T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BEYİN MR GÖRÜNTÜLERİ ÜZERİNDEN BEYİN TÜMÖRÜ TESPİTİ, SINIFLANDIRMASI VE SEGMENTASYONUNDA DERİN ÖĞRENME YÖNTEMLERİNİN İNCELENMESİ

Esma Nur GÜNEŞER

YÜKSEK LISANS TEZİ
Matematik Mühendisliği Anabilim Dalı
Matematik Mühendisliği Programı

Danışman
Doç. Dr. Mert BAL

Mart, 2026

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BEYİN MR GÖRÜNTÜLERİ ÜZERİNDEN BEYİN TÜMÖRÜ TESPİTİ,SINIFLANDIRMASI VE SEGMENTASYONUNDA DERİN ÖĞRENME YÖNTEMLERİNİN İNCELENMESİ

Esma Nur GÜNEŞER tarafından hazırlanan tez çalışması 10.03.2026 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Matematik Mühendisliği Anabilim Dalı Matematik Mühendisliği Programı **YÜKSEK LISANS TEZİ** olarak kabul edilmiştir.

Doç. Dr. Mert BAL
Yıldız Teknik Üniversitesi
Danışman

Jüri Üyeleri

Doç. Dr. Mert BAL, Danışman
Yıldız Teknik Üniversitesi _____

Doç. Dr. Ramazan TEKERCİOĞLU, Üye
Yıldız Teknik Üniversitesi _____

Dr. Öğr. Üyesi Ebu Yusuf GÜVEN, Üye
İstanbul Üniversitesi - Cerrahpaşa _____

Danışmanım Doç. Dr. Mert BAL sorumluluğunda tarafımca hazırlanan “Beyin MR Görüntüleri Üzerinden Beyin Tümörü Tespiti, Sınıflandırması ve Segmentasyonunda Derin Öğrenme Yöntemlerinin İncelenmesi” başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim.

Bu tez çalışması sürecinde, tez sahibinin kontrolü altında dil ve ifade bütünlüğünü güçlendirmek, teknik süreçleri desteklemek ve literatür taraması amacıyla Üretken Yapay Zekâ (ÜYZ) araçlarından Gemini ve Consensus kullanılmıştır.

ÜYZ araçlarının yanlış veya gerçek dışı bilgi (halüsinasyon) üretebilme riski taşıdığının farkında olduğumu, bu araçlardan elde edilen tüm çıktıları doğruluk, bütünlük, güvenilirlik ve etik standartlar açısından dikkatle kontrol ettiğimi beyan ederim.

ÜYZ kullanımı hiçbir aşamada yazarın entelektüel katkısının, bilimsel muhakemesinin ve özgün fikir üretiminin yerine geçmemiştir. Tez çalışmasında yer alan tüm bilgi, veri, analiz, iddia ve referansların doğruluğu, güncelliği ve bilimsel geçerliliğine ilişkin sorumluluk tamamen tarafıma aittir.

Beyanımın aksinin tespiti hâlinde doğabilecek her türlü idari ve hukuki sorumluluğu kabul ederim.

Esma Nur GÜNEŞER

İmza

Dedicated to my family
and my best friend

TEŞEKKÜR

Tez çalışmamın planlanmasından sonuçlanmasına kadar geçen süreçte, engin bilgi ve tecrübesiyle bana yol gösteren, her türlü desteği ve sabrı ile yanımda olan değerli danışman hocam Sayın Doç. Dr. Mert Bal'a en içten teşekkürlerimi sunarım.

Çalışmamın tıbbi değerlendirme süreçlerinde uzman görüşleriyle araştırmama büyük katkı sağlayan, tecrübelerini benden esirgemeyen İstanbul Medeniyet Üniversitesi Radyoloji Anabilim Dalı Öğretim Üyesi Sayın Doç. Dr. Begümhan Baysal'a değerli katkılarından dolayı teşekkür ederim.

Son olarak, eğitim hayatım boyunca maddi ve manevi hiçbir fedakârlıktan kaçınmayan, sabır, sevgi ve inançlarıyla her zaman en büyük destekçim olan aileme sonsuz minnet ve şükranlarımı sunarım.

Esma Nur GÜNEŞER

İÇİNDEKİLER

SİMGE LİSTESİ viii

KISALTMA LİSTESİ ix

ŞEKİL LİSTESİ xi

TABLO LİSTESİ xii

ÖZET xiii

ABSTRACT xv

1 GİRİŞ 1

1.1 Problem Tanımı ve Motivasyon 1

1.2 Tezin Amacı ve Kapsamı 2

1.3 Literatür Taraması 3

2 BEYİN TÜMÖRLERİ MEDİKAL ARKA PLAN VE MANYETİK REZONANS GÖRÜNTÜLEME 6

2.1 Beyin Tümörlerinin Tanımı, Epidemiyolojisi ve İstatistikler 6

2.2 Beyin Tümörlerinin Sınıflandırılması ve Derecelendirme Sistemleri 7

2.2.1 WHO Derecelendirme Sistemi (Grading) ve Moleküler Devrim 7

2.2.2 Doku Kökenine Göre Tümör Tipleri 8

2.2.3 Tanı ve Görüntüleme Yöntemleri 9

2.3 Manyetik Rezonans Görüntüleme 10

2.3.1 Manyetik Rezonans Görüntülemenin Tarihçesi 10

2.3.2 MRG Makinalarının Bileşenleri 11

2.3.3 Manyetik Rezonans Görüntü Süreci 12

3 DERİN ÖĞRENMENİN MATEMATİKSEL TEMELLERİ 17

3.1 Derin Öğrenmenin Tarihsel Gelişimi ve Kırılma Noktaları 18

3.2 Yapay Sinir Ağları (YSA) 20

3.2.1 Sinir Ağlarının Temel Bileşenleri 21

3.2.2 Öğrenme Süreci: Geri Yayılım ve Optimizasyon Algoritmaları . . . 23
3.2.3 Katmanlı Mimari ve Derin İleri Beslemeli Ağlar . . . 25
3.3 Evrişimli Sinir Ağları (CNN) Matematiği . . . 26
3.3.1 CNN Mimarisinin Katmanları . . . 27
3.3.2 CNN Tabanlı Sınıflandırma Mimarileri . . . 29
3.4 Transformer Mimarisi . . . 35
3.4.1 Mimarinin Yapısı . . . 36
3.4.2 Transformer Tabanlı Mimariler . . . 39
3.5 Transfer Öğrenme Teorisi . . . 42
3.5.1 Matematiksel Çerçeve ve Alan Tanımı . . . 42
3.5.2 ImageNet Veri Seti . . . 43
3.5.3 RadImageNet Veri Seti . . . 43
3.6 Semantik Segmentasyon ve Mimari Yapılar . . . 44
3.6.1 Mimari Elemanları . . . 44
3.6.2 Segmentasyon Mimari Modelleri . . . 46

4 METODOLOJİ VE ÖNERİLEN YAKLAŞIM 51
4.1 Donanım ve Yazılım Altyapısı . . . 51
4.2 Kullanılan Veri Setleri . . . 51
4.2.1 BR35H (Brain Tumor Detection 2020) . . . 52
4.2.2 Beyin Tümörü MRG Veri Seti (Nickparvar) . . . 52
4.2.3 BRISC 2025 (Brain Imaging Segmentation and Classification) 52
4.2.4 BraTS 2020 (Multimodal Brain Tumor Segmentation Challenge) . . . 53
4.3 Önerilen Sistem Mimarisi . . . 54
4.3.1 Veri Setlerinin Düzenlenmesi ve Hazırlık Süreçleri . . . 55
4.3.2 Sınıflandırma Veri Setleri: Birleştirme ve Algısal Temizleme 55
4.3.3 Segmentasyon Veri Setleri: 3D-2D Dönüşümü ve Dilimleme 56
4.3.4 Veri Bölümleme ve Tutarlılık Kontrolü . . . 56
4.4 Veri Ön İşleme ve Zenginleştirme . . . 57
4.4.1 Beyin Çıkarma (Brain Extraction / Skull-Stripping) . . . 57
4.4.2 Görüntü Boyutlandırma ve Kanal Dönüşümü . . . 57
4.4.3 Çevrimiçi Veri Artırma (On-the-fly Augmentation) . . . 58
4.4.4 Normalizasyon (Normalization) . . . 59
4.5 Deneysel Kurulum ve Eğitim Stratejileri . . . 59
4.5.1 Yazılım ve Donanım Konfigürasyonu . . . 59
4.5.2 Transfer Öğrenme ve Ağırlık Başlatma Stratejileri . . . 60
4.5.3 Hiperparametreler ve Optimizasyon . . . 60

4.6 Performans Değerlendirme Metrikleri 61
4.6.1 Sınıflandırma Metrikleri 61
4.6.2 Segmentasyon Metrikleri 61

5 DENEYSEL SONUÇLAR **63**
5.1 Sınıflandırma Sonuçları ve Yorumları 63
5.1.1 Grup 1: Temel CNN Modelleri 63
5.1.2 Grup 2: Transformer Modelleri ve Optimizasyon Etkisi . . . 65
5.1.3 Kör Test (Blind Test) Sonuçları 66
5.2 Segmentasyon Sonuçları 67
5.2.1 Tek Kanallı Segmentasyon Sonuçları (BRISC-2025) 67
5.2.2 Çok Sekans Segmentasyon Sonuçları (BraTS-2020) 69

6 TARTIŞMA **74**
6.1 Sınıflandırma Sonuçlarının Literatürdeki Benzer Çalışmalarla Karşılaştırmalı Değerlendirmesi 74
6.2 Literatürdeki En İyi Sınıflandırma Sonuçlarıyla Karşılaştırma 75
6.3 Segmentasyon Sonuçlarının Değerlendirilmesi 76
6.4 Tez Çalışması Kapsamında Elde Edilen En İyi Segmentasyon Performanslarının Özeti 77

7 SONUÇ **78**

KAYNAKÇA **80**

TEZDEN ÜRETİLMİŞ YAYINLAR **86**

SİMGE LİSTESİ

Ai	Activities of Daily Life
c	Alternate Step Test
C	Body Mass Index
CR	Cross Step moving on Four Stops
$fc(.)$	Dynamic Bayesian Networks
ΔH	Demura's Fall Risk Assessment Chart
λi	Electromyography
Ω	Faculdade de Engenharia da Universidade do Porto

KISALTMA LİSTESİ

BN	Batch Normalization
BOS	Beyin Omurilik Sıvısı
BT	Bilgisayarlı Tomografi
CNN	Evrişimli Sinir Ağları (Convolutional Neural Networks)
CNS	Merkezi Sinir Sistemi (Central Neural System)
DenseNet	Densely Connected Networks
EPI	Echo Planar Imaging
FCN	Fully Convolutional Networks
FFT	Fast Fourier Transform
FLAIR	Fluid Attenuated Inversion Recovery
GELU	Gaussian Error Linear Unit
IR	İnversiyon Recovery
MLP	Çok Katmanlı Algılayıcı (Multi-Layer Perceptron)
MR	Manyetik Rezonans
MRG	Manyetik Rezonans Görüntüleme
MSA	Multi-Head Self Attention
NAS	Neural Architecture Search
NLP	Doğal Dil İşleme (Natural Language Processing)
NMR	Nükleer Manyetik Rezonans
ReLU	Rectified Linear Unit
ResNet	Residual Networks
RF	Radyo Frekans
SGD	Stokastik Gradyan İniş (Stochastic Gradient Descent)

T1+CE	T1 Contrast Enhanced
TE	Time to Echo
TI	Time of Inversion
TR	Time to Repeat
ViT	Vision Transformer
YSA	Yapay Sinir Ağları

ŞEKİL LİSTESİ

Şekil 3.1 Eşitlik (3.1) de verilen perseptron modeli şekildeki gibi gösterilebilir. 21
Şekil 3.2 Şekilde CNN mimarisinin temsili bir modeli görülmektedir. Burada görüntü matrisinin tek boyuta inişi gözlemlenebilir [36]. . 27
Şekil 3.3 Sol: ResNet-34 için yapı taşı Sağ: ResNet50/101/152 için dar boğaz yapısı [3]. 30
Şekil 3.4 DenseNet mimarisinin yapısı şekildeki gibi özetlenebilir [37]. . . 32
Şekil 3.5 ViT mimarisinin temsili gösterimi [31] 40
Şekil 3.6 Unet mimarisinin temsili görseli [29]. 47
Şekil 5.1 UNet++ RAW için örnek bir segmentasyon çıktısı. Sol: orjinal görüntü, Orta: orjinal maske, Sağ: Modelin çizdiği maske 68
Şekil 5.2 UNet++ BE için örnek bir segmentasyon çıktısı. Sol: orjinal görüntü, Orta: orjinal maske, Sağ: Modelin çizdiği maske 69
Şekil 5.3 LinkNet BE için örnek bir segmentasyon çıktısı. Sol: orjinal görüntü, Orta: orjinal maske, Sağ: Modelin çizdiği maske 69
Şekil 5.4 UNet++ için sırayla maske ve pred: tahminler gösterilir. 70
Şekil 5.5 UNet++ için sırayla maske ve pred: tahminler gösterilir. 71
Şekil 5.6 UNet++ için sırayla maske ve pred: tahminler gösterilir. 71
Şekil 5.7 LinkNet için sırayla maske ve pred: tahminler gösterilir. 72
Şekil 5.8 UNet++ için sırayla maske ve pred: tahminler gösterilir. 72
Şekil 5.9 UNet++ için sırayla maske ve pred: tahminler gösterilir. 72

TABLO LİSTESİ

Tablo 2.1 WHO Tümör Derecelendirme Sistemi ve Özellikleri 8
Tablo 2.2 Sekansların özelliklerini özetleyen tablo 16
Tablo 4.1 Kullanılan Veri Setlerinin Genel Özellikleri 53
Tablo 5.1 Temel CNN Modellerinin İkili Sınıflandırma (Binary) Performansı (Test Accuracy) 64
Tablo 5.2 Temel CNN Modellerinin Dörtlü Sınıflandırma (C4C) Performansı (Test Accuracy) 64
Tablo 5.3 Hafif Modellerin İkili Sınıflandırma (Binary) Performansı 65
Tablo 5.4 Hafif Modellerin Dörtlü Sınıflandırma (C4C) Performansı 65
Tablo 5.5 Transformer Modellerinin İkili Sınıflandırma (Binary) Performansı 65
Tablo 5.6 Transformer Modellerinin Dörtlü Sınıflandırma (C4C) Performansı 66
Tablo 5.7 Kör Test - İkili Sınıflandırma (Binary) En Başarılı 10 Model . . . 66
Tablo 5.8 Kör Test - Dörtlü Sınıflandırma (4-Class) En Başarılı 10 Model . 67
Tablo 5.9 Ham (RAW) Veri Seti Segmentasyon Sonuçları 68
Tablo 5.10 Beyin Çıkarılmış (BE) Veri Seti Segmentasyon Sonuçları 68
Tablo 5.11 T1+CE (Kontrastlı T1) Sekansı Segmentasyon Performansı . . . 70
Tablo 5.12 FLAIR Sekansı Segmentasyon Performansı 71
Tablo 5.13 T2 Sekansı Segmentasyon Performansı 71
Tablo 5.14 T1 (Standart) Sekansı Segmentasyon Performansı 72
Tablo 6.1 Gómez-Guzmán [58] ile Bu Tezdeki Sınıflandırma Başarılarının Karşılaştırması 74
Tablo 6.2 Güncel Literatürdeki En İyi Sınıflandırma Sonuçları ile Karşılaştırma 75
Tablo 6.3 BraTS 2020 Veri Seti Üzerinde 2 Boyutlu Yaklaşımların Dice Skoru Karşılaştırması 76
Tablo 6.4 Farklı Segmentasyon Görevlerinde Elde Edilen En İyi Model ve Sekans Kombinasyonları 77

ÖZET

Beyin MR Görüntüleri Üzerinden Beyin Tümörü Tespiti, Sınıflandırması ve Segmentasyonunda Derin Öğrenme Yöntemlerinin İncelenmesi

Esma Nur GÜNEŞER

Matematik Mühendisliği Anabilim Dalı
Yüksek Lisans Tezi

Danışman: Doç. Dr. Mert BAL

Beyin tümörlerinin erken teşhisi ve cerrahi planlama için tümör sınırlarının hassas bir şekilde belirlenmesi, hastanın hayatta kalma oranını doğrudan etkileyen kritik bir süreçtir. Manyetik Rezonans (MR) görüntüleme, bu süreçte altın standart olsa da, görüntülerin manuel analizi zaman alıcıdır ve uzmanlar arası yorum farklılıklarına açıktır. Bu tez çalışmasında, beyin tümörlerinin otomatik tespiti (sınıflandırma) ve lezyon sınırlarının belirlenmesi (segmentasyon) amacıyla, kısıtlı donanım kaynaklarında (3GB VRAM) yüksek performans sağlayan uçtan uca bir derin öğrenme sistemi geliştirilmiştir.

Çalışma kapsamında Br35H, Nickparvar, BRISC-2025 ve BraTS-2020 veri setleri birleştirilerek kapsamlı bir veri havuzu oluşturulmuştur. Metodolojik bir yenilik olarak, görüntülerden kafatası ve beyin dışı dokuların temizlendiği özgün bir "Beyin Çıkarma (Brain Extraction - BE)" algoritması geliştirilmiş ve bu işlemin model başarısına etkisi analiz edilmiştir. Sınıflandırma aşamasında ResNet, DenseNet, Inception, EfficientNet gibi Konvolüsyonel Sinir Ağları (CNN) ile Swin Transformer ve Vision Transformer (ViT) mimarileri karşılaştırılmıştır. Literatürdeki standart yaklaşımların aksine, bu çalışmada genel nesne tanıma verileriyle eğitilmiş ImageNet ağırlıkları ile radyolojik görüntülerle eğitilmiş RadImageNet (RIN) ağırlıklarının transfer öğrenme performansları sistematik olarak kıyaslanmıştır. Ayrıca Transformer modellerinin optimizasyon hassasiyeti (Adam vs AdamW) deneysel olarak incelenmiştir.

Elde edilen sonuçlar, medikal alana özgü transfer öğrenmenin (RIN), özellikle karmaşık dörtlü sınıflandırma görevlerinde (Glioma, Meningioma, Hipofiz, Sağlıklı) standart ImageNet modellerine göre daha üstün olduğunu göstermiştir. InceptionV3 modeli, RIN ağırlıkları ve BE ön işlemesi ile %99.12 doğruluk oranına ulaşırken, hafif mimariye sahip EfficientNet-B0 modeli %99.78 başarı ile en verimli model olmuştur. Transformer tabanlı modellerde ise AdamW optimizasyonunun, ViT modelinin performansını %9 oranında artırdığı ve ön eğitimin (Pre-training) bu mimariler için zorunlu olduğu tespit edilmiştir. Segmentasyon modülünde ise Linknet mimarisi, BRISC veri setinde 0.86 Dice Skoru ve 3.15 piksel Hausdorff hatası ile en başarılı sonuçları vermiştir. Çok modlu BraTS veri setinde ise Kontrastlı T1 (T1CE) sekansının aktif tümör tespiti için en kritik modalite olduğu kanıtlanmıştır.

Sonuç olarak bu tez; doğru veri ön işleme, alana özgü transfer öğrenme ve optimize edilmiş hafif mimarilerin kullanımıyla, yüksek hesaplama gücü gerektirmeden klinik düzeyde başarılı bir tanı destek sistemi geliştirilebileceğini ortaya koymuştur.

Anahtar Kelimeler: Derin Öğrenme, MRI, Beyin Tümörü, Sınıflandırma, Segmentasyon, ImageNet, RadImageNet, CNN, Vision Transformer, Unet.

ABSTRACT

An Investigation of Deep Learning Approaches for Brain Tumor Detection, Classification, and Segmentation from Brain MRI Images

Esma Nur GÜNEŞER

Department of Mathematical Engineering
Master of Science Thesis

Supervisor: Assoc. Prof. Dr. Mert BAL

Early diagnosis of brain tumors and the precise delineation of tumor boundaries for surgical planning are critical processes that directly influence patient survival rates. Although Magnetic Resonance (MR) imaging is the gold standard in this domain, manual analysis of images is time-consuming and prone to inter-observer variability. In this thesis, an end-to-end, multi-stage deep learning system capable of high performance on limited hardware resources (3GB VRAM) has been developed for the automatic detection (classification) and lesion segmentation of brain tumors.

Within the scope of the study, a comprehensive data repository was constructed by combining Br35H, Nickparvar, BRISC-2025, and BraTS-2020 datasets. As a methodological innovation, a novel "Brain Extraction (BE)" algorithm utilizing morphological operations to remove skull and non-brain tissues was developed, and its impact on model performance was analyzed. In the classification phase, Convolutional Neural Networks (CNN) such as ResNet, DenseNet, Inception, and EfficientNet were compared with Swin Transformer and Vision Transformer (ViT) architectures. Contrasting with standard approaches in the literature, this study systematically compared the transfer learning performance of ImageNet weights (trained on general object recognition data) versus RadImageNet (RIN) weights (trained on radiological images). Additionally, the optimization sensitivity (Adam vs. AdamW) of Transformer models was experimentally investigated.

Experimental results demonstrated that medical domain-specific transfer learning

(RIN) is superior to standard ImageNet models, particularly in complex four-class classification tasks (Glioma, Meningioma, Pituitary, Healthy). The InceptionV3 model achieved 99.12% accuracy using RIN weights and BE preprocessing, while the EfficientNet-B0 model, with its lightweight architecture, emerged as the most efficient model with 99.78% accuracy. For Transformer-based models, it was determined that AdamW optimization improved ViT performance by 9%, and pre-training proved to be essential for these architectures. In the segmentation module, the Linknet architecture yielded the best results on the BRISC dataset with a 0.86 Dice Score and 3.15 pixel Hausdorff error. Furthermore, analyses on the multi-modal BraTS dataset proved that the Contrast-Enhanced T1 (T1CE) sequence is the most critical modality for active tumor detection.

In conclusion, this thesis demonstrates that a clinically viable diagnostic support system can be developed without requiring high computational power by employing robust data preprocessing, domain-specific transfer learning, and optimized lightweight architectures.

Keywords: Deep Learning, MRI, Brain Tumor, Classification, Segmentation, ImageNet, RadImageNet, CNN, Vision Transformer, Unet.

1
GİRİŞ

Tıbbi görüntüleme teknolojilerindeki gelişmeler ve yapay zeka alanındaki devrim niteliğindeki ilerlemeler, hastalıkların erken teşhisi ve tedavisinde yeni bir dönem başlatmıştır. Özellikle hayati öneme sahip organların incelenmesinde, Manyetik Rezonans Görüntüleme (MRG) gibi non-invaziv[1] yöntemlerin derin öğrenme algoritmalarıyla entegrasyonu, radyologlara karar destek mekanizması sağlama potansiyeli taşımaktadır. Ancak beyin tümörlerinin tespiti, sınıflandırılması ve segmentasyonu; salt bir bilgisayarlı görü problemi olmanın ötesinde, karmaşık anatomik ve radyolojik zorlukları barındıran çok katmanlı bir süreçtir.

1.1 Problem Tanımı ve Motivasyon

Beyin, homojen bir yapı olmayıp; beyaz madde[2], gri madde ve beyin omurilik sıvısı (BOS) gibi farklı yoğunluk ve doku özelliklerine sahip katmanlardan oluşan heterojen bir organdır. Tümör dokusunun bu karmaşık yapı içerisindeki konumu, sınırlarının belirlenmesini zorlaştırmaktadır. Ayrıca beynin kafatası kemiği ile çevrili kapalı bir kutu içerisinde yer alması, görüntü işleme süreçlerinde kafatası temizleme (skull stripping) gibi ek ön işleme adımlarını ve bu adımların model başarısı üzerindeki etkisinin analiz edilmesini zorunlu kılmaktadır.

Klinik pratikteki ve otomatize sistemlerdeki en büyük zorluklardan biri ayırt edicilik problemidir. Kafa içi yapılarda, tümör dışı diğer anomaliler (enfeksiyonlar, iskemik lezyonlar vb.) tümör dokusuyla benzer morfolojik özellikler gösterebilmektedir. Dahası, beyin tümörlerinin kendi alt tipleri görsel olarak birbirine oldukça yakındır. Bu durum, tek bir görüntü kesiti üzerinden çalışan yapay zeka modelleri için ciddi bir yanlış tanı riskini beraberinde getirmektedir.

[1] Vücuda herhangi bir fiziksel giriş yapılmadan dışarıdan uygulanan tanı veya tedavi yöntemlerini ifade eder.

[2] Sinir lifleri, miyelin, miyelin üreten oligodendrositler ve beyin dokusundaki düzenliliği koruyan astrositlerden oluşur.

Bu noktada veri modalitesinin (sekans bilgisinin) önemi ortaya çıkmaktadır. Radyolojik tanıda tek bir doğru görüntü kavramı yoktur; aksine, farklı doku karakteristiklerini ortaya çıkaran sekansların (T1, T2, FLAIR, T1+CE[3]) bir arada değerlendirilmesi esastır. Örneğin; FLAIR sekansı ödemi netleştirirken, T1+CE (kontrastlı) sekansı aktif tümör çekirdeğini vurgular. Sekans bilgisinin eksik olduğu, tekil kullanıldığı veya hangi sekansın hangi bilgiyi taşıdığının modele doğru aktarılmadığı senaryolarda, teşhisin güvenilirliği azalmaktadır. Literatür incelendiğinde; çalışmaların genellikle spesifik bir mimarinin ince ayarlarla (fine-tuning) maksimize edilmiş başarısına odaklandığı, ancak hangi duruma hangi veri setinin veya modelin uygun olduğuna dair kapsamlı bir rehberin eksik olduğu görülmektedir.

1.2 Tezin Amacı ve Kapsamı

Bu tez çalışmasının temel amacı; beyin tümörü tespiti ve segmentasyonu alanında yapılacak ileri düzey akademik çalışmalara zemin oluşturacak nitelikte, doğru model, doğru veri seti ve doğru sekans seçimi konusunda araştırmacılara yol gösterecek bir analiz ve karar destek altyapısı sunmaktır. Çalışma, modelleri hiperparametre optimizasyonu ile spesifik bir veriye uydurmak yerine; modellerin genel başarım kapasitelerini, farklı sınıflandırma zorlukları altındaki davranışlarını ve tıbbi gerçeklikle örtüşme düzeylerini deneysel olarak incelemeyi hedeflemiştir.

Çalışma kapsamında öncelikle sınıflandırma mimarilerinin ayırt edicilik yetenekleri, problemin zorluk derecesine göre analiz edilmektedir. Modellerin, sadece "tümör var/yok" şeklindeki temel (ikili/binary) ayrımı yapması ile görsel ve yapısal olarak birbirine oldukça benzeyen farklı tümör tiplerini (çoklu sınıf/multi-class) ayırt etmesi gereken senaryolar arasındaki performans farkları karşılaştırılmaktadır. Bu bağlamda, genel nesne tanıma görevlerinde standart kabul edilen ImageNet ağırlıkları ile radyolojik görüntülerle eğitilmiş RadImageNet ağırlıklarının ön eğitimleri kıyaslanarak; beyin tümörü gibi spesifik ve veri kısıtı olan bir alanda, alana özgü (domain-specific) transfer öğrenmenin, sıfırdan eğitime veya genel amaçlı modellere göre anlamlı bir fark yaratıp yaratmadığı sorgulanmaktadır.

Segmentasyon görevlerinde ise veri seti seçiminde sekans bilgisinin kritik rolü ve optimum kombinasyonlar mercek altına alınmıştır. Eldeki veri setinde sekans bilgisinin hiç bulunmaması, tek bir sekansın bulunması veya çoklu sekansların yer alması durumlarının model başarısını nasıl değiştirdiği incelenmektedir. Özellikle

[3]T1 Contrast Enhanced - Kontrast madde enjekte edilerek alınan T1 sekans görüntüsüdür.

"daha fazla veri her zaman daha iyidir" algısının aksine, tümör çekirdeği veya ödem bölgesi gibi spesifik alanların segmentasyonu için hangi sekans görüntülerinin (örneğin sadece FLAIR veya T1+CE kullanımı gibi) daha efektif sonuçlar verdiği deneysel olarak ortaya konulmaktadır.

Son olarak çalışma, derin öğrenme modellerinin şeffaflığını ve güvenilirliğini artırmak adına, modellerin tıbbi bilgi ile ne derece örtüştüğünü de değerlendirmektedir. Farklı tümör alt bölgelerini temsil eden maskeler oluşturulurken, modelin radyolojik gerçekliğe uygun öznitelik çıkarımı yapıp yapmadığı; yani T1+CE sekansını aktif tümörle, FLAIR sekansını ödemle eşleştirip eşleştirmediği analiz edilmektedir. Tüm bu süreçlerde, modellerin ince ayarlarla yapay olarak yükseltilmiş skorları yerine ham (baseline) başarımları esas alınmış; kafa tası temizleme gibi ön işleme adımlarının ve veri setinin yapısal özelliklerinin öğrenme dinamiği üzerindeki etkileri kapsamlı bir şekilde raporlanmıştır.

1.3 Literatür Taraması

Beyin tümörlerinin medikal görüntüler üzerinden otomatik analizi, son yıllarda bilgisayarlı görü ve derin öğrenme alanındaki gelişmelerle birlikte ivme kazanmış geniş bir araştırma sahasıdır. Literatürdeki çalışmalar genel olarak; tümörün varlığının tespiti (sınıflandırma), tümör alt bölgelerinin ayrıştırılması (segmentasyon) ve veri yetersizliği problemleriyle başa çıkma stratejileri (transfer öğrenme ve veri artırma) üzerine yoğunlaşmaktadır. Bu bölümde, tez kapsamında ele alınan yöntemlerin tarihsel gelişimi ve mevcut literatürdeki yeri detaylandırılmıştır.

Derin öğrenme tabanlı çalışmalar için yaygın kullanılan açık veri setleri şunlardır: BraTS (Brain Tumor Segmentation Challenge) çok modlu MR görüntüleri (T1, T1+CE, T2, FLAIR) ve uzman segmentasyon maskeleri içerir; Figshare gibi platformlarda paylaşılan veri setleri (ör. Jun Cheng'in paylaştığı 3064 adet T1 ağırlıklı görüntü kümesi); ve çeşitli Kaggle kümeleri (ör. Nick Parvar, Vishwa Patel'in setleri, BR35H) gibi kaynaklar. BraTS yıllardır glial tümör alt bölümlerini otomatik segmentasyon için uluslararası bir standart olmuş, 2023'te pediatrik glioma ve metastaz/menengiom gibi alt görevler içerecek şekilde genişletilmiştir [1, 2]. Literatürdeki derlemeler, en çok kullanılan kümelerin Figshare kaynaklı veri setleri olduğunu, BraTS'ın ise takip ettiğini göstermiştir. Bu ve benzeri açık veri setleri, model eğitimi ve karşılaştırmalar için önemli değerlendirme deneyleri olarak kullanılmaktadır.

Beyin tümörü sınıflandırmasında başta evrişimli sinir ağları (CNN) olmak üzere pek çok derin ağ mimarisi kullanılmıştır. Özellikle ResNet tabanlı modeller en yaygın olandır [1]. Örneğin He ve arkadaşlarının geliştirdiği ResNet, çok katmanlı evrişim filtreleri ve atlama (skip) bağlantıları ile MR görüntülerinden hiyerarşik özellikler çıkarıp tümör tiplerini sınıflandırır [3]. Bu CNN'ler genellikle ImageNet veya tıbbi görüntülerden ön-eğitilmiş ağırlıklarla kullanılır. Son yıllarda Vision Transformer (ViT) ve kendine dikkat mekanizmaları (attention) içeren modeller de yaygınlaşmaktadır [1]. ViT, görüntü içindeki uzak bölgeler arası ilişkileri öğrenmede güçlüdür fakat genellikle büyük veri gerektirir; bu nedenle halihazırda çok büyük veri seti olmayan beyin tümörü sınıflandırmasında gelecek vaat etmemektedir.

RadImageNet ön-eğitimi, tıbbi görüntülerden oluşan büyük bir veri kümesi ile hazırlanmış modeller sunar. Yang ve arkadaşlarının çalışmasında RadImageNet ön-eğitimli modeller, standart ImageNet modellerine göre sınıflandırmada %0.9–9.4 arasında AUC artışı göstermiştir [4]. Bu kapsamda, ImageNet veya RadImageNet ön eğitimi transfer öğrenme için kullanılmakta; RadImageNet'in lezyon segmentasyonu görevlerinde de (ör. tiroid ve meme görüntülerinde) ImageNet'e kıyasla anlamlı gelişmeler sağladığı belirtilmektedir.

Literatürde ayrı ayrı çalışan tek CNN modellerin yanı sıra hibrit yaklaşımlar da önerilmiştir. Örneğin ConvAttenMixer ve ARM-Net gibi çalışmalarda, CNN mimarilerine ek olarak dışadönük ve içsel dikkat mekanizmaları ekleyerek çoklu ölçekli özellikler çıkarılmıştır [2]. Bazı modeller CNN + RNN katmanları (ör. AlexNet-GRU gibi) ile uzaysal özelliklere ek olarak dilimsel veya zamansal bağımlılıkları öğrenmeye çalışmıştır. Yapısal olarak, graf tabanlı ağlar veya kapsül ağlar gibi alternatifler de incelenmiştir. Genellikle literatürde doğruluk değerleri %90'lar düzeyine ulaşmış olup, en yüksek başarıları güncel çalışmalarda dikkat mekanizmaları ve ön-eğitim kullanan derin modeller sergilemektedir [2].

Beyin tümörü segmentasyonunda en yaygın mimari U-Net'tir. Ronneberger ve ark. tarafından 2015'te önerilen U-Net, daralan (encoder) ve genişleyen (decoder) yolları ve katmanlar arası bağlantıları sayesinde segmentasyon maskesi oluşturur. U-Net, BraTS yarışmalarının kazanan modellerinin temeli olmuş ve nöroradyolojide sıklıkla tercih edilmiştir [1]. Çeşitli U-Net türevleri geliştirilmiştir: Örneğin Myronenko (2018) U-Net'e rezidüel bloklar ekleyerek başkanlık kazanmıştır; başka çalışmalarda çoklu kademeli (cascaded) U-Net'ler ile ardışık maskeler üretilmiştir; nnU-Net projesi ise otomatik ön işleme, hiper-parametre ayarı ve ansamble stratejisi ile her yarışmada üstün sonuç sağlamıştır. U-Net dışında LinkNet, U-Net++, DeepLab gibi encoder-decoder

ağları da beyin tümörü segmentasyonunda kullanılmıştır. Örneğin EfficientNetB7 tabanlı LinkNet modelleri doğuştan tercih edilmiş, çoklu atlamalı U-Net++ mimarileri denenmiştir. Hibrit olarak, CNN ve Transformatör'ün kombinasyonu (ör. TransBTS gibi) segmentasyon çalışılmış; çoklu-öznitelikli ve dikkat katmanlı U-Net varyasyonları tasarlanmıştır [1].

Beyin MR görüntülerinde genellikle normal beyin dokusunu ve kafatası gibi bölgesel olarak ilgili olmayan yapıları kaldırmak için beyin çıkarma (skull-stripping) uygulanır. Ancak bazı çalışmalar bu aşamanın hatalara yol açabileceğini göstermiştir. Pacheco ve ark. beyin çıkarma adımının tumor segmentasyonuna etkisini incelediklerinde, doğru BE yönteminin segmentasyonu %15'e varan oranda etkileyebileceğini; hatta bazen işlenmemiş (BE yapılmamış) MR görüntüleriyle eğitimin daha rekabetçi sonuçlar verebileceğini bulmuştur [5]. Bu sonuç, klinik uygulamalarda tümör segmentasyon modellerinin hata kaynaklarını azaltmak için BE adımını atlamanın bir strateji olabileceğini düşündürmektedir.

2
BEYİN TÜMÖRLERİ MEDİKAL ARKA PLAN VE MANYETİK REZONANS GÖRÜNTÜLEME

Bu bölüm, tez çalışmasında önerilen derin öğrenme modellerinin üzerinde çalıştığı biyolojik problemin doğasını ve verinin fiziksel oluşum süreçlerini kapsamlı bir şekilde ele almaktadır. Beyin tümörlerinin moleküler ve histopatolojik özellikleri, hastalığın klinik seyri ve bu patolojilerin dijital veriye dönüştürülmesini sağlayan Manyetik Rezonans Görüntüleme (MRG) teknolojisinin matematiksel temelleri aşağıda detaylandırılmıştır.

2.1 Beyin Tümörlerinin Tanımı, Epidemiyolojisi ve İstatistikler

Beyin tümörleri, Dünya Sağlık Örgütü'nün (WHO) Merkezi Sinir Sistemi (MSS) tümör sınıflamasında tanımlandığı üzere, beyin parankimi[1], meninksler[2], kranial sinirler veya hipofiz bezi gibi yapılardan köken alabilen ve moleküler/ histopatolojik özellikleriyle sınıflandırılan neoplazmlardır[3] [6]. Biyolojik düzeyde bu süreç, hücrenin programlı ölümü (apoptoz) mekanizmalarının devre dışı kalması ve proliferasyon[4] yeteneği kazanmasıyla başlar. Tümör dokusu, sağlıklı beyin parankimini işgal ederek[5] veya çevre dokulara baskı uygulayarak nörolojik fonksiyon kayıplarına yol açabilir.

Hastalığın küresel yükü ve epidemiyolojik dağılımı incelendiğinde, beyin tümörlerinin insidansının demografik ve coğrafi faktörlere göre belirgin farklılıklar gösterdiği görülmektedir. Uluslararası Kanser Araştırmaları Ajansı'nın (IARC) sağladığı GLOBOCAN 2022 verilerine göre, dünya genelinde her yıl yaklaşık 321.476 kişiye yeni beyin ve merkezi sinir sistemi tümörü tanısı konulmakta ve bu vakaların büyük bir kısmı mortalite ile sonuçlanmaktadır [7]. Gelişmiş

[1] işlevsel sinir dokusu
[2] Beyni saran zarlardır.
[3] Yeni ve anormal doku oluşumu.
[4] Kontrolsüz çoğalma.
[5] İnvazyon.

ülkelerde, tanısal görüntüleme yöntemlerine (MRG, Bilgisayarlı Tomografi (BT)) erişimin yaygın olması nedeniyle insidans oranları daha yüksek raporlanırken (Kuzey Avrupa ve Kuzey Amerika'da 100.000'de 5-6 vaka), tanı imkanlarının kısıtlı olduğu bölgelerde bu oranlar daha düşüktür. İstatistikler, beyin tümörlerinin her yaş grubunu etkileyebilmekle birlikte, insidansın yaşla birlikte arttığını ve özellikle 65 yaş üstü popülasyonda zirve yaptığını göstermektedir [7].

2.2 Beyin Tümörlerinin Sınıflandırılması ve Derecelendirme Sistemleri

Beyin tümörlerinin doğru sınıflandırılması, hastanın tedavi protokolünün (cerrahi, radyoterapi, kemoterapi) belirlenmesinde ve sağkalım beklentisinin (prognoz) öngörülmesinde en kritik aşamadır. Geleneksel olarak sadece mikroskop altındaki doku görünümüne dayanan sınıflandırma, günümüzde genetik ve moleküler belirteçlerin de dahil edildiği bütünleşmiş tanı sistemine evrilmiştir.

Beyin tümörleri diğer tümörler gibi primer, beyin dokusunda oluşan tümörler ve sekonder(metastatik), vücudun başka bir bölgesinde başlayan ve kan dolaşımı yoluyla beyne sıçrayan tümörler olabilir. Ayrıca büyüme hızı ve çevre dokularla olan etkileşimine göre (huyuna göre), benign (iyi huylu), yavaş büyüme hızına (düşük proliferasyon indeksi) sahip olan tümörler ve malign (kötü huylu), hızlı ve agresif bir büyüme paterni gösteren tümörler olarak sınıflandırılabilir.

Tümörler huylarına göre Dünya Sağlık Örgütü (WHO) derecelendirme sistemine göre derecelendirilerek inceleniler.

2.2.1 WHO Derecelendirme Sistemi (Grading) ve Moleküler Devrim

WHO, tümörün beklenen doğal seyri üzerinden klinik agresifliğini (büyüme, invazyon, nüks ve sağ kalım) standartlaştıran bir derecelendirme sistemi kullanır. WHO'nun merkezi sinir sistemi (CNS) derecelendirmesi CNS WHO Grade olarak adlandırılır. WHO CNS ile iki önemli değişiklik yapmıştır: dereceler Romen rakamları yerine Arap rakamlarıyla yazılır ve derecelendirme artık tümör tipleri arasında değil, her tümör tipinin kendi biyolojik spektrumu içerisinde verilir [6]. WHO derecelendirmesi tek bir bulguya değil; histopatolojik ve gerektiğinde moleküler belirteçlerin birlikte derecelendirmesine dayanır. Bu nedenle derecelendirme, pratikte çoğunlukla biyopsi/rezeksiyon materyalinin patoloji incelmesiyle yapılır; görüntüleme (MRG) derece hakkında ipucu verebilse de tek başına nihai derece belirtmez [3],[4]. Tablo 2.1 de farklı derecelerin özellikleri özetlenmiştir.

Tablo 2.1 WHO Tümör Derecelendirme Sistemi ve Özellikleri

WHO Derecesi	**Büyüme Hızı**	**Yayılım (İnvazyon)**	**Prognoz (Seyir)**
Grade 1	Çok Yavaş	Yok (İyi sınırlı)	Mükemmel (Cerrahi ile kür)
Grade 2	Yavaş	Var (İnfiltratif olabilir)	İyi (Ancak nüks riski var)
Grade 3	Hızlı	Belirgin	Orta/Kötü (Anaplastik özellikler)
Grade 4	Çok Hızlı	Yaygın İnvazyon	Kötü (Nekroz ve damarlanma var)

2.2.2 Doku Kökenine Göre Tümör Tipleri

Beyin ve omurilikten oluşan merkezi sinir sistemi (MSS) tümörleri; doğrudan sinir dokusunun kendisinden ya da bu dokuyu çevreleyen/destekleyen yapılardan gelişebilen, kökeni çok farklı olabilen bir hastalık grubudur. MSS içinde en sık kaynak dokular; glial destek hücreleri (nöronların çalışmasını sürdürebilmesi için altyapı ve bakım görevleri gören hücreler)i beyin zarları (meninksler; beyin ve omuriliğini saran koruyucu katmanlar), beyin boşluklarını döşeyen ependim tabakası (ventriküllerin iç yüzünü kaplayan hücresel hat) ve hipofiz bezi (beynin tabanında hormonların salınımını düzenleyen bez) gibi yapılardır. Bu nedenle beyin tümörü ifadesi, tek bir hastalıktan çok; oluştukları dokuya ve biyolojik davranışına göre sınıflana geniş bir tümör ailesini temsil eder [6].

Bu çalışmada kullanılan veri setlerinde bahsi geçen tümör tipleri olan gliom, Menenjiom ve pitnet beyin tümörü türleri aşağıdaki özelliklerdedir.

Gliomlar, beyin dokusunun yapısal ve metabolik desteğini sağlayan glia hücrelerinden (astrositler, oligodendrositler) köken alan tümörlerdir. Primer beyin tümörlerinin en büyük grubunu oluştururlar, Gliomlar çeşitli alt gruplarda incelenir. Yıldız şeklindeki astrositlerde türeyen astrositomlar, düşük dereceli (Grade 2) olabileceği gibi, yüksek dereceli (Grade 3-4) formlara evrilebilir. Oligodendrositlerden kaynaklanan oligodengliomlar, moleküler tanıda 1p/19q kodelasyonu varlığı ile tanımlanır. Astrosit kökenli tümörlerin en agresifi olan glioblastom (GBM- WHO Grade 4) iki temel özelliği vardır. Nekroz (tümörün hızlı büyümesi sonucu merkezdeki hücrelerin beslenemeyip ölmesi) ve Mikrovasküler Proliferasyon (tümörün kendi damar ağını oluşturması) Radyolojik olarak Ring Enhancement (halka tarzı parlaklaşma) gösterirler. İzositrat dehidrogenaz (IDH) mutasyonu, WHO CNS5 (2021) sınıflamasında özellikle erişkin tip difüz gliomaların ayrımında temel bir moleküler belirteçtir; bu grup üç ana tanı birimine ayrılır: Astrositom, IDH-mutant; Oligodendrogliom, IDH-mutant ve 1p/19q-kodelezyonlu[6]; Glioblastom, IDH-wildtype [6].

[6]Bir tümör hücresinin iki kromozomunun bir kısmını kaybettiği anlamına gelir;1. kromozomun kısa kolu (1p olarak adlandırılır.) ve 19. kromozomun uzun kolu (19q olarak adlandırılır)

Menenjiom (Menengioma), beyin dokusunun (parankim) içinden değil; beyni saran zarların araknoid tabakasındaki meningotelyal/araknoid kap hücrelerinden gelişen ve bu yüzden çoğu zaman beyin dokusu dışında (ekstraaksiyel) yerleşen bir tümördür. Klinik etkisi çoğunlukla içeriye sızmadan çok, bulunduğu yerde yer kaplayıp beyni itmesi/sıkıştırması (kitle etkisi) ile ortaya çıkar. WHO derecelendirmesinde olguların büyük kısmı Grade 1'dir (yaklaşık %90); daha küçük bir bölüm Grade 2 (atipik) ve Grade 3 (anaplastik) gruba girer [8].

Hipofiz bezinde (pituiter) görülen tümörlerin klinik açıdan en önemli grubu, ön hipofiz (adenohipofiz) kökenli ve güncel WHO yaklaşımında hipofizer nöroendokrin tümör (PitNET) olarak adlandırılan neoplazmlardır (eski terminolojide hipofiz adenomu). Bu tümörler iki ana yolla sorun çıkarır: (i) hormon salgılayarak (ör. prolaktin, büyüme hormonu, ACTH gibi) vücutta sistemik tablolar oluşturabilir; (ii) hormon üretimi belirgin olmasa bile kitle etkisiyle (özellikle makroadenom boyutlarına ulaştığında) çevre yapılara bası yaparak baş ağrısı, görme alanı kusuru gibi nörolojik yakınmalara yol açabilir. Klasik literatürde bu hormonal etki ve kitle etkisi ikiliğinin çok erken dönemde fark edildiği vurgulanır [9].

WHO 2022 çerçevesi PitNET'lerin biyolojik davranışının zararsız/intrasellar olgulardan, invaziv–tedaviye dirençli (agresif) olgulara ve çok nadiren metastaz yapan olgulara kadar geniş bir spektrumda değişebildiğini özellikle öne çıkarır; bu nedenle eski pituitary carcinoma (hipofiz karsinomu) yerine metastatik PitNET terimi önerilmiştir. Metastatik PitNET, lenf nodu veya uzak organ metastazı ya da MSS içinde kesintili yayılım gösteren PitNET olarak tanımlanır. Önemli bir pratik sonuç: tek bir klinik, radyolojik veya histolojik ölçüt tümörün gelecekteki büyüme/agresifleşme riskini güvenilir biçimde tek başına belirleyemez; ayrıca hipofiz tümörleri için diğer organ NET'lerinde olduğu gibi evrensel, oturmuş bir derecelendirme sistemi bulunmadığı da vurgulanır [9].

2.2.3 Tanı ve Görüntüleme Yöntemleri

Beyin tümörlerinde tanı yaklaşımı çoğunlukla klinik şüphe ile başlar (yeni başlayan nöbet, ilerleyici baş ağrısı, fokal nörolojik defisit, bilişsel/kişilik değişiklikleri, artmış kafa içi basınç bulguları vb.). İlk hedef, hastada acil bir tablo (kanama, herniasyon riski, hidrosefali) olup olmadığını hızlıca değerlendirmek ve uygun görüntülemeyi seçmektir. Acil durum şüphesinde BT[7], hız ve erişilebilirlik nedeniyle pratik bir ilk basamak olabilir; akut kanama ve kaba kitle etkisini hızlı gösterebilir [10]. Klinik stabil hastada veya tümör şüphesinin ön planda olduğu

[7]Computed Tomography- CT olarakda kaynaklarda görülür.

senaryoda ise temel inceleme kontrastlı/kontrastsız beyin MR'dır [11].

MRG değerlendirmesi tek görüntü değil, sekans-temelli bir protokolle yürür: Kontrast madde olmadan T1-ağırlıklı görüntüler anatomik referans sağlar; T2 ve FLAIR sekansları lezyonun varlığını, çevresel ödemi/kitle etkisini ve olası infiltratif bileşeni daha iyi ortaya koyar [10]. Şüpheli kitle saptandıktan sonra kontrastlı T1 (T1+CE) istenerek kan-beyin bariyeri bozulmasına bağlı kontrast tutulumu (solid/heterojen/halka tarzı vb.) değerlendirilir; bu bulgular tek başına kesin derece koydurmasa da lezyonun biyolojisi, cerrahi planlama ve biyopsi hedeflemesi için kritik ipuçları verir [10, 11]. Gerekli olgularda difüzyon (DWI/ADC), perfüzyon Manyetik Rezonans (MR) ve MR spektroskopi gibi ileri teknikler; tümör ve tümör taklidi durumlarının ayırıcı tanısına, daha agresif bölgelerin örneklenmesine ve tedavi sonrası değişikliklerin yorumlanmasına yardımcı ek bilgi sağlayabilir [10].

Görüntülemenin rolü yalnızca tümör var/yok değildir: Lezyonun yerleşimi (intraaksiyel–ekstraaksiyel), ventrikül ve damar yapılarıyla ilişkisi, kitle etkisi, ödem ve olası cerrahi riskler bu aşamada haritalanır ve çok disiplinli kararı yönlendirir [10]. Ancak güncel nöro-onkolojide nihai tanı, yalnızca görüntülemeye değil; histopatoloji + moleküler belirteçlerin birlikte değerlendirilmesine dayanır. Bu nedenle mümkünse rezeksiyon, uygun değilse stereotaktik biyopsi ile doku tanısı alınarak tanı kesinleştirilir ve tedavi planı buna göre şekillendirilir [11].

Bu tez çalışmasının ana konusu olan MRG süreci sıradaki bölümde daha ayrıntılı olarak anlatılmıştır.

2.3 Manyetik Rezonans Görüntüleme

MRG, vücuda iyonlaştırıcı radyasyon vermeden doku anatomisini ve işlevini incelemek için kullanılan klinik bir araçtır. Bu teknoloji dokuların yapısı ve biyokimyasal içeriği hakkında çok boyutlu bilgi sunmak için yapısal, difüzyon veya spektroskopik sekansları bir arada kullanabilir [12]. MRG, nesnenin doğrudan görsel kopyasını alan bir fotoğraf makinesi değil; dokudan gelen elektromanyetik verileri işleyerek karmaşık bir sinyal dönüştürücü ve rekonstrüksiyon sistemidir. Bu sistem, vücuttaki hidrojen moleküllerinin hareketlerinden oluşan sinyalleri üç boyutlu bir görüntü ortamına dönüştürür [13].

2.3.1 Manyetik Rezonans Görüntülemenin Tarihçesi

MRG teknolojisinin fiziksel temeli olan Nükleer Manyetik Rezonans (NMR), ilk kez 1946 yılında Felix Bloch [14] ve Edward Purcell [15] tarafından birbirinden

bağımsız olarak deneysel olarak tanımlanmıştır. NMR, belirli atom çekirdeklerinin (en yaygın olarak hidrojen protonlarının) güçlü bir statik manyetik alan içindeyken elektromanyetik radyofrekans enerjisini emme ve tekrar geri yayma durumudur. Bu aşamada maddelerin kimyasal bileşimini incelmeye yarayan bir spektroskopi yöntemiydi.

NMR'ın Görüntülemeye geçişteki devrimsel adım, 1973 yılında Paul Lauterbur'un manyetik alan gradyanlarını kullanmayı önermesiyle atılmıştır [16]. Bu yenilik, makinelerin sadece bir sinyal kaydetmek yerine, uzamsal bilgiyi (konum bilgisini) sinyalin içine kodlamasına olanak tanıyan bir mühendislik yapısına kavuşmasını sağlamıştır [13].

2.3.2 MRG Makinalarının Bileşenleri

Bir MRG cihazı temelde ana mıknatıs, çeşitli bobinler, soğutma sistemi ve dış frekans engelleyicilerden oluşur.

Ana mıknatıs sistemi, sistemin temel elemanıdır ve ihtiyaç duyulan manyetik alanı (B_0) oluşturur. Genellikle 0.5 ile 3 Tesla arasında statik ve homojen bir manyetik alan üretir. Vücuttaki hidrojen atomlarının (protonların) manyetik momentlerini bir ana alan yönünde hizalar [17].

Gradyan Bobinleri (G_x, G_y, G_z), ana mıknatısın içine yerleştirilmiş, üç dik eksende (x, y, z)) çalışan dirençli bobinlerdir. Manyetik alanı konuma göre kontrollü değiştirerek uzamsal kodlama (konum belirleme) yapılmasını sağlarlar. X ve Ygradyanları faz ve frekans kodlamasını gerçekleştirirken, Z gradyanı genellikle kesit seçimi için kullanılır. Bu sayede sinyalin vücudun hangi noktasından geldiği saptanır [17].

Radyofrekans (RF) bobinleri, sistemin antenleri olarak görev yaparlar. Verici bobinler (transmitter) protonları uyarmak için Larmor frekansında kısa süreli RF darbeleri (B_1 alanı) gönderir ve protonların yüksek enerji seviyesine geçmesini sağlarlar [17]. Alıcı bobinler (receiver) uyarılmış protonlar eski hallerine (dengeye) dönerken yaydıkları zayıf elektromanyetik sinyalleri (FID) yakalar.

Soğutma sistemleri (kriyostat), süper iletken mıknatıs sargılarının dirençsiz çalışabilmesi için gereken bir termos yapısıdır. Sargıları mutlak sıfıra yakın sıcaklıklarda tutmaya çalışır. Bu amaçla genellikle sıvı helyum kullanılır; helyum sargıları $-268,9°C$ civarında soğutarak akımın enerji kaybı olmadan çalışmasını sağlar.

RF zırhlama (faraday kafesi), dışarıdaki radyo dalgalarının MR sinyalleriyle karışmasını önler ve cihazın ürettiği RF sinyallerinin dışarı kaçmasını engeller.

Aktif/Pasif manyetik zırhlama, cihazın ürettiği güçlü manyetik alanın oda dışındaki elektronik cihazlara zarar vermesini engellemek için kullanılır.

2.3.3 Manyetik Rezonans Görüntü Süreci

MR görüntüleri ile çalışılırken görüntülerin nasıl oluştuğunun ve hangi bilgilerin okunduğunun anlaşılması gerekir. Görüntüleme süreci, hidrojen çekirdeklerinin statik manyetik alan altındaki davranışlarını yöneten kuantum mekaniksel fizik yasalarıyla başlar. Bu yasaların temelinde Larmor Presesyonu ve Relaksasyon (durulma) süreçleri yatar.

2.3.3.1 Larmor Eşitliği ve Rezonans

Güçlü bir statik manyetik alan (B_0) altına yerleştirilen protonlar, bu alanın yönü etrafında belirli bir açısal frekansla dönerler. MR fiziğinin temeli olan bu dönüş frekansı, Larmor Eşitliği ((2.1))ile ifade edilir [18]:

$$\omega_0 = \gamma B_0 \tag{2.1}$$

Burada ω_0 açısal frekans (MHz), B_0 manyetik alan şiddeti (Tesla) ve γ ise hidrojen çekirdeğine özgü jiromanyetik orandır (MHz/Tesla). Bir hidrojen protonu için jiromanyetik oran 42,6 MHz/T'dır. Klinik MR cihazları genellikle 1.5T veya 3T gücündedir; 1.5T bir cihazda protonlar yaklaşık 64 MHz civarında dönerler [18].

2.3.3.2 RF Uyarımı ve Rezonans

B_0 manyetik alanı içinde presesyon yapan protonlar, henüz dışarıdan ölçülebilir bir sinyal üretmezler. Sinyal alabilmek için sisteme dışarıdan enerji verilmesi gerekir. Bu enerji transferi, Larmor frekansıyla (ω_0) tam uyumlu bir Radyo Frekans (RF) darbesi gönderilerek sağlanır. Buna Rezonans denir [18].

RF darbesi uygulandığında iki temel fiziksel olay gerçekleşir:

Enerji emilimi, protonlar RF enerjisini emerek düşük enerji seviyesinden yüksek enerji seviyesine geçerler.

Faz uyumu (Phase Coherence), rastgele fazlarda dönen protonlar, RF darbesinin etkisiyle eşzamanlı (aynı fazda) dönmeye başlarlar.

Bu sürecin sonucunda, net manyetizasyon vektörü (M_z), B_0 yönünden ayrılarak enine düzleme (xy-düzlemi) doğru yatar. Bu yatma açısına devrilme açısı (flip angle) denir. Sadece enine düzlemde dönen manyetizasyon (M_{xy}), alıcı bobinlerde bir voltaj indükleyebilir ve MR sinyalini oluşturur [18].

2.3.3.3 Relaksasyon Süreçleri ve Dokular

RF uyarımı kesildiğinde, protonlar aldıkları enerjiyi geri vererek eski denge konumlarına dönmeye çalışırlar. Bu durulma sürecine Relaksasyon denir ve görüntüdeki kontrast dokuların durulma sürelerinin farklı olmasından kaynaklanır [18].

T1 relaksasyonu (boyuna durulma); Spinlerin enerjilerini çevre dokuya aktararak tekrar boyuna manyetizasyon dengesine (M_z) ulaşmasıdır. Matematiksel modeli eşitlik (2.2)'deki gibidir [19].

$$M_z(t) = M_0\left(1 - e^{-t/T1}\right) \tag{2.2}$$

Yağ doksu hızlı toparlanır (kısa T1), su yavaş toparlanır (uzun T1).

T2 relaksasyonu (enine durulma); spinlerin kendi aralarındaki etkileşimi sonucu sinyalin (M_{xy}) sönümlemesidir. Matematiksel modeli şu şekildedir [19].

$$M_{xy}(t) = M_{xy}(0)\, e^{-t/T2} \tag{2.3}$$

Su molekülleri sinyali uzun süre korur (uzun T2 relaksasyonu).

2.3.3.4 Görüntü Ağırlıklandırma Parametreleri (TR, TE ve TI)

T1 ve T2 relaksasyonları dokunun değişmez özellikleridir. Ancak hangi dokunun parlak görüneceğini belirlemek için MR cihazında zamanlama parametreleri ayarlanır. TR (Repetition Time – Tekrarkama Zamanı); Bir RF uyarımı bir sonraki RF uyarımı arasında geçen süresir. TR, dokuların ne kadar toparlamasına (T1 relaksasyonuna) izin verileceğini belirler. Kısa TR, dokuların tam toparlanmasına izin verilmez. Bu sayede hızşı toparlanan (yağ/beyaz cevher) ile yavaş toparlanan (ör. su) arasındaki fark maksimize edilir. Bu nedenle T1 sekanslı görüntülerde TR kısa tuutlur [18].

TE (Time to Echo – Ekho Zamanı); RF uyarımı yapıldığı an ile sinyalin (ekho) okunduğu an arasında geçen sürendir. TE, sinyalin ne kadar sönümlenmesine izin

verileceğini belirler. Uzun TE, sinyalin sönümlemesi için süre tanır. Sinyali hızlı kaybedeb dokular kararırken, suyu tutan (ödem/tümör) dokular hala parlak kalır. Bu nedenle T2 sekanslı görüntülerde TE uzun tutulur [18].

TI (Inversion Time – Ters Çevirme Zamanı); 180°'lik hazırlık darbesinden sonra görüntünün alınması için beklenen süredir. T1 süresi, baskılanmak istenen dokunun manyetizasyonunun sıfır noktasına geldiği ana göre ayarlanır. Böylece o dokudan sinyal alınmaz ve görüntüde siyah çıkar [10], [18].

2.3.3.5 Uzaysal Kodlama ve K-Uzayı

MR sinyali, bobin tarafından algılandığında tüm hacimden gelen toplam sinyali içerir ve uzaysal bilgi barındırmaz. Görüntü rekonstrüksiyonu için sinyalin x, $y$ve $z$eksenlerinde kodlanması gerekmektedir. Bu işlem, ana manyetik alan (B_0) üzerine eklenen ve manyetik alan şiddetini doğrusal (lineer) olarak değiştiren gradyan bobinleri ile gerçekleştirilir.

Bir noktadaki manyetik alanın konuma bağlı değişimi ve buna bağlı rezonans frekansı şu denklemle ifade edilir:

$$\omega(r) = \gamma(B_0 + G \cdot r) \tag{2.4}$$

Burada G gradyan şiddeti (mT/m) ve ise $r$konumdur. Bu fiziksel prensip kullanılarak uzaysal kodlama üç temel aşamada gerçekleştirilir [18, 20]:

Kesit Seçimi (Slice Selection - G_z): Görüntülenecek anatomik düzlemi belirlemek için, RF uyarımı sırasında z -ekseni gradyanı (G_z) aktif hale getirilir. Sadece belirli bir frekanstaki (konumdaki) protonlar rezonansa girer.

Faz kodlama (Phase Encoding -G_y): sinyalin satır bazında (y-ekseni) konumunu belirlemek için G_y gradyanı uygulanır. Bu işlem, protonlar arasında konuma bağlı bir faz farkı (phase shift) oluşturur.

Frekans kodlama (Frequency Encoding -G_x): Sinyalin okunması sırasında x -ekseni gradyanı (G_x) devreye sokulur. Böylece farklı sütunlardaki protonlar farklı frekanslarda sinyal yayar.

2.3.3.6 Sinyal İşleme ve Görüntü Oluşumu

Gradyan bobinleri ile kodlanan ve antenler tarafından toplanan ham sinyaller, doğrudan insan gözünün algılayabileceği bir görüntü değildir. Bu veriler önce

k-uzayı (k-space) adı verilen bir frekans haritasında depolanır. Görüntünün elde edilmesi için bu matematiksel verinin işlenmesi gerekir [18]. Bu görüntü oluşum süreci şu şekildedir:

Ters fourier dönüşümü (Inverse FFT); k-uzayındaki frekans verilerine ters hızlı fourier dönüşümü uygulanır. Bu işlem, frekans alanındaki karmaşık sayıları, uzaysal alandaki sinyal şiddeti değerlerine dönüştürür.

Piksel matrisinin oluşumu; fourier dönüşümü sonucunda, her bir hücresi (pikseli) belirli bir sayısal değere sahip olan 2 Boyutlu bir matris elde edilir. Bu matristeki her sayı, o noktadaki dokudan gelen MR sinyalinin gücünü (amplitüdünü) temsil eder.

Gri skala (Grayscale) haritalama; matristeki sayısal değerler, 0 (siyah) ile 255 (beyaz) arasında değişen gri tonlara dönüştürülür:

- Yüksek sinyal (yüksek sayısal değer): ekranda parlak/beyaz (hiperintens) görünür.
- düşük sinyal (düşük sayısal değer): ekranda koyu/gri (hipointens) görünür.
- Sinyal yok (sıfır): ekranda siyah (background) görünür.

Sonuç olarak; derin öğrenme modelinin işleyeceği görüntü, aslında protonların yoğunluğuna ve relaksasyon sürelerine (T1/T2) göre parlaklık değerleri atanmış bu sayısal matristir.

2.3.3.7 MRG Sekansları

Elde edilen bu piksel matrisindeki parlaklık dağılımı, çekim sırasında ayarlanan TR, TE ve TI sürelerine göre değişir. Derin öğrenme modelleri, aynı anatomik bölgenin farklı parametrelerle (farklı parlaklıklarla) elde edilmiş versiyonlarını kullanarak doku ve anamoli ayrımı yapar.

T1 sekansı; ödem, tümör, enfeksiyon, düşük proton ağırlığı, kalsifikasyonlar ve akımın olmadığı yerlerin koyu görünmesi, Yağ, melanin protein zengin sıvıların, yavaş akım olan dokuların ve paramanyetik maddelerin (gadolinyum, bakır, manganez) dokuların daha parlak görünmesi olarak gerçekleşir [20]. Gadolinyum gibi maddelerin bu sekansta daha parlak olmasından dolayı T1 sekansta ek olarak kontrastlı T1 görüntülerde alınır. Bu durum genellikle T1 sekansta gözlemlenen anamoliye göre uzman tarafından istenerek hastaya kontrast madde enjekte edilmesi ile gerçekleştirilir.

Tablo 2.2 Sekansların özelliklerini özetleyen tablo

Sekans Türü	**Parametreler**	**Fiziksel Amaç**	**Doku Görünümü (Tümör/Ödem)**
T1 Ağırlıklı	Kısa TR Kısa TE	Anatomik yapıyı göstermek.	Koyu (Hipointens)
T2 Ağırlıklı	Uzun TR Uzun TE	Su içeriğini (patolojiyi) parlatmak.	Parlak (Hiperintens)
FLAIR	Uzun TR Uzun TE Uzun TI	BOS'u silip, sadece ödemi parlatmak.	Parlak (Hiperintens)

T2 sekanslarda, düşük proton içeren yapılar, kalsifikasyon, fibröz doku, paramanyetik maddeler (deoksihemoglobin, methemoglobin (intraselüler[8]), ferritin, hemosiderin, melanin, proteinden zengin sıvılar ve akımın olduğu yerler koyu olarak, Ödem, tümör, enfeksiyon, inflamasyon, subdural birikim gibi yapılar parlak olarak görünür [20].

IR (Inversiyon Recovery), TI (inversiyon zamanı) ve TR uzun tutulduğunda sıvı baskılanır bu şekilde Fluid Attenuasyon Invertion Recovery -FLAIR sekansı elde edilir. Ödem bu sekansta daha kolay tespit edilir [20].

Bunlara ek olarak farklı dokuların ve hastalıkların tespiti için farklı sekanslarda kullanılır. Bu çalışmada bahsi geçecek olan sekansları sadece tanımlanmıştır.

[8]Hücre içi.

3
DERİN ÖĞRENMENİN MATEMATİKSEL TEMELLERİ

Derin öğrenme (deep learning), temelini yapay sinir ağlarının (YSA) çok katmanlı (derin) mimarilerinden alan ve temsili (özellikleri) veriden hiyerarşik biçimde öğrenmeyi hedefleyen makine öğrenmesinin bir alt dalıdır [21]. Bu yaklaşım, özellikle ham verinin (piksel, dalga formu, metin, çoklu-sensör ölçümleri vb.) doğrudan modele verilerek, problem için gerekli ara temsillerin (features/representations) insan eliyle tasarlanmasına olan bağımlılığı azaltmasıyla öne çıkar [21].

Makine öğrenmesi (machine learning), ayarlanabilir parametrelere sahip bir sistemin veriyle eğitilerek belirli bir görevdeki çıktısını iyileştirmesi ilkesine dayanır. En yaygın senaryo olan denetimli öğrenmede, her örnek bir etiketle birlikte toplanır; model eğitim sırasında girdiyi görür, bir çıktı üretir ve hedefle arasındaki hatayı azaltacak şekilde iç parametrelerini günceller. Böylece sistem, eğitimde görmediği yeni örneklere de anlamlı yanıt üretebilecek bir genelleme davranışı kazanır [21].

Derin öğrenme ile geleneksel makine öğrenmesi arasındaki temel farklardan biri, özellik bilgisinin (feature) edinilme biçimidir: Geleneksel yaklaşımlarda çoğu zaman ham veriden öğrenmeye uygun bir vektör elde etmek için, problem alanına özgü bir özellik çıkarıcı tasarlanır; bu adım dikkatli mühendislik ve alan bilgisi gerektirebilir. Derin öğrenme ise ham girdiden başlayıp katmanlar boyunca ilerleyen temsilleri öğrenerek, gerekli ayrıştırıcı bilgiyi büyük ölçüde modelin içinde üretir; böylece elle özellik tasarımına bağımlılık önemli ölçüde azalır (alan bilgisinin tamamen gereksiz olduğu iddia edilmez, fakat bağımlılık düşer) [21].

Derin öğrenme, veri hacmi ve hesaplama kapasitesindeki (özellikle paralel hesaplama tabanlı) artışla birlikte ölçeklenebilir hâle gelmiş; bu ölçeklenebilirlik, yöntemlerin yalnızca teorik bir çerçeve olmaktan çıkıp geniş uygulama sınıflarında

etkin biçimde kullanılabilmesini mümkün kılmıştır. Nitekim derin öğrenme yaklaşımları; bilgisayarlı görü (görsel nesne tanıma ve nesne tespiti), konuşma işleme ve doğal dil işleme gibi algısal ve sembolik verinin birlikte ele alındığı alanlarda güçlü sonuçlar üretmiş; bunun yanı sıra biyomedikal keşif odaklı problemlerde (ör. ilaç keşfi, genomik) de yaygın kullanım alanı bulmuştur [21].

Tıbbi görüntüleme ve sağlık bağlamında derin öğrenme, özellikle görüntü sınıflandırma, lezyon/tümör tespiti, segmentasyon ve kayıt/registration gibi klinik iş akışlarıyla doğrudan ilişkili görevlerde önemli faydalar sağlamıştır. Bu fayda, görüntü verisinin yüksek boyutlu ve karmaşık uzamsal örüntüler içermesi nedeniyle, çok katmanlı temsillerin klinik açıdan anlamlı ayrımları yakalamaya elverişli olmasından kaynaklanır. Bununla birlikte sağlık uygulamalarında yöntem seçimi ve değerlendirme tasarımı; genellenebilirlik, güvenilirlik ve uzmanların çıktıyı anlamlandırabilmesi gibi gereksinimlerle birlikte ele alınmak zorundadır [21].

Özellikle MR görüntülerinden beyin tümörü analizi bağlamında, tümör dokusunun heterojen yapısı ve farklı sekanslarda farklı biyofiziksel kontrastların ortaya çıkması; tespit ve ayrıştırma (tümör var/yok, alt bölge ayrımı, sınırların belirlenmesi) problemlerini hem klinik hem de hesaplamalı açıdan kritik hâle getirmektedir. Bu çalışmada, derin öğrenmenin MR görüntülerinden beyin tümörü tespiti ğroblemindeki kullanımını; veritemsili, model seçimi ve performans değerlendirme ölçitleri bağlamında eke akarak, yöntemlerşn klinik problem sınıfına sağladığı katkı tartışmalı bir çerçevede değerlendirmek amaçlanmaktadır.

3.1 Derin Öğrenmenin Tarihsel Gelişimi ve Kırılma Noktaları

Derin öğrenmenin tarihsel serüveni, insan beyninin bilgi işleme mekanizmasını matematiksel olarak modelleme çabalarıyla başlamış; donanım, veri ve algoritmaların eşzamanlı gelişimiyle günümüzdeki baskın formuna ulaşmıştır. Bu alandaki ilk somut adım, 1943 yılında Warren McCulloch ve Walter Pitts tarafından atılmış; biyolojik bir nöronun hep ya da hiç(all-or-none) prensibiyle çalışan mantıksal bir devre olarak modellenebileceği gösterilmiştir [22]. Bu teorik model, 1958 yılında Frank Rosenblatt tarafından Perseptron (Perceptron) adıyla öğrenilebilir bir algoritmaya dönüştürülmüştür [23]. Perseptron, il başlarda ümit vaadedici görünsede, 1969 yılında Minsky ve Papert'in, tek katmanlı ağların XOR (Özel Veya) gibi basit doğrusal olmayan problemleri çözemeyeceğini kanıtlamasıyla alan popülerliğini kaybetmiş ve bu alandaki yatırımların azalmasıyla, literatürde yapay zeka kışı olarak bilinen ilk duraklama

dönemine girmiştir [24]. Bu durgunluk, 1986 yılında Rumelhart ve ekibinin, hatanın ağ boyunca geriye doğru iletilmesini sağlayan geri yayılım (backpropagation) algoritmasını çok katmanlı ağlara uyarlamasıyla aşılmıştır [25].

CNN mimarisinin gelişimi, görüntü işlemede derin mimarilerin temeli, Hubel ve Wiesel'in kedilerin görme korteksindeki hiyerarşik yapıyı (basit ve karmaşık hücreler) keşfetmesine dayanır. Bu biyolojik keşfi matematiksel bir modele dönüştüren ilk kişi Kunihiko Fukushima olmuştur. Fukushima, 1980 yılında geliştirdiği Neocognitron modeli ile, bugün CNN'lerde kullanılan evrişim (convolution) ve havuzlama (pooling) katmanlarının çalışma prensibini literatüre kazandırmıştır [26]. Ancak Neocognitron, denetimli öğrenme için verimli bir hata düzeltme mekanizmasına sahip değildi. Modern CNN'lerin doğuşu, Fukushima'nın bu mimarisinin geri yayılım (backpropagation) algoritması ile birleştirilmesiyle gerçekleşmiştir. 1989 yılında Yann LeCun ve ekibi, geri yayılım algoritmasını evrişimli ağlara entegre ederek LeNet-5 mimarisini geliştirmiş ve el yazısı rakamların (MNIST) tanınmasında büyük başarı elde etmiştir [27]. LeCun'un çalışması, derin ağların verimli bir şekilde eğitilebileceğini kanıtlayan ilk somut örnek olmuştur.

Derin öğrenmenin modern çağını başlatan ve tıbbi görüntülemede standart hale gelmesini sağlayan asıl kırılma noktası 2012 yılında yaşanmıştır. ImageNet veri seti üzerinde düzenlenen yarışmada, Alex Krizhevsky ve arkadaşları tarafından geliştirilen AlexNet, GPU gücünü ve ReLU aktivasyon fonksiyonunu kullanarak hata oranını dramatik şekilde düşürmüş ve derin öğrenme patlamasını tetiklemiştir [28]. Bu başarıyı 2015 yılında Kaiming He ve ekibinin geliştirdiği ResNet (Residual Networks) izlemiştir. ResNet, atlamalı bağlantılar (Skip Connections) kullanarak çok derin ağlardaki kaybolan gradyan problemini çözmüş ve insan seviyesini aşan sınıflandırma performansı sunmuştur [3].

Sınıflandırma başarısı artarken, tıbbi görüntüleme için asıl ihtiyaç duyulan hastalığın sınırlarını çizme (segmentasyon) sorunu, yine 2015 yılında Olaf Ronneberger ve ekibi tarafından sunulan U-Net mimarisi ile çözülmüştür [29]. U-Net, kodlayıcı-kod çözücü (encoder-decoder) yapısı ve katmanlar arası atlamalı bağlantıları sayesinde, az sayıda etiketli medikal veri ile çok yüksek doğrulukta segmentasyon yapabilmektedir. Bu başarı, U-Net ve türevlerini günümüzde medikal görüntü segmentasyonunun altın standardı haline getirmiştir.

CNN mimarilerinin görüntü işlemedeki başarısına rağmen, temel yapı taşları olan konvolüsyon işleminin yerel alıcı alanlara (local receptive field) odaklanması, görüntü üzerindeki uzak mesafeli anlamsal ilişkilerin (long-range dependencies)

modellenmesinde sınırlılıklar yaratmaktadır. Bu kısıtlılık, derin öğrenme literatüründe yeni bir paradigma değişimini tetiklemiş ve 2017 yılında Vaswani ve ekibi tarafından Doğal Dil İşleme (NLP) alanı için önerilen Transformer mimarisi ile öz dikkat (self-attention) mekanizması bilgisayarlı görü alanına taşınmıştır [30].

Bu yaklaşımın görüntü analizindeki ilk başarılı uyarlaması, 2020 yılında Dosovitskiy ve ekibi tarafından geliştirilen Vision Transformer (ViT) modeli olmuştur. ViT, görüntüyü pikseller yerine yamalar (patches) dizisi olarak işleyerek global bağlamı yakalamada üstün performans sergilemiştir [31].

Günümüzde ise araştırmalar tek bir mimariye bağlı kalmak yerine; CNN'lerin yerel öznitelik çıkarma gücü (inductive bias) ile Transformer'ların global ilişki kurma yeteneğini birleştiren hibrit mimariler geliştirmek üzerine yoğunlaşmıştır. Beyin tümörü tespiti gibi hem yapısal detayların hem de anatomik bütünlüğün kritik olduğu alanlarda, bu hibrit yaklaşımlar ve geliştirilmeye devam eden yeni nesil modeller, tanısal doğruluk oranlarını daha ileriye taşımayı hedeflemektedir.

3.2 Yapay Sinir Ağları (YSA)

Derin öğrenme mimarilerinin temel yapı taşı ve en küçük işlem birimi, biyolojik sinir hücresinin (nöron) çalışma prensibinden esinlenerek geliştirilen Perseptron (yapay nöron)'dur. Bu kavramın teorik temeli, McCulloch ve Pitts (1943) tarafından önerilen ve bir nöronun hep ya da hiç prensibiyle çalışan mantıksal bir kapı olarak modellenebileceğini gösteren çalışmaya dayanır [22]. Ancak bu modelin öğrenilebilir parametrelerle matematiksel bir algoritmaya dönüştürülmesi, Rosenblatt (1958) tarafından geliştirilen Perseptron modeli ile gerçekleşmiştir [23].

Bir yapay nöron,her bir x girdi vektörünü her birine özgü bir ω ağırlık vektörü ile çarpar. Bu modele göre bir nöronun net girdisi, tüm giriş sinyallerinin apırlıklı toplamına eşittir. Ayrıca modele daha sonra aktivasyon eşiğini kontrol etmek için bir b sapma (bias) değeri eklenir [23]. Eşitlik (3.1) de n adet girdiye sahip bir nöronun gerçekleştirdiği bu doğrusal işlem ifade edilmektedir:

$$z = \sum_{i=1}^{n} \omega_0 x_i + b \tag{3.1}$$

Günümüzdeki derin öğrenme uygulamalarında, çok sayıda nöronun aynı anda işlenmesi gerektiğinden, bu işlem vektör ve matris çarpımı formunda daha kompakt bir şekilde gösterilir. LeCun ve ark. (2015) tarafından tanımlandığı üzere [21], modern gösterim eşitlikl (3.2) deki gibidir:

$$z = W^T X + b \tag{3.2}$$

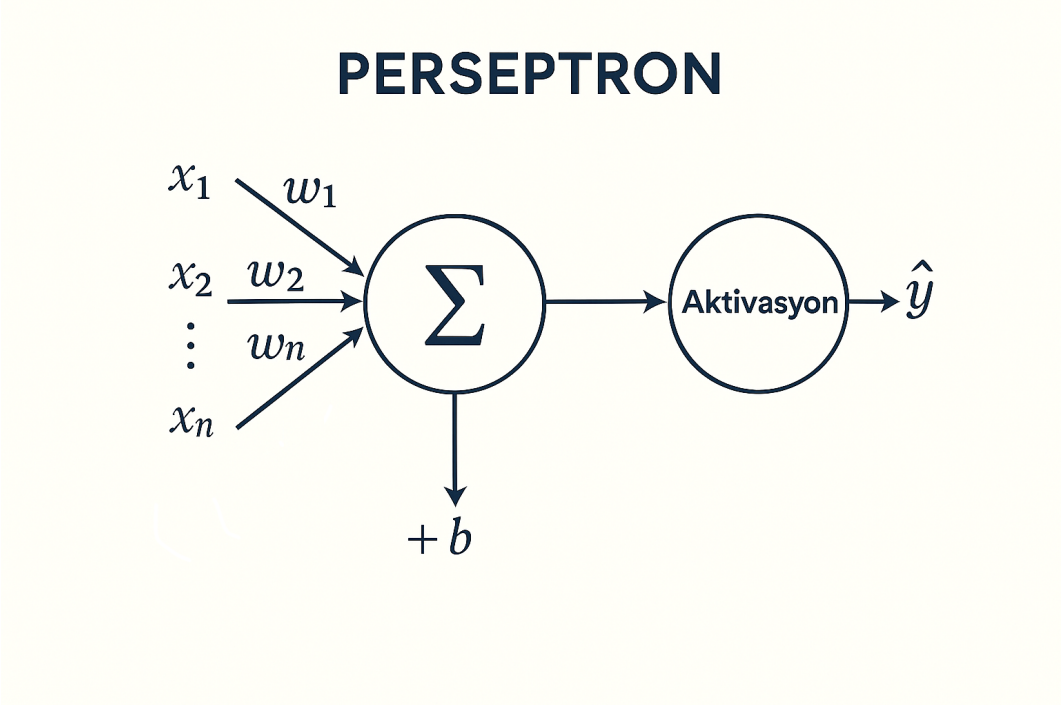
Burada W ağırlık vektörünü, X girdi vektörünü temsil eder. Ancak, sadece doğrusal toplama işlemi yapmak, ağın karmaşık problemleri (örneğin MR görüntüsündeki düzensiz tümör sınırlarını) modellemesi için yeterli değildir. Elde edilen z değeri, ağa lineer olmayan (non-linear) bir özellik kazandırmak amacıyla, türevlenebilir bir aktivasyon fonksiyonundan (f) geçirilir.

Nöronun nihai çıktısı (y), eşitlik (3.3)'teki genel eşitlik ile elde edilir [21]:

$$y = f(z) = f\left(W^T X + b\right) \tag{3.3}$$

Ağın öğrenme süreci, tahmin edilen çıktı ($\hat{y}$) ile gerçek etiket (y) arasındaki hatayı minimize eden W ve b parametrelerinin bulunması problemidir .

Şekil 3.2.1 Eşitlik 3.1 de verilen perseptron modeli şekildeki gibi gösterilebilir.



Şekil 3.1 Eşitlik (3.1) de verilen perseptron modeli şekildeki gibi gösterilebilir.

3.2.1 Sinir Ağlarının Temel Bileşenleri

Bir yapay sinir ağının sadece ağırlıklı toplama işlemi (doğrusal kombinasyon) yapması, onu matematiksel olarak tek katmanlı bir ağdan farksız kılar. Karmaşık

veri setlerini modelleyebilmek için ağın doğrusal olmayan (non-linear) bir yapıya kavuşması gerekir [21]. Bu özellik aktivasyon fonksiyonları ile sağlanır. Ayrıca, modelin öğrenme başarısı, tahmin edilen çıktı ile gerçek veri arasındaki farkı (hatayı) ölçen kayıp fonksiyonları (loss functions) ile değerlendirilir.

3.2.1.1 Aktivasyon Fonksiyonları

Aktivasyon fonksiyonları, nöronun ürettiği net girdiyi (z) belirli bir aralığa sıkıştırarak veya dönüştürerek ağın karar verme mekanizmasını belirler. Literatürde en sık kullanılan fonksiyonlar şunlardır:

1. Sigmoid Fonksiyonu: Girdi değerini (0,1) aralığına sıkıştırarak bir olasılık değeri üretir. Matematiksel modeli eşitlik (3.4)'te gösterildiği gibidir [21]:

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{3.4}$$

Ancak, çok derin ağlarda türev değerlerinin uç noktalarda sıfıra yaklaşması nedeniyle kaybolan gradyan (vanishing gradient) problemine yol açtığı için günümüzde genellikle sadece ikili sınıflandırma (binary classification) yapan çıkış katmanlarında tercih edilir.

2. ReLU (Rectified Linear Unit): Modern derin öğrenme mimarilerinin ve özellikle evrişimli sinir ağlarının (CNN) varsayılan aktivasyon fonksiyonudur. Nair ve Hinton (2010) tarafından önerilen bu fonksiyon, negatif girdileri sıfırlarken pozitif girdileri olduğu gibi geçirir [32]. Eşitlik (3.5) te görüldüğü gibidir:

$$f(x) = \max(0,x) \tag{3.5}$$

3. Softmax Fonksiyonu: Genellikle çok sınıflı sınıflandırma problemlerinde ağın son katmanında kullanılır. Çıktı vektöründeki değerleri, toplamı 1 olan bir olasılık dağılımına dönüştürür. Bir x vektörü için i. sınıfın olasılığı eşitlik (3.6)'daki gibi hesaplanır:

$$softmax(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{K} e^{x_j}} \tag{3.6}$$

Burada K, toplam sınıf sayısını temsil eder.

3.2.1.2 Kayıp Fonksiyonları (Loss Functions)

Modelin eğitimi sırasında, ağırlıkların güncellenmesi için modelin ne kadar hata yaptığının hesaplanması gerekir. Bu tez çalışmasında kullanılan temel kayıp fonksiyonları şunlardır:

1. Çapraz Entropi (Cross-Entropy Loss): Sınıflandırma problemlerinde standart olarak kullanılır. Modelin tahmin ettiği olasılık dağılımı (p) ile gerçek sınıf etiketi (y) arasındaki farkı (bilgi kaybını) ölçer . Eşitlik (3.7)'de ifade edilmiştir:

$$L_{CE} = -\sum_{c=1}^{M} y_c \log(p_c) \tag{3.7}$$

2. Dice Kaybı (Dice Loss): Tıbbi görüntü segmentasyonunda, özellikle tümörün görüntüde çok küçük bir alan kapladığı (dengesiz veri) durumlarda çapraz entropi yetersiz kalabilir. Dice kaybı, piksel bazlı doğruluktan ziyade, tahmin edilen alan ile gerçek alanın örtüşme oranına (overlap) odaklanır . Eşitlik (3.8)'de gösterilmiştir:

$$L_{Dice} = 1 - \frac{2\sum_i y_i p_i}{\sum_i y_i + \sum_i p_i + \varepsilon} \tag{3.8}$$

Burada ε, paydanın sıfır olmasını engelleyen küçük bir sabittir.

3.2.2 Öğrenme Süreci: Geri Yayılım ve Optimizasyon Algoritmaları

Bir derin öğrenme modelinin eğitimi, esasen bir optimizasyon problemidir. Temel amaç, ağın tahmin ettiği çıktı ile gerçek etiket arasındaki farkı ölçen kayıp fonksiyonunu (L) minimize edecek en uygun ağırlık (W) ve sapma (b) parametrelerini bulmaktır. Bu süreç, analitik olarak tek adımda çözülemediği için, türev tabanlı ve yinelemeli (iteratif) algoritmalar kullanılarak gerçekleştirilir. Öğrenme süreci iki temel aşamadan oluşur: Hatanın kaynağının bulunması (geri yayılım) ve parametrelerin bu hataya göre güncellenmesi (optimizasyon).

3.2.2.1 Geri Yayılım (Backpropagation) Algoritması

Çok katmanlı bir ağda, çıkış katmanında hesaplanan hatanın, giriş katmanına yakın bir nöronun ağırlığından ne kadar kaynaklandığını belirlemek karmaşık bir işlemdir. Bu sorun, Rumelhart, Hinton ve Williams (1986) tarafından popülerleştirilen geri yayılım algoritması ile çözülmüştür [25].

Geri yayılım, hatanın ağ boyunca geriye doğru (çıkıştan girişe) iletilmesini sağlar ve matematiksel olarak kalkülüsteki zincir kuralı (chain rule) prensibine dayanır.

Bir ağırlık parametresine (w) göre kayıp fonksiyonunun (L) türevi (gradyanı) eşitlik (3.9) da gösterildiği gibi hesaplanır:

$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial z} \cdot \frac{\partial z}{\partial w} \tag{3.9}$$

Bu gradyan değeri ($\partial\, L/\partial\, w$), ilgili ağırlığın hatayı artırma veya azaltma yönünü ve büyüklüğünü ifade eder.

3.2.2.2 Stokastik Gradyan İniş (SGD)

Hesaplanan gradyanlar kullanılarak ağırlıkların güncellenmesi işlemine gradyan iniş (gradient descent) denir. Ağırlıklar, hatayı azaltmak amacıyla gradyanın tersi yönünde güncellenir. Ancak tıbbi görüntüleme verileriyle çalışırken tüm veri setini tek seferde işlemek bellek açısından mümkün olmadığından, güncelleme işlemi küçük veri paketleri (mini-batch) üzerinden yapılır. Buna Stokastik Gradyan İniş (SGD) denir. SGD güncelleme kuralı eşitlik (3.10)'da gösterildiği gibidir [33]:

$$w_{t+1} = w_t - \eta \nabla L(w_t) \tag{3.10}$$

Burada η, öğrenme oranı (Learning Rate) olup adım büyüklüğünü belirler. SGD'nin yerel minimumlardan kurtulmasını sağlamak ve optimizasyona ivme kazandırmak için genellikle momentum terimi eklenir:

$$v_{t+1} = \mu\ v_t - \eta \nabla L(w_t) \tag{3.11}$$

$$w_{t+1} = w_t + v_{t+1} \tag{3.12}$$

3.2.2.3 Adaptif Moment Tahmini (Adam)

CNN ve Transformer gibi milyonlarca parametreye sahip modellerde, her parametre için sabit bir öğrenme oranı kullanmak verimsizdir. Kingma ve Ba (2014) tarafından önerilen Adam, her parametre için öğrenme oranını dinamik olarak ayarlayan adaptif bir yöntemdir. Adam, gradyanların hem birinci momentini (ortalama -m_t) (eşitlik (3.13)) hem de ikinci momentini (varyans -v_t) (eşitlik (3.14)) kullanarak çalışır [34]:

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)\, g_t \tag{3.13}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2)\, g_t^2 \tag{3.14}$$

Momentler hesaplandıktan sonra sapma düzeltmesi (bias correction) yapılır ve parametreler güncellenir:

$$w_t = w_{t-1} - \frac{\eta}{\sqrt{\hat{v}_t} + \varepsilon} \hat{m}_t \tag{3.15}$$

3.2.2.4 AdamW (Decoupled Weight Decay)

Standart Adam algoritması, L2 regülarizasyonunu kayıp fonksiyonuna ekleyerek uygular. Ancak Loshchilov ve Hutter (2019), adaptif yöntemlerde bu yaklaşımın ağırlık çürütme (weight decay) etkisini bozduğunu göstermiştir [35].

Bu sorunu çözmek için geliştirilen AdamW, ağırlık çürütme işlemini gradyan hesaplamasından ayırarak (decouple) doğrudan güncelleme adımına ekler. Bu yöntem, özellikle bu tezde kullanılan Vision Transformer (ViT) ve Swin Transformer modellerinin genelleme yeteneğini artırmada kritik öneme sahiptir. AdamW güncelleme kuralı eşitlik (3.16)'da gösterillir:

$$w_t = w_{t-1} - \eta \left(\frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \varepsilon} + \lambda w_{t-1} \right) \tag{3.16}$$

Burada λ, ağırlık çürütme katsayısıdır. Bu formülasyon, modelin aşırı öğrenmesini (overfitting) engelleyerek daha kararlı bir eğitim süreci sağlar.

3.2.3 Katmanlı Mimari ve Derin İleri Beslemeli Ağlar

Tek bir nöron (Perseptron), sadece doğrusal olarak ayrılabilen basit problemleri çözebilir. Ancak bu yapının karmaşık prblemleri çözmedeki yetersizliğinden çok katmanlı algılayıcılar (multi layer perceptron - MLP) çıkmıştır. Bu çok katmanlı yapılar derin öğrenmenin derin kısmını oluşturur. Çok katmanlı mimari sinyal akışının girişten çıkışa doğru tek yönlü olduğu üç temel katmandan oluşur:

1. Giriş Katmanı(Input Layer): Dış dünyadan alınan ham verinin ağa beslendiği katmandır. Bu katmandaki nöronlar herhangi bir hesaplama yapmaz, sadece veriyi

bir sonraki katmana iletir.

2. Gizli Katmanlar (Hidden Layers): Giriş ve çıkış katmanları arasında yer alan bu katmanlar, ağın dışarıdan doğrudan gözlemlenemeyen içsel temsillerini oluşturur. gizli olarak adlandırılmalarının nedeni budur. Ağın derinliği arttıkça, bu katmanlar verideki hiyerarşik öznitelikleri (kenar, doku, şekil gibi) öğrenir.

3. Çıkış Katmanı (Output Layer): Ağın işlenen veriye dair nihai kararını verdiği (sınıflandırma veya regresyon) son katmandır.

Katmanlı bir mimaride, bir nöronun çıktısı, kendisinden önceki katmandaki nöronların çıktılarının ağırlıklı toplamının bir aktivasyon fonksiyonundan geçirilmesiyle elde edilir. l. katmandaki bir nöronun işlemi matematiksel olarak eşitlik (3.17)'de ki gibi genelleştirilir [21]:

$$h^{(l)} = \varphi\left(W^{(l)}h^{(l-1)} + b^{(l)}\right) \tag{3.17}$$

Burada: $W^{(l)}$: Ağırlık matrisini,

$b^{(l)}$: Sapma (bias) vektörünü,

φ: Doğrusal olmayan aktivasyon fonksiyonunu (örneğin ReLU),

$h^{(l-1)}$: Bir önceki katmanın çıktısını temsil eder.

Bu yapı sayesinde sinir ağı, giriş uzayını (input space) gizli katmanlar aracılığıyla bükerek, veriyi çıkış katmanında doğrusal olarak ayrılabilir bir hale getirir.

3.3 Evrişimli Sinir Ağları (CNN) Matematiği

Geleneksel Çok Katmanlı Algılayıcı (MLP) mimarileri, görüntü verilerinin işlenmesinde yapısal bazı kısıtlılıklar barındırmaktadır. Standart bir Yapay Sinir Ağı modeli, giriş verisini işleyebilmek için 2 boyutlu (2D) görüntü matrisini tek boyutlu (1D) bir vektöre dönüştürerek düzleştirmek (flattening) zorundadır. Ancak bu işlem, pikseller arasındaki komşuluk ilişkilerinde saklı olan mekansal bilginin (spatial correlation) ve topolojik yapının kaybolmasına neden olmaktadır. Ayrıca görüntü boyutu büyüdükçe parametre sayısı yönetilemez boyutlara ulaşmaktadır [21].

Bu duruma çözüm, biyolojik görme sistemindeki yerel alıcı alan prensibinden ilham alan Evrişimli Sinir Ağları (CNN) ile sağlanmıştır. CNN mimarisi, görüntüyü

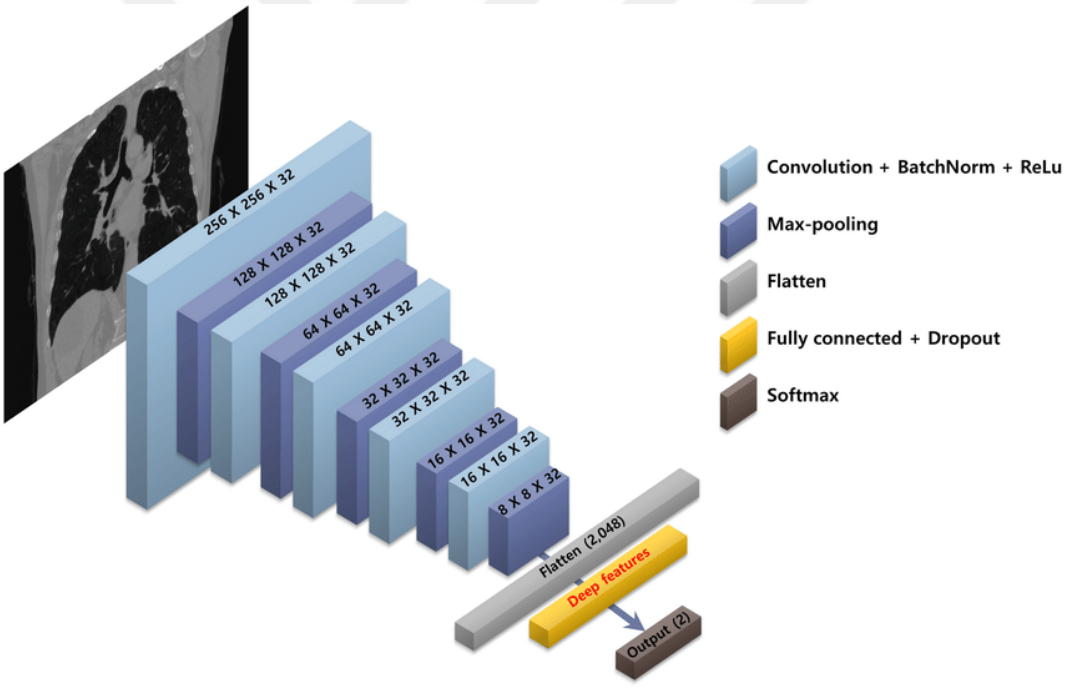
vektörize etmeden, ham matris formunda işleyerek mekansal özellikleri korur. Bölüm 3.1'de tarihsel gelişim süreci aktarılan CNN mimarisinin, bu bölümde çalışma prensibi ve matematiksel altyapısı incelenmiştir.

CNN yapıları temelde, her biri farklı bir amaca hizmet eden üç ana işlem bloğundan oluşur:

Evrişim (Convolution): Görüntüden öznitelik (kenar, doku vb.) çıkarma işlemi gerçekleştirilir.

Havuzlama (Pooling): Boyut azaltma ve hesaplama yükünü hafifletme işlemi yapılır.

Tam Bağlı Katmanlar (Fully Connected): Çıkarılan öznitelikler kullanılarak sınıflandırma yapılması hedeflenir. Şekil 3.2 de CNN mimarisinin katmanlarının temsili bir anlatımı görülmektedir.



Şekil 3.2 Şekilde CNN mimarisinin temsili bir modeli görülmektedir. Burada görüntü matrisinin tek boyuta inişi gözlemlenebilir [36].

Aşağıda bu katmanların matematiksel modelleri detaylandırılmıştır.

3.3.1 CNN Mimarisinin Katmanları

3.3.1.1 Evrişim İşlemi ve Filtreleme

Evrişim, görüntü üzerindeki anlamlı öznitelikleri tespit etmek için kullanılan doğrusal bir işlemdir. Bu süreçte, filtre veya çekirdek (kernel) adı verilen,

öğrenilebilir ağırlıklara sahip küçük boyutlu (örneğin 3×3) matrisler, giriş görüntüsü üzerinde kaydırılarak işlem yapılır [27]. Matematiksel olarak, 2 boyutlu bir I görüntüsü (giriş) ve K çekirdeği (filtre) arasındaki evrişim işlemi eşitlik (3.18)'de ifade edilmiştir:

$$S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(m,n) \cdot K(i-m, j-n) \tag{3.18}$$

Burada $S(i,j)$, işlem sonucunda oluşan öznitelik haritası (Feature Map) üzerindeki değerdir. Ağ, eğitim sürecinde geri yayılım algoritmasını kullanarak bu filtrelerin içindeki ağırlık değerlerini kendisi öğrenir; böylece örüntü tespiti için gerekli en uygun filtreler (kenar, doku vb.) otomatik olarak oluşturulur.

3.3.1.2 Havuzlama Katmanları

Evrişim katmanlarından elde edilen öznitelik haritaları genellikle yüksek boyutludur. Hesaplama maliyetini düşürmek ve modelin küçük ötelemelere karşı dayanıklılığını (translation invariance) artırmak için havuzlama işlemi uygulanır.

Literatürde en yaygın kullanılan yöntem maksimum havuzlama (Max Pooling) işlemidir [28]. Bu işlem, belirlenen bir pencere (örneğin 2×2) içerisindeki en büyük değeri seçerek çalışır ve Eşitlik (3.19)'deki gibi formüle edilir:

$$y_{i,j} = \max_{(p,q) \in R_{i,j}} \left(x_{p,q}\right) \tag{3.19}$$

Burada $R_{i,j}$, havuzlama penceresini temsil eder. Bu işlem sayesinde görüntüdeki gürültü filtrelenirken en baskın özellikler korunmuş olur.

3.3.1.3 Tam Bağlı Katmanlar (Fully Connected)

Evrişim ve havuzlama katmanları, görüntüden hiyerarşik öznitelikleri çıkarır ancak bu çıktı hala matris formundadır. Sınıflandırma işleminin yapılabilmesi için bu matrisler düzleştirme (Flattening) işlemi ile tek boyutlu bir vektöre dönüştürülür ve geleneksel bir YSA yapısına (tam bağlı katman) iletilir. Son katmanda, Bölüm 3.2.1.1'de değinilen Softmax veya Sigmoid fonksiyonları kullanılarak nihai sınıflandırma olasılıkları hesaplanır [21].

3.3.2 CNN Tabanlı Sınıflandırma Mimarileri

Konvolüsyonel Sinir Ağları (CNN), görüntüdeki yerel özellikleri (kenar, doku) öğrenme konusundaki başarıları sayesinde uzun süre alanın hakimi olmuştur. Ancak ağlar derinleştikçe ortaya çıkan eğitim zorluklarını aşmak için CNN tabanlı farklı mimari yaklaşımlar geliştirilmiştir.

3.3.2.1 ResNet (Residual Networks)

Derin öğrenme alanında ağlar derinleştikçe (katman sayısı arttıkça), modelin başarısının artması beklenirken, pratikte kaybolan gradyan (vanishing gradient) ve bozulma (degradation) problemleri nedeniyle eğitim zorlaşmakta ve doğruluk oranları düşmektedir. Bu sorunu çözmek amacıyla He ve arkadaşları (2016), ImageNet yarışmasında devrim yaratan ResNet (Residual Network) mimarisini önermiştir [3].

ResNet mimarisinin temel yeniliği, ağın öğrenmeye çalıştığı haritalama fonksiyonunu değiştirmesidir. Geleneksel düz (plain) ağlar, giriş verisi x'i doğrudan bir hedef çıktıya $H(x)$ eşlemeye çalışırken; ResNet, katmanların artık fonksiyonu (Residual Function) öğrenmesini hedefler.

Bu yapıda, bir önceki katmanın çıktısı (x), bir sonraki katmana sadece ağırlıklar üzerinden değil, aynı zamanda bir atlamalı bağlantı (Skip Connection) veya diğer adıyla kısa devre (Shortcut) üzerinden doğrudan aktarılır.

Matematiksel olarak, bir artık bloğun (Residual Block) çıktısı, şu eşitlik (3.20) ile ifade edilir:

$$y = \sigma\left(F\left(x, \{W_i\}\right) + x\right) \tag{3.20}$$

Bu denklemde:

x: Bloğa giren veri (giriş tensörü).

$F\left(x, \{W_i\}\right)$: Öğrenilecek artık fonksiyon (genellikle iki veya üç evrişim katmanı).

$\{W_i\}$: Katmanların ağırlıkları.

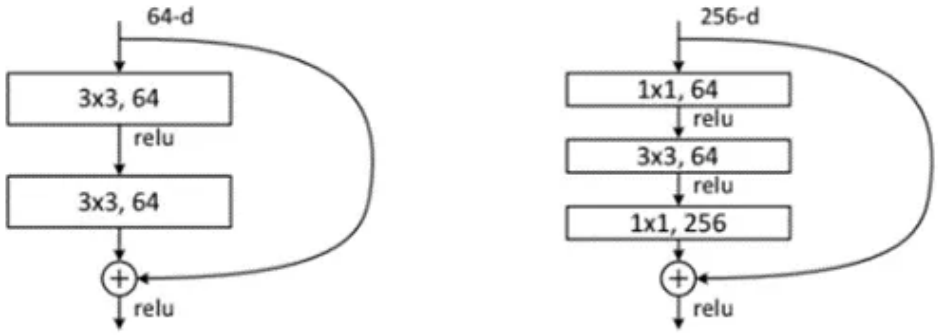
σ: ReLU (Rectified Linear Unit) aktivasyon fonksiyonu.

Eğer x ve $F(x)$boyutları aynı değilse (boyut indirgeme yapılıyorsa), boyutları eşitlemek için $x$üzerinde lineer bir projeksiyon (W_s) uygulanır:

$$y = \sigma\left(F\left(x, \{W_i\}\right) + W_s x\right) \tag{3.21}$$

Bu yapı sayesinde, geriye yayılım (backpropagation) sırasında gradyanlar, ağırlık katmanlarında kaybolmadan x yolu üzerinden doğrudan ağın ilk katmanlarına kadar akabilir. Bu da teorik olarak sonsuz derinlikte ağların eğitilmesine olanak tanır.

ResNet50, ResNet mimarisinin katman sayısına göre $(18, 34, 50, 101, 152)$ farklı varyasyonları bulunmaktadır. Bu tez çalışmasında, doğruluk ve hesaplama maliyetini dengeleyen ResNet50 modeli tercih edilmiştir.



Şekil 3.3 Sol: ResNet-34 için yapı taşı Sağ: ResNet50/101/152 için dar boğaz yapısı [3].

ResNet50, daha sığ olan ResNet18/34 modellerinden yapısal olarak ayrılır. Standart bloklar yerine, hesaplama verimliliğini artıran bottleneck (darboğaz) blokları kullanır. Bu yapılar arasındaki fark şekil 3.3 de yapsal olarak görülmektedir. Bir darboğaz bloğu, sırasıyla şu üç katmandan oluşur:

1×1 Konvolüsyon: Boyut (kanal sayısı) azaltma işlemi yaparak hesaplama yükünü düşürür.

3×3 Konvolüsyon: Asıl öznitelik çıkarma işlemini yapar.

1×1 Konvolüsyon: Boyutu tekrar artırarak (restore) bir sonraki bloğa hazırlar.

Bu daralt-işle-genişlet stratejisi, modelin 50 katmanlı olmasına rağmen donanım kaynaklarını verimli kullanmasını sağlar. Çalışmada kullanılan ResNet50 modeli, ImageNet ve RadImageNet ağırlıklarıyla başlatılarak son katmanı veri setlerindeki sınıf sayısına (2 veya 4) göre modifiye edilmiştir.

3.3.2.2 DenseNet (Densely Connected Networks)

ResNet mimarisinin başarısından ilham alan ancak bilgi akışını daha da iyileştirmeyi hedefleyen Huang ve arkadaşları (2017), DenseNet (Yoğun Bağlantılı Ağlar) mimarisini geliştirmişlerdir [37].

ResNet'lerdeki toplama (addition) işlemi, bilginin ağ boyunca akmasını sağlasa da, bazı özniteliklerin kaybolmasına veya değişmesine neden olabilir. DenseNet'in temel felsefesi, ağdaki her katmanın, kendisinden önceki tüm katmanlardan doğrudan bilgi almasıdır. Bu yapı, öznitelik tekrar kullanımı (Feature Reuse) sağlar ve ağın daha az parametre ile daha zengin desenleri öğrenmesine olanak tanır.

CNN'de l. katmanın çıkışı x_l, bir önceki katmanın çıktısı x_{l-1}'e bağlıdır: $x_l = H_l(x_{l-1})$.

ResNet bu ilişkiyi $x_l = H_l(x_{l-1}) + x_{l-1}$ şeklinde değiştirir.

DenseNet ise toplama yerine kanal birleştirme (Concatenation) işlemini kullanır. l. katman, giriş olarak $x_0, x_1, \ldots, x_{l-1}$ özellik haritalarının birleşimini alır. Bu durum eşitlik (3.42) ile gösterilir.

$$x_l = H_l([x_0, x_1, \ldots, x_{l-1}]) \tag{3.22}$$

Bu formülde:

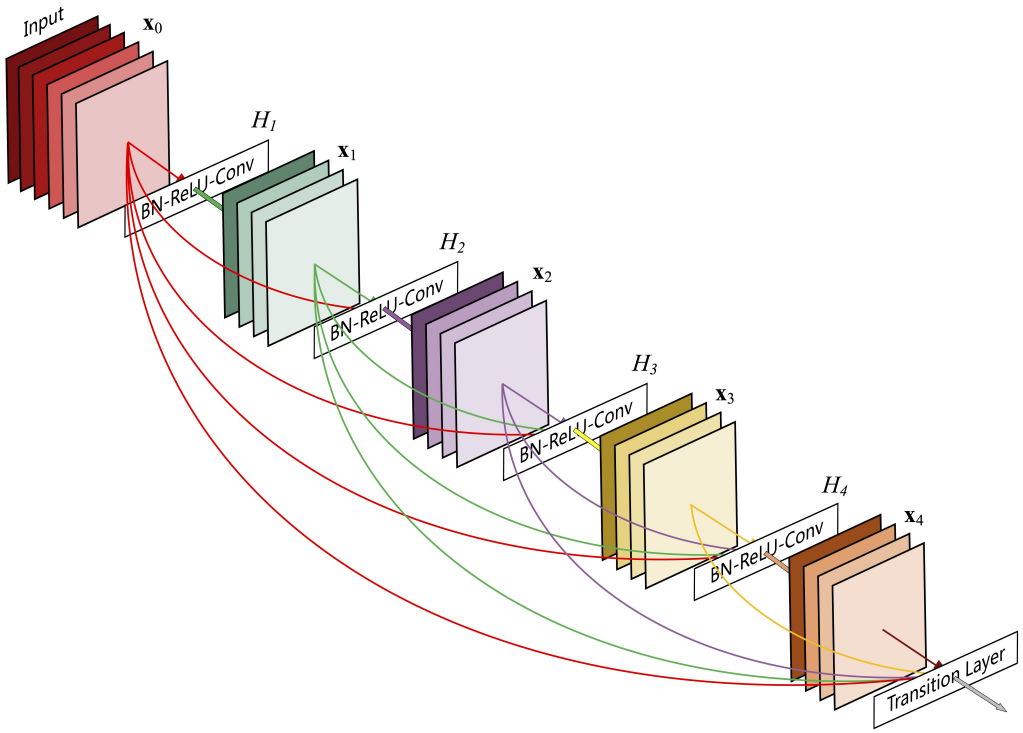
$[x_0, x_1, \ldots, x_{l-1}]$: Önceki katmanlardan gelen özellik haritalarının kanal ekseninde birleştirilmesini (concatenation) ifade eder.

$H_l(\bullet)$: Genellikle batch normalization (BN), ReLU ve Konvolüsyon işlemlerinden oluşan bileşik bir fonksiyondur.

Bu yoğun bağlantı yapısı sayesinde, ağın son katmanlarındaki hata sinyali (gradient), aradaki katmanlara uğramadan doğrudan ilk katmanlara kadar ulaşabilir (Deep Supervision etkisi), bu da eğitimin kararlılığını artırır.

DenseNet mimarisinin derinliğine göre (121, 169, 201, 264) çeşitli versiyonları bulunmaktadır. Bu tez çalışmasında DenseNet121 modeli tercih edilmiştir. DenseNet121 şu yapıtaşlarından oluşur:

Dense Bloklar: Katmanların yoğun şekilde bağlandığı ana bloklar. DenseNet121'de sırasıyla (6, 12, 24, 16) katmanlı 4 adet Dense Blok bulunur.



Şekil 3.4 DenseNet mimarisinin yapısı şekildeki gibi özetlenebilir [37].

Geçiş Katmanları (Transition Layers): Bloklar arasında özellik haritasının boyutunu (conv ile) ve çözünürlüğünü pooling ile) azaltan katmanlardır.

3.3.2.3 Inception (GoogLeNet / InceptionV3)

Konvolüsyonel Sinir Ağları tasarlanırken karşılaşılan en kritik zorluklardan biri, görüntü içerisindeki nesnelerin boyutlarının (scale) değişkenlik göstermesidir. Örneğin, bir beyin MR görüntüsünde tümör bazen milimetrik boyutta (küçük) olabilirken, bazen beynin büyük bir kısmını kaplayan (büyük) bir kitle halinde olabilir.

Geleneksel CNN mimarileri, her katmanda sabit bir filtre boyutu (çekirdek) seçmek zorundadır:

Küçük Filtreler (3×3): İnce detayları ve küçük nesneleri yakalar ama büyük resmi (bağlamı) kaçırır.

Büyük Filtreler (5×5 veya 7×7): Geniş alanı ve bağlamı görür ama ince detayları kaybeder ve hesaplama maliyeti çok yüksektir.

Szegedy ve arkadaşları (2016), bu ikilemden kurtulmak için şu soruyu sormuştur: "Neden seçmek zorunda kalalım? Hepsini aynı anda kullanamaz mıyız?" [38]. Bu

düşünceyle, ağın derinliğinden ziyade genişliğine odaklanan Inception mimarisini geliştirmişlerdir.

Inception mimarisinin önerdiği çözüm, paralel işleme mantığıdır. Bir Inception modülü, gelen veri üzerinde 1×1, 3×3 ve 5×5 konvolüsyon işlemleri ile 3×3 havuzlama (pooling) işlemini aynı andagerçekleştirir. Elde edilen tüm öznitelik haritaları (feature maps) daha sonra birleştirilir (concatenation). Böylece ağ, tümörün boyutu ne olursa olsun, onu yakalayabilecek uygun filtreyi kendi içinde barındırmış olur.

Ancak, büyük filtrelerin paralel kullanımı işlem yükünü aşırı artırır. InceptionV3, bunu çözmek için matematiksel bir optimizasyon olan filtre ayrıştırma (factorization) yöntemini kullanır. Büyük boyutlu bir $n \times n$ filtresi, iki adet tek boyutlu (asimetrik) filtreye ayrılır: $n \times n \rightarrow (1 \times n + n \times 1)$.

Örneğin, 5×5 boyutunda bir filtre yerine, önce 1×5, ardından 5×1 boyutunda iki filtre ardışık olarak uygulanır. Bu işlemin sağladığı hesaplama tasarrufu (gain) eşitlik (3.23)'de ifade edilir:

$$\frac{1 \times n + n \times 1}{n \times n} = \frac{2n}{n^2} = \frac{2}{n} \tag{3.23}$$

$n = 5$ için bu yöntem, parametre sayısını %60 oranında azaltırken ($\frac{2}{5} = 0.4$), ağın öğrenme kapasitesini (non-linearity) artırır.

Inception mimarisinin gelişimi sürecinde (V1, V2, V3, V4), bu tez çalışması için en dengeli versiyon olan InceptionV3 tercih edilmiştir.

3.3.2.4 MobileNet

Derin öğrenme modelleri geliştikçe artan parametre sayıları ve işlem yükleri (FLOPS), bu modellerin mobil cihazlar ve gömülü sistemlerde çalıştırılmasını zorlaştırmıştır. Bu sorunu çözmek amacıyla Google araştırmacıları Howard ve arkadaşları (2017), modelin doğruluğundan minimum ödün vererek işlem hızını artıran MobileNet mimarisini önermişlerdir [39].

Mimarinin zamanla gelişen versiyonları (V2, V3) arasından, bu tez çalışmasında en güncel optimizasyonları içeren MobileNetV3 (2019) varyasyonu tercih edilmiştir [40].

MobileNet mimarisinin temel yapı taşı, 2017'deki ilk makalede detaylandırılan derinlik ayrışımlı evrişim (Depthwise Separable Convolution) tekniğidir. Standart

konvolüsyon işlemi iki aşamaya ayrılarak işlem yükü hafifletilir:

Derinlik Evrişimi (Depthwise Convolution): Her giriş kanalı için ayrı bir uzaysal filtre (3×3) uygulanır. Kanallar arası etkileşim yoktur.

Noktasal Evrişim (Pointwise Convolution): 1×1 boyutundaki filtrelerle, ayrık kanalları birleştirir ve yeni özellikler oluşturur.

Bu yöntemin standart konvolüsyona göre sağladığı kazanç şu formülle ifade edilir:

$$Hesaplama\ Kazancı = \frac{1}{N} + \frac{1}{D_K^2} \tag{3.24}$$

Burada N çıkış kanalı sayısı, D_K filtre boyutudur. 3×3 filtreler için işlem yükü yaklaşık 9 kat azalır.

Çalışmada kullanılan MobileNetV3 versiyonu, 2019 yılında yayınlanan makalede önerilen şu yeniliklerle güçlendirilmiştir:

Hard-Swish Aktivasyonu: Standart Swish fonksiyonunun maliyetli e^{-x} hesaplaması yerine, parçalı doğrusal (piecewise linear) bir yaklaşım kullanılır. Eşitlik (3.25) de gösterilir:

$$h-swish\,(x) = x\frac{ReLU6\ (x+3)}{6} \tag{3.25}$$

Squeeze-and-Excitation (SE) Blokları: Ağın önemli kanallara odaklanmasını sağlayan dikkat mekanizmaları modele entegre edilmiştir.

NetAdapt: Katman sayısı ve filtre boyutları, Neural Architecture Search (NAS) algoritmalarıyla otomatik olarak optimize edilmiştir.

3.3.2.5 EfficientNet

Konvolüsyonel Sinir Ağları'nın başarısını artırmak için literatürde genellikle üç farklı ölçekleme stratejisi izlenmiştir:

Derinlik (Depth) Artırımı: Daha fazla katman eklemek (Örn: ResNet-152).

Genişlik (Width) Artırımı: Katmanlardaki kanal sayısını artırmak (Örn: WideResNet).

Çözünürlük (Resolution) Artırımı: Giriş görüntüsünün boyutunu büyütmek (Örn: 224×224 yerine 480×480).

Ancak Tan ve Le (2019), bu boyutların tek başına ve rastgele artırılmasının verimsiz olduğunu kanıtlamıştır. Örneğin, sadece çözünürlüğü artırmak, ağın derinliği (receptive field) artırılmazsa anlamsız kalmaktadır. Bu sorunu çözmek için, üç boyutu matematiksel bir dengeyle optimize eden EfficientNet mimarisini önermişlerdir [41].

EfficientNet'in temel yeniliği, bileşik ölçekleme (compound scaling) yöntemidir. Bu yöntemde, derinlik, genişlik ve çözünürlük parametreleri rastgele değil, belirli bir ϕ(phi) katsayısı ile orantılı olarak artırılır.

Ölçekleme işlemi eşitlik (3.26)'daki matematiksel kurala göre yapılır:

$$d = \alpha^{\phi}, \quad w = \beta^{\phi}, \quad r = \gamma^{\phi} \tag{3.26}$$

Bu denklemde d derinlik, w genişlik,r çözünürlük faktörüdür. α, β, γ sabitleri ise küçük bir model üzerinde yapılan şebeke araması (grid search) ile bulunur ve şu kısıtlayıcı denklemi sağlamak zorundadır:

$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2, \quad \alpha \geq 1, \beta \geq 1, \gamma \geq 1 \tag{3.27}$$

Formülde β (genişlik) ve , γ (çözünürlük) terimlerinin karesinin alınmasının sebebi, bu boyutların hesaplama maliyetini (FLOPS) karesel olarak artırmasıdır. Derinlik (α) ise maliyeti lineer artırır. Bu kısıt denklemi sayesinde, ϕkatsayısı her 1 birim arttığında, toplam hesaplama maliyeti yaklaşık $2^1 = 2$ kat artar; böylece modelin büyümesi kontrol altında tutulur.

3.4 Transformer Mimarisi

Bölüm 3.3'te incelenen Evrişimli Sinir Ağları (CNN), görüntü işlemede yüksek başarı sağlasa da, mimari yapısının temelini oluşturan konvolüsyon işleminin doğası gereği, verideki uzun mesafeli bağımlılıkları (long-range dependencies) modellemede teorik sınırlılıklara sahiptir. Konvolüsyon işlemi, yerel alıcı alan (local receptive field) prensibiyle çalışır; yani bir nöron sadece komşu piksellerle etkileşime girer. Görüntünün tamamına yayılan global bağlamın (global context) öğrenilebilmesi için ağın derinliğinin artırılması ve katmanlar boyunca alıcı

alanların genişletilmesi gerekir. Bu durum, bilgi akışında verimlilik kaybına yol açar ve uzaktaki pikseller arasındaki ilişkinin kurulmasını matematiksel olarak zorlaştırır.

Bu yapısal kısıtlılığı aşmak amacıyla, Vaswani ve ark. (2017) tarafından önerilen Transformer mimarisi ve bu yapının Dosovitskiy ve ark. (2020) tarafından görüntü alanına uyarlanan versiyonu olan Vision Transformer (ViT) geliştirilmiştir [30], [31]. ViT mimarisi, pikseller arası mesafeden bağımsız olarak, görüntüdeki tüm veri parçaları arasındaki ilişkisel bağıntıyı tek bir işlem adımında hesaplayabilen öz-dikkat (Self-Attention) mekanizmasına dayanır.

3.4.1 Mimarinin Yapısı

3.4.1.1 Görüntü Yamalama (Image Patching) ve Gömme

Transformer mimarisi, CNN'lerin aksine 2 boyutlu matrisler yerine 1 boyutlu dizileri işlemek üzere tasarlanmıştır. Ancak yüksek çözünürlüklü bir görüntünün doğrudan düzleştirilmesi, dikkat mekanizmasının hesaplama maliyetini karesel ($O\left(N^2\right)$) olarak artıracağından uygulanabilir değildir.

Vision Transformer bu sorunu, giriş görüntüsünü pikseller yerine yamalar (Patches) dizisine dönüştürerek çözer [31]. $H \times W$ boyutundaki giriş görüntüsü, $H \times P$ boyutunda sabit yamalara bölünür. Elde edilen yama dizisinin uzunluğu (N), Eşitlik (3.28) ile ifade edilir:

$$N = \frac{H \times W}{P^2} \tag{3.28}$$

Her bir yama düzleştirilerek doğrusal bir projeksiyon (linear projection) katmanından geçirilir ve D boyutunda gömme (embedding) vektörlerine dönüştürülür. Transformer yapısı, verinin uzaysal sırasına duyarlı olmadığından (permutation invariant), her bir yama vektörüne öğrenilebilir pozisyonel kodlama (Positional Encoding) vektörleri eklenerek modele verinin konumsal bilgisi sağlanır [31].

$$z_0 = \left[x_p^1 E; x_p^2 E; \cdots; x_p^N E\right] + E_{pos} \tag{3.29}$$

3.4.1.2 Öz-Dikkat (Self-Attention) Mekanizması

Transformer mimarisinin temel bileşeni olan öz dikkat mekanizması, girdideki her bir elemanın (yamanın) diğer tüm elemanlarla olan ilişkisini hesaplar. CNN'deki

yerel komşuluk kısıtlamasının aksine, Öz dikkat mekanizmasında herhangi iki yama arasındaki etkileşim mesafesi sabittir ve doğrudan hesaplanır.

Matematiksel olarak bu süreç; sorgu (Query - Q), anahtar (Key - K) ve değer (Value - V) matrisleri üzerinden yürütülür. Vaswani ve ark. (2017) tarafından tanımlanan ölçeklenmiş nokta çarpım dikkati (Scaled Dot-Product Attention), eşitlik 3.27'de gösterilmiştir [30]:

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{3.30}$$

Burada:

QK^T: Tüm yamaların birbirleriyle olan benzerliğini (ilişki katsayısını) içeren matris çarpımıdır.

$\sqrt{d_k}$: Gradyanların kararlılığını sağlamak için kullanılan ölçekleme faktörüdür.

Softmax: İlişki skorlarını toplamı 1 olan bir olasılık dağılımına dönüştürür.

Bu işlem sonucunda model, görüntünün belirli bir bölgesindeki bilginin, diğer herhangi bir bölgesindeki bilgiyle ne kadar ilişkili olduğunu (Attention Score) ağırlıklı bir toplam olarak elde eder.

3.4.1.3 Çok Başlı Dikkat (Multi-Head Self Attention -MSA)

Modelin, verinin farklı öznitelik alt uzaylarını (örneğin şekil bilgisi ve doku bilgisi gibi) paralel olarak öğrenebilmesi için çok başlı dikkat (Multi-Head Attention) yapısı kullanılır [30].

Girdi vektörleri h adet farklı parçaya bölünerek her biri için bağımsız dikkat mekanizmaları çalıştırılır ve sonuçlar birleştirilir (concatenation). Matematiksel ifadesi eşitlik (3.31)'te verilmiştir:

$$MultiHead(Q,K,V) = Concat(head_1,\ldots,head_h)W^O \tag{3.31}$$

Burada her bir dikkat kafası ($head_i$), Eşitlik (3.32)'deki gibi kendine özgü öğrenilebilir projeksiyon matrisleri (W_i^Q, W_i^K, W_i^V) ile hesaplanır:

$$head_i = Attention\left(QW_i^Q, KW_i^K, VW_i^V\right) \tag{3.32}$$

Sonuç olarak elde edilen vektörler birleştirilir (concatenation) ve W^O doğrusal projeksiyon matrisi ile çarpılarak nihai çıktı üretilir. Bu mekanizma, ViT'in global bağlamı çok yönlü olarak analiz etmesini sağlar.

3.4.1.4 Çok Katmanlı Algılayıcı (MLP) Blokları

Transformer kodlayıcı (encoder) katmanlarında, dikkat mekanizmasının (MSA) hemen ardından gelen ve özniteliklerin doğrusal olmayan dönüşümlerini gerçekleştiren yapı, literatürde konuma duyarlı ileri beslemeli ağ (Position-wise Feed-Forward Network) olarak da adlandırılan çok katmanlı algılayıcı (MLP) bloğudur [30].

Bu blok, dikkat mekanizmasından çıkan her bir yama vektörüne (token) bağımsız ve özdeş olarak uygulanır. Yani ağın ağırlıkları, dizideki tüm yamalar için ortaktır, ancak her yama kendi içinde işlenir. Bu yapı, CNN mimarilerindeki 1×1çekirdek boyutuna sahip konvolüsyon işlemine benzetilebilir.

ViT mimarisinde kullanılan MLP bloğu, iki adet tam bağlı (linear) katman ve bu katmanlar arasına yerleştirilmiş doğrusal olmayan bir aktivasyon fonksiyonundan oluşur. Bir giriş vektörü x için MLP bloğunun matematiksel işlemi eşitlik (3.33) ile ifade edilir:

$$MLP(x) = \sigma(xW_1 + b_1)W_2 + b_2 \tag{3.33}$$

Burada:

W_1: Giriş vektörünü daha yüksek bir boyuta (genellikle giriş boyutunun 4 katı, $4d_{model}$) genişleten birinci ağırlık matrisidir.

b_1, b_2: Öğrenilebilir sapma (bias) vektörleridir.

W_2: Genişletilmiş vektörü tekrar orijinal boyutuna (d_{model}) düşüren ikinci ağırlık matrisidir.

σ: Doğrusal olmayan aktivasyon fonksiyonudur.

Bu genişleme ve daralma yapısı (Bottleneck yapısının tersi), modelin verideki karmaşık desenleri daha yüksek boyutlu bir uzayda çözümlemesine olanak tanır.

3.4.1.5 GELU (Gaussian Error Linear Unit) Aktivasyon Fonksiyonu

Geleneksel CNN modellerinde (ResNet vb.) yaygın olarak ReLU fonksiyonu kullanılırken, ViT, BERT ve GPT gibi modern Transformer mimarilerinde Hendrycks ve Gimpel (2016) tarafından önerilen GELU aktivasyon fonksiyonu tercih edilmektedir [42].

GELU, nöronun girdisini (x), standart normal dağılımın kümülatif dağılım fonksiyonu ($\Phi(x)$) ile ağırlıklandırır. ReLU'nun deterministik yapısının aksine, stokastik bir düzenleyici (regularizer) gibi davranır. Matematiksel tanımı eşitlik (3.34)'de verilmiştir:

$$GELU(x) = x\Phi(x) \approx 0.5x\left(1+\tanh\left[\sqrt{\frac{2}{\pi}}\left(x+0.044715x^3\right)\right]\right) \tag{3.34}$$

GELU fonksiyonu, noktasında ReLU gibi keskin bir kırılma yaşamaz; bunun yerine türevlenebilir kavisli bir yapı sergiler. Bu özellik, özellikle çok derin Transformer ağlarının eğitiminde gradyan akışını iyileştirerek modelin daha hızlı ve kararlı yakınsamasını (convergence) sağlar.

3.4.2 Transformer Tabanlı Mimariler

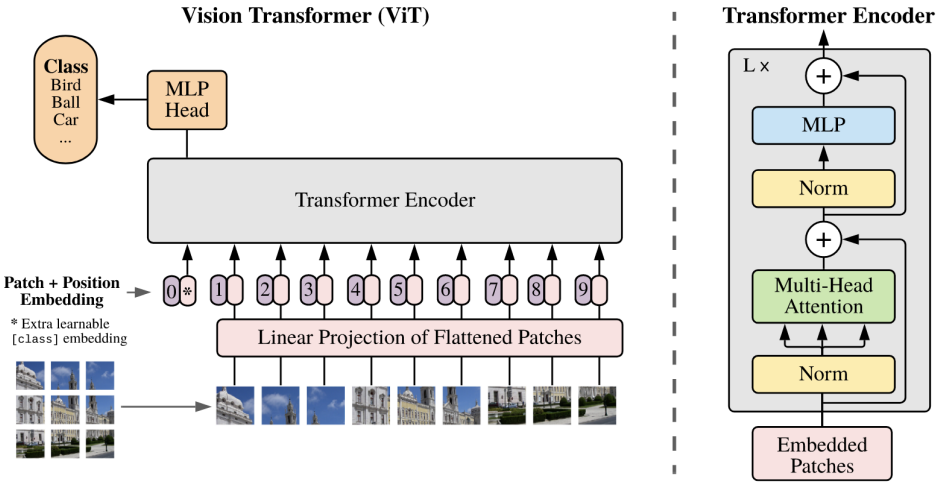
Doğal Dil İşleme (NLP) alanında "Attention is All You Need" (2017) makalesiyle devrim yaratan Transformer mimarisi [30], uzun süre sadece metin verileri üzerinde en iyi sonuçları vermiştir. Ancak Dosovitskiy ve arkadaşları (2021) [31], bu yapıyı minimum değişiklikle görüntü işlemeye uyarlayarak CNN mimarilerine güçlü bir alternatif sunmuşlardır.

3.4.2.1 Vision Transformer (ViT)

Dosovitskiy ve arkadaşları (2021) tarafından önerilen ViT mimarisi, standart Transformer yapısınndaki yapıltaşlarını şekil 3.5.1 de gösterilen modelde bir araya getirerek görüntüyü, bir metin gibi işleyen saf transformer yaklaşımını sunar[27].

ViT mimarisinin işleyişi, şema üzerinden takip edileceği üzere üç temel aşamadan oluşur:

1. Girdi Hazırlığı ve Sınıflandırma Vektörü (Class Token): İlk aşamada görüntü, Bölüm 3.4.1.1'de detaylandırıldığı gibi sabit boyutlu yamalara bölünür ve vektör uzayına izdüşürülür. Ancak CNN'lerin aksine, ViT'in görüntünün bütününe dair bir sınıflandırma kararı verebilmesi için, diziye harici bir parametre eklenir.



Şekil 3.5 ViT mimarisinin temsili gösterimi [31]

Şemada 0* ile gösterilen ve class Token (x_{class}) olarak adlandırılan bu öğrenilebilir vektör, dizinin en başına yerleştirilir. Ayrıca, yamaların görüntüdeki yerini modelin algılayabilmesi için tüm diziye pozisyonel kodlama vektörleri eklenir.

2. Transformer Kodlayıcı (Encoder) Bloğu: Hazırlanan vektör dizisi, kodlayıcı (Encoder) bloğuna girer. Bu blok, daha önce tanımladığımız iki ana mekanizmanın ardışık ve tekrarlı (L kez) kullanımından oluşur:

Çok Başlı Dikkat (MSA): Bölüm 3.4.1.3 te anlatılan bu yapı, x_{class} vektörünün görüntüdeki diğer tüm yamalarla etkileşime girerek global bilgiyi toplamasını sağlar.

Çok Katmanlı Algılayıcı (MLP): Bölüm 3.4.1.4'te açıklanan bu yapı, dikkat katmanından gelen özellikleri işleyerek doğrusal olmayan dönüşümleri gerçekleştirir.

Şemada görüldüğü üzere, her iki alt bloğun öncesinde katman Normalizasyonu (Layer Norm) uygulanır ve sonrasında giriş ile çıkışı toplayan artık bağlantılar (Residual Connections) bulunur.

3. Sınıflandırma (MLP Head): Kodlayıcı katmanları tamamlandığında, görüntü yamalarına ait vektörler göz ardı edilir. Sadece, tüm süreç boyunca görüntünün özet bilgisini toplayan 0* numaralı sınıflandırma vektörü (x_{class}) dışarı alınır. Bu vektör, son bir MLP katmanına (MLP Head) iletilerek nihai sınıf olasılıkları hesaplanır.

Bu çalışmada, ViT mimarisinin base (ViT-B/16) varyasyonu kullanılmıştır. Base ifadesi modelin 12 Kodlayıcı katmanına ve 768 gizli boyuta sahip olduğunu; 16

ifadesi ise giriş görüntüsünün piksellik yamalara bölündüğünü belirtir.

3.4.2.2 Swin Transformer

Standart Vision Transformer (ViT) mimarisi global bağlamı yakalamada başarılı olsa da, tüm yamaların birbirine bakması (global dikkat) nedeniyle hesaplama maliyeti görüntü çözünürlüğü ile karesel olarak artar. Bu durum, özellikle yüksek çözünürlüklü tıbbi görüntülerin işlenmesini zorlaştırır. Liu ve arkadaşları (2021), bu sorunu çözmek ve Transformer yapısını CNN'lerin hiyerarşik özelliğiyle birleştirmek amacıyla Swin Transformer mimarisini önermişlerdir [43].

Swin Transformer, ViT'ten iki temel yapısal yenilikle ayrılır:

1. Hiyerarşik Yapı ve Yama Birleştirme (Hierarchical Feature Maps): ViT boyunca yama boyutu sabit kalırken, Swin Transformer tıpkı bir CNN gibi derinleştikçe çözünürlüğü düşürür. Patch Merging katmanları sayesinde komşu yamalar birleştirilerek öznitelik haritasının boyutu küçültülür ve kanal sayısı artırılır. Bu yapı, modelin U-Net gibi segmentasyon mimarilerinde kodlayıcı (encoder) olarak kullanılabilmesine olanak tanır.

2. Kaydırılmış Pencere Dikkati (Shifted Window Attention): Hesaplama yükünü azaltmak için, dikkat mekanizması (MSA) tüm görüntü üzerinde değil, görüntüyü bölen yerel pencereler (Windows) içinde hesaplanır (W-MSA). Ancak bu durum, pencereler arasında bilgi akışını engeller.

Bu iletişimi sağlamak için, birbirini takip eden bloklarda pencereleme konumu kaydırılır (Shifted Window - SW-MSA). Böylece bir önceki katmanda farklı pencerelerde olan pikseller, bir sonraki katmanda aynı pencereye düşerek bilgi alışverişinde bulunur.

Matematiksel olarak, Swin Transformer bloğu her zaman ikili (çift) bir yapıdan oluşur. İlk blok standart pencereleme, ikinci blok ise kaydırılmış pencereleme uygular:

$$\widehat{z^l} = W - MSA\left(LN\left(z^{l-1}\right)\right) + z^{l-1} \tag{3.35}$$

$$z^l = MLP\left(LN\left(\widehat{z^l}\right)\right) + \widehat{z^l} \tag{3.36}$$

$$\widehat{z^{l+1}} = SW - MSA\left(LN\left(z^{l}\right)\right) + z^{l} \tag{3.37}$$

$$z^{l+1} = MLP\left(LN\left(\widehat{z^{l+1}}\right)\right) + \widehat{z^{l+1}} \tag{3.38}$$

Bu denklemlerde W-MSA standart pencere tabanlı dikkati,SW-MSA ise kaydırılmış pencere tabanlı dikkati temsil eder. MLP bloğu ise Bölüm 3.4.1.4'te tanımlanan yapı ile aynıdır.

Bu çalışmada, hesaplama verimliliği ve parametre sayısı (yaklaşık 29 milyon) dengesi nedeniyle Swin Transformer ailesinin en küçük üyesi olan Swin-T (Tiny) versiyonu tercih edilmiştir. Swin-T, pencere boyutlu 7×7 olan ve katman sayıları 2,2,6,2 şeklinde düzenlenen 4 aşamalı bir yapıya sahiptir.

3.5 Transfer Öğrenme Teorisi

Derin öğrenme modellerinin genelleme yeteneğinin yüksek olması ve aşırı öğrenme (overfitting) probleminden kaçınabilmesi için genellikle büyük ölçekli etiketli veri setlerine ihtiyaç duyulmaktadır. Ancak tıbbi görüntüleme alanında (beyin MR gibi), veri mahremiyeti ve uzman etiketleme maliyetleri nedeniyle büyük veri setlerine ulaşmak oldukça zordur. Bu kısıtlılığı aşmak amacıyla, kaynak bir görevde (source task) eğitilmiş modelin bilgisinin, ilişkili bir hedef göreve (target task) aktarılması prensibine dayanan Transfer Öğrenme yöntemi kullanılmaktadır [44].

3.5.1 Matematiksel Çerçeve ve Alan Tanımı

Pan ve Yang (2010) tarafından önerilen teorik çerçeveye göre, transfer öğrenme problemi alan (domain - D) ve görev (task -T) kavramlarıyla matematiksel olarak tanımlanır [44].

Bir alan D, iki bileşenden oluşur: Öznitelik uzayı X ve marjinal olasılık dağılımı $P(X)$.

$$D = \{X, P(X)\} \tag{3.39}$$

Burada X, Xuzayında tanımlı olan örneklem vektörlerini ifade eder ve matematiksel olarak $X = \{x_1, \ldots, x_n\} \in X$ şeklinde tanımlanır. Yani her bir veri örneği, tanımlı öznitelik uzayının bir elemanıdır.

Transfer öğrenme sürecinde temel amaç, kaynak alan (D_S) üzerindeki geniş veriyle optimize edilmiş model parametrelerinin (W), hedef alandaki (D_T) sınırlı veriyle güncellenmesidir. Bu çalışma kapsamında, Endüktif Transfer Öğrenme (Inductive Transfer Learning) stratejisi temel alınmıştır. Bu yaklaşımda, ağın kenar ve doku gibi evrensel görsel öznitelikleri öğrenen ilk katmanlarının ağırlıkları dondurularak (freezing) hedef modele aktarılmakta; son katmanlar ise hedef veri setine özgü soyut özellikleri öğrenmek üzere yeniden eğitilmektedir (fine-tuning).

3.5.2 ImageNet Veri Seti

Bilgisayarlı görü çalışmalarında altın standart olarak kabul edilen ImageNet, Deng ve ark. (2009) tarafından geliştirilen ve WordNet hiyerarşisine dayalı olarak düzenlenen devasa bir görsel veri tabanıdır. Veri seti, 1000 farklı nesne kategorisine (araçlar, hayvanlar, bitkiler vb.) ait 14 milyondan fazla yüksek çözünürlüklü etiketli görüntü içermektedir [45].

ImageNet'in derin öğrenme literatüründeki temel faydası, modellerin görme yetisini kazanmasını sağlamasıdır. Milyonlarca doğal görüntü üzerinde eğitilen evrişimli sinir ağları (CNN), görüntüdeki kenar, köşe, doku ve renk geçişleri gibi temel görsel öznitelikleri ayırt etmeyi öğrenir. Bu süreçte optimize edilen ağırlıklar, tıbbi görüntüleme dahil olmak üzere farklı alanlardaki problemler için güçlü bir özellik çıkarıcı (feature extractor) işlevi görür. Ancak ImageNet, RGB formatındaki doğal görüntülerden oluştuğu için, gri skala formatındaki radyolojik görüntülerle (MR, BT) arasında yapısal bir uyumsuzluk (alan kayması) barındırmaktadır.

3.5.3 RadImageNet Veri Seti

Tıbbi görüntüleme alanındaki veri azlığı sorununu çözmek ve doğal görüntülerden kaynaklanan uyumsuzluğu gidermek amacıyla Mei ve ark. (2022) tarafından RadImageNet veri seti literatüre sunulmuştur. RadImageNet, ImageNet'e kıyasla alana özgü (domain-specific) bir veri tabanıdır ve tamamen radyolojik görüntülerden oluşmaktadır [4].

Veri seti yaklaşık 1.35 milyon etiketli radyolojik görüntüden oluşmaktadır. İçeriğinde Beyin, göğüs, batın ve kas-iskelet sistemi dahil olmak üzere çok çeşitli antomik bölgeler ve patolojilere ait, bilgisayarlı tomografi (BT), MR ve ultrason görüntüleri bulunmaktadır.

RadImageNet'in en önemli avantajı, modelin eğitim sürecine anatomik yapıları ve radyolojik gürültü paternlerini tanıyarak başlamasını sağlamasıdır. Mei ve

arkadaşlarının yaptığı karşılaştırmalı analizlerde, RadImageNet ile ön eğitilmiş modellerin, ImageNet ile eğitilenlere kıyasla daha hızlı yakınsadığı ve daha yüksek doğruluk oranlarına ulaştığı raporlanmıştır.

Özellikle beyin patolojilerinin tespiti üzerine yapılan deneylerde (örneğin kafa içi kanama veya lezyon tespiti), RadImageNet ağırlıklarının kullanılması modelin ayırt ediciliğini artırmıştır. MR görüntülerindeki doku kontrastlarını ve tümör benzeri yapıları (nodül, kitle) tanıma yeteneğinin, doğal görüntülerle eğitilen modellere göre daha gelişmiş olduğu gözlemlenmiştir. Bu nedenle RadImageNet, beyin tümörü segmentasyonu ve sınıflandırması görevlerinde güçlü bir alternatif olarak değerlendirilmektedir [4].

3.6 Semantik Segmentasyon ve Mimari Yapılar

Bölüm 3.3 ve 3.4'te incelenen sınıflandırma modelleri, bir görüntüye tek bir etiket atamaktadır. Ancak cerrahi planlama ve radyoterapi gibi hassas tıbbi süreçlerde, patolojinin sadece varlığının bilinmesi yeterli değildir; tam konumunun ve sınırlarının belirlenmesi gerekmektedir. Bu gereksinim, görüntüdeki her bir pikselin ayrı ayrı sınıflandırıldığı Semantik Segmentasyon problemini ortaya çıkarmıştır.

Matematiksel olarak segmentasyon işlemi, bir giriş görüntüsü $I \in R^{H \times W \times C}$'nin, aynı uzaysal boyutlara sahip bir segmentasyon haritasına (maske) $M \in \{0,1\}^{H \times W}$ dönüştürülmesi fonksiyonu olarak tanımlanır. Burada $M_{i,j} = 1$değeri ilgili pikselin tümörlü dokuya, $M_{i,j} = 0$ değeri ise arka plana ait olduğunu ifade eder.

Bu dönüşümü sağlamak için literatürde genellikle kodlayıcı-kod çözücü (encoder-Decoder) tabanlı mimariler tercih edilmektedir.

3.6.1 Mimari Elemanları

3.6.1.1 Kodlayıcı-Kod Çözücü (Encoder-Decoder) Yapısı

Bu mimari paradigması, birbirini tamamlayan ve ardışık çalışan iki ana bloktan oluşur:

Kodlayıcı (Encoder): Görüntüden öznitelik çıkaran bölümdür. Genellikle CNN veya Transformer tabanlı katmanlardan oluşur. Havuzlama işlemleriyle görüntünün uzaysal boyutu (H,W) küçültülürken, anlamsal derinliği (kanal sayısı) artırılır. Bu süreçte elde edilen düşük çözünürlüklü ancak yoğun bilgi içeren tensöre gizli uzay (Latent Space) adı verilir.

Kod Çözücü (Decoder): Sıkıştırılmış öznitelik haritasını tekrar orijinal görüntü boyutuna ($H \times W$) genişleten bölümdür. Bu işlem için genellikle ters evrişim (Transposed Convolution) veya yukarı örnekleme (Upsampling) algoritmaları kullanılır.

3.6.1.2 U-Net Mimarisi ve Atlamalı Bağlantılar

Tıbbi görüntü segmentasyonu alanında devrim yaratan U-Net mimarisi, Ronneberger ve ark. (2015) tarafından geliştirilmiştir [29]. Standart kodlayıcı-kod çözücü yapılarında, boyut küçültme işlemleri sırasında piksel düzeyindeki hassas konum bilgisi (spatial information) kaybolma eğilimindedir. U-Net, bu sorunu çözmek için atlamalı bağlantılar (skip connections) mekanizmasını kullanır.

U-Net mimarisi, ismini aldığı "U" şeklinde simetrik bir yapıya sahiptir:

Daralan Yol (Contracting Path): Görüntünün bağlamını (context) ve özniteliklerini yakalar.

Genişleyen Yol (Expansive Path): Hassas konumlandırma (localization) yapar.

Matematiksel olarak, kod çözücüdeki l.katmanın girdisi ($h_{dec}^{(l)}$), sadece bir önceki katmandan gelen yukarı örneklenmiş veriden değil, aynı zamanda kodlayıcıdaki karşılık gelen katmandan ($h_{enc}^{(l)}$) gelen öznitelik haritasının birleştirilmesiyle (concatenation) oluşturulur. Bu işlem eşitlik (3.40) ile ifade edilir:

$$h_{dec}^{(l)} = Concat\left(h_{enc}^{(l)}, UpSample\left(h_{dec}^{(l-1)}\right)\right) \tag{3.40}$$

Bu bağlantılar sayesinde, kodlayıcının ilk katmanlarında öğrenilen yüksek çözünürlüklü kenar ve doku bilgileri, doğrudan kod çözücüye aktarılarak segmentasyon maskesinin sınırlarının (tümör kenarlarının) çok daha keskin ve doğru tahmin edilmesi sağlanır.

3.6.1.3 Piksel Tabanlı Sınıflandırma ve Kayıp Fonksiyonu

Ağın son katmanında, kanal sayısını sınıf sayısına indirgemek için 1×1boyutunda evrişim işlemi uygulanır. Her bir piksel ($x_{i,j}$) için bir olasılık değeri ($p_{i,j}$) üretmek amacıyla Sigmoid aktivasyon fonksiyonu (eşitlik (3.41)) kullanılır. Kullanılan fonksiyon eşitlik 3.25'teki gibidir:

$$p_{i,j} = \frac{1}{1 + e^{-z_{i,j}}} \tag{3.41}$$

Tıbbi görüntü segmentasyonunda, tümörlü bölgenin görüntü geneline oranla çok küçük yer kaplaması nedeniyle sınıf dengesizliği (class imbalance) problemi ortaya çıkmaktadır. Bu durum, standart hata fonksiyonlarının (örneğin Çapraz Entropi) modelin eğitiminde yetersiz kalmasına ve ağın arka planı ezberlemesine yol açabilmektedir.

Bu problemi aşmak ve segmentasyon başarımını maksimize etmek amacıyla, optimizasyon sürecinde eşitlik (3.8) ile matematiksel ifadesi verilen Dice Loss (kayıp fonksiyonu) kullanılır. Bu fonksiyon, piksel bazlı doğruluktan ziyade, tahmin edilen maske ile gerçek etiket (Ground Truth) arasındaki alansal örtüşmeyi maksimize etmeye odaklanarak modelin küçük ve düzensiz sınırlara sahip tümörleri daha hassas öğrenmesini sağlar.

3.6.2 Segmentasyon Mimari Modelleri

Segmentasyon mimarisi kullanılarak çeşitli modeller geliştirilmiştir. Bu çalışmada kullanılan modeller aşağıda tanıtılmıştır.

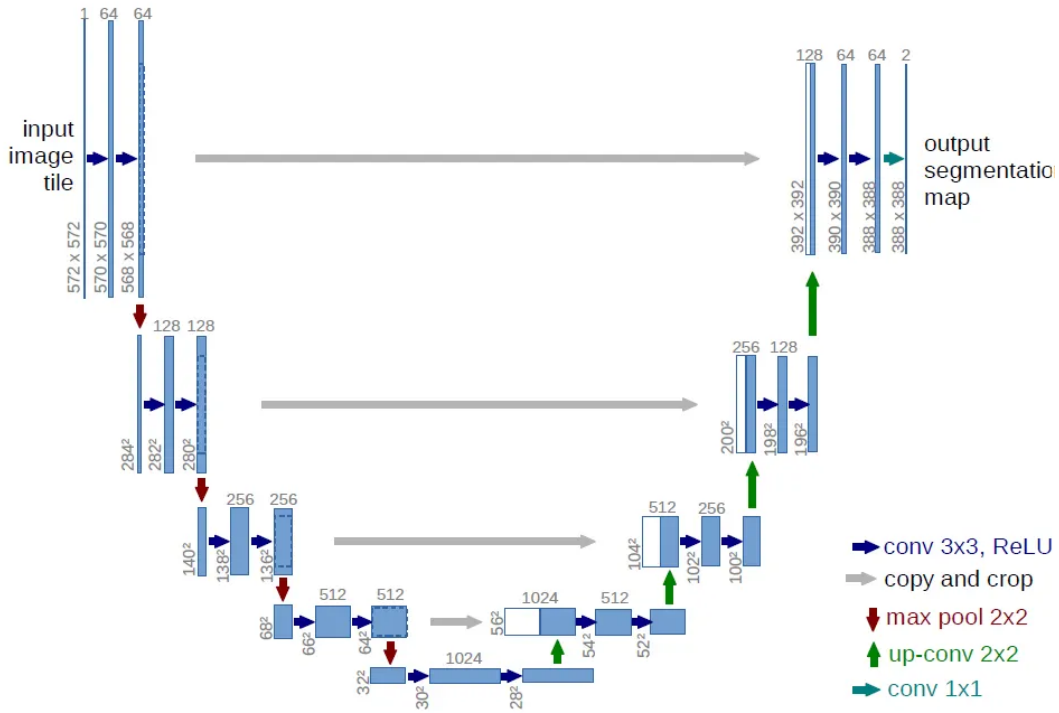
3.6.2.1 U-Net Mimarisi

Tıbbi görüntü analizinde, sadece bir görüntünün hasta olup olmadığını (sınıflandırma) belirlemek her zaman yeterli değildir; patolojinin tam konumunu ve sınırlarını belirlemek (segmentasyon) de hayati önem taşır. Bu ihtiyacı karşılamak için Ronneberger ve arkadaşları (2015), az sayıda etiketli görüntüyle dahi yüksek hassasiyetli segmentasyon yapabilen U-Net mimarisini geliştirmişlerdir [29].

Mimarinin Yapısı: Adını "U" harfine benzeyen simetrik yapısından alan bu mimari, temelde tam evrişimli ağlar (Fully Convolutional Networks - FCN) prensibine dayanır ancak önemli yapısal iyileştirmeler içerir.

Mimari iki ana yoldan oluşur:

Daralan Yol (Contracting Path / Encoder): Klasik bir CNN mimarisi gibidir. Görüntüden öznitelik çıkarmak için 3×3 konvolüsyonlar ve ReLU aktivasyonları kullanılır. Her adımda 2×2 Maksimum Havuzlama (Max Pooling) işlemi uygulanarak görüntünün uzaysal boyutu yarıya indirilirken, kanal sayısı (derinlik) iki katına çıkarılır. Bu kısım, görüntünün ne içerdiğini (bağlamı) öğrenir.



Şekil 3.6 Unet mimarisinin temsili görseli [29].

Genişleyen Yol (Expanding Path / Decoder): Bu kısım, sıkıştırılmış öznitelik haritasını tekrar orijinal görüntü boyutuna getirmeyi hedefler. Her adımda yukarı evrişim (Up-Convolution / Transposed Conv) işlemi uygulanarak uzaysal boyut iki katına çıkarılır. Bu kısım, nesnenin nerede olduğunu (lokalizasyonu) öğrenir.

U-Net'in başarısının sırrı, kodlayıcı (encoder) ile kod çözücü (decoder) katmanları arasındaki atlamalı bağlantılardır.

Normal bir FCN'de, havuzlama işlemleri sırasında mekansal (yerel) bilgi kaybolur. U-Net, kodlayıcı kısmındaki yüksek çözünürlüklü özellik haritalarını kopyalar ve bunları kod çözücüdeki karşılık gelen katmanları birleştirme (Concatenation) yoluyla aktarır.

Bir U-Net bloğundaki işlem akışı şu şekil (3.42) ile formülize edilebilir:

l.katmandaki genişleyen yolun çıktısı (x_l), bir önceki katmanın yukarı örneklenmiş hali ($Up(x_{l-1})$) ile kodlayıcıdan gelen karşılık gelen haritanın (c_l) birleşimidir:

$$x_l = \sigma(W_l * [Up(x_{l-1}); c_l] + b_l) \tag{3.42}$$

Burada:

$Up(\bullet)$: Yukarı örnekleme (Up-sampling) işlemi,

c_l: Atlamalı bağlantı ile gelen kodlayıcı öznitelik haritası,

$[;]$: Birleştirme (Concatenation) operatörü,

$*$: Konvolüsyon işlemini temsil eder.

3.6.2.2 UNet++ (Nested U-Net)

Standart U-Net'te, kodlayıcıdan gelen sığ (düşük seviyeli) özellikler ile kod çözücüdeki derin (yüksek seviyeli) özellikler, atlamalı bağlantılar (skip connections) aracılığıyla doğrudan birleştirilir. Ancak Zhou ve arkadaşları (2018), bu iki özellik haritası arasında anlamsal bir boşluk (semantic gap) olduğunu savunmuşlardır [46].

Bu boşluğu doldurmak için geliştirilen UNet++, kodlayıcı ve kod çözücü arasına yuvalanmış (Nested) yoğun konvolüsyon blokları yerleştirir.

UNet++'ın en belirgin özelliği, atlamalı bağlantıların yeniden tasarlanmış olmasıdır. Bir düğüm (node) , sadece altındaki katmandan değil, aynı seviyedeki önceki tüm konvolüsyon bloklarından girdi alır. Bu yapı dense convolution mantığına benzer.

Bir $x^{i,j}$düğümünün çıkışı eşitlik (3.43) ile hesaplanır:

$$x^{i,j} = H\left(\left[x^{i-1,j}, U\left(x^{i+1,j-1}\right), x^{i,0}, x^{i,1}, \ldots, x^{i,j-1}\right]\right) \tag{3.43}$$

Burada:

$H(\bullet)$: Konvolüsyon ve aktivasyon işlemi,

$U(\bullet)$: Yukarı örnekleme işlemi,

$[\bullet]$: Birleştirme (concatenation) işlemidir.

UNet++, ağın sadece sonundan değil, ara katmanlarından da çıktı alarak (pruning) eğitilebilir. Bu sayede, gradyan akışı iyileştirilir ve model hem kolay hem de zor segmentasyon görevlerini öğrenebilir.

3.6.2.3 DeepLabV3+

Segmentasyondaki en büyük zorluklardan biri, nesnelerin çok farklı ölçeklerde (çok küçük veya çok büyük tümörler) olabilmesidir. Google araştırmacıları Chen ve arkadaşları (2018), bu sorunu çözmek için DeepLabV3+ mimarisini geliştirmişlerdir [47]. Mimarinin kalbinde, çözünürlüğü düşürmeden alıcı alanı (receptive field) genişleten seyreltilmiş evrişim (Atrous/Dilated Convolution) yatar.

DeepLabV3+, iki ana bileşenden oluşur:

Kodlayıcı (ASPP Modülü): Atrous Spatial Pyramid Pooling (ASPP) modülü, görüntüyü farklı oranlarda ($r = 6,\ 2,\ 18$) seyreltilmiş filtrelerle tarayarak çok ölçekli bağlamı yakalar.

Kod Çözücü: DeepLabV3'ten farklı olarak V3+ versiyonu, basit bir yukarı örnekleme yerine, keskin nesne sınırlarını (boundary refinement) oluşturmak için özel bir kod çözücü modülü ekler.

Standart konvolüsyonda filtre bitişik piksellere uygulanırken, seyreltilmiş evrişimde filtre ağırlıkları arasına $r-1$adet sıfır eklenir. 1-boyutlu bir $y[i]$ sinyali için işlem şöyledir:

$$y[i] = \sum_k x[i + r \cdot k]\, w[k] \tag{3.44}$$

Burada r(rate), genişleme oranıdır. $r = 1$olduğunda işlem standart konvolüsyona döner. Bu teknik, parametre sayısını artırmadan ağın görüş açısını genişletir.

3.6.2.4 Linknet

Tıbbi görüntüleme cihazlarında veya mobil uygulamalarda gerçek zamanlı segmentasyon yapabilmek için hız ve verimlilik kritiktir. Chaurasia ve Culurciello (2017) tarafından önerilen Linknet, parametre sayısını minimumda tutarak yüksek performans vermeyi hedefler [48].

Linknet, kodlayıcı olarak genellikle ResNet-18 gibi hafif bir omurga (backbone) kullanır. Mimarinin yeniliği, kodlayıcı ve kod çözücü blokları arasındaki bağlantı şeklindedir.

U-Net'in aksine, Linknet özellik haritalarını birleştirmez (concatenation), toplar (addition). Bu işlem bellek kullanımını ciddi oranda azaltır. Kodlayıcı bloğunun çıktısı x olsun. Linknet kod çözücü bloğu şu işlemi yapar:

$$y = Decoder(x) + x \tag{3.45}$$

Bu artık bağlantı (residual connection) mantığı, kod çözücünün sıfırdan öğrenmek yerine sadece farkı veya kayıp detayları öğrenmesine odaklanmasını sağlar. Bu sayede model çok daha hızlı eğitilir (convergence) ve çalışma zamanında (inference) yüksek FPS değerlerine ulaşır.

3.6.2.5 MAnet (Multi-scale Attention Net)

Standart CNN tabanlı segmentasyon modelleri, yerel özelliklere odaklanırken global bağlamı ve kanallar arası ilişkileri bazen kaçırabilir. Fan ve arkadaşları (2020), bu sorunu çözmek için MAnet (Multi-scale Attention Net) mimarisini önermişlerdir [49]. Temel amaç, dikkat (Attention) mekanizmasını kullanarak ağın neye ve nereye odaklanması gerektiğini dinamik olarak belirlemektir.

MAnet, ResNet gibi bir omurga üzerine inşa edilir ancak iki kritik modül ekler:

Konum Bazlı Dikkat Bloğu (PAB): Görüntüdeki her pikselin diğer tüm piksellerle olan uzaysal ilişkisini (global context) modeller. (self-attention mekanizmasına benzer).

Çok Ölçekli Füzyon Dikkat Bloğu (MFAB): Farklı ölçeklerdeki özellik haritaları (low-level ve high-level) arasındaki anlamsal boşluğu, kanal bazlı dikkat mekanizmasıyla birleştirerek doldurur.

Dikkat mekanizması genellikle özellik haritaları (A) arasındaki benzerliği hesaplayarak bir dikkat matrisi (X) oluşturur:

$$x_{ji} = \frac{\exp\left(A_i \cdot A_j\right)}{\sum_{i=1}^{N} \exp\left(A_i \cdot A_j\right)} \tag{3.46}$$

Bu formül (Softmax), j. pikselin i. piksel üzerindeki etkisini (dikkat ağırlığını) belirler. MAnet, bu ağırlıkları kullanarak tümörlü bölgeyi arka plandan çok daha keskin bir şekilde ayırt edebilir.

4
METODOLOJİ VE ÖNERİLEN YAKLAŞIM

Bu bölümde, tez çalışması kapsamında kullanılan veri setleri, deneylerin gerçekleştirildiği donanım altyapısı, uygulanan veri ön işleme yöntemleri ve seçilen derin öğrenme modellerinin mimari detayları sunulmuştur.

4.1 Donanım ve Yazılım Altyapısı

Tez çalışması kapsamındaki tüm model eğitimleri ve test süreçleri, derin öğrenme hesaplamaları için optimize edilmiş bir iş istasyonu üzerinde gerçekleştirilmiştir. Donanım hızlandırıcı olarak NVIDIA GeForce GTX 1050 grafik işlem birimi (GPU) kullanılmıştır. Sahip olduğu 3GB VRAM bellek kapasitesi, özellikle 3 boyutlu verilerin işlenmesi ve büyük parametreli modellerin eğitimi noktasında çalışmanın sınırlarını belirleyen temel faktör olmuştur. Bu donanım kısıtı, metodoloji kısmında detaylandırılacak olan 3D'den 2D'ye dönüşüm ve optimize edilmiş yığın boyutu (batch size stratejilerinin geliştirilmesini zorunlu kılmıştır.

Yazılım tarafında, tüm algoritmalar Python programlama dili kullanılarak geliştirilmiştir. Derin öğrenme çerçevesi (framework) olarak, dinamik hesaplama grafiği yeteneği ve geniş topluluk desteği nedeniyle PyTorch kütüphanesi tercih edilmiştir. Görüntü işleme ve veri artırma operasyonları için OpenCV ve Albumentations kütüphanelerinden, medikal görüntü formatlarının (.nii, .dcm) okunması için ise Nibabel kütüphanesinden yararlanılmıştır.

4.2 Kullanılan Veri Setleri

Bu tez çalışmasında, önerilen derin öğrenme modellerinin başarımını farklı zorluk seviyelerinde ve farklı görüntüleme senaryolarında test etmek amacıyla, literatürde kabul görmüş temel veri kaynakları ve yeni hazırlanmış açık kaynaklı veri kaynakları kullanılmıştır. Bu veri setleri; tümör varlığının tespit edildiği ikili sınıflandırma, tümör tipinin belirlendiği çoklu sınıflandırma ve lezyon sınırlarının

belirlendiği segmentasyon görevlerine göre özelleştirilmiştir. Aşağıda, kullanılan veri setlerinin yapısal özellikleri ve içerikleri detaylandırılmıştır.

4.2.1 BR35H (Brain Tumor Detection 2020)

İkili sınıflandırma (Binary Classification) deneylerinde, modelin sadece hastalık var/yok ayrımındaki hassasiyetini ölçmek amacıyla BR35H veri seti kullanılmıştır. Hamada ve ark. tarafından yayımlanan bu veri seti, toplamda 3.000 adet 2 boyutlu (2D) MR görüntüsünden oluşmaktadır [50]. Veri seti, 1.500 adet tümörlü ve 1.500 adet tümörsüz (sağlıklı) görüntü içerecek şekilde mükemmel bir dengeye sahiptir. Segmentasyon maskesi içermeyen bu koleksiyon, özellikle sistemin ilk aşaması olan lezyon tespit modülünün eğitimi için ideal bir referans noktası oluşturmuştur. Birçok çalışmada tercih edilen veri setinin ağırlıkları karışık yapıdadır ve sadece yatay kesitli görüntülerden oluşmaktadır.

4.2.2 Beyin Tümörü MRG Veri Seti (Nickparvar)

Çoklu sınıflandırma (Multi-class Classification) görevi için ana veri kaynağı olarak, Nickparvar (2021) tarafından derlenen ve dört farklı sınıfı kapsayan geniş ölçekli veri seti tercih edilmiştir [51]. Toplam 7.023görüntüden oluşan bu set; 1.621 Glioma, 1.645 Meningioma, 1.757 Hipofiz (Pituitary) tümörü ve 2.000 sağlıklı (No Tumor) görüntü içermektedir. Bu veri seti aslında, literatürdeki diğer önemli kaynakların (Figshare, SARTAJ ve BR35H) birleştirilmesiyle oluşturulmuş kompozit bir yapıdır. Figshare veri setinden gelen 3.064 görüntü ve SARTAJ veri setinden gelen eklemeler, modelin genelleştirme yeteneğini artırmak için zengin bir çeşitlilik sunmaktadır. Ancak bu birleşim, metodoloji bölümünde belirtilen algısal hashleme ile veri temizliği yapılmasını zorunlu kılan mükerrer veri riskini de beraberinde getirmiştir. Sınıflar arası dengeli dağılımı sayesinde, özellikle transfer öğrenme tabanlı modellerin eğitiminde temel kaynak olarak kullanılmıştır. Ayrıca veri seti T1 ağırlıklı olacak şekilde ayarlanmış ve farklı ağırlıklardaki görüntüler temizlenmiştir.

4.2.3 BRISC 2025 (Brain Imaging Segmentation and Classification)

Çalışmanın hem sınıflandırma hem de segmentasyon aşamalarında kullanılan en güncel ve kapsamlı veri kaynağı BRISC 2025 veri setidir [52]. Toplam 6.000 adet yüksek çözünürlüklü T1-kontrastlı MR dilimi içeren bu veri seti; Glioma (1.401), Meningioma (1.635), Hipofiz (1.757) ve Tümörsüz (1.207) olmak üzere dört sınıfa ayrılmıştır. BRISC veri setini diğerlerinden ayıran en önemli özellik, görüntülerin sadece tek bir düzlemden değil; aksiyal, koronal ve sagital

düzlemlerden dengeli bir şekilde alınmış olmasıdır. Ayrıca her görüntü, uzman radyologlar tarafından hazırlanmış hassas segmentasyon maskeleri ile birlikte sunulmaktadır. Bu çalışma kapsamında BRISC 2025, hem çoklu sınıflandırma performansının doğrulanmasında hem de tek kanallı segmentasyon modellerinin (U-Net, Linknet vb.) eğitiminde kullanılmıştır.

4.2.4 BraTS 2020 (Multimodal Brain Tumor Segmentation Challenge)

Çok modlu (multi-modal) segmentasyon başarısının değerlendirilmesi amacıyla, tıbbi görüntüleme alanında altın standart olarak kabul edilen BraTS 2020 veri seti kullanılmıştır [53–57]. Veri seti, 293'ü Yüksek Dereceli Glioma (HGG) ve 76'sı Düşük Dereceli Glioma (LGG) olmak üzere toplam 369 hasta vakasından oluşmaktadır. Her bir vaka için T1, T1+CE (kontrastlı), T2 ve FLAIR olmak üzere dört farklı MR sekansı 3 boyutlu (3D) NIfTI formatında sunulmuştur. BraTS veri seti, tümörün alt bölgelerini tanımlayan üç farklı etiket (maske) içerir: Tüm tümör dokusunu kapsayan whole tumor (WT), nekrotik ve kistik alanları içeren tumor core (TC) ve aktif tümör dokusunu gösteren enhancing tumor (ET). Bu çalışmada, donanım kısıtları nedeniyle 3D hacimler 2D kesitlere ayrılarak kullanılmış ve farklı MR sekanslarının bu tümör alt bölgelerini tespit etmedeki başarısı analiz edilmiştir.

Tablo 4.1 Kullanılan Veri Setlerinin Genel Özellikleri

Veri Seti	**Sayı**	**Sınıf Sayısı /Türleri**	**Tip**	**Seg. Maskesi**	**Kullanım Amacı**
BR35H	3.000 G	2 Sınıf: Tümör Sağlıklı	2D MRG	Yok	İkili Sınıflandırma (Tespit)
Nickparvar	7.023 G	4 Sınıf: Glioma, Meningioma, Hipofiz, Sağlıklı	2D MRG (T1)	Yok	Çoklu Sınıflandırma
BRISC 2025	6.000 G	4 Sınıf: Glioma, Meningioma, Hipofiz, Sağlıklı	2D MRG (3 Düzlem)	Var (Tek Kanal)	Sınıflandırma ve Segmentasyon
BraTS 2020	369 Vaka	Glioma 2 Grade: HGG, LGG	3D MRG (4 Sekans)	Var (3 Kanal: WT, TC, ET)	Çok Modlu Segmentasyon

4.3 Önerilen Sistem Mimarisi

Bu tez çalışmasında, beyin tümörlerinin analizi için veri hazırlığından başlayıp nihai klinik çıktıların (sınıf etiketi ve tümör maskesi) üretilmesine kadar uzanan, hiyerarşik ve çok modüllü bir derin öğrenme mimarisi tasarlanmıştır. Sistem temel olarak, patolojik durumun tespit edildiği sınıflandırma modülü ve lezyon sınırlarının belirlendiği segmentasyon modülü olmak üzere birbirini tamamlayan iki ana aşamadan oluşmaktadır. Her iki modül de farklı veri kaynaklarını ve görüntüleme senaryolarını işleyecek şekilde özelleştirilmiş alt akışlara sahiptir.

Sistemin ilk aşaması olan Sınıflandırma Modülü, problemin karmaşıklığına göre iki kola ayrılmaktadır. İlk kolda, lezyonun varlığını veya yokluğunu tespit etmek amacıyla ikili (binary) sınıflandırma görevi yürütülürken; ikinci kolda Glioma, Meningioma ve Hipofiz tümörlerini ayırt eden çoklu sınıflandırma görevi gerçekleştirilmektedir. Mimarinin bu aşamasındaki en kritik tasarım tercihi, veri çeşitliliğini artırmak için farklı veri setlerinin (Nickparvar ve BRISC2025) birleştirilmesidir. Ancak bu birleştirme işlemi doğrudan yapılmamış, veri tekrarını önlemek ve modelin ezberlemesini (data leakage) engellemek amacıyla algısal hashleme (perceptual hashing) tabanlı bir filtreleme katmanı sisteme entegre edilmiştir. Temizlenen ve birleştirilen veriler, sistem içerisinde iki paralel hat üzerinden işlenmektedir: Bir hatta ham (RAW) görüntüler, diğer hatta ise kafatası ve beyin dışı dokuların temizlendiği (Brain Extracted - BE) görüntüler kullanılarak ön işlemenin model başarısına etkisi sistemsel olarak analiz edilmektedir.

Sistemin ikinci aşaması olan Segmentasyon Modülü ise tek kanallı ve çok kanallı görüntüleme senaryolarını kapsayacak şekilde kademeli bir yapıya sahiptir. İlk adımda, tek sekanslı görüntüler üzerinden tümörün genel sınırları tespit edilmektedir. İkinci ve daha ileri adımda ise BraTS veri seti kullanılarak T1, T2, FLAIR ve T1+CE sekanslarının bir arada değerlendirildiği çok modlu segmentasyon süreci işletilmektedir. Mevcut donanım altyapısının 3 boyutlu (3D) hacimsel verileri işleme kısıtını aşmak için sisteme özgün bir "Boyut İndirgeme ve Dilimleme" katmanı eklenmiştir. Bu katman, 3D MRG hacimlerini aksiyal, koronal ve sagital düzlemlerde 2D kesitlere ayırarak, 2 boyutlu ağların (U-Net vb.) 3 boyutlu uzaysal bilgiyi kaybetmeden eğitilebilmesini sağlamaktadır. Ayrıca bu modül, tümörün alt bölgelerini (tümör çekirdeği, ödem, genişleyen tümör) temsil eden 3 kanallı maske yapısı ile çıktı üreterek, hangi MR sekansının hangi tümör bölgesini daha iyi temsil ettiğinin belirlenmesine olanak tanımaktadır.

4.3.1 Veri Setlerinin Düzenlenmesi ve Hazırlık Süreçleri

Önerilen derin öğrenme mimarisinin başarısı, yüksek kaliteli ve doğru etiketlenmiş veri setlerinin varlığına doğrudan bağlıdır. Bu tez çalışmasında, literatürdeki mevcut veri setleri ham halleriyle kullanılmamış; veri birleştirme, temizleme, boyut indirgeme ve sızıntı önleme (leakage prevention) gibi bir dizi metodolojik işlemden geçirilerek sisteme entegre edilmiştir. Veri hazırlık süreci, sınıflandırma ve segmentasyon görevlerinin gereksinimlerine göre iki ana başlık altında kurgulanmıştır.

4.3.2 Sınıflandırma Veri Setleri: Birleştirme ve Algısal Temizleme

Sınıflandırma modülü için en büyük zorluk, derin mimarileri (ViT, Swin vb.) eğitebilecek çeşitlilikte ve büyüklükte bir veri havuzu oluşturmaktır. İkili sınıflandırma (Tümör Var/Yok) görevi için yapısı itibarıyla dengeli olan Br35H veri seti doğrudan kullanılırken, çoklu sınıflandırma (Glioma, Meningioma, Hipofiz, Tümörsüz) görevi için Nickparvar ve BRISC 2025 veri setlerinin birleştirilmesi stratejisi izlenmiştir. Farklı kaynaklardan gelen bu verilerin birleştirilmesi, eğitim setinin çeşitliliğini artırsa da veri tekrarı (duplication) riskini beraberinde getirmektedir. Aynı görüntünün hem eğitim hem de test setinde yer alması, modelin başarı skorlarını yapay olarak yükselteceğinden (veri sızıntısı), birleştirilen havuz üzerinde titiz bir temizlik işlemi uygulanmıştır.

Geleneksel dosya adı veya MD5 kontrolü, görüntü formatı değiştiğinde (örneğin JPG'den PNG'ye) işlevsiz kaldığından, bu çalışmada algısal hashleme (perceptual hashing - pHash) yöntemi kullanılmıştır. Bu yöntemle, her görüntünün görsel içeriğini temsil eden dijital parmak izleri (hash kodları) üretilmiş ve hamming mesafesi kullanılarak benzerlik analizi yapılmıştır. Görsel olarak birebir aynı veya aşırı benzer olan kopyalar tespit edilerek veri setinden çıkarılmış, böylece temizlenmiş birleşik veri seti (Combined 4 class - c4c) oluşturulmuştur.

Elde edilen bu temiz veri seti ve binaary veri seti üzerinden, ön işlemenin etkisini ölçmek amacıyla iki paralel veri hattı türetilmiştir:

Ham (RAW) Veri Seti: Görüntülerin sadece boyutlandırıldığı ve normalize edildiği set.

Kafatası Temizlenmiş (Brain Extracted - BE) Veri Seti:Görüntüden kafatası, göz ve beyin dışı dokuların temizlendiği set.

4.3.3 Segmentasyon Veri Setleri: 3D-2D Dönüşümü ve Dilimleme

Segmentasyon modülünde, özellikle BraTS 2020 veri setinin kullanımı, donanımsal kısıtlar nedeniyle özel bir düzenleme gerektirmiştir. BraTS veri seti, doğası gereği 3 boyutlu (3D) hacimsel MR verilerinden oluşmaktadır. Ancak çalışmada kullanılan NVIDIA GTX 1050 (3GB) donanımı, 3D Konvolüsyonel Sinir Ağlarını (3D U-Net vb.) eğitmek için yeterli bellek kapasitesine sahip değildir.

Bu engeli aşmak ve 3D verinin içerdiği uzaysal bilgiyi kaybetmemek için sistemsel bir boyut indirgeme ve dilimleme (Slicing) stratejisi uygulanmıştır. 3D MR hacimleri; aksiyal, koronal ve sagital olmak üzere üç farklı anatomik düzlem boyunca kesilerek 2 boyutlu (2D) dilimlere dönüştürülmüştür. Bu işlem sırasında, herhangi bir tümör bilgisi içermeyen maske bilgisi boş dilimler elenerek veri setinin verimliliği artırılmıştır. Ayrıca BraTS veri setinin karmaşık etiket yapısı (nekroz, ödem vb.), klinik olarak anlamlı üç kanallı bir maske yapısına (Tümör Çekirdeği, Genişleyen Tümör, Tüm Tümör) dönüştürülmüştür. Böylece 2D mimariler (U-Net, MAnet vb.) kullanılarak, 3D veri setinin zenginliğinden faydalanılması sağlanmıştır.

BraTS veriseti kafatası ayrılmış olarak bulunduğundan bir işlem yapılmazken, BRISC 2025 segmentasyon taskı içerisindeki görüntüler ham veri ve kafatası ayrılmış (BE) veri setileri hazılanmıştır.

4.3.4 Veri Bölümleme ve Tutarlılık Kontrolü

Hazırlanan tüm veri setleri (sınıflandırma ve segmentasyon), modelin genelleştirme yeteneğini doğru ölçmek amacıyla %70 eğitim (training), %15 doğrulama (validasion) ve %15 test oranlarında rastgele bölünmüştür.

Buradaki en kritik metodolojik detay, RAW ve BE veri setleri arasındaki bölümleme tutarlılığıdır (Split Consistency). Model performansını karşılaştırırken "Kafatası temizleme işlemi başarıyı artırdı mı?" sorusuna doğru cevap verebilmek için, her iki veri setinin de birebir aynı hasta görüntülerini içermesi gerekmektedir. Bu nedenle, önce ana veri seti indeks bazlı olarak bölünmüş, ardından bu indeksler hem RAW hem de BE veri setlerine uygulanmıştır. Örneğin, 105 numaralı hasta RAW setinde eğitim grubundaysa, BE setinde de mutlaka eğitim grubunda yer almıştır. Bu yaklaşım, sonuçların veri dağılımındaki rastgelelikten değil, uygulanan yöntemin (BE veya RAW) kendisinden kaynaklandığını garanti altına almıştır.

4.4 Veri Ön İşleme ve Zenginleştirme

Derin öğrenme modellerinin başarısı, sisteme beslenen verinin kalitesiyle doğrudan ilişkilidir. Ham MR görüntüleri; farklı çözünürlükler, yoğunluk farkları ve beyin dışı anatomik yapılar (kafatası, göz, yağ dokusu) içerdiğinden, modelin sadece ilgili patolojiye odaklanmasını sağlamak amacıyla çok aşamalı bir ön işleme hattı uygulanmıştır.

4.4.1 Beyin Çıkarma (Brain Extraction / Skull-Stripping)

Bu tez çalışmasında, modelin sınıflandırma ve segmentasyon performansını artırmak amacıyla, görüntüden kafatası ve beyin dışı dokuların temizlendiği Brain Extracted (BE) veri setleri oluşturulmuştur. Bu işlem için hazır kara kutu araçlar yerine, Python kütüphaneleri (scipy.ndimage, numpy) kullanılarak geliştirilen, morfolojik operasyonlara dayalı özgün bir algoritma kullanılmıştır.

Uygulanan Beyin Çıkarma algoritması şu adımları izler:

Gürültü Azaltma: İlk adımda, görüntü üzerindeki yüksek frekanslı gürültüleri bastırmak için Gaussian Bulanıklaştırma (Blur) filtresi uygulanır.

Otsu Eşikleme (Otsu's Thresholding): Görüntü histogramı analiz edilerek, ön plan (kafa yapısı) ile arka planı en iyi ayıran dinamik eşik değeri hesaplanır ve ikili (binary) maske oluşturulur.

Morfolojik İyileştirme: Oluşturulan maske üzerinde, beyin dokusu içindeki küçük delikleri kapatmak için Morfolojik Kapanma (Closing) ve beyin dışındaki gürültüleri (artefaktlar) temizlemek için Morfolojik Açılma (Opening) işlemleri uygulanır.

En Büyük Bağlantılı Bileşen Analizi (LCC): Eşikleme sonucunda maskede kafatası parçaları veya etiket yazıları kalabilmektedir. Algoritma, maske üzerindeki tüm bağımsız nesneleri tarar, piksel alanı en büyük olan nesneyi (beyin parankimi) korur ve diğer tüm küçük parçaları maskeden siler.

Elde edilen nihai maske, orijinal görüntü ile çarpılarak sadece beyin dokusunun kaldığı, arka planın tamamen siyah olduğu BE görüntüleri üretilir.

4.4.2 Görüntü Boyutlandırma ve Kanal Dönüşümü

Derin öğrenme modelleri, sabit boyutlu giriş tensörleri ile çalışacak şekilde tasarlandığından, farklı çözünürlüklere sahip olan MR görüntüleri standart bir

formata getirilmiştir. Veri setindeki (Br35H, Nickparvar, BRISC vb.) tüm görüntüler, kullanılan mimarinin giriş matris boyutuna uygun olarak yeniden ölçeklendirilmiştir. Bu işlemde bilgi kaybını minimize etmek amacıyla Bilineer Enterpolasyon (Bilinear Interpolation) yöntemi kullanılmıştır. ResNet, DenseNet, Swin Transformer ve ViT modelleri için standart boyut 224×224 piksel olarak belirlenirken, mimari yapısı daha geniş bir alıcı alan gerektiren InceptionV3 modeli için görüntüler 299×299 piksel boyutuna dönüştürülmüştür.

Boyutlandırmanın yanı sıra, önemli bir format dönüşümü de uygulanmıştır. MR görüntüleri doğası gereği tek kanallı (gri tonlamalı) verilerdir. Ancak bu tezde kullanılan transfer öğrenme modelleri (ImageNet ağırlıklı), üç kanallı (RGB) renkli görüntüler üzerinde eğitilmiştir. Ağırlık uyumsuzluğunu gidermek ve model mimarisini bozmamak amacıyla, tek kanallı gri görüntüler kanal ekseninde kopyalanarak yapay bir RGB (3 kanallı) formata dönüştürülmüştür. Bu işlem, Python görüntü işleme kütüphanesi (PIL) kullanılarak veri yükleme aşamasında otomatikleştirilmiştir.

4.4.3 Çevrimiçi Veri Artırma (On-the-fly Augmentation)

Sınırlı sayıda tıbbi görüntü ile eğitilen derin ağların, eğitim verisini ezberlemesini (overfitting) engellemek ve modelin genelleştirme yeteneğini artırmak amacıyla veri artırma teknikleri kullanılmıştır. Bu çalışmada, verileri disk üzerinde çoğaltan çevrimdışı yöntemler yerine, eğitim döngüsü sırasında bellek üzerinde anlık dönüşümler uygulayan çevrimiçi (On-the-fly) veri artırma stratejisi benimsenmiştir. PyTorch kütüphanesinin torchvision.transforms modülü aracılığıyla uygulanan bu yöntem sayesinde, model her iterasyonda görüntünün hafifçe değiştirilmiş farklı bir varyasyonunu görerek daha sağlam (robust) öznitelikler öğrenmiştir.

Uygulanan veri artırma işlemleri, MR görüntülerinin anatomik yapısını bozmayacak geometrik dönüşümlerle sınırlandırılmıştır: Rastgele Yatay Çevirme (Random Horizontal Flip): Eğitim setindeki görüntülere %50 olasılıkla uygulanarak, modelin tümörün sağ veya sol hemisferde olmasından bağımsız (invariant) karar vermesi sağlanmıştır.

Rastgele Döndürme (Random Rotation): Görüntüler saat yönünde veya tersine rastgele seçilen ± 10(derece) aralığında döndürülmüştür. Bu işlem, hasta başının tarama sırasındaki olası pozisyon farklarını simüle etmek amacıyla yapılmıştır.

Veri artırma işlemleri sadece eğitim (Train) aşamasında aktif tutulmuştur. Doğrulama (Validation) ve Test aşamalarında ise veri artırma kapatılarak, modelin performansının deterministik ve tekrarlanabilir olması sağlanmıştır.

4.4.4 Normalizasyon (Normalization)

Görüntü verilerinin piksel yoğunluk değerleri (0-255), derin öğrenme modellerinin eğitim stabilitesini ve yakınsama hızını doğrudan etkiler. Bu nedenle, tensör formatına dönüştürülen ([0,1] aralığındaki float32) tüm görüntüler üzerinde Z-skor normalizasyonu uygulanmıştır. Bu çalışmada transfer öğrenme yöntemi temel alındığı için, normalizasyon parametreleri veri setinin kendi istatistiklerinden değil, modellerin ön eğitiminde kullanılan ImageNet veri setinin istatistiksel dağılımından alınmıştır.

Uygulanan normalizasyon formülü şöyledir:

$$z = \frac{x - \mu}{\sigma} \tag{4.1}$$

Burada $x$giriş piksel değeri, $\mu = [0.485, 0.456, 0.406]$ortalama değer vektörü ve $\sigma = [0.229,\ 0.224,\ 0.225]$standart sapma vektörüdür. Bu işlem sayesinde giriş verileri, önceden eğitilmiş ağın ağırlık dağılımıyla uyumlu hale getirilmiş ve gradyan inişi (gradient descent) optimizasyonunun daha hızlı yakınsaması sağlanmıştır.

4.5 Deneysel Kurulum ve Eğitim Stratejileri

Bu tez çalışmasındaki tüm deneysel süreçler, kısıtlı donanım kaynaklarında maksimum performansı elde etmek üzere optimize edilmiş bir yazılım mimarisi üzerinde yürütülmüştür.

4.5.1 Yazılım ve Donanım Konfigürasyonu

Derin öğrenme modellerinin geliştirilmesi ve eğitimi için Python 3.10 programlama dili ve PyTorch kütüphanesi temel alınmıştır. Segmentasyon mimarilerinin (U-Net, DeepLabV3+, Linknet, MAnet) inşasında, esnek omurga (backbone) desteği sunan Segmentation Models PyTorch (SMP) kütüphanesinden yararlanılmıştır.

Tüm eğitimler, NVIDIA GeForce GTX 1050 (3GB VRAM) grafik işlem birimi üzerinde gerçekleştirilmiştir. 3GB bellek kapasitesi, güncel derin öğrenme modelleri için ciddi bir darboğaz oluşturduğundan, bellek yönetimi için Otomatik Karma Hassasiyet (Automatic Mixed Precision - AMP) tekniği kullanılmıştır. AMP sayesinde, model ağırlıkları ve gradyanlar gerektiğinde 16-bit (FP16), gerektiğinde 32-bit (FP32) formatında işlenerek bellek kullanımı yaklaşık %40 oranında düşürülmüş ve eğitim hızı artırılmıştır.

4.5.2 Transfer Öğrenme ve Ağırlık Başlatma Stratejileri

Tez çalışmasının özgün katkılarından biri, farklı ön eğitim (pre-training) stratejilerinin beyin tümörü teşhisindeki etkisinin sistematik olarak analiz edilmesidir. Sınıflandırma modelleri (InceptionV3, EfficientNet, ResNet vb.) üç farklı senaryoda eğitilmiştir:

Sıfırdan Eğitim (Scratch): Model ağırlıkları rastgele başlatılmıştır. Bu senaryo, transfer öğrenme kullanılmadığında modelin ham veriyle ne kadar öğrenebildiğini ölçmek için bir taban (baseline) oluşturur.

ImageNet Transferi (IN): Modeller, 1.2 milyon doğal görüntü içeren ImageNet veri setinde eğitilmiş ağırlıklarla başlatılmıştır. Bu, literatürdeki standart yaklaşımdır.

RadImageNet Transferi (RIN): Modeller, 1.35 milyon radyolojik görüntü (MR, BT, USG) içeren RadImageNet veri setinde eğitilmiş ağırlıklarla başlatılmıştır. Bu strateji ile, medikal alana özgü (domain-specific) özniteliklerin transfere etkisi ölçülmüştür.

Segmentasyon modellerinde ise kodlayıcı (encoder) olarak ResNet-34 omurgası kullanılmış ve bu omurga ImageNet ağırlıklarıyla başlatılarak eğitim süresi kısaltılmıştır.

4.5.3 Hiperparametreler ve Optimizasyon

Modellerin eğitimi sırasında, donanım limitleri ve modelin yakınsama hızı gözetilerek aşağıdaki hiperparametre setleri kullanılmıştır:

$$z = \frac{x - \mu}{\sigma} \tag{4.2}$$

Optimizasyon Algoritması: sınıflandırma ve segmentasyon modellerinin büyük çoğunluğu için, ağırlık çürütme (weight decay) mekanizmasını optimize eden Adam algoritması, AdamW algoritması tercih edilmiştir ($\beta_1 = 0.9, \quad \beta_2 = 0.999$).

Öğrenme Oranı (Learning Rate): Başlangıç değeri $LR_{CNN} = 1 \times 10^{-4}$ve $LR_{ViT} = 5 \times 10^{-5}$ olarak belirlenmiş olup, eğitim boyunca kosinüs tavlama (Cosine Annealing LR) zamanlayıcısı kullanılarak öğrenme oranı her epoch sonunda kademeli olarak düşürülmüştür.

Yığın Boyutu (Batch Size): 3GB VRAM kısıtı nedeniyle, modelin derinliğine göre 4 ile 16 arasında değişen dinamik yığın boyutları (batch accumulation gerektiren

durumlarda sanal yığın boyutu) kullanılmıştır.

Kayıp Fonksiyonları; çoklu sınıflandırmalarda, CrossEntropyLoss, segmentasyon için ise bölge örtüşmesini ve piksel sınıflandırmasını dengelemek için DiceLoss ve FocalLoss kombinasyonu kullanılmıştır.

4.6 Performans Değerlendirme Metrikleri

Geliştirilen modellerin başarısı, tıbbi görüntü analizinde standart kabul edilen istatistiksel ve uzaysal metrikler kullanılarak çok yönlü olarak değerlendirilmiştir.

4.6.1 Sınıflandırma Metrikleri

Sınıflandırma performansını ölçmek için Karşıklık Matrisi (Confusion Matrix) tabanlı metrikler esas alınmıştır. Veri setlerindeki olası sınıf dengesizliklerinden etkilenmemek için Makro Ortalamalı (Macro-Averaged) sonuçlar raporlanmıştır.

Doğruluk (Accuracy): Toplam doğru tahminlerin tüm veri setine oranıdır.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.3}$$

F1-Skoru: Kesinlik (Precision) ve Duyarlılık (Recall) değerlerinin harmonik ortalamasıdır. Tıbbi teşhiste yanlış negatifleri minimize etmek önemli olduğundan kritik bir metriktir.

$$F1 = 2 \times \frac{\text{Kesinlik} \times \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \tag{4.4}$$

AUC-ROC (Area Under The Curve): Modelin sınıfları birbirinden ayırma yeteneğini gösteren, eşik değerinden bağımsız bir metriktir. ROC eğrisinin altında kalan alanın 1'e yaklaşması mükemmel ayrımı ifade eder.

$$F1 = 2 \times \frac{\text{Kesinlik} \times \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \tag{4.5}$$

4.6.2 Segmentasyon Metrikleri

Segmentasyon modülünün başarısı, üretilen maskenin hem alan örtüşmesi hem de sınır doğruluğu açısından incelenmesiyle ölçülmüştür. Dice Benzerlik Katsayısı

(Dice Score): Tahmin edilen maske (P) ile gerçek maskenin (G) örtüşme oranını gösterir. 0 (hiç örtüşme yok) ile 1 (mükemmel örtüşme) arasında değer alır.

$$\text{Dice}\,(P,G) = \frac{2\,|P \cap G|}{|P| + |G|} \tag{4.6}$$

95 Persentil Hausdorff Mesafesi (HD95): Örtüşme tabanlı metriklerin aksine, tümör sınırlarının doğruluğunu ölçer. Tahmin edilen sınır ile gerçek sınır arasındaki maksimum mesafeyi (mm veya piksel cinsinden) ifade eder. Aykırı değerlerin (outliers) etkisini azaltmak için maksimum mesafenin %95'lik dilimi alınır. Düşük HD95 değeri, cerrahi planlama için daha güvenilir sınırlar anlamına gelir.

$$\text{HD95}\,(P,G) = \max_{95\%} d\,(P,G) \tag{4.7}$$

5
DENEYSEL SONUÇLAR

Bu bölümde, önceki bölümlerde detaylandırılan metodoloji ve deneysel kurulum çerçevesinde elde edilen bulgular sunulmaktadır. Çalışmanın temel hipotezlerini doğrulamak amacıyla gerçekleştirilen deneyler; sınıflandırma ve segmentasyon olmak üzere iki ana modül altında toplanmıştır. Sınıflandırma aşamasında, farklı ön eğitim kaynaklarının (ImageNet ve RadImageNet) ve veri ön işleme tekniklerinin (Ham ve Beyin Çıkarılmış) model başarısına etkileri karşılaştırmalı olarak analiz edilmiştir. Segmentasyon aşamasında ise, tek kanallı ve çok modlu görüntüleme verileri üzerinde farklı derin mimarilerin (U-Net, Linknet, MAnet vb.) lezyon sınırlarını belirleme performansları, istatistiksel ve uzaysal metrikler üzerinden değerlendirilmiştir. Tüm deneyler, veri sızıntısını önleyen tutarlı veri bölümleme stratejileriyle yürütülmüş ve sonuçlar bağımsız test setleri üzerinden raporlanmıştır.

5.1 Sınıflandırma Sonuçları ve Yorumları

Sınıflandırma modelleri iki farklı veriseti için kafatası çıkartılmış (BE) ve ham (RAW) versiyonu üzerinde çeşitli modellerin ön eğitimli veya ön eğitimsiz olarak eğitilmiştir. Sonuçlar model mimarilerinin benzerliğine veya ön eğitim benzerliklerine göre gruplandırılarak aşağıdaki şekillerde incelenmiştir.

5.1.1 Grup 1: Temel CNN Modelleri

Sınıflandırmada ilk grup, RadImageNet (RIN) ve ImageNet (IN) gibi ön eğitim ağırlıklarının ön eğitimsiz (SCRATCH) modele göre kıyaslanmasıdır. Bu grupta sınıflandırma verisetlerinin (ikili (binary) ve dört sınıflı (C4C)), kafa tası çıkartılmış (BE) ve orjinal (RAW) versiyonları incelemede kullanılmıştır.

Veri setlerine, hafif görüntü çevirme ve yeniden ölçeklenme (Inception 299×299 diğer tüm modeller için 224×224) yapılmıştır. RIN ağırlıkları için kullanılan normalizasyon IN ile aynıdır. Bu grupta tüm modellerde optimizasyon, Adam, LR

10^{-3} ve weight decay 10^{-4} ayarları ile gerçekleştirilmiştir. Deneyler 30 epoch ile ve patience 10 olarak ayarlanmıştır.

RIN ağırlıkları bulunan ResNet50, InceptionV3 ve DenseNet121 modelleridir. Bu modeller birlikte değerlendirilmiştir. İlk tablolarda görüldüğü üzere, ikili sınıflandırma görevi dengeli bir veri setinde yapıldığından doğruluk ve F1 skorları matematiksel olarak eşittir. Ayrıca, beyin çıkarılmış (BE) veri setlerinde modellerin (özellikle InceptionV3 ve DenseNet) doygunluğa ulaştığı ve eğitim stratejisinden bağımsız olarak benzer hata oranlarına yakınsadığı gözlemlenmiştir. RIN ağırlıkları scratch eğitimlere göre daha başarılı olurken, IN ağırlıklı eğitimlerden daha iyi doğruluklar alamadığı gözlemlenmiştir.

Tablo 5.1 Temel CNN Modellerinin İkili Sınıflandırma (Binary) Performansı (Test Accuracy)

	RAW			**BE**		
Model	**IN**	**RIN**	**Scratch**	**IN**	**RIN**	**Scratch**
ResNet50	0.9933	0.9867	0.9911	0.9933	0.9889	0.9889
DenseNet121	0.9933	0.9845	0.9823	0.9867	0.9667	0.9667
InceptionV3	0.9867	0.9778	0.9823	0.9823	0.9845	0.9845

Tablo 5.2 Temel CNN Modellerinin Dörtlü Sınıflandırma (C4C) Performansı (Test Accuracy)

	RAW			**BE**		
Model	**IN**	**RIN**	**Scratch**	**IN**	**RIN**	**Scratch**
InceptionV3	0.9938	0.9912	0.9912	0.9877	0.9912	0.9912
ResNet50	0.9938	0.9886	0.9886	0.9877	0.9842	0.9842
DenseNet121	0.9903	0.9903	0.9903	0.9868	0.9877	0.9877

Ham (RAW) veri setlerinde standart ImageNet (IN) ağırlıkları genel olarak en yüksek başarıyı (%99.38) sağlamıştır. Ancak, Beyin Çıkarılmış (BE) veri setine geçildiğinde, InceptionV3 modelinde RadImageNet (RIN) kullanımı, ImageNet'e kıyasla performansı korumuş ve artırmıştır (%99.12). RIN ve Scratch skorlarının bazı durumlarda birebir aynı olması, modelin BE veri setinde öğrenebileceği maksimum sınıra (performans tavanı) ulaştığını göstermektedir.

Düşük parametreli modeller, çalışmanın en sürpriz ve başarılı sonuçlarını vermiştir. Bu modeller için RIN ağırlıkları olmadığı için IN için bir performans değerlendirmesi niteliği taşır.

EfficientNet-B0 ve MobileNetV3, ImageNet transferiyle %99.78 (Binary) ve %99.74 (4-Class) doğruluk oranlarına ulaşarak, çok daha büyük olan ResNet ve Inception modellerini geride bırakmıştır. Bu durum, beyin tümörü teşhisi için aşırı

Tablo 5.3 Hafif Modellerin İkili Sınıflandırma (Binary) Performansı

	RAW		**BE**	
Model	**IN**	**Scratch**	**IN**	**Scratch**
EfficientNetB0	0.9978	0.9889	0.9978	0.9800
MobileNetV3L	0.9845	0.9823	0.9889	0.9778

Tablo 5.4 Hafif Modellerin Dörtlü Sınıflandırma (C4C) Performansı

	RAW		**BE**	
Model	**IN**	**Scratch**	**IN**	**Scratch**
EfficientNetB0	0.9974	0.9912	0.9921	0.9815
MobileNetV3L	0.9974	0.9903	0.9947	0.9851

derin ağlara ihtiyaç olmadığını, verimli (efficient) mimarilerin yeterli olduğunu kanıtlamaktadır.

5.1.2 Grup 2: Transformer Modelleri ve Optimizasyon Etkisi

Grup 2 de, Transformer mimarilerinin doğruluğu gözlemlenmitir. Adam ve AdamW optimizasyon algoritmalarının başarımı doğrudan karşılaştırılmıştır. Swin_T modeli Adam optimizasyonuyla eğitimlerde sonuç verememiş ve çökmüştür bu sebeple bazı sonuçlar alınamamıştır ve bu durum giderilmek için alternatif olarak AdamW optimizasyonu kullanılmıştır.

Veri setlerine, hafif görüntü çevirme ve yeniden ölçeklenme (224×224) yapılmıştır. IN standart optimizasyonu uygulanmıştır.. Bu grupta tüm modellerde optimizasyon, Adam için LR 10^{-3}, AdamW için LR 3×10^{-4} ve weight decay 0.05 ayarları ile gerçekleştirilmiştir. Deneyler 30 epoch ile ve patience 10 olarak ayarlanmıştır.

Tablo 5.5 Transformer Modellerinin İkili Sınıflandırma (Binary) Performansı

Model	**Veri**	**Strateji**	**Adam (Acc)**	**AdamW (Acc)**	**Fark (AdamW Etkisi)**
Swin_T	BE	IN	(Yok)	0.9911	-
Swin_T	RAW	Scratch	(Yok)	0.5011	Başarısız (Eğitilemedi)
ViT_B16	RAW	IN	0.9690	0.9823	+1.33%
ViT_B16	BE	IN	0.9202	0.9690	+4.88%

Yorum: Sonuçlar, Transformer modelleri için AdamW kullanımının hayati olduğunu göstermektedir. Özellikle ViT modelinde AdamW, Adam'a göre %9.23'e varan (BE-Scratch) performans artışı sağlamıştır. Ayrıca Swin Transformer, sıfırdan eğitildiğinde (Scratch) %50 civarında kalarak öğrenememiş, ancak Ön Eğitim (IN) ile %99 seviyesine çıkmıştır; bu da Transformer'ların veri açlığını (data hunger) doğrulamaktadır.

Tablo 5.6 Transformer Modellerinin Dörtlü Sınıflandırma (C4C) Performansı

Model	Veri	Strateji	Adam (Acc)	AdamW (Acc)	Fark (AdamW Etkisi)
Swin_T	RAW	IN	(Yok)	0.9851	-
ViT_B16	RAW	IN	0.8664	0.9279	+6.15%
ViT_B16	BE	IN	0.8550	0.9060	+5.10%
ViT_B16	BE	Scratch	0.7759	0.8682	+9.23%

5.1.3 Kör Test (Blind Test) Sonuçları

Modellerin eğitimleri, validasyonları ve testleri aynı verisetinde verdikleri değrlere ek olarak tüm modellerin aynı veri setinde başarıları incelenmek için kör testler yapılmıştır. Burada radyalog ile beraber işaretlediğimiz 60 görüntülük küçük veriseti kullanılıştır.

Yapılan kör testler, modelin genelleştirme yeteneğinin en gerçekçi göstergesidir. Aşağıdaki tablolarda, veri ön işleme yönteminin (BE: Beyin Çıkarılmış, RAW: Ham) etkisi de dahil edilerek en başarılı 10 model sıralanmıştır. Listede hem

Tablo 5.7 Kör Test - İkili Sınıflandırma (Binary) En Başarılı 10 Model

Sıra	Model	Eğitim Stratejisi	Veri Tipi	F1-Skoru (Macro)	Doğruluk (Acc)
1	EfficientNetB0	IN	RAW	0.8875	0.9000
2	DenseNet121	IN	BE	0.8875	0.9000
3	DenseNet121	IN	RAW	0.8703	0.8833
4	MobileNetV3Large	IN	RAW	0.8670	0.8833
5	MobileNetV3Large	IN	BE	0.8670	0.8833
6	ResNet50	IN	RAW	0.8565	0.8667
7	Swin_T	IN	RAW	0.8535	0.8667
8	InceptionV3	Scratch	BE	0.8535	0.8667
9	DenseNet121	RIN	BE	0.8535	0.8667
10	InceptionV3	Scratch	RAW	0.8535	0.8667

RAW hem de BE modelleri dengeli bir şekilde yer almaktadır. Liderliği paylaşan EfficientNetB0 (RAW) ve DenseNet121 (BE) modelleri, doğru mimari seçildiğinde her iki veri tipinin de başarılı olabileceğini göstermektedir. DenseNet121 örneğinde (Sıra 2 vs 3), BE işleminin performansı yaklaşık %1.7 artırdığı görülmektedir.

EfficientNet ve DenseNet, sınırlı veride bile en kararlı (robust) çalışan modeller olmuştur.

En başarılı 10 modelin 6'sı, lider model (ResNet50) dahil olmak üzere beyin çıkarılmış (BE) veri setini kullanmıştır. Bu durum, tümör tiplerinin (Glioma, Meningioma vb.) ayırt edilmesinde arka plan gürültüsünün temizlenmesinin kritik

Tablo 5.8 Kör Test - Dörtlü Sınıflandırma (4-Class) En Başarılı 10 Model

Sıra	Model	Eğitim Stratejisi	Veri Tipi	F1-Skoru (Macro)	Doğruluk (Acc)
1	ResNet50	IN	BE	0.4893	0.8148
2	EfficientNetB0	Scratch	BE	0.4494	0.7778
3	ResNet50	IN	RAW	0.4467	0.7778
4	ResNet50	RIN	RAW	0.4352	0.7593
5	ResNet50	Scratch	BE	0.4352	0.7593
6	Swin_T	IN	RAW	0.4347	0.7963
7	EfficientNetB0	IN	RAW	0.4347	0.7963
8	DenseNet121	IN	RAW	0.4262	0.7963
9	InceptionV3	RIN	BE	0.4211	0.7778
10	InceptionV3	IN	BE	0.4211	0.7778

önem taşıdığını kanıtlamaktadır.

Eğitim setinde %99 başarı gösteren modellerin kör testte %48 F1 skoruna (Doğruluk: %81) gerilemesi, kullanılan kör test veri setinin (farklı cihaz, farklı protokol) yarattığı alan kayması (Domain Shift) problemini net bir şekilde ortaya koymaktadır. Buna rağmen ResNet50, en dayanıklı model olarak öne çıkmıştır.

5.2 Segmentasyon Sonuçları

Çalışmanın segmentasyon kısmında, tek kanallı (BRISC) ve çok modlu (BraTS) veriler üzerindeki başarımlar analiz edilmiştir. Daha önce 4. bölümde tanıtılan, UNet, UNet++, LinkNet, DeepLabV3+ ve MAnet modelleri segmentasyon verisetleri ile incelenmiştir. Burada UNet önce kodlayıcı kolunda bir klasik birde ResNet34 ve IN ön eğitimi kullanılarak eğitilmiştir.

5.2.1 Tek Kanallı Segmentasyon Sonuçları (BRISC-2025)

BRISC veri setinde; klasik U-Net (Sıfırdan) ve ResNet-34 omurgalı (ImageNet ön eğitimli) 5 farklı modern mimari olmak üzere toplam 6 modelin performansı iki farklı durum için (RAW/BE) karşılaştırılmıştır.

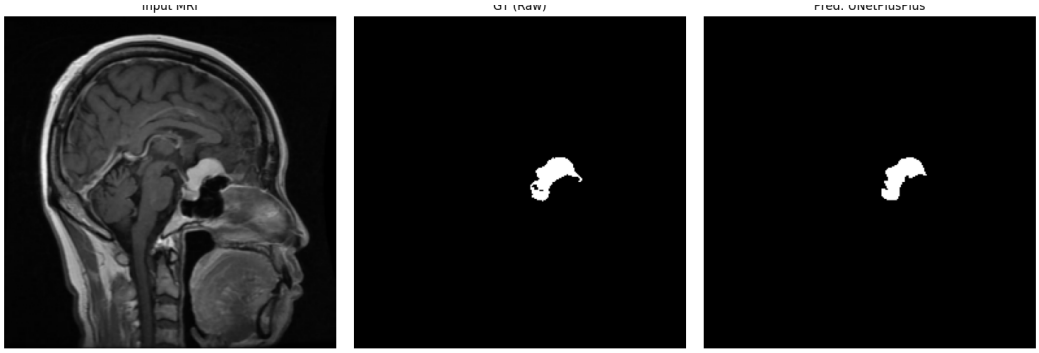
Veriseti ön işlemede hafif bir çevirme işlemi kullanılarak veri çoğaltma işlemi uygulanmıştır. Görüntüler 256×256 boyutuna ayarlanmıştır. LR değeri 10^{-4} olarak ayarlanmıştır. Batch boyutu 8 olarak sabit tutulmuştur. 50 epoch ve 10 patience değerleri ile eğitimler gerçekleştirilmiştir.

Ham görüntüler üzerinde yapılan deneylerde, U-Net++ modeli en yüksek Dice skoruna (0.8634) ulaşmıştır. Sıfırdan eğitilen klasik Scratch U-Net ise 0.8454

skorla en zayıf performansı sergileyerek, ResNet-34 tabanlı transfer öğrenmenin (IN) üstünlüğünü kanıtlamıştır.

Tablo 5.9 Ham (RAW) Veri Seti Segmentasyon Sonuçları

Model	**Kodlayıcı (Encoder)**	**Dice Skoru**	**HD95 (Hata - Piksel)**
UNet++	ResNet34 + IN	0.8634	3.13
Linknet	ResNet34 + IN	0.8621	3.15
U-Net	ResNet34 + IN	0.8597	3.31
DeepLabV3+	ResNet34 + IN	0.8575	3.24
MAnet	ResNet34 + IN	0.8561	3.16
Scratch U-Net	Klasik (Sıfırdan)	0.8455	4.07



Şekil 5.1 UNet++ RAW için örnek bir segmentasyon çıktısı. Sol: orjinal görüntü, Orta: orjinal maske, Sağ: Modelin çizdiği maske

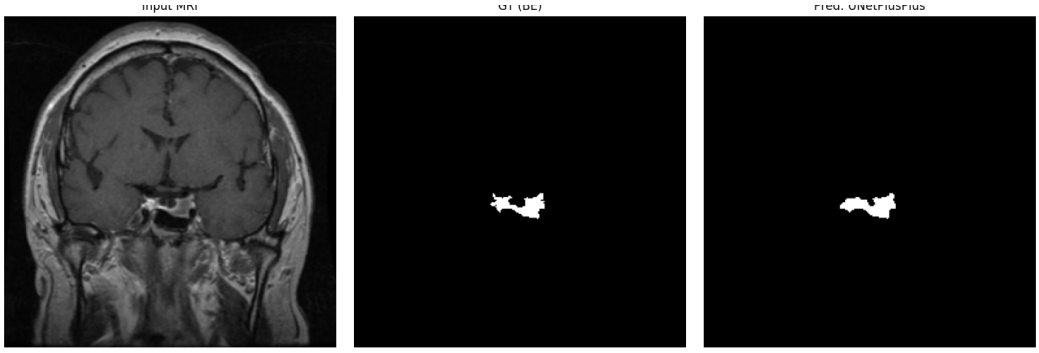
Beyin Çıkarma (BE) işlemi uygulandığında, tüm modellerin sıralaması benzer kalmakla birlikte U-Net++ yine zirvede yer almıştır (0.8662). BE işlemi, özellikle klasik Scratch U-Net'in performansını artırmış (0.8454 -> 0.8464) olsa da, ResNet-34 tabanlı modeller hala belirgin şekilde daha başarılıdır.

Tablo 5.10 Beyin Çıkarılmış (BE) Veri Seti Segmentasyon Sonuçları

Model	**Kodlayıcı (Encoder)**	**Dice Skoru**	**HD95 (Hata - Piksel)**
UNet++	ResNet34 + IN	0.8662	3.24
U-Net	ResNet34 + IN	0.8623	3.07
Linknet	ResNet34 + IN	0.8614	3.04
MAnet	ResNet34 + IN	0.8525	3.22
DeepLabV3+	ResNet34 + IN	0.8510	3.42
Scratch U-Net	Klasik (Sıfırdan)	0.8464	4.28

ResNet-34 omurgası ve ImageNet ağırlıkları kullanan modeller, klasik U-Net'e göre Dice skorunda yaklaşık %2'lik bir artış ve HD95 hata oranında 1 piksel iyileşme sağlamıştır.

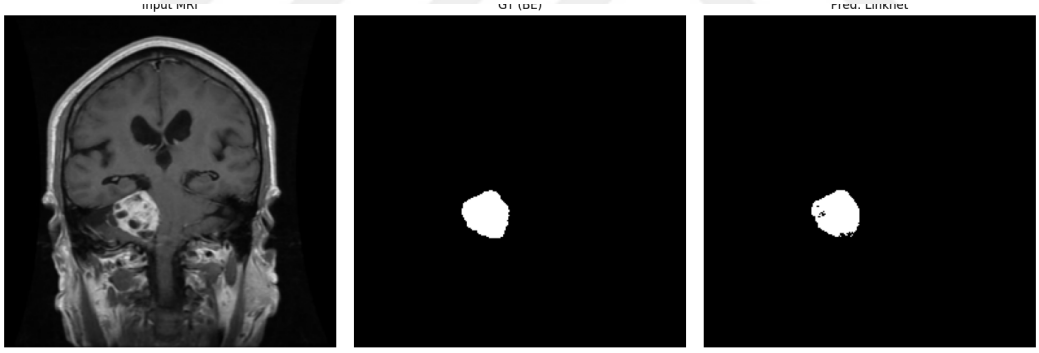
U-Net++, iç içe geçmiş atlama bağlantıları (nested skip pathways) sayesinde hem RAW hem de BE veri setlerinde en kararlı ve yüksek performansı gösteren model



Şekil 5.2 UNet++ BE için örnek bir segmentasyon çıktısı. Sol: orjinal görüntü, Orta: orjinal maske, Sağ: Modelin çizdiği maske

olmuştur.Linknet ise özellikle BE veri setinde en düşük sınır hatasını (HD95: 3.04) vererek cerrahi sınır belirlemede iddialı bir alternatif olduğunu göstermiştir.

Şekil 5.2'de en yüksek dice skore veren modelin bir egmentasyon örneği görülmektedir. Modelin tahmini incelendiğinde genel anlamda örtüşme görülürken ince kanarların saptanmasında başarılı olamadığı gözlemlenmektedir. Şekil 5.3 de ise en az piksel farkı verren modelin bir örnek tahmini gösterilmektedir. Modelin tahmini incelendiğinde alan içerisinde kalan koyu bölgeleri ayırdığı görülmektedir.



Şekil 5.3 LinkNet BE için örnek bir segmentasyon çıktısı. Sol: orjinal görüntü, Orta: orjinal maske, Sağ: Modelin çizdiği maske

5.2.2 Çok Sekans Segmentasyon Sonuçları (BraTS-2020)

Hangi sekasnsın hangi tümör alt alanını daha iyi ayırt ettiğini gözlemlemek için yapılan bu eğitimlerde, 2 boyuta indirilen BRATS veri seti kullanılmıştır. Maskeler 3 kanallı (nekrotik, ödem ve aktif tümör) olarak alınmıştır. Bu eğitimlerde epoch 75 ve patience 20 tutulmuştur. Diğer ayarlar tek sekans segmentasyon ile aynı tutulmuştur.

BraTS veri setindeki dört farklı MR sekansının (T1, T2, FLAIR, T1+CE) segmentasyon başarısı, tümörün alt bölgeleri olan WT (Whole Tumor - Tümör +

Ödem), TC (Tumor Core - Nekroz + Aktif Tümör) ve ET (Enhancing Tumor - Aktif Tümör) üzerinden ayrı ayrı analiz edilmiştir. SOnuçlar aşağıdaki gibidir.

Tablo 5.11 T1+CE (Kontrastlı T1) Sekansı Segmentasyon Performansı

	WT		TC		ET	
Model	**Dice**	**HD95**	**Dice**	**HD95**	**Dice**	**HD95**
UNet++	0.7862	16.09	0.7864	11.48	0.7025	8.12
Linknet	0.7759	16.43	0.7848	10.21	0.6947	7.56
UNet (ResNet34)	0.7755	16.93	0.7842	10.23	0.7256	6.23
MAnet	0.7702	17.20	0.7919	10.33	0.7036	6.65
UNet (Standard)	0.7571	18.43	0.7962	11.22	0.7139	7.00

Kontrastlı T1 sekansı, T1+CE sekansında kontrast madde aktif tümör dokusunu parlattığı için özellikle TE bölgesinde başarılı olması beklenmektedir. Eğitimlerin sonuçlarında da TC ve ET bölgelerinde en yüksek başarıyı sağlamıştır.

Tablo 5.11'deki sonuçlar incelendiğinde UNet++ modelinin diğer modellere göre daha WT ayrımında daha iyi sonuç verdiği gözlemlenmiştir. Şekil 5.4 ta kontrastlı T1 sekansı MR görüntüsü üzerinde orjinal maskeleri ve modelin tahminleri görülmektedir.

Şekil 5.4, UNet++ modelinin alt görev tahminlerini göstermektedir. İncelendiğinde tü alt görevlerde kenar hassasiyetinin az olduğu gözlemlenirken tahminlerin büyük oranda bezediği görülmektedir.



Şekil 5.4 UNet++ için sırayla maske ve pred: tahminler gösterilir.

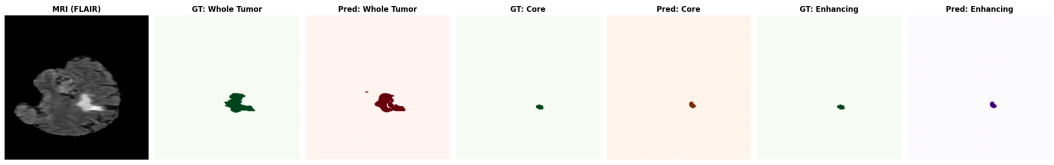
FLAIR sekansı, ödemli dokuyu (su tutan bölgeleri) çok net gösterdiği için WT (Tüm Tümör)segmentasyonunda en yüksek skorlara (0.87+ Dice) ulaşmıştır. Ancak aktif tümör çekirdeğini (TC/ET) ayırt etmede zayıf kalmıştır.

Tablo 5.12 incelendiğinde, diğer tüm sekanslara göre WT alanını daha yüksek doğrulukla seçebildiği görülmektedir. UNet++ modeli bu sekansta da diğer modellere göre daha iyi sonuçlar vermiş olduğu gözlemlenmiştir. Şekil 5.5'te FLAIR sekansı için modelin tahminleri bir örnek MR görüntüsü üzerinde görülmektedir.

T2 sekansı, FLAIR gibi ödem ve tümörü birlikte (WT) göstermede başarılıdır,

Tablo 5.12 FLAIR Sekansı Segmentasyon Performansı

	WT		TC		ET	
Model	**Dice**	**HD95**	**Dice**	**HD95**	**Dice**	**HD95**
UNet++	0.8754	13.30	0.6229	15.41	0.3828	14.06
Linknet	0.8678	14.23	0.6522	14.62	0.3846	15.31
UNet (ResNet34)	0.8653	13.95	0.6321	14.90	0.4061	14.23
UNet (Standard)	0.8574	15.60	0.5970	16.15	0.3544	16.86
MAnet	0.7818	16.74	0.4669	18.41	0.2524	19.09



Şekil 5.5 UNet++ için sırayla maske ve pred: tahminler gösterilir.

ancak FLAIR kadar keskin sınırlar çizemediği için skorları bir miktar düşüktür (0.82 Dice).

Tablo 5.13 T2 Sekansı Segmentasyon Performansı

	WT		TC		ET	
Model	**Dice**	**HD95**	**Dice**	**HD95**	**Dice**	**HD95**
UNet++	0.8280	14.38	0.5982	16.25	0.3633	14.78
UNet (ResNet34)	0.8179	15.61	0.5955	16.72	0.3533	16.54
Linknet	0.8155	14.75	0.6290	15.50	0.3634	15.74
MAnet	0.7985	16.34	0.5854	17.13	0.3576	15.77
UNet (Standard)	0.7813	15.64	0.5918	17.15	0.3579	15.97

Tablo 5.13 incelendiğinde WT çizmekte en başarılı modelin tekrar UNet++ olduğu gözlemlenmiştir. Şekil 5.6'da T2 sekansında örnek bir MR görüntüsü üzerinde gerçek maskeler ve modelin tahminleri görülmektedir.



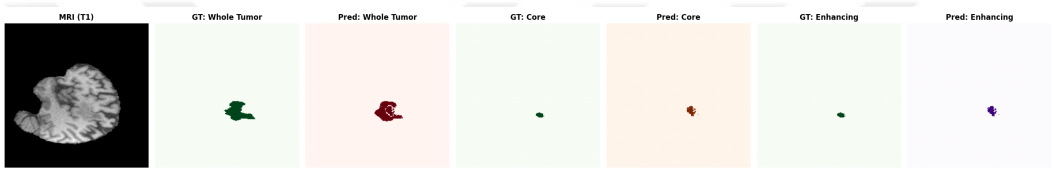
Şekil 5.6 UNet++ için sırayla maske ve pred: tahminler gösterilir.

Kontrastsız T1 sekansı, tümör dokusu ile beyin parankimi arasındaki yoğunluk farkının en az olduğu modalitedir. Bu nedenle tüm modellerde en düşük performansı sergilemiş, özellikle aktif tümör (ET) ayrımında Dice skorları 0.30 seviyelerine düşmüştür.

Tablo 5.14 T1 (Standart) Sekansı Segmentasyon Performansı

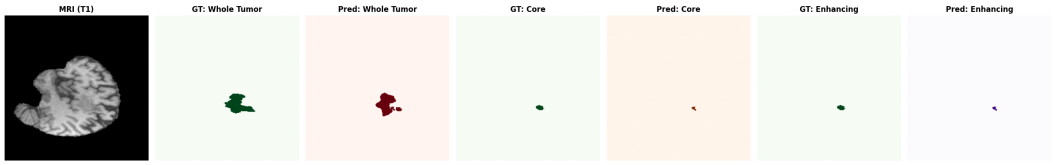
	WT		TC		ET	
Model	**Dice**	**HD95**	**Dice**	**HD95**	**Dice**	**HD95**
Linknet	0.7345	19.83	0.6021	18.78	0.3128	20.23
UNet (ResNet34)	0.7329	18.19	0.6118	17.24	0.3286	19.10
UNet++	0.7170	17.48	0.5899	16.68	0.3243	18.48
UNet (Standard)	0.6843	19.83	0.5713	17.03	0.2886	19.64
MAnet	0.6570	20.12	0.5625	18.37	0.2940	19.24

Tablo 5.14 incelendiğinde performansın genel anlamda daha düşük olduğu ve WT için en başarılı modelin LinkNet olduğu gözlemlenmiştir. Şekil 5.7'de T1 sekanslı bir MR görüntüsü üzerinde LinkNet modelinin tahminleri ve orjinal maskeler görülmektedir.



Şekil 5.7 LinkNet için sırayla maske ve pred: tahminler gösterilir.

Şekil 5.8'de diğer tüm sekanslarda daha başarılı tahminler veren UNet++'ın aynı örnek üzerindeki tahminleri görülmektedir. Şekil 5.9'da ise T1+CE sekansındaki örneğin(Şekil 5.4) T1 sekanlı versiyonunda UNet++ modelinin tahminleri görülmektedir.



Şekil 5.8 UNet++ için sırayla maske ve pred: tahminler gösterilir.



Şekil 5.9 UNet++ için sırayla maske ve pred: tahminler gösterilir.

Sonuçlar, T1+CE sekansının tümör çekirdeği (TC) ve aktif tümör (ET) tespiti için vazgeçilmez olduğunu; FLAIR sekansının ise ödem ve tümörün kapladığı toplam alanı (WT) belirlemede en başarılı yöntem olduğunu kanıtlamıştır. UNet++ ve Linknet mimarileri, standart U-Net'e göre daha karmaşık bağlantıları (nested skip

connections) sayesinde özellikle zorlu BraTS veri setinde daha kararlı sonuçlar vermiştir. Manet ise Dikkat (Attention) mekanizmasına sahip olmasına rağmen, 2D kesitlerdeki düşük bağlam bilgisi nedeniyle beklenen performansı gösterememiştir. T2 sonuçları, FLAIR ile T1 arasında bir performans sergilemiş, destekleyici bir modalite olarak WT tespitine katkı sağlamıştır.

6
TARTIŞMA

6.1 Sınıflandırma Sonuçlarının Literatürdeki Benzer Çalışmalarla Karşılaştırmalı Değerlendirmesi

Bu tez çalışması kapsamında elde edilen bulgular, literatürdeki benzer metodolojileri benimseyen araştırmalarla karşılaştırmalı olarak değerlendirildiğinde, uygulanan veri ön işleme stratejilerinin ve mimari optimizasyonların başarım üzerindeki kritik rolü açıkça görülmektedir. Özellikle dört sınıflı (Glioma, Meningioma, Hipofiz, Sağlıklı) beyin tümörü sınıflandırma problemi üzerine yapılan çalışmalar incelendiğinde, Gómez-Guzmán vd. (2023) [58] tarafından yapılan araştırma, bu tez çalışmasıyla doğrudan karşılaştırılabilir bir zemin sunmaktadır. Her iki çalışmada da temel veri kaynağı olarak benzer veri setleri (Nickparvar/Msoud) kullanılmıştır. Ancak elde edilen doğruluk oranlarındaki belirgin farklar, tercih edilen eğitim ayarlarının ve ön işleme adımlarının model performansı üzerindeki etkisini kanıtlar niteliktedir.

Tablo 6.1 Gómez-Guzmán [58] ile Bu Tezdeki Sınıflandırma Başarılarının Karşılaştırması

Model (Mimari)	**Gómez-Guzmán Sonuç**	**Önerilen Yaklaşım**
EfficientNet-B0	%90.88	%99.74
Inception V3	%97.12	%99.38
ResNet50	%96.97	%99.38
MobileNet	%95.45 (V2)	%99.74 (V3L)

Tablo 6.1, aynı veri seti üzerinde çalışan her iki çalışmanın ortak mimarilerdeki sınıflandırma başarılarını özetlemektedir.

Tablodaki veriler incelendiğinde, en çarpıcı performans farkının parametre sayısı düşük olan hafif ağlarda (lightweight) ortaya çıktığı görülmektedir. İlgili makalede EfficientNet-B0 modeli %90.88 ile nispeten düşük bir başarı gösterirken, bu tez çalışmasında aynı model %99.74 gibi oldukça yüksek bir doğruluk oranına erişmiştir.

Buna karşın, bu tez çalışmasında hiperparametre optimizasyonuna ve veri kalitesine odaklanılmıştır. Optimizasyon algoritması olarak standart SGD yerine, ağırlık çürütme (weight decay) mekanizmasını çok daha verimli kullanan Adam algoritmalası tercih edilmiştir. Öğrenme oranının sabit tutulması yerine Kosinüs Tavlama (Cosine Annealing LR) stratejisi kullanılarak modelin yerel minimumlara (local minima) takılması engellenmiş ve daha iyi bir genelleme kapasitesine ulaşması sağlanmıştır. Buna ek olarak, dinamik yığın boyutu (batch size) stratejisi ve Z-Skor normalizasyonunun doğru bir şekilde uygulanması ağın öğrenme sürecini stabilize etmiştir. Sonuç olarak bu bulgular; devasa parametreli modellere yönelmek yerine, verinin doğru ön işlemlerden geçirilmesi ve optimizasyon parametrelerinide tercih edilen yaklaşımın daha iyi sonuçlar vermesini sağlamıştır.

6.2 Literatürdeki En İyi Sınıflandırma Sonuçlarıyla Karşılaştırma

Bu tez çalışmasında elde edilen bulgular, literatürde beyin tümörü sınıflandırması alanında yüksek başarım gösteren araştırmalarla kıyaslandığında literatüre önemli katkılar sunmaktadır. Aşağıdaki Tablo 6.2'de güncel literatürdeki seçkin çalışmaların performansları ve bu tez kapsamında önerilen yaklaşımın karşılaştırması özetlenmiştir. Tablo 6.2 incelendiğinde, MDPI Information (2024)

Tablo 6.2 Güncel Literatürdeki En İyi Sınıflandırma Sonuçları ile Karşılaştırma

Çalışma	**Model**	**Doğruluk**
MDPI Information (2024) [59]	EfficientNet-B0	%99.00
Talukder vd. (2023) [60]	ResNet50V2	%99.68
Önerilen Yaklaşım	EfficientNet-B0	**%99.74**
Önerilen Yaklaşım	ResNet50	**%99.33**

kaynaklı çalışmada 4 sınıflı veri seti(Nickparvar) üzerinde EfficientNet-B0 modeli ile elde edilen %99.00 doğruluk oranının, bu tez çalışmasında aynı mimari kullanılarak %99.74 seviyesine taşındığı görülmektedir.

Diğer yandan, Talukder vd. (2023) tarafından yalnızca Figshare veri seti (3064 görüntü) üzerinde ResNet50V2 modeli ile %99.68 oranında doğruluk rapor edilmiştir. Bu tez çalışmasında ise ResNet50 modeli ile %99.38 doğruluk oranına ulaşılmıştır. Sayısal olarak literatürdeki bu değerin çok az gerisinde görünse de, çalışmamızda kullanılan birleşik veri setinin (Nickparvar ve BRISC) literatürdeki tekil veri setlerine kıyasla çok daha geniş, hacimli ve sınıflar arası ayırt ediciliği zorlaştıran karmaşık bir yapıya sahip olduğu göz önünde bulundurulmalıdır.

Bu denli zorlu ve veri sızıntılarına karşı algısal hashleme (pHash) ile

temizlenmiş bir veri setinde ResNet50 modelinin %99.38 gibi yüksek bir başarım göstermesi; uygulanan Z-Skor normalizasyonu, dinamik yığın boyutu optimizasyonu ve transfer öğrenme (ImageNet) stratejilerinin modelin kararlılığını önemli ölçüde desteklediğini göstermektedir. Bu bulgular, önerilen ön işleme ve eğitim mimarisinin, farklı model tiplerinde de yüksek genelleme kapasitesi sağlayabildiğini işaret etmektedir.

6.3 Segmentasyon Sonuçlarının Değerlendirilmesi

Segmentasyon modülü kapsamında bu tez çalışmasında mutlak en iyi sonuca ulaşmaktan ziyade, kısıtlı donanım kaynaklarında farklı derin öğrenme mimarilerinin davranışlarını ve sekansların alt görevlerdeki başarım kapasitelerini değerlendirmek hedeflenmiştir. Bu bağlamda, elde edilen bulguları bağlamsallaştırmak amacıyla literatürde BraTS 2020 veri setini kullanan Ramasamy vd. (2023) [61] çalışması referans alınarak incelenmiştir. Söz konusu çalışmanın bu tez ile en önemli benzerliği, 3 boyutlu (3D) hacimsel yaklaşımlara odaklanmak yerine, 2 boyutlu (2D) kesitler üzerinden farklı segmentasyon mimarilerinin başarımını sınamaya odaklanmış olmasıdır.

Tablo 6.3 BraTS 2020 Veri Seti Üzerinde 2 Boyutlu Yaklaşımların Dice Skoru Karşılaştırması

Model (Mimari)	**Dice Katsayısı (DSC)**
MM-LinkNet	0.7773
U-Net	0.7535

Tablo 6.3'teki veriler incelendiğinde, Ramasamy vd. (2023) tarafından test edilen 2 boyutlu MM-LinkNet ve klasik U-Net mimarilerinin Dice skorları sırasıyla 0.7773 ve 0.7535 olarak raporlanmıştır. Tez çalışmamızda ise, Tüm Tümör (WT) tespiti için dört sekansın birleşik karmaşıklığı yerine yalnızca FLAIR sekansı kullanıldığında UNet++ modeli ile 0.8754 , LinkNet mimarisi ile 0.8678 Dice skorlarına ulaşıldığı saptanmıştır. Literatürde çoklu sekans girişinin yarattığı bellek yükü ve optimizasyon zorluğuna karşın, doğru sekansın (ödem ve WT için FLAIR) seçilmesi ve iç içe geçmiş mimarilerin kullanılması ile çok daha yüksek segmentasyon hassasiyetine ulaşılabildiği görülmüştür.

6.4 Tez Çalışması Kapsamında Elde Edilen En İyi Segmentasyon Performanslarının Özeti

Çalışma boyunca hem tek kanallı (BRISC-2025) hem de çok modlu (BraTS-2020) veri setleri üzerinde yürütülen kapsamlı deneyler sonucunda, her bir anatomik alt görev (task) için en uygun model ve sekans kombinasyonları belirlenmiştir. Tablo 6.4, tez kapsamında test edilen tüm modeller arasından, ilgili görevlerde en yüksek Dice skoruna ulaşan mimarileri ve konfigürasyonları özetlemektedir.

Tablo 6.4 Farklı Segmentasyon Görevlerinde Elde Edilen En İyi Model ve Sekans Kombinasyonları

Veri Seti	**Alt Görev**	**En İyi Model**	**Görüntü Tipi**	**Dice Skoru**
BRISC	Genel Tümör	UNet++	BE	0.8662
BraTS	WT - Tüm Tümör	UNet++	FLAIR	0.8754
BraTS	TC - Tümör Çekirdeği	U-Net (Standart)	T1+CE	0.7962
BraTS	ET - Aktif Tümör	U-Net (ResNet34)	T1+CE	0.7256

Tablo 6.4 toplu olarak değerlendirildiğinde, beyin tümörü segmentasyonunda "tek bir ideal model veya sekans" yaklaşımının geçerli olmadığı açıkça görülmektedir. Ayrıca segmentasyonda kafatası çıkarımının bir şart olduğu görülmektedir. Sınıflandırma modellerinde çok büyük etkiler gözlemlenmesede segmentasyon için bu önemli bir işlemdir.

Genel lezyon sınırlarının belirlendiği görevlerde (BRISC-2025 ve BraTS WT bölgesi), UNet++ mimarisi belirgin bir üstünlük sağlamıştır ,. UNet++ modelinin sahip olduğu yoğun atlama bağlantıları, özellikle ödem gibi belirsiz sınırları haritalamada son derece başarılı olmuştur. Ayrıca bu bölgelerin tespiti için su tutulumunu (ödemi) en iyi gösteren FLAIR sekansının vazgeçilmez olduğu kanıtlanmıştır.

Diğer taraftan, tümörün iç yapısını karakterize eden Tümör Çekirdeği (TC) ve Aktif/Kontrast Tutan Tümör (ET) bölgelerinin analizinde ise, kontrast maddenin lezyon merkezini aydınlattığı T1+CE sekansı beklendiği üzere en yüksek başarıyı sunmuştur. İlginç bir bulgu olarak, sınırları daha net ancak daha küçük olan bu alt bölgelerin segmentasyonunda standart U-Net ve ResNet34 omurgalı U-Net modellerinin , daha karmaşık olan UNet++'tan kısmen daha kararlı sonuçlar (0.7962 ve 0.7256 Dice) ürettiği saptanmıştır. Bu durum, daha küçük ve belirgin patolojik hedeflerde ağın aşırı parametre karmaşasına girmeden, doğrudan lokalizasyona odaklanan geleneksel U-Net mimarisiyle daha kolay optimize olabildiğini işaret etmektedir.

7
SONUÇ

Bu tez çalışmasında, beyin tümörlerinin manyetik rezonans (MR) görüntüleri üzerinden otomatik tespiti, sınıflandırılması ve bölütlenmesi (segmentasyon) için uçtan uca, çok aşamalı bir derin öğrenme analizi yapılmıştır. Sınırlı donanım kaynakları (NVIDIA GTX 1050 - 3GB VRAM) altında maksimum performansı hedefleyen çalışma, veri odaklı yapay zeka (Data-Centric AI) yaklaşımını benimseyerek; veri temizleme, gelişmiş ön işleme (Beyin Çıkarma), transfer öğrenme stratejileri (RadImageNet vs ImageNet) ve model optimizasyonları (AdamW, AMP) üzerine yoğunlaşmıştır.

Çalışmanın en çarpıcı sonuçlarından biri, genel nesne tanıma veri setleri (ImageNet) yerine, radyolojik görüntülerle eğitilmiş RadImageNet (RIN) ağırlıklarının sağladığı katkıdır. Özellikle dört sınıflı (Glioma, Meningioma, Hipofiz, Sağlıklı) karmaşık sınıflandırma görevinde, InceptionV3 modeli RIN ağırlıkları ve Beyin Çıkarılmış (BE) verilerle eğitildiğinde %99.12 doğruluk oranına ulaşarak standart modelleri geride bırakmıştır. Bu durum, kaynak alan (Source Domain) ile hedef alan (Target Domain) uyumunun derin ağların başarısını doğrudan etkilediğini göstermiştir. Ancak diğer modellerde istenilen performans gözlemlenememiştir. Buda RIN ön eğitiminin doğru model seçimi ile kullanımında etkili olabileceğini ancak RIN verisetinin beyin tümörü için henüz yeterli bir alternatif olarak değerlendirilemeyeceği sonucuna varılmıştır.

Derin öğrenmede daha büyük model daha iyidir algısının aksine, parametre sayısı düşük olan EfficientNet-B0 ve MobileNetV3, devasa modellerle yarışan hatta onları geçen (%99.78 ikili sınıflandırma başarısı) sonuçlar üretmiştir. Bu bulgu, geliştirilen sistemin yüksek işlem gücü gerektirmeden klinikteki standart bilgisayarlarda veya mobil cihazlarda çalışabileceğini kanıtlamaktadır.

Kafatası ve beyin dışı dokuların temizlenmesi (Brain Extraction), ikili sınıflandırmada (Var/Yok) belirgin bir fark yaratmasa da, tümör tiplerinin ayrıldığı çoklu sınıflandırma ve kör testlerde modelin kararlılığını artırmıştır. BE işlemi,

modelin arka plandaki gürültüye değil, doğrudan patolojik dokuya odaklanmasını sağlamıştır.

Swin Transformer ve ViT modelleri, AdamW optimizasyonu ve ImageNet ön eğitimi ile desteklendiğinde CNN modelleriyle rekabet edebilmiştir. Ancak sıfırdan eğitimde (Scratch) başarısız olmaları, Transformer mimarilerinin veri açlığı (data hunger) teorisini doğrulamış ve küçük medikal veri setleri için ön eğitimin şart olduğunu ortaya koymuştur.

Klasik U-Net mimarisi yerine, kodlayıcı (encoder) kısmında ImageNet ön eğitimli ResNet-34 kullanan hibrit modellerin (Linknet, U-Net++, DeepLabV3+) daha başarılı olduğu görülmüştür. Özellikle Linknet ve Unet++, 0.86 ve üzeri Dice skoru ve düşük parametre sayısı ile en verimli segmentasyon modeli olarak öne çıkmıştır.

Çok modlu MR analizinde, T1+CE (Kontrastlı) sekansının aktif tümör ve tümör çekirdeğini belirlemede en kritik modalite olduğu (Dice: 0.58), standart T1 sekansının ise yetersiz kaldığı tespit edilmiştir. FLAIR sekansı ise ödem ve genel kitle tespitinde en yüksek başarıyı (Dice: 0.87) sağlamıştır.

3 boyutlu hacimsel verilerin (3D MRG) donanım kısıtı nedeniyle 2D kesitlere indirgenerek işlenmesi, BraTS veri setindeki segmentasyon skorlarının literatürdeki 3D modellerin (SOTA) gerisinde kalmasına neden olmuştur. Bu durum, z-eksenindeki (derinlik) uzaysal bilginin segmentasyon için hayati olduğunu göstermektedir.

Bu tez çalışması, kısıtlı kaynaklarla yüksek başarımlı bir sistemin mümkün olduğunu göstermiştir. Ancak çalışmanın kapsamını genişletmek ve klinik uygulanabilirliği artırmak adına gelecekte şu çalışmaların yapılması önerilmektedir:

Çalışmadaki en büyük kısıt olan GPU belleği sorunu aşılarak (örneğin A100 veya bulut tabanlı sistemler kullanılarak), BraTS verileri için 3D U-Net veya nnU-Net gibi hacimsel mimariler eğitilmelidir. Bu sayede tümörün 3 boyutlu uzaysal bütünlüğü korunarak segmentasyon başarısı artırılabilir.

Çalışmadaki başarılı sekasnların birlikte eğitimleri ile daha doğru segmentasyonlar yapılan model mimarileri geliştirlebilir. Örneğin T1+CE ve Flair ın birlikte kullanıldığı yapıda Tümör alt elemanları daha doğru bir şekilde saptana bilir.

Geliştirilen Python tabanlı modellerin, doktorların doğrudan kullanabileceği bir web arayüzüne veya PACS (Picture Archiving and Communication System) sistemlerine entegre bir eklentiye dönüştürülmesi, çalışmanın akademik bir çıktı olmaktan öteye geçip toplumsal faydaya dönüşmesini sağlayacaktır.

KAYNAKÇA

[1] F. J. Dorfner, J. B. Patel, J. Kalpathy-Cramer, E. R. Gerstner C. P. Bridge, "A review of deep learning for brain tumor analysis in MRI," *npj Precision Oncology*, c. 9, no. 1, s. 2, 3 Ocak 2025, ISSN: 2397-768X. DOI: 10.1038/s41698-024-00789-2.

[2] S. Bouhafra H. El Bahi, "Deep Learning Approaches for Brain Tumor Detection and Classification Using MRI Images (2020 to 2024): A Systematic Review," *Journal of Imaging Informatics in Medicine*, c. 38, no. 3, ss. 1403–1433, 1 Haz. 2025, ISSN: 2948-2933. DOI: 10.1007/s10278-024-01283-8.

[3] K. He, X. Zhang, S. Ren J. Sun, "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, ss. 770–778, Haz. 2016. DOI: 10.1109/CVPR.2016.90.

[4] X. Mei ve diğ., "RadImageNet: An Open Radiologic Deep Learning Research Dataset for Effective Transfer Learning," *Radiology: Artificial Intelligence*, c. 4, no. 5, e210315, Eyl. 2022. DOI: 10.1148/ryai.210315.

[5] B. M. Pacheco, G. d. S. e. Cassia D. Silva, "Towards fully automated deep-learning-based brain tumor segmentation: is brain extraction still necessary?" *Biomedical Signal Processing and Control*, c. 82, s. 104 514, Nis. 2023, ISSN: 17468094. DOI: 10.1016/j.bspc.2022.104514. arXiv: 2212.07497[eess].

[6] D. N. Louis ve diğ., "The 2021 WHO Classification of Tumors of the Central Nervous System: a summary," *Neuro-Oncology*, c. 23, no. 8, ss. 1231–1251, 2 Ağus. 2021, ISSN: 1522-8517, 1523-5866. DOI: 10.1093/neuonc/noab106.

[7] F. Bray ve diğ., "Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, c. 74, no. 3, ss. 229–263, May. 2024, ISSN: 0007-9235, 1542-4863. DOI: 10.3322/caac.21834.

[8] J. Watts, G. Box, A. Galvin, P. Brotchie, N. Trost T. Sutherland, "Magnetic resonance imaging of meningiomas: a pictorial review," *Insights into Imaging*, c. 5, no. 1, ss. 113–122, Şub. 2014, ISSN: 1869-4101. DOI: 10.1007/s13244-013-0302-4.

[9] O. Casar-Borota, P. Burman M. B. Lopes, "The 2022 WHO classification of tumors of the pituitary gland: An update on aggressive and metastatic pituitary neuroendocrine tumors," *Brain Pathology*, c. 35, no. 1, e13302, Ocak 2025, ISSN: 1015-6305, 1750-3639. DOI: 10.1111/bpa.13302.

[10] M. Martucci ve diğ., "Magnetic Resonance Imaging of Primary Adult Brain Tumors: State of the Art and Future Perspectives," *Biomedicines*, c. 11, no. 2, s. 364, 26 Ocak 2023, ISSN: 2227-9059. DOI: 10.3390/biomedicines11020364.

[11] M. Weller ve diğ., "EANO guidelines on the diagnosis and treatment of diffuse gliomas of adulthood," *Nature Reviews Clinical Oncology*, c. 18, no. 3, ss. 170–186, Mar. 2021, ISSN: 1759-4774, 1759-4782. DOI: 10.1038/s41571-020-00447-z.

[12] T. Çukur ve diğ., "A Tutorial on MRI Reconstruction: From Modern Methods to Clinical Implications," *IEEE Transactions on Biomedical Engineering*, ss. 1–20, 2025, ISSN: 0018-9294, 1558-2531. DOI: 10.1109/TBME.2025.3617575.

[13] D. B. Plewes W. Kucharczyk, "Physics of MRI: A primer," *Journal of Magnetic Resonance Imaging*, c. 35, no. 5, ss. 1038–1054, May. 2012, ISSN: 1053-1807, 1522-2586. DOI: 10.1002/jmri.23642.

[14] F. Bloch, "Nuclear Induction," *Physical Review*, c. 70, no. 7, ss. 460–474, 1 Ekim 1946, ISSN: 0031-899X. DOI: 10.1103/PhysRev.70.460.

[15] E. M. Purcell, H. C. Torrey R. V. Pound, "Resonance Absorption by Nuclear Magnetic Moments in a Solid," *Physical Review*, c. 69, no. 1, ss. 37–38, 1 Ocak 1946, ISSN: 0031-899X. DOI: 10.1103/PhysRev.69.37.

[16] P. C. Lauterbur, "Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance," *Nature*, c. 242, no. 5394, ss. 190–191, Mar. 1973, ISSN: 0028-0836, 1476-4687. DOI: 10.1038/242190a0.

[17] Ž. Đ. Vujović, "Magnets, Gradients, and RF Coils of MR Scanners," *International Journal of Physics Research and Applications*, c. 6, no. 2, ss. 128–135, 25 Tem. 2023, ISSN: 2766-2748. DOI: 10.29328/journal.ijpra.1001062.

[18] R. A. Pooley, "Fundamental Physics of MR Imaging," *RadioGraphics*, c. 25, no. 4, ss. 1087–1099, Tem. 2005, ISSN: 0271-5333, 1527-1323. DOI: 10.1148/rg.254055027.

[19] S. D. Serai, "Basics of magnetic resonance imaging and quantitative parameters T1, T2, T2*, T1rho and diffusion-weighted imaging," *Pediatric Radiology*, c. 52, no. 2, ss. 217–227, Şub. 2022, ISSN: 0301-0449, 1432-1998. DOI: 10.1007/s00247-021-05042-7.

[20] Oral Diagnoz ve Maksillofasiyal Radyoloji Derneği, "MR FİZİĞİ VE TEMEL KURALLAR SEKANSLAR," ODMFR Yayınları, Eğitim Materyali. erişim tarihi 16 Ocak 2026. erişim adresi: https://www.odmfr.org/storage/app/uploads/public/62e/373/ff8/62e373ff84628074890122.pdf.

[21] Y. LeCun, Y. Bengio G. Hinton, "Deep learning," *Nature*, c. 521, no. 7553, ss. 436–444, 28 May. 2015, ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature14539.

[22] W. S. McCulloch W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The Bulletin of Mathematical Biophysics*, c. 5, no. 4, ss. 115–133, Ara. 1943, ISSN: 0007-4985, 1522-9602. DOI: 10.1007/BF02478259.

[23] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain.," *Psychological Review*, c. 65, no. 6, ss. 386–408, 1958, ISSN: 1939-1471, 0033-295X. DOI: 10.1037/h0042519.

[24] M. Minsky S. A. Papert, *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA, USA: MIT Press, 15 Ocak 1969, 258 ss., ISBN: 978-0-262-63022-1.

[25] D. E. Rumelhart, G. E. Hinton R. J. Williams, "Learning representations by back-propagating errors," *Nature*, c. 323, no. 6088, ss. 533–536, Ekim 1986, ISSN: 1476-4687. DOI: 10.1038/323533a0.

[26] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, c. 36, no. 4, ss. 193–202, Nis. 1980, ISSN: 0340-1200, 1432-0770. DOI: 10.1007/BF00344251.

[27] Y. LeCun ve diğ., "Backpropagation Applied to Handwritten Zip Code Recognition," *Neural Computation*, c. 1, no. 4, ss. 541–551, Ara. 1989, ISSN: 0899-7667, 1530-888X. DOI: 10.1162/neco.1989.1.4.541.

[28] A. Krizhevsky, I. Sutskever G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing Systems*, c. 25, 2012.

[29] O. Ronneberger, P. Fischer T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells A. F. Frangi, ed., ss. 234–241, 2015. DOI: 10.1007/978-3-319-24574-4_28.

[30] A. Vaswani ve diğ., "Attention is All you Need," *Advances in Neural Information Processing Systems*, c. 30, 2017.

[31] A. Dosovitskiy ve diğ., *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*, Version Number: 2, 2020. DOI: 10.48550/ARXIV.2010.11929.

[32] V. Nair G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," *Proceedings of the 27th international conference on machine learning (ICML-10)*, 2010.

[33] L. Bottou O. Bousquet, "The Tradeoffs of Large Scale Learning," *Advances in Neural Information Processing Systems*, c. 20, 2007. erişim tarihi 20 Ocak 2026.

[34] D. P. Kingma J. Ba, *Adam: A Method for Stochastic Optimization*, 30 Ocak 2017. DOI: 10.48550/arXiv.1412.6980. arXiv: 1412.6980[cs]. erişim tarihi 20 Ocak 2026.

[35] I. Loshchilov F. Hutter, *Decoupled Weight Decay Regularization*, 4 Ocak 2019. DOI: 10.48550/arXiv.1711.05101. arXiv: 1711.05101[cs]. erişim tarihi 20 Ocak 2026.

[36] J. Yun ve diğ., "Deep radiomics-based survival prediction in patients with chronic obstructive pulmonary disease," *Scientific Reports*, c. 11, no. 1, s. 15 144, 26 Tem. 2021, ISSN: 2045-2322. DOI: 10.1038/s41598-021-94535-4.

[37] G. Huang, Z. Liu, L. Van Der Maaten K. Q. Weinberger, "Densely Connected Convolutional Networks," içinde maintitle+booktitle, Honolulu, HI: IEEE, Tem. 2017, ss. 2261–2269, ISBN: 978-1-5386-0457-1. DOI: 10.1109/CVPR.2017.243.

[38] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, ss. 2818–2826, Haz. 2016. DOI: 10.1109/CVPR.2016.308.

[39] A. G. Howard ve diğ., *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*, 17 Nis. 2017. DOI: 10.48550/arXiv.1704.04861. arXiv: 1704.04861[cs].

[40] A. Howard ve diğ., "Searching for MobileNetV3," *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, ss. 1314–1324, Ekim 2019. DOI: 10.1109/ICCV.2019.00140.

[41] M. Tan Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *Proceedings of the 36th International Conference on Machine Learning*, ss. 6105–6114, 24 May. 2019, ISSN: 2640-3498.

[42] D. Hendrycks K. Gimpel, *Gaussian Error Linear Units (GELUs)*, 6 Haz. 2023. DOI: 10.48550/arXiv.1606.08415. arXiv: 1606.08415[cs].

[43] Z. Liu ve diğ., "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, ss. 9992–10 002, Ekim 2021. DOI: 10.1109/ICCV48922.2021.00986.

[44] S. J. Pan Q. Yang, "A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering*, c. 22, no. 10, ss. 1345–1359, Ekim 2010, ISSN: 1558-2191. DOI: 10.1109/TKDE.2009.191.

[45] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," içinde maintitle+booktitle, Haz. 2009, ss. 248–255. DOI: 10.1109/CVPR.2009.5206848.

[46] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh J. Liang, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, D. Stoyanov ve diğ., ed., ss. 3–11, 2018. DOI: 10.1007/978-3-030-00889-5_1.

[47] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff H. Adam, "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," *Computer Vision – ECCV 2018*, V. Ferrari, M. Hebert, C. Sminchisescu Y. Weiss, ed., c. 11211, Series Title: Lecture Notes in Computer Science, Cham: Springer International Publishing, 2018, ss. 833–851. DOI: 10.1007/978-3-030-01234-2_49.

[48] A. Chaurasia E. Culurciello, "LinkNet: Exploiting encoder representations for efficient semantic segmentation," *2017 IEEE Visual Communications and Image Processing (VCIP)*, ss. 1–4, Ara. 2017. DOI: 10.1109/VCIP.2017.8305148.

[49] T. Fan, G. Wang, Y. Li H. Wang, "MA-Net: A Multi-Scale Attention Network for Liver and Tumor Segmentation," *IEEE Access*, c. 8, ss. 179 656–179 665, 2020, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.3025372.

[50] A. H. Hamada, *Br35H :: Brain Tumor Detection 2020*. DOI: 10.21227/T5K6-5R54.

[51] Msoud Nickparvar, *Brain Tumor MRI Dataset*. DOI: 10.34740/KAGGLE/DSV/2645886. erişim adresi: https://www.kaggle.com/dsv/2645886.

[52] Amirreza Fateh ve diğ., *BRISC 2025*. DOI: 10.34740/KAGGLE/DS/7632487.

[53] S. Bakas ve diğ., "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features," *Scientific Data*, c. 4, s. 170 117, 5 Eyl. 2017, ISSN: 2052-4463. DOI: 10.1038/sdata.2017.117.

[54] B. H. Menze ve diğ., "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE transactions on medical imaging*, c. 34, no. 10, ss. 1993–2024, Ekim 2015, ISSN: 1558-254X. DOI: 10.1109/TMI.2014.2377694.

[55] S. Bakas ve diğ., *Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge*, 23 Nis. 2019. DOI: 10.48550/arXiv.1811.02629. arXiv: 1811.02629[cs].

[56] S. Bakas ve diğ., *Segmentation Labels for the Pre-operative Scans of the TCGA-GBM collection*, ortak. yap. TCIA Team, ver. 1, 2017. DOI: 10.7937/K9/TCIA.2017.KLXWJJ1Q.

[57] S. Bakas ve diğ., *Segmentation Labels for the Pre-operative Scans of the TCGA-LGG collection*, ortak. yap. TCIA Team, ver. 1, 2017. DOI: 10.7937/K9/TCIA.2017.GJQ7R0EF.

[58] M. A. Gómez-Guzmán ve diğ., "Classifying Brain Tumors on Magnetic Resonance Imaging by Using Convolutional Neural Networks," *Electronics*, c. 12, no. 4, s. 955, Ocak 2023, ISSN: 2079-9292. DOI: 10.3390/electronics12040955.

[59] A. Iqbal, M. A. Jaffar R. Jahangir, "Enhancing Brain Tumour Multi-Classification Using Efficient-Net B0-Based Intelligent Diagnosis for Internet of Medical Things (IoMT) Applications," *Information*, c. 15, no. 8, s. 489, Ağus. 2024, ISSN: 2078-2489. DOI: 10.3390/info15080489.

[60] M. A. Talukder ve diğ., "An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning," *Expert Systems with Applications*, c. 230, s. 120 534, 15 Kas. 2023, ISSN: 0957-4174. DOI: 10.1016/j.eswa.2023.120534.

[61] G. Ramasamy, T. Singh X. Yuan, “Multi-Modal Semantic Segmentation Model using Encoder Based Link-Net Architecture for BraTS 2020 Challenge,” *Procedia Computer Science*, International Conference on Machine Learning and Data Engineering, c. 218, ss. 732–740, 1 Ocak 2023, ISSN: 1877-0509. DOI: 10.1016/j.procs.2023.01.053.

TEZDEN ÜRETİLMİŞ YAYINLAR

Konferans Bildirisi

1. Güneşer, E., & BAL, M. "A Comparative Analysis of RadImageNet and ImageNet Pretraining for Brain Tumor MRI Classification". 23TH International İstanbul Scientific Research Congress on Life, Engineering, Architecture and Mathematical Sciences Proceedings Book.