

**UZAKTAN ALGILAMA GÖRÜNTÜLERİNDE DERİN ÖĞRENME TEMELLİ
YAKLAŞIMLAR KULLANARAK NESNE TESPİTİ**

Nuri Erkin ÖÇER

DOKTORA TEZİ

Uzaktan Algılama ve Coğrafi Bilgi Sistemleri Anabilim Dalı

Danışman: Doç. Dr. Uğur AVDAN

Eskişehir

Eskişehir Teknik Üniversitesi

Lisansüstü Eğitim Enstitüsü

Şubat 2020

Bu tez çalışması BAP komisyonu tarafından kabul edilen 1707F461 no.lu proje kapsamında desteklenmiştir.

JÜRİ VE ENSTİTÜ ONAYI

Nuri Erkin ÖÇER'in UZAKTAN ALGILAMA GÖRÜNTÜLERİNDE DERİN ÖĞRENME TEMELLİ YAKLAŞIMLAR KULLANARAK NESNE TESPİTİ başlıklı tezi 18/02/2020 tarihinde aşağıdaki jüri tarafından değerlendirilerek "Eskişehir Teknik Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliği"nin ilgili maddeleri uyarınca, Uzaktan Algılama ve Coğrafi Bilgi Sistemleri Anabilim dalında Doktora tezi olarak kabul edilmiştir.

<u>Jüri Üyeleri</u>	<u>Unvanı Adı Soyadı</u>	<u>İmza</u>
Üye (Tez Danışmanı)	: Doç. Dr. Uğur AVDAN	
Üye	: Prof. Dr. Yasin Dursun SARI	
Üye	: Doç. Dr. Murat UYSAL	
Üye	: Dr.Öğr.Üyesi Emrah PEKKAN	
Üye	: Dr. Öğr. Üyesi Resul ÇÖMERT	

Prof. Dr. Murat TANIŞLI

Lisansüstü Eğitim Enstitüsü Müdürü

ÖZET

UZAKTAN ALGILAMA GÖRÜNTÜLERİNDE DERİN ÖĞRENME TEMELLİ YAKLAŞIMLAR KULLANARAK NESNE TESPİTİ

Nuri Erkin ÖÇER

Uzaktan Algılama ve Coğrafi Bilgi Sistemleri Anabilim Dalı

Eskişehir Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü, Şubat 2020

Danışman: Doç. Dr. Uğur AVDAN

Uzaktan algılama görüntülerinde nesne tespiti konusu özellikle yeryüzeyine dair alakalı gözlemlerde birçok alanda kendine yer bulmaktadır. Zengin detaya sahip uydu ve hava görüntülerinin elde edilmesiyle beraber insan yapımı birçok nesnenin algılanıp ayırt edilebilmesi mümkün hale gelmiştir. Bununla beraber, değişken nesne ölçekleri ve görünüşleri, ayrıca görüntülemeye dair bir takım etkenler bu görevi zorlu bir hale getirmektedir. Öte yandan, derin öğrenme alanında gerçekleşen ilerlemeler ve bunların uzaktan algılama alanına uyarlanması nesne algılama görevinde önemli iyileştirmeler sağlamıştır. Özellikle, konvolüsyonel sinir ağlarının kullanıldığı bölge-tabanlı nesne tespit araçları doğal görüntülerdeki tespit görevlerinde insan başarımını geride bırakmıştır. Bu yapıdaki bir model olan Mask R-CNN verilen görüntüdeki hedef nesnelere maskeleriyle beraber çıkarma kabiliyetindeki en güncel tespit sistemlerinden biridir. Bu tezde, tespit hedefi olarak gemi örneği ele alınarak uydu görüntülerinde yer alan bu sınıftaki nesnelere tespit için Mask R-CNN modelinin kullanımı önerilmiştir. Modelin farklı ölçeklerde de yüksek doğrulukta çalışabilmesi için özellik haritalarının yüksek semantik değere sahip üst katmanlarıyla mekansal çözünürlüğü fazla olan alt katmanlarının kaynaştırıldığı özellik piramit ağının kullanımı önerilmiştir. Ayrıca, üretilen yanlış alarmları azaltmak adına modelin negatif örnekleme eğitimi veya kayıp hesabında odak kayıp fonksiyonunun kullanımı önerilmiştir. Bu çalışma kapsamında sunulan öneriler, oluşturulan veri setleri üzerinde değerlendirilmiş ve doğruluk, kesinlik ve F1 skorları açısından gemi tespiti konusundaki en iyi performansı göstermiştir.

Anahtar Sözcükler: Nesne tespiti, Derin öğrenme, Uzaktan algılama, Konvolüsyonel sinir ağları

ABSTRACT

OBJECT DETECTION BY USING DEEP LEARNING BASED APPROACHES IN REMOTE SENSING IMAGES

Nuri Erkin ÖÇER

Department of Remote Sensing and Geographical Information Sciences

Eskişehir Technical University, Institute of Graduate Programs, February 2020

Supervisor: Associate Professor Doctor Uğur AVDAN

The issue of object detection in remote sensing images finds its place in many areas, especially in observations related to the earth. With the acquisition of richly detailed satellite and aerial images, it has become possible to detect and distinguish many man-made objects. However, variable object scales and appearances, as well as a number of display factors make this task challenging. On the other hand, advances in the field of deep learning and their adaptation to the field of remote sensing have provided significant improvements in the object detection task. In particular, region-based object detection tools using convolutional neural networks outperformed human performance in detection tasks in natural images. Mask R-CNN, which is a model in this structure, is one of the most up-to-date detection systems capable of extracting target objects in the given image with their masks. In this thesis, the use of the Mask R-CNN model for the detection of objects of this class in satellite images is proposed by considering the ship sample as a detection target. In order for the model to work at different scales with high accuracy, it is proposed to use the feature pyramid network, in which the upper layers of feature maps with high semantic values and the lower layers with high spatial resolution are fused. In addition, in order to reduce false alarms, training of the model with negative sampling or use of the focal loss function in the loss calculation is proposed. In order to perform a more quantitative error analysis, it is suggested that test images be classified according to their contents and the dimensions of the objects. Suggestions presented within the scope of this study were evaluated on the data sets created and showed the best performance in ship detection in terms of accuracy, precision and F1 scores.

Keywords: Object detection, Deep learning, Remote sensing, Convolutional neural networks

TEŐEKKÜR

Bu alıŐma s¼recinde birikimlerini benimle paylaŐan, yol g¼steren, her t¼rl¼ desteęi vermiŐ olan deęerli hocam ve danıŐmanım Do. Dr. Uęur AVDAN'a, yaptıkları y¼nlendirmelerle bu alıŐmada b¼y¼k emekleri gemiŐ Do. Dr. Murat UYSAL, Dr. Őęr. Üyesi Emrah PEKKAN, Prof. Dr. Yasin Dursun SARI ve Dr. Őęr. Üyesi Resul ÖMERT hocalarıma ve her daim destekleri benimle olan aileme teŐekk¼rlerimle.

Nuri Erkin ŐER



ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Bu tezin bana ait, özgün bir çalışma olduğunu; çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalarında bilimsel etik ve kurallara uygun davrandığımı; bu çalışma kapsamında elde edilen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi; bu çalışmanın Eskişehir Teknik Üniversitesi tarafından kullanılan “bilimsel intihal tespit programı”yla tarandığını ve hiçbir şekilde “intihal içermediğini” beyan ederim. Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçları kabul ettiğimi bildiririm.

Nuri Erkin ÖÇER

İÇİNDEKİLER

Sayfa

BAŞLIK SAYFASI	i
JÜRİ VE ENSTİTÜ ONAYI	ii
ÖZET	iii
ABSTRACT.....	iv
TEŞEKKÜR	v
ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ.....	vi
İÇİNDEKİLER	vii
TABLolar DİZİNİ	x
ŞEKİLLER DİZİNİ	xii
SİMGELER VE KISALTMALAR DİZİNİ	xvii
1. GİRİŞ	1
1.1. Literatür Özeti.....	3
1.2. Araştırma Kapsamı.....	13
1.3. Özgün Değer	13
1.4. Tez Yapısı.....	14
2. TEORİ.....	15
2.1. Bölge Tabanlı R-CNN Yöntemleriyle Nesne Tespiti: Faster R-CNN.....	15
2.1.1. Özellik çıkarımı	15
2.1.2. Bölge öneri ağı	17
2.1.2.1. Çapalar	17
2.1.2.2. RPN'de eğitim: hedefler ve kayıp fonksiyonları.....	21
2.1.2.3. Art işleme	24
2.1.3. İlgili bölgesi havuzlaması	25
2.1.4. Bölge-temelli konvolüsyonel sinir ağı	25
2.1.4.1. R-CNN modülünde hedefler ve kayıp fonksiyonu	27
2.1.4.2. Art işleme	28
2.1.4.3. Eğitim	28

3. YÖNTEM.....	30
3.1. Mask R-CNN Tabanlı Gemi Tespit Çerçevesi.....	31
3.1.1. Nesnenin maskeyle temsili: maske tahmin kafası	34
3.1.2. İlgi bölgelerinin (RoI) hizalanması	35
3.2.1. Aşağıdan yukarıya yol.....	37
3.2.2. Yukarıdan-aşağıya yol	38
3.2.3. FPN'nin bölge öneri ağında kullanımı	39
3.2.4. FPN'nin Mask R-CNN'de kullanımı	40
3.3. Hatalı Tespitlerin Azaltılması İçin Yapılan Düzenlemeler	41
3.3.1. Mask R-CNN-tabanlı gemi tespiti için negatif örneklem eğitimi.....	42
3.3.2. Gemi tespitinde Mask R-CNN ve odak kayıp fonksiyonunun kullanımı.....	44
3.3.2.1. <i>Eğitim</i>	45
3.4. Veri Kümesinin İçeriğe ve Gemi Büyüklüklerine Göre Tanzimi	46
4. VERİ SETİNE VE MODEL EĞİTİMİNE DAİR DETAYLAR	48
4.1. Veri Seti.....	48
4.2. Eğitim	53
4.3. Tespit Değerlendirme Metrikleri.....	54
5. DENEYLER VE BULGULAR	56
5.1. Nesne Güven Eşiğinin Tespit Üzerindeki Etkisi.....	56
5.3. Karşılaştırma: Önerilen Yöntemin Gemi Tespit Performansına Etkisi	59
5.3.1. Önerilen yöntemin diğer modellere göre performansı.....	60
5.3.2. Görüntü çözünürlüğünün tespit performansına etkisi	61
5.4. Farklı veri artırım yöntemleriyle eğitim: önerilen yöntemin performansı.....	65
5.5. Hatalı Tespitlerin Azaltılmasına Dair Deneyler	66
5.5.1. Mask R-CNN-tabanlı gemi tespiti için negatif örneklem eğitimi.....	66
5.5.2. Tespitte odak kayıp fonksiyonunun kullanımı: RetinaNet.....	75
5.5.3. Karşılaştırma: negatif örnekle eğitim ve odak kayıp fonksiyonunun hatalı tespitler üzerindeki etkisi	77
6. TARTIŞMA	79

	<u>Sayfa</u>
6.1. NMS Eşığı	79
6.2. RPN'de Dönük Sınır Kutuları Kullanımı	81
6.3. Gerçek-Zamanlı Tespit	82
7. SONUÇ VE ÖNERİLER.....	83
KAYNAKÇA	85
ÖZGEÇMİŞ	



TABLULAR DİZİNİ

Sayfa

Tablo 4.1. Çalışma kapsamında kullanılan veri kümesinin içeriği	48
Tablo 4.2. Veri kümesi içeriğindeki gemi sayıları	48
Tablo 4.3. Veri kümelerini oluşturan görüntülerin içerdikleri gemi sayılarına dair bilgiler	49
Tablo 4.4. Gemilerin görüntüler üzerinde kapladıkları alanlara dair istatistikler	51
Tablo 4.5. Görüntülerdeki gemilerin uzunluklarına dair istatistikler	52
Tablo 4.6. Eğitimde kullanılan öğrenme oranının turlara göre değişim takvimi	53
Tablo 5.1. Önerilen modelin (veri artırımı olmadan) farklı güven eşiği değerlerinde ürettiği gemi tespit sonuçları	57
Tablo 5.2. Çeşitli modeller için görüntü başına düşen hesaplama sürelerinin karşılaştırılması	59
Tablo 5.3. Önerilen model ile diğer modellerin test veri kümesi üzerindeki tespit performansının karşılaştırması	60
Tablo 5.4. Test veri kümesi (orijinal görüntü çözünürlüğü = 1m) kullanılarak yer örneklem aralıkları kabalaştırılmış veriler üzerinde üç farklı modelin ürettiği tespit neticeleri	62
Tablo 5.5. Artırılmış veri setleri ve bunlarda görüntülere uygulanan işlemler	66
Tablo 5.6. Farklı artırma yöntemleriyle oluşturulmuş veri kümelerinin kullanılmasıyla eğitilmiş olan önerilen modelin test kümesi üzerindeki tespit sonuçları	66
Tablo 5.7. Eğitim ve validasyon verilerine eklenen negatif örneklem görüntü sayılarına ve eğitime dair bilgiler	67
Tablo 5.8. Test amacıyla kullanılan büyük boyutlu görüntülere dair bilgiler	68
Tablo 5.9. BG-1 görüntüsü için farklı negatif örneklem oranlarında gemi tespit sonuçları ve harcanan işlem zamanı	71
Tablo 5.10. BG-2 görüntüsü için farklı pozitif/negatif örneklem oranlarında gemi tespit sonuçları ve harcanan işlem zamanı	73
Tablo 5.11. BG-3 görüntüsü için farklı negatif örneklem oranlarında gemi tespit sonuçları ve harcanan işlem zamanı	74
Tablo 5.12. BG-1 görüntüsü için farklı γ değerlerindeki tespit sonuçları ve model eğitiminde harcanan zaman	76

Tablo 5.13. BG-2 görüntüsü için farklı γ değerlerindeki tespit sonuçları ve model eğitiminde harcanan zaman.....	76
Tablo 5.14. BG-3 görüntüsü için farklı γ değerlerindeki tespit sonuçları ve model eğitiminde harcanan zamanlar	76
Tablo 5.15. Yanlış alarmları azaltmaya dönük olarak bu çalışmada kullanılan yöntemlerin birbirleriyle ve baz modellerle olan karşılaştırması.....	77



ŞEKİLLER DİZİNİ

Sayfa

- Şekil 1.1.** a) Sınıflandırma amacıyla kullanılan bir görüntü örneği; nesne görüntüyü kaplamaktadır. b) Tespit görevinde, nesnelere görüntü içerisinde konumlarıyla beraber çıkarılmaktadır; nesnelere görüntü içerisinde kapladıkları alan küçüktür. 5
- Şekil 1.2.** Gemi büyüklüklerinin ve görüntüleme açılarının a) doğal fotoğrafta ve b) uydu görüntüsünde karşılaştırmaları. 8
- Şekil 1.3.** Uydu görüntüsünde yer alan bir geminin a) dik sınırlayıcı kutu, b) dönük sınırlayıcı kutularla işaretlenmesi 8
- Şekil 1.4.** Dönük sınırlayıcı kutular dik emsallerine göre nesneyi daha iyi temsil etme becerisindedir. 9
- Şekil 1.5.** Uydu görüntüsündeki bir geminin ikili (binary) maskelemesi..... 9
- Şekil 1.6.** Bir görüntü piramidinin en alt katmanında orijinal çözünürlükteki görüntü yer alırken üst katmanlar bunun kabalaştırılmış hallerinden oluşmaktadır. Tahminler her ayrı düzey için yapılır..... 11
- Şekil 1.7.** Gemiyle benzerlik taşıyan ve model için kafa karıştırıcı nitelikte olan negatif örnekler 12
- Şekil 2.1.** Bir görüntü için özellik çıkarımı omurga ağ aracılığıyla gerçekleştirilir; derinlik (semantik değer) artarken görüntü boyutu küçülür. 16
- Şekil 2.2.** 5 bloktan oluşan ResNet ağ yapısında her bir blok arasında özellik haritalarının mekansal boyutları yarıya düşmekte, derinliği artmaktadır. 16
- Şekil 2.3.** Bölge Öneri Ağı'yla görüntü içerisindeki nesne olması muhtemel bölgeler belirlenmektedir. 17
- Şekil 2.4.** Oluşturulan çapaların görüntü üzerindeki konumu. Ağın 5. bloğundan çıkan özellik haritası kullanıldığı takdirde, çapalar orijinal görüntü üzerinde birbirinden $2^{5-1}=16$ piksel aralıkla oluşturulacaktır..... 18
- Şekil 2.5.** Faster R-CNN'de çapalar üç farklı büyüklük skalasında oluşturulur: a) 128x128 piksel, b) 256x256 piksel, c) 512x512 piksel. 19
- Şekil 2.6.** Özellik haritasındaki aynı konumda oluşturulan 3 farklı skaladaki ve 3 farklı en boy oranındaki çapalar. 19

Şekil 2.7. RPN’de, özellik haritasına uygulanan 3x3’lük filtrenin akabinde biri, her bir çapa için nesnellik skoru hesabını yapan, diğeri her bir çapa için yer-doğruluğuna göre gerekli düzeltmeleri gerçekleştiren iki 1x1’lik filtre kullanılmaktadır.....	20
Şekil 2.8. Çapa kutusuna dair parametreler, yer-doğruluğuna dair parametrelerle arasındaki farklara bakılarak ayarlanmaktadır.....	21
Şekil 2.9. RPN’de üretilen nesne önerilerinin yer-doğruluklarıyla olan IoU oranına bakılarak ön plan veya arka plan olarak ataması gerçekleştirilir.....	22
Şekil 2.10. Bir nesne için farklı çapalardan gelen önerilerin (a) birden fazla tespiti yol açmasını engellemek üzere birbirleriyle belirli bir oranın üzerinde çakışan önerilerden en yüksek nesnellik skoruna sahip öneri haricindekiler elenmektedir (b).....	24
Şekil 2.11. Faster R-CNN’nin RPN kısmında üretilen farklı boyutlardaki nesne önerileri havuzlama katmanında sabit boyuta getirilir.....	25
Şekil 2.12. Faster R-CNN’nin son modülünde nesne sınıfının belirlenmesi ve sınırlayıcı kutunun nesneye göre daha iyi ayarlanması gerçekleştirilir.....	27
Şekil 3.1. Çalışma sürecinin özeti	31
Şekil 3.2. Mask R-CNN modelinde işlem adımları.....	33
Şekil 3.3. Mask R-CNN’le girdi görüntüsü kullanılarak nesnelere dair sınıf, sınırlayıcı kutu ve maske tahminleri üretilmektedir.....	34
Şekil 3.4. Mask R-CNN modelinde üç çıktı alınmaktadır: Nesne sınıfı, nesne sınırlayıcı kutusu ve nesne maskesi.....	35
Şekil 3.5. RoI Hizalama: Mask R-CNN modelinde Faster R-CNN’den farklı olarak, özellik haritasının hücre değerlerinin enterpolasyonu kullanılarak öneri bölgelerinin nesnelere daha iyi oturması sağlanmaktadır.....	36
Şekil 3.6. Mask R-CNN makalesinde RoI Hizalamamın (Align) havuzlama katmanına göre nesne tespiti üzerindeki etkinliği gösterilmiştir.....	36
Şekil 3.7. Nesne tespitinde özellik piramit ağının kullanımı.....	37
Şekil 3.8. Özellik piramit ağında yukarıdan-aşağıya tanımlanan yeni yolda mekansal çözünürlüğü düşük, semantik değeri yüksek özellik haritaları,	

alt katmanla kurulan yanıl baęlantılar ve üst-örnekleme kullanılarak semantik değeri korunarak mekansal boyutu artırılmaktadır.....	38
Şekil 3.9. Yukarıdan-aşağıya yolda mekansal çözünürlük, alt katmanla yanıl baęlantının da kullanımıyla 2 katına çıkmaktadır.....	39
Şekil 3.10. Nesne tespitinde FPN kullanıldığında RPN'deki tahminler, oluşturulan yeni P_i katmanları kullanılarak gerçekleştirilir.	40
Şekil 3.11. Mask R-CNN modelinde RPN modülünün a) omurga aędan gelen özellik haritasıyla, b) FPN'de üretilen özellik haritalarıyla beslenmesi. FPN kullanıldığında piramitteki hangi katmanın kullanılacağı nesne boyutuna baęlıdır.	41
Şekil 3.12. Modelin eğitiminde kullanılan pozitif örneklem görüntüleri	43
Şekil 3.13. Modelin eğitiminde kullanılan negatif örneklem görüntüleri	44
Şekil 4.1. Veri kümelerindeki gemi sayılarının dağılımı.....	49
Şekil 4.2. Eğitim, validasyon ve test kümelerindeki gemilerin kapladıkları alanlara dair istatistikler.....	51
Şekil 4.3. Eğitim, validasyon ve test kümelerindeki gemilerin uzunluklarına dair istatistikler.....	52
Şekil 5.1. Önerilen modelin farklı güven eşięi değeriinde ürettięi gemi tespit sonuçlarının grafięi	58
Şekil 5.2. Farklı omurgalarla çalıştırılan modellerin test kümesinde ürettikleri tespitlerin F1 skorları cinsinden karşılaştırmaları.....	59
Şekil 5.3. Test veri kümesi üzerinde çalıştırılan farklı modellere ait kesinlik-geri getirme eğrileri.....	61
Şekil 5.4. Farklı görüntü çözünürlükleri için üç modelin geri getirme (doęruluk) performansı	63
Şekil 5.5. FPN kullanılmadığı durumda, farklı görüntü çözünürlükleri için, Mask R-CNN modelinin verdięi, farklı gemi uzunluklarına göre ayırımı yapılmış geri getirme değeri.	64
Şekil 5.6. Mask R-CNN modelinin FPN ile beraber kullanıldığı durumda, farklı görüntü çözünürlükleri için, farklı gemi uzunluklarına göre ayırımı yapılmış geri getirme değeri.	65

Şekil 5.7. Yanlış alarmların azaltılmasına dönük gerçekleştirilen deneylerde kullanılan, farklı kara-deniz içeriklerine sahip büyük ölçekli görüntüler.....	68
Şekil 5.8. Boyutu 5120x5120 pikselin üzerinde olan görüntüler bu büyüklükteki parçalara bölünerek modele girilmektedir.	69
Şekil 5.9. Sol sütun: Orijinal görüntüler. Sağ sütun: Sadece pozitif örneklerle eğitilmiş ağ tarafından üretilen tespit sonuçları	70
Şekil 5.10. BG-1 görüntüsü içinden bir kesitte a) eğitimde pozitif örneklem sayısı kadar negatif örneklemin kullanıldığı modelle yapılan tespit, b) eğitimde sadece pozitif örneklerin kullanıldığı modelle yapılan tespit	72
Şekil 5.11. BG-1 görüntüsü içinden bir kesitte, model eğitiminde pozitif/negatif örneklem oranı olarak a) $P/N=1:1$, b) $P/N=1:0$ kullanıldığında üretilen yanlış alarmlar.....	72
Şekil 5.12. BG-2 görüntüsü içinden bir kesitte a) eğitimde pozitif örneklem sayısı kadar negatif örneklemin kullanıldığı modelle yapılan tespit, b) eğitimde sadece pozitif örneklerin kullanıldığı modelle yapılan tespit	73
Şekil 5.13. BG-2 görüntüsü içinden bir kesitte, model eğitiminde pozitif/negatif örneklem oranı olarak a) $P/N=1:1$, b) $P/N=1:0$ kullanıldığında üretilen yanlış alarmlar.....	74
Şekil 5.14. BG-3 görüntüsü içinden bir kesitte a) eğitimde pozitif örneklem sayısı kadar negatif örneklemin kullanıldığı modelle yapılan tespit, b) eğitimde sadece pozitif örneklerin kullanıldığı modelle yapılan tespit	75
Şekil 5.15. Yanlış alarmları azaltmaya yönelik bu çalışmada kullanılan yöntemlerin birbirleriyle ve baz modelle olan karşılaştırması.....	78
Şekil 6.1. Orijinal görüntüsü a-sütununda verilen görüntüler için modelin ürettiği hatalı tespitler (b-sütunu)	79
Şekil 6.2. Şekil 6.1'deki görüntüde yer alan gemilere ait (a) yer doğrulukları (b) oluşturulmuş 64x64 piksel büyüklük skalasındaki komşu çapalar	80
Şekil 6.3. Şekil 6.2'de yer alan A gemisinin yer doğruluğunun Çapa-1'le olan kesişim/birleşim durumu, (b) B gemisinin yer doğruluğunun Çapa-1'le olan kesişim birleşim durumu	81

Şekil 6.4. Yang vd. tarafından (2018) önerilen, RPN’de nesne sınırının temsili için dik sınırlayıcı kutu yerine kullanılmak üzere tanımlanan dönük sınırlayıcı kutu 82



SİMGELER VE KISALTMALAR DİZİNİ

α	: Alpha
γ	: Beta
AVS	: Artırılmış Veri Seti
BG	: Büyük-Ölçekli Görüntü
CNN	: Convolutional Neural Networks
COCO	: Common Objects in Context
FN	: False Negative
FP	: False Positive
FPN	: Feature Pyramid Network
IoU	: Intersection over union
FN	: False Negative
FP	: False Positive
konv	: Konvolüsyonel özellik haritası
lr	: Learning Rate
NMS	: Non-Maximum Suppression
R-CNN	: Region-based Convolutional Neural Networks
reg	: Regresyon
ResNet	: Residual Networks
RoI	: Region of Interest
RPN	: Region Proposal Network
SAR	: Synthetic-aperture radar
SPP-net	: Spatial Pyramid Pooling Net
SSD	: Single Shot MultiBox Detector
TP	: True Positive
YOLO	: You Only Look Once

1. GİRİŞ

Uzaktan algılamada nesne tespiti, hava ve uydu görüntüleri analizi alanında en temel ve ilgi çekici konulardan biri olarak birçok uygulamada önemli rol oynamakta ve gün geçtikçe daha fazla ilgi toplamaktadır. Uzaktan algılamada nesne tespiti, eldeki hava veya uydu fotoğrafında yer alan hedef objelerin ait olduğu alakalı sınıfların belirlenmesini ve tahmini yapılan görüntüdeki objelerin konumlarının tespitini içermektedir. Bu kapsamdaki çalışmalarda “obje” terimi gemiler, uçaklar, taşıtlar, binalar gibi arka plan ortamından bağımsız, keskin sınırlarla ayrılmış insan yapımı nesnelere olabildiği gibi toprak parçaları, bitki örtüsü gibi kesin çizgileri olmayan sınırlara sahip ve arka planından veya içerikten ayrılması zor, tabii yeryüzü cisimlerini de ifade etmektedir. Uzaktan algılanmış görüntülerde nesne tespiti çevresel görüntüleme (Aryal & Josselin, 2019), jeolojik afet tespitine (Amit & Aoki, 2017), tarım ve bitki örtüsü haritalamasından (Blaschke, Johansen, & Tiede, 2011) coğrafi bilgi sistemlerinde güncellemelere (Ardeshir, Zamir, Torroella, & Shah, 2014), çevreye duyarlı tarımdan (Sadgrove & Falzon, 2018) şehir planlamaya (Serna & Marcotegui, 2014) kadar birçok alanda önemli bir rol oynamaktadır.

Uzaktan algılamada nesne tespiti, objelerin dış görünüşlerinde büyük değişikliklere yol açan görüntüleme noktası değişimi, görüntü blokajı, arka plan düzensizlikleri, aydınlatma, gölgeleme gibi çeşitli zorluklar içermektedir (Cheng, G., 2016). Ayrıca her geçen gün sayısı ve kalitesi artan hava ve uydu fotoğrafları ve yeni uygulama alanlarında duyulan çeşitli gereksinimler uzaktan algılamada nesne tespitini daha da önemli hale getirmektedir. Bu sorunlara çözüm arayışında olan coğrafi mekânsal obje tespiti konusu 80’li yıllardan beri yaygın olarak çalışılmaktadır. Landsat gibi düşük mekânsal çözünürlüklü eski nesil uydu görüntüleri genellikle insan yapımı veya doğal nesnelere tespit edilmesinde yetersiz kaldığından araştırmacılar çoğunlukla bunlardan ancak bölgesel özellik çıkarımı yapmaya çabalamışlardır. Uzaktan algılama teknolojisindeki ilerlemelerle çok yüksek çözünürlüğe sahip IKONOS, SPOT-5 ve Quickbird gibi uydu ve hava fotoğrafları daha detaylı konumsal ve dokusal bilgiye ulaşımı mümkün kılmıştır. Böylelikle, bölgesel özellik çıkarımı bir yana, insan yapımı birçok nesnenin algılanıp ayırt edilebilmeleri mümkün hale gelmiştir. Bu da nesnelere otomatik tespitinde yeni olanakların önünü açmıştır.

Uzaktan algılamada nesne tespiti alandaki çalışmaların büyük bir kısmı uydu ve hava görüntülerinde yer alan belirli türdeki objelerin algılanması için çeşitli yöntemlerin geliştirilmesi üzerine olmuştur, örneğin; yollar için (Xin, Zhang, Zhang, & Fang, 2019); (Shahi, Shafri, Taherzadeh, Mansor, & Muniandy, 2015), binalar için (Prathap & Afanasyev, 2018); (Akçay & Aksoy, 2010), gemiler konusunda (Yang & Sun, 2018); (Liu, Cui, Kuang, & Li, 2017), ağaçlar konusunda (Weinstein, Marconi, Bohlman, Zare, & White, 2019); (Santos & Junior, 2019), taşıtlar için (H., X., Liang, Wu, & Chen, 2019). Bu çalışmaların ortak özelliği, nesne tespitini uçak, gemi, bina, yol gibi belirli sınıflar için ele almalarıdır.

Bu çalışmada ise uzaktan algılama görüntülerinde nesne tespiti konusu gemi örneği üzerinden ele alınmaktadır. Gemilerin görüntülerde buldukları içerikler itibariyle tespit edilebilmeleri kimi bölgelerde geleneksel yöntemlerle çözülebilecek ölçüde basitken, kimi bölgelerde ve durumlarda klasik yaklaşımlar tamamen yetersiz kalmaktadır. Gemiler nesne büyüklükleri açısından da geniş bir yelpazededirler. Ayrıca bazı bölgelerde birbirlerine yakın, bazı yerlerde aralıklı konumlanmaları uzaktan algılama görüntülerinde yer alan diğer sınıflardaki nesnelere de ortak paydalarındandır. Bu açılarından, gemilerin uzaktan algılama görüntülerinde tespiti problemi benzer türdeki nesnelere tespit için de önem arz etmektedir.

Kullanıldığı alanlara bakıldığında, uzaktan algılama görüntülerinde gemi tespiti konusu ülke karasularının güvenliği, deniz gözetimi, liman yönetimi, deniz taşımacılığı ve insani kurtarma faaliyetleri gibi birçok alanda kendine yer bulmaktadır. Bu bağlamda gemi tespiti, eldeki hava veya uydu fotoğrafında yer alan gemilerin görüntü içerisinde ayırte edilmesini ve konumlarının tespitini içermektedir. Son yıllarda uzaktan algılama teknolojisindeki gelişmelerle birlikte sayısı her geçen gün artan yüksek çözünürlükteki görüntüler gemi tespitinde önemli fırsatlar sağlamıştır. Niteliksel ve niceliksel olarak giderek büyümekte olan bu veriler geliştirilen yeni algılama modelleriyle beraber gemilerin otomatik olarak çıkarımının önünü açmıştır. Bununla beraber, görüntülerde yer alan gemilerin sahip oldukları ince uzun yapı, boyutlarının geniş bir yelpazede farklılık göstermesi, göreceli olarak karmaşık, zorlayıcı arka planlarda/çevrede yer almaları ve yönelimlerinin herhangi bir açıda olabilmesi gibi yapısal karakteristikler ve ayrıca görüntüleme noktası değişimi, görüntü blokajları, ışık problemleri, gölgelemeler gibi görüntülemeyle alakalı etkenler gemi tespitini zorlu bir görev haline getirmektedir.

1.1. Literatür Özeti

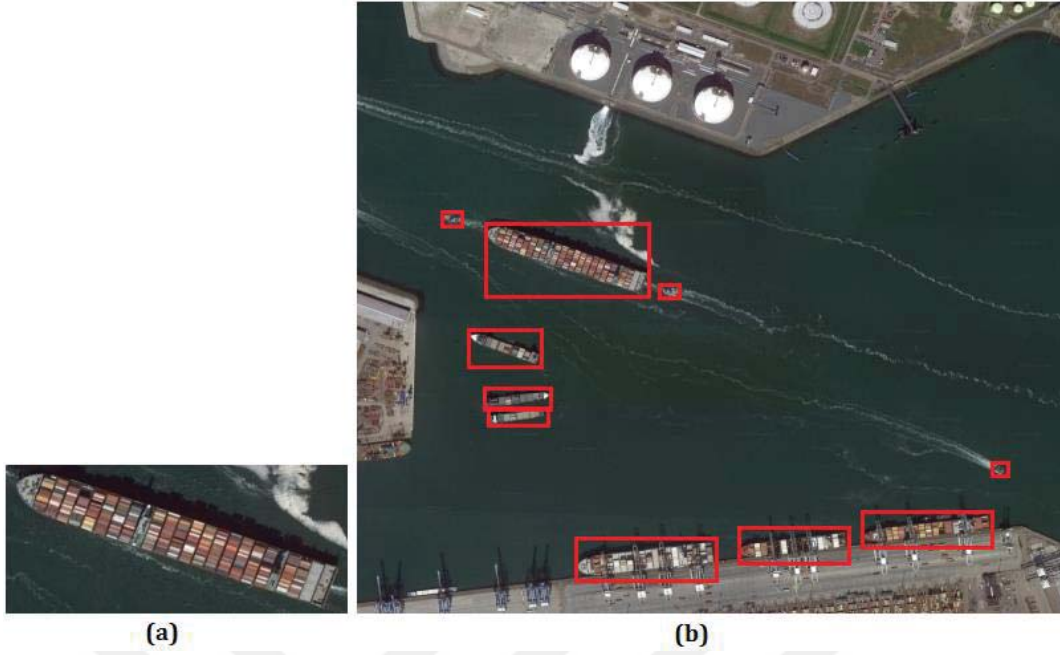
Literatürdeki hem eski hem yeni yaklaşımlarda gemi tespiti görevi temel olarak nesnelere dair özelliklerin çıkarılmasından ve sınıflandırılmasından oluşmaktadır. Diğer nesne algılama uygulamalarında olduğu gibi tespit performansı, gemilere dair özelliklerin tanımlanmasındaki yeterlilik ile orantılıdır. Bu nedenle özellik çıkarımı, başarılı bir gemi tespit sistemi oluşturmada önemli bir adımdır. Özellik çıkarımını gerçekleştiren özneye göre literatürdeki gemi tespit çalışmaları (1) insan eliyle çıkarılan özelliklere dayanan yöntemler ve (2) makine öğrenimiyle çıkarılan özelliklere dayalı olanlar olmak üzere iki gruba ayrılabilir.

İnsan eliyle çıkarılan özellikler söz konusu olduğunda gemiler hakkında önceden edinilmiş bilgilerden (uzman bilgilerinden) faydalanarak bazı özellik çıkarma yöntemleri önerilmiştir. Bu yaklaşımda çoğunlukla şekil, kenar ve doku özellikleri ön plandadır (Zhu, Zhou, Wang, & Guo, 2010). Ancak, uydu görüntülerindeki gemilerin bir kısmının sadece birkaç piksel boyutunda olması veya görüntü netliğinin kötü olması gibi durumlar söz konusu olduğundan bu tür nesnelere insan tarafından çıkarılan özellikler aracılığıyla ayırtılabilmesi mümkün olmamaktadır. Ayrıca, bu yolla çıkarılan özelliklerin genelleştirme kabiliyetinin zayıf kaldığı, nesne için belirli koşullar altında türetilen bilgilerin farklı türdeki gemilerde veya görüntüleme şartlarında işe yaramayarak algılama performansının düştüğü gözlenmiştir (Antelo, Ambrosio, Gonzalez, & Galindo, 2009). Konu üzerine bir çalışmada (Shi, Yu, Jiang, & Li, 2014) gemi boyutlarının çeşitliliğinin performans üzerindeki olumsuz etkisi gözlenmiştir. Benzer şekilde, HOG-temelli sınıflandırıcıların da kalabalık gemi içeriğinde veya karışık içerikteki durumlarda başarısız olduğu bilinmektedir (Dalal & Triggs, 2005), (Lu, Liu, & Wang, 2015).

Diğer taraftan, makine-öğrenmesi alanında yaşanan gelişmelerle beraber (Hinton & Salakhutdinov, 2006) nesne tespiti bir sınıflandırma problemi olarak ele alınmaya başlanmış ve önemli ilerlemeler kaydedilmiştir (Krizhevsky, Sutskever, & Hinton, 2012). Bu yaklaşım, sahip olunan etiketli bilgiden yararlanarak nesne özelliklerinin otomatik olarak öğrenilmesini ve insan algılamasının tasvir etmeye yetmediği nesne karakteristiklerinin ve ayrıştırma kurallarının bulunmasını sağlamaktadır. Nesne görüntülerindeki değişimleri, farklı görüş açılarını, çeşitli ışık/hava koşullarındaki görüntüleri içeren eğitim verisinin sınıflandırıcı tarafından denetimli veya yarı-denetimli altyapılarda öğrenilmesiyle gerçekleşen makine öğrenme temelli nesne tespiti kendinden

önceki yöntemlere göre daha geniş bir görüntüleme yelpazesinde daha hızlı ve doğru sonuçlar üretmiştir (Han, Zhang, Cheng, Guo, & Ren, 2015). Ancak, SVM (Destek Vektör Makineleri) gibi geleneksel makine öğrenme algoritmalarında girdilerin eğitim standartından sapması durumunda yapılması gereken hesaplamalar ve işlem zamanı üssel olarak artmakta ve performans etkilenmektedir (Xia, Wan, & Yue, 2011).

Özellikle konvolüsyonel sinir ağları (Convolutional Neural Networks-CNN) aracılığıyla sınıflandırma alanındaki yetkinliğini gösteren derin öğrenme yöntemi en revaçta makine öğrenimi dallarından biridir. Krizhevsky vd. (2012) tarafından önerilen CNN yapısı üzerinde sağlanan gelişmelerle son yıllarda görüntü sınıflandırması ve nesne tespitinde önemli ilerlemeler sağlanmıştır. Detayları 2. Bölüm’de ifade edilen çeşitli mimarileri kullanan CNN, eğitiminde kendisine öğretilen nesne yapıları dışında tespiti yapılması istenen yeni girdilerin belirli derecede mekansal kaymalardan, ölçeğinin değişmesinden veya biçiminin bozulmasından etkilenmeyen, doğru tahminler üretebilme yeteneğinde bir yapıdır (Cheng, Zhou, & Han, 2016). Bu yapıdaki (Krizhevsky, Sutskever, & Hinton, 2012), GoogleNet (Szegedy, Liu, & Jia, 2014) ve ResNet (He, Zhang, Ren, & Sun, 2015) gibi modeller kabiliyetlerini, içerdiği görüntülerde 1000 ayrı sınıftan nesnenin yer aldığı ImageNet gibi veri setleri üzerinde yüksek sınıflandırma doğrulukları sağlayarak göstermişlerdir. Sınıflandırma kabiliyeti yüksek olan bu yöntemin görüntü içerisindeki nesnelere tespit edebilmesi ise bir takım zorluklar içermektedir. Bu modeller, nesnenin içeriği büyük oranda kapladığı ve boyutları sabit olan görüntüleri işleme (sınıflandırma) yeteneğindedirler (Şekil 1.1.a). Ancak söz konusu, boyutları ağır işleyebileceğinden büyük olan bir görüntü içerisinde yer alıp içerikte küçük yer kaplayan nesnelere tespit edildiğinde (Şekil 1.1.b) belirli stratejilerin takip edilmesine ihtiyaç duyulmuştur.



Şekil 1.1. a) Sınıflandırma amacıyla kullanılan bir görüntü örneği; nesne görüntüyü kaplamaktadır. b) Tespit görevinde, nesnelere görüntü içerisinde konumlarıyla beraber çıkarılmaktadır; nesnelere görüntü içerisinde kapladıkları alan küçüktür.

Başlangıçta bu sorun, görüntünün ağırlık işleyebileceği boyuttaki yamalar halinde parçalara ayrılıp ağa beslenmesi ve görüntü parçaları üzerinde sınıflandırmaların tek tek yapılarak kesik parçaların koordinatlarının elde edilmesiyle çözülmeye çalışılmıştır. Kayan pencereler (Sliding Windows) adı verilen bu yaklaşımla hem optik hem SAR görüntülerinde birçok nesnenin tespiti gerçekleştirilmiştir (Lee, Bang, & Yang, 2017), (Huang & Huang, 2016). Lakin bu yöntemde nesne boyutları çeşitlilik arz ettiğinde tek bir tarama pencere boyutu ve en-boy oranı yeterli olmamakta ve çok türde ve sayıda pencerelerin görüntü üzerinde gezdirilmesi ve pencerelerin kaydırma oranlarının çok dikkatli belirlenmesi gerekmektedir. Oldukça fazla işlem zamanına mal olan bu yaklaşım derin öğrenmenin potansiyelinden yeteri kadar faydalanmamaktadır ve pratiklikten uzaktır.

Yukarıda ifade edilen sorunun çözümüne dönük olarak CNN-tabanlı iki farklı türde hedef tespit yaklaşımı günümüzde yaygın olarak kullanılmaktadır. Bunlardan biri, iki aşamalı bir süreç içinde önce görüntüler içerisindeki nesne olma potansiyeli yüksek bölgelerin önerilip sonrasında sınıflandırma ve sınırlayıcı kutular kümesinin üretildiği bölge-tabanlı CNN yaklaşımıdır. Faster R-CNN (Ren, He, Girshick, & Sun, 2016) veya Mask R-CNN (He, G., Dollár, & Girshick, 2018) gibi modeller bu türün örnekleri arasındadır. İkinci yaklaşım ise nesne tespitini bir regresyon problemi olarak ele alan ve

sınırlayıcı kutuların koordinatlarını ve sınıf olasılıklarını doğrudan görüntü özelliklerinden tahmin eden YOLO (You Only Look Once) (J., Divvala, Girshick, & A., 2015) ve SSD (Single Shot MultiBox Detector) (Liu, ve diğerleri, 2016) gibi modellerin dâhil olduğu tek-kademeli, bölge öneri ağına ihtiyaç duymayan yöntemlerdir.

Bölge önerisine dayalı konvolüsyonel sinir ağlarının ilk örneği olan R-CNN (Girshick, Donahue, Darrell, & Malik, 2014) yönteminde, görüntüde yer alan olası nesnelere için aday kutuların üretilmesi öğrenmeyi içermeyen yollarla, örneğin Seçmeli Arama (Selective Search) (Uijlings, Sande, Gevers, & Smeulders, 2013) veya EdgeBox (Dollár & Zitnick, 2014) gibi araçlar kullanılarak gerçekleştirilmektedir. SPP-net (He K. Z., 2015), Fast R-CNN (Girshick, Fast R-CNN, 2015), Faster R-CNN ve Mask R-CNN gibi sonraki çalışmalarda aday bölgelerin belirlenmesi ve özellik çıkarımının tamamı konvolüsyonel katmanlar aracılığıyla gerçekleştirilmektedir. Böylelikle, R-CNN yönteminden farklı olarak, diğer bahsi geçen yöntemlerde konvolüsyon işleminin her aday bölge için tekrar edilmesi gerekliliği ortadan kalkmıştır. Fast, Faster ve Mask R-CNN yöntemlerinde, tespit sürecinin ilk evresinde önerilen nesne konumlarının düzeltilmesi için ayrıca bir regresyon öğrenmesi de uyarlanmıştır. Fast R-CNN yöntemine, bölgeleri daha yüksek doğrulukta ve hızda önermek üzere bölge öneri ağının (Region Proposal Network - RPN) eklenmesiyle Faster R-CNN yöntemi geliştirilmiş ve kendinden önceki tüm modellerden daha hızlı ve doğru sonuçlar elde edilmiştir. Mask R-CNN ise Faster R-CNN mimarisi üzerine inşa edilmiş bir yapı olup görüntüler üzerindeki nesnelere sınırlayıcı kutular yerine maskelerle işaretlemek üzere tasarlanmıştır. Bu yöntemde ayrıca ilgi bölgesi hizalaması (ROI Align) kullanılarak nesnelere konumlarının kendinden önceki yöntemlere göre daha doğru bir şekilde belirlenmesini sağlamıştır.

Öte yandan, YOLO varyantları ve SSD gibi tek-kademeli yöntemlerde ise görüntü içerisinde yer alan nesnelere kategorileri ve pozisyonları bir ağ tarafından tek aşamada üretilmektedir. YOLO modeli, ortaya konduğu dönemde diğer tüm CNN tabanlı nesne tespit modellerinden daha hızlı olmasıyla ön plana çıkmıştır. SSD ise YOLO'dan farklı olarak, tam bağı katmanlar kullanmak yerine özellik haritalarına küçük konvolüsyon filtreleri uygulamaktadır. Gerçek-zamanlı tespit kabiliyetiyle hem optik hem SAR görüntülerinde gemi tespiti alanında kullanılagelen bir yaklaşım olan tek aşamalı detektörlerin performansı küçük nesnelere söz konusu olduğunda düşmektedir (Nie, Zhang, Niu, Dou, & Xia, 2017), (Ma, Chen, Liu, & Yang, 2018). Girdi olarak belirli

boyutlarda görüntüleri işleyebilen bu yaklaşımda büyük görüntülerde tespit yapabilmek için yine kayan pencere gibi görüntü kesme-kaydırma işlemine veya büyük görüntünün 300x300 piksel gibi boyutlara ölçeklendirilmesine gereksinim duyulmaktadır. Bu görüntü kabalaştırma işlemi ise küçük veya birbirine yakın nesnelere gözden kaçmalara sebebiyet vermektedir.

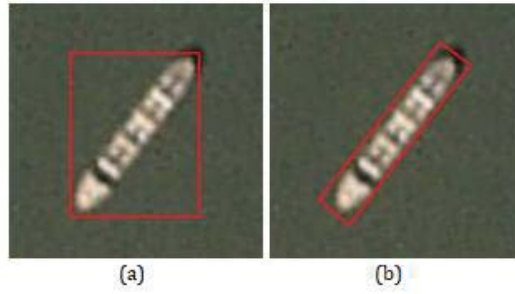
CNN-tabanlı nesne detektörlerinin omurgası VGG-16, ResNet-50/101 gibi derin mimarili ağlardır. Görevi girdi görüntüsüne dair özellikleri çıkartmak olan bu omurga ağlarının eğitilebilmesi için milyonlarca görüntüye ihtiyaç duyulmaktadır. Doğal görüntülerde nesne tespiti söz konusu olduğunda içerisinde yine aynı türde, belirli sınıflardaki nesnelere ait yeteri sayıda görüntüyü barındıran Pascal, COCO (Common Objects in Context) gibi veri setleriyle eğitilen bu ağların, bu setlerde tanımlı olmayan türlerdeki nesnelere tespitinde kullanılabilmesi için öğrenim aktarımı (transfer learning) yöntemi kullanılmaktadır. Öğrenim aktarımı, bir görev için geliştirilen bir modelin farklı bir görevdeki bir model için başlangıç noktası olarak yeniden kullanıldığı bir makine öğrenme yöntemidir. Gemiler söz konusu olduğunda özellikle optik banttaki uydu ve hava görüntüleri doğal görüntülerle ortak görsel özellikleri paylaşmakta olduğundan öğrenim aktarımı yöntemi literatürdeki çeşitli tespit çalışmalarında kullanılan bir araç olmuştur (Nie, Zhang, Niu, Dou, & Xia, 2017), (Tayara & Chong, 2018).

Ne var ki, özellikle görüntüleme açılarındaki ve sensör-nesne arası uzaklıklardaki farklılıklar dolayısıyla gemilerin görünüşleri doğal ve uzaktan algılama görüntülerinde birbirinden farklılık arz etmektedir (Şekil 1.2.a ve 1.2.b). Öncelikle, uydu ve hava görüntüleri doğal olanlara kıyasla farklı görüntüleme açılarından sahneler içermekte ve çok daha geniş bir alanı kapsamaktadırlar. Gemi boyutları tipik görüntülerdeki emsallerine göre oldukça küçük kalmakta ayrıca kimi örneklerde çeşitli (bulutluluk veya köprü gibi yüksek yapıların altında kalma gibi) sebepler dolayısıyla gemiler kısmen görünür durumda veya gölgede kalabilmektedir. Bu ve benzeri sebeplerle uydu görüntülerinde gemi tespiti doğal görüntülerdekine göre farklı zorluklar arz etmektedir.



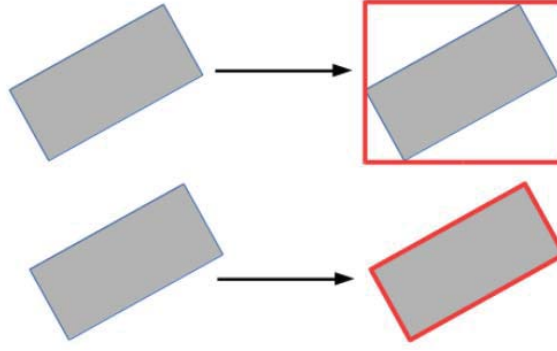
Şekil 1.2. Gemi büyüklüklerinin ve görüntüleme açılarının a) doğal fotoğrafta ve b) uydu görüntüsünde karşılaştırmaları.

Küçük hedefler için daha uygun olmaları dolayısıyla uydu ve hava görüntülerinde nesne tespiti konusunda önerilmiş yöntemlerin birçoğu bölge-temelli CNN'leri kullanmaktadır (Yang X. , ve diğerleri, 2018), (Liu, Kuang, & Li, 2017), (Gao, He, Sun, Jia, & Zhang, 2019), (Tayara & Chong, 2018). Bu yolla gerçekleştirilen güncel çalışmaların çoğunda ise araştırmacılar Faster R-CNN yöntemini tercih etmişlerdir. Bu yöntemde nesne çıkarımları ve etiketlemeleri sınırlayıcı kutular aracılığıyla gerçekleştirilmektedir. İşaretlemeler nesnenin alt ve üst uç sınırlarını, sağ ve sol uç sınırlarını ancak içine alan dikdörtgen kutular vasıtasıyla yapılır (Şekil 1.3.a). Bu yolla, tespiti istenen nesneyle beraber nesneye ait olmayan arka plan bölgeleri de ister istemez işaretli alan içerisinde kalmaktadır. Nesnelere dair çıkarılan özellikleri etkileyen bu sorunu hafifletmek adına önerilen yeni bir yaklaşımda (Yang X. , ve diğerleri, 2018), (Gao, He, Sun, Jia, & Zhang, 2019) görüntülerdeki nesnelere, dönük sınırlayıcı kutularla işaretlenmeye başlanmıştır (Şekil 1.3.b).



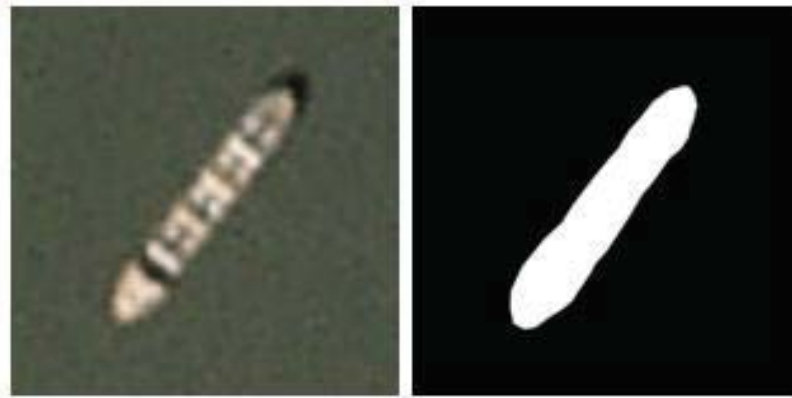
Şekil 1.3. Uydu görüntüsünde yer alan bir geminin a) dik sınırlayıcı kutu, b) dönük sınırlayıcı kutularla işaretlenmesi

Şekil 1.4'te temsil edildiği gibi, işaretlenen nesnenin dikdörtgen olması halinde nesneyi tam olarak çevreleme/işaretleme kabiliyeti olan bu yaklaşımın Şekil 1.3.b'deki gibi eğri yüzeylere sahip cisimlerdeyse nesne harici belirli miktarda arka plan içermesi, bunun da modelin nesneye dair edineceği bilgiyi belli oranda yanlış etkilemesi kaçınılmazdır.



Şekil 1.4. Dönük sınırlayıcı kutular dik emsallerine göre nesneyi daha iyi temsil etme becerisindedir.

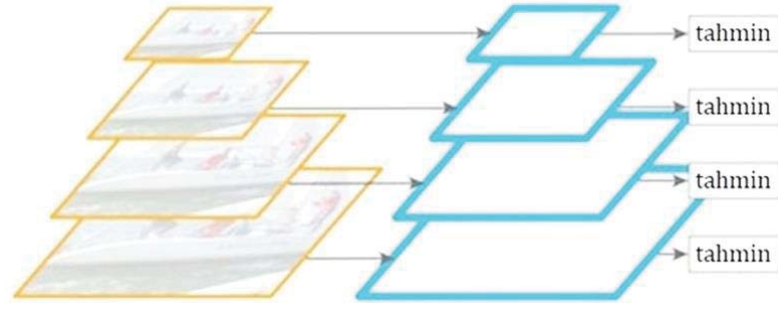
Bir nesneye dair uzaysal yapının gerçeğe sadık şekilde temsil edilebilmesinin ve tespit sistemi tarafından nesne modelinin tam doğru şekilde öğrenilebilmesinin bir yolu maske kullanımudur. Maske, kullanıcı tarafından objenin görünür dış hatlarını temsil etmek üzere oluşturulan bir poligondur (Şekil 1.5). Boolean türündeki bu maskelemeye nesneye ait pikseller, binary (çiftli) temsilde 1 değerini alırken, nesne harici bölgeler 0 değerini almakta ve siyahla temsil edilmektedirler. Bu yöntem, dönük olsun olmasın sınırlayıcı kutulardan farklı olarak, nesnenin uzaysal yapısını gerçekte görüldüğü haliyle temsil etme becerisindedir. Dolayısıyla, etiketlemesi bu şekilde yapılmış nesnelere özellik çıkarımı sınırlayıcı kutu kullanan yöntemlere göre daha başarılıdır. Tespiti bu yolla yapan Mask R-CNN, ilgi bölgesi hizalama katmanının da katkısıyla orijinal çalışmada da (He, G., Dollar, & Girshick, 2018) ifade edildiği üzere nesne tespit performansını Faster R-CNN yönteminden yukarı taşımıştır.



Şekil 1.5. Uydu görüntüsündeki bir geminin ikili (binary) maskelemesi

Nesne tespit modellerinin işlevselliğine dair bir diğer önemli konu detektörün görüntü çözünürlüğündeki/ölçeğindeki değişimlere karşı gösterdiği başarıdır. Derin öğrenme temelli yaklaşımlarda eğitim görüntüleri çeşitli şekillerde ele alınmaktadır.

Çoğunlukla, eğitim verilerinin tamamı aynı ölçekteki numunelerden hazırlanmakta, sonrasında nesne çıkarımı da aynı ölçek üzerinden gerçekleştirilmektedir. Tespitin eğitimden farklı bir ölçek üzerinde gerçekleştirilmesi halinde ise doğruluk oranları bundan olumsuz etkilenmektedir (Nie, Zhang, Niu, Dou, & Xia, 2017). Ancak, insansız hava araçları, uçaklar ve helikopterler gibi irtifası değişken olan görüntüleme platformları dikkate alındığında bunlar aracılığıyla görüntülenen herhangi bir nesnenin kayıt içerisindeki boyutu platformun nesneye olan uzaklığına bağlı olarak değişecektir ve görüntüde belirli tek bir ölçekten bahsetmek mümkün olmayacaktır. Dolayısıyla bu koşullarda herhangi bir hedef nesneye dair yüksek doğrulukta tespit istendiğinde nesneye dair geniş bir görüntüleme ölçeği aralığında bilgiye ihtiyaç olacaktır. Benzer şekilde söz konusu, dünya yörüngesinde dolanan yer gözetleme uyduları olduğunda da belirli bir uydudan alınmış görüntülerle eğitilen bir modelin, ayrıca bir eğitim sürecine gidilmeden çözünürlüğü farklı olan başka bir uydu verisinde de kullanılması arzu edilebilir. Bu türdeki talepleri karşılama adına aynı görüntünün farklı ölçeklere örneklenmiş hallerinden oluşturulan görüntü piramitleri (Şekil 1.6), modellerin eğitiminde kullanılmış ve performansta iyileştirmeler sağlanmıştır (Dollar, Appel, Belonie, & Perona, 2014). Ne var ki, bu tür bir yaklaşım kullanıldığında özellikle küçük nesnelere ayırt edilmesi söz konusuysa piramitin üst katmanlarında görüntünün kabalaşması/çözünürlüğün düşmesiyle nesnelere dair detaylar giderek kaybolmakta ve akabinde bu görüntülerin ağa beslenmesiyle beraber art arda konvolüsyon+havuzlama işlemlerine maruz kalan nesneye dair mekânsal bilgi tamamen kaybolma noktasına gelmektedir. Bu problemin giderilmesi adına Bilgisayarlı Görme (Computer Vision) alanında geliştirilmiş olan Özellik Piramit Ağı (Feature Pyramid Network-FPN) (Lin, ve diğerleri, 2017) önemli bir araç olmuş ve doğal görüntülerde yer alan nesnelere söz konusu olduğunda performansta önemli artış sağlamıştır. FPN, ağ içerisindeki yüksek çözünürlüklü ancak sığ semantik değere sahip katman özelliklerini semantik değeri yüksek derin katman özellikleriyle birleştirmek üzere geliştirilmiş yanal bağlantılara sahip, bölge-temelli CNN sistemlerine uyumlu, bir özellik çıkarıcı olarak çeşitli doğal görüntü uygulamalarında kendini kanıtlamış bir mimaridir. Eğitim verisinden farklı çözünürlükteki girdilerde tespitte ihtiyaç duyulduğunda veya küçük objeler üzerindeki doğruluk oranlarının yüksek olması istendiğinde görüntü piramitlerinin hesaplanmasına gerek bırakmayan ve daha doğru sonuçlar üretmeye yardımcı olan pratik bir çözümdür.



Şekil 1.6. Bir görüntü piramidinin en alt katmanında orijinal çözünürlükteki görüntü yer alırken üst katmanlar bunun kabalaştırılmış hallerinden oluşmaktadır. Tahminler her ayrı düzey için yapılır.

Büyük-ölçekli, yüksek çözünürlüklü uzaktan algılama görüntülerindeki gemilerin tespitinde yaşanan sorunlardan bir başkası, özellikle karasal bölgeleri içeren sahnelerde homojen olmayan karasal kısımdan gemilere benzer doku ve parlaklık dağılımlarına sahip yapıların gemi sanılarak yanlış alarmlara yol açmasıdır. Mevcut gemi tespit yöntemlerinin birçoğunda bu sorunu çözmek için gemi özelliklerini çıkarmadan önce genellikle deniz-kara segmentasyonuna başvurulmaktadır (Hu, Du, Jiang, & Ming, 2015), (Cheng, Meng, Xiang, & Pan, 2016). Segmentasyonu takiben, kara bölgelerinden arındırılmış görüntüden gemi özellikleri derin ağlar aracılığıyla çıkarılmakta böylece karada yer alıp gemilerle benzerlik taşıyan yapıların yanlış alarm üretmesinin önüne geçilmektedir. Bu yöntemin en büyük handikapı ise karaya yanaşmış gemilerin ve büyük gemilerin de segmentasyonda kara parçaları olarak değerlendirilip gözden kaçırılmasıdır (An, Pan, & You, 2018). Diğer bir dezavantaj ise tespitten tamamen ayrı bir işlem sürecine gerek olmasıdır.

Zorlayıcı içerikten kaynaklanan hatalı nesne tespitlerine mani olmanın diğer bir yolu eğitimde negatif örnekleme başvurmaktır. Özellik çıkarımının çok sayıda verinin kullanılmasıyla sağlandığı derin öğrenme için uygun bir yoldur. Bu yaklaşımda, hedef dışı objeler de negatif örnekler olarak ağa tanıtılarak bunların yapısal özelliklerinin öğrenilmesi sağlanmaktadır. Böylece, aynı kategorideki hedef olmayan nesnelere daha sonra karşısına geldiğinde, ağ onları arka plan olarak tanıyacak şekilde yapılandırılmış olur. Negatif örnek olarak gemi gibi görünen, gemi-harici nesnelere kullanılmaktadır. Geçmiş çalışmalarda rıhtımlar, adacıklar ve diğer gemi benzeri kıyı yapıları yanlış alarmlara mani olmak üzere negatif eğitim örnekleri olarak kullanılmışlardır (Liu, ve diğerleri, 2013), (He, Lin, Chen, Tai, & Yin, 2017) (Şekil 1.7). Bu yöntemde hassas negatif örneklerin özel olarak seçilmesine ve etiketlenmesine gerek duyulmaktadır.



Şekil 1.7. *Gemiyle benzerlik taşıyan ve model için kafa karıştırıcı nitelikte olan negatif örnekler*

Konuya farklı yaklaşan diğer bir çalışmada (Gao, He, Sun, Jia, & Zhang, 2019) tespit için kara bölgelerinden kaynaklı yanlış alarmları önlemek üzere ağ eğitiminde, nesne içeren görüntülere ilave olarak, hassas örnekleri özel olarak seçip eklemek yerine, yalnızca karasal alanları içeren çok sayıda rastgele negatif veri örneği eklenmiş ve deniz-kara segmentasyonuna göre başarılı sonuçlar elde edilmiştir. Aynı çalışmada ayrıca büyük ölçekteki görüntüler içerik durumlarına (farklı kara-deniz dağılımlarına) göre de sınıflandırılmış böylece yanlış alarm kaynakları ve gözden kaçan gemiler daha nicel bir şekilde değerlendirilebilmiştir.

Literatürdeki yeni bir yaklaşımda, eğitim verisine herhangi bir ekleme yapmaya gidilmeden verideki arkaplan içeriğinin hedef nesnelere olan benzeşiminden kaynaklanan yanlış alarmların üretimine mani olmak ve böylece tespit doğruluğunu artırmak üzere odak kayıp kavramının tanımlandığı RetinaNet (Lin, Goyal, Girshick, He, & Dollár, 2018) yöntemi önerilmiştir. Mask R-CNN, Faster R-CNN gibi tespit yöntemlerinin iç işleyişinde görüntüdeki obje sayısından birkaç kat fazla sayıda nesne önerisi üretilmektedir. Bir kısmı gerçek nesnelere tekabül eden bu önerilerin kalan kısmı ise arka plan nesnelere aittir. Eğitimde kayıp fonksiyonunun hesaplanmasında gerçek nesnelere kaynaklı kayıplar yanında bu nesne-harici bölgelerden gelen kayıplar da kullanılmaktadır. Bu yöntemde tanımlanmış odak kayıp kavramıyla modelin gerçek nesnelere gelen kayıplara ağırlık vermesi sağlanarak daha başarılı özellik çıkarımı elde edilebilmektedir. Bu yöntem işlevi ve ürettiği sonuçlar itibarıyla negatif eğitim örneği yaklaşımına bir alternatif olmuştur.

Görüntülerden nesne çıkarımını derin öğrenme yöntemiyle ele alan çalışmaların birçoğunda elde edilen eğitim verisinden daha iyi faydalanabilmek için verinin sentetik olarak artırımına (data augmentation) başvurulmaktadır. Veri artırımı, uygulayıcıların yeni veri toplamadan eğitim modelleri için mevcut verinin çeşitliliğini önemli ölçüde artırmasını sağlayan bir stratejidir. Görüntü döndürme, yaklaştırma/uzaklaştırma, kırpma ve parlaklık/kontrast değiştirme gibi veri artırma teknikleri, derin sinir ağlarını eğitmek için

yaygın olarak kullanılmaktadır. Nesne tespitini derin öğrenmeyle gerçekleştiren çalışmalarda da (Tayara & Chong, 2018) bu teknikten faydalanılmış olup, CNN temelli bir detektörün farklı veri artırma yöntemleri altındaki performansı karşılaştırılmıştır.

1.2. Araştırma Kapsamı

Bu tezde, uydu görüntülerinde yer alan gemilerin tespiti için,

- Nesne işaretlemesinde sınırlayıcı kutular yerine maske kullanımının,
- Görüntülerdeki ölçek değişikliklerine karşı tespit doğruluğunu korumak üzere özellik piramit ağı kullanımının,
- Yanlış tespitleri azaltmak üzere; modelin negatif örnekleme eğitilmesinin veya modelde odak kayıp fonksiyonunun kullanımının,
- Çeşitli eğitim verisi artırma tekniklerinin ve farklı omurga ağlarının

tespit sonuçları üzerindeki etkileri incelenmiştir.

Araştırma süreci, oluşturulan veri setleri üzerinde farklı tespit modellerinin, çeşitli omurga ağlarının ve veri artırma tekniklerinin kullanılarak kapsamlı deneylerin gerçekleştirilmesine ve bulguların analizine dayanmaktadır. Önerilen yöntemin performansı çeşitli metrikler kullanılarak incelenmiş ve bu alanda sıkça tercih edilen Faster R-CNN gibi bir algılama modeliyle mukayesesi yapılmıştır.

1.3. Özgün Değer

Çalışma, uzaktan algılama görüntülerinde gemi tespiti çalışmalarına aşağıdaki katkıları sunma iddiasındadır:

- Önceki algılama modellerinden daha iyi performansa sahip, zorlayıcı içeriklerde yer alan gemileri ayırma/algılama kapasitesi yüksek ve tespit nesneyle alakasız kısımları içermeyen sonuçlar üretebilen, Mask R-CNN yöntemine dayalı yeni bir gemi tespit çerçevesi oluşturulması,
- Özellik piramit ağının algılama modeliyle beraber kullanılmasıyla tespit performansının sadece tek ölçekteki görüntüler ve nesnelere üzerinde değil çoklu ölçekte (farklı çözünürlüklü görüntülerde) yüksek kalmasının sağlanması,
- Deniz-kara segmentasyonuna başvurmadan modelde negatif örnekleme eğitimi veya RetinaNet/odak kayıp yaklaşımının kullanılmasıyla karasal bölgelerden kaynaklı yanlış alarmların azaltılması,

- Test verilerinin, (1) içerikleri açısından farklı düzeylerde tespit zorluklarına yol açan üç kategoriye ayrılması ve (2) içeriklerdeki nesnelerin büyüklüklerine göre tanzim edilip değerlendirilmesiyle hataların daha nicel analizini gerçekleştirmenin mümkün hale gelmesi,
- Çeşitli veri artırım yöntemlerinin ve farklı omurga ağlarının tespit performansına etkisinin sistematik olarak değerlendirilmesi.

1.4. Tez Yapısı

Bu tezin kalan kısmı şu şekilde düzenlenmiştir: İkinci bölümde, tespit modelinin teorisi, literatürdeki en yaygın olarak kullanılan ve bölge-tabanlı bir konvolüsyonel nesne tespit modeli olan Faster R-CNN modeli üzerinden sunulmaktadır. Önerilen yöntemin ayrıntıları üçüncü bölümde yer almaktadır. Dördüncü bölüm, çalışmada kullanılan uzaktan algılama veri kümesine ve model eğitime dair detayları, beşinci bölüm ise önerilen çerçevenin etkinliğini doğrulamak üzere yapılan deneyleri ve bulguları sunmaktadır. Yöntemin ürettiği sonuçlar altıncı bölümde tartışılmakta ve son bölümde çalışmanın değerlendirmesi yapılmaktadır.

2. TEORİ

Bu çalışma kapsamında önerilen gemi tespit modeli, bölge tabanlı konvolüsyonel bir ağ olan Mask R-CNN'dir. Bu model, bölge tabanlı bir nesne tespit modeli olan Faster R-CNN üzerine geliştirilmiştir. Faster R-CNN modeli, detayları bu kısımda ifade edilecek olan çeşitli aşamalardan oluşmaktadır. Model, girdi görüntüsünün omurga ağı girilmesiyle başlayan işlem sürecinde çıktı olarak görüntü içerisindeki hedef nesnelere ve konumlarını sınırlayıcı kutular aracılığıyla vermektedir. Bu model üzerindeki bir takım geliştirmelerle gemi tespitinde kullanımını önerdiğimiz Mask R-CNN modeli oluşturulmaktadır.

2.1. Bölge Tabanlı R-CNN Yöntemleriyle Nesne Tespiti: Faster R-CNN

Bu bölümde, bu tez kapsamında gemi tespitinde kullanımı önerilen Mask R-CNN yönteminin üzerine inşa edilmiş olduğu ve uzaktan algılamada nesne ve gemi tespit literatüründe sıklıkla kullanılan bir yöntem olan Faster R-CNN modeline dair detaylar verilecektir. Burada modelin işleyişi ifade edilecek olup önerilen yöntemin katkılarının hangi zemin üzerine yapıldığını aktarmak amaçlanmaktadır.

Çeşitli bileşenlerden oluşan Faster R-CNN ile girdi görüntüsünden başlayan süreçte şu çıktılar üretilmektedir:

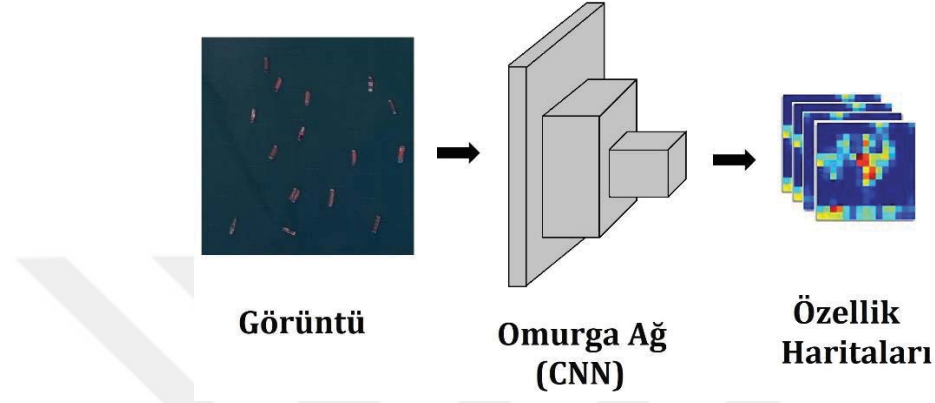
- Nesnelere içeren sınırlayıcı kutular listesi
- Her sınırlayıcı kutunun etiketlenilmesi (ilgili sınıfa atanması)
- Her sınırlayıcı kutu için olasılık değerleri

Süreç, ön-eğitilmiş bir CNN'den geçirilen $Yükseklik \times Genişlik \times Derinlik$ tensörü formatındaki girdi görüntülerinden konvolüsyonel özellik haritalarının elde edilmesiyle başlamaktadır. Bir sonraki kısım için bu, özellik çıkarıcı işlevindedir.

2.1.1. Özellik çıkarımı

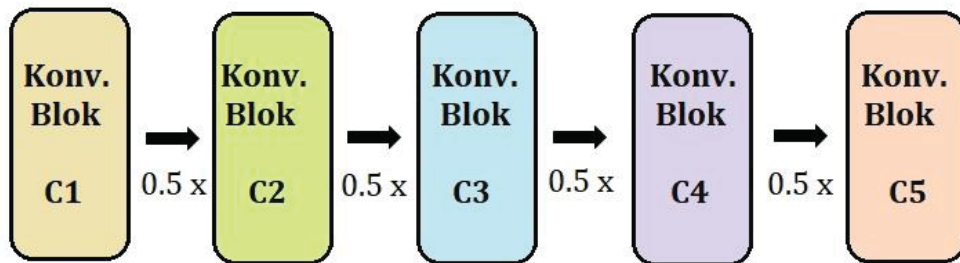
Özellik çıkarımı elde edilen veriden içeriğe dair bilginin derinleştirilip (daha kullanışlı veri gruplarına evriltilip) mekansal boyutlarının indirildiği bir süreçtir. Mask R-CNN'de özellik çıkarımı konvolüsyonel sinir ağları aracılığıyla sağlanmaktadır. Bu ağlar çok katmandan meydana gelmekte ve her konvolüsyonel katman kendinden önceki bilgilere dayanarak soyutlamalar oluşturmakta, özellik haritaları üretmektedir. İlk katmanlar genellikle kenar, doku gibi özellikleri öğrenmekteyken sonrakiler daha

karmaşık şekilleri, ilişkileri bulmak üzere özelleşmektedir. Sonunda, orijinal görüntüden çok daha küçük boyutlara sahip, ancak daha derin yapıda, konvolüsyonel bir özellik haritası elde edilmiş olunur. Özellik haritasının genişliği ve yüksekliği, konvolüsyonel katmanlar arasında uygulanan havuzlama nedeniyle azalmakta ve öğrendiği filtre sayısına bağlı olarak derinliği artmaktadır (Şekil 2.1).



Şekil 2.1. Bir görüntü için özellik çıkarımı omurga ağ aracılığıyla gerçekleştirilir; derinlik (semantik değer) artarken görüntü boyutu küçülür.

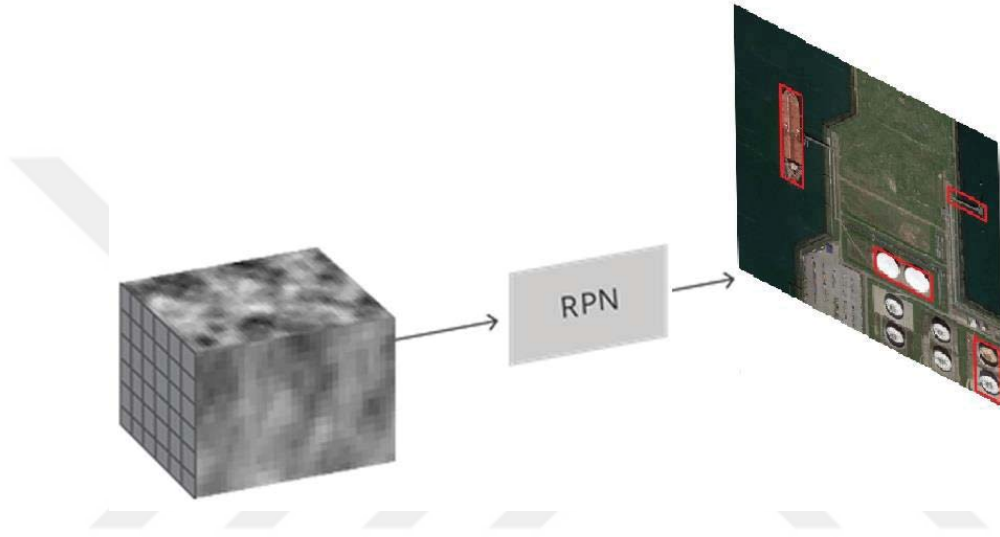
Bu çalışmada özellik çıkarıcı olarak iki farklı ağ kullanılmıştır: ResNet-50 ve ResNet-101. 50 katmandan oluşan ResNet-50 toplamda 23.5 milyon parametreye sahipken 101 katmanlı ResNet-101’de toplam parametre sayısı 42.5 milyondur. ResNet, iç işleyişinde artık bağlantı (residual connection), kestirme bağlantı (shortcut connection) ve yığın normalizasyonun (batch normalization) kullanıldığı, derin olmasına rağmen az veriyle de eğitimin sağlanabildiği bir mimaridir. Ağ içerisindeki katmanlar 5 blok olarak değerlendirilmektedir (Şekil 2.2). Her bir bloktan çıkan özellik haritasının mekansal boyutları bir önceki blok çıktısına göre yarıya inerken, derinliği (semantik değeri) artmaktadır.



Şekil 2.2. 5 bloktan oluşan ResNet ağ yapısında her bir blok arasında özellik haritalarının mekansal boyutları yarıya düşmekte, derinliği artmaktadır.

2.1.2. Bölge öneri ağı

Süreçte bir sonraki aşama, özellik haritalarının Bölge Öneri Ağına (RPN) girilmesidir. CNN tarafından çıkarılan özellikleri kullanarak, nesne içermesi muhtemel bölgeleri özellik haritasından çıkarmak üzere kullanılmaktadır (Şekil 2.3). Haritalarda çapalar oluşturmayla başlayıp paralel yürütülen iki konvolüsyon işlemiyle görüntüde yer alan olası nesnelere dair sınırlayıcı kutular çıkarılmakta ve bu kutuları görüntüdeki gerçek nesnelere oturtmak üzere düzeltme işlemleri uygulanmaktadır.

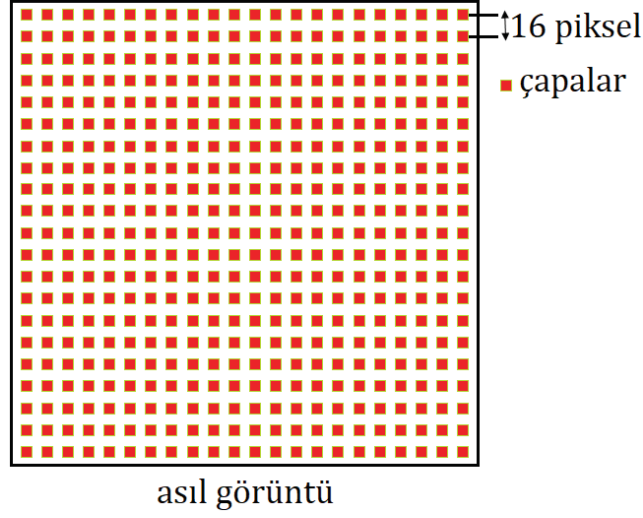


Şekil 2.3. Bölge Öneri Ağı'yla görüntü içerisindeki nesne olması muhtemel bölgeler belirlenmektedir.

2.1.2.1. Çapalar

Bölge tabanlı CNN yöntemlerinde görüntüdeki nesne içerebilecek sınırlayıcı kutuların bulunması çapa adı verilen yapılar aracılığıyla gerçekleştirilmektedir. Çapalar görüntü içerisinde ağ tarafından eşit aralıklarla yerleştirilmiş dikdörtgen şeklindeki sınırlama kutularıdır.

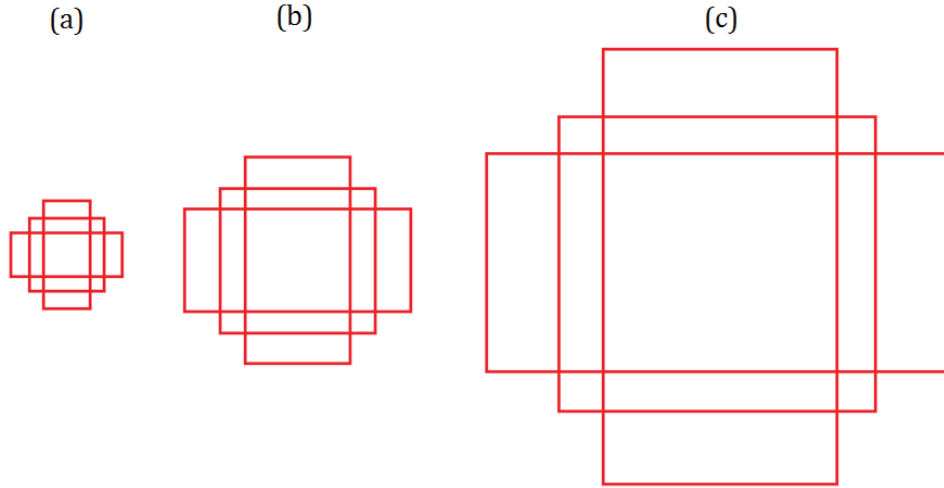
$konv_{genişlik} \times konv_{yükseklik} \times konv_{derinlik}$ Boyutunda bir konvolüsyonel özellik haritası üzerinde çalışıldığında, haritadaki $konv_{genişlik} \times konv_{yükseklik}$ 'deki her nokta için asıl resim üzerinde (farklı boyut ve en boy oranlarına sahip) bir çapa kümesi oluşturulmaktadır (Şekil 2.4). Her ne kadar tanımlamaları konvolüsyonel özellik haritası temelinde yapılmış olsa da çapalar nihai olarak orijinal görüntüye referans olmaktadır.



Şekil 2.4. Oluşturulan çapaların görüntü üzerindeki konumu. Ağın 5. bloğundan çıkan özellik haritası kullanıldığı takdirde, çapalar orijinal görüntü üzerinde birbirinden $2^{5-1}=16$ piksel aralıkla oluşturulacaktır.

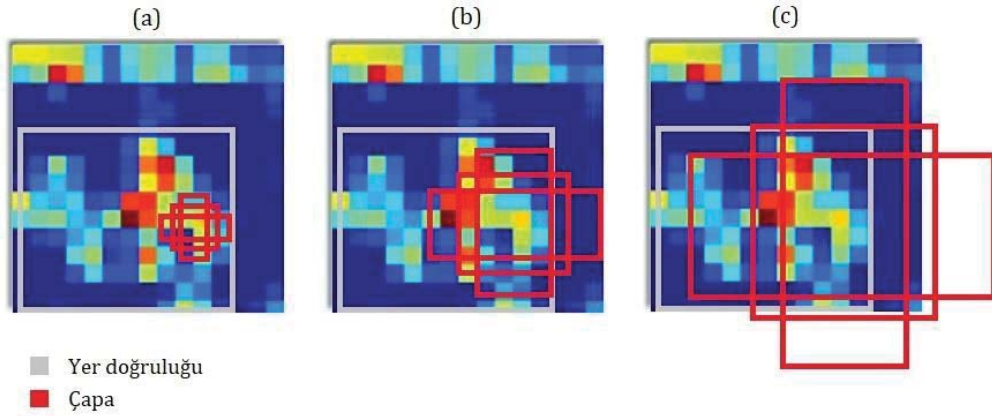
ResNet gibi sadece konvolüsyonel ve havuzlayıcı katmanların yer aldığı bir ağda üretilen özellik haritasının boyutları bloklar arasında yarıya inmektedir. Matematiksel olarak, eğer görüntü $w \times h$ boyutunda ise özellik haritasının boyutu ($r = 2^{blok\ no - 1}$: alt örnekleme oranını temsil etmek üzere) $(w/r) \times (h/r)$ olacaktır. Özellik haritasının her bir mekansal konumu başına bir çapa kümesi tanımlandığından asıl görüntüde r pikselle birbirinden ayrılmış çapalar kümesi elde edilmiş olmaktadır. ResNet'in 5. konvolüsyonel bloğundan çıkan özellik haritası kullanıldığı takdirde $r = 16$ olacaktır ve dolayısıyla asıl görüntüde 16 piksel aralıklarla çapalar oluşturulacaktır.

Faster R-CNN'nin orijinal çalışmasında (Ren, He, Girshick, & Sun, 2016) nesne önerisi oluşturma sürecinde görüntü içerisinde yer alabilecek olan farklı büyüklüklerde ve en boy oranlarına sahip nesnelere göz önüne alınarak üç farklı büyüklük skalasında ve her bir skala için üç farklı en/boy oranında çapalar üretilmektedir. En küçük skaladaki çapalar 128×128 piksel içerirken, orta skaladakiler 256×256 piksel, büyük skaladakiler ise 512×512 piksel içermektedir. Her bir skala için oluşturulan çapaların en/boy oranları $\{1:1, 1:2, 2:1\}$ şeklindedir. En/boy oranları $1:2$ veya $2:1$ olduğunda oluşturulan çapalar; uzun kenarı, kısa kenarının iki katı olmak üzere büyüklük skalasına göre toplamda 128×128 , 256×256 veya 512×512 piksel içermektedir. Bu şekilde, özellik haritasındaki her bir nokta için toplam 9 adet çapa üretilmiş olunur (Şekil 2.5).



Şekil 2.5. *Faster R-CNN'de çapalar üç farklı büyüklük skalasında oluşturulur: a) 128x128 piksel, b) 256x256 piksel, c) 512x512 piksel.*

Şekil 2.6'da bir gemi içeren özellik haritasındaki bir nokta için oluşturulmuş 9 çapa ve bunların yer doğruluğuna göre durumları görülmektedir.



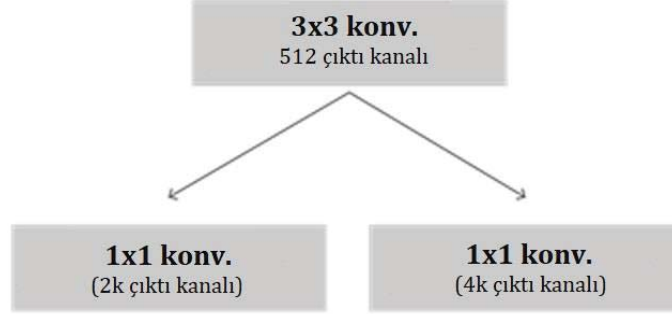
Şekil 2.6. *Özellik haritasındaki aynı konumda oluşturulan 3 farklı skaladaki ve 3 farklı en boy oranındaki çapalar.*

Çapalar aracılığıyla nesne olabilecek bölgelerin önerileri hazırlanırken, sorun iki parçaya ayrılmaktadır. Her çapa için şunlar ele alınıp, gerekli hesaplamalar yapılmaktadır:

- Bu çapa herhangi bir nesne (veya ön plan) içermekte midir?
- Bu çapayı nesneye daha doğru oturtmak için nasıl bir ayarlama gerekir?

RPN'de bu işlemler tamamen konvolüsyonel bir yolla hesaplanmaktadır. Önce, 512 kanaldan oluşan, çekirdek boyutu 3×3 olan bir konvolüsyonel filtre, özellik haritaları üzerine uygulanmakta, buradan gelen verilerse daha sonra paralel iki dalda 1×1 çekirdek

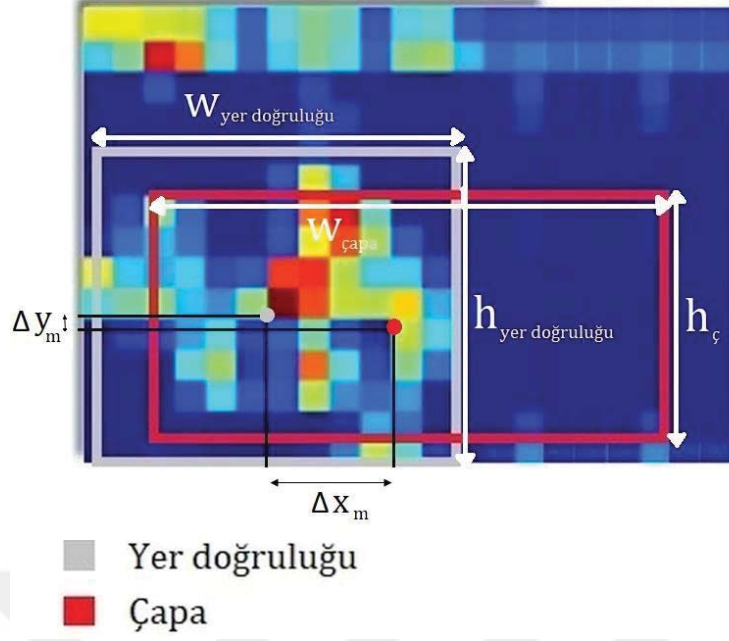
boyutundaki konvolüsyonel iki aşdan geçirilerek, biri $2k$ (k: konum başına çapa sayısı) diğeri $4k$ kanallı iki çıktı üretmektedir (Şekil 2.7).



Şekil 2.7. RPN’de, özellik haritasına uygulanan 3×3 ’lük filtrenin akabinde biri, her bir çapa için nesnellik skoru hesabını yapan, diğeri her bir çapa için yer-doğruluğuna göre gerekli düzeltmeleri gerçekleştiren iki 1×1 ’lik filtre kullanılmaktadır.

Sınıflandırma katmanında, çapa başına iki tahmin üretilmektedir: Çapanın arka plan ve ön plan (nesne) olma olasılıkları. Regresyon ya da sınırlayıcı kutu ayar katmanında ise nesne önerisinin görüntü içerisindeki konumuna dair önerileri son haline getirmek için çapalara uygulanacak olan “ Δx_{merkez} , Δy_{merkez} , $\Delta w(\text{genişlik})$, $\Delta h(\text{yükseklik})$ ” öteleme değerlerine karşılık gelen çapa başına 4 tahmin çıktısı alınmaktadır.

Herbir çapa için yer-doğruluğuna göre gerekli düzeltmeleri (ofsetleri) yapmak üzere “ x_{merkez} , y_{merkez} , w , h ” parametreleriyle tanımlanan çapa ele alınarak bunun alakalı olduğu yer-doğruluğunun sınır kutusuna göre “ Δx_{merkez} , Δy_{merkez} , Δw , Δh ” değerleri hesaplanmaktadır (Şekil 2.8). Böylece çapalar üzerindeki son düzeltmeler ve “nesnellik” skorlarının kullanılmasıyla, tespiti hedeflenen nesnelere için bir öneri seti elde edilmiş olmaktadır.

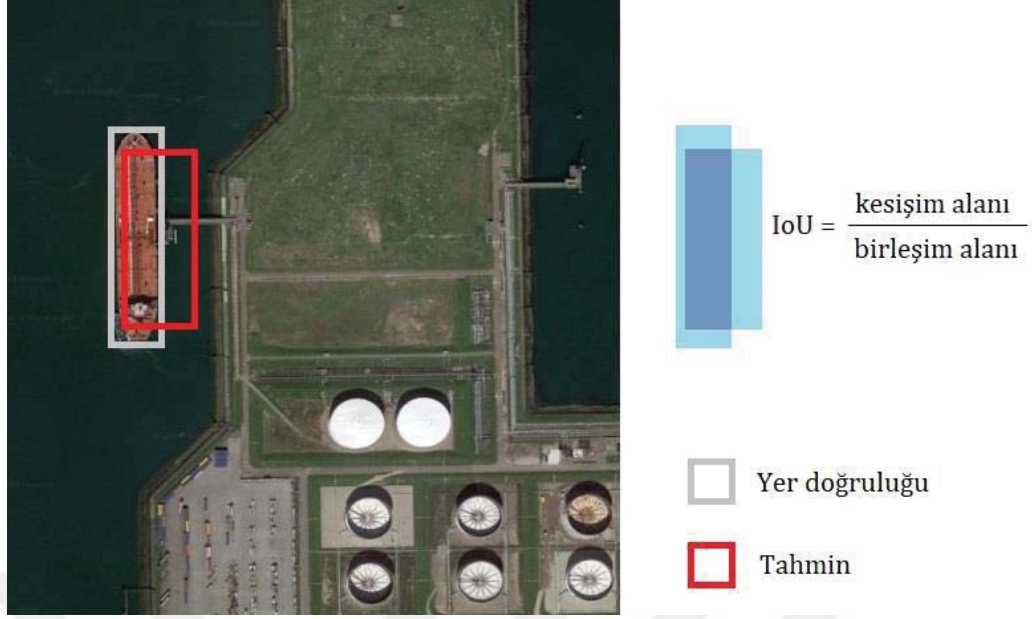


Şekil 2.8. Çapa kutusuna dair parametreler, yer-doğruluğuna dair parametrelerle arasındaki farklara bakılarak ayarlanmaktadır.

2.1.2.2.RPN’de eğitim: hedefler ve kayıp fonksiyonları

RPN’de her bir çapa için iki farklı tahmin alınmış durumdadır. Bunlar, ikili (binary) sınıflandırmalar (nesne/arka plan) ve tahmini yapılmış sınırlayıcı kutulardır.

Eğitimde tüm öneriler yer doğruluğu kutularıyla olan kesişim/birleşim oranlarına (Intersection over Union- IoU) bakılarak iki kategoriye ayrılmaktadır (Şekil 2.9). Etiketli bir nesneye, yani yer doğruluğuna göre IoU değeri 0,5’ten büyük olan öneriler (veya hiçbir çapanın bu değeri sağlamadığı koşulda bir yer doğruluğuna göre en yüksek IoU değerine sahip öneri) “ön plan” veya pozitif hedefler olarak kabul edilmektedirler. Herhangi bir etiketli nesneyle/yer doğruluğuyla çakışmayan veya çakışıp IoU değeri 0.3’ten küçük olanlar “arka plan” olarak değerlendirilmektedir.



Şekil 2.9. RPN’de üretilen nesne önerilerinin yer-doğruluklarıyla olan IoU oranına bakılarak ön plan veya arka plan olarak ataması gerçekleştirilir.

Toplamda çok sayıda (768x768 piksellik bir görüntü ve özellik haritası olarak ResNet’in 5. blok çıktısı kullanıldığı takdirde 20736 adet) olan bu örneklerden, ön plan ve arka planlar arasında dengeli bir dağılım oranı (genelde varsayılan değer olarak: 1 x ön plan, 3 x arka plan) kullanılarak görüntü başına 64’ü ön plan, 192’si arka plan olmak üzere toplamda 256 öneri kutusundan oluşan mini-yığın oluşturulmaktadır.

RPN’de sınıflandırma kaybı, mini yığını oluşturan tüm çapalar üzerinde ikili çapraz entropi (binary cross entropy) kullanımıyla hesaplanmaktadır. Bir öneri kutusu (çapa) için üretilen nesnellik skoru p ile ifade edilirse buna dair kayıp şöyle hesaplanmaktadır:

$$L_{cls}(p, u) = -\log p_u \quad (2.1)$$

Bu denklemde u , öneri kutusunun içeriğiyle alakalı olup, söz konusu kutu eğer gemi olarak görülmekteyse $p_u = p$ olarak, arka plan olarak görülmekteyse $p_u = (1 - p)$ olarak alınır. Regresyon kaybı ise mini yığında sadece ön plan olarak işaretlenmiş çapalar için hesaplanmaktadır. Regresyon hesaplamasında, bir ön plan çapası ve bağlantılı olduğu yer-doğruluğu nesnesi kullanılarak çapa çerçevesini nesneye dönüştürmek için gereken doğru ötelemeler (Δ) hesaplanmaktadır. Kaynak çalışmada (Ren, He, Girshick, & Sun, 2016), regresyon hatası için düzgün (smooth) L1 kaybının kullanımı önerilmiştir. Bu kayıp fonksiyonu, L1 hata fonksiyonunu $-\sigma$ ile tanımlanmış- yeterince küçük bir değere ulaştığında hatasız kabul etmektedir ve kaybın daha hızlı bir şekilde düşürülmesine

yardımcı olmaktadır. Yer doğruluğuna dair sınırlayıcı kutu $v = (v_x, v_y, v_w, v_h)$ ile ve tahmini yapılmış öneri kutusu $t = (t_x, t_y, t_w, t_h)$ ile tanımlanmak üzere regresyon kaybı şöyle hesaplanmaktadır (Ren, He, Girshick, & Sun, 2016):

$$L_{reg}(t, v) = \sum_{i \in \{x, y, w, h\}} smooth_{L1}(t_i - v_i) \quad (2.2)$$

Burada

$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & \text{eğer } |x| < 1 \\ |x| - 1 & \text{diğer türlü,} \end{cases} \quad (2.3)$$

RPN'deki toplam hata ise şu şekilde hesaplanmaktadır:

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (2.4)$$

Burada, i mini-yığın içindeki çapaya ait indistir, p_i ise i . çapanın nesne olmasına dair olasılık tahminidir. Yer-doğruluğu etiketi p_i^* , eğer çapa pozitifse 1, negatifse 0 değerini almaktadır. t_i Tahmini yapılmış sınırlayıcı kutuya ait 4 parametreyi, t_i^* ise pozitif çapalar için yer-doğruluğuna dair parametrik koordinat bilgisini ifade etmektedir. Normalizasyon sabitlerinden N_{cls} için mini-yığın büyüklüğünün sayısı (256) kullanılırken, N_{reg} için çapaların konum sayısı (768x768 piksellik görüntü ve ResNet 5.blok çıktısı kullanıldığında 2304) kullanılmaktadır.

Sınırlayıcı kutu regresyonunda kullanılan 4 parametre ise şöyle tanımlanmaktadır:

$$t_x = (x - x_a)/w_a, \quad t_y = (y - y_a)/h_a, \quad (2.5)$$

$$t_w = \log\left(\frac{w}{w_a}\right), \quad t_h = \log\left(\frac{h}{h_a}\right), \quad (2.6)$$

$$t_x^* = (x^* - x_a)/w_a, \quad t_y^* = (y^* - y_a)/h_a, \quad (2.7)$$

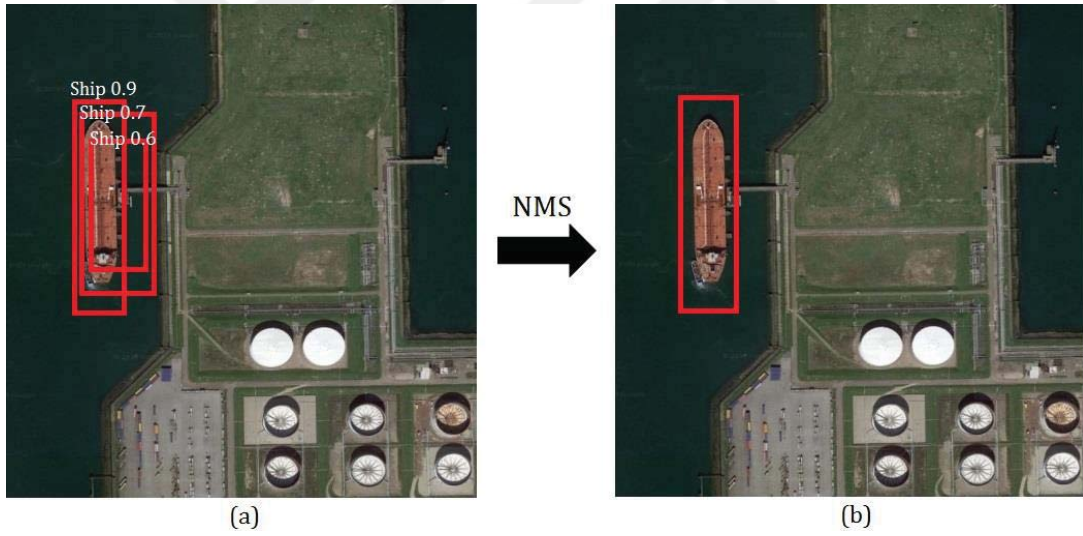
$$t_w^* = \log\left(\frac{w^*}{w_a}\right), \quad t_h^* = \log\left(\frac{h^*}{h_a}\right), \quad (2.8)$$

Burada x , y , w ve h sırasıyla kutunun merkezinin koordinatlarını ve kutunun genişliğiyle yüksekliğini temsil etmektedir. x , x_a ve x^* sırasıyla, tahmini yapılan kutuya, çapa kutusuna ve yer doğruluğu kutusuna dair değişkenlerdir (aynı durum y , w , h değişkenleri için de geçerlidir). Bu süreç bir çapa kutusundan yer-doğruluğu kutusuna dönüşüm olarak düşünülebilir.

Yukarıda tanımlanan hedeflerin ve kayıp fonksiyonlarının kullanılmasıyla RPN, geri-yayılım ve stokastik gradyan inişi (Stochastic Gradient Descent – SGD) yöntemleriyle eğitilebilir hale gelmiş olmaktadır.

2.1.2.3. Art işleme

Çapaların ve oluşturulan önerilerin birbirleriyle çakışmalarından dolayı, aynı nesne için farklı çapalardan gelmiş önerilerde üst üste binmeler söz konusu olabilmektedir. Aynı nesne için çok sayıda tespitle yüz yüze kalmama adına Maksimum Olmayanı Bastırma (Non-Maximum Supression - NMS) olarak adlandırılan basit bir algoritmik yaklaşım kullanılmaktadır. NMS, nesnellik skorlarına göre sıralanmış nesne öneri listesini (nesnellik skoru ve çapa kutusu konum tahminlerini içeren) ele alarak, birbirleriyle belirli bir oranın üzerinde çakışan öneriler için en yüksek nesnellik skoruna sahip öneri haricindekileri elemektedir (Şekil 2.10). Belirlenen bu oran, NMS eşiği olarak adlandırılmaktadır.



Şekil 2.10. Bir nesne için farklı çapalardan gelen önerilerin (a) birden fazla tespitiyle yol açmasını engellemek üzere birbirleriyle belirli bir oranın üzerinde çakışan önerilerden en yüksek nesnellik skoruna sahip öneri haricindekileri elemektedir (b).

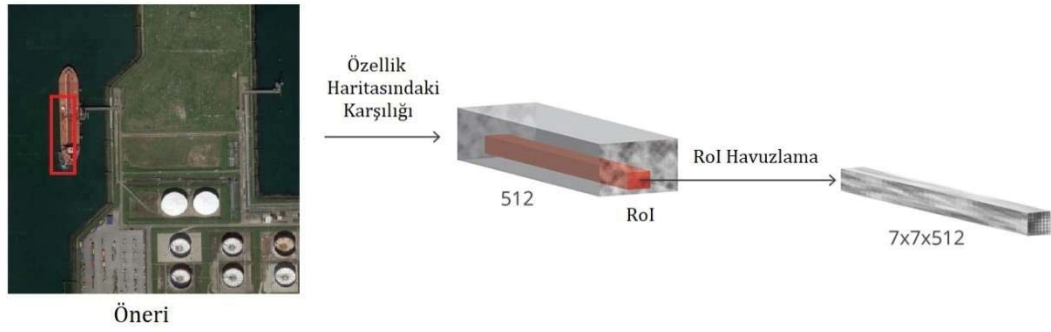
Basit görünmekle birlikte, NMS eşiğinin doğru belirlenmesi önem arz etmektedir. Düşük tutulduğu takdirde aynı nesne için çok fazla sayıda öneriyle karşılaşma durumu doğabilecekken; yüksek tutulduğu takdirde yetersiz öneri sayısına veya birbirine yakın konumlanmış nesnelere gözden kaçmalara neden olabilir. Yaygın olarak kullanılan değer 0.6'dır.

NMS'nin uygulanmasından sonra, nesnellik skorlarına göre sıralanmış en yüksek N teklif sürecin bundan sonraki aşamasında kullanılır. (Ren, He, Girshick, & Sun, 2016)'de $N = 2000$ kullanılmasına rağmen bu sayının 50'ye kadar düşürülmesi halinde bile iyi sonuç almanın mümkün olduğu belirtilmiştir.

2.1.3. İlgi bölgesi havuzlaması

RPN aracılığıyla girdi görüntüsünde ve dolayısıyla özellik haritalarında yer alan muhtemel nesnelere ve konumlar elde edilmiş bulunmaktadır. İlgi Bölgeleri (Region of Interest - RoI) olarak adlandırılan bu öneriler bu aşamada görüntüden çıkarılarak bir havuzda toplanmaktadır (RoI Havuzlama).

Bu katmanda RPN'den gelen her öneri bölgesi, konvolüsyonel özellik haritasından kırılıp çıkarılmakta ve sonrasında tüm parçalar bilineer enterpolasyon kullanılarak $14 \times 14 \times \text{konv_derinliđi}$ boyutlarına getirilmektedir. Bundan sonra, her teklif için $7 \times 7 \times \text{konv_derinliđi}$ boyutunda nihai bir özellik haritası elde etmek üzere 2×2 'lik bir çekirdekle en büyük değer havuzlaması (max-pooling) uygulanmaktadır. Bu şekilde, tekliflerin tümü belirli bir kalıba uydurulmakta böylece çıktılar, tam-bađlantılı katmanların kullanıldıđı bir sonraki modülde kullanıma hazır hale gelmektedirler (Şekil 2.11).



Şekil 2.11. *Faster R-CNN'nin RPN kısmında üretilen farklı boyutlardaki nesne önerileri havuzlama katmanında sabit boyuta getirilir.*

2.1.4. Bölge-temelli konvolüsyonel sinir ađı

Bölge temelli konvolüsyonel sinir ađı (R-CNN), Faster R-CNN ve Mask R-CNN modellerindeki son işlem adımıdır. Görüntüden konvolüsyonel özellik haritasının oluşturulmasıyla başlanıp, bunun RPN ile nesne önerileri üretmek üzere kullanımı sonrasında bu önerilerin her biri için RoI havuzlaması (Mask R-CNN'de RoI hizalaması) vasıtasıyla özellikler çıkarılmış durumdadır. Sabit bir formata boyutlandırılmış bu bölge

önerilerinin sayısı da belirli olduğundan tam-bağlantılı bir yapıda olan son modülde işlenmeye hazır haldedirler.

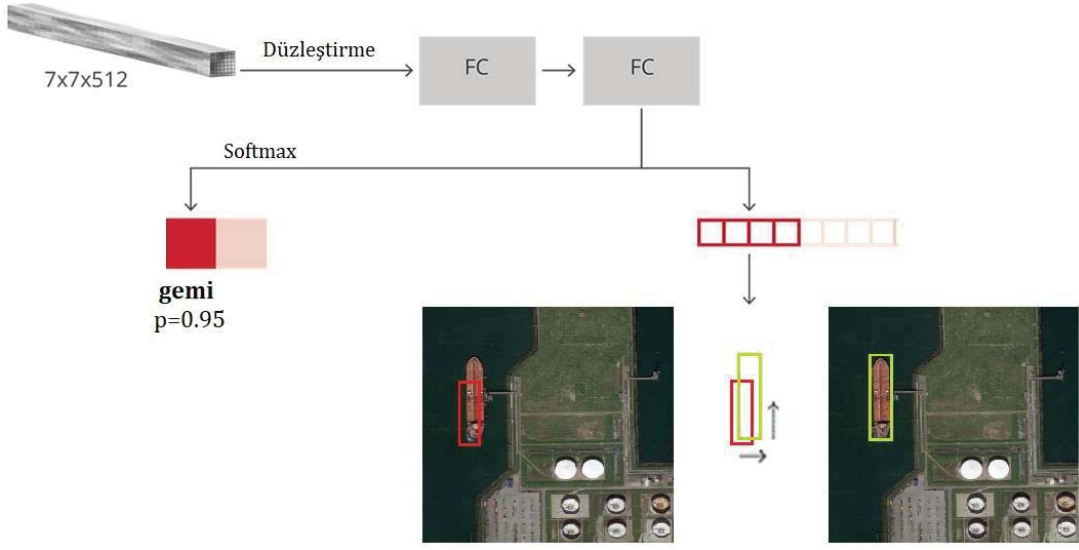
R-CNN iki farklı amaca sahiptir:

- Teklifleri tanımlanmış sınıflardan birine atamak (Not: Kötü teklifleri elemek üzere bu sınıflardan biri 'arka plan'dır)
- Tahmini yapılmış sınıfa göre, önerilmiş sınırlayıcı kutuyu yer doğruluğuna göre daha doğru ayarlamak

Orijinal çalışmada R-CNN, kendine gelen ve kalıbı belli olan her öneri için özellik haritalarını alarak bunları düzleştirmekte (flatten) ve ReLU aktivasyon fonksiyonlu 4096 düğümden oluşan tam-bağlantılı iki katmanda (sınıflandırıcı ve regresör) kullanmaktadır: Bu katmanlar ve işlevleri şöyledir:

- N toplam sınıf sayısını ve haricen bir diğeri 'arka plan' sınıfını göstermek üzere, $N + 1$ düğümlü bir tam-bağlantılı katman.
- Tanımlı N sınıftan herbiri için kutu regresyonu tahmininde hesaplanması gereken Δx_{merkez} , Δy_{merkez} , $\Delta genişlik$, $\Delta yükseklik$ 'e karşılık gelen, $4N$ düğümden oluşan bir tam-bağlantılı katman.

Böylelikle, RPN'den gelmiş olan ilgili öneri kutusu içerisindeki nesnenin türü belirlenmiş ve buna dair sınırlayıcı kutu üzerinde düzeltmeler yapılmış olacaktır (Şekil 2.12).



Şekil 2.12. *Faster R-CNN'nin son modülünde nesne sınıfının belirlenmesi ve sınırlayıcı kutunun nesneye göre daha iyi ayarlanması gerçekleştirilir.*

2.1.4.1. R-CNN modülünde hedefler ve kayıp fonksiyonu

R-CNN hedefleri, RPN hedeflerine benzer yolla hesaplanmaktadır. RPN'de önerilerin nesne olup olmadığının değerlendirilmesi yapılırken R-CNN'de gelen pozitif örneklerin nesne sınıfları belirlenmekte ve sınırlayıcı kutular sınıf kategorileri de göz önünde bulundurularak bir kez daha düzeltmeye tabi tutulmaktadır.

Bu modülde herhangi bir yer doğruluğu kutusuna göre IoU değeri 0,5'den büyük olan ve NMS sonrası filtrelenmiş öneriler o yer doğruluğu için nesne olarak atanmaktadır. IoU değeri 0.3 ila 0.5 arasında olanlarsa arka plan olarak etiketlenmektedir. RPN için hedefleri oluştururken yapılanın aksine, yer doğruluğuyla kesişimi olmayan öneriler dikkate alınmamaktadır. Bunun nedeni, RPN sonrası bu aşamada eldeki önerilerin daha rafine olması ve artık daha zor vakaların ele alınabilir duruma gelmesidir.

Sınırlayıcı kutu regresyonunda parametrelerin (ötelemelerin) tahmini, IoU eşiği baz alınarak bir sınıfa ataması gerçekleşmiş teklifler ve onlara karşılık gelen yer doğrulukları arasında hesaplanmaktadır.

Genellikle 64 özellik haritasından oluşturulan bu modüldeki mini-yığınlar, mevcut öneriler arasından rasgele şekilde fakat %25'i ön plan teklifi (sınıflarıyla beraber), %75'i arka plan olmak üzere dengeli bir yapıda örneklenmektedir.

RPN'deki kayıpta izlenen yola benzer şekilde, sınıflandırma kaybı çok-sınıflı çapraz entropi kullanılarak ölçülmektedir. Benzer şekilde, yer doğruluklarıyla örtüşen ve

nesne olarak önerilmiş %25'lik kısım için sınırlayıcı kutu regresyon kaybının hesaplanmasında yine smooth-L1 fonksiyonu kullanılmaktadır. R-CNN'nin sınırlayıcı kutu regresyonunda her bir sınıf için ayrı tahmin üretiyor olmasıyla beraber, regresyon kaybının hesaplanmasında yalnızca o öneri için olasılığı en yüksek çıkmış olan sınıf kategorisine dair tahmin dikkate alınmaktadır.

2.1.4.2. Art işleme

Bu noktada, RPN'ye benzer şekilde, sınıflarına atanmış ve sınırlayıcı kutuları oluşturulmuş ancak üzerinde bir takım ek işlemlere ihtiyaç olan bir grup nesne teklifine erişilmiş bulunmaktadır.

Nihai nesnelere elde edilip, arka plan olarak öngörülenlerin elenmesinden sonra, nesnelere sınıfları da göz önüne alınarak bir kez daha NMS uygulanmaktadır. Bu işlem, nesnelere sınıflarına göre gruplandırılması, olasılıklarına göre sıralanması ve sonra belirli bir IoU eşiği belirlenerek her ayrı gruba NMS'nin uygulanması usulüyle yapılmaktadır.

2.1.4.3. Eğitim

Orijinal çalışmada Faster R-CNN, uçtan-uca değil çok-adımlı bir yaklaşım kullanılarak eğitilmiştir. RPN ve R-CNN modülleri birbirinden bağımsız olarak eğitilmiş sonrasında öğrenilen ağırlıkların birleştirilmesiyle topluca tekrar eğitime tabi tutulmuştur. Sonraki çalışmalarda eğitimi uçtan-uca gerçekleştirmenin daha iyi sonuçlar verdiği gösterilmiştir (He, G., Dollar, & Girshick, 2018).

Model bir araya getirildiğinde ikisi RPN ve ikisi R-CNN için olmak üzere 4 farklı kayıp fonksiyonu bulunmaktadır. Bu kısımlar dışında, yeteri kadar veriye sahip olduğunda takdirde omurga ağına da haricen eğitilmesi (ince ayar/fine tune yapılabilmesi) mümkündür.

Temel ağı eğitme kararı, tespiti istenen nesnelere doğasına ve eldeki hesaplama gücüne bağlıdır. Eğer, saptanması istenen nesnelere omurga ağı üzerinde eğitilmiş olduğu hazır veri setinde bulunanlara benzediği takdirde ince ayara ihtiyaç olmayacaktır. Aksi takdirde bu ağı üzerinde de ayarlama yapılması gerekmektedir.

Dört farklı türdeki kayıp, ağırlıklı bir toplam kullanılarak birleştirilmektedir. Böylelikle bu faktörlerden arzu edilenlerin etkisini diğerlerine göre artırıp/azaltmak mümkündür. Örneğin, regresyona göre sınıflandırma kayıplarına daha fazla ağırlık

(önem) vermek veya R-CNN kayıplarının RPN'den daha etkili olmasını sağlamak istenirse bunlarla alakalı ağırlık katsayılarının daha büyük tutulması bu amaca hizmet edecektir.

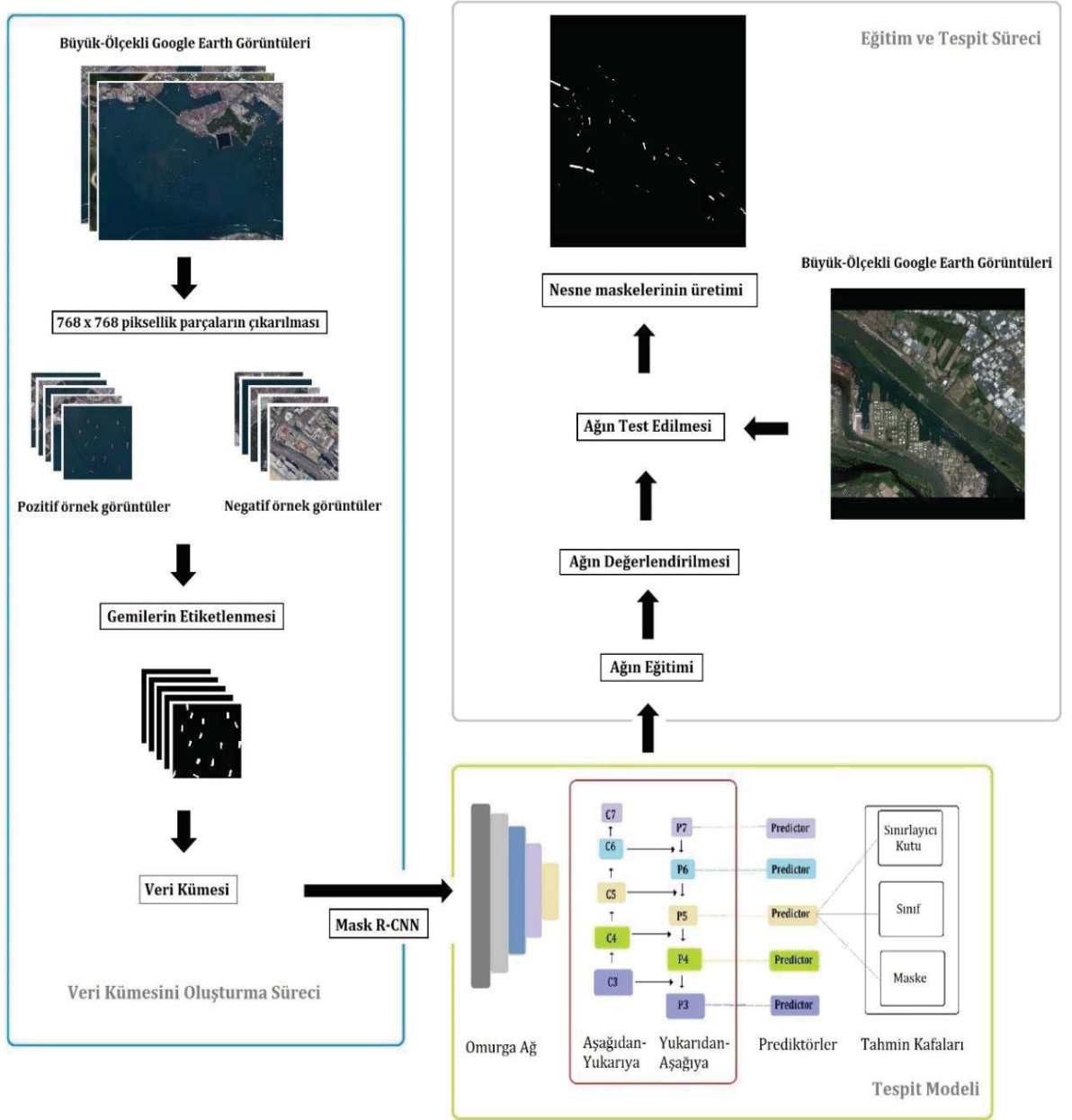
Orijinal çalışmada (Ren, He, Girshick, & Sun, 2016) öğrenme algoritması olarak, geri-yayılım algoritmasıyla beraber momentumlu Stokastik Gradyan İnişi (SGD with momentum) kullanılmıştır.



3. YÖNTEM

Bu bölümde uzaktan algılama görüntülerinde yer alan gemilerin tespiti için önerdiğimiz Mask R-CNN tabanlı algılma modeline dair detaylar ifade edilecektir. Önerilen yöntemin işlem adımları Şekil 3.1'de gösterilmektedir. Öneriler/katkılar dört ana başlık altında toplanmıştır. Özetle;

- Görüntülerdeki gemilerin tespit edilip maskelerinin ve sınırlayıcı kutularının çıkarılmasını sağlamak üzere Mask R-CNN tabanlı yeni bir gemi tespit modeli (Şekil 3.1; yeşil kutu)
- Bu modelin farklı görüntü ölçeklerinde etkinliğini korumasını sağlamak amacıyla özellik piramit ağının kullanımı (Şekil 3.1; kırmızı kutu)
- Yanlış tespitleri azaltmak amacıyla, eğitimde negatif örneklem veya kayıp fonksiyonunda odak kayıp/RetinaNet kullanımı
- Hataların görüntü içeriği ve nesne boyutları açısından daha nicel değerlendirilmesine olanak sağlayacak test veri kümesinin tanzimi önerilmektedir.



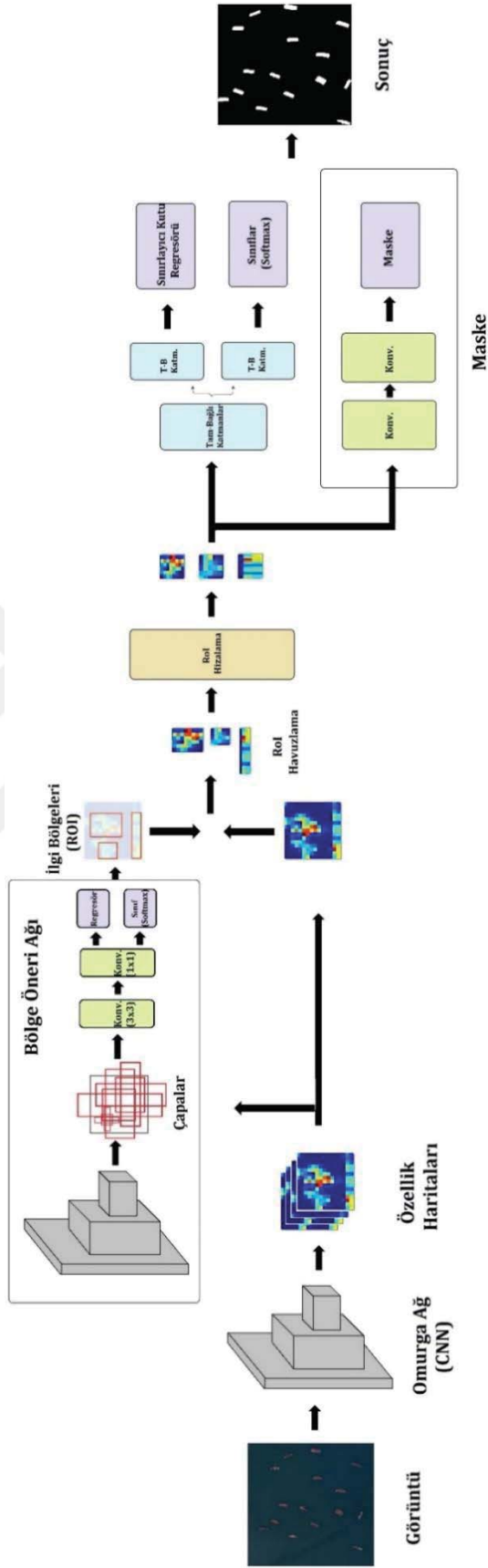
Şekil 3.1. Çalışma sürecinin özeti

3.1. Mask R-CNN Tabanlı Gemi Tespit Çerçevesi

Gemi tespitinde yaygın olarak kullanılan Faster R-CNN tespit çerçevesinden farklı olarak bu çalışmada yararlanılan Mask R-CNN modeliyle tespit, nesneyi çevreleyen sınırlayıcı kutuyla değil, o gemiye ait piksel gruplarının segmentasyonu ile yani maskelenmesiyle çıkarılmaktadır. Faster R-CNN üzerine geliştirilmiş olan Mask R-CNN yönteminde de özellik çıkarma, bölge teklifi hazırlama, maske+sınırlayıcı kutu regresyonu ve sınıflandırma tek bir model tarafından gerçekleştirilmektedir. Bu çalışma

kapsamında önerilen ve Mask R-CNN modelinin temel alındığı gemi tespit modelinin akış diyagramı Şekil 3.2’de gösterilmektedir.



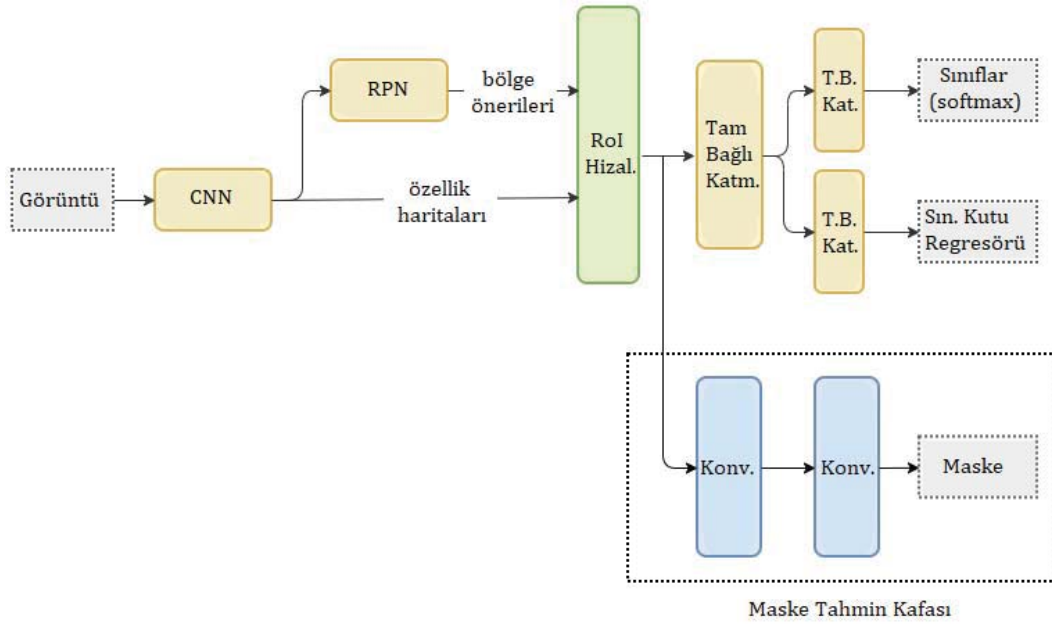


Şekil 3.2. Mask R-CNN modelinde işlem adımları

Bu modelde, RGB bantlarındaki uydu görüntüleri girdi olarak kullanılarak Mask R-CNN aracılığıyla şu çıktılar üretilmektedir:

- Gemileri içeren maskeler + sınırlayıcı kutular
- Her maskenin etiketlenilmesi (gemi olarak sınıflandırılması)
- Her gemi (maske) için nesnellik değerleri (tespit edilen nesnenin ne oranda gemi olduğunun modelce hesaplanan değeri)

Mask R-CNN, Faster R-CNN tespit yöntemi gibi iki-aşamalı işlemektedir, RPN kısmı aynen kullanılmaktadır. İkinci aşamada, RPN'in akabinde Faster R-CNN'den farklı olarak İlgi Bölgesi Hizalaması (RoI Align) yapılmakta ve sonrasında tahminleri paralel yürütülen sınıf ve kutu ayarlamalarına ilave olarak her RoI için ayrıca ikili (binary) maske çıktısı üretilmektedir (Şekil 3.3).



Şekil 3.3. Mask R-CNN'le girdi görüntüsü kullanılarak nesnelere dair sınıf, sınırlayıcı kutu ve maske tahminleri üretilmektedir.

3.1.1. Nesnenin maskeyle temsili: maske tahmin kafası

Mask R-CNN'de, sınıflandırma ve sınırlayıcı kutu kayıplarına ilave olarak maskelemenin de model tarafından öğrenilip gerçekleştirilebilmesi için eğitimde örneklenen her RoI için maske kaybını da içeren bir çoklu-görev kaybı tanımlanmıştır: $L = L_{cls} + L_{box} + L_{mask}$. Sınıflandırma kaybı (L_{cls}) ve sınırlayıcı-kutu kaybı (L_{box}) sırasıyla Denklem 2.1 ve Denklem 2.2 ile aynıdır. Her RoI için tahmin kafasının maske

dalı m^2 -boyutlu bir maske çıktısı üretmektedir. Maskeye dair kayıp (L_{mask}) ortalama ikili çapraz entropiyle tanımlanmıştır ve bu işlemden önce her girdi pikseline sigmoid fonksiyonu uygulanmaktadır.

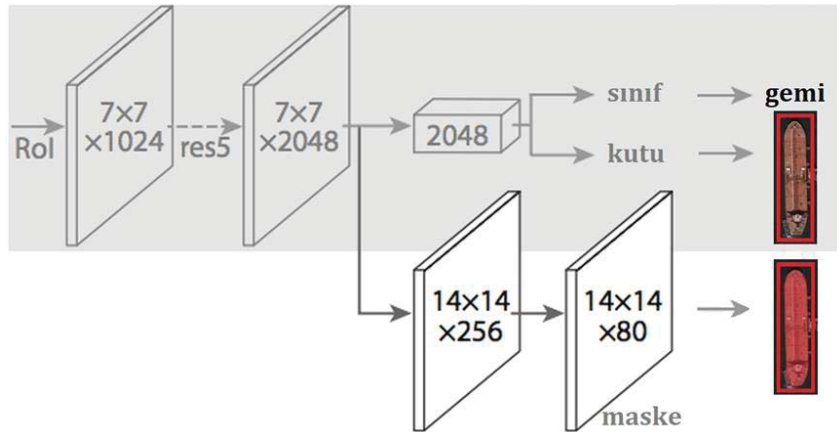
$$CE = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)) \quad (3.1)$$

$$sigmoid \text{ fonksiyonu: } f(x) = \frac{1}{1+e^{-x}} \quad (3.2)$$

Denklem 3.1’de $N=mxm$ ’dir ve y_i gemi olarak sınıflandırılan pikseller için 1, gemi olmayanlar için 0 değerini almaktadır.

Bir girdi nesnesinin mekansal şeklini görüldüğü haliyle temsil etme kabiliyetinde olan maskeleme işlemi, tam-bağlantılı (fully-connected, fc) katmanların kullanımını gerektiren sınıf etiketlendirmesi veya sınırlayıcı kutu regresyonundan farklı olarak konvolüsyon işlemiyle ele alınabilmektedir. Mask R-CNN’de tam konvolüsyonel ağlar kullanılarak her RoI için mxm (genellikle $14x14$ veya $28x28$) piksel içeren maske tahmini yapılmaktadır.

Süreç sonunda model tarafından her ilgi bölgesi için sınıflandırma, sınırlayıcı-kutu ve maske tahminleri üretilmiş olur (Şekil 3.4).

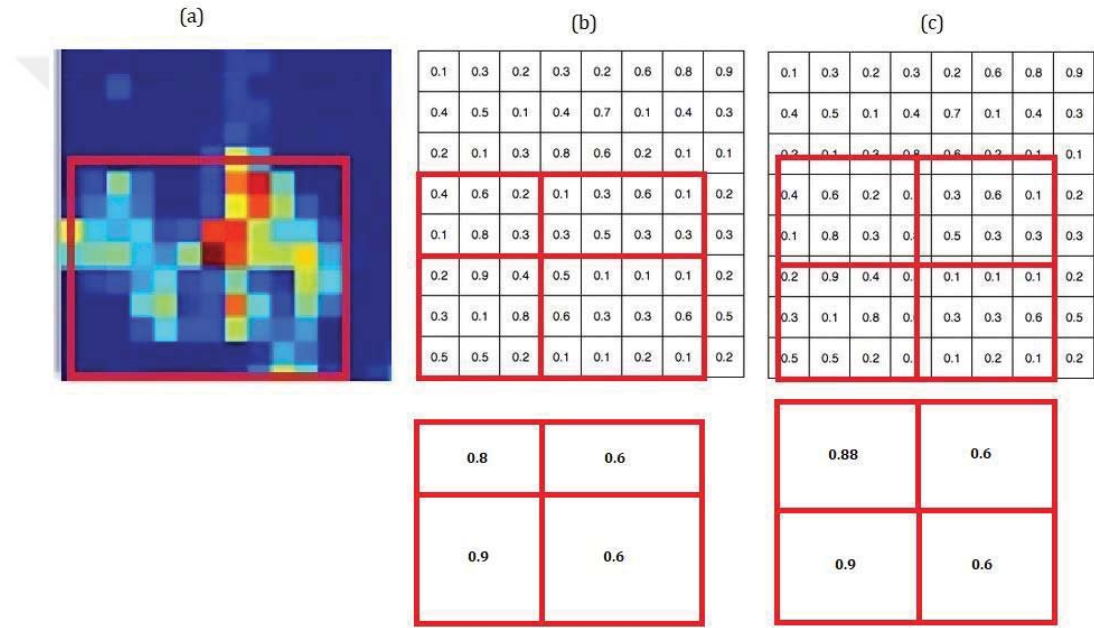


Şekil 3.4. Mask R-CNN modelinde üç çıktı alınmaktadır: Nesne sınıfı, nesne sınırlayıcı kutusu ve nesne maskesi.

3.1.2. İlgi bölgelerinin (RoI) hizalanması

Mask R-CNN'nin diğer bölge-tabanlı tespit modellerine göre önemli bir artışı, RoI havuzlaması üzerinde yapılan iyileştirmededir. İkinci bölümde detayları verilmiş olan Faster R-CNN’de kullanılan RoI havuzlamasında hedef özellik haritasında muhtemel

nesneyi temsil eden hücre sınırları (Şekil 3.5.a sütunu), girdi özellik haritasının sınırlıya (pikselleriyle) yeniden hizalanmaya zorlanmaktadır yani sınırlar dijitalleştirilmektedir (quantization) (Şekil 3.5.b sütunu). Bu nedenle, sınırlaylayıcı kutuyu meydana getiren her hedef hücreler farklı boyutlarda olabilmektedir. Mask R-CNN’de ise hücrelerin sınırını dijitalleştirmeyen (Şekil 3.5.c üst diyagram) RoI Hizalaması (RoI Align) kullanılmakta ve her hedef hücrenin aynı boyuta (Şekil 3.5.c alt diyagram) olması sağlanmaktadır. Ayrıca ilgi bölgesi hücrelerinin içindeki değerleri daha iyi hesaplamak için enterpolasyon uygulanmaktadır. Orijinal çalışmada da belirtildiği gibi RoI Hizalaması doğruluk üzerinde önemli iyileştirmeler sağlamaktadır (Şekil 3.6).



Şekil 3.5. RoI Hizalama: Mask R-CNN modelinde Faster R-CNN’den farklı olarak, özellik haritasının hücre değerlerinin enterpolasyonu kullanılarak öneri bölgelerinin nesnelere daha iyi oturması sağlanmaktadır.

	AP	AP ₅₀	AP ₇₅	AP ^{bb}	AP ₅₀ ^{bb}	AP ₇₅ ^{bb}
<i>RoIPool</i>	23.6	46.5	21.6	28.2	52.7	26.9
<i>RoIAlign</i>	30.9	51.8	32.1	34.0	55.3	36.4
	+7.3	+ 5.3	+10.5	+5.8	+2.6	+9.5

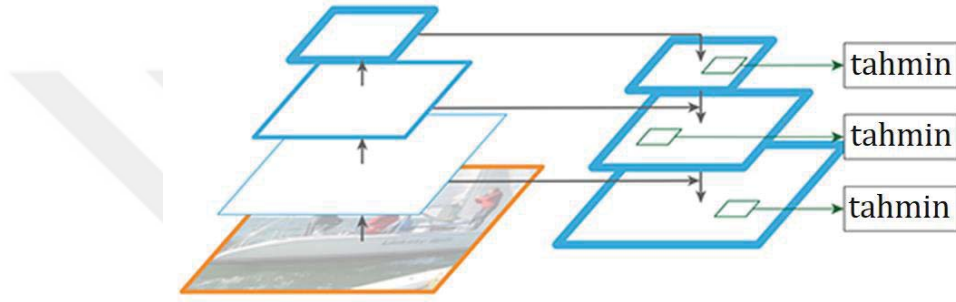
Şekil 3.6. Mask R-CNN makalesinde RoI Hizalamanın (Align) havuzlama katmanına göre nesne tespiti üzerindeki etkinliği gösterilmiştir.

3.2. Gemi Tespitinde Mask R-CNN’nin Özellik Piramit Ağıyla Kullanımı

Değişken ölçekli görüntülerde gemi tespiti için kullanılan yollardan biri, görüntünün çeşitli ölçeklere yeniden boyutlandırılmasıyla katmanlar halinde oluşturulan,

aşağıdan yukarıya doğru çözünürlüğün azaldığı görüntü piramitlerinin oluşturulmasıdır (Şekil 3.7). Bu yolla nesnelerin farklı ölçeklerdeki özelliklerinin çıkarılması sağlansa da çözünürlüğün düşmesiyle beraber küçük gemilerin veya nesnelerin gözden kaçırılması kaçınılmaz olmaktadır.

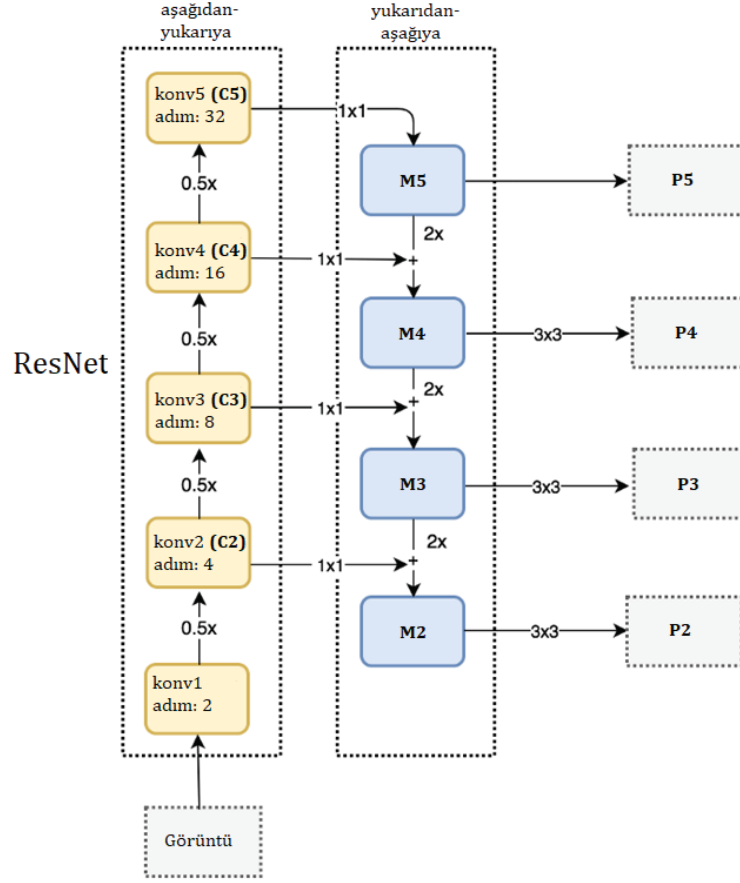
Ayrıca, bu katmanlardaki herbir görüntü, özelliklerinin çıkarılması için CNN'e girildiğinde ağ içerisindeki her blok sonrasında üretilen özellik haritalarının derinliği artarken çözünürlüğü düşmektedir. Ancak üst katmanlar semantik açıdan zenginken alt katmanlar düşük semantik bilgi taşımaktadır.



Şekil 3.7. Nesne tespitinde özellik piramit ağının kullanımı.

3.2.1. Aşağıdan yukarıya yol

Bu çalışmada aşağıdan yukarıya yol oluşturulurken (özellik haritaları çıkarılırken) ResNet blokları kullanılmıştır (Şekil 3.8). Toplam 5 adet olan bloklar, çeşitli sayılarda konvolüsyonel katman içermektedir. Yukarı çıkıldıkça, ardışık blok katmanları arasında mekansal boyut yarı yarıya azalmakta, böylece adım (stride) sayısı iki katına çıkmaktadır. Her konvolüsyon bloğunun çıktısı C_i ile gösterilmektedir. Bu bloklardan gelen çıktılar yukarıdan-aşağıya yolu inşa etmede kullanılmaktadır.

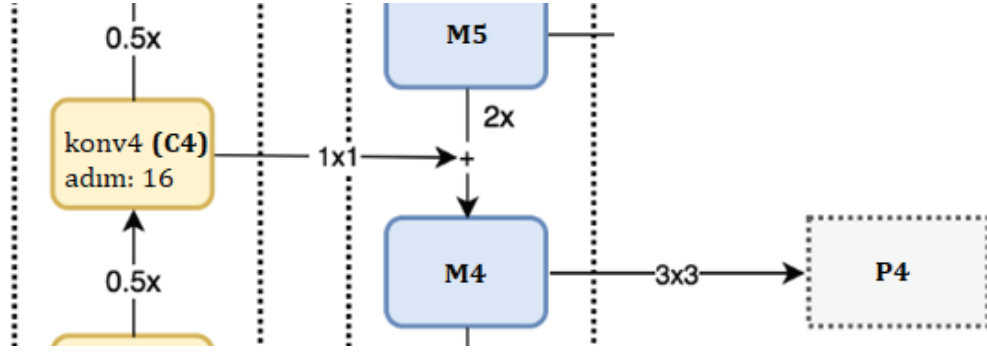


Şekil 3.8. Özellik piramit ağında yukarıdan-aşağıya tanımlanan yeni yolda mekansal çözünürlüğü düşük, semantik değeri yüksek özellik haritaları, alt katmanla kurulan yanak bağlantılar ve üst-örnekleme kullanılarak semantik değeri korunarak mekansal boyutu artırılmaktadır.

3.2.2. Yukarıdan-aşağıya yol

Yukarıdan-aşağıya yolda oluşturulan ilk özellik haritası katmanı M5'tir. M5'i oluşturmak üzere, 1×1 'lik bir konvolüsyon filtresi uygulanarak C5'in kanal derinliği 256 kanala düşürülmektedir. Bu, nesne tahmininde kullanıma hazır ilk özellik haritası katmanıdır.

Yukarıdan aşağıya inilirken, en yakın komşuluk üst-örnekleme (nearest neighbors upsampling) kullanılarak üst katmanın mekansal çözünürlüğü 2 katına çıkarılmaktadır (Şekil 3.9). Alt katmanda karşılık gelen ilgili özellik haritasına 1×1 'lik konvolüsyon uygulanmasından sonra mekansal boyutlarıyla beraber derinlikleri de denk hale getirilen iki katman birbirlerine eklenmektedir. Bozucu (aliasing) etkileri hafifletmek üzere birleştirilmiş tüm katmanlara ayrıca 3×3 'lük bir konvolüsyon da uygulanmaktadır.

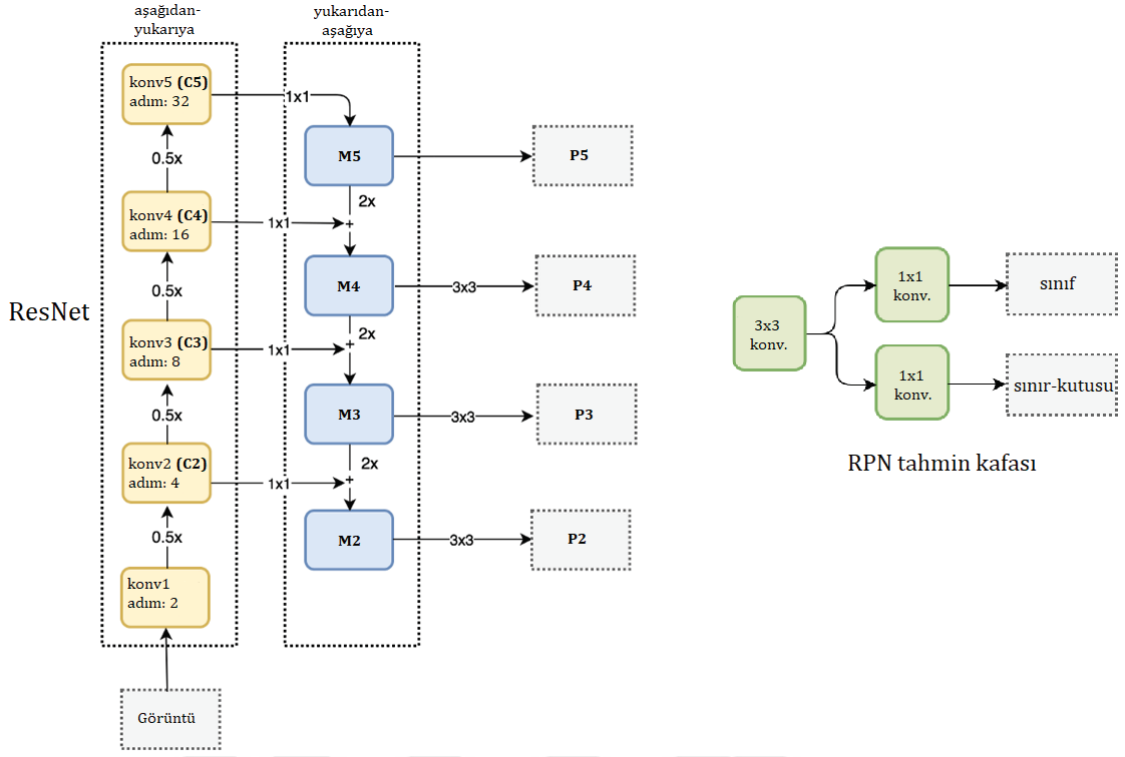


Şekil 3.9. Yukarıdan-aşağıya yolda mekansal çözünürlük, alt katmanla yanıl bağlantının da kullanımıyla 2 katına çıkmaktadır.

Aynı işlemler P3 ve P2 için de tekrar edilmektedir. Süreç, C1'in mekansal boyutu çok büyük olduğundan ve işlem sürecini yavaşlattığından P2'de durdurulmaktadır. Her özellik haritası çıktısı için son (tam-bağlı katman yapısında olan) R-CNN modülünde aynı sınıflandırıcı ve kutu regresörü kullanıldığından tüm piramit özellik haritalarının {P5, P4, P3, P2} derinlikleri aynı ($d=256$) tutulmaktadır.

3.2.3. FPN'nin bölge öneri ağında kullanımı

Kendi başına bir nesne dedektörü olmayan FPN, nesne dedektörlerinin içerisinde çalışan bir özellik algılayıcısıdır. Bu çalışmada FPN, yukarıdan-aşağıya yolla semantik açıdan zengin yüksek çözünürlüklü özellik harita katmanlarının oluşturulup akabinde gemiler için bölge önerileri üretmede yani RPN'yi beslemede kullanılmıştır (Şekil 3.10).

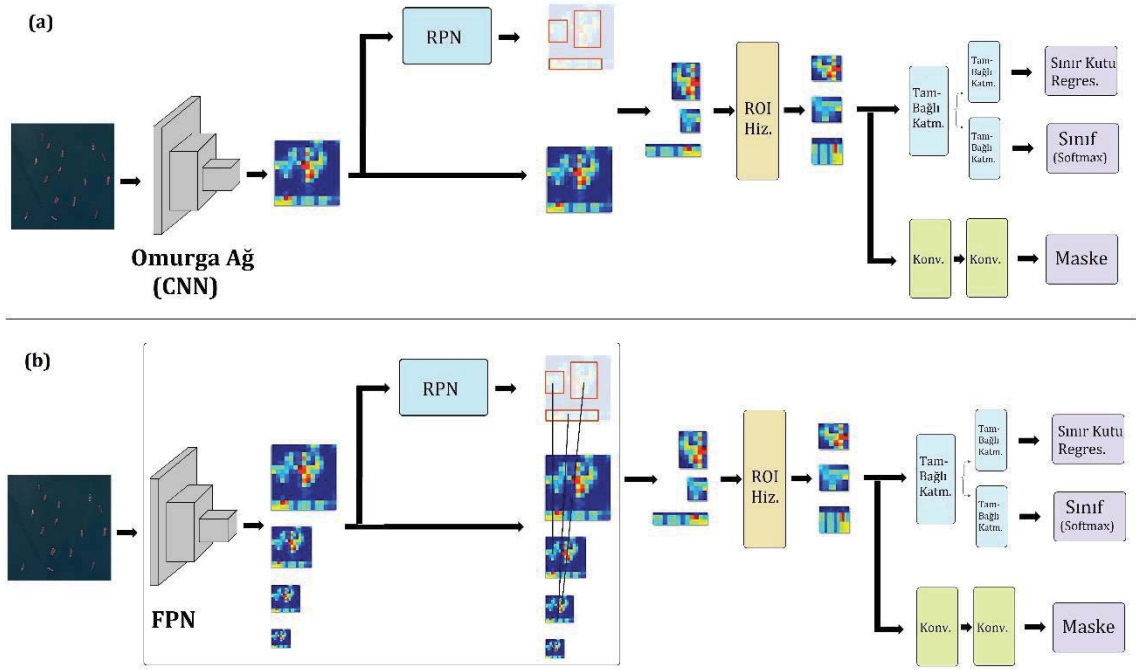


Şekil 3.10. Nesne tespitinde FPN kullanıldığında RPN'deki tahminler, oluşturulan yeni P_i katmanları kullanılarak gerçekleştirilir.

3.2.4. FPN'nin Mask R-CNN'de kullanımı

Mask R-CNN'e dair veri akış diyagramı Şekil 3.11'de gösterilmektedir. FPN kullanılmadığında, ROI'lerin oluşturulmasında omurga ağının son bloğundan gelen tek özellik katmanı kullanılmaktadır (Şekil 3.11.a).

Modelde FPN kullanıldığındaysa özellik haritalarından oluşan bir piramit üretilmektedir (Şekil 3.11.b). ROI'leri oluşturmak için uygulanan RPN sürecinde özellik yamalarını çıkarırken ROI'nin boyutuna bağlı olarak en uygun ölçekteki özellik harita katmanı seçilmektedir.



Şekil 3.11. Mask R-CNN modelinde RPN modülünün a) omurga ağdan gelen özellik haritasıyla, b) FPN'de üretilen özellik haritalarıyla beslenmesi. FPN kullanıldığında piramitteki hangi katmanın kullanılacağı nesne boyutuna bağlıdır.

Genişliği w , yüksekliği h ile temsil edilen bir RoI için piramitteki hangi özellik haritasınının seçilmesi gerektiği şu eşitlik kullanılarak belirlenmektedir:

$$k = [k_0 + \log_2(\frac{\sqrt{wh}}{224})] \quad (3.3)$$

Burada $k_0 = 4$ olup k ise RoI'nin sınır kutusunun büyüklüğüne göre değeri belirlenen ve FPN'de özellik yaması üretmekte kullanılacak olan katman seviyesini (P_k) temsil etmektedir. Yukarıdaki işlem neticesinde, örneğin, k 'nın 3'e eşit olması durumunda, o RoI için özellik haritası olarak P3 seçilmektedir. Akabinde, tahmin sürecini tamamlamak üzere, elde edilen yamalar RoI havuzlamasına/hizalamasına tabi tutulmakta ve sürecin kalanı Mask R-CNN'le aynı şekilde devam etmektedir.

3.3. Hatalı Tespitlerin Azaltılması İçin Yapılan Düzenlemeler

Derin öğrenmenin kullanıldığı nesne tespit modellerinde üretilen yanlış alarmların büyük çoğunluğu hedef nesneyle benzerlik taşıyan öğelerin detektörce nesne sanılmasından kaynaklanmaktadır. Gemi tespitinde de hatalı alarmların büyük bölümü karasal bölgelerde yer alan ve şekilsel olarak gemilerle benzerlik arz eden çatılar, ince uzun betonarme bloklar gibi yapılardan oluşmaktadır. Bölge-tabanlı herhangi bir modelde bu türde hatalı tespitlerin önüne geçebilmek için modelin sınıflandırmadaki nesnellik

eşiğinin yüksek tutulması bir çözüm olmasına rağmen bunu yapmak diğer taraftan bazı gerçek gemilerin gözden kaçmasına yol açmakta ve tespitteki doğruluk oranını düşürmektedir. Bu soruna çözüm sağlamak üzere iki türde yaklaşım kullanılmaktadır. Birincisi, eğitim veri setlerinin hatalı tespitlerin önüne geçecek şekilde yeniden düzenlenmesini içeren yöntemler, diğeri ise veriye dokunmadan modelin kayıp fonksiyonu üzerinde değişikliğe giderek tespit performansını artırma amacıyla olanlar.

Bu yaklaşımların ilki negatif örneklem eğitimi olarak isimlendirilmektedir. Temel olarak, veri setine model için “kafa karıştırıcı” örneklerin dahil edilmesine dayanmaktadır. İki farklı şekilde ele alınmaktadır. Zor negatif örneklem eğitimi denen birinci yol, görüntülerde yer alan ve hedef nesnelere benzerlik taşıyan, kafa karıştırıcı örneklerin hassasiyetle hazırlanıp eğitim verisine eklenmesine ve modelin bu örneklerle eğitilmesine dayanmaktadır. Diğer yoldaysa modelin eğitimi, gemi içeren görüntülerle hazırlanmış standart veri setine ilave olarak gemi içermeyen karasal bölgelerden rastgele hazırlanan görüntü parçalarının da herhangi bir etiketleme yapılmadan kullanılmasıyla gerçekleştirilmektedir ve negatif örneklemle eğitim olarak adlandırılmaktadır. İlk yaklaşım kadar hassasiyet gerektirmemesi sebebiyle özellikle bölge-tabanlı tespit modellerinde kullanılmıştır (Gao, He, Sun, Jia, & Zhang, 2019). Çalışmanın bu kısmında da yanlış alarmları azaltmak adına kullanılması önerilen ilk yöntem negatif örneklem eğitimi olmuştur.

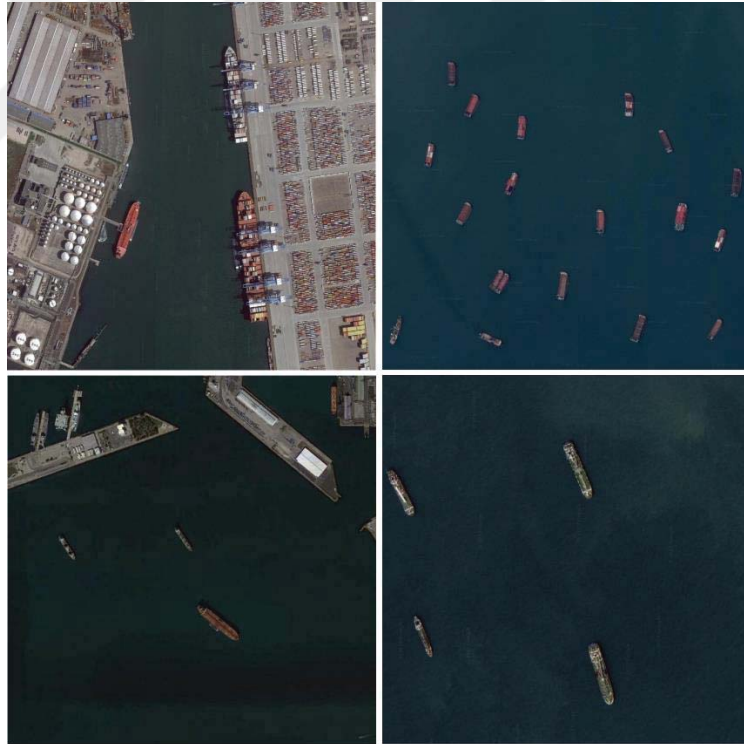
Yanlış tespitlerin azaltılmasını hedefleyen diğer yaklaşımdaysa eğitim verisine dokunulmamakta, modelin kayıp fonksiyonu uygun şekilde modifiye edilmektedir. RetinaNet adı verilen bu yöntemde, modelin kayıp hesaplamasında kullandığı, RPN’de üretilen çok sayıda nesne önerileri arasından zor olanların etkisi odak kayıp fonksiyonu kullanılarak artırılmakta böylelikle model, basit örnekler yerine zor öneriler üzerine odaklanarak ayırım kabiliyetini artırmakta ve daha az hataya düşmektedir.

3.3.1. Mask R-CNN-tabanlı gemi tespiti için negatif örneklem eğitimi

Özellik çıkarımının çok sayıda görüntünün kullanılmasıyla gerçekleştirildiği derin öğrenme yönteminde ve dolayısıyla bu yaklaşımı kullanan tespit sistemlerinde veri sayısının artırılması performansı çoğunlukla olumlu etkilemektedir. Modelin hem hedef nesnelere hem de hedef olmayanları tanıması bu veriler sayesinde mümkündür. Çalışmanın bu kısmında, modelin eğitim verisi için, gemi içeren standart verilere ilave olarak karasal alan görüntüleri negatif örnekler olarak eklenmekte böylece modelin

bunlara dair özellikleri öğrenmesi amaçlanmaktadır. Bu yolla, daha sonra karasal alanlardan gelen öneriler karşısına çıktığında, ağ bunları gemi olmayan arka plan olarak tanımak üzere eğitilmiş olmaktadır.

Bu çalışmada kara-deniz segmentasyonuna başvurmadan, karasal bölgelerden kaynaklı yanlış alarmları önlemek amacıyla gemi içeren standart eğitim verisine (Şekil 3.12) ilave olarak, içeriğini karasal alanların oluşturduğu, gemi bulunmayan çok sayıda rastgele negatif görüntü örneğinin eklenmesi ve modelin bunlarla eğitilerek tespit için kullanılması birinci seçenek olarak önerilmektedir. Zor negatif örneklemeden farklı olarak, Şekil 3.13'te gösterilen negatif örnekler etiketlenmeye gerek kalmayacak şekilde ağ eğitmek için hedef-dışı örnekler olarak kullanılmaktadır. Verinin bunlar haricindeki kısmında gemi hedefleri etiketlendiğinde, negatif örnek görüntüler pozitif örnek görüntülerden ayrılmış olmaktadır. Dolayısıyla negatif örnek görüntülerin model eğitiminde kullanılabilmesi için yalnızca boş etiketlerin oluşturulması yeterli olmaktadır.



Şekil 3.12. Modelin eğitiminde kullanılan pozitif örneklem görüntüleri



Şekil 3.13. Modelin eğitiminde kullanılan negatif örneklem görüntüleri

3.3.2. Gemi tespitinde Mask R-CNN ve odak kayıp fonksiyonunun kullanımı

Hem bu çalışma kapsamında incelenmiş olan iki-aşamalı, bölge-tabanlı nesne tespit yöntemlerinin hem de tek-aşamalıların paylaştığı ortak yön, veri dengesizliği sorununa yol açan, yöntemin işleyişi sırasında üretilen ve nesne olmayan arka plan kutularının sayısının fazla olmasıdır. Tespit modelleri bu girdilerle öğrenmeyi sağlamak üzere tüm öneri kutularını eş şekilde ele alarak hata miktarını ölçmektedirler. Buradaki sorun ele alınan zor örneklerle, çözümü kolay örneklerin öğrenme üzerindeki etkisinin aynı olması sebebiyle modelin zor örneklerle karşılaştığında yeterli çözümleme kabiliyetinden uzak olmasıdır. RetinaNet yönteminde tanımlanan odak kayıp fonksiyonuyla modelin bu tür zor örneklere ilişkin hata ağırlıkları artırılarak odaklanması sağlanmaktadır.

Bu tez kapsamında, gemi tespit sonuçlarını iyileştirmek üzere odak kayıp fonksiyonunun kullanımı da önerilmiştir. Bu amaçla tespit için RetinaNet yaklaşımı kullanılmış ve kayıp fonksiyonundaki ayarlanabilir odak parametresi değiştirilerek sonuca etkisi gözlenmiştir. Çalışmamızda, (Lin, Goyal, Girshick, He, & Dollar, 2018) tarafından önerilen odak kayıp fonksiyonu kullanılmıştır. Kolay ve iyi sınıflandırılmış örneklerle alakalı kaybın ağırlığını azaltıp, ayrımı zor olan örneklerinkini artıracak şekilde değiştirilerek çapraz entropi kaybını düzenlemektedir.

3.3.2.1. Eğitim

Mask R-CNN’de olduğu gibi bu yöntemde de kayıp fonksiyonu, sınırlayıcı kutu regresyonu, sınıflandırma ve maske kayıp fonksiyonlarının toplamından oluşmaktadır.

Yer-doğruluğu kutusu ve alakalı çapa arasındaki birbirlerine göre düzeltme hesaplamaları (Girshick, Donahue, Darrell, & Malik, 2014), (Lin, Goyal, Girshick, He, & Dollar, 2018)’de verildiği şekliyle yapılmaktadır. Sınırlayıcı kutu kaybının hesaplanmasında kutu merkezlerinin koordinatları yerine kutuların sol-üst ve sağ-alt köşe noktalarının koordinatları kullanılmaktadır. (X_{1b}, Y_{1b}) ve (X_{2b}, Y_{2b}) , sırasıyla, yer-doğruluğu kutusunun sol-üst ve sağ-alt köşeleri ve (X_{1a}, Y_{1a}) ve (X_{2a}, Y_{2a}) , sırasıyla, çapa kutusunun sol-üst ve sağ-alt köşelerini temsil etmektedir. Hedefler şu şekilde hesaplanmaktadır:

$$W_a = (X_{2a} - X_{1a}) \quad (3.4)$$

$$H_a = Y_{2a} - Y_{1a} \quad (3.5)$$

$$X_{1t} = (X_{1b} - X_{1a})/W_a \quad (3.6)$$

$$Y_{1t} = (Y_{1b} - Y_{1a})/H_a \quad (3.7)$$

$$X_{2t} = (X_{2b} - X_{2a})/W_a \quad (3.8)$$

$$Y_{2t} = (Y_{2b} - Y_{2a})/H_a \quad (3.9)$$

Burada W_a ve H_a çapanın sırasıyla genişliğini ve yüksekliğini ve (X_{1t}, Y_{1t}) ve (X_{2t}, Y_{2t}) ise hedef kutusunun sırasıyla sol-üst ve sağ-alt köşelerini temsil etmektedir. (X_{1p}, Y_{1p}) ve (X_{2p}, Y_{2p}) ise tahmin edilen sınırlayıcı kutunun sırasıyla sol-üst ve sağ-alt köşelerini temsil etmektedir ve regresyon kaybı, smooth-L1 fonksiyonu kullanılarak hesaplanmaktadır:

$$L_{reg}(t_i, p_i) = \text{smooth}_{L1}(t_i - p_i) \quad (3.10)$$

$$\text{smooth}_{L1}(d) = \begin{cases} 0.5d^2 & \text{eğer } |d| < 1 \\ |d| - 1 & \text{diğer türlü,} \end{cases} \quad (3.11)$$

Arka plan örneklerinin ön plandakilerden daha fazla olmasından kaynaklanan model içindeki veri işlem sürecinde oluşan sınıf dengesizliği ile başa çıkmak ve doğruluk oranını artırabilmek için RetinaNet yaklaşımında Denklem 3.10’da ifade edilen hata fonksiyonu yerine odak kayıp fonksiyonu tanımlanmıştır. Bu yaklaşımda γ , ayarlanabilir

odaklama parametresi olmak üzere, çapraz entropi kaybına, $(1 - p_t)^\gamma$ ile tanımlanan bir hassas ayar faktörü eklenmektedir. Odak kayıp fonksiyonu şöyle tanımlanmaktadır:

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma(p_t) \quad (3.12)$$

Negatif/pozitif örneklerin önemini dengeleyen α ağırlıklandırma faktörüyle modelin zor örnekler üzerine odaklanması sağlanmaktadır. Odaklanma oranı γ parametresiyle değiştirilmektedir.

Bu kayıp fonksiyonu, p_t 'nin küçük değerler aldığı hatalı sınıflandırılmış örnekler ve α 'nın bire yakın olduğu durumlarda çapraz-entropi kaybına eşit olmaktadır. İyi sınıflandırılmış örnekler p_t 'yi 1'e ve hassas-ayar faktörünü 0'a yakınsatmaktadır. Böylece, iyi sınıflandırılmış örneklerde kaybın ağırlığı azaltılmış olmaktadır. Kaybın ağırlığının azaltılması γ ile belirlenmektedir. Deneylerimizde α değeri 0.25'te sabit tutulmuş ve γ değerlerindeki oynamaların tespiti etkisi incelenmiştir.

3.4. Veri Kümesinin İçeriğe ve Gemi Büyüklüklerine Göre Tanzimi

Uzaktan algılama görüntülerinde gemi tespiti literatüründeki önemli eksiklerden biri tespit modellerinin performans değerlendirmelerinin gemi boyutlarına ve farklı zorluk seviyelerindeki içeriklere göre nicel değerlendirmelerinin genellikle yapılmıyor olmasıdır. Bir çalışmada (Liu, Cui, Kuang, & Li, 2017) gemi boyutlarına göre bir değerlendirme söz konusuysen, bir diğesinde (Gao, He, Sun, Jia, & Zhang, 2019) ise içerik zorluğu açısından bir değerlendirme geniş-ölçekli görüntüler üzerinden yapılmıştır.

Bilgisayarlı Görme alanında geliştirilen nesne tespit modelleri, performansları değerlendirilmek üzere COCO gibi veri setleri üzerinde çalıştırıldıklarında tespiti konu olan her nesne sınıfı, büyüklüklerine göre küçük, orta ve büyük olmak üzere gruplandırılmaktadırlar. COCO'da, görüntüdeki nesnenin yer-doğruluğu maskesinin içerdiği piksel sayısına göre; $32 \times 32 (= 1024)$ 'den az olduğu takdirde nesne 'küçük' büyüklük sınıfına, 32×32 ile $96 \times 96 (= 9216)$ arasında olduğunda 'orta' büyüklük sınıfına, 96×96 'nın üzerindeyse 'büyük' kategorisine girmektedir. Gemi tespiti alanında da bu yönde bir ayırımın ve analizin yapılması tespit sonuçlarının daha sağlıklı incelenmesine yardımcı olacaktır.

Bu çalışmada oluşturulan veri setlerindeki gemiler hem COCO standartlarına göre büyüklük gruplarına hem de uzunluklarına göre $0-50m$, $50-100m$, $100-150m$, $150-200m$, $200-250m$, $250-300m$ ve $300m-$ olacak şekilde gruplandırılmışlardır.

Çalışmamızda veriler ayrıca içerik açısından da kategorize edilmişlerdir. Özellikle geniş-ölçekli görüntülerde karasal alanların yoğun olduğu bölgelerde gemilere benzer yapıların yanlış alarmlara yol açtığı birçok araştırma tarafından ortaya konmuş durumdadır. Veri setlerini oluşturan görüntülerin bir kısmında karasal alanlar yer almazken kimisinde karasal alanların baskın olduğu iç sular veya bazılarındaysa deniz ve karanın daha dengeli yer aldığı liman/rıhtım yakınları gibi içeriklerle karşılaşmak mümkündür. Gemi tespitinde tamamen farklı zorluklar içeren bu içeriklerin gruplandırılması hata analizinin daha sağlıklı yapılmasına yardımcı olacaktır.

Bu göz önüne alınarak çalışmamızda verilerin içeriklerine göre de sınıflandırılması önerilmiş olup eldeki görüntüler

- Açık sular
- Liman/rıhtım yakınları
- İç sular

olmak üzere üç gruba ayrılmıştır. Bu gruplandırmanın diğer bir yararı ise bu üç ayrı koşulda bulunan gemilerin birbirine göre yakınlık durumlarının da değişmesidir. Açık sulardaki gemiler genellikle birbirinden uzakta bulunurken diğer iki sahnede çoğunlukla birbirlerine yakın konumlanmış haldedirler. Bu da üzerinde nesne tespit modellerinin performansının değerlendirilmesi gereken diğer bir önemli faktördür.

4. VERİ SETİNE VE MODEL EĞİTİMİNE DAİR DETAYLAR

4.1. Veri Seti

Veri setimiz, Google Earth tarafından sağlanan 1 metre mekansal çözünürlüğe sahip toplam 9714 uydu görüntüsünden oluşmaktadır (Tablo 4.1). SAS Planet yazılımı aracılığıyla indirilmişlerdir. Bu görüntüler kırmızı, yeşil ve mavi olmak üzere 3 bantta bilgi taşımaktadırlar. Enlem ve boylam bilgilerinin bulunduğu Geotif formatındadırlar. Görüntüler sivil limanlar, askeri üsler, açık sular, iç sular gibi sahneleri içermektedir. Görüntü boyutları 768x768 pikseldir. Görüntülerde en az bir gemi bulunmaktadır. Eğitim, validasyon ve test veri setlerini oluşturan görüntülerin sayılarının birbirine göre oranları yaklaşık olarak, sırasıyla, 7:2:2'dir.

Tablo 4.1. Çalışma kapsamında kullanılan veri kümesinin içeriği

Eğitim Kümesindeki Görüntü Sayısı	Validasyon Kümesindeki Görüntü Sayısı	Test Kümesindeki Görüntü Sayısı	Toplam Görüntü Sayısı	Eğitim, Validasyon ve Test Görüntülerinin Oranı (Yaklaşık)	Görüntü Kaynağı	Görüntü Boyutu (piksel x piksel)	Mekansal Çözünürlük
6217	1727	1770	9714	7:2:2	Google Earth	768 x 768	1.0 m

Tüm setlerdeki görüntülerde toplam 11534 gemi yer almaktadır. İçerilen gemi sayılarının herbir kümeye göre dağılımı Tablo 4.2'de gösterilmiştir. Görüntü sayısında olduğu gibi, setlerin gemi sayılarının oranı da yaklaşık 7:2:2'dir.

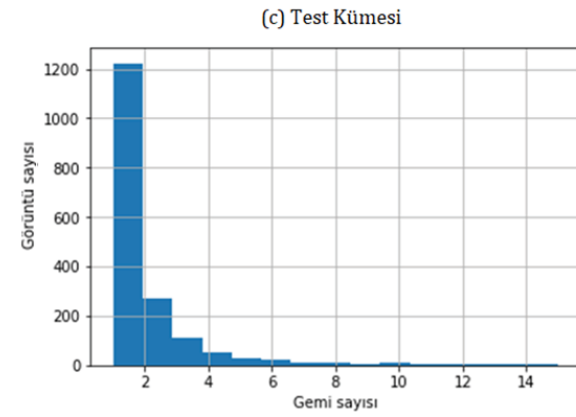
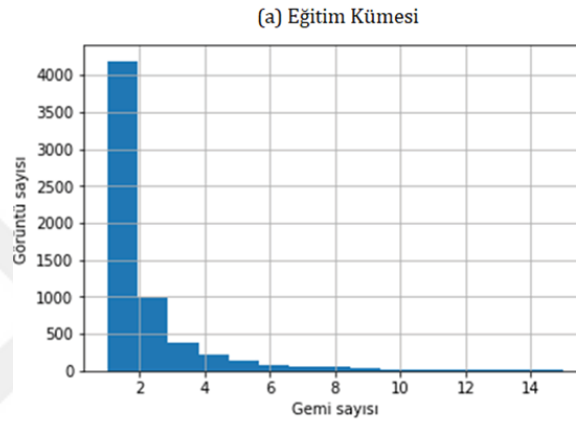
Tablo 4.2. Veri kümesi içeriğindeki gemi sayıları

Eğitim Kümesindeki Gemi Sayısı	Validasyon Kümesindeki Gemi Sayısı	Test Kümesindeki Gemi Sayısı	Toplam Gemi Sayısı	Eğitim, Validasyon ve Test Kümelerindeki Gemi Sayılarının Oranı (Yaklaşık)
11534	3197	3180	17911	7:2:2

Görüntülerdeki gemi sayılarına dair bazı bilgiler Tablo 4.3'de verilmiştir. Görüleceği üzere, hemen her kümede benzer istatistikler söz konusudur. Gemi sayılarının görüntülere göre dağılımları Şekil 4.1'de daha detaylı gösterilmektedir.

Tablo 4.3. Veri kümelerini oluşturan görüntülerin içerdikleri gemi sayılarına dair bilgiler

	Görüntü Sayısı	Görüntü Başına Düşen Gemi Sayısının Ortalaması	Görüntü Başına Düşen Gemi Sayısının Standart Sapması	Bir Görüntüde Yer Alan En Az Gemi Sayısı	Bir Görüntüde Yer Alan En Çok Gemi Sayısı
Eğitim Kümesi	6217	1.86	1.89	1	15
Validasyon Kümesi	1727	1.85	1.96	1	14
Test Kümesi	1770	1.80	1.87	1	15



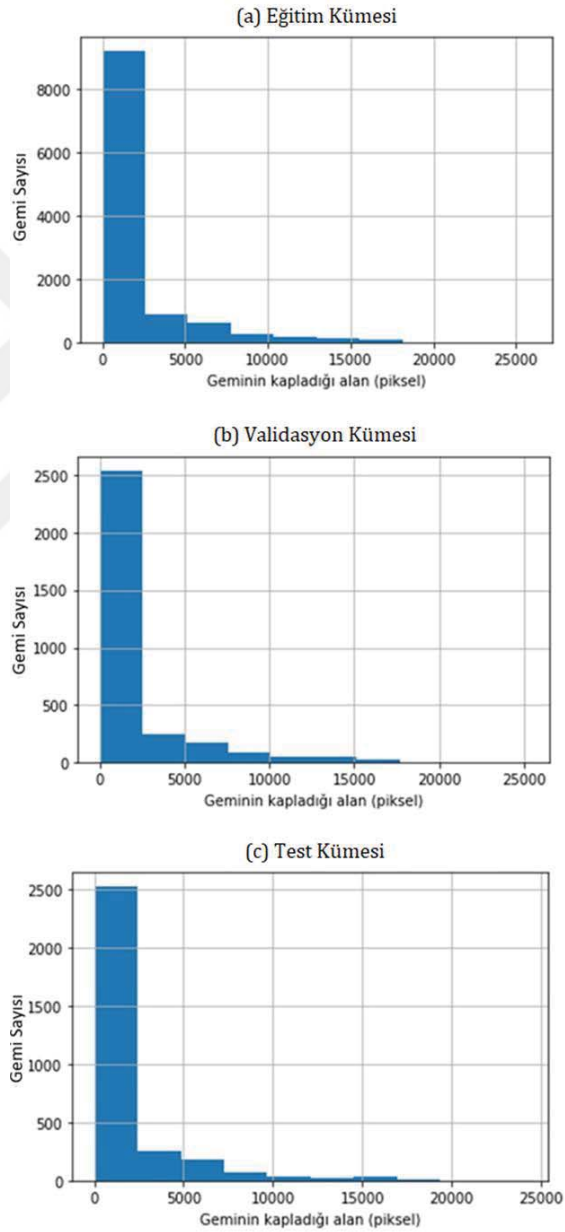
Şekil 4.1. Veri kümelerindeki gemi sayılarının dağılımı

Görüntülerde yer alan gemiler, uzunluğu 300 metreden fazla olan kargo gemilerinden, 200 metrenin üzerindeki uzunluklarıyla tanker gemilerine, ulaşımda kullanılan gemilerden, uzunluğu birkaç metreyi aşmayan küçük balıkçı teknelerine kadar geniş bir büyüklük yelpazesindedir. Gemilerin görüntü içerisinde kapladıkları alanlar da doğal olarak aynı çeşitliliktedir (Tablo 4.4 ve Şekil 4.2). Gemi uzunluklarına dair bilgiler de Tablo 4.5’te ve Şekil 4.3’te verilmiştir.



Tablo 4.4. Gemilerin görüntüler üzerinde kapladıkları alanlara dair istatistikler

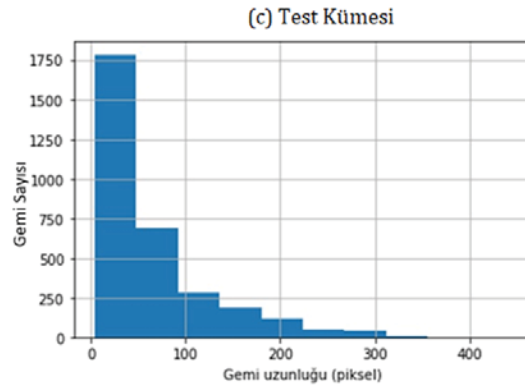
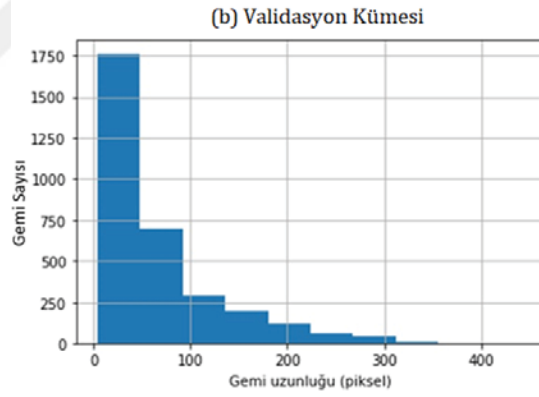
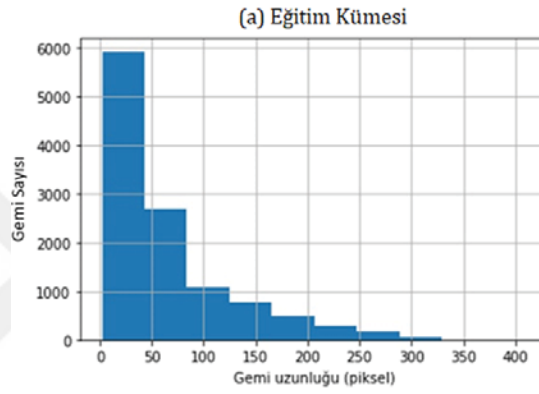
	Gemi Sayısı	Gemilerin Kapladığı Alanların Ortalaması (Piksel)	Gemilerin Kapladığı Alanların Standart Sapması (Piksel)	En Küçük Geminin Kapladığı Alan (Piksel)	En Büyük Geminin Kapladığı Alan (Piksel)
Eğitim Kümesi	11534	1953	3335	6	25904
Validasyon Kümesi	3197	1976	3401	8	25245
Test Kümesi	3180	1869	3249	6	24216



Şekil 4.2. Eğitim, validasyon ve test kümelerindeki gemilerin kapladıkları alanlara dair istatistikler

Tablo 4.5. Görüntülerdeki gemilerin uzunluklarına dair istatistikler

	Gemi Sayısı	Gemi Uzunluklarının Ortalaması (Piksel)	Gemi Uzunluklarının Standart Sapması (Piksel)	En Küçük Geminin Uzunluğu (Piksel)	En Büyük Geminin Uzunluğu (Piksel)
Eğitim Kümesi	11534	64.53	64.66	3	411
Validasyon Kümesi	3197	65.12	65.50	4	444
Test Kümesi	3180	63.31	64.53	3	424



Şekil 4.3. Eğitim, validasyon ve test kümelerindeki gemilerin uzunluklarına dair istatistikler

4.2. Eğitim

Bu çalışmadaki deneylerin Faster R-CNN modeli ile alakalı kısımlarında yhenon-GitHub (Henon, 2017), Mask R-CNN ile alakalı olanlarında Matterport-GitHub (Abdulla & Clauss, 2018) RetinaMask ile alakalı deneylerdeyse Fizyr-GitHub (Gaiser, Carpani, & Liscio, 2018) uyarlamaları kullanılmıştır. Bunlar Tensorflow ve Keras derin öğrenme çerçevelerinde çalışmaktadırlar. Modeller uçtan-uca eğitilmişlerdir. Önerilen model, takvimi Tablo 4.6’da verilen ve başlangıç değeri 0.001 olan öğrenme oranları kullanılarak toplamda 80 bin iterasyonda eğitilmiştir. Eğitimde 0.0001’lik bir ağırlık bozunumu (weight decay) ve 0.9’luk momentum değerleriyle momentumlu Stokastik Gradyan İniş yöntemi ve geri-yayılım algoritmaları kullanılmıştır.

Tablo 4.6. Eğitimde kullanılan öğrenme oranının turlara göre değişim takvimi

Tur Sayısı	Öğrenme Oranı
1-2	$lr \times 2$
3-4	lr
15-24	$lr/2$
25-34	$lr/4$
35-40	$lr/6$
41-50	$lr/8$
51-60	$lr/10$
61-70	$lr/12$
71-80	$lr/15$
81-90	$lr/20$
91-100	$lr/25$

$$lr = 0.001$$

RPN aşamasında, mini-yığınların eğitimi için toplamda 512 çapa örneklenmiş ve ikinci aşamaya 256 öneri gönderilmiştir. Pozitif örneklerin negatiflere oranı 1:3’tür. IoU eşiği olarak 0.5 kullanılmıştır. Deneyler kısmında görüleileceği gibi gemi güven eşiğinin belirlenmesi için de deneyler yapılmış ve deneylerin kalan kısmı için nesne güven eşiği 0.5 olarak belirlenmiştir.

Omurga ağları olarak ResNet-50 ve ResNet-101 kullanılmıştır. Ağ ağırlıkları, bu modellerin Common Objects in Context (COCO) veri seti kullanılarak ön-eğitim almış değerleriyle başlatılmıştır.

Ağların eğitilmesinde veri artırımına da başvurulmuştur. Bunun için farklı görüntü işleme yöntemlerinin sonuca etkisi gözlenmiş ve neticede orijinal görüntülere,

verilerin (1) 90 veya 270 derecelik döndürülmüş, (2) yukarı-aşağı veya sağ-sol döndürülmüş, (3) parlaklığı veya kontrastı yüzde 10 artırılmış veya azaltılmış ve (4) keskinliği veya bulanıklığı yüzde 10 artırılmış halleri de ilave edilmiştir. Veri artırımı, eğitim veri kümesindeki her örnek görüntüden yapay olarak değiştirilmiş yeni görüntüler oluşturma işlemidir. Bu teknik, eğitim verisi miktarının büyütülmesine ve veri sayısının az+ağdaki parametre sayısının yüksek olduğu durumlarda karşılaşılan aşırı-uyumu (over-fitting) önlemeye ve önerilen modelin performansını artırmayı sağlamaktadır.

Çapa büyüklükleri {8, 16, 32, 64, 128} ve adım sayıları {4, 8, 16, 32, 64} olarak tanımlanmıştır. Her büyüklükteki çapalar için üç farklı en boy oranı {1:1, 1:2, 2:1} belirlenmiştir.

Odak kayıp fonksiyonunun kullanıldığı model eğitiminde β parametresi için 0.01 değeri kullanılmıştır.

4.3. Tespit Değerlendirme Metrikleri

Önerilen modelin tespit performansını nicel olarak değerlendirmek için yaygın olarak kullanılan keskinlik, geri getirme, F1-skoru ve keskinlik-geri getirme eğrisi kullanılmıştır. Bu metrikler vasıtasıyla tespit modelinin performansı sadece doğru tespitlerle (True positives - TP) değil, üretilen yanlış alarmlar (False Positives - FP) ve gözden kaçan nesnelere (False Negatives - FN) de göz önüne alınarak hesaplanmaktadır.

Keskinlik, doğru tahminlerin toplam tahminler içerisindeki oranını (TP'lerin toplam pozitif tahminlere oranını) göstermektedir. Bu metrik, üretilen yanlış alarmların nicel olarak değerlendirilmesine yardımcı olmaktadır. Şu şekilde tanımlanmıştır:

$$Keskinlik = \frac{TP}{TP+FP} \quad (4.1)$$

Geri getirme ise yapılan tespitlerdeki doğru pozitiflerin toplam yer doğruluklarına oranıdır ve tespit doğruluğu olarak da geçmektedir. Tanımından da anlaşılacağı üzere bu metrik, tespiti hedeflenen nesnelere ne oranda belirlendiğini nicel olarak değerlendirmeye yardımcı olmaktadır. Tanımı şöyledir:

$$Geri\ getirme = \frac{TP}{TP+FN} \quad (4.2)$$

Bu iki niceliğin harmonik ortalaması olması hasabiyle F1 skoru, tespit performansına dair tek metrikte daha etraflı bir değerlendirme imkanı vermektedir ve şöyle tanımlanmıştır:

$$F1 = 2 \cdot \frac{\text{kesinlik} \cdot \text{geri getirme}}{\text{kesinlik} + \text{geri getirme}} = \frac{2TP}{2TP + FN + FP} \quad (4.3)$$

Bu çalışma kapsamında, modellerce üretilen herhangi bir tespit tahmini, eğer yer-doğruluğuyla olan IoU değeri 0.5'ten büyükse gerçek doğru (TP), aksi halde yanlış doğru/tespit (FP) olarak değerlendirilmiştir.



5. DENEYLER VE BULGULAR

Bu tez kapsamında gerçekleştirilen deneyler ve bulgular bu bölümde beş kısım halinde sunulmaktadır. Bölüm 5.1’de nesne sınıflandırması için belirlenen farklı güven eşiklerinin tespit performansı üzerindeki etkileri gösterilmektedir. Bölüm 5.2’de farklı derinliklerdeki omurga ağlarıyla gerçekleştirilen deneyler yer almaktadır. Önerilen yöntemin gemi tespiti üzerindeki performansı farklı modellerle Bölüm 5.3’te mukayeseli olarak sunulmaktadır. Bu kısımda, ayrıca, modelin farklı ölçeklerdeki görüntülerde ve farklı gemi boyutlarına göre performansının değerlendirilmesine olanak sağlayan deneyler yer almaktadır. Bölüm 5.4’te çeşitli veri artırım tekniklerinin kullanılmasıyla elde edilen sonuçlar sunulmaktadır. Bölüm 5.5, modelin ürettiği yanlış alarmları azaltmak üzere kullanımı önerilen iki farklı yöntemin tespit sonuçlarına etkisi üzerinedir. Bu bölümde, söz konusu yöntemlerin büyük ölçekteki görüntüler üzerinde çalıştırıldığı deneyler de yer almaktadır.

Deneyler, Ubuntu 16.04 işletim sisteminde Intel i7-6850k işlemcili, 64 GB RAM’e sahip, Nvidia GeForce GTX 1080 Ti ekran kartlı bir bilgisayar kullanılarak gerçekleştirilmiştir.

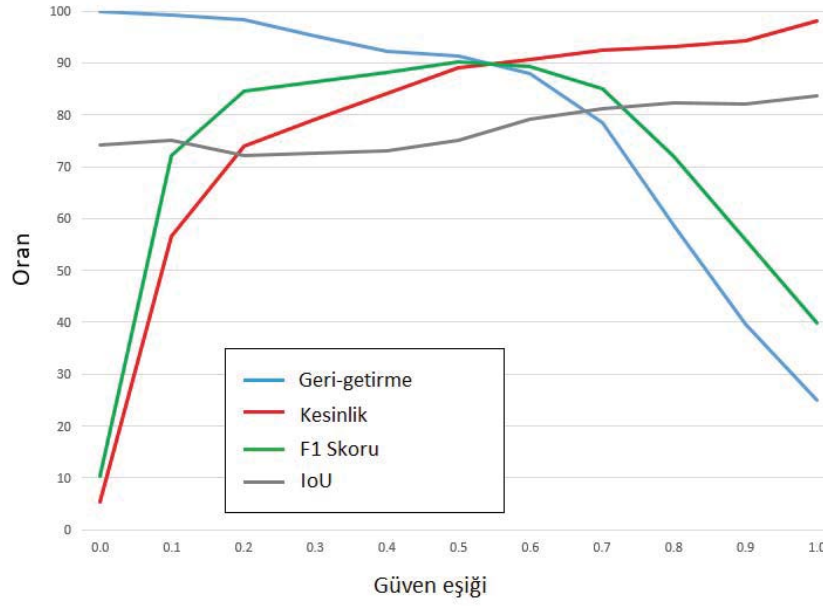
5.1. Nesne Güven Eşiğinin Tespit Üzerindeki Etkisi

Başlangıçta, önerilen modellerde nesne güven eşiği olarak hangi değer kullanılması gerektiğine karar vermek üzere deneyler yapılmıştır. Nesne güven eşiği, detayları Bölüm 3.1’de ifade edilen, modelin R-CNN katmanında RPN’den gelmiş önerinin “nesne midir?” veya “nesne değil midir?” sorusuna cevap veren sınıflandırma tahmin kafasında belirlenen nesnellik skor eşiğidir. Gelen öneri için bu kafanın ürettiği skor bu eşiğin altında olduğu takdirde öneri “arka plan” olarak görülmekte, üstünde olduğu takdirde “gemi” olarak atanmaktadır. Yapılan deneylerde bu eşik 0’dan başlatılmış ve 0.1’lik adımlarla artırılmıştır. Bu güven düzeyleri için modelin sağladığı tespit sonuçları Tablo 5.1’de sunulmuştur. Eşik düşük tutulduğunda gelen hemen her öneri “gemi” olarak değerlendirilmiş ve çok sayıda yanlış alarm üretilmiştir. Toplamda 3180 adet geminin yer aldığı test kümesindeki ikisi haricindeki tüm gemiler tespit edilmiş ancak bununla beraber toplamda 55332 yanlış alarm üretilmiştir. Eşiğin 0.1 puan artırılmasıyla beraber yanlış alarmlar keskin şekilde düşmüş bunun yanında toplam 25 yer-doğruluğu gözden kaçmıştır. Eşiği arttırdıkça bu trend devam etmektedir. Güven düzeyi 1.0 olarak belirlendiğinde yanlış alarmların sayısı 77 düşmüş, ancak yer-

doğruluklarının yarısından fazlası tespit edilememiştir. Çalışmanın bundan sonraki kısmında kullanmak üzere, optimal güven eşiği olarak bu tablodaki en yüksek F1-skorunu veren 0.5 değeri seçilmiştir. Eldeki veri seti için doğru tespitler ve yanlış alarmlar arasında en dengeli tespit sonuçlarını bu değerin verdiği düşünülmektedir. Bu değer altında tespitlerin yer-doğruluklarıyla çakışımında yüzde 75'lik bir IoU değeri söz konusudur. Farklı güven eşiklerindeki geri-getirme, kesinlik, F1 skoru ve IoU değerlerinin değişimi Şekil 5.1'de gösterilmiştir.

Tablo 5.1. Önerilen modelin (veri artırımı olmadan) farklı güven eşiği değerlerinde ürettiği gemi tespit sonuçları

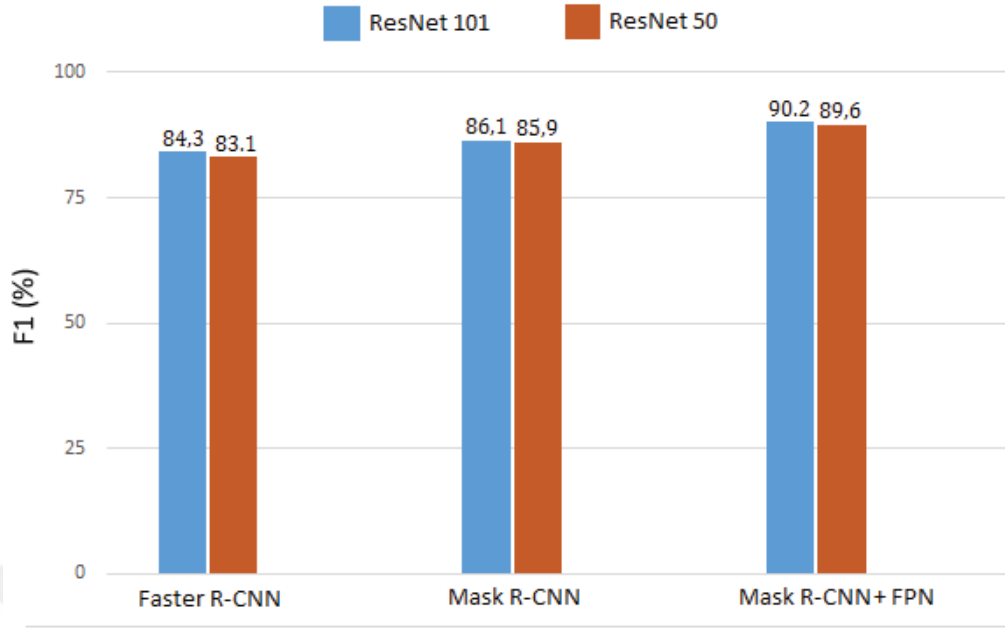
Güven Düzeyi	Yer Doğrulukları	Doğru Tespitler	Yanlış Alarmlar	Geri Getirme (%)	Kesinlik (%)	F1 (%)	IoU
0.0	3180	3178	55332	99,94	5,43	10,30	74,26
0.1	3180	3155	2418	99,21	56,61	72,09	75,18
0.2	3180	3129	1096	98,40	74,06	84,51	72,14
0.3	3180	3025	796	95,13	79,17	86,42	72,54
0.4	3180	2936	550	92,33	84,22	88,09	73,12
0.5	3180	2904	352	91,32	89,19	90,24	75,16
0.6	3180	2796	290	87,92	90,60	89,24	79,04
0.7	3180	2497	204	78,52	92,45	84,92	81,12
0.8	3180	1866	138	58,68	93,11	71,99	82,24
0.9	3180	1264	77	39,75	94,26	55,92	82,16
1.0	3180	794	16	24,97	98,02	39,80	83,78



Şekil 5.1. Önerilen modelin farklı güven eşiği değerlerinde ürettiği gemi tespit sonuçlarının grafiği

5.2. Farklı Derinliklerdeki Omurga Ağlarının Tespit Üzerindeki Etkisi

Önerilen yöntem ve karşılaştırmaya imkan tanımak için Faster R-CNN yöntemi farklı derinliklerdeki omurga ağları kullanılarak gemi veri kümesi üzerinde test edilmişlerdir. Omurga ağları olarak 50 ve 101 katmandan oluşan ResNet-50 ve ResNet-101 kullanılmıştır. Her iki omurga ağı için de Mask R-CNN yöntemi hem FPN'li hem de FPN'siz olarak Faster R-CNN yönteminden daha yüksek F1-skoru sağlamıştır (Şekil 5.2). Sonuçlardan görüleceği üzere 101-katmanlı ResNet-101 50-katmanlı muadiline göre daha yüksek F1-skorlu tespitler sağlamıştır. Bu sonuçlar göz önüne alınarak deneylerin bundan sonraki kısmında omurga ağı olarak ResNet-101 tercih edilmiştir. Bu modeller ve ağlar için görüntü başına harcanan işlem zamanları Tablo 5.2'de gösterilmektedir. Her üç model için de daha derin omurga ağ kullanımının daha fazla işlem zamanına mal olduğu görülmektedir. Ayrıca, Mask R-CNN'de FPN'nin kullanılması yaklaşık dörtte bir oranında bir yavaşlamaya neden olmuştur.



Şekil 5.2. Farklı omurgalarla çalıştırılan modellerin test kümesinde ürettikleri tespitlerin F1 skorları cinsinden karşılaştırmaları

Tablo 5.2. Çeşitli modeller için görüntü başına düşen hesaplama sürelerinin karşılaştırılması

Yöntem	Görüntü Başına İşlem Süresi (s)
Faster R-CNN (ResNet 50)	0.226
Faster R-CNN (ResNet 101)	0.238
Mask R-CNN (ResNet 50)	0.098
Mask R-CNN (ResNet 101)	0.105
Önerilen Yöntem (ResNet 50)	0.124
Önerilen Yöntem (ResNet 101)	0.131

5.3. Karşılaştırma: Önerilen Yöntemin Gemi Tespit Performansına Etkisi

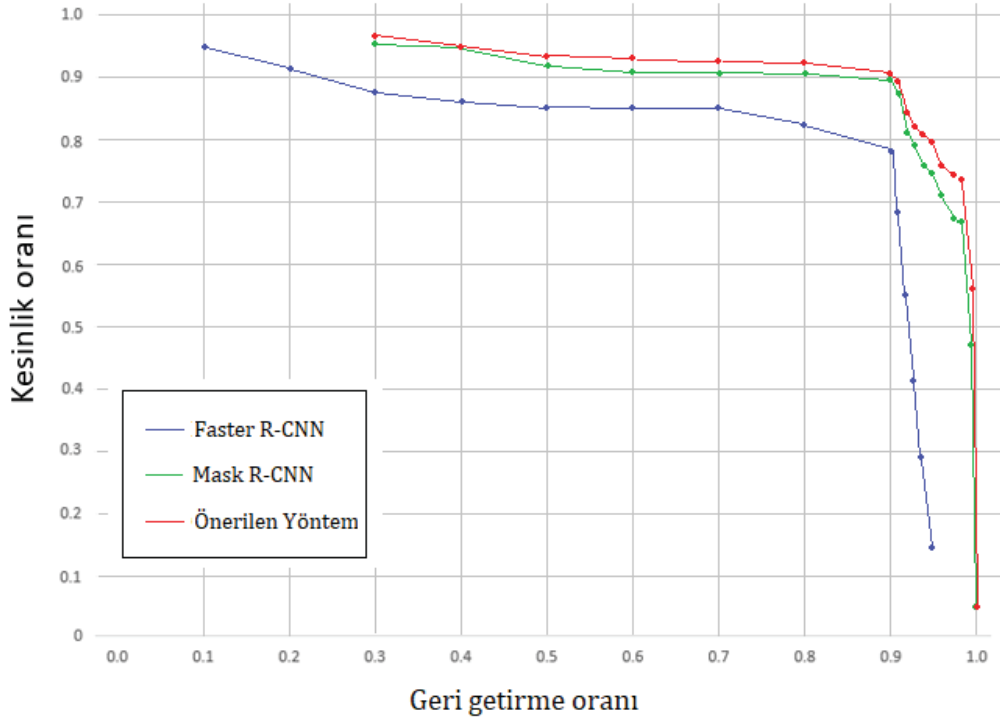
Bu bölüm iki başlığa ayrılmaktadır. İlk kısımda tek ölçekli test verileri üzerinde, önerilen modelle diğer yöntemler karşılaştırılmıştır. İkinci kısımda ise test verilerinin çözünürlükleri düşürülerek görüntüler kontrollü bir şekilde kabalaştırılmış ve söz konusu tüm yöntemlerin bu yeni küme üzerindeki performansları değerlendirilmiştir. Ayrıca tespitte nesne boyutlarının etkisi de gözlemlenmiştir.

5.3.1. Önerilen yöntemin diğer modellere göre performansı

Bu çalışma kapsamında önerilen gemi tespit modelinin performansı iki farklı modelin de aynı test verisinde kullanılmasıyla karşılaştırmalı olarak test edilmiştir. Bu karşılaştırma Bölüm 4.3'te ifade edilen değerlendirme ölçütleri kullanılarak yapılmıştır. 3180 adet yer-doğruluğunun (geminin) yer aldığı 1 metre mekansal çözünürlüğe sahip uydu görüntüleri için yapılan doğru tespitler, üretilen yanlış alarmlar ve gözden kaçan nesnelere kullanılarak geri-getirme, kesinlik ve F1-skorlarına bakılmıştır (Tablo 5.3). Ayrıca modellerin Kesinlik (precision -P)-Geri Getirme (Recall - R) (P-R) eğrileri de farklı nesne güven eşikleri kullanılarak oluşturulmuştur (Şekil 5.3).

Tablo 5.3. Önerilen model ile diğer modellerin test veri kümesi üzerindeki tespit performansının karşılaştırması

Yöntem	Özellik Piramit Ağı	ROI Hizalama	Yer Doğrulukları	Doğru Tespitler	Yanlış Alarmlar	Geri Getirme (%)	Kesinlik (%)	F1 (%)
Faster R-CNN (ResNet-101)	X	X	3180	2674	492	84,09	84,46	84,27
Faster R-CNN (ResNet-50)	X	X	3180	2591	464	81,48	84,81	83,11
Mask R-CNN (ResNet-101)	X	√	3180	2752	474	86,54	85,31	85,92
Mask R-CNN (ResNet-50)	X	√	3180	2765	480	86,95	85,21	86,07
Önerilen Yöntem(ResNet101)	√	√	3180	2904	352	91,32	89,19	90,24
Önerilen Yöntem (ResNet-50)	√	√	3180	2878	364	90,50	88,77	89,63



Şekil 5.3. Test veri kümesi üzerinde çalıştırılan farklı modellere ait kesinlik-geri getirme eğrileri

Üç model arasında en yüksek geri-getirme, kesinlik ve F1-skor değerleri, FPN'le çalıştırılan Mask R-CNN tarafından sağlanmıştır. Faster R-CNN'den farklı olarak bünyesinde ROI hizalama katmanı, maske tahmin kafası ve Özellik Piramit Ağı'nı içeren bu model yardımıyla aynı test kümesi için tespit F1-skoru üzerinde yaklaşık 6 puanlık (yaklaşık yüzde 7'lik) bir geliştirme sağlanmıştır.

Diğer taraftan, aynı modellerin farklı nesne güven eşikleriyle çalıştırılmasıyla yaklaşık 20 farklı kesinlik-geri getirme çifti elde edilmiştir. Şekil 5.3'te ResNet-101 omurgası kullanan önerilen modelin diğer yöntemlerle olan karşılaştırması görülmektedir. Bu metrik, tespit modelinin etkinlik ve stabilitesini gösteren önemli bir değerlendirme aracıdır. Y eksenini kesinliği ve X eksenini geri-getirmeyi temsil etmektedir. Eğrinin üstte olması daha iyi performansa işaret etmektedir. Önerilen FPN+Mask R-CNN modeli deneylerimizde kullandığımız tüm güven eşiklerinde diğer yaklaşımlardan iyi sonuçlar vermiştir.

5.3.2. Görüntü çözünürlüğünün tespit performansına etkisi

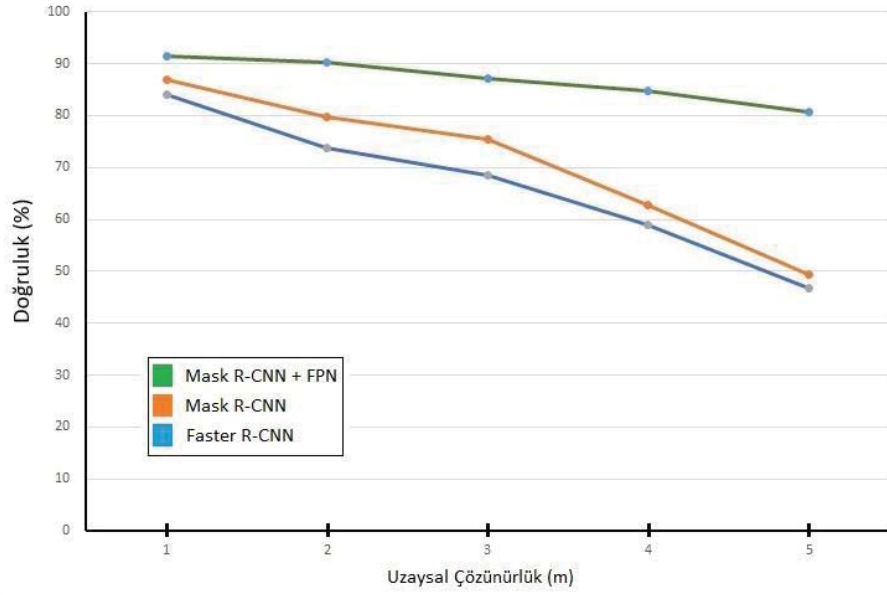
Görüntü çözünürlüğünün önerilen yöntemin performansı üzerindeki etkisini incelemek üzere, yer örnekleme aralığı 1.0 m olan test verileri farklı çözünürlüklere

indirgenerek kapsamlı deneyler yapılmıştır. Lanczos filtrelemesi (yeniden örnekleme) yöntemi kullanılarak test veri kümesindeki her görüntünün 4 farklı çözünürlükte (*yer örnekleme aralığı* = 2.0m, 3.0m, 4.0m, 5.0m) yeni örnekleri oluşturulmuştur. Deneyler, bu görüntülerin modellere beslenmesiyle gerçekleştirilmiştir.

Tablo 5.4’de üç farklı modelin 5 ayrı çözünürlükte ortaya koydukları performans, geri-getirme, kesinlik ve F1 skorları cinsinden verilmiştir. Sonuçlar, önerilen FPN+Mask R-CNN tespit yönteminin farklı çözünürlüklere karşı sahip olduğu stabiliteyi göstermektedir. Çözünürlüğün iki uç arasında 5 kat oynadığı bu aralıkta, geri-getirme değerleri açısından bakıldığında, diğer iki yöntem nazaran önerilen yöntemin performansını çok az kaybettiği görülebilir (Şekil 5.4). Önerilen model için en yüksek geri-getirme %91.32’lik değerle orijinal çözünürlükteyken, en düşük değeri %80.57’lik geri-getirmeyle 5 metre çözünürlükte vermiştir. Diğer yöntemlere bakıldığında, FPN’nin kullanılmadığı maske ve Faster R-CNN yaklaşımlarında ise geri-getirme performansı 37 puandan fazla düşmüştür. Bununla beraber, yanlış alarmlarla bağlantılı olan kesinlik metriğinin aynı oranda korunmadığı da sonuçlarda görülebilir. Çözünürlükteki iki uç arası için önerilen modelde kesinlik 29 puan düşerken diğer yöntemlerde bu metriktaki düşüş 40 puandan fazla olmuştur.

Tablo 5.4. Test veri kümesi (*orijinal görüntü çözünürlüğü = 1m*) kullanılarak *yer örnekleme aralıkları kabalaştırılmış veriler üzerinde üç farklı modelin ürettiği tespit neticeleri*

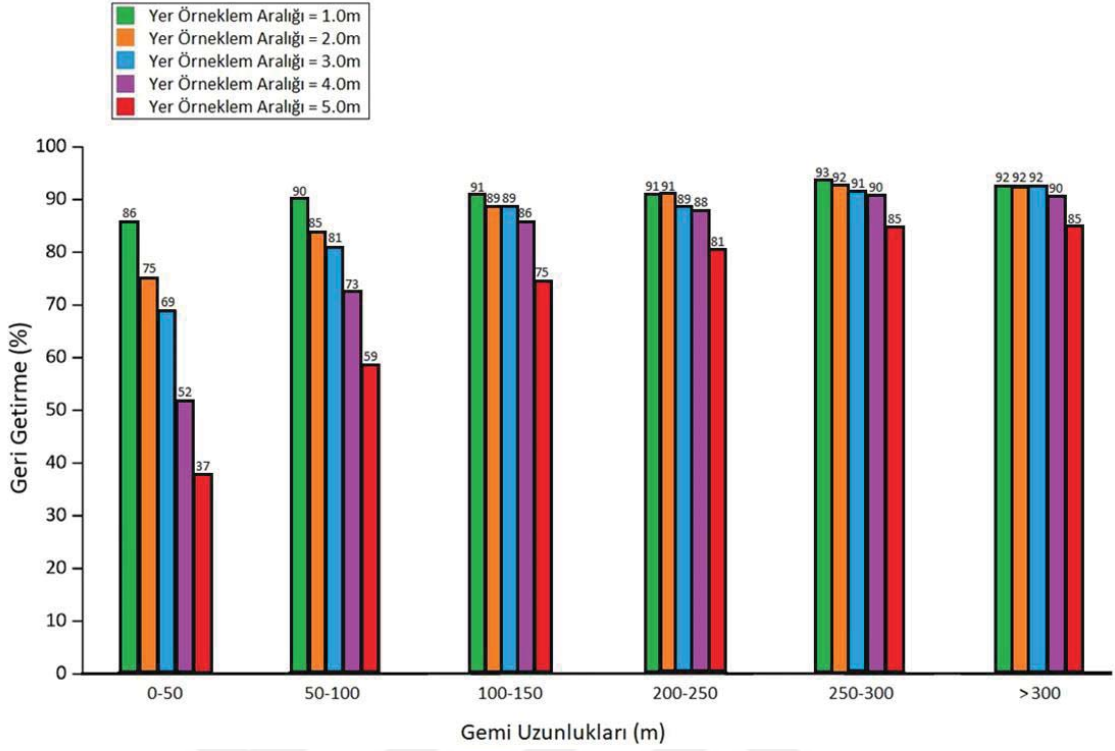
	Yer Örnekleme Aralığı = 1.0m			Yer Örnekleme Aralığı = 2.0m			Yer Örnekleme Aralığı = 3.0m			Yer Örnekleme Aralığı = 4.0m			Yer Örnekleme Aralığı = 5.0m		
	Faster R-CNN	Mask R-CNN	Mask R-CNN + FPN	Faster R-CNN	Mask R-CNN	Mask R-CNN + FPN	Faster R-CNN	Mask R-CNN	Mask R-CNN + FPN	Faster R-CNN	Mask R-CNN	Mask R-CNN + FPN	Faster R-CNN	Mask R-CNN	Mask R-CNN + FPN
Geri getirme	84,09	86,95	91,32	73,65	79,75	90,25	68,46	75,35	87,11	58,84	62,70	84,72	46,73	49,34	80,57
Kesinlik	84,46	85,21	89,19	64,38	65,84	77,86	58,70	60,54	64,90	52,37	53,80	62,90	43,17	45,02	60,06
F1 Skoru	84,27	86,07	90,24	68,70	72,13	83,60	63,21	67,14	74,38	55,42	57,91	72,20	44,88	47,08	68,82



Şekil 5.4. Farklı görüntü çözünürlükleri için üç modelin geri getirme (doğruluk) performansı

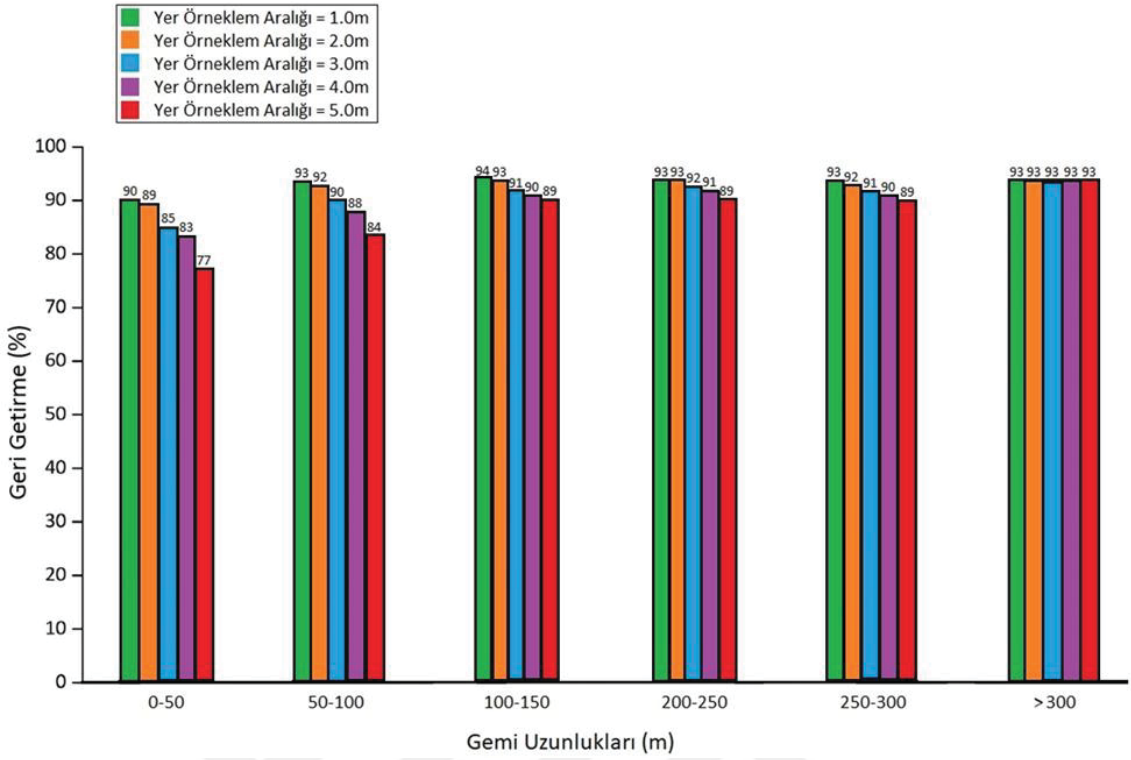
FPN'nin kullanılmadığı modellerde azalan çözünürlükle beraber doğruluğun keskin bir şekilde düşmesi dikkat çekicidir. Mask R-CNN modelinde FPN kullanımı bu düşüş oranını azaltmaktadır. Bu gözlemi daha detaylı incelemek ve hata sebeplerine dair daha elle tutulur doneler elde edebilmek üzere görüntülerde yer alan gemiler uzunluklarına göre 6 gruba ayrılmış ve her grup için geri-getirme değerleri hem FPN'siz çalışan Mask R-CNN modeli hem de FPN'le çalışan Mask R-CNN modeli kullanılarak hesaplanmıştır.

Şekil 5.5, FPN'nin kullanılmadığı Mask R-CNN modelinin her uzunluk grubundaki gemiler için farklı mekansal çözünürlüklerde ürettiği geri-getirmeleri göstermektedir. Grafikte, azalan çözünürlükle, uzunlukları 100 metreden az olan gemiler için geri-getirme oranının keskin bir şekilde düştüğü görülmektedir. Orijinal görüntüler içerisinde sadece birkaç piksele sahip gemilerin olduğu düşünüldüğünde, bu gemilerin görsel özellikleri önce çözünürlüğün düşürülmesi arkasından ağ içerisinde birçok katmandan geçmeyle kaybolmakta ve neticesinde düşük geri-çağırma oranıyla sonuçlanmaktadır. Bu çözünürlük aralığında, uzunluğu 250 metrenin üzerinde olan gemilerdeyse tespit performansındaki oynamaların giderek azaldığı görülmektedir.



Şekil 5.5. FPN kullanılmadığı durumda, farklı görüntü çözünürlükleri için, Mask R-CNN modelinin verdiği, farklı gemi uzunluklarına göre ayrılmış geri getirme değerleri.

Mask R-CNN modelinin FPN'le beraber kullanıldığı durumdaysa tüm boyut skalalarında geri-getirmeler iyileşmekle beraber, özellikle uzunlukları 200 metrenin altında olan gemilerin tespitinde büyük iyileşmeler sağlanmaktadır (Şekil 5.6). 5 metre çözünürlüğündeki görüntülerde uzunlukları 50 metreye kadar olan gemiler için geri-getirme 40 puan, 100 metreye kadar olanlar için 25 puan ve 150 metreye kadar olanlarda 14 puan artmıştır. Bu çözünürlük için geri-getirmedeki toplam artış 30 puanın üzerinde olmaktadır (%80.57'ye karşılık %49.34). Bu durum, Özellik Piramit ağındaki semantik değeri yüksek katmanla mekansal çözünürlüğü fazla olan alt katmanın kaynaştırılmasının sonuçlar üzerindeki olumlu etkisini göstermektedir.



Şekil 5.6. Mask R-CNN modelinin FPN ile beraber kullanıldığı durumda, farklı görüntü çözünürlükleri için, farklı gemi uzunluklarına göre ayrılmış geri getirme değerleri.

5.4. Farklı veri artırım yöntemleriyle eğitim: önerilen yöntemin performansı

Bu bölümde eğitim verilerindeki görüntülere farklı yollarla uygulanan dönüşümler neticesinde oluşturulan artırılmış veri setleriyle eğitilen önerilen modelin tespit performansı incelenmiştir. Deneyleerde dört temel görüntü işleme yöntemi ve bunların çeşitli kombinasyonları kullanılarak toplam 9 adet artırılmış veri seti oluşturulmuştur (Tablo 5.5). Deneyleer aynı koşullar altında gerçekleştirilmiştir. Tablo 5.6’da farklı artırma yöntemleriyle eğitilmiş modelin tespit sonuçları sunulmaktadır. Buradan görüldüğü üzere uygulanan her artırım yöntemi sonuçlarda iyileşmeye yol açmıştır. Sonuçlardaki en yüksek iyileştirme, modelin dört görüntü dönüşümünün beraber uygulanarak oluşturulduğu AVS9 setiyle aldığı eğitimle sağlanmıştır. Eğitimde sadece ham görüntülerin kullanıldığı model, 3180 geminin yer aldığı test verisi için 2904 doğru tespitte bulunurken AV9 setiyle eğitilen model 3024 gemiyi doğru tespit etmiştir. Yanlış alarm sayısı da 352’den 279’a düşmüştür.

Tablo 5.5. Artırılmış veri setleri ve bunlarda görüntülere uygulanan işlemler

Artırılmış Veri Seti	Görüntüye Uygulanan İşlem
AVS 1	90 veya 270 derecelik döndürmeden biri
AVS 2	Yukarı-aşağı veya sağ-sol döndürme
AVS 3	Parlaklık veya kontrastın yüzde 10 artırılması veya azaltılmasından biri
AVS 4	Keskinliğin veya bulanıklığın yüzde 10 artırılmasından biri
AVS 5	AVS 1 + AVS 2
AVS 6	AVS 2 + AVS 3
AVS 7	AVS 2 + AVS 4
AVS 8	AVS 2 + AVS 3 + AVS 4
AVS 9	AVS 1 + AVS 2 + AVS 3 + AVS 4

Tablo 5.6. Farklı artırma yöntemleriyle oluşturulmuş veri kümelerinin kullanılmasıyla eğitilmiş olan önerilen modelin test kümesi üzerindeki tespit sonuçları

Veri Seti	Test Kümesindeki Toplam Gemi Sayısı	Doğru Tespitler	Yanlış Alarmlar	Geri Getirme (%)	Keskinlik (%)	F1 (%)
Orijinal	3180	2904	352	91,32	89,19	90,24
AVS 1	3180	2988	316	93,96	90,44	92,17
AVS 2	3180	2986	324	93,90	90,21	92,02
AVS 3	3180	2971	336	93,43	89,84	91,60
AVS 4	3180	2964	344	93,21	89,60	91,37
AVS 5	3180	2984	322	93,84	90,26	92,01
AVS 6	3180	3002	296	94,40	91,02	92,68
AVS 7	3180	2996	304	94,21	90,79	92,47
AVS 8	3180	3006	301	94,53	90,90	92,68
AVS 9	3180	3024	279	95,09	91,55	93,29

5.5. Hatalı Tespitlerin Azaltılmasına Dair Deneyler

Bu bölümde modelin ürettiği yanlış alarmları azaltmak için önerilen iki farklı yaklaşıma dair deneyler aktarılacaktır.

5.5.1. Mask R-CNN-tabanlı gemi tespiti için negatif örneklem eğitimi

Çalışmanın bu kısmında, model eğitiminde farklı sayılarda negatif örneklem kullanmanın tespit sonuçları üzerindeki etkileri incelenmektedir.

Pozitif örneklerin negatiflere oranının tespit performansına etkisini görmek amacıyla FPN+Mask R-CNN modeli pozitif örnekler aynı kalmak kaydıyla farklı sayılardaki negatif örnekler kullanılarak eğitime tabi tutulmuştur (Tablo 5.8). Pozitif ve negatif görüntü örnekleri sırasıyla Şekil 3.12’de ve Şekil 3.13’de verildiği gibidir. Dört veri kümesi için de modelin eğitimine kayıp fonksiyonunun büyük oranda düşmesine kadar devam edilmiş ve nihayetinde tespit model ağırlıkları olarak en düşük validasyon hatasına sahip tur ağırlıkları kullanılmıştır. Eğitimle alakalı bilgiler Tablo 5.7’de özetlenmektedir.

Tablo 5.7. Eğitim ve validasyon verilerine eklenen negatif örneklem görüntü sayılarına ve eğitime dair bilgiler

Pozitif Örneklem Görüntülerin Negatiflere Oranı	Negatif Örneklem Görüntülerin Sayısı	Eğitim Adım Sayısı	En Düşük Validasyon Kaybı	Eğitimde Harcanan Zaman (saat)
1:0	0	87400	1.073	11.5
2:1	3972	114000	0.972	15
1:1	7944	190000	1.044	24
1:2	15888	228000	1.120	29

Önerilen negatif örnekle eğitim yönteminin karasal bölgelerde karşılaşılan yanlış alarmları azaltıp azaltmadığını test etmek amacıyla test verisi olarak farklı türde deniz-kara içeriklerine sahip geniş ölçekli, 1.0 metre çözünürlüğe sahip 3 ayrı Google Earth görüntüsü kullanılmıştır (Şekil 5.7). Bu görüntüler,

- karasal alanın büyük yer kapladığı iç suları,
- kara ve denizin dengeli yer aldığı liman/rıhtım yakınlarını,
- karasal alan içermeyen açık suları

içericek şekilde seçilmiştir. Bu şekilde, karasal alanlardan kaynaklanan yanlış alarmların analizinin daha detaylı yapılması mümkün hale gelmiştir. Görüntülere dair bilgiler Tablo 5.8’de verilmiştir.

Tablo 5.8. Test amacıyla kullanılan büyük boyutlu görüntülere dair bilgiler

Görüntü Adı	Görüntü Boyutu (piksel)	Görüntü Kaynağı	Mekansal Çözünürlük	Gemi Sayısı	Görüntü Konumu
BG-1	13632 x 11456	Google Earth	1m	319	Rotterdam, Hollanda (51° 54' K, 4° 11' D)
BG-2	5568 x 4480	Google Earth	1m	288	Victoria Limanı, Hong Kong (22° 18' K, 114° 07' D)
BG-3	11776 x 4032	Google Earth	1m	138	Singapur Limanı, Singapur (1° 15' K, 103° 53' D)



Şekil 5.7. Yanlış alarmların azaltılmasına dönük gerçekleştirilen deneylerde kullanılan, farklı kara-deniz içeriklerine sahip büyük ölçekli görüntüler.

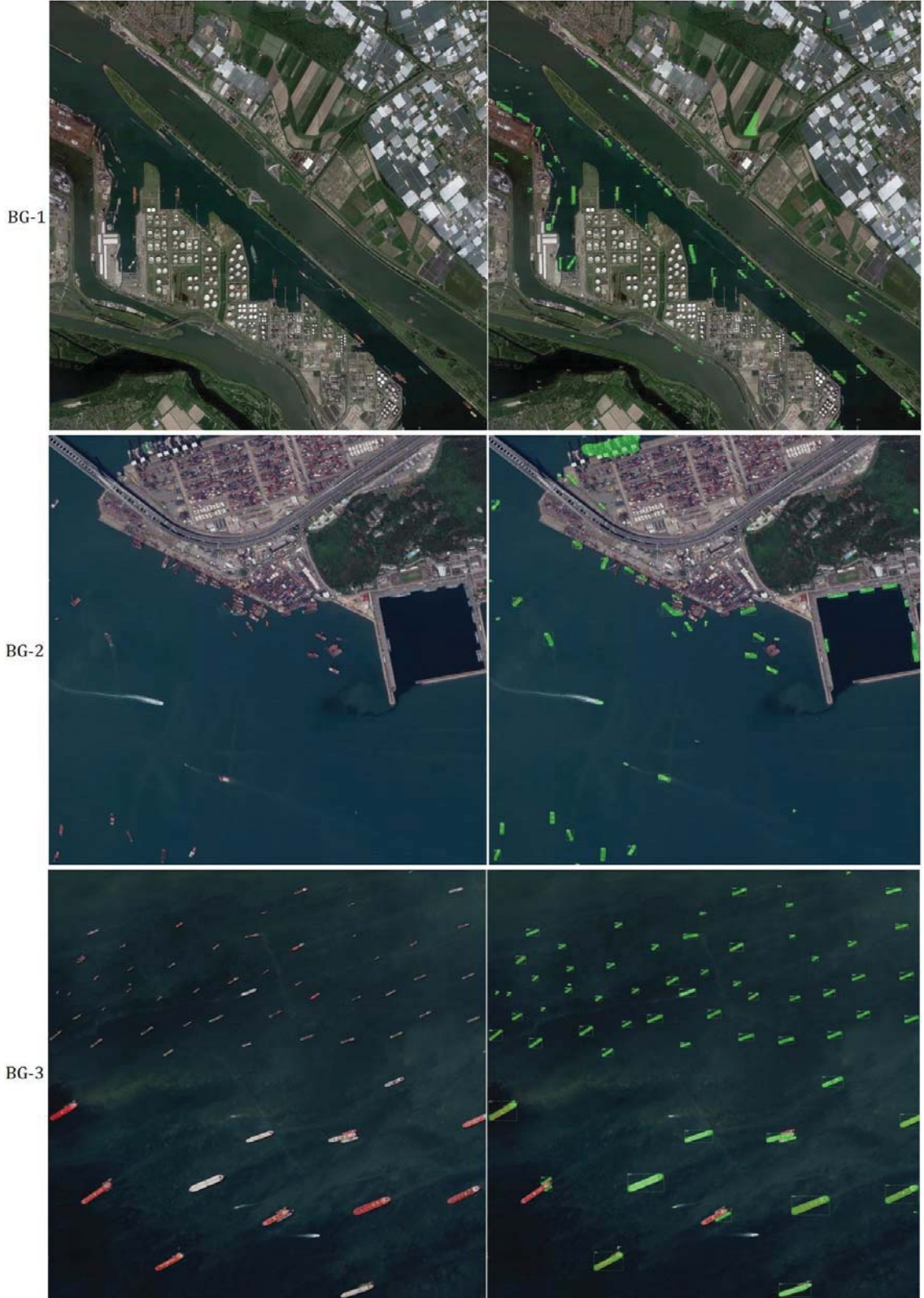
Bu çalışmada kullanılan grafik işlemcinin hafızası, önerilen tespit modelinde boyutu en fazla 5120x5120 piksele kadar olan görüntülerle çalışmaya imkan vermektedir. Bu sebeple boyutu bundan büyük görüntülerde gemileri tespit ederken görüntünün

bloklar halinde çıkarılıp ayrı ayrı işlenmesi gerekmektedir. Deneylerimizde, büyük görüntüler blok büyüklükleri 5120×5120 piksel olan parçalara ayrılıp işlenmişlerdir. Bunu yaparken gemilerin bloklar arasındaki sınırlarda gözden kaçmasını önlemek için her blok, komşusuyla 200 piksellik kesişime sahip olacak şekilde oluşturulmuştur. Resim dışına çıkan blok parçalarındaki piksellere siyah renk atanmıştır (Şekil 5.8).



Şekil 5.8. Boyutu 5120×5120 pikselin üzerinde olan görüntüler bu büyüklükteki parçalara bölünerek modele girilmektedir.

Üç büyük ölçekli görüntüden alınan kesitler için deney sonuçları Şekil 5.9'da gösterilmiştir.



Şekil 5.9. Sol sütun: Orijinal görüntüler. Sağ sütun: Sadece pozitif örneklerle eğitilmiş ağ tarafından üretilen tespit sonuçları

Diğer deneylerde olduğu gibi bu kısımda da gemi tespit performansı geri-getirme, kesinlik ve F1-skorlarının ölçülmesiyle hesaplanmıştır. Ayrıca, işlem zamanları da ölçülmüştür. Tablo 5.9 tespit performanslarını ve işlem sırasında harcanan zamanları göstermektedir.

(1) **BG-1:** İç suları içeren, karasal alanların ağırlıkta olduğu 13632×11456 piksel boyutundaki görüntü.

BG-1 görüntüsü Hollanda'nın Rotterdam şehrinde bulunan, denizin kara içlerine sokularak rıhtımlara ulaşımı sağladığı bir içeriği göstermektedir. Diğer görüntülerle karşılaştırıldığında modelin tespit performansı göreceli olarak kötüdür. Görüntüde karasal bölgeler ağırlıkta olduğundan negatif örnekleme eğitim yöntemi karasal yanlış alarmları azaltmaya yardımcı olmaktadır. Tablo 5.9 ve Şekil 5.10'da görüldüğü gibi negatif örneklemin kullanılmasıyla eğitim almış Mask R-CNN tabanlı gemi tespit modeli, sadece pozitif örneklerle eğitilen ağa kıyasla yanlış alarmları büyük ölçüde azaltmıştır.

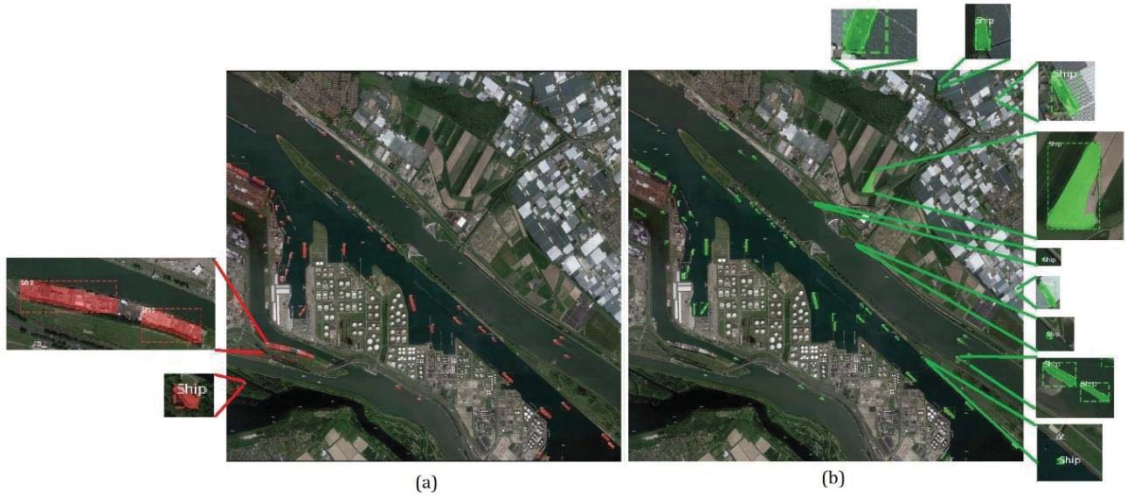
Uygulanan farklı pozitif/negatif örnekleme oranları arasında en yüksek F1-skoru her iki sınıf için de eşit sayıda görüntünün kullanıldığı 1:1 oranında elde edilmiştir. Ancak en az yanlış alarm $P/N = 1:2$ kullanıldığında üretilmiş olup tespitin kesinlik değeri en yüksek bu oranda alınmıştır. İlk durumda toplam 53 yanlış alarm üreten model, pozitif/negatif örnekleme oranı 1:2 olan veriyle eğitildiğinde bu sayıyı 22'ye düşürmüştür. Bu P/N oranında hatadaki azalma miktarı yüzde 65.6'dır. Ayrıca, bu yolla sadece hatalı tespitler azaltılmamış, yapılan doğru tespitler de modelin genelleştirme kabiliyetinin artmasıyla 274'ten 293'e (yüzde 6.9) yükselmiştir. Şekil 5.11'deki BG-1 görüntüsünden alınan kesitte de kara kökenli 9 yanlış alarmın 3'e düştüğü görülmektedir.

Tablo 5.9. BG-1 görüntüsü için farklı negatif örnekleme oranlarında gemi tespit sonuçları ve harcanan işlem zamanı

Pozitif/Negatif Örneklem	Zaman (s)	Yer Doğrulukları	Doğru Tespitler	Gözden Kaçan Gemiler	Yanlış Alarmlar	Geri getirme	Kesinlik	F1 Skoru
1:0	255	319	274	45	64	0,86	0,81	0,83
2:1	253	319	287	32	30	0,90	0,91	0,90
1:1	260	319	293	26	25	0,92	0,92	0,92
1:2	254	319	281	38	22	0,88	0,93	0,90



Şekil 5.10. BG-1 görüntüsü içinden bir kesitte a) eğitimde pozitif örnek sayısı kadar negatif örneğin kullanıldığı modelle yapılan tespit, b) eğitimde sadece pozitif örneklerin kullanıldığı modelle yapılan tespit



Şekil 5.11. BG-1 görüntüsü içinden bir kesitte, model eğitiminde pozitif/negatif örnek oranı olarak a) $P/N=1:1$, b) $P/N=1:0$ kullanıldığında üretilen yanlış alarmlar

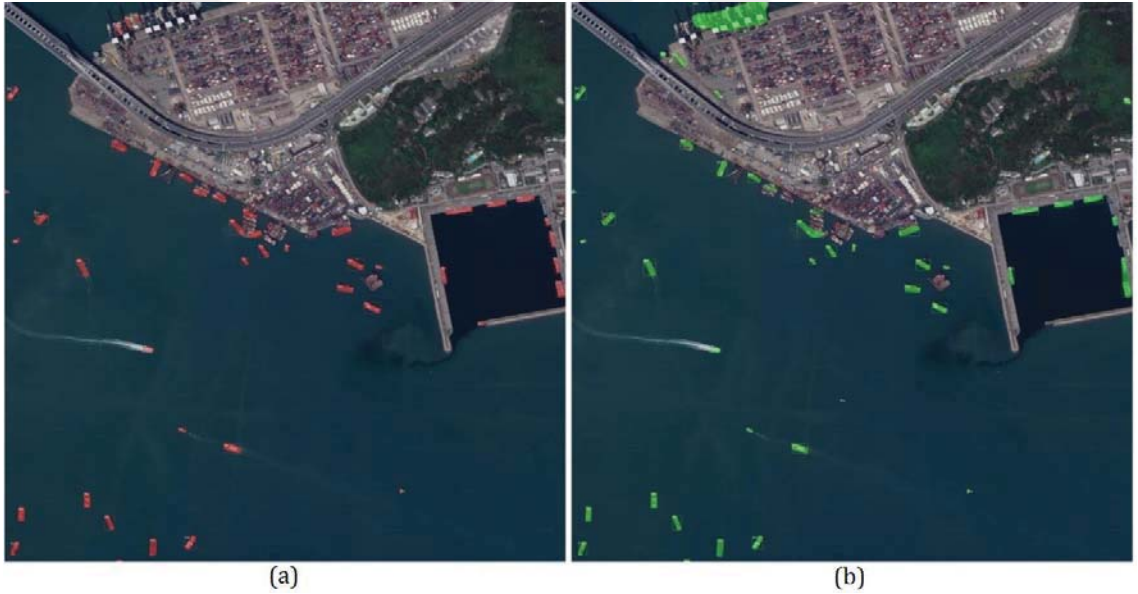
(2) BG-2: Hong Kong'da Victoria Limanı civarını içeren 5568×4480 piksel boyutundaki görüntü.

Bu görüntüde deniz ve kara arasındaki ayırım açık vaziyettedir. Açıkta bulunan gemilerin yanı sıra karaya yanaşmış gemiler de sahnede yer almaktadır. Karasal bölge ve deniz dengeli orandadır. İlk görüntüde olduğu gibi burada da, negatif örnekle eğitim tespitinin kesinlik performansında yaklaşık 10 puanlık bir iyileştirme sağlamış durumdadır (Tablo 5.10 ve Şekil 5.12). Yalnızca pozitif örneklerle eğitilen model toplamda 53 yanlış

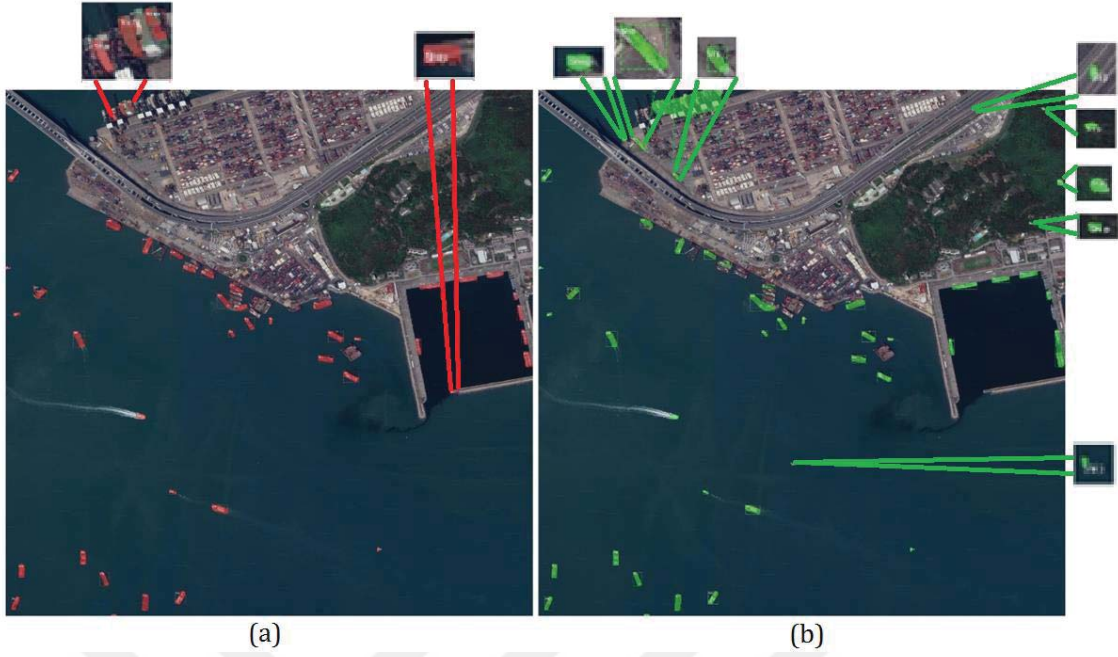
alarına yol açmışken, P/N = 1:1 oranında negatif örnek kullanıldığında yaklaşık yüzde 58.5’luk bir iyileşmeyle bu sayı 22’ye düşmüş aynı zamanda doğru tespitler üzerinde de 29 adetlik (yüzde 12.3’lük) bir kazanım sağlanmıştır. Şekil 5.13’teki BG-2 görüntüsünden alınan kesitte başlangıçtaki kara kökenli 8 hatanın 4’e düştüğü görülmektedir.

Tablo 5.10. BG-2 görüntüsü için farklı pozitif/negatif örneklem oranlarında gemi tespit sonuçları ve harcanan işlem zamanı

Pozitif/Negatif Örneklem	Zaman (s)	Yer Doğrulukları	Doğru Tespitler	Gözden Kaçan Gemiler	Yanlış Alarmlar	Geri getirme	Kesinlik	F1 Skoru
1:0	50	288	235	53	53	0,82	0,82	0,82
2:1	51	288	254	34	30	0,88	0,89	0,89
1:1	53	288	264	24	22	0,92	0,92	0,92
1:2	53	288	255	33	22	0,89	0,92	0,90



Şekil 5.12. BG-2 görüntüsü içinden bir kesitte a) eğitimde pozitif örneklem sayısı kadar negatif örneklemin kullanıldığı modelle yapılan tespit, b) eğitimde sadece pozitif örneklerin kullanıldığı modelle yapılan tespit



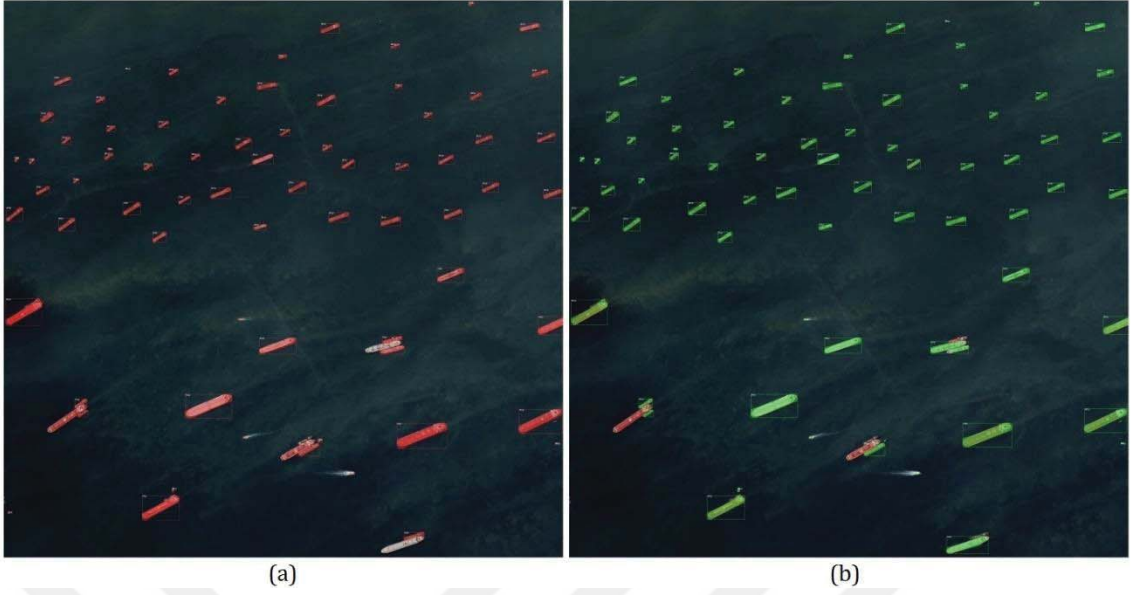
Şekil 5.13. BG-2 görüntüsü içinden bir kesitte, model eğitiminde pozitif/negatif örneklem oranı olarak a) $P/N=1:1$, b) $P/N=1:0$ kullanıldığında üretilen yanlış alarmlar

(3) **BG-3:** Singapur Limanı açıklarını içeren 11776×4032 piksel boyutundaki görüntü.

Herhangi bir karasal alan içermeyen BG-3 görüntüsünde çeşitli boyutlardaki gemiler yer almaktadır. Sahnede toplam 138 gemi bulunmaktadır. Yalnızca pozitif örneklerle aldığı eğitimle model iki yanlış alarm üretmektedir ve bunlar denizde yer alan, boyutları çok küçük nesnelere. Eğitimde farklı oranlarda negatif örnekler kullanıldığında yanlış alarm sayısı en fazla 1'e düşmüştür. Negatif örnekle eğitimin doğru tespitler üzerindeki etkisi de sınırlı olmuştur (122'den 128'e : Yüzde 4.7). Deneysel sonuçlara bakıldığında, karanın yer almadığı bu görüntü için eğitime negatif örnek eklemenin sonuçlar üzerine etkisinin az olduğu görülmektedir (Tablo 5.11).

Tablo 5.11. BG-3 görüntüsü için farklı negatif örneklem oranlarında gemi tespit sonuçları ve harcanan işlem zamanı

Pozitif/Negatif Örneklem	Zaman (s)	Yer Doğrulukları	Doğru Tespitler	Gözden Kaçan Gemiler	Yanlış Alarmlar	Geri getirme	Kesinlik	F1 Skoru
1:0	100	138	122	16	2	0,88	0,98	0,93
2:1	98	138	126	12	1	0,91	0,99	0,95
1:1	99	138	128	10	1	0,93	0,99	0,96
1:2	102	138	125	13	1	0,91	0,99	0,95



Şekil 5.14. BG-3 görüntüsü içinden bir kesitte a) eğitimde pozitif örneklem sayısı kadar negatif örneklemin kullanıldığı modelle yapılan tespit, b) eğitimde sadece pozitif örneklerin kullanıldığı modelle yapılan tespit

5.5.2. Tespitte odak kayıp fonksiyonunun kullanımı: RetinaNet

Bu kısımda bir önceki bölümde yapılanın aksine eğitim veri kümesinde herhangi bir değişiklik yapmaya gidilmeden, modelin kayıp fonksiyonunu odak kayıp fonksiyonuyla değiştirerek tespit sonuçlarının nasıl etkilendiği incelenmektedir. Bu amaçla deneyler, kayıp fonksiyonu için ağırlıklandırma faktörü sabit tutularak ($\alpha=0.25$) farklı odaklama parametrelerinde ($\gamma= 1.0$, $\gamma= 2.0$, $\gamma= 5.0$) gerçekleştirilmiş ve odak kayıp fonksiyonunun kullanılmadığı durumla ($\alpha=1.0$, $\gamma=0$) karşılaştırması yapılmıştır. RetinaNet'in kullanıldığı model için eğitim süresi kullanılan odaklama değerine göre 1.5-2.5 saat arasında uzamıştır.

BG-1 görüntüsüyle gerçekleştirilen ilk deneyde tüm değerlendirme metriklerinde en iyi sonuçlar, odaklandırma parametresi için $\gamma= 2.0$ değeri kullanıldığında elde edilmiştir. Bu değerde geri-getirme 7 puan, kesinlik 11 puan yükseliş göstermiştir, bu değerlerin harmonik ortalaması olan F-1 skorunda da 9 puanlık bir artış olmuştur (Tablo 5.12).

Tablo 5.12. BG-1 görüntüsü için farklı γ değerlerindeki tespit sonuçları ve model eğitiminde harcanan zaman

γ ($\alpha=0.25$)	Eğitimde Harcanan Zaman (saat)	Geri getirme	Kesinlik	F1 Skoru
0 ($\alpha=1.0$)	11.5	0,86	0,81	0,83
1.0	14	0,88	0,87	0,87
2.0	14	0,93	0,92	0,92
5.0	13	0,87	0,85	0,86

BG-2 görüntüsüyle gerçekleştirilen deneylerde de en iyi sonuçlar, odaklandırma parametresi için $\gamma = 2.0$ değeri kullanıldığında elde edilmiştir. Sadece pozitif görüntülerle eğitilmiş, kayıpta odaklamanın kullanılmadığı duruma göre geri-getirme 7 puan, kesinlik 12 puan, F1-skoru 9 puan artmıştır (Tablo 5.13).

Tablo 5.13. BG-2 görüntüsü için farklı γ değerlerindeki tespit sonuçları ve model eğitiminde harcanan zaman

γ ($\alpha=0.25$)	Eğitimde Harcanan Zaman (saat)	Geri getirme	Kesinlik	F1 Skoru
0	11.5	0,82	0,82	0,82
1.0	14	0,85	0,88	0,86
2.0	14	0,92	0,94	0,93
5.0	13	0,85	0,90	0,87

BG-3 görüntüsü için modelde odak kayıp fonksiyonunun kullanılmasının sonuçlar üzerindeki etkisi BG-1 ve BG-2 düzeyinde olmamıştır. Üretilen yanlış alarmlar/kesinlik bu değişiklikten hemen hiç etkilenmezken doğru tespitlerde/geri-getirmede bir miktar ($\gamma = 2.0$ için 3 puan) iyileşme gözlenmiştir (Tablo 5.14).

Tablo 5.14. BG-3 görüntüsü için farklı γ değerlerindeki tespit sonuçları ve model eğitiminde harcanan zamanlar

γ ($\alpha=0.25$)	Eğitimde Harcanan Zaman (saat)	Geri getirme	Kesinlik	F1 Skoru
0	11.5	0,88	0,98	0,93
1.0	14	0,90	0,98	0,94
2.0	14	0,91	0,98	0,94
5.0	13	0,90	0,98	0,94

5.3.3. Karşılaştırma: negatif örnekle eğitim ve odak kayıp fonksiyonunun hatalı tespitler üzerindeki etkisi

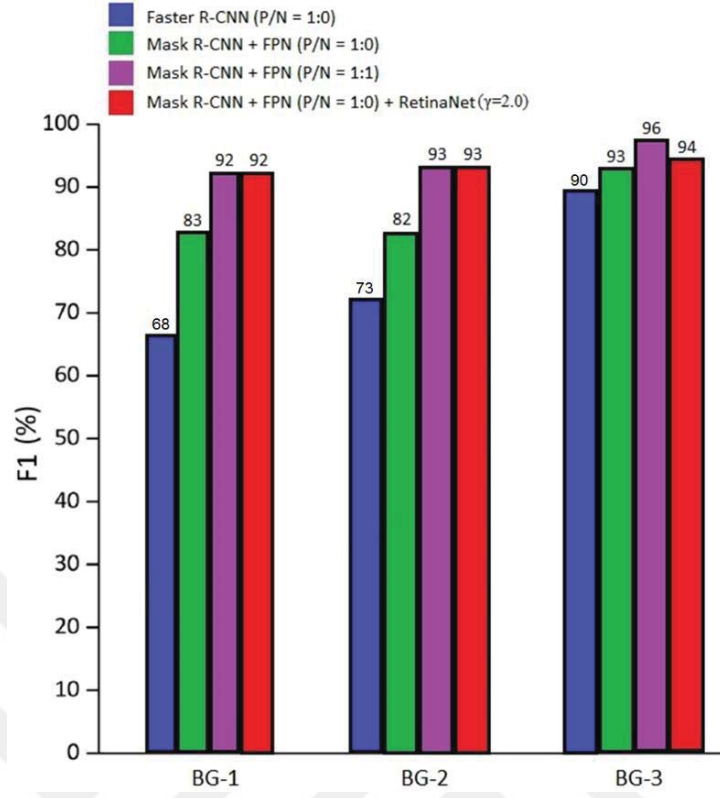
Bu alt başlıkta, Bölüm 5.1 ve 5.2’de ilgili deneylerin sunulmuş olduğu, tespit sonuçlarının iyileştirilmesi ve özellikle yanlış alarmların azaltılması kapsamında önerilmiş olan iki yöntemin karşılaştırmasına yer verilmektedir.

Üç geniş-ölçekli görüntü için 4 farklı modelle gerçekleştirilen deney sonuçları Tablo 5.15’de F1-skorlarıyla özetlenmektedir. Görüntüler, Faster R-CNN (P:N=1.0) , FPN+Mask R-CNN (P:N=1.0), FPN+Mask R-CNN (P:N=2.0) ve RetinaNet+Mask ($\gamma=2.0$) modelleriyle işlenmişlerdir.

Sonuçlardan görüleceği üzere bu görüntüler söz konusu olduğunda gemi tespitindeki en düşük F1-skorunu üreten model, nesne işaretlemelerinin sınır kutularıyla yapıldığı ve ROI hizalama katmanının kullanılmadığı Faster R-CNN yöntemi olmuştur. Aynı verilerle eğitilen FPN’li Mask R-CNN modeli tüm görüntüler için Faster R-CNN’den yüksek performans vermiştir. Eğitimde pozitif örneklerle aynı oranda, karasal bölgeden toplanan negatif örnekler kullanıldığında aynı FPN’li Mask R-CNN modeli için tespit sonuçları tüm görüntülerde iyileşmiştir. Bilhassa karasal alanların içerikteki oranı arttıkça (BG-1 ve BG-2 için) yanlış alarmların hızla azalmasıyla beraber F1 skorunda 9 puana kadar yükseliş olmuştur. Eğitim verilerine negatif örneklerin eklenmesi yerine modelde odak kayıp fonksiyonunun kullanılması da benzer etkiyi göstermiş ve modelin zor örnekler üzerine odaklanmasını sağlayarak hatalı alarmları azaltmaya ve doğru tespitleri artırmaya yardımcı olmuştur. İçeriğinde karasal bölgelerin yer almadığı görüntüde (BG-3) uygulanan iki yöntemin de etkisinin sınırlı kaldığı gözlenmiştir. Sonuçlar grafiksel olarak Şekil 15’te özetlenmiştir.

Tablo 5.15. Yanlış alarmları azaltmaya dönük olarak bu çalışmada kullanılan yöntemlerin birbirleriyle ve baz modellerle olan karşılaştırması

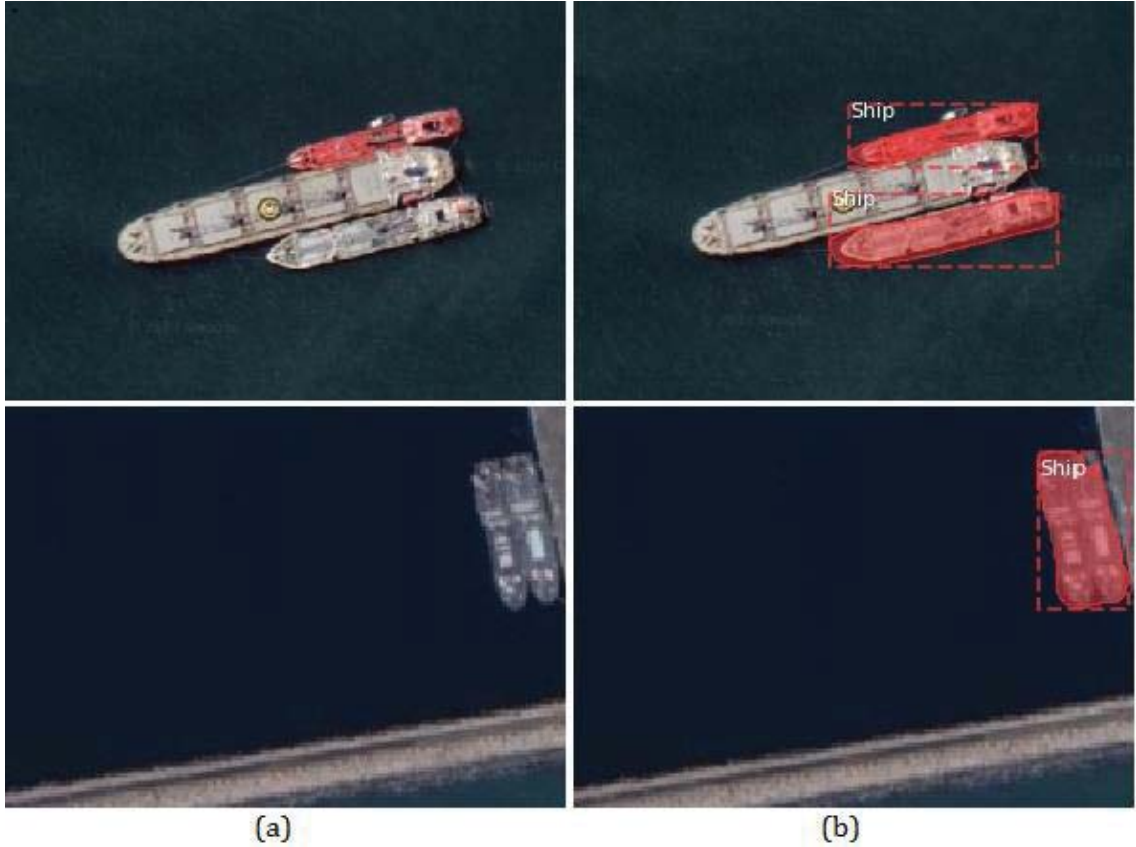
	BG-1				BG-2				BG-3			
	Faster R-CNN	FPN (P:N = 1:0)	FPN (P:N = 1:1)	RetinaNet ($\gamma=2.0$)	Faster R-CNN	FPN (P:N = 1:0)	FPN (P:N = 1:1)	RetinaNet ($\gamma=2.0$)	Faster R-CNN	FPN (P:N = 1:0)	FPN (P:N = 1:1)	RetinaNet ($\gamma=2.0$)
F1	0,68	0,83	0,92	0,92	0,73	0,82	0,93	0,93	0,90	0,93	0,96	0,94



Şekil 5.15. Yanlış alarmları azaltmaya yönelik bu çalışmada kullanılan yöntemlerin birbirleriyle ve baz modellerle olan karşılaştırması

6. TARTIŞMA

Bu çalışma kapsamında gerçekleştirilen deneylerin sonuçlarının karşılaştırılması ve analiz edilmesiyle, önerilen yöntemin katkıları ortaya konmuştur. Uydu görüntülerinde gemi tespitinde, negatif numunelerin de yer aldığı verilerle eğitim alan FPN+Mask R-CNN modeli hem zorlayıcı içeriklerde hem de çoklu ölçeklerde selefi Faster R-CNN yöntemine göre üstün performans vaad etmektedir. Buna karşın, test sonuçları incelendiğinde yöntemin birbirine yakın konumlanmış gemiler söz konusu olduğunda hataya meyilli olduğu göze çarpmaktadır. Bu tür durumlarda ya gemilerden sadece biri işaretlenmekte ya da gemi grubunun tümü tek bir gemi olarak atanmaktadır (Şekil 6.1). Bu türdeki yanlış sınıflandırmalar için iki hata kaynağı olduğu değerlendirilmektedir:



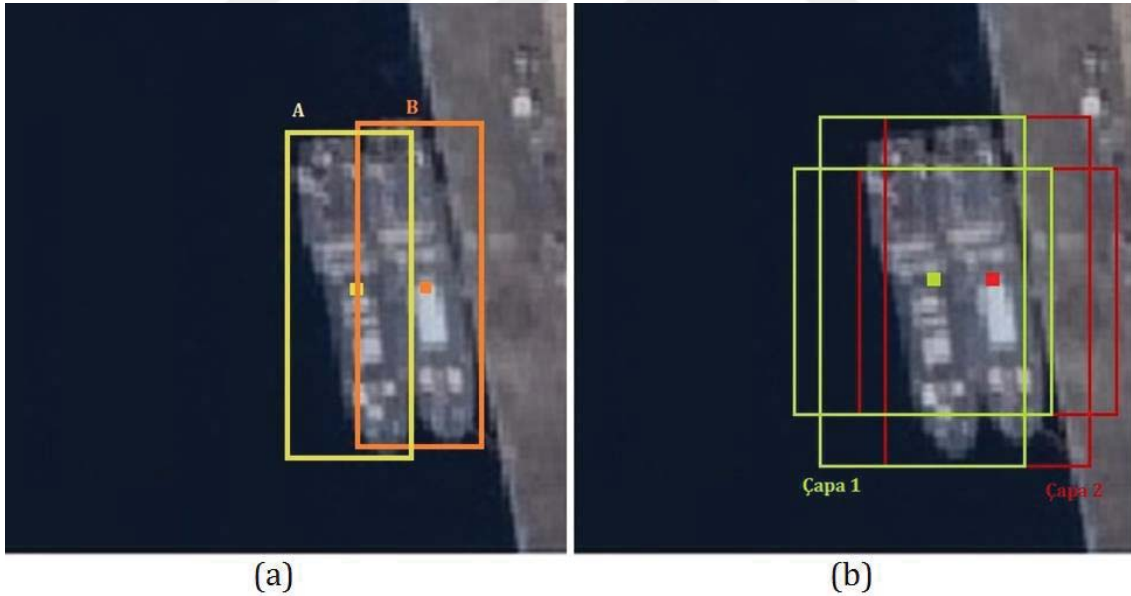
Şekil 6.1. Orijinal görüntüsü a-sütununda verilen görüntüler için modelin ürettiği hatalı tespitler (b-sütunu)

6.1. NMS Eşiği

Bölüm 2.2.2.c ve 2.2.4.b'de detayları ifade edilen, görüntü içerisinde oluşturulan çapaların kullanılarak nesne önerilerinin oluşturulması sürecinde kullanılan maksimum olmayan öneriyi bastırma (NMS) yöntemi tek bir nesne için çok sayıda tespiti yol açılmasını engellemek üzere işlev görmektedir. Gemilerin birbirine yakın olmadığı

durumlarda oluşturulan çapalar ve bunlara bağlı olan nesne önerilerinde gemilerin birbirinden ayırt edilmesinde bir zorluk söz konusu olmamaktadır.

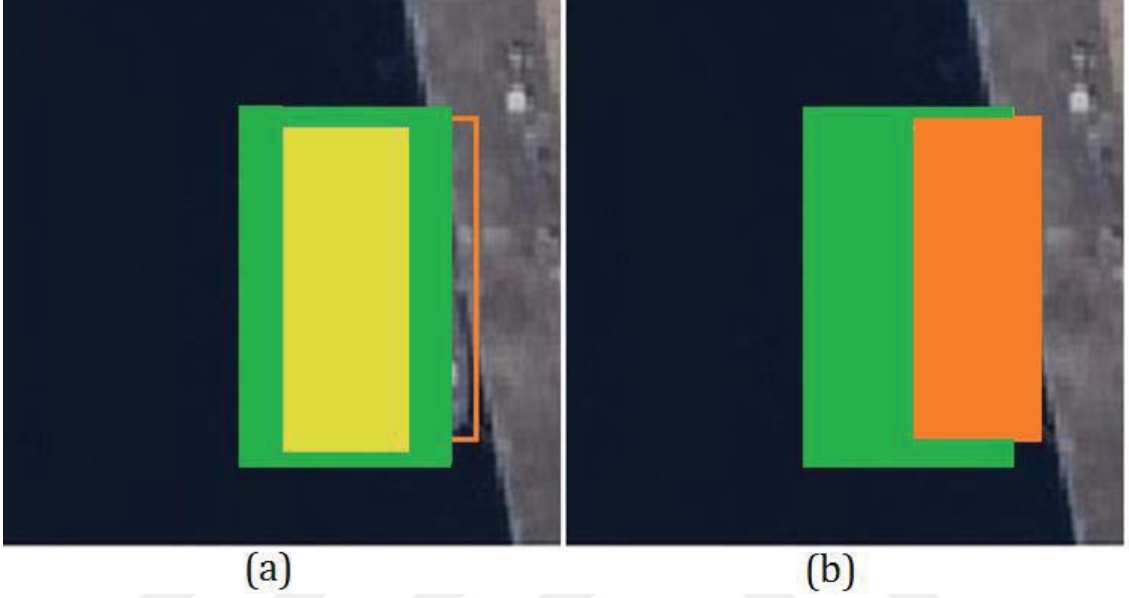
Şekil 6.2.a’da birbirine yakın konumlanmış gemilerin yer doğrulukları gösterilmektedir. Bu görüntünün model tarafından işlenmesi sürecinde, omurga ağın son katmanından gelen özellik haritasının kullanıldığı varsayılırsa, RPN katmanı aracılığıyla gerçek görüntü üzerinde 16 piksel aralıklarla çapalar oluşturulacaktır. Bu çalışma kapsamında, görüntü üzerinde çapaların oluşturulduğu her noktada çapalar 5 farklı büyüklük skalasına {8, 16, 32, 64, 128} ve 3 farklı en boy oranına {1:1, 2:1, 1:2} sahip olacak şekilde oluşturulmuştur. Burada büyüklük skalasındaki sayılar 1:1 en boy oranında oluşturulan çapaların kenar piksel uzunluğudur. 64 piksellik büyüklük skalasına bakarsak en boy oranı 1:2 olduğunda çapaların kenarları, uzun kenarı kısa kenarının iki katı olmak üzere “uzun kenarı x kısa kenarı = $64 x 64 = 4096$ ” olacak şekilde (45 piksel:91 piksel) oluşturulmaktadır. İncelenen örnekte bu skalada ve en boy oranında oluşturulmuş komşu çapaların durumu Şekil 6.2.b’de temsil edilmiştir.



Şekil 6.2. Şekil 6.1’deki görüntüde yer alan gemilere ait (a) yer doğrulukları (b) oluşturulmuş 64x64 piksel büyüklük skalasındaki komşu çapalar

Aynı nesne için farklı çapalardan gelen çok sayıda tespiti engelleme amacıyla belirlenen NMS eşiğiyle, değeri bu eşik üzerinde olan öneriler arasında nesnellik skoru en yüksek olan öneri haricinde diğerlerinin elenmesi amaçlanmaktadır. Bununla beraber, bu örnek için ardışık çapalarca birden fazla nesnenin içerilmesi zorluğa yol açmakta (Şekil 6.3), aynı çapa içerisinde IoU değerleri yüksek birden fazla geminin yer alması söz

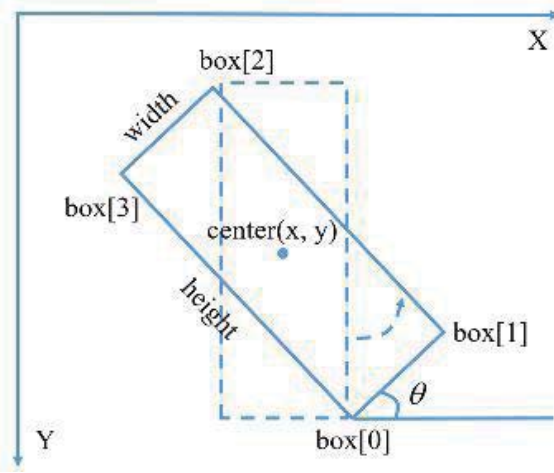
konusu olabilmekte ve kimi zaman iki geminin tek bir gemi olarak görülüp işaretlemesinin yapılmasına kimi zaman da gemilerden birinin veya birkaçının NMS tarafından elenmesine sebebiyet vermektedir. Bu sorun NMS eşik değerinin daha hassas belirlenmesiyle hafifletilebilir olmakla beraber, sınırlayıcı kutular üzerine de eğilimesi gerekmektedir.



Şekil 6.3. Şekil 6.2'de yer alan A gemisinin yer doğruluğunun Çapa-1'le olan kesişim/birleşim durumu, (b) B gemisinin yer doğruluğunun Çapa-1'le olan kesişim birleşim durumu

6.2. RPN'de Dönük Sınır Kutuları Kullanımı

Kullandığımız nesne algılama modeli gemi tespit çıktılarını maske formatında verme kabiliyetinde olmasına rağmen iç işleyişinde (RPN safhasında) nesne önerilerini oluştururken Faster R-CNN modelinde olduğu gibi dikdörtgen sınır kutularını kullanarak gemilerin yer-doğruluklarıyla olan çakışma durumlarını belirlemektedir. Nesne haricindeki alanları da bünyesinde barındıran bu temsil yöntemi yukarıda ifade edildiği gibi, bitişik gemilerin algılanması söz konusu olduğunda yanlış tespitlere, gözden kaçmalara sebebiyet vermektedir. Bu sorunun çözümü için (Yang & Sun, 2018)'de Faster R-CNN'le gemi tespitinde, nesnenin dönüklük açısıyla yönlendirilmiş, standart kutulara göre gemilere çok daha doğru oturma kabiliyetindeki dönük sınır kutularının kullanımı önerilmiştir (Şekil 6.4). Bu yaklaşımın burada önerilen yöntemde de kullanılmasının sonuçları olumlu etkilemesi beklenmektedir.



Şekil 6.4. Yang vd. tarafından (2018) önerilen, RPN’de nesne sınırının temsili için dik sınırlayıcı kutu yerine kullanılmak üzere tanımlanan dönük sınırlayıcı kutu

6.3. Gerçek-Zamanlı Tespit

Bu tezde önerilen ve çoklu-ölçekte yüksek başarımlı sağlama iddiasındaki FPN+Mask R-CNN yöntemiyle saniyede en fazla 8 adet 768x768 piksel boyutundaki görüntü işlenebilmektedir. Küçük gemiler de dahil olmak üzere yüksek tespit performansı vaad eden yöntem, işleyişinde bölge önerisi oluşturmayı da içeren 2-aşamalı bir model olduğundan dolayı bu haliyle gerçek-zamanlı tespitten yani saniyede 24 görüntü işlemekten uzaktadır. Gerçek-zamanlı gemi tespiti gerektiren uygulamalar için yöntem bu koşullarda uygun gözükmemektedir. Bunu sağlayabilmek için (1) girdi görüntüsünün boyutunun küçültülmesi, (2) bölge öneri ağı içermeyen tek-aşamalı modellerin kullanılması önerilebilir.

7. SONUÇ VE ÖNERİLER

Bu tez kapsamında çoklu-ölçekteki ve farklı zorlayıcı içeriklere sahip uzaktan algılama görüntülerinde gemi tespitini yüksek doğrulukla başarmak üzere, bünyesinde Özellik Piramit Ağının yer aldığı Mask R-CNN modelinin kullanımı önerilmektedir. Önceki yaklaşımlardan farklı olarak gemiler görüntülerde sınırlayıcı kutulara ilave olarak gerçek görünümüne daha sadık biçimde maskelerle işaretlenmektedir. Model içinde barındırdığı, kendinden önceki bölge-önerisine dayalı yöntemlerde bulunmayan ve nesne konumunun daha doğru şekilde belirlenmesini sağlayan RoI hizalama katmanının yardımıyla gemi tespitinde daha yüksek doğruluk sağlamaktadır. Bu modelin FPN'le beraber çalıştırılmasıyla omurga ağın farklı bloklarından gelen semantik açıdan zengin, çözünürlüğü düşük özellik haritaları çözünürlüğü yüksek, semantik içeriği zayıf olanlarla kaynaştırılmıştır. Böylelikle model, farklı çözünürlüklerdeki görüntülerde yüksek tespit performansını korumuştur. Önceki çalışmalardan farklı olarak, verinin sentetik olarak artırılmasının sonuçlar üzerindeki etkisi daha nicel bir şekilde ele alınmış ve veri seti için en uygun görüntü işleme yöntemleri bir arada kullanılmıştır. Modelin test görüntüleri üzerinde ürettiği, genellikle karasal alanlarda yer alan ve şekilsel açıdan gemilere benzeyen yapılardan kaynaklanan yanlış tespitleri azaltmak amacıyla negatif örnekleme eğitim ve odak kayıp fonksiyonunun kullanımı önerilmiştir.

Araştırma sürecinde, oluşturulan veri setleri üzerinde farklı tespit modellerini Resnet-50 ve Resnet-101 gibi farklı omurga ağlarıyla ve çeşitli veri artırma teknikleriyle kullanarak kapsamlı deneyler yapılmıştır. Bu alanda sıkça tercih edilen Faster R-CNN gibi bir algılama modeliyle mukayese edildiğinde gemi tespiti için önerilen yöntemin hem ürettiği F1-skoru hem de hesaplama süresi açısından daha başarılı sonuçlar verdiği görülmektedir. Deneyler, negatif örnekleme eğitim almış FPN+Mask R-CNN modelinin gemi tespitinde en yüksek doğrulukta hedefleri belirleyip en az yanlış alarm üreten yöntem olduğunu ortaya koymuştur.

Çalışmayla uzaktan algılama görüntülerinde gemi tespiti görevine aşağıdaki katkılar sunulmaktadır:

- Önceki algılama modellerinden daha iyi performansa sahip, zorlayıcı içeriklerde yer alan gemileri ayırma/algılama kapasitesi yüksek ve tespitte nesneyle alakasız kısımları içermeyen sonuçlar üretebilen, Mask R-CNN yöntemine dayalı yeni bir gemi tespit çerçevesi oluşturulmuştur.

- Özellik piramit ağının algılama modeliyle beraber kullanılmasıyla tespit performansının sadece tek ölçekteki görüntüler ve nesnelere üzerinde değil çoklu ölçekte (farklı çözünürlüklü görüntülerde) yüksek kalması sağlanmaktadır.
- Deniz-kara segmentasyonuna başvurmadan modelde negatif örnekle eğitim veya RetinaNet/odak kaybı yaklaşımının kullanılmasıyla karasal bölgelerden kaynaklı yanlış alarmlar azaltılmaktadır.
- Test verilerinin, (1) içerikleri açısından farklı düzeylerde tespit zorluklarına yol açan üç kategoriye ayrılması ve (2) içeriklerdeki nesnelere büyüklüklerine göre tanzim edilip değerlendirilmesiyle hataların daha nicel analizini gerçekleştirmek mümkün hale gelmiştir.
- Çeşitli veri artırma yöntemlerinin ve farklı omurga ağlarının tespit performansına etkisi sistematik olarak değerlendirilmiştir.

Önerilen yöntemin mevcut tespit algoritmalarına nazaran sağladığı performans iyileştirmelerine rağmen, özellikle birbirine yakın gemiler söz konusu olduğunda bazı hedeflerin gözden kaçırıldığı, bunun da tespit performansının daha fazla iyileştirilmesine engel olduğu gözlenmiştir. Bu soruna çare getirmek adına, modelin RPN modülünde nesne işaretlemesinde kullanılan dik sınırlayıcı kutuların dönük sınırlayıcı kutularla değiştirilip hata fonksiyonlarının buna göre tanımlanması ve model eğitiminin bu yolla gerçekleştirilmesinin uygun olduğu düşünülmektedir.

KAYNAKÇA

- Akçay, G., & Aksoy, S. (2010). Building Detection Using Directional Spatial Constraints. *2010 IEEE International Geoscience and Remote Sensing Symposium*, pp. 1932-1935.
- Amit, S. N., & Aoki, Y. (2017). Disaster detection from aerial imagery with convolutional neural network. *International Electronics Symposium on Knowledge Creation and Intelligent Computing, IES-KCIC 2017*, pp. 239-245.
- An, Q., Pan, Z., & You, H. (2018). Ship Detection in Gaofen-3 SAR Images Based on Sea Clutter Distribution Analysis and Deep Convolutional Neural Network. *Sensors*, 18, 334.
- Antelo, J., Ambrosio, G., Gonzalez, J., & Galindo, C. (2009). Ship detection and recognition in high-resolution satellite images. *In 2009 IEEE International Geoscience and Remote Sensing Symposium*, pp. 509-514.
- Ardeshir, S., Zamir, A. R., Torroella, A., & Shah, M. (2014). GIS-Assisted Object Detection and Geospatial Localization. *Computer Vision – ECCV 2014: 13th European Conference*, pp. 602-617. Zurich, Switzerland: ECCV.
- Aryal, J., & Josselin, D. (2019). Environmental Object Recognition in a Natural Image: An Experimental Approach Using Geographic Object-Based Image Analysis (GEOBIA). *In Geospatial Intelligence: Concepts, Methodologies, Tools, and Applications*, IGI Global, pp. 1353-1371.
- Blaschke, T., Johansen, K., & Tiede, D. (2011). Object-based image analysis for vegetation mapping and monitoring. *Advances in Environmental Remote Sensing: Sensors, Algorithms and Applications*. United States: CRC Press Taylor & Francis Group, pp. 241-271.
- Cheng, D., Meng, G., Xiang, S., & Pan, C. (2016). Efficient sea-land segmentation using seeds learning and edge directed graph cut. *Neurocomputing*, 207, 36-47.
- Cheng, Gong, and Junwei Han. (2016) "A survey on object detection in optical remote sensing images." *ISPRS Journal of Photogrammetry and Remote Sensing*, 117, 11-28.

- Cheng, G., Zhou, P., & 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 - 20.
- Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *In 2005 IEEE computer society conference on computer vision and pattern recognition*, pp. 886-893.
- Dollár, C., & Zitnick, L. (2014). Edge Boxes: Locating Object Proposals from Edges. *European Conference on Computer Vision, ECCV*, pp. 391-405.
- Dollár, P., Appel, R., Belongie, S., & Perona, P. (2014). Fast feature pyramids for object detection. *IEEE transactions on pattern analysis and machine intelligence*, 36(8), 1532-1545.
- Gao, L., He, Y., Sun, X., Jia, X., & Zhang, B. (2019). Incorporating negative sample training for ship detection based on deep learning. *Sensors*, 19 (3), 684.
- Girshick, R. (2015). Fast r-cnn. *In Proceedings of the IEEE international conference on computer vision*, pp. 1440-1448.
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 580-587.
- Han, J., Zhang, D., Cheng, G., Guo, L., & Ren, J. (2014). Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning. *IEEE Transactions on Geoscience and Remote Sensing*, 53(6), 3325-3337.
- He, H., Lin, Y., Chen, F., Tai, H. M., & Yin, Z. (2017). Inshore ship detection in remote sensing images via weighted pose voting. *IEEE Transactions on Geoscience and Remote Sensing*, 55(6), 3091-3107.
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(9), 1904-1916.

- He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. *In Proceedings of the IEEE International Conference on Computer Vision*, pp. 2961-2969.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778.
- Hinton, G., & Salakhutdinov, R. (2006). Reducing the dimensionality of data with neural networks. *Science* 313, 504-507.
- Hu, Q., Du, Y., Jiang, Y., & Ming, D. (2015, October). A new method of inshore ship detection in high-resolution optical remote sensing images. *In AOPC 2015: Image Processing and Analysis*, International Society for Optics and Photonics, 9675, pp. 23 - 29).
- Huang, S., Huang, W., & Zhang, T. (2016). A New SAR image segmentation algorithm for the detection of target and shadow regions. *Scientific reports*, 6, 38596.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779-788.
- Koo, J., Seo, J., Jeon, S., Choe, J., & Jeon, T. (2018). RBox-CNN: rotated bounding box based CNN for ship detection in remote sensing image. *SIGSPATIAL/GIS*, pp. 420-423.
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. *International Conference on Neural Information Processing Systems Curran Associates Inc.*, pp. 1097-1105.
- Gao, L., He, Y., Sun, X., Jia, X., & Zhang, B. (2019). Incorporating negative sample training for ship detection based on deep learning. *Sensors*, 19(3), 684.
- Lee, J., Bang, J., & Yang, S. I. (2017, October). Object detection with sliding window in images including multiple similar objects. *In 2017 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 803-806.

- Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2117-2125.
- Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. *In Proceedings of the IEEE international conference on computer vision*, pp. 2980-2988.
- Liu, G., Zhang, Y., Zheng, X., Sun, X., Fu, K., & Wang, H. (2013). A new method on inshore ship detection in high-resolution satellite images using shape and context information. *IEEE geoscience and remote sensing letters*, 11(3), 617-621.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). Ssd: Single shot multibox detector. *In European conference on computer vision*, Springer, Cham, pp. 21-37.
- Liu, Y., Cui, H. Y., Kuang, Z., & Li, G. Q. (2017). Ship detection and classification on optical remote sensing images using deep learning. *In ITM Web of Conferences EDP Sciences*, pp. 05 - 12.
- Lu, G., Liu, P., & Wang, L. (2015). Rotation Sliding Window of the Hog Feature in Remote Sensing Images for Ship Detection. *International Symposium on Computational Intelligence and Design*, pp. 401-404.
- Ma, M., Chen, J., Liu, W., & Yang, W. (2018). Ship classification and detection based on CNN using GF-3 SAR images. *Remote Sensing*, 10(12), 2043.
- Nie, G. H., Zhang, P., Niu, X., Dou, Y., & Xia, F. (2017). Ship detection using transfer learned single shot multi box detector. *In ITM Web of Conferences, EDP Sciences*, pp. 01006 - 1010.
- Prathap, G., & Afanasyev, I. (2018). Deep Learning Approach for Building Detection in Satellite Multispectral Imagery. *International Conference on Intelligent Systems (IS)*, pp. 461-165.
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *In Advances in neural information processing systems*, pp. 91-99.

- Sadgrove, E. J., Falzon, G., Miron, D., & Lamb, D. W. (2018). Real-time object detection in agricultural/remote environments using the multiple-expert colour feature extreme learning machine (MEC-ELM). *Computers in Industry*, 98, 183-191.
- Santos, A. A. D., Marcato Junior, J., Araújo, M. S., Di Martini, D. R., Tetila, E. C., Siqueira, H. L., & Feitosa, R. Q. (2019). Assessment of CNN-Based Methods for Individual Tree Detection on Images Captured by RGB Cameras Attached to UAVs. *Sensors*, 19(16), 3595.
- Serna, A., & Marcotegui, B. (2014). Detection, segmentation and classification of 3D urban objects using mathematical morphology and supervised learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93, 243-255.
- Shahi, K., Shafri, H., Taherzadeh, E., Mansor, S., & Muniandy, R. (2015). A novel spectral index to automatically extract road networks from WorldView-2 satellite imagery. *The Egyptian Journal of Remote Sensing and Space Science*, pp. 27-33.
- Shi, Z., Yu, X., Jiang, Z., & Li, B. (2014). Ship detection in high-resolution optical imagery based on anomaly detector and local shape feature. *IEEE Trans. Geosci. Remote Sens.*, pp. 4511-4523.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., & Rabinovich, A. (2015). Going deeper with convolutions. *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1-9.
- Tayara, H., & Chong, K. T. (2018). Object detection in very high-resolution aerial images using one-stage densely connected feature pyramid network. *Sensors*, 18(10), 3341.
- Uijlings, J. R., Van De Sande, K. E., Gevers, T., & Smeulders, A. W. (2013). Selective search for object recognition. *International journal of computer vision*, 104(2), 154-171.
- Weinstein, B. G., Marconi, S., Bohlman, S., Zare, A., & White, E. (2019). Individual tree-crown detection in RGB imagery using semi-supervised deep learning neural networks. *Remote Sensing*, 11(11), 1309.

- Yang, X., Sun, H., Fu, K., Yang, J., Sun, X., Yan, M., & Guo, Z. (2018). Automatic ship detection in remote sensing images from google earth of complex scenes based on multiscale rotation dense feature pyramid networks. *Remote Sensing*, 10(1), 132-144.
- Xia, Y., Wan, S., & Yue, L. (2011). A novel algorithm for ship detection based on dynamic fusion model of multi-feature and support vector machine. *6th International Conference on Image and Graphics (ICIG)*, pp. 521-526.
- Xin, J., Zhang, X., Zhang, Z., & Fang, W. (2019). Road Extraction of High-Resolution Remote Sensing Images Derived from DenseUNet. *Remote Sensing*, 11(21), 2499.
- Yang, X., Sun, H., Fu, K., Yang, J., Sun, X., Yan, M., & Guo, Z. (2018). Automatic ship detection in remote sensing images from google earth of complex scenes based on multiscale rotation dense feature pyramid networks. *Remote Sensing*, 10(1), 132.
- Zhu, C., Zhou, H., Wang, R., & Guo, J. (2010). A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features. *IEEE Transactions on geoscience and remote sensing*, 48(9), 3446-3456.

- http-1: Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow, Abdulla, W., & Clauss, C. (2018), GitHub: https://github.com/matterport/Mask_RCNN (Eriřim Tarihi 15 Mart 2019)
- http-2: Gaiser, H., Carpani, V., & Liscio, E. (2018), fizyr /keras-maskrcnn, https://github.com/fizyr/kerasmaskrcnn/blob/master/keras_maskrcnn/models/retinanet.py adresinden alındı (Eriřim Tarihi 15 Mart 2019).
- http-3: Henon, Y. (2017), Keras-rcnn, https://github.com/yhenon/keras-rcnn/tree/master/keras_rcnn/layers adresinden alındı. (Eriřim Tarihi 15 Mart 2019)



ÖZGEÇMİŞ

Adı-Soyadı: Nuri Erkin ÖÇER

Yabancı Dil: İngilizce

Doğum Yeri ve Yılı: Kütahya, 1982

E-posta: neocer@eskisehir.edu.tr

Eğitim Özgeçmişi:

- 2005, Ankara Üniversitesi, Fen Fakültesi, Fizik Bölümü / Lisans
- 2010, Ortadoğu Teknik Üniversitesi, Havacılık ve Uzay Müh. / Yüksek Lisans

Yayınlar ve Bilimsel Faaliyetler:

- Taşar G., Öçer, N. E., Özgen, S., Uzol, O, Near and Far Field Measurements of Circular and Non-Circular Turbulent Jets. "3rd European Conference for Aerospace Sciences", (2009).
- Taşar, G., Öçer, N. E., Miu, L. M., Uzol, O., Özgen, S, A Low-Cost Miniature Two-Color Particle Image Velocimetry (PIV) System for Fluid Mechanics Measurements. "Ankara International Aerospace Conference", (2007).
- Ocer, N. E., Tasar, G., Uzol, O., Ozgen, S., Dairesel ve Dairesel Olmayan Jetlerin Yakın Bölgelerinde Akış Yapısı ve Türbülans. "Kayseri VII. Havacılık Sempozyumu", (2008).
- Ozgen, S., Uzol, O., Öçer, N.E., Taşar, G., Sıvı Jetlerin Akış Kararsızlığı Özelliklerinin Ve Bunlara Etki Eden Faktörlerin Parametrik Bir Çalışma İle Sayısal Olarak Belirlenmesi. "16. Ulusal Mekanik Kongresi, Konferans Bildirisi, Teorik ve Uygulamalı Mekanik Türk Milli Komitesi, (2010)
- N. E. Öçer, U. Avdan, Uzaktan Algılanmış Görüntülerde Obje Tespitinde Ve Sınıflandırmasında Derin Öğrenme Temelli Yaklaşım, TUFUAB IX. Teknik Sempozyumu, (2017)
- Kaplan, G., Öçer, N.E., Erdem, F., Matcı, D., Avdan, U., Tree Detection Using UAV Data And Deep Learning Method: First Step Towards Green Cadastre In Eskişehir Technical University Campus, Geobalcanica (2020)