

# Göktürk-2 ve Hyperion EO-1 Uydu Görüntülerinden Rastgele Orman Sınıflandırıcısı ve Destek Vektör Makineleri ile Arazi Kullanım Haritalarının Üretilmesi

Özlem AKAR<sup>\*1</sup>, Esra Tunç GÖRMÜŞ<sup>2</sup>

<sup>1</sup>Erzincan 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

<sup>2</sup>Karadeniz Teknik Üniversitesi, Mühendislik Fakültesi, Harita Mühendisliği Bölümü, Trabzon (etuncgormus@ceng.ktu.edu.tr) ORCID ID 0000-0002-7667-199X

## Öz

Çalışmada spektral özellikleri birbirine yakın arazi sınıflarını birbirinden ayırarak, Göktürk-2 uydu görüntülerinden daha doğru bir arazi kullanım haritasının üretilmesi amaçlanmıştır. Bunun için Hyperion EO-1 hiperspektral uydu görüntüsünün, yüksek spektral çözünürlüğünden yararlanılmıştır. Çalışma alanı olarak spektral özellikleri birbirine yakın arazi sınıflarına sahip olan Trabzon Akçaabat ilçesinin Büyükbaba yaylası seçilmiştir. Çalışmada Göktürk-2 Multispektral (GMS), Göktürk-2 Pankromatik (GPAN) ve Hyperion EO-1 hiperspektral uydu görüntüleri kullanılmıştır. Öncelikle Hyperion EO-1 hiperspektral uydu görüntüsü için atmosferik ve radyometrik düzeltmeler yapılmış, bozuk ve kullanılmayan bantların temizlenmesi için bant indirgeme işlemleri uygulanmıştır. Bant indirgeme işlemi için dalgacık tabanlı Ampirik Kip Ayırıştırma (AKA) yöntemi kullanılmıştır. Sonrasında tüm görüntüler rektifiye edilerek aynı koordinat sisteminde olması sağlanmıştır. Görüntüler ön işlemden geçirildikten sonra GPAN, GMS ve indirgenmiş Hyperion EO-1 (DHYP) görüntüleri ile Gram Schmidt (GS) ve Principle Component (PC) gibi görüntü kaynaştırma yöntemleri kullanılarak kaynaştırılmıştır. Kaynaştırma yöntemleriyle elde edilen kaynaştırılmış görüntüler üzerinden sınıflandırmada kullanılacak arazi kullanım sınıfları belirlenmiştir. Bu görüntüler yüksek sınıflandırma doğruluğu veren Rastgele Orman (RO) ve Destek Vektör Makineleri (DVM) yöntemleriyle sınıflandırılmıştır. Her bir sınıflandırma sonucu için doğruluk analizleri yapılmış ve elde edilen doğruluklar karşılaştırılmıştır. Çalışmanın sonuçları incelendiğinde, en yüksek sınıflandırma doğruluğunun, PC kaynaştırma yöntemine göre kaynaştırılmış ve RO sınıflandırıcısı ile sınıflandırılmış görüntülerden elde edildiği gözlenmiştir. PC ile kaynaştırılmış GPAN ve GMS görüntüsü ile GPAN ve DHYP' nin kaynaştırılmış görüntüsünün RO ile sınıflandırılması sonucu genel sınıflandırma doğrulukları sırasıyla %72.13 ve %83.06 olarak elde edilmiştir. Bu sonuçlara göre sınıflandırma doğruluğu % 11 oranında artırılmıştır. Son olarak en yüksek sınıflandırma doğruluğuna sahip olan tematik görüntü kullanılarak arazi kullanım haritası üretilmiştir.

**Anahtar Kelimeler:** Göktürk-2, Hyperion EO-1, Rastgele Orman, Ampirik Kip Ayırıştırma, Arazi Kullanım Haritası

\* Sorumlu Yazar

# **Producing Land Use Maps With Random Forest Classifier and Suport Vector Machines From Göktürk-2 and Hyperion EO-1 Satellite Images**

## **Abstract**

In this study it is aimed to produce a more accurate land use map from Gokturk-2 by separating the land classes with very similar spectral properties. Hyperion EO-1 hyperspectral satellite image with high spectral resolution was exploited in order to do that.. Büyükoba highland in Akcaabat which has land classes with very similar spectral properties was chosen as working field. Göktürk-2 Multispectral (MS), Göktürk-2 Pancromatik(PAN) and Hyperion EO-1 hyperspectral satellite images were used in this study. First of all atmospheric and radiometric corrections of hyperspectral images were done together with dimensionality reduction of unused and irrelevant bands. Wavelet based empirical mode decomposition method was used for dimensionality reduction. Then all the images were synchronized in the same coordinate system by rectification. After preprocessing steps Gokturk-2 MS and reduced Hyperion EO-1 (DHYP) were fused with Gokturk-2 Pan image by methods like Gram Schmidt (GS) and Principle Component (PC). Land classes were determined on the fused images. Then these images were classified by Random Forest (RF) and Support Vector Machine (SVM). Accuracy analysis were done for each classification results and then compared with each other. In the end it was seen that, the best classification accuracy was obtained by using RF classification method and PC fusion method. RF overall classification accuracies obtained from both fusing GPAN and GMS image with PC, and fusing GPAN and DHYP image with PC were 72.13 % and 83.06%, respectively. According to these results, the overall classification accuracy was increased by 11%. Finally, a land use map was produced using the thematic image with the highest classification accuracy.

**Keywords:** *Göktürk-2, Hyperion EO-1, Random Forest, Empirical Mode Decomposition, Land Use Map.*

## **1. GİRİŞ**

Arazi kullanım haritaları, birçok bilimsel ve çevresel uzaktan algılama uygulamaları için önemli bilgiler temin etmektedir. Teknolojinin ilerlemesi ve uyduların çoğalmasıyla birlikte uydu görüntüleri bu haritaların hızlı ve güvenilir bir şekilde yapılmasına yardımcı olmaktadır (Belward,2014). Bu anlamda arazi kullanım haritalarının hızlı ve güvenilir olarak üretilmesinde görüntü sınıflandırma işlemi en yaygın kullanılan uzaktan algılama tekniklerinden biridir. Son yıllarda sağladığı yüksek doğruluk nedeniyle sıklıkla tercih edilen Rastgele Orman (RO) ve Destek vektör makineleri (DVM) gibi ileri sınıflandırma yöntemleri ile optik uydu görüntülerinden, daha doğru ve büyük ölçekli arazi kullanım haritaları oluşturulabilmektedir (Tardy vd.,2017, Srivastava vd., 2012).

Görüntü sınıflandırma mevcut spektral verileri kullanarak belirli bir spektral sınıfın parçası

olan pikselleri etiketleme işlemidir (Richards ve Jia, 2006). Sınıflandırma işlemi iki adımdan oluşur. Sınıflandırmada birinci adım, arazi yüzeyi olarak düşünüldüğünde orman, su, çayır-çimen alanları, toprak gibi farklı arazi yüzeylerini içeren sınıfları tanımlamaktır. İkinci aşama ise seçilen sınıflara göre piksellerin etiketlenmesi işlemidir. Bu işlemlere sınıflandırma ve etiketleme denilmektedir (Mather, 2004).

Sınıflandırmada en yaygın olarak karşılaşılan sorunlardan biri de optik görüntülerin spektral çözünürlüklerinin multispektral görüntülerdeki benzer spektral özelliklere ait sınıfları birbirinden ayırmada yetersiz olmasıdır. Bu tarz problemlerin cevabı başka kaynaklardan gelen uydu görüntülerinin özelliklerini tamamlayıcı bilgi olarak kullanarak, multispektral görüntülerin sınıflandırılmasının iyileştirilmesidir. Buna en iyi örnek spektral

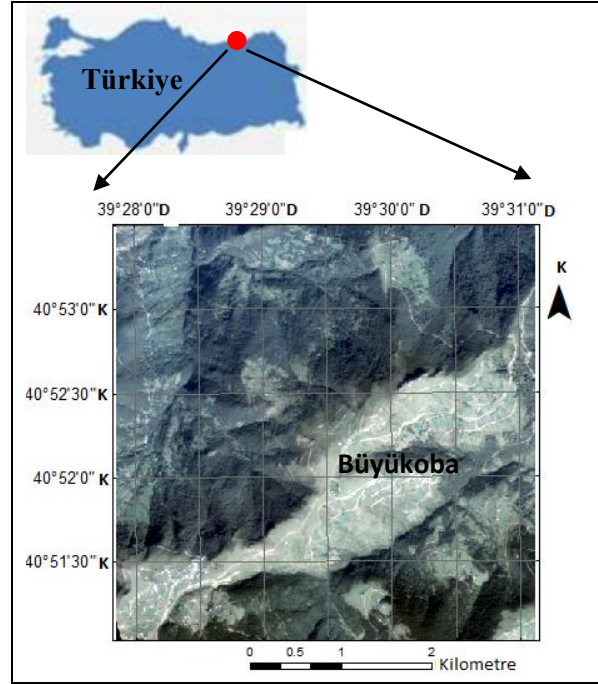
çözünürlüğü yüksek olan hiperspektral uydu görüntüleridir. Farklı uydu görüntülerinin kaynaştırılması ile tek bir görüntüden elde edilemeyecek bilgiler elde edilir, bilgi çıkartımı yapılır ve karar verme mekanizmaları geliştirilir (Lahat vd., 2015).

Literatüre bakıldığında arazi kullanım haritalarının üretimi için farklı uydu görüntüleri kullanılarak bir çok çalışmanın yapıldığı görülmektedir. Örneğin, optik uydu görüntülerinden Landsat5 ve 8' e ait olan görüntüler sıkça hassas arazi kullanım haritası üretmek için kullanılmaktadır (İnglada vd., 2015). Bunun yanı sıra Sentinel 2'ye ait görüntülerin de başarı ile kullanıldığı görülmektedir (Immitzer vd., 2016).

Çalışmada kullanılan Göktürk-2 uydu görüntüsünü sağlayan uydu, Türkiye'nin ikinci uydusudur ve 18 Aralık 2012 tarihinde fırlatılmıştır (Teke, 2014). Şimdiye kadar detay çıkartımı (Başak ve Yıldırım, 2017) ve bitki ayırımı (Kalkan vd. 2015) gibi birçok uygulamada kullanılmıştır. Bu çalışmada ise Göktürk-2 uydu görüntüsünün arazi kullanım haritalarının üretimindeki başarısı test edilmiştir. Bunun için öncelikle, küçük ve birbirine yakın spektral özellikteki objeleri birbirinden daha iyi ayırt edebilmek için hem konumsal hem de spektral çözünürlüğü iyi görüntü elde etmek amacıyla, konumsal çözünürlüğü 2.5 m olan Göktürk-2 PAN uydu görüntüsü, spektral çözünürlüğü yüksek, indirgenmiş Hyperion EO-1(DHYP) hiperspektral uydu görüntüsü ile kaynaştırılmıştır. Daha sonra yüksek sınıflandırma doğruluğu sağlayan RO sınıflandırma yaklaşımı ile spektral özellikleri birbirine yakın arazi sınıflarını birbirinden ayırarak arazi kullanım haritası üretilmiştir.

### 1.1. Çalışma Alanı ve Veri setleri

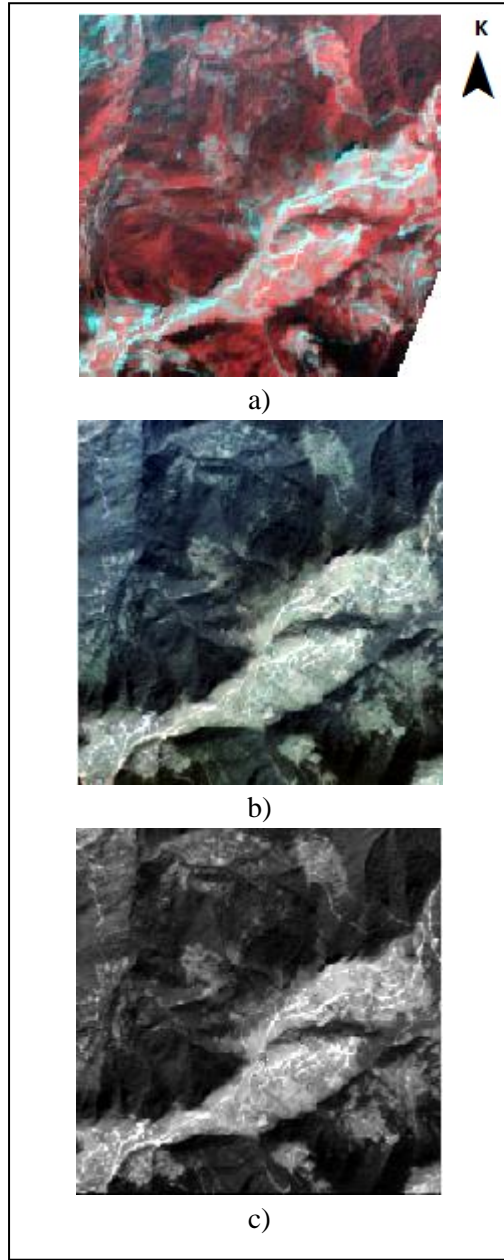
Çalışma alanı olarak seçilen Büyükoba yaylası, Türkiye' nin Doğu Karadeniz Bölgesinde yer alan Trabzon ili Akçaabat ilçesinde yer almaktadır. Çalışma alanı 2091 hektarlık bir alanı kapsamaktadır (Şekil 1). Çalışmada Göktürk-2 Multispektral (MS), Göktürk-2



Şekil 1. Çalışma Alanı

Pankromatik (PAN) ve Hyperion EO-1 hiperspektral uydu görüntüleri kullanılmıştır (Şekil 2). Çalışmada kullanılan Göktürk-2 uydu görüntüsü, 5 m konumsal çözünürlüklü 4 adet Multispektral banda (Kırmızı, Yeşil, Mavi ve Kızılötesi) ve 2.5 m konumsal çözünürlüklü 1 adet PAN banda sahiptir. Kızılötesi bant çalışma alanını kapsamadığı için kullanılmamıştır. Hyperion EO-1 hiperspektral uydu görüntüsü ise 220 adet spektral banda ve 30 konumsal çözünürlüğe sahiptir.

Çalışma alanı, mera ve orman, toprak yollar ve toprak, binalar ve beton yollar olmak üzere birbirine spektral olarak benzer özellikteki arazi kullanım sınıflarını içermektedir. Göktürk-2 uydu görüntüsü gibi bant sayısı düşük bir görüntüyle bu sınıfları birbirinden ayırmak zor olmaktadır. Bant sayısı fazla ve konumsal çözünürlüğü iyi bir görüntü kullanılması halinde yakın spektral özellikli sınıfları birbirinden ayırt edebilmek daha mümkün olmaktadır. Bu nedenle yakın spektral özellikteki sınıfları birbirinden daha iyi ayırarak sınıflandırma doğruluğunu artırmak için Hyperion EO-1 hiperspektral uydu görüntüsünün spektral bantlarından ve Göktürk-2 uydu görüntüsünün konumsal çözünürlüğünden yararlanılmıştır.



Şekil 2. Kullanılan uydu görüntüleri a) Hyperion EO-1, b) Göktürk-2 MS, c) Göktürk-2 PAN

## 2. YÖNTEM

Bu bölümde çalışmada boyut azaltmak ve sınıflandırma yapmak için kullanılan yöntemler kısaca açıklanmıştır.

### 2.1.Dalgacık Dönüşümü (DD)

Dalgacık dönüşümü görüntüler içerisindeki değişimleri analiz eden etkili bir yöntemdir. Geleneksel yöntemler ile karşılaştırıldığında sinyalin orijinal yapısını bozmadan kolayca

işlem yaptığı görülmektedir. Dalgacık analizi bir sinyale ölçek ve zaman perspektifinden bakmayı sağlamaktadır (Abbak, 2007). Ölçek ile sinyalin yerel düzenliliği ve zaman ile değişimin oluşum anı hakkında bilgi verdiği için fourier ve gabor gibi diğer dönüşümlere göre daha avantajlı olduğu görülmektedir. Dalgacık dönüşümü sürekli ve ayrık olmak üzere iki ana başlığa ayrılmaktadır. Dalgacık fonksiyonunun kaydırılıp bir ölçekle çarpıldıktan sonra zaman alanı boyunca toplanmasıyla sürekli dalgacık dönüşümü gerçekleştirilir. Tüm ölçek aralığında yapılacak dalgacık analizi sonucu çok büyük veri yığınları oluşmaması için analist belirli ölçek grupları tespit eder ve bu aralıkta analizleri yapmaktadır. Çoğunlukla en pratik ve kullanışlı yol, ölçek ve konum değerleri ikinin kuvveti olacak şekilde seçilmesidir. Bu işleme ayrık dalgacık dönüşümü adı verilmektedir (Abbak, 2007).

Bu çalışmada ayrık dalgacık dönüşümü'nü hesaplarken ana dalgacık olarak Daubechies 3 (db3) (Chen,2009) ve indirgeme seviyesi olarak 4 alınmıştır. Dalgacık dönüşümünden sonraki sinyalin uzunluğu dalgacık filtresinin ve orijinal sinyalin uzunluğuna bağlıdır. Dalgacık dönüşümü tabanlı boyut azaltımı hangi bileşenlere ait katsayıları (alçak frekans, yüksek frekans ya da her ikisi) seçecekse, ona göre değişmektedir. Bu çalışmada Dalgacık dönüşümü yapıldıktan sonra boyut azaltım amaçlı katsayılar seçimi doğrusal dalgacık öznelik çıkartımına göre yapılmıştır (Tunç Görmüş vd., 2012) .

### 2.2.Ampirik Kip Ayrıştırma (AKA)Yöntemi

AKA dönüşümü diğer dönüşümlerin aksine bir sınırlama parametresinin ve veri hakkında önceden bir varsayımı olmamasından dolayı görüntü analizinde tercih edilmeye başlanmıştır. AKA ile görüntü İçkin Kip Fonksiyonlarına (İKF) ve bir kalana ayrılır (Huang ve ark. 1998). Daha sonra elde edilen İKF' leri, kalan ile birlikte toplayarak bilgi kaybı ve hata olmadan orijinal işaret elde edilir.

AKA, görüntünün konumsal frekansını İKF' ler ile ifade eder. İlk İKF görüntünün en yüksek lokal konumsal frekanslarını göstermektedir. İkinci İKF, ikinci en yüksek konumsal frekansları göstermektedir ve bu

böyle devam etmektedir. İKF olabilmek için 2 şart vardır. Birincisi, sıfır geçişlerinin sayısının, uç noktasının sayısına eşit olması ya da aralarındaki farkın 1 olmasıdır. Diğeri ise, lokal maksimum ve lokal minimum değerlerinin belirlediği zarfların ortalama değerlerinin herhangi bir noktada sıfır olmasıdır. Üst zarf, işaretin yerel maksimum noktalarının ara değerlemesi ile, alt zarf yerel minimum noktalarının ara değerlemesi ile bulunmaktadır. AKA, sinyalin istatistiklerine bağlı değildir ve sadece sinyalin lokal özelliklerine dayanmaktadır. Frekanslar içi ve frekanslar arası birçok öznelik oluşturmaktadır.

Bir bandın İKF' lerini bulan iteratif işlemlere 'eleme işlemi' adı verilmektedir. Bu işlem bandın kendisi ( $B_1(i,j)$ ) ile başlamaktadır.

$$I_{i,1}^{(1)} = B_1(i,j) \cdot I_{i,m}^{(n)} \quad [1]$$

[1]' de (n) iterasyon sayısını, l=1,2,...L band sayısını, L hiperspektral görüntüdeki toplam görüntü sayısını, (i,j) uzamsal boyutu, m=1:M, M İKF sayısını temsil etmektedir.

Buna göre 2D-AKA'nın nasıl çalıştığı, hiperspektral görüntünün 1 numaralı

bandı  $B_l(i,j)$  için, birinci İKF'nin ( $\dot{I}KF_{l,m}$ ) nasıl bulunduğu aşağıdaki gibi açıklanmıştır;

1) Girdi görüntüde tüm yerel maksimum ve yerel minimum noktalarının 2B genlik ve pozisyonlarının elde edilmesi.

2) Eğri ara değerlendirilmesi ile yerel maksimum noktalarından 2B üst zarf  $E_{max}(i,j)$  ve yerel minimum noktalarından 2B alt zarf  $E_{min}(i,j)$  oluşturulması.

3) Her (i,j) uzamsal pozisyon için üst zarf ve alt zarfın ortalamasının hesaplanması [2]

$$EM_m^n(i,j) = (E_{max}(i,j) + E_{min}(i,j))/2 \quad [2]$$

4) Ortalama zarf işaretinin, giriş işaretinden çıkarılması [3]

$$S_m^n(i,j) = I_{l,m}^n(i,j) - EM_m^n(i,j) \quad [3]$$

Buraya kadarki adımlar eleme işleminin bir iterasyonunu göstermektedir.

5) Bu adımda yapılacak olan, 4. adımda elde edilen görüntünün İKF olup olmadığını test etmektir. İKF olma şartı zarf ortalama işaretinin sıfıra yeterince yakın olup olmamasına bağlıdır[4] (Demir ve Erturk, 2010).

$$\sum_{i=1}^P \sum_{j=1}^R |EM^n(i,j)| / (P * R) < \tau \quad [4]$$

Burada P ve R ortalama zarfın boyutları,  $\tau$  ise sıfıra yakın bir değerdir. Eğer durdurma kriteri iterasyonda sağlandıysa (n=N), ilk İKF elde

edilmiş olur,  $\dot{I}KF_{l,m}(i,j) = S_m^N(i,j)$ . Eğer bu kriter sağlanmadı ise, ilk İKF bulunana kadar, adım 4'te elde edilen görüntü adım 1'de girdi görüntü olarak ele alınıp işlemler tekrar edilir [5],

$$I_{l,m}^{n+1}(i,j) = S_m^n(i,j) \quad [5]$$

6) Eğer İKF başarılı bir şekilde elde edildiyse, kalan sinyal hesaplanır

$R_m(i,j) = I_{l,m}^n(i,j) - \dot{I}KF_{l,m}(i,j)$ . Eğer kalan sinyal daha fazla uç nokta içermiyorsa AKA işlemi sona erer. Eğer içeriyorsa, bir sonraki İKF'nin hesaplanmasına, bu kalan görüntünün birinci adımda girdi görüntü olarak ele alınması ile devam edilir [6],

$$I_{l,m}^{n+1}(i,j) = R_m(i,j) \quad [6]$$

Orijinal bant  $B_l(i,j)$ , tüm İKF'lerin ve en son kalanın toplanması ile tam olarak geri elde edilmektedir [7].

$$B_l(i,j) = \sum_{m=1}^M \dot{I}KF_{l,m}(i,j) + R_M(i,j) \quad [7]$$

Bu çalışmada 2D-AKA, 0.006 durdurma kriteri ve 9x9 pencere kullanılarak yapılmıştır.

Boyut indirmek için ise bu iki yöntem birlikte kullanılmış, sonuçta ana veriyi daha iyi özetleyen yeni bir bant alt kümesi oluşturulmuştur. İndirgeme yöntemi kısaca aşağıda özetlenmiştir (Tunç Gormus vd., 2017).

1. 2B-AKA'nın her bir spektral banda uygulanması ve M derece İKF özniteliklerinin oluşturulması,
2. Elde edilen her bir İKF özniteliklerindeki piksellere spektral uzayda 1B-DD uygulanması, DİKF'lerin elde edilmesi,
3. Düşük seviye DİKF'leri toplayarak yeni dalgacık tabanlı İKF'lerin oluşturulması. Örneğin DİKF1 ve DİKF2 toplayarak DİKF12 elde edilmesi [8],

$$WIKF(1,..M)(i,j) = \sum_{m=1}^M WIKF_m(i,j) \quad [8]$$

4. Son olarak, dalgacık dönüşümünün tipine göre istenildiği kadar öznitelik seçilerek sınıflandırmaya girecek DİKF özniteliklerinin belirlenmesi Tunç Görmüş vd., (2012).

### 2.3. Rastgele Orman (RO) Sınıflandırıcısı

Günümüzde RO algoritması, sınıflandırmada çok iyi performans sergilediği için görüntü sınıflandırmada sıklıkla tercih edilmektedir (Akar ve Güngör, 2015). RO, ağaç tipi sınıflandırıcılar topluluğudur (Breiman, 2001). Breiman ve Cutler (2005) RO'yu, şuan ki algoritmalar arasında doğruluğu eşsiz olan bir sınıflandırıcı olarak tanımlamıştır. Ayrıca hızlı ve belirli bir kalıbı olmayan bir yöntem olduğunu söylemektedir. Bu yöntemde ne kadar istenirse o kadar ağaçla çalışılır (Breiman ve Cutler 2005). RO yönteminde,  $\{h(x, \theta_k) \ k = 1, \dots\}$  şeklinde ağaç tipi sınıflandırıcılar kullanılmaktadır. Burada,  $x$ , girdi verisini;  $\theta_k$ , rastgele vektörü temsil etmektedir (Breiman, 2001; Breiman, 2009).

RO yönteminde, öncelikle gerçek veri setinden yer değiştirmeli olarak yeni bir eğitim veri seti oluşturulur. Ardından, rastgele özellik seçimi kullanılarak yeni eğitim setinden bir ağaç geliştirilir. Geliştirilen ağaçlar budanmaz (Archer, 2008; Breiman, 2001).

RO sınıflandırıcısı ile bir ağaç üretmek için kullanıcı tarafından tanımlanan 2 parametre gereklidir. Bu parametreler, en iyi bölünmeyi belirlemek için her bir düğümde kullanılan değişkenlerin sayısı ( $m$ ) ve geliştirilecek ağaçların sayısı ( $N$ )'dir (Pal, 2005).Kullanıcı tarafından başlangıç  $m$  değeri rastgele seçilir

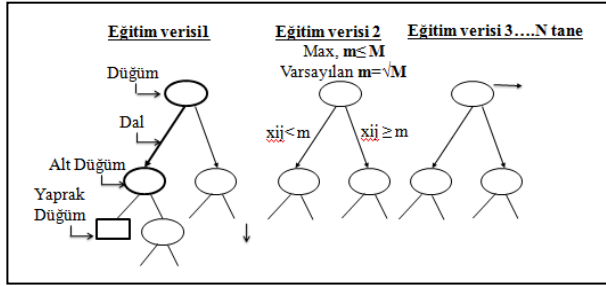
sonraki  $m$ 'ler genelleştirilmiş hatalara (OBB) göre artırılır ya da azaltılır.  $m$  azalınca korelasyon ve güç azalır,  $m$  artınca korelasyon ve güç artar. Bu şekilde en uygun  $m$  bulunur ve sınıflandırma duyarlılığı artar, hata azalır. Breiman, (2002)'a göre,  $m$  değişken değeri seçilirken,  $m$  değerinin  $M$  (toplam değişken sayısı) değişkeninin kareköküne eşit olarak alınması genellikle optimum sonuçlar verir. RO' da tüm girdi değişkenleri kullanılmamaktadır. RO algoritmasının hesap karmaşıklığını sadeleştirmek ve ağaçlar arasındaki korelasyonu azaltmak amacıyla bir  $m$  değeri seçimi yapılmaktadır. Bu  $m$  değeri, sınıflandırmada önemli olan değişken sayısını ifade etmektedir. (Chen, 2008).  $T$  eğitim verisinden  $T_k$  yer değiştirmeli yeni eğitim verisi üretilir. Yeni eğitim veri seti kullanılarak  $h(x, T_k)$  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. (Breiman, 2001).

RO en büyük boyutta ağaç geliştirmek için kullandığı CART (Classification and Regression Tree) algoritmasında, bir düğümde belirli bir ölçüt uygulanarak bölünme işlemi gerçekleştirilir (Özkan, 2008). Bölünme işlemlerinde homojen sınıf dağılımına sahip düğümler tercih edilir. Düğüm homojenliğinin ölçümünde RO yöntemi, Gini indeksini kullanmaktadır. Verilen bir  $T$  eğitim veri seti için rastgele bir örnek (piksel) seçilsin ve bu örnek  $C_i$  sınıfına ait olsun. Bu duruma göre Gini indeksi şöyle ifade edilir (Eşitlik 9);

$$\sum_{j \neq i} f(C_i, T) / |T| \quad [9]$$

Eşitliğinde [1],  $f(C_i, T) / |T|$  seçilen örneğin  $C_i$  sınıfına ait olma olasılığını gösterir (Pal,2005). Gini indeksi büyüdükçe sınıf heterojenliği artarken, Gini İndeksi azaldıkça sınıf homojenliği artar. Bir alt düğümün Gini indeksi bir üst düğümün Gini indeksinden daha az olduğunda o dal başarılıdır. Gini indeksi sifıra ulaşıncaya yani her bir yaprak düğümde bir

sınıf kaldığında ağaç dallanma işlemi sonlanır (Watts, vd., 2011). Kaç tane ağaç üretmek istenirse her düğüm için en iyi dal belirlenerek o kadar ağaç üretilir (Liaw ve Wiener,2002). Şekil3, RO sınıflandırıcısının ağaç yapısını özetlemektedir. Şekildeki  $x_{ij}$ , girdi verilerini temsil etmektedir.



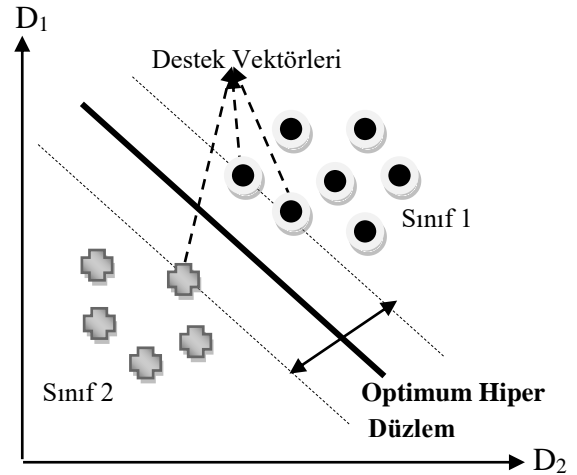
Şekil 3. RO sınıflandırıcısına ait ağaç yapısı

Ağaç gelişiminden sonraki aşama, girdi verisinin sınıflandırılması işlemidir. Bu işlemde RO, girdi verisini ormandaki her bir ağaca yerleştirir. Belirlenen ağaçlar arasında oylama yapılır ve en çok oyu alan ağaç bir sınıfa atanır (Liaw ve Wiener, 2002).

#### 2.4. Destek Vektör Makineleri

Son yıllarda sınıflandırma doğruluğunu ciddi derecede artırdığı için uzaktan algılama uygulamalarında makine öğrenme yöntemi gibi parametrik olmayan yöntemler sıklıkla tercih edilmektedir ki Destek Vektör Makineleri (DVM) bunlardan biridir (Thanh Noi ve Kappas, 2018). Bu yöntem, hem doğrusal olarak birbirinden ayrılabilen (lineer olan) hem de doğrusal olarak birbirinden ayrılamayan (lineer olmayan) verileri sınıflandırabilir. Amaç sınıfları birbirinden ayıran optimum hiper düzlem belirlenmesidir (Vapnik, 1995). Eğer şekil 4'te gösterilen sınıflar birbirinden doğrusal olarak ayrılabilirse, 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 ayrılamıyor ise sınıflandırma hatasını minimum, düzlemler arası mesafeyi maksimum yapmayı sağlayacak pozitif bir  $C$

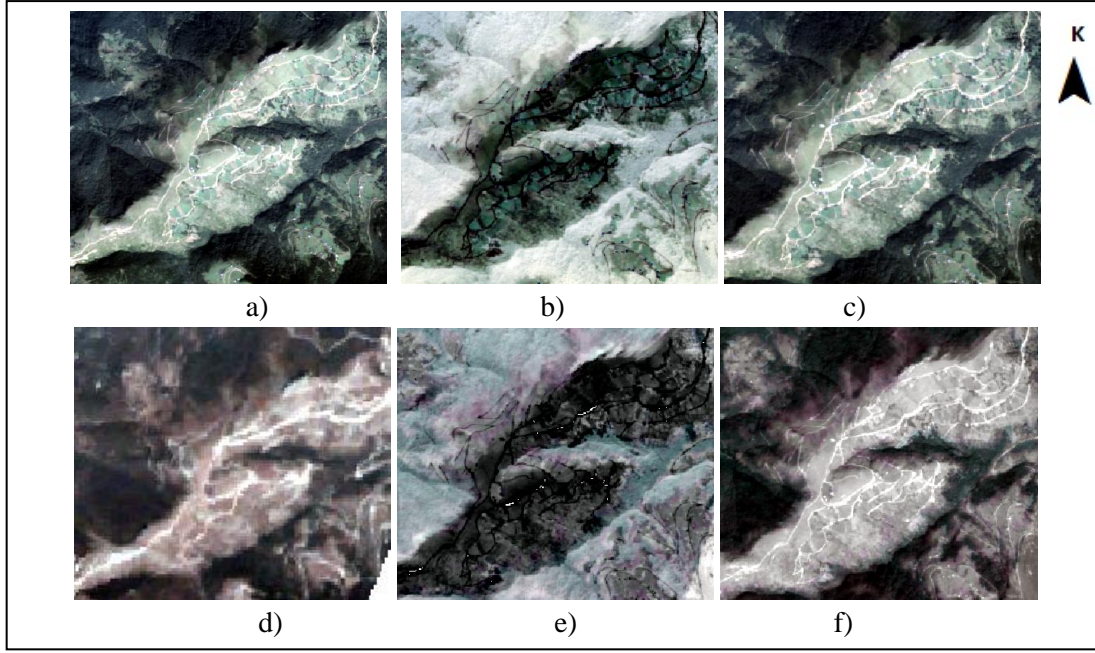
parametresi ve kernel fonksiyonları kullanılarak sınıfların doğrusal olarak ayrılabilir oldukları daha yüksek boyutlu başka bir uzaya taşınırlar. Bu uzayda sınıflandırma işlemi gerçekleşir (Özkan, 2008; (Tso ve Mather, 2009; Stephens ve Diesing; 2014, Çölkesen ve Yomralıoğlu, 2014). İyi performans gösterdiği için en yaygın kullanılan kernel fonksiyonu Radyal tabanlı fonksiyondur. (Thanh Noi ve Kappas, 2018; Kavzoglu ve Çölkesen 2009).



Şekil 4. Destek vektör makineleri

### 3. UYGULAMA

Hyperion EO 1 hiperspektral görüntüsü, kaynaştırma işlemlerinde kullanılmadan önce L1R formatında indirilerek ilgili çalışma alanı kesilmiştir. 242 banttan bozuk bantlar çıkartıldıktan sonra 175 bant kalmıştır. ENVI yazılımı kullanılarak sırasıyla Radyometrik düzeltme ile radyans ve FLAASH düzeltmesi ile reflektans formuna dönüştürülmüştür.



Şekil 5. Orjinal ve Kaynaştırılmış görüntüler a) GMS b) PC\_GMS\_GPAN, c) GS\_GMS\_GPAN, d) DHYP e) PC\_DHYP\_GPAN, f) GS\_DHYP\_GPAN

Bozuk bantların manuel olarak çıkartılmasından sonra, hiperspektral görüntüye boyut indirgeme yöntemi uygulanmıştır. Boyut indirgeme işlemi Hiperspektral görüntülerin işlenmesi sırasında yapılması gereken ön işlem adımlarından biridir. Çok sayıda komşu ve dar bantlara sahip olması görüntünün çok sayıda tekrar eden bilgi içermesine neden olmaktadır. Bu durum hem kullanıcıyı, hem de uygulamanın başarımını olumsuz etkilemektedir. Dolayısı ile bu çalışmada da boyut indirgeme yöntemleri kullanarak uygulamanın sonucunu iyileştirecek en iyi bantlar seçilmiş ve onlar ile işlemlere devam edilmiştir. Bu çalışmada Tunç Görmüş vd. (2012) önerdiği Dalgacık dönüşümü tabanlı Ampirik Kip Ayrıştırma (AKA) yöntemi (DAKA) kullanılarak hiperspektral görüntülerin boyutu azaltılmış ve görüntüyü en iyi özetleyen öznelik alt kümesi bulunmuştur.

İki boyutlu (2B) AKA dönüşümü kullanarak konumsal bilgiden, bir boyutlu (1B) ayrık dalgacık dönüşümü (DWT) kullanarak ise spektral bilgiden yararlanılmaktadır. Diğer yöntemlere göre başarılı olmasının nedeni de hem spektral, hem konumsal bilgiden aynı anda yararlanıyor olmasıdır. Bu indirgeme işlemiyle birlikte Hyperion EO-1 görüntüsü 14 banda indirgenmiştir.

Daha sonraki aşamada, kullanılacak görüntülerin aynı koordinat sisteminde olmaları için ERDAS Imagine yazılımı kullanılarak Göktürk-2 MS ve Hyperion EO-1 görüntüleri Göktürk-2 PAN uydu görüntüsüne göre geometrik olarak düzeltilmiştir. Bu şekilde uydu görüntülerinin ön işlemi tamamlanmıştır.

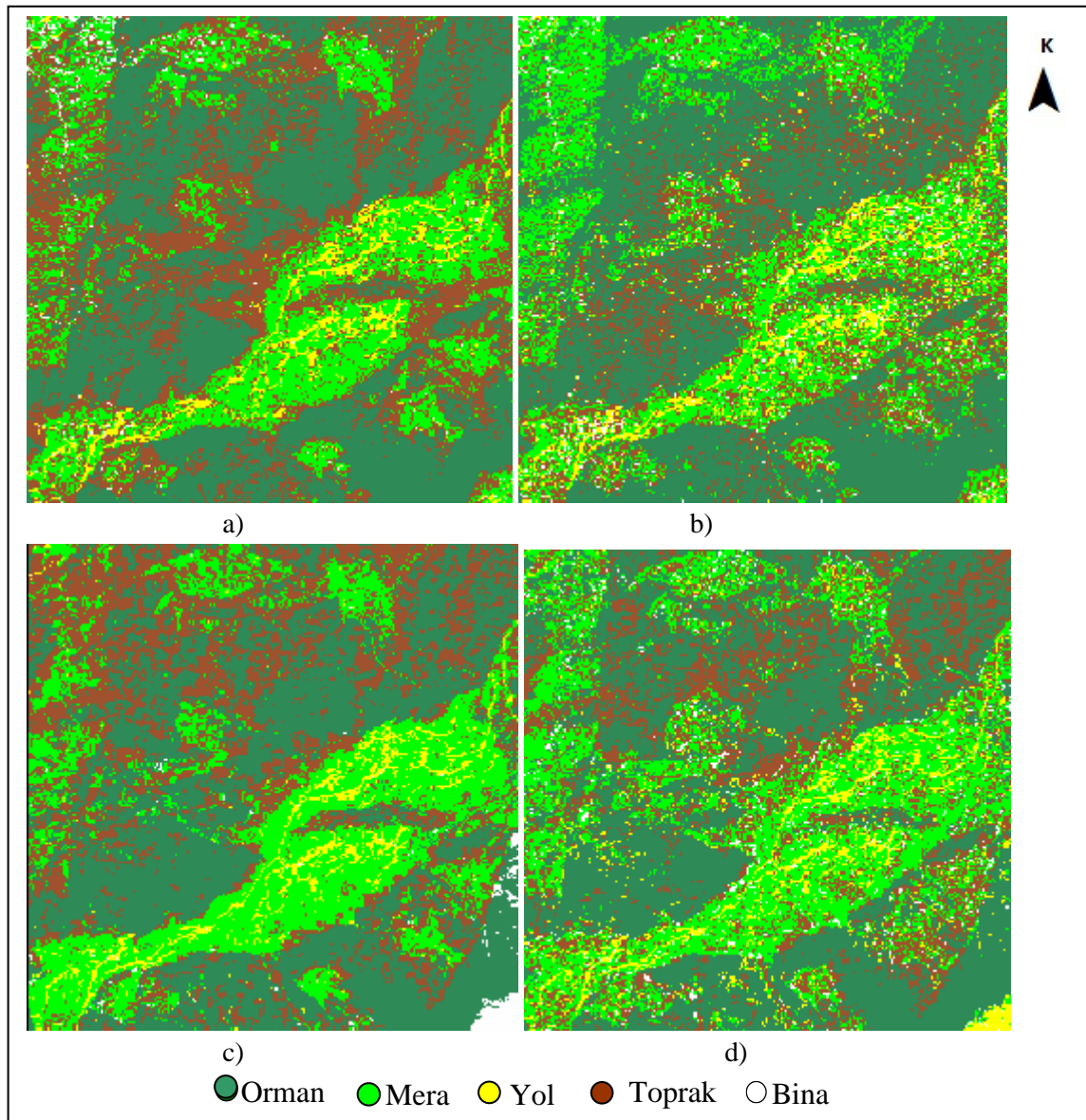
Ön işlemden sonra diğer bir aşama da Şekil 5'de görülen görüntülerin kaynaştırılması işlemidir. Görüntülerdeki arazi kullanım sınıflarını daha iyi yorumlayabilmek için, ENVI yazılımında 5m konumsal çözünürlüklü Göktürk-2 MS ve 30m konumsal çözünürlüklü Hyperion EO-1 uydu görüntüleri, Göktürk-2 PAN görüntüsüyle, GS ve PC gibi görüntü kaynaştırma yöntemleri kullanılarak kaynaştırılmıştır. Bu işlem sonucunda 2.5 m konumsal çözünürlüklü görüntüler elde edilmiştir.

Sınıflandırma işlemi için kaynaştırılmış görüntüler üzerinden 6 adet arazi kullanım sınıfı (mera, yol, orman, toprak, gölge ve bina) belirlenmiştir. ENVI yazılımında her sınıf için görüntüler üzerinden sınıflara ait örnek pikseller toplanmıştır. Mera, yol, orman ve toprak sınıfları için yaklaşık 1300 piksel, çalışma alanında daha az kullanım alanına sahip olan gölge ve bina sınıfları için ise daha az sayıda piksel toplanmıştır. Toplamda



görüntü üzerinden 6408 eğitim pikseli toplanmıştır. Sonrasında ENVI yazılımında DVM sınıflandırıcısı yardımıyla optimum sınıflandırma parametresine ( $C$ ) ve Radyal tabanlı fonksiyona göre tüm kaynaştırılmış görüntüler sınıflandırılmıştır.  $C$  parametresi kullanıcı tarafından görüntüye göre denenerek 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 6408 pikselin 3204'ü eğitim verisi, kalan 3204'ü test verisi olarak seçilmiştir ve RO sınıflandırıcısı ile görüntüler, kullanıcı tarafından belirlenen optimum parametrelere ( $m$  ve  $N$ ) göre sınıflandırılmıştır (Örn. Şekil 6).



Şekil 6. Sınıflandırılmış görüntüler a) DVM\_PC\_GMS\_GPAN, b) RO\_PC\_GMS\_GPAN, c) DVM\_PC\_DHYP\_GPAN, d)RO\_PC\_DHYP\_GPAN

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 5 sınıf için, her bir görüntü üzerinde, sınıfların kapladıkları alanlara göre nokta üreten stratified random yöntemi kullanılarak her sınıf için en az 30 nokta olacak şekilde toplamda 183 nokta üretilmiştir. Bu noktalar kullanılarak sınıflandırılmış tematik görüntülerin doğrulukları analiz edilmiştir. Her bir

sınıflandırma sonucu için hata matrisleri oluşturulmuştur (Tablo 1).

#### 4. BULGULAR ve TARTIŞMA

RO sınıflandırıcısının DVM yöntemine göre performansını değerlendirmek için doğruluk analizleri sonucu elde edilen hata matrisleri incelenmiştir. Öncelikle PC\_GMS\_GPAN ve GS\_GMS\_GPAN kaynaştırılmış görüntüleri sınıflandıran RO'nun hata matrislerine bakıldığında sırasıyla %72.13 ve % 69.40 genel sınıflandırma doğrulukları ve  $\kappa$  (0.72),  $\kappa$  (0.69) kappa olarak elde edilmiştir. DVM'nin hata matrislerine bakıldığında aynı görüntüler için sırasıyla %65.57 ve % 65.03 genel sınıflandırma doğrulukları ve  $\kappa$  (0.65),  $\kappa$  (0.65) kappa olarak elde edilmiştir. Bu sonuçlara göre RO'nun, DVM'den daha iyi performans sergilediği görülmüştür. Ayrıca PC kaynaştırma yönteminin de GS'den daha iyi sınıflandırma sonucu verdiği gözlemlenmiştir.

Tablo 1. Sınıflandırma sonucu elde edilen hata matrisleri a) RO\_PC\_GMS\_GPAN b) RO\_PC\_DHYP\_GPAN

		Referans							Ü.D., (%)	K.D.(%)
		Mera	Yol	Orman	Toprak	Bina	$\Sigma$ Sütun			
Sınıflar	Mera	23	2	0	6	4	35	56.10	65.71	
	Yol	1	21	0	0	3	25	80.77	84.00	
	Orman	0	0	60	4	0	64	75.95	93.75	
	Toprak	16	2	19	15	3	55	57.69	27.27	
	Bina	1	1	0	1	1	4	9.09	25.00	
$\Sigma$ Satır		41	26	79	26	11	183			
<b>Genel Doğruluk = %72.13 <math>\kappa=0.7208</math></b>										

a)

		Referans							Ü.D., (%)	K.D.(%)
		Mera	Yol	Orman	Toprak	Bina	$\Sigma$ Sütun			
Sınıflar	Mera	38	4	0	1	2	45	92.68%	84.44%	
	Yol	0	21	1	1	3	26	80.77%	80.77%	
	Orman	1	0	70	4	0	75	88.61%	93.33%	
	Toprak	1	0	8	20	3	32	76.92%	62.50%	
	Bina	1	1	0	0	3	5	27.27%	60.00%	
$\Sigma$ Satır		41	26	79	26	11	183			
<b>Overall Accuracy = %83.06 <math>\kappa=0.8303</math></b>										

b)

Bunun yanında indirgenmiş hiper bantlı görüntü için aynı analiz yapıldığında PC\_DHYP\_GPAN ve GS\_DHYP\_GPAN kaynaştırılmış görüntüleri sınıflandıran RO'nun hata matrislerinde sırasıyla %83.06 ve %81.97 genel sınıflandırma doğrulukları ve  $\kappa$  (0.83),  $\kappa$  (0.82) kappa olarak elde edilmiştir. DVM'nin hata matrislerine bakıldığında aynı görüntüler için sırasıyla %70.49 ve %71.58 genel sınıflandırma doğrulukları ve  $\kappa$  (0.70),  $\kappa$  (0.71) kappa olarak elde edilmiştir. Benzer şekilde bu RO daha iyi performans sergilemiştir. Çalışma alanının benzer spektral özellikteki sınıfları barındırması sınıflandırma sonuçlarını olumsuz etkilemiş; ancak genel olarak analiz sonuçlarından da anlaşıldığı gibi hiper bantlı görüntünün kullanılması genel sınıflandırma doğruluğunu önemli oranda (%11) artırmıştır (Tablo 2).

Hata matrisleri incelendiğinde; spektral özellikleri birbirine yakın sınıflarda karışmalar olduğu gözlenmiştir. Bölge toprak ve beton yollar, beton binalar çimenlik mera alanları, topraklaşmaya yüz tutmuş mera alanları, sık ve seyrek orman alanları gibi kendi içinde farklı karakteristikteki sınıfları barındırmaktadır. Bu nedenle de yol sınıfı, bina sınıfı ile, toprak sınıfı orman, mera ve yol ile, mera sınıfı orman ile karışmalar göstermiştir.

Tablo 2. Genel Sınıflandırma Doğrulukları a) DHYP\_GPAN, b) GMS\_GPAN

		Genel Doğruluk (%)	Kappa
DVM	GS	71.58	0.7158
	PC	70.49	0.7047
RO	GS	81.97	0.8194
	PC	83.06	0.8303

a)

		Genel Doğruluk (%)	Kappa
DVM	GS	65.03	0.6497
	PC	65.57	0.6353
RO	GS	69.40	0.6933
	PC	72.13	0.7208

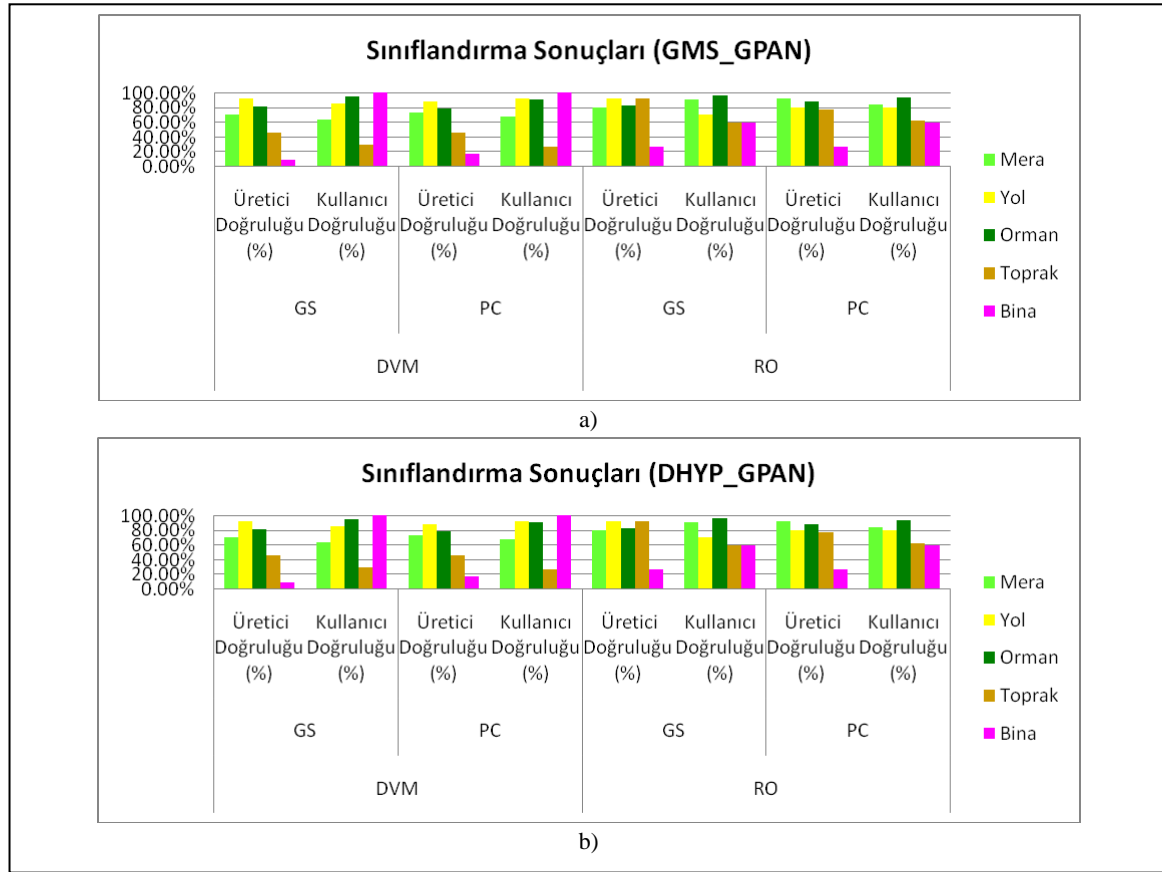
b)

Özellikle Göktürk-2 görüntüsünün bant sayısının az olması bu sınıfları birbirinden ayırmada yeterli olmamıştır. Bu nedenle sınıflandırma doğruluğu daha düşük olarak elde edilmiştir. Hiper bantlı görüntü kullanımı ile bu sınıflar birbirinden daha iyi ayırt edilmiştir. Göktürk-2 ve Hiper bantlı görüntülerinin RO'ya göre elde edilen üretici ve kullanıcı doğrulukları incelendiğinde; hiper bantlı görüntü birbiriyle en çok karışan mera sınıfını %15, yol sınıfını %13 ve toprak sınıfını %48 daha iyi temsil etmiştir. Kullanıcı doğruluklarında ise yine hiper bantlı görüntü mera sınıfını %23, toprak sınıfını %33, orman sınıfını %7 ve bina sınıfını %5 daha doğru sınıflandırmıştır (Şekil 7). Dolayısıyla en yüksek sınıflandırma doğruluğuna sahip tematik görüntüden ArcGIS yazılımında üretilen sonuç ürün olan arazi kullanım haritasının kalitesi artırılmıştır (Şekil 8).

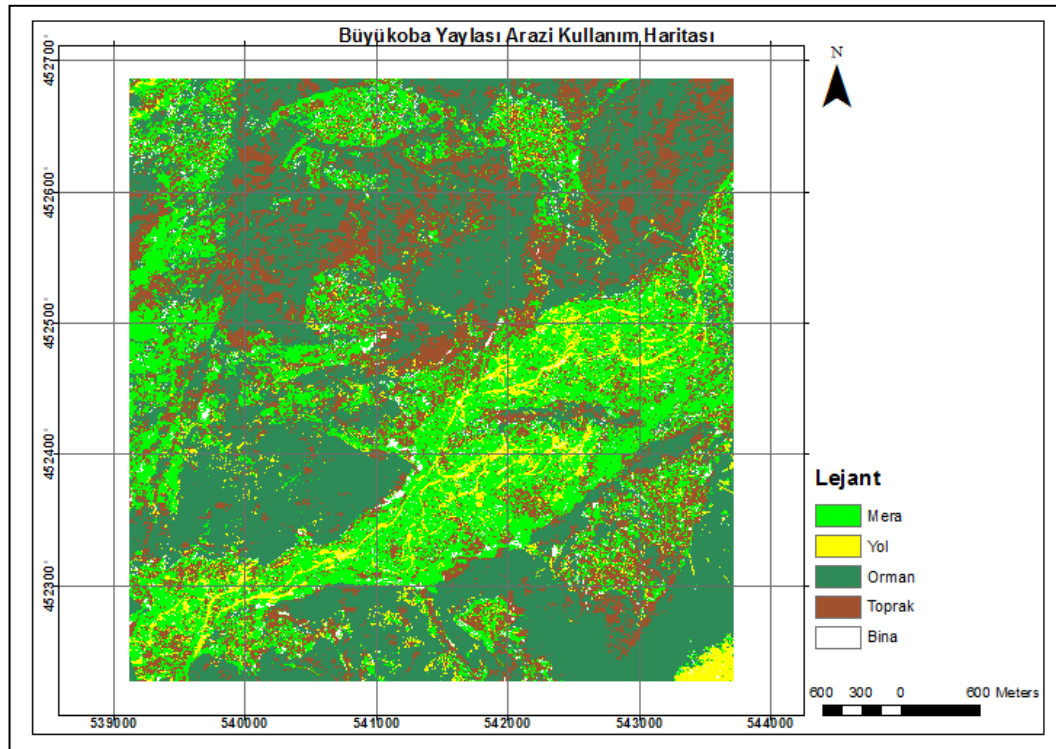
## 5. SONUÇLAR

Çalışmada en yüksek sınıflandırma doğruluğu, PC kaynaştırma yöntemine göre kaynaştırılmış görüntülerin, RO sınıflandırıcısı ile sınıflandırılması sonucu elde edilmiştir. PC ile kaynaştırılmış Göktürk-2 MS ve PAN görüntüsünün RO ile sınıflandırılması sonucu doğruluk %72.13 olarak elde edilmiştir. PC ile kaynaştırılmış indirgenmiş Hyperion EO-1 ve Göktürk-2 PAN görüntüsünün, RO sınıflandırıcısı ile de %83.06 genel sınıflandırma doğruluğuyla sınıflandırıldığı gözlenmiştir. Bu sonuçlara göre %11 oranında sınıflandırma doğruluğu artırılmıştır.

Sonuç olarak bu çalışma özellikle spektral olarak yakın özellikteki alanların arazi kullanım haritalarının oluşturulmasında, Göktürk-2 ve hiper bantlı Hyperion EO-1 uydu görüntülerinin RO sınıflandırıcısı ile birlikte kullanımındaki başarısını ortaya koymaktadır.



Şekil 7. Üretici ve kullanıcı doğrulukları a) GMS\_GPAN, b) DHYP\_GPAN



Şekil 8. Arazi kullanım haritası

**Teşekkür**

Çalışmada kullanılan Göktürk-2 uydu görüntülerinin temin edildiği Hava Kuvvetleri Komutanlığı'na ve Worldview-2 uydu görüntüsünü sağlayan Karadeniz Teknik Üniversitesi Harita Mühendisliği bölümüne teşekkür ederim.

**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.
- 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.
- 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](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](http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm) (15.11.2009).
- 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.
- 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.
- 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.
- 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..
- 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.
- 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.
- 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.
- 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.
- 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