

Araç Park Yerlerinin Doluluk Durumlarının Derin Öğrenme Yöntemi ile Tespit Edilmesi

Parking Space Occupancy Detection Using Deep Learning Methods

Fatih Can Akıncı
Bilgisayar Mühendisliği
Atılım Üniversitesi
06836, Ankara, Türkiye
akincifca@gmail.com

Murat Karakaya
Bilgisayar Mühendisliği
Atılım Üniversitesi
06836, Ankara, Türkiye
murat.karakaya@atilim.edu.tr

Özetçe—Bu çalışmada, park yerleri doluluk bilgilerinin gerçek zamanlı olarak tespit edilmesinde gömülü sistem üzerinde çalışan bir Evrişimsel Sinir Ağı (ESA) uygulaması geliştirilmiştir. Daha önce bu konuda yapılan çalışmalarda kullanılan sinir ağı modellerine göre çok daha düşük maliyetli gömülü sistem üzerinde hızlı ve yüksek doğrulukla çalışabilen bir ESA geliştirilmiştir. Bu düşük maliyetli sistemde kullanılan ağın başarısı, kendini görüntü sınıflandırma yarışmalarındaki başarısıyla kanıtlamış ağlardan biriyle, değişik ışık ve hava koşullarında elde edilmiş binlerce görüntü üzerinde yapılan testlerle karşılaştırılmıştır.

Anahtar Kelimeler —Akıllı Şehirler, Bilgisayarlı Görü, Derin Öğrenme, Evrişimsel Sinir Ağları

Abstract—This paper presents an approach for gathering information about the availability of the parking lots using Convolutional Neural Network (CNN) for image processing running on an embedded system. By using an efficient neural network model, we made it possible to use a very low cost embedded system compared to the ones used in previous works on this topic. This efficient model's performance is compared to one of the models that proved its accuracy in image classification competitions. In these tests, we used datasets that has thousands of different images taken from parking lots in different light and weather conditions.

Keywords —Smart Cities, Computer Vision, Deep Learning, Convolutional Neural Networks

I. GİRİŞ

Kişisel otomobillerin sayısının gün geçtikçe artmasıyla, sürücülerin şehir merkezlerinde park yeri bulmaları zorlaşmış ve her gün karşılaşmak zorunda oldukları bir problem haline dönüşmüştür. Uzun süreler boyunca park yeri arayan sürücüler trafik yoğunluğunun artmasına, gereksiz yakıt tüketimine ve çevre kirliliğine neden olmaktadır. Daha önce yapılan çalışmalarda trafik sıkışıklığına sebep olan araçların %30'nun park yeri aradığı belirtilmiştir [1]. Park yerlerindeki boş veya dolu alanların belirlenip, sürücülere bu bilgilerin aktarılmasıyla sürücülerin park yeri arama sürelerinin kısaltılması mümkün

olacaktır. Bu maksatla birçok sistem önerilmiş ve geliştirilmiştir [15].

Günümüzde kapalı park yerlerindeki park alanlarının doluluk durumlarının belirlenmesi için çeşitli sensörler kullanılmaktadır. Akıllı şehirler kavramıyla bu çözümleri şehir içindeki açık alanlara ve tüm sokaklara yaymak istediğimiz de ise bu sensörlerin kullanılması kurulum ve bakım maliyetleri gibi konular dikkate alındığında uygun olamamaktadır [2]. Bu yüzden yapılan araştırmalarda park alanları doluluk bilgilerine ulaşmak için kamera görüntülerini kullanan teknikler ön plana çıkmaktadır [3], [4]. Bu çalışmalarda, çeşitli senaryolar için özel olarak geliştirilmiş farklı görüntü işleme teknikleri üzerinde durulmuştur. Ancak, önerilen görüntü işleme tekniklerinin değişik park yerlerini gerçek zamanlı gözlemlenmede kullanılması için geliştirilebilir olmaları gerekmektedir. Bu sebeple bu araştırmada park yeri bilgilerini görüntü ile elde etmede daha verimli, doğruluğu yüksek ve en önemlisi değişik park alanları için geliştirilebilir olması [5] nedeniyle Derin Evrişimsel Sinir Ağı (DESA) yöntemi kullanılmıştır. DESA ile üretilen çözüm, görüntülerde oluşabilecek çeşitli bozulmalara (değişik ışık koşulları, gölgeler, görüntü üzerindeki kısmi engeller) rağmen yüksek doğrulukta sonuç vermektedir [5], [10].

Eğitim süreleri kullanılan DESA'nın yapısına göre farklılık göstermektedir. Görüntü sınıflandırmada GoogLeNet [6] (Inception) gibi yüksek doğrulukta başarı elde edebilen DESA'nın eğitilmesi için gereken süreler günümüz bilgisayarlarında saatler sürmektedir. Bu çalışmada, hem Inception DESA hem de onun kadar başarılı ancak daha kısa sürelerde eğitilebilen, daha düşük hesaplama gücüne ihtiyaç duyan MobileNet [7] DESA kullanılmıştır. Çalışmamızda, MobileNet ve Inception, aynı veri setleriyle eğitilmiş, bu aşamalarda geçen süreler ve alınan başarı oranları karşılaştırılmıştır.

Inception modelini kullanan daha önceki bazı çalışmalarda, örneğin [5], eğitilen DESA, kamera ve buna bağlı Raspberry Pi 2 [11] gibi cihazlar üzerinde çalıştırılmıştır. Söz konusu çalışmada önerilen kamera ve Raspberry Pi 2 modülünden oluşan akıllı kamera sisteminin maliyeti yaklaşık 100 dolar olmaktadır. Bu maliyet, yer sensörleri kullanan sistemlere göre düşük maliyetlidir. Ancak, çalışmamızda bu maliyeti daha da düşürmek amacıyla; sadece 5 dolar civarında satılan Raspberry Pi Zero [12] (RPi Zero) üzerinde çalışabilecek iki DESA çözümü (Inception ve MobileNet) üzerinde yoğunlaştık. Öğrenmenin transferi (Transfer Learning) tekniği kullanarak Inception ve MobileNet DESA ile elde ettiğimiz yeni DESA modellerini RPi Zero üzerinde çalıştırdık ve her iki modelin çalışma sürelerini karşılaştırdık. Sonuçlara bakıldığında, yeniden eğitilen MobileNet DESA kullanılarak oluşturulan modelin, RPi Zero üzerinde Inception DESA'na göre çok daha hızlı ve pratikte de kullanılabilir olduğu görülmüştür. Bu sayede RPi Zero ve buna bağlı bir kamera modülü sistemi ile 30'a yakın aracın park yerleri doluluk bilgileri gerçek zamanlı elde edilebilir. Bu da park yeri başına 1 dolardan çok daha az maliyet getirecektir.

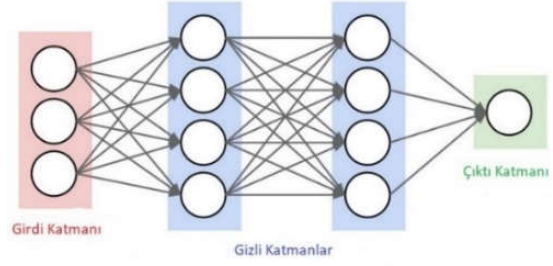
II. BOŞ PARK YERLERİNİ TESPİT ETMEK İÇİN DERİN ÖĞRENME YÖNTEMİNİN KULLANILMASI

Bu çalışmada park yerlerinin boş veya dolu olma durumlarının belirlenmesi amacıyla Derin Öğrenme (*Deep Learning*) yöntemi kullanılmıştır. Bilgisayarlar insanların gerçekleştiremediği karmaşık hesaplamaları yüksek hızda ve doğrulukta gerçekleştirebilmektedir. Buna karşın insanlar için çok kolay olan bazı işlemler (işitilen sözler ile cümle oluşturup anlama, görülen bir kişinin tanınması, bir fotoğraftaki cisimlerin belirlenmesi gibi) bilgisayarlar için zorlayıcı olmaktadır. Derin Öğrenme, Yapay Zekâ'nın alt dalı olarak, bilgisayarların insan seviyesine yakın algılama becerileri kazanması için geliştirilen bir yöntemdir. Günümüzde popülerleşen sürücüsüz arabalar, insansız hava araçları ve gerçek zamanlı obje takibi konularında derin öğrenme yöntemleri sıklıkla kullanılmakta ve önemini artırmaktadır.

Görüntü analizi ve bilgisayarlı görü gibi alanlarda DESA'ların kullanılması etkili sonuçlar vermektedir [5]. Bir DESA'da çok sayıda gizli katman (*hidden layer*) bulunmaktadır. Bu katmanlara girdi olarak verilen değerler matematiksel işlemlerden geçerek çıktı oluştururlar. Bu çıktılar da bir sonraki katmana girdi olarak verilir (Şekil-1). Görüntü analizi ve sınıflandırılması için belirlenmiş sınıflar üzerinden DESA eğitilir. Bu eğitim için DESA'nın üzerinde çalışacağı, daha önceden etiketlenmiş bir veri seti gereklidir. Görüntülerin analizinin yüksek hesaplama gücü gerektirmesinden dolayı, DESA çözümünün çalışacağı sistemin yapılandırılması başarımlar açısından çok önemlidir.

Bu çalışmada iki farklı DESA öğrenmenin transferi (transfer learning) tekniği kullanılarak tekrar eğitilmiştir. Yeniden eğitilen DESA, kendisine verilen park yerinein sayısal

görüntüsünü inceleyerek park yerlerinin boş ya da dolu olduğuna karar vermektedir. Önerilen çözümü başarımlarını ölçmek için iki farklı veri seti üzerinde çalışma yapılmıştır. Bu veri setleri hakkında detaylı bilgi aşağıda verilmiştir.



Şekil 1: DESA Yapısı [9]

A. Kullanılan Veri Setleri

Bu çalışmada iki farklı veri seti (PKLot [8] ve CNRPark [5]) kullanılmıştır. PKLot, değişik hava koşullarında (güneşli, bulutlu, yağmurlu) farklı park alanlarından elde edilmiş 700.000'den fazla park yeri görüntüsünden oluşmaktadır (Şekil-2). CNRPark ise iki farklı kameradan, değişik gün ve ışık koşullarında elde edilmiş yaklaşık 13.000 adet park yeri görüntüsünden oluşturulmuştur. Her iki veri setinde bulunan park yeri görüntüleri, boş veya dolu olacak şekilde etiketlenmiştir. DESA'nın eğitilmesi için bu veri setlerindeki görüntüler; eğitim (75%), doğrulama (15%) ve test (15%) olacak şekilde sınıflandırılmıştır. Böylece, eğitilen sinir ağını başarımının ölçülmesi esnasında daha önceden hiç görmediği test resimleri üzerinde çalışması gerekecektir. Eğer eğitimde ve test aşamasında aynı resimler kullanılırsa, sinir ağı veri setinin özelliklerini ezberleyeceği (overfitting) için başarı oranı gerçeği yansıtmayacak kadar yüksek çıkabilecektir. Birbirinden farklı görüntülerden oluşan üç veri setinin kullanımı ile DESA'nın sınıflandırma başarımı hakkında daha gerçekçi sonuçlar elde edilmek istenmiştir.



Şekil 2: Veri setlerinde bulunan park yeri görüntülerinden örnekler

B. Derin Evrimsel Sinir Ağları

Bu çalışmada Inception (v3) [6] ve MobileNet [7] olarak iki adet DESA kullanılmıştır. Inception, 2012 yılındaki ImageNet Geniş Görsel Tanıma Yarışması [13] verileri üzerinde önceden eğitilmiş bir DESA'dır ve "kedi", "bulaşık makinesi" veya "uçak" gibi yaklaşık 1000 farklı sınıfa ayrılmış görselleri birbirinden ayırt edebilmektedir. MobileNet ise Inception'a yaklaşan başarı yüzdesiyle ve eğitimi için Inception'a göre çok daha düşük hesaplama gücü gerektirmesi sayesinde, mobil cihazlar gibi daha az işlem kapasitesi bulunan cihazlar üzerinde çalıştırılabilen bir modeldir. Bu çalışmamızda MobileNet kullanılarak, Inception gibi daha karmaşık yapıdaki bir

DESA'nın başarısına yakın, eğitim süresi kısa ve dosya boyutu daha düşük bir model ortaya çıkarılmıştır. Bu sayede daha düşük işlem kabiliyetine sahip cihazlar üzerinde sınıflandırma işleminin gerçekleştirilmesi ve olası uygulamalarda sistemin maliyetinin düşürülmesi sağlanmıştır.

DESA'ların yeniden eğitimi için Intel i7 6700K işlemcili, 16 GB Ram'e sahip, 250 GB SSD ve GeForce GTX 1060 6GB ekran kartlı bilgisayar ve bu bilgisayar üzerinde çalışan Ubuntu 16.04 işletim sistemi ile Tensorflow 1.4 [14] uygulaması kullanılmıştır. Tensorflow açık kaynak olarak yayınlanan, sayısal hesaplamalar yapmaya yarayan bir yazılım kütüphanesidir. Kullanılan sinir ağlarının her ikisi için de analiz edilecek görüntü boyutu 224x224 piksel olarak belirlenmiştir. Veri setlerinde bulunan görüntüler, işlemlerden önce 224x224 boyutuna otomatik olarak ölçeklendirilmiştir. PKLot veri setindeki 700.000 adet görüntüden rastgele seçilen değişik sayıda görüntülerle dört farklı grup oluşturulmuştur. CNRPark veri seti ise iki gruba ayrılmış toplamda yaklaşık 13.000 görüntüden oluştuğundan bu veri seti Pklot ile kıyaslanabilir olması açısından daha fazla bölünmemiştir. Her iki veri seti için oluşturulan grupların detayları Tablo-1'de gösterilmiştir.

Tablo 1: PKLot ve CNRPark veri setlerinden oluşturulan gruplardaki dolu ve boş park yeri görüntüsü sayıları

	Dolu	Boş
PKLot-1	9687	11519
PKLot-2	9712	11478
PKLot-3	4611	5994
PKLot-4	4610	5951
CNRPark-A	3621	2550
CNRPark-B	4781	1632

Bu tabloda görüldüğü gibi gruplara ayrılan boş ve dolu park yeri görüntüleri, DESA'ların eğitilmesi için kullanılmıştır. DESA'lar ardışık olarak bulunan birden fazla katmandan meydana gelmektedir. Bu katmanlar önceden görüntü verileri üzerinde eğitilmiştir ve görüntüler arasında sınıflandırma yapmasını sağlayan bilgileri barındırmaktadır.



Şekil-3: Inception modeline ait bazı katmanlar [14]

Bu çalışmada, transfer öğrenmesi yöntemiyle sadece son katmanın eğitilmesi ve önceden eğitilen diğer katmanların bilgilerinin kullanılması amaçlanmıştır. Transfer öğrenmesi sırasında, eğitim dizinine ayırdığımız, boş veya dolu olacak şekilde etiketlenmiş park yeri görüntüleri kullanılarak sinir ağı eğitilir.

Şekil-3'te Inception modelinde bulunan art arda konumlandırılmış katmanlardan bazıları gösterilmiştir. Bu

katmanlar önceden eğitilmiş ve görüntüleri sınıflandırmada kullandığı birçok bilgiyi barındırmaktadır. Bu çalışmada kullanılan öğrenmenin transferi yöntemiyle sadece son katman Tablo-2'de gruplandırılmış veri setleriyle eğitilir.

Tablo 2: MobileNet/Inception DESA'ları performans karşılaştırması

	Başarı (%)	Geçen Süre (dk:sn)	Dosya Boyutu (MB)
PKLot-1	99,8 / 99,5	01:23 / 08:43	5,5 / 87,7
PKLot-2	99,9 / 99,6	00:50 / 05:11	5,5 / 87,7
PKLot-3	99,8 / 99,6	00:46 / 05:44	5,5 / 87,7
PKLot-4	99,7 / 99,8	00:49 / 04:28	5,5 / 87,7
CNRPark-A	99,9 / 99,7	00:34 / 02:50	5,5 / 87,7
CNRPark-B	97,3 / 99,6	00:37 / 02:56	5,5 / 87,7

Tablo-1'deki gruplara ayrılmış PKlot ve CNRPark veri setlerinden elde edilen görüntüler üzerinde MobileNet ve Inception modelleri transfer öğrenmesi yöntemiyle 500 adım boyunca beşer defa eğitilmiş ve test başarı yüzdelerinin ortalaması alınmıştır. Inception ve MobileNet'in eğitimi için geçen süreler, oluşturulan model dosya boyutu ve set içindeki test başarı oranları ortalamaları Tablo-2'de karşılaştırılmıştır.

Tablo 2'de görüldüğü gibi MobileNet tüm veri setlerinde çok daha hızlı bir şekilde eğitilmiş ve daha az yer kaplayan model dosyası oluşturulmuştur. Buna karşın Inception ile benzer başarıyı yakaladığı görülmektedir. MobileNet modeli kullanılarak CNRPark-A veri grubu kullanılarak yapılan eğitim sırasındaki başarı oranı grafiği Şekil-4'te gösterilmiştir.



Şekil-4: MobileNet kullanılarak yapılan eğitim sırasındaki sınıflandırma başarısının adım sayısına göre grafiği [14]

Bilgisayarda kısa sürelerde eğitip elde ettiğimiz DESA modelleri, RPi Zero üzerine kurulan Tensorflow ile kullanılarak, sınıflandırma işlemi test edilmiştir. Bu test sırasında bir park yeri görüntüsünün dolu ya da boş olma durumuna karar vermesinde geçen süreler kıyaslanmıştır (Tablo-3).

Tablo 3: Raspberry Pi Zero üzerinde DESA'ların çalışma performansı (1 adet park yeri görüntüsünün sınıflandırılması sırasında geçen süre)

Kullanılan DESA	Inception	MobileNet
Geçen Süre (sn)	72	9
Dosya Boyutu (MB)	87,7	5,5

Tablo-3'te görüldüğü gibi Inception temelli modellerin RPi Zero üzerinde sınıflandırma işlemi yaparken, MobileNet temelli modellere göre çok daha uzun süre gerektirdiği gözlemlenmiştir. Tek bir park yeri görüntüsünün analizi için gereken süreler bakıldığında Inception modelinin RPi Zero üzerinde çalıştırılması halinde yaklaşık 30 araçlık bir park alanındaki tüm park yerlerinin analizi için yaklaşık 36 dakika gerekmektedir. Bu süre de gerçek zamanlı bir sistem için pratikte uygun değildir. MobileNet kullanan sistem ise aynı sayıda park yeri için 5 dakikadan az bir süre gerekmektedir. Bu süreler göz önüne alındığında MobileNet kullanan RPi Zero ve kamera modüllü sisteminin pratik uygulamalarda kullanılması mümkün gözükmektedir.

III. SONUÇLAR VE GELİŞTİRİLEBİLİR YÖNLER

Trafikte park yeri arayan araçların yaşadığı sıkıntıları daha aza indirmek için sürücelere boş veya dolu park yerlerinin bilgilerinin aktarılması doğrultusunda park yerlerine yerleştirilecek olan kameralardan alınan görüntülerin derin öğrenme yöntemleriyle analizinin pratiğe uygunluğu değerlendirilmiştir. Bu değerlendirme yapılırken maliyet, gerçek zaman analiz süreleri gibi kriterler göz önünde bulundurulmuştur. Park yeri doluluk bilgilerinin çeşitli hava ve ışık koşullarına rağmen derin öğrenme yöntemleriyle yüksek doğrulukta tespit edilebileceği görülmüştür. Bu çalışmada ayrıca daha önce bu konuda derin öğrenme yöntemini kullanan çalışmalara göre doğruluk oranı aynı seviyede olmasına rağmen çok daha az maliyetli bir sistem meydana getirilmiştir.

Bu çalışmaya ek olarak elde edilecek park yeri video görüntülerinden sürekli olarak boş park yerlerini kontrol eden ve bu bilgiyi geliştirilecek mobil cihaz uygulamaları ile sürücülere görsel olarak aktaran bütün bir sistem geliştirmek gelecek çalışmalara bırakılmıştır.

IV. BİLGİLENDİRME

A. Fatih Can Akıncı

ODTÜ Elektrik-Elektronik Mühendisliği 2014 yılı mezunu. Şu anda Atılım Üniversitesi Bilgisayar Mühendisliği bölümünde yüksek lisans eğitimine devam ediyor. İlgi alanları arasında yapay zeka ve derin öğrenme konuları bulunuyor. Yazılım mühendisi olarak özel bir şirkette çalışmaya devam etmektedir.



B. Murat Karakaya

KHO Elektrik-Elektronik Mühendisliğinden lisans, Bilkent Üniversitesinden Bilgisayar Mühendisliğinden Yüksek Lisans ve Doktora derecelerini sırasıyla 1991, 2000 ve 2008 yılında almıştır. Şu anda Atılım Üniversitesi Bilgisayar Mühendisliği bölümünde Doç. Dr. olarak araştırmalarını devam ettirmektedir. Araştırma alanları arasında Sensör Ağları, Doğal Hesaplama, Makina Öğrenmesi, Nesnelerin İnterneti, ve Görüntü İşleme konuları bulunmaktadır.



V. KAYNAKLAR

- [1] Tsias C, Hobi L, Hofstetter F, Liniger S, Stiller B. Park IT smart: Minimization of cruising for parking. 24th International Conference on Computer Communications and Networks (ICCCN); University of Zurich, Switzerland. 2015. p. 1-8.
- [2] J. Yang, J. Portilla and T. Riesgo, "Smart parking service based on Wireless Sensor Networks," IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society, Montreal, QC, 2012, pp. 6029-6034.
- [3] N. Dan, "Parking management system and method," Jan. 2002, uS Patent App. 10/066,215.
- [4] Q. Wu, C. Huang, S.-y. Wang, W.-c. Chiu, and T. Chen, "Robust parking space detection considering inter-space correlation," in Multimedia and Expo, IEEE International Conference on. IEEE, 2007, pp. 659-662.
- [5] Amato, G., Carrara, F., Falchi, F., Gennaro, C. and Vairo, C. "Car parking occupancy detection using smart camera networks and Deep Learning," Computers and Communication (ISCC), 2016 IEEE Symposium on. IEEE, 2016.
- [6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097-1105.
- [7] Howard Andrew G. et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications" 2017.
- [8] P. R. de Almeida, L. S. Oliveira, A. S. Britto, E. J. Silva, and A. L. Koerich, "Pklot—a robust dataset for parking lot classification," Expert Systems with Applications, vol. 42, no. 11, pp. 4937-4949, 2015.
- [9] A. Karpathy, CS231n: Convolutional Neural Networks for Visual Recognition, Lecture Notes, 2017.
- [10] C. K. Ng, S. N. Cheong, E. Hajimohammadhosseinmemar and W. J. Yap, "Mobile outdoor parking space detection application," 2017 IEEE 8th Control and System Graduate Research Colloquium (ICSGRC), Shah Alam, 2017, pp. 81-86.
- [11] Raspberry Pi Zero web sitesi, <https://www.raspberrypi.org/products/raspberrypi-2-model-b>
- [12] Raspberry Pi Zero web sitesi, <https://www.raspberrypi.org/products/raspberrypi-zero>
- [13] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al. Imagenet large scale visual recognition challenge. arXiv preprint arXiv:1409.0575, 2014.
- [14] M. Abadi A. Agarwal P. Barham E. Brevdo Z. Chen C. Citro G.S. Corrado A. Davis J. Dean M. Devin S. Ghemawat I. Goodfellow A. Harp G. Irving M. Isard Y. Jia R. Jozefowicz L. Kaiser M. Kudlur J. Levenberg D. Mané R. Monga S. Moore D. Murray C. Olah M. Schuster J. Shlens B. Steiner I. Sutskever K. Talwar P. Tucker V. Vanhoucke V. Vasudevan F. Viégas O. Vinyals P. Warden M. Wattenberg M. Wicke Y. Yu X. Zheng "TensorFlow: Large-scale machine learning on heterogeneous systems", 2015.
- [15] F.C. Akıncı, M. Karakaya, "Şehirlerin Dijital Dönüşümü: Görüntü İşleme Yöntemlerinin Boş Park Yerlerinin Tespitinde Kullanılması", TBD 34. Ulusal Bilişim Kurultayı (BİLİŞİM 2017), 20-21 Aralık 2017.