



Derin Öğrenme Yaklaşımlarıyla Çevresel İzlemeye Yönelik Çok Sınıflı Sahne Sınıflandırma

Şaziye Özge Atik^{1*}

^{1*} Gebze Teknik Üniversitesi, Mühendislik Fakültesi, Harita Mühendisliği Bölümü, İstanbul, Türkiye, (ORCID: 0000-0003-2876-040X), soatik@gtu.edu.tr

(İlk Geliş Tarihi 14 Ocak 2022 ve Kabul Tarihi 20 Kasım 2022)

(DOI: 10.31590/ejosat.1057643)

ATIF/REFERENCE: Atik, S. O. (2022). Derin Öğrenme Yaklaşımlarıyla Çevresel İzlemeye Yönelik Çok Sınıflı Sahne Sınıflandırma. *Avrupa Bilim ve Teknoloji Dergisi*, (41), 307-314.

Öz

Bu çalışmada uzaktan algılama sahne sınıflandırma veri setlerinden UC Merced ve WHU-RS19 üzerinde Alexnet, Densenet, ResNet, VGG derin ağlarının farklı modellerinin performans analizleri yapılmıştır. Kullanılan modellerin test aşamasında derin öğrenme modellerinin sınıflandırma sonuçlarına ait değerlendirme ölçütleri hesaplanmıştır. Kullanılan ESA modellerinin veri setlerindeki performansları genel doğruluk ölçütünde değerlendirilmiştir. DenseNet201 modelinin, UC Merced ve WHU-RS19 veri setlerinin her ikisinde de yapılan testlerde daha yüksek performanslı sonuçlara sahip olduğu gözlemlenmiştir. ESA modelleri uygulanarak elde edilen sonuçlar, literatürde bu veri setlerini kullanan diğer çalışmaların sonuçlarıyla karşılaştırılmıştır. UC Merced veri setindeki uygulamada %98.81 genel doğruluk ile bu çalışmada kullanılan DenseNet201 modelinin literatürde kıyası yapılan diğer çalışmalardan daha yüksek performansa sahip olduğu gözlemlenmiştir. Ayrıca, her iki veri setinde benzer olan arazi kullanım sınıfları belirlenmiş ve en iyi performans gösteren algoritmadaki sınıflandırma doğruluğu sonuçları yorumlanmıştır. İki veri setinde de bulunan benzer sınıfların yapılan testlerde sınıflandırılması kesinlik, duyarlılık ve F_1 skoru ölçütleri kullanılarak değerlendirilmiştir. Son olarak, iki veri setindeki ortak sınıflar için WHU-RS 19 veri seti kullanılarak eğitilen modellerin testleri UC Merced veri setinde ve UC Merced veri setinde eğitilen modellerin testleri WHU-RS 19 veri setinde yapılarak model transfer edilebilirliği araştırılmıştır.

Anahtar Kelimeler: Derin öğrenme, Sahne sınıflandırma, Çevresel izleme, Evrimsel Sinir Ağları

Multi-Class Scene Classification for Environmental Monitoring with Deep Learning Approaches

Abstract

In this study, performance analyzes of different models of Alexnet, Densenet, ResNet, VGG deep networks were performed on UC Merced and WHU-RS19, which are remote-sensing scene classification datasets. In the testing phase of the models used, the evaluation criteria of the classification results of the deep learning models were calculated. The performances of the CNN models used in the data sets were evaluated in the general accuracy criterion. It has been observed that the DenseNet201 model has higher performance results in tests performed on both UC Merced and WHU-RS19 datasets. The results obtained by applying ESA models were compared with the results of other studies using these datasets in the literature. It has been observed that the DenseNet201 model used in this study has a higher performance than other studies compared to the literature, with an overall accuracy of 98.81% in the application in the UC Merced dataset. In addition, land use classes that are similar in both data sets were determined and the classification accuracy results in the best-performing algorithm were interpreted. Classification of similar classes in both datasets in tests was evaluated using precision, sensitivity, and F1 score criteria. Finally, for the common classes in the two datasets, the models trained using the WHU-RS 19 dataset were tested on the UC Merced dataset and the models trained in the UC Merced dataset were tested on the WHU-RS 19 dataset, and the model transferability was investigated.

Keywords: Deep learning, Scene classification, Environmental monitoring, Convolutional Neural Networks

* Sorumlu Yazar: soatik@gtu.edu.tr

1. Giriş

Yer gözlem uygulamaları ve çevresel izleme, özellikle son on yılda tüm dünyada yaygın bir şekilde yayılmıştır. Bu uygulamalar, dünyadaki sınırlı doğal ve diğer kaynakların sınırlı olması sebebiyle her yıl daha önemli hale gelmektedir. Hava aracı kaynaklı (Atik vd., 2018) ve uydu kaynaklı (Dönmez ve Tunç, 2016) sensör görüntüleri gibi uzaktan algılanan verileri kullanmak, geo-konumsal, yer gözlemleri ve daha fazla uygulama için ana veri kaynaklarıdır. Derin öğrenme teknikleri, bilgisayarlı görü ile birçok endüstride her yıl daha popüler hale gelmektedir. Derin öğrenme yöntemlerinin kullanıldığı uygulamalar daha az zaman alması ve daha düşük maliyete sahip olması özellikleri ile ön plana çıkmaktadır. Derin öğrenme yaklaşımları, görüntü işleme ve bilgisayarla görmede birçok aşama için hızlı hesaplamalar ve analizler sağlar. Son yıllarda, uzaktan algılama teknolojileri kullanılarak hava görüntülerinin sınıflandırması (Weng vd., 2017; Bi vd., 2019), semantik görüntü segmentasyonu (Diakogiannis vd., 2020; Atik ve Ipbuker, 2021), nesne tespiti (Napierkowska vd., 2018) gibi çok sayıda çalışma yapılmaktadır.

Yapay zeka kullanımının yüksek işlem performansına sahip donanım ihtiyacı gelişen bilgisayar sistemleri sayesinde giderilebilmektedir. Yapay zeka teknolojilerinin kullanımı bir çok alanda olduğu gibi, uzaktan algılama uygulamaları için otomatik ve yarı otomatik yaklaşımlara olanak sağlar. Pek çok sınıflandırma görevi için pekiştirmeli öğrenme modelleri kullanılmaktadır (Donmez ve Ipbuker, 2018). Örneğin, kentsel büyümenin çevresel etkilerinin değerlendirilmesi ve arazi kullanımının tespiti, uzaktan algılama verileri kullanılarak mümkündür. Uzaktan algılama ile otomatik sınıflandırma ve etiketleme yöntemlerinin geliştirilmesi ve iyileştirilmesi için birçok veri seti bulunmaktadır. Semantik bilgilerin çıkarılması ve bölgesel izlemenin gerçekleştirilebilmesi için bu veri setleri kullanılarak oldukça çeşitli uygulamalar gerçekleştirilebilmektedir. Bu veri setlerine örnek olarak; WHU-RS19 (Sheng vd., 2012), UC Merced (UCM) (Yang ve Newsam, 2010), AID (Xia vd., 2017), NPWU-RFSISC45 (Cheng vd., 2017), Patternet (Zhou vd., 2018), Category Ground Scene (Brown ve Süsstrunk, 2011), Brazilian Coffee Scene (Penatti vd., 2015) ve BigEarthNet (Sumbul vd., 2019) verilebilir. Literatürde birçok çalışmada, hava fotoğrafları ve uydu görüntüleri kullanılarak derin öğrenme yöntemleri ile bilgi çıkarımı gerçekleştirilmiştir. Petrovska vd. (2020) UC Merced ve WHU-RS veri setlerinde SVM algoritması ile oluşturulan konsantr öznelikler üzerinden ESA-bazlı sınıflandırma yapmışlardır. Özyurt vd. (2020) geliştirdikleri optimize edilmiş Wavelet Entropi algoritması ile evrimsel sinir ağları (ESA) özneliklerinin boyutunu azaltarak UC Merced veri setini sınıflandırmışlardır. Muhammed vd. (2018) yaptıkları çalışmada önceden eğitilmiş VGGNet mimarilerini kullanarak UCM ve WHU-RS veri setlerinde 4-katmanlı SVM sınıflandırıcısıyla sınıflandırma yapmışlardır. Stivaktakis vd. (2019) UC Merced veri seti üzerinde veri artırımı yaparak çalışmada önerilen yöntem ile ESA kullanarak güncel yöntemlerden F1 skor için % 6 daha iyi sonuç elde etmişlerdir. Marmanis vd. (2015) çalışmalarında UC Merced veri setinde özellik birleştirme algoritması uygulayarak derin öğrenme ağlarını kullanmışlardır.

Literatürdeki uygulamalarda arazi kullanım sınıflarına sahip veri setleri kullanılarak farklı görevler gerçekleştirilebilmektedir. Görüntülerin sahne sınıflandırması Şekil 1'de gösterildiği gibi üç ana grup altında toplanabilir. İkili sınıflandırmanın iki sınıfı vardır ve görüntüler bunlardan yalnızca birine atanabilir. Çok sınıflı

sınıflandırmada birçok sınıf vardır ve görüntüler etiketlerden biri olarak etiketlenir. Bunun yanısıra çoklu etiket sınıflandırması için görüntüler birden fazla etikete sahip olabilir ve sınıf numarası ikiden fazladır. Aynı zamanda, hava fotoğrafları ve uydu görüntüleri içeren veri setleri, çok sınıflı sınıflandırma veya çok etiketli sınıflandırma için kullanılabilir. Sonuç olarak, sınıflandırma türü, veri kümesinin kapsamına ve kategori türüne bağlıdır.

Tekil Sınıflandırma	Çoklu-Sınıf Sınıflandırması	Çoklu-Etiket Sınıflandırması
Etiket 1 ✓	Etiket 1	Etiket 1 ✓
Etiket 2	Etiket 2 ✓	Etiket 2
	Etiket 3	Etiket 3 ✓
	Etiket 4	Etiket 4 ✓

	Etiket	Etiket

Şekil 1. Sahne sınıflandırması çeşitleri (Figure 1. Types of scene classification)

Stivaktakis vd. (2019) UC Merced veri seti üzerinde veri artırımı yaparak çalışmada önerilen yöntem ile ESA kullanarak güncel yöntemlerden F1 skor için % 6 daha iyi sonuç elde etmişlerdir. Marmanis vd. (2015) çalışmalarında UC Merced veri setinde özellik birleştirme algoritması uygulayarak derin öğrenme ağlarını kullanmışlardır. Hu vd. (2022) WHU-RS veri setini kullanarak MINet olarak adlandırdıkları ESA tabanlı algoritma ile sahne sınıflandırması yapmışlardır. UC Merced, WHU-RS19, RSSCN7 ve AID veri setlerini kullanarak ESA modelleri yardımıyla sahn sınıflandırması yapan Anwer vd. TEX-Net modeli ile daha yüksek genel doğruluk değerlerine ulaşmışlardır. Şen ve Keleş (2020) yaptıkları çalışmada, farklı ESA mimarilerini kullanarak NPWU-RESISC45 veri setinde eğittikleri modelleri başka beş yaygın kullanılan veri setlerini sınıflandırmadaki başarımını ampirik olarak gözlemlemişlerdir.

Bu çalışmada, farklı veri kümelerinde aynı sınıflar için 14 derin öğrenme modeli ile veri kümelerindeki arazi kullanım sınıflarının sahne sınıflandırması amacıyla otomatik etiketlendirilmesi amaçlanmıştır. On dört adet ESA modeli ile çok sınıflı sınıflandırma görevinde UCM ve WHU-RS19 veri setleri kullanılmıştır (Şekil 2). Elde edilen sonuçlar literatürdeki benzer çalışmalarla karşılaştırıldığında bu çalışmadaki UC Merced veri setinde yapılan uygulamada daha iyi doğruluk performansına ulaşılmıştır. Bu çalışmada kullanılan veri setleri kumsal, orman, otopark, akarsu ve liman şeklinde benzer sınıflara sahiptir (Şekil 3). Uygulama kısmında, ESA modellerinin sınıflandırma doğruluğuna ait performansları ve eğitim-test süreleri karşılaştırılmıştır. En iyi sınıflandırma sonucuna sahip algoritmanın sınıflandırması F1 skoru, kesinlik ve duyarlılık açısından bu iki veri seti için karşılaştırmalı olarak gösterilmiştir. Aynı zamanda iki veri setindeki ortak sınıflar için, WHU-RS 19 veri setinin görüntüleri kullanılarak eğitilen ESA modeller, UC Merced veri setinin görüntülerinde test edilmiştir. Aynı zamanda UC Merced veri setinde eğitilen modeller, WHU-RS 19 veri setinde test edilmiştir. Bu sayede model transfer edilebilirliği bir çok sayıda ESA modeli ile karşılaştırmalı olarak araştırılmıştır.

2. Materyal ve Metot

2.1. Veri Setleri

Çalışmada, farklı ESA ile çok sınıflı sınıflandırma için UCM ve WHU-RS19 veri setleri kullanılmıştır. UCM veri setinde 21 sınıf ve WHU-RS19 veri setinde 19 sınıf bulunmaktadır. UCM 256 x 256 piksel boyutlarından oluşan RGB görüntülerden oluşmaktadır. Görüntülerde mekânsal çözünürlük yaklaşık 0.3m (1 ft) şeklindedir. Her sınıfa ait 100 görüntü bulunmaktadır. WHU-RS veri seti ise, 600 x 600 boyutlarında her sınıftan 50 adet olmak üzere toplamda 1005 görüntüden oluşmaktadır. WHU-RS 19 veri seti, mekânsal çözünürlükleri 0.5 m'ye kadar değişim gösteren farklı mekânsal çözünürlükteki görüntülerden oluşmaktadır. Bu veri setlerinde kumsal, orman, otopark ve nehir akarsu sınıfları her ikisi için de benzer sınıflardır. Yerleşim sınıfı, UC Merced veri setinde üç sınıfa ayrılmıştır. Ancak yerleşim sınıfı, WHU-RS19 veri setinde tek bir sınıfta toplanmıştır. Genel olarak bu iki veri seti birçok yapay yapı ve doğal sınıf içermektedir. Bu çalışmada, UC Merced veri seti %20 test, %80 eğitim ve WHU-RS19 veri seti ise %40 test, %60 eğitim şeklinde rastgele olarak bölünmüştür. Bu şekilde bölümlendirilmesinin sebebi, çalışmanın sonuçlarını literatürdeki diğer yapılan bazı çalışmaların sonuçlarıyla karşılaştırılmak istenmesidir. Karşılaştırması yapılan çalışmalarda da (Tablo 7 ve Tablo 8) bu veri setleri benzer şekilde bölümlendirilmiştir.



Şekil 2. UC-Merced (Yang ve Newsam,2010) veri seti (üst) ve WHU-RS veri seti (alt) örnekleri (Figure 2. UC-Merced (Yang and Newsam,2010) dataset (top) and WHU-RS data set (bottom) samples)

UC Merced Veri Seti		WHU-RS Veri Seti	
Tarım	Liman	Havaalanı	Dağ
Uçak	Kavşak	Kumsal	Park
Beyzbol sahası	Orta yoğunlukta konut	Köprü	Otopark
Kumsal	Karavan parkı	Ticari bölge	Gölet
Binalar	Üst geçit	Çöl	Liman
Çalı	Otopark	Çiftlik	Tren istasyonu
Yoğun konut	Akarsu	Futbol sahası	Yerleşim
Orman	Koşu yolu	Orman	Akarsu
Otoban	Seyrek konut	Endüstriyel alan	Viyadük
Golf sahası	Depolama tankı	Çayır	
Tenis kortu			

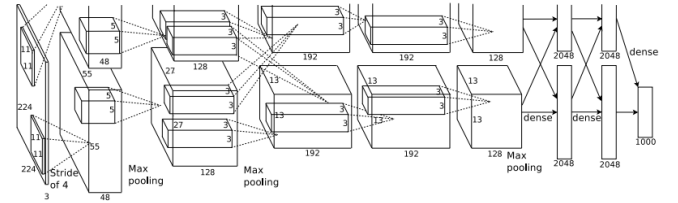
UC Merced WHU-RS Ortak Sınıflar

Şekil 3. Uygulamada kullanılan veri setlerinin sınıf gösterimleri (Figure 3. Class representations of datasets used in the study)

2.2. ESA Modelleri

Uygulamalarda, performansların gerçekleştirildiği sistem olarak Google Colaboratory (Colab) kullanılmıştır. Bu bulut hizmeti, tek bir 12 GB NVIDIA Tesla K80 GPU'lu Jupyter Notebook arayüzünü temel almaktadır. Uygulamalarda kullanılan ESA modelleri; Alexnet (Krizhevsky vd., 2012), ResNet-18 (He vd., 2016), ResNet-34 (He vd., 2016), VGG-19 (Simonyan ve Zissersman, 2014), ResNet-152 (He vd., 2016), ResNet-50 (He vd., 2016), ResNet101 (He vd., 2016), DenseNet-169 (Huang vd., 2017), VGG-11 (Simonyan ve Zissersman, 2014), VGG-13 (Simonyan ve Zissersman, 2014), DenseNet-161 (Huang vd., 2017), DenseNet-201 (Huang vd., 2017), VGG-16 (Simonyan ve Zissersman, 2014), DenseNet-121 (Huang vd., 2017) şeklindedir. Uygulamalarda kullanılan derin öğrenme mimarileri kısaca aşağıda açıklanmıştır.

AlexNet. Krizhevsky vd. (2012) ImageNet ILSVRC-2010 yarışmasında 1000 farklı sınıf için 1.2 milyon yüksek çözünürlüklü görüntü eğiten bir ESA mimarisidir.



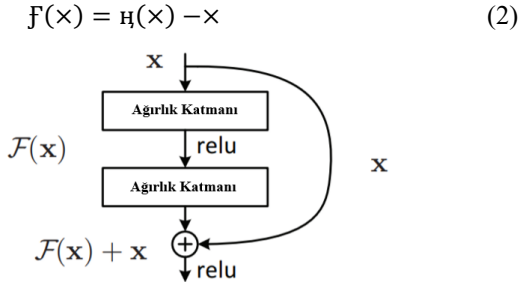
Şekil 4. AlexNet mimarisine ait bir görselleştirme (Figure 4. A visualization of the AlexNet architecture)(Krizhevsky vd., 2012)

Mimaride (Şekil 4) yerel normalizasyon şeması genelleştirmeye yardımcı olmaktadır (eşitlik 1). Burada $a_{(x,y)}^i$ nöronun aktivitesidir ve i çekirdeği x,y konumu ifade etmektedir. ReLU (Rektifiye Edilmiş Doğrusal Birim) içermekte ve $b_{(x,y)}^i$ uygulanması ile bu normalleştirilmiş bir aktivitedir. İfadedeki $k=2$, $n=5$, $\alpha=10$ ve $\beta=0.75$ şeklindedir.

$$b_{x,y}^i = a_{x,y}^i / \left(k + \alpha \sum_{\min(N-1, i+\frac{n}{2})}^{\min(N-1, i-\frac{n}{2})} (\alpha_{x,y}^j)^2 \right)^\beta \quad (1)$$

Mimarinin etkinliği, CIFAR-10 veri setinde de doğrulanmıştır.

ResNet. Şekil 5' te artık öğrenme şeması gösterilmektedir. $H(\times)i$ 'nin formülü eşitlik 2' de gösterilmiştir.

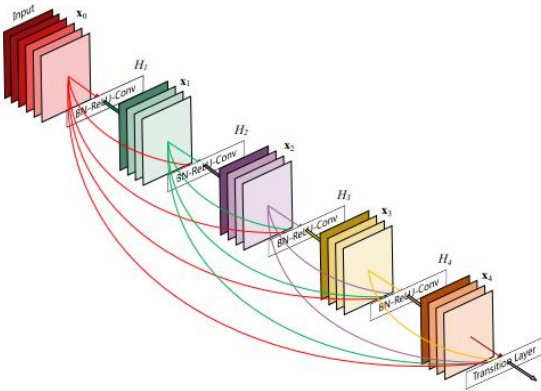


Şekil 5. Artık öğrenmenin temel gösterimi (Figure 5. Basic notation of residual learning)(He vd., 2016).

$F(x)+x$ formülünde ileri beslemeli sinir ağının kısayol bağlantıları vardır. Eşitlikte yapı bloğu tanımlanır ve x ile F boyutunun eşit olması gerekir. x ve y , katmanların giriş ve çıkış vektörlerini ifade eder. Algoritmada çoklu evrişim katmanlarını mevcuttur (He vd., 2016). İçerdiği katman sayısı ile bağlantılı olarak ResNet50 ve ResNet101 gibi isimler almaktadır.

VGG. ILSVRC-2012 veri setinde Simonyan ve Zisserman (2014) tarafından eğitilmiş ConvNet mimarisi 1000 sınıf içermektedir. Yarışmaya ait veri seti 1,3 M görüntü eğitimi, doğrulama için 50 K görüntü ve test için 100 K görüntü olarak kullanılır (Aksoy vd., 2020). Bu mimari de derinlik artırılarak çok küçük (3×3) evrişim filtreleri ile mimarinin kullanımına katkı sağlanmıştır. Deneysel olarak 11, 16 ve 19 ağırlık katmanları mimaride VGG 11, 16 ve 19 modelleri olarak kullanılabilir. Evrişimli katman parametreleri, 16 ağırlık katmanı için örnek olarak conv3-64 gibi conv (alıcı alan boyutu)-(kanal sayısı) şeklindedir.

DenseNet. Huang vd. (2017) Yoğun Evrişimsel Ağ (DenseNet) olarak bir ESA oluşturmuşlardır. Her katmanı diğer katmanlara ileri beslemeli tip ile bağlar (Şekil 6). Ağın kullandığı dört trend performans ölçütünde değerlendirilmiştir. Bunlar; CIFAR-10, CIFAR-100, SVHN ve ImageNet şeklindedir.



Şekil 6. Büyüme oranı $k = 4$ olan 5 katmanlı yoğun bloğun bir gösterimi (Figure 6. A representation of the 5-layer dense block with growth rate $k = 4$)

DenseNet mimarisi, kaybolan gradyan problemini hafifletmek, özellik yayılımını güçlendirmek, özelliklerin yeniden kullanımını kolaylaştırmak ve parametre sayısını önemli ölçüde azaltmak gibi birçok avantaja sahip olduğu için oldukça yaygın olarak kullanılmaktadır (Qi vd., 2020). DenseNet içerdiği katman sayısına bağlı olarak 121, 169 ve 201 isimlerini almaktadır.

Uygulamada sınıflandırma için kullanılan derin öğrenme ağlarının performanslarını değerlendirmek için, genel doğruluk

(eşitlik 3) ölçütü kullanılmıştır. En iyi performansı gösteren modelin sınıf bazında sonuçları ise kesinlik, duyarlılık, F1 skoru (eşitlik 4-6) ölçütleri ile gösterilmiş ve benzer sınıflar arasında bir karşılaştırma yapılmıştır. Eşitliklerde TP: doğru pozitif, FN: yanlış negatif, FP: yanlış pozitif ve N: toplam sınıflandırma sayısı şeklindedir.

$$Genel\ Doğruluk = \left(\frac{TN+TP}{N} \right) \quad (3)$$

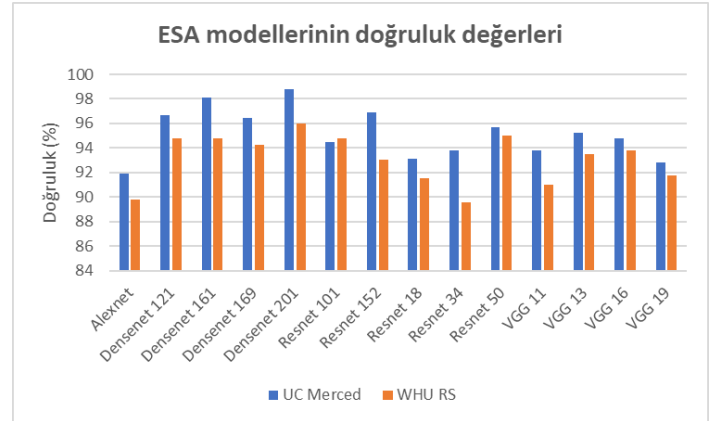
$$Kesinlik = \frac{TP}{TP+FP} \quad (4)$$

$$Duyarlılık = \frac{TP}{TP+FN} \quad (5)$$

$$F1\ skoru = 2 * \frac{Kesinlik * Duyarlılık}{Kesinlik + Duyarlılık} \quad (6)$$

2. Araştırma Sonuçları ve Tartışma

Bu çalışmada, iki farklı veri seti kullanılarak farklı derin öğrenme modelleri ile çoklu-sınıflı sahne sınıflandırması yapılmıştır. Derin öğrenme algoritmaları önceden eğitilmiş modeller ile uygulanmıştır. Test sonunda değerlendirme metrikleri hesaplanır ve birbirleriyle karşılaştırılır. İki veri kümesi için en iyi ve düşük performansı gösteren modellerde benzerlik gözlenmiştir (Şekil 7). Tablo 1'de ESA modellerinin genel doğrulukları gösterilmektedir. Tablo üzerinde değerler % şeklinde verilmiştir.



Şekil 7. ESA modellerinin iki veri setindeki uygulamalarda test sonuçları (Test results on applications of ESA models on two datasets)

Modellerin eğitim aşamasında önceden eğitilmiş model ağırlıkları başlangıç ağırlığı olarak kullanılmıştır. ESA modellerinde epok sayısı deneysel olarak belirlenerek yedi olarak seçilmiştir ve tüm ESA modelleri için aynı epok sayısı (7) kullanılmıştır. Her model için hata oranı izlenmiştir. Uygulamalarda ESA modellerinin eğitim ve test aşamalarındaki toplam süreler Tablo 1 ve 2'de gösterilmiştir. Buna göre her iki veri setinde en uzun süre Densenet 161 modeline ve en kısa süre de Alexnet modeline aittir (Şekil 8). UC Merced veri setinin boyutunun daha fazla olmasının öğrenme ve test sürelerini bazı mimariler için daha uzun olmasına neden olduğu gözlenmiştir.

UCM veri seti ile yapılan deneyde, DenseNet201, yapılan testlerde en iyi performans modeli olarak %98.81 doğruluk elde edilmiştir. AlexNet modeli kullanarak ise %91.90 ile en düşük test doğruluğu elde edilmiştir.

edilmiştir. İki veri setindeki benzer sınıflar için performanslarının veri setlerine göre değiştiği gözlenmiştir.

Tablo 1. UC Merced veri setinde ESA modellerinin doğruluk ve süre değerleri (Table 1. Accuracy and duration values of ESA models in UC Merced dataset)

Model	UC -Merced (20)	Süre (sn)
Alexnet	91.90	79
VGG 19	92.86	94
Resnet 18	93.10	78
Resnet 34	93.80	90
VGG 11	93.81	90
Resnet 101	94.52	183
VGG 16	94.76	101
VGG 13	95.24	94
Resnet 50	95.71	117
Densenet 169	96.43	164
Densenet 121	96.67	135
Resnet 152	96.90	255
Densenet 161	98.10	270
Densenet 201	98.81	215

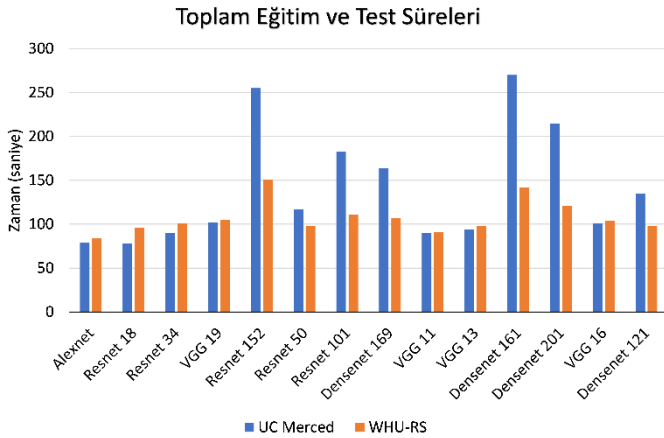
Tablo 2. WHU RS veri setinde ESA modellerinin doğruluk ve süre değerleri (Table 2. Accuracy and duration values of ESA models in WHU-RS dataset)

Model	WHU-RS (40)	Süre(sn)
Resnet 34	89.55	101
Alexnet	89.80	84
VGG 11	91.04	91
Resnet 18	91.54	96
VGG 19	91.79	105
Resnet 152	93.04	141
VGG 13	93.53	98
VGG 16	93.78	104
Densenet 169	94.28	107
Densenet 121	94.78	98
Densenet 161	94.78	142
Resnet 101	94.78	111
Resnet 50	95.02	98
Densenet 201	96.02	121

Şekil 8. ESA modellerinin iki veri setindeki eğitim ve test süreleri (Figure 8. Training and testing times of ESA models on two datasets)

Tablo 3. DenseNet 201 modelinin UC Merced veri setindeki sonuçları (Table 3. Results of the DenseNet 201 model in the UC Merced dataset)

UC Merced Sınıfları	Kesinlik	Duyarlılık	F1 skoru
Tarım	1	0.96	0.98
Uçak	1	1	1
Beyzbol sahası	1	1	1
Kumsal	1	1	1
Binolar	1	1	1
Çalı	1	1	1
Yoğun konut	0.92	0.75	0.83
Orman	0.88	1	0.93
Otoban	1	1	1
Golf sahası	1	1	1
Liman	1	1	1
Kavşak	1	1	1
Orta yoğunluk. konut	0.89	0.92	0.91
Karavan parkı	0.94	1	0.97
Üst geçit	1	1	1
Otopark	1	1	1
Akarsu	1	1	1
Koşu yolu	1	1	1
Seyrek konut	0.94	0.94	0.94
Depolama tankı	1	1	1
Tenis kortu	1	1	1



WHU-RS19 veri setindeki uygulamada, DenseNet201 modeli ile %96.02 ile en yüksek test doğruluğu elde edilmiştir. Her iki veri setinde yapılan çoklu-sınıflı sahne sınıflandırması uygulamasında DenseNet 201 modeli diğer ESA modellerinden daha yüksek performansa sahip olduğu görülmüştür. UC Merced veri setindeki uygulamada olduğu gibi, WHU-RS 19 veri setindeki uygulamada da en düşük performans değeri %89.80 olarak AlexNet modeli ile elde edilmiştir. Bu durum AlexNet mimarisinin bu alandaki sınıflandırmalar için ilk geliştirilen modellerden biri olduğu (2012) göz önünde bulundurulduğunda beklenmektedir. Daha sonra bu mimarinin katmanları daha derin katmanlı öğrenme yapacak şekilde geliştirilmeye devam

Kumsal ve akarsu sınıfları en iyi performans gösteren DenseNet201 modeli ile yapılan uygulamada tüm test görüntülerini doğru sınıflandırmıştır. Liman ve otopark sınıfları için, aynı model UCM veri setinde daha yüksek doğrulukla sınıflandırılırken, orman sınıfı WHU-RS19 veri setinde aynı model ile kesinlik, duyarlılık ve F1 skoru ölçütlerinde daha yüksek doğrulukla sınıflandırılmıştır (Tablo 3-4). Bu sonuçlar, kullanılan veri setlerinin görüntülerindeki özelliklerden kaynaklanabilmektedir. WHU-RS19 veri setinin görüntülerinde farklı ışıklandırma özellikleri ve çeşitli mekânsal çözünürlükler söz konusudur. Aynı zamanda dünyanın çeşitli yerlerinden toplanmış olmasının da sınıflandırma performansına etkisi söz konusudur.

Tablo 4. DenseNet 201 modelinin WHU-RS19 veri setindeki sonuçları (ortak sınıflar **kalin** yazıyla gösterilmiştir) (Table 4. Results of the DenseNet 201 model in the WHU-RS19 dataset (common classes shown in bold))

WHU-RS Sınıfları	Kesinlik	Duyarlılık	F1 skoru
Hava alanı	0.96	0.96	0.96
Kumsal	1	1	1
Köprü	1	1	1
Ticari bölge	0.81	0.85	0.83
Çöl	1	1	1
Çiftlik	0.95	1	0.98
Orman	0.96	1	0.98
Endüstriyel alan	0.91	0.84	0.87
Çayır	1	0.96	0.98
Dağ	1	0.95	0.97
Park	0.85	0.94	0.89
Otopark	0.94	1	0.97
Gölet	1	1	1
Liman	1	0.94	0.97
Yerleşim	0.95	0.9	0.93
Akarsu	1	1	1
Viyadük	0.95	1	0.97
Futbol sahası	1	1	1
Tren istasyonu	1	0.94	0.97

Uygulamada ek olarak, ESA modellerinin transfer edilebilirliği de araştırılmıştır. Eğitilen bu modeller tamamen farklı sernsöre ait iki veri setinden birinin görüntüleri modelleri ortak sınıfları sınıflandırmada eğitmek için, diğer veri setinin görüntülerini de test etmek için kullanılmıştır. Böylece çok sınıflı sahne sınıflandırması alan kaydırma (domain-shift) uygulaması gerçekleştirilmiştir. Bu uygulamanın her modele ait sonuçları Tablo 5 ve 6'da beş ortak sınıf için gösterilmiştir. Bu amaçla gerçekleştirilen ilk uygulamada; WHU-RS 19 veri setinin görüntüleri kullanılarak eğitilen modeller UC Merced veri setinin görüntüleri üzerinde ortak sınıflar için test edilmiştir. En yüksek ortalama değer % 76,6 olarak Resnet 18 modeli ile ve en düşük

ortalama değer % 60.2 ile Resnet 101 modeli ile elde edilmiştir (Tablo 5). İkinci aşamada UC merced veri setinin görüntüleri kullanılarak eğitilen ESA modelleri, WHU-RS 19 veri setinde test edilmiştir. Ortak beş sınıf için gerçekleştirilen bu uygulamada, en yüksek doğruluk %79.3 ile DenseNet 161 modeline aittir. En düşük doğruluk değeri ise % 69.5 olarak VGG 11 modeli ile elde edilmiştir (Tablo 6).

Tablo 5. WHU-RS19 Eğitim & UC Merced Test Uygulaması. Değerler % olarak verilmiştir. (Table 5. WHU-RS19 Education & UC Merced Test. Values are given in %.)

Model	Akarsu	Orman	Otopark	Liman	Kumsal	Ort.
Alexnet	23	64	88	80	75	66
Resnet 18	26	91	92	79	95	76,6
Resnet 34	7	96	73	58	91	65
VGG 19	28	93	85	68	76	70
Resnet 152	37	78	78	46	97	67,2
Resnet 50	39	62	87	32	94	62,8
Resnet 101	22	57	98	51	73	60,2
Densenet 169	34	69	100	79	90	74,4
VGG 11	27	97	98	67	76	73
VGG 13	32	99	95	48	87	72,2
Densenet 161	28	85	96	76	94	75,8
Densenet 201	17	87	94	69	84	70,2
VGG 16	27	94	89	51	89	70
Densenet 121	34	89	93	76	85	75,4

Tablo 6. UC Merced Eğitim & WHU-RS19 Test Uygulaması. Değerler % olarak verilmiştir. (Table 6. UC Merced Education & WHU-RS19 Test. Values are given in %.)

Model	Akarsu	Orman	Otopark	Liman	Kumsal	Ort.
Alexnet	67	62	82	51	100	72,4
Resnet 18	73	62	94	34	94	71,5
Resnet 34	88	57	98	30	98	74,1
VGG 19	80	72	98	26	92	73,7
Resnet 152	54	85	92	28	80	67,8
Resnet 50	77	53	96	32	100	71,5
Resnet 101	64	57	94	34	100	69,8
Densenet 169	91	70	96	42	98	79,3
VGG 11	52	79	90	26	100	69,5
VGG 13	77	74	84	47	84	73,1
Densenet161	59	85	98	42	100	76,7
Densenet 201	73	55	90	51	100	73,8
VGG 16	77	81	94	32	98	76,4
Densenet 121	57	55	98	52	94	71,2

Model transferinin doğruluk sonuçlarında eğitim ve test olarak kullanılan veri setlerine göre değişiklik gösterdiği gözlenmiştir. Aynı ESA modelleri kullanılarak yapılan model

transfer edilebilirliği potansiyellerinin değiştiği görülmüştür. Eğitim ve test aşamasında aynı veri setlerinin kullanıldığı uygulamaların sonuçlarına göre, alan kaydırma uygulamasının sonuçlarında ESA modellerinin doğruluk değerlerinde farklılıklar gözlenmiştir.

4. Sonuç

Bu çalışmada yapılan deneylerde her iki veri seti için kullanılan on dört farklı ESA modelinin performansları niceliksel olarak karşılaştırılmıştır. Çalışmada kullanılan ESA mimarilerinin çeşitli katman sayısına ait farklı versiyonlarına yer verilmiştir (Densenet 161,169,201, Resnet 18,34,50 ve VGG11,13,16,19). Her iki veri setinde de benzer olan sınıfların doğrulukları karşılaştırılmış ve sonuçlar yorumlanmıştır. Söz konusu deneylerde en iyi sınıflandırma sonucuna sahip algoritma Densenet201 modeli ile yapılan sınıflandırmanın F1 skoru, kesinlik ve duyarlılık açısından karşılaştırmalı olarak sunulmuştur. Aynı zamanda semantik özellikler olarak, çalışmada kullanılan veri setlerinde kumsal, orman, otopark, nehir ve liman şeklinde benzer sınıflar mevcuttur (Şekil 3). Eğitilen modellerin farklı bölge ve özelliklere ait benzer sınıfların etiketlenilmesinde kullanılabilirliği de test edilmiştir. CNN modellerinin geniş kapsamlı karşılaştırmalı analizi gerçekleştirilmiştir. Ayrıca eğitim ve test verilerinin, birbirinden farklı özellikteki veri kaynaklarından elde edildiği durumlarda derin öğrenme mimarilerinin sınıflandırma performansları alan kaydırma uygulamaları ile incelenmiştir.

Bu çalışma kapsamında elde edilen sonuçlar, literatürdeki diğer çalışmaların sonuçlarıyla karşılaştırılmıştır (Tablo 7,8). Bu karşılaştırmada UC Merced veri seti için, bu çalışmadaki sonuç diğer sonuçlardan daha yüksektir (Tablo 7).

Tablo 7. UC Merced veri setinin diğer çalışmalarla karşılaştırılması. Değerler % olarak verilmiştir. (Table 7. Comparison of the UC Merced dataset with other studies. Values are given in %.)

Kullanılan Model	GD	Referans
CNN Data Augmentation	82.29	Stivaktakis vd. (2019)
CNN with Overfeat feature	92.40	Marmanis vd. (2016)
VLAT	94.30	Negrel vd. (2014)
Bu çalışma (DenseNet201 uygulaması)	98.81	

Tablo 8. WHU-RS19 veri setinin diğer çalışmalarla karşılaştırılması. Değerler % olarak verilmiştir. (Table 8. Comparison of the WHU-RS19 dataset with other studies. Values are given in %.)

Kullanılan Model	GD	Referans
MINet	97.57	Hu vd. (2021)
Tex-Net-LF	98.48	Anwer vd. (2018)
Bu çalışma (DenseNet201 uygulaması)	96.02	

Gelecekteki çalışmalarda, veri setlerinin gürültü azaltılmış versiyonu (Aksoy vd., 2020) kullanılabilir. Ayrıca, sonuçların varyasyonların gözlemlenmesi için bu veri setleri için, çoklu-sınıf sınıflandırmasına ek olarak çoklu-etiketli sınıflandırma uygulamaları gerçekleştirilebilir. Arazi kullanımı sınıflarının otomatik etiketlenilmesinde derin öğrenme modelleri her geçen gün geliştirilmektedir. Yapılan çalışmalarda geliştirilen tekniklerle uzaktan algılama verilerinin sınıflandırılmasında verimlilik artışı sağlanmaktadır.

Kaynakça

- Aksoy, A. K., Ravanbakhsh, M., Kreuziger, T., & Demir, B. (2020). CCML: A Novel Collaborative Learning Model for Classification of Remote Sensing Images with Noisy Multi-Labels. arXiv preprint arXiv:2012.10715.
- Anwer, R. M., Khan, F. S., van de Weijer, J., Molinier, M., & Laaksonen, J. (2018). Binary patterns encoded convolutional neural networks for texture recognition and remote sensing scene classification. ISPRS journal of photogrammetry and remote sensing, 138, 74-85.
- Atik, M. E., Donmez, S. O., Duran, Z., & İpbüker, C. (2018). Comparison Of Automatic Feature Extraction Methods For Building Roof Planes By Using Airborne Lidar Data And High Resolution Satellite Image. Proceeding Book of 7th International Conference on Cartography and GIS, 18-23 June 2018, Sozopol, Bulgaria.
- Atik, S. O., & İpbüker, C. (2021). Integrating Convolutional Neural Network and Multiresolution Segmentation for Land Cover and Land Use Mapping Using Satellite Imagery. Applied Sciences, 11(12), 5551.
- Bi, Q., Qin, K., Zhang, H., Xie, J., Li, Z., & Xu, K. (2019). APDC-Net: Attention pooling-based convolutional network for aerial scene classification. IEEE Geoscience and Remote Sensing Letters, 17(9), 1603-1607.
- Brown, M., & Süsstrunk, S. (2011, June). Multi-spectral SIFT for scene category recognition. In CVPR 2011 (pp. 177-184). IEEE.
- Cheng, G., Han, J., & Lu, X. (2017). Remote sensing image scene classification: Benchmark and state of the art. Proceedings of the IEEE, 105(10), 1865-1883.
- Diakogiannis, F. I., Waldner, F., Caccetta, P., & Wu, C. (2020). ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. ISPRS Journal of Photogrammetry and Remote Sensing, 162, 94-114.
- Donmez, S.O.; İpbüker, C. Investigation on Agent Based Models for Image Classification of Land Use and Land Cover Maps. In Proceedings of the 39th Asian Conference on Remote Sensing (ACRS): Remote Sensing Enabling Prosperity, Kuala Lumpur, Malaysia, 15–19 October 2018; pp. 2005–2008.
- Dönmez, Ş. Ö., & Tunc, A. (2016). Transformation methods for using combination of remotely sensed data and cadastral maps. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 4, 587-589.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE

- conference on computer vision and pattern recognition (pp. 770-778).
- Hu, J., Shu, Q., Pan, J., Tu, J., Zhu, Y., & Wang, M. (2021). MINet: Multilevel Inheritance Network-Based Aerial Scene Classification. *IEEE Geoscience and Remote Sensing Letters*, 19, 1-5.
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- Bapu, J., & Florinabel, J. (2021). An Improved RF-PNN Based Annotation Model With Texture and Edge Based Feature Descriptors.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 1097-1105.
- Laban, N., Abdellatif, B., Ebied, H. M., Shedeed, H. A., & Tolba, M. F. (2017, September). Performance enhancement of satellite image classification using a convolutional neural network. In *International Conference on Advanced Intelligent Systems and Informatics* (pp. 673-682). Springer, Cham.
- Marmanis, D., Datcu, M., Esch, T., & Stilla, U. (2015). Deep learning earth observation classification using ImageNet pretrained networks. *IEEE Geoscience and Remote Sensing Letters*, 13(1), 105-109.
- Muhammad, U., Wang, W., Chattha, S. P., & Ali, S. (2018, August). Pre-trained VGGNet architecture for remote-sensing image scene classification. In *2018 24th International Conference on Pattern Recognition (ICPR)* (pp. 1622-1627). IEEE.
- Napiorkowska, M., Petit, D., & Marti, P. (2018, July). Three applications of deep learning algorithms for object detection in satellite imagery. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium* (pp. 4839-4842). IEEE.
- Negrel, R., Picard, D., & Gosselin, P. H. (2014, June). Evaluation of second-order visual features for land-use classification. In *2014 12th International Workshop on Content-Based Multimedia Indexing (CBMI)* (pp. 1-5). IEEE.
- Özyurt, F., Ava, E., & Sert, E. (2020). UC-merced image classification with cnn feature reduction using wavelet entropy optimized with genetic algorithm.
- Penatti, O. A., Nogueira, K., & Dos Santos, J. A. (2015). Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 44-51).
- Petrovska, B., Zdravevski, E., Lameski, P., Corizzo, R., Štajduhar, I., & Lerga, J. (2020). Deep learning for feature extraction in remote sensing: A case-study of aerial scene classification. *Sensors*, 20(14), 3906.
- Qi, X., Zhu, P., Wang, Y., Zhang, L., Peng, J., Wu, M., ... & Mathiopoulos, P. T. (2020). MLRSNet: A multi-label high spatial resolution remote sensing dataset for semantic scene understanding. *ISPRS Journal of Photogrammetry and Remote Sensing*, 169, 337-350.
- Sen, O., & Keles, H. Y. (2020). On the evaluation of CNN models in remote-sensing scene classification domain. *PFG-Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 88(6), 477-492.
- Sheng, G., Yang, W., Xu, T., & Sun, H. (2012). High-resolution satellite scene classification using a sparse coding based multiple feature combination. *International journal of remote sensing*, 33(8), 2395-2412.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Sumbul, G., Charfuelan, M., Demir, B., & Markl, V. (2019, July). Bigearthnet: A large-scale benchmark archive for remote sensing image understanding. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium* (pp. 5901-5904). IEEE.
- Stivaktakis, R., Tsagkatakis, G., & Tsakalides, P. (2019). Deep learning for multilabel land cover scene categorization using data augmentation. *IEEE Geoscience and Remote Sensing Letters*, 16(7), 1031-1035.
- Weng, Q., Mao, Z., Lin, J., & Guo, W. (2017). Land-use classification via extreme learning classifier based on deep convolutional features. *IEEE Geoscience and Remote Sensing Letters*, 14(5), 704-708.
- Xia, G. S., Hu, J., Hu, F., Shi, B., Bai, X., Zhong, Y., ... & Lu, X. (2017). AID: A benchmark data set for performance evaluation of aerial scene classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(7), 3965-3981.
- Verma, Y. (2019). Diverse image annotation with missing labels. *Pattern Recognition*, 93, 470-484.
- Yang, Y., & Newsam, S. (2010, November). Bag-of-visual-words and spatial extensions for land-use classification. In *Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems* (pp. 270-279).
- Zhao, B., Zhong, Y., Xia, G. S., & Zhang, L. (2015). Dirichlet-derived multiple topic scene classification model for high spatial resolution remote sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 54(4), 2108-2123.
- Zhou, W., Newsam, S., Li, C., & Shao, Z. (2018). PatternNet: A benchmark dataset for performance evaluation of remote sensing image retrieval. *ISPRS journal of photogrammetry and remote sensing*, 145, 197-209.