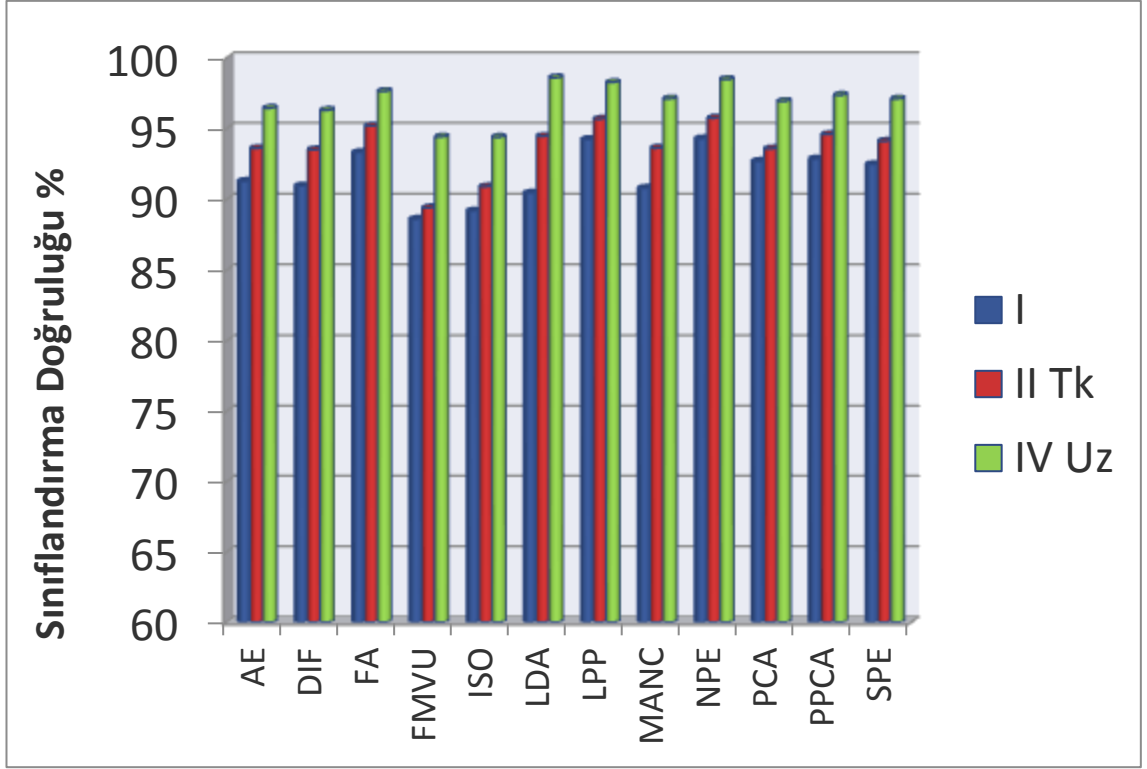
Çizelge 6.5 İlk aşama çoğunluk oylamanın sınıflandırma sonucuna etkisi

	Sınıflandırma Doğruluğu <sup>a</sup>			
	Indian Pines		Pavia University	
	II-I	IV-I	II-I	IV-I
AE	7,82	8,22	2,33	5,17
DIF	4,74	7,66	2,55	5,34
FA	2,97	5,74	1,88	4,33
FMVU	3,87	6,77	0,80	5,81
ISO	3,55	7,10	1,72	5,22
LDA	8,62	15,76	4,00	8,14
LPP	2,63	5,17	1,44	4,03
MANC	2,10	6,13	2,84	6,32
NPE	3,28	5,46	1,46	4,19
PCA	3,95	7,13	0,88	4,23
PPCA	2,73	5,95	1,74	4,52
SPE	4,66	7,43	1,64	4,63
Ortalama	4,24	7,38	1,94	5,16

<sup>a</sup> %



Şekil 6.5 Indian Pines, ilk aşama çoğunluk oylaması yöntemleri



Şekil 6.6 Pavia University, ilk aşama çoğunluk oylaması yöntemleri

Bu verilere göre, her iki çoğunluk oylamasında da tüm yöntemler için pozitif değerlere ulaşılmış, çoğunluk oylaması yöntemlerinin sınıflandırma sonuca etkisinin olumlu olduğu görülmüştür. Tekrarlı çoğunluk oylamasının etkisi Indian Pine için %4,24 ve Pavia University verisi için %1,94 iken; uzaysal çoğunluk oylamasının etkisi sırası ile %7,38 ve %5,16 olmuştur. İlk aşama sınıflandırmada uzaysal çoğunluk oylamasının hem tarımsal hem de kentsel verilerde tekrarlı çoğunluk oylamasına göre çok daha başarılı olduğu görülmüştür. Sonuç olarak en dikkat çekici yöntem LDA'dir. LDA kullanılarak yapılan testin ilk aşama çoğunluk oylaması sonrası sınıflandırma doğruluğundaki artış hem tekrarlı hem uzaysal çoğunluk oylaması için, ortalama artışın iki katıdır.

Bunun yanında, hem tekrarlı hem de uzaysal çoğunluk oylamasının Indian Pines verisinde Pavia University verisine göre daha başarılı olmasının sebebi, Indian Pines verisi sınıflandırma doğruluğunun daha düşük olması ve bu çoğunluk oylaması yöntemlerinin düzeltebileceği daha fazla hata bulunmasından kaynaklanmaktadır.

Birinci aşama oylamanın ardından yapılan ikinci aşama oylamanın sonuçları tablo olarak çizelge 6.6, grafiksel olarak şekil 6.7 - 6.8 ile verilmiştir.

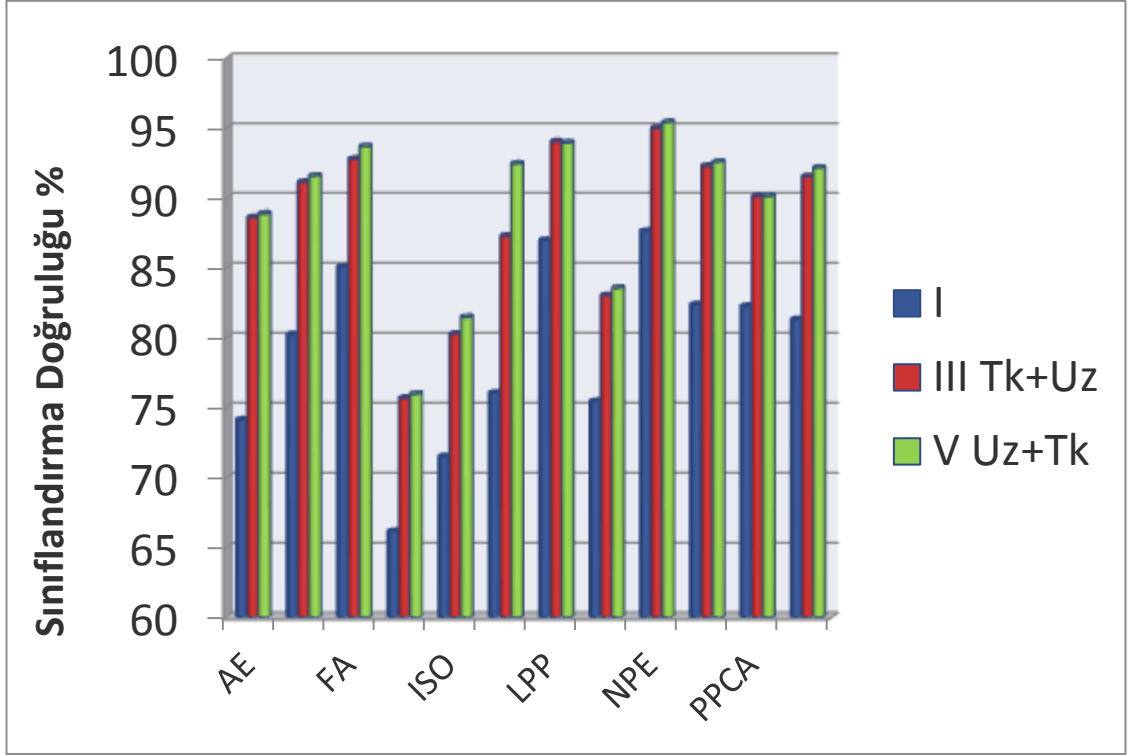
Çizelge 6.6 İkinci aşama çoğunluk oylamanın sınıflandırma sonucuna etkisi

	Sınıflandırma Doğruluğu <sup>a</sup>			
	Indian Pines		Pavia University	
	III-II	V-IV	III-II	V-IV
AE	6,64	6,50	3,86	1,14
DIF	6,16	3,67	4,02	1,06
FA	4,75	2,88	3,31	1,05
FMVU	5,65	3,04	5,28	0,30
ISO	5,22	2,87	4,30	0,80
LDA	2,60	0,64	2,82	0
LPP	4,43	1,80	3,17	0,76
MANC	5,47	1,97	4,77	1,41
NPE	4,08	2,33	3,41	0,73
PCA	5,95	3,06	4,00	0,66
PPCA	5,14	1,91	3,40	0,86
SPE	5,58	3,42	3,60	0,69
Ortalama	5,14	2,84	3,83	0,79

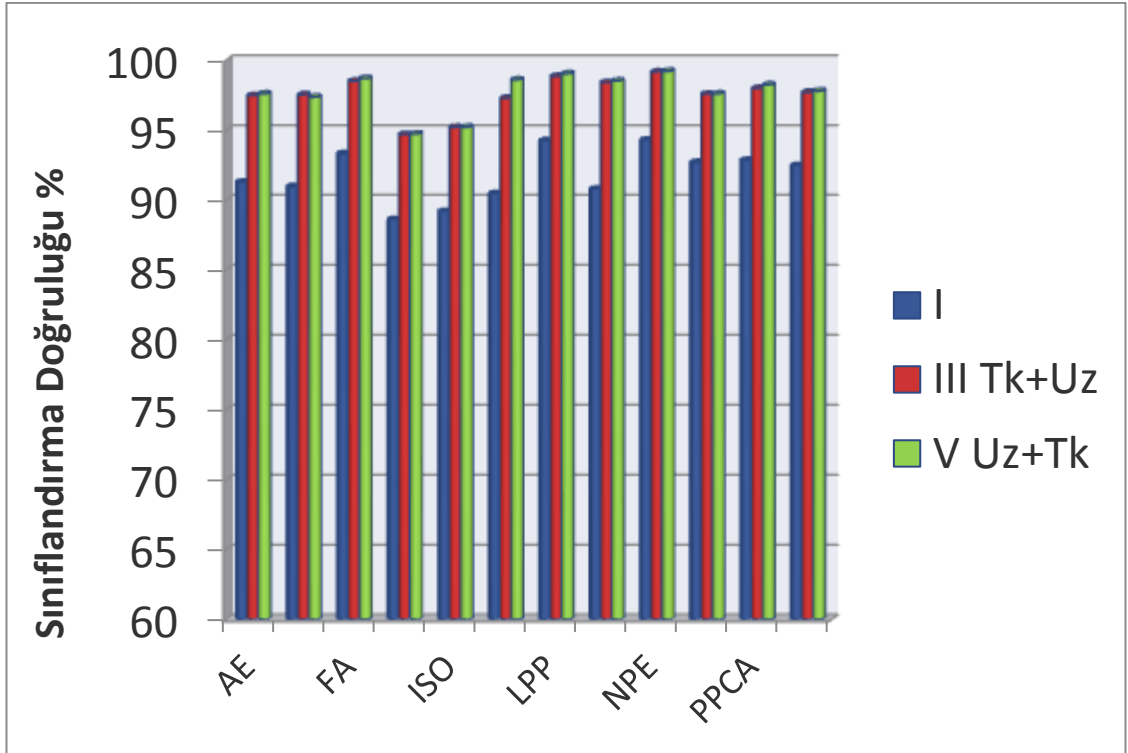
<sup>a</sup> %

Çizelge 6.6 incelendiğinde; III. adımda yapılan uzaysal çoğunluk oylamasının etkisi ikinci adımda bulunan sonuçtan çıkartılarak (III-II), Indian Pines verisi için ortalama %5,14; Pavia University verisi için %3,83 bulunmuştur. V. adımda yapılan tekrarlı çoğunluk oylamasının etkisi ise Indian Pines verisi için ortalama %2,84, Pavia University verisi için %0,79'dur. Her iki görüntü için de ikinci aşama sınıflandırmada III. adımda yapılan uzaysal çoğunluk oylamasının etkisi, V. adımda yapılan tekrarlı çoğunluk oylamasına göre fazladır.





Şekil 6.7 Indian Pines, ikinci aşama çoğunluk oylaması yöntemleri



Şekil 6.8 Pavia University, ikinci aşama çoğunluk oylaması yöntemleri

Uzaysal çoğunluk oylamasının tek bir test ile yapılabildiği, tekrarlı çoğunluk oylamasında ise testlerin sınıflandırma aşamasının çoğunluk oylamasına girdi sağlayacak veri kadar tekrar edilmesi gerektiği düşünüldüğünde tüm veri kümeleri için hem sınıflandırma doğruluğu hem de süre yönünden uzaysal çoğunluk oylamasının çok daha verimli olduğu görülmektedir.

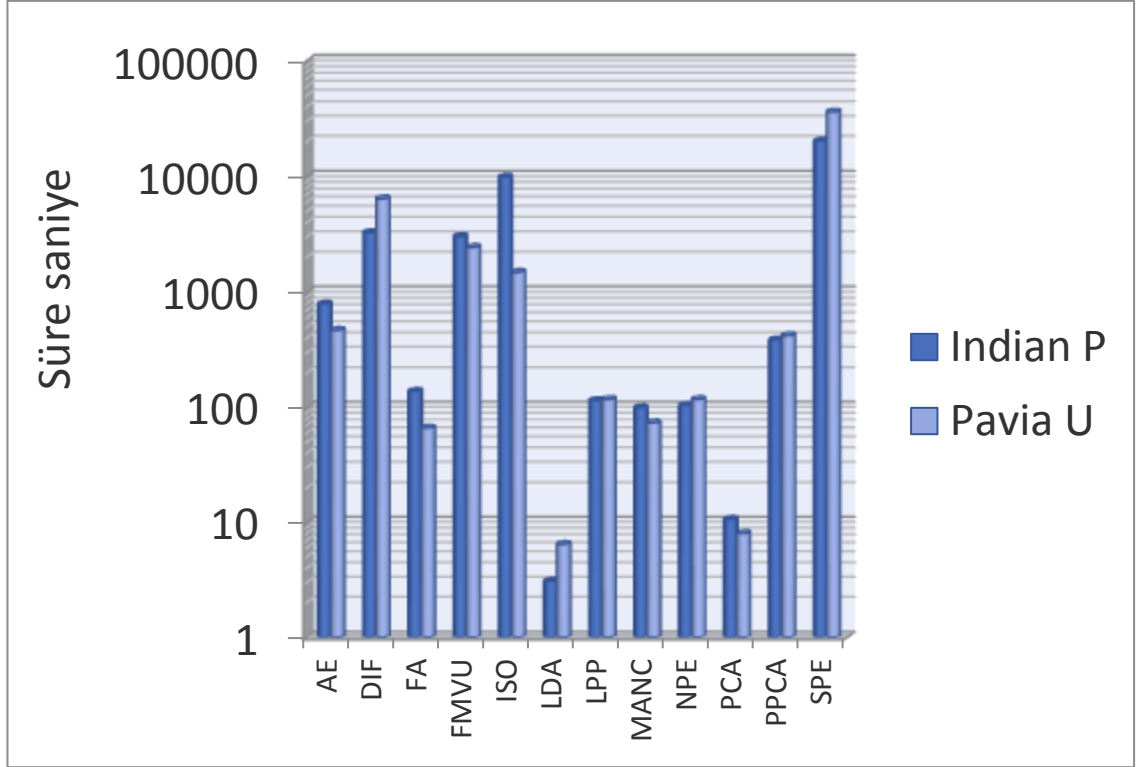
Test 2'ye süre olarak bakıldığında çizelge 6.7 ve şekil 6.9 ile verildiği gibi, MANC hariç doğrusal olmayan yöntemler süre performansı açısından kötü olup uygulamalarda kullanılmaya elverişli değildir.

Çizelge 6.7 Boyut indirgeme yöntem sürelerinin karşılaştırılması Test 2

	Test Süresi <sup>a</sup>	
	Indian Pines	Pavia University
AE	791,87	469,21
DIF	3269,11	6434,92
FA	138,35	66,40
FMVU	3025,62	2448,82
ISO	9904,13	1489,73
LDA	3,08	6,46
LPP	115,00	117,74
MANC	99,56	73,64
NPE	104,29	117,79
PCA	10,65	8,00
PPCA	386,87	422,02
SPE	20223,29	36131,14

<sup>a</sup> s

Yöntemlerin bir yarısı, daha az veri içeren Indian Pines testini daha kısa zamanda tamamlarken, diğer yarısı daha ayrılabilir sınıflardan oluşan Pavia University verisini daha kısa zamanda sınıflandırmıştır. Bunun sebebi farklı tekniklerin kullandığı farklı metrikler ve lokal ya da global hesaplamaların etkisidir. Bir kısmı, sınıflar arasında daha fazla fark olan ve daha düşük boyutlu olan Pavia University verisini daha çabuk hesaplarken, diğer kısım daha az veri noktası içeren Indian Pines verisinde zaman yönünden daha başarılı olmuştur.

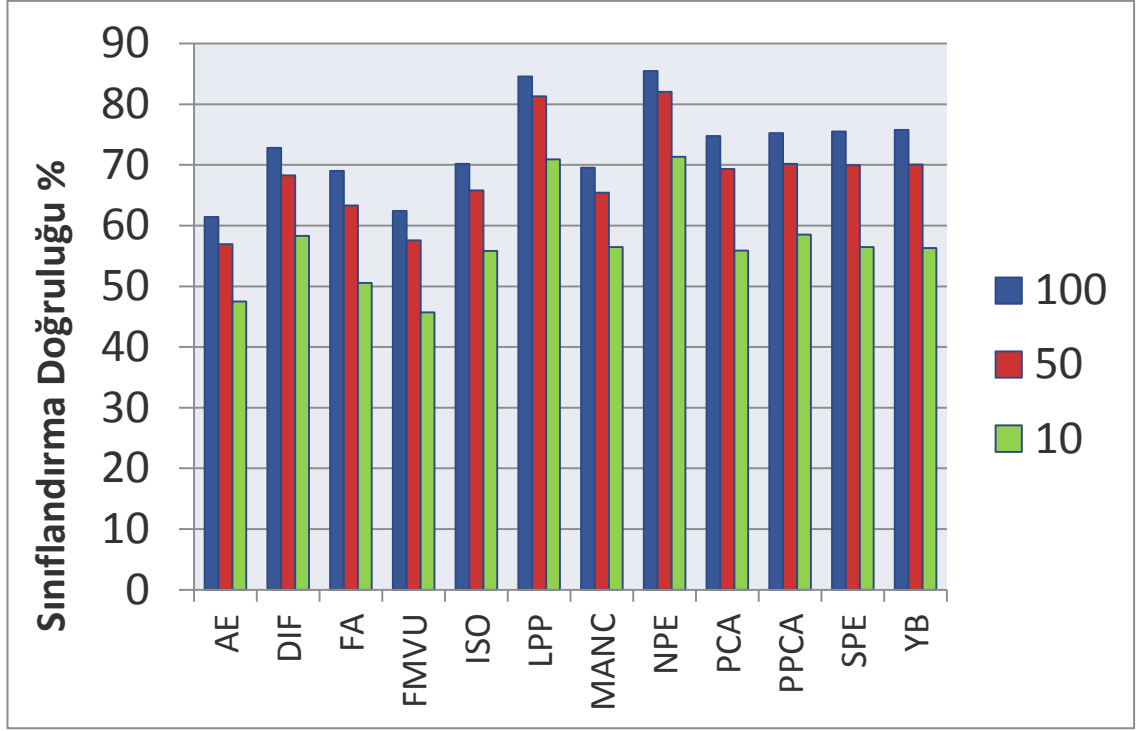


Şekil 6.9 Test 2 SVM, boyut indirgeme yöntem sürelerinin karşılaştırılması

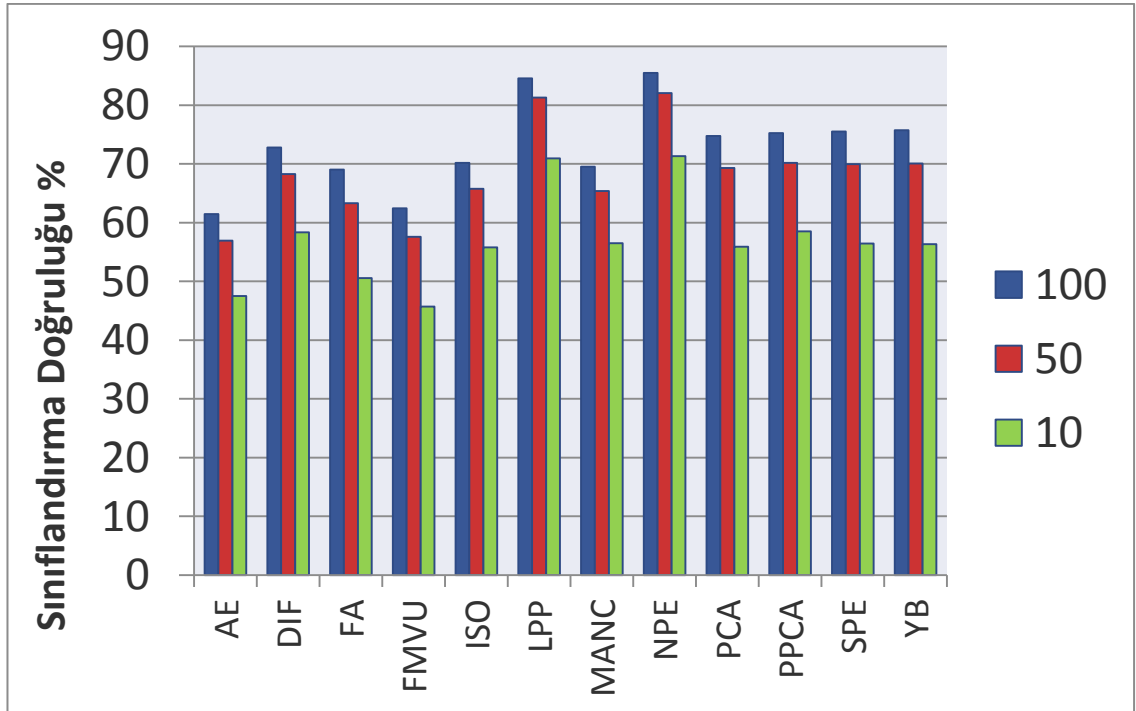
Hughes olgusunun etkisini ölçmek için yapılan ve çizelge 5.8 ile verilen testin sonuçları grafik olarak şekil 6.10 - 6.11 ile gösterilmiştir.

Grafikte en sağ veri olarak “YB” ile etiketlenmiş veri “Yüksek Boyut”u ifade etmektedir ve boyutu indirgenmemiş veri ile yapılan testleri gösterir. KNN eğitim kümesi olarak sınıf başına 100, 50 ve 10 verilik değerler ayrı ayrı gösterilmiştir.

Daha az sayıda özneliğe sahip Indian Pines verisinde yüksek boyuta göre daha fazla boyut indirgeme yöntemi iyi sonuç vermiştir. Boyut indirgenmemiş duruma göre sadece %4,5 oranında veri kullanılmasına rağmen 100 eğitim noktası için LPP ve NPE yüksek boyuttan daha iyi sonuçlar verirken eğitim noktası sayısı 10’a düşürüldüğünde yöntemlerinin çoğunun %4,5 oranında veri ile daha iyi sonuçlar verdiği gözlemlenmiştir.



Şekil 6.10 Indian Pines, Test 3 sonuç karşılaştırması



Şekil 6.11 Pavia University, Test 3 sonuç karşılaştırması

Pavia University verisinin Indian Pines verisine göre sahip olduğu daha düşük boyuttan

dolayı hem veri oranı hem de öznitelik başına düşen eğitim kümesi oranı Indian Pines verisine göre düşüktür. Bu yüzden Hughes etkisi bu veriye göre daha az gözlemlenmiştir.

Bu veride toplam verinin %8,7'si kullanılmış, 100'er nokta verildiğinde Hughes olgusu gözlemlenmezken, 10 eğitim noktasına düştüğünde ISO, PPCA ve SPE yüksek boyuttan daha iyi sonuçlar vermiştir. PCA ve FA ile ise yakın sonuçlar alınmıştır.

Beklendiği gibi, boyut başına düşen veri noktası azaldıkça verilerin sınıflandırılması Hughes olgusunun belirttiği şekilde zorlaşmaktadır. Bu sebeple eğitim noktası azaldıkça boyut indirgeme yöntemleri %10'un daha altında veri kullanarak, boyut indirgemesiz veriden daha iyi sonuçlar vermeye başlamıştır.

## 7. SONUÇ

Tüm çalışmaya sonuç olarak bakıldığında, doğru parametreler kullanılarak eğitilen SVM'nin KNN'e göre tüm yöntemler için üstün sonuç verdiği görülmüştür. Dolayısı ile hiperspektral görüntülerin sınıflandırılmasında SVM, KNN'e göre tercih edilmelidir. SVM'nin dezavantajı, her veri ve her yonteme göre  $C$  ceza parametresi ve radyal taban fonksiyonu  $\gamma$  parametresinin deęişkenlik göstermesidir. Dolayısı ile en başta her veri için Bölüm 5.1.1'de anlatıldığı gibi bir parametre taraması yapmak gereklidir. Bu da sınıflandırma aşamasını defalarca tekrarlamak anlamına gelmektedir.

Sınıflandırma öncesi alt örnekleme yaparak boyut indirgemek, sınıflandırma yöntemine ve uygulamaya göre seçilmelidir. Bazı yöntemler için sürede büyük bir kazanım, doğrulukta ise küçük bir düşüş olurken, bazı yöntemlerin sınıflandırma doğruluklarında %9,6'lık veri kullanıldığında yüksek düşüşler yaşanmıştır. Uygulamadaki hız ihtiyacı ve sınıflandırma performansındaki düşüş kıstaslarını değerlendirerek uygun olan veri kümesini seçmek lazımdır.

En yüksek sınıflandırma doğrulukları LPP ve NPE, ardından FA ve onun ardından PCA ve PPCA yöntemlerinde elde edilmiş olup bunların tamamı doğrusal yöntemlerdir. Bu yöntemlerin tamamı süre açısından iyi performans vermiştir. PCA'in süre performansı açısından diğerlerine göre üstünlüğü vardır. Tüm bunlar göz önüne alındığında SVM ile sınıflandırmada en yüksek sınıflandırma doğruluğuna sahip NPE ardından da ona çok yakın sonuçlar veren LPP yöntemi önerilmektedir. Bu iki yöntem, süre performansı yönünden de benzerlik göstermektedir. Ayrıca sınıflandırma doğruluğu bunların bir derece altında olsa da süre açısından çok daha iyi performans gösteren PCA de hiperspektral görüntülerin SVM ile sınıflandırılmasında kullanılabilir.

Sınıflandırıcıyı eğitmek için elde ne kadar az veri varsa hiperspektral verilerde boyut indirgemek o kadar elzem hale gelmektedir. Boyut indirgemek, işlemci gücü ve işlem zamanından kazanç sağlamanın yanında, azalan örnek sayısı ve yüksek boyut yani yüksek öznitelik sayısı durumlarında sınıflandırma performansını artırmaktadır.

Görüntü verisine göre, çoğunluk oylaması kullanıp kullanılmamaya karar verilebilir.

Uzaysal çoğunluk oylaması, hem performansı yönünden çok kullanışlıdır, hem de testlere ek bir yük getirmemektedir. Ancak dezavantaj olarak, veri içerisinde büyük sınıfların içine gizlenmiş küçük sınıflar olduğunda uzaysal çoğunluk oylamasının bu sınıfların doğru olarak sınıflandırılmasını engelleme ihtimali vardır. Tekrarlı çoğunluk oylaması ise aynı testi belli sayıda tekrar edip bunların sonuçlarının oylanmasına dayandığı için hem tekrar dolayısı ile test süresi yönünden dezavantajlıdır hem de sınıflandırma doğruluğu başarımına etkisi uzaysal çoğunluk oylamasına göre düşüktür. Bunun yanında sınıflandırma sonuçları için pozitif bir etki yaptığı düşünüldüğünde, testteki tekrardan kaynaklanan zaman kaybı göz önüne alınarak sınıflandırma doğruluğunun yükseltilmesi için kullanılabilir.

Bu tezde operasyonel bir hiperspektral sınıflandırma yönteminin geliştirilmesi için kıyaslamalı çalışmalar yapılmıştır. Farklı alanlarda kullanılan çeşitli boyut indirgeme tekniklerinin, hiperspektral verilerin sınıflandırmasındaki performansları incelenmiştir. Bunun yanında bu yöntemlerin hiperspektral görüntülerdeki alt örneklenmiş veri setine tepkisi sınıflandırma başarımı yönünden ölçülmüştür. Ayrıca farklı yöntemlerin KNN ve SVM sınıflandırıcıları ile başarımları incelenmiş, Tekrarlı Çoğunluk Oylaması ve Uzaysal Çoğunluk Oylaması kullanarak sınıflandırma başarımları yükseltilmiştir.

## KAYNAKLAR

- Agrafiotis, D. K. 2003. Stochastic proximity embedding. Journal of computational chemistry, 24(10), 1215-1221.
- Anonymous. 2013. Pavia University scene. Web Sitesi: [http://www.ehu.es/ccwintco/index.php/Hyperspectral Remote Sensing Scenes](http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes), Eriřim Tarihi: 15.06.2013.
- Anonymous. 2013. Indian Pines 220-band data set. Web Sitesi: <ftp://ftp.ecn.purdue.edu/biehl/MultiSpec/92AV3C.tif.zip>, Eriřim Tarihi: 08.02.2013.
- Anonim. 2014. İleri Görüntüleme Teknolojileri (TUYGUN) Projesi. Web Sitesi: <http://www.ssm.gov.tr/anasayfa/projeler/Sayfalar/proje.aspx?projeID=219>, Eriřim Tarihi: 15.03.2014
- Anonim. 2014. İMECE Projesi. Web Sitesi: <http://www.uzay.tubitak.gov.tr/?q=tr/uydu-uzay/imece>, Eriřim Tarihi: 01.01.2014
- Anonymous. 2013. AVIRIS and Other Imaging Spectrometers. Web Sitesi: [http://rst.gsfc.nasa.gov/Sect13/Sect13\\_9.html](http://rst.gsfc.nasa.gov/Sect13/Sect13_9.html), Eriřim Tarihi: 19.05.2013.
- Anonymous. 2013. AVIRIS Moffett Field Image Cube. Web Sitesi: [http://aviris.jpl.nasa.gov/data/image\\_cube.html](http://aviris.jpl.nasa.gov/data/image_cube.html), Eriřim Tarihi: 12.09.2013.
- Anonymous. 2013. Electromagnetic Spectrum. Web Sitesi: [http://earthobservatory.nasa.gov/Features/RemoteSensing/remote\\_03.php](http://earthobservatory.nasa.gov/Features/RemoteSensing/remote_03.php), Eriřim Tarihi: 15.12.2013.
- Anonymous. 2013. How Landsat Images Are Made. Web Sitesi: <http://landsat.gsfc.nasa.gov/wp-content/uploads/2012/12/How2make.pdf>, Eriřim Tarihi: 16.12.2013.
- Anonymous. 2013. More than a Pretty Picture The Making of Landsat Images. Web Sitesi: [http://landsat.gsfc.nasa.gov/wp-content/uploads/2013/05/Landsat\\_MoreThanPrettyPicture\\_2013Apr11.ppt](http://landsat.gsfc.nasa.gov/wp-content/uploads/2013/05/Landsat_MoreThanPrettyPicture_2013Apr11.ppt), Eriřim Tarihi: 17.12.2013.
- Anonymous. 2013. The Electromagnetic Spectrum Image. Web Sitesi: [http://mynasadata.larc.nasa.gov/images/EM\\_Spectrum3-new.jpg](http://mynasadata.larc.nasa.gov/images/EM_Spectrum3-new.jpg), Eriřim Tarihi: 10.12.2013.
- Anonim. 2014. TÜBİTAK UZAY Uydu Uygulamaları. Web Sitesi: <http://www.uzay.tubitak.gov.tr/tubitakUzay/tr/projects/spaceApplications.php>, Eriřim Tarihi: 21.05.2014
- Bazi, Y. and Melgani, F. 2006. Toward an optimal SVM classification system for hyperspectral remote sensing images. Geoscience and Remote Sensing, IEEE Transactions on, 44(11), 3374-3385.
- Brand, M. 2002. Charting a manifold. Advances in neural information processing systems 15, MIT Press 985-992.



- Byrnes, J. S. 2009. Unexploded Ordnance Detection and Mitigation. Springer, 286, Dordrecht.
- Campbell, J. B. 2002. Introduction to remote sensing. The Guilford Press, 642, New York.
- Camps-Valls, G. Gomez-Chova, L. Calpe-Maravilla, J. Martin-Guerrero, J. D. Soria-Olivas, E. Alonso-Chorda, L. Moreno, J. 2004. Robust support vector method for hyperspectral data classification and knowledge discovery. *Geoscience and Remote Sensing, IEEE Transactions on*, 42(7), 1530-1542.
- Chang, C. and Lin, C. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2, 27:1--27:27.
- Cortes, C. and Vapnik, V. 1995. Support-vector networks. *Machine learning*, 20(3), 273-297.
- Cover, T. and Hart, P. 1967. Nearest neighbor pattern classification. *Information Theory, IEEE Transactions on*, 13(1), 21-27.
- Demir, B. 2010. Hiperspektral görüntülerin yüksek doğruluklu sınıflandırılması. Doktora Tezi, Kocaeli Üniversitesi, Fen Bilimleri Enstitüsü, Elektronik ve Haberleşme Mühendisliği Anabilim Dalı, Kocaeli.
- Fodor, I. K. 2002. A survey of dimension reduction techniques. A survey of dimension reduction techniques. Technical Report UCRL-ID-148494, Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, Livermore, CA, ABD.
- Folkman, M. A. Pearlman, J. Liao, L. B. Jarecke, P. J. 2001. EO-1/Hyperion hyperspectral imager design, development, characterization, and calibration. Second International Asia-Pacific Symposium on Remote Sensing of the Atmosphere, Environment, and Space, Proc. SPIE 40-51.
- Foody, G. McCulloch, M. Yates, W. 1995. The effect of training set size and composition on artificial neural network classification. *International Journal of Remote Sensing*, 16(9), 1707-1723.
- Fukunaga, K. 1990. Introduction to statistical pattern recognition. Academic Press, 592, San Diego, CA, ABD.
- Ghahramani, Z. Hinton, G. E. 1996. The EM algorithm for mixtures of factor analyzers. Technical Report CRG-TR-96-1, University of Toronto, Toronto, Kanada.
- Globerson, A. and Roweis, S. T. 2005. Metric learning by collapsing classes. *Advances in neural information processing systems* 18, MIT Press, 451-458.
- Goldberger, J. Roweis, S. Hinton, G. Salakhutdinov, R. 2004. Neighbourhood components analysis. *Advances in neural information processing systems* 17, MIT Press, 513-520.
- He, X. Niyogi, P. 2004. Locality preserving projections. *Advances in neural information processing systems* 15, MIT Press, 153-160.
- He, X. Cai, D. Yan, S. Zhang, H.J. 2005. Neighborhood preserving embedding. *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, 2, 1208-1213.

- Hinton, G. E. and Salakhutdinov, R. R. 2006. Reducing the dimensionality of data with neural networks. *Science*, 313(5786), 504-507.
- Hinton, G. E. Osindero, S. Teh, Y.W. 2006. A fast learning algorithm for deep belief nets. *Neural computation*, 18(7), 1527-1554.
- Hsu, C.W. and Lin, C.J. 2002. A comparison of methods for multiclass support vector machines. *Neural Networks, IEEE Transactions on*, 13(2), 415-425.
- Huang, C. Davis, L. Townshend, J. 2002. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, 23(4), 725-749.
- Johnson, R. A. and Wichern, D. W. 2007. *Applied multivariate statistical analysis*. Pearson Prentice Hall, 773, NJ ABD.
- Kozal, A. Teke, M. Ilgin, H. 2013a. Comparative analysis of hyperspectral dimension reduction methods. *Signal Processing and Communications Applications Conference (SIU) 21st*, 2013, 1-4.
- Kozal, A. Ö. Teke, M. Ilgin, H. A. 2013b. The Effect of Sub-sampling on Hyperspectral Dimension Reduction. *Nostradamus 2013: Prediction, Modeling and Analysis of Complex Systems*, 529-537.
- Lafon, S. and Lee, A. B. 2006. Diffusion maps and coarse-graining: A unified framework for dimensionality reduction, graph partitioning, and data set parameterization. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(9), 1393-1403.
- Landgrebe, D. 2003. *Signal Theory Methods in Multispectral Remote Sensing*. 2003. Wiley Hoboken, 528, NJ, ABD.
- Levina, E. and Bickel, P. J. 2004. Maximum likelihood estimation of intrinsic dimension. *Advances in neural information processing systems 17*, MIT Press, 777-784.
- Melgani, F. and Bruzzone, L. 2002. Support vector machines for classification of hyperspectral remote-sensing images. *Geoscience and Remote Sensing Symposium, 2002. IGARSS'02. 2002 IEEE International*, 1, 506-508.
- Rodarmel, C. and Shan, J. 2002. Principal component analysis for hyperspectral image classification. *Surveying and Land Information Science*, 62(2), 115-122.
- Roweis, S. 1998. EM algorithms for PCA and SPCA. *Advances in neural information processing systems 10*, MIT Press, 626-632.
- Rubin, D. B. and Thayer, D. T. 1982. EM algorithms for ML factor analysis. *Psychometrika*, 47(1), 69-76.
- Schowengerdt, R. A. 2006. *Remote sensing: Models and Methods for Image Processing*. Academic Press, 560, ABD.
- Shaw, G. A. and Burke, H. K. 2003. Spectral imaging for remote sensing. *Lincoln Laboratory Journal*, 14(1), 3-28.
- Shippert, P. 2003. Introduction to hyperspectral image analysis. *Online Journal of Space Communication*, 3.

- Tenenbaum, J. B. 1998. Mapping a manifold of perceptual observations. *Advances in neural information processing systems* 10, MIT Press, 682-688.
- Theodoridis, S. Pikrakis, A. Koutroumbas, K. and Cavouras, D. 2010. *Introduction to Pattern Recognition: A Matlab Approach*. Academic Press, 231, ABD.
- Tipping, M. E. and Bishop, C. M. 1999. Probabilistic principal component analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(3), 611-622.
- Van der Maaten, L. 2007. *An introduction to dimensionality reduction using matlab*. Report MICC 07-07, Universiteit Maastricht.
- Van der Maaten, L. Postma, E. and Van Den Herik, H. 2009. Dimensionality reduction: A comparative review. *Journal of Machine Learning Research*, 10, 1-41.
- Weinberger, K. Q. Saul, L. K. 2006. Unsupervised learning of image manifolds by semidefinite programming. *International Journal of Computer Vision*, 70(1), 77-90.

## ÖZGEÇMİŞ

Adı Soyadı : Ali Ömer KOZAL

Doğum Yeri : Denizli

Doğum Tarihi : 12.03.1978

Medeni Hali : Evli

Yabancı Dili : İngilizce

### Eğitim Durumu (Kurum ve Yıl)

Lise : Bursa Fen Lisesi (1996)

Lisans : Orta Doğu Teknik Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü (2001)

Yüksek Lisans : Ankara Üniversitesi Fen Bilimleri Enstitüsü, Elektrik-Elektronik Mühendisliği Anabilim Dalı (Eylül 2008 – Mayıs 2014)

### Çalıştığı Kurumlar

TÜBİTAK UZAY, Uydu Teknolojileri Bölümü (2006 - )

BEKO Elektronik AŞ, Ar-Ge (2004 - 2006)

Vestel Komünikasyon AŞ, Dijital Ar-Ge (2001 - 2003)

### Uluslararası Kongreler

Kozal, A. Ö. Teke, M. Ilgın, H. A. 2013. The Effect of Sub-sampling on Hyperspectral Dimension Reduction. Nostradamus 2013: Prediction, Modeling and Analysis of Complex Systems, 529-537.

Durna, M. Atar, O. Ceylan, M. Cakmakci, Y. Demirci, M. Kozal, A. Ö. Oturak, M. Ozdemir, A. Turhan, O. 2011. On Board Data Handling Subsystem Featuring BiLGE. Recent Advances in Space Technologies (RAST), 2011 5th International Conference on, 932 - 937

Durna, M. Urhan, H. Turhan, O. Kozal, Ö. Gürün, M. 2007. A New Generation On-Board Computer and Solid State Data Recorder suitable for SpaceWire Platforms, Recent Advances in Space Technologies (RAST), 2007. 3rd International Conference on, 429 - 432

### Ulusal Kongreler

Kozal, A. Teke, M. Ilgın, H. 2013. Comparative analysis of hyperspectral dimension reduction methods. Signal Processing and Communications Applications Conference (SIU) 21st, 2013, 1-4.