

T.C.
ÇUKUROVA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI

**YAPAY ZEKA DESTEKLİ FCN MODEL PERFORMANSININ
DEĞERLENDİRİLMESİ: KIBT GÖRÜNTÜLERİ ÜZERİNDE
NAZOPALATİN KANALIN OTOMATİK SEGMENTASYON
UYGULAMASI**

Hazal DUYAN YÜKSEL

BİYOİSTATİSTİK YÜKSEK LİSANS PROGRAMI

YÜKSEK LİSANS TEZİ

DANIŞMANI

Prof. Dr. Gülşah SEYDAOĞLU

ADANA-2025

T.C.
ÇUKUROVA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI

**YAPAY ZEKA DESTEKLİ FCN MODEL PERFORMANSININ
DEĞERLENDİRİLMESİ: KIBT GÖRÜNTÜLERİ ÜZERİNDE
NAZOPALATİN KANALIN OTOMATİK SEGMENTASYON
UYGULAMASI**

Hazal DUYAN YÜKSEL

BİYOİSTATİSTİK YÜKSEK LİSANS PROGRAMI

YÜKSEK LİSANS TEZİ

DANIŞMANI

Prof. Dr. Gülşah SEYDAOĞLU

Bu çalışma Çukurova Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğü tarafından TYL-2024-16632 no'lu proje olarak desteklenmiştir.

ADANA-2025

T.C. ÇUKUROVA ÜNİVERSİTESİ SAĞLIK BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS TEZ ÇALIŞMASI

ETİK BEYANI

Çukurova Üniversitesi Bilimsel Araştırma ve Yayın Etiği Yönergesini okuduğumu ve anladığımı ve Çukurova Üniversitesi Sağlık Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmasında;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik kurallarına uygun olarak sunduğumu,
- Tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde ve ortaya çıkan sonuçlarda herhangi bir değişiklik yapmadığımı,
- Tez olarak sunduğum bu çalışmanın özgün olduğunu,

bildirim, aksi bir durumda bu konuda hakkımda yapılacak tüm yasal işlemleri ve aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim. 20/10/2025

Hazal DUYAN YÜKSEL

Kayıtlı olunan Program : Yüksek Lisans

Tezin Konusu : Yapay Zeka Destekli FCN Model Performansının

Değerlendirilmesi: KIBT Görüntüleri Üzerinde Nazoplatin Kanalin Otomatik Segmentasyon Uygulaması

Tezin Türü : Yüksek Lisans : ☒ Doktora: ☐

Danışmanın Adı-Soyadı : Prof. Dr. Gülşah SEYDAOĞLU

Danışmanın İletişim Bilgileri

Telefon :

E-Posta :

Öğrencinin İletişim Bilgileri

Telefon :

E-Posta :

Adresi :

**Bu belgenin Lisansüstü eğitim tezleri savunmaya alınmadan önce öğrenci tarafından doldurulup*

imzalanarak Enstitü Müdürlüğüne teslim edilmesi gerekmektedir.

TEŞEKKÜR

Yüksek lisans eğitimim sürecince tecrübesi ile beni destekleyen değerli tez danışmanım Sayın Prof. Dr. Gülşah SEYDAOĞLU'na,

Tez çalışmamda yardım ve katkıları ile bilgi birikimlerini benimle paylaşan Sayın Dr. Öğr. Üyesi Yusuf Kemal ARSLAN'a ve Sayın Prof. Dr. İbrahim Şevki Bayrakdar'a,

Yüksek lisans eğitimim boyunca desteklerini esirgemeyen Sayın Doç. Dr. Yaşar SERTDEMİR'e, Sayın Doç. Dr. İlker ÜNAL'a, Sayın Doç. Dr. Burcu EVLİCE'ye, Sayın Doç. Dr. Damla SOYDAN ÇABUK'a, Arş. Gör. Hülya BİNOKAY'a, Sayın Arş. Gör. Sevinç Püren YÜCEL'e, Sayın Arş. Gör. Nazlı TOTİK'e,

Destekleri için Çukurova Üniversitesi Sağlık Bilimleri Enstitüsü ve Bilimsel Araştırma Proje Birimi'ne,

Bu günlere gelmemde hiçbir yardım ve fedakarlıktan kaçınmayan, her türlü desteklerini benden hiçbir zaman esirgemeyen sevgili annem Miyase DUYAN'a, sevgili babam Hüseyin DUYAN'a, canım kardeşlerim Ahmet DUYAN ve Çağan DUYAN'a,

Beni her zaman destekleyen ve yanımda olan sevgili eşim Dr. Yüksel YÜKSEL'e ve gelmesini sabırsızlıkla beklediğimiz canım kızıma,

Sevgi, saygı ve tüm içtenliğimle teşekkürlerimi sunarım.

Hazal DUYAN YÜKSEL

İÇİNDEKİLER

KABUL VE ONAY	ii
TEŞEKKÜR	iv
İÇİNDEKİLER	v
ŞEKİLLER DİZİNİ	vii
ÇİZELGELER DİZİNİ	viii
KISALTMALAR DİZİNİ	ix
ÖZET	x
ABSTRACT	xi
1. GİRİŞ	1
2. GENEL BİLGİLER	3
2.1. Yapay Zeka	3
2.1.1. Makine Öğrenimi	4
2.1.2. Derin Öğrenme	5
2.1.3. Hiperparametre Optimizasyonu	9
2.1.4. Yapay Sinir Ağları ile Yaklaşımda Sınıflandırma	16
2.1.5. U-Net	18
2.2. Modellerin Değerlendirilmesinde Kullanılan Performans Metrikleri	19
2.2.1. Karışıklık Matrisi	20
2.2.2. Karışıklık Matrisinden Türetilen Diğer Metrikler	21
2.2.3. Diğer Değerlendirme Metrikleri	23
2.2.4. Receiver Operating Characteristics (ROC)	25
2.2.5. Eğri Altında Kalan Alan (EAA)	26
2.3. Nazopalatin Kanal	30
2.4. Konik Işınlı Bilgisayarlı Tomografi	31

2.4.1. KIBT'nin BT'ye Göre Üstünlükleri	32
2.4.2. KIBT'nin BT'ye Kıyasla Dezavantajları.....	32
3. GEREÇ ve YÖNTEMLER.....	34
3.1. Etik Kurul Onayı.....	34
3.2. Veri Toplama ve Hasta Seçimi	34
3.3. KIBT Analizi.....	35
3.4. Manuel Etiketleme (Ground Truth)	35
3.5. Derin Öğrenme Modelinin Geliştirilmesi	36
3.6. Örneklem Büyüklüğü	38
3.7. İstatistiksel Analiz.....	38
4. BULGULAR	39
5. TARTIŞMA	48
6. SONUÇLAR VE ÖNERİLER.....	60
KAYNAKLAR	61
ÖZGEÇMİŞ	72

ŞEKİLLER DİZİNİ

Şekil No:

Şekil 2. 1. Derin öğrenme sürecinde hiperparametre optimizasyonu	10
Şekil 2. 2. Model eğitimi sırasında epoch ve batch için örnek bir şematizasyon	12
Şekil 2. 13. ROC ve EAA örneği.....	27
Şekil 3. 1. Konik ışınlı bilgisayarlı tomografi görüntülerinde aksiyal, koronal, sagittal kesitlerde nazopalatin kanalın manuel olarak etiketlenmesi ve 3B görünümü.....	36
Şekil 3. 2. Derin öğrenme modeli geliştirmek ve eğitmek için kullanılan metodolojik iş akışı.....	37
Şekil 4. 1. Aksiyel KIBT kesitlerinde nazopalatin kanalın orijinal görüntüleri ve nnU-Net v2 derin öğrenme modelinin tahmin (prediction) görüntüleri	39
Şekil 4. 2. Test setinde hasta bazlı Dice katsayılarının grafiksel gösterimi.....	43
Şekil 4. 3. Test setinde hasta bazlı Jaccard indeksi (IoU) skorlarının grafiksel gösterimi	44
Şekil 4. 4. Test setinde hasta bazlı kesinlik değerlerinin grafiksel gösterimi.....	44
Şekil 4. 5. Test setinde hasta bazlı duyarlılık (recall) değerlerinin grafiksel gösterimi .	45
Şekil 4. 6. ROC eğrisi ve EAA	45
Şekil 4. 7. Döngü (epoch) boyunca Dice katsayısı (a), kayıp fonksiyonları, epoch süresi (b) ve öğrenme oranı (c)	47

ÇİZELGELER DİZİNİ

Çizelge No:

Çizelge 2. 1. İkili Sınıflandırma Görevi için Karışıklık Matrisi.....	20
Çizelge 2. 2. EAA değerlerinin ayırt edilebilirlik seviyeleri	28
Çizelge 2. 3. Segmentasyon performans ölçütleri tablosu.....	29
Çizelge 3. 1. nnU-Net v2 modelinin eğitim parametreleri	37
Çizelge 4. 1. nnU-Net v2 modelinin performans metrikleri ve değerleri	40
Çizelge 4. 2. nnU-Net v2 modelinin ve test seti hastalarının performans metrikleri ve değerleri	42

KISALTMALAR DİZİNİ

BT	: Bilgisayarlı Tomografi
Adam	: Adaptive Moment Estimation
CNN	: Convolutional Neural Networks (Evrişimsel Sinir Ağları)
DK	: Dice Katsayısı
EAA	: Eğri Altında Kalan Alan
fCN	: Fully Convolutional Neural Network (Tam Evrişimsel Sinir Ağı)
FOV	: Field of View (Görüş Alanı)
GP	: Gerçek Pozitif
GN	: Gerçek Negatif
IoU	: Intersection over Union (Kesişim / Birleşim Oranı, Jaccard İndeksi)
KIBT	: Konik Işınlı Bilgisayarlı Tomografi
MRG	: Manyetik Rezonans Görüntüleme
NifTI	: Neuroimaging Informatics Technology Initiative
NPK	: Nazopalatin Kanal
OPG	: Ortopantomografi
ROC	: Receiver Operating Characteristic
SGD	: Stochastic gradient descent
USG	: Ultrason Görüntüleme
YN	: Yanlış Negatif
YP	: Yanlış Pozitif

ÖZET

Yapay Zeka Destekli FCN Model Performansının Değerlendirilmesi: KIBT Görüntüleri Üzerinde Nazopalatin Kanalı Otomatik Segmentasyon Uygulaması

Derin öğrenme, tıbbi görüntü analizinde önemli bir çerçeve haline gelmiş olup, ilgi bölgelerinin segmentasyonunu otomatikleştirmekte ve standartlaştırmaktadır. Bu süreç, bilgisayar destekli tanı, girişimsel işlemler ve tedavi için kritik öneme sahiptir. Konik ışınli bilgisayarlı tomografi (KIBT), maksillofasiyal yapıların ayrıntılı üç boyutlu görüntülerini sunarak dental görüntüleme için çığır açmış olmasına rağmen, KIBT kullanılarak NPK segmentasyonunun çeşitli zorlukları mevcuttur. Bu çalışmada, derin öğrenme tabanlı nnU-Net v2 model performansına dayalı KIBT görüntülerinde NPK'nın otomatik segmentasyonunun değerlendirilmesi amaçlanmıştır.

Çalışma retrospektif olarak yürütülmüş, veriler eğitim (n=82) ve test (n=8) seti olmak üzere iki gruba ayrılmıştır. Otomatik segmentasyon, klinik radyoloji ve yapay zekâ alanında yaygın olarak kullanılan performans metrikleri (gerçek pozitif, yanlış pozitif, yanlış negatif, doğruluk, kesinlik, duyarlılık, F1-skoru, Dice katsayısı, Jaccard indeksi [IoU], %95 Hausdorff mesafesi [mm] ve Hausdorff mesafesi [mm]) ile değerlendirilmiştir. Ayrıca ROC eğrisi elde edilerek eğri altında kalan alan (EAA) hesaplanmıştır.

NPK segmentasyonu için kullanılan nnU-Net v2 mimarisi sonucunda elde edilen karmaşıklık matrisine göre gerçek pozitif, yanlış pozitif, yanlış negatif, doğruluk, kesinlik, duyarlılık ve F1-skoru değerleri sırasıyla 914, 282,13, 131,25, 3,09, 0,11, 0,99, 0,74, 0,89 ve 0,81 olarak bulunmuştur. Ayrıca Dice katsayısı, Jaccard indeksi (IoU), %95 Hausdorff mesafesi (mm) ve Hausdorff mesafesi (mm) sırasıyla 0,80, 0,67, 1,02 mm ve 1,90 mm olarak hesaplanmıştır. EAA ise 0,94 olarak analiz edilmiştir.

Bu sonuçlar, nnU-Net v2 modelinin nazopalatin kanal segmentasyonunda yüksek doğruluk ve güvenilirlik sağladığını göstermektedir.

Anahtar Kelimeler: Derin öğrenme, konik ışınli bilgisayarlı tomografi, nazopalatin kanal, nnU-Net v2, performans metrikleri

ABSTRACT

Evaluation of Artificial Intelligence Assisted FCN Model Performance: Application of Automatic Segmentation of the Nasopalatine Canal on CBCT Images

Deep learning has become an important framework in medical image analysis, automating and standardizing the segmentation of regions of interest. This process is crucial for computer-aided diagnosis, interventional procedures, and treatment. Although cone-beam computed tomography(CBCT) has revolutionized dental imaging by providing detailed three-dimensional views of maxillofacial structures, several challenges remain in segmenting the nasopalatine canal(NPC) using CBCT. This study aimed to evaluate the automatic segmentation of the NPC on CBCT images based on the performance of a deep learning–based nnU-Net v2 model.

The study was conducted retrospectively, and the data were divided into two groups as training (n=82) and test (n=8) sets. Automatic segmentation was evaluated using performance metrics commonly employed in clinical radiology and artificial intelligence, including true positives, false positives, false negatives, accuracy, precision, recall, F1-score, Dice coefficient, Jaccard index (IoU), 95% Hausdorff distance (mm), and Hausdorff distance (mm). In addition, the receiver operating characteristic (ROC) curve was generated, and the area under the curve (AUC) was calculated.

According to the confusion matrix obtained with the nnU-Net v2 architecture for NPC segmentation, the values of true positives, false positives, false negatives, accuracy, precision, recall, and F1-score were found to be 914, 282.13, 131.25, 3.09, 0.11, 0.99, 0.74, 0.89, and 0.81, respectively. Furthermore, the Dice coefficient, Jaccard index (IoU), 95% Hausdorff distance (mm), and Hausdorff distance (mm) were calculated as 0.80, 0.67, 1.02 mm, and 1.90 mm, respectively. The AUC value derived from the ROC curve was 0.94.

These findings demonstrate that the nnU-Net v2 model provides high accuracy and reliability in the segmentation of the nasopalatine canal.

Keywords: Cone-beam computed tomography, deep learning, nasopalatine canal, nnU-Net v2, performance metrics

1. GİRİŞ

Sağlık alanında ve ağız, diş ve çene radyolojisinde güncel ve giderek daha fazla ilgi gören konulardan biri yapay zekâdır. Yapay zekâ, bilgisayar sistemlerinin genellikle “zeki davranış” olarak adlandırılan olguyu anlamasını sağlayan bir bilim ve mühendislik alanıdır¹. Bilgisayar tabanlı tanı sistemlerinin kullanımı, insan gözü tarafından fark edilemeyecek lezyonları saptama ve tanılama kapasitesi sayesinde giderek artmaktadır. Makine öğrenmesi, yapay zekânın bir alt dalıdır ve bilgisayarların belirli bir görevi yerine getirebilmek için örnek veri kümelerinden kurallar çıkararak kendi kendine öğrenen sistemler oluşturmaya dayanır. Bu yaklaşımda, doğru ve klinik olarak faydalı modeller geliştirmek için doğrudan kod yazmak yerine, algoritmalar örnek veri setlerinden kendi mantıklarını üretir². Derin öğrenme, tıbbi görüntü analizinde önemli bir çerçeve haline gelmiş olup, ilgi bölgelerinin segmentasyonunu otomatikleştirmekte ve standartlaştırmaktadır. Bu süreç, bilgisayar destekli tanı, girişimsel işlemler ve tedavi için kritik öneme sahiptir³. Bunun önemine rağmen, manuel görüntü segmentasyonu zahmetli ve zaman alıcı bir iş olabilmekte ve kullanıcılar arasında değişken sonuçlara yol açabilmektedir^{4,5}. Tam otomatik segmentasyon ise zaman tasarrufu sağlamak ve daha tutarlı sonuçlar üretmektedir^{4,5}. Convolutional neural networks (CNN) anatomik yapıların veya patolojilerin etiketlenmesi gibi segmentasyon görevlerinde dikkate değer başarılar elde etmiştir⁶⁻⁹.

Baş ve boyun görüntüleme alanında, yapay zekâ; dijital radyografi, bilgisayarlı tomografi (BT), konik ışınlı bilgisayarlı tomografi (KIBT), ultrason görüntüleme (USG) ve manyetik rezonans görüntüleme (MRG) gibi sistemlere entegre edilerek, olağandışı küçük sapmaları öğrenme yeteneği sayesinde tanıyabilir ve anatomik belirteçlerin tanımlanması, kistik ve tümöral lezyonların sınıflandırılması, ayrıca diş çürükleri, periodontal ve periapikal patolojiler ile temporomandibular eklem bozukluklarının tanısına yardımcı olmak gibi görevlerde yaygın olarak kullanılmaktadır.¹⁰⁻¹⁵.

Nazopalatin kanal (NPK), ağız boşluğu ve burun boşluğunu birbirine bağlayan içinden arter, sinir gibi yapıların geçtiği önemli bir anatomik yapıdır^{16,17}. NPK’nın morfolojisi hakkında bilgi sahibi olmak özellikle patoloji ya da implant cerrahisi gibi durumlar için kritik önem taşımaktadır¹⁸⁻²¹. KIBT, maksillofasiyal yapıların ayrıntılı üç

boyutlu görüntülerini sunarak dental görüntülemeye çıđır açmış olmasına rağmen^{22,23}, KIBT kullanılarak NPK segmentasyonunun çeşitli zorlukları mevcuttur. Bunlar arasında maksillofasiyal yapıların karmaşık anatomisi, görüntü gürültüsü, belirsiz sınırlar ve sınırlı sayıda veri kümesi sayılabilir^{24,25}. Bu amaçla KIBT görüntülerinden derin öğrenme modeli ile NPK'yı yüksek performans ile segmente edilebilmesini sağlayan çalışmaların geliştirilmesi yarar sağlayacaktır. Mevcut literatürdeki boşluğu doldurmayı amaçlayan bu araştırma, H1 hipotezine dayanmakta olup, KIBT görüntülerinden derin öğrenme modeli ile NPK'yı yüksek performans ile segmente edilebileceđi yönündedir.

Bu çalışmada, derin öğrenme tabanlı nnU-Net v2 model performansının değerlendirilerek bu modelin KIBT görüntülerinde NPK'yı otomatik olarak segmentlemek için kullanımının değerlendirilmesi amaçlanmıştır.

2. GENEL BİLGİLER

2.1. Yapay Zeka

Yapay zeka, insanlara özgü bilişsel becerileri (problem çözme, öğrenme, karar verme ve muhakeme) taklit edebilen makineler ve teknolojiler bütünü olarak tanımlanmaktadır³⁷. Bu alandaki ilk önemli adımlar Alan Turing tarafından atılmıştır. Turing, 1950’de yayımladığı çalışmasında “Makineler düşünebilir mi?” sorusunu ortaya koymuş ve “Taklit Oyunu” adını verdiği, günümüzde ise Turing Testi olarak bilinen yaklaşımı tanımlamıştır. Bu testte bir bilgisayarın, insanla etkileşim sırasında ayırt edilemeyecek düzeyde yanıtlar verebilmesi, onun zekâya sahip olduğunun göstergesi kabul edilmektedir³⁸⁻⁴⁰.

Yapay zekâ 1990’ların sonunda IBM’in Deep Blue sistemi, dünya satranç şampiyonu Garry Kasparov’u yenerek yapay zekânın potansiyelini tüm dünyaya göstermiştir⁴¹. 2000’lerden itibaren artan işlem gücü, büyük veri kümelerinin erişilebilirliği ve gelişmiş algoritmalar sayesinde yapay zekâ, makine öğrenmesi, derin öğrenme, doğal dil işleme ve robotik gibi alt dallarda hızla gelişim göstermiştir. 2012 yılında derin sinir ağları kullanılarak ImageNet’te elde edilen başarı, derin öğrenmenin modern yapay zekâ uygulamalarındaki merkezi rolünü pekiştirmiştir⁴².

Yapay zekâ, doğal zekâdan belirli yönleriyle ayrılmaktadır. İnsan zekâsı sürekli öğrenme, deneyim biriktirme, yaratıcılık ve sezgi gibi özelliklere sahipken, yapay zekâ yalnızca kendisine öğretilen veri setleriyle sınırlı bir şekilde çalışır. Bu nedenle, yapay zekâ bilgiyi depolama ve aktarma açısından kalıcı ve hızlı avantajlar sunsa da, insan zekâsındaki esneklik ve yaratıcılıktan yoksundur^{43,44}.

Sonuç olarak, yapay zeka günümüzde hem teorik hem de pratik boyutlarıyla hızla gelişen, farklı disiplinlerin kesişiminde yer alan ve geleceğin teknolojik dönüşümünde kritik rol üstlenen bir alandır.

Yapay zekâ, karmaşık ve zor verilerin bilgisayarlı algoritmalar aracılığıyla değerlendirildiği bir teknolojidir. Tıbbi görüntülemeye çok sayıda uygulama alanı

bulmaktadır. Çıkış değişkeni sürekli olduğunda regresyon temelli görevler, kategorik olduğunda ise sınıflandırma temelli görevler için denetimli makine öğrenmesi kullanılmaktadır⁴⁵.

Yapay zekânın tıbbi görüntülemeye sunduğu uygulamalar arasında tanısal görüntüleme en önemli yaklaşımlardandır. Çalışmalar, bu bilgisayarlı algoritmaların radyografik anormallikleri tespit etmede yüksek doğruluk, duyarlılık ve iyi özgüllüğe sahip olduğunu göstermektedir. Radyomik özellik çıkarımı ve seçimi, patoloji tespiti için yapay zekâ yaklaşımları arasında değerlendirilmektedir⁴⁶.

Yapay zekânın tıbbi görüntülemeye kullanımı hâlen yoğun şekilde araştırılmaktadır. Yüksek duyarlılık ve doğruluk göstermesi nedeniyle, doku temelli tespit ve karakterizasyon için de uygun olabileceği öngörülmektedir. Bununla birlikte, özellikle radyologlar, yapay zekâ tarafından yerlerinin alınabileceğinden endişe duymaktadır. Doğal seçim ve doğal genetik mekanizmaları taklit eden stokastik arama algoritmaları bu bağlamda geliştirilmiştir⁴⁷.

2.1.1. Makine Öğrenimi

Makine öğrenmesi, yapay zekânın temel bir alt dalı olup, bu yaklaşım, geçmiş deneyimlerden öğrenmeyi hedefleyen ve elde edilen verilerden istatistiksel çıkarımlar yaparak tahmin veya karar üreten algoritmaların geliştirilmesine dayanmaktadır⁴⁸.

Makine öğrenimi algoritmaları, giriş ve çıkış verileri arasındaki ilişkiyi kurmak üzere eğitilmiş matematiksel modeller kullanır. Eğitim sürecinde sistem, etiketlenmiş ya da etiketlenmemiş veri kümeleri üzerinden öğrenir ve bu öğrenmeye dayalı olarak yeni durumlarda doğru tahminlerde bulunabilir^{49,50}. Örneğin, radyologlar tarafından normal ve anormal olarak etiketlenen binlerce görüntü ile eğitilen bir makine öğrenimi algoritması, yeni bir görüntüdeki lezyonu normal veya anormal olarak tanımlayabilir^{51,52}.

Makine öğrenimi farklı kategorilere ayrılmaktadır. Denetimli öğrenme, her bir giriş için doğru çıktının bilindiği etiketli verilerle çalışır ve çoğunlukla sınıflandırma ile regresyon problemlerinde kullanılır. Bu yöntem, tanısal görüntülerin sınıflandırılmasında radyologların etiketleme yapmasını gerektirir. Denetimsiz öğrenme, etiketlenmemiş verilerden örüntüleri keşfetmeye odaklanır ve kümeleme, dağılım tahmini gibi

uygulamalarda öne çıkar. Yarı denetimli öğrenme, az miktarda etiketlenmiş veri ile daha büyük miktardaki etiketlenmemiş veriyi birlikte kullanır. Pekiştirmeli öğrenme ise sistemin dinamik bir ortamda pozitif veya negatif geri bildirimlerle öğrenmesini sağlar ve robotik, oyun tasarımı ile bilgisayarlı görme gibi alanlarda uygulanır^{49,50,53,54}.

Makine öğreniminde sürecin başarısı verilerin hazırlanma kalitesine bağlıdır. Veri toplama, ön işleme, eğitim ve test gruplarına ayırıştırma, uygun formatlara dönüştürme ve model eğitimi temel aşamalardır. Bu adımların ardından algoritmalar, eğitim verilerindeki örüntülere dayanarak tahminler üretir. Model performansı değerlendirilir ve istenilen başarı sağlanmadığında yeniden eğitilerek geliştirilir⁵⁵.

Makine öğrenimi, insan zihninin işlevlerini taklit edebilmesi bakımından önemli avantajlar sunmaktadır. Görüntü tanıma, sınıflama, kendi kendini düzeltme ve büyük veri kümeleri üzerinde öğrenme gibi işlevler, makine öğreniminin insan bilişsel süreçlerine benzer şekilde çalışmasını sağlamaktadır⁵⁶. Veri miktarı arttıkça algoritmalar daha doğru ve güvenilir tahminler üretir. Bu özellik, makine öğrenmesini özellikle tıp gibi büyük veri kullanılan alanlarda güçlü bir araç haline getirmektedir⁵⁷.

2.1.2. Derin Öğrenme

Derin öğrenme, yapay sinir ağlarından geliştirilmiş ve günümüzde makine öğrenmesinin en yaygın kullanılan alt dallarından biri haline gelmiştir⁵⁸. Çok sayıda gizli katman içeren bu modeller, hiyerarşik bir yapı kullanarak verilerin düşük seviyeli özelliklerini daha soyut ve karmaşık özelliklere dönüştürebilmektedir. Böylece, sığ makine öğrenmesi yöntemlerine göre daha güçlü bir özellik temsili sağlar. Geleneksel makine öğrenmesinde başarı büyük ölçüde insan uzmanların deneyimlerine dayanırken, derin öğrenme doğrudan verilere dayalı bir öğrenme süreci yürütmektedir⁵⁸.

Derin öğrenmenin kökleri 1980'lerdeki sinir ağı araştırmalarına dayanır. Derin öğrenme ile klasik makine öğrenmesinin ortak noktası veriden öğrenmeleridir; ancak derin öğrenme, çoklu doğrusal-olmayan dönüşümleri birleştirerek büyük veriden yüksek düzey soyutlamalar ve karmaşık özellikler türetebilmesiyle ayrışır^{59,60}. Bu hiyerarşik temsil yaklaşımı tıpta hastalıkların görüntülerden tanımlanması, sınıflandırılması ve prognoz öngörüsü gibi klinik görevlerde güçlü performanslar üretir⁶⁰.

Derin öğrenme modelleri, katmanlar arası akışta bir katmanın çıktısını bir sonrakinin girdisi yapan hiyerarşiler kurar. Alt katmanlar girişe yakın temel nitelikleri, üst katmanlar giderek daha karmaşık birleşimleri öğrenir; bu da verinin çok düzeyli ve anlamlı bir temsilini ortaya çıkarır^{58,61}.

Derin öğrenme, tıp alanında geniş bir uygulama yelpazesine sahip bir alandır. Derin öğrenmenin çok yönlülük, yüksek genelleme kapasitesi, disiplinler arası kullanım ve yüksek performans gibi çeşitli özel özellikleri vardır. Bu özellikler, tüm bilim topluluğunun derin öğrenmenin sunduğu olanaklara odaklanmasına neden olmuştur. Derin öğrenme kavramı, yapay sinir ağlarının basit yönleriyle başlar ve ardından dikkat modelleri, evrimsel yapılar ve yinelemeli yapılar ile genişler. Özellikle tıbbi görüntülemelerde kullanılan derin öğrenmenin önemli yönleri, tıbbi görüntü sınıflandırma ve segmentasyon süreçleridir. Günümüzde derin öğrenme, tanı alanında büyük bir etkiye sahiptir. Tıbbi görüntülemelerde anormallikler tespit edilebilir, analiz edilebilir ve bunlar farklı hastalık türlerine göre sınıflandırılabilir. Derin öğrenmenin yardımıyla, daha kısa değerlendirme süresinde doğru sonuçlara ulaşmak mümkündür. Özellikle radyoloji alanında bu teknikler kritik bir rol oynamaktadır⁶².

2.1.2.1. Yapay Sinir Ağları (ANN)

Yapay sinir ağı (ANN), en bilinen yöntemlerden biri olup, birbirine tam bağlı ve paralel çalışan temel işlem birimlerinden (perceptronlar) oluşan bir yapıdır. Her perceptron, önceki katmanlardan aldığı verileri kullanarak tek bir çıktı üretir ve bu çıktı fonksiyonu her adımda doğrusal olmayan bir dönüşüm sağlar. Perceptronlar arasındaki bağlantılar, ağırlıklar ile tanımlanır. Sinir ağları, görüntü verileri üzerinde gözetimli şekilde eğitilerek bu ağırlıkları öğrenir⁶³. Klasik bir sinir ağında genellikle tek bir gizli katman bulunurken, derin sinir ağlarında bu katman sayısı oldukça fazladır. Daha fazla hesaplama gücü gerektirmelerine rağmen, işlemci maliyetlerinin düşmesi ve grafik işlemcilerde (GPU) derin öğrenme yöntemlerinin kullanılabilir hale gelmesi, bu ağların popülaritesini artırmıştır.

2.1.2.2. Evrimsel Sinir Ağları (CNN)

Derin sinir ağlarının genel yapısına evrim (convolution) katmanlarının eklenmesiyle birlikte, evrimsel sinir ağları (CNN) geliştirilmiştir. Geleneksel sinir

ağları yalnızca tam bağlı katmanlardan oluşurken, tipik bir CNN mimarisi; evrişim ve havuzlama (pooling) katmanlarının ardından tam bağlı katmanların yer aldığı bir yapıya sahiptir. CNN'lerin görüntü analizi için yaygın şekilde tercih edilmesinin nedeni, evrişim ve havuzlama katmanlarının parametre sayısını önemli ölçüde azaltması, ağırlık paylaşımına olanak tanınması ve derin ağlarda büyük boyutlu girişlerin verimli şekilde işlenmesini sağlamasıdır⁶⁴. CNN'ler, tam bağlı katmanların çıktılarından özellik çıkarıcı olarak kullanılabileceği gibi, doğrudan sınıflandırma modeli olarak da görev yapabilir. Görüntü verilerinde yüksek düzeyde karmaşık özellikleri öğrenebilme kapasiteleri sayesinde⁶⁵, son yıllarda görüntü sınıflandırma⁶⁶⁻⁶⁸ ve nesne tespiti⁶⁹ alanlarında oldukça başarılı sonuçlar elde etmişlerdir. Güncel literatürde bilgisayarla görme uygulamalarında en sık kullanılan derin öğrenme yöntemleri büyük ölçüde CNN tabanlıdır⁷⁰.

2.1.2.3. Tam Evrişimli Ağlar (FCN)

Tam evrişimli ağlar (FCN), Long ve arkadaşları⁷¹ tarafından önerilen CNN tabanlı bir yaklaşım olup, yalnızca evrişim, havuzlama (ve gerektiğinde dekonvolüsyon ile yeniden örnekleme) katmanlarından oluşur ve sonunda tam bağlı katman içermez. FCN çıktıları çoğunlukla giriş boyutuyla aynı büyüklüğe sahiptir. Özellikle görüntü segmentation görevlerinde, bu yoğun tahmin modelleri hem doğruluk hem de verimlilik açısından önemli gelişmeler sağlamıştır.

2.1.2.4. Derin Öğrenme ile Görüntü Sınıflandırma Süreci

2.1.2.4.1. Görüntü Sınıflandırma

Görüntü sınıflandırma süreci, örneğin patoloji tipinin belirlenmesiyle başlar. Farklı boyutlarda ve çeşitli hastalık paternlerine sahip görüntüler bulunduğundan, bu tür paternlerle CNN modelini doğrudan eğitmek karmaşık hale gelmektedir. Verilere daha derinlemesine bir bakış sağlamak için, geleneksel istatistiksel analiz yapay sinir ağları ile birlikte kullanılabilir^{62,72}.

Derin öğrenmedeki en büyük zorluk, aşırı uyumlamayı (overfitting) azaltmak ve performansı artırmak için büyük miktarda veriye ihtiyaç duyulmasıdır. Az sayıda tıbbi görüntünün analiz edilmesi zordur. Küçük veri setleriyle başa çıkabilmek için en iyi tekniklerin seçilmesi gerekmektedir⁶².

2.1.2.4.2. Nesne Tanımlama

Nesnelerin bulunması ve kategorize edilmesi süreci, nesne tanımlama olarak adlandırılır. Biyomedikal görüntülerin tespiti sürecinde, hastanın lezyonlarını belirlemek için kutu koordinatları kullanılır. Bu teknik genel olarak iki kategoriye ayrılmaktadır: Bölge önerisine dayalı algoritma ve regresyon yöntemi olarak tek-aşamalı ağ.

Bölge önerisine dayalı algoritmada, giriş görüntülerinden seçici arama algoritması yardımıyla çeşitli yamalar çıkarılır. Bu işlemin tamamlanmasının ardından, eğitilmiş model, her bir yama içerisinde birden fazla görüntünün varlığına dair karar verir ve sınıflandırma ilgi bölgesine göre yapılır. Bu tekniğin avantajı, tespit hızını artırmasıdır⁶².

Regresyon yöntemi–tek aşamalı ağ tekniğinde, kutu koordinatları ve sınıf olasılıkları doğrudan tüm görüntü piksellerinden elde edilir ve tespit edilir. Bölge önerisi yaklaşımı doğruluk açısından daha iyi kabul edilirken, regresyon yöntemi yani tek aşamalı ağ hem hız hem de doğruluk açısından verimli bulunmuştur.

2.1.2.4.3. Görüntü segmentasyonu ve kaydı

Segmentasyon ve kayıt süreçlerinde derin öğrenme tekniklerinin çok daha verimli olduğu kanıtlanmıştır. Hedeflenen segmentasyonun başarımını artırmak için bağımlı eşikleme ve kapalı-kontur yöntemleri kullanılmaktadır. Aynı zamanda kayıt işlemleri de segmentasyona uygulanmaktadır. Tümörlerin ve farklı bölgelerdeki diğer yapıların segmentasyonu için çok sayıda girişimde bulunulmuştur. Segmentasyonu geliştirmek amacıyla derin öğrenme ağları sürekli olarak geliştirilmektedir⁶².

2.1.2.4.4. Görüntü üretimi

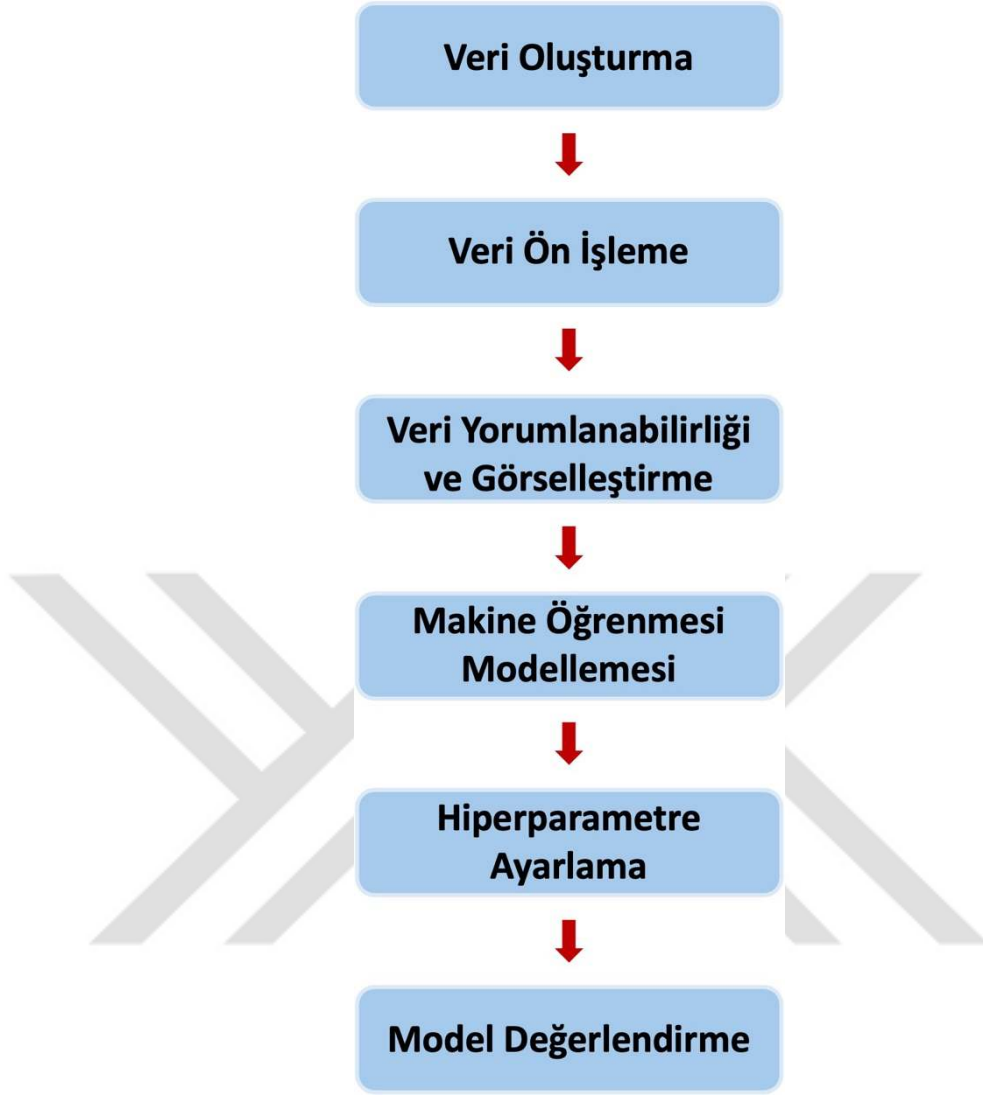
Çeşitli etiketlere sahip yüksek kaliteli ve dengeli veri setleri elde etmek zordur. Araştırmalarda üretilen görüntüler, veri artırma sürecinde kullanılmaktadır. Tıbbi görüntülemelerde anomali tespiti, üretici çekişmeli ağların en dikkat çekici uygulamasıdır. Sentetik görüntülerde sınıflandırma ve segmentasyon doğruluklarının yüksekliği büyük önem taşır⁶².

2.1.2.4.5. Görüntü dönüşümü

Son çalışmalar, CNN'in en sık kullanılan yöntem olduğunu göstermektedir. Derin öğrenme geliştirme çerçevesi (framework), hızla büyüyen bir yazılım geliştirme ortamıdır. Derin öğrenme, insan müdahalesi gerektirmeyen her türlü veri seti üzerinde çalışabilen bir teknolojidir. Özellikle CNN yapılarının evrimiyle birlikte 2012 yılından itibaren derin öğrenme alanında çok büyük değişimler yaşanmıştır⁷³. Bu gelişmelerin ardından, derin öğrenme çeşitli görüntü problemlerinin çözümünde yaygın olarak kullanılmaya başlanmıştır. Bu ilerlemeler sayesinde birçok hekim tıbbi görüntüleme işlemlerini kullanmakta, tanı ve sınıflandırma işlemleri de çeşitli makine öğrenmesi teknikleriyle desteklenmektedir. Günümüzde derin öğrenme teknolojileri ağırlıklı olarak tıbbi görüntü sınıflandırma ve segmentasyon amaçlarıyla kullanılmaktadır. Tüm bunlara dayanarak, derin öğrenme, görüntü analizi sürecinde en önemli araçlardan biri haline gelmiştir. Bu teknoloji; tespit, segmentasyon, kayıt ve sınıflandırma alanlarında başarılı bir şekilde uygulanmaktadır. Derin öğrenmenin çok sayıda farklı yaklaşımı vardır. Tıp alanındaki çeşitli özellikler nedeniyle büyük data'lar için geniş bir potansiyel söz konusudur ve bu da derin öğrenmeyi bu uygulamalar için daha uygun hale getirmektedir. Derin öğrenme; oftalmoloji, nörogörüntüleme ve ultrasonografi gibi pek çok tıbbi alanda kullanılmaktadır⁶². Küçük veri setlerinin yetersizliğini aşmak ve büyük veri setlerini hızla büyütmek için bariz bir çözüm, verilerin tıbbi merkezler arasında paylaşılmasıdır. Ancak mevcut veri gizliliği politikaları nedeniyle bu durumun uygulanması zor görünmektedir⁷⁴.

2.1.3. Hiperparametre Optimizasyonu

Hiperparametre optimizasyonu, makine öğrenmesi modellerinin başarımını yükseltmek amacıyla modeldeki hiperparametrelerin en uygun değerlerinin belirlenmesi sürecidir. Bu süreç, modelin öğrenme sürecini yönlendiren ve yapısını şekillendiren parametrelerin en iyi kombinasyonlarının bulunmasını kapsar. Hiperparametreler, model eğitimi sırasında otomatik olarak öğrenilmeyen; bunun yerine eğitimden önce elle ya da çeşitli optimizasyon yöntemleriyle belirlenmesi gereken değişkenlerdir. Derin öğrenme iş akışında hiperparametre optimizasyonunun genel yapısı Şekil 2.1'de gösterilmiştir. Makine öğrenmesi modellerinde performansı artıran hiperparametrelerin belirlenmesi genellikle modelin eğitilme süresinden daha uzun süren bir optimizasyon aşamasını gerektirir⁷⁵.



Şekil 2. 1. Derin öğrenme sürecinde hiperparametre optimizasyonu⁷⁶

Optimizasyon süreci, belirli bir arama uzayı içerisinde farklı hiperparametre kombinasyonlarının sistematik olarak test edilmesi ve her bir kombinasyonun performansının doğruluk, F1 skoru veya AUC gibi değerlendirme metrikleri aracılığıyla ölçülmesiyle gerçekleştirilir. Derin öğrenme modellerinde, genellikle optimize edilen başlıca hiperparametreler arasında öğrenme oranı, parti boyutu (batch size), epoch sayısı ve kullanılan optimizasyon algoritması gibi değişkenler bulunmaktadır⁷⁷.

Hiperparametre uzayının kapsamlı biçimde incelenmesi, her bir hiperparametre kombinasyonu için modelin yeniden eğitilmesini gerektirdiğinden oldukça yüksek hesaplama maliyetine sahip bir süreçtir. Bu optimizasyon işlemi, hem zaman hem de donanım kaynakları açısından yoğun bir yük oluşturur. Bu nedenle, mevcut kaynakların

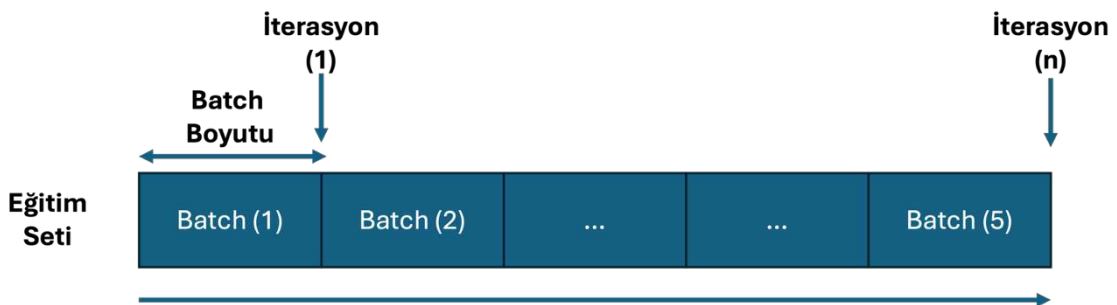
daha verimli kullanılabilmesi amacıyla erken durdurma, aşamalı eleme ve paralel hesaplama gibi stratejiler sıklıkla tercih edilmektedir⁷⁸.

2.1.3.1. Epoch Sayısı

Epoch sayısı, modelin tüm eğitim veri setini kaç kez baştan sona işleyeceğini ifade eden temel bir hiperparametredir. Modelin öğrenme sürecinde epoch kavramı Şekil 2.2’de örnek olarak görselleştirilmiştir. Eğitim sırasında veri seti küçük partilere (batch) bölünür ve her parti model tarafından işlenerek ağırlıklar geriye yayılım yöntemiyle güncellenir. Bu işlem, her eğitim adımında yinelenir ve modelin ağırlıkları kademeli olarak optimize edilir. Derin öğrenme modellerinde, ilk epoch’larda modelin başarımı genellikle düşüktür; çünkü ağ, veri örüntülerini yeni öğrenmeye başlamıştır. Epoch sayısı arttıkça modelin performansı genellikle yükselir, ancak belirli bir noktadan sonra öğrenme hızı azalır ve modelin aşırı uyum riski artabilir. Ayrıca epoch sayısının fazla olması, eğitim süresinin uzamasına yol açar⁷⁹.

2.1.3.2. Batch Boyutu

Batch boyutu, her bir eğitim adımında modele sunulan veri örneği sayısını belirleyen önemli bir hiperparametredir. Modelin öğrenme sürecinde parti boyutunun her adımda nasıl işlendiği Şekil 2.2’de örnek olarak görselleştirilmiştir. Batch boyutu, modelin hem eğitim hızını hem de genelleme performansını doğrudan etkiler. Büyük batch boyutları, daha hızlı öğrenmeye olanak tanır ve gradyan tahminlerindeki gürültüyü azaltarak daha kararlı bir optimizasyon sağlar. Ancak, yüksek bellek kullanımına neden olur ve modelin genelleme kabiliyetini olumsuz yönde etkileyebilir. Küçük batch boyutları ise daha az bellek gerektirir ve genellikle daha iyi genelleme sonuçları sunabilir, ancak öğrenme sürecini yavaşlatır ve hesaplama maliyetini artırabilir⁸⁰.



Şekil 2. 2. Model eğitimi sırasında epoch ve batch için örnek bir şematizasyon

2.1.3.3. Öğrenme Oranı

Öğrenme oranı, modelin ağırlıklarının her bir eğitim adımında ne kadar hızlı güncelleneceğini belirleyen temel bir hiperparametredir. Bu değerin çok yüksek olması, modelin eğitim verilerine aşırı uyum göstermesine ve yeni veriler üzerinde düşük genelleme performansı sergilemesine neden olabilir. Buna karşılık, çok düşük bir öğrenme oranı, modelin veri setindeki örüntüleri yeterince öğrenememesine ve dolayısıyla zayıf bir performans ortaya koymasına yol açar. Aşırı düşük bir öğrenme oranı, modelin yerel minimumlarda sıkışıp kalmasına neden olurken; aşırı yüksek bir öğrenme oranı, modelin küresel minimuma ulaşmasını engelleyebilir. En uygun öğrenme oranı, başlangıçta nispeten yüksek bir değerle modelin küresel minimuma hızlıca yaklaşmasını, ardından giderek azalan bir oranla bu minimum çevresinde daha kararlı bir yakınsama sağlamayı amaçlar⁸¹⁻⁸³.

2.1.3.4. Optimizasyon fonksiyonu

Optimizasyon fonksiyonu, modelin öğrenme sürecini en verimli şekilde yönlendirmek ve hata fonksiyonunu minimize etmek için kullanılan algoritmadır. Doğrusal olmayan problemlerde, modelin küresel minimuma ulaşmasını hedefleyen optimizasyon yöntemleri tercih edilir. Derin öğrenme uygulamaları, çok boyutlu ve karmaşık fonksiyonların optimizasyonunu gerektirdiğinden, bu alanda en yaygın kullanılan algoritmalar genellikle stochastic gradient descent (SGD) ve onun türevleridir. SGD'nin farklı varyantları olan Adagrad, Adadelat, adaptive moment estimation (Adam) ve Adamax, öğrenme oranını dinamik olarak ayarlayarak her parametre için farklı hızlarda güncelleme yapılmasına olanak tanır. Adagrad, geçmiş gradyanların karelerinin kümülatif toplamını dikkate alarak öğrenme oranını zamanla azaltır. Adadelat, bu yaklaşımı geliştirerek yalnızca belirli bir geçmiş aralığındaki gradyanların ortalamasını hesaba katar. Adam ve Adamax, momentum kavramını kullanarak daha hızlı yakınsama sağlar; Adam ikinci moment istatistiğini hesaba katarken, Adamax sonsuz normu temel alır. Bu algoritmaların yakınsama hızı ve genel performansı, veri setinin yapısı, problem türü ve seçilen hiperparametre değerlerine göre değişiklik gösterebilir^{84,85}.

Adam, derin öğrenmede en sıklıkla uygulanan optimizasyon algoritmalarındandır. Adam algoritmasının bazı önemli özellikleri vardır. İlk olarak, momentum doğrudan gradyanın birinci dereceden momentinin bir tahmini olarak dâhil edilmiştir. Ayrıca Adam, hem birinci hem de ikinci moment tahminlerinin başlangıçta sıfırdan başlatılmalarını dengelemek için önyargı düzeltmelerini içerir. Algoritma, gradyanın üstel hareketli ortalamalarını (m_t) ve gradyanın karesinin üstel hareketli ortalamalarını (v_t) günceller. Burada hiperparametreler ρ_1 ve ρ_2 , bu hareketli ortalamaların üstel azalma hızlarını kontrol eder. Bu hareketli ortalamalar, sırasıyla gradyanın birinci momentinin (ortalama) ve ikinci ham momentinin (merkezlenmemiş varyans) tahminleridir. Adam algoritması, birinci ve ikinci moment değişkenleri olan m ve v 'yu gerektirir. Gradyan hesaplandıktan sonra, zaman adımı t 'de sırasıyla önyargılı birinci ve ikinci moment tahminleri güncellenir. Ardından, birinci ve ikinci momentlerdeki önyargı düzeltilir. Düzeltilmiş moment tahminleri kullanılarak parametre güncellemeleri hesaplanır ve uygulanır⁸⁶⁻⁸⁸.

2.1.3.5. Aktivasyon fonksiyonu

Aktivasyon fonksiyonları, sinir ağlarının doğrusal olmayan ilişkileri öğrenmesini sağlayan temel bileşenlerdir. Bu fonksiyonlar, giriş verisini daha ayrılabilir hâle getirerek ağın soyut özellikleri öğrenmesine olanak tanır. İyi bir aktivasyon fonksiyonu; gradyan akışını kesmeden, hesaplama yükünü artırmadan ve veri dağılımını koruyarak eğitim sürecini iyileştirmelidir. Farklı türde AF'ler (ör. Sigmoid/Tanh, ReLU, ELU, adaptif ve karma fonksiyonlar) belirli avantaj ve sınırlamalara sahiptir. Bu nedenle ağın mimarisi ve veri tipi, hangi aktivasyon fonksiyonunun seçileceğini belirleyen en önemli faktörlerdir. Aktivasyon fonksiyonları, yapay sinir ağlarında giriş değerini belirli bir matematiksel dönüşümden geçirerek hangi nöronların aktif olacağını belirler. Logistic Sigmoid / Tanh tabanlı fonksiyonlar, ReLU tabanlı fonksiyonlar, üstel tabanlı fonksiyonlar, öğrenilebilir / adaptif fonksiyonlar ve diğer fonksiyonlar olmak üzere başlıklara ayrılabilir. Bu başlıkların altında çok fazla sayıda fonksiyon bulunmakta olup, derin öğrenme modellerinde sıklıkla kullanılan fonksiyonlar şu şekildedir^{85,89-93}:

2.1.3.5.1. Logistic Sigmoid / Tanh Tabanlı Fonksiyonlar

a) Sigmoid Fonksiyonu: Bu fonksiyon, gerçekte sayı değerlerini girdi olarak alır ve çıktıyı 0 ile 1 aralığında sınırlar. Doğrusal olmayanlık kazandırmak için kullanılan klasik bir aktivasyon fonksiyonudur. Ancak büyük pozitif ve negatif girişlerde gradyan sönmesi problemi görülür ve sıfır merkezli olmaması nedeniyle ağırların yakınsaması zayıflar.

Matematiksel formülü:

$$f(x)_{\text{sigm}} = \frac{1}{(1 + e^{-x})}$$

b) Tanh (Hyperbolic Tangent) Fonksiyonu: Benzer şekilde çalışır fakat sıfır merkezli olup çıktıyı -1 ile +1 aralığına sıkıştırır. Tanh, Sigmoid'e göre daha iyi yakınsama sağlasa da büyük girişlerde gradyan sönmesi ve hesaplama maliyeti sorunlarını paylaşır.

Matematiksel formülü:

$$f(x)_{\text{tanh}} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

2.1.3.5.2. Rectified Linear Unit (ReLU) Tabanlı Fonksiyonlar

a) ReLU: Derin öğrenmede en yaygın kullanılan aktivasyon fonksiyonudur. Pozitif girişlerde doğrusal, negatiflerde sıfır değer üretir. Eğitim sürecini hızlandırır ve gradyan kaybolma problemini azaltır. Basit yapısı ve düşük hesaplama maliyetiyle eğitimi hızlandırır; ancak negatif girdilerde gradyan sıfırlandığı için “ölü nöron” problemi ortaya çıkabilir.

Matematiksel formülü:

$$f(x)_{\text{ReLU}} = \text{üst sınır}(0, x)$$

b) Leaky ReLU: Klasik ReLU'nun negatif değerleri tamamen yok sayma sorununu gidermek için geliştirilmiştir. ReLU fonksiyonundaki ölü nöron problemini azaltmak için negatif girişlerde küçük bir eğim (sabit bir katsayı) bırakır. Yani negatif

değerleri sıfırlamak yerine, bu değerlere çok küçük bir çarpan (örneğin 0.01) uygular. Böylece gradyan tamamen kaybolmaz ve nöronlar “ölmez”. ReLU’ya benzer şekilde hesaplama açısından verimlidir.

Matematiksel formülü:

$$f(x)_{LeakyReLU} = \begin{cases} x, & \text{eğer } x > 0 \\ \alpha x, & \text{eğer } x \leq 0 \end{cases}$$

2.1.3.5.3. Üstel Tabanlı Fonksiyonlar

a) Exponential Linear Unit (ELU): ReLU’nun negatif değerleri kullanmaması sorununu gidermek için geliştirilmiştir. Pozitiflerde doğrusal, negatiflerde üstel biçimdedir. Negatif bölgede küçük bir eğim bırakarak gradyan sönmesini azaltır ve gürültüye karşı dayanıklılığı artırır.

Matematiksel formülü:

$$f(x)_{ELU} = \begin{cases} x, & \text{eğer } x > 0 \\ \alpha(e^x - 1), & \text{eğer } x \leq 0 \end{cases}$$

b) (Scaled Exponential Linear Unit (SELU): ELU’nun ölçeklendirilmiş hâlidir ve self-normalization özelliği kazandırır. Ağın çıktısını otomatik olarak sıfır ortalama ve birim varyansa yakın tutar, böylece eğitim stabilitesini artırır.

Matematiksel formülü:

$$f(x)_{SELU} = \lambda \begin{cases} x, & \text{eğer } x > 0 \\ \alpha(e^x - 1), & \text{eğer } x \leq 0 \end{cases}$$

2.1.3.5.4. Öğrenilebilir / Adaptif Fonksiyonlar

a) Swish: Google tarafından önerilen, ReLU ile Sigmoid’in birleşimi olan modern bir aktivasyon fonksiyonudur. Öğrenilebilir β parametresi sayesinde fonksiyon şekli veri yapısına uyum sağlar. Yumuşak geçişli ve non-monotonic yapısı, özellikle derin ağlarda ReLU’ya göre daha iyi performans sağlayabilir.

Matematiksel formülü:

$$f(x)_{Swish} = x \cdot \sigma(x) = \frac{x}{1 + e^{-x}}$$

b) Exponential Swish (E-Swish): E- Swish, Swish'in geliştirilmiş versiyonudur ve pozitif yöndeki eğimi kontrol etmek için ek bir katsayı içerir. Bu form, ağırlık pozitif bölgede daha büyük gradyanlar üretmesini sağlar ve eğitim hızını artırabilir. ReLU'ya göre daha yüksek doğruluk, ancak biraz daha fazla hesaplama maliyeti sunar.

Matematiksel formülü:

$$f(x)_{E-Swish} = \beta x \cdot \sigma(x) = \beta x \cdot \frac{1}{1 + e^{-x}}$$

2.1.3.5.5. Diğer Fonksiyonlar

a) Softplus: ReLU'nun yumuşak bir versiyonudur ve diferansiyellenebilir yapısı sayesinde gradyan akışını iyileştirir. Çıktısı her zaman pozitiftir ve küçük girişlerde ReLU'ya benzer, büyük girişlerde ise doygunlaşır. Pürüzsüz yapısı sayesinde istikrarlı optimizasyon sağlar ancak hesaplama maliyeti ReLU'dan yüksektir.

Matematiksel formülü:

$$f(x)_{Softplus} = \ln(1 + e^x)$$

b) Mish: Son yıllarda popülerleşen, Swish ve Softplus fonksiyonlarını birleştiren bir aktivasyon fonksiyonudur. Sürekli türevlenebilir ve non-monotonic bir yapıya sahiptir. YOLOv4 gibi ağlarda kullanılmıştır.

Matematiksel formülü:

$$f(x)_{Mish} = x \cdot \tanh(\ln(1 + e^x))$$

2.1.4. Yapay Sinir Ağları ile Yaklaşımda Sınıflandırma

Tıbbi görüntülemelerde sınıflandırmanın iki yöntemi vardır: biri geleneksel modellere, diğeri ise derin öğrenme modellerine dayanmaktadır. Görüntü sınıflandırması için kullanılan tüm derin öğrenme yöntemleri içinde CNN en yaygın kullanılan yapılardan biridir. Krizhevsky ve arkadaşları tarafından 2012 yılında önerilen AlexNet

tabanlı CNN modeli, görüntü sınıflandırmada dikkat çekici sonuçlar elde etmiştir⁷³. Bu tarihten itibaren görüntü sınıflandırma uygulamalarında derin öğrenmenin yükselişi başlamıştır. 2014 yılında GoogLeNet ve VGGNet modelleri, ImageNet veri setinin genel doğruluğunu artırmıştır. Sonraki yıllarda GoogLeNet'in birkaç yeni versiyonu daha geliştirilmiş ve modelin performansı daha da artırılmıştır⁹⁴.

AlexNet algoritması, sekiz katmanlı bir ağ yapısı kullanmaktadır: beş evrişim katmanı ve üç tam bağlı katman. Evrişim katmanlarında evrişim işlemi uygulanarak pooling en üst düzeye çıkarılır ve veri boyutu azaltılır. Beş katmandaki evrişim tamamlandıktan sonra, çıktı özellik matrisi biçiminde tam bağlı katmanlara gönderilir. İlk tam bağlı katmanda 4.096 çekirdek ile dropout işlemi uygulanır. Son iki katman ise kayan noktalı veri çıktısı üreterek tahmin sonuçlarını verir⁶².

Sınıflandırma adımları, bir şeyin belirli bir sınıfa ait olup olmadığının ayrıştırılması veya tahmin edilmesini içerir. Sınıflandırma türleri şunlardır:

-İkili sınıflandırma (Binary classification): Bir örneğin iki kategoriden hangisine ait olduğunu tahmin etmek.

-Çok sınıflı sınıflandırma (Multiclass classification): İki den fazla kategoriye sahip sınıflandırma.

-Çok etiketli sınıflandırma (Multilabel classification): Tek bir örneğin birden fazla etikete sahip olabilmesi durumu.

Eğitim verilerinde bulunan benzer paternler, veri korelasyonu, analiz ve kümeleme algoritmaları yardımıyla incelenmektedir^{62,95}.

Sınıflandırmada Yer Alan Adımlar:

1. Gerekli kütüphaneleri içe aktarma,
2. Veri setini oluşturma,
3. Seçilen veri setine göre etiketleri tanımlama (örnek: ikili sınıflandırma (1 veya 0),

4. Verilerin görselleştirilmesi: böylece hangi tür tahmin modelinin inşa edileceği anlaşılır,

5. Sinir ağı inşa edilirken, giriş (input) ve çıkış (output) özellik boyutlarının aynı olmasına dikkat edilmesi,

6. Sinir ağları, model oluşturma, derleme ve uyarlama adımlarıyla modellenir.

Adımlar tamamlandıktan sonra, model doğruluğu kontrol edilmelidir. Eğer performans düşükse, veriler yeniden görselleştirilmeli ve katman sayısı veya optimizier fonksiyonu deęiştirilerek en iyi performans elde edilene kadar süreç tekrarlanmalıdır⁶².

2.1.5. U-Net

CNN, derin öğrenmenin en yaygın kullanılan algoritmalarından biridir^{96,97}. Özellikle görüntü işleme alanında önemli avantajlar sunar; hedef nesne tespiti ve görüntü segmentasyonu gibi görevlerde sıkça tercih edilir. CNN, çok büyük veri kümelerinde görüntü tanıma problemlerini aşamalı olarak basitleştirerek sonrasında öğrenilebilir bir hale getirir. Aslında uzun süredir bilinen bir yöntem olmasına rağmen, yeterince büyük eğitim veri setlerinin bulunmaması ve ağların karmaşıklığı nedeniyle başarı düzeyi başlangıçta sınırlı kalmıştır.

2015 yılında Olaf Ronneberger ve ark.⁸, tıbbi görüntü analizi için uçtan uca eğitim yaklaşımını benimseyen U-Net mimarisini geliştirmiştir. CNN tabanlı en önemli anlamsal segmentasyon modellerinden biri olarak kabul edilen U-Net, günümüzde tıbbi görüntülemelerde kritik bir rol oynamaktadır^{98,99}.

U-Net yapısı temel olarak iki bölümden oluşur: aşağı örnekleme ve yukarı örnekleme. Bu mimarinin ana fikri, sabit boyutlu görüntülerin çözünürlüğünü azaltarak daha küçük önizleme görüntüleri oluşturmak ve buradan daha derin seviyelerdeki özellikleri çıkarmaktır. Ardından, yukarı örnekleme aşamasıyla görüntü yeniden büyütülür. Aşağı örnekleme sırasında elde edilen özellikler, yukarı örnekleme katmanlarına kopyalama ve kırpma yöntemleriyle aktarılır. Son aşamada konvolüsyon katmanı, bu bilgileri kullanarak daha doğru tahminler üretmeyi öğrenir. U-Net'in öne çıkan avantajlarından biri, yukarı örnekleme bölümünde çok sayıda özellik kanalına sahip olmasıdır. Bu özellik, bağlamsal bilgilerin yüksek çözünürlüklü katmanlara taşınmasına

imkân verir. U-net, biyomedikal segmentasyon için oldukça esnek bir yapı sunar, farklı boyutlardaki işlenmesine görüntülerin farklı görüntüleme yöntemindeki uygulamalarına olanak tanır¹⁰⁰.

2.1.5.1. nnU-Net

U-Net, biyomedikal görüntülerin doğru ve verimli biçimde segmentasyonu için geliştirilmiş özel bir fCN mimarisidir⁸. nnU-Net ise, U-Net'in ileri düzeyde geliştirilmiş bir uyarlamasıdır ve çeşitli veri çoğaltma (data augmentation) yöntemleri, farklı kayıp fonksiyonları ile çoklu görüntü ölçeklendirme olanakları gibi yenilikler sunmaktadır. nnU-Net'in en belirgin özelliği, farklı veri türleri ve boyutlarına kolayca adapte olabilmesi, dolayısıyla her segmentasyon görevi için özel modellerin oluşturulabilmesidir. Bu yapı, veri kümesine özgü hiperparametreleri tanımlamak amacıyla “veri parmak izi” yaklaşımını kullanır. Planlama parametreleri (kayıp fonksiyonu, optimizasyon algoritmaları, ağ mimarisi), çıkarım parametreleri (yeniden örnekleme, normalizasyon, batch size ve patch size) ile veri özellikleri bir araya getirilerek bu parmak izleri oluşturulur. Elde edilen bilgiler ışığında 2B, 3B ve kademeli 3B U-Net modelleri için en uygun konfigürasyon belirlenir ve bu en iyi ayarlar test veri setine uygulanır¹⁰¹.

2.2. Modellerin Değerlendirilmesinde Kullanılan Performans Metrikleri

Modellerin veri setleri üzerindeki başarımını ölçmek ve genel performanslarını değerlendirmek amacıyla çeşitli performans metrikleri kullanılmaktadır. Bu metrikler, modelin eğitim sürecinde ortaya çıkabilecek sorunların belirlenmesine ve en uygun modelin seçilmesine yardımcı olur. Her bir metriğin model performansının farklı bir yönünü yansıttığı göz önünde bulundurulmalı ve değerlendirme sürecinde yalnızca tek bir metriğe bağlı kalınmamalıdır. En uygun performans metriği, çözülmek istenen problemin türüne ve kullanılan veri kümesinin yapısına göre değişiklik gösterebilir. Elde edilen sonuçlar, istatistiksel olarak anlamlı bir biçimde raporlanarak model seçim sürecine bilimsel dayanak oluşturur.

2.2.1. Karışıklık Matrisi

Bir modelin veya algoritmanın performansını ölçmek için kullanılan nicel ölçütlerdir. Bu ölçütler, bir modelin belirli bir görevde ne kadar iyi performans gösterdiğini belirlemeye yardımcı olur. Önemleri, farklı algoritmalar arasında karşılaştırma yapılmasını kolaylaştırmalarından ve bir modelin pratik bağlamda doğruluğunu sağlamada hayati rol oynamalarından kaynaklanır. Bu metrikler, modelin tahmin edilen sonuçlarını ground truth (gerçek sonuçlarla) karşılaştırarak performansı değerlendirir. Ground truth, tıbbi uzmanlar tarafından manuel olarak elde edilen gerçek akciğer segmentasyon sonuçlarını ifade eder. Aşağıda tıbbi görüntülerde segmentasyonlar için yaygın olarak kullanılan bazı metriklerin özeti verilmiştir:

Karışıklık matrisi (confusion matrix), ikili sınıflandırıcının (binary classifier) performansını görselleştirmek için kullanılan temel bir araçtır. Tahmin edilen ve gerçek etiketlerin her kombinasyonunun kaç kez gerçekleştiğini kaydeder. İkili sınıflandırmada (olasılıkla gerçek etiketler {Pozitif, Negatif}), karışıklık matrisi Çizelge 2.1'de gösterildiği gibi 2×2 'lik bir tablo şeklindedir¹⁰².

Çizelge 2. 1. İkili Sınıflandırma Görevi için Karışıklık Matrisi

		Tahmin Edilen	
		Pozitif	Negatif
Gerçek Sınıf	Pozitif	Gerçek Pozitif	Yanlış Negatif
	Negatif	Yanlış Pozitif	Gerçek Negatif

2.2.1.1. Gerçek Pozitif

Model, pozitif bir örneği doğru şekilde pozitif olarak sınıflandırır. NPK segmentasyonu bağlamında, modelin görüntü piksellerini NPK pikseli olarak doğru şekilde sınıflandırmasıdır. Sınıflandırma fonksiyonu $f(x)$, gerçek etiket y olmak üzere formülü:

$$\text{Gerçek pozitif} = \text{Frekans} [f(x) = 1 \wedge y = 1]$$

2.2.1.2. Gerçek Negatif

Model, negatif bir örneği doğru şekilde negatif olarak sınıflandırır. NPK segmentasyonu bağlamında, sistemin NPK bölgesine ait olmayan pikselleri doğru şekilde NPK dışı piksel olarak sınıflandırmasıdır. Sınıflandırma fonksiyonu $f(x)$, gerçek etiket y olmak üzere formülü:

$$\text{Gerçek negatif} = \text{Frekans} [f(x) = 0 \wedge y = 0]$$

2.2.1.3. Yanlış Pozitif

Tip I hata olarak da bilinir. Model, negatif bir örneği hatalı olarak pozitif sınıflandırır (“yanlış alarm”). NPK segmentasyonu bağlamında, modelin NPK bölgesine ait olmayan pikselleri yanlışlıkla NPK pikseli olarak sınıflandırmasıdır. Sınıflandırma fonksiyonu $f(x)$, gerçek etiket y olmak üzere formülü:

$$\text{Yanlış pozitif} = \text{Frekans} [f(x) = 1 \wedge y = 0]$$

2.2.1.3. Yanlış Negatif

Tip II hata olarak da bilinir. Model, pozitif bir örneği hatalı olarak negatif sınıflandırır (“kaçırma”). Sınıflandırma fonksiyonu $f(x)$, gerçek etiket y olmak üzere formülü:

$$\text{Yanlış negatif} = \text{Frekans} [f(x) = 0 \wedge y = 1]$$

Birçok performans metriği doğrudan gerçek pozitif, gerçek negatif, yanlış pozitif, yanlış negatif değerlerinden türetilir. Karışıklık matrisini incelemek, sınıflandırıcının en sık hangi durumlarda hata yaptığını ortaya çıkararak modele yönelik hedeflenmiş iyileştirmeler yapılmasını sağlar.

2.2.2. Karışıklık Matrisinden Türetilen Diğer Metrikler

Karışıklık matrisinden türetilen diğer metrikler¹⁰²:

2.2.2.1. Doğruluk

Doğruluk (accuracy), genel piksel doğruluğu olarak da adlandırılır ve görüntüdeki piksellerin yüzde kaçının doğru sınıflandırıldığını gösterir. Temel bir performans ölçütü

olmasına rağmen, sınıf dengesizliği olduğunda segmentasyon performansını temsil etmede kısıtlıdır¹⁰³. Segmentasyon sınıfları arasında dengesizlik olduğunda, bir sınıf diğerlerini baskılama eğilimindedir. Eğer baskın sınıfın doğruluğu diğer sınıflara göre daha yüksekse, sonuçlarda yanlılık (bias) oluşabilir. Bu nedenle, görüntüde sınıf dengesizliği olmadığında doğruluk, segmentasyon performansının ölçülmesinde kullanılmalıdır. Doğruluğun alternatif bir tanımı ise sınıf başına doğruluk'tur. Bu ölçüt, her sınıftaki doğru sınıflandırılmış piksellerin oranının ortalamasını alır. Özellikle sınıf dengesizliği bulunan görüntüler için faydalıdır. Sınıf dengesizliği durumunda hem ortalama sınıf doğruluğu hem de sınıf başına doğruluk dikkate alınmalıdır. Ancak ortalama sınıf doğruluğunun, tek tek sınıfların güvenilirlik ölçümünü azaltma gibi bir sınırlılığı vardır. Ayrıca, belirli bir sınıfa ait örneklerin az olduğu durumlarda yüksek varyans görülebilir ve bu da sonuçların güvenilirliğini etkileyebilir^{62,102,104-107}. Gerçek pozitif GP, yanlış pozitif YP, gerçek negatif GN ve yanlış negatif YN olmak üzere formülü:

$$\text{Doğruluk} = \frac{GP + GN}{GP + GN + YP + YN}$$

2.2.2.2. Kesinlik (Precision)

Kesinlik, aynı zamanda doğru pozitiflik değerlendirmesi olarak da bilinir ve doğru pozitiflerin, doğru pozitifler ile yanlış pozitiflerin toplamına oranı ile hesaplanır.

Kesinlik, tahminlerin başarısını değerlendirmek için değerli bir ölçüttür. Pozitif tahminlerin ne kadarının güvenilir olduğunu doğrudan belirtir. Daha yüksek kesinlik seviyeleri, modelin daha doğru ve kesin sonuçlar ürettiğini gösterir. Yüksek kesinlik değeri, verilen verilerle iyi eğitilmiş bir mimarının göstergesidir^{62,102,104-107}. Gerçek pozitif GP ve yanlış pozitif YP olmak üzere formülü:

$$\text{Kesinlik} = \frac{GP}{GP + YP}$$

2.2.2.3. Duyarlılık (Recall / Sensitivity)

Duyarlılık, doğru pozitiflerin, doğru pozitifler ile yanlış negatiflerin toplamına oranı olarak tanımlanır. Özünde, duyarlılık, doğru pozitiflerin ne kadarının başarıyla

tespit edildiğini gösterir. Kesinlik ile benzer şekilde, duyarlılık da tahmin başarısını değerlendirmede değerli bir ölçüttür. Özellikle sınıf dengesizliğinin bulunduğu durumlarda önemli hale gelir. Daha yüksek duyarlılık değerleri, modelin pozitif sonuçların çoğunu doğru şekilde yakaladığını ve etiketlediğini ifade eder^{62,102,104–107}. Gerçek pozitif GP ve yanlış negatif YN olmak üzere formülü:

$$\text{Duyarlılık} = \frac{GP}{GP + YN}$$

2.2.2.4. F1 Skoru

Bu ölçüt, bir modelin doğruluğunu değerlendirmek için gereklidir. Kesinlik ve duyarlılığı, harmonik ortalama kullanarak tek bir ölçü altında birleştirir. Geleneksel ortalama yerine harmonik ortalama kullanılması, aşırı uç değerlerin etkisini sınırlar.

Bu ölçüm, duyarlılık ve kesinlik arasında denge kurmak açısından faydalıdır. F1 skoru 0 ile 1 arasında değer alır; 1 en iyi performansı, 0 ise en kötü performansı ifade eder. F1 skoru 1'e yakın olan modeller ideal kabul edilir çünkü yanlış pozitif (false positive) ve yanlış negatif (false negative) sayıları düşüktür.

Bu ölçüt, hem ikili (binary) hem de çok sınıflı (multi-class) sınıflandırma ve segmentasyon problemleri için uygulanabilir^{62,102,104–107}. Formülü:

$$F1 = \frac{2 \times \text{Kesinlik} \times \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}}$$

2.2.3. Diğer Değerlendirme Metrikleri

2.2.3.1. Dice Katsayısı

Dice katsayısı, gerçek (ground truth) bölge ile segmentasyon sonucu elde edilen bölgenin ne kadar örtüştüğünü değerlendirir. Hesaplama, segmentasyon ile ground truth bölgelerinin kesişiminin iki katının, bu bölgelerin boyutları toplamına oranlanmasıyla yapılır. Daha yüksek bir Dice katsayısı değeri, daha iyi segmentasyon doğruluğunu gösterir.

Genel piksel doğruluğu ile karşılaştırıldığında, Dice katsayısı daha tercih edilebilir bir ölçüttür çünkü her sınıftaki yanlış alarmları (false positive) ve kaçırılan

değerleri (false negative) dikkate alır. Dice katsayısının avantajı sadece piksellerin doğru etiketlenip etiketlenmediğini kontrol etmesi değil, aynı zamanda segmentasyon sınırlarının doğru çizilip çizilmediğini de denetlemesidir. Dice katsayısı, 0 ile 1 arasında bir değer alır. 1'e yakın bir değer, tahmin edilen ve gerçek bölgelerin çok yüksek oranda örtüştüğünü, yani modelin başarılı bir segmentasyon yaptığını gösterir. 0'a yakın bir değer ise, bölgeler arasında çok az veya hiç örtüşme olmadığını ve modelin başarısız olduğunu belirtir. Basitçe ifade etmek gerekirse, formül iki bölgenin kesişimini, bu bölgelerin toplam büyüklüğüne göre normalize eder. Bu sayede, tahminin ne kadar doğru olduğunu sayısal bir şekilde görmemizi sağlar^{9,104}. Gerçek pozitif GP, yanlış pozitif YP ve yanlış negatif YN olmak üzere formülü:

$$\text{Dice Katsayısı} = \frac{2 \times GP}{2 \times GP + YP + YN}$$

2.2.3.2. Jaccard indeksi (Intersection over Union – IoU)

Jaccard indeksi (Intersection over Union, IoU), aynı zamanda kesişim/birleşim olarak da adlandırılır. Beklenen segmentasyon ile gerçek (ground truth) segmentasyonun ne kadar örtüştüğünü ölçer. Bu ölçüm, iki segmentasyonun kesişim alanının, birleşim alanına oranı alınarak yapılır.

Jaccard indeksi, Dice katsayısı ile benzerlik gösterir çünkü her ikisi de monoton ya da pozitif korelasyon sergiler¹⁰⁴. Gerçek pozitif GP, yanlış pozitif YP ve yanlış negatif YN olmak üzere formülü:

$$\text{Jaccard indeksi} = \frac{GP}{GP + YP + YN}$$

2.2.3.3. Hausdorff Mesafesi

Tahmin edilen segmentasyon sınırları ile ground truth (gerçek) segmentasyon sınırları üzerindeki herhangi iki nokta arasındaki en büyük mesafeyi ölçer. Daha düşük Hausdorff mesafesi, segmentasyonlar arasında daha iyi hizalanma (uyum) olduğunu gösterir. Hausdorff mesafesi, kümeler arasındaki en kötü eşleşmeyi ölçer. Dışarıda kalan tek bir uç nokta (outlier) mesafeyi ciddi şekilde yükseltebilir. İki küme tamamen aynıysa (segmentasyon ve gerçek sınırlar birebir örtüşüyorsa) mesafe sıfır olur. Formülü:

$$HM(A, B) = \text{üst sınır}(\text{üst sınır}_{a \in A} \text{ alt sınır}_{b \in B} |a - b|, \text{üst sınır}_{b \in B} \text{ alt sınır}_{a \in A} |b - a|)$$

2.2.3.4. %95 Hausdorff Mesafesi

Standart Hausdorff mesafesi tek bir uzak nokta tarafından aşırı etkilenebildiği için pratik uygulamalarda %95 Hausdorff mesafesi tercih edilir. A ve B kümeleri için tüm noktaların birbirine olan en yakın mesafeleri hesaplanır. Bu mesafelerin %95’lik yüzdelik değeri alınır. Yani en uzak %5’lik uç noktalar göz ardı edilir. Gürültü ve uç noktalara karşı daha dayanıklıdır. Segmentasyon doğruluğunu daha anlamlı bir şekilde ölçer. Bu yaklaşım, özellikle tıbbi görüntü segmentasyonu gibi alanlarda daha dengeli ve gerçeğe daha yakın bir hata değerlendirmesi sunar. İki küme tamamen aynıysa (segmentasyon ve gerçek sınırlar birebir örtüşüyorsa) mesafe sıfır olur.

Herhangi bir tıbbi görüntü segmentasyon görevi için kullanılacak değerlendirme ölçütlerinin seçimi, segmentasyonun özel amaçları, veri kümesinin özellikleri ve değerlendirme metriğinde aranan özel nitelikler gibi çeşitli faktörlere bağlıdır^{9,104,105}.

2.2.4. Receiver Operating Characteristics (ROC)

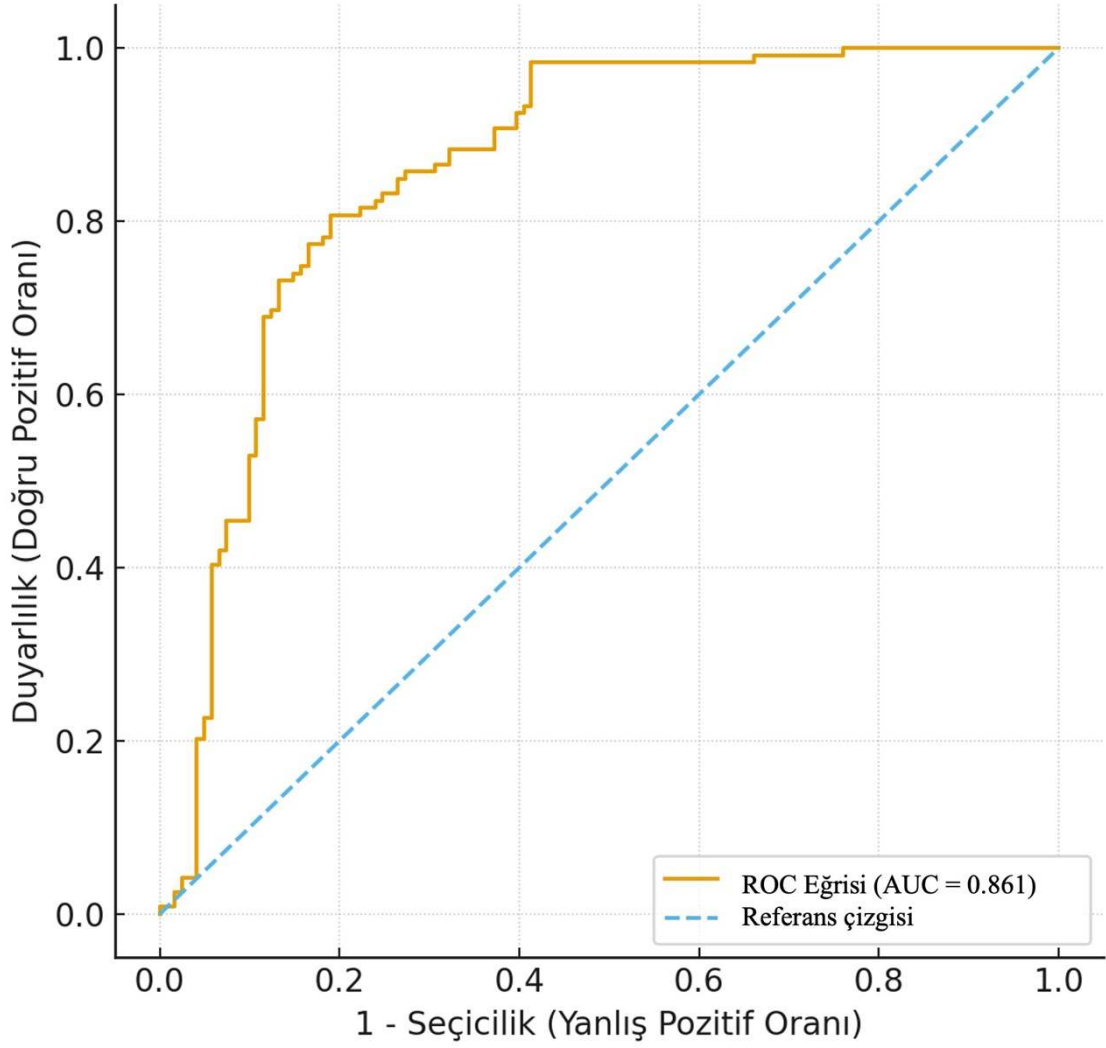
Receiver operating characteristics (ROC) eğrisi kavramı ilk olarak II. Dünya Savaşı döneminde radar sinyallerinin çözümlemesinde ortaya çıkmıştır. Radar sinyallerinin yardımıyla dost ve düşman ayrımının yapılması, aldatıcı sinyallerin oranlarının hesaplanması ve bu sinyallerin azaltılması için yürütülen çalışmalar ROC yönteminin geliştirilmesine ilham vermiştir¹⁰⁸. Tıp alanında ise ROC analizinin görüntüleme yöntemlerinin değerlendirilmesinde kullanımı 1960’lı yılların sonlarına uzanır^{109,110}. 1980’li yıllara dek ROC analizi ile ilgili yayımlanan çalışmaların sayısı yavaş da olsa artış göstermiştir. 1980’lerden itibaren ROC teknikleri radyografi, kredi puanlama ve benzeri birçok alanda analitik bir araç olarak benimsenmiş, bu dönemle birlikte literatürde belirgin bir artış gözlenmiştir¹⁰⁸. Günümüzde ROC eğrileri, özellikle radyoloji gibi birçok alanda tanısal testlerin değerlendirilmesinde yaygın biçimde kullanılmaktadır. Tanısal testlerin doğruluğunu saptamak açısından ROC eğrisi oldukça yararlı bir yöntem hâline gelmiştir^{111,112}.

ROC analizi, tanı testlerinin doğruluğunu istatistiksel olarak değerlendirmek için kullanılan bir tekniktir. Bu yöntemle testin duyarlılığı (sensitivity, yani doğru pozitif

oranı) ile seçiciliği (specificity, yani doğru negatif oranı) arasındaki ilişki en iyi biçimde ortaya konabilir¹¹³. Klinik uygulamalarda, bir testin hastaları doğru şekilde tanıma yeteneği “duyarlılık”, sağlıklı bireyleri doğru biçimde ayırt etme yeteneği ise “seçicilik” olarak adlandırılır ve ROC analizi bu iki ölçütü değerlendirmede önemli bir rol oynar¹¹⁴. Uygulamada çoğunlukla iki sınıfın (örneğin hasta/sağlam, evet/hayır, doğru/yanlış, kabul/red) ayrımı söz konusu olmakla birlikte, ikiden fazla grubun ayrılması gereken durumlar da olabilir. Çoklu sınıflandırma gereken koşullarda ise genellikle bu gruplar ikili ayrımlar hâline getirilebilir¹¹⁵. ROC eğrileri, bir tahmin edicinin gerçek sonuçla nasıl karşılaştırıldığını göstermek için kullanılır. ROC analizinin en büyük avantajlarından biri eşik değerinden bağımsız olmasıdır; bir tahmin edicinin performansı belirli bir eşik değeri olmadan tahmin edilir ve ayrıca belirli bir maliyet fonksiyonuna veya hedefe dayalı olarak en uygun eşiği seçmek için bir kriter sunar. Tipik olarak, bir ROC analizi duyarlılığın (gerçek pozitif oranı) farklı eşik değerleri için değişen özgülük (gerçek negatif oranı veya 1 - yanlış pozitif oranı) ile nasıl değiştiğini gösterir. Analizler ayrıca genellikle yanlış pozitifleri ve yanlış negatifleri eşit şekilde tartar. ROC analizlerinde, bir değişkenin tahmin yetenekleri genellikle eğri altında kalan alan (EAA) ile özetlenir ve bu, doğru parçalarının altındaki alanların entegre edilmesiyle bulunabilir¹¹⁶.

2.2.5. Eğri Altında Kalan Alan (EAA)

ROC eğrisini oluşturan her bir noktanın birleştirilmesiyle ortaya çıkan geometrik alanların toplamı, eğri altında kalan alanı (EAA) verir. EAA, ROC analizinde en sık kullanılan istatistiksel ölçütlerden biridir¹¹⁷. ROC eğrisi altında kalan alan, farklı olası eşik değerleri dikkate alınarak testin genel performansını ortaya koyan bir ölçüttür. Bu değer, hastalık prevalansından bağımsız olduğundan tanı testlerinin doğruluğunu özetlemek için güvenilir bir göstergedir¹¹⁸. Tanısal testlerin başarısını değerlendirerek hasta ve sağlıklı bireyleri ayırt etmede kullanılan önemli göstergelerden biri de bu EAA değeridir. ROC eğrisinin çizimi için testin duyarlılık ve seçicilik değerlerine ihtiyaç vardır. Grafikte x eksenini 1-seçicilik, y eksenini ise duyarlılığı gösterecek şekilde ROC eğrisi oluşturulur ve Şekil 2.2’de rastgele bir ROC grafiği örnek olarak gösterilmiştir.



Şekil 2. 3. ROC ve EAA örneği

Şekil 2.2.'de de görülebileceği gibi ROC eğrisi (0,0) ile (1,1) noktaları arasında yer alır. EAA değeri 0 ile 1 arasında değişir. Bu alan, rastgele seçilen iki bireyden biri hasta diğeri sağlıklı olduğunda, testin hasta olan bireye daha yüksek bir değer atama olasılığını ifade eder. Dolayısıyla EAA'nın 1'e yaklaşması, testin ayırt etme gücünün yüksek olduğunu; 1'e eşit olması ise testin kusursuz bir performansa sahip olduğunu gösterir. Anlamı, test, hasta ve sağlıklı bireyleri %100 doğru ayırmakta olup, duyarlılık ve seçicilik %100'dür; hiç yanlış pozitif ya da yanlış negatif yoktur. Eğri sol üst köşeden (0,1) geçer ve neredeyse köşeyi kaplayan bir alan oluşturur. EAA'nın 0.5'e eşit olması ayırt edemeyen test anlamına gelmekte olup, şansa dayalı tahmin yapmaktan daha iyi değildir. Grafiği (0,0) ile (1,1) arasında giden çapraz bir doğru çizgi şeklindedir. EAA'nın 0 olması tamamen ters sınıflandırma anlamına gelip, testin çıktıları tamamen yanlıştır;

yani hasta olanlara sürekli sağlıklı, sağlıklılara sürekli hasta demektedir. ROC eğrisi (0,1) ile (1,0) arasında, yani ana köşegenin altındaki alanı temsil eder. Bu nedenle tanısal bir testin güvenilirliğini değerlendirmek için EAA'nın 1'e yakın olması istenir¹¹⁸⁻¹²⁰. EAA değerlerinin ayırt edilebilirlik seviyeleri Çizelge 2.2.'de gösterilmektedir¹²¹. Doğru pozitif duyarlılığı DPD, yanlış pozitif oranı YPO olmak üzere formülü:

$$EAA = \int_0^1 DPD(YPO) dYPO$$

Çizelge 2. 2. EAA değerlerinin ayırt edilebilirlik seviyeleri

EAA değeri	Seviye
$0.9 \leq AUC$	Mükemmel
$0.8 \leq AUC < 0.9$	İyi
$0.7 \leq AUC < 0.8$	Orta düzey
$0.6 \leq AUC < 0.7$	Zayıf
$0.5 \leq AUC < 0.6$	Başarısız

Bu değerler, segmentasyon sonucunda elde edilen çıktı kümesiyle, gözlemci tarafından oluşturulan ve gerçek referans kümesi (ground-truth) olarak kabul edilen maske bölgelerinin üst üste bindirilmesiyle ilgili bölgelerdeki piksel sayılarının hesaplanmasıyla elde edilmiştir.

Segmentasyon için kullanılan performans ölçütleri Çizelge 2.3'te gösterilmiştir.

Çizelge 2. 3. Segmentasyon performans ölçütleri tablosu

Metrik	Formül	Açıklama
Doğruluk	$\text{Doğruluk} = (TP + TN) / (TP + TN + FP + FN)$	Tüm pikseller içinde doğru sınıflandırılanların oranıdır. Sınıf dengesizliği varsa yanıltıcı olabilir.
Kesinlik	$\text{Kesinlik} = TP / (TP + FP)$	Pozitif tahminlerin ne kadarının doğru olduğunu ölçer. Yanlış pozitiflerin maliyetli olduğu durumlarda önemlidir.
Duyarlılık (Recall)	$\text{Duyarlılık} = TP / (TP + FN)$	Modelin pozitif örnekleri ne kadar yakalayabildiğini ölçer. Yanlış negatiflerin kritik olduğu durumlarda önemlidir.
Dice katsayısı	$DK = 2TP / (2TP + FP + FN)$	Gerçek bölge ile tahmin edilen segmentasyonun örtüşme derecesini ölçer. Yanlış pozitif ve yanlış negatifleri dikkate alır.
F1 skoru	$F1 = 2 \times K \times D / (K + D)$	Kesinlik ve duyarlılığın harmonik ortalamasıdır. 0–1 arasında değer alır.
Dice katsayısı	$(2 \times GP) / (2 \times GP + YP + YN)$	İki bölgenin kesişimini, bu bölgelerin toplam büyüklüğüne göre normalize eder.
Jaccard indeksi (IoU)	$\text{IoU} = TP / (TP + FP + FN)$	Tahmin edilen segmentasyon ile gerçek segmentasyonun kesişim alanının birleşim alanına oranıdır.
Hausdorff mesafesi	$H(A, B) = \max \left\{ \max_{s \in B} [d(s, A)], \max_{s \in A} [d(s, B)] \right\}$	Bir kümedeki her noktanın diğer kümeye olan en yakın mesafesi bulunur ve bu mesafelerin en büyüğü alınır.

D: Duyarlılık, DK: Dice katsayısı, GP: gerçek pozitif, K: Kesinlik, YP: yanlış pozitif, YN: yanlış negatif

2.3. Nazopalatin Kanal

NPK, sert damağın ön kısmında yer alır ve burun boşluğu ile ağız boşluğunu birbirine bağlayan bir kanaldır¹⁶. Bu kanalın içinde nazopalatin arterin terminal dalları, nazopalatin sinir, küçük tükürük bezleri, yağ dokusu ve bağ dokusu bulunur^{16,17}. Kanal, ağız boşluğunda santral kesici dişlerin posteriorunda tek bir insiziv foramen olarak ve burun boşluğunda genellikle iki adet olan Stenson forameni olarak devam eder. Bu foramenden her birinden, sfenopalatin arterin posterior septal dalı ve büyük palatin sinir ile sırasıyla bağlantı kurmak üzere, inen palatin arterin terminal dalı ve nazopalatin sinir geçer. NPK'nın anatomik görünümü ve varyasyonları hakkında ayrıntılı bilgi, implant yerleştirme ve anterior maksillada lokal anestezi gibi cerrahi girişimlerden önce hayati öneme sahiptir. Maksiller santral diş bölgesine uygulanan implantlar ile NPK'nın konumu arasındaki ilişkiye dair zorluklar ve anatomik sınırlamalar bildirilmiştir¹⁸. İmplant rehabilitasyonu komplikasyonları arasında, sinir dokusu ile temas sonucu implantın kemik ile osseointegrasyonun sağlanamaması veya duysal fonksiyon bozukluğu yer almaktadır¹⁹. Ayrıca, implantların güvenilir bir şekilde yerleştirilebilmesi amacıyla insiziv foramen yakınında augmentasyon yapıldığı da bildirilmiştir²⁰. Nazopalatin kanalın karakteristik tanımı ile içeriğinin ve değişken görünümünün bilinmesi, cerrahi planlamanın optimize edilmesi ve bu anestezi açısından hassas bölgede implant yerleştirme sırasında komplikasyonların önlenmesi açısından son derece önemlidir²¹.

NPK kisti ya da diğer adıyla insisiv kanal kisti, NPK'nın embriyonik epitel kalıntılarından gelişir²⁶. Çene kistlerinin yaklaşık %10'unu oluşturur ve çoğunlukla 40–60 yaş aralığında tanı konur. Görüntülemede NPK kistleri genellikle radyolusent lezyonlar şeklinde izlenir ve NPK'nın içinde ya da yakınında yer alır. Anatomik olarak ön maksillada, orta hatta ve insisiv foramenin üzerinde bulunur; çoğu zaman rutin radyografik incelemeler sırasında tesadüfen saptanır. Klinik olarak genellikle asemptomatiktir, ancak ön maksillanın palatal bölgesinde şişlik, akıntı veya ağrı görülebilir. Ayırıcı tanıda radiküler kist, median palatal kist ve primordial kist dikkate alınmalıdır²⁷. İnsisiv kanalın yatay çapı 10 mm'ye kadar ulaşabileceğinden, NPK kistlerinin genişlemiş bir insisiv foramen ile karıştırılmaması önemli bir noktadır¹⁷.

Ortopantomografi (OPG) ve intraoral X-ışınları gibi geleneksel radyografik teknikler, alveolar kemiğin bukko-lingual genişliği, anatomik yapıların durumu veya ilgili bölgenin 3 boyutlu (3B) yapısı hakkında herhangi bir bilgi sağlamaz²⁸. OPG ve ağız içi görüntülerde tespiti zor olan çeşitli yapıların 3B görüntülenmesi için KIBT önerilir²⁹. KIBT'deki doğrusal ölçümler ile altın standart olan maksillofasiyal yapıların doğrudan ölçümleri arasında belirgin bir fark olmayıp³⁰, KIBT kullanılarak kanal konumunun ve boyutsal özelliklerinin değerlendirilmesinin klinik durumlarda kullanılabilecek ayrıntılı bilgiler sağladığı gösterilmiştir³¹.

2.4. Konik Işınlı Bilgisayarlı Tomografi

KIBT görüntüleme, OPG'nin klinik kullanıma girmesinden bu yana maksillofasiyal görüntüleme alanındaki en önemli teknolojik yeniliklerden biri olarak öne çıkmaktadır. KIBT, ilk olarak 1980'li yılların başında anjiyografik incelemeler için ticari kullanıma sunulmuştur. Bu yöntemde, hacimsel veri elde edebilmek amacıyla koni ya da piramit biçiminde iyonlaştırıcı radyasyon kaynağı ile döner bir kızak üzerine sabitlenmiş iki boyutlu alan dedektörü entegre edilerek çok sıralı iletim görüntüleri üretilir. 1990'ların başlarında gerçekleşen çeşitli teknik gelişmeler, diş hekimliği muayenelerinde maksillofasiyal görüntüleme için yeterince kompakt ve ekonomik KIBT ünitelerinin üretilmesini mümkün kılmıştır³².

KIBT teknolojisinin klinik alanda kullanılmaya başlaması, maksillofasiyal görüntülemede iki boyutlu incelemelerden üç boyutlu hacimsel yaklaşıma geçişin öncüsü olmuştur. Bu yöntem; görüntü elde etme, görüntülerin görselleştirilmesi ve elde edilen verilerin yorumlanması olmak üzere üç temel aşamadan oluşmaktadır³². İlgili bölge etrafında tek bir rotasyonla görüntü alınması, ışınlama süresini kısaltarak hastanın maruz kaldığı radyasyon miktarını azaltır³³. KIBT ile yapılan görüntülemede, radyasyon dozu konvansiyonel BT'ye kıyasla yaklaşık 15 kat daha düşüktür, OPG'ye göre ise 4 ila 15 kat arasında daha yüksektir. Ayrıca 10–70 saniyelik kısa tarama süresi sayesinde, hastanın hareketine bağlı artefaktların oluşma olasılığı da önemli ölçüde düşer^{33,34}.

KIBT kraniofasiyal bölgede pek çok endikasyonda kullanılmaktadır. Bu yöntem; implant cerrahisinin planlanması ve sonrasındaki takibi, ortognatik operasyonların hazırlığı, çeşitli patolojik oluşumların incelenmesi, temporomandibular eklemin kemik

bileşenlerinin değerlendirilmesi, travma olgularının analizi ve gömülü dişlerin konum tespiti gibi durumlarda yaygın biçimde tercih edilmektedir^{32,35}.

2.4.1. KIBT'nin BT'ye Göre Üstünlükleri

Daha Düşük Radyasyon Dozu: KIBT'de hastanın aldığı radyasyon miktarı BT'ye kıyasla belirgin biçimde daha azdır. Çoğu cihazda, inceleme alanının genişliğine bağlı olarak X-ışını şiddetini ayarlayabilen sistemler bulunur. Bu sayede hem ışın demetinin boyutu değiştirilebilir hem de hastanın gereksiz radyasyona maruz kalması engellenir³².

Ekonomik ve Kompakt Yapı: BT ünitelerine göre hem cihaz maliyetleri daha düşüktür hem de KIBT sistemleri daha küçük boyutlara sahiptir.

Kısa Tarama Süresi: Daha hızlı görüntüleme yapılabildiğinden, hasta hareketine bağlı artefakt oluşma olasılığı belirgin şekilde azalır.

Yüksek Detaylı İnceleme: Maksillofasiyal bölge için daha ince kesitler elde edilerek çok daha ayrıntılı bir değerlendirme yapılmasına olanak tanır.

Kolay ve Uygun Maliyetli İşleme: Tarama verilerinin incelenmesi için gerekli yazılımlara dönüştürme süreci daha basittir ve maliyeti düşüktür. Bu durum, görüntülerin yorumlanma süresini de kısaltır.

Üstün Voksel Özellikleri: BT'de voksel yapıları dikdörtgen prizması biçiminde ve üç düzlemde farklı boyutlardayken, KIBT'de voksel boyutları küp şeklindedir, izotropiktir ve hacmi yaklaşık 0,07–0,4 mm³ arasındadır (BT'de bu değer yaklaşık 0,3 mm³'tür)^{45,77}. Bu izotropik voksel yapısı, görüntü kalitesini artırır.

Metal Artefaktlarının Azlığı: Metal restorasyonlardan kaynaklanan artefaktlar KIBT'de daha az görülür^{32,33,36}.

2.4.2. KIBT'nin BT'ye Kıyasla Dezavantajları

Yumuşak Doku Sınırlılığı: KIBT, yumuşak dokuların görüntülenmesinde daha kısıtlı kapasiteye sahiptir ve kontrast çözünürlüğü BT'ye göre daha düşüktür.

Hounsfield Skalası Eksikliği: Kesitsel BT’de kullanılan Hounsfield birimlerinin KIBT’de bulunmaması en önemli dezavantajlardan biridir.

Tarama Alanının Kısıtlılığı: Daha küçük dedektörlerle çalışması nedeniyle incelenebilecek anatomik bölge boyutu sınırlıdır^{32,33,35}.



3. GEREÇ ve YÖNTEMLER

Bu retrospektif çalışma, KIBT görüntüleri üzerinde NPK otomatik segmentasyonunu ile yapay zeka destekli bir FCN modeli kullanarak modelin klinik performansını değerlendirmek amacıyla tasarlanmıştır.

3.1. Etik Kurul Onayı

Bu çalışma, Çukurova Üniversitesi Tıp Fakültesi Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu'ndan 8 Aralık 2023 tarih ve 139 toplantı sayılı, 8 numaralı karar ile onaylanmış ve Helsinki Deklarasyonu'nun güncel ilkelerine uygun olarak gerçekleştirilmiştir.

3.2. Veri Toplama ve Hasta Seçimi

Bu retrospektif çalışma, 2010–2024 yılları arasında Çukurova Üniversitesi Diş Hekimliği Fakültesi Ağız, Diş ve Çene Radyolojisi Anabilim Dalı arşivinden elde edilen, 45'i kadın, 45'i erkek olmak üzere toplam 90 hastaya ait KIBT görüntülerini kapsamaktadır. Toplumun genelini temsil edebilmesi için 18-62 yaş arası her bir yaştan bir kadın ve bir erkek seçilmiştir. Her hasta seçiminde ilgili yaş ve cinsiyetteki dahil edilme kriterlerine uyan tüm hastalar listelenerek aralarından rastgele seçim yapılmıştır.

Dahil etme kriterleri:

- 18 yaş ve üzeri bireyler
- Nazopalatin kanal anatomisinin KIBT'de belirgin olarak izlenebilmesi
- KIBT görüntüsünde belirgin artefakt olmaması

Hariç tutma kriterleri:

- KIBT taramalarında hareket / metal artefaktları veya eksik veri bulunan bireyler,
- 18 yaşından küçük bireyler,
- Kraniyofasiyal travma veya cerrahi işlem hikayesi olan bireyler,

- Kraniyofasiyal patolojisi veya anomalisi / sendromu bulunan bireyler ve
- Sistemik olarak sağlıklı olmayan (özellikle metabolik kemik hastalıkları) bireyler.

3.3. KIBT Analizi

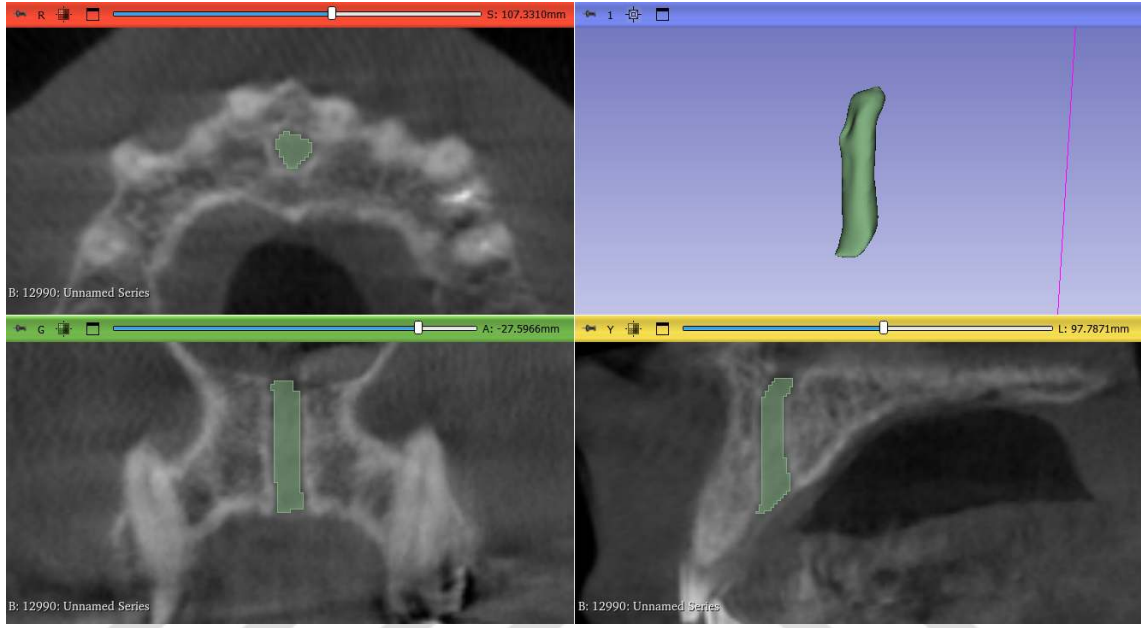
Tüm KIBT verileri Planmeca ProMax® 3D Mid cihazı (Planmeca, Helsinki, Finlandiya) kullanılarak elde edilmiştir. Cihaz ayarları 90 kV, 10 mA, 27 saniye tarama süresi, 0,4 mm voksel boyutu ve 20 × 17 cm görüş alanı (FOV) şeklindedir. DICOM (Digital Imaging and Communications in Medicine) formatındaki ham veriler, daha sonra manuel segmentasyon için açık kaynaklı 3D Slicer (sürüm 5.8.1 ; MIT, Cambridge, MA, ABD) programına aktarılmıştır ve tüm görüntüler koronal, aksiyel ve sagittal kesitlerde ve 0.4 mm kesitler üzerinden değerlendirilmiştir.

Veriler maksillofasiyal bölgeyi ve kafa kaidesini içermektedir. Hastalara yutkunmaksızın, burundan yavaş solunum yapmaları söylenmiş ve hastaların dişlerini maksimum interküspidasyonda sabit bir şekilde pozisyonlandırması sağlanmıştır. Rutin tomografik kayıtlar alınırken baş pozisyonunun stabilizasyonu amacıyla özel bir sabitleyici ve birbirine dikey geçen iki lazer ışın hüzmesi kullanılmaktadır. Vertikal hüzme, hastanın zahiri orta hattından geçecek şekilde, horizontal hüzme ise, Frankfurt horizontal düzlemine paralel olacak şekilde ayarlanmaktadır.

3.4. Manuel Etiketleme (Ground Truth)

Nazopalatin kanalın manuel etiketlemesi, bir araştırma görevlisi [Ağız Diş ve Çene Radyolojisi alanında 2.5 yıllık deneyim, (BB)] tarafından 3D Slicer'ın "Segment Editor" modülü kullanılarak gerçekleştirilmiştir. Tıbbi görüntü işleme ve analizi için güçlü bir araç olan 3D Slicer programı, farklı düzlemlerde hassas segmentasyona olanak tanır. Kanal sınırları "Paint" aracı ile çizilmiş ve "Draw" ve "Erase" araçları ile hassaslaştırılmıştır. Segmentasyon, anatomik sınırlara sıkı sıkıya bağlı kalınarak, NPK sınırlarının aşılmasına ve kanal içinde boşluk bırakılmamasına dikkat edilerek, her hasta için aksiyel, sagittal ve koronal düzlemlerde doğrulanarak kontrol edilmiştir (Şekil 3.1). Görsel yorgunluğu en aza indirmek amacıyla, her seans en fazla 5 görüntü ile sınırlandırılmıştır.

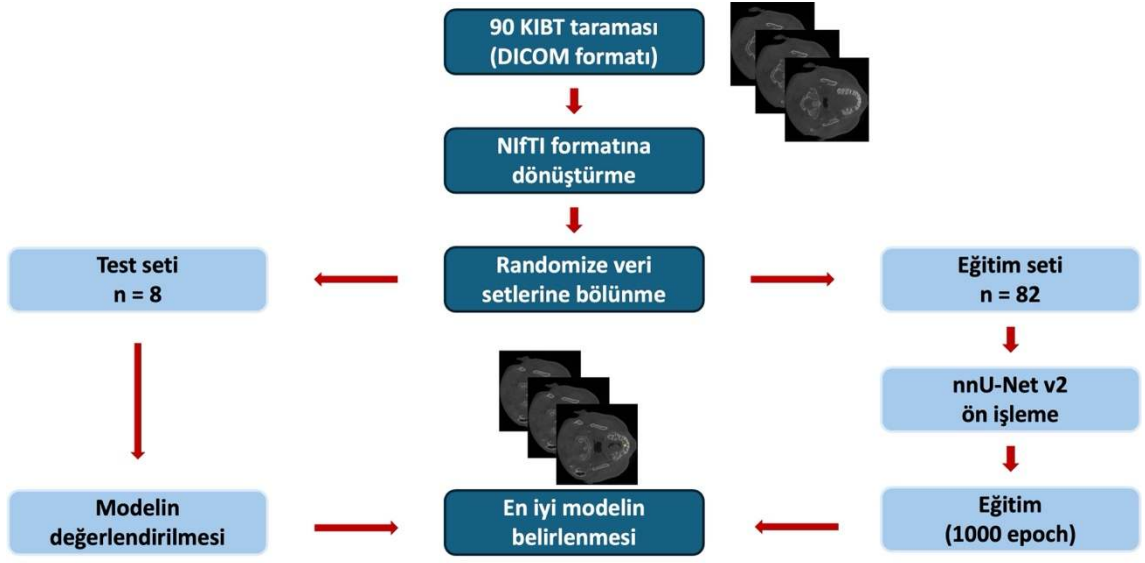
Segmentasyon süreci tamamlandıktan sonra, ağız dişi ve çene radyolojisi alanında 7 yıllık deneyime sahip kıdemli uzman (HDY) tarafından segmentasyonlar kontrol edilmiştir ve tüm etiketler üzerinde tam bir fikir birliğine varılmıştır. Bu doğrulama süreci, modelin gerçek dünya verileriyle ne kadar uyumlu olduğunu değerlendirmek için kritik öneme sahiptir.



Şekil 3. 1. Konik ışınlı bilgisayarlı tomografi görüntülerinde aksiyal, koronal, sagittal kesitlerde nazopalatin kanalın manuel olarak etiketlenmesi ve 3B görünümü

3.5. Derin Öğrenme Modelinin Geliştirilmesi

CranioCatch AI Aracı (CranioCatch, Eskişehir, Türkiye), Bayrakdar ve arkadaşları tarafından açıklanan derin öğrenme modeli geliştirme ve eğitim sürecinde kullanılmıştır¹²². 3B NPK otomatik tespiti için manuel olarak etiketlenmiş veriler, daha ileri işleme ve gelişmiş analiz için uygun hale getirilmek amacıyla, neuroimaging informatics technology initiative (NIfTI) formatına dönüştürülerek nnU-Net v2 derin öğrenme mimarisini temel alan bir yapay zeka modeli eğitmek için kullanılmıştır. Bu format dönüşümü, yüksek çözünürlüklü ve veri kaybı olmadan segmentasyon sağlar. Toplam 90 hastanın verisi, randomize olarak, 82'si eğitim seti ve 8'i test seti olmak üzere ikiye ayrılmıştır. Yapay zeka modeli oluşturma sürecinde kullanılan metodolojik iş akışı Şekil 3.2'de gösterilmiştir.



Şekil 3. 2. Derin öğrenme modeli geliştirmek ve eğitmek için kullanılan metodolojik iş akışı

Derin öğrenme modelleri tasarlanırken, modelde kullanılan algoritmaların performansını belirleyen çeşitli parametreler vardır. Bu parametreler, kullanılan soruna, veri kümesine veya sisteme göre seçilebilen değişkenlerdir. Bu değişkenlere hiperparametreler denir. Bu çalışmada eğitim için kullanılan hiperparametreler Çizelge 3.1’de gösterilmiştir.

Çizelge 3. 1. nnU-Net v2 modelinin eğitim parametreleri

Parametre	Değer
Model	nnU-Net v2
Döngü (epoch)	1000
Batch size	2
Öğrenme Oranı	0,01
Optimizasyon	ADAM
Aktivasyon fonksiyonu	LeakyReLU

Yeniden üretilebilirliği sağlamak için tüm rastgelelik değerleri sabitlenmiştir ve eğitim metrikleri ile kayıp eğrileri sürecin tamamında ayrıntılı olarak kaydedilmiştir. nnU-Net v2 tabanlı otomatik NPK 3B segmentasyon algoritması:

-GPU: Tesla V100-PCIE-16GB

-RAM: 228 GB

-CPU: Intel(R) Xeon(R) Gold 5218 teknik detaylarına sahip bir iş istasyonu kullanılarak yapılmıştır. Eğitim sürecinde ortalama epoch süresi 61,65 saniye olarak kaydedilmiş olup, pytorch, scipy, numpy, scikit-learn, SimpleITK, pandas, matplotlib, scikit-image, nibabel ve pydicom gibi kütüphaneler kullanılmıştır.

3.6. Örneklem Büyüklüğü

Bu çalışmada, literatürde nnU-Net v2 derin öğrenme modeli kullanılarak gerçekleştirilen segmentasyon araştırmalarına benzer şekilde, örneklem büyüklüğü hesaplamasında bağımlı örneklem *t*-testi yaklaşımı esas alınmıştır. Bu yöntem, gelecekteki değerlendirme adımlarında manuel ve otomatik segmentasyon performansları arasında istatistiksel olarak anlamlı fark olup olmadığının belirlenebilmesi için gerekli en az örneklem sayısını öngörmek amacıyla tercih edilmiştir¹²²⁻¹²⁴. Örneklem büyüklüğü analizi G*Power 3.1 yazılımı kullanılarak, çift yönlü hipotez testi varsayımı altında, %95 güç ($1-\beta = 0.95$), %5 anlamlılık düzeyi ($\alpha = 0.05$) ve orta düzeyde etki büyüklüğü ($d = 0.3$) kabul edilmiştir. Bu hesaplama sonucunda çalışmada gerekli minimum örneklem sayısı 54 olarak belirlenmiş; olası veri kayıpları ve dışlama kriterleri dikkate alınarak istatistiksel gücü koruyabilmek amacıyla 90 hasta çalışmaya dahil edilmiştir.

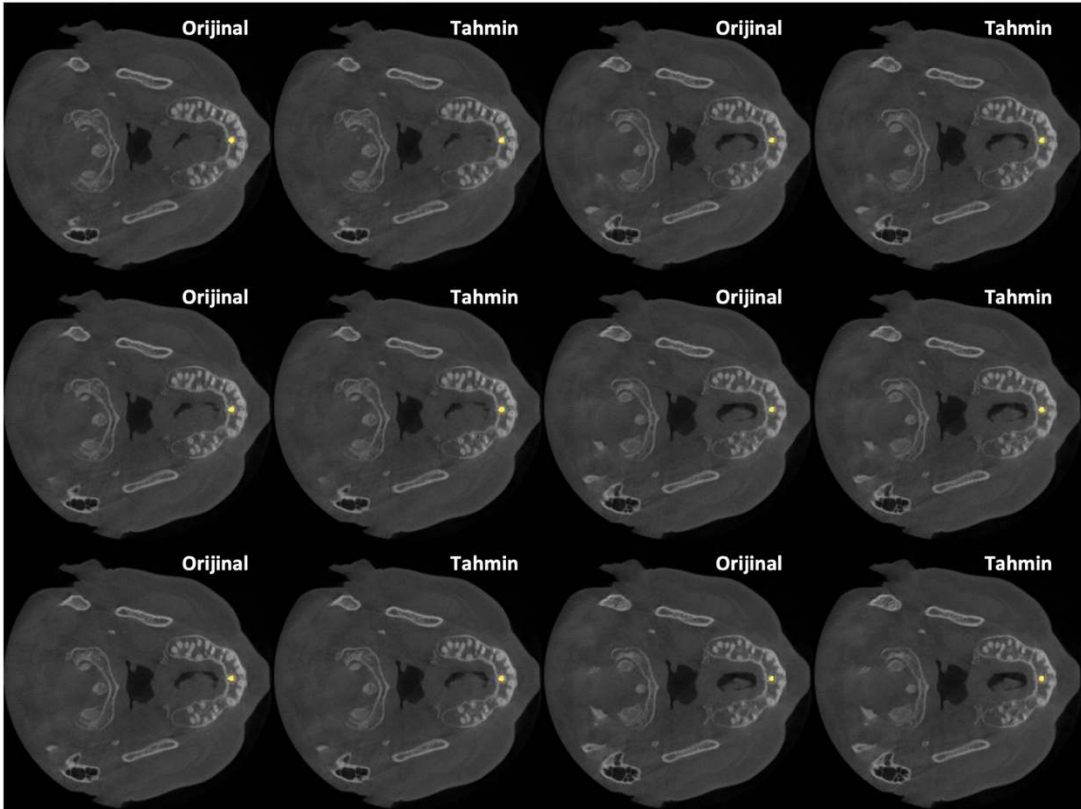
3.7. İstatistiksel Analiz

Çalışmada elde edilen otomatik segmentasyon sonuçlarının değerlendirilmesi için, klinik radyoloji ve yapay zeka alanında yaygın olarak kullanılan çeşitli performans metrikleri kullanılmıştır. Buna göre öncelikle karmaşıklık matrisinden, gerçek pozitif, yanlış pozitif, yanlış negatif oranı, doğruluk, kesinlik, duyarlılık ve F1-skoru hesaplanmıştır. Ayrıca modelin performansı Dice katsayısı, Jaccard indeksi (IoU), %95 Hausdorff mesafesi (mm) ve Hausdorff mesafesi (mm) ile değerlendirilmiştir. Ek olarak, ROC eğrisi oluşturularak EAA değeri hesaplanmıştır. EAA değerleri, $AUC \geq 0.9$ ise mükemmel, $0.8 \leq AUC < 0.9$ ise yüksek, $0.7 \leq AUC < 0.8$ ise orta, $0.6 \leq AUC < 0.7$ ise zayıf ve $0.5 \leq AUC < 0.6$ ise başarısız düzeyde ayırt edici olarak kabul edilmiştir¹²¹.

4. BULGULAR

Bu retrospektif çalışmaya yaş ortalaması: $40,0 \pm 13,1$ (alt sınır: 18, üst sınır: 62) olan yaş ve cinsiyet eşleştirilmiş 45'i kadın, 45'i erkek olmak üzere toplam 90 hasta dahil edilmiştir.

Veri seti: 82 hasta eğitim seti için, 8 hasta test seti için ayrıldı. Geliştirilen derin öğrenme ağı olan nnU-Net v2 modelinin performansı, test setinde elde edilen segmentasyon sonuçları üzerinden çeşitli metriklerle değerlendirilmiştir. Eğitim süresi: Ortalama 1000 epoch boyunca gerçekleştirilmiştir. KIBT kesitlerinde NPK'nın orijinal görüntüleri ve nnU-Net v2 derin öğrenme modeli kullanılarak oluşturulan tahmin (prediction) görüntüleri Şekil 4.1'de gösterilmiştir.



Şekil 4. 1. Aksiyel KIBT kesitlerinde nazopalatin kanalının orijinal görüntüleri ve nnU-Net v2 derin öğrenme modelinin tahmin (prediction) görüntüleri

NPK segmentasyonu için kullanılan nnU-Net v2 mimarisi sonucunda elde edilen karmaşıklık matrisine göre gerçek pozitif, yanlış pozitif, yanlış negatif, yanlış pozitif oranı, yanlış negatif oranı, doğruluk, kesinlik, duyarlılık ve F1-skoru değerleri sırasıyla

914, 282,13, 131,25, 3,09, 0,11, 0,99, 0,74, 0,89 ve 0,81 olarak analiz edilmiştir. Ayrıca Dice katsayısı, Jaccard indeksi (IoU) değeri, %95 Hausdorff mesafesi (mm) ve Hausdorff mesafesi (mm) sırasıyla 0,80, 0,67, 1,02 mm ve 1,90 mm olarak hesaplanmıştır (Çizelge 4.1.).

Çizelge 4. 1. nnU-Net v2 modelinin performans metrikleri ve değerleri

Performans metrikleri	NPK
Gerçek pozitif	914
Yanlış pozitif	282,13
Yanlış negatif oranı	0,11
Doğruluk (accuracy)	0,99
Kesinlik (precision)	0,74
Duyarlılık (recall)	0,89
F1-skoru	0,81
Dice katsayısı	0,80
Jaccard indeksi (IoU)	0,67
%95 Hausdorff mesafesi (mm)	1,02
Hausdorff mesafesi (mm)	1,90

Test setindeki 8 hasta için performans değerleri Çizelge 4.2’te sunulmuştur. Hasta bazlı değerlendirmelerde metrikler incelendiğinde, kesinlik değerlerinin bazı olgularda diğer performans göstergelerine kıyasla düşük olduğu görülmüştür. Bu durum, ilgili vakalarda yanlış pozitif oranının yüksek olmasından kaynaklanmaktadır. Özellikle test setindeki 6 nolu hastada kesinlik değerinin düşük olmasına karşın duyarlılık ve Dice katsayısı değerlerinin yüksek olması, modelin gerçek kanal alanlarını doğru biçimde yakaladığı ancak çevre dokulardan bazı bölgeleri de pozitif olarak sınıflandırdığına işaret etmektedir. Benzer şekilde, 8. hastada da kesinlik ve F1-skorlarında düşüş gözlenmiştir. Bununla birlikte, genel olarak hastalar arasında Dice katsayısı ve F1-skorlarının benzer

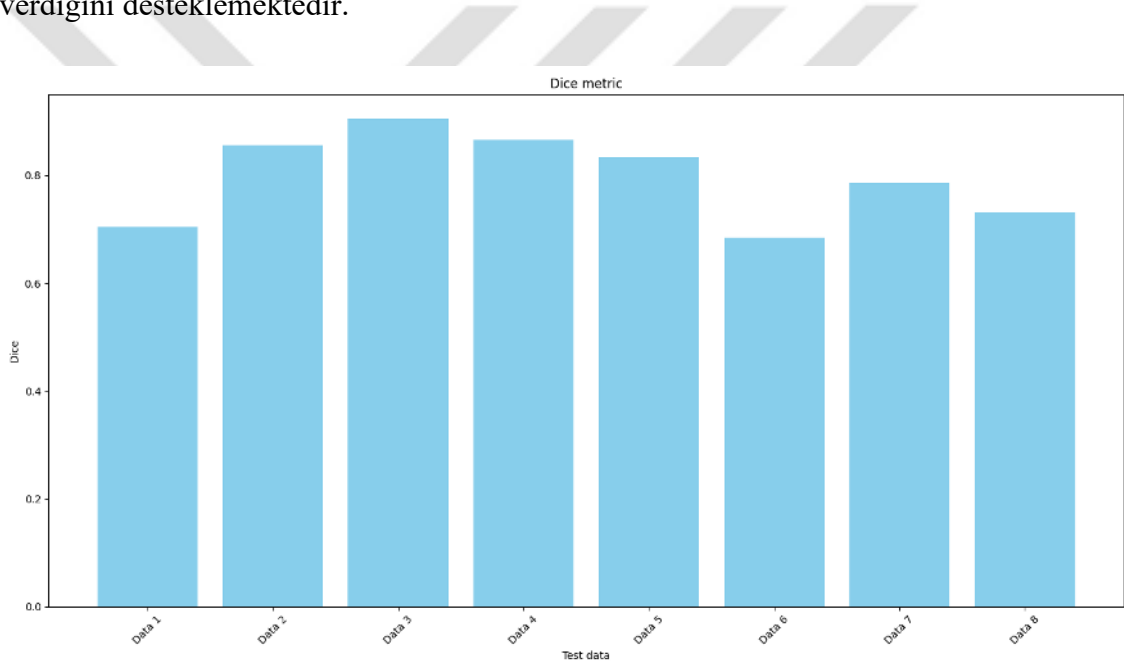
eğilimler göstermesi, modelin segmentasyon performansının vaka bazında tutarlı olduğunu göstermektedir.



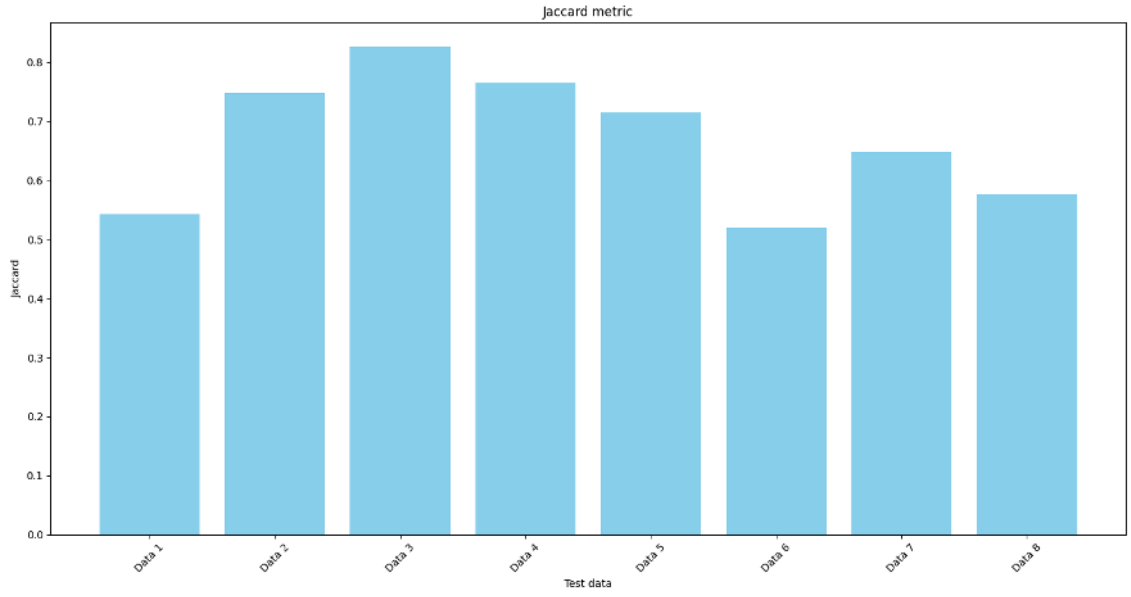
Çizelge 4. 2. nnU-Net v2 modelinin ve test seti hastalarının performans metrikleri ve değerleri

		Hasta no							
Metrik	Model	H1	H2	H3	H4	H5	H6	H7	H8
Gerçek pozitif	914	508	390	1513	1934	872	472	758	865
Yanlış pozitif	282,13	398	94	152	267	265	424	162	495
Yanlış negatif oranı	0,11	0,06	0,09	0,10	0,15	0,09	0,02	0,25	0,14
Doğruluk (accuracy)	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99
Kesinlik (precision)	0,74	0,56	0,81	0,91	0,88	0,77	0,53	0,82	0,64
Duyarlılık (recall)	0,89	0,94	0,91	0,9	0,85	0,91	0,98	0,75	0,86
F1-skoru	0,81	0,70	0,86	0,9	0,87	0,83	0,68	0,79	0,73
Dice katsayısı	0,80	0,70	0,86	0,9	0,87	0,83	0,68	0,79	0,73
Jaccard indeksi (IoU)	0,67	0,54	0,75	0,83	0,76	0,72	0,52	0,65	0,58
%95 Hausdorff mesafesi (mm)	1,02	2,32	0,40	0,40	0,57	0,57	0,89	0,57	2,43
Hausdorff mesafesi (mm)	1,90	3,35	0,69	1,26	1,70	1,20	1,79	1,44	3,77

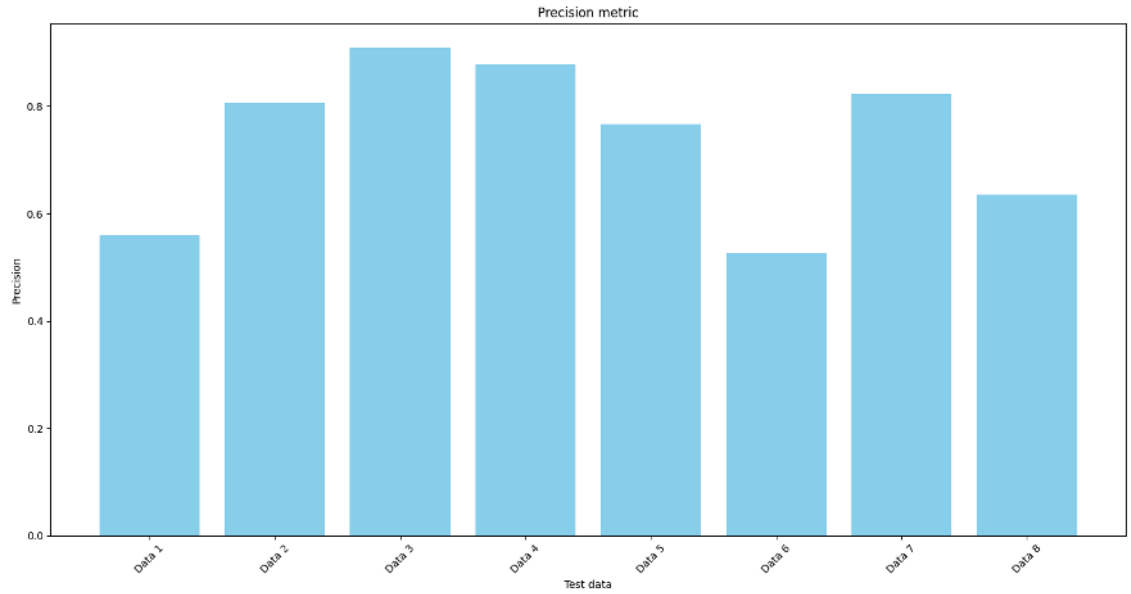
Şekil 4.2–4.5’te test setindeki her bir hastaya ait Dice katsayısı, Jaccard indeksi, kesinlik ve duyarlılık değerleri grafiksel olarak sunulmuştur. Grafikler incelendiğinde, çoğu hastada metrikler arasında yüksek korelasyon olduğu gözlenmiştir. Özellikle Dice katsayısı ve Jaccard indeksi değerlerinin paralel seyri, bu iki metriğin segmentasyon başarısını benzer biçimde yansıttığını göstermektedir. Bazı hastalarda duyarlılığın kesinlikten yüksek olması, modelin kanal sınırlarını geniş biçimde tahmin ettiğini, buna karşın aşırı segmentasyon eğilimi olabileceğini düşündürmektedir. Ayrıca, kesinliğin daha yüksek olduğu örneklerde modelin daha konservatif bir segmentasyon yaptığı görülmüştür. Hasta bazlı dağılımın genel olarak homojen olması, modelin genelleme kapasitesinin güçlü olduğunu ve farklı morfolojik varyasyonlarda da istikrarlı çıktılar verdiğini desteklemektedir.



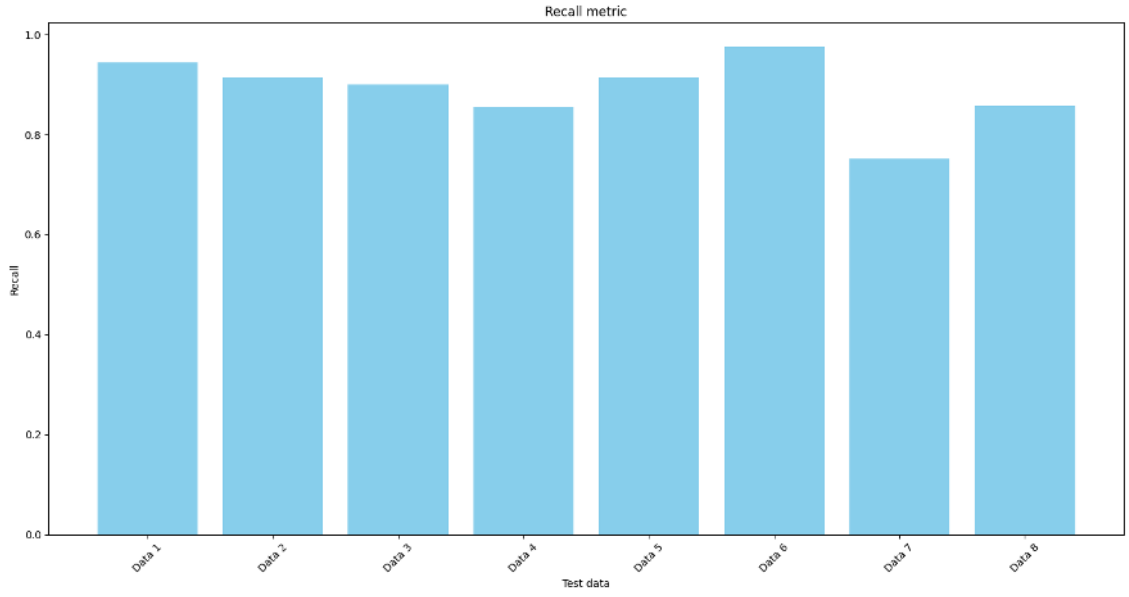
Şekil 4. 2. Test setinde hasta bazlı Dice katsayılarının grafiksel gösterimi



Şekil 4. 3. Test setinde hasta bazlı Jaccard indeksi (IoU) skorlarının grafiksel gösterimi

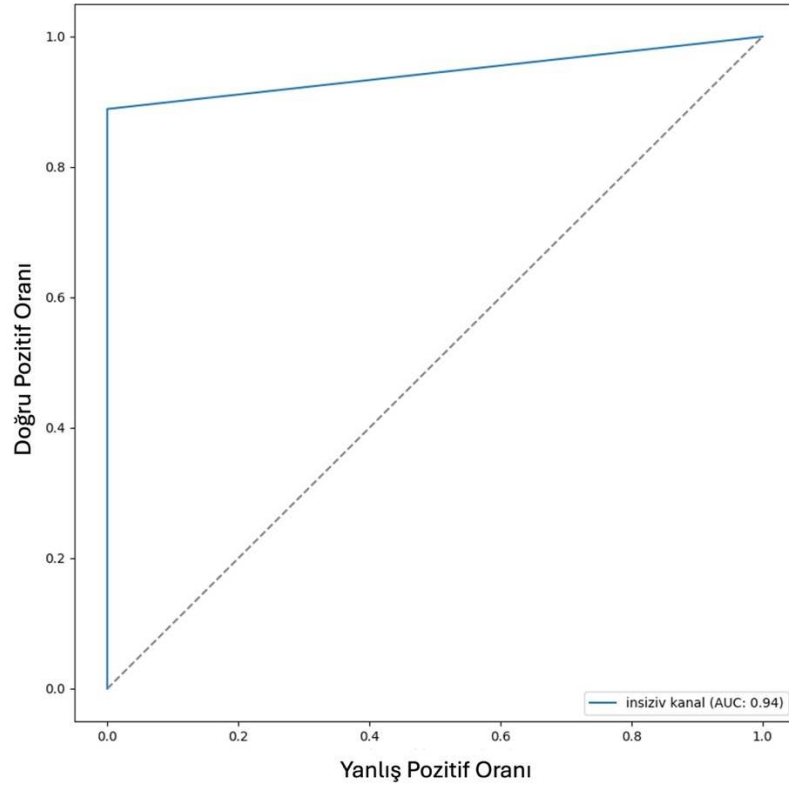


Şekil 4. 4. Test setinde hasta bazlı kesinlik değerlerinin grafiksel gösterimi



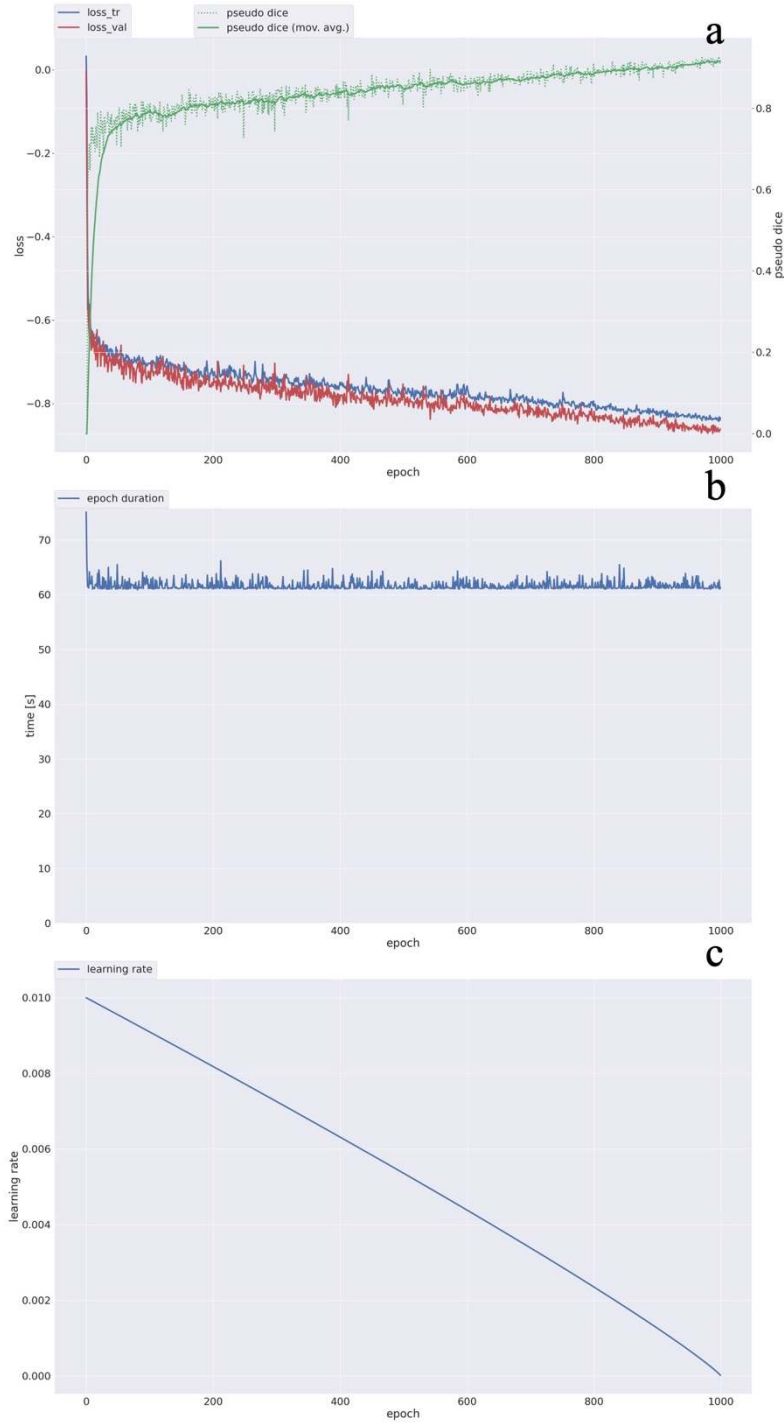
Şekil 4. 5. Test setinde hasta bazlı duyarlılık (recall) değerlerinin grafiksel gösterimi

ROC eğrisi Şekil 4.6’da gösterilmiştir. EAA değeri ise 0,94 olarak hesaplanmış olup, bu değer mükemmel¹²¹ ayırt ediciliği göstermektedir.



Şekil 4. 6. ROC eğrisi ve EAA

Modelin eğitimi hakkında daha fazla bilgi edinebilmek amacıyla, modelin başlangıcından itibaren her bir döngüde (epoch) elde edilen Dice katsayısı, kayıp fonksiyonu, döngü (epoch) süresi ve öğrenme oranı değerlerini gösteren grafik Şekil 4.7’de yer almaktadır. Dice katsayısı için görselleştirilen grafikte (Şekil 4.7.a) başlangıçtaki (ilk 50 epoch civarı) keskin düşüşle, modelin hızla öğrenmeye başladığı, sonrasında (yaklaşık 200. epoch’tan itibaren) hem eğitim hem doğrulama kaybının birlikte azalmaya devam ettiği ancak eğrinin giderek daha yatay hale geldiği görülmektedir. Bu, modelin artık veriden maksimum bilgiyi öğrendiğini ve yakınsama dönemine girdiğini göstermektedir. Pseudo Dice (segmentasyon başarısı) başta 0 civarında iken hızla yükselmiş, sonrasında küçük dalgalanmalarla artmaya devam etmiştir. Hareketli ortalamanın (yeşil düz çizgi) sürekli yukarı yönlü olması, modelin performansının epoch ilerledikçe kararlı biçimde arttığını, 800. epoch civarından sonra artış hızının yavaşlaması, platonun (doyum noktası) başlamış olabileceğini göstermektedir. Epoch Süresi için görselleştirilen grafikte (Şekil 4.7.b) ortalama epoch süresi ~60–65 saniye, ilk birkaç epoch’ta biraz daha yüksek (~75 saniye) ve 50. Epoch civarından sonra süre stabil hale gelmiş, yani ciddi bir dalgalanma yok olarak görülmektedir. Bu durum da zamanla sistemin optimize olarak, sabit hızda ilerlediğini ve modelin eğitiminin sorunsuz ve dengeli çalıştığına işarettir. Şekil 4.7.c’de öğrenme oranı başlangıçta 0.01 civarındayken epoch ilerledikçe lineer olarak azalarak 1000. epoch sonunda 0’a yaklaşmıştır. Başta yüksek oranla hızlı öğrenme ve sonlara doğru düşük oranla ağırlıkların hassas ayarlanması, kayıp eğrilerinin düzgün azalmasını ve dice katsayısının istikrarlı yükselmesini desteklemiştir.



Şekil 4. 7. Döngü (epoch) boyunca Dice katsayısı (a), kayıp fonksiyonları, epoch süresi (b) ve öğrenme oranı (c)

5. TARTIŞMA

NPK, nazal kavite ile oral kaviteyi birbirine bağlayan ve içinden arter ve sinir gibi önemli yapıların geçtiği kritik bir anatomik bölge olup, özellikle patolojik oluşumların değerlendirilmesi ya da implant cerrahisi planlaması gibi durumlarda, anatomik ve morfolojik özelliklerinin bilinmesi büyük klinik önem taşır¹⁶⁻²¹. Literatürde NPK ile ilgili derin öğrenme tabanlı yapay zeka çalışmaları incelendiğinde, NPK kisti ile ayırt edilmesine yönelik 2B görüntüleme yöntemlerine dayanan çalışmalar dikkati çekmektedir¹²⁵⁻¹²⁷. NPK ve NPK kisti, anatomik olarak maksiller anterior orta hatta yer almakta ve sıklıkla benzer radyografik görünüm sergileyerek görüntüleme ile ayırt edilmelerini özellikle güçleştirmektedir. Özellikle küçük veya asemptomatik NPK kisti olgularında, genişlemiş bir insisiv foramen ile karıştırılma olasılığı mevcuttur^{16,17,27}. Bu benzerlikler tanısal belirsizliğe yol açabilir ve NPD ile NPK kistini doğru biçimde ayırt edebilecek nesnel ve nicel yöntemlere duyulan ihtiyacı vurgular.

Kise ve ark. OPG görüntülerinde NPK kisti, radiküler kistler ve normal grupları ayırt etmek için GoogLeNet mimarisinden türetilmiş bir DetectNet modeli geliştirmiştir. Bu çalışmada NPK kisti için duyarlılık, kesinlik ve F1-skorları sırasıyla 0,83, 0,92 ve 0,87 olarak bulunmuş; ayrıca NPK kisti, radiküler kistler ve normal gruplar için EAA değerleri (0,894, 0,913 ve 0,913) asistan hekimlerden elde edilen değerlerden daha yüksek analiz edilmiştir (sırasıyla 0,794, 0,856 ve 0,863)¹²⁶. Benzer şekilde, Ito ve ark., OPG görüntülerinde CNN modellerini uygulamış ve LeNet'in (doğruluk: 0,853, duyarlılık: 0,720, özgüllük: 0,920) ve VGG16'nın (doğruluk: 0,880, duyarlılık: 0,760, özgüllük: 0,940) gözlemcilerden daha iyi performans gösterdiğini bildirmiştir (radyoloji uzmanı: doğruluk: 0,800, duyarlılık: 1,000, özgüllük: 0,700; radyolog: doğruluk: 0,747, duyarlılık: 0,760, özgüllük: 0,740)¹²⁵. Bir başka güncel çalışmada Lee ve ark., OPG görüntülerinde NPK kisti ve periapikal kistlerin otomatik tespit ve sınıflandırmasını değerlendirmiştir. EfficientDet-D3, en yüksek ortalama hassasiyet (mAP: %93,8) değerine ulaşmış ve iki radyoloji eğitimi almamış dış hekimine kıyasla daha yüksek performans göstermiştir (doğruluk: %94,4, duyarlılık: %94,4, özgüllük: %97,2, pozitif prediktif değer: %94,6, negatif prediktif değer: %97,2)¹²⁷. Bu çalışmaların tümü 2B görüntüleme yöntemlerine dayanmakta olup, CNN tabanlı derin öğrenmenin NPK kisti

tespitinde güçlü bir potansiyele sahip olduğunu ortaya koymaktadır. Deniz ve ark. ise KIBT kesitlerinde NPK'ya dik dilimlerden elde edilen aksiyal düzlemde kesitler içeren iki boyutlu veri kümeleri ile yaptıkları segmentasyon çalışmasında, bir yapay zeka çerçevesi olan "You only look once"ın (YOLO), YOLOv5x-seg versiyonunu kullanmışlardır. Buna göre duyarlılık, kesinlik, F1 skoru, IoU, mAP ve EAA değerlerini, NPK furkasyonunun olmadığı grup için 0,9680, 0,9953, 0,9815, 0,9636, 0,7930 ve 0,8841 ve furkasyonunun olduğu grup için 0,9827, 0,9975, 0,9900, 0,9803, 0,9637 ve 0,9510 olarak bildirmişlerdir¹²⁸.

Derin öğrenme tıbbi görüntüleri işlerken daha iyi performans sağlayabilse de, bazı sınırlamaları vardır ve bu sınırlamalar diğer klinik tıp alanlarına uygulanmasını kısıtlayabilir. Bir derin öğrenme ağı büyük miktarda veri kümesi ve güçlü hesaplama gücü gerektirir. Veri kümesi çok küçükse ve hesaplama gücü yetersizse, ağı eğitmek için daha fazla zamana ihtiyaç duyulur. CNN gibi çoğu derin öğrenme ağı, tıbbi görüntülerin etiketlenmesini gerektirir ve bu etiketleme işlemi manuel olarak yapıldığında oldukça zordur. Ancak, artan hesaplama gücü, veri depolama imkânları, tıbbi görüntülerin dijital depolanması ve derin ağ mimarilerindeki gelişmeler sayesinde bu sınırlamalar giderek aşılmaktadır. CNN'in gelişimiyle birlikte, manuel etiketlemedeki düşük hassasiyet sorunu da belli ölçüde çözülmüştür¹⁰⁰.

U-net, CNN'in en önemli anlamsal segmentasyon çerçevelerinden biridir. Yapısı nispeten açıktır ve "U" şeklinde bir mimariyi göstermektedir. U-net, biyomedikal alanda büyük etki yaratmıştır. Günümüzde tıbbi görüntü segmentasyonunu çözmek için ilk tercih haline gelmiştir^{100,101}. Aşağı örnekleme, çevresel bilgileri kademeli olarak göstermek için kullanılırken, yukarı örnekleme her katmandaki ayrıntılı bilgiyi geri yüklemek için kullanılmakta ve görüntü doğruluğu kademeli olarak artırılmaktadır. Günümüzde çoğu U-net ağı, görüntüden daha derin özellikler çıkarabilmek için birçok katmandan oluşmaktadır. Ayrıca, her bir mimari giderek daha karmaşık hale gelmekte ve optimizasyon ile olasılık modellerinden yararlanmaktadır. Görüntüleme sistemindeki uygulamalardan görüldüğü üzere, çoğu çalışma 2B U-net ağı kullanılarak yapılmıştır. Her modelin hesaplama açısından avantajları ve dezavantajları vardır. Ancak iyi bir genelleme elde edebilmek için, görüntü kalitesi ve sınıf dengesizliği dikkate alınarak en

uygun eğitim modu ve hiperparametrelerin seçildiği, optimizasyon katmanına sahip bir mimari gerekmektedir¹⁰⁰.

BT görüntülerinin segmentasyonu, patolojilerin doğru şekilde belirlenebilmesi açısından kritik öneme sahiptir. Bu işlem sonucunda tümörler, organların göreceli konumları veya anormal yapılar gibi klinik açıdan önemli veriler elde edilir. Bu hedef bilgilerin daha güvenilir çıkarılabilmesi için klasik U-net mimarisine dayalı çeşitli iyileştirmeler yapılmıştır. Derin öğrenme yöntemleri ile elde edilen hedef veri türleri oldukça çeşitlidir. Mesanenin doğru şekilde segmentasyonunu hedefleyen bir çalışmada CNN ile 3B tam bağlantılı rastgele alan özyineli sinir ağı (CRF-RNN) birleştirilmiştir¹²⁹. Performansı artırmak amacıyla iki kanallı yeni bir ön işleme yaklaşımı önerilmiştir. Klinik hedef hacim ve riskli organların segmentasyonu için derin dilate CNN kullanılmıştır. Eğitim setinde 218, doğrulama setinde ise 60 hasta yer almış; Dice benzerlik katsayısı klinik hedef hacim için %87,7, mesane için %93,4, sol femur başı için %92,1, sağ femur başı için %92,3 olarak bulunmuştur¹³⁰. Parotis bezinin ayrıştırılmasında 3B U-net kullanılmıştır. 2015 MICCAI baş-boyun oto-segmentasyon organizasyonunun test verilerinde 2B modelin Dice skoru %85,6 iken, 3D U-net ile %88'e ulaşılmıştır¹³¹. Akciğer kanseriyle ilişkili riskli organların segmentasyonu için CNN ve atlas tabanlı yöntemler uygulanmış; kalp, akciğer ve karaciğer için Dice skorları %87–95 arasında elde edilmiştir¹³². Karaciğer segmentasyonu amacıyla SBS-CNN yöntemi 100 hasta üzerinde test edilmiş ve $97,31 \pm 0,36$ doğruluk oranı rapor edilmiştir¹³³. Osteosarkom için çoklu denetimli FCN-MSFCN modeli geliştirilmiş ve 405 hasta üzerinde %87,80 doğruluk sağlanmıştır¹³⁴. Rektal kanser segmentasyonunda CNN kullanılarak 70 T2-ağırlıklı MR ve 100 planlama BT görüntüsü işlenmiş, %85 doğruluk elde edilmiştir¹³⁵. Osteosarkom segmentasyonu için çoklu denetimli artık ağ (residual network) yaklaşımı 1900 BT görüntüsü ile eğitilmiş, 405 BT görüntüsü üzerinde test edilmiş ve ortalama Dice skoru %89,22 bulunmuştur¹³⁶.

MRG verilerinin manuel olarak segmentasyonu oldukça zahmetli ve zaman alıcıdır; ayrıca büyük veri setlerinde doğruluk oranı düşebilmektedir. Bu nedenle derin öğrenme tabanlı yöntemlerle manuel segmentasyon, otomatik segmentasyona dönüştürülmüştür. Özellikle U-net mimarisi, farklı organ ve yapılar için etkili sonuçlar vermektedir. Çeşitli çalışmalarda kalp segmentasyonu, sol ve sağ ventrikül ayrımı için

üçlü U-net yapıları, çok görevli U-net ağırları veya U-seg-net gibi geliştirilmiş modeller kullanılmıştır¹³⁷⁻¹⁴².

Rektal kanser segmentasyonunda iki farklı yöntem önerilmiştir. Bunlardan ilki, T2-ağırlıklı MRG görüntülerinde 2D U-net tabanlı otomatik segmentasyon algoritmasıdır¹⁴³. Model, tümör tanıma ve tümör segmentasyonu olmak üzere iki aşamalı olarak eğitilerek verimlilik artırılmıştır. Benzer şekilde 2D U-net, kıkırdak ve menisküs segmentasyonunda da kullanılmış; relaksasyon süreleri ve morfolojik özelliklerin çıkarılmasında osteoartrit tanı ve takibine katkı sağlamıştır^{144,145}.

Kolorektal tümör segmentasyonunda normalizasyon tabanlı yöntemlerle VGG-16 kullanılmış ve pozitif prediktif değer %82,67, özgüllük %96,75, duyarlılık %87,85 olarak rapor edilmiştir. Bu yöntem, klasik U-net'e kıyasla anlamlı ölçüde daha başarılı bulunmuştur ($p < 0.05$)¹⁴⁶. Benzer şekilde, beyin ekstraksiyonunda¹⁴⁷, diz eklemi ve akciğer segmentasyonunda^{148,149}, femur¹⁵⁰, karaciğer damarları¹⁵¹, meme dokusu¹⁵² ve omurga disklerinde¹⁵³ U-net veya benzer CNN tabanlı yaklaşımlar başarıyla uygulanmıştır. Ayrıca prostat kontur segmentasyonunda kaskat FCN kullanılarak klasik U-net'e göre belirgin bir doğruluk artışı sağlanmıştır ($p < 0,001$)¹⁵⁴.

Genel olarak, U-net tabanlı derin öğrenme yöntemleri farklı anatomik yapıların MRG üzerinde otomatik segmentasyonunda yüksek doğruluk oranları sağlamaktadır. Bu bulgular, U-net'in tıbbi görüntülemelerde en önemli anlamsal segmentasyon çerçevelerinden biri olduğunu ve gelecekte görüntü analizi ve işleme süreçlerinde kritik rol oynamaya devam edeceğini göstermektedir.

Ultrason görüntüleme (USG), akustik, tıp, optik ve elektronik bilimlerinin birleşiminden doğmuş olup, vücuda gönderilen ses dalgalarının yansıma sinyallerinin işlenmesiyle iç organların görüntülerini üretmektedir. Bu görüntülerde, hedef yapıların doğru şekilde segmentasyonu kritiktir. Ancak manuel segmentasyon, yüksek uzmanlık gerektirmesine rağmen subjektif farklılıklara bağlı olarak çoğu zaman hatalara yol açmaktadır. Bu nedenle USG'de otomatik segmentasyon yöntemleri önemli bir ihtiyaç haline gelmiştir¹⁰⁰.

Geleneksel U-net yapısı, uçtan uca bir ağ mimarisi olup, yukarı örnekleme ve aşağı örnekleme aşamalarıyla özellik çıkarımı yaparak segmentasyon

gerçekleştirmektedir. Bununla birlikte, akustik gölgeler, düşük kontrast veya belirgin olmayan hedefler gibi durumlarda başarısı sınırlı kalmakta, Dice katsayısı düşük değerlerde seyretmektedir¹⁰⁰.

Alternatif iyileştirmelerden biri, CNN’de doğrusal olmayan aktivasyon fonksiyonu olarak scaling exponential linear elementin (SELU) kullanılmasıdır. Bu yöntem, 3B US hacimlerinden 2B görüntülerde levator kası boşluğunu otomatik olarak ayırmada başarılı bulunmuştur¹⁵⁵. Yakın dönemde, derin denetimli FCN modelleri¹⁵⁶, residual ağ yapıları (ResNet)¹⁵⁷ ve genişletilmiş konvolüsyon teknikleri¹⁵⁸ USG’de segmentasyon doğruluğunu artırmak için geliştirilmiştir. Bu yöntemler, zaman serisi bilgilerini daha etkin kullanarak birçok yapıda başarılı segmentasyon sağlamıştır.

Genel olarak, U-net ve türevleri, ultrason görüntülerinde akustik gölge, düşük kontrast, speckle gürültüsü ve eksik sınır gibi zorlukların üstesinden gelmede etkili bulunmuştur. Geleneksel makine öğrenmesi yöntemlerine kıyasla, çok katmanlı konvolüsyon yapıları sayesinde derin özellik çıkarımı yapabilmesi ve nihai segmentasyon sonucunu iyileştirmesi, U-net’i ultrason görüntülerinde tercih edilen yöntemlerden biri haline getirmiştir¹⁰⁰.

U-net aşağı örnekleme sonularını yukarı örnekleme sürecinde kullanır, bu da evriřim giriřini daha zengin hale getirir ve sonuların orijinal görüntü bilgisini daha iyi yansıtmasını saėlar. Ancak U-net aė yapısı yalnızca tek bir ölekte tahmin yapabilmektedir. Bu nedenle boyut deėiřimlerini iyi řekilde iřleyemez. Dolayısıyla geleneksel U-net aėının geliştirilmesi, bu sorunların etkin biimde çözölmesine yardımcı olabilir. U-Net, biyomedikal görüntülerin etkin segmentasyonu için yaygın řekilde kullanılan özel tasarlanmış bir fCN’dir⁸. Bu alıřmada kullanılan nnU-Net, U-Net’in geliştirilmiş bir versiyonudur ve: eřitli veri artırma teknikleri, farklı kayıp fonksiyonları, oklu görüntü öleklendirmesi gibi yenilikler sunar. nnU-Net’in öne ıkan özelliėi, farklı segmentasyon görevlerine uyarlanabilme kapasitesidir. Bu sayede farklı veri tipleri ve boyutlarına özel modeller geliştirilebilir. nnU-Net, veri setine özel hiperparametreleri belirlemek için “veri parmak izi” (data fingerprint) yöntemini kullanır. Bu parmak izleri; planlama parametreleri (kayıp fonksiyonu, optimizasyon, mimari), ıkarım parametreleri (yeniden örnekleme, normalizasyon, batch size, patch size), veri özellikleri ile

birleştirilerek elde edilir. Sonuçta, 2B, 3B ve 3B kademeli U-Net modelleri için en iyi konfigürasyon belirlenir ve test veri setine uygulanır¹⁰¹.

Bu çalışmada, tıbbi görüntü segmentasyonu için geliştirilen 3B derin öğrenme modeli olan nnU-Net v2 modeli kullanılmıştır. nnU-Net'in ilk versiyonu segmentasyon görevlerinde olağanüstü bir performans göstererek, birçok tıbbi görüntüleme uygulamasında standart haline gelmiştir¹⁵⁹. nnU-Net v2 modeli, önceki versiyonuna kıyasla daha esnek, daha hızlı ve daha performanslı hale getirilerek tıbbi görüntü segmentasyonunda en iyi sonuçları elde etmek için otomatik bir çözüm sağlamıştır.

Çalışmada, modelin güvenilirliğini artırmak amacıyla rastgelelik kaynakları kontrol edilmiş ve eğitim sürecine ait metrikler düzenli olarak izlenmiştir. Erken durdurma ve öğrenme oranı ayarlayıcılarının kullanımı aşırı uyum (overfitting) riskini azaltmış, düzenli kontrol noktaları ise en iyi modellerin korunmasına katkı sağlamıştır. Eğitim sonrasında, modelin genelleme yeteneği bağımsız test seti üzerinde de doğrulanmış ve yüksek segmentasyon başarısını sürdürdüğü görülmüştür. Bu durum, geliştirilen modelin klinik uygulamalarda güvenilir şekilde kullanılabileceğini ve pratikte önemli bir değer sunabileceğini göstermektedir.

Literatürde maksillofasiyal bölgede nnU-Net v2 mimarisi kullanılarak gerçekleştirilen birçok segmentasyon çalışması bulunmaktadır. Gümüşsoy ve ark., bifid mandibular kanalın otomatik olarak segmentasyonu amacıyla yürüttükleri araştırmada mandibular kanal ve bifid kanal için sırasıyla doğruluk oranlarını %99 ve %99, Dice katsayılarını 0,823 ve 0,469, kesinlik değerlerini 0,850 ve 0,701, duyarlılık değerlerini ise 0,801 ve 0,420 olarak bildirmiştir¹⁰¹. Hyoid kemik segmentasyonuna odaklanan diğer bir çalışmada, Dice katsayısını 0,9439, Intersection over Union (IoU) değerini 0,8941, F1-skorunu 0,9446, %95 Hausdorff mesafesini 1,9998 mm ve EAA 0,98 olarak rapor edilerek modelin yüksek hassasiyet ve doğrulukta çalıştığını vurgulanmıştır¹²³. Infraorbital kanalın otomatik segmentasyonunun değerlendirildiği bir çalışmada ise Dice katsayısı 0,7792, IoU 0,6402, F1-skoru 0,7837, %95 Hausdorff mesafesi 0,7661 mm ve EAA 0,91 değerleri elde edilmiştir¹⁶⁰. Haylaz ve ark.nın nasolakrimal kanal segmentasyonu incelenen bir başka çalışmada Dice katsayısı 0,8465, IoU 0,7341, F1-skoru 0,8480, %95 Hausdorff mesafesi 0,9460 mm ve EAA 0,96 olarak bildirilmiş, bu sonuçların özellikle maksiller sinüs ön duvarı cerrahileri için klinik önem taşıdığı

belirtilmiştir¹²⁴. Paranasal sinüslerin tam otomatik 3B segmentasyonunu hedefleyen bir çalışmada frontal, sfenoid, maksiller ve etmoid sinüsler için Dice katsayıları sırasıyla 0,95, 0,96, 0,97 ve 0,89; Jaccard indeksleri sırasıyla 0,90, 0,92, 0,95 ve 0,80; %95 Hausdorff mesafeleri sırasıyla 0,52 mm, 0,86 mm, 0,51 mm ve 1,18 mm; EAA değerleri sırasıyla 0,97, 0,97, 0,99 ve 0,95 olarak bulunmuş ve tüm sinüslerde doğruluk oranı %99'un üzerinde rapor edilmiştir¹⁶¹. Özudoğru ve ark.nın çocuklarda karma dişlenme dönemine yönelik dişlerin otomatik tespiti ve numaralandırılmasını değerlendirdikleri çalışmasında, doğruluk 0,99, kesinlik 0,86, duyarlılık 0,84, Dice katsayısı 0,81, Jaccard indeksi 0,81 ve %95 Hausdorff mesafesi 1,93 mm olarak bildirilmiştir¹⁶². Yüce ve ark., nazal fossa, maksiller sinüs, nazopalatin kanal, mandibular kanal, foramen mentale ve foramen mandibula olmak üzere altı farklı anatomik yapıyı değerlendirmiş ve her yapı için doğruluk değeri %99 olarak rapor edilmiştir. Bu altı anatomik yapı için Dice katsayısı 0,73 ile 0,98, kesinlik 0,73 ile 0,98, duyarlılık 0,68 ile 0,98, Jaccard indeksi 0,60 ile 0,96 ve %95 Hausdorff mesafesi 0,39 mm ile 6,04 mm aralığında bulunmuştur¹⁶³. Ünal ve Namdar Pekiner mandibular üçüncü molar ile mandibular kanal ilişkisini derin öğrenme ile değerlendirmişlerdir. Buna göre tüm yapılar için doğruluk değeri 0,99 olarak bildirilmiş olup, duyarlılık, kesinlik, Dice katsayısı, Jaccard indeksi ve EAA değerleri, gömülü mandibular üçüncü molar için sırasıyla 0,90, 0,85, 0,85, 0,78 ve 0,80; mandibular kanal için sırasıyla 0,75, 0,78, 0,76, 0,62 ve 0,80; mental foramen için sırasıyla 0,64, 0,66, 0,34, 0,47 ve 0,74 ve mandibular foramen için sırasıyla ise 0,79, 0,68, 0,71 ve 0,57 olarak rapor edilmiştir¹⁶⁴. Gülşen ve ark., sfenoid sinüs ve orta kafa tabanı yapılarının otomatik segmentasyonunu ele aldıkları çalışmada, süperior orbital fissür, sfenoidopalatinal foramen, foramen rotundum, optik kanal, inferior orbital fissür, foramen spinozum, pterigoid (vidian) kanal, foramen ovale ve sfenoid sinüsü incelemişlerdir. Bu anatomik yapılar için Dice katsayısı 0,398 ile 0,964, kesinlik 0,387 ile 0,956, duyarlılık 0,443 ile 0,972, Jaccard indeksi 0,262 ile 0,930 ve %95 Hausdorff mesafesi 0,761 mm ile 10,027 mm aralığında bulunmuştur¹⁶⁵.

Modelin doğruluğu, kesinliği ve duyarlılığı genel performansını değerlendirmek için kullanılan üç temel ölçüttür. Duyarlılık, modelin gerçek pozitif vakaları doğru bir şekilde belirleme yeteneğini ölçerken, kesinlik modelin tahmin edilen pozitif vakaları ne kadar doğru sınıflandırdığını gösterir. Bu çalışmada, modelin doğruluğu, kesinliği, duyarlılığı ve F1-puanı sırasıyla 0,99, 0,74, 0,89 ve 0,81 olarak hesaplanmıştır. Bu

sonular, modelin %74 oranında doęru ngrlerde bulunduęunu ve %89 oranında pozitif sınıfı doęru bir ekilde sınıflandırdıęını ve %81 oranında doęruluk ile kesinlik arasındaki dengenin iyi olduęunu gstermektedir.

Mevcut alıřmada, modelin eęitimi ve performansı hakkında daha fazla bilgi edinmek amacıyla, modelin ilk dngsnden bařlayarak her bir dngde elde edilen ve Dice katsayısı, kayıp fonksiyonu ve Jaccard indeksi eęrisi metrikleri deęerlendirildi. Elde edilen sonular modelin segmentasyon bařarısını belirlemede nemli bir gstergeydi. 0,80 deęerindeki Dice katsayısı, modelin gerek pozitif rneklerle %80 oranında rtřtęn gstermektedir. Azalan bir kayıp fonksiyonu eęrisi, modelin zaman iinde daha iyi ęrendięini ve hata oranının azaldıęını gsterir. Mevcut alıřmadaki kayıp fonksiyonu grafięi, eęitim kaybının dnem sayısına gre nasıl deęiřtięini gstermektedir. Eęitim kaybındaki dzenli azalma, modelin ęrenme srecinin bařarılı bir ekilde ilerledięini gstermektedir. Jaccard indeksi 0,67 olarak hesaplandı ve tahmin edilen alan ile gerek alan arasındaki keřiřim oranının %67 olduęunu gstermektedir. %95 Hausdorff mesafesi ise 1,02 mm olarak hesaplandı ve bu sonu tahmin edilen sınırların, gerek anatomik sınırlarla ok yakından hizalandıęını ve gerek deęerlerle ok byk lde uyumlu olduęunu ortaya koymaktadır.

Bu alıřmada elde edilen kesinlik deęerlerinin dięer metriklere gre grece dřk ıkmasının temel nedeni, modelin yanlış pozitif deęerinin yksek olmasıdır. Kesinlik metrięi, modelin pozitif olarak sınıflandırdıęı rneklerin ne kadarının gerekten pozitif olduęunu gsterdięinden, yanlış pozitif sayısındaki artıř bu deęeri doęrudan dřrmektedir. Model, bazı rneklerde kanal sınırlarını evreleyen dokuların yoęunluk benzerlięi nedeniyle bu blgeleri de kanalın bir parası olarak yanlış biimde etiketlemiř olabilir. Bu durum zellikle sınır kontrastının zayıf olduęu ve evre dokuların gri yoęunluęunun kanala yakın deęerlere sahip olduęu vakalardan kaynaklanabilir. Bu bulgu, kk anatomik yapıların segmentasyonunda komřu dokuların yoęunluk geiřlerinin model aısından hata kaynaęı olabileceęini dřndrmektedir. Hasta bazlı metriklerin yer aldıęı tabloda da benzer bir eęilim izlenmiřtir. rneęin test setindeki 6 nolu hastada yanlış pozitif oranı yksek olduęu iin kesinlik deęeri belirgin biimde dřmřtir. Buna karřın aynı vakada duyarlılık ve Dice katsayısı deęerleri yksek kalmıřtır; bu durum modelin gerek kanal blgelerinin tamamını doęru biimde

yakaladığı, ancak fazladan alanları da pozitif olarak sınıflandırdığı anlamına gelir. Diğer vakalarda kesinlik ile birlikte Dice değerinde de düşüş gözlenmiş, bu da yanlış pozitifliğin yalnızca kesinliği değil genel benzerlik skorunu da olumsuz etkilediğini göstermektedir. kesinlik ile F1-skorları arasında güçlü bir ilişki bulunmaktadır; çünkü F1-skoru kesinlik ve duyarlılığın harmonik ortalaması olarak tanımlanır. Bu nedenle kesinlik değerinin düşmesi, F1-skorunun da doğrudan azalmasına neden olur. Ancak duyarlılık değerinin yüksekliği, F1-skorundaki düşüşü kısmen dengelemiştir. Her iki metrik de kesişim oranına dayalı benzerliği ölçmekte ve küçük farklılıklar dışında aynı eğilimi göstermektedir. Bu çalışmada da F1-skoru ve Dice katsayısı değerleri paralel değişim göstermiş, bu durum modelin genel segmentasyon performansının istikrarlı olduğunu desteklemiştir. Bazı olgularda kesinlik değerleri düşük olmasına rağmen, doğruluk değerleri çok yüksek (0.99) bulunmuştur. Modelin büyük çoğunluğunun gerçek negatif örneklerden oluşan vokselleri doğru sınıflandırması, toplam doğruluk oranını yüksek tutmuştur. Ancak pozitif sınıfa ait vokselle sayısının görece az olması, yanlış pozitiflerin kesinlik değerleri üzerinde düşürücü bir etkiye yol açmıştır. Bu nedenle doğruluk değeri, her zaman tek başına güvenilir bir performans göstergesi olmayıp, tüm performans metriklerinin birlikte değerlendirilmesi, modelin gerçek ayırt edici başarısını daha doğru yansıtabilmektedir.

EAA ise 0,94 olarak hesaplanmıştır. ROC eğrisi, bir sınıflandırma modelinin performansını değerlendirmek için kullanılan önemli bir araç olup, gerçek pozitif oranı ve yanlış pozitif oranının grafiğe dökülmesiyle elde edilir ve bu sayede duyarlılık ile özgüllük arasındaki ilişkiyi yansıtır. EAA değeri, bu eğrinin altında kalan alanı temsil eder ve modelin ayırım gücünü gösterir. 0,94 EAA değerine sahip bir model iyi bir performansa sahip olarak yorumlanır. Bu değer, modelin rastgele tahminler yapan bir sistemden ($EAA = 0,5$) daha başarılı olduğunu ve genellikle yüksek bir doğruluk oranına sahip olduğunu gösterir. EAA değeri 1'e yaklaştıkça, modelin pozitif ve negatif sınıfları ayırt etme yeteneği artar. Bu nedenle, 0,94 gibi yüksek bir EAA değeri, modelin yanlış sınıflandırmaları en aza indirdiğini ve doğru tahminler yaparak güvenilir bir sınıflandırma sağladığını gösterir.

Bu değerler, semantik segmentasyon sonucu elde edilen çıktı kümesi ile gözlemci tarafından oluşturulan ve gerçek referans (ground-truth) olarak kabul edilen maske

bölgelerinin üst üste getirilmesi sonucunda, ilgili alanlardaki piksel sayılarının hesaplanmasıyla belirlenmiştir.

Bu çalışmada kullanılan hiperparametre değerleri (epoch sayısı, öğrenme oranı, batch size vb.) literatürde benzer derin öğrenme tabanlı tıbbi görüntü segmentasyon çalışmalarında sıklıkla kullanılan parametre aralıkları dikkate alınarak belirlenmiştir. Modelin tüm eğitim veri setini kaç kez baştan sona işleyeceği anlamına gelen epoch sayısı, mevcut çalışmada 1000 olarak belirlenmiştir. Küçük veri kümelerinde modelin yeterli genelleme kapasitesine ulaşabilmesi için yüksek epoch sayısının tercih edilir. Küçük örneklemli veri setlerinde düşük epoch değerleri modelin öğrenme sürecini sınırlayarak erken durdurma riskini artırabilirken, yüksek epoch sayıları modelin veri örüntülerini daha kapsamlı biçimde öğrenmesine olanak tanımaktadır. Ayrıca, eğitim süreci boyunca kayıp fonksiyonunun stabil hale gelmesi ve Dice katsayısının plato eğrisine ulaşması, modelin aşırı uyum göstermeden doyumluğa ulaştığını göstermiştir. Bu nedenle epoch sayısının artırılması, özellikle validasyon seti ayrılmamış küçük örneklemli çalışmalarda, modelin genelleme kapasitesini destekleyen uygun bir strateji olarak değerlendirilmiştir^{79,166}. Epoch sayısı, modelin öğrenme eğrisinin (loss fonksiyonu) stabil hale geldiği nokta dikkate alınarak 1000 olarak sabitlenmiştir. Öğrenme oranı, modelin erken evrede hızlı ancak kararlı şekilde öğrenmesini sağlamak amacıyla 0.01 olarak seçilmiştir. Bu değer, literatürdeki benzer derin öğrenme tabanlı segmentasyon çalışmalarında kullanılan değerden yüksektir. Bunun temel nedeni, mevcut veri setinin sınırlı büyüklüğe sahip olması ve modelin öğrenme sürecinde erken evrede daha hızlı yakınsama sağlamasının hedeflenmesidir. Düşük öğrenme oranı, özellikle az sayıda örnek içeren veri kümelerinde modelin daha uzun sürede öğrenmesine ve yerel minimumlara takılmasına neden olabilmektedir. Yüksek öğrenme oranı başlangıçta daha geniş bir ağırlık güncellemesi sağlayarak ağırlık genel eğilimini daha hızlı yakalamasına olanak tanır. Bununla birlikte, aşırı yüksek değerlerin kararsızlığa yol açmaması için eğitim sırasında kayıp (loss) eğrisi dikkatle izlenmiş ve 1000 epoch sonunda modelin hata oranının stabil hale geldiği gözlemlenmiştir. Batch size değeri ise sistemin donanımsal kısıtları ve bellek kapasitesi göz önünde bulundurularak 2 olarak belirlenmiştir. Bu parametreler, nnU-Net v2 mimarisinde önerilen varsayılan değerlere ve literatürde raporlanan benzer çalışmaların ayarlarına paralellik göstermektedir^{101,123,124,160,161}.

Adam, derin öğrenmede en yaygın kullanılan optimizasyon algoritmalarından biridir⁸⁶. Derin öğrenmede kullanılan optimizasyon algoritmalarının karşılaştırıldığı bir çalışmada uygulama sonucunda optimizasyon algoritmalarının sınıflamadaki başarı oranları sgd için %64.5, momentum için %85.56, adam için %92.31, adagrad için %71.25, rmsprop için %40.26 ve adadelata için %86.88 olarak hesaplanmıştır⁸⁵. Adam adı, uyarlanabilir moment kestirimi (adaptive moment estimation) ifadesinden türetilmiştir çünkü gradyanların birinci ve ikinci momentlerinin tahminlerinden farklı parametreler için bireysel uyarlanabilir öğrenme oranlarını hesaplar. Adam, seyrek gradyanlarda iyi çalışan AdaGrad'in ve çevrimiçi ile durağan olmayan ortamlarda etkili olan RMSProp'un avantajlarını birleştirir. Adam'ın birçok avantajı vardır. Her şeyden önce, öğrenme oranı için çok az ayarlama gerektirir. Ayrıca, uygulanması kolay ve gradyanların diyagonal yeniden ölçeklendirmesine karşı değişmez bir yöntemdir. Adam durağan olmayan hedefler ve oldukça gürültülü ya da seyrek gradyanlara sahip problemler için uygundur. Bunun yanında, hem hesaplama açısından verimli, hem de çok az bellek gerektirir ⁸⁶⁻⁸⁸. Tüm bu sebeplere dayanarak ve ayrıca literatürdeki nnU-net modelinin kullanıldığı birçok segmentasyon çalışmasında da olduğu gibi^{101,123,124,160}, bu çalışmada optimizasyon fonksiyonu olarak Adam algoritması seçilmiştir.

Leaky ReLU, en başarılı aktivasyon fonksiyonu olduğu düşünülen⁸⁵ ve ayrıca literatürdeki nnU-net modelinin kullanıldığı birçok segmentasyon çalışmasında da kullanılan ReLU fonksiyonunun^{101,123,124,160} negatif değerleri tamamen sıfıra indirerek nöronları “ölü” hale getirme problemini çözmek için geliştirilmiştir. Bu sayede türev hiçbir zaman sıfır olmaz ve gradyan akışı tamamen kesilmez. Böylece nöronlar öğrenmeye devam edebilir. Leaky ReLU'nun bazı durumlarda, klasik ReLU'dan daha kararlı ve verimli sonuçlar verdiği belirtilmiştir. Leaky ReLU'nun “Dead Neuron” problemini büyük ölçüde azaltması, gradyan sönmesi riskini düşürmesi, negatif bölgede türev $\neq 0$ olduğu için öğrenmenin tamamen durmaması, özellikle yüksek öğrenme oranlarında, ReLU'ya kıyasla daha iyi performans gösterebilmesi gibi avantajlarından dolayı¹⁶⁷⁻¹⁶⁹, bu çalışmada aktivasyon fonksiyonu olarak Leaky ReLU seçilmiştir. Gelecekteki çalışmalarda daha geniş veri setleri ve farklı parametre kombinasyonlarıyla hiperparametre optimizasyonu yapılması modelin performansını daha da iyileştirebilir.

Bu çalışmada ayrı bir validasyon seti ve çapraz doğrulama yöntemi uygulanmamıştır. Bu durum teorik olarak derin öğrenme modellerinde hiperparametre optimizasyonu ve modelin aşırı uyum göstermemesi açısından önemli olmakla birlikte, mevcut modelin performans metrikleri incelendiğinde, test seti üzerinde elde edilen yüksek Dice katsayısı ve düşük Hausdorff mesafesi değerleri, modelin yalnızca eğitim verisine özgü bir öğrenme göstermediğini ve genelleme kapasitesinin yeterli olduğunu düşündürmektedir. Ayrıca eğitim süreci boyunca kayıp fonksiyonunun stabil hale gelmesi, modelin doyumluğa ulaştığını ve öğrenme sürecinin aşırı uyuma gitmeden dengelendiğini göstermektedir. Büyük olmayan örneklem büyüklüğüne sahip veri setlerinde validasyon veya çapraz doğrulama yapılmaması modelin genellenebilirliğini kısmen sınırlandırırsa da, bu yöntem eğitim verisi miktarını maksimum düzeyde tutarak öğrenme kapasitesini artırmayı hedeflemiş ve eğitim için ayrılacak veri miktarının azalması önleyerek modelin öğrenme kapasitesinin sınırlanabilmesinin önüne geçilmek hedeflenmiştir. Bu nedenle mevcut veri kümesinin tamamı modelin eğitiminde ve testinde kullanılmıştır. Gelecekteki çalışmalarda daha geniş veri kümeleri ile validasyon ve çapraz doğrulama yöntemlerinin birlikte uygulanması, modelin genellenebilirliğini ve performans güvenilirliğini güçlendirecektir.

6. SONUÇLAR VE ÖNERİLER

Bu çalışma, NPK'nın yapay zeka modeli ile segmente edilmesinde önemli bulgular ortaya koysa da, bazı kısıtlılıkların dikkate alınması gerekir. KIBT görüntüleri tek bir cihazla, tek bir merkezde ve sabit görüntüleme parametreleri altında elde edilmiştir. Bu durum, modelin genellenebilirliğini sınırlandırabilir. Ayrıca, modelin gerçek klinik ortamlardaki etkinliğinin değerlendirilmesi için ek doğrulama çalışmaları gereklidir ve farklı cihazlardan elde edilmiş çok merkezli veri setlerinde test edilmesi genellenebilirliği artıracaktır.

Gelecek çalışmalarda, hasta sayısının artırılması ve farklı populasyonlarda uygulanabilirliğin sınanması önerilmektedir. Ayrıca eğitim sürecinde validasyon setinin ayrılması ve çapraz doğrulama yöntemlerinin eklenmesi, modelin aşırı uyum riskini azaltacaktır. Veri azlığı sorununu aşmak amacıyla farklı veri artırma tekniklerinin kullanılması ve farklı derin öğrenme mimarilerinin performanslarının karşılaştırılması da ileriki araştırmalar için değerli olacaktır.

Sonuç olarak bu kapsamlı performans metrikleri, modelin genel doğruluğunu, kesinliğini ve genelleme yeteneğini değerlendirmek için kullanıldı ve elde edilen değerler nnU-Net v2 modelinin başarılı bir performans sergilediğini, NPK'yı segmente etme açısından klinik uygulamalarda kullanılabilecek güvenilir, etkili ve kararlı bir yapay zeka aracı olduğunu-göstermektedir.

KAYNAKLAR

1. **Alexander B, John S, Aralamoodu P.** Artificial intelligence in dentistry: current concepts and apeek into the future. *Int J Adv Res (Indore)*, **2018**; 6:1105-1108.
2. **Gezer N.** Interstitial Lung Disease, Quantitative CT Analysis and Artificial Intelligence Applications, Radiomics. *Turk Radiol Semin*, **2024**; 12:162-176.
3. **Sharma N, Ray AK, Shukla KK, Sharma S, Pradhan S, Srivastva A, et al.** Automated medical image segmentation techniques. *J Med Phys*, **2010**; 35(1):3-14.
4. **Ermiş E, Jungo A, Poel R, Blatti-Moreno M, Meier R, Knecht U, et al.** Fully automated brain resection cavity delineation for radiation target volume definition in glioblastoma patients using deep learning. *Radiation Oncology*, **2020**; 15(1).
5. **Thomson D, Boylan C, Liptrot T, Aitkenhead A, Lee L, Yap B, et al.** Evaluation of an automatic segmentation algorithm for definition of head and neck organs at risk. *Radiation Oncology*, **2014**; 9(1).
6. **Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O.** 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI*, **2016**; 9901(Springer):424-432.
7. **Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH.** nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods*, **2021**; 18(2):203-211.
8. **Ronneberger O, Fischer P, Brox T.** U-net: Convolutional networks for biomedical image segmentation. *Springer O Ronneberger, P Fischer, T Brox International Conference on Medical image computing and computer-assisted*, **2015**; 9351:234-241.
9. **Celaya A, Diaz A, Balsells A, Riviere B, Fuentes D.** A Weighted Normalized Boundary Loss for Reducing the Hausdorff Distance in Medical Imaging Segmentation. *arXiv preprint arXiv*, **2023**; 2302.03868.
10. **Setzer FC, Shi KJ, Zhang Z, Yan H, Yoon H, Mupparapu M, et al.** Artificial Intelligence for the Computer-aided Detection of Periapical Lesions in Cone-beam Computed Tomographic Images. *J Endod*, **2020**; 46(7):987-993.
11. **Santos GNM, da Silva HEC, Ossege FEL, de Souza Figueiredo PT, de Santos Melo N, Stefani CM, et al.** Radiomics in bone pathology of the jaws. *Dentomaxillofacial Radiology*, **2023**; 52(1).
12. **Lee JH, Kim DH, Jeong SN, Choi SH.** Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. *J Periodontal Implant Sci*, **2018**; 48(2):114-123.
13. **Lee JH, Kim DH, Jeong SN, Choi SH.** Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J Dent*, **2018**; 77:106-111.
14. **Duyan Yüksel H, Orhan K, Evlice B, Kaya Ö.** Evaluation of temporomandibular joint disc displacement with MRI-based radiomics analysis. *Dentomaxillofac Radiol*, **2025**; 54(1):19-27.

15. **Polat S, Öksüzler FY, Öksüzler M, Çoban Ö, Tunç M, Yüksel HD, et al.** Temporomandibular joint and masticatory muscles morphometry and morphology in healthy subjects and individuals with temporomandibular dysfunction: An anatomical, radiological, and machine learning application study. *Medicine (United States)*, **2024**; 103(50):e40846.
16. **Yadav U, Shenoy N, Ahmed J, Sujir N, M A, Gupta A.** Assessment of variations in the nasopalatine canal on CBCT: considerations from an anatomical point of view. *J Periodontal Implant Sci*, **2025**; 55(1).
17. **Lam E.** Cysts. In: Mallya S, Lam E, eds. *White and Pharoah's Oral Radiology: Principles and Interpretation*. 8th ed. Mosby; 2018:400-402.
18. **Mardinger O, Namani-Sadan N, Chaushu G, Schwartz-Arad D.** Morphologic changes of the nasopalatine canal related to dental implantation: a radiologic study in different degrees of absorbed maxillae. *J Periodontol*, **2008**; 79(9):1659-1662.
19. **Liang X, Jacobs R, Martens W, Hu Y, Adriaenssens P, Quirynen M, et al.** Macro- and micro-anatomical, histological and computed tomography scan characterization of the nasopalatine canal. *J Clin Periodontol*, **2009**; 36(7):598-603.
20. **Raghoobar GM, Hartog L Den, Vissink A.** Augmentation in proximity to the incisive foramen to allow placement of endosseous implants: A case series. *Journal of Oral and Maxillofacial Surgery*, **2010**; 68(9):2267-2271.
21. **Thakur AR, Burde K, Guttal K, Naikmasur VG.** Anatomy and morphology of the nasopalatine canal using cone-beam computed tomography. *Imaging Sci Dent*, **2013**; 43(4):273-281.
22. **Choudhury IR, Rajan K.** Have conventional dental radiographs lost its charm to modern techniques? *J Dent Spec*, **2021**; 9:3-6.
23. **Stratemann SA, Huang JC, Maki K, Hatcher DC, Miller AJ.** Evaluating the mandible with cone-beam computed tomography. *American Journal of Orthodontics and Dentofacial Orthopedics*, **2010**; 137(4 SUPPL.).
24. **Lv J, Zhang L, Xu J, Li W, Li G, Zhou H.** Automatic segmentation of mandibular canal using transformer based neural networks. *Front Bioeng Biotechnol*, **2023**; 11.
25. **Abdolali F, Zoroofi RA, Abdolali M, Yokota F, Otake Y, Sato Y.** Automatic segmentation of mandibular canal in cone beam CT images using conditional statistical shape model and fast marching. *Int J Comput Assist Radiol Surg*, **2017**; 12(4):581-593.
26. **Tanaka S, Iida S, Murakami S, Kishino M, Yamada C, Okura M.** Extensive nasopalatine duct cyst causing nasolabial protrusion. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*, **2008**; 106(4).
27. **Cicciù M, Battista Grossi G, Borgonovo A, Santoro G, Pallotti F, Maiorana C.** Rare bilateral nasopalatine duct cysts: a case report. *Open Dent J*, **2010**; 4(1):8-12.
28. **Tyndall DA, Brooks SL.** Selection criteria for dental implant site imaging: a position paper of the American Academy of Oral and Maxillofacial radiology. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*, **2000**; 89(5):630-637.

29. **Asaumi R, Kawai T, Sato I, Yoshida S, Yosue T.** Three-dimensional observations of the incisive canal and the surrounding bone using cone-beam computed tomography. *Oral Radiol*, **2010**; 26(1):20-28.
30. **Chatriyanuyoke P, Lu CI, Suzuki Y, Lozada JL, Rungcharassaeng K, Kan JYK, et al.** Nasopalatine canal position relative to the maxillary central incisors: a cone beam computed tomography assessment. *J Oral Implantol*, **2012**; 38(6):713-717.
31. **Panjnoush M, Norouzi H, Kheirandish Y, Shamshiri AR, Mofidi N.** Evaluation of Morphology and Anatomical Measurement of Nasopalatine Canal Using Cone Beam Computed Tomography. *J Dent (Tehran)*, **2016**; 13(4):287-294.
32. **White S, Pharoah M.** Oral radiology: Principles and interpretation. *Elsevier Health Sciences*, Published online 2014.
33. **Harorlı A, Akgül H, Yılmaz A, Bilge O.** Ağız, Diş ve Çene Radyolojisi. *İstanbul: Nobel Tıp Kitabevleri*, Published online 2014.
34. **Scarfe W, Farman A, Sukovic P.** Clinical applications of cone-beam computed tomography in dental practice. *J Can Dent Assoc*, **2006**; 72(1):75-80.
35. **Brook I.** Sinusitis of odontogenic origin. *Otolaryngol Head Neck Surg*, **2006**; 135(3):349-355.
36. **Hauman CHJ, Chandler NP, Tong DC.** Endodontic implications of the maxillary sinus: a review. *Int Endod J*, **2002**; 35(2):127-141.
37. **Khanagar SB, Al-ehaideb A, Maganur PC, Vishwanathaiah S, Patil S, Baeshen HA, et al.** Developments, application, and performance of artificial intelligence in dentistry – A systematic review. *J Dent Sci*, **2021**; 16(1):508-522.
38. **Turing AM. I.**—Computing machinery and intelligence. *Mind*, **1950**; LIX(236):433-460.
39. **Kaul V, Enslin S, Gross SA.** History of artificial intelligence in medicine. *Gastrointest Endosc*, **2020**; 92(4):807-812.
40. **Ding H, Wu J, Zhao W, Matinlinna JP, Burrow MF, Tsoi JKH.** Artificial intelligence in dentistry—A review. *Frontiers in Dental Medicine*, **2023**; 4.
41. **Campbell M, Hoane AJ, Hsu FH.** Deep Blue. *Artif Intell*, **2002**; 134(1-2):57-83.
42. **Krizhevsky A, Sutskever I, Hinton GE.** ImageNet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst*, **2012**; 25.
43. **Shabbir J, Anwer T.** Artificial Intelligence and its Role in Near Future. *arXiv preprint arXiv*, **2018**; 1804.01396.
44. **Pandl K, Thiebes S, Schmidt-Kraepelin M, Sunyaev A.** On the convergence of artificial intelligence and distributed ledger technology: A scoping review and future research agenda. *IEEE Access*, **2020**; 8:57075-57095.
45. **Amoroso N, Rocca M La, Bellantuono L, Diacono D, Fanizzi A, Lella E, et al.** Deep learning and multiplex networks for accurate modeling of brain age. *Front Aging Neurosci*, **2019**; 11(MAY).
46. **Lopez Torres E, Fiorina E, Pennazio F, Peroni C, Saletta M, Camarlinghi N, et al.** Large scale validation of the M5L lung CAD on heterogeneous CT datasets. *Med Phys*, **2015**; 42(4):1477-1489.

47. **Militello C, Vitabile S, Rundo L, Russo G, Midiri M, Gilardi MC.** A fully automatic 2D segmentation method for uterine fibroid in MRgFUS treatment evaluation. *Comput Biol Med*, **2015**; 62:277-292.
48. **Simeone O.** A brief introduction to machine learning for engineers. *Foundations and Trends in Signal Processing*, **2018**; 12(3-4):200-431.
49. **Chauhan N, Singh K.** A review on conventional machine learning vs deep learning. *International conference on computing, power and communication technologies (GUCON)*, **2018**; IEEE(2018):347-352.
50. **Kufel J, Bargiel-Lączek K, Kocot S, Koźlik M, Bartnikowska W, Janik M, et al.** What Is Machine Learning, Artificial Neural Networks and Deep Learning?—Examples of Practical Applications in Medicine. *Diagnostics*, **2023**; 13(15).
51. **Khanagar SB, Al-chaideb A, Maganur PC, Vishwanathaiah S, Patil S, Baeshen HA, et al.** Developments, application, and performance of artificial intelligence in dentistry – A systematic review. *J Dent Sci*, **2021**; 16(1):508-522.
52. **Arsiwala-Scheppach LT, Chaurasia A, Müller A, Krois J, Schwendicke F.** Machine Learning in Dentistry: A Scoping Review. *J Clin Med*, **2023**; 12(3).
53. **Bishop C, Nasrabadi N.** Pattern recognition and machine learning. Published online 2006.
54. **Ongsulee P.** Artificial intelligence, machine learning and deep learning. *15th international conference on ICT and knowledge engineering (ICT&KE)*, **2017**; IEEE.
55. **Skansi S.** Introduction to Deep Learning: from logical calculus to artificial intelligence. Published online 2018.
56. **Sen D, Chakrabarti R, Chatterjee S, Grewal DS, Manrai K.** Artificial intelligence and the radiologist: the future in the Armed Forces Medical Services. *J R Army Med Corps*, **2020**; 166(4):254-256.
57. **Kohli M, Prevedello LM, Filice RW, Geis JR.** Implementing machine learning in radiology practice and research. *American Journal of Roentgenology*, **2017**; 208(4):754-760.
58. **Du X, Cai Y, Wang S, Zhang L.** Overview of deep learning. *31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, **2016**; IEEE(2016):159-164.
59. **Najafabadi MM, Villanustre F, Khoshgoftaar TM, Seliya N, Wald R, Muharemagic E.** Deep learning applications and challenges in big data analytics. *J Big Data*, **2015**; 2:1-21.
60. **Suzuki K.** Overview of deep learning in medical imaging. *Radiol Phys Technol*, **2017**; 10(3):257-273.
61. **Bengio Y.** Learning deep architectures for AI. *Foundations and Trends® in Machine Learning*, **2009**; 2(1):1-127.
62. **Soufiene B, Chakraborty C.** *Machine Learning and Deep Learning Techniques for Medical Image Recognition*. CRC Press Taylor & Francis Group; 2023.
63. **Schmidhuber J.** Deep Learning in neural networks: An overview. *Neural Networks*, **2015**; 61:85-117.

64. **Guo Y, Liu Y, Oerlemans A, Lao S, Wu S, Lew MS.** Deep learning for visual understanding: A review. *Neurocomputing*, **2016**; 187:27-48.
65. **Zeiler MD, Fergus R.** Visualizing and Understanding Convolutional Networks. *Computer Vision – ECCV, Springer*, Published online 2014:818-833.
66. **Krizhevsky A, Sutskever I, Hinton GE.** Imagenet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*. 2012:1097-1105.
67. **Simonyan K, Zisserman A.** Very deep convolutional networks for largescale image recognition. *arXiv preprint*, **2014**; 1409:1556.
68. **Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al.** Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Published online 2015:1-9.
69. **Girshick R, Donahue J, Darrell T, Malik J.** Rich feature hierarchies for accurate object detection and semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014:580-587.
70. **Gao Y, Rong W, Shen Y, Xiong Z.** Convolutional neural network based sentiment analysis using adaboost combination. In: *Neural Networks (IJCNN), International Joint Conference*. 2016:1333-1338.
71. **Long J, Shelhamer E, Darrell T.** Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Published online 2015:3431-3440.
72. **Balthazar P, Harri P, Prater A, Safdar NM.** Protecting Your Patients' Interests in the Era of Big Data, Artificial Intelligence, and Predictive Analytics. *Journal of the American College of Radiology*, **2018**; 15(3):580-586.
73. **Krizhevsky A, Sutskever I, Hinton GE.** ImageNet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst*, **2012**; 25(6):1097-1105.
74. **Rieke N, Hancox J, Li W, Milletari F, Roth HR, Albarqouni S, et al.** The future of digital health with federated learning. *NPJ Digit Med*, **2020**; 3(1).
75. **Jin H.** Hyperparameter Importance for Machine Learning Algorithms. *arXiv preprint*, **2022**; 05132.
76. **Ippolito PP.** Hyperparameter Tuning. In: Egger R, ed. *Applied Data Science in Tourism. Tourism on the Verge*. Vol Part F1051. Springer, Cham; 2022:231-251.
77. **Snoek J, Larochelle H, Adams RP.** Practical bayesian optimization of machine learning algorithms. *Adv Neural Inf Process Syst*, **2012**; 25.
78. **Li L, Jamieson K, Rostamizadeh A, Gonina E, Ben-Tzur J, Hardt M, et al.** A system for massively parallel hyperparameter tuning. *Proceedings of Machine Learning Systems*, **2020**; 2:230-246.
79. **Shafi S, Assad A.** Exploring the Relationship Between Learning Rate, Batch Size, and Epochs in Deep Learning: An Experimental Study. In: *Lecture Notes in Networks and Systems*. Vol 547. Springer, Singapore; 2023:201-209.

80. **Masters D, Luschi C.** Revisiting small batch training for deep neural networks. *arXiv preprint arXiv*, **2018**; 07612.
81. **Iiduka H.** Appropriate Learning Rates of Adaptive Learning Rate Optimization Algorithms for Training Deep Neural Networks. *IEEE Trans Cybern*, **2022**; 52(12):13250-13261.
82. **Lewkowycz A, Bahri Y, Dyer E, Sohl-Dickstein J, Gur-Ari G.** The large learning rate phase of deep learning: the catapult mechanism. *arXiv preprint arXiv*, **2020**; 2003:02218.
83. **Li Z, Arora S.** An Exponential Learning Rate Schedule for Deep Learning. *arXiv preprint arXiv*, **2019**; 1910:07454.
84. **Ruder S.** An overview of gradient descent optimization algorithms. *arXiv preprint arXiv*, Published online 2016:04747.
85. **Seyyarer E, Ayata F, Uçkan T, Karci A.** Derin öğrenmede kullanılan optimizasyon algoritmalarının uygulanması ve kıyaslanması. *Computer Science*, **2020**; 5(2):90-98.
86. **Soydaner D.** Derin öğrenme için optimizasyon algoritmalarının karşılaştırılması. *Uluslararası Desen Tanıma ve Yapay Zeka Dergisi*, **2020**; 34(13):2052013.
87. **Goodfellow I, Bengio Y, Courville A.** Deep Learning. *The MIT Press*, Published online 2016.
88. **Kingma D, Ba J.** Adam: A method for stochastic optimization. *arXiv preprint arXiv*, **2014**; 1412(6980).
89. **Dubey SR, Singh SK, Chaudhuri BB.** Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing*, **2022**; 503:92-108.
90. **Agostinelli F, Hoffman M, Sadowski P, Baldi P.** Learning Activation Functions to Improve Deep Neural Networks. *arXiv preprint arXiv: 3rd International Conference on Learning Representations, ICLR 2015 - Workshop Track Proceedings*, **2014**; 1412:6830.
91. **Kılıçarslan S, Adem K, Çelik M.** An overview of the activation functions used in deep learning algorithms. *Journal of New Results in Science*, **2021**; 10(3):75-88.
92. **Dugas C, Bengio Y, Bélisle F, Nadeau C, Garcia R.** Incorporating Second-Order Functional Knowledge for Better Option Pricing. *Adv Neural Inf Process Syst*, **2001**; 13:472-478.
93. **Misra D.** Mish: A self-regularized non-monotonic activation function. *arXiv preprint arXiv*, **2019**; 1908:08681.
94. **Cai L, Gao J, Zhao D.** A review of the application of deep learning in medical image classification and segmentation. *Ann Transl Med*, **2020**; 8(11):713-713.
95. **Miranda E, Aryuni M, Irwansyah E.** A survey of medical image classification techniques. *International Conference on Information Management and Technology (ICIMTech)*, Published online 2016.
96. **Garcia-Garcia A, Orts-Escolano S, Oprea S, Villena-Martinez V, Martinez-Gonzalez P, Garcia-Rodriguez J.** A survey on deep learning techniques for image and video semantic segmentation. *Appl Soft Comput*, **2018**; 70:41-65.

97. **Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al.** Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans Med Imaging*, **2016**; 35(5):1285-1298.
98. **Fabijańska A.** Segmentation of corneal endothelium images using a U-Net-based convolutional neural network. *Artif Intell Med*, **2018**; 88:1-13.
99. **Sevastopolsky A.** Optic disc and cup segmentation methods for glaucoma detection with modification of U-Net convolutional neural network. *Pattern Recognition and Image Analysis*, **2017**; 27(3):618-624.
100. **Du G, Cao X, Liang J, Chen X, Zhan Y.** Medical image segmentation based on U-Net: A review. *Journal of Imaging Science and Technology*, **2020**; 64(2):1.
101. **Gumusoy I, Demirezer K, Duman SB, Haylaz E, Bayrakdar IS, Celik O, et al.** AI-powered segmentation of bifid mandibular canals using CBCT. *BMC Oral Health*, **2025**; 25(1):1-9.
102. **Terven J, Cordova-Esparza DM, Ramirez-Pedraza A, Chavez-Urbiola EA, Romero-Gonzalez JA.** Loss Functions and Metrics in Deep Learning. *arXiv preprint arXiv*, **2023**; 2307.02694.
103. **Hossin M, Sulaiman MN.** A review on evaluation metrics for data classification evaluations. *International journal of data mining knowledge management process*, **2015**; 5(2):1.
104. **Bhat S, Rehman A, Abulaish M.** *Deep Learning Applications in Medical Image Segmentation: Overview, Approaches, and Challenges.* (Spurgeon S, ed.). IEEE Press Wiley; 2025.
105. **Rodriguez R, Kannan H, Shaikh K, Bekal S.** Deep learning approaches for early diagnosis of neurodegenerative diseases. *IGI Global Publisher of Timely Knowledge*, Published online 2024:1-325.
106. **Ray D, Pinti O, Oberai A.** *Deep Learning and Computational Physics.* Springer Nature Switzerland; 2024.
107. **Singh N, Kose U, Gochhayat S.** *Deep Learning in Biomedical Signal and Medical Imaging.* CRC Press Taylor & Francis Group; 2024.
108. **Kılıç S.** Klinik karar vermede ROC analizi. *Journal of Mood Disorders*, **2013**; 3(3):135-140.
109. **Keçeoğlu Ç.** Kesme puanının belirlenmesinde lojistik regresyon analizi, diskriminant analizi ve ROC eğrisi yöntemlerinin karşılaştırılması. *Yüksek Lisans Tezi, Hacettepe Üniversitesi Ankara*, Published online 2012.
110. **Koyuncu M.** Psikolojik ölçeklerde Roc analizi yöntemiyle standart belirleme. *Yüksek lisans tezi, Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara*, Published online 2015.
111. **Greiner M, Pfeiffer D, Smith RD.** Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Prev Vet Med*, **2000**; 45(1-2):23-41.
112. **Zweig MH, Campbell G.** Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem*, **1993**; 39(4):561-577.
113. **Metz CE.** Basic principles of ROC analysis. *Semin Nucl Med*, **1978**; 8(4):283-298.
114. **Fawcett T.** An introduction to ROC analysis. *Pattern Recognit Lett*, **2006**; 27(8):861-874.

115. **Krzanowski WJ, Hand DJ.** ROC curves for continuous data. *ROC Curves for Continuous Data*, Published online May 21, 2009:1-241.
116. **Muschelli J.** ROC and AUC with a Binary Predictor: a Potentially Misleading Metric. *J Classif*, **2020**; 37(3):696-708.
117. **Hanley JA, McNeil BJ.** A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*, **1983**; 148(3):839-843.
118. **Obuchowski NA.** Receiver operating characteristic curves and their use in radiology. *Radiology*, **2003**; 229(1):3-8.
119. **Bamber D.** The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *J Math Psychol*, **1975**; 12(4):387-415.
120. **Hanley JA, McNeil BJ.** The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, **1982**; 143(1):29-36.
121. **Çorbacioğlu ŞK, Aksel G.** Receiver operating characteristic curve analysis in diagnostic accuracy studies: A guide to interpreting the area under the curve value. *Turk J Emerg Med*, **2023**; 23(4):195.
122. **Bayrakdar IS, Elfayome NS, Hussien RA, Gulsen IT, Kuran A, Gunes I, et al.** Artificial intelligence system for automatic maxillary sinus segmentation on cone beam computed tomography images. *Dentomaxillofacial Radiology*, **2024**; 53(4):256-266.
123. **Gümüşsoy I, Haylaz E, Duman SB, Kalabalik F, Say S, Celik O, et al.** Automated 3D segmentation of the hyoid bone in CBCT using nnU-Net v2: a retrospective study on model performance and potential clinical utility. *BMC Med Imaging*, **2025**; 25(1).
124. **Haylaz E, Gumussoy I, Duman SB, Kalabalik F, Eren MC, Demirsoy MS, et al.** Automatic Segmentation of the Nasolacrimal Canal: Application of the nnU-Net v2 Model in CBCT Imaging. *J Clin Med*, **2025**; 14(3).
125. **Ito K, Hirahara N, Muraoka H, Sawada E, Tokunaga S, Komatsu T, et al.** Graphical user interface-based convolutional neural network models for detecting nasopalatine duct cysts using panoramic radiography. *Sci Rep*, **2024**; 14(1):7699.
126. **Kise Y, Kuwada C, Mori M, Fukuda M, Ariji Y, Ariji E.** Deep learning system for distinguishing between nasopalatine duct cysts and radicular cysts arising in the midline region of the anterior maxilla on panoramic radiographs. *Imaging Sci Dent*, **2024**; 54(1):33-41.
127. **Lee HS, Yang S, Han JY, Kang JH, Kim JE, Huh KH, et al.** Automatic detection and classification of nasopalatine duct cyst and periapical cyst on panoramic radiographs using deep convolutional neural networks. *Oral Surg Oral Med Oral Pathol Oral Radiol*, **2024**; 138(1):184-195.
128. **Deniz HA, Bayrakdar İŞ, Nalçacı R, Orhan K.** Segmentation of the nasopalatine canal and detection of canal furcation status with artificial intelligence on cone-beam computed tomography images. *Oral Radiol*, **2025**; 41(3):403-413.
129. **Xu X, Zhou F, Liu B.** Automatic bladder segmentation from CT images using deep CNN and 3D fully connected CRF-RNN. *Int J Comput Assist Radiol Surg*, **2018**; 13(7):967-975.

130. **Men K, Dai J, Li Y.** Automatic segmentation of the clinical target volume and organs at risk in the planning CT for rectal cancer using deep dilated convolutional neural networks: *Med Phys*, **2017**; 44(12):6377-6389.
131. **Hänsch A, Schwier M, Gass T, Morgas T.** Evaluation of deep learning methods for parotid gland segmentation from CT images. *Journal of Medical Imaging*, **2018**; 6(01):1.
132. **Zhu J, Zhang J, Qiu B, Liu Y, Liu X, Chen L.** Comparison of the automatic segmentation of multiple organs at risk in CT images of lung cancer between deep convolutional neural network-based and atlas-based techniques. *Acta Oncol (Madr)*, **2019**; 58(2):257-264.
133. **Qin W, Wu J, Han F, Yuan Y, Zhao W, Ibragimov B, et al.** Superpixel-based and boundary-sensitive convolutional neural network for automated liver segmentation. *Phys Med Biol*, **2018**; 63(9).
134. **Huang L, Xia W, Zhang B, Qiu B, Gao X.** MSFCN-multiple supervised fully convolutional networks for the osteosarcoma segmentation of CT images. *Comput Methods Programs Biomed*, **2017**; 143:67-74.
135. **Men K, Boimel P, Janopaul-Naylor J, Zhong H, Huang M, Geng H, et al.** Cascaded atrous convolution and spatial pyramid pooling for more accurate tumor target segmentation for rectal cancer radiotherapy. *Phys Med Biol*, **2018**; 63(18).
136. **Zhang R, Huang L, Xia W, Zhang B, Qiu B, Gao X.** Multiple supervised residual network for osteosarcoma segmentation in CT images. *Computerized Medical Imaging and Graphics*, **2018**; 63:1-8.
137. **Chen Y, Shi B, Wang Z, Zhang P, Smith CD, Liu J.** Hippocampus segmentation through multi-view ensemble ConvNets. *Proceedings - International Symposium on Biomedical Imaging*, Published online June 15, 2017:192-196.
138. **Vigneault DM, Xie W, Ho CY, Bluemke DA, Noble JA.** Ω -Net (Omega-Net): Fully automatic, multi-view cardiac MR detection, orientation, and segmentation with deep neural networks. *Med Image Anal*, **2018**; 48:95-106.
139. **Tao Q, Yan W, Wang Y, Paiman EHM, Shamonin DP, Garg P, et al.** Deep learning-based method for fully automatic quantification of left ventricle function from cine MR images: A multivendor, multicenter study. *Radiology*, **2019**; 290(1):81-88.
140. **Yang X, Zeng Z, Yi S.** Deep convolutional neural networks for automatic segmentation of left ventricle cavity from cardiac magnetic resonance images. *IET Computer Vision*, **2017**; 11(8):643-649.
141. **Zotti C, Luo Z, Lalande A, Jodoin PM.** Convolutional Neural Network with Shape Prior Applied to Cardiac MRI Segmentation. *IEEE J Biomed Health Inform*, **2019**; 23(3):1119-1128.
142. **Zheng Q, Delingette H, Duchateau N, Ayache N.** 3-D Consistent and Robust Segmentation of Cardiac Images by Deep Learning With Spatial Propagation. *IEEE Trans Med Imaging*, **2018**; 37(9):2137-2148.
143. **Wang J, Lu J, Qin G, Shen L, Sun Y, Ying H, et al.** Technical Note: A deep learning-based autosegmentation of rectal tumors in MR images. *Med Phys*, **2018**; 45(6):2560-2564.

144. **Pedoia V, Norman B, Mehany SN, Bucknor MD, Link TM, Majumdar S.** 3D convolutional neural networks for detection and severity staging of meniscus and PFJ cartilage morphological degenerative changes in osteoarthritis and anterior cruciate ligament subjects. *Journal of Magnetic Resonance Imaging*, **2019**; 49(2):400-410.
145. **Norman B, Pedoia V, Majumdar S.** Use of 2D U-net convolutional neural networks for automated cartilage and meniscus segmentation of knee MR imaging data to determine relaxometry and morphometry. *Radiology*, **2018**; 288(1):177-185.
146. **Jian J, Xiong F, Xia W, Zhang R, Gu J, Wu X, et al.** Fully convolutional networks (FCNs)-based segmentation method for colorectal tumors on T2-weighted magnetic resonance images. *Australas Phys Eng Sci Med*, **2018**; 41(2):393-401.
147. **Mohseni Salehi SS, Erdogmus D, Gholipour A.** Auto-Context Convolutional Neural Network (Auto-Net) for Brain Extraction in Magnetic Resonance Imaging. *IEEE Trans Med Imaging*, **2017**; 36(11):2319-2330.
148. **Pröve PL, Jopp-van Well E, Stanczus B, Morlock MM, Herrmann J, Groth M, et al.** Automated segmentation of the knee for age assessment in 3D MR images using convolutional neural networks. *Int J Legal Med*, **2019**; 133(4):1191-1205.
149. **Tustison NJ, Avants BB, Lin Z, Feng X, Cullen N, Mata JF, et al.** Convolutional Neural Networks with Template-Based Data Augmentation for Functional Lung Image Quantification. *Acad Radiol*, **2019**; 26(3):412-423.
150. **Zeng G, Zheng G.** Deep learning-based automatic segmentation of the proximal femur from MR images. *Adv Exp Med Biol*, **2018**; 1093:73-79.
151. **Huang Q, Sun J, Ding H, Wang X, Wang G.** Robust liver vessel extraction using 3D U-Net with variant dice loss function. *Comput Biol Med*, **2018**; 101:153-162.
152. **Kim S, Bae WC, Masuda K, Chung CB, Hwang D.** Fine-grain segmentation of the intervertebral discs from MR spine images using deep convolutional neural networks: BSU-Net. *Applied Sciences (Switzerland)*, **2018**; 8(9).
153. **Dalmış MU, Litjens G, Holland K, Setio A, Mann R, Karssemeijer N, et al.** Using deep learning to segment breast and fibroglandular tissue in MRI volumes: *Med Phys*, **2017**; 44(2):533-546.
154. **Zhu Y, Wei R, Gao G, Ding L, Zhang X, Wang X, et al.** Fully automatic segmentation on prostate MR images based on cascaded fully convolution network. *Journal of Magnetic Resonance Imaging*, **2019**; 49(4):1149-1156.
155. **Bonmati E, Hu Y, Sindhvani N, Dietz H, D'hooge J, Barratt D, et al.** Automatic segmentation method of pelvic floor levator hiatus in ultrasound using a self-normalizing neural network. *J Med Imaging (Bellingham)*, **2018**; 5(2):1.
156. **Mishra D, Chaudhury S, Sarkar M, Soin AS.** Ultrasound Image Segmentation: A Deeply Supervised Network With Attention to Boundaries. *IEEE Trans Biomed Eng*, **2019**; 66(6):1637-1648.
157. **Anas EMA, Mousavi P, Abolmaesumi P.** A deep learning approach for real time prostate segmentation in freehand ultrasound guided biopsy. *Med Image Anal*, **2018**; 48:107-116.

158. **Anas EMA, Nouranian S, Mahdavi SS, Spadinger I, Morris WJ, Salcudean SE, et al.** Clinical target-volume delineation in prostate brachytherapy using residual neural networks. *Medical Image Computing and Computer Assisted Intervention – MICCAI*, **2017**; 10435 LNCS:365-373.
159. **Savjani R.** NnU-net: Further automating biomedical image autosegmentation. *Radiol Imaging Cancer*, **2021**; 3(1).
160. **Gumusoy I, Haylaz E, Duman SB, Kalabalık F, Eren MC, Say S, et al.** Automatic Segmentation of the Infraorbital Canal in CBCT Images: Anatomical Structure Recognition Using Artificial Intelligence. *Diagnostics*, **2025**; 15(13).
161. **Kaygısız Yiğit M, Pınarbaşı A, Etöz M, Duman ŞB, Bayrakdar İŞ.** Artificial intelligence based fully automatic 3D paranasal sinus segmentation. *Dentomaxillofac Radiol*, Published online July 25, 2025.
162. **Ozudogru S, Gulsen E, Mahyaddinova T, Kizilay FN, Gulsen IT, Kuran A, et al.** Artificial intelligence system for automatic tooth detection and numbering in the mixed dentition in CBCT. *Eur J Paediatr Dent*, **2025**; 26(2):140-146.
163. **Yuce F, Buyuk C, Bilgir E, Çelik Ö, Bayrakdar İŞ.** Deploying a novel deep learning framework for segmentation of specific anatomical structures on cone-beam CT. *Oral Radiol*, **2025**; 41(4).
164. **Ünal SY, Namdar Pekiner F.** Evaluation of the mandibular canal and the third mandibular molar relationship by CBCT with a deep learning approach. *Oral Radiol*, **2025**; 41(2):222-230.
165. **Gülşen İT, Kuran A, Evli C, Baydar O, Dinç Başar K, Bilgir E, et al.** Deep learning model for automated segmentation of sphenoid sinus and middle skull base structures in CBCT volumes using nnU-Net v2. *Oral Radiol*, Published online August 1, 2025.
166. **Afaq S, Rao S.** Significance Of Epochs On Training A Neural Network. *Int J Sci Technol Res*, **2020**; 9(6):485-488.
167. **Xu J, Li Z, Du B, Zhang M, Liu J.** Reluplex made more practical: Leaky ReLU. *Proc IEEE Symp Comput Commun*, **2020**; 2020-July:1.
168. **Maas AL, Hannun AY, Ng AY.** Rectifier nonlinearities improve neural network acoustic models. *Proceedings of the 30th International Conference on Machine Learning*, **2013**; 30:1-6.
169. **He K, Zhang X, Ren S, Sun J.** Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Published online 2015:1026-1034.

ÖZGEÇMİŞ

