

Headwall Hyperspec VNIR Kamerası İle Elde Edilen Hiperspektral Hava Fotoğrafı İçin Boyut İndirgeme Yöntemlerinin Performanslarının Analizi

Esra Tunç GÖRMÜŞ^{*1}, Özlem AKAR²

¹Karadeniz Teknik Üniversitesi, Mühendislik Fakültesi, Harita Mühendisliği Bölümü, Trabzon
(etuncgormus@ceng.ktu.edu.tr) ORCID ID 0000-0002-7667-199X

²Erzincan Binali Yıldırım Üniversitesi, Meslek Yüksek Okulu, Mimarlık ve Şehir Planlama Bölümü, Harita ve Kadastro
Programı, Erzincan
(oakar@erzincan.edu.tr) ORCID ID 0000 – 0001 – 6381 – 4907

Öz

Çalışmada yaygın olarak kullanılan farklı boyut indirgeme yöntemlerinin performanslarının Headwall Hyperspec VNIR Kamerası ile elde edilen hiperspektral hava fotoğrafı için analizi amaçlanmıştır. Bunun için Harita Genel Müdürlüğü tarafından, BHKPK-Bilimsel Araştırma ve Koordinasyon Komisyonunca (BARKOK) 2017 yılında Bergama bölgesinde 2100 metre ortalama yükseklikten Headwall Hyperspec VNIR kamerası ile elde edilen hiperspektral hava fotoğrafı kullanılmıştır. Çalışma alanı olarak, çekimin gerçekleştiği Bergama bölgesinde kentsel ve kırsal arazi kullanım sınıflarını barındıran pilot bir bölge seçilmiştir. Çalışmada öncelikli olarak hiperspektral hava fotoğrafına atmosferik ve radyometrik düzeltmeler yapılmıştır. Ardından bozuk ve kullanılmayan bantların temizlenmesi için yaygın olarak kullanılan Temel Bileşen Analizi (PCA), Ayrık Dalgacık Dönüşümü (DWT), Faktör Analizi, Çok Boyutlu Ölçekleme yöntemi (MDS – Multidimensional Scaling) ve En Fazla Daralan Metrik Öğrenme (MCML-Maximally Collapsing Metric Learning) gibi 5 farklı yöntemle bant indirgeme işlemi uygulanmıştır. Sonrasında tüm görüntülerin geometrik düzeltmeleri yapılarak aynı koordinat sisteminde olması sağlanmıştır. Görüntüler ön işlemden geçirildikten sonra indirgeme yöntemlerinin performanslarının analizi için yaygın olarak kullanılan ve yüksek sınıflandırma doğruluğu sağlayan makine öğrenme yöntemlerinden Destek Vektör Makineleri (DVM) ve Rastgele Orman (RO) sınıflandırıcıları kullanılmıştır. Sınıflandırma işlemi sonucunda elde edilen tematik görüntüler için doğruluk analizleri yapılmış ve elde edilen doğruluklar karşılaştırılmıştır. Çalışmanın sonuçları incelendiğinde, en yüksek sınıflandırma doğruluğunu % 91.67 (RO) doğrulukla DWT boyut indirgeme yönteminin sağlandığı gözlemlenmiştir. Genel olarak bakıldığında RO ile elde edilen sonuçlarda genel sınıflandırma doğrulukları yüksekten düşüğe sırasıyla DWT, MCML, MD, FACTORNAL ve PCA olarak elde edilmiştir. DVM sonuçlarına göre bu sıranın ise DWT, MCML, MD, PCA ve FACTORNAL olduğu gözlemlenmiştir. Buna göre her iki sınıflandırma yönteminde de en yüksek doğruluğu DWT'nin, en düşük sınıflandırma doğruluğunu da FACTORNAL ve PCA yöntemlerinin verdiği görülmüştür. Çalışma sonucu göstermektedir ki, bu hiperspektral hava fotoğrafı için denenilen yöntemler arasında en uygun boyut indirgeme yöntemi DWT'dir.

Anahtar Kelimeler: Hiperspektral hava fotoğrafı, Ayrık Dalgacık Dönüşümü, Destek vektör makineleri, Rastgele Orman, Boyut indirgeme

* Esra Tunç GÖRMÜŞ

Investigating the Performance Of Dimensionality Reduction Methods for Hyperspectral Aerial Photo Obtained from Headwall Hyperspec VNIR Camera

Abstract

In this study it is aimed to analyse the performances of dimensionality reduction methods in processing the hyperspectral aerial photos that obtained from Headwall Hyperspec VNIR camera. Data have been produced by General Directorate of Maps in 2017 in Bergama region. Aerial photos have been taken from 2100 m height by Headwall Hyperspec VNIR camera. As the exact area is very big, only one part of the area which includes both urban and rural classes was chosen as a study area in this work. First of all radiometric and atmospheric corrections were conducted. Then, 5 of commonly used dimensionality reduction methods like Principle Component Analysis (PCA), Discrete Wavelet Transform (DWT), Factor Analysis, Multi-Dimensional Scaling (MDS) and Maximally Collapsing Metric Learning (MCML) were exploited. Next, the photos were geometrically corrected in order to get them in the same coordinate system. After pre-processing steps, two different classification methods belong to the machine learning systems, namely; Support Vector Machine (SVM) and Random Forest (RF) were exploited to investigate the performances of dimensionality reduction methods. Accuracy assessment was carried out and compared according to the classification performance results. When the results were investigated, it was seen that DWT method has the highest classification accuracy of 91.67% with RO. In general classification accuracies of RO method were listed as; DWT, MCML, MD, Factor Analysis and PCA in decreasing order. Whereas in SVM method, classification accuracies were ordered as DWT, MCML, MD, PCA and Factor Analysis in decreasing order. According to these results it was seen that DWT is the dimensionality reduction method with the highest classification accuracy for both classification methods and PCA and Factor Analysis methods are the least successful ones. Consequently, this study showed that DWT is the best dimensionality reduction method for hyperspectral aerial photos amongst other methods used in this study.

Keywords: *Hyperspectral Aerial Photo, Discret Wavelet Transform, Support Vector Machine, Random Forest, Dimensionality Reduction,*

1. GİRİŞ

Uzaydan ya da havadan çekilen hiperspektral görüntüler son yıllarda diğer optik görüntüler yerine ya da onları tamamlayıcı olarak sıkça kullanılmaya başlanmıştır. Yüksek spektral çözünürlüklü hiperspektral görüntüler yeryüzündeki geniş alanların izlenmesi, birbirine benzeyen tarımsal ürünlerin sınıflandırılması, jeolojik materyallerin ayırılması, orman yönetimi ve şehir planlaması (Somers,2013; Liu vd., 2017) gibi birçok uzaktan algılama uygulamalarında başarılı bir şekilde kullanılmaktadır. Ancak verinin yüksek boyutlu olması, yeteri kadar eğitim verisinin

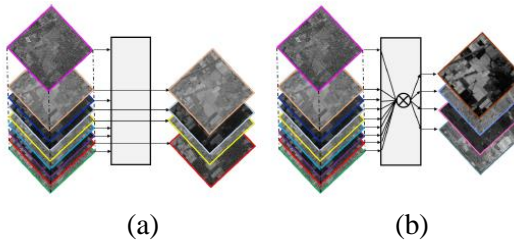
olmaması, katışimli pikseller, farklı geometrik ve atmosferik bozukluklar gibi bazı faktörler hiperspektral görüntülerin mevcut yöntemler ile etkin bir biçimde işlenmesini ve analiz edilmesini zorlaştırmaktadır (Ghamisi P. vd.,2017). Belirtilen faktörlerin her biri araştırmacıların üzerinde çalıştıkları geniş başlıklardır. Bu çalışmada ise hiperspektral görüntülerin yüksek boyutlu olmasından dolayı meydana gelen olumsuzlukları ortadan kaldırmak ve veriyi sonraki işlem adımları için hazırlayabilmek için hiperspektral görüntülerin boyutlarının indirgenmesi ele alınacaktır.

Hiperspektral görüntülerin boyut indirgemesi, yüksek boyutlu verinin özetlenerek, orijinali

kadar anlamlı, daha düşük bir boyutta temsil edilmesini sağlayan dönüşümdür (Maaten vd.,2007). Verilerin işlenebilmesi için gerekli olan ön işlem adımlarındandır.

Spektral bantların sayısı eğitim örneklerinin sayısından çok fazla olursa, Hughes etkisi denilen (Pal ve Foody, 2010) çok boyutluluk problemi ile karşılaşmaktadır. Elde edilen indirgenmiş verinin, orijinal verinin içindeki çoğu bilgiyi içermesi (Jia vd., 2013), saklamayı kolaylaştırması sınıflandırma, spektral katışım, hedef tanıma gibi bir çok uygulamanın performansını artırması ve uzaktan algılanmış verilerin görüntülenmesinde kolaylık sağlaması (Jia vd., 2013; Kuybeda ve Barzohar,2007; Ren vd, 2014) boyut indirgeme işleminin hiperspektral görüntülerin işlenmesinde ne kadar yararlı olduğunu göstermektedir

Boyut indirgeme yöntemleri literatürde farklı şekillerde gruplandırılmaktadır. En genel hali Şekil 1'de gösterildiği gibi öznitelik seçme (ÖS) ve öznitelik çıkartma (ÖÇ) yöntemleridir. ÖS, sınıflandırma, hedef tanımlama gibi belli bir uygulamanın performansını daha çok artırmak için, belirlenen bir seçme kriterine göre, orijinal öznitelik kümesinden belli sayıda öznitelik seçilmesidir. ÖÇ'nin amacı ise yüksek boyutlu veri kümesini çoğu bilgisini koruyabilecek ve bu veriyi orijinal boyuttan daha küçük bir veri kümesine dönüştürecek bir dönüşüm fonksiyonu bulmaktır. Bu dönüşüm doğrusal olabilir ya da doğrusal olmayabilir. ÖS yöntemlerinin aksine, ÖÇ yüksek boyutlu orijinal veriyi daha küçük sayıda yeni öznitelikler bulabilmek için sıkıştırılmaktadır (Ghamisi P. vd.,2017).



Şekil 1: Hiperspektral görüntülerde boyut azaltımı a) Öznitelik Seçme yöntemi b) Öznitelik Çıkartma yöntemi

Boyut indirgeme yöntemleri en son yapılan genel tarama çalışmasında (Ghamisi P. vd.,2017) eğitilmiş, eğitimsiz ve yarı eğitilmiş olmak üzere üçe ayrılmıştır.

Etiketli örneklerin olmadığı durumlarda ilgilene eğitimsiz boyut indirgeme yöntemleri, verilen bazı kriterleri sağlayarak daha az sayıda bant kümesini bulmayı amaçlarlar.

Eğitimsiz yöntemlerden olan doğrusal temel bileşen analizi (PCA) (Jliffe I.,1986), verideki en yüksek varyansı bularak, bağımsız bileşen analizi (ICA) (Hyvrinen,2001) istatistiksel bağımsızlığı maksimum yapıp projeksiyon matrisini bularak ve en az gürültü giderme dönüşümü (MNF) (Green vd., 1988) sinyal-gürültü oranı (SNR) kriterini kullanarak indirgenmiş öznitelikleri elde ederler.

Eğitilmiş boyut indirgeme yöntemleri sınıfların ayrılabilirliklerini arttırmak için, etiketli örnekleri kullanarak boyut indirgemesi gerçekleştirmektedir. Çok kullanılan birkaç eğitilmiş boyut indirgeme yöntemi Lineer Diskriminant analizi (LDA) (Chang ve Ren,2000), ve parametrik olmayan ağırlıklı öznitelik çıkartma (NWFE) (Kuo ve Landgrebe 2004), Jeffries-Matsushita mesafesini (Swain ve Davis, 1978) ve karşılıklı bilgiyi kullanan (Battiti,1994) bant seçme yöntemidir. Bu yöntemlerin uzantısı olan diğer yöntemler, değiştirilmiş Fishers LDA ve düzenlenmiş LDA, konumsal ve spektral bilgiyi kullanan değiştirilmiş NWFE ve Kernel NWFE yöntemleridir (Ghamisi P. vd.,2017) . Son zamanlarda çıkan eğitilmiş boyut indirgeme yöntemleri lokal komşuluk özelliklerini kullanmaktadır. Li vd. (Li, 2012) ilgili çok modelli yapıyı koruyarak boyutu indirgemek için Fishers Discriminant analiz yöntemini kullanmaktadır.

Gerçek uygulamalarda etiketli veriler çok az bulunmaktadır. Yeterli olmadığı durumlarda da üretilmesi uzun süre almakta ve çok maliyetli olmaktadır. Diğer yandan çok az maliyet ile bolca etiketsiz veri bulunmaktadır. Bu yüzden hem etiketsiz, hem de az bulunan etiketli verileri kullanarak, sınıflandırma başarımını iyileştirmeyi amaçlayan yarı eğitilmiş metotlar

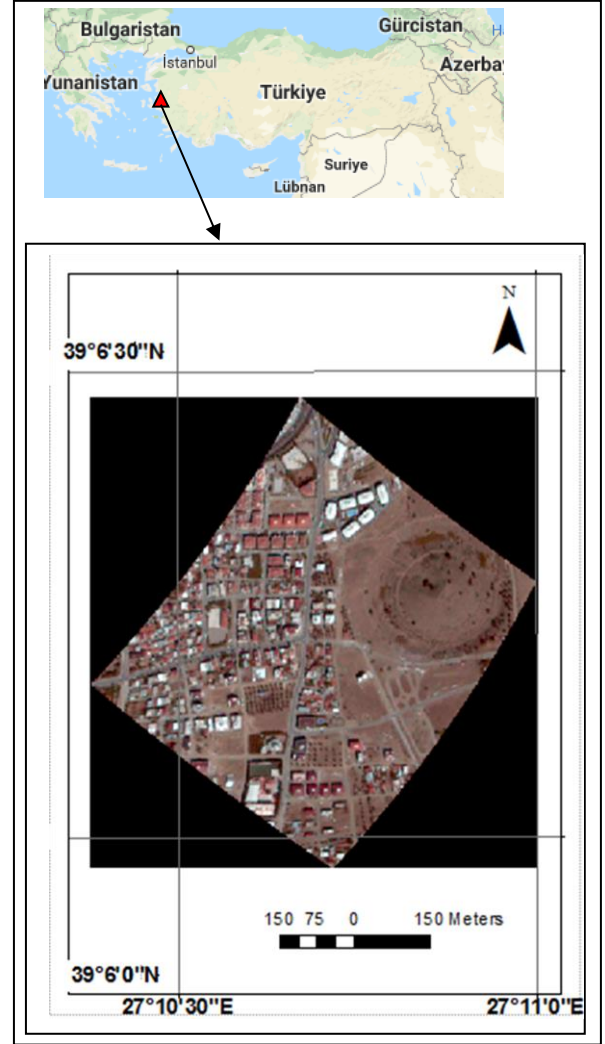
geliştirilmeye başlanmıştır (Blum ve Mitchell, 1998; Bruzzone, Chi vd., 2006). Bu yöntemlerden bazıları grafik tabanlı yarı eğitilmiş öğrenme yöntemleridir. Bazıları da verinin kesin potansiyel özelliklerini korumak için düzenleyici bir terim ekleyerek öznelik çıkartan yöntemlerdir. Bu yöntemler sınırlı sayıda etiketli örnekleri sınıf ayrımlarını maksimum yapmak için ve hem etiketli hem etiketsiz örnekleri verinin lokal özelliklerini saklamak için kullanılırlar. Yöntemler hakkında daha ayrıntılı bilgi Ghamisi vd. yazdığı tarama yayınında yer almaktadır (Ghamisi P. vd.,2017).

Bu çalışmada Harita Genel Müdürlüğü tarafından, araştırmacıların kullanımına sunulan, Headwall Hyperspec VNIR kamerası ile Bergama ili üzerinde elde edilen hiperspektral hava fotoğrafları kullanılmıştır. Uçaktan çekilen Bergama'ya ait hava fotoğraflarına uygulanan boyut indirgeme yöntemlerinin sınıflandırma başarılarına olan etkileri araştırılmıştır.

1.1. Çalışma Alanı ve Veri Setleri

Çalışma alanı olarak Bergama Bölgesi seçilmiştir (Şekil 2). Bu bölgede 2100 metre ortalama yükseklikten Headwall Hyperspec VNIR kamerası ile görüntü alımı gerçekleştirilmiştir (Web2, 2019).

Hyperspec kamerası ile 90 cm çözünürlükte, 380-1000 nm dalgaboyu aralığında ve 400 spektral bant içeren hiperspektral görüntüler çekilmiştir. Bu çalışmada 1004 piksel yükseklik, 5500 piksel genişlik bilgisayarda kullanım açısından çok büyük olacağı için 512x512 piksel boyutunda bir alt görüntü kesilmiş ve işlemler onun üzerinden yapılmıştır (Şekil 2).



Şekil 2: Çalışma alanı

2. YÖNTEM

Sınıflandırma, boyut indirgeme yöntemlerinin performanslarını ölçmek için en çok tercih edilen yöntemdir. Dolayısı ile bu çalışmada da 5 farklı yöntem ile hiperspektral görüntülerin boyutları indirgindikten sonra, 2 farklı sınıflandırma yöntemi ile boyut indirgeme yöntemlerinin performansları ölçülmüştür.

2.1. Boyut İndirgeme Yöntemleri

Bu çalışmada kullanılan boyut indirgeme yöntemleri Temel Bileşen Analizi (PCA), Ayrık Dalgacık Dönüşümü (DWT), Faktör Analizi, Çok Boyutlu Ölçekleme yöntemi (MDS – Multidimensional Scaling) ve En Fazla Daralan Metrik Öğrenme (MCML-Maximally

Collapsing Metric Learning) yöntemidir. Bu yöntemlerden PCA, Faktör Analizi, MDS ve MCML Maaten (Maaten, 2007) tarafından Matlab'ta geliştirilen boyut indirgeme araç kutusunda gerçekleştirilmiştir (Web1,2019). DWT yönteminde ise Matlab'ta DWT2 fonksiyonu kullanılmıştır.

Temel Bileşenler Analizi (PCA), veriyi daha az bir boyuttaki doğrusal alt uzaya yerleştirdiği için doğrusal bir boyut indirgeme yöntemidir (Hotelling,1933). PCA verileri varyanslarının en yüksek oldukları değerlere göre birbirleri ile kolerasyonsuz bir uzaya izdüşüm yapan boyut indirgeme yöntemidir (Saralioglu ve Gormus, 2018). Birinci temel bileşen değişimin en fazla olduğu bileşen, ikinci temel bileşen değişimin ikinci en fazla olduğu bileşen olmak üzere istenen boyut sayısı kadar sıralanmaktadır.

Çok Boyutlu Ölçekleme (MDS) analizi, n nesne ya da birim arasındaki p değişkene göre belirlenen uzaklıklara dayalı olarak nesnelerin k boyutlu $k < p$ bir uzayda gösterimini elde etmeyi amaçlayan, böylece nesneler arasındaki ilişkileri belirlemeye yarayan bir yöntemdir. Verilerin sınıflandırılması ve gruplandırılması amacıyla geliştirilmiş bir analiz olan çok boyutlu ölçekleme analizi, bu açıdan kümeleme analizi ve diskriminant analizi ile benzerlik taşımaktadır. Çok boyutlu ölçekleme analizi çok boyutlu uzayda verilerin ilişki yapısını grafiksel olarak ortaya koyarken birimlerin benzerlik veya farklılık değerlerini dikkate almaktadır (Bulbul ve Kose, 2010). Temel bileşenlere benzer bir yaklaşım kullanılmaktadır. n tane nesne arasındaki uzaklıklar matrisi S verildiğinde bu nesne ($n-1$) boyutlu uzayda n nokta ile gösterilmektedir. Buradaki S_{ij} ' ler tüm i, j ve k için $S_{ij} < S_{ik} + S_{kj}$ koşulunu sağlamaktadır. Bu koşulu sağlayan her n nesnenin geometrik modeli kurulabilmektedir (Filiz ve Çemrek,2005).

Faktör Analizi yöntemi, başta sosyal bilimler olmak üzere pek çok alanda sıklıkla kullanılmaktadır. Faktör analizinde amaç, p

değişkenli bir olayda, birbiriyle ilişkili değişkenlerden faktör adı verilen az sayıda yeni ilişkisiz değişken bulmaktır. Bu analiz boyut indirgemeye ve bağımlılık yapısını gidermeye çalışır (Filiz ve Çemrek,2005). $X_{p \times n}$ ham veri matrisi standartlaştırılarak $Z_{p \times n}$ standartlaştırılmış veri matrisi kullanılır. Z_j değişkenleri ile ortak faktörler arasındaki ilişkiyi gösteren doğrusal model aşağıdaki gibi ifade edilmektedir:

$$z_j = a_{j1}f_1 + a_{j2}f_2 + \dots + a_{jm}f_m + b_j u_j \quad (1)$$

$$j = 1, 2, \dots, p$$

Burada; a_{jm} , j . değişkenin m 'inci faktör üzerindeki yükü yada ağırlığı, f_m : m . faktörü, u_j : artık faktörü, b_j : artık faktöre ilişkin katsayıyı göstermektedir.

(1) denklemini matris formunda aşağıdaki gibi yazmak da mümkündür:

$$Z = AF + BU \quad (2)$$

Faktör Analizi uygulandıktan sonra faktör dönüşümüyle değişkenler arasındaki korelasyonları en iyi açıklayan en az sayıda faktörler tanımlanarak, boyut indirgenmesi gerçekleştirilmiş olur.

En Fazla Daralan Metrik Öğrenme (MCML), Mahalanobis uzaklığını kullanarak sınıflandırma yapan, aynı sınıftaki verilerin metrik olarak birbirine en yakın olacağını kabul eden ve buna göre verileri daraltan bir öğrenme yöntemidir (Saralioglu ve Görmüş, 2018). MCML algoritması benzerlik ve farklılık sınırlamaları tabanlı Mahalanobis mesafesi metriğini öğrenmeyi amaçlamaktadır. Burada aynı sınıftaki tüm örneklerin tek bir noktaya yığılmasını sağlayan ve diğer sınıflardaki örneklerin ise mümkün olduğunca uzağa iten bir optimizasyon problemi kurulmuştur. Benzer etiketlenen ve farklı etiketlenen örnekler arasında lokal komşuluk kısıtlamalarını kullanmaktadır (Globerson, A., Roweis, 2005). DWT metodu ile boyut indirgeme yöntemi, istenilen seviyede ve istenilen dalgacık filtresi ile iki boyutlu dalgacık dönüşümü yapmaktadır.

Girdi verisini yaklaşım ve detay katsayıları şekline matrislere ayrıştırır. Daha sonra yaklaşım katsayılar matrisinden ilk d tane alınan bantlar d boyutlu indirgenmiş yeni veriyi oluşturmaktadır. Bu uygulamada dalgacık dönüşümü olarak Daubechies 3 (db3) ve ayrışım seviyesi olarak 4 kullanılmıştır.

2.2. Makine Öğrenme Yöntemleri

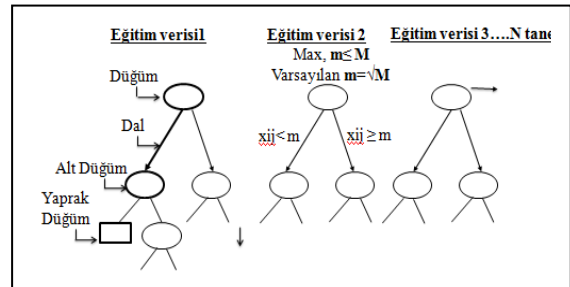
Son yıllarda sınıflandırma doğruluğunu ciddi derecede artırdığı için uzaktan algılama uygulamalarında Destek Vektör Makineleri (DVM) ve Rastgele Orman (RO) gibi makine öğrenme yöntemleri sıklıkla tercih edilmektedir (Thanh Noi ve Kappas, 2018).

Toplu öğrenme sınıflandırıcıları arasında yer alan ve doğruluğu eşsiz olarak tanımlanan RO, hızlı ve belirli bir kalıbı olmayan bir yöntemdir (Breiman, 2001; Breiman ve Cutler 2005; Akar ve Güngör, 2015). İstenilen sayıda ağaç üreterek sınıflandırma işlemini gerçekleştirir (Breiman ve Cutler 2005). RO yönteminin kullandığı ağaç tipi sınıflandırıcılar $\{h(x, \theta_k) k = 1, \dots\}$ şeklindedir. Burada, x , girdi verisini; θ_k , rastgele vektörü temsil etmektedir (Breiman, 2001; Breiman, 2009). Ağaç gelişimi için T gerçek eğitim veri setinden yer değiştirmeli olarak ve rastgele özellik seçimi kullanılarak T_k yeni bir eğitim veri seti oluşturulur. Bu yeni veri setine göre ağaçlar budanmadan geliştirilir ve $h(x, T_k)$ sınıflandırıcısı oluşturulur. Sınıflandırıcı ile çantaya atılmış tahminlerden oylama yapılır. Eğitim verisindeki her x , y için sadece bu sınıflandırıcı ile oylama gerçekleşir. T_k , x ve y 'yi içermez (Archer, 2008; Breiman, 2001). Ağaç geliştirme sırasında kullanıcı tarafından değişkenlerin sayısı (m) ve geliştirilecek ağaçların sayısı (N) olmak üzere 2 parametre tanımlanmalıdır (Pal, 2005). m değeri ağacın nasıl oluşturulacağını belirleyen bir parametredir. Kullanıcı tarafından başlangıç m değeri rastgele seçilir sonraki m 'ler geliştirilmiş hatalara (OBB) göre tespit edilir. m değişken değeri seçilirken, m değerinin M (toplam değişken sayısı) değişkeninin kareköküne eşit olarak alınması

genellikle optimum sonuçlar verir (Breiman, 2002). Kısaca m değeri, sınıflandırmada önemli olan değişken sayısını ifade etmektedir. (Chen, 2008). Bunun yanında RO en büyük boyutta ağaç geliştirmek için kullandığı CART (Classification and Regression Tree) algoritmasında, düğümlerde homojenliği ölçen Gini indeksini kullanır (Özkan, 2008). T eğitim veri seti için C_i sınıfına ait rastgele bir piksel seçildiğinde Gini indeksi Eşitlik 3 ile ifade edilir.

$$\sum_{j \neq i} (f(C_i, T) / |T|) (f(C_j, T) / |T|) \quad (3)$$

Eşitliğinde (3), $f(C_i, T) / |T|$ seçilen örneğin C_i sınıfına ait olma olasılığını gösterir (Pal, 2005). Gini indeksi sıfıra ulaşınca yani her bir yaprak düğümde bir sınıf kaldığında ağaç dallanma işlemi sonlanır (Watts, vd., 2011). Kullanıcı tarafından tanımlanan sayıda ağaç üretilir (Liaw ve Wiener, 2002). RO sınıflandırıcısının ağaç yapısını Şekil 3' teki gibidir.

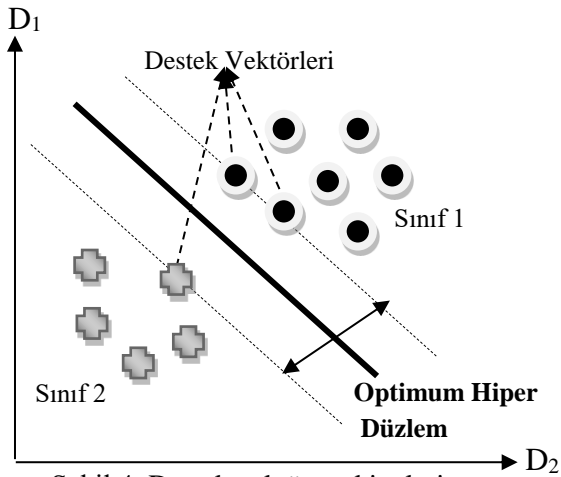


Şekil 3. RO sınıflandırıcısına ait ağaç yapısı

Şekildeki x_{ij} , girdi verilerini temsil etmektedir. İstenilen sayıda ağaçlar üretildikten sonra girdi verisi ormandaki her bir ağaca yerleştirilir ve ağaçlar arasında oylama yapılır ve o piksel ağaçlar arasında en çok hangi sınıfa atanmışsa o sınıfın etiketiyle etiketlenir (Liaw ve Wiener, 2002).

DVM sınıflandırıcısı, hem doğrusal olarak birbirinden ayrılabilen hem de doğrusal olarak birbirinden ayrılamayan verileri sınıflandırabilir. Amaç sınıfları birbirinden ayıran optimum hiper düzlem belirlenmesidir (Vapnik, 1995). Eğer sınıflar birbirinden doğrusal olarak ayrılabiliriyorsa, sınıfları birbirinden ayıran düzlemlerden mesafesi en

büyük olan düzlemleri belirler ve bu düzlemleri kullanarak doğrusal bir ayırt edici fonksiyon oluşturur. Lineer fonksiyonlarla sınıflar birbirinden ayrılır. Bu sınıflar doğrusal olarak ayırlamıyor ise sınıflandırma hatasını minimum, düzlemler arası mesafeyi maksimum yapmayı sağlayacak pozitif bir C parametresi ve kernel fonksiyonları kullanılarak sınıfların doğrusal olarak ayrılabilirdikleri daha yüksek boyutlu başka bir uzaya taşınırlar. Bu uzayda sınıflandırma işlemi gerçekleşir (Özkan, 2008; (Tso ve Mather, 2009; Stephens ve Diesing; 2014, Çölkesen ve Yomralıoğlu, 2014). İyi performans gösterdiği için en yaygın kullanılan kernel fonksiyonu Radyal tabanlı fonksiyondur. (Thanh Noi ve Kappas, 2018; Kavzoglu ve Çölkesen 2009). Şekil 4'te DVM yönteminin doğrusal problemlerde çalışma prensibi gösterilmiştir.

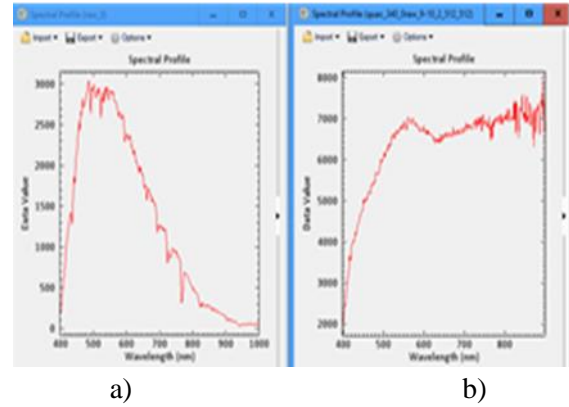


Şekil 4. Destek vektör makineleri

3. UYGULAMA

HGK'dan alınan veriler öncelikle ENVI programında manuel olarak incelenerek genellikle sonda ve başta bulunan çok bozuk ve çok gürültülü bantlar çıkartılmıştır. Kötü bantların çıkartılmasından sonra 400 banttan 340 bant kalmıştır. Daha sonra spektral özellikleri incelendiğinde veri formatının radyans olduğu görülmüştür. Radyans formatında görüntü hem alıcıdan hem de atmosferdeki gazlardan etkilenerek gerçek yeryüzündeki yansımaya değerlerinden farklı

spektral özelliğe sahip olmaktadır. Bu Şekil 5-a'da yansıma grafiğinde görülmektedir. Su buharı etkisinden dolayı belli dalga boylarında yansımaların çok düşük olduğu görülmektedir. Dolayısı ile diğer görüntü işleme adımlarına geçmeden önce radyans değerlerinin yeryüzü yansımaya değerlerine yani reflektans değerlerine dönüştürülmesi gerekmektedir. Bunun için de ENVI 5.2 programında Atmosfer Düzeltmesi (Quick Atmospheric Correction (QUAC)) yapılmıştır. Elde edilen reflektans görüntüsüne ait yeşil alan olmayan bir noktanın spektral yansıma grafiği Şekil 5-b'de gösterilmiştir.



Şekil 5 : a) İşlenmemiş görüntünün spektral yansımasının radyans olduğu, b) Görüntünün QUAC atmosferik düzeltmeden sonra spektral yansımasının reflektans formatına dönüşmesi

Atmosferik düzeltmeden sonra reflektans görüntüsü PCA, DWT, MDS, Faktör analizi ve MCML yöntemleri ile 20 bant indirgenmiş 5 ayrı görüntü elde edilmiştir. Seçilen 5 farklı yonteme göre hiperspektral hava fotoğrafının boyut indirgeme işlemi tamamlandıktan sonra elde edilen her bir görüntüye geometrik düzeltme işlemi uygulanarak görüntü düzeltilmiştir. Google Earth üzerinden seçilen 8 kontrol noktası yardımıyla ERDAS Imagine yazılımı kullanılarak polynomial yontemine göre piksel altı doğrulukla (x yönünde 0.65 m, y yönünde 0.18 m) toplamda 0.67 m hatayla geometrik düzeltme işlemi gerçekleştirilmiştir. Bu şekilde hiperspektral hava görüntülerinin ön işlemleri tamamlanmıştır.

Ön işlemlerden sonra boyut indirgeme yöntemlerinin performanslarının analizi için görüntüler sınıflandırma işlemine tabi tutulmuştur. Sınıflandırma işlemi için görüntüler üzerinden toprak, yol, bitki, bina (beton çatılı), bina (kiremit çatılı) ve gölge olmak üzere 6 adet arazi kullanım sınıfı belirlenmiştir. ENVI yazılımında her sınıf için görüntüler üzerinden sınıflara ait örnek pikseller toplanmıştır. Her bir sınıf için yaklaşık eşit sayıda piksel toplanmıştır.

Gölge sınıfı görüntüde az yer kapladığı için diğer sınıflara göre daha az sayıda piksel toplanmıştır. Toplamda görüntü üzerinden 9332 eğitim pikseli toplanmıştır. Sonrasında ENVI yazılımında DVM sınıflandırıcısı yardımıyla optimum sınıflandırma parametresine (C) ve Radyal tabanlı fonksiyona göre tüm indirgenmiş görüntüler sınıflandırılmıştır. (C) parametresi kullanıcı tarafından görüntülere göre deneyerek belirlenmiştir. Sonrasında Matlab yazılımında her bir görüntü üzerinde seçilen bu örnek alanlara göre eğitim ve test verileri Rastgele Özellik Seçim Metodu kullanılarak oluşturulmuştur. Görüntüler üzerinden toplanan 9332 pikselin 4666'ü eğitim verisi, kalan 4666'ü test verisi olarak seçilmiştir ve RO sınıflandırıcısı ile görüntüler, kullanıcı tarafından belirlenen optimum parametrelere (m) ve (N) göre sınıflandırılmıştır (Örn. Şekil 3).

DVM ve RO için aynı eğitim alanları kullanılmıştır. Gölge sınıfı sadece ormanlık alanlarda olduğu için sınıflandırma sonucu elde edilen 8 adet tematik görüntüde orman sınıfı gölge sınıfıyla birleştirilmiştir.

Her bir sınıflandırma sonucunun doğruluğu, en yaygın kullanılan sınıflandırma sonrası doğruluk analizleri yöntemlerinden biri olan hata matrisi kullanılarak değerlendirilmiştir. Bu matris kullanılarak, bilinen referans veriler (yer gerçeği) ile buna karşılık gelen otomatik sınıflandırma sonucu elde edilen sonuçlar arasındaki ilişkiler kategorik olarak birbirleriyle karşılaştırılabilmektedir. (Lillesand, vd. 2004). Çalışmada 6 sınıf için, her bir görüntü üzerinde, sınıfların kapladıkları alanlara göre nokta

üreten stratified random yöntemi kullanılarak her sınıf için en az 20 nokta olacak şekilde toplamda 120 nokta üretilmiştir. Bu noktalar kullanılarak sınıflandırılmış tematik görüntülerin doğrulukları analiz edilmiştir. Her bir sınıflandırma sonucu için elde edilen genel, üretici ve kullanıcı doğrulukları Tablo 1' de verilmiştir.

4. BULGULAR ve TARTIŞMA

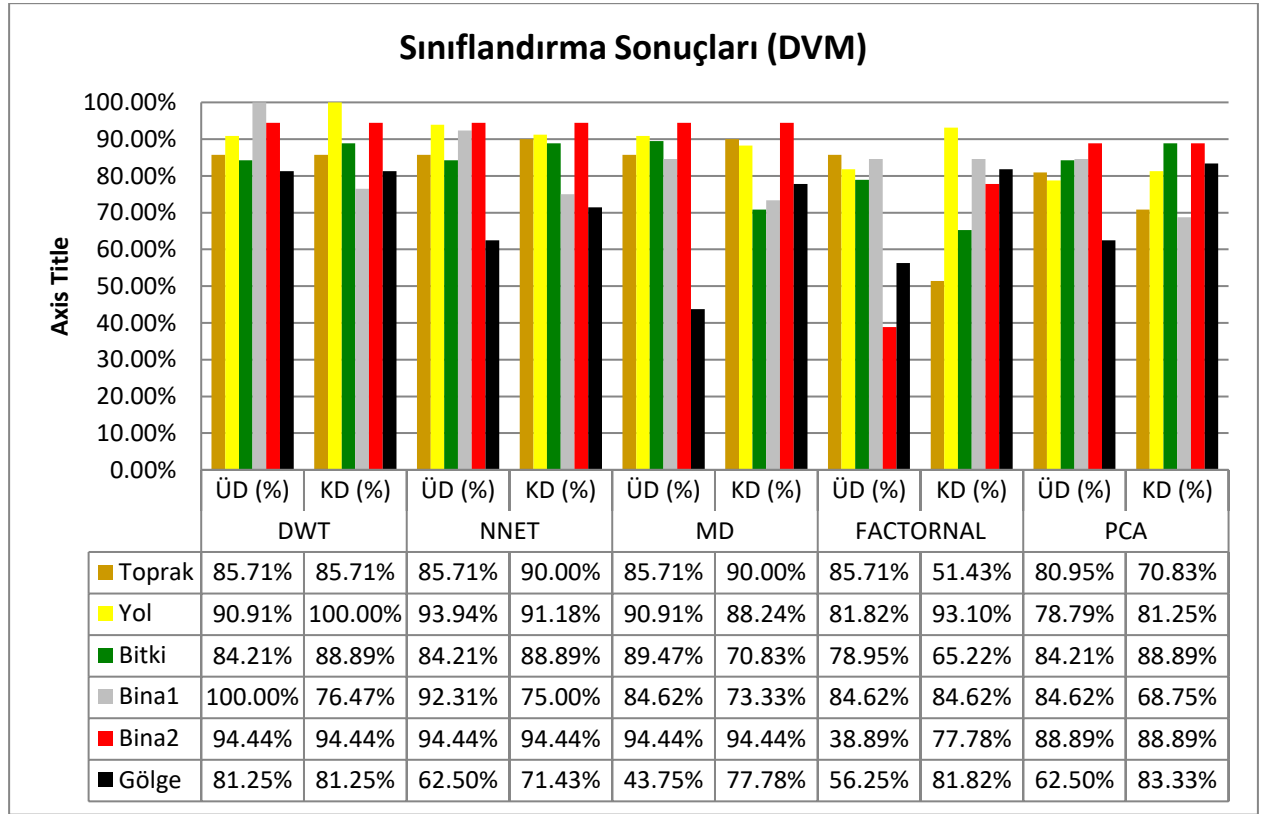
Çalışmada kullanılan boyut indirgeme yöntemlerinin DVM ve RO ile sınıflandırılmaları neticesinde performanslarını değerlendirmek için, doğruluk analizleri sonucu elde edilen hata matrisleri incelenmiştir.

Buna göre en yüksek sınıflandırma doğruluğunu RO ile sınıflandırma sonucu, %91.67 genel doğruluk ve 0.91 κ değeri ile DWT boyut indirgeme yönteminin sağladığı gözlemlenmiştir. Genel olarak bakıldığında RO ile elde edilen sonuçlarda genel sınıflandırma doğrulukları yüksekte düşüğe sırasıyla DWT, MCML, MDS, FACTOR ve PCA olarak elde edilmiştir. DVM sınıflandırma sonuçlarına göre bu sıranın ise DWT, MCML, MDS, PCA ve FACTOR olduğu gözlemlenmiştir. Buna göre her iki sınıflandırma yönteminde de en yüksek sınıflandırma doğruluğunu DWT yönteminin, en düşük sınıflandırma doğruluğunu da FACTOR ve PCA yöntemlerinin verdiği görülmüştür. Tablo 2'de elde edilen sınıflandırma doğrulukları Genel Doğruluk (%) ve Kappa katsayısı cinsinden gösterilmiştir.

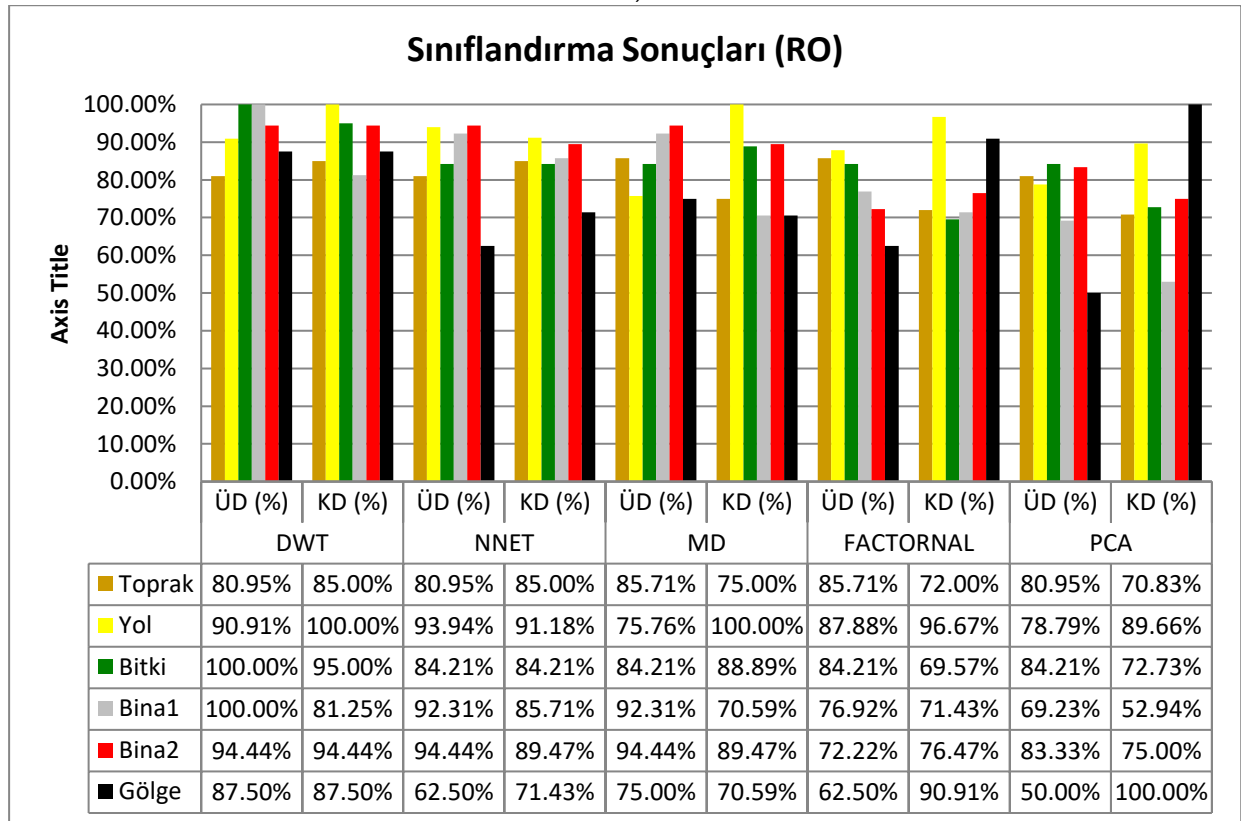
Hiperspektral hava fotoğrafının RO ve DVM'e göre elde edilen üretici ve kullanıcı doğrulukları incelendiğinde çoğu sınıflarda en yüksek başarımın DWT yöntemi ile 20 banda indirgenen görüntünün sınıflandırılması sonucu elde edildiği görülmüştür.

Her iki sınıflandırmada da (RO ve DVM) Bina1 sınıfı %100, Bina2 %94.44 ÜD ile DWT yöntemi ile elde edilmiştir.

Tablo 1. Üretici Doğrulukları (ÜD) ve Kullanıcı Doğrulukları (KD) a) DVM b) RO



a)



b)

Tablo 2: DWT ve RO sınıflandırıcıları için Genel Doğruluk tablosu

Sınıflandırıcı	İndirgeme Yöntemi	Genel Doğruluk (%)	Kappa
DVM	DWT	89.17	0.8892
	PCA	80.00	0.7954
	MDS	83.33	0.8295
	FACTOR	72.50	0.7219
	MCML	86.67	0.8636
RO	DWT	91.67	0.9147
	PCA	75.83	0.7521
	MDS	83.33	0.8293
	FACTOR	80.00	0.7957
	MCML	85.83	0.8549

DVM'ye göre Toprak sınıfının en yüksek başarımı %87.71 ÜD, RO'ya göre Bitki ve Gölge sınıflarının en yüksek başarımları sırasıyla %100 ve %87.50 ÜD ile yine DWT yöntemi ile indirgenmiş görüntünün sınıflandırılması sonucu elde edilmiştir. DWT'nin yanısıra MCML ve MDS yöntemleri ile de DWT'ye yakın veya eşit başarımların elde edildiği Tablo1'de görülmektedir. Sadece Yol sınıfının en yüksek başarımı her iki sınıflandırıcıya göre %93.94 ÜD ile MCML yöntemine göre indirgenmiş görüntünün sınıflandırılması sonucu elde edilmiştir. Diğer yandan Yol sınıfı KD'e göre en yüksek başarımı (%100) DWT yöntemi ile elde etmiştir. KD'ye göre başarımlara bakıldığında, Yol, bitki ve Bina2 sınıflarının DWT ile en yüksek başarımları verdiği görülmektedir. Buna göre, Bina2 sınıfının her iki sınıflandırmada da %94.44, Bitki sınıfının da RO'ya göre %95 olduğu görülmektedir. Binal ve Gölge sınıflarında en yüksek KD RO'ya göre sırasıyla MCML (%85.71) ve PCA (%100) indirgeme yöntemleri, Toprak sınıfında ise DVM'de MCML (%90) yöntemi ile elde edilmiştir. DVM ve RO yöntemleri ile sınıflandırılan her bir indirgenmiş görüntünün tematik görüntüleri sırasıyla Şekil 6'da ve Şekil 7'de verilmiştir.

Çalışma sonucunda, bu hiperspektral hava fotoğrafı için denenen yöntemler arasında en yüksek sınıflandırma başarımı, RO sınıflandırıcısı ve DWT boyut indirgeme yöntemi kullanılması ile gerçekleştirilmiştir.

5. SONUÇLAR

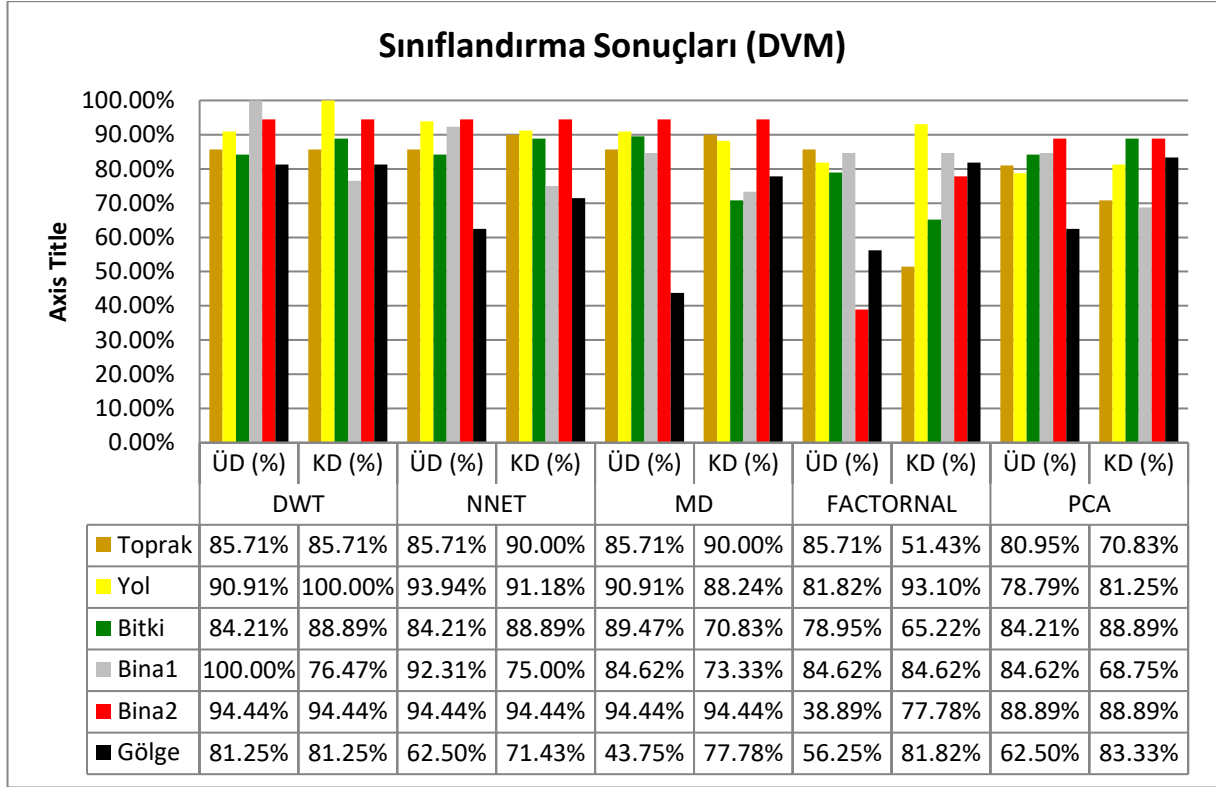
Bu çalışmada Harita Genel Müdürlüğü tarafından, BHİKPK-Bilimsel Araştırma ve Koordinasyon Komisyonunca (BARKOK) 2017 yılında Bergama bölgesinde Headwall Hyperspec VNIR kamerası ile elde edilen hiperspektral hava görüntüleri kullanılmıştır. Bu görüntülere yaygın olarak kullanılan farklı boyut indirgeme yöntemleri uygulanmıştır. Çalışma kapsamında boyut indirgeme yöntemlerinin performanslarının kullanılan Headwall Hyperspec VNIR görüntüleri için analizi gerçekleştirilmiştir. Uygulamada 5 farklı indirgeme yöntemi (DWT, MD, FACTORNAL, PCA ve MCML) kullanılmış ve bu yöntemlerin performansı 2 farklı sınıflandırıcı (RO ve DVM) ile test edilmiştir. Elde edilen sonuçlara bakıldığında DWT yönteminin hem RO hem de DVM tarafından en yüksek başarımı sağlayan indirgeme yöntemi olarak seçildiği görülmüştür. Bunun nedeni olarak DWT metodunun boyut indirgeme sırasında hem spektral hem de konumsal bilgiyi kullanıyor olması söylenebilir.

Bu çalışma ile Bergama bölgesine ait olan veri setinin ileriki görüntü işleme aşamaları için ön işlem aşamaları hakkında araştırmacılara yol gösterilmiş ve herkese açık olan bu veri setinin daha çok kullanılması teşvik edilmiştir.

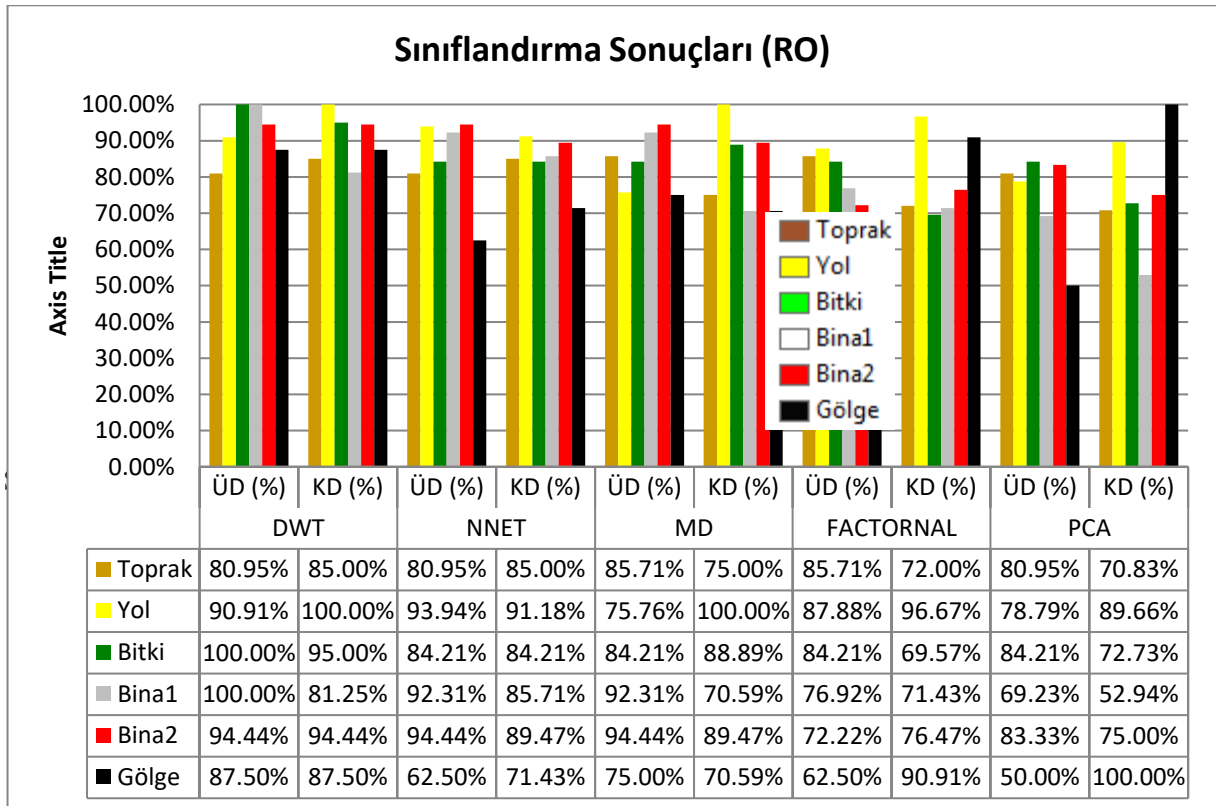
Teşekkür

Çalışmada kullanılan Headwall Hyperspec VNIR Kamerası ile elde edilen Hiperspektral Hava Fotoğrafı Harita Genel Komutanlığı'ndan temin edilmiştir.

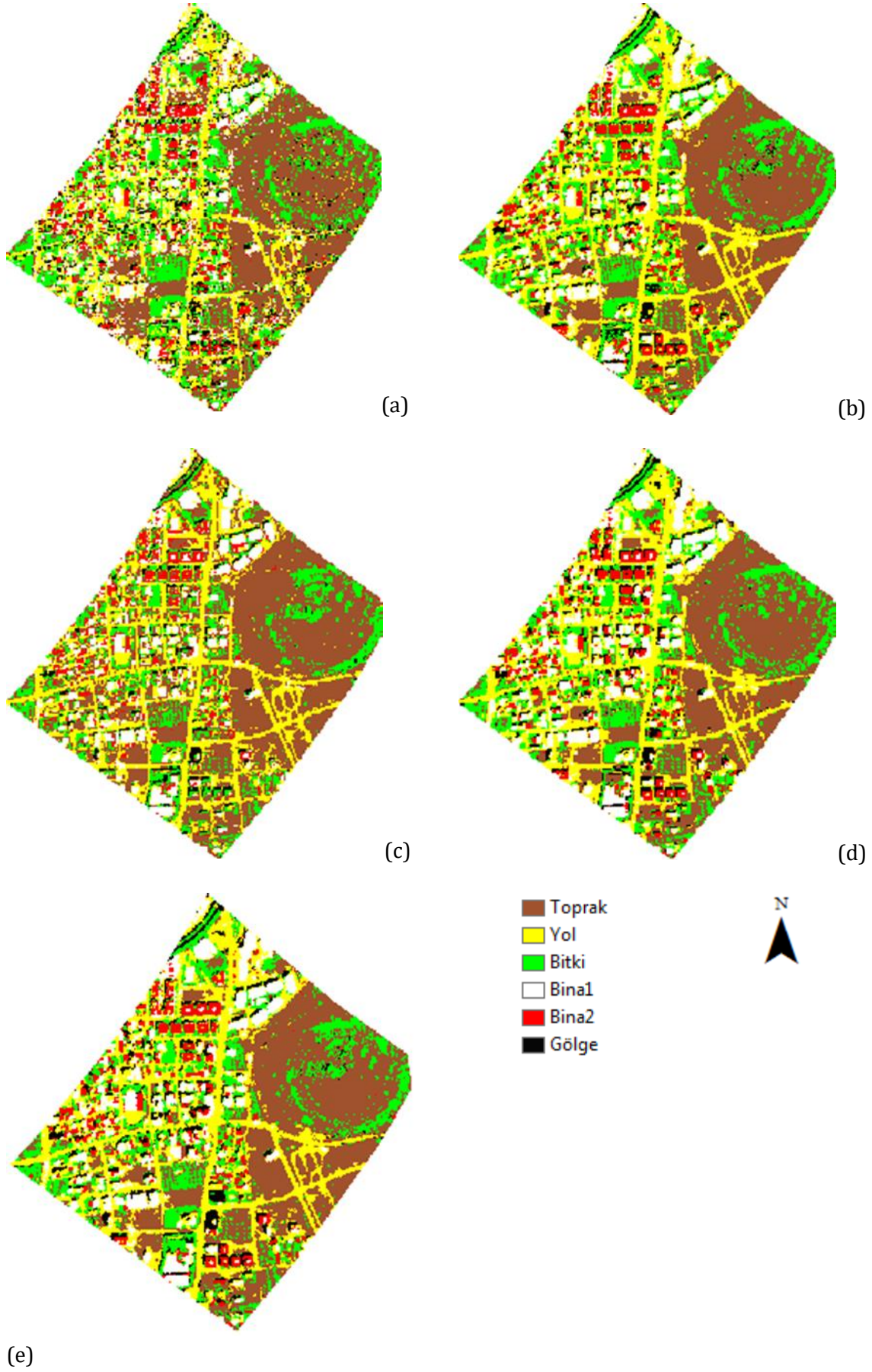
Tablo 1. a) DVM ve b) RO sonucu elde edilen Üretici Doğrulukları (ÜD) ve Kullanıcı Doğrulukları (KD)



a)



b)



Şekil 6: DVM ile elde edilen tematik görüntüler a)PCA b)MDS c)Faktör Analiz d)DWT e)MCML



Şekil 7: RO ile elde edilen tematik görüntüler a)PCA b)MDS c)Faktör Analiz d)DWT e)MCML

KAYNAKÇA

- Abbak, A. (2007). Jeodezide Zaman Dizilerinin Dalgacık (Wavelet) Analizi. Doktora Semineri. Selçuk üniversitesi, Jeodezi ve Fotogrametri Mühendisliği Anabilim Dalı Konya.
- Akar, Ö., ve Güngör, O. (2015). Integrating multiple texture methods and NDVI to the random forest classification algorithm to detect tea and hazelnut plantation areas in northeast Turkey, *International Journal of Remote Sensing*, 36, 442–464.
- Archer, K. J. (2008). Empirical characterization of random forest variable importance measure, *EDMputational Statistics & Data Analysis*, 52(4), 2249-2260.
- Başak, H., Yıldırım, M.A. (2017). Göktürk-2 Uydu Görüntüsünün Otomatik Detay Çıkarımında Kullanılabilirliğinin Araştırılması "Eskişehir İli Örneği" TMMOB Harita ve Kadastro Mühendisleri Odası, 16. Türkiye Harita Bilimsel ve Teknik Kurultayı, 3-6 Mayıs 2017, Ankara.
- Battiti, R. (1994). "Using mutual information for selecting features in supervised neural net learning," *IEEE Trans. Neural Networks*, vol. 5, no. 4, pp. 537–550, Jul.
- Belward A.S. ve Skoien, J.O. (2015) "Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 103, no. May 2015, 115-128, 2014.
- Blum A., ve Mitchell T., (1998) "Combining labeled and unlabeled data with co-training," *Proceedings of the 11th Annual Conference on Computational Learning Theory*, pp. 92–100, 1998.
- Breiman, L. (2001). Random Forests, Machine learning, *Kluwer Academic Publishers*, 45(1), 5-32.
- Breiman, L. (2002). Manual On Setting Up, Using, And Understanding Random Forests V3.1, http://oz.berkeley.edu/users/breiman/Using_random_forests_V3.1.pdf (06.11.2009)
- Breiman, L. ve Cutler, A.. (2005). Random Forest, http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm (15.11.2009).
- Bruzzone L., Chi M., ve Marconcini M., (2006) "A novel transductive svm for semisupervised classification of remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3363 – 3373.
- Bulbul S. ve Kose A., (2010). "Türkiye’de bölgelerarası iç göç hareketlerinin çok boyutlu ölçekleme yöntemi ile incelenmesi", *İstanbul Üniversitesi İşletme Fakültesi Dergisi*. Cilt/Vol:39, Sayı/No:1, 75-94.
- Chang C.-I. ve Ren C.-I. (2000). "An experiment-based quantitative and comparative analysis of target detection and image classification algorithms for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 2, pp. 1044–1063.
- Chen, C. H. (2008). *Image Processing For Remote Sensing*, CRS Press, Taylor&Francis Group, USA.
- Çölkesen, İ. ve Yomralıoğlu, T. (2014). Arazi Örtüsü ve Kullanımının Haritalanmasında WorldView-2 Uydu Görüntüsü ve Yardımcı Verilerin Kullanımı, *Harita Dergisi*, Temmuz 2014 (152): 12-24.
- Filiz Z., Çemrek F., (2005). Avrupa Birliğine Üye Ülkeler ile Türkiye'nin Karşılaştırılması, İstanbul Üniversitesi VII.Ulusal Ekonometri ve İstatistik Sempozyumu.
- Ghamisi P. vd. (2017), "Advances in Hyperspectral Image and Signal Processing: A Comprehensive Overview of the State of the Art," in *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 37-78.
- Globerson, A., Roweis, S., (2005). 'Metric learning by collapsing classes'. In: *Advances in Neural Information Processing Systems (NIPS)*.
- Green A., Berman M., Switzer P., and Craig M., (1988). "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. and Remote Sens.*, vol. 26, no. 1, pp. 65–74, Jan.
- Hotelling H., (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24:417–441.
- Hyvriinen A., Karhunen J., and Oja E., (2001). *Independent Component Analysis*. New York: Wiley.
- Immitzer, M., Vuolo, F. ve Atzberger, C. (2016) First Experience With Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sens.* 2016, 8, 166.
- Inglada, J., Arias, M., Tardy, B., Hagolle, O., Valero, S., Morin, D., Dedieu, G., Sepulcre, G., Bontemps, S., Defourny, P., (2015). Assessment of an Operational System for Crop Type Map Production Using High Temporal and Spatial Resolution Satellite Optical Imagery. *Remote Sens.* 2015, 7, 12356–12379.
- Jia X., Kuo B, and Crawford M., (2013). "Feature mining for hyperspectral image classification," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 676–697, March.
- Jliffe I., (1986). *Principal Component Analysis*. New York:Springer-Verlag.
- Kalkan, K., Orhun, Ö., Filiz, B. ve Teke, M. (2015). Vegetation Discrimination Analysis from Göktürk-2, in *Recent Advances in Space Technologies (RAST)*, 2015 7th International Conference on, 2015.
- Kavzoglu, T. ve Colkesen, I. (2009), A kernel functions analysis for support vector machines for

- land cover classification, *International Journal of Applied Earth Observation and Geoinformation*, 11(5): 352- 359.
- Kuo B. ve Landgrebe D.,(2004) “Nonparametric weighted feature extraction for classification,” IEEE Trans. Geosci. and Remote Sens., vol. 42, no. 5, pp. 1096–1105, May.
- Kuybeda D. M. O and Barzohar M., (2007). “Rank estimation and redundancy reduction of high dimensional noisy signals with preservation of rare vectors,” IEEE Trans. on Signal Processing, vol. 55, no. 12, pp. 5579–5592.
- Lahat, D., Adali, T. ve Jutten, C., (2015). Multimodal Data Fusion: An Overview of Methods, Challenges, and Prospects, in Proceedings of the IEEE, vol. 103, no. 9, pp. 1449-1477, Sept..
- Li W, Prasad S., Fowled J. E., ve Bruce L. M., (2012).“Locality-preserving dimensionality reduction and classification for hyperspectral image analysis,” IEEE Trans. Geosci. Remote Sens., vol. 50, no. 4, pp. 1185–1198.
- Liaw, A. ve Wiener, M. (2002). Classification And Regression By Random Forest, R News, Vol.2/3, December.
- Lillesand, T. M., Kiefer R. W. ve Chipman J. W., 2004. Remote Sensing and Image Interpretation, Wiley, United States of America, 804.
- Liu N. W. A. L., Coops N. C. and Pang Y., (2017). “Mapping urban tree species using integrated airborne hyperspectral and lidar remote sensing data,” Remote Sensing of Environment, vol. 200, pp. 170–182.
- Maaten L.J.P. van der (2007). An introduction to dimensionality reduction using Matlab. Technical Report 07-06, MICC-IKAT, Maastricht University, Maastricht, The Netherlands.
- Maaten, L.J.P. van der, Postma E.O., and Herik H.J. van den (2007). Dimensionality reduction: A comparative review. Preprint.
- Mather, P. M. (2004). EDMputer Processing of Remotely-Sensed Images: An Introduction, Third edition, Wiley, USA, ISBN 0-470-84918-5.
- Özkan, Y. (2008). Veri Madenciliği Yöntemleri, Papatya Yayıncılık, İstanbul.
- Pal, M. (2005). Random Forest Classifier For Remote Sensing Classification, *International Journal of Remote Sensing*, 26(1) , 217-222.
- Pal M. ve Foody G. M. (2010), "Feature Selection for Classification of Hyperspectral Data by SVM," in IEEE Transactions on Geoscience and Remote Sensing, vol. 48, no. 5, pp. 2297-2307, May.
- Ren J., Zabalza J., Marshall S., and Zheng J.,(2014). “Effective feature extraction and data reduction in remote sensing using hyperspectral imaging,” IEEE Signal Processing Magazine, vol. 31, no. 4, pp. 149–154, June.
- Richards, J.A. ve Jia, X. (2006). Remote Sensing Digital Image Analysis: An Introduction 4th Edition, Springer-Verlag, Germany, ISBN-10: 3-540-25128-6.
- Saralioğlu E., Tunç Görmüş E.,(2018). "Doku Bilgisinin Hiperspektral Görüntülerin Sınıflandırılmasına Olan Etkisinin Araştırılması", 26. IEEE Sinyal İşleme Ve İletişim Uygulamaları Kurultayı, İzmir, Türkiye, 2-5 Mayıs, pp.1-4
- Somers G. P. B., (2013). “Multi-temporal hyperspectral mixture analysis and feature selection for invasive species mapping in rainforests,” Remote Sensing of Environment, vol. 136, pp. 14–27.
- Srivastava, P.K., Han, D., Rico-Ramirez, M.A., Bray, M. ve Islam, T. (2012) Selection of Classification Techniques for Land Use/land Cover Change Investigation. Adv. Space Res., 50, 1250–1265.
- Stephens, D. ve Diesing, M. (2014). A Comparison of Supervised Classification Methods for the Prediction of Substrate Type Using Multibeam Acoustic and Legacy Grain-Size Data. *Magar V, ed. PLoS ONE.*;9(4):e93950.
- Swain, P. H. ve Davis S. M., (1978) Remote Sensing: The Quantitative Approach. New York: McGraw-Hill.
- Tardy, B., Inglada, J. ve Michel, J. (2017). Fusion Approaches for Land Cover Map Production Using High Resolution Image Time Series without Reference Data of the Corresponding Period, Remote Sensing, 9, 1151.
- Teke, M. (2016). Satellite Image Processing Workflow for Rasat and Göktürk-2, Journal of Aeronautics and Space Technologies, Volume 9 Number 1.
- Thanh Noi, P. ve Kappas, M. (2018). Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery, *Sensors (Basel, Switzerland)*. 2018;18(1):18.
- Vapnik, V.N. (1995). The Nature of Statistical Learning Theory, Springer-Verlag, New York.
- Watts, J. D., Powell, S.L., Lawrence, R. L. ve Hilker, T. (2011). Improved Classification of Conservation Tillage Adoption Using High Temporal And Synthetic Satellite Imagery, *Remote Sensing of Environment* 115 (2011) 66–75
- Web1 (2019), <http://lvdmaaten.github.io/drtoolbox/>
- Web2 (2019), <https://www.harita.gov.tr/images/kurumsal/ec160682cb3586d.pdf>