



T.C.
KONYA TEKNİK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



FARKLI OMURGA MİMARİLERİNE SAHİP
FASTER R-CNN MODELİ İLE BÖBREK
DOKU GÖRÜNTÜLERİNDE GLOMERÜL
TESPİTİ VE SAYIMI

Ramazan KARABACAK

YÜKSEK LİSANS TEZİ

Bilgisayar Mühendisliği Anabilim Dalı

Aralık-2025
KONYA
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Ramazan KARABACAK tarafından hazırlanan “**Farklı Omurga Mimarilerine Sahip Faster R-CNN Modeli ile Böbrek Doku Görüntülerinde Glomerül Tespiti ve Sayımı**” adlı tez çalışması 05/12/2025 tarihinde aşağıdaki jüri tarafından oy birliği ile Konya Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Başkan

Doç. Dr. Özkan İNİK
(Tokat Gaziosmanpaşa Üniversitesi)

.....

Danışman

Doç. Dr. İsmail KOÇ
(Konya Teknik Üniversitesi)

.....

Üye

Dr. Öğr. Üyesi Burak YILMAZ
(Konya Teknik Üniversitesi)

.....

Yukarıdaki sonucu onaylarım.

Prof. Dr. Mevlüt UYAN
Enstitü Müdürü

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Ramazan KARABACAK

Tarih: 05/12/2025

ÖZET

YÜKSEK LİSANS TEZİ

FARKLI OMURGA MİMARİLERİNE SAHİP FASTER R-CNN MODELİ İLE BÖBREK DOKU GÖRÜNTÜLERİNDE GLOMERÜL TESPİTİ VE SAYIMI

Ramazan KARABACAK

**Konya Teknik Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı**

Danışman: Doç. Dr. İsmail KOÇ

2025, 58 Sayfa

Jüri

**Doç. Dr. İsmail KOÇ
Doç. Dr. Özkan İNİK
Dr. Öğr. Üyesi Burak YILMAZ**

Bu çalışmada, histolojik böbrek doku görüntülerinde glomerüllerin otomatik tespiti ve sayımı için Faster R-CNN mimarisi kullanılarak farklı omurga (backbone) mimarilerinin performansları karşılaştırmalı olarak değerlendirilmiştir. AlexNet, ResNet-50, ResNet-101, GoogLeNet (Inception-v3), VGG-16, DenseNet-121, EfficientNet-B0 ve MobileNet-V2 olmak üzere sekiz omurga (backbone) aynı veri kümesi ve eğitim protokolü altında değerlendirilmiştir. Toplam 2500 adet 500×500 piksel çözünürlüğündeki histolojik görüntü %70 eğitim, %15 doğrulama ve %15 test olmak üzere üç gruba ayrılmıştır. Tüm modeller ImageNet veri seti üzerinde ön eğitilmiş ağırlıklarla transfer öğrenme yaklaşımı kullanılarak eğitilmiştir. Performans değerlendirmesi kesinlik (precision), duyarlılık (recall), F1-skor, 0,50 Birleşim Üzerinden Kesişim (IoU) eşliğinde Genel Ortalama Hassasiyet mAP50 ve 0,50-0,95 IoU aralığında Genel Ortalama Hassasiyet (mAP50-95) metrikleri kullanılarak gerçekleştirilmiştir. Test seti sonuçlarına göre, MobileNet-V2 genel performans açısından en yüksek F1-skor değerini 0,973 ile öne çıkmıştır. Kesinlik metriğinde ResNet-50 0,954 en başarılı olurken, duyarlılık metriğinde DenseNet-121 0,994 en yüksek değeri göstermiştir. Nesne tespit kalitesinin değerlendirilmesinde mAP50 metriğinde VGG-16 0,992 öne çıkarken, mAP50-95 metriğinde VGG-16 0,708 en iyi performansı göstermiştir.

Anahtar Kelimeler: Evrimsel Sinir Ağları, Derin Öğrenme, Faster R-CNN, Nesne Tespiti, Omurga Mimarileri

ABSTRACT

MS THESIS

DETECTION AND COUNTING OF GLOMERULI IN KIDNEY TISSUE IMAGES USING FASTER R-CNN MODEL WITH DIFFERENT BACKBONE ARCHITECTURES

Ramazan KARABACAK

**Konya Technical University
Institute of Graduate Studies
Department of Computer Engineering**

Advisor: Assoc. Prof. Dr. İsmail KOÇ

2025, 58 Pages

**Jury
Assoc. Prof. Dr. İsmail KOÇ
Assoc. Prof. Dr. Özkan İNİK
Asst. Prof. Dr. Burak YILMAZ**

In this study, the performance of different backbone architectures was comparatively evaluated using the Faster R-CNN architecture for the automatic detection and counting of glomeruli in histological kidney tissue images. Eight backbones-AlexNet, ResNet-50, ResNet-101, GoogLeNet (Inception-v3), VGG-16, DenseNet-121, EfficientNet-B0, and MobileNet-V2-were evaluated under the same dataset and training protocol. A total of 2500 histological images with a resolution of 500×500 pixels were divided into three groups: 70% training, 15% validation, and 15% testing. All models were trained using a transfer learning approach with pre-trained weights on the ImageNet dataset. Performance evaluation was performed using the metrics of precision, recall, F1-score, Mean Average Precision at 0,50 Intersection over Union (IoU) threshold (mAP50), and Mean Average Precision in the 0,50-0,95 IoU range (mAP50-95). According to the test set results, MobileNet-V2 stood out with the highest F1-score value of 0,973 in terms of overall performance. ResNet-50 was the most successful in the precision metric with 0,954, while DenseNet-121 showed the highest value in the recall metric with 0,994. In evaluating object detection quality, VGG-16 stood out with 0,992 in the mAP50 metric, while VGG-16 showed the best performance with 0,708 in the mAP50-95 metric.

Keywords: Convolutional Neural Networks, Deep Learning, Faster R-CNN, Object Detection, Backbone Architectures

ÖNSÖZ

Bu tez çalışmasının yürütülmesi sürecinde bilgi, birikim ve tecrübelerini benimle paylaşarak araştırma konusunun belirlenmesine öncülük eden değerli danışmanım Doç. Dr. İsmail KOÇ'a en içten teşekkürlerimi sunarım. Çalışmada kullanılan veri setinin temini konusundaki katkıları ve tez savunması sürecindeki yapıcı değerlendirmeleri için tez savunma jürisi üyesi Doç. Dr. Özkan İNİK'e ayrıca savunmadaki değerli görüş ve önerileri için jüri üyesi Dr. Öğr. Üyesi Burak YILMAZ'a teşekkür ederim. Bu süreç boyunca beni destekleyen aileme de teşekkürlerimi sunarım.

Ramazan KARABACAK
KONYA-2025

İÇİNDEKİLER

ÖZET	iv
ABSTRACT.....	v
ÖNSÖZ	vi
İÇİNDEKİLER.....	vii
SİMGELER VE KISALTMALAR.....	ix
1. GİRİŞ.....	1
1.1. Tezin Amacı ve Önemi	2
1.2. Tezin Organizasyonu	2
2. KAYNAK ARAŞTIRMASI	3
3. MATERYAL VE YÖNTEM	11
3.1. Evrişimsel Sinir Ağı (ESA) Mimarisi.....	11
3.1.1. Giriş katmanı	13
3.1.2. Evrişim katmanı	13
3.1.3. Düzeltilmiş doğrusal birim (ReLU) katmanı	15
3.1.4. Havuzlama Katmanı	16
3.1.5. Tam Bağlantılı Katman.....	18
3.1.6. Sınıflandırma Katmanı.....	18
3.2. Evrişimsel Sinir Ağı (ESA) Modelinin Eğitim Parametreleri	19
3.2.1. Devir sayısı (Epoch)	19
3.2.2. Optimizasyon algoritması	19
3.2.3. Öğrenme hızı (Learning rate)	19
3.2.4. Momentum katsayısı	20
3.2.5. Paket boyutu (Mini-batch size).....	20
3.2.6. Erken durdurma (Early stopping)	20
3.3. Nesne Sınıflandırmada Kullanılan Derin Öğrenme Mimarileri	21
3.3.1. AlexNet.....	21
3.3.2. ResNet.....	22
3.3.3. GoogLeNet	23
3.3.4. VGGNet.....	25
3.3.5. EfficientNet.....	27
3.3.6. DenseNet.....	27
3.3.7. MobileNet	28
3.4. Nesne Tespitinde Kullanılan Derin Öğrenme Modelleri.....	29
3.4.1. R-CNN	29
3.4.2. Fast R-CNN	30
3.4.3. Faster R-CNN	31
3.5. Önerilen Yöntem	32

3.5.1. Veri Seti	33
3.5.2. Model Mimarisi ve Omurga Seçimi	34
3.5.3. Model Eğitimi	34
3.5.4. Eğitim Ortamı	35
3.5.5. Hiper parametreler ve Optimizasyon Stratejisi.....	35
3.5.6. Eğitim Süreci ve Kayıp Fonksiyonları	36
3.6. Performans Değerlendirme Metrikleri.....	40
3.6.1. Kesinlik.....	42
3.6.2. Duyarlılık.....	42
3.6.3. F1-Skor	42
3.6.4. Ortalama Hassasiyet	42
3.6.5. Genel Ortalama Hassasiyet.....	42
4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA.....	44
4.1. Doğrulama (Validation) Sonuçları.....	44
4.2. Test Sonuçları	46
4.2.1. Kesinlik (Precision) performansı.....	47
4.2.2. Duyarlılık (Recall) performansı.....	47
4.2.3. F-1 Skor performansı	48
4.2.4. Ortalama Hassasiyet (Average Precision) performansı	48
4.2.5. Modellerin TP, FP ve FN Değerlerinin Karşılaştırılması	50
5. SONUÇLAR VE ÖNERİLER	54
5.1 Sonuçlar	54
5.2 Öneriler	55
KAYNAKLAR	56

SİMGELER VE KISALTMALAR

Kısaltmalar

AP	: Ortalama Hassasiyet (Average Precision)
AUC	: Eğri Altındaki Alan (Area Under the Curve)
BÖA	: Bölge Öneri Ağı (Region Proposal Network - RPN)
CenterNet Object Detection)	: Merkez Nokta Tabanlı Nesne Tespiti (Center Point Based
DenseNet Networks)	: Yoğun Bağlantılı Ağ (Densely Connected Convolutional
EK	: Evrişim Katmanı
ESA	: Evrişimsel Sinir Ağları (Convolutional Neural Networks)
Fast R-CNN	: Hızlı Bölge Tabanlı Evrişimsel Sinir Ağı (Fast Region- based Convolutional Neural Network)
Faster R-CNN	: Daha Hızlı Bölge Tabanlı Evrişimsel Sinir Ağı (Faster Region-based Convolutional Neural Network)
FN	: Yanlış Negatif (False Negative)
FP	: Yanlış Pozitif (False Positive)
GPU	: Grafik İşlem Birimi (Graphics Processing Unit)
ILSVRC	: Büyük Ölçekli Görsel Tanıma Yarışması (ImageNet Large Scale Visual Recognition Challenge)
ImageNet	: Görüntü Ağı (Image Networks)
IoU	: Birleşim Üzerinden Kesişim (Intersection over Union)
mAP	: Genel Ortalama Hassasiyet (Mean Average Precision)
MobileNet	: Taşınabilir Ağ (Mobile Network)
MR	: Manyetik Rezonans
R-CNN	: Bölge Bazlı Evrişimsel Sinir Ağı (Region-Based Convolutional Neural Networks)
ReLU	: Düzleştirilmiş Doğrusal Birim (Rectified Linear Unit)
ResNet	: Kalıntı Ağı (Residual Network)
ROC	: Alıcı İşletim Karakteristiği (Receiver Operating Characteristics)
RoI	: Bölge Önerisi Ortaklaması (Region of Interest Pooling)
SA	: Seçici Arama (Selective Search)
SGD	: Stokastik Gradyan İniş (Stochastic Gradient Descent)

SGDM	: Momentumlu Stokastik Gradyan İniş (Stochastic Gradient Descent with Momentum)
SSD	: Tek Atışta Tespit (Single Shot Detection)
TN	: Doğru Negatif (True Negative)
TP	: Doğru Pozitif (True Positive)
VGGNet	: Görsel Geometri Grubu Ağı (Visual Geometry Group Network)
YOLO	: Yalnızca Bir Kere Bak (You Only Look Once)
ZFNet	: Zeiler ve Fergus Evrişimsel Ağı (Zeiler and Fergus Convolutional Network)



1. GİRİŞ

Görüntü analizi, bilgisayarlı görünün önemli araştırma alanlarından biridir. Günümüzde yapay zekâ ve makine öğrenmesi yöntemleri bu alanda yaygın olarak kullanılmaktadır. Geleneksel makine öğrenmesi yaklaşımlarında, görüntülerin analiz edilebilmesi için özellik çıkarımı adı verilen bir ön işlem adımı gereklidir. Bu aşamada, görüntülerden ayırt edici özelliklerin belirlenmesi ve seçilmesi sürecinde uzman bilgisine ihtiyaç duyulmaktadır.

Derin sinir ağları ise ham veriden özellikleri otomatik olarak öğrenir. Bu sayede uzman müdahalesine olan ihtiyaç azalmaktadır. Evrişimsel Sinir Ağları (ESA), görüntü sınıflandırma ve nesne tespiti görevlerinde etkili sonuçlar vermektedir. Günümüzde derin öğrenme, tıbbi görüntü segmentasyonu, kanser teşhisi ve evrelemesi, doku sınıflandırması, lezyonların otomatik tespiti gibi sağlık alanlarında yaygın olarak kullanılmaktadır. Ayrıca derin öğrenme, sağlık dışında otonom sürüş, endüstriyel üretim, doğal dil işleme ve finans gibi pek çok alanda da başarıyla uygulanmaktadır.

Böbrek hastalıkları, dünya genelinde önemli sağlık sorunlarına yol açmaktadır. Bu hastalıkların erken teşhisi, tedavi sürecinin başarısı açısından kritik öneme sahiptir. Glomerüller, böbreklerde kanın süzülmesini sağlayan temel fonksiyonel birimlerdir. Glomerüllerin sayısı, boyutu ve yapısal bütünlüğü, böbrek hastalıklarının teşhisinde önemli parametreler olarak değerlendirilmektedir. Ancak glomerüllerin manuel olarak tespiti ve sayımı, uzmanlık gerektiren ve zaman alıcı bir süreçtir. Bu süreç ayrıca insan hatasına açıktır.

Yapılan bu tez çalışması kapsamında histolojik böbrek görüntülerinden glomerüllerin otomatik tespiti ve sayımı için Daha Hızlı Bölge Tabanlı Evrişimsel Sinir Ağı (Faster Region-based Convolutional Neural Network-Faster R-CNN) nesne tespit algoritması kullanılmıştır. Sekiz farklı omurga (backbone) mimarisi sistematik olarak karşılaştırılmıştır. Çalışma kapsamında omurga mimarisi olarak AlexNet, ResNet-50, ResNet-101, GoogLeNet (Inception-v3), VGG-16, DenseNet-121, EfficientNet-B0 ve MobileNet-V2 modelleri değerlendirilmiştir. Araştırmanın temel amacı farklı omurga mimarilerinin glomerül tespiti görevindeki performansını değerlendirmektir.

1.1. Tezin Amacı ve Önemi

Bu tezde amaç, histolojik böbrek görüntülerinde glomerüllerin otomatik tespiti için Faster R-CNN mimarisini kullanmak ve omurga seçiminin bu görevdeki etkisini karşılaştırma ile ortaya koymaktır. Bunun için AlexNet, ResNet-50, ResNet-101, GoogLeNet (Inception-v3), VGG-16, DenseNet-121, EfficientNet-B0 ve MobileNet-V2 omurgaları aynı veri kümesi aynı eğitim protokolü ve hiper parametreler altında değerlendirilmiştir. Böylece performans farklarının yalnızca mimariden kaynaklanması hedeflenmiştir. Glomerüllerin manuel tespiti zaman alıcı uzmanlığa bağımlı ve hataya açık olduğundan geliştirilen sistemin patologların iş yükünü azaltarak daha verimli çalışmalarına katkı sağlaması hedeflenmektedir. Bu sayede daha fazla hastanın daha kısa sürede değerlendirilmesi amaçlanmaktadır.

1.2. Tezin Organizasyonu

Yapılan bu tez çalışması kapsamında tez beş bölümden oluşmaktadır. Bölüm 1’de tezin amacı ve önemi açıklanmıştır. Bölüm 2’de kaynak araştırması yapılmış olup Faster R-CNN mimarisinin omurga seçimi, eğitim stratejileri ve değerlendirme yaklaşımları literatürle karşılaştırmalı olarak incelenmiştir. Bölüm 3’te öncelikle evrimsel sinir ağlarının temel prensipleri açıklanmaktadır. Ardından derin öğrenme, nesne sınıflandırmada ve nesne tespitinde kullanılan derin öğrenme modelleri ele alınmaktadır. Bu bölümde ayrıca tezde kullanılan materyal ve metotlar hakkında bilgilendirmeler yapılmıştır. Veri kümesinin özellikleri, Faster R-CNN mimarisi ve kullanılan model mimarileri ve eğitim protokolü açıklanmaktadır. Bu bölümün sonunda modellerin performans değerlendirmek için kullanılan metrikler detaylıca açıklanmaktadır. Bölüm 4’te deneysel sonuçlar ve karşılaştırmalar ayrıntılı biçimde sunulmuştur. Bölüm 5’te çalışmanın genel değerlendirmesi yapılmakta sonuçlar verilip gelecek çalışmalar için öneriler sunulmuştur.

2. KAYNAK ARAŞTIRMASI

Derin öğrenme, son yıllarda görüntü tabanlı uygulamalarda yaygın bir şekilde kullanılmakta olup, bu alanda özellikle Evrişimsel Sinir Ağları (ESA) öne çıkmaktadır. Bu tez çalışması kapsamında histolojik böbrek görüntülerinden glomerüllerin otomatik tespiti amacıyla Faster R-CNN mimarisini temel alan bir derin öğrenme modeli önerilmektedir. Konuyla ilgili benzer çalışmalar bu bölümünde sunulmuştur.

Jasitha ve Pournami (2023), böbrek hastalıklarının teşhisinde önemli bir rol oynayan glomerül tespiti için Hızlandırılmış Bölge Tabanlı Evrişimsel Sinir Ağı (Faster R-CNN) ve Merkez Nokta Tabanlı Nesne Tespiti (CenterNet) modellerini karşılaştırmalı olarak ele almıştır. Biyopsi görüntüleri yamaları üzerinde yapılan deneylerde, Faster R-CNN modelinin 0,657 genel ortalama hassasiyet (mAP) değerleriyle CenterNet modeline göre daha üstün performans sergilediği gösterilmiştir. Bu çalışma, glomerül tespiti için derin öğrenme tabanlı yaklaşımların etkinliğini vurgulamaktadır (Jasitha ve Pournami, 2023).

Traoré ve diğ. (2020), akciğer nodüllerinin tespitinde Faster R-CNN dahil olmak üzere farklı derin öğrenme modellerinin performansını karşılaştırmışlardır. Çalışmada LUNA16 veri seti kullanılarak mAP metriğiyle değerlendirme yapılmıştır. Faster R-CNN modeli, Resnet-50 ve Resnet-101 omurga mimarileriyle test edilmiştir. Sonuçlar, Resnet-101 omurgasının kullanıldığı Faster R-CNN modelinin 0,357 mAP50 ile daha yüksek performans gösterdiğini ortaya koymuştur (Traoré ve ark., 2020).

Akyol (2021), diyabetik ayak ülserlerinin tespiti konusunda Faster R-CNN modelini uygulamıştır. Çalışmada ResNet-50 ve MobileNet-V2 omurga mimarileri kullanılarak DFUC2020 veri seti üzerinde karşılaştırmalı bir analiz gerçekleştirilmiştir. Deneysel sonuçlar, ResNet-50 omurgasının 0,43 mAP sağlarken, MobileNet-V2 omurgasının 0,34 oranında mAP elde ettiğini göstermektedir. Elde edilen sonuçlara göre ResNet-50 mimarisi diyabetik ülser tespitinde yüksek performans gösterdiğini ortaya koymuştur (Akyol).

Ma ve diğ. (2020), servikal omurga manyetik rezonans (MR) görüntülerinde disk dejenerasyonu hastalıkları ile travmatik omurilik hasarı lezyonlarını tespit etmeye yönelik derin öğrenme tabanlı bir yöntem önermiştir. Çalışmada, omurga mimarisi olarak ResNet-50 ve VGG-16 ağlarının performansı karşılaştırmalı olarak değerlendirilmiştir. Deneysel sonuçlar ResNet-50 mimarisi ile entegre edilen Faster R-CNN modelinin 0,886 mAP değerine ulaştığını göstermiştir. Bununla birlikte omurga olarak VGG-16 mimarisi

kullanılan model ile yapılan deneylerde 0,723 mAP değeri elde ettiği ve tespit performansının ResNet-50 tabanlı modele kıyasla daha düşük kaldığı belirlenmiştir (Ma ve ark., 2020).

Çelik (2022), panoramik radyograflerde alt çenede gömülü dişlerin otomatik tespiti amacıyla derin öğrenme yöntemleri uygulanmıştır. Araştırmada iki aşamalı bir yöntem olan Faster R-CNN (ResNet-50, AlexNet ve VGG-16 omurgaları ile) ve tek aşamalı bir yöntem olan YOLOv3 modeli kullanılarak karşılaştırmalı bir analiz yapılmıştır. Deneysel sonuçlar, YOLOv3 modelinin 0,96 mAP değeri ile tüm modeller arasında en yüksek tespit performansını sağladığını göstermektedir. Faster R-CNN modelinde ise ResNet-50 omurga ağı 0,91 oranında mAP değeri ile AlexNet ve VGG-16'nın sırasıyla 0,86 ve 0,87 oranındaki performanslarından daha üstün sonuçlar ortaya koymuştur (Celik, 2022).

Huynh ve diğ. (2022), akne lezyonlarının tespiti ve şiddet derecelendirmesine yönelik bir derin öğrenme tabanlı sistem geliştirmiştir. Bu çalışmada, akne lezyonlarının tespiti için Faster R-CNN mimarisi kullanılmış ve omurga ağı olarak ResNet-50 tercih edilmiştir. Model, akne tespitinde 0,54 mAP değeri ile yüksek bir performans ortaya koymuştur (Huynh ve ark., 2022).

Sert (2021), biber ve patates yapraklarında hastalık tespiti için Faster R-CNN ve GoogLeNet mimarilerini birleştiren yeni bir nesne tespit yaklaşımı (Faster R-CNN-GC) önermiştir. Önerilen yaklaşım Faster R-CNN modeline kıyasla daha yüksek doğruluk oranları elde etmiştir (Sert, 2021).

Öztürk (2022), derin öğrenme yöntemleri kullanarak bina bölgelerine ait çatı konumlarının tespiti ve çatı tiplerinin sınıflandırılmasına yönelik bir çalışma gerçekleştirmiştir. Bu çalışmada, AlexNet, GoogLeNet ve DarkNet-53 modelleri kullanılmıştır. Uydu görüntüleri üzerinde yapılan testler sonucunda AlexNet modelinin 0,923, GoogLeNet modelinin ise 0,871 doğruluk oranıyla çatı bölgelerini tespit ettiği görülmüştür. Bu sonuçlar AlexNet modelinin GoogLeNet modelinden daha başarılı sonuçlar elde ettiğini ortaya koymuştur (Öztürk, 2022).

Sorgun (2022), X-ışını görüntülerinde yasaklı nesnelerin tespiti amacıyla Faster R-CNN tabanlı bir yaklaşım önermiştir. Çalışmada SIXray veri seti kullanılarak eğitilen ResNet-50, ResNet-101, ResNet-152 ve Inception-ResNet-v2 omurga mimarilerine dayalı dört farklı Faster R-CNN modeli kıyaslanmıştır. ResNet-101 mimarisi ile yapılandırılmış Faster R-CNN modelinin 0,906 doğrulukla en yüksek tespit performansını sağladığı görülmüştür (Sorgun, 2022).

Kawazoe ve diğ. (2018), böbrek biyopsisi görüntülerinde glomerül tespitine yönelik Faster R-CNN tabanlı bir yaklaşım geliştirmiştir. Çalışmada Tokyo Üniversitesi Hastanesinden toplanan 800 biyopsi görüntüsü kullanılmıştır. Omurga mimarisi olarak Inception-ResNet ile önceden eğitilmiş bir model uygulanmıştır. Toplam yaklaşık 33.000 glomerül içeren veri kümesi dört farklı boyama türüne göre değerlendirilmiştir. Deneysel sonuçlar, dört farklı boyama türü için sırasıyla 0,876 ile 0,928 arasında değişen F-1 skor değerleri ortaya koymuştur. Bu bulgular Faster R-CNN mimarisinin farklı boyama türlerine karşı güçlü bir performans gösterdiğini ve dijital patoloji uygulamalarında otomatik glomerül tespitinin uygulanabilir olduğunu göstermiştir (Kawazoe ve ark., 2018).

Pandey ve Raghatat (2025), gerçek zamanlı trafik işareti tespit ve tanıma sistemleri için Faster R-CNN mimarisini VGG-16 omurgası ile birleştiren bir yöntem geliştirmiştir. Çalışma German Traffic Sign Detection Benchmark (GTSDb) veri kümesi üzerinde gerçekleştirilmiştir. Test veri kümesinde 0,934 kesinlik, 0,921 duyarlılık, 0,927 F1-skoru ve 0,945 mAP değerleri elde edilmiştir (Pandey ve Raghatat, 2025).

Sood ve diğ. (2025), buğday bitkilerindeki mantar hastalıklarının otomatik tespitine yönelik derin öğrenme tabanlı nesne tespit yöntemlerinin karşılaştırmalı analizini gerçekleştirmiştir. Çalışmada WFD2020 veri kümesinden türetilen ve sarı pas, yaprak pası ile külleme hastalıklarını kapsayan 350 görüntüden oluşan bir veri seti kullanılmıştır. Faster R-CNN ve Tek Atışta Tespit (SSD) mimarileri ResNet-50, ResNet-101 ve ResNet-152 omurga ağları ile değerlendirilmiş olup, en yüksek performans ResNet-50 omurga yapısı kullanılan Faster R-CNN ile 0,62 oranında mAP değeri ile elde edilmiştir. (Sood ve ark., 2025).

Hakim ve diğ. (2024), trafik uygulamalarında araç tespiti için Faster R-CNN algoritmasının farklı omurga mimarilerinin performansını değerlendirmiştir. Araştırmada ResNet-50, ResNet-50-v2, MobileNet-V3-Large ve MobileNet-V3-Large-320 omurga mimarileri aynı eğitim protokolü altında karşılaştırılmıştır. Çalışmada, otobüs, otomobil, motosiklet, pikap, kamyon ve kamyon kutusu olmak üzere altı araç sınıfından oluşan 1104 resimlik Endonezya trafik veri seti kullanılmıştır. ResNet-50 mimarisi 0,5 IoU eşliğinde 0,966 oranında mAP değeri ile en yüksek performansı göstermiştir. Çalışmanın sonuçları, omurga mimarisi seçiminin araç tespit performansı üzerinde önemli bir etkisi olduğunu ortaya koymakta ve ResNet-50 modelinin bu uygulama için en uygun mimarı olduğunu göstermektedir (Hakim ve ark., 2024).

Ma ve diğ. (2024), insansız hava araçları platformunda hedef takip ve tespiti için optimize edilmiş HRNet omurga ağını Faster R-CNN mimarisinde kullanarak yeni bir yaklaşım önermişlerdir. Çalışmada, VisDrone2019 veri seti kullanılarak geliştirilen modelin performansı farklı derin öğrenme tabanlı nesne tespit yöntemleriyle karşılaştırılmıştır. Bu kapsamda, ResNet-50 ve ResNet-101 omurga yapısına sahip Faster R-CNN, VGG tabanlı SSD, Darknet-53 omurgalı YOLOv3 ve EfficientNet-B0 tabanlı YOLOv8 modelleri değerlendirilmiştir. Deneysel sonuçlar, HRNet omurgalı Faster R-CNN modelinin 0,369 mAP değeri ile diğer modellere göre üstün performans gösterdiğini ortaya koymuştur (Ma ve ark., 2024).

Bhatt ve diğ. (2025), akciğer bilgisayarlı tomografi (BT) görüntülerinde nodül tespiti amacıyla SSD ve Faster R-CNN modellerinin performansını karşılaştırmıştır. Araştırmada, modeller 14 farklı veri seti üzerinde eğitilmiştir. Deneysel sonuçlar, MobileNet-V2 omurgalı SSD modelinin 0,438 doğruluk oranına ulaştığını, buna karşın Inception-V2 tabanlı Faster R-CNN modelinin 0,958 doğruluk oranı elde ederek yüksek performans gösterdiğini ortaya koymuştur (Bhatt ve ark., 2025).

Kharb ve Chaudhary (2024), beyin tümörlerinin tespiti ve sınıflandırılması amacıyla EfficientNet ve Faster R-CNN mimarilerini birleştiren hibrit bir derin öğrenme modeli önermiştir. Çalışmada 3064 MR görüntüsünden oluşan veri bir seti kullanılmıştır. Geliştirilen hibrit modelde Faster R-CNN mimarisine omurga olarak EfficientNet-B0 entegre edilmiştir. Modelin performansı tek başına kullanılan Faster R-CNN ve EfficientNet modelleriyle karşılaştırılmıştır. Deneysel sonuçlar hibrit modelin 0,992 test doğruluğu elde ederek yüksek bir sınıflandırma başarımı gösterdiğini ortaya koymuştur (Kharb ve Chaudhary, 2024).

Hasan ve diğ. (2022), yol yüzeyindeki hasarların otomatik tespiti ve sınıflandırılması amacıyla derin sinir ağı tabanlı bir yaklaşım sunmuştur. RDD-2020 veri seti kullanılarak gerçekleştirilen araştırmada EfficientDet, MobileNet-V1 ve MobileNet-V2 omurgalı SSD modelleri ile ResNet-50 omurgalı SSD ve ResNet-101 omurgalı Faster R-CNN modellerinden oluşan beş farklı ön eğitilmiş model karşılaştırmalı olarak değerlendirilmiştir. Deneysel sonuçlar ResNet-101 omurgalı Faster R-CNN modelinin diğer modellere kıyasla daha yüksek performans gösterdiğini ortaya koymuştur (Hasan ve ark., 2022).

Yang ve diğ. (2021), iletim hatlarındaki gövde kusurlarının otomatik tespiti amacıyla Faster R-CNN tabanlı bir kusur tanıma algoritması geliştirmiştir. Çalışmada, özellik çıkarma ağırları olarak ZFNet ve VGG-16 mimarileri kullanılmış ve bu iki model

farklı iterasyon sayılarıyla karşılaştırmalı olarak değerlendirilmiştir. Deneysel sonuçlar, aynı iterasyon sayısında VGG-16 ağının mAP değerinin ZFNet modeline kıyasla daha yüksek olduğunu göstermiştir (Yang ve ark., 2021).

Wang ve Qi (2019), domates hastalıklarının tespiti ve sınıflandırılması için Faster R-CNN tabanlı bir sistem geliştirmiştir. Çalışmada sağlıklı domates meyveleri ile altı farklı hastalık türünü içeren toplam 207 orijinal görüntüden oluşan veri seti dijital aydınlatma ayarları, otomatik beyaz dengesi ve yüksek ISO gürültü azaltma gibi veri artırma teknikleriyle 1035 görüntüye çıkarılmıştır. Faster R-CNN mimarisi omurga ağı olarak VGG-16, ResNet-50 ve ResNet-101 modelleriyle yapılandırılmış ve performansları karşılaştırılmıştır. Deneysel sonuçlar, ResNet-101 tabanlı modelin 0,908 mAP değeri ile diğer mimariler arasında en yüksek performansı sergilediğini göstermiştir (Wang ve Qi, 2019).

Somasundaram ve Basak (2025), kan yayma görüntülerinden lösemi evresinin tespiti ve sınıflandırılması amacıyla VGG-11 omurgalı Faster R-CNN mimarisini değerlendirmiştir. Çalışmada 6600 eğitim, 623 doğrulama ve 312 test örneğinden oluşan bir veri seti kullanılmış ve transfer öğrenme yöntemi uygulanmıştır. Elde edilen sonuçlara göre, VGG-11 omurgalı modelin mAP değeri 0,972, AlexNet omurgalı modelin ise 0,961 olarak bulunmuştur. Bu sonuçlar, omurga mimarisi seçiminin aynı deneysel koşullar altında tespit ve sınıflandırma performansı üzerinde belirgin bir etkiye sahip olduğunu göstermektedir (Somasundaram ve Basak, 2025).

Albahli ve diğ. (2021), diyabetik retinopati (DR) lezyonlarının tanınması ve şiddet düzeyine göre sınıflandırılması amacıyla DenseNet-65 omurgalı özelleştirilmiş bir Faster R-CNN modeli önermiştir. Kaggle platformundan derlenen 88.704 retina görüntüsü üzerinde yürütülen çalışmada diyabetik retinopatinin beş aşaması otomatik olarak tespit edilmiş ve sınıflandırılmıştır. Model 0,972 doğruluk oranına ulaşmıştır (Albahli ve ark., 2021).

Devi ve diğ. (2024), iç su yüzeylerinde yüzen plastik şişelerin otomatik tespiti amacıyla Faster R-CNN mimarisine dayalı nesne tanıma modellerini sistematik bir şekilde değerlendirmiştir. FloW-Img veri kümesi üzerinde yürütülen araştırmada 1200 eğitim ve 800 test görüntüsü kullanılmış; ResNet-50, VGG-16, DenseNet-121 ve MobileNet-V3 omurgaları Stokastik Gradyan İniş (Stochastic Gradient Descent - SGD) ve Adam optimize edicileri ile çeşitli kombinasyonlarda değerlendirilmiştir. Deneysel sonuçlar, ResNet50 omurgası ile SGD optimize edicisinin kombinasyonunun en yüksek performansı gösterdiğini ve 0,180 mAP değerine ulaştığını ortaya koymuştur. Aynı

omurgada Adam optimize edicisi kullanıldığında performans 0,130 mAP değerine düşmüştür. VGG-16 tabanlı modeller ise daha düşük performans sergilemiş olup SGD ile 0,050 ve Adam optimize edicisi ile 0,048 mAP değerleri elde edilmiştir (Devi ve ark., 2024).

Akshatha ve diğ. (2022), havadan çekilmiş termal görüntülerde insan tespiti amacıyla Faster R-CNN ve SSD modellerini karşılaştırmıştır. Çalışmada ResNet-50, Inception-v2 ve MobileNet-V1 omurgaları kullanılmıştır. Çalışma farklı ortamlardan toplanmış havadan çekilmiş termal görüntüler içeren OSU ile AAU PD-T veri kümeleri üzerinde yürütülmüştür. Faster R-CNN ResNet-50 modelinin en yüksek tespit doğruluğunu gösterdiği tespit edilmiştir. Model OSU test verilerinde 1,0 mAP değerine, AAU PD-T test verilerinde ise 0,557 mAP değerine ulaşmıştır (Akshatha ve ark., 2022).

Khan ve diğ. (2023), meme kanseri histopatolojik görüntülerde mitotik çekirdeklerin tespitine yönelik Faster R-CNN tabanlı SMDetector adlı bir yöntem önermiştir. Çalışmada, ResNet-101 omurga mimarisine genişletilmiş konvolüsyon (dilated convolution) katmanları entegre edilerek modelin özellik çıkarma kapasitesi artırılmıştır. Önerilen model ICPR 2014 veri seti üzerinde CenterNet ve EfficientDet gibi yöntemlerle karşılaştırmalı olarak değerlendirilmiştir. Deneysel sonuçlar, özelleştirilmiş ResNet-101 omurgalı modelin 0,503 mAP değerine ulaşarak diğer yöntemlere kıyasla daha yüksek tespit başarısı sergilediğini göstermiştir (Khan ve ark., 2023).

Amjoud ve Amrouch (2020), nesne tespiti görevlerinde kullanılan evrimsel sinir ağı omurga mimarilerinin performansını karşılaştırmalı olarak incelemiştir. Çalışmada AlexNet, VGG-16, GoogLeNet, ResNet-50, ResNet-101 ve Inception-ResNet-V2 mimarileri ImageNet, Pascal VOC ve MS COCO veri setleri üzerinde farklı nesne tespit modelleriyle entegre edilerek değerlendirilmiştir. Pascal VOC 2007 veri seti üzerinde elde edilen test sonuçlarında, CoupleNet modeli ile entegre edilen ResNet-101 mimarisinin 0,827 mAP değerine ulaşarak en yüksek performansı sergilediği belirlenmiştir (Benali Amjoud ve Amrouch, 2020).

Wang ve diğ. (2023), uzay aracı bileşenlerinin tespitine yönelik olarak Faster R-CNN mimarisine RegNet tabanlı bir omurga yapısı entegre eden bir yöntem önermiştir. Çalışmada, 2517 görüntüden oluşan veri seti kullanılarak RegNet omurgalı modelin performansı, ResNet ve ResNeXt tabanlı Faster R-CNN varyasyonları ile SSD ve YOLOX algoritmalarıyla karşılaştırmalı olarak değerlendirilmiştir. Deneysel sonuçlar, RegNet tabanlı Faster R-CNN modelinin 0,894 mAP değeri ile en yüksek tespit başarısına ulaşmıştır. YOLOX algoritmasının 0,836 mAP, ResNeXt tabanlı modelin 0,798 mAP,

SSD algoritmasının 0,790 mAP ve ResNet tabanlı modelin 0,623 mAP değerlerine ulaştığı görülmüştür (Wang ve ark., 2023).

Kumar ve diğ. (2019), uzaktan algılama görüntülerinde nesne tespitine yönelik Faster R-CNN tabanlı bir yöntem önermiştir. Çalışmada, uçak, gemi, depolama tankı, beyzbol sahası, tenis kortu, basketbol sahası, koşu pisti, liman, köprü ve araç olmak üzere on farklı sınıfı içeren NWPU VHR-10 veri seti üzerinde ResNet-50 ve ResNet-101 omurga yapılarının performansı karşılaştırmalı olarak değerlendirilmiştir. Deneysel sonuçlar, ResNet-101 mimarisi ile entegre edilen modelin 0,871 mAP değerine, ResNet-50 omurgası kullanılan model ile yapılan deneylerde 0,857 mAP değerine ulaştığını göstermiştir (Kumar ve ark., 2019).

Zhang ve diğ. (2020), çilek bitkilerinin otomatik olarak tespit edilmesi ve sayılmasına yönelik Faster R-CNN tabanlı bir yöntem geliştirmiştir. Çalışmada, omurga mimarisi olarak VGG-16 ve ResNet-50 ağlarının performansı karşılaştırmalı olarak değerlendirilmiştir. Deneysel sonuçlar, ResNet-50 mimarisi ile yapılandırılan Faster R-CNN modelinin 0,988 mAP değerine, VGG-16 omurga mimarisini kullanan model ise 0,963 mAP değerine ulaştığını göstermiştir (Zhang ve ark., 2020).

Chen ve diğ. (2022), iletim hatlarındaki cıvata kusurlarının tespiti amacıyla Faster R-CNN mimarisi tabanlı bir yöntem önermiştir. Çalışmada omurga mimarisi olarak VGG-16, ResNet-50 ve ResNeSt ağlarının performansları kıyaslanmıştır. Deneysel sonuçlar, ResNeSt mimarisi ile entegre edilen Faster R-CNN modelinin 0,874 mAP değerine ulaşarak en yüksek tespit başarımını gösterdiğini ortaya koymuştur. ResNet-50 tabanlı modelde 0,798 mAP, VGG-16 tabanlı modelde ise 0,773 mAP değerlerine ulaşıldığı görülmüştür (Chen ve ark., 2022).

Gong ve Zhang (2021), lomber intervertebral disk deformitelerinin otomatik tespit amacıyla Faster R-CNN mimarisi tabanlı bir yöntem geliştirmiştir. Çalışmada sınırlı sayıda örnek nedeniyle veri artırma teknikleri uygulanarak veri seti genişletilmiştir ve modelin başarımını değerlendirmek üzere VGG-16, ResNet-50 ve MobileNet omurga mimarileri kıyaslanmıştır. Deneysel sonuçlar, ResNet-50 omurgası ile yapılandırılan Faster R-CNN modelinin 0,904 mAP değerine ulaştığını göstermiştir. VGG-16 omurgalı modelde 0,827 mAP, MobileNet omurgalı modelde ise 0,728 mAP değerleri elde edilmiştir (Gong ve Zhang, 2021).

Bihanda ve diğ. (2023), trafik akışının izlenmesi ve araçların otomatik tespiti amacıyla Faster R-CNN tabanlı bir yöntem önermiştir. Çalışmada 12 araç sınıfını içeren özel bir CCTV veri seti üzerinde ConvNeXt-v1 ve MobileNetV3-Large omurga

mimarilerinin performansı kıyaslamıştır. Deneysel sonuçlar, ConvNeXt-v1 omurgası ile yapılandırılan Faster R-CNN modelinin 0,818 mAP değerine ulaştığını göstermiştir. MobileNetV3-Large omurgası kullanılan model ile yapılan deneylerde 0,303 mAP değeri elde edildiği görülmüştür (Bihanda ve ark., 2023).

Popescu ve diğ. (2022), tarımsal üretim için ciddi bir tehdit oluşturan Halyomorpha halys türü böceklerin otomatik tespiti ve sınıflandırılmasına yönelik farklı omurga yapılarıyla yapılandırılmış çeşitli derin öğrenme tabanlı nesne tespit modellerini karşılaştırmalı biçimde değerlendirmiştir. Çalışmada, Maryland Biyoçeşitlilik veri seti kullanılarak YOLOv5s, SSD (MobileNet-V1, MobileNet-V2 ve ResNet-50 omurgalı), Faster R-CNN (ResNet-50 omurgalı) ve EfficientDet-D0 modellerinin performansları analiz edilmiştir. Deneysel bulgular, YOLOv5s modelinin 0,86 mAP değeri ile en yüksek doğruluğu sağlayan model olduğunu ortaya koymuştur (Popescu ve ark., 2022).

Middya ve diğ. (2021), sokak atıklarının otomatik olarak tespit edilmesi ve sınıflandırılmasına yönelik olarak Faster R-CNN tabanlı bir model geliştirmiştir. Çalışmada on farklı atık sınıfını içeren 1015 görüntüden oluşan bir veri seti kullanılarak Inception-V2 ve ResNet-101 omurga mimarilerinin performansı kıyaslanmıştır. Elde edilen sonuçlara göre Inception-V2 omurgası ile yapılandırılan Faster R-CNN modelinin 0,92 mAP değerine ulaşarak en yüksek tespit doğruluğunu sağladığını ortaya koymuştur (Middya ve ark., 2021).

Deng ve diğ. (2021), çeltik tarlalarında verim kayıplarına neden olan Alternanthera philoxeroides türü yabancı otların fide aşamasında tespit edilmesine yönelik Faster R-CNN tabanlı bir yöntem geliştirmiştir. Çalışmada transfer öğrenme yaklaşımı kullanılarak VGG-16 ve AlexNet omurga mimarilerinin performansı kıyaslanmıştır. Deneysel sonuçlar, VGG-16 omurgası ile yapılandırılan Faster R-CNN modelinin 0,824 doğruluk oranına ulaştığını göstermiştir (Deng ve ark., 2021).

Manjari ve diğ. (2023), görme engelli bireylere yönelik farklı yol dokularının tespit edilmesi için Faster R-CNN tabanlı bir yöntem geliştirmiştir. Çalışmada transfer öğrenme yaklaşımı kullanılarak Inception-V2 ve ResNet-50 omurga mimarilerinin performansı karşılaştırmalı olarak değerlendirilmiştir. Elde edilen sonuçlara göre Inception-V2 omurgasının kullanıldığı Faster R-CNN modelinin 0,305 mAP ile daha yüksek performans gösterdiğini ortaya koymuştur (Manjari ve ark., 2023).

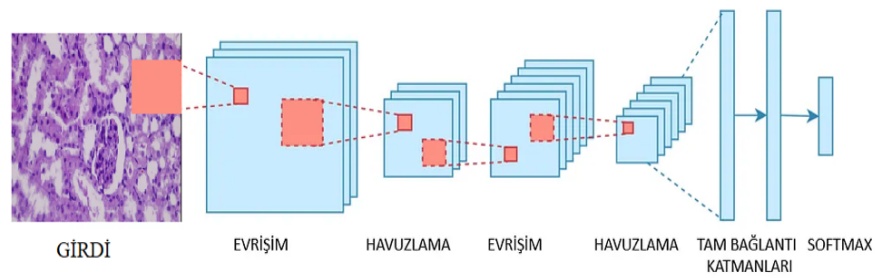
3. MATERYAL VE YÖNTEM

Bu bölümde görüntü işleme ve analiz uygulamalarında yaygın olarak kullanılan Evrişimsel Sinir Ağları (ESA) mimarisi ayrıntılı biçimde açıklanmıştır. Görüntü sınıflandırma ve nesne tespiti görevlerinde kullanılan ESA tabanlı yaklaşımlar ele alınmıştır. Görüntü sınıflandırmada kullanılan başlıca derin öğrenme modelleri ve bu modellerin mimari özellikleri incelenmiştir. Bu tez kapsamında değerlendirilen AlexNet, ResNet-50, ResNet-101, GoogLeNet (Inception-v3), VGG-16, DenseNet-121, EfficientNet-B0 ve MobileNet-V2 olmak üzere sekiz omurga mimarisi detaylı olarak anlatılmıştır.

Bu bölümde ayrıca nesne tespiti için geliştirilmiş olan derin öğrenme yaklaşımları ve bu yaklaşımların temel çalışma mekanizmaları açıklanmıştır. Çalışma kapsamında kullanılan histolojik böbrek görüntüleri veri setinin özellikleri tanıtılmıştır. Son olarak tezde uygulanan eğitim protokolü, hiper parametre seçimleri ve deneysel kurulum parametreleri açıklanmıştır. Ayrıca model performansının değerlendirilmesinde kullanılan metrikler ve karşılaştırma yöntemleri ayrıntılı bir şekilde anlatılmıştır.

3.1. Evrişimsel Sinir Ağı (ESA) Mimarisi

Evrişimsel sinir ağı mimarisinin temelleri, LeCun tarafından 1998 yılında LeNet adı verilen bir modelle atılmıştır (LeCun ve ark., 1998). ESA mimarisi, Şekil 3.1'de gösterilmiştir. Her katman alt başlıklar halinde anlatılmıştır.



Şekil 3.1. Evrişimsel sinir ağı mimarisi

ESA görüntü işleme ve bilgisayarlı görü alanında yaygın olarak kullanılan derin öğrenme mimarilerinden biridir. Bu ağlar, katmanlı bir yapıya sahip olup her katman görüntüden farklı seviyelerde özellikler çıkarmakta ve bu özellikleri sınıflandırma veya tespit işlemlerinde kullanmaktadır. Bir ESA mimarisinin temel bileşenleri ve çalışma prensibi birbiri ardına gelen katmanlar aracılığıyla gerçekleştirilmektedir. ESA mimarisinin ilk aşamasında görüntü ağa giriş verisi olarak verilmektedir. Giriş görüntüsü,

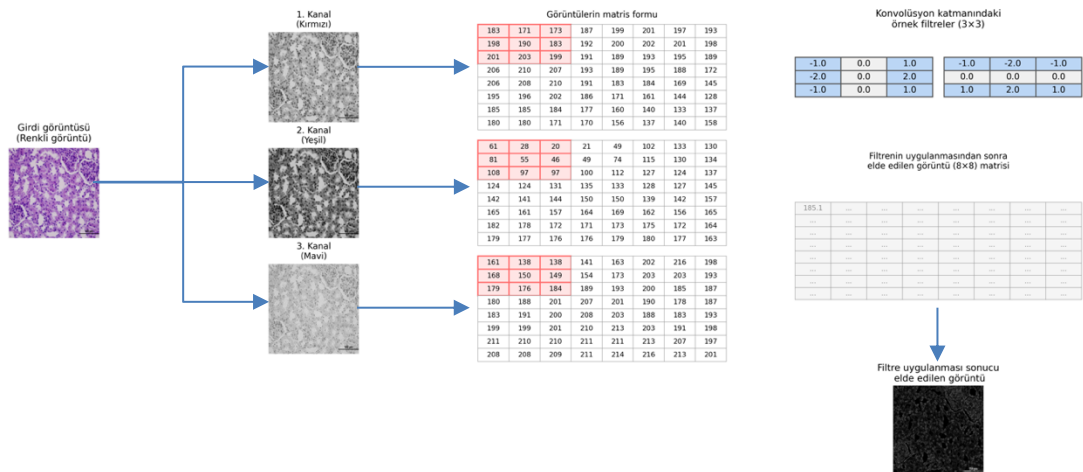
tek kanallı gri seviye formatında veya üç kanallı RGB renkli formatında olabilmektedir. Gri seviye görüntüler tek bir kanal üzerinden piksel yoğunluk değerlerini taşıırken RGB formatındaki görüntüler kırmızı, yeşil ve mavi olmak üzere üç ayrı renk kanalından oluşmaktadır. Bu katmandan sonra evrişim katmanı gelir. Bu katmanda, görüntüden anlamlı özelliklerin çıkarılması işlemi gerçekleştirilmektedir. Evrişim işlemi, çekirdek veya ağırlık matrisi olarak da adlandırılan filtreler aracılığıyla yapılmaktadır. Bu filtreler, literatürde yaygın olarak 2×2 , 3×3 , 5×5 veya 7×7 gibi farklı boyutlarda tanımlanabilmektedir. Evrişim katmanında kullanılan filtreler, eğitim süreci boyunca öğrenilen parametreleri içermektedir. Bu parametreler ağırlık görüntüdeki kenarlar, köşeler, dokular ve daha karmaşık yapısal özellikleri tanımasını sağlamaktadır. Her bir filtre, belirlenen adım genişliği (stride) değerine göre giriş görüntüsü üzerinde sistematik olarak kaydırılarak özellik haritaları oluşturmaktadır. Adım genişliği parametresi, filtrenin her hamlede kaç piksel ilerleyeceğini belirlemekte ve sonuç olarak üretilen özellik haritasının boyutlarını etkilemektedir. Evrişim işlemi sonucunda, kullanılan filtre sayısı kadar çıktı özellik haritası meydana gelmektedir. Her bir özellik haritası, ilgili filtrenin tespit ettiği spesifik özellikleri temsil etmektedir. Evrişim işleminin ardından aktivasyon katmanı gelmektedir. Bu katmanın temel amacı, ağırlık doğrusal olmayan özellikler kazandırarak daha karmaşık örüntülerin öğrenilmesini sağlamaktır. Aktivasyon fonksiyonları arasında en yaygın kullanılanı Düzleştirilmiş Doğrusal Birim (Rectified Linear Unit-ReLU) fonksiyonudur. ReLU aktivasyon fonksiyonu, özellik haritalarında bulunan negatif değerleri sıfıra dönüştürürken, pozitif değerleri herhangi bir değişiklik yapmadan bir sonraki katmana iletmektedir. Aktivasyon katmanını takiben havuzlama katmanı uygulanmaktadır. Havuzlama katmanının işlevi özellik haritalarının boyutlarını küçülterek hesaplama yükünü azaltmaktır. Bir veya birden fazla evrişim, aktivasyon ve havuzlama katmanı dizisinin ardından, tam bağlantılı (fully connected) katman yapısına geçilmektedir. Bu katmanda, önceki katmanlardan gelen tüm nöronlar birbirleriyle tam bir bağlantı kurmaktadır. Tam bağlantılı katmanın görevi, evrişim katmanlarında çıkarılan düşük ve orta seviye özelliklerden yüksek seviyeli özellik kombinasyonlarını öğrenmektir. Bu katman, geleneksel yapay sinir ağlarında kullanılan yapıya benzer şekilde çalışmakta ve ağırlık sınıflandırma kararını vermesinde kritik rol oynamaktadır. ESA mimarisinin son katmanını oluşturan sınıflandırma katmanı, önceki katmanlardan gelen öznitelik haritalarını kullanarak nihai sınıf tahminini gerçekleştirir. Bu katmandan gelen çıktıları Softmax aktivasyon fonksiyonu uygulanarak olasılık dağılımına dönüştürme işlemi yapılmaktadır.

3.1.1. Giriş katmanı

Giriş katmanı, adından da anlaşılacağı üzere modelin ilk bileşenini oluşturmaktadır. Bu katmanda ham veri doğrudan ağı iletilir ve ağı işlemeye başladığı ilk aşamayı temsil eder. Tasarlanan modelin genel başarımı açısından bu katmana sunulan verinin boyutu oldukça önemlidir. Giriş görüntüsünün boyutunun yüksek seçilmesi daha fazla bellek gereksinimi, uzun eğitim süresi ve test aşamasında her bir görüntü için daha fazla işlem süresi anlamına gelmektedir. Bununla birlikte daha yüksek çözünürlüklü girdiler kullanılması modelin öğrenme kapasitesini artırarak genel doğruluk oranına olumlu katkı sağlayabilmektedir. Diğer yandan giriş boyutunun küçük seçilmesi, bellek kullanımını azaltırken eğitim süresini de kısalmaktadır. Ancak bu durum ağı öğrenme kapasitesini sınırlandırabileceği için performans düşüklüğüne neden olabilir. Bu nedenle görüntü tabanlı derin öğrenme uygulamalarında; model derinliği, donanımsal kaynakların yeterliliği ve hedeflenen doğruluk oranı göz önünde bulundurularak uygun bir giriş boyutu seçimi yapılmalıdır (İnik ve Ülker, 2017).

3.1.2. Evrişim katmanı

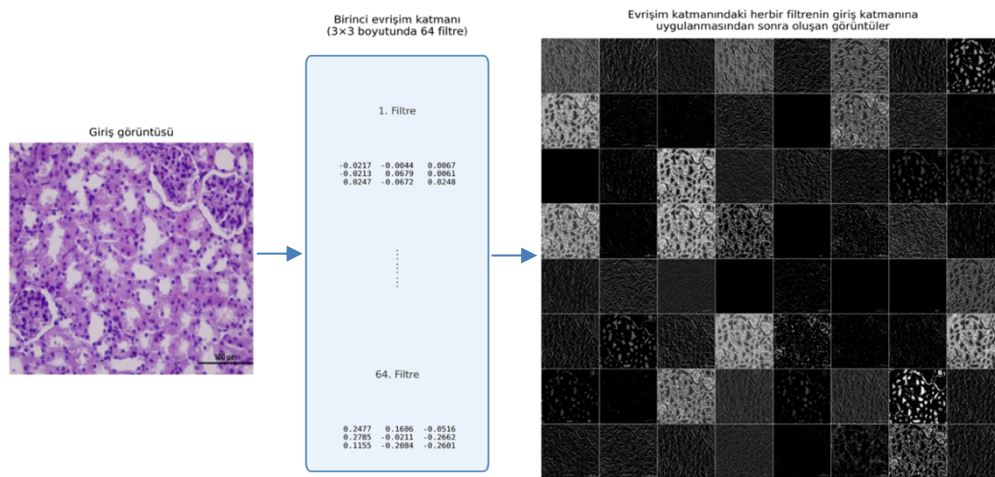
Evrişim katmanı, ESA mimarisinin temel bileşenlerinden biridir. Bu katman görüntü verileri üzerinde otomatik özellik çıkarımı gerçekleştiren ana yapıyı oluşturur. Başka bir ifadeyle giriş verisindeki önemli örüntülerin tespiti bu katmanda sağlanır. Evrişim işlemi, belirli boyutlara sahip filtrelerin giriş görüntüsü üzerinde adım adım kaydırılması ve her konumda matematiksel olarak evrişim işleminin uygulanması prensibine dayanır. Bu işlemler sonucunda aktivasyon (özellik) haritası adı verilen çıktılar elde edilir. Aktivasyon haritaları, kullanılan her bir filtrenin görüntü üzerinde hangi bölgelerde belirgin özellikleri yakaladığını gösterir. Böylece ağı, özellik çıkarımı sürecinde verinin hangi bölgelerinin daha önemli olduğunu öğrenir.



Şekil 3.2. Görüntü üzerinde filtrenin uygulanması

3×3 boyutundaki bir filtrenin görüntü üzerine uygulanması Şekil 3.2’de verilmiştir. Histolojik görüntüler renkli görüntü olup 3 kanaldan oluşmaktadır. Bu kanallardaki görüntülerin bir parçası matris formunda ifade edilmiştir. Filtre matrisi görüntünün ilgili piksel değerleriyle çarpılıp toplamları ile elde edilir. Bu işlem her bir görüntü için yapıldıktan sonra elde edilen değerlerin ortalaması alınarak evrişim sonucu oluşan görüntünün ilgili pikseline yazılır. Bu filtreleme işlemi belli bir adım sayısına göre görüntü üzerinde uygulanır.

Şekil 3.3’te giriş görüntüsüne VGG-16 modelinin birinci evrişim katmanındaki filtre uygulanması sonucu oluşan özellik haritasına ait örnek bir gösterim verilmiştir. VGG-16 modelinin birinci evrişim katmanında 3×3 boyutunda 64 filtre olduğundan her bir filtrenin giriş görüntüsüne uygulanması sonucu farklı görüntüler elde edilmiştir. Şekil 3.3’e bakıldığında her bir filtrenin giriş görüntüsüne etkisinin farklı olduğu görülmektedir.



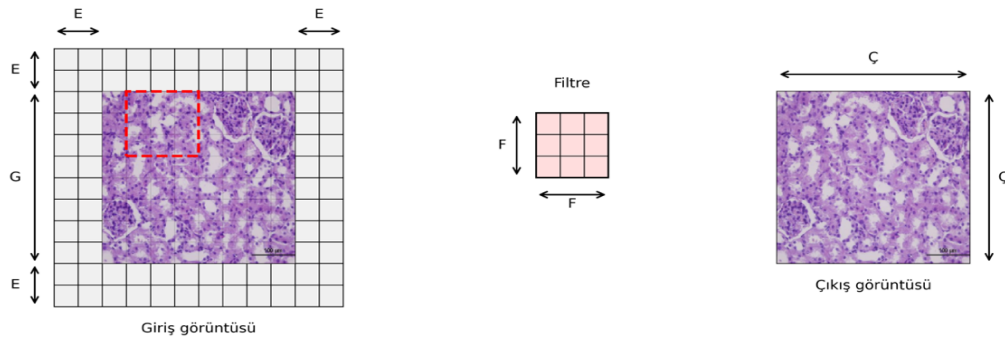
Şekil 3.3. VGG-16 modelinin birinci evrişim katmanının örnek görsel uygulanması

Evrişimsel sinir ağının ilk katmanında gerçekleşen özellik çıkarma süreci Şekil 3.3’te detaylı olarak gösterilmiştir. Sol tarafta 224×224 piksel boyutundaki giriş görüntüsü yer almaktadır. Ortada birinci evrişim katmanının yapısı şematik olarak sunulmuştur. Bu katman 3×3 boyutunda 64 farklı filtre içermektedir. Görselleştirme amacıyla katmandaki ilk filtre ve son 64 nolu filtrenin ağırlık değerleri sayısal olarak gösterilmiştir. Her bir filtre RGB giriş kanalları üzerinden ortalama alınarak 3×3 matris formunda sunulmuştur. Sağ tarafta ise 64 filtrenin her birinin giriş görüntüsüne uygulanması sonucunda oluşan özellik haritaları 8×8 ızgara (grid) formatında düzenlenmiştir. Bu ızgara yapısı, katmanın çıktı olarak ürettiği 64 farklı özellik haritasının tamamını eşzamanlı olarak görselleştirmektedir. Her bir özellik haritası ReLU aktivasyon fonksiyonu uygulandıktan sonra minimum-maksimum normalizasyonu ile

[0,1] aralığına ölçeklendirilerek gri tonlamalı görüntüler halinde sunulmuştur. Bu görselleştirme tek bir evrişim katmanının giriş görüntüsünden çok sayıda farklı özellik çıkarabileceğini ve her filtrenin görüntünün farklı yapısal özelliklerine duyarlı olduğunu göstermektedir.

Evrişim sonrasında oluşan özellik haritasının boyutları filtre (çekirdek) boyutu, adım (stride) ve dolgu (padding) parametrelerine bağlıdır ve bu ilişki Eşitlik 3.1 ile verilmektedir. Eşitlikte yer alan değişkenlerin tanımı ve konumları Şekil 3.4'te şematik olarak gösterilmiştir.

$$Ç = \frac{G - F + 2 * E}{A} + 1 \quad (3.1)$$



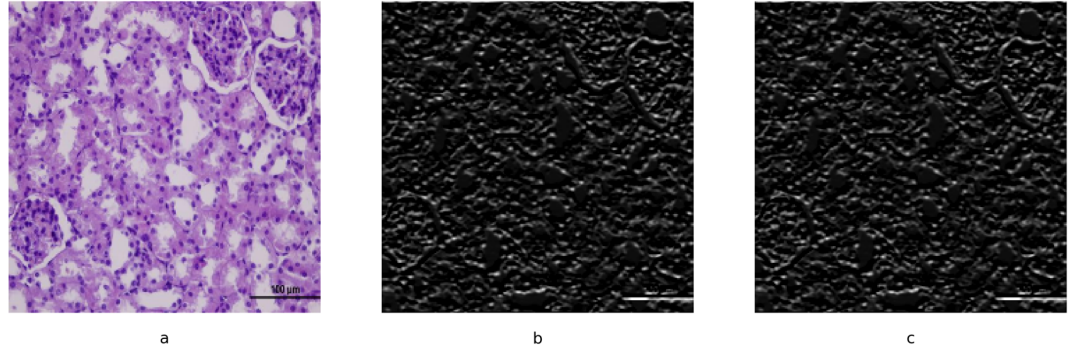
Şekil 3.4. Evrişim katmanında boyut hesabı gösterimi

G, giriş görüntüsünün boyutunu; F, uygulanan çekirdeğin (filtrenin) boyutunu; E, giriş matrisinin kenarlarına eklenen sıfır dolgu miktarını; A ise filtrenin giriş üzerinde ilerleme miktarını belirleyen adım (stride) değerini ifade eder.

3.1.3. Düzeltilmiş doğrusal birim (ReLU) katmanı

ReLU, evrişimsel ve tam bağlantılı katmanların ardından kullanılan bir aktivasyon katmanıdır. Evrişim katmanında yapılan işlemler matematiksel olarak doğrusal bir karakter taşır. Bu katman, ağırlı doğrusal olmayan bir yapıya dönüştürülmesi için uygulanır. Böylece modelin daha zengin ve karmaşık örüntüleri yakalayabilmesini sağlar. ReLU katmanına gelen özellik matrisindeki tüm negatif değerler sıfır ile değiştirilirken pozitif değerler herhangi bir değişikliğe uğramadan doğrudan çıkışa aktarılır.

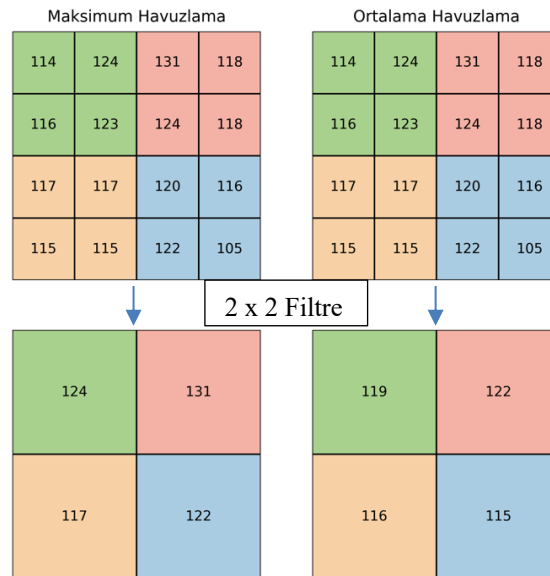
Evrişimsel sinir ağlarında aktivasyon fonksiyonunun etkisi Şekil 3.5'te gösterilmiştir. Bu görselleştirme için VGG-16 modelinin ilk evrişim katmanı kullanılmış ve 64 özellik haritası arasından varyansı en yüksek olan filtre otomatik olarak seçilmiştir. ReLU katmanının giriş görüntüsü üzerine yaptığı etki Şekil 3.5'te gösterilmiştir. Şekil 3.5 (a) giriş görüntüsü, (b) evrişim katmanında 3x3'lük bir filtre uygulanması sonucu oluşan görüntü ve (c) ise ReLU katmanından çıkan görüntüyü göstermektedir.



Şekil 3.5. (a) Giriş görüntüsü, (b) VGG-16 modelinin ilk evrişim katmanının giriş görüntüsüne uygulandıktan sonra çıkan görüntü, (c) ReLu katmanından çıkan görüntü

3.1.4. Havuzlama Katmanı

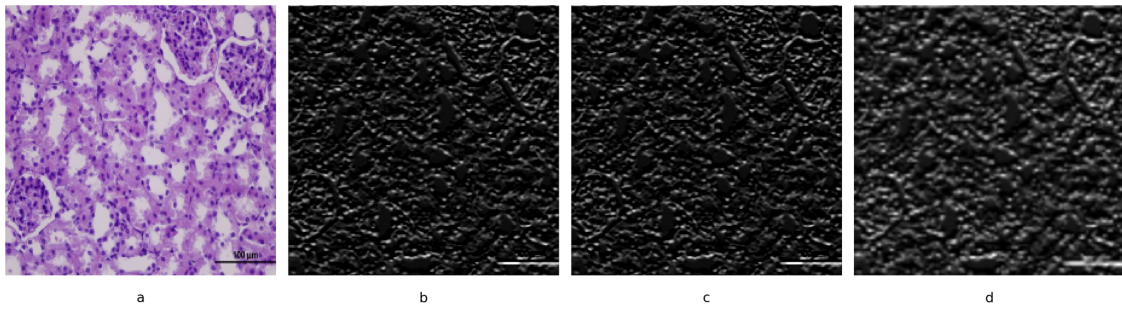
Havuzlama katmanı özellik haritalarının boyutunu küçültmek hesaplama yükünü azaltmak, ağıın performansını iyileştirmek ve modelin genelleme kabiliyetini artırmak amacıyla kullanılmaktadır. Bu katman, bir önceki katmandan elde edilen özellik haritaları üzerinde belirlenen pencere boyutu ve adım genişliği kullanılarak boyut küçültme işlemi gerçekleştirmektedir. Boyut küçültme sürecinde, özellik haritalarının üzerine belirli aralıklarla pencereler yerleştirilmekte ve her pencere içindeki değerlerden maksimum, minimum ya da ortalama seçilmektedir. Bu işlem Şekil 3.6'da gösterilmektedir. Yapılan bu işlem sonucunda görüntünün genişlik ve yükseklik boyutları azalırken kanal sayısı sabit kalmaktadır. Havuzlama katmanında öğrenilen parametre bulunmamakta, yalnızca boyut küçültme gerçekleştirilmektedir. Boyut küçültme sırasında bazı detay bilgileri kaybedilmektedir.



Şekil 3.6. Havuzlama katmanı örneği

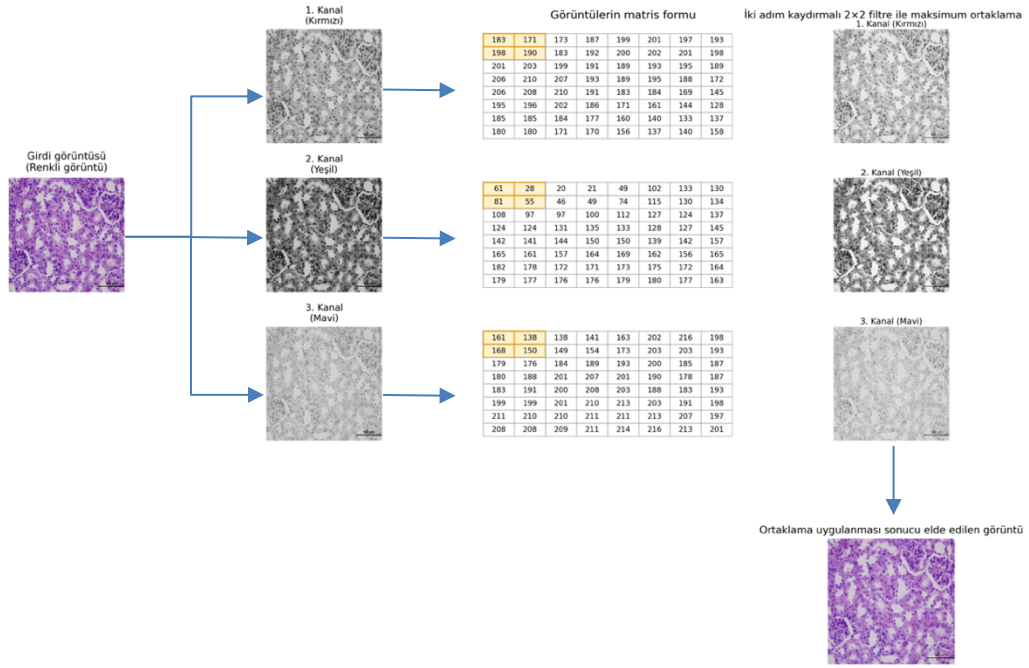
Evrişimsel sinir ağıında kullanılan ardışık katmanların giriş görüntüsü üzerine yaptığı etkiler Şekil 3.7'de gösterilmiştir. Görselleştirme için VGG-16 modelinin ilk

evrişim katmanını kullanılmış ve tek bir özellik haritası izlenmiştir. Şekil 3.7(a) orijinal giriş görüntüsünü, Şekil 3.7(b) birinci evrişim katmanında 3×3 boyutundaki bir filtrenin aktivasyon öncesi (pre-activation) çıktısını göstermektedir. Şekil 3.7(c)'de ReLU aktivasyon fonksiyonu uygulanması sonrası elde edilen görüntü sunulmuştur. ReLU fonksiyonu, negatif değerleri sıfırlayarak özellik haritalarında doğrusal olmayan dönüşüm sağlamaktadır. Son olarak Şekil 3.7(d), ortalama havuzlama (average pooling) katmanından çıkan görüntüyü göstermektedir. Havuzlama katmanında 2×2 filtre boyutu ve 2 adım kayma değeri kullanılmıştır. Veri boyutundaki küçülme nedeniyle havuzlama sonrası görüntü daha az uzamsal bilgi içermekte, ancak görselleştirme amacıyla orijinal boyutlara yeniden ölçeklendirilmiştir.



Şekil 3.7. Bazı (a) Giriş görüntüsü, (b) VGG16 modelinin ilk evrişim katmanının giriş görüntüsüne uygulandıktan sonra çıkan görüntü, (c) ReLu katmanından çıkan görüntü (d) Havuzlama katmanından çıkan görüntü

Maksimum havuzlama katmanının giriş görüntüsü üzerine yaptığı etki Şekil 3.8'de gösterilmiştir. Giriş görüntüsü öncelikle RGB kanallarına (kırmızı, yeşil ve mavi) ayrılmış ve her kanal üzerinde ayrı ayrı işlem uygulanmıştır. Görüntülerin matris formunda gösteriminde her kanaldan örnek olarak 8×8 boyutunda bölgeler seçilmiş ve piksel değerleri sayısal olarak sunulmuştur. Havuzlama katmanında 2×2 filtre boyutu ve 2 adım kayma sayısı kullanılarak maksimum havuzlama yöntemi uygulanmıştır. Havuzlama işlemi her 2×2 piksel bloğundan maksimum değer seçilmesi prensibine dayanmaktadır. Veri boyutundaki azalma nedeniyle havuzlama sonrası görüntü orijinal giriş görüntüsüne kıyasla daha düşük çözünürlükte ve daha az piksel içermektedir.



Şekil 3.8. Giriş görüntüsüne maksimum havuzlama işleminin uygulanması

3.1.5. Tam Bağlantılı Katman

Evrişimsel sinir ağı mimarilerinde ardışık evrişim, aktivasyon ve havuzlama katmanları aracılığıyla elde edilen yüksek seviyeli özniteliklerin nihai sınıflandırma veya regresyon çıktısına dönüştürülmesi tam bağlantılı katman vasıtasıyla gerçekleştirilmektedir. Bu katman türü, klasik yapay sinir ağlarındaki nöronlara benzer şekilde bir önceki katmanda bulunan her bir nöronla doğrudan bağlantı kurmaktadır. Tam bağlantılı katmana geçiş yapılmadan önce, özellik çıkarma aşamasının son katmanından gelen çok boyutlu tensör yapısının tek boyutlu bir vektöre dönüştürülmesi gerekmektedir. Bu dönüşüm düzleştirme işlemi olarak adlandırılmakta ve ağı ilerleyen katmanlarında bilgi akışının sağlanmasını mümkün kılmaktadır. Tam bağlı katmanlar, evrişimsel sinir ağı mimarisinde parametrelerin yoğun olarak kullanıldığı bölgelerden birini oluşturmaktadır. Dolayısıyla hesaplama yükü ve eğitim süresini sınırlamak için tam bağlantılı katmanların parametre sayısı özenle belirlenmelidir.

3.1.6. Sınıflandırma Katmanı

Evrişimsel sinir ağı mimarisinin son aşamasını oluşturan sınıflandırma katmanı, önceki katmanlardan elde edilen özellik vektörünü belirli sınıflara atayan işlevi yerine getirmektedir. Bu katmandaki çıkış nöron sayısı, problemdeki sınıf sayısına eşit olmaktadır. Örneğin iki sınıflı bir problemde iki çıkış nöronu, beş sınıflı bir problemde beş çıkış nöronu kullanılmaktadır. Sınıflandırma katmanında kullanılan aktivasyon fonksiyonu, modelin nihai sonucunun şekli üzerinde doğrudan etkili olmaktadır. Bu amaçla çeşitli aktivasyon fonksiyonları bulunmakla birlikte, Softmax fonksiyonu görüntü

sınıflandırma görevlerinde en yaygın olarak tercih edilmektedir. Softmax fonksiyonu, çıkış katmanındaki her değeri $[0, 1]$ arasında normalize ederek olasılık değerlerine dönüştürmektedir. Bu sayede her sınıfın olasılığı açık ve anlaşılır bir şekilde ifade edilmekte, tüm olasılıkların toplamı 1 olacak şekilde düzenlenmektedir. Sınıflandırma kararı, üretilen olasılık değerleri içinden en yüksek değere sahip sınıfın seçilmesi prensibine dayanmaktadır. Bu yaklaşım hem sonuçların anlaşılabilir olmasını sağlamakta hem de modelin her tahmin için ne kadar emin olduğunu gösteren güven puanı vermektedir.

3.2. Evrişimsel Sinir Ağı (ESA) Modelinin Eğitim Parametreleri

ESA modellerinin eğitimi sırasında kullanılan devir sayısı (epoch), momentum katsayısı, paket boyutu (mini-batch size), öğrenme hızı (learning rate), kullanılan optimizasyon yöntemi gibi parametreler eğitim parametreleri arasında yer almaktadır.

3.2.1. Devir sayısı (Epoch)

Devir sayısı (epoch), eğitim verisinin tamamının model üzerinde kaç kez işleneceğini gösteren temel bir hiper parametredir. Her bir devir, veri setindeki tüm örneklerin paket boyutu (mini-batch size) kadar gruplar halinde modele sunulmasını ve ileri-geri yayılım mekanizmasıyla ağırlıkların güncellenmesini ifade etmektedir.

3.2.2. Optimizasyon algoritması

Derin öğrenme modellerinin eğitim süreci, doğrusal olmayan bir optimizasyon probleminin çözümü olarak ele alınmaktadır. Bu süreçte temel amaç, hata fonksiyonunun minimum değerine ulaşabilecek optimal ağırlık değerlerini bulmaktır. Global minimum noktasına ulaşmak için farklı optimizasyon algoritmaları kullanılmaktadır. Problem karmaşıklığı arttıkça model derinliği, parametre sayısı, optimizasyon yöntemi ve hesaplama gereksinimleri de buna paralel olarak değişir. Stokastik Gradyan İnişi (SGD) ve momentum tabanlı türevi (SGDM), derin sinir ağlarının eğitiminde sıklıkla tercih edilen temel algoritmalar arasında yer almaktadır. SGD özellikle momentum katsayısıyla birlikte kullanıldığında güçlü genelleme özellikleri sunmakta ve yerel minimumlarda takılma riskini azaltmaktadır.

3.2.3. Öğrenme hızı (Learning rate)

Öğrenme hızı, model ağırlıklarının geriye yayılım algoritması ile her iterasyonda ne ölçüde güncelleneceğini belirleyen en kritik hiper parametrelerden biridir. Bu parametre, modelin genel performansını doğrudan etkilemektedir. Ağırlık güncellemesi sürecinde hata fonksiyonunun türevi alınarak hesaplanan gradyan değeri öğrenme hızı katsayısı ile çarpılır. Elde edilen bu değer mevcut ağırlıktan çıkarılarak yeni ağırlık değeri

hesaplanır. Öğrenme hızının büyüklüğü, eğitim sürecinin hem hızını hem de başarımlarını belirlemektedir. Yüksek öğrenme hızı değerleri kullanıldığında model hata fonksiyonunun minimumuna daha hızlı yakınsayabilir. Ancak bu durum optimal ağırlık değerlerinin atlanmasına ve küresel minimum noktasının kaçırılmasına neden olabilir. Düşük öğrenme hızı değerlerinde ise ağırlıklar daha küçük adımlarla güncellenir. Bu da eğitim süresinin önemli ölçüde uzamasına ve yerel minimum bölgelerinde takılma riskine yol açabilmektedir. Öğrenme hızı, eğitim süreci boyunca sabit tutulabileceği gibi dinamik olarak da değiştirilebilir. Sabit öğrenme hızı kullanımı basit ve hesaplama açısından düşük maliyetlidir. Ancak öğrenme hızı planlayıcıları gibi dinamik stratejiler genellikle daha etkili sonuçlar vermektedir. Bu stratejiler, eğitim sürecinin belirli aşamalarında öğrenme hızını artırarak veya azaltarak modelin daha dengeli ve kararlı öğrenmesini sağlar. Bu çalışmada kullanılan kosinüs çürüme ile ısınma (cosine decay with warmup) yaklaşımı başlangıçta düşük öğrenme hızı ile başlar. Ardından belirli bir devir sayısına kadar kademeli olarak artırılır. Daha sonra kosinüs eğrisi boyunca sistematik bir şekilde azaltılır. Bu yöntem, eğitim süreci boyunca optimal öğrenme dengesini kurmaya yardımcı olur.

3.2.4. Momentum katsayısı

Momentum katsayısı, modelin ağırlık güncellemeleri sırasında önceki iterasyonlardaki gradyan bilgisini dikkate alarak öğrenme sürecinin daha kararlı ve hızlı gerçekleşmesini sağlayan bir hiper parametredir. Geriye yayılım algoritması her iterasyonda hata fonksiyonunun gradyanına göre ağırlıkları günceller. Momentum katsayısı genellikle 0,9 olarak seçilmektedir.

3.2.5. Paket boyutu (Mini-batch size)

Paket boyutu, modelin her bir eğitim adımında ağırlıklarının güncellenmesi için kullanılan örnek sayısını ifade eden bir hiper parametredir. Eğitim verisinin tamamının tek seferde değil, belirli büyüklükteki küçük gruplara (mini-batch) bölünerek işlenmesi hem hesaplama verimliliğini artırmakta hem de öğrenme sürecinin daha kararlı ilerlemesini sağlamaktadır. Paket boyutu değeri belirlenirken modelin yapısı, veri setinin büyüklüğü, eğitim hızı ve kullanılan donanımın bellek kapasitesi dikkate alınmalıdır.

3.2.6. Erken durdurma (Early stopping)

Erken durdurma, derin öğrenme modellerinin aşırı öğrenmesini (overfitting) önlemek amacıyla kullanılan bir düzenleme tekniğidir. Eğitim sürecinde modelin doğrulama verisi üzerindeki performansı sürekli izlenir ve belirli bir süre boyunca doğrulama kaybında iyileşme gözlenmezse eğitim durdurulur. Bu yöntem modelin eğitim

verisine aşırı uyum sağlamasını engelleyerek genelleme yeteneğini artırır. Early stopping uygulamalarında genellikle patience (sabır) adı verilen bir parametre kullanılır. Bu parametre doğrulama performansında iyileşme görülmediği ardışık devir sayısını belirler. Eğitim tamamlandığında, son iterasyondaki ağırlıklar yerine doğrulama başarımının en yüksek olduğu noktadaki ağırlıklar tercih edilir ve en iyi kaydedilmiş parametrelere geri dönülür. Bu sayede model, gereksiz uzun eğitimlerden kaçınılırken en yüksek doğrulama başarımına sahip ağırlıklarla kullanıma hazır hâle getirilmiş olur.

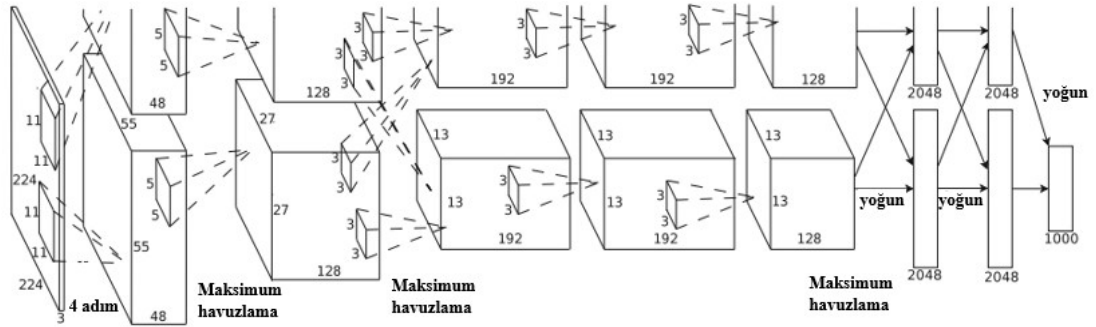
3.3. Nesne Sınıflandırmada Kullanılan Derin Öğrenme Mimarileri

3.3.1. AlexNet

AlexNet, 2012 yılında Krizhevsky ve arkadaşları tarafından geliştirilmiştir (Krizhevsky ve ark., 2012). Aynı yıl ILSVRC (ImageNet Large Scale Visual Recognition Challenge) yarışmasında ILSVRC 2012 test seti üzerinde %15.3 top-5 hata oranı ile birinci olmuştur. Girdi katmanı 3 kanallı 224x224x3 boyutunda görüntülerden oluşmaktadır. AlexNet mimarisi, 5 evrişim katmanı ve 3 tam bağlı katman olmak üzere 8 katman içermektedir. İlk evrişim katmanında 11x11 boyutunda 96 adet filtre ve 4 birimlik adım (stride) değeri kullanılmaktadır. Aktivasyon için ReLU ve çıkış için Softmax fonksiyonları kullanılmıştır. Aşırı uyum sorununu azaltmak için sönümleme (dropout) metodu uygulanmıştır. Eğitim sırasında bazı nöronların rastgele devre dışı bırakılmasının, modelin genelleme yeteneğini artırdığı bilinmektedir. Veri artırma (data augmentation) teknikleri ile eğitim verisinin çeşitliliği artırılmış, böylece modelin genelleme yeteneği güçlendirilmiştir. Çıkış katmanında 1000 sınıf için Softmax aktivasyon fonksiyonu kullanılarak çok sınıflı sınıflandırma gerçekleştirilmektedir. AlexNet mimarisi Şekil 3.9'da ve Çizelge 3.1'de verilmiştir.

Çizelge 3.1. AlexNet mimarisi

Tür	Filtre Sayısı	Filtre Boyutu/Adım Sayısı	Dolgu Boyutu	Çıktı Boyutu
Evrişim Katmanı	96	11 x 11 / 4		55 x 55 x 96
Maksimum Havuzlama Katmanı		3 x 3 / 2		27 x 27 x 96
Evrişim Katmanı	256	5 x 5 / 1	2	27 x 27 x 256
Maksimum Havuzlama Katmanı		3 x 3 / 2		13 x 13 x 256
Evrişim Katmanı	384	3 x 3 / 1	1	13 x 13 x 384
Evrişim Katmanı	384	3 x 3 / 1	1	13 x 13 x 384
Evrişim Katmanı	256	3 x 3 / 1	1	13 x 13 x 256
Maksimum Havuzlama Katmanı		3 x 3 / 2		6 x 6 x 256
Tam Bağlantılı Katman				4096
Tam Bağlantılı Katman				4096
Tam Bağlantılı Katman				1000
Softmax Aktivasyon Fonksiyonu				

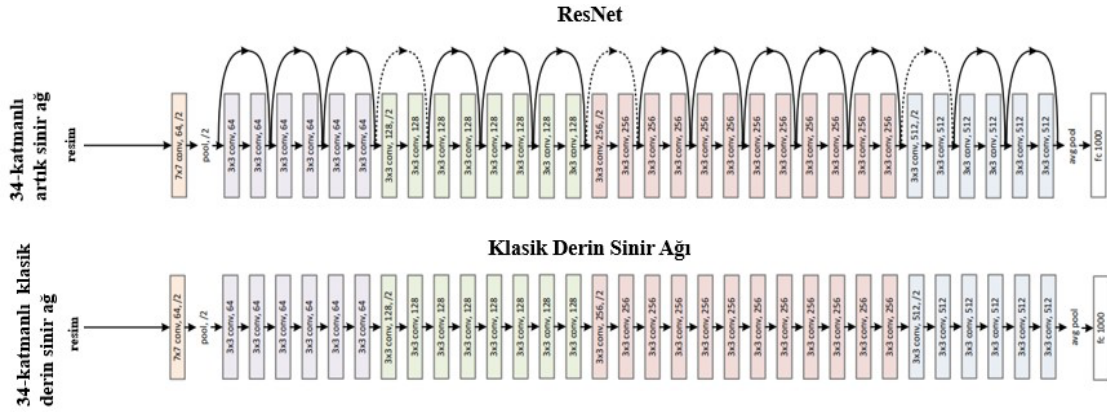


Şekil 3.9. AlexNet mimarisi(Krizhevsky ve ark., 2012)

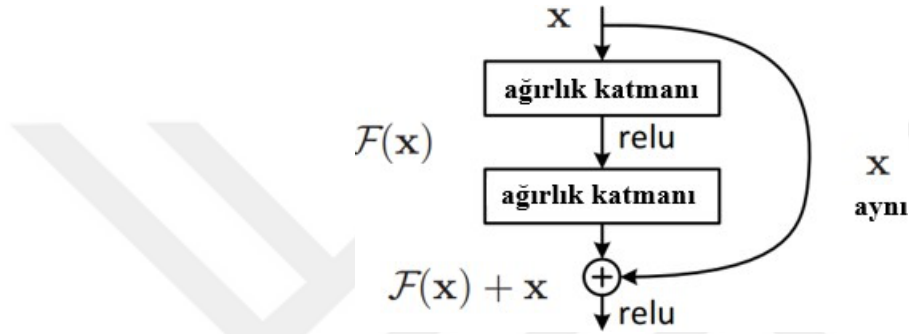
3.3.2. ResNet

ResNet, 2016 yılında Microsoft Research'ten He ve ekibi tarafından geliştirilmiştir (He ve ark., 2016). Aynı yıl ILSVRC (ImageNet Large Scale Visual Recognition Challenge) yarışmasında %3.57 top-5 hata oranı ile birinci olmuştur. Ağ, 224x224 boyutunda renkli görüntüleri giriş olarak alır. ResNet mimarisi, derin sinir ağlarında karşılaşılan gradyan yok olması (vanishing gradient) ve performans kaybı (degradasyon) sorunlarını çözmek için artık öğrenme (residual learning) yaklaşımını kullanmaktadır. ResNet ailesi, farklı derinliklerde birden çok sürümden oluşur: ResNet-18 (18 katman), ResNet-34 (34 katman), ResNet-50 (50 katman), ResNet-101 (101 katman) ve ResNet-152 (152 katman).

ResNet mimarisinin temel yapı taşı olan artık bloklar (residual blocks), bir katmanın girişini, birkaç katman sonrasına doğrudan ileten atlama bağlantıları (skip connections) veya kısayol bağlantıları (shortcut connections) içermektedir. Bu bağlantılar, gradyan akışını iyileştirerek derin ağların etkili bir şekilde eğitilmesini sağlamaktadır. Mimari, ardışık 3x3 evrişim filtreleri kullanılarak tasarlanmıştır ve her evrişim katmanından sonra toplu normalizasyon (batch normalization) ve ReLU aktivasyon fonksiyonu uygulanmaktadır. Artık bloklarda, boyut farklılıklarını gidermek için 1x1 evrişimler ve 2 birimlik adım boyutu (stride) olan katmanlar kullanılmaktadır. Ağın son kısmında, global ortalama havuzlama (global average pooling) işlemi ile özellik haritaları tek bir vektöre dönüştürülmekte ve ardından 1000 sınıflı softmax katmanına bağlanmaktadır. Eğitim sürecinde, veri artırma, ağırlık azaltımı (weight decay) ve momentum gibi teknikler kullanılmıştır. ResNet mimarisi Şekil 3.10'da ve artık blok yapısı Şekil 3.11'de verilmiştir.



Şekil 3.10. Microsoft ResNet ilk 34 katmanının ağ mimarisi (He ve ark., 2016)



Şekil 3.9. Residual blok(He ve ark., 2016)

3.3.3. GoogLeNet

GoogLeNet, 2014 yılında Szegedy ve Google araştırmacıları tarafından geliştirilmiştir (Szegedy ve ark., 2015). Aynı yıl ILSVRC (ImageNet Large Scale Visual Recognition Challenge) yarışmasında %6.7 top-5 hata oranı ile birinci olmuştur. Ağ, 224x224 boyutunda renkli (RGB) görüntüleri giriş olarak alır. GoogLeNet mimarisi toplamda 22 ağırlıklı katmandan oluşmakta ve yaklaşık 6.8 milyon parametre içermektedir. Modelin temelini oluşturan Inception modülleri, farklı boyutlardaki (1x1, 3x3, 5x5) evrişim filtrelerini ve havuzlama işlemlerini paralel olarak uygulayarak çoklu ölçekte özellik çıkarımı yapmaktadır. 1x1 evrişimler, hesaplama maliyetini azaltmak için boyut indirgeme katmanları (bottleneck) olarak kullanılmaktadır. Evrişim işlemleri sonrası ReLU aktivasyon fonksiyonu uygulanmakta, ara katmanlarda ise maksimum havuzlama işlemi ile özellik haritalarının boyutları küçültülmektedir. Model, gradyan yok olması problemini azaltmak ve düzenleme sağlamak amacıyla farklı seviyelere yardımcı sınıflandırıcılar (auxiliary classifiers) eklemiştir. Bu yardımcı sınıflandırıcılar sadece eğitim sırasında kullanılmakta ve test aşamasında devre dışı bırakılmaktadır. Ağın son kısmında tam bağlı katman sayısını azaltmak amacıyla global ortalama havuzlama (global average pooling) işlemi uygulanmakta ve ardından 1000 sınıflı bir Softmax fonksiyonu

ile sınıflandırma gerçekleştirilmektedir. Eğitim sırasında sönümleme (dropout) ve veri artırma yöntemleri kullanılmıştır.

Inception v3, orijinal GoogLeNet mimarisinin geliştirilmiş versiyonu olarak Szegedy ve arkadaşları tarafından 2016 yılında ortaya konulmuştur (Szegedy ve ark., 2016). Bu mimari, temel Inception yaklaşımını korurken hesaplama verimliliği ve doğruluk açısından önemli iyileştirmeler getirmektedir. Inception v3 mimarisi 48 evrişimsel katmandan oluşan derin bir yapıya sahiptir ve 299x299 boyutunda görüntü girdisi almaktadır. Inception-v3 mimarisinin en önemli yeniliklerinden biri, büyük evrişim filtrelerinin faktörize evrişimler kullanılması ile değiştirilmesidir. Örneğin, 5x5 boyutundaki bir filtre iki ardışık 3x3 filtre ile, 7x7 boyutundaki bir filtre ise asimetrik 1x7 ve 7x1 filtrelerin kombinasyonu ile yer değiştirilmiştir. Bu faktörize evrişim yaklaşımı, parametre sayısını ve hesaplama gereksinimlerini önemli ölçüde azaltırken ağın ifade gücünü korumaktadır. Inception v3'te kullanılan diğer önemli teknikler arasında tüm evrişim katmanlarına uygulanan toplu normalizasyon (batch normalization), etiket düzgünleştirme (label smoothing) düzenlileştirme ve Karekök Ortalama Yayılım (Root Mean Square Propagation -RMSProp) optimizasyon algoritması bulunmaktadır. Bu teknikler modelin eğitim kararlılığını artırmakta ve genelleme performansını iyileştirmektedir. ImageNet veri seti üzerinde Inception v3, %21.2 top-1 ve %5.6 top-5 hata oranları elde etmiştir. Inception v3 mimarisinin diyagramı Çizelge 3.2'de gösterilmektedir.

Çizelge 3.2. GoogLeNet (Inception-v3) mimarisi

Tür	Yama Boyutu / Adım Sayısı	Çıktı Boyutu	Derinlik	1 x 1 Filtre	3 x 3 İndirgeme	3 x 3 Filtre	5 x 5 İndirgeme	5 x 5 Filtre	Havuzlama	Parametre Sayısı	İşlem Sayısı
Evrişim Katmanı	7 x 7 / 2	112 x 112 x 64	1							2.7K	34M
Maksimum Havuzlama Katmanı	3 x 3 / 2	56 x 56 x 64	0								
Evrişim Katmanı	3 x 3 / 1	56 x 56 x 192	2		64	192				112K	360M
Maksimum Havuzlama Katmanı	3 x 3 / 2	28 x 28 x 192	0								
Inception Modülü(3a)		28 x 28 x 256	2	64	96	128	16	32	32	159K	128M
Inception Modülü(3b)		28 x 28 x 480	2	128	128	192	32	96	64	380K	304M
Maksimum Havuzlama Katmanı	3 x 3 / 2	14 x 14 x 480	0								
Inception Modülü(4a)		14 x 14 x 512	2	192	96	208	16	48	64	364K	73M
Inception Modülü(4b)		14 x 14 x 512	2	160	112	224	24	64	64	437K	88M
Inception Modülü(4c)		14 x 14 x 512	2	128	128	256	24	64	64	463K	100M
Inception Modülü(4d)		14 x 14 x 528	2	112	144	288	32	64	64	580K	119M
Inception Modülü(4e)		14 x 14 x 832	2	256	160	320	32	128	128	840K	170M
Maksimum Havuzlama Katmanı	3 x 3 / 2	7 x 7 x 832	0								
Inception Modülü(5a)		7 x 7 x 1024	2	256	160	320	32	128	128	1072K	54M
Inception Modülü(5b)		7 x 7 x 1024	2	384	192	384	48	128	128	1388K	71M
Ortalama Havuzlama	7 x 7 / 1	1 x 1 x 1024	0								
Düşürme (40%)			0								
Doğrusal Katman		1 x 1 x 1000	1							1000K	1M
Softmax Aktivasyon		1 x 1 x 1000	0								

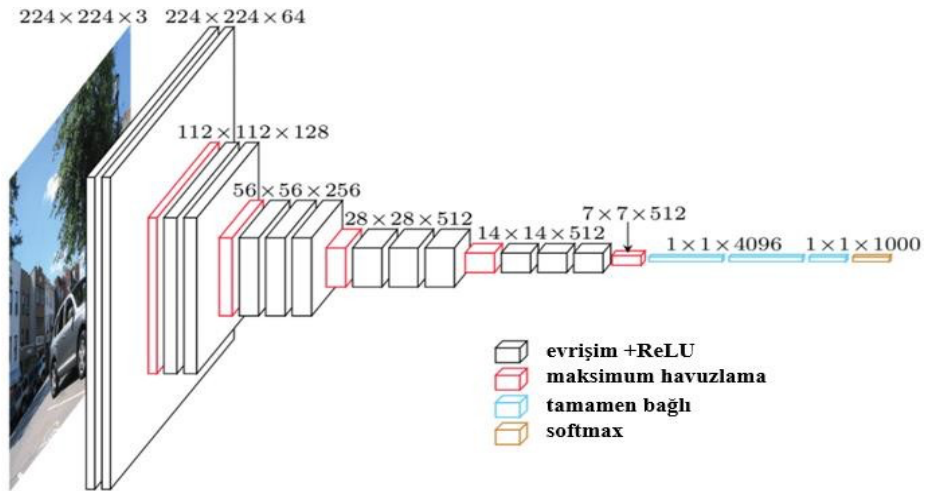
3.3.4. VGGNet

VGGNet, 2014 yılında Simonyan ve Zisserman tarafından Oxford Üniversitesi'nde geliştirilmiştir (Simonyan ve Zisserman, 2014). Aynı yıl ILSVRC (ImageNet Large Scale Visual Recognition Challenge) yarışmasında %7.3 top-5 hata oranı elde ederek AlexNet'in performansını önemli ölçüde aşmıştır. Giriş katmanı, 224x224 boyutunda renkli (RGB) görüntülerden oluşmaktadır. VGGNet mimarisi, iki farklı versiyona sahiptir. VGG-16, toplamda 16 katmandan (13 evrişim katmanı + 3 tam bağlı katman) ve VGG-19, toplamda 19 katmandan (16 evrişim katmanı + 3 tam bağlı katman) oluşmaktadır. Ağ, tüm katmanlarda aktivasyon fonksiyonu olarak ReLU kullanırken,

çıkış katmanında ise Softmax fonksiyonu tercih edilmiştir. Aşırı uyum sorununu önlemek için eğitim sırasında sönümleme (dropout) ve L2 regularizasyonu teknikleri uygulanmıştır. VGGNet'in en ayırt edici özelliği, 3x3 boyutundaki küçük filtreleri kullanarak derin bir mimari oluşturmastır. Bu küçük filtreler, ağı daha fazla katmanla derinleştirilmesini sağlarken aynı zamanda daha ayrıntılı özelliklerin öğrenilmesine olanak tanımaktadır. Ardışık 3x3 evrişim katmanlarının kullanımı, daha büyük alıcı alanlar (receptive fields) elde edilmesini sağlarken parametre sayısını ve hesaplama maliyetini düşürmektedir. Evrişim katmanlarından sonra 2x2 boyutunda maksimum havuzlama (max pooling) işlemleri uygulanarak özellik haritalarının boyutları azaltılmaktadır. Eğitim sürecinde, veri artırma teknikleri kullanılarak modelin genelleme kabiliyeti geliştirilmiş ve test sırasında çoklu ölçek değerlendirmesi uygulanarak performans artırılmıştır. VGGNet, derinlik ve performans arasındaki dengeyi kurarak modern sinir ağı mimarilerinin gelişimine önemli katkılar sağlamıştır. VGGNet mimarisi Şekil 3.11'de ve Çizelge 3.3'te verilmiştir.

Çizelge 3.3. VGG-16 mimarisi

Tür	Filtre Sayısı	Filtre Boyutu/Adım Sayısı	Çıktı Boyutu
Evrişim Katmanı	64	3 x 3 / 1	224 x 224 x 64
Evrişim Katmanı	64	3 x 3 / 1	224 x 224 x 64
Maksimum Havuzlama Katmanı		2 x 2 / 2	112 x 122 x 64
Evrişim Katmanı	128	3 x 3 / 1	112 x 122 x 128
Evrişim Katmanı	128	3 x 3 / 1	112 x 122 x 128
Maksimum Havuzlama Katmanı		2 x 2 / 2	56 x 56 x 128
Evrişim Katmanı	256	3 x 3 / 1	56 x 56 x 256
Evrişim Katmanı	256	3 x 3 / 1	56 x 56 x 256
Evrişim Katmanı	256	3 x 3 / 1	56 x 56 x 256
Maksimum Havuzlama Katmanı		2 x 2 / 2	28 x 28 x 256
Evrişim Katmanı	512	3 x 3 / 1	28 x 28 x 512
Evrişim Katmanı	512	3 x 3 / 1	28 x 28 x 512
Evrişim Katmanı	512	3 x 3 / 1	28 x 28 x 512
Maksimum Havuzlama Katmanı		2 x 2 / 2	14 x 14 x 512
Evrişim Katmanı	512	3 x 3 / 1	14 x 14 x 512
Evrişim Katmanı	512	3 x 3 / 1	14 x 14 x 512
Evrişim Katmanı	512	3 x 3 / 1	14 x 14 x 512
Maksimum Havuzlama Katmanı		2 x 2 / 2	7 x 7 x 512
Tam Bağlantılı Katman			4096
Tam Bağlantılı Katman			4096
Tam Bağlantılı Katman			1000
Softmax Aktivasyon Fonksiyonu			



Şekil 3.11. VGGNet mimari yapısı (Simonyan ve Zisserman, 2014)

3.3.5. EfficientNet

EfficientNet, model ölçeklendirme stratejilerinde yenilikçi bir yaklaşım sunan bir derin öğrenme mimarisidir. Bu mimari 2019 yılında Google Araştırma (Google Research) bünyesinde Tan ve Le tarafından geliştirilmiştir. (Tan ve Le, 2019). Bu mimari, bileşik ölçeklendirme (compound scaling) yöntemi kullanarak ağ genişliği, derinliği ve çözünürlüğü arasında optimal bir denge kurma prensibine dayanmaktadır. Bu çalışmada kullanılan EfficientNet-B0, ailenin temel modeli olup 224×224 piksel giriş çözünürlüğünde çalışmakta ve yaklaşık 5.3 milyon parametre içermektedir. Model, Ters Dar Boğaz Evrişimi (Mobile Inverted Bottleneck Convolution-MBConv) bloklarından oluşan bir yapıya sahiptir ve bu bloklar içerisinde derinlik tabanlı ayrışabilir evrişim, sıkıştırma-uyarma (squeeze-and-excitation) mekanizması ve ters artık bağlantıları kullanılmaktadır. Sıkıştırma-uyarma mekanizması, kanal bazında özellik haritalarının önemini dinamik olarak ağırlıklandırarak modelin dikkat mekanizmasını güçlendirmektedir. EfficientNet-B0 mimarisi, yedi ana MBConv bloğundan oluşmakta ve her blok farklı genişletme oranları ile yapılandırılmaktadır. Swish aktivasyon fonksiyonunun kullanımı, ReLU aktivasyon fonksiyonuna kıyasla gradyan akışını iyileştirerek modelin eğitim sürecinde daha kararlı bir performans sergilemesini sağlamaktadır. Bu çalışmada EfficientNet-B0'ın tercih edilmesinin temel nedenleri arasında düşük parametre sayısına rağmen yüksek performans göstermesi, hesaplama verimliliği ve transfer öğrenme için güçlü özellik temsil kabiliyeti yer almaktadır.

3.3.6. DenseNet

DenseNet (Densely Connected Convolutional Networks), 2017 yılında Huang ve arkadaşları tarafından IEEE Computer Vision and Pattern Recognition (CVPR)

konferansında sunulan ve katmanlar arası yoğun bağlantı yapısıyla dikkat çeken bir evrişimsel sinir ağı mimarisidir (Huang ve ark., 2017). Bu çalışmada kullanılan DenseNet-121 versiyonu, toplam 121 katmandan oluşmakta ve dört ana yoğun blok içermektedir. Mimari, 224×224 piksel boyutundaki RGB görüntüleri girdi olarak almakta ve her katmanın kendisinden önceki tüm katmanların özellik haritalarını doğrudan girdi olarak kullandığı yoğun bağlantı prensibiyle çalışmaktadır. Bu bağlantı mekanizması, özellik haritalarının toplama yerine kanal boyutunda birleştirme (concatenation) yöntemiyle gerçekleştirilmektedir. DenseNet-121'in yapısında ilk evrişim katmanı 7×7 boyutunda 64 filtre içermekte, bunu maksimum havuzlama katmanı takip etmektedir. Dört yoğun blok sırasıyla 6, 12, 24 ve 16 katman içermekte ve her blok arasında boyut azaltma işlemini gerçekleştiren geçiş katmanları bulunmaktadır. Büyüme oranı (growth rate) parametresi $k=32$ olarak ayarlanmış olup, her katman 32 yeni özellik haritası üretmektedir. Geçiş katmanları, 1×1 evrişim ile kanal sayısını azaltmakta ve 2×2 ortalama havuzlama ile uzaysal boyutları küçültmektedir. DenseNet-121'in yoğun bağlantı (dense connectivity) yapısı, gradyan yok olması problemini minimize ederek derin ağların eğitimini kolaylaştırmaktadır. Bu yapı ayrıca, özellik yeniden kullanımını (feature reuse) teşvik ederek parametre verimliliğini artırmakta ve sağladığı güçlü düzenleme (regularization) etkisiyle aşırı öğrenmeyi azaltmaktadır.

3.3.7. MobileNet

MobileNet-V2, özellikle kaynak kısıtlı ortamlar (resource-constrained environments) için optimize edilmiş hafif bir evrişimsel sinir ağı mimarisidir. Bu mimari 2018 yılında Google Araştırma ekibinden Sandler ve arkadaşları tarafından geliştirilmiştir (Sandler ve ark., 2018). Bu mimari, 53 katmandan oluşmakta ve 224×224 piksel standart giriş boyutunu desteklemektedir. MobileNet-V2'nin temel yeniliği, ters artık blokları (inverted residual blocks) ve doğrusal darboğaz (linear bottleneck) katmanlarının kullanımında yatmaktadır. Her ters artık bloğu üç temel bileşenden oluşmaktadır. İlk bileşen olan genişletme katmanı, 1×1 nokta tabanlı evrişim kullanarak düşük boyutlu girdileri yüksek boyutlu bir uzaya genişletmektedir. Genişletme faktörü tipik olarak 6 olarak ayarlanmıştır. İkinci bileşen, 3×3 derinlik tabanlı ayrışabilir evrişim katmanıdır ve her kanal üzerinde bağımsız uzaysal filtreleme gerçekleştirmektedir. Üçüncü bileşen ise doğrusal projeksiyon katmanıdır ve ReLU6 aktivasyonu yerine doğrusal aktivasyon kullanarak bilgi kaybını önlemektedir. MobileNet-V2'de toplam 17 ters artık bloğu bulunmaktadır. Bu bloklar farklı adım (stride) değerleri ile çözünürlük azaltma işlemini gerçekleştirmektedir. Derinlik çarpanı (width multiplier) ve çözünürlük

çarpanı (resolution multiplier) parametreleri ile modelin kapasitesi ve hesaplama yükü ayarlanabilmektedir. Bu çalışmada standart MobileNetV2 konfigürasyonu kullanılmış olup, yaklaşık 3.5 milyon parametre içermektedir. MobileNetV2'nin ters artık yapısı, düşük boyutlu darboğazlarda bilginin korunmasını sağlarken, yüksek boyutlu uzaylarda doğrusal olmayan dönüşümlerin gerçekleştirilmesine olanak tanımaktadır. Bu tasarım, modelin hem hesaplama verimliliğini artırmakta hem de temsil gücünü korumaktadır. Böylece histolojik görüntü analizi gibi hassas görsel tanıma görevlerinde etkili bir performans sergilemektedir.

3.4. Nesne Tespitinde Kullanılan Derin Öğrenme Modelleri

Nesne tespiti, bir görüntü içerisindeki nesnelerin hem ne olduğunun belirlenmesi (sınıflandırma) hem de nerede bulunduğunun tespiti (konumlandırma) görevlerini birlikte gerçekleştiren bir bilgisayarlı görü problemidir. Derin öğrenmenin gelişmesiyle birlikte özellikle evrimsel sinir ağları tabanlı yöntemler bu alanda ilerlemeler kaydetmiştir. Bu bölümde, nesne tespitinde yaygın olarak kullanılan derin öğrenme modellerinin mimarileri ve çalışma prensipleri detaylı olarak ele alınacaktır.

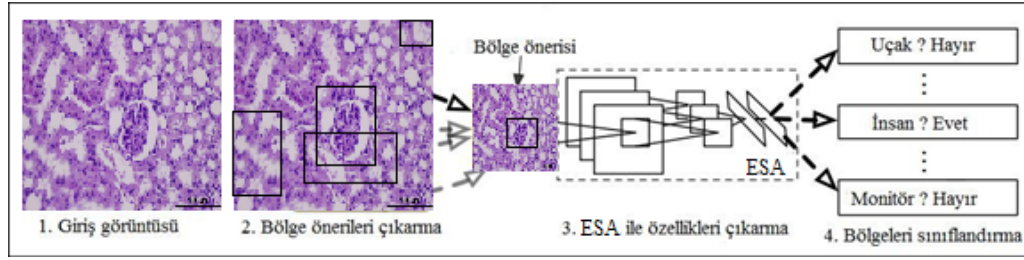
3.4.1. R-CNN

R-CNN mimarisi, Girshick ve arkadaşları (2014) tarafından önerilmiştir (Girshick ve ark., 2014). R-CNN mimarisi dört temel aşamadan oluşmaktadır. İlk olarak, modele giriş olarak bir görüntü verilir. İkinci aşamada, bu görüntü üzerinde olası nesne bölgeleri tespit edilir. Bu amaçla Seçici Arama (Selective Search) yöntemi kullanılarak, nesne içermeye olasılığı yüksek yaklaşık iki bin aday bölge (region proposal) oluşturulur. Üçüncü adımda, elde edilen her bir bölge, bir evrimsel sinir ağı modeline aktarılır ve bu ağıdan elde edilen aktivasyonlar kullanılarak her bölge için bir öznitelik vektörü (feature vector) çıkarılır.

Son aşamada ise çıkarılan bu öznitelikler, Destek Vektör Makineleri (DVM) aracılığıyla sınıflandırılarak nesnenin ait olduğu kategori belirlenir. Ayrıca, bölge konumlarının daha doğru biçimde saptanabilmesi için lineer regresyon temelli bir sınır kutusu (bounding box) iyileştirme işlemi uygulanır. Böylece model hem nesnenin sınıfını hem de konumunu yüksek doğrulukla tahmin edebilmektedir.

Bu yöntemin temel dezavantajı, her bir görüntüde seçici arama tarafından önerilen yaklaşık iki bin bölgenin ayrı ayrı işlenmesi gerektiğinden hem eğitim hem de test aşamalarında önemli ölçüde zaman gerektirmesidir. Bunun yanında, seçici arama algoritması öğrenmeye dayalı bir yöntem olmadığından, bölge önerileri model performansına göre uyarlanamaz. Bu durum, bazı durumlarda düşük kaliteli veya ilgisiz

bölge tahminlerinin üretilmesine neden olabilmektedir. R-CNN mimarisi Şekil 3.12’de verilmiştir.



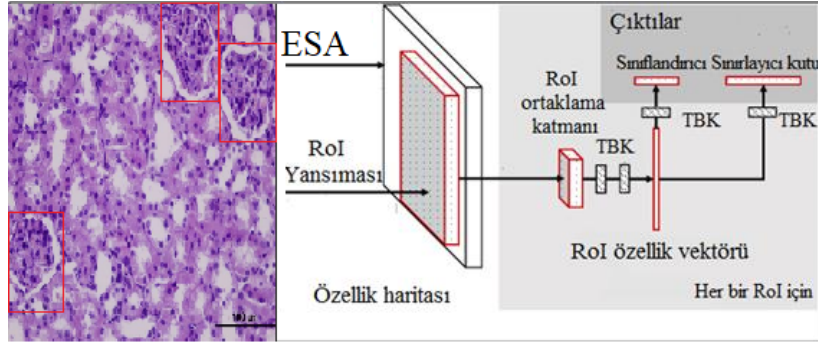
Şekil 3.12. R-CNN mimarisi

3.4.2. Fast R-CNN

Fast R-CNN, Girshick (2015) tarafından, R-CNN modelinin getirdiği yüksek hesaplama maliyeti ve düşük hız sorunlarını çözmek amacıyla geliştirilmiştir (Girshick, 2015). Bu mimarinin temel yeniliği, R-CNN'in aksine her bölge önerisini (~2000 adet) ayrı ayrı evrişim ağına sokmak yerine tüm giriş görüntüsünü yalnızca bir kez bir evrişimsel sinir ağı katmanlarından geçirerek ortak bir özellik haritası elde etmesidir.

Mimarinin çalışma prensibi, bu ortak özellik haritası üzerine kuruludur. Görüntüden elde edilen aday bölgeler (Region of Interest - RoI), bu ortak harita üzerine yansıtılır. Mimarinin ikinci temel yeniliği olan RoI Havuzlama (RoI Pooling) katmanı, bu yansıtılan ve değişken boyutlarda olan bölgelerden, sabit boyutlu (örn: 7x7) özellik vektörleri çıkarır. Bu özellik vektörleri daha sonra tam bağlı katmanlara aktarılır ve burada iki paralel işlem hattına ayrılır. Birinci hat Softmax aktivasyonu aracılığıyla nesne sınıfını tahmin ederken ikinci hat sınır kutusu regresyonu ile nesnenin uzamsal koordinatlarını hassaslaştırır.

Fast R-CNN'in en önemli avantajı, tüm görüntü üzerinde bir kez evrişim işlemi gerçekleştirerek hesaplama paylaşımını (computation sharing) sağlamasıdır. Sınıf tahmini ve sınır kutusu düzeltmesinin tam bağlı katmanlardan elde edilen ortak özellikler kullanılarak tek bir ağda gerçekleştirilmesi modeli daha verimli hale getirmektedir. Sonuç olarak, Fast R-CNN, R-CNN nesne tespit yöntemlerine kıyasla hem hız hem de doğruluk açısından daha üstün bir performans sergilemektedir.



Şekil 3.13. Fast R-CNN mimarisi

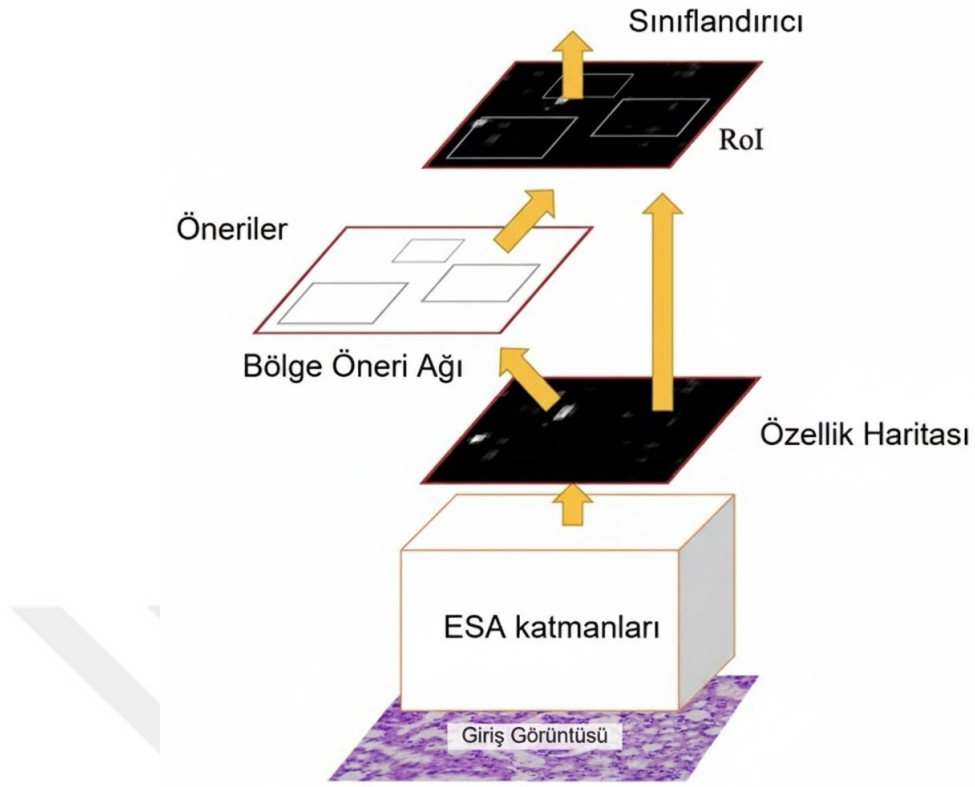
3.4.3. Faster R-CNN

Faster R-CNN, Ren ve arkadaşları tarafından geliştirilen Fast R-CNN mimarisinin bir gelişmiş versiyonudur. (Ren ve ark., 2015). Bu model, R-CNN ve Fast R-CNN yöntemlerinde bulunan seçici arama adımının hesaplama açısından yavaş olması sorununu ortadan kaldırmak amacıyla tasarlanmıştır.

Modelin en önemli yeniliği, Bölge Öneri Ağı (Region Proposal Network - RPN) adlı alt ağıdır. RPN'in temel başarısı, ana ağ (backbone) ile aynı evrimsel özellik haritasını (feature map) paylaşarak olası nesne bölgelerini otomatik olarak önermektedir. Böylece harici bir algoritmaya ihtiyaç duymadan bölge önerileri hızlı bir şekilde üretilmektedir. RPN, paylaşılan özellik haritası üzerinde kayan bir pencere (sliding window) yöntemiyle her konumda nesne olma olasılığını ve bu nesnelere ait sınır kutularını (bounding boxes) tahmin etmektedir. Her pencere konumu, farklı ölçek ve en-boy oranlarına sahip çapa kutuları (anchor boxes) ile ilişkilendirilir. Bu yapıda her çapa kutusu için iki paralel tahmin yapılır. Bölgenin nesne içerip içermediğini belirleyen sınıflandırma skoru ve tahmini kutunun konumunu ve boyutlarını hassaslaştıran regresyon değerleri tahmin edilir.

RPN'den elde edilen yüksek olasılıklı bölge önerileri (RoI'lar), daha sonra doğrudan Fast R-CNN'in tespit mekanizmasına girdi olarak verilir. Bu aşama, Fast R-CNN modelindeki süreçle aynıdır. Bölge önerileri, RoI Havuzlama (RoI Pooling) katmanından geçirilir ve ardından Softmax ile konum ayarlaması (regresyon) yapan paralel çıkış katmanlarına beslenir.

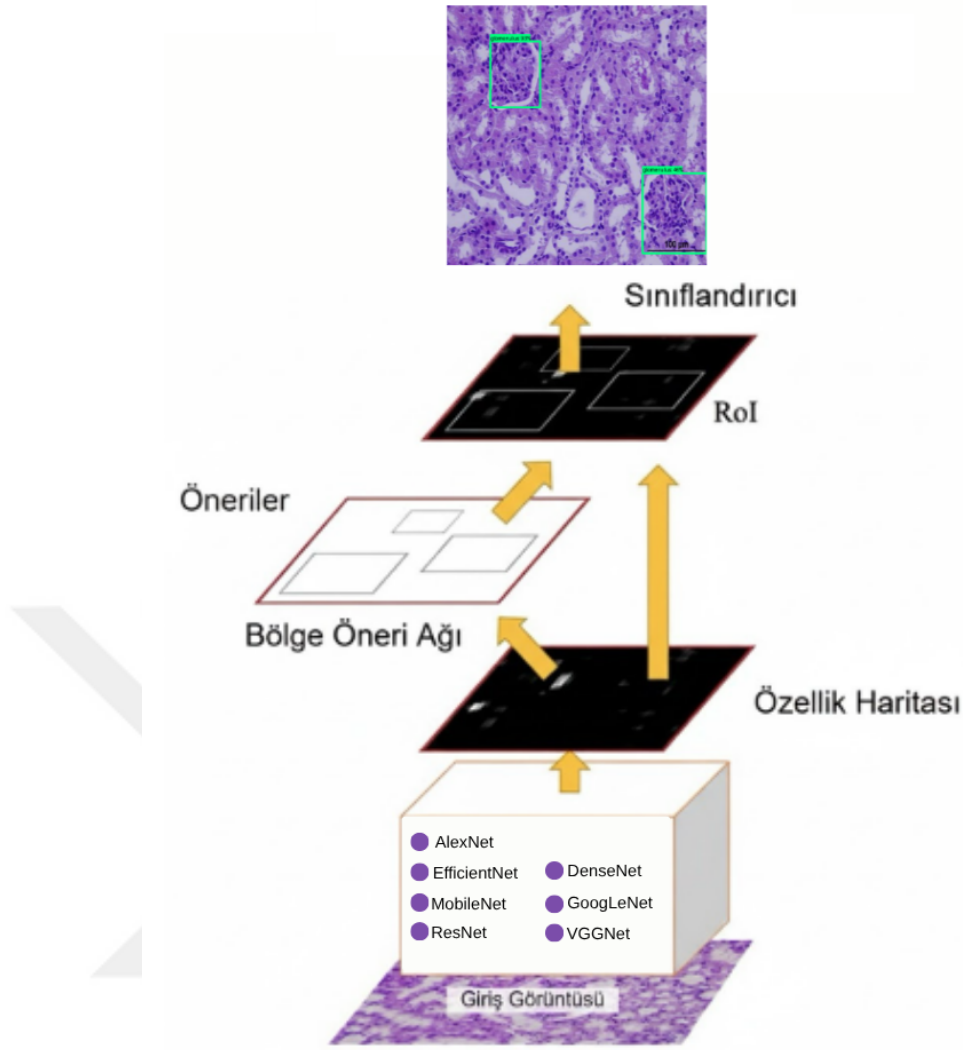
Faster R-CNN, bölge önerilerini doğrudan ESA yapısı içerisine entegre ettiği için önceki yöntemlerdeki gibi dışsal algoritmalara ihtiyaç duymaz. Böylece, tüm nesne tespit süreci tek bir birleşik ağ yapısında gerçekleştirilir. Bu bütünleşik mimari hem eğitim hem de test süresini önemli ölçüde azaltırken, tespit doğruluğunu da artırır. Sonuç olarak Faster R-CNN, önceki bölge tabanlı yöntemlere kıyasla daha hızlı, verimli ve yüksek doğrulukta nesne algılama performansı sunmaktadır.



Şekil 3.13. Faster R-CNN mimarisi

3.5. Önerilen Yöntem

Bu çalışmada, histolojik böbrek görüntülerinden glomerüllerin otomatik tespiti için Faster R-CNN mimarisi temel alınarak bir derin öğrenme modeli geliştirilmiş ve farklı omurga modelleriyle performansı karşılaştırmalı olarak analiz edilmiştir. Geliştirilen yöntemin genel mimarisi Şekil 3.14'te özetlenmiştir. Bu bölümde çalışmanın metodolojik çerçevesi, kullanılan donanım, veri seti yapılandırması, model mimarisi, eğitim süreci ve performans değerlendirme kriterleri başlıkları altında ayrıntılı olarak sunulmaktadır.



Şekil 3.14. Önerilen yöntem

3.5.1. Veri Seti

Bu çalışmada kullanılan veri kümesi, modelin öğrenme ve genelleme başarımını değerlendirmek amacıyla eğitim (train), doğrulama (validation) ve test (test) olmak üzere 3 alt kümeye ayrılmıştır. Veri kümesinin bu şekilde bölünmesi hem modelin yeterli sayıda örnek üzerinde öğrenmesini sağlamak hem de bağımsız veri gruplarıyla değerlendirerek aşırı uyum (overfitting) riskini azaltmak amacıyla planlanmıştır. Bu kapsamda, toplam 2500 verinin %70'i eğitim, %15'i doğrulama ve %15'i test setine ayrılmıştır.

Veri yükleme ve eğitim sürecinde yığın (batch) boyutu 16 olarak belirlenmiştir. Her bir yığında 16 görüntü işlenmiştir. Eğitim sırasında karıştırma (shuffle) işlemi etkinleştirilerek her devirde (epoch) örneklerin rastgele sıralanması sağlanmıştır. Bu yaklaşım veri sırasından kaynaklanabilecek öğrenme yanlılığını (bias) azaltarak modelin daha sağlam genelleme yeteneği kazanmasına katkı sağlamaktadır. Eğitim ve değerlendirme aşamaları arasında tutarlı ve karşılaştırılabilir sonuçlar elde edilmesi hedeflenmiştir.

3.5.2. Model Mimarisi ve Omurga Seçimi

Bu çalışmada kullanılan nesne tespit modelinin temelini Girshick ve arkadaşları tarafından önerilen iki aşamalı nesne tespit algoritması olan Faster R-CNN mimarisi oluşturmaktadır. Bu mimari, bir özellik çıkarım ağı (omurga), bir Bölge Öneri Ağı (Region Proposal Network-RPN) ve bir sınıflandırıcı/regresör başlığından oluşur.

Bu çalışmanın ana odak noktası, Faster R-CNN'in performansını farklı omurga ağları ile karşılaştırmalı olarak değerlendirmektir. Bu kapsamda, derin öğrenme literatüründe yaygın olarak kullanılan sekiz farklı mimari incelenmiştir. Bunlar:

- AlexNet
- ResNet-50
- ResNet-101
- GoogLeNet (Inception-v3)
- VGG-16
- DenseNet-121
- EfficientNet-B0
- MobileNet-V2

Tüm omurga modellerinin eğitimi sürecinde transfer öğrenme yaklaşımından faydalanılmıştır. Modeller, ImageNet veri kümesi üzerinde önceden eğitilmiş ağırlıklarla başlatılmıştır. Bu durum modelin glomerül tespiti gibi özel görevlere daha hızlı ve etkili şekilde uyum sağlamasına yardımcı olur. Ayrıca, eğitim süresini kısaltır ve özellikle sınırlı veriye sahip tıbbi görüntüleme alanında performansı artırır.

3.5.3. Model Eğitimi

Bu çalışmada Faster R-CNN mimarisi, farklı omurga (backbone) modelleri kullanılarak sistematik biçimde değerlendirilmiştir. Modellerin adil biçimde karşılaştırılabilmesi için eğitim sürecinde yalnızca omurga mimarisi değişken bırakılmıştır. Öğrenme oranı, optimizasyon algoritması, devir sayısı, yığın (batch) boyutu, veri bölünme oranları ve erken durdurma gibi temel ayarlar sabit tutulmuştur. Bu koşullar altında, eğitim sürecinde değişken olarak yalnızca omurga (backbone) mimarisi kullanılmıştır. Böylece tüm deneyler aynı veri kümesi üzerinde aynı optimizasyon stratejisi ve özdeş hiper parametrelerle yürütülmüştür. Bu bölümde kullanılan donanım altyapısı, veri seti yapılandırması, hiper parametre seçimi ve optimizasyon yaklaşımı ile eğitim sürecinin işleyişi ayrıntılı olarak sunulmaktadır. Çizelge 3.4'te, her bir omurga

modeline ait omurga ağlarının yapısal özellikleri, giriş boyutları, ön eğitim bilgileri ve parametre sayıları sunulmuştur.

Çizelge 3.4. Çalışmada kullanılan farklı omurga mimarilerinin giriş boyutu, ön eğitim durumu ve parametre sayılarına ilişkin karşılaştırma

Model	Girdi boyutu	Ön eğitim	Omurga parametre sayısı	Eğitilebilir parametre sayısı
AlexNet	500×500×3	ImageNet-1K	2,47 M	16,98 M
EfficientNet-B0	500×500×3	ImageNet-1K	3,6 M	21,66 M
DenseNet-121	500×500×3	ImageNet-1K	6,95 M	68,91 M
GoogLeNet (Inception-v3)	500×500×3	ImageNet-1K	21,79 M	163,51 M
MobileNet-V2	500×500×3	ImageNet-1K	2,22 M	82,35 M
ResNet-50	500×500×3	ImageNet-1K	23,51 M	165,23 M
ResNet-101	500×500×3	ImageNet-1K	42,5 M	184,23 M
VGG-16	500×500×3	ImageNet-1K	14,71 M	43,86 M

3.5.4. Eğitim Ortamı

Tüm derin öğrenme modellerinin eğitimi, Google Colab Pro+ platformu üzerinde gerçekleştirilmiştir. Eğitim sürecinde kullanılan NVIDIA Tesla A100 GPU (80 GB VRAM) donanımı, yüksek çözünürlüklü histolojik görüntülerin işlenmesi ve derin öğrenme modellerinin optimize edilmesi için gerekli hesaplama gücünü sağlamıştır. Eğitim süreci, Python programlama dili temel alınarak yürütülmüştür.

3.5.5. Hiper parametreler ve Optimizasyon Stratejisi

Bu çalışmada modelin optimizasyonu için Momentumlu Stokastik Gradyan İnişi (Stochastic Gradient Descent with Momentum – MSGD) tercih edilmiştir. MSGD, özellikle Faster R-CNN gibi nesne tespiti odaklı modellerde yaygın kullanılan bir yaklaşımdır. Başlangıç öğrenme oranı (initial learning rate) 0,005, momentum katsayısı (momentum) 0,9 ve ağırlık çürümesi (weight decay) 0,0005 olarak belirlenmiştir. Bu değerler, literatürde önerilen aralıklarla uyumlu olup veri kümesinin özellikleri ve modelin karmaşıklığı dikkate alınarak seçilmiştir.

Öğrenme oranının zaman içindeki değişimini düzenlemek için ısıtılmalı kosinüs azaltma (cosine annealing with warmup) stratejisi uygulanmıştır. Eğitimin ilk beş devirde ısınma (warmup) aşaması olarak tanımlanmış, bu sürede öğrenme oranı kademeli olarak artırılarak başlangıçtaki ani ağırlık güncellemeleri engellenmiştir. Isınma sürecinin ardından öğrenme oranı, kosinüs fonksiyonu düzenli biçimde azaltılmıştır. Bu yaklaşım erken aşamalarda daha hızlı öğrenmeyi ileri aşamalarda ise daha hassas ağırlık güncellemelerini mümkün kılarak modelin yakınsama (convergence) başarımını artırmıştır.

Aşırı öğrenmeyi (overfitting) önlemek amacıyla erken durdurma (early stopping) kullanılmıştır. Doğrulama setindeki mAP50 metriği izlenmiştir. Bu metrikte 10 ardışık devir sayısı boyunca iyileşme görülmediğinde eğitim sonlandırılmıştır. Maksimum devir sayısı 50 olarak belirlenmiş olsa da erken durdurma sayesinde çoğu model bu sınıra ulaşmadan eğitimi tamamlamıştır. Böylece doğrulama performansı en yüksek noktada eğitim durdurulmuş, gereksiz hesaplama maliyeti azaltılmıştır. Çalışmada kullanılan omurga mimarilerinin eğitim hiper parametreleri Çizelge 3.5'te verilmiştir.

Çizelge 3.5. Çalışmada kullanılan omurga mimarilerinin eğitim parametreleri

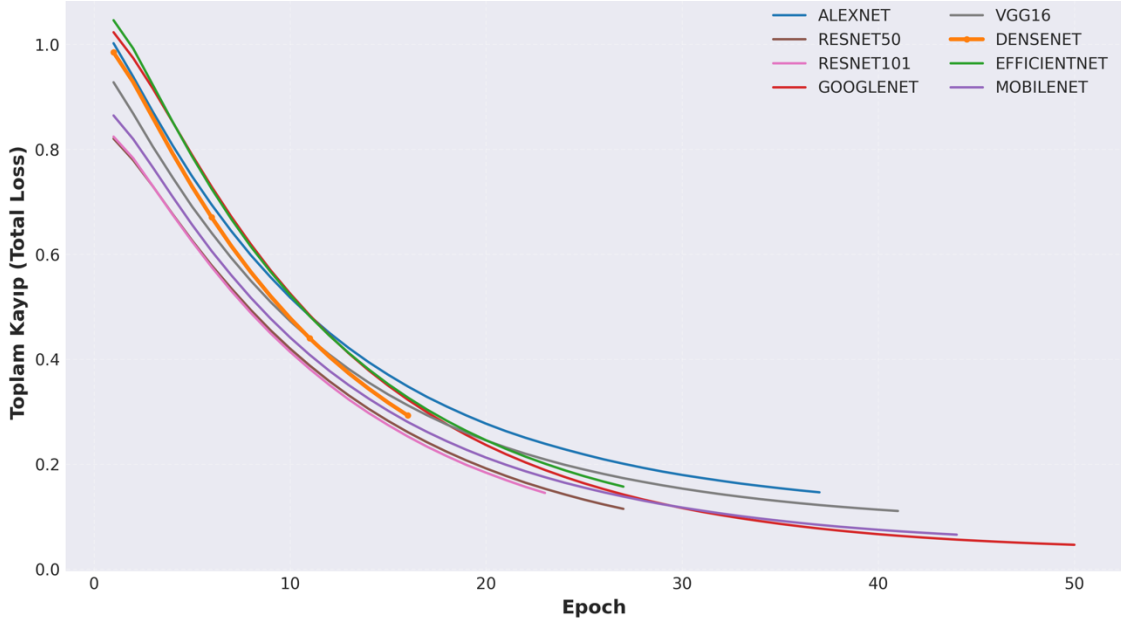
Parametre	Değer
Optimizasyon algoritması	Momentumlu Stokastik Gradyan İnişi
Başlangıç öğrenme oranı	0,005
Momentum	0,9
Ağırlık çürümesi (weight decay)	0,0005
Devir sayısı	50
Erken durdurma (patience)	10
Bir devirde kullanılacak görüntü sayısı	16
Öğrenme oranı zamanlaması	Isıtmalı kosinüs azaltma
Çalışma ortamı	Google Colab Pro+

3.5.6. Eğitim Süreci ve Kayıp Fonksiyonları

Faster R-CNN mimarisinin eğitim süreci, dört temel kayıp bileşeninin optimizasyonunu içerir: RPN nesne varlık (objectness) kaybı, RPN kutu regresyon kaybı, R-CNN sınıflandırma kaybı ve R-CNN kutu regresyon kaybı. Her bir bileşen, modelin farklı alt görevlerini hedefleyen biçimde tanımlanmıştır. RPN bileşenleri, görüntü üzerinde potansiyel nesne bölgelerini önermektedir. Objectness kaybı, aday bir bölgenin nesne içerme olasılığını eniyilemeyi amaçlar. RPN kutu regresyonu, bu bölgelerin sınırlarını hassaslaştırır. R-CNN bileşenleri, önerilen bölgeleri sınıflandırır ve konumlarını son kez düzeltir. Sınıflandırma kaybı, glomerül-arka plan ayrımının öğrenilmesini sağlar. Kutu regresyon kaybı, tespit edilen glomerüllerin sınırlayıcı kutularının konum ve boyutlarının piksel düzeyinde iyileştirilmesine katkı sağlamaktadır. Bu dört bileşenin eşzamanlı eniyilenmesi, bölge önerisi, sınıflandırma ve konumlandırma adımlarının dengeli ve tutarlı biçimde öğrenilmesini sağlar.

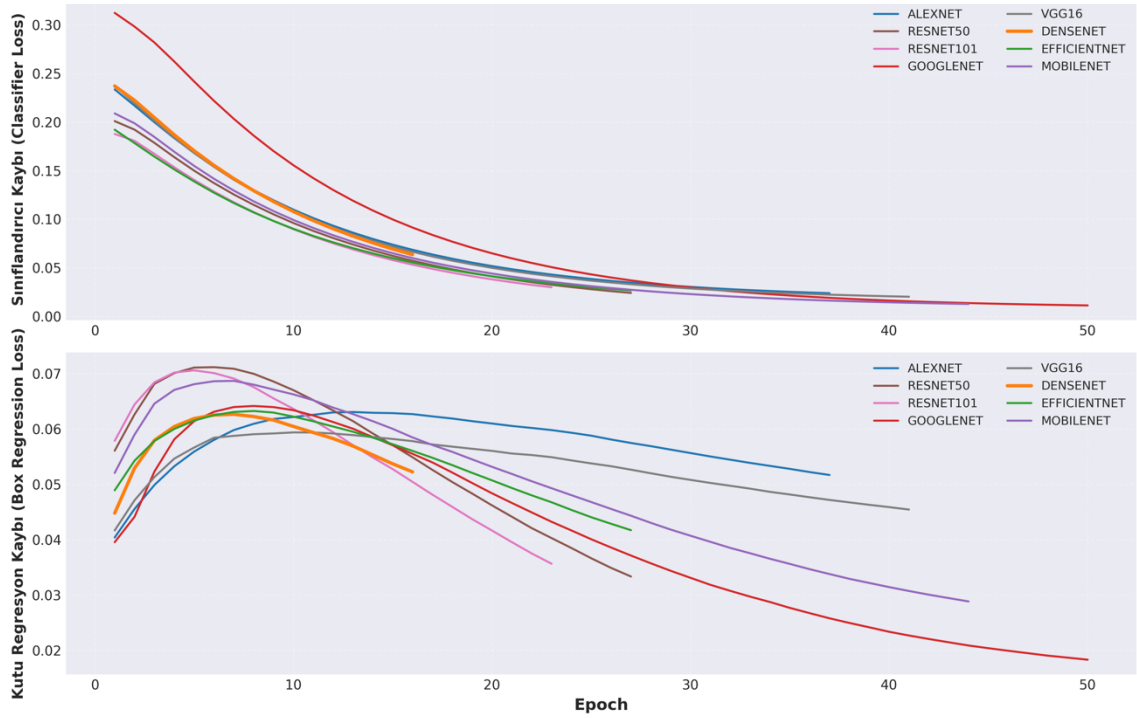
Eğitim sürecinde, her bir eğitim devrinin (epoch) sonunda dört temel kayıp bileşeninin (RPN sınıflandırma kaybı, RPN kutu regresyon kaybı, ROI sınıflandırma kaybı ve ROI kutu regresyon kaybı) mini-yığın (mini-batch) ortalamaları hesaplanmış ve toplam kayıp değeri kaydedilmiştir. Tüm modellerin eğitim süreci, aşırı uyum (overfitting) riskini azaltmak ve hesaplama kaynaklarının verimli kullanımını sağlamak

amacıyla erken durdurma mekanizması ile denetlenmiştir. Erken durdurma stratejisi, doğrulama kümesi üzerinde hesaplanan mAP50 metriği temel alınarak uygulanmıştır. Doğrulama kümesinde mAP50 metriğinde ardışık 10 devir sayısı (epoch) boyunca iyileşme gözlenmediğinde eğitim sonlandırılmıştır. Bu yaklaşım aşırı öğrenme riskini azaltırken modellerin en iyi genelleme düzeyinde durdurulmasını sağlamıştır. Sonuç olarak, her bir omurga ağı için elde edilen nihai model doğrulama kümesi üzerinde en yüksek mAP50 değerini veren devir sayısına ait ağırlıklardır.



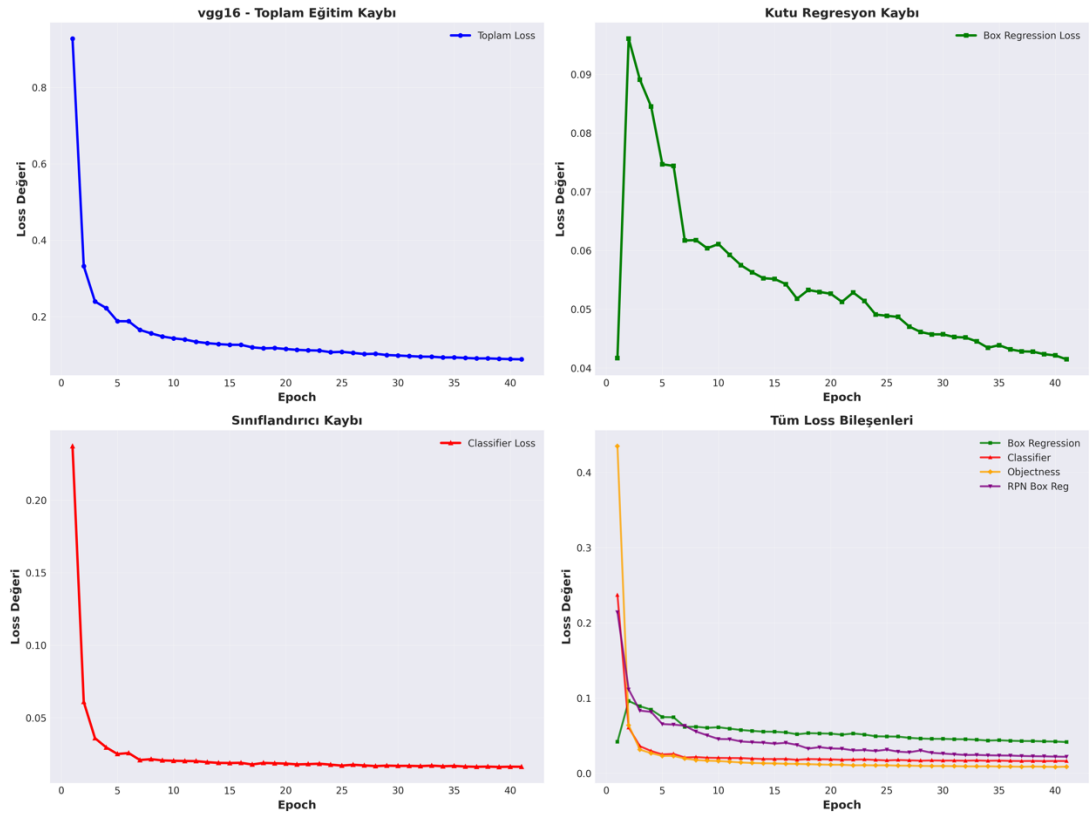
Şekil 3.15. Modellerin eğitim kaybı grafiği

Şekil 3.15'te sekiz farklı omurga ağı (AlexNet, ResNet-50, ResNet-101, GoogLeNet, VGG-16, DenseNet-121, EfficientNet-B0 ve MobileNet-V2) için toplam eğitim kaybının devir sayısına (epoch) göre değişimini karşılaştırmalı olarak göstermektedir. Şekil 3.16, sınıflandırma ve kutu regresyon kayıplarının tüm modeller için toplu karşılaştırmalarını göstermektedir.

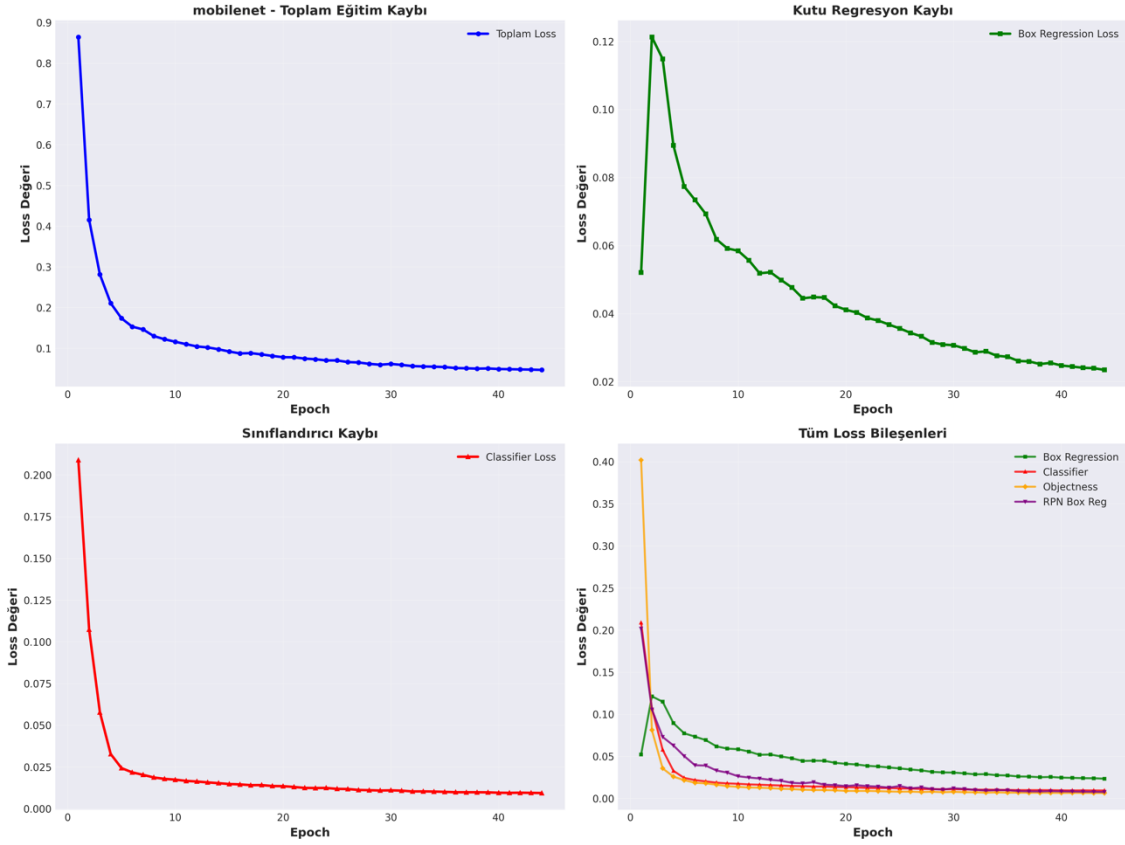


Şekil 3.16. Modellerin sınıflandırma kaybı ve kutu regresyon kaybı grafiği

Eğitim kayıplarının incelenmesi, modellerin öğrenme kapasitelerini ve eğitim verisi üzerindeki performanslarını değerlendirmek için önemli bir gösterge olmakla birlikte modellerin nihai başarımları test kümesi üzerindeki performans metrikleri ile değerlendirilmiştir.



Şekil 3.17. VGG16 modeline ait kayıp grafikleri



Şekil 3.18. MobileNetV2 modeline ait kayıp grafikleri

Şekil 3.17 ve Şekil 3.18'te VGG-16 ve MobileNet-V2 omurga ağları için toplam eğitim kaybı, kutu regresyon kaybı, sınıflandırıcı kaybı ve tüm kayıp bileşenlerinin epoch değerlerine göre değişimi detaylı olarak sunulmuştur. Bu iki örnek, mimari farklılıklara karşın sınıf ayırımının eğitim boyunca düzenli biçimde iyileştiğini ortaya koymaktadır.

Tüm omurga modelleri (AlexNet, ResNet-50, ResNet-101, GoogLeNet (Inception-v3), VGG-16, DenseNet-121, EfficientNet-B0 ve MobileNet-V2) için gerçekleştirilen eğitim süreçlerinde, toplam kayıp (total loss) ile kayıp bileşenlerinin devir sayısı (epoch) ilerledikçe istikrarlı biçimde azaldığı gözlemlenmiştir. Doğrulama setinde hesaplanan mAP50 metriğinin artış eğilimi, kayıp fonksiyonlarının düşüş eğrileriyle paralel bir yakınsama davranışı sergilemiştir. Bu durum, eğitim sürecinin dengeli ve kararlı bir şekilde ilerlediğini, modellerin aşırı öğrenme (overfitting) veya yetersiz öğrenme (underfitting) problemleriyle karşılaşmadığını göstermektedir.

Eğitim sürecinde RPN (Region Proposal Network) bileşenlerine ait kayıplar özellikle objectness loss ve RPN box regression loss erken dönemlerde hızlı bir azalma göstermiştir. Bu bulgu, bölge öneri mekanizmasının kısa sürede kararlı hâle geldiğini ve modelin potansiyel nesne bölgelerini güvenilir biçimde tanımayı erken dönemde öğrendiğini ortaya koymaktadır.

Bileşen bazlı hataların devir sayısı ilerledikçe sistematik biçimde azalması, modelin optimizasyon sürecinin dengeli biçimde yürütüldüğünü ve öğrenme dinamiklerinin tutarlı olduğunu göstermektedir. Ayrıca erken durdurma (early stopping) mekanizması sayesinde tüm modeller doğrulama performanslarının en yüksek olduğu noktalarda eğitimi tamamlamış ve böylece hesaplama kaynakları verimli biçimde kullanılmıştır.

3.6. Performans Değerlendirme Metrikleri

Derin öğrenme modellerinin performansını değerlendirmek ve diğer modellerle karşılaştırmak amacıyla çeşitli ölçütler kullanılmaktadır. Nesne tespit problemlerinde Birleşim Üzerinden Kesişim (IoU - Intersection over Union) değerine göre tespitlerin gerçek etiket (ground truth) ile birebir eşleştirilmesine dayalı TP/FP/FN sayımlarını temel aldığı için karşılaştırmalar nesne tespiti bazlı doğruluk (mAP), kesinlik-duyarlılık eğrileri ve IoU tabanlı metrikler üzerinden yapılmaktadır. Bu tezde, model performansı mAP, kesinlik (precision), duyarlılık (recall) ve F1-skor metrikleri kullanılarak değerlendirilmiştir.

IoU, modelin tahmin ettiği sınırlayıcı kutu (bounding box) ile gerçek etiket (ground truth) kutusu arasındaki kesişim alanının, bu iki alanın birleşim alanına oranını ifade etmektedir. IoU, modelin tahmininin gerçek etikete ne ölçüde yakın olduğunu değerlendirmek için kullanılır. Bu metrik 0 ile 1 arasında bir değer alır; değer büyümeye, model tahmininin gerçek etiketle daha fazla örtüştüğünü gösterir. IoU değerinin 1 olması, tahmin ile gerçek etiketin tamamen örtüştüğü anlamına gelmektedir.

IoU değerleri ve bu değerleri sınıflandırmak için belirlenen eşik değeri kullanılarak bir hata matrisi (confusion matrix) oluşturulur. Hata matrisi, Çizelge 3.6'da gösterildiği üzere, Doğru Pozitif (TP), Yanlış Pozitif (FP), Doğru Negatif (TN) ve Yanlış Negatif (FN) olmak üzere dört temel bileşenden oluşur. Bu bileşenler, modelin tahmin performansını ve sınıflandırma doğruluğunu ayrıntılı biçimde değerlendirmeye olanak sağlamaktadır.

Çizelge 3.6. Karmaşıklık matrisi

		GERÇEK	
		Pozitif	Negatif
TAHMİN	Pozitif	Doğru Pozitif (TP)	Yanlış Pozitif (FP)
	Negatif	Yanlış Negatif (FN)	Doğru Negatif (TN)

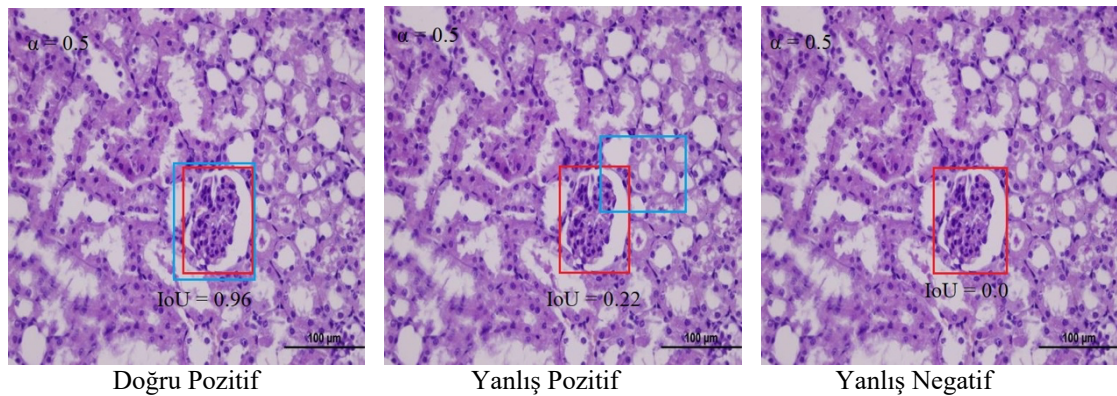
Doğru Pozitif, modelin doğru şekilde tespit ettiği nesneleri ifade eder. Bir tahminin TP olarak değerlendirilmesi için tahmin edilen sınırlayıcı kutunun gerçek referans kutusuyla IoU değerinin belirlenen eşik değerinin (genellikle 0,5) üzerinde olması gerekir. TP sayısı, modelin başarılı tespit sayısını göstermektedir.

Yanlış Pozitif, modelin gerçekte nesne olmayan bir bölgeyi nesne olarak etiketlemesi durumudur. Bir tahmin, IoU değerinin eşik değerinin altında kalması veya herhangi bir gerçek referans kutusuyla eşleşmemesi halinde FP olarak değerlendirilir. FP sayısının artması kesinlik (precision) metriğini olumsuz etkiler.

Yanlış Negatif, görüntüde mevcut olduğu halde model tarafından tespit edilemeyen nesneleri ifade eder. FN sayısı, modelin kaçırdığı nesne sayısını gösterir ve duyarlılık (recall) metriğini olumsuz etkiler.

Doğru Negatif (TN) kavramı, sınıflandırma problemlerinde anlamlı ve yaygın olarak kullanılan bir performans metriği olmakla birlikte nesne tespitinde pratik bir kullanım alanı bulunmamaktadır. Bu durumun temel nedeni nesne tespiti görevlerinde arka plan bölgelerinin tanımının belirsiz olması ve hangi bölgelerin (çapa kutuları/bölge önerileri) negatif olarak değerlendirileceğine dair standart bir ölçütün bulunmamasıdır.

Pratikte IoU değeri genellikle belirli bir eşik değeri (threshold) üzerinden değerlendirilir. Bu eşik, bir tespitin doğru kabul edilip edilmeyeceğini belirler. Çoğunlukla 0,5 değeri eşik olarak kullanılmakta olup, IoU değeri 0,5'in üzerinde olan tespitler doğru (başarılı) kabul edilir. Şekil 3.19'da 0,5 eşik değeri kullanılarak IoU'nun sağladığı sınıflandırma örneği gösterilmektedir.



Şekil 3.19. 0,5 eşik değerine göre IoU örneği

Glomerül tespit probleminde, Doğru Pozitif (TP) değeri modelin doğru olarak tespit ettiği glomerül sayısını, Yanlış Pozitif (FP) değeri gerçekte glomerül olmayan ancak model tarafından glomerül olarak işaretlenen bölge sayısını, Yanlış Negatif (FN) değeri gerçekte

glomerül olan ancak model tarafından tespit edilemeyen glomerül sayısını temsil etmektedir.

3.6.1. Kesinlik

Kesinlik (precision), modelin doğru tespit ettiği pozitif sonuçların modelin tespit ettiği tüm pozitif sonuçlara oranıdır. Modelin ne sıklıkla doğru tahminde bulunduğunu gösterir. Bu sayede modelin pozitif tahminlerinin güvenilirliği sorgulanabilir. Yüksek kesinlik değeri, modelin yaptığı tespitlerin büyük çoğunluğunun doğru olduğunu gösterir.

$$Kesinlik = \frac{TP}{TP+FP} \quad (3.2)$$

3.6.2. Duyarlılık

Gerçek pozitif oran olarak da bilinen duyarlılık (recall) modelin doğru tespit ettiği pozitif sonuçların pozitif olarak tanımlanması gereken sonuçlara oranıdır. Modelin tespit etmesi gereken nesnelerin ne kadarını tespit edebildiğini gösterir. Bu doğrultuda kaçırılan tespitler duyarlılığı düşürmektedir. Yüksek duyarlılık değeri, modelin gerçek nesnelerin büyük çoğunluğunu başarıyla tespit ettiğini gösterir.

$$Duyarlılık = \frac{TP}{TP+FN} \quad (3.3)$$

3.6.3. F1-Skor

F1-Skor, duyarlılık ve kesinlik arasındaki dengeyi hesaplar. F1-Skor ile duyarlılık ve kesinlik arasında doğru orantı vardır. Eğer F1-Skor yüksekse duyarlılık ve kesinlik de yüksek, düşük ise düşüktür.

$$F - 1 \text{ Skor} = \frac{2 * Duyarlılık * Kesinlik}{Duyarlılık + Kesinlik} \quad (3.4)$$

3.6.4. Ortalama Hassasiyet

Ortalama Hassasiyet (Average Precision) kesinlik-duyarlılık eğrisinin altındaki alanı kapsar. Kesinlik-duyarlılık eğrisi her iki ölçümün etkisini de en üst düzeye çıkarır. Modelin genel doğruluğu hakkında daha iyi bir fikir verir.

$$AP = \int_0^1 p(r) dr \quad (3.5)$$

3.6.5. Genel Ortalama Hassasiyet

Genel Ortalama Hassasiyet (Mean Average Precision - mAP), nesne algılama modellerinin doğruluğunu değerlendirmek için kullanılan en yaygın performans

ölçütlerinden biridir. Bu metrik, modelin farklı IoU (Intersection over Union) eşik değerlerinde elde ettiği Ortalama Hassasiyet (Average Precision - AP) değerlerinin ortalamasını ifade eder. Başka bir deyişle mAP, modelin farklı eşik seviyelerinde doğru tespitler yapma yeteneğini bütüncül biçimde değerlendirir. Bu sayede yalnızca tek bir eşik değerine bağlı kalmadan, modelin genel tespit performansı daha güvenilir biçimde ölçülür.

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (3.6)$$



4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA

Bu bölümde, Faster R-CNN mimarisinde farklı omurga yapıları kullanılarak eğitilen modellerin doğrulama ve test verileri üzerindeki performansları kapsamlı olarak analiz edilmiştir. Performans değerlendirmesi, nesne tespit literatüründe yaygın olarak kullanılan çoklu metrikler üzerinden gerçekleştirilmiş ve modeller arasında detaylı karşılaştırmalar yapılmıştır.

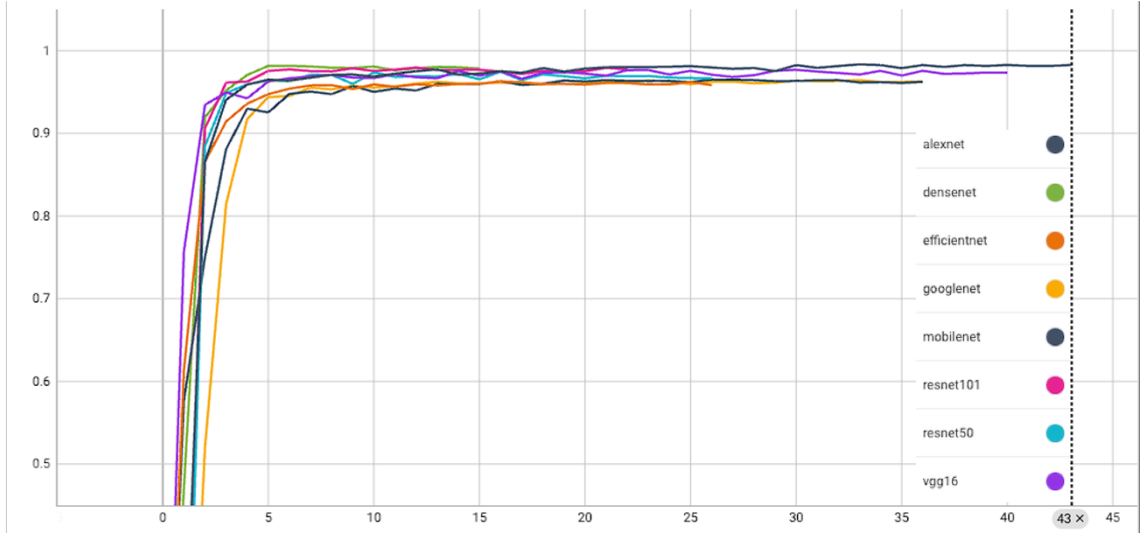
4.1. Doğrulama (Validation) Sonuçları

Doğrulama seti üzerinde gerçekleştirilen performans değerlendirmesi, modellerin eğitim sürecindeki gelişimini izlemek ve en iyi model ağırlıklarını belirlemek amacıyla her dönem sonunda gerçekleştirilmiştir. TensorBoard kullanılarak, tüm doğrulama (validation) metrikleri gerçek zamanlı olarak kaydedilmiş ve görselleştirilmiştir. Bu aşamada elde edilen metrikler, modellerin eğitim sırasındaki öğrenme dinamiklerini ve yakınsama davranışlarını anlamak için kritik öneme sahiptir.

Doğrulama metrikleri, eğitim sürecinde her dönem sonunda hesaplanmaktadır ve modelin genelleme yeteneğinin bir göstergesidir. Bu metrikler, eğitim verisi dışında ayrılmış doğrulama kümesi üzerindeki başarıyı belirler ve olası aşırı uyum belirtilerini izlemeye imkân tanır. Erken durdurma mekanizması, doğrulama (validation) mAP50 metriğine göre çalıştığından, bu bölümde sunulan sonuçlar aynı zamanda modellerin optimal durma noktalarını da belirlemektedir.

Tüm omurga modelleri için gerçekleştirilen eğitim süreçlerine ait doğrulama performans grafikleri Şekil 4.1 ve Şekil 4.2'de sunulmaktadır. Bu grafikler, sekiz farklı omurga modelinin (AlexNet, ResNet-50, ResNet-101, GoogLeNet (Inception-v3), VGG-16, DenseNet-121, EfficientNet-B0 ve MobileNet-V2) doğrulama performansını devir sayısı bazında göstermektedir ve karşılaştırmalı analiz yapılmasına olanak tanımaktadır.

Şekil 4.1'de görülen doğrulama mAP50 metrikleri, modellerin 0,5 Kesişim üzeri Birleşim (IoU) eşliğindeki ortalama hassasiyet performanslarını göstermektedir. Bu metrik, nesne tespit sistemlerinde yaygın olarak kullanılan temel performans göstergelerinden biri olup tespit edilen sınırlayıcı kutuların gerçek değerle en az %50 örtüşme oranına sahip olduğu durumları değerlendirir.

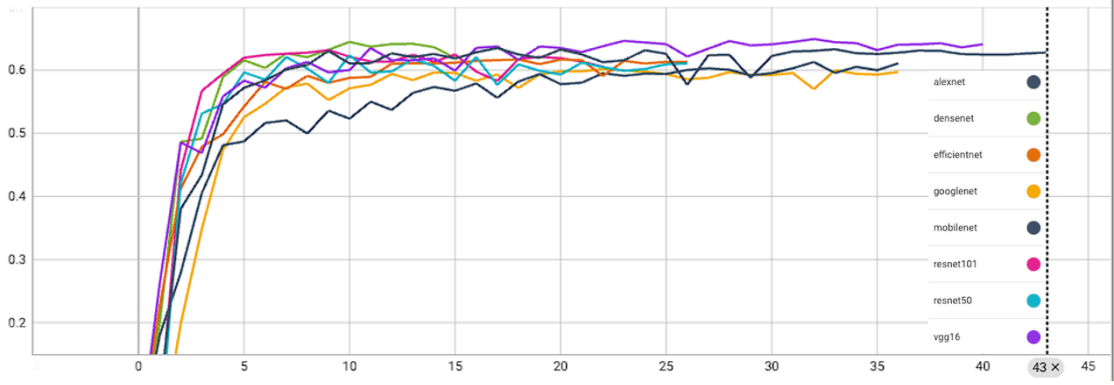


Şekil 4.1. Doğrulama kümesi üzerinde hesaplanan mAP50 (IoU=0,5) değerleri

Çizelge 4.1. Doğrulama kümesi üzerinde hesaplanan mAP50 (IoU=0,5) değerleri

Parametre	Değer	Devir sayısı
AlexNet	0,962	36
EfficientNet-B0	0,958	26
DenseNet-121	0,978	15
GoogLeNet (Inception-v3)	0,962	36
MobileNet-V2	0,982	43
ResNet-50	0,966	26
ResNet-101	0,977	22
VGG-16	0,973	40

Şekil 4.1'de sunulan doğrulama kümesi mAP50 değerlerinin devir sayılarına göre değişimi ve Çizelge 4.1'de yer alan nihai sonuçlar, farklı omurga ağlarının performanslarını karşılaştırmalı olarak ortaya koymaktadır. Elde edilen bulgulara göre, MobileNetV2 mimarisi 0,982 mAP50 değeri ile en yüksek performansı sergilemiştir. Bu mimariyi, 0,978 değeri ile DenseNet-121, 0,973 değerleri ile ResNet-101 ve VGG-16 mimarileri çok yakın sonuçlarla takip etmektedir. Tüm modellerin eğitimin ilk beş devir sayısında mAP50 metriğinde belirgin bir artış gösterdiği, bu artışın sonraki devir sayılarında yavaşlayarak stabil değerlere ulaştığı gözlemlenmiştir. Bu durum, transfer öğrenme yaklaşımının modellerin ImageNet üzerinde önceden öğrenilmiş ağırlıklardan etkin biçimde yararlandığını göstermektedir. Ayrıca, erken durdurma mekanizmasının etkin şekilde çalıştığı ve gereksiz eğitim iterasyonlarının başarılı bir şekilde önleniği gözlemlenmiştir. Genel olarak değerlendirildiğinde tüm modellerin doğrulama kümesinde son derece yüksek mAP50 performansı sergilediği görülmektedir. Elde edilen değerler 0,958 ile 0,982 aralığında yer almaktadır. Bu sonuçlar, IoU eşik değeri 0,5 olarak belirlendiğinde tüm modellerin glomerül tespitinde son derece başarılı olduğunu ortaya koymaktadır.



Şekil 4.2. Doğrulama kümesi üzerinde hesaplanan AP@0.50:95 (IoU=0.5) değerleri

Şekil 4.2'de görülen doğrulama mAP50-95 metrikleri, farklı IoU eşikleri (0,5'ten 0,95'e kadar 0,05 artışlarla) üzerindeki ortalama hassasiyet performanslarını göstermektedir. Bu metrik, COCO değerlendirme standardında kullanılan kapsamlı bir performans göstergesi olup, modellerin farklı lokalizasyon hassasiyeti seviyelerindeki performansını değerlendirmektedir.

Çizelge 4.2. Doğrulama kümesi üzerinde hesaplanan mAP50-95 (IoU=0,5-0,95) değerleri.

Parametre	Değer	Devir sayısı
AlexNet	0,610	36
EfficientNet-B0	0,613	26
DenseNet-121	0,619	15
GoogLeNet (Inception-v3)	0,597	36
MobileNet-V2	0,628	43
ResNet-50	0,610	26
ResNet-101	0,605	22
VGG-16	0,641	40

Şekil 4.2'de sunulan doğrulama kümesi mAP50-95 metriğinin devir sayılarına göre değişimi ve Çizelge 4.2'de yer alan nihai sonuçlar, farklı omurga ağlarının performanslarını karşılaştırmalı olarak ortaya koymaktadır. mAP50-95 metriğindeki genel performans seviyelerinin mAP50 sonuçlarına göre belirgin şekilde daha düşük olması beklenmekte olup literatürde yaygın olarak gözlemlenen bir durumdur. Bu durum, metriğin daha yüksek IoU eşiklerinde de doğru tespit yapılmasını gerektirmesinden kaynaklanmaktadır. Analiz sonuçları, modellerin mAP50-95 performanslarının 0,59 ile 0,64 aralığında yoğunlaştığını göstermektedir. Elde edilen bulgulara göre, VGG-16 mimarisi 0,641 mAP50-95 değeri ile en yüksek performansı sergilemiştir.

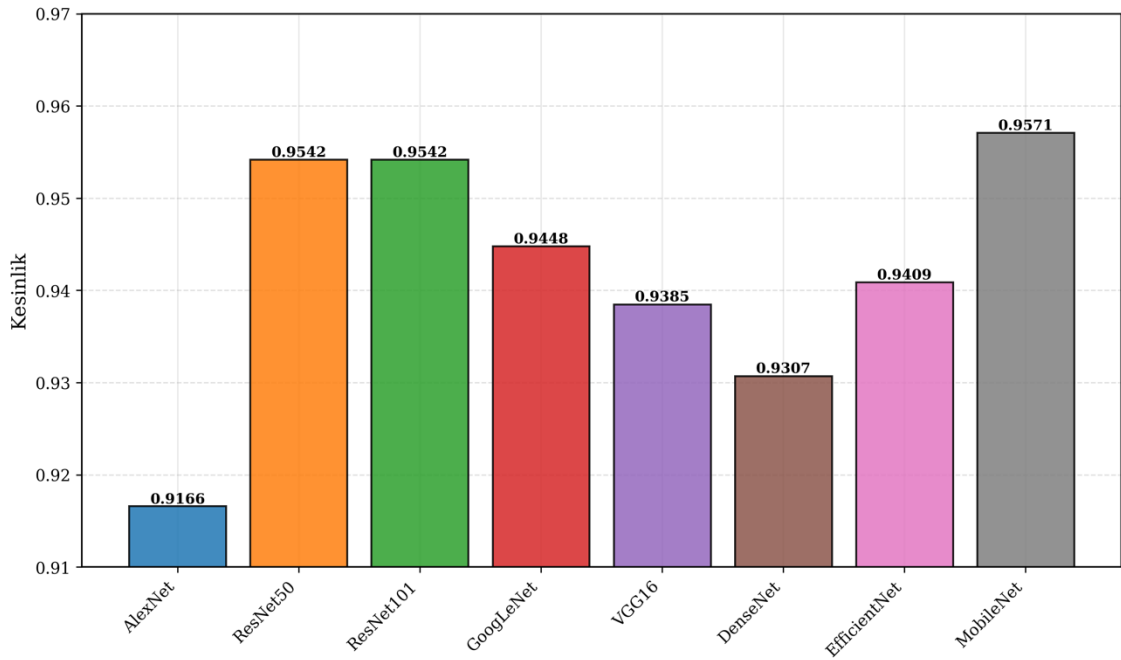
4.2. Test Sonuçları

Eğitim ve doğrulama süreçleri tamamlanan modellerin gerçek performanslarını değerlendirmek amacıyla hiçbir aşamada kullanılmamış bağımsız bir test veri seti üzerinde değerlendirme yapılmıştır. Her model için erken durdurma mekanizması

tarafından belirlenen en iyi ağırlıklar kullanılarak test seti üzerinde çıkarım yapılmış ve ilgili performans metrikleri hesaplanmıştır.

4.2.1. Kesinlik (Precision) performansı

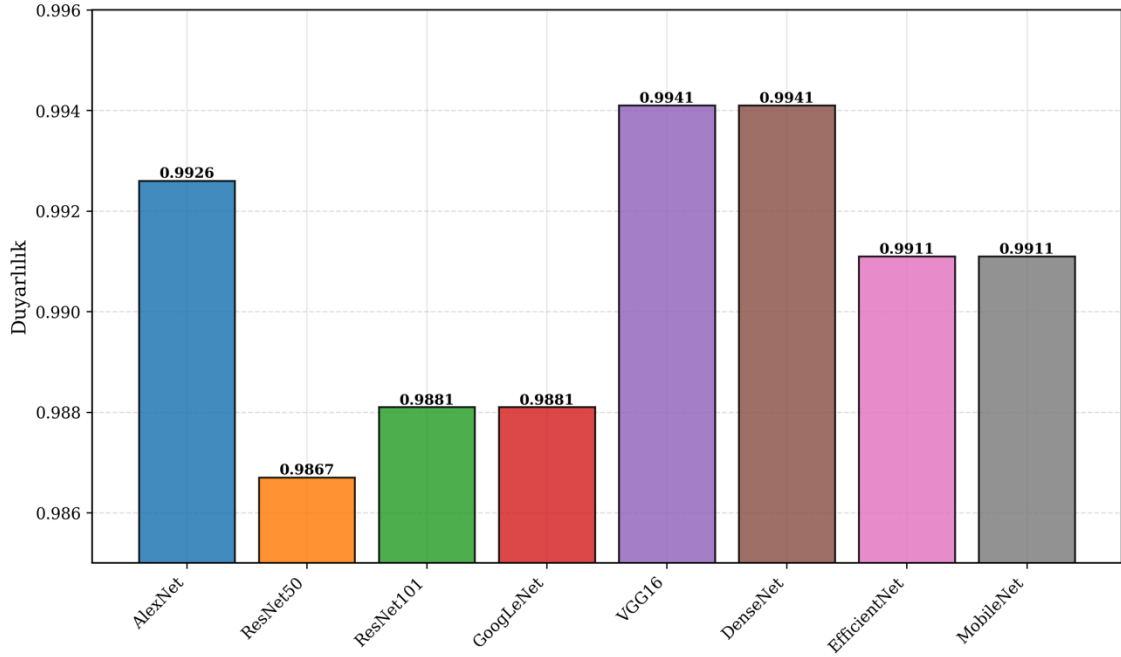
Kesinlik metriği, modelin pozitif olarak sınıflandırdığı örnekler arasında gerçekten pozitif olanların oranını ifade etmektedir. Şekil 4.3'te sunulan kesinlik değerleri incelendiğinde MobileNet-V2 modelinin 0,957 ile en yüksek kesinlik değerine sahip olduğu görülmektedir. Bu değeri 0,9542 değeri ile ResNet-50 ve ResNet-101 modelleri takip etmektedir.



Şekil 4.3. Modellerin kesinlik değerleri

4.2.2. Duyarlılık (Recall) performansı

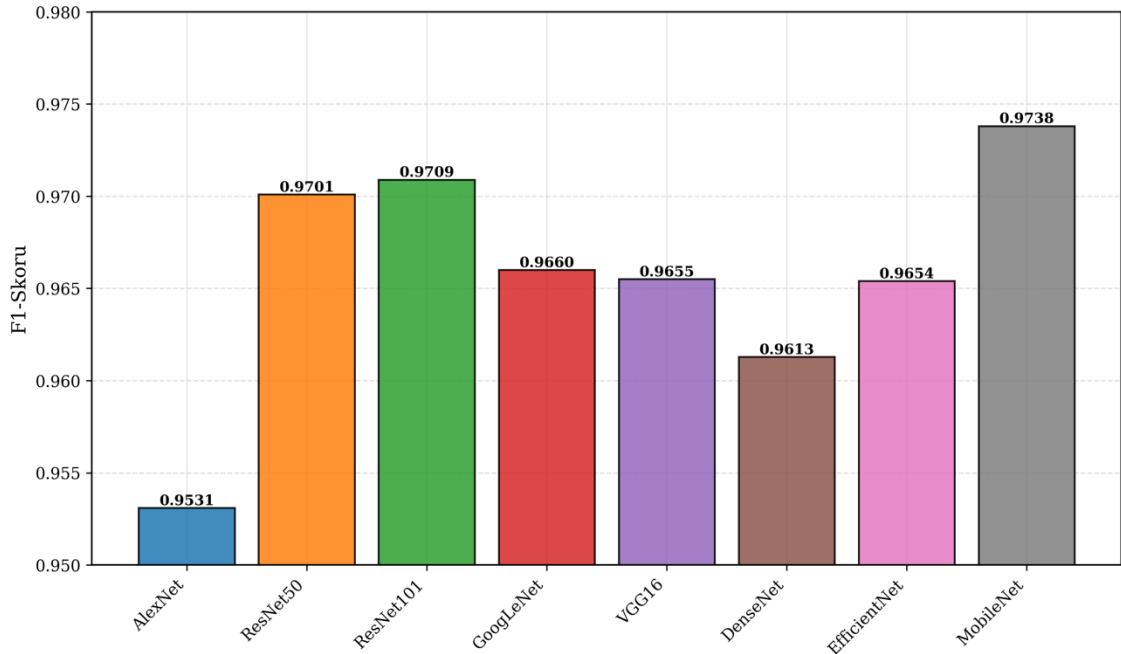
Duyarlılık metriği, gerçekte pozitif olan (glomerül içeren) örneklerin model tarafından doğru şekilde tespit edilen oranını gösterir. Şekil 4.4'te sunulan duyarlılık değerleri, DenseNet-121 ve VGG-16 modelinin 0,994 ile en yüksek duyarlılığa sahip olduğunu ortaya koymaktadır.



Şekil 4.4. Modellerin duyarlılık değerleri

4.2.3. F-1 Skor performansı

F1-skor, kesinlik ve duyarlılık metriklerinin harmonik ortalaması olup, her iki metriğin dengeli bir şekilde değerlendirilmesini sağlamaktadır. Şekil 4.5'te görüldüğü üzere MobileNet-V2 modeli 0,973 ile en yüksek F1-skor değerine ulaşmıştır.

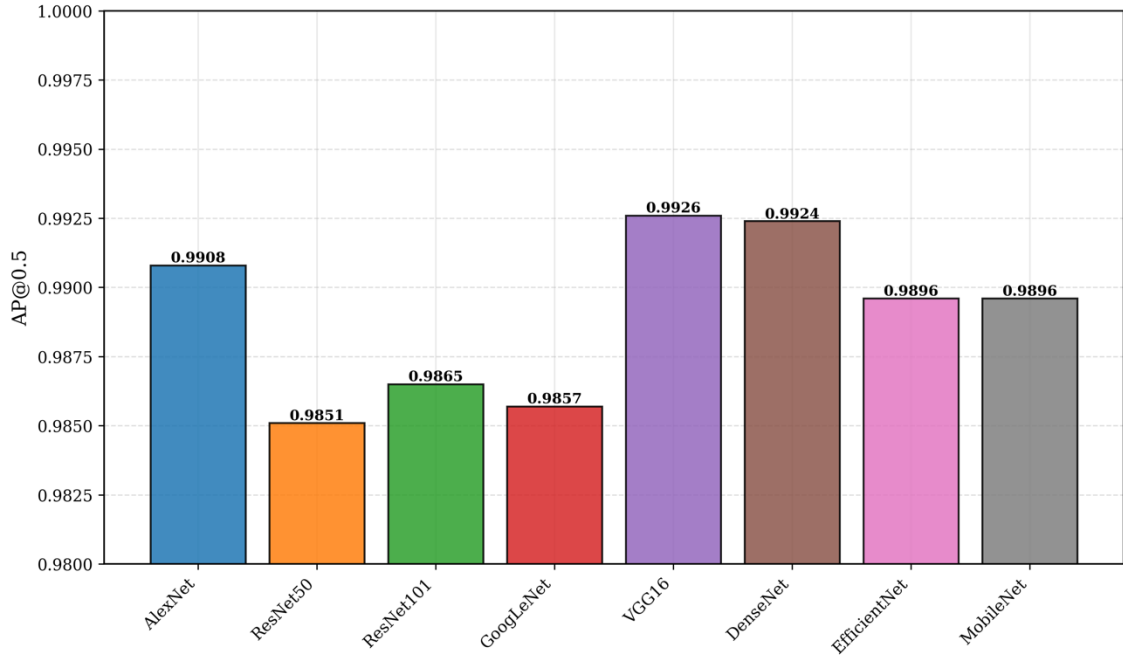


Şekil 4.5. Modellerin f-1 skor değerleri

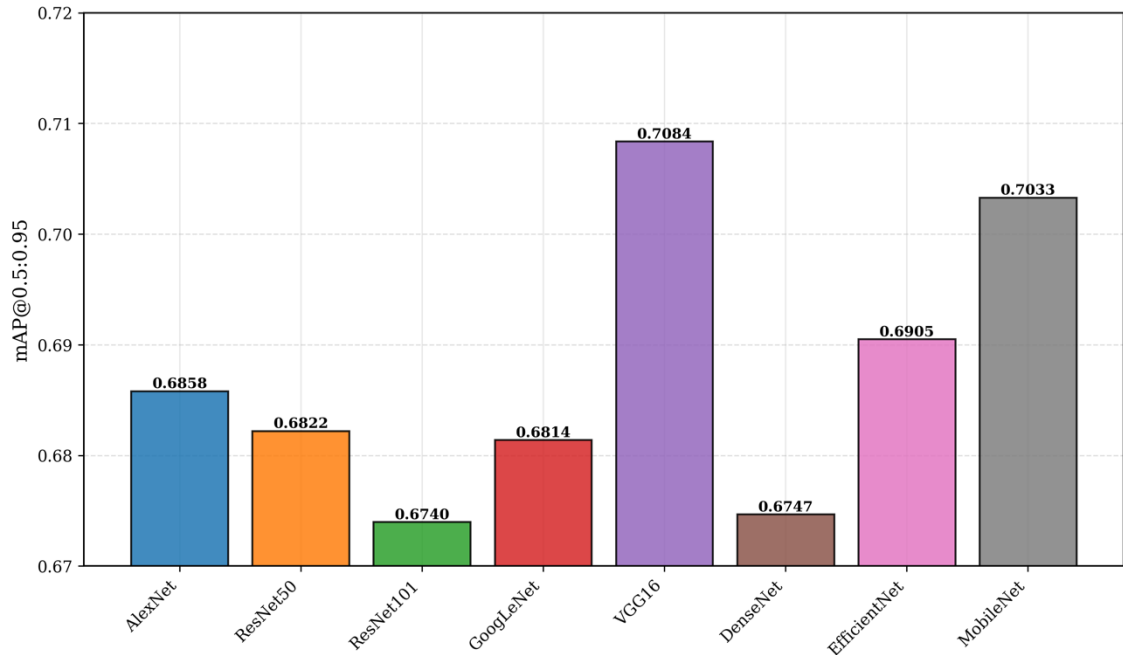
4.2.4. Ortalama Hassasiyet (Average Precision) performansı

Şekil 4.6 ve Şekil 4.7'de modellerin farklı eşik değerlerinde mAP50 ve mAP50-95 performansları karşılaştırılmıştır. mAP50 metriği, 0,5 IoU eşikinde precision-recall

eğrisinin altında kalan alanı temsil ederken, mAP50-95 metriği 0,5 ile 0,95 arasındaki farklı IoU eşiklerinin ortalamasını vermektedir.



Şekil 4.6. Modellerin mAP50 değerleri



Şekil 4.7. Modellerin mAP50-95 değerleri

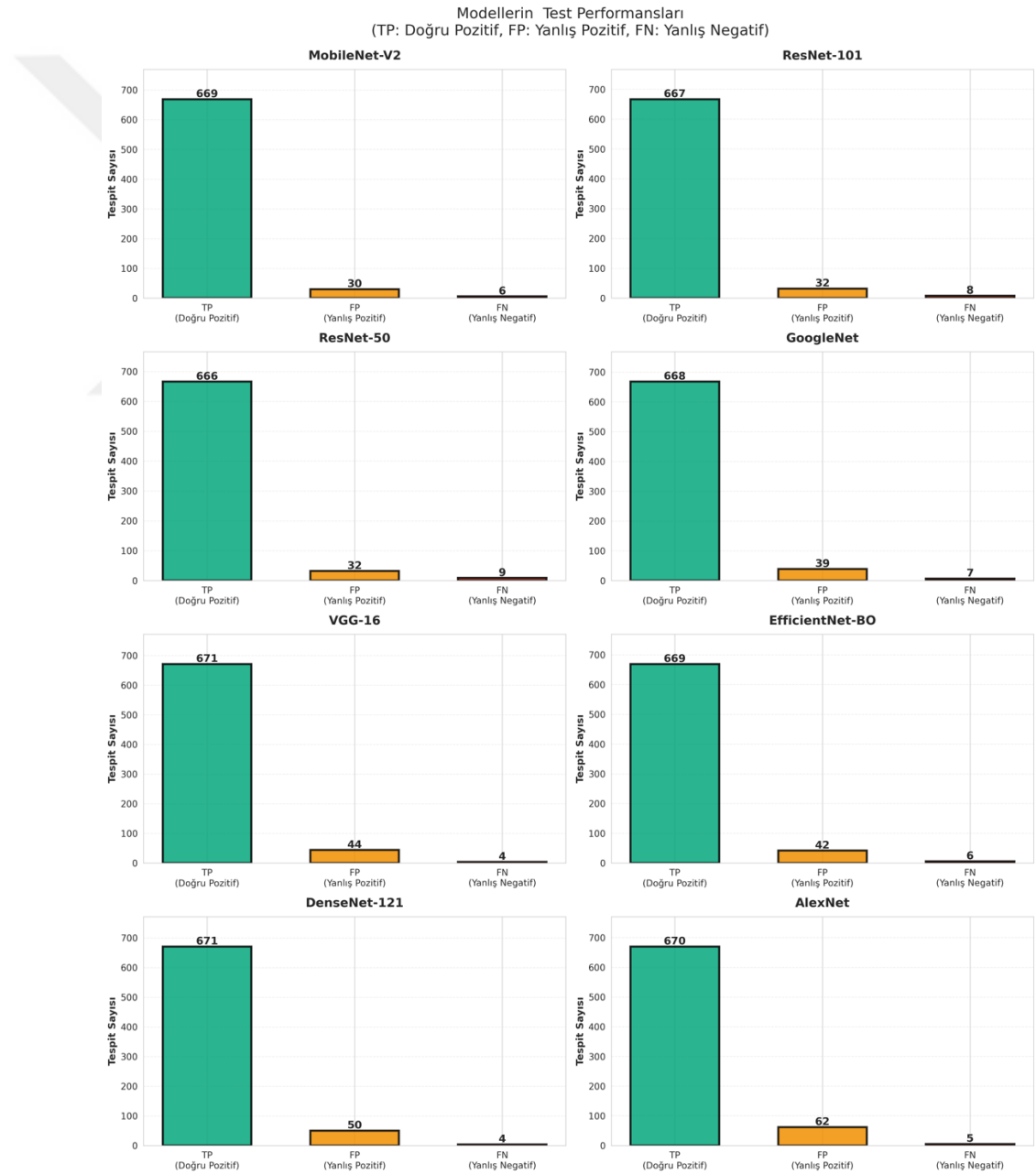
Şekil 4.6 ve Şekil 4.7'de geliştirilen sekiz farklı Faster R-CNN modelinin mAP50 ve mAP50-95 metriklerine göre performans karşılaştırması sunulmaktadır.

mAP50 performans karşılaştırması incelendiğinde en yüksek performansın 0,992 değeri ile VGG-16 modelinde elde edildiği görülmektedir. Elde edilen sonuçlara göre tüm modellerin IoU eşiği 0,5 olduğunda oldukça yüksek tespit performansı sergilediğini ortaya koymaktadır.

mAP50-95 metriği açısından değerlendirildiğinde performans dağılımının daha geniş bir aralıkta gerçekleştiği gözlenmektedir. VGG-16 modeli 0,708 değeri ile en yüksek mAP50-95 performansını gösterirken, MobileNet-V2 0,703 modeli de benzer bir performans sergilemiştir.

4.2.5. Modellerin TP, FP ve FN Değerlerinin Karşılaştırılması

Bu çalışmada, glomerül tespiti için geliştirilen Faster R-CNN modellerinin performansı 8 farklı omurga mimarisi kullanılarak test veri seti üzerinde kapsamlı bir şekilde değerlendirilmiştir. Test veri seti, toplam veri setinin %15'ini oluşturmakta olup, 675 adet gerçek etiket (ground truth) glomerül içermektedir. Model değerlendirmelerinde eşik değeri (confidence threshold) 0,5 ve IoU eşiği 0,5 olarak belirlenmiştir.



Şekil 4.8. Modellerin tespit sayıları

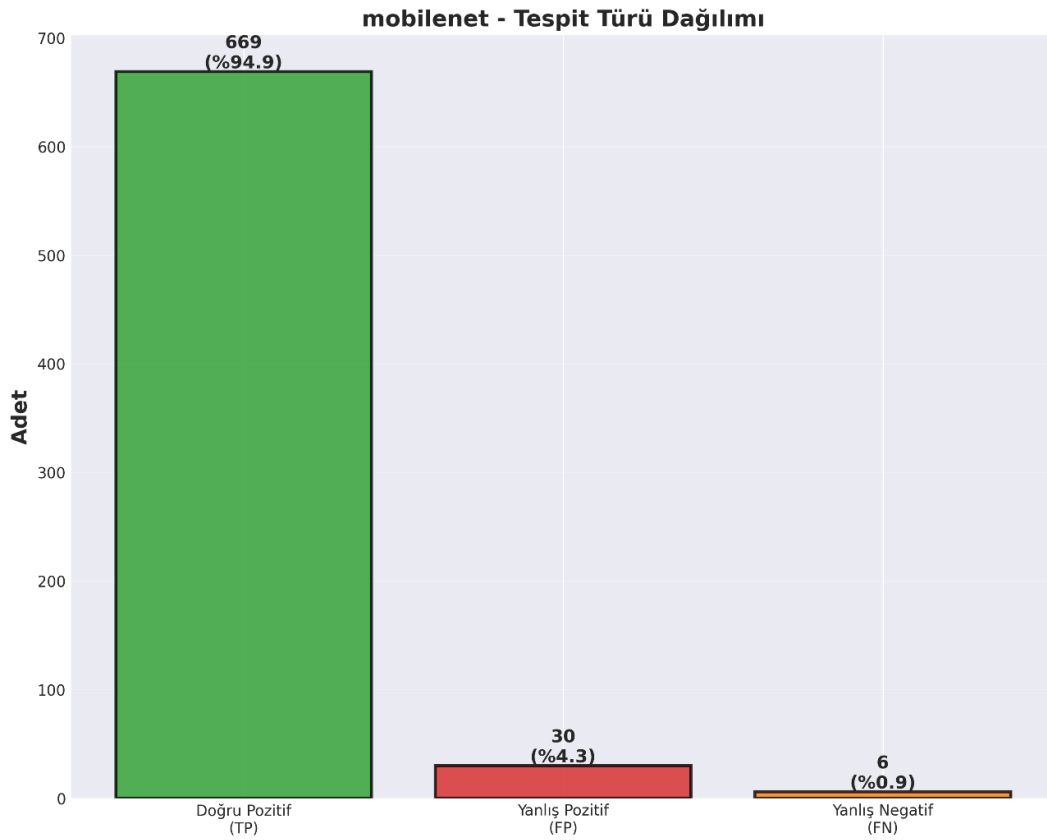
Doğru pozitif, model tarafından doğru şekilde tespit edilen glomerül sayısını ifade ederken, yanlış pozitif model tarafından hatalı olarak glomerül olarak sınıflandırılan bölge sayısını göstermektedir. Yanlış negatif ise model tarafından tespit edilemeyen gerçek glomerül sayısını temsil etmektedir. Çizelge 4.3'te sunulan sonuçlar incelendiğinde, geliştirilen tüm modellerin test veri seti üzerinde yüksek düzeyde performans sergiledikleri görülmektedir. En yüksek F1-skoru 0,973 ile MobileNet-V2 omurga mimarisini kullanan model tarafından elde edilmiştir. Bu modeli sırasıyla 0,970 değerleri ile ResNet-50 ve ResNet-101 modelleri takip etmektedir. MobileNet-V2 modeli 669 TP, 30 FP ve 6 FN değerleriyle 0,957 kesinlik ve 0,991 duyarlılık elde etmiştir. ResNet-101 modeli 667 TP 32 FP ve 8 FN değerleriyle 0,954 kesinlik ve 0,988 duyarlılık göstermiştir. ResNet-50 modeli ise 666 TP, 32 FP ve 9 FN değerleriyle 0,954 kesinlik ve 0,986 duyarlılık sağlamıştır. GoogLeNet (Inception-v3) modeli 668 TP, 39 FP ve 7 FN ile 0,944 kesinlik, 0,989 duyarlılık ve 0,966 F1-skoru elde etmiştir. VGG-16 ve EfficientNet-B0 modelleri 0,965 F1-skoru ile aynı performans seviyesinde yer almış olup, VGG-16 modeli 671 TP, 44 FP ve 4 FN ile 0,938 kesinlik ve 0,994 duyarlılık gösterirken, EfficientNet-B0 modeli 669 TP, 42 FP ve 6 FN ile 0,941 kesinlik ve 0,991 duyarlılık sergilemiştir. DenseNet-121 modeli 671 TP, 50 FP ve 4 FN ile 0,930 kesinlik, 0,994 duyarlılık ve 0,961 F1-skoru elde etmiştir. AlexNet modeli ise 670 TP, 62 FP ve 5 FN ile 0,915 kesinlik, 0,992 duyarlılık ve 0,952 F1-skoru ile en düşük performansı göstermiştir.

Çizelge 4.3. Test kümesi üzerinde hesaplanan performans değerlendirme metrikleri

Parametre	TP	FP	FN	Kesinlik (%)	Duyarlılık (%)	F-1 Skor (%)
AlexNet	670	62	5	91,53	99,26	95,25
EfficientNet-B0	669	42	6	94,10	99,11	96,54
DenseNet-121	671	50	4	93,07	99,41	96,14
GoogLeNet (Inception-v3)	668	39	7	94,48	98,96	96,67
MobileNet-V2	669	30	6	95,71	99,11	97,38
ResNet-50	666	32	9	95,41	98,66	97,01
ResNet-101	667	32	8	95,42	98,81	97,08
VGG-16	671	44	4	93,84	99,41	96,54

Test sonuçları bakıldığında tüm modellerin yüksek tespit başarı oranı sergilediği görülmüştür. Modellerin tamamı %98,66'nın üzerinde duyarlılık değeri elde etmiştir. Bu sonuçlar veri setindeki glomerüllerin büyük çoğunluğunun başarıyla tespit edildiğini göstermektedir. En yüksek duyarlılık değeri 0,994 ile VGG-16 ve DenseNet-121 modellerinde gözlenmiştir. Bu modeller gerçek glomerüllerin neredeyse tamamını tespit edebilmiştir. Düşük yanlış pozitif oranlarının elde edilmesi, çalışmanın bir diğer önemli bulgusunu oluşturmaktadır. MobileNet-V2 modeli yalnızca 30 yanlış pozitif tespit ile en düşük FP değerini elde etmiştir. Bu bulgu, söz konusu modelin glomerül olmayan doku

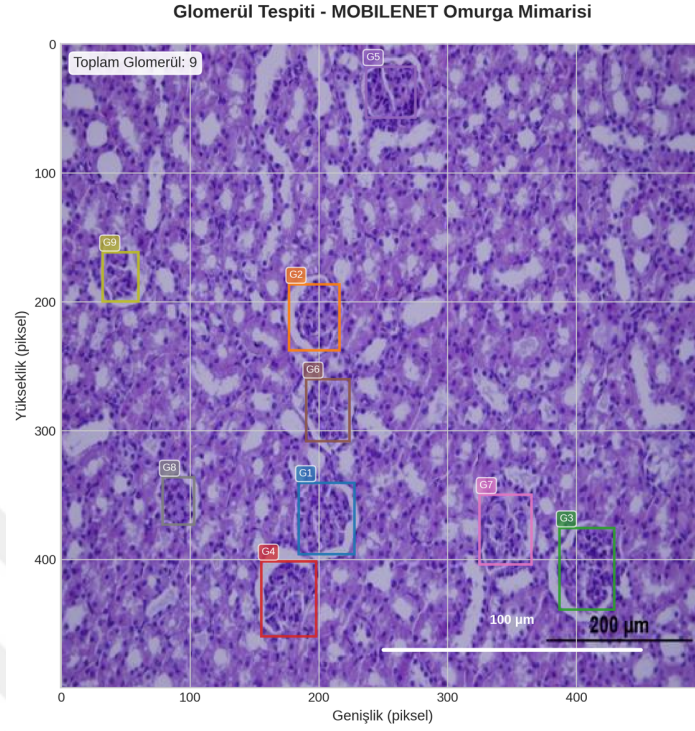
bölgelerini hatalı olarak glomerül olarak sınıflandırma eğiliminin düşük olduğunu ortaya koymaktadır. Kesinlik-duyarlılık dengesi açısından MobileNet-V2, ResNet-50 ve ResNet-101 modellerinin hem kesinlik hem de duyarlılık metriklerinde dengeli ve yüksek performans sergilediği tespit edilmiştir. Bu modeller %95'in üzerinde kesinlik ve %98'in üzerinde duyarlılık değerleri ile kesinlik-duyarlılık (precision-recall) dengesini başarıyla sağlamışlardır. AlexNet modelinin 0,915 kesinlik değeri ile en düşük kesinliğe sahip model olduğu belirlenmiştir. Bu modelin 62 FP tespit sayısı diğer modellere kıyasla belirgin şekilde yüksektir ve modelin ayırt edicilik kapasitesinin görece sınırlı olduğunu göstermektedir.



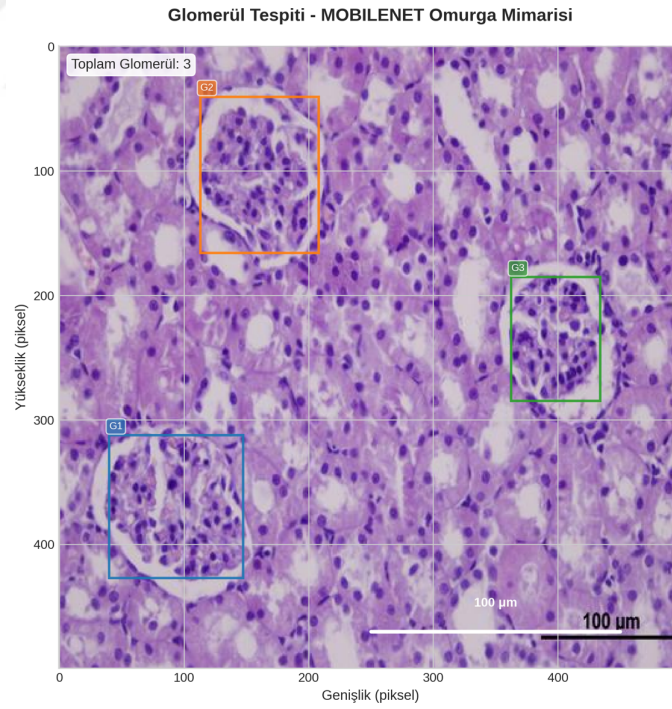
Şekil 4.9. MobileNet-V2 modelinin tespit sayıları

Şekil 4.9'da MobileNet-V2 modelinin test veri seti üzerindeki performansının detaylı dağılımı görselleştirilmiştir. Grafik incelendiğinde, modelin 669 TP tespit gerçekleştirdiği görülmektedir. Buna karşılık, model tarafından FP olarak sınıflandırılan bölge sayısı 30 olup, bu değer doğru pozitif tespit sayısına kıyasla oldukça düşüktür. Ayrıca, FN değeri yalnızca 6 olarak belirlenmiştir. Bu sonuçlar, MobileNet-V2 modelinin glomerül tespitinde yüksek duyarlılığa ve genel olarak başarılı bir performansa sahip olduğunu göstermektedir. Test veri setinde gerçekleştirilen glomerül tespit sonuçları

Şekil 4.10 ve Şekil 4.11’de sunulmuştur. Örnek görseller MobileNet-V2 omurga mimarisi kullanılarak elde edilmiştir.



Şekil 4.10. MobileNet modelinin tespit sonuçlarının görseli



Şekil 4.11. MobileNet modelinin tespit sonuçlarının görseli

5. SONUÇLAR VE ÖNERİLER

Bu bölümde histolojik böbrek doku görüntülerinde glomerül tespiti için sekiz farklı omurga mimarisi kullanılarak gerçekleştirilen çalışmadan elde edilen bulgular özetlenmekte ve değerlendirilmektedir. Ayrıca gelecekte yapılacak araştırmalar için öneriler sunulmaktadır. Sonuç olarak, bu çalışma glomerül tespiti görevinde farklı omurga mimarilerinin sistematik bir karşılaştırmasını sunmuş ve klinik uygulamalar için model seçiminde dikkate alınması gereken temel faktörleri ortaya koymuştur.

5.1. Sonuçlar

Bu tez çalışmasında, histolojik böbrek doku görüntülerinde glomerüllerin otomatik tespiti ve sayımı için Faster R-CNN mimarisi kullanılmıştır. Bu modelin performansı, sekiz farklı omurga (backbone) mimarisi kullanılarak karşılaştırmalı olarak incelenmiştir. Tüm deneyler, 2500 adet 500x500 piksel çözünürlüklü görüntüden oluşan veri kümesi üzerinde aynı eğitim yapılandırılması ve hiper parametreler ile yürütülmüştür. Bu standart yaklaşım sayesinde, omurga mimarisi seçiminin performans üzerindeki temel etkisi incelenmiştir. Modellerin performansını kapsamlı bir şekilde analiz etmek için kesinlik (precision), duyarlılık (recall), F1-skoru ve Ortalama Hassasiyet (Average Precision) metrikleri incelenmiştir.

Kesinlik metriği modelin yaptığı pozitif tahminlerin ne kadarının doğru olduğunu göstermektedir. Bu açıdan yapılan değerlendirmede, MobileNet-V2 modeli 0,957 değeri ile en yüksek kesinlik oranına ulaşmıştır. Bu sonuç, modelin 699 pozitif tespit içerisinden yalnızca 30 tanesinin FP olduğunu göstermektedir. ResNet-101 0,954 ve ResNet-50 0,954 modelleri de benzer kesinlik değerleri sergilemiştir. Yüksek kesinlik oranları modellerin glomerül olmayan bölgeleri yanlışlıkla glomerül olarak işaretleme eğiliminin düşük olduğunu ortaya koymaktadır. Bu durum klinik uygulamalarda yanlış alarm sayısının azaltılması açısından önemlidir.

Duyarlılık metriği modelin veri setindeki gerçek glomerüllerin ne kadarını tespit edebildiğini göstermektedir. Bu metrikte DenseNet-121 ve VGG-16 modelleri 0,994 ile en yüksek performansı göstermiştir. Her iki model de toplam 675 gerçek glomerülden yalnızca 4 tanesini tespit edememiştir. AlexNet modeli 0,992 duyarlılık oranı ile bu modelleri yakından takip etmiştir. Yüksek duyarlılık değerleri, modellerin mevcut glomerüllerin büyük çoğunluğunu başarılı bir şekilde tespit ettiğini göstermektedir.

Kesinlik ve duyarlılık metrikleri arasındaki dengeyi ölçen F1-skoru metriği de analiz edilmiştir. Bu değerlendirmede, MobileNet-V2 modeli 0,973 değeriyle en dengeli performansı göstermiştir. Bu modeli sırasıyla 0,970 F1-skoru değeriyle ResNet-50 ve

ResNet-101 modelleri benzer başarı göstermiştir. Bu sonuçlar, bu modellerin hem yanlış pozitif hem de yanlış negatif oranlarını başarılı bir şekilde optimize ettiğine işaret etmektedir. Özellikle MobileNet-V2'nin hem kesinlik hem de F1-skoru metriklerinde üstün performans göstermesi kısıtlı ya da yetersiz kaynağı olduğu ortamlarda kullanılabilirlik açısından önemli bir avantaj sunmaktadır.

Ortalama hassasiyet performans karşılaştırması incelendiğinde en yüksek performansın 0,992 değeri ile VGG-16 modelinde elde edildiği görülmektedir. Elde edilen sonuçlar, tüm modellerin IoU eşiği 0,5 olduğunda oldukça yüksek tespit performansı sergilediğini ortaya koymaktadır. mAP50-95 metriği açısından değerlendirildiğinde, performans dağılımının daha geniş bir aralıkta gerçekleştiği gözlenmektedir. Bu metrik daha hassas lokalizasyon doğruluğu gerektiren ve farklı IoU eşiklerini (0,5'ten 0,95'e) kapsayan bir ölçümdür. VGG-16 modeli 0,708 değeri ile en yüksek mAP50-95 performansını gösterirken, MobileNet-V2 0,703 modeli değeri ile de benzer bir performans sergilemiştir. Bu durum, modellerin tespit başarısının beklenen lokalizasyon hassasiyetine göre değiştiğini göstermektedir.

Sonuç olarak bu çalışma glomerül tespiti görevinde farklı omurga mimarilerinin sistematik bir karşılaştırmasını sunmuş ve klinik uygulamalar için model seçiminde dikkate alınması gereken temel faktörleri ortaya koymuştur.

5.2. Öneriler

Bu çalışmada elde edilen sonuçlar doğrultusunda gelecekteki araştırmalar için öncelikli olarak geliştirilen modelin genelleme kabiliyetinin artırılması amacıyla, farklı boyama teknikleri ve hastalık evrelerini içeren daha kapsamlı veri setleri üzerinde test edilmesi ve gerçek klinik ortamlarda doğrulanması önerilmektedir. Klinik kabul ve güvenilirlik açısından, Grad-CAM, SHAP ve LIME gibi açıklanabilir yapay zekâ tekniklerinin uygulanması, modellerin karar verme süreçlerinin görselleştirilmesine ve patolojlara daha güvenilir bir karar destek sistemi sunulmasına olanak tanıyacaktır. Bu öneriler doğrultusunda yapılacak gelecek çalışmalar, histolojik böbrek görüntü analizi alanında daha güvenilir, etkili ve klinik olarak kullanılabilir sistemlerin geliştirilmesine önemli katkılar sağlayacağı öngörülmektedir.

KAYNAKLAR

- Akshatha, K., Karunakar, A. K., Shenoy, S. B., Pai, A. K., Nagaraj, N. H. ve Rohatgi, S. S., 2022, Human detection in aerial thermal images using faster R-CNN and SSD algorithms, *Electronics*, 11 (7), 1151.
- Akyol, M. C., Derin öğrenme yöntemleri ile görüntü üzerinde diyabetik ayak ülserinin tespit edilmesi, *Fen Bilimleri Enstitüsü*.
- Albahli, S., Nazir, T., Irtaza, A. ve Javed, A., 2021, Recognition and Detection of Diabetic Retinopathy Using Densenet-65 Based Faster-RCNN, *Computers, Materials & Continua*, 67 (2).
- Benali Amjoud, A. ve Amrouch, M., 2020, Convolutional neural networks backbones for object detection, *International conference on image and signal processing*, 282-289.
- Bhatt, S. D., Astik, M. B. ve Soni, H. B., 2025, Faster R-CNN vs SSD: A Deep Learning Approach for Nodule Detection in Lung Cancer Diagnosis, *2025 3rd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT)*, 410-414.
- Bihanda, Y. G., Fatichah, C. ve Yuniarti, A., 2023, Comparative Analysis of ConvNext and Mobilenet on Traffic Vehicle Detection, *2023 IEEE 8th International Conference On Software Engineering and Computer Systems (ICSECS)*, 101-105.
- Celik, M. E., 2022, Deep learning based detection tool for impacted mandibular third molar teeth, *Diagnostics*, 12 (4), 942.
- Chen, Z., Wang, L., Li, B., Chen, S., Bian, J., Zheng, F. ve Deng, Y., 2022, An improved faster R-CNN transmission line bolt defect detection method, *2022 World Automation Congress (WAC)*, 82-85.
- Deng, X., Liang, S., Xu, Y., Gong, K., Zhong, Z., Chen, X. ve Chen, Y., 2021, Object detection of alternanthera philoxeroides at seedling stage in paddy field based on faster R-CNN, *2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 1125-1129.
- Devi, B. S., Nagaraja, K. ve Singh, R. P., 2024, Optimization Enhancements for Faster R-CNN in Floating Bottle Detection, *2024 IEEE 11th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, 1-7.
- Girshick, R., Donahue, J., Darrell, T. ve Malik, J., 2014, Rich feature hierarchies for accurate object detection and semantic segmentation, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 580-587.
- Girshick, R., 2015, Fast r-cnn, *Proceedings of the IEEE international conference on computer vision*, 1440-1448.
- Gong, J. ve Zhang, K., 2021, Automatic diagnosis system for lumbar intervertebral disc, *2021 3rd International Conference on Natural Language Processing (ICNLP)*, 267-271.
- Hakim, L., Hendrawan, A. ve Khoiriyah, R., 2024, Traffic Vehicle Detection Using Faster R-CNN: A Comparative Analysis of Backbone Architectures, *International Journal of Artificial Intelligence and Science*, 1 (1), 50-62.
- Hasan, M. M., Sakib, S. ve Deb, K., 2022, Road damage detection and classification using deep neural network, *2022 4th International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)*, 1-6.
- He, K., Zhang, X., Ren, S. ve Sun, J., 2016, Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770-778.

- Huang, G., Liu, Z., Van Der Maaten, L. ve Weinberger, K. Q., 2017, Densely connected convolutional networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4700-4708.
- Huynh, Q. T., Nguyen, P. H., Le, H. X., Ngo, L. T., Trinh, N.-T., Tran, M. T.-T., Nguyen, H. T., Vu, N. T., Nguyen, A. T. ve Suda, K., 2022, Automatic acne object detection and acne severity grading using smartphone images and artificial intelligence, *Diagnostics*, 12 (8), 1879.
- İnik, Ö. ve Ülker, E., 2017, Derin öğrenme ve görüntü analizinde kullanılan derin öğrenme modelleri, *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, 6 (3), 85-104.
- Jasitha, P. ve Pournami, P., 2023, Glomeruli Detection Using Faster R-CNN and CenterNet, *2023 3rd Asian Conference on Innovation in Technology (ASIANCON)*, 1-6.
- Kawazoe, Y., Shimamoto, K., Yamaguchi, R., Shintani-Domoto, Y., Uozaki, H., Fukayama, M. ve Ohe, K., 2018, Faster R-CNN-based glomerular detection in multistained human whole slide images, *Journal of Imaging*, 4 (7), 91.
- Khan, H. U., Raza, B., Shah, M. H., Usama, S. M., Tiwari, P. ve Band, S. S., 2023, SMDetector: Small mitotic detector in histopathology images using faster R-CNN with dilated convolutions in backbone model, *Biomedical Signal Processing and Control*, 81, 104414.
- Kharb, A. ve Chaudhary, P., 2024, Designing efficient brain tumor classifier using hybrid EfficientNet-faster R-CNN deep learning model, *Engineering Research Express*, 6 (3), 035216.
- Krizhevsky, A., Sutskever, I. ve Hinton, G. E., 2012, Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, 25.
- Kumar, D., Zhang, X., Su, H. ve Wei, S., 2019, Accurate object detection based on faster R-CNN in remote sensing imagery, *2019 6th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, 1-6.
- Ma, J., Li, H., Xi, J., Zha, G., Yue, Y. ve Yang, H., 2024, UAV target tracking and detection based on faster r-cnn improved networks, *2024 IEEE 2nd International Conference on Control, Electronics and Computer Technology (ICCECT)*, 751-755.
- Ma, S., Huang, Y., Che, X. ve Gu, R., 2020, Faster RCNN-based detection of cervical spinal cord injury and disc degeneration, *Journal of applied clinical medical physics*, 21 (9), 235-243.
- Manjari, K., Verma, M. ve Singal, G., 2023, Path Texture Detection for Visually Impaired using Transfer Learning, *2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC)*, 726-731.
- Middya, A. I., Chattopadhyay, D. ve Roy, S., 2021, Garbage detection and classification using faster-RCNN with inception-V2, *2021 IEEE 18th India Council International Conference (INDICON)*, 1-6.
- Öztürk, E., 2022, Uydu görüntülerinden evrimsel sinir ağları ile bina bölge tespiti ve çatı tipi sınıflandırması, *Başkent Üniversitesi Fen Bilimler Enstitüsü*.
- Pandey, D. R. ve Raghatat, K. S., 2025, Real-Time Traffic Sign Detection and Recognition Using Faster R-CNN with VGG16 Backbone, *2025 International Conference on Automation and Computation (AUTOCOM)*, 279-284.
- Popescu, D., Ichim, L., Dimoiu, M. ve Trufealea, R., 2022, Comparative Study of Neural Networks Used in Halyomorpha Halys Detection, *2022 30th Mediterranean Conference on Control and Automation (MED)*, 182-187.

- Ren, S., He, K., Girshick, R. ve Sun, J., 2015, Faster r-cnn: Towards real-time object detection with region proposal networks, *Advances in neural information processing systems*, 28.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. ve Chen, L.-C., 2018, Mobilenetv2: Inverted residuals and linear bottlenecks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4510-4520.
- Sert, E., 2021, A deep learning based approach for the detection of diseases in pepper and potato leaves, *Anadolu Tarım Bilimleri Dergisi*, 36 (2), 167-178.
- Simonyan, K. ve Zisserman, A., 2014, Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556*.
- Somasundaram, S. ve Basak, S., 2025, Detection and Classification of the Stage of Leukemia From Blood Smear Images Using a VGG11 Backbone-Based Faster R-CNN Network, *2025 9th International Conference on Inventive Systems and Control (ICISC)*, 1-7.
- Sood, S., Singh, H., Khan, S. B. ve Almusharraf, A., 2025, Leveraging the WFD2020 Dataset for Multi-Class Detection of Wheat Fungal Diseases with YOLOv8 and Faster R-CNN, *Computers, Materials & Continua*, 84 (2).
- Sorgun, Ö., 2022, X-ray görüntülerinde faster R-CNN kullanılarak yasaklı nesne tespiti, *Sakarya Uygulamalı Bilimler Üniversitesi*.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. ve Rabinovich, A., 2015, Going deeper with convolutions, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1-9.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. ve Wojna, Z., 2016, Rethinking the inception architecture for computer vision, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2818-2826.
- Tan, M. ve Le, Q., 2019, Efficientnet: Rethinking model scaling for convolutional neural networks, *International conference on machine learning*, 6105-6114.
- Traoré, A., Ly, A. O. ve Akhloufi, M. A., 2020, Evaluating deep learning algorithms in pulmonary nodule detection, *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 1335-1338.
- Wang, Q. ve Qi, F., 2019, Tomato diseases recognition based on faster RCNN, *2019 10th International Conference on Information Technology in Medicine and Education (ITME)*, 772-776.
- Wang, Z., Cao, Y. ve Li, J., 2023, A detection algorithm based on improved faster R-CNN for spacecraft components, *2023 IEEE International Conference on Image Processing and Computer Applications (ICIPCA)*, 1-5.
- Yang, D., Huang, C. ve Tan, H., 2021, Study on Transmission Line Image Defect Recognition Algorithm based on Faster-RCNN, *2021 2nd International Conference on Electronics, Communications and Information Technology (CECIT)*, 1190-1193.
- Zhang, H., Lin, P., He, J. ve Chen, Y., 2020, Accurate strawberry plant detection system based on low-altitude remote sensing and deep learning technologies, *2020 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD)*, 1-5.