

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**DERİN ÖĞRENMEYE DAYALI NESNE TESPİTİ
MODELLERİ İLE CİLT KANSERİ TESPİTİ**

YÜKSEK LİSANS TEZİ

Sarah Atheer Abbas AL-MUSAWI

Bilişim Sistemleri Mühendisliği Anabilim Dalı

EYLÜL 2025

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**DERİN ÖĞRENMEYE DAYALI NESNE TESPİTİ
MODELLERİ İLE CİLT KANSERİ TESPİTİ**

YÜKSEK LİSANS TEZİ

Sarah Atheer Abbas AL-MUSAWI

Bilişim Sistemleri Mühendisliği Anabilim Dalı

Tez Danışmanı: Prof. Dr. İsmail Hakkı CEDİMOĞLU

EYLÜL 2025

Sarah Atheer Abbas AL-MUSAWI tarafından hazırlanan “Derin Öğrenmeye Dayalı Nesne Tespiti Modelleri ile Cilt Kanseri Tespiti” adlı tez çalışması 08.09.2025 tarihinde aşağıdaki jüri tarafından oy birliği/oy çokluğu ile Sakarya Üniversitesi Fen Bilimleri Enstitüsü Bilişim Sistemleri Mühendisliği Anabilim Dalı’nda Yüksek Lisans tezi olarak kabul edilmiştir.

Tez Jürisi

Jüri Başkanı :

Jüri Üyesi :

Jüri Üyesi :



ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Sakarya Üniversitesi Fen Bilimleri Enstitüsü Lisansüstü Eğitim-Öğretim Yönetmeliğine ve Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesine uygun olarak hazırlamış olduğum “DERİN ÖĞRENMEYE DAYALI NESNE TESPİTİ MODELLERİ İLE CİLT KANSERİ TESPİTİ” başlıklı tezin bana ait, özgün bir çalışma olduğunu; çalışmamın tüm aşamalarında yukarıda belirtilen yönetmelik ve yönergeye uygun davrandığımı, tezin içerdiği yenilik ve sonuçları başka bir yerden almadığımı, tezde kullandığım eserleri usulüne göre kaynak olarak gösterdiğimi, bu tezi başka bir bilim kuruluna akademik amaç ve unvan almak amacıyla vermediğimi ve 20.04.2016 tarihli Resmi Gazete’de yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince Sakarya Üniversitesi’nin abonesi olduğu intihal yazılım programı kullanılarak Enstitü tarafından belirlenmiş ölçütlere uygun rapor alındığını, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun ortaya çıkması halinde doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi beyan ederim.

(08/09/2025).

Sarah Atheer Abbas AL-MUSAWI





Bu tezi anneme ve babama ithaf ediyorum.



TEŞEKKÜR

Yüksek lisans eğitimim boyunca bilgi birikimi, yönlendirmeleri ve ilham verici bakış açısıyla bana rehberlik eden saygıdeğer danışmanım Prof.Dr. İsmail Hakkı CEDİMOĞLU'na en derin şükranlarımı sunarım.

Bu çalışmanın ilerlemesinde gösterdiği özverili katkılar, sabrı ve sürekli akademik desteğiyle bana büyük güven veren, her adımda yanımda olup sadece bir akademik danışman değil, aynı zamanda moral ve motivasyon kaynağı olan değerli hocam Dr. Öğretim Üyesi Merve Şişçi'ye sonsuz teşekkür ederim.

Tezimin savunmasına katılarak değerli görüş ve katkılarını sunan Prof. Dr. Khalid ÖZTEKİN'e en içten teşekkürlerimi sunarım. Ayrıca Sakarya Üniversitesi Bilgi Sistemleri Mühendisliği Anabilim Dalı'ndaki hocalarıma sürekli destekleri ve bilimsel yönlendirmeleri için teşekkür ederim.

Tez çalışmam süresince daima yanımda olan, sevgileri, anlayışları ve destekleriyle bana güç veren aileme en içten teşekkürlerimi sunarım. Bu süreç boyunca başarıma olan inancı ve göstermiş olduğu özveri için değerli anneme şükranlarımı arz ederim. Varlığı ve duaları, bu çalışmanın tamamlanmasında en önemli motivasyon kaynağım olmuştur. Ayrıca, zorlu anlarda yanımda olup beni motive eden tüm arkadaşlarıma gönülden teşekkür ederim.

Sarah Atheer Abbas AL-MUSAWI



İÇİNDEKİLER

Sayfa

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ	v
TEŞEKKÜR	ix
İÇİNDEKİLER	xi
KISALTMALAR	xiii
SİMGELER	xv
TABLO LİSTESİ	xvii
ŞEKİL LİSTESİ	xix
ÖZET	xxi
SUMMARY	xxiii
1. GİRİŞ	1
2. LİTERATÜR	5
2.1. Deri Kanseri	5
2.1.1. İyi huylu lezyonlar (benign)	6
2.1.2. Kötü huylu lezyonlar (malignant)	6
2.2. Deri Kanserini Tespit Etmek için Uygulanan Genel Yöntem	7
2.2.1. ABCDE	7
2.3. Deri Kanserinin Sınıflandırılması ile İlgili Yapılan Çalışmalar	8
2.4. Farklı Algoritmaların Sınıflandırılması ile ilgili Yapılan Çalışmalar	11
3. YÖNTEM	15
3.1. Nesne Tespiti	15
3.2. Yolov Algoritması	18
3.2.1. YOLOv8 algoritması	25
3.2.2. YOLOv11 algoritması	28
3.2.3. YOLOv12 algoritması	31
3.3. RT-DETR Algoritması	34
4. METODOLOJİ	39
4.1. Veri Toplama	39
4.2. Veri Etiketleme	40
4.3. Modellerin Eğitilmesi	40
4.4. Değerlendirme Ölçütleri	43
5. SONUÇLAR VE BULGULAR	49
5.1. Performans Ölçütü Değerlerinin Analizi	49
5.1.1. YOLOv8 sonuçları	49
5.1.2. YOLO11 sonuçları	50
5.1.3. YOLOv12 sonuçları	51
5.1.4. RT-DETR	52
5.2. Karmaşıklık Matrisi ve Kayıp Fonksiyonlarının Analizi	52
5.2.1. YOLOv8	52
5.2.1.1. YOLOv8 karmaşıklık matris	52
5.2.1.2. YOLOv8 kayıp fonksiyonunun analizi	55
5.2.2. YOLOv11	57

5.2.2.1. YOLOv11 karmaşıklık matris	57
5.2.2.2. YOLOv11 kayıp fonksiyonun analizi	59
5.2.3. YOLOv12.....	61
5.2.3.1. YOLOv12 karmaşıklık matris	61
5.2.3.2. YOLOv12 kayıp fonksiyonun analizi	62
5.2.4. RT-DETR.....	64
5.2.4.1. RT-DETR algoritmanın karmaşıklık matris.....	64
5.2.4.2. RT-DETR algoritmanın kayıp fonksiyonun analiz	65
6. SONUÇ VE ÖNERİLER.....	67
KAYNAKLAR.....	69
ÖZGEÇMİŞ.....	79



KISALTMALAR

CAD	: Bilgisayar Destekli Tanı
CNNs	: Evrişimli Sinir Ağları
COCO	: Bağlamda Yaygın Nesneler
CSPNet	: Çapraz Aşamalı Kısmi Ağ
FN	: Yanlış Negatif
FP	: Yanlış Pozitif
FPN	: Özellik Piramidi Ağı
FPN	: Özellik Piramidi Ağı
GPU	: grafik işlemci ünitesi
IoU	: Kestirilmiş bölgeler
ISIC	: Uluslararası Cilt Görüntüleme İşbirliği
mAP	: Ortalama Doğruluk Oranı
MITF	: Modelin bir görüntü başına çıkarım süresi
NMS	: Maksimun Olmayan Bastırma
PANet	: Yol Toplama Ağı
R-CNN	: Bölgesel Tabanlı Evrişimlisel Sinir Ağı
RT-DETR	: Gerçek Zamanlı Tespit Dönüştürücüsü
SPPF	: Uzamsal Piramit Havuzlama Hızlı
SSD	: Tek Atış Dedektörü
TP	: Gerçek Pozitif
YOLOv11	: Yalnızca Bir Kere Bak — Sürüm 11
YOLOv12	: Yalnızca Bir Kere Bak — Sürüm 12
YOLOv8	: Yalnızca Bir Kere Bak — Sürüm 8
YZ	: Yapay Zeka



SİMGELER

- $\emptyset(\mathbf{x})$: Aktivasyon fonksiyonu
 λ_{koord} : Kayıp fonksiyonunda yer alan sabit
 $\lambda_{\text{nesneyok}}$: Kayıp fonksiyonunda yer alan sabit





TABLO LİSTESİ

Sayfa

Tablo 4.1. Verisetinin bölünmesi	39
Tablo 4.2. YOLOv8 modellerinin katman ve parametre sayısı.....	41
Tablo 4.3. YOLOv11 modellerinin katman ve parametre sayısı.....	41
Tablo 4.4. YOLOv12 modellerinin katman ve parametre sayısı.....	41
Tablo 4.5. RT-DETR modelinin katman ve parametre sayısı.....	41
Tablo 4.6. Deneysel kurulum için yapılandırma detayları.	42
Tablo 4.7. Geliştirilen modellerin eğitim parametreleri ve değerleri.....	42
Tablo 5.1. YOLOv8 sonuçları.....	50
Tablo 5.2. YOLOv11 sonuçları.....	51
Tablo 5.3. YOLOv12 sonuçları.....	51
Tablo 5.4. RT-DETR sonuçları.....	52



ŞEKİL LİSTESİ

Sayfa

Şekil 2.1. İyi huylu deri kanseri.	6
Şekil 2.2. Kötü huylu deri kanseri.	7
Şekil 3.1. Tek aşamalı (Single-stage) ve iki aşamalı (Two-stage) dedektörler.	16
Şekil 3.2. Deri kanser tespiti.	17
Şekil 3.3. YOLO algoritmasının yıllar içindeki gelişmi.	19
Şekil 3.4. YOLO'nun genel mimarisi (Alif ve ark., 2024).	20
Şekil 3.5. YOLO'nun nesne tespit sistemi (Redom ve ark., 2016).	22
Şekil 3.6. YOLO modeli, her ızgara hücresi için sınırlayıcı kutular ve sınıf olasılıklarını tahmin eder.	24
Şekil 3.7. YOLOv8 mimarisi (Niu ve ark., 2024)	26
Şekil 3.8. YOLOv11 mimarisi (Mao ve Hong, 2024).	29
Şekil 3.9. YOLOv12 mimarisi (Nguyen ve ark., 2025).	32
Şekil 3.10. R-ELAN tasarımı (Tian ve ark., 2025).	33
Şekil 3.11. RT-DETR mimarisi (Jun, 2024).	35
Şekil 4.1. .txt dosyasından oluşan bir sınıfın etiketlemesi	40
Şekil 4.2. IoU değerine göre tespit kalitesi örnekleri.	44
Şekil 4.3. Birleşim üzerinden kesişim (IoU).	46
Şekil 5.1. YOLOv8 modellerine ait karmaşıklık matrisleri (a) YOLOv8n, (b) YOLOv8s, (c) YOLOv8m, (d) YOLOv8l, (e) YOLOv8x	54
Şekil 5.2. YOLOv8n modelinin kayıp fonksiyonu.	55
Şekil 5.3. YOLOv8s modelinin kayıp fonksiyonu.	55
Şekil 5.4. YOLOv8m modelinin kayıp fonksiyonu.	56
Şekil 5.5. YOLOv8l modelinin kayıp fonksiyonu.	56
Şekil 5.6. YOLOv8x modelinin kayıp fonksiyonu.	57
Şekil 5.7. YOLOv11 modellerine ait karmaşıklık matrisleri (a) YOLOv11n, (b) YOLOv11s, (c) YOLOv11m, (d) YOLOv11l, (e) YOLOv11x	58
Şekil 5.8. YOLOv11n modelinin kayıp fonksiyonu.	59
Şekil 5.9. YOLOv11s modelinin kayıp fonksiyonu.	59
Şekil 5.10. YOLOv11m modelinin kayıp fonksiyonu.	60
Şekil 5.11. YOLOv11l modelinin kayıp fonksiyonu.	60
Şekil 5.12. YOLOv11x modelinin kayıp fonksiyonu.	61
Şekil 5.13. YOLOv12 modellerine ait karmaşıklık matrisleri (a) YOLOv12n, (b) YOLOv12s, (c) YOLOv12m, (d) YOLOv12l, (e) YOLOv12x	62
Şekil 5.14. YOLOv12n modelinin kayıp fonksiyonu.	63
Şekil 5.15. YOLOv12s modelinin kayıp fonksiyonu.	63
Şekil 5.16. YOLOv12m modelinin kayıp fonksiyonu.	64
Şekil 5.17. YOLOv12l modelinin kayıp fonksiyonu.	64
Şekil 5.18. RT-DETR-L modelinin karmaşıklık matris.	65
Şekil 5.19. RT-DETR algoritmanın kayıp fonksiyonu.	65



DERİN ÖĞRENMEYE DAYALI NESNE TESPİTİ MODELLERİ İLE CİLT KANSERİ TESPİTİ

ÖZET

Cilt kanseri, en ciddi cilt hastalıklarından biridir ve çoğu vakada ölümcül olabilir. Küresel istatistikler, her yıl bir milyondan fazla yeni vakanın kaydedildiğini ve bu kanser türünün giderek artan sağlık risklerini ortaya koyduğunu göstermektedir. Vaka sayısındaki bu artış, özellikle uzun süre doğrudan güneş ışığına maruz kalma ve aşırı solaryum kullanımı gibi çeşitli sosyal ve estetik faktörlere bağlanmaktadır. Solaryumlar, özellikle gençler arasında cilt kanserinin önde gelen nedenlerinden biridir. Bu çalışmada, derin öğrenmeye dayalı mimarilerden Yalnızca Bir Kere Bak (You Only Look Once-YOLO) algoritmasının YOLOv8, YOLOv11, YOLOv12 olmak üzere çeşitli sürümleri (n-nano, s-small, m-medium, l-large, x-xlarge) ve Gerçek Zamanlı Algılama Dönüştürücüsü (Real-Time Detection Transformer-RT-DETR) algoritmasının cilt kanserinin görüntüler üzerinde erken teşhisinde kullanılabilirliğinin değerlendirilmesi amaçlanmıştır. Bu algoritmaların her biri, nesne tespitinde onları öne çıkaran farklı farklı özelliklere sahiptir. YOLOv8, segmentasyonu ve nokta tespitini iyileştirmek için C2f'yi modülünü entegre etmiştir. YOLOv11, C3k2 ve C2PSA kullanarak omurga ve boyun arasındaki katmanları güçlendirilmiştir. YOLOv12 ise gerçek zamanlı performansı artırmak için bölgesel dikkat mekanizması (A^2) ve (R-ELAN) yapısını sunmuştur ve son olarak RT-DETR, bir maksimum olmayan bastırma (Non-Maximum Suppression-NMS)'ya ihtiyaç duymadan Transformatör mimarisine dayanılır ve bu da onu gerçek zamanlı tespitte yeni bir standart haline getirilmektedir.

Çalışmada, lezyonlar iyi huylu ve kötü huylu olmak üzere iki ana sınıfa ayrılmıştır. Görsel değerlendirmede ABCDE kuralı esas alınmıştır. Kullanılan veri seti ISIC arşivine dayalıdır ve toplamda 3297 görüntü içermektedir. Etiketleme işlemi LabelImg aracıyla manuel olarak yapılmış, veriler YOLO formatında kaydedilmiş ve tüm modeller için ortak YAML dosyası hazırlanmıştır. Eğitim, RTX 3060 grafik işlemci ünitesi (Graphics Processing Unit, GPU) ve i7 işlemcili Windows 10 sistem üzerinde gerçekleştirilmiş; Python 3.12.4, PyTorch 2.3.1, CUDA 12.5 ve cuDNN 9.2 gibi güncel araçlar kullanılmıştır. Hiperparametre optimizasyonunda AdamW algoritması tercih edilmiştir.

Elde edilen deneysel sonuçlara göre, YOLOv8m serisinde, YOLOv8m genel doğruluk açısından, %81.93 ortalama kesinlik değeri (mean Average Precision- mAP) ve %77.64 F1-Skor ile en yüksek performansı göstermiştir. YOLOv8n ise 4.1 ms/görüntü ile en hızlı model olarak dikkat çekmiş, dengeli performans sunmuştur. YOLOv8x modeli ise zayıf doğruluk ve yavaşlık nedeniyle pratikte sınırlı kalmıştır. YOLOv11s modeli, %83.03 mAP ve %78,69 F1 Skoru ile genel anlamda en iyi sonuçları sunmuştur. YOLOv11m %80.59 ile en yüksek duyarlılığı göstermiştir. Hafif yapılı YOLOv11n modeli kaynak dostu uygulamalarda uygun performans sergilemiştir. Daha büyük modeller, özellikle YOLOv11x, beklenen performans artışını sağlayamamıştır. YOLOv12 serisinde, YOLOv12m en iyi genel sonuçları elde etmiş,

%81.36 mAP ve 0.7791 F1 skoruyla öne çıkmıştır. YOLOv12n %82.86 ile en yüksek duyarlılığı göstermiştir. Büyük modeller (YOLOv12l, YOLOv12x) beklentileri karşılamamıştır. RT-DETR modeli ise diğer modellere göre %64 civarında değeri ile daha düşük bir doğruluk göstermiştir.

Karmaşıklık matrisi analizine göre, YOLOv8n, YOLOv8s, YOLOv8m ve YOLOv8l modelleri iyi huylu lezyonlarda %86 doğrulukla benzer performanslar sunmuştur. Kötü huylu lezyonlarda ise, YOLOv8n %74 doğrulukla en iyi sonucu verirken, YOLOv8l %42 ile en düşük sonucu almıştır. Arka plan ile karışıklık oranı kötü huylu lezyonlarda %37–49, iyi huylularda %51–63 arası değişmiştir. Lezyonlar arası karışıklık oranı genelde %15–19 civarındayken, YOLOv8x modelinde %30’a kadar çıkmıştır. YOLOv11 modelleri daha dengeli sonuçlar sunmuştur. İyi huylu lezyonlar %82 üzeri doğrulukla sınıflandırılmış, kötü huylularda en yüksek doğruluk %74 (YOLOv11s) olarak ölçülmüştür. Arka plan karışıklıkları genel olarak daha düşük seyretmiştir.

Farklı çalışmalarda elde edilen sonuçlar, Cirkovic ve ark. (2024) çalışmasında YOLOv8 ile %78.2 doğruluk, Datta ve ark. (2023) YOLOv7-tiny ile %82.1 mAP ve Sutaji ve ark. (2023)’nin YOLOv8-large ile %79.6 mAP sonuçları gibi önceki literatürle karşılaştırıldığında rekabetçi hatta üstün performans göstermektedir.

SKIN CANCER DETECTION WITH DEEP LEARNING-BASED OBJECT DETECTION MODELS

SUMMARY

This thesis aims to evaluate the application of artificial intelligence-supported object detection approaches in the early diagnosis of skin cancer and the effectiveness of these approaches. Skin cancer stands out as a serious public health problem with the increasing number of cases worldwide and high mortality rates due to delayed diagnosis. Especially in cases of malignant melanoma, delays in diagnosis have serious effects on the life expectancy and quality of life of patients. Therefore, the need for faster, more accurate and objective diagnostic tools is increasing.

This study relied on classifying skin lesions into two main categories: benign and malignant. Benign lesions are often stable in size, shape, and color and do not pose an immediate threat to the patient's life. Common moles and mild skin pigmentation are prominent examples. Medical intervention is limited or cosmetic unless worrisome changes occur. In contrast, malignant lesions are characterized by distinctive visual characteristics such as asymmetry, irregular borders, color variation, rapid increase in size, or persistent change. This necessitates early diagnosis and prompt intervention to increase the chance of survival. Melanoma is one of the most dangerous malignant lesions.

For early clinical assessment purposes, the ABCDE rule is an important clinical tool for skin cancer detection. This rule is based on five visual characteristics of lesions: asymmetry, border irregularity, color variation, lesion diameter, and evolution over time. These criteria are currently used as the basis for many artificial intelligence systems that analyze skin images and automatically diagnose cancer.

To meet this need, systems based on artificial intelligence (AI) have been developed, revolutionizing the world of technology, particularly computer vision and deep learning techniques. These technologies have achieved remarkable success in analyzing digital skin images and detecting lesions within this framework.

YOLO (You Only Look Once) is a leading AI algorithm for object detection. It can process an entire image and perform predictions in a single step, unlike traditional convolutional networks (CNNs), which rely on two stages (object segmentation and classification). CNNs are used to extract features from images using layers of convolutions and pooling and form the backbone of many modern applications, including YOLO.

The YOLO architecture is based on three main components: the backbone, the neck, and the head. The backbone is used to extract initial features from the image and typically consists of a CNN such as CSPDarknet. The neck helps enhance the transfer of information between different feature levels using techniques such as FPN (Feature Pyramid Network) and PAN (Path Aggregation Network), which helps detect small objects. Finally, the head is responsible for locating objects using bounding boxes, determining each object's class, and assessing its confidence level.

This study focused on four popular models in this context: YOLOv8, YOLOv11, YOLOv12, and RT-DETR (Real-TimeDetection Transformer), analyzing their architecture, processing speed, and performance accuracy.

YOLOv8 was released on January 10, 2023, by the Ultralytics team. This version excels at object detection, as demonstrated in previous work, with impressive results. YOLOv11 was developed by Rahima Khanam and others and released on October 23, 2024. The newest version of YOLO to date is YOLOv12, released in 2025. It is an advanced version that incorporates improvements to the backbone and attention mechanisms. Finally, the RT-DETR algorithm, released in June 2023 by ByteDance, is the first real-world transformer model that operates without an NMS (Non-Maximum Suppression) and boasts high timescale performance.

These algorithms have seen significant architectural advancements in recent years. YOLOv8 introduced a lighter and faster architecture compared to previous versions, with improved contextual understanding through the C2f module. YOLOv11 leveraged the C3k2 and C2PSA modules to improve performance in detecting small objects. YOLOv12 featured a more advanced design, incorporating the R-ELAN architecture, separate 7x7 convolutional layers, and FlashAttention technology to achieve a better balance between speed and accuracy. RT-DETR relied on the Transformer architecture without the need for post-processing steps such as NMS, and stood out for its integration of the AIFI and CCFF modules to efficiently handle multi-scale features.

The dataset used in this study is actually source data based on the ISIC archive, obtained from the Kaggle platform. The dataset consists of 3,297 skin images. The training dataset contains 1,440 images of benign lesions and 1,197 images of malignant lesions. The testing dataset contains 360 images of benign lesions and 300 images of malignant lesions. The dataset was divided into 50% training, 30% validation, and 20% testing.

In this study, the labeling process for all images was performed using LabelImg, an open-source platform widely used for data learning. During the annotation phase, the locations of skin lesions within each image were manually identified and saved in the YOLO format (center_x, center_y, width, height) after normalization. Each lesion was also classified in .txt text files, where each line represents a single object in the image. The images were divided into two classes: class (0) for benign lesions, and class (1) for malignant lesions. These classifications were included within the annotation files to improve the learning process of the models during training, and all files underwent careful manual review to ensure data validity. A YAML file containing class names and data structures was also prepared, enabling different models to easily run on the same architecture. The captioned images were used to train four object detection algorithms: YOLOv8, YOLOv11, YOLOv12, and RT-DETR. All models were trained on the same dataset to provide accurate and unbiased comparison of their performance.

Training for all models was performed on a Windows 10 computer equipped with an Intel Core i7-12700H processor clocked at 2.30 GHz and an NVIDIA GeForce RTX 3060 GPU(Graphics Processing Unit). The development environment used recent software tools, including Python 3.12.4 and PyTorch 2.3.1, with CUDA 12.5 and cuDNN 9.2 adopted to fully utilize the capabilities of graphics processing. Each model was trained for a maximum of 300 epochs, with a training batch size of 8. The automatic parameter tuning tool available in the Ultralytics library was used, which selected the AdamW algorithm as the optimal optimization option. The initial learning

rate was set at 0.001667 and the momentum coefficient at 0.9, which contributed to enhancing training stability and improving the overall performance of the models.

As detailed results on the dataset demonstrate, the YOLOv8m model achieved the highest performance in terms of overall accuracy, scoring high across all metrics, such as mAP at 81.93% and F1-Score at 77.64%. In contrast, the YOLOv8n model emerged as the fastest model in terms of processing time, taking only 4.1 milliseconds per image, while maintaining a good balance in overall performance. YOLOv8s offered a balanced solution between performance and speed, making it a suitable choice for real-time applications. YOLOv8x, on the other hand, performed poorly compared to the other models, despite its large architecture, and had the slowest response time, limiting its effectiveness in practical environments.

Test results of the YOLOv11 series of models showed that YOLOv11s delivered the best overall performance, achieving a mAP@0.5 of 83.03% and an F1-Score of 0.7869, reflecting a strong balance in detecting various skin lesion types. The model also demonstrated clear superiority across object sizes. YOLOv11m demonstrated strong recall performance, reaching 0.8059, demonstrating its ability to detect most actual lesions without overlooking a significant number of positive cases. In contrast, YOLOv11n delivered competitive performance despite its simple architecture, making it suitable for resource-constrained applications. Larger versions such as YOLOv11x and YOLOv11l demonstrated relatively lower performance, indicating that increasing model size does not always lead to improved accuracy.

The YOLOv12m model performed best overall, achieving the highest accuracy mAP 81.36% and the highest F1-score %77.91. YOLOv12n model had the highest recall (82.86%), making it suitable for full detection even with false alarms. YOLOv12s model performed poorly compared to the other models but achieved a good balance between accuracy and speed.

While YOLOv12l and YOLOv12X performed poorly, indicating that increasing model size does not guarantee better performance.

Finally, we have the RT-DETR model, which is considered the least accurate among the models, as the accuracy rate in RT-DETR Large reached 64% of mAP.

Compared to previous studies, the results obtained in this work demonstrate remarkably competitive performance. For example, Cirkovic et al. (2024) achieved an accuracy of 78.2% using YOLOv8, while Datta et al. (2023) recorded a mAP of 82.1% using YOLOv7-tiny on ISIC data. Sutaji and Yıldız reported that YOLOv8-large achieved a mAP of 79.6% on HAM10000 data. In Nie et al. (2019)'s study, YOLOv2 achieved an accuracy of 83% and YOLOv3 an accuracy of 77% on a relatively small dataset (only 200 images). Hasya et al. (2021) and colleagues also demonstrated a consistent accuracy of 96% in non-temporal cases and 80% in real-time applications using YOLOv3 and a CNN.

In this study, we faced a significant challenge with the dataset due to the similarity between the lesions and the background, which led to competitive results in the confusion matrix. The confusion matrix results showed that the YOLOv8 models outperformed the benign lesion classification with an accuracy rate of 86%, while the performance of YOLOv8x declined in this category. In classifying malignant lesions, YOLOv8n achieved the highest accuracy at 74%, compared to 42% for YOLOv8l. High confusion rates were observed between the background and malignant lesions, ranging from 37% to 49%, while for benign lesions, it ranged from 51% to 63%.

Confusion between the two categories ranged from 15% to 19%, rising to 30% for YOLOv8x.

In contrast, YOLOv11 demonstrated improved performance, exceeding 82% in all versions for benign disease classification. YOLOv11s achieved the highest accuracy for malignant lesion classification at 74%, while the lowest was 67%. Confusion rates between background and malignant lesions ranged from 52% to 58%, and between background and benign lesions ranged from 42% to 50%, while the confusion rate between the two classes was between at 15% and 19%.



1. GİRİŞ

Cilt kanseri, dünyada en yaygın kanser türlerinden biri olup, küresel ölçekte giderek büyüyen bir halk sağlığı sorunu haline gelmiştir (Alwahaibi ve ark., 2025; American, 2023). Küresel düzeyde yapılan son istatistikler, bu hastalığın önemli bir yük oluşturduğunu ortaya koymaktadır. 2022 yılında yaklaşık 331.722 cilt kanserinin en tehlikeli türü olan malign melanom vakası ve yaklaşık 58.667 ölüm kaydedilmiştir (Zhang ve ark., 2025). Ayrıca, bazal hücreli ve skuamöz hücreli karsinom gibi malign olmayan cilt kanserleri, yılda yaklaşık 1.23 milyon vaka ile dünya çapında en sık rastlanan beşinci kanser türü konumundadır ve bu türlerin neden olduğu ölümler yılda 69.000'i aşmaktadır (Zhang ve ark., 2025). Bu veriler, cilt kanserinin halk sağlığı üzerindeki etkisini açıkça ortaya koymakta ve erken teşhisin önemini vurgulamaktadır. Özellikle erken evrede tespit edilen melanom vakalarında beş yıllık sağkalım oranı %99'un üzerine çıkabilirken (Skin Cancer Foundation, 2023), ileri evrede, kanser lenf düğümleri veya diğer organlara yayıldığında bu oran ciddi şekilde düşmektedir.

Erken teşhisin önemine rağmen, geleneksel cilt kanseri tanı yöntemleri birçok zorlukla karşı karşıyadır. Genellikle ilk tanı, çıplak gözle yapılan klinik muayene veya dermatoskop aracılığıyla yapılır, ardından kesin tanı için biyopsi alınarak laboratuvarında incelenir. Ancak biyopsi, ağrılı ve zaman alıcı bir invaziv işlem olup, hastada yara izi veya enfeksiyona yol açabilir (Smith ve ark, 2022). Ayrıca, yalnızca gözle muayene ve klinik tecrübeye dayanmak subjektif yorumlara neden olabilir, aynı deri lezyonunu değerlendiren doktorlar arasında tanı farklılıkları da oluşabilir. Bu da erken evre veya atipik görünümlü vakalarda, iyi huylu lezyonların kötü huylu olarak değerlendirilmesi anlamına gelen aşırı tanı ya da kötü huylu lezyonların gözden kaçırılması olarak ifade tanımlanan yetersiz tanı gibi sorunlara yol açabilir (Ergün ve ark., 2021), bazı bölgelerde dermatolog eksikliği veya randevu gecikmeleri erken tanıyı engelleyebilir ve tedavinin gecikmesine neden olabilir (Ergün ve ark., 2021).

Bu zorluklarla başa çıkmak amacıyla yapay zeka (Artificial Intelligence - YZ) ve bilgisayarlı görme teknikleri, cilt kanseri tanısını iyileştirmek için umut verici araçlar olarak öne çıkmaktadır. Son yıllarda, deri lezyonlarını analiz edebilen ve yüksek

doğrulukla tanı koyabilen derin öğrenme (LeCun ve ark., 2015) algoritmalarında önemli gelişmeler yaşanmıştır (Albahli, 2025). Bilgisayar destekli tanı sistemleri (Computer Aided Diganosis-CAD), görüntülerin nesnel ve hızlı analizini mümkün kılarak insan hatasını en aza indirir. Güncel çalışmalar, YZ tekniklerinin geleneksel yöntemlerle karşılaştırıldığında, cilt kanseri tespitinde doğruluk ve hassasiyet oranlarını artırabileceğini göstermektedir (Abdallah ve ark., 2024). Bu sistemler, dijital deri görüntülerini işleyerek tümörlerin belirleyici özelliklerini daha kararlı ve tutarlı şekilde ayırt edebilmektedir. Böylece bilgisayarlı görme, doktorlara daha kesin ve anlık tanı bilgileri sağlayarak erken teşhis şansını artırmakta ve klinik karar sürecini hızlandırmaktadır (Jones ve ark., 2021).

Nesne tespitinin temel amacı, farklı konum ve boyutlarda ortaya çıkabilecek bir veya birden fazla sınıfa ait nesneleri doğru şekilde tespit etmektir. Bu görev, özellikle cilt kanseri gibi karmaşık tıbbi görüntülerde, lezyon ile arka plan arasındaki yüksek görsel benzerlik nedeniyle doğru lezyon algılamanın zorlaşmasıyla birlikte daha da kritik hale gelmektedir. (Yu ve ark., 2016; Zhao ve ark., 2019). Bu alandaki en dikkat çekici YZ modelleri arasında, görüntüleri gerçek zamanlı analiz etme kapasitesiyle öne çıkan Yalnızca Bir Kere Bak (You Only Look Once-YOLO) algoritmaları yer almaktadır. YOLO algoritması ilk kez 2016 yılında Redmon ve arkadaşları tarafından tanıtılmış ve görüntünün tamamını tek seferde işleyerek gerçek zamanlı nesne tespitini mümkün kılan sinir ağı mimarisi ile bir dönüm noktası olmuştur. Son yıllarda geliştirilen YOLO ailesine ait nesne algılama modelleri, gerçek zamanlı analiz kabiliyetleri ve yüksek doğruluk oranları ile dikkat çekmektedir (Redmon ve ark, 2016). Gerçek Zamanlı Algılama Dönüştürücüsü (Real-Time Detection Transformer - (RT- DETR)), Transformer mimarisine dayanan ve Haziran 2023'te kullanıma giren, son teknoloji bir nesne algılama modelidir. Bu model, DETR'nin doğruluğunu YOLO'nun gerçek zamanlı işleme yetenekleriyle birleştirerek, nesne algılama görevleri için hız ve doğruluk açısından etkili bir çözüm haline getirir. En dikkat çekici özelliklerinden biri, Maksimum Olmayan Bastırma (NMS) gibi yaygın son işlem adımları gerektirmemesi ve sonuçları doğrudan üretebilmesidir (Carion ve ark., 2020). Söz konusu modellerin gelişmiş mimarisi, tespit performansını artırmış ve işlem hızını yüksek doğruluk seviyesini koruyarak iyileştirmiştir (Zhang ve ark, 2023). Bu iyileştirmeler, YOLO algoritmalarını tıbbi görüntülerde cilt lezyonlarını hızlı ve hassas bir şekilde tespit edebilen güçlü araçlar hâline getirmiştir. Örneğin, YOLOv8'e dayanan modern bir

model, melanom lezyonlarını %98.6 ortalama kesinlik değeri (mean Average Precision-mAP) oranı ile tespit etmiş ve görüntü başına yalnızca 12.5 milisaniyede işlem gerçekleştirmiştir (Albahli, 2025) Bu doğruluk ve hız kombinasyonu, modeli klinik uygulamalarda, hem uzman dermatoloji merkezlerinde hem de mobil sağlık uygulamaları aracılığıyla tarama ve ön tanı desteği sağlayabilecek biçimde uygun hale getirmektedir.

Bu çalışmada, cilt kanserinin erken teşhisine yardımcı olmak için gerçek dünya klinik ortamlarında cilt lezyonlarının tespitinin hızlandırılması ve doğruluğunun artırılması amaçlanmıştır. Bu doğrultuda, YOLOv8 (Ultralytics, 2023), YOLOv11, YOLOv12 ve RT-DETR modelleri kullanılarak derin öğrenmeye dayalı farklı nesne tespiti yaklaşımları değerlendirilmiştir. Modellerin farklı sürümleri, cilt kanserinin tespitinde literatürde yaygın olarak kullanılan Uluslararası Cilt Görüntüleme İşbirliği (International Student Identity Card – ISIC) veriseti üzerinde geliştirilmiştir. Gerçekleştirilen deneylerde, modeller doğruluk, duyarlılık (recall) ve doğruluk derecesi (precision) gibi metrikler açısından karşılaştırılmış ve en uygun model belirlenmeye çalışılmıştır (Wang ve ark, 2021). Elde edilen bulgular, YZ destekli tıbbi görüntü analizi uygulamalarının geleceği için önemli çıkarımlar sunmaktadır (Patel, 2023). Gelişmiş bilgisayarlı görme algoritmalarının, klinik uygulamalarla bütünleştirilerek, hastaların sağlık sonuçlarını iyileştirmeye katkı sağlayacağı öngörülmektedir.

Tez çalışmasının sonraki bölümleri şu şekilde düzenlenmiştir: İkinci bölümde, deri kanseri ve deri kanserini tespit etmek için uygulanan genel yöntem ele alınmakta, bu araştırma kapsamında literatürde benzer çalışmalar incelenmektedir. Üçüncü bölümde, nesne tespiti araştırma alanı ve tez çalışmasında kullanılan YOLO algoritmaları ve RT-DETR algoritması hakkında detaylı bilgi verilmektedir. Dördüncü bölümde, çalışma kapsamında uygulanan metodoloji kapsamlı bir şekilde sunulmaktadır. Beşinci bölümde, incelenen derin öğrenmeye dayalı nesne tespiti modellerinin performanslarının karşılaştırmalı bir analizi yürütülmektedir. Son olarak altıncı bölümde, sonuçlar değerlendirilerek tartışılmaktadır.



2. LİTERATÜR

2.1. Deri Kanseri

Cilt kanseri, dünya genelinde en sık görülen kanser türlerinden biridir. Genellikle güneş ışığı veya solaryum gibi yapay kaynaklardan gelen ultraviyole radyasyona aşırı maruz kalmanın sonucu olarak ortaya çıkan, cilt hücrelerinin anormal büyümesiyle karakterize bir hastalıktır. Bu kanser türünün şiddeti ve etkisi, türüne ve tespit aşamasına bağlı olarak değiştiğinden, değişen şiddette olduğu düşünülmektedir. (Raimondi ve ark., 2020; Kliniec ve ark., 2023; Narayanan ve ark., 2010)

İstatistikler, son yıllarda cilt kanseri oranlarının önemli ölçüde arttığını göstermektedir. Bu artış, güneşe maruz kalma alışkanlıklarındaki değişiklikler ve ultraviyole radyasyon üreten yapay cihazların kullanımının artmasıyla tetiklenmektedir. Yapılan araştırmalar, açık tenli kişilerin, ailesinde bu hastalığın öyküsü bulunanların veya bağışıklık sistemi zayıf olanların enfeksiyona daha yatkın olduğunu göstermiştir. Ayrıca, birden fazla veya anormal benin varlığı, bu lezyonlardan bazılarının malign melanom'a dönüşme olasılığını artırmaktadır (Garbe ve ark., 2016).

Cilt kanserinin tehlikesi tanıda gecikmeden kaynaklanmaktadır. Cilt lezyonları erken evrelerinde iyi huylu benlere benzeyebilmekte ve bu durum hastalar ve hatta bazı sağlık uygulayıcıları tarafından göz ardı edilmelerine veya önemsenmemelerine yol açabilmektedir. Ancak yapılan çalışmalar erken teşhisin, özellikle müdahaleye metastatik evreden önce başlanırsa, iyileşme oranlarını %95'in üzerine çıkardığını göstermiştir (Tschandl ve ark., 2020). Bu nedenle cilt kanserinin erken teşhisi için etkili araç ve tekniklerin geliştirilmesi çağdaş tıp alanında temel bir konudur.

Cilt kanserleri birkaç ana türe ayrılır; bunların en önemlileri şunlardır: Bazal Hücreli Karsinom (Basal Cell Carcinoma), Skuamöz Hücreli Karsinom (Squamous Cell Carcinoma) ve Melanom (Melanoma). Bu türler arasında melanom, erken teşhis edilip tedavi edilmediği takdirde vücudun diğer organlarına yayılma potansiyeli yüksek olduğu için en tehlikelisi olarak kabul edilir (American Cancer Society, 2023).

Cilt kanserine bağılı cilt lezyonları iyi huylu lezyonlar ve kötü huylu lezyonlar olmak üzere iki ana türe ayrılmaktadır. İki tür arasındaki temel farklılık bu hücrelerin biyolojik ve histolojik davranışlarından kaynaklanmaktadır. (Firnhaber, 2020; Helfand ve ark., 2001)

2.1.1. İyi huylu lezyonlar (benign)

Şekil 2.1 'de iyi huylu lezyonlar gösterilmiştir. Hücrelerin çevre dokulara yayılmadan veya vücudun diğer bölgelerine yayılmadan orijinal yerlerinde kaldığı sınırlı anormal büyüme ile karakterizedir. Bu lezyonlar genellikle boyut, şekil ve renk olarak sabittir ve hastanın yaşamı için acil bir tehdit oluşturmazlar. İyi huylu lezyonlara örnek olarak yaygın benler (nevüsler) ve basit cilt pigmentasyonları verilebilir ve tedavi genellikle yalnızca kozmetiktir veya endişe verici değişiklikler ortaya çıkmadığı sürece tıbbi müdahale gerektirmez (Marghoob ve ark., 2019).

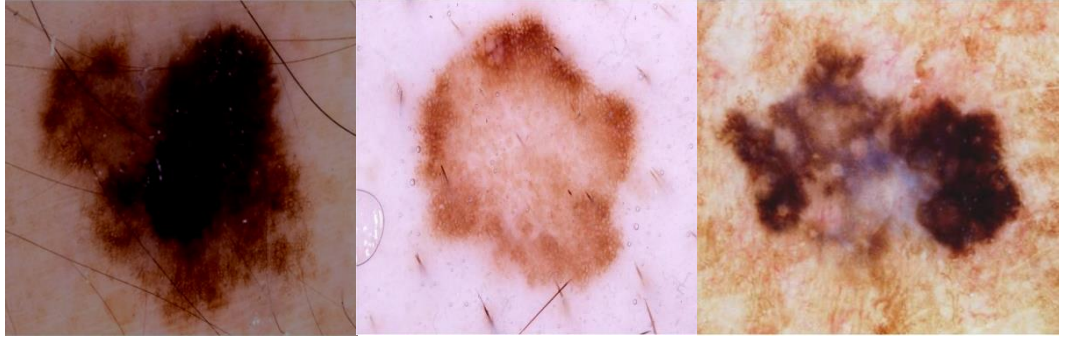


Şekil 2.1. İyi huylu deri kanseri.

2.1.2. Kötü huylu lezyonlar (malignant)

Kötü huylu lezyonlar agresif davranışlarıyla bilinirler. Kanser hücreleri, hızla bölünme, çevre dokulara yayılma ve kan damarları veya lenf damarları yoluyla vücuttaki diğer organlara yayılma yetenekleriyle karakterize edilirler. Kötü huylu cilt kanserinin en belirgin örneklerinden biri olan melanom, yüksek yayılma hızı ve ileri evrelerinde kontrol altına alınmasının zor olması nedeniyle en tehlikeli cilt kanseri türü olarak kabul edilir (Garbe ve Leiter, 2016). Kötü huylu lezyonlar Şekil 2.2 'de belirttiği gibi asimetri, düzensiz kenarlar, renk çeşitliliği, hızlı boyut artışı veya sürekli değişim gibi belirgin görsel özelliklerle karakterize olduğundan, erken teşhis ve hızlı müdahale

hayatta kalma şansını artırmak için çok önemlidir (Friedman ve ark., 1985; Firnhaber, 2020; Mishraa & Celebi, 2016; Lucas ve ark., 2016).



Şekil 2.2. Kötü huylu deri kanseri.

2.2. Deri Kanserini Tespit Etmek için Uygulanan Genel Yöntem

2.2.1. ABCDE

ABCDE kuralı, melanom ve diğer cilt kanserlerinin erken teşhisini desteklemek amacıyla geliştirilmiş önemli bir klinik araçtır. Bu kural, benlerin ve cilt lezyonlarının beş görsel özelliğinin değerlendirilmesine dayanmaktadır: Asimetri (Asymmetry), Sınır Düzensizliği (Border Irregularity), Renk Çeşitliliği (Color Variation), Çap (Diameter) ve Zaman İçinde Değişim (Evolving) (Senan ve Jadhav, 2021). 2004'te, melanomun dinamik doğasını vurgulamak için E harfi (evrimleşme) eklenmiştir. ABCDE mnemoniği, son otuz yıldır halk ve sağlık topluluğu tarafından yaygın olarak kullanılmaktadır (Abbasi ve ark., 2004).

Değerlendirmeye lezyonun iki yarısının yaklaşık olarak aynı olduğu asimetri kriteri ile başlanır. İyi huylu vakalarda benler genellikle şekil olarak nispeten simetriktir, buna karşın kanserli lezyonlar simetride belirgin düzensizlik gösterir ve kontrolsüz kötü huylu büyümeyi yansıtır. Sınır düzensizliği kriteri lezyonun dış kenarlarının şekline işaret eder. İyi huylu benler genellikle düzgün, düzenli sınırlarla çevrilidir, melanomlar ise engebeli, düzensiz sınırlara sahiptir ve bu da kanser hücrelerinin çevre dokuya nüfuz ettiğini gösterir (Lattoofi ve ark., 2020).

Renk çeşitliliği açısından, normal benler genellikle tek tip bir renge (kahverengi veya açık siyah) sahipken, melanom, kanser hücrelerini karakterize eden hücresel heterojenliği yansıtan siyah, kahverengi, kırmızı, beyaz ve hatta mavi renklerin değişen tonlarını içerebilir. Ayrıca çap da önemli bir göstergedir. Yapılan çalışmalar, çapı 6 milimetreden büyük olan benlerin dikkatli incelenmesi gerektiğini gösterse de,

çapı daha küçük olan melanom vakalarına da rastlanmaktadır. Son olarak, değerlendirmede evrim kriteri önemli bir faktördür, çünkü lezyonun boyutunda, biçiminde, renginde, dokusunda veya davranışında (kanama veya kaşıntı gibi) yakın zamanda meydana gelen değişiklikleri içerir. Stabilitate iyi huylu benlerin bir özelliği iken, değişim kötü huylu olma ihtimalini gösteren endişe verici bir işarettir (Thanh ve ark., 2019; Moura ve ark., 2018).

Son araştırmalar, ABCDE ilkesinin yalnızca doğrudan klinik muayeneyle sınırlı olmadığını, cilt görüntülerini analiz etmek ve kanseri otomatik olarak teşhis etmek için kullanılan birçok modern yapay zeka sisteminin temeli haline geldiğini gösteriyor. Evrimsel sinir ağı (Convolutional Neural Network - CNN) gibi derin öğrenme algoritmaları kullanılarak simetri, sınırlar, renkler, çap ve zamansal değişimlerle ilgili özellikler çıkarılabilir ve bu da özellikle ISIC ve HAM10000 gibi büyük veri kümelerinin kullanılabilirliği ile bazı durumlarda geleneksel tanılamadan daha iyi performans gösterebilen modellerin oluşturulmasına yol açabilir (Kestek ve ark., 2023; Salomi ve ark., 2024).

ABCDE kuralı, uzman dermatologlar tarafından yararlanılan seviyelere yakın seviyelerden otomatik algılama sistemlerinde artan duyarlılık ve doğruluk oranları ile karakterizedir. Bazı çalışmalar, ABCDE özelliklerinin gelişmiş sınıflandırma algoritmaları ile birleştirilmesinin, iyi huylu ve kötü huylu lezyonları ayırt etmede %95'ten fazla bir doğruluk oranıyla sonuçlandığını göstermiştir. Bu teknolojik gelişme sadece erken teşhisi destekleyip iyileşme oranlarını arttırmakla kalmaz, aynı zamanda doktorların yanlış tespit edildiği ve gelişmekte olan bölgelerde erken teşhis araçlarının mevcudiyetini artırır (Salomi ve ark., 2024; Lattoofi ve ark., 2020).

Bu nedenle ABCDE kuralı hem geleneksel klinik düzeyde hem de yapay zekanın (YZ) modern uygulamalarıyla cilt lezyonlarının değerlendirilmesinde temel bir dayanak olmaya devam ediyor ve erken ve sistematik kullanımının erken teşhis ve hayat kurtarma şansını artırmada en önemli faktörlerden biri olduğu belirtiliyor (Senan ve ark., 2021).

2.3. Deri Kanserinin Sınıflandırılması ile İlgili Yapılan Çalışmalar

Bu bölümde bilgisayarlı görme yöntemlerine dayalı deri kanserinin sınıflandırılması konusu ile ilgili literatürde yürütülen çalışmalar incelenmiştir.

Nie ve ark. (2019), YOLOv2 ve YOLOv3 modellerini Bölgesel Tabanlı Evrimsel Sinir Ağı (Region Based Convolutional Neural Network - R-CNN) ve Tek Atış Dedektörü (Single Shot Detecto - SSD) ile karşılaştırmışlardır. YOLOv2 modeli %83 doğruluk elde ederken, YOLOv3 modeli %77 doğruluk elde etmiştir. Çalışmada, mobil uygulamalarda kullanılabilecek bir sistem geliştirilmesi amaçlanmış ve ISIC veri setinden 200 görüntü kullanılmıştır.

Sutaji ve Yıldız. (2023), YOLOv8'in nano, medium ve large versiyonlarının performansını analiz etmişlerdir. Çalışmada, cilt kanseri tespitinde doğruluğu optimize etmek amaçlanmış ve HAM10000 veri seti kullanılmıştır. Modelin large versiyonunda mAP değeri %79,6 olarak ölçülmüştür.

Aishwarya ve ark. (2023), YOLOv3-tiny ve YOLOv8s modellerini test ederek bir füzyon stratejisi uygulamışlardır. Çalışmada ISIC 2020 veri seti kullanılmıştır. YOLOv8s doğruluğu %93,4 iken füzyon sonucu %94,3 doğruluk elde etmiştir.

Tchema ve ark. (2024), YOLOv7 modelini kullanarak cilt lezyonlarını tespit etmiş ve sınıflandırmışlardır. Modelin geliştirilmesinde MED-NODE veri seti kullanılmıştır. Ortalama doğruluk %89,65, duyarlılık %85 ve kesinlik %91,90 olarak ölçülmüştür.

Datta ve ark. (2023), YOLOv7-tiny modelini düşük maliyetli cihazlarda test etmişlerdir. Çalışmada ISIC 2017 veri setinden 2750 görüntü kullanılmıştır. Deneyler sonucunda mAP değeri %82,1 olarak ölçülmüştür.

Rusydiyah ve ark. (2023), Android tabanlı bir cilt kanseri tespit uygulamasında YOLOv5'in farklı versiyonlarını test etmiştir. Çalışma ISIC 2019 veri setini kullanarak sekiz cilt hastalığı kategorisi üzerinde yürütülmüştür. YOLOv5l modeli %84,4 mAP doğruluğu elde ederken, YOLOv5x modeli %89,0 doğruluk elde etmiştir.

Benzer şekilde, Albahli ve ark. (2022) YOLOv4-DarkNet tabanlı bir model geliştirerek cilt kanseri lezyonlarını doğru bir şekilde sınıflandırmayı ve segmentlere ayırmayı amaçlamıştır. Çalışmada kullanılan CSPDarkNet-53 modeli ile Jaccard indeksi %98 ve duyarlılık %93,2 olarak bildirilmiştir. Çalışmada modelin etkinliği ISIC 2016 ve ISIC 2018 veri kümeleri kullanılarak test edilmiştir.

Hum ve ark. (2022) AlexNet tabanlı bir derin öğrenme modeli geliştirmiş ve bunu cilt kanseri teşhisinde kullanmışlardır. Özellik çıkarma CNN ile yapılmış ve sınıflandırma ECOC SVM ile yapılmıştır. Model skuamöz hücreli karsinom için %95,1 doğruluk ve

aktinik keratoz için %98,9 duyarlılık elde etmiştir. Eğitim için internetten toplanan RGB cilt lezyonu görüntüleri kullanılmıştır.

Malo ve ark. (2022), Keras ve TensorFlow kullanarak bir sistem geliştirmişlerdir. VGG-16 modeli transfer öğrenme yöntemi ile optimize edilmiş ve %87,6 doğruluk oranı elde edilmiştir. ISIC veri seti kullanılmıştır.

Tabrizchi ve ark. (2023), VGG-16 mimarisini iyileştirerek %93 AUC doğruluğu elde etmişlerdir. Model, ResNet50 ve AlexNet gibi modellerden daha iyi performans göstermiştir. Çalışmada 33.126 görüntü kullanılmıştır.

Tuncer ve ark. (2024), daha az eğitim parametresiyle TurkerNet modelini geliştirmişlerdir. Model, %92,12 test doğruluğu elde etmiştir. Kaggle veri seti kullanılmıştır.

Aishwarya ve ark. (2023), YOLOv3 ve YOLOv4 kullanarak cilt kanseri sınıflandırması gerçekleştirmiştir. YOLOv3, %88,03'lük bir mAP elde ederken, YOLOv4, %86,52'lik bir mAP elde etmiştir. Çalışmada ISIC veri seti kullanılmıştır.

Hasya ve ark. (2021), gerçek zamanlı cilt kanseri tespiti için YOLOv3 ve CNN kullanmışlardır. Statik doğruluk %96 ve gerçek zamanlı doğruluk %80 olarak ölçülmüştür. Çalışmada HAM10000 veri seti kullanılmıştır.

Gouda ve ark. (2022), ESRGAN ile görüntü kalitesini iyileştirerek cilt kanserini tespit etmiştir. InceptionV3 modeli %85,7 doğruluk oranıyla en iyi sonucu vermiştir. Modelin eğitimi için ISIC 2018 veri seti kullanılmıştır.

AlSadhan ve ark. (2024), YOLOv3, YOLOv4, YOLOv5 ve YOLOv7 modellerini karşılaştırmışlardır. YOLOv7, %75.4 mAP ve %86.3 IoU ile en iyi performansı sergilemiştir. Çalışmada ISIC 2017 veri seti kullanılmıştır.

Nersisson ve ark. (2021), YOLOv2'yi geleneksel özelliklerle birleştirerek %94 doğruluk elde etmişlerdir. Çalışmada ISBI Melanoma veri seti kullanılmıştır.

Demir (2021), MobileNetV2 ve SVM tabanlı bir yöntemle dermatoskopik görüntülerden kötü huylu cilt kanserini tespit etmek amaçlanmıştır. 3.297 görüntüden oluşan veri setinde, modelin doğruluk oranı %88,35 olup, özellik seçimi ile %2,31 oranında artırılmıştır. Yöntemin, düşük hesaplama maliyeti ile klinik kullanıma uygun olduğu bildirilmiştir. Gelecekte, modelin geliştirilmesi ve mobil platformlara uyarlanması amaçlanmaktadır.

Riyadi ve ark. (2024), YOLOv8n algoritmasını kullanarak cilt kanseri tespiti ve sınıflandırması üzerine bir çalışma yürütmüştür. Model, ISIC 2019 veri setinden 4289 görüntü kullanılarak eğitilmiştir. YOLOv8n, özellikle bazal hücreli karsinom ve vasküler lezyon kategorilerinde %93,5 doğruluk, %93,7 duyarlılık ve %93,5 F1 skoru ile %100 doğruluk elde etmiştir. Çalışma, hızlı ve hassas cilt kanseri tespiti için etkili bir yöntem sağlamayı amaçlamıştır.

Ćirković ve ark. (2024), YOLOv8 algoritmasını kullanarak 3000 görüntüden oluşan bir veri setini incelemiş ve %78,2'lik genel bir doğruluk oranına ulaşmışlardır. Çalışmanın amacı dermatologlara ön bir tarama aracı sağlamaktır.

2.4. Farklı Algoritmaların Sınıflandırılması ile ilgili Yapılan Çalışmalar

Bu bölümde, çalışmada kullanılan algoritmalarla ilgili önceki çalışmalar incelenerek, bunların çeşitli alanlardaki uygulamaları vurgulanmaktadır.

Zhao ve ark., (2024) RT-DETR algoritmasına dayanan geliştirilmiş bir tespit modeli geliştirmişler ve buna RT-DETR-Tomato adını vermişlerdir. Çalışmanın amacı, özellikle dengesiz aydınlatma, üst üste binen yapraklar ve küçük hedef boyutu gibi zorluklar ışığında, karmaşık tarımsal ortamlarda domates tespitinin doğruluğunu ve etkinliğini artırmaktır. Yaklaşım, orijinal RT-DETR modelinin mimarisini, ResNet-50 yerine temel ağı bir Swin Transformer entegre ederek ve kodlayıcıya bir BiFormer Attention modülü ekleyerek değiştirmeyi içermektedir. Deneyler, Shandong Tarım Üniversitesi'ne bağlı bir çiftlikten toplanan ve LabelImg aracı kullanılarak etiketlenmiş 425 adet yüksek çözünürlüklü görüntü içeren gerçek dünya veri seti üzerinde gerçekleştirilmiştir. Veriler %80 eğitim, %20 test için ayrılmıştır. Deneysel sonuçlar, hibrit RT-DETR-Tomato-BS modelinin %88,7'lik en yüksek tespit doğruluğuna (mAP@0,5) ve %84,1'lik F1 puanına ulaşarak daha iyi performans sergilediğini gösterdi.

Zhang ve ark (2024) tarafından gerçekleştirilen çalışmada, RT-DETR algoritmasının doğrudan geliştirilmiş bir versiyonu olan Drone-DETR adlı yeni bir model, drone görüntülerindeki küçük nesneleri tespit etmek için geliştirilmiştir. Model, çok ölçekli dikkat mekanizmalarına dayanan EDF-FAM modülünü kullanarak özellikleri entegre etmek ve MPD ve SFEM gibi yüksek verimli modülleri kullanarak özellikleri çıkarmak için ESDNet'i kullanıyor. VisDrone2019 veri setindeki deneysel sonuçlar, önerilen modelin önceki sürümlerden önemli ölçüde daha iyi performans gösterdiğini,

daha az parametreyle %53,9'luk bir mAP50 değerine ulaştığını ve doğru ve hızlı algılama gerektiren gerçek dünya ortamlarında yüksek verimlilik sağladığını göstermektedir.

Wang ve ark. (2024), doğal ortamlarda domates olgunlaşmasının tespiti için PDSI-RTDETR adı verilen geliştirilmiş bir model önermektedir. Model, RT-DETR altyapısına dayanmaktadır; ancak geleneksel bir ağ yerine hafif bir PConv_Block modülünün kullanımı, Ölçek İçi Özellik Etkileşimi (AIFI-DAT) modülünün ve çarpık dikkat mekanizmasının tanıtımı gibi önemli mimari iyileştirmeler içermektedir. Boyun aşaması da, geleneksel kayıp fonksiyonunun İç-EIoU fonksiyonuyla değiştirilmesinin yanı sıra, ince boyunlu SSFF tasarımı kullanılarak geliştirilmiştir. Geliştirilmiş model, mAP50 doğruluğu (%3,9 artış), mAP50:95 doğruluğu (%4,1 artış) ve işlem hızında (FPS) %38,7 artış bakımından temel RT-DETR modeline kıyasla daha yüksek performans elde etmiştir.

Hedaya ve ark. (2024), süs balığı hastalıklarının otomatik tespitinin zorluklarını ele almak için, YOLOv9 algoritmasını RT-DETR mimarisiyle birleştirmeyi önermiştir. Çalışma, karmaşık bir su ortamında görsel hastalık göstergelerini tespit edebilen bir sistem geliştirmeyi amaçlamıştır. Sonuçlar, birleşik modelin %97,53'lük üzerinde bir sınıflandırma doğruluğuna ulaştığını ve gerçek zamanlı görevlerde yüksek verimliliğini kanıtladığını göstermektedir.

Wang ve ark. (2024), hiyerarşik uzamsal kümelemeyi simüle eden ve seviyeler ve özellikler arasında etkileşim yeteneğini artıran doğrusal bir hiyerarşik uzamsal dikkat (LSPA) mekanizmasını içeren RT-DETR algoritmasına mimari bir gelişme. Araştırmacılar ayrıca mekansal temsili geliştirmek ve paraziti azaltmak için bir mekansal koordinasyon filtresi (SCF) geliştirmişlerdir. Hesaplama yükünü azaltmak ve performans doğruluğunu korumak için kısmi evrişim yöntemi (PConv) kullanılmıştır. Deneyler BDD100K ve KITTI grupları üzerinde gerçekleştirilmiş ve sonuçlar, önerilen modelin doğruluk açısından etkinliğini (mAP@50) ve parametre sayısı açısından YOLO ve DETR versiyonlarından daha iyi performans gösterdiğini, karmaşık trafik ortamlarında ve gerçek dünya zorluklarında etkinliğini vurgulamıştır.

Hoque ve ark. (2024) beyin tümörlerinin doğru bağlamsal tespitine adanmış RPC-DETR olarak bilinen geliştirilmiş bir model önermiştir. Bu model, RT-DETR'nin temel mimarisine dayanılmıştır ancak bağlamsal dikkat modülü, ince taneli özellik

işlemcisi ve eğitim sırasında denetim kaybını azaltan bir mekanizma gibi gelişmiş bileşenleri içermiştir. Bu model, BraTS2020 veri setinde, bu model MRI görüntülerinden beyin tümörlerini tespit etmek için kullanıldı ve sonuçlar RPC-DETR'nin konvansiyonel YOLOv5, Hızlı bölge tabanlı evrişimli sinir ağı (Faster R-CNN) ve DETR modellerinden ortalama doğrulukla yaklaşık %91,12'lik daha iyi performans gösterdiğini gösterilmiştir. Bu çalışma, özellikle yüksek bağlamsal anlayış ve gelişmiş konumsal doğruluk gerektiğinde, karmaşık tıbbi görüntüleme zorluklarının ele alınmasında iyileştirilmiş dönüştürücü modellerin yüksek potansiyelini yansıtmaktadır.





3. YÖNTEM

3.1. Nesne Tespiti

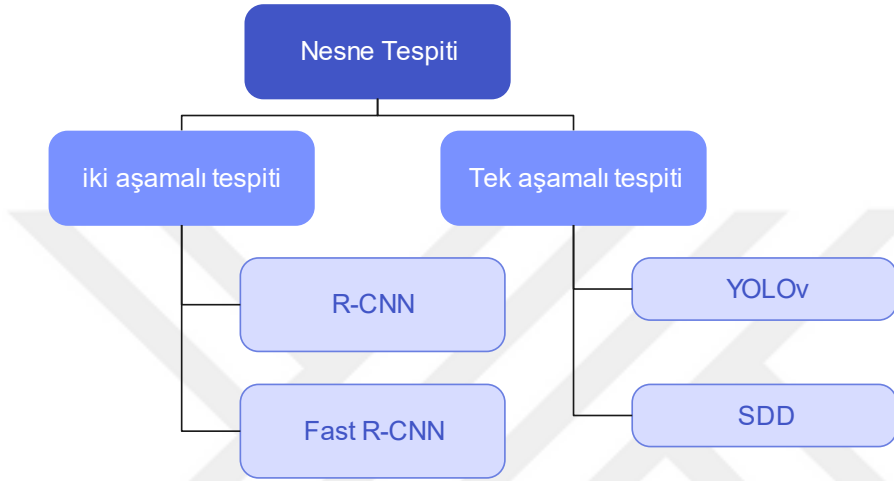
Nesne tespiti, bir görüntüdeki nesneleri aynı anda bulup (yerleştirmeyi) ve sınıflandırmayı (tür) amaçlayan bilgisayarlı görme alanındaki en büyük zorluklardan biridir. Bu teknik, görüntüdeki her görünür nesnenin etrafına çevreleyen kareler çizmeye ve ardından her kareye bir etiket veya kategori atamaya dayanır. Sadece “Görüntüde ne var?” sorusunu cevaplayan sınıflandırma algoritmalarının aksine, nesne tespiti “Nerede?” sorusuna da cevap verir.

Tarihsel olarak, bu alandaki araştırmaların çoğu yalnızca bir sınıfı (genellikle yüzleri veya yayaları) veya birkaç sınıfı tespit etmeye odaklanmıştır (Liu ve ark., 2020). Ancak son birkaç yıldır araştırma topluluğu, insan performansına yakın bir şekilde çok çeşitli nesneleri tanıyabilen kapsamlı algılama sistemleri inşa etme gibi daha zorlu bir hedefe doğru ilerlemeye başlamıştır (Krizhevsky ve ark., 2012).

Geleneksel nesne tespiti yöntemleri üç temel adımdan oluşur: öncelikle görüntüde nesne içermesi muhtemel bölgeler belirlenir (bölge önerileri), sonra bu bölgelerden özellikler çıkarılır (özellik çıkarma) ve son olarak bu bilgiler kullanılarak nesneler sınıflandırılır ve konumlandırılır. Bu yapı, R-CNN gibi iki aşamalı modellerde yaygındır ve yüksek doğruluk sağlar. Ancak uzun işlem süresi nedeniyle gerçek zamanlı uygulamalarda etkili olmamaktadır (Carranza-García ve ark., 2021; Zhang ve ark., 2021).

Derin öğrenme tekniklerinin gelişmesiyle birlikte, evrişimli sinir ağlarına dayalı modeller, nesne tespiti alanında önceki modellere kıyasla önemli ölçüde üstünlük sağlamıştır. Bu modeller, görüntüyü parçalara ayırmak ve işlenebilen sayılara ve dijital birimlere dönüştürmek için çok katmandan oluşan derin bir mimariye dayanır. Bu modeller renk ve farklı ışıklandırma açısından büyük miktarda veri kullanılarak eğitilir (Önder ve Karan, 2024) . Tanıma ve sınıflandırmada güçlü bir performans elde edilmesi hedeflenir. Bu modelleri geliştirmek ve tespit süreçlerini daha da iyileştirmek için birçok tespit modeli önerilmiştir; bunların en önemlileri Şekil 3.1’de gösterildiği gibi iki ana türe ayrılır (Karbouj ve ark., 2024): Tek aşamalı (Single-stage) ve iki

aşamalı (Two-stage) dedektörler. Tek aşamalı dedektör, doğrudan görüntü üzerinde tahmin yapar. Görüntüyü belirli ızgara adı verilen bölgelere böler ve her hücre üzerinde nesne sınıflarını ve sınırlayıcı kutuları (bounding box) tahmin eder. İki aşamalı dedektör ise, görüntü içinde nesne olabilecek aday bölgeleri belirlemek için Aday Bölge Öneri Ağını (Region Proposal Network - RPN) kullanır ve bu bölgelerin sınıflandırılması ile konumlarının hassas biçimde belirlenmesine dayanır (Carranza-García ve ark., 2021).



Şekil 3.1. Tek aşamalı (Single-stage) ve iki aşamalı (Two-stage) dedektörler.

2014 yılında Ross Gershik ve Microsoft araştırma ekibi, iki aşamalı nesne tespitinde derin öğrenmeyi uygulamak için ilk başarılı yaklaşım olan R-CNN'yi tanıtmıştır (Girshick ve ark., 2014). Model, görüntülerdeki nesne konumlarını ve sınıflarını tespit etmek için derin sinir ağlarını bölge öneri algoritmalarıyla etkili bir şekilde birleştirmiş ve bu da yüksek doğrulukla sonuçlanmıştır. Sonraki yıllarda Gershik, R-CNN'nin hesaplama etkisini azaltmak için Rapid R-CNN'yi geliştirmiştir (Girshick ve ark., 2014). Hızlı R-CNN ağının sunduğu iyileştirmelere ek olarak, Gershik ve ekibi, alan önerisi aşamasını entegre eden ve böylece tek bir ağa dönüşen Alan Teklif Ağı'nı (RPN) tanıtmıştır (Ren ve ark., 2016). RPN, farklı boyut ve oranlardaki kutuları otomatik olarak önererek Faster R-CNN'nin hızını ve doğruluğunu büyük ölçüde geliştirmiştir.

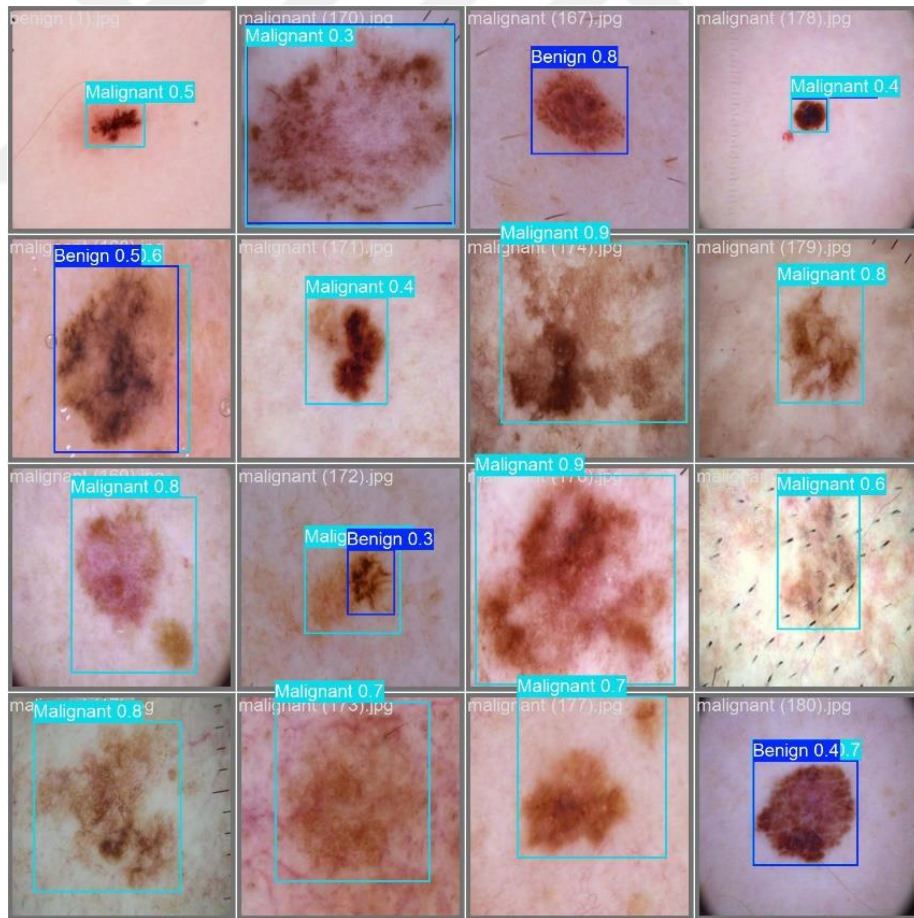
Tek aşamalı (one-stage) algılayıcılar ise, nesne tespitini ayrı bir bölge öneri aşamasına ihtiyaç duymadan tek adımda gerçekleştirmeyi amaçlar. Bu yaklaşımda SSD (Liu ve ark., 2016) ve YOLO (Redmon ve ark. 2016) gibi modeller öne çıkmaktadır. Bu çalışmada temel algoritma olarak YOLO tercih edilmiştir; çünkü ilk sürümü olan YOLOv1, nesne tespiti alanında özellikle hız açısından iki aşamalı modellere kıyasla

önemli bir ilerleme sağlamıştır. Ancak bu sürüm, doğruluk açısından bazı sınırlamalarla karşı karşıya kalmıştır (Redmon ve ark. 2016).

Aynı yıl tanıtılan bir diğer model olan SSD (Liu ve ark., 2016), işlem hızını korurken performans açısından bazı iyileştirmeler sunmuş, ancak küçük nesnelerin tespitinde beklenen başarıyı gösterememiştir.

2017 ile 2025 yılları arasında YOLO (Redmon ve ark. 2016) ailesi (Borisyuk ve ark., 2018; Bochkovskiy ve ark., 2020) mimari açıdan önemli gelişmeler kaydetmiş ve bu gelişmeler sayesinde tespit doğruluğunda ciddi artışlar elde edilmiştir (Wang ve Liao, 2024; Wang ve ark., 2024). Özellikle ağ yapısı, başlık katmanları ve eğitim stratejilerinde yapılan iyileştirmeler bu başarıda etkili olmuştur.

Şekil 3.2 'de, cilt kanseri hastalıklarının tespiti sürecinde YOLO algoritmasının kullanım şekli gösterilmekte ve bu algoritmanın tıbbi görüntüleme alanındaki potansiyel uygulamalarından biri örneklenmektedir.



Şekil 3.2. Deri kanser tespiti.

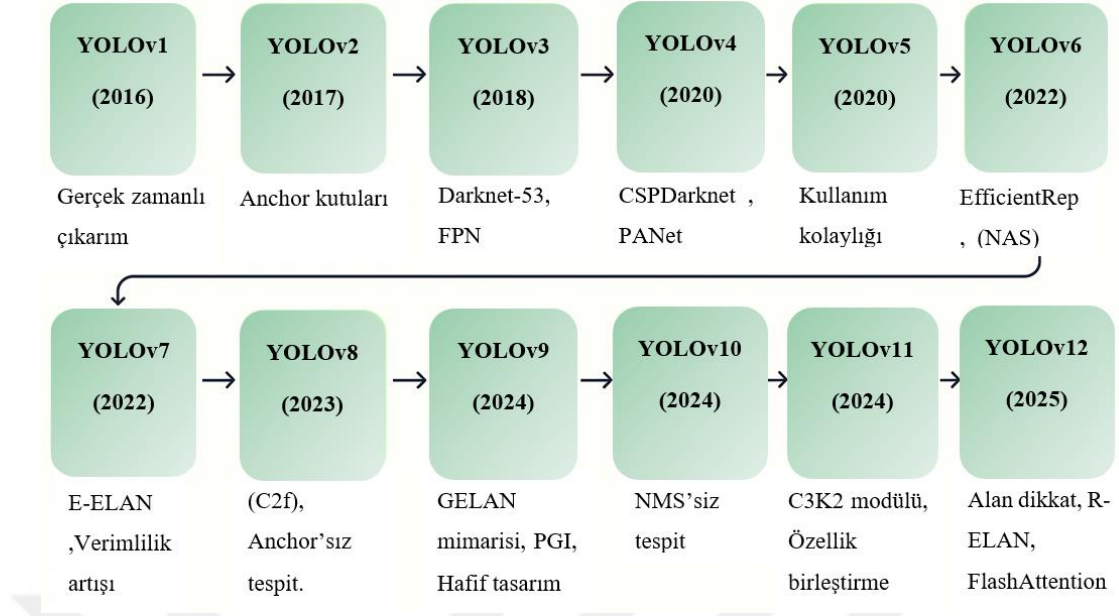
Hem kötü huylu hem de iyi huylu hastalıklar, etraflarında boş kutucuklar bulunmasıyla karakterize edilir. Bu kutular, nesne tespit sisteminin bulunduğu düşünülen yerin koordinatlarıdır. Bu kutulara bilgisayarlı görme alanında tahmin kutuları veya seçim kutuları denir. Bu kutucuklarda yüzdeler ve içerdikleri nesnenin adı yer alır. Bu yüzdeler, modelin tahminin doğruluğuna olan güvenini göstermektedir. Yukarıdaki görseldeki kareler, mavi karenin iyi huylu tümörü, açık mavi karenin ise kötü huylu tümörü gösterdiğini göstermektedir.

Tıbbi alanında nesne tespiti, cilt kanseri gibi hastalıkların erken teşhisinde önemli rol oynamakta ve hastaların yaşamını kurtarabilecek daha hızlı ve doğru kararların alınmasına yardımcı olmaktadır (Daghrir ve ark., 2020).

3.2. YOLO Algoritması

YOLO, "You Only Look Once" ifadesinin kısaltmasıdır ve ilk olarak 2016 yılında Joseph Redmon tarafından tanıtılan gerçek zamanlı bir nesne algılama algoritmasıdır (Redmon ve ark. 2016). YOLO, görüntüyü tek bir evrişimli sinir ağı içinde gerçekleştiren bir yapı sunar. Bu, YOLO'nun en ayırt edici özelliğidir ve onu bölgesel önerilere (region proposal) dayanan R-CNN tabanlı gibi yöntemlerden farklı kılar (Albahar, 2019; Redmon ve ark. 2016). Görüntüyü alt bölgelere ayıran geleneksel yöntemlerin aksine, YOLO görüntüyü bütünsel olarak ele alır ve algılama sürecini basitleştirir. Bu yöntem, nesnelerin konumunu belirlemek ve sınıflandırmak için sınırlayıcı kutular kullanır (Riyadi ve ark. 2024; Redmon ve ark. 2016). YOLO ailesi son yıllarda büyük ölçüde gelişmiştir (Soni ve Rai, 2024). Bu sayede, hız ve doğruluk açısından gerçek zamanlı nesne tespiti için en güçlü algoritmalarından biri haline gelmiştir. Bu çalışmada YOLO ailesinin son versiyonları olan YOLOv8 ve YOLOv11, YOLOv12 kullanılmıştır.

YOLO algoritması, 2016 yılında ilk kez tanıtılmasından bu yana sürekli olarak evrilmiş ve her sürüm, mimari yapı ve hesaplama doğruluğu açısından önemli iyileştirmeler sunmuştur. Şekil 3.3' te YOLO sürümlerinin gelişimi gösterilmektedir.



Şekil 3.3. YOLO algoritmasının yıllar içindeki gelişimi.

YOLOv1 (Redmon ve ark., 2016), doğruluktan ödün vererek hız elde etmek için tek aşamalı algılama yöntemini tanıtmıştır.

YOLOv2 (Redmon ve ark., 2017), bağlantı kutuları, çok ölçekli eğitim stratejileri ve Darknet-19 omurgasını tanıtarak yerleştirme ve tespit doğruluğunu iyileştirmiştir.

YOLOv3 (Redmon ve ark., 2017), çok seviyeli algılama ve çoklu sınıflandırmayı desteklemek için Darknet-53 omurgasını ve Özellik Piramidi Ağı'nı (Feature Pyramid Network -FPN) kullanmıştır.

YOLOv4 (Bochkovskiy ve ark., 2020), CSPDarknet53 ve mozaik veri artırma gibi teknikleri birleştirerek, hesaplama maliyetini artırmadan tespit performansını geliştirmiştir.

YOLOv5 (Glenn ve ark., 2020), PyTorch üzerine kurulu, eğitim ve dağıtım süreçlerini basitleştirerek ve otomatik çapa öğrenimini desteklemiştir.

YOLOv6–v7 (Li ve ark., 2022; Wang ve ark., 2023), Neural Architecture Search (NAS), E-ELAN ve yeniden temsil edilen evrişim katmanları gibi yöntemlerle mimari verimliliğe odaklanmıştır.

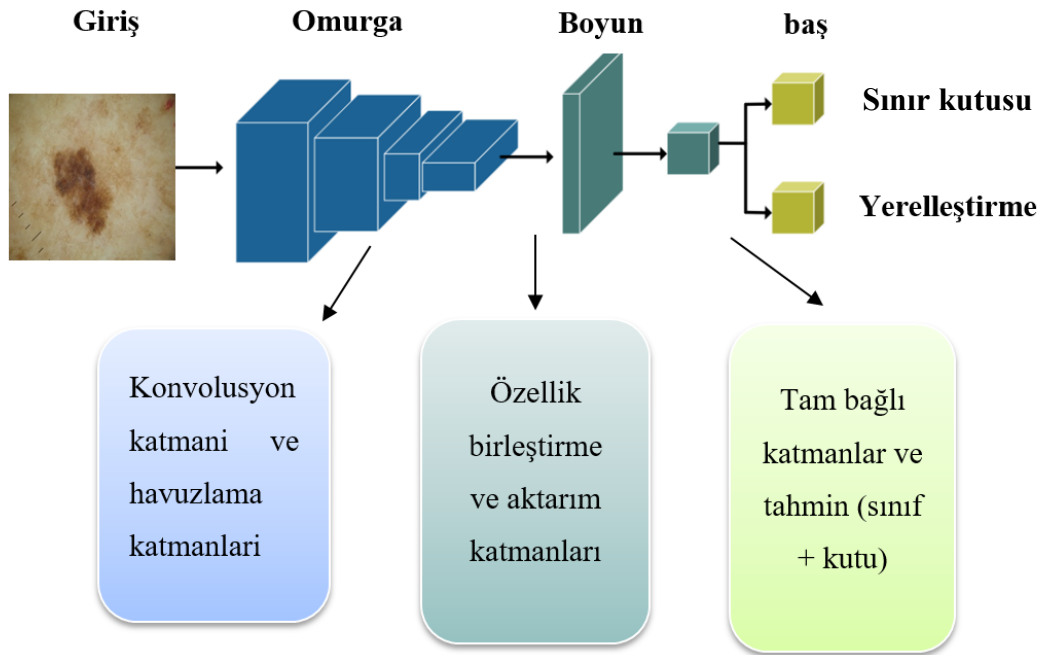
YOLOv8 (Glenn ve ark., 2023), çapa (anchor)sız algılama, panoramik segmentasyon ve anahtar nokta tespiti için Cross-stage Feature Fusion (C2f) modülünü entegre etmiş ve daha esnek bir yapı sunmuştur.

YOLOv9–v10 (Wang ve ark., 2024; Wang ve ark., 2024), hafif mimariler, maksimum olmayan bastırma (Non-Maximum Suppression-NMS) uçtan uca eğitim ve çift etiket ataması gibi yenilikler getirmiştir.

YOLOv11 (Khanam ve ark., 2024), Gelişmiş özellik çıkarımı için (C3k2) ve (C2PSA) modüllerini entegre ile modelin genel verimliliğini artırmıştır.

YOLOv12 (Tian, 2025), Gerçek zamanlı performansta Evrişimli Sinir Ağları'na (CNN'ler) uyum sağlamak için Bölgesel Dikkat Mekanizması (A^2) ve Residual-ELAN'ı (R-ELAN) yapılarını tanıtmıştır.

YOLOv1 mimarisi, nesne tespiti görevleri için özel olarak tasarlanmış, tam evrişimli bir sinir ağı yapısına sahiptir ve Şekil 3.4'te genel ağ yapısı sunulmaktadır. Bu mimarinin temel yapı taşlarını oluşturan omurga (backbone), boyun (neck) ve baş (head) bileşenleri, modelin girişten çıktıya kadar olan tüm bilgi işleme sürecini kapsamaktadır (Redmon ve ark. 2016).



Şekil 3.4. YOLO'nun genel mimarisi (Alif ve ark., 2024).

Omurga kısmı, CNN mimarisine dayanmakta olup, modelin giriş görüntüsünden üst düzey görsel özellikler çıkarmasını sağlar. Bu yapı, sınıflandırma alanında başarı göstermiş GoogLeNet mimarisinden ilham alarak geliştirilmiştir. Ancak Redmon ve arkadaşlarının (2016) çalışmasında önerdiği GoogLeNet'te kullanılan Inception modüllerinin aksine, YOLOv1'de 1×1 indirgeme katmanları ve 3×3 evrişim

katmanları dönüşümlü olarak kullanılmıştır. Özellikle bu 1×1 katmanlar, hesaplama yükünü azaltmak için öznitelik boyutlarını sıkıştırırken, ardından gelen 3×3 katmanlar uzamsal bilgileri yakalayarak yerel bağlamı modellemeye yardımcı olur. Bu omurga yapısı toplamda 24 evrişimli katman ve 2 tam bağlantılı katmandan oluşur. Ayrıca belirli aşamalarda kullanılan maksimum havuzlama (max-pooling) katmanları, özniteliklerin soyut seviyelere taşınması ve boyutların düşürülmesi için kullanılmıştır (Redmon ve ark. 2016).

Boyun kısmı, omurgada çıkarılan özellikleri daha soyut ve anlamlı temsillere dönüştürülmesini sağlayan bağlantı katmanlarını (fully connected layers) içerir. Bu katmanlar, özellikle sınırlayıcı kutu koordinatlarının ve sınıf olasılıklarının tahmini için gerekli olan yüksek seviyeli temsilleri öğrenmekle görevlidir. YOLOv1'de boyun kısmı, iki adet tam bağlantılı katmandan oluşur ve bu katmanlar, 4096 nöronluk geniş yapısıyla modelin son katmanına güçlü öznitelikler iletir (Redmon ve ark., 2016)

YOLO mimarisinde "boyun" katmanı ilk olarak YOLOv3'te kullanılmış olup, burada çok ölçekli özellikleri birleştirmek için FPN yapısı kullanılmıştır. YOLOv4'te boyun tabakasını daha da güçlendirmek için Yol Toplama Ağı (Path Aggregation Network - PAN) eklenmiştir. Daha sonraki YOLOv5, YOLOv8, YOLOv11 ve YOLOv12 versiyonlarında Uzamsal Piramit Havuzlama-Hızlı (Spatial Pyramid Pooling - Fast-SPPF), C2f modülü ve Dikkat Mekanizmaları gibi modüller eklenerek boyun kısmı geliştirilmiştir.

YOLOv1 ve YOLOv2'de boyun tabakası bulunmamaktaydı, yapı doğrudan omurgadan başa gidiyordu. Bu nedenle boyun kısmı sadece YOLOv3 ve sonrası versiyonlarda kullanılmaktadır. Bu katman önemlidir çünkü modelin küçük nesneleri daha doğru bir şekilde algılamasına ve bağlamsal bilgileri daha verimli bir şekilde işlemesine yardımcı olur (Redmon ve ark., 2016).

Modelin eğitimi öncesinde, evrişimsel katmanlar 224×224 çözünürlüğe sahip görüntüler ile ImageNet sınıflandırma görevinde ön eğitime (pretraining) tabi tutulmuştur. Nesne tespiti görevine geçildiğinde, ağın giriş çözünürlüğü iki katına çıkarılarak 448×448 piksele yükseltilmiştir. Bu sayede, daha detaylı görsel bilgilerin yakalanması mümkün hale gelmiştir.

Tespit aşamasında, Şekil 3.4 'te gösterilen YOLO algoritması, giriş görüntüsü $S \times S$ Boyutlarında ızgara bölgelerine ayrılır. Her ızgara hücresi, sinir ağı üzerinden işlenir

ve ilgili hücreye ait tahminler üretilir. Bu aşamadaki temel amaç, ızgara hücrelerinde yer alan nesneyi doğru şekilde tespit etmek ve onu sınır kutusu (bounding box) içerisinde sınıflandırmaktır. Eğer bir nesnenin merkezi belirli bir ızgara hücrelerine denk geliyorsa, o hücre ile ilgili nesnenin konum bilgilerini (genişlik ve yükseklik), ait olduğu sınıfı ve tahmine ilişkin güven skorunu hesaplamaktan sorumludur (Redom ve ark., 2016).

YOLO baş tespiti, modelin son katmanını temsil eder; burada omurga ve boyundan gelen özellikler analiz edilerek görüntüdeki nesnelerin yeri, boyutu ve sınıflandırması belirlenir. Başlık, sınırlayıcı kutuları çıkarmak ve nesne sınıflarını tanımlamak için özellik haritalarını işler. Dolayısıyla, bu katmanda her ızgara hücresi için bir veya daha fazla sınır kutusu tahmini yapılır. Şekil 3.5, YOLO tespit sistemine örnek olarak verilmiştir (Redom ve ark., 2016).



Şekil 3.5. YOLO'nun nesne tespit sistemi (Redom ve ark., 2016).

Sınır kutuları için güven (confidence score) skorları ile tahmin eder. Güven skoru iki ana öğeden meydana gelmektedir: bir kutunun bir nesne içermesi olasılığı, yani modelin o alanda bir nesnenin bulunduğu ne kadar güvendiğini ifade eder ve kutunun konumunun doğruluğu, tahmin edilen konumun nesnenin gerçek konumuyla ne kadar uyumlu olduğunu Kestirilmiş bölgelerin (Intersection over Union- IOU) puanı ile göstermektedir (Jang ve ark., 2022; Redom., 2016). Denklem 3.1'deki gibi hesaplanabilmektedir.

$$\text{Güven Skoru} = \text{Pr}(\text{nesne}) \times \text{IOU}_{\text{tahmin}}^{\text{doğru}} \quad (3.1)$$

$\text{Pr}(\text{nesne})$, ızgara içinde nesnenin bulunma olasılığı $\text{IOU}_{\text{tahmin}}^{\text{doğru}}$ ise gerçekte nesnenin bulunduğu kutu ile tahmin edilen kutunun kesişim değerini ifade etmektedir.

Öncelikle nesne olasılığı $Pr(\text{nesne})$ olarak tanımlanır, bu kutunun bir nesne içerip içermediğini gösterir. Eğer kutuda hiç nesne yoksa $Pr(\text{nesne})=0$ olarak tanımlanır; eğer içinde bir nesne varsa, o zaman $Pr(\text{nesne}) = 1$ 'dir.

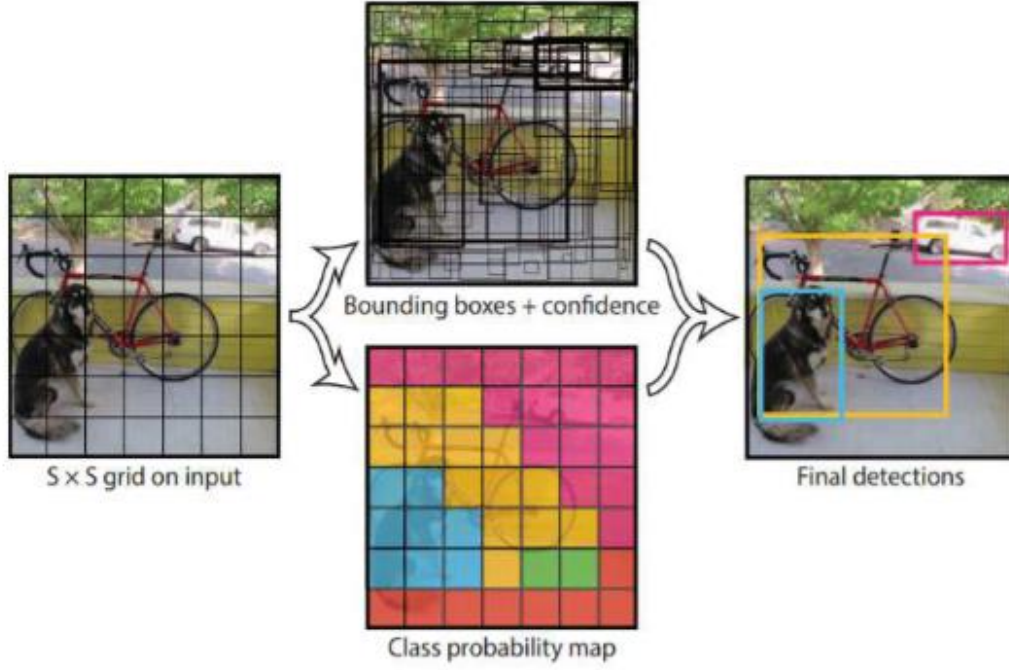
Her bir sınırlayıcı kutu için x, y, w, h ve güven skoru(g) olan 5 parametreden oluşur; burada (x,y) koordinatları, karenin ızgara hücreesindeki konumuna göre merkezini göstermektedir. Genişlik (w) ve yükseklik (h), tam görüntünün boyutlarının yüzdesi olarak tahmin edilmektedir. Son olarak güven skoru, tahmin edilen karenin gerçek kare (IOU) ile ne kadar yakından kesiştiğini göstermektedir (Diwan ve ark., 2022).

Her bir grid hücresi, ayrıca C adet koşullu sınıf olasılığı da tahmin eder: $Pr(\text{Sınıf}_i | \text{Nesne})$ ile gösterilir. B sınırlayıcı kutu sayısından bağımsız olarak, her ızgara hücresi için yalnızca bir sınıf olasılık kümesi tahmin edilir. Yani, olasılıklar nesneyi içeren ızgara hücresine bağlıdır. Denklem 3.2, koşullu sınıf olasılıkları ile test bölümündeki her bir kutunun güven tahminlerinin çarpımını temsil eder. Bu denklem, her sınırlayıcı kutu için sınıfa özgü bir güven puanı verir. Bu puan, ilgili sınıfın kutuda bulunma olasılığını ve tahmin edilen kutunun nesneyle ne kadar yakın eşleştiğini gösterir.

$$\begin{aligned} Pr(\text{sınıf}_i | \text{Nesne}) \times Pr(\text{Nesne}) \times IOU_{\text{tahmin}}^{\text{doğru}} \\ = Pr(\text{sınıf}_i) \times IOU_{\text{tahmin}}^{\text{doğru}} \end{aligned} \quad (3.2)$$

Sonuç olarak ortaya çıkan skor, her bir sınırlayıcı kutunun hem nesne içirme olasılığını hem de sınıf tahmin güvenini birlikte temsil eder (Liu ve ark., 2018).

Her hücrenin olasılık değerlerini tahmin etmek için C adet sınıfı kullanarak (x, y, w, h, g) değerlerini tahmin etmektedir. Başarıkları her bir hücre için tahminde bulundukları, sınıf değerleri C olmak üzere toplamda C sayıda değer tahmini yapmalıdır. Bu durumda, her hücre toplamda $B \times 5 + C$ sayı değer tahmin eder. Giriş görüntüsü $S \times S$ hücrelerden meydana gelen bir ızgaraya bölündüğünde, Şekil 3.6 'da gösterilen nihai çıktı $S \times S \times (B \times 5 + C)$ şeklinde bir tensör formatında verilir. Burada B , tahmin edilen sınır kutusu sayısını, C sınıf sayısını, S , görüntünün bölündüğü parça (hücre) sayısını temsil eder. Bu, yatay ve dikey olarak bölüm sayısını ifade eder (Redom ve ark., 2016).



Şekil 3.6. YOLO modeli, her ızgara hücresi için sınırlayıcı kutular ve sınıf olasılıklarını tahmin eder.

Sınıf olasılığı ve hedefe ait sınırlayıcı kutu koordinatları, kutunun genişliği ve yüksekliği görüntü boyutlarına göre normalleştirilir. Daha sonra x ve y koordinatları, ilişkili ızgara hücresi konumunun oranları olarak tanımlanır, böylece değerler 0 ile 1 aralığında kalır. Son katmanda doğrusal bir aktivasyon fonksiyonu kullanılırken, kalan katmanlar Leaky ReLU fonksiyonuna dayanır. Leaky ReLU fonksiyonu denklem 3.3'te gösterilmiştir (Redmon ve ark., 2016).

$$\phi(x) = \begin{cases} x, & x > 0 \\ 0.1x, & \text{aksi halde} \end{cases} \quad (3.3)$$

YOLOv1 modeli, eğitim sırasında konum ve sınıf tahmin hatalarını birlikte değerlendiren kareli hata toplamı (sum-squared error) tabanlı bir kayıp fonksiyonu kullanır. Ancak bu fonksiyon, tüm hatalara eşit ağırlık verir ve bu durum modelin doğruluğunu olumsuz etkileyebilir. Bu sorunu dengelemek için, konum hatalarını artıran $\lambda_{koordinat} = 5$ ve nesne içermeyen hücrelerdeki güven skoru hatalarını azaltan $\lambda_{nesneyok} = 0.5$ katsayıları denklem 3.4'teki gibi uygulanır. Ayrıca küçük kutulardaki hataların etkisini azaltmak amacıyla genişlik ve yükseklik yerine karekökleri kullanılır. Her hücreden yalnızca en yüksek IoU değerine sahip kutu sorumlu seçilir.

Bu yapı, modelin istikrarlı şekilde öğrenmesini ve daha doğru sonuçlar üretmesini sağlar (Redom ve ark., 2016).

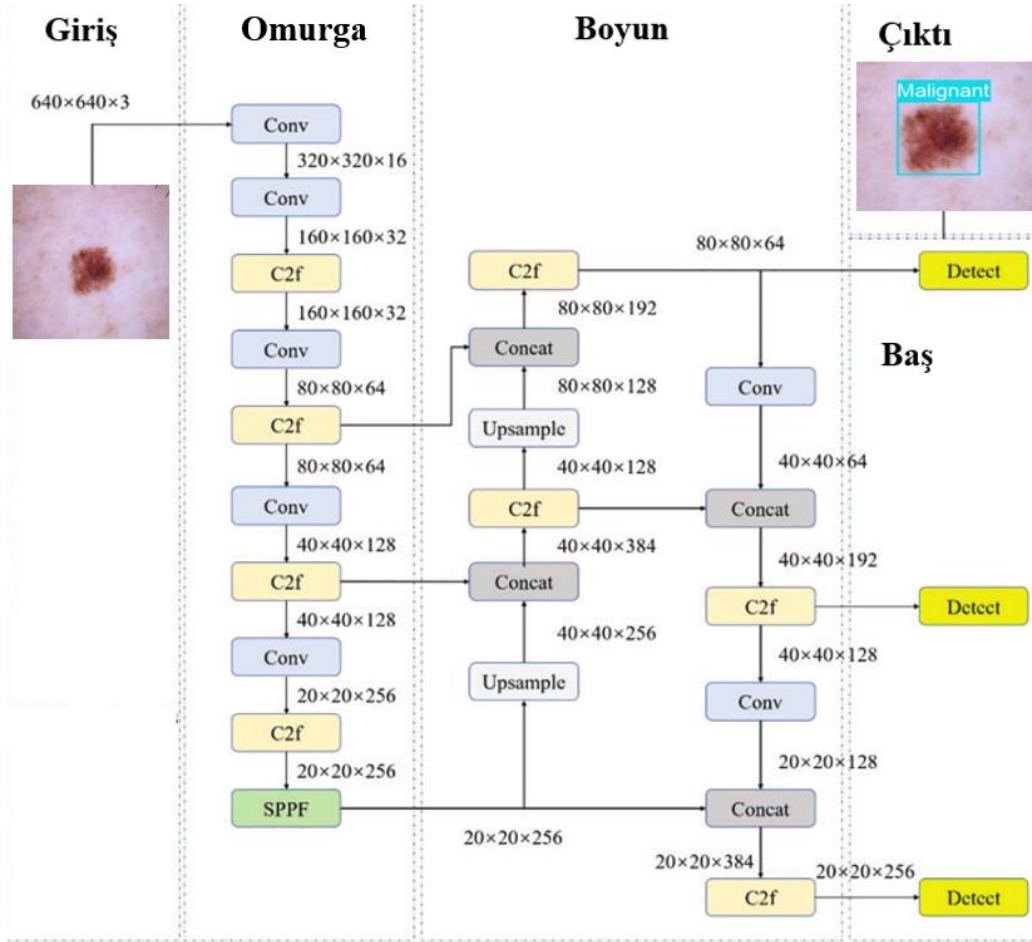
$$\begin{aligned}
& \lambda_{koord} \sum_{i=0}^{s^2} \sum_{j=0}^B j_{ij}^{obj} [(x_i - x'_i) + (y_i - y'_i)] \\
& + \lambda_{koord} \sum_{i=0}^{s^2} \sum_{j=0}^B j_{ij}^{nesne} \left[\left(\sqrt{\omega_i} - \sqrt{\omega'_i} \right)^2 \right. \\
& \left. + \left(\sqrt{h_i} - \sqrt{h'_i} \right)^2 \right] \\
& + \sum_{i=0}^{s^2} \sum_{j=0}^B j_{ij}^{nesne} (C_i - C'_i)^2 \\
& + \lambda_{nesneyok} \sum_{i=0}^{s^2} \sum_{j=0}^B j_{ij}^{nesneyok} [(C_i - C'_i)^2] \\
& + \sum_{i=0}^{s^2} j_{ij}^{nesne} \sum_{c \in \text{sınıf}}^B (p_i(c) - p'_i(c))^2
\end{aligned} \tag{3.4}$$

YOLOv1 algoritmasında, bir hücre i için (x_i, y_i) değerleri, sınırlayıcı kutunun o hücreye göre normalize edilmiş merkez koordinatlarını ifade eder. (w_i, h_i) ise kutunun genişlik ve yüksekliğinin, tüm görüntü boyutuna göre normalize edilmiş değerleridir. C_i modeli tarafından tahmin edilen güven skoru olarak tanımlanır. 1_{ij}^{obj} değeri, ilgili hücrede gerçekten bir nesne bulunduğunu belirtirken, 1_{ij}^{obj} ifadesi, j . sınırlayıcı kutu tahmincisinin, o hücredeki nesneye ilişkin tahminde bulunduğunu gösterir (Diwan ve ark., 2022).

3.2.1. YOLOv8 algoritması

YOLOv8, Ultralytics tarafından geliştirilmiş ve Ocak 2023'te yayınlanmıştır (Reis ve ark. 2023). Nesne tespiti alanında önceki YOLO sürümlerine göre birçok yapısal iyileştirme ve teknik yenilik sunan güncel bir algoritmadır. Nesne tespiti, segmentasyon, poz tahmini, izleme ve sınıflandırma dahil olmak üzere çok çeşitli bilgisayarlı görme görevlerini desteklerken, gerçek zamanlı uygulamalarda kullanılmaktadır. Bu model, hem tespit doğruluğunda hem de hız olarak kayda değer gelişmeler göstermiştir (Reis ve ark. 2023; Terven ve ark. 2023). Şekil 3.7 'de

belirtildiği gibi YOLOv8 mimarisi genel olarak üç ana bileşen üzerine kuruludur: omurga (Backbone), boyun (Neck) ve baş (Head).



Şekil 3.7. YOLOv8 mimarisi (Niu ve ark., 2024)

CSPDarkNet, YOLOv8'in YOLOv8, çapa ihtiyacını ortadan kaldırarak çapa bağımsız (anchor-free) bir yapı sunmaktadır (Ragab ve ark., 2024; Soni ve Rai., 2024) ve bu da onu daha esnek ve verimli hale getirmiştir. YOLOv8, geleneksel modellerin aksine görüntüdeki her nokta için birden fazla tahmin üretmeye dayanmaz; bunun yerine nesnenin merkezini doğrudan belirlemeye odaklanarak hesaplama karmaşıklığını önemli ölçüde azaltır (Farooq ve ark., 2024). Bu değişim, özellikle küçük nesneler için algılama doğruluğunu artırarak YOLOv8'i mevcut en verimli ve hızlı modellerden biri haline getirmiştir (Sohan ve ark., 2024).

Sinir ağının ana öznelilik çıkarma unsuru omurga ağıdır. Nesnelerin tanınmasını ve görüntülerin aşağı yönde sınıflandırılmasını sağlayacak mühim özellikleri girdi verileri sayesinde, örneklerin ayırt edilmesi için çıkarır. Eğitim esnasında modelin performans kalitesini kayıp fonksiyonu ölçer. Yolov8, Çapraz Aşamalı Kısmi Ağ 'ı

(Cross Stage Partial Network'ü - CSPNet)(Wang ve ark., 2020) Darknet'e entegre ederek omurgası olarak CSPDarknet'i oluşturur. CSPNet, büyük ölçekli omurga ağlarında sıkça karşılaşılan tekrar eden gradyan bilgisi sorununu, gradyan değişimlerini özellik haritasına entegre ederek çözmektedir (Radovanovic ve ark., 2024).

Bu sürümde, YOLOv5'te kullanılan C3 yerine C2f modülü tanıtılmıştır. Bu yeni modül, veriyi daha derin katmanlara aktarırken bilgi kaybını azaltır ve gradyan akışını optimize eder. Bu sayede yüksek düzeyli özellikleri bağlamsal bilgilerle bütünleştirerek algılama doğruluğunu artırmakta ve ince ayrıntıları ve karmaşık desenleri yakalamaya yardımcı olmaktadır (Han ve ark., 2024).

Boyun mimarisi, omurgadan bilgi toplayan ve ağdaki birden fazla seviyeden çıkarılan özellikleri entegre ederek tespit sürecini iyileştirmeye yardımcı olan ara katmanı temsil eder. YOLOv8, çeşitli boyutlardaki nesneleri tespit etmede modelin etkinliğini artırmak için FPN ve PAN mimarisine dayanır (Sohan ve ark., 2024)..

Nesne tespiti, boyut, renk ve parlaklık gibi nesne özelliklerindeki büyük farklılıklar nedeniyle zorluklarla karşı karşıyadır. Evrimsel ağlar hiyerarşik özellik temsilleri üretse de, bazı özelliklerin etkinliğini azaltan anlamsal boşluklardan muzdariptir. Bu zorlukların üstesinden gelmek için YOLOv8, özellikle küçük nesnelerin algılanmasında performansı artırırken yürütme hızını koruyan birden fazla özellik katmanını birleştiren FPN tekniğine güvenmektedir (Yaseen., 2024).

FPN iki ana bölümden oluşmaktadır. İlk yol (Aşağıdan Yukarıya Yol), özelliklerin hiyerarşisinin sınır ağı tarafından hesaplandığı yoldur. İkinci yol ise (Yanal Bağlantılar), iki farklı yoldan gelen özellikleri birleştirerek nesne tespiti sırasında yerleştirme ve sınıflandırma sürecini iyileştirir (Lin ve ark., 2017).

Öte yandan, özellikle bilgisayarlı görme alanındaki en zorlu görevlerden biri olan, görüntüdeki her pikselin belirli bir sınıfa sınıflandırıldığı anlamsal segmentasyon görevlerinde, algılama sonuçlarını iyileştirmek ve sınıflandırma doğruluğunu artırmak için ek bir geliştirme olarak PAN önerilmiştir. PAN, dermatoloji görüntülerinin analizinde, farklı hastalık sınıfları arasında ayırım yapmanın son derece zor olduğu büyük bir sorun olan, görüntülerdeki boyut değişiminin ele alınmasını iyileştirmeyi amaçlamaktadır (Lin ve ark., 2017; Ghiasi ve ark., 2019).

PAN yapısı Özellik Piramidi Dikkat modülü ve Global Attention Upsample modülünü içermektedir. Özellik Piramidi Dikkat modülü derin özelliklere hiyerarşik bir dikkat katmanı ekleyerek tahminleri iyileştirmeye yardımcı olur (Li ve ark., 2018). Global Attention Upsample modülü düşük seviyeli özellikleri yönlendirerek ince detayları çıkaran ve sınıflandırmayı kolaylaştırır.

PAN'ın performansı PASCAL VOC 2012 veri setinde üzerinde, PSPNet ve DeepLabv2 gibi rakip modellerle karşılaştırıldığında en yüksek değer olan %84,0 mIoU'ya ulaşmıştır. Performansı ayrıca 5.000 adet doğru şekilde açıklamalı görsel içeren bir veri setinde karşılaştırıldığında, PSPNet ve DeepLabv2 gibi modellerden %78,6 oranında daha iyi performans gösterdiği görülmüştür (Li ve ark., 2018).

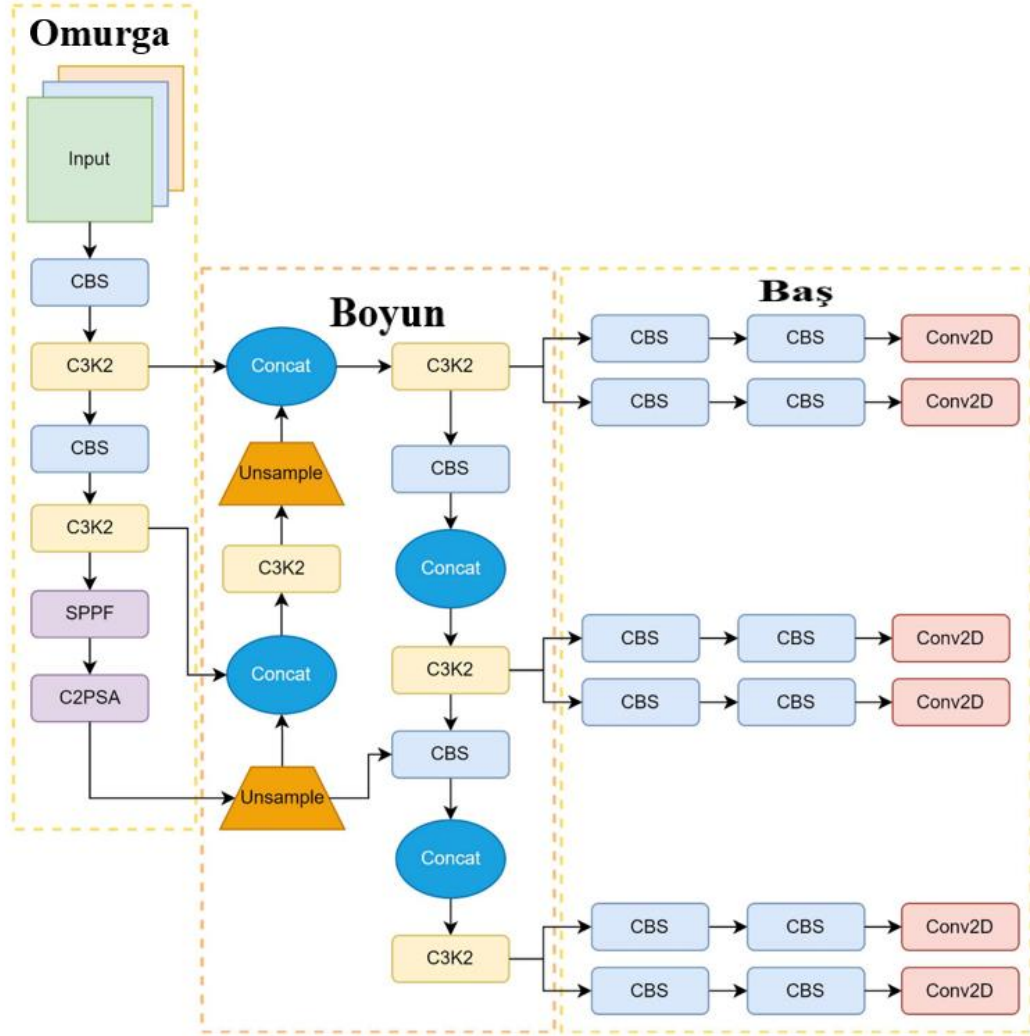
YOLOv8, modelinin özellikleri birden fazla ölçekte işleyerek performansı artıran bir SPPF modülünü içermektedir. Paralel MaxPooling'e dayalı geleneksel SPP, tahminlerin kalitesinden ödün vermeden hesaplama karmaşıklığını azaltan, sıralı ve paralel havuzlamaya dayalı SPPF'ye iyileştirildi (Sohan ve ark., 2024).

YOLO'nun baş kısmı ise, hesaplama karmaşıklığını azaltarak ve modelin küçük nesneleri doğru bir şekilde tespit etme yeteneğini geliştirerek, çapa gerektirmeyecek şekilde optimize edilmiştir. Ayrıca, aynı nesneye ait birden fazla tahmini ortadan kaldırarak doğruluk oranını artıran NMS algoritması kullanılmıştır (Hussain, 2024). YOLOv4 ve YOLOv5 gibi önceki sürümlerde algılama, COCO gibi belirli bir veri dağıtımına uyacak şekilde tasarlanmış, ancak özel veri kümelerine iyi ölçeklenmeyen bağlantı kutularına dayanıyordu. Bu, hedef nesnelerle eşleşecek şekilde kare boyutlarının dikkatli bir şekilde manuel olarak ayarlanmasını gerektirdiğinden önemli bir zorluktu ve bu da hesaplama karmaşıklığını artırıyordu. (Zhang ve ark., 2020).

3.2.2. YOLOv11 algoritması

YOLOv11, Ultralytics tarafından geliştirilen YOLO serisinin en yeni sürümüdür (Jocher ve ark., 2023). Bu sürüm, YOLOv8 ve YOLOv10'daki önceki geliştirmelere dayanarak mimari iyileştirmeler ve gelişmiş algılama doğruluğu sunar. YOLOv11, düşük güçlü cihazlardan bulut sistemlerine kadar çok çeşitli uygulamalarda yaygın olarak kullanılabilir. Algoritma, hem doğruluk hem de verimlilik açısından bilgisayarlı görme uygulamalarında üstün performans göstermektedir (Khanam ve ark., 2024; Sazak ve ark., 2024).

YOLOv11 modelinde omurga, ağın çekirdek bileşeni olarak, giriş görüntüsünden çok katmanlı ve çok ölçekli özellik haritalarını çıkarmakla sorumludur. Şekil 3.8 'de gösterildiği gibi model, özellik çıkarma ve işlemeye katkıda bulunan C3k2 (Jocher, 2024) ve C2PSA (Tsoi, 2024) blokları gibi yeni bileşenler sunar ve daha az parametreyle gelişmiş performans sağlamaktadır (Jegham ve ark. 2024; Khanam ve Hussain, 2024). Başlangıç olarak, C3k2 bloğu YOLOv8'deki C2f bloğunun geliştirilmiş bir versiyonudur ve hem omurga hem de boyun yapılarında alternatif olarak kullanılmaktadır. Bu blok, CSP mimarisi üzerine kurulmuş olup daha verimli ve hassas çalışmaktadır. C3k2 bloğu, büyük boyutlu evrişimler yerine iki küçük evrişim kullanır ve çıkarılan özelliklerin kalitesini korurken korumayı başarmaktadır (Alif, 2024).



Şekil 3.8. YOLOv11 mimarisi (Mao ve Hong, 2024).

C3k2, hesaplama verimliliğini ve çıkarım hızını artıran C3k ve Bottleneck olmak üzere iki yapıdan oluşur. C3k2 yapısı C3k adlı bir Boole parametresi tarafından kontrol edilir. C3k, FALSE olarak ayarlandığında, modül standart Bottleneck yapısını kullanarak öncülü C2f gibi davranır. TRUE olarak ayarlanırsa, C3k yapısıyla değiştirilir (Jegham ve ark., 2025; Khanam ve Hussain, 2024)

YOLOv11 modeli, çoklu ölçeklerde mekansal bilgileri entegre eden SPPF modülünü kullanmaya devam etmektedir. Bu yapı, modelin farklı boyutlardaki nesneleri, özellikle küçük nesneleri, daha etkin bir şekilde algılamasını sağlar (Khanam ve Hussain, 2024).

Buna ek olarak, YOLOv11, SPPF katmanının ardından YOLOv10'dan alınan C2PSA (Cross-Stage Partial with Spatial Attention) modülünü entegre etmektedir. C2PSA, iki temel yaklaşıma dayanmaktadır: birincisi olarak, Çapraz Aşamalı Kısmi (CSP) yapısı, özellik haritalarının yalnızca bir kısmını derin katmanlar üzerinden geçirerek hesaplama verimliliğini artırır. İkincisi ise, Mekansal dikkat (Spatial Attention) mekanizması ise modelin görüntü üzerindeki önemli bölgelere odaklanmasını sağlayarak, özellikle küçük veya üst üste binen nesnelerin tespitinde doğruluğu artırır.

Dolayısıyla, Bu modül, kanal ve mekansal bilgileri çok başlı dikkat (multi-head attention) mekanizmaları aracılığıyla birleştirerek özellik çıkarımını daha etkili hale getirir. Sonuç olarak, modelin özellikle karmaşık sahnelerde veya yüksek hassasiyet gerektiren durumlarda algılama başarımı önemli ölçüde geliştirilir. (Sazak & Kotan, 2025; Satya Mallick, LearnOpenCV, 2024).

YOLOv11, bu birleşimi daha verimli hâle getirmek ve özellikle önemli özelliklere odaklanırken önemsiz bilgileri baskılamak amacıyla C3k2 bloğunu boyun mimarisi içerisine entegre etmiştir. Bu sayede, hem özellik toplama işleminin hızı ve doğruluğu artırılmış hem de genel algılama performansı iyileştirilmiştir (Al Rabbani Alif, 2024).

Boyun fazı, çoklu çözünürlük seviyelerinde oluşturduğu zenginleştirilmiş özellik haritalarını daha sonra baş (head) yapısına ileterek, nesne sınıflandırması ve konum tespiti işlemleri için temel veri sağlar.

Son olarak, modelin baş (head) bileşeni, nesne algılama ve sınıflandırmaya ilişkin nihai tahminlerin üretilmesinden sorumludur. Özellik haritalarının kalitesini artırmak ve çok ölçekli derinliklerde etkili bir şekilde işlemek amacıyla, baş mimarisine birden fazla C3k2 bloğu entegre edilmiştir. Her bir algılama dalından sonra, çıkarılan

özellikleri gerekli çıktı formatına dönüştürmek için iki boyutlu evrişim katmanları (Conv2D) kullanılır. Bu katmanlar, sınırlayıcı kutuların koordinatlarını, nesnenin varlığına ilişkin güven skorunu ve ait olduğu sınıfa dair olasılıkları üretir. Son tahmin katmanı, önceki aşamalarda oluşturulan bilgileri birleştirerek, algılanan nesne için nihai çıktı olan konum, güven değeri ve sınıf tahminini üretir (He ve ark., 2025).

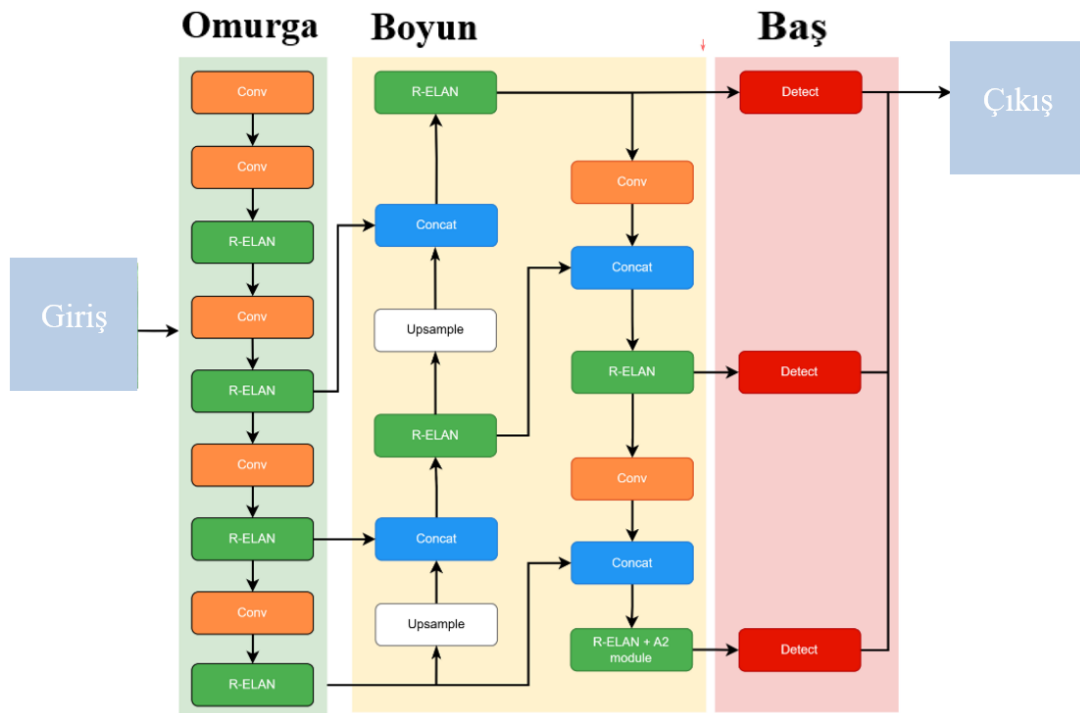
YOLOv11 mimarisinde yapılan iyileştirmeler, C3k2 bloklarının işlem verimliliği, C2PSA mekânsal dikkat mekanizmasının seçici odaklanma yeteneği ve gelişmiş özellik çıkarma modüllerinin katkısıyla, modeli YOLO serisindeki en verimli ve çok yönlü çözümlerden biri hâline getirmiştir.

3.2.3. YOLOv12 algoritması

YOLOv12, 2025 yılında YOLO modellerinin evriminde mimari açıdan büyük bir adımı temsil etmektedir. Bu modelin, omurga (backbone), boyun (neck) ve baş (head) bileşenlerini yeniden yapılandırıp, temel mimarisini gözleden geçirip geleneksel olarak sadece evrişimli sinir ağlarına (CNN) dayanan eski sürümlerden farklı olarak dikkat (attention) mekanizmalarına dikkat eden bir tasarıma geçilmiştir. YOLOv10 ve YOLOv11 gibi diğer sürümlerle kıyasla edildiğinde en önemli fark, bu dikkat tabanlı yapının olduğu gözlemlenir. YOLOv12'nin omurgasında kullanılan R-ELAN yapısı, katman düzeyinde artık bağlantılar (residual connections) ve ölçekleme teknikleri ile modelin daha derin yapılarda kararlılıkla eğitilmesini sağlamış, önemli bilgi kayıplarını önlemiş ve öğrenme sürecindeki gradyan akışını güçlendirerek eğitimi genel anlamda daha stabil hale getirmiştir (Tian ve ark., 2025).

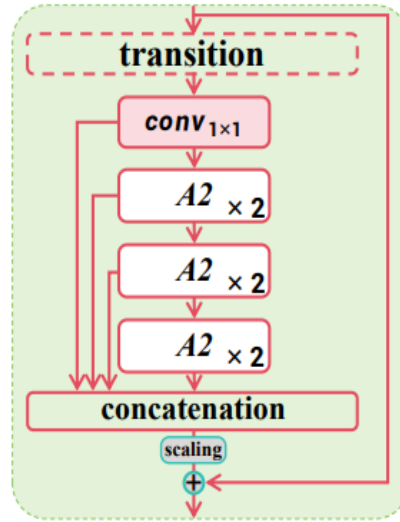
R-ELAN modülünde ayrıca dar boğaz benzeri bir yapı ile öznitelik haritalarının boyutu geçici olarak küçültülüp ardından genişletilerek daha karmaşık nesnelerin daha verimli şekilde temsil edilmesi sağlanmıştır. Bunun yanı sıra, A^2 (alan dikkat) olarak adlandırılan yeni bir dikkat mekanizması sayesinde öznitelik haritaları örtüşmeyen küçük bölümlere ayrılarak hem daha geniş alansal bağlam modellenmiş hem de klasik dikkat yöntemlerine kıyasla hesaplama maliyeti ciddi oranda azaltılmıştır. Modelin mekânsal kodlama kapasitesini artırmak için 7×7 ayırık evrişimler (separable convolutions) tercih edilmiş ve bu sayede parametre sayısı düşürülerek verimlilik artırılmıştır. Ayrıca, grafik işlemci ünitesi (Graphics Processing Unit-GPU) belleğini daha etkin kullanan ve hesaplama gecikmesini azaltan FlashAttention teknolojisi de entegre edilmiştir (Alif ve ark., 2025; Tian ve ark., 2025).

Bu, özellikle gerçek zamanlı nesne tespiti uygulamaları için büyük önem taşımaktadır. YOLOv12'nin genel yapısı geleneksel üç bölümlü düzenini (omurga, boyun, baş) korumaktadır. Şekil 3.9 'da gösterilen omurga; öznetelik çıkarımını, boyun; çok ölçekli öznetelik birleştirmeyi, baş; sınıflandırma ve konum tahmini görevlerini üstlenmektedir. COCO veri seti üzerinde yapılan mukayese testlerinde YOLOv12 modelleri, hem doğruluk (mAP) hem de çıkarım hızı açısından YOLOv10 ve YOLOv11 modellerinden daha da üstün performans göstermiştir. YOLOv12-N ve YOLOv12-S gibi daha küçük ölçekli sürümleri bile, RT-DETR gibi gelişmiş modelleri daha düşük kaynak kullanımıyla geride bırakmıştır (Sapkota ve ark., 2025).



Şekil 3.9. YOLOv12 mimarisi (Nguyen ve ark., 2025).

YOLOv12, önceki YOLO sürümlerinde kullanılan Verimli Katman Birleştirme Ağı (ELAN) mimarisine dayanan, ancak bazı geliştirmeler içeren ve Şekil 3.10 'da yapısı gösterilen R-ELAN adında geliştirilmiş bir omurga mimarisi sunmaktadır. Bu yapı, katman düzeyinde artık bağlantılar (residual connections) ve ölçekleme teknikleri ile birlikte gelir. Ayrıca, yeniden tasarlanmış özellik birleştirme yöntemiyle şişe boynu (bottleneck) benzeri bir yapı oluşturur. Bu sayede gradyan darboğazları azaltılır, özelliklerin entegrasyonu iyileştirilir ve özellikle dikkat (attention) temelli daha büyük modellerdeki optimizasyon zorluklarının üstesinden gelinmesine yardımcı olunur (Khanam ve ark., 2025).



Şekil 3.10. R-ELAN tasarımı (Tian ve ark., 2025)

YOLOv12'nin omurga yapısında 7×7 boyutunda ayırık evrişim katmanlarının kullanıldığı görülmektedir. Bu yapı, klasik büyük çekirdekli evrişimler ya da pozisyonel kodlama teknikleri yerine geçerek uzamsal bilgilerin daha verimli şekilde temsil edilmesini sağlar. YOLOv12 modelinde FlashAttention bileşeni, dikkat mekanizmalarındaki bellek erişim sorunlarını azaltmak için dizayn edilmiştir. Bu teknoloji, GPU belleğini daha verimli kullanarak çıkarım süresini hızlandırır. Böylece modelin, gerçek zamanlı uygulamalarda daha hızlı ve etkili çalışması sağlanmış olur. FlashAttention, YOLOv12'nin hız ve doğruluk dengesini korurken daha az kaynakla daha yüksek performans sağlamasına katkı sağlar (Tian ve ark., 2025).

YOLOv12'deki boynun, omurganın farklı katmanlarından gelen özellik haritalarını etkili bir şekilde birleştirerek farklı ölçeklerdeki giriş görüntülerinin geniş çaplı bir temsilini sağlayarak özellik toplamada kritik bir rol oynamaya devam ettiği gösterilmiştir. Başlangıçta omurgada tanıtılan R-ELAN mimarisi boyuna kadar uzanmaktadır. R-ELAN içindeki yeniden tasarlanan özellik toplama yöntemi, farklı seviyelerdeki özelliklerin etkili bir şekilde birleştirilmesini sağlayarak, modelin çeşitli boyutlardaki nesneleri daha iyi doğruluk ve verimlilikle tespit etme yeteneğine fayda sağlar. YOLOv12'nin boyun boyunca yukarı ve aşağı örnekleme yapmak için esnek süreçler kullanabileceği bildirilmektedir. Bu uyarlanabilir ölçekleme tekniklerinin, çok ölçekli özelliklerin birleştirilmesini geliştirmesi ve modelin giriş görüntüsündeki boyutlarından bağımsız olarak nesneleri daha iyi uyarlamasına ve doğru bir şekilde tespit etmesine imkan kılması muhtemeldir. YOLOv12'deki birebir güncellemeler, özellikle R-ELAN prensiplerinin uygulanması ve potansiyel olarak daha esnek

ölçekleme işlemleri, çok ölçekli özelliklerin verimli ve etkili bir şekilde bir araya getirilmesine odaklanmaktadır. Bu, çok çeşitli nesne boyutlarında sağlam nesne algılama performansı için kritik bir husustur (Alif ve ark., 2025; Khanam ve ark., 2025).

YOLOv12 başının, omurga ve boyun tarafından işlenen özelliklere dayalı nesne lokalizasyonu (sınırlayıcı kutu regresyonu) ve nesne sınıflandırmasının son görevlerinden sorumlu olduğu tanımlanmaktadır. Bu temelde, kafa mimarisi içerisinde daha büyük alıcı alanların potansiyel kullanımı tartışılmaktadır. Daha geniş bir alıcı alan, tahmin kafasının algılanan nesneler etrafındaki daha geniş bağlamı hesaba katmasına olanak tanır; bu da özellikle tıkanıklıkların veya kalabalık arka planların olduğu karmaşık sahnelerde yerleştirme ve sınıflandırmada daha iyi doğruluk sağlayabilir (Tian ve ark., 2025).

YOLOv12'deki güncellemeler nihai tahmin aşamasını iyileştirmeye odaklanmaktadır. Daha büyük alıcı alanlar ve SiLU (sigmoid doğrusal birim) gibi gelişmiş aktivasyon fonksiyonlarının dahil edilmesiyle model, daha iyi yerleştirme ve sınıflandırma yetenekleriyle daha doğru ve güvenilir nesne tespiti sonuçları elde etmeyi hedeflemektedir (Tian ve ark., 2025).

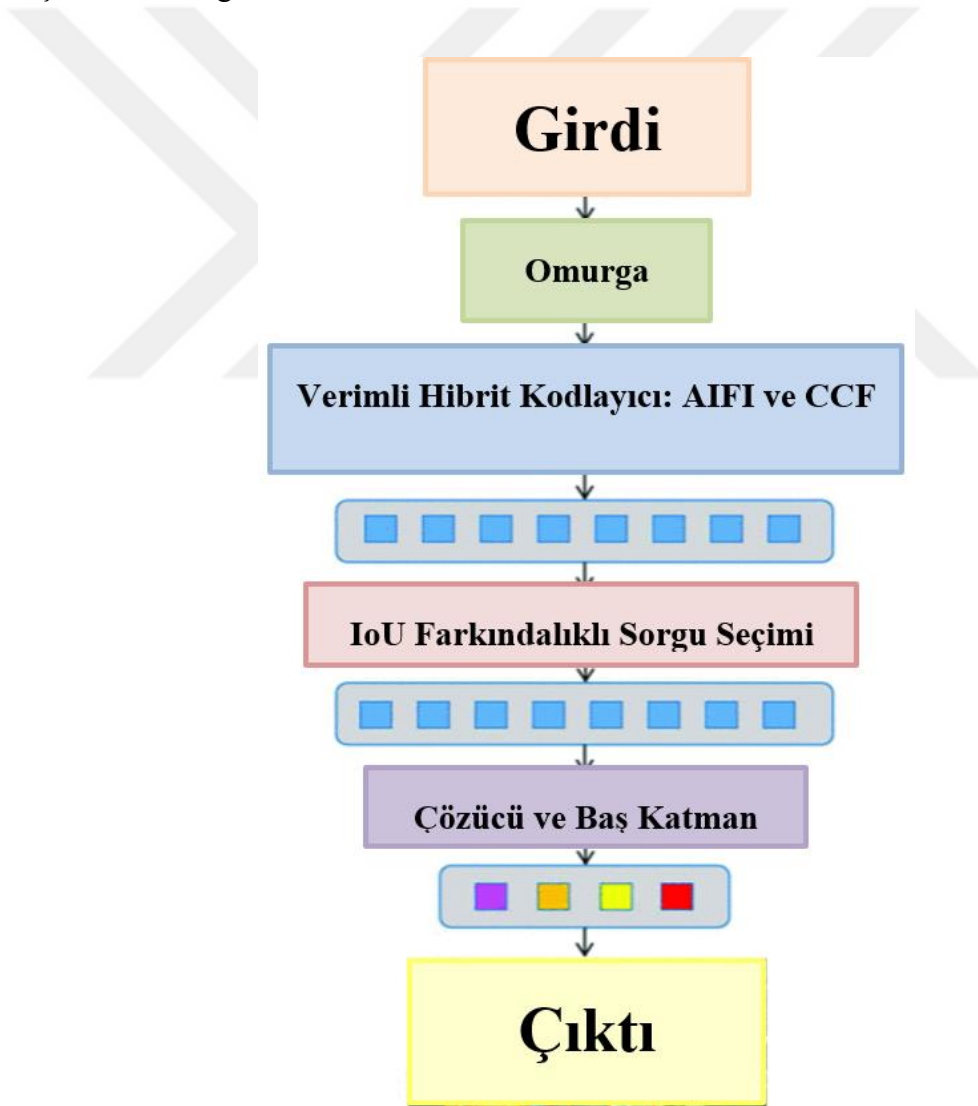
3.3. RT-DETR Algoritması

RT-DETR (Real-Time Detection Transformer), DETR modelinin sağladığı doğruluğu genişletmek üzere geliştirilen yeni nesil bir nesne tespiti modelidir (Zhao ve ark., 2024). RT-DETR, DETR modelinin doğruluğunu korurken aynı zamanda gerçek zamanlı işlem kapasitesi sunduğu için YOLO mimarisiyle benzerlik gösterir ve bu durum da nesne tespiti görevlerinde hem hassasiyet hem de hızda büyük gelişmeler sağlamaktadır. Yine de şu anda nesne tespiti modelleri, işlem sonrası adımlar uygulamadan doğrudan çıktıları kullanamaz, örneğin NMS (Carion ve ark., 2020; Redmon ve ark., 2016). Bu yaklaşım, tıbbi görüntüleme gibi küçük ve yoğun nesnelerin algılanmasında kritik öneme sahip olan RT-DETR tarafından benimsenen bir yeniliktir (Sun ve ark., 2023).

Klasik DETR mimarisi, elle tanımlanmış statik ön tanımlayıcılar yerine doğrudan öğrenilen sorguları kullanarak çalışır, böylece veri ön işleme sürecini basitleştirir. Bununla birlikte, özellikle küçük boyutlu nesneler için yüksek hesaplama maliyeti ve DETR'nin düşük algılama başarısı, uygulamasındaki faktörleri sınırlamaktaydı. RT-

DETR, Zhang ve ark., (2023) tarafından bu sorunları ele almak için yapısal verimlilik ve doğruluk için geliştirilmiş ve optimize edilmiştir ve 2024'te CVPR konferansında sunulmuştur (Zhang ve ark., 2023). Bu, özellikle cilt lezyonları gibi ayrıntıları gerektiren senaryolarda, yüksek doğruluk oranlarında üstün model başarısı ile sonuçlandı.

Modelin temelinde, çok ölçekli özellik çıkarımı için ResNet-50 veya ResNet-101 gibi derin evrişimli sinir ağları (CNN) kullanılmaktadır. Bu özellik haritaları genellikle S3, S4 ve S5 katmanlarından elde edilir. Alt katmanlar (örneğin S3), pigmentasyon değişiklikleri gibi ince ayrıntıları yakalarken, üst katmanlar (S5) daha soyut ve bağlamsal bilgiler sağlar (Zhang ve ark., 2023). RT-DETR algoritmanın mimarisi Şekil 3.11'de göstermektedir.



Şekil 3.11. RT-DETR mimarisi (Jun, 2024).

RT-DETR'nin en dikkat çekici bileşenlerinden biri olan Hibrit Kodlayıcı, hem işlem yükünü azaltmakta hem de çok ölçekli özellikleri etkin bir şekilde işlemektedir. Bu yapı iki ana bileşenden oluşur: Dikkat tabanlı ölçek içi etkileşim (AIFI) ve evrişimsel ağ tabanlı ölçekler arası füzyon (CCFF) modülü. Yalnızca S5 katmanında uygulanan ve self-attention mekanizmasıyla zengin semantik temsiller oluşturan AIFI (Attention-based Intra-scale Feature Interaction) modülüdür. AIFI modülü, tek bantlı bir transformatör kodlayıcı kullanarak yalnızca beşinci özellik düzeyinde (S5) bant içi etkileşim gerçekleştirir. S5, görüntüdeki nesneler arasındaki kavramsal ilişkileri tanımlamaya katkıda bulunan zengin anlamsal anlama sahip yüksek düzeyli özellikler içerdiği için seçilmiştir. Daha düşük düzeyler (S1–S4) bu kavramsal ilişkileri açıkça içermez, bu nedenle dikkat mekanizması, yedekliliği önlemek ve hesaplama kaynaklarından tasarruf etmek için bunlara uygulanmaz. Bu modül, düşük seviyeli katmanları dışarda bırakarak hesaplama süresini %35 azaltırken doğruluk oranında %0.4 artış sağlamaktadır (Chen ve ark., 2022). İkincisi ise CCFF (Cross-scale Feature Fusion) modülüdür. CCFF modülü, çoklu seviyelerden (S3, S4 ve AIFI'den elde edilen F5) çıkarılan özellikleri evrişim (convolution) katmanları kullanarak birleştirmeyi amaçlar. CCFF içerisinde, her biri iki bitişik özellik seviyesini daha bütünleşik bir temsile dönüştüren birden fazla birleştirme bloğu (Fusion Block) bulunur. Özellikle küçük, kümelenmiş veya üst üste binmiş cilt lezyonlarının doğru ayrıştırılmasını sağlar (Wu ve ark., 2025; Mingji ve ark., 2024).

Model ayrıca, IoU farkındalığına sahip bir sorgu seçme mekanizması kullanarak sabit sayıda özellik temsili seçer ve bunları kod çözücü (decoder) için başlangıç sorguları olarak kullanır. Bu sorgular, her yinelemede iyileştirilerek nihai tahmin kutuları ve güven skorları üretilir (Zhang ve ark., 2023). Ek olarak, RT-DETR modeli, “belirsizliği en düşük sorgu seçimi” (uncertainty-minimal query selection) yöntemini önererek arka planla benzer renge sahip veya pigmentasyonu düşük lezyonlarda hata oranını azaltır (Huang ve ark., 2024; He ve ark., 2025).

Nihai tahminler, özel olarak yapılandırılmış algılama kafaları (detection heads) aracılığıyla elde edilir. Bu aşamada, iki farklı evrişimsel katman kullanılır: Conv1x1 s1 (1x1 evrişim, Batch Normalization ve SiLU aktivasyonu) ve Conv3x3 s2 (3x3 evrişim, Batch Normalization ve SiLU aktivasyonu). Bu yapı, mekânsal ayrıntıların hassas şekilde işlenmesini sağlar (Zhang ve ark., 2023).

RT-DETR'nin temel mimarisi (geniř omurga ađı, hibrit kodlayıcı ve transformer tabanlı decoder) korunmakla birlikte, RT-DETRv2 ve RT-DETRv3 sűrűmleri denetim stratejileri, sorgu seęimi teknikleri ve ok lekli zellik etkileřimlerini geliřtirerek hem daha yűksek dođruluk hem de daha hızlı ıkarım sűresi sađlamayı amalamaktadır (Lv ve ark., 2024; Wang ve ark., 2024).

Bu geliřmiř mimari sayesinde RT-DETR, medikal tanı gibi hassas uygulamalarda gűvenilir ve hızlı nesne algılama yetenekleri sunan gűlű bir model haline gelmiřtir.





4. METODOLOJİ

4.1. Veri Toplama

Veri setleri, cilt kanseri erken teşhis çalışmalarının temel bir bileşenidir. Bu veri setleri gerçek klinik verilerden alınır ve farklı boyut, kontrast ve renk seçeneklerine sahip görüntülerden oluşur. Bu veri kümeleri, derin öğrenmeye dayalı otomatik tespit ve sınıflandırma sistemlerinin geliştirilmesinde referans kaynağı olarak kullanılır. Bu sistemler, tanı doğruluğunu artırmaya ve manuel değerlendirmeye olan bağımlılığı azaltmaya katkıda bulunarak, yapay zekâ tabanlı tıbbi uygulamaları destekleme yolunda önemli bir adım teşkil eder (Dildar ve ark., 2018).

Bu çalışmada kullanılan veri seti, ISIC arşivine dayanan ve Kaggle platformu üzerinden temin edilen bir kaynak veridir (ISIC Archive; Kaggle (2018)). Veriler, eğitim ve test olmak üzere iki ana gruba ayrılmıştır. Tablo 4.1 'de verilen eğitim veri setinde 1440 adet iyi huylu ve 1197 adet kötü huylu lezyon görüntüsü yer almaktadır. Test veri setinde ise 360 adet iyi huylu ve 300 adet kötü huylu lezyon görüntüsü bulunmaktadır. Bu doğrultuda, toplam veri seti 3297 dermoskopik görüntüden oluşmaktadır.

Çalışma kapsamında veri kümesi dengeli olacak şekilde %50 eğitim, %30 doğrulama ve %20 test oranlarında bölünmüştür. Bu bölme işlemi sırasında her bir alt kümede iyi huylu ve kötü huylu örneklerin dengeli şekilde temsil edilmesine özen gösterilmiştir.

Tablo 4.1. Verisetinin bölünmesi

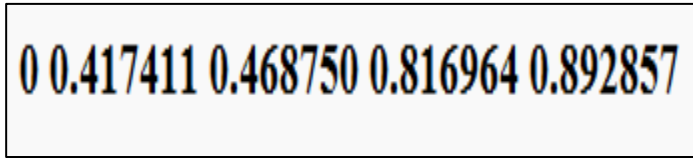
Sınıf	Eğitim	Doğrulama	Test	Toplam
Benign	1008	432	360	1800
Malignant	838	359	300	1497
Toplam	1846	791	660	3297

4.2. Veri Etiketleme

Bu çalışmada kullanılan tüm görüntülerin etiketleme işlemleri, yaygın olarak tercih edilen açık kaynaklı bir etiketleme aracı olan LabelImg (Labelimg, 2024) platformu aracılığıyla gerçekleştirilmiştir. Etiketleme sürecinde (Tousch AM ve ark., 2012), her bir görüntüde yer alan cilt lezyonları manuel olarak işaretlenmiş ve YOLO formatıyla uyumlu olacak şekilde, sınıf etiketi ile normalize edilmiş sınırlayıcı kutu koordinatları (center_x, center_y, width, height) biçiminde .txt formatında kaydedilmektedir.

Veri kümesindeki görüntüler iki sınıfa ayrılmıştır, 0 iyi huylu lezyonu, 1 ise kötü huylu lezyonu temsil etmektedir. Bu sınıf etiketleri, modelin sınıflandırma sürecinde doğru öğrenmeyi sağlamak amacıyla etiket dosyalarına entegre edilmiştir. Etiketleme tamamlandıktan sonra tüm veri kümesi, eğitim öncesinde manuel olarak gözle kontrol edilerek doğruluk açısından denetlenmiştir.

Etiketleme işlemi tamamlandıktan sonra, her görüntü için sınıf bilgisi ve normalize edilmiş koordinatları içeren Şekil 4.1 'deki gibi .txt dosyaları oluşturulmuştur. Her satır bir nesneyi temsil eder. Şekilde gösterilen ifade, görüntüdeki iyi huylu bir lezyonun konum ve boyutunu göstermektedir. Ayrıca, eğitim sürecinde bu etiketleri tanımlamak için model mimarisi ve sınıf bilgilerini içeren bir YAML dosyası kullanılmıştır. Bu yapı, modellerin farklı veri kümelerine kolayca uyarlanmasını sağlamaktadır.



Şekil 4.1. .txt dosyasından oluşan bir sınıfın etiketlemesi

Etiketlenmiş görüntüler, YOLOv8, YOLOv11, YOLOv12 ve RT-DETR olmak üzere dört farklı nesne tespiti algoritmasının eğitimi için kullanılmıştır. Tüm modeller aynı veri kümesi üzerinde eğitilmiş, bu sayede mimari farklılıklara rağmen objektif ve tutarlı bir karşılaştırma ortamı sağlanmıştır.

4.3. Modellerin Eğitilmesi

Bu çalışmada, cilt kanseri görüntülerinin nesne tespiti amacıyla YOLO mimarisine ait üç farklı sürüm (YOLOv8, YOLOv11 ve YOLOv12) ile Transformer tabanlı RT-DETR algoritması kullanılarak toplam 15 farklı model geliştirilmiştir. Geliştirilen

modeller sürümlerine göre sırasıyla YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, YOLOv8x, YOLOv11n, YOLOv11s, YOLOv11m, YOLOv11l, YOLOv11x ve YOLOv12n, YOLOv12s, YOLOv12m, YOLOv12l ve RT-DETR Large olarak adlandırılmaktadır. Modellerin katman sayıları ve toplam parametre miktarları Tablo 4.2, 4.3, 4.4, 4.5’de sunulmaktadır. RT-DETR-L modeli, 681 katman ile en derin model olurken, 68.154.534 parametre sayısı ile en büyük model ise YOLOv8x olmuştur.

Tablo 4.2. YOLOv8 modellerinin katman ve parametre sayısı.

Model	Katman Sayısı	Parametre Sayısı
YOLOv8n	225	3,011,238
YOLOv8s	225	11,136,374
YOLOv8m	295	25,857,478
YOLOv8l	365	43,631,382
YOLOv8x	365	68,154,534

Tablo 4.3. YOLOv11 modellerinin katman ve parametre sayısı.

Model	Katman Sayısı	Parametre Sayısı
YOLOv11n	319	2,590,230
YOLOv11s	319	9,428,566
YOLOv11m	409	20,054,550
YOLOv11l	631	25,312,022
YOLOv11x	631	56,876,086

Tablo 4.4. YOLOv12 modellerinin katman ve parametre sayısı.

Model	Katman Sayısı	Parametre Sayısı
YOLOv12n	225	3,011,238
YOLOv12s	225	11,136,374
YOLOv12m	295	25,857,478
YOLOv12l	365	43,631,382

Tablo 4.5. RT-DETR modelinin katman ve parametre sayısı.

Model	Katman Sayısı	Parametre Sayısı
RT-DETR-L	681	32,810,186

Tüm modellerin eğitimi Windows 10 işletim sistemine sahip, Intel Core i7-12700H 2.30GHz işlemcili ve NVIDIA GeForce RTX 3060 GPU'lu bir donanım üzerinde gerçekleştirilmiştir. Eğitim sürecinde Python 3.12.4 ve PyTorch 2.3.1 kütüphaneleri kullanılmıştır. CUDA sürümü olarak 12.5, cuDNN sürümü olarak 9.2 tercih edilmiştir. Modellerin eğitimi sırasında kullanılan konfigürasyon Tablo 4.6'te verilmiştir.

Tablo 4.6. Deneysel kurulum için yapılandırma detayları.

Özellik	Yapılandırma
Programlama Dilli	Python 3.12.4
Derin Öğrenme Çatısı	PyTorch 2.3.1
CUDA Sürüm	CUDA 12.5
cuDNN Sürüm	cuDNN 9.2
İşletim Sistemi	Windows 10
İşlemci	Intel Core i7-12700H @ 2.30 GHz
Başlangıç Öğrenme Oranı	0.001667
Momentum	0.9

Tablo 4.7.'de eğitim aşamasında kullanılan parametreler ve değerleri gösterilmektedir. Eğitim süreci boyunca her model için maksimum 300 dönem sayısı ile eğitim yapılmış, yığın boyutu değeri 8 olarak belirlenmiştir. Ultralytics kütüphanesinin otomatik hiperparametre optimizasyon aracı, modeller için AdamW (Loshchilov ve Hutter, 2017) algoritmasını seçmiştir. Başlangıç öğrenme oranı 0.001667, momentum değeri ise 0.9 olarak atanmıştır.

Tablo 4.7. Geliştirilen modellerin eğitim parametreleri ve değerleri.

Parametre	Değer	Parametre	Değer
Dönem (Epoch)	300	Erken durdurma için beklenecek dönem sayısı	100
Yığın sayısı (Batch size)	8	Çalışan sayısı	8
Momentum	0.9	Optimizasyon algoritması	AdamW
Giriş görüntü boyutu	224x224	Ağırlık Azaltma	0.0005
Isınma dönemi sayısı	3	İlk Öğrenme Oranı	0.001667
Isınma momentumu	0.8	Son Öğrenme Oranı	0.01
Isınma bias öğrenme oranı	0.1		

Aşırı öğrenmeyi (overfitting) önlemek amacıyla, her modelde erken durdurma (early stopping) (Prechelt, 2002) tekniği uygulanmıştır. Bu yöntemde, doğrulama verisi üzerindeki başarı ölçütlerinde 100 ardışık dönem boyunca iyileşme sağlanamazsa eğitim durdurulmuştur. Bu yaklaşımla, bazı modeller belirlenen 300 dönem sayısına ulaşamazken, bazıları maksimum dönem sayısına kadar eğitilmiştir. Modellerin tamamladığı dönem sayıları sırasıyla 300, 284, 295, 139, 194, 281, 295, 300, 300 ve 300'dür. Eğitim süreleri ise yaklaşık 4.4 saat (YOLOv11n) ile 141.7 saat (YOLOv11x) arasında değişmektedir. Modellerin eğitimi sırasıyla 4.979, 5.649, 9.002, 8.752, 74.814, 4.445, 5.155, 9.187, 14.359 ve 141.770 saat sürmüştür.

4.4. Değerlendirme Ölçütleri

Nesne tespiti, bilgisayarla görü alanının temel araştırma konularından biridir ve görüntü içerisindeki nesnelerin tanımlanması ve konumlandırılmasını hedeflemektedir. Bu çalışmada, YOLOv8, YOLOv11 ve YOLOv12 algoritmalarının 5 farklı modellerinin ve RT-DETR algoritmasının bir modelini kullanılarak geliştirilen modellerin doğrulama veri seti ve test veri seti üzerindeki performanslarının değerlendirilmesinde nesne tespiti modellerinin değerlendirilmesinde yaygın olarak kullanılan kesinlik (precision), duyarlılık (recall), F1-Skor, mAP (Mean Average Precision) ve Kare Başına Ortalama Çıkartım Süresi (melanocyte inducing transcription factor-MITF) performans ölçütleri kullanılmıştır.

Modelin bir nesneyi “doğru tespit ettiğini” söyleyebilmek için, tahmin edilen sınırlayıcı kutunun (bounding box) gerçek kutu ile yeterli düzeyde örtüşmesi gerekir. Bu örtüşme genellikle Intersection over Union (IoU) ile ölçülür (Yu ve ark., 2016). Eğer IoU belirli bir eşik değerinin üzerindeyse (örneğin 0.5), bu tahmin doğru kabul edilir. Bu noktada doğru pozitif (TP), yanlış pozitif (FP) ve yanlış negatif (FN) kavramları devreye girer. Bunlar üzerinden çeşitli metrikler hesaplanarak model başarısı sayısal olarak ifade edilir (Zhang ve ark., 2024).

Gerçek Pozitif (TP), Gerçek kutusuyla eşleşen bir kutunun doğru tespit edilmesidir. Yanlış Pozitif (FP), var olmayan bir nesnenin yanlış tespit edilmesi veya var olan bir nesnenin yerinin yanlış tespit edilmesi anlamına gelir. Son olarak, yanlış negatif (FN), model tarafından tespit edilemeyen gerçek bir kutunun varlığıdır. Ayrıca nesne tespiti bağlamında "Gerçek Negatif (TN)" kavramının kullanılmadığını da belirtelim, çünkü

herhangi bir görüntüde tespit edilmemesi gereken sonsuz sayıda kutu bulunmaktadır (Sahafi., 2024).

Tahmin edilen	
Gerçekleşen	Gerçek Pozitif (TP)
	Yanlış Negatif (FN)
	Yanlış Pozitif (FP)
	Gerçek Negatif (TN)

Şekil 4.2. IoU değerine göre tespit kalitesi örnekleri.

Kesinlik (Denklem 4.1’de), modelin tespit ettiği nesnelerden kaç tanesinin gerçekten doğru olduğunu gösterir. Başka bir deyişle, modelin “bu bir deri kanseri” dediği kaç görüntünün gerçekten doğru olduğunu ifade eder. Kesinlik, modelin pozitif olarak tanımladığı tüm nesneler arasından doğru tanımlananların oranını ölçer. Kesinlik değeri, denklemi 4.1 ile hesaplanmaktadır (Önder ve Karan., 2024).

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

Duyarlılık metriği, özellikle yanlış negatif (gerçek pozitif olduğu hâlde modelin negatif tahmin ettiği) hataların maliyetinin yüksek olduğu durumlarda kritik öneme sahiptir (Önder ve Karan., 2024). Duyarlılık hesaplama denklem 4.2’de verilmiştir.

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

Kesinlik ve duyarlılık genellikle birbirini dengeleyen iki ölçüttür. Model çok seçici davranırsa kesinlik değeri yükselirken duyarlılık değeri düşebilir. Tersine, model daha geniş tahminlerde bulunursa duyarlılık artarken kesinlik düşebilir. Bu nedenle bu iki metrik arasındaki dengeyi ifade eden F1-Skor kullanılır. Denklem 4.3’te verilen F1-Skor, kesinlik ve duyarlılığın harmonik ortalamasıdır ve her iki değer dengeli olması gereken durumlarda faydalı bir ölçüttür. Ancak, YOLO gibi çok sınıflı nesne tespiti modellerinde, sınıf dengesizliklerinden dolayı F1-Skor tek başına yeterli bir gösterge olmayabilir.

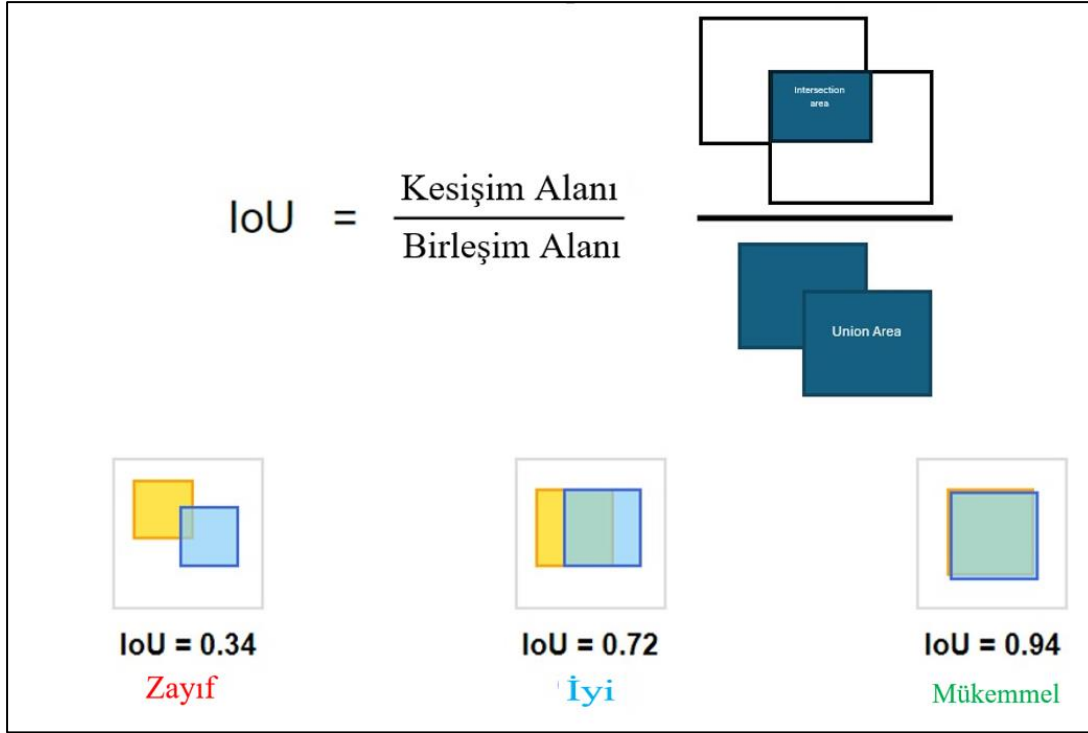
$$F1 - Skor = \frac{2 * Kesinlik * Duyarlilik}{Kesinlik + Duyarlilik} \quad (4.3)$$

Yukarıdaki tanımlara dayalı olarak gerçek ve yanlış tespiti belirlemek için kullanılan yaygın bir yöntem "Kesişim ve Birleşim (IoU)" kriterini kullanmaktır. Bu kriter, iki veri seti arasındaki benzerliği ölçmek için Jaccard Endeksi'ne dayanmaktadır. IoU (Intersection over Union), tahmin edilen kutu ile gerçek kutunun ne kadar örtüştüğünü gösteren temel konumlama metriğidir. Yüksek bir IoU değeri, modelin yalnızca “ne” olduğunu değil aynı zamanda “nerede” olduğunu da doğru tahmin ettiğini gösterir. Denklem 4.4'te görüldüğü gibi IoU, Nesne tespiti işleminin ardından, modelin nesnenin konumunu ne derece doğru tahmin ettiğini değerlendirmek amacıyla, tahmin edilen sınırlayıcı kutu (B_1) ile gerçek sınırlayıcı kutu (B_2) arasındaki birlik-kesişim oranı (Intersection over Union - IoU) hesaplanır. Bu oran, denklem (4.4)'te ifade edilmektedir.

$$J(B_1, B_2) = IOU = \frac{alan(B_1 \cap B_2)}{alan(B_1 \cup B_2)} \quad (4.4)$$

Elde edilen IoU değeri, önceden belirlenmiş bir eşik değeri t ile karşılaştırılarak tespitin doğruluğu değerlendirilir. Eğer $IoU(B_1, B_2) \geq t$, ise bu tespit doğru (true positive) olarak kabul edilir. Aksi halde ise, Eğer $IoU(B_1, B_2) < t$, ise tespit yanlış (false positive) olarak değerlendirilir (Sapkota ve ark., 2025).

Şekil 4.3' de, IoU değeri farklı olan örnekler gösterilmektedir. Tahmin edilen kutu ile gerçek kutu arasındaki mesafe büyükse, IoU değeri düşüktür ve bu da tahmin doğruluğunun zayıf olduğunu gösterir. Kutular birbirine ne kadar yakınsa ve aralarındaki örtüşme ne kadar fazlaysa, IoU değeri o kadar yüksek olur ve bu da modelin daha yüksek tespit doğruluğuna işaret eder.



Şekil 4.3. Birleşim üzerinden kesişim (IoU).

AP (Ortalama Kesinlik) ölçüsü, tahmin edilen fonların kalitesini değerlendirmek için "birleşmelerin kesişimi (IoU)" ilkesini ortaya koymaktadır. AP, modelin belirli bir sınıf için farklı güven eşiği (confidence threshold) değerleri altında elde ettiği kesinlik ve duyarlılık çiftlerinden oluşturulan eğrinin (Precision-Recall Curve) altında kalan alanı temsil eder ve bu değer denklem 4.5'teki gibi hesaplanır.

$$AP = \int_0^1 precision(Recall)d(Recall) \quad (4.5)$$

AP, modelin hem yüksek kesinlikte hem de yüksek duyarlılıkta tahminler yapabilme kabiliyetini gösterir. Bu açıdan AP, belirli bir sınıf için modelin genel başarımını yansıtan güçlü bir metriktir.

mAP (Ortalama Kesinliğin Ortalaması), veri kümesindeki tüm sınıflar için ayrı ayrı hesaplanan AP değerlerinin aritmetik ortalamasıdır. Bu metrik, bir modelin tüm sınıflar üzerindeki genel başarı düzeyini tek bir sayısal değerle özetlemeye olanak tanır. mAP değeri Denklem 4.6'da gösterildiği gibi hesaplanmaktadır.

$$mAP = \frac{1}{N} \sum_{k=1}^N AP_k \quad (4.6)$$

Burada N , veri kümesindeki toplam sınıf sayısını; AP_k ise her bir sınıfa ait ortalama doğruluk değerini temsil etmektedir.

Son olarak MITF performans ölçütü sınıflandırma modelinin tek bir görüntü üzerinde harcadığı çıkarım süresidir.





5. SONUÇLAR VE BULGULAR

5.1. Performans Ölçütü Değerlerinin Analizi

Bu çalışmada, YOLOv8, YOLOv11 ve YOLOv12 algoritmalarının beş farklı versiyonu (n, s, m, l, x) ve RT-DETR-L modeli, test veri seti kullanılarak kapsamlı bir şekilde analiz edilmiştir. Bu değerlendirmede, nesne algılama alanındaki yaygın performans ölçütleri (Kesinlik, duyarlılık, F1 ölçümü, mAP@0.5, mAP@0.5: 0.95 ve görüntü başına çıkarım süresi) kullanılarak gerçekleştirilmiştir. Amacımız, her modelin doğruluk ve yürütme süresi açısından performansını belirlemek ve farklı mimarilerin (CNN tabanlı YOLO modelleri ile Transformatör tabanlı RT-DETR modelleri) avantaj ve sınırlamalarını karşılaştırmaktır.

5.1.1. YOLOv8 sonuçları

Tablo 5.1, YOLOv8 mimarisi kullanılarak test veri kümesi üzerinde elde edilen performans ölçütü değerlerini göstermektedir. Tüm modeller değerlendirildiğinde, YOLOv8m %75,60 kesinlik, %79,79 duyarlılık, %77,64 F1-Skor, %81,93 mAP@0,5 ve %55,75 mAP@0,5: 0,95 ile en yüksek genel doğruluk performansına sahip olduğu görülmektedir. Ancak, işlem süresi açısından, YOLOv8n görüntü başına ortalama 4,1 ms ile en hızlı çalışan model; Ayrıca, diğer modellere göre, %80,37 duyarlılık ve %77,65 F1- score ile en yüksek performans elde edilmiştir. Buna karşılık YOLOv8m modeli küçük ölçekli modellere kıyasla daha düşük sonuçlar elde edilmiştir.

Daha büyük yapıya sahip olan YOLOv8l, %75,76 oranla diğer modellere göre en yüksek kesinlik değerine ulaşmıştır. Ancak 25,7 ms'lik işlem süresi, bu modelin daha güçlü donanımlarda tercih edilmesini gerektirmektedir.

Öte yandan, YOLOv8x modeli %57,56 kesinlik, %69,54 duyarlılık, %62,99 F1 skoru, %62,00 mAP@0.5 ve yalnızca %36,32 mAP@0.5:0.95 değeri ile diğer modellere göre daha düşük performans sergilemiştir. 49,3 ms'lik yüksek işlem süresi nedeniyle pratik kullanım açısından sınırlı bir verimlilik sunmaktadır.

Tablo 5.1. YOLOv8 sonuçları.

Model	Kesinlik	Duyarlılık	F1-Skor	mAP@0.5	mAP@0.5:0.95	MITF (ms/görüntü)
YOLOv8n	0,7510	0,8037	0,7765	0,8076	0,5558	4.1
YOLOv8s	0,7308	0,7807	0,7550	0,7986	0,5514	7.6
YOLOv8m	0,7560	0,7979	0,7764	0,8193	0,5575	16.6
YOLOv8l	0,7576	0,7763	0,7668	0,8024	0,5463	25.7
YOLOv8x	0,5756	0,6954	0,6299	0,6200	0,3632	49.3

Elde edilen bu sonuçlar, YOLOv8 mimarisinde, model seçiminde hız ve doğruluk arasında bir denge kurulması gerektiğini göstermektedir.

5.1.2. YOLO11 sonuçları

Tablo 5.2’de, YOLOv11 modelinin test veri seti üzerindeki performans ölçütü değerleri verilmektedir. Değerler incelendiğinde, YOLOv11s, genel olarak en yüksek performansı sergileyen model olarak öne çıkmaktadır. Bu versiyon, %77 kesinlik, %80 duyarlılık ve %78 F1-Skor değerleriyle, diğer tüm versiyonlardan daha iyi sonuçlar vermiştir. Özellikle mAP@0.5 değeri %80 ile en yüksek ortalama doğruluk değerine ulaşılmış ve mAP@0.5:0.95 değeri %57 ile daha zorlu IoU eşiklerinde de üstünlük sağlanmıştır. Bu durum, modelin hem genel nesne tespit başarısının yüksek olduğunu hem de farklı sınıflar ve nesne boyutları arasında dengeli bir performans gösterdiğini ifade etmektedir.

Diğer taraftan, YOLOv11n ve YOLOv11m versiyonları, sırasıyla 0.7712 ve 0.7708 değerleri ile F1-Skor ölçütü açısından birbirine oldukça yakın performanslar elde edilmiştir. Ayrıca YOLOv11n, modeli görüntü başına 3.5 ms ile en hızlı çalışan model olarak öne çıkmaktadır. YOLOv11n, daha hafif bir model olmasına rağmen, oldukça rekabetçi sonuçlar üretmiştir.

YOLOv11l ve YOLOv11x modelleri ise diğer versiyonlara göre daha düşük F1-Skor, mAP değerleri ve en yüksek işlem süresine sahiptir. Bu durum, modelin daha büyük mimarisinin her zaman daha iyi sonuç vermediğini ve parametre artışına rağmen tespit doğruluğunda belirgin bir kazanım sağlamadığını ortaya koymaktadır. Özellikle YOLOv11x, en düşük mAP@0.5:0.95 değeri olan 0.5478 ile en alt sırada yer almıştır.

Tablo 5.2. YOLOv11 sonuçları.

Model	Kesinlik	Duyarlılık	F1-Skor	mAP@0.5	mAP@0.5:0.95	MITF (ms/görüntü)
YOLOv11n	0.7477	0.7963	0.7712	0.8116	0.5554	3.5
YOLOv11s	0.7718	0.8025	0.7869	0.8303	0.5733	6.8
YOLOv11m	0.7386	0.8060	0.7708	0.8073	0.5505	14.3
YOLOv11l	0.7407	0.7631	0.7517	0.7949	0.5487	19.1
YOLOv11x	0.7340	0.7862	0.7592	0.8047	0.5478	39.0

Deneysel sonuçlara göre, YOLOv11s versiyonu YOLOv11 mimarileri arasında, hem genel doğruluk hem de dengeli performans açısından en etkili model olarak değerlendirilebilmektedir.

5.1.3. YOLOv12 sonuçları

Tablo 5.3.'te YOLOv12 modelinin sonuçları yer almaktadır. YOLOv12m modeli %76 kesinlik, %77 F1-Skor, %81,36 mAP@0.5 ve %61 mAP@0.5: 0.95 değerleri ile genel performans açısından diğer modellerden daha iyi performans göstermiştir. Bu sonuçlar, modelin tespit doğruluğu ve genelleme kalitesi arasında etkili bir denge sağlama yeteneğini göstererek, cilt lezyonu tespitinde mükemmel performansını ortaya koymaktadır. Buna karşılık, YOLOv12n modeli %82 duyarlılık ve 6.1 ms'lik yüksek işlem süresine ulaşarak çok sayıda nesneyi tespit etmek için mükemmel bir sonuç elde etmiştir.

YOLOv8s' de olduğu gibi, küçük bir boyut olmasına rağmen YOLOv12s'de de sonuçlar düşük çıkmıştır.

Karmaşık mimarisine rağmen YOLOv12l modeli ile beklenenden daha düşük performans elde edilmiştir. Bu sonuçlar YOLOv12l modelinin küçük nesneler için daha az uygun olduğunu göstermektedir.

Tablo 5.3. YOLOv12 sonuçları.

Model	Kesinlik	Duyarlılık	F1-Skor	mAP@0.5	mAP@0.5:0.95	MITF (ms/görüntü)
YOLOv12n	0.7032	0.8286	0,7608	0.8105	0.5429	6.1
YOLOv12s	0.7620	0.7901	0,7758	0.8072	0.5570	10.4
YOLOv12m	0.7695	0.7890	0,7791	0.8136	0.6110	20.8
YOLOv12l	0.7388	0.7296	0,7341	0.7936	0.5398	29.0

YOLOv12n hızlı tepki süresine sahiptir ve YOLOv12m ile %81 oranında karşılaştırılabilir performans göstermiştir. YOLOv12n, gömülü veya mobil sistemler gibi kaynak kısıtlı ortamlarda çalışan gerçek zamanlı uygulamalar için idealdir.

5.1.4. RT-DETR

Tablo 5.4'te RT-DETR Large modelinin performans ölçütü değerleri sunulmaktadır. Model ile %70 kesinlik, 69% duyarlılık, %69,91 mAP@0.5 ve %42 ile mAP@0.5:0.95 elde edilmiştir. Değerler incelendiğinde, YOLOv8x modeli hariç tüm modeller arasında RT-DETR modeli ile en düşük değerlerin elde edildiği görülmektedir. Dolayısıyla RT-DETR large modelinin büyük boyutu olmasına rağmen düşük kesinlik, duyarlılık, mAP@0.5 ve mAP@0.5:0.95 değerleri elde edilmiştir. Bu sonuçlar RT-DETR Large modelinin cilt kanseri teşhisi uygulamalarında uygun olmadığını göstermektedir.

Tablo 5.4. RT-DETR sonuçları.

Model	Kesinlik	Duyarlılık	F1-Skor	mAP@0.5	mAP@0.5:0.95	MITF (ms/görüntü)
RT-DETR Large	0.7007	0.6975	0,6991	0.6416	0.4293	23.1

5.2. Karmaşıklık Matrisi ve Kayıp Fonksiyonlarının Analizi

Bu bölümde çalışmada geliştirilen her bir model için elde edilen karmaşıklık matrisleri ve kayıp fonksiyonları incelenmektedir.

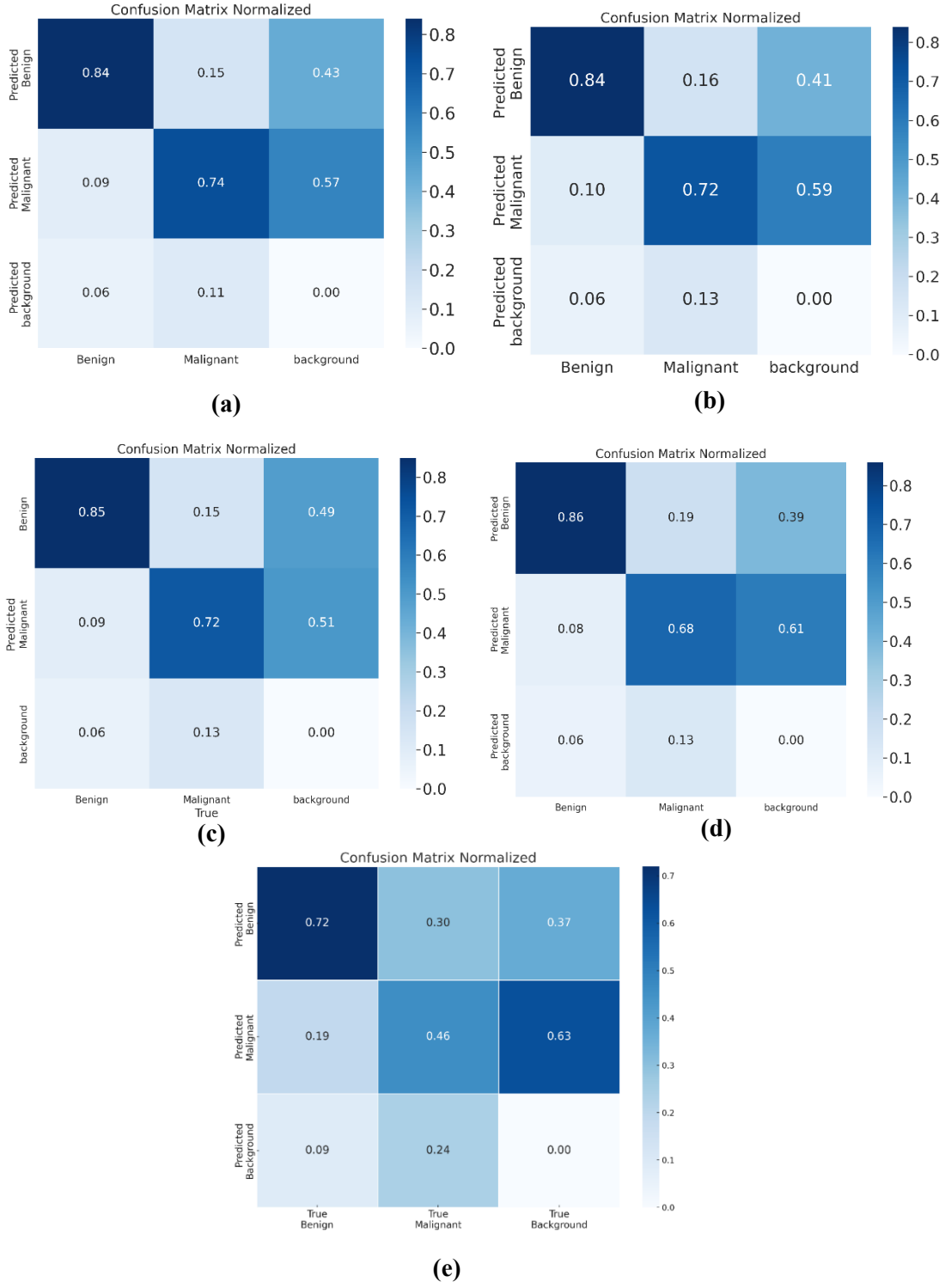
5.2.1. YOLOv8

5.2.1.1. YOLOv8 karmaşıklık matris

Şekil 5.1'de YOLOv8 modeli versiyonlarına modeline ait karmaşıklık matrisleri gösterilmektedir. Şekilde iyi huylu, kötü huylu ve arka plan olmak üzere üç sınıf bulunmaktadır. Yolov8n (a) küçük hacimli olmasına rağmen, %84 doğruluk oranı ile iyi huylu sınıfını oldukça iyi tahmin ederken, %74 doğruluk oranı ile kötü huylu sınıfında ve %43 doğruluk oranı ile arka plan sınıfında karışıklık yaşamaktadır. YOLOv8s (b) modeli, bir önceki versiyona benzer bir performans elde edilmiştir. Kötü huylu %72 sınıfta daha fazla yanlış sınıflama söz konusudur ve bu örneklerin çoğu arka plan olarak değerlendirilmiştir. Bunun sebebi ise düşük kontrastlı veya küçük boyutlu lezyonların model tarafından ayırt edilemediğini düşündürmektedir. Yolov8m modeli iyi huylu sınıfta %85 oranla başarılı bir sonuç verirken, %72 doğruluk oranına

sahip kötü huylu sınıfı, önceki versiyonları gibi %51 oranını ile arka plan sınıfına karışma eğilimi göstermektedir. Ancak arka plan oranı daha düşüktür, bu da sınıf ayrımını daha iyi yaptığına işaret eder. En büyük parametreye sahip model YOLOv8x (e) modeli olmasına rağmen, şekilde kötü huylu sınıf başarımında en düşük performansı göstermektedir. İyi huylu sınıfta bile ciddi karışıklıklar yaşanmıştır.

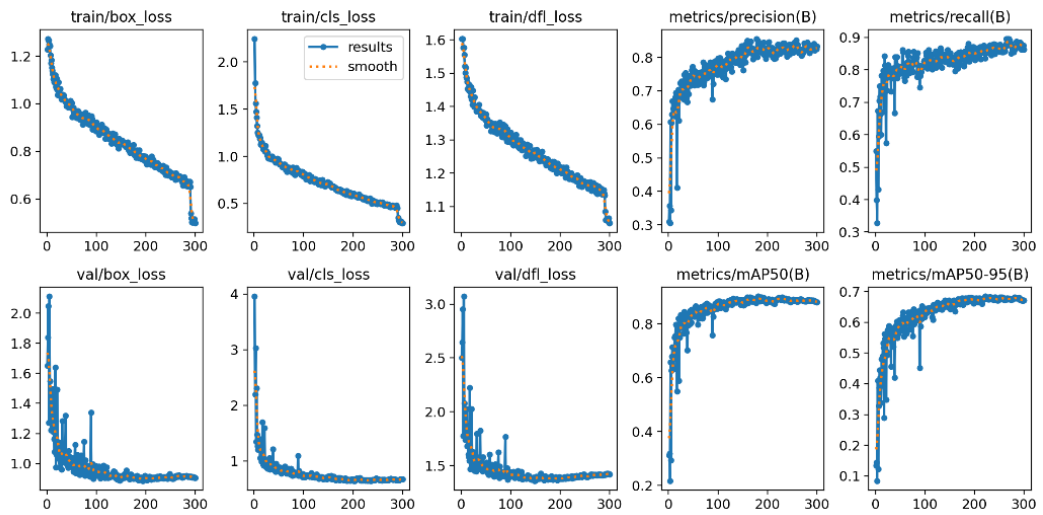




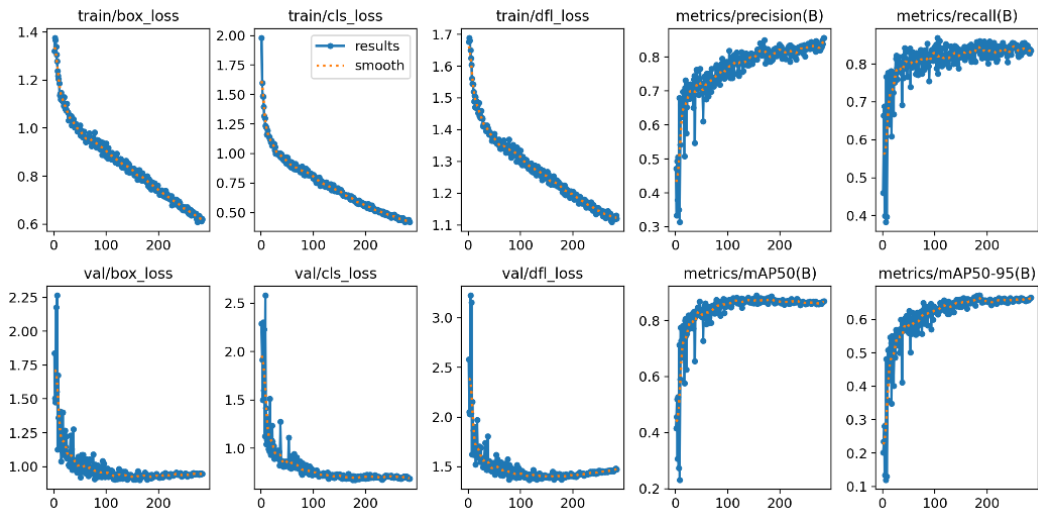
Şekil 5.1. YOLOv8 modellerine ait karmaşıklık matrisleri (a) YOLOv8n, (b) YOLOv8s, (c) YOLOv8m, (d) YOLOv8l, (e) YOLOv8x

5.2.1.2. YOLOv8 kayıp fonksiyonunun analizi

Şekil 5.2’de YOLOv8n modelinin, Şekil 5.3’te YOLOv8s modelinin eğitim süresi boyunca kaydedilen kayıp fonksiyonu grafikleri gösterilmektedir. Grafikler incelendiğinde, YOLOv8s ve YOLOv8n modelleri eğitim dönemi boyunca benzer sonuçlar elde ettiği görülmektedir. YOLOv8n, sabitlenmeden önce erken aşamalarda bazı dalgalanmalar yaşarken, doğrulama kaybının kararlılığında hafif bir iyileşme gözlemlenmiştir. Kesinlik ve duyarlılık değerleri kademeli olarak 0,85’in üzerine çıkmış; mAP@0.5 yaklaşık 0,85’e ve mAP@0.5:0.95 yaklaşık 0,67’ye ulaşmıştır. YOLOv8s için ise mAP metrikleri kademeli olarak iyileşmeler göstermiş ve sırasıyla yaklaşık 0,84 ve 0,63’e ulaşmıştır. Sonuçlar, her iki modelin de model boyutu ve doğruluk arasında iyi bir denge sunduğunu göstermektedir.

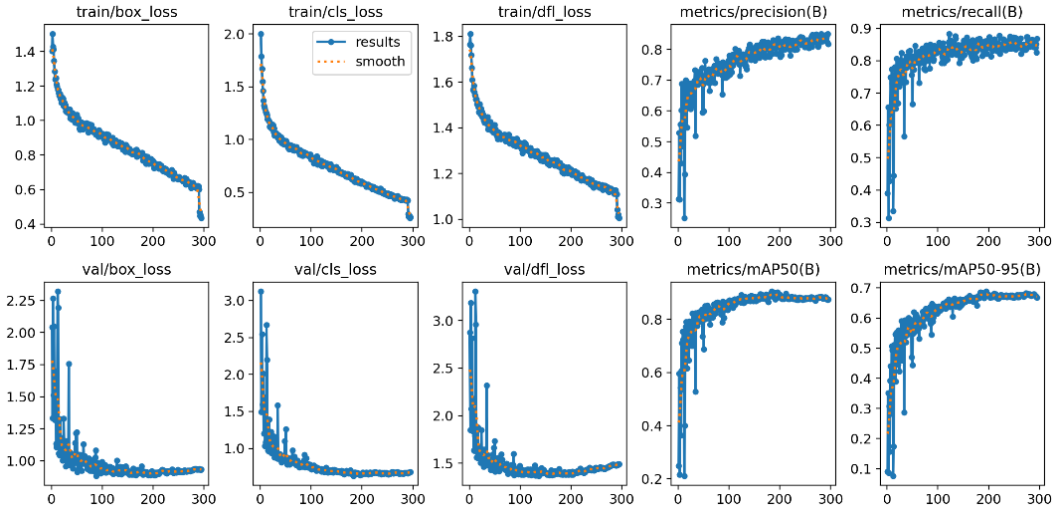


Şekil 5.2. YOLOv8n modelinin kayıp fonksiyonu.



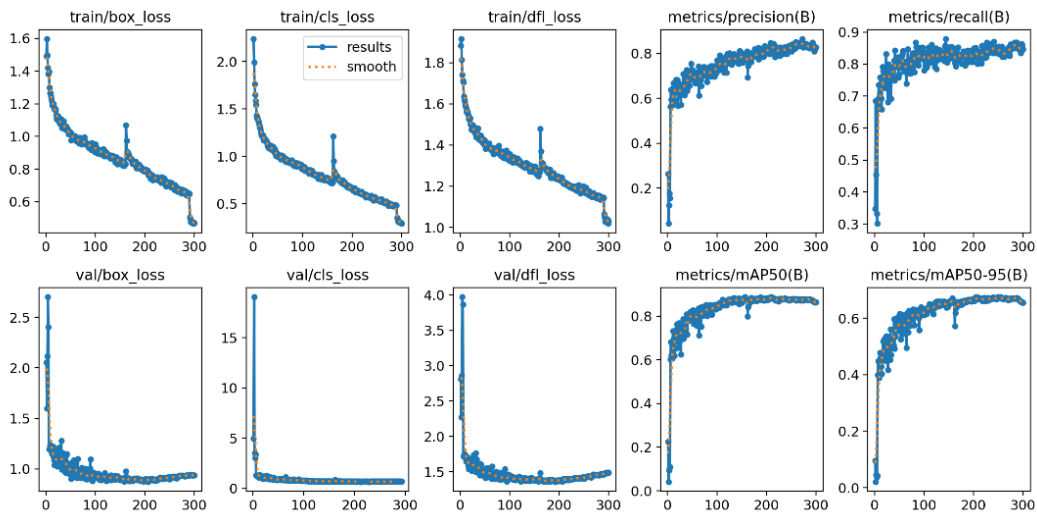
Şekil 5.3. YOLOv8s modelinin kayıp fonksiyonu.

Şekil 5.4'te YOLOv8m modelinin kayıp fonksiyonu grafiği verilmektedir. Model kararlılık ve tahmin doğruluğu açısından olağanüstü bir performans göstermiştir. Tüm kayıplar istikrarlı bir şekilde azalmış ve doğrulama eğrileri daha küçük modellere göre daha az değişkenlik göstermiştir. $mAP@0.5$ yaklaşık 0.85 ve $mAP@0.5:0.95$ yaklaşık 0.68 ile kaydedilen en yüksek sonuçlar arasında olmuştur. Kesinlik ve duyarlılık değerleri de 0.88'i aşmış, bu da modelin iyi genelleme yapma yeteneğini göstermiştir.



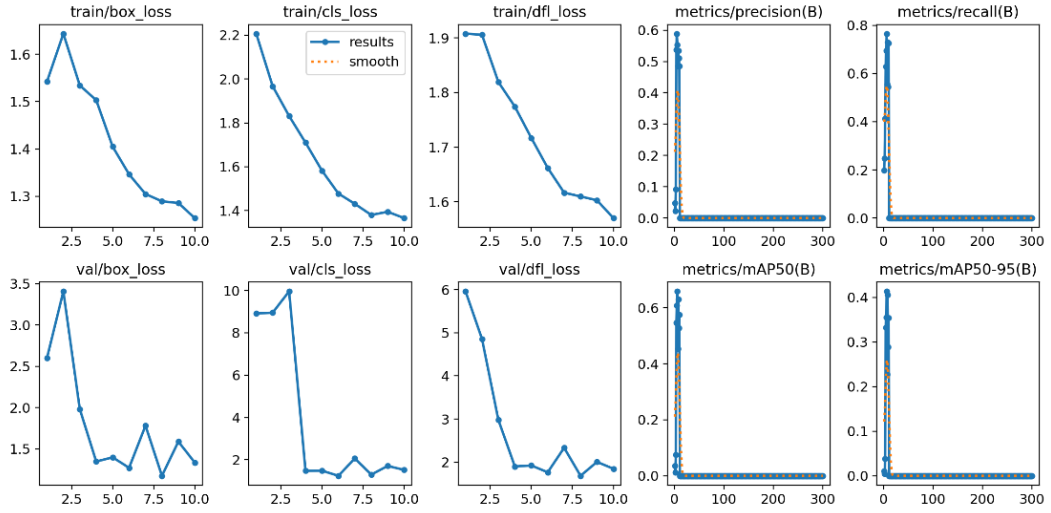
Şekil 5.4. YOLOv8m modelinin kayıp fonksiyonu.

Şekil 5.5'te YOLOv8l modelinin kayıp fonksiyonu grafiği sunulmaktadır. Model nispeten yüksek kayıp seviyelerinde eğitime başlamış ve kademeli olarak azalmıştır. Son mAP değerleri 0.83 ve 0.63 civarına ulaşmış ve kesinlik ve duyarlılık değerleri 0.85'i aşmıştır. Ancak bazı kayıplar dalgalanma göstermiştir.



Şekil 5.5. YOLOv8l modelinin kayıp fonksiyonu.

Şekil 5.6 YOLOv8x modeline ait kayıp fonksiyonu grafiğini göstermektedir. Beklentilerin aksine, YOLOv8x modeli verilerin verimli bir temsilini öğrenmede başarısız olmuştur. Kayıp eğrileri önemli ölçüde azalmamış ve performans ölçümleri (mAP, kesinlik, duyarlılık) eğitim süresi boyunca sıfıra yakın seviyelerde kalmıştır.

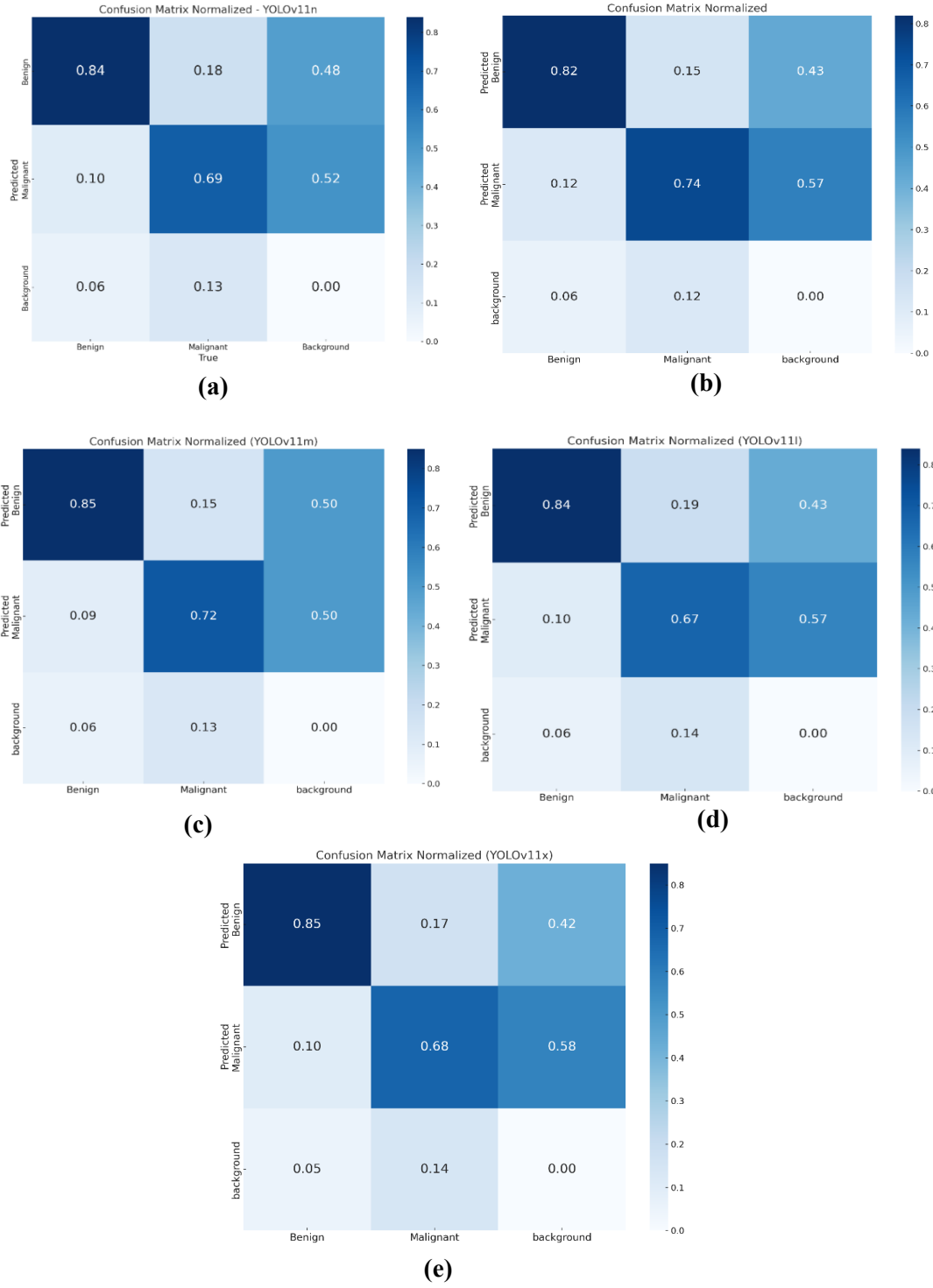


Şekil 5.6. YOLOv8x modelinin kayıp fonksiyonu.

5.2.2. YOLOv11

5.2.2.1. YOLOv11 karmaşıklık matris

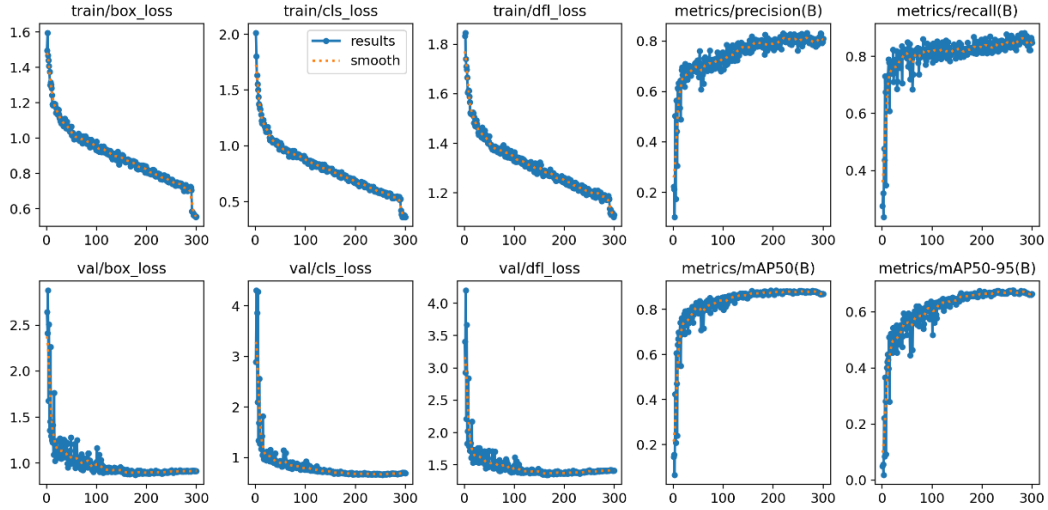
Şekil 5.7'de YOLOv11 model ailesine ait küçükten büyüğe beş farklı modelin karmaşıklık matrislerini göstermektedir. Şekil 5.7 (a)'da YOLOv11n modeli, iyi huylu sınıflandırmada %84'e ulaşan yüksek doğruluk göstermiştir. Ancak, kötü huylu sınıfı sınıflandırmada %69 doğruluk elde edilmiş ve %52 oranında arka plan ile karışıklık göstermiştir. Şekil 5.7 (b)'de YOLOv8s modeli ise, önceki modele göre benzer bir performans göstermiştir. Ancak kötü huylu sınıfı %74'e yükselterek iyi bir gelişme sağlamıştır. Bununla birlikte, aynı sınıfın arka planla karışıklık oranı %57'lik bir doğrulukla daha fazla yanlış sınıflandırma gözlemlenmiştir. Öte yandan, Şekil 5.7 (c)'de YOLOv11m modeli, iyi huylu sınıfta %85'lik bir doğruluk, kötü huylu sınıfta %72'lik bir doğruluk ve arka plan örtüşmesinde azalma sağlayarak diğer modellere kıyasla etkileyici bir başarı ve iyi bir denge elde etmiş, daha iyi sınıf ayrımcılığını göstermiştir. Şekil 5.7 (d)'de YOLOv11l modeli ve Şekil 5.7 (e)'de YOLOv11x modeli, iyi huylu sınıftaki önceki modellere benzer şekilde iyi bir sınıflandırma doğruluğu elde edilmiştir ancak daha büyük boyutlarına rağmen, kötü niyetli sınıfı sınıflandırmada birçok hatayla ve yüksek bir arka plan örtüşmesiyle karşılaşmıştır.



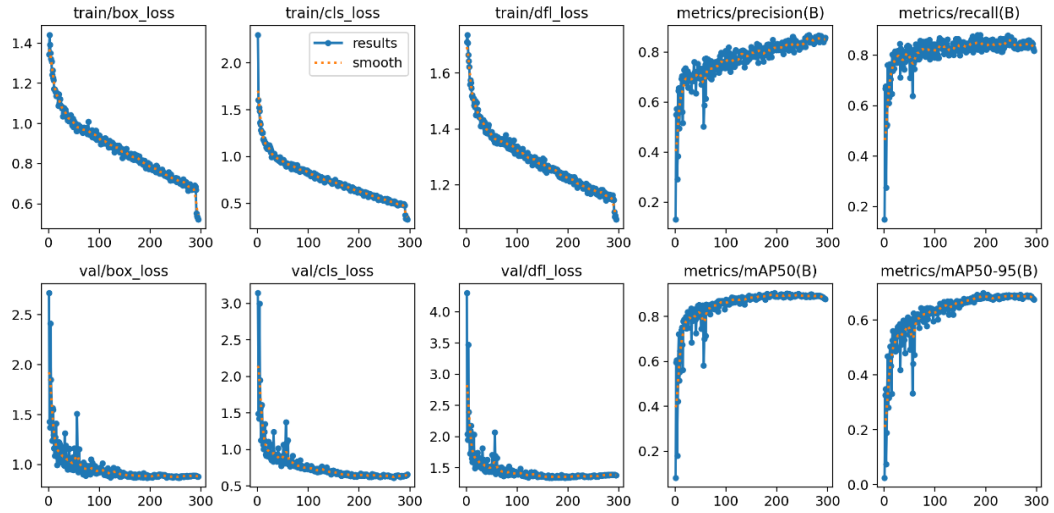
Şekil 5.7. YOLOv11 modellerine ait karmaşıklık matrisleri (a) YOLOv11n, (b) YOLOv11s, (c) YOLOv11m, (d) YOLOv11l, (e) YOLOv11x

5.2.2.2. YOLOv11 kayıp fonksiyonunun analizi

YOLOv11 modelleri hem eğitim hem de doğrulama için kayıp değerlerinde önemli iyileştirmeler göstermektedir. YOLOv11n modeli ve YOLOv11s modelinin kayıp fonksiyonu grafikleri sırasıyla Şekil 5.8 ve Şekil 5.9,'da sunulmaktadır. Grafikler, modellerin kayıp fonksiyonunda bozulma olmadan istikrarlı bir şekilde öğrendiğini göstermektedir. Ayrıca, kayıp eğitim ve doğrulama arasında dengelendiğinden aşırı uyum belirtisi de göstermektedir.



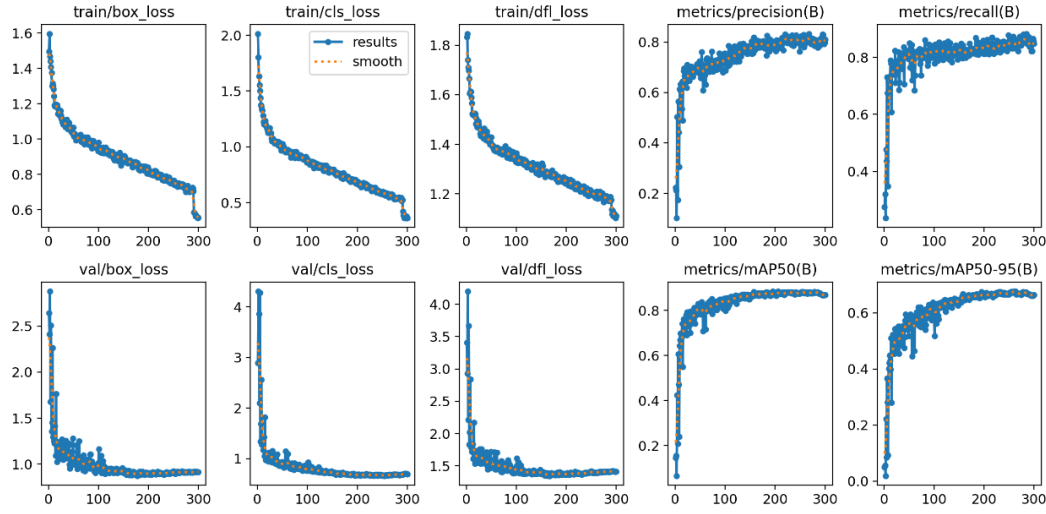
Şekil 5.8. YOLOv11n modelinin kayıp fonksiyonu.



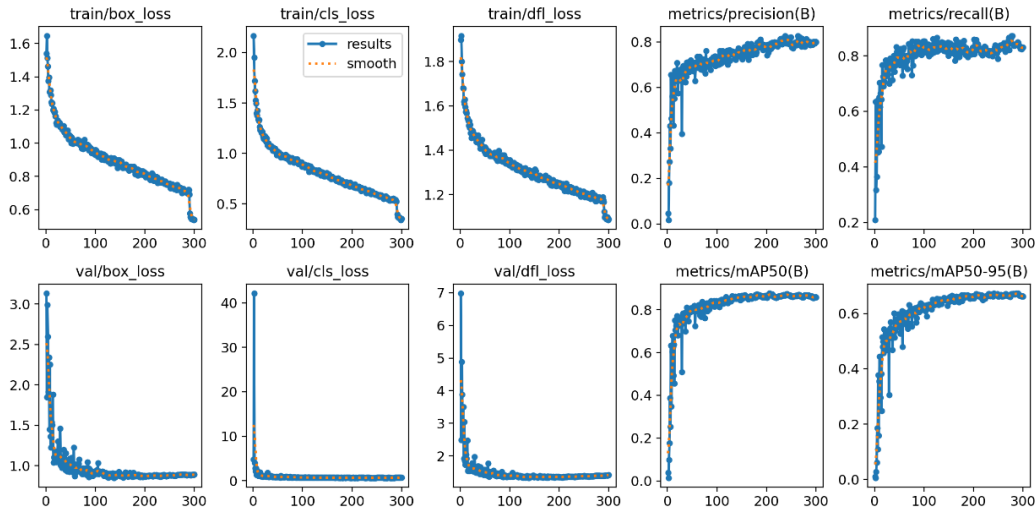
Şekil 5.9. YOLOv11s modelinin kayıp fonksiyonu.

Sırasıyla Şekil 5.10 ve Şekil 5.11'de verilen YOLOv11m ve YOLOv11l modellerinin kayıp fonksiyonları incelendiğinde, eğitim hızı ve doğruluk arasında iyi bir denge olduğu görülmektedir. Kayıp sürekli azalırken, kesinlik ve duyarlılık değerleri

yükselmiş ve dönem sayısı ile artış göstermiştir. Bu da modellerin verilerdeki çeşitli desenlere iyi yanıt verdiğini göstermiştir.

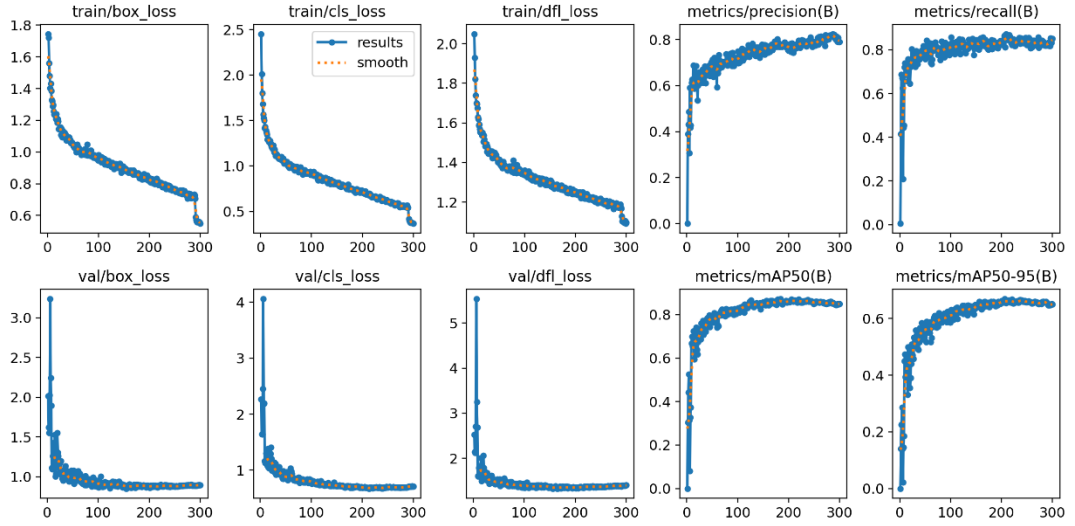


Şekil 5.10. YOLOv11m modelinin kayıp fonksiyonu.



Şekil 5.11. YOLOv11l modelinin kayıp fonksiyonu.

Şekil 5.12’de YOLOv11x modelinin kayıp fonksiyonu grafikleri gösterilmektedir. Grafikler eğitim sürecinin oldukça başarılı ilerlediğini ifade etmektedir. Eğitim ve doğrulama kayıpları düzenli olarak azalmış, bu da modelin istikrarlı öğrenme sağladığını göstermektedir. Kesinlik ve duyarlılık değerleri zamanla artarak yaklaşık 0.9 seviyelerine ulaşmıştır. YOLOv11x modelinin, yüksek doğrulukta tahminler yapabildiğini ve genelleme yeteneğinin güçlü olduğunu göstermektedir.

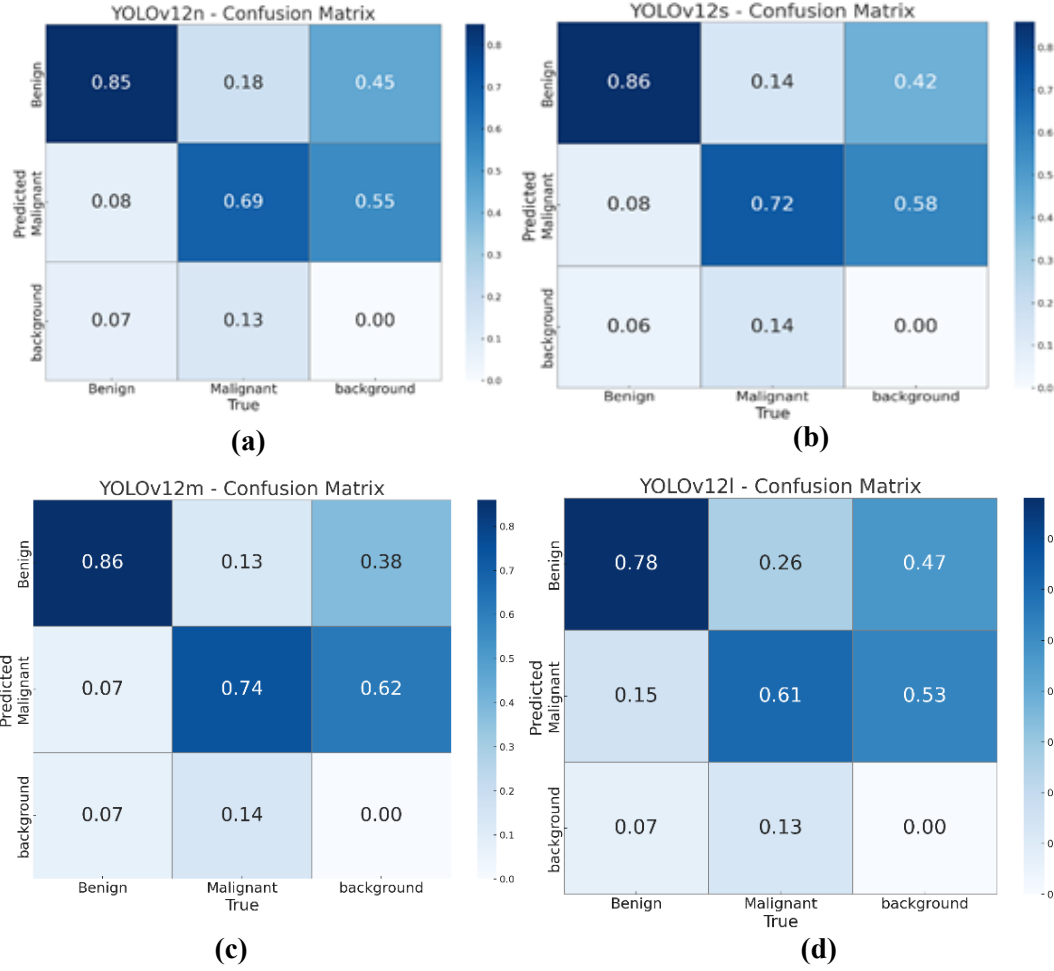


Şekil 5.12. YOLOv11x modelinin kayıp fonksiyonu.

5.2.3. YOLOv12

5.2.3.1. YOLOv12 karmaşıklık matris

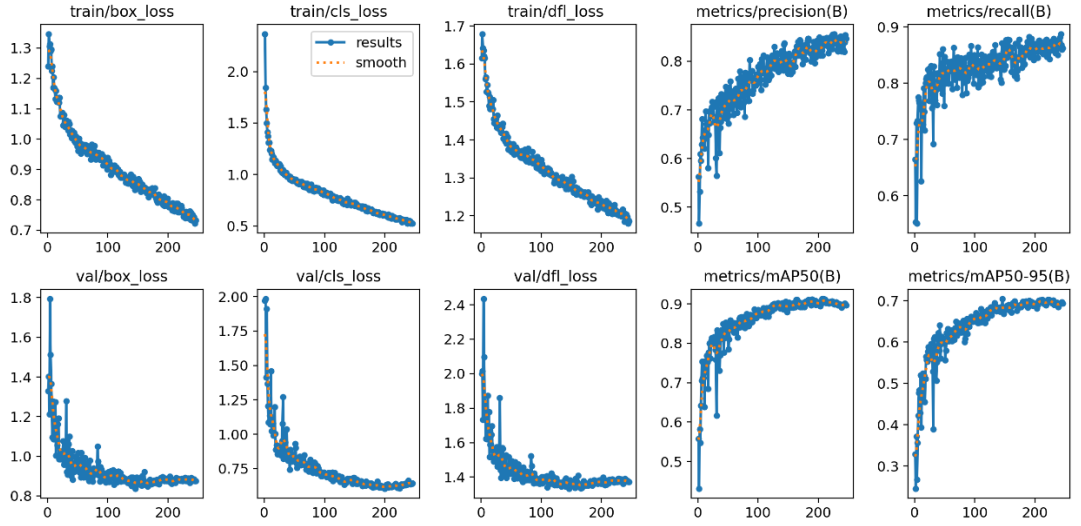
Şekil 5.13’ te YOLOv12 modellerinin karmaşıklık matrisleri verilmektedir. YOLOv12n (Şekil 5.13 (a)), YOLOv12s (Şekil 5.13 (b)) ve YOLOv12m (Şekil 5.13 (c)) modellerinin iyi huylu sınıfı sınıflandırmada güçlü bir performans gösterdiği ve %85 ile yüksek bir doğruluk oranına ulaştığı anlaşılmaktadır. YOLOv12n kötü huylu sınıfta %69 doğruluk ve arka plan olarak yanlış sınıflandırmada %55’lik bir örtüşme sorunu yaşamıştır. YOLOv12s, önceki modele kıyasla kötü huylu sınıflandırmada %72 doğruluk oranla yükseltme sağlamıştır. Şekil 5.13(c)’de karmaşıklık grafiği gösterilen YOLOv12m modeli kötü huylu sınıfı %62 oran ile arka plan olarak tahmin ederken önceki modellere göre daha yüksek karışıklılık elde edilmiştir. Şekil 5.13(d)’de YOLOv12l, daha büyük boyutuna rağmen, diğer modellere göre daha iyi sonuçlar elde edememiş, iyi huylu sınıfı %78 oran ile doğru sınıflandırırken kötü huylu sınıf tahmininde %61 doğruluk oranı ile düşük performans elde edilmiştir. Kötü huylu sınıfta %53’lük bir arka plan karışıklığı oranı elde etmiştir ve arka planın ele alınmasında kusurlu bir dengeyi yansıtmıştır.



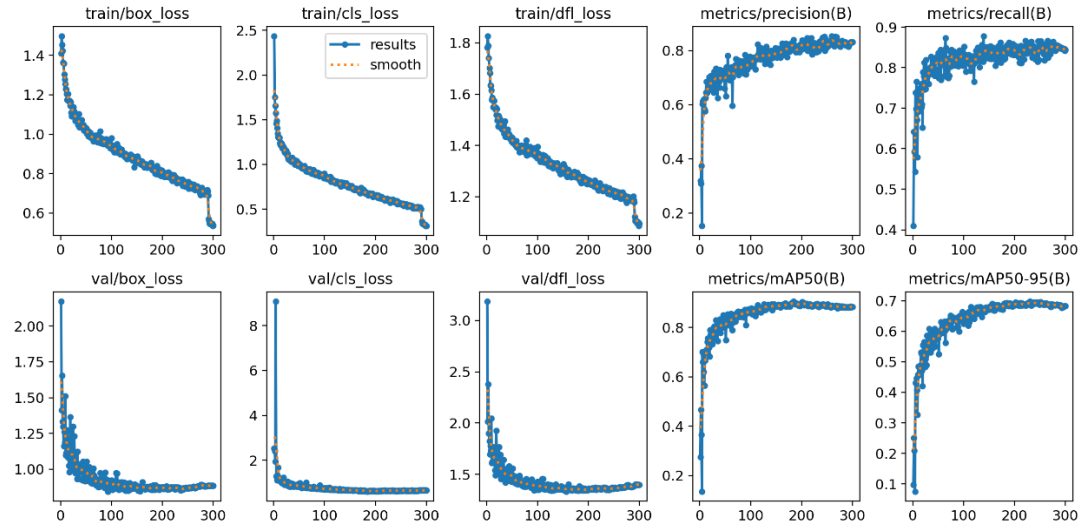
Şekil 5.13. YOLOv12 modellerine ait karmaşıklık matrisleri (a) YOLOv12n, (b) YOLOv12s, (c) YOLOv12m, (d) YOLOv12l, (e) YOLOv12x

5.2.3.2. YOLOv12 kayıp fonksiyonun analizi

Şekil 5.14 'te YOLOv12n modelinin ve Şekil 5.15'te YOLOv12s modelinin kayıp fonksiyonu grafikleri sunulmaktadır. İki model, ilk dönemlerde daha hızlı öğrenme hızları göstermiştir. YOLOv12s, özellikle kesinlik ve mAP performans ölçütünde YOLOv12n'den daha yüksek doğruluk elde etmiştir. YOLOv12n, hız ve çeviklik gerektiren görevler için idealdir ancak ince ayrıntıları tespit etmedeki performansı daha düşüktür. YOLOv12s, boyut ve performans arasında uygun bir denge sunmuştur.

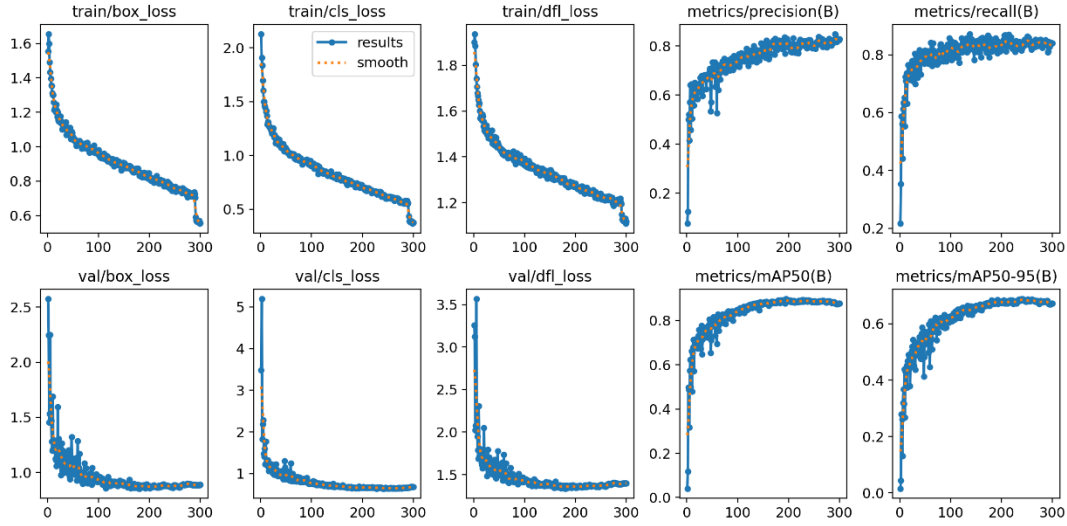


Şekil 5.14. YOLOv12n modelinin kayıp fonksiyonu.

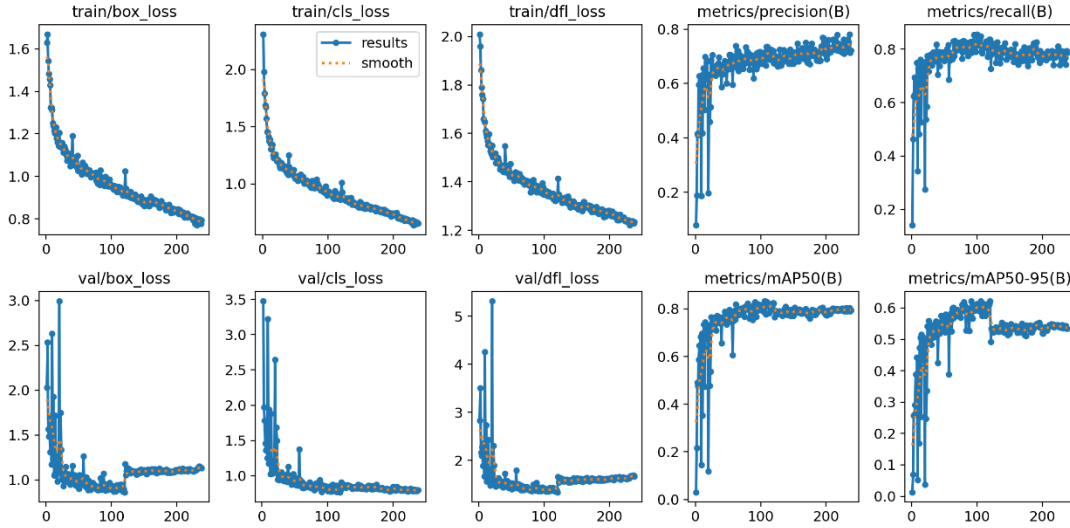


Şekil 5.15. YOLOv12s modelinin kayıp fonksiyonu.

Şekil 5.16’da gösterilen YOLOv12m modelinin kayıp fonksiyonu eğitim ve doğrulama arasındaki kayıp eğrilerinde mükemmel bir yakınsama göstermiştir. Kesinlik ve duyarlılık eğrileri kararlı ve yüksektir. Şekil 5.17’de sunulan YOLOv12l modelinin kayıp fonksiyonu ise doğruluk ve genelleme metriklerinde daha iyi performans göstermiştir ancak kararlılığa ulaşması daha uzun sürmüştür. Her ikisi modelin de aşırı uyum olmadan kararlı eğriler göstermiş, kayıp ve mAP’de kademeli iyileştirmeler olmaktadır.



Şekil 5.16. YOLOv12m modelinin kayıp fonksiyonu.

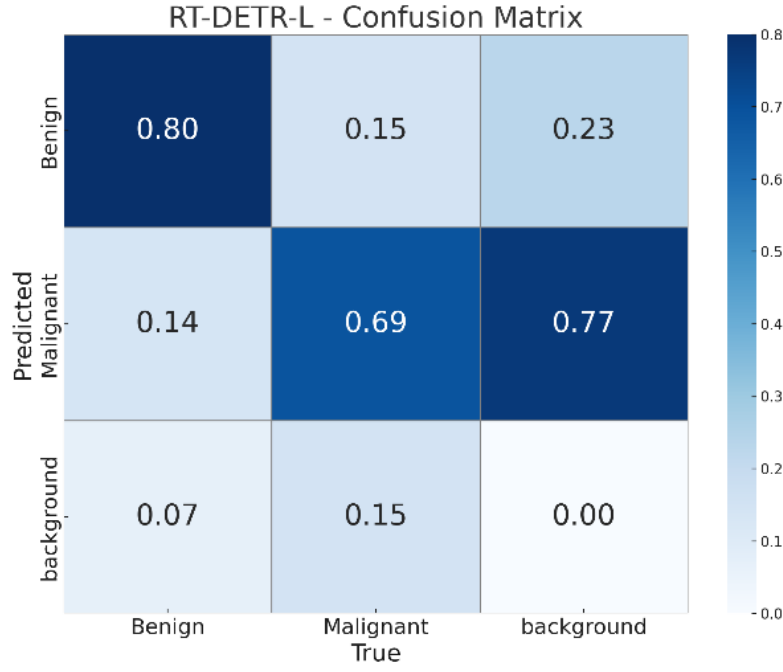


Şekil 5.17. YOLOv12l modelinin kayıp fonksiyonu.

5.2.4. RT-DETR

5.2.4.1. RT-DETR algoritmanın karmaşıklık matris

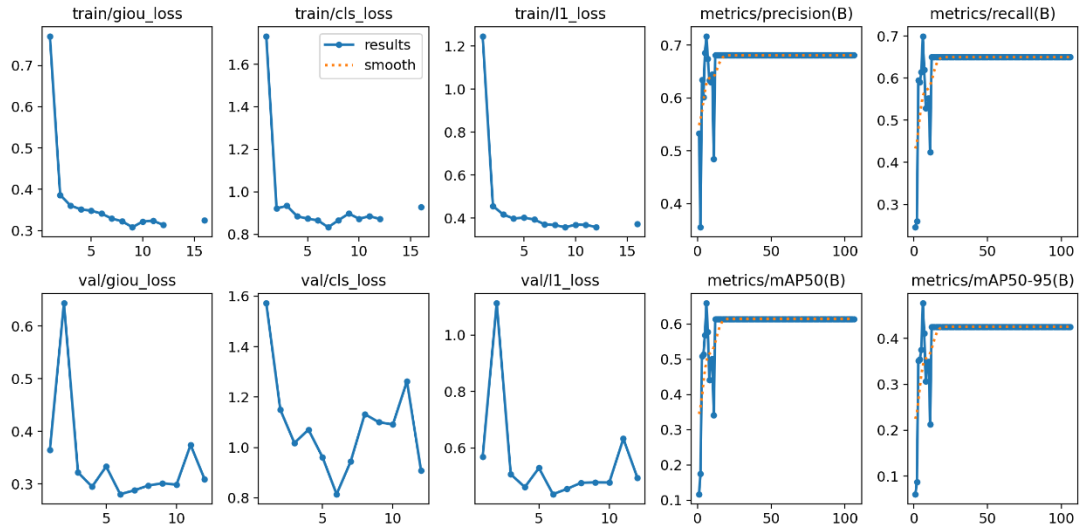
Şekil 5.18’de sunulan RT-DETR modeli karmaşıklık matrisinde iyi huylu sınıftaki örnekler, kötü huylu sınıftaki örneklere göre daha iyi sınıflandırılmıştır. Kötü huylu sınıfta %69 doğruluk oranı elde edilmiştir. Kötü huylu sınıf bazında arka planla %77 karışıklık oranı elde edilmiştir. Dolayısıyla model kötü huylu sınıflandırmada başarısız bulunmuştur. İyi huylu sınıfta ise düşük oranda ciltle karışıklık elde edilmiş ve diğer algoritmalarla kıyasla özellikle YOLOv12m modeli %38 sınıflandırma yapılırken RT-DETR modeli ile % 23 oranı elde edilmiştir.



Şekil 5.18. RT-DETR-L modelinin karmaşıklık matris

5.2.4.2. RT-DETR algoritmanın kayıp fonksiyonunun analiz

Şekil 5.19’da, RT-DETR modelin genel kayıp fonksiyonunun özetini göstermektedir. Eğitim süresinde kayıp eğrileri aşağı yönlü bir eğilim göstermektedir; bu, eğitim sırasında kayıpların en aza düştüğü anlamına gelmektedir. Bu, kutu doğruluğu ve nesne yerelleştirmesinde önemli bir iyileşme olduğunu göstermektedir. Doğrulama sırasında kayıp eğilimi dalgalanmış olsa da genel olarak düşüş eğilimi göstermiştir. Metrik eğrileri yukarı yönlü bir eğilim göstermektedir; bu da modelin performansının eğitim süresi boyunca iyileştiği anlamına gelmektedir.



Şekil 5.19. RT-DETR algoritmanın kayıp fonksiyonu.



6. SONUÇ VE ÖNERİLER

Dünya genelinde cilt kanserinin giderek artan yaygınlığı, aşırı güneşe maruz kalma ve solaryum kullanımı gibi kanserle ilişkilendirilen yaygın günlük alışkanlıklar ile açıklanmaktadır. Ayrıca tıp merkezlerinde yapılan taramaların yüksek maliyetli olması ve teşhisin çoğu zaman gecikmesi, etkili teknolojik çözümlere olan ihtiyacı artırmaktadır. Bu noktada yapay zekâ, özellikle derin öğrenme algoritmaları, erken teşhis sürecini desteklemede ve hastalık riskini azaltmada önemli bir rol üstlenmektedir.

Bu çalışmada, cilt kanseri tespiti için görüntüler üzerinden nesne tespiti amacıyla YOLOv8, YOLOv11, YOLOv12 ve RT-DETR olmak üzere dört modern derin öğrenme algoritması kullanılmıştır. Her bir YOLO versiyonu, farklı boyutlarda beş modelden (n, s, m, l, x) oluşmaktadır. Tüm YOLO modelleri ve RT-DETR algoritmasına ait large modeli kullanımı ile eğitim gerçekleştirilmiştir.

Kullanılan her bir modelin kendine özgü yapısal iyileştirmeleri bulunmakta olup, bu iyileştirmeler elde edilen sonuçlara doğrudan yansımıştır. Örneğin, YOLOv8 sürümünde yer alan çapa (anchor) bağımsız yapı, YOLOv11 sürümünde ise boyun (neck) bölümündeki mimari iyileştirmeler, C3K2 bloğu ve C2PSA modülü gibi bileşenler, YOLO serisinin daha verimli ve çok yönlü bir nesne tespit çözümü haline gelmesine katkı sağlamaktadır. YOLOv12 sürümünde ise, Bölgesel Dikkat Mekanizması (A^2) ve R-ELAN yapıları geliştirilmiştir. RT-DETR algoritması, Hibrit kodlayıcı ve AIFI ile yüksek düzeyde semantik temsiller oluşturulurken, CCFF modülü, farklı seviyelerden gelen özellikleri birleştirerek geliştirilmiştir.

Bu çalışmada, ISIC verisetine dayanarak geliştirilen ve Kaggle sitesinden elde edilebilen yaklaşık 3297 cilt lezyon görüntüsünden oluşan veri seti kullanılmıştır. YOLOv formatında eğitebilmek için sınırlayıcı kutular çizilmiş ve bilgiler .txt formatında kaydedilmiştir. Geliştirilen modellerin her biri farklı katman ve parametre sayısı ile eğitilmiş ve her modelin hiperparametreleri eğitim sürecinde optimize edilmiştir. Aşırı uyumu önlemek için erken durdurma kullanılmış ve eğitim süresi model türüne bağlı olarak 4,4 saat ile 141,7 saat arasında değişmiştir.

Cilt lezyon görüntülerinin, gerçekleştirilen performans değerleri nesne tespiti alanındaki yaygın performans değerlendirme ölçütleri (kesinlik, duyarlılık, F1 Skor, mAP@0.5, mAP@0.5: 0.95 ve MITF) kullanılmıştır. YOLOv modellerinde, YOLOv8m %81,93 mAP, YOLOv11s %83,03 mAP, YOLOv12m %81,36 mAP değerleriyle en yüksek performans elde edilmiştir. RT-DETR-L modeli ile %64,16 mAP değerine ulaşılmıştır. Elde edilen sonuçlara göre, YOLOv11 modeli diğer modellere kıyasla %83 mAP değerinin yanı sıra %77,18 kesinlik, %78,69 F1-Skor, 3,5 MITF ve değerleriyle de en yüksek performansı göstermiştir. YOLOv12 modeli performans değerlendirme ölçütlerinde %82,86 duyarlılık, %61,10 mAP@0.5:0.95 olarak diğer modeller arasında ikinci en iyi performansı sağlayan model olmuştur.

Sonuç olarak, küçük hacme sahip modellerin etkili sonuçlar gösterdiği değerlendirilebilmektedir. Özellikle Karmaşıklık Matrislerinde iyi huylu sınıfın sınıflandırılmasında yüksek performans elde edilmiştir. Kötü huylu sınıfta ise arka planla karışıklık yaşanmıştır ve düşük performans elde edilmesine sebep olmuştur. Bunun sebebi de düşük kontrastlı veya çok küçük boyutlu lezyonların model tarafından ayırt edilemediğini düşündürmektedir.

Elde ettiğimiz sonuçlar, kullanılan veri sayısı ile karşılaştırıldığında, önceki çalışmalara kıyasla en iyi ve en yüksek sonuçlar arasında yer almaktadır.

Gelecekte, daha büyük ve çeşitli veri kümeleri üzerinde eğitim yapılması ve daha gelişmiş modellerin karşılaştırılması planlanmaktadır. Ayrıca cilt kanserinin güneş ışığı mı yoksa solaryum kullanımı mı ile daha fazla ilişkili olduğunu araştıracağız. Güneş ve solaryum maruziyetini ölçen temel değişkenleri toplayacağız ve hangi nedenin risk katkısının daha yüksek olduğunu açık biçimde göstereceğiz. Çevrimiçi tespit sistemi, görüntü yükleme, otomatik ön işleme ve anlık sonuç sunan basit bir web arayüzü geliştirilecektir. Donanımın etki analizi, Modellerin tahmin doğruluğu ve hızına donanım gücünü artırmanın etkisi incelenecektir. Bu sayede, gerçek ortamlar için daha yüksek doğruluk ve daha kararlı sonuçlar elde edilmesi hedeflenmektedir.

KAYNAKLAR

- A. Soni and A. Rai, "YOLO for Medical Object Detection (2018–2024)," in *2024 IEEE 3rd International Conference on Electrical Power and Energy Systems (ICEPES)*, June 21–22, 2024. DOI: 10.1109/ICEPES60647.2024.10653506.
- Abbasi NR, Shaw HM, Rigel DS, et al. Early diagnosis of cutaneous melanoma: revisiting the ABCD criteria. *JAMA*. 2004; 292(22): 2771-6. Epub 2004/12/09. PMID:15585738. <http://dx.doi.org/10.1001/jama.292.22.2771>
- Aishwarya, N., Prabhakaran, M. K., Debebe, F. T., Reddy, M. S. S. A., & Pranavee, P. (2023). Skin Cancer diagnosis with Yolo Deep Neural Network. *Procedia Computer Science*. <https://doi.org/10.1016/j.procs.2023.03.083>
- Albahli, S. (2025). A robust YOLOv8-based framework for real-time melanoma detection and segmentation with multi-dataset training. *Diagnostics*, 15(6), 691. <https://doi.org/10.3390/diagnostics15060691>
- Alif, M. A. R. (2024). *YOLOv11 for vehicle detection: Advancements, performance, and applications in intelligent transportation systems* (arXiv:2410.22898v1 [cs.CV]). arXiv. <https://arxiv.org/abs/2410.22898>
- Alif, M. A. R., & Hussain, M. (2024, June 14). *YOLOv1 to YOLOv10: A comprehensive review of YOLO variants and their application in the agricultural domain* (arXiv:2406.10139v1) [Computer Vision and Pattern Recognition]. arXiv. <https://doi.org/10.48550/arXiv.2406.10139>
- Alif, M. A. R., & Hussain, M. (2025, February 20). *YOLOv12: A breakdown of the key architectural features*. arXiv. <https://arxiv.org/abs/2502.14740>
- AlSadhan, N. A., Alamri, S. A., Ben Ismail, M. M., & Bchir, O. (2024). Skin Cancer Recognition Using Unified Deep Convolutional Neural Networks. *Cancers*, 16(7), 1246. <https://doi.org/10.3390/cancers16071246>
- Alwahaibi N, et al. Mini review on skin biopsy: AI and CNN in dermatopathology. *Frontiers in Medicine*, 2025
- American Cancer Society (2023). Skin Cancer Facts & Statistics.
- Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection.
- Borisyyuk, F., Gordo, A., & Sivakumar, V. (2018). Rosetta. Proceedings of the 24th ACM SIGKDD.
- Carranza-García, M., Torres-Mateo, J., Lara-Benítez, P., & García-Gutiérrez, J. (2021). On the Performance of One-Stage and Two-Stage Object Detectors in Autonomous Vehicles Using Camera Data. *Remote Sensing*, 13(1), 89. <https://doi.org/10.3390/rs13010089>
- D. Reis, J. Hong, J. Kupec, and A. Daoudi, "Real-Time Flying Object Detection with YOLOv8," Georgia Institute of Technology.

- Daghrir, J., Tlig, L., Bouchouicha, M., & Sayadi, M. (2020, September). Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach. In *2020 5th international conference on advanced technologies for signal and image processing (ATSIP)* (pp. 1-5). IEEE.
- Deni Sutaji, Oktay Yildiz, "YOLOv8's Head-Layer Performance Comparison for Skin Cancer Detection," *Proceeding of International Conference of Advance Transportation Engineering and Applied Social Science*, November 2023. DOI: 10.46491/icateas.v2i1.
- Dhruba Datta, Harsh Prakash, Priya Singh, "Skin Cancer Detection with Edge Devices Using YOLOv7 Deep CNN," *Springer Lecture Notes in Electrical Engineering*, November 2023. DOI: 10.1007/978-981-99-6550-2_5.
- Dildar, M., Akram, S., Irfan, M., Khan, H. U., Ramzan, M., Mahmood, A. R., ... & Mahnashi, M. H. (2021). Skin cancer detection: a review using deep learning techniques. *International journal of environmental research and public health*, 18(10), 5479.
- Evi Fatimatur Rusydiyah, Dian Candra Rini Novitasari, Faris Mushlihull Amin, Anisa Nur Azizah, Huda Febrianto Nurrohman, Dina Zatusiva Haq, Elen Riswana Safila Putri, "Skin Cancer Diagnosis System on Object Detection Using Various CNN YOLOv5 in Android Mobile," *Conference Paper*.
- F. Demir, "Derin Öğrenme Tabanlı Yaklaşımla Kötü Huylu Deri Kanserinin Dermatoskopik Görüntülerden Saptanması," *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 2021. DOI: 10.35234/fumbd.900170.
- Firnhaber, J. M. (2020). Basal cell and cutaneous squamous cell carcinomas: Diagnosis and treatment. *American Family Physician*, 102(6), 339–346.
- Friedman, R. J., Rigel, D. S., & Kopf, A. W. (1985). Early detection of malignant melanoma: The role of physician examination and self-examination of the skin. *CA: A Cancer Journal for Clinicians*, 35(3), 130–151.
- Garbe, C., Peris, K., Hauschild, A., Saiag, P., Middleton, M., Bastholt, L., Grob, J.J., Malvehy, J., Newton-Bishop, J., Stratigos, A.J., Pehamberger, H., & Eggermont, A.M. (2016). *Diagnosis and treatment of melanoma: European consensus-based interdisciplinary guideline – Update 2016*. *European Journal of Cancer*, 63, 201–217. <https://doi.org/10.1016/j.ejca.2016.05.005>
- Ghiasi, G., Lin, T.-Y., & Le, Q. V. (2019). *NAS-FPN: Learning Scalable Feature Pyramid Architecture for Object Detection*. Google Brain.
- Girshick, R. (2015). Fast R-CNN. *ICCV*. <https://arxiv.org/abs/1504.08083v2>
- Gouda, W., Us Sama, N., Al-Waakid, G., Humayun, M., & Jhanjhi, N. Z. (2022). Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning. *Healthcare*, 10(7), 1183. <https://doi.org/10.3390/healthcare10071183>
- Han, G., Wang, R., et al. (2024). *Enhancing Semantic and Spatial Information in YOLOv8*. SSRN. <https://ssrn.com/abstract=4797816>
- Hasya, H. F., Nuha, H. H., & Abdurrohman, M. (2021). Real Time-based Skin Cancer Detection System using Convolutional Neural Network and YOLO. *2021 4th International Conference of Computer and Informatics Engineering (IC2IE)*. <https://doi.org/10.1109/IC2IE53219.2021.9649224>

- He, W., Zhang, Y., Xu, T., An, T., Liang, Y., & Zhang, B. (2025). Object Detection for Medical Image Analysis: Insights from the RT-DETR Model. *arXiv preprint*, arXiv:2501.16469.
- Helfand, M., Mahon, S. M., Eden, K. B., Frame, P. S., & Orleans, C. T. (2001). Screening for skin cancer. *American Journal of Preventive Medicine*, [https://doi.org/10.1016/S0749-3797\(01\)00258-6](https://doi.org/10.1016/S0749-3797(01)00258-6)
- <https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign/data>
- Huang, J., & Wang, H. (2024). Small Object Detection by DETR via Information Augmentation and Adaptive Feature Fusion. *arXiv preprint*, arXiv:2401.08017.
- Huda, D. N., Romdoni, M. R., Safitri, L., Winarni, A., & Rahman, A. (2024). Real-time Detection Transformer (RT-DETR) of Ornamental Fish Diseases with YOLOv9 using CNN Algorithm. *Journal of Applied Informatics and Computing*, 8(2), 463–471. <https://doi.org/10.30871/jaic.v8i2.8561>
- Hum, Y. C., Tan, H. R., Tee, Y. K., Yap, W. S., Tan, T. S., Salim, M. I. M., & Lai, K. W. (2022). The development of skin lesion detection application in smart handheld devices using deep neural networks. *Multimedia Tools and Applications*, 81, 41579–41610. <https://doi.org/10.1007/s11042-021-11013-9>
- Hussain, M. (2024). *YOLOv5, YOLOv8 and YOLOv10: The Go-To Detectors for Real-time Vision*. arXiv:2407.02988.
- ISIC, "ISIC Archive : The International Skin Imaging Collaboration: Melanoma Project," ISIC, 5 Jan 2016. [Online]. Available: <https://isic-archive.com/#>. [Accessed 20 Jan 2016].
- J. Farooq, M. Muaz, K. K. Jadoon, N. Aafaq, and M. K. A. Khan, "An improved YOLOv8 for foreign object debris detection with optimized architecture for small objects," *Multimedia Tools and Applications*, vol. 83, 2023. DOI: 10.1007/s11042-023-17838-w.
- J. Terven, D.-M. Córdova-Esparza, and J.-A. Romero-González, "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," *Machine Learning and Knowledge Extraction*, vol. 5, no. 4, pp. 1680–1716, Nov. 2023. DOI: 10.3390/make5040083.
- Jang, W.S.; Kim, S.; Yun, P.S.; Jang, H.S.; Seong, Y.W.; Yang, H.S.; Chang, J.S. Accurate detection for dental implant and peri-implant tissue by transfer learning of faster R-CNN: A diagnostic accuracy study. *BMC Oral Health* 2022, 22, 591
- Jegham, N., Koh, C. Y., Abdelatti, M., & Hendawi, A. (2025, February 25). *YOLO Evolution: A Comprehensive Benchmark and Architectural Review of YOLOv12, YOLO11, and Their Previous Versions* (v3). arXiv. <https://arxiv.org/abs/2411.00201>
- Jocher, G.; Chaurasia, A.; Qiu, J. Ultralytics YOLO. 2023. Available online: <https://github.com/ultralytics/ultralytics> (accessed on 20 October 2024).
- Jones, M., & Brown, L. (2021). Deep Learning in Dermatology: A Comprehensive Review. *Artificial Intelligence in Medicine*, 18(2), 67-89.

- Jun, E. L. T., Tham, M.-L., & Kwan, B.-H. (2024). *A comparative analysis of RT-DETR and YOLOv8 for urban zone aerial object detection*. IEEE. <https://doi.org/10.1109/I2CACIS61270.2024.10649836>
- Karbouj, B., Topalian-Rivas, G. A., & Krüger, J. (2024). Comparative Performance Evaluation of One-Stage and Two-Stage Object Detectors for Screw Head Detection and Classification in Disassembly Processes. *Procedia CIRP*. <https://doi.org/10.1016/j.procir.2024.01.077>
- Kestek, E., Aktan, M. E., & Akdogan, E. (2023). Skin Lesion Classification Using Convolutional Neural Network and ABCD Rule. *Turkish Journal of Modeling and Computational Sciences*, 3(1), Article 1249300. <https://doi.org/10.47000/tjmcs.1249300>
- Khanam, R., & Hussain, M. (2024, October 23). *YOLOv11: An Overview of the Key Architectural Enhancements*. arXiv. <https://doi.org/10.48550/arXiv.2410.17725>
- Khanam, R., & Hussain, M. (2025, April 16). *A review of YOLOv12: Attention-based enhancements vs. previous versions*. arXiv. <https://arxiv.org/abs/2504.11995>
- Kliniec, K., Tota, M., Zalesińska, A., Łyko, M., & Jankowska-Konsur, A. (2023). Skin cancer risk, sun-protection knowledge and behavior in athletes—A narrative review. *International Journal of Environmental Research and Public Health*, PMID: 37444391.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems (NIPS)*.
- Labelimg (2022), <https://github.com/HumanSignal/labelImg>. Accessed 03 Dec 2024
- Lattoofi, N. F., Al-sharuee, I. F., Kamil, M. Y., & Obaid, A. H. (2020). Melanoma Skin Cancer Detection Based on ABCD Rule. *2019 International Conference on Computer Applications and Information Engineering (CAS)*. <https://doi.org/10.1109/CAS47993.2019.9075465>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Li, H., et al. (2018). *Pyramid Attention Network for Semantic Segmentation*. arXiv:1805.10180.
- Li, P., Zhao, T., & Wang, Y. (2022). Medical Image Analysis Using Deep Learning Techniques. *International Journal of Biomedical Imaging*, 24(6), 1123-1135.
- Lin, T.-Y., et al. (2017). *Feature Pyramid Networks for Object Detection*. CVPR.
- Liu, C., Tao, Y., Liang, J., Li, K., & Chen, Y. (2018, December). Object detection based on YOLO network. In *2018 IEEE 4th information technology and mechatronics engineering conference (ITOEC)* (pp. 799-803). IEEE.
- Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., & Pietikäinen, M. (2020). Deep Learning for Generic Object Detection: A Survey. *International Journal of Computer Vision*, 128, 261–318. <https://doi.org/10.1007/s11263-019-01247-4>
- Liu, W., et al. (2016). SSD: Single Shot Multibox Detector. ECCV.

- Loshchilov, I., & Hutter, F. (2017). Fixing weight decay regularization in adam. *arXiv preprint arXiv:1711.05101*, 5.
- Lucas, A. S., Chung, E., Marchetti, M. A., & Marghoob, A. A. (2016). Early diagnosis of cutaneous melanoma: revisiting the ABCD criteria. *JAMA*, 292(22), 2771. <https://doi.org/10.1001/jama.292.22.2771>
- Lucas, A. S., Chung, E., Marchetti, M. A., & Marghoob, A. A. (2016). A guide for dermatology nurses to assist in the early detection of skin cancer. *Journal of Nursing Education and Practice*, 6(10), 71. <https://doi.org/10.5430/jnep.v6n10p71>
- Lv, W., Zhao, Y., Chang, Q., Huang, K., Wang, G., & Liu, Y. (2024). RT-DETRv2: Improved Baseline with Bag-of-Freebies for Real-Time Detection Transformer. *arXiv preprint*, arXiv:2407.17140.
- M. A. Albahar, "Skin Lesion Classification Using Convolutional Neural Network With Novel Regularizer," *Ibn Rushd College for Management Science, Abha, Saudi Arabia*, 2019. DOI: 10.1109/ACCESS.2019.2906241.
- M. A. Riyadi, A. Ayuningtias, and R. R. Isnanto, "Detection and Classification of Skin Cancer Using YOLOv8n," in *2024 11th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 2024. DOI: 10.1109/EECSI63442.2024.10776505
- M. G. Ragab, S. J. Abdulkadir, A. Muneer, A. Alqushaibi, E. H. Sumiea, R. Qureshi, S. M. AISelwi, and H. Alhussian, "A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023)," *IEEE Access*, 2024. DOI: 10.1109/ACCESS.2024.3386826.
- Malo, D. C., Khan, M. M., Rahman, M. M., & Mahbub, J. (2022). Skin Cancer Detection using Convolutional Neural Network. *2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)*. <https://doi.org/10.1109/CCWC54503.2022.9720751>
- Mao, M., & Hong, M. (2024). YOLO object detection for real-time fabric defect inspection in the textile industry: A review of YOLOv1 to YOLOv11. *Sensors*, 25(7), 2270. <https://doi.org/10.3390/s25072270>
- Mingji, Y., Xu, R., Yang, C., & Wu, H. (2024). Hybrid-DETR: A Differentiated Module-Based Model for Object Detection in Remote Sensing Images. *Electronics*, 13(24), 5014. <https://doi.org/10.3390/electronics13245014>
- Mishraa, N. K., & Celebi, M. E. (2016). An overview of melanoma detection in dermoscopy images using image processing and machine learning. *arXiv preprint*, arXiv:1601.07843. <https://doi.org/10.48550/arXiv.1601.07843>
- Moura, N., Veras, R., Aires, K., Machado, V., Silva, R. M., Araujo, F., & Mai'la, C. (2018). ABCD rule and pre-trained CNNs for melanoma diagnosis. *Multimedia Tools and Applications*, 1–20.
- Mupparaju, S., Thotakura, S. R., & Reddy, V. R. Ch. (2024). A Review on YOLOv8 and its Advancements. In *Data Intelligence and Cognitive Informatics* (pp. 529–545). https://doi.org/10.1007/978-981-99-7962-2_39
- N. Aishwarya, G.S. Yaythish Kannaa, Karthik Seemakurthy, "Skin Cancer Diagnosis with YOLO Deep Neural Network," *Procedia Computer Science*, vol. 219, pp. –, 2023. DOI: 10.1016/j.procs.2023.03.083.

- N. Jegham, C. Y. Koh, M. Abdelatti, and A. Hendawi, "Evaluating the Evolution of YOLO (You Only Look Once) Models: A Comprehensive Benchmark Study of YOLO11 and Its Predecessors," *arXiv preprint arXiv:2411.00201v1*, Oct. 31, 2024.
- Narayanan, D. L., Saladi, R. N., & Fox, J. L. (2010). Review: Ultraviolet radiation and skin cancer. *International Journal of Dermatology*, <https://doi.org/10.1111/j.1365-4632.2010.04474.x>
- Nersisson, R., Iyer, T. J., Raj, A. N. J., & Rajangam, V. (2021). A Dermoscopic Skin Lesion Classification Technique Using YOLO-CNN and Traditional Feature Model. *Arabian Journal for Science and Engineering*, 46, 9797–9808. <https://doi.org/10.1007/s13369-021-05571-1>
- Nguyen, T. (2025). Improve Underwater Object Detection through YOLOv12 Architecture and Physics-informed Augmentation. *arXiv preprint arXiv:2506.23505*.
- Niu, C., Han, D., Han, B., & Wu, Z. (2024). *SAR-LtYOLOv8: A lightweight YOLOv8 model for small object detection in SAR ship images*. *Computers, Materials & Continua*, 78(3), xx–xx. <https://doi.org/10.32604/csse.2024.056736>
- Önder, Ö., & Karan, Y. (2024). Çay ve eğrelti otunun YOLOv5 ve YOLOv8 algoritmaları ile karşılaştırmalı tespiti. *Recep Tayyip Erdoğan Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 5(1), 74–88. <https://doi.org/10.53501/rteufemud.1402167>
- Patel, K., Sharma, R., & Kumar, N. (2023). The Role of AI in Early Diagnosis of Skin Cancer. *Machine Learning in Medicine*, 27(4), 567-582.
- Paul Paul Tsoi. YOLO11: The cutting-edge evolution in object detection — a brief review of the latest in the yolo series. <https://medium.com>, October 2024. Accessed: 2024-10-17.
- Prechelt, L. (2002). Early stopping-but when?. In *Neural Networks: Tricks of the trade* (pp. 55-69). Berlin, Heidelberg: Springer Berlin Heidelberg.
- R. Khanam and M. Hussain, "YOLOV11: An Overview of the Key Architectural Enhancements," *arXiv preprint arXiv:2410.17725v1*, Oct. 23, 2024.
- Radovanovic, M., Petrovski, A., Behlic, A., Chaari, M. Z., Hashimov, E. G., Fellner, R., & Agbeyangi, A. (2024). Unleashing Autonomous Forces: Integrating AI-DRIVEN Drones in Modern Military Strategy. *Системи управління, навігації та зв'язку. Збірник наукових праць*, 3(77), 55-69.
- Raimondi, S., Suppa, M., & Gandini, S. (2020). Melanoma epidemiology and sun exposure. *Acta Dermato-Venereologica*, PMID: 32346751. <https://doi.org/10.2340/00015555-3491>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *CVPR*, 779-788.
- Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6), 1137-1149.

- Rodrigue Bogne Tchema, Marios Nestoros, Anastasis C. Polycarpou, "Computer-Aided Classification of Skin Cancer Based on the YOLO Algorithm," *Proceedings of the 2024 13th International Conference on Modern Circuits and Systems Technologies (MOCASST)*, IEEE, 2024. DOI: 10.1109/MOCASST61810.2024.10615645.
- S. Ćirković and N. Stanić, "Application of the YOLO algorithm for Medical Purposes in the Detection of Skin Cancer," in *Proceedings of the 2024 Conference on Computer Sciences and Information Technology*, Kragujevac, Serbia, Sep. 2024. DOI: 10.46793/TIE24.083C.
- Sahafi, A., Koulaouzidis, A., & Lalinia, M. (2024). Polypoid lesion segmentation using YOLO-V8 network in wireless video capsule endoscopy images. *Diagnostics*, 14(5), 474.
- Saleh Albahli, Nudrat Nida, Aun Irtaza, Muhammad Haroon Yousaf, Muhammad Tariq Mahmood, "Melanoma Lesion Detection and Segmentation Using YOLOv4-DarkNet and Active Contour," *IEEE Access*, November 2020. DOI: [Available online November 12, 2020].
- Salomi, M., Daram, G., & Harshitha, S. S. (2024). Early Skin Cancer Detection Using CNN-ABCD Rule Based Feature Extraction Classification and K-Means Clustering Algorithm Through Android Mobile Application. *2024 International Conference on Emerging Trends in Information Technology and Engineering (IC-ETITE)*. <https://doi.org/10.1109/IC-ETITE58242.2024.10493828>
- Santos, C., Aguiar, M., Welfer, D., & Belloni, B. (2022). A new approach for detecting fundus lesions using image processing and deep neural network architecture based on YOLO model. *Sensors*, 22(17), 6441.
- Sapkota, R., Meng, Z., Churuvija, M., Du, X., Ma, Z., & Karkee, M. (2025, February 25). *Comprehensive performance evaluation of YOLOv12, YOLOv11, YOLOv10, YOLOv9 and YOLOv8 on detecting and counting fruitlet in complex orchard environments* (Version 7) [Preprint]. arXiv. <https://arxiv.org/abs/2407.12040>
- Sazak, H., & Kotan, M. (2025). *Otomatik Kan Hücresi Tespiti ve Mikroskopik Görüntülerde Sınıflandırma: YOLOv11 ve Optimizasyonlu Ağırlıklarla*. *Diagnostics*, 15(1), 22. <https://doi.org/10.3390/diagnostics15010022>
- Senana, E. M., & Jadhav, M. E. (2021). Analysis of dermoscopy images by using ABCD rule for early detection of skin cancer. *Global Transitions Proceedings*. <https://doi.org/10.1016/j.gltp.2021.01.001>
- Shao, D., Jiang, J., Ma, L., Lai, H., & Yi, S. (2024). Real-time medical lesion screening: Accurate and rapid detectors. *Journal of Real-Time Image Processing*, 21(4), Article 134. <https://doi.org/10.1007/s11554-024-01512-x>
- Skin Cancer Foundation. Skin Cancer Facts & Statistics, 2023 (Skin Cancer Facts & Statistics).
- Smith, J., Doe, A., & Green, R. (2022). Advances in Early Skin Cancer Detection. *Journal of Oncology Research*, 15(3), 123-135.
- Solawetz, J., & Francesco. (2023). What is YOLOv8? The Ultimate Guide. Yayın tarihi: 30 Nisan 2023.

- Tabrizchi, H., Parvizpour, S., & Razmara, J. (2023). An Improved VGG Model for Skin Cancer Detection. *Neural Processing Letters*, 55, 3715–3732. <https://doi.org/10.1007/s11063-022-10927-1>
- Thanh, D. N. H., Prasath, V. B. S., Hieu, L. M., & Hien, N. N. (2019). Melanoma Skin Cancer Detection Method Based on Adaptive Principal Curvature, Colour Normalisation and Feature Extraction with the ABCD Rule. *Journal of Digital Imaging*. <https://doi.org/10.1007/s10278-019-00316-x>
- Tian, Y., Ye, Q., & Doermann, D. (2025, February 18). *YOLOv12: Attention-centric real-time object detectors*. arXiv. <https://arxiv.org/abs/2502.12524>
- Tousch AM, Herbin S, Audibert JY (2012) Semantic hierarchies for image annotation: A survey. *Patt Recog* 45(1):333–345
- Tuncer, T., Barua, P. D., Tuncer, I., Dogan, S., & Acharya, U. R. (2024). A lightweight deep convolutional neural network model for skin cancer image classification. *Applied Soft Computing*. <https://doi.org/10.1016/j.asoc.2024.111794>
- Ultralytics (2023). YOLOv8: Real-Time Object Detection and Classification. Technical Report. (Ultralytics, 2023)
- Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). YOLOv10: Real-Time End-to-End Object Detection. arXiv preprint arXiv:2405.14458.
- Wang, C. Y., Liao, H. Y. M., Wu, Y. H., Chen, P. Y., Hsieh, J. W., & Yeh, I. H. (2020). CSPNet: A new backbone that can enhance learning capability of CNN. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops* (pp. 390-391).
- Wang, C.-Y., & Liao, H.-Y. M. (2024). YOLOv9: Learning what you want to learn using programmable gradient information.
- Wang, L., & Lee, C. (2021). Performance Evaluation of Deep Learning Models in Medical Image Classification. *Medical AI Journal*, 19(5), 98-115.
- Wang, S., Jiang, H., Yang, J., Ma, X., Chen, J., Li, Z., & Tang, X. (2024). Lightweight tomato ripeness detection algorithm based on the improved RT-DETR. *Frontiers in Plant Science*, 15, Article 1415297. <https://doi.org/10.3389/fpls.2024.1415297>
- Wang, S., Xia, C., Lv, F., & Shi, Y. (2024). RT-DETRv3: Real-time End-to-End Object Detection with Hierarchical Dense Positive Supervision. *arXiv preprint*, arXiv:2409.08475.
- Wang, Y., Xu, S., Wang, P., Liu, L., Li, Y., & Song, Z. (2025). Vehicle detection algorithm based on improved RT-DETR. *The Journal of Supercomputing*, 81(1), Article 290. <https://doi.org/10.1007/s11227-024-06766-7>
- Wu, M., Qiu, Y., Wang, W., Su, X., Cao, Y., & Bai, Y. (2025). Improved RT-DETR and its application to fruit ripeness detection. *Frontiers in Plant Science*. <https://doi.org/10.3389/fpls.2025.1423682>
- Yali Nie, Paolo Sommella, Mattias O’Nils, Consolatina Liguori, Jan Lundgren, "Automatic Detection of Melanoma with Yolo Deep Convolutional Neural Networks," *Proceedings of the IEEE International Conference on E-Health and Bioengineering (EHB)*, November 2019. DOI: 10.1109/EHB47216.2019.8970033.

- Yaseen, M. What is YOLOv8: An in-depth exploration of the internal features of the next-generation object detector. arXiv 2024. arXiv preprint arXiv:2408.15857.
- Yu, J., Jiang, Y., Wang, Z., Cao, Z., & Huang, T. (2016, October). Unitbox: An advanced object detection network. In Proceedings of the 24th ACM international conference on Multimedia (pp. 516-520).
- Zhang L, et al. Recent global patterns in skin cancer incidence, mortality, and prevalence. Chinese Medical Journal, 2025
- Zhang, S., et al. (2020). *Bridging the Gap Between Anchor-based and Anchor-free Detection via Adaptive Training Sample Selection*. arXiv:1912.02424.
- Zhang, Y., Li, X., Wang, F., Wei, B., & Li, L. (2021). A Comprehensive Review of One-stage Networks for Object Detection. IEEE ICSPCC, 9564613. <https://doi.org/10.1109/ICSPCC52875.2021.9564613>
- Zhang, Y., Liu, X., & Chen, W. (2023). YOLOv11: Next Generation Object Detection Model. IEEE Transactions on Image Processing, 32, 2105-2118.
- Zhao, Z., Chen, S., Ge, Y., Yang, P., Wang, Y., & Song, Y. (2024). RT-DETR-Tomato: Tomato Target Detection Algorithm Based on Improved RT-DETR for Agricultural Safety Production. *Applied Sciences*, 14(14), 6287. <https://doi.org/10.3390/app14146287> (MDPI)
- Zhou, C., He, H., Zhou, H., Ge, F., & Yu, P. (2025). MSRT-DETR: A novel RT-DETR model with multi-scale feature sequence for cell detection. *Biomedical Signal Processing and Control*, 103, 107378. <https://doi.org/10.1016/j.bspc.2024.107378>
- Zhou, C., He, H., Zhou, H., Ge, F., & Yu, P. (2025). MSRT-DETR: A novel RT-DETR model with multi-scale feature sequence for cell detection. *Biomedical Signal Processing and Control*, 103, 107378. <https://doi.org/10.1016/j.bspc.2024.107378>



ÖZGEÇMİŞ

Ad-Soyad : Sarah Atheer Abbas AL-MUSAWI

ÖĞRENİM DURUMU:

- Lisans** : 2021 yılında mezun, Sakarya Üniversitesi, Fen Edebiyat Fakültesi, Matematik Bölümü
- Yükseklisans** : 2025 yılında mezun, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Bilişim Sistemleri Mühendisliği

MESLEKİ DENEYİM

- 2020-2025 yılları arasında öğretmen olarak çalıştı.