

Eğitici Sınıflandırma Yöntemlerinin Dengeli ve Dengesiz SAR Veri Kümelerindeki Başarımlarının Karşılaştırılması

Comparing the Classification Performances of Supervised Classifiers with Balanced and Imbalanced SAR Data Sets

Mustafa Üstüner*, Ünsal Gökdağ^{†‡}, Gökhan Bilgin[†], Füsün Balık Şanlı*

*Department of Geomatic Engineering, Yıldız Technical University, 34220 Istanbul

[†]Department of Computer Engineering, Yıldız Technical University, 34220 Istanbul

[‡]Seven Bridges Genomics, 34220 Istanbul

{mustuner,gbilgin,fbalik}@yildiz.edu.tr, unsal.gokdag@ieee.org

Özetçe —Bu çalışmada, çift polarizasyona sahip Sentinel-1B SAR (Sentetik Açıklıklı Radar) uydu görüntüsünün sınıflandırmasında dört farklı (iki adet dengeli ve iki adet dengesiz yapıdaki) eğitim veri kümesinin sınıflandırma doğruluklarına olan etkileri incelenmiştir. Yer gerçekliği verisinin toplanmasında bazen bazı sınıf türleri için çok sayıda örnek piksel toplamak mümkün olabilirken, bazı sınıflar için bu mümkün olmamaktadır ve bu sınıflar az sayıda piksel ile temsil edilmektedir. Bu tarz durumlarda da dengesiz yapıda veri kümeleri oluşmaktadır. Eğitici sınıflandırma işlemi doğası gereği sınıfları birbirinden ayıran karar yüzeylerinin belirlenmesinde birden fazla girdi parametreyi birlikte kullanmaktadır. Seçilecek modelden ziyade, eğitim verisinin boyutu, dağılımı ve sınıfları ne doğrulukta temsil edip etmediği, kullanılacak sınıflandırma modeli ile eğitim verisinin uyumu vb. parametreler sınıflandırma doğruluğunu etkilemektedir. Çalışmamızda görüntülerin sınıflandırma işleminde iki tane parametrik (Naive Bayes sınıflandırıcı ve doğrusal ayırma analizi) ve iki tane de parametrik olmayan (destek vektör makineleri ve rastgele orman algoritması) sınıflandırma yöntemi seçilmiştir. Deneysel sonuçlarımız, kullanılan parametrik yöntemler ve destek vektör makineleri yöntemi için sınıflandırma doğruluklarında eğitim verisine bağlı olarak ciddi bir değişim olmadığını, rastgele orman algoritmasında ise dengesiz yapıdaki eğitim veri kümesinin kullanılması ile sınıflandırma doğruluğunda bir artış olduğunu göstermiştir. Ayrıca en yüksek sınıflandırma doğruluğu (%89.94) destek vektör makineleri ile elde edilmiştir.

Anahtar Kelimeler—Uzaktan algılama, dengesiz yapıda eğitim verisi, sentetik açıklıklı radar (SAR), sınıflandırma.

Abstract—In this study, the classification accuracies of four different classification methods with two balanced and two imbalanced data sets for the classification of Sentinel-1B SAR (Synthetic Aperture Radar) data were comparatively evaluated and the impacts of training data sets into the accuracy were investigated. In some circumstances, it is possible to collect high number of ground truth samples for some classes however not possible for some other classes which are represented by less

number of ground truth samples. In such cases, the imbalanced data set is the issue. Supervised classifiers, by its nature, employ many different input parameters in consideration of the decision surface separating the two classes. More than the classification model itself, purity, size and allocation of ground truth samples as well as the adaptation between the training data and adopted classifier are of key importance in accuracy of image classification. In our study, two parametric (Naive Bayes and Linear Discriminant Analysis) and two non-parametric (Support Vector Machines and Random Forests) supervised classification methods were implemented. Our experimental results demonstrated that there were not any significant change in classification accuracies of parametric classifiers and support vector machines however an increase in classification accuracy of random forest with imbalanced dataset. Furthermore, highest classification accuracy of this study (89.94%) was obtained by Support Vector Machines classification.

Keywords—Remote sensing, imbalanced training data, synthetic aperture radar (SAR), classification.

I. GİRİŞ

Avrupa Uzay Ajansı'nın (ESA) Kopernik çevre izleme programı kapsamında fırlattığı ve 6 günlük tekrar görüntüleme süresine (dünya üzerinde aynı noktanın tekrar görüntülenme süresi) sahip Sentinel-1A ve Sentinel-1B uydu çifti ile Sentetik Açıklıklı Radar (SAR) görüntüleri uzaktan algılama alanındaki araştırmacıların büyük ilgisini çekmekte ve önemli avantajlar sağlamaktadır. Ücretsiz olarak elde edilebilen Sentinel-1 C bant SAR görüntüleri, sık görüntüleme sıklığına sahip olması ve beraberinde ücretsiz olarak sağlanan SNAP yazılımı ile birlikte arazi örtüsü/kullanımı, tarım, ormancılık, jeoloji ve deprem ile ilgili çalışmalarda sıklıkla kullanılmaktadır. Ayrıca arazi örtüsü/kullanımındaki zamansal değişim ve tarım alanlarının izlenmesi vb. gibi sık zaman aralıklı görüntülemeye ihtiyaç duyulan çalışmalar için özellikle tercih edilmek-

tedir. İklim koşullarından (bulut ve sis vb.) etkilenmeden gece/gündüz görüntüleme özelliğine sahip olması nedeniyle aktif uzaktan algılama sistemleri, pasif sistemlerdeki kısıtlamaların önüne geçmektedir. Pasif sistemler ile (optik uzaktan algılama) yüksek enleme sahip veya tropikal bölgelerdeki yüksek bulutluluk oranı nedeniyle görüntü alımında kısıtlamalar yaşanmaktadır ve bu noktada aktif sistemler önemli avantajlar sağlamaktadır [1]. Yeryüzüne ilişkin arazi örtüsünün hızlı, güvenilir ve ucuz bir şekilde görüntülenmesi ve bilgi elde edilmesinde önemli avantajlar sağlayan uzaktan algılama teknolojisi, son yıllarda artan uydu sayısı ve yapay zekâ alanındaki veri işleme hızındaki/kapasitesindeki gelişmeler ile birlikte giderek önem kazanmaktadır. Uzaktan algılama da arazi örtüsü ve kullanımına ilişkin bilginin çıkarılmasında kullanılan yöntemlerin başında eğitici sınıflandırma gelmektedir. Görüntü sınıflandırma, piksel veya piksel gruplarının yeryüzünde karşılık geldikleri arazi örtüsü/kullanımı sınıfına atanması işlemidir. SAR görüntülerinde benekler (speckle) görüntü kalitesini düşürdüğünden, sınıflandırma işlemini ve doğruluğunu olumsuz yönde etkilemektedir [2].

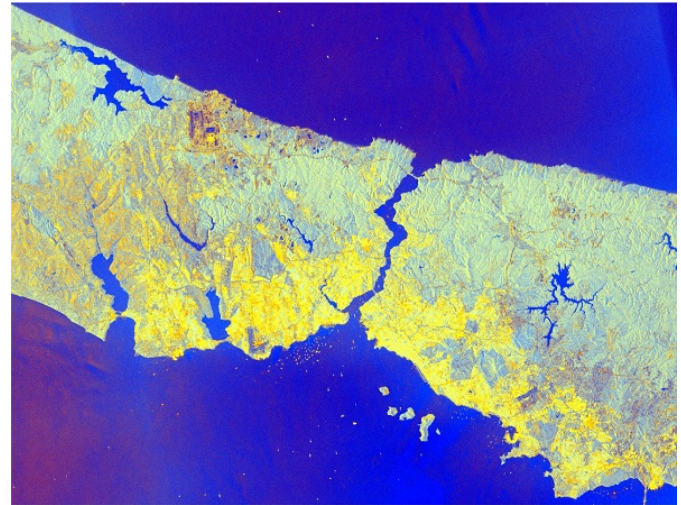
Eğitici sınıflandırma algoritmaları hem algoritmanın öğrenme kısmında hem de sınıflandırılmış görüntülerin doğruluklarının test edilmesi aşamasında doğru referans veriye (yer gerçekliği bilgisi) ihtiyaç duymaktadır [3]. İleri düzeyde sınıflandırma algoritmaları (aktif öğrenme, derin öğrenme, yarı-eğitici öğrenme vb.) bulunmasına rağmen, sınıflandırma işlemi doğası gereği sınıfları birbirinden ayıran karar yüzeylerinin belirlenmesinde birden fazla girdi parametreyi birlikte kullandığından dolayı sınıflandırma işlemi halen istenilen basit bir yapıya ulaşamamıştır. Görüntünün çözünürlüğü, eğitim verisinin boyutu ve sınıfları ne doğrulukta temsil edip etmediği, arazi örtüsü sınıflarının heterojen yapısı, kullanılacak sınıflandırma modelinin karakteristik özelliği ve eğitim verisi ile uyumu vb. parametreler sınıflandırma doğruluğunu etkilemektedir [4]. Bu parametreler içinde eğitim verisinin kalitesi, boyutu ve arazi örtüsü sınıflarını ne derece doğru temsil ettiği sınıflandırma işleminde büyük öneme sahiptir. Sınıflandırma modelinin doğruluğu, modelin ne derece doğru eğitildiği ile ilgilidir. Bu noktada kullanılacak sınıflandırma algoritmasının/modelinin yapısından ziyade eğitim verisinin karakteristik özellikleri (kalitesi, boyutu, sadeliği) ve kullanılacak olan sınıflandırma algoritması ile olan uyumu ön plana çıkmaktadır [5]. Eğitici sınıflandırma yöntemleri, sınıflandırma algoritmasının öğrenme kısmında eğitim verisindeki bilgileri kullanım şekillerine göre parametrik (istatistiksel) ve parametrik olmayan (yapay sinirsel) yöntemler olmak üzere ikiye ayrılmaktadır. Parametrik yöntemler (en yakın komşuluk, maksimum olasılık vb.) eğitim verisindeki örneklerden elde edilmiş bazı istatistiksel değerleri (ortalama değer, değişim vb.) kullanmaktadır ve verinin dağılımına ilişkin bazı varsayımlar (eğitim verisinin normal/Gauss dağılımında olduğu vb.) yapmaktadır. Parametrik olmayan yöntemler (destek vektör makineleri, rastgele orman algoritması, karar ağaçları vb.) ise parametrik yöntemlerin aksine verinin dağılımına ilişkin herhangi bir varsayımda bulunmayıp eğitim verisinin kendisini kullanmaktadır [5]. Eğitim veri kümesinin toplanması işlem adımı bazen bazı sınıf türleri için çok sayıda örnek piksel toplamak mümkün olabilirken, bazı sınıflar için bu mümkün olmamaktadır ve bu sınıflar az sayıda piksel ile temsil edilmektedir. Bu tarz durumlarda da dengesiz yapıda veri kümeleri (farklı sayıda örnek piksellere sahip sınıflardan

oluşan veri kümesi) oluşmaktadır. Dengesiz yapıdaki veri kümesinde sınıflara ait örnek piksel sayılarının arasındaki fark fazla ise, bu durum parametrik sınıflandırma yöntemlerinde hem aşırı uyum gösterme (over-fitting) problemine hem de işlem süresinin uzamasına neden olabilir. Eğitici öğrenme algoritmalarının öğrenme sürecindeki belirsiz ve karmaşık yapısından dolayı eğitim verisinde sınıf türlerine ilişkin en uygun (optimum) sayının önceden bilinmesi mümkün değildir [6] [7].

Bu çalışmada, çift polarizasyona sahip Sentinel-1B SAR uydu görüntüsünün sınıflandırmasında dört farklı (iki adet dengeli ve iki adet dengesiz yapıdaki) eğitim veri kümesinin sınıflandırma doğruluklarına olan etkileri incelenmiştir. Çalışma kapsamında, sadece orijinal bantlar (DY, DD, D: Düşey, Y: Yatay) kullanılmıştır. Görüntülerin sınıflandırma işleminde ise iki tane parametrik (Naive Bayes sınıflandırıcı ve doğrusal ayırma analizi) ve iki tane de parametrik olmayan (destek vektör makineleri ve rastgele orman algoritması) sınıflandırma yöntemi seçilmiştir. Dengeli ve dengesiz veri kümelerinin sınıflandırma doğrulukları üzerinde nasıl bir etkisi olduğu izlenmiş ve elde edilen sonuçlar karşılaştırılarak analiz edilmiştir.

II. ÇALIŞMA ALANI VE KULLANILAN VERİLER

Hızlı nüfus artışı ve çarpık kentleşme ile birlikte, özellikle son yıllarda metropol şehirlerde arazi örtüsü ve kullanımına ilişkin ciddi değişimler olmaktadır. Şehirlere ilişkin kentsel planlama, çevre yönetimi ve doğal kaynakların (ormanlık alanlar, sulak alanlar vb.) doğru, hızlı ve etkin yönetimi için güncel, doğru ve ucuz elde edilebilen arazi örtüsü/kullanımı bilgisine ihtiyaç duyulmaktadır [8], [9]. Sinoptik görüş özelliği ve sık zaman aralıklarında hedef bölgeye ilişkin veri sağlayabilmesi nedeniyle uzaktan algılama teknolojisi bu noktada önemli avantajlar sağlamaktadır. Çalışma alanı olarak İstanbul seçilmiştir.



Şekil 1: Çalışma Alanı (3 Ocak 2018-Sentinel-1B).

Çalışma kapsamında çift polarizasyona (DY, DD) sahip Sentinel-1B SAR uydu görüntüsü kullanılmıştır. Görüntü alım tarihi 3 Ocak 2018 ve görüntü interferometrik geniş aralık (interferometric wide swath) modunda ESA'dan ücretsiz olarak

elde edilmiştir. Çalışmada kullanılan SAR görüntüsü işleme seviyesine göre ESA'dan Seviye-1 GRD (ground range detected) olarak indirilmiştir. İndirilen orijinal veri 10 metre piksel aralığına (pixel spacing) sahiptir. Görüntünün orijinal uzamsal (spatial) çözünürlüğü azimut yönünde 22 metre ve menzil yönünde 20 metredir (bakış sayısı: 5×5). Seviye-1 GRD verileri, orijinal SAR görüntüsüne çoklu bakış (multi-looking) işlemi uygulanmış halidir ve görüntüde faz bilgisi mevcut değildir.

Çalışma alanına ait arazi örtüsü/kullanımı sınıf bilgileri CORINE 2012 arazi örtüsü verisinden alınmış ve daha sonra 5 farklı sınıf grubunda (tarım alanları, çıplak alanlar, ormanlık alanlar, kent ve su) birleştirilmiştir. Yer gerçekliği (ground-truth) bilgisi Google Earth (GE)'den sağlanan yüksek çözünürlüklü uydu görüntülerinden faydalanarak toplanmıştır. Toplam yer gerçekliği veri kümesinin %70'i test verisi ve %30'u da eğitim verisi olarak ayrılmıştır. Toplam eğitim verisi üzerinden de dört farklı eğitim veri kümesi oluşturulmuştur. Oluşturulan eğitim veri kümelerinden iki adet dengeli yapıda (her bir sınıf için eşit sayıda piksel sayısına sahip olma durumu) ve iki adet de dengesiz yapıda (her bir sınıf için farklı sayıda piksel sayısına sahip olma durumu) veri kümesi oluşturulmuştur. Yer gerçekliği veri kümesine ilişkin bilgiler Tablo I'de görülmektedir.

TABLO I: Yer Gerçekliği Bilgisi

Sınıf No	Sınıflar	Piksel sayıları
1	Tarım Alanları	14697
2	Çıplak Alanlar	5655
3	Ormanlık Alanlar	19453
4	Yerleşim	22101
5	Su	80765

Toplam eğitim veri kümesinin farklı sayı ve oranda alınarak oluşturulmuş olan dengeli ve dengesiz yapıdaki veri kümeleri de Tablo II'de sunulmaktadır. Pikseller rastgele örnekleme (random sampling) tekniği ile seçilmiştir.

TABLO II: Dengeli ve Dengesiz Yapıdaki Eğitim Verileri

Veri Kümesi No	Yapı	Piksel Bilgisi (Açıklama)
1	Dengeli	500 piksel/sınıf
2	Dengeli	1000 piksel/sınıf
3	Dengesiz	Mevcut verinin %10'u oranında piksel
4	Dengesiz	Mevcut verinin %20'si oranında piksel

III. SAR VERİLERİNİN ÖNİŞLENMESİ VE SINIFLANDIRILMASI

Bu bölümde Sentinel-1B SAR görüntüsünün ön işleme adımları ve ardından veri sınıflandırma işlemi anlatılacaktır. Sınıflandırma işleminde dört farklı (iki parametrik ve iki parametrik olmayan) yöntem (Naive Bayes sınıflandırıcı, doğrusal ayırma analizi, destek vektör makineleri ve rastgele orman algoritması) kullanılmıştır.

A. SAR Verilerinin Ön İşleme Adımları

Birinci işleme seviyesinde (GRD) indirilen SAR uydu görüntülerine, bazı ön işleme adımlarının (radyometrik kalibrasyon, arazi düzeltmesi vb.) uygulanması gerekmektedir. Bu işlem adımları Avrupa Uzay Ajansı'nın ücretsiz olarak

sağladığı SNAP (Sentinel Application Platform) yazılımı ile gerçekleştirilmiştir. Sentinel-1 SAR görüntüleri ESA tarafından SAFE (Standard Archive Format for Europe) formatında kullanıcılara sunulmaktadır. SAFE formatında SNAP yazılımına aktarılan görüntü için ilk olarak verinin radyometrik kalibrasyonu yapılmıştır ve sigma nought (σ^0) kalibrasyon modeli kullanılmıştır. Görüntüdeki benek (speckle) etkisini azaltmak amacıyla görüntüye 5×5 boyutunda Lee filtresi uygulanmıştır. Kullanılan Seviye 1 (GRD) görüntü çoklu bakış uygulanmış (multilooked) bir veri olduğu için görüntüye tekrar çoklu bakış (multilooking) işlemi uygulanmamıştır. Bu işlem adımlarının ardından görüntüye Range-doppler arazi düzeltmesi (range-doppler terrain correction) uygulanmıştır. Arazi düzeltmesi işlem adımı ise 30 metrelik SRTM (Shuttle Radar Topography Mission) sayısal yükseklik modeli kullanılmıştır. Son işlem adımı olarak ise bantlardaki piksel dijital değerleri, radar gerisayılım (dB) değerine dönüştürülmüştür. Elde edilen görüntüler 10 metre den 20 metreye yeniden örneklenmiştir. Görüntü çalışma alanını kaplayacak şekilde kesilmiş ve Geo-TIFF formatında kaydedilmiştir.

B. Sınıflandırma

Bu çalışmada kapsamında dengeli ve dengesiz yapıdaki veri kümelerinin sınıflandırma algoritmaları üzerindeki etkilerini incelemek amacıyla iki tane parametrik (Naive Bayes sınıflandırıcı ve doğrusal ayırma analizi) ve iki tane de parametrik olmayan (destek vektör makineleri ve rastgele orman algoritması) sınıflandırma yöntemi seçilmiştir.

Parametrik yöntemler eğitim verisinden üretilmiş bazı istatistiksel parametreleri kullanır iken parametrik olmayan yöntemler eğitim verisinin kendisini kullanmaktadır. O nedenle eğitim verisinin sınıf türlerini ne derece doğru temsil ettiği ve boyutu sınıflandırma algoritmaları için önem arz etmektedir. Naive Bayes (NB) sınıflandırma algoritması Bayes teoremi ve en büyük sonsal (maximum posteriori) hipotezine dayalıdır. Basit yapısı ile görüntü tanımadaki sıklıkla kullanılan eğitimci bir sınıflandırma tekniğidir. Bu yöntemde, piksellerin hangi sınıfa hangi olasılıkla ait oldukları (sınıf üyelik olasılığı) hesaplanır ve verilerin normal (Gauss) dağılımı oldukları varsayılır. Pixel en yüksek sonsal olasılığa (posteriori probability) sahip olduğu sınıfa atanır [10], [11]. Doğrusal ayırma analizi (LDA) görüntü sınıflandırma amacıyla sıklıkla kullanılan bir yöntemdir ve bu yöntemde ayırma fonksiyonları Fisher oranına (sınıflar arası dağılım matrisinin sınıf içi dağılım matrisine olan oranına (Raleigh katsayısı) dayalı olarak oluşturulur. Yöntem yüksek boyutlu veriyi daha düşük boyutlu uzaya taşır ve bu sayede sınıfların birbirinden iyi derecede ayrılabilmesi sağlanır [12], [13] Topluluk öğrenme algoritmalarından biri olan rastgele orman (random forest) algoritması ağaç tabanlı bir yöntemdir. Hızlı, gülbüz ve uygulanabilirliği kolay bir yöntem olması nedeniyle uzaktan algılama da sınıflandırma işleminde sıklıkla tercih edilmektedir. Bu yöntemde, sınıflandırma işleminde birden fazla karar ağacı kullanılır. Rastgele orman (RO) algoritması, girdi veriyi (eğitim verisi ve görüntü) rastgele olarak alt kümelerle böler ve birden fazla sınıflandırma sonucu elde eder. Sonuç sınıflandırmayı ise oy çokluğuna dayalı sistemden (majority voting) faydalanarak üretir [14], [15]. Makine öğrenme algoritmalarından birisi olan Destek Vektör Makineleri (DVM), iki sınıfı birbirinden ayırabilen en uygun hiperdüzlemin belirlenmesi esasına dayanan çekirdek (kernel) tabanlı bir sınıflandırma yöntemidir ve çok boyutlu verilerin

TABLO III: Veri Kümelerine Göre Sınıflandırma Doğrulukları

Yöntem	Veri Kümesi No	Toplam Doğruluk (%) (Başarım Oranı)	Kappa Değerleri
NB	1	88.65	0.8188
	2	88.39	0.8147
	3	88.74	0.8193
	4	88.73	0.8192
LDA	1	87.97	0.8078
	2	88.05	0.8091
	3	88.88	0.8209
	4	88.94	0.8220
RO	1	86.46	0.7845
	2	86.66	0.7879
	3	88.04	0.8082
	4	88.21	0.8109
DVM	1	88.87	0.8229
	2	89.11	0.8264
	3	89.91	0.8382
	4	89.94	0.8385

sınıflandırılması işleminde sıklıkla tercih edilmektedir. İki sınıfın doğrusal düzlemler ile ayrılabilmesi durumunda, DVM çekirdeklerden faydalanarak veriyi yüksek boyutlu uzaya taşıyarak ve doğrusal düzlemler ile ayrılmasına olanak sağlar [16], [17]. Çalışma kapsamında radyal tabanlı fonksiyon çekirdeği kullanılmıştır. Radyal tabanlı fonksiyon çekirdeği için ceza (cost) parametresi 1 ve gamma parametresi 1/2 olarak alınmıştır. Optimum parametrelerin bulunmasında deneme-yanılma (trial-error) metodundan faydalanılmıştır.

IV. SONUÇLAR VE DEĞERLENDİRME

Dengeli ve dengesiz yapıdaki veri kümelerinin dört farklı sınıflandırma yöntemi üzerindeki sınıflandırma başarımları Tablo III de gösterilmiştir. Tablo III genel olarak incelendiğinde dengesiz yapıdaki veri kümeleri ile dengeli yapıdaki veri kümelerine kıyasla daha yüksek başarımlar elde edildiği gözlenmiştir.

Bu çalışmada, çift polarizasyona sahip Sentinel-1B SAR (Sentetik Açıklıklı Radar) uydusu görüntüsünün sınıflandırılmasında dört farklı (iki adet dengeli ve iki adet dengesiz yapıdaki) eğitim veri kümesinin sınıflandırma doğruluklarına olan etkileri incelenmiştir. En yüksek sınıflandırma doğruluğu DVM yöntemi ile elde edilirken (%89.94), en düşük sınıflandırma doğruluğu RO algoritması ile (%86.46) elde edilmiştir. Kullanılan parametrik yöntemler ve DVM yöntemi için sınıflandırma doğruluklarında eğitim verisine bağlı olarak ciddi bir değişim olmadığı gözlenmiştir. Rastgele orman algoritmasında ise dengesiz yapıdaki eğitim veri kümesinin kullanılması ile sınıflandırma doğruluğunda bir artış olduğunu tespit edilmiştir. Yer gerçekliği verisinin dengesiz yapıda olduğu durumlarda, parametrik olmayan yöntemlerden DVM ile en yüksek sınıflandırma doğruluğu elde edilmiş ve bu yöntem için sınıflandırma doğruluğunun eğitim veri kümesine aşırı bağlı olmadığı sonucuna varılmıştır. Bu sonucun asıl gerekçesi, DVM yönteminin kendi karakteristik özelliği olan iki sınıfı ayırırken birbirine paralel hiperdüzlemler üzerindeki destek vektörlerinden faydalanmasıdır. Gelecek çalışmalarda dengeli ve dengesiz yapıdaki eğitim veri kümelerinin sınıflandırma doğruluğuna olan etkilerinin sınıf sayısı daha fazla olan tematik çalışma alanlarında (tarımsal ürünler, farklı ağaç türleri, toprak sınıfları vb.) farklı algoritmalar ile test edilmesi hedeflenmektedir.

Sentinel 1B SAR uydusu görüntüsünün ücretsiz temini için Avrupa Uzay Ajansına (ESA) teşekkür ederiz.

KAYNAKÇA

- [1] H. Balzter, B. Cole, C. Thiel, and C. Schmullius, "Mapping CORINE land cover from Sentinel-1A SAR and SRTM digital elevation model data using Random Forests," *Remote Sensing*, vol. 7, no. 11, pp. 14 876–14 898, 2015.
- [2] T. Idol, B. Haack, and R. Mahabir, "Radar speckle reduction and derived texture measures for land cover/use classification: a case study," *Geocarto International*, vol. 32, no. 1, pp. 18–29, 2017.
- [3] G. M. Foody, M. Pal, D. Rocchini, C. X. Garzon-Lopez, and L. Bastin, "The sensitivity of mapping methods to reference data quality: training supervised image classifications with imperfect reference data," *ISPRS International Journal of Geo-Information*, vol. 5, no. 11, p. 199, 2016.
- [4] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *International Journal of Remote Sensing*, vol. 28, no. 5, pp. 823–870, 2007.
- [5] T. Kavzoglu, "Increasing the accuracy of neural network classification using refined training data," *Environmental Modelling & Software*, vol. 24, no. 7, pp. 850–858, 2009.
- [6] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.
- [7] I. Khosravi, A. Safari, S. Homayouni, and H. McNairn, "Enhanced decision tree ensembles for land-cover mapping from fully polarimetric SAR data," *International Journal of Remote Sensing*, vol. 38, no. 23, pp. 7138–7160, 2017.
- [8] F. Calò, S. Abdikan, T. Görüm, A. Pepe, H. Kiliç, and F. Balık Şanlı, "The space-borne SBAS-DInSAR technique as a supporting tool for sustainable urban policies: The case of Istanbul megacity, Turkey," *Remote Sensing*, vol. 7, no. 12, pp. 16 519–16 536, 2015.
- [9] M. Üstüner, F. B. Şanlı, G. Bilgin, and S. Abdikan, "Land use and cover classification of Sentinel-1A SAR imagery: A case study of Istanbul," in *Signal Processing and Communications Applications Conference (SIU)*, 2017 25th. IEEE, 2017, pp. 1–4.
- [10] G. H. John and P. Langley, "Estimating continuous distributions in bayesian classifiers," in *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann Publishers Inc., 1995, pp. 338–345.
- [11] M. Weinmann, B. Jutzi, S. Hinz, and C. Mallet, "Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 105, pp. 286–304, 2015.
- [12] Q. Du and H. Ren, "Real-time constrained linear discriminant analysis to target detection and classification in hyperspectral imagery," *Pattern Recognition*, vol. 36, no. 1, pp. 1–12, 2003.
- [13] Q. Du and C.-I. Chang, "A linear constrained distance-based discriminant analysis for hyperspectral image classification," *Pattern Recognition*, vol. 34, no. 2, pp. 361–373, 2001.
- [14] V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez, "An assessment of the effectiveness of a random forest classifier for land-cover classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 67, pp. 93–104, 2012.
- [15] B. Waske, J. A. Benediktsson, and J. R. Sveinsson, "Random forest classification of remote sensing data," *Signal and Image Processing for Remote Sensing*, p. 365, 2012.
- [16] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 8, pp. 1778–1790, 2004.
- [17] B. Dixon and N. Candade, "Multispectral landuse classification using neural networks and support vector machines: one or the other, or both?" *International Journal of Remote Sensing*, vol. 29, no. 4, pp. 1185–1206, 2008.