




## Hiperspektral Görüntülerin Sınıflandırılmasında Farklı Boyut İndirgeme Yöntemlerinin Karşılaştırılması

\*<sup>1</sup>Mehmet Zahid Yıldırım, <sup>2</sup>Caner Özcan, <sup>3</sup>Okan Ersoy

<sup>1</sup>Karabük Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Karabük, Türkiye  
m.zahidyildirim@karabuk.edu.tr, 

<sup>2</sup>Karabük Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Karabük, Türkiye  
canerozcan@karabuk.edu.tr, 

<sup>3</sup>Purdue Üniversitesi, Elektrik ve Bilgisayar Mühendisliği Fakültesi, Indiana, Amerika Birleşik Devletleri  
ersoy@purdue.edu, 

Araştırma Makalesi

Geliş Tarihi: 14.05.2020

Kabul Tarihi: 25.08.2020

### Öz

Günümüzde gittikçe önem kazanan uzaktan algılamada, araştırmacılar çeşitli spektral imzalar arasındaki ilişkileri bulmak için dünyanın yüzeyini temsil eden yüksek boyutlu verileri kullanırlar. Özellikle görüntüler, farklı malzemelerin özelliklerini yansıtan yüzlerce yüksek çözünürlüklü banttan oluşabilirler. Bununla birlikte, yüksek boyutlu uzayda çok sayıda farklı bantların bulunması, bu özelliklerin yorumlanmasını zorlaştırabilmektedir. Uzaktan algılama verilerinin ön-ışlemesi için boyutsallık probleminin bağlı olarak çeşitli zorluklar ile karşılaşmaktadır. Bu alanda ortaya çıkan araştırmalar, bunun zor bir problem olduğunu ve tüm sorunlara tek bir çözüm olmadığını ortaya koymaktadır. Bununla birlikte, son çalışmalar katmanlı uzay öğrenme tekniklerinin hiperspektral görüntülerin ön işleminde çok önemli bir çözüm olduğunu göstermektedir. Bu çalışmada, en güncel katmanlı uzay yerleştirme yöntemlerinin hiperspektral veriler üzerindeki performansı karşılaştırmalı olarak analiz edilmiştir. Her bir yöntemin bu alanda en çok kullanılan iki farklı veri seti kullanılarak boyut indirgeme uygulaması gerçekleştirilmiş ve en yakın komşu (1NN) sınıflandırması ile performansı doğrulanmıştır. Elde edilen sonuçlara göre karşılaştırılan katmanlı uzay yerleştirme yöntemlerinin hiperspektral verilerin sınıflandırılmasında sınıf bazlı farklılıklar olsa da başarılı sonuçlar verdiği görülmektedir. Ayrıca her bir yöntemin çalışma zamanı grafik olarak sunulmuş ve hangi yöntemin daha hızlı çalıştığı sebepleriyle birlikte açıklanmıştır.

**Anahtar Kelimeler:** Boyut indirgeme, Çok katmanlı alan yerleştirme, Hiperspektral görüntü, Sınıflandırma

## Comparison of Different Dimension Reduction Methods in Classification of Hyperspectral Images

\*<sup>1</sup>Mehmet Zahid YILDIRIM, <sup>2</sup>Caner ÖZCAN, <sup>3</sup>Okan ERSOY

<sup>1</sup>Karabuk University, Faculty of Engineering, Department of Computer Engineering, Karabuk, Turkey  
m.zahidyildirim@karabuk.edu.tr

<sup>2</sup>Karabuk University, Faculty of Engineering, Department of Computer Engineering, Karabuk, Turkey  
canerozcan@karabuk.edu.tr

<sup>3</sup>Purdue University, School of Electrical and Computer Engineering, Indiana, United States  
ersoy@purdue.edu

### Abstract

In remote sensing, which is becoming increasingly important today, researchers use high-dimensional data representing the surface of the earth to find relationships between various spectral signatures. In particular, images can consist of hundreds of high-resolution bands that reflect the properties of different materials. However, the presence of a large number of different bands in high-dimensional space can make interpretation of these features difficult. Various difficulties are encountered due to dimensionality problem for pre-processing of remote sensing data. Research in this area reveals that this is a difficult problem and not a single solution to all problems. However, recent studies show that manifold learning techniques are a very important solution in the preprocessing of hyperspectral images. In this study, the performance of the state-of-the-art manifold embedding

\*Sorumlu yazar: Karabük Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Karabük, m.zahidyildirim@karabuk.edu.tr, +90 507 486 1453

methods on hyperspectral data is analyzed comparatively. The dimension reduction application of each method has been carried out by using two different data sets that are used most in this field and their performance have been verified by the nearest neighbor (1NN) classification. Even though there are class-based differences in the classification of hyperspectral data, it is seen that manifold embedding methods, which are compared according to the obtained results, yield successful results. In addition, the runtime of each method is presented graphically and explained along with the reasons for which method works faster.

**Keywords:** Dimension reduction, Multidimensional field embedding, Hyperspectral image, Classification

## 1. GİRİŞ

Birçok gerçek dünya nesnesi, görüntüler, konuşma sinyalleri, videolar ve metin belgeleri yüksek boyutlu gösterimlerle ilişkilendirilebilir. Bu büyük veri setlerini analiz etmek ve işlemek, araştırmacılar için önemli bir çalışma alanıdır. Büyük veri kümeleri, karakter özelliklerini koruyarak daha küçük kümelere dönüştürülmelidir. Yüksek boyutlu bir özellik alanında temsil edilen anlamlı veri kalıplarını bulmak oldukça zordur. Bu zorlukları ele almak için, temel bileşenler analizi [1]-[4], çok boyutlu ölçekleme [5], yerel Fisher diskriminant analizi [6] ve yarı denetimli yerel Fisher diskriminant analizi [7] gibi geleneksel doğrusal temelli boyut indirgeme teknikleri geliştirilmiştir.

Görüntüleme sistemlerinde artan veri boyutları ve detaylar, doğrusal yöntemlerin bu zorlukları ele almak için çoğunlukla yeterli olmadığını göstermiştir. Bu gelişmeler sonucunda, doğrusal olmayan yöntemler ortaya çıkmıştır. Yerel doğrusal yerleştirme yöntemi [8], doğrusal bir dönüşümle komşuluk ilişkisini temsil eder ve her bir komşu bölge için veri noktalarının yeniden yapılandırılmasını sağlar. İzometrik özellik haritalama [9], hem verilerin yerel komşuluğunda hem de parçalı kombinasyonlarında doğrusallık varsayılarak küresel doğrusal olmayan yapılar oluşturmak için önerilmiştir. İzometrik özellik haritalama ilk önce yerel doğrusal yerleştirmede olduğu gibi yerel bir çözüm üretir ve daha sonra çok boyutlu ölçekleme üzerinden doğrusal olmayan veriler için küresel optimal sistemi elde eder. Laplas özharitalar [10], bir katmanlı uzay üzerindeki laplas Beltrami operatörü ve ısı denklemi bağlantıları arasındaki ilişkiyi kullanır. Laplas özharitalar çözüm alanlarının yanı sıra yerel doğrusal yerleştirme, lineer transformasyon ve laplas matrislerinin izleyen öz vektörlerini içerir. Lokal tanjant uzay hizalamada [11], katmanlı uzayın yerel geometrisi, her bir veri noktasında teğet uzayı için bir yaklaşımın oluşturulmasıyla öğrenilir ve küresel katmanlı uzay, daha sonra temel noktaya göre veri noktalarının global koordinatlarını vermek üzere hizalanır. Boyutların azaltılması için etkili bir keşif olan maksimum açılma varyansı [12], yerel mesafelerin korunmasına bağlı varyans yerleştirme haritalarının en üst düzeye çıkarılmasıyla verilerin düşük boyutlu bir şekilde temsil edilmesini sağlar. Yapay gözetimsiz bir sinir ağı olan öz-düzenlemeli özellik haritası [13], yüksek boyutlu verilerin görselleştirilmesi ve analizi için de önerilmektedir.

Son zamanlarda, olasılıksal yöntemler için verimli algoritmaların geliştirilmesine büyük ilgi gösterilmiştir. Rastlantısal komşu yerleştirme (Stochastic neighbor embedding, SNE) [14], veri noktaları arasındaki koşullu

olasılıklar olarak yüksek boyutlu mesafeleri temsil ederek benzerlikleri tanımlamaktadır. t-dağılımlı rastlantısal komşu yerleştirme (t-distribution stochastic neighbor embedding, t-SNE) [15] SNE' nin daha küçük boyutlu uzayda verilerin yerel yapısını koruyan ve önemli bir küresel yapı ortaya koyan bir varyasyondur. Küresel rastlantısal komşu yerleştirme (The spherical stochastic neighbor embedding, sSNE) [16], sabit bir kavisi koordinat sistemi göz önünde bulundurularak, öklit olmayan bir katmanlı uzay içine veri katmayı amaçlayan başka bir SNE varyasyonudur. Çok boyutlu yapay alan yerleştirme (Multidimensional artificial field embedding-MAFE) [17], yüksek boyutlu alandaki komşuluk ilişkilerini korumak için minimum enerji konfigürasyonuna yol açan yeni bir boyut indirgeme çatisıdır. Tüm bu yöntemler, doğrusal olmayan katmanlı uzay verisinde bir koordinat temsili arar ve hiperspektral veriler dahil olmak üzere farklı uygulama alanlarının analizinde önemli bir rol oynar [18]-[20].

Algoritma verimliliği ve belirli veri uygulamalarındaki farklılıklara rağmen, boyut indirgeme yöntemleri, sınıflandırıcı girdi özelliklerini azaltma, daha iyi sıkıştırma ve daha iyi görselleştirme gibi bazı özellikleri miras alır. Her yerleştirme tekniği, doğrusal olmayan katmanlı uzay verisinde bir koordinat temsili bulmaya çalışır.

Bu çalışmada, olasılıksal temelli yöntemlerden SNE ve t-SNE yöntemleri ile kuvvet alan formülasyonuna dayanan çok boyutlu bir alan yerleştirme algoritması olan MAFE yöntemi ve MAFE yönteminin geliştirilmiş bir varyasyonu olan Enhanced-MAFE yöntemleri kullanılarak analizler gerçekleştirilmiştir. Bu yöntemler hiperspektral görüntülerden elde edilmiş olan Botswana Hyperion ve Kennedy Space Center (KSC) verileri üzerine uygulanmıştır. Tüm yöntemlerin kodları Matlab ortamında yazılarak uygulanmıştır. Ayrıca Enhanced-MAFE yönteminin KSC verisi üzerindeki performans değerlendirmesi ilk kez bu çalışmada gerçekleştirilmiştir.

## 2. YÖNTEMLER

Doğrusal olmayan katmanlı uzay yerleştirme yöntemleri hiperspektral görüntülerdeki spektral imzaları daha düşük boyutlarda karakterize etmek için oldukça kullanışlı tekniklerdir. Bu teknikler yüksek boyutlu uzaydaki veriye özgü geometrik imzaları koruyarak bunu daha düşük boyutlu bir uzayda temsil etmeye çalışır. Çalışmada kullanılan tüm yöntemler girdi uzayındaki komşuluk ilişkilerini analiz eder. Bu ilişkiler korunarak daha düşük boyutlardaki çıktı uzayı elde edilir. Bu çıktının girdi uzayı ile aynı sonuçları vermesi hedeflenmektedir.

## 2.1. Rastlantısal Komşu Yerleştirme

SNE algoritması yüksek boyutlu uzayda bulunan verilerin daha düşük bir düzlemde birbirlerine olan konumlarını korumayı amaçlamaktadır [14]. Yüksek boyutlu veri setindeki her bir  $i$  noktasının ve tüm potansiyel komşu  $j$  noktalarının olasılık yoğunluk fonksiyonu,

$$p_{ij} = \frac{\exp(-d_{ij}^2)}{\sum_{k \neq i} \exp(-d_{ik}^2)} \quad (1)$$

$$d_{ij}^2 = \frac{\|x_i - x_j\|^2}{2\sigma_i^2} \quad (2)$$

şeklinde verilmektedir. Burada  $\sigma$  standart sapmayı ifade etmektedir. Eğer veri içerisinde iki nokta birbirlerine yakın ise olasılık yoğunluk fonksiyonu değeri yüksektir. Benzer şekilde düşük boyutlu çıkış uzayındaki her  $i$  noktasının  $j$  komşulukları ise denklem (3)' deki gibi verilmektedir.

$$q_{ij} = \frac{\exp(-\|y_i - y_j\|^2)}{\sum_{k \neq i} \exp(-\|y_i - y_k\|^2)} \quad (3)$$

Yüksek boyutlu veri setindeki veriler, düşük boyuta indirildiğinde (1) ve (3) denklemlerinin birbirine eşit olması beklenmektedir. Bu, her bir nesne için komşular üzerinde orijinal ( $p_{ij}$ ) ve indirgenmiş ( $q_{ij}$ ) dağılımları arasındaki Kullback-Leibler (KL) sapmalarının toplamı olan bir maliyet fonksiyonunu en aza indirerek elde edilir:

$$C = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} = \sum_i KL(P_i || Q_i) \quad (4)$$

Bu denklemin gradyan iniş algoritması kullanılarak indirgenmiş hali ise denklem (5) de verilmektedir.

$$\frac{\delta C}{\delta y_i} = 2 \sum_j (y_i - y_j)(p_{ij} - q_{ij} + p_{ji} - q_{ji}) \quad (5)$$

## 2.2. T-Dağılımlı Rastlantısal Komşu Yerleştirme

t-SNE yöntemi, SNE yönteminin daha düşük boyutlu uzayda verilerin yerel yapılarını koruyan ve önemli bir küresel yapı ortaya koyan bir varyasyonudur [15]. Yüksek boyutlu uzay için SNE ile aynı formülasyonu kullanan t-SNE çıkış uzayı için denklem (6)' daki formülasyonu kullanmaktadır.

$$q_{ij} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq i} (1 + \|y_i - y_k\|^2)^{-1}} \quad (6)$$

SNE yönteminin gauss dağılımından farklı olarak burada t-dağılımı kullanılmaktadır. Maliyet fonksiyonun gradyan iniş algoritması ile indirgenmiş hali ise denklem (7)' de verilmektedir.

$$\frac{\delta C}{\delta y_i} = 4 \sum_j (p_{ij} - q_{ij})(y_i - y_j)(1 + \|y_i - y_j\|^2)^{-1} \quad (7)$$

## 2.3. Çok Boyutlu Yapay Alan Yerleştirme

MAFE, yüksek boyutlu alandaki komşuluk ilişkilerini korumak için minimum enerji konfigürasyonuna yol açan yeni bir boyut indirgeme çatısıdır [17]. Potansiyel bir alan için çok boyutlu alan yerleştirme veya kuvvet alanı yöntemi, en uygun enerji çözümünü bulmak için toplam çekme ve itme potansiyel fonksiyonlarını en aza indirmek olarak yorumlanır. Her yüksek boyutlu veri noktasının daha düşük boyutlu bir alanda bir haritaya sahip olduğu varsayılmaktadır. Optimum yerleştirme haritası, her iki kuvvet alanları tarafından belirlenen grafiğin denge durumuna karşılık gelir. Bir dizi kriter, kuvvetlerin etkileşimi sırasında potansiyel işlevlerin seçimini yönlendirir. İtme potansiyel fonksiyonu kısa mesafelerde baskındır ve bu nedenle yerel ilişkileri temsil ederken çekme potansiyel fonksiyonu uzun mesafelerde baskındır ve küresel ilişkileri temsil eder. Bu şekilde, boyutsallık azaltma sorunu, yerleştirme ilişkilerini koruyan minimum enerji konfigürasyonunu bulmak olarak formüle edilmektedir.

Boyut indirgeme yöntemleri,  $Y = \{y_1, y_2, \dots, y_m\}$  yüksek boyutlu veri kümesini  $Z = \{z_1, z_2, \dots, z_n\}$  daha düşük boyutlu veriye dönüştürür. Kuvvet alanı formülasyonlarına göre, her bir harita için hareket dinamikleri şu şekilde yazılabilir:

$$\dot{z}_i = - \sum_{j \neq i} \{ \nabla_{z_i} U_a^{ij}(\|z_i - z_j\|) - \nabla_{z_i} U_r^{ij}(\|z_i - z_j\|) \} \quad (8)$$

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

$$U_i(Z) = \sum_{j \neq i} \{ U_a^{ij}(\|z_i - z_j\|) - U_r^{ij}(\|z_i - z_j\|) \} \quad (9)$$

olarak ifade edilebilir. Komşuluk grafiğindeki toplam üst üste gelme potansiyeli ise şöyle tanımlanır:

$$U(Z) = \sum_{i=1}^N U_i(Z) \quad (10)$$

Kararlı durum için genel optimizasyon problemi ise,

$$Z^* = \operatorname{argmin}_Z U(Z) \quad (11)$$

olarak ifade edilebilir. Burada  $Z^*$  daha düşük boyutlu uzayda optimum  $Z$  olur. Minimum enerji yakınsaması sonucunda, itme ve çekme kuvvetleri dengelenmiştir.

Maliyet fonksiyonunu en aza indirmek için, her bir piksel için yüksek boyutlu ve düşük boyutlu uzaylar arasındaki KL sapmalarının toplamının gradyanı kullanılır. Bu, yüksek boyutlu uzay ile düşük boyutlu uzay arasındaki bağlantıyı verir. KL ayrılmalarının gradyanı ile birlikte maliyet fonksiyonu,

$$\dot{z}_i = - \sum_{j \neq i} (z_i - z_j) \xi_a w_{ij} p \|z_i - z_j\|^{p-2} - \xi_r q \|z_i - z_j\|^{q-2} \exp\left\{-\frac{\|z_i - z_j\|^q}{\sigma}\right\} \quad (12)$$

olarak ifade edilmektedir. Burada  $\xi_a$  ve  $\xi_r$  sırasıyla çekme ve itme kuvvetlerinin büyüklük parametreleridir.

#### 2.4. Geliştirilmiş Çok Boyutlu Yapay Alan Yerleştirme

Enhanced-MAFE [21], MAFE yönteminin bir varyasyonudur. Yüksek boyutlu veri kümesi MAFE ile aynı şekilde kullanılırken, düşük boyutlu çıktı uzayındaki formülasyondaki geliştirme ile MAFE yöntemine göre Botswana verileri üzerinde daha yüksek doğruluk ve daha iyi görselleştirme sağlanmıştır. Maliyet fonksiyonu düşük boyutlu uzaydaki değişim ile beraber şu şekilde tanımlanmaktadır.

$$\dot{z}_i = -2 \sum_{j \neq i} (z_i - z_j) \left\{ \|z_i - z_j\|^{-1} (\xi_a p_{ij} - \xi_r q_{ij}) \right\} \quad (13)$$

burada çekme ve itme kuvvetlerinin büyüklük parametreleri deneysel olarak 0 ile 1 arasında seçilmektedir. Bu aralık, bu kuvvetler arasında güçlü bir ilişki sağlamaktadır. Önerilen yöntemin üstünlüğü bu yeni maliyet fonksiyonundan gelmektedir.

### 3. DENEYSEL SONUÇLAR

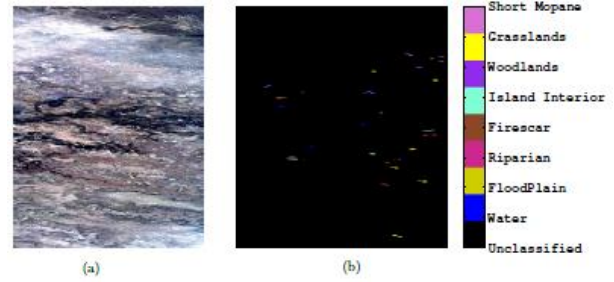
Sunulan yöntemlerin performanslarını değerlendirmek için, 9 farklı sınıfa sahip Botswana Hyperion verileri kullanılmıştır.

**Tablo 1.** Botswana verisi için sınıf sayıları

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

Genel sınıf grupları Tablo 1' de gösterilmektedir. NASA EO-1 uydusu tarafından 2001-2004 yılları arasında elde edilen görüntülerde, içerisinde yanlış kalibre edilmiş detektörlerden gelen veriler ve gürültülü bantlar

çıkarıldıktan sonra, sınıflandırma çalışmaları için 145 banttan oluşan gözlemler sunulmaktadır.

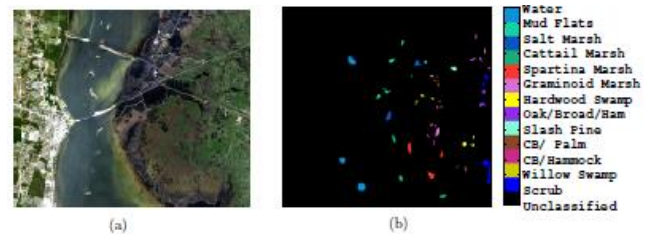


**Şekil 1.** (a) Okavango deltası, Botswana (b) Doğrulama verisi

Şekil 1' de Güney Afrika ülkelerinden biri olan Botswana üzerinde bulunan Okavango deltasına ait uydu ve doğrulama verisi görselleri gösterilmektedir. Botswana verisi ile birlikte, 13 farklı sınıfa sahip KSC verileri kullanılmıştır. Sınıflandırma grupları Tablo 2'de verilmektedir. NASA AVIRIS uçağı ile elde edilen görüntülerde, Su absorpsiyonu ve düşük çözünürlüklü bantlar çıkarılarak, sınıflandırma çalışmalarında 176 banttan oluşan veriler kullanılmaktadır.

**Tablo 2.** KSC verisi için sınıf sayıları

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



**Şekil 2.** (a) KSC renkli görüntüsü (b) Doğrulama verisi

Şekil 2' de Amerika'nın Florida eyaletinde bulunan uzay merkezine ait uydu ve doğrulama verisi görselleri gösterilmektedir. Deneysel sonuçlar her bir sınıf için ayrı olarak sunulmaktadır. Her algoritmanın parametreleri, en iyi sınıflandırma doğruluğunu elde etmek için belirlenmiştir. Bu çalışmada, tanımlanmış nesne sınıflarının veya toprak arazi

örtüsünün, alt boyut uzayında ayrılabilir hacimler kullanıp kullanmadığını ve başarılı bir şekilde ayırt edilemeyeceğini belirlemek için Matlab’ da tanımlı en yakın komşu (1NN) sınıflandırma performans doğruluğu kullanılmıştır.

Deneysel çalışmalarda farklı yerleştirme uzayları kullanılmış ve gömülü haritalar rastgele eğitim (%70) ve test (%30) kümelerine bölünmüştür. Tablo 3 ve Tablo 4’ de verilerde bulunan tüm sınıflara ait doğruluk oranları 1NN yöntemi kullanılarak elde edilmiştir. Ayrıca toplam doğruluk oranları için Kappa istatistiği (KS) [22] kullanılarak sınıflandırma doğruluklarının ortalaması alınmıştır. KS şu şekilde hesaplanmaktadır:

$$KS = \frac{N \sum_{c=1}^{|C|} t_{cc} - \sum_{c=1}^{|C|} t_{c+} t_{+c}}{N^2 - \sum_{c=1}^{|C|} t_{c+} t_{+c}} \quad (14)$$

Burada  $N$  test örneklerinin sayısını,  $t_{cc}$  C sınıfındaki doğru sınıflandırılmış örnek sayısını,  $t_{c+}$  C sınıfı olarak etiketlenmiş test örneklerinin sayısını ve  $t_{+c}$  C sınıfına ait olduğu tahmin edilen örnek sayısını göstermektedir.  $|C|$ , verilerdeki toplam sınıf sayısını belirtmektedir. KS tesadüfen doğru sınıflandırma olasılığını hesaba kattığı için basit yüzdelik doğruluk hesaplamasından daha güçlü bir ölçü olarak bilinir.

**Tablo 3.** Botswana verisi için sınıflandırma sonuçları

Etiket	Doğruluk Oranları (%)			
	SNE	t-SNE	MAFE	Enhanced MAFE
C1	<b>100</b>	<b>100</b>	99.89	<b>100</b>
C2	<b>100</b>	<b>100</b>	<b>100</b>	99.58
C3	76.19	90.91	90.48	<b>97.70</b>
C4	94.74	94.74	79.46	<b>100</b>
C5	94.44	88.89	<b>99.17</b>	96.76
C6	68.42	90	94.19	<b>100</b>
C7	96.18	93.75	93.24	<b>100</b>
C8	91.67	<b>100</b>	<b>100</b>	<b>100</b>
C9	91.67	90.91	90.91	<b>99.13</b>
<b>KS</b>	89.26	93.55	94.27	<b>99.07</b>

Botswana verisi için Tablo 3’ de verilen sonuçlar incelendiğinde En iyi KS sonucunu Enhanced-MAFE yönteminin verdiği görülmektedir. Ayrıca 9 sınıfın 7’ sinde yine Enhanced-MAFE diğer yöntemlerden daha başarılı sonuçlar vermektedir. Diğer iki sınıfta ise kıyaslanan yöntemlere yakın sonuçlar aldığı görülmektedir. Sınıf ayrımlarının en zor olduğu sınıflar C3 ve C6 sınıflarıdır. C1, C2 ve C8 sınıfları ise tüm yöntemler için etiketlenmesi kolay sınıflar olarak görünmektedir.

Tablo 4’ te verilen KSC verileri için KS oranı tüm yöntemlerde yakın olmakla beraber en başarılı yöntem Enhanced-MAFE olarak görülmektedir. Enhanced-MAFE 13 sınıfın 8’ i için en iyi sınıflandırma doğruluk oranlarına sahiptir. C1, C7 ve C8 sınıfları için SNE, C2 sınıfında t-SNE, C4 sınıfında ise MAFE yöntemi en yüksek doğruluk

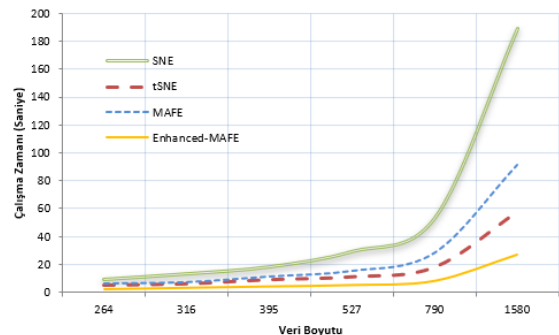
oranlarını vermektedir. Sınıf ayrımlarının en zor olduğu sınıflar C4, C5, C6 ve C8 sınıflarıdır. C13 sınıfı tüm sınıfların tam doğruluk oranı gösterdiği tek sınıftır.

**Tablo 4.** KSC verisi için sınıflandırma sonuçları

Etiket	Doğruluk Oranları (%)			
	SNE	t-SNE	MAFE	Enhanced MAFE
C1	<b>91.20</b>	90.34	65.26	68.54
C2	85.34	<b>87.04</b>	79.03	77.42
C3	80.40	78.33	94.52	<b>98.61</b>
C4	51.90	45.61	<b>68.75</b>	64.62
C5	41.20	41.03	93.33	<b>95.24</b>
C6	39.30	35.71	26.98	<b>44.26</b>
C7	<b>82.77</b>	80.77	35.29	67.65
C8	<b>63.62</b>	60	43.75	42.98
C9	93.56	91.74	<b>99.30</b>	<b>99.30</b>
C10	93.68	93.68	95.08	<b>97.56</b>
C11	93.81	93.81	99.11	<b>100</b>
C12	81.91	80.51	97.60	<b>98.39</b>
C13	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
<b>KS</b>	81.90	80.10	79.92	<b>82.51</b>

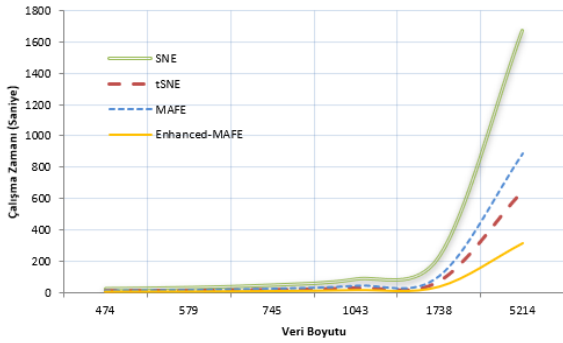
Tablo 3 ve Tablo 4’ te verilen sonuçlar için literatürdeki diğer çalışmalarda olduğu gibi Botswana verisi 1/3, KSC verisi ise 1/11 oranında örneklenerek sonuçlar elde edilmektedir.

Her iki tabloda da sınıflar bağımsız olarak incelendiğinde ağaçlık, dere kenarı ve bataklık bölgelerdeki sınıflandırmanın zor olduğu görülmektedir. Su ise ayırt edilebilmesi en kolay sınıf olarak karşımıza çıkıyor. Ayrıca her iki veri seti içinde Enhanced-MAFE yönteminin sınıflar arasındaki ayrımı oldukça başarılı yapabildiği gözlemlenmektedir. Özellikle ayrımı zor sınıflar için en yüksek doğruluk oranlarını büyük çoğunluğunda Enhanced-MAFE yöntemi vermektedir.



**Şekil 3.** Botswana verisi için çalışma zamanı

Şekil 3’ de verilen veri boyutuna bağlı çalışma zamanı grafiği incelendiğinde, yöntemlerin veri boyutları arttıkça giderek ivmelenen bir çalışma zamanı artış grafiği izledikleri görülmektedir. SNE yönteminin büyük boyutlardaki veri gruplarını analiz etmede diğer yöntemlere kıyasla çalışma zamanı oldukça yüksektir. Enhanced-MAFE diğer yöntemlere göre daha az etkilenen bir grafik sunmaktadır.



Şekil 4. KSC verisi için çalışma zamanı

Şekil 4’ te verilen KSC verileri de Botswana ile benzer şekilde SNE yönteminin çalışma zamanı, veri boyutu arttıkça artan bir ivme ile yükselmektedir. Enhanced-MAFE, KSC verileri içinde en hızlı yöntem olarak karşımıza çıkmaktadır. KSC veri boyutunun Botswana’ya göre daha büyük olmasına rağmen karşılaştırılan yöntemlerin çalışma zamanlarının benzer özellikler gösterdiği görülmektedir.

Uygulanan yöntemlerin her birinde yüksek boyuttaki veriler, gömülü uzayda gradyan iniş algoritması ile optimum uzaklıklara erişinceye kadar yakınsamaktadır. Optimum yakınsama için her yöntem farklı iterasyonlarda çalışmaktadır. Yöntemlerin çalışma zamanları hesaplanırken her bir yöntem için en yüksek doğruluk oranlarına ulaştıkları iterasyon sayıları kullanılmıştır. Bu anlamda incelendiğinde Enhanced-Mafe yöntemi çok daha düşük iterasyon sayısı ile maksimum doğruluk oranına ulaşmaktadır. Bu da çalışma zamanına olumlu bir katkı sağlamaktadır. Verilerin örnekleme doğruluk oranlarını etkileyen en önemli faktörlerden biridir. Örneklem yapmadan tüm verilerin seçilmesi ise çalışma zamanını oldukça arttırmaktadır. Hem Şekil 3 hem de Şekil 4’ te verilen grafikler incelendiğinde yöntemlerin düşük boyutlu veriler kullanıldığında benzer çalışma zamanı gerektirdikleri, ancak boyutlar arttıkça bu durumun değiştiği görülmektedir.

#### 4. SONUÇLAR VE TARTIŞMA

Bu çalışmada hiperspektral veri sınıflandırması için potansiyel alanlara yönelik katmanlı uzay yerleştirme algoritmalarının karşılaştırmalı performansları incelenmiştir. Deneysel çalışmalarda Botswana ve KSC hiperspektral veri setleri kullanılmıştır. Veri setleri üzerinde 4 farklı boyut indirgeme yönteminin doğruluk oranı ve çalışma zamanı performansları gözlemlenmektedir. Her iki veri setinde de Enhanced-MAFE yönteminin diğer yöntemlerden daha yüksek sınıflandırma başarısı elde ettiği gözlemlenmektedir. Veri setlerindeki sınıflar bağımsız değerlendirildiğinde ise ayırt edilmesi en zor olan ağaçlık, bataklık ve dere kenarı bölgelerde yine Enhanced-MAFE yönteminin sınıflandırma başarısı diğer yöntemlere kıyasla daha iyi sonuçlar vermektedir.

Hiperspektral veriler özellikle büyük boyutları sebebiyle işlenmesi zor ve zaman almaktadır. Bu anlamda yöntemlerin veri boyutuna bağlı olarak çalışma zamanlarında artan

oranlarla yükselmeler gözlemlenmiştir. Burada Enhanced-Mafe yöntemi en iyi çalışma zamanı değerlerini elde etmiştir. Bunun en önemli sebebi yöntemin en yüksek doğruluk oranına çok daha az iterasyon ile ulaşabilmesidir. Bunlara ek olarak, çalışma zamanı grafikleri her iki veri setinde de oransal olarak benzer sonuçlar elde edilmiştir. Bu sonuçlar bize yöntemlerin çalışma zamanının veriden bağımsız olduğunu da göstermektedir.

Çalışmada gözlemlediğimiz en önemli hususlardan birisi de yöntemlerde kullanılan parametrelerdeki ufak değişimlerin sonuçları çok fazla etkilemesidir. Gelecekteki çalışmalarda bu parametrelerin optimizasyonunu gerçekleştiren yaklaşımlar yöntemlerin başarı oranlarında ve çalışma zamanlarında daha iyi sonuçlar elde edilmesine olanak tanıyacaktır.

#### KAYNAKÇA

- [1] K. Pearson, “On lines and planes of closest fit to systems of points in space,” *Philosophical Magazine*, vol. 2, no. 11, pp. 559-572, 1901.
- [2] H. Hotelling, “Analysis of a complex of statistical variables into principal components,” *Journal of Educational Psychology*, vol. 24, pp. 417-441, 1933.
- [3] I. T. Jolliffe, “Principal component analysis in regression analysis,” *Springer*, vol. 2, pp. 129-155, 1986.
- [4] A. M. Martinez and A.C. Kak, “Pca versus lda,” *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 228-223, 2001.
- [5] W. Torgerson, “Multidimensional scaling I: theory and method,” *Psychometrika*, vol. 17, pp. 401-419, 1952.
- [6] M. Sugiyama, “Dimensionality reduction of multimodal labeled data by local fisher discriminant analysis,” *Journal of Machine Learning Research*, vol. 8, pp. 1027-1061, 2007.
- [7] M. Sugiyama, T. Ide, S. Nakajima and J. Sese, “Semi-supervised local fisher discriminant analysis for dimensionality reduction,” *Machine Learning*, vol. 78, no. 1-2, pp. 35-61, 2010.
- [8] S. Roweis and L. K. Saul, “Nonlinear dimensionality reduction by locally linear embedding,” *Science*, vol. 290, pp. 2323-2326, 2000.
- [9] J. B. Tenenbaum, V. De Silva and J. Langford, “A global geometric framework for nonlinear dimensionality reduction,” *Science*, vol. 290, pp. 2319-2323, 2000.
- [10] M. Belkin and P. Niyogi, “Laplacian eigenmaps for dimensionality reduction and data representation,” *Neural Computation*, vol. 15, no. 6, pp. 1373-1396, 2003.
- [11] Z. Zhag and H. Zha, “Principal manifolds and nonlinear dimensionality reduction via tangent space alignment,” *SIAM Journal on Scientific Computing*, vol. 26, no. 1, pp. 313-338, 2004.
- [12] L. Song, A. Smola, K. Borgwardt and A. Gretton, “Colored maximum variance unfolding,” *Advances in Neural Information Processing Systems*, vol. 21, pp. 1385-1392, 2008.
- [13] T. Kohonen and T. Honkela, “Kohonen network,” *Scholarpedia*, vol. 2, no. 1, pp. 1568, 2007.

- [14] G. Hinton and S. Roweis, "Stochastic neighbor embedding," *Advances in Neural Information Processing System*, vol. 15, pp. 857-864, 2002.
- [15] L. V. D. Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of Machine Learning Research*, vol. 9, pp. 2579-2605, 2008.
- [16] D. Lungu and O. Ersoy, "Spherical stochastic neighbor embedding of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 2, pp. 857-871, 2013.
- [17] D. Lungu and O. Ersoy, "Multidimensional artificial field embedding with spatial sensitivity," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 2, pp. 1518-1532, 2014.
- [18] C. M. Bachmann, T. L. Ainsworth and R. A. Fusina, "Exploiting manifold geometry in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 441-454, 2005.
- [19] H. Huang and Y. Huang, "Improved discriminant sparsity neighborhood preserving embedding for hyperspectral image classification," *Neurocomputing*, vol. 136, no. 1, pp. 224-234, 2014.
- [20] H. Huang, F. Luo, J. Liu and Y. Yang, "Dimensionality reduction of hyperspectral images based on sparse discriminant manifold embedding," *ISPRS Journal of Photogrammetry Remote Sensing*, vol. 106, pp. 42-54, 2015.
- [21] C. Ozcan and O. Ersoy, "Enhanced multidimensional field embedding method by potential fields for hyperspectral image classification and visualization," *Electronics Letters*, vol. 54, no. 12, pp. 756-758, 2018.
- [22] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37-46, 1960.