

T.C.
ERCIYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

UZAKTAN ALGILAMA GÖRÜNTÜLERİNDE NESNE
TESPİTİ VE SAYIMI İÇİN EVRİŞİMLİ SINIR AĞLARI
TABANLI YENİ MODELLERİN GELİŞTİRİLMESİ

Hazırlayan
Ersin KILIÇ

Danışman
Doç. Dr. Serkan ÖZTÜRK

Doktora Tezi

TEMMUZ 2024
KAYSERİ

T.C.
ERCIYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

UZAKTAN ALGILAMA GÖRÜNTÜLERİNDE NESNE
TESPİTİ VE SAYIMI İÇİN EVRİŞİMLİ SINIR AĞLARI
TABANLI YENİ MODELLERİN GELİŞTİRİLMESİ
(Doktora Tezi)

Hazırlayan
Ersin KILIÇ

Danışman
Doç. Dr. Serkan ÖZTÜRK

Bu çalışma Erciyes Üniversitesi Bilimsel Araştırma Projeleri Birimi
tarafından FDK-2018-8624 numaralı proje ile desteklenmiştir.

TEMMUZ 2024
KAYSERİ

BİLİMSEL ETİĞE UYGUNLUK

Bu çalışmadaki tüm bilgilerin, akademik ve etik kurallara uygun bir şekilde elde edildiğini beyan ederim. Aynı zamanda bu kural ve davranışların gerektirdiği gibi, bu çalışmanın özünde olmayan tüm materyal ve sonuçları tam olarak aktardığımı ve referans gösterdiğimi belirtirim.

Ersin KILIÇ

“Uzaktan Algılama Görüntülerinde Nesne Tespiti ve Sayımı için Evrişimli Sinir Ağları Tabanlı Yeni Modellerin Geliştirilmesi” adlı Doktora Tezi, Erciyes Üniversitesi Lisansüstü Tez Önerisi ve Tez Yazma Yönergesi’ne uygun olarak hazırlanmıştır.

Hazırlayan
Ersin KILIÇ

Danışman
Doç. Dr. Serkan ÖZTÜRK



Bilgisayar Mühendisliği ABD Başkanı
Prof. Dr. Veysel ASLANTAŞ

TEŞEKKÜR

"Uzaktan Algılama Görüntülerinde Evrişimli Sinir Ağları İle Nesne Tespiti ve Sayımında Yeni Yaklaşımlar" konulu tez çalışmamın gerçekleştirilmesinde desteklerini esirgemeyen danışman hocam Doç. Dr. Serkan ÖZTÜRK'e teşekkürü bir borç bilirim. Tez süresince çalışmalarımı izleyen ve katkıda bulunan saygıdeğer hocalarım Doç. Dr. Mete ÇELİK ve Doç. Dr. Nurettin ÜSTKOYUNCU'ya teşekkür ederim.

Başta Ahmet Nusret TOPRAK, Ömür ŞAHİN ve Tayyip ÖZCAN olmak üzere tüm değerli çalışma arkadaşlarıma teşekkürlerimi sunarım.

Hayatım boyunca desteğini ve sevgisini esirgemeyen babam Galip KILIÇ başta olmak üzere, annem Nebahat KILIÇ ve ablam Esra KILIÇ'a tüm kalbimle teşekkür ederim.

Başarıma inancı hep var olan daimi yoldaşım Eda KILIÇ'a ve hayata sınıksız sarılmamı, her türlü zorlukla mücadelede güçlü olmamı sağlayan yaşam kaynağım canım kızlarım Mihra ve İlve'ye sonsuz teşekkür ederim.

UZAKTAN ALGILAMA GÖRÜNTÜLERİNDE NESNE TESPİTİ VE SAYIMI İÇİN EVRİŞİMLİ SINIR AĞLARI TABANLI YENİ MODELLERİN GELİŞTİRİLMESİ

Ersin KILIÇ

Erciyes Üniversitesi, Fen Bilimleri Enstitüsü

Doktora Tezi, TEMMUZ 2024

Danışman : Doç. Dr. Serkan ÖZTÜRK

ÖZET

Uydu ve insansız hava aracı teknolojisindeki gelişmeler uzaktan algılayıcılardan yüksek miktarda görüntü elde edilmesini sağlamış ve bu görüntülerin analizi bir çok alanda olumlu gelişmelerin önünü açmıştır. Son zamanlarda uzaktan algılama görüntülerinde nesne tespiti ve sayımı bu görüntülerin analizinde popüler hale gelmiştir. Bu tez çalışmasında uzaktan algılama görüntülerinde nesne tespiti ve sayımına yönelik Alt Sınıf Destekli Evrişimli Sinir Ağı (SSCNN), Isı Haritası Öğrenen Evrişimli Sinir Ağı (HLCNN) ve CNessNet adları verilen üç yeni model önerilmiştir. SSCNN, nesnelere daha iyi temsil eden ayırt edici özellikler çıkarımını sağlamak için nesnelere alt sınıflara ayırarak temsil eden bir eğitim stratejisi içeren yeni bir modeldir. Yapılan deneysel çalışmalarda en uygun alt sınıf sayısı belirlenmiş olup, nesne tespiti başarımları incelenmiştir. HLCNN, ısı haritası bağlantısını gerçekleştiren tek atışlı evrişimli sinir ağı modelidir. Önerilen modelin insansız hava aracı görüntülerinde araba sayımı başarımlarını değerlendirmek için literatürde sıklıkla kullanılan CARPK ve PUCPR+ veri setleri üzerinde deneysel çalışmalar yapılmıştır. Önerilen model için önemli bir parametre olan alt örnekleme oranının başarıma etkisi incelenmiştir. Ayrıca, önerilen modelin düşük maliyetli işaretleme yöntemi ile eğitilmesini sağlayan bir model tanıtılmıştır ve başarımları CARPK ve PUCPR+ veri setleri üzerinde değerlendirilmiştir. CNessNet, insansız hava araçları görüntülerinden araç tespiti gerçekleştirilen evrişimli sinir tabanlı yeni bir modeldir. Önce omurga ağ seçiminin önerilen modelin başarımlarına etkisi CARPK veri seti üzerinde değerlendirilmiştir. Daha sonra omurga ağı olarak Kum Saati ağları içeren modelin çok sınıflı araç tespiti başarımları VAID veri seti ile değerlendirilmiştir.

Anahtar Kelimeler: Derin öğrenme, Evrişimli Sinir Ağları, Nesne tespiti, Görüntü işleme, Uzaktan algılama, Araba tespiti, Araba sayımı, Araç tespiti.

DEVELOPMENT OF NEW MODELS BASED ON CONVOLUTIONAL NEURAL NETWORK FOR OBJECT DETECTION AND COUNTING IN REMOTE SENSING IMAGES

Ersin KILIÇ

Erciyes University, Graduate School of Natural and Applied Sciences

Master Thesis, JULY 2024

Supervisor: Assoc. Prof. Dr. Serkan ÖZTÜRK

ABSTRACT

Advancements in satellite and unmanned aerial vehicle technologies have enabled the acquisition of large volumes of remote sensing images, and the analysis of these images has facilitated significant progress across various fields. Recently, object detection and counting in remote sensing images have become popular topics in image analysis. This thesis proposes three new models for object detection and counting in remote sensing images: Subclass Supported Convolutional Neural Network (SSCNN), Heatmap Learning Convolutional Neural Network (HLCNN), and CNessNet. SSCNN is a new model that includes a training strategy where objects are divided into subclasses to extract distinguishing features that better represent the objects. Experimental studies have determined the optimal number of subclasses and examined object detection performance. HLCNN is a single-shot convolutional neural network model that performs heatmap regression. Experimental studies have been conducted on CARPK and PUCPR+ datasets, which are frequently used in the literature, to evaluate the performance of the proposed model in counting cars on unmanned aerial vehicle images. The effect of the subsampling rate which is an important parameter for the proposed model on the performance has been investigated. Additionally, a method for training the proposed model using a low-cost annotation technique has been introduced and its effectiveness has been evaluated on the CARPK and PUCPR+ datasets. CNessNet is a new convolutional neural network-based model for vehicle detection from UAV images. First, the effect of the backbone network selection on the performance of the proposed model has been evaluated on the CARPK dataset. Then, the method incorporating Hourglass Networks as the backbone network was evaluated for multi-class vehicle detection effectiveness using the VAID dataset.

Keywords: Deep learning, Convolutional neural network, Object detection, Image processing, Remote sensing, Car detection, Car counting, Vehicle detection.

İÇİNDEKİLER

UZAKTAN ALGILAMA GÖRÜNTÜLERİNDE NESNE TESPİTİ VE SAYIMI İÇİN EVRİŞİMLİ SİNİR AĞLARI TABANLI YENİ MODELLERİN GELİŞTİRİLMESİ

BİLİMSEL ETİĞE UYGUNLUK SAYFASI	ii
YÖNERGEYE UYGUNLUK SAYFASI	iii
KABUL VE ONAY	iv
TEŞEKKÜR	v
ÖZET	vi
ABSTRACT	vii
İÇİNDEKİLER	viii
KISALTMALAR	xi
TABLolar LİSTESİ	xiii
ŞEKİLLER LİSTESİ	xiv
GİRİŞ	1

1. BÖLÜM

GENEL BİLGİLER VE LİTERATÜR ÇALIŞMASI

1.1. Evrişimli Sinir Ağları	6
1.1.1. Evrişim Katmanı	7
1.1.2. Ortaklama Katmanı	8
1.1.3. Yiğın Normalizasyonu	9
1.1.4. Transfer Öğrenme	9
1.2. CNN Tabanlı Nesne Tespiti Modelleri	10
1.2.1. Çapa Tabanlı Nesne Tespiti Modelleri	10
1.2.1.1. Tek Atış Dedektör	10
1.2.1.2. Faster R-CNN	12
1.2.1.3. RetinaNet	14
1.2.2. Çapasız Nesne Tespiti Modelleri	17

1.2.2.1. CornerNet	17
1.2.2.2. CenterNet	17
1.3. Uzaktan Algılama Görüntülerinde Nesne Tespiti ve Sayımı	19
1.3.1. Uzaktan Algılama Görüntülerinde Nesne Tespiti Üzerine Yapılan Çalışmalar	19
1.3.2. İHA Görüntülerinden Araç Sayımı ve Tespiti Üzerine Yapılan Çalışmalar	25

2. BÖLÜM

ALT SINIF DESTEKLİ EVRİŞİMLİ SİNİR AĞLARI (SSCNN)

2.1. Önerilen SSCNN Modeli	30
2.1.1. SSCNN Mimarisi ve Veri Hazırlığı	31
2.1.2. Lokalizasyon Değerlendirme Fonksiyonu (LEF)	32
2.2. Deneysel Çalışmalar	33
2.2.1. Veri Setleri	33
2.2.2. Değerlendirme Kriterleri	34
2.2.3. Alt Sınıf Sayısının Belirlenmesi	34
2.2.4. SSCNN ve R-CNN Modellerinin Karşılaştırılması	36
2.2.5. SSCNN Özelliklerinin Etkisinin Belirlenmesi	45
2.3. Değerlendirme	46

3. BÖLÜM

İNSANSIZ HAVA ARACI GÖRÜNTÜLERİNDE ARABA SAYIMI VE TESPİTİ

3.1. Isı Haritası Öğrenen CNN (HLCNN) ile Araç Sayımı ve Tespiti	48
3.1.1. Önerilen HLCNN Modeli	50
3.1.1.1. Tepe Nokta Haritası Oluşturma Algoritması	52
3.1.2. Deneysel Çalışmalar	52
3.1.2.1. Veri Setleri	54
3.1.2.2. Değerlendirme Kriterleri	56
3.1.2.3. CARPK Veri Seti Sonuçları	56
3.1.2.4. PUCPR+ Veri Seti Sonuçları	62
3.1.3. Değerlendirme	65

3.2. Merkez Nokta İşaretleme Tabanlı HLCNN ile Araç Sayımı ve Tespiti . . .	66
3.2.1. Önerilen Merkez Nokta İşaretleme Tabanlı HLCNN Modeli . . .	67
3.2.1.1. Merkez Nokta İşaretleme Tabanlı HLCNN Modelinin Eğitimi	67
3.2.2. Deneysel Çalışmalar	68
3.2.3. Değerlendirme	71

4. BÖLÜM

İNSANSIZ HAVA ARACI GÖRÜNTÜLERİNDE ÇOK SINIFLI ARAÇ TESPİTİ

4.1. Önerilen CNessNet Modeli	72
4.1.1. Merkeziyet Skoru ile Haritalama Yöntemi	74
4.1.2. CNessNet Modelinin Eğitimi ve Kayıp Fonksiyonu	75
4.2. Deneysel Çalışmalar	78
4.2.1. Veri Setleri	79
4.2.2. Araba Tespiti	79
4.2.3. Çok Sınıflı Araç Tespiti	81
4.3. Değerlendirme	82

5. BÖLÜM

TARTIŞMA, SONUÇ ve ÖNERİLER

5.1. Tartışma ve Sonuçlar	83
5.2. Öneriler	85
KAYNAKÇA	86
ÖZGEÇMİŞ	99

KISALTMALAR

CNN	:	Evrişimli sinir ağı
HOG	:	Yönlü gradyanlar histogramı
SSD	:	Tek Atışlı Dedektör
RetinaNet	:	Retina Ağı
R-CNN	:	Bölgesel Evrişimli Sinir Ağı
Faster R-CNN	:	Daha Hızlı Bölgesel Evrişimli Sinir Ağı
RPN	:	Bölge Teklif Ağı
SS	:	Seçici Arama
TL	:	Transfer Öğrenme
Fast R-CNN	:	Hızlı Bölgesel Evrişimli Sinir Ağı
FPN	:	Özellik Piramit Ağı
CornerNet	:	Köşe Ağı
CenterNet	:	Merkez Ağı
FCOS	:	Tam Evrişimli Tek-aşamalı Nesne Tespiti Yöntemi
RICNN	:	Rotasyon Bağımsız Evrişimli Sinir Ağı
NMS	:	Maksimum Olmayanı Baskılama
USB-BBR	:	Denetimsiz Puan Tabanlı Sınırlayıcı Kutu Regresyon
PSB	:	Konuma Duyarlı Dengeleme
arcNMS	:	En-boy Oranı Kısıtlı Maksimum Olmayanı Baskılama
HRPN	:	Hiper Bölge Teklif Ağı
VC-CNN	:	Araç-sayar Evrişimli Sinir Ağı
GANet	:	Rehberli Dikkat Ağı
EB	:	Kenar Kutuları
LEF	:	Lokalizasyon değerlendirme fonksiyonu
SSCNN	:	Alt-sınıf Destekli Evrişimli Sinir Ağı
AP	:	Ortalama Hassasiyet
mAP	:	Ortalama Hassasiyetlerin Ortalaması
TP	:	Doğru Pozitif
FP	:	Yanlış Pozitif
FN	:	Yanlış Negatif

HLCNN	:	Isı Haritası Öğrenen Evrişimli Sinir Ağı
GAP	:	Global Ortalamaya Ortaklama
GAM	:	Gaussian Aktivasyon Haritası
BN	:	Yığın Normalizasyonu
MAE	:	Ortalama Mutlak Hata
RMSE	:	Kök Ortalama Kare Hatası



TABLOLAR LİSTESİ

Tablo 2.1.	ERUCV-RS4 veri seti için SSCNN'nin farklı alt sınıf modellerinin AP metrik sonuçları	36
Tablo 2.2.	Eğitilen SSCNN modellerinin parametreleri	37
Tablo 2.3.	Eğitilen R-CNN modellerinin parametreleri	37
Tablo 2.4.	SS algoritmasının parametreleri	37
Tablo 2.5.	ERUCV-RS4 veri seti için SSCNN ve R-CNN modellerinin AP metrik sonuçları	40
Tablo 2.6.	NWPU-VHR10 veri seti için SSCNN ve farklı modellerin AP metrik sonuçları	46
Tablo 3.1.	HLCNN modelinin alt örnekleme oranının CARPK veri setinde araba sayımı başarımına etkisi	57
Tablo 3.2.	Önerilen HLCNN modelinin için veri artırmanın ve BN katmanının başarım CARPK veri setinde araba sayımı başarımına etkisi	57
Tablo 3.3.	CARPK veri seti için önerilen HLCNN modelinin karşılaştırma sonuçları	63
Tablo 3.4.	HLCNN modelinin alt örnekleme oranının PUCPR+ veri setinde araba sayımı başarımına etkisi	63
Tablo 3.5.	PUCPR+ veri seti için HLCNN modelinin karşılaştırma sonuçları	64
Tablo 3.6.	CARPK veri seti için önerilen merkez nokta işaretleme tabanlı HLCNN modelinin karşılaştırma sonuçları	69
Tablo 4.1.	VAID veri setinde kategori başına araç sayıları	79
Tablo 4.2.	CARPK veri seti için CNessNet modellerinin karşılaştırma sonuçları	79
Tablo 4.3.	VAID veri seti için sınıf ağırlıkları	81
Tablo 4.4.	VAID veri seti için önerilen CNessNet-4 modelinin karşılaştırma sonuçları	81

ŞEKİLLER LİSTESİ

Şekil 1.1.	2 boyutlu veri üzerinde evrişim işlemi örneği	8
Şekil 1.2.	En büyüğe ortaklama örneği	9
Şekil 1.3.	SSD mimarisi	11
Şekil 1.4.	SSD mimarisinin farklı boyutlarında özellik haritalarında çapaların temsili yerleşimi (a) temsili sonuç görüntüsü, (b) 8×8 , çapa yerleşimi (c) 4×4 çapa yerleşimi	12
Şekil 1.5.	R-CNN genel blok şeması (a) giriş görüntüsü, (b) SS ile üretilen aday nesne önerileri, (c) CNN özellik çıkarımı, (d) bölgelerin sınıflandırılması [1]	13
Şekil 1.6.	Fast R-CNN genel blok şeması [2]	13
Şekil 1.7.	Faster R-CNN genel blok şeması [3]	14
Şekil 1.8.	Faster R-CNN çapaları [3]	14
Şekil 1.9.	Odak kayıp fonksiyonu [4]	15
Şekil 1.10.	RetinaNet mimarisi (a) ResNet, (b) Özellik Piramit ağı, (c) sınıflandırıcı ve çerçeve regresyonu [4]	16
Şekil 1.11.	CornerNet nesne sınırlayıcı kutu gösterimi [5]	18
Şekil 1.12.	CornerNet mimarisi [5]	18
Şekil 1.13.	CenterNet mimarisi [6]	19
Şekil 2.1.	Önerilen SSCNN modelinin genel blok şeması (a) test görüntüsü, (b) SS algoritması ile üretilen aday nesne önerileri, (c) SSCNN ile özellik çıkarımı ve sınıflandırma skorlarının hesaplanması, (d) alt sınıfları ile birlikte nesne sınıfları, (e) LEF, (f) NMS, kesişen nesne bölgelerini elimine edilmesi, (g) Tespit edilen nesnelere	30
Şekil 2.2.	SSCNN mimarisi ve eğitim şeması	32
Şekil 2.3.	Pozitif görüntü örnekleri için nesne alt bölgeleri	32

Şekil 2.4.	0.5 ve 0.7 IoU değerleri için örnek sınırlayıcı kutu örtüşme gösterimleri [7]	35
Şekil 2.5.	SSCNN-1 için görsel nesne tespiti sonuçları (a) uçak ve depolama tankı, (b) uçak, (c) futbol sahası ve beyzbol sahası, (d) depolama tankı, (e) beyzbol sahası, (f) futbol sahası ve beyzbol sahası, (g) uçak, (h) futbol sahası ve beyzbol sahası, (i) depolama tankı	38
Şekil 2.6.	SSCNN-1 modelinin conv5 katmanının seçilen özellik haritası ve giriş görüntüsü çiftleri (a) futbol sahası, (b) beyzbol sahası, (c) depolama tankı, (d) uçak	39
Şekil 2.7.	Eğitilen modellerin uçak PRC değerleri (a) IoU=0.3, (b) IoU=0.5, (c) IoU=0.6, (d) IoU=0.7	41
Şekil 2.8.	Eğitilen modellerin beyzbol sahası PRC değerleri (a) IoU=0.3, (b) IoU=0.5, (c) IoU=0.6, (d) IoU=0.7	42
Şekil 2.9.	Eğitilen modellerin futbol sahası PRC değerleri (a) IoU=0.3, (b) IoU=0.5, (c) IoU=0.6, (d) IoU=0.7	43
Şekil 2.10.	Eğitilen modellerin depolama tankları PRC değerleri (a) IoU=0.3, (b) IoU=0.5, (c) IoU=0.6, (d) IoU=0.7	44
Şekil 2.11.	Eğitilen Faster R-CNN modellerinin her bir nesne için Duyarlılık değerleri	46
Şekil 3.1.	Önerilen HLCNN modelinin genel blok şeması	51
Şekil 3.2.	Önerilen HLCNN modelinin mimarisi	51
Şekil 3.3.	CARPK veri setinden örnek görüntü [8]	55
Şekil 3.4.	Renk tabanlı veri artırma işleminin etkisini gözlemlemek amacıyla HLCNN-WOBN modeli için tahmin edilen ısı haritaları (a), (b) Örnek görüntüler, (c), (d) ısı haritaları	58
Şekil 3.5.	Renk tabanlı veri artırma işleminin etkisini gözlemlemek amacıyla HLCNN modeli için tahmin edilen ısı haritaları (a), (b) Örnek görüntüler, (c), (d) ısı haritaları	59
Şekil 3.6.	HLCNN-WOBNAU modeli için tahmin edilen ısı haritaları (a), (b) Farklı örnek görüntüler, (c), (d) ısı haritaları	60

Şekil 3.7.	HLCNN modeli için tahmin edilen ısı haritaları (a), (b) Farklı örnek görüntüler, (c), (d) ısı haritaları	61
Şekil 3.8.	CARPK veri seti için HLCNN modeli ile elde edilen sonuçlar , (a), (b) Örnek görüntüler, (c), (d) ısı haritaları	62
Şekil 3.9.	PUCPR+ veri seti için HLCNN model ile elde edilen sonuçlar (a), (b) Örnek görüntüler, (c), (d) ısı haritaları	64
Şekil 3.10.	Farklı görüntü işaretleme yöntemleri örnekleri (a) araba, (b) araba [merkez nokta], (c) araba (çerçeve), (d) araba (piksel)	66
Şekil 3.11.	Önerilen merkez işaretleme tabanlı HLCNN modeli ile elde edilen görsel sonuçlar	70
Şekil 4.1.	CNessNet modelinin genel blok şeması	73
Şekil 4.2.	Merkeziyet skoru ile oluşturulmuş örnek harita görseli [9]	74
Şekil 4.3.	Regresyon katmanı için örnek kesin referans haritası (a) Temsili görüntü ve nesne, (b) X eksenini için kesin referans değerleri, (c) Y eksenini için kesin referans değerleri	76
Şekil 4.4.	CARPK veri seti için CNessNet-2 modeli ile elde edilen sonuçlar	80
Şekil 4.5.	VAID veri seti için CNessNet-4 modeli ile elde edilen sonuçlar	82

GİRİŞ

İnsansız hava aracı (İHA), uydu ve sensör teknolojilerinde yaşanan gelişmeler sayesinde yüksek çözünürlüklü uzaktan algılama görüntülerinin kalitesi ve erişebilirliği gün geçtikçe artmaktadır. Bu görüntüler yeryüzü hakkında çok fazla bilgi taşıması sayesinde sivil ve askeri bir çok uygulamada kullanılmaktadır. Uzaktan algılama görüntülerinin analizi, tarım, ormancılık, şehirleşme, biyoçeşitlilik, haritalama ve benzeri bir çok farklı alanda yeni uygulamaların ortaya çıkmasını sağlamaktadır [10]. Askeri alanda taktiksel analiz, hedef tespiti, otomatik gözetim ve keşif önemli uygulama örnekleridir. Tarımsal alanda ekili alanların tespiti, ürün çeşit analizi, üretim tahminlemesi, ürün sağlığı tespiti ve ürün kalite analizi gibi uygulamalar bulunmaktadır. Kaçak yapıların tespiti, yeni yerleşim alanların tespiti ve planlanması şehircilik alanında uygulamalardır. Afet sonrası oluşan hasar tespiti, afet riski bulunan alanların tespiti ve afet müdahale planı hazırlığı gibi uygulamalar afet yönetimi için önemlidir. Çevre izleme alanında bölgesel tür envanteri çıkarımı, su havzası belirlenmesi ve obruk tespiti gibi uygulamalar bulunmaktadır. Coğrafi tehlike haritalaması, litolojik haritalama, maden arama ve sedimantasyon haritalama gibi jeoloji uygulamaları bulunmaktadır.

Yukarıda bahsedilen alanlar, uzaktan algılama görüntülerinin yoğunlukla kullanıldığı uygulama alanları olarak kabul görmektedir. Ayrıca, İHA teknolojisindeki gelişmeler, şehir yönetimi, trafik izleme [11], çiftlik yönetimi ve park yeri kullanımı [12] gibi havadan görüş kameraları için yeni potansiyel uygulama alanları ortaya çıkarmıştır. Genel olarak, bu uygulamalar farklı nesnelerin sayısı ve konumları ile ilgilenmektedir. Özellikle şehir yönetimi ve trafik izleme uygulamalarında yoldaki veya park alanlarındaki arabaların [8, 13–16] sayısının tespit edilmesi oldukça önemlidir.

Uzaktan algılama görüntülerinde nesne tespiti, bahsedilen tüm uygulamalarda öncelikli gereksinimlerden birisidir. Nesnelerin görünüm farklılıkları, arka plan düzensizlikleri, aydınlatma zayıflığı ve gölgeler nedeniyle görüntüde bulunan nesnelerin tespiti genellikle

en zorlayıcı, fakat bu uygulamaların en temel parçasıdır. Görüntüden anlamlı bilgi çıkarılması için görüntüde bulunan nesnelerin tespiti önemlidir. Bu nedenle bir çok araştırmacının uzun yılladır ilgisini çekmektedir. Basit arkaplanlarda net olarak ayırt edilebilen nesnelerin tespiti üzerine bir çok başarılı yöntem geliştirilmiştir [17]. Fakat, uzaktan algılama görüntülerinde arkaplan oldukça karmaşıktır. Ayrıca sensör gürültüsü, hava durumu, aydınlatma gibi bir çok faktör uzaktan algılama görüntülerinde nesne tanımayı daha da zorlaştırmaktadır [18]. Landsat gibi mekansal çözünürlük değeri düşük olan uydu görüntüleri nesnelerin tespitini mümkün kılmamaktadır. Bu sebeple Landsat görüntülerini kullanan araştırmacılar sadece bölgesel özelliklerin tespiti gibi çalışmalar yapmıştır. Özellikle Quickbird ve IKONOS gibi yüksek çözünürlüklü uydular insan yapımı ve doğal nesnelerin tespitini mümkün hale getirmiştir.

Uzaktan algılama görüntülerinin uzamsal çözünürlüğünün düşük olması ve nesnelerin özelliklerinden dolayı bir çok etken nesne tespitini zorlaştırmaktadır. Bu faktörler görüş açısı farklılıkları, ışıklandırma koşulları, ölçek farklılıkları, deformasyon, üstüste geçme ve sınıf içi görünüş farklılığıdır. 3 boyutlu uzaydaki bir nesnenin 2 boyutlu görüntü uzayına izdüşümünden dolayı aynı nesne görüş açısına göre farklı görülebilir. Işıklandırma koşulları önemli bir etkidir. Siyah bir nesne aşırı ışıpta beyaz, beyaz bir nesne ışığın az olduğu ortamda siyah görünebilir. Ölçek farklılıkları sebebiyle objektife olan uzaklığından dolayı aynı boyutlardaki iki nesne farklı boyutlarda görüntülenebilir. Nesnenin deformasyona uğramış olması nesneyi farklılaştırmaktadır. Nesnelerin üst üste geçmesi ile sahnede nesnenin belirli bölümleri gözükmebilir. Arkaplanın karmaşıklığı özellikle nesne ile aynı desene sahip olması önemli bir dezavantajdır. Sınıf içi görünüş farklılığı önemli bir etmendir. Nesnenin birden fazla şekilde olması ve bu örneklerin veri setinde eşit sayıda bulunmaması nesne tespiti yöntemlerinin başarımını etkilemektedir.

Tezin Motivasyonu

Nesne sayımı, bir görüntüdeki belirli bir nesnenin veya nesnelerin sayısını otomatik olarak belirleme işlemidir. Nesne tespiti, görüntüde belirli bir sınıftaki nesnelere algılamakla birlikte nesnenin görüntü üzerinde koordinatlarının da bulunmasını içerir. Bulunan koordinatlar sayesinde nesnenin bir çerçeve ile içine alınacağı alan da tespit edilmiş olur. Nesne sayımı ve tespiti için birçok yöntem geliştirilmiştir. Bu

yöntemler, geleneksel görüntü işleme tekniklerinden derin öğrenmeye kadar çeşitli teknikleri kullanmaktadır. Son yıllarda derin öğrenme yöntemlerinin görüntü analizi alanında göstermiş olduğu başarı, nesne sayımı ve tespiti alanında da olumlu gelişmeleri tetiklemiştir. Özellikle, evrişimli sinir ağları (Convolutional neural network, CNN) tabanlı yöntemler bu alanda yaygın olarak kullanılmakta ve yüksek başarımlarına ulaşan yöntemler olarak ön plana çıkmaktadırlar.

Derin öğrenme tabanlı nesne sayımı ve tespiti araştırmalarında hedef genellikle hesaplama karmaşıklığı düşük ve doğruluk oranı yüksek yöntemlerin geliştirilmesi olmuştur. Ayrıca son yıllarda, veri hazırlama sürecini kolaylaştırmak, eğitim süresini kısaltmak ve yöntemlerin karar verme süreçlerini açıklanabilir hale getirmek gibi hedeflere ulaşmak için çalışmalar yürütülmektedir. Bu hedeflere ulaşmak için genelleştirme yeteneği yüksek karmaşıklığı düşük mimariler geliştirmeye ihtiyaç duyulmaktadır. Son yıllarda bu mimarileri geliştirmek için yeni katman türleri öneren ve veri gösterimini farklı şekilde ele alan yenilikçi çalışmalar bulunmaktadır. Bahsedilen tüm hedeflere ulaşmak için bu çalışmalara ek olarak efektif kayıp fonksiyonlarının ve yeni eğitim stratejilerinin araştırılması gerekmektedir.

Tezin Amacı ve Katkıları

Bu tezin amacı, uzaktan algılama görüntülerinde nesne sayımı ve tespiti yöntemlerinin geliştirilmesine katkıda bulunmaktır. Tez çalışması kapsamında uzaktan algılama görüntülerinde nesnelere üç alt sınıfta temsil eden yeni bir nesne tespiti yöntemi, araba sayımı için yeni bir CNN mimarisi, araba sayımı için önerilen yöntemin eğitiminde kullanılan verinin hazırlanma sürecini kolaylaştıran yeni bir eğitim stratejisi ve çok sınıflı araç tespiti başarımlarını artırmak için yeni bir CNN tabanlı yöntem önerilmektedir. Tez çalışması sonucunda elde edilen temel katkılar aşağıda sıralanmaktadır:

- Nesne tespiti başarımlarını artırmak için daha efektif özellik çıkaran CNN tabanlı yeni bir alt sınıf tabanlı sınıflandırıcı yöntemi önerilmiştir [19].
- ERUCV-RS4 adı verilen 4 sınıftan nesne içeren uydu görüntülerinde nesne tespiti çalışmalarında kullanılmak üzere yeni bir veri seti oluşturulmuştur [19].

- İHA görüntülerinde araba sayımı için HLCNN adı verilen tek atışlı yeni bir nesne sayımı yöntemi önerilmiştir [20].
- Önerilen HLCNN yönteminin veri hazırlama sürecini kolaylaştıran yeni bir eğitim stratejisi önerilmiştir [21].
- İHA görüntülerinde araç tespiti için CNessNet adı verilen yeni bir CNN tabanlı yöntem önerilmiştir. Ayrıca, önerilen yöntemin eğitimi sırasında veri dengesizliğini azaltmak amacıyla yeni bir kayıp fonksiyonu sunulmuştur.

Tezin Organizasyonu

Uzaktan algılama görüntülerinde CNN ile nesne tespiti ve sayımında yeni yaklaşımlar öneren bu tezin organizasyonu şu şekildedir.

Birinci bölümde, CNN tabanlı nesne tespiti ve sayımı yöntemleri hakkında genel bilgiler verilmektedir. Ayrıca, mevcut nesne tespiti ve sayımı yöntemleri sınıflandırılarak tanıtılmaktadır.

İkinci bölümde, nesnelere bir bütün yerine farklı alt sınıflarda tanımlayarak CNN eğitimi gerçekleştiren alt sınıf destekli CNN adı verilen yöntem tanıtılmaktadır. Deneysel çalışmalar kapsamında tanımlanan alt sınıf sayısı analiz edilmektedir. Ayrıca transfer öğrenme tekniği ile önerilen yöntem ile oluşturulan özelliklerin nesne tespitine olan etkisi değerlendirilmektedir. Elde edilen sonuçlar, literatürde bulunan mevcut yöntemlerle karşılaştırmalı olarak sunulmaktadır.

Üçüncü bölümde, araba sayımı ve tespitine yönelik iki farklı yöntem tanıtılmaktadır. İlk olarak araba sayımı ve tespiti için CNN tabanlı yeni bir yöntem tanıtılmaktadır. Deneysel çalışmalar kapsamında CARPK ve PUCPR+ veri setleri ile elde edilen sonuçlar sunulmaktadır. İkinci olarak, bu bölümde önerilen CNN mimarisinin eğitimi için etiketleme maliyeti düşük eğitim stratejisi yöntemi tanıtılmaktadır.

Dördüncü bölümünde, araç tespitine yönelik CNN tabanlı çapasız nesne tespiti modeli tanıtılmaktadır. Deneysel çalışmalar kapsamında CARPK ve VAID veri setleri ile elde edilen sonuçlar literatürde bulunan mevcut yöntemlerle karşılaştırmalı olarak sunulmaktadır.

Beşinci bölümde ise tez çalışması kapsamında elde edilen sonuçlar değerlendirilmekte ve gelecek çalışmalara yönelik önerilerden bahsedilmektedir.



1. BÖLÜM

GENEL BİLGİLER VE LİTERATÜR ÇALIŞMASI

Bu bölümde tez çalışması esnasında kullanılan yöntem ve kavramlar tanıtılacaktır. Literatürde uzaktan algılama görüntülerinde nesne tespiti, araç sayımı ve tespiti konularında gerçekleştiren çalışmalar incelenecektir.

1.1. Evrişimli Sinir Ağları

Derin öğrenme mimarilerinden Evrişimli Sinir Ağları (Convolutional Neural Network, CNN), görüntü gibi yapısal dizileri işlemek için tasarlanmış yapay sinir ağı mimarileridir. CNN bilgisayarlı görü uygulamalarında sıklıkla kullanılmakta ve nesne tespiti, görüntü bölütleme ve görüntü sınıflandırma gibi birçok alanda en yüksek başarıma ulaşmaktadır. CNN mimarisi insanlarda bulunan görsel korteksin yapısından esinlenmektedir. Görsel korteksteki sinir hücreleri görüş alanının alt bölgelerine duyarlıdır. CNN bu yapıdan esinlenerek giriş verisinin yerel bilgilerini kullanmaktadır.

LeCun ve ark. tarafından 1989 yılında el yazısı ile yazılmış rakamların tanınması üzerine yapılan çalışmada ilk kez CNN mimarisi kullanılmıştır [22]. CNN en az bir adet evrişim katmanı içeren yapay sinir ağıdır. Evrişim katmanı, CNN mimarisine yerel etkileşim ve paylaşımlı parametre özelliklerini kazandırmaktadır. CNN'in diğer temel bileşenleri; ortaklama katmanı (pooling layer), aktivasyon fonksiyonları (activation functions), tam bağlantılı katman (fully connected layer), kayıp fonksiyonu (loss function), düzenleme (regularization) ve optimizasyondur (optimization) [23]. Bu bölümün alt başlıklarında CNN mimarisinin temel bileşenleri hakkında bilgi verilecektir.

1.1.1. Evrişim Katmanı

Evrişim katmanı, CNN mimarisinin otomatik özellik çıkarma mekanizmasının temel bileşenidir. Bu katmanda giriş sinyaline eğitilebilir filtreler kullanılarak evrişim işlemi uygulanmaktadır. Evrişim işlemi sonucunda özellik haritası adı verilen çıktılar oluşmaktadır. Katmanlar boyunca evrişim işlemi uygulanan bu haritalar CNN tarafından çıkarılan özellikleri taşımaktadır. Giriş sinyali, f , ve kernel, g , için evrişim işlemi Eşitlik 1.1'de verilmiştir.

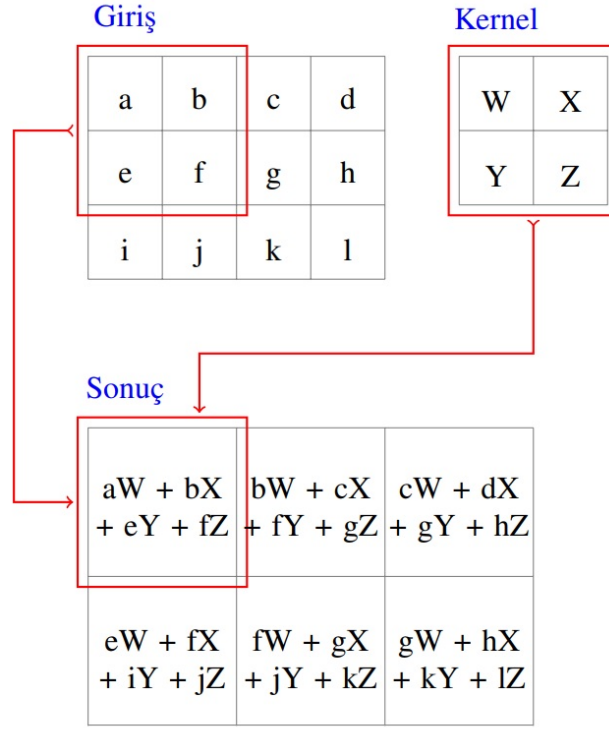
$$s(t) = \int f(\tau)g(t - \tau)d\tau \quad (1.1)$$

Ayrık zamanlı sinyallerde evrişim işlemi Eşitlik 1.2'de verilmiştir.

$$s[t] = \sum_{-\infty}^{+\infty} f[\tau]g(t - \tau) \quad (1.2)$$

Kernel, eğitilebilir parametrelerden oluşmaktadır ve CNN mimarisine eğitilebilir özellik çıkarma özelliği kazandırmaktadır [24]. 2 boyutlu giriş verisi ve 2 boyutlu kernel için örnek evrişim işlemi Şekil 1.1'de verilmektedir. Şekilden görüleceği üzere 4×3 boyutlarındaki giriş matrisi ve 2×2 boyutlarındaki kernel matrisi için sonuç 3×2 boyutlarında matris olarak elde edilmektedir. Evrişim katmanı için önemli parametreler; filtre genişliği (kernel boyutu), filtre sayısı (katman derinliği, depth), adım aralığı (stride) ve dolgu boyutu (padding size) olarak sıralanabilir.

Filtre genişliği, kullanılan kernel boyutunu temsil etmektedir. Genel olarak 3×3 olarak tercih edilmektedir. Filtre sayısı, evrişim işlemi için kullanılacak öğrenilebilir filtre sayısını ve sonuçta oluşacak özellik haritası sayısını belirtmektedir. Adım aralığı, evrişim işlemi sırasında her adımda filtrenin kaç birim kaydırılacağını belirtmektedir. Evrişim işlemi uygulanırken çıktı özellik haritasının boyutunu kontrol etmek amacıyla giriş sınırları sıfır ile dolgu yapılarak genişletilmektedir. Bu genişlemenin boyutunu belirleyen parametre dolgu boyutudur. Genel olarak, dolgu boyutu sıfır seçilmektedir.



Şekil 1.1. 2 boyutlu veri üzerinde evrişim işlemi örneği

1.1.2. Ortaklama Katmanı

Ortaklama katmanı, CNN mimarisini giriş sinyalinde oluşan küçük pozisyon değişimlerine karşı dayanıklı hale getiren ve genelleme yeteneğini artıran bir yapıdır. Ayrıca, özellik haritalarının boyutunu azaltarak ağır hesaplama maliyetini düşürmektedir. En sık kullanılan ortaklama yöntemleri en büyüğe ortaklama (max pooling) ve ortalamaya ortaklamadır (average pooling). Bunlardan en büyüğe ortaklama işleminde giriş sinyalini oluşturan değerlerin en büyüğü seçilmekte, ortalamaya ortaklama işleminde ise giriş sinyali değerlerinin ortalaması alınmaktadır. Burada Şekil 1.2, örnek bir en büyüğe ortaklama işlemini göstermektedir. 6×6 boyutlu giriş matrisine 2×2 pencere boyutu ile en büyüğe ortaklama işlemi uygulanmakta ve 3×3 boyutlu sonuç matrisi elde edilmektedir.

Giriş

84	12	85	64	52	128
122	200	201	86	86	143
135	179	250	56	88	35
255	45	12	21	166	46
12	125	97	101	52	89
88	32	78	46	65	125

Sonuç

200	201	143
255	250	166
125	101	125

Şekil 1.2. En büyüğe ortaklama örneği

1.1.3. Yığın Normalizasyonu

Yapay sinir ağlarının eğitimi genellikle geri yayılım (backpropagation) yöntemi ile ağırlıkların güncellenmesine dayanmaktadır. Geri yayımlı güncelleme hesaplanan hata ile ağ ağırlıklarının son katmandan ilk katmana doğru yeniden hesaplanması ile gerçekleştirilmektedir. Sıralı katmanlarda yapılan güncellemeler birbirini takip eden katmanlarda değişime neden olmaktadır. Bu sorun iç değişken kayması (covariate shift) olarak bilinir. İç değişken kaymasını önleyen ve yapay sinir ağlarının eğitimini daha hızlı ve daha kararlı hale getirmek için kullanılan yöntemlerden birisi de yığın normalizasyonudur. Ioffe ve ark. tarafından [25] 2015 yılında tanıtılan yığın normalizasyonu yönteminde katmanların girdileri, yeniden merkezleme ve yeniden ölçeklendirme yoluyla normalize edilmektedir.

1.1.4. Transfer Öğrenme

Transfer öğrenme (Transfer learning, TL), en geniş ifadesiyle, derin sinir ağlarının belirli problem için eğitim aşamasında öğrendiği bilgiyi saklayıp farklı bir problemin çözümünde kullanmasına dayalı bir yaklaşımdır. Transfer öğrenme yaklaşımında belirli bir problem için eğitilen CNN mimarisinin ağırlıkları farklı problemler için de kullanılmaktadır. Transfer öğrenme yönteminde önceden eğitilmiş CNN modeli

ön-eğitilmiş ağ olarak adlandırılmaktadır. Ön-eğitilmiş ağ olarak genellikle ImageNet [26], COCO [27] gibi büyük veri setleri ile eğitilmiş ağlar kullanılmaktadır.

CNN eğitimi için genellikle oldukça fazla veriye ve hesaplama aracına sahip olmak gerekmektedir. Buna karşın birçok problemde istenilen sayıda veri bulmak oldukça zordur. Bu durumlarda transfer öğrenme yöntemi ile büyük veri setleri ile eğitilmiş ön-eğitilmiş ağların özellik çıkarımı yeteneklerinin kısıtlı veriye sahip olunan problemlere aktarılması sağlanmaktadır.

1.2. CNN Tabanlı Nesne Tespiti Modelleri

CNN tabanlı nesne tespiti modelleri çapa tabanlı (anchor-based) ve çapasız (anchor-free) olmak üzere iki gruba ayrılmaktadır. Bu bölümde, literatürde sıklıkla kullanılan temel çapa tabanlı ve çapasız CNN tabanlı nesne tespiti modelleri tanıtılmaktadır.

1.2.1. Çapa Tabanlı Nesne Tespiti Modelleri

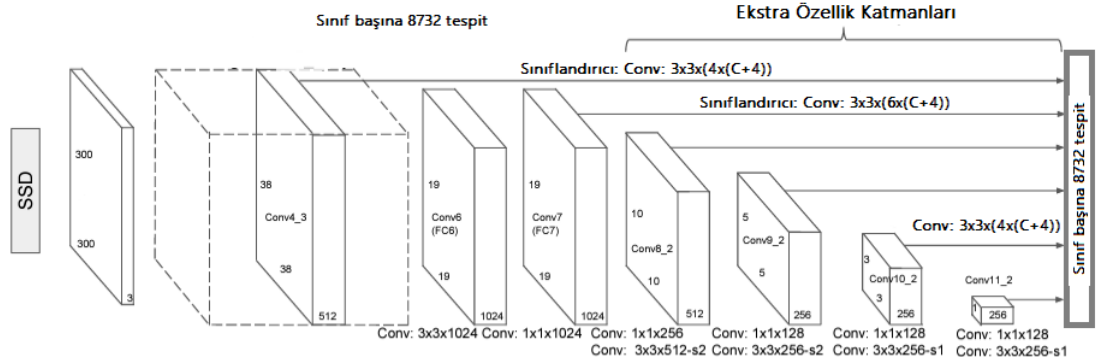
Çapa tabanlı nesne tespiti modelleri tek aşamalı ve iki aşamalı olmak üzere iki bölümde incelenmektedir. Tek aşamalı modellere Tek Atışlı Dedektör (Single Shot MultiBox Detector, SSD) ve Retina Ağı (RetinaNet) modelleri, iki aşamalı modellere ise Bölgesel Evrişimli Sinir Ağı (R-CNN) ailesi örnek olarak verilebilmektedir.

Bu modellerin çapa tabanlı olarak adlandırılmasının sebebi, nesne lokalizasyonunu gerçekleştirmek amacıyla önceden tanımlı sınırlayıcı kutular kullanmalarındadır. Bu bölümde CNN tabanlı nesne tespiti modellerinden SSD, Daha Hızlı R-CNN (Faster R-CNN) ve RetinaNet detaylı olarak tanıtılacaktır.

1.2.1.1. Tek Atışlı Dedektör

SSD, Liu ve ark. tarafından 2015 yılında önerilen tek aşamalı bir nesne tespiti yöntemidir [28]. SSD mimarisi iki temel bölümden oluşmaktadır. İlk bölüm nesne tespiti için çok ölçekli özellik haritaları olarak adlandırılmaktadır. İkinci bölüm ise, nesne tespiti için evrişimsel tahminicilerdir. Şekil 1.3'de verilen SSD mimarisinde omurga ağ olarak VGG-16 [29] kullanılmıştır. VGG-16 ağının devamına eklenen evrişim katmanları ile

yeni özellik haritaları oluşturulmaktadır. Bu özellik haritalarının boyutları aşamalı olarak küçülmektedir ve bu sayede çok ölçekli nesne tespitine olanak sağlanmaktadır.

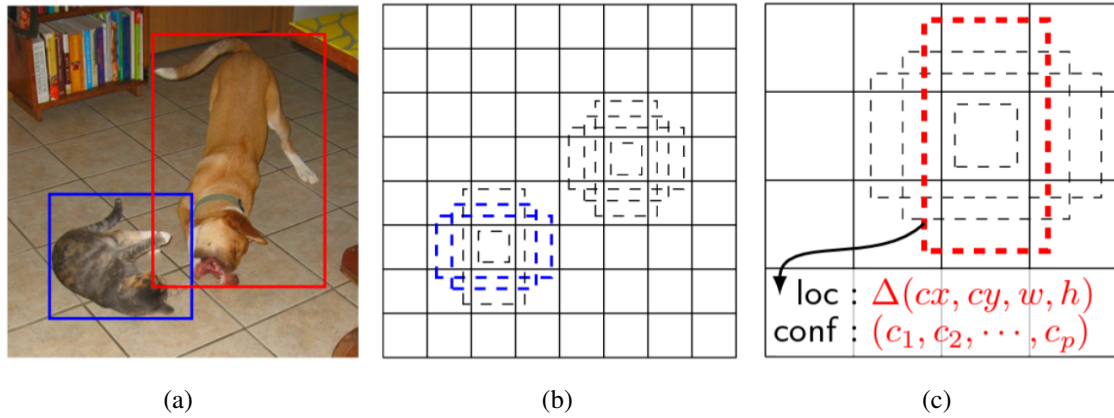


Şekil 1.3. SSD mimarisi

Eklenen her bir özellik haritası katmanı ya da omurga ağdan herhangi bir katman bir dizi evrişim işlemi ile birlikte sabit sayıda nesne tespiti tahmini oluşturmaktadır. Bunlar Şekil 1.3'de ağ mimarisinin üzerinde gösterilmektedir. p kanallı $m \times n$ boyutunda bir özellik katmanı için nesne kategorisi ve sınırlayıcı kutu kayma parametreleri tahmini $3 \times 3 \times p$ boyutlu bir kernel kullanılarak hesaplanmaktadır. Kernelin uygulandığı $m \times n$ konumlarının her birinde bir çıktı değeri üretilmektedir. Sınırlayıcı kutu kayma parametreleri her bir özellik haritası katmanı için önceden belirlenen çapa sınırlayıcı kutuları ile birlikte tahmin edilen sınırlayıcı kutuları belirlemektedir.

SSD mimarisine giriş olarak 300×300 boyutunda bir görüntü verildiğinde nesne tespiti için kullanılan ilk özellik haritası boyutu 38×38 olmaktadır. Çapalar 38×38 boyutundaki özellik haritasının her bir noktasına farklı boyut ve en boy oranlarında hesaplanarak yerleştirilmektedir. SSD mimarisinde her bir hücreye dört farklı çapa yerleştirilmektedir. Şekil 1.4'de farklı boyutta ve en boy oranlarındaki çapaların örnek bir hücreye yerleşimi görülmektedir. Çıktı olarak her bir çapa için dört adet sınırlayıcı kutu kayma parametreleri ve her bir nesne sınıfı için güven skorları hesaplanmaktadır. Nesne sınıf sayısı c olarak temsil edilirse, $m \times n$ boyutlu bir özellik haritası için $(c+1)k \times m \times n$ adet çıktı üretilmektedir. $(c+1)$ arkaplanın eklenmesi sonucunda oluşmaktadır.

SSD mimarisi tüm çapa tabanlı nesne tespiti modelleri gibi çapa sayısı, boyutları ve en boy oranları gibi ek hiperparametreler içermektedir. Bu hiperparametrelerin



Şekil 1.4. SSD mimarisinin farklı boyutlarında özellik haritalarında çapaların temsili yerleşimi (a) temsili sonuç görüntüsü, (b) 8×8 , çapa yerleşimi (c) 4×4 çapa yerleşimi

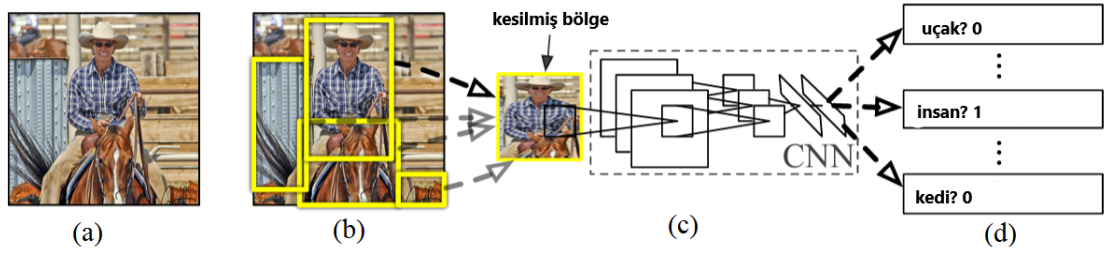
optimizasyonu eğitilen modelin başarımını ve çalışma zamanını etkilemektedir.

1.2.1.2. Faster R-CNN

Faster R-CNN, R-CNN [1–3] ailesinin en gelişmiş versiyonudur. Faster R-CNN, derin öğrenme tabanlı nesne tespiti uygulamalarında sıklıkla kullanılan iki aşamalı nesne tespiti modelidir. Bu iki aşama RPN ve nesne tespit ağıdır.

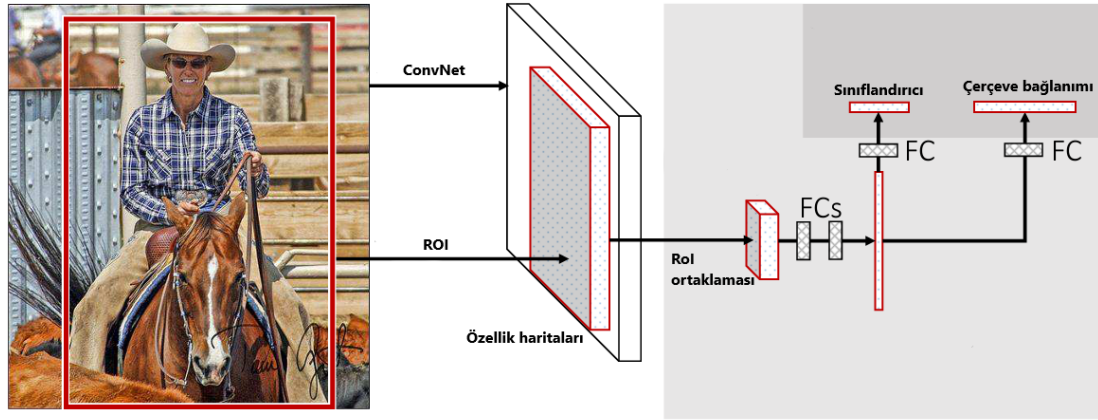
İlk R-CNN [1] modeli sadece nesne tespit ağından oluşmaktadır. RPN yerine, nesne önerileri yöntemlerinden Seçici Arama (Selective Search, SS) [30] kullanılmaktadır. SS girdi olarak bir görüntü almakta ve nesne olabilecek belirli sayıda görüntü alt bölgesini çıktı olarak sunmaktadır. R-CNN, SS'den gelen nesne bölgesi tekliflerini sınıflandırıcı ve sınırlayıcı kutu regresyonuna tabi tutarak nesne tespitini gerçekleştirmektedir. Şekil 1.5'de R-CNN'nin genel blok şeması görülmektedir. R-CNN, SS tarafından yaklaşık olarak iki bin adet nesne bölgesi önerisi almaktadır ve her bir alt görüntüyü özellik çıkarımı aşamasını takip eden sınıflandırıcı ve sınırlayıcı kutu regresyonu aşamalarına tabi tutmaktadır. Bu nedenle hesaplama maliyeti ve zamanı oldukça yüksektir.

Fast R-CNN [2], her bir alt görüntü üzerinden özellik çıkarımı yapmak yerine tüm görüntüye bir defa özellik çıkarımı uygulayarak hesaplama maliyetini ve zamanını oldukça azaltmaktadır. Şekil 1.6'de verilen Fast R-CNN genel blok şemasında özellik çıkarımı tüm görüntüye bir defa uygulanmakta ve SS tarafından önerilen nesne



Şekil 1.5. R-CNN genel blok şeması (a) giriş görüntüsü, (b) SS ile üretilen aday nesne önerileri, (c) CNN özellik çıkarımı, (d) bölgelerin sınıflandırılması [1]

bölgelerinin özelliklerini sınıflandırıcı ve sınırlayıcı kutu regresyonuna iletmek için ilgili Bölge Ortaklama Katmanını (ROI Pooling) kullanmaktadır.

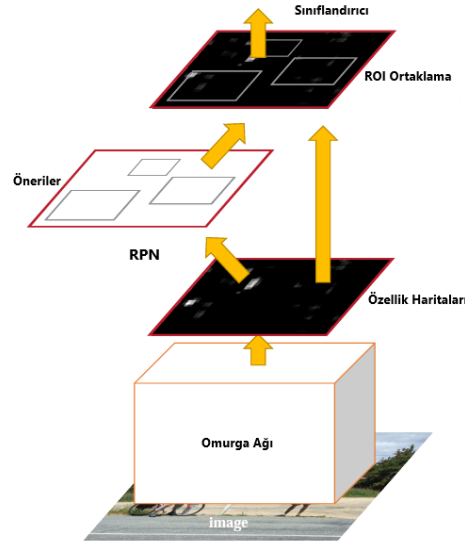


Şekil 1.6. Fast R-CNN genel blok şeması [2]

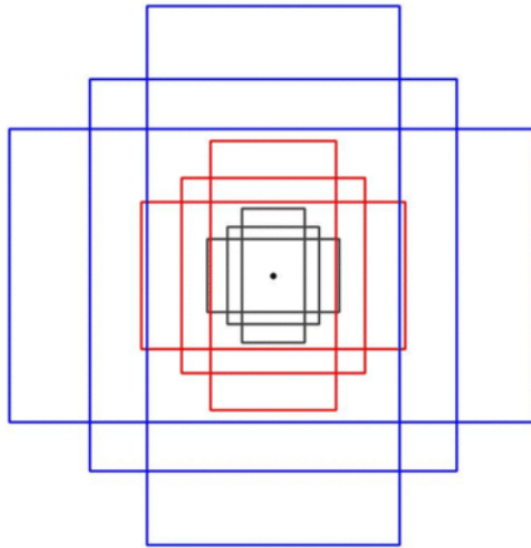
Faster R-CNN'i seleflerinden ayıran en önemli özelliği SS yerine RPN kullanarak baştan sona eğitilebilir bir mimariye sahip olmasıdır. Şekil 1.7'de verilen Faster R-CNN genel blok şemasında nesne bölgesi önerileri RPN tarafından sınıflandırıcı ve sınırlayıcı kutu regresyonu gerçekleştiren nesne tespiti ağına iletilmektedir. RPN, SSD'ye benzer bir şekilde çapaları kullanarak sınıftan bağımsız olarak nesnelere tespit etmeye çalışan bir CNN modelidir. Faster R-CNN modellerinde genellikle üç farklı boyutta ve üç farklı en boy oranlarında dokuz adet çapa kullanılmaktadır. Bu çapalar Şekil 1.8'de görülmektedir.

RPN, Faster R-CNN'i baştan uca eğitilebilir yapmasının yanında eğitim aşamasında arka plan ve ön plan bölgelerinin kullanımlarını ayarlayabilen bir parametre içermesi sebebiyle ön plan-arka plan sınıf dengesizliğini de önlemektedir. Faster R-CNN modellerinin eğitiminde genellikle üç arka plan, bir ön plan bölgesi kullanılmaktadır.

İki aşamalı mimarisinin özellikle ön plan-arka plan sınıf dengesizliğini önlemesi



Şekil 1.7. Faster R-CNN genel blok şeması [3]



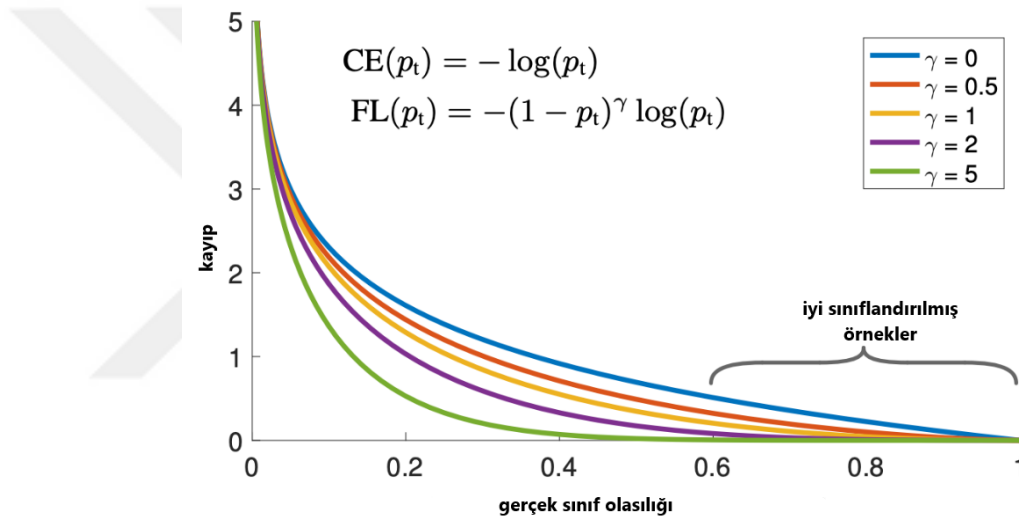
Şekil 1.8. Faster R-CNN çapaları [3]

sayesinde Faster R-CNN, SSD gibi tek aşamalı yöntemlere göre daha başarılı nesne tespiti gerçekleştirebilmektedir. Fakat hesaplama maliyetinin yüksek olması gerçek zamanlı uygulamalarda kullanılmasını engellemektedir.

1.2.1.3. RetinaNet

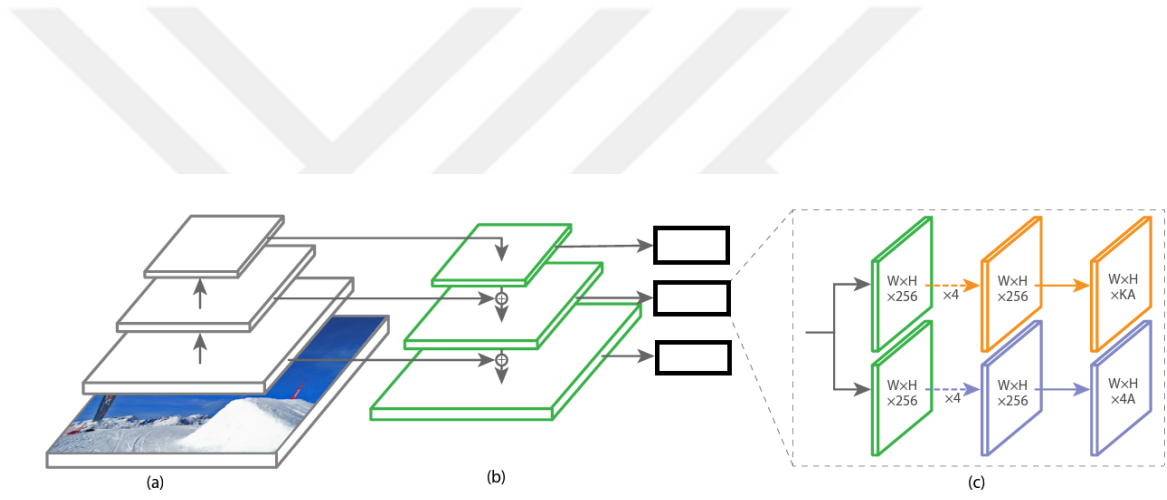
RetinaNet [4], CNN tabanlı tek aşamalı bir nesne tespiti modelidir. Odak kaybı ve Özellik piramit ağı kullanarak nesne tespiti başarımını artırmaktadır.

Tek aşamalı nesne tespiti modellerinin başarımlarına dair en önemli problemlerden biri ön plan-arka plan sınıflarındaki dengesizliktir. RetinaNet modelinde bu problemi önlemek amacıyla çapraz entropi kayıp fonksiyonunun bir türevi olan odak kayıp fonksiyonu kullanılmaktadır. Odak kayıp fonksiyonu bu dengesizliği önlemek amacıyla zor ya da kolay yanlış sınıflandırılan örneklerin ağırlığını artırmakta ya da kolay sınıflandırılan örneklerin ağırlığını azaltmaktadır. Kısaca, yanlış sınıflandırılan örneklerin ağ hatasına katkısını artırıp, daha çok arkaplan örneklerinde kolayca sınıflandırılan örneklerin hataya katkılarını azaltmaktadır. Bu durum Şekil 1.9'da verilen odak kayıp fonksiyonuna ait grafikte görülmektedir [4].



Şekil 1.9. Odak kayıp fonksiyonu [4]

CNN tabanlı nesne tespiti modelleri, derin katmanlara ilerledikçe görüntü alt örneklemeden kaynaklı olarak birçok anlamsal bilgiyi kaybetmektedir. Bu nedenle özellikle küçük nesnelerin tespitinde oldukça başarısız sonuçlar ortaya koyabilmektedirler. Bu durumu önlemek amacıyla Lin ve ark. tarafından Özellik piramit ağı [31] önerilmiştir. Şekil 1.10 (b)'de gösterilen, RetinaNet modelinde, özellik piramit ağı ile yukarıdan aşağıya geçiş ve yanal bağlantı yollarıyla; düşük çözünürlüklü ve semantik açıdan güçlü özellikler ile yüksek çözünürlüklü ve semantik açıdan zayıf olan özellikleri birleştiren bir mimari ortaya koyulmuştur [32]. Tek aşamalı nesne tespiti modeli olan RetinaNet ağı, nesne tespiti başarımının iki aşamalı modelleri geride bırakması ile son yıllarda nesne tespiti problemlerinde yaygın olarak kullanılmaktadır.



Şekil 1.10. RetinaNet mimarisi (a) ResNet, (b) Özellik Piramit ağı, (c) sınıflandırıcı ve çerçeve regresyonu [4]

1.2.2. Çapasız Nesne Tespiti Modelleri

Son yıllarda CNN tabanlı çapa içermeyen nesne tespiti modelleri literatüre kazandırılmıştır. CornerNet [5], CenterNet [6] ve Tam Evrişimsel Tek Aşamalı (Fully Convolutional One-Stage, FCOS) ağ [9] sıklıkla kullanılan çapasız nesne tespiti modelleridir. Çapa tabanlı nesne tespiti modellerinde çapa sayısı, boyutu ve en boy oranları gibi ayarlanması gereken birçok parametre bulunmaktadır. Çapasız nesne tespiti modelleri bu parametreleri ortadan kaldırmaktadır. Bu bölümde CornerNet ve CenterNet detaylı olarak tanıtılmaktadır.

1.2.2.1. CornerNet

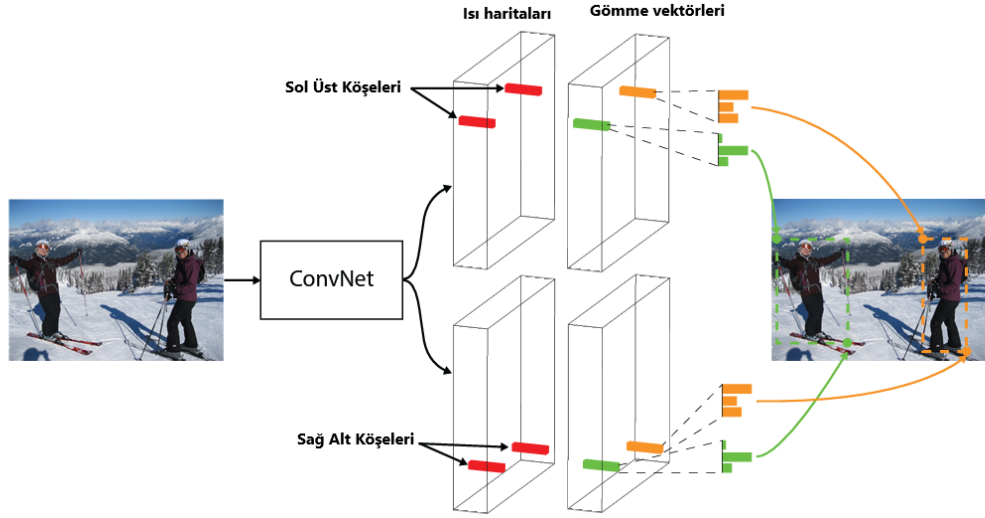
CornerNet [5], Law ve Deng tarafından 2019 yılında tanıtılmıştır. Çapa içermeyen tek aşamalı CNN tabanlı bir nesne tespiti modelidir. Nesne sınırlayıcı kutularını sol üst ve sağ alt köşe olmak üzere bir çift anahtar nokta olarak algılayan bir nesne tespiti modelidir. Şekil 1.11’de yöntemin nesne sınırlayıcı kutuları anahtar nokta olarak gösterilmiştir.

Şekil 1.11’de gösterilen CornerNet modelinde CNN, tüm sol üst ve tüm sağ alt köşeler için iki adet ısı haritası ve algılanan her köşe için bir gömme vektörü çıkarmaktadır. Gömme vektörleri, bir görüntüde birden fazla nesnenin bulunduğu durumlarda tespit edilen köşelerin doğru şekilde eşleştirilmesi ve böylece sınırlayıcı kutuların oluşturulması için kullanılmaktadır. Ağ, aynı nesneye ait olan köşeler için benzer gömme vektörlerini tahmin etmek üzere eğitilmiştir.

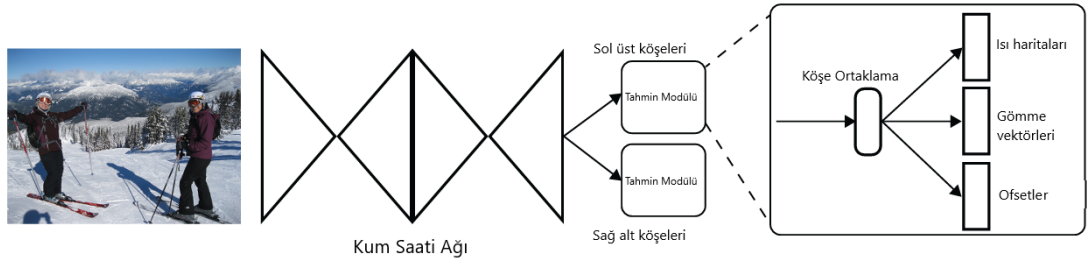
Şekil 1.12, CornerNet mimarisini göstermektedir. CornerNet, omurga ağ olarak Kum Saati ağlarını [33] kullanmaktadır. Ayrıca, ağın köşeleri daha iyi yerelleştirmesine yardımcı olan köşe ortaklama katmanı kullanılmaktadır.

1.2.2.2. CenterNet

CenterNet [6], Duan ve ark. tarafından 2019 yılında tanıtılmıştır. CenterNet, çapa içermeyen, her nesneyi bir çift anahtar nokta yerine üçlü anahtar nokta olarak temsil



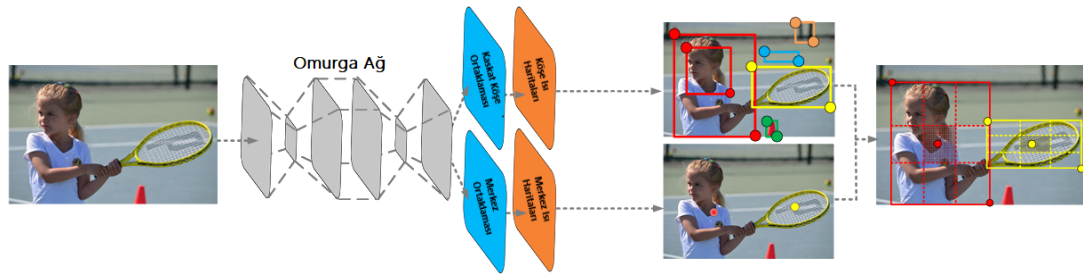
Şekil 1.11. CornerNet nesne sınırlayıcı kutu gösterimi [5]



Şekil 1.12. CornerNet mimarisi [5]

eden tek aşamalı bir nesne tespiti modelidir. Bu ağ CornerNet'in devamı niteliğindedir. CornerNet nesnelere ve sınırlarını belirlemede sınırlı başarıya sahiptir. CenterNet, nesneyi temsil etmek için köşe noktalarının yanında merkez noktasını da kullanmaktadır.

CenterNet, iki adet özelleşmiş yapı içermektedir. Bunlar sırasıyla köşe bilgilerini zenginleştirme ve merkez bölgelerde daha tanımlayıcı bilgiler toplama rollerini üstlenen kademeli köşe ortaklama ve merkez ortaklama katmanlarıdır. Şekil 1.13'de CenterNet mimarisi görülmektedir. CenterNet, omurga ağ olarak Kum Saati ağlarını [33] kullanmaktadır. Özelleşmiş ortaklama katmanları ile birlikte köşe anahtar noktaları ve merkez anahtar noktaları için ısı haritaları üretmektedir. CenterNet, tahmin edilen bir sınırlayıcı kutunun, kesin doğru sınırlayıcı kutu ile yüksek bir çakışmaya sahip olması durumunda, merkez bölgesindeki merkez anahtar noktasının aynı sınıf olarak tahmin edilmesini beklemektedir. Bu sayede yanlış tespit oranını azaltmaktadır.



Şekil 1.13. CenterNet mimarisi [6]

1.3. Uzaktan Algılama Görüntülerinde Nesne Tespiti ve Sayımı

Bu bölümde, tez kapsamında yapılan çalışmalar da göz önünde bulundurularak iki ana bölüm üzerine gerçekleştirilen literatür taraması verilmektedir. İlk bölümde uzaktan algılama görüntülerinde nesne tespiti alanında yapılan çalışmalar tanıtılmaktadır. İkinci bölümde ise İHA görüntülerinden araç sayımı ve tespiti üzerine yapılan çalışmalar sunulmaktadır.

1.3.1. Uzaktan Algılama Görüntülerinde Nesne Tespiti Üzerine Yapılan Çalışmalar

Literatürde uzaktan algılama görüntülerinde nesne tespiti üzerine yapılan çalışmalar, dört ana kategoride gruplandırılmaktadır. Bunlar; şablon eşleştirme, bilgi tabanlı nesne tespiti yöntemleri, nesne tabanlı nesne tespiti yöntemleri ve makine öğrenmesi tabanlı yöntemlerdir. Makine öğrenmesi tabanlı nesne algılama yöntemleri diğer yöntemlere göre daha güçlü ve başarılı yaklaşımlara sahiptir. Bu yaklaşımlar, genellikle nesne algılama görevini bir sınıflandırma problemi olarak ele alır. Makine öğrenmesi tabanlı nesne algılama yöntemleri genel olarak özellik çıkarımı ve sınıflandırıcı olmak üzere iki ana bölüme ayrılmaktadır. Özellik çıkarım aşamasında kullanılan teknik, nesne tespiti başarısını oldukça etkilemektedir. Ölçeklemeden Bağımsız Özellik Dönüşümü (SIFT) [34], Hızlandırılmış Sağlam Özellikler (SURF) [35] ve Yönlendirilmiş Gradyanların Histogramı (HOG) [36] gibi özellik çıkarım yöntemleri araştırmacılar tarafından sıklıkla kullanılmasına rağmen son yıllarda eğitilebilir olmalarının avantajı ile derin öğrenme tabanlı özellik çıkarımı yöntemlerinin daha başarılı sonuçlar ürettiği görülmektedir.

Sirmacek ve Unsalan 2009 yılında yaptıkları çalışmada, uydu görüntülerinde binaları ve kentsel alanları tespit etmek için graf teorisi ile SIFT algoritmasını kullanmıştır

[37]. Teurmer ve ark. yoğun kentsel alanlarda araç tespiti için HOG özelliklerini kullanmışlardır [38]. Lu ve ark. uzaktan algılama görüntülerde gemi tespitini başarılı bir şekilde gerçekleştirmek amacıyla HOG özelliklerini dönen kayan pencere yöntemi ile birlikte kullanmışlardır [39]. Yüksek uzamsal çözünürlüğe sahip uzaktan algılama görüntüleri, karmaşık arka plan bilgisi ve ayrıntılı nesne bilgisi içermektedirler. Bu durum uzaktan algılama görüntülerinde bölgenin uzamsal ve yapısal özelliklerinin doğru anlaşılmasını zor bir iş haline getirmektedir. SIFT ve HOG yöntemlerin çıkardığı yerel özelliklerin bu sorunu çözmek için yeterli olmadığı ve daha güçlü uzamsal özelliklerin gerekli olduğu görülmektedir. Bu sebeple, literatürde derin mimari yapıları sayesinde güçlü özellik çıkarımı yapabilen CNN modelleri sıklıkla kullanılmaktadır.

Salberg, havadan uzaktan algılama görüntülerinde fok tespiti için özellik çıkarma aşamasında önceden eğitilmiş derin CNN kullanmıştır. Bu yöntem, potansiyel nesne bölgelerinin tespiti, CNN tabanlı özellik çıkarımı ve doğrusal destek vektör makinesi ile potansiyel nesnelerin sınıflandırılması olmak üzere üç aşamadan oluşmaktadır [40]. Chen ve ark., uzaktan algılama görüntülerindeki değişken boyutlu nesnelere daha iyi tespit edebilmek için görüntülerin çoklu ölçek özelliklerini çıkarmak amacıyla CNN'nin son evrişim ve ortaklama katmanını çok değişkenli bloklara ayıran hibrit bir CNN modeli önermektedir [41]. Jiang ve ark. yaptıkları çalışmada, ilk olarak uzaktan algılama görüntülerinde araç konumunu tespit etmek için görüntü parçaları üreten graf tabanlı süper piksel segmentasyonu yöntemi kullanmaktadırlar. Ardından bu bölgelerin araç içerip içermediğini tespit etmek için CNN tabanlı bir nesne tespit ağı önermişlerdir [42]. Li ve ark. uzaktan algılama görüntülerinde araç tespiti için kısıtlı eğitim verisi ile eğitilen CNN modellerinin aşırı uyum sorununu aşmak amacıyla gürültü giderici Yığınlanmış Otomatik Kodlayıcı kullanan bir yöntem önermişlerdir [43].

Zhu ve ark. uzaktan algılama görüntülerinde nesne tespitinin nesne renklerinden, en-boy oranlarının varyansından, arka plan çeşitliliği ve nesne yönelimlerinin değişkenliğinden dolayı zorlaştığına değinmişlerdir. Çalışmalarında nesne yönelimine dayanıklılığı artırmak için CNN özelliklerini birleştiren bir yöntem önermişlerdir [44]. Cheng ve ark. CNN mimarilerinin nesne tespitinde dikkat çekici başarısına rağmen, nesne döndürme, sınıf içi çeşitlilik ve sınıflar arası benzerlik dahil olmak üzere nesne tespiti başarımını etkileyen birkaç büyük zorluk olduğundan bahsetmişlerdir [45]. Bu sorunları ele almak için

birbirini takip eden iki farklı çalışma sunmuşlardır. Cheng ve ark., nesne tespiti başarımını geliştirmek için dönme ile değişmeyen CNN (RICNN) modelini eğitmek için yeni ve etkili bir yaklaşım önermişlerdir [46]. Çalışmada, geleneksel CNN modellerinin eğitiminden farklı olarak, RICNN modeli, eğitim örneklerinin özelliklerini döndürmeden önce ve sonra açıkça zorlayan bir düzenleme kısıtlaması uygulayan yeni bir kayıp fonksiyonu ile eğitilmektedir. Cheng ve ark., RICNN modelini daha başarılı hale getirmek amacıyla dönme ile değişmeyen ve Fisher ayırmacı CNN modellerini eğitmek için basit ama etkili bir yöntem önermişlerdir [45]. Bu çalışmada CNN özelliklerinde dönme ile değişmeyen düzenleyici ve bir Fisher ayırım düzenleyicisi uygulayan yeni bir kayıp fonksiyonu önerilmektedir. Spesifik olarak, ilk düzenleyici, dönme değişmezliği elde etmek için döndürmeden önce ve sonra eğitim örneklerinin CNN özellik temsillerinin birbirine yakın şekilde eşlenmesine zorlamaktadır. İkinci düzenleyici, CNN özelliklerini sınıf içi küçük dağılıma, ancak sınıflar arası büyük ayrıma sahip olacak şekilde kısıtlamaktadır [45].

Long ve ark., yüksek çözünürlüklü uzaktan algılama görüntülerinde tespit edilen nesnelerin otomatik olarak doğru bir şekilde konumlandırılması sorununu çözmeye odaklanmışlardır [18]. Bunun için bölge önerisi, sınıflandırma ve doğru nesne yerleştirme aşamasından oluşan bir yöntem önermişlerdir. İlk olarak, görüntülerdeki tüm ilgi çekici nesne olabilecek bölgeleri tespit etmek amacıyla aday bölgeler oluşturmak için bir bölge önerme yöntemi kullanmışlardır. Daha sonra, her bölge önerisine karşılık gelen yerel görüntüden genel görüntü özelliklerini, iki boyutlu indirgeme CNN'lerin bir kombinasyon modeli ile çıkarmışlardır. Son olarak, konum doğruluğunu iyileştirmek için, nesne olarak tespit edilen bölgelerin sınırlayıcı kutularını optimize etmek için Maksimum Olmayan Baskılama (NMS) algoritmasıyla birleştirilmiş Denetimsiz Puan Tabanlı Sınırlayıcı Kutu Regresyon algoritmasını önermişlerdir [18].

Zhang ve ark. uzaktan algılama görüntülerinden uçak tespiti için zayıf denetimli bir yöntem önermişlerdir [47]. Önce nesnelerin yüksek seviyeli özelliklerini ve hiyerarşik özellik gösterimlerini çıkarmak için CNN tabanlı bir yöntem geliştirmişlerdir. Ardından, orijinal görüntüden eğitim veri kümesini otomatik olarak çıkarmak ve artırmak için yinelemeli, zayıf denetimli bir öğrenme çerçevesi önermişlerdir. Önerileri çıkarmak ve büyük ölçekli uzaktan algılama görüntülerinde bile daha verimli şekilde nesnelerin

konumlarını doğru bulmak için aday bölge teklif ağı ve yerelleştirme ağını birleştiren bir CNN yöntemi önermişlerdir.

Li ve Wang [48], daha az aday bölge önerebilen ve daha fazla ayırt edici özellik çıkarabilen verimli bir çerçeve önermişlerdir. Önerilen çerçeve, iyi tasarlanmış ve güçlü iki adet sekiz katmanlı CNN'den oluşmaktadır. İki adet CNN'i kaba-ince olmak üzere iki aşamada kullanan bir nesne tespiti yöntemi önermişlerdir. Kaba adımda, nesnelere içerebilecek aday görüntü alt bölgeleri çıkarılması amaçlanmıştır. İnce adımda, ince aday bölgeler kaba aday bölgelerden kırılarak ve nesne veya arka plan olarak sınıflandırılmaktadır.

Literatürdeki birkaç araştırmada uzaktan algılama görüntülerinde nesne tespiti doğruluğunu ve hızını artırmak için Faster R-CNN mimarisini uyarlayan yeni yöntemler önerilmiştir. Bu çalışmalardan birinde, Li ve ark. [49], mevcut derin öğrenmeye dayalı yöntemlerin çoğunun, döndürme varyasyonları ve görünüm belirsizliği gibi jeo-uzamsal nesne tespiti için karşılaşılan zorluklarla etkili bir şekilde başa çıkmasının zor olduğunu öngörmüşlerdir. Bu sorunları ele almak için çalışmada, uzaktan algılama görüntüleri için tasarlanmış RPN ve Yerel-Bağlamsal Özellik Birleştirme Ağını içeren yeni bir derin öğrenme tabanlı nesne algılama çerçevesi önermektedirler. Zhong ve ark. [50], uzaktan algılama görüntülerindeki nesnelere dağılımının konuma göre değişken ve karmaşık olması sorununu aşmak amacıyla daha iyi nesne önerileri oluşturmak için Konuma Duyarlı Dengeleme (PSB) çerçevesi önermişlerdir. Xu ve ark. [51], nesne tespitinde geometrik modellemeyi ele almak için deforme olabilen ConvNet adlı bir CNN yapısını tanıtmışlardır. Çalışmada, normal evrişim katmanı, deforme olabilen bir evrişim katmanı ile değiştirilerek, Deforme Olabilen Bölge Tabanlı Evrişimli Ağlar (R-FCN) oluşturulmuştur. Ayrıca, yanlış bölge önerilerindeki artışı önlemek amacıyla En-boy Oranı Kısıtlı Maksimum Olmayan Baskılama yöntemi tanıtılmıştır. Tang ve ark. [52], Faster R-CNN mimarisinin iki temel sorun sebebiyle uzaktan algılama görüntülerinde küçük nesnelere tespit zorlandığından bahsetmişlerdir: "(1) Faster R-CNN'deki bölge teklif ağı, nispeten kaba özellik haritaları nedeniyle küçük boyutlu nesnelere doğru bir şekilde konumlandırılması için düşük başarıma sahiptir. (2) RPN'den sonraki sınıflandırıcı, nesnelere ve karmaşık arka planları iyi ayırt edememektedir." Çalışmada bu sorunları aşmak amacıyla hiyerarşik özellik haritalarının bir kombinasyonu ile araç

gibi küçük nesnelere çıkarmak için bir hiper bölge teklif ağı önerilmişlerdir. Ayrıca, aday bölgeleri başarılı şekilde sınıflandırmak için sınıflandırıcıyı, negatif örnek madenciliği ile yanlış algılamayı azaltmayı amaçlayan bir dizi güçlendirilmiş sınıflandırıcı ile değiştirmişlerdir.

Gerçek zamanlı nesne tespiti elde etmek için, birkaç çalışma, doğal sahne görüntüleri için geliştirilen tek aşamalı nesne tespiti yöntemlerini uzaktan algılama görüntülerine aktarmaya çalışmıştır. Ahmed ve ark. [53], uzaktan algılama görüntülerinde küçük nesnelere tespit edebilmek amacıyla düşük katman semantik bilgisini ve üst katman uzamsal çözünürlüğünü iyileştirmek için FPN yerine yeni bir bağlam füzyon modülüne sahip gelişmiş bir RetinaNet modeli önermişlerdir. Lu ve ark. [54], tek aşamalı yöntemlerin uzaktan algılama görüntülerinde nesne tespitinde başarımını daha da iyileştirmek amacıyla Dikkat ve Özellik Füzyonu SSD adlı yeni bir ağ önermişlerdir. İlk olarak, sık özelliklerin anlamsal bilgisini geliştirmek için çok katmanlı bir özellik birleştirme yapısı tasarlamışlardır. Ardından, özellik bilgilerini taramak için çift yönlü bir dikkat modülü tanıtmışlardır. Bu modül, arka plan gürültüsünü bastırmak ve temel özelliği vurgulamak için uzamsal dikkat ve kanal dikkati içermektedir. Ayrıca kayıp fonksiyonu, pozitif ve negatif örnekler arasındaki dengesizliği hafifletmek için optimize edilmiştir. Qin ve ark. [55], küçük nesnelere özellik ve semantik bilgilerinin kombinasyonunu anlayarak ve analiz ederek yüksek çözünürlüklü görüntülerin uzamsal bilgilerini çıkarmaya odaklanan özel olarak optimize edilmiş tek aşamalı bir ağ (SOON) önermişlerdir. SOON modeli, özellik geliştirme, çok ölçekli algılama ve özellik birleştirmeden oluşmaktadır.

Son yıllarda uzaktan algılama görüntülerinde nesnelere yönlerinin tespit edilmesi de birçok çalışmada araştırma konusu olmuştur. Böyle bir çalışmada Tang ve ark. [56], uzaktan algılama görüntülerinde nesnelere yön bilgisi ile birlikte tespit edebilmek için yönlendirilmiş sınırlayıcı kutu tahmini gerçekleştirebilen SSD tabanlı uçtan uca tek bir CNN önermişlerdir. Nesnelere tahmin etmek amacıyla her özellik haritası konumunda çeşitli ölçeklerde bir dizi varsayılan sınırlayıcı kutu kullanan yöntem açısı parametresini içeren yönlendirilmiş sınırlayıcı kutuların nesne ile daha iyi örtüşmesi amacıyla her varsayılan kutu için ofset değerleri de tahmin etmektedirler. Liu ve ark. [57], geleneksel sınırlayıcı kutuları döndürülebilir sınırlayıcı kutular ile değiştirmişlerdir. Tanımlanan

döndürülebilir sınırlayıcı kutuyu kullanan yeni bir tespit yöntemi önermişlerdir. Önerilen dedektör (DRBox), nesnelerin oryantasyon açılarının keyfi olduğu durumu etkili bir şekilde ele alabilmektedir. Ayrıca önerilen eğitim çerçevesi sayesinde DRBox nesnelerin doğru oryantasyon açısını öğrenmeye zorlanmaktadır, böylece dönme ile değişmeyen özelliği elde edilmektedir. Liu ve ark. [58], yönlendirilmiş gemileri yön bilgisi ile tespit edebilmek için YOLOv2 [59] mimarisini kullanarak döndürülmüş/yönlendirilmiş sınırlayıcı kutuları doğrudan tahmin edebilen bir yöntem önermişlerdir.

Barut, yaptığı tez çalışmasında [60], uzaktan algılama görüntülerini kullanarak kara kullanım analizi ve mekânsal nesne tespit analizi gerçekleştirmek amacıyla CNN tabanlı yöntem sunmuştur. Sümbül yaptığı tez çalışmasında uzaktan algılama görüntülerinde ince taneli nesne tespiti üzerine bir yöntem sunmuştur. Çalışmada farklı kaynaklardan edinilen görüntüleri doğru bir şekilde karşılaştırmayı ve sınıflandırma kurallarını aynı anda tek bir çerçevede öğrenmeyi amaçlayan bir sinir ağı modeli önermiştir [61]. Soydaş, yaptığı tez çalışmasında, uydu görüntülerinden uçak tespiti üzerine yeni yaklaşımlar üzerine yöntemler tanıtmıştır [62]. Geleneksel yöntemler ile derin öğrenme tekniğine dayalı farklı sinir ağı mimarilerini eğitmiş ve test etmiştir. Ayrıca, büyük ölçekli uzaktan algılama görüntülerinde hızlı tespit ve yüksek başarımlar için yeni bir yöntem önermiştir, farklı mimarilerin kullanımı ve eğitim yöntemlerinin başarımlarına etkilerini incelemiştir. Önün, derin öğrenme ile yüksek çözünürlüklü hava görüntülerinde yolların tespit edilmesi üzerine yaptığı çalışmada [63], büyük boyutlu hava görüntülerinin işlenebilir resim boyutuna gelmesi için bölümlenerek işlenmesi ve veri seti elde edilmesini sağlamıştır. Hazırlanan bu veri seti eğitilerek daha sonra test amaçlı ayrılan verilerin CNN tabanlı ağ modeli ile yolların tespiti gerçekleştirilmiştir. Körez yaptığı tez çalışmasında [64], uzaktan algılama görüntülerindeki nesnelerin tespit edilmesi için iki farklı nesne tespit modeli önermiştir. Önerilen ilk modelde, Faster R-CNN modelinin insansız hava araçlarından çekilen görüntülerdeki nesne tespit başarımının artırılmasına yönelik geliştirmeler yapmıştır. Önerilen ikinci modelde ise, üç farklı derin öğrenme nesne tespit modelinin optimize edilmiş katsayı ağırlıklı nesne tespiti yapmasına olanak sağlayan topluluk öğrenimi tabanlı bir nesne tespitini gerçekleştirmiştir.

1.3.2. İHA Görüntülerinden Araç Sayımı ve Tespiti Üzerine Yapılan Çalışmalar

Literatürde uzaktan algılama görüntülerinde nesne tespiti üzerine birçok yaklaşım bulunmaktadır. Bu bölümde derin öğrenme yöntemleri kullanılarak uzaktan algılama görüntülerinde araç tespiti gerçekleştiren çalışmalar özetlenmektedir. İHA görüntülerinden araç sayımı gerçekleştiren yöntemler, regresyon tabanlı, yoğunluk haritası tabanlı ve nesne tespiti tabanlı olmak üzere üç ana sınıfta kategorize edilmektedir.

Regresyon tabanlı yöntemler, zorlu nesne yerleştirme problemi ile uğraşmadan nesne sayısını tahmin etmeyi amaçlamaktadır. Regresyon tabanlı yöntemler uzamsal bilgiyi kullanmaması sebebiyle, istenilen düzeyde nesne sayımı başarımına ulaşmamaktadır. Bu yöntemler özellikle değişen ölçeklerdeki nesnelere ve kümelenmiş dağılım ile ilgili zorluklara sahiptir [65]. Son yıllarda nesnelere uzamsal bilgisini kullanarak daha güçlü özellikler çıkarmasını sağlayan yoğunluk haritası tabanlı yöntemler sıklıkla kullanılmakta ve en iyi nesne sayımı başarımını göstermektedir. Bu sebeple literatürde bulunan çalışmalar yoğunluk haritası tabanlı yöntemler ve nesne tespiti tabanlı yöntemler olmak üzere iki ana başlıkta tanıtılmıştır.

Tayara ve ark. [66], havadan çekilen görüntülerde araçların küçük boyutları, farklı türleri ve binalardaki klima üniteleri, çöp kutuları ve yol işaretleri gibi diğer bazı nesnelere görsel görünümündeki benzerlikler nedeniyle uzamsal yoğunluk haritasını öğrenmek için evrişim sinir ağını kullanmışlardır. Aich ve Stavness yaptıkları çalışmada [67], tek bakışlı regresyon modellerinin nesne sayımı başarımını artırmak amacıyla basit ve etkili bir yöntem sunmuşlardır. Çalışmada, basit nokta etiketleri üzerinden oluşturulan kesin referans aktivasyon haritaları ile ağın son evrişim katmanının aktivasyon haritalarını düzenleyerek tek bakışlı regresyon sayma modellerinin nesne sayma başarımını geliştiren bir yöntem sunmuşlardır. Kan ve ark. [68], karayollarındaki yoğunluğu ölçmek amacıyla araç görünümündeki değişkenlik ve düşük çözünürlük sorunlarına odaklanarak araç sayımı gerçekleştiren CNN tabanlı Araç-sayar CNN (VC-CNN) isimli bir yöntem önermişlerdir. VC-CNN, yoğunluk haritaları oluşturularak doğru araç sayısı tahminleri üretmektedir.

Cai ve ark. [69], nesne tespiti ve sayımı için Rehberli Dikkat Ağı (GANet) ismini verdikleri CNN tabanlı bir yöntem önermişlerdir. GANet, doğru araç lokalizasyonunu kolaylaştırmak için oluşturulan arka plan dikkati ve ön plan dikkati modülleri sayesinde geliştirilmiş özellik gösterimi sağlamaktadır. Ayrıca, çeşitli karmaşık sahnelerde sağlam bir model eğitmek için tasarlanan yeni veri artırımı stratejisi tanıtmışlardır. Goldman ve ark. [70], yoğun ve bitişik nesnelerin bulunduğu sahnelerde NMS yönteminin sadece sınıflandırma skoru kullanarak çakışan tespitler arasında yeterli düzeyde ayırım yapamadığını ortaya koymuşlardır. Bu çalışmada bu sorunu aşmak amacıyla, Soft-IoU adı verilen bir katman eklenerek NMS'de kullanılmak üzere ek bir değer tahmin edilmektedir. Tanıtılan nesne tespiti ile araç sayımı yöntemlerine ek olarak literatürde nesne tespiti için kullanılan Faster R-CNN, RetinaNet, CenterNet ve CornerNet modelleri de kullanılmaktadır [67].

Liu ve ark. SSD yöntemini uyarlayarak DLR-MVDA ve VEDAI veri setlerinde araç tespiti gerçekleştirmişlerdir [71]. Lavanya ve ark. [72], uzaktan algılama görüntülerinde küçük nesnelerin tespiti, özellikle de araç tespiti için Çok Ölçekli Swift Tespit Sistemi adı verilen bir yöntem önermişlerdir. Önerilen yöntem, YOLO [73]'ün özelleştirilmiş bir versiyonu olan YOLT [74]'ü içermektedir. Bu yaklaşım, birden fazla sensör üzerinde daha az eğitim verisi ile farklı ölçeklerdeki nesneleri hızlı bir şekilde tespit edebilmektedir. Zhou ve ark. [75], uzaktan algılama görüntülerinde arabaları tespit etmek için özel olarak tasarlanmış IR-PANet adlı yeni bir hedef tespit algoritması önermişlerdir. IR-PANet, üst düzey ağ içindeki anlamsal bilgiyi geliştirmeyi amaçlamaktadır. Bunu başarmak için, üst örnekleme yapıldıktan sonra, algoritma derinlemesine ayrılabilir konvolüsyon kullanmaktadır. Bu yaklaşım, sığ ağdaki konvolüsyon işlemi sırasında önemli küçük hedef özellik bilgilerinin kaybolması sorununu etkili bir şekilde azaltmaktadır. Kong ve ark. [76] iki ana soruna odaklanan iki aşamalı bir nesne algılama yöntemi önermişlerdir. İlk olarak, ölçek değişimi sorununu etkili bir şekilde ele almak için farklı alıcı alanlara sahip konvolüsyonel katmanlardan yararlanan bir Paralel RPN tasarlamışlardır. İkinci olarak, eğitim aşamasında yoğun kümelenmiş araçlarla başa çıkmak için yoğunluğa dayalı bir örnek atayıcı geliştirmişlerdir. Bu atayıcı, düşük kaliteli ve tıkalı pozitif örneklerin varlığını azaltarak genel algılama başarımını iyileştirmektedir. Li ve ark. [77] Çok Dikkatli Özellik Piramit Ağı (MA-FPN) ve daha hassas bir foveal alan (MPFA)

kullanan çapasız bir model tanıtmışlardır. Bu model, özellik haritası üzerinde doğrudan regresyon ve piksel tahmini yoluyla uzaktan algılama görüntülerinde doğru ve hızlı araç tahmini sağlamaktadır. Yu ve ark. [78] uzaktan algılama görüntülerinde küçük boyutlu araçları sağlam bir şekilde tespit etmek için tasarlanmış çapasız bir araç algılama ağı (AVD-kpNet) sunmuşlardır. Zhou ve ark. [79], bağlamsal anlamsal bilgileri dahil ederek özellik çıkarımını geliştiren yeni bir çapasız küçük nesne algılama algoritması önermişlerdir.

Guo ve ark. Grup ve Graf Dikkat ağı (GGANet) [80] ismini verdikleri CNN tabanlı yeni model önermişlerdir. GGANet, bir grup kanal dikkati (GCA) modülü ve öğrenilebilir bir graf dikkati (LGA) modülünü içeren bir kodlayıcı-kod çözücü mimarisidir. Dört adet kalabalık sayma veri seti, iki adet araç sayma veri seti, bir adet uzaktan algılama görüntülerinde nesne sayma veri seti ve bir adet birkaç atışlı nesne sayma veri seti üzerinde deneyler gerçekleştirmişlerdir. Chen ve ark. Seçici Uzamsal-Frekans Piramit ağı (SSFPNet) [81] adını verdikleri piramit dikkat (PA) modülü ve hibrit özellik piramidi (HFP) modülü olmak üzere iki çekirdek modülden oluşan bir model önermişlerdir. PA modülü, dört paralel dal üzerinde çalışarak hedef bölgeleri doğru bir şekilde çıkarır ve arka plan gürültüsünü ortadan kaldırmaktadır. HFP modülü, sayımın doğruluğunu artırmak amacıyla nesne sayımı için farklı alanlardan gelen ölçek bilgilerinden yararlanarak uzamsal ve frekans alanı bilgilerini birleştirmek için önerilmiştir. Priyanka ve ark. [82] trafik yoğunluğunu tahmin etmek amacıyla Çok Ölçekli Evrişimli Sinir ağları (MSCNN) ve Tek Ölçekli Evrişimli Sinir ağları (SSCNN) adlı iki farklı CNN modeli önermişlerdir. Önerilen modelin başarımı, mevcut derin sinir ağlarına kıyasla UCDS veri setinde daha yüksektir.

Tsai ve ark. [83] çok küçük araçları ve kalabalık insan gruplarını saymak amacıyla yeni bir CNN modeli önermişlerdir. Sabit olmayan İHA kameralarından kaynaklı görüntü titremesi, kötü aydınlatma koşulları ve küçük nesne boyutları gibi zorlukları aşmak amacıyla Çoklu Veri Artırma ve Çoklu Derin Öğrenme (MDAMD) adını verdikleri modeli önermişlerdir. Deneysel çalışmaları, kalabalık gruplarda insan tespiti ve motosiklet tespiti veri setleri ile yürütmüşlerdir. D'Alessandro ve ark. [84], nesne sayımı veri setlerinin etiketleme maliyetine dikkat çekerek Her Şey için Etiketsiz Sayma (AFreeCA) adını verdikleri denetimsiz nesne sayımı yöntemi önermişlerdir. Görüntüyü

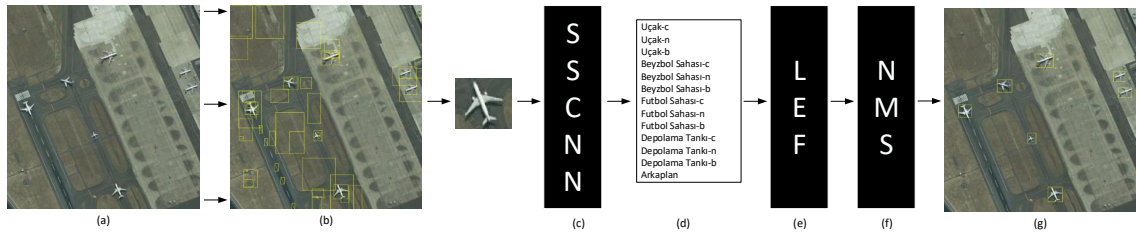
güvenilir bir şekilde sayılabilecek nesnelere bölme için yoğunluk sınıflandırıcı destekli bir yöntem sunmuşlardır. Yöntemin başarımını ölçmek için deneysel çalışmaları kalabalık ve araç sayımı veri setleri üzerinde yürütmüşlerdir. Liu ve ark. [85], trafik sıkışıklığının yoğun olduğu senaryolarda yüksek kaliteli araç yoğunluğu haritalarının oluşturulması için özellik piramidi bölünmüş dikkat mekanizmasını temel alan bir araç nesne sayma ağı önermişlerdir. 3 farklı araç veri setleri üzerinde yapılan deneylerde, tespit veya yoğunluk tahminine dayalı mevcut sayma yöntemlerinin çoğundan daha iyi başarımlar elde etmişlerdir. Wang ve ark. [86], uzaktan algılama sayma görüntülerinde, küçük ölçekli nesnelere sayımı sorununun üstesinden gelmek amacıyla, uzaktan algılama nesne sayımı için HKINet adı verilen yeni bir hiyerarşik çekirdek etkileşim ağı önermişlerdir. HKINet, aynı anda yüksek çözünürlüklü özellikleri korumak ve derin katman anlamsal bilgilerini çıkarmak için çeşitli hiyerarşik çekirdek etkileşim modüllerinden oluşmaktadır. CARPK ve PUCPR+ veri setlerinin içinde bulunduğu 4 uzaktan algılama nesne sayımı veri seti üzerinde kapsamlı deneyler yürütmüşlerdir ve deneysel sonuçlar, HKINet'in iyi performans elde ettiğini göstermiştir. Fu ve ark. [87], trafik sahnelerindeki yoğun nesnelere sayımı için, özellik geliştirmeye dayalı yeni bir çok sınıflı yoğun nesne sayma yöntemi önermişlerdir. Önerilen model, regresyon alt ağından ve sınıflandırma alt ağından oluşmaktadır. Regresyon alt ağı öncelikle, esas olarak başlangıç özellik katmanını ve özellik geliştirme katmanını içeren, iki kanallı tahmin edilen yoğunluk haritalarını oluşmaktadır. Sınıflandırma alt ağı, regresyon alt ağlarının görevine yardımcı olmak için esas olarak yoğun araçların ve insanların iki özellik kanalında sınıflandırılmasını sağlamaktadır. Deneysel çalışmaları yoğun araç sayımı ve yoğun insan sayımı veri setleri üzerinde yürütmüşlerdir. Wang ve ark. [88], nesnelere sınıf bağımsız saymayı amaçlayan ve yalnızca sınırlı sayıda örnek nesneye dayanan sınıftan bağımsız sayma problemini çözmek amacıyla örnek nesnelere görsel ve ölçek bilgilerini ve çapraz dikkati kullanarak örnekler ve sorgular arasındaki model benzerliğini birleştiren yeni, ölçeğe duyarlı transformatör tabanlı özellik birleştirme modülü önermişlerdir. SATCount adını verdikleri modelin başarımını FSC-147 ve CARPK üzerinde yaptıkları deneysel çalışmalar ile ölçmüşlerdir.

2. BÖLÜM

ALT SINIF DESTEKLİ EVRİŞİMLİ SINIR AĞLARI (SSCNN)

CNN görüntüleri sınıflandırmada oldukça başarılıdır. Binlerce farklı sınıftan nesnelere yüksek doğrulukla sınıflandırabilmektedir. Bunu üstün özellik çıkarma yetenekleri sayesinde başarmaktadır. Mevcut nesne tespiti yaklaşımları özellik çıkarımı ve sınıflandırıcı bölümlerinden oluşmaktadır. Nesnelere arkaplandan ayırt etmek için güçlü özelliklerin çıkarılması yöntemlerinin başarımını doğrudan etkilemektedir. HOG [89] ve SIFT [34] en sık kullanılan el yapımı özellik çıkarıcı algoritmalarıdır. Bu özellik çıkarıcıları kullanan yöntemler umut verici sonuçlar elde etmesine rağmen istenilen başarıya ulaşamamıştır. CNN modellerinin derin mimarileri istenilen veriden otomatik olarak güçlü özellikler elde edilmesini sağlamaktadır. CNN tarafından oluşturulan özellikler görüntü sınıflandırmada önemli gelişmeleri ortaya çıkarmıştır [90, 91]. Derin CNN mimarileri görüntüleri sınıflandırmanın ötesinde nesnenin konumu ve boyutunu da doğru şekilde bulabilen nesne tespiti yöntemlerinin en önemli parçası olmuştur. Görüntü üzerinde nesnenin konumunun tespiti için yapılan ilk çalışmalarda hesaplama maliyeti ve zamanı yüksek olan kayan pencere yöntemi kullanılmıştır. Kayan pencere yöntemi görüntü üzerinde her noktada belirli ölçekteki alt görüntülerin sınıflandırılması şeklinde uygulanmaktadır. Görüntü üzerindeki her noktanın araştırılması yerine görüntüden belirli sayıda nesne bölgesi önerileri sunan SS [30] ve Kenar Kutuları (Edge Boxes, EB) [7] algoritmaları nesne önerileri yöntemleri olarak adlandırılmaktadır. SS ve CNN mimarisini kullanan R-CNN [1] yakaladığı başarı ile nesne tespiti problemi için önemli bir dönüm noktası olmuştur.

Tezin bu bölümünde CNN mimarisinin daha ayırt edici özellik çıkarmasını sağlamak amacıyla nesnelere temsilini alt bölgelere ayıran Alt-sınıf Destekli Evrişimli Sinir Ağları (Subclass Supported Convolutional Neural Network, SSCNN) tanıtılmaktadır.



Şekil 2.1. Önerilen SSCNN modelinin genel blok şeması (a) test görüntüsü, (b) SS algoritması ile üretilen aday nesne önerileri, (c) SSCNN ile özellik çıkarımı ve sınıflandırma skorlarının hesaplanması, (d) alt sınıfları ile birlikte nesne sınıfları, (e) LEF, (f) NMS, kesişen nesne bölgelerini elimine edilmesi, (g) Tespit edilen nesneler

SSCNN nesneleri kenar, yakın-merkez, ve merkez gibi alt sınıflara bölerek temsil etmektedir. Bu yaklaşım ile nesnelerin alt bölgelerini tahmin etmeye zorlanan CNN mimarisinin daha ayırt edici özellikler çıkarması amaçlanmaktadır. Tezin bu bölümünde önerilen SSCNN üç aşamadan oluşmaktadır. İlk aşamada, SS algoritması [30] ile nesne önerileri oluşturulmaktadır. İkinci aşamada, SSCNN ile önerilen nesne bölgeleri sınıflandırılmaktadır. Son aşamada, lokalizasyon değerlendirme fonksiyonu, sınıflandırma skorları ile nesne konumu belirlemektedir. Önerilen yeni yaklaşım oluşturulan 4 sınıflı veri seti ile test edilmiştir. Başarımı R-CNN ile kıyaslanmıştır. Buna ek olarak, SSCNN tarafından oluşturulan özelliklerin etkinliğini değerlendirmek amacıyla, Faster R-CNN mimarisi SSCNN tarafından oluşturulan özellikler ile TL yaklaşımı ile eğitilmiştir.

2.1. Önerilen SSCNN Modeli

Önerilen nesne tespiti modeli, nesne önerilerinin oluşturulması, sınıflandırma ve lokalizasyon değerlendirilmesi olmak üzere 3 aşamadan oluşmaktadır. Şekil 2.1'de önerilen yöntemin genel blok şeması görülmektedir. Şekilde c , n ve b sırasıyla merkez, yakın merkez ve kenarı temsil etmektedir. İlk aşamada, SS algoritması ile nesne önerileri oluşturulmaktadır. İkinci aşamada, önerilen SSCNN modeli kullanılarak özellik çıkarımı ve sınıflandırma skorları hesaplanmaktadır. Son aşamada ise lokalizasyon değerlendirme fonksiyonu kullanılarak nesnelerin konumları değerlendirilmektedir. Ayrıca, sınıflandırma skorları yardımıyla kesişen nesne bölgelerinin eliminasyonu NMS yöntemi ile gerçekleştirilmektedir. Önerilen yöntemin sözde kodları Algoritma 2.1'de

verilmiştir.

```

Giriş:  $\mathbf{I}$ ,  $\alpha = 5$  and  $\beta = 3$ 
Çıkış: Tespit edilen nesnelere sınırlayıcı kutuları(LocatedBB)
 $\mathbf{P} \leftarrow \{p_1, p_2, p_3, \dots, p_n\}$ ; // nesne önerilerinin üretilmesi
while tüm nesne önerileri işlenmedi ise do
    /* Nesne önerisi için özellik çıkarımı ve
       sınıflandırma skorunun hesaplanması. */
     $\mathbf{C}_i \leftarrow \text{SSCNN}(\mathbf{I}, p_i)$ ;
    /* Nesne önerisi için lokalizasyonun hesaplanması.
       */
     $\mathbf{L}_i \leftarrow \text{LEF}(\mathbf{C}_i, \alpha, \beta)$ ;
end
LocatedBB  $\leftarrow \text{NMS}(\mathbf{P}, \mathbf{L})$ ; // kesişen nesne önerilerinin
elimine edilmesi

```

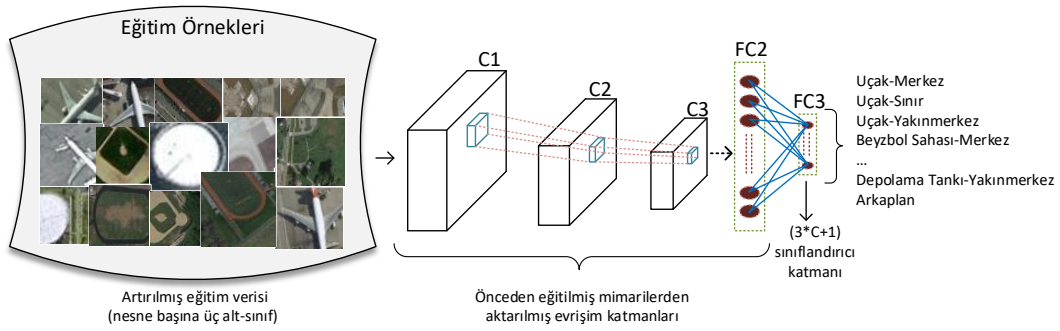
Algoritma 2.1: Önerilen SSCNN modelinin sözde kodları

2.1.1. SSCNN Mimarisi ve Veri Hazırlığı

Önerilen yöntem sınıflandırma katmanında her bir nesneyi üç alt sınıf ile temsil etmektedir. SSCNN modeli, CNN tabanlı görüntü sınıflandırıcı modellerin çıkış katmanlarının $(3 \times C + 1)$ haline getirilmesi ile oluşmaktadır. C , veri setinde bulunan nesne sayısını temsil etmektedir. 3 alt sınıf sayısını ve 1 ise arkaplan sınıfını temsil etmektedir. Eğitim şeması ve SSCNN model mimarisi Şekil 2.2’de görülmektedir.

CNN tabanlı klasik görüntü sınıflandırıcı modellerden farklı şekilde nesnelere temsil etmesi sebebiyle eğitim aşamasında kullanılan nesne görüntüleri üç alt sınıfa uygun şekilde hazırlanmaktadır. Verinin bu yapıya uygun şekilde hazırlanması veri artırma aşamasına dahil edilmektedir. SSCNN eğitimi için veri artırma aşamasında pozitif ve negatif (arkaplan) görüntüler eğitim veri seti kullanılarak elde edilmektedir. Pozitif eğitim örnekleri, her bir nesne için nesne merkezi, nesne yakın merkezi ve nesne sınırı olmak üzere üç alt sınıfta temsil edilen nesne sınıflarından oluşmaktadır.

Veri artırımı aşamasında alt sınıf nesnelere ait görüntülerin elde edilmesi için nesnelere alt bölgelere ayıran daireler kullanılmaktadır. Şekil 2.3’de bir nesnenin alt bölgelerini oluşturan daireler görülmektedir. Merkez pikseli alt bölge içerisinde bulunan alt görüntüler rastgele bir şekilde toplanarak artırılmış veri setine eklenmektedir. Bu şekilde pozitif görüntü örnekleri toplanmaktadır.



Şekil 2.2. SSCNN mimarisi ve eğitim şeması



Şekil 2.3. Pozitif görüntü örnekleri için nesne alt bölgeleri

Nesneyi alt bölgelere ayıran dairelerin yarıçapı nesnenin genişliğine (kısa kenar, O_w) ve bölge belirleme katsayısına (k_c) göre belirlenmektedir. Nesnenin merkezindeki bölge nesne merkez bölgesi olarak adlandırılır (Şekil 2.3'de turuncu bir daire ile gösterilmiştir) ve yarıçapı $R_c = k_c \times O_w$ olarak hesaplanmaktadır. Bu bölgeye bitişik bölge, merkeze yakın bölge olarak adlandırılır (Şekil 2.3'de kırmızı bir daire ile gösterilmiştir) ve yarıçapı $R_{nc} = k_{nc} \times O_w$ olarak hesaplanmaktadır. En dıştaki bölge nesne sınır bölgesi olarak adlandırılır (Şekil 2.3'de yeşil bir daire ile gösterilmiştir) ve yarıçapı $R_b = k_b \times O_w$ olarak hesaplanmaktadır. Veri artırma aşamasında eğitim setinde bulunan her nesne için bölge başına iki örnek görüntü alınmaktadır. Ayrıca, bu eğitim görüntüleri 90° , 180° ve 270° döndürülerek artırılmıştır. Yani artırılmış eğitim seti içerisinde, her nesne başına 24 görüntü bulunmaktadır.

2.1.2. Lokalizasyon Değerlendirme Fonksiyonu (LEF)

LEF, her bir nesne önerisi için tahmin edilen sınıflandırma skorlarını kullanarak

lokalisasyon skorunu Eşitlik 2.1 ile hesaplamaktadır.

$$E_L = \frac{\alpha \times P_{\text{merkez}} + \beta \times P_{\text{yakınmerkez}} + P_{\text{sınır}}}{\alpha}, \{\alpha, \beta > 1; \alpha > \beta\} \quad (2.1)$$

P_{merkez} , $P_{\text{yakınmerkez}}$ ve $P_{\text{sınır}}$, sırasıyla nesne merkezi, nesne yakın merkez ve nesne sınır bölgesinin sınıflandırma skorlarını temsil etmektedir. α ve β alt sınıfların lokalizasyon değerlemesine olan etkilerini belirleyen katsayılardır. $\alpha > \beta$ olması nesne sınırlayıcı kutusunu daha iyi temsil eden nesne önerisinin değerlendirme skorunun yüksek olmasını sağlamaktadır.

2.2. Deneysel Çalışmalar

Önerilen yöntemin başarımını ölçmek için iki farklı veri seti kullanılmıştır. Bu veri setleri Bölüm 2.2.1'de tanıtılmaktadır. Önerilen yöntemin nesne tespitinde başarısını değerlendirmenin yanında alt sınıf sayısının başarıya olan etkisi ve SSCNN mimarisinin çıkardığı özelliklerin etkinliğinin belirlenmesi üzerine de deneyler yapılmıştır.

2.2.1. Veri Setleri

Önerilen SSCNN modelinin başarımını ölçmek için 4 sınıftan nesne içeren yeni bir uzaktan algılama görüntü veri seti oluşturulmuştur. Bu veri seti ERUCV-RS4¹ olarak adlandırılmaktadır. Veri seti 592 adet görüntüden oluşmaktadır. Bu görüntülerde uçak, depolama tankı, beyzbol sahası ve futbol sahası olmak üzere dört kategoriden nesnelere etiketlenmiştir. Bu veri seti 976 uçak, 623 beyzbol sahası, 837 depolama tankı ve 222 futbol sahası içermektedir. Görüntülerin %60'ı eğitim ve %40'ı test için kullanılmak üzere rastgele seçilmiştir.

SSCNN modelinin özelliklerinin etkinliğini göstermek için, 10 sınıflı uzaktan algılama nesne algılama veri seti olan NWPU VHR-10 [91] kullanılmıştır. NWPU VHR-10 veri seti, on farklı nesne türünde 800 uzaktan algılama görüntüsü içeren çok çözünürlüklü ve çok kaynaklı nesne algılama verileridir. Bu veri seti, uçak, gemi, depolama tankı, beyzbol sahası, tenis kortu, basketbol sahası, futbol sahası, liman, köprü ve araç sınıflarından nesnelere içermektedir. Görüntülerin %40'ı eğitim ve %60'ı test için kullanılmak üzere

¹<https://drive.google.com/drive/folders/15c9bfXEJBjH-6kLB00bZdr6aBR79kK8E>

rastgele seçilmiştir.

2.2.2. Değerlendirme Kriterleri

Önerilen yöntemin nesne tespitini sayısal metriklerle nicel olarak değerlendirmek amacıyla literatürde yaygın olarak kullanılan Kesinlik, Duyarlılık, Kesinlik-Duyarlılık Eğrisi (Precision-Recall Curve, PRC) ve Ortalama Kesinlik (Average Precision, AP) kullanılmıştır. Kesinlik, Doğru Pozitif (True Positive, TP) edilen nesnelerin, toplam tahmin edilen nesnelere oranıdır. Duyarlılık, doğru tahmin edilen nesnelerin, toplam gerçek nesne sayısına oranını göstermektedir. Bu metrikler önerilen yöntemin, nesne tespiti başarımını sadece doğru tespitlerle (TP) değil, yanlış tespit (False Positives - FP) edilen nesnelerin ve tespit edilemeyen (False Negatives - FN) nesnelerin de göz önüne alınarak hesaplanmasını sağlamaktadır. Kesinlik ve Duyarlılık metrikleri aşağıdaki eşitlikler ile hesaplanmaktadır.

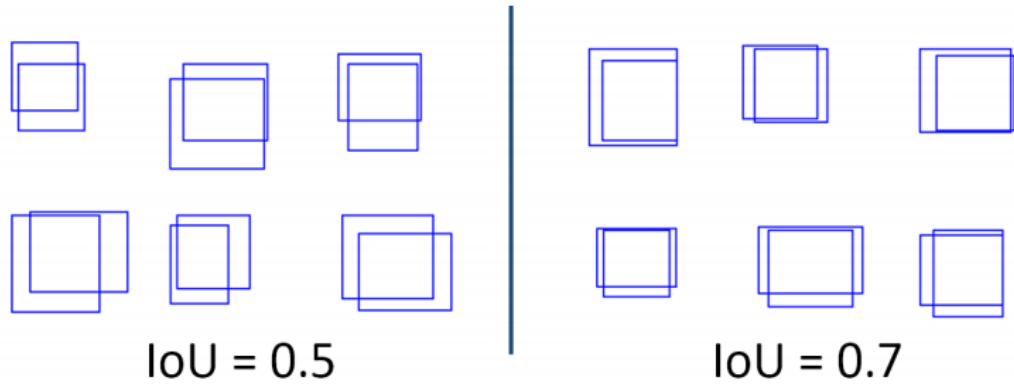
$$\text{Kesinlik} = \frac{TP}{TP + FP} \quad (2.2)$$

$$\text{Duyarlılık} = \frac{TP}{TP + FN} \quad (2.3)$$

AP, 0 ile 1 arasındaki Duyarlılık değerleri için her bir Kesinlik değerinin ortalamasıdır [46]. PRC'nin altında kalan alan olarak tanımlanabilir. Birleşim üzerinden Kesişim, (Intersection over Union, IoU) tahmin edilen sınırlayıcı kutu ile tahmin edilen sınırlayıcı kutu arasındaki örtüşme oranını belirlemektedir [46]. Bu oranın 1 olması tam örtüşme demektir. Genellikle nesne tespiti problemlerinde tespitin başarılı olarak kabul edilmesi için IoU değerinin 0.5'ten büyük ya da eşit olması gerekmektedir. 0.5'ten küçük değerler için tahmin edilen sınırlayıcı kutu yanlış tespit olarak kabul edilmektedir. Ayrıca, aynı nesne için birden fazla IoU değeri 0.5'ten büyük sınırlayıcı kutular var ise, sadece bir tanesi doğru tespit diğerleri ise yanlış tespit olarak kabul edilmektedir. Şekil 2.4'de 0.5 ve 0.7 IoU değerleri için örnek sınırlayıcı kutu örtüşme gösterimleri görülmektedir.

2.2.3. Alt Sınıf Sayısının Belirlenmesi

Önerilen yöntemin en önemli parametresi olan alt sınıf sayısını belirlemek için yapılan deneylerde ERUCV-RS4 veri seti kullanılmıştır. Bu deneylerde AlexNet [92] mimarisinin



Şekil 2.4. 0.5 ve 0.7 IoU değerleri için örnek sınırlayıcı kutu örtüşme gösterimleri [7]

son katmanı düzenlenerek SSCNN modeli oluşturulmuştur. Deneyle, MATLAB R2017a kullanılarak NVidia GTX 960 GPU ve 16 GB RAM üzerinde gerçekleştirilmiştir.

Alt sınıf sayısını belirlemek amacıyla önerilen SSCNN modeli C parametresi 1 ile 5 arasında tüm değerlerde eğitilerek test edilmiştir. Bu beş model, önceden eğitilmiş AlexNet kullanılarak eğitilmiştir. 5 alt sınıf modeli (AlexNet-5), 4 alt sınıf modeli (AlexNet-4), 3 alt sınıf modeli (SSCNN mimarisi olarak adlandırılan AlexNet-3), 2 alt sınıf modeli (AlexNet-2) ve alt sınıf içermeyen model (AlexNet-1) karşılaştırılmıştır.

AlexNet-5, 3 adet yakın merkez bölgesine sahiptir ve k_c , k_{nc1} , k_{nc2} , k_{nc3} , ve k_b parametreleri sırasıyla 0.15, 0.3, 0.5, 0.7 ve 1 olarak seçilmiştir. AlexNet-4, 2 adet yakın merkez bölgesine sahiptir ve k_c , k_{nc1} , k_{nc2} , ve k_b parametreleri sırasıyla 0.2, 0.4, 0.6 ve 1 olarak seçilmiştir. k_c , k_{nc} , ve k_b parametreleri AlexNet-3 için sırasıyla 0.2, 0.5 ve 1 olarak belirlenmiştir. AlexNet-2, SSCNN'nin nesne sınırı alt sınıfının kaldırılmasıyla oluşturulmuştur. AlexNet-1 için k_c parametresi 0.5 olarak seçilmiştir. Bu modellerin eğitimi sırasında veri hazırlığı esnasında, her nesne için 25 negatif görüntü örneği eklenmiştir. Her nesne alt sınıf bölgesi için 8 pozitif görüntü örneği oluşturulmuştur.

Eğitilen beş model, ERUCV-RS4 veri setinde bulunan 237 görüntü ile test edilmiştir. Tablo 2.1, eğitilen modellerin başarımlarını farklı IoU değerleri için AP metriği üzerinden karşılaştırmalı olarak sunmaktadır. Elde edilen sonuçlar, AlexNet-3 modelinin diğer modellere kıyasla nesne tespiti için daha etkili görüntü sınıflandırıcı oluşturmak için daha iyi başarıma sahip olduğunu açıkça göstermektedir. Sonuçlar, alt sınıfların sayısının artırılmasının tespit başarımını artırmadığını da göstermektedir.

Tablo 2.1. ERUCV-RS4 veri seti için SSCNN'nin farklı alt sınıf modellerinin AP metrik sonuçları

IoU	Nesne Sınıfları	AlexNet-1	AlexNet-2	AlexNet-3	AlexNet-4	AlexNet-5
0.3	Uçak	0.7105	0.8216	0.9035	0.8850	0.8963
	Beyzbol Sahası	0.7696	0.9092	0.9452	0.9445	0.9348
	Futbol Sahası	0.8733	0.9372	0.9717	0.9583	0.9792
	Depolama Tankı	0.6267	0.8977	0.9028	0.9201	0.8790
	mAP	0.7450	0.8914	0.9308	0.9269	0.9223
0.5	Uçak	0.6868	0.7859	0.8845	0.8610	0.8760
	Beyzbol Sahası	0.7428	0.9061	0.9419	0.9422	0.9310
	Futbol Sahası	0.8677	0.9372	0.9717	0.9583	0.9792
	Depolama Tankı	0.6118	0.8903	0.8919	0.9051	0.8790
	mAP	0.7272	0.8798	0.9225	0.9171	0.9163
0.6	Uçak	0.5792	0.7121	0.8260	0.8002	0.8205
	Beyzbol Sahası	0.6686	0.8906	0.9349	0.8799	0.9218
	Futbol Sahası	0.8052	0.9277	0.9717	0.9222	0.9784
	Depolama Tankı	0.5768	0.8365	0.8431	0.8602	0.8289
	mAP	0.6574	0.8417	0.8939	0.8656	0.8874
0.7	Uçak	0.2353	0.4995	0.5556	0.5866	0.5881
	Beyzbol Sahası	0.3861	0.8276	0.8400	0.8354	0.8271
	Futbol Sahası	0.4107	0.8692	0.9480	0.9382	0.9289
	Depolama Tankı	0.3962	0.4939	0.5274	0.4864	0.5663
	mAP	0.3570	0.6725	0.7178	0.7116	0.7276

2.2.4. SSCNN ve R-CNN Modellerinin Karşılaştırılması

Önerilen yöntemin başarımını karşılaştırmalı olarak değerlendirmek için, 3 farklı SSCNN ve 2 farklı R-CNN [1] farklı parametrelerle eğitilmiştir. SSCNN'lerin parametreleri Tablo 2.2'de ve R-CNN'lerin parametreleri ise Tablo 2.3'te gösterilmektedir. Eğitim görüntülerindeki her nesne için 25 negatif örnek ve 24 pozitif görüntü örneği oluşturulmuştur. Eğitim veri seti yaklaşık 75000 örnek içermektedir. SSCNN'leri eğitmek için küçük yığın boyutu kullanılmış ve değeri 64 olarak seçilmiştir. SSCNN modelleri 30 epok için eğitilmiştir. Öğrenme oranı ilk 15 epok için 0.0001 ve son 15 epok için 0.00001 olarak belirlenmiştir. R-CNN modellerini eğitmek için EB algoritması bir nesne önerme fonksiyonu olarak kullanılmıştır. İlk öğrenme hızı ve küçük yığın boyutu sırasıyla 0.00001 ve 64 olarak ayarlanmış ve modeller 100 için epok eğitilmiştir.

Eğitilen SSCNN ve R-CNN modelleri, ERUCV-RS4 veri setinde bulunan 237 görüntü ile test edilmiştir. Nesne öneri bölgelerini çıkarmak için SS algoritması kullanılmıştır. SS

Tablo 2.2. Eğitilen SSCNN modellerinin parametreleri

	k_c	k_{nc}	k_b
SSCNN-1	0.2	0.5	1
SSCNN-2	0.1	0.4	1
SSCNN-3	0.3	0.5	1

Tablo 2.3. Eğitilen R-CNN modellerinin parametreleri

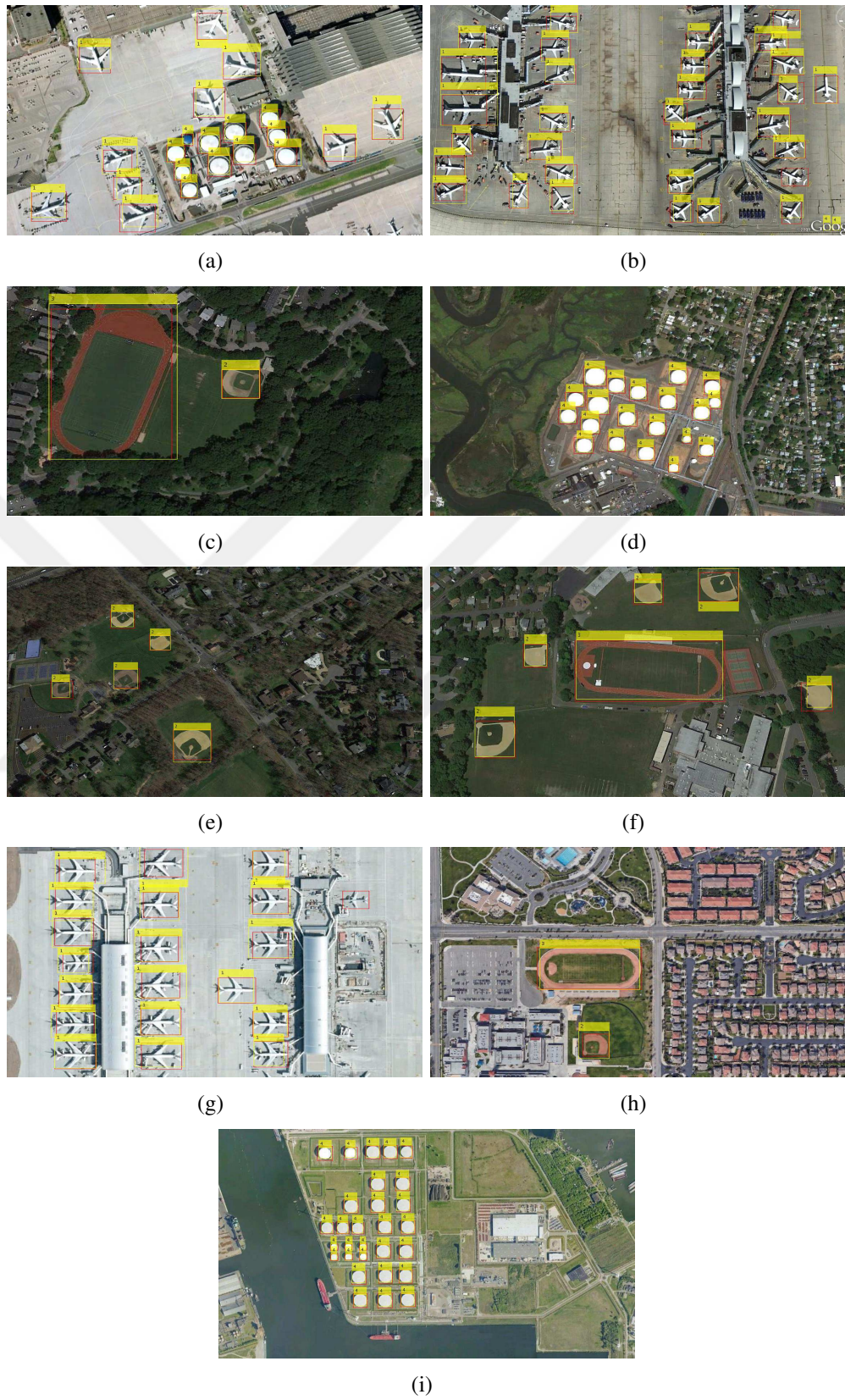
	Pozitif örtüşme	Negatif örtüşme
R-CNN-1	[1,0.5]	[0.5,0.1]
R-CNN-2	[1,0.6]	[0.6,0]

algoritmasının parametreleri Tablo 2.4'te görülmektedir. Sınıflandırmayı gerçekleştirmek için aday nesne bölgeleri kesilmiş ve eğitilmiş SSCNN modeli ile değerlendirilmiştir. LEF, bu bölgelerin tespit başarımını artırmak için sınıflandırma skorları üzerinden nesne önerilerini değerlendirmiştir. Örtüşen sınırlayıcı kutuların sayısını azaltmak için NMS uygulanmıştır. Bir nesne için birkaç sınırlayıcı kutu algılandığında, NMS üst üste gelen kutulardan en yüksek puanlı kutuyu seçmektedir. LEF parametrelerinden α ve β , sırasıyla 5 ve 3 olarak ayarlanmıştır.

Tespit sonuçlarının başarımını değerlendirmek için PRC kullanılmıştır. Önerilen yöntemin deneysel sonuçlarını değerlendirmek için farklı IoU değerleri kullanılmıştır. IoU için 0.3, 0.5, 0.6 ve 0.7 değerleri seçilmiştir. Şekil 2.5'te farklı uzaktan algılama görüntülerinde SSCNN-1 modelinin nesne tespiti başarımının görsel sonuçları görülmektedir. Kesin sonuçlar kırmızı sınırlayıcı kutu ve SSCNN-1 tarafından tahmin edilen sınırlayıcı kutular sarı renk ile temsil edilmektedir. Şekilde, uçak, beyzbol sahası, futbol sahası ve depolama tankı nesne kategorileri için etiketler sırasıyla 1, 2, 3 ve 4 ile gösterilmiştir. Nesnelerin boyut ve yönelimindeki büyük farklılıklara rağmen,

Tablo 2.4. SS algoritmasının parametreleri

Parametre	Değer
Benzerlik	Renk (C), Doku (T), Boyut (S), Doldurma (F)
Renk uzayları	Hue, Doygunluk, and Değer (HSV) Parlaklık ve iki renk kanalı a ve b (Lab) HSV renk uzayından Hue kanalı (H) Kırmızı, yeşil ve intensite (rgI) Gri seviye görüntü intensitesi (I)
Eşik değer	100, 150, 200



Şekil 2.5. SSCNN-1 için görsel nesne tespiti sonuçları (a) uçak ve depolama tankı, (b) uçak, (c) futbol sahası ve beyzbol sahası, (d) depolama tankı, (e) beyzbol sahası, (f) futbol sahası ve beyzbol sahası, (g) uçak, (h) futbol sahası ve beyzbol sahası, (i) depolama tankı



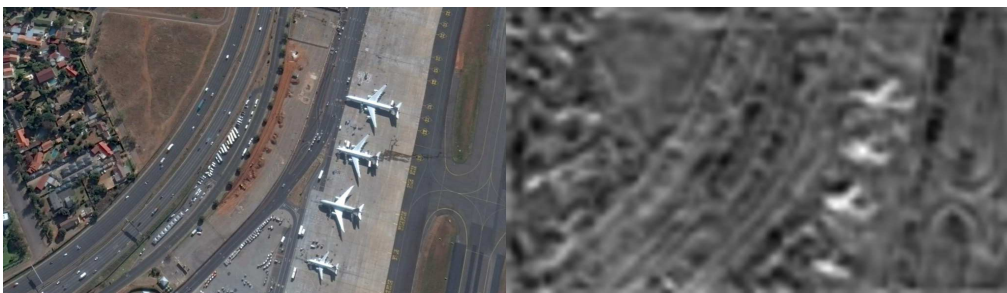
(a)



(b)



(c)



(d)

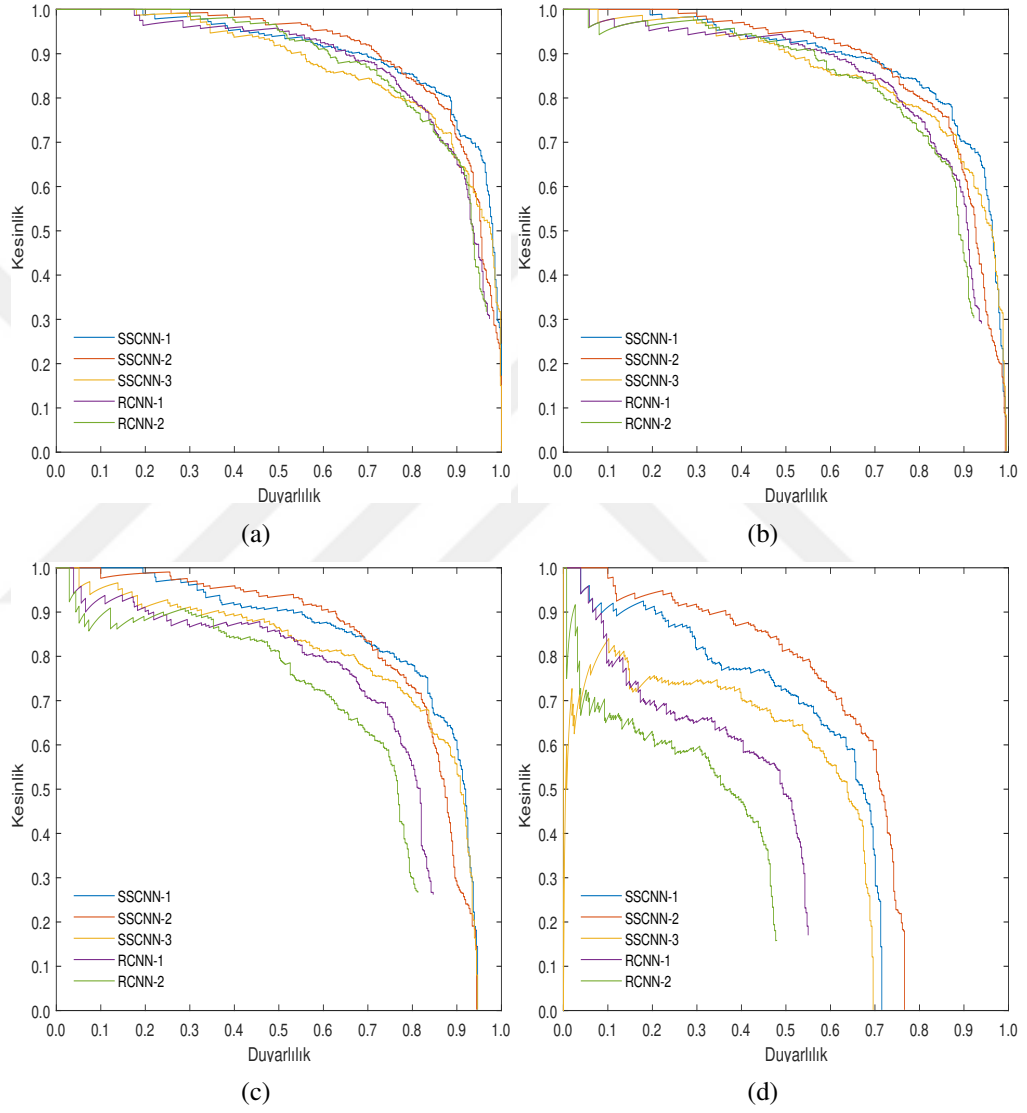
Şekil 2.6. SSCNN-1 modelinin conv5 katmanının seçilen özellik haritası ve giriş görüntüsü çiftleri (a) futbol sahası, (b) beyzbol sahası, (c) depolama tankı, (d) uçak

Tablo 2.5. ERUCV-RS4 veri seti için SSCNN ve R-CNN modellerinin AP metrik sonuçları

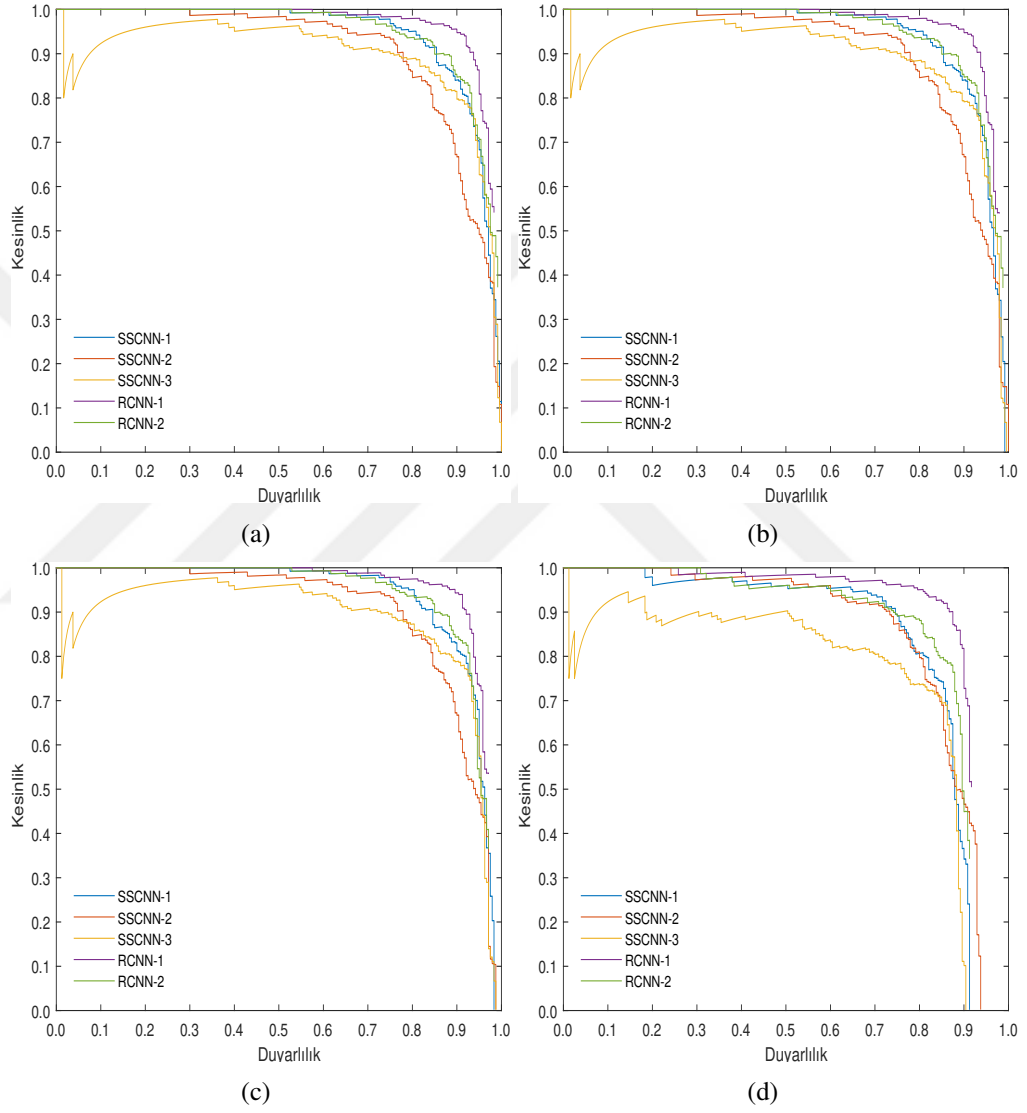
IoU	Nesne Sınıfı	SSCNN-1	SSCNN-2	SSCNN-3	R-CNN-1	R-CNN-2
0.3	Uçak	0.9035	0.9041	0.8723	0.8672	0.8655
	Beyzbol Sahası	0.9452	0.9063	0.8975	0.9631	0.9485
	Futbol Sahası	0.9717	0.9436	0.9227	0.9671	0.9531
	Depolama Tankı	0.9028	0.7745	0.9399	0.8549	0.8346
	mAP	0.9308	0.9183	0.9081	0.9130	0.9004
0.5	Uçak	0.8845	0.8724	0.8570	0.8217	0.8060
	Beyzbol Sahası	0.9419	0.9042	0.8922	0.9590	0.9452
	Futbol Sahası	0.9717	0.9434	0.9227	0.9671	0.9506
	Depolama Tankı	0.8919	0.7548	0.9279	0.8162	0.8151
	mAP	0.9225	0.8696	0.8999	0.8910	0.8792
0.6	Uçak	0.8260	0.8131	0.7805	0.6936	0.6397
	Beyzbol Sahası	0.9349	0.8998	0.8799	0.9497	0.9327
	Futbol Sahası	0.9717	0.9379	0.9227	0.9671	0.9506
	Depolama Tankı	0.8431	0.7221	0.8602	0.7874	0.7906
	mAP	0.8939	0.8432	0.8608	0.8494	0.8284
0.7	Uçak	0.5556	0.6305	0.4693	0.3755	0.2836
	Beyzbol Sahası	0.8400	0.8489	0.7554	0.8903	0.8562
	Futbol Sahası	0.9480	0.8964	0.7342	0.8652	0.8807
	Depolama Tankı	0.5274	0.4866	0.4959	0.6709	0.6683
	mAP	0.7178	0.7156	0.6137	0.7005	0.6722

önerilen yöntemin nesnelere başarılı bir şekilde tespit ettiği görülmektedir. Ayrıca, Şekil 2.6, giriş görüntüsü çiftlerini ve SSCNN-1 modelinin son evrişimli katmanı olan conv5 katmanının seçilen özellik haritalarını göstermektedir. Şekil 2.6'da CNN'lerin derin mimarisinin nesnelere son evrişimli katmanlara doğru başarılı bir şekilde ortaya çıkardığı görülmektedir.

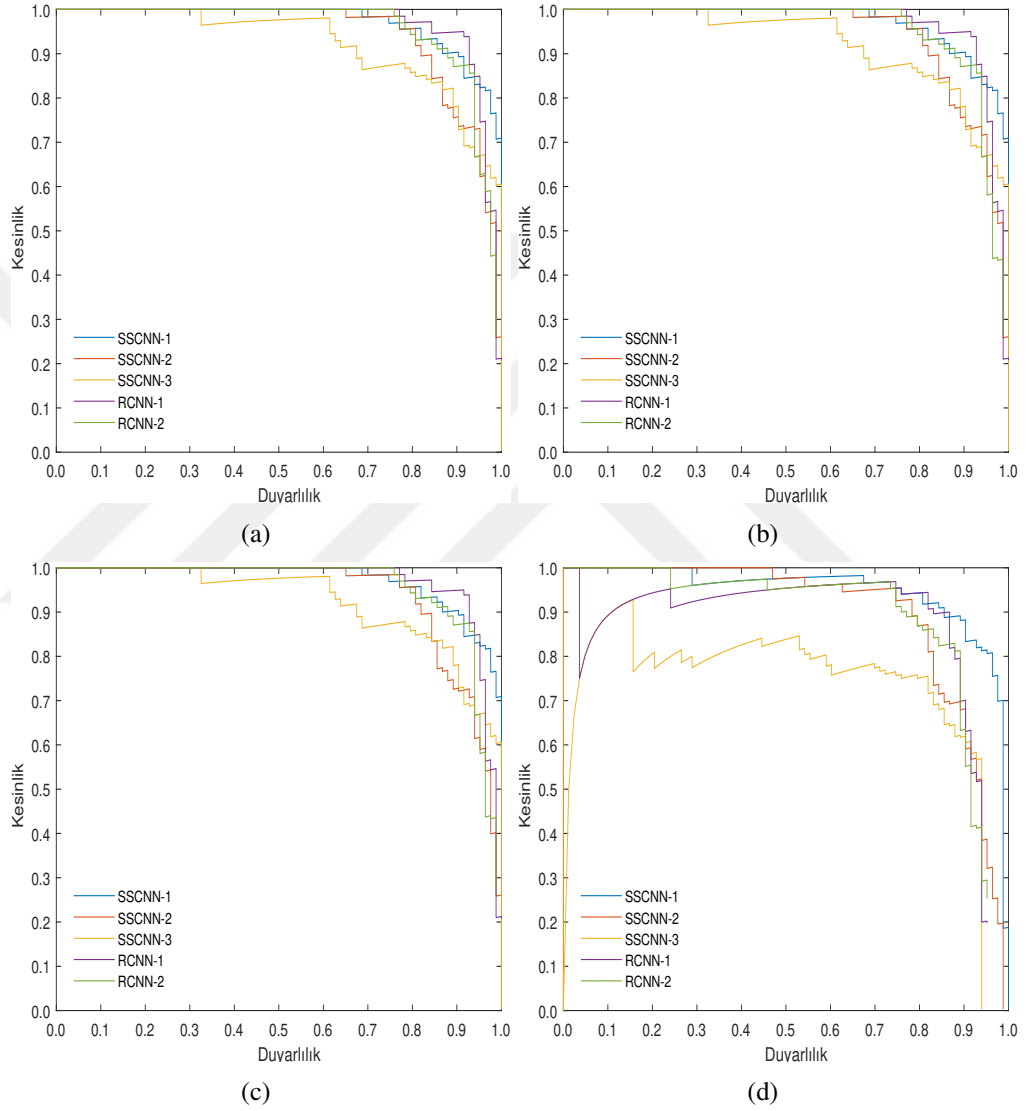
Şekil 2.7, 2.8, 2.9 ve 2.10 eğitilen modellerin farklı IoU seviyelerinde PRC eğrilerini göstermektedir. Tablo 2.5, eğitilen modellerin başarımlarını farklı IoU değerleri için AP metriği üzerinden karşılaştırmalı olarak sunmaktadır. Sonuçlar incelendiğinde önerilen yöntemin nesne sınıflarının çoğunda R-CNN'den daha iyi başarımlar gösterdiği anlaşılmaktadır. Her ağ farklı nesnelere algılanmasında öne çıkarsa da, SSCNN-1 Ortalama Kesinliklerin Ortalaması (mAP) açısından tüm IoU seviyesinde diğer tüm ağlardan daha iyi başarımlar göstermiştir. Her nesnenin kendine özgü görsel özellikleri nedeniyle, ağlar, farklı eğitim parametrelerine göre farklı nesnelere daha başarılı bir şekilde tespit başarımlarını sergilemiştir.



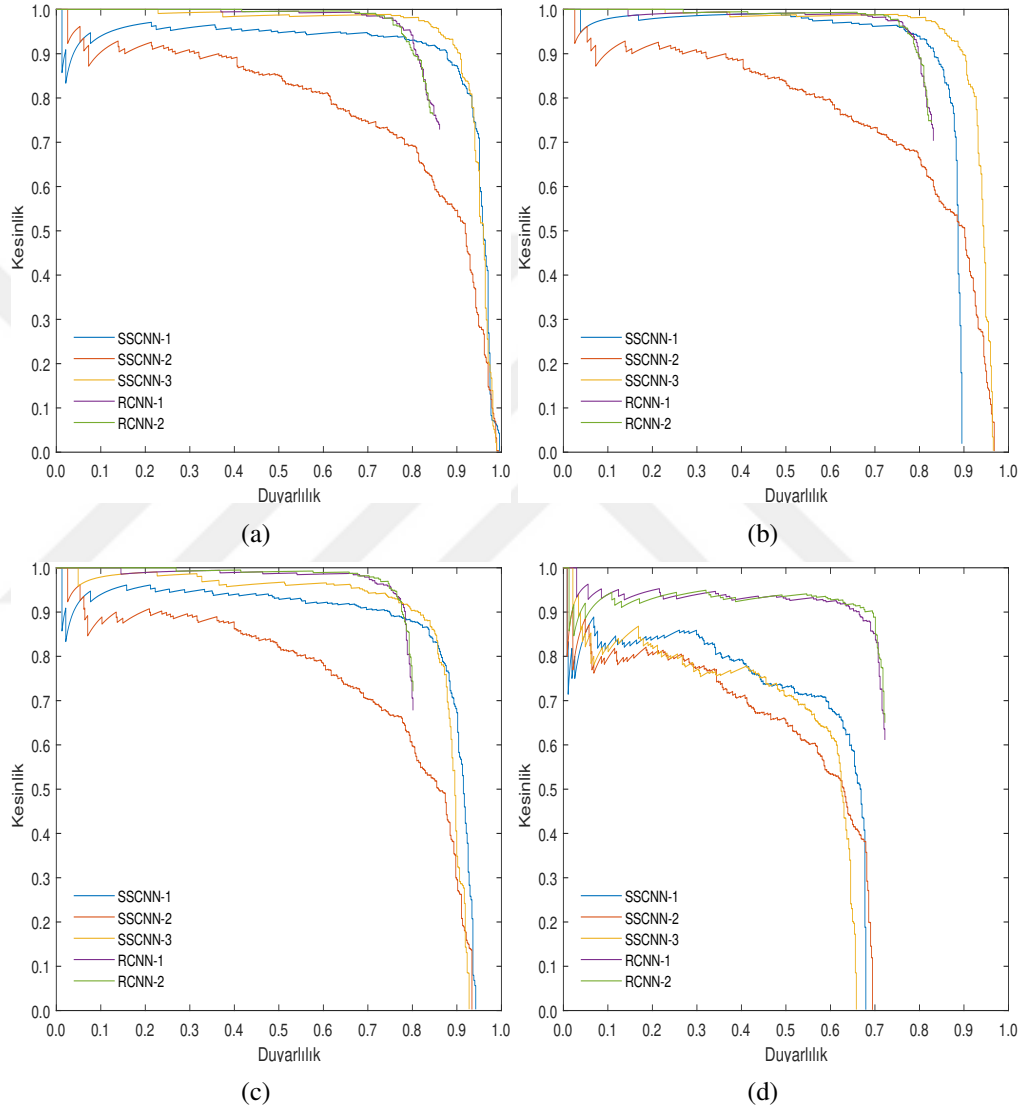
Şekil 2.7. Eğitilen modellerin uçak PRC değerleri (a) IoU=0.3, (b) IoU=0.5, (c) IoU=0.6, (d) IoU=0.7



Şekil 2.8. Eğitilen modellerin beyzbol sahaları PRC değerleri (a) IoU=0.3, (b) IoU=0.5, (c) IoU=0.6, (d) IoU=0.7



Şekil 2.9. Eğitilen modellerin futbol sahası PRC değerleri (a) IoU=0.3, (b) IoU=0.5, (c) IoU=0.6, (d) IoU=0.7



Şekil 2.10. Eğitilen modellerin depolama tankları PRC değerleri (a) IoU=0.3, (b) IoU=0.5, (c) IoU=0.6, (d) IoU=0.7

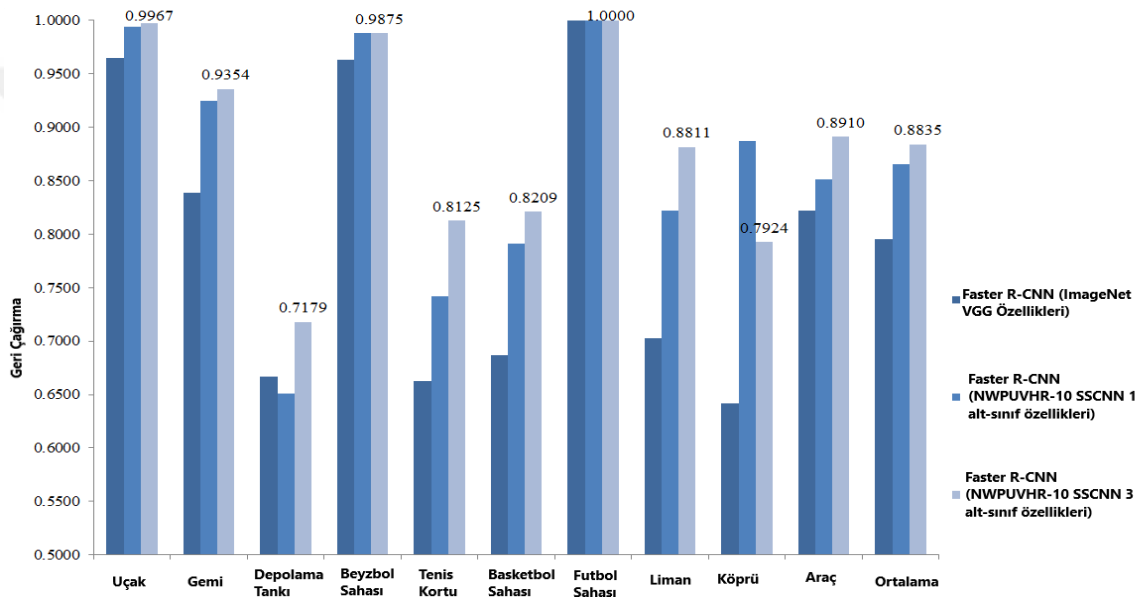
2.2.5. SSCNN Özelliklerinin Etkisinin Belirlenmesi

Alt sınıf tabanlı özelliklerin etkisini göstermek için 3 alt sınıf ile eğitilmiş SSCNN özellikleri kullanılarak Faster R-CNN'ler eğitilmiştir. Eğitim için iki aşamalı bir yöntem kullanılmıştır. İlk aşamada VGG-16 [29], 3 alt sınıf SSCNN olarak eğitilmiştir. İkinci aşamada, Faster R-CNN'ler, 3 alt sınıflı SSCNN özellikleri ile transfer eğitim yapılarak eğitilmiştir.

SSCNN özelliklerinin başarımını değerlendirmek amacıyla Faster R-CNN modeli 3 farklı şekilde eğitilmiştir. İlk olarak ImageNet [92] veri seti ile eğitilmiş VGG-16 özellikleri kullanılarak TL yapılmıştır. İkinci olarak NPWU VHR-10 veri seti ile eğitilmiş VGG-16 özellikleri kullanılarak TL yapılmıştır. Son olarak NPWU VHR-10 veri seti ile eğitilmiş SSCNN VGG-16 özellikleri kullanılarak TL yapılmıştır. Faster R-CNN modelleri, 0.001 öğrenme oranı ve 0.0005 ağırlık azalması ile 15 bin iterasyonda eğitilmiştir. ChainerCV [93] derin öğrenme kütüphanesi, Faster R-CNN modellerinin eğitimi için kullanılmıştır. Eğitilen modeller karşılaştırılarak test edilmiştir. Karşılaştırma sonuçları NPWU VHR-10 ile eğitilmiş 3 alt sınıflı VGG SSCNN özellikleri ile eğitilmiş Faster R-CNN'in en iyi nesne tespiti başarımını sergilediğini göstermektedir. Ayrıca, önerilen Faster R-CNN'lerin sonuçları, RICNN [46], PSB (ResNet101-end2endohem) [50] ve RP-Faster R-CNN [94] ile karşılaştırılmıştır. Deneyle sonuçları Tablo 2.6'da gösterilmektedir. Tablo 2.6'da, 3 alt sınıflı SSCNN özellikleri ile eğitilen Faster R-CNN'nin tüm yöntemler arasında 0.8440'lık en iyi mAP değerini elde ettiği görülmektedir. 3 alt sınıflı SSCNN özellikleri ile eğitilen Faster R-CNN, uçak, gemi, tenis kortu, liman ve araç nesne sınıfları için en iyi AP değerleri elde etmiştir. Beyzbol sahası, futbol sahası ve köprü sınıfları için, NPWU VHR-10 ile eğitilmiş VGG özellikleri ile eğitilmiş Faster R-CNN, diğer yöntemlere göre daha iyi başarımlar elde etmiştir. RICNN yöntemi depolama tankı için daha iyi sonuç vermiştir ve PSB (ResNet101-end2endohem) yöntemi basketbol sahası için diğer yöntemlere göre çok daha iyi başarımlar göstermiştir. Deneyle elde edilen sonuçlar, omurga ağının eğitim veri seti için ön eğitiminin, nesne tanımda Faster R-CNN'in başarımını artırdığını göstermektedir. Şekil 2.11'de eğitilen modellerin Duyarlılık değerleri detaylı olarak sunulmaktadır.

Tablo 2.6. NWPU-VHR10 veri seti için SSCNN ve farklı modellerin AP metrik sonuçları

Nesne Sınıfları	Faster R-CNN (ImageNet VGG özellikleri)	Faster R-CNN (NWPU VHR-10 VGG özellikleri)	Faster R-CNN (NWPU VHR-10 VGG SSCNN özellikleri)	RICNN	PSB (ResNet-101 -end2end-ohem)	R-P Faster R-CNN
Uçak	0.9587	0.9928	0.9962	0.8835	0.907	0.904
Gemi	0.7810	0.8942	0.9119	0.7734	0.806	0.750
Depolama tankı	0.5626	0.5616	0.6042	.8527	0.803	0.444
Beyzbol sahası	0.9353	0.9715	0.9606	0.8812	0.899	0.899
Tenis kortu	0.6118	0.7152	0.7828	0.4083	0.755	0.797
Basketbol sahası	0.6536	0.7152	0.7829	0.5845	0.816	0.776
Futbol sahası	0.9894	0.9986	0.9964	0.8673	0.865	0.877
Liman	0.5846	0.7366	0.8293	0.6860	0.785	0.791
Köprü	0.5420	0.8739	0.7322	0.6151	0.772	0.682
Araç	0.7337	0.7920	0.8440	0.7110	0.710	0.732
mAP	0.7353	0.8298	0.8440	0.7263	0.812	0.765



Şekil 2.11. Eğitilen Faster R-CNN modellerinin her bir nesne için Duyarlılık değerleri

2.3. Değerlendirme

Bu bölümde, CNN mimarisinin özellik çıkarma sürecini daha verimli hale getirmek ve nesnelere daha iyi temsil eden özelliklerin çıkarımını sağlamak için alt sınıf tabanlı bir eğitim stratejisi içeren yeni bir model tanıtılmıştır. Önerilen SSCNN modelinin en önemli parametresi nesnelere temsil edildiği alt sınıf sayısının belirlenmesidir. Bu bağlamda, önerilen yöntem farklı alt sınıf sayıları ile 4 sınıflı veri seti ile eğitilmiş ve en iyi alt sınıf sayısı 3 olarak belirlenmiştir. mAP açısından önerilen SSCNN'nin R-CNN ile başarımlarını karşılaştırması da gerçekleştirilmiş ve SSCNN daha başarılı sonuçlar ortaya çıkarmıştır. Ek olarak, alt sınıf tabanlı CNN mimarisinin oluşturduğu özelliklerin

etkinliđini göstermek için, Faster R-CNN ve SSCNN özellikleri, nesneleri daha doğru algılamak için entegre edilmiştir. Bu yöntemin başarımı, 10 sınıflı uzaktan algılama görüntü veri seti kullanılarak PSB, R-P-Faster R-CNN ve RICNN ile karşılaştırılmıştır. Karşılaştırma sonuçları, önerilen yöntemin diğer yöntemlerden daha iyi nesne algılama başarımına sahip olduğunu göstermiştir.



3. BÖLÜM

İNSANSIZ HAVA ARACI GÖRÜNTÜLERİNDE ARABA SAYIMI VE TESPİTİ

Bu bölümde, İHA görüntülerinde araba sayımı ve tespiti için derin öğrenme tabanlı iki yeni yöntem tanıtılmaktadır. İlk model Isı Haritası Öğrenen CNN (Heatmap Learner CNN, HLCNN) adı verilen Gaussian tabanlı nesne ısı haritası öğrenen yeni bir CNN mimarisidir. İkinci model ise HLCNN mimarisinin zayıf denetimli öğrenme yöntemi ile eğitilmesini sağlayan bir modeldir. Bu model HLCNN mimarisinin veri setinde bulunan nesnelerin sınırlayıcı kutu işaretlemesine ihtiyaç duymadan sadece merkez nokta işaretlemesi ile eğitilebilmesini sağlamaktadır.

3.1. Isı Haritası Öğrenen CNN (HLCNN) ile Araç Sayımı ve Tespiti

Nesne sayımı problemi, bir görüntüde bulunan nesnelerin sayısının belirlenmesidir. Nesne sayımı, nesne tespiti ve bölütleme problemlerinin bir alt dalı olarak tanımlanmaktadır. Nesne tespiti, bir görüntüde bulunan nesnelerin konum ve sınıfının bulunmasıdır. Nesne bölütleme ise, bir görüntünün piksel tabanlı olarak nesne bazlı sınıflandırılmasıdır. Bir görüntüde bulunan nesnelerin tespiti veya piksel tabanlı sınıflandırılması gerçekleştirilse görüntüde bulunan nesnelerin sayısı otomatik olarak tespit edilmektedir [95]. Bu nedenle nesne tespiti ve bölütleme yöntemleri nesne sayımı problemlerinin çözümü için sıklıkla kullanılmaktadır.

Nesne sayımı için önerilen ilk yaklaşımlar, pozitif tam sayı tahmininde bulunan bağlanım tabanlı yöntemlerdir. Nesne sayısını tahmin eden tek çıkışlı CNN modelleri, bu amaçla sıklıkla kullanılmaktadır. Bu CNN tabanlı yaklaşımlar, tek atışlı bağlanım modelleri olarak adlandırılmaktadır ve nesne konumlarını tespit etmeden nesne sayısını tahmin

etmektedirler. Nesne tespiti tabanlı yaklaşımlar ise bağlanım tabanlı yöntemlerden farklı olarak nesne konumlarını da tespit ederek nesnelere saymaktadır. Faster R-CNN [3], SSD [28], Yolo [96], RetinaNet [4] gibi çapa tabanlı nesne tespiti modelleri ve CornerNet [97] ve CenterNet [98] gibi çapasız nesne tespiti modelleri, birçok nesne tespiti probleminde üstün başarımlar göstermektedir [99]. Ayrıca, bu yöntemler ve varyantları, İHA görüntülerinde araba sayımında başarılı sonuçlar elde etmektedirler [8, 69, 70, 100]. Nesne konumlarını tespit etmeksizin nesne sayımını gerçekleştirmek yüzlerce farklı sınıftan nesnelere sınıflandırabilen CNN mimarileri için aslında zor bir problem değildir. Bu problem için eğitilen bağlanım tabanlı modellerin öznelik haritalarının görselleştirmeleri, öznelik haritalarının araba konumları hakkında bilgi içerdiğini göstermektedir [67]. Aich ve Stavness [67], CNN tabanlı tek atışlı bağlanım modellerinin başarımlarını iyileştirmek amacıyla CNN'nin son evrişimli katmanında oluşan öznelik haritalarını geliştirmişlerdir. CNN modellerini daha iyi özellik haritaları oluşturması amacıyla sınıf etkinleştirme haritası (Class Activation Map, CAM) görselleştirmelerini kullanarak gözetimli öğrenme ile eğitmişlerdir.

CAM, görüntü sınıflandırıcı CNN modellerinin giriş görüntüsünde karar verme sürecinde kullandığı sınıfa özgü ayırt edici bölgeleri görselleştiren bir yöntemdir. Başka bir deyişle CAM, CNN modellerinin tahminini oluştururken her bir sınıf için görüntü üzerinde hangi bölgeleri kullandığını çözümlenerek görselleştirmektedir. CAM görselleştirmeleri sınıfa özgü ayırt edici bölgeleri ısı haritası olarak sunmaktadır. Zhou ve ark. [101] CNN modellerinde Global Ortalamaya Ortaklama (Global Average Pooling, GAP) kullanarak CAM görselleştirme sürecini tanıtmaktadırlar. Bu çalışma, CNN'nin hem görüntüyü sınıflandırmak için hem de sınıfa özgü ayırt edici bölgelerin tespit edilmesini sağlamak için GAP katmanında basit bir değişiklik önermektedir. CAM görselleştirmeleri CNN modellerinin bir nesnenin görüntü üzerindeki konumunu doğru şekilde gösteren ısı haritasını oluşturabileceğini göstermektedir.

Bu bölümde tanıtılan çalışmada, CNN tabanlı tek atışlı nesne sayma ve tespiti için yeni bir yaklaşım önerilmektedir. Önerilen CNN, Gaussian fonksiyonu kullanılarak oluşturulan Gaussian Aktivasyon Haritasını (Gaussian Activation Map, GAM) tahmin etmeyi öğrenen bir bağlanım modelidir. GAM, nesnelere ısı haritasını oluşturmaktadır. Bu ısı haritaları gözetimli öğrenme ile CNN modelinin eğitimi aşamasında kullanılmaktadır. Önerilen

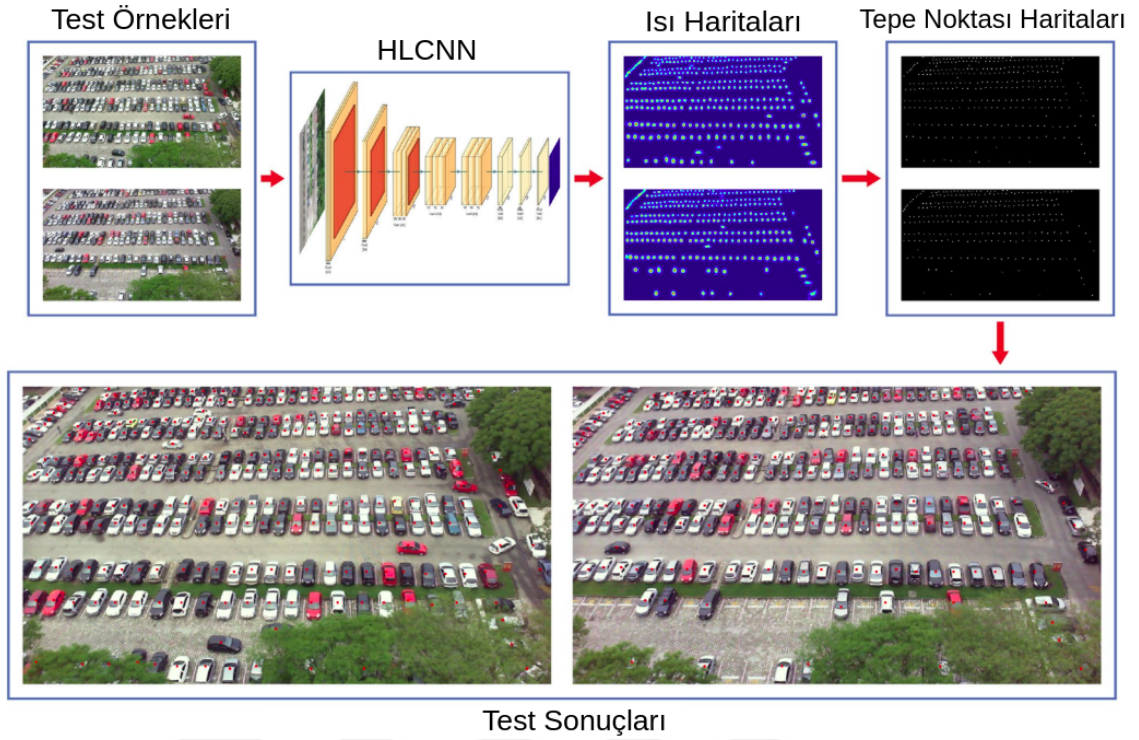
HLCNN modeli, nesne temsili için sınırlayıcı kutu kullanmaması sebebiyle çapasız nesne tespiti modelleri arasında yer almaktadır. Sınırlayıcı kutu temsilinin kutu sayısı, boyutu ve en boy oranı olmak üzere üç ana hiper parametresi bulunmaktadır [97]. Bu parametrelerin optimum şekilde ayarlanması modelin başarımını doğrudan etkilemektedir. HLCNN modelinin daha az parametreye sahip olması, eğitimini kolaylaştırmaktadır. Önerilen model, nesne sayımında pozitif tamsayı olarak nesne sayısını tahmin eden tek atışlı bağlanım modellerinden farklı olarak nesnelerin merkez noktalarını da tespit etmektedir. Bu da önerilen modelin nesne sayımı sonucunu daha açıklayıcı hale getirmektedir. Ayrıca, modelin nasıl çalıştığının ve nerede hata yaptığı da ortaya çıkmaktadır. Önerilen modelin başarımını değerlendirmek amacıyla CARPK ve PUCPR+ [8] olmak üzere iki araba sayma veri seti kullanılmıştır.

3.1.1. Önerilen HLCNN Modeli

Önerilen nesne sayımı ve tespiti yöntemi iki ana aşamadan oluşmaktadır. İlk aşamada HLCNN modeli ile ısı haritaları tahmin edilmektedir. İkinci aşamada ise tahmin edilen ısı haritası üzerinden araba konumlarını bulmak için yerel maksimumları tespit ederek arabaların merkez noktaları tespit edilmektedir. Şekil 3.1’de önerilen yöntemin genel blok şeması görülmektedir. Şekilde arabaların tespit edilen merkez noktaları, örnek çıktılarında kırmızı nokta ile gösterilmektedir.

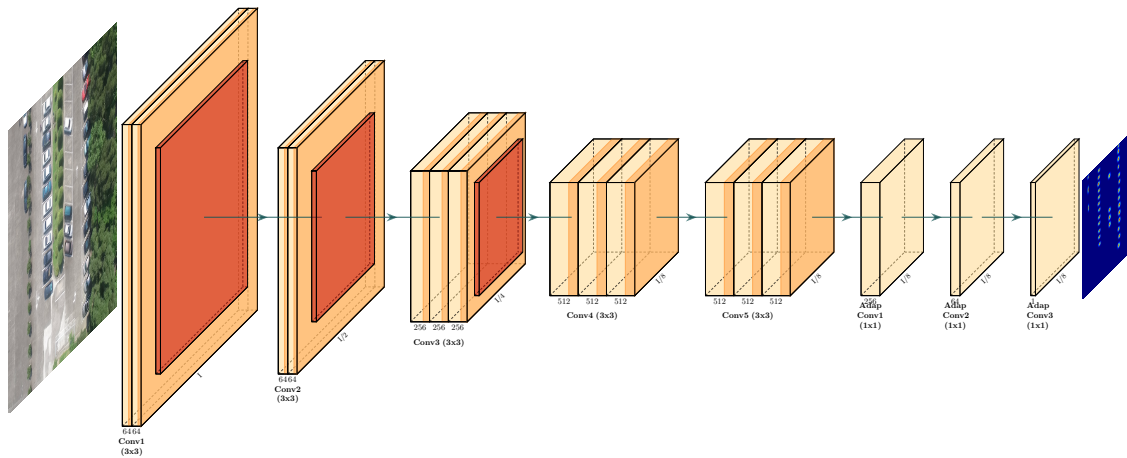
HLCNN modeli bir giriş görüntüsüne karşılık olarak nesne ısı haritası tahmini yapan CNN tabanlı bir bağlanım modelidir. Giriş görüntüsü herhangi bir boyutta olabilmektedir. HLCNN modeli giriş görüntüsüne karşılık çıktı olarak üretilen ısı haritası boyutu giriş görüntüsü boyutu, modelin alt örnekleme oranı ve sınıf sayısına bağlı olarak belirlenmektedir. Alt örnekleme oranı r olan bir model, $w \times h$ boyutlarında bir giriş görüntüsü için $w/r \times h/r \times C$ boyutlarında bir çıktı üretmektedir. Burada C sınıf sayısını temsil etmektedir. HLCNN modeli, her bir nesne sınıfı için bir 2B ısı haritası tahmin etmektedir.

HLCNN mimarisi omurga ağı ve adaptasyon katmanları olmak üzere iki bölümden oluşmaktadır. Görüntü sınıflandırıcı CNN mimarileri tam bağlı katmanların çıkartılması sonucunda omurga ağı olarak kullanılabilir. Adaptasyon katmanları, farklı sayıda



Şekil 3.1. Önerilen HLCNN modelinin genel blok şeması

öznitelik haritasına sahip 1×1 filtre boyutlu evrişim katmanlarından oluşmaktadır. Adaptasyon katmanlarında bulunan evrişim katmanlarını sızıntılı ReLU [102] aktivasyon fonksiyonu takip etmektedir. Adaptasyon katmanları, nesne kategori sayısına bağlı olarak bir dizi ısı haritası oluşturmaktadır. Ek olarak, omurga ağ çıkışından sonra yığın normalizasyonu katmanı (BN) [103] bulunmaktadır.



Şekil 3.2. Önerilen HLCNN modelinin mimarisi

Önerilen HLCNN modelinde bulunan ortaklama katmanları ve evrişim katmanlarının parametreleri alt örnekleme oranını belirlemektedir. VGG-16, ResNet-18 gibi bir

çok CNN modelinde alt örnekleme oranları 32'dir. Alt örnekleme oranının, HLCNN modelinin başarımına olan etkisi oldukça önemlidir ve sayılacak nesnenin giriş görüntüsündeki büyüklüğüne bağlı olarak uygun şekilde belirlenmelidir. Alt örnekleme oranı, modelin başarımını önemli ölçüde etkilemektedir. Bu parametre, çapasız CNN tabanlı nesne tespiti modellerinde genellikle 4 olarak belirlenmektedir. Şekil 3.2'de omurga ağ olarak VGG-16 [29] mimarisi kullanan ve 3 evrişim katmanından oluşan adaptasyon katmanı içeren örnek bir HLCNN mimarisi gözükmemektedir. Örnek olarak bu yapı 720×1280 boyutunda bir giriş görüntüsü için, tüm sınıflara özel 90×160 boyutunda ısı haritaları oluşturmaktadır. Alt örnekleme boyutunu ayarlamak için kullanılan omurğa ağın ortaklama katmanının ve evrişim katmanlarının parametreleri uyarlanmaktadır. HLCNN modelinin araba sayımında eğitimi için tahmin edilen ısı haritası (H_{pred}) ve GAM'ın L_1 kayıp hatası Eşitlik 3.1 kullanılarak geri yayılmaktadır.

$$L_1 = \frac{\sum_i^w \sum_j^h | H_{pred_{ij}} - GAM_{ij} |}{w * h} \quad (3.1)$$

Eşitlikte w ve h sırasıyla GAM'ın genişliğini ve yüksekliğini temsil etmektedir.

3.1.1.1. Tepe Nokta Haritası Oluşturma Algoritması

Bu bölümde, ısı haritasından nesnelere merkez noktalarını tespit etmek için basit bir algoritma önerilmektedir. Önerilen algoritma, tahmini bir ısı haritasını giriş olarak almaktadır ve yerel bir maksimum filtre kullanarak tepe noktalarını tespit etmektedir. Maksimum filtre, bir görüntünün yerel bir bölgesindeki en yüksek piksel değerini algılamaktadır. Önerilen tepe nokta haritası oluşturma yönteminin temel adımları Algoritma 3.1'de verilmiştir.

3.1.2. Deneysel Çalışmalar

Önerilen modelin başarımını ölçmek için iki farklı veri seti üzerinde deneysel çalışmalar yürütülmüştür. Önerilen yöntemin başarımına alt örnekleme oranının etkisini gözlemlemek amacıyla 3 farklı alt örnekleme oranı ile başarımler ölçülmüştür. Ayrıca, veri artırma tekniğinin ve BN katmanının başarıma olan etkisi detaylı şekilde araştırılarak her

Giriş: \mathbf{H} (Isı Haritası), t (*EsikDeğer*)
Çıkış: \mathbf{P} (Tepe Nokta Haritası) and **LocatedOC**(Nesnelerin merkez noktalarının koordinatları)

```

 $\mathbf{H} \leftarrow \mathbf{H} / \max(\mathbf{H})$ ; // Tahmin edilen ısı haritasının
normalizasyonu
 $\mathbf{H} \leftarrow \mathbf{H}[\mathbf{H} < t] = 0$ ; // Eşik değeri altında kalan değerlerin
sıfır olarak ayarlanması.
 $\mathbf{LM} \leftarrow \text{maximum}(\mathbf{H}) == \mathbf{H}$ ; // Maksimum filtresi ile yerel
tepe noktalarının belirlenmesi.
 $\mathbf{BG} \leftarrow \mathbf{H} == 0$ ; // Isı haritası arkaplanının elde edilmesi.
 $\mathbf{EBG} \leftarrow \text{erosion}(\mathbf{BG})$ ; // Arkaplanın morfolojik aşındırma ile
temizlenmesi.
 $\mathbf{P} \leftarrow (\mathbf{LM} \oplus \mathbf{EBG})$ ; // Özelveya fonksiyonu uygulanarak ile
tepe noktalarının hesaplanması
LocatedOC  $\leftarrow \mathbf{P} == 1$ ; // Tepe noktası koordinatlarının elde
edilmesi.

```

Algoritma 3.1: Önerilen tepe nokta haritası oluşturma algoritmasının sözde kodları

bir veri seti için sunulmuştur.

3.1.2.1. Veri Setleri

Önerilen araba sayımı ve tespiti modelinin başarımını ölçmek amacıyla PUCPR+ ve CARPK olmak üzere iki farklı araba sayımı ve tespiti veri seti kullanılmıştır.

PUCPR+ veri seti PUCPR veri setinin uyarlanmış bir versiyonudur. Veri seti bir otopark alanında bulunan araba görüntülerinden oluşmaktadır. PUCPR veri seti, yüksek bir konuma sabitlenmiş bir kamerayla çekilen görüntülerden oluşmaktadır. Bu görüntüler birbirine çok benzemektedir. Bu nedenle PUCPR veri seti derin öğrenme modellerini eğitmek için uygun gözükmemektedir. Hsieh ve ark. [8] tarafından PUCPR veri setinin yeniden düzenlenmesi ile PUCPR+ veri seti oluşturulmuştur. PUCPR+ veri seti toplamda 17000 araba içermektedir.

CARPK veri seti 2017 yılında Heish ve ark. [8] tarafından tanıtılmıştır. Bu veri seti, 4 farklı otoparktan çekilmiş farklı sahneler içeren 1448 adet görüntüden oluşmaktadır. Görüntülerde toplamda 98777 araba bulunmaktadır. Görüntüler İHA kullanılarak yaklaşık olarak 40 metre yükseklikten çekilmiştir. Veri setinde bulunan 989 görüntü eğitim, 459 görüntü ise test amaçlı kullanılmaktadır. Şekil 3.3'te CARPK veri setinden örnek bir görüntü gözükmektedir.



Şekil 3.3. CARPK veri setinden örnek görüntü [8]

3.1.2.2. Değerlendirme Kriterleri

Önerilen modelin nesne sayımında başarımını sayısal metriklerle nicel olarak değerlendirmek amacıyla Ortalama Mutlak Hata (Mean Absolute Error, MAE) ve Kök Ortalama Kare Hatası (Root Mean Square Error, RMSE) kullanılmıştır. MAE, iki sürekli değişken arasındaki farkın değerini ölçmektedir. MAE, bağlanım problemlerinde sıkça kullanılmaktadır. MAE ve RMSE sırasıyla Eşitlik 3.2 ve 3.3 kullanılarak hesaplanmaktadır.

$$MAE = \frac{\sum_i |x_i - y_i|}{N} \quad (3.2)$$

$$RMSE = \sqrt{\frac{\sum_i (x_i - y_i)^2}{N}} \quad (3.3)$$

Eşitliklerde N görüntü sayısını, x_i i . görüntüde işaretlenen araba sayısını ve y_i i . görüntüde tahmin edilen araba sayısını temsil etmektedir.

3.1.2.3. CARPK Veri Seti Sonuçları

Önerilen modelin CARPK veri setinde başarımını ölçmek için 4, 8 ve 16 olmak üzere 3 farklı alt örnekleme oranına sahip HLCNN modelleri eğitilmiştir. Alt örnekleme oranının ayarlanması için VGG-16 mimarisinin ortaklama katman parametreleri düzenlenmiştir. Alt örnekleme oranı 4 olan modelde, VGG-16 mimarisinin ilk iki ortalama ortaklama katmanı hariç diğer tüm maksimuma ortalama katmanlarının adım aralığı ve filtre boyutu parametreleri 1 olarak ayarlanmıştır. Tüm modeller Stokastik Optimizasyon İçin Bir Yöntem (Adam: A Method for Stochastic Optimization, ADAM) optimizasyon algoritması kullanılarak 30 epok boyunca eğitilmiştir. İlk öğrenme oranı 0.0001 olarak ayarlanmıştır. 15. ve 25. epokta öğrenme oranı 0.1 çarpanı ile küçültülmüştür. Veri seti görüntüleri 540×960 boyutuna dönüştürülerek eğitilmiştir. Alt örnekleme oranının başarıma etkisini belirlemek için eğitilen 3 modelin test sonuçları Tablo 3.1'de verilmiştir. Alt örnekleme oranı 8 olan model, MAE ve RMSE metriklerinde sırasıyla 2.12 ve 3.02

skorları ile başarımı en yüksek model olmuştur. Bu model 720×1280 boyutundaki giriş görüntüsü için 90×160 boyutunda ısı haritası oluşturmaktadır. Tepe nokta haritası oluşturma aşamasında t parametresi 0.05 olarak ayarlanmıştır.

Tablo 3.1. HLCNN modelinin alt örnekleme oranının CARPK veri setinde araba sayımı başarımına etkisi

Alt örnekleme oranı	MAE	RMSE
4	5.29	7.15
8	2.12	3.02
16	6.00	9.80

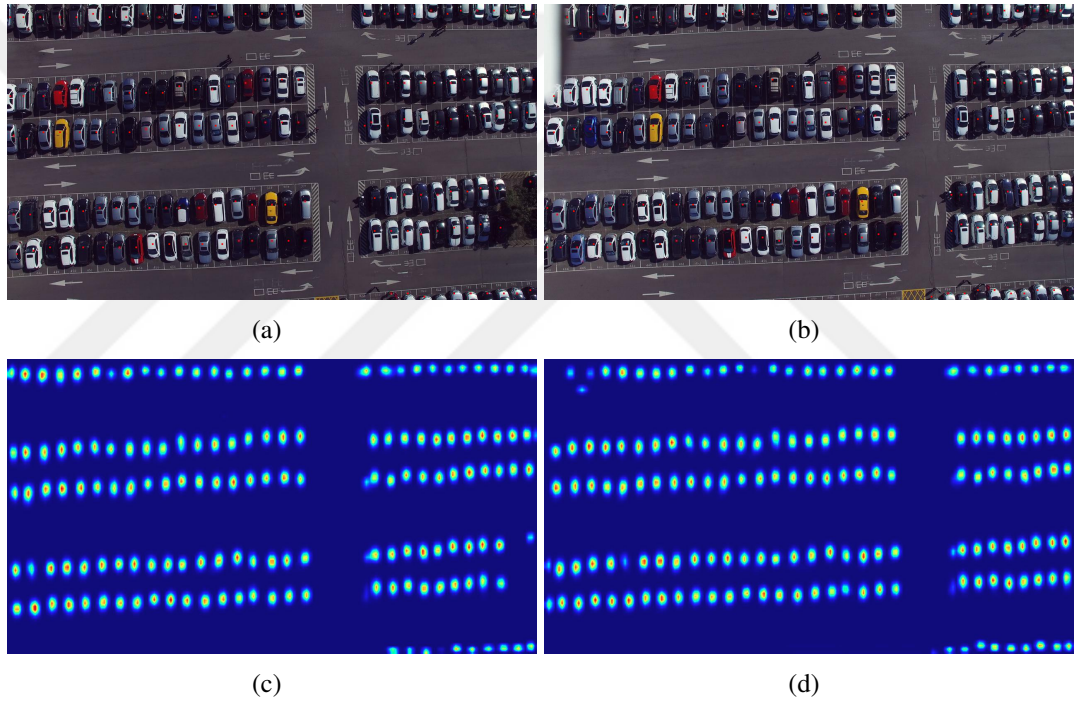
Modellerin eğitimi esnasında görüntü döndürme ve renk değiştirme veri artırma teknikleri kullanılmıştır. Renk değiştirme veri artırma yöntemi 0.5 olasılıkla görüntünün Yoğunluk kanalında rastgele değişimler ile uygulamaktadır. Görüntü döndürme veri artırma yönteminde 0.2 olasılıkla görüntü 0 ila 45 derecelik rastgele belirlenen bir açıyla döndürülmektedir.

Veri artırma tekniklerinin ve BN katmanının başarımına etkisini ölçmek için alt örnekleme oranı 8 olan model, veri artırma ve BN katmanı kullanılarak veya kullanılmayarak eğitilmiştir. Elde edilen sonuçlar Tablo 3.2’de gösterilmiştir. İlk olarak HLCNN-WOBNAU adı verilen model, veri artırma ve BN katmanı olmadan eğitilmiştir. HLCNN-WOBNAU, MAE ve RMSE metrikleri ile sırasıyla 3.19 ve 4.33 hata oranlarına ulaşmıştır. İkinci olarak HLCNN-WOBN adı verilen model, BN katmanı kaldırılarak eğitilmiştir. HLCNN-WOBN, MAE metriği ile 2.61 ve RMSE metriği ile 3.52 hata oranına ulaşmıştır. Son olarak HLCNN-WOAU adı verilen model, veri artırma yapılmadan eğitilmiştir. HLCNN-WOAU, MAE ve RMSE metrikleri ile sırasıyla 2.73 ve 3.93 hata oranlarına ulaşmıştır. Elde edilen sonuçlar, veri artırmanın ve BN katmanının modelin başarımına olumlu katkılar sunduğunu göstermiştir. En iyi başarımları veri artırma ve BN katmanının ortak kullanıldığı model elde etmiştir.

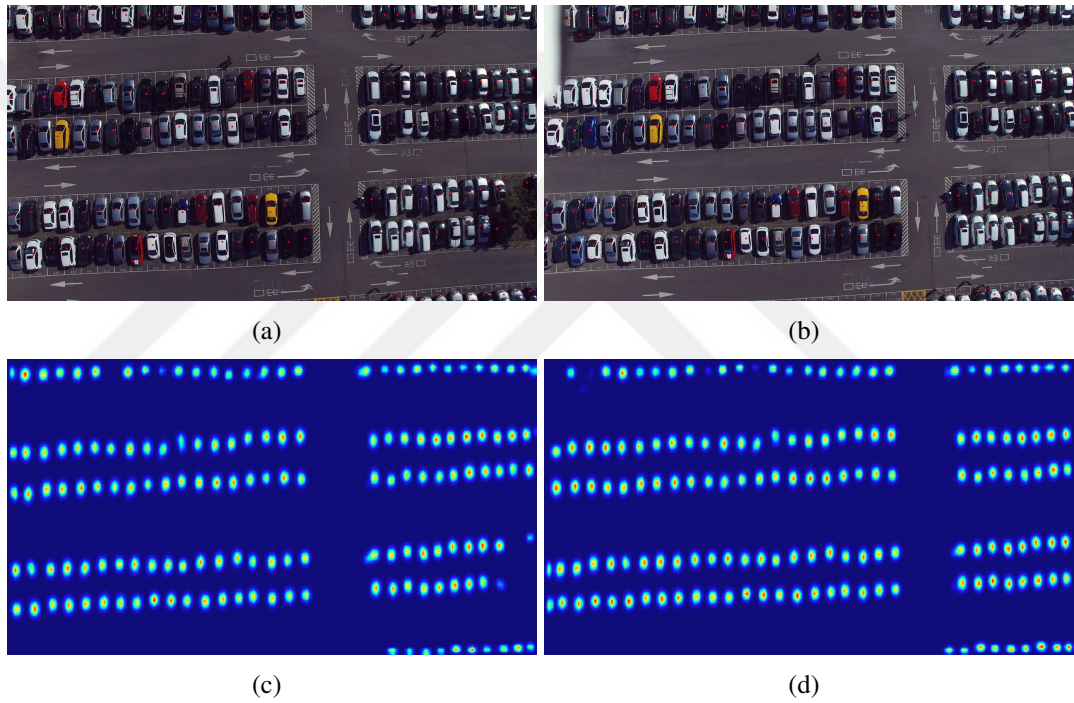
Tablo 3.2. Önerilen HLCNN modelinin için veri artırmanın ve BN katmanının başarımları CARPK veri setinde araba sayımı başarımına etkisi

Model	MAE	RMSE
HLCNN-WOBNAU	3.19	4.33
HLCNN-WOAU	2.73	3.93
HLCNN-WOBN	2.61	3.52
HLCNN	2.12	3.02

Veri artırmanın ve BN katmanının başarımına olan etkileri görsel sonuçların detaylı olarak analizi ile değerlendirilmiştir. Değerlendirme sonucunda renk tabanlı veri artırma işleminin, araba gölgelerinden dolayı yapılan yanlış tespitleri azalttığı görülmüştür. HLCNN ve HLCNN-WOBN modellerinin aynı giriş görüntüsüne verdikleri çıktılar karşılaştırılarak renk tabanlı veri artırmanın olumlu etkisi görülmüştür. Bu modellerin ilgili görsel çıktıları Şekil 3.4 ve Şekil 3.5'te verilmiştir. Şekillerde kırmızı noktalar tespit edilen merkez noktaları göstermektedir. Bu şekiller karşılaştırmalı olarak incelendiğinde, veri artırmanın araç gölgelerinin neden olduğu yanlış tespit sayısını azalttığı görülmektedir.

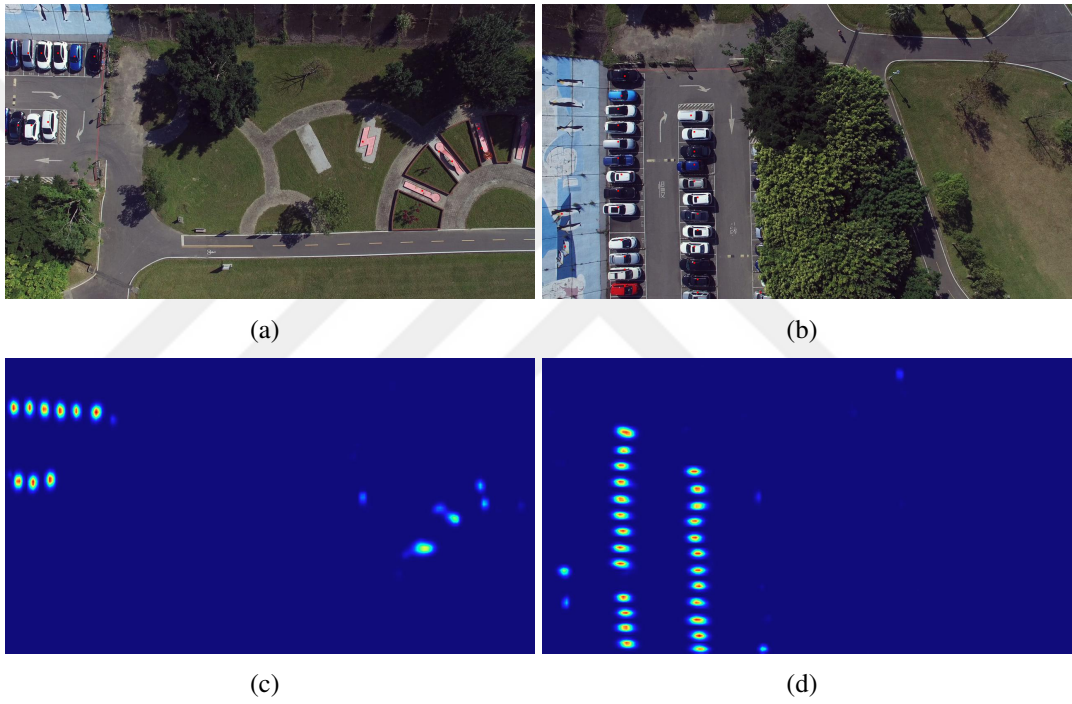


Şekil 3.4. Renk tabanlı veri artırma işleminin etkisini gözlemlemek amacıyla HLCNN-WOBN modeli için tahmin edilen ısı haritaları (a), (b) Örnek görüntüler, (c), (d) ısı haritaları

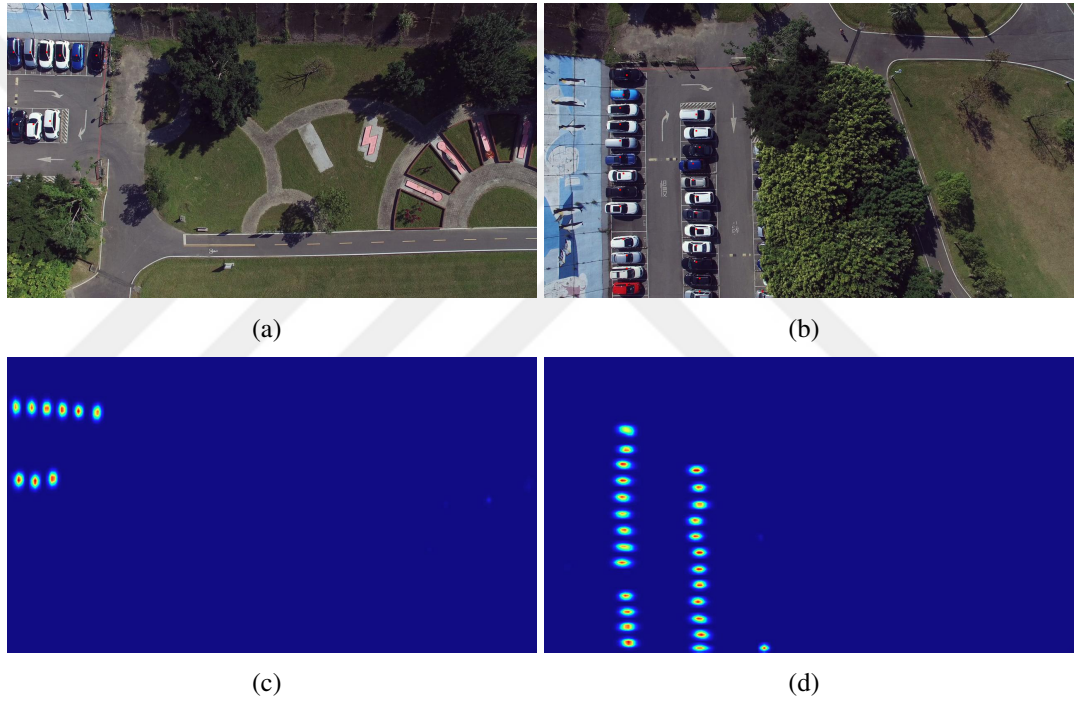


Şekil 3.5. Renk tabanlı veri artırma işleminin etkisini gözlemlemek amacıyla HLCNN modeli için tahmin edilen ısı haritaları (a), (b) Örnek görüntüler, (c), (d) ısı haritaları

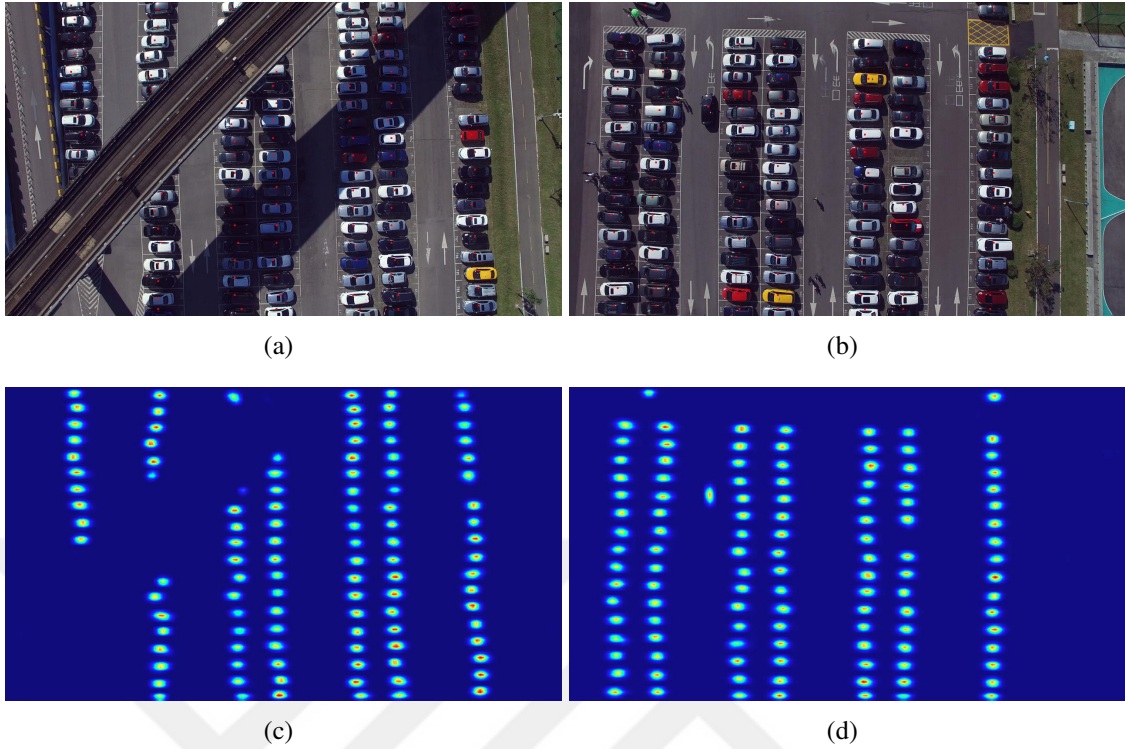
BN katmanı ve veri artırma işlemini, önerilen yöntemin yanlış tespit oranını düşürmüştür. Şekil 3.5 ve 3.6'da HLCNN-WOBNAU ve HLCNN modellerinin aynı giriş görüntüsüne verdikleri çıktılar sunulmuştur. Şekillerde sunulan görsel sonuçlar karşılaştırıldığında grafiti ve park alanında bulunan oturma alanı ve park oyuncakları gibi yanlış tespit edilen nesnelere yapılan hataların azaldığı görülmektedir. Ayrıca, önerilen modelin en çok hata yapmasına neden olan görüntü özelliği kısmen ağaçların altında kalan arabalar ve köprü gölgesi gibi alanlarda bulunan arabaların bulunduğu görüntülerdir. Özellikle kısmen görülen arabalar tespit edilemeyen hedef nesnelere oluşturmaktadır.



Şekil 3.6. HLCNN-WOBNAU modeli için tahmin edilen ısı haritaları (a), (b) Farklı örnek görüntüler, (c), (d) ısı haritaları



Şekil 3.7. HLCNN modeli için tahmin edilen ısı haritaları (a), (b) Farklı örnek görüntüler, (c), (d) ısı haritaları



Şekil 3.8. CARPK veri seti için HLCNN modeli ile elde edilen sonuçlar , (a), (b) Örnek görüntüler, (c), (d) ısı haritaları

Önerilen HLCNN modelinin literatürde iyi sonuç almış yöntemlerle başarımlarını karşılaştırmaları Tablo 3.3'de sunulmuştur. Karşılaştırma sonuçları önerilen yöntemin MAE ve RMSE metriklerinde en iyi skoru elde ettiğini göstermektedir. Şekil 3.8'de önerilen model ile elde edilmiş görsel sonuçlar verilmiştir. Şekilde test görüntüleri üstte, önerilen modelin tahmin ettiği ısı haritaları altta sunulmuştur.

3.1.2.4. PUCPR+ Veri Seti Sonuçları

Önerilen HLCNN modelinin başarımını farklı bir veri setinde ölçmek amacıyla literatürde sıklıkla kullanılan PUCPR+ veri seti kullanılmıştır. Bu veri setinde yapılan çalışmalar için omurga ağ olarak VGG-16 kullanılmıştır. Alt örnekleme oranının başarıma etkisini araştırmak için 4, 8 ve 16 alt örnekleme oranına sahip HLCNN mimarileri eğitilerek test edilmiştir. Tüm modellerin eğitimi 100 epokta gerçekleştirilmiştir. İlk öğrenme oranı 0.0001 olarak ayarlanmıştır. 60. ve 80. epokta öğrenme oranı 0.1 çarpanı ile küçültülmüştür. Tepe nokta haritası oluşturma aşamasında t parametresi 0.45 olarak ayarlanmıştır.

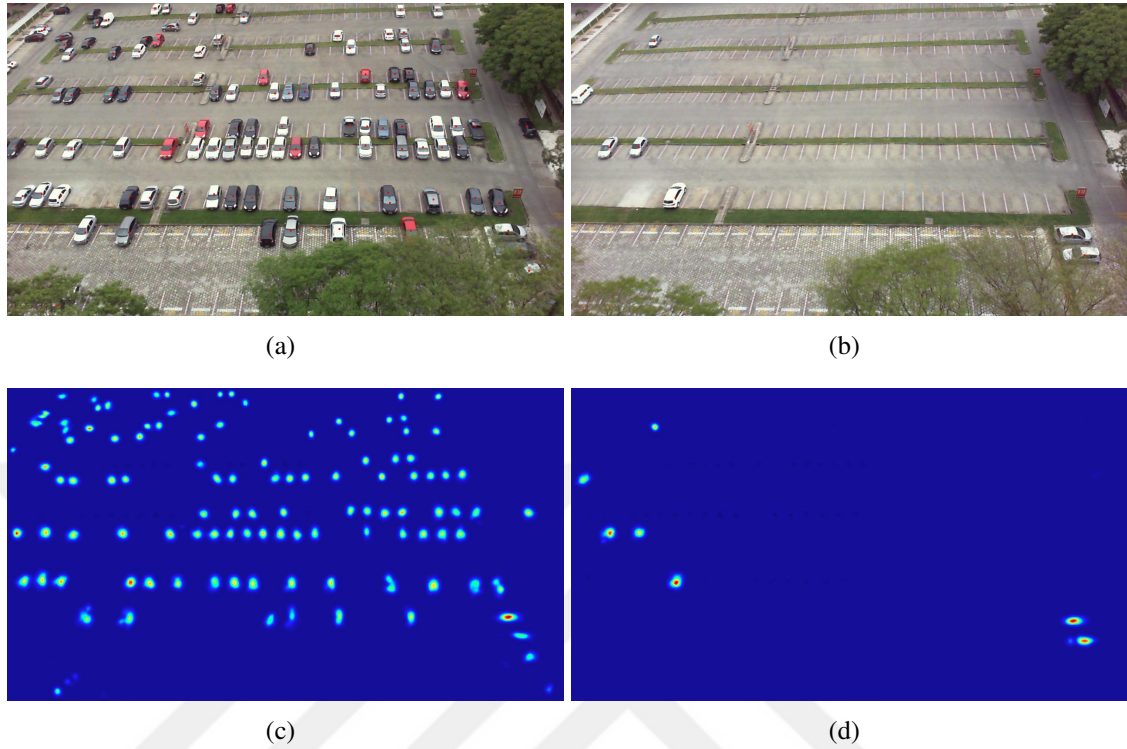
Tablo 3.3. CARPK veri seti için önerilen HLCNN modelinin karşılaştırma sonuçları

Yöntem	MAE	RMSE
Faster R-CNN [3, 8]	24.32	37.62
YOLO [8, 96]	48.89	57.55
One-Look Regression [8, 104]	59.46	66.84
LPN [8]	23.80	36.79
RetinaNet [4]	16.62	22.30
IEP Counting [105]	51.83	
IoUNet [70]	6.77	8.52
YOLOv3 [73]	7.92	11.08
VGG-GAP-HR [67]	5.88	9.30
GSP-224 [106]	5.46	8.09
SA+CF+CRT [100]	5.42	7.38
GANet (VGG-16) [69]	4.80	6.9
Arruda ve ark. [107]	4.45	6.18
HLCNN (VGG-16)	2.12	3.02

Tablo 3.4. HLCNN modelinin alt örnekleme oranının PUCPR+ veri setinde araba sayımı başarımına etkisi

Alt örnekleme oranı	MAE	RMSE
4	2.52	3.40
8	2.64	4.05
16	29.92	41.62

Alt örnekleme oranının etkisini ölçmek için yapılan deneylerin sonuçları Tablo 3.4’de sunulmuştur. Alt örnekleme oranı 4 ile eğitilen HLCNN model, RMSE metriği ile 3.40 hata oranı değeri elde etmiştir. Önerilen model literatürde bulunan bir çok modelle karşılaştırılmıştır ve sonuçlar Tablo 3.5’de sunulmuştur. Tablodaki sonuçlar önerilen yöntemin literatürde bilinen en iyi başarımı elde ettiğini göstermektedir. Ayrıca, Şekil 3.9’da önerilen model ile elde edilen sonuçlar verilmiştir.



Şekil 3.9. PUCPR+ veri seti için HLCNN model ile elde edilen sonuçlar (a), (b) Örnek görüntüler, (c), (d) ısı haritaları

Tablo 3.5. PUCPR+ veri seti için HLCNN modelinin karşılaştırma sonuçları

Yöntem	MAE	RMSE
Faster R-CNN [3, 8]	39.88	47.67
YOLO [8, 96]	156.00	200.42
One-Look Regression [8, 104]	21.88	36.73
LPN [8]	22.76	34.46
RetinaNet [4]	24.58	33.12
IEP Counting [105]	15.17	
IoUNet [70]	7.16	12.00
VGG-GAP-HR [67]	5.24	6.67
YOLOv3 [73]	5.24	7.14
SA+CF+CRT [100]	3.92	5.06
GANet (VGG-16) [69]	3.68	5.47
Arruda ve ark. [107]	3.16	4.39
HLCNN (VGG-16)	2.52	3.40

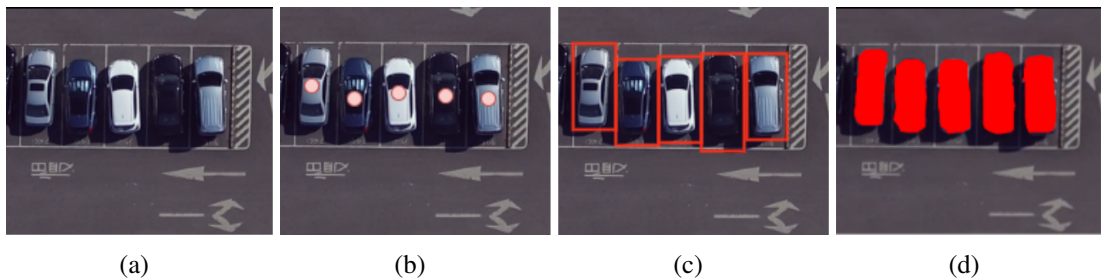
3.1.3. Değerlendirme

Bu bölümde, İHA görüntülerinden araba sayımı ve tespiti amacıyla yeni bir CNN tabanlı model önerilmiştir. Önerilen model, görüntüde bulunan arabaların sayısını belirlemenin yanında merkez noktalarını da tespit etmektedir. HLCNN olarak adlandırılan model giriş görüntüsü üzerinden araba ısı haritası tahmin etmektedir. HLCNN modeli, ısı haritası bağlanımı gerçekleştiren tek atışlı CNN modeli olarak tanımlanmaktadır. Literatürde CNN tabanlı nesne sayısı bağlanımı gerçekleştiren yöntemler nesne sayımı için sıklıkla kullanılmaktadır. Önerilen modeli literatürde bulunan bu yöntemlerden ayıran en büyük özellik nesne sayısının bağlanımı yerine ısı haritası bağlanımı gerçekleştirmesidir. Bu sayede, nesne sayısını tahmin etmesinin yanında nesne konumlarını da tespit etmektedir. Önerilen yöntemin başarımı, veri artırma tekniklerinin ve BN katmanının etkisini ölçmek için detaylı analiz yapılarak sunulmuştur. BN katmanı ve veri artırma tekniklerinin başarımına olumlu etkisi görülmüştür. Önerilen modelin literatürde sıklıkla kullanılan CARPK ve PUCPR+ araba veri setleri üzerinde başarımı değerlendirilmiştir. Elde edilen sonuçlar, önerilen yöntemin literatürde bilinen en iyi araba sayma başarımına ulaştığını göstermiştir. HLCNN modelinin alt örnekleme oranına olan bağımlılığı yöntemin en büyük dezavantajı olarak görülmüştür. Yapılan deneylerde alt örnekleme oranının 4 veya 8 seçilmesinin uygun olduğu değerlendirilmiştir.

3.2. Merkez Nokta İşaretleme Tabanlı HLCNN ile Araç Sayımı ve Tespiti

CNN tabanlı nesne sayımı ve tespiti yöntemleri genellikle denetimli öğrenme yaklaşımı ile eğitilmektedir. Bölüm 3.1’de önerilen HLCNN modeli de denetimli öğrenme yaklaşımı ile eğitilmektedir. Bu yaklaşımla eğitilen modeller çözülen probleme özgü olarak farklı etiketleme ve işaretleme bilgilerine ihtiyaç duymaktadır. Özellikle nesne tespiti ve bölütlemesi problemlerinde görüntüde bulunan nesnelerin piksel ya da çerçeve tabanlı işaretleme ihtiyacı duyulmaktadır.

Görüntü etiketleme genellikle 4 farklı şekilde yapılmaktadır. Bunlar; düşük seviyeden yüksek seviyeye doğru olmak üzere görüntü etiketleme, nokta işaretleme, çerçeve işaretleme ve piksel işaretlemedir. Görüntü etiketlemede, sadece görüntüde bulunan nesnelerin sınıfları kaydedilmektedir. Nokta işaretlemede, görüntüde bulunan nesnelerin sınıfları ve nesnenin merkez noktasının koordinatları kaydedilmektedir. Çerçeve işaretlemede, görüntüde bulunan nesnenin sınıfları ve nesnenin boyutlarını 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 işaretlemede, görüntüde bulunan her piksel için bir nesne sınıfı kaydedilmektedir. Şekil 3.10’da 4 farklı görüntü etiketleme yöntemi için örnekler görülmektedir. Düşük seviyeden yüksek seviyeye ilerledikçe görüntü etiketleme ve işaretlemenin zaman maliyeti artmaktadır. Görüntü etiketleme de nesne başına 1 saniye harcanırken bu süre nokta işaretleme için 2.4 saniye, çerçeve işaretleme için 10 saniye, piksel işaretleme için ise 78 saniyedir [108]. Bu sebeple, düşük seviye görüntü etiketlemesi ile eğitilebilen başarılı nesne sayımı yöntemlerine ihtiyaç duyulmaktadır.



Şekil 3.10. Farklı görüntü işaretleme yöntemleri örnekleri (a) araba, (b) araba [merkez nokta], (c) araba (çerçeve), (d) araba (piksel)

Bu bölümde yapılan çalışma HLCNN modelinin nokta işaretleme ile eğitilmesidir. Bölüm 3.1’de HLCNN modeli çerçeve etiketleri üzerinden eğitilmiştir ve araba sayısını belirlemekte oldukça başarılı olmuştur. Bu bölümde ise sadece merkez nokta işaretlemeleri üzerinden eğitilen HLCNN mimarisinin araba sayımı başarımı test edilmektedir. Bu şekilde eğitilen model çerçeve etiketleme ile eğitilen modele benzer şekilde görüntüde bulunan araba sayısı ve arabaların merkez noktalarını tahmin etmektedir. Bu bölümde yapılan çalışmada HLCNN mimarisi nesne ısı haritası yerine nesne merkez nokta haritası tahmin etmektedir. HLCNN modelinin bu merkez nokta haritasını tahmin edebilir şekilde eğitilmesi amacıyla yeni bir kayıp fonksiyonu tanımlanmıştır. Bu bölümde önerilen yöntemin başarımını ölçmek amacıyla CARPK veri seti kullanılmıştır.

3.2.1. Önerilen Merkez Nokta İşaretleme Tabanlı HLCNN Modeli

Önerilen model 2 temel aşamadan oluşturmaktadır: İlk aşamada, giriş görüntüsü üzerinden nesne ile arka planı ayırt eden bir nesne merkez nokta harita oluşturulmaktadır. İkinci aşamada ise, tahmin edilen haritadan nesne konum ve sayılarını tespit edilmektedir. Nesne merkez nokta haritasını tahmin etmek amacıyla HLCNN mimarisi kullanılmıştır. HLCNN modelinin eğitimi için nokta etiketleme ile oluşturulan kesin referans haritaları kullanılmaktadır. Nesne haritasında arka plan 0 ile temsil edilmektedir. Nesnelere ise 255 değeri ile temsil edilmektedir. Nesnelere merkez noktaları ve istenilen kadar komşu pikselleri 255 değeri ile doldurularak nesne haritaları oluşturulmaktadır.

3.2.1.1. Merkez Nokta İşaretleme Tabanlı HLCNN Modelinin Eğitimi

Merkez nokta işaretleme tabanlı HLCNN modelinin eğitiminin gerçekleştirilmesi amacıyla model hatasını arka plan ve ön plan pikselleri için ayrı ayrı hesaplayan iki parçalı bir kayıp fonksiyonu kullanılmıştır. Genellikle görüntülerde nesnelere arka planlara göre çok az yer kaplamaktadır. Arka plan ve ön plan piksel sayılarındaki bu dengesizlik, derin öğrenme modellerinin eğitimini zorlaştırmaktadır. Eğitilen modellerin tarafı 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 sistemi arka plan sınıfını daha fazla tahmin etme eğilimine itmektedir. Bunun önüne geçmek amacıyla 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 Eşitlik 3.4, 3.5 ve 3.6'da verilmektedir.

$$L_{bg} = \frac{\sum_i^n |x_i - y_i|}{n} \quad (3.4)$$

$$L_{fg} = \frac{\sum_i^{k \times m} |x_i - y_i|}{k} \quad (3.5)$$

$$L = L_{bg} + L_{fg} \quad (3.6)$$

Eşitliklerde, 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 hatasının 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.

3.2.2. Deneysel Çalışmalar

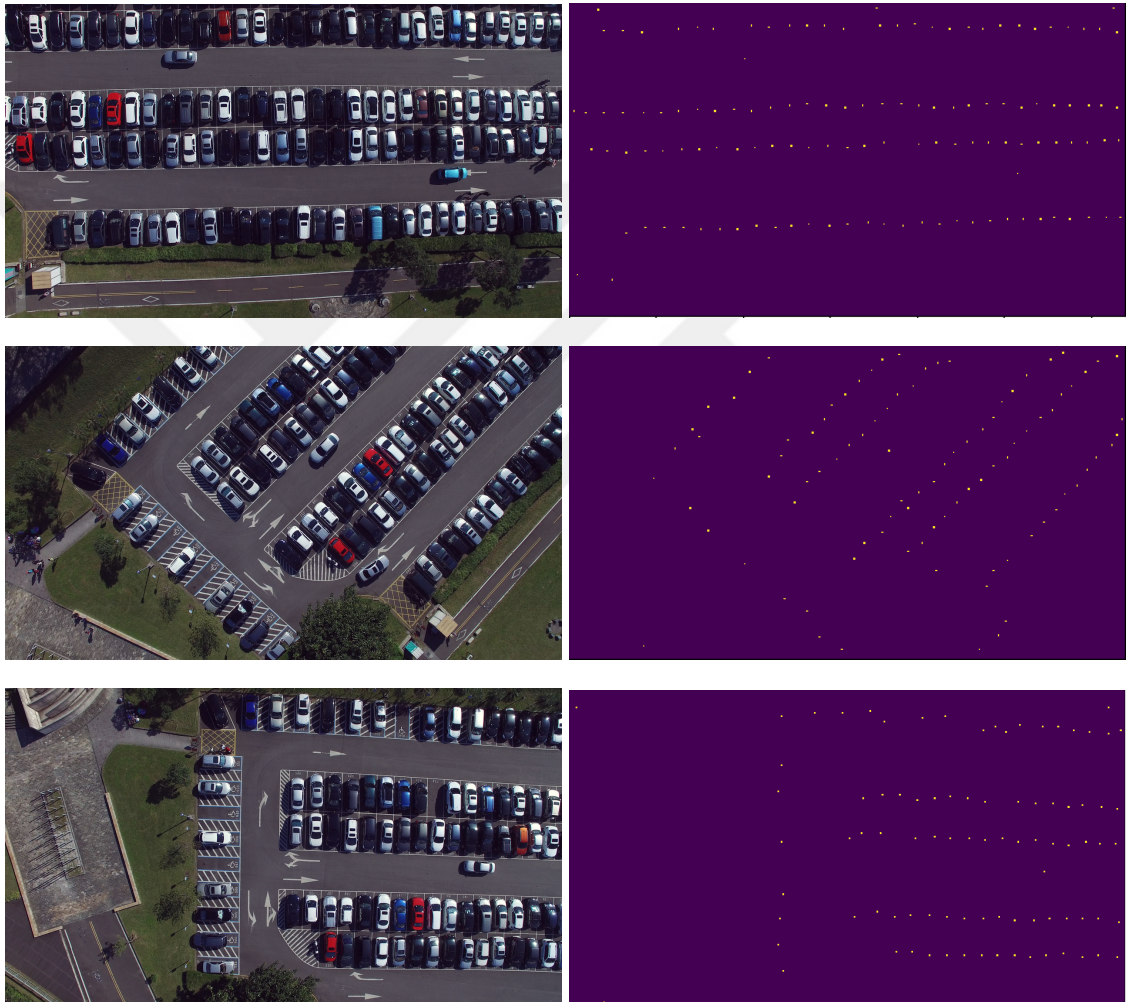
Önerilen yöntemin başarımının değerlendirilmesi için CARPK veri seti üzerinde yapılan deneysel çalışmalarda omurga ağı olarak ResNet-50 mimarisi kullanılmıştır. Omurga ağını takip eden adaptasyon katmanında 3 adet evrişim katmanı bulunmaktadır. Bu katmanlar omurga ağından gelen 512 adet özellik haritasını sırasıyla 64, 16 ve 1'e indirmektedir. Model için alt örnekleme oranı 4 olarak belirlenmiştir. Eğitim esnasında

görüntüler 540×960 boyutuna indirilmiştir. 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 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 gerçekleştirilmiştir. Bu sayede sistemin hatalı sayım oranı azaltılmıştır. Model 60 ile epok eğitilmiştir. Eğitim oranı 0.0001 olarak seçilmiştir ve her 20 epokta bir 0.1 ile çarpılarak azaltılmıştır. Optimizasyon algoritması olarak ADAM [109] kullanılmıştır.

Tablo 3.6. CARPK veri seti için önerilen merkez nokta işaretleme tabanlı HLCNN modelinin karşılaştırma sonuçları

Yöntem	Etiketleme Yöntemi	MAE	RMSE
Faster R-CNN [3, 8]	çerçeve	24.32	37.62
YOLO [8, 96]	çerçeve	48.89	57.55
One-Look Regression [8, 104]	Nesne sayısı	59.46	66.84
LPN [8]	çerçeve	23.80	36.79
RetinaNet [4]	çerçeve	16.62	22.30
IoUNet [70]	çerçeve	6.77	8.52
VGG-GAP-HR [67]	çerçeve	7.88	9.30
GANet [69]	çerçeve	4.80	6.90
HLCNN	merkez nokta	9.30	13.44

Eğitilen model farklı yöntemlerin CARPK veri seti üzerindeki başarımları ile kıyaslanmıştır. Kıyaslama sonuçları ve yöntemlerin kullandığı etiketleme yöntemleri Tablo 3.6'da görülmektedir. Önerilen model MAE kriteri üzerinden 9.30 ve RMSE kriteri üzerinden 13.40 hata oranını elde etmiştir. Elde edilen sonuçlar nesnelere nokta etiketlerini kullanarak eğitilen önerilen modelin nesne çerçeve etiketlerini kullanarak eğitilen bir çok yöntemden daha başarılı sonuçlar elde ettiğini ve diğer yöntemler ile yarışabilecek düzeyde başarımlar gösterdiğini ortaya koymuştur. Etiketleme maliyetini azaltan yöntemin sağladığı verimlilik nesne sayımının kalitesini artırmaktadır. Yöntemin nesnelere merkez noktalarını tespit ederek çıktı olarak sunması da bir özelliğidir. Deneysel çalışmalar sonucunda elde edilen görsel sonuçlar Şekil 3.11'de görülmektedir.



Şekil 3.11. Önerilen merkez işaretleme tabanlı HLCNN modeli ile elde edilen görsel sonuçlar

3.2.3. Deęerlendirme

Bu bölümde görüntü etiketleme maliyeti düşük nokta işaretleme ile eğitilebilen yeni bir CNN tabanlı araba sayımı modeli önerilmektedir. HLCNN kullanılarak elde edilen deneysel sonuçların ışığında çerçeve etiketleme kullanan bir çok yöntemden daha iyi sonuç elde etmiştir. Nokta işaretleme 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.



4. BÖLÜM

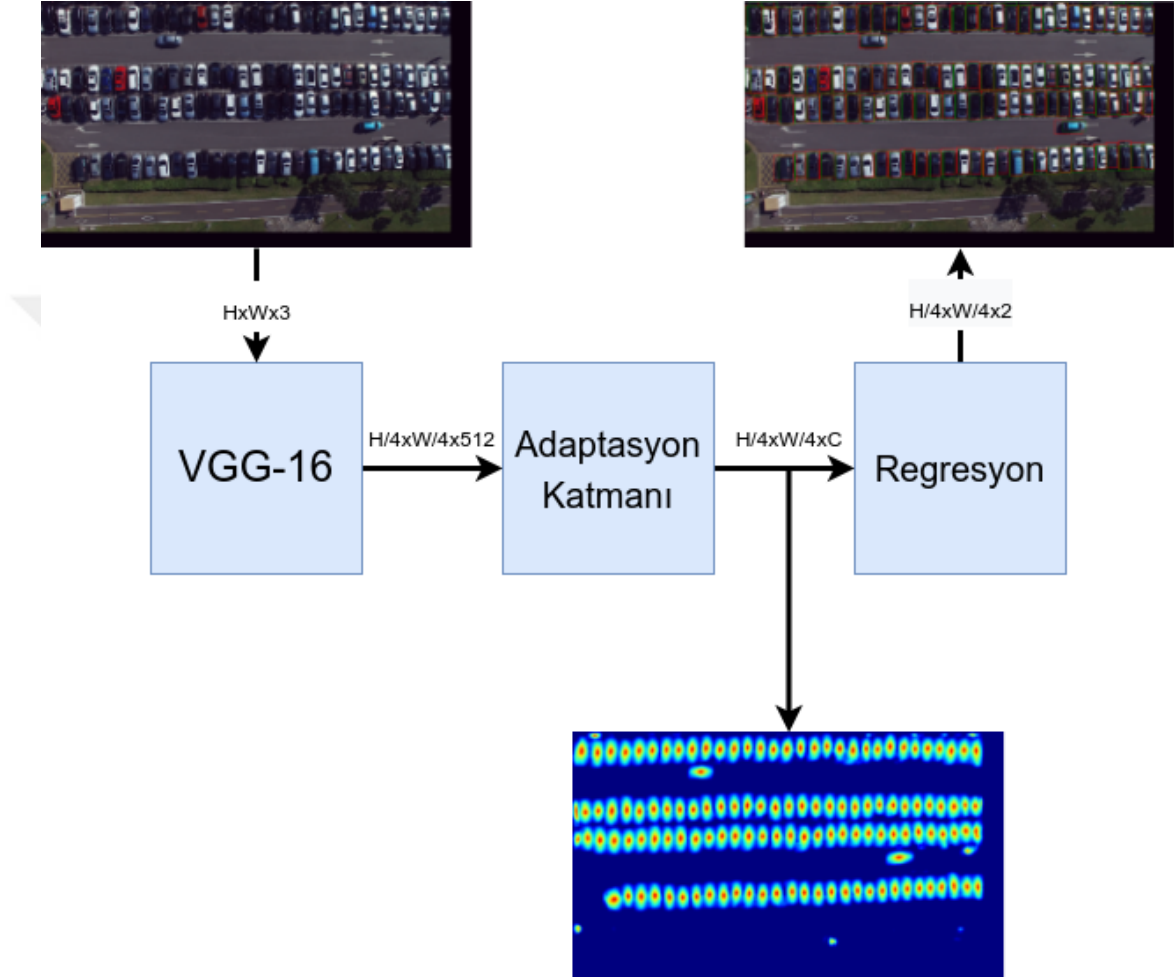
İNSANSIZ HAVA ARACI GÖRÜNTÜLERİNDE ÇOK SINIFLI ARAÇ TESPİTİ

Bu bölümde, İHA görüntülerinde çok sınıflı araç tespiti üzerine derin öğrenme tabanlı yeni bir model tanıtılmaktadır. Bir önceki bölümde tanıtılan HLCNN mimarisinin nesne ısı haritası oluşturmada gösterdiği başarıdan esinlenerek bir model geliştirilmiştir. HLCNN mimarisine nesne en boyunu tahmin eden bir katman eklenmiştir. Nesnelerin merkez noktalarını tahmin edebilen HLCNN mimarisine nesne büyüklüğünü tespit etme yeteneği kazandırılmıştır. Önerilen yöntem CNessNet olarak adlandırılmıştır. Bu bölümde nesne ısı haritaları oluşturulması için GAM yerine Bölüm 4.1.1’de tanıtılan merkezîyet skoru ile haritalama kullanılmıştır. Önerilen yöntemin eğitimi için kullanılan yeni bir kayıp fonksiyonu tanıtılmıştır. Bu bölümde tanıtılan modelin başarısını değerlendirmek için iki farklı deneysel çalışma yapılmıştır. İlk aşamada CARPK veri seti kullanılarak HLCNN mimarisinin başarımı çapasız nesne tespiti modellerinde sıklıkla kullanılan kum saati ağları ile karşılaştırılmıştır. İkinci aşamada ise önerilen yöntemin başarımı VAID veri seti üzerinde ölçülmüştür. Önerilen yöntemin araç tespiti üzerine başarılı olduğu karşılaştırılmalı olarak sunulmuştur.

4.1. Önerilen CNessNet Modeli

Önerilen CNessNet modelinin genel blok şeması Şekil 4.1’de gösterilmiştir. CNessNet omurga ağ olarak ısı haritası tahmin edilebilen modelleri kullanan ve bu modellere ilave nesne en ve boylarını tahmin eden 5 adet konvolüsyon katmanı içeren çerçeve tespiti katmanından oluşmaktadır. CNessNet tek atışlı çapasız nesne tespiti modelidir. Bir çok nesne tespiti modeli sınıflandırıcı ve regresyon olmak üzere birden fazla kafadan oluşmaktadır. Bu kafalar genellikle aynı özellikleri paylaşarak kullanırlar.

Önerilen yöntemde ise regresyon kafası sınıflandırıcı olarak kullanılan ısı haritalarını özellik haritası olarak kullanmaktadır. Bu özelliği ile literatürde bulunan yöntemlerden ayrılmaktadır. CNessNet mimarisinde alt örnekleme oranı CornerNet [5], CenterNet [6] gibi çapasız nesne tespiti yöntemlerine benzer şekilde 4 olarak belirlenmiştir.



Şekil 4.1. CNessNet modelinin genel blok şeması

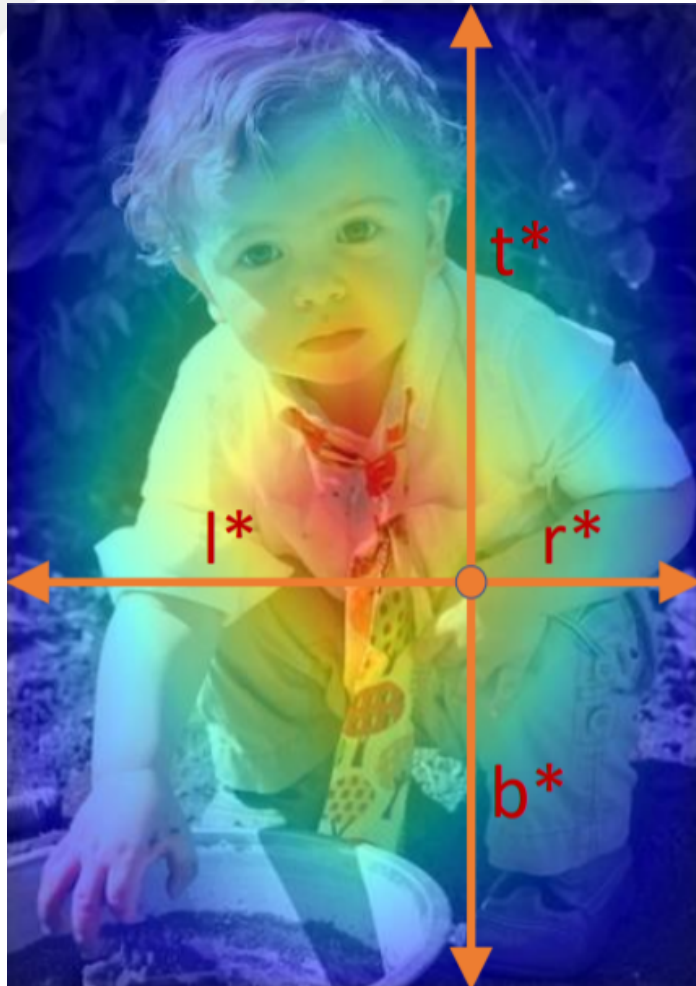
Çapa tabanlı nesne tespiti modellerinde regresyon katmanı çıktıları belirli bir sayıdaki çapa için oluşmaktadır. Sınıflandırma skoru en yüksek belirli sayıdaki çapalar nihai tespit edilen nesnelere oluşturulmaktadır. Çapasız nesne tespiti yöntemlerinde genellikle tahmin edilen ısı haritası üzerindeki lokal maksimum noktaları nesnelere kenar veya köşelerini temsil etmektedir. Bunlara benzer şekilde CNessNet mimarisinde lokal maksimumlar nesnelere merkez noktalarını temsil etmektedir. Aynı zamanda, lokal maksimum noktasındaki değer sınıflandırma skorunu oluşturulmaktadır. CNessNet regresyon katmanı çıktısı nesnelere en ve boy değerlerinin tespiti için iki adet harita oluşturulmaktadır. Bu haritalar çerçeve haritası olarak adlandırılmaktadır. Isı haritasında tespit edilen lokal

maksimum noktasının koordinatlarının çerçeve haritasındaki karşılıkları nesnenin en ve boy değerlerini tahmin etmektedir.

Önerilen model iki parçalı bir kayıp fonksiyonu ile eğitilmektedir. Kayıp fonksiyonu tahmin edilen nesne ısı haritası hatası ve tahmin edilen nesne en ve boyu hatasından oluşmaktadır. Tezin bu bölümünde merkeziyet skoru ile harita oluşturma yöntemi ve önerilen kayıp fonksiyonu detaylı olarak tanıtılmaktadır.

4.1.1. Merkeziyet Skoru ile Haritalama Yöntemi

Merkeziyet skoru ile haritalama ilk olarak Tian ve ark. [9] tarafından tanıtılmıştır. Merkeziyet skoru Şekil 4.2’de gösterildiği gibi lokasyonun nesne merkezine olan normleştirilmiş uzaklığı ifade etmektedir. Merkeziyet skoru 4 parametreye bağlı olarak Eşitlik 4.1 ve 4.2’ye göre hesaplanmaktadır.



Şekil 4.2. Merkeziyet skoru ile oluşturulmuş örnek harita görseli [9]

$$l^* = x - x_0^{(i)}, t^* = y - y_0^{(i)}, r^* = x_1^{(i)} - x, b^* = y_1^{(i)} - y. \quad (4.1)$$

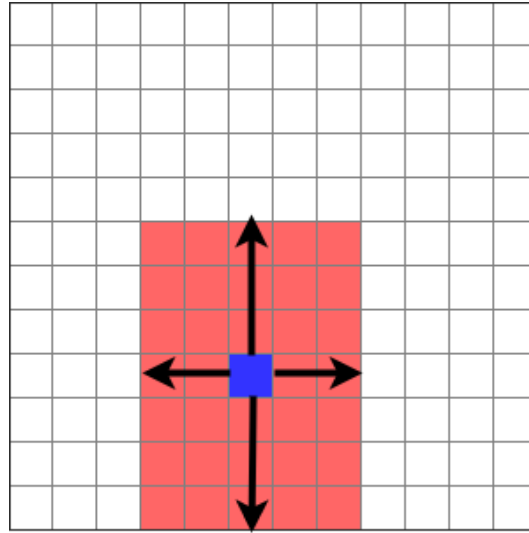
$$centerness^* = \sqrt{\frac{\min(l^*, r^*)}{\max(l^*, r^*)} \times \frac{\min(t^*, b^*)}{\max(t^*, b^*)}} \quad (4.2)$$

4.1.2. CNessNet Modelinin Eğitimi ve Kayıp Fonksiyonu

CNessNet modelinin eğitim aşamasında tahmin edilen ısı haritası ve nesne en ve boylarında yapılan hatanın geri yayılımı yapılmaktadır. Adaptasyon katman çıkışında oluşan ısı haritası hatası ve çerçeve katmanı çıkışında oluşan hata ayrı ayrı hesaplanarak geri yayılmaktadır.

CNessNet eğitimi aşamasında kullanılan nesne ısı haritaları merkeziyet skoru haritalama yöntemi ile oluşturulmaktadır.

Eğitim aşamasında kullanılan kesin referans çerçeve haritalarının oluşturulması için kullanılan yöntem kabaca Şekil 4.3'de görülmektedir. Şekil 4.3 (a)'da temsili olarak 12×12 boyutunda bir görüntü bulunmaktadır. Kırmızı ile boyanmış pikseller nesneyi temsil etmektedir. Nesnenin merkez noktası mavi ile boyanmıştır. Nesnenin merkez noktasından 4 kenara olan maksimum uzaklıklar tahmin edilmesi gereken değerleri oluşturmaktadır. Şekil 4.3 (b)'de X eksenini için çerçeve tespiti katmanının tahmin etmesi beklenen değerler bulunmaktadır. Şekil 4.3 (c)'de Y eksenini için çerçeve tespiti katmanının tahmin etmesi beklenen değerler bulunmaktadır.



(a)

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(b)

(c)

Şekil 4.3. Regresyon katmanı için örnek kesin referans haritası (a) Temsili görüntü ve nesne, (b) X eksenini için kesin referans değerleri, (c) Y eksenini için kesin referans değerleri

Önerilen yöntemin tahmin ettiği merkez noktada kaymaların oluşması muhtemeldir. Merkez noktada oluşacak kaymalar çerçeve tespiti katmanında nesnelerin en ve boylarının doğru şekilde tahminini olumsuz yönde etkileyebilmektedir. Bu durumun önüne geçmek amacıyla çerçeve haritalarında sadece nesnenin merkez noktasına karşılık gelen koordinat en ve boy değerleri ile doldurmak yerine bir k_x parametresine bağlı olarak merkez noktanın k_x derinlikte 8 komşusu en ve boy değerleri ile doldurulmaktadır. Şekil 4.3'te k_x 1 olarak belirlenerek çerçeve haritaları oluşturulmuştur. k_x parametresinin 0 olduğu

durumda sadece merkez noktaya karşılık gelen koordinat nesne en ve boy değerleri ile doldurulmaktadır.

Çerçeve katmanında en ve boyun 0 olduğu arka plan bölgelerinin 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 Eşitlik 4.3, 4.4 ve 4.5'te görülmektedir.

$$L_{bg} = \frac{\sum_i^n |x_i - y_i|}{n} \quad (4.3)$$

$$L_{fg} = \frac{\sum_i^{k \times m} |x_i - y_i|}{k} \quad (4.4)$$

$$L_{dim} = L_{bg} + L_{fg} \quad (4.5)$$

y model tarafından tahmin edilen haritayı, x kesin referans çerçeve katmanı haritasını ve k görüntüde bulunan nesne sayısını temsil etmektedir. k_x parametresinin büyüklüğüne göre nesne en ve boy değeri 0 olmayan bölgelerde yapılan hatalar, 0 olan bölgelere göre daha çok katkı sağlayacak bir kayıp fonksiyonu tanımlanmıştır.

Nesne konumlarının ısı haritasını tahmin eden adaptasyon katmanı çıkışı hatasını hesaplayan yeni bir kayıp fonksiyonu önerilmiştir.

$$err = |pred - gt| \quad (4.6)$$

$$gt_r = gt + \theta \quad (4.7)$$

$$gtR = \begin{cases} gt_r & gt_r \leq 1 \\ 1 & gt_r > 1 \end{cases} \quad (4.8)$$

$$L_1 = \sum_c \sum_i \log(1 - err_i) \times \log(1 - gtR_i) \times \sqrt{w_c^{-\beta}} \quad (4.9)$$

$$L_2 = \sum_c \sum_k -\log(1 - err_k) \times \alpha \times \sqrt{w_c}^\beta \quad (4.10)$$

$$L_3 = \sum_c \sum_n 2 \times \alpha \times \sqrt{w_c}^\beta \quad (4.11)$$

$$L_{map} = L_1 + L_2 + L_3 \quad (4.12)$$

Önerilen kayıp fonksiyonu hesaplanan hata err ve kesin referans harita gt değerlerine göre 3 aşamada hesaplanmaktadır. Bu eşitlikte gt kesin referans değerlerini ve $pred$ ise tahmin edilen nesne ısı haritası çıkışını temsil etmektedir. w sınıf ağırlıklarını temsil eder. Sınıf dengesizliğinin etkisini azaltmak için, çoğunluk sınıfına daha düşük bir ağırlık atanarak modelin tahmin süreci sırasında daha az sıklıkta görülen sınıflara odaklanması teşvik edilebilir. Sınıf dengesizliğini ortadan kaldırmak için veri kümesindeki her bir sınıf için sınıf piksel sayısı ile ters orantılı ağırlıklar kullanılır. En çok piksel içeren sınıfın ağırlığı 1 olarak ayarlanır. gt değeri 1 olmadığında ve err değeri 1 olmadığında her sınıf için L_1 hesaplanır. gt değeri 1 ve err değeri 1 olmadığında her sınıf için L_2 hesaplanır. gt değeri 1 ve err değeri 1 olduğunda her sınıf için L_3 hesaplanır. Toplam hata Eşitlik 4.5 ve Eşitlik 4.12'nin toplamı olarak hesaplanmaktadır.

4.2. Deneysel Çalışmalar

Deneysel çalışmalar iki kısımdan oluşmaktadır. İlk kısımda, önerilen CNessNet mimarisinin farklı omurga ağı ile başarımının analizi için CARPK veri seti ile deneyler yapılmıştır. HLCNN ve HG olmak üzere 2 farklı omurga ağı ile testler yapılmıştır. HLCNN ve HG omurga ağı kullanan CNessNet mimarisileri mAP metriğiyle sırasıyla 0.899 ve 0.9292 başarımlar elde etmiştir. Araba tespitinde omurga ağı olarak HG kullanan CNessNet mimarisinin daha başarılı olduğu tespit edilmiştir. İkinci kısımda omurga ağı olarak HG kullanan CNessNet mimarisinin sınıf dengesizliğini önleyen yeni kayıp fonksiyonunun çok sınıflı araç tespitinde başarımını analiz edilmiştir.

4.2.1. Veri Setleri

Tablo 4.1. VAID veri setinde kategori başına araç sayıları

Kategori	Araç Sayısı
Araba	40330
Minibüs	501
Kamyonet	3187
Pikap	3011
Otobüs	580
Çimento Kamyonu	191
Tır	804

Önerilen yöntemin araba tespitinde başarımı ölçmek amacıyla Bölüm 3.1.2.3’de tanıtılan CARPK veri seti kullanılmıştır. Çok sınıflı araç tespiti için VAID veri seti [110] kullanılmıştır. Araba, minibüs, kamyonet, pikap, otobüs, çimento kamyonu ve tır olmak üzere 7 sınıfta araç içeren veri seti 5989 adet RGB görüntüden oluşmaktadır. 1512 adet görüntü eğitim için 2939 adet görüntü ise test için kullanılmıştır. Tablo 4.1’de kategorilere ait araç sayıları bulunmaktadır.

4.2.2. Araba Tespiti

Önerilen yöntemin araba tespitinde başarımını ölçmek ve omurga ağ olarak HLCNN ile HG-2 ağının kıyaslanması için CARPK veri seti kullanılmıştır. Bölümde HLCNN omurga ağı ile eğitilen model CNessNet-1, HG-2 omurga ağı ile eğitilen model CNessNet-2 olarak adlandırılmaktadır.

Tablo 4.2. CARPK veri seti için CNessNet modellerinin karşılaştırma sonuçları

Yöntem	Araba
SSD [28, 69]	0.687
YOLOv3 [69, 73]	0.853
SA+CF+CRT [100]	0.898
GANet (VGG-16) [69]	0.902
GANet (ResNet-50) [69]	0.901
CNessNet-1	0.899
CNessNet-2	0.929

CNessNet-1 ve CNessNet-2 için omurga ağ alt örnekleme oranı 4 olarak belirlenmiştir.



Şekil 4.4. CARPK veri seti için CNessNet-2 modeli ile elde edilen sonuçlar

Modeller 8 epok için eğitilmişlerdir. Öğrenme katsayısı 0.001 olarak belirlenmiş ve 4. epokta 0.1 oranında azaltılmıştır. θ , 0.5 olarak belirlenmiştir. α , 15 olarak kullanılmıştır. Tek sınıf olduğu için β 0 olarak belirlenmiştir. Elde edilen sonuçlar karşılaştırmalı olarak Tablo 4.2’de verilmektedir. HG-2 omurga ağı kullanan CNessNet-2 modeli literatürde bilinen en iyi sonucu elde etmiştir. CNessNet-2 modelinin araba tespiti görsel sonuçları Şekil 4.4’de görülmektedir. Yeşil sınırlayıcı kutular kesin referansları, kırmızı sınırlayıcı kutular önerilen yöntemin tahminini temsil etmektedir.

4.2.3. Çok Sınıflı Araç Tespiti

CNessNet mimarisinin çok sınıflı araç tespiti başarımını ölçmek için bu bölümde VAID veri seti üzerinde eğitim ve testler gerçekleştirilmiştir. İHA görüntülerinde çok sınıflı araç tespitinde zorlayıcı noktalara ek olarak VAID veri setinde sınıflar arasında örnek dengesizliği de bulunmaktadır. Bu sebeple Eşitlik 4.12’de önerilen kayıp fonksiyonuna yöntemin sınıf dengesizliğine karşı gürbüz olması için yeni bir terim eklenmiştir ve β parametresi ile kontrol edilmesi sağlanmıştır. İlgili terimde bulunan w ağırlıkları ve ağırlıkları hesaplamak için kullanılan sınıf bazlı piksel sayıları Tablo 4.3’te sunulmuştur. Eklenen bu terimin başarımını değerlendirmek amacıyla 2 farklı CNessNet modeli benzer koşullarda eğitilmiştir.

Tablo 4.3. VAID veri seti için sınıf ağırlıkları

Sınıf	Nesne Sayısı	Piksel Sayısı	w
Araba	10362	11731240	1
Minibüs	128	325375	36.05
Kamyonet	818	1584865	7.4
Pikap	767	1253934	9.36
Otobüs	150	1047350	11.2
Çimento Kamyonu	51	185936	63.1
Tır	216	1473409	7.96

Tablo 4.4. VAID veri seti için önerilen CNessNet-4 modelinin karşılaştırma sonuçları

Yöntem	M.F. R-CNN [110]	YOLOv4 [111]	RefineDet [112]	CNessNet-4
Araba	0.9022	0.9849	0.8908	0.9577
Minibüs	0.9080	0.9604	0.9436	0.9471
Kamyonet	0.8934	0.9644	0.8346	0.8895
Pikap	0.8893	0.5725	0.8280	0.8703
Otobüs	0.9087	0.9703	0.9784	0.923
Çimento Kamyonu	0.9096	0.6994	0.9124	0.9138
Tır	0.8975	0.9545	0.8074	0.8473
mAP	0.9012	0.8723	0.8626	0.907

İlk eğitilen model (CNessNet-3) β parametresi 0 olarak eğitilmiştir. İkinci modelde (CNessNet-4) β parametresi son 10 epokta 1 olarak eğitilmiştir. Eğitimler 25 epokta yapılmıştır. Öğrenme katsayısı 0.001 olarak belirlenmiştir ve her 10 epokta 0.1 oranında azaltılmıştır. θ ve α sırasıyla 0.5 ve 15 olarak belirlenmiştir. CNessNet-3 ve CNessNet-4, çok sınıflı araç tespiti için VAID veri setinde sırasıyla 0.8871 ve 0.9070 mAP



Şekil 4.5. VAID veri seti için CNESSNet-4 modeli ile elde edilen sonuçlar

başarımına ulaşmıştır. Elde edilen sonuçlarda önerilen yeni kayıp fonksiyonun CNESSNet başarımına olumlu yönde katkı verdiği ortaya çıkmıştır. Elde edilen sonuçlar Tablo 4.4'te karşılaştırılmalı olarak sunulmuştur.

CNESSNet-4 modelinin araç tespiti görsel sonuçları Şekil 4.5'de görülmektedir. Yeşil , sarı, mavi, kırmızı, siyah, beyaz, pembe ve mor renkli çerçeveler sırasıyla kesin referans, araba, minibüs, kamyonet, pikap, otobüs, çimento kamyonu ve traktörleri temsil etmektedir.

4.3. Değerlendirme

Bu bölümde, İHA görüntülerinden araç tespiti amacıyla yeni bir CNN tabanlı model önerilmiştir. Önerilen model, tek atışlı çapasız bir nesne tespiti yöntemidir ve omurga ağ olarak nesne ısı haritası öğrenebilen CNN mimarileri içermektedir. Ayrıca bu bölümde ısı haritası eğitimi için yeni bir kayıp fonksiyonu önerilmiştir. Önerilen kayıp fonksiyonu eğitim setinde bulunan arkaplan önplan dengesizliği ve sınıflar arası dengesizliği ayarlamak için yapılar içermektedir. Önerilen yöntemin, başarımını literatürde sıklıkla kullanılan CARPK ve VAID araç veri setleri üzerinden değerlendirilmiştir. Elde edilen sonuçlar, önerilen yöntemin literatürde bilinen en iyi araç sayma başarısına ulaştığını göstermiştir.

5. BÖLÜM

TARTIŞMA, SONUÇ ve ÖNERİLER

5.1. Tartışma ve Sonuçlar

Bu tez kapsamında çoklu-ölçekteki ve farklı zorlayıcı içeriklere sahip uydu ve İHA görüntülerinde nesne tespiti ve sayımına yönelik yeni modeller önerilmiştir. Önerilen modeller son yıllarda literatürde başarılı sonuçlar elde eden CNN tabanlı yöntemlerle karşılaştırılmıştır.

İlk olarak, uydu görüntülerinde nesnelere daha doğru bir şekilde tespit etmek için CNN tabanlı yeni bir alt sınıf tabanlı sınıflandırıcı önerilmiştir. Önerilen sınıflandırıcı, SSCNN olarak adlandırılmakta ve daha etkin özellik çıkarıcı elde etmek için nesnelere temsilini nesne merkezinin mesafesine bağlı olarak yakın merkez, merkez ve sınır gibi alt sınıflara ayırmak için kullanılmaktadır. Önerilen sınıflandırıcının başarımını değerlendirmek için üç aşamalı bir nesne tanıma çerçevesi kullanılmıştır. Bu aşamaların ilkinde, SS algoritması görüntüden nesne önerileri üretmiştir. Ardından, önerilen SSCNN, teklifleri sınıflandırmıştır. Son olarak, sınıflandırma sonuçları ile nesnenin lokalizasyonunu hesaplamak için alt sınıf tabanlı yerleştirme değerlendirme fonksiyonu önerilmiştir. Sınırlı sayıda uydu görüntüsü örneği nedeniyle, önceden eğitilmiş AlexNet, etkili öznetelik çıkarıcı oluşturmak için transfer öğrenme yaklaşımıyla kullanılmıştır. Önerilen model, 411 uçak, 240 beyzbol sahası, 468 depolama tankı ve 83 futbol sahasından oluşan dört sınıflı bir uzaktan algılama test veri setinde R-CNN ile karşılaştırılmıştır. Ayrıca Faster R-CNN, SSCNN özellikleri ile eğitilmiş ve eğitilmiş Faster R-CNN'lerin başarımları, 10 sınıflı uzaktan algılama görüntü veri seti üzerinde karşılaştırmalı olarak değerlendirilmiştir. Deney sonuçları, önerilen modelin daha ayırt edici özellik çıkarımı sağladığını göstermiştir.

İkinci olarak, İHA görüntülerinde arabaları tespit etmek ve saymak için basit ve etkili nesne tespiti modeli önerilmiştir. HLCNN olarak adlandırılan önerilen model, hedef araba örneklerinin ısı haritasını tahmin etmek için kullanılmıştır. Hedef arabaların ısı haritasını öğrenmek için, tam bağlantılı katmanlar yerine uyarlama katmanları olarak adlandırılan üç evrişim katmanı ekleyerek HLCNN mimarisini oluşturulmuştur. VGG-16, önerilen modelde bir omurga ağ olarak kullanılmıştır. Önerilen model, başarılı bir şekilde araba sayısını belirlemiş ve hedef arabaların merkezini hassas bir şekilde tespit etmiştir. İki farklı araba veri seti (PUCPR+ ve CARPK) üzerinde yapılan deneyler, önerilen modelin mevcut yöntemlerle karşılaştırıldığında son teknoloji sayma ve yerleştirme başarısını elde ettiğini göstermiştir. Ayrıca, veri büyütme ve yığın normalleştirilmesinin önerilen modelin başarısı üzerindeki etkisini incelemek için deneyler yapılmıştır. Daha sonra, HLCNN modelinin işaretleme maliyeti daha düşük merkez nokta işaretleme üzerinden eğitimi gerçekleştiren bir yöntem önerilmiştir. HLCNN modelin eğitimi için nokta etiketleme verilerinin yeterli olması veri seti oluşturma aşamasını oldukça kolaylaştırmıştır. Nokta etiketleme ile CNN modellerin eğitilmesi için yeni bir kayıp fonksiyonu önerilmiştir. Önerilen modelin başarımının değerlendirilmesi amacıyla CARPK veri seti üzerinde deneysel çalışmalar yapılmıştır. Elde edilen sonuçlar çerçeve etiketleme ile eğitilen modeller ile kıyaslandığında önerilen modelin bu modeller ile yarışabilecek düzeyde başarımlar elde ettiği göstermiştir.

Son olarak, İHA görüntülerinde çok sınıflı araç tespiti üzerine CNN tabanlı yeni bir model önerilmiştir. HLCNN mimarisinin nesne ısı haritası oluşturmada gösterdiği başarıdan esinlenerek, HLCNN mimarisine nesne boyutlarını tahmin eden yapı eklenerek oluşturulmuştur. Nesnelerin merkez noktalarını tahmin edebilen HLCNN mimarisine nesne büyüklüğünü tespit etme yeteneği kazandırılmıştır. CNessNet olarak adlandırılan önerilen model, HLCNN gibi ısı haritası öğrenebilen mimarilere, nesne boyutlarını tahmin edebilmesi için 5 konvolüsyon katmanının eklenmesi ile oluşturulmuştur. Yapılan deneysel çalışmalarda ilk olarak CNessNet yönteminde omurga ağın başarımına etkisi CARPK veri seti üzerinden değerlendirilmiştir. HG'nin daha başarılı sonuçlar elde ettiği görülmüştür. Deneysel çalışmaların ikinci bölümünde önerilen yöntemin çok sınıflı araç tespitinde başarımını ölçmek amacıyla omurga ağ olarak HG kullanan CNessNet eğitilmiştir. Ayrıca, sınıf dengesizliğini önleyen yeni kayıp fonksiyonu

önerilmiştir. Önerilen yeni kayıp fonksiyonun çok sınıflı araç tespitinde başarımı analiz edilmiştir. Elde edilen sonuçlarda önerilen kayıp fonksiyonun daha iyi sonuçlar elde ettiği görülmüştür.

5.2. Öneriler

Uzaktan algılama görüntülerinde nesne tespiti ve sayımı hâlâ araştırılması gereken açık bir alandır. Her ne kadar önerilen algoritmalar literatürde oldukça iyi sonuçlar elde ediyor olsa da verinin hızla arttığı günümüz dünyasında daha iyi sonuç üreten ve daha hızlı yöntemlere ihtiyaç bulunmaktadır. Özellikle belirli bir veri seti üzerinde yüksek başarımlar elde eden yöntemlerin, gerçek dünyada başarısız olduğu senaryolar mevcuttur. Bu tez çalışması ile bağlantılı olarak gelecekte aşağıda bulunan konuların araştırılması önerilmektedir:

HLCNN ve CNessNet modelleri araba ve araç sayımında literatürde bilinen en iyi başarı oranını yakalamışlardır. Bu modellerin insan, gemi, ağaç ve hayvan gibi farklı nesnelere tespiti ve sayımı üzerine çalışmalarda başarımlarının değerlendirilmesi önerilmektedir. Problemler özelinde farklı omurga ağlarının ve veri artırma tekniklerinin başarımlarına etkisinin incelenmesi yapılabilir. HLCNN modeli sınıf bağımsız nesne sayımı problemine uygulanarak yeni yöntemler geliştirilebilir. HLCNN ve CNessNet modellerinin, son aşamada NMS içermemesi uç cihazlarda gerçek zamanlı çalışabilme potansiyelini artırmaktadır. Uç cihazlarda gerçek zamanlı çalışabilen modellerin geliştirilmesi için önerilen modellerin daha az parametre içeren omurga ağlar ile başarımlarının artırılmasına yönelik çalışmalar yapılabilir.

KAYNAKÇA

1. Girshick, R., Donahue, J., Darrell, T., and Malik, J., 2014. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, *2014 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE.
2. Girshick, R., 2015. Fast R-CNN, *International Conference on Computer Vision (ICCV)*.
3. Ren, S., He, K., Girshick, R., and Sun, J., 2017. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, **IEEE Transactions on Pattern Analysis and Machine Intelligence**, **39(6)**:1137–1149.
4. Lin, T.Y., Goyal, P., Girshick, R., He, K., and Dollár, P., 2020. Focal Loss for Dense Object Detection, **IEEE Transactions on Pattern Analysis and Machine Intelligence**, **42(2)**:318–327.
5. Law, H. and Deng, J., 2019. CornerNet: Detecting Objects as Paired Keypoints, **International Journal of Computer Vision**, **128(3)**:642–656.
6. Duan, K., Bai, S., Xie, L., Qi, H., Huang, Q., and Tian, Q., 2019. CenterNet: Keypoint Triplets for Object Detection, *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE.
7. Zitnick, C.L. and Dollár, P., 2014. Edge Boxes: Locating Object Proposals from Edges, *Computer Vision – ECCV 2014*, 391–405, Springer International Publishing.
8. Hsieh, M.R., Lin, Y.L., and Hsu, W.H., 2017. Drone-based Object Counting by Spatially Regularized Regional Proposal Networks, *The IEEE International Conference on Computer Vision (ICCV)*, IEEE.
9. Tian, Z., Shen, C., Chen, H., and He, T., 2020. FCOS: A Simple and Strong Anchor-free Object Detector, **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 1–1.

10. Lu, D. and Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance, **International Journal of Remote Sensing**, **28**(5):823–870.
11. Wang, J., Liu, C., Fu, T., and Zheng, L., 2019. Research on automatic target detection and recognition based on deep learning, **Journal of Visual Communication and Image Representation**, **60**:44–50.
12. Chen, H., Zhang, L., Ma, J., and Zhang, J., 2019. Target heat-map network: An end-to-end deep network for target detection in remote sensing images, **Neurocomputing**, **331**:375–387.
13. Mundhenk, T.N., Konjevod, G., Sakla, W.A., and Boakye, K., 2016. A Large Contextual Dataset for Classification, Detection and Counting of Cars with Deep Learning, *ECCV*.
14. Sun, M., Wang, Y., Li, T., Lv, J., and Wu, J., 2017. Vehicle counting in crowded scenes with multi-channel and multi-task convolutional neural networks, **Journal of Visual Communication and Image Representation**, **49**:412–419.
15. Mauro, D.D., Furnari, A., Patanè, G., Battiato, S., and Farinella, G., 2019. Estimating the occupancy status of parking areas by counting cars and non-empty stalls, **Journal of Visual Communication and Image Representation**, **62**:234–244.
16. Chen, W., Qiao, Y., and Li, Y., 2020. Inception-SSD: An improved single shot detector for vehicle detection, **Journal of Ambient Intelligence and Humanized Computing**.
17. Han, J., Zhang, D., Cheng, G., Liu, N., and Xu, D., 2018. Advanced Deep-Learning Techniques for Salient and Category-Specific Object Detection: A Survey, **IEEE Signal Processing Magazine**, **35**(1):84–100.
18. Long, Y., Gong, Y., Xiao, Z., and Liu, Q., 2017. Accurate Object Localization in Remote Sensing Images Based on Convolutional Neural Networks, **IEEE Transactions on Geoscience and Remote Sensing**, **55**(5):2486–2498.

19. Kilic, E. and Ozturk, S., 2019. A subclass supported convolutional neural network for object detection and localization in remote-sensing images, **International Journal of Remote Sensing**, **40**(11):4193–4212.
20. Kilic, E. and Ozturk, S., 2021. An accurate car counting in aerial images based on convolutional neural networks, **Journal of Ambient Intelligence and Humanized Computing**, **14**(2):1259–1268.
21. Kılıç, E. and Öztürk, S., 2020. İnsansız Hava Aracı Görüntülerinde Evrişimli Sinir Ağları Tabanlı Araba Sayımı Yöntemi, **Erciyes Üniversitesi Fen Bilimleri Enstitüsü Fen Bilimleri Dergisi**, **36**(2):170–177.
22. LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., and Jackel, L., 1990. Handwritten Digit Recognition with a Back-Propagation Network, *D. Touretzky, Ed.; Advances in Neural Information Processing Systems*, volume 2, Morgan-Kaufmann.
23. Özcan, T., 2020. Derin Öğrenme ile İnsan Edimlerinin Tanınması, Doktora Tezi, Erciyes Üniversitesi, Kayseri, TR.
24. Goodfellow, I., Bengio, Y., and Courville, A., 2016. Deep Learning, MIT Press, <http://www.deeplearningbook.org>.
25. Ioffe, S. and Szegedy, C., 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37, ICML'15*, 448–456, JMLR.org.
26. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., and Fei-Fei, L., 2009. ImageNet: A large-scale hierarchical image database, *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255.
27. Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollar, P., and Zitnick, L., 2014. Microsoft COCO: Common Objects in Context, *ECCV*, European Conference on Computer Vision.

28. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., and Berg, A.C., 2016. SSD: Single Shot MultiBox Detector, *B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., Computer Vision – ECCV 2016*, 21–37, Springer International Publishing, Cham.
29. Simonyan, K. and Zisserman, A., 2015. *Very deep convolutional networks for large-scale image recognition*, 1–14, Computational and Biological Learning Society.
30. Uijlings, J.R.R., van de Sande, K.E.A., Gevers, T., and Smeulders, A.W.M., 2013. Selective Search for Object Recognition, **International Journal of Computer Vision**, **104**(2):154–171.
31. Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., and Belongie, S., 2017. Feature Pyramid Networks for Object Detection, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 936–944.
32. Körez, A., 2020. DERİN ÖĞRENME KULLANARAK UZAKTAN ALGILAMA GÖRÜNTÜLERİNDEKİ NESNELERİN TESPİTİ, Doktora Tezi, Gazi Üniversitesi, Ankara, TR.
33. Newell, A., Yang, K., and Deng, J., 2016. Stacked Hourglass Networks for Human Pose Estimation, *B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., Computer Vision – ECCV 2016*, 483–499, Springer International Publishing, Cham.
34. Lowe, D.G., 2004. Distinctive Image Features from Scale-Invariant Keypoints, **International Journal of Computer Vision**, **60**(2):91–110.
35. Bay, H., Tuytelaars, T., and Van Gool, L., 2006. SURF: Speded Up Robust Features, *A. Leonardis, H. Bischof, and A. Pinz, Eds., Computer Vision – ECCV 2006*, 404–417, Springer Berlin Heidelberg, Berlin, Heidelberg.
36. Dalal, N. and Triggs, B. Histograms of Oriented Gradients for Human Detection, *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE.

37. Sirmacek, B. and Unsalan, C., 2009. Urban-Area and Building Detection Using SIFT Keypoints and Graph Theory, **IEEE Transactions on Geoscience and Remote Sensing**, **47**(4):1156–1167.
38. Tuermer, S., Kurz, F., Reinartz, P., and Stilla, U., 2013. Airborne Vehicle Detection in Dense Urban Areas Using HoG Features and Disparity Maps, **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, **6**(6):2327–2337.
39. Gan, L., Liu, P., and Wang, L., 2015. Rotation Sliding Window of the Hog Feature in Remote Sensing Images for Ship Detection, *2015 8th International Symposium on Computational Intelligence and Design (ISCID)*, volume 1, 401–404.
40. Salberg, A.B., 2015. Detection of seals in remote sensing images using features extracted from deep convolutional neural networks, *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 1893–1896.
41. Chen, X., Xiang, S., Liu, C.L., and Pan, C.H., 2014. Vehicle Detection in Satellite Images by Hybrid Deep Convolutional Neural Networks, **IEEE Geoscience and Remote Sensing Letters**, **11**(10):1797–1801.
42. Jiang, Q., Cao, L., Cheng, M., Wang, C., and Li, J., 2015. Deep neural networks-based vehicle detection in satellite images, *2015 International Symposium on Bioelectronics and Bioinformatics (ISBB)*, 184–187.
43. Li, H., Fu, K., Yan, M., Sun, X., Sun, H., and Diao, W., 2016. Vehicle detection in remote sensing images using denoizing-based convolutional neural networks, **Remote Sensing Letters**, **8**(3):262–270.
44. Zhu, H., Chen, X., Dai, W., Fu, K., Ye, Q., and Jiao, J., 2015. Orientation robust object detection in aerial images using deep convolutional neural network, *2015 IEEE International Conference on Image Processing (ICIP)*, 3735–3739.
45. Cheng, G., Han, J., Zhou, P., and Xu, D., 2019. Learning Rotation-Invariant and Fisher Discriminative Convolutional Neural Networks for Object Detection, **IEEE Transactions on Image Processing**, **28**(1):265–278.

46. Cheng, G., Zhou, P., and Han, J., 2016. Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing Images, **IEEE Transactions on Geoscience and Remote Sensing**, **54**(12):7405–7415.
47. Zhang, F., Du, B., Zhang, L., and Xu, M., 2016. Weakly Supervised Learning Based on Coupled Convolutional Neural Networks for Aircraft Detection, **IEEE Transactions on Geoscience and Remote Sensing**, **54**(9):5553–5563.
48. Li, X. and Wang, S., 2017. Object Detection Using Convolutional Neural Networks in a Coarse-to-Fine Manner, **IEEE Geoscience and Remote Sensing Letters**, **14**(11):2037–2041.
49. Li, K., Cheng, G., Bu, S., and You, X., 2018. Rotation-Insensitive and Context-Augmented Object Detection in Remote Sensing Images, **IEEE Transactions on Geoscience and Remote Sensing**, **56**(4):2337–2348.
50. Zhong, Y., Han, X., and Zhang, L., 2018. Multi-class geospatial object detection based on a position-sensitive balancing framework for high spatial resolution remote sensing imagery, **ISPRS Journal of Photogrammetry and Remote Sensing**, **138**:281–294.
51. Xu, Z., Xu, X., Wang, L., Yang, R., and Pu, F., 2017. Deformable ConvNet with Aspect Ratio Constrained NMS for Object Detection in Remote Sensing Imagery, **Remote Sensing**, **9**(12):1312.
52. Tang, T., Zhou, S., Deng, Z., Zou, H., and Lei, L., 2017. Vehicle Detection in Aerial Images Based on Region Convolutional Neural Networks and Hard Negative Example Mining, **Sensors**, **17**(2):336.
53. Ahmed, M., Wang, Y., Maher, A., and Bai, X., 2022. Fused RetinaNet for small target detection in aerial images, **International Journal of Remote Sensing**, **43**(8):2813–2836.
54. Lu, X., Ji, J., Xing, Z., and Miao, Q., 2021. Attention and Feature Fusion SSD for Remote Sensing Object Detection, **IEEE Transactions on Instrumentation and Measurement**, **70**:1–9.

55. Qin, H., Li, Y., Lei, J., Xie, W., and Wang, Z., 2021. A Specially Optimized One-Stage Network for Object Detection in Remote Sensing Images, **IEEE Geoscience and Remote Sensing Letters**, **18**(3):401–405.
56. Tang, T., Zhou, S., Deng, Z., Lei, L., and Zou, H., 2017. Arbitrary-Oriented Vehicle Detection in Aerial Imagery with Single Convolutional Neural Networks, **Remote Sensing**, **9**(11):1170.
57. Liu, L., Pan, Z., and Lei, B., 2017. Learning a Rotation Invariant Detector with Rotatable Bounding Box, **arXiv preprint, arXiv:1711.09405**.
58. Liu, W., Ma, L., and Chen, H., 2018. Arbitrary-Oriented Ship Detection Framework in Optical Remote-Sensing Images, **IEEE Geoscience and Remote Sensing Letters**, **15**(6):937–941.
59. Redmon, J. and Farhadi, A., 2017. YOLO9000: Better, Faster, Stronger, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
60. Barut, O., 2018. Geospatial object recognition using deep networks for satellite images, Yüksek Lisans Tezi, Middle East Technical University.
61. Gencer, S., 2018. Fine-Grained Object Recognition In Remote Sensing Imagery, Yüksek Lisans Tezi, Bilkent Üniversitesi Fen Bilimleri Enstitüsü.
62. Soydaş, M., 2019. Büyük Ölçekli Uzaktan Algılama Görüntülerinden Derin Öğrenme Teknikleriyle Uçak Tespiti, Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü.
63. Önün, F., 2019. Derin Öğrenme ile Yüksek Çözünürlüklü Hava Görüntülerinde Yolların Tespit Edilmesi, Yüksek Lisans Tezi, Fırat Üniversitesi Fen Bilimleri Enstitüsü.
64. Körez, A., 2020. Derin Öğrenme Kullanarak Uzaktan Algılama Görüntülerindeki Nesnelerin Tespiti, Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü.

65. Moon, J., Lim, S., Lee, H., Yu, S., and Lee, K.B., 2022. Smart Count System Based on Object Detection Using Deep Learning, **Remote Sensing**, **14**(15):3761.
66. Tayara, H., Gil Soo, K., and Chong, K.T., 2018. Vehicle Detection and Counting in High-Resolution Aerial Images Using Convolutional Regression Neural Network, **IEEE Access**, **6**:2220–2230.
67. Aich, S. and Stavness, I., 2018. Improving Object Counting with Heatmap Regulation, **arXiv preprint**, **arXiv:1803.05494**.
68. Kan, Y., Chen, W., Ma, Z., Bai, S., Gong, Y., and Wu, W., 2021. A Congestion Detection Framework based on Vehicle-Counter CNN and Self-Learning Critical Density Approach, *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 1809–1814.
69. YuanQiang, C., Du, D., Zhang, L., Wen, L., Wang, W., Wu, Y., and Lyu, S., 2020. Guided Attention Network for Object Detection and Counting on Drones, *Proceedings of the 28th ACM International Conference on Multimedia*, MM '20, ACM.
70. Goldman, E., Herzig, R., Eisenschtat, A., Goldberger, J., and Hassner, T., 2019. Precise Detection in Densely Packed Scenes, *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE.
71. Tang, T., Zhou, S., Deng, Z., Lei, L., and Zou, H., 2017. Arbitrary-Oriented Vehicle Detection in Aerial Imagery with Single Convolutional Neural Networks, **Remote Sensing**, **9**(11).
72. K, L., Karnick, S., Ghalib, M.R., Shankar, A., Khapre, S., and Tayubi, I.A., 2022. A novel method for vehicle detection in high-resolution aerial remote sensing images using YOLT approach, **Multimedia Tools and Applications**, **81**(17):23551–23566.
73. Redmon, J. and Farhadi, A., 2018. YOLOv3: An Incremental Improvement, **arXiv preprint**, **arXiv:1804.02767**.

74. Etten, A.V., 2018. You Only Look Twice: Rapid Multi-Scale Object Detection In Satellite Imagery, **arXiv preprint, arXiv:1805.09512**.
75. Zhou, L., Zheng, C., Yan, H., Zuo, X., Qiao, B., Zhou, B., Fan, M., and Liu, Y., 2021. Vehicle Detection in Remote Sensing Image Based on Machine Vision, **Computational Intelligence and Neuroscience, 2021:1–12**.
76. Kong, X., Zhang, Y., Tu, S., Xu, C., and Yang, W., 2023. Vehicle Detection in High-Resolution Aerial Images with Parallel RPN and Density-Assigner, **Remote Sensing, 15(6)**.
77. Li, X., Men, F., Lv, S., Jiang, X., Pan, M., Ma, Q., and Yu, H., 2021. Vehicle Detection in Very-High-Resolution Remote Sensing Images Based on an Anchor-Free Detection Model with a More Precise Foveal Area, **ISPRS International Journal of Geo-Information, 10(8)**.
78. Yu, L., Zhi, X., Hu, J., Jiang, S., Zhang, W., and Chen, W., 2021. Small-Sized Vehicle Detection in Remote Sensing Image Based on Keypoint Detection, **Remote Sensing, 13(21)**.
79. Zhou, W., Shen, J., Liu, N., Xia, S., and Sun, H., 2022. An Anchor-Free Vehicle Detection Algorithm in Aerial Image Based on Context Information and Transformer, **IEEE Geoscience and Remote Sensing Letters, 19:1–5**.
80. Guo, X., Gao, M., Zou, G., Bruno, A., Chehri, A., and Jeon, G., 2023. Object Counting via Group and Graph Attention Network, **IEEE Transactions on Neural Networks and Learning Systems, 1–12**.
81. Chen, J., Gao, M., Guo, X., Zhai, W., Li, Q., and Jeon, G., 2023. Object counting in remote sensing via selective spatial [U+2010] frequency pyramid network, **Software: Practice and Experience**.
82. Priyanka, G., Senthil Kumar, J., and Veena, S., 2023. Deep learning based video surveillance for predicting vehicle density in real time scenario, **Journal of Ambient Intelligence and Humanized Computing, 14(4):4371–4383**.

83. Tsai, C.M., Hsieh, J.W., Chang, M.C., Chen, P.Y., and He, G.L., 2023. Counting Very Small Vehicles and Crowds in Uav Images Using Mdamdl Framework.
84. D'Alessandro, A., Mahdavi-Amiri, A., and Hamarneh, G., 2024, AFreeCA: Annotation-Free Counting for All.
85. Liu, M., Wang, Y., Yi, H., and Huang, X., 2023. Vehicle object counting network based on feature pyramid split attention mechanism, **The Visual Computer**, **40**(2):663–680.
86. Wang, H., Zeng, X., Zhang, T., Wei, J., Hou, X., Wu, H., and Zhang, K., 2024. Hierarchical Kernel Interaction Network for Remote Sensing Object Counting, **IEEE Transactions on Geoscience and Remote Sensing**, **62**:1–13.
87. Fu, Q., Min, W., Sheng, W., and Peng, C., 2024. Counting dense object of multiple types based on feature enhancement, **Frontiers in Neurorobotics**, **18**.
88. Wang, Y., Yang, B., Wang, X., Liang, C., and Chen, J., 2024. SATCount: A scale-aware transformer-based class-agnostic counting framework, **Neural Networks**, **172**:106126.
89. Dalal, N. and Triggs, B., 2005. Histograms of oriented gradients for human detection, *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, 886–893 vol. 1.
90. Zhao, W., Guo, Z., Yue, J., Zhang, X., and Luo, L., 2015. On combining multiscale deep learning features for the classification of hyperspectral remote sensing imagery, **International Journal of Remote Sensing**, **36**(13):3368–3379.
91. Cheng, G., Yang, C., Yao, X., Guo, L., and Han, J., 2018. When Deep Learning Meets Metric Learning: Remote Sensing Image Scene Classification via Learning Discriminative CNNs, **IEEE Transactions on Geoscience and Remote Sensing**, **56**(5):2811–2821.
92. Krizhevsky, A., Sutskever, I., and Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks, *F. Pereira, C.J.C. Burges, L. Bottou, and*

K.Q. Weinberger, Eds.: Advances in Neural Information Processing Systems, volume 25, Curran Associates, Inc.

93. Niitani, Y., Ogawa, T., Saito, S., and Saito, M., 2017. ChainerCV: a Library for Deep Learning in Computer Vision, *ACM Multimedia*.
94. Han, X., Zhong, Y., and Zhang, L., 2017. An Efficient and Robust Integrated Geospatial Object Detection Framework for High Spatial Resolution Remote Sensing Imagery, **Remote Sensing**, **9**(7):666.
95. Kang, D., Ma, Z., and Chan, A.B., 2019. Beyond Counting: Comparisons of Density Maps for Crowd Analysis Tasks—Counting, Detection, and Tracking, **IEEE Transactions on Circuits and Systems for Video Technology**, **29**(5):1408–1422.
96. Redmon, J., Divvala, S., Girshick, R., and Farhadi, A., 2016. You Only Look Once: Unified, Real-Time Object Detection, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE.
97. Law, H. and Deng, J., 2018. CornerNet: Detecting Objects as Paired Keypoints, *The European Conference on Computer Vision (ECCV)*.
98. Zhou, X., Wang, D., and Krähenbühl, P., 2019. Objects as Points, **ArXiv preprint**, **ArXiv:1904.07850**.
99. Cazzato, D., Cimarelli, C., Sanchez-Lopez, J.L., Voos, H., and Leo, M., 2020. A Survey of Computer Vision Methods for 2D Object Detection from Unmanned Aerial Vehicles, **Journal of Imaging**, **6**(8):78.
100. Li, W., Li, H., Wu, Q., Chen, X., and Ngan, K.N., 2019. Simultaneously Detecting and Counting Dense Vehicles From Drone Images, **IEEE Transactions on Industrial Electronics**, **66**(12):9651–9662.
101. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., and Torralba, A., 2016. Learning Deep Features for Discriminative Localization, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE.

102. Xu, B., Wang, N., Chen, T., and Li, M., 2015, Empirical Evaluation of Rectified Activations in Convolutional Network.
103. Ioffe, S. and Szegedy, C., 2015, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.
104. Mundhenk, T.N., Konjevod, G., Sakla, W.A., and Boakye, K., 2016. A Large Contextual Dataset for Classification, Detection and Counting of Cars with Deep Learning, *B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds.; Computer Vision – ECCV 2016*, 785–800, Springer International Publishing, Cham.
105. Stahl, T., Pinteá, S.L., and van Gemert, J.C., 2019. Divide and Count: Generic Object Counting by Image Divisions, **IEEE Transactions on Image Processing**, **28**(2):1035–1044.
106. Aich, S. and Stavness, I., 2018. Global Sum Pooling: A Generalization Trick for Object Counting with Small Datasets of Large Images, **arXiv preprint**, arXiv–1805.
107. dos Santos de Arruda, M., Osco, L.P., Acosta, P.R., Gonçalves, D.N., Junior, J.M., Ramos, A.P.M., Matsubara, E.T., Luo, Z., Li, J., de Andrade Silva, J., and Gonçalves, W.N., 2021, Counting and Locating High-Density Objects Using Convolutional Neural Network.
108. Bearman, A., Russakovsky, O., Ferrari, V., and Fei-Fei, L., 2016. What’s the Point: Semantic Segmentation with Point Supervision, *Computer Vision – ECCV 2016*, 549–565, Springer International Publishing.
109. Kingma, D.P. and Ba, J., 2015. Adam: A Method for Stochastic Optimization, *Y. Bengio and Y. LeCun, Eds.; 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.
110. Lin, H.Y., Tu, K.C., and Li, C.Y., 2020. VAID: An Aerial Image Dataset for Vehicle Detection and Classification, **IEEE Access**, **8**:212209–212219.

111. Bochkovskiy, A., Wang, C., and Liao, H.M., 2020. YOLOv4: Optimal Speed and Accuracy of Object Detection, **arXiv preprint, arXiv:2004.10934**.
112. Zhang, S., Wen, L., Bian, X., Lei, Z., and Li, S.Z., 2017. Single-Shot Refinement Neural Network for Object Detection, **arXiv preprint, arXiv:1711.06897**.



ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı, Soyadı : Ersin Kılıç
Uyruğu : Türkiye (T.C.)

EĞİTİM

Derece	Kurum	Mez.Yılı
Lise	Nevşehir Anadolu Lisesi, NEVŞEHİR	2006
Lisans	KOÜ Elektronik ve Haberleşme Müh., KOCAELİ	2011
Yüksek Lisans	ERÜ Fen Bilimleri Ens. Bilgisayar Müh. ABD, KAYSERİ	2016

İŞ DENEYİMİ

Yıl	Kurum	Görev
2014-2018	ERÜ Mühendislik Fakültesi Bilgisayar Müh. Bölümü, KAYSERİ	Araştırma Görevlisi
2018-2019	ZBEÜ Mühendislik Fakültesi Bilgisayar Müh. Bölümü, ZONGULDAK	Araştırma Görevlisi
2019-2021	Akgün Bilgisayar, ANKARA	Bilgisayarlı Görü Mühendisi
2022-Halen	Turkcell Teknoloji, İSTANBUL	Bilgisayarlı Görü Mühendisi

YAYINLAR**SCI, SSCI, AHCI İndekslerine Giren Dergilerde Yayınlanan Makaleler**

- **E. Kılıç, S. Oztürk**, "A subclass supported convolutional neural network for object detection and localization in remote-sensing images", International Journal of Remote Sensing , 40:11, 4193-4212, DOI: 10.1080/01431161.2018.1562260.
- **E. Kılıç, S. Oztürk**, "An accurate car counting in aerial images based on convolutional neural networks", J Ambient Intell Human Comput (2021). <https://doi.org/10.1007/s12652-021-03377-5>.

TR Dizinli Dergilerde Yayınlanan Makaleler

- **E. Kılıç, S. Oztürk**, "İnsansız Hava Aracı Görüntülerinde Evrişimli Sinir Ağları Tabanlı Araba Sayımı Yöntemi", Erciyes Üniversitesi Fen Bilimleri Enstitüsü Fen Bilimleri Dergisi, vol. 36, no. 2, pp. 170-177, Aug. 2020