

LANDSAT-8 Uydu Görüntüsü ile Arazi Örtüsü Sınıflandırmasında Makine Öğrenme Algoritmalarının Kullanımı

Mustafa Üstüner^{1,*}, Fusun Balık Şanlı¹

¹Yıldız Teknik Üniversitesi, İnşaat Fakültesi, Harita Mühendisliği Bölümü, 34220, İstanbul.

Özet

Bu çalışmada, Landsat-8 uydu görüntüsünün arazi örtüsü ve kullanımı amacıyla sınıflandırmasında makine öğrenme algoritmalarının kullanılabilirliği ve karşılaştırmalı analizi (sınıflandırma doğruluğu) yapılmıştır. Çalışma kapsamında Landsat-8 uydu görüntüsü yüzey yansımaya (surface reflectance) verileri kullanılmıştır. Çalışma alanı olarak İstanbul seçilmiştir. Çalışma alanında, su, orman, tarım alanları, çıplak alanlar ve kentsel alanlar olmak üzere beş farklı sınıf belirlenmiştir. Çalışma kapsamında sınıflandırma amacıyla, üç farklı makine öğrenme algoritması (Destek Vektör Makineleri (DVM), Rastgele Orman (RO) ve K-En Yakın Komşuluk(K-EYK)) ve en çok benzerlik (EÇB) yöntemi kullanılmıştır. En yüksek sınıflandırma doğruluğu DVM yöntemi (%96.82) ile elde edilirken, en düşük sınıflandırma doğruluğu ise EÇB yöntemi (%92.63) ile elde edilmiştir. Çalışma kapsamında kullanılan makine öğrenme algoritmaları ile %95'in üzerinde sınıflandırma doğruluğu elde edilmiştir.

Anahtar Sözcükler

Landsat-8, Sınıflandırma, Makine Öğrenmesi, Destek Vektör Makineleri, Rastgele Orman

Abstract

In this study, the usability and comparative analysis of machine learning algorithms for the land use/cover classification of Landsat-8 satellite imagery was carried out. Surface reflectance data of Landsat-8 were used in this study. Istanbul was selected as the study region. The study area covers five land cover classes which are water, forestry, agricultural areas, barelands and urban areas. Three different machine learning algorithms (SVM, RF and K-NN) and maximum likelihood classifier were implemented in this study. Results shows that highest classification accuracy (96.82%) of this study was obtained by SVM classification. With three different machine learning algorithms implemented in this study obtained the classification accuracy over 95%.

Keywords

Landsat-8, Classification, Machine Learning, Support Vector Machines, Random Forests

1. Giriş

Uzaktan algılama teknolojisi, son yıllarda artan uydu sayısı ile birlikte yeryüzüne ilişkin hızlı, güvenilir ve güncel bilgiye ulaşmamızda önemli kolaylıklar sağlamaktadır ve doğal kaynakların yönetiminden iklim değişikliğine kadar geniş bir kullanım alanına sahiptir (Khatami vd. 2016). Uydu görüntülerinden bilgi elde etmek amacıyla, uzaktan algılamada kullanılan en yaygın yöntemlerin başında görüntü sınıflandırması gelmektedir. Elde edilecek bilginin güvenilirliği, sınıflandırma doğruluğu ile ilişkilidir ve sınıflandırma doğruluğu; eğitim/test verilerinin dağılımı, kullanılan bantların spektral ve mekânsal çözünürlüğü, sınıflandırma algoritmasının karakteristik özelliği gibi birçok parametreye bağlıdır (Lu ve Weng,2007; Kavzoglu 2009). Günümüzde parametrik ve parametrik olmayan birçok sınıflandırma yöntemi mevcuttur. Özellikle son yıllarda makine öğrenme algoritmaları (piksel ve nesne tabanlı) uydu görüntülerinin sınıflandırılmasında sıklıkla kullanılmaktadır.

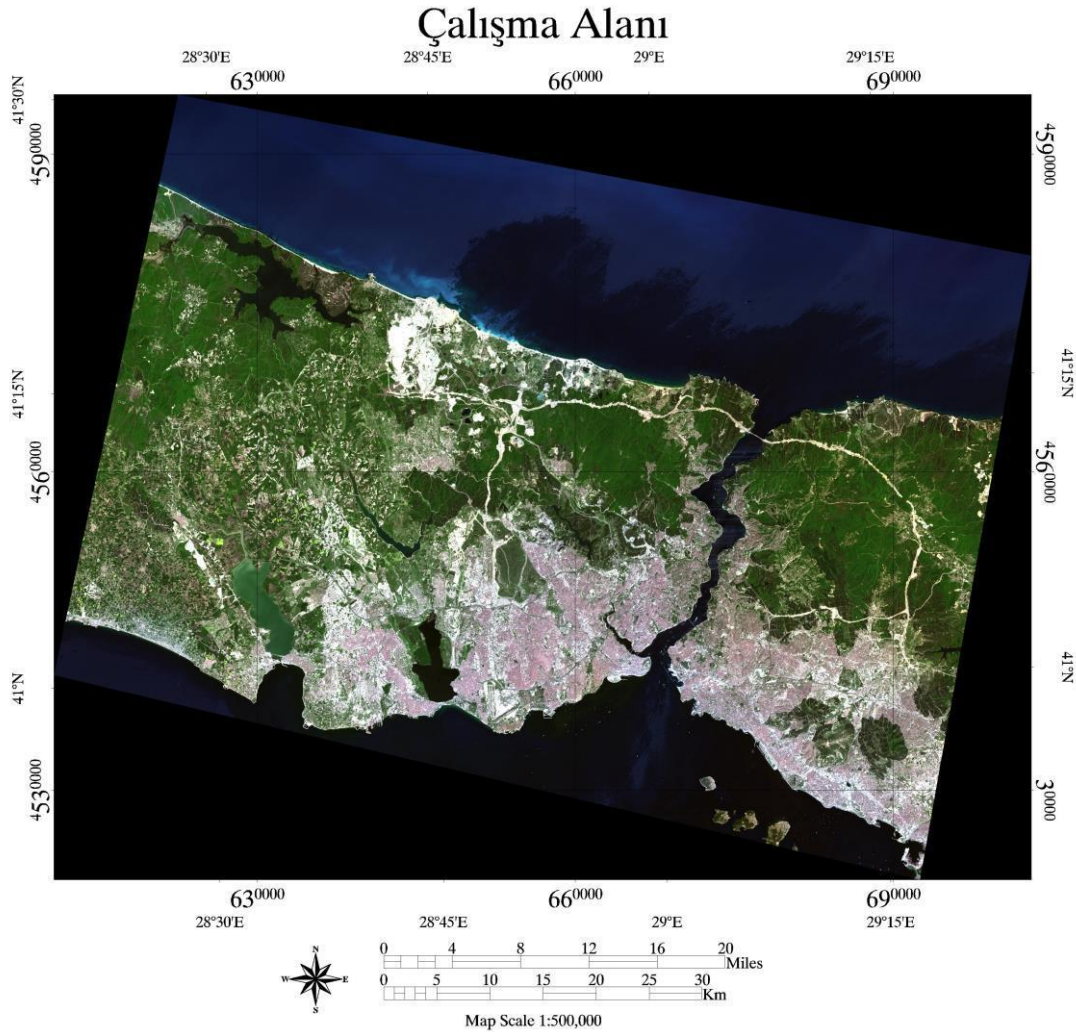
Hızlı nüfus artışı ve kentsel alanlardaki büyüme oranı ile birlikte İstanbul'un arazi örtüsü yapısı giderek değişiklik göstermektedir ve bu değişimin izlenmesi gerekmektedir (Şekil 1). Sinoptik görüş özelliği ve farklı zamanlarda görüntü sağlayabilmesi avantajları ile uzaktan algılama bu amaçla sıklıkla kullanılmaktadır (Calo vd.2015) . Bu çalışma kapsamında son yıllarda uydu görüntülerinin sınıflandırılmasında sıklıkla kullanılan ve başarılı sonuçlar veren üç farklı makine öğrenme algoritması (Destek Vektör Makineleri, Karar Ağaçları ve K-En yakın komşuluk) ve en çok benzerlik sınıflandırma yöntemi kullanılmıştır. Sınıflandırma işlemleri piksel tabanlı olarak gerçekleştirilmiştir. Çalışma alanında su, orman, tarım alanları, çıplak alanlar ve kentsel alanlar olmak üzere beş farklı sınıf belirlenmiştir. Çalışma kapsamında Landsat-8 uydu görüntüsü yüzey yansımaya verileri kullanılmıştır ve 30 metre çözünürlüklü yedi adet spektral bant (kıyı aerosol, mavi, yeşil, kırmızı, yakın kızılötesi, kısa dalga kızılötesi-1, kısa dalga kızılötesi-2) içermektedir.

Bu çalışmada, Landsat-8 uydu görüntüsünün arazi örtüsü ve kullanımı amacıyla sınıflandırmasında makine öğrenme algoritmalarının kullanılabilirliği ve karşılaştırmalı analizi (sınıflandırma doğruluğu) yapılmıştır. İstanbul'a ilişkin arazi örtüsü kullanımı sınıf bilgileri Google Earth'den elde edilen yüksek çözünürlüklü uydu görüntülerinden elde edilmiştir.

* Sorumlu Yazar E-posta: mustuner@yildiz.edu.tr (Mustafa Üstüner)

2. Sınıflandırma ve Çalışma Alanı

Uydu görüntülerinin sınıflandırılması işleminde kullanılacak yöntemin seçimi; eğitim/test verilerinin karakteristik özelliği ve spektral bantların bilgi içeriğine göre değişiklik gösterebilmektedir. Çalışma kapsamında parametrik olmayan yapıları nedeniyle üç farklı makine öğrenme algoritması ve parametrik bir sınıflandırıcı olan en çok benzerlik algoritması kullanılmıştır. Destek vektör makineleri sınıflandırma yöntemi, iki sınıftı birbirinden ayırabilen optimum hiperdüzlemin belirlenmesi esasına dayanır ve az sayıda eğitim verisi ile yüksek doğrulukta sınıflandırma doğruluğu elde edebilmesi nedeniyle optik (multispektral ve hiperspektral) ve radar görüntülerinin sınıflandırılmasında sıklıkla tercih edilmektedir (Huang vd. 2002; Kavzoglu ve Colkesen 2009). DVM sınıflandırılmasında, radyal tabanlı fonksiyon kerneli kullanılmıştır. Topluluk öğrenme algoritmalarında birisi olan rastgele orman algoritması, ağaç tabanlı bir yöntemdir ve karar yüzeyinin belirlenmesinde birden fazla ağaçtan faydalanır (Waske vd. 2012). Ağaç sayısı $N=150$ olarak belirlenmiştir. K-en yakın komşuluk yönteminde ise piksele en yakın k sayıdaki piksel belirlenir ve ilgili piksel, k sayıdaki pikselden baskın olan sınıfa atanır (Franco vd.2001).



Şekil 1: Çalışma Alanı

3. Doğruluk Analizi ve Sonuçlar

Sınıflandırılmış görüntülerin doğruluk analizi, toplam doğruluk ve kappa katsayısına göre yapılmıştır. Çalışmada en yüksek sınıflandırma doğruluğu DVM yöntemi (%96.82) ile elde edilirken, en düşük sınıflandırma doğruluğu ise EÇB yöntemi (%92.63) ile elde edilmiştir. Çalışma kapsamında kullanılan makine öğrenme algoritmaları ile %95'in üzerinde sınıflandırma doğruluğu elde edilmiştir (Tablo 1). Bu çalışma ile makine öğrenme algoritmalarının Landsat-8 uydu görüntülerinin arazi örtüsü/kullanımı sınıflandırılmasında etkin bir şekilde kullanılabileceği ve yüksek sınıflandırma doğruluğu elde edilebileceği görülmüştür. Ayrıca parametrik olmayan yapıları ile makine öğrenme algoritmaları, parametrik yapıda bir sınıflandırıcı olan en çok benzerlik algoritmasına kıyasla yüksek sınıflandırma doğruluğu elde etmişlerdir.

Tablo 1: Sınıflandırma Doğrulukları

Sınıflandırma	Açıklama	Doğruluk	
		Toplam (%)	Kappa
1	Destek Vektör Makineleri	96.82	0.9598
2	Rastgele Orman	95.84	0.9474
3	K-En Yakın Komşuluk	95.51	0.9434
4	En çok benzerlik	92.63	0.9071

Teşekkür

Landsat 8 uydu görüntüsünün ücretsiz temini için Amerika Birleşik Devletleri Jeoloji Araştırmaları Kurumuna (USGS) teşekkür ederiz.

Kaynaklar

Dergi İçinde Makale:

Calò, F., Abdikan, S., Görüm, T., Pepe, A., Kiliç, H., and Balık Şanlı, F. (2015) *The Space-Borne SBAS-DInSAR Technique as a Supporting Tool for Sustainable Urban Policies: The Case of Istanbul Megacity, Turkey*, Remote Sensing, 7, (12), pp. 15842

Franco-Lopez, H., Ek, A.R. ve Bauer, M.E (2001) “*Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method*”, Remote Sensing of Env., 77(3):251-274,

Huang, C., Davis, L. S., ve Townshend, J. R. G. (2002). *An assessment of support vector machines for land cover classification*. International Journal of Remote Sensing, 23(4), 725-749.

Kavzoglu, T. (2009) *Increasing the accuracy of neural network classification using refined training data*, Environmental Modelling & Software, 24, (7), pp. 850-858

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. doi: <http://dx.doi.org/10.1016/j.jag.2009.06.002>

Khatami, R. Mountrakis, G. ve Stehman, S.V.(2016): *A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research*, Remote Sensing of Environment, 177, pp. 89-100

Lu, D. ve Weng, Q. (2007) *A survey of image classification methods and techniques for improving classification performance*, International Journal of Remote Sensing, 28, (5), pp. 823-870

Waske, B., Benediktsson, J., ve Sveinsson, J. (2012) “Random Forest Classification of Remote Sensing Data”, Signal and Image Processing for Remote Sensing, Second Edition, CRC Press, pp. 365-374