



Model ve Veri Odaklı Yaklaşımlar ile Nesne Tespit Başarısının Arttırılmasına Yönelik ESA ve Veri Seti Optimizasyonları

CNN and Dataset Optimizations to Improve Object Detection Performance with Model and Data-Centric Approaches

¹Önder ALPARSLAN , ²Ahmet ÖZCAN , ³Ömer ÇETİN 

^{1,2,3}Milli Savunma Üniversitesi, Hezârfen Havaçılık ve Uzay Teknolojileri Enstitüsü, Yeşilyurt/İstanbul, Türkiye

¹oalparslan@hho.msu.edu.tr, ²iletisim@ahmetozcan.com,

³omer_cetin@outlook.com.tr

Araştırma Makalesi/Research Article

ARTICLE INFO

Article history

Received : 20 April 2022

Accepted : 29 June 2022

Keywords:

Object Detection and Classification,
Model-Centric Approaches,
Data-Centric Approaches

ABSTRACT

To improve the accuracy of object detection and classification methods, model and data-centric approaches have been frequently studied by researchers in recent years. Most researchers propose methods based on their problem or present new methods by changing the existing models. On the other hand, the number of studies on the data used in the training process is quite small. In this study, the effects of model and data-centric approaches on an existing object detection problem are compared. The accuracy and performance effect obtained with YOLOv4 network, which has widespread usage, and reprocessing of the dataset are compared and evaluated. While the accuracy of object detection increased by about 4% with the change in the network structure, an average of 8% decrease occurred in the computational speed. On the other hand, reprocessing the data and running the object detection algorithm provided 6% gain, but did not cause a change in the calculation cost. Although it is not considered enough today, it is observed that the preparations of the data have a significant effect on the classification accuracy.

© 2022 Bandırma Onyedi Eylül University, Faculty of Engineering and Natural Science. Published by Dergi Park. All rights reserved.

MAKALE BİLGİSİ

Makale Tarihleri

Gönderim : 20 Nisan 2022

Kabul : 29 Haziran 2022

Anahtar Kelimeler:

Nesne Tanıma ve Sınıflandırma
Model Odaklı Yaklaşımlar,
Veri Odaklı Yaklaşımlar.

ÖZET

Nesne tespit ve sınıflandırma yöntemlerinin başarısının artırılması amacıyla model odaklı ve veri odaklı yaklaşımlar araştırmacılar tarafından son yıllarda sıklıkla çalışılmaktadır. Araştırmacıların birçoğu problemlere özgü model önerilerinde bulunmakta ve mevcut modeller üzerinde değişimler önermektedir. Öte yandan, eğitim sürecinde kullanılmakta olan veri üzerinde yapılan çalışmaların sayısı oldukça azdır. Bu çalışmada, mevcut bir tanıma ve sınıflandırma problemi üzerinde, model ve veri odaklı yaklaşımların etkileri kıyaslanmıştır. Yaygın kullanıma sahip olan YOLOv4 ağı üzerinde yapılan ağ yapısı değişikliğinin başarı ve performansa etkisiyle, veri setinde kullanılan verilerin yeniden hazırlanmasıyla elde edilen başarı karşılaştırılarak yorumlanmıştır. Ağ yapısının değişimi ile nesne tanıma başarısı yaklaşık %4 oranında artarken, hesaplama hızında ortalama %8'lik düşüş meydana gelmiştir. Öte yandan verilerin yeniden hazırlanarak nesne tanıma algoritmasının çalıştırılması %6 oranında kazanç sağlarken, hesaplama maliyetinde değişime neden olmamıştır. Günümüzde yeteri kadar dikkate alınmasa da veri üzerindeki hazırlıkların sınıflandırma doğruluğuna önemli derecede etki yaptığı gözlemlenmiştir.

© 2022 Bandırma Onyedi Eylül Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi. Dergi Park tarafından yayınlanmaktadır. Tüm Hakları Saklıdır.

ORCID ID: ¹0000-0001-8803-1597

²0000-0002-1098-7078

³0000-0001-5176-6338

1. GİRİŞ

Günümüzde robotik, güvenlik, sağlık ve doğa bilimleri gibi birçok alanda etkinlikle kullanılmakta olan nesne tespit ve sınıflandırma yöntemlerinin başarı ve hız performanslarını artırmaya yönelik yoğun çalışmalar sürdürülmektedir. Bu çalışmalara ait yaklaşımları iki kategoride inceleyebiliriz. Birinci kısımda eğitim ve sınıflandırma modelinin güncellenmesi yer alırken, diğer kısımda verilerin güncellenmesi yer almaktadır. Yapay zekâ araştırmacıları, matematiksel model ve optimizasyon yöntemlerindeki gelişmeler sayesinde model odaklı yaklaşımlardan daha sık faydalanmaktadır [1-3]. Öte yandan, veri odaklı yaklaşımların ele alınması da gündeme gelmeye başlanmıştır [4].

Nesne sınıflandırma yöntemlerinde sıklıkla evrimsel sinir ağları tercih edilmektedir [5]. Son yıllarda, mobil cihazlarda ve araçlarda nesne sınıflandırmasının kullanımı ile hızlı ve etkin sonuçlar üretebilen düşük maliyetli yöntemler daha popüler hale gelmiştir. Faster R-CNN, YOLO, SSD, DSSD, RetinaNet ve SPP-Net [6-11] gibi algoritmalar, giriş verisini evrimsel bir sinir ağı modeli içerisinde sınıflandırmaktadırlar. Görüntü üzerinde nesne sınıflandırma için kullanılan bu yöntemler, birden çok işleme katmanından oluşan hesaplama modellerinin, birden çok soyutlama düzeyi sayesinde verilerin temsillerinin öğrenmesini sağlar [12]. Krizhevsky'nin çalışmasıyla [13] oldukça popüler hale gelen derin öğrenme teknikleri, özellikle mobil robotlarda ve cihazlarda yüksek performans elde edilebilmek için tek bir tespit ağı kullanımını ve ağı küçültülmesini içeren yöntemlere evrilmektedir. Ancak işlem yükünü azaltan bu yöntemler, tespit başarısının düşmesine sebep olduğundan araştırmacılar hesaplama maliyetini artırmadan tespit doğruluğunu artıracak geliştirmeleri araştırmaktadırlar.

Nesne tespit modelinin başarısının artırılabilmesi için giriş verisi üzerinde ön işlemler gerçekleştirilebilir veya ağ yapısında değişikliklere gidilebilmektedir. Giriş verisi üzerinde yapılacak özellik çıkarımları (SIFT, HOG) resim üzerinde aranan nesnelere ağ için daha belirgin hale getirebilmektedir [14, 15]. Bunun yanında, ağ hiperparametrelerinin ve ağ modelinin değişimleri de pozitif getiriler sağlayabilmektedir [16].

Veri setinin hazırlanmasında homojen ortamlar ve kategoriler kullanılamaması, kullanılan ekipmanların veri setini kullanacak kişilerden çok farklı olması, çevresel koşulların ve engelleyici durumların yeterince işlenmemesi ve kullanıcının sınıflandırma yapacağı ekipmanların ya da mekânın sıra dışı olması sebebiyle, sınıflandırma başarısı memnun edici seviyede olmayabilmektedir. Bu nedenlerle ihtiyaca özgü veri seti oluşturulması ya da mevcut veriler üzerinde çalışma yapılması da sınıflandırma başarısına katkı sağlayabilmektedir.

Bu çalışma ile nesne tespit ve sınıflandırma başarısının artırılabilmesi için kullanılacak model odaklı ve veri odaklı yaklaşımlar irdelenerek, mevcut bir model ve veri seti üzerinde yapılan değişikliklerle başarı sınaması yapılması ve sonuçların kıyaslanması amaçlanmıştır. Model odaklı çözümler ve veri odaklı yaklaşımlara duyulan ihtiyaç ikinci bölümde, mevcut bir sistem üzerinde yapılan değişiklikler üçüncü bölümde, yapılan değişikliklerin başarıya olan etkileri ve değerlendirmeleri ise dördüncü bölümde yer almaktadır.

2. MODEL ODAKLI VE VERİ ODAKLI YAKLAŞIMLAR

Derin öğrenme yöntemlerinin günümüzdeki karmaşık yapısı ve yüksek hesaplama maliyetlerine rağmen zorlu koşullarda başarısı güvenilir seviyelerin altındadır. Mevcut model mimarisini, kullanılan fonksiyonların ya da veri işleme tekniklerinin güncellenmesi ile daha yüksek başarı elde edilmeye çalışılmaktadır [1-3]. En sık kullanılan modellerden birisi olan YOLO için farklı geliştiriciler tarafından YOLOv4 [17] ve YOLOv5 [18] önerilmiştir. Bu modeller ve yapılan geliştirmelerle yöntemlerin başarısı artırılmakla birlikte, ağ daha yavaş çalışabilmekte ya da farklı örnekler için başarı oranı düşmektedir [1]. Kendi problemlerini ve modelin çalışma yöntemini iyi anlayabilen araştırmacılar, mevcut model üzerinde girdiler yaparak ihtiyaca özgü bir model oluşturup daha iyi sonuçlar elde edebilmektedir.

Literatürde yapılan çalışmaların çok büyük bir bölümünün model odaklı yaklaşımlar üzerine olduğu görülmektedir [19]. Ancak, yapay zekâ çalışmalarında kullanılan veri setlerinin birçoğu yetersiz, kirli hatta kimi zaman yanlış bilgilerle doludur [20]. Nesne tespit ve sınıflandırması için kullanılan veri setleri araştırmacılar tarafından özgün ya da açık kaynak verilerin toplanmasıyla oluşturulmakta ve Kaggle, Github gibi ortamlarda paylaşılmaktadır. Yaygın kullanılan hazır veri setlerinden alınan veriler; yanlış, eksik ya da uzmanlar tarafından bile ne olduğu konusunda uzlaşamayan etiketlere sahiptir. Şekil 1'de CIFAR-100'da verilen istakoz örneği gerçekte yengeç, Caltech-256'da verilen yunus örneği gerçekte kano, ImageNet'de verilen beyaz leylek örneği gerçekte siyah leylektir. CIFAR-100'da yer alan hamster, Caltech-256'da laptop ve ImageNet'de peygamber devesi örnekleri doğru örnekler olsa da görüntülerde yer alan kupa, insan ve fens teli nesnelere işaretlenmemiştir. Öte yandan şekilde en altta yer alan örneklerin ne olduğu konusunda ise uzlaşamamaktadır.

Northcutt ve arkadaşlarına [4] göre geniş kullanım alanına sahip veri setlerinin %3,4'ü yanlış etiketlenmiş durumdadır ve daha büyük veri setleri de bu hatalardan etkilenmektedir. Yukarıda verilen örneklerden görüleceği üzere kullanımda olan hazır veri setleri eğitimi olumsuz etkileyebilecek olan örneklere ve etiketlere sahiptir. Bu sebeple de veri hazırlama sürecine önem verilerek veriler ayıklanmalı ve etiketlemeleri hassas bir şekilde gerçekleştirilmelidir.

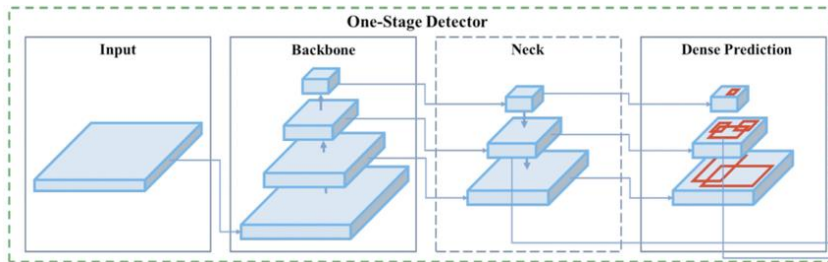
CIFAR-100 Caltech-256 ImageNet



Şekil 1. Hazır veri setlerine ait olumsuz etiketleme örnekleri [4].

3. YÖNTEM

Model ve veri odaklı yaklaşımların nesne tespiti başarısına olan etkilerini ölçmek için, kapalı ortamlarda yer alan kapı, pencere ve merdivenlerin tespit ve sınıflandırılması problemi ele alınmıştır. Bu nesnelere, robotik sistemlerin kapalı ortamda otonom yol planlamasında robotun diğer alanlara ya da bina dışına gitmesi için geçiş noktaları olarak kullanılabilir. Bu nesnelere teşhisi için mobil robotlarda kullanılacak performans ve tespit başarısı gösterebileceği değerlendirilen yöntemler incelenmiştir. Mobil robotlar için güncel tespit ve yöntemlerinin karşılaştırıldığı çalışmalar [21, 22] da incelenerek en optimal çözüm olacağı değerlendirilen YOLOv4 yöntemi tercih edilmiştir. Nesne tespit yöntemine ait modelin daha hassas sonuçları üretmesi amacıyla ağ yapısı değiştirilmiştir ve veri seti düzenlenmiştir.



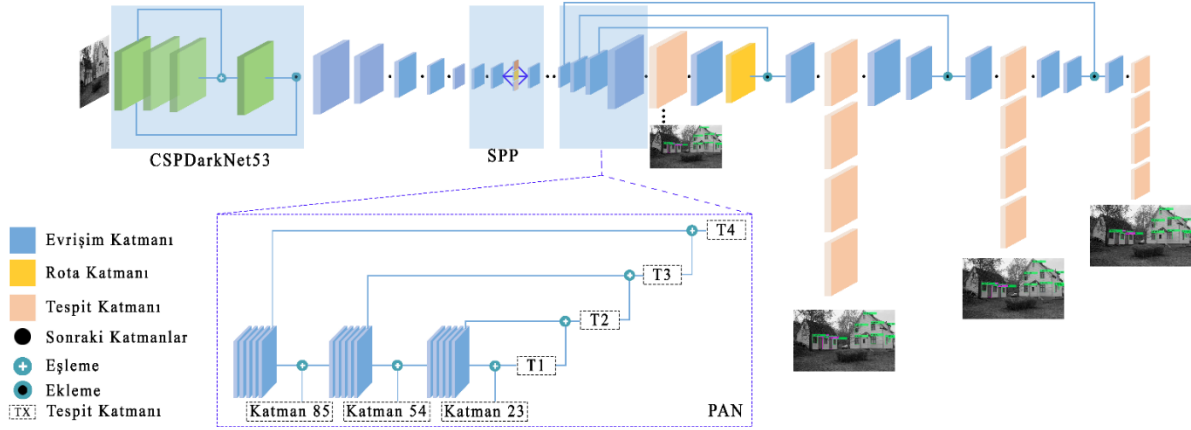
Şekil 2. YOLOv4 mimarisi [17].

3.1. Ağ Yapısının Değişimi

Günümüzün popüler nesne tespit yöntemlerinden YOLOv4'ü inceleyecek olursak, Şekil 2'de görülen bir ağ yapısıyla karşılaşmaktayız. Ağdaki omurga (backbone) ve boyun (neck) kısımlarında CSPDarknet53 ağı, Spatial Pyramid Pooling (SPP) ve Path Aggregation Network (PAN) yöntemlerinden faydalanılmaktadır. CSPDarknet53 ağı, darknet53 üzerine kurulmuş ve böl/birleştir stratejisiyle daha fazla gradyan akışına izin vermektedir. SPP [11] ve PAN kavramı alanı artırmak ve yanıl bağlantılarla özellik haritaları arasında özellik piramidi inşa etmek için kullanılmaktadır. PAN, ilk olarak eleman bazında maksimum işlemi kullanarak tüm katmanlardan gelen bilgileri birleştirir. Daha sonra, kısa yoldan tanıma katmanlarına ulaşarak daha hassas yerleştirilmiş bilgileri üst katmanlar için kullanılabilir hale getirir. Bu çok özellikli ağ yapısına rağmen, YOLOv4, farklı ölçeklerdeki (çok küçük ya da çok büyük) nesnelere sınıflandırmakta güçlük çekmektedir. Bu zorluğu aşmak adına ağ modeline bir eklenti yapılması tasarlanmıştır.

Görüntü üzerinde yer alan küçük boyutlu nesnelere tespit başarısını artırabilmek adına, PAN ve tanıma katmanları üzerinde güncellemelere gidilmiştir. Evrimsel sinir ağı (ESA), düşük seviyeli özellik haritasında zengin anlamsal

bilgiye ve yüksek seviyeli özellik haritasında zengin konum bilgisine sahiptir. Düşük seviyedeki anlam bilgisini artırmak ve nesnelerin görüntü üzerindeki yer bilgisini daha iyi çıkarabilmek amacıyla, giriş katmanındaki özellik çıkarımlarından daha fazla faydalanmayı sağlayacak bir füzyon katmanı PAN'a eklenmiştir. Şekil 3'de görüleceği üzere PAN içerisindeki görüntü verisi, omurgada yer alan 23, 54 ve 85'inci katmanlardaki görüntülerle eşlenerek tanıma katmanlarına sunulmaktadır. Ayrıca, PAN'dan sonra klasik YOLOv4'ten farklı olarak dört farklı ölçekte tespit işlemi gerçekleştirilmektedir. YOLOv4'te en büyük 64x64 (Giriş verisinin 512x512 olması durumunda) olan tespit ölçeğinin çok küçük ve çok büyük nesneleri tanımda yetersiz olabileceği düşünüldükten dolayı iki katı büyüklüğünde, küçük ve büyük nesneleri daha etkinlikle tespit edebilecek ilave bir tanıma katmanı oluşturulmuştur. Böylece daha büyük boyutlu sayısal görüntüler üzerinde de tespit ve sınıflandırma yapılarak standart ağ modeliyle eksik ya da yanlış yerde tespit edilen nesnelerin daha doğru bir şekilde tespit edilmesi amaçlanmıştır.



Şekil 3. Özelleştirilen ağ modeli.

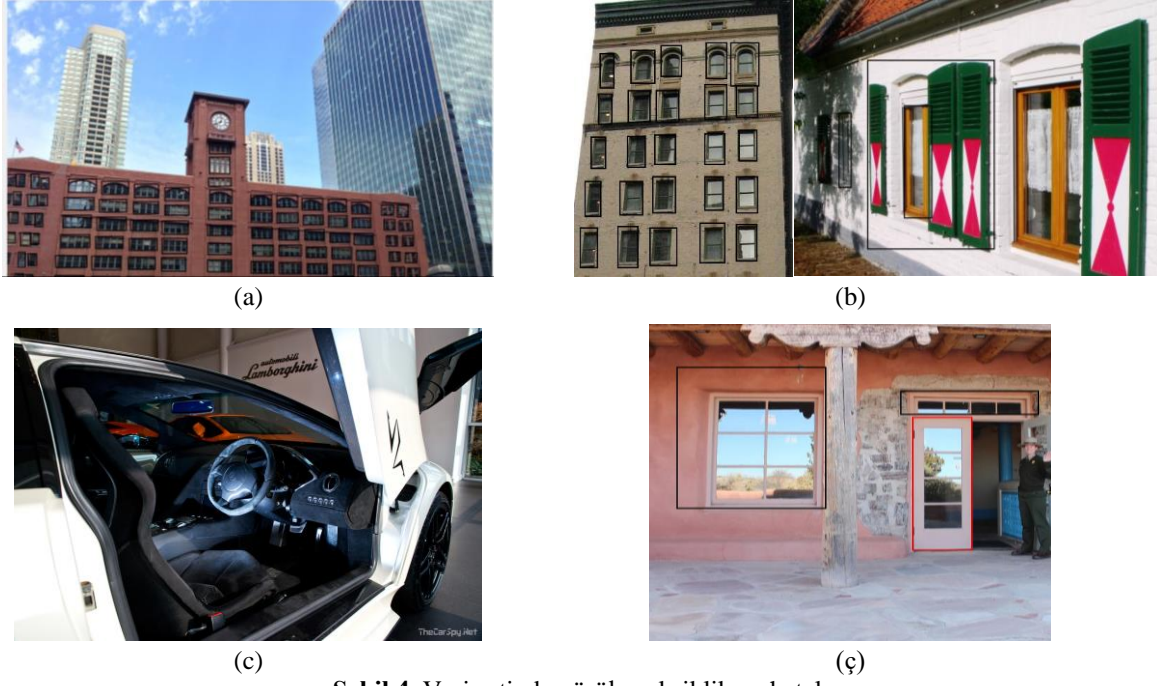
3.2. Veri Setinin Düzenlenmesi

Tespit ağının eğitiminde kullanılan verilerin kalitesi tespit modelinin başarısına doğrudan katkı sağlamaktadır. Veriler, gerçek hayattaki olası tüm varyasyonları içermeli, verilerin aynı korelasyonda olmasına izin verilmemelidir. Örneğin araba sınıflandırması yapan bir modeldeki verilerin birçoğunun otobandan alınan görüntü olması, yapay sinir ağının asfalt arka planla araba arasında bir bağıntı kurmasına sebep olabilir. Bunun yanında, farklı ışık ve netlik koşullarından veri elde edilmesi de değişen hava koşullarında tespit başarısının sağlanmasını kolaylaştırır. Elde edilen verilerin hatalı olup olmadığının bulunması ve bu verilerin temizlenmesiyle de eğitim sürecine katkıda bulunmaktadır.

Çalışma kapsamında kullanılan veri seti, üç farklı veri setine [23-25] ait görüntüleri ve internet kaynaklarından elde edilen görüntüleri içermektedir. Oluşturulan veri setine internet bağlantı adresinden (www.kaggle.com/dataset/7f4acf71b4f8c8c970c6d8f4a5bc8f4c97ab9bc22f00b38cf5985d7847b772a4) erişilebilmektedir. Bu veri setinin düzenlenebilmesi amacıyla, öncelikle veri seti tek tek gözden geçirilmiştir. Veri seti içinde yer alan hatalı veriler, eksik işaretlemeler, işaretlemede standart dışılık ve bazı ortamlara ait veri miktarının çok kısıtlı olması gibi nedenlerle eğitimin olumsuz etkileyeceği değerlendirilmiştir.

Verilerdeki en büyük zorluklardan birisi, farklı ölçeklerdeki kapı ve pencerelerdir. Örneğin, veri setinin orijinal yapısında Şekil 4 (a)'da görülen çok sayıda küçük pencerenin bir kısmı işaretli iken, bir kısmının işaretli olmadığı görülmüştür. Şekil 4 (b)'de yer alan örneklerde, pencereleri çevreleyen kutuların bazı pencereler için geniş bazı pencereler içinse dar çizildiği anlaşılmaktadır. Bu tip standarttan uzak işaretlemeler IoU (Intersection over Union-birleşim kesişimi) hatalarına sebep olabilmekte ve tespit başarısını azaltmaktadır. Farklı kültür mimarilerine ait kapı ve pencere şekilleri, büyük bir çeşitlilikle veri setine aktarılarak gerçek dünya varyasyonları için iyi bir hazırlık yapıyor olsa da Şekil 4 (c)'de yer alan araba kapısı çok ender rastlanan bir kapı şeklidir ve veri seti genelinde de başka örneği çok azdır. Bu yüzden veri setinin eğitimini olumsuz etkileyeceği değerlendirilmiştir.

Şekil 4 (ç)'de yer alan bir pencereye ait çerçeveyici kutu ise nesneden oldukça büyüktür. Veri seti incelendiğinde bu tip örneklerin çok sayıda olduğu ve görüntü üzerine konum bilgisini oluşturan verilerin dönüşüm hatasından kaynaklandığı anlaşılmıştır. Bu sebeplerle nesne sınıflandırmasına olumlu katkı sağlamayacağı ya da mevcut başarının da düşmesine sebep olacağı düşünülen veriler üzerinde eksiltme, artırma ve yeniden işaretleme yöntemleriyle güncellemeye gidilmiştir.



Şekil 4. Veri setinde görülen eksiklik ve hatalar.

4. BULGULAR

Mevcut tanıma ve sınıflandırma probleminin mobil robotlarda kullanılacak olması nedeniyle, test çalışmaları bir mobil geliştirme birimi olan NVIDIA Jetson AGX Xavier üzerinde gerçekleştirilmiş ve yöntem kısmında belirtildiği üzere kapı, pencere ve merdiven görüntülerini içeren bir veri seti üzerinde değerlendirilmiştir. Veri setinde yer alan görüntülerdeki nesnelerin sınıf dağılımları Tablo 1’de verilmiştir.

Tablo 1. Görüntüler üzerindeki örneklerin sınıf dağılımları.

	Kapı	Merdiven	Pencere	Toplam
Örnek Sayısı	3140	1584	9312	14036
Toplam Görüntü	5012			

Mevcut veri setinin orijinal ağ modeli ile ve modelin değişimi sonrasında eğitimi gerçekleştirilmiştir. Veri setinde yer alan 5012 adet görüntünün %10’luk kısmı rassal olarak ayrılarak modellerin başarısının ölçümü için kullanılmıştır. Modellerin eğitim başarısı ve kayıp değeri her 5000 iterasyonda sınanarak, en iyi başarımların ulaşılan 55000 iterasyona kadar gerçekleştirilmiştir. Tablo 2’de görülen karşılaştırma verisi, her bir sınıf için farklı birleşim kesişimi değerlerinde (0.25, 0.50 ve 0.75) ortalama hassasiyet değerlerini (Average Precision-AP), ortalama doğruluk değerlerini (accuracy) genel ortalama hassasiyet değerlerini (Mean Average Precision-MAP), ortalama birleşim kesişimi değerlerini (average Intersection over Union- avg IoU), ve iki farklı yöntemde saniyede işlenebilen çerçeve sayısını (FPS) içermektedir. Daha iyi bir ölçüm işlemi için 5-katlı çapraz doğrulama gerçekleştirilerek sonuçların ortalaması, en düşük ve en yüksek değerleri tabloya yansıtılmıştır. Modeldeki değişiklik tespit başarısını hemen hemen her karşılaştırmada artırarak hassasiyet değerlerinde ortalama %4’lük bir iyileştirme sağlamış, çerçeveleyici kutular yaklaşık %12 oranında daha isabetli bulunmuştur. Bununla birlikte, ağa eklenen katmanların getirdiği hesaplama maliyeti, saniyede işlenebilen çerçeve sayısının (FPS) yaklaşık %8 oranında düşmesine neden olmuştur

Tablo 2. Ağ yapısı özelleştirmesine ait test sonuçları.

	Klasik YOLOv4			Özelleştirilmiş YOLOv4		
	AP@0,25	AP@0,50	AP@0,75	AP@0,25	AP@0,50	AP@0,75
Kap_{lmin}	%52,86	%45,54	%28,65	%59,17	%48,03	%33,18
Kap_{lort}	%58,56	%50,838	%35,04	%64,864	%55,58	%38,93
Kap_{lmaks}	%61,45	%56,15	%44,32	%68,92	%60,54	%46,77
Merdive_{nmin}	%76,23	%75,69	%45,23	%94,31	%88,08	%39,15
Merdive_{nort}	%82,64	%77,97	%49,58	%95,10	%91,45	%56,40
Merdive_{nmaks}	%88,45	%81,12	%53,02	%95,55	%94,78	%86,66
Pencere_{min}	%45,44	%42,16	%17,60	%39,02	%32,53	%18,37
Pencere_{ort}	%56,83	%47,90	%19,15	%50,44	%40,74	%22,37
Pencere_{maks}	%66,24	%51,42	%21,45	%64,75	%49,01	%24,9
mAP	%66,01	%58,90	%39,51	%70,13	%62,59	%39,24
F1 Score	0,55	0,49	0,28	0,61	0,55	0,35
avg iou	%42,99	%43,96	%35,84	%53,95	%56,82	%58,05
FPS	17,6			16,2		

Veri setinde yapılan düzenlemeler sonucunda veri setinde mevcut 5012 görüntüde hatalı verilerin çıkarılması ve yanlış etiketlemelerin düzenlenmesi sonucunda Tablo 3'teki sınıf dağılımları elde edilmiştir.

Tablo 3. Düzenlenen görüntüler üzerindeki örneklerin sınıf dağılımı.

	Kapı	Merdiven	Pencere	Toplam
Örnek Sayısı	3484	1975	13170	18629
Toplam Görüntü	4767			

Düzenlenen veri setindeki görüntülerin %10'luk kısmı sınama için ayrılarak kalan görüntülerle ağız eğitimi gerçekleştirilmiştir. Eğitim sırasında oluşan kayıp değeri ve her 5000 iterasyonda yapılan başarımlarını dikkate alınarak 55000 iterasyona kadar eğitime devam edilmiştir. Daha iyi bir ölçüm işlemi için 5-katlı çapraz doğrulama gerçekleştirilerek eğitim ve test işlemi beş kez tekrar edilmiştir. Eğitim sonunda her sınıf için ortalama hassasiyet, doğruluk ve ortalama birleşim kesişimi değerleri kıyaslanarak, sonuçlar Tablo 4'te gösterilmiştir. Veri setindeki değişimlerin ortalama hassasiyeti yaklaşık %6 artırdığı, nesne konumunun ise yaklaşık %4 oranında daha iyi tespit ettiği anlaşılmıştır.

Tablo 4. Düzenlenen veri setine ait test sonuçları.

	Klasik YOLOv4			Düzenlenen Veri Seti		
	AP@0,25	AP@0,50	AP@0,75	AP@0,25	AP@0,50	AP@0,75
Kap_{lmin}	%52,86	%45,54	%28,65	%61,32	%46,99	%29,09
Kap_{lort}	%58,56	%50,838	%35,04	%63,63	%53,42	%34,22
Kap_{lmaks}	%61,45	%56,15	%44,32	%66,89	%58,08	%38,91
Merdive_{nmin}	%76,23	%75,69	%45,23	%86,43	%82,06	%48,66
Merdive_{nort}	%82,64	%77,97	%49,58	%90,02	%87,29	%64,98
Merdive_{nmaks}	%88,45	%81,12	%53,02	%92,53	%91,38	%81,40
Pencere_{min}	%45,44	%42,16	%17,60	%55,14	%45,69	%17,30
Pencere_{ort}	%56,83	%47,90	%19,15	%63,64	%54,20	%21,44
Pencere_{maks}	%66,24	%51,42	%21,45	%72,70	%62,39	%26,14
mAP	%66,01	%58,90	%39,51	%72,43	%64,97	%40,21
F1 Score	0,55	0,49	0,28	0,54	0,51	0,36
avg iou	%42,99	%43,96	%35,84	%45,31	%48,26	%39,55

5. DEĞERLENDİRME VE SONUÇ

Bu çalışmada, nesne sınıflandırma yöntemlerinin başarısına artırmaya yönelik literatürdeki çalışmalar incelenerek model ve veri odaklı yaklaşımların tespit başarısına olan etkileri değerlendirilmiştir. Son yıllarda çok sayıda çalışmada sınıflandırma modelinin üzerinde değişiklikler yapılıyor olsa da araştırmacılar veri üzerindeki hazırlıkların da önemine vurgu yapmaktadır. Veri ön işleme süreci ve doğru verilerle eğitim yapılmasının da sınıflandırma başarısı üzerinde büyük etkisi vardır. Model odaklı ve veri odaklı yaklaşımların etkilerini incelemek üzere halihazırda kapı, pencere ve merdiven nesnelerini tespit etmekte kullanılan veri seti ve YOLOv4 ağı üzerinde düzenlemelere gidilmiştir. İki yaklaşımla da tespit başarısının artırılabilirdiği gözlemlenirken, model odaklı yaklaşımda ağı büyütülmesi nedeniyle gerçek zamanlı çalışmada hesaplama performansında yaklaşık %8'lik bir yavaşlamaya neden olduğu görülmüştür. Öte yandan veri setinde yer alan 5012 adet görüntüden gereksiz olduğu değerlendirilen 245 adedinin silinmesi, 2348 görüntü üzerindeki örneklere ait işaretlemelerin değişimi (ekleme, silme ve yeniden işaretleme) ile ağa ek yük getirmeden tespit başarısının yaklaşık %6 oranında artırılması sağlanmıştır. Elde edilen sonuçlar, sıklıkla ihmal edilen ya da bu alanda yeterince profesyonel olmayan üçüncü kişilere yaptırılan veri hazırlama sürecinin sınıflandırma başarısına önemli derecede etkisini olduğunu net olarak göstermektedir. Gelecek çalışmalar ile, çok sayıda içerik barındıran ve geniş kullanım alanına sahip olan COCO veri seti üzerinde düzenleme çalışmaları yapılması planlanmaktadır.

Yazar Katkıları

Yazarlar çalışmaya eşit oranlı katkı sunmuşlardır.

Çıkar Çatışması

Makale yazarları aralarında herhangi bir çıkar çatışması olmadığını beyan ederler

KAYNAKÇA

- [1] A. Kumar, Z. J. Zhang, and H. Lyu, "Object detection in real time based on improved single shot multi-box detector algorithm." EURASIP Journal on Wireless Communications and Networking, vol.1, pp. 1-18, 2020.
- [2] Y. Cai, T. Luan, H. Gao, H. Wang, L. Chen, Y. Li, M.Á. Sotelo, and Z. Li, "YOLOv4-5D: An Effective and Efficient Object Detector for Autonomous Driving." IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-13, 2021.
- [3] S. Zhai, D. Shang, S. Wang, and S. Dong, "DF-SSD: An improved SSD object detection algorithm based on DenseNet and feature fusion", IEEE Access, vol. 8, pp. 24344-24357, 2020.
- [4] C. G. Northcutt, A. Athalye, and J. Mueller, "Pervasive label errors in test sets destabilize machine learning benchmarks", arXiv preprint arXiv:2103.14749, 2021.
- [5] K. Li, W. Ma, U. Sajid, Y. Wu, and G. Wang, "Object detection with convolutional neural networks", In Deep Learning in Computer Vision, pp. 41-62, 2020.
- [6] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 2017.
- [7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified Real-Time Object Detection", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788, 2016.
- [8] W. Liu, D. Anguelov, D. Erhan et al., "SSD: Single Shot MultiBox Detector", Lecture Notes in Computer Science, vol. 9905, pp. 21-37, 2016.
- [9] C. Y. Fu, W. Liu, A. Ranga, et al., "DSSD: Deconvolutional Single Shot Detector", arXiv, Computer Vision and Pattern Recognition, 2017.
- [10] T. Y. Lin, P. Goyal, R. Girshick, et al., "Focal Loss for Dense Object Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, no. 99, pp. 2999-3007, 2017.
- [11] P. Purkait, C. Zhao and C. Zach, "SPP-Net: Deep absolute pose regression with synthetic views", arXiv preprint arXiv:1712.03452, 2017.
- [12] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning. Nature", vol. 521 no. 7553, pp. 436-444, 2015.
- [13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks", Advances in neural information processing systems, vol. 25, pp. 1097-1105, 2012.
- [14] S. S. Shanta, S. T. Anwar, and M. R. Kabir, "Bangla sign language detection using sift and CNN", In 2018 9th international conference on computing, communication and networking technologies (ICCCNT), pp. 1-6, 2018.
- [15] B. T. Hung, "Face recognition using hybrid HOG-CNN approach", In Research in Intelligent and Computing in Engineering, pp. 715-723, 2021.
- [16] Y. Tian, G. Yang, Z. Wang, H. Wang, E. Li, and Z. Liang, "Apple detection during different growth stages in orchards using the improved YOLO-V3 model", Computers and electronics in agriculture, vol. 157, pp. 417-426, 2019.
- [17] A. Bochkovskiy, W. Chien-Yao, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv preprint arXiv:2004.10934, 2020.

- [18] Ultralytics, "YOLOv5". Available: <https://github.com/ultralytics/yolov5>. (Erişim Tarihi: 02.01.2022)
- [19] S. E. Whang, Y. Roh, H. Song, and J. G. Lee, "Data Collection and Quality Challenges in Deep Learning: A Data-Centric AI Perspective", arXiv preprint arXiv:2112.06409, 2021.
- [20] Y. Lin, Y. Guan, A. Asudeh, H. V. Jagadish, "Identifying insufficient data coverage in databases with multiple relations. Proceedings of the VLDB Endowment", vol. 13, no. 11, 2020.
- [21] Ö. Alparslan, Ö. Çetin, "Comparison of Object Detection and Classification Methods For Mobile Robots", Sakarya University Journal of Science, vol. 25, no. 3, pp. 764-778, 2021.
- [22] S. Manzoor, S. H. Joo, and T. Y. Kuc, "Comparison of object recognition approaches using traditional machine vision and modern deep learning techniques for mobile robot", 19th International Conference on Control, Automation and Systems (ICCAS), pp. 1316-1321, 2019.
- [23] A. Kuznetsova, H. Rom, N. Alldrin, J. Uijlings, I. Krasin, J. PontTuset, et al. "The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale", International Journal of Computer Vision, vol. 128, pp. 1956-1981, 2020.
- [24] M. Arduengo, T. Carme and L. Sentis, "Robust and adaptive door operation with a mobile manipulator robot", arXiv preprint arXiv:1902.09051, 2019.
- [25] U. Patil et al., "Deep Learning Based Stair Detection and Statistical Image Filtering for Autonomous Stair Climbing", 2019 Third IEEE International Conference on Robotic Computing (IRC), pp. 159-166, 2019.