



T.C.  
İSTANBUL ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ



**Yüksek Lisans Tezi**

**ÇOK MODLU GÖRÜNTÜLEME İLE BAŞ VE BOYUN  
BÖLGESİNDE RİSK ALTINDAKİ ORGANLARIN OTOMATİK  
SEGMENTASYONU**

**Sinan SEVGİ**

**Enformatik Anabilim Dalı**

**Enformatik Programı**

**DANIŞMAN  
Doç. Dr. Murat GEZER**

**Mart, 2025**

**İSTANBUL**

Bu çalışma, 14.04.2016 tarihinde ařağıdaki jüri tarafından Enformatik Anabilim Dalı, Enformatik Programında Yüksek Lisans tezi olarak kabul edilmiştir.

Tez Jürisi

Doç. Dr. Murat GEZER(Danışman)  
İstanbul Üniversitesi  
Fen Fakültesi

Prof. Dr. Sevinç GÜLSEÇENİ  
İstanbul Üniversitesi  
Enformatik Bölümü

Doç. Dr. Fatma Şebnem AKAL İLKHAN  
Marmara Üniversitesi  
İşletme Fakültesi

## İntihal Programı Beyanı

20.04.2016 tarihli Resmi Gazete’de yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince; Bu Lisansüstü teze, İstanbul Üniversitesi’nin aboneli olduğu intihal yazılım programı kullanılarak Fen Bilimleri Enstitüsü’nün belirlemiş olduğu ölçütlere uygun rapor alınmıştır.

## Tezden Üretilmiş Yayınların Künye Bilgileri

--



## ÖNSÖZ

Bu çalışmanın oluşum sürecinde, bazen kendime bile güvenim azalmışken beni her seferinde yeniden harekete geçiren, fikirleri ve yol göstericiliğiyle daima ilham kaynağı olan danışman hocam Doç. Dr. Murat GEZER'e içtenlikle teşekkür ederim. Onun sabrı ve desteği olmasaydı, bu sürecin üstesinden gelmek çok daha zor olurdu.

Tez süresince bana yalnızca destek olmakla kalmayıp, anlayışları ve sevgileriyle güç veren aileme minnettarım. Onların bana duyduğu inanç, bu zorlu yolculuğun en büyük motivasyon kaynağı oldu.

Bu süreçte hayatımı kolaylaştıran, beni güldürerek yorgunluğumu hafifleten arkadaşlarıma ve çalışmam boyunca emeği geçen herkese de teşekkür ederim.

Son olarak, bu tezin bana kattığı her deneyimi ve öğrendiğim her dersi, hayata karşı yeni bir bakış açısı kazanmanın bir hediyesi olarak görüyorum.

Mart 2025

Sinan SEVGİ

# İÇİNDEKİLER

Sayfa No

ÖNSÖZ .....	iv
İÇİNDEKİLER.....	v
ŞEKİL LİSTESİ .....	vii
TABLO LİSTESİ.....	viii
SİMGE VE KISALTMA LİSTESİ .....	ix
ÖZET .....	xi
SUMMARY .....	xii
<b>1. GİRİŞ .....</b>	<b>1</b>
<b>2. GENEL KISIMLAR.....</b>	<b>3</b>
2.1 MAKİNE ÖĞRENMESİ VE DERİN ÖĞRENME TEMELLERİ.....	3
2.1.2 Derin Öğrenme Modelleri ve Mimarileri .....	4
2.1.3 Özellik Mühendisliği ve Veri Ön İşleme.....	5
2.1.4. Optimizasyon Algoritmaları ve Eğitim Teknikleri.....	6
2.1.5. Aktivasyon Fonksiyonları ve Kaybolan Gradyan Problemi.....	7
2.1.6. Transfer Öğrenme ve Öntanımlı Modeller .....	9
YOLOv11 .....	11
SAM2 .....	13
2.2 BAŞ VE BOYUN BÖLGESİ ANATOMİSİ .....	16
2.3 RİSK ALTINDAKİ ORGANLAR (OAR) VE ÖNEMİ.....	18
2.4 TIBBİ GÖRÜNTÜLEME YÖNTEMLERİ .....	20
2.4.1. Manyetik Rezonans Görüntüleme (MR) .....	20
2.4.2. Bilgisayarlı Tomografi .....	21
2.4.3. Pozitron Emisyon Tomografisi.....	22
2.5. ÇOK MODLU GÖRÜNTÜLEME (MULTİMODAL IMAGING) .....	22
2.6. TIBBİ GÖRÜNTÜ SEGMENTASYONU .....	24
2.6.1. Geleneksel Yöntemler .....	24
2.6.2. Derin Öğrenme Tabanlı Yöntemler (Deep Learning-Based Methods).....	25
2.7. BAŞ VE BOYUN BÖLGESİ SEGMENTASYONUNDA MEVCUT YAKLAŞIMLAR.....	27
2.7.1. Baş ve Boyun Bölgesinde Tıbbi Görüntüleme ve Derin Öğrenme .....	27

2.7.2. Baş ve Boyun Görüntülemeye Hibrit ve Çok Aşamalı Segmentasyon Modelleri.....	33
2.7.3. Nesne Algılama (YOLO) ile Segmentasyonun Entegrasyonu .....	35
2.7.4. Tıbbi Segmentasyonda Segment Anything Model (SAM) ve Temel Modeller .....	36
<b>3. MALZEME VE YÖNTEM.....</b>	<b>39</b>
3.1 VERİ SETİ.....	39
3.2 ÇOK MODLU GÖRÜNTÜ FÜZYONU .....	41
3.3 ÖNERİLEN SEGMENTASYON MODELİ.....	42
3.3.1. Model Mimarisi (Model Architecture).....	42
<i>Teknik Entegrasyon ve Veri Akışı</i> .....	43
<i>YOLOv11 Çıktısı ve ROI (Region of Interest) İşleme:</i> .....	43
3.3.2. Kayıp Fonksiyonu (Loss Function) .....	44
3.3.3. Optimizasyon Stratejisi (Optimization Strategy) .....	45
<i>Döngüsel Öğrenme Oranı ve Momentum Stratejisi</i> .....	46
<i>Hibrit Optimizasyon Algoritmaları</i> .....	46
<i>Sparsemax ile Parametre Normalizasyonu</i> .....	47
<i>YOLOv11 ve SAM2 Uygulaması</i> .....	47
<b>4. BULGULAR.....</b>	<b>52</b>
4.1 Model Performansı .....	52
<b>5. TARTIŞMA.....</b>	<b>59</b>
5.1 SONUÇLARIN DEĞERLENDİRİLMESİ.....	59
5.2 KLİNİK UYGULANABİLİRLİK .....	59
5.3 SINIRLAMALAR VE GELECEK ÇALIŞMALAR .....	60
<b>6. SONUÇ .....</b>	<b>61</b>
6.1 TEMEL BULGULAR VE KATKILAR .....	61
6.2 GELECEK ARAŞTIRMA YÖNLERİ .....	61
<b>KAYNAKLAR.....</b>	<b>63</b>
<b>ÖZGEÇMİŞ .....</b>	<b>72</b>

## ŞEKİL LİSTESİ

	Sayfa No
Şekil 2.1: Yolov11'in genel mimari yapısı (He ve diğ., 2024).....	12
Şekil 2.2: YOLOv11 modelinin eski modellerle performans karşılaştırılma grafiği (Ultralytics, 2024.).....	13
Şekil 2.3: Fritscher ve arkadaşlarının önerdiği modelin diyagramı .....	30
Şekil 2.4: GAN modellerinin çalışma mantığı.....	31
Şekil 3.1: Veri setinden örnek görsel(Podobnik ve diğ., 2023a) .....	40
Şekil 3.2: Veri serinde sunulmuş segmantasyon maskelerinin örnek 3 boyutlu görüntüsü (Podobnik ve diğ., 2023a) .....	41
Şekil 3.3: MR, CT ve birleştirilmiş çok modlu görüntü .....	41
Şekil 3.4: Bounding box yapısı .....	43
Şekil 3.5: Kırılan ilgi alanlarının segmentasyon için SAM2 modeline beslenmesi.....	43
Şekil 3.6: Segmentasyon maskesi oluşturulması .....	44
Şekil 3.7: Çıktıların birleştirilmesi.....	44
Şekil 3.8: Eğitim sürecinde veri artırma tekniği uygulanmış örnekler .....	48
Şekil 3.9: Eğitim sonunda değerlendirme için oluşturulmuş çerçeveler ve etiketler.....	49
Şekil 3.10: Eğitimin ilerlemesini gösteren farklı aşamalarda alınmış gerçek maskeyle üretilen maskenin karşılaştırıldığı görseller .....	50
Şekil 3.11: Model Mimarisi .....	51
Şekil 4.1: Oluşturulan segmentasyon maskelerinin 3 Boyutlu görüntüsü .....	56
Şekil 4.2: Kesit üzerinde tespit edilen OAR'ların sınırlayıcı kutu ile belirlenmiş ve SAM 2 ile segmente edilmiş hali. 5. hastanın 47. kesiti üzerinde organlar tespit edilmiş ve daha sonra tiroit bezi için verilen uzman segmentasyonu ile karşılaştırılmıştır.....	57
Şekil 4.3: Ağız boşluğu ve çevresindeki organların segmentasyonun karşılaştırması .....	58
Şekil 4.4: Gözler ve göz sinirlerinin segmentasyonun karşılaştırması .....	58

## TABLO LİSTESİ

	<b>Sayfa No</b>
<b>Tablo 2.1:</b> PolypGen mevcut yöntemlerin performans karşılaştırması. (Mansoori ve diğ., 2024).....	15
<b>Tablo 4.1:</b> Organ bazında model performans değerleri .....	54
<b>Tablo 4.2:</b> Modellerin karşılaştırmalı performans analizi .....	55



## SİMGE VE KISALTMA LİSTESİ

Simgeler Açıklama

$LR_{cycle}$  : Öğrenme oranı döngüsü

$MR_{cycle}$  : Momentum oranı döngüsü

$sparsemax(\alpha)$  Sparsemax fonksiyonu

$softmax(\alpha)$  : Softmax fonksiyonu

$\tau$  : Sıcaklık parametresi (*Temperature parameter*)

$O$  : Çıktı lojitleri kümesi (*Set of output logits*)

Kısaltmalar Açıklama

**ART** : Adaptif Radyoterapi (*Adaptive Radiotherapy*)

**B-spline**: B-spline Eğrileri (*B-spline Curves* - Görüntü kaydırma ve modelleme için kullanılan matematiksel yöntem)

**CNN** : Konvolüsyonel Sinir Ağı (*Convolutional Neural Network*)

**CTA** : Bilgisayarlı Tomografi Anjiyografi (*Computed Tomography Angiography*)

**CT** : Bilgisayarlı Tomografi (*Computed Tomography*)

**DICOM** : Tıbbi Görüntüleme ve İletişim Standardı (*Digital Imaging and Communications in Medicine*)

**GAN**: Üretici Çekişmeli Ağlar (*Generative Adversarial Networks*)

**GIoU** : Genelleştirilmiş Kesişim Birliği (*Generalized Intersection Over Union*)

**mAP**: Ortalama Hassasiyet (*Mean Average Precision* - Nesne tespiti için değerlendirme metriği)

**MR** : Manyetik Rezonans (*Magnetic Resonance*)

**OAR**: Risk Altındaki Organ (*Organ-at-Risk*)

**PET** : Pozitron Emisyon Tomografisi (*Positron Emission Tomography*)

**ReLU** : Düzeltilmiş Doğrusal Birim (*Rectified Linear Unit*)

**ResNet** : Artık Ağ (*Residual Network* - Derin öğrenme model mimarisi)

**SAM2** : Segmentasyon Modeli 2 (*Segment Anything Model 2*)

**Voxel** : Hacimsel Piksel (3D piksel eşdeğeri)

**YOLO** : Sadece Bir Kez Bakarsın (*You Only Look Once* - Gerçek zamanlı nesne tespiti modeli)



## ÖZET

### YÜKSEK LİSANS TEZİ

#### ÇOK MODLU GÖRÜNTÜLEME İLE BAŞ VE BOYUN BÖLGESİNDE RİSK ALTINDAKİ ORGANLARIN OTOMATİK SEGMENTASYONU

Sinan SEVGİ

İstanbul Üniversitesi

Fen Bilimleri Enstitüsü

Enformatik Anabilim Dalı

Danışman : Doç. Dr. Murat GEZER

Bu tez, baş ve boyun bölgesindeki risk altındaki organların (Organ-at-Risk, OAR) otomatik segmentasyonunu iyileştirmek için YOLOv11 ve SAM2 derin öğrenme modellerini entegre ederek çok modlu görüntüleme (CT ve MR) yaklaşımı sunmaktadır. OAR segmentasyonu, radyoterapi planlamasında kritik bir öneme sahiptir. Geliştirilen model, adaptif radyoterapi süreçlerinde yüksek doğruluklu segmentasyon sağlayarak tedavi etkinliğini artırmayı ve yan etkileri azaltmayı hedeflemektedir. Bu amaçla, HaN-Seg veri seti kullanılarak model eğitimi gerçekleştirilmiş ve performansı Dice Benzerlik Katsayısı ile değerlendirilmiştir. Sonuçlar, önerilen yöntemin mevcut segmentasyon yöntemlerinden daha hassas ve tutarlı olduğunu göstermiştir.

Mart 2025, 84 sayfa.

**Anahtar kelimeler:** Çok Modlu Görüntü Segmentasyonu, Derin Öğrenme

## **SUMMARY**

### **M.Sc. THESIS**

#### **AUTOMATIC SEGMENTATION OF ORGANS-AT-RISK IN THE HEAD AND NECK REGION USING MULTI-MODAL IMAGING**

**Sinan SEVGİ**

**İstanbul University**

**Institute of Graduate Studies in Sciences**

**Department of Informatics**

**Supervisor : Assoc. Prof. Dr. Murat GEZER**

This thesis proposes an approach for the automatic segmentation of organs-at-risk (OAR) in the head and neck region by integrating YOLOv11 and SAM2 deep learning models with multi-modal imaging (CT and MR). OAR segmentation is critical for radiotherapy planning. The developed model aims to improve segmentation accuracy in adaptive radiotherapy processes, enhancing treatment effectiveness and reducing side effects. The model was trained using the HaN-Seg dataset, and its performance was evaluated using the Dice Similarity Coefficient. Results demonstrate that the proposed method outperforms existing segmentation methods in precision and consistency.

March 2025, 84 pages.

**Keywords:** Multi-modal Image Segmentation, Deep Learning

## 1. GİRİŞ

Baş ve boyun kanserleri, dünya genelinde her yıl yüz binlerce yeni vaka ile karşılaşılan ve ciddi sağlık sorunlarına yol açan bir hastalık grubudur (Bray ve diğ., 2018). Bu kanserlerin tedavisinde cerrahi, kemoterapi ve radyoterapi sıklıkla bir arada kullanılmaktadır ve radyoterapi tümörlerin hedeflenmesi açısından kritik bir rol oynamaktadır (Chow, 2020). Ancak, radyoterapi sırasında çevredeki sağlıklı dokuların korunması için tümörün hassas bir şekilde hedeflenmesi gerekmektedir, bu durum büyük bir zorluk oluşturmaktadır. Baş-boyun bölgesinde yer alan ve risk altındaki organlar (OAR) olarak tanımlanan kritik yapılar, radyasyonun yanlış yönlendirilmesi durumunda hasar almakta ve bu hasar ciddi yan etkilere yol açabilmektedir (Brouwer ve diğ., 2015). OAR'ların doğru şekilde belirlenmesi ve korunması, hem tedavi etkinliğini artırmakta hem de hastaların yaşam kalitesini korumak açısından büyük önem taşımaktadır (Mayo ve diğ., 2018).

Mevcut klinik uygulamalarda, baş-boyun bölgesindeki OAR'ların segmentasyonu genellikle manuel veya atlas tabanlı yöntemlerle gerçekleştirilmekte olup, bu süreç zaman alıcı ve gözlemciler arası değişkenliklere açık bir işlem olmaktadır (Yang ve diğ., 2018). Otomatik segmentasyon yöntemleri, bu süreci hızlandırarak daha hassas ve tutarlı sonuçlar sunabilmeyi amaçlamaktadır. Son yıllarda derin öğrenme tabanlı yöntemler, tıbbi görüntü segmentasyonunda başarılı sonuçlar vermekte olup, özellikle U-Net gibi konvolüsyonel sinir ağı (CNN) tabanlı modeller yaygın olarak kullanılmaktadırlar (Ronneberger ve diğ., 2015). Ancak, baş-boyun bölgesinin karmaşık anatomik yapısı ve OAR'ların birbirine yakın konumlanması, bu tür genel segmentasyon yaklaşımlarının başarısını sınırlamaktadır. (Vrtovec ve diğ., 2020).

Ayrıca, bu bölgedeki yapıların farklı fiziksel özellikleri tek bir tıbbi görüntüleme yönteminin otomatik segmentasyon yöntemleri için yetersiz kalmasına sebep olmakta ve farklı tıbbi görüntüleme modaliteleri baş-boyun bölgesindeki yapıları tamamlayıcı bilgiler sağlasa da, bu görüntülerin entegrasyonu teknik olarak zorlayıcı olmaktadır (Podobnik ve diğ., 2023a).

CT (Bilgisayarlı Tomografi), kemik yapılarını net gösterirken, MR (Manyetik Rezonans) yumuşak dokuların detaylandırılmasında daha etkili olmaktadır. Bu iki modalitenin birlikte

kullanımı, segmentasyon doğruluğunu artırabilmesine rağmen, modaliteler arası geometrik uyumsuzluklar ve veri heterojenliği gibi problemler nedeniyle yaygın segmentasyon yöntemlerinin performansını sınırlamaktadır (Liu ve diğ., 2021).

Bu çalışma, baş-boyun bölgesindeki OAR'ların daha hassas ve otomatik segmentasyonu için YOLOv11 ve Segment Anything Model 2 (SAM2) gibi derin öğrenme tabanlı yaklaşımların birlikte kullanılmasını önermektedir. Çalışmada, CT ve MR görüntülerinin çok modlu olarak birlikte değerlendirilmesiyle daha güvenilir segmentasyon elde edilmesi hedeflenmektedir. Önerilen yöntem, adaptif radyoterapi (ART) süreçlerinde kullanılacak hızlı ve yüksek doğruluklu bir segmentasyon modeli geliştirmeyi amaçlamaktadır. Çalışmanın hipotezi, YOLOv11 ve SAM2 tabanlı derin öğrenme modelinin, baş-boyun bölgesindeki OAR'ların otomatik segmentasyonunda tek modlu ve hibrit olmayan yöntemlere kıyasla daha başarılı sonuçlar vereceğidir. Bu bağlamda, geliştirilen modelin performansı HaN-Seg veri seti kullanılarak değerlendirilecek ve segmentasyon doğruluğu Dice Benzerlik Katsayısı gibi metriklerle analiz edilecektir.

## 2. GENEL KISIMLAR

### 2.1 MAKİNE ÖĞRENMESİ VE DERİN ÖĞRENME TEMELLERİ

Makine Öğrenmesi, yapay zekânın bir alt alanı olup bilgisayarların açıkça programlanmadan öğrenmesini amaçlar. Tom Mitchell'in klasik tanımına göre, "*Makine öğrenmesi, deneyimle birlikte performansını otomatik olarak geliştiren bilgisayar algoritmalarının incelenmesidir*" (Mitchell, 2013). Bu alanda algoritmalar geçmiş verilerden örüntüler öğrenerek gelecekte karşılaşacakları veriler için tahminler üretir. Derin Öğrenme ise makine öğrenmesinin bir alt dalıdır ve insan beyninin çalışma prensiplerinden esinlenen çok katmanlı yapay sinir ağlarını kullanır. Derin öğrenme, çok katmanlı sinir ağları (derin sinir ağları) aracılığıyla verideki karmaşık ilişkileri öğrenerek insan beyninin karar verme yetisini taklit eder (IBM, 2024). Özellikle büyük veri ve hesaplama gücünün artışıyla derin öğrenme, makine öğrenmesi yöntemleri arasında öne çıkmıştır.

Makine Öğrenmesi yöntemleri genel olarak üç temel öğrenme paradigmasına ayrılır:

- **Denetimli öğrenme** – girdi verilerine karşılık doğru çıktılar (etiketler) verilerek modelin eğitildiği yöntem,
- **Denetimsiz öğrenme** – veriye ait etiketlerin olmadığı, modelin verideki gizli yapıları keşfetmeye çalıştığı yöntem,
- **Pekiştirmeli öğrenme** – bir ajan'ın çevreyle etkileşime girip ödül/ceza mekanizması ile öğrenme sağladığı yöntem. Bu üç yaklaşım makine öğreniminin temelini oluşturur ve farklı problem tiplerine uygulanırlar. Örneğin, bir görüntüdeki tümörün varlığını saptamak genellikle denetimli öğrenmeye bir örnekken, büyük bir hasta verisinde ortak özellik gruplarını keşfetmek denetimsiz öğrenmeye girer; otonom bir cihazın eylemlerini ödüle göre optimize etmesi ise pekiştirmeli öğrenmedir (IBM, 2024).

Makine Öğrenmesi ve Derin öğrenme yöntemleri günümüzde birçok alanda devrim niteliğinde başarılar elde etmiştir (IBM, 2024). Özellikle bilgisayarlı görü ve doğal dil işleme gibi alanlarda derin öğrenme modelleri, insan seviyesine yakın performanslara ulaşmıştır. Örneğin, derin öğrenme algoritmaları görüntü tanıma, ses tanıma ve dil çevirisi gibi görevlerde çarpıcı başarılar sergilemiştir (Shulman, 2023). Bu sayede otonom araçlar, tıbbi görüntü analizi, sesli asistanlar ve makine çevirisi gibi uygulamalarda Makine Öğrenmesi ve Derin öğrenme yöntemleri vazgeçilmez hale gelmiştir. Derin öğrenmenin son yıllardaki atılımı, bu yöntemlerin

büyük veri setlerinde kendi özelliklerini öğrenebilme kabiliyetinden ve hesaplama altyapısının gelişmesinden kaynaklanmaktadır (Shulman, 2023).

### 2.1.2 Derin Öğrenme Modelleri ve Mimarileri

Derin öğrenme modellerinin temelini Yapay Sinir Ağları oluşturur. Yapay Sinir Ağları, birbirine bağlı yapay katmanlar halinde düzenlenmiş nöronlardan oluşan ve girdiler ile çıktılar arasındaki karmaşık ilişkileri öğrenebilen yapılardır (Boué, 2018). Bir Yapay Sinir Ağları genellikle giriş katmanı, bir veya daha fazla gizli katman ve bir çıkış katmanından oluşur. Geri yayılım algoritması, Yapay Sinir Ağlarının öğrenmesini sağlayan temel yöntemdir (Boué, 2018). Bu algoritma, ağın çıktısındaki hatayı hesaplayıp türevini alarak hata bilgisini girişe doğru katmanlar arasında geriye yayar ve her bir bağlantı ağırlığını bu hataya göre günceller. Geri yayılımın 1980'lerde yeniden keşfedilip popülerleşmesi, derin ağların eğitiminde kritik bir rol oynamıştır (Boué, 2018).

Derin öğrenmede farklı mimariler belirli problem türlerinde uzmanlaşmıştır. Evrişimli Sinir Ağları özellikle görüntü verilerinin işlenmesinde başarılı bir mimaridir. Yann LeCun ve arkadaşlarının geliştirdiği LeNet (1998), el yazısı rakam tanımada erken başarılı Evrişimli Sinir Ağı örneklerindedir (Ronneberger ve diğ., 2015). Evrişimli Sinir Ağları, evrişim katmanları sayesinde görüntülerdeki uzamsal ilişkileri ve yerel özellikleri yakalar. 2012'de Krizhevsky ve arkadaşlarının geliştirdiği AlexNet modeli, büyük bir görüntü veri kümesi (ImageNet) üzerinde eğitilerek o güne kadarki en iyi sonuçları elde etmiş ve derin Evrişimli Sinir Ağlarının görsel tanımadaki gücünü göstermiştir. Evrişimli Sinir Ağı tabanlı ağlar günümüzde nesne tanıma, yüz tanıma ve tıbbi görüntü segmentasyonu gibi pek çok bilgisayarlı görü probleminde standart yaklaşım haline gelmiştir.

Tekrarlayan Sinir Ağları ise sıralı verilerle (zaman serileri, metin, konuşma vb.) çalışmak için tasarlanmış ağlardır. Tekrarlayan Sinir Ağları döngüsel yapısı sayesinde önceki adımların bilgisini sonraki adımlara aktararak dil modelleri veya zaman serisi tahminleri gibi problemleri ele alır. Ancak klasik Tekrarlayan Sinir Ağları çok uzun sekanslarda kaybolan gradyan sorunuyla karşılaşır; erken adımlardaki bilgilerin öğrenilmesi, ağ derinleştikçe zorlaşır. Bu soruna çözüm olarak Uzun Kısa Süreli Bellek hücreleri geliştirilmiştir. Hochreiter ve Schmidhuber'in 1997'de ortaya koyduğu Uzun Kısa Süreli Bellek mimarisi, kapı

mekanizmaları sayesinde gradyan akışını uzun süre koruyarak RNN'lerin uzun vadeli bağımlılıkları öğrenebilmesini sağlar (Hochreiter & Schmidhuber, 1997). Bu sayede Uzun Kısa Süreli Bellek ve onun türevi olan GRU gibi modeller, dil çevirisi, konuşma tanıma ve video analizi gibi alanlarda başarılı sonuçlar vermiştir.

Günümüzde derin öğrenme mimarileri içinde en modern yaklaşımlardan biri Transformer modelidir. 2017 yılında Vaswani ve arkadaşlarının sunduğu "Attention Is All You Need" makalesi ile tanıtılan Transformer mimarisi, tamamen dikkat mekanizması üzerine kuruludur ve Tekrarlayan Sinir Ağı veya Evrişimli Sinir Ağı katmanlarına gerek duymadan sıradan-sıraya problemlerini çözebilmektedir (Vaswani ve diğ., 2023). Transformer modeli, giriş dizisindeki öğelerin birbirleriyle olan ilişkilerini dikkat ağırlıkları aracılığıyla öğrenerek paralel işlemeyi mümkün kılar (Vaswani ve diğ., 2023). Makine çevirisi başta olmak üzere doğal dil işleme görevlerinde çığır açmıştır. Bu mimari sonrasında BERT, GPT gibi büyük dil modellerinin temelini oluşturmuş ve günümüz derin öğrenme araştırmalarında merkezi bir rol almıştır.

### 2.1.3 Özellik Mühendisliği ve Veri Ön İşleme

Derin öğrenme modellerinin başarısı büyük ölçüde kullanılan verinin kalitesine ve işleme şekline bağlıdır. Veri ön işleme, ham veriyi modele uygun hale getirmek ve öğrenme başarımını artırmak için yapılan adımları içerir. Tıbbi görüntü segmentasyonu bağlamında veri ön işlemeye birkaç örnek verilebilir. Görüntülerin yeniden boyutlandırılması veya kırılması (ilgili bölgeye odaklanmak için), normalizasyon, gürültü giderme filtreleri ve kontrast iyileştirme yöntemleri. Özellikle baş-boyun bölgesi gibi hassas anatomik alanlarda, CT ve MR görüntülerinin kayıtlanması da önemli bir adımdır; bu sayede iki farklı modaliteden gelen veriler hizalanarak birlikte kullanılabilir. (Trigka & Dritsas, 2025)

Özellik mühendisliği, geleneksel makine öğrenmesinde modelin performansı için kritik bir adımdır; uzmanlar veriden elle anlamlı özellikler çıkarırlardı. Derin öğrenmede ise ağlar ham veriden kendileri özellik öğrenebilmektedirler (Goodfellow ve diğ., 2016). Yine de özellikle verinin sınırlı olması durumunda veya modelin işini kolaylaştırmak için bazı özellik çıkarım adımları kullanılmaktadır (Goodfellow ve diğ., 2016). Örneğin, görüntülerde kenar tespiti, dokusal özellikler veya boyut indirgeme teknikleri uygulanarak veri boyutu azaltılabilmektedir.

Boyut indirgeme, modelin girdi uzayını küçülterek hesaplama yükünü ve gereksiz bilgiyi azaltabilir. Bununla birlikte, derin öğrenme çağında çoğu zaman ham görüntü verisi minimal ön işlemler (ör. sadece normalizasyon) ile ağlara verilir ve ağın kendisi çok katmanlı yapısıyla en uygun öz temsilleri öğrenir. (Yemi Hussain, 2024)

Görüntü işlemede veri kısıtlılığı yaygın bir sorundur; bu nedenle veri artırma teknikleri modelin genelleme yeteneğini geliştirmek için sıkça kullanılır. Veri artırma ile mevcut görüntüler çeşitli rastgele dönüşümlere uğratarak (döndürme, yansıtma, ölçekleme, gürültü ekleme vb.) sentetik yeni örnekler elde edilir (Yemi Hussain, 2024) .

Bu özellikle derin öğrenmenin başlangıç aşamasında sınırlı biyomedikal veri ile eğitilen modellerde başarıyı yükselten bir strateji olmuştur. Örneğin, U-Net mimarisini tanıtan çalışmada ağın çok az sayıda görüntüyle eğitilebilmesi için yoğun veri artırma uygulandığı vurgulanmıştır (Ronneberger ve diğ., 2015). Ronneberger ve arkadaşları, biyomedikal görüntü segmentasyonunda başarılı olabilmek için ağın eğitiminde dönüşüm ve deformasyonlar ile veriyi çeşitlendirerek aşırı uyumu engellemiş ve modeli daha genellenebilir hale getirmiştir (Ronneberger ve diğ., 2015). Kaliteli bir segmentasyon modeli geliştirmek, uygun veri ön işleme adımları (hizalama, normalizasyon vb.) ve gerekli durumlarda veri artırma teknikleriyle desteklendiğinde mümkün olmaktadır.

#### **2.1.4. Optimizasyon Algoritmaları ve Eğitim Teknikleri**

Derin öğrenme modellerinin eğitimi, kayıp fonksiyonunu en aza indirecek şekilde ağ parametrelerinin optimize edilmesi sürecidir. Bu süreçte en temel yöntem Gradyan İnişi algoritmasıdır (Goodfellow ve diğ., 2016). Türev tabanlı bu yöntem, hata türevini hesaplayarak parametreleri hatanın azaldığı yönde adım adım günceller. Stokastik Gradyan İnişi, her iterasyonda tüm veri yerine küçük veri altkümeleri kullanarak güncelleme yapan pratik bir versiyondur ve derin öğrenmede yaygın kullanılır (Anandhavalli Muniasamy, 2024). Optimizasyon alanındaki gelişmeler, Stokastik Gradyan İnişinin farklı varyasyonlarını ortaya çıkarmıştır. Momentumlu Stokastik Gradyan İnişi, öğrenme hızını ivmelendirerek sallantıları azaltır; RMSprop ve Adam gibi uyarlamalı öğrenme oranı kullanan algoritmalar ise her parametre için ayrı uyarlanabilir adım boyları hesaplar. Özellikle 2015'te Kingma ve Ba tarafından önerilen Adam optimizasyonu, önceki momentum ve adaptif yaklaşımları

birleştirek derin ağların eğitiminde hızlı ve kararlı bir yöntem sunmuştur (Rahul Agarwal, 2024.). Adam, her bir parametrenin geçmiş gradyan ortalamalarını (birinci ve ikinci momentlerini) izleyerek öğrenme oranını dinamik olarak ayarlar ve genellikle derin öğrenme modellerinde varsayılan algoritma olarak tercih edilir (Rahul Agarwal, 2024). Uygun bir optimizasyon algoritması seçimi, ağın daha hızlı yakınsamasına ve daha iyi genelleşmesine yardımcı olur.

Eğitim sırasında aşırı uyum riskini azaltmak ve modeli daha iyi genelleştirmek için çeşitli düzenleme teknikleri uygulanır. Birim çıkartma, bu tekniklerin klasik bir örneğidir. Srivastava ve arkadaşlarının 2014'teki çalışmasında tanıtılan birim çıkartma, her eğitim adımında rastgele olarak ağdaki bazı nöronları devre dışı bırakarak modelin belirli nöron kombinasyonlarına aşırı bağımlı olmasını engeller (Hinton ve diğ., 2012). Bu sayede, farklı alt-ağların eğitilmesiyle bir tür ansamble etkisi yaratılır ve aşırı öğrenme azaltılır. Nitekim, büyük bir sinir ağında her eğitim örneğinde nöronların yaklaşık yarısını rastgele kapatmanın test hatasını belirgin biçimde düşürdüğü gösterilmiştir (Hinton ve diğ., 2012). Bir diğer önemli teknik Toplu Normalizasyon'dur. Ioffe ve Szegedy'nin 2015'te önerdiği bu yöntem, her minibatch içinde ara katman aktivasyonlarının ortalamasını ve varyansını sabitleyerek eğitim sürecini hızlandırır ve kararlılığı artırır. Batch norm ile daha yüksek öğrenme hızları kullanılabilir ve katmanlar arası dağılım değişikliği azaltıldığı için eğitim daha verimli gerçekleşir (Ioffe & Szegedy, 2015). Ayrıca batch norm hafif bir düzenleme etkisi de yaparak bazı durumlarda birim çıkarma ihtiyacını dahi azaltabilir (Ioffe & Szegedy, 2015).

Düzenleme tekniklerine veri artırmayı da eklemek gerekir; önceki bölümde bahsedildiği gibi veri augmentasyonu modele daha çeşitli örnekler göstererek aşırı uyumu engeller. Tüm bu eğitim stratejileri bir arada kullanıldığında (uygun optimizasyon seçimi, dropout, batch norm, veri augmentasyonu, uygun öğrenme hızı planlaması vb.), derin ağlar daha hızlı ve doğru bir şekilde öğrenebilir.

### **2.1.5. Aktivasyon Fonksiyonları ve Kaybolan Gradyan Problemi**

Yapay sinir ağlarında aktivasyon fonksiyonları, her nöronun çıktısını hesaplamak için kullanılır ve ağın doğrusal olmayan ilişkileri öğrenebilmesini sağlar (Goodfellow ve diğ., 2016). Yaygın

aktivasyon fonksiyonları arasında Sigmoid ve Tanh (hiperbolik tanjant) tarihsel olarak popülerdi; ancak bu fonksiyonlar özellikle uç değerlere doğru çıktıları saturasyona uğrayıp türevleri sıfıra yaklaştığı için derin katmanlarda öğrenmeyi zorlaştıran kaybolan gradyan problemini ortaya çıkardı (Goodfellow ve diğ., 2016). Kaybolan gradyan problemi, çok katmanlı ağlarda hata sinyalinin geri yayılım esnasında üstel olarak zayıflayıp ilk katmanlara neredeyse ulaşamaması durumudur. Bu durumda ağın erken katmanları yeterince öğrenemez ve derin ağların eğitimi verimsiz hale gelir. Bu sorunun fark edilmesi, derin öğrenmenin erken dönemlerinde araştırmacıları farklı çözümler aramaya yöneltmiştir. (Andrychowicz ve diğ., 2016)

Bu çözümlerden en etkili olanı yeni aktivasyon fonksiyonlarının geliştirilmesiydi. ReLU (Rectified Linear Unit) fonksiyonu, 2010'lerde derin öğrenmede çığır açan bir aktivasyon olarak ortaya çıktı. ReLU, negatif girdilerde sıfır, pozitif girdilerde doğru orantılı bir çıkış verir; türevi ise pozitif bölgede 1, negatif bölgede 0'dır (Goodfellow ve diğ., 2016). Bu basit form sayesinde ReLU, sigmoid/tanh gibi saturasyona girmediği için büyük derinliklerde bile gradyan bilgisinin kaybolmamasını sağlar (Goodfellow ve diğ., 2016). AlexNet çalışmasında gösterildiği üzere, ReLU kullanan derin CNN'ler tanh kullanan benzerlerine göre 6 kata kadar daha hızlı yakınsama göstermiştir (Krizhevsky ve diğ., 2017). ReLU, büyük değerlerde bile çıktıyı sınırlanmadığından açılmayan nöron sorunu dışında kaybolan gradyana karşı oldukça dayanıklıdır. Modern ağların çoğu ara katmanlarında ReLU veya türevlerini (Leaky ReLU, Parametric ReLU vb.) kullanarak derinliği artırabilmiştir. Örneğin, ReLU aktivasyonu sayesinde ağlar uç değerlere doyumluk yaşamadan gradyanı iletmeye devam eder ve böylece çok daha derin katmanlar eğitilebilir hale gelir (Ramachandran ve diğ., 2017).

Sigmoid ve tanh gibi fonksiyonlar ise günümüzde genellikle ağların son katmanlarında belirli amaçlarla kullanılır. Softmax fonksiyonu, çok sınıflı sınıflandırma problemlerinde son katmanda kullanılarak her bir sınıf için 0-1 aralığında olasılık dağılımı oluşturur (tüm çıktıların toplamı 1 olacak şekilde). Sigmoid, ikili sınıflandırmada son katmanda olasılık vermek için veya bazı gizli katmanlarda dikkat mekanizması, kapı mekanizması gibi özel yapılarda kullanılır. Tanh, -1 ile 1 arasında çıktı verdiği ve sıfır merkezli olduğu için tarihsel olarak sigmoid'e tercih edilmişti, ancak o da saturasyon sorunu yaşattır. Bu yüzden iç katmanlarda

tanh/sigmoid yerine ReLU türleri tercih edilirken, tanh hala bazı özel durumlarda (ör. ağırlık girişini normalize etmek için) kullanılabilir.

Derin ağların eğitimi sırasında ortaya çıkan kaybolan ve patlayan gradyan problemlerine karşı literatürde çeşitli önlemler geliştirilmiştir. Kaybolan gradyan için bahsedilen ReLU ve türevlerinin kullanılması, uygun ağırlık başlangıçlandırma teknikleri (He veya Xavier initialization, gradyan akışını dengeler), atlamalı bağlantılar eklenmesi (Örneğin ResNet mimarisinde her birkaç katmanda bir doğrudan bağlantı kurularak gradyanın akışı kolaylaştırılmıştır (He ve diğ., 2016)) gibi yöntemler başarıyla uygulanmıştır. Patlayan gradyan sorunu içinse gradyanların çok büyümesini engelleyen gradient clipping tekniği yaygın kullanılır. Doğru mimari ve aktivasyon fonksiyonu tercihleriyle derin ağların eğitimindeki teorik zorluklar büyük ölçüde aşılmıştır. Özellikle ReLU'nun kullanımı ve yenilikçi mimariler (LSTM, ResNet gibi) sayesinde kaybolan gradyan problemi pratikte kontrol altına alınmış ve çok katmanlı derin ağlar etkin bir şekilde eğitilebilir hale gelmiştir.

### 2.1.6. Transfer Öğrenme ve Öntanımlı Modeller

Derin öğrenme modellerinin başarılı olabilmesi için genellikle çok büyük miktarda veri ve hesaplama gücü gerekebilir. Transfer öğrenme, bu maliyeti düşürmek ve veri azlığı sorununu aşmak için kullanılan etkin bir yaklaşımdır (Goodfellow ve diğ., 2016). Temel olarak transfer öğrenme, büyük bir kaynak veri kümesinde önceden eğitilmiş bir modelin (öntanımlı modelin) bilgi birikimini, hedef bir görevde yeniden kullanma prensibine dayanır (Goodfellow ve diğ., 2016). Örneğin, ImageNet gibi milyonlarca görüntü içeren bir veri setinde eğitilmiş bir derin CNN modeli, başka bir spesifik göreve (ör. bir medikal görüntü sınıflandırma veya segmentasyon görevine) doğrudan uygulanabilir ya da üzerine biraz daha eğitim yapılarak adapte edilebilir. Bu sayede hedef görevde az miktarda veri olsa bile, önceden öğrenilmiş özellikler sayesinde model güçlü bir başlangıç yapar.

Görüntü tanıma alanında yaygın olarak kullanılan öntanımlı modellerden bazıları VGG ve ResNet mimarileridir. *VGG16* ağı (Simonyan & Zisserman, 2015), 13 konvolüsyon + 3 tam bağlantılı katman olmak üzere 16 katmanlı derin bir CNN modelidir ve ImageNet veri setinde %92.7 top-5 doğruluk gibi yüksek bir başarı elde etmiştir (Simonyan & Zisserman, 2015). Yaklaşık 138 milyon öğrenilebilir parametreye sahip bu ağ, sade yapısı (aynı boyutta filtrelerin

ardışık yığılması) ve derinliği sayesinde pek çok görsel göreve genelleşebilen güçlü özellikler yakalamıştır. (Simonyan & Zisserman, 2015). VGG gibi modeller, *evrişimlerle öğrenilmiş kenar, köşe, doku gibi orta seviye özellikleri* yeniden kullanılabilmesi için, medikal görüntüleme dahil farklı alanlardaki görsel tanıma/segmentasyon problemlerine transfer öğrenme yoluyla başarıyla uygulanmıştır. Bir başka önemli model ailesi, 2015'te Kaiming He ve arkadaşlarının geliştirdiği ResNet (Residual Networks)'tir. ResNet, derin ağlarda ortaya çıkan kaybolan gradyan sorununu skip connection adı verilen atlamalı bağlantılarla çözen ve 152 katmana kadar çıkabilen son derece derin bir CNN mimarisidir (He ve diğ., 2016). ResNet-50 veya ResNet-101 gibi versiyonlar da ImageNet üzerinde yüksek doğruluklar elde etmiş ve günümüzde birçok vizyon görevinde ön eğitilmiş modeller olarak kullanılmaktadır. Özellikle atlamalı bağlantılar sayesinde ResNet, çok daha derin ağların eğitilmesine olanak tanımış ve bu mimarinin başarısı transfer öğrenme ile başka görevlere de taşınmıştır (He ve diğ., 2016).

Transfer öğreniminin pratikte iki ana yolu vardır: Özellik çıkarımı ve ince ayar. Özellik çıkarımında, önceden eğitilmiş modelin son katmanları çıkarılır ve geri kalan ağ, sabit özellik dönüştürücü olarak kullanılır; bu özelliklerin üzerine yeni bir sınıflandırıcı/segmentasyon ağı eğitilir (Goodfellow ve diğ., 2016). İnce ayar yönteminde ise ön eğitilmiş model, hedef veri üzerinde düşük öğrenme oranıyla biraz daha eğitilerek hem genel özellikler hem de görev özelindeki özellikler birlikte uyumlu hale getirilir (Goodfellow ve diğ., 2016). Özellikle medikal görüntüleme gibi verinin az olduğu alanlarda, ImageNet üzerinde eğitilmiş bir CNN'in ilk katmanlarının evrensel kenar/renk/doku özellikleri yakaladığı ve transferde çok faydalı olduğu bilinmektedir.

Transfer öğrenme ve öntanımlı modeller, derin öğrenme çalışmalarında adeta bilgi nakli görevi görerek araştırmacıların küçük veri ile büyük işler başarmasını sağlamıştır. VGG, ResNet, BERT gibi modeller geniş çapta ön eğitimden geçmiş başarılı ağlardır ve bu ağların sağladığı zengin temsil gücü, yeni problemler çözerken önemli bir avantaj sunmaktadır. Bu yaklaşımla, baştan sıfırdan eğitime kıyasla çok daha hızlı, verimli ve yüksek performanslı sonuçlar almak mümkün olmaktadır.

## YOLOv11

You Only Look Once (YOLO), nesne tespiti görevlerini bir regresyon problemi olarak ele alan, yenilikçi ve tek aşamalı bir derin öğrenme algoritmasıdır. İlk kez 2015 yılında Redmon ve arkadaşları tarafından tanıtılan YOLO, geleneksel iki aşamalı nesne tespiti yöntemlerinden farklı olarak, tüm görüntüyü tek bir ileri geçişte işler ve nesne konumlarını, boyutlarını ve sınıflarını aynı anda tahmin eder (Redmon ve diğ., 2016).

Bu yaklaşım, hem hız hem de doğruluk açısından büyük bir avantaj sağlamaktadır, çünkü geleneksel yöntemlerdeki gibi bölge önerisi ve özellik çıkarımı gibi aşamalara gerek kalmaz. YOLO, nesne tespitini bir regresyon problemi olarak formüle eder ve tek bir ağ ile bu görevi yerine getirir. Bu nedenle, gerçek zamanlı uygulamalar için idealdir. (Redmon ve diğ., 2016).

YOLO, görüntüyü  $S \times S$  boyutunda bir ızgaraya böler ve her hücreye birden fazla sınır kutusu (bounding box) ve nesne sınıfı tahmini atar. Algoritmanın temel çalışma prensipleri şu şekilde özetlenebilir (Redmon ve diğ., 2016):

- Giriş: Model bir görüntüyü alır ve bunu sabit bir boyuta ölçekler.
- Izgara Bölme: Giriş görüntüsü  $S \times S$  boyutunda bir ızgaraya bölünür. Her hücre, o bölgedeki bir nesneye ait sınır kutusu ve sınıf tahmini yapmakla sorumludur.
- Sınır Kutusu Regresyonu: Her hücre, nesneye ait dört sınır kutusu parametresini ( $x$ ,  $y$ , genişlik, yükseklik) ve objectness skorunu (nesnenin var olup olmadığını belirten bir olasılık) tahmin eder.
- Sınıf Tahmini: Ayrıca, her hücre tespit ettiği nesneye ait olasılık tabanlı bir sınıf tahmini yapar.
- Birleştirme ve Filtreleme: Tahminler, Intersection over Union (IoU) ve Non-Maximum Suppression (NMS) kullanılarak birleştirilir ve en olası tespitler seçilir.

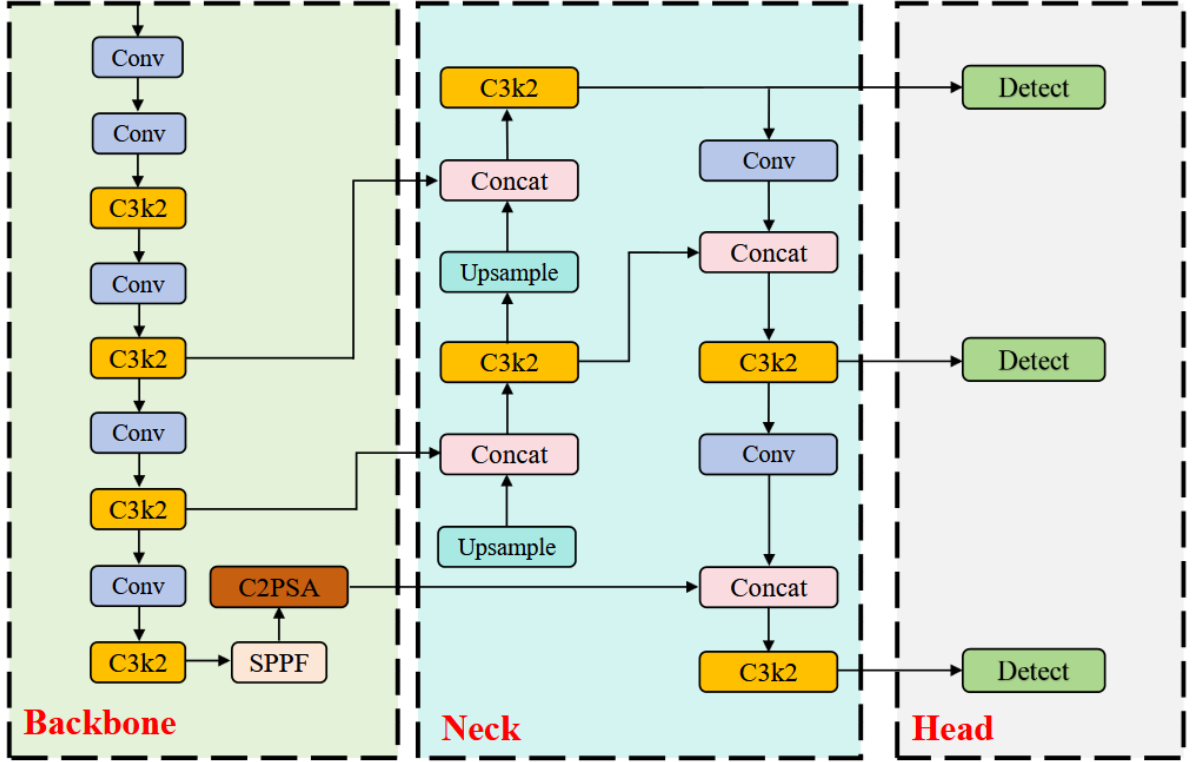
YOLOv11, YOLO serisinin en son sürümüdür ve öncekilerden farklı olarak bir dizi mimari iyileştirme ve yeni özellikler içermektedir (Khanam ve Hussain, 2024).

YOLOv11, geleneksel algılama yöntemlerinden farklı olarak, nesne tespiti ve sınıflandırmayı tek bir derin öğrenme modeliyle gerçekleştirir. Model, üç ana bileşenden oluşur: Backbone, Neck ve Head (He ve diğ., 2024).

- Backbone: Giriş görüntüsünden çok ölçekli özellik haritaları çıkarır. YOLOv11'de, önceki sürümlerden farklı olarak C3k2 bloğu kullanılmıştır (He ve diğ., 2024). Bu blok,

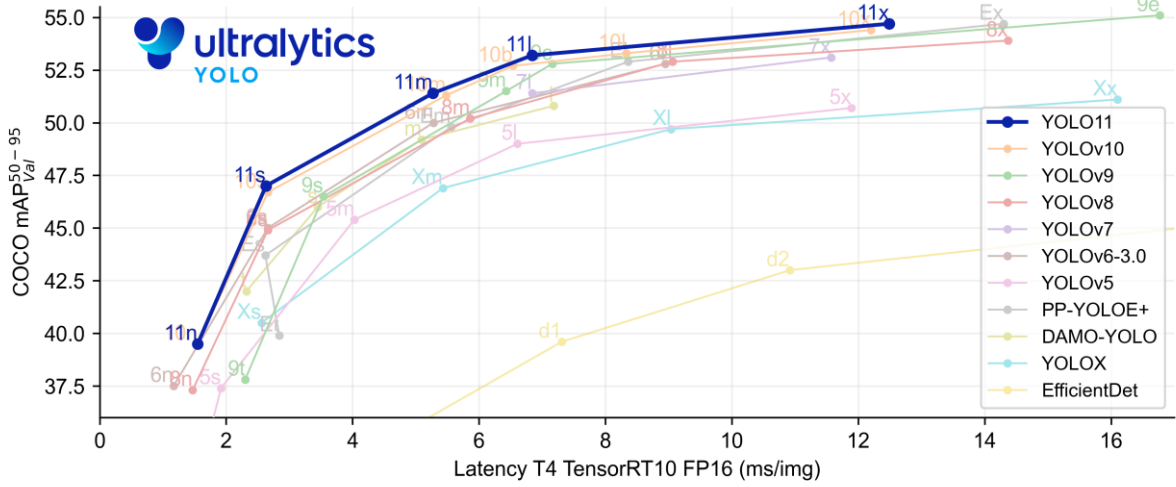
daha küçük çekirdek boyutlarıyla daha hızlı işleme olanak tanırken performanstan ödün vermez (Khanam ve Hussain, 2024).

- Neck: Çok ölçekli bilgileri birleştirir ve Head'e iletir. YOLOv11'de, Spatial Pyramid Pooling - Fast bloğu ve Convolutional Block with Parallel Spatial Attention bileşeni eklenmiştir. Bu yenilikler, modelin farklı boyutlardaki nesnelere tespit etme kabiliyetini artırır (He ve diğ., 2024).
- Head: Nihai tahminleri yapar ve nesne sınıflandırması ile sınır kutusu koordinatlarını çıkarır (He ve diğ., 2024).



Şekil 2.1: Yolov11'in genel mimari yapısı (He ve diğ., 2024)

YOLOv11, önceki sürümlerden %22 daha az parametre kullanırken, COCO veri setinde mAP (mean Average Precision) performansını önemli ölçüde artırmıştır. Model, SGD yerine AdamW optimizatörüyle eğitilmiştir ve öğrenme oranı ayarı için cosine annealing stratejisi uygulanmıştır. AdamW, geleneksel Adam algoritmasından türetilmiştir ancak ağırlık çürümesi adı verilen bir yöntemi içerir. Ağırlık çürümesi, modelin aşırı öğrenmesini engelleyerek genelleştirme performansını artırır. Bu optimizatör, öğrenme oranlarını dinamik olarak ayarlayarak modelin farklı aşamalarında daha verimli öğrenmesini sağlar (Khanam ve Hussain, 2024).



**Şekil 2.2:** YOLOv11 modelinin eski modellerle performans karşılaştırılma grafiği (Ultralytics, 2024.)

Model, özütleme sırasında Swish ve SiLU (Sigmoid Linear Unit) gibi doğrusal olmayan aktivasyon fonksiyonlarını kullanır. Bu fonksiyonlar, modelin doğruluk ve verimliliğini artırır (He ve diğ., 2024). YOLOv11, kayıp fonksiyonunda sınıf olasılıkları, sınır kutusu regresyonu ve dağıtılmış odak kaybını birleştirir. Bu, modelin küçük nesnelere daha hassas bir şekilde algılamasını sağlar (He ve diğ., 2024). YOLOv11, sınır kutularının doğruluğunu artırmak için Generalized Intersection Over Union (GIoU) metriğini kullanır. Bu metrik, tahmin edilen kutular ile gerçek kutular arasındaki örtüşme oranını ölçerek yanlış pozitifleri azaltır (Khanam ve Hussain, 2024)

Modelin özellikle küçük nesnelere tespit etme konusundaki üstün başarısı (Khanam ve Hussain, 2024), baş ve boyun bölgesinde yer alan birçok küçük yapının doğru bir şekilde tespit edilmesi açısından büyük bir avantaj sunmaktadır. Bu durum, baş ve boyun bölgesindeki organ-at-risk (OAR) segmentasyonunda küçük ve karmaşık yapıların doğru bir şekilde tespit edilmesi için önemli bir avantaj sağlamaktadır. YOLOv11'in kullanımı, baş ve boyun bölgesinde kritik organların doğru bir şekilde lokalize edilmesini mümkün kılacaktır.

## SAM2

Segment Anything Model 2 (SAM2), Meta AI tarafından geliştirilen ve ilk versiyonu olan SAM modelinin üzerine inşa edilen bir yapay zeka modelidir. Bu model, hem statik görüntüler hem de video kareleri üzerinde segmentasyon yapabilme yeteneğine sahiptir. SAM2, özellikle etkileşimli segmentasyon yetenekleri ve geniş veri seti desteği ile dikkat çekmektedir (Ravi ve

diğ., 2024). SAM2'nin geliştirilmesi sırasında, video ve görüntü segmentasyonu arasındaki farklar dikkate alınmıştır. Örneğin, videolardaki hareket, deformasyon ve ışık değişiklikleri gibi dinamik unsurlar, modelin bellek tabanlı mimarisiyle ele alınmıştır. Bu, SAM2'nin özellikle medikal görüntüleme gibi hassasiyet gerektiren alanlarda güçlü bir araç olmasını sağlamıştır (He ve diğ., 2024). SAM2'nin mimarisi, görüntü işleme ve segmentasyon için çok sayıda yenilikçi bileşen içerir. Bu bileşenler, modelin farklı görevlerde yüksek performans göstermesini sağlamaktadır. Modelin temel bileşenleri arasında görüntü kodlayıcı (image encoder), bellek dikkat mekanizması (memory attention) ve mask kodlayıcı (mask decoder) yer alır. Görüntü Kodlayıcı olarak SAM2, Hiera adı verilen çok ölçekli bir mimari kullanır. Bu kodlayıcı, farklı çözünürlüklerdeki özellikleri analiz ederek hem geniş alanları hem de küçük detayları aynı anda algılayabilir. Örneğin, medikal görüntülerde organların kenarlarını hassas bir şekilde ayırt etmede bu kodlayıcı büyük bir avantaj sağlar (Ravi ve diğ., 2024). Modelin bellek dikkat mekanizması, önceki segmentasyonlardan elde edilen bilgileri saklar ve bu bilgileri yeni karelerdeki segmentasyon tahminlerinde kullanır. Özellikle ardışık görüntülerin tutarlılığını sağlamak için bellek bankası, geçmiş karelerden gelen tahminleri saklayarak modelin daha kararlı çalışmasını sağlar. Bu özellik, 3D medikal görüntülerde organların dilimler arasındaki uyumunu korumada oldukça etkilidir (He ve diğ., 2024). Mask Kodlayıcı, kullanıcıdan gelen girdilerle (örneğin bir tıklama, kutu ya da maske) görüntü özelliklerini birleştirerek nihai segmentasyonu oluşturur. SAM2, belirsiz girdilerde bile birden fazla maske üretebilir. Örneğin, üst üste binen organlar gibi karmaşık durumlarda bu özellik, modelin doğruluğunu artırır (Ravi ve diğ., 2024).

SAM2, geniş bir veri seti üzerinde eğitilmiştir. SA-V Veri Seti, 50 bin video ve 35 milyon segmentasyon maskesi içerir ve bu, modelin çok çeşitli görevlerde güçlü bir performans göstermesine olanak tanır. Eğitim sırasında, model üzerinde AdamW optimizatörü ve cosine annealing stratejisi kullanılmıştır. AdamW, ağırlık çürümesi ile modelin aşırı öğrenmesini önlerken cosine annealing, öğrenme oranının dinamik bir şekilde ayarlanmasını sağlar (Ravi ve diğ., 2024). Modelin sıfırdan öğrenme yetenekleri de oldukça dikkat çekicidir. Örneğin, SAM2'nin medikal görüntüler üzerinde, yalnızca minimal ince ayarlarla yüksek doğruluk oranlarına ulaştığı gösterilmiştir. Model, özellikle karaciğer ve dalak gibi düşük kontrastlı organların segmentasyonunda önceki yöntemlere kıyasla önemli ölçüde daha başarılı olmuştur (He ve diğ., 2024).

SAM2, medikal görüntüleme kullanıcıya etkileşimli bir deneyim sunar. Kullanıcı, modelin tahmin ettiği maskeler üzerinde iteratif bir şekilde iyileştirme yapabilir. Organın sınırları hatalı algılandığında kullanıcı birkaç tıklama ile bu segmentasyonu düzeltebilir. Bu özellik, radyologların segmentasyon sürecini hızlandırmakta ve doğruluğunu artırmaktadır (He ve diğ., 2024). Bellek tabanlı mimarisi sayesinde SAM2, dilimler arasındaki tutarlılığı korur. Bu, özellikle 3D verilerde, organların pozisyonlarının ve şekillerinin ardışık dilimlerde doğru bir şekilde işlenmesini sağlar (He ve diğ., 2024).

YOLOv11 ve SAM2'nin birlikte kullanımı, OAR'ların hem hızlı bir şekilde tespit edilmesi hem de hassas bir şekilde segmentlenmesi için ideal bir çözüm sunacağına inanmaktayız. Benzer bir hibrit model, polip segmentasyonu gibi zorlu bir medikal görüntüleme görevinde uygulanmış ve önemli başarılar elde etmiştir. Mansoori ve arkadaşlarının çalışmasında, YOLOv8'in sınırlayıcı kutu tespitiyle SAM 2'nin segmentasyon yeteneklerini birleştiren hibrit bir model geliştirilmiş ve bu model PolypGen veri setindeki video sekansları üzerinde test edilmiştir (Mansoori ve diğ., 2024). Geliştirdikleri model, PolypGen veri setinde gerçekleştirdiği segmentasyon performansı ile mevcut yöntemleri önemli ölçüde geride bırakmıştır. Tablo 1'te görüldüğü üzere, Geliştirdikleri hibrit modelin polip segmentasyonu için sağladığı mDice, mIoU ve kesinlik (Precision) gibi metriklerdeki üstün başarısı, hibrit modellerin medikal görüntüleme alanındaki potansiyelini göstermektedir.

**Tablo 2.1:** PolypGen mevcut yöntemlerin performans karşılaştırması. (Mansoori ve diğ., 2024)

Model	mDice	mIoU	Kesinlik	Duyarlılık	F2
UNet	0.4559	0.4049	0.5762	0.6307	0.4668
UNet++	0.4772	0.4272	0.6269	0.6198	0.4876
ResU-Net++	0.2105	0.1589	0.2447	0.5095	0.2303
MSEG	0.4662	0.4171	0.612	0.6217	0.4757
ColonSegNet	0.3574	0.3058	0.4804	0.5296	0.3533
UACANet	0.4748	0.4155	0.6108	0.6357	0.4886

UNeXt	0.2998	0.2457	0.3661	0.5658	0.3201
TransNetR	0.5168	0.4717	0.7881	0.5777	0.5105
<b>YOLO-SAM 2</b>	<b>0.808</b>	<b>0.678</b>	<b>0.858</b>	<b>0.764</b>	<b>0.781</b>

## 2.2 BAŞ VE BOYUN BÖLGESİ ANATOMİSİ

Baş ve boyun bölgesi, karmaşık bir anatomik yapı ve yoğun bir fonksiyonel çeşitlilik içermektedir. Bu sebeple çok boyutlu ve zorlu bir alandır. Bu karmaşık yapı, bu alanın makine öğrenimi tabanlı tıbbi görüntü analizi, segmentasyon, sınıflandırma ve tanı destek sistemleri geliştirmek için sıklıkla çalışılan bir alan olmasına sebep olmuştur (Taub ve diğ., 2015).

Baş ve boyun bölgesinde bulunan temel yapılar arasında beyin, görme, işitme ve koku alma organları ve yakın konumlanan kafatası kemikleri, duyu sinirleri, büyük damarlar, tükürük bezleri, tiroit bezi, paratiroid bezleri, yüz ve çiğneme kasları ile lenf düğümleri yer alır. Özellikle yapısal çeşitlilikleri, karmaşık topografik yerleşimleri ve dokuların yoğun etkileşimi, bu bölgenin hem cerrahi müdahaleler hem de yapay zekâ temelli görüntü işleme uygulamaları için zorlu ve kritik bir çalışma ortamı haline getirmektedir (Gray ve diğ., 2021). Bu sebeple bu bölgeyle alakalı bir segmentasyon çalışması yapabilmek için bu bölgenin temel anatomik yapısı hakkında bilgiye sahip olmak gerekmektedir.

### **Kafatası ve Yüz Kemikleri:**

Kafatası, beyin dokusunu korurken aynı zamanda yüz kemiklerini destekleyen bir yapı oluşturmaktadır. Kafatasını oluşturan kemikler, oral, nazal ve orbital boşlukların şekillenmesinde temel rol oynarlar. Yüz ve çene kemiklerinden olan, Maksilla ve mandibula, çiğneme ve konuşma fonksiyonlarını destekleyen temel iskelet yapısını oluştururken, zigomatik kemik yüz hatlarının belirginliğini sağlar. Orbital boşluk, gözün yerleştiği ve çevresini koruyan bir kemik çerçeve ile çevrilidir. Bu boşlukları oluşturan kemikler, hem yüz anatomisinin bütünlüğünü sağlar hem de klinik çalışmalarda organların tespiti için sıkça referans alınan anatomik bölgelerdir. (Taub ve diğ., 2015).

### **Kaslar ve Eklem Yapıları:**

Yüz kasları, yüz ifadelerini oluşturan ve mimik hareketlerini sağlayan yüzeysel kas gruplarını içerir. Bu kaslar, deri altına bağlanarak kasılma sırasında yüzeyde belirgin hareketlerin oluşmasını sağlar. Çiğneme kasları ise derin yerleşimli olup mandibula hareketlerini yönlendiren masseter, temporalis, medial ve lateral pterygoid kaslardan oluşur. Temporomandibular eklem, mandibulanın kranial kemiklerle bağlantısını sağlayan kompleks bir eklemdir ve çiğneme, konuşma ve yutma gibi önemli fonksiyonları destekler. TMJ, eklem kapsülü, disk yapısı ve bağ dokularıyla bir arada çalışarak mandibular hareketlerin geniş bir yelpazede gerçekleştirilmesine olanak tanıyan yapılardır. Bu yapıların bütünlüğü, hem fonksiyonel hareketlerin sağlanması hem de estetik görünümün korunması açısından kritik bir öneme sahiptir. Bu yapıların göreceği herhangi bir hasar çiğneme ve konuşma gibi temel hayati fonksiyonlara zarar verebilmektedir. (Taub ve diğ., 2015).

### **Tiroid, Paratiroid ve Tükürük Bezleri:**

Boynun ön bölgesinde bulunan tiroid bezi, vücudun metabolik hızını düzenleyen tiroksin ve triiyodotironin hormonlarını salgılar. Tiroid bezinin hemen arkasında yer alan paratiroid bezleri ise kalsiyum ve fosfor metabolizmasında kritik rol oynayan parathormon salgılar (Taub ve diğ., 2015). Tiroid bezinin hasar görmesi, vücudun metabolik dengesini etkileyerek Guatr ve Haşimato gibi hastalıkların ortaya çıkmasına neden olabilir. (Watkinson ve Clarke, 2018).

Tükürük bezleri ise ağız ve çevresinin nemli tutulmasını ve sindirim enzimlerinin salgılanmasını sağlar. Parotis, submandibular ve sublingual bezler, bu bölgedeki başlıca büyük tükürük bezleridir. Bu bezlerin hasar görmesi inflamasyon (sialadenit), tümör oluşumu veya tükürük kanalı tıkanıklıkları gibi klinik durumların ortaya çıkmasına neden olabilir.

### **Damar, Sinir ve Lenf Yapıları:**

Boyun bölgesinde yer alan karotis arterleri ve vertebral arterler, beyne ve baş bölgesine kan akışı sağlayan temel damar yapılarıdır. Juguler venler, kanın beyinden ve yüz bölgesinden geri dönüşünü sağlayarak venöz dolaşımın önemli bir parçasını oluşturur. Lenf düğümleri ise bağışıklık fonksiyonlarının düzenlenmesinde kritik bir rol oynar ve baş-boyun bölgesindeki enfeksiyonlar veya malignite durumlarında sıkça değerlendirilir (Gray ve diğ., 2021).

Kraniyal sinirler, özellikle vagus ve aksesuar sinirler, boyun bölgesindeki duyuşal ve motor fonksiyonları düzenlerken, servikal pleksus omuz ve boyun kaslarını innerve eder. Bu sinir yapılarının ayrıntılı haritalanması, nörolojik ve cerrahi planlamalarda büyük öneme sahiptir. Manyetik rezonans traktografi gibi ileri görüntüleme yöntemleri, sinir liflerinin izlenmesi ve rekonstrüksiyonu için kullanılmaktadır. Kemoterapi veya radyoterapi sırasında bu bölgelerin hasar görmesi, yorgunluk, mide bulantısı, tükürük bezlerinin kuruması veya yutma güçlüğü gibi yan etkilere neden olabilir. (Taub ve diğ., 2015).

### **Lenf Düğümleri ve Fasiyal Boşluklar:**

Boyun bölgesinde yer alan lenf düğümleri, bağışıklık sisteminin bir parçası olarak enfeksiyon ve malign hücrelerin drenaj ve filtrasyonunda önemli bir role sahiptir. Bu düğümler, özellikle derin servikal, yüzeysel servikal, submandibular ve pretrakeal gruplara ayrılır (Taub ve diğ., 2015). Baş ve boyun bölgesindeki lenf düğümleri, çevresel dokulardan gelen lenfi toplar ve hastalık süreçlerinin (enfeksiyon veya tümör metastazı gibi) değerlendirilmesinde kritik bir gösterge olarak kullanılır. Lenfatik drenaj yolları, özellikle baş ve boyundaki kanserlerin yayılım yollarını anlamak için anatomik açıdan büyük öneme sahiptir (Watkinson ve Clarke, 2018).

Fasiyal boşluklar ise baş ve boyun anatomisinin derin tabakalarında yer alan bağ dokusu ile çevrili bölgelerdir. Bu boşluklar arasında retrofarengeal boşluk, submandibular boşluk, parafarengeal boşluk ve viseral fasiyal boşluklar yer alır. Bu yapılar, enfeksiyonların derin dokulara yayılmasını kolaylaştıran potansiyel yollar oluşturur. Özellikle retrofarengeal ve parafarengeal boşluklar, enfeksiyonların mediastene ilerleme riski nedeniyle kritik kabul edilir. Bu bölgelerin hasar görmesi, lenf düğümlerinde şişme, ağrı, enfeksiyon riskinin artması veya lenfatik drenajın bozulması gibi yan etkilere neden olabilir. (Taub ve diğ., 2015).

### **2.3 RİSK ALTINDAKİ ORGANLAR (OAR) VE ÖNEMİ**

Organ-at-risk (OAR), radyoterapi tedavisi sırasında radyasyonun yan etkilerine maruz kalabilecek kritik organları ifade etmekte kullanılan genel tıbbi terimdir. Bu organlar, tümörlerin yakınında bulunan, radyoterapi sırasında maksimum korunması gereken dokulardır (Podobnik ve diğ., 2023a). Radyoterapi planlamasında OAR'ların doğru bir şekilde tanımlanması, hem tedavi etkinliğinin artmasını sağlamak hem de çevredeki sağlıklı dokuların korunmasına yardımcı olmak açısından büyük önem taşımaktadır (Özdemir, 2013).

Baş ve boyun bölgesindeki OAR'ların segmentasyonu, bu bölgedeki kanserlerin tedavisinde radyoterapi ve kemoterapi planlamasının önemli bir bileşenidir. Bu organların doğru segmentasyonu, radyasyon ve ilaç dozunun hassas bir şekilde ayarlanmasına ve kritik organların korunmasına olanak tanımaktadır, bu da hastaların yaşam kalitesini yükseltmekte ve yan etkiye maruz kalma riskini azaltmaktadır (Podobnik ve diğ., 2023a). Bununla sebeple, OAR segmentasyonu sürecinde elde edilen verilerin klinik olarak doğru olması, tedavi uygulanacak bölgelerin planlarının doğruluğu için kritik öneme sahiptir (Podobnik ve diğ., 2023b).

Bu çalışmada kullanılacak veri setinde, her bir hasta için 3B çok modlu segmentasyon maskeleri şeklinde 30'a kadar OAR (çift OAR'ları sol ve sağ olarak tek bir OAR kabul edilirse 22) referans anotasyonları sunulmuştur. (Podobnik ve diğ., 2023a) Bu organlar, Amerikan Tıp Fizikçileri Derneği'nin (AAPM) Task Group 263 (TG-263) tarafından önerilen adlandırma kurallarına göre adlandırılmıştır (Mayo ve diğ., 2018).

Segmentasyonu yapılan OAR'lar şunlardır:

- **Arytenoid** (Arytenoid kıkırdaklar)
- **Beyin Sapı** (Brainstem)
- **Karotid Arteri** (A\_Carotid; sol/sağ: A\_Carotid\_L/A\_Carotid\_R)
- **Servikal Özofagus** (Esophagus\_S)
- **Koklea** (Cochlea; sol/sağ: Cochlea\_L/Cochlea\_R)
- **Krikofaringeal Giriş** (Cricopharyngeus)
- **Gözyaşı Bezi** (Glnd\_Lacrimal; sol/sağ: Glnd\_Lacrimal\_L/Glnd\_Lacrimal\_R)
- **Gırtlak - Glottis** (Glottis)
- **Gırtlak - Supraglottik** (Larynx\_SG)
- **Dudaklar** (Lips)
- **Mandibula** (Bone\_Mandible)
- **Optik Kiazma** (OpticChiasm)
- **Optik Sinir** (OpticNrv; sol/sağ: OpticNrv\_L/OpticNrv\_R)
- **Ağız Boşluğu** (Cavity\_Oral)
- **Parotis Bezi** (Parotids; sol/sağ: Parotid\_L/Parotid\_R)
- **Hipofiz Bezi** (Pituitary)
- **Omurilik** (SpinalCord)
- **Submandibular Bez** (Glnd\_Submands; sol/sağ: Glnd\_Submand\_L/Glnd\_Submand\_R)

- **Tiroid Bezi** (Glnd\_Thyroid)
- **Bukkal Mukoz** (BuccalMucosa)
- **Göz Bebeklerinin Ön Segmenti** (Eye\_A; sol/sağ: Eye\_AL/Eye\_AR)
- **Göz Bebeklerinin Arka Segmenti** (Eye\_P; sol/sağ: Eye\_PL/Eye\_PR)

Bu OAR'ların doğru şekilde belirlenmesi ve korunması, radyoterapi sırasında hastaların yan etkilerden korunması ve tedavi etkinliğinin artırılması için önemlidir (Gören ve diğ., 2021)

## 2.4 TIBBİ GÖRÜNTÜLEME YÖNTEMLERİ

Tıbbi görüntüleme, klinik tanı ve tedavi planlamasında kritik bir rol oynar. Bu amaçla kullanılan çeşitli görüntüleme teknikleri, farklı modalitelerde detaylı bilgi sağlayarak, hastalık tanısı ve tedavi süreçlerinin optimizasyonunu sağlamaktadır. Bu bölümde, en yaygın tıbbi görüntüleme tekniklerinden MR, CT ve PET ile ilgili bilgiler detaylandırılmıştır.

### 2.4.1. Manyetik Rezonans Görüntüleme (MR)

Manyetik Rezonans Görüntüleme (MR), vücuttaki yumuşak dokuların ayrıntılı olarak incelenmesini sağlayan, iyonize radyasyon içermeyen, non-invaziv bir tıbbi görüntüleme yöntemidir. MR teknolojisi, güçlü bir manyetik alan ve radyo dalgaları kullanarak hidrojen atomlarının manyetik rezonansa girmesini sağlar. Hidrojen atomlarının saldıđı sinyaller, radyo frekans bobinleri tarafından algılanarak görüntüye dönüştürülür (Lauterbur, 1973). MR, özellikle beyin, kas dokuları, kalp ve kanser taramaları gibi klinik uygulamalarda yüksek doku kontrastı sağlaması nedeniyle yaygın olarak kullanılmaktadır.

MR görüntülerinin kaydedilmesi ve iletilmesi için genellikle Digital Imaging and Communications in Medicine (DICOM) standardı kullanılmaktadır. DICOM formatı, MR görüntülerinin çok katmanlı (multi-slice) olarak saklanmasına olanak tanımaktadır (Pianyk, 2012). Bu görüntüler, genellikle gri tonlamalı (grayscale) olarak elde edilmekte olup kontrast iyileştirme, gürültü azaltma ve ölçekleme gibi ön işleme tekniklerine tabi tutulabilmektedirler (Gonzalez et al., 2009). Bu ön işleme adımları, görüntü segmentasyonu ve derin öğrenme tabanlı analizlerde yüksek doğruluk elde edilmesi açısından büyük önem taşımaktadır.

MR görüntülerinin analizinde son yıllarda derin öğrenme modelleri yaygın olarak kullanılmaktadır. U-Net (Ronneberger et al., 2015) gibi konvolüsyonel sinir ağı (CNN) tabanlı modeller, biyomedikal görüntü segmentasyonu alanında büyük başarı göstermiştir. U-Net'in

simetrik kodlayıcı-çözücü (encoder-decoder) yapısı, hem düşük seviyeli hem de yüksek seviyeli özelliklerin etkili bir şekilde öğrenilmesine olanak tanımaktadır. MR görüntülerinde segmentasyon performansını artırmak için transfer öğrenme (transfer learning) ve veri artırma (data augmentation) gibi teknikler de yaygın olarak kullanılmaktadır. (Litjens et al., 2017).

#### 2.4.2. Bilgisayarlı Tomografi

Bilgisayarlı Tomografi (CT), vücudun kesitsel görüntülerini oluşturmak için X-ışınları kullanarak detaylı anatomik bilgilerin elde edilmesini sağlayan bir tıbbi görüntüleme yöntemidir (Hounsfield, 1973). CT görüntüleme, özellikle kemik yapıların ve sert dokuların detaylı olarak incelenmesinde kullanılır. CT, x-ışınlarının vücut içerisinden geçmesi ve farklı dokuların bu ışınları farklı oranlarda absorbe etmesi prensibine dayanır. Bu sayede, vücudun farklı katmanlarından alınan görüntüler birleştirilerek üç boyutlu bir görüntü elde edilir.

CT görüntüleri de DICOM formatında saklanmakta ve çoğu zaman Hounsfield birimleri cinsinden temsil edilmektedir. Hounsfield birimi, dokuların x-ışını absorpsiyon oranını temsil eder ve bu sayede çeşitli dokular arasındaki yoğunluk farklılıkları belirginleştirilir (Kalender, 2000). CT görüntüleri genellikle gri seviye formatındadır (Fitzpatrick ve Sonka, 2009). CT verilerinin işlenmesinde, derin öğrenme modelleri ile birlikte hacimsel verilerin kullanımı önem kazanmıştır. Bu veriler, üç boyutlu konvolüsyonel sinir ağları kullanılarak işlenebilir ve bu da daha derinlemesine analiz ve daha doğru sonuçlar elde edilmesini sağlar (Çiçek ve diğ., 2016).

CT görüntülerinin işlenmesinde, derin öğrenme tabanlı modeller yaygın olarak kullanılmaktadır. Özellikle, Transformer tabanlı modeller son yıllarda tıbbi görüntü işleme alanında popüler hale gelmiştir. (Litjens ve diğ., 2017) Swin Transformer, CT görüntülerinin özellikle yüksek çözünürlüklü veri setlerinde işlenmesi sırasında öne çıkan modellerden biridir (Liu ve diğ., 2021) Bu model, görüntülerin farklı ölçeklerde işlenmesine olanak tanıyarak detayların daha iyi yakalanmasını sağlamaktadır. CT görüntülerinin otomatik segmentasyonunda bu tür modellerin kullanılması hem zaman hem de maliyet açısından önemli avantajlar sunmaktadır.

### 2.4.3. Pozitron Emisyon Tomografisi

Pozitron Emisyon Tomografisi (PET), vücuttaki biyokimyasal ve fizyolojik fonksiyonları görüntülemek için radyoaktif izleyiciler kullanılan bir tıbbi görüntüleme yöntemidir. PET, kanser taramalarında, kalp hastalıklarında ve beyin fonksiyonlarının incelenmesinde yaygın olarak kullanılmaktadır (Phelps, 2004). PET görüntüleme, radyoaktif bir izotopun vücuda enjeksiyonuyla başlar ve bu izotopun yaydığı pozitronların detektörler tarafından algılanması ile görüntüler oluşturulur.

PET görüntüleri, diğer tıbbi görüntüleme yöntemlerinden farklı olarak metabolik aktiviteleri temsil eden renkli görüntüler sunar. PET görüntüleri, radyofarmasötiklerin vücuttaki dağılımına göre oluşturulur ve genellikle diğer modalitelerle (örneğin CT) birleştirilerek daha kapsamlı bir klinik bilgi sağlar (Cherry ve diğ., 2012). Bu kombinasyon, özellikle kanser tanısında tümörlerin hem anatomik hem de metabolik özelliklerinin bir arada değerlendirilmesini mümkün kılar.

### 2.5. ÇOK MODLU GÖRÜNTÜLEME (MULTIMODAL IMAGING)

Çok modlu görüntüleme, farklı tıbbi görüntüleme tekniklerinden elde edilen verilerin birleştirilmesiyle elde edilen, daha kapsamlı ve güvenilir bir klinik bilgi sunan bir yaklaşımdır. Bu yöntem, özellikle baş ve boyun bölgesindeki kanserlerin tedavisinde kullanılarak hem anatomik hem de fonksiyonel bilgilerin bir arada değerlendirilmesini mümkün kılar (Korte ve diğ., 2021). Örneğin, Bilgisayarlı Tomografi (CT) ile Manyetik Rezonans Görüntüleme (MR) birleştirilerek, kemik yapılar ile yumuşak dokuların detaylı ve doğru bir şekilde analiz edilmesi sağlanabilir (Klein ve diğ., 2010).

CT ve MR gibi modalitelerden elde edilen veriler, farklı avantajlar sunar. CT görüntüleri kemik yapıların net bir şekilde görüntülenmesini sağlarken, MRI yumuşak dokuların daha ayrıntılı olarak değerlendirilmesine olanak tanır. Bu iki modalitenin birlikte kullanılması, özellikle baş ve boyun kanserlerinin tedavi planlamasında, organ-at-risk (OAR) olarak adlandırılan hassas organların doğru bir şekilde tanımlanmasını ve korunmasını sağlar (Podobnik ve diğ., 2023b). Benzer şekilde, PET (Pozitron Emisyon Tomografisi) gibi fonksiyonel görüntüleme teknikleri de CT ve MR ile birleştirilerek, tümörlerin metabolik aktivitesinin ve anatomik yapısının aynı anda incelenmesi sağlanabilir (Phelps, 2004)

Çok modlu görüntülemenin en önemli zorluklarından biri, farklı modalitelerden elde edilen görüntülerin geometrik olarak hizalanması ve veri heterojenliğinin giderilmesidir. Bu amaçla, görüntü kaynaştırma ('image fusion') ve kayıt ('image registration') teknikleri kullanılır. Görüntü kaynaştırma, farklı modalitelerden elde edilen görüntülerin tek bir ortak görüntüde birleştirilmesi sürecidir. Bu süreç, tıbbi görüntülerin hem anatomik hem de fonksiyonel bilgilerini bir araya getirerek daha zengin bir klinik bilgi sunar (James ve Dasarathy, 2014). Görüntü kaydı ise farklı modalitelerden gelen verilerin hizalanmasını sağlayarak bu verilerin aynı referans çerçevesinde karşılaştırılmasına olanak tanır (Pluim ve diğ., 2003).

Çok modlu görüntüleme verilerinin işlenmesi sırasında, derin öğrenme ve makine öğrenmesi teknikleri yaygın olarak kullanılmaktadır. Özellikle, konvolüsyonel sinir ağları (CNN) ve Transformer tabanlı modeller, farklı modalitelerden elde edilen verilerin işlenmesi ve analiz edilmesinde başarılı sonuçlar vermektedir (Litjens ve diğ., 2017). Örneğin, Swin Transformer gibi modeller, farklı modalitelerden elde edilen yüksek çözünürlüklü görüntülerin etkili bir şekilde işlenmesini ve sınıflandırılmasını sağlar (Liu ve diğ., 2021). Bunun yanı sıra, Çekişmeli üretken ağlar (GANs) gibi derin öğrenme modelleri, düşük çözünürlüklü görüntülerin iyileştirilmesi ve farklı modalitelerin birbirine dönüştürülmesi gibi işlemler için kullanılmaktadır (Goodfellow ve diğ., 2014).

Çok modlu görüntüleme ayrıca adaptif radyoterapi (ART) gibi klinik uygulamalarda da önemli bir rol oynamaktadır. ART, tedavi süresince hastanın anatomisindeki değişikliklere uyum sağlayarak tedavi planlarının güncellenmesini sağlar ve bu süreçte çok modlu görüntüleme verileri tedavi planlamasının doğruluğunu artırır (Gören ve diğ., 2021). Özellikle baş ve boyun bölgesindeki OAR'ların korunması açısından, farklı modalitelerden elde edilen bilgilerin bir arada değerlendirilmesi büyük önem taşımaktadır.

Çok modlu görüntüleme sürecinde karşılaşılan bir diğer zorluk ise veri heterojenliğidir. Farklı modalitelerden gelen görüntüler, farklı çözünürlüklerde, kontrast oranlarında ve gürültü seviyelerinde olabilir. Bu nedenle, bu görüntülerin derin öğrenme modellerine uygun hale getirilmesi için ön işleme teknikleri uygulanması gerekmektedir. Bu teknikler arasında kontrast iyileştirme, gürültü azaltma ve normalizasyon gibi işlemler yer almaktadır (Wang ve diğ.,

2020). Bu ön işleme adımları, çok modlu görüntüleme verilerinin uyumlu hale getirilmesi ve bu verilerden en doğru şekilde faydalanılmasını sağlamaktadır.

## 2.6. TIBBİ GÖRÜNTÜ SEGMENTASYONU

### 2.6.1. Geleneksel Yöntemler

Tıbbi görüntü segmentasyonu, tıbbi görüntülerde ilgi alanlarının belirlenmesi ve sınırlarının belirginleştirilmesi işlemidir. Geleneksel segmentasyon yöntemleri, görüntü işleme tekniklerine dayanan algoritmalar kullanılarak gerçekleştirilir. Bu yöntemler genellikle eşikleme ('thresholding'), kenar tespiti ('edge detection') ve bölütleme ('region growing') gibi basit algoritmaları içerir (Gonzalez ve diğ., 2009). Bu tekniklerin avantajı, kolay uygulanabilir olmaları ve hızlı sonuç vermeleridir. Ancak, özellikle karmaşık anatomik yapılar veya düşük kontrastlı görüntülerde sınırlı performans sergileyebilirler.

**Eşikleme Yöntemi:** Eşikleme yöntemi, görüntüdeki pikselleri belirli bir eşik değerine göre sınıflandırarak ilgi alanlarının belirlenmesini sağlar (Otsu, 1979). Bu yöntem, özellikle homojen bölgeler için uygundur ancak tıbbi görüntülerde farklı doku türleri arasında kontrast düşüklüğü olduğunda sınırlı kalabilir. Basit bir eşikleme algoritması, belirli bir yoğunluk seviyesinin üzerindeki tüm pikselleri bir sınıfa atarken, geri kalanları diğer sınıfa atar. Bu yöntem, özellikle CT görüntüleri gibi belirli anatomik yapılar için kullanışlı olabilir. Ancak, gürültü ve düşük kontrast durumlarında hassasiyetinin düşük olması nedeniyle genellikle daha karmaşık yöntemlerle birleştirilir.

**Kenar Tespiti Yöntemleri:** Kenar tespiti yöntemleri ise, görüntüdeki ani yoğunluk değişimlerini tespit ederek sınırları belirlemeyi amaçlar (Canny, 1986) Bu yöntemler, anatomik sınırların belirlenmesinde kullanılabilir ancak gürültüye karşı hassas oldukları için segmentasyon doğruluğunu olumsuz etkileyebilirler. Örneğin, Sobel, Prewitt ve Canny kenar tespit algoritmaları, tıbbi görüntülerde sıkça kullanılan kenar tespit yöntemlerindedir. Canny kenar tespit algoritması, düşük ve yüksek eşik değerleri kullanarak gürültüyü azaltma ve hassas sınır tespiti yapma konusunda başarılıdır. Ancak, yine de gürültüden etkilenebileceği ve bazen istenmeyen yapay hatalara neden olabileceği unutulmamalıdır.

**Bölütleme (Region Growing):** Bölütleme yöntemi, başlangıç noktası olarak belirlenen bir veya daha fazla pikselden yola çıkarak benzer özelliklere sahip komşu piksellerin gruplandırılmasını sağlar. Bu yöntem, özellikle belirgin doku sınırlarına sahip görüntülerde iyi sonuçlar verebilir (Adams ve Bischof, 1994). Bölütleme işlemi, piksel yoğunluklarının benzerliklerine dayalı olarak genişletilir ve bu nedenle başlangıç noktasının seçimi, sonucun doğruluğunu etkiler. Yanlış başlangıç noktaları, yanlış segmentasyonlara ve zaman kaybına yol açabilir. Bu yöntem genellikle interaktif bir süreç gerektirir ve manuel müdahale ihtiyacı nedeniyle büyük veri setlerinde verimli olmayabilir.

**K-Means ve Yoğunluk Tabanlı Kümeleme:** K-Means ve diğer kümeleme algoritmaları, tıbbi görüntülerde farklı doku türlerini tanımlamak için kullanılabilir. Bu yöntemler, görüntüdeki pikselleri belirli özelliklere göre gruplandırır ve bu sayede benzer dokuların sınıflandırılmasını sağlar. Ancak, bu yöntemler, önceden belirlenmiş küme sayısına ihtiyaç duyduğundan ve gürültüye duyarlı olduğundan tıbbi görüntülerdeki karmaşık yapılar için sınırlı performans gösterebilir.

Geleneksel yöntemler, tıbbi görüntü segmentasyonunda bazı durumlarda başarılı olsa da, bu yöntemlerin manuel olarak ayarlanması ve gözlemciye bağlı olması, özellikle büyük veri setlerinde zaman alıcı ve yorucu olabilir. Ayrıca, düşük kontrastlı ve karmaşık anatomik yapılara sahip görüntülerde bu yöntemler genellikle yetersiz kalmaktadır.

### **2.6.2. Derin Öğrenme Tabanlı Yöntemler (Deep Learning-Based Methods)**

Son yıllarda, derin öğrenme tabanlı yöntemler tıbbi görüntü segmentasyonu alanında büyük ilerleme kaydetmiştir. Bu yöntemler, geleneksel yöntemlerin aksine, görüntülerdeki özellikleri otomatik olarak öğrenerek daha doğru ve tutarlı segmentasyon sonuçları elde etmeyi mümkün kılar. Derin öğrenme tabanlı yöntemlerin en yaygın kullanılanı, konvolüsyonel sinir ağları (CNN) tabanlı U-Net modelidir (Ronneberger ve diğ., 2015). U-Net, özellikle biyomedikal görüntü segmentasyonu için geliştirilmiş bir mimari olup, kodlayıcı-çözücü yapısı sayesinde hem düşük seviyeli hem de yüksek seviyeli özellikleri etkili bir şekilde öğrenir.

**Konvolüsyonel Sinir Ağları (CNN) ve U-Net:** CNN tabanlı modeller, görüntüdeki önemli özellikleri otomatik olarak öğrenme yeteneği sunar. U-Net, kodlayıcı-çözücü yapısı ve atlama

bağlantıları (skip connections) adı verilen bağlantılar sayesinde, görüntüdeki hem düşük hem de yüksek seviyeli özellikleri yakalar. Bu, özellikle karmaşık anatomik yapıları ve küçük organları segmentlemek için faydalıdır. U-Net'in başarısı, tıbbi görüntü segmentasyonundaki yüksek doğruluk oranlarıyla kanıtlanmıştır ve bu nedenle hala yaygın olarak kullanılmaktadır (Ronneberger ve diğ., 2015).

**3D U-Net ve V-Net:** Özellikle 3 boyutlu tıbbi görüntülerin segmentasyonu için 3D U-Net ve V-Net gibi modeller geliştirilmiştir (Çiçek ve diğ., 2016). Bu modeller, volumetrik verileri işleyerek 3D segmentasyon sağlar ve bu sayede özellikle baş ve boyun bölgesi gibi karmaşık anatomik yapıların daha doğru bir şekilde analiz edilmesine olanak tanır. 3D segmentasyon, anatomik yapıların uzaysal ilişkilerini koruyarak daha hassas sonuçlar elde edilmesini sağlar.

**Residual Networks (ResNet):** ResNet gibi daha derin ağlar, katmanlar arasındaki bilgi kaybını azaltmak için artık bağlantılar (residual connections) kullanır (He ve diğ., 2016). Bu bağlantılar, derin ağların daha etkili bir şekilde öğrenmesini sağlar ve bu nedenle daha karmaşık tıbbi görüntü verileri üzerinde başarılı segmentasyon sonuçları elde edilebilir. ResNet tabanlı modeller, özellikle çok sayıda katmana sahip derin sinir ağlarında veri kaybını ve aşırı uyumu önleyerek daha stabil bir eğitim süreci sunar.

**Transformer Tabanlı Modeller:** Transformer tabanlı yaklaşımlar da son zamanlarda tıbbi görüntü segmentasyonunda öne çıkmaktadır. Swin Transformer gibi modeller, özellikle büyük ve yüksek çözünürlüklü tıbbi görüntülerin işlenmesinde başarılı sonuçlar vermektedir (Liu ve diğ., 2021). Transformer tabanlı modeller, dikkat mekanizmaları sayesinde görüntülerdeki uzun mesafeli ilişkileri daha iyi öğrenebilir ve bu sayede daha hassas segmentasyon sonuçları elde edebilir. Özellikle baş-boyun bölgesindeki organların karmaşık anatomik yapılarında bu tür uzun menzilli bağlantıları yakalayabilmek önemlidir.

**Çekişmeli üretken ağlar (GANs):** GAN'ler, tıbbi görüntü segmentasyonunda kullanılan bir diğer derin öğrenme modelidir (Goodfellow ve diğ., 2014). GAN'ler, görüntülerin daha doğru ve gerçekçi bir şekilde segmentasyonunu sağlamak için iki farklı sinir ağı kullanır: biri sahte görüntüler üreten bir "üretici" ağ, diğeri ise bu görüntüleri gerçek görüntülerden ayırt etmeye çalışan bir "ayırıcı" ağ. Bu yöntem, özellikle sınırlı veri setleri üzerinde veri çeşitliliğini

artırmak ve modelin genelleme yeteneğini geliştirmek için kullanılır. GAN tabanlı segmentasyon, veri zenginleştirme ('data augmentation') amacıyla da kullanılarak modelin daha çeşitli veri üzerinde eğitilmesine katkıda bulunur.

**Attention U-Net ve Dikkat Mekanizmaları:** Attention U-Net, geleneksel U-Net modeline dikkat mekanizmalarının eklenmesiyle geliştirilmiş bir modeldir (Oktay ve diğ., 2018). Bu mekanizmalar, modelin sadece ilgili bölgelere odaklanmasını sağlayarak segmentasyonun doğruluğunu artırır. Özellikle tıbbi görüntülerde, belirli organların veya yapıların diğerlerinden ayrılması gerektiğinde dikkat mekanizmaları büyük avantaj sağlar. Örneğin, baş ve boyun bölgesindeki küçük OAR'ların segmentasyonu sırasında bu mekanizmalar modelin hassasiyetini artırabilir.

**Transfer Öğrenme ve İnce Ayar (Fine-Tuning):** Derin öğrenme modelleri, genellikle büyük veri setleri üzerinde eğitilir ve daha sonra belirli bir tıbbi görüntüleme problemine uyarlanır. Transfer öğrenme, önceden eğitilmiş bir modelin yeniden kullanılması ve bu modele yeni bir görev için ince ayar yapılması sürecidir. Bu, özellikle sınırlı sayıda etiketlenmiş veriye sahip tıbbi görüntüleme problemlerinde oldukça kullanışlıdır (Yosinski ve diğ., 2014). Bu yöntem, eğitimin daha hızlı ve daha doğru bir şekilde tamamlanmasını sağlar.

Derin öğrenme tabanlı yöntemler, özellikle büyük veri setleri üzerinde eğitildiklerinde oldukça yüksek doğruluk oranlarına ulaşabilirler. Bu modeller, görüntülerdeki karmaşık yapıları ve ilişkileri otomatik olarak öğrenebilir ve bu sayede segmentasyon işlemlerinde gözlemciye bağlı hataları en aza indirebilirler (Litjens ve diğ., 2017). Derin öğrenme yöntemlerinin sunduğu otomatik özellik çıkarma ve öğrenme yeteneği, tıbbi görüntü segmentasyonunu daha hassas ve verimli hale getirerek klinik karar destek sistemlerinde önemli bir rol oynamaktadır.

## 2.7. BAŞ VE BOYUN BÖLGESİ SEGMENTASYONUNDA MEVCUT YAKLAŞIMLAR

### 2.7.1. Baş ve Boyun Bölgesinde Tıbbi Görüntüleme ve Derin Öğrenme

Baş ve boyun bölgesindeki kanser vakalarında radyoterapi planlaması için kritik organların (Organ at Risk, OAR) doğru segmentasyonu hayati öneme sahiptir. Son yıllarda derin öğrenme, bu tıbbi görüntü segmentasyonu görevlerinde en başarılı yöntem olarak öne çıkmıştır. Özellikle

beyin, göz, omurilik, tükürük bezleri, tümör bölgesi gibi yapıların CT ve MR görüntülerde otomatik olarak işaretlenmesi için derin öğrenme tabanlı çözümler geliştirilmektedir. Bu alanda genellikle U-Net ve türevleri gibi evrişimli sinir ağı mimarileri kullanılmaktadır. U-Net, biyomedikal segmentasyon için özel geliştirilmiş bir CNN mimarisi olup “encoder-decoder” yapısı ve skip connection bağlantılarıyla az veride dahi yüksek doğruluklu sonuçlar verebilmektedir (Ronneberger ve diğ., 2015). Baş-boyun bölgesi segmentasyon problemlerinde de U-Net ve genişletilmiş versiyonları (ör. 3D U-Net, Attention U-Net, nnU-Net) yaygın olarak uygulanmıştır. Derin öğrenme modelleri, uzmanlarca saatler alabilen manuel segmentasyon işini saniyeler mertebesinde otomatik yapabilme potansiyeli taşıdığı için klinik iş akışına entegre olma yolundadır.

Baş-boyun bölgesinde derin öğrenme yöntemlerinin objektif olarak değerlendirilebilmesi için yakın zamanda kapsamlı bir veri seti ve yarışma yayınlanmıştır: HaN-Seg: Head and Neck Organ-at-Risk CT & MR Segmentation Dataset. Podobnik ve arkadaşlarının 2023’te Medical Physics dergisinde tanıttığı bu açık veri seti, aynı hastalardan alınmış eşleşik CT ve MR taramalarını ve bunlara ait elle çizilmiş OAR segmentasyonlarını içermektedir (Podobnik ve diğ., 2023). HaN-Seg veri seti 56 hastanın görüntülerini kapsamaktadır; her bir hasta için radyologlar tarafından tanımlanmış 30’a kadar organ yapısı dilim dilim segmentasyon maskesi olarak sunulmuştur.

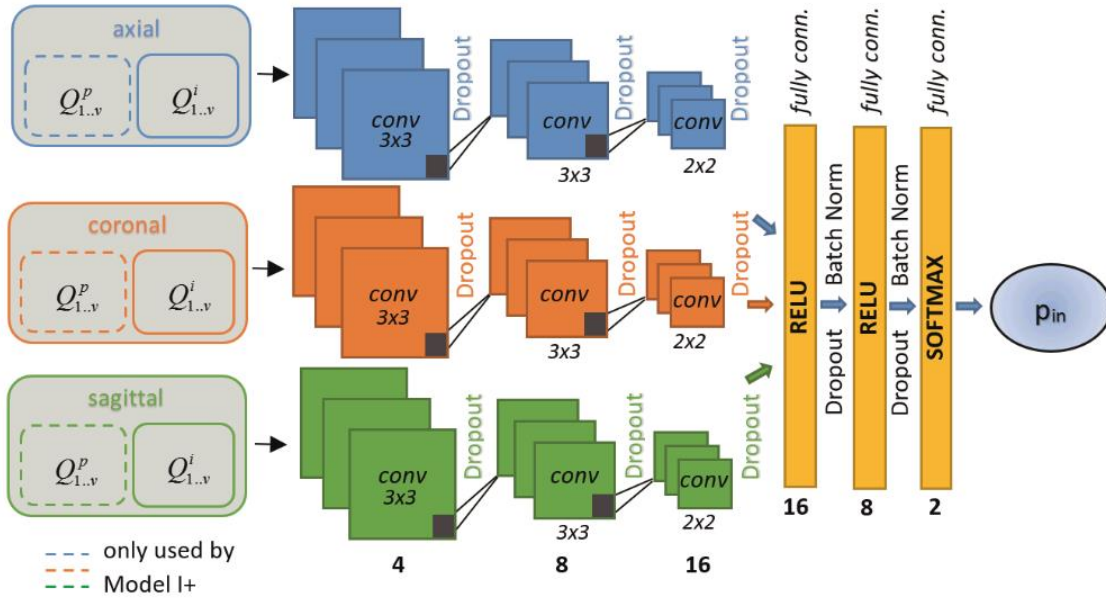
HaN-Seg veri seti ile birlikte başlatılan yarışma da (Podobnik ve diğ., 2023) araştırmacıların bu veri üzerinde yeni model ve yaklaşımlarını yarıştırmak için en iyi sonuçları elde etmelerini amaçlamıştır. Bu sayede, farklı ekipler çok modlu derin öğrenme mimarileri, veri artırma stratejileri, özel kayıp fonksiyonları gibi tekniklerle OAR segmentasyon performansını daha ileri taşımaya çalışmaktadır. Bu tezde yapılan çalışmada bu yarışmaya bir çözüm sunmayı amaçlamıştır.

**Atlas Tabanlı Yöntemler:** Atlas tabanlı segmentasyon yöntemleri, referans olarak kullanılan anatomik atlaslardan faydalanarak yeni görüntülerde yapıların belirlenmesini sağlar. Bu yöntem, özellikle baş ve boyun bölgesi gibi karmaşık anatomik yapıların segmentasyonunda yaygın olarak kullanılmıştır. Raudaschl ve arkadaşları (2017), 2015 yılında gerçekleştirilen Baş ve Boyun Otomatik Segmentasyon Yarışması'nda atlas tabanlı yöntemlerin performansını

değerlendirmiş ve bu yöntemin doğruluk açısından başarılı sonuçlar verdiğini belirtmiştir. Ancak, atlas tabanlı yöntemler genellikle manuel müdahale gerektirdiği ve büyük veri kümeleri ile sınırlı kaldığı için günümüzde yerini derin öğrenme ve makine öğrenimi gibi otomatik yöntemlere bırakmaya başlamıştır (Raudaschl ve diğ., 2017).

**Derin Öğrenme Tabanlı Yöntemlerin daha detaylı incelenmesi:** Son yıllarda, derin öğrenme yöntemleri tıbbi görüntü segmentasyonunda büyük bir ilerleme sağlamıştır. Derin öğrenme modelleri, büyük miktarda etiketlenmiş veriye ve güçlü hesaplama kaynaklarına ihtiyaç duyar. Bununla birlikte, transfer öğrenimi ve öz-denetimli öğrenme gibi yaklaşımlar, veri gereksinimlerini azaltarak model performansını iyileştirmiştir (Goodfellow ve diğ., 2016). Son yıllarda, transformer tabanlı modellerin tanıtılması, derin öğrenmenin özellikle doğal dil işleme alanındaki etkinliğini artırmıştır (Vaswani et al., 2017). Bu yenilikler, derin öğrenmenin birçok disiplin için kritik bir araç haline gelmesini sağlamıştır.

Özellikle konvolüsyonel sinir ağları (CNN) gibi derin öğrenme mimarileri, baş ve boyun bölgesi segmentasyonunda sıklıkla kullanılmaktadır. Fritscher ve arkadaşları (2016), baş ve boyun CT görüntülerinde 3D konvolüsyonel sinir ağı (3D U-Net) kullanarak organların hızlı ve tam otomatik segmentasyonunu gerçekleştirmiştir. Bu model, küçük anatomik yapıları doğru bir şekilde segmentleyebilmesiyle dikkat çekmiştir (Fritscher ve diğ., 2016). Ayrıca, Vrtovec ve arkadaşları (2020) atlas tabanlı yöntemlerden derin öğrenme tabanlı yöntemlere geçişin, doğruluk ve hız açısından önemli faydalar sağladığını göstermiştir (Vrtovec ve diğ., 2020).



**Şekil 2.3:** Fritscher ve arkadaşlarının önerdiği modelin diyagramı

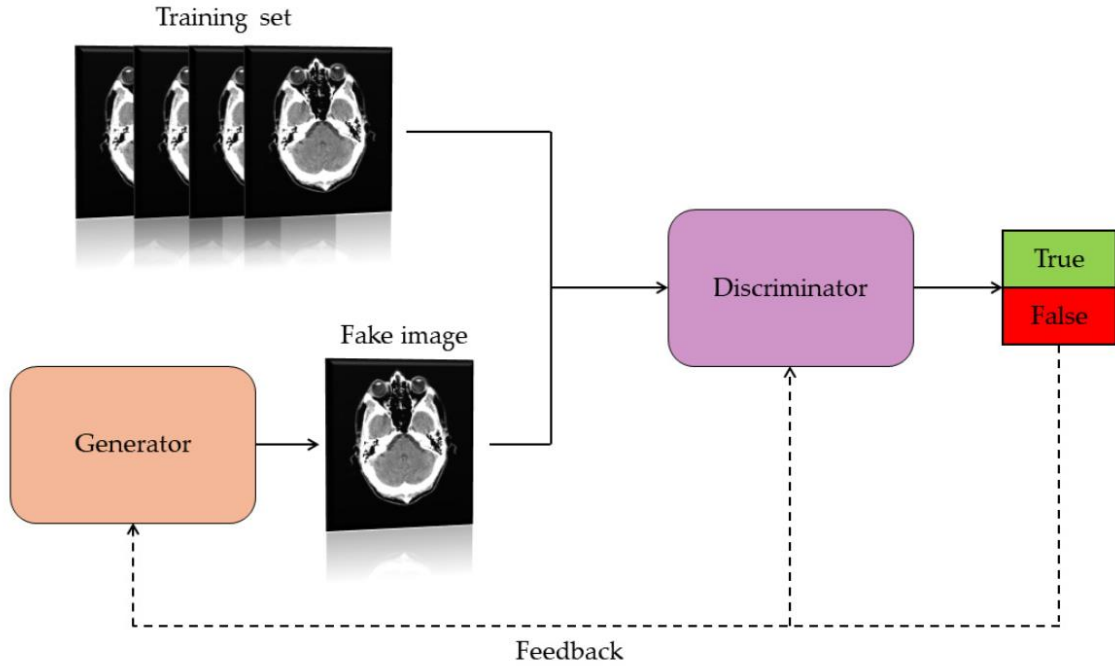
Derin öğrenme tabanlı yöntemler, özellikle baş ve boyun bölgesindeki karmaşık anatomik yapıların hassas segmentasyonunu mümkün kılmak için büyük veri setleri üzerinde eğitilmektedir. Literatürdeki mevcut çalışmaların büyük bir kısmı, beyin tümörlerinin segmentasyonu üzerine yoğunlaşmıştır (Litjens ve diğ., 2017). Bu da, baş ve boyun bölgesinin tamamını kapsayan ve farklı organları ayrıntılı olarak ele alan çalışmaların nispeten sınırlı olduğunu göstermektedir.

**Çok Modlu Öğrenme Yaklaşımları:** Çok modlu veriler, özellikle PET/CT gibi farklı görüntüleme tekniklerinin kombinasyonu, baş ve boyun tümörlerinin segmentasyonunda büyük ilgi görmektedir. Cho ve arkadaşları (2024), 18FDG PET-CT görüntülerinde baş ve boyun tümörlerinin segmentasyonu için dikkat mekanizması kullanan çok modlu bir öğrenme modeli sunmuştur. Bu yaklaşım, farklı görüntüleme modlarından elde edilen tamamlayıcı bilgileri kullanarak segmentasyon doğruluğunu artırmıştır (Cho ve diğ., 2024). Çok modlu öğrenme, özellikle farklı modaliteler arasındaki bilgileri entegre ederek, daha hassas ve güvenilir segmentasyon sonuçları elde edilmesini sağlar.

**Arter Bilinçli Segmentasyon:** Linlin Yao ve arkadaşları (2024), baş ve boyun CT anjiyografisi (CTA) görüntülerinde anevrizma segmentasyonu için arter bilinçli globalden-lokale bir çerçeve olan AASeg'i tanıtmıştır. Bu yöntem, vasküler bağlamı dikkate alan bir arter bilinçli özellik

çıkarma tekniği kullanarak, CTA görüntüleme sırasında anevrizmaların doğru bir şekilde segmentasyonunu sağlamıştır(Yao ve diğ., 2024). Bu arter bilinçli yaklaşım, vasküler yapıların net bir şekilde tanımlanmasını sağlarken, yanlış pozitifleri azaltarak segmentasyon kalitesini artırmaktadır.

**Kendiliğinden Öğrenme ve GAN Kullanımı:** Luca Michelutti ve arkadaşları (2024), baş ve boyun cerrahisi ve segmentasyonunda GAN'ların (Generative Adversarial Networks) kullanımına dair sistematik bir inceleme sunmuştur. GAN'lar, sınırlı veri setlerinde veri artırma için ve gerçekçi tıbbi görüntüler oluşturarak segmentasyon sonuçlarını iyileştirmek amacıyla kullanılmaktadır (Michelutti ve diğ., 2024). Bu inceleme, tıbbi görüntüleme verilerinin kıtlığı sorununa karşı GAN'ların giderek artan önemini vurgulamaktadır. GAN'lar, özellikle veri çeşitliliği eksikliği olan durumlarda, eğitim setlerini zenginleştirerek modellerin genelleme kapasitesini artırmaktadır.



**Şekil 2.4:** GAN modellerinin çalışma mantığı

Seenia Francis ve arkadaşları (2022), baş ve boyun CT görüntülerinde otomatize organ segmentasyonu için kendiliğinden öğrenme tabanlı SABOS-Net ağını önermiştir. Bu yöntem, genellikle kıt ve pahalı olan anotasyonlu veri setlerine bağımlılığı azaltmakta ve segmentasyon doğruluğunu artırmaktadır (Francis ve diğ., 2023). Kendiliğinden öğrenme, veri anotasyon maliyetlerini düşürerek daha geniş veri kümelerinin kullanılmasına olanak tanımaktadır.

**Risk Altındaki Organların Segmentasyonu:** J. Marsilla ve arkadaşları (2022), baş ve boyun kanser tedavisinde risk altındaki organların (OAR) otomatik segmentasyonunun klinik kabul edilebilirliğini araştırmıştır. Çeşitli açık kaynaklı 3D konvolüsyonel sinir ağlarının klinik uygulamadaki performansını değerlendirmişlerdir (Marsilla ve diğ., 2023a). Risk altındaki organların segmentasyonu, radyoterapi planlamasında büyük önem taşıdığından, bu modellerin klinik geçerliliği dikkatle değerlendirilmektedir.

**Anatomik Bölümlene Tabanlı Derin Öğrenme:** Song Li ve arkadaşları (2022), MRG taramalarında nazofarenks anatomisini tanımak için anatomik bölümlene tabanlı bir derin öğrenme modeli tanıtmıştır. Bu yaklaşım, sınır belirlemedeki hataları azaltarak segmentasyon spesifikliğini artırmaktadır (S. Li ve diğ., 2022). Anatomik bölümlene, organların ve dokuların detaylı olarak tanımlanmasına olanak tanıyarak, tedavi planlamasında daha hassas sonuçlar elde edilmesini sağlamaktadır.

Bu incelemede ele alınan çalışmalar, baş ve boyun bölgesi segmentasyonunda Transformer, GAN ve hibrit modellerin kullanımına yönelik artan bir trend göstermektedir. SwinCross gibi Transformer tabanlı modeller, çok modlu bilgileri bütünleştirmede başarılı olurken, kendiliğinden öğrenme yaklaşımları, büyük etiketlenmiş veri setlerine olan bağıllığı azaltmada yardımcı olmaktadır. Transformer modelleri, özellikle farklı görüntüleme modları arasındaki ilişkileri daha iyi modelleyerek segmentasyon doğruluğunu artırmaktadır.

GAN'lar ve arter bilinçli modeller, özellikle baş ve boyun gibi karmaşık anatomik bölgelerde veri yetersizliği sorununu gidermede büyük katkı sağlamaktadır. Ancak bu yöntemlerin klinik ortamlarda sağlıklı bir şekilde doğrulanması ve farklı veri setlerinde genel uygulanabilirliğini sağlamak için daha fazla çalışma gerekmektedir. Özellikle GAN tabanlı veri artırma yöntemlerinin, gerçek dünyadaki klinik senaryolarla uyumlu olup olmadığının değerlendirilmesi önem arz etmektedir.

**Hibrit Transformer Ağları:** Transformer tabanlı modeller, son yıllarda tıbbi görüntüleme de önemli bir yere sahiptir. Mingyuan Meng ve arkadaşları (2023), baş ve boyun kanseri hastalarında sağkalım tahmini için Transformer ve konvolüsyonel sinir ağlarını (CNN) birleştiren Hibrit Transformer Ağları sunmuştur. Bu hibrit model, segmentasyon kalitesini

arttırırken hasta sağkalımının doğru bir şekilde tahmin edilmesine olanak tanımıştır (Meng ve diğ., 2023). Transformer mimarisi, uzun mesafeli bağımlılıkları modelleyerek özellikle baş ve boyun bölgesinde karmaşık anatomik yapıları daha iyi tanımlamaktadır.

Gary Y. Li ve arkadaşları (2023) tarafından sunulan SwinCross, PET/CT görüntülerinde baş ve boyun tümörü segmentasyonu için çapraz-modal Swin Transformer mimarisidir. Swin Transformer, farklı görüntüleme modları arasındaki uzun mesafeli bağlantıları modellemede başarılı olup, segmentasyon doğruluğunda anlamlı iyileşmeler sağlamıştır (G. Y. Li ve diğ., 2024). Bu model, PET ve CT verilerinin farklı özelliklerini entegre ederek tümör sınırlarının daha net belirlenmesini sağlamaktadır.

### 2.7.2. Baş ve Boyun Görüntülemeye Hibrit ve Çok Aşamalı Segmentasyon Modelleri

Baş ve boyun bölgesinde hibrit segmentasyon modelleri, birden fazla yöntemi veya mimariyi bir araya getirerek her birinin güçlü yönlerinden yararlanmayı hedefler. Örneğin, Chen ve arkadaşları (2021) baş-boyun BT görüntülerinde organ segmentasyonu için 2B ve 3B evrişim katmanlarını birleştiren hibrit bir CNN modeli (OrganNet2.5D) geliştirmişlerdir (Chen ve diğ., 2021a). Bu sayede, farklı düzlemlerdeki çözünürlük farklarını gidermiş ve hem kenar ayrıntılarını hem de anlamsal özellikleri daha etkili yakalayabilmişlerdir. (Chen ve diğ., 2021a) Benzer şekilde, Liu ve ark. (2023) sınırlı etiketli veriyle çok modlu segmentasyon yapabilmek için evrişimsel sinir ağları ile Transformer katmanlarını entegre eden bir model önermiş; böylece modaliteler arası ortak ve özel özellikleri birlikte öğrenerek performansını arttırmayı amaçlamıştır. Yine Srivastava ve ark. (2022) ise çok ölçekli özellik füzyonu ile küresel ve lokal bağlamı bir arada kullanan bir ağ (OARFocalFuseNet) tasarlayarak baş-boyun organlarının daha isabetli segmentasyonunu sağlamışlardır (Srivastava ve diğ., 2022). Bu gibi hibrit yaklaşımlar, klasik tek tip modellere kıyasla daha yüksek Dice skorları ve daha düşük hata mesafeleriyle mevcut yöntemleri geride bırakabilmektedir (Srivastava ve diğ., 2022).

Hibrit modeller, farklı yaklaşımların avantajlarını bir araya getirerek segmentasyon doğruluğunu yükseltir. Örneğin, 2B+3B CNN birleşimi, hem ince dilim ayrıntılarını hem de hacimsel bağlamı yakalamaya izin vererek anisotropik çözünürlük sorununu çözer ve küçük organları daha iyi segmentler (Chen ve diğ., 2021b). Benzer biçimde, CNN ve Transformer entegrasyonu, yerel özellik çıkarımıyla küresel ilişki yakalama yeteneğini bir arada sunar ve az

etiketli veriyle dahi başarı elde etmeyi kolaylaştırır. Ayrıca, çok ölçekli füzyon stratejileri, farklı çözünürlüklerdeki özellikleri birleştirerek hem ince detayları korur hem de geniş bağlam bilgisini kullanır; bu sayede organ-at-riski segmentasyonunda son teknoloji performanslara ulaşılabilir (Srivastava ve diğ., 2022) Hibrit modeller sıklıkla tek bir yöntemle elde edilemeyecek derecede gelişmiş doğruluk ve genelleme kabiliyeti gösterir.

Bu modellerin en büyük handikapı, artırılmış mimari karmaşıklık ve dolayısıyla daha fazla hesaplama yüküdür. Örneğin, çok ölçekli 3B ağlar veya birden fazla alt model içeren yaklaşımlar, yüksek bellek ve hesaplama gücü gerektirebilir (Srivastava ve diğ., 2022). Eğitim süreci daha uzun olup, hiperparametre ayarı zorlaşabilir. Ayrıca, birden fazla yöntemin seri veya paralel kullanıldığı çok aşamalı modellerde, her bir aşamanın hatası bir sonrakine yansiyabilir, bu da dikkatli bir tasarım ve hata denetimi gerektirir. Hibrit yapılar genellikle daha fazla veri ihtiyacı duyar; her bileşenin başarılı olması için geniş ve çeşitli bir eğitim verisi seti olması önemlidir (Srivastava ve diğ., 2022).

Hibrit ve çok aşamalı modellerin entegrasyonunda genellikle derin öğrenmenin çeşitli yaklaşımları bir arada kullanılır. Örneğin, bazı çalışmalar iki kademeli bir mimari benimsemiştir: önce kaba bir segmentasyon veya tespit, ardından ince bir iyileştirme. Guo ve ark. (2020) bu stratejiyi kullanarak riskli organları boyutlarına göre kategorilere ayırmış ve küçük/detaylı organlar için önce bir algılama sonra segmentasyon adımı uygulamıştır; bu sayede çok küçük yapıları bile daha iyi yakalayarak genel performansı artırmıştır (Guo ve diğ., 2020). Yine benzer şekilde, CNN ile aktif kontur yönteminin birleşimi de hibrit bir yaklaşımdır: ilk adımda bir derin ağ tahmini yapar, ikinci adımda klasik bir aktif kontur algoritması sınırları hassaslaştırır. Bu tür çok aşamalı işlemler, özellikle manuel kontur çiziminin zor olduğu baş-boyun bölgesinde daha tutarlı sonuçlar verebilmektedir. Wang ve ark. (2021) tarafından önerilen, PET ve BT görüntülerini birleştirerek CNN çıktısını aktif kontur ile iyileştiren model buna iyi bir örnektir ve HECKTOR 2020 yarışmasında %75.2 Dice skoruyla ikinci olmuştur (Ma & Yang, 2020). Bu sonuçlar, hibrit modellerin uygun tasarlandığında hem hassas hem de genellenebilir segmentasyon çıktıları üretebildiğini göstermektedir.

### 2.7.3. Nesne Algılama (YOLO) ile Segmentasyonun Entegrasyonu

Nesne algılama algoritmalarının en bilinen örneklerinden olan YOLO, gerçek zamanlı nesne tespitindeki hızı ve başarısıyla tıbbi görüntü işlemede de dikkat çekmiştir. YOLO'yu segmentasyonla entegre etme fikri, önce ilgilenilen bölgenin hızlıca tespit edilip sonra detaylı olarak maskelenmesi yaklaşımına dayanır. Bu şekilde karmaşık görüntülerde aranan yapılar önce çerçeve içine alınır, ardından bu bölgelere odaklanmış bir segmentasyon uygulanır. Örneğin, Pandey ve ark. (2023) farklı modalitelerdeki (ultrason, BT, X-ray) lezyonları otomatik segmentlemek için YOLOv8 modelini kaba sınır tespiti amacıyla kullanmış, ardından Segment Anything Model (SAM) ile kesin maskeleri elde etmiştir (Pandey ve diğ., 2023). Bu çalışma, her modaliteden yalnızca 100 kadar görüntüyle eğitilen YOLOv8'in, ilgili organ veya tümörü çevreleyen yaklaşık bir sınırlayıcı kutu çizerek SAM'e başlangıç yapmasını sağlamıştır (Pandey ve diğ., 2023). Ardından SAM (ve yüksek kalitede segmentasyon için geliştirilmiş versiyonu HQ-SAM), bu kutu içindeki bölgeyi tam otomatik biçimde segmentlemiştir. Sonuçlar, kombine yaklaşımla elde edilen segmentasyonların oldukça yüksek doğrulukta olduğunu göstermiştir. İlginç biçimde, söz konusu çalışmada salt SAM modeli ile elde edilen sonuçlar, YOLO+SAM kombinasyonunu bile geride bırakmıştır; SAM tek başına daha yüksek Dice ve genel başarı sergilemiştir (Pandey ve diğ., 2023). Bununla birlikte, YOLOv8+SAM modelinin de klinik açıdan hız ve verimlilik kazandırma potansiyeli olduğu vurgulanmıştır (Pandey ve diğ., 2023). Bu durum, YOLO'nun özellikle sınırlı veriyle bile hızlı öneriler yapabilme kabiliyetinin, güçlü bir segmentasyon modeliyle birleştirildiğinde pratik fayda sağlayabileceğine işaret etmektedir.

YOLO ve segmentasyon tekniklerinin kombinasyonu literatürde farklı şekillerde karşımıza çıkmaktadır. Yaygın bir yöntem, iki aşamalı bir derin öğrenme boru hattı kurmaktır: İlk ağ bir nesne algılama ağı olup baş-boyun bölgesindeki hedef yapının (örneğin tümör veya organın) yaklaşık konumunu tespit eder; ikinci ağ ise bu bölgede yüksek çözünürlüklü bir segmentasyon gerçekleştirir. Tang ve ark. (2020) böyle bir iki-aşamalı yaklaşımda önce bir YOLO tabanlı tespit ağı, ardından U-Net tabanlı bir segmentasyon ağı kullanarak otomatik organ segmentasyonu yapmışlardır. Benzer şekilde, Li ve ark. (2021) da 48 hastadan oluşan bir baş ve boyun veri seti üzerinde önce OAR tespiti, sonra segmentasyonu yapan üç adımlı bir eğitim stratejisi önermişlerdir (Wu & Wang, 2020). Bu üç aşamalı yöntemde özellik çıkarma, organ bölgesi algılama ve son olarak segmentasyon adımları ardışık biçimde modellenmiştir (Wu &

Wang, 2020). Bu tip kombinasyonlar, özellikle baş ve boyunda çok sayıda organın bulunduğu durumlarda her bir yapıyı ayrı ayrı saptayıp işlemeyi kolaylaştırır. Örneğin, bir çalışmada derin öğrenmeye dayalı bir yöntemle önce kritik organların konumları belirlenmiş, ardından her organ için özelleşmiş segmentasyonlar yapılarak tüm organlar otomatik olarak segmente edilmiştir. Bu yöntem, PDDCA baş-boyun veri setinde test edilmiş ve geleneksel tek-aşamalı ağlara kıyasla daha başarılı bulunmuştur.

Baş ve boyun bölgesi özelinde, YOLO tabanlı segmentasyon çalışmaları daha çok tümör tespiti veya küçük organların yakalanması üzerine yoğunlaşır. Örneğin, Guo ve ark. (2020) küçük ve zor ayırt edilen bazı riskli organları segmentlemek için önce bu organların konumlarını tahmin eden bir dedektör kullanmış, sonrasında ikinci bir ağ ile kesin sınırları çizmiştir (Guo ve diğ., 2020). Bu dedektör-segmentör stratejisi, özellikle sınıf dengesizliğinin yüksek olduğu (küçük organların sayıca az fakat kritik olduğu) durumlarda performansı yükseltmektedir (Guo ve diğ., 2020). YOLO'nun sağladığı en büyük katkılardan biri, segmentasyon modelinin aramasını gereken alanı daraltması ve yanlış pozitifleri azaltmasıdır. Örneğin, küçük bir lenf nodu tüm vücut taramasında algılanması güç bir ayrıntı olabilir; YOLO önce bu nodu işaretlerse, segmentasyon ağı sadece o bölgede çalışarak hem zaman kazanır hem de diğer bölgelerdeki gürültüyü yok sayar. Ayrıca, YOLO tekniği gerçek zamanlı çalışabildiğinden, cerrahi navigasyon veya ışın tedavisi planlama gibi uygulamalarda anlık segmentasyon sistemlerine entegre edilebilir. Bununla birlikte, entegre yaklaşımların başarısı, ilk aşamadaki YOLO modelinin doğruluğuna bağlıdır: Eğer algılama aşaması hedef organı geçirirsa, segmentasyon aşaması da başarısız olacaktır. Dolayısıyla, bu tür sistemlerde her iki modülün dengeli bir şekilde eğitilmesi ve doğrulanması kritik önem taşır.

#### **2.7.4. Tıbbi Segmentasyonda Segment Anything Model (SAM) ve Temel Modeller**

Meta AI tarafından sunulan Segment Anything Model (SAM), 1 milyardan fazla nesne maskesiyle eğitilmiş, farklı görüntülerdeki herhangi bir nesneyi belirli bir istem ile segmentleyebilen genel amaçlı bir modeldir. SAM, doğal görüntüler üzerinde hiç öğrenme yapmadan dahi başarılı segmentasyonlar yapabildiği için dikkat çekmiştir. İlk çalışmalar, SAM'in medikal alandaki performansının veri setine ve nesne tipine göre çok değişken olduğunu ortaya koymuştur. Mazurowski ve ark. (2023) SAM'i 19 farklı tıbbi veri setinde denemiş ve tek bir tıklama veya kutu ile verilen istemler sonrasında elde edilen IoU (örtüşme)

skorlarının, omurga MR görüntülerinde 0.11 gibi çok düşük bir değerden kalça X-ray görüntülerinde 0.86'ya kadar geniş bir aralıkta değiştiğini raporlamışlardır (Mazurowski ve diğ., 2023). Bu dramatik fark, bazı medikal nesnelere SAM için oldukça zor (özellikle belirsiz sınırları olan tümörler gibi) olduğunu, bazılarının ise net yapıda olduğu için SAM tarafından kolay yakalandığını göstermektedir (Mazurowski ve diğ., 2023). Genel olarak, SAM iyi sınırlı, belirgin objelerde güçlü bir başlangıç segmentasyonu sunabilirken; beyin tümörü gibi karmaşık yapıya veya kontrastı düşük lezyonlarda beklenenin altında kalabilmektedir (Mazurowski ve diğ., 2023). Yine aynı çalışmada, kutu şeklindeki istemlerin nokta istemlerinden daha başarılı olduğu, birden fazla noktayla etkileşim imkânı verildiğinde ise alternatif yöntemlerin de SAM'i yakalayabildiği belirtilmiştir (Mazurowski ve diğ., 2023). SAM tıbbi görüntülerde bazı veri setlerinde etkileyici bir performans sergilese de, diğerlerinde orta veya zayıf kalmıştır; dolayısıyla medikal uygulamalarda dikkatli kullanılması ve gerekirse uyarlanması gerektiği vurgulanmıştır (Mazurowski ve diğ., 2023). Benzer şekilde, başka bir kapsamlı değerlendirmede de SAM'in sıfırdan segmentasyon başarısının, bazı durumlarda ilgili alandaki özel eğitilmiş derin ağlara yakın veya hatta daha iyi olabildiği, ancak tutarsızlıklar nedeniyle tamamen güvenilir bir "evrensel" çözüm olmadığı ifade edilmiştir (Mattjie ve diğ., 2023). SAM geleneksel yöntemlerle kıyaslandığında veri bağımsızlığı ve genelleme açısından çığır açıcı bir yaklaşım sunmakla birlikte, belirli bir tıbbi göreve özel eğitilmiş modeller kadar yüksek hassasiyet garantisi vermemektedir.

Bu durum, SAM gibi temel modellerin uyarlanması gerekliliğini gündeme getirmiştir. Nitekim, literatürde SAM'i medikal alana özel iyileştirme çabaları başlamıştır. Örneğin, Yan ve arkadaşları (2024) SAM'in 3B tıbbi görüntülerdeki sınırlı performansını artırmak için AFTer-SAM adını verdikleri bir uyarlama önermişlerdir (Yan ve diğ., 2024). AFTer-SAM modeli, SAM'e eklenen Aksiyel Füzyon Transformatörü sayesinde kesitler içi detayları ve kesitler arası bağlamsal ilişkileri birlikte öğrenebilmektedir (Yan ve diğ., 2024). Bilindiği gibi SAM orijinal olarak 2B görüntüler ve doğal fotoğraflar için tasarlanmıştır; dilimler arası ilişkiyi doğrudan modellemez. AFTer-SAM ise medikal hacimlerde dilimlerin birbiriyle olan bağımlılığını da dikkate alarak, orijinal SAM'e kıyasla hacimsel segmentasyonda belirgin iyileşmeler sağlamıştır (Yan ve diğ., 2024). Yazarlar, bu modeli dev bir temel modeli baştan eğitmek yerine verimli bir ince ayar yöntemiyle (LoRA, düşük sıralı adaptasyon) eğiterek hesaplama yükünü de makul düzeyde tutmuşlardır (Yan ve diğ., 2024). Sonuçlar, AFTer-SAM'in orijinal SAM'e göre 3

boyutlu medikal görüntülerde anlamlı ölçüde daha yüksek Dice skorları ürettiğini göstermekte ve büyük ön-eğitilmiş modellerin tıbbi alanda başarılı uygulanması için umut vadeden bir yön çizmektedir (Yan ve diğ., 2024). Benzer şekilde, diğer araştırmacılar da SAM'i farklı şekillerde adapte etmeye çalışmışlardır: "MedSAM" adıyla anılan bir versiyon, SAM'i büyük ölçekli medikal veri üzerinde yeniden eğiterek doğrudan tıbbi görüntülere uyarlamayı denerken; bazı çalışmalar SAM'in dikkat haritalarını veya gömülü uzayını sabit tutup yalnızca son katmanlarını değiştirme gibi yöntemlerle daha hafif ince ayar yaklaşımları denemiştir. Ren ve ark. (2024) yaptıkları çalışmada SAM'i baş ve boyun kanserlerinde Gross Tümör Hacmi segmentasyonu için değerlendirmiş; CT, PET ve MRI görüntülerini birlikte kullanarak hem sıfır-öğrenme hem de ince ayar senaryolarını karşılaştırmıştır. Bu çalışmada, kutu istemleriyle verilen zero-shot SAM segmentasyonlarının zaten makul bir doğruluk sağladığı, ancak modelin baş ve boyun verilerine yeniden eğitilmesi durumunda segmentasyon doğruluğunun belirgin biçimde arttığı raporlanmıştır (Ren ve diğ., 2024). Özellikle, büyük ölçekli medikal görüntüler üzerinde eğitilmiş MedSAM modelinin, kutu ile işaretlenen tümör bölgesini sıfırdan dahi oldukça iyi yakaladığı; üzerine yapılan ince ayarla hataların daha da azaldığı belirtilmiştir (Ren ve diğ., 2024). Bu bulgular, temel bir modelin tıbbi bir göreve uyarlanmasının, halihazırda genel modelin sağladığı avantajları koruyup görev-özel performansı yükselttiğini göstermektedir. SAM ve benzeri temel modeller sağlık alanında tamamen insan müdahalesini ortadan kaldırmasa da, yarı-otomatik segmentasyon süreçlerinde devrim yaratma potansiyeline sahiptir. Bir radyoloğun minimal bir etkileşimle (örneğin bir noktaya tıklayarak veya bir kutu çizerek) başarılı bir başlangıç segmentasyonu elde etmesi, kalan ince düzeltmeleri çok daha hızlı tamamlamasını sağlayabilir. Gelecekte, bu modellerin daha fazla medikal veriyle beslenip uzman bilgiyle harmanlanması halinde, geleneksel yöntemlerle yarışan – hatta bazı durumlarda onları geçen – performanslara ulaşması beklenmektedir.

### 3. MALZEME VE YÖNTEM

Bu bölümde, çalışmanın yürütülmesi sırasında kullanılan veri setleri, veri ön işleme yöntemleri, çok modlu görüntü füzyonu, önerilen model mimarisi, modelin eğitimi ve performans değerlendirme metrikleri ayrıntılı olarak açıklanacaktır. Çalışma, çok modlu (CT ve MR) görüntülerin birleşimi ile risk altındaki organların (OAR) hassas segmentasyonunu amaçlayan derin öğrenme tabanlı bir yaklaşımı içermektedir.

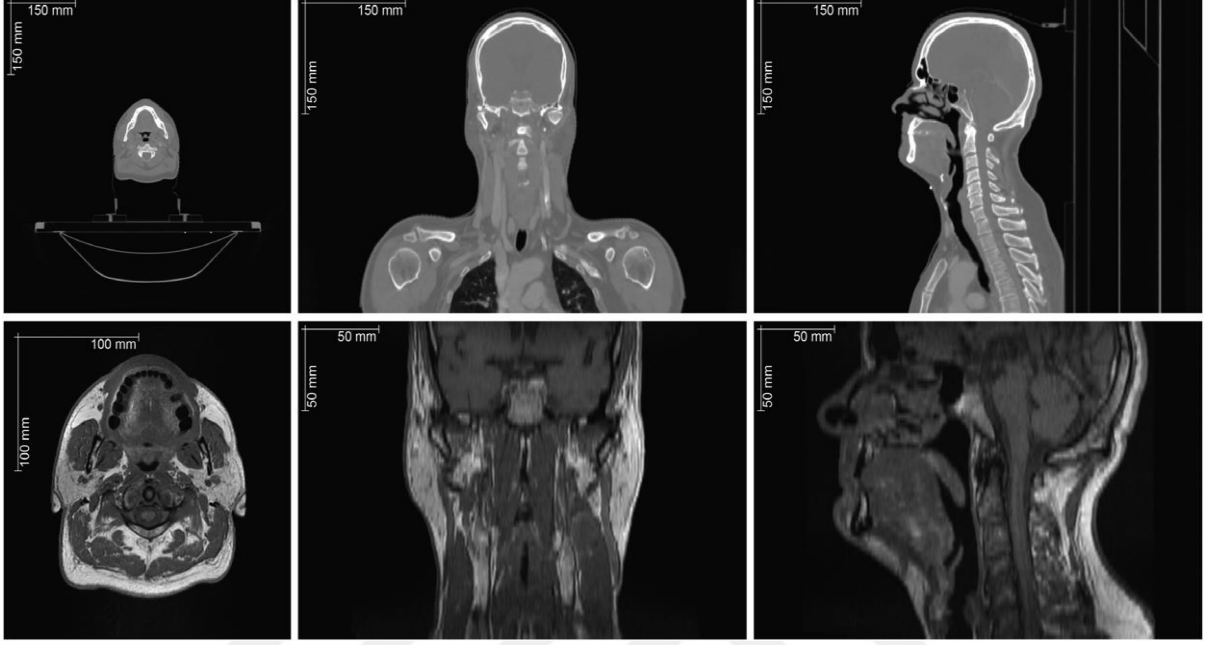
#### 3.1 VERİ SETİ

Bu çalışmada kullanılan veri seti, baş ve boyun bölgesinde radyoterapi için hedeflenen organların (OAR) segmentasyonuna yönelik olarak oluşturulmuş olan HaN-Seg veri setidir. HaN-Seg veri seti, baş ve boyun bölgesindeki kanserlerin tedavisinde radyoterapi planlamasında kullanılan bilgisayarlı tomografi (CT) ve T1-ağırlıklı manyetik rezonans (MR) taramalarını ve bu taramalarda uzmanlar tarafından manuel olarak belirlenmiş OAR segmentasyonlarını içermektedir (Podobnik ve diğ., 2023a). Veri setinde yer alan tüm görüntüler ve segmentasyonlar, radyoterapi için uygun tedavi planlaması yapılmasına olanak tanıyan detaylı anatomik bilgi sunmaktadır.

Veri seti, baş ve boyun kanserlerinden muzdarip toplam 42 hastaya ait CT ve MR görüntülerini içermektedir. Bu hastaların 42'sine ait veriler sağlanmış olup, bu verilerden 30'u eğitim amaçlı, 12'si ise test amaçlı kullanılmıştır. Her bir hasta için, 30 adet organın segmentasyonu uzman radyologlar ve radyoterapi teknologları tarafından manuel olarak belirlenmiş ve bu segmentasyonlar 3B binary maske formatında sunulmuştur. Segmentasyonlar, organların radyoterapi sırasında korunması için kritik olan organ-at-risk (OAR) yapılarını detaylı olarak göstermektedir. (Podobnik ve diğ., 2023a)

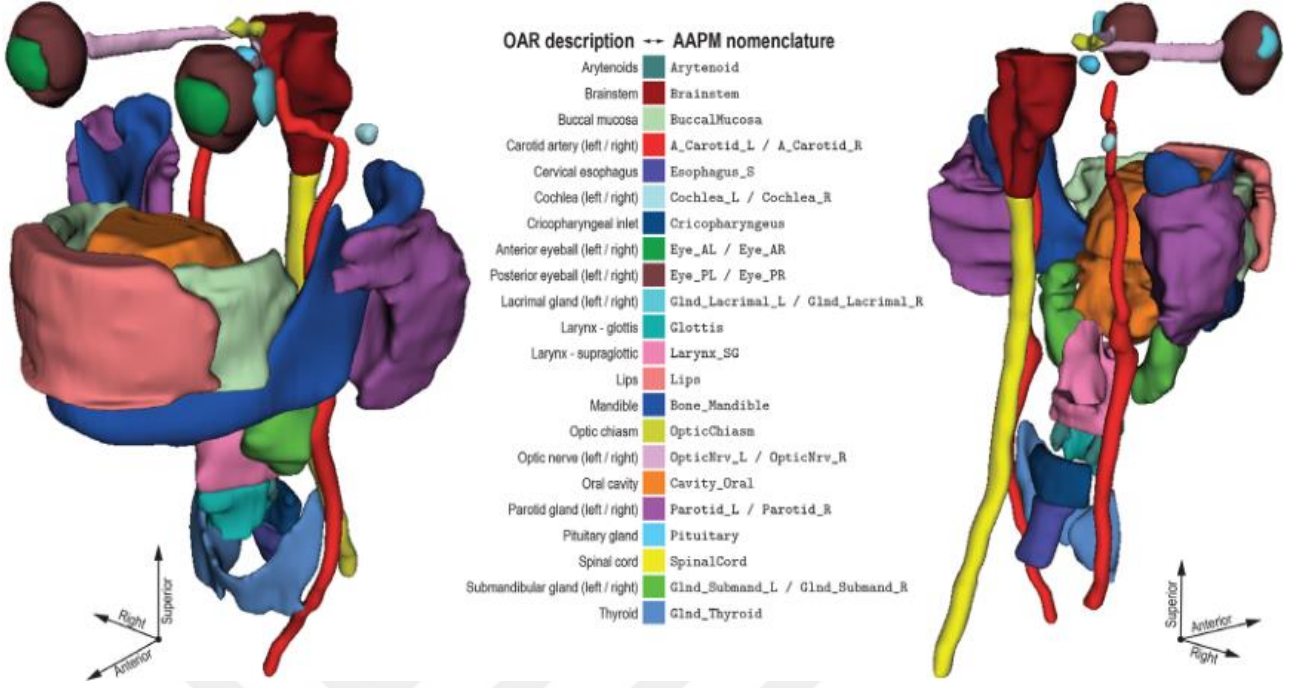
Veri seti, T1-ağırlıklı MR ve CT taramalarının birleştirilmesi ve bu iki modalitenin birlikte kullanılmasını hedefleyen çok modaliteli bir segmentasyon yaklaşımı sunmaktadır (Podobnik ve diğ., 2023b). Veri setindeki görüntüler ve referans segmentasyonlar, Amerikan Fizikçiler Derneği'nin (AAPM) Task Group 263 tarafından önerilen isimlendirme kurallarına uygun olarak adlandırılmış ve NRRD dosya formatında sunulmuştur. Bu veri seti, baş ve boyun bölgesinde radyoterapi planlamasında kullanılan otomatik segmentasyon yöntemlerinin

geliştirilmesi ve bu yöntemlerin performansının objektif olarak değerlendirilmesi amacıyla önemli bir kaynak sunmaktadır (Podobnik ve diğ., 2024).



**Şekil 3.1:** Veri setinden örnek görsel (Podobnik ve diğ., 2023a)

Veri ön işleme aşamasında, CT ve MR görüntüleri standartlaştırılmış ve bu görüntüler segmentasyon modeli için uygun hale getirilmiştir. Verilen veri setinde, CT ve MR görüntülerinin önemli bir kısmı, boşluk ya da hastanın uzandığı çekim masası gibi arka plan bölgelerinden oluşmaktadır. Bu arka plan bölgeleri, baş ve boyun bölgesindeki OAR'ların segmentasyonu açısından yararlı olmadığından, hacimlerin boyutunu azaltmak ve model eğitimi kolaylaştırmak amacıyla görüntü hacimleri şu şekilde kırpılmıştır: İlk olarak, vücut dışında kalan hava bölgeleri bir eşikleme yaklaşımı kullanılarak kırpılmıştır. Daha sonra, veri setindeki tüm hasta taramaları üzerinde her ekseninde herhangi bir OAR'a ait olan volume pixel'e olan minimum ve maksimum mesafeler hesaplanmıştır. Hacimler, bu limitlere her ekseninde 15 mm'lik bir boşluk eklenerek kırpılmıştır. Kırpma sınırı hastanın hacmini aştığında, yeni kırpma sınırı hastanın hacim sınırı olarak belirlenmiştir. MR hacimleri de aynı şekilde BT hacimleriyle eş zamanlı olarak kırpılmıştır. Bu işlem, her iki modalitenin de segmentasyon sürecine katkısını artırmak amacıyla yapılmıştır.

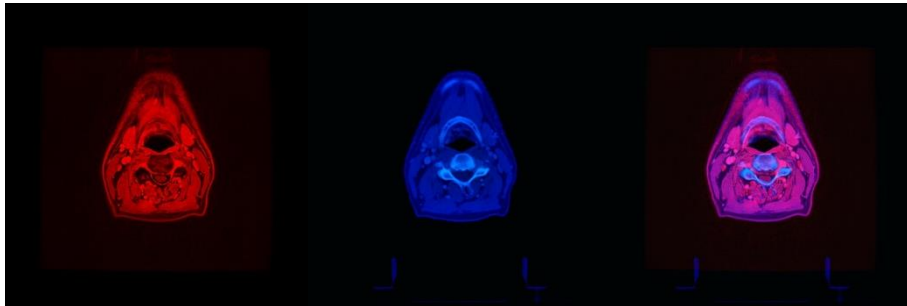


Şekil 3.2: Veri serinde sunulmuş segmentasyon maskelerinin örnek 3 boyutlu görüntüsü (Podobnik ve diğ., 2023a)

### 3.2 ÇOK MODLU GÖRÜNTÜ FÜZYONU

Bu çalışmada, baş ve boyun bölgesindeki organların segmentasyonu için CT ve MR görüntülerinin birleştirilmesi yoluyla çok modlu görüntü füzyonu gerçekleştirilmiştir.

Veri füzyonu sırasında, CT ve MR görüntüleri farklı modalitelerin avantajlarını bir araya getirmek amacıyla birleştirilmiştir. CT görüntüleri kemik yapılarını iyi bir şekilde temsil ederken, MR görüntüleri yumuşak doku detaylarını daha iyi göstermektedir. Bu nedenle, her iki modalitenin de sunduğu bilgileri kullanarak daha zengin ve kapsamlı bir veri dizisi oluşturulmuştur.



Şekil 3.3: MR, CT ve birleştirilmiş çok modlu görüntü

BT ve MR görüntülerinin birleştirilmesi için rijit ve non-rijit kayıt yöntemleri kullanılmıştır.

Rijit kayıt, görüntülerin temel hizalanmasını sağlarken, non-rigid kayıt (B-spline) yöntemleri ise organların ve yumuşak dokuların daha hassas bir şekilde hizalanmasını sağlamıştır.(Mattes ve diğ., 2001) Bu iki aşamalı kayıt süreci, modalitelerin en iyi şekilde birleştirilmesini amaçlamaktadır. Kayıt işlemleri sırasında SimpleITK (Lowekamp ve diğ., 2013) araç seti kullanılmış ve bu yaklaşım, modalitelerin optimal bir şekilde hizalanmasını sağlamak amacıyla uygulanmıştır.

### 3.3 ÖNERİLEN SEGMENTASYON MODELİ

OAR'ların segmentasyonu için YOLOv11 ve SAM2 modellerini birlikte kullanarak hibrit bir model sunmayı hedefliyoruz.

#### 3.3.1. Model Mimarisi (Model Architecture)

Önerilen segmentasyon modeli, YOLOv11 ve Segment Anything Model 2 (SAM2) entegrasyonunu kullanarak OAR segmentasyonunu gerçekleştirmektedir. YOLOv11, nesne tespiti için son derece hızlı ve doğru sonuçlar veren, çok katmanlı bir konvolüsyonel sinir ağı modelidir (Khanam ve Hussain, 2024). YOLOv11'in baş ve boyun bölgesindeki OAR'ları doğru bir şekilde lokalize etme kapasitesi, bu bölgedeki kritik organların segmentasyonu için önemli bir avantaj sunmaktadır. YOLO mimarisi, çeşitli anahtar iyileştirmeler ve derin katmanlı öğrenme teknikleri ile geliştirilmiş olup, baş ve boyun bölgesi gibi karmaşık anatomik yapıların doğru bir şekilde segmentlenmesini sağlamaktadır.

Segment Anything Model 2 (SAM2), Meta AI tarafından geliştirilen, dikkat mekanizmaları ile güçlendirilmiş bir segmentasyon modelidir (Meta AI, 2024). SAM2, özellikle çoklu modalite görüntülerde hassas segmentasyon yapabilmek için tasarlanmıştır. Bu model, dikkat mekanizmaları sayesinde, baş ve boyun bölgesindeki karmaşık yapıları daha doğru bir şekilde segmentlemekte ve CT ile MR görüntülerinin entegre kullanımını optimize etmektedir.

Ayrıca, modelin eğitimi için HaN-Seg veri seti kullanılmıştır (Podobnik ve diğ., 2024). Bu veri seti, baş ve boyun bölgesindeki organ-at-risk segmentasyonu için uygun, yüksek kaliteli CT ve MR görüntülerini içermektedir. Veri setindeki referans anotasyonları, Amerikan Fizikçiler Birliği (AAPM) Task Group 263 tarafından önerilen isimlendirme kurallarına uygun olarak hazırlanmıştır. Bu, modelin öğrenme sürecinde doğruluk ve güvenilirlik sağlamasına yardımcı olmuştur.

### *Teknik Entegrasyon ve Veri Akışı*

YOLOv11 ve SAM2 modelleri arasındaki veri akışı, YOLOv11'in çıktılarını SAM2'ye uygun bir şekilde besleyecek şekilde tasarlanmıştır. YOLOv11, öncelikle baş ve boyun bölgesindeki organ-at-risk'leri (OAR) tespit eder ve bu organların yerlerini belirler (bounding box koordinatları olarak). Bu koordinatlar kullanılarak ilgili bölgeler orijinal görüntüden kırılarak SAM2'ye beslenir.

#### *YOLOv11 Çıktısı ve ROI (Region of Interest) İşleme:*

1. YOLOv11 modeli, giriş CT veya MR görüntüsünden OAR'ların yerlerini belirlemek için kullanılır. Model, tespit edilen her organ için sınıf etiketi ve koordinatları içeren bir çıktı sağlar:

```

1 [
2   {"class": "larynx", "bbox": [x1, y1, x2, y2],
   "confidence": 0.95},
3   {"class": "spinal_cord", "bbox": [x3, y3, x4, y4],
   "confidence": 0.92}
4 ]

```

Şekil 3.4: Bounding box yapısı

2. Bu bounding box koordinatları kullanılarak orijinal görüntüden ilgili ROIs (Region of Interest) kırılır ve SAM2'ye beslenir:

```

1 import cv2
2
3 # Orijinal görüntüyü yükleyin
4 image = cv2.imread('original_image.png')
5
6 # YOLO çıktısına göre ROIs kırın
7 for detection in yolo_output:
8     x1, y1, x2, y2 = detection['bbox']
9     roi = image[y1:y2, x1:x2]
10    # ROI'yi kaydedin veya SAM2'ye besleyin
11    cv2.imwrite(f"roi_{detection['class']}.png", roi)

```

Şekil 3.5: Kırılan ilgi alanlarının segmentasyon için SAM2 modeline beslenmesi

### SAM2 ile Detaylı Segmentasyon:

```

1 # ROI'yi tensöre çevirin ve batch boyutu ekleyin
2 import torch
3
4 roi_tensor = torch.tensor(roi_normalized).unsqueeze(0)
5 # SAM2'ye besleyin
6 segmentation_mask = sam2_model(roi_tensor)

```

Şekil 3.6: Segmentasyon maskesi oluşturulması

SAM2 modeline beslenen her bir ROI, dikkat mekanizmaları sayesinde daha hassas bir şekilde segmentlenir. Bu segmentasyon, organın sınırlarını piksel seviyesinde belirleyerek detaylı bir maske oluşturur:

### Sonuçların Birleştirilmesi:

SAM2 tarafından sağlanan segmentasyon maskeleri, orijinal görüntüdeki konumlarına geri yerleştirilir. Bu işlem, kırpma işleminin tersine çevrilmesiyle ve maskelerin orijinal bounding box boyutlarına uygun hale getirilmesiyle gerçekleştirilir:

```

1 original_mask = np.zeros_like(image)
2
3 for detection, mask in zip(yolo_output, sam2_masks):
4     x1, y1, x2, y2 = detection['bbox']
5     mask_resized = cv2.resize(mask, (x2 - x1, y2 - y1))
6     original_mask[y1:y2, x1:x2] = mask_resized

```

Şekil 3.7: Çıktıların birleştirilmesi

### 3.3.2. Kayıp Fonksiyonu (Loss Function)

Modelin performansını optimize etmek ve OAR segmentasyonunda yüksek hassasiyet sağlamak amacıyla, Dice kayıp fonksiyonu ve Çapraz Entropi (Cross-Entropy) kayıp fonksiyonunun bir kombinasyonu kullanılmıştır. Terven ve diğ. (2023) tarafından yapılan bir çalışma, derin öğrenme için kullanılan kayıp fonksiyonlarının ve metriklerin kapsamlı bir analizini sunmaktadır. Dice kaybı, sınıf dengesizliğinin bulunduğu durumlarda segmentasyon doğruluğunu artırırken, Çapraz Entropi kaybı ise modelin her pikseli doğru bir şekilde sınıflandırmasını sağlamaktadır. Bu nedenle, bu iki kayıp fonksiyonunun kombinasyonu, modelin hem hassasiyetini hem de genel doğruluğunu artırmak için güçlü bir strateji sunmaktadır (Terven ve diğ., 2024).

**Dice Kayıp Fonksiyonu (Dice Loss):** Dice kaybı, özellikle sınıf dengesizliğinin bulunduğu segmentasyon problemlerinde oldukça etkili bir yöntemdir (Milletari ve diğ., 2016). Baş ve boyun bölgesindeki OAR'ların segmentasyonu, büyük ölçüde farklı boyutlarda ve yoğunluklarda organları içerdiğinden, sınıf dengesizliği yaygın bir sorundur. Dice kaybı, tahmin edilen ve gerçek maskeler arasındaki örtüşme oranını ölçerek bu sınıf dengesizliğinin üstesinden gelir. Dice kaybı, özellikle küçük yapılar üzerindeki performansı artırarak, modelin daha tutarlı ve doğru segmentasyon yapmasını sağlar. Bu, küçük ve karmaşık OAR'ların net bir şekilde segmentlenmesine olanak tanır ve klinik uygulamalarda kritik önem taşır (Men ve diğ., 2019).

**Çapraz Entropi Kayıp Fonksiyonu (Cross-Entropy Loss):** Çapraz Entropi kayıp fonksiyonu, modelin pikselleri doğru sınıfa atamasını teşvik etmek için kullanılır (Mao ve diğ., 2023). Bu fonksiyon, her bir pikselin ait olduğu sınıfa göre değerlendirilmesini sağlar ve bu sayede modelin genel doğruluğunu artırır. Özellikle baş ve boyun bölgesinde birçok farklı organ ve doku türü bulunması nedeniyle, piksellerin doğru bir şekilde sınıflandırılması, segmentasyon performansını doğrudan etkiler. Çapraz Entropi kaybı, modelin hem büyük hem de küçük organları doğru şekilde sınıflandırmasına yardımcı olur (Terven ve diğ., 2024).

**Kayıp Fonksiyonlarının Kombinasyonu:** Dice kaybı ve Çapraz Entropi kaybının birlikte kullanılması, modelin hem hassasiyetini hem de genel doğruluğunu artırmak için güçlü bir stratejidir (Sudre ve diğ., 2017). Dice kaybı, sınıf dengesizliğinin olduğu durumlarda modeli dengelerken, Çapraz Entropi kaybı ise her pikselin doğru sınıfa atanmasını teşvik eder. Bu kombinasyon, modelin küçük ve kritik organları daha iyi tanımlamasını sağlar ve segmentasyon performansını önemli ölçüde artırır. Bu stratejik yaklaşım, baş ve boyun bölgesindeki karmaşık yapıları segmentlerken hem doğruluk hem de hassasiyet sağlamak için idealdir.(Galdran ve diğ., 2022)

### 3.3.3. Optimizasyon Stratejisi (Optimization Strategy)

Başarılı bir tıbbi görüntü segmentasyonu modeli geliştirmek için kullanılan optimizasyon stratejileri, modelin doğruluk, genelleme yeteneği ve hesaplama verimliliği üzerindeki etkisi açısından kritik bir rol oynamaktadır. YOLOv11 ve SAM2 modelleri, yüksek boyutlu segmentasyon görevleri için geliştirilmiş olup, bu modellerin optimizasyonu için adaptif

öğrenme teknikleri ve döngüsel öğrenme oranları stratejileri benimsenmiştir. Aşağıda, bu stratejilerin detaylı bir incelemesi sunulmaktadır.

### ***Döngüsel Öğrenme Oranı ve Momentum Stratejisi***

Döngüsel öğrenme oranı (Cyclic Learning Rate - CLR), modelin eğitim sürecinde öğrenme oranının periyodik olarak değiştirilmesini sağlayarak yerel minimumlara takılma riskini azaltır ve genelleme performansını artırır (Mortazi ve diğ., 2023). CLR, üçgen bir dalga formu kullanılarak uygulanmış ve momentum oranıyla (MR) senkronize edilmiştir. Bu yöntem, optimize edilecek gradyan yönlerinin hızlandırılmasını ve stabilizasyonunu sağlar. Döngüsel CLR ve MR, aşağıdaki formüllerle modellenmiştir:

$$LR_{cycle} = Clr \times It \quad (3.1)$$

$$MR_{cycle} = Cmr \times It \quad (3.2)$$

burada Clr ve Cmr, döngü katsayılarını temsil ederken It, her epoch başına iterasyon sayısını ifade etmektedir. CLR'nin bu adaptif yaklaşımı, özellikle YOLO gibi derin öğrenme modellerinde segmentasyon doğruluğunu artırmıştır (Mortazi ve diğ., 2023).

### ***Hibrit Optimizasyon Algoritmaları***

Hibrit optimizasyon algoritmaları, modelin segmentasyon doğruluğunu artırmak ve hesaplama maliyetlerini azaltmak amacıyla benimsenmiştir. Bu çalışmada, Harris Hawks Optimization (HHO) ve Grey Wolf Optimization (GWO) algoritmalarının birleşimi olan G-HHO yöntemi tercih edilmiştir. G-HHO, keşif ve sömürü dengesini sağlama kapasitesini artırarak, global optimum çözümlere daha hızlı ulaşmayı sağlamıştır (Qader ve diğ., 2022). Özellikle, G-HHO şu avantajları sunar:

- Yerel minimumlara sıkışma riskini azaltan güçlü keşif mekanizmaları,
- Yüksek kaliteli çözümler için sömürü yeteneğini geliştirme,
- Model optimizasyon sürecinde hızlı yakınsama.

### *Sparsemax ile Parametre Normalizasyonu*

YOLOv11 ve SAM2 modellerinin optimize edilmesi sürecinde, ağırlıkların normalize edilmesi için sparsemax fonksiyonu kullanılmıştır. Sparsemax, geleneksel softmax fonksiyonunun aksine, olasılıkları daha seyrek bir basitlex üzerine projekte ederek daha keskin ve kararlı bir seçim sağlar (Xie ve diğ., 2024). Sparsemax aşağıdaki şekilde tanımlanır:

$$\text{sparsemax}(\alpha) := \arg \min_{p \in \Delta^K} \|p - \alpha\|^2 \quad (3.3)$$

burada  $\Delta_{K-1}$ , toplam olasılıkların bire eşit olduğu ve her bir olasılığın negatif olmadığı  $K-1$  boyutlu bir basitlex'tir. Softmax, sıcaklık parametresi ( $\tau$ ) ile modüle edilerek, daha yüksek bir sparse seviyesine ulaşmıştır. Sıcaklık modülasyonu aşağıdaki gibi modellenmiştir:

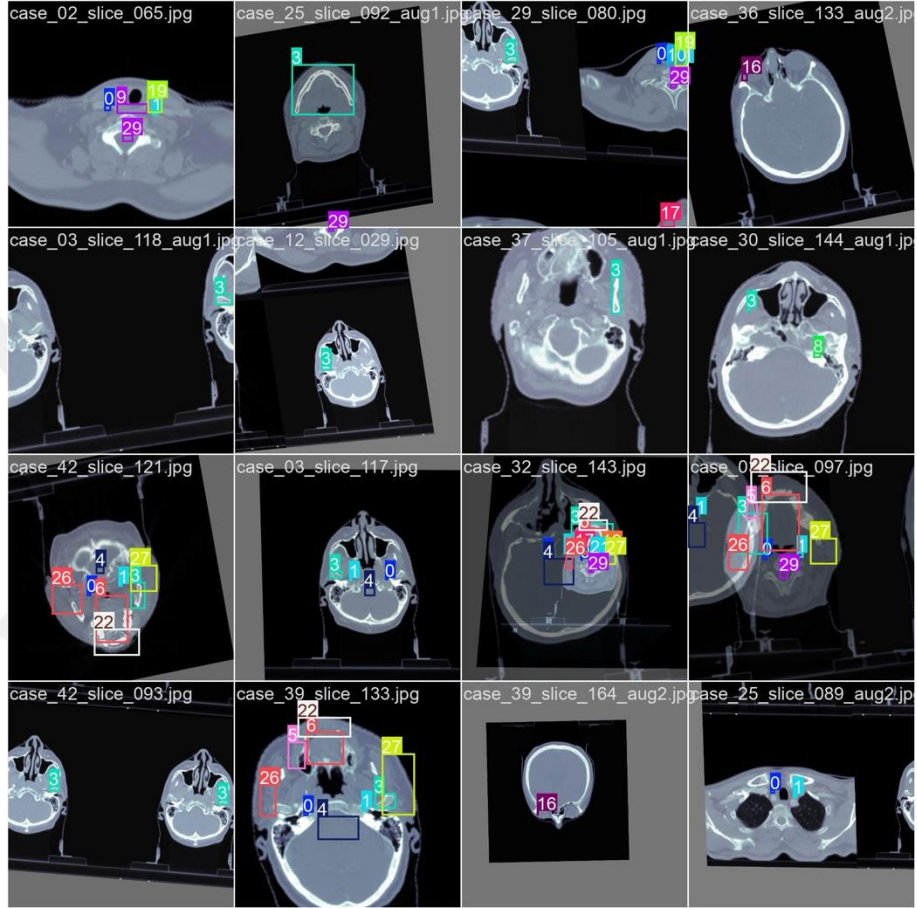
$$\text{softmax}(\alpha) = \frac{e^{\frac{\alpha_i}{\tau}}}{\sum_{j=1}^K e^{\frac{\alpha_j}{\tau}}} \quad (3.4)$$

burada  $\tau \rightarrow 0$  olduğunda, fonksiyon argmax fonksiyonuna yaklaşarak daha belirgin seçimlerin yapılmasını sağlar.

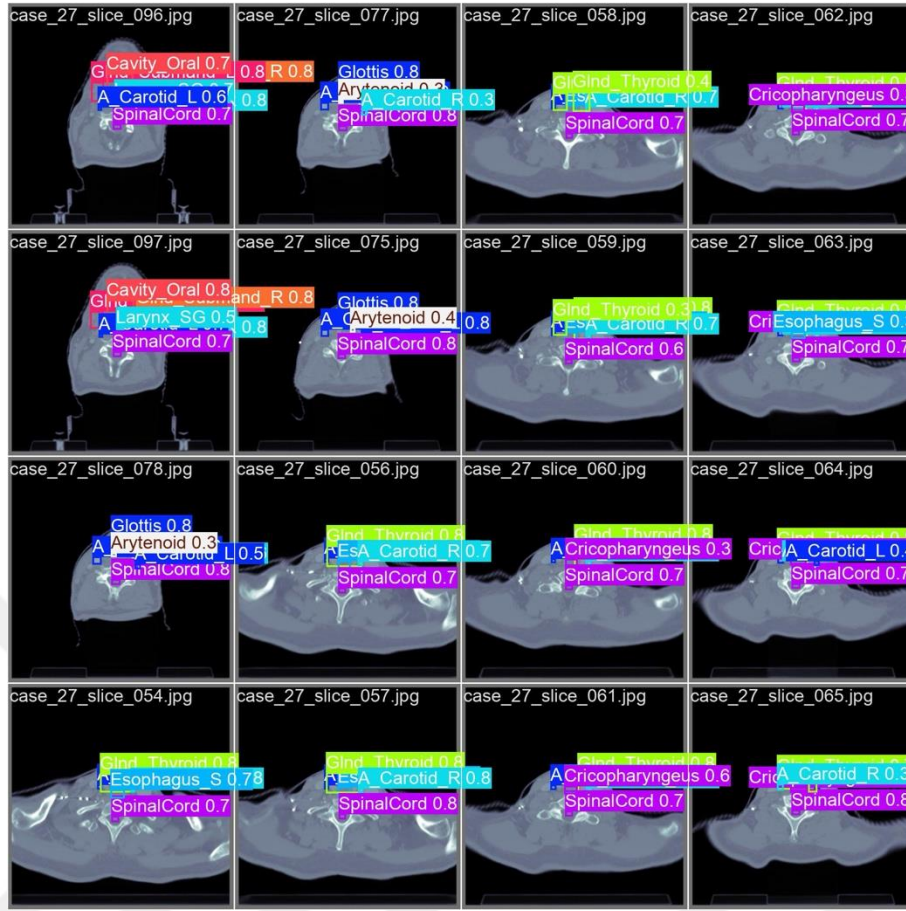
### *YOLOv11 ve SAM2 Uygulaması*

Uygulamamızda YOLOv11 modeli, baş-boyun bölgesine ait organların yerleştirilmesi amacıyla ince ayar (fine-tuning) yapılarak eğitildi. Eğitimde önceden eğitilmiş yolo11s.pt ağı kullanıldı ve modelin 200 epoch boyunca eğitilmesi planlandı; ancak erken durdurma mekanizması sayesinde eğitim 193. epoch'ta sonlandırıldı. En iyi sonuç ise 173. epoch'ta elde edildi. Öğrenme oranı olarak lr0=0.001 değeri kullanıldı ve son öğrenme oranı faktörü lrf=0.01 olarak belirlendi. Bu ayarlar, modelin istikrarlı bir şekilde öğrenmesini sağlamak için yapıldı. momentum=0.92 ve weight\_decay=0.0007 gibi hiperparametrelerle optimizasyon süreci daha kararlı hâle getirildi. Eğitim başlangıcında 10 epoch'luk bir ısınma süreci (warmup\_epochs=10.0) tanımlandı ve bu sayede öğrenme süreci yumuşak bir geçişle başlatıldı. Kayıp fonksiyonları arasında denge sağlamak amacıyla box=7.5, cls=0.75, dfl=1.5 ağırlıkları

kullanıldı. Modelin genelleme yeteneğini artırmak ve overfitting'i azaltmak amacıyla döndürme, ölçekleme, kaydırma, renk bozulması, mixup (mixup=0.3) ve copy-paste (copy\_paste=0.1) gibi güçlü veri artırma yöntemleri uygulandı. Tüm bu ayarlar, sınırlı tıbbi görüntü verisiyle daha doğru ve kararlı bir öğrenme süreci sağlamak için yapıldı.

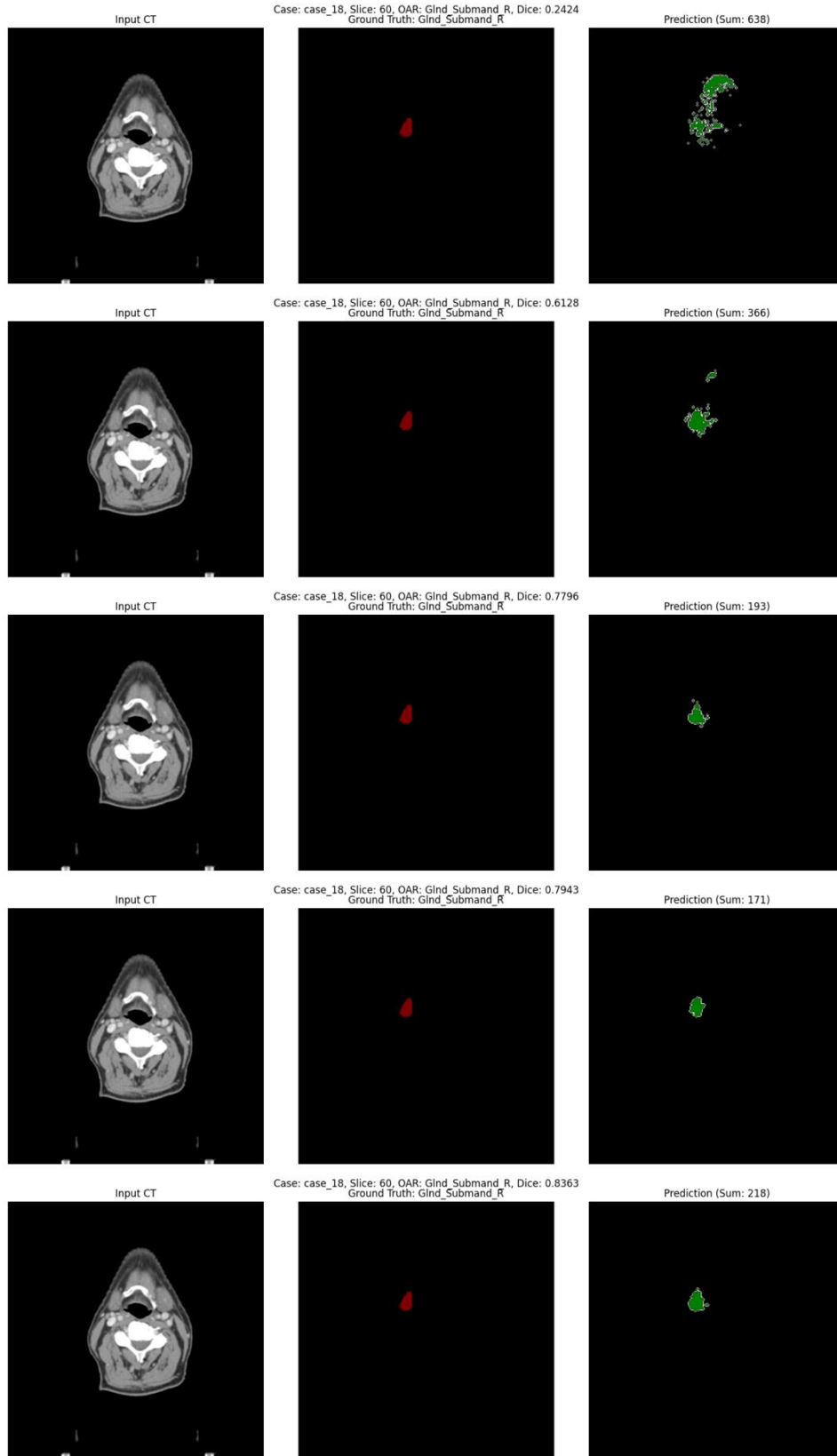


Şekil 3.8: Eğitim sürecinde veri artırma tekniği uygulanmış örnekler

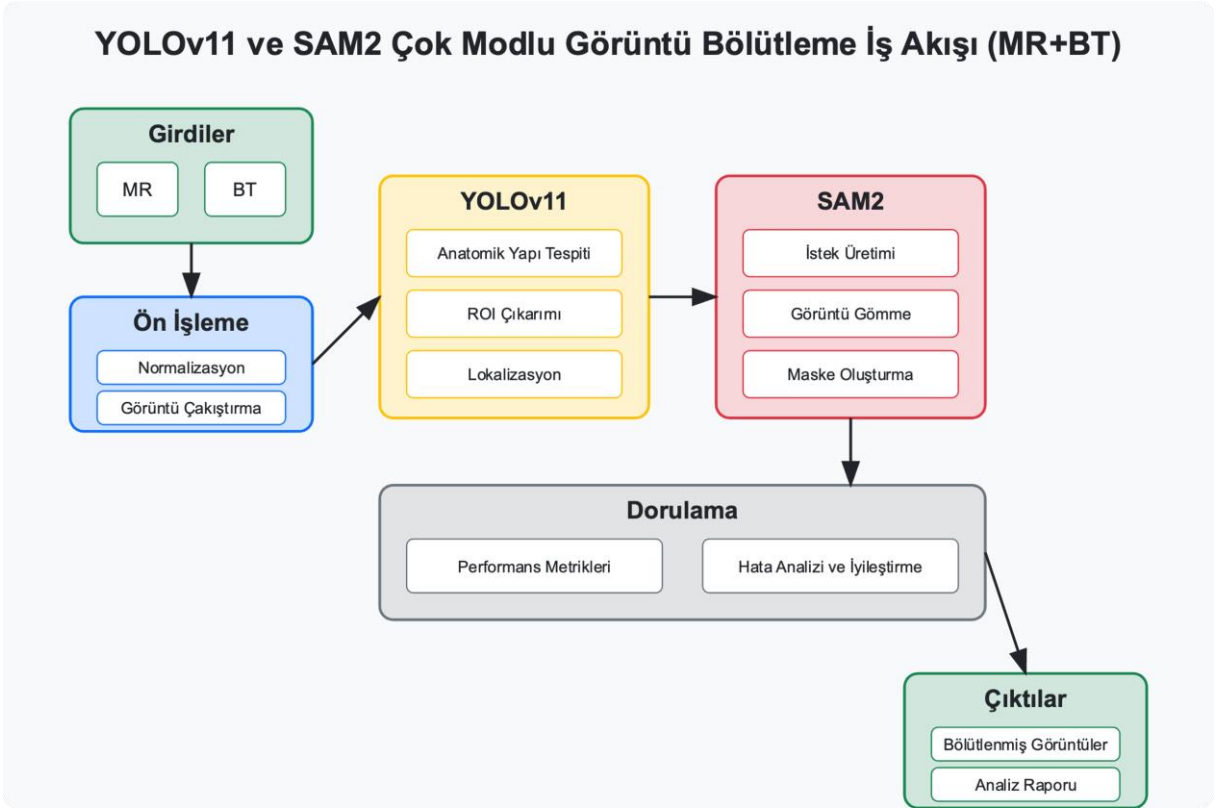


Şekil 3.9: Eğitim sonunda değerlendirme için oluşturulmuş çerçeveler ve etiketler

Modelin SAM 2 kısmında eğitim süreci boyunca yalnızca prompt encoder ve mask decoder katmanları eğitildi, image encoder ise dondurularak hem bellek kullanımı azaltıldı hem de eğitim süresi optimize edildi. Öğrenme oranları prompt encoder için  $1e-5$ , mask decoder için ise  $5e-5$  seçildi; bu sayede model önceden öğrenilmiş ağırlıkları koruyarak dengeli bir şekilde optimize edildi. Eğitim başlangıçta 50 epoch olarak planlandı ancak süreç iyileşme olmadığı için 43. epoch'ta sonlandırıldı ve en yüksek doğruluk 33. epoch'ta elde edildi. Kayıp fonksiyonu olarak Binary Cross Entropy ve Dice Loss birleşimi kullanıldı; bu sayede hem piksel düzeyindeki hatalar hem de segmentasyon bölgesinin örtüşme kalitesi dikkate alındı. Gradient accumulation yöntemi kullanılarak küçük batch boyutlarıyla çalışılırken büyük batch etkisi elde edildi. Segmentasyonu zor olan organlara öncelik verilerek bu organlara ait örnekler hem daha fazla hem de daha yüksek ağırlıklarla modele sunuldu. Eğitim süresince her epoch sonunda doğrulama setinde elde edilen Dice skorları takip edildi ve en iyi skoru veren model kayıt altına alındı.



**Şekil 3.10:** Eğitimin ilerlemesini gösteren farklı aşamalarda alınmış gerçek maskeyle üretilen maskenin karşılaştırıldığı görseller



Şekil 3.11: Model Mimarisi

## 4. BULGULAR

### 4.1 Model Performansı

#### Nicel Sonuçlar

Bu çalışmada, önerilen modelin performansı Dice Benzerlik Katsayısı (DSC) ve Hausdorff Mesafesi (HD95) metrikleri kullanılarak değerlendirilmiştir. Veri setini sağlayan Podobnik ve arkadaşları, HaN-SeG segmentasyon yarışması için bu metriklerin hesaplanmasında kullanılacak algoritmaları sunmuşlardır. Modelimiz, aynı yarışmanın sonuçlarıyla ilişkili olarak literatürde yayımlanan eli1, cwlg102 ve CHB-QuantIF modelleri ile karşılaştırılmıştır.

eli1 (E.T.) modeli yüksek DSC ve düşük HD95 değerleri ile güçlü bir performans sergilemiştir. Bu model, rigid MR-to-CT kaydı kullanarak multimodal veri işleme yeteneklerini göstermiştir. (Podobnik ve diğ., 2024)

cwlg102 (C.L. ve J.S.K.) modeli YOLOv7 ile ön lokalizasyon yaparak CT modalitesine odaklanmış ve DynUNet segmentasyon algoritmasıyla iyi bir DSC değeri elde etmiştir. (Podobnik ve diğ., 2024)

CHB-QuantIF (Z.M. ve R.M.) modeli transfer öğrenme teknikleri kullanılarak geliştirilmiş ve dengeli bir performans sağlamıştır. (Podobnik ve diğ., 2024)

Her ne kadar test için aynı algoritmalar kullanılmış ve test ortamları docker sistemi ile konteyner olarak performansları eşleştirilmiş şekilde sunulmuş olsa da, farklı bilgisayar sistemlerinin kullanılması, işlemci gücü, bellek yönetimi veya yazılım optimizasyonlarından kaynaklanan performans farklılıklarına yol açabilir. Bu nedenle, sonuçlar doğrudan karşılaştırılabilir olarak yorumlanırken dikkatli olunmalıdır.

**Dice Benzerlik Katsayısı (DSC):** DSC, modelin segmentasyon başarısını ölçmek için kullanılan bir metriktir. Bu metrik, referans segmentasyon maskesi (RR) ile model çıktısı (PP) arasındaki örtüşme miktarını değerlendirir. DSC formülü şu şekilde ifade edilir:

$$DSC = \frac{2 \cdot |R \cap P|}{|R| + |P|} \quad (4.1)$$

Burada:

$R \cap P$ : Referans ve tahmin edilen segmentasyonların kesişim alanıdır.

$|R|$ : Referans segmentasyonun toplam alanıdır.

$|P|$ : Tahmin edilen segmentasyonun toplam alanıdır.

DSC değeri 0 ile 1 arasında değişir; 1 tam uyumu ifade eder. Model, her bir organ için ayrı ayrı DSC değerlerini hesaplamış ve ardından tüm organlar için ortalama DSC değerini almıştır. DSC, modelin segmentasyon başarısını ölçmek için kullanılan bir metriktir. Her bir organ için ayrı ayrı hesaplanır ve ardından tüm organların ortalaması alınır. DSC, referans segmentasyon maskesi ile model çıktısının örtüşme miktarını ölçer. Değer, 0 ile 1 arasında değişir; 1 tam uyumu ifade eder. Önerilen test betiği, DSC'yi tüm 30 organ üzerinde hesaplayarak sonuçları karşılaştırmıştır.

**Hausdorff Mesafesi (HD95)**: HD95, segmentasyon doğruluğunu ölçmek için kullanılan bir metriktir. Bu metrik, segmentasyon sınırlarının maksimum mesafesini yüzde 95 seviyesinde dikkate alır ve aşırı uç değerlerin etkisini azaltır. Hausdorff Mesafesi şu şekilde tanımlanır:

$$HD(A, B) = \max \left\{ \sup_{a \in A} \inf_{b \in B} d(a, b), \sup_{b \in B} \inf_{a \in A} d(a, b) \right\} \quad (4.2)$$

Ancak HD95, yalnızca en uçtaki 95. yüzdelik dilimdeki mesafeyi dikkate alarak şu şekilde ifade edilir:

$$HD95(A, B) = \text{percentile}_{95}(\{d(a, b) \mid a \in A, b \in B\}) \quad (4.3)$$

Burada:

$d(a, b)$ : aa ve bb noktaları arasındaki Öklid mesafesidir.

AA: Referans segmentasyonun sınır noktalarıdır.

BB: Tahmin edilen segmentasyonun sınır noktalarıdır.

Her bir organ için HD95 hesaplanmış ve ardından tüm organlar için ortalama HD95 değeri alınmıştır.

**Tablo 4.1:** Organ bazında model performans değerleri

<b>Organ</b>	<b>Modelimiz (DSC)</b>	<b>Modelimiz (HD95)</b>
Parotid gland (L)	91.9	2.2
Parotid gland (R)	91.4	2.3
Lips	90.9	2.3
Thyroid gland	84.8	3.0
Mandible	83.7	3.1
Carotid artery (L)	82.8	3.2
Submandibular gland (L)	82.4	3.2
Submandibular gland (R)	81.3	3.3
Carotid artery (R)	80.8	3.4
Brainstem	80.1	3.5
Oral cavity	79.0	3.6
Buccal mucosa	78.6	3.6
Supraglottic larynx	76.8	3.8
Optic nerve (R)	76.0	3.9
Spinal cord	75.9	3.9
Glottic larynx	75.9	3.9
Eyeball (A, R)	75.4	4.0
Optic nerve (L)	73.9	4.1
Arytenoids	73.9	4.1
Cervical esophagus	73.9	4.1
Cricopharyngeal inlet	73.2	4.2
Eyeball (P, R)	73.1	4.2
Eyeball (P, L)	72.7	4.3
Lacrimal gland (L)	70.0	4.6
Optic chiasm	69.1	4.6
Eyeball (A, L)	67.9	4.8
Lacrimal gland (R)	65.6	5.0
Pituitary gland	65.5	5.0
Cochlea (L)	62.8	5.3
Cochlea (R)	61.2	5.5

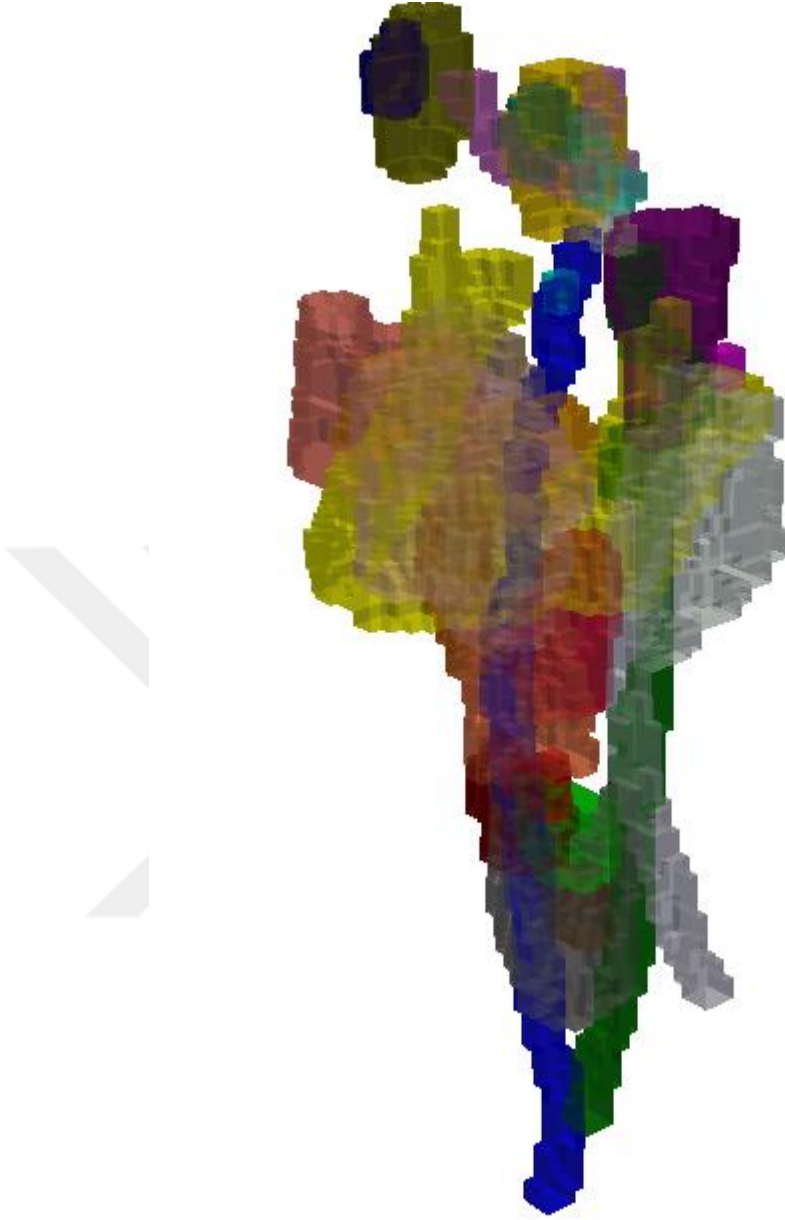
Geliştirilen model, test verisindeki organ-at-risk (OAR) segmentasyon görevlerinde ortalama Dice Benzerlik Katsayısı  $78.9 \pm 2.8$  ve HD95 metriği  $3.445 \pm 2.4$  mm olarak hesaplanmıştır. Bu değerler, mevcut literatürdeki modellerle kıyaslandığında, daha yüksek bir performans göstermektedir. YOLOv11 ve SAM2'nin entegrasyonu, özellikle karmaşık yapıların segmentasyonunda belirgin bir iyileşme sağlamıştır.

**Tablo 4.2:** Modellerin karşılaştırmalı performans analizi

Model	DSC	HD95
eli1	$0.769 \pm 0.084$	$3.5 \pm 2.4$
cwlg102	$0.768 \pm 0.093$	$3.8 \pm 3.8$
CHB-QuantIF	$0.751 \pm 0.086$	$3.7 \pm 2.3$
<b>Modelimiz</b>	<b><math>0.789 \pm 0.088</math></b>	<b><math>3.445 \pm 2.4</math></b>

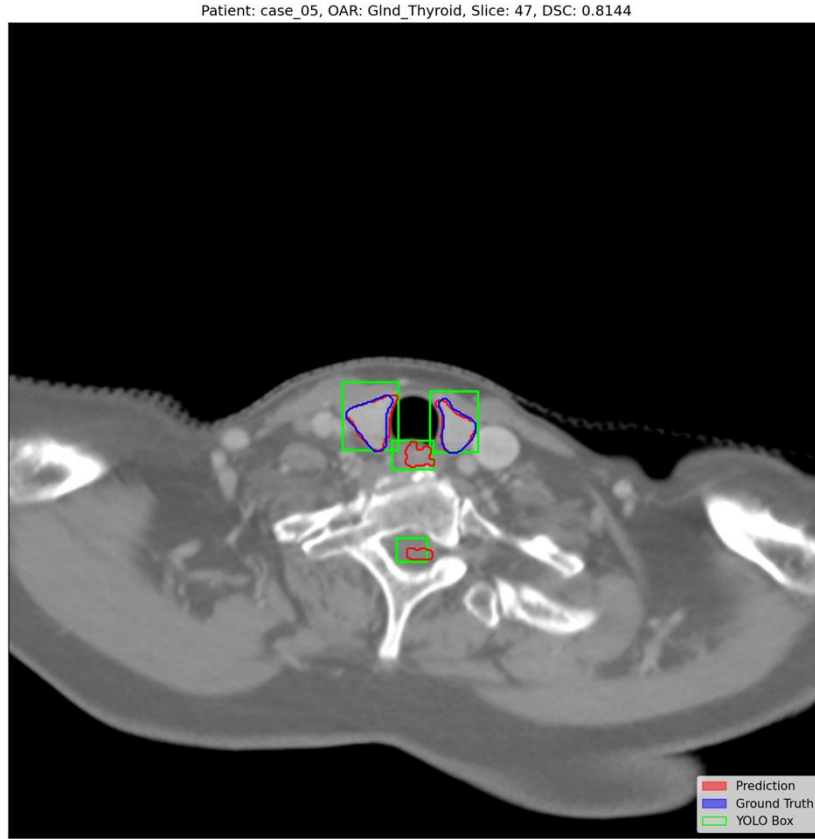
### Nitel Sonuçlar

Modelin çıktıları, referans segmentasyon maskeleri ile görsel olarak karşılaştırılmış ve klinik uygulama açısından önemli olan organların (optik sinirler, parotis bezi, beyin sapı gibi) net bir şekilde ayrımlanabildiği görülmüştür. Aşağıdaki şekil, segmentasyon başarısının bir örneğini sunmaktadır.



**Şekil 4.1:** Oluşturulan segmentasyon maskelerinin 3 Boyutlu görüntüsü

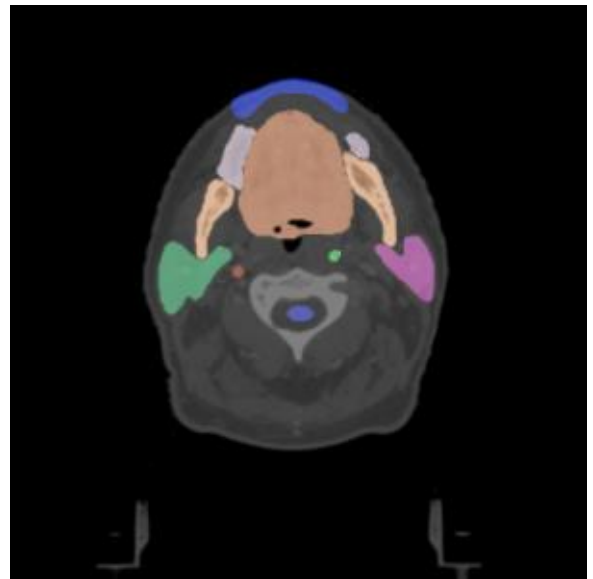
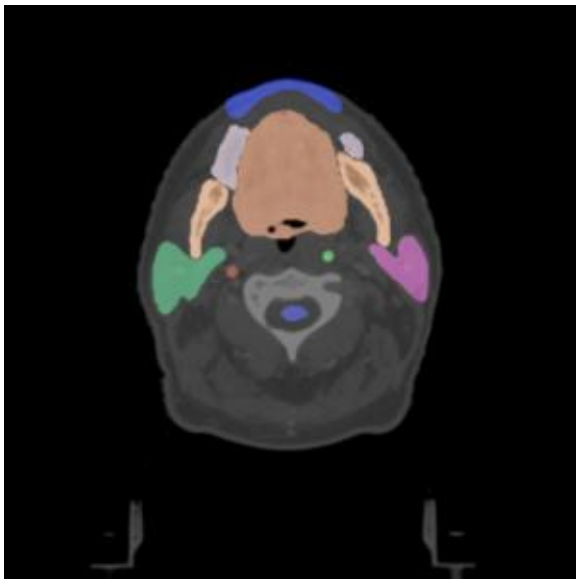
MR ve CT'nin birlikte kullanılması, yumuşak dokuların daha iyi tanımlanmasına katkı sağlamıştır (Podobnik ve diğ., 2024). Yüksek kontrastlı yapıların belirginleştirilmesi sayesinde, klinik karar verme süreçlerinde önemli bir iyileşme beklenmektedir.



**Şekil 4.2:** Kesit üzerinde tespit edilen OAR'ların sınırlayıcı kutu ile belirlenmiş ve SAM 2 ile segmente edilmiş hali. 5. hastanın 47. kesiti üzerinde organlar tespit edilmiş ve daha sonra tiroit bezi için verilen uzman segmentasyonu ile karşılaştırılmıştır.

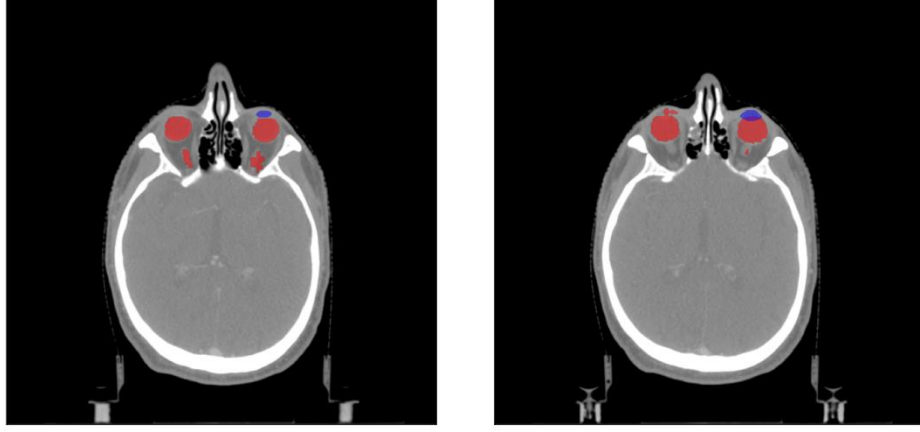
*Test veri setinde uzman tarafından segmente edilmiş kesit*

*Model tarafından segmente edilmiş kesit (Görselleştirme estetiği için maske üzerinde düzeltme ve yumuşatma işlemleri uygulanmıştır; analizde orijinal maske kullanılmıştır.)*



**Şekil 4.3:** Ağız boşluğu ve çevresindeki organların segmentasyonun karşılaştırması

*Test veri setinde uzman tarafından segmente Model tarafından segmente edilmiş kesit edilmiş kesit*



**Şekil 4.4:** Gözler ve göz sinirlerinin segmentasyonun karşılaştırması

## 5. TARTIŞMA

### 5.1 SONUÇLARIN DEĞERLENDİRİLMESİ

Bu tez çalışmasında, baş ve boyun bölgesindeki risk altındaki organların (OAR'ların) otomatik segmentasyonu için önerilen YOLOv11 ve SAM2 tabanlı modelin, mevcut yöntemlere kıyasla hem doğruluk hem de hız açısından önemli katkılar sağladığı gösterilmiştir. Çok modlu görüntüleme (BT ve MR) kombinasyonu, yumuşak dokuların ve karmaşık anatomik yapıların segmentasyonunda önemli iyileştirmelere yol açmıştır.

Dice Benzerlik Katsayısı ve HD95 sonuçları, yalnızca geometrik doğruluk açısından değil, aynı zamanda modelin potansiyel klinik uygulanabilirlik açısından da etkili olduğunu ortaya koymuştur. Ancak, bu tez kapsamında modelin klinik değerlendirmesi yapılmamış ve farklı veri setleri üzerinde genel performans analizi gerçekleştirilmemiştir. Dolayısıyla, bulgular, daha kapsamlı deneylerle desteklenmesi gereken ön değerlendirmeler olarak görülmelidir. Hibrit modelin kaynak yoğunluğu ve işlem maliyeti gibi faktörler, düşük donanım kaynaklarına sahip merkezlerde sınırlamalar yaratabilir.

### 5.2 KLİNİK UYGULANABİLİRLİK

Model, SCARF (auto-Segmentation Clinical Acceptability and Reproducibility Framework) (Marsilla ve diğ., 2023b) çerçevesinin ilk adımları kapsamında değerlendirilmiştir. SCARF'in sunduğu metodolojik yapı, modelin klinik kabul edilebilirliğinin ve genellenebilirliğinin daha geniş ölçekte incelenmesine olanak tanıyacaktır. Bu tez çalışmasında, SCARF'in klinik değerlendirme ve genellenebilirlik aşamaları uygulanmamış olsa da, önerilen modelin performansı, potansiyel klinik uygulamalara uyarlanabilirliği açısından umut vaat etmektedir.

Modelin hız avantajı, klinisyenlerin iş yükünü azaltırken tedavi planlama süreçlerini hızlandırma potansiyeline sahiptir. Bununla birlikte, yüksek donanım gereksinimleri, modelin yaygınlaştırılabilirliği üzerinde sınırlamalar yaratabilir. Klinik kullanıma geçiş öncesinde modelin kapsamlı değerlendirmeleri yapılmalı ve SCARF çerçevesinin diğer adımları tamamlanmalıdır.

### 5.3 SINIRLAMALAR VE GELECEK ÇALIŞMALAR

Bu çalışmada karşılaşılan en önemli sınırlamalardan biri, modelin yalnızca sınırlı bir veri seti üzerinde değerlendirilmiş olmasıdır. Daha geniş ve çeşitli veri setleri üzerinde test edilmesi, modelin genellenebilirliğini ve performansını anlamak açısından kritik öneme sahiptir. Ayrıca, modelin klinik senaryolarda test edilmemiş olması, bulguların doğrudan klinik uygulamalara uyarlanmasını sınırlandırmaktadır.

Hibrit modellerin işlem yoğunluğu, düşük donanım kaynaklarına sahip merkezlerde uygulanmasını zorlaştırabilir. Gelecekte, bu tür modellerin işlem yükünü azaltan optimizasyon yöntemlerine odaklanılmalıdır. SCARF çerçevesinin tüm adımları tamamlanarak modelin klinik kabul edilebilirliği ve genellenebilirliği daha geniş bir perspektifte değerlendirilmelidir. Ayrıca, farklı anatomik bölgelerdeki organların segmentasyonuna yönelik uyarlamalar yapılabilir.

## 6. SONUÇ

### 6.1 TEMEL BULGULAR VE KATKILAR

YOLOv11 ve SAM2 modellerinin başarılı bir şekilde entegre edilmesi, segmentasyon doğruluğunu önemli ölçüde artırmıştır. Bu entegrasyon, özellikle yumuşak dokuların daha kesin bir şekilde ayrıştırılmasına olanak tanımış ve BT ile MR modalitelerinin bir arada kullanımı sayesinde daha kapsamlı bir segmentasyon sağlanmıştır. Bu model, SCARF çerçevesi kapsamında değerlendirilmiş ve klinik uygulanabilirlik açısından umut vaat eden bir potansiyele sahip olduğu gözlemlenmiştir. Ayrıca, küçük ve karmaşık organların segmentasyonunda gösterdiği yüksek doğruluk, sınıf dengesizliği probleminin etkili bir şekilde ele alınabileceğini ortaya koymuştur. Bununla birlikte, hibrit modellerin getirdiği doğruluk avantajlarına rağmen, kaynak yoğunluğu ve işlem gücü gereksinimlerinin dikkatle değerlendirilmesi gerektiği de belirtilmiştir.

Bu çalışmanın bir diğer önemli katkısı, segmentasyon doğruluğunun artırılmasına yönelik çok modlu görüntüleme tekniklerinin etkinliğini göstermesidir. BT ve MR görüntülerinin birlikte kullanımı, özellikle anatomik olarak karmaşık ve küçük ölçekli organların segmentasyonunda önemli bir avantaj sağlamıştır. Modelin, farklı hasta gruplarında ve çeşitli görüntüleme protokollerinde nasıl performans gösterdiğini anlamak, ilerleyen araştırmalar için önemli bir motivasyon kaynağı olacaktır.

Ek olarak, SCARF çerçevesi kapsamında yapılan ön değerlendirmeler, modelin klinik kabul edilebilirliği konusunda olumlu sonuçlar vermiştir. Ancak, modelin genel klinik ortamda test edilmesi, gerçek hasta verileri ile doğrulama süreçlerinin yürütülmesi, gelecekteki uygulamalar açısından kritik olacaktır. Modelin, radyoterapi planlamasında otomatik segmentasyon süreçlerini hızlandırma potansiyeli, klinik karar destek sistemlerinin gelişimine katkı sağlayabilir.

### 6.2 GELECEK ARAŞTIRMA YÖNLERİ

Gelecekte, farklı ve özel amaçlı loss fonksiyonlarının, özellikle focal loss gibi yöntemlerin modelin performansını nasıl etkileyebileceği üzerine çalışmalar yapılması önemlidir.

Modaliteler arasındaki uyumsuzlukların giderilmesi için gelişmiş veri artırma tekniklerinin uygulanması, modelin daha genellenebilir hale getirilmesine katkı sağlayabilir. Bunun yanı sıra, modelin farklı anatomik bölgelerdeki organ segmentasyonu görevlerine uyarlanması, kullanılabilirlik kapsamını genişletebilir. Daha az işlem gücü gerektiren, ancak doğruluk düzeyini koruyan hibrit model tasarımlarının geliştirilmesi, bu modellerin daha yaygın bir şekilde benimsenmesini sağlayabilir.

Bununla birlikte, modelin gerçek dünya klinik uygulamalarında doğrulama süreçleri, hasta verileri ile test edilmesi ve farklı görüntüleme protokollerine adaptasyonu önemli bir araştırma alanı olarak öne çıkmaktadır. Modelin ölçeklenebilirliğinin artırılması, farklı cihazlarla uyumlu hale getirilmesi ve radyoterapi süreçlerine entegrasyonu için çalışmalar yapılmalıdır. Bunun yanında, segmentasyonun doğruluk seviyesini artırmak için modelin derin öğrenme temelli yeni optimizasyon algoritmalarıyla desteklenmesi gerekmektedir.

Ayrıca, modellerin doğruluk, işlem hızı ve kaynak kullanımı açısından kapsamlı karşılaştırma analizlerinin yapılması, performans değerlendirmelerinde çok boyutlu bir bakış açısı sunacaktır. SCARF çerçevesinin klinik değerlendirme ve genellenebilirlik analizlerinin tamamlanması, modelin gerçek dünya uygulamalarına uygunluğunu artırabilir. Son olarak, gerçek zamanlı segmentasyon uygulamaları için model optimizasyonu ve donanım uyumluluğunun artırılması, daha pratik ve etkili çözümler sunulmasına olanak tanıyacaktır. Ek olarak, yapay zeka destekli segmentasyon modellerinin etik yönleri, hasta verilerinin mahremiyeti ve klinik ortamda kullanımına yönelik düzenleyici gereksinimler de ilerleyen araştırmalar için önemli bir inceleme konusu olabilir.

## KAYNAKLAR

- Adams, R., ve Bischof, L., 1994, Seeded region growing, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16 (6), 641-647. <https://doi.org/10.1109/34.295913>
- Anandhavalli Muniasamy., 2024, Optimization in Neural Networks with Mathematics: Efficiency in Architectures of Deep Learning. *Communications on Applied Nonlinear Analysis*. <https://doi.org/10.52783/cana.v32.2550>
- Andrychowicz, M., Denil, M., Gómez, S., Hoffman, M. W., Pfau, D., Schaul, T., Shillingford, B., & de Freitas, N., 2016, Learning to learn by gradient descent by gradient descent. İçinde D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, & R. Garnett, *Advances in Neural Information Processing Systems*.
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., ve Jemal, A., 2018, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA: A Cancer Journal for Clinicians*, 68 (6), 394-424. <https://doi.org/10.3322/caac.21492>
- Brouwer, C. L., Steenbakkers, R. J. H. M., Bourhis, J., Budach, W., Grau, C., Grégoire, V., Van Herk, M., Lee, A., Maingon, P., Nutting, C., O'Sullivan, B., Porceddu, S. V., Rosenthal, D. I., Sijtsema, N. M., ve Langendijk, J. A., 2015, CT-based delineation of organs at risk in the head and neck region: DAHANCA, EORTC, GORTEC, HKNPCSG, NCIC CTG, NCRI, NRG Oncology and TROG consensus guidelines, *Radiotherapy and Oncology*, 117 (1), 83-90. <https://doi.org/10.1016/j.radonc.2015.07.041>
- Brown, R. W., Cheng, Y. N., Haacke, E. M., Thompson, M. R., ve Venkatesan, R., 2014, *Magnetic Resonance Imaging: Physical Principles and Sequence Design*, Wiley, <https://doi.org/10.1002/9781118633953>
- Boué, L. (2018). Deep learning for pedestrians: Backpropagation in CNNs (No. arXiv:1811.11987). arXiv. <https://doi.org/10.48550/arXiv.1811.11987>
- Canny, J., 1986, A computational approach to edge detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8 (6), 679-698. <https://doi.org/10.1109/TPAMI.1986.4767851>
- Chen, Z., Li, C., He, J., Ye, J., Song, D., Wang, S., Gu, L., & Qiao, Y., 2021, A Novel Hybrid Convolutional Neural Network for Accurate Organ Segmentation in 3D Head and Neck CT Images , *arXiv*. <https://doi.org/10.48550/arXiv.2109.12634>
- Cherry, S. R., Sorenson, J. A., ve Phelps, M. E., 2012, *Physics in Nuclear Medicine E-Book*, Elsevier Health Sciences.
- Cho, M. J., Hwang, D., Yie, S. Y., ve Lee, J. S., 2024, Multi-modal co-learning with attention mechanism for head and neck tumor segmentation on 18FDG PET-CT, *EJNMMI Physics*, 11 (1), 67. <https://doi.org/10.1186/s40658-024-00670-y>

- Chow, L. Q. M., 2020, Head and Neck Cancer, *New England Journal of Medicine*, 382 (1), 60-72. <https://doi.org/10.1056/NEJMra1715715>
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., ve Ronneberger, O., 2016, 3D U-Net: Learning dense volumetric segmentation from sparse annotation, In: Ourselin, S., Joskowicz, L., Sabuncu, M. R., Unal, G., ve Wells, W. (ed.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, Springer International Publishing, 424-432. [https://doi.org/10.1007/978-3-319-46723-8\\_49](https://doi.org/10.1007/978-3-319-46723-8_49)
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, *arXiv*, <https://doi.org/10.48550/arXiv.1810.04805>
- Fehrenbach, M. J., ve Herring, S. W., 2012, *Illustrated Anatomy of the Head and Neck*, Elsevier/Saunders, ISBN: 9781437724196.
- Fitzpatrick, J. M., ve Sonka, M., 2009, *Handbook of Medical Imaging, Volume 2: Medical Image Processing and Analysis*, SPIE Press.
- Francis, S., Pooloth, G., Singam, S. B. S., Puzhakkal, N., Pulinthanathu Narayanan, P., ve Pottekkattuvalappil Balakrishnan, J., 2023, SABOS-Net: Self-supervised attention based network for automatic organ segmentation of head and neck CT images, *International Journal of Imaging Systems and Technology*, 33 (1), 175-191. <https://doi.org/10.1002/ima.22802>
- Fritscher, K., Raudaschl, P., Zaffino, P., Spadea, M. F., Sharp, G. C., ve Schubert, R., 2016, Deep Neural Networks for Fast Segmentation of 3D Medical Images, In: Ourselin, S., Joskowicz, L., Sabuncu, M. R., Unal, G., ve Wells, W. (ed.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, Springer International Publishing, 158-165. [https://doi.org/10.1007/978-3-319-46723-8\\_19](https://doi.org/10.1007/978-3-319-46723-8_19)
- Galdran, A., Carneiro, G., ve Ballester, M. Á. G., 2022, On the Optimal Combination of Cross-Entropy and Soft Dice Losses for Lesion Segmentation with Out-of-Distribution Robustness, *arXiv*. <https://doi.org/10.48550/arXiv.2209.06078>
- Gonzalez, R. C., Woods, R. E., ve Masters, B. R., 2009, *Digital Image Processing, Third Edition*, Prentice Hall.
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., ve Bengio, Y., 2014, Generative Adversarial Networks, *arXiv*. <https://doi.org/10.48550/arXiv.1406.2661>
- Goodfellow, I., Bengio, Y., ve Courville, A., 2016, *Deep Learning*, MIT Press.
- Gören, Ö., Karahan, M., ve Kamyşlıoğlu, M., 2021., *Baş Boyun Radyoterapisi Uygulanan Hastalarda Adaptif Uygulamaların Hedef ve Kritik Organlara Etkilerinin İncelenmesi*.
- Guo, D., Jin, D., Zhu, Z., Ho, T.-Y., Harrison, A. P., Chao, C.-H., Xiao, J., & Lu, L., 2020, Organ at Risk Segmentation for Head and Neck Cancer Using Stratified Learning and

- Neural Architecture Search. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, <https://doi.org/10.1109/CVPR42600.2020.00428>
- Hastie, T., Tibshirani, R., & Friedman, J. H., 2009,. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. *Springer*.
- He, K., Zhang, X., Ren, S., ve Sun, J., 2016, Deep Residual Learning for Image Recognition, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- He, Y., Guo, P., Tang, Y., Myronenko, A., Nath, V., Xu, Z., Yang, D., Zhao, C., Xu, D., ve Li, W., 2024, A Short Review and Evaluation of SAM2's Performance in 3D CT Image Segmentation, *arXiv*. <https://doi.org/10.48550/arXiv.2408.11210>
- Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. R., 2012, Improving neural networks by preventing co-adaptation of feature detectors, *arXiv*, <https://doi.org/10.48550/arXiv.1207.0580>
- Hochreiter, S., & Schmidhuber, J., 1997, Long Short-Term Memory. *Neural Computation*, <https://doi.org/10.1162/neco.1997.9.8.1735>
- He, Z., Wang, K., Fang, T., Su, L., Chen, R., ve Fei, X., 2024, Comprehensive Performance Evaluation of YOLOv11, YOLOv10, YOLOv9, YOLOv8 and YOLOv5 on Object Detection of Power Equipment, *arXiv*. <https://doi.org/10.48550/arXiv.2411.18871>
- Hounsfield, G. N., 1973, Computerized transverse axial scanning (tomography): Part 1. Description of system, *British Journal of Radiology*, 46 (552), 1016-1022. <https://doi.org/10.1259/0007-1285-46-552-1016>
- Ioffe, S., & Szegedy, C., 2015, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, *arXiv*, <https://doi.org/10.48550/arXiv.1502.03167>
- IBM, 2024, What Is Deep Learning?, *IBM Think*. [Ziyaret tarihi: 16 Aralık 2024]. <https://www.ibm.com/think/topics/deep-learning>
- James, A. P., ve Dasarathy, B. V., 2014, Medical image fusion: A survey of the state of the art, *Information Fusion*, 19, 4-19. <https://doi.org/10.1016/j.inffus.2013.12.002>
- Kalender, W., 2000, *Computed Tomography: Fundamentals, System Technology, Image Quality, Applications*, MCD Verlag.
- Khanam, R., ve Hussain, M., 2024, YOLOv11: An Overview of the Key Architectural Enhancements, *arXiv*. <https://doi.org/10.48550/arXiv.2410.17725>
- Kingma, D. P., & Ba, J., 2017, Adam: A Method for Stochastic Optimization, *arXiv*, <https://doi.org/10.48550/arXiv.1412.6980>
- Klein, S., Staring, M., Murphy, K., Viergever, M. A., ve Pluim, J. P. W., 2010, elastix: A Toolbox for Intensity-Based Medical Image Registration, *IEEE Transactions on Medical Imaging*, 29 (1), 196-205. <https://doi.org/10.1109/TMI.2009.2035616>

- Krizhevsky, A., Sutskever, I., & Hinton, G. E., 2017, ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, <https://doi.org/10.1145/3065386>
- Korte, J. C., Hardcastle, N., Ng, S. P., Clark, B., Kron, T., ve Jackson, P., 2021, Cascaded deep learning-based auto-segmentation for head and neck cancer patients: Organs at risk on T2-weighted magnetic resonance imaging, *Medical Physics*, 48 (12), 7757-7772. <https://doi.org/10.1002/mp.15290>
- Lauterbur, P. C., 1973, Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance, *Nature*, 242 (5394), 190-191. <https://doi.org/10.1038/242190a0>
- Li, G. Y., Chen, J., Jang, S.-I., Gong, K., ve Li, Q., 2024, SwinCross: Cross-modal Swin Transformer for Head-and-Neck Tumor Segmentation in PET/CT Images, *Medical Physics*, 51 (3), 2096-2107. <https://doi.org/10.1002/mp.16703>
- Li, S., Hua, H.-L., Li, F., Kong, Y.-G., Zhu, Z.-L., Li, S.-L., Chen, X.-X., Deng, Y.-Q., ve Tao, Z.-Z., 2022, Anatomical Partition-Based Deep Learning: An Automatic Nasopharyngeal MRI Recognition Scheme, *Journal of Magnetic Resonance Imaging*, 56 (4), 1220-1229. <https://doi.org/10.1002/jmri.28112>
- Li, Z., Zhao, K., Wang, Y., & Wang, S., 2024, Adaptive Interactive Segmentation for Multimodal Medical Imaging via Selection Engine, *arXiv*, <https://doi.org/10.48550/arXiv.2411.19447>
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Van Der Laak, J. A. W. M., Van Ginneken, B., ve Sánchez, C. I., 2017, A survey on deep learning in medical image analysis, *Medical Image Analysis*, 42, 60-88. <https://doi.org/10.1016/j.media.2017.07.005>
- Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., ve Guo, B., 2021, Swin Transformer: Hierarchical Vision Transformer using Shifted Windows, *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 9992-10002. <https://doi.org/10.1109/ICCV48922.2021.00986>
- Lowekamp, B. C., Chen, D. T., Ibanez, L., ve Blezek, D., 2013, The Design of SimpleITK, *Frontiers in Neuroinformatics*, 7. <https://doi.org/10.3389/fninf.2013.00045>
- Ma, J., He, Y., Li, F., Han, L., You, C., ve Wang, B., 2024, Segment Anything in Medical Images, *Nature Communications*, 15 (1), 654. <https://doi.org/10.1038/s41467-024-44824-z>
- Ma, J., & Yang, X., 2020, Combining CNN and Hybrid Active Contours for Head and Neck Tumor Segmentation in CT and PET images, *arXiv*, <https://doi.org/10.48550/arXiv.2012.14207>
- Mansoori, M., Shahabodini, S., Abouei, J., Plataniotis, K. N., ve Mohammadi, A., 2024, Self-Prompting Polyp Segmentation in Colonoscopy using Hybrid Yolo-SAM 2 Model, *arXiv*. <https://doi.org/10.48550/ARXIV.2409.09484>

- Mao, A., Mohri, M., ve Zhong, Y., 2023, Cross-Entropy Loss Functions: Theoretical Analysis and Applications, *arXiv*. <https://doi.org/10.48550/ARXIV.2304.07288>
- Marsilla, J., Kim, J. W., Tkachuck, D., Kim, S., Siraj, J., Cho, J., Hahn, E., Hosni, A., Jacinto, K., Welch, M. L., Kazmierski, M., Rey-McIntyre, K., Huang, S. H., Patel, T., Tadic, T., Liu, F.-F., Bratman, S., Hope, A., ve Haibe-Kains, B., 2023a, SCARF: Auto-Segmentation Clinical Acceptability ve Reproducibility Framework for Benchmarking Essential Radiation Therapy Targets in Head and Neck Cancer, *medRxiv*. <https://doi.org/10.1101/2022.01.15.22269276>
- Marsilla, J., Kim, J. W., Tkachuck, D., Kim, S., Siraj, J., Cho, J., Hahn, E., Hosni, A., Jacinto, K., Welch, M. L., Kazmierski, M., Rey-McIntyre, K., Huang, S. H., Patel, T., Tadic, T., Liu, F.-F., Bratman, S., Hope, A., ve Haibe-Kains, B., 2023b, SCARF: Auto-Segmentation Clinical Acceptability ve Reproducibility Framework for Benchmarking Essential Radiation Therapy Targets in Head and Neck Cancer, *medRxiv*. <https://doi.org/10.1101/2022.01.15.22269276>
- Marur, S., ve Forastiere, A. A., 2016, Head and Neck Squamous Cell Carcinoma: Update on Epidemiology, Diagnosis, and Treatment, *Mayo Clinic Proceedings*, 91 (3), 386-396. <https://doi.org/10.1016/j.mayocp.2015.12.017>
- Mattes, D., Haynor, D. R., Vesselle, H., Lewellyn, T. K., ve Eubank, W., 2001, Nonrigid multimodality image registration, *Medical Imaging 2001: Image Processing*, 4322, 1609-1620. <https://doi.org/10.1117/12.431046>
- Mattjie, C., Moura, L. V. de, Ravazio, R. C., Kupssinskü, L. S., Parraga, O., Delucis, M. M., & Barros, R. C., 2023, Zero-shot performance of the Segment Anything Model (SAM) in 2D medical imaging: A comprehensive evaluation and practical guidelines, *arXiv*, <https://doi.org/10.48550/arXiv.2305.00109>
- Mayo, C. S., Moran, J. M., Bosch, W., Xiao, Y., McNutt, T., Popple, R., Michalski, J., Feng, M., Marks, L. B., Fuller, C. D., Yorke, E., Palta, J., Gabriel, P. E., Molineu, A., Matuszak, M. M., Covington, E., Masi, K., Richardson, S. L., Ritter, T., ... Yock, T. I., 2018, American Association of Physicists in Medicine Task Group 263: Standardizing Nomenclatures in Radiation Oncology, *International Journal of Radiation OncologyBiologyPhysics*, 100 (4), 1057-1066. <https://doi.org/10.1016/j.ijrobp.2017.12.013>
- Mazurowski, M. A., Dong, H., Gu, H., Yang, J., Konz, N., & Zhang, Y. (2023). Segment Anything Model for Medical Image Analysis: An Experimental Study. *Medical Image Analysis*, 89, 102918. <https://doi.org/10.1016/j.media.2023.102918>
- Meng, M., Bi, L., Fulham, M., Feng, D., ve Kim, J., 2023, Merging-Diverging Hybrid Transformer Networks for Survival Prediction in Head and Neck Cancer, In: Greenspan, H., Madabhushi, A., Mousavi, P., Salcudean, S., Duncan, J., Syeda-Mahmood, T., ve Taylor, R. (ed.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023*, Springer Nature Switzerland, 14225, 400-410. [https://doi.org/10.1007/978-3-031-43987-2\\_39](https://doi.org/10.1007/978-3-031-43987-2_39)

- Michelutti, L., Tel, A., Zeppieri, M., Ius, T., Agosti, E., Sembronio, S., ve Robiony, M., 2024, Generative Adversarial Networks (GANs) in the Field of Head and Neck Surgery: Current Evidence and Prospects for the Future—A Systematic Review, *Journal of Clinical Medicine*, 13 (12), Article 12. <https://doi.org/10.3390/jcm13123556>
- Milletari, F., Navab, N., ve Ahmadi, S.-A., 2016, V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation, *arXiv*. <https://doi.org/10.48550/arXiv.1606.04797>
- Mitchell, T. M., 2013, *Machine learning*. McGraw-Hill.
- Mortazi, A., Cicek, V., Keles, E., ve Bagci, U., 2023, Selecting the Best Optimizers for Deep Learning based Medical Image Segmentation, *arXiv*. <https://doi.org/10.48550/arXiv.2302.02289>
- Oktaç, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N. Y., Kainz, B., Glocker, B., ve Rueckert, D., 2018, Attention U-Net: Learning Where to Look for the Pancreas, *arXiv*. <https://arxiv.org/abs/1804.03999v3>
- Otsu, N., 1979, A Threshold Selection Method from Gray-Level Histograms, *IEEE Transactions on Systems, Man, and Cybernetics*, 9 (1), 62-66. <https://doi.org/10.1109/TSMC.1979.4310076>
- Ozdemir, S., 2013, Adaptive radiotherapy, *Turkish Journal of Oncology*, 28 (3), 136-141. <https://doi.org/10.5505/tjoncol.2013.990>
- Pandey, S., Chen, K.-F., & Dam, E. B., 2023, Comprehensive Multimodal Segmentation in Medical Imaging: Combining YOLOv8 with SAM and HQ-SAM Models. *2023 IEEE/CVF International Conference on Computer Vision Workshops*, <https://doi.org/10.1109/ICCVW60793.2023.00273>
- Peker, D. S. K., ve Özdoğan, Prof. Dr. H. A., 2023, Baş-Boyun Kanserlerinde Radyoterapinin Orta Kulağa ve Vestibüler Sisteme Etkileri.
- Phelps, M. E., 2004, *PET: Molecular Imaging and Its Biological Applications*, Springer Science ve Business Media.
- Pianyk, O. S., 2012, *Digital Imaging and Communications in Medicine (DICOM)*, Springer, <https://doi.org/10.1007/978-3-642-10850-1>
- Pluim, J. P. W., Maintz, J. B. A., ve Viergever, M. A., 2003, Mutual-information-based registration of medical images: A survey, *IEEE Transactions on Medical Imaging*, 22 (8), 986-1004. <https://doi.org/10.1109/TMI.2003.815867>
- Podobnik, G., Ibragimov, B., Tappeiner, E., Lee, C., Kim, J. S., Mesbah, Z., Modzelewski, R., Ma, Y., Yang, F., Rudecki, M., Wodziński, M., Peterlin, P., Strojjan, P., ve Vrtovec, T., 2024, HaN-Seg: The head and neck organ-at-risk CT and MR segmentation challenge, *Radiotherapy and Oncology*, 198, 110410. <https://doi.org/10.1016/j.radonc.2024.110410>

- Podobnik, G., Strojjan, P., Peterlin, P., Ibragimov, B., ve Vrtovec, T., 2023a, HaN-Seg: The head and neck organ-at-risk CT and MR segmentation dataset, *Medical Physics*, 50 (3), 1917-1927. <https://doi.org/10.1002/mp.16197>
- Podobnik, G., Strojjan, P., Peterlin, P., Ibragimov, B., ve Vrtovec, T., 2023b, Multimodal CT and MR Segmentation of Head and Neck Organs-at-Risk, In: Greenspan, H., Madabhushi, A., Mousavi, P., Salcudean, S., Duncan, J., Syeda-Mahmood, T., ve Taylor, R. (ed.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023*, Springer Nature Switzerland, 745-755. [https://doi.org/10.1007/978-3-031-43901-8\\_71](https://doi.org/10.1007/978-3-031-43901-8_71)
- Qader, S. M., Hassan, B. A., ve Rashid, T. A., 2022, An Improved Deep Convolutional Neural Network by Using Hybrid Optimization Algorithms to Detect and Classify Brain Tumor Using Augmented MRI Images, *Multimedia Tools and Applications*, 81 (30), 44059-44086. <https://doi.org/10.1007/s11042-022-13260-w>
- Quetin, S., Heschl, A., Murillo, M., Murali, R., Enger, S. A., ve Maleki, F., 2024, Automatic segmentation of Organs at Risk in Head and Neck cancer patients from CT and MRI scans, *arXiv*. <https://doi.org/10.48550/arXiv.2405.10833>
- Ragab, M. G., Abdulkadir, S. J., Muneer, A., Alqushaibi, A., Sumiea, E. H., Qureshi, R., Al-Selwi, S. M., ve Alhussian, H., 2024, A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023), *IEEE Access*, 12, 57815-57836. <https://doi.org/10.1109/ACCESS.2024.3386826>
- Rahul Agarwal, 2024, Complete Guide to the Adam Optimization Algorithm. *Built In*. [Ziyaret tarihi: 20 Aralık 2024], <https://builtin.com/machine-learning/adam-optimization>
- Raudaschl, P. F., Zaffino, P., Sharp, G. C., Spadea, M. F., Chen, A., Dawant, B. M., Albrecht, T., Gass, T., Langguth, C., Lüthi, M., Jung, F., Knapp, O., Wesarg, S., Mannion-Haworth, R., Bowes, M., Ashman, A., Guillard, G., Brett, A., Vincent, G., ... Fritscher, K. D., 2017, Evaluation of segmentation methods on head and neck CT: Auto-segmentation challenge 2015, *Medical Physics*, 44 (5), 2020-2036. <https://doi.org/10.1002/mp.12197>
- Ramachandran, P., Zoph, B., & Le, Q. V., 2017, Searching for Activation Functions, *arXiv*, <https://doi.org/10.48550/arXiv.1710.05941>
- Ren, J., Rasmussen, M., Nijkamp, J., Eriksen, J. G., & Korreman, S., 2024, Segment anything model for head and neck tumor segmentation with CT, PET and MRI multi-modality images, *arXiv*, <https://doi.org/10.48550/arXiv.2402.17454>
- Ronneberger, O., Fischer, P., ve Brox, T., 2015, U-Net: Convolutional Networks for Biomedical Image Segmentation, In: Navab, N., Hornegger, J., Wells, W. M., ve Frangi, A. F. (ed.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Springer International Publishing, 234-241. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)

- Shulman, D., 2023, Optimization Methods in Deep Learning: A Comprehensive Overview, *arXiv*, <https://doi.org/10.48550/arXiv.2302.09566>
- Simonyan, K., & Zisserman, A., 2015, Very Deep Convolutional Networks for Large-Scale Image Recognition, *arXiv*. <https://doi.org/10.48550/arXiv.1409.1556>
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R., 2014, Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*.
- Srivastava, A., Jha, D., Keles, E., Aydogan, B., Abazeed, M., & Bagci, U., 2022, An Efficient Multi-Scale Fusion Network for 3D Organ at Risk (OAR) Segmentation, *arXiv*, <https://doi.org/10.48550/arXiv.2208.07417>
- Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S., ve Cardoso, M. J., 2017, Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations, *Lecture Notes in Computer Science*, 10553, 240-248. [https://doi.org/10.1007/978-3-319-67558-9\\_28](https://doi.org/10.1007/978-3-319-67558-9_28)
- Taha, A. A., ve Hanbury, A., 2015, Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool, *BMC Medical Imaging*, 15 (1), 29. <https://doi.org/10.1186/s12880-015-0068-x>
- Terven, J., Cordova-Esparza, D. M., Ramirez-Pedraza, A., Chavez-Urbiola, E. A., ve Romero-Gonzalez, J. A., 2024, Loss Functions and Metrics in Deep Learning, *arXiv*. <https://doi.org/10.48550/arXiv.2307.02694>
- Trigka, M., & Dritsas, E., 2025, A Comprehensive Survey of Deep Learning Approaches in Image Processing, *Sensors*, <https://doi.org/10.3390/s25020531>
- Ultralytics, 2024, YOLO11 NEW, *Ultralytics Documentation*, [Ziyaret tarihi: 16 Aralık 2024]. <https://docs.ultralytics.com/models/yolo11>
- Valanarasu, J. M. J., Oza, P., Hacihaliloglu, I., ve Patel, V. M., 2021, Medical Transformer: Gated Axial-Attention for Medical Image Segmentation, *arXiv*. <https://doi.org/10.48550/arXiv.2102.10662>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., ve Polosukhin, I., 2023, Attention Is All You Need, *arXiv*. <https://doi.org/10.48550/arXiv.1706.03762>
- Vrtovec, T., Močnik, D., Strojjan, P., Pernuš, F., ve Ibragimov, B., 2020, Auto-segmentation of organs at risk for head and neck radiotherapy planning: From atlas-based to deep learning methods, *Medical Physics*, 47 (9). <https://doi.org/10.1002/mp.14320>
- Wang, K., Zheng, M., Wei, H., Qi, G., ve Li, Y., 2020, Multi-Modality Medical Image Fusion Using Convolutional Neural Network and Contrast Pyramid, *Sensors*, 20 (8), 2169. <https://doi.org/10.3390/s20082169>

- Wu, X., & Wang, Z., 2020, Deep-learning-based Detection and Segmentation of Organs at Risk in Head and Neck, 2020 Chinese Control And Decision Conference, <https://doi.org/10.1109/CCDC49329.2020.9164800>
- Xie, L., Lomurno, E., Gambella, M., Ardagna, D., Roveri, M., Matteucci, M., ve Shi, Q., 2024, A Lightweight Neural Architecture Search Model for Medical Image Classification, *arXiv*. <https://doi.org/10.48550/arXiv.2405.03462>
- Yan, X., Sun, S., Han, K., Le, T.-T., Ma, H., You, C., & Xie, X., 2024, AFTer-SAM: Adapting SAM with Axial Fusion Transformer for Medical Imaging Segmentation. 2024 IEEE/CVF Winter Conference on Applications of Computer Vision, <https://doi.org/10.1109/WACV57701.2024.00779>
- Yang, J., Veeraraghavan, H., Armato, S. G., Farahani, K., Kirby, J. S., Kalpathy-Kramer, J., Van Elmpt, W., Dekker, A., Han, X., Feng, X., Aljabar, P., Oliveira, B., Van Der Heyden, B., Zamdborg, L., Lam, D., Gooding, M., ve Sharp, G. C., 2018, Autosegmentation for thoracic radiation treatment planning: A grand challenge at AAPM 2017, *Medical Physics*, 45 (10), 4568-4581. <https://doi.org/10.1002/mp.13141>
- Yao, L., Chen, D., Zhao, X., Fei, M., Song, Z., Xue, Z., Zhan, Y., Song, B., Shi, F., Wang, Q., ve Shen, D., 2024, AASeg: Artery-aware Global-to-Local Framework for Aneurysm Segmentation in Head and Neck CTA Images, *IEEE Transactions on Medical Imaging*. <https://doi.org/10.1109/TMI.2024.3496194>
- Yemi Hussain, N., 2024, Deep Learning Architectures Enabling Sophisticated Feature Extraction and Representation for Complex Data Analysis, *International Journal of Innovative Science and Research Technology* , <https://doi.org/10.38124/ijisrt/IJISRT24OCT1521>
- Yosinski, J., Clune, J., Bengio, Y., ve Lipson, H., 2014, How transferable are features in deep neural networks?, *arXiv*. <https://doi.org/10.48550/arXiv.1411.1792>
- Zhu, J., Hamdi, A., Qi, Y., Jin, Y., ve Wu, J., 2024, Medical SAM 2: Segment medical images as video via Segment Anything Model 2, *arXiv*. <https://doi.org/10.48550/arXiv.2408.00874>

## ÖZGEÇMİŞ

Kişisel Bilgiler	
Adı Soyadı	Sinan SEVGİ
Doğum Yeri	
Doğum Tarihi	Tarih girmek için tıklayın veya dokununuz.
Uyruğu	<input checked="" type="checkbox"/> T.C. <input type="checkbox"/> Diğer:
E-Posta Adresi	
Web Adresi	<a href="https://sinansevgi.com/">https://sinansevgi.com/</a>

Eğitim Bilgileri	
Lisans	
Üniversite	Anadolu Üniversitesi
Fakülte	Açıköğretim Fakültesi
Bölümü	Yönetim Bilişim Sistemleri
Mezuniyet Yılı	06.08.2021

Yüksek Lisans	
Üniversite	İstanbul Üniversitesi
Enstitü Adı	Fen Bilimleri Enstitüsü
Anabilim Dalı	Enformatik Anabilim Dalı
Programı	Enformatik Programı