Erciyes Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
Cilt 36, Sayı 2 , 2020

Erciyes University
Journal of Institue Of Science and Technology
Volume 36, Issue 2, 2020

İnsansız Hava Aracı Görüntülerinde Evrişimli Sinir Ağı Kullanarak Araç Sayımı için Yeni Bir Haritalama Yöntemi

Ersin KILIÇ*[1] , Serkan ÖZTÜRK[2],

*[1] Erciyes Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği, KAYSERİ
[2] Erciyes Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği, KAYSERİ

(Alınış / Received: 24.04.2020, Kabul / Accepted: 04.08.2020, Online Yayınlanma / Published Online: 17.08.2020)

Anahtar Kelimeler
Evrişimli Sinir Ağları,
Araba Sayımı,
Havadan Görüntüleme

Öz: İnsansız hava araçlarından (İHA) elde edilen görüntüleri yapay zeka sistemleri ile analiz eden otomatik sistemlerin sayısı giderek artmaktadır. Bu çalışmaların örneklerini şehircilik ve trafik uygulamalarında da görmekteyiz. Bu uygulamalar için görüntüde bulunan arabaların sayısının belirlenmesi oldukça önemlidir. Veri hazırlama ve etiketleme işlemi yapıldığı yönteme bağlı olarak çok zahmetli olmaktadır. Özellikle çerçeve etiketleme ile verilerin hazırlanması uzun zaman almaktadır. Nokta etiketleme ile verilerin hazırlanması çerçeve etiketleme ile hazırlanan süreyi 4 kat azaltmaktadır. Bu çalışmada nokta etiketleme ile hazırlanan veri ile İHA görüntülerinden araç sayımı öğrenebilen derin öğrenme modeli önerilmektedir. Nokta etiketleme ile derin modellerin eğitilmesi için yeni bir kayıp fonksiyonu önerilmiştir. Önerilen yöntemin başarımının değerlendirilmesi amacıyla CARPK veri seti üzerinde deneysel sonuçlar yapılmıştır. Elde edilen sonuçlar çerçeve etiketleme ile eğitilen modeller ile kıyaslandığında önerilen yöntemin bu modeller ile yarışabilecek düzeyde başarım elde ettiği görülmektedir.

A Car Counting Method in Aerial Images Based on Convolutional Neural Network

Keywords
Convolutional Neural Networks, Car Counting, Aerial Imaging

Abstract: Numbers of automated artificial intelligence systems that process images captured from unmanned aerial vehicles (UAV) are gradually increasing. Examples of these studies have been performed in urbanization and traffic applications. Determining the number of cars in the image is very crucial for these applications. Data preparation and labeling process is very laborious depending on the method it is performed. It takes a long time to prepare the data, especially with bounding box annotation. Preparation of data with point annotation reduces the time prepared with bounding box labeling by 4 times. In this study, a novel deep learning model that can learn vehicle counting from UAV images with data prepared with point labeling is proposed. A novel loss function has been proposed for the training of the model with point annotation. Experiments on the CARPK dataset show the competitive counting and localizing performance of the proposed method compared with existing methods that were been trained with bounding box annotations.

*İlgili Yazar, email: ersinkilic@erciyes.edu.tr

1. Giriş

İnsansız hava araçlarının (İHA) gelişimi ve yaygınlaşması sayesinde havadan çekilmiş görüntülerin analizi çok büyük önem kazanmıştır. Tarım, hayvancılık, şehir planlaması, trafik kontrolü, orman yönetimi ve benzeri birçok alanda yeni uygulamaların ortaya çıkmasına olanak sağlamıştır [1]. Bu alanlarda yürütülen uygulamalar genellikle görüntünün analizine ihtiyaç duymaktadır ve en önemli problemlerden birisi görüntüde bulunan nesnelerin

sayısının belirlenmesidir. Yolda ya da park halindeki arabaların sayılması şehir planlaması ve trafik kontrolü gibi alanlarda geliştirilen uygulamalarda sıklıkla karşılaşılan bir problemdir [2].

Nesne sayımı problemi, nesne tespiti probleminin nesne tespiti ise nesne segmentasyonu probleminin bir alt problemi olarak tanımlanabilir. Nesne segmentasyonu, görüntünün tüm piksellerinin belirli nesne sınıfları ile etiketlenmesidir. Bu da görüntüde bulunan tüm nesnelerin sınırlarının detaylı bir şekilde bulunmasını sağlamaktadır. Nesne tespiti, görüntüde bulunan belirli nesnelerin konum ve sınıflarının belirlenmesidir. Genellikle nesne tespitinde, nesneler dikdörtgen çerçeveler ile temsil edilmektedir. Nesne konumunun bulunması, nesneyi temsil eden çerçevenin en az iki noktasının koordinatlarının tespitidir. Dolasıyla, nesne segmentasyonu ve tespiti yöntemleri nesne sayımı probleminin çözümünde kullanılabilir. Nesne segmentasyonu ve tespitinde olduğu gibi nesne sayımında da evrişimli sinir ağları (CNN) tabanlı yöntemler son yıllarda en yüksek başarımı gösteren yöntemlerdir [3, 4, 5]. Gelişmiş CNN modelleri nesne segmentasyonu ve tespiti problemleri için denetimli öğrenme yaklaşımı ile eğitilmektedir. Bu modellerin eğitimi için piksel etiketleme, çerçeve etiketleme gibi yüksek seviyede imge etiketlemeye ihtiyaç duyulmaktadır.

İmge etiketleme genellikle 4 farklı şekilde yapılmaktadır. Bunlar; düşük seviyeden yüksek seviyeye doğru olmak üzere görüntü etiketleme, nokta etiketleme, çerçeve etiketleme ve piksel etiketlemedir. Görüntü etiketlemede, sadece görüntüde bulunan nesnelerin sınıfları kaydedilmektedir. Nokta etiketleme, görüntüde bulunan nesnelerin sınıfları ve nesnenin merkez noktası kaydedilmektedir. Çerçeve etiketleme, görüntüde bulunan nesnelerin sınıfları ve nesnenin sınırlarını belirleyen bir dikdörtgen çerçeve kaydedilmektedir. Dikdörtgenler, 2 nokta veya 4 nokta ile temsil edilmektedir. Eğer 4 nokta ile temsil ediliyorsa, nesnenin doğrultusu hakkında da bilgi taşımaktadır. Piksel etiketleme, görüntüde bulunan her piksel için bir nesne sınıfı atanarak kaydedilmesidir. Şekil 1'de 4 farklı imge etiketleme yöntemi için örnekler gözükmektedir. Düşük seviyeden yüksek seviyeye ilerledikçe imge etiketlemenin zaman maliyeti artmaktadır. Görüntü etiketleme de nesne başına 1 sn. harcanırken bu süre nokta etiketleme de 2.4 sn., çerçeve etiketleme de 10 sn., piksel etiketleme de ise 78 saniyedir [6]. Bu sebeple, düşük seviye imge etiketlemesi ile eğitilebilen başarılı nesne sayımı yöntemlerine ihtiyaç duyulmaktadır.



Şekil 9. Farklı imge etiketleme yöntemleri. a-) görüntü etiketleme, b-) nokta etiketleme, c-) çerçeve etiketleme, d-) piksel etiketleme

Nesne sayımında başarılı olan ilk CNN modelleri genellikle görüntü sınıflandırıcı ağların, görüntüde bulunan nesne sayınını tahmin eden tek çıkışlı ağlara dönüştürülmesi ile elde edilmiştir. Bu CNN modelleri çok boyutlu doğrusal olmayan regresyon modelleri olarak sınıflandırılmaktadır ve eğitimleri için görüntüde bulunan nesne sayısı kullanılmaktadır. Nesnelerin konum bilgisi olmadan eğitilen bu CNN modellerinin aktivasyon haritaları görselleştirildiğinde modellerin nesnelerin konumlarına odaklandığı görülmektedir [7]. Bu aktivasyon haritaları nesnelerin sayısını tahmin eden CNN modellerinin nesnelerin konumları hakkında da bilgi içerdiğini kanıtlamaktadır. Aslında onlarca farklı sınıftan nesnelerin tespiti ve segmentasyonunu başarılı şekilde gerçekleştiren CNN'ler için bu şaşırtıcı bir özellik değildir. Regresyon ya da nesne tespiti ile nesne sayımı gerçekleştiren modellerin başarı performanslarını artırmak için aktivasyon haritalarının regülasyonuna dayalı yöntemler kullanılmıştır [7]. Bu modellerin eğitimi için zahmetli ve zaman maliyeti yüksek bir işlem olan çerçeve etiketleme yapılması gerekmektedir.

Son yıllarda, havadan çekilmiş görüntülerde araba sayımı için yapılmış başarılı çalışmalar derin öğrenme yöntemlerini kullanmaktadırlar. Özellikle görüntü sınıflandırma, nesne tespiti ve segmentasyonu problemlerinde oldukça başarılı olan CNN modelleri kullanılmaktadır. Nesne sayımı problemi üzerine yapılan çalışmalar 2 farklı grupta sınıflandırılabilir: 1- Nesne tespiti yöntemleri ile nesne sayımı, 2- Regresyon yöntemleri ile nesne sayımı.

Son yıllarda nesne tespiti probleminde CNN modellerinin gösterdiği başarı diğer tüm yöntemleri gölgede bırakmıştır. En çok bilinen mimari R-CNN ailesidir. Gelişim sırasıyla birbirine takip eden 3 adet üyesi bulunmaktadır: R-CNN [8], Fast R-CNN [9] ve Faster R-CNN [3]. Faster R-CNN iki aşamalı nesne tespiti mimarilerinin en çok bilinen örneğidir. Faster R-CNN ve varyantları araba sayımı için kullanılmıştır [10]. YOLO [4] ve RetinaNet [5] gibi tek aşamalı nesne tespiti yöntemleri de araba sayımı problemi için kullanılmaktadır. Bu

yöntemlerin genel nesne tespiti problemlerinde başarılı olmalarına rağmen yoğun ve küçük nesneler içeren havadan çekilmiş görüntülerde nesne tespiti ve sayımı için revize edilmeleri yöntemlerin başarısını artırmaktadır. Hsieh ve ark. [10] 'Layout Proposal Network' (LPN) isimli nesnelerin uzamsal bilgilerini göz önünde bulundurarak nesne önerisi yapan yeni bir mimari önermişlerdir. Goldman ve ark. [11] yoğun nesne içeren görüntülerde nesne tespiti için oluşturulan CNN mimarilerinde son işlem aşaması olarak kullanılan maksimum olmayanı baskılama (non-maximum supression - NMS) yönteminin başarısız performansının oluşturduğu olumsuz etkiyi yok etmek için nesnelerin jaccard indekslerini öğrenmesini sağlayan Soft-IoU isimli yeni bir katman içeren mimariyi önermişlerdir. Li ve ark. [12] yoğun ve küçük nesnelerin tespiti için yeni ölçek uyarlamalı çerçeve üretim yöntemi ve tahmin edilen çerçevelerin kesin referans çerçevelere daha benzer olmalarını baskılayan yeni bir kayıp fonksiyonu önermişlerdir. Cai ve ark. [13] İHA görüntülerinde nesne tespiti ve sayımı için yeni bir çerçevesiz derin öğrenme modeli önermişlerdir. Nesnelerin merkez ve köşe noktalarını referans olarak alan bu yöntem önerdiği yeni bir modül ile nesnelerin küresel ve yerel görünüşlerini dikkate alarak lokalizasyon doğruluğunu artırmıştır. Önerilen yöntem eğitim aşamasında çerçeve etiketlemeye ihtiyaç duymamasıyla nesne tespitine dayalı diğer yöntemlerden ayrılmaktadır. Veri seti hazırlama süresini ve etiketleme zahmetini azaltmaktadır.

Regresyon tabanlı nesne sayımı yöntemleri nesnelerin konumlarını tespit etmeden sadece nesne sayısını bulmaya odaklanmaktadır [14, 15, 16, 17, 2]. Direkt olarak hedefe ulaşmaya çalışan bu yöntemler alt düzey görüntü özelliklerini kullanarak nesne sayısını tahmin etmektedir [14, 17]. Bazı yöntemler yoğunluk haritası üzerinden nesne sayısını tahmin etmeye çalışmaktadır. Lempitsky ve Zisserman [14] tasarladıkları nesne sayımı modelini eğitmek için nokta etiketleme ile gauss maskesi kullanarak kesin referans yoğunluk haritası oluşturma yöntemi önermişlerdir. Aich ve Stavness [7] bu kesin referans yoğunluk haritasını kullanarak CNN tabanlı modellerinin aktivasyon haritalarının regülasyonunu gerçekleştirerek regresyon tabanlı yöntemin daha başarılı olmasını sağlamışlardır. Regresyon tabanlı nesne sayımı yöntemlerinin istenilen düzeyde başarılı olmaları için çerçeve etiketlerini kullanan eğitim süreçlerine ihtiyaç duydukları yapılan çalışmalarda görülmektedir. Önerilen yöntem nokta etiketleme kullanarak regresyon tabanlı yöntemler ile yarışabilecek düzeyde başarım göstermektedir. Ayrıca, yöntemimiz nesnelerin merkez noktalarını da tespit ederek nesne konumu hakkında da bilgi oluşturmaktadır.

Bu çalışmada, havadan çekilmiş görüntülerde araba sayımı için nokta etiketleme ile eğitilen basit ve efektif yeni bir evrişimli sinir ağı tabanlı derin öğrenme modeli önerilmiştir. Önerilen yöntem nesnelerin merkez noktalarını tespit eden tek aşamalı bir nesne tespiti modeli içermektedir. Bu model nesnelerin merkez noktalarını ve görüntünün arka planını ayırt eden bir haritanın regresyonunu gerçekleştirmektedir. Önerilen yöntemde modelin eğitimi için veri setinde bulunan görüntülerin nokta etiketleme yöntemi ile etiketlenmesi yeterlidir. Çıkarsama aşamasını model tek bir ileri geçiş ile tamamlanmaktadır ve sonuç olarak 2-boyutlu bir harita üretmektedir. Tahmin edilen bu haritadan nesne sayısı ve merkez noktalarını belirlemek için son bir işlem aşaması uygulanmaktadır. Önerilen yöntemin başarısını ölçmek için CARPK [10] araba sayımı veri seti kullanılmaktadır. Önerilen yöntem oldukça basittir ve farklı alanlarda nesne sayımı ve tespiti için kullanılma potansiyeline sahiptir. Bu çalışmanın temel katkıları aşağıdaki gibi özetlenebilir:

1. Modelin eğitimi için nesnelerin merkez noktalarını ve görüntünün arka planını gösteren yeni bir harita oluşturma yöntemi ve bu haritanın eğitiminde kullanılacak yeni bir kayıp fonksiyonu önerilmiştir.
2. Evrişimli sinir ağları tabanlı derin öğrenme modelleri ile nesne sayımında nokta etiketleme kullanarak eğitilen, çerçeve etiketleme ile eğitilmiş yöntemler ile yarışabilecek düzeyde başarılı bir yöntem önerilmiştir.

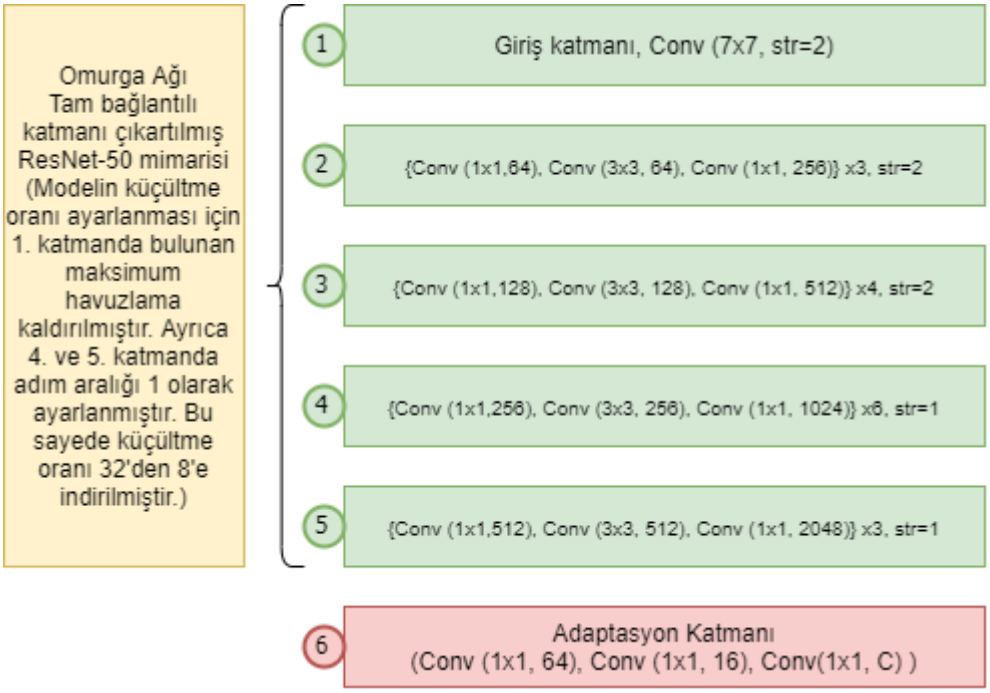
3. Önerilen yöntemin başarımı CARPK veri seti vasıtasıyla İHA kullanılarak çekilen görüntüler üzerinde test edilmiştir. Nesneler hakkında daha fazla bilgi kullanılarak eğitilen modeller ile yarışabilecek düzeyde başarı elde edilmiştir. Bu başarı performansı ileride yapılabilecek çalışmalar için umut vericidir.

2. Materyal ve Metot

Önerilen yöntem 2 temel aşamadan oluşturmaktadır: 1. Giriş görüntüsü üzerinden nesne ile arka planı ayırt eden bir harita oluşturan CNN modelidir. 2. Tahmin edilen haritadan nesne konum ve sayılarını tespit eden son işlem aşamasıdır. Nesne haritası tahmin eden CNN modelinin eğitimi için nokta etiketleme ile oluşturulan kesin referans haritaları kullanılmaktadır. Nesne haritasında arka plan 0 ile temsil edilmektedir. Nesneler ise 255 değeri temsil edilmektedir. Nesnelerin merkez noktaları ve istenilen kadar komşu pikselleri 255 değeri ile doldurularak nesne haritaları oluşturulmaktadır.

Önerilen CNN modeli arzu edilen her boyutta giriş verisini işleyebilmektedir. Çıkışta oluşan tahmini nesne haritası ise CNN modelinin küçültme oranına (kr) göre değişmektedir. Örnek olarak kr=8 ise 512x512 boyutunda giriş görüntüsü için çıkış olarak 64x64 boyutunda nesne haritası oluşmaktadır. CNN modeli tam bağlantılı (fully

connected) katmanı çıkartılan omurga ağ ve adaptasyon katmanı olmak üzere için ana parçadan oluşmaktadır. Omurga ağı olarak literatürde kullanılan AlexNet [18], VGG [19], ResNet [20] gibi herhangi bir omurga ağı kullanılabilmektedir. Omurga ağını takip eden adaptasyon katmanları 1x1 filtre boyutlu evrişim katmanlarından oluşmaktadır. Adaptasyon katmanında evrişim katmanlarının çıkışlarına sızdıran relu (leaky relu) aktivasyon fonksiyonu uygulanmıştır. Adaptasyon katmanı omurga ağ tarafından oluşturulan özellik haritalarının sayısını nesne sınıfı sayısına indirmektedir. Önerilen modelin girişi w x h x3 boyutunda ise çıkış w/kr x h/kr x C boyutunda olmaktadır. Burada C veri setinde bulunan nesne sınıfı sayısını temsil etmektedir. CNN modeli her bir nesne sınıfı için ayrı nesne haritası tahmin etmektedir. Omurga ağı olarak ResNet-50 kullanan modelin genel yapısı şekil 3'te gözükmektedir. Modelin küçültme oranı modelin içerdiği havuzlama katman sayısı ve adım aralığı özelliğine bağlıdır. Ayrıca modelin içerdiği evrişim katmanlarında adım aralığı sistemin küçültme oranını belirlemektedir.



Şekil 10. Omurga Ağı olarak ResNet-50 kullanan modelin genel mimarisi ('Conv' evrişim işlemini ve 'str' adım aralığını temsil etmektedir.)

Genellikle görüntülerde nesneler arka planlara göre çok az yer kaplamaktadır. Bu da oluşturulan veri setlerinde arka plan ve ön plan piksel sayılarında bir dengesizliğe yol açmaktadır. Bu da eğitilen modellerin taraflı olmalarına neden olmaktadır. Bu sebeple, oluşturulan sistemin kesin referans nesne haritasını daha kolay öğrenebilmesi için nesne (ön plan) bölgelerinde yapılan hatayı daha ön plana çıkaran bir kayıp fonksiyonu önerilmiştir. Arka plan piksellerinin sayısının nesne pikselleri sayısından çok daha fazla olması, eğitim verisinde dengesizlik oluşturmaktadır. Bu da sistemin arka plan sınıfını daha fazla tahmin etme eğilimine itmektedir. Bunun önüne geçmek için ortalama mutlak hata (MAE) revize edilmiştir. Kesin referans nesne haritasından arka plan pikselleri ve ön plan piksellerinde yapılan hatalar ayrı ayrı hesaplanıp toplanarak toplam hata elde edilmektedir. Arka plan hatasının hesaplanması için standart ortalama mutlak hata olduğu gibi kullanılmıştır. Ön plan hatası ise toplam mutlak hatanın nesne sayısına bölünmesiyle elde edilmektedir. Önerilen kayıp fonksiyonunun formülleri 1, 2 ve 3. Denklemde gözükmektedir. y model tarafından tahmin edilen haritayı, x kesin referans nesne haritasını ve k görüntüde bulunan nesne sayısını temsil etmektedir. Kesin referans nesne haritasında bir nesne m adet piksel ile temsil ediliyorsa; ön plan hatası arka plan hatasına göre m kat daha önemli olması bu şekilde sağlanmaktadır. Bu da doğal olarak dengesiz olan veriden eğitilen modelin aşırı öğrenmeden kaçınmasına ve daha kolay şekilde eğitilmesine katkı sağlamaktadır.

$$L_{bg} = \frac{\sum_{i=1}^{n} |yi - xi|}{n} \quad (1)$$

$$L_{fg} = \frac{\sum_{i=1}^{m} |yi - xi|}{k} \quad (2)$$

$$L = L_{bg} + L_{fg} \quad (3)$$

Tahmin edilen haritadan nesne merkezlerinin ve sayının belirlenmesi için ilk aşamada harita ikili görüntüye dönüştürülmektedir. Bu önceden belirlenmiş bir t eşik değerine göre yapılmaktadır. İkili görüntüde oluşan nesne bölgelerinin merkez noktalarını ve sayısını belirlemek amacıyla bağlı bileşen analizi [21] yöntemi kullanılmaktadır.

3. Bulgular

İHA görüntülerinden araba sayımı sisteminin başarımının değerlendirilmesi için CARPK veri seti kullanılmıştır. Değerlendirme kriteri olarak ortalama mutlak hata (MAE) ve ortalama hata kareleri kökü (RMSE) kriterleri kullanılmıştır.

3.1. Veri Seti

CARPK veri seti 989 adet eğitim görüntüsü ve 454 adet test görüntüsünden oluşmaktadır. Veri setinde bulunan görüntüler İHA ile 4 farklı otoparktan elde edilmiştir. Görüntülerde toplamda 89777 adet araba bulunmaktadır ve bu arabalar çerçeve etiketleme yöntemi ile etiketlenmiştir. Görüntüler yaklaşık olarak 40 metre uzaklıktan elde edilmiştir. Şekil 4'te veri setinden alınan örnek görüntüler gözükmektedir.

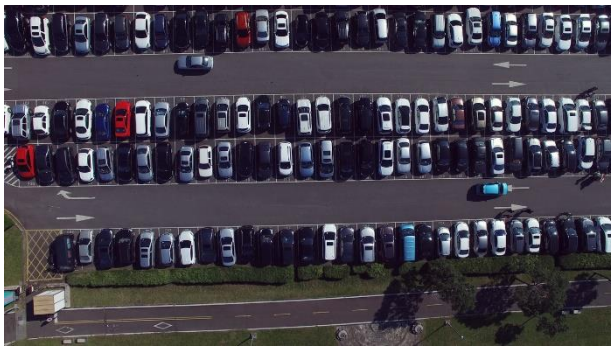
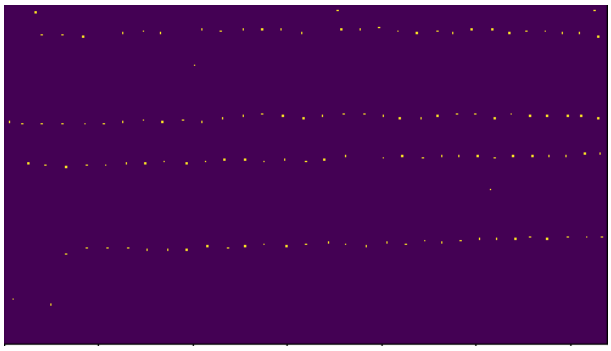
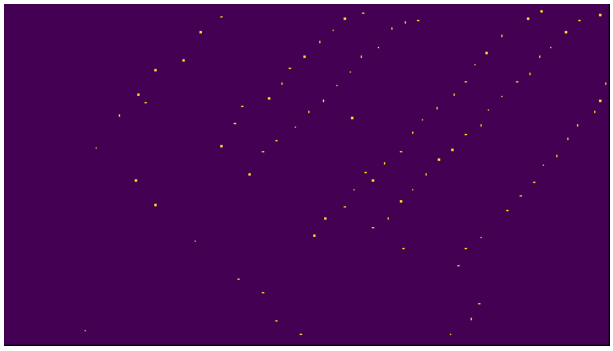
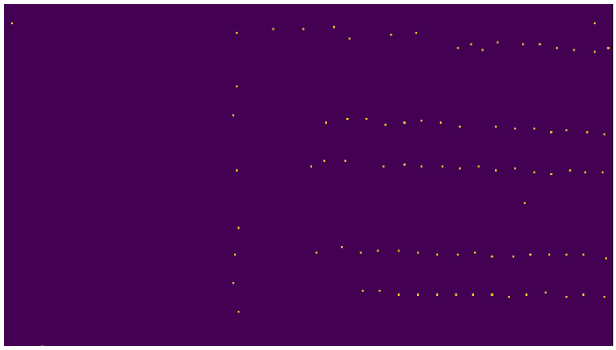
3.2. Deneysel Bulgular

Önerilen yöntemin başarımının değerlendirilmesi için CARPK üzerinden yapılan deneysel çalışmalarda omurga ağı olarak ResNet-50 mimarisi kullanılmıştır. Omurga ağı takip eden adaptasyon katmanında 3 adet evrişim katmanı bulunmaktadır. Bu katmanlar omurga ağından gelen 512 adet özellik haritasını sırasıyla 64, 16 ve 1'e indirmektedir. Model için kr 4 olarak belirlenmiştir. Eğitim esnasında görüntüler 540x960 boyutuna getirilmiştir. Eğitim için oluşturulan kesin referans nesne haritasında her bir nesne 2 piksel ile temsil edilmiştir. Modelin aşırı öğrenmeden kaçınması amacıyla veri artırma tekniklerinden döndürme, HSV (Hue, Saturation, Value – Renk Özü, Doygunluk, Parlaklık) uzayında renk değiştirme ve arka plan görüntüsü ekleme teknikleri kullanılmıştır. Döndürme veri artırma tekniği 0.2 olasılık oranında görüntünün -45 ila 45 derece arasında döndürülmesi ile uygulanmıştır. HSV uzayında renk değiştirme 0.5 olasılıkla uygulanmıştır. Arka plan ekleme veri artırma tekniği ise hiçbir nesne içermeyen 19 adet görüntü eklenerek uygulanmıştır. Bu sayede sistemin hatalı sayım oranı azaltılmıştır. Model 60 epok eğitilmiştir. Eğitim oranı 0.0001 olarak seçilmiştir ve her 20 epokta bir 0.1 ile çarpılarak azaltılmıştır. Eğitim için ADAM (A Method for Stochastic Optimization - Stokastik Optimizasyon için Bir Yöntem) optimizasyon algoritması kullanılmıştır.

Tablo 1. Önerilen yöntem ve kıyaslanan yöntemlerin CARPK veri seti üzerindeki MEA ve RMSE kriterleri ile başarımları

Yöntem	Etiketleme Yöntemi	MAE	RMSE
Faster R-CNN [10, 3]	çerçeve	24,32	37,62
YOLO [10, 4]	çerçeve	48,89	57,55
One-Look Regression [2, 10]	Nesne sayısı	59,46	66,84
ShuffleDet [22]	çerçeve	26,75	38,46
LPN [10]	çerçeve	23,80	36,79
RetinaNet [5]	çerçeve	16,62	22,30
IoUNet [11]	çerçeve	6,77	8,52
VGG-GAP-HR [7]	çerçeve	7,88	9,30
GANet [13]	çerçeve	4,80	6,90
Amato vd. [23]	çerçeve	3,72	5,11
Önerilen Yöntem	merkez nokta	9,30	13,44

Tablo 2. Örnek görsel sonuçlar.

Giriş Görüntüsü	Elde edilen nesne haritası
	
	
	

Eğitilen modelin başarı performansının etkinliğinin değerlendirilmesi için farklı yöntemlerin CARPK veri seti üzerindeki başarımları ile kıyaslanmıştır. Kıyaslama sonuçları ve yöntemlerin kullandığı etiketleme yöntemleri Tablo 1'de gözükmektedir. Önerilen model MAE kriteri üzerinden 9.30 ve RMSE kriteri üzerinden 13.40 hata oranını elde etmiştir. Elde edilen sonuçlar nesnelerin nokta etiketlerini kullanarak eğitilen önerilen yöntemin nesne çerçeve etiketlerini kullanarak eğitilen bir çok yöntemden daha başarılı sonuçlar elde ettiğini ve diğer yöntemler ile yarışabilecek düzeyde başarım gösterdiğini ortaya koymuştur. Etiketleme maliyetini azaltan yöntemin sağladığı verimlilik nesne sayımı başarımının kıymetini artırmaktadır. Yöntemin nesnelerin merkez noktalarını tespit ederek çıktı olarak sunması da artı bir özelliğidir. Deneysel çalışmalar sonucunda elde edilen görsel sonuçlar Tablo 2'de gözükmektedir.

4. Tartışma ve Sonuç

Bu çalışmada imge etiketleme maliyeti düşük nokta etiketleme ile eğitilebilen yeni bir CNN tabanlı araba sayımı derin öğrenme modeli önerilmektedir. Model elde edilen deneysel sonuçların ışığında çerçeve etiketleme kullanan birçok yöntemden daha iyi sonuç elde etmiştir. Nokta etiketleme ile eğitilen sistemin kolay ve verimli bir şekilde eğitilebilmesi için yeni bir kayıp fonksiyonu önerilmiştir. Önerilen yöntem çerçeve etiketleme kullanan nesne tespiti modelleri ile yarışabilen bir yöntem olduğu deneysel sonuçlarla gösterilmiştir. Gelecek çalışmalarda farklı nesneler için yöntemin uygulanması amaçlanmaktadır.

Teşekkür

Bu çalışma Erciyes Üniversitesi tarafından Bilimsel Araştırma Projeleri kapsamında desteklenmiştir. (FDK-2018-8624)

Kaynakça

[1] C. Huai, Z. Libao, M. Jie ve Z. Jue, «Target heat-map network: An end-to-end deep network for target detection in remote sensing images,» *Neurocomputing,* no. 331, pp. 375-387, 2019.

[2] N. Mundhenk, G. Konjevod, W. Sakla ve B. Kofi, «A Large Contextual Dataset for Classification, Detection and Counting of Cars with Deep Learning,» %1 içinde *The 14th European Conference on Computer Vision*, Amsterdam, 2016.

[3] S. Ren, K. He, R. Girshick ve J. Sun, «Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,» *IEEE Transactions on Pattern Analysis and Machine Intelligence,* cilt 39, no. 6, pp. 1137 - 1149, 2017.

[4] J. Redmon, S. Divvala, R. Girshick ve A. Farhadi, «You Only Look Once: Unified, Real-Time Object Detection,» *CoRR,* cilt abs/1506.02640, pp. 1-15, 2015.

[5] T.-Y. Lin, P. Goyal, R. Girshick, K. He ve P. Dollar, «Focal Loss for Dense Object Detection,» *CoRR,* cilt abs/1708.02002, pp. 1-12, 2017.

[6] A. Bearman, O. Russakovsky, V. Ferrari ve L. Fei-Fei, «What's the Point: Semantic Segmentation with Point Supervision,» %1 içinde *European Conference on Computer Vision*, SANTIAGO, CHILE, 2015.

[7] I. Stavness ve S. Aich, «Improving Object Counting with Heatmap Regulation,» *CoRR,* pp. 1-17, 2018.

[8] R. Girshick, J. Donahue, T. Darrell ve J. Malik, «Rich feature hierarchies for accurate object detection and semantic segmentation,» *CoRR,* cilt abs/1311.2524, pp. 1-12, 2013.

[9] R. Girshick, «Fast R-CNN,» *CoRR,* cilt abs/1504.08083, pp. 1-15, 2015.

[10] M.-R. Hsieh, Y.-L. Lin ve W. Hsu, «Drone-based Object Counting by Spatially Regularized Regional Proposal Networks,» %1 içinde *International Conference on Computer Vision*, Venice, 2017.

[11] E. Goldman, R. Herzig, O. Ratzon, I. Levi, J. Goldberger ve T. Hassner, «Precise Detection in Densely Packed Scenes,» *CoRR,* cilt abs/1904.00853, pp. 1-15, 2019.

[12] W. Li, H. Li, Q. Wu, X. Chen ve K. N. Ngan, «Simultaneously Detecting and Counting Dense Vehicles From Drone Images,» *IEEE Transactions on Industrial Electronics,* cilt 66, no. 12, pp. 9651 - 9662, 2019.

[13] Y. Cai, D. Du, L. Zhang, L. Wen, W. Wang, Y. Wu ve S. Lyu, «Guided Attention Network for Object Detection and Counting on Drones,» *CoRR,* pp. 1-15, 2019.

[14] V. Lempitsky ve A. Zisserman, «Learning To Count Objects in Images,» %1 içinde *Advances in Neural Information Processing Systems 23*, Curran Associates, Inc., 2010, pp. 1324--1332.

[15] L. Boominathan, S. S. S. Kruthiventi ve R. V. Babu, «CrowdNet: A Deep Convolutional Network for Dense Crowd Counting,» %1 içinde *24th ACM international conference on Multimedia*, Amsterdam , 2016.

[16] W. Xie, J. A. Noble ve A. Zisserman, «Microscopy cell counting and detection with fully convolutional regression networks,» *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization,* cilt 6, no. 3, pp. 283-292, 2018.

[17] L. Fiaschi, R. Nair, U. Köthe ve F. Hamprecht, «Learning to count with regression forest and structured labels,» %1 içinde *21st International Conference on Pattern Recognition*, Tsukuba Science City, 2012.

[18] A. Krizhevsky, I. Sutskever ve G. E. Hinton, «ImageNet Classification with Deep Convolutional Neural Networks,» %1 içinde *Advances in Neural Information Processing Systems 25*, Curran Associates, Inc., 2012, pp. {1097--1105.

[19] K. Simonyan ve A. Zisserman, «Very Deep Convolutional Networks for Large-Scale Image Recognition,» *CoRR,* pp. 1-11, 2014.

[20] K. He, X. Zhang, S. Ren ve J. Sun, «Deep Residual Learning for Image Recognition,» %1 içinde *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 2016.

[21] L. He, Y. Chao, K. Suzuki ve K. Wu, «Fast connected-component labeling,» *Pattern Recognition,* cilt 42, no. 9, pp. 1977--1987, 2009.

[22] S. M. Azimi, «ShuffleDet: Real-Time Vehicle Detection Network in On-Board Embedded UAV Imagery,» %1 içinde *European Conference on Computer Vision*, Munich, 2018.

[23] G. Amato, L. Ciampi, F. Falchi ve C. Gennaro, «Counting Vehicles with Deep Learning in Onboard UAV Imagery,» %1 içinde *IEEE Symposium on Computers and Communications*, Barcelona, 2019.