

**T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

HİPERSPEKTRAL GÖRÜNTÜLERDE BOYUT İNDİRGEME

BURAK AKGÜL

**YÜKSEK LİSANS TEZİ
ELEKTRONİK VE HABERLEŞME MÜHENDİSLİĞİ ANABİLİM DALI
HABERLEŞME PROGRAMI**

**DANIŞMAN
PROF. DR. ABDULLAH BAL**

İSTANBUL, 2015

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

HİPERSPEKTRAL GÖRÜNTÜLERDE BOYUT İNDİRGEME

Burak AKGÜL tarafından hazırlanan tez çalışması 30.11.2015 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Elektronik ve Haberleşme Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Prof. Dr. Abdullah BAL
Yıldız Teknik Üniversitesi

Jüri Üyeleri

Prof. Dr. Abdullah BAL
Yıldız Teknik Üniversitesi

Doç. Dr. Kayhan GÜLEZ
Yıldız Teknik Üniversitesi

Yrd. Doç. Dr. Yusuf ŞAHİN
Marmara Üniversitesi

Bu tez çalışması, 112E207 nolu “Hiperspektral Görüntülerde Uzamsal ve Spektral Öznitelikler Kullanarak Yüksek Hızda Hedef Tespiti ve Sınıflandırma Yapan Algoritmaların Geliştirilmesi” isimli TÜBİTAK projesi kapsamında desteklenmiştir.

ÖNSÖZ

Öncelikle, görüntü işleme alanıyla tanışmamı sağlayan ve çalışmalarımın her aşamasında ilgi ve tecrübesini yanımda hissettiğim tez danışmanım Prof. Dr. Abdullah Bal'a değerli katkılarından ve desteklerinden dolayı şükranlarımı sunuyorum.

Çalışmalarım boyunca değerli zamanını ve tecrübesini benden esirgemeyerek çalışmalarına büyük katkı sağlayan kıymetli hocam Doç. Dr. Olcay Kurşun'a desteklerinden dolayı teşekkürü bir borç bilirim.

Tezimle ilgili yaşadığım tüm sıkıntılarda arkamda duran ve beni daima motive eden kıymetli eşime, tezimi tamamlamam konusunda heyecanımı benimle paylaşan anne ve babama tüm kalbimle teşekkür ediyorum.

Kasım, 2015

Burak AKGÜL

İÇİNDEKİLER

	Sayfa
SİMGE LİSTESİ.....	vii
KISALTMA LİSTESİ.....	viii
ŞEKİL LİSTESİ.....	ix
ÇİZELGE LİSTESİ	xi
ÖZET	xii
ABSTRACT.....	xiv
BÖLÜM 1	
GİRİŞ.....	1
1.1 Literatür Özeti.....	1
1.2 Tezin Amacı.....	3
1.3 Hipotez.....	4
BÖLÜM 2	
HİPERSPEKTRAL GÖRÜNTÜLER (HSG) VE ÖZELLİKLERİ	5
2.1 Hiperspektral Görüntülerin Yapısı	5
2.2 Hiperspektral Görüntüleme Sistemi ve Spektral İmza	5
2.3 Hiperspektral Görüntüleme Kullanım Alanları	8
BÖLÜM 3	
BOYUT İNDİRGEME VE SINIFLANDIRMA YÖNTEMLERİ.....	10
3.1 Özellik Çıkarımı	10
3.1.1 Temel Bileşen Analizi	10
3.2 Band Seçimi	11
3.2.1 Ayrık Temel Bileşen Analizi	11
3.3 Eğitici Öğrenme	12

3.3.1	K En Yakın Komşuluk.....	13
3.3.2	Destek Vektör Makineleri	13
3.4	Eğitici-siz Öğrenme	15
3.4.1	K-Ortalama Kümeleme	15
3.4.2	Hiyerarşik Kümeleme.....	16
3.5	Ayarlanmış Rasgelelik Endeksi (Adjusted Rand Index) ile Kümeleme Performansı Ölçümü.....	18
3.6	Karışıklık Matrisi ile Sınıflandırma Performansı Ölçümü	18
BÖLÜM 4		
FARKLI VERİ SETLERİ İLE BOYUT İNDİRGEME VE SINIFLANDIRMA		21
4.1	Yöntem	21
4.2	AVIRIS - Indian Pines Görüntülerinde Boyut İndirgeme ve Sınıflandırma ..	22
4.2.1	Doğrudan Sınıflandırma Uygulanması	24
4.2.2	Yerel Maksimumları Seçilmiş Ayrık Temel Bileşen Analizi Uygulanması	27
4.2.3	Temel Bileşen Analizi Uygulanması	31
4.2.4	Klasik Ayrık Temel Bileşen Analizi Uygulanması	35
4.2.5	Gruplama Uygulanmış Ayrık Temel Bileşen Analizi Uygulanması	36
4.2.6	Sonuçların Karşılaştırılması	40
4.3	AVIRIS - KSC Görüntülerinde Boyut İndirgeme ve Sınıflandırma	48
BÖLÜM 5		
SONUÇ VE ÖNERİLER		55
KAYNAKLAR		58
ÖZGEÇMİŞ		60

SİMGE LİSTESİ

λ	Genelleştirilmiş özdeğerler
β	Yakınsama parametresi
α	İlk temel bileşenler
X	Giriş verisi
U	Birim matris
V	Birim özellikleri tutan matris
D	Tekil değerlerin tutulduğu köşegen matris
w	Hiperdüzlemin normali (ağırlık vektörü)
a	Aynı kümeye ait piksel çiftlerinin sayısı (1) verisinde aynı (2) verisinde farklı kümelere ait piksel
b	çiftlerinin sayısını (2) verisinde aynı (1) verisinde farklı kümelere ait piksel
c	çiftlerinin sayısını
d	İki veri için de farklı kümelere ait olan piksel çiftlerinin sayısı

KISALTMA LİSTESİ

ATBA	Ayrık Temel Bileşen Analizi
ATBA0	Yerel maksimumlara bakılmış Ayrık Temel Bileşen Analizi
ATBA25	Gruplama uygulanmış Ayrık Temel Bileşen Analizi(25 adım aralıklı)
ATBA40	Gruplama uygulanmış Ayrık Temel Bileşen Analizi(40 adım aralıklı)
ARE	Ayarlanmış Rasgelelik Endeksi
AVIRIS	Airborne Visible InfraRed Imaging Spectrometer
AYHB	Ağırlık yaklaşımli hiyerarşik bağlantı
DN	Doğru Negatif
DP	Doğru Pozitif
DVM	Destek Vektör Makinaları
HSG	Hiperspektral Görüntüler
HSI	Hyperspectral Imagery
IEEE	Institute of Electrical and Electronics Engineers
KATMA	Klasik Ayrık Temel Bileşen Analizi
KEYK	K en yakın komşuluk
KO	K ortalama
KSC	Kennedy Space Center
nm	nanometre
OYHB	Ortalama yaklaşımli hiyerarşik bağlantı
ROSIS	Reflective Optics System Imaging Spectrometer
TBA	Temel Bileşen Analizi
TBA0	Klasik Temel Bileşen Analizi
TBA1	Her temel bileşenin en büyüğünün seçildiği TBA
TBA2	Temel bileşenler toplamının en büyüğünün seçildiği TBA
VYHB	Varyans yaklaşımli hiyerarşik bağlantı
YN	Yanlış Negatif
YP	Yanlış Pozitif
BS	Seçilecek özellik ya da band sayısı

ŞEKİL LİSTESİ

	Sayfa
Şekil 1.1 Hiperspektral veri küpü	2
Şekil 1.2 Hiperspektral görüntü özellikleri [3]	2
Şekil 1.3 Hiperspektral Görüntü İşleme Adımları	3
Şekil 2.1 Elektromanyetik Dalga Spektrumu	6
Şekil 3.1 Gauss Çekirdeğine Sahip DVM	14
Şekil 3.2 Destek Vektör Makinesi.....	14
Şekil 3.3 Hiyerarşik Kümeleme Dendogramı.....	17
Şekil 3.4 Farklı Hiyerarşik Bağlantı Yöntemleri İle Kümeleme [16]	18
Şekil 4.1 Indian Pines örnek band görüntüsü ve doğruluk verisi bilgisi.....	22
Şekil 4.2 Indian Pines sınıflarına ait örnek spektral imzalar.....	24
Şekil 4.3 Orijinal veri seti kümeleme sonuçları	25
Şekil 4.4 Orijinal veri setine DVM uygulanması ile oluşturulan karışıklık matrisi	26
Şekil 4.5 Orijinal veri seti karışıklık matrisi performans sonuçları	26
Şekil 4.6 Orijinal veri seti sınıflarının F1_skor değerleri.....	27
Şekil 4.7 Indian_Pines Datası Ayrıklık katsayıları ve yerel maksimum noktaları	28
Şekil 4.8 Yerel maksimumları seçilmiş ATBA uygulanması sonucu kümeleme başarısı .	28
Şekil 4.9 Yerel maksimumları seçilmiş ATBA uygulanması sonucu karışıklık matrisi.....	29
Şekil 4.10 Yerel maksimumları seçilmiş ATBA uygulanması sonrası karışıklık matrisi performans değerleri	29
Şekil 4.11 Yerel maksimumları seçilmiş ATBA uygulanması sonucu sınıfların F1_skor değerleri	30
Şekil 4.12 k=1, 3, 5 için KEYK sınıflandırma sonuçlarının karşılaştırılması (Orijinal - Yerel maksimumları seçilmiş ATBA için)	30
Şekil 4.13 Klasik TBA(BS=3) uygulanması sonucu kümeleme başarısı	31
Şekil 4.14 Klasik TBA(BS=3) uygulanması sonucu karışıklık matrisi	32
Şekil 4.15 Klasik TBA(BS=3) uygulanması sonrası karışıklık matrisi performans değerleri	32
Şekil 4.16 Klasik TBA(BS=3) uygulanması sonucu sınıfların F1_skor değerleri	33
Şekil 4.17 TBA ile band seçilmesi sonucu kümeleme başarısı (a)TBA1 (b)TBA2.....	34
Şekil 4.18 BS=3 için TBA ile band seçilmesi sonrası F1_skor değerleri (a) TBA1 (b) TBA2	35
Şekil 4.19 Klasik ATBA uygulanması sonucu kümeleme başarıları	36
Şekil 4.20 Gruplama uygulanmış 25 adım aralıklı ATBA için kümeleme sonuçları	37
Şekil 4.21 Gruplama uygulanmış 40 adım aralıklı ATBA için kümeleme başarısı	37

Şekil 4.22 BS=3 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) VYHB (c) OYHB (d) AYHB (e) KEYK (f) DVM (devamı).....	40
Şekil 4.23 BS=5 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) VYHB (c) OYHB (d) AYHB (e) KEYK (f) DVM (devamı).....	43
Şekil 4.24 BS=10 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) VYHB (c) OYHB (d) AYHB (e) KEYK (f) DVM (devamı).....	46
Şekil 4.25 KSC verisine ait band görüntüsü ve doğruluk verisi	48
Şekil 4.26 KSC verisine ait spektral imzalar	49
Şekil 4.27 Boyut sınırlaması olmadan uygulanmış sınıflandırıcı sonuçlarının karşılaştırılması (a) KO (b) KEYK	50
Şekil 4.28 BS=3 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) KEYK	51
Şekil 4.29 BS=5 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) KEYK	52
Şekil 4.30 BS=10 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) KEYK	53

ÇİZELGE LİSTESİ

	Sayfa
Çizelge 2.1 Multispektral Sensör Karakteristikleri [5]	7
Çizelge 2.2 Hiperspektral Sensör Karakteristikleri [5]	7
Çizelge 3.1 İkili sınıflandırma için karışıklık matrisi	19
Çizelge 3.2 Çok sınıflı örnek için karışıklık matrisi	19
Çizelge 3.3 Sınıflandırma performans kıstasları ve hesaplamaları [20]	20
Çizelge 4.1 Indian Pines görüntüsünde kullanılan sınıflar ve örnek sayıları [21]	23
Çizelge 4.2 Indian Pines verisi için karşılaştırmalı sonuçlar	47
Çizelge 4.3 Sınıflandırma işlem sürelerinin karşılaştırılması	48
Çizelge 4.4 KSC verisi için karşılaştırmalı sonuçlar	54

HİPERSPEKTRAL GÖRÜNTÜLERDE BOYUT İNDİRGEME

Burak AKGÜL

Elektronik ve Haberleşme Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Tez Danışmanı: Prof. Dr. Abdullah BAL

Gelişen sayısal görüntüleme teknolojisi, günümüzde birçok alanda hayatı kolaylaştırmanın yanında bir ihtiyaç haline gelmiştir. Teknolojinin gelişmesi ile birlikte görüntüleme teknolojileri de ilerlemiş, hiperspektral görüntüleme gibi ileri seviyede görüntüleme imkânları oluşmuştur.

Hiperspektral görüntülerin kullanılmasıyla birlikte spektral düzlemde dar band aralıklarına sahip çok büyük verilerle işlemler yapılmaya başlanmış, sınıflandırma başarımlarının yükselmesine karşılık yüksek hesaplama maliyetleri ortaya çıkmıştır. Bunun için de boyut indirgeme ihtiyacı doğmuştur.

Bu çalışmada farklı boyut indirgeme yöntemleri hiperspektral veri setlerine uygulanmış ve farklı sınıflandırıcılar kullanılarak sınıflandırma başarımları analiz edilmiştir. Temel Bileşen Analizi ve Ayrık Temel Bileşen Analizi yöntemleri ile özellik çıkarımı ya da band seçimi yoluyla boyut indirgeme yapılmış ardından, indirgenmiş veriler çeşitli eğitici ve eğitici olmayan öğrenme yöntemleri ile eğitilerek sınıflandırılmışlardır. Öğrenme yöntemi olarak k En Yakın Komşuluk, Destek Vektör Makineleri, K-Ortalama ve Hiyerarşik Kümeleme yöntemleri seçilmiştir.

Bu çalışmada Ayrık Temel Bileşen Analizi yönteminin sonucunu iyileştirme hedeflenmiş, buradan yola çıkılarak bantları gruplama yoluna gidilmiştir. Klasik yöntemden farklı olarak ayrıklık katsayılarının tepe yaptığı yerel maksimum noktalara bakılmış ve bu noktalar belli adım aralıkları ile gruplandırılarak band aralıkları içerisinden seçilmiştir.

Band seçme işlemi 3, 5 ve 10 band seçimi için denenmiştir. Eğitici sınıflandırıcı sonuçlarının başarıları karışıklık matrisi yardımıyla belirlenmiştir. Tüm sistem başarısı için doğruluk hesabı yapılmış, ayrıca kesinlik, geri çağırım değerlerinden F1_skor değeri hesaplanarak sınıf başarıları hesaplanmıştır. Kümeleme sonuçları için ise Ayarlanmış Rasgelelik Endeksi yöntemi ile performans ölçümü yapılmıştır.

Tüm boyut indirgeme ve sınıflandırma işlemleri AVIRIS sensörü ile alınmış Indian Pines ve KSC görüntülerine uygulanmıştır. Ayrık Temel Bileşen Analizine yapılan farklı yaklaşımın sonuçlara olumlu etki ettiği her iki veri setinde de görülmüştür.

Anahtar Kelimeler: Hiperspektral Görüntüler, Boyut İndirgeme, Band Seçme, Ayrık Temel Bileşen Analizi

DIMENSION REDUCTION IN HYPERSPECTRAL IMAGES

Burak AKGÜL

Department of Electronics and Communications Engineering

MSc. Thesis

Advisor: Prof. Dr. Abdullah BAL

Besides making life easier in many areas, developed digital imaging technology has become a necessity. By the development of technology, imaging technology has also forwarded and advanced imaging opportunities, like hyperspectral imaging, have occurred.

By the use of hyperspectral imaging, while large sized data with narrow spectral bands has been processed and the success of classification has increased, calculation costs has started to increase. The result was the need to dimension reduction.

In this study, different dimensionality reduction methods have been applied to hyperspectral data sets and the classification performance on different classifiers is analyzed. Using PCA and Sparse PCA methods, feature extraction or band selection was applied to reduce dimension and then classified with different types of classifiers. K-nn, SVM, K-means and hierarchical clustering methods has been chosen as learning methods.

In this study, the bands are grouped in order to rehabilitate of the results of the Sparse PCA. Unlike classical SPCA method, local maximum points where coefficient of discreteness peaked have been observed. Those points were chosen from band groups that were grouped specific range.

Band selection process was applied for 3, 5 and 10 bands. Confusion matrix was used to measure the results of supervised classifiers. Accuracy calculation was implemented for the whole system. Also the success of each class was calculated by calculating F1_

scores. Adjusted Rand Index method was used for the performance of measurement of unsupervised classifiers.

All dimension reduction and classification processes has been applied to AVIRIS Indian Pines and KSC data sets. It was seen that grouping approach to Sparse PCA has yield successful results.

Key words: Hyperspectral images, Dimension Reduction, Band Selection, Sparse PCA

1.1 Literatür Özeti

Hiperspektral görüntüleme, uzaktan algılama alanında hızla ilerleyen bir teknolojidir. Son yıllarda bu alanda birçok çalışma yer almıştır. Dünya yüzeyini inceleyen jeodezi bilimi, savunma sanayinde görüntüleme vb. alanlarda bizlere geniş bir algı imkânı sunmuştur. Hiperspektral görüntüleme, görüntünün elektromanyetik spektrumunu incelemektedir. Hedeften yansıyan elektromanyetik ışımının kamera tarafından ölçülerek hiperspektral karakteristiğinin çıkarılması işlemi yapılmaktadır. Bilinen görüntülerde yer alan ve x-y koordinat bilgisini içeren uzamsal (spatial) bilginin yanı sıra z-ekseninde spektral bilgi de yer almaktadır. Görüntüde yer alan her bir piksel bir spektral imzaya karşılık düşmektedir.

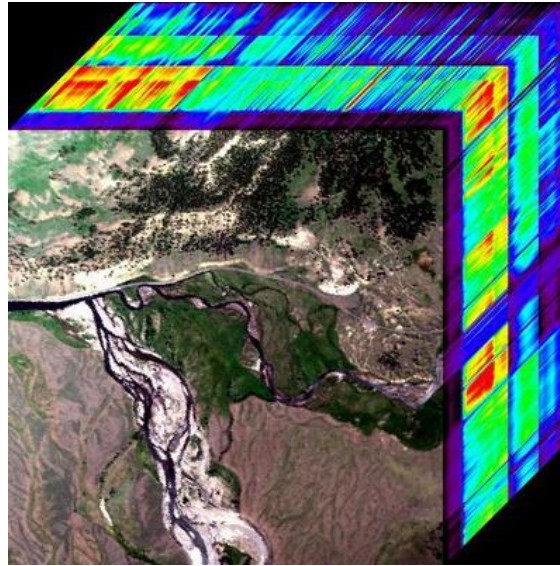
Hiperspektral görüntülemeyi diğer görüntüleme işlemlerinden ayıran özelliği spektral bandlarıdır. UV bandlardan termal bandlara kadar farklı dalga boylarını içermektedir. Bandlara ait dalga boylarının dar olması, spektral çözünürlüğün iyi olduğu anlamına gelmektedir. Buradaki çözünürlük sayesinde nesneler daha detaylı ayırt edilebilmekte, hedeflerin keşfedilebilmesi imkânı artmaktadır. Hiperspektral yaygın olarak atmosferik olaylarda, jeolojik incelemelerde, zirai haritalama uygulamalarında ve savunma sanayinde kullanılmaktadır.

Hiperspektral veri küpleri spektral düzlemde getirdiği yüksek boyut ile görüntü işleme alanında fazlaca hesaplama maliyeti katmıştır. Eldeki verinin azaltılması ilgilenilmesi gereken farklı bir konu olarak ortaya çıkmıştır. Hiperspektral görüntülerde de diğer görüntülerde olduğu gibi veri belirli bir yöntemle farklı düzleme iz düşürülerek

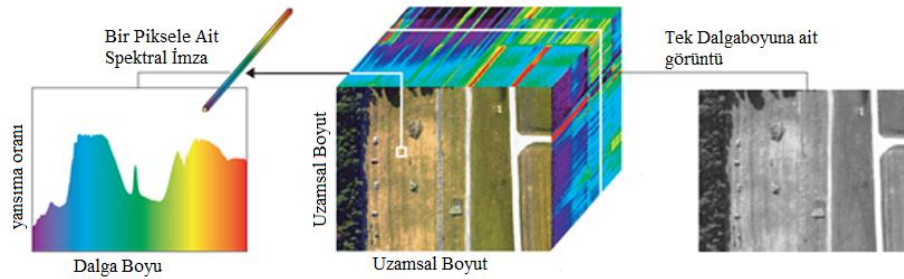
görüntünün varyansını büyük oranda karşılayan daha küçük miktarlara ulaşabilmektedir.

Hiperspektral verilerin analizi çok kolay olamamaktadır. Bunun nedeni ise atmosferik bozulmalar, yüksek boyut problemi, yüzey üzerinde olabilecek spektral imza çeşitliliği gibi faktörlerdir.

Boyut indirgeme ile ilgili çalışmalar yüksek boyutlu verilerin analizi için önemli bir aşamadır. Yüksek boyutlardaki verilerde veriler son derece seyrek olabilmektedirler. İlgisiz, tekrarlayan ya da gürültü içeren özelliklerin elenmesi boyutun azaltılmasında etkili olabilmektedir [1].



Şekil 1.1 Hiperspektral veri küpü [2]



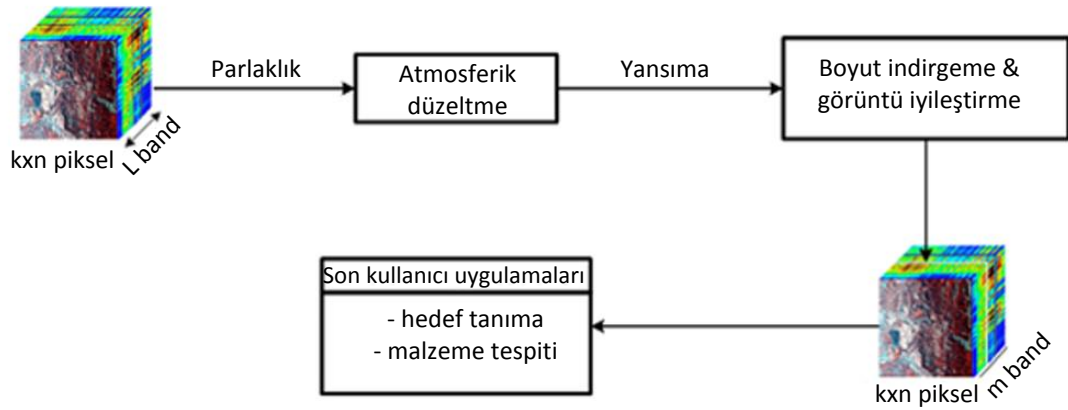
Şekil 1.2 Hiperspektral görüntü özellikleri [3]

Uzaktan algılama yöntemlerinde boyut indirgeme yapılırken özellik çıkarımı ve özellik seçimi olmak üzere iki farklı yaklaşım vardır. Özellik çıkarımında orijinal veri seti farklı bir düzleme geçirilerek daha düşük boyutta bir veri setine dönüştürülmektedir. Özellik

seiminde ise korelasyon oranı yksek, gereksiz ve grltl bilgiler elenmiř ve nemli bilgileri ierdiėi dřnlen zellikler seilerek diėerlerinden ayrılmaktadırlar.

Hiperspektral verilerin iřlenmesi genel olarak ařaėıdaki adımlarla tanımlanabilmektedir:

- Toplanan verinin yansıma ve ıřıma deėerlerinin dnřtrlmesi,
- Atmosferik dzeltme,
- Boyut indirgeme,
- İstenen veriye ait spektral imzanın belirlenmesi,
- Seilen imza rnekleri ile sınıflandırıcının eėitilmesi.



řekil 1.3 Hiperspektral Grnt İřleme Adımları

1.2 Tezin Amacı

Grnt iřleme, sinyal iřleme alanında nemli bir bařlıktır. İlerleyen teknolojiyle birlikte artan iřlemci hızları yeni grnt iřleme tekniklerini mmkn kılmaktadır. Birok alanda grntleme sistemleri kullanıldıėı iin bu grntler zerinden bilgi ıkartımı hayati nem tařımaktadır. Otomasyon alanında bilgisayarlı grme konusu son zamanlarda yoėun ilgilenilen bir alandır. Bununla birlikte endstriyel uygulamalarda hatalı rn tespiti, bilgisayarlı modelleme, sınırda hedef tespiti gibi birok sistemde grnt iřleme kullanılmaktadır.

Geleneksel grntleme sistemleri bazı uygulamalar iin yetersiz kalmaktadır. Bunun iin yeni nesil grntleme cihazları geliřtirilmiřtir. Bunlardan en yaygın kullanılanı

hiperspektral kameralardır. Yeni görüntüleme sistemlerinin geliştirilmesi ile birlikte yeni görüntü işleme algoritmalarının da oluşturulması kaçınılmazdır. Özellikle iyi spektral çözünürlüklere sahip görüntülerin büyük boyutlara sahip olması yeni algoritmaların geliştirilmesini tetiklemiştir. Bu görüntülerin işlenmesi için boyut indirgeme algoritmaları önemli yer tutmaktadır. Literatürde pek çok boyut indirgeme algoritması mevcuttur. Bu tezde klasik olan TBA gibi özellik çıkarım yöntemlerinin yanı sıra Ayrık Temel Bileşen Analizi yöntemini de kullanarak özellik seçimi uygulanmıştır.

1.3 Hipotez

Bu tezde, hiperspektral görüntülerde boyut indirgeme yöntemi olarak Ayrık Temel Bileşen Analizi (ATBA) önerilmiştir. Klasik Ayrık Temel Bileşen Analizi yönteminde veriye TBA uygulandıktan sonra seçilecek band sayısı dışındakilerin katsayıları sıfır olacak şekilde bandlara ayrıklık katsayıları verilmektedir [4]. Klasik ATBA'dan farklı olarak, bandların özelliklerine göre onlara ağırlıklar vererek birbirlerinden ayrılmasını sağlayan bu yöntemin uygulamasında grupta gibi farklı yaklaşımlar önerilmiştir. Önerilen yöntem ile birlikte farklı sınıflandırma yöntemleri uygulanarak daha başarılı sonuçlar elde edilmesi beklenmiştir.

HİPERSPEKTRAL GÖRÜNTÜLER (HSG) VE ÖZELLİKLERİ

2.1 Hiperspektral Görüntülerin Yapısı

Spektral görüntüler uydular ya da havayolu araçları ile toplanabilmektedirler. Görüntü düz şeritler halinde taramalar sonucu x-y düzleminde veri elde edilmesiyle oluşmaktadır.

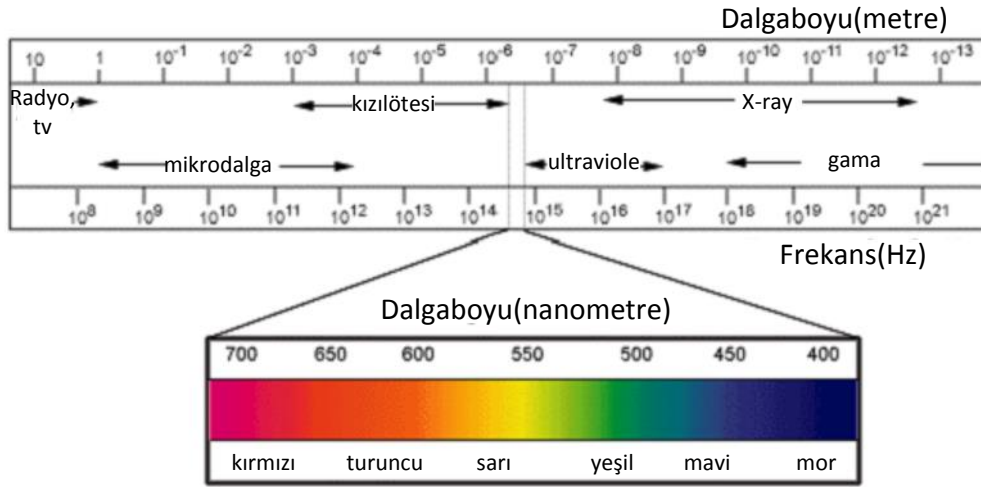
Hiperspektral sensörler elektromanyetik yansımayı ölçen aletlerdir. Farklı dalga boylarından gelen yansımalar veri olarak elimize ulaşırlar. Hiperspektral sensörler, her bir piksel için birbirlerine bitişik şekilde dar dalga boyu aralığına sahip bandların parlaklık değerini ölçerek detaylı bir spektral imza elde ederler. Yeterli spektral özelliğin elde edilebilmesi için spektral bandların dalga boyu aralıklarının yeterince dar olabilmesi gerekmektedir. Bu aralık yaklaşık 1-10 nm değerlerindedir.

2.2 Hiperspektral Görüntüleme Sistemi ve Spektral İmza

Uzaktan algılama, bir nesne ile ilgili bilgiyi o nesneye temas etmeden elde etme teknolojisidir. İstenilen hedefin atmosfer, uzay, uçak vb. platformlardan gelişmiş görüntüleme sistemleri yardımı ile alınarak analizini ve tespitini içerir. Bunun yanında hedef ile herhangi bir uzaklıktan yapılan ölçümler ile bilgi edinme bilimi olarak da ifade edilebilmektedir.

Kapsadığı spektruma göre uzaktan algılama üç temel başlığa ayrılır. Bunlar görünür ve kızılötesi dalga boyunda uzaktan algılama, termal uzaktan algılama ve mikrodalga uzaktan algılamadır. Hiperspektral sensörler elektromanyetik spektrumda UV, görünür,

kızılötesi ve mikrodalga aralığını içeren kısmı ölçebilmektedir. Dalga boyu aralığı ve çözünürlük ise sensörlerin özelliklerine göre değişkenlik göstermektedir.



Şekil 2.1 Elektromanyetik Dalga Spektrumu

Hiperspektral görüntüleme yüzey cisimlerinden yansımanın bitişik dar dalga boyları seviyesinde ölçümüdür. Hiperspektral; çok sayıda spektral banttandır. Hiperspektral sensörler multispektral sensörler ile kıyaslandıklarında band sayısında büyük artış olduğu görülmektedir. Ancak hiperspektral sensörlerin en önemli özelliği band sayısı değil band çözünürlüğüdür. Cisimlerden yansıyan veya yayılan ışığın farklı dalga boylarına dağılımlarının ölçülebilmesi ile bu teknoloji ortaya çıkmıştır. Yüzeyden gelen enerjinin algılanmasında her bir bant için ayrı bir algılayıcı kullanılmakta, algılayıcı sayısı ile orantılı olarak da ölçülebilen dalga boyu aralığı genişlemektedir. Hiperspektral kamera ile elde edilen görüntülerde iki boyut uzamsal, üçüncü boyut ise spektral bilgi içermektedir. Üç boyutlu hiperspektral veri Hiperküp olarak da adlandırılabilir. Normal koşullarda, bir cisme ulaşan ışığın bir kısmı yansımakta bir kısmı ise emilmektedir. Işığın yansıyan miktarı 'yansıma' (reflektans) olarak ifade edilmektedir. Nesneden yansıyan ışığın farklı dalga boylarındaki enerjisinin ölçülmesi ile yansıma spektrumu elde edilir. Cisimlerin yansıma ve emilme miktarları karakteristiklerini gösterir. Bu karakteristikler 'spektral imza' olarak tanımlanmakta ve cisimlerin sınıflanabilmelerini sağlamaktadır. Çizelge 2.1 ve Çizelge 2.2'de multispektral ve hiperspektral sensörlerin karakteristikleri verilmiştir [5].

Çizelge 2.1 Multispektral Sensör Karakteristikleri [5]

Sensör	Tarih	Band Sayısı	Spektral Mesafe(nm)	Band Genişliği(nm)
Airbone Thematic Mapper ATM-Daedalus	1986	11	420-13000	420-450, 450-520, 620-600, 600-620, 630-690, 690-750, 760-900, 910-1050, 1550-1750, 2080-2350, 8500-13000
AMTIS – Multi-angle Thermal/Visible Imaging System	-	3	400-13000	400-700, 700-900 7500-13000
GER EPS-M serisi	1994	3	300-12000	300-400, 600-700 8000-12000

Çizelge 2.2 Hiperspektral Sensör Karakteristikleri [5]

Sensör	Tarih	Band Sayısı	Spektral Mesafe(nm)	Band Genişliği(nm)
AAHIS Advanced Airbone HSI Sensör	1994	288	432-832	6
AHI Airbone HSI	1994	256	7500-11700	100
AIS-1/2	1982-1985 1985-1987	128	900-2100 800-2400	9.3 10.6
AISA+	1997	244	400-970	2.9
AISA Eagle Airbone	2002	-	-	-
Aisa Hawk	2003	240	1000-2400	8
ARGUS	-	400 300	370-2500 780-2500	5-10
ASAS	1987	62	400-1060	11.5

Çizelge 2.2 Hiperspektral Sensör Karakteristikleri (devamı) [5]

Sensör	Tarih	Band Sayısı	Spektral Mesafe(nm)	Band Genişliği(nm)
MIVIS Multispectral Infrared Visible Spectrometer	1993	20 8 64 10	433-833 1150-1550 200-2500 8200-12700	20 50 800 400-500
OMNIS	1999	64 16 32 8 8	460-1100 1060-1700 200-2500 3000-5000 8000-12500	10 40 15 250 500
PROBE-1	-	100-200	440-2543	11-18
ROSIS	1993	128	440-850	5
SASI	2002	160	850-2450	10
SFSI	1994	22-120	1230-2380	10.3
VIFIS	1994	64	420-870	10.12

2.3 Hiperspektral Görüntüleme Kullanım Alanları

Hiperspektral sensörler ardışık bandlardan gelen yansımaları yüksek hızda örnekleyerek görüntüyü elde ederler. Bu örnekleme sonucu geniş frekans aralığına sahip 100 MB'tan daha büyük boyutlarda veriler ortaya çıkar. Verinin büyüklüğü yüksek hesaplama maliyetlerini de beraberinde getireceğinden analiz edilmesi de zorlaşmaktadır. Datanın büyüklüğünün yanında atmosferik etkiler ve gürültü de analiz işlemlerini zorlaştıran diğer faktörlerdir.

Hedef tanıma hiperspektral görüntülemenin başlıca ilgi alanlarındanıdır. Hedef tanıma işlemi askeri hedef tanıma, tarım alanlarında ürün tespiti gibi birçok alanda kullanılmaktadır [6]. Uzamsal çözünürlüğün düşük olduğu durumlarda bir piksel birden fazla cismi içinde barındırabilmektedir. Bu da içeride birden fazla spektral imzanın karışımı olduğu anlamına gelir ki bu da karışık (mixed) sinyal olarak adlandırılmaktadır. Buradaki hedefi bulma işlemi ise alt piksel analizi ile yapılmaktadır.

Hiperspektral görüntülemenin bir diğ er kullanım alanı da madde tanımlamadır. İlgili bölgede bulunan ve bilinmeyen maddelerin haritasını çıkarmada kullanılmaktadır. Örneğın jeologlar tarafından ekonomik değeri yüksek minerallerin haritasını çıkarmada kullanılmaktadırlar.

Yüzey tuzluluğ u, nemlilik, cismin organik içeriğinin belirlenmesi gibi konularda da spektral imzaların kullanılması oldukça etkili yöntemlerdir. Bu özelliklerin belirlenmesi yalnızca görün r bandı kullanan klasik görünt leme yöntemleri ile oldukça zordur.

BOYUT İNDİRGEME VE SINIFLANDIRMA YÖNTEMLERİ

3.1 Özellik Çıkarımı

Boyut indirgeme aşaması, hiperspektral görüntülerin yüksek boyutlarından kaynaklanan problemlerin önemli bilgilerin kaybedilmeden giderilmesini hedefleyen bir ön işlemdir. Hiperspektral görüntülerde boyut indirgeme konusunda yeni yöntemler geliştirmeye yönelik yapılan çalışmaların yanında bu yöntemlerin performanslarının analizi de ayrı bir çalışma konusu olmuştur [7]. Hiperspektral görüntüler üç boyutlu bir veri küpüdür. Bunun iki boyutu uzamsal koordinatları bir boyutu ise spektral derinliği içerir. Hiperspektral görüntülerde boyut indirgeme iki farklı yaklaşımla yapılabilmektedir. Bunlardan biri spektral düzlemdeki bandların seçilmesini kapsayan band seçme yöntemi diğeri ise özellik çıkartma yöntemidir.

Özellik çıkarım yönteminde orijinal bandlar üzerinden doğrusal ve doğrusal olmayan dönüşümler uygulanmaktadır. Spektral düzlemde indirgeme yapmak yerine uzamsal düzlemdeki verilerin farklı düzlemlere iz düşürülmesi ile boyut azaltma yoluna gidilmektedir. Özellik çıkarım yöntemlerinden en bilineni Temel Bileşen Analizi yöntemidir.

3.1.1 Temel Bileşen Analizi

Aralarında yüksek korelasyon ilişkisi olan çok değişkenli verilerin yeni bir koordinat sistemine doğrusal olarak dönüştürülmesi işlemidir. TBA mümkün olduğunca varyansı en yüksek olan verinin düşük boyutlu gösterimini oluşturur.

Temel Bileşenler Analizinde yeni değişkenler orijinal korelasyonlu değişkenlerin özvektörlerle çarpılarak yeni uzaya dönüştürülmesi ile elde edilir. Özvektörler bulunurken ise kovaryans matrisi kullanılmaktadır.

TBA'nın temel ilkesi; görüntü vektör uzayında, veriler arasındaki korelasyon ilişkisinin en aza indiği yeni bir koordinat sisteminin bulunmasıdır.

Temel Bileşenler Analizi uzaktan algılamada;

- Görüntü sıkıştırmada,
- Görüntü iyileştirmede,
- Değişim belirlemede,
- Band seçiminde kullanılmaktadır.

3.2 Band Seçimi

Hiperspektral sensörlerin geniş dalga boyu aralıklarındaki ışıma değerlerini ölçebilmesi ile birlikte çok büyük boyutlarda veriler elde edilmeye başlanmıştır. Bu çok boyutluluğun çözümü için veri kaybını en aza indirecek şekilde doğru bandların seçilerek boyut indirgeme yapılması ilgi gören çalışma alanları içerisinde.

Band seçim yöntemlerinde özellik çıkarımındaki gibi dönüşüm işlemleri yapılmaksızın görüntü özelliklerinin birbirlerinden en iyi şekilde ayırtedilebildiği bandların bulunması amaçlanır. Böylelikle orijinal bandlardan önemli olanları seçilerek veri işleme başarısını düşürmeden hesaplama yükünün azaltılması amaçlanmaktadır. Band seçim yönteminin özellik çıkarımına göre üstünlüğü uzamsal boyuttaki özelliklerin kaybedilmemesidir[8].

Hiperspektral verilerde band seçimi ile ilgili yöntemlere literatürde rastlanmıştır. Hongtao Du, Bağımsız Bileşen Analizi yaklaşımı ile band seçimi yöntemi geliştirmiştir [9]. Bunun dışında varyans yaklaşımı, Ayırık Temel Bileşen Analizi yaklaşımı gibi birçok farklı yaklaşım ile band seçimi uygulanmış ve en yüksek başarıları verecek bandlara ulaşabilmek hedeflenmiştir [10].

3.2.1 Ayırık Temel Bileşen Analizi

Temel Bileşen Analizi (TBA) görüntü işleme ve boyut indirgemedede sıklıkla kullanılan bir yöntemdir. Ancak, TBA'nın lineer bir yöntem olması ve temel bileşenlerin verinin lineer

birleşimleri olmasından dolayı sonucun analizi zor olabilmektedir. Ayırık (Sparse) TBA bu dezavantajın üstesinden gelebilmek için normal TBA'nın üzerine ayrıklık (sparsity) yaklaşımı eklemektedir. Temel Bileşenlerden bazılarının ağırlıklarını sıfıra gitmeye zorlayarak çıkan sonuçlarının analizini kolaylaştırmaktadır [11].

Genel ATBA yaklaşımı şu şekildedir:

1. TBA yöntemi uygulanır. Temel bileşenler $A = [\alpha_1, \dots, \alpha_k]$ ile temsil edilmektedir.

2. (3.1) denklemi çözülerek ayrıklık yaklaşımı uygulanır.

$$B_j = \arg \min \beta^{*T} (X^T X + \lambda) \beta^* - 2\alpha_j^T X^T X \beta^* + \lambda_{1j} |\beta^*| \quad (3.1)$$

3. Her β değeri için SVD dönüşümü ($X^T X \beta = U D V^T$) uygulanır ve $\alpha = U V^T$ güncellenir.

4. 2-3 nolu adımlar yakınsama olana kadar tekrarlanarak $V_j = \frac{B_j}{|B_j|}, j = 1, \dots, k$ normalize edilir.

Bu adımlar uygulandıktan sonra oluşan matrisin her bandda ne sıklıkla görüldüğüne bakılarak her band için ayrıklık katsayıları elde edilir [12].

3.3 Eğitici Öğrenme

Eğitim ve test işlemlerinden oluşan eğitici sınıflandırma yöntemlerinde görüntünün hangi sınıflara ayrılacağı ya da görüntüden hangi sınıfların elde edilmek istendiği önceden bilinmektedir. Eğitici öğrenme için eğitim ve test veri setleri bulunmaktadır. Eğitim aşamasında eğitim verileri kullanılarak geliştirilen model sayesinde daha önceden görülmemiş giriş verilerinin sınıfı belirlenebilmektedir. Test aşamasında eğitim verileri kesinlikle kullanılmamakta, sadece elde edilen parametreler kullanılmaktadır. Eğitici öğrenmede yaygın olarak K en yakın komşu sınıflandırıcı ve Destek Vektör Makineleri bilinmektedir. Destek Vektör Makineleri hiperspektral verilerin sınıflandırılmasında kullanılan önemli bir sınıflandırıcıdır [13].

3.3.1 K En Yakın Komşuluk

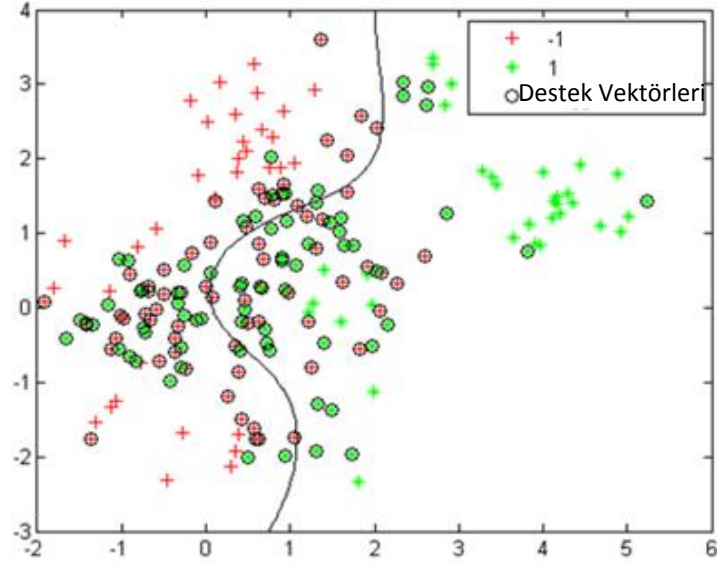
K En Yakın Komşuluk(k-nn) yöntemi, eğitici öğrenme yöntemleri arasında yer alır. Yöntemde; sınıflandırma yapılacak verilerin öğrenme kümesindeki verilere benzerliklerine bakılarak en yakın k verinin ortalamasıyla, belirlenen eşik değere göre sınıflara atamaları yapılır. Önemli olan, her bir sınıfın özelliklerinin önceden net bir şekilde belirlenmiş olmasıdır. Yöntemin performansını k en yakın komşu sayısı, eşik değeri, benzerlik kriteri ve öğrenme kümesindeki eğitim verilerinin sayısı belirlemektedir.

3.3.2 Destek Vektör Makineleri

Destek Vektör Makineleri(DVM) son zamanlarda hiperspektral görüntülerde regresyon ve sınıflandırma problemlerini çözmek için uygulanmıştır [14]. DVM sınıflandırmada yüksek başarı elde etmek için yüksek boyut özelliklerine sahip çekirdek fonksiyonları kullanılmaktadır. DVM lineer olmayan bir yöntemdir. Karar ağacı, maksimum benzetim, yapay sinir ağları gibi sınıflandırma yöntemleri ile karşılaştırıldığında DVM'in daha başarılı olduğu söylenebilmektedir.

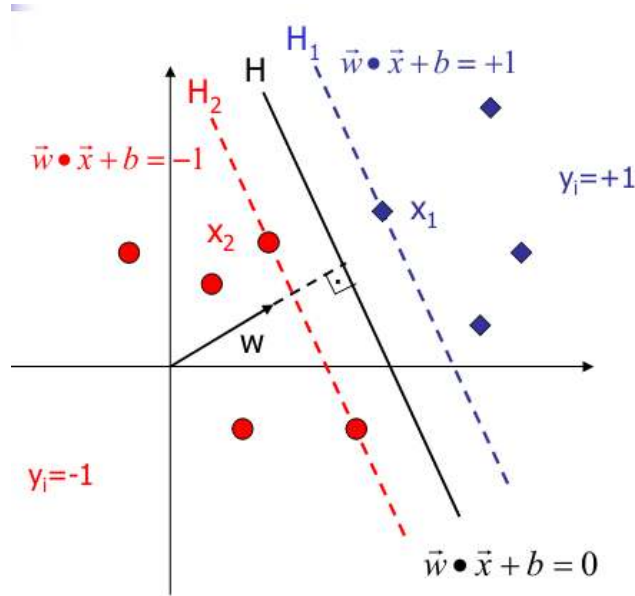
DVM'nin dayandığı temel fikir, öznelik uzayın içerisinde iki sınıfı en uygun olarak ayırabilecek bir hiperdüzlem bulmaktır. DVM'in diğer doğrusal öğrenme yöntemlerinden farkı, test verisi için yanlış sınıflandırma olasılığını en aza indirecek çözümler sunmasıdır [15].

DVM'de ilk olarak veri uygun bir çekirdek fonksiyonu yardımıyla öznelik uzayına taşınır. Sonrasında bir hiperdüzlem yardımıyla orijinden ayrılır. Şekil 3.1'de Gauss çekirdek(kernel) için DVM örneği verilmiştir.



Şekil 3.1 Gauss Çekirdeğine Sahip DVM

Destek Vektör Makinası yönteminde sistemdeki verilerin dağılımına bakılarak uygun bir çekirdek fonksiyonu seçilir. Şekil 3.2’de lineer çekirdeğe sahip DVM verilmiştir.



Şekil 3.2 Destek Vektör Makinesi

$$(w * x) + b \geq +1, y_i = +1, (w * x) + b \leq -1, y_i = -1 \quad (3.2)$$

$$y_i(w * x + b) \geq +1 \quad (3.3)$$

$$(w * x_1) + b = y_i = +1, (w * x_2) + b = y_i = -1 \quad (3.4)$$

$$w * (x_1 - x_2) = 2 \Rightarrow \frac{w}{||w||} * (x_1 - x_2) = \frac{2}{||w||} \quad (3.5)$$

Hiperdüzlemin her iki tarafa da en uzak mesafede olabilmesi amaçlanmaktadır. DVM ile sınıflar belirlendikten sonra sınıflandırmak istenilen herhangi bir vektör seçilir ve düzlemin hangi tarafında kalıyorsa $\{-1,+1\}$ olarak sınıflandırılır.

3.4 Eğiticişiz Öğrenme

Eğiticişiz öğrenme, piksellerin, kullanıcı müdahalesi olmadan algoritmalar yardımı ile otomatik olarak kümelendirilmesi mantığına dayanmaktadır. Eğiticişiz öğrenmede, sınıflandırılacak bölgenin tüm pikselleri kullanılarak küme grupları elde edilir. Görüntüdeki verinin tanımlanamadığı ya da yeterli veri bulunamadığı zamanlarda bu yöntemle başvurulur.

Eğiticişiz öğrenme yöntemleri içinde en yaygın olanı K-ortalama (K-means) kümeleme yöntemidir. Bu öğrenme yönteminde, her bir kümenin ortalaması hesaplanmaktadır. Piksel değerlerinin her bir küme ortalamasına uzaklığı bulunmakta ve buna göre en yakın kümeye yerleştirilmektedirler. Öğrenme için oluşturulan veri kümelerinin sayısına göre kümeler oluşturulmakta ve her kümenin ortalaması alınarak işlemler tekrar edilmektedir.

3.4.1 K-Ortalama Kümeleme

K-Ortalama yöntemi, kümeleme problemini çözen en basit eğiticişiz öğrenme yöntemleri arasında yer alır. Algoritmanın genel mantığı n adet veriden oluşan bir veri kümesini, daha önceden belirlenmiş olan k adet kümeye bölümlenektir. Amaç, gerçekleştirilen bölümlenme işlemi sonunda, küme içi benzerlikler maksimuma çıkarken, kümeler arası benzerliklerinin minimuma düşmesidir. Yöntemin performansını k küme sayısı, başlangıç olarak seçilen küme merkez değerleri ve benzerlik ölçümü kriterleri etkilemektedir.

K-ortalama algoritmasının işlem basamakları şu şekildedir:

- İlk olarak küme merkezleri belirlenir.
- Her verinin seçilen merkez noktaya uzaklığı hesaplanır. Elde edilen verilere göre en yakın kümeye yerleştirme işlemi yapılır.
- Oluşan kümelerin yeni merkez değerleri hesaplanır ve ortalamalar güncellenir.

- Merkez noktalar deęiřmeye kadar bu adımlar tekrarlanır.

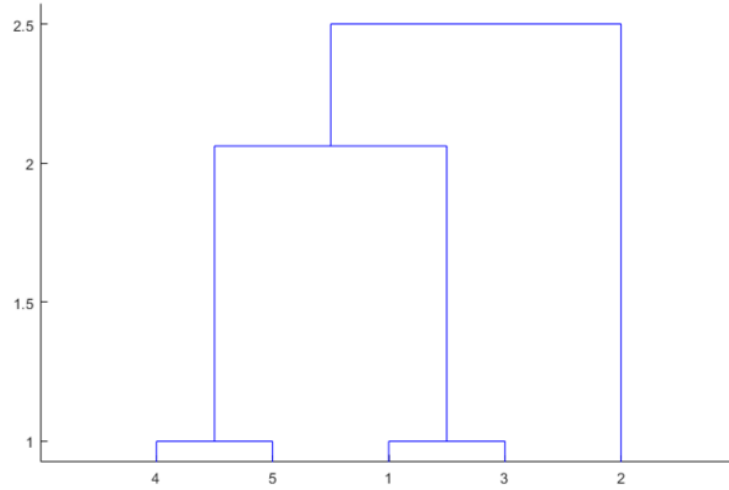
K-ortalama algoritmasının en büyük dezavantajı sınıf sayısını bilememesidir. Bu nedenle başarılı bir sonuç elde edebilmek için farklı K deęerleri ile denemeler yapılması gerekmektedir [16].

3.4.2 Hiyerarşik Kümeleme

Hiyerarşik kümeleme yöntemleri örnekleri birbirleri ile farklı aşamalarda bir araya getirerek ardışık kümeler belirleme temeline dayanır. Belirlenen uzaklık kıstasları ile kümeler belirlenir ve hangi elemanların yerleştirileceğine karar verilir. Hiyerarşik kümelemede, örneklerin birbirlerine yakınlıklarına göre hiyerarşileri çıkarılır. Yığılmalı (agglomerative) ve bölünmeli (divisive) olmak üzere iki yaklaşıma sahiptir. Yığılmalı yaklaşımda, başlangıçta örneklerin her birini bir küme olarak kabul edip her bir adımında kümelerin birbirlerine uzaklığını hesaplayıp o anda birbirine en yakın olan iki kümeyi birleştirir. Bölünmeli yaklaşımda ise başlangıçta tüm örnekleri tek bir kümeye koyup her seferinde birbirine en uzak iki küme bulmaya çalışmaktadır. Her iki işlem sırasında hiyerarşik bir ağaç ortaya çıkar. Örnekler arasındaki mesafenin bilinmesi bu algoritma için yeterlidir.

Hiyerarşik kümeleme, kümeleme algoritmalarını iç içe birleştirerek başarılı sonuçlar elde etmeyi amaçlar. Kümelerin hiyerarşisi hiyerarşi ağacı (dendogram) olarak temsil edilmektedir. Ağacın temelinde tüm örnekleri kapsayan tek bir küme bulunmaktadır. Ağacın dalları ilerledikçe kümelemeler oluşmaktadır.

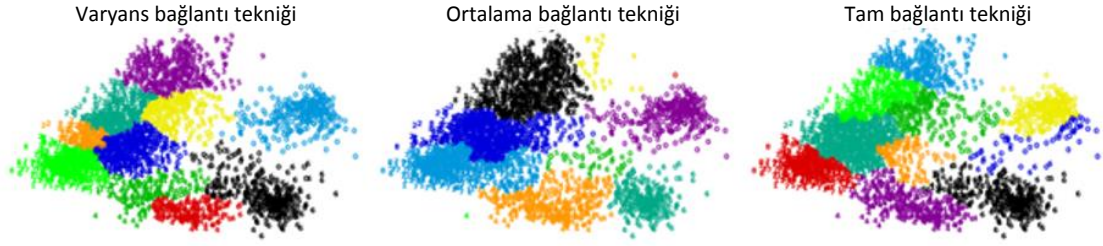
Dendogram örneęi Şekil 3.3’de verilmiştir. Burada girilecek kıstaslar ile küme sayısı, eşik miktarı, kümeleme yöntemi seçilmekte ve sınıflar oluşturulmaktadır [17].



Şekil 3.3 Hiyerarşik Kümeleme Dendogramı

Hiyerarşik kümeleme yöntemi uygulanırken küme sayısı farklı değerlerde uygulanmıştır. Bunun yanında kümeleme yöntemleri olarak varyans (ward) tekniği, ortalama (average), maksimum, tam (complete), tek (single) bağlantı yöntemleri seçilmiştir. Bu yöntemlerin stratejisi şu şekildedir:

Varyans yaklaşımlı hiyerarşik bağlantı tekniği, kümelerin farklarının karelerinin toplamını minimize etmektedir. Ortalama yaklaşımında, kümeler arası mesafelerin ortalamalarını aynı tutmaya çalışılır. Tek bağlantı yöntemi, en kısa mesafe esasına dayanmaktadır. Birbirine en yakın iki gözlemi bularak ilk aşamaya oturtmakta ve daha sonra yine birbirine yakın gözlemleri bularak kümeyi genişletmektedir. Tam bağlantı tekniğinde ise tek bağlantıdan farklı olarak en uzak gözlemlerin kümelenmesinden başlanmaktadır. Ortalama yöntemi, tek bağlantı ve tam bağlantı tekniği ile paralel olarak, kümeler arasındaki ortalama uzaklığın en düşük değerini temel almaktadır. Maksimum bağlantı yöntemi ise en uzun küme aralığını minimize etmektedir. Farklı hiyerarşik bağlantı yöntemleri sonucu ortaya çıkan gruplamalar Şekil 3.4'de gösterilmiştir [16].



Şekil 3.4 Farklı Hiyerarşik Bağlantı Yöntemleri İle Kümeleme [16]

3.5 Ayarlanmış Rasgelelik Endeksi (Adjusted Rand Index) ile Kümeleme Performansı Ölçümü

Ayarlanmış Rasgelelik Endeksi kümeleme sonucu ile doğru kümelemeyi karşılaştırmak ve başarı hesaplamak için kullanılmıştır. Kümeleme başarılarının ölçümünde sıklıkla kullanılan başarılı bir yöntemdir [16]. T verisinin doğru kümelemeyi; C verisinin, kümeleme algoritmaları tarafından üretilen kümeleme sonuçlarını ifade ettiği varsayılırsa ARE değeri denklem (3.6)'da ifade edilen fonksiyonla C ve T verisi kullanılarak hesaplanmaktadır [18].

$$ARE(T, C) = \frac{2(ad-bc)}{(a+b)(b+d)+(a+c)(c+d)} \quad (3.6)$$

Burada a , hem T hem C verisinde aynı kümeye ait olan piksel çiftlerinin sayısını temsil etmektedir. b , T verisinde aynı kümeye fakat C verisinde farklı kümelerine ait olan piksel çiftlerinin sayısını, c T verisinde farklı kümelerine fakat C verisinde aynı kümeye ait olan piksel çiftlerinin sayısını, d ise her iki veride de farklı kümelerine ait olan piksel çiftlerinin sayısını temsil etmektedir.

ARE değeri $\{0,1\}$ aralığında değer almaktadır. Bu değer yüksek olması C verisinin T verisine daha benzer olduğunu ifade etmektedir [16].

3.6 Karışıklık Matrisi ile Sınıflandırma Performansı Ölçümü

Karışıklık (confusion) matrisi sistemin sınıflama başarısını ölçmek için kullanılmaktadır. Örnek kümesindeki gerçek sınıf etiketi ile doğru ya da yanlış tahmin edilen sınıf etiket sayılarını içerir [19]. Çizelge 3.1'de iki sınıflandırma için karmaşıklık matrisi örneği verilmiştir. Burada DP doğru pozitif, YP yanlış pozitif, YN yanlış negatif, DN doğru

negatif demektir. Pozitif olan bir sonuç pozitif olarak sınıflanmışsa DP olarak, negatif olarak sınıflanmışsa YP olarak temsil edilir. Aynı şekilde negatif olması gereken doğru olarak negatif seçilmişse DN, yanlış seçilmişse YN olarak karışıklık matrisinde yerini alır.

Çizelge 3.1 İkili sınıflandırma için karışıklık matrisi

Karışıklık Matrisi		Tahmin Edilen	
		p	n
Gerçek	p	DP	YP
	n	YN	DN

Çizelge 3.2’de ise çok sınıflı bir sınıflandırma işlemi için karışıklık matrisi verilmektedir. Burada her bir sınıf için ayrı ayrı sınıflandırma başarıları hesaplanabilir. Örneğin, B sınıfı için DP miktarı 7’dir. Bununla birlikte 3 örnek ise yanlış sınıflara verilmiştir. Yani YN oranı 3’tür. B sınıfına ait olmadığı halde A sınıfına ait bir örnek sınıflandırma sonucu B olarak etiketlenmiştir. Buna bakılarak da B sınıfının YP miktarı bir olarak hesaplanmaktadır.

Çizelge 3.2 Çok sınıflı örnek için karışıklık matrisi

	Sınıf Nosu	Tahmin Edilen Sınıf			
		A	B	C	D
Gerçek Sınıf	A	9	1	0	0
	B	1	7	0	2
	C	0	0	10	0
	D	2	0	2	6

Karışıklık matrisi çıkarıldıktan sonra sınıflandırıcının performansının ölçülebilmesi için bazı kıstaslar belirlenmiştir. Bunlar doğruluk (accuracy), kesinlik (precision), geriçağırım (recall) ve F1_skor değerleridir. Doğruluk sınıflandırıcıda doğru sınıflandırılmış(DP+DN) örnek sayısının tüm örnek sayısına oranıdır. Kesinlik ise DP değerlerin pozitif değere sınıflandırılanların toplamına oranıdır. Geriçağırım değeri, DP değerinin gerçek değeri pozitif olanlara(DP+YN) oranı ile hesaplanmaktadır. F1_skor değeri ise kesinlik ve geriçağırım değerlerinin harmonik ortalamasının bulunarak tek bir ölçütte

değerlendirilmesini sağlamaktadır. Çizelge 3.3’de performans kıstasları ve formülleri verilmektedir [20].

Çizelge 3.3 Sınıflandırma performans kıstasları ve hesaplamaları [20]

Ölçüm	Formül
Doğruluk	$\frac{dp + dn}{dp + dn + yp + yn}$
Kesinlik	$\frac{dp}{dp + yp}$
Geriçağırım	$\frac{dp}{dp + yn}$
F_{β_skor}	$\frac{(\beta^2 + 1)dp}{(\beta^2 + 1)dp + \beta^2 yn + yp}$
F_1_skor	$\frac{2 * kesinlik * geriçağırım}{kesinlik + geriçağırım}$

FARKLI VERİ SETLERİ İLE BOYUT İNDİRGEME VE SINIFLANDIRMA

4.1 Yöntem

Bu tez çalışmasında, ATBA yöntemi ile band seçimi yapılması ve başarılı sonuçlar elde edilmesi amaçlanmıştır. Çalışma kapsamında hiperspektral veri seti üç aşamadan geçmektedir. İlk aşamada veri setine çeşitli boyut indirgeme yöntemleri uygulanarak özellik çıkarımı ya da band seçimi yapılmıştır. Daha sonra ise çeşitli öğrenme algoritmaları ile sınıflandırma yapılmıştır. Bu sınıflandırmaların sonucunda ise karışıklık matrisi yöntemi ya da Ayarlanmış Rasgelelik Matrisi yöntemleri ile sınıflandırma başarıları ölçülmüş ve kıyaslamalar yapılmıştır.

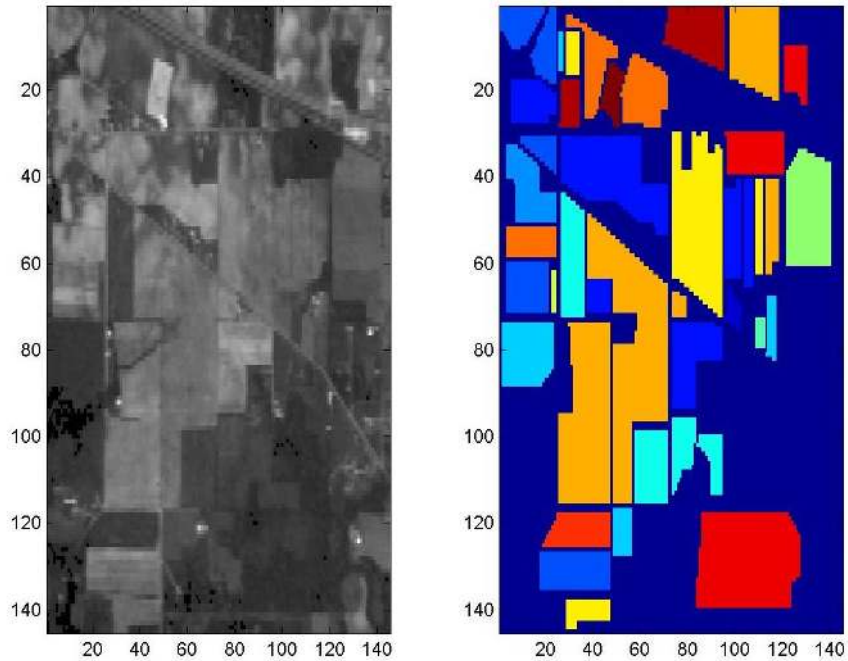
Boyut indirgeme işlemlerine geçmeden veri setindeki her bir sınıftan eşit sayıda örnekler alınmış, eğitim ve test sınıfları oluşturulmuştur. Örneklerin alınacağı sınıfın en az 100 adet imzaya sahip olmasına dikkat edilmiş, eğer yeterli örnek yok ise sınıf elenmiştir. Örnekler alındıktan sonra ilk adım olan boyut indirgeme adımı uygulanmıştır. Bu adımda klasik Temel Bileşen Analizi özellik çıkarımında ve band seçiminde kullanılırken Ayrık Temel Bileşen Analizi yöntemi farklı yaklaşımlar ile band seçimi için uygulanmıştır. Klasik Temel Bileşen Analizi yönteminde band seçimi yapılırken temel bileşen değerlerinin en yüksek olduğu bandlar seçilmiştir. Ayrık Temel Bileşen Analizi yönteminde ise sıfır dışında ağırlık verilmiş olan bandlar seçilerek sınıflandırma aşamasına geçilmiştir.

Sınıflandırma aşamasında farklı öğrenme yöntemleri kullanılmıştır. Çıkarılan özellikler ya da seçilmiş bandlardan elde edilen veriler hiyerarşik kümeleme ve K-ortalama kümeleyicilerine, k en yakın komşuluk ve Destek Vektör Makinesi sınıflandırıcılarına

sokulmuştur. Kümeleyici çıkışlarının başarısı Ayarlanmış Rasgelelik Endeksi ile sınıflayıcı çıkışlarınınki ise karışıklık matrisi oluşturulması sonucu yapılan başarı hesaplamaları ile belirlenmiştir.

4.2 AVIRIS - Indian Pines Görüntülerinde Boyut İndirgeme ve Sınıflandırma

İlk olarak NASA'nın AVIRIS sensörü ile çekilmiş olan "Indian Pines" veri seti üzerinde boyut indirgeme ve sınıflandırma işlemleri uygulanmıştır. AVIRIS sensörü 10 nm spektral çözünürlüğe sahiptir ve 400-2500 nm aralığındaki dalga boylarını görüntüleyebilmektedir. Indian Pines görüntüsü ağırlıklı olarak tarım ve ormanlık alanları kapsamaktadır. 145x145 uzamsal çözünürlüğün yanında 224 band derinliğine sahip olan Indian Pines verisine ait görüntü ve bu görüntüye ait sınıfları içeren doğruluk verisi Şekil 4.1'de verilmiştir. Bu görüntüdeki 4 bandın gürültülü ve bozuk bandlar olduğu bilindiğinden dolayı çıkarılmış ve çalışmamızda 220 band olarak kullanılmıştır. Veri setinde bulunan sınıflar ve her bir sınıfa ait piksel sayısı Çizelge 4.1'de verilmiştir.



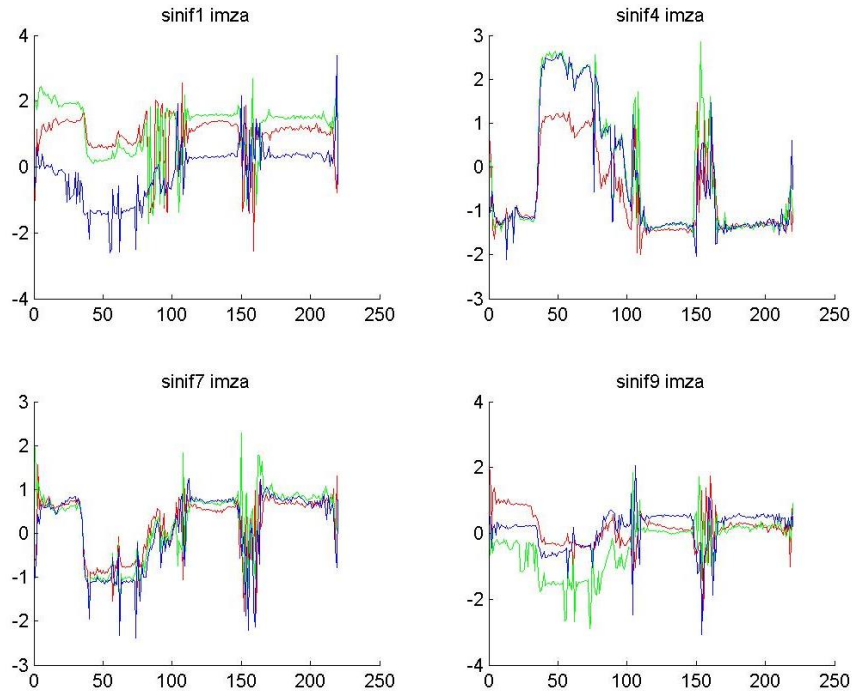
Şekil 4.1 Indian Pines örnek band görüntüsü ve doğruluk verisi bilgisi

Çizelge 4.1 Indian Pines görüntüsünde kullanılan sınıflar ve örnek sayıları [21]

#	Sınıf	Örnek Sayısı
1	Yonca	46
2	Mısır(işlenmemiş)	1428
3	Mısır(az işlenmiş)	830
4	Mısır	237
5	Çimen-otlak	483
6	Çimen-ağaç	730
7	Çimen-otlak(biçilmiş)	28
8	Saman	478
9	Yulaf	20
10	Soya fasulyesi(işlenmemiş)	972
11	Soya fasulyesi(az işlenmiş)	2455
12	Soya fasulyesi	593
13	Buğday	205
14	Odun	1265
15	Bina-ağaç-araç	386
16	Taş-çelik-kuleler	93

Şekil 4.2’de bazı sınıflarına ait spektral imzaları verilmiş olan Indian Pines veri setine ait örnekler ilk olarak hiçbir boyut indirgeme işlemine sokulmadan doğrudan sınıflandırıcılara verilmiştir. Örnekler işleme sokulmadan önce örnek sayısı yüzü geçemeyen 1, 7, 9, 16 nolu sınıflar elenerek eğitim ve test setleri belirlenmiştir. Sınıflandırıcı olarak k en yakın komşuluk sınıflayıcısı ve Destek Vektör Makinaları kullanılmıştır. Ayrıca K-ortalama kümeleyicisi ve hiyerarşik kümeleme yöntemleri kullanılmıştır. Hiyerarşik kümeleme yöntemlerinin hepsi denenmiştir. Ancak en başarılı

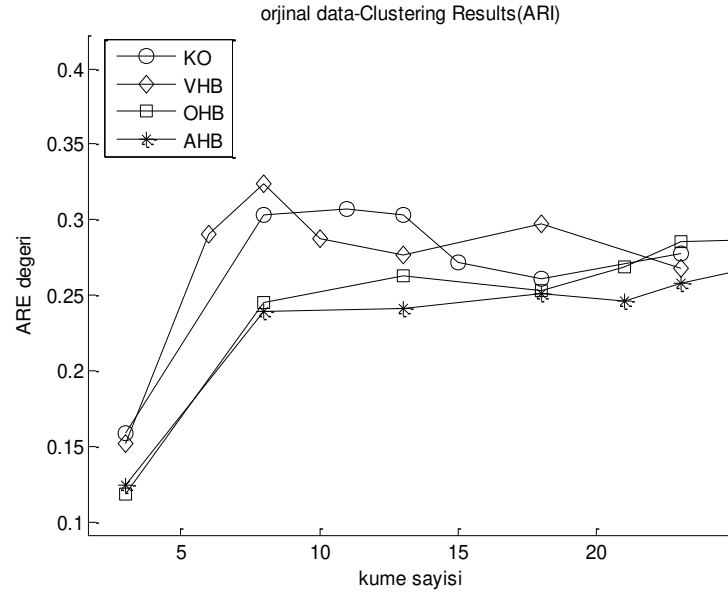
sonuçları 'varyans', 'ortalama', 'ağırlık' yaklaşım teknikleri verdiğiinden dolayı çalışmamızda yalnızca bunların sonuçları yer almaktadır.



Şekil 4.2 Indian Pines sınıflarına ait örnek spektral imzalar

4.2.1 Doğrudan Sınıflandırma Uygulanması

220 bandlı orjinal Indian Pines verisine K-ortalama ve hiyerarşik kümeleme yöntemleri uygulanmış ve ARE başarı hesaplaması ile başarıları hesaplanmıştır. Boyut indirgeme yapılmadan doğrudan sınıflandırma yapılan bu adım tezin kalan kısımlarında 'Orijinal' olarak adlandırılmaktadır. Sınıf sayısı bilinmediğinden dolayı farklı sınıf sayıları için kümeleme başarıları ölçülmüş ve sonuçları Şekil 4.3'de verilmiştir.



Şekil 4.3 Orijinal veri seti kümeleme sonuçları

Kümeleme yöntemlerinden farklı olarak k en yakın komşuluk ve Destek Vektör Makinaları eğitici sınıflandırma yöntemleri de orijinal veriden alınan örneklerle uygulanmıştır. Elde edilen sonuçlardan doğru tahmin edilenlerin tüm örnek sayısına bölünmesi ile 'doğruluk' başarı sonucu hesaplanmıştır.

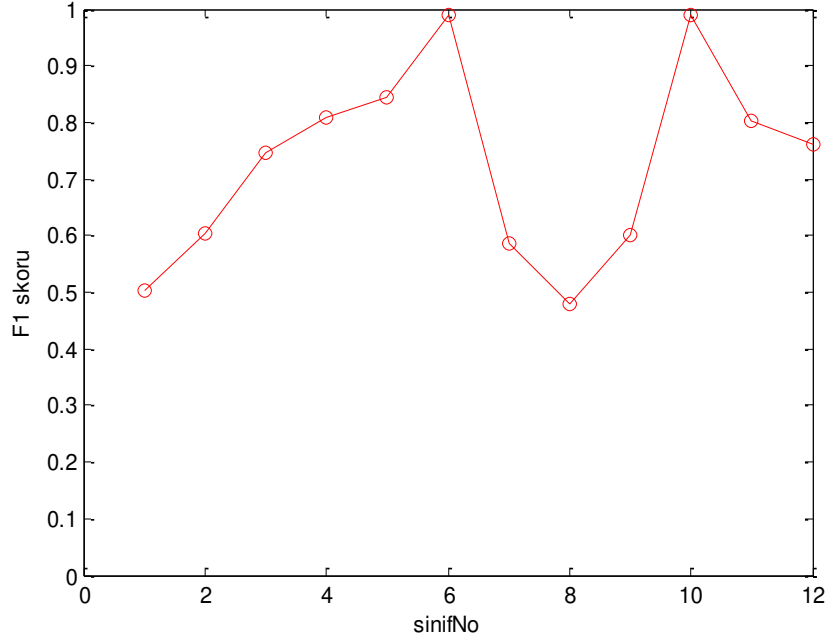
K en yakın komşuluk ve DVM sınıflandırıcılarının sonuçları alınarak karışıklık matrisi oluşturulmuş ve genel başarının hesaplanması dışında her bir sınıf için ayrı ayrı başarı performansları hesaplanmıştır. DVM için elde edilen karışıklık matrisi Şekil 4.4'teki gibidir. Karışıklık matrisinden elde edilen sınıflandırma başarı sonuçları ise Şekil 4.5 ve Şekil 4.6'de verilmektedir.

		TahminEdilenSinif											
		1	2	3	4	5	6	7	8	9	10	11	12
GercekSinif	1	58.0	6.0	8.0	0.0	0.0	0.0	20.0	4.0	4.0	0.0	0.0	0.0
	2	10.0	64.0	4.0	0.0	0.0	0.0	4.0	6.0	12.0	0.0	0.0	0.0
	3	16.0	6.0	74.0	0.0	0.0	0.0	0.0	0.0	4.0	0.0	0.0	0.0
	4	0.0	0.0	0.0	84.0	6.0	0.0	4.0	0.0	0.0	0.0	4.0	2.0
	5	0.0	0.0	2.0	10.0	82.0	0.0	0.0	0.0	0.0	0.0	0.0	6.0
	6	0.0	0.0	0.0	2.0	0.0	98.0	0.0	0.0	0.0	0.0	0.0	0.0
	7	14.0	2.0	2.0	0.0	0.0	0.0	58.0	24.0	0.0	0.0	0.0	0.0
	8	18.0	20.0	2.0	0.0	0.0	0.0	8.0	44.0	8.0	0.0	0.0	0.0
	9	14.0	14.0	6.0	0.0	0.0	0.0	4.0	6.0	56.0	0.0	0.0	0.0
	10	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	98.0	0.0	0.0
	11	0.0	0.0	0.0	6.0	0.0	0.0	0.0	0.0	0.0	0.0	78.0	16.0
	12	0.0	0.0	0.0	6.0	4.0	0.0	0.0	0.0	2.0	0.0	12.0	76.0

Şekil 4.4 Orijinal veri setine DVM uygulanması ile oluşturulan karışıklık matrisi

sinifNo			
	Kesinlik	Gericagirim	F1_Skoru
1	0.4462	0.5800	0.5043
2	0.5714	0.6400	0.6038
3	0.7551	0.7400	0.7475
4	0.7778	0.8400	0.8077
5	0.8723	0.8200	0.8454
6	1.0000	0.9800	0.9899
7	0.5918	0.5800	0.5859
8	0.5238	0.4400	0.4783
9	0.6512	0.5600	0.6022
10	1.0000	0.9800	0.9899
11	0.8298	0.7800	0.8041
12	0.7600	0.7600	0.7600

Şekil 4.5 Orijinal veri seti karışıklık matrisi performans sonuçları

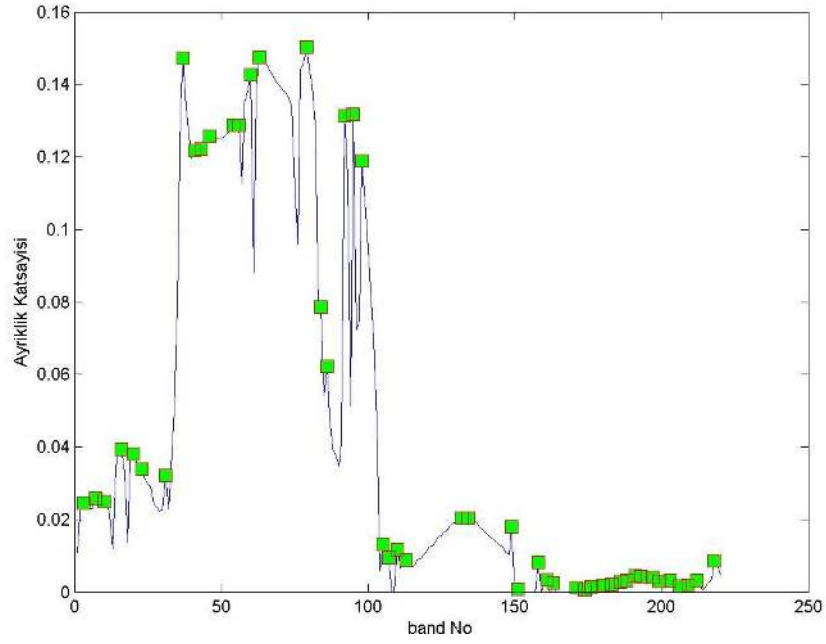


Şekil 4.6 Orijinal veri seti sınıflarının F1_skor değerleri

Her bir sınıfa ait doğru pozitif, doğru negatif, yanlış pozitif, yanlış negatif değerleri karışıklık matrisinde gösterilmiştir. Elde edilen veriler ile kesinlik (precision) ve geriçağırım (Recall) değerleri hesaplanarak F1_skor değeri bulunmuştur. F1_skor değeri sınıfların sınıflandırma başarılarını ölçmede kullanılan önemli bir parametredir.

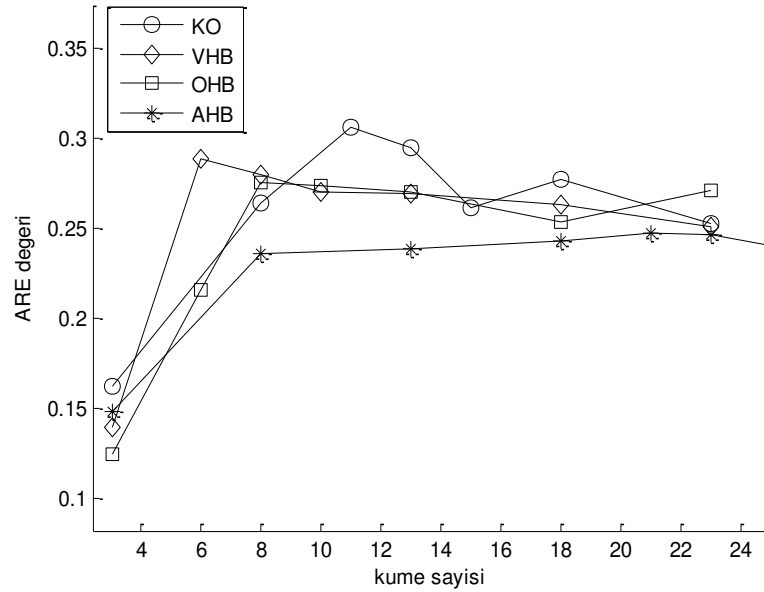
4.2.2 Yerel Maksimumları Seçilmiş Ayırık Temel Bileşen Analizi Uygulanması

Çalışmamızda bu yöntem “ATBA0” olarak adlandırılmıştır. Band sayısı sınırlaması olmadan uygulanan bu yöntemde ilk olarak klasik ATBA yöntemi Indian Pines verisine uygulanmıştır. Elde edilen ayrıklık verilerine genel yaklaşımdan farklı bir yaklaşımla yaklaşılmış ve band numaraları bir eksen de olacak şekilde ayrıklık değerleri çizdirildiğindeki tepe noktaları bulunmuştur. Böylece bandlara yerel bölgeler şeklinde yaklaşılmış ve bu bölgelerdeki en yüksek değerlerin olduğu bandlar seçilmiştir. Ayrıklık değerlerini ve tepe noktalarını gösteren grafik Şekil 4.7’de verilmektedir.



Şekil 4.7 Indian_Pines Datası Ayrıklık katsayıları ve yerel maksimum noktaları

Yalnızca yerel maksimum noktaların olduğu bandların seçilip kalan bandların elenmesi ile Indian Pines veri seti için band sayısı 220'den 48'e düşmüştür. Bu 48 bandın seçilmesi ile 1200x48'lik yeni veri setimiz elde edilmiştir. Band seçme işlemi tamamlandıktan sonra daha önce boyut indirgenmemiş veriye uygulanan sınıflandırma ve kümeleme yöntemleri uygulanarak başarı sonuçları hesaplanmıştır. ATBA0 verisinin kümeleme performans sonuçları Şekil 4.8'de verilmektedir.



Şekil 4.8 Yerel maksimumları seçilmiş ATBA uygulanması sonucu kümeleme başarısı

Veri için genel performans değerlendirmesinden sonra sınıflar seviyesinde başarı ölçümü yapabilmek için Şekil 4.9'deki karışıklık matrisi oluşturulmuştur.

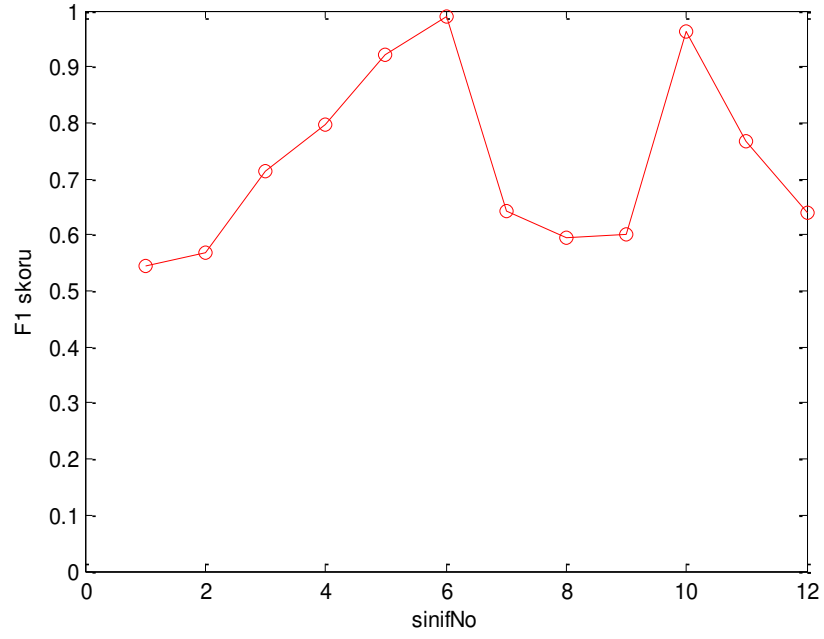
	TahminEdilenSinif											
	1	2	3	4	5	6	7	8	9	10	11	12
GerçekSinif	1	60.0	10.0	8.0	0.0	0.0	0.0	10.0	10.0	2.0	0.0	0.0
	2	4.0	58.0	6.0	0.0	0.0	0.0	12.0	10.0	10.0	0.0	0.0
	3	10.0	12.0	72.0	0.0	0.0	0.0	0.0	2.0	4.0	0.0	0.0
	4	0.0	0.0	0.0	82.0	2.0	0.0	4.0	0.0	0.0	6.0	6.0
	5	0.0	0.0	0.0	4.0	92.0	0.0	0.0	0.0	0.0	0.0	4.0
	6	0.0	0.0	0.0	2.0	0.0	98.0	0.0	0.0	0.0	0.0	0.0
	7	16.0	6.0	0.0	0.0	0.0	0.0	68.0	6.0	4.0	0.0	0.0
	8	18.0	10.0	0.0	0.0	0.0	0.0	12.0	56.0	4.0	0.0	0.0
	9	12.0	8.0	16.0	0.0	0.0	0.0	6.0	4.0	54.0	0.0	0.0
	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0
	11	0.0	0.0	0.0	6.0	0.0	0.0	0.0	0.0	0.0	0.0	72.0
	12	0.0	0.0	0.0	12.0	6.0	0.0	0.0	0.0	2.0	8.0	10.0

Şekil 4.9 Yerel maksimumları seçilmiş ATBA uygulanması sonucu karışıklık matrisi

Karışıklık matrisi sonucunda her bir sınıfa ait sınıflandırma başarısı Şekil 4.10'deki gibi hesaplanmış ve F1_skor değerleri Şekil 4.11'de verilmiştir.

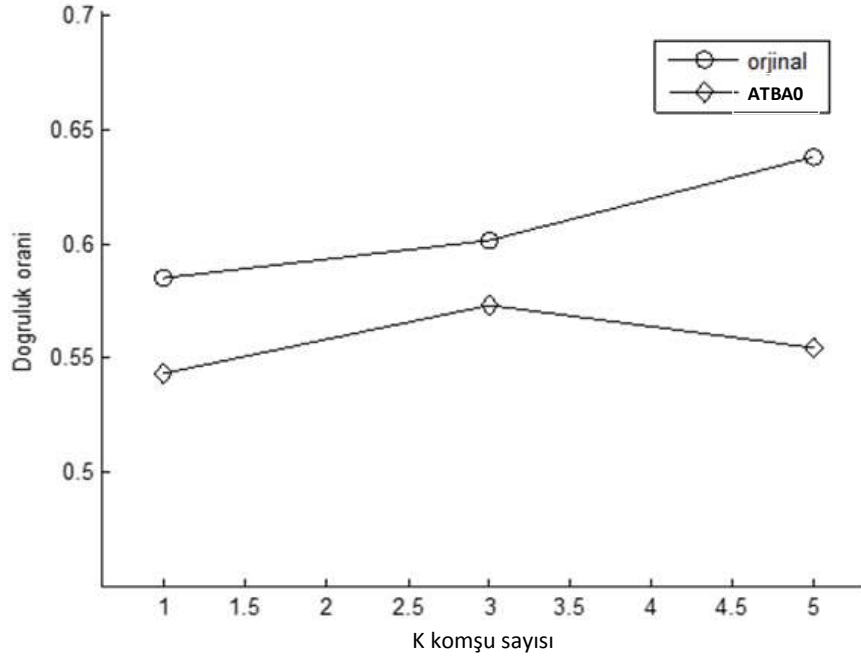
	Kesinlik	Gericagirim	F1_Skoru	
sinifNo	1	0.5000	0.6000	0.5455
	2	0.5577	0.5800	0.5686
	3	0.7059	0.7200	0.7129
	4	0.7736	0.8200	0.7961
	5	0.9200	0.9200	0.9200
	6	1.0000	0.9800	0.9899
	7	0.6071	0.6800	0.6415
	8	0.6364	0.5600	0.5957
	9	0.6750	0.5400	0.6000
	10	0.9259	1.0000	0.9615
	11	0.8182	0.7200	0.7660
	12	0.6596	0.6200	0.6392

Şekil 4.10 Yerel maksimumları seçilmiş ATBA uygulanması sonrası karışıklık matrisi performans değerleri



Şekil 4.11 Yerel maksimumları seçilmiş ATBA uygulanması sonucu sınıfların F1_skor değerleri

Şekil 4.12’de Orijinal ve ATBA0 yöntemlerinin sınıflandırma sonuçları gösterilmektedir. Band sayısının 48’a düşmesi ile 78% oranında azaltılmış boyut için sınıflandırma başarılarının aynı seviyelerde olduğu görülmektedir.

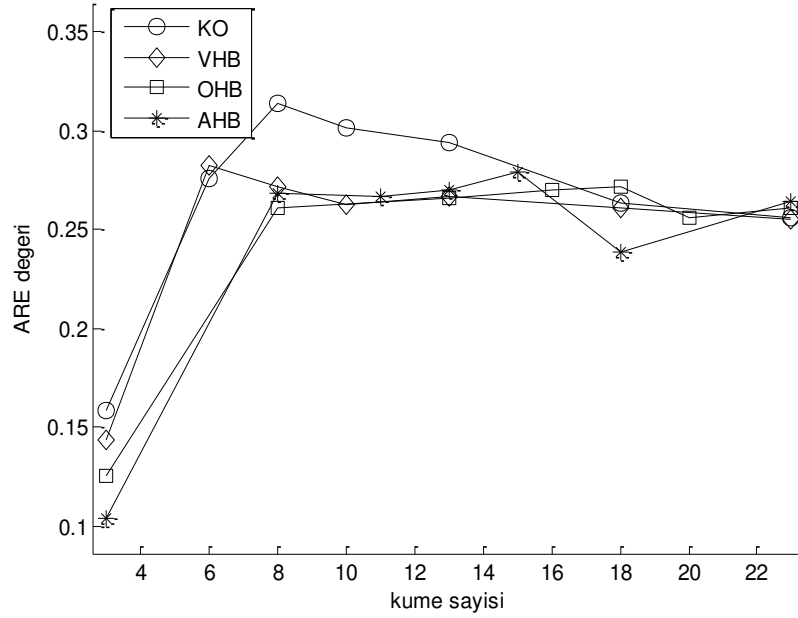


Şekil 4.12 k=1, 3, 5 için KEYK sınıflandırma sonuçlarının karşılaştırılması (Orijinal - Yerel maksimumları seçilmiş ATBA için)

4.2.3 Temel Bileşen Analizi Uygulanması

İlk olarak klasik Temel Bileşen Analizi yöntemi uygulanmıştır. Bu yöntem çalışmamızda “PCA0” olarak adlandırılmaktadır. Bu ve bundan sonraki boyut indirgeme yöntemlerinde seçilecek band sayısı ya da çıkarılacak özellik boyutu için sınır belirtilmiştir. Bu sınır 3, 5, 10 olarak denenmiş ve BS olarak temsil edilmiştir.

BS değeri 3 olarak seçildiğinde TBA yardımı ile veri farklı bir uzayda 1200x3'lük (12 sınıf için 100'er örnek) veriye dönüştürülmüştür. Bu elde edilen boyutu indirgenmiş veriye sınıflandırma işlemleri teker teker uygulandığında elde edilen sonuçlar aşağıda verilmiştir.



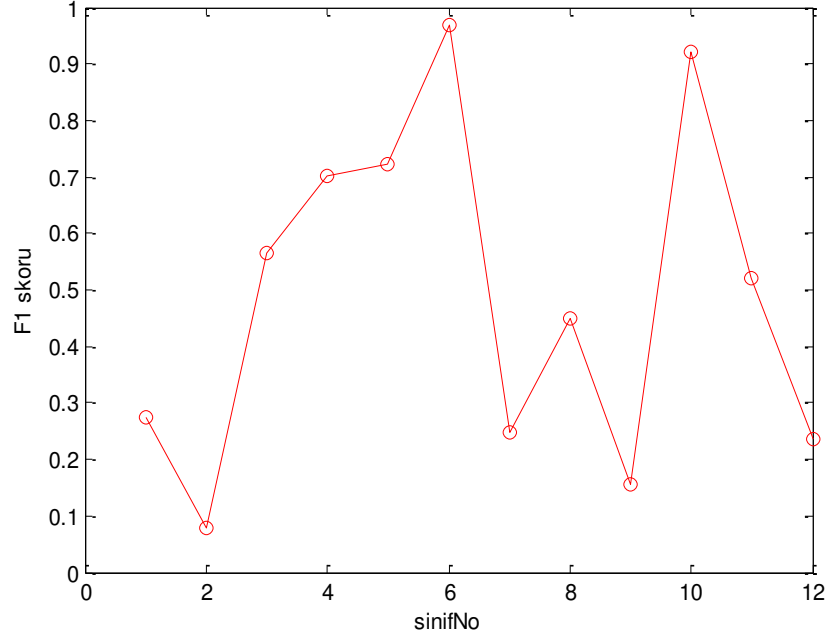
Şekil 4.13 Klasik TBA(BS=3) uygulanması sonucu kümeleme başarısı

		TahminEdilenSinif											
		1	2	3	4	5	6	7	8	9	10	11	12
GercekSinif	1	28.0	4.0	12.0	0.0	0.0	0.0	24.0	20.0	12.0	0.0	0.0	0.0
	2	30.0	8.0	4.0	0.0	0.0	0.0	28.0	16.0	14.0	0.0	0.0	0.0
	3	22.0	2.0	52.0	0.0	6.0	0.0	18.0	0.0	0.0	0.0	0.0	0.0
	4	0.0	0.0	0.0	78.0	6.0	4.0	0.0	2.0	0.0	0.0	10.0	0.0
	5	0.0	0.0	0.0	2.0	88.0	0.0	0.0	0.0	0.0	0.0	2.0	8.0
	6	0.0	0.0	0.0	2.0	0.0	98.0	0.0	0.0	0.0	0.0	0.0	0.0
	7	4.0	54.0	0.0	0.0	0.0	0.0	28.0	10.0	4.0	0.0	0.0	0.0
	8	4.0	22.0	2.0	0.0	0.0	0.0	10.0	50.0	12.0	0.0	0.0	0.0
	9	16.0	16.0	12.0	0.0	2.0	0.0	18.0	24.0	12.0	0.0	0.0	0.0
	10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	92.0	0.0	8.0
	11	0.0	0.0	0.0	40.0	4.0	0.0	0.0	0.0	0.0	0.0	52.0	4.0
	12	0.0	0.0	2.0	0.0	38.0	0.0	0.0	0.0	0.0	8.0	36.0	16.0

Şekil 4.14 Klasik TBA(BS=3) uygulanması sonucu karışıklık matrisi

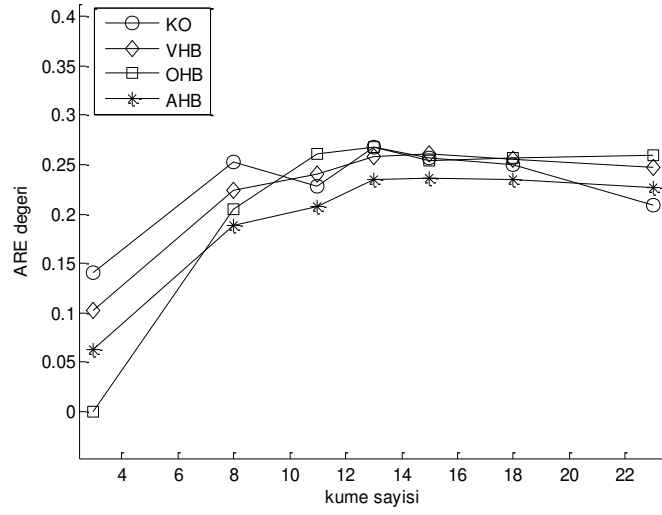
	Kesinlik			Gericagirim			F1_Skoru		
	1	2	3	4	5	6	7	8	9
sinifNo	1	0.2692	0.2800	0.2745					
	2	0.0755	0.0800	0.0777					
	3	0.6190	0.5200	0.5652					
	4	0.6393	0.7800	0.7027					
	5	0.6111	0.8800	0.7213					
	6	0.9608	0.9800	0.9703					
	7	0.2222	0.2800	0.2478					
	8	0.4098	0.5000	0.4505					
	9	0.2222	0.1200	0.1558					
	10	0.9200	0.9200	0.9200					
	11	0.5200	0.5200	0.5200					
	12	0.4444	0.1600	0.2353					

Şekil 4.15 Klasik TBA(BS=3) uygulanması sonrası karışıklık matrisi performans değerleri

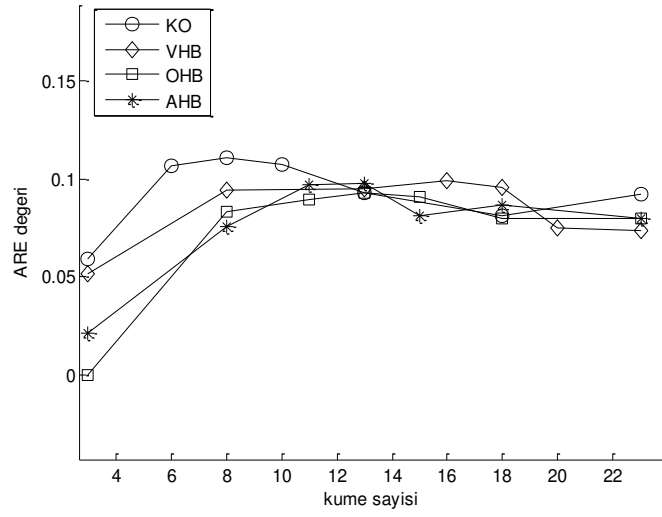


Şekil 4.16 Klasik TBA(BS=3) uygulanması sonucu sınıfların F1_skor değerleri

Klasik TBA ile özellik çıkarımı yapıldıktan sonra farklı bir yaklaşım sunulmuş ve band seçimi yapılmıştır. Çalışmamızda bu yöntemler “TBA1” ve “TBA2” olarak adlandırılmışlardır. Burada ilk olarak, TBA yöntemi ile seçilmiş olan temel bileşen değerlerine bakılmış ve temel bileşenlerin en büyük olduğu band değerleri seçilerek boyut indirgeme yoluna gidilmiştir. “TBA1” yönteminde 220x3'lük temel bileşen matrisinde her bir 220x1'lik vektörün en yüksek değeri aldığı satır band numarası olarak seçilmiştir. “TBA2” yönteminde ise 220x1'lik vektörler toplanarak elde edilen 220x1'lik vektörün en yüksek 3(BS adet) değerine karşılık gelen bandlar işaretlenerek band seçimi yapılmıştır. Boyut indirgemenin ardından sınıflandırma yöntemleri uygulanarak aşağıdaki sonuçlar elde edilmiştir.

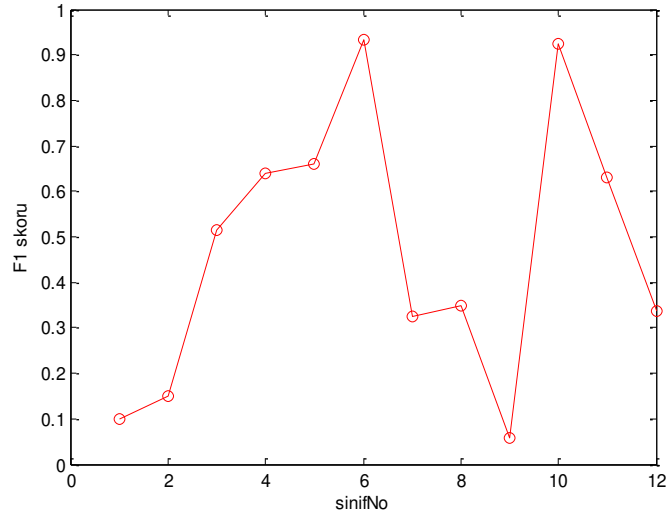


(a)

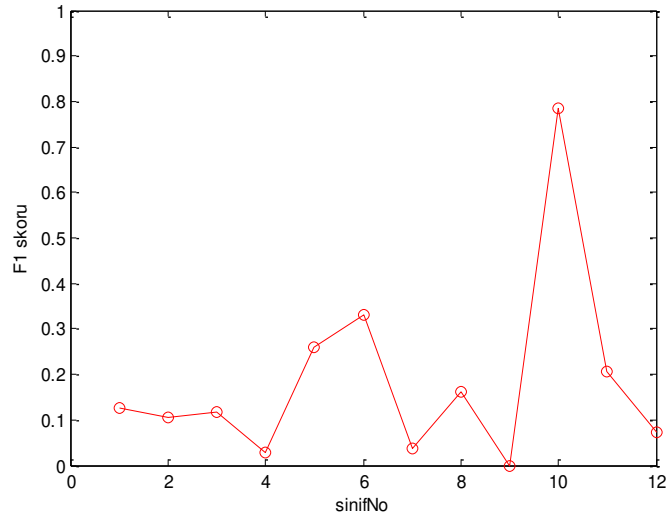


(b)

Şekil 4.17 TBA ile band seçilmesi sonucu kümeleme başarısı (a)TBA1 (b)TBA2



(a)

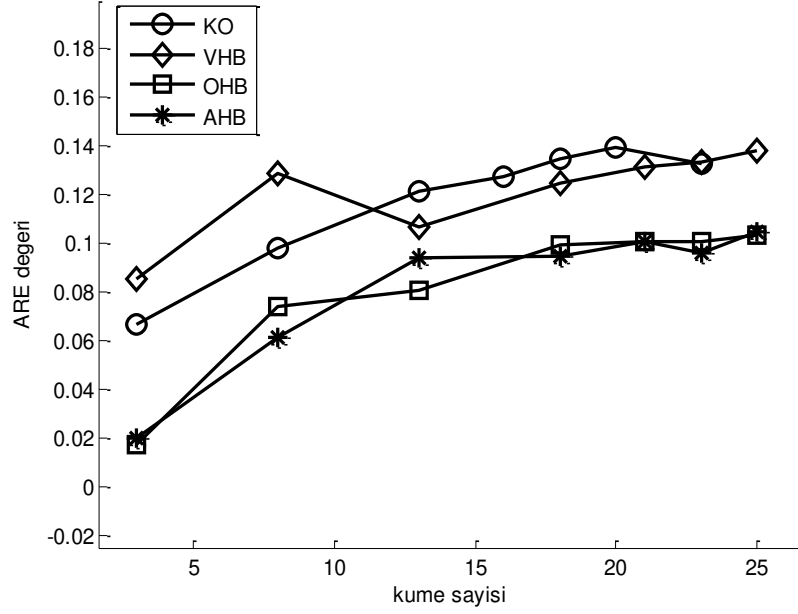


(b)

Şekil 4.18 BS=3 için TBA ile band seçilmesi sonrası F1_skor değerleri (a) TBA1 (b) TBA2

4.2.4 Klasik Ayırık Temel Bileşen Analizi Uygulanması

Çalışmamızda bu yöntem “KATBA” olarak adlandırılmıştır. Bu yöntemde TBA yöntemine ayırıklık yaklaşımı ile yaklaşılarak elde edilen katsayılara bakılmış ve band seçimi yapılmıştır [12]. KATBA yönteminde ilk olarak Şekil 4.7’deki gibi her bir band için ayırıklık katsayıları belirlenmektedir. Daha sonra BS=3 için en yüksek değerlere karşılık gelen 3 band seçilmekte ve bu bandlar oluşan 145x145x3’lük görüntüye ait örneklerden oluşan 1200x3’lük veri sınıflandırma işlemine sokulmaktadır. Kümeleme sonrası elde edilen grafik Şekil 4.19’de verilmektedir.

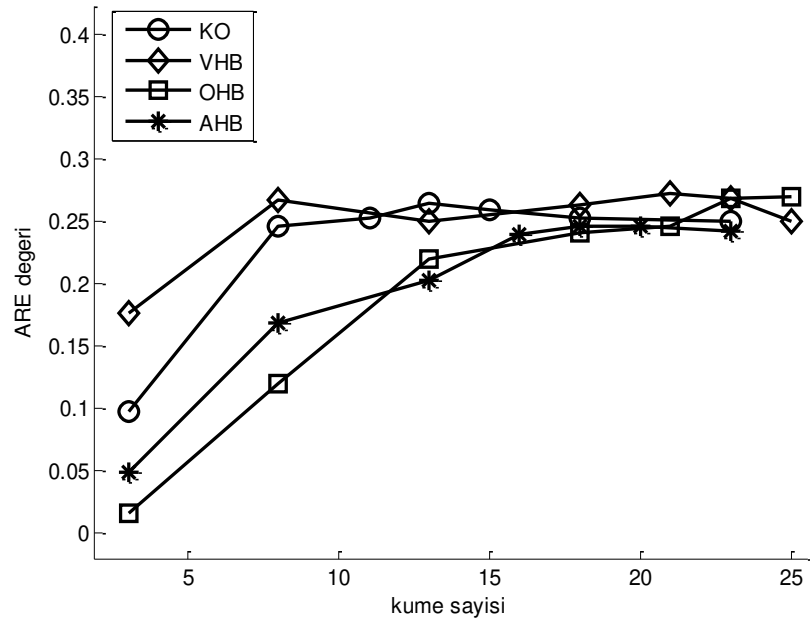


Şekil 4.19 Klasik ATBA uygulanması sonucu kümeleme başarıları

4.2.5 Gruplama Uygulanmış Ayırık Temel Bileşen Analizi Uygulanması

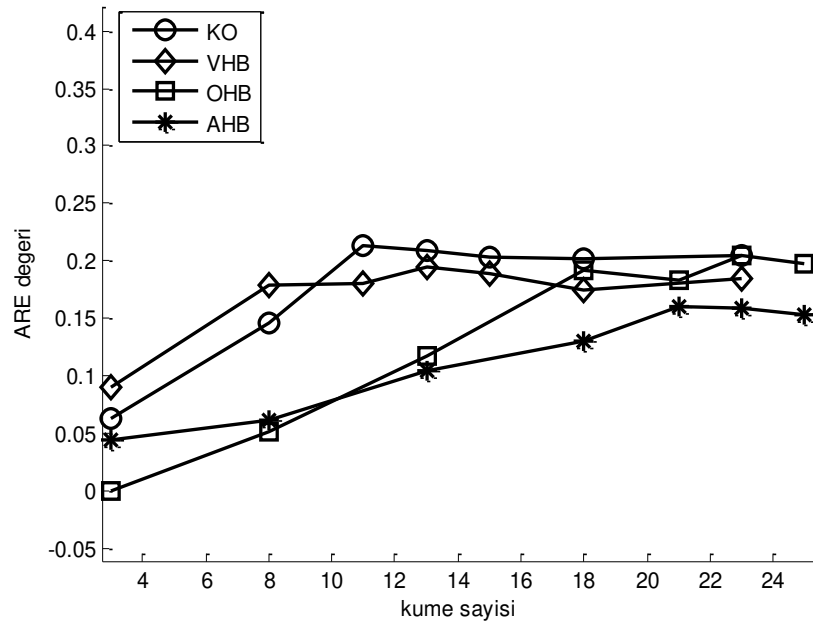
Çalışmamızda bu yöntemler “ATBA25” ve “ATBA40” olarak adlandırılmıştır. Seçilecek band sayısının(BS) belli olduğu bu yöntemde ilk olarak ATBA0 yönteminde anlatıldığı gibi yerel maksimum noktalar bulunmaktadır. Daha sonrasında ise ATBA0’dan farklı olarak tüm bandlar seçilmemekte ya da KATBA’daki gibi doğrudan en yüksek bandlar alınmamaktadır. Bunların yerine bandlar, belirlenmiş adım aralıkları ile taranmakta ve her aralıktaki en yüksek ayrıklık katsayısına sahip yerel maksimum noktası band olarak seçilmektedir. Böylelikle farklı dalga boylarındaki özelliklerin de kaybedilmemesi hedeflenmektedir.

ATBA25 için adım aralığı 25’tir. Yani sırasıyla 1-25, 26-50, 51-75, ..., 176-200, 201-220 band aralıkları taranmış ve her aralıktan en yüksek değere sahip bandlar (16 37 63 79 106 133 153 193) olarak bulunmuştur. Daha sonra bunlar arasından en yüksek değere sahip 3 (BS adet) band (79 63 37) seçilerek boyut indirgeme adımı tamamlanmıştır. Seçilen bu bandlara kümeleme işlemleri uygulandıktan sonra elde edilen sonuçlar Şekil 4.20’deki gibidir.



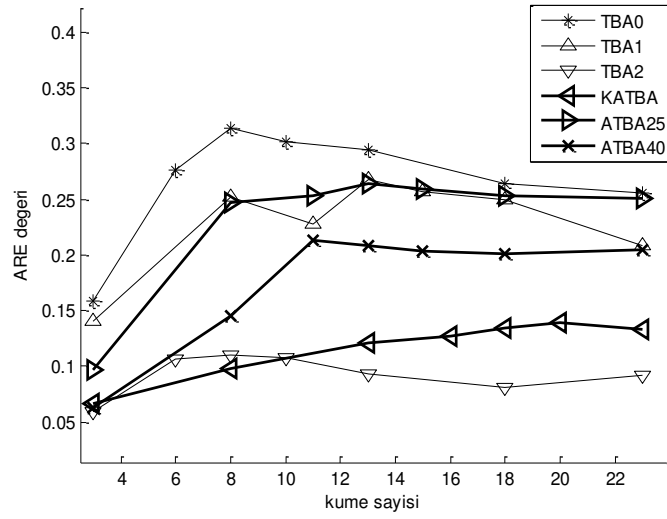
Şekil 4.20 Gruplama uygulanmış 25 adım aralıklı ATBA için kümeleme sonuçları

Tarama adım aralığı 40 yapılarak uygulanan yöntem ATBA40 olarak adlandırılmıştır. Adım aralığının artmasına bağlı olarak elde edilen kümeleme sonucu Şekil 4.21’de verilmiştir.

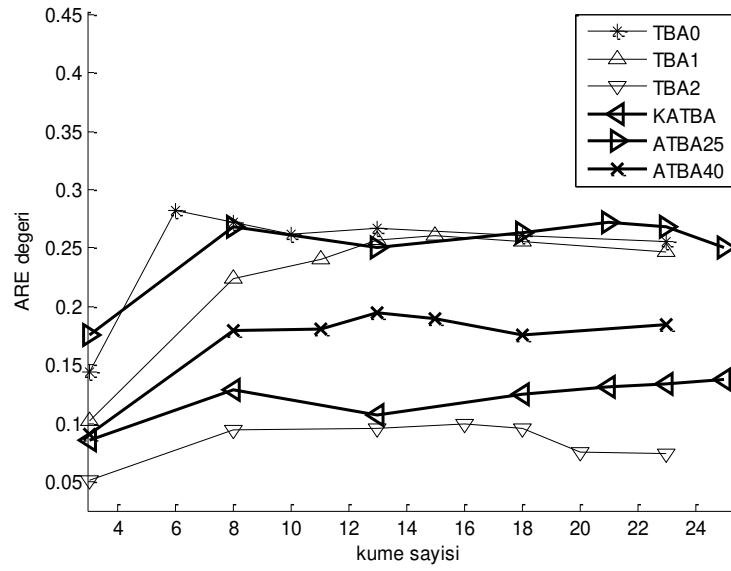


Şekil 4.21 Gruplama uygulanmış 40 adım aralıklı ATBA için kümeleme başarıları

BS=3 için uygulanan boyut indirgeme yöntemleri sonrası elde edilen verilerin farklı sınıflandırma yöntemleri sonucu başarılarının kıyaslanması Şekil 4.22’de verilmektedir.

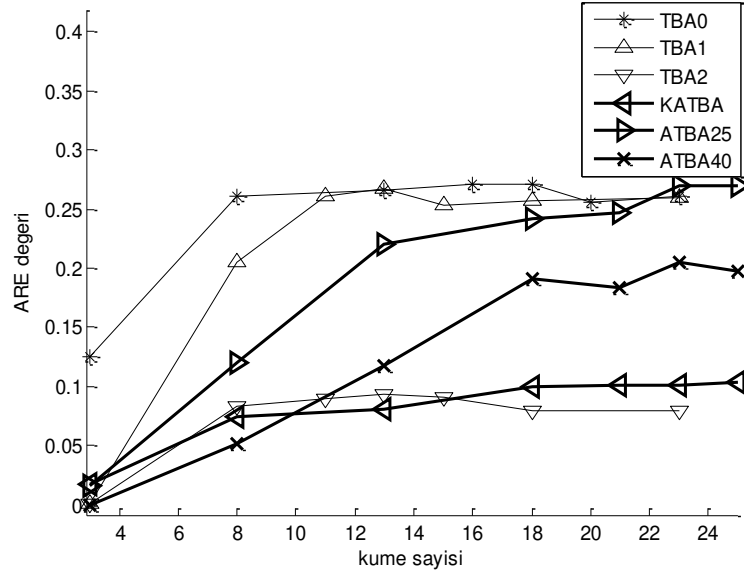


(a)

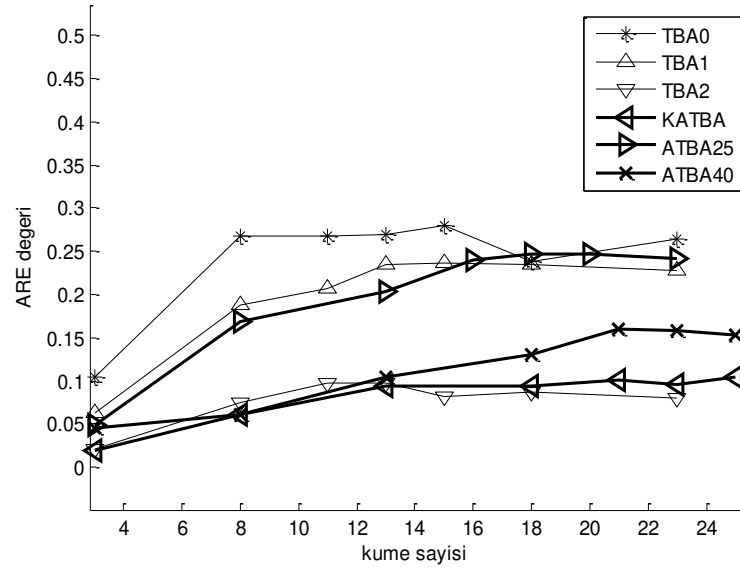


(b)

Şekil 4.22 BS=3 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) VYHB (c) OYHB (d) AYHB (e) KEYK (f) DVM

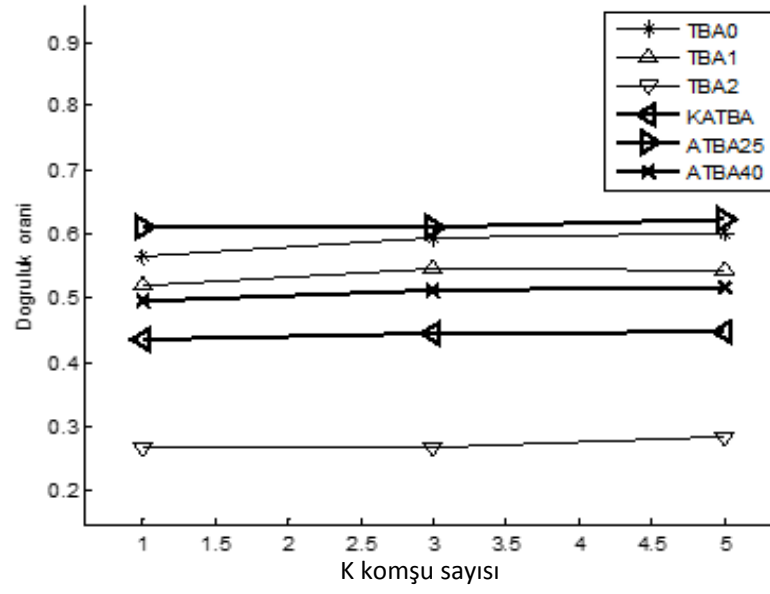


(c)

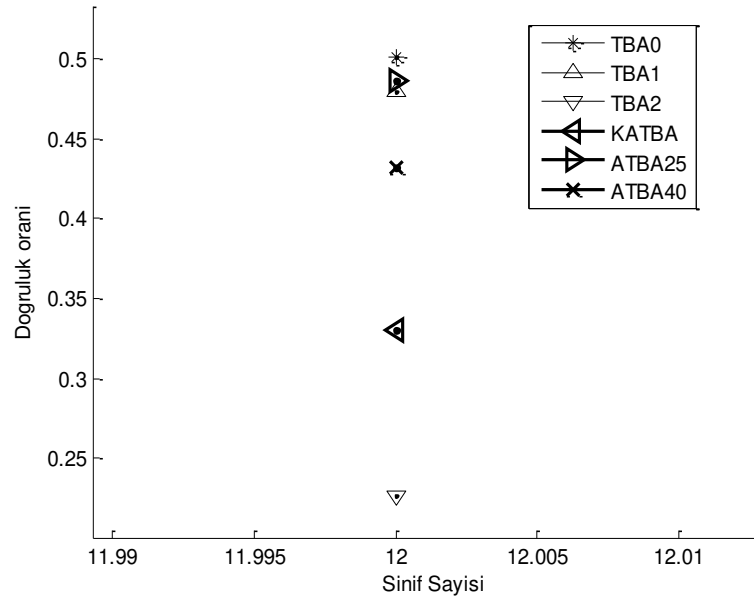


(d)

Şekil 4.22 BS=3 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) VYHB (c) OYHB (d) AYHB (e) KEYK (f) DVM (devamı)



(e)

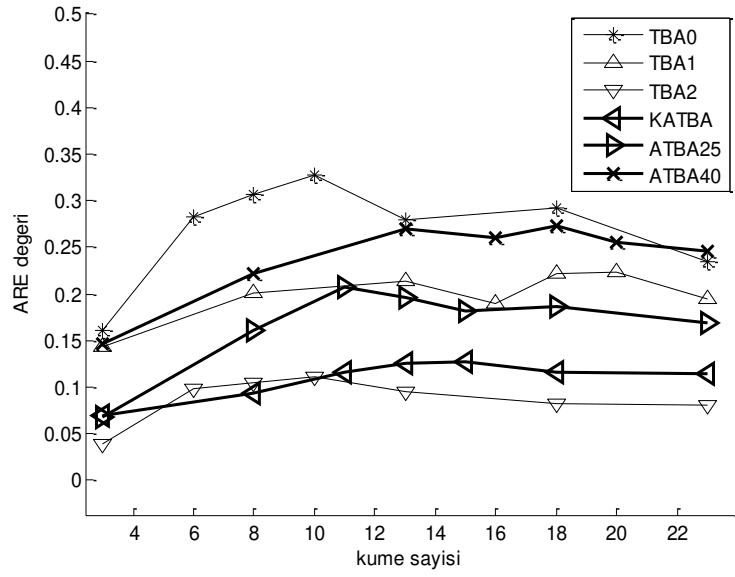


(f)

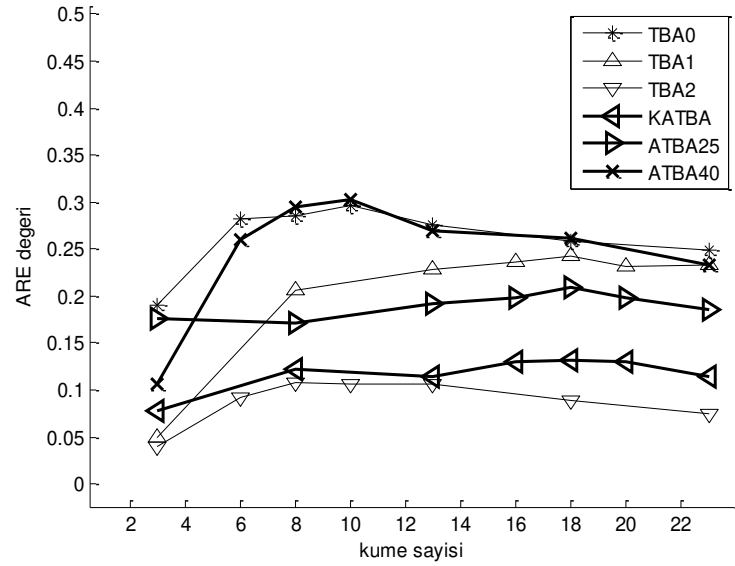
Şekil 4.22 BS=3 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) VYHB (c) OYHB (d) AYHB (e) KEYK (f) DVM (devamı)

4.2.6 Sonuçların Karşılaştırılması

BS seçilecek band sayısı 3 olacak şekilde sonuçlar bulunduktan sonra band sayısı 5 ve 10 için de aynı işlemler tekrarlanmıştır. Elde edilen sonuçlar aşağıdaki gibidir.

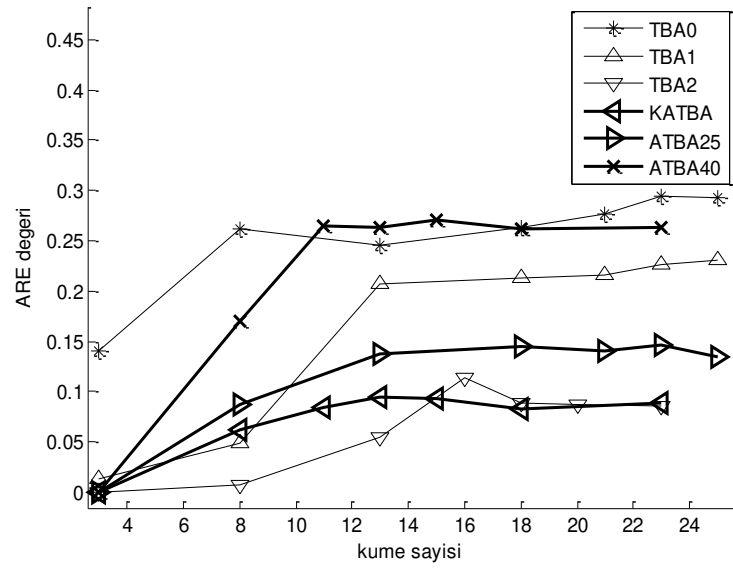


(a)

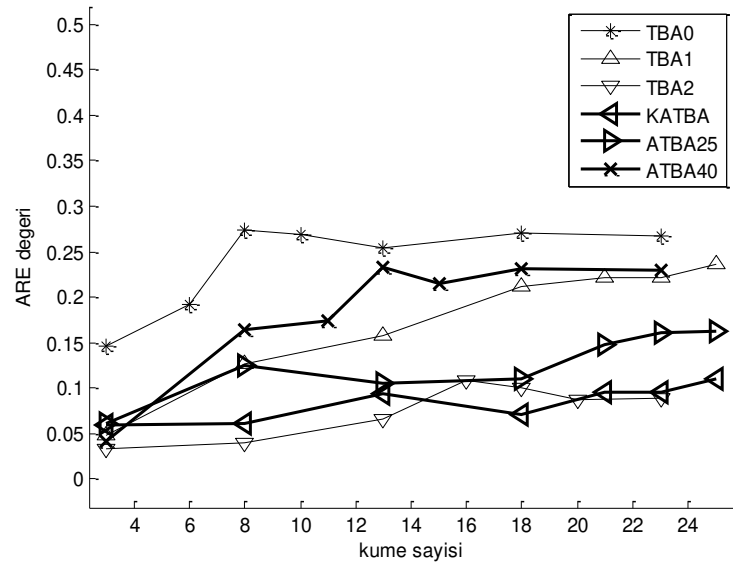


(b)

Şekil 4.23 BS=5 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) VYHB (c) OYHB (d) AYHB (e) KEYK (f) DVM

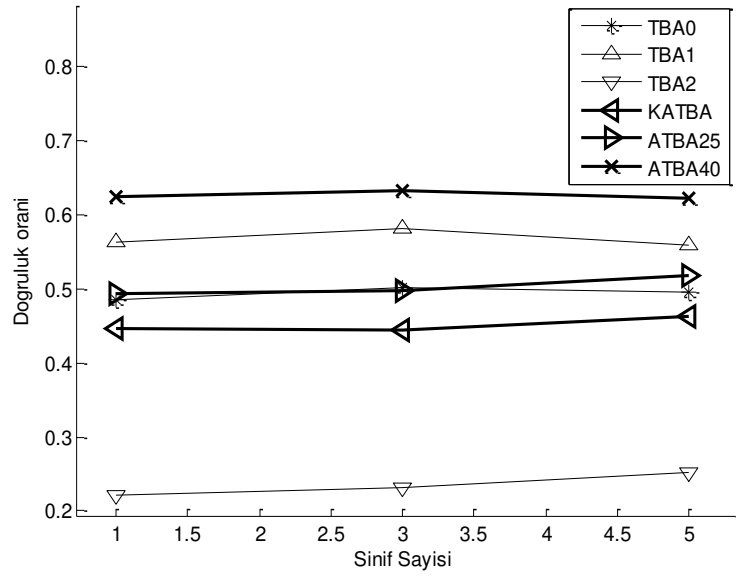


(c)

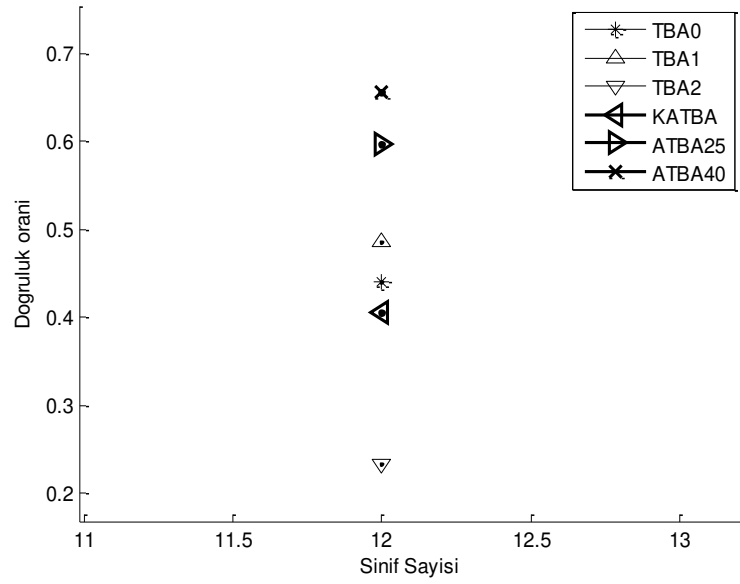


(d)

Şekil 4.23 BS=5 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) VYHB (c) OYHB (d) AYHB (e) KEYK (f) DVM (devamı)

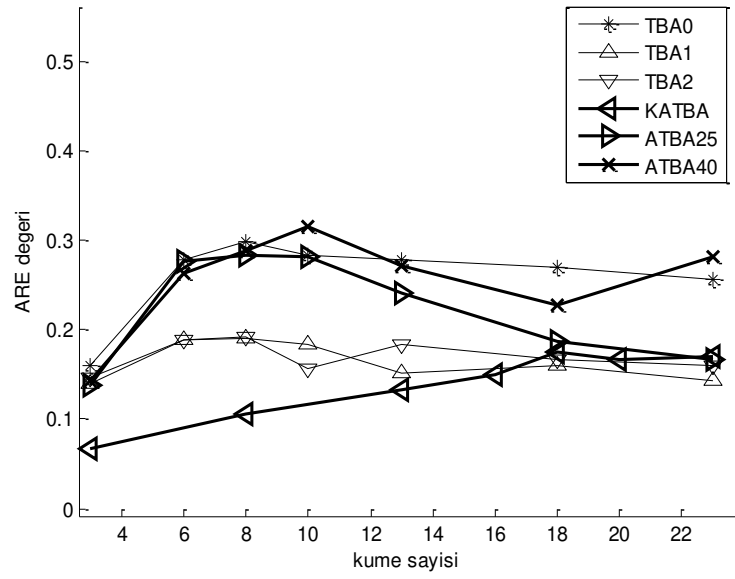


(e)

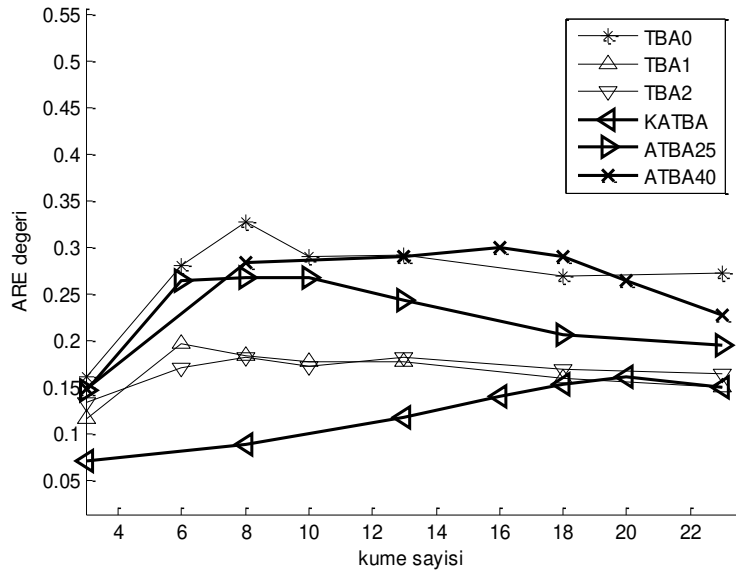


(f)

Şekil 4.23 BS=5 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) VYHB (c) OYHB (d) AYHB (e) KEYK (f) DVM (devamı)

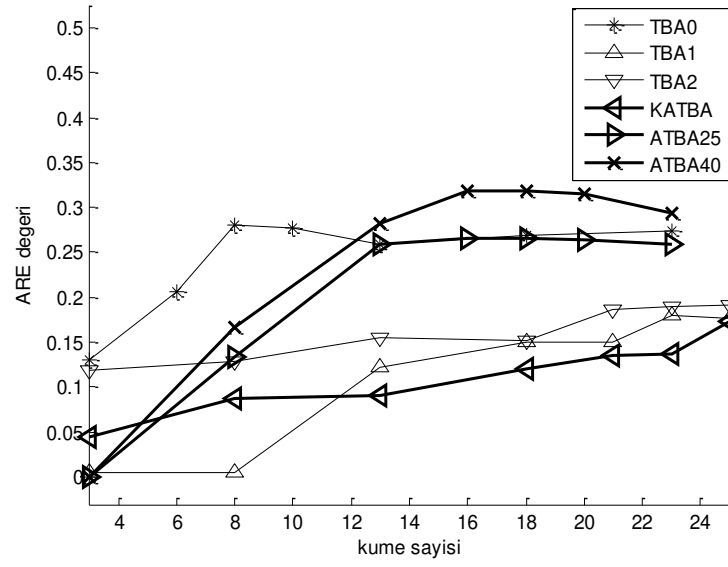


(a)

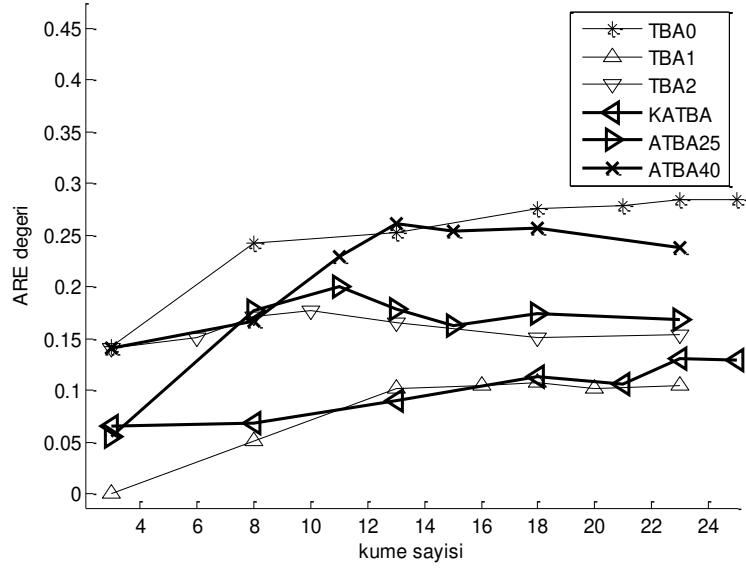


(b)

Şekil 4.24 BS=10 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) VYHB (c) OYHB (d) AYHB (e) KEYK (f) DVM

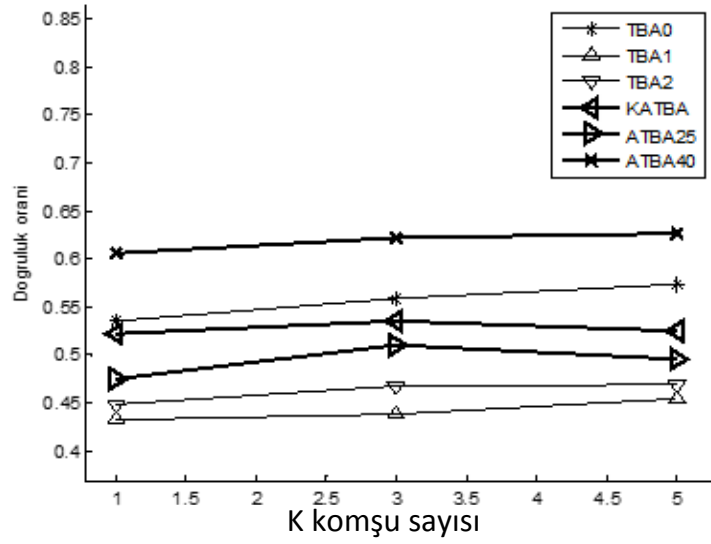


(c)

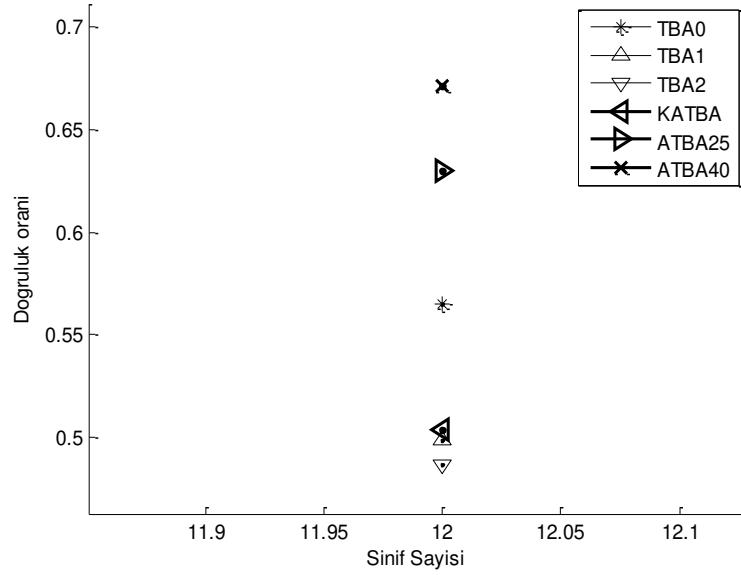


(d)

Şekil 4.24 BS=10 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) VYHB (c) OYHB (d) AYHB (e) KEYK (f) DVM (devamı)



(e)



(f)

Şekil 4.24 BS=10 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) VYHB (c) OYHB (d) AYHB (e) KEYK (f) DVM (devamı)

Indian Pines veri setine uygulanan tüm boyut indirgeme ve sınıflandırma işlemleri sonucunda elde edilen başarı değerleri Çizelge 4.2’de karşılaştırmalı olarak gösterilmektedir.

Çizelge 4.2 Indian Pines verisi için karşılaştırmalı sonuçlar

Veri Seti Indian_pines			Kümeleme(ARE)				Sınıflama(Doğruluk)	
			KO	Hiyerarşik Bağlantı			KEYK	DVM
				VHYB	OHYB	AHYB		
	Orijinal(220)		0,307546	0,323577	0,286535	0,265753	0,585	0,725
	ATBA0(48)		0,306714	0,288476	0,275489	0,247278	0,543333	0,728333
BS=3	Özellik Çıkarma	TBA0	0,31375	0,282839	0,271539	0,279317	0,565	0,501667
	Band Seçme	TBA1	0,267624	0,260984	0,267765	0,235965	0,52	0,48
		TBA2	0,110544	0,099216	0,093246	0,097915	0,266667	0,226667
		KATBA	0,139495	0,13798	0,103371	0,104597	0,435	0,33
		ATBA25	0,264381	0,272519	0,270368	0,246774	0,61	0,486667
		ATBA40	0,213201	0,195194	0,20495	0,160357	0,496667	0,431667
BS=5	Özellik Çıkarma	TBA0	0,327516	0,29589	0,293866	0,274155	0,485	0,44
	Band Seçme	TBA1	0,223566	0,24233	0,230565	0,235821	0,563333	0,485
		TBA2	0,110371	0,1077	0,113406	0,107851	0,221667	0,233333
		KATBA	0,1278	0,131668	0,095092	0,110355	0,446667	0,405
		ATBA25	0,206724	0,209418	0,146983	0,162177	0,493333	0,596667
		ATBA40	0,27359	0,303506	0,271386	0,233445	0,623333	0,655
BS=10	Özellik Çıkarma	TBA0	0,298691	0,327343	0,281147	0,284676	0,535	0,565
	Band Seçme	TBA1	0,189647	0,196966	0,179209	0,107477	0,431667	0,498333
		TBA2	0,192228	0,183051	0,191974	0,177613	0,448333	0,486667
		KATBA	0,174576	0,162355	0,173091	0,130171	0,521667	0,503333
		ATBA25	0,282909	0,268274	0,266312	0,199764	0,475	0,63
		ATBA40	0,31543	0,301283	0,318779	0,261653	0,605	0,671667

Boyut indirgeme sonrası DVM sınıflayıcılarında orijinal ve indirgenmiş verilerin eğitim ve sınıflandırma işlemleri için harcanan toplam süre intel i5-3317U işlemcili ve 6 gb RAM'e sahip Windows 8 kurulu bilgisayarda test edilmiştir. MATLAB ortamında "tic-toc" komutu yardımı ile süreler ölçülmüş ve Çizelge 4.3'teki sonuçlar elde edilmiştir.

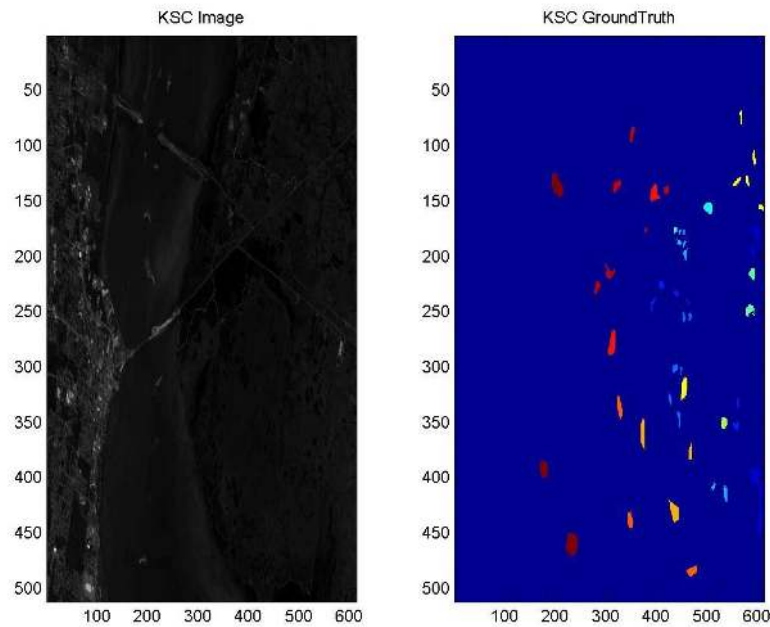
Çizelge 4.3 Sınıflandırma işlem sürelerinin karşılaştırılması

Yöntem	DVM için geçen Süre(sn)	Yöntem	DVM için geçen Süre(sn)
<i>orjinal</i>	0.6154	<i>SR5PCA2</i>	0.0824
<i>SPCA0</i>	0.1931	<i>SR5spcaC</i>	0.0604
<i>SR3PCA0</i>	0.0791	<i>SR5SPCA25</i>	0.0551
<i>SR3PCA1</i>	0.0415	<i>SR5SPCA40</i>	0.0484
<i>SR3PCA2</i>	0.0467	<i>SR10PCA0</i>	0.0831
<i>SR3spcaC</i>	0.0639	<i>SR10PCA1</i>	0.0712
<i>SR3SPCA25</i>	0.0468	<i>SR10PCA2</i>	0.0819
<i>SR3SPCA40</i>	0.0479	<i>SR10spcaC</i>	0.0668
<i>SR5PCA0</i>	0.0707	<i>SR10SPCA25</i>	0.0448
<i>SR5PCA1</i>	0.0521	<i>SR10SPCA40</i>	0.0463

4.3 AVIRIS - KSC Görüntülerinde Boyut İndirgeme ve Sınıflandırma

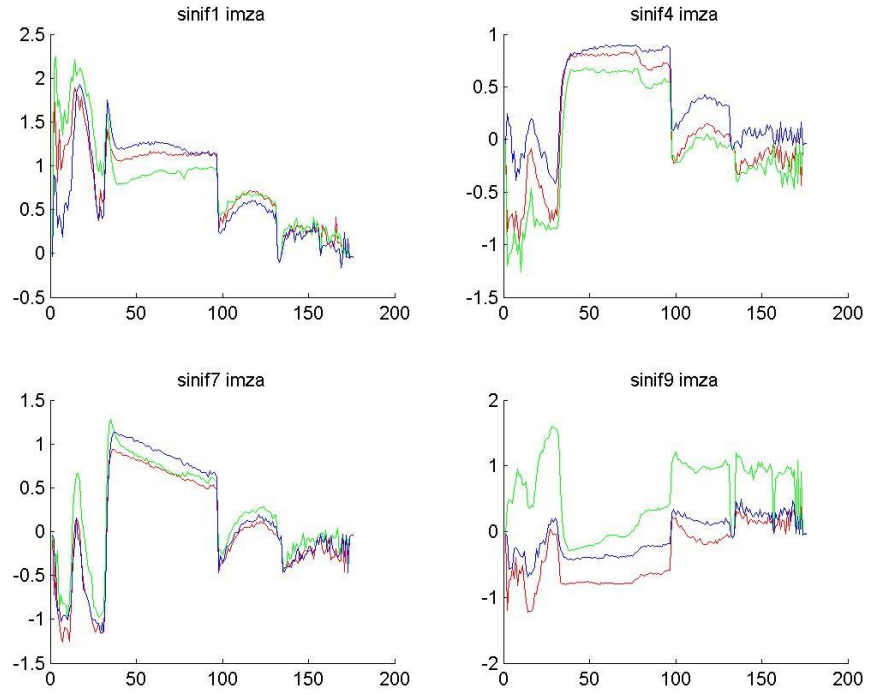
KSC görüntüsü 23 Mart 1996 yılında NASA tarafından AVIRIS sensörü ile çekilmiştir. 400-2500 nm spektral aralığa sahip bu sensör ile alınmış 224 bandlık görüntüdeki gürültülü bandlar çıkarıldığında geriye 176 anlamlı band kalmaktadır.

Şekil 4.25’de KSC verisinin 15’nci bandına ait görüntü ve doğruluk verisi verilmektedir.



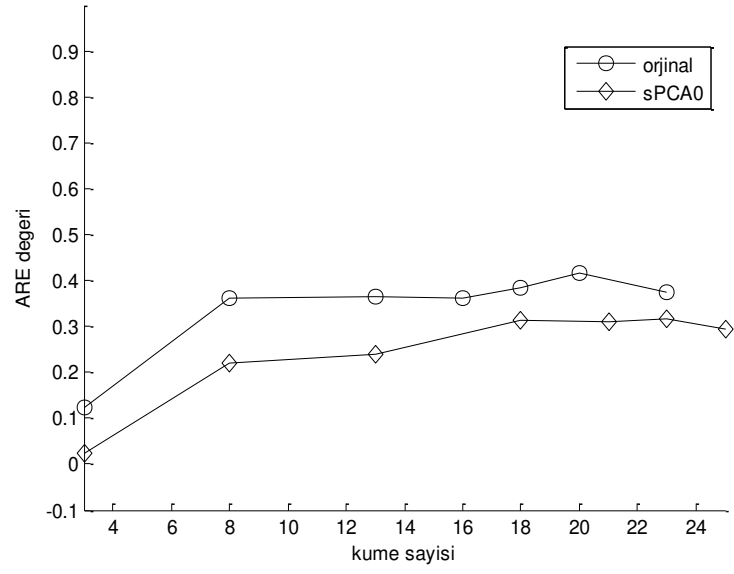
Şekil 4.25 KSC verisine ait band görüntüsü ve doğruluk verisi

KSC verisinden rasgele seçilen örnek ve sınıflardan alınan imzalar Şekil 4.26’de verilmiştir.

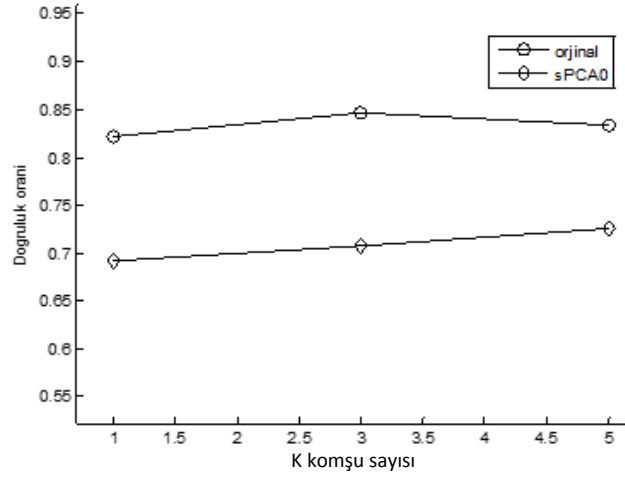


Şekil 4.26 KSC verisine ait spektral imzalar

Indian Pines veri setine uygulanan tüm boyut indirgeme ve sınıflandırma işlemleri KSC verisine de uygulanmıştır. Analizler sonucunda elde edilen sonuçlar Şekil 4.27-Şekil 4.30'deki gibidir.

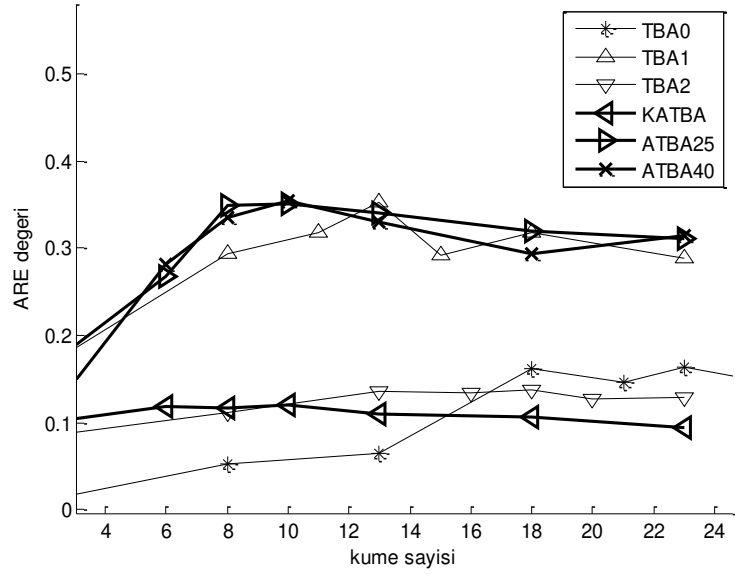


(a)

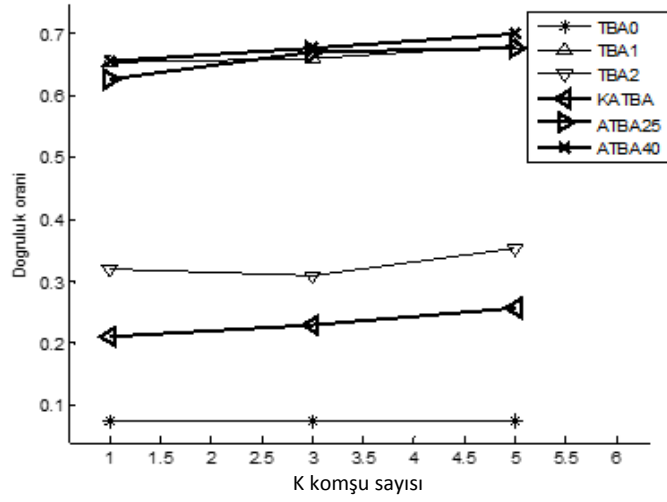


(b)

Şekil 4.27 Boyut sınırlaması olmadan uygulanmış sınıflandırıcı sonuçlarının karşılaştırılması (a) KO (b) KEYK

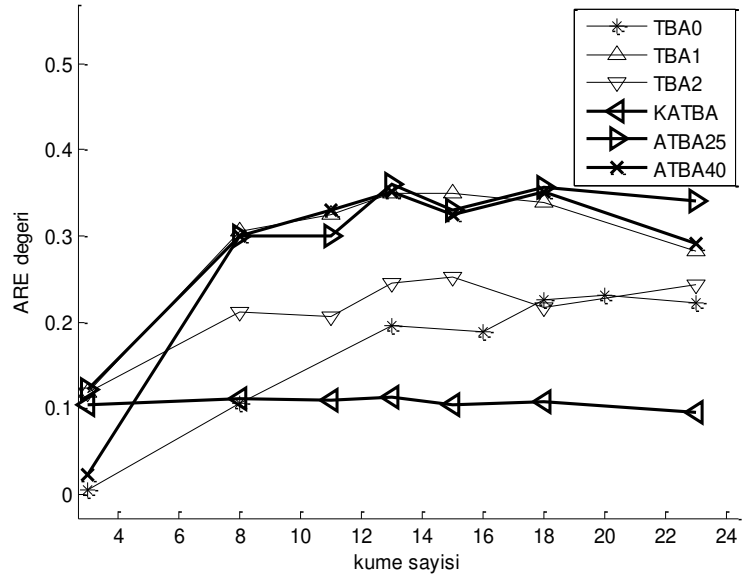


(a)

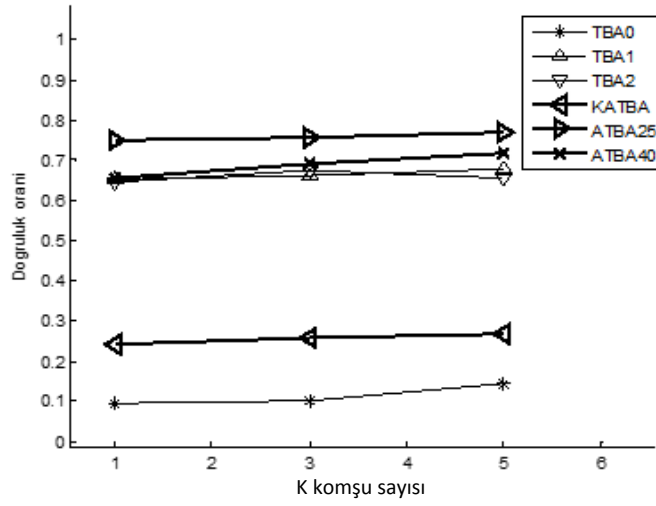


(b)

Şekil 4.28 BS=3 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) KEYK

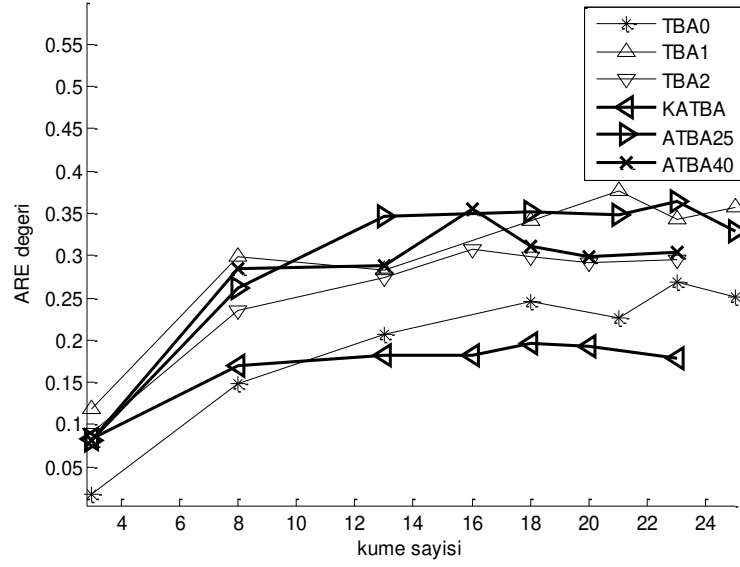


(a)

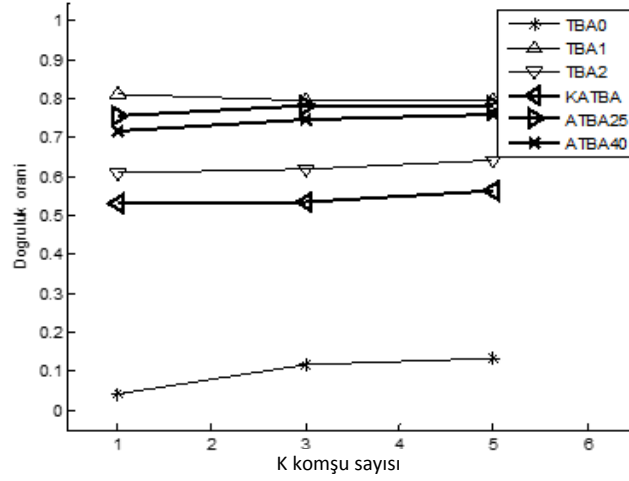


(b)

Şekil 4.29 BS=5 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) KEYK



(a)



(b)

Şekil 4.30 BS=10 için boyut indirgeme yöntemlerinin farklı sınıflandırıcılarda karşılaştırılması (a) KO (b) KEYK

BS değeri 3, 5 ve 10 için verilmiş grafiklerde de görüldüğü gibi gruplama uygulanmış ATBA sonuçları diğer yöntemlerden daha iyi sonuçlar vermiştir. Indian Pines verisinde kısmen iyi sonuçlar veren TBA yöntemi uygulandığında ise çok yetersiz sonuçlar elde edilmiştir.

KSC verilerine uygulanan yöntemlerin karşılaştırmalı sonuçları Çizelge 4.4'de verilmektedir.

Çizelge 4.4 KSC verisi için karşılaştırmalı sonuçlar

Veri Seti KSC			Kümeleme(ARE)				Sınıflama(Doğruluk)	
			KO	Hiyerarşik Bağlantı			KEYK	DVM
				VHYB	OHYB	AHYB		
	Orijinal(176)		0,414644	0,329756	0,253536	0,25626	0,821538	0,870769
	ATBA0(11)		0,316941	0,30663	0,18672	0,209325	0,692308	0,741538
BS=3	Özellik Çıkarma	TBA0	0,163191	0,068571	0,046502	0,050217	0,073846	0,073846
	Band Seçme	TBA1	0,353375	0,300815	0,287616	0,260954	0,655385	0,543077
		TBA2	0,136738	0,134212	0,101154	0,115778	0,32	0,281538
		KATBA	0,120015	0,110925	0,109498	0,099582	0,210769	0,215385
		ATBA25	0,350518	0,337282	0,343835	0,276319	0,627692	0,592308
		ATBA40	0,353891	0,330676	0,265593	0,262897	0,656923	0,549231
BS=5	Özellik Çıkarma	TBA0	0,231013	0,234025	0,197438	0,173531	0,095385	0,087692
	Band Seçme	TBA1	0,349301	0,342333	0,292558	0,267456	0,653846	0,616923
		TBA2	0,25192	0,253129	0,203632	0,179646	0,644615	0,629231
		KATBA	0,111949	0,113333	0,107404	0,101796	0,241538	0,224615
		ATBA25	0,360163	0,348493	0,240142	0,226703	0,749231	0,718462
		ATBA40	0,351325	0,304986	0,250755	0,284111	0,656923	0,616923
BS=10	Özellik Çıkarma	TBA0	0,269158	0,242201	0,052786	0,126684	0,043077	0,084615
	Band Seçme	TBA1	0,377116	0,401375	0,250618	0,288725	0,810769	0,770769
		TBA2	0,307387	0,328695	0,227553	0,253848	0,609231	0,592308
		KATBA	0,196105	0,18322	0,1528	0,14907	0,530769	0,567692
		ATBA25	0,363276	0,35545	0,255539	0,273392	0,755385	0,775385
		ATBA40	0,355512	0,333587	0,227819	0,264558	0,718462	0,723077

Çizelge 4.4'teki sonuçlara bakıldığında ATBA25 ve ATBA40 yöntemleriyle birlikte TBA1 boyut indirgeme yönteminin de başarılı bandları seçtiği görülmektedir.

SONUÇ VE ÖNERİLER

Boyut indirgemedeki en önemli amaç işlem maliyetini azaltmak ve bu esnada sınıflama başarısını koruyabilmek ya da artırmaktır. Boyut miktarı azaldıkça işlem yükü de azalacak, dolayısıyla sınıflama ya da hedef tanıma gibi son adımdaki işlemlerin gerçekleşme süreleri de kısılacaktır.

Bu tez çalışmasında hiperspektral görüntülere Temel Bileşen Analizi ve Ayrık Temel Bileşen Analizi yöntemleri farklı yaklaşımlarla uygulanmış ve band seçimi ile boyut indirgeme yapılmıştır. İşlem parametreleri değiştirilerek adımlar tekrarlanmış ve yöntemler detaylıca analiz edilmiştir. Bu detaylı analiz seçilecek band sayısının 3, 5 ve 10 olarak değiştirilmesini, Indian Pines ve KSC olmak üzere iki farklı veri seti kullanılmasını ve K-en yakın komşuluk, DVM, K-ortalama, Hiyerarşik bağlantı öğrenme metotları ile sınıflandırılarak performans ölçümünü kapsamaktadır.

Indian Pines veri seti hiçbir boyut indirgeme yöntemi uygulanmadan DVM sınıflandırıcıya sokulduğunda Çizelge 4.2’de görüldüğü üzere 72.5% başarı elde edilmiştir. Boyut indirgeme yöntemi olarak ATBA uygulandığında boyut miktarı 78% azalırken sınıflandırma başarısının korunduğu görülmektedir. Seçilecek band sayısı daha da aşağı çekilerek 3, 5 ve 10 banda düşürüldüğünde uygulanan klasik TBA ve klasik ATBA sonuçlarında ise önemli miktarda düşüş yaşanmıştır. Bu düşüşü iyileştirmek için ise ATBA yöntemi üzerinde farklı yaklaşımlar denenmiş ve sonuçların iyileşmesi hedeflenmiştir.

ATBA yönteminin başarısını artırmaya yönelik ilk çalışma yerel maksimum noktaların bulunması olmuştur. Bandlar için ayrıklık katsayıları hesaplandıktan sonra oluşan vektörde yerel maksimum noktalar hesaplanmış ve tepelerin bulunduğu bandlar

seçilerek boyut indirgeme yapılmıştır. Bu yöntem sonucunda Indian Pines verisinde 48 band seçilerek sınıflama yapılmış ve tüm bandların kullanıldığı sınıflama ile neredeyse aynı başarı yakalanmıştır.

Yerel maksimum noktalar belirlendikten sonra da band sayısını azaltmak için çalışmalar devam etmiştir. Bu aşamada bandları gruplama yoluna gidilmiştir. 220 bandlı Indian Pines verisi 25'erli bandlar şeklinde gruplanmış ve her grubun en yüksek değerli yerel maksimum bandı seçilmiştir. Daha sonra seçilen bu bandlardan BS değeri 3, 5 ve 10 olacak şekilde en yüksek ayırıklık değerine sahip bandlar seçilmiştir. Daha sonra gruplama adım aralığı 40 olarak değiştirilmiş ve aynı işlemler ATBA40 yöntemi için denenmiştir. Gruplama yaklaşımı sayesinde farklı dalga boylarının özelliklerini seçme avantajı kullanılmıştır. Çizelge 4.2'ye bakıldığında klasik ATBA yöntemine kıyasla ATBA25 ve ATBA40 yöntemlerinde başarı oranlarının yükseldiği açık şekilde görülebilmektedir. Örneğin BS değeri 5 için KATBA sonrası DVM sınıflama başarısı 0.405 iken ATBA40 başarısı 0.655'dir.

Analizler tekrar tekrar yapıldığında gruplama uygulanmış ATBA yönteminin diğer yöntemlerden daha iyi sonuçlar verdiği görülmüştür. TBA yöntemi sonuçları bazen iyi sonuçlar vermekle beraber bazen de kabul edilemeyecek büyüklükte yetersiz sonuçlar vermektedir. Örneğin KSC verisi sonuçlarına bakıldığında TBA0 sonrası sınıflama başarılarının sıfıra yakın oldukları görülmektedir.

Daha önce de belirtildiği gibi boyut indirgemenin amacı işlem yükünü azaltmak böylece daha hızlı sonuca ulaşabilmektir. Çizelge 4.3'de Indian Pines verisinin farklı boyut indirgeme yöntemleri sonrası sınıflama işlem süreleri verilmiştir. Boyut indirgeme uygulanmadığında geçen işlem süresi 0.499 saniye iken ATBA ile boyut indirgendiğinde geçen DVM sınıflama süresi 0.07 saniyeye düşmüştür. Boyut indirgeme sonrasında işlem yükünün azaldığı buradan da gözlemlenebilmektedir.

Yapılan çalışmada yukarıda değinilen sonuçlar elde edilmekle beraber, çalışma sonrasında ileriye dönük üzerinde çalışılabilecek yeni konuların da ortaya çıktığı görülmüştür. Öncelikle hesaplama maliyetleri açısından boyut indirgemenin gerekliliği gözardı edilemeyecek kadar kritiktir. Bununla birlikte yapılan analizlerin sonucunda boyut indirgenmiş verilerin orjinal veri ile aynı sonuçları yakalayabildiği, hatta bazı

durumlarda da geçtiđi gör÷lmektedir. Analizler daha da ilerletilerek sensöre, dalga boyuna, arazi yapısı ya da sınıf çeşitliliđine göre seçilecek band sayısı, gruplama uygulanmış ATBA tarama adım aralığı gibi değerler daha iyi belirlenerek başarının belirli bir değerin üzerinde olacağı garanti edilebilir.

KAYNAKLAR

-
- [1] Fukunaga, K., (1990). Introduction to Statistical Pattern Recognition, Secon Edition, Academic Press, San Diego
- [2] e-Laser Forestry and Environmental Mapping Solutions, Hyperspectral, www.e-laser.it/eng/research/hyperspectral.php, 10 Ekim 2015
- [3] Manolakis, D., Marden D., ve Shaw, G., (2003). "Hyperspectral Image Processing for Automatic Target Detection Applications", Lincoln Lab. Journal, 14:79-116.
- [4] Nielsen, A., Larsen, R., ve Vestergaard, J., (2011), "Sparse Principal Component Analysis in Hyperspectral Change Detection", SPIE Remote Sensing, 81800S-81800S-6, 6 Ekim 2011.
- [5] Varshney, P., ve Arora, M., (2004). Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data, 1st Edition, Springer-Verlag, Berlin
- [6] Shippert, P., (2004). "Why Use Hyperspectral Imagery?", Photogrammetric Engineering & Remote Sensing, 70(4):377:380
- [7] Burgers, K., Fessehatsion, Y., ve Rahmani, S., (2009). "A Comparative Analysis of Dimension Reduction Algorithms on Hyperspectral Data", LAMBDA Research Group, 7 Ağustos 2009
- [8] Xijun, L., ve Jun, L., (2009). "An Adaptive Band Selection Algorithm for Dimension Reduction of Hyperspectral Images", Image Analysis and Signal Processing, 11-12 April 2009, Taizhou
- [9] Du, H., Qi, H., Wang, X., Ramanath, R., ve Snyder, W.E., (2003). "Band selection using independent component analysis for hyperspectral image processing", Applied Imagery Pattern Recognition Workshop, 15-17 Oct. 2003
- [10] Martinez-Uso, A., Pla, F., Sotoca, J.M., ve Garcia, P., (2007). "Clustering-Based Hyperspectral Band Selection Using Information Measures", Geoscience and Remote Sensing, IEEE Transactions, 45(12):4158-4171.
- [11] Zou, H., Hastie, T., ve Tibshirani, R., (2006). "Sparse Principal Component Analysis", Journal of Computational and Graphical Statistics, 15(2), 265-286
- Hui Zou, "Sparse Principal Component Analysis"
- [12] Shuangjiang, L., ve Hairong, Q., (2011). "Sparse Representation Based Band Selection For Hyperspectral Images", Image Processing (ICIP), 2011 18th IEEE International Conference, 11-14 Sept. 2011, Brussels.

- [13] Plaza, A., Plaza, J., Martin, G., ve Sanchez, S. (2011). *Hyperspectral Data Processing Algorithms*; Derleyen: Prasad, S., Thenkabail, J., ve Huete, J., (2011). *Hyperspectral Remote Sensing of Vegetation*, CRC Press, New York.
- [14] Melgani, F., ve Bruzzone, L., (2004). "Classification of Hyperspectral Remote Sensing Images With Support Vector Machines", *Geoscience and Remote Sensing, IEEE Transactions*, 42(8):1778-1790.
- [15] Landgrebe, D., (1997). "On Information Extraction Principles for Hyperspectral Data"
- [16] Scikit Learn, Clustering, scikit-learn.org/stable/modules/clustering.html, 10 Ekim 2015
- [17] Mathworks, Hierarchical Clustering, www.mathworks.com/help/stats/hierarchical-clustering.html, 11 Ekim 2015
- [18] Mukhopadhyay, A. ve Maulik, U., (2011), "A multiobjective approach to MR brain image segmentation", *Applied Soft Computing*, 11(1):872–880.
- [19] Provost, F., ve Kohavi, R., (1998). "On Applied Research in Machine Learning", *Machine Learning - Special issue on applications of machine learning and the knowledge discovery process*, 30:127-132
- [20] Sokolova, M., ve Lapalme, G., (2009). "A systematic analysis of performance measures for classification tasks", *Information Processing & Management*, 45(4):427:437.
- [21] University of Basque Country, *Hyperspectral Remote Sensing Scenes*, www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes, 10 Ekim 2015.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Burak AKGÜL
Doğum Tarihi ve Yeri : 20.09.1988 - BURDUR
Yabancı Dili : İngilizce
E-posta : akgulb@yildiz.edu.tr

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Lisans	Elektronik Müh.	İstanbul Teknik Üniversitesi	2011
Lise	Fen / Matematik	İzmir Fatih Fen Lisesi	2006

İŞ TECRÜBESİ

Yıl	Firma/Kurum	Görevi
2015 - ...	TUBİTAK Uzay Tekn. Arş. Enst.	Araştırmacı
2014 - 2015	C Tech Bilişim A.Ş.	Donanım Mühendisi
2012 - 2014	Yıldız Teknik Üniversitesi	Araştırma Görevlisi
2011 - 2011	PAVO Tasarım	Tasarım Mühendisi