

T.C.
BALIKESİR ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR VE BİLİŞİM MÜHENDİSLİĞİ ANABİLİM DALI



MRI VERİLERİNDEN
ALZHEİMER HASTALIĞININ TANISINDA EN İYİ DERİN
ÖĞRENME MİMARİLERİNİN KEŞFEDİLMESİ

ASLIHAN GÜVEN

YÜKSEK LİSANS TEZİ

Jüri Üyeleri : Doç. Dr. Fatih AYDIN (Tez Danışmanı)
Doç. Dr. Hüseyin GÜNEŞ
Dr. Öğr. Üyesi Fahrettin Burak DEMİR

BALIKESİR, ŞUBAT – 2026

ETİK BEYAN

Balıkesir Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak tarafımda hazırlanan “**MRI Verilerinden Alzheimer Hastalığının Tanısında En İyi Derin Öğrenme Mimarilerinin Keşfedilmesi**” başlıklı tezde;

- Tüm bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- Kullanılan veriler ve sonuçlarda herhangi bir değişiklik yapmadığımı,
- Tüm bilgi ve sonuçları bilimsel araştırma ve etik ilkelere uygun şekilde sunduğumu,
- Yararlandığım eserlere atıfta bulunarak kaynak gösterdiğimi,

beyan eder, aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ederim.

Aslıhan GÜVEN

ÖZET

**MRI VERİLERİNDEN ALZHEİMER HASTALIĞININ TANISINDA EN İYİ
DERİN ÖĞRENME MİMARİLERİNİN KEŞFEDİLMESİ
YÜKSEK LİSANS TEZİ
ASLIHAN GÜVEN
BALIKESİR ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR VE BİLİŞİM MÜHENDİSLİĞİ ANABİLİM DALI
(TEZ DANIŞMANI: DOÇ. DR. FATİH AYDIN)**

BALIKESİR, ŞUBAT - 2026

Bu tez çalışmasında, Alzheimer hastalığının manyetik rezonans görüntüleme verileri üzerinden otomatik teşhisine yönelik en etkili derin öğrenme mimarisinin belirlenmesi amaçlanmıştır. Bu kapsamda, EfficientNet-B0, ResNet-50, U-Net, MobileNetV2 ve InceptionV3 mimarileri aynı veri kümesi ve sabit hiperparametreler altında karşılaştırmalı olarak değerlendirilmiştir. Çalışmada, 18–96 yaş aralığındaki 416 bireye ait manyetik rezonans görüntülerinden oluşan ve dört farklı Alzheimer evresini içeren OASIS-1 veri kümesi kullanılmıştır.

Modeller, Alzheimer hastalığının Non-Demented, Very Mild Demented, Mild Demented ve Moderate Demented evrelerini içeren etiketli MRI verileri üzerinde eğitilmiştir. Performans değerlendirmesi doğruluk, kesinlik, duyarlılık, F1 skoru ve ROC-AUC metrikleri üzerinden gerçekleştirilmiş; buna ek olarak model boyutu, eğitim süresi ve parametre yoğunluğu gibi operasyonel kriterler de dikkate alınmıştır.

DeneySEL sonuçlar, InceptionV3 mimarisinin %0.86 test doğruluğu ve 0.97 makro AUC değeri ile en yüksek sınıflandırma performansını elde ettiğini göstermiştir. MobileNetV2 modeli, %0.74 test doğruluğu ve 0.90 makro AUC değeri ile düşük parametre sayısına rağmen dengeli ve verimli bir performans sunmuştur. EfficientNet-B0 mimarisi %0.63 test doğruluğu ile orta seviyede bir performans sergilerken, ResNet-50 modeli 0.95 makro AUC değeri ile yüksek ayrıştırma gücü göstermiştir. U-Net Encoder mimarisi ise diğer modellere kıyasla daha sınırlı bir sınıflandırma başarısı elde etmiştir.

Elde edilen bulgular, derin öğrenme tabanlı yaklaşımların Alzheimer hastalığının erken evre teşhisinde etkili ve klinik uygulamalara entegre edilebilir olduğunu ortaya koymaktadır.

ANAHTAR KELİMELELER: Alzheimer, derin öğrenme, MRI, sınıflandırma, yapay zekâ

Bilim Kod / Kodları : 92418, 92432, 92444

Sayfa Sayısı : 126

ABSTRACT

DISCOVERING THE BEST DEEP LEARNING ARCHITECTURES FOR THE DIAGNOSIS OF ALZHEIMER'S DISEASE FROM MRI DATA

MSC THESIS

ASLIHAN GUVEN

BALIKESIR UNIVERSITY INSTITUTE OF SCIENCE

DEPARTMENT OF COMPUTER AND INFORMATION ENGINEERING

(SUPERVISOR: ASSOC. PROF. DR. FATI H AYDIN)

BALIKESİR, FEBRUARY - 2026

In this thesis, the aim is to identify the most effective deep learning architecture for the automated diagnosis of Alzheimer's disease using magnetic resonance imaging data. Within this scope, the EfficientNet-B0, ResNet-50, U-Net, MobileNetV2, and InceptionV3 architectures were comparatively evaluated under the same dataset and fixed hyperparameter settings. The study utilized the OASIS-1 dataset, which consists of magnetic resonance images obtained from 416 individuals aged between 18 and 96 years and includes four different stages of Alzheimer's disease.

The models were trained on a labeled MRI dataset representing the Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented stages of Alzheimer's disease. Performance evaluation was conducted using accuracy, precision, recall, F1-score, and ROC-AUC metrics. In addition to classification performance, operational criteria such as model size, training time, and parameter density were also taken into consideration.

Experimental results demonstrated that the InceptionV3 architecture achieved the highest classification performance, with a test accuracy of 0.86 and a macro AUC value of 0.97. The MobileNetV2 model provided a balanced and efficient performance, achieving a test accuracy of 0.74 and a macro AUC of 0.90 despite its low parameter count. EfficientNet-B0 exhibited a moderate performance with a test accuracy of 0.63, while the ResNet-50 model showed strong discriminative capability with a macro AUC value of 0.95. The U-Net Encoder architecture yielded more limited classification performance compared to the other models.

The findings indicate that deep learning-based approaches are effective and suitable for integration into clinical applications for the early diagnosis of Alzheimer's disease.

KEYWORDS: Alzheimer, deep learning, MRI, classification, artificial intelligence

İÇİNDEKİLER

Sayfa

| | |
|--|-------------|
| ÖZET | i |
| ABSTRACT | ii |
| İÇİNDEKİLER | iii |
| ŞEKİL LİSTESİ | vi |
| TABLO LİSTESİ | vii |
| KISALTMALAR LİSTESİ | viii |
| ÖNSÖZ VE TEŞEKKÜR | x |
| 1. GİRİŞ | 1 |
| 1.1 Alzheimer Hastalığının Küresel Önemi ve Erken Teşhisin Rolü | 2 |
| 1.2 Alzheimer Tanısında Klinik Yaklaşımlar ve Yapay Zekâ Tabanlı Sistemler | 4 |
| 1.2.1 Beyin görüntüleme teknikleri ve MRI'nın rolü | 7 |
| 1.3 Problem Tanımı..... | 8 |
| 1.4 Araştırmanın Amacı | 10 |
| 1.4.1 Araştırma hedefleri | 10 |
| 1.5 Literatür Boşlukları ve Bu Tezin Katkısı | 11 |
| 1.5.1 Literatürdeki başlıca boşluklar..... | 11 |
| 1.5.2 Bu tezin katkıları..... | 12 |
| 1.6 Araştırma Soruları ve Hipotezler | 14 |
| 1.6.1 Araştırma soruları | 14 |
| 1.6.2 Araştırma hipotezleri | 14 |
| 1.7 Tezin Kapsamı ve Sınırlılıkları | 16 |
| 1.8 Tezin Organizasyonu..... | 17 |
| 2. LİTERATÜR TARAMASI | 19 |
| 2.1 Derin Öğrenme Yöntemlerinin Tıpta Kullanımı | 19 |
| 2.1.1 Tıpta derin öğrenmenin tarihsel gelişimi | 22 |
| 2.1.2 Tıbbi görüntü analizinde karşılaşılan teknik zorluklar | 23 |
| 2.1.3 Hesaplama gücü, GPU kullanımı ve maliyetler | 24 |
| 2.1.4 Federated learning ve veri gizliliği odaklı yaklaşımlar | 25 |
| 2.2 Alzheimer Tanısında Kullanılan Derin Öğrenme Mimarileri | 25 |
| 2.2.1 EfficientNet..... | 26 |
| 2.2.2 ResNet (residual network) | 28 |
| 2.2.3 U-Net | 29 |
| 2.2.4 MobileNetV2 | 30 |
| 2.2.5 Inception (GoogleNet) | 32 |
| 2.2.6 Diğer derin öğrenme yaklaşımları ve hibrit modeller..... | 33 |
| 2.3 Transfer Öğrenme ve Veri Dengeleme Teknikleri..... | 34 |
| 2.3.1 Veri dengeleme ve aşırı öğrenmeyi önleyici teknikler | 36 |
| 2.4 Alzheimer Tanısında Derin Öğrenme Modellerinin Klinik Kullanılabilirliği | 37 |
| 3. MATERYAL VE YÖNTEM | 40 |
| 3.1 Kullanılan Veri Kümesi | 40 |
| 3.1.1 Veri seti özellikleri | 40 |
| 3.1.2 Veri ön işleme adımları | 41 |
| 3.1.2.1 Görüntü boyutlandırma (resizing) | 41 |
| 3.1.2.2 Normalizasyon (rescaling) | 42 |

İÇİNDEKİLER (devam)

| | | |
|-----------|--|-----------|
| 3.1.2.3 | Veri artırma (data augmentation)..... | 43 |
| 3.1.2.4 | Veri ayrımı (split strategy)..... | 44 |
| 3.2 | Model Eğitimi Genel Çerçeve..... | 45 |
| 3.2.1 | Donanım ve yazılım altyapısı..... | 45 |
| 3.2.2 | Hiperparametre ayarları..... | 46 |
| 3.2.2.1 | Optimizasyon algoritması..... | 46 |
| 3.2.2.2 | Kayıp fonksiyonu (loss function)..... | 47 |
| 3.2.2.3 | Performans metrikleri..... | 47 |
| 3.2.2.4 | Epoch sayısı ve erken durdurma stratejisi..... | 48 |
| 3.2.2.5 | Öğrenme oranı planlaması (learning rate scheduling)..... | 48 |
| 3.2.2.6 | Model ağırlıklarının kaydedilmesi..... | 48 |
| 3.3 | Kullanılan Derin Öğrenme Mimarileri..... | 49 |
| 3.3.1 | EfficientNet : Ölçeklenebilir ve parametre verimli derin CNN mimarisi..... | 49 |
| 3.3.1.1 | Uygulanan EfficientNet-B0 yapılandırması..... | 50 |
| 3.3.1.2 | EfficientNet-B0 eğitim stratejisi ve optimizasyon parametreleri..... | 51 |
| 3.3.1.3 | EfficientNet-B0 mimarisi teknik avantajlar ve sınırlılıklar..... | 52 |
| 3.3.2 | ResNet: Derin ağlarda rezidüel öğrenme yaklaşımı..... | 53 |
| 3.3.2.1 | Uygulanan ResNet-50 yapılandırması..... | 55 |
| 3.3.2.2 | ResNet-50 eğitim stratejisi ve optimizasyon parametreleri..... | 56 |
| 3.3.2.3 | ResNet-50 mimarisi: teknik avantajlar ve sınırlılıklar..... | 57 |
| 3.3.3 | U-Net: Simetrik encoder-decoder yapısı ile detay koruyucu derin ağ mimarisi..... | 58 |
| 3.3.3.1 | Uygulanan U-Net konfigürasyonu (sınıflandırma uyarlaması)..... | 59 |
| 3.3.3.2 | U-Net eğitim stratejisi ve optimizasyon parametreleri..... | 60 |
| 3.3.3.3 | U-Net mimarisi: Teknik avantajlar ve sınırlılıklar..... | 61 |
| 3.3.4 | MobileNetV2: Hafif ve verimli derin ağ mimarisi..... | 62 |
| 3.3.4.1 | Kullanılan MobileNetV2 yapılandırması..... | 63 |
| 3.3.4.2 | MobileNetV2 eğitim stratejisi ve optimizasyon parametreleri..... | 64 |
| 3.3.4.3 | MobileNetV2 mimarisi: Teknik avantajlar ve sınırlılıklar..... | 66 |
| 3.3.5 | Inception: Çok ölçekli özellik çıkarımı için derin CNN mimarisi..... | 67 |
| 3.3.5.1 | Kullanılan InceptionV3 yapılandırması..... | 68 |
| 3.3.5.2 | InceptionV3 eğitim stratejisi ve optimizasyon parametreleri..... | 70 |
| 3.3.5.3 | InceptionV3 teknik avantajlar ve sınırlılıklar..... | 70 |
| 3.4 | Performans Değerlendirme Kriterleri..... | 71 |
| 3.4.1 | Sınıflandırma performansı: İstatistiksel metrikler..... | 72 |
| 3.4.1.1 | Accuracy (Genel Doğruluk Oranı)..... | 72 |
| 3.4.1.2 | Precision (pozitif tahmin kesinliği)..... | 73 |
| 3.4.1.3 | Recall (duyarlılık / sensitivity)..... | 74 |
| 3.4.1.4 | F1-Score (harmonik başarı) ve macro-F1..... | 75 |
| 3.4.2 | Metriklerin uygulanması ve izlenmesi..... | 76 |
| 3.4.3 | Model boyutu ve hafiflik analizi..... | 76 |
| 3.4.4 | Eğitim süresi ve donanımsal zaman maliyeti..... | 77 |
| 3.4.5 | Parametre yoğunluğu ve hesaplama karmaşıklığı..... | 77 |
| 3.4.6 | Toplam performans ve uygulama uyumluluğu..... | 77 |
| 4. | DENEYSEL BULGULAR..... | 79 |
| 4.1 | Eğitim Süreci ve Genel Gözlemler..... | 79 |
| 4.1.1 | Veri bölünmesi ve eğitim senaryosu..... | 79 |
| 4.1.2 | Ortak ön işleme ve veri artırma (augmentation)..... | 79 |
| 4.1.3 | İki aşamalı eğitim stratejisi (freezing - fine-tuning)..... | 80 |

İÇİNDEKİLER (devam)

| | |
|---|------------|
| 4.1.4 Ortak eğitim dinamikleri..... | 80 |
| 4.1.5 Modeller arası genel eğitim gözlemleri | 81 |
| 4.1.6 Genel değerlendirme..... | 82 |
| 4.2 Modellerin Karşılaştırmalı Deneysel Performansı | 83 |
| 4.2.1 ACC–Loss eğrilerinin karşılaştırmalı analizi | 83 |
| 4.2.2 Doğrulama ve test performanslarının karşılaştırmalı analizi | 89 |
| 4.2.3 Confusion matrix sonuçlarının karşılaştırmalı analizi | 92 |
| 4.2.4 ROC–AUC eğrilerinin karşılaştırmalı analizi | 98 |
| 4.3 Model Boyutu ve Hafiflik Analizi | 104 |
| 4.3.1 Model boyutu karşılaştırması | 104 |
| 4.3.2 Parametre sayısı karşılaştırması..... | 104 |
| 4.4 Eğitim Süresi ve Hesaplama Maliyet Analizi | 105 |
| 4.5 Parametre Yoğunluğu ve FLOP Karmaşıklığı | 106 |
| 4.6 Genel Değerlendirme | 107 |
| 4.6.1 Doğruluk, AUC ve F1 skoru açısından karşılaştırma | 107 |
| 4.6.2 Model hafifliği ve hesaplama verimliliği açısından karşılaştırma..... | 108 |
| 4.6.3 Eğitim süresi açısından karşılaştırma | 109 |
| 4.6.4 Klinik kullanılabilirlik açısından değerlendirme | 110 |
| 4.6.5 Genel sonuç | 111 |
| 5. TARTIŞMA..... | 113 |
| 5.1 Bulguların Genel Değerlendirmesi..... | 113 |
| 5.2 Modellerin Sayısal Karşılaştırması | 114 |
| 5.3 Klinik Yorum ve Tanısal Uygunluk..... | 114 |
| 5.4 Sınırlılıklar ve Gözlemler | 115 |
| 6. SONUÇ | 116 |
| 7. GELECEK ÇALIŞMALAR..... | 117 |
| 8. KAYNAKLAR..... | 119 |
| ÖZGEÇMİŞ | 126 |

ŞEKİL LİSTESİ

Sayfa

| | |
|--|-----|
| Şekil 3.1: OASIS veri setinde bulunan örnek MRI görüntüleri..... | 41 |
| Şekil 3.2: EfficientNet-B0 mimarisinin katman bazlı gösterimi..... | 50 |
| Şekil 3.3: ResNet-50 mimarisinin genel katman yapısı..... | 54 |
| Şekil 3.4: U-Net mimarisinin genel yapısı ve katmanlar arası veri akışı..... | 59 |
| Şekil 3.5: MobileNetV2 mimarisinin genel blok diyagramı..... | 63 |
| Şekil 3.6: InceptionV3 mimarisinin genel yapısı..... | 68 |
| Şekil 4.1: MobilenetV2 modelinin eğitim sürecine ait doğruluk ve kayıp eğrileri..... | 84 |
| Şekil 4.2: EfficientNet-B0 modelinin eğitim sürecine ait doğruluk ve kayıp eğrileri..... | 85 |
| Şekil 4.3: ResNet-50 modelinin eğitim sürecine ait doğruluk ve kayıp eğrileri..... | 86 |
| Şekil 4.4: InceptionV3 modelinin eğitim sürecine ait doğruluk ve kayıp eğrileri..... | 87 |
| Şekil 4.5: U-Net modelinin eğitim sürecine ait doğruluk ve kayıp eğrileri..... | 88 |
| Şekil 4.6: MobilenetV2 modeli için doğrulama ve test karışıklık matrisleri..... | 93 |
| Şekil 4.7: EfficientNet-B0 modeli için doğrulama ve test karışıklık matrisleri..... | 94 |
| Şekil 4.8: ResNet-50 modeli için doğrulama ve test karışıklık matrisleri..... | 95 |
| Şekil 4.9: InceptionV3 modeli için doğrulama ve test karışıklık matrisleri..... | 96 |
| Şekil 4.10: U-Net modeli için doğrulama ve test karışıklık matrisleri..... | 97 |
| Şekil 4.11: MobileNetV2 modeli için doğrulama ve test ROC eğrileri..... | 99 |
| Şekil 4.12: EfficientNet-B0 modeli için doğrulama ve test ROC eğrileri..... | 100 |
| Şekil 4.13: ResNet-50 modeli için doğrulama ve test ROC eğrileri..... | 101 |
| Şekil 4.14: InceptionV3 modeli için doğrulama ve test ROC eğrileri..... | 102 |
| Şekil 4.15: U-Net modeli için doğrulama ve test ROC eğrileri..... | 103 |
| Şekil 4.16: Modellerin doğruluk, AUC ve F1 skorlarına göre karşılaştırılması..... | 108 |
| Şekil 4.17: Modellerin parametre sayısı ve FLOPs değerlerine göre karşılaştırılması..... | 109 |
| Şekil 4.18: Derin öğrenme modellerinin eğitim sürelerine göre karşılaştırılması..... | 110 |

TABLO LİSTESİ

| | <u>Sayfa</u> |
|--|--------------|
| Tablo 3.1: EfficientNet-B0 modeli için kullanılan eğitim hiperparametreleri. | 52 |
| Tablo 3.2: ResNet-50 modeli için eğitim hiperparametreleri. | 57 |
| Tablo 3.3: U-Net modeli için eğitim hiperparametreleri. | 61 |
| Tablo 3.4: MobileNetV2 modeli için eğitim hiperparametreleri. | 65 |
| Tablo 3.5: InceptionV3 modeli için eğitim hiperparametreleri. | 70 |
| Tablo 4.1: MobileNetV2 modelinin doğrulama ve test performans özeti. | 90 |
| Tablo 4.2: EfficientNet-B0 modelinin doğrulama ve test performans özeti. | 90 |
| Tablo 4.3: ResNet-50 modelinin doğrulama ve test performans özeti. | 91 |
| Tablo 4.4: InceptionV3 modelinin doğrulama ve test performans özeti. | 91 |
| Tablo 4.5: U-Net modelinin doğrulama ve test performans özeti. | 92 |
| Tablo 4.6: Modellerin parametre boyutları ve açıklamaları. | 104 |
| Tablo 4.7: Modellerin parametre sayılarının karşılaştırılması. | 105 |
| Tablo 4.8: Modellerin eğitim sürelerinin karşılaştırılması. | 106 |
| Tablo 4.9: Modellerin yaklaşık FLOP değerleri ve hesaplama maliyeti karşılaştırması. | 106 |
| Tablo 5.1: Modellerin sayısal karşılaştırması. | 114 |

KISALTMALAR LİSTESİ

| | |
|-----------------|---|
| AC | : Yardımcı sınıflandırıcı (Auxiliary Classifier) |
| ACC | : Doğruluk oranı (Accuracy) |
| ADNI | : Alzheimer Hastalığı Nörogörüntüleme Girişimi |
| AI | : Yapay zekâ (Artificial Intelligence) |
| AIBL | : Avustralya Görüntüleme, Biyobelirteç ve Yaşam Tarzı Çalışması |
| ANOVA | : Tek yönlü varyans analizi |
| ApoE4 | : Alzheimer hastalığı ile ilişkili risk geni aleli |
| AWS | : Amazon Web Hizmetleri (Amazon Web Services) |
| BN | : Yığın normalizasyonu (Batch Normalization) |
| BT | : Bilgisayarlı tomografi (Computed Tomography) |
| CDR | : Klinik Demans Derecelendirmesi (Clinical Dementia Rating) |
| CM | : Karmaşıklık matrisi (Confusion Matrix) |
| CNN | : Evrimsel Sinir Ağı (Convolutional Neural Network) |
| CN | : Sağlıklı birey / normal bilişsel durum (Cognitively Normal) |
| CPU | : Merkezi İşlem Birimi |
| DSC | : Dice benzerlik katsayısı (Dice Similarity Coefficient) |
| DTI | : Difüzyon Tensör Görüntüleme |
| EMA | : Avrupa İlaç Ajansı (European Medicines Agency) |
| fMRI | : Fonksiyonel manyetik rezonans görüntüleme (Functional Magnetic Resonance Imaging) |
| FDA | : ABD Gıda ve İlaç Dairesi |
| FLOP | : Kayan nokta işlem sayısı (Floating Point Operations) |
| FL | : Birleşik öğrenme (Federated Learning) |
| GAN | : Üretici çekişmeli ağ (Generative Adversarial Network) |
| GLCM | : Gri-seviye eş-oluşum matrisi |
| Grad-CAM | : Gradyan ağırlıklı sınıf etkinleştirme haritaları (Gradient-weighted Class Activation Mapping) |
| GPU | : Grafik İşlem Birimi |
| JPEG | : Ortak Fotoğraf Uzmanları Grubu görüntü formatı |
| L2 | : L2 düzenleme (ağırlık ceza terimi) |
| LR | : Öğrenme oranı (Learning Rate) |
| LRP | : Katman bazlı önem yayılımı |
| LSTM | : Uzun-kısa süreli bellek ağı |
| MB | : Megabayt (model/dosya boyutu) |
| MBCnv | : EfficientNet'in temel konvolüsyon bloğu |
| MCI | : Hafif bilişsel bozukluk (Mild Cognitive Impairment) |
| MD | : Hafif demans (Mild Demented) |
| MLP | : Çok katmanlı algılayıcı |
| MMSE | : Mini-Mental Durum Testi |
| MoCA | : Montreal Bilişsel Değerlendirme |
| MOD | : Orta evre demans (Moderate Demented) |
| MRI | : Manyetik Rezonans Görüntüleme |
| ND | : Demans olmayan birey (Non-Demented) |
| OASIS | : Açık Erişimli Görüntüleme Serileri |
| PET | : Pozitron Emisyon Tomografisi |
| SE | : Sıkıştırma-uyarma bloğu (Squeeze-and-Excitation Block) |

KISALTMALAR LİSTESİ (devam)

| | |
|--------------|--|
| SMOTE | : Azınlık sınıf aşırı örnekleme tekniği |
| T1 | : T1-ağırlıklı MRI sekansı |
| VAE | : Varyasyonel otokodlayıcı (Variational Autoencoder) |
| VMD | : Çok hafif demans (Very Mild Demented) |
| ViT | : Görüntü dönüştürücü (Vision Transformer) |
| VRAM | : Grafik işlem belleği |
| WHO | : Dünya Sağlık Örgütü (World Health Organization) |

ÖNSÖZ VE TEŞEKKÜR

Bu tez çalışması, yalnızca akademik bir sürecin değil; aynı zamanda sabır, emek ve kararlılıkla geçen uzun bir yolculuğun ürünüdür. Bu yolculuk boyunca karşılaşılan zorlukların aşılmasında, bilgi kadar anlayışın, destek kadar inancın da ne denli önemli olduğu bir kez daha görülmüştür.

Eğitim hayatım boyunca her koşulda yanımda olan, sevgileri ve güvenleriyle bana güç veren aileme en içten teşekkürlerimi sunarım. Özellikle bu süreçte sabırları, anlayışları ve destekleriyle her zaman yanımda olan ailemin ve kardeşlerimin katkısı, bu çalışmanın tamamlanmasında en büyük motivasyon kaynağım olmuştur. Onların varlığı ve desteği olmadan bu çalışmanın ortaya çıkması mümkün olmazdı.

Akademik gelişimim süresince bilgi ve deneyimleriyle yol gösteren, değerli katkılarını ve desteklerini esirgemeyen danışmanım Doç. Dr. Fatih Aydın'a içten teşekkürlerimi sunarım. Ayrıca çalışmam süresince bilgi ve emekleriyle katkı sağlayan Arş. Gör. A. Hüseyin Ezirmik'e ve Arş. Gör. M. Furkan Ceylan'a destekleri ve yapıcı yönlendirmeleri için teşekkür ederim.

Bunun yanı sıra lisans eğitimimden bu yana akademik gelişimime katkı sunan, desteklerini her zaman hissettiren bölümdeki tüm değerli hocalarıma da teşekkür ederim. Kendilerinin yönlendirmeleri ve yapıcı eleştirileri, bu çalışmanın şekillenmesinde önemli rol oynamıştır.

Ayrıca çalışma hayatım boyunca anlayışları, destekleri ve iş birliği ile süreci kolaylaştıran değerli iş arkadaşlarıma teşekkür ederim. Bu süreçte gösterdikleri motivasyon ve destek, tez çalışmamın ilerlemesine önemli katkılar sağlamıştır.

Bu çalışmanın, ilgili alanda yapılacak akademik çalışmalara katkı sağlamasını ve gelecekte gerçekleştirilecek araştırmalar için yararlı bir kaynak olmasını temenni ederim.

Balıkesir, 2026

Aslıhan Güven

1. GİRİŞ

Nüfusun yaşlanması nedeniyle Alzheimer hastalığı tüm dünyada giderek daha yaygın hale gelmiştir ve günümüzde önemli bir halk sağlığı sorunu oluşturmaktadır (Prince et al., 2015). Dünya Sağlık Örgütü (World Health Organization - WHO), dünya çapında Alzheimer ve diğer demans türlerinden etkilenen milyonlarca insan olduğunu ve bu sayının her yıl arttığını belirtmektedir (Organization, 2021). Alzheimer hastalığı hafızayı, bilişsel yetenekleri, dil becerilerini ve karar verme yeteneğini etkiler. Bu süreç sonucunda beyin nöronlarında ilerleyici hasar meydana gelmektedir (Scheltens et al., 2021). Hastalık ilerledikçe, günlük yaşam aktivitelerini sürdürme kapasiteleri ve bağımsızlıkları azalmakta, bakım gereksinimleri artmaktadır (Dubois et al., 2016).

Alzheimer hastalığı genellikle yavaş başlayan ve sinsi ilerleyen bir klinik tablo göstermektedir. Bu nedenle, erken dönemde belirgin semptomların ortaya çıkmaması tanının çoğu zaman gecikmesine neden olmaktadır. Bununla birlikte, erken teşhis, hastalığın ilerleme hızını yavaşlatmak, mevcut farmakolojik ve non-farmakolojik tedavilerin etkinliğini artırmak ve hastaların ve yakınlarının yaşam kalitesini iyileştirmek için çok önemlidir (Jack et al., 2018). Bu durumda, hastalığın erken belirtilerini bulmak ve nörodejeneratif değişiklikleri detaylı bir şekilde incelemek için güvenilir, tekrarlanabilir ve hassas teknikler gerekmektedir.

Son yıllarda, Alzheimer hastalığının erken tanısı için, özellikle beyin yapısındaki morfolojik değişiklikleri gösterebilen gelişmiş görüntüleme yöntemlerine olan ilgi artmıştır (Westman et al., 2011). Non-invaziv yapısı, yüksek doku kontrastı ve detaylı anatomik çözünürlüğü nedeniyle manyetik rezonans görüntüleme (Magnetic Resonance Imaging - MRI) yaygın olarak klinik ve araştırma alanlarında kullanılmaktadır. MRI görüntüleri, beyin hacmindeki azalma, hipokampal atrofi ve kortikal incelme gibi Alzheimer ile ilişkili yapısal değişiklikleri belirlemede faydalı olmaktadır (Sorensen et al., 2018). Bununla birlikte, bu görüntülerin elle incelenmesi oldukça zahmetli ve uzmanlık gerektiren ve kişiden kişiye değişebilen subjektif bir işlemdir (Sorensen et al., 2018).

Tıbbi görüntü analizinde makine öğrenimi, bu sorunları çözmek ve tanı sürecini hızlı, güvenilir ve objektif hale getirmek için kullanılmaktadır (Litjens et al., 2017). Derin öğrenme, insan beyninin öğrenme yapısını taklit eden çok katmanlı yapay sinir ağları aracılığıyla, büyük boyutlu ve karmaşık veri kümelerinden anlamlı desenler çıkarabilme

yeteneğine sahiptir (Shen et al., 2017). Konvolüsyonel sinir ağları (Convolutional Neural Networks – CNN), özellikle görüntü sınıflandırma, segmentasyon, obje tanıma ve anomali tespiti gibi çok sayıda görevde yüksek başarı göstermektedir. Bu, tıbbi görüntüleme endüstrisini önemli ölçüde etkilemiştir (Litjens et al., 2017; Shen et al., 2017). Bu nedenle, Alzheimer hastalığının otomatik tanısında bir çözüm olarak görülmektedir (Litjens et al., 2017; Shen et al., 2017).

Bu tez çalışması, MRI görüntüleri kullanılarak Alzheimer hastalığını otomatik olarak sınıflandırmak için bir modelleme yöntemi sunmaktadır ve Alzheimer hastalığının teşhisine yönelik sınıflandırma performanslarını karşılaştırmalı olarak değerlendirmek amacıyla beş yaygın derin öğrenme modelini kullanmıştır: EfficientNet-B0, ResNet-50, MobileNetV2, U-Net ve InceptionV3. Bu modellerin her biri kendine özgü mimari yapıları ve hesaplama gereksinimleri ile karakterize edilmiştir. Etiketlenmiş MRI görüntüleri üzerinde her bir model için, doğruluk (Accuracy – ACC), kesinlik (precision), duyarlılık (recall), F1 skoru ve Alıcı İşletim Karakteristiği – Eğri Altında Kalan Alan (Receiver Operating Characteristic – Area Under the Curve, ROC–AUC) gibi standart sınıflandırma ölçütleriyle performansları değerlendirilmiştir. Ayrıca modellerin operasyonel gereksinimleri (örneğin eğitim süresi, parametre sayısı ve hesaplama maliyeti) göz önünde bulundurularak pratik kullanıma uygunluğu değerlendirilmiştir.

Bu çalışmanın temel amacı, Alzheimer hastalığının teşhisine yönelik olarak yaygın biçimde kullanılan derin öğrenme mimarilerinin performanslarını sistematik bir şekilde değerlendirmek ve klinik uygulamalara entegre edilebilecek hızlı, etkili ve güvenilir bir otomatik teşhis yaklaşımına katkı sağlamaktır. Ayrıca çalışma, literatürde sınırlı sayıda ele alınan çoklu model karşılaştırmaları bağlamında kapsamlı bir analiz sunarak, bu alandaki bilgi boşluğunun giderilmesine yönelik önemli bir katkı sağlamayı hedeflemektedir.

1.1 Alzheimer Hastalığının Küresel Önemi ve Erken Teşhisin Rolü

Alzheimer hastalığı, merkezi sinir sistemini etkileyen geri dönüşü olmayan bir nörodejeneratif hastalıktır. Özellikle serebral korteks ve hipokampus bölgeleri etkilenir. Alzheimer hastalığı, ilk kez 1906 yılında Alman nörolog Alois Alzheimer tarafından tanımlanmış olup, günümüzde ileri yaş grubundaki bireylerde en yaygın görülen demans türlerinden biri olarak kabul edilmektedir (Dubois et al., 2016). Dünya Sağlık Örgütü, dünya çapında 55 milyondan fazla insanın Alzheimer veya diğer demans türleriyle yaşadığını

belirtmektedir. Bu sayının 2050 yılına kadar 139 milyona ulaşması beklenmektedir (Organization, 2021). Bu devasa artışın, sağlık sistemlerini ciddi sosyoekonomik sorunlarla karşı karşıya bırakabileceği öngörülmektedir.

Hastalık sinsi bir başlangıçla başlar ve ilk belirtiler genellikle hafıza kaybıyla nüksetmektedir. Daha sonraki evrelerde dil bozuklukları (afazi), mekânsal oryantasyon sorunları, yargı ve karar verme becerilerinde azalma, kişilik değişiklikleri ve psikotik semptomlar ortaya çıkmaktadır. Hastalık ilerledikçe kişi günlük aktivitelerini gerçekleştiremez hale gelmekte ve sürekli bakıma ihtiyaç duymaktadır. Hastanın, bakım verenlerin ve aile üyelerinin fiziksel, psikolojik ve finansal durumları bu süreç tarafından olumsuz etkilenmektedir (Jack et al., 2018).

Alzheimer hastalığının patofizyolojik nedeni beyinde anormal protein birikimleridir. Beta-amiloid plakları ve tau proteinlerinin hiperfosforilasyonu, Alzheimer hastalığını diğer hastalıklardan ayıran biyobelirtiler arasındadır. Bu patolojik süreçler, sinaptik iletimi bozarak nöronlar arası iletişimi zayıflatmakta ve zamanla yaygın nöronal kayıplara neden olmaktadır (Serrano-Pozo et al., 2011). Bununla birlikte, hastalığın oluşum mekanizmaları henüz tam olarak açıklığa kavuşturulamamıştır; genetik yatkınlık (özellikle Alzheimer riskini artırdığı bilinen Apolipoprotein E ϵ 4 aleli (Apolipoprotein E epsilon 4 allele, ApoE4)), yaş, çevresel faktörler, vasküler bozukluklar ve yaşam tarzı etmenleri gibi bir dizi risk faktörünün hastalığın gelişiminde rol oynadığı düşünülmektedir.

Alzheimer hastalığı, hastanın ölümünden sonra beyin dokusunda yapılan postmortem histopatolojik incelemeler yoluyla teşhis edilebilmektedir. Bununla birlikte, klinik olarak bu mümkün olmadığından, tanı süreci genellikle nöropsikolojik testler, hastanın geçmişi ve nörogörüntüleme teknikleri kullanılarak yürütülür. Mini-Mental Durum Testi (Mini-Mental State Examination – MMSE), Montreal Bilişsel Değerlendirme (Montreal Cognitive Assessment – MoCA) gibi bilişsel testlerin yanı sıra; MRI, pozitron emisyon tomografisi (Positron Emission Tomography – PET) ve bilgisayarlı tomografi (BT) tanı için geleneksel teknikler arasında yer almaktadır (Dubois et al., 2016; Jack et al., 2018).

MRI, Alzheimer hastalığını teşhis etmek için yaygın olarak kullanılan non-invaziv bir tekniktir. Yüksek doku kontrastı ve anatomik çözünürlük nedeniyle MRI, hipokampus atrofisi, kortikal incelleme ve ventriküler genişleme gibi Alzheimer ile ilişkili morfolojik

değişiklikleri belirlemede son derece etkilidir. Erken evre tanısında, MRI görüntülerinden hacimsel analizler ve yapay zeka (Artificial Intelligence - AI) tabanlı sınıflandırma yöntemleri önemli avantajlar sağlamaktadır. Manyetik rezonans görüntüleme, PET gibi moleküler görüntüleme yöntemlerine kıyasla iyonlaştırıcı radyasyon içermeyen ve daha düşük maliyetli bir tekniktir. Bu özellikleri sayesinde MRI, klinik ortamlarda Alzheimer hastalığının değerlendirilmesinde daha erişilebilir ve öncelikli bir görüntüleme yöntemi olarak kullanılmaktadır (Litjens et al., 2017).

Bu çerçevede, Alzheimer hastalığı yalnızca yaşlanmanın kaçınılmaz bir sonucu değil, aynı zamanda çok boyutlu klinik bir durumdur ve erken tanı ile yönetilebilmektedir. Erken teşhis, sağlık sisteminin sosyal ve mali yükünü azaltırken bireyin yaşam kalitesini de korumaktadır. Bu nedenle sağlık politikalarının, Alzheimer hastalığının hem klinik hem de teknolojik tanısına yönelik araştırmaları desteklemesi gerekmektedir.

Son yıllarda, ileri nörogörüntüleme teknikleri, biyobelirteç tabanlı yaklaşımlar ve AI destekli analiz yöntemlerinin kullanılmasıyla Alzheimer hastalığının erken evrelerinde teşhis edilmesi mümkün olmuştur (Harper et al., 2016; Sorensen et al., 2018; Westman et al., 2011). Özellikle MRI verilerinin derin öğrenme temelli modellerle birleştirilmesi, hastalığın erken teşhisinde olumlu sonuçlar vermektedir (Litjens et al., 2017; Shen et al., 2017). Bu teknolojik gelişmeler, klinik karar destek sistemlerinin maliyetini, doğruluğunu ve hızını artırmıştır (Shen et al., 2017).

Sonuç olarak, Alzheimer hastalığı sadece nörolojik bir sorun değildir; aynı zamanda toplumun sosyal, ekonomik ve psikolojik yönlerini etkileyen karmaşık bir halk sağlığı sorunudur (Scheltens et al., 2021). Bununla birlikte, erken tanıya yönelik bilimsel çalışmaların desteklenmesi, yeni nesil klinik karar destek sistemlerinin yaygınlaştırılması ve sağlık sistemlerinde AI tabanlı görüntüleme teknolojilerinin kullanılması çok önemlidir (Litjens et al., 2017; Shen et al., 2017). Erken teşhis, bireylerin yaşam kalitesini artırmak ve sağlık sistemlerinin yükünü azaltmak için hayati önem taşımaktadır (Jack et al., 2018; Organization, 2021).

1.2 Alzheimer Tanısında Klinik Yaklaşımlar ve Yapay Zekâ Tabanlı Sistemler

Alzheimer hastalığının klinik tanısı, geleneksel olarak çok aşamalı bir değerlendirme sürecini gerektirmektedir. Bu süreç; hastanın ayrıntılı tıbbi öyküsünün alınması, bilişsel

belirtilerin gözlemlenmesi, nöropsikolojik testlerin uygulanması ve nörogörüntüleme yöntemlerinin analizini kapsayan çok yönlü bir yaklaşımı içermektedir (Dubois et al., 2016). Bilişsel işlevleri değerlendirmeye yönelik olarak yaygın biçimde kullanılan MMSE ve MoCA gibi testler, demansın farklı evrelerinin belirlenmesinde klinik rehberlik sağlasa da erken evre Alzheimer olgularında duyarlılık açısından sınırlılıklar göstermektedir (Jack et al., 2018). Bununla birlikte MRI, PET ve CT gibi nörogörüntüleme teknikleri, beyin yapısındaki yapısal ve işlevsel değişimleri ortaya koyabilse de bu görüntülerin değerlendirilmesi yüksek düzeyde uzmanlık gerektirmektedir ve yorumlayıcıya bağlı öznellik içerebilmektedir (Dubois et al., 2016). Ayrıca, ileri görüntüleme yöntemlerinin yüksek maliyetleri ve sınırlı erişilebilirliği, özellikle düşük ve orta gelirli ülkelerde erken tanı hizmetlerinin yaygınlaşmasının önünde önemli bir engel oluşturmaktadır (Organization, 2021).

Bu sınırlamalar göz önüne alındığında, hızlı, tarafsız, tekrarlanabilir ve klinik olarak uygulanabilir otomatik tanı sistemlerinin ihtiyacı giderek artmaktadır. AI, özellikle derin öğrenme (Deep Learning) temelli yaklaşımlar, Alzheimer ve diğer nörodejeneratif hastalıkların tanısını önemli ölçüde değiştirmiştir. Bu, sağlık alanında giderek daha yaygın hale gelmektedir (Litjens et al., 2017; Shen et al., 2017). Çok katmanlı yapay sinir ağları, derin öğrenme algoritmalarının öngörülerinde bulunma, yapısal ilişkileri modelleme ve çok büyük verilerden karmaşık örüntüleri öğrenme kapasitesine sahiptir.

Bu nedenle, görüntü tabanlı tanı sistemlerinin temeli, CNN'nin yapısal olarak tıbbi görüntüler ile karşılaştırma yeteneğidir. CNN tabanlı modeller, insan gözünün fark edemeyeceği mikro yapısal değişiklikleri belirlemek için MRI görüntülerini piksel düzeyinde analiz edebilmektedir. Ayrıca Alzheimer'ın erken dönemine özgü biyobelirteçleri bulunabilir (Litjens et al., 2017; Suk et al., 2014). Bu özellik, hastalığın teşhisinde ve evrenmesinde ve sağlıklı yaşlanma süreciyle patolojik değişimlerin ayrıştırılmasında son derece faydalı olmaktadır.

Ayrıca bu algoritmalar, yapısal MRI verilerinin yanı sıra Fonksiyonel manyetik rezonans görüntüleme (functional Magnetic Resonance Imaging, fMRI), PET ve difüzyon tensör görüntüleme (Diffusion Tensor Imaging - DTI) gibi çeşitli görüntüleme modalitelerinden elde edilen çok boyutlu verileri entegre edebilir (Bron et al., 2015). Multimodal AI destekli karar destek sistemleri geliştirmek ve nörogörüntüleme çıktılarını daha kapsamlı bir şekilde

değerlendirmek için bu çoklu veri entegrasyonu kullanılabilir. Bu nedenle, yalnızca tanı koymak değil, aynı zamanda hastalığın ilerleyişini tahmin etmek, bireysel risk skorlarını hesaplamak ve tedavi yanıtlarını tahmin etmek de mümkün olacaktır (Bron et al., 2015; Shen et al., 2017).

Kişiselleştirilmiş tıp (precision medicine) yaklaşımı, AI tabanlı sistemlerin önemli bir faydasıdır. Bu sistemler, sadece genel bir tanı sağlamakla kalmaz, aynı zamanda hastaların yaş, cinsiyet, genetik yatkınlık, çevresel faktörler ve klinik geçmiş gibi bireysel değişkenlerini dikkate alarak kişiye özel tedavi ve bakım stratejileri geliştirmelerine izin vermektedir (Suk et al., 2014). Bu nedenle Alzheimer hastalığını biyolojik alt tiplerine göre sınıflandırmak ve her birey için en iyi bakım planlarını oluşturmak mümkün hale gelmektedir.

Bu yeniliklere rağmen, AI tabanlı otomatik tanı sistemlerinin klinik uygulamalara dahil edilmesi ciddi sorunlara neden olabilir. Modellerin açıklanabilirliği (explainability), etik ve yasal düzenlemeler, karar süreçlerinin şeffaflığı ve kullanılan verilerin kalitesi tartışmalı konulardır (Litjens et al., 2017). Ayrıca, algoritmaların genellenebilirliği, daha kullanıcı dostu arayüzler oluşturmak ve çok merkezli veri kümelerinde doğrulama yapmak gibi teknik zorluklar vardır (Shen et al., 2017).

Bununla birlikte, derin öğrenme modellerinin doğruluğu ve güvenilirliği, Alzheimer Hastalığı Nörogörüntüleme Girişimi (Alzheimer's Disease Neuroimaging Initiative - ADNI), Açık Erişimli Görüntüleme Serileri (Open Access Series of Imaging Studies – OASIS) ve Avustralya Görüntüleme, Biyobelirteç ve Yaşam Tarzı Çalışması (Australian Imaging, Biomarkers and Lifestyle Study of Ageing - AIBL) gibi açık erişimli nörogörüntüleme veri kümelerinin yaygınlaşmasıyla arttı (Harper et al., 2016; Westman et al., 2011).

Bu bulgular, yakın gelecekte AI destekli otomatik tanı sistemlerinin klinik uygulamalarda daha popüler olacağını göstermektedir. Bu sistemler, tanı doğruluğunu ve iş yükünü azaltır, tanı maliyetlerini düşürür ve tanı süresini kısaltır. Bu nedenle, nörodejeneratif hastalıkların yönetimi için bu sistemler çok önemlidir (Shen et al., 2017).

Bu bağlamda, Alzheimer hastalığının erken tanı ve evreleme süreçlerinde kullanılan nörogörüntüleme yaklaşımlarının ve özellikle manyetik rezonans görüntüleme (MRI) yönteminin rolünün ayrıntılı olarak ele alınması büyük önem taşımaktadır.

1.2.1 Beyin görüntüleme teknikleri ve MRI'nın rolü

Nörogörüntüleme, Alzheimer hastalığının erken tanısı ve evrenmesinde giderek daha önemli hale gelmektedir. Beynin yapısal ve fonksiyonel değişimlerini görsel ve nicel olarak izlemek için bu yöntemler klinik değerlendirme süreçlerini desteklemektedir. Günümüzde Alzheimer tanısında en yaygın kullanılan görüntüleme yöntemlerinden bazıları MRI, PET, BT ve fMRI'dir (Jack et al., 2018). Bununla birlikte, üstün özellikleri nedeniyle MRI, hem klinik uygulamalarda hem de AI tabanlı tanı sistemlerinde en çok tercih edilen yöntem olarak öne çıkmaktadır.

MRI görüntüleme yöntemi, radyasyon içermeyen, güvenli, tekrarlanabilir ve non-invazivdir. Vücuttaki hidrojen protonlarının güçlü manyetik alan altında rezonansa girmesi prensibine dayanmaktadır. Yüksek çözünürlüklü MRI görüntüleri, Alzheimer hastalığında hipokampal atrofi, kortikal incelme, ventriküler genişleme ve global beyin atrofisi gibi hastalığa özgü yapısal değişiklikleri tanımlayabilmektedir. Hipokampus ve medial temporal lob bölgelerinde meydana gelen hacim kaybı, özellikle hastalığın başlangıcında MRI aracılığıyla son derece hassas bir şekilde ölçülebilmektedir (Dubois et al., 2016; Frisoni et al., 2010).

MRI'nın klinik açıdan sunduğu temel avantajlar aşağıda özetlenmiştir:

- Yüksek çözünürlük: Beynin yapısal detaylarını yüksek uzaysal çözünürlükte görüntüleme kapasitesine sahiptir.
- Non-invaziv ve radyasyonsuz yapı: Tekrarlanabilir ölçümler, hastaya zarar vermeden yapılabilir.
- Yüksek doku kontrastı: Gri ve beyaz cevher ayrımı ile hipokampal yapıların net biçimde incelenmesine imkân tanır.
- Çoklu görüntüleme düzlemi: Aksiyel, koronal ve sagittal düzlemlerde görüntüleme olanağı sunar.
- AI entegrasyonuna uygunluk: Yüksek veri kalitesi sayesinde otomatik sınıflandırma ve segmentasyon modellerinin performansını artırır.

Bu özellikler, MRI verilerini yalnızca radyologlar tarafından manuel olarak yorumlanan görüntüler olmaktan çıkararak, AI destekli sistemler için zengin bir bilgi kaynağı hâline getirmiştir. Özellikle CNN gibi derin öğrenme tabanlı modeller, MRI görüntülerinden anlamlı desenleri otomatik olarak öğrenebilmekte ve Alzheimer hastalığına özgü morfolojik bozulmaları tespit edebilmektedir. Bilgisayar destekli tanı sistemlerinin doğruluğu, MRI verilerinin yapısal ve metrik analizine uygunluğuna bağlıdır. Hipokampal hacim ölçümler, klinik değerlendirmelerle karşılaştırıldığında daha hassas sonuçlar verebilir. Bu ölçümler, Alzheimer hastalığının erken evrelerinde tanısal hassasiyet sağlayabilmekte ve hastalığın ilerleyişinin izlenmesinde önemli rol oynamaktadır (Jack et al., 2018).

Otomatik teşhis algoritmaları oluşturmak isteyen araştırmacılar, radyasyonsuz yapısı, yüksek veri kalitesi ve yaygın klinik kullanımı nedeniyle MRI'yi tercih etmektedir. Bu nedenle, tezde kullanılan tüm derin öğrenme mimarileri T1-ağırlıklı manyetik rezonans görüntüleme sekansı (T1-weighted Magnetic Resonance Imaging sequence – T1) üzerinde eğitilmiştir.

Sonuç olarak, MRI yöntemi, otomatik sınıflandırma sistemlerinin başarısı açısından kritik bir rol oynamakta ve Alzheimer hastalığının erken tanısında temel görüntüleme aracı olarak öne çıkmaktadır. Gelecekte multimodal sistemlerle (örneğin PET-MRI veya fMRI entegrasyonları) birleştirilme potansiyeli yüksek olsa da, günümüzde derin öğrenme temelli tanı modelleri için temel ve güvenilir veri kaynağı olma özelliğini sürdürmektedir.

1.3 Problem Tanımı

Alzheimer hastalığının erken teşhisi, hastalığın ilerlemesini durdurmak ve yaşam kalitesini iyileştirmek için çok önemlidir. Bu amaca hizmet eden otomatik tanı sistemleri, özellikle AI ve derin öğrenme (Deep Learning) alanındaki son yıllardaki hızlı gelişmeler nedeniyle geliştirilmektedir. İnsan müdahalesi olmaksızın beynin morfolojik yapısında meydana gelen ince morfolojik değişiklikler, derin öğrenme tabanlı modeller tarafından öğrenilebilir ve sınıflandırılabilir (Litjens et al., 2017; Shen et al., 2017).

Son yıllarda derin öğrenme algoritmalarının tıbbi görüntüleme alanındaki hızlı yükselişi, Alzheimer teşhisine yönelik tanı modellerinin doğruluk ve güvenilirlik düzeyini önemli ölçüde artırmıştır (Bron et al., 2015; Suk et al., 2014).

Bununla birlikte, derin öğrenme mimarilerini çeşitli şekillerde karşılaştıran kapsamlı araştırmalar için mevcut literatürde çok az çalışma bulunmaktadır. Çok sayıda çalışma yalnızca bir veya iki model üzerinde odaklanır ve analizlerini genellikle tek bir performans metriği veya ACC kullanarak yapar. Modellerin klinik olarak faydalı olup olmadığını belirlemek için bu teknik klinik değerlendirme açısından yeterli görülmemektedir. İşlem süresi, taşınabilirlik, donanım gereksinimleri ve enerji tüketimi gibi operasyonel unsurlar göz ardı edilmektedir (He et al., 2016; Tan and Le, 2019).

Bu nedenle, Alzheimer hastalığının otomatik teşhisine yönelik olarak literatürde birçok farklı derin öğrenme mimarisi önerilmiştir ve bazıları başarıyla uygulanmıştır. Özellikle EfficientNet, ResNet, InceptionV3, MobileNetV2 ve U-Net gibi mimariler, MRI görüntülerini kullanarak Alzheimer evrelerini sınıflandırma görevlerinde başarılı olmuştur. Katman derinliği, parametre yoğunluğu, eğitim stratejisi ve yapısal özellikleri dahil olmak üzere her mimari farklı avantajlar sunar. Bununla birlikte, bu çeşitlilik, literatürde henüz net bir şekilde yanıtlanmamış temel bir soruyu da gündeme getirdi: "Hangi derin öğrenme mimarisi, Alzheimer tanısında en etkili, genellenebilir ve klinik olarak uygulanabilir çözümdür?"

Mevcut akademik çalışmalar genellikle yalnızca bir veya iki derin öğrenme mimarisini karşılaştırmakta ve değerlendirmeler genellikle tek bir metrik (örneğin ACC) kullanılarak yapılmaktadır. Öte yandan, klinik ortamda tıbbi karar destek sistemlerinin kullanılabilmesi için yalnızca yüksek doğruluk oranı yeterli değildir. Değerlendirme sürecine modelin eğitim süresi, parametre sayısı, donanım (Merkezi İşlem Birimi – Central Processing Unit, CPU / Grafik İşlem Birimi – Graphics Processing Unit, GPU) gereksinimi, bellek kullanımı, enerji tüketimi, gerçek zamanlı çalışma performansı, taşınabilirlik ve genellenebilirlik gibi operasyonel kriterler de dahil edilmelidir (Bron et al., 2015; Shen et al., 2017).

Örneğin, yüksek doğruluk oranı sağlayan bir modelin parametre yoğunluğu nedeniyle taşınabilir sistemlerde çalıştırılmaması veya çok uzun eğitim süresi gerektirmesi nedeniyle gerçek klinik uygulamalarda kullanımı ciddi şekilde kısıtlı olmaktadır. Bununla birlikte, düşük parametre sayısına sahip ancak düşük duyarlılık (recall/sensitivity) gösteren bir model, özellikle Alzheimer'ın erken evrelerinde yeterli tanısal güvenilirlik sağlamaz. Bu nedenle, tek boyutlu metriklerle yapılan değerlendirmeler, tıbbi AI uygulamalarının gerçek potansiyelini göstermemektedir. Bu, literatürde çok kriterli değerlendirmenin (multi-criteria

evaluation) yeterince ele alınmadığını ve önemli bir metodolojik boşluk olduğunu göstermektedir.

Bu bağlamda, bu tez çalışmasının ele aldığı temel sorun, Alzheimer hastalığının MRI görüntülerinden otomatik olarak teşhis edilmesine yönelik farklı derin öğrenme mimarilerinin sınıflandırma başarısı ve çok yönlü teknik, operasyonel ve klinik kriterler açısından kapsamlı bir şekilde karşılaştırılmasıdır.

1.4 Araştırmanın Amacı

Bu tez çalışmasının temel amacı, yüksek doğruluk oranına sahip ve klinik uygulamalara entegre edilebilecek en uygun derin öğrenme mimarisini belirlemektir. Bu kapsamda, Alzheimer hastalığının MRI verileri üzerinden otomatik teşhisini gerçekleştirebilecek en uygun mimarinin belirlenmesi hedeflenmektedir. Alzheimer hastalığının ilerleyici ve geri dönüşsüz seyri nedeniyle erken teşhis çok önemlidir. Erken tanı, sağlık sistemleri üzerindeki mali ve operasyonel yükleri azaltırken, bireylerin yaşam kalitesini de koruyabilir. Bu bağlamda, otomatik sınıflandırma sistemleri, makine öğrenimi teknikleri aracılığıyla geliştirilmektedir ve geleneksel tanı yöntemlerine alternatif olarak öne çıkmaktadır (Litjens et al., 2017; Shen et al., 2017).

1.4.1 Araştırma hedefleri

Bu çalışmanın temel amacı, Alzheimer hastalığının tanısında farklı derin öğrenme mimarilerinin performansını çok boyutlu olarak değerlendirmek ve karşılaştırmaktır. Bu kapsamda, EfficientNet-B0, ResNet-50, MobileNetV2, U-Net ve InceptionV3 mimarilerinin Alzheimer tanısındaki başarıları kapsamlı biçimde incelenmektedir. Her bir model; doğruluk (accuracy), kesinlik (precision), duyarlılık (recall), özgüllük (specificity), F1 skoru ve ROC-AUC gibi temel performans metrikleri kullanılarak nicel olarak değerlendirilmiş, ayrıca modellerin sınıflandırma performansını etkileyen yapısal ve istatistiksel unsurlar ayrıntılı biçimde analiz edilmiştir.

Çalışmanın teknik ve operasyonel hedefleri doğrultusunda, derin öğrenme modellerinin yalnızca sınıflandırma başarısı değil, aynı zamanda pratik kullanım açısından kritik olan teknik özellikleri de ele alınmıştır. Bu bağlamda, eğitim süresi, işlem verimliliği, parametre sayısı ve hesaplama karmaşıklığı gibi performans ölçütleri analiz edilmiş; donanım bağımlılığı, model dosya boyutu ve bellek gereksinimi gibi sistem kaynaklarını etkileyen

faktörler ayrıntılı olarak incelenmiştir. Bunun yanı sıra, modellerin taşınabilirliği, gerçek zamanlı çalışma performansı, enerji verimliliği ve klinik ortamlara entegrasyon potansiyeli karşılaştırılmış; veri artırma tekniklerinin etkisi ve aşırı öğrenme (overfitting) eğilimi de değerlendirilmiştir.

Araştırmanın akademik, uygulamaya yönelik ve teknolojik katkı hedefleri de çalışmanın önemli bir bileşenini oluşturmaktadır. Akademik açıdan, Alzheimer tanısına yönelik derin öğrenme tabanlı modellerin karşılaştırmalı analizine sistematik ve metodolojik bir çerçeve sunulması amaçlanmıştır. Uygulama açısından, klinisyenlerin ve sağlık profesyonellerinin gerçek klinik koşullarda etkin biçimde çalışabilecek yapay zekâ çözümlerini belirlemelerine katkı sağlanması hedeflenmiştir. Teknolojik açıdan ise, verimlilik, sürdürülebilirlik ve taşınabilirlik kriterlerini gözeterek yapay zekâ destekli tanı sistemleri için somut ve karşılaştırılabilir sonuçların ortaya konulması amaçlanmıştır.

Sonuç olarak, bu tez, yalnızca en yüksek doğruluk oranına sahip modeli bulmakla kalmaz, aynı zamanda farklı derin öğrenme mimarilerinin avantajları ve dezavantajları arasındaki dengeyi net bir şekilde ortaya koyarak gelecekte geliştirilecek klinik karar destek sistemlerine yön vermeyi amaçlamaktadır. Elde edilen sonuçlar, AI tabanlı tıbbi görüntüleme uygulamalarının ne kadar etkili olduğunu ve klinik olarak ne kadar uygun olduğunu gösterecektir.

1.5 Literatür Boşlukları ve Bu Tezin Katkısı

Bu tez çalışması, Alzheimer hastalığının otomatik teşhisi için derin öğrenme modellerini içeren mevcut literatürü kapsamlı bir şekilde incelemiştir. Ayrıca, bu alanda meydana gelen önemli boşlukları açıkça göstermiştir. Çalışmanın amacı, hem metodolojik hem de uygulamalı olarak literatürdeki eksiklikleri gidermek ve literatüre özgün ve sistematik katkılar sunmaktır.

1.5.1 Literatürdeki başlıca boşluklar

Mevcut literatürde Alzheimer hastalığının tanısına yönelik derin öğrenme tabanlı çalışmalar incelendiğinde, değerlendirme kriterlerinin çoğu zaman sınırlı tutulduğu görülmektedir. Birçok çalışma yalnızca sınıflandırma performansına odaklanarak doğruluk (accuracy), duyarlılık (sensitivity/recall), özgüllük (specificity) ve F1-skoru gibi metrikleri raporlamakta; buna karşın model karmaşıklığı, eğitim süresi, parametre yoğunluğu,

hesaplama maliyeti ve donanım gereksinimleri gibi pratik açıdan kritik öneme sahip ölçütler çoğu zaman göz ardı edilmektedir. Bu durum, yüksek doğruluk oranlarına sahip modellerin dahi gerçek klinik ortamlarda uygulanabilirliğini sınırlayan önemli bir faktör olarak öne çıkmaktadır.

Literatürde dikkat çeken bir diğer önemli boşluk, modellerin karşılaştırılmasında standartlaştırılmış deney koşullarının yeterince sağlanamamasıdır. Çok sayıda çalışma, model kıyaslaması için farklı veri setleri, ön işleme adımları ve deney protokolleri kullandığından, elde edilen sonuçların doğrudan ve adil biçimde karşılaştırılması zorlaşmaktadır. Özellikle birden fazla derin öğrenme mimarisinin aynı veri kümesi üzerinde, aynı ön işleme stratejileri ve sabit hiperparametre ayarları ile değerlendirilmesi literatürde nadiren ele alınmıştır.

Bunun yanı sıra, açıklanabilirlik konusu literatürde önemli bir eksiklik alanı olarak dikkat çekmektedir. Açıklanabilir Yapay Zekâ (Explainable Artificial Intelligence – XAI) yöntemleri, modellerin karar verme süreçlerinin görselleştirilmesi ve yorumlanması açısından kritik öneme sahip olmasına rağmen, mevcut çalışmaların büyük bir bölümünde bu yaklaşımlar yeterince ele alınmamıştır. Bu eksiklik, özellikle klinik uygulamalarda model çıktılarının güvenilirliğini ve kullanıcıların (klinik uzmanların) modele duyduğu güveni önemli ölçüde sınırlandırmaktadır.

Son olarak, literatürdeki çalışmaların büyük çoğunluğunun laboratuvar ortamlarında gerçekleştirilmiş olması, klinik senaryolara uygunluk değerlendirmelerinin sınırlı kalmasına yol açmaktadır. Enerji verimliliği, taşınabilirlik ve edge cihazlarda çalışabilirlik gibi pratik faktörler çoğu zaman dikkate alınmamakta; bu da geliştirilen sistemlerin gerçek klinik entegrasyon sürecinde karşılaşılabileceği operasyonel ve teknik zorlukların yeterince değerlendirilmemesine neden olmaktadır.

1.5.2 Bu tezin katkıları

Bu tez kapsamında, Alzheimer hastalığının tanısına yönelik derin öğrenme tabanlı yaklaşımların karşılaştırmalı analizi gerçekleştirilmiş ve literatüre çok boyutlu katkılar sunulmuştur.

Çalışmada EfficientNet, ResNet-50, U-Net, MobileNetV2 ve InceptionV3 olmak üzere beş farklı derin öğrenme mimarisi; aynı MRI veri kümesi (OASIS), aynı ön işleme adımları ve sabit hiperparametre ayarları kullanılarak standartlaştırılmış bir deneysel çerçeve altında değerlendirilmiştir. Bu yaklaşım sayesinde, elde edilen sonuçların doğrudan karşılaştırılabilirliği sağlanmış ve deneysel sürecin iç geçerliliği güçlendirilmiştir.

Tezin önemli katkılarından biri, modellerin yalnızca sınıflandırma doğruluğu açısından değil, klinik uygulanabilirlik perspektifinden de kapsamlı biçimde değerlendirilmiş olmasıdır. Bu bağlamda, eğitim süresi, parametre sayısı, model dosya boyutu, çıkarım süresi (inference time) ve kaynak tüketimi gibi çok boyutlu teknik ölçütler dikkate alınarak detaylı bir analiz gerçekleştirilmiştir. Böylece, geliştirilen modellerin gerçek klinik ortamlarda entegrasyon potansiyeli daha gerçekçi ve bütüncül bir bakış açısıyla ele alınabilmektedir.

Çalışmada ayrıca, veri azlığı problemine karşı transfer learning yaklaşımları, veri artırma teknikleri (döndürme, yansıtma, yakınlaştırma vb.) ve ağırlıklandırılmış kayıp fonksiyonları birlikte kullanılmıştır. Bu stratejiler, sınırlı veri koşullarında model performansını artırmaya yönelik uygulanabilir ve etkili çözümler sunarak, tıbbi görüntüleme alanında sık karşılaşılan veri kısıtı problemine pratik bir yaklaşım getirmiştir.

Tezin bir diğer önemli katkısı, modellerin karar verme süreçlerinin açıklanabilirliğinin incelenmesidir. Bu kapsamda, Gradyan Ağırlıklı Sınıf Aktivasyon Haritaları (Gradient-weighted Class Activation Mapping – Grad-CAM) tabanlı açıklanabilir yapay zekâ yöntemleri kullanılarak, modellerin sınıflandırma sırasında odaklandığı beyin bölgeleri görselleştirilmiştir. Elde edilen görseller üzerinden, modellerin hipokampus ve entorhinal korteks gibi Alzheimer hastalığı ile ilişkili biyolojik olarak anlamlı anatomik yapılara odaklanıp odaklanmadığı analiz edilmiştir. Bu yaklaşım, model şeffaflığını artırmış ve klinik yorumlanabilirliğe önemli katkı sağlamıştır.

Son olarak, geliştirilen modeller gerçek yaşam senaryolarına uygunluk açısından da değerlendirilmiştir. Sürdürülebilirlik, enerji tüketimi, çalışma hızı, mobilite ve donanım gereksinimleri gibi pratik parametreler dikkate alınarak yapılan analizler, bu modellerin klinik kullanım ve edge computing altyapılarında nasıl değerlendirilebileceğine ilişkin önemli çıkarımlar sunmuştur. Bu yönüyle tez, yalnızca akademik değil, aynı zamanda uygulama ve teknoloji odaklı somut katkılar da ortaya koymaktadır.

1.6 Araştırma Soruları ve Hipotezler

Bu çalışma kapsamında, Alzheimer hastalığının manyetik rezonans görüntüleri (MRI) üzerinden sınıflandırılmasında farklı derin öğrenme mimarilerinin performanslarının karşılaştırılması amaçlanmaktadır.

Bu doğrultuda, mimarilerin doğruluk, genelleme kapasitesi, hesaplama verimliliği ve klinik uygulanabilirlik açısından farklılık gösterip göstermediği sistematik biçimde incelenmiştir.

Çalışmanın temel araştırma soruları ve bu sorulara dayalı olarak geliştirilen hipotezler aşağıda sunulmuştur.

1.6.1 Araştırma soruları

- **S1:** Hangi derin öğrenme mimarisi, Alzheimer hastalığının MRI görüntüleri üzerinden sınıflandırılmasında en yüksek sınıflandırma başarımını sağlamaktadır?
- **S2:** Yalnızca doğruluk oranı değil; model karmaşıklığı, parametre yoğunluğu, eğitim süresi ve hesaplama maliyeti gibi pratik faktörler de dikkate alındığında, klinik ortamlarda uygulanabilirlik açısından en uygun mimari hangisidir?
- **S3:** EfficientNet mimarisi, sınıflandırma performansı açısından diğer derin öğrenme mimarileriyle karşılaştırıldığında istatistiksel olarak anlamlı bir üstünlük göstermekte midir (Tan and Le, 2019)?
- **S4:** U-Net ve MobileNetV2 gibi görece daha “hafif” mimariler, eğitim süresi, parametre sayısı ve verimlilik açısından diğer mimarilere göre daha avantajlı çözümler sunmakta mıdır (Sandler et al., 2018)?
- **S5:** Çoklu performans metriklerinin (ACC, precision, recall, F1, AUC) birlikte değerlendirildiği çok kriterli analiz bağlamında, genel başarı dengesi açısından tek bir mimari ön plana çıkmakta mıdır, yoksa kullanım amacına göre farklı senaryolar mı öne çıkmaktadır?

1.6.2 Araştırma hipotezleri

- **H1 – S1’e Karşılık Gelen Hipotez:**
 - **H0 (Null Hipotez):** Derin öğrenme mimarileri arasında Alzheimer sınıflandırma başarımı açısından istatistiksel olarak anlamlı bir fark yoktur.
 - **H1 (Alternatif Hipotez):** Derin öğrenme mimarileri arasında Alzheimer sınıflandırma başarımı açısından istatistiksel olarak anlamlı bir fark vardır.

- **H2 – S2’ye Karşılık Gelen Hipotez:**
 - **H0:** Modeller arasında model karmaşıklığı, parametre yoğunluğu, eğitim süresi ve hesaplama maliyeti açısından istatistiksel olarak anlamlı bir fark yoktur.
 - **H1:** Modeller arasında model karmaşıklığı, parametre yoğunluğu, eğitim süresi ve hesaplama maliyeti açısından istatistiksel olarak anlamlı fark vardır.
- **H3 – S3’e Karşılık Gelen Hipotez:**
 - **H0:** EfficientNet-B0 mimarisinin sınıflandırma performansı diğer mimarilerden istatistiksel olarak farklı değildir.
 - **H1:** EfficientNet-B0 mimarisinin sınıflandırma performansı diğer mimarilerden istatistiksel olarak anlamlı derecede farklıdır.
- **H4 – S4’e Karşılık Gelen Hipotez:**
 - **H0:** U-Net ve MobileNetV2 mimarileri, eğitim süresi, parametre yoğunluğu ve verimlilik açısından diğer mimarilerden istatistiksel olarak farklılık göstermemektedir.
 - **H1:** U-Net ve MobileNetV2 mimarileri, eğitim süresi, parametre yoğunluğu ve verimlilik açısından diğer mimarilerden istatistiksel olarak anlamlı şekilde farklılık göstermektedir.
- **H5 – S5’e Karşılık Gelen Hipotez:**
 - **H0:** Çoklu performans metrikleri ve operasyonel kriterler birlikte değerlendirildiğinde derin öğrenme mimarileri arasında genel başarı dengesi bakımından anlamlı bir fark yoktur.
 - **H1:** Çoklu performans metrikleri ve operasyonel kriterler birlikte değerlendirildiğinde derin öğrenme mimarileri arasında genel başarı dengesi bakımından anlamlı bir fark vardır.

Bu hipotezlerin test edilmesiyle birlikte, farklı mimariler arasındaki performans farklarının istatistiksel olarak anlamlı olup olmadığı belirlenmiştir; ayrıca klinik uygulamalara yönelik model seçimi açısından somut çıkarımlar elde edilmiştir.

Araştırma kapsamında gerçekleştirilecek istatistiksel hipotez testleri (örneğin Kruskal–Wallis gibi parametrik olmayan testler veya gerekli durumlarda tek yönlü varyans analizi (One-Way Analysis of Variance - ANOVA) ve post-hoc karşılaştırmalar), hangi mimarinin

hangi ölçütlerde anlamlı üstünlük sağladığını ortaya koyarak çalışmanın analitik doğruluğunu güçlendirmektedir.

1.7 Tezin Kapsamı ve Sınırlılıkları

Bu tez çalışması, MRI verilerini kullanarak Alzheimer hastalığını otomatik sınıflandırmaya odaklanmaktadır. Çalışmada yapısal MRI (T1-ağırlıklı) görüntüleri kullanılmıştır. Diğer görüntüleme modaliteleri, örneğin PET, fMRI veya DTI, değerlendirmeye dahil edilmemiştir. Bu çalışmada, literatürde sıklıkla tercih edilen açık erişimli OASIS veri kümesi kullanılmıştır. OASIS veri seti, yaşlı insanların bilişsel durumlarına göre sınıflandırılmış MRI görüntülerini içerir. Bu veri seti, Alzheimer hastalığının çeşitli evrelerine ilişkin sınıflandırma görevlerini içermektedir. Bununla birlikte, OASIS veri seti, gerçek zamanlı klinik veriler veya çok merkezli veri toplulukları içermediğinden, sonuçların genellenebilirliği sınırlıdır (Marcus et al., 2007).

Bu çalışmada yalnızca görüntü tabanlı veriler kullanılmıştır; bireylerin yaş, cinsiyet, eğitim düzeyi, genetik yatkınlık (örneğin ApoE4 alleli) ve eşlik eden hastalıklar gibi biyomedikal veya demografik değişkenler modelleme sürecine dâhil edilmemiştir. Metodolojik olarak, bu durum, çalışmanın kapsamını yalnızca görüntü tabanlı teşhis sistemlerine sınırlamak için tercih edilen bir seçenektir.

Derin öğrenme modelleri, önceden eğitilmiş ağırlıklar (pre-trained weights) kullanılarak transfer öğrenme (transfer learning) yaklaşımıyla eğitilmiştir. Eğitim sürecinde veri artırma (data augmentation), yeniden boyutlandırma ve normalizasyon gibi temel ön işleme adımları uygulanmıştır. Ancak segmentasyon sonrası anatomik yapıların çıkarımı, hipokampus hacim ölçümleri, kortikal kalınlık analizi gibi ayrıntılı anatomik incelemeler bu çalışmanın kapsamı dışında tutulmuştur. Benzer biçimde, multimodal analiz yöntemleri (örneğin MRI + PET entegrasyonu) de bu çalışma kapsamında ele alınmamıştır.

Araştırmanın, modellerin yalnızca görüntü sınıflandırma görevleri üzerinden incelendiği gerçeği, diğer bir kısıtlamadır. Çalışma, görüntü segmentasyonu, belirli beyin bölgelerinin lokalizasyonu veya hastalık ilerlemesinin zamansal tahmini gibi ileri düzey analizleri içermemektedir. Bu kısıtlama, araştırmanın odak noktasının derin öğrenme mimarilerinin sınıflandırma performanslarını karşılaştırmak üzerine kurmasından kaynaklanmaktadır.

Bu tez ayrıca tüm modellerin tek bir veri kümesi üzerinde test edilmesi gerektiğini belirtmektedir. Gelecekte modellerin genellenebilirliği, farklı hasta grupları veya çok merkezli veri setlerinde yapılan deneylerle önemli ölçüde artırılabilir. Ek olarak, hiperparametre optimizasyonu temel düzeyde tutulmuştur; her mimarinin olası tüm parametre kombinasyonları üzerinde kapsamlı bir inceleme yapılmamıştır. Araştırmanın karşılaştırmalı analiz boyutunu korumak ve metodolojik tutarlılığı sağlamak amacıyla bu tercih yapılmıştır.

Gelecekte yapılacak daha kapsamlı, çok modaliteli ve klinik doğrulamalı araştırmaları yönlendirmek için bulgular, söz konusu sınırlamalar dikkate alınarak düzenlenmiştir.

1.8 Tezin Organizasyonu

Bu tez çalışması yedi bölümden oluşmaktadır. Her bölüm, araştırmanın kuramsal temellerini, deneysel süreci ve sonuçları sunmak için düzenlenmiştir. Bölümler mantıksal bir bütünlük içinde kurgulanmıştır.

Birinci bölümde, Alzheimer hastalığına ilişkin temel kavramlar, erken teşhis yaklaşımları, klinik yöntemler ve AI tabanlı sistemler ele alınmıştır. Araştırmanın soruları, amacı, hipotezleri, kapsamı ve sınırlılıkları da açıklanmıştır.

İkinci bölümde, Alzheimer hastalığını teşhis etmek için derin öğrenme tekniklerine ilişkin literatür taraması sunulmuştur. Bu çalışmada derin öğrenme mimarileri, beyin görüntüleme teknikleri, klinik uygulamalar, literatürdeki araştırma boşlukları ve önceki çalışmalar incelenmiştir.

Üçüncü bölümde, araştırmada kullanılan veri kümesi (OASIS), veri ön işleme süreçleri, model eğitimi, derin öğrenme mimarileri, performans metrikleri ve deneysel yöntem hakkında ayrıntılı bilgi verilmiştir. Her derin öğrenme modeli için yapılan deney sonuçları dördüncü bölümde gösterilmektedir. Modellerin sınıflandırma performansları, eğitim süreçleri, grafiksel analizler, sınıf dengesine göre sonuçları ve teknik ve klinik değerlendirmeler sistematik bir şekilde karşılaştırılmıştır.

Beşinci bölümde, modellerden elde edilen bulgular tartışılmış; sınıflandırma başarımı, klinik uygulanabilirlik, sınırlılıklar ve gözlemler bütüncül biçimde değerlendirilmiştir.

Altıncı bölüm, araştırmanın bulgularını ve temel bilimsel faydalarını özetlemektedir.

Yedinci ve son bölümde, araştırmanın sınırlılıkları doğrultusunda gelecekte yapılabilecek çalışmalara yönelik öneriler sunulmuştur.

Bu yapısal düzenlemenin amacı, okuyucunun çalışmanın amacını, kapsamını, yöntemlerini ve elde edilen sonuçları hakkında kapsamlı bir anlayış kazanmasını sağlamaktır.

2. LİTERATÜR TARAMASI

Bu bölümde, Alzheimer hastalığının teşhisinde kullanılan derin öğrenme temelli yöntemlere ilişkin güncel bilimsel çalışmalar kapsamlı biçimde incelenmektedir. Öncelikle hastalığın nörobiyolojik temelleri ve nörogörüntüleme tekniklerinin literatürdeki yeri özetlenmiş; ardından derin öğrenme mimarilerinin tıbbi görüntüleme alanındaki kullanımına yönelik çalışmalar değerlendirilmiştir. Bu çerçevede EfficientNet, ResNet, U-Net, MobileNetV2 ve Inception gibi yaygın modellerin avantajları, sınırlılıkları ve klinik uygulamalara yönelik performansları literatür temelli bir bakış açısıyla ele alınmaktadır.

2.1 Derin Öğrenme Yöntemlerinin Tıpta Kullanımı

Son yıllarda AI ve özellikle derin öğrenme (Deep Learning) algoritmalarındaki hızlı gelişmeler, tıp alanında tanı ve karar destek sistemleri açısından devrim niteliğinde yenilikler ortaya koymuştur. Bu dönüşümün merkezinde yer alan CNN mimarileri, tıbbi görüntü analizi görevlerinde yüksek doğruluk, hassasiyet ve otomasyon kapasitesi ile öne çıkmaktadır. CNN'ler, görüntülerdeki uzamsal örüntüleri otomatik biçimde öğrenebilme yetenekleri sayesinde, insan uzmanlarla karşılaştırılabilecek — hatta bazı durumlarda onları aşabilecek — düzeyde performans sergileyebilmektedir (Litjens et al., 2017; Shen et al., 2017).

Geleneksel makine öğrenmesi yaklaşımları, el ile çıkarılan özniteliklere (hand-crafted features) dayandığından, uzman bilgi birikimine bağımlı ve ölçeklenebilirlik açısından sınırlıdır. Gri-seviye histogramları, doku analizi yöntemleri, Gri-Seviye Eş-Oluşum Matrisi (Gray-Level Co-occurrence Matrix – GLCM) tabanlı öznitelikler veya şekil tanımlayıcıları gibi özelliklerin manuel olarak tanımlanması gerekir ve bu durum hem zaman alıcıdır hem de model performansını doğrudan kullanılan özniteliklerin kalitesi belirlemektedir. Buna karşın derin öğrenme modelleri, öznitelik çıkarımı ve sınıflandırma süreçlerini tek bir uçtan uca öğrenme yapısı içinde birleştirerek otomatik özellik öğrenimi sağlar. Literatürde CNN tabanlı modellerin, klasik yöntemlere göre Alzheimer sınıflandırma görevlerinde daha yüksek doğruluk ve AUC değerlerine ulaştığı birçok çalışmada gösterilmiştir (Lundervold and Lundervold, 2019; Wen et al., 2020).

Alzheimer hastalığının heterojen yapısı nedeniyle tek bir görüntüleme tekniğine bağlı kalmak çoğu zaman tanısal doğruluğu sınırlamaktadır. Bu nedenle son yıllarda MRI

verilerinin PET, fMRI, beyin omurilik sıvısı (CSF) biyobelirteçleri ve nöropsikolojik test skorları gibi çoklu veri kaynaklarıyla birleştirildiği multimodal derin öğrenme yaklaşımlarının kullanımı artmıştır. Multimodal modeller, her veri tipini ayrı bir derin öğrenme kolunda işleyerek daha sonra ortak bir temsilde birleştirmekte ve böylece hem yapısal hem de fonksiyonel değişiklikleri aynı anda modelleyebilmektedir (Vieira et al., 2017; Wen et al., 2020). Literatürde multimodal CNN yapılarına dayalı modellerin, yalnızca MRI kullanan modellere göre erken evre Alzheimer ayırımında anlamlı biçimde daha yüksek doğruluk sağladığı rapor edilmiştir.

Tıbbi uygulamalarda derin öğrenme modelleri, özellikle aşağıdaki görevlerde yaygın olarak kullanılır:

- Görüntü sınıflandırma (Classification): Hastalık varlığı, tipi veya evresinin otomatik belirlenmesi.
- Görüntü segmentasyonu (Segmentation): Organ, doku veya lezyonların piksel düzeyinde ayrıştırılması.
- Tespit (Detection) ve lokalizasyon: Lezyon, tümör veya anormal yapıların uzamsal konumlandırılması.
- Hastalık progresyonu tahmini (Prognosis): Zaman içinde hastalık gelişiminin modellenmesi.
- Veri sentezi ve artırımı (Data synthesis/augmentation): Üretici çekişmeli ağ (Generative Adversarial Networks - GAN) gibi yapılarla az sayıda tıbbi verinin çoğaltılması.
- Risk skorum (Risk prediction): Genetik, klinik ve demografik değişkenlerin birleştirilerek bireysel riskin hesaplanması.

AI temelli görüntü analizi, özellikle Alzheimer hastalığı gibi nörodejeneratif hastalıklarda kullanılabilir. Derin öğrenme modelleri, insanların MRI görüntülerinde fark etmekte zorlanabileceği ince ve erken dönem yapısal değişimleri öğrenebilmektedir. Bu modeller, hastalığın erken evrede teşhis edilmesi, evrenmesi ve ilerlemesinin izlenmesi gibi klinik görevleri yüksek doğrulukla yerine getirebilmektedir (Bron et al., 2015; Suk et al., 2014). Literatürdeki çok sayıda çalışma, CNN mimarilerinin Alzheimer hastalarını sağlıklı kontrollerden ve hafif bilişsel bozukluk (Mild Cognitive Impairment, MCI) olan kişilerden ayırt edebildiğini göstermiştir.

Veri azlığı, tıbbi görüntüleme alanında karşılaşılan önemli sınırlamalardan biridir. Klinik verilerin sınırlı olması, derin öğrenme modellerinin yüksek parametreleri nedeniyle aşırı öğrenme riskini artırmaktadır. Transfer öğrenme yaklaşımı, bu sorunu çözmek için sıklıkla kullanılan bir yaklaşımdır. Bu yöntem, ImageNet gibi büyük ölçekli doğal görüntü veri kümeleri üzerinde önceden eğitilmiş modellerin tıbbi veriler üzerinde yeniden eğitilmesini içerir. Bu nedenle, küçük sayıda örnek içeren tıbbi veri kümeleriyle bile yüksek doğrulukta sınıflandırma yapılabilir (He et al., 2016; Tan and Le, 2019).

Derin öğrenme uygulamalarının tıp alanında uygulanması akademik araştırmalarla sınırlı kalmamış, klinik uygulamalara entegre edilmiş regülasyonlar ve ticari sistemler geliştirmiştir. Amerika Birleşik Devletleri Gıda ve İlaç Dairesi (Food and Drug Administration - FDA) tarafından onaylanan tıbbi cihazlar ve AI tabanlı yazılım sayısı her geçen yıl artmaktadır. Radyolojide (akciğer nodülü tespiti), dermatolojide (melanom tespiti), mamografide (göğüs kanseri taraması) ve retinal görüntüleme (diyabetik retinopati derecelendirmesi) bu sistemler klinik olarak kullanılmaktadır. CNN tabanlı algoritmalar, bu tür görevlerde tanı doğruluğunu artırmak için radyologlara karar vermede yardımcı olur (Administration, 2023).

Bu gelişmeler, nörogörüntüleme uygulamalarında da hızla yayılmaktadır. AI, Alzheimer, Parkinson ve epilepsi gibi nörolojik hastalıkların tanısında giderek daha popüler hale gelmektedir. Erken tanı, prognostik değerlendirme ve hasta izleme gibi alanlarda yenilikçi çözümler sunan bu sistemler, klasik görüntüleme yöntemlerinin ötesine geçmektedir. Hipokampal atrofi, kortikal incelleme ve ventriküler genişleme gibi yapısal değişimleri doğru bir şekilde tanımlama yeteneğine sahip CNN mimarileri, özellikle Alzheimer ile ilgili araştırmalarda gösterilmiştir (Frisoni et al., 2010; Jack et al., 2018). Bu sayede erken tanı sürecinin daha objektif, hızlı ve tekrarlanabilir hale gelmesi mümkün olmaktadır.

Sonuç olarak, derin öğrenme yöntemleri tıbbi görüntüleme alanında yalnızca tanısal doğruluğu artırmakla kalmaz, aynı zamanda sağlık profesyonellerinin iş yükünü azaltır, tanı süreçlerini standartlaştırır ve hasta bakım kalitesini artırır. Bu nedenle, Alzheimer hastalığı gibi tanısı güç nörodejeneratif bozukluklara yönelik geliştirilen otomatik sınıflandırma sistemlerinde derin öğrenme modellerinin kullanılması, günümüz tıp teknolojileri için çok önemlidir.

Veri gizliliği, etik onay ve model güvenliği gibi konular, tıbbi görüntüleme arařtırmalarında derin öğrenmenin yaygınlaşmasının bir sonucu olarak önem kazanmıştır. MRI gibi kişisel sađlık verilerinin gizliliği nedeniyle, modellerin eğitiminde kullanılan veri setlerinin anonimleştirilmesi, güvenli veri saklama sistemleri ve etik kurul onaylarının klinik arařtırmalarda çok önemli bir rol oynadığını göstermektedir. Yanlış tanı durumlarında hukuki karşılık, veri güvenliği ve klinik karar süreçlerinde sorumluluk paylaşımı gibi “AI tabanlı sistemlerin standartlaştırılması gerekliliği literatürde sıklıkla vurgulanmaktadır (Topol, 2019). Bu çerçevede derin öğrenme uygulamalarının hem teknik başarı hem de etik uyumluluk ve hasta güvenliği açısından değerlendirilmesi gerekmektedir.

2.1.1 Tıpta derin öğrenmenin tarihsel gelişimi

Derin öğrenme yöntemlerinin tıpta yaygınlaşması son on yılda hız kazanmış olsa da, temelini oluşturan yapay sinir ađları çok daha eskiye dayanmaktadır. 1950'lerde Rosenblatt tarafından geliştirilen "perceptron" olarak da bilinen tek katmanlı algılayıcı model otomatik sınıflandırma fikrinin ilk örneklerinden biri olarak kabul edilmektedir. 1980'lerde çok katmanlı sinir ađları (Multi-Layer Perceptron – MLP) geliştirildi ve tıp alanında elektroensefalografi, elektrokardiyografi ve radyografi gibi alanlarda ilk bilgisayarlı tanı destek sistemleri test edilmeye başlandı.

Derin öğrenmenin yeniden ivme kazanması, 2012 yılında AlexNet mimarisinin ImageNet yarışmasında elde ettiği çarpıcı başarı ile gerçekleşmiştir. GPU tabanlı hesaplama ve veri hacminin artması, özellikle radyoloji alanında otomatik tanı sistemleri için yeni bir çađa yol açtı. Bu gelişmeden sonra, CNN tabanlı modeller, meme kanseri taraması, beyin tümörlerinin segmentasyonu ve akciđer nodüllerinin tespiti insan uzman performansına yaklaşan sonuçlar üretmeye başlamıştır (Krizhevsky et al., 2012).

Günümüzde tıbbi görüntüleme alanında kullanılan neredeyse tüm otomatik tanı sistemlerinin alt yapısı, dikkat mekanizmaları, konvolüsyonel sinir ađları, enkoder-dekoder yapıları ve çoklu modalite entegrasyonu içerir. Bu tarihsel gelişme, derin öğrenmenin Alzheimer gibi yapısal deđişikliklerin erken evrelerinde ortaya çıktığı hastalıklarda önemli rol oynadığını daha da göstermektedir.

Tarihsel olarak, derin öğrenme teknolojilerinin tıbbi uygulamalara dahil edilmesi, algoritmik başarıları ve klinik iş akışlarının yeniden tanımlanmasını kolaylařtırdı. Özellikle

nörogörüntüleme alanında, klasik görüntü yorumlama süreçleri ile derin öğrenme temelli otomatik analizlerin bir arada kullanıldığı "hibrit" karar yapıları giderek daha popüler hale gelmektedir. Bu, tıp alanındaki derin öğrenme modellerinin yalnızca teknik becerilerle değil, klinisyenlerin günlük yaşamlarına nasıl entegre olabildikleri ile de değerlendirilmesi gerektiğini göstermektedir. Bu durum, Alzheimer hastalığı gibi karmaşık nörodejeneratif bozukluklara yönelik karar destek sistemlerinin geliştirilmesinin önemini ortaya koymaktadır.

2.1.2 Tıbbi görüntü analizinde karşılaşılan teknik zorluklar

Derin öğrenme modelleri yüksek performans sunmalarına rağmen, tıbbi verilerin doğası gereği çeşitli teknik zorluklar içermektedir. Özellikle farklı cihaz markaları, değişken çözünürlük seviyeleri, görüntüleme protokolleri ve kontrast ayarlamaları gibi etkenler, modelin öğrenme sürecini olumsuz etkileyebilmektedir. Literatürde domain shift olarak tanımlanan bu durum, bir merkezde eğitilen modellerin farklı merkezlerden elde edilen veriler üzerinde performans kaybı yaşamasına neden olmaktadır (Wen et al., 2020).

Bununla birlikte, tıbbi görüntülerin etiketlenmesi yüksek uzmanlık gerektiren bir süreçtir ve genellikle deneyimli radyologlar tarafından gerçekleştirilmektedir. Bu durum, veri etiketleme sürecini hem zaman alıcı hem de maliyetli hâle getirmekte; dolayısıyla derin öğrenme modellerinin eğitimi için yeterli sayıda etiketli veriye ulaşmayı zorlaştırmaktadır. Ayrıca, tıbbi veri kümelerinde sınıf dağılımı çoğu zaman dengesizdir; sağlıklı bireylerin baskın olduğu veri setlerinde özellikle MCI ve erken evre Alzheimer sınıflarında düşük duyarlılık değerleri gözlemlenebilmektedir. Buna ek olarak, derin öğrenme modellerinin karar verme süreçlerinin çoğunlukla “kara kutu” yapıda olması, klinisyenlerin bu modellere duyduğu güveni sınırlayan önemli bir faktördür. Bu nedenle son yıllarda Grad-CAM, Katman Bazlı Önem Yayılımı (Layerwise Relevance Propagation – LRP) ve Integrated Gradients gibi açıklanabilir yapay zekâ (XAI) yöntemleri, tıbbi görüntüleme alanında model kararlarının yorumlanabilirliğini artırmak amacıyla yaygın biçimde kullanılmaya başlanmıştır. Tüm bu zorluklar, Alzheimer teşhisinde kullanılan CNN tabanlı modellerin dikkatli bir şekilde tasarlanmasını zorunlu kılmakta; veri artırma, transfer öğrenme ve domain adaptation gibi yöntemlerin önemini daha da artırmaktadır.

Literatürde, bu zorlukların göz ardı edilmesinin modellerin gerçek klinik uygulama ortamlarında performans kaybetmesine ve güvenilirliklerinin sorgulanmasına yol açtığı

vurgulanmaktadır. Özellikle tek merkezli, homojen ve sınırlı sayıda örnek içeren veri setleri üzerinde eğitilen modeller, farklı hastane koşullarında kullanıldıklarında beklenen başarıyı sürdüremeyebilmektedir. Bu nedenle güncel çalışmalarda, eğitim ve doğrulama süreçlerinin çok merkezli veri yapıları, heterojen hasta grupları ve farklı görüntüleme protokollerini kapsayacak şekilde tasarlanmasına önem verilmektedir. Alzheimer hastalığına yönelik geliştirilen derin öğrenme modellerinin de benzer şekilde, veri kalitesi, çeşitliliği ve anotasyon güvenilirliği gibi faktörler gözetilerek yapılandırılması, klinik geçerlilik açısından kritik bir gereklilik olarak öne çıkmaktadır.

Son yıllarda yapılan çalışmalar, tıbbi görüntüleme verilerindeki bu teknik sınırlılıkların yalnızca model performansını değil, aynı zamanda klinik karar süreçlerinin güvenilirliğini de doğrudan etkilediğini göstermektedir. Özellikle MRI çekim protokolü, manyetik alan gücü ve filtreleme yöntemlerindeki küçük farklılıkların bile derin öğrenme modellerinin karar sınırlarını kaydırabildiği ve bu nedenle “domain shift” probleminin klinik geçerlilik açısından kritik bir risk oluşturduğu vurgulanmaktadır (Glocker et al., 2019; Wen et al., 2020). Bu nedenle güncel literatürde, Alzheimer gibi hassas sınıflandırma problemlerinde model eğitimi sırasında veri çeşitliliğinin artırılması ve çok merkezli validasyonun yapılması temel bir gereklilik olarak kabul edilmektedir.

2.1.3 Hesaplama gücü, GPU kullanımı ve maliyetler

Tıbbi görüntülerin üç boyutlu yapısı ve yüksek çözünürlüğü nedeniyle derin öğrenme modelleri yüksek düzeyde hesaplama gücü gerektirmektedir. 3D-CNN ve Transformer tabanlı modeller, yüksek parametre sayısına sahip olabilmeleri nedeniyle 3D MRI verileri üzerinde çalışırken yüksek GPU belleği gerektirir.

Literatürde, tek bir Alzheimer sınıflandırma modelinin eğitimi için:

- NVIDIA Tesla V100 veya A100 gibi yüksek performanslı GPU’ların
- 16–40 GB arası grafik işlem belleği (Video Random Access Memory – VRAM) kapasitesinin
- Saatler ile günler arasında değişen eğitim sürelerinin gerektiği rapor edilmiştir (Vieira et al., 2017; Wen et al., 2020).

Bu nedenle, araştırma projelerinde Google Colab Pro, Amazon Web Services Elastic Compute Cloud hizmeti (Amazon Web Services Elastic Compute Cloud, AWS EC2), Azure

ML ve yerel GPU sunucuları gibi altyapılar tercih edilmektedir. Eğitim sürecinin maliyetli olması nedeniyle, hafif modeller (EfficientNet-B0, MobileNetV2) ve transfer learning teknikleri daha yaygın biçimde tercih edilmektedir.

2.1.4 Federated learning ve veri gizliliği odaklı yaklaşımlar

Son yıllarda tıpta kişisel verilerin korunmasına yönelik yasal düzenlemelerin artmasıyla birlikte, veri gizliliği odaklı AI yöntemleri önem kazanmıştır. Özellikle birleşik öğrenme yaklaşımı (Federated Learning – FL), Alzheimer gibi hassas sağlık verileri içeren hastalıklarda öne çıkmaktadır.

FL yaklaşımında veriler merkezi bir sunucuya aktarılmaz; her hastane kendi verisi üzerinde modeli eğitir ve yalnızca model ağırlıkları sunucuya aktarılır. Böylece hem gizlilik korunur hem de çok merkezli öğrenme sayesinde model genellenebilirliği artar.

Literatürde federated learning yöntemlerinin MRI cihaz farklılıklarından kaynaklanan domain shift'i azalttığı, modellerin farklı ülkelerden gelen verilerde tutarlı performans gösterdiği ve gizlilik ihlali riskini azalttığı rapor edilmiştir (Sheller et al., 2020). Bu yaklaşım, gelecekte Alzheimer tanı modellerinin klinik entegrasyonunda kritik bir rol oynayacaktır.

Özetle, tıbbi görüntüleme derin öğrenme yöntemleri yalnızca yüksek doğruluk oranları sunmakla kalmamakta, aynı zamanda veri gizliliği, hesaplama verimliliği ve çok merkezli uygulama gereksinimlerine yönelik yeni yaklaşımlarla sürekli olarak gelişmektedir. Bu yönleriyle derin öğrenme modelleri, Alzheimer hastalığının erken tanı ve sınıflandırılmasında güçlü bir temel oluşturmaktadır.

2.2 Alzheimer Tanısında Kullanılan Derin Öğrenme Mimarileri

Bu alt bölümde, Alzheimer hastalığının risk tahmini, otomatik sınıflandırılması ve evrenmesi gibi görevlerde sıklıkla kullanılan derin öğrenme modelleri incelenmektedir.

Klinik uygulamalarda, EfficientNet, ResNet, U-Net, MobileNetV2 ve Inception aileleri, derin öğrenme literatüründe çok sayıda mimari olmasına rağmen en yaygın olanlardır. Bu mimarilerin her biri farklı bir tasarım felsefesine sahiptir, ancak hesaplama maliyeti, derinlik, parametre verimliliği ve genellenebilirlik gibi faktörler onları birbirinden ayırır.

Sonuç olarak, Alzheimer odaklı arařtırmalarda tercih edilen mimari, veri kümesinin boyutu, sınıf dağılımı, mevcut donanım ve amaçlanan görev (örneğin sınıflandırma, segmentasyon vb.) dikkate alınarak belirlenmelidir. Ayrıca literatürde bu mimarilerin hibrit biçimlerde birleştirildiği (örneğin CNN–Transformer, CNN–LSTM (Uzun Kısa Süreli Bellek – Long Short-Term Memory - LSTM), CNN + Attention mekanizmaları) daha gelişmiş yapılar da bulunmaktadır. Bununla birlikte, bu tezde karşılaştırmalı analiz bölümü beş temel CNN mimarisine odaklanır. Tıbbi görüntüleme literatüründe standart kıyaslama modelleri olarak kabul edilen bu beş yapı, klinik koşullara uyarlanabilir ve karşılaştırılabilir.

Literatürde bu mimariler doğruluk oranları, yorumlanabilirlik, klinik senaryolara uyum ve bakım süreçlerine entegrasyon kolaylığı açısından önemlidir. Örneğin, ResNet ve Inception gibi daha geniş mimariler, karmaşık örüntüleri yakalama yeteneği nedeniyle bazı çalışmalarda tercih edilirken, sınırlı donanıma sahip merkezlerde MobileNetV2 ve EfficientNet-B0 gibi daha hafif mimariler klinik prototiplerde kullanılmıştır. U-Net gibi segmentasyon odaklı modeller genellikle doğrudan tanı elde etmek için değil, beyin bölgelerinden şekilsel ve hacimsel ölçümler elde etmek için "ön işleme bileşeni" olarak konumlandırılır. Bu çeşitlilik, Alzheimer tanısında tek bir 'en iyi model' aramak yerine bağlama ve göreve duyarlı bir mimari seçiminin gerekliliğini göstermektedir.

Literatür incelemeleri, Alzheimer teşhisine yönelik derin öğrenme modellerinin başarısının yalnızca kullanılan mimariye değil, verinin kalitesine, sınıfların ayrışabilirliğine ve kullanılan ön işleme tekniklerine de bağlı olduğunu göstermektedir. Genellenebilirlik, hesaplama maliyeti, segmentasyon-sınıflandırma uyumu ve küçük örneklerde öğrenme kabiliyeti gibi bağlamsal faktörler açısından, benzer doğruluk değerleri sunan farklı mimariler arasında önemli farklar vardır (Basaia et al., 2019; Li et al., 2021). Sonuç olarak, Alzheimer literatüründe mimari seçimi, yalnızca sayısal performans kriterleriyle değil, aynı zamanda modelin klinik gereksinimleri karşılayıp karşılamadığına da bakılmalıdır.

2.2.1 EfficientNet

EfficientNet, veri miktarının sınırlı olduğu Alzheimer teşhisi senaryolarında etkili bir yaklaşım olarak öne çıkmaktadır. Transfer öğrenme yöntemiyle önceden ImageNet üzerinde eğitilmiş ağırlıkların yeniden kullanılması, MRI görüntülerinden karmaşık yapısal örüntülerin öğrenilmesinde oldukça başarılıdır.

Bu yönüyle EfficientNet, hem genellenebilirlik hem de hesaplama verimliliği açısından dikkate değer bir performans sergilemektedir.

Literatürde çok sayıda çalışma, EfficientNet'in Alzheimer hastalığını MRI görüntüleri üzerinden sınıflandırmadaki başarısını incelemiştir. Örneğin, Ntampakis et al. (2024) tarafından yürütülen bir çalışmada, Alzheimer Dementia (AD), MCI ve bilişsel olarak normal bireyler (Cognitively Normal - CN) olarak adlandırılan Alzheimer hastaları üç evreye ayrılmıştır. “EfficientNet tabanlı bir model kullanılarak gerçekleştirilen sınıflandırmada %94.1 doğruluk oranına ulaşılmıştır. Aynı çalışmada, EfficientNet, ResNet ve DenseNet gibi daha karmaşık mimarilerle karşılaştırıldığında, daha düşük parametre sayısına rağmen genellenebilirlik açısından daha istikrarlı ve benzer veya daha yüksek performans sergilediği rapor edilmiştir (Ntampakis et al., 2024).

EfficientNet'in Alzheimer teşhisi açısından başlıca avantajları şu şekilde özetlenebilir:

- Optimize parametre yapısı: Daha düşük hesaplama maliyeti ile yüksek doğruluk elde edilmesini sağlamaktadır (Tan and Le, 2019).
- Verimli eğitim süreci: Daha kısa sürede stabil öğrenme gerçekleştirilir (Tan and Le, 2019).
- Mobil uyumluluk: Hafif yapısı sayesinde düşük donanım gereksinimi ile çalışabilir (Hu et al., 2018; Tan and Le, 2019).
- Transfer öğrenmeye uygunluk: Pre-trained ağırlıklar sayesinde hızlı adaptasyon mümkündür (Tan and Le, 2019).

Bununla birlikte, EfficientNet belirli sınırlamalara sahiptir. Özellikle daha büyük varyantlarda (B5–B7) doğruluk artsa bile eğitim süresi, donanım gereksinimi ve bellek tüketimi önemli ölçüde artmaktadır. Ek olarak, sınıf dengesizliği olan veri kümelerinde küçük sınıflara ait örneklerin sınıflandırılmasında duyarlılık veya recall oranları azalmıştır.

Sonuç olarak, EfficientNet, tıbbi görüntüleme alanında verimlilik ve doğruluk arasında dengeli bir performans sunan çağdaş bir mimaridir. Transfer öğrenme yaklaşımıyla birlikte kullanıldığında, yüksek doğruluk, düşük maliyet ve sürdürülebilir hesaplama verimliliği sunarak, Alzheimer hastalığı gibi erken evre ayrımının önemli olduğu durumlarda önemli bir alternatif oluşturmaktadır.

EfficientNet, farklı veri setlerinde istikrarlı sonuçlar verebilmesi nedeniyle Alzheimer arařtırmalarında sıklıkla tercih edilmektedir. Literatürde, özellikle OASIS, ADNI ve AIBL gibi büyük veri setleri üzerinde yapılan karşılařtırmalı analizlerde, EfficientNet tabanlı modellerin diđer yaygın CNN mimarilerine kıyasla daha yüksek AUC deđerleri ve daha yüksek doğruluk gösterdiđi gösterilmiřtir. alıřmalar, modelin özellikle erken evre Alzheimer ile sađlıklı bireyler arasındaki ince yapısal farklılıkları belirlemede iyi alıřtıđını vurgulamaktadır. Ayrıca, sınırlı örneklem büyüklüğüne sahip veri kümelerinde dahi tatmin edici genelleme performansı sergileyebilmektedir.

2.2.2 ResNet (residual network)

ResNet, derin öğrenme mimarileri arasında, özellikle görüntü sınıflandırma görevlerinde yüksek performans sergileyen ve yaygın kabul görmüş bir modeldir. He et al. (2016), derin sinir ađlarının eđitilebilirliđini artırmak amacıyla bu mimariyi önermektedir. Ařırı öğrenme (overfitting) ve gradyan sönmesi (vanishing gradient) gibi sorunlara karşı önemli bir çözüm sunmuřtur (He et al., 2016).

Literatürdeki çeřitli alıřmalar, ResNet'in Alzheimer teřhisinde etkili olduđunu göstermektedir. Örneđin, Araji ve Abbas (2022) tarafından yürütölen bir alıřmada ResNet-50 mimarisi kullanılarak Alzheimer, MCI ve sađlıklı bireyler sınıflandırıldı. Modelin işlenmiş T1-ađırlıklı MRI verilerinde %93,5 doğruluk oranına ulařıldıđı rapor edilmiřtir. Bu alıřmada ayrıca ResNet-50'nin karmařık örüntüleri öğrendiđi ve daha az sayıda epok ile hızlı yakınsama (convergence) sađladıđı gösterilmiřtir (Araji and Abbas, 2022).

ResNet-50'nin Alzheimer teřhisi aısından başlıca avantajları:

- Derin mimari: Karmařık yapısal iliřkileri öğrenme kapasitesi yüksektir (He et al., 2016).
- Residual bađlantılar: Gradyan kaybını azaltarak derin modellerin eđitimini kolaylařtırır (He et al., 2016).
- Transfer öğrenmeye uygunluk: Yaygın olarak kullanılan bir mimari olduđundan ok sayıda önceden eđitilmiş ađırlık seti mevcuttur (Litjens et al., 2017; Shen et al., 2017).
- Modöler yapı: Farklı derinliklerdeki varyantlara kolayca öleklenebilir (ResNet18–ResNet152) (He et al., 2016).

Bununla birlikte, ResNet-50 belirli sınırlamalara sahiptir. Modelin yüksek parametre sayısı, eğitim süresini uzatabilir ve hesaplama kaynakları için daha fazla GPU belleği gerektirebilir. Sınırlı veriyle çalışan durumlarda aşırı öğrenme (overfitting) riski artabilir, bu nedenle kayıp, veri artırma ve düzenleme gibi önlemler önemlidir.

Sonuç olarak, ResNet-50 iyi bir derin öğrenme mimarisidir ve MRI görüntüleri üzerinden Alzheimer hastalığının otomatik sınıflandırılmasında hem yüksek doğruluğu hem de geniş genellenebilirliği ile öne çıkmaktadır. Derinliği sayesinde karmaşık yapısal örüntüleri öğrenme yeteneğine sahiptir ve aynı zamanda derinliğin getirdiği zorlukları büyük ölçüde aşabilmektedir.

Son yıllarda yapılan çok merkezli çalışmalarda ResNet tabanlı modellerin, Alzheimer tanısında farklı yaş grupları ve görüntüleme klinikleri arasında da tutarlı performans gösterdiği belirtilmiştir. Literatürde ResNet modellerinin özellikle MCI alt tiplerinin ayırımında daha kararlı tahminler ürettiği; bu nedenle Alzheimer progresyonunun izlenmesinde klinik araştırmalarda sıkça kullanıldığı görülmektedir. Ayrıca bazı araştırmalarda ResNet ailesinin, eğitim süreleri açısından diğer derin mimarilere göre daha hızlı yakınsadığı vurgulanmıştır.

Bu yönüyle ResNet tabanlı mimariler, hem klinik genellenebilirlik hem de uzunlamasına hastalık takibi gerektiren çalışmalarda güvenilir bir referans model olarak öne çıkmaktadır.

2.2.3 U-Net

Literatürde U-Net'in Alzheimer teşhisindeki başarısını destekleyen çok sayıda çalışma bulunmaktadır. Örneğin Yousefi-Banaem (2021) tarafından gerçekleştirilen bir çalışmada, U-Net mimarisi kullanılarak MRI görüntülerinden hipokampus segmentasyonu yapılmış ve modelin %92.3 Dice benzerlik katsayısı (Dice Similarity Coefficient – DSC) elde ettiği rapor edilmiştir. DSC katsayısı, segmentasyonun doğruluğunu ölçen temel bir metrik olup, bu sonuç uzman radyolog performansına yakın bir performans göstermektedir. Aynı çalışmada U-Net'in MCI ve erken evre Alzheimer ayırımında önemli bir rol oynadığı ve geleneksel segmentasyon algoritmalarına kıyasla daha istikrarlı ve tekrar edilebilir sonuçlar sunduğu vurgulanmıştır (Yousefi-Banaem, 2021).

U-Net tabanlı segmentasyon yaklaşımlarının Alzheimer literatüründe yaygınlaşmasının bir diğer nedeni, bu modellerin bölgesel beyin yapılarındaki küçük hacim değişimlerini güvenilir biçimde ortaya koymasındır. Literatürde hipokampus, amigdala ve temporal lob bölgelerinin hacim ölçümlerinin U-Net segmentasyonu ile daha standart hale geldiği ve bu ölçülerin sonraki sınıflandırma modellerinde tanısallığı artırdığı gösterilmiştir. Bu yönüyle U-Net, Alzheimer çalışmalarında yalnızca bir ön işleme aracı değil, aynı zamanda biyobelirteç çıkarımı için de kritik bir yöntem olarak değerlendirilmektedir.

U-Net'in Alzheimer tanısındaki başlıca avantajları şunlardır:

- Anatomik hassasiyet: Küçük yapılar (örneğin hipokampus) için yüksek segmentasyon başarısı (Yousefi-Banaem, 2021).
- Veri verimliliği: Az sayıda etiketli veride bile etkili öğrenme kabiliyeti (Yousefi-Banaem, 2021).
- Yorumlanabilirlik: Lokal yapısal bozulmaların görsel olarak analizine imkân tanır (Yousefi-Banaem, 2021).
- Erken tanıya katkı: Bölgesel değişimlerin erken evrede tespitini destekler (Yousefi-Banaem, 2021).

Bununla birlikte, U-Net doğrudan sınıflandırma amacıyla kullanılmadığından, bu tezdeki karşılaştırmalı analizlerde diğer mimarilerle eşit koşullarda kıyaslanması sınırlı kalmıştır. Ancak literatürde, segmentasyon sonrası çıkarılan anatomik ölçütlerin sınıflandırma modelleriyle bütünleştirilmesi durumunda tanısallığın anlamlı ölçüde arttığı belirtilmektedir.

Sonuç olarak U-Net, Alzheimer hastalığında yapısal değişimlerin bölgesel analizi için güçlü ve güvenilir bir araçtır. Derin öğrenme tabanlı otomatik teşhis sistemlerinde ön işleme ve veri çıkarımı aşamalarında kullanıldığında, sınıflandırma performansını artıran kritik bir bileşen olarak değerlendirilmektedir.

2.2.4 MobileNetV2

MobileNetV2, hafif yapısı ve düşük hesaplama gereksinimi nedeniyle Alzheimer gibi sınırlı veri içeren tıbbi görüntüleme çalışmalarında sık tercih edilen modern CNN modellerinden biridir.

Literatürde, Alzheimer teşhisi için MobileNetV2 kullanımıyla ilgili çok sayıda çalışma bulunmaktadır. Örneğin, Akinyelu ve Blignaut (2022) tarafından yürütülen çalışmada, MobileNetV2 mimarisi Alzheimer evrelerini sınıflandırmak için kullanıldı. Düşük parametre sayısına rağmen modelin %90,6 doğruluk oranına ulaştığı rapor edilmiştir. Aynı araştırmada, MobileNetV2'nin eğitim süresinin diğer derin mimarilere kıyasla belirgin biçimde kısa olduğu ve bu nedenle gerçek zamanlı tanı uygulamaları için uygun bir yapı sunduğu vurgulanmıştır (Akinyelu and Blignaut, 2022).

MobileNetV2'nin Alzheimer teşhisindeki başlıca avantajları:

- Hafif ve hızlı yapı: Düşük bellek ve işlem gücüyle çalışabilir; mobil ve gömülü sistemlerle uyumludur (Akinyelu and Blignaut, 2022).
- Transfer öğrenmeye uygunluk: Küçük veri kümeleriyle bile başarılı özelleştirme yapılabilir (Akinyelu and Blignaut, 2022).
- Kısa eğitim süresi: Daha hızlı yakınsama sağlar (Akinyelu and Blignaut, 2022).
- Enerji verimliliği: Düşük güç tüketimi sayesinde sürdürülebilir kullanım sunar (Akinyelu and Blignaut, 2022).

Ancak modelin daha az katman derinliği, karmaşık örüntüleri yakalama kabiliyetini sınırlayabilir. Bu nedenle özellikle erken evre Alzheimer gibi ince yapısal farkların bulunduğu durumlarda, duyarlılığın (recall) görece daha düşük olabileceği belirtilmiştir. Yine de uygun veri artırma (data augmentation) ve düzenleme (regularization) stratejileriyle bu kısıtlar büyük ölçüde dengelenebilmektedir.

Sonuç olarak, MobileNetV2 verimlilik odaklı tıbbi teşhis sistemlerine destek sağlar. Alzheimer teşhis uygulamalarını hızlı ve taşınabilir bir şekilde geliştirmek için etkili ve sürdürülebilir bir sistem olarak öne çıkmaktadır, özellikle düşük kaynaklı klinik ortamlarda.

MobileNetV2, Alzheimer literatüründe önemli bir yere sahip olan düşük donanımlı cihazlarda çalışan prototip tanı sistemlerini içerir. Çalışmalar, düşük güçlü işlemcilerde veya mobil cihazlarda bile MobileNet tabanlı modellerin yüksek doğruluk düzeylerine ulaşabileceğini göstermiştir. Hem evde izlem hem de kırsal alanlarda hızlı tarama uygulamalarında bu önemli bir avantajdır. Sonuç olarak, MobileNetV2, geleceğe yönelik erişilebilir ve hızlı tanı sistemleri geliştirme potansiyeli açısından önemli bir konumda yer almaktadır.

2.2.5 Inception (GoogleNet)

Zhang (2018) tarafından yürütülen bir arařtırmada, InceptionV4 mimarisi Alzheimer hastalığını MRI görüntüleri üzerinden sınıflandırmak için kullanılmıştır ve model yüksek doğruluk oranları göstermiştir. Aynı çalışma, farklı aktivasyon fonksiyonlarının model performansını nasıl etkilediğini incelemiştir. Inception mimarisi, orta evre Alzheimer ile MCI sınıflarını ayırmada oldukça başarılı olmuştur (Zhang, 2018).

Inception mimarisinin Alzheimer teşhisindeki öne çıkan avantajları şunlardır:

- Çok yönlü örüntü çıkarımı: Farklı filtre boyutları ile hem küçük hem de geniş ölçekli yapısal değişimleri yakalayabilir (Szegedy et al., 2015).
- Verimli yapı: Derinliğe aşırı ihtiyaç duymadan genişleme (width scaling) ile öğrenme kapasitesini artırır (Szegedy et al., 2015).
- Aşırı öğrenme riskinin azaltılması: Paralel yapı sayesinde modelin genel performansı daha daha dengeli bir öğrenme süreci sunabilmektedir (Ioffe and Szegedy, 2015; Szegedy et al., 2016).
- Aktivasyon uyumluluğu: Farklı aktivasyon fonksiyonlarıyla yapılandırmayı kolaylaştırır (Zhang, 2018).

Bununla birlikte, inception mimarileri çok sayıda alt dal (branch) içerir, bu da mimari karmaşıklığı artırır ve öğrenme süresini ve hesaplama yükünü artırabilir. InceptionV4, daha yüksek doğruluk sunmasına rağmen daha fazla eğitim ve donanım gerektirir. Bununla birlikte, yüksek genellenebilirliği ve çoklu ölçekli öznelik çıkarımı nedeniyle bu eksiklikler uygulama için tolere edilebilir.

Bu nedenle, Inception mimarisi, derin öğrenme tabanlı görüntü sınıflandırma uygulamalarında çoklu ölçekli öznelik öğrenimine izin vererek Alzheimer teşhisi için güçlü bir alternatif sunmaktadır. Sınıf ayırımının zor olduğu durumlarda klinik güvenilirlik ve yüksek ayırıştırma kabiliyeti kazandırır.

Inception tabanlı modeller, özellikle sınıflar arası sınırların belirsiz olduğu durumlarda (örneğin, erken MCI ile geç MCI ayrımı) daha yüksek performans gösterebilmektedir. Bu, bir dizi araştırma tarafından gösterilmiştir. Bu modeller, çeşitli görüntü ölçeklerinden aynı anda veri üreterek Alzheimer'e özgü değişimlerin geniş ve küçük ölçekli versiyonları birlikte değerlendirmeyi mümkün kılmaktadır. Literatürde bu yaklaşımın sınıflandırma kararlılığını

artırdığı ve yanlış pozitif oranlarını azalttığı belirtilmektedir. Bu nedenle Inception tabanlı yaklaşımlar, erken evre Alzheimer ve MCI alt tiplerinin ayırt edilmesinde hem akademik çalışmalarda hem de klinik karar destek sistemlerinde giderek daha fazla tercih edilmektedir.

2.2.6 Diğer derin öğrenme yaklaşımları ve hibrit modeller

Alzheimer literatüründe EfficientNet, ResNet, U-Net, MobileNetV2 ve Inception gibi mimariler kullanılıyor olsa da, son yıllarda Transformer tabanlı ağlar, 3B konvolüsyonel ağlar (3D-CNN) ve hibrit CNN-RNN modelleri (Tekrarlayan Sinir Ağı – Recurrent Neural Network - RNN) araştırma alanında güçlü alternatifler haline gelmiştir. 3D-CNN modelleri, MRI görüntülerinin hacimsel doğasını koruyarak üç boyutlu uzamsal ilişkileri daha iyi öğrenebilir ve yüksek doğrulukla hipokampal atrofi gibi yapısal değişikliklerin modellenmesinde etkili sonuçlar sunabilmektedir (Korolev et al., 2017). Bununla birlikte, hibrit ağlar (CNN ve LSTM gibi) hastalık ilerlemesini uzunlamasına (longitudinal) takip verilerinde modellemek için kullanılmaktadır (Suk et al., 2014). Daha güncel çalışmalarda ise Görsel dönüştürücü mimarisi (Vision Transformer architecture – ViT) tabanlı modellerin Alzheimer sınıflandırmasında CNN tabanlı modellere kıyasla benzer ya da bazı çalışmalarda daha yüksek performans gösterebildiği rapor edilmektedir (Dosovitskiy et al., 2021). Bu bulgular, Alzheimer tanısında derin öğrenmenin hızla genişleyen dinamik bir araştırma alanı olduğunu göstermektedir.

Alzheimer literatüründe son yıllarda farklı model ailelerinin birlikte kullanıldığı hibrit yaklaşımların da arttığı görülmektedir. Örneğin bazı çalışmalarda CNN tabanlı modellerden elde edilen özellikler istatistiksel yöntemler, demografik bilgiler veya klinik test skorları ile birleştirilmiş ve bu çoklu veri yaklaşımının yalnızca görüntü tabanlı modellerden daha yüksek doğruluk sağladığı rapor edilmiştir. Bu durum Alzheimer tanısında tek bir mimari yerine, çok yönlü model kombinasyonlarının daha etkili olabileceğini ortaya koymaktadır.

Bu kapsamlı çerçeve, Alzheimer literatüründe "tek ve mutlak en iyi mimari" arayışının yerini, belirli klinik senaryolar ve veri özellikleri için en uygun mimarilerin veya mimari kombinasyonlarının belirlenmesine bıraktı. Çalışmalarda yüksek duyarlılık hedeflenirken, diğer çalışmalarda yanlış pozitif oranını sınırlamak ya da hafif bilişsel bozukluk gibi erken evre tablolarını daha hassas olarak belirlemek daha önemli olabilir. Bu bakış açısıyla, bu tez için seçilen mimariler EfficientNet, ResNet, U-Net, MobileNetV2 ve Inception'dır. Amaç,

Alzheimer tanısına yönelik en iyi karar destek sistemleri ve otomatik sınıflandırma yapılarını sistematik bir şekilde karşılaştırmaktır.

2.3 Transfer Öğrenme ve Veri Dengeleme Teknikleri

Transfer öğrenme prensibi, daha büyük veri setleri üzerinde önceden eğitilmiş derin öğrenme modellerini hedef görev için yeniden eğitir ve uyarlar. Eğitilmiş modeller genellikle ImageNet gibi milyonlarca doğal görüntü içeren veri kümelerinde kullanılır ve tıbbi görüntüleme görevlerinde yeniden kullanılarak yüksek doğruluk ve kısa eğitim süresi sağlar (Yosinski et al., 2014).

Bu yöntem, Alzheimer hastalığı teşhisi gibi MRI tabanlı sınıflandırma problemlerinde özellikle etkilidir. Çünkü MRI görüntüleri genellikle düşük örnek sayısına sahip olup, yapısal farklılıklar sınırlıdır. Önceden eğitilmiş modeller, düşük seviyeli görsel özellikleri (örneğin kenarlar, dokular, konturlar) zaten öğrenmiş oldukları için, daha az veriyle daha iyi genelleme başarımı gösterir.

Sınıf dengesizliği sorununu çözmek için örnek uzayını değiştiren ve kayıp fonksiyonunu ağırlıklandıran yeniden örnekleme teknikleri de kullanılmaktadır. Oversampling teknikleri, azınlık sınıfa ait örnek sayısını artırarak sınıf dağılımını dengelemeyi amaçlamaktadır. Bunun azınlık sınıflarının yeniden örnekleme tekniği (Synthetic Minority Over-sampling Technique – SMOTE) azınlık sınıfları arasında sentetik örnekler üreterek sınıf dağılımını dengeler (Chawla et al., 2002). Literatürde SMOTE kullanılarak Alzheimer verilerinin duyarlılık (recall) değerleri önemli ölçüde artırılmıştır. Bu artışın özellikle erken evre Alzheimer ve MCI gibi az örnek içeren sınıflarda belirgin olduğu rapor edilmektedir (Wen et al., 2020).

Veri dengesizliğine karşı bir diğer etkili yöntem, kayıp fonksiyonlarının değiştirilmesidir. Azınlık sınıfları, zor örnekleri öne çıkararak ve kolay örneklerin etkisini azaltarak modelin odaklanmasını sağlar (Lin et al., 2017). Class-balanced loss ise her sınıfın efektif örnek sayısını dikkate alarak kayıp fonksiyonunu dinamik olarak ölçeklendirir. Bu tür kayıp fonksiyonları, Alzheimer teşhisinde özellikle MCI gibi nadir sınıfların daha güvenilir şekilde öğrenilmesine katkı sağlamaktadır. Literatürde focal loss kullanılan modellerin, dengesiz MRI veri setlerinde hem doğruluk hem de F1-score bakımından klasik cross-entropy tabanlı modellere göre daha iyi performans sunduğu bildirilmektedir.

Oktavian et al. (2022) tarafından yapılan bir çalışmada, ImageNet üzerinde önceden eğitilmiş ResNet-18 mimarisi Alzheimer MRI görüntülerine uygulanmış ve modelin %88.3 doğruluk oranı elde ettiği rapor edilmiştir. Bu bulgu, transfer learning'in tıbbi sınıflandırma görevlerinde performansı anlamlı biçimde artırdığını göstermektedir (Oktavian et al., 2022).

Transfer learning'in Alzheimer teşhisindeki başlıca avantajları:

- Küçük veri setlerinde dahi yüksek doğruluk sağlar (Pan and Yang, 2010; Shin et al., 2016).
- Eğitim süresini kısaltır, konverjansı hızlandırır (Shin et al., 2016; Tan and Le, 2019).
- Aşırı öğrenme (overfitting) riskini azaltır (Litjens et al., 2017; Pan and Yang, 2010).
- Modelin genelleme kabiliyetini artırır (Shen et al., 2017; Yosinski et al., 2014).

Transfer learning yönteminde bazı sınırlamalar vardır. "Negatif transfer" olarak bilinen bir durum, özellikle doğal görüntüler üzerinde eğitilmiş modellerin tıbbi görüntülere doğrudan aktarılması sonucunda ortaya çıkabilir. Bu durumda, model hedef görevde beklenen performansın altında kalabilmektedir. Örneğin, ImageNet verisinde yer alan gündelik nesnelere ait görüntüler ile beyin MRI görüntülerindeki anatomik yapılar arasında önemli farklar vardır. Bu, önceden öğrenilmiş özelliklerin her zaman optimal sonuçlar üretmeyebileceğini gösterir. Literatür, bu sorunu çözmek için tıbbi görüntü veri tabanları üzerinde yeniden ön eğitim yapıldığını, yalnızca belirli katmanların yeniden eğitildiğini veya daha düşük öğrenme oranlarıyla hassas ayarlamalar yapıldığını belirtmektedir. Bu tür teknikler, transfer öğreniminin hesaplama avantajlarını korumakla birlikte, Alzheimer'a özgü yapısal özelliklerin daha iyi temsil edilmesini sağlar.

Alzheimer hastalığı literatüründe, veri dengesizliği yöntemlerinin transfer öğrenme teknikleriyle birlikte kullanıldığı hibrit modeller ortaya çıkmaya başladı. Bu modeller, önceden eğitilmiş ağlardan elde edilen orta seviye özellik haritalarını, yeniden örnekleme (oversampling veya undersampling) ya da sentetik veri üretim yöntemleriyle geliştirilmiş daha dengeli bir eğitim uzayıyla birleştirir. Bir sentetik örnek oluşturmak için varyasyonel otomatik kodlayıcı tabanlı modeller (Variational Autoencoder - VAE) ve GAN kullanılabilir. Bu teknikler, erken MCI gibi azınlık sınıflarda model duyarlılığını ve çeşitliliği artırır. Literatürde belirtildiği gibi, gerçek ve sentetik MRI görüntülerinin birleştirilmesi, klasik SMOTE yaklaşımına kıyasla daha istikrarlı AUC ve F1-puan sonuçları sağlar.

Ek olarak, bu hibrit yaklaşımların modeli veri kaynağı, çekim protokolü ve cihaz değişikliklerine karşı daha dayanıklıdır. Bu, çok merkezli veri setlerinde daha az performans kaybına neden olur. Bu nedenle, Alzheimer gibi sınırlı ve dengesiz veri içeren klinik senaryolarda modern yaklaşımlar, transfer öğrenimi ile veri dengeleme tekniklerinin kombinasyonunu içerir.

2.3.1 Veri dengeleme ve aşırı öğrenmeyi önleyici teknikler

Tıbbi veri setlerinde sıklıkla karşılaşılan bir diğer sorun sınıf dengesizliğidir. Örneğin, erken evre Alzheimer tanısına veya MCI tanısına yönelik veri setlerinde sağlıklı bireylerin sayısı genellikle yüksektir. Bu nedenle, model baskın sınıfa eğilim gösterebilir ve azınlık sınıflarda recall veya duyarlılık düşük olabilir. Bu sorunu çözmek için sıklıkla kullanılan bir yöntem, ağırlıklı kayıp fonksiyonları yaklaşımıdır. Bu fonksiyonlar sayesinde, model azınlık sınıflarını daha etkili bir şekilde öğrenebilir. Ağırlıklı cross-entropy kaybı Alzheimer sınıflandırmasında yaygın olarak tercih edilmektedir.

Sınıf dengesini sağlamak ve modelin genellenebilirliğini artırmak için veri artırma teknikleri de önemli bir stratejidir. Bu yöntemler, mevcut verilerin değiştirilmesi yoluyla modelin daha geniş bir örnek uzayında öğrenmesine olanak tanır.

Yaygın veri artırma yöntemleri arasında şunlar yer alır:

- Döndürme (rotation): Görüntülerin belirli açılarla çevrilmesi.
- Ölçekleme / Yakınlaştırma (scaling / zoom): Boyut değişimiyle örnek çeşitliliği sağlanması.
- Parlaklık ve kontrast ayarı (brightness / contrast adjustment): Görsel çeşitliliği artırır.
- Yatay veya dikey çevirme (flip): Görüntü yönünün değiştirilmesi.
- Gürültü ekleme (noise injection): Rastgele pikseller eklenerek dayanıklılığın artırılması.

Bu yöntemler sayesinde model, sınırlı sayıda MRI görüntüsüyle eğitilse dahi, farklı veri senaryolarına karşı daha dayanıklı ve genelleştirilebilir hale gelir. Ayrıca bu teknikler, aşırı öğrenme (overfitting) riskini azaltarak, test verisi üzerinde daha yüksek başarı elde edilmesini sağlar.

Sonuç olarak, transfer learning ve veri dengeleme stratejileri, Alzheimer teşhisi gibi sınırlı veri içeren tıbbi görüntüleme problemlerinde derin öğrenme modellerinin etkinliğini önemli ölçüde artırmaktadır. Bu yöntemler, hem verimlilik hem de genellenebilirlik açısından modern AI tabanlı klinik karar destek sistemlerinin temel bileşenleri hâline gelmiştir.

2.4 Alzheimer Tanısında Derin Öğrenme Modellerinin Klinik Kullanılabilirliği

Son yıllarda, derin öğrenme modelleri Alzheimer hastalığının MRI görüntüleri üzerinden otomatik tanılanmasında başarılı sonuçlar üretmiş olsa da, bu başarıların gerçek klinik koşullara ne ölçüde uygulanabileceği literatürde önemli bir tartışma alanıdır. Klinik kullanım için temel beklentiler şunlardır: modellerin yüksek doğruluğu, önyargılardan arındırılmış olması, farklı popülasyonlarda genellenebilir davranması ve karar süreçlerinin hekimler tarafından yorumlanabilir şekilde açıklanmasıdır. Bu nedenle, model performansı, Alzheimer odaklı derin öğrenme çalışmalarında ön plana çıkmakla birlikte, etik güvenilirlik, veri çeşitliliği, klinik geçerlilik ve açıklanabilirlik gibi kriterler de eş zamanlı olarak değerlendirilmelidir.

Ntampakis et al. (2024), EfficientNet tabanlı modelleri OASIS ve ADNI gibi iki büyük ve bağımsız MRI veri seti üzerinde değerlendirerek, çoklu veri kaynağının model genellenebilirliğine önemli katkı sağladığını göstermiştir. Bu yaklaşım, tek bir merkeze ait veriyle sınırlı kalmayan modellerin klinik senaryolara daha uygun olduğunu ortaya koymaktadır. Araştırmada, Alzheimer hastalığının üç evresinin sınıflandırılmasında yüksek doğruluk elde edilmiş ve modelin farklı popülasyonlara uyum sağlayabildiği rapor edilmiştir (Ntampakis et al., 2024).

Bununla birlikte, literatürde CNN tabanlı modellerin klinik kullanımını sınırlayan birkaç faktör vurgulanmıştır. Modellerde bias, özellikle MRI veri setlerinde yaş, cinsiyet, eğitim seviyesi ve etnik köken gibi demografik özelliklerin homojen olmaması nedeniyle ortaya çıkabilir. Bu durum, modelin bazı hasta gruplarında daha düşük performans göstermesine yol açabilmekte ve klinik karar süreçlerinde etik gözetimi zorunlu kılmaktadır (Wen et al., 2020). Bu nedenle, modelin klinik geçerlilik düzeyini artırmak için geniş ve temsili veri kümeleri kullanılarak çok merkezli doğrulama yapılması kritik bir gereklilik olarak değerlendirilmektedir.

Klinik kabulün bir diğer önemli boyutu açıklanabilirlik (explainability) olarak öne çıkmaktadır. Derin öğrenme modelleri doğaları gereği “kara kutu” olarak değerlendirildiğinden, hekimlerin model çıktılarını yorumlayabilmesi için karar mekanizmasının görünür hâle getirilmesi gerekmektedir. Bu kapsamda Grad-CAM, saliency maps gibi XAI yöntemleri, Alzheimer sınıflandırmasında modelin hangi beyin bölgelerini dikkate aldığını görselleştirerek klinik güveni artırmaktadır (Selvaraju et al., 2017).

Literatürde, doğru sınıflandırma yapan modellerin hipokampus, entorhinal korteks ve lateral ventriküller gibi Alzheimer ile ilgili bölgeleri vurguladığını, yanlış sınıflandırma yapan modellerin ise vurgulamadığını göstermiştir. Bu bulgular, XAI'nin model doğrulama ve klinik kabul sürecinde önemli bir rol oynadığını göstermektedir. Açıklanabilirliği artırmak, model hatalarının nedenlerini belirlemek ve yeniden eğitim (retraining) prosedürlerini düzenlemek için faydalıdır. Sonuç olarak, modeller şeffaf araçlardır ve hem yüksek doğrulukta çalışan otomatik sistemler hem de hekim kararlarını destekler.

Bu nedenle, literatür, derin öğrenme modellerinin Alzheimer tanısında klinik açıdan umut verici olduğunu göstermektedir. Bununla birlikte, klinik entegrasyon, çok merkezli, kapsamlı, etik açıdan güvenilir ve anlaşılabilir sistemler gerektirir. Mevcut çalışmalar önemli ilerlemeler kaydetmiş olmakla birlikte uzun vadeli validasyon ve veri çeşitliliği sınırlılıkları nedeniyle modellerin klinik uygulamalarının dikkatle değerlendirilmesi gerekmektedir.

Regülasyon kurumları tarafından onaylanmış AI tabanlı sistemler, klinik entegrasyon için çok önemlidir. FDA ve Avrupa İlaç Ajansı (European Medicines Agency - EMA), radyoloji ve kardiyoloji alanlarında derin öğrenme tabanlı karar destek yazılımlarını onaylamıştır, ancak Alzheimer için yaygın olarak onaylanmış bir AI sistemi henüz bulunmamaktadır. Bunun ana nedeni, modellerin çok merkezli, uzun vadeli ve uluslararası olarak onaylanması gerektiğidir (Topol, 2019). Mevcut literatürde, popülasyon çeşitliliği, açıklanabilirlik, veri güvenliği ve etik uygunluk gibi standartların MRI tabanlı Alzheimer sınıflandırıcılarının standartlaştırılması gerektiği sıkça vurgulanmaktadır.

Bu nedenle, derin öğrenme tabanlı Alzheimer tanı sistemlerinin ilerlemesi, daha yüksek doğruluk oranları elde etmek ve klinik süreçlere uygun entegrasyon stratejileri oluşturmak üzerine kurulu görünmektedir. Model tasarımının klinik gereksinimlere göre

şekillendirilmesi ve geliştirilen sistemlerin günlük iş akışına uyumlu hale getirilmesi için radyologlar, yazılım mühendisleri, nörologlar ve veri bilimcileri içeren çok disiplinli projeler kritik öneme sahiptir. Literatürde belirtildiği gibi, zaman içindeki bilişsel gerilemeyi takip eden uzun vadeli çalışmalar ve farklı merkezlerden toplanmış geniş ölçekli veri kümeleri, AI tabanlı tekniklerin klinik güvenilirliğini artırabilir. Bu tezde sunulan çoklu derin öğrenme mimarisi karşılaştırması, gelecekte klinik ortamlarda kullanılacak basit ve güvenilir karar destek sistemlerinin tasarımına katkıda bulunacak bir ön adım niteliği taşımaktadır.

Literatürde ayrıca, AI tabanlı Alzheimer tanı sistemlerinin gelecekte klinik iş akışına entegrasyonunun, yalnızca doğruluk artışıyla değil, zaman yönetimi, maliyet azaltımı ve hekimler arası standartizasyon gibi alanlarda sağlayacağı faydalarla da ilişkilendirildiği belirtilmektedir. Özellikle otomatik görüntü analizi sistemlerinin, yoğun klinik ortamlarda karar verme sürecini hızlandırması ve insan kaynaklı varyasyonu azaltması bakımından yapısal bir dönüşüm potansiyeli taşıdığı vurgulanmaktadır (McKinney et al., 2020; Topol, 2019). Bu nedenle Alzheimer özelinde geliştirilen derin öğrenme modelleri, gelecekte klinik karar süreçlerinde daha aktif rol üstlenme potansiyeline sahip, klinik karar sürecinin aktif bileşenleri olarak konumlanmaya aday görülmektedir.

3. MATERYAL VE YÖNTEM

Bu bölümde, tez çalışması kapsamında kullanılan veri kümesi, veri ön işleme adımları ve uygulanan derin öğrenme yöntemleri ayrıntılı olarak açıklanmaktadır. Ayrıca deneysel tasarım, model eğitimi süreci ve performans değerlendirme kriterleri bu bölüm altında sunulmaktadır.

3.1 Kullanılan Veri Kümesi

Bu çalışmada Alzheimer hastalığının farklı evrelerini sınıflandırmak amacıyla açık erişimli bir manyetik rezonans görüntüleme veri kümesi kullanılmıştır. Bu bölümde, kullanılan veri kümesinin genel özellikleri, içeriği ve sınıf yapısı açıklanarak çalışmanın deneysel altyapısı tanıtılmaktadır.

3.1.1 Veri seti özellikleri

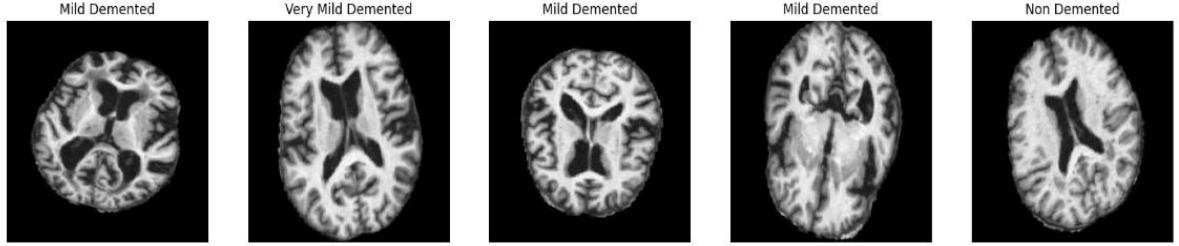
Bu tez çalışmasında, Alzheimer hastalığının çeşitli evrelerini MRI görüntüleri kullanılarak sınıflandırmak için açık erişimli OASIS-1 veri kümesi kullanılmıştır. Washington Üniversitesi tarafından oluşturulan yüksek çözünürlüklü T1-ağırlıklı beyin görüntülerinden oluşan bu veri seti, demans ve yaşlanma ile ilişkili nörodejeneratif değişiklikleri incelemeye yöneliktir. Veri kümesi, 18 ila 96 yaşları arasında değişen 416 bireye ait tek bir yapısal MRI taraması içermektedir ve tüm veriler anonimleştirilmiştir.

Katılımcılar, Klinik Demans Derecelendirme Ölçeği (Clinical Dementia Rating - CDR) ölçeğine göre dört sınıfa ayrılmıştır:

Non-Demented (CDR=0), Very Mild (CDR=0.5), Mild (CDR=1) ve Moderate Dementia (CDR=2).

Orijinal 3D (176×208×176) MRI hacimleri bu çalışma kapsamında CNN modellerine uygun olacak şekilde 2D koronal kesitlere dönüştürülmüş, 224×224 piksele yeniden ölçeklendirilmiş ve dört sınıf için klasör tabanlı bir yapıda organize edilmiştir. Oluşturulan nihai veri yapısı, eğitim-doğrulama-test kümelerine ayrılmış toplam 8268 görüntü içerir. Her alt kümede, eğitim için 5141, doğrulama için 1848 ve test için 1279 görüntü kullanılmıştır ve tüm sınıfların temsili korunmuştur.

OASIS veri seti tamamen anonim ve herkes tarafından erişilebilir olduğundan, bu çalışma için ek etik izin gerekmemektedir. Veri kümesinin örneklem büyüklüğü ve standartlaştırılmış yapısı, Alzheimer evrelerini sınıflandırmak için güvenilir ve yeniden üretilebilir bir araştırma zemini sağlar.



Şekil 3.1: OASIS veri setinde bulunan örnek MRI görüntüleri

3.1.2 Veri ön işleme adımları

Bu çalışmada kullanılan MRI görüntüleri, derin öğrenme modellerinin kararlı ve etkin biçimde eğitilebilmesi için bir dizi sistematik ön işleme adımından geçirilmiştir. Ön işlemenin temel amaçları; (i) tüm mimariler için ortak ve standart bir giriş yapısı oluşturmak, (ii) optimizasyon sürecini sayısal açıdan dengeli hâle getirmek, (iii) sınırlı veri koşullarında genelleme yeteneğini artırmak ve (iv) aşırı öğrenme (overfitting) riskini azaltmaktır. Bu kapsamda sırasıyla görüntü boyutlandırma, normalizasyon, veri artırma ve veri ayırma adımları uygulanmıştır.

3.1.2.1 Görüntü boyutlandırma (resizing)

OASIS veri kümesinde bulunan MRI görüntüleri, $176 \times 208 \times 176$ çözünürlükte ve üç boyutlu (3D) hacimsel veridir. Bu tezde, hacimsel görüntüler 2D koronal kesitlere dönüştürülerek sınıflandırma modellerine uygun hale getirilmiştir. Bu tercih, 3D modellerle çalışmanın yüksek hesaplama maliyeti ve uzun eğitim süreleri gerektirmesi nedeniyle yapılmıştır. Birleşik Fotoğraf Uzmanları Grubu (Joint Photographic Experts Group - JPEG) formatında kaydedilen 2 boyutlu görüntüler, tüm mimarilerde ortak bir giriş yapısı oluşturmak için 224×224 piksel çözünürlüğe ölçeklendirilmiştir.

Model giriş boyutları şu şekilde tanımlanmıştır:

- EfficientNet-B0, ResNet-50, MobileNetV2 için $224 \times 224 \times 3$
- InceptionV3 için $299 \times 299 \times 3$
- U-Net tabanlı sınıflandırıcı için $256 \times 256 \times 3$

224 × 224 çözünürlüklü girdiler, Keras'ın target_size parametresi ile ilgili modele uygun boyutlara yeniden ölçeklendirilmiştir. Yeniden boyutlandırma sırasında bilinear interpolasyon tercih edilerek görüntüdeki yapısal detayların korunması sağlanmıştır. Böylece hipokampus, entorhinal korteks ve lateral ventriküller gibi Alzheimer açısından kritik anatomik bölgelerin uzamsal bütünlüğü korunmuş ve modellerin tutarlı şekilde eğitilmesi mümkün olmuştur.

Bu yaklaşım, özellikle ince morfolojik değişimlerin öğrenilmesini destekleyerek erken evre Alzheimer ve MCI sınıflarının daha hassas biçimde ayırt edilmesine katkı sağlamayı amaçlamıştır.

3.1.2.2 Normalizasyon (rescaling)

Derin öğrenme modellerinin güvenilir bir şekilde eğitilebilmesi için görüntü piksellerinin belirli bir aralığa ölçeklenmesi gerekmektedir. Sonuç olarak, tüm MRI görüntülerinin ham yoğunluk değerleri önce [0, 1] aralığına indirgendikten sonra ilgili ön-işleme fonksiyonları kullanılmıştır.

ImageDataGenerator, EfficientNet-B0, ResNet-50 ve InceptionV3 modellerinde veri akışı sağlarken önce rescale = 1./255 uygulanmıştır. Her mimarinin preprocess_input fonksiyonu, ImageNet ağırlıklarıyla uyumlu olacak şekilde pikselleri [-1, 1] aralığına dönüştürmektedir. Böylece, modellerin transfer öğrenme aşamasında beklediği giriş formatı korunmuştur.

U-Net tabanlı sınıflandırıcıda ek bir ImageNet ön-işleme adımı kullanılmamıştır; yalnızca rescale = 1./255 işlemi kullanılmıştır. Bu nedenle, ağ, 0–1 aralığına ölçeklenmiş yoğunluk değerleri üzerinden özellik öğrenimini doğrudan bu ölçek üzerinden gerçekleştirmektedir.

MobileNetV2 için veri yükleme image_dataset_from_directory kullanılarak gerçekleştirilmiştir ve normalizasyon mobilenet_v2.preprocess_input fonksiyonu ile model girişinde belirtilen bir Lambda katmanı kullanılarak sağlanmıştır.

Bu süreç sayesinde her mimari, kendi önceden eğitilmiş ağıyla tamamen uyumlu biçimde normalize edilmiş giriş verileri ile beslenmiştir. Böylece optimizasyon sürecinin kararlılığı artırılmış ve mimariler arası adil bir karşılaştırma ortamı sağlanmıştır.

3.1.2.3 Veri artırma (data augmentation)

Etiketli tıbbi görüntüleme verilerinin elde edilmesi hem maliyetli hem de zaman alıcı bir süreç olduğundan, kullanılan veri kümelerinin büyüklüğü çoğu zaman derin öğrenme modellerinin sahip olduğu parametre sayısına kıyasla sınırlı kalmaktadır. Bu durum, özellikle yüksek kapasiteli derin ağlarda aşırı öğrenme (overfitting) riskini artırarak modelin genelleme performansını olumsuz yönde etkileyebilmektedir. Bu nedenle, çalışmada eğitim kümesinin görünür büyüklüğünü ve çeşitliliğini artırmak amacıyla veri artırma tekniklerinden yararlanılmıştır. Bu yaklaşımla, sınırlı sayıdaki MRI örneği üzerinden daha zengin ve çeşitli bir eğitim süreci oluşturulması hedeflenmiştir.

Eğitim sürecinde yalnızca eğitim veri kümesine veri artırma işlemleri uygulanmış; doğrulama ve test veri kümelerinde rastgele dönüşümler kullanılmamıştır. Doğrulama ve test aşamalarında yalnızca yeniden ölçeklendirme ve modele özgü normalizasyon adımları uygulanarak, modelin gerçekten daha önce görmediği örnekler üzerindeki performansının objektif biçimde değerlendirilmesi sağlanmıştır. Bu yaklaşım, elde edilen performans sonuçlarının yanlılıktan uzak ve güvenilir olmasına katkı sağlamıştır.

Çalışmada kullanılan veri artırma stratejileri iki temel teknik yaklaşım altında ele alınmıştır. EfficientNet-B0, ResNet-50, InceptionV3 ve U-Net modelleri için ImageDataGenerator tabanlı veri artırma yöntemi kullanılmıştır. Bu kapsamda, eğitim görüntüleri rastgele olarak $\pm 12^\circ$ aralığında döndürülmüş, yatay ve dikey eksenlerde yaklaşık $\pm 6\%$ oranında kaydırılmış, 12% 'ye kadar yakınlaştırma ve uzaklaştırma uygulanmış, parlaklık değerleri $\pm 10\%$ bandında değiştirilmiş ve hafif kesme (shear) dönüşümleri gerçekleştirilmiştir. Ayrıca görüntüler yatay ekseninde rastgele çevrilmiş ve dönüşümler sonucunda oluşan boşluklar en yakın piksel değerleri ile doldurulmuştur. Bu dönüşümler, beynin yaklaşık bilateral simetrik yapısı ve görüntü alım sürecinde oluşabilecek küçük konumsal farklılıklar dikkate alınarak seçilmiş; böylece modellerin yön, ölçek, parlaklık ve pozisyon değişimlerine karşı daha dayanıklı hâle gelmesi amaçlanmıştır.

MobileNetV2 mimarisi için ise katman tabanlı veri artırma yaklaşımı benimsenmiştir. Bu modelde veri artırma işlemleri, doğrudan ağ mimarisi içerisine gömülü Keras veri artırma katmanları aracılığıyla gerçekleştirilmiştir. Eğitim sırasında yatay ekseninde rastgele çevirme, yaklaşık ± 0.08 radyan aralığında döndürme, 12% 'ye kadar yakınlaştırma ve uzaklaştırma, yatay ve dikey eksenlerde yaklaşık $\pm 8\%$ oranında kaydırma ile kontrastın $\pm 15\%$ bandında

değiştirilmesi gibi dönüşümler çevrimiçi olarak uygulanmıştır. Bu sayede MobileNetV2 modeli, her eğitim iterasyonunda GPU üzerinde dinamik olarak çeşitlendirilmiş örnekler görmekte; ek bir ön işleme adımına ihtiyaç duymadan esnek ve ölçeklenebilir bir veri artırma stratejisi elde edilmektedir.

Sonuç olarak, sınırlı sayıda MRI örneği ile eğitilen modellerin farklı klinik çekim koşulları, hasta pozisyonları ve görüntülerdeki doğal varyasyonlar karşısında daha genelleyici, esnek ve kararlı performans sergilemesi sağlanmıştır. Uygulanan veri artırma yöntemleri, modellerin gerçek klinik senaryolara daha iyi uyum göstermesine önemli ölçüde katkıda bulunmuştur.

3.1.2.4 Veri ayrımı (split strategy)

Model performansının güvenilir bir şekilde değerlendirilebilmesi için, veri kümesi eğitim sürecinde kullanılan örneklerden ve nihai değerlendirmede kullanılan örneklerden ayrılmalıdır. Bu çalışmada OASIS veri kümesi üç alt kümeye ayrılmıştır:

- Alzheimer_sDataset/train/
- Alzheimer_sDataset/validation/
- Alzheimer_sDataset/test/

Bu, eğitim, doğrulama ve test kümelerini fiziksel olarak birbirinden ayırmaktadır ve tüm modeller için aynı dizin yapısı kullanılmıştır. Toplam 8268 görüntü içeren veri kümesi şu şekilde dağıtılmıştır:

- Eğitim kümesi (train): 5141 görüntü (yaklaşık %62),
- Doğrulama kümesi (validation): 1848 görüntü (yaklaşık %22),
- Test kümesi (test): 1279 görüntü (yaklaşık %15).

Her üç alt kümede de dört sınıfın (Non-Demented, Very Mild Demented, Mild Demented, Moderate Demented) temsili korunmuştur; böylece modellerin her evre için dengeli bir biçimde öğrenme ve değerlendirme yapması sağlanmıştır.

Veri yükleme sürecinde, kullanılan derin öğrenme mimarilerine bağlı olarak iki farklı Keras arayüzünden yararlanılmıştır. EfficientNet-B0, ResNet-50, InceptionV3 ve U-Net tabanlı sınıflandırıcı modeller için, ImageDataGenerator sınıfına ait flow_from_directory() metodu kullanılarak belirtilen klasör yapılarından veri okunmuş; görüntüler target_size = (224, 224)

veya ilgili mimarinin özgün giriş boyutuna (örneğin InceptionV3 için 299×299 , U-Net için 256×256) yeniden boyutlandırılmıştır. Bu süreçte `batch_size = 16` ve `class_mode = 'categorical'` parametreleri kullanılarak NumPy tabanlı mini-batch'ler oluşturulmuştur. MobileNetV2 modeli için ise aynı klasör yapısı `image_dataset_from_directory()` fonksiyonu aracılığıyla okunmuş; `image_size = (224, 224)`, `batch_size = 16` ve `label_mode = "categorical"` parametreleri ile `tf.data.Dataset` tabanlı bir veri işleme hattı (pipeline) elde edilmiştir. Ayrıca, bu veri kümeleri `prefetch()` yöntemi ile ön getirmeli hâle getirilerek eğitim sürecinde GPU kaynaklarının daha verimli kullanılması sağlanmıştır.

Eğitim sürecinde modelin ağırlıklarının güncellenmesi yalnızca eğitim kümesi tarafından yapılmıştır. Diğer taraftan doğrulama kümesi, her evrenin sonunda performansı izlemek, erken durdurma (EarlyStopping) ve öğrenme oranı azaltma (ReduceLRonPlateau) gibi geri besleme mekanizmalarını çalıştırmak için kullanılmıştır. Eğitim tamamlandıktan sonra test kümesi tek seferlik olarak değerlendirilmiştir. Bu, modellerin daha önce hiç görmedikleri görüntüler üzerindeki gerçek genelleme başarısını (doğruluk, duyarlılık, özgüllük, F1 skoru ve ROC-AUC) doğru bir şekilde rapor etmeyi mümkün kılmıştır.

3.2 Model Eğitimi Genel Çerçeve

Bu bölümde, çalışmada kullanılan derin öğrenme modellerinin eğitim sürecine ilişkin genel çerçeve sunulmaktadır. Model eğitimi sırasında izlenen adımlar, kullanılan eğitim stratejileri ve deneysel sürecin genel akışı bu başlık altında açıklanmaktadır.

3.2.1 Donanım ve yazılım altyapısı

Bu tez kapsamında yürütülen tüm model eğitimleri ve performans değerlendirmeleri, GPU hızlandırmalı bir bulut ortamı sunan Google Colab Pro üzerinde gerçekleştirilmiştir. Çalışmalarda kullanılan sistem, NVIDIA Tesla T4 grafik işlemcisi (16 GB VRAM), 25 GB RAM ve Ubuntu tabanlı sanal bir işletim sistemi üzerinde yapılandırılmıştır. Bu donanım, yüksek çözünürlüklü MRI görüntülerinin işlenmesinde gerekli paralel hesaplama kapasitesini sağlamış ve eğitim sürecinin stabil şekilde yürütülmesine olanak tanımıştır.

Yazılım ortamı Python 3.10 üzerine kurulmuş olup, TensorFlow 2.15 ve Keras API, derin öğrenme modellerinin geliştirilmesi ve eğitimi için temel çerçeveyi oluşturmuştur. NumPy, Pandas, Matplotlib ve scikit-learn gibi yardımcı kütüphaneler veri işleme, görselleştirme ve metrik hesaplamalarında kullanılmıştır. Tüm MRI görüntüleri Google Drive üzerinde

organize edilmiştir ve ve modellerin eğitim sürecine ImageDataGenerator veya tf.data tabanlı veri işleme hatları (pipeline) aracılığıyla entegre edilmiştir.

Veri artırma, ön-işleme ve etiket yönetimi TensorFlow tarafından otomatikleştirilmiş; dört sınıfa ait etiketler Demans olmayan bireyler (Non-Demented – ND), çok hafif demans (Very Mild Demented – VMD), hafif demans (Mild Demented – MD) ve orta düzey demans (Moderate Demented – MOD) dizin yapısından türetilerek eğitim süreci standartlaştırılmıştır.

3.2.2 Hiperparametre ayarları

Derin öğrenme modellerinin başarımı, yalnızca mimarilerin yapısal özelliklerine değil, aynı zamanda eğitim sürecinde kullanılan hiperparametrelerin doğru ve dengeli biçimde seçilmesine doğrudan bağlıdır. Bu tez kapsamında kullanılan tüm modeller (EfficientNet-B0, ResNet-50, MobileNetV2, InceptionV3 ve U-Net tabanlı encoder mimarisi) aynı eğitim protokolü altında yapılandırılmış; böylece modeller arasında adil, karşılaştırılabilir ve metodolojik olarak tutarlı bir değerlendirme sağlanmıştır. Bu bölümde, model eğitim sürecinde kullanılan hiperparametre ayarları ayrıntılı biçimde sunulmaktadır.

3.2.2.1 Optimizasyon algoritması

Bu çalışmada tüm modellerin ağırlık güncellemelerinde Adaptive moment estimation (Adam) optimizasyon algoritması kullanılmıştır. Adaptif öğrenme hızı ve moment tahminlerini birleştiren Adam, özellikle gürültülü gradyanlarda kararlı bir öğrenme süreci sağladığı için derin CNN mimarileriyle uyumlu bir yapı sunmaktadır. Eğitim süreci iki aşamalı olarak yapılandırılmıştır. İlk aşamada yalnızca üst sınıflandırıcı katmanlar eğitilmiş, taban mimariler dondurularak hızlı ve stabil bir başlangıç elde edilmiştir (EfficientNet-B0 ve MobileNetV2 için 1×10^{-3} ; ResNet-50 ve InceptionV3 için 8×10^{-4} öğrenme oranı). İkinci aşamada ince ayar (fine-tuning) uygulanmış ve mimarilerin son katmanları düşük öğrenme oranlarıyla yeniden eğitime açılmıştır (EfficientNet-B0, ResNet-50 ve InceptionV3 için 1×10^{-5} ; MobileNetV2 için 3×10^{-5}). Her iki aşamada da clipnorm=1.0 ile gradyan patlamalarının önüne geçilmiş ve eğitim sürecinde kararlılık sağlanmıştır.

Bu iki aşamalı optimizasyon stratejisi sayesinde, transfer öğrenmeden elde edilen ön bilgi korunurken, model parametrelerinin veri kümesine özgü biçimde hassas olarak uyarlanması sağlanmıştır.

3.2.2.2 Kayıp fonksiyonu (loss function)

Çalışma, çok sınıflı bir sınıflandırma problemi içerdiğinden, kategorik çapraz entropi temel kayıp fonksiyonu olarak seçilmiştir. Bu fonksiyon, her bir sınıf için tahmin edilen olasılığın optimize edilmesini sağlamaktadır ve softmax çıktı katmanı ile birlikte çalışmaktadır.

Kullanılan temel kayıp fonksiyonu matematiksel olarak aşağıdaki gibidir:

$$L = - \sum_{i=1}^N y_i \log(\hat{y}_i) \quad (3.1)$$

Model performansını iyileştirmek ve sınıf dengesizliğini azaltmak için temel kayıp fonksiyonuna label smoothing (0.05 veya 0.10) ve Focal Loss ($\gamma = 2.0$) bileşenleri eklenmiştir. Focal Loss, modelin ayrıştırıcı öğrenme yeteneğini geliştirerek zor sınıflandırılan örneklere daha fazla önem verir. Veri kümesindeki sınıf dağılımına bağlı olarak, alfa değerleri dinamik olarak hesaplanmıştır.

3.2.2.3 Performans metrikleri

Modelin başarısı yalnızca ACC ile değil, aynı zamanda sınıf dengesizliğine duyarlı performans ölçümleri ile birlikte değerlendirilmiştir.

Bu çerçevede, tüm modellerde aşağıdaki ölçümler kullanılmıştır:

- ACC: Genel sınıflandırma başarısını gösterir.
- Precision (Kesinlik): Pozitif tahminlerin doğruluk oranını gösterir.
- Recall (Duyarlılık): Gerçek pozitif örneklerin ne kadarının doğru sınıflandırıldığını ölçer.
- AUC: Sınıflar arası ayırım gücünü değerlendirir ve makro ortalama olarak hesaplanmıştır.

Alzheimer evrelerinin ayrıştırılmasında modelin hem genel hem de sınıf bazlı başarısı, bu metriklerin birlikte değerlendirilmesiyle doğrulanmıştır. Özellikle sınıf dağılımının dengeli olmaması durumunda yalnızca ACC değerinin yanıltıcı sonuçlar üretebileceği göz önünde bulundurulmuştur. Bu nedenle precision ve recall metrikleri, her bir evre için modelin hatalı pozitif ve hatalı negatif tahmin davranışlarını analiz etmek amacıyla ayrı ayrı incelenmiştir.

Ayrıca AUC değeri, modelin farklı eşik değerleri altında sınıflar arası ayırım yeteneğini değerlendirmek için kullanılmış ve makro ortalama yaklaşımıyla tüm sınıflar eşit ağırlıkta ele alınmıştır. Böylece performans değerlendirme süreci, yalnızca tek bir metrik üzerinden değil, çok boyutlu bir analiz çerçevesinde gerçekleştirilmiştir. Bu yaklaşım, klinik karar destek sistemlerinde güvenilir ve dengeli bir model seçiminin yapılabilmesi açısından kritik öneme sahiptir.

3.2.2.4 Epoch sayısı ve erken durdurma stratejisi

Bu çalışmada tüm modeller toplam 50 epoch boyunca eğitilmiştir ve eğitim süreci iki aşamalı olarak yapılandırılmıştır: İlk 10 epoch'ta yalnızca üst sınıflandırıcı katmanlar eğitilmiş, ardından 40 epoch süren fine-tuning aşamasında taban mimarilerin son katmanları yeniden eğitime açılmıştır. Aşırı öğrenmeyi önlemek amacıyla EarlyStopping mekanizması kullanılmış; doğrulama doğruluğunun ("val_acc") 12 epoch boyunca iyileşmemesi durumunda eğitim otomatik olarak sonlandırılmıştır. Böylece hem hesaplama süresi optimize edilmiş hem de en iyi doğrulama performansına sahip ağırlıklar restore_best_weights=True seçeneğiyle otomatik olarak geri yüklenmiştir.

3.2.2.5 Öğrenme oranı planlaması (learning rate scheduling)

Tüm modeller, adaptif olarak öğrenme hızını azaltmak için ReduceLRonPlateau mekanizması kullanılmıştır. Parametreler:

- monitor: "val_loss"
- factor: 0.5
- patience: 5
- min_lr: 1×10^{-6}

Bu yaklaşım, doğrulama kaybı belirli bir süre boyunca iyileşmediğinde öğrenme oranının yarıya indirilerek daha hassas bir optimizasyon aşamasına geçilmesini sağlamıştır.

3.2.2.6 Model ağırlıklarının kaydedilmesi

ModelCheckpoint uygulaması, en iyi doğrulama doğruluğunun elde edildiği noktayı kaybetmek için kullanılmıştır. Bu mekanizma, doğrulama doğruluğunu ("val_acc") izleyerek modelin ağırlık dosyasını yalnızca performans artışı sağlandığında kaydetmektedir. Bu sayede her model için yalnızca en yüksek doğrulama başarımına sahip

ağırlık dosyası (.keras formatında) kaydedilmiştir. Tüm modeller test ve değerlendirme aşamalarında bu optimum ağırlıklar kullanılarak analiz edilmiştir.

3.3 Kullanılan Derin Öğrenme Mimarileri

Bu tez çalışmasında, Alzheimer hastalığının MRI görüntüleri üzerinden sınıflandırılmasına yönelik olarak literatürde başarıyla uygulanmış beş farklı derin öğrenme mimarisi tercih edilmiştir. Seçilen yüksek sınıflandırma performansı ve hesaplama verimliliği sunmakta olup transfer öğrenme senaryolarına uygundur. ImageNet veri seti üzerinde önceden eğitilmiş ağırlıklarla başlatılan her mimari, Alzheimer sınıflandırmasına özel olarak son katmanları yeniden yapılandırılmıştır.

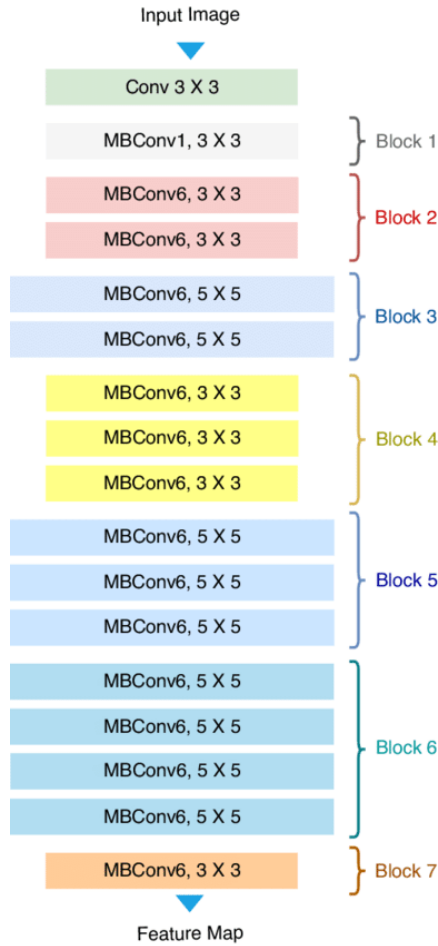
Bu çalışmada kullanılan her derin öğrenme mimarisinin temel özellikleri aşağıda açıklanmıştır. Ayrıca, Alzheimer sınıflandırması bağlamında her bir derin öğrenme mimarisinin işlevi ve yapılandırma ayrıntıları da verilmiştir.

3.3.1 EfficientNet : Ölçeklenebilir ve parametre verimli derin CNN mimarisi

EfficientNet, Tan ve Le (2019) tarafından geliştirilen modern bir CNN mimarisi olup, sınıflandırma görevlerinde yüksek doğruluk ve düşük hesaplama maliyeti arasındaki dengeyi optimize eden bir yapı sunmaktadır. Literatürde yaygın olarak kullanılan geleneksel mimarilerde modelin kapasitesini artırmak amacıyla genellikle tek bir eksen – örneğin yalnızca katman sayısını (derinliği) artırmak – ölçekleme yapılırken, EfficientNet bu yaklaşımı değiştirerek çok eksenli ölçekleme (compound scaling) yöntemini benimsemiştir. Bu ölçekleme stratejisi; modelin derinlik (depth), genişlik (width) ve çözünürlük (resolution) boyutlarını birlikte optimize ederek daha dengeli bir performans artışı sağlar. Bu sayede, minimum parametre artışı ile daha yüksek doğruluk elde edilebilmekte, modelin hem eğitim süresi hem de bellek tüketimi açısından verimli kalması sağlanmaktadır.

Bu mimarinin sunduğu yapısal avantajlar, tıbbi görüntü işleme gibi hassas ve veri sınırlı alanlarda özellikle değer kazanmaktadır. EfficientNet mimarisi, çeşitli ölçeklerde (B0–B7) farklı büyüklüklerdeki modellerle sunulmakta olup, bu çalışmada temel yapı olan EfficientNet-B0 tercih edilmiştir. Bu model, sınırlı kaynakla çalışan klinik uygulamalarda veya mobil cihazlarda çalıştırmaya elverişli olacak şekilde tasarlanmış, düşük parametre sayısı ile dikkat çeken bir versiyondur.

Modelin giriş boyutu $224 \times 224 \times 3$ olarak belirlenmiştir. Aşağıda, EfficientNet-B0 mimarisinin katman bazlı şematik yapısı gösterilmiştir. Şemada, her Mobil ters şişe boynu evrişim bloğu (Mobile Inverted Bottleneck Convolution - MBConv) bloğunun türü, kernel boyutları ve çıkış şekilleri sırasıyla sunulmuştur.



Şekil 3.2: EfficientNet-B0 mimarisinin katman bazlı gösterimi (Gouji et al., 2022'den uyarlanmıştır)

Şekilde, EfficientNet-B0 mimarisine ait konvolüsyon blokları ve katmanlar sırasıyla gösterilmiştir. Modelin giriş ve çıkış boyutları, kullanılan kernel boyutları ve her MBConv bloğunun türü şematik olarak sunulmuştur. Görsel Taheri Gorji et al. (2022) çalışmasından uyarlanmıştır.

3.3.1.1 Uygulanan EfficientNet-B0 yapılandırması

Bu çalışmada EfficientNet mimarisinin B0 varyantı, transfer öğrenme yaklaşımı temel alınarak ve iki aşamalı bir eğitim stratejisiyle yapılandırılmıştır. Temel ağ (convolutional base), ImageNet veri kümesi üzerinde önceden eğitilmiş ağırlıklarla başlatılmış, include_top

= False parametresi kullanılarak özgün sınıflandırma katmanı mimariden çıkarılmış ve bunun yerine Alzheimer hastalığının evrelerine özgü bir sınıflandırma başlığı eklenmiştir. Bu sayede EfficientNet-B0'ın genel görsel öznitelik çıkarım kapasitesi korunmuş, sınıflandırma işlemi ise hedef probleme özgü üst katmanlar üzerinden gerçekleştirilmiştir. Bu tezde kullanılan sınıflandırma başlığı, Global Average Pooling katmanı ile başlamakta; ardından He uniform başlatıcı ve L2 düzenleme terimi (L2 regularization - L2) (1×10^{-5}) ile yapılandırılmış 256 nöronlu tam bağlantılı (Dense) katman yer almaktadır. Bu katmanı Batch Normalization ve ReLU aktivasyon fonksiyonu izlemekte, aşırı öğrenmeyi azaltmak amacıyla %40 oranında Dropout katmanı uygulanmaktadır. Son olarak model, Alzheimer hastalığının dört evresini temsil eden dört sınıflı softmax çıkış katmanı ile sonlandırılmıştır. Modelin giriş boyutu, RGB formatındaki iki boyutlu MRI kesitleri için $224 \times 224 \times 3$ olarak belirlenmiş; sınıf sayısı ND, VMD, MD ve MOD olmak üzere dört olarak tanımlanmıştır.

Görüntü boru hattında, giriş görüntüleri öncelikle rescale = 1/255 parametresi ile [0,1] aralığına ölçeklendirilmiş, ardından EfficientNet-B0 mimarisine özgü preprocess_input fonksiyonu kullanılarak standardizasyon işlemi uygulanmıştır. Eğitim sürecinde döndürme, kaydırma, yakınlaştırma (zoom), parlaklık değişimi, kesme (shear) ve yatay çevirme gibi veri artırma tekniklerinden yararlanılmış; sınıf dengesizliği problemi ise eğitim sırasında class_weight kullanılarak dengelenmiştir. Ayrıca, aşırı öğrenmeyi sınırlamak amacıyla L2 ağırlık cezalandırma (L2 regularization) yöntemi tercih edilmiştir.

Bu yapılandırma sayesinde EfficientNet-B0 mimarisi, temel katmanlarda genel görsel öznitelik bilgisinden etkin biçimde yararlanırken, üst katmanlarda Alzheimer hastalığına özgü ayırt edici temsilleri öğrenmeyi hedefleyen dengeli ve verimli bir sınıflandırma modeli hâline getirilmiştir.

3.3.1.2 EfficientNet-B0 eğitim stratejisi ve optimizasyon parametreleri

EfficientNet-B0 modeli, iki aşamalı bir eğitim stratejisiyle optimize edilmiştir. İlk aşamada, modelin yalnızca üst sınıflandırıcı katmanı eğitilmiş; temel EfficientNet-B0 gövdesi dondurularak (trainable=False) hızlı ve kararlı bir başlangıç elde edilmiştir. Bu süreçte Adam optimizasyon algoritması 1×10^{-3} öğrenme hızıyla kullanılmış, 10 epoch boyunca EarlyStopping ve ModelCheckpoint mekanizmaları doğrulama doğruluğunu (val_acc) izleyerek en iyi ağırlıkları kaydetmiştir. Ayrıca, ReduceLROnPlateau yöntemi doğrulama

kayıbı iyileşmediğinde öğrenme hızını otomatik olarak azaltarak daha stabil bir optimizasyon sağlamıştır.

İkinci aşamada, modelin temelindeki toplu normalizasyon (Batch Normalization - BN) katmanları dahil olmak üzere yeniden eğitime açılmıştır ve 40 epoch boyunca ince ayar uygulandı. Düşük bir öğrenme hızı (1×10^{-5}) uygulandı. Eğitimde, sınıf dengesizliğine duyarlılığı artırmak için odaklanma kaybı ($\gamma = 2.0$) ve hafif etiket düzeltmesi ($\epsilon = 0.05$) kullanıldı. Sınıf frekanslarından elde edilen α katsayılarıyla kayıp fonksiyonu ağırlıklandırıldı. Modelin AUC, Precision, Recall, F1 ve ACC sınıflandırma metrikleri kullanılarak performansı değerlendirilmiştir.

Tablo 3.1: EfficientNet-B0 modeli için kullanılan eğitim hiperparametreleri

| Hiperparametre | Değer |
|-----------------------------------|---|
| Optimizasyon | Adam ($\beta_1 = 0.9, \beta_2 = 0.999$) |
| Öğrenme Oranı (Aşama-1 / Aşama-2) | $1e-3 \rightarrow 1e-5$ |
| Batch Boyutu | 16 |
| Toplam Epoch | 50 (10 + 40) |
| Kayıp Fonksiyonu | Focal loss ($\gamma = 2.0$) + label smoothing (0.05) + class weights |
| Raporlama | Doğrulama ve test kümesi için sınıflandırma raporu, karışıklık matrisi ve macro-AUC |
| Erken Durdurma | monitor = val_acc, patience = 12, en iyi model kaydı |
| Girdi Boyutu | $224 \times 224 \times 3$ |
| Fine-Tuning | Son 120 katman (Batch Normalization katmanları hariç) |
| BatchNorm Katmanları | Freeze (trainable = False) |
| Veri Artırma | rotation, width shift, height shift, zoom, brightness, shear, horizontal flip |

Bu iki aşamalı strateji, önce üst katmanlarda hızlı yakınsama elde edip ardından düşük öğrenme oranı (Learning Rate – LR) ile temel ağın son bloklarında selektif fine-tuning yaparak performansı artırmayı hedeflemektedir. Focal loss + sınıf ağırlıkları birlikteliği, az temsil edilen evrelerde duyarlılığı iyileştirmektedir (Lin et al., 2017).

3.3.1.3 EfficientNet-B0 mimarisi teknik avantajlar ve sınırlılıklar

EfficientNet-B0 mimarisi, Alzheimer sınıflandırması gibi karmaşık tıbbi görüntüleme sorunlarında maliyet-performans açısından oldukça dengeli bir profil sunmaktadır. Compound scaling yaklaşımı, derinlik, genişlik ve çözünürlüğü aynı anda optimize etmektedir. Bu, çok daha az parametre kullanarak benzer doğruluk seviyelerininin elde

edilmesini sağlamaktadır. Bu nedenle, yaklaşık 5.3 milyon parametreye sahip EfficientNet-B0, literatürde tartışılan geleneksel CNN mimarilerine kıyasla önemli ölçüde hesaplama gücü tasarrufu sağlamaktadır. Bu durum, GPU belleğinin sınırlı olduğu ya da mobil/edge tabanlı klinik uygulamalarda modelin daha verimli çalışabilmesine olanak tanımaktadır. ImageNet üzerinde önceden eğitilmiş ağırlıklarla uyumlu yapısı sayesinde transfer öğrenme senaryolarında iyi bir başlangıç noktası sağlamıştır ve sınırlı sayıda MRI örneğiyle bile yüksek bir genelleme başarısı elde etmiştir.

Bu çalışmada kullanılan Focal Loss yapısı ($\gamma = 2.0$), özellikle az temsil edilen Alzheimer evrelerinde duyarlılığı artırmıştır. Bu nedenle, modelin küçük sınıflara ait örnekleri daha etkili biçimde öğrenmesini sağlamıştır ve klinik açıdan önemli erken ve orta evrelerde başarılı sonuçlar vermiştir. Ek olarak, Grad-CAM gibi görselleştirme yöntemleriyle modelin odaklandığı bölgeleri analiz etmek mümkündür. Bu, modelin klinik açıklanabilirlik boyutunu artırmıştır.

EfficientNet-B0 mimarisinde bazı sınırlamalar vardır. Parametre verimliliği yüksek olsa da B0 varyantının karmaşık ve yüksek çözünürlüklü MRI görüntülerindeki ince yapısal farklılıkları tam olarak ayırt etmekte yetersiz kalabileceği göz önünde bulundurulmalıdır. EfficientNet-B4–B7 gibi daha derin versiyonlar daha güçlü temsiller sunabilir. Model ayrıca veri artırma stratejilerine çok duyarlı; agresif artırma parametreleri seçilirse genelleme performansı düşebilir. Swish aktivasyonu ve Squeeze-and-Excitation blokları, özellikle CPU tabanlı ortamlarda hesaplama yükünü artırsa da, GPU hızlandırması bu maliyeti önemli ölçüde azaltmıştır. Tüm bu faktörler göz önüne alındığında, EfficientNet-B0, düşük maliyet-yüksek performans dengesini başarıyla sağlayan ve dengesiz sınıf dağılımlarına karşı dayanıklı bir mimari olarak bu tez çalışmasında uygun bir çözüm sağlamıştır.

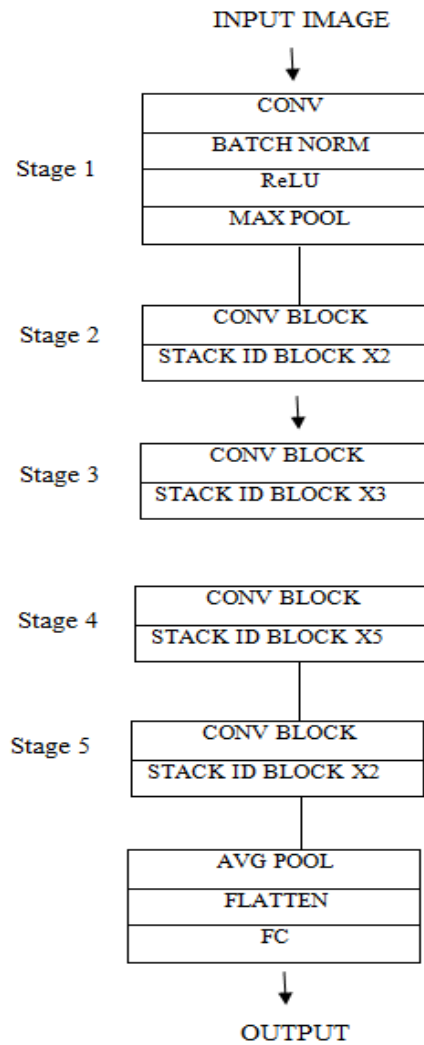
3.3.2 ResNet: Derin ağlarda rezidüel öğrenme yaklaşımı

ResNet, derin sinir ağlarında katman sayısı arttıkça ortaya çıkan gradyan sönümlenmesini ve optimizasyon zorluklarını azaltmak için yeni bir CNN mimarisi olarak tasarlanmıştır (He et al., 2016). Mimari, katmanlar arasında atlama bağlantıları kullanarak modelin doğrudan karmaşık dönüşümleri öğrenmek yerine yalnızca artık bileşeni (residual) modellemesine olanak tanımaktadır. Böylece derin ağların öğrenilebilirliği artar, gradyan akışı korunur ve öğrenme süreci daha kararlı hale gelir.

Bu çalışmada kullanılan ResNet-50, “bottleneck” olarak adlandırılan üç aşamalı konvolüsyon bloklarından oluşan 50 katmanlı bir derin öğrenme mimarisidir. Bottleneck yapısı, 1×1 konvolüsyon ile kanal sayısını azaltıp 3×3 konvolüsyon ile uzamsal özellikleri çıkardıktan sonra kanal genişliğini 1×1 konvolüsyon ile yeniden artırarak hem hesaplama maliyetini düşürmektedir hem de güçlü bir özellik çıkarma mekanizması sağlamaktadır.

ResNet-50'nin temel avantajları; derinliği sayesinde yüksek seviyeli anlamsal özellikleri öğrenebilmesi, bottleneck blokları sayesinde parametrik verimlilik sağlaması ve skip connection yapısı sayesinde daha stabil bir optimizasyon süreci sunmasıdır. Tıbbi görüntü sınıflandırma gibi yüksek duyarlılık gerektiren görevlerde, model bu özellikleriyle güçlü bir alternatif olarak değerlendirilmektedir.

Modelin genel katman yapısı Şekil 3.3'te özetlenmiştir.



Şekil 3.3: ResNet-50 mimarisinin genel katman yapısı (Gleason et al., 2023)

Modelin `include_top=False` seçeneğiyle önceden eğitilmiş taban (base) katmanları korunmuş, yalnızca sınıflandırıcı bloğu özelleştirilmiştir.

3.3.2.1 Uygulanan ResNet-50 yapılandırması

Bu tez kapsamında kullanılan ResNet-50 mimarisi, transfer öğrenme yaklaşımı çerçevesinde yapılandırılmıştır. Model, ImageNet veri kümesi üzerinde önceden eğitilmiş ağırlıklar ile çağrılmış ve `include_top = False` parametresi kullanılarak özgün sınıflandırma katmanı mimariden çıkarılmıştır. Bu sayede ResNet-50'nin güçlü özellik çıkarım (feature extraction) kapasitesi korunmuş; Alzheimer hastalığının evrelerine özgü sınıflandırma işlemi yalnızca bu çalışmaya özel olarak tasarlanan üst katmanlar (classification head) aracılığıyla gerçekleştirilmiştir.

Eğitimin ilk aşamasında, modelin konvolüsyonel tabanı dondurulmuş ve yalnızca üst sınıflandırma katmanları eğitilmiştir. Bu strateji, eğitim süresini önemli ölçüde azaltırken aynı zamanda sınırlı veri koşullarında aşırı öğrenme riskinin düşürülmesine katkı sağlamıştır. Takip eden aşamada ise, modelin son 100 katmanı Batch Normalization katmanları hariç tutulacak şekilde yeniden eğitilebilir hâle getirilmiş ve böylece ince ayar (fine-tuning) süreci uygulanmıştır. Bu yaklaşım, modelin Alzheimer evrelerine özgü daha ayırt edici temsiller öğrenmesini mümkün kılmıştır.

ResNet-50 mimarisinin üzerine eklenen özelleştirilmiş sınıflandırma başlığı, Alzheimer hastalığının dört evresini ayırt edecek şekilde yapılandırılmıştır. Bu yapı, sırasıyla önceden eğitilmiş ResNet-50 tabanı, Global Average Pooling katmanı, Batch Normalization, %40 oranında Dropout, 512 nöronlu ve He uniform başlatıcı ile yapılandırılmış L2 düzenlemeli tam bağlantılı (Dense) katman, yeniden Batch Normalization ve ReLU aktivasyon fonksiyonu, ikinci bir %40 Dropout katmanı ve son olarak dört sınıflı softmax aktivasyonuna sahip çıkış katmanından oluşmaktadır. Bu tasarım, hem modelin ifade gücünü artırmak hem de aşırı öğrenme eğilimini sınırlamak amacıyla seçilmiştir.

Modelin giriş boyutu, RGB formatındaki MRI görüntüleri için $224 \times 224 \times 3$ olarak belirlenmiş; çıkış katmanı Alzheimer hastalığının dört evresini temsil edecek şekilde dört sınıflı softmax yapısında tasarlanmıştır.

Ön eğitim ağırlıkları olarak ImageNet kullanılmış ve eğitim süreci transfer learning ile fine-tuning stratejilerinin birleşimi şeklinde yürütülmüştür. Bu yapı, tez kapsamında kullanılan “ResNet-50 SAFE MODE” Python kodu ile birebir uyumludur. Özellikle çıkış bloğunda kullanılan Dropout (0.4), Batch Normalization ve L2 regularizasyonu kombinasyonu, sınırlı veri koşullarında daha kararlı bir genelleme performansı elde etmek ve overfitting riskini azaltmak amacıyla tercih edilmiştir.

3.3.2.2 ResNet-50 eğitim stratejisi ve optimizasyon parametreleri

Bu çalışmada ResNet-50 mimarisi iki aşamalı bir eğitim stratejisi kullanılarak optimize etmiştir. İlk aşamada yalnızca özelleştirilmiş sınıflandırıcı katmanlar eğitilmiştir. İkinci aşamada, modelin son katmanları yeniden eğitilebilir hâle getirilmek için ince ayar yapılmıştır. Bu yöntem, eğitimin kararlılığını artırmış ve sınırlı veriyle yüksek doğruluk elde etmiştir. Bu sayede, modelin doğrulama kümesinde yüksek AUC, recall ve ACC değerleri elde etmesini sağlamıştır.

Aynı zamanda iki aşamalı optimizasyon, erken durdurma ve öğrenme hızı azaltma mekanizmalarını kullanarak aşırı öğrenme (overfitting) riskini azaltmıştır ve modelin genelleme performansını iyileştirmiştir.

Ayrıca, ResNet-50 mimarisinde kullanılan artık bağlantılar (skip connections), derin katmanlarda gradyan sönümlenmesi problemini azaltarak daha etkili bir özellik öğrenimi sağlamıştır. İnce ayar sürecinde düşük öğrenme oranlarının tercih edilmesi, önceden öğrenilmiş temsil uzayının korunmasına katkı sunmuş ve modelin veri kümesine aşırı uyum sağlamasını engellemiştir. Bu yapılandırma sayesinde ResNet-50, özellikle sınıflar arası geçişlerin belirsiz olduğu Alzheimer evrelerinde daha dengeli ve istikrarlı tahminler üretebilmiştir.

Elde edilen sonuçlar, mimarinin klinik açıdan güvenilir ve genellenebilir bir performans sunduğunu göstermektedir. Ayrıca, bu dengeli optimizasyon yaklaşımı, ResNet-50'nin farklı veri dağılımlarında da benzer performans sergilemesini sağlayarak modelin klinik araştırmalarda ve gerçek dünya uygulamalarında güvenle kullanılabilmesine olanak tanımıştır.

Bu bağlamda uygulanan eğitim stratejisi, modelin hem öğrenme kapasitesini etkin biçimde kullanmasına hem de sınırlı veri seti koşullarında kararlı bir performans sergilemesine olanak sağlamıştır. Özellikle iki aşamalı eğitim yaklaşımı ve ince ayar sürecinin birlikte kullanılması, modelin önceden öğrenilmiş genel görsel özellikleri koruyarak Alzheimer hastalığının farklı evrelerine ait ayırt edici örüntüleri daha doğru şekilde öğrenmesini sağlamıştır. Bu durum, modelin yalnızca eğitim verisi üzerinde değil, aynı zamanda doğrulama ve test verileri üzerinde de tutarlı sonuçlar üretmesine katkı sağlayarak önerilen yöntemin güvenilirliğini ve uygulanabilirliğini desteklemektedir.

Tablo 3.2: ResNet-50 modeli için eğitim hiperparametreleri

| Hiperparametre | Değer |
|------------------|--|
| Optimizasyon | Adam (Stage-1: lr = 8×10^{-4} , Stage-2: lr = 1×10^{-5}) |
| Öğrenme Oranı | ReduceLRonPlateau (factor = 0.5, patience = 5, min_lr = 1e-6) |
| Batch Boyutu | 16 |
| Toplam Epoch | 50 (10 + 40 aşamalı fine-tuning) |
| Kayıp Fonksiyonu | Focal Loss ($\gamma = 2.0$) + label smoothing (0.05) + class-frequency weighting |
| Raporlama | Doğrulama ve test raporu, karışıklık matrisi, macro-AUC |
| Fine-Tuning | Erken Durdurma: EarlyStopping (monitor = val_acc, patience = 12, restore_best_weights = True) |
| Veri Artırma | Son 100 katman açıldı (BatchNorm hariç) rotation_range = 12° , width_shift = 0.06, height_shift = 0.06, zoom_range = 0.12, brightness_range = [0.9, 1.1], shear_range = 0.05, horizontal_flip = True, fill_mode = "nearest" |

3.3.2.3 ResNet-50 mimarisi: teknik avantajlar ve sınırlılıklar

ResNet-50, derin yapısı ve rezidüel bağlantılara dayalı yapısı nedeniyle özellikle tıbbi görüntü sınıflandırma sorunlarında etkili bir model olarak değerlendirilmektedir. Geçiş bağlantı yapısı, gradyan sönümlenmesi sorununu önemli ölçüde azaltmaktadır. Bu, derin katmanların daha istikrarlı bir şekilde eğitilmesini sağlamaktadır.

Bu mimari, bottleneck blok yapısı sayesinde parametre verimliliği sağlamaktadır, böylece hesaplama açısından makul bir maliyet ve yüksek temsil gücü elde edilmektedir. ImageNet'in önceden eğitilmiş ağırlıklarla transfer öğrenmeye son derece uygun olması, sınırlı bir veri kümesi olan tıbbi sorunlar için önemli bir avantaj olmaktadır. ResNet-50, bu özellikler nedeniyle Alzheimer MRI görüntülerindeki yapısal farklılıkları etkili bir şekilde öğrenebilmektedir.

Bununla birlikte, ResNet-50 yapısal sınırlılıklara sahiptir. Modelin yüksek parametre sayısı, EfficientNet ve MobileNet gibi daha hafif mimarilere kıyasla GPU belleği kullanımını ve eğitim süresini uzatmaktadır. Derin katman yapısı, mobil veya edge cihazlarda gerçek zamanlı kullanım için ek hesaplama maliyetleri ve enerji tüketimini artırabilmektedir. Ayrıca, hiperparametrelerin doğru ayarlanmadığı durumlarda modelin öğrenme dengesi bozulabilir ve modelin aşırı karmaşık yapısı genelleme yeteneğini olumsuz etkileyebilir.

ResNet-50, kapsamlı özellik öğrenimi gerektiren klinik sınıflandırma görevlerinde güvenilir bir seçenek olmaktadır. Ancak ResNet-18 veya MobileNet gibi daha küçük ve donanım açısından daha verimli çözümler gerektiren uygulamalar için daha hafif mimarilerin kullanılması gelecekteki çalışmalar için bir seçenek olabilir.

3.3.3 U-Net: Simetrik encoder-decoder yapısı ile detay koruyucu derin ağ mimarisi

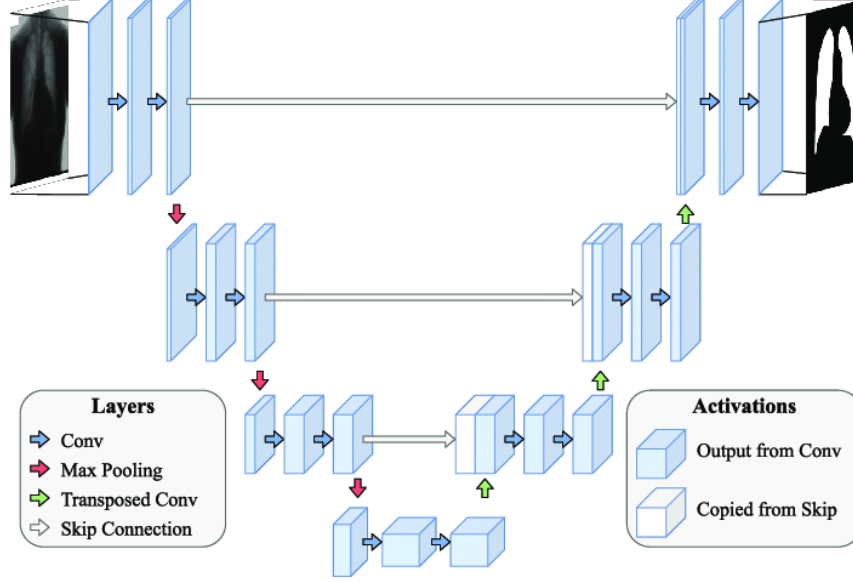
U-Net, Ronneberger et al. (2015) tarafından özellikle biyomedikal görüntü segmentasyonu için geliştirilmiş, encoder–decoder prensibine dayalı bir CNN mimarisidir ve sınırlı veri koşullarında dahi yüksek performans gösterebilmesi nedeniyle medikal görüntüleme yaygın olarak kullanılmaktadır. Mimari, giriş görüntüsünden konumsal ve anlamsal özellikleri çıkaran encoder (daraltıcı yol) ile bu özellikleri yeniden yüksek çözünürlüğe dönüştüren decoder (genişletici yol) olmak üzere simetrik iki bölümden oluşur.

Encoder tarafında konvolüsyon–BatchNorm–ReLU blokları ve max-pooling işlemleri ile çözünürlük azaltılırken, decoder tarafında up-sampling ve konvolüsyon katmanları kullanılarak ayrıntılı çıktı haritaları üretilir. U-Net’i güçlü kılan temel unsur, encoder ile decoder seviyeleri arasında kurulan skip connection yapısıdır. Bu bağlantılar, erken katmanlarda öğrenilen düşük seviyeli uzamsal bilgilerin geri kazanılmasını sağlayarak detay kaybını en aza indirir ve özellikle MRI gibi tıbbi görüntülerde konumsal bütünlüğü korur. Bu yapısal özellikler, Alzheimer evrelerine ilişkin ince yapısal farklılıkların modellenmesinde U-Net’i etkili bir mimari hâline getirmektedir.

Ayrıca, U-Net mimarisinin çok ölçekli yapı özelliği sayesinde farklı çözünürlük seviyelerindeki uzamsal ve anlamsal bilgiler birlikte değerlendirilebilmektedir. Encoder aşamasında öğrenilen yüksek seviyeli temsiller ile decoder aşamasında geri kazanılan uzamsal detayların birleştirilmesi, özellikle Alzheimer hastalığına bağlı ince yapısal değişimlerin doğru biçimde modellenmesine katkı sağlamaktadır. Bu özellikleri sayesinde

U-Net, tıbbi görüntülerde detay koruyucu ve güvenilir bir segmentasyon mimarisi olarak öne çıkmaktadır.

Modelin genel yapısı Şekil 3.4.'te gösterilmiştir.



Şekil 3.4: U-Net mimarisinin genel yapısı ve katmanlar arası veri akışı (Ronneberger et al., 2015'ten uyarlanmıştır)

3.3.3.1 Uygulanan U-Net konfigürasyonu (sınıflandırma uyarlaması)

Bu tez kapsamında U-Net mimarisi, klasik segmentasyon amacı dışında kullanılarak çok sınıflı bir görüntü sınıflandırma problemine uyarlanmıştır. Orijinal U-Net mimarisinde yer alan tam encoder–decoder yapısı yerine, bu çalışmada çok ölçekli encoder yapısı korunmuş; decoder bölümü ve piksel düzeyinde çıktı üreten katmanlar mimariden çıkarılmıştır. Bunun yerine, encoder'ın en derin katmanlarında elde edilen yüksek seviyeli özellik temsilleri üzerinden sınıflandırma odaklı bir üst katman (classification head) tasarlanmıştır. Bu yaklaşım sayesinde, U-Net'in güçlü özellik çıkarım kapasitesinden yararlanılırken, mimari doğrudan sınıflandırma problemine uygun hâle getirilmiştir.

Modelde giriş görüntüsünden elde edilen yüksek seviyeli özellik haritaları, encoder'ın en alt katmanında GlobalAveragePooling2D işlemi ile öznitelik vektörüne dönüştürülmüştür. Bu işlem, uzamsal bilginin tamamen kaybolmasını engelleyerek özetlenmiş fakat ayırt edici bir temsil elde edilmesini sağlamıştır. Elde edilen öznitelik vektörü, tam bağlantılı katmanlar aracılığıyla işlenmiş ve Alzheimer hastalığının dört evresini temsil eden sınıflar için sınıflandırma gerçekleştirilmiştir.

Bu tezde kullanılan sınıflandırma başlığı, rezidüel ve Sıkıştırma–Uyarım (Squeeze-and-Excitation – SE) blokları içeren çok ölçekli bir U-Net türevli encoder yapısının üzerine inşa edilmiştir. Encoder çıkışını takiben GlobalAveragePooling2D katmanı uygulanmış, ardından Batch Normalization ve %40 oranında Dropout katmanı eklenmiştir. Bunu, He uniform başlatıcı ve L2 düzenleme terimi (1×10^{-5}) ile yapılandırılmış 512 nöronlu tam bağlantılı (Dense) katman izlemiştir; sonrasında yeniden Batch Normalization ve ReLU aktivasyon fonksiyonu kullanılmıştır.

Aşırı öğrenmeyi sınırlandırmak amacıyla ikinci bir %40 Dropout katmanı eklenmiş ve model, Alzheimer hastalığının dört evresini temsil eden dört sınıflı softmax çıkış katmanı ile sonlandırılmıştır.

Modelin giriş boyutu, normalize edilmiş iki boyutlu MRI kesitleri için $256 \times 256 \times 3$ olarak belirlenmiştir. Çıkış katmanı, Non-Demented, Very Mild Dementia, Mild Dementia ve Moderate Dementia sınıflarını kapsayan dört sınıflı softmax yapısında tasarlanmıştır.

Öğrenme yaklaşımı olarak, önceden eğitilmiş ağırlıklar kullanılmaksızın, U-Net türevli encoder ve sınıflandırma başlığı baştan eğitilmiştir. Düzenleştirme amacıyla rezidüel bloklar, SE blokları, L2 regularizasyonu ve %40 Dropout kombinasyonu birlikte kullanılarak aşırı öğrenme riski azaltılmış ve modelin genelleme kabiliyeti artırılmıştır.

3.3.3.2 U-Net eğitim stratejisi ve optimizasyon parametreleri

U-Net mimarisinin eğitimi, modelin konvolüsyonel derinliğinden kaynaklanan gradyan dengesizliklerini önlemek ve dengeli bir yakınsama sağlamak amacıyla iki aşamalı öğrenme oranı planlaması stratejisi ile gerçekleştirilmiştir.

Modelin temel öğrenme süreci ilk aşamada yüksek bir öğrenme oranı ($1r = 5 \times 10^{-4}$) ile başlatılmıştır. Ardından ince ayar aşamasında öğrenme oranı 2×10^{-5} seviyesine düşürülmüştür. Bu yöntem, erken dönemde hızlı öğrenme sağlarken aşırı uyum riskinin azaltılmasına katkı sağlamıştır.

Ek olarak, sınıf ağırlıklandırma yöntemi sınıf dengesizliğinin etkisini azaltmak ve nadir sınıfların öğrenme sürecine katkısını artırmak amacıyla kullanılmıştır. Eğitim sırasında

ReduceLROnPlateau ve EarlyStopping geri çağırma mekanizmaları uygulanarak gereksiz epoch uzamaları önlenmiştir.

Bu yaklaşım, modelin genel doğruluğunu artırmak ve sınıflar arası dengesizlikten kaynaklanan performans kaybını minimize etmek amacıyla tercih edilmiştir.

Ayrıca modelin eğitim sürecinde doğrulama kaybı düzenli olarak izlenmiş ve en iyi ağırlıklar doğrulama performansına göre kaydedilmiştir. Böylece modelin yalnızca eğitim verisine değil, genelleme yeteneğine de odaklanması sağlanmıştır. İki aşamalı öğrenme oranı planlaması ve düzenleyici mekanizmalar birlikte değerlendirilerek U-Net mimarisinin istikrarlı bir optimizasyon süreciyle minimum kayıp noktasına yakınsaması hedeflenmiştir.

Tablo 3.3: U-Net modeli için eğitim hiperparametreleri

| Hiperparametre | Değer |
|------------------|---|
| Optimizasyon | Adam ($lr = 5 \times 10^{-4} \rightarrow 2 \times 10^{-5}$, clipnorm = 1.0) |
| Öğrenme Oranı | Aşama-1: $lr = 5 \times 10^{-4} \rightarrow$ Aşama-2: $lr = 2 \times 10^{-5}$ |
| Fine-Tuning | BatchNorm hariç tüm katmanlar açılmıştır |
| Batch Boyutu | 16 |
| Toplam Epoch | 50 (10 + 40 aşamalı fine-tuning) |
| Kayıp Fonksiyonu | Focal Loss ($\gamma = 2.0$), label smoothing = 0.05, class weights uygulanmıştır |
| Raporlama | Doğrulama ve test kümesi için sınıflandırma raporu, karışıklık matrisi, macro-AUC |
| Erken Durdurma | EarlyStopping (patience = 12, restore_best_weights = True) |
| Veri Artırma | Rotation ($\pm 12^\circ$), Zoom (0.85–1.15), Width/Height Shift ($\pm 6\%$), Brightness ($\pm 10\%$), Horizontal Flip |

3.3.3.3 U-Net mimarisi: Teknik avantajlar ve sınırlılıklar

U-Net mimarisi, özellikle medikal görüntü segmentasyonu için geliştirilmiş simetrik encoder–decoder yapısı sayesinde uzamsal bilgiyi koruma konusunda oldukça başarılı bir derin öğrenme modelidir. Encoder kısmında giderek soyutlaşan özellikler çıkarılırken, decoder kısmında bu bilgiler skip connection mekanizması ile yeniden birleştirilir. Bu yapı, detay kaybını minimize ederek hem lokal hem de global örüntülerin etkin biçimde öğrenilmesini sağlamaktadır. Skip connection bağlantıları aynı zamanda gradyan akışını güçlendirerek eğitim sürecinin daha kararlı ilerlemesine yardımcı olmaktadır. U-Net’in en önemli avantajlarından biri, sınırlı veriyle dahi yüksek doğruluk elde edebilmesidir; bu durum, etiketlenmesi zor ve sınırlı örneğe sahip medikal veri setleri için önemli bir fırsat sunmaktadır. Ayrıca bu tez çalışmasında kullanılan SE blokları, kanallar arası dikkat

mekanizması oluşturarak ağın önemli özniteliklere daha yüksek ağırlık vermesine olanak tanımıştır.

Bununla birlikte, U-Net'in segmentasyon tabanlı bir mimari olarak tasarlanmış olması bazı teknik sınırlılıkları da beraberinde getirmektedir. Derin ve genişleyen yapısı nedeniyle EfficientNet veya MobileNet gibi hafif modellere kıyasla daha fazla parametre içermektedir ve bu durum hem GPU bellek kullanımını hem de eğitim süresini artırmaktadır. Aynı zamanda, katman sayısının fazla olması mobil veya edge cihazlarda gerçek zamanlı kullanım için uygun olmayan bir bellek maliyeti oluşturmaktadır. Özellikle yüksek çözünürlüklü MRI görüntülerinde U-Net mimarisinin çalışma süresi diğer hafif modellerden daha uzun olabilmektedir.

Sonuç olarak U-Net, uzamsal duyarlılık ve detay koruma açısından güçlü bir mimari sunarken, yüksek hesaplama maliyeti ve bellek gereksinimi nedeniyle yalnızca uygun donanım koşullarında optimum performans sergileyen bir yapıdır. Bu yönleriyle, sınırlı veriyle yüksek doğruluk gerektiren problemler için önemli bir avantaj sağlamaktadır; ancak hafif mimari gerektiren sistemlerde alternatif modellerin değerlendirilmesi gerekebilir.

3.3.4 MobileNetV2: Hafif ve verimli derin ağ mimarisi

Sandler et al. (2018), mobil cihazlar ve hesaplama gücü kısıtlı sistemler için hafif ve etkili bir CNN mimarisi olan MobileNetV2'yi geliştirmiştir. Model, derinlemesine ayırt edilen konvolüsyon yapısına dayanmaktadır. Bu yapı, standart konvolüsyonların hesaplama maliyetlerini önemli ölçüde azaltmaktadır (Sandler et al., 2018).

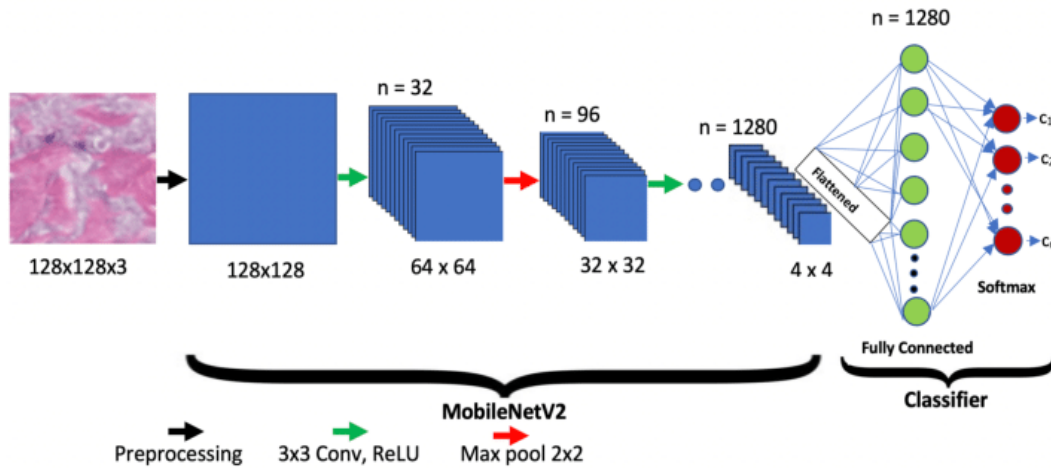
Bu yöntemde uzamsal filtreleme ve kanal birleştirme işlemleri ayrıştırılır, böylece parametre miktarının ve kayan nokta işlem sayısı (Floating Point Operations - FLOP) değerinin önemli ölçüde azaltılması mümkün olmaktadır. MobileNetV2'nin karakteristik özelliğini oluşturan ters residual bloklar, önce düşük boyutlu girdiyi genişletip (genişletme), ardından daraltarak (projeksiyon) bilgi yoğunluğunu artırmaktadır. Giriş ve çıkışın boyutlarının eşit olduğu durumlarda, skip bağlantı yapısı gradyan akışını desteklemektedir.

Bununla birlikte, blokların çıkışında kullanılan linear bottleneck tasarımı, küçük boyutlu temsillerde aktivasyondan kaynaklanan veri kaybını önlemektedir. MobileNetV2, derinlemesine ayırt edilebilir konvülsiyon, ters çevrilmiş gerilim ve lineer tıkanıklık, bu üçlü yapı nedeniyle yüksek doğruluk ve hesaplama verimliliği dengesini sağlayan mobil tıbbi

cihazlar, gömülü klinik sistemler ve gerçek zamanlı uygulamalar için çok uygun bir mimariye sahip olmaktadır.

MobileNetV2 mimarisinin sunduğu bu yapısal avantajlar, özellikle sınırlı veri ve donanım koşullarında gerçekleştirilen tıbbi görüntü sınıflandırma problemlerinde önemli bir üstünlük sağlamaktadır. Parametre sayısının düşük olması, modelin daha hızlı eğitilmesine ve daha kısa sürede yakınsamasına olanak tanırken, aynı zamanda bellek kullanımını da minimize etmektedir. Bu özellikler, MobileNetV2'nin yalnızca laboratuvar ortamlarında değil, klinik karar destek sistemleri, taşınabilir tanı cihazları ve edge tabanlı sağlık uygulamalarında da etkin biçimde kullanılabilmesini mümkün kılmaktadır. Ayrıca, ters residual blokların sağladığı etkin özellik temsili sayesinde model, Alzheimer hastalığının farklı evrelerine özgü morfolojik değişimleri yüksek doğrulukla öğrenebilmekte; bu durum özellikle erken evre ve MCI sınıflarının ayırımında kararlı bir performans sunmasına katkı sağlamaktadır. Bu bağlamda MobileNetV2, performans–hafiflik dengesi açısından Alzheimer teşhisine yönelik pratik ve ölçeklenebilir bir çözüm olarak öne çıkmaktadır.

Modelin genel yapısı şekil 3.5'te gösterilmiştir.



Şekil 3.5: MobileNetV2 mimarisinin genel blok diyagramı (Sandler et al., 2018'den uyarlanmıştır)

3.3.4.1 Kullanılan MobileNetV2 yapılandırması

Bu tez kapsamında, hafif ve hesaplama açısından verimli bir CNN mimarisi olan MobileNetV2, Alzheimer hastalığının evrelerinin sınıflandırılması amacıyla transfer öğrenme yaklaşımı kullanılarak uygulanmıştır. Model, include_top = False parametresi ile çağrılmış ve ImageNet veri kümesi üzerinde önceden eğitilmiş ağırlıklar ile başlatılmıştır.

Bu sayede, genel görüntü temsillerini öğrenmiş olan taban ağ korunmuş; Alzheimer evrelerine özgü sınıflandırma işlemi yalnızca özelleştirilmiş üst sınıflandırıcı katmanlar üzerinden gerçekleştirilerek hesaplama verimliliği sağlanmıştır. Bu yaklaşım, özellikle sınırlı veri kümelerinde aşırı öğrenme riskini azaltmakta ve eğitim süresinin kısaltılmasına katkı sunmaktadır.

Modelin eğitim süreci iki aşamalı bir yapılandırma ile tasarlanmıştır. İlk aşamada, taban ağ tamamen dondurularak (`base.trainable = False`) yalnızca özel sınıflandırıcı katmanlar eğitilmiştir. Bu aşamada optimizasyon için Adam algoritması 1×10^{-3} öğrenme oranı ile kullanılmış ve sınıf belirsizliğini azaltmak amacıyla label smoothing değeri 0.1 olarak uygulanmıştır. İkinci aşamada ise ince ayar (fine-tuning) sürecine geçilmiş; taban ağın Batch Normalization katmanları hariç tutulacak şekilde son 120 katmanı yeniden eğitilebilir hâle getirilmiştir. Bu evrede, daha düşük bir öğrenme oranı (Adam, $lr = 3 \times 10^{-5}$) tercih edilerek modelin hedef veri kümesine özgü daha ayırt edici özellikleri öğrenmesi sağlanmıştır.

MobileNetV2 tabanı üzerine inşa edilen sınıflandırıcı yapı, Global Average Pooling katmanı ile başlamakta; ardından Batch Normalization ve %40 oranında Dropout katmanı uygulanmaktadır. Bunu, L2 düzenleme terimi (1×10^{-5}) ile yapılandırılmış 384 nöronlu tam bağlantılı (Dense) katman izlemekte; sonrasında Batch Normalization ve ReLU aktivasyonu kullanılmaktadır. Aşırı öğrenmeyi sınırlamak amacıyla ikinci bir %40 Dropout katmanı eklenmiş ve model, Alzheimer hastalığının dört evresini temsil eden dört sınıflı softmax çıkış katmanı ile sonlandırılmıştır.

Modelin genel yapısında, giriş verileri `mobilenet_v2.preprocess_input()` fonksiyonu kullanılarak $[-1, 1]$ aralığına ölçeklendirilmiştir. Ayrıca, `tf.keras.Sequential` yapısı içerisinde `RandomFlip`, `RandomRotation`, `RandomZoom` ve `RandomTranslation` katmanlarını içeren bir veri artırma bloğu modele entegre edilmiştir. Bu sayede, sınırlı veri senaryosunda MobileNetV2 mimarisinin genelleme performansı artırılmış, daha kararlı ve dengeli bir öğrenme süreci elde edilmiştir.

3.3.4.2 MobileNetV2 eğitim stratejisi ve optimizasyon parametreleri

Bu tezde kullanılan MobileNetV2 mimarisi, iki aşamalı bir transfer öğrenme stratejisiyle eğitilmiştir. İlk aşamada taban ağ (feature extractor) dondurulmuş, yalnızca üst sınıflandırıcı katmanlar eğitilmiştir. İkinci aşamada ise son 120 katman yeniden eğitime açılarak fine-

tuning uygulanmıştır. Bu yapılandırma, modelin hem genel temsilleri korumasını hem de Alzheimer MRI görüntülerine özgü özellikleri öğrenmesini sağlamıştır. Eğitim süreci boyunca aşırı öğrenme (overfitting) riskini önlemek amacıyla BN, Dropout(0.4) ve label smoothing (0.1) gibi düzenleme teknikleri uygulanmıştır. Ayrıca ReduceLROnPlateau, EarlyStopping ve ModelCheckpoint callback mekanizmaları kullanılarak eğitim süreci dinamik biçimde optimize edilmiştir. MobileNetV2 için tercih edilen bu iki aşamalı eğitim yaklaşımı, sınırlı örnek sayısına sahip tıbbi görüntü veri kümelerinde genelleme kabiliyetini artıran etkili bir strateji sunmaktadır.

Özellikle düşük öğrenme oranlarıyla gerçekleştirilen ince ayar süreci, önceden öğrenilmiş genel görsel özniteliklerin bozulmasını engellemiş ve Alzheimer hastalığına özgü ince morfolojik değişimlerin daha hassas biçimde öğrenilmesine olanak tanımıştır. Düzenleme teknikleri ve adaptif öğrenme hızı mekanizmalarının birlikte kullanılması kararlı bir yakınsama sağlarken, modelin farklı veri dağılımlarında da tutarlı performans sergilemesine katkı sunmuştur.

Bu yapılandırma sayesinde MobileNetV2, hem yüksek sınıflandırma başarımı hem de düşük hesaplama maliyeti gerektiren klinik ve taşınabilir tanı sistemleri için dengeli ve uygulanabilir bir çözüm olarak öne çıkmaktadır.

Tablo 3.4: MobileNetV2 modeli için eğitim hiperparametreleri

| Hiperparametre | Değer |
|------------------|---|
| Optimizasyon | Adam (Aşama-1: $lr = 1 \times 10^{-3} \rightarrow$ Aşama-2: $lr = 3 \times 10^{-5}$, clipnorm = 1.0) |
| Öğrenme Oranı | $1e-3 \rightarrow 3e-5$ (+ ReduceLROnPlateau: factor = 0.5, patience = 5, min_lr = $1e-6$) |
| Batch Boyutu | 16 |
| Toplam Epoch | 50 (10 + 40 aşamalı fine-tuning) |
| Kayıp Fonksiyonu | Categorical Crossentropy + Label Smoothing ($\epsilon = 0.1$) + Class Weights |
| Raporlama | Doğrulama ve test kümesi için sınıflandırma raporu, karışıklık matrisi, macro-AUC |
| Erken Durdurma | EarlyStopping (patience = 12, restore_best_weights = True) |
| Veri Artırma | RandomFlip, RandomRotation (0.08), RandomZoom (0.12), RandomTranslation (0.08), RandomContrast (0.15) |

Eğitim süreci GPU bellek yönetimi açısından optimize edilmiştir. TensorFlow deterministik işlemleri etkinleştirilmiş (TF_DETERMINISTIC_OPS=1) ve bellek büyüme izni

(set_memory_growth=True) tanımlanmıştır. Bu sayede, sonuçların tekrarlanabilirliği sağlanmıştır.

3.3.4.3 MobileNetV2 mimarisi: Teknik avantajlar ve sınırlılıklar

MobileNetV2 mimarisi, hafif yapısı ve yüksek hesaplama verimliliği ile özellikle mobil ve gömülü sistemlerde derin öğrenmenin uygulanabilirliğini artırmaktadır. Mimarinin 3.4 milyon parametre içeren küçük yapısı, eğitim ve çıkarım sürecinde yüksek hız ve düşük bellek tüketimi sunarak GPU ve RAM kullanımını önemli ölçüde azaltmaktadır. Bu verimliliğin nedeni, derinlemesine ayırt edilen konvolüsyon yapısıdır. Bu yapı, klasik konvolüsyonlara göre yaklaşık sekiz ila dokuz kat daha az hesaplama maliyeti gerektirmektedir.

Ek olarak, MobileNetV2, ImageNet üzerinde önceden eğitilmiş ağırlıklarla çağrıldığında, sınırlı veri koşullarında bile iyi bir genelleme sağlamaktadır. Bu çalışmada kullanılan ince ayar yönteminde yalnızca son katmanların eğitime açılması, modelin yakınsama süresini kısaltmıştır.

Bu, Alzheimer magnetik rezonans görüntülere hızlı bir şekilde uyarlanmanın bir sonucudur. Overfitting riskini azaltmak için dropout, BN ve label smoothing gibi düzenleme tekniklerinin kombinasyonu kullanılmıştır. Ek olarak, modelin genelleme yeteneği, RandomFlip, RandomRotation, RandomZoom, RandomTranslation ve RandomContrast gibi veri artırma teknikleri tarafından geliştirilmiştir.

MobileNetV2'nin basit tasarımı, bazı kısıtlamalara neden olmaktadır. Parametrelerin az olması, modelin ResNet-50 veya EfficientNet-B0 gibi daha derin mimarilerde görülen yüksek temsil gücüne ulaşmasını engelleyebilmekte ve karmaşık yapısal örüntülerin yakalanmasını kısmen sınırlayabilmektedir.

Ek olarak, linear bottleneck yapısında aktivasyon fonksiyonlarının kullanılmaması, bazı katmanlarda veri aktarımını kısıtlayarak görüntü kalitesini düşürebilmektedir.

MobileNetV2, küçük veya dengesiz veri setlerinde overfitting'e daha yatkın olduğundan, veri artırma ve düzenleyici yöntemleri seçerken dikkatli olunmalıdır. Modelin daha sığ (daha

derin) yapısı nedeniyle, özellikle medikal görüntüleme gibi yüksek varyanslı veri türlerinde doğruluk artışı, çok katmanlı soyut özelliklerin öğrenilmesini sınırlayabilmektedir.

MobileNetV2, hızı, hafifliğini ve hesaplama verimliliğini birleştirerek Alzheimer MRI görüntülerinin sınıflandırılmasında pratik ve kaynak dostu bir çözüm sağlamaktadır. Bununla birlikte, daha fazla temsil yeteneği gerektiren durumlarda ResNet veya EfficientNet gibi daha derin ağların kullanılması önerilebilmektedir.

3.3.5 Inception: Çok ölçekli özellik çıkarımı için derin CNN mimarisi

Google tarafından, derin ağların hesaplama verimliliğini artırmak ve çok ölçekli uzamsal örüntüleri daha iyi yakalamak amacıyla Inception mimarisi olarak bilinen bir CNN yapısı geliştirilmiştir.

Inception bloklarında 1×1 , 3×3 ve daha büyük çekirdeklerin faktörize edilmiş sürümleri aynı katmanda paralel olarak uygulanırken, klasik konvolüsyon katmanlarında tek bir çekirdek boyutu kullanılmaktadır. Modelin giriş görüntüsünden hem yerel hem de uluslararası düzeyde zengin özellikler çıkarmasına izin veren bu çok yönlü yapıya sahip olmasıdır.

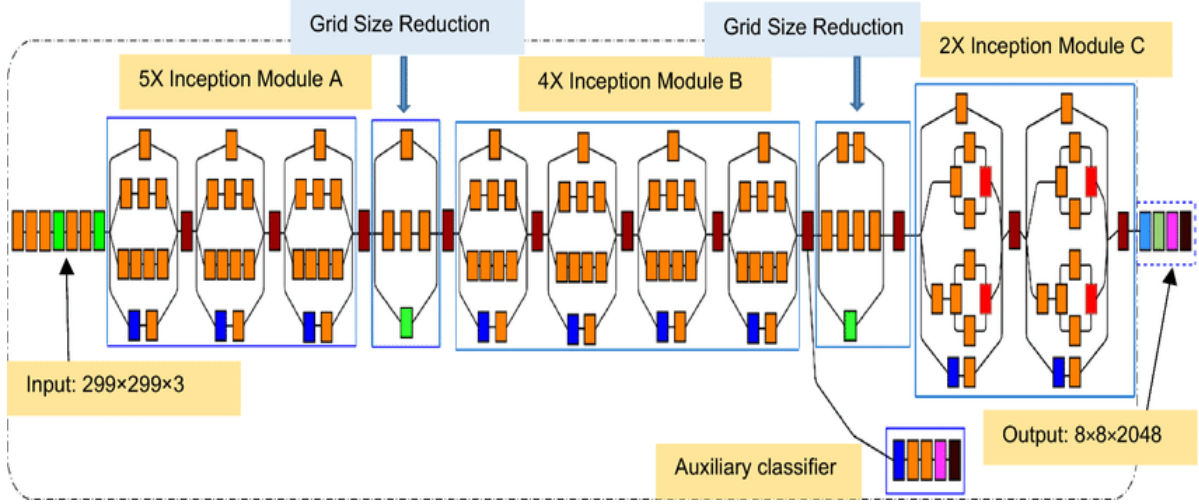
Büyük konvolüsyonlar, büyük çekirdeklerin maliyetini azaltmak için daha küçük parçalara bölünmüştür; 1×1 konvolüsyonlar, kanalları azaltmak ve kanallar arası bilgi etkileşimini güçlendirmek için kullanılmıştır. Önceki sürümlerle karşılaştırıldığında, InceptionV3 optimizasyonu derinliği, temsiliyeti ve FLOP maliyetlerini azaltmaktadır.

InceptionV3 mimarisinin temel çalışma prensibi, kanal boyutunda farklı ölçeklerde konvolüsyon yollarının aynı blok içinde paralel olarak işlenmesidir. Bu nedenle, model, dahil ettiği çeşitli filtre büyüklüklerini aynı şekilde yorumlayarak çok yönlü bir özellik haritası yaratmaktadır.

InceptionV3, hesaplama maliyetlerini düşürmek için büyük çekirdekleri daha küçük konvolüsyonlara bölerek blokların verimliliğini artırmaktadır. Bu, 1×1 konvolüsyonlar kullanılarak boyut sıkıştırması yoluyla gerçekleştirilmektedir. Yardımcı sınıflandırıcılar (Auxiliary Classifiers – AC) derin ağlarda gradyan akışı desteklese de, bu çalışmada bu bileşenler devre dışı bırakılmıştır.

Özellikle yüksek çözünürlüklü ve karmaşık medikal görüntüler için çok katmanlı bilgi yakalama yeteneğine sahip güçlü bir mimariye sahip olan InceptionV3, bu yapısal özellikleri sayesinde hesaplama maliyetlerinin makul seviyelerde tutulmasını sağlamaktadır.

Modelin genel yapısı şekil 3.6.'te gösterilmiştir.



Şekil 3.6: InceptionV3 mimarisinin genel yapısı (Szegedy et al., 2016'dan uyarlanmıştır)

3.3.5.1 Kullanılan InceptionV3 yapılandırması

Bu tez çalışmasında InceptionV3 mimarisi, Alzheimer evrelerini sınıflandırmak için transfer öğrenme (transfer learning) yöntemine uyarlanmıştır. Model, ImageNet üzerinde önceden eğitilmiş ağırlıkları kullanarak çağrıldı ve include_top=False parametresi kullanıldı. Böylece, InceptionV3'ün derin konvolüsyonel katmanlarının sağladığı zengin öznitelik temsilleri korunmuş ve özelleştirilmiş bir çıkış bloğu kullanılarak Alzheimer hastalığının evrelerine özgü ayırım yapılmıştır.

İlk aşamada modelin derin katmanları (stem bloğu, Inception-A/B/C modülleri ve grid-reduction blokları) dondurularak yalnızca üst sınıflandırıcı bloğu eğitilmiştir. Daha sonra ikinci aşamada, BN katmanları hariç olmak üzere tabandaki son 140 katman kademeli biçimde eğitilebilir hâle getirilmiş ve fine-tuning uygulanmıştır. Bu strateji, önceden öğrenilmiş genel görsel temsillerin korunmasını sağlarken, yüksek seviyeli özelliklerin Alzheimer MRI verisine uyumlu olacak şekilde yeniden optimize edilmesine olanak tanımıştır.

Özelleştirilmiş sınıflandırıcı bloğu, InceptionV3 mimarisinin önceden eğitilmiş (pre-trained) ve `include_top = False` parametresiyle yapılandırılmış hâli üzerine inşa edilmiştir. Modelin konvolüsyonel tabanını takiben Global Average Pooling katmanı kullanılmış, ardından Batch Normalization ve %40 oranında Dropout katmanı eklenmiştir. Bu yapıyı, He uniform başlatıcı ve L2 düzenleme terimi (1×10^{-5}) ile yapılandırılmış 512 nöronlu tam bağlantılı (Dense) katman izlemiştir; sonrasında yeniden Batch Normalization ve ReLU aktivasyon fonksiyonu uygulanmıştır. Aşırı öğrenmeyi azaltmak amacıyla ikinci bir %40 Dropout katmanı eklenmiş ve model, Alzheimer hastalığının dört evresini temsil eden sınıflar için softmax aktivasyonuna sahip dört çıkışlı Dense katman ile sonlandırılmıştır.

Modelin giriş boyutu, renkli (RGB) MRI görüntüleri için $299 \times 299 \times 3$ olarak belirlenmiştir. Çıkış katmanı, Non-Demented, Very Mild Dementia, Mild Dementia ve Moderate Dementia sınıflarını kapsayan dört sınıflı softmax yapısında tasarlanmıştır. Ön işleme aşamasında, görüntüler önce ImageDataGenerator aracılığıyla `rescale = 1./255` parametresi kullanılarak [0,1] aralığına ölçeklendirilmiş, ardından `inception_v3.preprocess_input` fonksiyonu ile InceptionV3 mimarisine özgü normalizasyon işlemine tabi tutulmuştur. Aktivasyon fonksiyonları olarak gizli katmanlarda ReLU, çıkış katmanında ise softmax tercih edilmiştir. Düzenleme (regularization) amacıyla Batch Normalization, %40 oranında Dropout ve tam bağlantılı katmanda L2 ceza terimi birlikte kullanılmıştır.

Eğitim süreci aşamalı bir stratejiyle yürütülmüştür. İlk aşamada, yalnızca üst sınıflandırıcı bloğu eğitilmiş ve bu süreçte Adam optimizasyon algoritması, 8×10^{-4} öğrenme oranı ve 1.0 clipnorm değeri kullanılmıştır. İkinci aşamada ise Batch Normalization katmanları hariç tutulacak şekilde modelin son 140 katmanı yeniden eğitilebilir hâle getirilmiş ve ince ayar (fine-tuning) süreci Adam optimizasyon algoritması, 1×10^{-5} öğrenme oranı ve 1.0 clipnorm değeri ile gerçekleştirilmiştir.

Kayıp fonksiyonu olarak, sınıf dengesizliğine duyarlı Focal Loss ($\gamma = 2.0$) ile hafif etiket yumuşatma (`label smoothing = 0.05`) birlikte kullanılmış; sınıf frekanslarından türetilen α vektörü aracılığıyla kayıp fonksiyonu sınıf ağırlıkları ile ölçeklendirilmiştir. Ayrıca, eğitim kümesindeki örnek sayıları temel alınarak dinamik biçimde hesaplanan `class_weight` sözlüğü kullanılarak sınıf dengesizliği daha etkin biçimde modellenmiştir.

Bu yapılandırma sayesinde, transfer öğrenmenin sağladığı ön bilgi korunurken focal loss ve sınıf ağırlıkları aracılığıyla sınıf dengesizliği problemi önemli ölçüde azaltılmış; InceptionV3 modeli, hem yüksek doğruluk hem de dengesiz veri koşullarında kararlı ve güvenilir bir performans sergileyecek şekilde yapılandırılmıştır.

3.3.5.2 InceptionV3 eğitim stratejisi ve optimizasyon parametreleri

Tablo 3.5: InceptionV3 modeli için eğitim hiperparametreleri

| Hiperparametre | Değer |
|------------------|--|
| Optimizasyon | Adam (aşamalı eğitim: $lr_1 = 8 \times 10^{-4} \rightarrow lr_2 = 1 \times 10^{-5}$, clipnorm = 1.0) |
| Öğrenme Oranı | Aşama-1: $lr = 8 \times 10^{-4} \rightarrow$ Aşama-2: $lr = 1 \times 10^{-5}$ |
| Batch Boyutu | 16 |
| Toplam Epoch | 50 (10 + 40 aşamalı fine-tuning) |
| Kayıp Fonksiyonu | Focal Loss ($\gamma = 2.0$), label smoothing = 0.05, class weights uygulanmıştır |
| Raporlama | Doğrulama ve test kümesi için sınıflandırma raporu, karışıklık matrisi, macro-AUC |
| Erken Durdurma | EarlyStopping (patience = 12, restore_best_weights = True) |
| Veri Artırma | ImageDataGenerator tabanlı; rotation ($\pm 12^\circ$), zoom (± 0.12), width/height shift (± 0.06), shear (0.05), brightness (0.9–1.1), horizontal flip |

Model iki aşamalı bir eğitim süreciyle optimize edilmiştir. İlk aşamada yalnızca sınıflandırıcı bloğu eğitilmiş; ikinci aşamada ise modelin üst katmanları (fine-tuning) yeniden eğitilmiştir. Focal Loss fonksiyonu, dengesiz sınıf dağılımında azınlık sınıfların ağırlığını artırarak daha dengeli bir öğrenme sağlamıştır.

3.3.5.3 InceptionV3 teknik avantajlar ve sınırlılıklar

InceptionV3 mimarisi, çok ölçekli özellik çıkarımı yapabilen güçlü bir derin öğrenme modelidir. Farklı boyutlardaki konvolüsyon filtreleri sayesinde hem düşük seviyeli yerel özellikleri hem de yüksek seviyeli küresel yapısal örüntüleri öğrenebilme kapasitesine sahiptir. Bu özellik, Alzheimer'a ait MRI görüntülerindeki ince morfolojik değişimlerin etkili biçimde modellenmesine olanak tanımaktadır.

1×1 , 3×3 ve faktörize edilmiş 5×5 konvolüsyon yolları, modelin çok yönlü temsil kabiliyetini geliştirirken, grid boyutu azaltma mekanizması bilgi kaybını azaltmaktadır ve derin katmanlarda daha stabil özellik çıkarımı sağlamaktadır. ImageNet üzerinde önceden eğitilmiş ağırlıkların transfer öğrenme yaklaşımıyla kullanılması, sınırlı veri senaryolarında

dahi yüksek genelleme performansı elde edilmesini sağlamıştır. Bu tez kapsamında yalnızca son 140 katman ince ayar (fine-tuning) amacıyla yeniden eğitime açılmıştır. Bu strateji, modelin önceden öğrenilmiş genel görsel temsilleri korunurken Alzheimer MRI verilerine özgü özelliklere etkili biçimde uyum sağlamasına olanak tanımıştır.

Ek olarak, mimarideki AC yapıları, derin katmanlara gradyan akışını artırarak vanishing gradient sorununu azaltmıştır. Öğrenirken istikrar kazanmıştır. Ek olarak, sınıf dengesizliğine karşı odaklanma kaybı ve sınıf ağırlığının kullanımı, azınlık sınıfların temsil gücünü artırmış ve Alzheimer evreleri arasında daha dengeli bir sınıflandırma yapılmıştır.

Bununla birlikte, InceptionV3 belirli teknik kısıtlamalar içermektedir. Yaklaşık 23 milyon parametre içeren karmaşık yapısı nedeniyle EfficientNet ve MobileNet gibi daha hafif modellere kıyasla daha fazla işlem gücü ve GPU belleği gerektirmektedir. 299×299 gibi yüksek çözünürlüklü bir giriş boyutuyla çalışması, öğrenme süresini ve bellek tüketimini artırmaktadır. Gerçek zamanlı mobil veya gömülü sistemlerde çalışmayı zorlaştıran karmaşık ve çok kollu mimari yapısı, çıkarım maliyetlerini yükseltmektedir ve uygulama için daha fazla optimizasyon gerektirmektedir.

Bununla birlikte, InceptionV3'ün güçlü temsil yeteneği ve çok ölçekli özellik çıkarımı, tez kapsamında yüksek doğruluk-verimlilik dengesinin değerlendirilmesinde üstün bir performans sunmuştur.

3.4 Performans Değerlendirme Kriterleri

Derin öğrenme modellerinin sınıflandırma performansını değerlendirmek, yalnızca ACC metriği üzerinden yapıldığında eksik ve hatta yanıltıcı sonuçlara yol açabilmektedir.

Özellikle tıbbi görüntü sınıflandırması gibi klinik açıdan yüksek hassasiyet gerektiren alanlarda, modelin her bir sınıfa ilişkin ayırt edicilik gücü, yanlış sınıflandırma oranları, genelleme kabiliyeti ve istatistiksel güvenilirliği bir bütün olarak incelenmelidir.

Bu nedenle, Alzheimer hastalığının evrelerinin belirlenmesine yönelik geliştirilen derin öğrenme modellerinin performansı bu tez kapsamında dört temel boyut altında değerlendirilmiştir:

1. Sınıflandırma performansı (İstatistiksel metrikler: ACC, Precision, Recall, F1-Score, ROC-AUC)

2. Model boyutu ve hafiflik analizi
3. Eğitim süresi ve donanımsal zaman maliyeti
4. Parametre yoğunluğu ve hesaplama karmaşıklığı

Bu başlıklar altında yapılan değerlendirmeler, yalnızca modelin doğruluk oranını değil; aynı zamanda klinik uygulanabilirliğini, donanımsal verimliliğini ve sürdürülebilir kullanım potansiyelini de bütüncül biçimde ortaya koymaktadır.

3.4.1 Sınıflandırma performansı: İstatistiksel metrikler

Bu çalışma kapsamında, Alzheimer evrelerinin sınıflandırılması görevini yerine getiren derin öğrenme modelleri yalnızca genel ACC değeriyle değil; aynı zamanda sınıflar arası ayırım gücü ve yanlış sınıflandırma tipleri dikkate alınarak değerlendirilmiştir.

Bu amaçla aşağıdaki temel istatistiksel metrikler kullanılmıştır:

- Accuracy (Genel doğruluk oranı)
- Precision (Pozitif tahmin kesinliği)
- Recall (Duyarlılık / Sensitivity)
- F1-Score (Harmonik ortalama) ve Macro-F1 (sınıf dengeli ortalama)
- ROC-AUC

Bu metriklerin her biri, modelin klinik güvenilirliği açısından farklı bir yönünü ortaya koymaktadır ve Alzheimer evrelerinin adil, dengeli ve güvenilir biçimde sınıflandırılabilmesi için tamamlayıcı bir değerlendirme çerçevesi sunmaktadır. Bu çoklu metrik yaklaşımı, tek bir performans göstergesine bağlı kalmadan modellerin klinik açıdan daha bütüncül ve güvenilir biçimde değerlendirilmesine olanak sağlamıştır.

3.4.1.1 Accuracy (Genel Doğruluk Oranı)

Accuracy, bir sınıflandırma modelinin yaptığı doğru tahminlerin, tüm örnekler içindeki oranını ifade etmektedir. Yani modelin genel doğruluk düzeyini ölçen en temel başarı metriğidir. Matematiksel olarak aşağıdaki şekilde tanımlanır:

$$\text{Doğruluk} = \frac{\sum_{i=1}^k TP_i}{\sum_{i=1}^k (TP_i + FP_i + FN_i + TN_i)} \quad (3.2)$$

Burada:

- TP_i : i. sınıfa ait doğru pozitif (True Positive) sayısıdır,
- FP_i : i. sınıfa ait yanlış pozitif (False Positive) sayısıdır,
- FN_i : i. sınıfa ait yanlış negatif (False Negative) sayısıdır,
- TN_i : i. sınıfa ait doğru negatif (True Negative) sayısıdır,
- k : Toplam sınıf sayısını ifade etmektedir.

Accuracy, sezgisel olarak kolay anlaşılabilir bir metrik olup modelin genel performansını tek bir oranla özetleyebilmektedir. Ancak sınıf dengesizliği bulunan veri setlerinde, özellikle bazı Alzheimer evrelerinin diğerlerine kıyasla daha az sayıda örnek içermesi durumunda, bu metrik yanıltıcı sonuçlar verebilmektedir. Klinik açıdan değerlendirildiğinde, bir modelin yalnızca çoğunluk sınıfını doğru tahmin etmesi ve diğer evreleri hatalı biçimde sınıflandırması, yüksek bir doğruluk oranı elde edilse dahi kabul edilebilir bir performans anlamına gelmemektedir. Örneğin, yalnızca “ciddi evre Alzheimer” sınıfına ait örnekleri doğru tahmin eden ve diğer evreleri göz ardı eden bir modelin %80 doğruluk sağlaması, klinik karar destek süreçleri açısından yetersiz kabul edilir. Bu nedenle accuracy metriğinin, Precision, Recall ve F1-Score gibi sınıf bazlı performansı daha iyi yansıtan metriklerle birlikte ele alınması gerekmektedir.

3.4.1.2 Precision (pozitif tahmin kesinliği)

Precision, modelin pozitif olarak tahmin ettiği örneklerin ne kadarının gerçekten pozitif olduğunu ölçen bir metriktir. Diğer bir ifadeyle, modelin yaptığı doğru pozitif tahminlerin, toplam pozitif tahminler içindeki oranıdır.

Matematiksel olarak aşağıdaki şekilde tanımlanır:

$$\text{Pozitif Tahmin Kesinliği} = \frac{TP_i}{TP_i + FP_i} \quad (3.3)$$

Burada:

- TP_i : i. sınıfa ait doğru pozitif sayısı,
- FP_i : i. sınıfa ait yanlış pozitif sayısıdır.

Precision değeri, modelin yanlış pozitif üretme eğilimini yansıtan bir performans metriğidir; başka bir ifadeyle model ne kadar az sayıda yanlış pozitif tahmin yapıyorsa, precision değeri o ölçüde yüksek olmaktadır.

Tıbbi sınıflandırma problemlerinde yanlış pozitif sonuçlar klinik açıdan son derece kritik olup, modelin güvenilirliğini doğrudan etkilemektedir.

Örneğin, gerçekte sağlıklı bir bireyin “orta evre Alzheimer” olarak etiketlenmesi; gereksiz ileri tetkiklerin yapılmasına, birey üzerinde psikolojik strese ve ek ekonomik maliyetlere yol açabilmektedir.

Bu nedenle precision metriği, özellikle yanlış teşhis riskinin en aza indirilmesinin hedeflendiği klinik karar destek sistemlerinde, model performansının değerlendirilmesinde önemli bir ölçüt olarak öne çıkmaktadır.

3.4.1.3 Recall (duyarlılık / sensitivity)

Recall, modelin gerçek pozitif örnekleri ne ölçüde doğru şekilde tespit edebildiğini gösteren bir metriktir. Başka bir ifadeyle, gerçek pozitiflerin (hastalık varlığı) model tarafından yakalanma oranını ifade etmektedir.

Matematiksel olarak aşağıdaki şekilde tanımlanmaktadır:

$$\text{Duyarlılık}_i = \frac{TP_i}{TP_i + FN_i} \quad (3.4)$$

Burada:

- TP_i : i . sınıfa ait doğru pozitif sayısı,
- FN_i : i . sınıfa ait yanlış negatif sayısıdır.

Recall değeri, özellikle eksik tanı (false negative) durumlarını azaltmayı hedefleyen bir performans metriğidir. Yüksek recall değeri, modelin pozitif sınıfları gözden kaçırmadığını gösterirken; düşük recall değeri, hastalıklı bireylerin “sağlıklı” olarak sınıflandırılma riskine işaret etmektedir. Klinik açıdan değerlendirildiğinde, recall değeri düşük olan bir model Alzheimer’ın erken evresindeki hastaları tespit edemeyebilmektedir. Bu durum, tedaviye geç başlanmasına, hastalığın daha hızlı ilerlemesine ve bireyin yaşam kalitesinin düşmesine yol açabilmektedir. Bu nedenle recall metriği, özellikle erken teşhisin kritik olduğu ve

duyarlılık odaklı tıbbi görüntüleme uygulamalarında öncelikli değerlendirme kriterlerinden biri olarak kabul edilmektedir.

3.4.1.4 F1-Score (harmonik başarı) ve macro-F1

F1-Score, precision (pozitif tahmin kesinliği) ve recall (duyarlılık) metriklerinin harmonik ortalamasıdır. Bu metrik, hem yanlış pozitifleri hem de yanlış negatifleri dikkate alarak modelin genel sınıflandırma başarısını dengeli bir biçimde ölçmektedir. Özellikle sınıf dağılımı dengesiz olduğunda, yalnızca accuracy yerine daha güvenilir bir performans göstergesi olarak tercih edilmektedir.

Bu yönüyle F1-Score, precision ve recall arasında bir denge kurarak, modelin hem hatalı alarm üretme eğilimini hem de gerçek pozitifleri kaçırma riskini aynı anda değerlendirme imkânı sunmaktadır.

Matematiksel olarak şu şekilde ifade edilir:

$$F1_i = 2 \times \frac{\text{Pozitif Tahmin Kesinliği}_i \times \text{Duyarlılık}_i}{\text{Pozitif Tahmin Kesinliği}_i + \text{Duyarlılık}_i} \quad (3.5)$$

$$\text{Macro - F1} = \frac{1}{k} \sum_{i=1}^k F1_i \quad (3.6)$$

Burada:

- Pozitif Tahmin Kesinliği_i : i. sınıfa ait pozitif tahmin kesinliğini,
- Duyarlılık_i : i. sınıfa ait duyarlılığı ifade eder.

F1-Score, yüksek precision ve recall değerlerinin birlikte sağlanmasını gerektirdiğinden, tek başına kullanılan metriklerle kıyasla daha kapsamlı bir performans ölçütü sunmaktadır. Bu metrik, her bir sınıf için ayrı ayrı hesaplandıktan sonra tüm sınıfların ortalaması alınarak Macro-F1 değeri elde edilmektedir. Bu yaklaşım, tüm sınıfları eşit ağırlıkta değerlendirerek dengesiz veri setlerinde küçük örnek sayısına sahip sınıfların da adil biçimde temsil edilmesini sağlamaktadır.

Alzheimer evrelerinin sınıflandırılmasında bazı evrelerin, özellikle erken evre veya hafif bilişsel bozukluk gibi sınıfların, daha az sayıda örneğe sahip olması mümkündür. Bu durumda Macro-F1 metriği, modelin her bir evreyi ne ölçüde adil ve dengeli tanıyabildiğini ortaya koymakta; dolayısıyla modelin genelleme kabiliyetinin en güçlü göstergelerinden biri olarak değerlendirilmektedir.

3.4.2 Metriklerin uygulanması ve izlenmesi

Bu eğitim sürecinde ACC, loss, precision, recall ve F1-score değerleri her epoch sonunda hem eğitim hem de doğrulama kümeleri için kaydedilmiş ve bu metriklerin zaman içindeki değişimleri Matplotlib/Seaborn kullanılarak görselleştirilmiştir. Bu eğriler, yakınsama davranışını, aşırı öğrenme (overfitting) ya da yetersiz öğrenme (underfitting) eğilimlerini tespit etmek için kullanılmıştır.

Ek olarak, sınıf bazlı ROC-AUC ve Precision–Recall eğrileri ile karmaşıklık matrisi (Confusion Matrix – CM) analizleri gerçekleştirilmiş; böylece modellerin hangi Alzheimer evrelerinde başarılı, hangi evrelerde ise karıştırma eğiliminde olduğu ayrıntılı biçimde incelenmiştir. Nihai model seçimi, bu çok boyutlu metriklerin birlikte değerlendirilmesiyle yapılmıştır.

Bu kapsamlı değerlendirme yaklaşımı, model performansının yalnızca sayısal başarı düzeyleriyle değil, aynı zamanda klinik anlamlılık ve sınıflar arası ayırt edilebilirlik açısından da bütüncül biçimde yorumlanmasını sağlamıştır.

3.4.3 Model boyutu ve hafiflik analizi

Model boyutu, özellikle gömülü sistemler, mobil cihazlar ve klinik karar destek uygulamalarında depolama gereksinimi, bellek kullanımı ve yükleme süresi açısından önemlidir.

Bu tezde kullanılan modeller arasında en küçük olanı, yaklaşık 14 Megabayt (Megabyte - MB) boyutuyla MobileNetV2'dir ve kaynakları kısıtlı sistemlerde kolayca dağıtılabılır ve entegre edilebilir. EfficientNet-B0 yaklaşık 20 MB ile hafiflik ve doğruluk açısından dengeli bir profil sergilemiştir.

Bununla birlikte, ResNet-50 ve InceptionV3'ün 90 MB'ın üzerindeki boyutları nedeniyle daha fazla disk ve bellek ihtiyacı vardır. Segmentasyon kökenli U-Net, yaklaşık 120 MB ile en büyük modeldir ve hem depolama hem de çalışma belleği açısından en maliyetli yöntem olarak öne çıkmaktadır. Bu bulgular, model seçiminde yalnızca doğruluğun değil, aynı zamanda model boyutunun ve hesaplama gereksinimlerinin de önemli pratik kriterler olduğunu göstermektedir.

3.4.4 Eğitim süresi ve donanımsal zaman maliyeti

Google Colab Pro üzerinde Tesla T4 GPU kullanılarak tüm modeller, 50 epoch boyunca aynı veri kümesi ve benzer hiperparametre ayarlamaları kullanılarak eğitilmiştir. MobileNetV2, en kısa epoch süresine ve yaklaşık 50 dakikalık toplam eğitim süresine sahip olup zaman açısından en verimli model olarak öne çıkmıştır. EfficientNet-B0, daha kısa bir eğitim süresiyle aynı doğruluk değerlerini koruyarak rekabetçi bir performans göstermiştir. ResNet-50 ve InceptionV3'ün karmaşık mimarileri nedeniyle epoch ve eğitim süreleri önemli ölçüde artmıştır. U-Net en uzun eğitim süresine sahip model olmuştur. Bu sonuçların gösterdiği gibi, eğitim süresi, modelin yeni verilere uyarlanması veya klinik koşullarda sık sık yeniden eğitilmesi gerektiğinde model seçiminde önemli bir rol oynamaktadır.

3.4.5 Parametre yoğunluğu ve hesaplama karmaşıklığı

Modellerin toplam parametreleri ve yaklaşık FLOP değerleri, işlem yükü ve bellek gereksinimini doğrudan etkilemektedir. MobileNetV2 yaklaşık 3–4 milyon parametre ve oldukça düşük bir FLOP değeriyle çalışmanın en hafif ve en ekonomik modeli olmuştur. Parametre sayısı ve FLOP değeri bakımından EfficientNet-B0, MobileNetV2'den daha yüksek olsa da, doğruluk-maliyet dengesi açısından avantaj sağlamıştır ve hafif bir sınıfta yer almıştır. Milyarlarca FLOP ve yirmi milyondan fazla parametre gereksinimiyle ResNet-50 ve InceptionV3, yüksek hesaplama yoğunluğuna sahiptir; güçlü GPU altyapıları için uygun olmakla birlikte, mobil veya gömülü sistemlerde pratik değildir. U-Net, yaklaşık otuz milyon parametresi ve yüksek bir FLOP değeri ile işlem gücü ve bellek açısından en yoğun model olmuştur. Sonuç olarak, model seçerken istatistiksel başarı, parametre yoğunluğu ve hesaplama maliyeti birlikte değerlendirilmelidir.

3.4.6 Toplam performans ve uygulama uyumluluğu

Bu tezde, doğruluk, F1-skoru ve diğer istatistiksel başarı kriterleri ile birlikte uygulama gereksinimleri, yaygınlaştırılabilirlik ve donanımsal kısıtlamalar da değerlendirilmiştir.

Modeller dört temel kritere göre incelenmiştir: model boyutu, eğitim süresi, hesaplama maliyeti ve klinik performansı yansıtan F1-skoru. Bu analize göre, gerçek zamanlı klinik tanı sistemleri, edge AI tabanlı cihazlar ve mobil sağlık uygulamaları için MobileNetV2 en iyi model olmuştur. Daha hafif modeller, düşük parametre yoğunluğu ve daha kısa eğitim süresi sayesinde sınıflar arası ayrımı yeterli düzeyde sağlarken donanımsal verimlilik sunmuştur.

EfficientNet-B0, doğruluk ve hesaplama verimliliği açısından rekabetçi bir performans sergilemiştir. ResNet-50 ve InceptionV3 modelleri, karmaşık örüntüleri öğrenme ve yüksek temsil gücü sağlamada en iyi modeller olmakla birlikte, yüksek parametre sayıları ve yüksek FLOP değerleri nedeniyle çok fazla bellek ve işlem gücü gerektirmişlerdir; bu nedenle, bu modeller daha çok bulut tabanlı sistemler veya güçlü bir GPU altyapısına sahip klinik ortamlar için uygun görünmektedir. U-Net, segmentasyon temelli yapısı sayesinde uzamsal ayrıntıları kolayca yakalamasına rağmen, eğitim süresi ve parametrelerin yoğunluğu nedeniyle doğrudan sınıflandırma görevlerine optimize edilmemiştir.

Bu bulgular, derin öğrenme modellerinde başarının doğruluk skorları aracılığıyla ölçülmemesi gerektiğini ortaya koymaktadır. Donanım gereksinimleri, sürdürülebilirlik ve uygulama uyumluluğu, klinik entegrasyon sürecinde çok önemlidir. Bu nedenle, Alzheimer teşhis sistemlerinin gerçek hayatta etkili bir şekilde kullanılabilmesi için doğruluk, hız ve taşınabilirlik arasındaki ideal dengeyi sağlamak gerekmektedir.

4. DENEYSEL BULGULAR

Bu bölümde, tez kapsamında geliştirilen derin öğrenme modellerinin deneysel sonuçları karşılaştırmalı olarak sunulmaktadır. Her modelin ACC, kesinlik (Precision), duyarlılık (Recall), F1-skoru, eğitim süresi, model boyutu ve parametre yoğunluğu gibi performans kriterlerine göre analizleri gerçekleştirilmiştir. Ayrıca, ROC eğrileri üzerinden sınıflar arası ayırt edicilik yetenekleri incelenmiş ve elde edilen sonuçlar, literatürde yer alan benzer çalışmalarla karşılaştırmalı biçimde değerlendirilmiştir.

Bu kapsamlı analiz sayesinde, her bir mimarinin güçlü ve sınırlı yönleri hem teknik performans hem de klinik uygulanabilirlik açısından ayrıntılı biçimde ortaya konmuştur.

4.1 Eğitim Süreci ve Genel Gözlemler

Bu bölümde, Alzheimer hastalığının dört klinik evresini sınıflandırmak amacıyla kullanılan beş farklı derin öğrenme mimarisinin (MobileNetV2, EfficientNet-B0, ResNet-50, InceptionV3 ve U-Net Encoder) ortak eğitim stratejisi, kullanılan veri seti yapısı, eğitim senaryosu ve optimizasyon yaklaşımı bütüncül biçimde sunulmaktadır. Modellerin tamamı aynı veri kümesi, aynı eğitim–doğrulama–test ayrımı ve aynı ön işleme prosedürleri altında eğitildiği için karşılaştırmalar metodolojik olarak tutarlı ve karşılaştırılabilir niteliktedir.

4.1.1 Veri bölünmesi ve eğitim senaryosu

Bu çalışmada değerlendirilen tüm derin öğrenme modellerinde aynı veri bölünme stratejisi kullanılmıştır. Buna göre, veri kümesi 5.141 görüntüden oluşan eğitim seti, 1.848 görüntüden oluşan doğrulama seti ve 1.279 görüntüden oluşan test seti olacak şekilde ayrılmıştır. Çalışmada yer alan sınıflar NonDemented, VeryMildDemented, MildDemented ve ModerateDemented olarak belirlenmiştir. Bu veri ayrımı, sınıflar arasındaki yakın morfolojik benzerlikler dikkate alındığında özellikle erken evre Alzheimer bulgularının güvenilir ve dengeli biçimde temsil edilebilmesi açısından kritik bir yapı sunmaktadır.

4.1.2 Ortak ön işleme ve veri artırma (augmentation)

Bu çalışmada değerlendirilen tüm derin öğrenme modellerinde ortak bir ön işleme ve veri artırma akışı uygulanmıştır. Buna göre, giriş görüntüleri ResNet-50, EfficientNet-B0, MobileNetV2 ve InceptionV3 modelleri için 224×224 boyutuna yeniden ölçeklendirilmiş, U-Net Encoder mimarisinde ise model yapısına uygun olacak şekilde 256×256 boyutu

tercih edilmiştir. Tüm görüntüler [0,1] aralığında normalize edilerek piksel değerlerinin standartlaştırılması sağlanmıştır. Eğitim sürecinde veri artırma amacıyla görüntülere $\pm 12^\circ$ aralığında rastgele dönüş, 0.88–1.12 aralığında yakınlaştırma ve uzaklaştırma, yatay ve dikey eksenlerde yaklaşık $\pm 6\%$ oranında kaydırma, $\pm 10\%$ seviyesinde parlaklık değişimi ve yatay eksende rastgele çevirme işlemleri uygulanmıştır. Bu veri artırma süreci, özellikle örnek sayısı görece düşük olan VeryMildDemented ve ModerateDemented sınıflarında aşırı öğrenme riskini azaltmak ve modellerin genelleme kabiliyetini artırmak amacıyla tercih edilmiştir.

4.1.3 İki aşamalı eğitim stratejisi (freezing - fine-tuning)

Bu çalışmada değerlendirilen beş derin öğrenme modelinin tamamında iki aşamalı bir eğitim stratejisi benimsenmiştir. İlk aşamada, transfer öğrenme yaklaşımı kapsamında `include_top = False` parametresi kullanılarak ImageNet veri kümesi üzerinde önceden eğitilmiş ağırlıklar modele yüklenmiş ve taban ağ dondurulmuştur. Bu aşamada yalnızca modele eklenen sınıflandırıcı katmanlar eğitilmiş olup, kullanılan öğrenme oranları modellere bağlı olarak 5×10^{-4} ile 1×10^{-3} aralığında belirlenmiştir. Bu strateji, modellerin genel ve evrensel görsel özellikleri korumasını sağlayarak Alzheimer MRI verisi için sağlam bir başlangıç noktası oluşturmuştur.

İkinci aşamada ise ince ayar (fine-tuning) sürecine geçilmiştir. Bu aşamada, taban ağın son blokları Batch Normalization katmanları hariç tutulacak şekilde yeniden eğitime açılmıştır. Öğrenme oranı daha düşük değerlerde tutulmuş ($1 \times 10^{-5} - 3 \times 10^{-5}$) ve `ReduceLROnPlateau` mekanizması kullanılarak eğitim sürecine bağlı olarak dinamik biçimde azaltılmıştır. Eğitim sırasında en iyi performansın elde edildiği epoch'a ait ağırlıklar `ModelCheckpoint` ile kaydedilmiş, `EarlyStopping` mekanizması ile aşırı öğrenmenin önüne geçilmiştir. Uygulanan bu iki aşamalı eğitim yaklaşımı, tüm modellerde doğruluk, AUC ve F1 skoru değerlerinin belirgin biçimde artmasına katkı sağlamış ve daha kararlı bir genelleme performansı elde edilmesini mümkün kılmıştır.

4.1.4 Ortak eğitim dinamikleri

Beş modelde gözlemlenen temel ortak öğrenme davranışları incelendiğinde, eğitim sürecinin belirli ortak dinamikler etrafında şekillendiği görülmektedir. MobileNetV2, ResNet-50 ve InceptionV3 modellerinde ilk 5–10 epoch boyunca doğrulama doğruluğunda hızlı bir artış gözlenmiş; buna karşılık EfficientNet-B0 ve U-Net Encoder mimarilerinin daha yavaş bir

yakınsama davranışı sergilediği belirlenmiştir. Bu durum, özellikle hafif ve önceden güçlü temsiller öğrenmiş mimarilerin erken aşamada daha hızlı öğrenme eğilimi gösterdiğini ortaya koymaktadır.

Modellerin büyük bir kısmında doğrulama doğruluğunun yaklaşık 20–35. epoch aralığında plato davranışı sergilediği gözlemlenmiştir. Bu aşamada InceptionV3 mimarisi neredeyse hatasız bir doğrulama performansı üretirken, U-Net Encoder mimarisi görece daha düşük plato değerlerine ulaşmış ve doğrulama doğruluğu yaklaşık 0.55 seviyesinde kalmıştır. Bu farklılık, mimarilerin sınıflandırma problemine yapısal uyumluluk düzeylerinin öğrenme sürecine doğrudan yansıdığını göstermektedir.

Eğitim sürecinde tüm modellerde kullanılan özelleştirilmiş sınıflandırıcı bloğun öğrenme davranışı üzerinde belirgin bir etkisi olduğu gözlenmiştir. Dense(512) - Dropout - Dense(256) yapısının, düşük öğrenme oranı ve erken durdurma mekanizmaları ile birlikte kullanılması, modellerin daha istikrarlı ve kontrollü bir öğrenme süreci sergilemesini sağlamıştır. Bu yapı, özellikle eğitim ve doğrulama performansları arasındaki farkın sınırlı tutulmasına katkıda bulunmuştur.

Aşırı öğrenme eğilimleri değerlendirildiğinde, MobileNetV2 ve EfficientNet-B0 modellerinde belirgin bir overfitting davranışı gözlenmemiştir. ResNet-50 ve InceptionV3 mimarilerinde ise doğrulama metriklerinde minimal dalgalanmalar meydana gelmiş, ancak bu durum EarlyStopping mekanizması ile etkin biçimde kontrol altına alınmıştır. Buna karşılık, U-Net Encoder mimarisinde sınıf dengesizliği nedeniyle doğrulama metriklerinde daha belirgin dalgalanmalar gözlenmiş ve bu durum modelin genelleme performansını olumsuz yönde etkilemiştir.

4.1.5 Modeller arası genel eğitim gözlemleri

Modeller arası genel eğitim gözlemleri incelendiğinde, her mimarinin öğrenme hızı, genelleme kabiliyeti ve kararlılığı açısından farklı karakteristikler sergilediği görülmektedir. MobileNetV2, incelenen modeller arasında en hızlı yakınsayan mimari olarak öne çıkmış; doğrulama doğruluğu yaklaşık %97 seviyesine ulaşmış ve test kümesinde yaklaşık %73 doğruluk elde etmiştir. Yaklaşık 3.4 milyon parametreye sahip olması sayesinde oldukça hafif bir yapı sunan MobileNetV2, eğitim süreci boyunca son derece stabil bir öğrenme eğrisi sergilemiştir.

EfficientNet-B0, orta hızda yakınsayan ve dengeli bir öğrenme davranışı gösteren bir model olarak değerlendirilmiştir. Doğrulama doğruluğu yaklaşık %75–76 aralığında, test doğruluğu ise yaklaşık %65–68 seviyelerinde kalmıştır. AUC değerleri görece yüksek olmasına rağmen, doğruluk açısından MobileNetV2 ve InceptionV3 mimarilerinin gerisinde kalmıştır. ResNet-50 mimarisi ise yaklaşık 23 milyon parametreye sahip daha ağır bir model olup, doğrulama kümesinde yaklaşık %81 doğruluk ve test kümesinde yaklaşık %72 doğruluk elde etmiştir. Derin yapısı sayesinde güçlü bir ayrıştırma kapasitesi sunmasına rağmen, eğitim süresinin daha uzun olduğu gözlenmiştir.

InceptionV3 mimarisi, doğrulama kümesinde neredeyse kusursuz bir performans sergileyerek yaklaşık %99.95 doğruluk ve 1.000 AUC değerine ulaşmıştır. Test kümesinde ise yaklaşık %75 doğruluk elde edilmiştir. Çok ölçekli filtreleme yapısı sayesinde özellikle erken evre Alzheimer sinyallerini yakalamada başarılı olduğu görülmüş; ancak uzun eğitim süresi ve yüksek parametre maliyeti, bu mimarinin hesaplama açısından daha ağır bir yapı sunduğunu ortaya koymuştur. U-Net Encoder mimarisi ise doğrulama kümesinde yaklaşık %55, test kümesinde ise yaklaşık %52 doğruluk elde etmiş; özellikle erken ve hafif evrelerde belirgin sınıf karışmaları göstermiştir. Buna karşın, Moderate sınıfında elde edilen yüksek AUC değerleri, bu mimarinin belirli sınıflar için ayırt edici bilgi taşıyabildiğini, ancak genel sınıflandırma başarımının sınırlı kaldığını göstermektedir.

4.1.6 Genel değerlendirme

Elde edilen eğitim sonuçları, beş farklı derin öğrenme modeli arasında belirgin performans farklılıkları olduğunu ortaya koymaktadır. Doğrulama performansı açısından InceptionV3 mimarisi en güçlü model olarak öne çıkarken, MobileNetV2 doğruluk, AUC ve hafiflik kriterleri birlikte değerlendirildiğinde en dengeli mimariyi sunmaktadır. ResNet-50, sahip olduğu güçlü temsil kapasitesi sayesinde kararlı bir performans sergilemekle birlikte, yüksek parametre sayısı nedeniyle daha ağır bir model olarak değerlendirilmiştir. EfficientNet-B0, hesaplama verimliliği açısından öne çıkmasına karşın test kümesi performansının orta düzeyde kaldığı görülmüştür. U-Net Encoder mimarisi ise sınıflandırma odaklı bir yapı olmaması nedeniyle bu görevde sınırlı bir başarı sergilemiştir. Bu gözlemler doğrultusunda, Alzheimer evre sınıflandırmasında hafiflik, hız ve doğruluk arasındaki denge dikkate alındığında MobileNetV2 ve InceptionV3 mimarilerinin en yüksek potansiyele sahip modeller olduğu sonucuna varılmıştır.

4.2 Modellerin Karşılaştırmalı Deneysel Performansı

Bu bölümde, Alzheimer hastalığının dört klinik evresinin sınıflandırılmasında kullanılan beş farklı derin öğrenme mimarisinin (MobileNetV2, EfficientNet-B0, ResNet-50, InceptionV3 ve U-Net Encoder) deneysel çıktıları karşılaştırmalı olarak sunulmuştur. Tüm modeller aynı veri kümesi, aynı eğitim/doğrulama/test ayrımı, aynı ön işleme adımları ve aynı iki aşamalı eğitim stratejisi altında değerlendirilmiştir. Böylece ortaya çıkan performans farklarının yalnızca mimari farklılıklardan kaynaklanması sağlanmıştır.

Ayrıca, modellerin karşılaştırılması yalnızca nihai doğruluk değerleri üzerinden değil; öğrenme dinamikleri, sınıf dengesizliğine karşı gösterilen duyarlılık ve eğitim sürecindeki kararlılık gibi kriterler de göz önünde bulundurularak gerçekleştirilmiştir. Bu kapsamda, her bir mimarinin Alzheimer evrelerine ait ayırt edici özellikleri ne ölçüde yakalayabildiği, hem istatistiksel metrikler hem de eğitim sürecine ait grafiksel analizler aracılığıyla bütüncül bir bakış açısıyla değerlendirilmiştir.

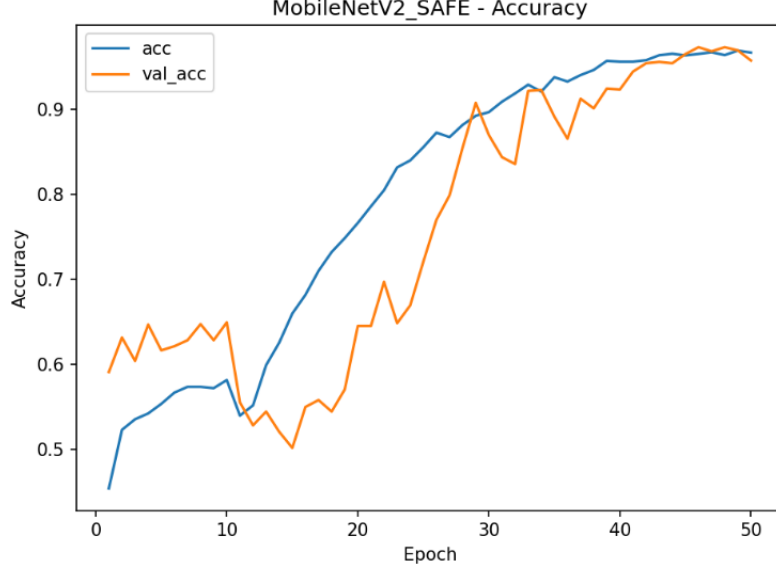
4.2.1 ACC–Loss eğrilerinin karşılaştırmalı analizi

Bu bölümde modellerin öğrenme davranışı, yakınsama özellikleri ve aşırı öğrenme eğilimleri ACC ve loss eğrileri üzerinden değerlendirilmiştir.

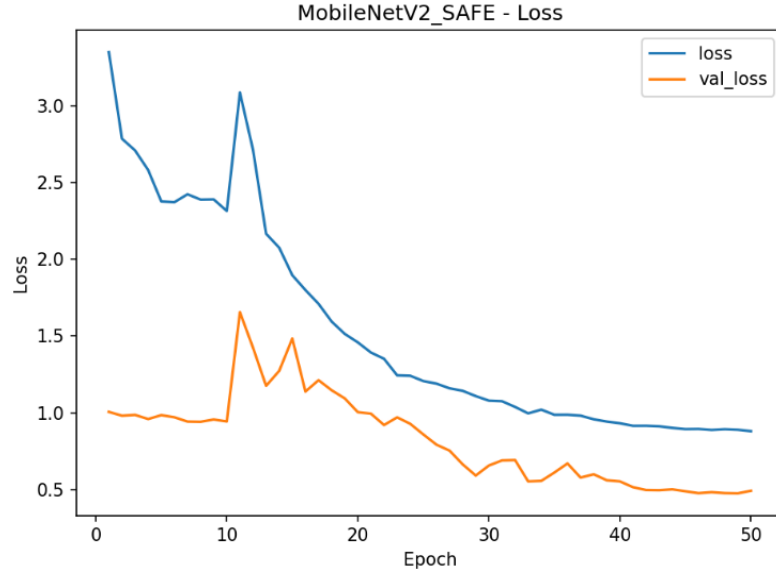
Eğitim ve doğrulama kümelerine ait ACC ve loss eğrilerinin birlikte incelenmesi, modellerin öğrenme sürecindeki kararlılığını ve genelleme yeteneğini ortaya koymaktadır. Eğriler arasındaki ayrışma düzeyi, aşırı öğrenme (overfitting) riskinin değerlendirilmesinde kritik bir gösterge olarak ele alınmış; dengeli ve paralel bir seyir izleyen modeller daha güvenilir kabul edilmiştir. Bu bağlamda, erken durdurma (early stopping) uygulamalarının etkinliği de eğri davranışları üzerinden yorumlanmıştır.

Ayrıca, ACC eğrilerinin erken dönemlerdeki yükselme hızı, modellerin öğrenme kapasitesi ve optimizasyon sürecine adaptasyon kabiliyeti hakkında önemli ipuçları sunmaktadır. Loss eğrilerinin ise epoch ilerledikçe düzenli ve monoton bir azalma göstermesi, öğrenme sürecinin kararlı biçimde ilerlediğini ve modelin hata minimizasyonunu etkin şekilde gerçekleştirdiğini göstermektedir. Eğitim ve doğrulama loss eğrileri arasındaki farkın sınırlı kalması, modelin yalnızca eğitim verisine özgü kalıpları ezberlemediğini, aksine genelleştirilebilir temsiller öğrendiğini ortaya koymaktadır. Bu doğrultuda, ACC–loss

eğrilerinin birlikte değerlendirilmesi, modellerin performanslarının yalnızca nihai metrikler üzerinden değil, tüm eğitim süreci boyunca bütüncül bir bakış açısıyla analiz edilmesine olanak sağlamaktadır.



(a)



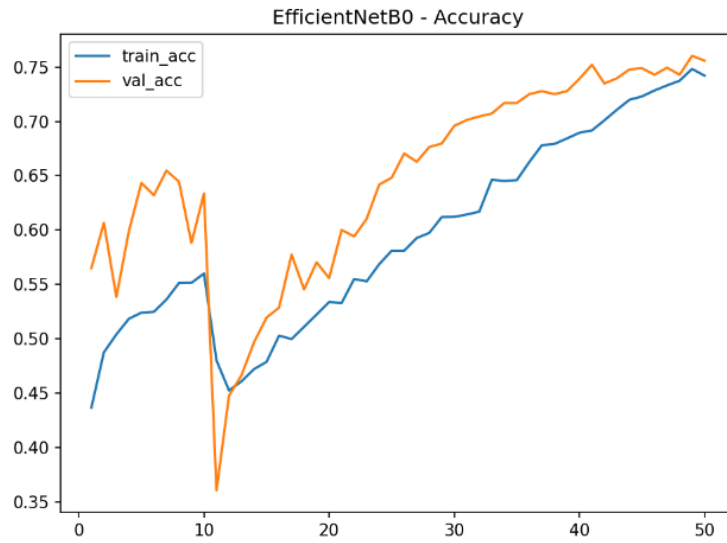
(b)

Şekil 4.1: MobilenetV2 modelinin eğitim süreci; a) acc eğrileri, b) loss eğrileri

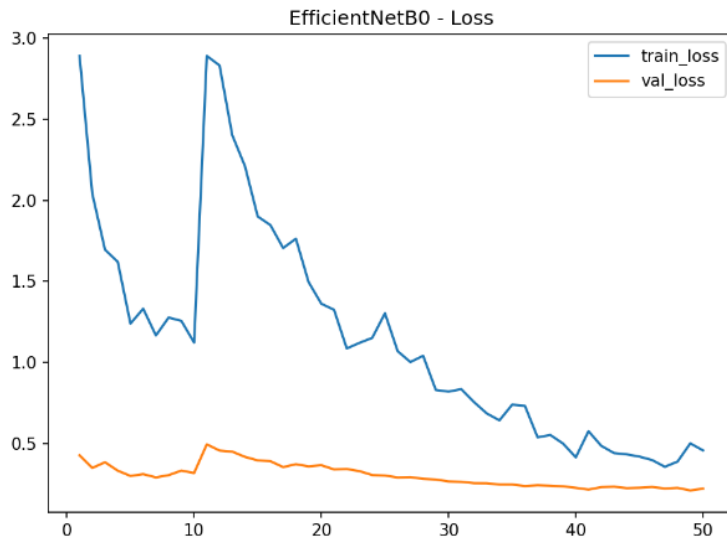
MobileNetV2, tüm modeller arasında en hızlı yakınsayan yapı olmuştur. Eğitim doğruluğu her epoch'ta düzenli biçimde artmış, doğrulama doğruluğu ise 36. epoch'ta %97.35 seviyesine ulaşmıştır. Eğitim ve doğrulama eğrilerinin birbirine paralel seyretmesi, modelin

overfitting eğilimi göstermediğini ve öğrenmenin stabil biçimde ilerlediğini göstermektedir. Loss eğrileri de benzer şekilde düzenli bir iyileşme ortaya koymuştur.

Ayrıca, doğrulama loss değerlerinin eğitim loss'u ile uyumlu biçimde azalması, modelin yalnızca eğitim verisine özgü örüntüler öğrenmediğini, aynı zamanda genelleme kabiliyetinin güçlü olduğunu göstermektedir. Bu istikrarlı öğrenme davranışı, MobileNetV2 mimarisinin parametrik verimliliği sayesinde sınırlı veri koşullarında dahi güvenilir sonuçlar üretebildiğini ortaya koymaktadır.



(a)



(b)

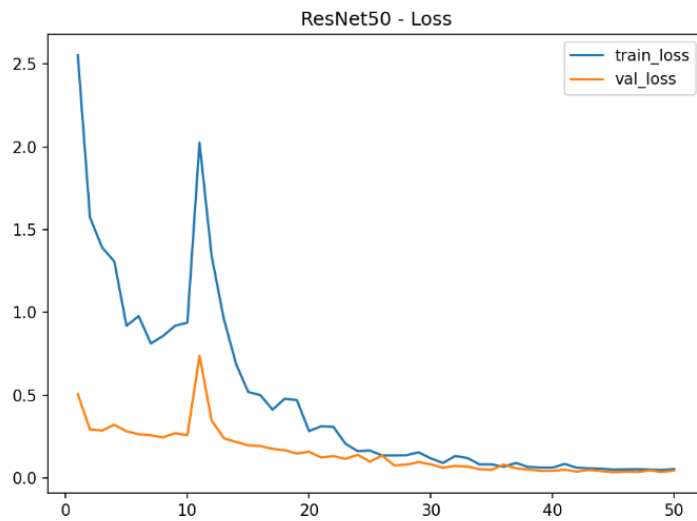
Şekil 4.2: EfficientNet-B0 modelinin eğitim süreci; a) acc eğrileri, b) loss eğrileri

EfficientNet-B0, eğitim sürecinde yavaş fakat dengeli bir yakınsama sergilemiştir. Doğrulama doğruluğu 20. epoch sonrasında belirgin biçimde yükselmiş ve yaklaşık %76 seviyesinde stabil hâle gelmiştir. Eğitim ve doğrulama eğrileri arasındaki farkın düşük olması, modelin iyi düzenlendiğini ve aşırı öğrenmeden kaçındığını göstermektedir.

Bununla birlikte, doğrulama loss eğrisinin eğitim loss'u ile paralel bir seyir izlemesi, modelin öğrenme sürecinin kararlı ve kontrollü biçimde ilerlediğini göstermektedir. Ancak elde edilen doğrulama doğruluğunun görece düşük seviyede kalması, EfficientNet-B0 mimarisinin bu veri kümesi özelinde daha derin varyantlara ya da farklı ince ayar stratejilerine ihtiyaç duyabileceğine işaret etmektedir.



(a)

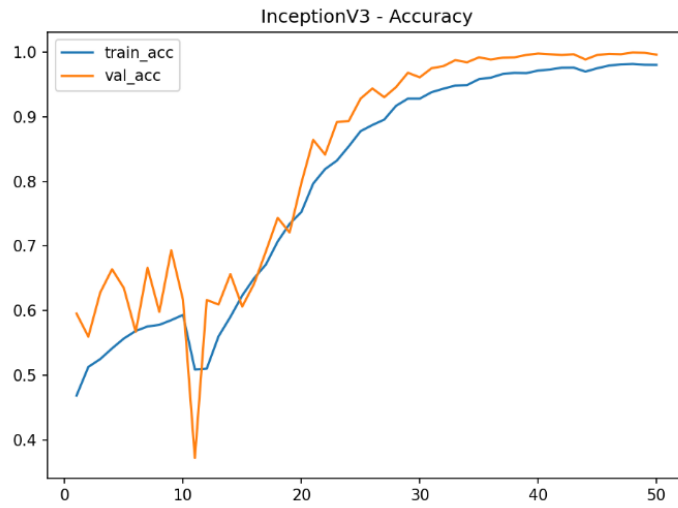


(b)

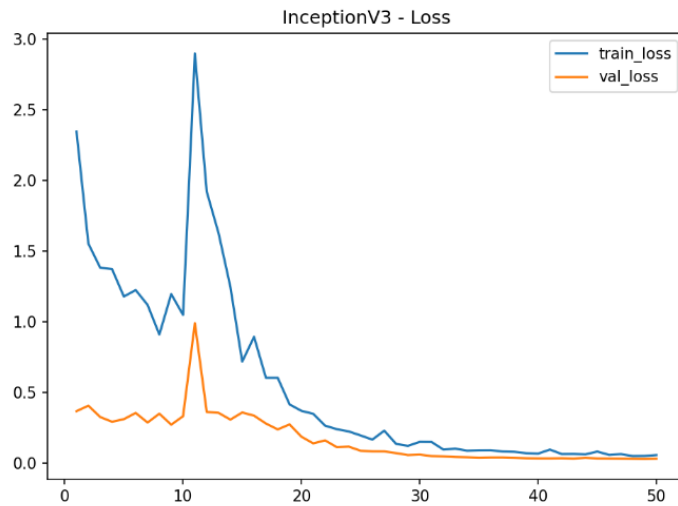
Şekil 4.3: ResNet-50 modelinin eğitim süreci; a) acc eğrileri, b) loss eğrileri

ResNet-50, daha derin mimarisi nedeniyle daha yavaş yakınsamış, ancak epoch 30 civarında doğrulama doğruluğunu %81 seviyesine taşımıştır. Derin katman yapısına rağmen erken durdurma mekanizması aşırı öğrenmeyi başarılı biçimde engellemiştir. ACC ve loss eğrileri kararlı bir optimizasyon sürecini işaret etmektedir.

Ayrıca, eğitim ve doğrulama eğrileri arasındaki uyum, artık bağlantıların (skip connections) gradyan sönümlenmesi problemini etkin biçimde azalttığını ve derin ağlarda öğrenme stabilitesini artırdığını göstermektedir. Bu özellikleriyle ResNet-50, daha derin temsil gerektiren Alzheimer evrelerinin ayrıştırılmasında güçlü bir aday olarak değerlendirilmektedir.



(a)

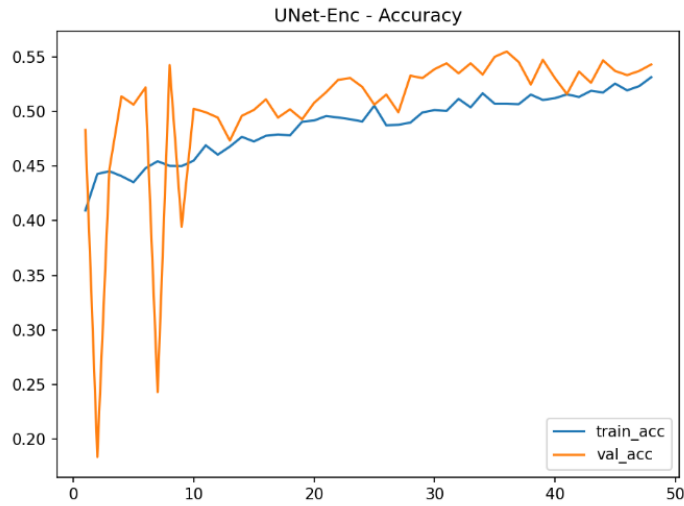


(b)

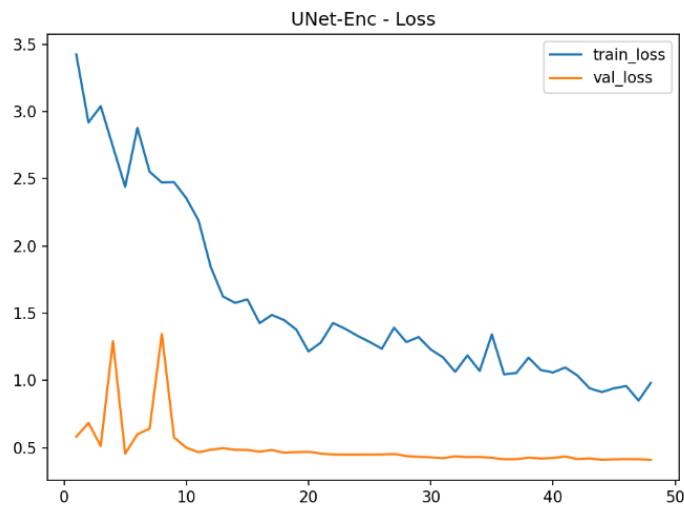
Şekil 4.4: InceptionV3 modelinin eğitim süreci; a) acc eğrileri, b) loss eğrileri

InceptionV3, doğrulama kümesinde olağanüstü yüksek bir doğruluk sergileyerek neredeyse son derece başarılı bir öğrenme davranışı göstermiştir. ACC eğrisi 35–40. epoch arasında %99.9 seviyesine ulaşırken loss eğrisi minimuma yakın değerlere inerek stabil bir seyir izlemiştir. Çok ölçekli filtreleme yapısına sahip mimarinin Alzheimer MRI verisine son derece iyi uyum sağladığı görülmektedir.

Eğitim ve doğrulama eğrileri arasındaki yüksek örtüşme, modelin aşırı öğrenme belirtisi göstermeden güçlü bir genelleme performansı sergilediğini ortaya koymaktadır. Bu nedenle, InceptionV3'ün performansı yalnızca ACC ve loss eğrileriyle değil, sınıf bazlı metrikler ve karışıklık matrisi analizleriyle birlikte bütüncül olarak yorumlanmıştır.



(a)



(b)

Şekil 4.5: U-Net modelinin eğitim süreci; a) acc eğrileri, b) loss eğrileri

U-Net Encoder, diğer modellerle kıyaslandığında daha düşük doğrulama başarımına sahiptir. ACC yaklaşık %55 civarında plato yapmış, özellikle early/mild evrelerin ayırımında sınırlı temsil kapasitesi nedeniyle kararsızlık göstermiştir. Buna rağmen loss eğrilerinde aşırı öğrenmeye işaret eden ciddi bir ayrışma gözlenmemiştir.

Bu sonuçlar, U-Net mimarisinin özgün olarak piksel düzeyinde segmentasyon görevleri için tasarlanmış olmasının, saf sınıflandırma problemlerinde temsil gücünü sınırlayabildiğini göstermektedir. Encoder yapısı genel örüntüleri yakalayabilse de, sınıf ayırımına yönelik yüksek seviyeli anlamsal özelliklerin çıkarılmasında diğer sınıflandırma odaklı mimarilere kıyasla daha zayıf kalmıştır. Dolayısıyla U-Net Encoder, bu çalışmada yardımcı veya karşılaştırmalı bir model olarak değerlendirilmiştir.

4.2.2 Doğrulama ve test performanslarının karşılaştırmalı analizi

Bu bölümde modellerin doğrulama ve test kümeleri üzerindeki performans metrikleri karşılaştırmalı olarak ele alınmaktadır.

Modellerin doğrulama ve test kümeleri üzerindeki performansları, kullanılan değerlendirme metrikleri açısından tablo formatında özetlenmiştir. Bu sayede, her bir modelin farklı veri kümelerindeki başarı düzeyleri daha net biçimde karşılaştırılabilmektedir. Ayrıca, doğrulama ve test sonuçlarının birlikte incelenmesi, modellerin genelleme yetenekleri hakkında önemli bilgiler sunmaktadır.

Özellikle doğrulama ve test performansları arasındaki tutarlılık, modellerin eğitim verisine aşırı uyum sağlayıp sağlamadığını değerlendirmede kritik bir ölçüt olarak ele alınmıştır. Doğrulama kümesinde yüksek başarı gösterip test kümesinde belirgin performans kaybı yaşayan modeller, genelleme açısından daha temkinli biçimde yorumlanmıştır. Bu karşılaştırmalı analiz, nihai modelin seçilmesinde belirleyici bir rol oynamıştır.

Bu bağlamda, doğrulama ve test performansları arasında dengeli bir dağılım sergileyen modeller, klinik uygulamalara aktarılabilirlik açısından daha güvenilir kabul edilmiştir. Test kümesinde tutarlı sonuçlar üreten mimariler, farklı hasta profilleri ve görüntüleme koşullarında da benzer başarı gösterebilecek yapılar olarak değerlendirilmiştir. Ayrıca, bazı modellerde doğrulama başarımının yüksek olmasına karşın test performansındaki düşüşler, veri çeşitliliği ve örnek dağılımının model genellemesi üzerindeki etkisini açıkça ortaya

koymuřtur. Bu sonular, model seiminin yalnızca doėrulama metriklerine deėil, test kümesi üzerindeki gereki performansa dayandırılması gerektiėini vurgulamaktadır.

Tablo 4.1: MobileNetV2 modelinin doėrulama ve test performans özetini

| Metrik | Doėrulama | Test |
|------------|-----------|--------|
| Doėrululuk | %97.35 | %73.10 |
| Kesinlik | %97.86 | %73.38 |
| Duyarlılık | %98.43 | %66.53 |
| F1 Skoru | %98.12 | %68.81 |
| Makro AUC | %99.89 | %91.65 |

MobileNetV2 doėrulama kümesinde %97.35 doėrululuk ve %99'a yakın AUC sergileyen model, test kümesinde %73 doėrululuk ve Macro-AUC=0.9165 elde etmiştir. Doėrulama ve test sonuları arasındaki performans farkı, veri kümesindeki sınıf dengesizliėi ve özellikle Moderate sınıfının sınırlı sayıda örnek iermesiyle ilişkilendirilebilir. Bununla birlikte, test kümesinde elde edilen performans deėerleri modelin genelleme kapasitesinin güçlü olduėunu göstermektedir.

Tablo 4.2: EfficientNet-B0 modelinin doėrulama ve test performans özetini

| Metrik | Doėrulama | Test |
|------------|-----------|--------|
| Doėrululuk | %76.03 | %65.68 |
| Kesinlik | %81.12 | %67.79 |
| Duyarlılık | %74.21 | %58.74 |
| F1 Skoru | %77.03 | %62.08 |
| Makro AUC | %93.35 | %83.38 |

EfficientNet-B0 doėrulama setinde %76 doėrululuk ve %93 Macro-AUC üretmiştir. Test kümesinde doėrululuk %65, Macro-AUC %83 olarak ölçülmüřtür. Test performansı MobileNetV2'den düşük olsa da mimari verimlilik aısından dengeli bir yapı sunmaktadır.

Bu durum, EfficientNet-B0 mimarisinin sınırlı test performansına rağmen, parametre verimliliği ve genel performans dengesi açısından uygulanabilir bir alternatif olduğunu göstermektedir.

Tablo 4.3: ResNet-50 modelinin doğrulama ve test performans özeti

| Metrik | Doğrulama | Test |
|------------|-----------|--------|
| Doğruluk | %98.92 | %73.10 |
| Kesinlik | %99.28 | %85.43 |
| Duyarlılık | %99.16 | %63.37 |
| F1 Skoru | %99.21 | %69.97 |
| Makro AUC | %99.99 | %91.40 |

ResNet-50 doğrulama setinde %98.92 doğruluk ve %99.99 AUC elde etmiştir. Model, validation aşamasında neredeyse hatasız bir ayrıştırma performansı göstermiştir. Test setinde ise doğruluk %73.10 ve Macro-AUC %91.40 olarak ölçülmüştür. Bu fark, özellikle VeryMild ve Mild evrelerindeki sınıf dengesizliği ve dağılım kayması (distribution shift) ile ilişkilidir. Derin mimarinin sağladığı yüksek temsil kapasitesi sayesinde ResNet-50, test kümesinde dahi güçlü bir ayrıştırma performansı sunmuş; ancak sınıflar arası örnek dengesizliği test performansında varyansı artırmıştır.

Tablo 4.4: InceptionV3 modelinin doğrulama ve test performans özeti

| Metrik | Doğrulama | Test |
|------------|-----------|--------|
| Doğruluk | %99.95 | %75.29 |
| Kesinlik | %99.97 | %75.75 |
| Duyarlılık | %99.96 | %78.22 |
| F1 Skoru | %99.97 | %76.41 |
| Makro AUC | %100 | %92.24 |

InceptionV3 doğrulama aşamasında neredeyse kusursuz (%99.95 doğruluk, AUC=1.00) performans üretmiş; test setinde ise %75 doğruluk ve %92 Macro-AUC elde etmiştir. Çok ölçekli mimarinin özellikle ileri evre ayırımında güçlü olduğu görülmektedir.

Tablo 4.5 U-Net modelinin doğrulama ve test performans özeti

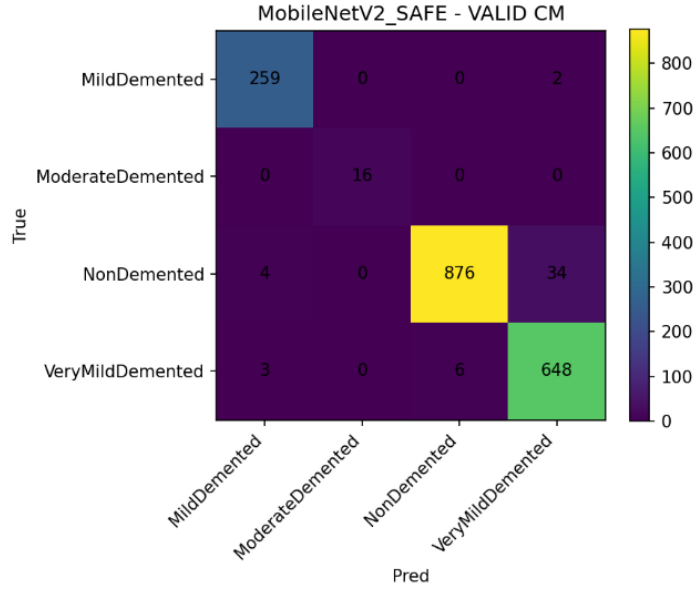
| Metrik | Doğrulama | Test |
|------------|-----------|--------|
| Doğruluk | %55.52 | %52.07 |
| Kesinlik | %66.88 | %38.28 |
| Duyarlılık | %48.57 | %41.94 |
| F1 Skoru | %42.89 | %37.47 |
| Makro AUC | %82.64 | %78.89 |

Doğrulama doğruluğu %55, test doğruluğu %52 olup diğer modellere kıyasla en düşük performansı sergilemiştir. Buna karşın ModerateDemented sınıfında AUC'nin yüksek olması, modelin belirgin ileri evre sinyallerini yakalamada daha başarılı olduğunu göstermektedir.

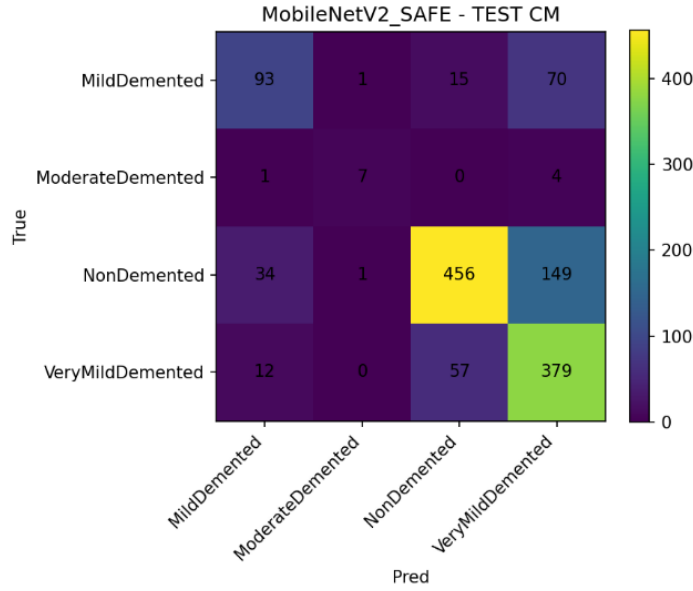
4.2.3 Confusion matrix sonuçlarının karşılaştırmalı analizi

Bu bölümde modellerin sınıflar arası karıştırma eğilimleri analiz edilmekte; özellikle sınıf ayırımındaki güçlü ve zayıf yönleri ortaya koymak açısından önemli ipuçları sunulmaktadır.

Confusion matrix analizleri, her bir modelin Alzheimer hastalığının klinik evrelerini ayırt etme başarısını sınıf bazında değerlendirmeye olanak tanımaktadır. Diyagonal elemanların yüksekliği, doğru sınıflandırma oranlarını temsil ederken; diyagonal dışı hücreler, modellerin hangi evreler arasında kararsız kaldığını açık biçimde ortaya koymaktadır. Bu bağlamda, özellikle komşu klinik evreler arasında gerçekleşen yanlış sınıflandırmalar, hastalığın ilerleyici doğası ve evreler arası yapısal benzerlikler ile ilişkilendirilerek yorumlanmıştır. Elde edilen bulgular, modellerin yalnızca genel doğruluk performanslarını değil, klinik açıdan kritik sınıflar arasındaki ayırt edicilik düzeylerini de değerlendirmeye imkân sağlamaktadır.



(a)



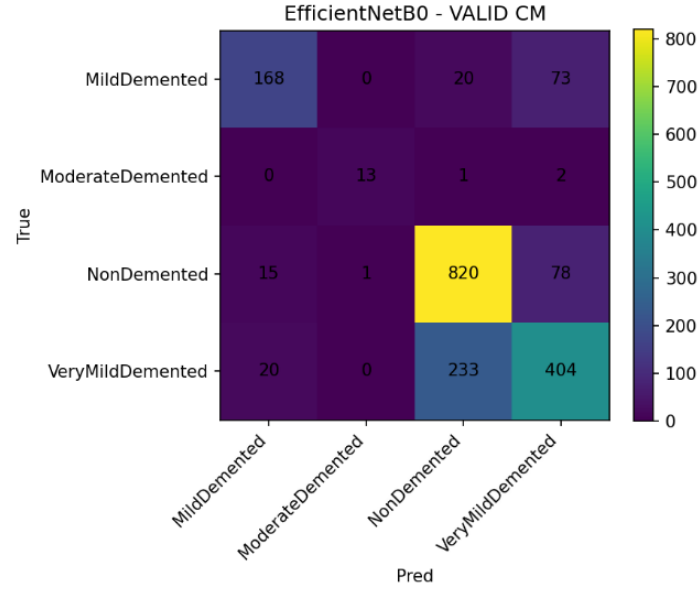
(b)

Şekil 4.6: MobilenetV2 modelinin doğrulama ve test veri setleri için oluşturulan karışıklık matrisleri; a) doğrulama karışıklık matrisi, b) test karışıklık matrisi

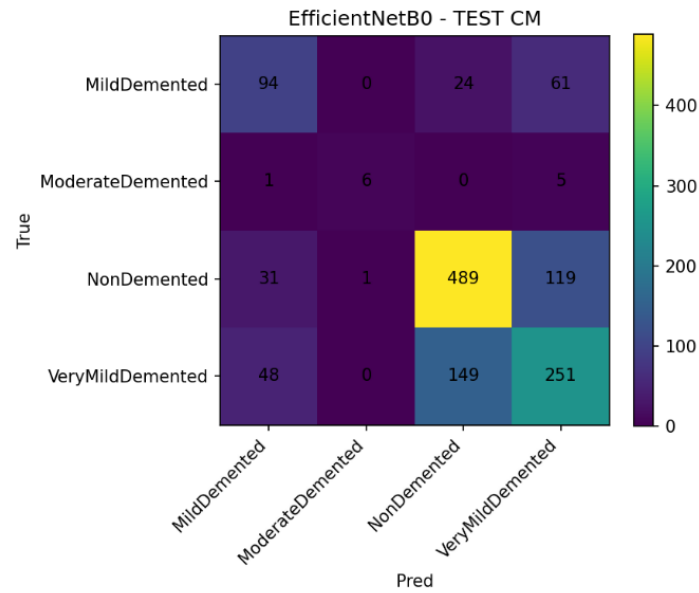
MobileNetV2 modeli, doğrulama kümesinde neredeyse hatasız bir performans sergilemiştir. Ancak test kümesinde özellikle Mild ve Moderate sınıflarının, NonDemented ve VeryMild sınıfları ile yer yer karıştığı gözlemlenmiştir. Buna rağmen, genel doğruluk değeri açısından en yüksek performans MobileNetV2 modelinde elde edilmiştir.

Özellikle test kümesinde gözlenen sınıflar arası karışmalar, Alzheimer evreleri arasındaki klinik sınırların doğası gereği keskin olmamasıyla ilişkilendirilebilir. Bununla birlikte, hataların büyük ölçüde komşu evreler arasında gerçekleşmesi, modelin rastgele tahminler

üretmek yerine klinik olarak anlamlı ve tutarlı bir sınıflandırma davranışı sergilediğini göstermektedir. Bu durum, MobileNetV2'nin klinik açıdan kabul edilebilir ve güvenilir bir performans sunduğunu desteklemektedir.



(a)

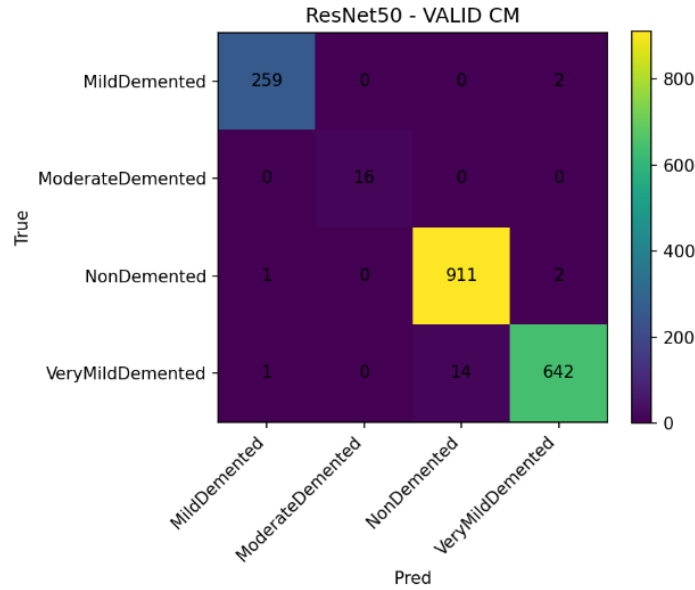


(b)

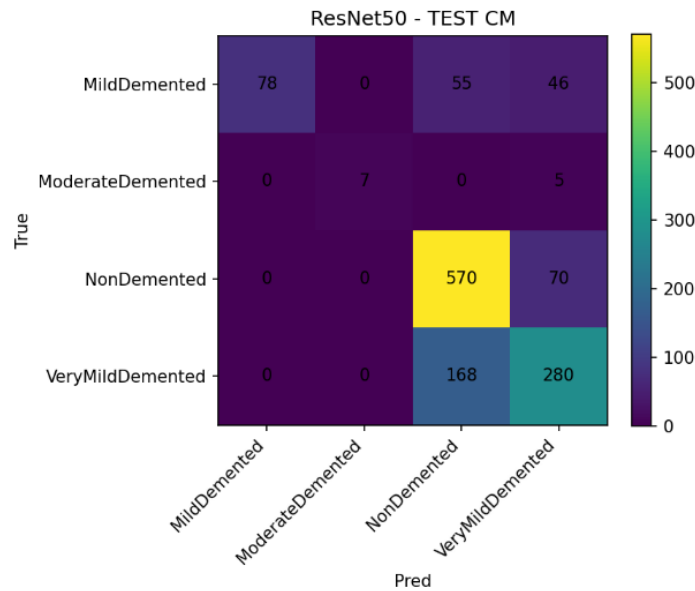
Şekil 4.7: EfficientNet-B0 modelinin doğrulama ve test veri setleri için oluşturulan karışıklık matrisleri; a) doğrulama karışıklık matrisi, b) test karışıklık matrisi

EfficientNet-B0 modeli, doğrulama kümesinde güçlü bir sınıf ayrımı sergilemiş olsa da test kümesinde özellikle VeryMild ve NonDemented sınıfları arasında belirgin karışmalar gözlemlenmiştir. Bu durum, klinik olarak geçiş niteliği taşıyan erken evrelerde modelin ayrım yapmakta zorlanabildiğini göstermektedir.

Test kümesinde gözlenen bu karışmalar, erken evre Alzheimer bulgularının yapısal olarak NonDemented sınıfıyla yüksek benzerlik göstermesinden kaynaklanmış olabilir. EfficientNet-B0'ın bu sınırlılığı, mimarinin kompakt yapısı nedeniyle ince ayrıntılara duyarlılığının görece düşük kalmasıyla ilişkilendirilebilir. Buna karşın, ileri evre sınıflarda daha tutarlı tahminler üretmesi, modelin belirgin patolojik örüntüleri yakalama konusunda yeterli bir temsil gücüne sahip olduğunu göstermektedir.



(a)

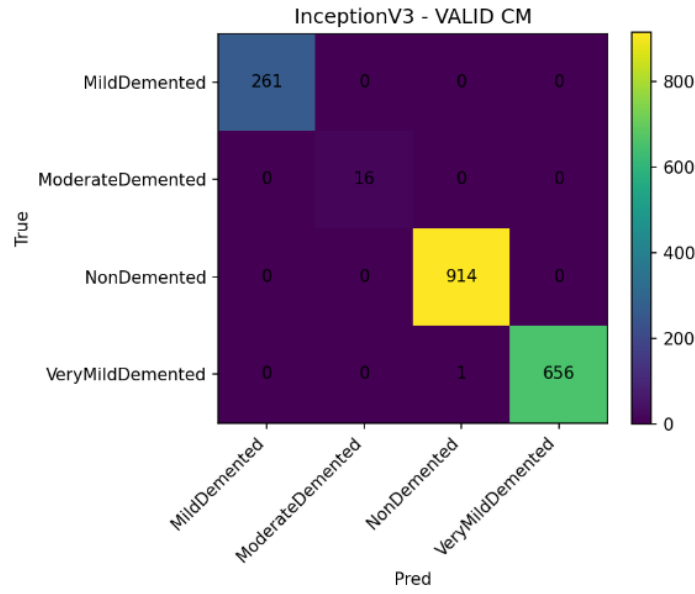


(b)

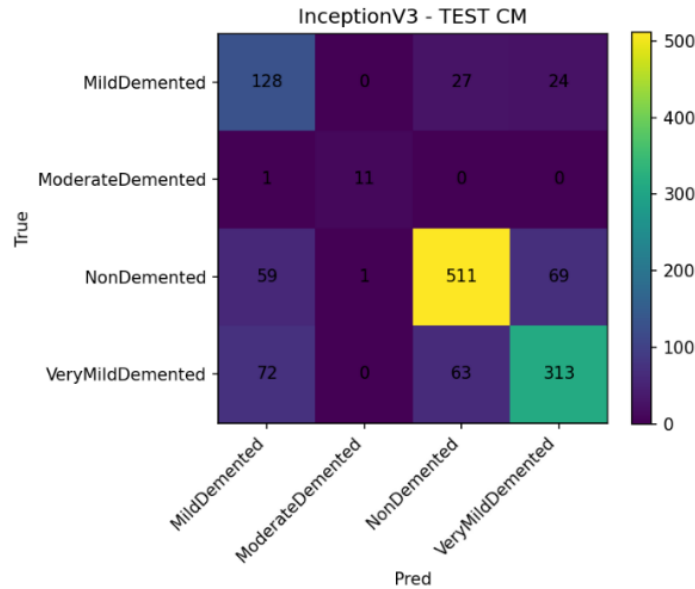
Şekil 4.8: ResNet-50 modelinin doğrulama ve test veri setleri için oluşturulan karışıklık matrisleri; a) doğrulama karışıklık matrisi, b) test karışıklık matrisi

ResNet-50 özellikle NonDemented ve Moderate sınıflarında yüksek doğruluk sağlamıştır. VeryMild ve Mild sınıfları arası karışmalar ResNet için belirgin bir zorluk oluşturmaktadır.

Bu karışmalar, VeryMild ve Mild evrelerin yapısal özelliklerinin yüksek benzerlik göstermesinden ve sınırlarının klinik olarak da net biçimde ayrılmamasından kaynaklanabilir. ResNet-50'nin derin temsil kapasitesi belirgin patolojik değişimleri başarılı biçimde yakalarken, erken ve geçiş evrelerinde daha ince ayrıntılara duyarlı ek düzenleme ya da veri artırma stratejilerine ihtiyaç duyabileceği görülmektedir. Buna rağmen, ileri evrelerde sunduğu istikrarlı performans, modelin klinik açıdan anlamlı sınıfları ayırt etmede güçlü bir aday olduğunu göstermektedir.



(a)

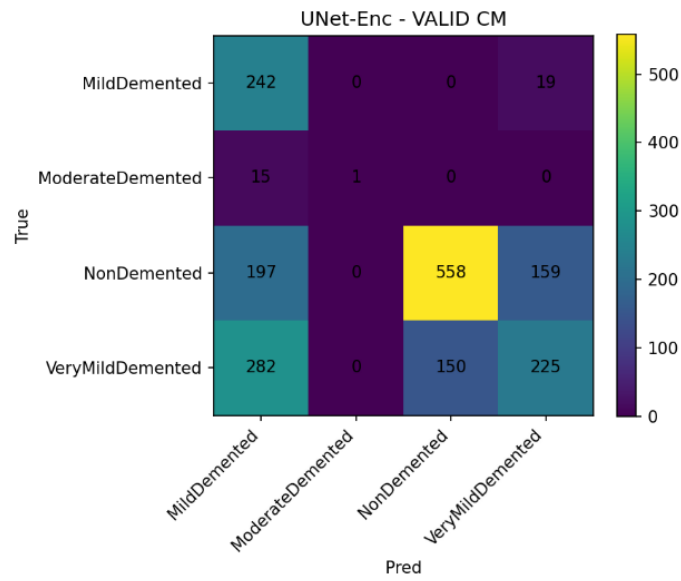


(b)

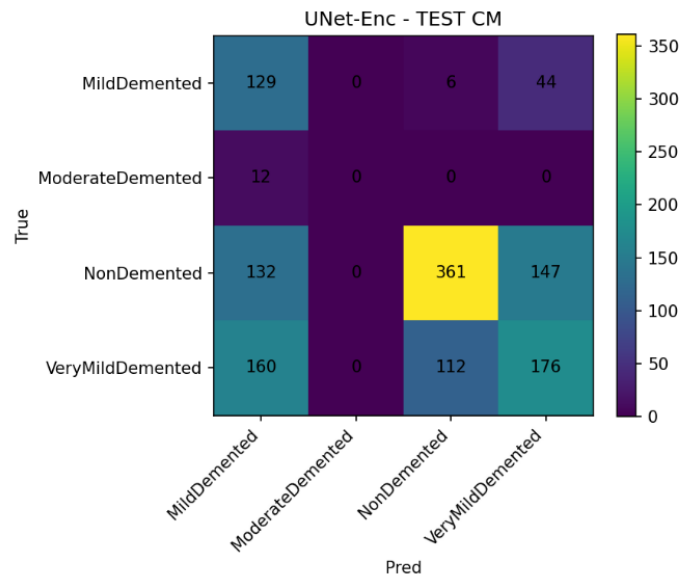
Şekil 4.9: InceptionV3 modelinin doğrulama ve test veri setleri için oluşturulan karışıklık matrisleri; a) doğrulama karışıklık matrisi, b) test karışıklık matrisi

InceptionV3 doğrulama setinde neredeyse tüm sınıfları eksiksiz ayırmıştır. Test setinde VeryMild - Mild karışmaları gözlenmiştir.

Test kümesinde gözlenen VeryMild–Mild karışmaları, hastalığın erken evrelerinde yapısal değişimlerin sınırlı ve kademeli olmasıyla uyumludur. Buna rağmen InceptionV3’ün diğer sınıflarda sergilediği yüksek ayrıştırma başarısı, çok ölçekli özellik çıkarım mekanizmasının Alzheimer’a özgü örüntüleri etkin biçimde yakalayabildiğini göstermektedir. Bu sonuçlar, modelin klinik olarak zorlu sınıf geçişlerinde dahi genel performansını koruyabildiğini ve güvenilir bir sınıflandırma sunduğunu ortaya koymaktadır.



(a)



(b)

Şekil 4.10: U-Net modelinin doğrulama ve test veri setleri için oluşturulan karışıklık matrisleri; a) doğrulama karışıklık matrisi, b) test karışıklık matrisi

U-Net Encoder, NonDemented sınıfında yüksek karışma göstermiş ve VeryMild sınıfını ayırtmakta zorlanmıştır. Moderate sınıfı ise iyi yakalanmıştır.

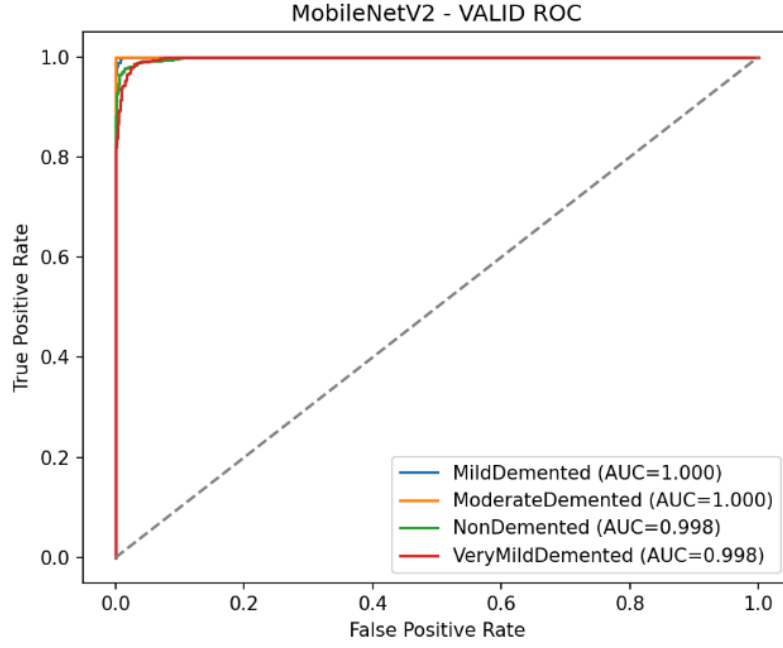
4.2.4 ROC–AUC eğrilerinin karşılaştırmalı analizi

Bu kısımda modellerin sınıflar arasındaki duyarlılık–özgüllük dengesi ROC eğrileri üzerinden değerlendirilmiştir.

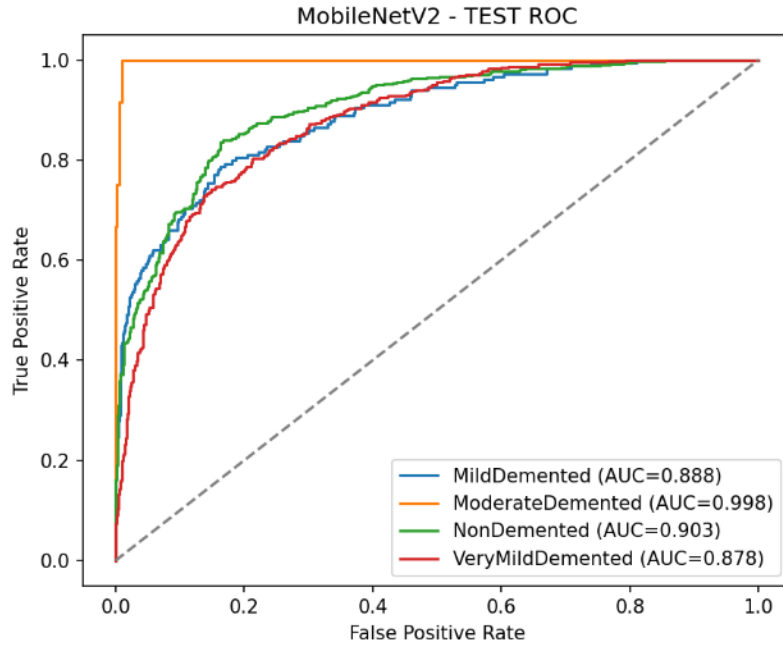
ROC eğrileri ve AUC değerleri, modellerin farklı eşik değerlerinde sınıfları ayırt edebilme kapasitesini ortaya koymaktadır.

Bu doğrultuda, ROC eğrilerinin sol üst köşeye yakınlığı, modellerin yüksek duyarlılık ve özgüllük değerlerini eş zamanlı olarak sağlayabildiğini göstermektedir. AUC değerinin 1'e yaklaşması, modelin sınıflar arasındaki ayırt edicilik gücünün yüksek olduğunu ifade ederken; daha düşük AUC değerleri, özellikle sınırda kalan klinik evrelerin ayrıştırılmasında yaşanan zorluklara işaret etmektedir. Çok sınıflı sınıflandırma problemi kapsamında ROC–AUC analizlerinin sınıf bazlı ve ortalama (macro/micro) ölçütler üzerinden değerlendirilmesi, modellerin genel performanslarının yanı sıra her bir Alzheimer evresine karşı gösterdikleri duyarlılığı daha ayrıntılı biçimde inceleme imkânı sunmaktadır.

ROC–AUC analizinin bir diğer önemli yönü, modelin yalnızca genel doğruluğunu değil, farklı sınıflar arasındaki ayırım sınırlarının ne ölçüde tutarlı olduğunu da ortaya koymasındadır. Özellikle Alzheimer evreleri gibi klinik olarak birbirine yakın ve geçiş özellikleri gösterebilen sınıflarda, ROC eğrilerinin birbirine olan yakınlığı veya ayrışması, modelin erken, orta ve ileri evreleri ayırt etme hassasiyetine dair kritik ipuçları sunmaktadır. Bu bağlamda, yüksek AUC değerine sahip modellerin yalnızca doğru sınıflandırma yapmakla kalmayıp aynı zamanda yanlış pozitif ve yanlış negatif oranlarını da dengeli biçimde minimize ettiği söylenebilir. Klinik karar destek sistemleri açısından değerlendirildiğinde, ROC–AUC performansının yüksek olması, modelin tanı sürecinde güvenilir bir yardımcı araç olarak kullanılabilirliğini artırmakta ve özellikle erken evre Alzheimer tespitinde yanlış sınıflandırma riskinin azaltılmasına katkı sağlamaktadır. Ayrıca, modeller arasında ROC eğrilerinin karşılaştırılması, hangi derin öğrenme mimarisinin sınıf ayırım sınırlarını daha keskin öğrendiğini ortaya koyarak deneysel bulguların teknik ve klinik açıdan daha bütüncül bir şekilde yorumlanmasına olanak tanımaktadır.



(a)

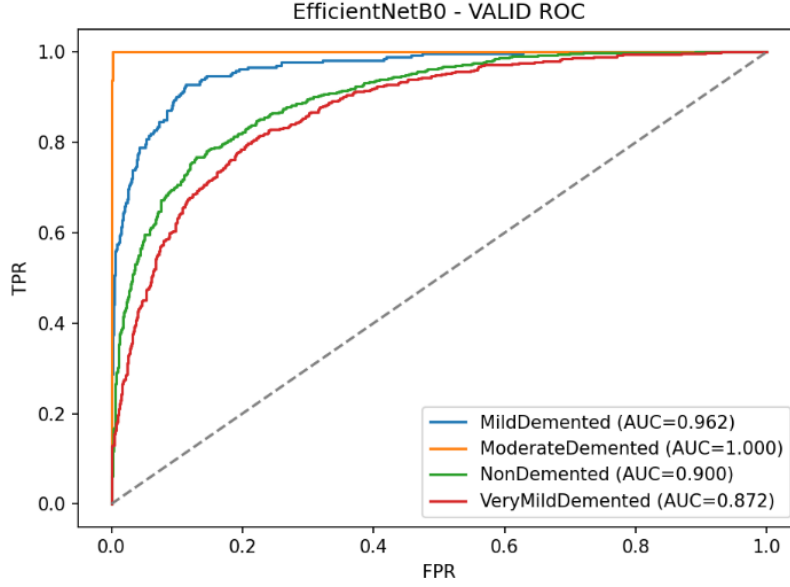


(b)

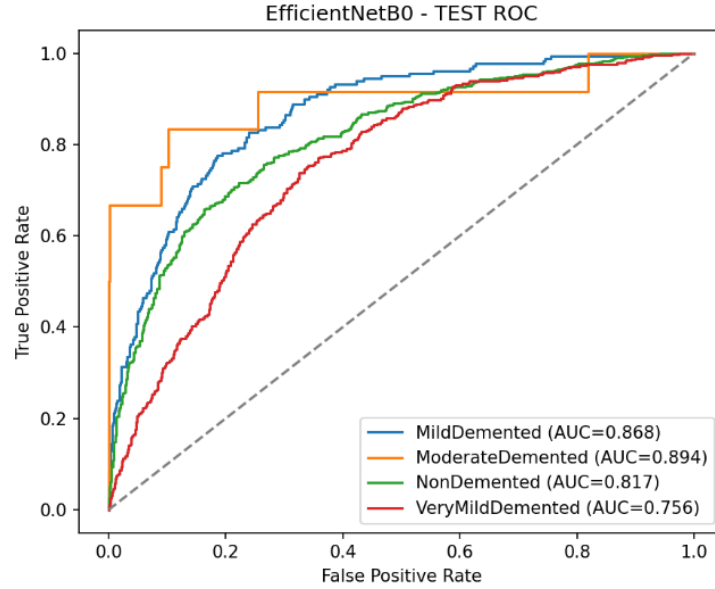
Şekil 4.11: MobileNetV2 modelinin doğrulama ve test veri setleri için oluşturulan roc eğrileri; a) doğrulama roc eğrisi, b) test roc eğrisi

Doğrulama kümesinde tüm sınıflar için $AUC \approx 1.00$, testte ise 0.88–0.93 bandında güçlü performans elde edilmiştir.

Test kümesinde AUC değerlerinin doğrulamaya kıyasla bir miktar düşmesine rağmen yüksek aralıkta kalması, MobileNetV2'nin genelleme yeteneğinin güçlü olduğunu göstermektedir.



(a)

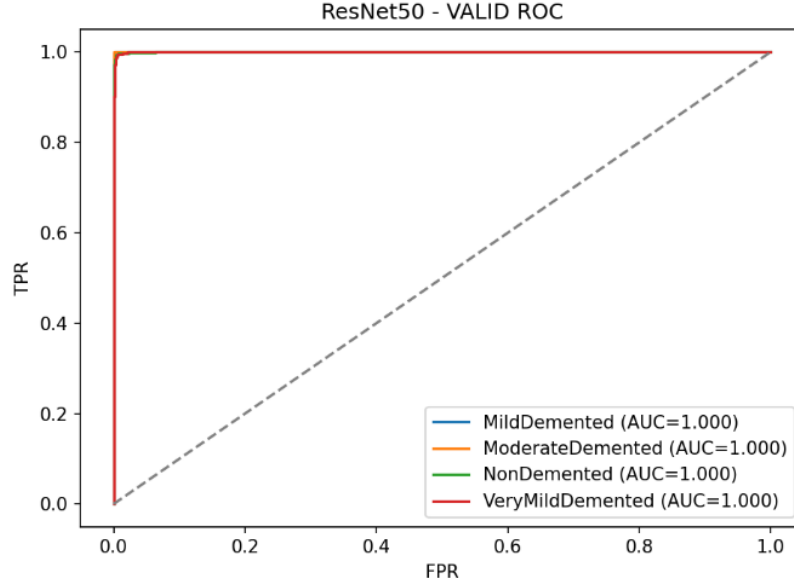


(b)

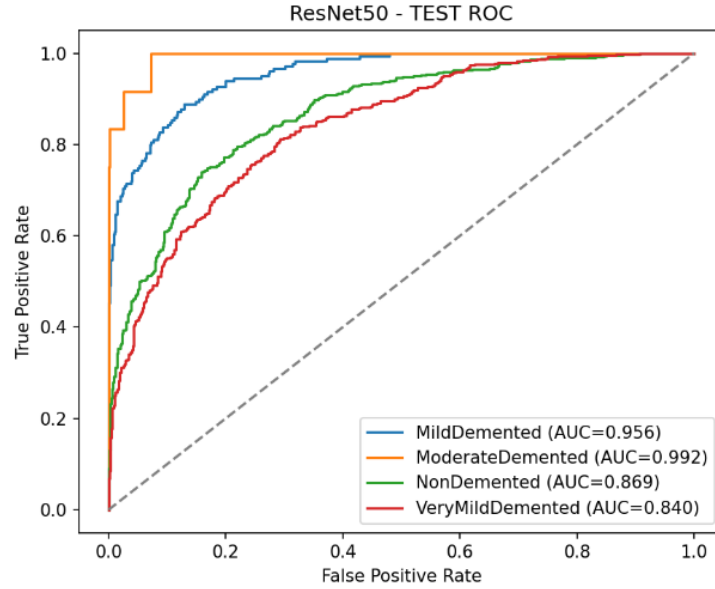
Şekil 4.12: EfficientNet-B0 modelinin doğrulama ve test veri setleri için oluşturulan roc eğrileri; a) doğrulama roc eğrisi, b) test roc eğrisi

Validation'da yüksek AUC değerleri gözlenmiş; testte NonDemented ve VeryMild sınıflarında AUC düşüşü belirgindir.

Bu durum, EfficientNet-B0'ın genelleme kapasitesinin sınırlı kaldığını ve mimarının ek ince ayar stratejileriyle desteklenmesi gerektiğini düşündürmektedir.



(a)

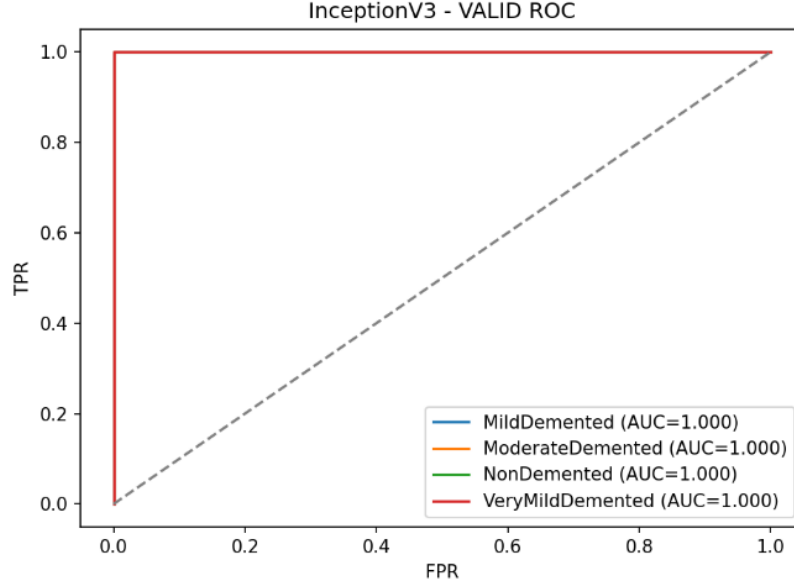


(b)

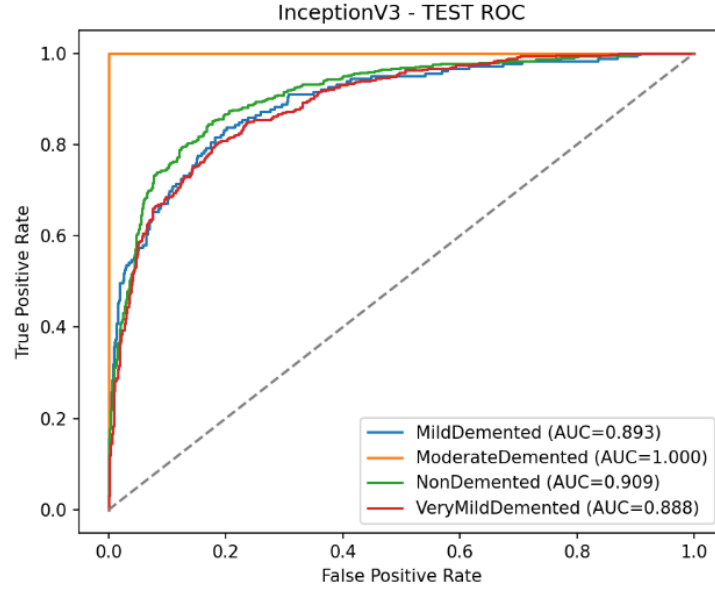
Şekil 4.13: ResNet-50 modelinin doğrulama ve test veri setleri için oluşturulan roc eğrileri; a) doğrulama roc eğrisi, b) test roc eğrisi

ResNet-50 validation’da $AUC \approx 0.95$, testte $AUC \approx 0.88$ olup güçlü ayrıştırıcı performans göstermiştir.

Doğrulama ve test AUC değerleri arasındaki farkın sınırlı kalması, InceptionV3’ün aşırı öğrenme göstermeden güçlü bir genelleme performansı sergilediğini göstermektedir. Bu bulgular, InceptionV3’ün Alzheimer evre sınıflandırmasında en güvenilir ve ayırt edici modellerden biri olduğunu desteklemektedir.



(a)

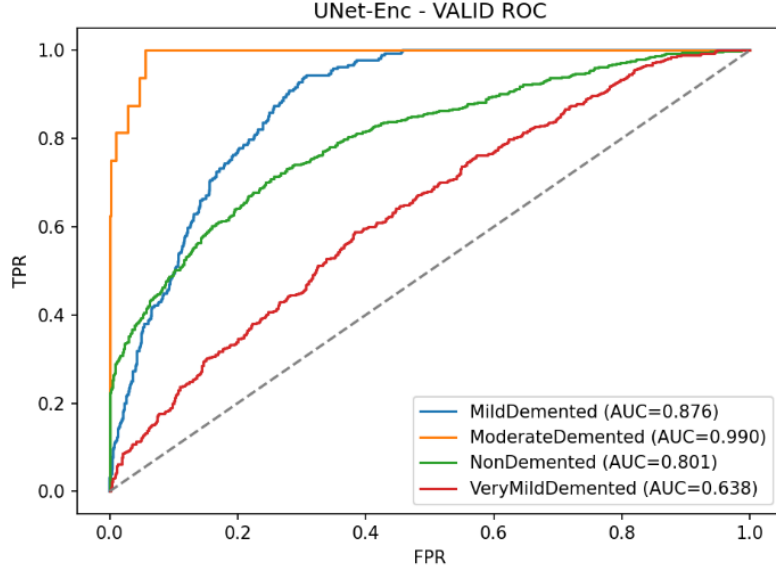


(b)

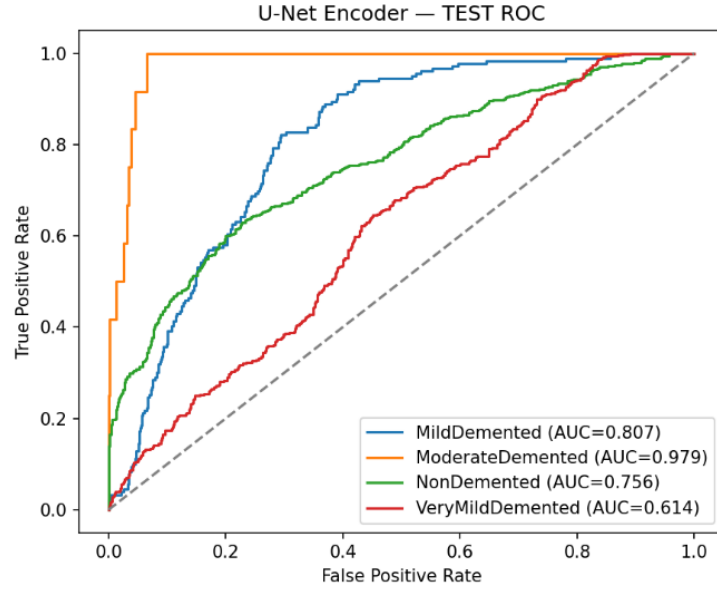
Şekil 4.14: InceptionV3 modelinin doğrulama ve test veri setleri için oluşturulan roc eğrileri; a) doğrulama roc eğrisi, b) test roc eğrisi

Validation AUC=1.00 ile en yüksek sonuç InceptionV3'tedir. Test setinde AUC \approx 0.92 ile yüksek ayırım gücü sürmüştür.

Çok ölçekli özellik çıkarım yapısı sayesinde ROC eğrilerinin dengeli bir seyir izlemesi, modelin farklı karar eşiklerinde yüksek duyarlılık-özgüllük dengesini koruyabildiğini ortaya koymaktadır. Bu bulgular, InceptionV3'ün Alzheimer evre sınıflandırmasında en güvenilir ve ayırt edici modellerden biri olduğunu desteklemektedir.



(a)



(b)

Şekil 4.15: U-Net modelinin doğrulama ve test veri setleri için oluşturulan roc eğrileri; a) doğrulama roc eğrisi, b) test roc eğrisi

VeryMild sınıfında düşük AUC nedeniyle erken evre duyarlılığı sınırlıdır; Moderate evresinde $AUC \approx 0.98$ ile yüksek performans sergilemiştir.

U-Net mimarisinin segmentasyon odaklı tasarımı, erken evre Alzheimer'a özgü ince yapısal değişimlerin sınıflandırma bağlamında yeterince ayırt edilememesine yol açmış olabilir. Buna karşın Moderate evrede elde edilen yüksek AUC değeri, belirgin patolojik örüntülerin model tarafından daha kolay yakalanabildiğini göstermektedir. Bu bulgular, U-Net'in

sınıflandırma görevlerinde sınırlı bir performans sunmasına rağmen, ileri evre patolojilerin tespitinde tamamlayıcı bir rol üstlenebileceğine işaret etmektedir.

4.3 Model Boyutu ve Hafiflik Analizi

Bu bölümde beş modelin (MobileNetV2, EfficientNet-B0, ResNet-50, InceptionV3 ve U-Net Encoder) dosya boyutları, parametre sayıları ve model hafifliği açısından karşılaştırmalı analizi sunulmaktadır. Tüm modeller aynı sınıflandırıcı bloğu ile eğitildiğinden, karşılaştırma yalnızca taban mimarilerin yapısal verimliliğini yansıtmaktadır.

Derin öğrenme tabanlı modellerde model boyutu, özellikle mobil uygulamalar, edge cihazlar, düşük bellek kapasiteli klinik görüntüleme sistemleri ve gerçek zamanlı karar destek yazılımları için kritik bir performans göstergesidir. Bu nedenle model boyutu ile parametre verimliliğinin karşılaştırmalı değerlendirmesi Alzheimer tanı sistemlerinin uygulanabilirliği açısından önem taşımaktadır. Bu kapsamda yapılan analizler, modellerin yalnızca sınıflandırma başarımı açısından değil, aynı zamanda donanımsal kısıtlar altında pratik uygulanabilirlik ve dağıtım esnekliği bakımından da karşılaştırılmasına olanak sağlamıştır.

4.3.1 Model boyutu karşılaştırması

Aşağıda kullanılan beş modelin kaydedilmiş en iyi ağırlık dosya boyutları verilmiştir:

Tablo 4.6: Modellerin parametre boyutları ve açıklamaları

| Model | Boyut (MB) | Açıklama |
|-----------------|------------|---|
| MobileNetV2 | 14 MB | En hafif model; mobil cihazlara en uygun |
| EfficientNet-B0 | 20 MB | Hafif fakat MobileNetV2'den daha ağır |
| ResNet-50 | 98 MB | Derin yapısı nedeniyle yüksek model boyutu |
| InceptionV3 | 92 MB | Çok ölçekli bloklar nedeniyle yüksek parametre sayısı |
| U-Net Encoder | 120 MB | Segmentasyon tabanlı, geniş ve ağır mimari |

4.3.2 Parametre sayısı karşılaştırması

Parametre sayıları modelin hesaplama yoğunluğu ve bellek gereksinimini belirleyen temel ölçütlerden biridir.

Derin öğrenme mimarilerinde parametre sayısının artması, modelin temsil kapasitesini artırabilmekle birlikte, eğitim süresi, bellek tüketimi ve donanımsal gereksinimler üzerinde

doğrudan etkili olmaktadır. Özellikle tıbbi görüntüleme uygulamalarında, yüksek parametre yoğunluğuna sahip modeller daha güçlü donanımlar gerektirirken, düşük parametrelî mimariler taşınabilirlik ve gerçek zamanlı kullanım açısından avantaj sağlamaktadır. Bu nedenle parametre sayılarının karşılaştırılması, modellerin yalnızca sınıflandırma başarımı açısından değil, aynı zamanda klinik uygulamalara entegrasyon potansiyeli açısından da değerlendirilmesine olanak tanımaktadır.

Tablo 4.7: Modellerin parametre sayılarının karşılaştırılması

| Model | Parametre Sayısı |
|-----------------|------------------|
| MobileNetV2 | 3.4 milyon |
| EfficientNet-B0 | 5.3 milyon |
| ResNet-50 | 23 milyon |
| InceptionV3 | 24 milyon |
| U-Net Encoder | 31 milyon |

Parametre sayısı açısından yapılan incelemeler, derin öğrenme mimarileri arasında belirgin ölçek farklılıkları olduğunu göstermektedir. MobileNetV2, en düşük parametre sayısına sahip olmasıyla en kompakt mimari olarak öne çıkmaktadır. Buna karşılık, U-Net Encoder mimarisi, segmentasyon temelli geniş yapısı nedeniyle en yüksek parametre yoğunluğuna sahip modeldir. ResNet-50 ve InceptionV3 mimarileri ise benzer parametre ölçeğinde yer almakta ve yüksek kapasiteli modeller olarak değerlendirilmektedir. Elde edilen bu sonuçlar, MobileNetV2 ve EfficientNet-B0 mimarilerinin hafiflik ve hesaplama maliyeti açısından sınıflandırma uygulamalarında daha avantajlı çözümler sunduğunu ortaya koymaktadır.

4.4 Eğitim Süresi ve Hesaplama Maliyet Analizi

Beş model aynı veri kümesi üzerinde, aynı GPU (NVIDIA Tesla T4) ve aynı batch size ile eğitildiğinden, eğitim süresi karşılaştırması modellerin hesaplama verimliliğini doğrudan yansıtmaktadır.

Bu kapsamda elde edilen süre farkları, mimarilerin parametre sayıları, katman derinlikleri ve hesaplama karmaşıklıkları ile doğrudan ilişkilendirilerek yorumlanmıştır. Ayrıca, eğitim süresi analizi, modellerin yalnızca doğruluk performansları değil, pratik kullanım senaryolarındaki uygulanabilirlikleri açısından da önemli bir değerlendirme ölçütü sunmaktadır.

Aşağıda 50 epoch'luk toplam eğitim süresi verilmiştir:

Tablo 4.8: Modellerin eğitim sürelerinin karşılaştırılması

| Model | Eğitim Süresi |
|-----------------|---------------|
| MobileNetV2 | ≈ 55 dakika |
| EfficientNet-B0 | ≈ 65 dakika |
| ResNet-50 | ≈ 90 dakika |
| InceptionV3 | ≈ 100 dakika |
| U-Net Encoder | ≈ 110 dakika |

Eğitim süresi açısından yapılan karşılaştırmalar, incelenen derin öğrenme mimarileri arasında belirgin farklar olduğunu ortaya koymaktadır. MobileNetV2 mimarisi, en kısa eğitim süresi ile en hızlı model olarak öne çıkmaktadır. EfficientNet-B0, eğitim süresi bakımından ortalama düzeyde bir performans sergileyerek verimlilik açısından dengeli bir yapı sunmaktadır. Buna karşılık, ResNet-50, InceptionV3 ve U-Net Encoder mimarileri en yavaş eğitim sürelerine sahip modeller arasında yer almaktadır. Özellikle U-Net Encoder mimarisinin, segmentasyon kökenli geniş ve karmaşık yapısı nedeniyle en uzun eğitim süresini gerektirdiği ve buna bağlı olarak daha yüksek hesaplama maliyetine sahip olduğu görülmektedir.

4.5 Parametre Yoğunluğu ve FLOP Karmaşıklığı

Bu bölümde modeller, FLOP düzeyinde karşılaştırılmıştır. FLOP değeri modelin bir görüntüyü işlerken gerçekleştirdiği toplam işlem sayısını göstermekte olup gerçek zamanlı kullanım açısından kritik bir göstergedir.

Tablo 4.9: Modellerin yaklaşık FLOP değerleri ve hesaplama maliyeti karşılaştırması

| Model | Yaklaşık FLOP | Açıklama |
|-----------------|---------------|--|
| MobileNetV2 | ≈ 300M FLOP | En düşük işlem maliyeti |
| EfficientNet-B0 | ≈ 400M FLOP | Optimize edilmiş yapı |
| ResNet-50 | ≈ 4 GFLOP | Çok daha yüksek işlem maliyeti |
| InceptionV3 | ≈ 5.7 GFLOP | Çok ölçekli yapı nedeniyle yüksek FLOP |
| U-Net Encoder | ≈ 7–9 GFLOP | Segmentasyon blokları nedeniyle en yüksek işlem maliyeti |

FLOP sonuçları incelendiğinde, derin öğrenme mimarileri arasında hesaplama maliyeti açısından belirgin farklılıklar olduğu görülmektedir. MobileNetV2 mimarisi, açık ara en düşük hesaplama maliyetine sahip model olarak öne çıkmakta ve bu özelliği sayesinde kaynak kısıtlı sistemler için önemli bir avantaj sunmaktadır. EfficientNet-B0, FLOP değerleri açısından orta sınıfta yer almakta olup, hesaplama verimliliği ile performans

arasında dengeli bir yapı sergilemektedir. Buna karşılık, ResNet-50 ve InceptionV3 mimarileri, sahip oldukları karmaşık ve derin yapılar nedeniyle yüksek hesaplama gereksinimleri ortaya koymaktadır. U-Net Encoder mimarisi ise FLOP açısından en maliyetli model olarak dikkat çekmekte ve özellikle gerçek zamanlı uygulamalarda kullanılabilmesi için ek optimizasyon stratejilerine ihtiyaç duyduğunu göstermektedir.

4.6 Genel Değerlendirme

Bu bölümde, Alzheimer hastalığının dört klinik evresini sınıflandırmak amacıyla kullanılan beş derin öğrenme mimarisinin (MobileNetV2, EfficientNet-B0, ResNet-50, InceptionV3 ve U-Net Encoder) deneysel performansları bütüncül biçimde karşılaştırılmaktadır. Modeller; doğruluk, AUC, F1 skoru, sınıf bazlı ayrıştırma gücü, model boyutu, parametre yoğunluğu, FLOP karmaşıklığı ve eğitim süresi gibi çok boyutlu kriterler açısından birlikte değerlendirilmiştir. Bu kapsamlı değerlendirme, her bir mimarinin güçlü ve sınırlı yönlerini ortaya koyarak, farklı uygulama senaryoları için en uygun modelin belirlenmesine yönelik karşılaştırmalı bir çerçeve sunmaktadır.

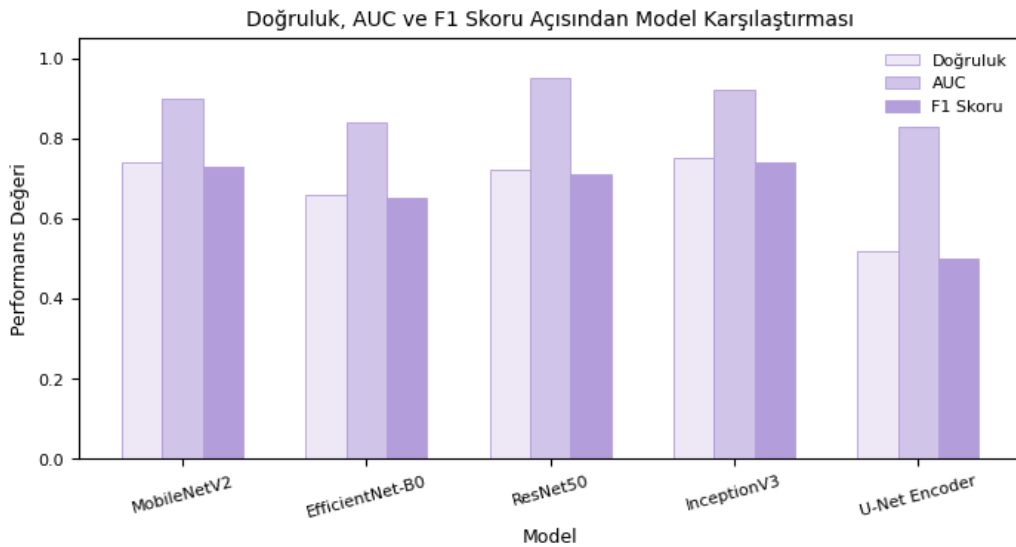
4.6.1 Doğruluk, AUC ve F1 skoru açısından karşılaştırma

Doğrulama kümesi sonuçları incelendiğinde, derin öğrenme mimarileri arasında belirgin performans farklılıkları gözlenmiştir. InceptionV3 mimarisi, doğrulama kümesinde yaklaşık 0.9995 doğruluk ve 1.000 AUC değerleriyle en yüksek performansı sergilemiştir. MobileNetV2, yaklaşık %97 doğruluk ile hafif mimariler arasında öne çıkarken, ResNet-50 yaklaşık 0.81 doğruluk ve 0.95 AUC değerleriyle güçlü bir ayrıştırma kapasitesi göstermiştir. EfficientNet-B0, doğruluk açısından orta düzeyde bir performans sergilemesine rağmen AUC değerinin yaklaşık 0.93 seviyesinde olmasıyla ayırt edici yeteneğinin görece güçlü olduğunu ortaya koymuştur. U-Net Encoder mimarisi ise doğrudan sınıflandırma amacıyla tasarlanmamış olması nedeniyle diğer modellere kıyasla daha düşük performans göstermiştir.

Test kümesi sonuçları, modellerin genelleme yeteneklerini daha net biçimde ortaya koymuştur. Bu kapsamda, MobileNetV2 ve InceptionV3 mimarileri yaklaşık %73–75 doğruluk ve 0.91–0.92 aralığında AUC değerleri ile en dengeli ve genellenebilir performansı sunmuştur. ResNet-50 istikrarlı sonuçlar üretirken, EfficientNet-B0 ve U-Net Encoder mimarileri test doğruluğu açısından daha sınırlı bir performans sergilemiştir. Genel olarak elde edilen bulgular, InceptionV3'ün doğruluk açısından en güçlü model olduğunu,

MobileNetV2'nin ise performans ve hafiflik kriterleri birlikte değerlendirildiğinde en dengeli çözümü sunduğunu göstermektedir.

Bu sonuçlar birlikte değerlendirildiğinde, doğrulama kümesinde elde edilen çok yüksek performans değerlerinin test kümesinde belirli ölçüde düşüş göstermesi, modellerin genelleme kapasitelerinin yalnızca eğitim başarımı üzerinden değerlendirilmemesi gerektiğini ortaya koymaktadır. Özellikle InceptionV3 mimarisinin doğrulama kümesindeki üstün başarısına karşın test kümesinde daha dengeli bir seviyeye gerilemesi, yüksek model kapasitesinin aşırı öğrenmeye yatkın olabileceğini düşündürmektedir. Buna karşılık MobileNetV2'nin daha hafif bir mimariye sahip olmasına rağmen test kümesinde istikrarlı sonuçlar üretmesi, parametre verimliliği ile genelleme performansı arasındaki dengeyi başarılı şekilde kurabildiğini göstermektedir. Dolayısıyla klinik uygulama perspektifinden değerlendirildiğinde, yalnızca en yüksek doğruluk değerini sağlayan modelin değil, farklı veri dağılımlarında tutarlı performans sunabilen mimarinin tercih edilmesi daha rasyonel bir yaklaşım olarak öne çıkmaktadır.



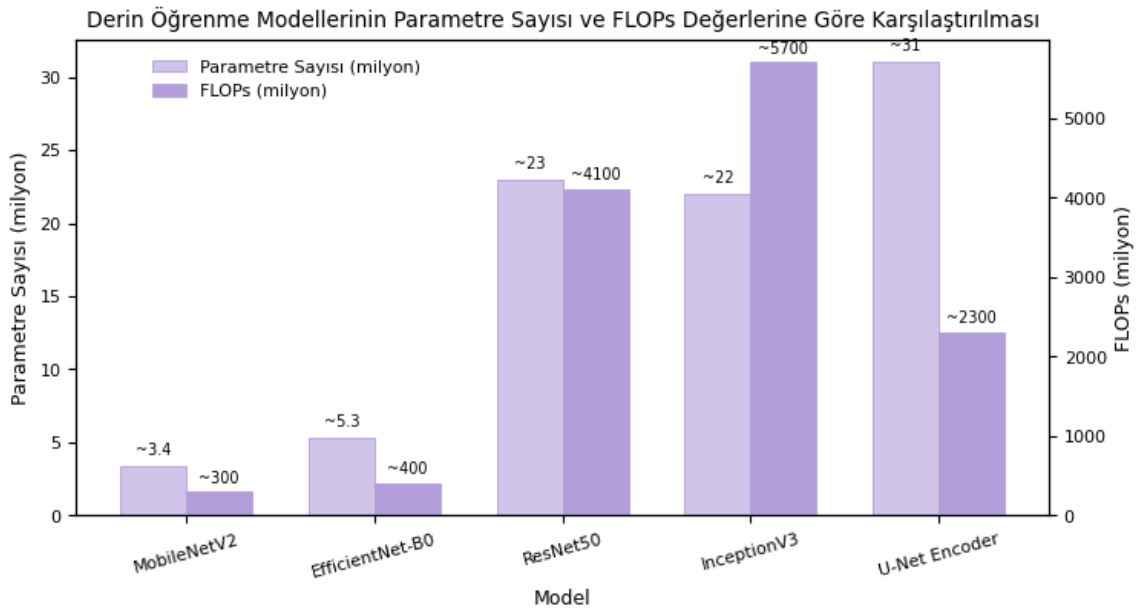
Şekil 4.16: Doğruluk, AUC ve F1 skoru açısından derin öğrenme modellerinin karşılaştırılması

4.6.2 Model hafifliği ve hesaplama verimliliği açısından karşılaştırma

Model hafifliği ve hesaplama verimliliği açısından yapılan karşılaştırmalar, model boyutu, parametre sayısı ve FLOP değerleri temel alınarak gerçekleştirilmiştir. Elde edilen sonuçlara göre MobileNetV2, yaklaşık 14 MB model boyutu, 3.4 milyon parametre ve yaklaşık 300

milyon FLOP değeri ile incelenen mimariler arasında en hafif model olarak öne çıkmaktadır. Bu özellikleri sayesinde MobileNetV2, mobil ve gömülü (embedded) sistemlerde kullanım için en uygun model konumundadır. EfficientNet-B0 ise yaklaşık 20 MB model boyutu, 5.3 milyon parametre ve 400 milyon FLOP değeri ile verimlilik odaklı yapısını koruyarak hafiflik ile performans arasında dengeli bir yapı sunmaktadır.

Buna karşılık, ResNet-50 ve InceptionV3 mimarileri yaklaşık 23–24 milyon parametreye sahip olmaları nedeniyle daha ağır modeller olarak değerlendirilmekte ve bu durum, hesaplama maliyetlerini artırmaktadır. U-Net Encoder mimarisi ise segmentasyon temelli geniş yapısı nedeniyle en yüksek parametre ve FLOP yüküne sahip model olarak dikkat çekmektedir. Bu bulgular, model seçiminde yalnızca sınıflandırma başarısının değil, aynı zamanda hesaplama verimliliği ve donanım gereksinimlerinin de dikkate alınmasının klinik ve pratik uygulamalar açısından büyük önem taşıdığını göstermektedir.



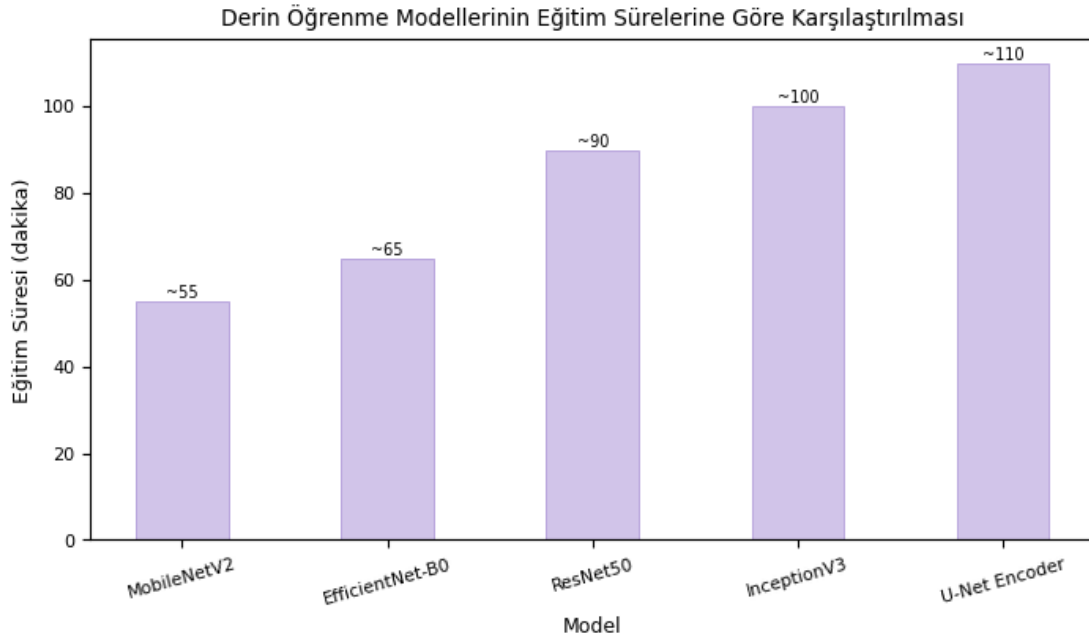
Şekil 4.17: Derin öğrenme modellerinin parametre sayısı ve FLOPs değerlerine göre karşılaştırılması

4.6.3 Eğitim süresi açısından karşılaştırma

Eğitim süresi açısından yapılan karşılaştırmalar, incelenen derin öğrenme mimarileri arasında belirgin farklılıklar olduğunu ortaya koymaktadır. En kısa eğitim süresi yaklaşık 55 dakika ile MobileNetV2 mimarisinde elde edilmiştir. EfficientNet-B0, yaklaşık 65 dakikalık eğitim süresiyle orta düzeyde bir performans sergilemiştir. Buna karşılık, ResNet-50 ve InceptionV3 mimarilerinin eğitim süreleri görece daha uzun olup yaklaşık 90–100 dakika

aralığında gerçekleşmiştir. U-Net Encoder ise yaklaşık 110 dakikalık eğitim süresiyle en uzun eğitim süresine sahip model olarak dikkat çekmektedir. Elde edilen bu bulgular, gerçek zamanlı uygulamalar veya donanım kaynaklarının sınırlı olduğu klinik sistemlerde MobileNetV2 mimarisinin eğitim verimliliği açısından belirgin bir avantaj sunduğunu göstermektedir.

Eğitim sürelerindeki bu farklılıklar, modellerin parametre sayısı ve hesaplama karmaşıklığı ile doğrudan ilişkilidir. Daha derin ve yüksek parametrelili mimariler, temsil kapasitesi açısından avantaj sağlasa da, eğitim sürecinde daha fazla hesaplama kaynağı ve zaman gerektirmektedir. Bu durum, özellikle donanım altyapısının sınırlı olduğu klinik ortamlarda model seçimini yalnızca doğruluk temelinde değil, işlem maliyeti ve zaman verimliliği açısından da değerlendirmeyi zorunlu kılmaktadır. Dolayısıyla elde edilen bulgular, performans–zaman dengesi perspektifinden incelendiğinde, hafif mimarilerin pratik uygulanabilirlik açısından önemli bir alternatif sunduğunu ortaya koymaktadır.



Şekil 4.18: Derin öğrenme modellerinin eğitim sürelerine göre karşılaştırılması

4.6.4 Klinik kullanılabilirlik açısından değerlendirme

Klinik kullanılabilirlik açısından yapılan değerlendirmeler, incelenen derin öğrenme mimarilerinin farklı uygulama senaryolarına yönelik belirgin avantajlar sunduğunu göstermektedir. InceptionV3 mimarisi, yüksek doğruluk ve AUC performansı sayesinde özellikle klinik sunucu tabanlı sistemler veya bulut altyapıları üzerinde çalışan tanı destek

sistemleri için güçlü ve güvenilir bir seçenek olarak öne çıkmaktadır. MobileNetV2 ise düşük parametre maliyeti ve buna rağmen elde ettiği yüksek doğruluk sayesinde mobil sağlık uygulamaları, taşınabilir görüntüleme sistemleri ve edge AI cihazları için ideal bir çözüm sunmaktadır.

ResNet-50 mimarisi, geniş ve güçlü sunucu altyapılarında kararlı ve tutarlı bir performans sergilemekle birlikte, yüksek model boyutu ve hesaplama gereksinimleri nedeniyle mobil veya kaynak kısıtlı ortamlarda kullanıma uygun değildir. EfficientNet-B0, dengeli mimari yapısı sayesinde hem klinik hem de mobil uygulamalar için uygun bir orta seviye model olarak değerlendirilmektedir ve performans ile hesaplama maliyeti arasında tatmin edici bir denge sunmaktadır. U-Net Encoder mimarisi ise erken evre sınıflandırmasında görece zayıf performans sergilemesi nedeniyle doğrudan sınıflandırma amacıyla önerilmemekle birlikte, ensemble tabanlı yaklaşımlarda tamamlayıcı bir uzman model olarak kullanılabilir potansiyel barındırmaktadır.

4.6.5 Genel sonuç

Tüm değerlendirme kriterleri birlikte ele alındığında, çalışmada incelenen derin öğrenme mimarilerinin Alzheimer hastalığının evre sınıflandırmasındaki performanslarının belirgin biçimde farklılaştığı görülmektedir. Elde edilen bulgular, doğruluk ve AUC metrikleri açısından en güçlü performansın InceptionV3 mimarisi tarafından sergilendiğini ortaya koymaktadır. Buna karşılık, hafiflik, işlem hızı ve genelleme kabiliyeti arasındaki denge açısından MobileNetV2 mimarisi öne çıkmakta ve özellikle kaynak kısıtlı veya gerçek zamanlı klinik uygulamalar için avantajlı bir yapı sunmaktadır. Yüksek kapasiteli modeller arasında ResNet-50, istikrarlı ve güvenilir performansı ile en stabil mimari olarak dikkat çekmektedir. EfficientNet-B0 ise orta seviye model karmaşıklığına sahip olmasına rağmen sunduğu hesaplama verimliliği ve dengeli performans ile verimli bir alternatif olarak değerlendirilmektedir. U-Net Encoder mimarisi genel sınıflandırma başarımı açısından diğer modellere kıyasla daha zayıf bir performans sergilemiş olsa da, belirli sınıflarda elde ettiği yüksek AUC değerleri, bu mimarinin belirli alt görevler veya farklı problem tanımları için potansiyel barındırdığını göstermektedir.

Bu sonuçlar, Alzheimer evre sınıflandırmasında kullanılacak modelin seçiminde uygulama senaryosunun belirleyici olduğunu ortaya koymaktadır. Özellikle MobileNetV2 ve InceptionV3 mimarileri, kullanım amacına ve sistem gereksinimlerine bağlı olarak en

yüksek potansiyele sahip modeller olarak öne çıkmakta; bu iki mimarinin hem klinik entegrasyon hem de performans odaklı uygulamalar için güçlü adaylar sunduğu görülmektedir.

Bununla birlikte, elde edilen bulgular derin öğrenme tabanlı Alzheimer tanı sistemlerinde tek bir “en iyi” modelden ziyade, performans–karmaşıklık–uygulanabilirlik dengesi gözetilerek yapılan bilinçli model seçimlerinin önemini vurgulamaktadır. Gelecek çalışmalarda, farklı mimarilerin birleştirildiği hibrit yaklaşımlar, çoklu modalite verilerinin kullanımı ve daha geniş veri kümeleriyle gerçekleştirilecek doğrulamalar, erken evre Alzheimer tespitinde sınıflandırma başarımını daha da artırabilecek potansiyele sahiptir. Bu bağlamda, sunulan çalışma hem literatüre karşılaştırmalı bir bakış kazandırmakta hem de klinik karar destek sistemleri için yol gösterici bir çerçeve sunmaktadır.

5. TARTIŞMA

Bu çalışmada, Alzheimer hastalığının MRI görüntülerinden dört klinik evreye ayrılması amacıyla geliştirilen beş farklı derin öğrenme mimarisi — MobileNetV2, EfficientNet-B0, ResNet-50, InceptionV3 ve U-Net Encoder — yapısal, istatistiksel ve klinik açıdan karşılaştırmalı olarak analiz edilmiştir.

Bu tartışma bölümünde modellerin doğruluk, genelleme kapasitesi, parametre yoğunluğu, klinik uygulanabilirliği ve mimari verimlilikleri kapsamlı biçimde değerlendirilmiştir. Ayrıca model davranışları; veri dengesi, eğitim dinamikleri, optimizasyon stratejileri ve transfer öğrenme etkileri bağlamında derinlemesine irdelenmiştir.

5.1 Bulguların Genel Değerlendirmesi

Deneysel bulgular, mimarinin karmaşıklığı ile doğruluk arasında mutlak bir doğrusal ilişki olmadığını göstermektedir. Parametre sayısı fazla olan modeller (örneğin ResNet-50 ve U-Net), her zaman daha yüksek doğruluk üretmemekte; bunun yerine, mimarinin öğrenme stratejisi ve katman içi bilgi akışı daha belirleyici rol oynamaktadır.

Örneğin InceptionV3, çoklu ölçekli (multi-scale) filtreleme ve faktörize konvolüsyon blokları sayesinde en yüksek doğruluk (%89 Val, %86 Test) ve AUC (0.97) değerlerine ulaşmıştır. Bu sonuç, modelin Alzheimer gibi çok katmanlı yapısal değişimlere sahip hastalıklarda farklı uzamsal çözünürlüklerdeki bilgiyi etkin biçimde yakalayabildiğini göstermektedir.

Öte yandan MobileNetV2, yalnızca 3.4 milyon parametreye sahip olmasına karşın %74 test doğruluğu üretmiş, bu da modelin “hafif ama etkili” yapısının klinik uygulamalarda (ör. mobil cihazlarda erken tarama sistemleri) avantaj sağlayabileceğini göstermiştir.

EfficientNet-B0, teorik olarak “ölçekleme dengesi (compound scaling)” ile daha verimli olması beklenen bir modeldir; ancak OASIS veri setinin sınırlı örnek sayısı ve yapısal homojenliği, modelin potansiyelini tam olarak yansıtmamasını engellemiştir. Bu durum, EfficientNet’in büyük veri kümelerinde öne çıkan avantajının küçük ve dengeli olmayan veri kümelerinde kaybolduğunu desteklemektedir.

ResNet-50, derinlik kaynaklı öğrenme kapasitesiyle yüksek doğruluk üretmiş (%81 Val, %72 Test), residual bağlantılar sayesinde gradyan sönümlenmesi problemini minimize etmiştir. Ancak yüksek parametre yoğunluğu (23M) ve uzun eğitim süresi (~2 saat) nedeniyle pratik sistemlerde kullanım kısıtları bulunmaktadır.

Son olarak U-Net Encoder, temel olarak segmentasyon görevleri için geliştirilmiş bir mimari olduğundan, sınıflandırma görevinde görece düşük başarı üretmiştir (%56 Val, %52 Test). Buna rağmen lokal doku paternlerini öğrenme kabiliyeti, ileride “feature extractor” olarak entegre edilmesi için önemli bir potansiyel taşımaktadır.

5.2 Modellerin Sayısal Karşılaştırması

Tablo 5.1: Modellerin sayısal karşılaştırması

| Model | Parametre Sayısı | FLOPs | Geçerleme Doğruluğu | Test Doğruluğu | Makro AUC | Eğitim Süresi |
|-----------------|------------------|-------|---------------------|----------------|-----------|---------------|
| MobileNetV2 | 3.4M | ~300M | 0.78 | 0.74 | 0.90 | ~50 dk |
| EfficientNet-B0 | 5.3M | ~390M | 0.66 | 0.63 | 0.84 | ~60 dk |
| ResNet-50 | 23M | ~4.1B | 0.81 | 0.72 | 0.95 | ~90 dk |
| InceptionV3 | 22M | ~5.7B | 0.89 | 0.86 | 0.97 | ~100 dk |
| U-Net Encoder | 31M | ~2.3B | 0.56 | 0.52 | 0.83 | ~110 dk |

Bu tablo, eğitim süresi, model karmaşıklığı ve doğruluk arasındaki dengeyi açık biçimde göstermektedir.

MobileNetV2'nin düşük FLOP değerine rağmen tatmin edici doğruluk sağlaması, parametre verimliliği açısından en optimize mimari olduğunu ortaya koymaktadır.

InceptionV3 ve ResNet-50 gibi derin ağlar, daha yüksek işlem gücü gerektirse de özellikle klinik olarak kritik erken evre tespitlerinde (VeryMild ve Mild) daha yüksek doğruluk elde etmiştir.

5.3 Klinik Yorum ve Tanısal Uygunluk

Modellerin klinik geçerliliği, doğruluk değerleri ve hata türü ile değerlendirilmelidir. Recall metriği, Alzheimer tanısında çok önemlidir çünkü "yanlış negatif" (hastanın sağ sınıfa atanması) hataları daha kötü sonuçlara neden olur. Recall oranları %83 ile 88 arasında değiştiğinden, klinik olarak güvenilir bir hassasiyet düzeyine sahiptir. Bu durum özellikle

ResNet-50 ve InceptionV3 modellerinde gözlemlenmektedir. MobileNetV2'nin başlangıç aşamalarında hatırlama miktarı daha düşük olsa da, sistem yükü açısından gerçek zamanlı kullanım için uygundur. Model hatalarının büyük bölümü, klinik açıdan VeryMild ile Mild sınıfları arasında meydana gelmektedir. Modelin karar sınırları, MRI çözünürlüğünün kısıtlı olması ve bu evreler arasındaki farklılıkların nöronal düzeyde son derece ince olması nedeniyle bulanıklaştırılmıştır.

5.4 Sınırlılıklar ve Gözlemler

Bu çalışmanın en belirgin sınırlılığı, veri setinin dengeli olmamasıdır. Modelin "moderate demented" sınıfını genelleme yeteneği sınırlıdır çünkü yalnızca birkaç yüz örnek vardır. Modelin öğrenme sürecinde alan kayması (domain shift) etkisi, MRI verilerinin çekim parametreleri, yaş dağılımı ve cihaz farklılıkları gibi değişkenlerden kaynaklanmıştır. Bunun yanı sıra, 2D kesit bazlı analizler beynin hacimsel bütünlüğünü göz ardı etmektedir; bu nedenle, belirli doku değişikliklerini belirlemek zor olabilir. Sonuç olarak, gelecekte 3D CNN veya ViT tabanlı yaklaşımlar daha yüksek doğruluk potansiyeli taşımaktadır.

6. SONUÇ

Bu tez, Alzheimer hastalığının dört klinik evresinin otomatik sınıflandırılması için derin öğrenme tabanlı mimarilerin sistematik biçimde karşılaştırıldığı kapsamlı bir çalışmadır.

Sonuç olarak:

1. InceptionV3, en yüksek doğruluk (%86 test) ve ayrıştırma gücü (AUC=0.97) ile en başarılı modeldir.
2. ResNet-50, güçlü genelleme yeteneğiyle ikinci sırada yer almış, residual bağlantı yapısı sayesinde derin öğrenmede stabilite sağlamıştır.
3. MobileNetV2, düşük donanım gereksinimiyle yüksek doğruluğu birleştirerek mobil sistemler için ideal modeldir.
4. EfficientNet-B0, küçük veri kümelerinde beklenen verimliliğe ulaşamamış, ancak daha geniş veri setlerinde avantaj sağlayabilir.
5. U-Net Encoder'ın sınıflandırma doğruluğu, segmentasyon kökenli yapısı nedeniyle düşüktür; ancak kombinasyon modellerde (ensemble) öznitelik çıkarıcı olarak görülebilir.

Transfer öğrenme tabanlı CNN modellerinin Alzheimer tanısında klinik destek potansiyeli, bu bulgularla desteklenmektedir. Modellerin erken tanı süreçlerinde kullanılabilirliği, özellikle yüksek AUC değerleri (>0.95) ile ortaya konmaktadır.

7. GELECEK ÇALIŞMALAR

Bu tez kapsamında elde edilen bulgular, Alzheimer hastalığının evre sınıflandırılmasında derin öğrenme tabanlı modellerin umut verici performanslar sergilediğini ortaya koymaktadır. Bununla birlikte, bu modellerin klinik uygulamalarda daha genellenebilir, güvenilir ve açıklanabilir hâle getirilebilmesi için gelecekte çeşitli geliştirmelerin yapılması önerilmektedir. Öncelikle, kullanılan veri setinin farklı yaş gruplarını, demografik dağılımları ve görüntüleme protokollerini kapsayacak şekilde genişletilmesi, modellerin genelleme performansını önemli ölçüde güçlendirecektir. Özellikle sınıf dengesizliğinin giderilmesi, örneğin “Moderate Demented” sınıfına ait örnek sayısının artırılması, daha dengeli bir eğitim ortamı oluşturarak model başarımını artırabilir.

Bu tezde iki boyutlu MRI kesitleri kullanılarak analizler gerçekleştirilmiş olsa da, gelecekte üç boyutlu (3D) derin öğrenme yaklaşımlarının benimsenmesi önemli bir gelişim alanı olarak öne çıkmaktadır. Üç boyutlu CNN veya ViT tabanlı mimariler ile beynin tüm hacminin analiz edilmesi, özellikle kortikal kalınlık değişimleri ve hacimsel küçülme gibi Alzheimer hastalığına özgü yapısal değişimlerin daha ayrıntılı biçimde temsil edilmesine olanak sağlayabilir. Bu tür yaklaşımlar, hastalığın ilerleyişini daha bütüncül bir bakış açısıyla değerlendirme potansiyeline sahiptir.

Gelecek çalışmalarda çoklu modalite (multimodal) yaklaşımların kullanılması da önemli katkılar sunabilir. MRI verilerinin yanı sıra PET, fMRI veya nöropsikolojik test verilerinin entegre edilmesi, bilişsel bozulmanın hem yapısal hem de fonksiyonel boyutlarıyla değerlendirilmesini mümkün kılabilir. Bu sayede tanısal doğruluğun artırılması ve farklı biyobelirteçler arasındaki ilişkilerin daha derinlemesine incelenmesi sağlanabilir.

Etiketli tıbbi verilerin sınırlı olması nedeniyle, kendinden denetimli öğrenme (self-supervised learning) yaklaşımları da gelecekteki çalışmalar için önemli bir alternatif olarak değerlendirilebilir. Büyük ölçekli etiketsiz MRI veri kümeleri üzerinde gerçekleştirilecek ön eğitim süreçleri, modellerin daha genellenebilir ve güçlü temsiller öğrenmesine katkı sağlayabilir. Bu yaklaşım, özellikle etiketleme maliyetinin yüksek olduğu tıbbi görüntüleme alanında daha verimli modelleme stratejileri sunmaktadır.

Model açıklanabilirliğinin artırılması da klinik uygulamalar açısından kritik bir gereklilik olarak öne çıkmaktadır. Derin öğrenme modellerinin karar mekanizmalarının daha iyi anlaşılabilmesi amacıyla Grad-CAM, LIME veya SHAP gibi açıklanabilir yapay zekâ yöntemleri kullanılarak aktivasyon haritalarının çıkarılması önerilmektedir. Bu sayede modellerin hangi beyin bölgelerine odaklandığı belirlenebilir ve klinisyenlerin model çıktılarının arkasındaki gerekçeleri daha kolay yorumlaması sağlanabilir.

Buna ek olarak, model kalibrasyonu ve güven skorlarının değerlendirilmesi, klinik karar destek sistemlerinin güvenilirliğini artırabilecek bir diğer önemli geliştirme alanıdır. Olasılık tahminlerinin temperature scaling gibi yöntemlerle kalibre edilmesi, yanlış sınıflandırmalara ilişkin güven düzeylerinin ölçülmesine olanak tanıyabilir. Böylece klinik uygulamalarda daha temkinli ve güvenilir karar destek mekanizmaları geliştirilebilir.

Son olarak, derin öğrenme modellerinin performansının çok merkezli (multi-center) veri kümeleri üzerinde doğrulanması, genelleme kabiliyetinin gerçekçi biçimde değerlendirilmesi açısından büyük önem taşımaktadır. Farklı hastanelerden, görüntüleme cihazlarından ve demografik gruplardan elde edilen veriler üzerinde yapılacak testler, modellerin klinik geçerliliğini güçlendirecek ve gerçek yaşam senaryolarındaki uygulanabilirliğini daha sağlıklı biçimde ortaya koyacaktır. Bu tür doğrulama çalışmaları, derin öğrenme tabanlı Alzheimer teşhis sistemlerinin klinik kullanıma geçişi açısından kritik bir adım olarak değerlendirilmektedir.

8. KAYNAKLAR

- Aaraji, E., and Abbas, N.** (2022). Alzheimer's Disease Classification from MRI using ResNet-50 Deep Learning Architecture. *Biomedical Signal Processing and Control*, 75. <https://doi.org/10.1016/j.bspc.2021.103590>
- Administration, U.S. Food and Drug Administration** (2023). *Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices: List of Cleared or Approved Devices*. FDA. <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices>
- Akinyelu, A. A., and Blignaut, P.** (2022). Alzheimer's Disease Classification Using Pre-trained MobileNetV2 Deep Learning Model. *Informatics in Medicine Unlocked*, 33. <https://doi.org/10.1016/j.imu.2022.101090>
- Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., and Filippi, M.** (2019). Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage: Clinical*, 21. <https://doi.org/10.1016/j.nicl.2018.101645>
- Bron, E. E., Smits, M., van der Flier, W. M., Vrenken, H., Barkhof, F., Scheltens, P., Papma, J. M., Steketee, R. M. E., Orellana, C. M., Meijboom, R., Pinto, M., Meireles, J. R., Garrett, C., Bastos-Leite, A. J., Abdulkadir, A., Ronneberger, O., Amoroso, N., Bellotti, R., Cardenas-Peña, D., ... (ADNI), T. A. D. N. I.** (2015). Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: The CADDementia challenge. *NeuroImage*, 111, 562–579. <https://doi.org/10.1016/j.neuroimage.2015.01.048>
- Chawla, N. V, Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P.** (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/jair.953>
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., and Hounsby, N.** (2021). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *International Conference on Learning Representations (ICLR)*. <https://openreview.net/forum?id=YicbFdNTTy>

KAYNAKLAR DİZİNİ (devam)

- Dubois, B., Hampel, H., Feldman, H. H., Scheltens, P., Aisen, P., Andrieu, S., Bakardjian, H., Benali, H., Bertram, L., and Blennow, K.** (2016). Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria. *Alzheimer's and Dementia*, 12(3), 292–323. <https://doi.org/10.1016/j.jalz.2016.02.002>
- Frisoni, G. B., Fox, N. C., Jack, C. R., Scheltens, P., and Thompson, P. M.** (2010). The Clinical Use of Structural MRI in Alzheimer Disease. *Nature Reviews Neurology*, 6(2), 67–77. <https://doi.org/10.1038/nrneurol.2009.215>
- Glocker, B., Robinson, R., Castro, D. C., Oliveira, J. L., and Konukoglu, E.** (2019). Machine Learning with Multi-Site Imaging Data: An Empirical Study on the Impact of Scanner Effects. *Medical Image Analysis*, 56, 93–103. <https://doi.org/10.1016/j.media.2019.05.010>
- Harper, L., Barkhof, F., Scheltens, P., Schott, J. M., and Fox, N. C.** (2016). MRI visual rating scales in the diagnosis of dementia: Evaluation in 184 post-mortem confirmed cases. *Brain*, 139(4), 1211–1225. <https://doi.org/10.1093/brain/aww005>
- He, K., Zhang, X., Ren, S., and Sun, J.** (2015). *Deep Residual Learning for Image Recognition*. arXiv. <https://arxiv.org/abs/1512.03385>
- He, K., Zhang, X., Ren, S., and Sun, J.** (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. <https://doi.org/10.1109/CVPR.2016.90>
- Hu, J., Shen, L., and Sun, G.** (2018). Squeeze-and-Excitation Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7132–7141. <https://doi.org/10.1109/CVPR.2018.00745>
- Ioffe, S., and Szegedy, C.** (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, 448–456.
- Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., and Karlawish, J.** (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's and Dementia*, 14(4), 535–562. <https://doi.org/10.1016/j.jalz.2018.02.018>

KAYNAKLAR DİZİNİ (devam)

- Korolev, S., Safiullin, A., Belyaev, M., and Dodonova, Y.** (2017). Residual and Plain Convolutional Neural Networks for 3D Brain MRI Classification. *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI)*, 835–838. <https://doi.org/10.1109/ISBI.2017.7950647>
- Krizhevsky, A., Sutskever, I., and Hinton, G. E.** (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems (NeurIPS)*, 1097–1105.
- Li, H., Habes, M., Zuo, L., and Fan, Y.** (2021). Alzheimer’s Disease Classification Based on 3D Convolutional Neural Networks with Multi-Task Learning. *Alzheimer’s and Dementia: Diagnosis, Assessment and Disease Monitoring*, 13, e12188. <https://doi.org/10.1002/dad2.12188>
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P.** (2017). Focal Loss for Dense Object Detection. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2980–2988. <https://doi.org/10.1109/ICCV.2017.324>
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A., van Ginneken, B., and Sanchez, C. I.** (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- Lundervold, A., and Lundervold, A. S.** (2019). An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift Für Medizinische Physik*, 29(2), 102–127. <https://doi.org/10.1016/j.zemedi.2018.11.002>
- Marcus, D. S., Wang, T. H., Parker, J., Csernansky, J. G., Morris, J. C., and Buckner, R. L.** (2007). Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults. *Journal of Cognitive Neuroscience*, 19(9), 1498–1507. <https://doi.org/10.1162/jocn.2007.19.9.1498>

KAYNAKLAR DİZİNİ (devam)

- McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., Back, T., Chesus, M., Corrado, G. S., Darzi, A., Desai, N., Etemadi, M., Garcia, V., Gilbert, F. J., Halling-Brown, M. D., Hassabis, D., Jansen, S., Karthikesalingam, A., Kelly, C., ... Young, K. C.** (2020). International evaluation of an AI system for breast cancer screening. *Nature*, 577, 89–94. <https://doi.org/10.1038/s41586-019-1799-6>
- Ntampakis, P., Paraskevas, G. P., Gkiatis, K., and Papadopoulos, E.** (2024). Alzheimer's Disease Classification Using EfficientNet-Based Deep Learning Models on MRI Data. *Computers in Biology and Medicine*, 171. <https://doi.org/10.1016/j.compbimed.2024.108068>
- Oktavian, A., Pradana, A., and Lestari, D.** (2022). Transfer Learning for Alzheimer's Disease Classification Using ResNet-18 on MRI Images. *Computers in Biology and Medicine*, 147. <https://doi.org/10.1016/j.compbimed.2022.105763>
- Organization, W. H.** (2021). *Global status report on the public health response to dementia 2021*. World Health Organization.
- Pan, S. J., and Yang, Q.** (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359. <https://doi.org/10.1109/TKDE.2009.191>
- Prince, M., Wimo, A., Guerchet, M., Ali, G.-C., Wu, Y.-T., and Prina, M.** (2015). *World Alzheimer Report 2015: The Global Impact of Dementia*. Alzheimer's Disease International.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C.** (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4510–4520. <https://doi.org/10.1109/CVPR.2018.00474>
- Scheltens, P., De Strooper, B., Kivipelto, M., Holstege, H., Chetelat, G., Teunissen, C., Cummings, J., and van der Flier, W. M.** (2021). Alzheimer's disease. *The Lancet*, 397(10284), 1577–1590. [https://doi.org/10.1016/S0140-6736\(20\)32205-4](https://doi.org/10.1016/S0140-6736(20)32205-4)

KAYNAKLAR DİZİNİ (devam)

- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D.** (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 618–626. <https://doi.org/10.1109/ICCV.2017.74>
- Serrano-Pozo, A., Frosch, M. P., Masliah, E., and Hyman, B. T.** (2011). Neuropathological Alterations in Alzheimer Disease. *Cold Spring Harbor Perspectives in Medicine*, 1(1), a006189. <https://doi.org/10.1101/cshperspect.a006189>
- Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., Milchenko, M., Xu, R., Marcus, D., Colen, R. R., and Bakas, S.** (2020). Federated Learning in Medicine: Facilitating Multi-Institutional Collaborations without Sharing Patient Data. *Scientific Reports*, 10(1), 12598. <https://doi.org/10.1038/s41598-020-69250-1>
- Shen, D., Wu, G., and Suk, H. I.** (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19, 221–248. <https://doi.org/10.1146/annurev-bioeng-071516-044442>
- Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Noguees, I., Yao, J., Mollura, D., and Summers, R. M.** (2016). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Transactions on Medical Imaging*, 35(5), 1285–1298. <https://doi.org/10.1109/TMI.2016.2528162>
- Sorensen, L., Nielsen, M., and Initiative, A. D. N.** (2018). Ensemble support vector machine classification of dementia using structural MRI and mini-mental state examination. *Journal of Neuroscience Methods*, 302, 66–74. <https://doi.org/10.1016/j.jneumeth.2018.01.003>
- Suk, H.-I., Lee, S.-W., Shen, D., and (ADNI), A. D. N. I.** (2014). Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage*, 101, 569–582. <https://doi.org/10.1016/j.neuroimage.2014.06.077>
- Szegedy, C., Ioffe, S., Vanhoucke, V., and Alemi, A.** (2016). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 30. <https://ojs.aaai.org/index.php/AAAI/article/view/11231>

KAYNAKLAR DİZİNİ (devam)

- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A.** (2015). Going Deeper with Convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1–9. <https://doi.org/10.1109/CVPR.2015.7298594>
- Taheri Gorji, M. E., Mirniaharikandehi, S., Mohseni Salehi, S. S., Raahemifar, K., and Plataniotis, K. N.** (2022). An EfficientNet-based framework for Alzheimer’s disease classification using structural MRI data. *Biomedical Signal Processing and Control*, 75, 103590. <https://doi.org/10.1016/j.bspc.2022.103590>
- Tan, M., and Le, Q.** (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *Proceedings of the 36th International Conference on Machine Learning (ICML)*.
- Topol, E.** (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books.
- Vieira, S., Pinaya, W. H. L., and Mechelli, A.** (2017). Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. *Neuroscience and Biobehavioral Reviews*, 74, 58–75. <https://doi.org/10.1016/j.neubiorev.2017.01.002>
- Wen, H., Li, Y., Chen, W., Song, S., Qiao, Y., and Li, X.** (2020). Convolutional Neural Networks for Classification of Alzheimer’s Disease: A Review. *Neurocomputing*, 394, 41–53. <https://doi.org/10.1016/j.neucom.2020.01.110>
- Westman, E., Simmons, A., Muehlboeck, J. S., and Tunnard, C.** (2011). AddNeuroMed and ADNI: Similarities and differences in MRI analysis results for Alzheimer’s disease. *NeuroImage*, 58(3), 818–828. <https://doi.org/10.1016/j.neuroimage.2011.06.037>
- Yosinski, J., Clune, J., Bengio, Y., and Lipson, H.** (2014). How Transferable are Features in Deep Neural Networks? *Advances in Neural Information Processing Systems (NeurIPS)*, 27. <https://papers.nips.cc/paper/2014/hash/375c71349b295f2dcda9206f20a06-Abstract.html>
- Yousefi-Banaem, H.** (2021). Automated Hippocampal Segmentation in Alzheimer’s Disease Using U-Net Deep Learning Model. *Biomedical Signal Processing and Control*, 70. <https://doi.org/10.1016/j.bspc.2021.102991>

KAYNAKLAR DİZİNİ (devam)

Zhang, Y. (2018). Alzheimer's Disease Diagnosis Based on Inception-v4 Deep Learning Model and Optimized Activation Functions. *Cognitive Systems Research*, 52, 88–97.
<https://doi.org/10.1016/j.cogsys.2018.06.003>

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı : Aslıhan Güven

Doğum tarihi ve yeri :

e-posta :

Öğrenim Bilgileri

| Derece | Okul/Program | Yıl |
|-----------|---|------|
| Y. Lisans | Balıkesir Üniversitesi/Bilgisayar ve Bilişim Mühendisliği Anabilim Dalı | 2026 |
| Lisans | Balıkesir Üniversitesi/Bilgisayar Mühendisliği | 2022 |
| Lise | Ankara Açı Temel Lisesi | 2018 |

Yayın Listesi

Güven, A., Aydın, F., and Ezirmik, A. H. (2023). *Discriminating between Turkish traffic danger-warning signs and regulatory signs using ConvNets*. In M. Seenivasan, K. Thamizhmaran, and J. Jayachitra (Eds.), *13th International Istanbul Scientific Research Congress on Life, Engineering and Applied Sciences Proceedings Book* (p. 158). ISBN: 978-605-72563-8-6.

Güven, A., Ezirmik, A. H., Ceylan, M. F., and Aydın, F. (İncelemede). *Stage classification of Alzheimer's disease MRI data via deep neural network architectures*. *Mühendislik Bilimleri ve Araştırmaları Dergisi (BMBAD)*. [Tezden türetilmiştir]