

SULTAN SAZLIĞI MİLLİ PARKI RAMSAR BÖLGESİ ARAZİ ÖRTÜSÜNÜN BELİRLENMESİ: BOOSTING SINIFLANDIRMA YAKLAŞIMI

M. H. Kesikoğlu^a, U. H. Atasever^a, A. Kesikoğlu^a, A. E. Karkınlı^a, C. Ozkan^a, E. Beşdok^a

^a Erciyes Üniversitesi, Mühendislik Fakültesi, Harita Mühendisliği Bölümü, 38039 Melikgazi / Kayseri, Turkey - (hayrikesikoglu, uhatasever, akedikoglu, aekarkinli, cozkan, ebesdok)@erciyes.edu.tr,

ANAHTAR KELİMELER: Boosting sınıflandırma, RUSBoost, LPBoost, Karar ağaçları, Aster, Sultan sazlığı

ÖZET

Sulak alanlar pek çok canlı türünü bir arada barındıran, onların yaşamsal faaliyetlerini sürdürebilecek ortamın sağlandığı doğal mekânlardır. Bu tür alanlar yalnızca buldukları ülkenin değil tüm dünyanın doğal zenginlik kaynakları olarak bilinmektedir. Türkiye'nin en önemli sulak alanlarından bir tanesi olan ve RAMSAR sözleşmesi kapsamında koruma altına alınan Sultan Sazlığı Milli Park alanındaki arazi örtüsünün incelenmesi, bu çevrede yaşayan canlı türlerinin devamlılığı ve doğal yapının korunması açısından oldukça önemlidir. Bu çalışmada Sultan Sazlığı Milli Parkı ve RAMSAR alanı arazi örtüsünün uzaktan algılama yöntemleriyle elde edilmesi amaçlanmıştır. Bu kapsamda, 2012 yılına ait Aster uydu görüntüsü kullanılmıştır. Arazi örtüsünün doğru bir şekilde belirlenmesi görüntü sınıflandırma yönteminin doğruluğuna bağlıdır. Bu nedenle, çalışmada temel karar ağaçları ve sınıflandırma yöntemlerinin doğruluklarını artırmak için kullanılan RUSBoost ve LPBoost hızlandırma yöntemlerinin her biriyle ayrı ayrı iyileştirilmiş karar ağaçları tabanlı kontrollü sınıflandırma yöntemleri kullanılarak görüntü sınıflandırılmıştır. Kontrollü sınıflandırma yöntemlerinin doğrulukları karşılaştırılmış ve en yüksek sınıflandırma doğruluğunun LPBoost yöntemiyle elde edildiği görülmüştür. Bu nedenle arazi örtüsünü belirlemede LPBoost sınıflandırma yöntemiyle desteklenmiş Karar Ağaçları metodu tercih edilmiştir.

KEY WORDS: Boosting classification, RUSBoost, LPBoost, Decision tree, Aster, Sultan Marshes

ABSTRACT

Wetlands which host many species are natural areas. They are also an environment that can sustain vital activities of livings. These areas are either the country 's or the world 's natural wealth resources. Sultan Marshes National Park protected by RAMSAR agreement is one of the most important wetlands. Investigation of Sultan Marshes National Park(SMNP) landcover is very important for resuming the life of livings and protecting the natural environment. In this study, it is aimed that obtaining of SMNP landcover by using remote sensing techniques. In this scope, Aster satellite image belonging to 2012 year was used. The most accurate way of determining landcover varies linearly with accuracy of image classification methods. Therefore, satellite image was classified separately by using decision tree, RUSBoost and LPBoost boosting methods using to increase accuracy of classification methods. Classification accuracy of supervised classification methods were compared and it was observed that the highest accuracy was obtained with LPBoost method. For this reason, LPBoost method based on decision tree was preferred to determine landcover classification.

1. GİRİŞ

Arazi yüzeyi üzerinde yapılan çalışmalarda yeryüzüne ait elde edilen bilgilerin doğru ve hızlı bir şekilde tespit edilmesi büyük önem taşımaktadır. Bu bilgilerin tespit edilebilmesi için yersel ölçümlerin yanı sıra uzaktan algılama teknikleri kullanılarak yapılan ölçümlerle de bilgi elde etme işlemi gerçekleştirilebilmektedir. Bu sayede yüz ölçümü çok büyük olan alanlarda daha kolay ve hızlı bir şekilde sonuca ulaşmak mümkün olmaktadır. Ayrıca aynı alana ait farklı zamanlarda elde edilen veriler kullanılarak zaman içerisinde meydana gelen farklı oluşumların da tespiti mümkün hale gelmektedir.

Uzaktan algılama teknolojisinin sağlamış olduğu avantajlar bu teknolojinin kullanıldığı ölçümlerin önemini arttırmaktadır. Uzaktan algılama görüntü verilerinden arazi örtüsü ve arazi kullanımının belirlenmesindeki en önemli aşama görüntü sınıflandırmasıdır. Bir görüntüde benzer ya da aynı spektral özelliği taşıyan piksellerin arazideki karşılığını belirleyecek şekilde otomatik olarak sınıflara ayırmak ya da benzer özellik gösteren piksellerden oluşan kümelere ayırma işlemi

sınıflandırma olarak ifade edilmektedir (Özgen, 2009; Siewe, 2007). Çoğu uygulamada bir nesneyi veya bir veri grubunu belirli bir sınıfa atamak oldukça zordur. Bu işlem kontrollü sınıflandırma olarak adlandırılmaktadır (Bailly vd., 2007). Arazi yüzeyini temsil eden örnek alanların görüntü verisine uygulanarak her bir görüntü verisinin en çok benzediği veri grubunun bulunduğu sınıfa atanması işlemi kontrollü sınıflandırma olarak ifade edilir. Kontrollü sınıflandırma işlemi birden fazla değişkene sahip veri setleri üzerinde bir karar verme kural mekanizması kullanılarak gerçekleştirilmektedir. Bu doğrultuda sınıflar arası ayrımı en iyi şekilde sağlayacak olan özellikler seçilir ve sonrasında sınıflandırma yöntemi görüntüye uygulanarak her bir pikselin en çok benzediği sınıf saptanarak pikselin o sınıfın adıyla etiketlenmesi sağlanır. Son olarak sonuç tematik harita şeklinde görselleştirilir ve yine eğitim ve test alanları kullanılarak doğruluk değerlendirmesi gerçekleştirilir (Richards and Jia, 2005; Kesikoğlu, 2013). Kontrollü sınıflandırma yaklaşımlarına bakıldığında literatürde birçok sınıflandırma algoritmasının olduğu görülür (Li and etc., 2014; Lu and Weng, 2007). Günümüzde kullanılan bu algoritmalarından bir tanesi de karar ağaçları (KA) yöntemidir.

KA yöntemi uzaktan algılanmış görüntülerin hızlı ve etkili bir şekilde sınıflandırılmasını sağlayan bir yöntemdir (Breiman vd., 1984). Sınıflandırmadaki temel problem sınıflar içerisindeki dengesiz veri dağılımıdır. Bir diğer ifadeyle, eğitim verisi içerisindeki her bir farklı sınıfı gösteren verilerin temsilindeki dengesizlik bir problem oluşturmaktadır (Seiffert vd., 2008). Çoğu görüntü sınıflandırma uygulamasında bu problemle karşılaşmaktadır (Kumar, 2012). Sınıflardaki veri dağılımı dengesizliğini gidermek için hızlandırma (boosting) algoritmaları kullanılabilir (Govindaraj and Lavanya, 2013). Hızlandırma yöntemi iyi bilinen makine öğrenme tekniklerinden birisidir. Sınıflandırma doğruluğunu arttırmada kullanılan en önemli yöntemlerdendir. Arazi örtüsü sınıflandırma, medikal sınıflandırma, hesaplamalı biyoloji, bilgisayarla görme ve daha birçok alanda kullanılmaktadır (Liu and Vemuri, 2011). Literatür incelendiğinde AdaBoost, TotalBoost, SMOTEBoost, LogitBoost vb. birçok hızlandırma yöntemi kullanılarak sınıflandırma işlemi gerçekleştirildiği görülmektedir (Özkan vd., 2008; Kumar, 2012).

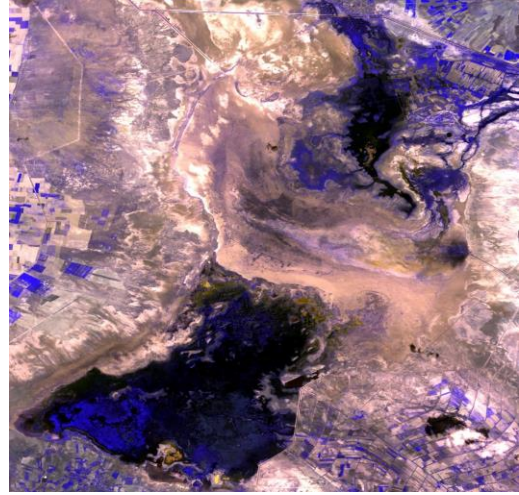
Bu çalışmada karar ağaçları ve sınıflandırma yöntemlerinin doğruluklarını arttırmak için kullanılan RUSBoost ve LPBoost hızlandırma yöntemlerinin her biriyle ayrı ayrı iyileştirilmiş KA tabanlı hibrit sınıflandırma yöntemleri kullanılarak görüntü sınıflandırılmıştır. Kontrollü sınıflandırma yöntemlerinin doğrulukları karşılaştırılmış ve en yüksek sınıflandırma doğruluğunun LPBoost yöntemiyle elde edildiği görülmüştür. Bu nedenle arazi örtüsünü belirlemede LPBoost sınıflandırma yöntemiyle desteklenmiş KA metodu tercih edilmiştir.

2. ÇALIŞMA ALANI

Sultan Sazlığı Milli parkı Kayseri ili sınırları içerisinde bulunan Develi, Yahyalı ve Yeşilhisar ilçeleri arasında bulunmaktadır. Coğrafi konumu itibarıyla 38° 12' 14" - 38° 25' 49" Kuzey enlemleri ile 35° 09' 20" - 35° 22' 20" Doğu boylamları arasında yer almaktadır. Dünya üzerinde az rastlanan tatlı ve tuzlu su ekosistemini bir arada barındıran sulak alanlardan bir tanesi olarak bilinmektedir. Ayrıca Afrika ve Avrupa arasında bulunan kuş göç yolunun kesişme noktasında bulunması da bu alanın büyük öneme sahip olduğunu göstermektedir. Yaklaşık olarak 130.000 civarında kuş türünün uğrak alanı olması ve toplamda 301 çeşit kuş türüne ev sahipliği yapması sebebiyle 1994 yılında " Özellikle Su Kuşları Yaşama Ortamı Olarak Uluslararası Önemli Sahip Sulak Alanların Korunması Sözleşmesi (RAMSAR) " alanı olarak belirlenmiştir. Böylece uluslararası öneme sahip sulak alanlardan biri haline gelmiştir (URL-1). Çalışma alanına ait çekilmiş bir fotoğraf ve kullanılan Aster uydu görüntüsü sırasıyla Şekil 1 ve 2 de görülmektedir.



Şekil 1: Sultan Sazlığı Milli Parkı Alanı görüntüsü (URL-2)

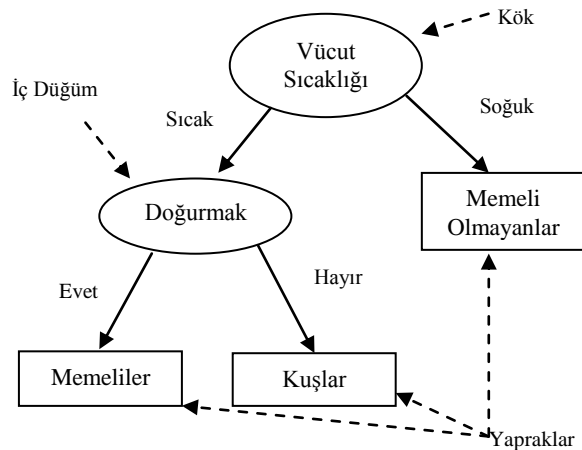


Şekil 2: Sultan Sazlığı Milli Parkı Alanı Aster uydu görüntüsü

3. YÖNTEMLER

3.1 Karar Ağaçları Yöntemi

KA analizi uzaktan algılamayla elde edilen multispektral ve hiperspektral görüntülerin sınıflandırılmasında etkin kullanım sağlayan yöntemlerden birisidir (Friedl ve Brodley, 1997; Lawrence ve Labus, 2003). 1984 yılında Breiman ve arkadaşları (Breiman vd., 1984) tarafından üretilen görece istatistiksel bir tekniktir. KA analizi sıralı karar verme modeli olarak tanımlanabilir. Bir KA şemasına bakıldığında meydana gelebilecek olaylar, alınacak kararlar, ayrıca olaylar ve kararlarla ilişkilendirilmiş sonuç değerleri görülür. İhtimaller olaylara atanır ve her bir çıktı için değerler belirlenir. Karar ağaçları yönteminin temel amacı en doğru kararı verebilmektir (URL-3). Bir KA analizinde ağaç yapısı kökten yapraklara doğru aşağı yönlü olarak karar ağacı sınıflandırıcısı tarafından sınıflandırılır. Bu ağaç yapısında kök, yaprak ve dal olarak ifade edilen üç kısım mevcuttur. Kök, veri girişi sınırlaması olmayan ya da fazlaca çıkış sınırları olan ve KA mekanizmasının en tepesinde bulunan kısımdır. Şemadaki iç düğüm olarak adlandırılan kısım bir tek girişi ve bir ya da daha fazla sınırları belli çıkışı olan birimdir. Sınırları belli bir girişi olan fakat başka çıkışı bulunmayan kısım yaprak olarak adlandırılır. Kök ve yapraklar arasındaki kısımlar ise dal olarak ifade edilir (Tan vd., 2006; Quinlan, 1993). Karar ağacında her yaprağa bir sınıf etiketi atanır. Kök ve diğer iç düğüm noktaları veriyi sınıflara ayırmaya yarayacak test durumlarını yani sınırlamaları içerirler.



Şekil 3: Örnek bir karar ağacı şeması

Şekil 3'te üç sınıfa ait iki boyutlu öznelik değerlerinden oluşan bir ağaç yapısı mevcuttur. Şekildeki vücut sıcaklığı ve doğum yapma öznelik değerlerini oluştururken sıcak, soğuk, evet, hayır gibi kriterler dallanmaları oluşturan eşik değerlerini; memeliler, memeli olmayanlar ve kuşlar ise sınıf etiketlerini oluşturmaktadırlar. Şekil 3 'teki sıcakkanlı hayvanlarla soğuk kanlı sürüngenleri birbirlerinden ayırt edebilmek için oluşturulmuş KA şemasına baktığımızda vücut ısısı soğuk olan hayvanların memeli olmayanlar etiketiyle etiketlendirilmiş yaprak kısmında toplanacağı görülmektedir. Vücut ısısı sıcak ise bir alttaki şarta geçilerek doğum yapıp yapmama durumuna yanıt aranmıştır. Böylelikle memeliler ve diğer sıcakkanlı hayvanlar grubuna dâhil olan kuşlar birbirlerinden ayırt edilmiş olur. Bir karar ağacı yapısı oluşturulunca test verisi de kolaylıkla sınıflandırılabilir. KA yapısına test verisi ağacın kök kısmından verilir. Oluşturulan karar mekanizmasına göre test verisi aşağı yönlü hareket ettirilip bir yaprağa ulaşana kadar alt birimlere ayrılarak sınıflandırma işlemi tamamlanır (Tan vd., 2006). Ayrıca Şekil 3'teki örnekte görüldüğü gibi KA sınıflandırma yöntemi "Eğer-ise" mantığıyla çalışmaktadır. Yöntemin temel amacı "Eğer-ise" mekanizmasıyla oluşturulan şartlara göre verinin uygun sınıflara en hızlı ve doğru bir şekilde ayrılmasını sağlamaktır. Karar ağaçlarında en önemli sorun dallanma olayının hangi şartlara göre yapılacağına belirlenmesidir. Bu sorunun çözümü için dallanma kriterinde "sınıflandırma ve regresyon ağaçları, entropiye dayalı algoritmalar ve bellek tabanlı sınıflandırma algoritmaları" olmak üzere üç farklı temel karar ağacı algoritması vardır. Sınıflandırma ve regresyon ağaçlarına "Twoing" ve "Gini indeksi", entropiye dayalı algoritmalara "C4.5" ve bellek tabanlı algoritmalara da "K-en yakın komşuluk yöntemi" örnek olarak gösterilebilir. Bu çalışmada dallanma kriteri olarak Gini indeksi tercih edilmiştir.

3.2 Hızlandırma (Boosting) Yöntemi

Literatürde son zamanlarda yapılan çalışmalara bakıldığında, çoğunda sınıflandırma yöntemlerinin birleştirilip yeni yaklaşımlar oluşturularak görüntü sınıflandırmalarının yapıldığı görülmektedir. Farklı yöntemlerin bir araya getirilmesiyle elde edilen sınıflandırıcılarla yapılan sınıflandırmanın genel olarak daha yüksek doğrulukta sonuçlar ürettiği görülmektedir (Zhou ve Tang, 2003; Pal ve Mather, 2003). Farklı yöntemlerin bir araya getirilmesiyle elde edilen sınıflandırma yaklaşımlarından birisi de hızlandırma (boosting) yöntemidir. Hızlandırma yöntemleri kontrollü sınıflandırmaya dayanan makine öğrenme algoritmaları olarak tanımlanmaktadır (Kumar, 2012). Hızlandırma yöntemleri zayıf sınıflandırıcıların doğrusal birleşimiyle güçlü bir sınıflandırıcı üretirler. Hızlandırma yöntemlerinin sınıflandırma doğruluğunu artırması nedeniyle bilgisayarda görme, görüntü işleme, hesaplamalı biyoloji vb. alanlarda bu yöntemlerin kullanımı artmıştır (Perrotton vd., 2009; Saffari vd., 2010).

Hızlandırma yöntemleri ağırlıklandırılmış eğitim veri setlerini kullanırlar. Başlangıçta her bir hızlandırma iterasyonunda bir zayıf sınıflandırıcı belirlenir ve bu zayıf sınıflandırıcı yanlış sınıflandırılmış vektörler göz önünde bulundurularak farklı örnek dağılımları kullanılarak eğitilirler. Daha sonrada ağırlıklandırılmış bir sınıflandırıcıyla birleştirilerek yeni bir sınıflandırıcı oluşturulur. Bu kombinasyonla oluşturulmuş sınıflandırma yaklaşımı sınıflandırıcının performansını geliştirerek yanlış sınıflandırmaları minimum düzeye düşürmeyi sağlamaktadır (Özkan vd., 2008). Literatüre bakıldığında birçok hızlandırma yönteminin olduğu görülmektedir. Bu çalışmada RUSBoost ve LPBoost yöntemleri kullanılmıştır.

3.2.1 RUSBoost Hızlandırma Yöntemi

RUSBoost Hızlandırma yöntemi görüntüyü sınıflandırırken meydana gelen sınıflar arasındaki dengesizlikleri gideren ve zayıf sınıflandırıcıların sınıflandırma performansını arttıran yöntemlerden birisidir. Sınıflar arası dengesizliği giderip sınıflandırma doğruluğunun artırılması veri örnekleme ve hızlandırma olmak üzere iki şekilde yapmak mümkündür. Veri örneklemede bu durum ya verisi az olan sınıfa veri eklemeye (oversampling) ya da verisi çok olan sınıftan veri çıkarmaya (undersampling) gerçekleştirilmektedir. Veri eklemeye ya da çıkarmaya ilgili birkaç yöntem önerilmektedir. Bu yöntemlerden bir tanesi de rastgele yeniden örneklemedir (random resampling). RUSBoost yöntemindeki RUS (random undersampling) rastgele veri çıkarmayı ifade etmektedir. Yani istenilen dengeli sınıf dağılımına ulaşana kadar yöntem rastgele olarak eğitim verisinden veri silmektedir (Seiffert vd., 2010; Weiss, 2004). RUSBoost yöntemi bir çeşit geliştirilmiş veri örnekleme tekniği olarak da kabul edilmektedir. Hızlandırma işlemi hem yeniden ağırlıklandırma hem de yeniden örneklemeyle yapılabilmektedir. Bu yöntemde eğitim veri setindeki örneklere verilen ağırlıklara göre yeniden örnekleme yapılmaktadır. RUSBoost yöntemi basitliğinin yanı sıra iyi bir performans göstermektedir. En temel avantajı yöntemlerin bir araya gelmesiyle oluşturulan modeller arasında en hızlı sonuç veren yöntemlerden biri olmasıdır. En temel dezavantajı ise veri çıkarma özelliğinden dolayı veri kayıplarına sebep olmasıdır (Weiss, 2004; Hulse, 2007).

3.2.2 LPBoost Hızlandırma Yöntemi

LPBoost hızlandırma yöntemi doğrusal programlamayı esas alan bir yöntemdir. LPBoost yöntemindeki LP ifadesi lineer (doğrusal) programlamayı ifade etmektedir. Yöntemde çok sayıda değişken olması sebebiyle ilk etapta fazla kullanımı tercih edilmeyen bir yöntem iken doğrusal programlamanın sağladığı etkin çözümlerin anlaşılmasıyla kullanımı artmaya başlamıştır. LPBoost yöntemi destek vektör makineleri yöntemine benzemektedir. Dolayısıyla destek vektör makinesi problemleri için kullanılan analiz şeması LPBoost yöntemi problemleri için de uygulanabilir (Ratsch vd., 2000). LPBoost yöntemi ikili sınıflandırma yapan yalın algoritmaların bir çeşididir. Fakat destek vektör makinelerinde olduğu gibi çok sınıflı sınıflandırma problemlerinde kullanılmak üzere uyarlanmıştır (Grove ve Schuurmans, 1998). Bu yöntemde kontrollü sınıflandırıcılar sınıflar arasındaki sınırları maksimize edecek şekilde verileri birbirinden ayırırlar. Bu sınıflandırma algoritmasında her bir iterasyonda döngüye bir zayıf sınıflandırıcı eklenir. Böylece her bir iterasyon sonucunda maliyet fonksiyonu güncellenerek sınıflandırma fonksiyonu oluşturulur. Diğer bir ifadeyle sınıflandırma fonksiyonu ağırlıklandırılmış zayıf sınıflandırıcıların doğrusal birleşiminden oluşmaktadır. Ağırlıklandırma işleminde doğruluğu net bir şekilde bilinen ya da doğru olarak sınıflandırılmış yerlerin ağırlıkları artırılırken yanlış olarak sınıflandırılmış yerlerin ağırlıkları azaltılmaktadır. Ağırlıklandırmada çoğunluğu elde eden sınıf tahmin edilen sınıf olarak ele alınır (Demiriz vd., 2002). Böylece sınıflandırma işlemi gerçekleştirilmiş olur.

4. UYGULAMA VE BULGULAR

Bu çalışma kapsamında Sultan Sazlığı Milli Park alanını içeren 2012 yılına ait Aster uydu görüntüsü kullanılmıştır. Uydu görüntüsü üzerinde tespit edilecek arazi örtüsü sınıflarının belirlenebilmesi için arazideki durumu en iyi şekilde gösteren eğitim alanları yersel çalışmalarla belirlenmiştir. Belirlenen

eğitim seti bataklık, bitki örtüsü, kahverengi toprak, kumsal, su, tarım alanı ve bulut olmak üzere yedi sınıftan oluşmaktadır. Sınıflandırma yöntemlerinin doğruluğunu belirlemek üzere test verilerinden yararlanılmıştır.

Bu çalışmada karar ağacı sınıflandırmasında dallanma kriteri olarak daha yüksek doğruluk elde edilebileceği için Gini indeksi tercih edilmiştir. Karar ağacı sınıflandırması ile çeşitli dallanmalar sonucunda 19 adet yaprak üretilmiş ve %95.52 genel doğruluk değerine ulaşılmıştır. Aynı veri seti kullanılarak RUSBoost ve LPBoost sınıflandırma yöntemlerinin her biriyle ayrı ayrı 2000 farklı kombinasyonda sınıflandırma işlemi gerçekleştirilmiş ve bunlara ait en iyi hangi kombinasyonla sınıflandırma sonucu elde edilebileceği tespit edilmiştir. Yapılan sınıflandırmalara ilişkin parametreler ise Tablo 1 ve Tablo 2' de gösterilmektedir.

Gözlem Numarası	Yaprak Sayısı	Zayıf Öğrenici	Öğrenme Oranı	Genel Doğruluk (%)
1	163	60	0.240	97.88
2	151	100	0.004	97.35
3	172	40	0.779	97.82
4	115	90	0.009	97.44
5	189	130	0.007	97.75
6	152	60	0.016	97.42
7	111	60	0.032	97.84
8	118	10	0.631	97.84
9	196	110	0.813	97.82
10	200	10	0.058	97.35

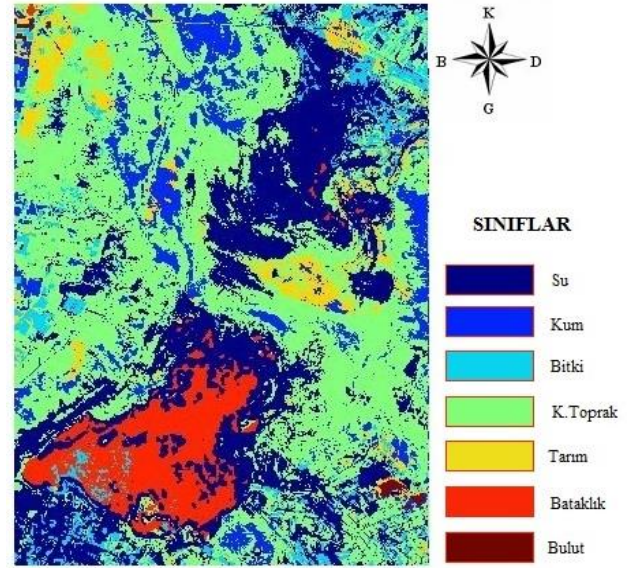
Tablo 1. RUSBoost sınıflandırmaya ait parametre değerleri

Gözlem Numarası	Yaprak Sayısı	Zayıf Öğrenici	Öğrenme Oranı	Genel Doğruluk (%)
1	70	60	0.950	98.60
2	142	250	0.301	96.80
3	96	30	0.552	96.81
4	157	60	0.994	97.31
5	172	10	0.827	97.25
6	151	80	0.963	97.37
7	126	30	0.378	98.00
8	186	400	0.959	98.02
9	85	450	0.985	98.53
10	42	310	0.912	98.49

Tablo 2. LPBoost sınıflandırmaya ait parametre değerleri

Tablo 1 ve Tablo 2 'de RUSBoost ve LPBoost sınıflandırma yöntemlerine ilişkin iterasyonlardan rastgele seçilen on farklı kombinasyon görülmektedir. Bu değerlere bakıldığında yaprak sayısı, zayıf öğrenici (weak learner) ve öğrenme oranındaki değişimler genel sınıflandırma doğruluğunu etkilediği görülmektedir. Her iki tablo da incelendiğinde Tablo 2' deki LPBoost yöntemiyle elde edilen gözlemlere göre %98.60 ile en

yüksek genel doğruluğun elde edildiği görülmektedir. Böylece uygulamamızda LPBoost yöntemiyle uygun parametre değerleri kullanıldığında genel doğruluğun temel karar ağaçları yöntemine göre yaklaşık olarak % 3, RUSBoost yöntemine göre % 1 oranında artış gösterebildiği görülmüştür. Bu nedenle arazi örtüsünü belirlemesi LPBoost yönteminin en iyi sonuç veren kombinasyonu kullanılarak gerçekleştirilmiştir. Arazi örtüsü tematik sınıflandırılmış görüntüsü Şekil 4' de gösterilmektedir.



Şekil 4. Sultan Sazlığı Milli Parkı Ramsar Bölgesi Arazi Örtüsünün sınıflandırılması ile üretilen tematik harita

5. SONUÇLAR

Arazi örtüsünün belirlenmesi doğal kaynakların korunması ve yönetimi, kent ve bölge planlaması gibi çevresel uygulamalar için büyük önem arz eden ve sıklıkla güncellenmesi gereken çalışmalardır. Uzaktan algılamayla yapılan görüntülemeyle farklı konumsal ve zamansal frekanslarda büyük coğrafi alanlara ait görüntüler elde edilebilmektedir. Bu durum bize arazi örtüsü bilgisinin görüntüler üzerinde yapılan sınıflandırma işlemleriyle elde edilebilmesine olanak sağlamaktadır. Yapılan çalışmaların çoğunda kullanılan sınıflandırma yönteminin doğruluğunun önemine değinilmiş ve kullanılacak sınıflandırma yönteminin seçiminde en önemli kriterlerden biri olduğu vurgulanmıştır. Günümüzde görüntü işleme alanında meydana gelen gelişmeler ışığında sınıflandırma yöntemlerinin doğruluğunu arttırmak üzere birleştirilmiş ya da hibrit algoritmalar olarak adlandırabileceğimiz yöntemler üretilmiştir. Bu yöntemlerden birisi de hızlandırma (boosting) yöntemidir. Bu çalışmada karar ağaçları sınıflandırma yöntemi, RUSBoost ve LPBoost hızlandırma algoritmalarıyla uydu görüntülerinin tek tek sınıflandırma doğrulukları incelenmiştir. Sınıflandırma yöntemlerinin sahip oldukları sınıflandırma kriterlerinde birçok farklı kombinasyonlar uygulanmış ve en yüksek doğruluğu veren sınıflandırma kombinasyonu araştırılmıştır. Bunun sonucunda en yüksek doğruluğun karar ağacı tabanlı bir yöntem olan LPBoost sınıflandırma algoritmasıyla elde edildiği görülmüştür. Yapılan çalışmanın sonucunda Sultan Sazlığı Milli Parkı ve Ramsar alanı arazi örtüsünün belirlenmesinde sınıflandırma algoritması olarak LPBoost yöntemi kullanılarak arazi örtüsü bilgisi temin edilmiş ve LPBoost sınıflandırma yönteminin uygun sınıflandırma parametreleri kullanıldığında sınıflandırma doğruluğunu arttırdığı ortaya konmuştur.

TEŞEKKÜR

Yazarlar, Erciyes Üniversitesi Bilimsel Araştırmalar Birimi'ne teşekkür eder. Bu araştırma Erciyes Üniversitesi EUBAP FYL-2013-4312 proje kodu ve Sultan Sazlığı Milli Parkı ve RAMSAR Alanı Kıyı Değişiminin Uydu Görüntü Analizleriyle İncelenmesi adlı proje ile desteklenmiştir.

KAYNAKLAR

- Bailly J. S., Amaud M. and Puech C., 2007. Boosting: a classification method for remote sensing, *Int. J. Remote Sensing*, 28(7), 1687-1710.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. 1984. *Classification and regression trees*. Belmont, CA: Wadsworth International Group.
- Demiriz A., Bennett K. P., Taylor J. S., 2002. Linear programming boosting via column generation. *Machine Learning.*, 46, 225-254.
- Friedl, M. A. and Brodley, C. E. 1997. Decision tree classification of landcover from remotely sensed data. *Remote Sensing of Environment*, 61, 399 – 409.
- Govindaraj M. and Lavanya S., 2013. A combined boosting and sampling approach for imbalanced data classification, *Int. J. Adv. Res. in Data Mining and Cloud Computing*, 1(1), 44-50.
- Grove A. ve Schuurmans D. 1998. Boosting in the limit: Maximizing the margin of learned ensembles. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence*, AAAI-98, USA.
- Khoshgoftaar T. M., Hulse J. V. and Napolitano A., 2007. Experimental perspectives on learning from imbalanced data, 24th International Conference on Mach. Learn., Corvallis, USA, ss. 935-942.
- Kesikoğlu M. H., (2013). Sultan Sazlığı Milli Parkı ve Ramsar Alanı Kıyı Değişiminin Uydu Görüntü Analizleriyle İncelenmesi. Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Kayseri, 74 s.
- Kumar B. S., 2012. Boosting techniques on rarity mining. *Int. J. Adv. Res. in Computer Science and Software Engineering*, 2(10), 27-35.
- Lawrence, R., and Labus, M. 2003. Early detection of douglas-fir beetle infestation with subcanopy resolution hyperspectral imagery. *Western Journal of Applied Forestry*, 18, 202 – 206.
- Liu M. and Vemuri B. C., 2011. Robust and efficient regularized boosting using total bregman divergence, *IEEE Conference on Computer Vision and Pattern Recognition*, Colorado Springs, USA, ss. 2897-2902.
- Lu D., Weng Q., 2007. A survey of image classification methods and techniques for improving classification performance, *Int. J. Remote Sensing*, 28(5), 823-870.
- Özgen K., 2009. Dikili-Kaynarca Ve Bergama-Ovacık Hidrotermal Alterasyon Zonlarının Uzaktan Algılama Ve Coğrafi Bilgi Sistemleri İle Değerlendirilmesi. Dokuz Eylül Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, İzmir, 89 s.
- Özkan C., Sunar F., Berberoğlu S. ve Dönmez C., 2008. Effectiveness of boosting algorithms in forest fire classification, *ISPRS Conference Vol. XXXVII. Part B7*. Beijing, China, ss. 625-630.
- Pal M., Mather P.M., 2003. An assessment of the effectiveness of decision tree methods for land cover classification, *Remote Sensing of Environment*, 86(4), 554-565.
- Perrotton X., Sturzel M. ve Roux M., 2010. Implicit hierarchical boosting for multi-view object detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Francisco, CA, USA, ss. 958 - 965.
- Quinlan J. R., 1993. *C4.5: Programs for Machine Learning*, Morgan Kaufmann, San Mateo, CA.
- Ratsch, G., Scholkopf, B., Smola, A., Muller, K.-R., Onoda, T., & Mika, S., 2000. V-Arc ensemble learning in the presence of outliers. *Advances in neural information processing systems 12*. Cambridge, MA: MIT Press.
- Richards J. A., Jia X., 2005. *Remote Sensing Digital Image Analysis 4th Edition*, Springer, Germany. 439 pp.
- Saffari A., Leistner C. ve Bischof H., 2009. Regularized multi-class semi-supervised boosting. *IEEE Conference on CVPR*, Miami, FL, USA, ss. 967-974.
- Seiffert C., Khoshgoftaar T. M., Hulse J. V. and Napolitano A., 2010. Rusboost: A hybrid approach to alleviating class imbalance, *IEEE Transactions On Systems*, 40(1), 185-197 pp.
- Siewe S. S., 2007. *Change Detection Analysis Of The Land use and Land cover Of The Fort Cobb Reservoir Watershed*. Oklahoma State University, Master of Science, USA, 74 pp.
- Seiffert C., Khoshgoftaar T. M., Hulse J. V., Napolitano A., 2008. RUSBoost: Improving classification performance when training data is skewed, *19th International Conference on Pattern Recognition*, Tampa, FL, USA, ss. 1-4.
- Tan P-N., Steinbach M. and Kumar V., 2006. *Introduction to Data Mining*. Addison-Wesley Companion Book Site, USA, pp. 150-155.
- URL-1: Sultan Sazlığı Milli Parkı, http://www.ormansu.gov.tr/osb/Dosya/Sulltan_Sazligi_Milli_Parki.pdf, (Erişim tarihi: Temmuz 2013).
- URL-2: Sultan Sazlığı, http://www.kayseri.gov.tr/default_b0.aspx?content=1025 (Erişim tarihi: Temmuz 2013).
- URL-3: Introduction to Decision Trees, <http://www.treeplan.com/chapters/introduction-to-decision-trees.pdf>, (Erişim tarihi: Mart 2015).
- Weiss G.M., 2004. Mining with rarity: A unifying framework, *ACM SIGKDD Explorations Newsletter.*, 6(1), 7-19.
- Zhou Z.H., Tang W., 2003. Selective ensemble of decision trees *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing Lecture Notes in Computer Science*, 2639, ss. 476-483.