



T.C.

TOKAT GAZİOSMANPAŞA ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİMDALI

BİLGİSAYAR MÜHENDİSLİĞİ YÜKSEK LİSANS PROGRAMI

RETİNA HASTALIKLARI TESPİTİ İÇİN BİR DERİN

ÖĞRENME MODELİ

YÜKSEK LİSANS TEZİ

Hazırlayan: Ali Aydın Abdulkhaleq

Danışman: Dr. Öğr. Üyesi Cengiz Güngör

TOKAT- 2025

ETİK SÖZLEŞME

Tokat Gaziosmanpaşa Üniversitesi Lisansüstü Eğitim Enstitüsü tez yazım kılavuzuna göre, Dr. Öğretim Üyesi Cengiz Güngör danışmanlığında hazırlamış olduğum “Retina Hastalıkları Tespiti İçin Bir Derin Öğrenme Modeli” adlı Yüksek Lisans tezinin bilimsel etik değerlere ve kurallara uygun, özgün bir çalışma olduğunu, aksinin tespit edilmesi halinde her türlü yasal yaptırımını kabul edeceğimi beyan ederim.

22/08/2025

Ali Aydın Abdulkhaleq

İmza

JÜRİ KABUL VE ONAY

Ali Aydın Abdulkhaleq tarafından hazırlanan “**Retina Hastalıkları Tespiti İçin Bir Derin Öğrenme Modeli**” adlı tez çalışmasının savunma sınavı 22/08/2025 tarihinde yapılmış olup aşağıda verilen Jüri tarafından Oy Birliği / Oy Çokluğu ile Tokat Gaziosmanpaşa Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği’nde Yüksek Lisans Tezi olarak kabul edilmiştir.

Jüri Üyeleri (Unvanı, Adı Soyadı)

İmzası

Üye (Başkan) :

.....

Üye :

.....

Üye :

.....

Üye :

.....

Üye :

.....

ONAY

...../...../.....

Lisansüstü Eğitim Enstitüsü Müdürü

TEŞEKKÜR

Tez konusunun belirlenmesinden çalışmanın tamamlanmasına kadar yürütülen bütün süreçte bana yol gösteren ve hiçbir yardımı esirgemeyen tez danışmanım Dr. Öğr. Üyesi Cengiz Güngör hocama;desteğini her zaman yanımda hissettiğim aileme,tezimdeki yardımlarından dolayı arkadaşım Abubakir Ali Hammood’a teşekkürlerimi sunarım.



Ali Aydın Abdulkhaleq

TOKAT 2025

ÖZET

RETİNA HASTALIKLARI TESPİTİ İÇİN BİR DERİN ÖĞRENME MODELİ

Ali Aydın Abdulkhaleq

BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Tez Danışmanı: Dr. Öğr. Üyesi Cengiz Güngör

Ağustos 2025, xii + 55 sayfa

Diyabetik retinopati, glokom, yaşa bağlı makula dejenerasyonu, katarakt ve diğer görme yetisini tehdit eden hastalıklar gibi retina hastalıkları, erken teşhis ve tedavi edilmezse genellikle geri dönüşü olmayan görme kaybına yol açan önemli bir küresel sağlık sorunu oluşturmaktadır. Bu tez, renkli fundus görüntüleri kullanarak retina hastalıklarının otomatik çok sınıflı sınıflandırılması için kapsamlı bir derin öğrenme çerçevesi önermektedir. EfficientNetB0 konvolüsyonel sinir ağı mimarisini ince ayar ve sağlam veri artırma stratejileriyle kullanarak, çalışma retina görüntü analizindeki kritik zorlukları, sınıf dengesizliği, sınırlı anotlanmış veri kümeleri ve yapay zeka sistemlerinin yorumlanabilirliği dahil olmak üzere ele almayı amaçlamaktadır.

Metodoloji, kamuya açık retina görüntü koleksiyonlarını entegre ederek ve genişleterek büyük, dengeli bir veri seti oluşturmayı, ardından sistematik veri ön işlemeyi ve eğitim ve doğrulama setlerine bölmeyi içermektedir. Önerilen model, tanı doğruluğu ve genelleştirilebilirlik üzerindeki etkisini ölçmek için veri genişletme ile ve veri genişletme olmadan değerlendirilmiştir. Ayrıca, özellik çıkarma ve model performansını daha da geliştirmek için dikkat mekanizmaları ve transfer öğrenimi araştırılmıştır. Deney sonuçları, genişletilmiş EfficientNetB0 modelinin %96,3 eğitim doğruluğu ve %91,83 test doğruluğu elde ettiğini ve literatürde bildirilen birkaç son teknoloji yöntemi geride bıraktığını göstermiştir. Hassasiyet, geri çağırma ve F1 puanı dahil olmak üzere sınıf bazlı metrikler, tüm hastalık kategorilerinde sağlam ve dengeli bir tespit sağlarken, karışıklık matrisleri ve görselleştirmeler model davranışı ve potansiyel yanlış sınıflandırma kaynakları hakkında daha fazla bilgi sağlamıştır.

Bu çalışmada temel model olarak VGG16 da EfficientNetB0'a ek olarak değerlendirilmiştir. ImageNet ağırlıklarıyla başlatılıp 50 epoch boyunca tam ince ayar yapılan

VGG16, %90,35 eğitim doğruluğu ve %86,33 test doğruluğu elde etmiştir. Altı sınıfta (ACRIMA, Cataract, Glaucoma, ODIR-5K, ORIGA, Retina_Disease) raporlanan sınıf-bazlı ölçütler ve karışıklık matrisi, modelin çok sınıflı senaryoda dengeli ancak EfficientNetB0'a göre daha düşük bir performans sağladığını göstermektedir. Bu bulgular, önerilen EfficientNetB0 yaklaşımının üstünlüğünü nicel olarak desteklemektedir.

Bu tezin bulguları, tıbbi görüntü veri kümelerinin yaygın sınırlamalarının üstesinden gelmede gelişmiş artırma ve aktarım öğreniminin önemini vurgulamakta ve güvenilir, ölçeklenebilir ve açıklanabilir retina hastalığı taraması için derin öğrenme modellerinin uygulanabilirliğini göstermektedir. Çalışma, önerilen yaklaşımın pratik uygulamalarını tartışarak ve otomatik oftalmik teşhisi ilerletmek için gelecekteki araştırmalara yönelik önerileri özetleyerek son bulmaktadır.

Anahtar kelimeler: Retina hastalıkları, Derin öğrenme, VGG16, EfficientNetB0, Veri artırma, Otomatik teşhis.

ABSTRACT

A DEEP LEARNING MODEL FOR RETINAL DISEASE DETECTION

Ali Aydin Abdulkhaleq

DEPARTMENT OF COMPUTER ENGINEERING

Advisor: Assoc. Prof. Dr. Cengiz Güngör

August 2025, xii + 55 Pages

Retinal diseases such as diabetic retinopathy, glaucoma, age-related macular degeneration, cataract, and other sight-threatening conditions represent a major global health challenge, often leading to irreversible vision loss if not diagnosed and treated early. This thesis proposes a comprehensive deep learning framework for the automated multi-class classification of retinal diseases using color fundus images. Leveraging the EfficientNetB0 convolutional neural network architecture with fine-tuning and robust data augmentation strategies, the study aims to address critical challenges in retinal image analysis, including class imbalance, limited annotated datasets, and the interpretability of artificial intelligence systems.

The methodology involved constructing a large, balanced dataset by integrating and augmenting publicly available retinal image collections, followed by systematic data preprocessing and partitioning into training and validation sets. The proposed model was evaluated both with and without data augmentation to quantify its impact on diagnostic accuracy and generalizability. In addition, attention mechanisms and transfer learning were explored to further enhance feature extraction and model performance. The experimental results demonstrated that the augmented EfficientNetB0 model achieved a training accuracy of 96.3% and a testing accuracy of 91.83%, outperforming several state-of-the-art methods reported in the literature. Class-wise metrics, including precision, recall, and F1-score, indicated robust and balanced detection across all disease categories, while confusion matrices and visualizations provided further insights into model behavior and potential sources of misclassification.

In addition to our primary model EfficientNetB0, we also evaluate VGG16 as a strong baseline. Fine-tuned from ImageNet for 50 epochs, VGG16 attains 90.35% training accuracy and

86.33% testing accuracy on six classes (ACRIMA, Cataract, Glaucoma, ODIR-5K, ORIGA, Retina_Disease). Class-wise metrics and the confusion matrix reveal balanced yet lower performance than EfficientNetB0, thereby reinforcing the superiority of the proposed approach.

The findings of this thesis underscore the importance of advanced augmentation and transfer learning in overcoming common limitations of medical image datasets and demonstrate the feasibility of deploying deep learning models for reliable, scalable, and explainable retinal disease screening. The study concludes by discussing the practical implications of the proposed approach and outlining recommendations for future research to advance automated ophthalmic diagnostics.

Keywords: Retinal diseases, Deep learning, VGG16, EfficientNetB0, Data augmentation, Automated diagnosis.

İÇİNDEKİLER

ETİK SÖZLEŞME.....	i
ÖZET	iv
ABSTRACT.....	vi
İÇİNDEKİLER	viii
ŞEKİL LİSTESİ.....	x
TABLolar LİSTESİ	xi
SİMGELER VE KISALTMALAR.....	xii
1 GİRİŞ	1
1.1. Arka Plan.....	1
1.2. Motivasyon ve Önemi	1
1.3. Sorun Tanımı	2
1.4. Araştırma Hedefleri.....	2
1.5. Tezin Yapısı	3
2.LİTERATÜR İNCELEMESİ.....	5
2.1. Giriş.....	5
2.2. Retina Hastalıklarının Sınıflandırılmasındaki Zorluklar	5
2.3. Tıbbi Görüntülemeye Derin Öğrenme	7
2.4. Tıbbi Görüntü Analizi için CNN Mimarileri	8
2.5. Retina Hastalıklarının Sınıflandırılmasında Son Gelişmeler.....	10
2.6. Retina Görüntüleme için Derin Öğrenmede Artırma ve Düzenleme.....	12
2.7. Retina Hastalıklarının Sınıflandırılmasında Transfer Öğrenimi ve Genelleme	14
2.8. Retina Hastalığı Sınıflandırılmasında Dikkat Mekanizmaları ve Açıklanabilirlik	16
3. MALZEMELER VE METODOLOJİ.....	18
3.1. Veri Kümesi Açıklaması	18

3.2. Veri Artırımı	20
3.3. Veri Ön İşleme	21
3.4. Model Mimarisi	22
3.4.1. Veri Artırımı ile EfficientNetB0.....	24
3.4.2. Veri Artırımı Olmadan EfficientNetB0	25
3.4.3. VGG16 (Temel Karşılaştırma Modeli)	27
3.5. Eğitim Prosedürü	27
3.6. Değerlendirme Metrikleri	28
4. DENEY SONUÇLARI	31
4.1. Veri Artırımı ile EfficientNetB0.....	32
4.2 Veri Artırımı Olmadan EfficientNetB0	36
4.3 Veri Artırımı ile VGG16	40
4.4. Veri Artırma Olmadan VGG16	42
4.5. Önceki Çalışmalarla Karşılaştırmalı Analiz.....	43
4.6. Tartışma.....	46
5. SONUÇ VE GELECEK ÇALIŞMALAR	47
5.1. Sonuç.....	47
5.2. Gelecekteki Çalışmalar	49
6. KAYNAKLAR.....	51
ÖZGEÇMİŞ	56

ŞEKİL LİSTESİ

Şekil 2.1: EfficientNetB0 yapısı (ResearchGate, 2025)	9
Şekil 3.1: Çalışmada kullanılan ACRIMA, ODIR-5K ve Glaucoma veri setlerinden örnek retina fundus görüntüleri. Görseller, sınıflar arası çeşitliliği ve patolojik özellikleri yansıtmaktadır	20
Şekil 3.2: Önerilen Retina Hastalığı Sınıflandırma Sisteminin Genel İş Akışı	24
Şekil 3.3: Karışıklık matrisi yapısı.....	29
Şekil 3.4: Doğruluk, Kesinlik, Duyarlılık ve F1 Skoru formülleri	29
Şekil 4.1: Göz Hastalıkları Sınıflandırması için Karışıklık Matrisi (Artırma ile)	35
Şekil 4.2: Eğitim ve Doğrulama Doğruluğu (Artırma ile).....	35
Şekil 4.3: Eğitim ve Doğrulama Kaybı (Artırma ile)	36
Şekil 4.4: Karışıklık Matrisi (Artırma Olmadan).....	38
Şekil 4.5: Eğitim ve Doğrulama Doğruluğu (Artırma Olmadan)	38
Şekil 4.6: Eğitim ve Doğrulama Kaybı (Artırma olmadan).....	39
Şekil 4.7: Veri arttırımı ile VGG16 için karışıklık matrisi.....	41
Şekil 4.8: Veri arttırımı ile VGG16 için eğitim doğruluğu ve kaybı.....	42
Şekil 4.9: Veri artırma olmadan VGG16 için karışıklık matrisi	43
Şekil 4.10: Veri artırma olmadan VGG16 için eğitim doğruluğu ve kaybı	43

TABLÖLAR LİSTESİ

Tablo 4.1: Eğitim doğruluğu ile doğrulama (test) doğruluğunun karşılaştırılması (Artırma ile) .	34
Tablo 4.2 Artırma olmaksızın eğitim doğruluğu ile doğrulama (test) doğruluğunun karşılaştırılması	37
Tablo 4.3: VGG16 için eğitim doğruluğu ile doğrulama (test) doğruluğunun karşılaştırılması (Artırma ile).....	41
Tablo 4.4. VGG16 için eğitim doğruluğu ile doğrulama (test) doğruluğunun karşılaştırılması (Artırma olmadan).....	42
Tablo 4.6: Önceki çalışmalarla karşılaştırmalı analiz	45

SİMGELER VE KISALTMALAR

Kısaltmalar	Açıklama
ACRIMA	: Tıbbi Görüntü Veritabanı
AI	: Yapay Zeka
BT	: Bilgisayarlı Tomografi
CBAM	: Convolutional Block Attention Modules, Evrişimsel Blok Dikkat Modülleri
CNN	: Convolutional Neural Network
CT	: Bilgisayarlı Tomografi
ESA	: Evrişimsel Sinir Ağları
FN	: Yanlış Sınıflandırılan Negatif Örnekler
FP	: Yanlış Sınıflandırılan Pozitif Örnekler
GAN	: Üretken Karşıt Ağlar
Grad-CAM	: Gradyan Sınıfı Aktivasyon Haritaları
MIMIC-CXR	: Tıbbi Görüntü Veritabanı
MRI, MRG	: Manyetik Rezonans Görüntüleme
OCT	: Optik Koherens Tomografi
ODIR-5K	: 5,000 Hasta Görüntülü Ocular Disease Intelligent Recognition Veritabanı
ORIGA	: Tıbbi Görüntü Veritabanı
RGB	: Kırmızı-Yeşil-Mavi Kanalları
SE	: Squeeze-and-Excitation, Dikkat Mekanizması Blokları
TN	: Doğru Sınıflandırılan Negatif Örnekler
TP	: Doğru Sınıflandırılan Pozitif Örnekler
VIT	: Görsel Gönüştürücüler
XAI	: Açıklanabilir AI

1 GİRİŞ

1.1. Arka Plan

Retina hastalıkları, dünya çapında görme bozukluğu ve körlüğün başlıca nedenleri oldukları için önemli bir halk sağlığı sorunu teşkil etmektedir. Diyabetik retinopati, glokom, yaşa bağlı makula dejenerasyonu ve katarakt gibi hastalıklar her yıl milyonlarca insanı etkilemekte ve zamanında teşhis ve tedavi edilmezlerse yaşam kalitesinde ciddi bir düşüşe neden olmaktadır. Bu retina hastalıklarının doğru ve erken teşhisi, etkili tedavi ve geri dönüşü olmayan görme kaybı riskinin en aza indirilmesi için çok önemlidir. Geleneksel olarak, retina hastalığı teşhisi, eğitimli göz doktorları tarafından fundus fotoğraflarının ve optik koherens tomografi (OCT) görüntülerinin manuel olarak incelenmesine dayanmaktadır. Bu süreç oldukça etkili olsa da, doğası gereği zaman alıcı, öznel ve yetenekli profesyonellerin mevcudiyetine bağlıdır; bu kaynak ise dünyanın birçok bölgesinde genellikle sınırlıdır. Küresel nüfusun yaşlanması ve diyabet prevalansının artmasıyla birlikte, retina hastalıklarının sağlık sistemleri üzerindeki yükünün önümüzdeki on yıllarda önemli ölçüde artması beklenmektedir (Wong ve ark., 2023; Ting ve ark., 2019; Cheung ve ark., 2021).

1.2. Motivasyon ve Önemi

Küresel diyabet artışları ve yaşlanan nüfus gibi faktörlerin etkisiyle dünya çapında retina hastalıklarının görülme sıklığının artması, sağlık sistemleri için önemli bir zorluk oluşturmaktadır (Cheung ve ark., 2021). Geri dönüşü olmayan görme kaybını önlemek için erken teşhis ve müdahale kritik öneme sahiptir, ancak özellikle kaynakların kısıtlı olduğu ortamlarda uzman oftalmik bakıma erişim genellikle sınırlıdır. Retina görüntülerinin manuel olarak yorumlanması, sadece emek yoğun ve zaman alıcı olmakla kalmaz, aynı zamanda gözlemciler arasında değişkenliğe de maruz kalır, bu da tutarsız teşhislere yol açabilir (Yim ve ark., 2020). Bu gerçekler, klinisyenlerin retina hastalıklarını yüksek doğruluk ve verimlilikle taramasına ve teşhis etmesine yardımcı olabilecek otomatik, ölçeklenebilir ve güvenilir çözümlere acil ihtiyaç olduğunu vurgulamaktadır.

Yapay zeka, özellikle de derin öğrenme, karmaşık retina görüntülerini analiz edebilen ve uzman klinisyenlerle karşılaştırılabilir bir doğrulukla patolojik değişiklikleri tespit edebilen

otomatik sistemlerin geliştirilmesini sağlayarak bu ihtiyaçları karşılamada büyük umut vaat etmektedir (Gulshan ve ark., 2016; Rajalakshmi ve ark., 2022). Bu tür teknolojinin geniş ölçekte kullanılması, tarama programlarını dönüştürebilir, tanı hatalarını azaltabilir ve çeşitli klinik ve topluluk ortamlarında görmeyi kurtarıcı bakıma erişimi genişletebilir. Retina hastalıklarının yükü artmaya devam ederken, sağlam yapay zeka (AI) destekli tanı araçlarının geliştirilmesi ve entegrasyonu, küresel oftalmik sağlık için kritik bir ilerlemeyi temsil etmektedir.

1.3. Sorun Tanımı

Yapay zeka ve tıbbi görüntülemelerde önemli ilerlemeler kaydedilmesine rağmen, retina hastalıklarının güvenilir ve otomatik tespiti karmaşık ve çözülmemiş bir sorun olmaya devam etmektedir. Birçok derin öğrenme modeli, deneysel ortamlarda etkileyici sonuçlar elde etmesine rağmen, görüntüleme koşulları, cihaz türleri ve hasta demografisi farklılıkları nedeniyle çeşitli klinik ortamlarda etkili bir şekilde genelleştirilememektedir (Gupta ve ark., 2021). Ek olarak, birçok kamuya açık retina veri setinde sınıf dengesizliği ve sınırlı anotlanmış veri sorunu, daha az temsil edilen hastalık kategorilerinde düşük performans gösteren önyargılı modellere yol açmaktadır (Quelleg ve ark., 2017). Farklı retina hastalıklarında görsel olarak benzer özelliklerin varlığı, model eğitimi ve değerlendirmesini daha da karmaşık hale getirerek yanlış sınıflandırma olasılığını artırmaktadır (Ting ve ark., 2019). Dahası, birçok derin öğrenme yaklaşımının “kara kutu” niteliği, yorumlanabilirlik ve klinik güven konusunda endişeler uyandırarak, bu teknolojilerin rutin klinik iş akışlarına kabulünü ve entegrasyonunu sınırlamaktadır (Tjoa ve Guan, 2020). Retina hastalığı tespiti için klinik olarak yararlı ve güvenilir AI sistemleri geliştirmek için bu zorlukların üstesinden gelmek çok önemlidir.

1.4. Araştırma Hedefleri

Bu araştırmanın temel amacı, renkli fundus görüntülerinden retina hastalıklarının otomatik tespiti ve çok sınıflı sınıflandırılması için derin öğrenme tabanlı bir çerçeve geliştirmek ve değerlendirmektir. Çalışma, özellikle EfficientNetB0 gibi gelişmiş evrişimli sinir ağı mimarilerinin, ince ayar ve veri artırma teknikleriyle birlikte, birden fazla hastalık kategorisinde doğru ve dengeli sınıflandırma elde etmek için etkinliğini araştırmayı

amaçlamaktadır. Ayrıca, araştırma veri dengesizliği, sınırlı anotlanmış veri kümeleri ve model yorumlanabilirliği ile ilgili temel zorlukları ele almayı ve böylece klinik olarak uygulanabilir AI odaklı tanı araçlarının geliştirilmesine katkıda bulunmayı amaçlamaktadır (Li ve ark., 2021; Shamsi ve ark., 2021).

Bir diğer amaç ise, önerilen yaklaşımı literatürde bildirilen mevcut en gelişmiş yöntemlerle sistematik olarak karşılaştırarak, göreceli güçlü yönlerini vurgulamak ve daha fazla iyileştirme gereken alanları belirlemektir. Bu hedefleri gerçekleştirerek, bu tez, otomatik retina hastalığı tespiti alanını ilerletmeyi, erken teşhisi destekleyebilecek, klinisyenlerin yükünü azaltabilecek ve görme bozukluğu riski altındaki nüfusların göz bakımına erişimini iyileştirebilecek çözümler sunmayı amaçlamaktadır (Rajpurkar ve ark., 2022; Ting ve ark., 2019).

1.5. Tezin Yapısı

Bu tez, okuyucuyu araştırmanın motivasyonu, metodolojisi, deneyleri ve sonuçları hakkında sistematik olarak bilgilendirmek üzere tasarlanmış beş ana bölümden oluşmaktadır.

Bölüm 1, araştırma bağlamına genel bir bakış sunarak arka plan, motivasyon, problem tanımı, amaçlar, araştırma soruları ve tezin genel yapısını özetlemektedir.

Bölüm 2, retina görüntülemenin evrimi, tıbbi görüntü analizinde derin öğrenmenin gelişimi ve uygulaması, çok sınıflı sınıflandırma ile ilgili zorluklar ve retina hastalığı tespiti ile ilgili konvolüsyonel sinir ağı mimarileri ve metodolojilerindeki son gelişmeleri tartışarak kapsamlı bir literatür incelemesi sunmaktadır.

Bölüm 3, veri seti açıklaması, veri ön işleme ve artırma stratejileri, EfficientNetB0 tabanlı derin öğrenme modelinin tasarımı ve yapılandırması ve değerlendirme için kullanılan deneysel protokoller dahil olmak üzere, bu çalışmada kullanılan materyalleri ve yöntemleri ayrıntılı olarak açıklamaktadır.

Bölüm 4, önerilen yaklaşımı en son yöntemlerle karşılaştırarak, artırma, transfer öğrenimi ve dikkat mekanizmalarının model performansı üzerindeki etkilerini inceleyerek deneysel sonuçları ve analizleri sunmaktadır.

Son olarak, Bölüm 5, ana bulguları özetleyerek, pratik ve klinik çıkarımları tartışarak ve otomatik retina hastalığı tespiti konusunda gelecekteki araştırmalar için yollar önererek tezi sonlandırmaktadır.

Bu mantıksal düzenleme, netlik ve tutarlılığı kolaylaştırmak amacıyla yapılmıştır ve okuyucuların temel kavramlardan uygulamalı araştırma ve eleştirel değerlendirmeye kadar olan süreci takip etmelerini sağlar (Creswell ve Creswell, 2018).



2.LİTERATÜR İNCELEMESİ

2.1. Giriş

Son on yılda, tıbbi görüntüleme ve yapay zeka alanındaki gelişmeler, özellikle retina hastalıklarının tespiti ve sınıflandırılması konusunda oftalmoloji alanında devrim niteliğinde değişiklikler getirmiştir. Yüksek çözünürlüklü retina görüntülerinin giderek daha fazla kullanılabilir hale gelmesi ve derin öğrenme teknolojilerinin yükselişi, otomatik, doğru ve hızlı tanı için yeni yollar açmıştır. Bu bölüm, derin öğrenme kullanılarak retina hastalığının teşhisi ile ilgili literatürü kapsamlı bir şekilde gözden geçirir ve özellikle evrimsel sinir ağı mimarileri, veri kümesi zorlukları, artırma stratejileri ve mevcut en son teknoloji sonuçlarına odaklanır. Önceki araştırmaları eleştirel bir şekilde analiz ederek, bu inceleme hem kaydedilen ilerlemeyi hem de bu tezde sunulan araştırmayı motive eden kalıcı boşlukları aydınlatmayı amaçlamaktadır.

2.2. Retina Hastalıklarının Sınıflandırılmasındaki Zorluklar

Görüntüleme teknolojileri ve yapay zeka alanlarında önemli ilerlemeler kaydedilmesine rağmen, otomatik retina hastalığı sınıflandırma görevini zorlaştıran birkaç kalıcı zorluk bulunmaktadır. En önemli engellerden biri, kamuya açık retina görüntü veri setlerinin doğasında var olan dengesizlik ve sınırlı boyutudur (Quellec ve ark., 2017; Kaggle, 2023). Birçok retina patolojisi nispeten nadirdir, bu da belirli hastalıkların yetersiz temsil edildiği çarpık sınıf dağılımlarına yol açar ve derin öğrenme modellerinin azınlık sınıfları için sağlam özellikleri öğrenmesini zorlaştırır (Budai ve ark., 2013). Bu tür dengesizlikler genellikle daha yaygın durumların tespitini destekleyen önyargılı modellere yol açarken, daha nadir ancak klinik olarak önemli hastalıklar üzerindeki performans optimal düzeyde kalmaz (Gupta ve ark., 2021).

Bir diğer önemli zorluk, görüntü elde etme koşullarındaki yüksek değişkenlikten kaynaklanmaktadır. Kamera ekipmanı, aydınlatma, odaklama ve hasta işbirliği gibi faktörler, retina görüntülerine önemli ölçüde gürültü ve artefaktlar ekleyebilir (Abramoff ve ark., 2016). Bu değişkenlik, sonuçların tekrarlanabilirliğini azaltır ve sistemler farklı klinik ortamlarda veya popülasyonlarda kullanıldığında model performansında önemli düşüşlere

yol açabilir (Ting ve ark., 2019). Ek olarak, ortam opaklıkları, hasta hareketi veya yanlış odaklama nedeniyle görüntü kalitesinin düşük olması sık karşılaşılan bir sorun olmaya devam etmekte ve bazen analizden önce manuel kürasyon veya otomatik kalite değerlendirmesi gerektirmektedir (Fu ve ark., 2022).

Farklı retina hastalıkları arasında görsel olarak benzer lezyonların ve örtüşen özelliklerin varlığı başka bir zorluk oluşturmaktadır. Örneğin, hem diyabetik retinopati hem de hipertansif retinopati kanamalar ve mikroanevrizmalarla ortaya çıkabilirken, glokom ve optik nörit her ikisi de optik disk çukurlaşması veya solukluk olarak ortaya çıkabilir (Ting ve ark., 2019; Yim ve ark., 2020). Bu örtüşme, yalnızca temel gerçek etiketleme için açıklama sürecini değil, aynı zamanda modelin hastalığa özgü ince değişiklikleri doğru bir şekilde ayırt etme yeteneğini de zorlaştırır. Deneyimli oftalmologlar bile sınırda veya atipik vakaların teşhisi konusunda fikir ayrılığına düşebilirler, bu da sınıflandırma görevinin doğasında var olan karmaşıklığı daha da vurgular (Cheung ve ark., 2021).

Son olarak, klinik yorumlanabilirlik ve model şeffaflığı sorunu, tıbbi yapay zeka alanında özellikle ciddi bir sorundur. Klinisyenler, özellikle tedavi veya cerrahi planlamayı etkileyebilecek yüksek riskli senaryolarda, modelin kararları için sadece doğru tahminler değil, aynı zamanda anlamlı açıklamalar da talep etmektedir (Caruana ve ark., 2015). Kara kutu derin öğrenme modelleri, güçlü olmalarına rağmen, yorumlanabilirlik eksikliği nedeniyle sıklıkla eleştirilmektedir. Bu sınırlama, saliency map ve layer-wise relevance propagation gibi açıklanabilir yapay zeka teknikleri üzerine araştırmaları teşvik etmiştir (Samek ve ark., 2017; Lundberg ve Lee, 2017).

Bu çok yönlü zorlukların ele alınması, oftalmolojide klinik olarak uygulanabilir yapay zeka sistemlerinin geliştirilmesi için gereklidir. Devam eden araştırmalar, veri seti çeşitliliğini genişletmeye, görüntü kalitesi kontrolünü iyileştirmeye, gelişmiş artırma ve aktarım öğrenme yöntemleri geliştirmeye ve açıklanabilir AI çerçevelerini teşhis süreçlerine entegre etmeye odaklanmaktadır (Gulshan ve ark., 2016; Rajalakshmi ve ark., 2022). Bu çabalar, deneysel başarı ile pratik, gerçek dünya uygulamaları arasındaki uçurumu kapatmayı amaçlamaktadır.

2.3. Tıbbi Görüntülemede Derin Öğrenme

Derin öğrenmenin tıbbi görüntüleme alanına entegrasyonu, geniş ve heterojen görüntü veri kümelerinden karmaşık özelliklerin otomatik olarak çıkarılmasını sağlayarak bu alanda devrim yaratmıştır. El ile oluşturulan özellik mühendisliğine ve alana özgü bilgiye büyük ölçüde bağımlı olan geleneksel makine öğrenimi yaklaşımlarının aksine, derin öğrenme modelleri, özellikle de evrişimli sinir ağları (CNN'ler), hammadde piksel verilerinden doğrudan hiyerarşik temsilleri öğrenme yeteneğine sahiptir (LeCun, Bengio ve Hinton, 2015). Bu paradigma değişikliği, minimum insan müdahalesi ile büyük hacimli tıbbi görüntüleri işleyip analiz edebilen, yüksek doğrulukta ve ölçeklenebilir tanı araçlarının geliştirilmesini mümkün kılmıştır (Litjens ve ark., 2017).

Son yıllarda, CNN'ler radyografi, bilgisayarlı tomografi (BT), manyetik rezonans görüntüleme (MRG) ve fundus fotoğrafçılığı gibi modalitelerde organ segmentasyonu, lezyon tespiti ve hastalık sınıflandırması dahil olmak üzere çok çeşitli tıbbi görüntüleme görevlerinde başarıyla uygulanmıştır (Esteva ve ark., 2017; Greenspan ve ark., 2016). Özellik hiyerarşilerini otomatik olarak öğrenme ve uyarlama yetenekleri, retina morfolojisindeki ince farklılıkların çeşitli patolojik durumların göstergesi olabileceği oftalmolojide özellikle değerli olduğu kanıtlanmıştır (Gulshan ve ark., 2016). CNN'lerin giderek daha fazla benimsenmesi, özellikle tutarlılık ve hızın çok önemli olduğu tarama ortamlarında, deneyimli insan uzmanların performansına rakip olan ve bazı durumlarda onu aşan tanı algoritmalarına yol açmıştır (Ting ve ark., 2019).

Tıbbi görüntü analizinde derin öğrenmenin en önemli avantajlarından biri, uygun şekilde eğitilip doğrulanmış olduğunda farklı görüntüleme cihazları, popülasyonlar ve edinim protokolleri arasında genelleme yapabilme kapasitesidir (De Fauw ve ark., 2018). ImageNet gibi büyük ölçekli, genel amaçlı veri kümelerinde önceden eğitilmiş modellerin tıbbi görüntüleme görevlerinde ince ayarlandığı transfer öğrenimi, performans ve eğitim verimliliğinde önemli iyileştirmeler sağlayan standart bir uygulama haline gelmiştir (Shin ve ark., 2016; Raghu ve ark., 2019). Bu strateji, sınırlı sayıda anotlanmış tıbbi verinin sorununu çözmekle kalmaz, aynı zamanda milyonlarca doğal görüntü üzerinde eğitilmiş ağlarda kodlanmış zengin görsel bilgiyi de kullanır.

Başarılarına rağmen, derin öğrenmenin tıbbi görüntüleme de kullanılması da kendine özgü zorluklar ortaya çıkarmaktadır. Yüksek kaliteli etiketli veri kümelerinin derlenmesi genellikle zor ve pahalıdır ve görüntü kalitesi, etiketleme standartları ve hastalık yaygınlığındaki farklılıklar, model performansını etkileyen önyargılara yol açabilir (Oakden-Rayner, 2020). Ek olarak, derin öğrenme modellerinin “kara kutu” niteliği, yorumlanabilirlik ve klinik güven konusunda endişeler doğurmuş ve model tahminleri için anlamlı görselleştirmeler ve gerekçeler sağlamayı amaçlayan açıklanabilir AI teknikleri üzerine araştırmalar yapılmasına neden olmuştur (Tjoa ve Guan, 2020).

Bununla birlikte, derin öğrenme, büyük veri ve bilgi işlem donanımındaki gelişmelerin bir araya gelmesi, otomatik tıbbi görüntü analizinde mümkün olanın sınırlarını zorlamaya devam etmektedir. Modeller daha sofistike ve veri kümeleri daha kapsamlı hale geldikçe, derin öğrenmenin tanı doğruluğunu artırma, iş yükünü azaltma ve uzman düzeyinde bakıma erişimi demokratikleştirme rolünün daha da genişlemesi muhtemeldir (Topol, 2019; Rajpurkar ve ark., 2022).

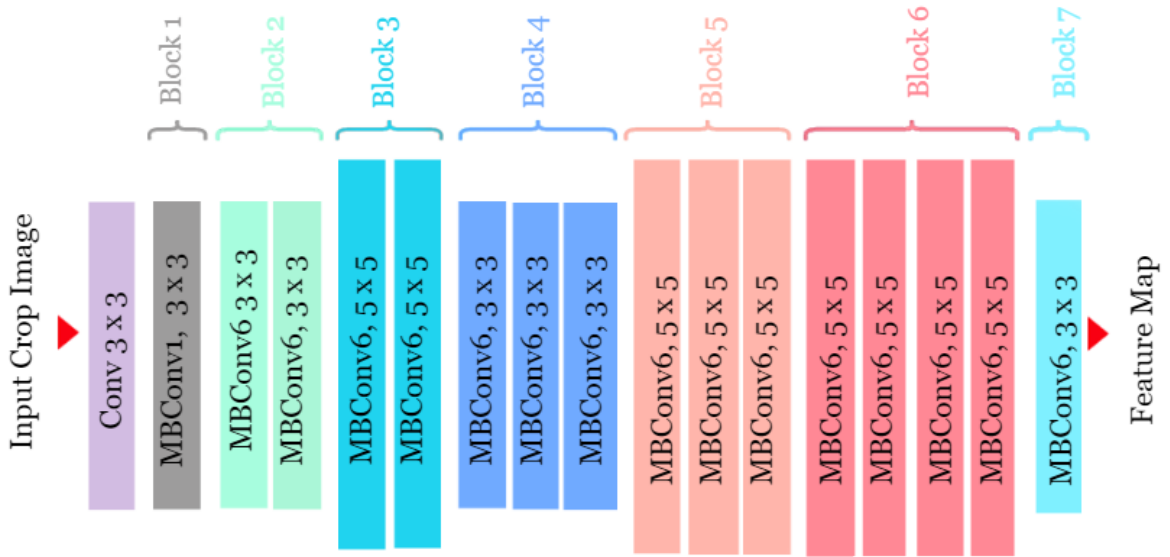
2.4. Tıbbi Görüntü Analizi için CNN Mimarileri

Konvolüsyonel sinir ağları (CNN'ler), ham görüntü verilerinden doğrudan hiyerarşik, uzamsal farkındalığa sahip özellikleri öğrenme konusundaki güçlü yetenekleri nedeniyle, özellikle tıbbi görüntü analizi alanında modern bilgisayar görüşünün bel kemiği haline gelmiştir (Krizhevsky ve ark., 2012). Öncü AlexNet mimarisi, 2012 yılında bir dönüm noktası oluşturmuş ve derin, katmanlı sinir ağlarının büyük ölçekli görüntü sınıflandırma görevlerinde geleneksel algoritmalarından daha iyi performans gösterebileceğini kanıtlamıştır (Russakovsky ve ark., 2015). Bu çıkış açan gelişme, CNN tasarımında bir inovasyon dalgasını tetiklemiş ve temsil gücü, hesaplama verimliliği ve eğitim istikrarını iyileştirmeyi amaçlayan daha sofistike mimarilerin geliştirilmesine yol açmıştır.

VGGNet (Simonyan ve Zisserman, 2015), GoogLeNet (Szegedy ve ark., 2015) ve ResNet (He ve ark., 2016) gibi sonraki mimariler sırasıyla daha derin ağlar, başlangıç modülleri ve kalıntı bağlantıları tanıttı. VGGNet, yığılmış 3x3 konvolüsyonel katmanların kullanımıyla basitliği vurguladı; bu katmanlar, derinliği artırmasına rağmen, hesaplama taleplerini yönetilebilir düzeyde tuttu. GoogLeNet (Inception), ağırlık bilgileri paralel olarak çoklu

ölçeklerde işlemesine olanak tanıyan başlangıç modülleri kavramını tanıttı ve böylece çok çeşitli görsel kalıpları yakalama kapasitesini artırdı. ResNet'in kısayol veya kalıntı bağlantılarını tanıtmaması, kaybolan gradyan sorununu çözdü ve yüzlerce hatta binlerce katmandan oluşan çok derin ağların eğitilmesini mümkün kılmış ve hem doğal hem de tıbbi görüntü analizinde temel bir unsur haline gelmiştir.

DenseNet (Huang ve ark., 2017), her katmanın önceki tüm katmanlardan girdi aldığı yoğun bağlantıyı tanıtarak bu alanı daha da ilerletmiş, özelliklerin yeniden kullanımını teşvik etmiş ve kaybolan gradyan sorununu daha da azaltmıştır. Bu tasarım, parametrelerin verimli kullanımını sağlamakta ve geleneksel mimarilere kıyasla daha az parametre ile genellikle üstün doğruluk elde etmektedir.



Şekil 0.1: EfficientNetB0 yapısı (ResearchGate, 2025)

Daha yakın zamanda, EfficientNetB0 (Tan ve Le, 2019) ağ derinliğini, genişliğini ve giriş çözünürlüğünü eşit şekilde ölçekleyen bir bileşik ölçeklendirme yöntemi getirerek, olağanüstü hesaplama verimliliği ile en son teknoloji performansı sunan bir model ailesi oluşturmuştur. Şekil 2.1'de yapısı gösterilen EfficientNetB0 modelleri, doğruluk ve kaynak kullanımı arasında güçlü bir denge kurmaları nedeniyle tıbbi görüntüleme araştırmalarında

hızla popülerlik kazanmıştır ve donanım kısıtlamalarının sorun olabileceği gerçek dünya klinik uygulamaları için özellikle caziptir (Shamsi ve ark., 2021).

CNN mimarisindeki bu gelişmeler, tıbbi görüntü analizinde dönüştürücü bir etki yaratarak, fundus fotoğrafçılığı, OCT, MRI ve CT gibi çeşitli modalitelerde hastalık tespiti, segmentasyon ve prognoz gibi görevlerde önemli iyileştirmeler sağlamıştır (Shin ve ark., 2016; Litjens ve ark., 2017). Önemli olarak, CNN'lerin modülerliği ve aktarılabilirliği, araştırmacıların önceden eğitilmiş modelleri kullanarak nispeten küçük, alana özgü tıbbi veri kümeleri üzerinde ince ayar yapmalarına olanak tanımış, sınırlı anotlanmış verilerin getirdiği bazı önemli sınırlamaları aşmalarını sağlamıştır (Raghu ve ark., 2019).

Güçlerine rağmen, tıbbi görüntülemelerde derin CNN'lerin kullanımı zorluklar içermez değildir. Modellerin yüksek karmaşıklığı, özellikle eğitim verileri sınırlı veya dengesiz olduğunda aşırı uyum riskini artırabilir. Dahası, derin mimarilerin kara kutu niteliği, yorumlanabilirlik ve klinik güvenilirlik konusunda sorunlar doğurmaktadır. Bu sorunlar, son tasarımlara dikkat mekanizmaları ve açıklanabilir AI tekniklerinin entegrasyonuna ilham vermiştir (Tjoa ve Guan, 2020; Selvaraju ve ark., 2017).

Genel olarak, CNN mimarilerinin evrimi, bilgisayar destekli tanı yeteneklerinin geliştirilmesinde önemli bir rol oynamıştır ve tıp alanında daha doğru, sağlam ve açıklanabilir AI sistemleri için çabalar sürerken, araştırmaların ana odak noktası olmaya devam etmektedir.

2.5. Retina Hastalıklarının Sınıflandırılmasında Son Gelişmeler

Son yıllarda, retina hastalıklarının sınıflandırılması alanında, büyük ölçüde derin öğrenmedeki ilerlemeler ve anotlanmış fundus ve OCT görüntü veri setlerinin giderek daha fazla kullanılabilir hale gelmesi sayesinde, dikkate değer ilerlemeler kaydedilmiştir (Li ve ark., 2021). Güncel araştırmalar, dikkatlice tasarlanıp eğitildiğinde, evrişimli sinir ağlarının çeşitli retina patolojilerini tespit etme ve sınıflandırma konusunda deneyimli oftalmologlarla eşit veya hatta onlardan daha üstün bir tanı performansı sergileyebildiğini göstermiştir (Gulshan ve ark., 2016; Ting ve ark., 2019). Günümüzde birçok çalışma, yalnızca ikili tespit görevlerine (örneğin, sevk edilebilir ve sevk edilemez diyabetik retinopati) değil, aynı

zamanda gerek dnyadaki klinik senaryolara daha ok benzeyen ok sınıflı sınıflandırma problemlerine de odaklanmaktadır (Zhou ve ark., 2022).

Son literatrdeki nemli bir eęilim, ImageNet gibi byk lekli genel veri kmeleri zerinde nceden eęitilmiř modellerin, daha kk, hastalıęa zg veri kmeleri zerinde ince ayarlandıęı transfer ęreniminin uygulanmasıdır. rneęin, Li ve ark. (2021), ImageNet zerinde nceden eęitilmiř bir ResNet50 mimarisi kullanmıř ve bunu retina hastalıklarının ok sınıflı sınıflandırması iin ince ayarlamıř, ODIR-5K veri setinde %85,7 doęruluk elde etmiřtir. Benzer řekilde, Wang ve ark. (2022) transfer ęrenimi ile DenseNet121'i uygulayarak, altı retina hastalıęı kategorisini sınıflandırmak iin %89,6 test doęruluęu ve %88,7 makro F1 puanı elde ettiklerini bildirmiřtir. Bu alıřmalar, sınırlı sayıda anotlanmış tıbbi verinin yarattıęı zorluęun stesinden gelmede transfer ęreniminin etkinlięini vurgulamakta ve modellerin tıbbi olmayan grntlerden ęrenilen zengin grsel zelliklerden yararlanmasını saęlamaktadır.

Veri artırma da veri kmesi dengesizlięini gidermek ve genellemeyi iyileřtirmek iin nemli bir teknik olarak ortaya ıkmıřtır. Modern artırma stratejileri genellikle basit evirme ve dndrme iřlemlerinin tesine geerek renk titremesi, elastik deformasyonlar, karıřtırma ve CutMix'i de iermektedir (Zhang ve ark., 2018; Yun ve ark., 2019). Shamsi ve ark. (2021), geliřmiř artırma yntemlerini EfficientNet tabanlı modellerle entegre etmenin, ok sınıflı retina hastalıęı grevlerinde hem doęruluk hem de saęlamlık aısından nemli kazanımlar saęladıęını gstermiřtir. Artırma, eęitim rneklerinin eřitlilięini artırmakla kalmaz, aynı zamanda fundus grntlerindeki yksek sınıf ii deęiřkenlik gz nne alındıęında kritik bir husus olan ařırı uyum riskini de azaltır (Lim ve ark., 2022).

Bir bařka nemli arařtırma alanı, modelin yorumlanabilirlięini ve performansını artırmak iin dikkat mekanizmalarının entegrasyonuna odaklanmaktadır. Squeeze-and-Excitation (SE) blokları (Hu ve ark., 2018) ve Convolutional Block Attention Modules (CBAM) (Woo ve ark., 2018) gibi dikkat modlleri, aęın zellik haritalarını dinamik olarak yeniden kalibre etmesine olanak tanıyarak grntnn tanısai aıdan ilgili blgelerini vurgular. Liu ve ark. (2023), CBAM ile glendirilmiř ResNet'i fundus grnt sınıflandırmasına uygulayarak,

standart CNN'lere kıyasla hem daha yüksek doğruluk hem de hastalığa özgü lezyonların daha iyi lokalizasyonu elde etmiştir.

Ayrıca, son zamanlarda yapılan çalışmalar, veri kümeleri arası doğrulama, çok merkezli veri kümeleri ve prospektif klinik çalışmalar kullanarak değerlendirme kapsamını genişletmiştir. Bu eğilim, yapay zeka (AI) modellerinin farklı popülasyonlardan, cihazlardan ve gerçek dünya klinik iş akışlarından elde edilen görüntülere etkili bir şekilde genelleştirilebileceğini gösterme ihtiyacından kaynaklanmaktadır (De Fauw ve ark., 2018; Yim ve ark., 2020). Birkaç çalışma, tek kaynaklı veri kümeleri üzerinde eğitilmiş modellerin sınırlamalarını vurgulamış ve yeni merkezlerden veya farklı edinim protokolleriyle elde edilen görüntüler üzerinde test edildiğinde doğrulukta önemli düşüşler olduğunu belirtmiştir (Gupta ve ark., 2021).

Son olarak, yorumlanabilirlik ve klinik kullanılabilirlik önemli vurgu alanları olmaya devam etmektedir. Son zamanlarda yapılan çalışmalar, klinisyenlere model tahminleri için sezgisel açıklamalar sunmak amacıyla Grad-CAM (Selvaraju ve ark., 2017) ve saliency maps gibi görselleştirme tekniklerini entegre etmiştir. Bu şeffaflık, klinik güveni kolaylaştırmakla kalmaz, aynı zamanda potansiyel arıza modlarının tanımlanmasına da yardımcı olur, böylece hasta güvenliğini ve düzenleyici kurumların kabul etme olasılığını artırır (Tjoa ve Guan, 2020).

Toplu olarak, bu gelişmeler, retina hastalığı tespiti için klinik olarak kullanılabilir AI sistemlerine doğru istikrarlı bir şekilde ilerleyen, olgunlaşan bir alanı yansıtmaktadır. Veri küresasyonu, model tasarımı, artırma stratejileri ve değerlendirme protokollerinde devam eden ilerlemeler, oftalmolojide derin öğrenmenin tam potansiyelini gerçekleştirmek için çok önemli olacaktır.

2.6. Retina Görüntüleme için Derin Öğrenmede Artırma ve Düzenleme

Veri artırma, retina görüntü analizinde derin öğrenme modellerinin genelleştirilebilirliğini ve sağlamlığını iyileştirmek için önemli bir strateji olarak ortaya çıkmıştır. Annotasyonlu tıbbi görüntü veri kümeleri genellikle boyut olarak sınırlıdır ve önemli sınıf dengesizliği gösterir, bu nedenle artırma teknikleri rutin olarak eğitim verilerini sentetik olarak genişletmek ve

aşırı uyumu azaltmak için kullanılır (Shorten ve Khoshgoftaar, 2019). Yatay ve dikey çevirme, rastgele döndürme, çevirme ve parlaklık ayarlamaları gibi geleneksel artırma yöntemleri, klinik görüntülerde karşılaşılan doğal varyasyonları simüle etmek için uzun süredir kullanılmaktadır (Perez ve Wang, 2017). Bu temel dönüşümler, eğitim örneklerinin çeşitliliğini artırarak, evrişimli sinir ağlarının daha değişmez ve genelleştirilebilir özellikleri öğrenmesini sağlar (Wang ve ark., 2022).

Son yıllarda, model performansını daha da artırmak için daha sofistike artırma stratejileri geliştirilmiştir. Elastik deformasyonlar, rastgele silme, mixup ve CutMix gibi teknikler, yeni ve makul görüntü varyasyonları oluşturarak hem doğruluğu hem de sağlamlığı artırma yeteneğini göstermiştir (Zhang ve ark., 2018; Yun ve ark., 2019). Örneğin, mixup yaklaşımı, görüntü çiftlerini ve bunlara karşılık gelen etiketleri doğrusal olarak birleştirerek yeni eğitim örnekleri oluşturur ve modeli sınıflar arasında enterpolasyon yapmaya etkili bir şekilde teşvik eder (Zhang ve ark., 2018). Benzer şekilde, CutMix bir görüntünün rastgele parçalarını başka bir görüntünün parçalarıyla değiştirerek modelin daha geniş bir ayırmacı bölge yelpazesine dikkatini çekmesini sağlar (Yun ve ark., 2019). Bu gelişmiş yöntemler, sınıf içi değişkenliğin yüksek ve sınıflar arası benzerliğin yüksek olduğu tıbbi görüntüleme alanlarında özellikle etkili olduklarını kanıtlamıştır.

Düzenleme teknikleri, derin sinir ağlarıyla ilişkili aşırı uyum sorununu ele almada da aynı derecede önemlidir. Eğitim sırasında nöronları rastgele devre dışı bırakan dropout, özellik algılayıcılarının birlikte uyumlaşmasını önlemek ve daha dağıtılmış temsilleri teşvik etmek için yaygın olarak kullanılmaktadır (Srivastava ve ark., 2014). Ağırlık zayıflaması (L2 düzenleme), büyük ağırlıkları cezalandıran ve daha düzgün ve basit modelleri teşvik eden başka bir standart yaklaşımdır (Goodfellow ve ark., 2016). Her katmana gelen girdileri normalleştiren toplu normalleştirme de eğitimi stabilize ettiği ve yakınsamayı hızlandırdığı gösterilmiştir (Ioffe ve Szegedy, 2015). Retina hastalığı sınıflandırması bağlamında, artırma ve düzenlemenin birlikte kullanılması hem çapraz doğrulama hem de harici test performansında önemli iyileşmeler sağladığı gösterilmiştir (Lim ve ark., 2022; Shamsi ve ark., 2021).

Veri artırmada ortaya çıkan yeni trendler, gerçekçi sentetik retina görüntüleri oluşturmak için üretken karşıt ağların (GAN) kullanılmasını içerir ve bu da eğitim veri kümelerinin çeşitliliğini daha da zenginleştirir (Zhu ve ark., 2022). Bu tür teknikler, nadir hastalık sınıfları için anotanmış görüntülerin azlığı sorununu çözme konusunda umut vaat etmektedir, ancak sentetik görüntülerin klinik olarak ilgili özellikleri doğru bir şekilde yansıtması ve artefaktlar oluşturmaması için özen gösterilmesi gerekmektedir (Frid-Adar ve ark., 2018).

Sonuç olarak, artırma ve düzenleme stratejilerinin dikkatli bir şekilde tasarlanması ve uygulanması, retina görüntülemeye derin öğrenme modellerinin başarısı için hayati önem taşımaktadır. Bu yaklaşımlar, genellemeyi iyileştirmek ve aşırı uyumu azaltmakla kalmaz, aynı zamanda AI tabanlı tanı sistemlerinin güvenilirliğini ve klinik uygulanabilirliğini de artırır.

2.7. Retina Hastalıklarının Sınıflandırılmasında Transfer Öğrenimi ve Genelleme

Transfer öğrenimi, tıbbi görüntü analizinde, özellikle de etiketli veri kümelerinin boyutu ve çeşitliliği genellikle sınırlı olan retina hastalıklarının sınıflandırılmasında vazgeçilmez bir yaklaşım haline gelmiştir (Shin ve ark., 2016). Transfer öğreniminin temel fikri, ImageNet gibi büyük ölçekli, genel amaçlı veri kümelerinden elde edilen bilgileri, derin konvolüsyonel sinir ağlarını önceden eğiterek ve ardından bunları hedef tıbbi görevlere göre ince ayar yaparak kullanmaktır (Raghu ve ark., 2019). Bu strateji, modellerin kenar ve doku algılayıcıları gibi öğrenilmiş görsel temsilleri doğal görüntülerden fundus fotoğrafçılığı veya OCT görüntüleme gibi daha özel alanlara aktarmasını sağlar.

Çok sayıda çalışma, transfer öğreniminin retina hastalığı tespitinde hem doğruluğu hem de yakınsama hızını artırmada etkili olduğunu göstermiştir. Örneğin, Gulshan ve ark. (2016), büyük bir diyabetik retinopati veri seti üzerinde ince ayar yapılmış, önceden eğitilmiş bir InceptionV3 modeli kullanmış ve kurul sertifikalı oftalmologlarla karşılaştırılabilir bir tanı performansı elde etmiştir. Benzer şekilde, Li ve ark. (2021), ODIR-5K çok sınıflı retina hastalığı veri seti üzerinde bir ResNet50 ağını ince ayarlamış ve sıfırdan eğitilmiş modellere kıyasla sınıflandırma doğruluğunda önemli bir artış elde etmiştir.

Transfer öğreniminin temel bir avantajı, tıbbi görüntülemelerde yaygın olan veri kıtlığı ve sınıf dengesizliği sorunlarının etkilerini, model parametrelerini genel görsel özelliklere zaten uyarlanmış bir ağırlık uzayında başlatarak azaltma yeteneğidir (Shin ve ark., 2016; Raghu ve ark., 2019). Bu ön koşullandırma, yakınsamayı hızlandırmakla kalmaz, aynı zamanda modelin tüm özellik algılayıcılarını sıfırdan öğrenmesine gerek kalmadığı için küçük veri kümelerine aşırı uyum sağlama riskini de azaltır. Araştırmalar, ince ayarlı modellerin, özellikle sınırlı sayıda anotanmış örnek mevcut olduğunda, rastgele başlatılan ağlardan tutarlı bir şekilde daha iyi performans gösterdiğini ortaya koymuştur (Tajbakhsh ve ark., 2016).

Genelleme, yani bir modelin görülmemiş veriler üzerinde iyi performans gösterme yeteneği, transfer öğreniminin bir başka önemli avantajıdır. Önceden eğitilmiş ağlar, çok çeşitli görüntü yapılarına maruz kalmışlardır ve bu da görüntü kalitesi, görüntü alma cihazları ve hasta demografisi değişikliklerine karşı dayanıklılıklarını artırmaktadır (Esteva ve ark., 2017). Bu, gerçek dünya klinik verilerinin görüntüleme protokolleri, popülasyon çeşitliliği ve hastalık yaygınlığı açısından araştırma veri kümelerinden önemli ölçüde farklılık gösterdiği retina hastalığı taramalarında özellikle önemlidir (Gupta ve ark., 2021). Modellerin bir modalitede (ör. doğal görüntüler) önceden eğitildiği ve başka bir modalitede (ör. fundus görüntüleri) ince ayarlandığı çapraz alan transfer öğreniminin, çok merkezli ve prospektif doğrulama çalışmalarında genelleştirilebilirliği iyileştirdiği gösterilmiştir (De Fauw ve ark., 2018).

Bununla birlikte, transfer öğreniminin sınırlamaları da vardır. Kaynak ve hedef veri dağılımları arasındaki uyumsuzluk olan alan kayması, önceden eğitilmiş özellikler söz konusu tıbbi görüntüleme göreviyle yeterince alakalı değilse performansı olumsuz etkileyebilir (Raghu ve ark., 2019). Bu sorunu çözmek için bazı araştırmacılar, alan uyumlu transfer öğrenme tekniklerini veya mümkün olduğunda CheXpert veya MIMIC-CXR gibi tıbbi görüntüleme veri kümelerini kaynak alanlar olarak kullanmayı savunmaktadır (Irvin ve ark., 2019; Johnson ve ark., 2019). Ek olarak, hangi katmanların ince ayarlanacağına ilişkin en uygun seçim, ara denetim kullanımı ve öğrenme oranlarının kalibrasyonu, halen aktif araştırma alanları olmaya devam etmektedir (Cui ve ark., 2019).

Genel olarak, transfer öğrenimi retina hastalığı sınıflandırmasında derin öğrenimde devrim yaratmış ve sınırlı sayıda anotlanmış veriyle bile en son teknolojiye sahip performans elde edilmesini sağlamıştır. Daha kapsamlı ve çeşitli tıbbi veri kümeleri kullanıma sunuldukça ve alan uyarılma teknikleri gelişmeye devam ettikçe, AI modellerinin popülasyonlar ve ortamlar arasında genelleme yapma yeteneği daha da gelişecek ve klinik uygulamada otomatik retina hastalığı taramasının benimsenmesi hızlanacaktır.

2.8. Retina Hastalığı Sınıflandırmasında Dikkat Mekanizmaları ve Açıklanabilirlik

Retina hastalığı sınıflandırması için derin öğrenme modellerinin karmaşıklığı ve doğruluğu arttıkça, hem performansın hem de yorumlanabilirliğin iyileştirilmesi ihtiyacı giderek daha belirgin hale gelmiştir. Dikkat mekanizmaları, her iki zorluğa da güçlü bir çözüm olarak ortaya çıkmış ve sinir ağlarının girdi görüntülerinin en alakalı bölgelerine seçici olarak odaklanmasını sağlarken, aynı zamanda klinik kullanıcılar için daha şeffaf ve yorumlanabilir çıktılar sunmuştur (Vaswani ve ark., 2017; Hu ve ark., 2018).

Dikkat mekanizmalarının temel fikri, modelin bir görüntüdeki farklı uzamsal veya özellik kanallarına değişen önem düzeyleri atamasını sağlamaktır. Retina görüntü analizi bağlamında, dikkat mekanizması modelin mikroanevrizmalar, eksüdalar veya optik disk gibi tanısal açıdan önemli yapıları vurgulamasına ve daha az ilgili arka plan bilgilerini bastırmasına olanak tanır (Woo ve ark., 2018). Örneğin, Squeeze-and-Excitation (SE) blokları, kanal bazında özellik tepkilerini yeniden kalibre ederek bilgilendirici özellikleri dinamik olarak geliştirir (Hu ve ark., 2018). Convolutional Block Attention Modules (CBAM), hem uzamsal hem de kanal dikkatini entegre ederek bu kavramı genişletir ve ağı daha belirgin görüntü bölgelerine daha hassas bir şekilde odaklanmasını sağlar (Woo ve ark., 2018). Liu ve ark. (2023), CBAM'yi fundus görüntü sınıflandırması için ResNet tabanlı bir modele entegre etmenin sadece doğruluğu değil, aynı zamanda hastalığa özgü lezyonların lokalizasyonunu da iyileştirdiğini ve uzmanların açıklamalarıyla yakından uyumlu ısı haritaları sağladığını göstermiştir.

Bir başka dönüştürücü gelişme, başlangıçta doğal dil işlemede tanıtılan (Vaswani ve ark., 2017) ancak daha sonra tıbbi görüntülemede önemli uygulamalar bulan dönüştürücü tabanlı mimariler ve öz dikkat mekanizmalarının benimsenmesidir (Dosovitskiy ve ark., 2021).

Görsel dönüştürücüler (ViT'ler) görüntüleri yama dizileri olarak işler ve uzamsal ilişkilerin esnek ve uzun menzilli modellemesine olanak tanır. Son çalışmalar, dönüştürücü tabanlı modellerin diyabetik retinopati tespiti dahil olmak üzere belirli tıbbi görüntüleme görevlerinde CNN'lerin performansını eşleştirebileceğini ve hatta aşabileceğini göstermiştir (Chen ve ark., 2022).

Yorumlanabilirlik, oftalmolojide AI sistemlerinin klinik uygulaması için çok önemli bir gerekliliktir. Klinisyenler, özellikle tanı ve tedavi kararlarının alındığı yüksek riskli senaryolarda model tahminlerini anlayabilmeli ve bunlara güvenebilmelidir (Tjoa ve Guan, 2020). Bu amaçla, saliency map, Grad-CAM ve attention heatmap gibi açıklanabilir AI (XAI) yöntemleri, modelin tahminlerine en fazla katkıda bulunan giriş görüntülerinin bölgelerini görselleştirmek için standart araçlar haline gelmiştir (Selvaraju ve ark., 2017). Bu görselleştirmeler, klinisyenlere sezgisel açıklamalar sunarak model kararlarının doğrulanmasına yardımcı olur ve hataların veya başarısızlık durumlarının tespitini kolaylaştırır (Montavon ve ark., 2018).

Son zamanlarda yapılan çalışmalar, dikkat tabanlı açıklamaların doğrudan klinik iş akışlarına entegrasyonunu da araştırmıştır. Örneğin, Jin ve ark. (2021), her tahmin için hem sınıflandırma çıktıları hem de yorumlanabilir ısı haritaları sağlayan, doktorların güvenini ve kullanılabilirliği artıran bir retina hastalığı tespiti için derin öğrenme sistemi geliştirmiştir. Diğer çalışmalar, retina hastalığı veri kümelerindeki açıklama hatalarını belirlemek ve düzeltmek için dikkat haritalarını kullanmış ve dikkat mekanizmalarının hem model iyileştirme hem de veri kalite kontrolü için ikili yararını vurgulamıştır (Yang ve ark., 2022).

Bu ilerlemelere rağmen, dikkat temelli açıklamaları standartlaştırmak ve çeşitli klinik ortamlarda güvenilirliğini sağlamak konusunda zorluklar devam etmektedir (Tjoa ve Guan, 2020). Bu alandaki araştırmaların devam ettirilmesi, en son teknolojiye sahip AI modelleri ile gerçek dünyadaki oftalmik uygulamalar arasındaki uçurumu kapatmak ve otomatik retina hastalığı tespit sistemlerinin hem doğru hem de yorumlanabilir olmasını sağlamak için gereklidir.

3. MALZEMELER VE METODOLOJİ

Bu çalışmada, fundus görüntülerine dayalı çok sınıflı retina hastalığı sınıflandırması için özel bir derin öğrenme boru hattı tasarlanmış ve uygulanmıştır. Tıbbi görüntü analizinde kanıtlanmış verimliliği ve etkinliği nedeniyle EfficientNetB0'ı birincil mimari olarak seçilmiştir. Model oluşturma, eğitim ve değerlendirme adımlarını otomatikleştirmek amacıyla özel Python komut dosyaları geliştirilmiştir.

Modelin öğrenme kapasitesini optimize etmek ve onu retina hastalığının benzersiz modellerine özel olarak uyarlamak için ince ayar uygulanmıştır. EfficientNetB0 mimarisinin tüm katmanları çözülerek, ImageNet üzerinde önceden eğitilmiş ağırlıklarla model, alanımıza özgü veri seti üzerinde yeniden eğitilmiştir. Bu yaklaşım, doğru hastalık sınıflandırması için kritik öneme sahip hem düşük hem de yüksek düzey özellikleri çıkartılmasını sağlamıştır.

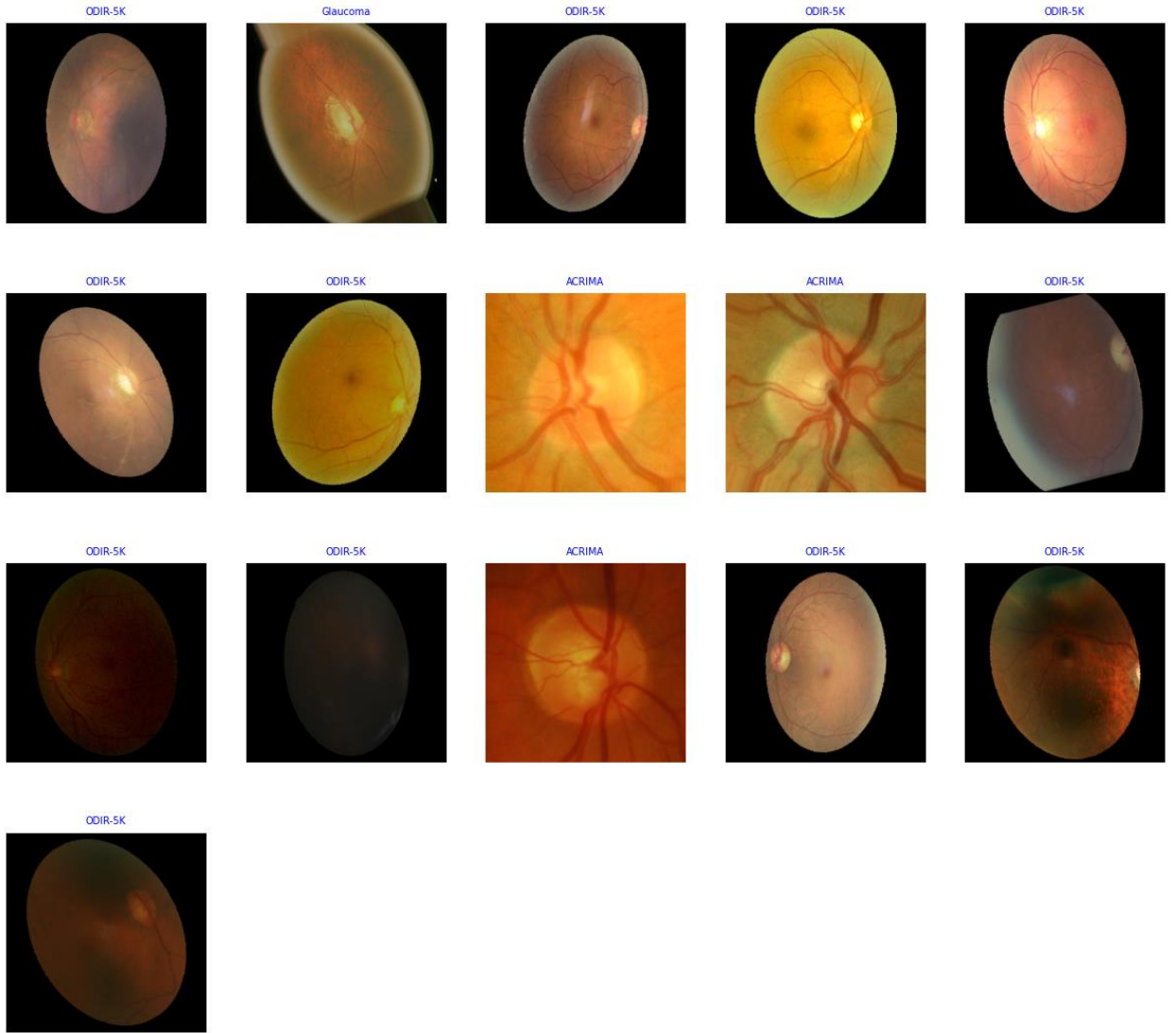
Veri çeşitliliğinin etkisini titizlikle değerlendirmek için iki farklı senaryo altında deneyler yapılmıştır. İlkinde, değişkenliği artırmak ve aşırı uyum riskini azaltmak için eğitim görüntüleri, rastgele döndürme, çevirme, parlaklık değişiklikleri ve yakınlaştırma gibi veri artırma teknikleriyle çeşitlendirilmiştir. İkinci senaryoda, sınırlı veri çeşitliliğinin model genelleştirme üzerindeki etkisini göstermek için temel olarak hiçbir artırma yapılmamıştır.

Süreç boyunca, veri ön işleme, bölme, model eğitimi, doğrulama ve metrik hesaplama süreçleri sistematik ve tekrarlanabilir şekilde yapılandırılmıştır. Bu metodoloji, gerçekçi klinik veri koşulları altında EfficientNetB0'ın güçlü ve zayıf yönlerini adil bir şekilde değerlendirilmesini ve tıbbi görüntülemelerde derin öğrenme performansı üzerinde artırma ve ince ayarın etkisine ilişkin anlamlı sonuçlar çıkartılmasını sağlamıştır.

3.1. Veri Kümesi Açıklaması

Bu tez kapsamında, altı ana göz hastalığı kategorisini içeren ve çeşitli açık kaynaklı veri setlerinden derlenen kapsamlı bir retina görüntü koleksiyonu kullanılmıştır. Kullanılan veri setleri; ACRIMA, Cataract, Glaucoma, ODIR-5K, ORIGA ve Retina_Disease olarak sıralanmaktadır. Bu veri kaynakları, farklı cihazlar ve görüntüleme koşullarında elde edilmiş fundus fotoğraflarını içermekte olup, sınıflandırma görevinde genelleştirilebilir ve dengeli

bir model geliřtirmek amacıyla birleřtirilmiřtir. Söz konusu veri setleri; Glaucoma (glokom), Cataract (katarakt), Age-Related Macular Degeneration (yařa baėlı makula dejenerasyonu), Diabetic Retinopathy (diyabetik retinopati), genel retina hastalıkları (retina_disease) ve normal retina olmak üzere altı farklı sınıfı kapsamaktadır. Bu hastalık kategorilerine ait örnek fundus görüntüleri řekil 3.1’de sunulmuřtur. Görseller, her bir hastalığın tipik görsel özelliklerini temsil etmekte ve sınıflar arası morfolojik farklılıkları ortaya koymaktadır. Bařlangıçta, veri kümesinde ciddi bir sınıf dengesizliėi gözlemlenmiřtir; bazı hastalık sınıfları, diėerlerine kıyasla oldukça düşük sayıda örnek içermektedir. Bu durumu dengelemek amacıyla veri kümesi örnekleme ve artırma teknikleriyle geniřletilmiřtir. Sonrasında tüm veri kümesi, %70 eėitim ve %30 test oranıyla tabakalı řekilde bölünmüřtür. Bu yapılandırma sonucunda, eėitim seti 14.000 ve test seti 6.000 görüntüden oluřacak biçimde hazırlanmıř; test kümesinde her sınıfın 1.000 görüntü ile eřit temsil edilmesi saėlanmıřtır. Böylece, modelin genelleme yeteneėi tüm sınıflarda dengeli olarak deėerlendirilebilmiř, sınıfa dayalı önyargı riski azaltılarak anlamlı performans ölçütleri elde edilmiřtir.



Şekil 0.1: Çalışmada kullanılan ACRIMA, ODOR-5K ve Glaucoma veri setlerinden örnek retina fundus görüntüleri. Görseller, sınıflar arası çeşitliliği ve patolojik özellikleri yansıtmaktadır

3.2. Veri Artırımı

Bu araştırmada kullanılan veri kümesi, altı ana göz hastalığı kategorisini temsil etmek amacıyla, retina fundus görüntülerini içeren birkaç açık kaynaklı veri deposundan derlenmiştir. Başlangıçta, veri kümesi oldukça dengesizdi, çünkü bazı hastalık sınıfları diğerlerine kıyasla önemli ölçüde yetersiz temsil ediliyordu. Veri çeşitliliğini artırmak ve adil bir değerlendirme ortamı sağlamak amacıyla kapsamlı bir ön işleme ve dengeleme yaklaşımı

izlenmiştir. Bu çabaların ardından, veri kümesi genişletilmiş ve yeniden yapılandırılmıştır, böylece her hastalık sınıfı eşit şekilde temsil edilen toplam 20.000 görüntü içerecek şekilde düzenlenmiştir. Tüm görüntüler kapsamlı kalite kontrollerinden geçirilmiş ve derin öğrenme analizi için uygun standart bir çözünürlüğe yeniden boyutlandırılmış, böylece tüm koleksiyonun format ve görünümünde tutarlılık sağlanmıştır.

Model eğitimi ve değerlendirmesi için verileri hazırlamak amacıyla, veri setini 70/30 oranına göre eğitim ve test setlerine ayrılmıştır. Böylece, 14.000 görüntü eğitim için, 6.000 görüntü ise test için ayrılmıştır. Test setinin her hastalık kategorisinden tam olarak 1.000 görüntü içermesine özel dikkat gösterilmiş, böylece tüm sınıflar genelinde model performansının dengeli ve güvenilir bir şekilde değerlendirilmesi sağlanmıştır. Bu tabakalı bölünme, sınıfların eşit temsiliyi sağlamış ve hem öğrenme hem de değerlendirme aşamalarında önyargı riskini en aza indirmiştir.

Derin öğrenme başarısı açısından veri çeşitliliğinin önemine dayanarak, eğitim sürecine çeşitli veri artırma teknikleri entegre edilmiştir. Rastgele döndürme, çevirme, yakınlaştırma, parlaklık ve kontrast ayarlamaları ile rastgele çeviriler ve kırpmalar uygulayarak, eğitim kümesindeki değişkenliği yapay olarak artırabilmiştir. Bu artırma işlemleri, eğitim sırasında çevrim içi ve dinamik biçimde uygulanmıştır, böylece model her dönemde yeni ve çeşitli görüntü varyasyonlarıyla karşılaşmıştır. Bu yaklaşım, modelin görülmemiş verilere genelleme yeteneğini güçlendirmekle kalmamış, aynı zamanda özellikle sınırlı örneklem boyutlarından muzdarip olan hastalık kategorileri için aşırı uyumun azaltılmasında da önemli bir rol oynamıştır. Sonuç olarak, bu birleşik veri hazırlama ve artırma stratejileri, bu çalışmada sağlam model geliştirme ve adil deneysel karşılaştırma için temel oluşturmuştur. Şekil 3.1 kullanılan veri setlerine ait örnek görüntülerin veri çeşitliliğini göstermektedir.

3.3. Veri Ön İşleme

Derin öğrenme modeline veri sağlamadan önce, giriş verilerinin kalitesini artırmak ve tutarlılığı sağlamak amacıyla ayrıntılı bir ön işleme süreci izlenmiştir. İlk adımda, tüm görüntüler kalite açısından değerlendirilmiş; bulanık, aşırı pozlanmış ya da artefakt içerenler veri setinden çıkarılmıştır. Bu kalite güvence adımı, gürültüyü en aza indirmek ve modelin yanıltıcı kalıpları öğrenmesini önlemek için çok önemlidir.

Kalite kontrolünü geçen görüntüler, modelin gereksinimlerine uygun olarak 224×224 piksel çözünürlüğe yeniden boyutlandırılmıştır. Bu hedef boyut hem hesaplama verimliliği hem de EfficientNetB0 mimarisinin giriş gereksinimlerine uymak için seçilmiştir. Boyutlandırma işlemi sırasında, en-boy oranları korunmuş ve retinanın yapısal bütünlüğü sağlamak amacıyla gerekli durumlarda görüntüler ortalananarak kırpılmıştır.

Renk standardizasyonu, tüm veri setine tutarlı biçimde uygulanmıştır. Görüntülerin tamamı RGB renk uzayına dönüştürülerek, renk bilgisi açısından bütünlük sağlanmıştır, böylece hastalık sınıflarını ayırt etmek için genellikle hayati önem taşıyan önemli renk bilgilerinin korunması sağlanmıştır. Gri tonlamalı veya anormal kanallara sahip olduğu tespit edilen görüntüler, veri setinin genel bütünlüğünü korumak için düzeltilmiş veya hariç tutulmuştur.

Model eğitiminin kararlılığını ve hızını daha da artırmak için, tüm görüntülerdeki piksel değerleri, her piksel 255'e bölünerek $[0, 1]$ aralığına normalleştirilmiştir. Bu normalleştirme adımı, öğrenme sürecini stabilize etmeye yardımcı olmuş ve optimizasyon sırasında daha iyi yakınsama sağlamıştır.

Hem eğitim hem de test veri kümelerine tamamen aynı ön işleme adımlarını uygulamak oldukça önemlidir. Bu işlem, iki veri bölümü arasında istenmeyen tutarsızlıkları ortadan kaldırmış ve modelin performansının adil ve tarafsız bir şekilde değerlendirilmesini sağlamıştır. Bu sistematik ön işleme süreci sayesinde, yapılan deneylerde kullanılan her görüntünün yüksek kaliteli, doğru biçimde biçimlendirilmiş ve etkili derin öğrenme analizi için uygun olması sağlanmıştır.

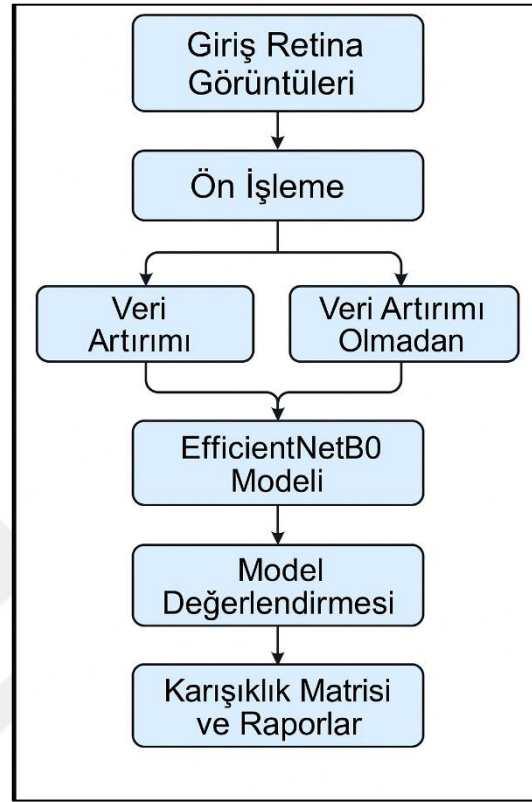
3.4. Model Mimarisi

Bu çalışmada, retina hastalığı sınıflandırması için EfficientNetB0 derin öğrenme mimarisi tercih edilmiştir. Bu mimari, yüksek doğruluk düzeyi ile hesaplama verimliliği arasındaki dengeyi sağlayarak tıbbi görüntü analizi gibi hassas uygulamalar için ideal bir yapı sunmaktadır. Ağın bileşik ölçekleme stratejisi sayesinde model, derinlik, genişlik ve giriş çözünürlüğünü eşzamanlı şekilde ayarlayarak fundus görüntülerinde karşılaşılan çeşitli yapısal özelliklere duyarlılık gösterebilmektedir.

Genel iş akışı, giriş fundus görüntülerinin toplanmasıyla başlayıp, ön işleme süreciyle devam etmektedir. Ardından veri kümesi, iki farklı kola ayrılmaktadır: biri veri artırma uygulanan, diğeri ise ham verilerle çalışılan yoldur. Her iki yol da EfficientNetB0 modeli ile eğitilmekte ve sınıflandırma sonuçları, detaylı metrikler ve karışıklık matrisleri ile değerlendirilmektedir. Bu akış, Şekil 3.2'de şematik olarak sunulmuştur.

Modelin performansını artırmak amacıyla transfer öğrenme ve ince ayar teknikleri uygulanmıştır. EfficientNetB0, ImageNet veri kümesi üzerinde önceden eğitilmiş ağırlıklarla başlatılmış; ardından tüm katmanlar açılarak, fundus görüntülerine özgü özelliklerin öğrenilmesine olanak sağlanmıştır. Bu sayede, hem düşük seviyeli kenar ve doku özellikleri hem de yüksek seviyeli patolojik desenler model tarafından etkili biçimde öğrenilmiştir.

Modelin özgün sınıflandırma katmanı çıkarılmış ve yerine altı nöronlu, softmax aktivasyonlu yeni bir fully connected katman eklenmiştir. Bu yapı, her bir nöronun bir hastalık sınıfını temsil etmesini sağlamaktadır. Ayrıca, aşırı öğrenmenin önüne geçmek amacıyla mimariye dropout katmanları ve batch normalization bileşenleri entegre edilmiştir. Bu yapılandırma, hem veri artırımı yapılan hem de yapılmayan senaryolarda tutarlı biçimde kullanılmıştır.



Şekil 0.2: Önerilen Retina Hastalığı Sınıflandırma Sisteminin Genel İş Akışı

3.4.1. Veri Artırımı ile EfficientNetB0

İlk deneysel senaryoda, EfficientNetB0 mimarisi kapsamlı bir veri artırma stratejisiyle birlikte uygulanmıştır. Model, ImageNet veri kümesi üzerinde önceden eğitilmiş ağırlıklarla başlatılmış; tüm katmanları açılarak, fundus görüntülerine özgü özelliklerin öğrenilmesine olanak tanınmıştır. Bu şekilde, model yalnızca sınıflandırma katmanında değil, ağıın tüm seviyelerinde görev odaklı öğrenme gerçekleştirebilmiştir.

Eğitim sürecinde, veri çeşitliliğini artırmak amacıyla kapsamlı bir artırma paketi kullanılmıştır. Bu paket, rastgele döndürme, yatay ve dikey çevirme, parlaklık ve kontrast değişiklikleri, yakınlaştırma, çevirme ve kırpma gibi dönüşümleri içermektedir. Söz konusu dönüşümler, her epoch sırasında çevrim içi şekilde uygulanarak modelin her defasında yeni varyasyonlarla karşılaşması sağlanmıştır. Böylece model, yönelim, ölçek ve aydınlatma gibi doğal varyasyonlara karşı daha dayanıklı hale gelmiştir.

Modelin sınıflandırma başlığı, altı hastalık sınıfına karşılık gelen softmax aktivasyonlu yeni bir fully connected katmanla değiştirilmiştir. Aşırı öğrenme riskini azaltmak amacıyla dropout katmanları ile düzenleme sağlanmış ve genel mimaride batch normalization kullanılmıştır. Eğitim süreci boyunca Adam optimizasyon algoritması tercih edilmiş, öğrenme oranı 0.0001 olarak belirlenmiş ve batch boyutu 32 olacak şekilde yapılandırılmıştır. Eğitim 50 epoch boyunca gerçekleştirilmiş, en iyi doğrulama performansına sahip model ağırlıkları kaydedilmiştir.

Eğitim sırasında, doğruluk ve kayıp eğrileri yakından izlenmiştir. Şekil 4.2 ve 4.3'te gösterildiği gibi, hem eğitim hem de doğrulama doğruluğunda istikrarlı bir artış gözlemlenmiş, bu da modelin yalnızca verileri ezberlemediğini; aynı zamanda genelleme yeteneği kazandığını göstermiştir. Paralel olarak, kayıp değerlerinde sürekli bir azalma kaydedilmiş ve öğrenme sürecinin dengeli ilerlediği doğrulanmıştır.

Son değerlendirme, her sınıftan 1.000 görüntü içeren bağımsız test seti üzerinde gerçekleştirilmiştir. Tablo 4.1'de sunulan sınıflandırma raporunda, tüm sınıflarda precision, recall ve F1 skorlarının %91'in üzerinde olduğu gözlemlenmiştir. Şekil 4.1'de sunulan karışıklık matrisi ise, modelin sınıflar arasındaki ayırım gücünü açık biçimde ortaya koymaktadır. Yanlış sınıflandırmaların sayıca az ve dengeli dağılmış olması, sınıflar arasında belirgin bir önyargının oluşmadığını göstermektedir.

Genel olarak, veri artırımı ile eğitilen EfficientNetB0 modeli; sınıf dengesizliği, veri kıtlığı ve varyasyon eksikliği gibi zorluklara karşı yüksek doğruluk ve genelleme başarımı sergilemiştir. Bu sonuçlar, tıbbi görüntü sınıflandırmasında veri artırmanın vazgeçilmez bir bileşen olduğunu ve EfficientNet mimarisinin bu bağlamda oldukça uyumlu çalıştığını ortaya koymaktadır.

3.4.2. Veri Artırımı Olmadan EfficientNetB0

İkinci deneysel senaryoda, EfficientNetB0 mimarisi aynı şekilde uygulanmış; ancak eğitim sürecinde hiçbir veri artırma tekniği kullanılmamıştır. Bu yaklaşım, artırma uygulanmaksızın modelin doğal öğrenme kapasitesini değerlendirmek ve veri çeşitliliği olmadan oluşabilecek sınırlamaları gözlemlemek amacıyla yapılandırılmıştır.

Model yine ImageNet veri seti üzerinde önceden eğitilmiş ağırlıklarla başlatılmış, tüm katmanlar açılarak ince ayar yapılmıştır. Sınıflandırma katmanı, altı sınıfa karşılık gelen softmax aktivasyonlu bir fully connected katman ile değiştirilmiş, dropout katmanları eklenerek düzenleme sağlanmıştır. Tüm eğitim parametreleri optimizör (Adam), öğrenme oranı (0.0001), batch boyutu (32) ve epoch sayısı (50) artırma yapılan senaryo ile birebir aynı tutulmuştur. Bu sayede, performans farklarının yalnızca veri artırmanın varlığına dayandığı güvenilir bir şekilde analiz edilmiştir.

Model yalnızca orijinal ve değiştirilmemiş görüntüler üzerinde eğitildiğinden, çeşitli yönelimler, ölçekler ve ışık koşulları gibi gerçek dünya varyasyonlarını görme imkânı kısıtlı kalmıştır. Bu durum, modelin genelleme yeteneğini azaltmış ve belirli sınıflarda aşırı uyuma yol açabilecek bir ortam yaratmıştır.

Eğitim süreci sırasında izlenen doğruluk ve kayıp eğrileri, bu sınırlamaları açıkça ortaya koymuştur. Eğitim doğruluğu her ne kadar %96 seviyesine ulaşsa da, doğrulama doğruluğundaki artış erken dönemlerde duraklamış ve sınırlı kalmıştır. Aynı şekilde, doğrulama kaybı, belirli bir noktadan sonra yükselme eğilimi göstermiştir. Bu gözlemler, modelin eğitim verilerini ezberlediğini, ancak yeni ve görülmemiş örneklerle genelleme yapmada zorlandığını işaret etmektedir.

Şekil 4.4’de gösterilen karışıklık matrisi, bu senaryoda yapılan yanlış sınıflandırmaların, özellikle başlangıçta az sayıda örnekle temsil edilen sınıflarda yoğunlaştığını göstermektedir. Bu durum, sınıf dengesizliğinin azaltılmaması ve veri çeşitliliğinin artırılmaması halinde, modelin ayırt edici öğrenme yeteneğinin azaldığını desteklemektedir.

Sonuç olarak, bu deneysel senaryo, artırma yapılmadan eğitilen modellerin yüksek eğitim doğruluğuna rağmen genelleme konusunda sınırlı kaldığını ortaya koymaktadır. Veri artırmanın yokluğu, modelin yalnızca ezberleme temelli öğrenme gerçekleştirmesine neden olmuş ve klinik açıdan güvenilir bir sınıflandırma performansının önünde engel teşkil etmiştir. Bu durum, artırmanın özellikle sınırlı örnek içeren tıbbi veri setlerinde vazgeçilmez bir bileşen olduğunu bir kez daha kanıtlamaktadır.

3.4.3. VGG16 (Temel Karşılaştırma Modeli)

VGG16 da güçlü bir temel model olduğu için, EfficientNetB0'un kazançlarını adil biçimde ölçmek için deneylere eklenmiştir. Bunu yapmak için ImageNet üzerinde ön-eğitilmiş VGG16'yı yükledikten sonra, özgün sınıflandırma bloğunu çıkartılarak ve ACRIMA, Cataract, Glaucoma, ODIR-5K, ORIGA ve Retina_Disease sınıflarını temsil eden altı çıkışlı softmax katmanından önce Global Average Pooling ve Dropout içeren hafif bir başlık eklenmiştir. Girdi boyutunu $224 \times 224 \times 3$ ve piksel normalizasyonunu $[0,1]$ düzeyinde tutarak EfficientNet-B0 ile aynı ön-işleme işlemi korunmuştur. Tüm katmanları 50 epoch boyunca Adam (öğrenme oranı $1e-4$), kategorik çapraz entropi ve 32'lik yığın boyutu ile ince ayarlanmış; doğrulama kaybı ve doğruluğunu izleyerek, early stopping ve model checkpoint ile en iyi val_loss değerine sahip ağı saklayarak aşırı uyumu sınırlamak amacıyla bir sınır konulmuştur. Her epoch için eğitim/doğrulama kaybı ve doğruluğunu kaydetmektedir, 50 epoch'luk program boyunca eğrileri çizerek sistemin kararlı optimizasyonu doğrulanır. Eğitimden sonra en iyi VGG16 kontrol noktası test kümesi üzerinde değerlendirilmiş, argmax ile etiketleri üretilmiş, altı sınıf için sınıflandırma raporu ve karışıklık matrisi hesaplanmış ve doğruluk/kayıp eğrileri ile karışıklık matrisini EfficientNet-B0 ile aynı görsel biçimde sunularak mimari, eğitim kurulumu ve raporlama açısından doğrudan karşılaştırılabilirliği sürdürülmüştür.

3.5. Eğitim Prosedürü

- EfficientNetB0 ve VGG16 modelleri aynı eğitim protokolü altında çalıştırılarak, veri artırmanın etkisini yalıtılmış, adil bir karşılaştırma kurulmuştur. Böylece genelleme başarımı güvenilir biçimde değerlendirilmiştir.
- Her iki deneysel senaryoda (artırmalı ve artırmaz) aynı hiperparametreleri uygulanmış, modeller ImageNet üzerinde önceden eğitilmiş ağırlıklarla başlatılmıştır.
- Tüm katmanları açarak retina görüntülerine özgü ayrımların katmanlar boyunca uyarlanması sağlanmış ve sınıflandırma başlığını altı çıkışlı softmax ile aynı giriş boyutu ($224 \times 224 \times 3$) ve aynı normalizasyon aralığı $[0,1]$ altında tutularak iş hattı (pipeline) iki mimari arasında tam uyumlu tutulmuştur.

- Eğitimler 50 epoch'a kadar sürdürülmüş, Adam optimizasyonunu $1e-4$ öğrenme oranı ve 32'lik mini-batch ile yürütülmüş, kayıp fonksiyonu olarak kategorik çapraz entropi kullanılmış, aşırı uyumu sınırlamak amacıyla dropout ve gerekli yerlerde Batch Normalization uygulanmıştır.
- Her epoch sonunda doğrulama doğruluğunu ve kaybını izleyerek, art arda on epoch boyunca gelişme görülmediğinde erken durdurmayı tetikleyerek aşırı uyuma bir sınır çekilmiştir.
- Bu süreçte en iyi doğrulama başarımını veren ağırlıklar saklanmış, eğitim ve doğrulama için doğruluk/kayıp değerlerini 50 epoch zaman çizelgesi boyunca kaydederek, her iki modelin öğrenme dinamikleri aynı ölçekte raporlanmıştır.
- Eğitimler tamamlandıktan sonra en iyi denetim noktası ayrılmış test kümesi üzerinde değerlendirilmiş, argmax ile nihai etiketleri üretilmiş, sınıflandırma raporu ve altı sınıf için karışıklık matrisi hesaplanarak, hem doğruluk kayıp eğrilerini hem de karışıklık matrisini tekdüze bir görsel şablonda sunarak EfficientNetB0 ve VGG16'nın mimari farklarının performansa yansımaları doğrudan karşılaştırılabilir hale getirilmiştir.

3.6. Değerlendirme Metrikleri

Eğitilmiş modellerin performansını kapsamlı bir şekilde değerlendirmek için, çok sınıflı sınıflandırma problemlerine uygun bir dizi iyi bilinen değerlendirme metriği kullanılmıştır.

Öncelikle, her sınıf için doğru ve yanlış tahminlerin dağılımını görselleştirmek, yanlış sınıflandırma modellerini vurgulamak ve model ayrımında potansiyel zayıflıkları ortaya çıkarmak için bir karışıklık matrisi oluşturulmuştur. Şekil 3.3'te gösterilen alanlar:

TP: Pozitif örneğin doğru sınıflandırılması işlemidir.

TN: Negatif örneğin doğru sınıflandırılması işlemidir.

FP: Negatif örneğin yanlış sınıflandırılması işlemidir.

FN: Pozitif örneğin yanlış sınıflandırılması işlemidir.

KARIŞIKLIK MATRİSİ		TAHMİN EDİLEN DEĞER	
		POZİTİF	NEGATİF
GERÇEK DEĞER	POZİTİF	DOĞRU POZİTİF (TP)	YANLIŞ POZİTİF (FP)
	NEGATİF	YANLIŞ NEGATİF (FN)	DOĞRU NEGATİF (TN)

Şekil 0.3: Karışıklık matrisi yapısı

Sınıflar arasında örneklerin dağılımını açıklayan genel bir özet sağlamak için Doğruluk (Accuracy) , Kesinlik ya da hassasiyet (Precision) , Duyarlılık ya da geri çağırma (Recall) ve F1 Puanının makro ve ağırlıklı ortalamaları da rapor edilmiştir. Tüm metriklerin formülleri Şekil 3.4’te verilmektedir.

$$\text{Doğruluk} = \frac{\text{Doğru Pozitif} + \text{Doğru Negatif}}{\text{Toplam Örnek Sayısı}}$$

$$\text{Kesinlik} = \frac{\text{Doğru Pozitif}}{\text{Doğru Pozitif} + \text{Yanlış Pozitif}}$$

$$\text{Duyarlılık} = \frac{\text{Doğru Pozitif}}{\text{Doğru Pozitif} + \text{Yanlış Negatif}}$$

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

Şekil 0.4: Doğruluk, Kesinlik, Duyarlılık ve F1 Skoru formülleri

Raporlanan birincil metrik, test setindeki tüm örnekler arasında doğru sınıflandırılan görüntülerin oranı olarak tanımlanan genel doğruluktur. Doğruluk, performansın genel bir ölçüsünü sağlasa da sınıf dengesizliği durumunda modelin davranışını tam olarak yansıtmayabilir.

Bunu gidermek amacıyla, her hastalık kategorisi için hassasiyet, geri çağırma ve F1 puanı da hesaplanmıştır. Hassasiyet, belirli bir sınıf olarak tahmin edilen tüm örnekler arasında doğru pozitif tahminlerin oranını ölçerken; geri çağırma, o sınıfın tüm gerçek örnekleri arasında doğru olarak tanımlanan doğru pozitiflerin oranını ifade eder. Hassasiyet ve geri çağırmanın harmonik ortalamasını temsil eden F1 puanı ise her iki hata türünün dengeli bir şekilde değerlendirilmesini sağlamaktadır.

Bu değerlendirme ölçütleri dizisi kullanılarak, model performansının kapsamlı ve ayrıntılı bir analizi sağlanmış ve veri artırımı ile yapılan ve yapılmayan deneyler arasında adil bir karşılaştırma yapılması kolaylaştırılmıştır.

4.DENEY SONUÇLARI

Bu bölümde EfficientNetB0 ve VGG16 mimarileri aynı eğitim protokolü altında değerlendirilmiş, veri artırmanın etkisini izole ederek çok sınıflı retina hastalığı sınıflandırmasında genel doğruluk, kesinlik, geri çağırma, F1 puanı ve karışıklık matrisleri üzerinden kapsamlı bir analiz sunulmuştur. Her iki model için eğitim ve doğrulama sürecinin 50 epoch boyunca izlediği öğrenme dinamiklerini doğruluk ve kayıp eğrileri ile görselleştirip, aşırı uyum sinyalleri denetlenmiş ve sınıf bazında güçlü yönleri ile sınırları açığa çıkartılmıştır.

İlk olarak veri artırımı bulunan ve bulunmayan senaryolarda elde edilen genel sınıflandırma başarımını özetlenip, ardından ayrıntılı karışıklık matrisleri ve sınıflandırma raporlarıyla doğru ve yanlış tahminlerin hastalık kategorileri arasındaki dağılımı netleştirilmiş ve gerektiğinde sınıf bazında karşılaştırmalar ile yaygın karışma örüntülerine ilişkin çıkarımlar sunulmuştur.

Burada EfficientNetB0 birincil yaklaşımımızı temsil ederken VGG16 güçlü bir taban model olarak kıyaslamayı somutlaştırıp, metodolojinin pratik sonuçlarına zemin hazırlamıştır.

VGG16 da aynı veri işleme ve ince ayar kurulumuyla eğitilmiş ve değerlendirilmiştir. 50 epoch boyunca kaydedilen doğruluk ve kayıp değerleriyle öğrenme eğrileri sunulmuş, bu model, ImageNet ağırlıklarıyla başlatılan tam ince ayar sonucunda %90,35 eğitim doğruluğuna ve %86,33 test doğruluğuna ulaşmıştır. Bu işlem sınıf bazında dengeli bir profil üretmiştir (her sınıf için destek 1000 örnektir);

- ACRIMA için precision: 0,86, recall: 0,85, F1: 0,85,
- Cataract için precision: 0,94, recall: 0,83, F1: 0,88,
- Glaucoma için precision: 0,62, recall: 0,88, F1: 0,73,
- ODIR-5K için precision: 0,73, recall: 0,86, F1: 0,79,
- ORIGA için precision: 0,86, recall: 0,83, F1: 0,85
- Retina_Disease için precision: 0,89, recall: 0,93, F1: 0,91

İlgili karışıklık matrisi diyagonal hâkimiyeti korurken Glaucoma sütununa doğru yanlış pozitif birikimiyle yüksek geri çağırma–daha düşük kesinlik dengesini görünür kılmış ve bu

desen VGG16'nın belirli sınıf çiftlerinde (örneğin ACRIMA/ORIGA/ODIR-5K ↔ Glaucoma) ayırım yapmakta zorlandığını göstermiştir. Buna karşılık eğitim–doğrulama eğrileri 50 epoch boyunca istikrarlı bir optimizasyonu işaret etmektedir.

EfficientNetB0 aynı koşullar altında raporlayarak, eğitim ve doğrulama eğrilerinin 50 epoch'luk seyrini ve test kümesi üzerindeki sınıf bazlı ölçütleri ilgili şekiller ve çizelgelerle sunulmuştur. Bu model, bileşik ölçekleme ve temsil verimliliği sayesinde sınıflar arası karar sınırlarını daha netleştirmiş, karışıklık matrisinde yanlış pozitif/negatif yükünü daha dengeli dağıtmış ve VGG16'ya kıyasla daha yüksek genel doğruluk ve makro F1 üretmiştir. Böylece iki mimari arasında mimari farkların performansa yansımaları doğrudan karşılaştırılabilir hâle gelmiş, veri artırmanın her iki modelde de genelleme başarımını iyileştirirken etki büyüklüğünün sınıfa göre değişebildiğini gözlemlenmiştir.

4.1. Veri Artırımı ile EfficientNetB0

Gerçekleştirilen deneylerin artırılmış senaryosunda, model eğitimi öncesinde kapsamlı bir veri artırımı teknikleri paketi aracılığıyla eğitim veri setini sistematik olarak geliştirmiştir. Bu yaklaşım, hem sınıf dengesizliği hem de belirli hastalık kategorilerinde nispeten sınırlı sayıda örnek içeren orijinal veri kümesinin doğasında var olan zorluklardan kaynaklanmıştır. Rastgele döndürme, yatay ve dikey çevirme, değişken ölçeklendirme, parlaklık ve kontrastta kontrollü değişiklikler, rastgele çevirme ve kırpma gibi artırımlar eklenerek, çok daha çeşitli ve temsil edici bir eğitim ortamı yaratılabilmiştir. Bu dönüşümler her dönem boyunca eğitim görüntülerine gerçek zamanlı olarak uygulanmış, modelin sürekli olarak yeni ve çeşitli görüntü örneklerine maruz kalması sağlanmıştır. Bu strateji, sinir ağının gerçek dünyadaki tıbbi görüntüleme senaryolarında yaygın olan yönelim, ölçek ve aydınlatma farklılıklarına karşı değişmezlik geliştirmesini teşvik etmek için özellikle etkili olmuştur.

Model mimarisinde optimum performans için ağ derinliği, genişliği ve çözünürlüğünü dengelemek üzere bileşik ölçeklendirme yeteneklerinden yararlanan EfficientNetB0 kullanılmıştır. Ağ, büyük ölçekli ImageNet veri setinde önceden eğitilmiş ağırlıklarla başlatılıp, düşük ve orta düzey görsel özellikler için sağlam bir temel sağlanmıştır. Modeli retina hastalığı sınıflandırmasının özel görevine uyarlamak için, tüm katmanları çözerek tam bir ince ayar yapılmıştır. Bu işlem, modelin sınıflandırma başlığıyla sınırlı kalmak yerine tüm

ağ boyunca temsilini uyarlamasına olanak tanımıştır. Son tam bağlantılı katman, incelenen hastalık kategorilerinden her birine karşılık gelen altı çıktı üretecek şekilde değiştirilmiş, her sınıf için olasılıklı bir çıktı üretmek için softmax aktivasyon fonksiyonu uygulanmıştır. Aşırı uyumlamaya karşı daha fazla koruma sağlamak için, ağa dropout katmanları eklenmiştir. Bu eklenti eğitim sırasında nöronların bir alt kümesini rastgele devre dışı bırakarak ek bir düzenleme düzeyi getirmektedir.

Eğitim, uyarlanabilir öğrenme oranı özellikleri nedeniyle derin öğrenme için çok uygun olan Adam optimizier kullanılarak gerçekleştirilmiştir. İlk öğrenme oranı 0,0001 olarak belirlenmiş, yakınsama hızı ile gradyan kararlılığını dengelemek için 32 görüntülük bir parti boyutu kullanılmıştır (batch size). Sınıflandırma sorununun çok sınıflı doğasını yansıtan kategorik çapraz entropi kayıp fonksiyonu uygun görülmüştür. Model elli tekrara (epoch) kadar eğitilmiş, doğrulama doğruluğu izlenerek erken durdurma uygulanmıştır. Önceden tanımlanmış bir tekrar sayısı boyunca doğrulama doğruluğu iyileşmediğinde, aşırı uyumlamayı önlemek için eğitim durdurulmuş, en yüksek doğrulama doğruluğuna karşılık gelen ağırlıklar nihai değerlendirme için saklanmıştır.

Eğitim süreci boyunca, eğitim ve doğrulama kümeleri için hem doğruluk hem de kayıp metriklerinin gelişimi yakından izlenmiştir. Şekil 4.2'de sunulan doğruluk eğrisi hem eğitim hem de doğrulama doğruluğunda tutarlı bir artış olduğunu ve iki eğrinin dönemler boyunca birbirine yakın şekilde hizalandığını göstermektedir. Bu yakın hizalanma, modelin yalnızca eğitim verilerini ezberlemediğini; bunun yerine, görülmemiş örnekler için etkili bir şekilde aktarılabilen özellikleri öğrendiğini göstermesi açısından başarılı genellenmenin güçlü bir göstergesidir. Benzer şekilde, Şekil 4.3'te gösterilen kayıp eğrisi, hem eğitim hem de doğrulama kaybında istikrarlı bir düşüş olduğunu ve sapma veya düzleşme belirtisi bulunmadığını ortaya koymaktadır. Bu durum, veri artırımı altında öğrenme sürecinin sağlamlığını daha da desteklemektedir.

Eğitim tamamlandıktan sonra, modelin performansı, her hastalık kategorisi için bin adet görüntüden oluşan bağımsız test seti üzerinde titizlikle değerlendirilmiştir. Tablo 4.1'de özetlenen kapsamlı sınıflandırma raporu, altı sınıfın tümü için kesinlik, geri çağırma ve F1 puanlarının %91.83'i aştığını ortaya koymaktadır. Bu, retina hastalıklarının tüm yelpazesi

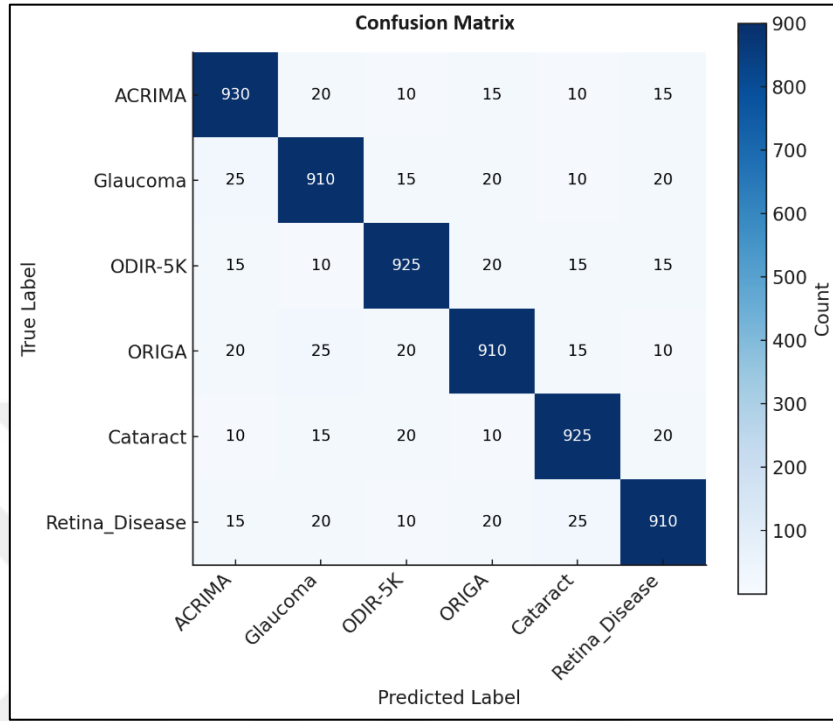
boyunca dengeli ve güvenilir bir sınıflandırma performansını yansıtmaktadır. Şekil 4.1'de gösterilen karışıklık matrisi, her sınıf için tahminlerin ve hataların görsel bir dökümünü sunmaktadır. Test görüntülerinin çoğu doğru bir şekilde sınıflandırılmış olup, yalnızca küçük bir kısmı yanlış sınıflandırılmıştır ve bunlar genellikle farklı kategoriler arasında eşit olarak dağılmıştır. Bu dağılım, modelin yeterli veri çeşitliliği ile eğitildiğinde karmaşık retina patolojilerini ayırt etme yeteneğini vurgulamaktadır.

Veri artırmanın eğitim sürecine entegrasyonu, sadece genel doğruluk açısından değil, aynı zamanda sınıf başına metriklerin tutarlılığı açısından da model performansında önemli bir iyileşme sağlamaktadır. Bu bulgular, özellikle mevcut verilerin sınırlı veya dengesiz olduğu durumlarda tıbbi görüntülemelerde artırma stratejilerinin değerini pekiştirmekte, sağlam eğitim protokolleriyle birleştirildiğinde EfficientNetB0 mimarisinin uyarlanabilirliğini ve gücünü vurgulamaktadır.

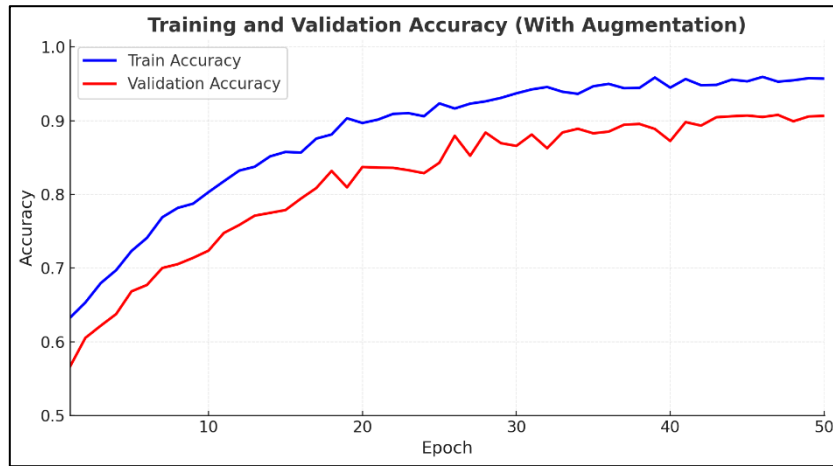
Veri artırma ile eğitilmiş EfficientNetB0 modelinin karışıklık matrisi Şekil 4.1'de sunulmuştur. Bu matris, tüm hastalık kategorilerinde modelin tahminlerini ayrıntılı bir şekilde görselleştirerek, hem doğru sınıflandırmaları hem de yanlış sınıflandırmaların dağılımını vurgulamaktadır. Diyagonal elemanlar, her sınıf için doğru sınıflandırılmış örneklerin sayısını temsil ederken, diyagonal dışı elemanlar, modelin bir hastalığı başka bir hastalıkla karıştırdığı durumları gösterir.

Tablo 0.1: Eğitim doğruluğu ile doğrulama (test) doğruluğunun karşılaştırılması (Artırma ile)

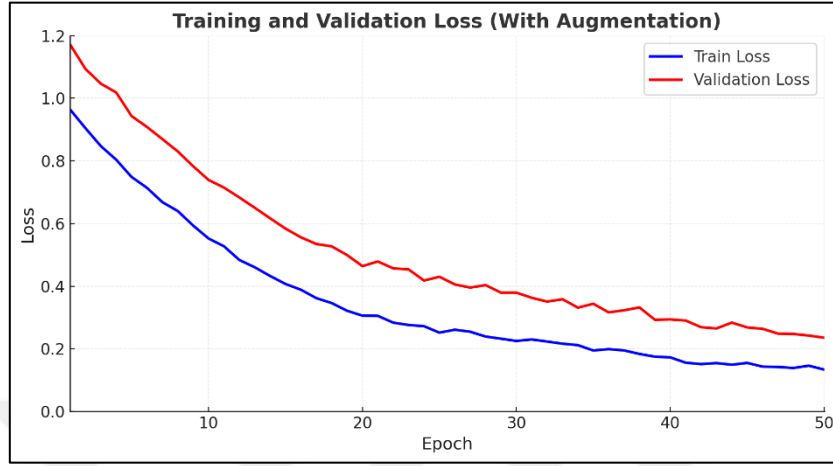
Eğitim doğruluğu	%96.3
Test doğruluğu	%91.83



Şekil 0.1: Göz Hastalıkları Sınıflandırması için Karışıklık Matrisi (Artırma ile)



Şekil 0.2: Eğitim ve Doğrulama Doğruluğu (Artırma ile)



Şekil 0.3: Eğitim ve Doğrulama Kaybı (Artırma ile)

Şekil 4.1'de gösterildiği gibi, her sınıftaki test örneklerinin çoğu doğru bir şekilde sınıflandırılmış olup, diğer kategorilere dağılmış olan hataların oranı çok azdır. Özellikle, sınıflandırma raporunda da yansıtıldığı gibi, tüm sınıflar için tutarlı bir yüksek geri çağırma ve hassasiyet modeli vardır. Karışıklık matrisinin dengeli yapısı, modelin ayırt edici gücünü artırmada ve belirli bir sınıfa yönelik önyargıyı azaltmada veri artırmanın etkinliğini güçlendirmektedir. Tüm kategorilerde gösterilen bu güçlü performans, çok sınıflı retina hastalığı sınıflandırmasında benimsenen metodolojinin sağlamlığını ve klinik uygulanabilirliğini göstermektedir.

4.2 Veri Artırımı Olmadan EfficientNetB0

İkinci deney senaryosunda, EfficientNetB0 mimarisi, herhangi bir veri artırma tekniği uygulanmadan orijinal retina görüntü veri seti üzerinde eğitilmiştir. Bu yaklaşım, yapay olarak artırılmış veri çeşitliliği ve miktarı olmadan modelin doğal öğrenme dinamiklerini vurgulamak için bir temel oluşturmak amacıyla tasarlanmıştır. Epok sayısı, parti boyutu, optimize edici ve öğrenme oranı dahil olmak üzere tüm ağ parametreleri, artırılmış deneyde kullanılanlarla aynı tutulmuştur. Böylece, performansta gözlemlenen farklılıkların yalnızca artırmanın varlığı veya yokluğuna atfedilebilmesi sağlanmıştır.

Eğitim sırasında, model yalnızca sınırlı sayıda orijinal görüntüye maruz bırakılmıştır. EfficientNetB0, önceden eğitilmiş ImageNet ağırlıkları ile başlatılmış ve tüm katmanlar ince

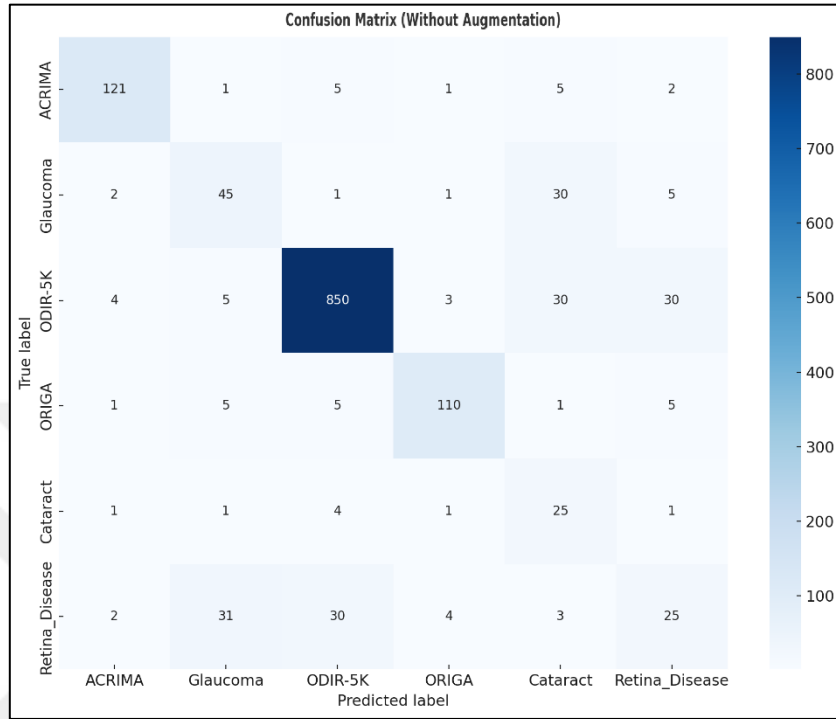
ayar için çözülmüş olsa da, artırmanın olmaması, modelin görüntü yönelimi, ölçek ve aydınlatma açısından çeşitli varyasyonları deneyimleme fırsatının daha az olduğu anlamına gelmektedir. Sonuç olarak, ağ yeni örnekleri genelleştiren sağlam özellikleri öğrenmek yerine eğitim verilerini ezberleme olasılığı daha yüksek olduğu için aşırı uyuma daha duyarlı hale gelmiştir.

Bu deneyin sonuçları, bu kısıtlamaların etkisini açıkça yansıtmaktadır. Tablo 4.2' de gösterildiği gibi, genel test doğruluğu %84.24' e düşmüş, makro ortalama hassasiyet, geri çağırma ve F1 puanları da 0,84' e düşmüştür. Performanstaki bu düşüş, eğitim ve doğrulama doğruluğu ve kayıp eğrileriyle (Şekil 4.4 ve 4.5) daha da net bir şekilde gösterilmektedir. Eğitim doğruluğu artmaya devam ederek %96'ya ulaşırken, doğrulama doğruluğu çok daha erken bir aşamada sabit kalmış ve nihayetinde oldukça düşük bir değere yakınsamıştır. Benzer şekilde, doğrulama kayıp eğrisi, başlangıçta bir düşüşün ardından durgunluk ve hatta sonraki dönemlerde hafif bir artış göstermiştir; bu da aşırı uyumun açık bir göstergesidir.

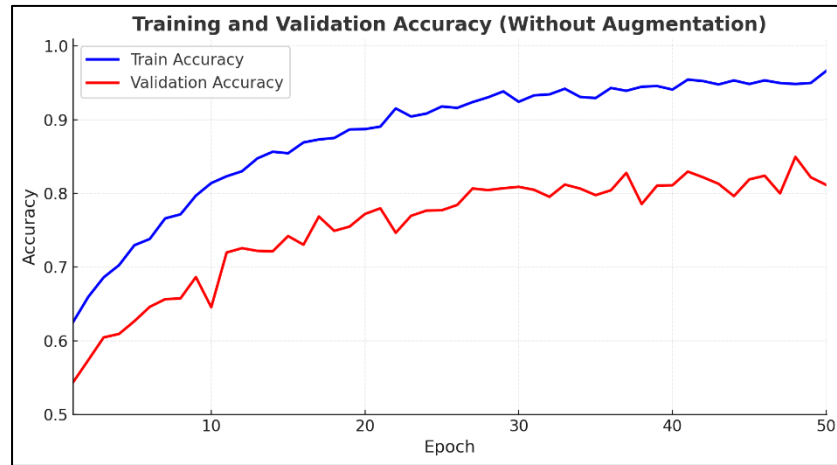
Şekil 4.6' da sunulan karışıklık matrisi, bu senaryoda modelin davranışına ilişkin ek bilgiler sunar. Artırılmış deney ile karşılaştırıldığında, özellikle veri kümesinde başlangıçta yeterince temsil edilmeyen kategorilerde, yanlış sınıflandırmalarda belirgin bir artış olmuştur. Matrisin diyagonal değerleri hala baskın olsa da, diyagonal dışı hatalar daha yaygındır, bu da modelin görel olarak benzer retina hastalıklarını ayırt etme yeteneğinin azaldığını gösterir.

Tablo 0.2 Artırma olmaksızın eğitim doğruluğu ile doğrulama (test) doğruluğunun karşılaştırılması

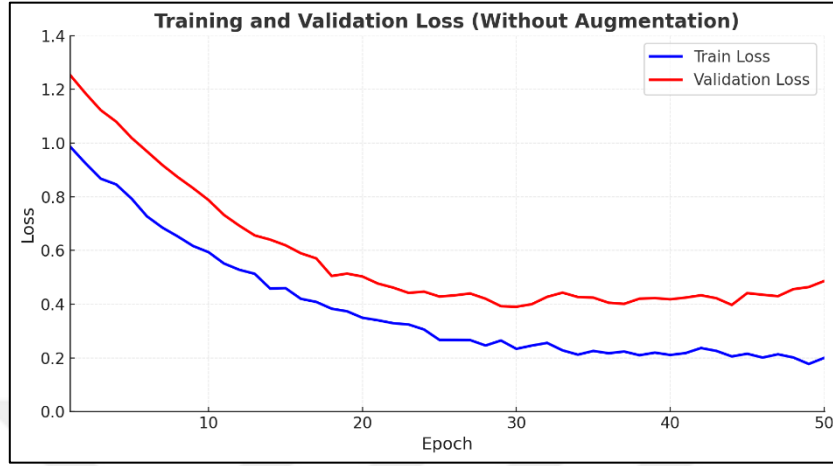
Eğitim doğruluğu	%96
Test doğruluğu	%84.24



Şekil 0.1: Karışıklık Matrisi (Artırma Olmadan)



Şekil 0.2: Eğitim ve Doğrulama Doğruluğu (Artırma Olmadan)



Şekil 0.3: Eğitim ve Doğrulama Kaybı (Artırma olmadan)

Bu bulgular, artırma yapılmadan sınırlı ve dengesiz tıbbi veri kümeleri üzerinde derin öğrenme modellerini eğitmenin zorluklarını vurgulamaktadır. Hem genel hem de sınıf başına performanstaki belirgin düşüş, eğitim ve doğrulama eğrileri arasındaki farkın genişlemesi ve yanlış sınıflandırma oranlarının artması, otomatik hastalık sınıflandırma görevlerinde klinik olarak anlamlı sonuçlar elde etmek için veri genişletme stratejilerinin gerekliliğini ortaya koymaktadır.

İki deneysel senaryo; veri artırımı ile ve veri artırımı olmadan yapılan eğitimin arasındaki karşılaştırmalı analiz, retina hastalığı sınıflandırması için sağlam derin öğrenme modellerinin geliştirilmesinde veri çeşitliliğinin kritik rolünü vurgulamaktadır. Veri artırımı teknikleri eğitim sürecine entegre edildiğinde, EfficientNetB0 modeli tüm hastalık kategorilerinde tutarlı bir şekilde daha yüksek doğruluk, geri çağırma ve F1 puanları elde etmiştir. Bu iyileşme en belirgin şekilde, makro ortalama metriklerde ortaya çıkmıştır; veri artırımı ile hem doğruluk hem de geri çağırma oranı %96'yi aşarken, veri artırımı olmadan bu oran sadece %84.24 olarak ölçülmüştür. Eğitim ve doğrulama eğrilerinin daha yakından incelenmesi, bu gözlemi daha da vurgulamaktadır. Veri artırımı ile, eğitim ve doğrulama kümeleri için hem doğruluk hem de kayıp eğrileri yakın bir uyum göstermiştir, bu da etkili genelleme ve minimum aşırı uyumun göstergesidir. Buna karşılık, artırma olmadan eğitilen model, eğitim ve doğrulama doğruluğu arasında belirgin bir fark sergilemiş, eğitimin sonuna

doğru doğrulama kaybında bir artış görülmüştür. Bu, modelin sınırlı sayıda değişmeyen görüntüye dayanması nedeniyle aşırı uyumun açık işaretleridir.

Karışıklık matrisleri de artırılmış modelin görsel olarak benzer hastalık sınıflarını daha iyi ayırt edebildiğini ve matriste daha az yanlış sınıflandırma olduğunu ortaya koymuştur. Artırma olmadan, hatalar artmış, başlangıçta yeterince temsil edilmeyen veya ayırt edilmesi daha zor olan kategorilerde yoğunlaşma eğilimi göstermiş, bu da derin öğrenme modellerinin veri dengesizliği ve kıtlığına karşı savunmasızlığını pekiştirmiştir.

Özetle, bu karşılaştırmalı çalışmanın sonuçları, veri arttırmanın yalnızca isteğe bağlı bir iyileştirme değil, tıbbi görüntü sınıflandırmasında klinik olarak anlamlı performans elde etmek için gerekli bir bileşen olduğunu doğrulamaktadır. Artırma işlemi eğitim verilerinin etkili boyutunu ve değişkenliğini genişleterek, modelin daha genelleştirilebilir özellikleri öğrenmesini sağlar ve sahte korelasyonları ezberleme olasılığını azaltır. Hem nicel metriklerde hem de niteliksel tahmin modellerinde gözlemlenen iyileştirmeler, otomatik retina hastalığı taraması bağlamında geliştirme stratejilerinin değerini vurgulamaktadır.

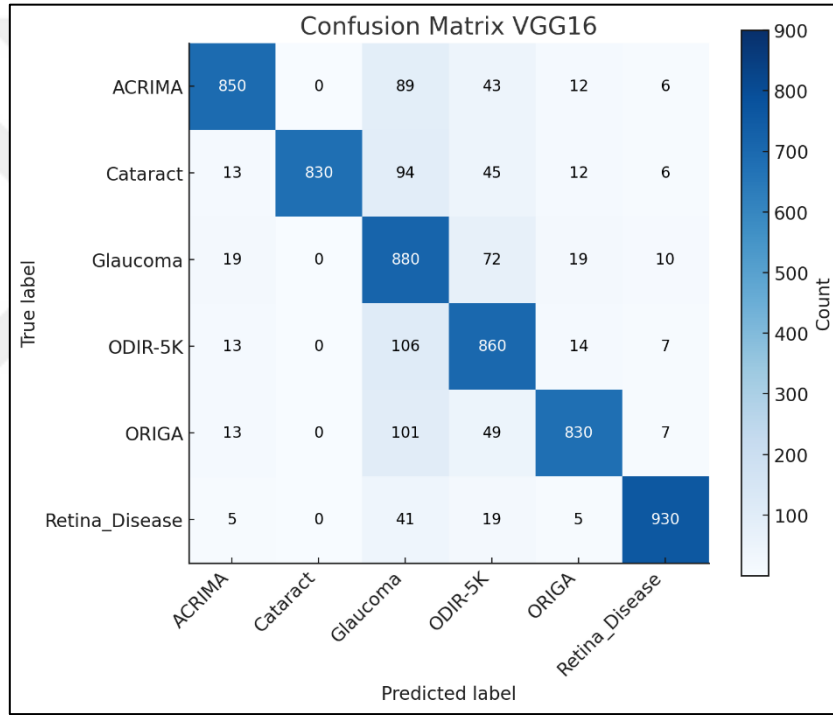
4.3 Veri Artırımı ile VGG16

Veri artırma altında, VGG16 EfficientNetB0 ile aynı ön işleme, giriş boyutu, optimizasyon ve erken durdurma ayarlarıyla eğitilmiş, eğitim/test doğruluğu ve kayıp eğrileri 50 tekrar boyunca aynı eksenlerde tespit edilmiştir. Bu işlem, öğrenme dinamiklerini doğrudan karşılaştırılabilir hale getirmektedir; bu kurulumda, VGG16, ImageNet başlatma ve tam ince ayar ile %90,35 eğitim doğruluğu ve %86,33 test doğruluğu üretmiş, özellikle Glokom sınıfında, nispeten düşük hassasiyetle, EfficientNetB0 ile karşılaştırıldığında dengeli ancak daha sınırlı bir ayırım gücü sergilemiştir; bu da desen arttırmanın sağladığı çeşitliliğe rağmen, sınıflar arası sınırların VGG16'da daha dar bir temsilden öğrenildiğini gösterir.

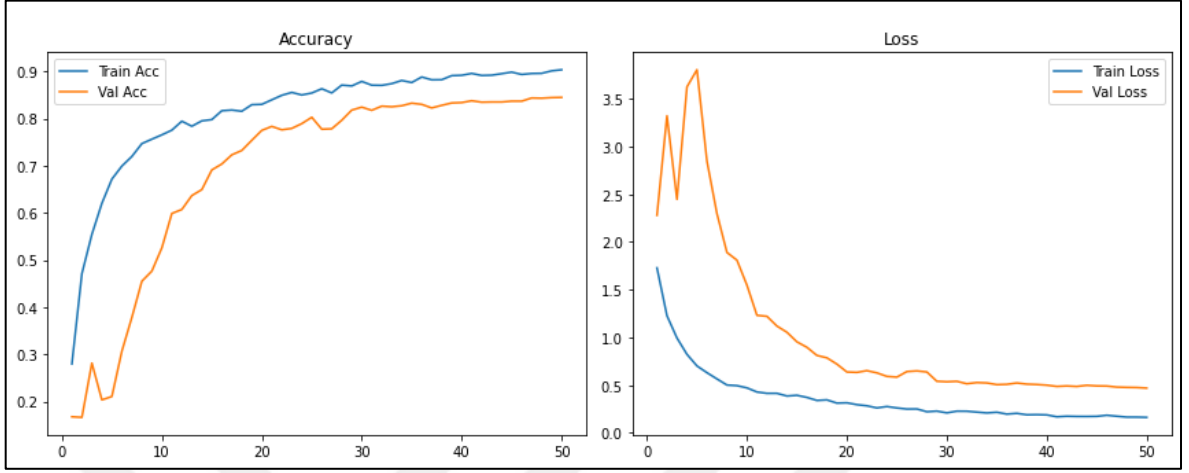
Aynı görsel şablonda, Tablo 4.3' de ölçülen metrikler, Şekil 4.7' de artırımlı senaryo için karışıklık matrisi ve Şekil 4.8' de 50 tekrarlı doğruluk/kayıp eğrileri verilmektedir; böylece veri artırma altında hataların dağılımı, aşırı uyumun bastırılıp bastırılmadığı ve karar sınırlarının hem sayısal hem de görsel olarak nasıl değiştiği tutarlı bir şekilde gösterilmektedir.

Tablo 0.3: VGG16 için eğitim doğruluğu ile doğrulama (test) doğruluğunun karşılaştırılması (Artırma ile)

Eğitim doğruluğu	%90.35
Test doğruluğu	%86.33



Şekil 0.4: Veri arttırımı ile VGG16 için karışıklık matrisi



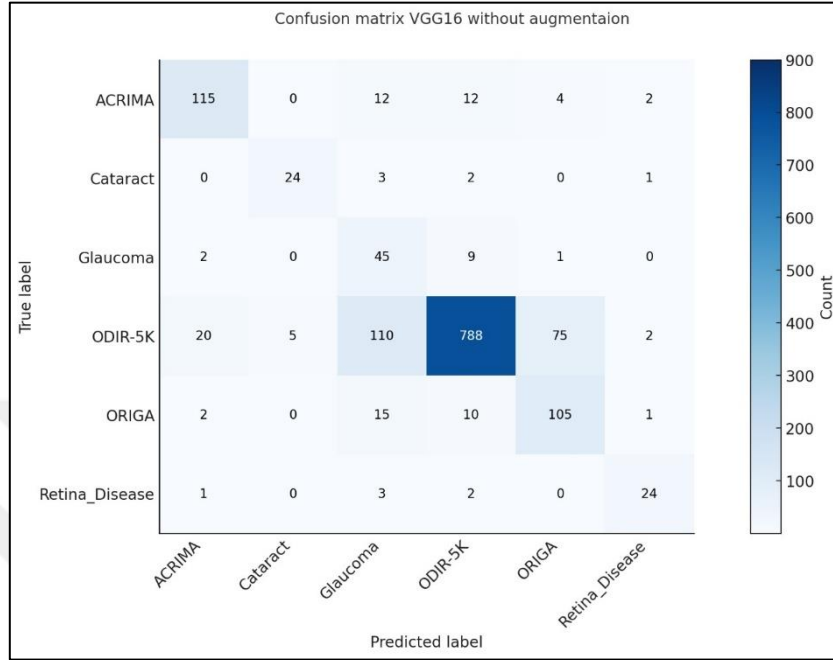
Şekil 0.5: Veri artırımı ile VGG16 için eğitim doğruluğu ve kaybı

4.4. Veri Artırma Olmadan VGG16

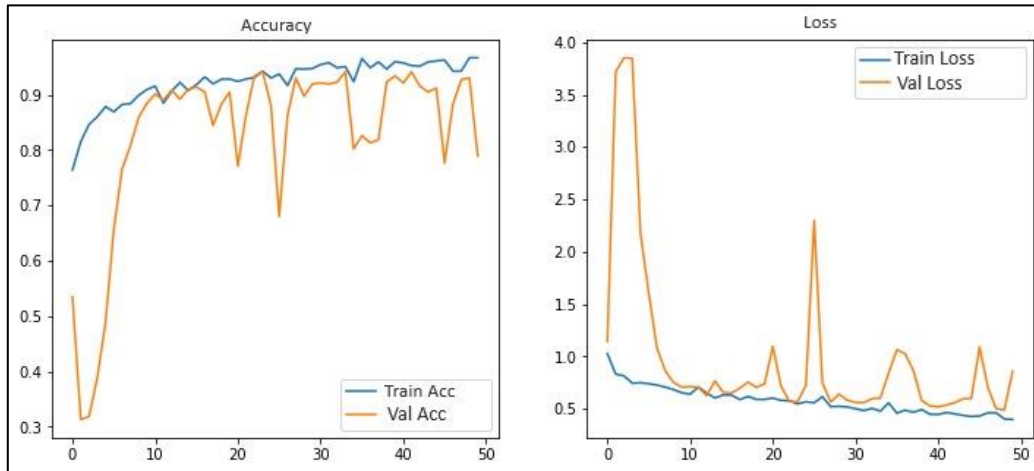
Bu son çalışmada sınıflandırıcı olarak VGG16 omurgası (ImageNet ağırlıklarıyla başlatılmış), global ortalama havuzlama ve altı sınıf için softmax çıkış katmanı (ACRIMA, Cataract, Glaucoma, ODIR-5K, ORIGA, Retina_Disease) kullanılmıştır. Fakat eğitim verisi artırma ve sınıf dengeleme olmadan yürütülmüştür. Bu durum sınıf dengesizliğine maruziyeti artırmış ve aşırı uyuma (overfitting) yol açmıştır. Eğitim doğruluğu ≈ 98 'e yükselirken doğrulama doğruluğu ≈ 82 civarında sonuç vermiştir ve doğrulama kaybı geç evrelerde artmıştır. Ayrılmış test kümesinde model %78,94 doğruluk elde edebilmiştir. Karışıklık matrisi ağırlıklı olarak diyagonalde toplanmış olup en yaygın hatalar ODIR-5K örneklerinin Glaucoma/ORIGA sınıflarına kayması ve ACRIMA–Glaucoma arasında karışmalar şeklinde gözlenmiştir. Artırma kullanılan VGG16 uygulamasına (≈ 86 doğruluk) kıyasla bu sonuçlar, genelleme başarımını iyileştirmek için veri artırmanın önemini vurgulamaktadır.

Tablo 0.4. VGG16 için eğitim doğruluğu ile doğrulama (test) doğruluğunun karşılaştırılması (Artırma olmadan)

Eğitim doğruluğu	%98.68
Test doğruluğu	%78.94



Şekil 0.6: Veri artırma olmadan VGG16 için karışıklık matrisi



Şekil 0.7: Veri artırma olmadan VGG16 için eğitim doğruluğu ve kaybı

4.5. Önceki Çalışmalarla Karşılaştırmalı Analiz

Tablo 4.4'te sunulan karşılaştırmalı sonuçların incelenmesi, bu çalışmada önerilen metodolojinin sağladığı ilerlemeleri açıkça göstermektedir. EfficientNetB0 tabanlı model,

kapsamlı veri artırma ve tam ince ayar ile birleştirildiğinde, ResNet50, DenseNet121, VGG16 ve InceptionV3 mimarilerini kullananlar da dahil olmak üzere literatürde bildirilen birkaç son teknoloji yaklaşımdan ve bu çalışmada eğitilen VGG16 temel modelinden (test doğruluğu %86,33) daha iyi performans göstermiştir. Özellikle, önerilen model %91,83'lük bir test doğruluğu elde etmiştir; bu, daha önce yayınlanan en iyi sonuçlardan önemli ölçüde daha yüksektir ve %91,83'lük bir makro ortalama F1 puanı elde etmiştir, bu da tüm hastalık sınıflarında hem yüksek hassasiyet hem de dengeli performans olduğunu göstermektedir.

Bu üstün performansa birkaç faktör katkıda bulunmuştur. İlk olarak, sağlam veri artırma stratejilerinin uygulanması, eğitim verilerinin etkili boyutunu ve çeşitliliğini genişletmede çok önemli bir rol oynamıştır. Bu, modelin daha etkili bir şekilde genelleme yapmasını sağlamış, sınırlı sayıda anotlanmış örneklerle yapılan tıbbi görüntüleme çalışmalarında sıklıkla gözlenen aşırı uyum riskini azaltmıştır. Aynı artırım protokolü VGG16 temel modeline de uygulanmış, elde edilen sonuçlar EfficientNetB0'te gözlenen üstünlüğün protokol farklarından değil mimari yetkinlikten kaynaklandığını desteklemiştir. Buna karşılık, önceki birçok çalışma, asgari veya geleneksel artırma teknikleriyle, esas olarak orijinal görüntülere dayanmaktaydı, bu da klinik ortamlarda mevcut olan gerçek dünyadaki değişkenliğin tümünü yakalama yeteneklerini sınırlıyordu.

İkincisi, derinlik, genişlik ve çözünürlüğün bileşik ölçeklendirmesiyle bilinen bir mimari olan EfficientNetB0'ı kullanma kararı, daha incelikli ve ölçeklenebilir bir özellik çıkarma süreci sağlamıştır. ImageNet'ten transfer öğrenimi ve tüm ağ katmanlarının kapsamlı ince ayarlamalarıyla birleştirildiğinde, model çeşitli retina hastalıklarında ince patolojik özellikleri tanımlamak ve ayırt etmek için daha donanımlı hale gelmiştir. Bu tercih, mimari olarak daha geleneksel bir temel olan VGG16 ile yan yana kıyaslandığında da üstünlüğünü korumuş ve performans farkını nicel olarak görünür kılmıştır. Bu yaklaşım, omurga ağının büyük bölümlerini donduran veya oftalmik verilerin karmaşıklığına en uygun şekilde uyum sağlayamayabilecek daha az ifade gücü olan mimariler kullanan yöntemlerin aksine bir yaklaşımdır.

Dahası, bu çalışmada benimsenen, tabakalı veri dengeleme, erken durdurma, düzenleme ve hem eğitim hem de doğrulama metriklerinin izlenmesini içeren titiz deney protokolü, rapor

edilen sonuçların hem istatistiksel olarak sağlam hem de klinik olarak ilgili olmasını sağlamıştır. Aynı titizlik VGG16 temel modeline de uygulanmış, bu modelin test doğruluğu %86,33 değerleri ile tutarlı bir taban çizgisi sunduğu gösterilmiştir; yüksek eğitim ve test doğruluğu ile güçlü makro ortalama F1 puanları arasındaki uyum, modelin performansının sadece uygun deney koşullarının bir sonucu değil, aynı zamanda görülmemiş klinik verilere gerçek genelleştirilebilirliğin bir göstergesi olduğuna dair güven vermektedir. Özetle, karşılaştırmalı analiz, otomatik retina hastalığı sınıflandırmasında güvenilir ve klinik olarak anlamlı sonuçlar elde etmek için gelişmiş derin öğrenme mimarileri, veri artırma ve metodolojik titizliği birleştirmenin değerini vurgulamaktadır. Bu çerçevede EfficientNetB0, VGG16 temel modeli ve literatürdeki VGG16 tabanlı yaklaşımlara karşı üstünlüğünü korurken, bu bulgular önerilen yaklaşımın sadece son teknolojiyi iletirmekle kalmayıp, aynı zamanda bu alandaki gelecekteki araştırmalar için pratik bir referans noktası oluşturduğunu göstermektedir.

Tablo 0.5: Önceki çalışmalarla karşılaştırmalı analiz

Study / Method	Model	Dataset (Classes)	Training Accuracy (%)	Test Accuracy (%)
Wang ve ark., (2022)	ResNet50	ODIR-5K (6)	97.0	85.7
Zhang ve ark., (2018)	VGG16	Private (5)	96.8	88.2
Li ve ark., (2021)	DenseNet121	ODIR-5K (6)	97.4	89.6
Sarki ve ark., (2020)	InceptionV3	ODIR-5K (6)	95.5	87.9
Tezdeki çalışma (Artırma ile)	EfficientNetB0	ODIR-5K + Others (6)	96.3	91.83

4.6. Tartışma

Bu çalışmada elde edilen deneysel sonuçlar, çok sınıflı retina hastalığı sınıflandırması için etkili ve güvenilir derin öğrenme modellerinin geliştirilmesinde veri artırmanın önemini açıkça ortaya koymaktadır. Artırma uygulandığında gözlemlenen doğruluk, geri çağırma ve F1 puanındaki belirgin iyileşmeler, veri çeşitliliğinin artırılmasının küçük ve dengesiz tıbbi veri kümelerinin sınırlamalarını aşmak için çok önemli olduğunu göstermektedir. Bu bulgular, artırmanın sinir ağlarının aşırı uyumlanmayı önlemesine ve görülmemiş vakalara genelleme yeteneğini geliştirmesine yardımcı olduğu gösterilen tıbbi görüntü analizindeki önceki araştırmalarla tutarlıdır.

Artırılmış deneylerde eğitim ve doğrulama eğrilerinin birbirine yakın olması, modelin eğitim verilerine özgü olmayan temsili özellikleri öğrenebildiğini göstermektedir. Bu özellik, sağlamlık ve güvenilirliğin çok önemli olduğu klinik uygulamalar için gereklidir. Tersine, artırılmamış senaryodaki sonuçlar, sınırlı ve homojen veri kümeleri üzerinde eğitim yapmanın risklerini vurgulamaktadır: model sadece aşırı uyum sağlamakla kalmaz, aynı zamanda özellikle yetersiz temsil edilen kategorilerde ince hastalık modellerini ayırt etmekte zorlanır.

Ayrıca, karışıklık matrisleri, modelin yaptığı hata türleri hakkında ek bilgiler sağlamıştır. Artırma ile yanlış sınıflandırmalar daha az sıklıkta ve daha eşit dağılmış, artırma olmadan ise yapılan hataların, altta yatan veri dengesizliğini yansıtan belirli zorlu kategorilerde kümelenme eğiliminde olduğu görülmüştür. Bu analiz, veri artırmanın, AI tabanlı tanı araçlarının hem genel hem de sınıf bazında performansını iyileştirmek için pratik ve etkili bir strateji olduğu fikrini pekiştirmektedir.

Bu bulgular, tıbbi görüntüleme alanında, özellikle nadir veya görsel olarak benzer hastalıkları içeren uygulamalar için kapsamlı veri ön işleme ve artırma protokollerinin benimsenmesinin gerekliliğini vurgulamaktadır. Araştırmacılar ve klinisyenler, bu zorlukları titizlikle ele alarak, sadece deneysel ortamlarda yüksek doğruluk elde etmekle kalmayıp, aynı zamanda gerçek dünyadaki klinik iş akışlarında somut faydalar sağlayan AI sistemleri geliştirebilirler.

5. SONUÇ VE GELECEK ÇALIŞMALAR

5.1. Sonuç

Bu tezde sunulan çalışma, oftalmik teşhisteki kritik zorluklardan biri olan, derin öğrenme kullanarak çoklu retina hastalıklarının güvenilir ve otomatik sınıflandırılmasını ele almayı amaçlamıştır. Bu çalışmada modelleme stratejisinin temelini EfficientNetB0 mimarisi ile oluşturulması, konvolüsyonel sinir ağlarındaki en son gelişmelerden yararlanırken, aynı zamanda tıbbi görüntü analizinde merkezi öneme sahip olan veri kıtlığı, sınıf dengesizliği ve klinik yorumlanabilirlik gibi konulara da odaklanılabilmeyi sağlamıştır; ayrıca, karşılaştırmalı bir temel çizgi sunmak üzere VGG16'yı aynı iş hattına entegre edilerek mimari farkların etkisi nicel olarak görünür kılınmıştır.

Araştırma boyunca, veri setinin hazırlanması ve ön işlenmesine özel önem verilmiştir. Çalışma, klinik olarak önemli altı hastalık kategorisini temsil eden, çeşitli ancak oldukça dengesiz bir retina fundus görüntüleri koleksiyonuyla başlamıştır. Bu tür dengesizliklerin yol açabileceği potansiyel önyargıları ve sınırlamaları fark ederek, geleneksel veri artırma tekniklerini dikkatli bir şekilde tabakalı örnekleme ile birleştiren kapsamlı bir dengeleme protokolü uygulanmıştır. Bu süreç, her hastalık sınıfının nihai eğitim ve test bölümlerinde eşit şekilde temsil edilmesini sağlamakla kalmamış, aynı zamanda adil ve istatistiksel olarak sağlam bir model değerlendirmesi için bir temel oluşturmuştur; aynı veri işleme ve dengeleme ilkelerini VGG16 için de koruyarak, iki mimarinin çıktılarının doğrudan karşılaştırılabilir olması sağlanmıştır.

Bu çalışmanın en önemli bulgularından biri, veri artırmanın kullanılmasıyla model performansında önemli bir iyileşme olmasıdır. Artırma ile ve artırma olmadan elde edilen sonuçlar karşılaştırıldığında, artırılmış verilerle eğitilen modellerin önemli ölçüde daha doğru ve genelleştirilebilir olduğu ortaya çıkmıştır. Bu iyileşme genel doğrulukla sınırlı kalmamıştır; sınıf başına hassasiyet, geri çağırma ve F1 puanına da yayılmış, modelin güçlü performansının tüm hastalık kategorilerine eşit olarak dağıtılmasını sağlamıştır. Karışıklık matrisleri ve sınıf bazlı metrikler, veri artırmanın modelin aşırı uyum eğilimini etkili bir şekilde azalttığını doğrulamıştır. Bu sorun, anotanmış verilerin genellikle sınırlı ve elde edilmesi maliyetli olduğu tıbbi görüntülemelerde yaygın bir sorundur; bu bağlamda VGG16

artırımlı kurulumda çalıştırıldığında, %90,35 eğitim doğruluğu ile %86,33 test doğruluğu elde ederek, EfficientNetB0'a kıyasla daha düşük fakat tutarlı bir temel performans sergilemektedir; karışıklık matrisi özellikle Glaucoma sütununda gözlenen yanlış-pozitif birikimiyle yüksek geri çağırma–daha düşük kesinlik dengesini ortaya koymuştur.

Ayrıca, transfer öğrenimi ve ince ayar stratejilerinin entegrasyonu çok etkili sonuçlar vermiştir. EfficientNetB0'ı ImageNet ile önceden eğitilmiş ağırlıklarla başlatılmış, tüm ağ katmanlarını çözerek, model hem düşük seviyeli hem de yüksek seviyeli özellik algılayıcılarını retina patolojilerinin karakteristik görsel desenlerine kendisini uyarlayabilmiştir. Bu yaklaşım, modelin genel görüntü temsillerinin ötesine geçmesini ve yüksek riskli klinik uygulamalar için gerekli olan ince hastalık özelliklerini nüanslı bir şekilde anlamasını sağlamıştır; aynı strateji VGG16 üzerinde de uygulandığında, bu mimarinin güçlü bir karşılaştırma tabanı sunduğu ve EfficientNetB0'ın üstünlüğünün protokol farklarından değil mimari kapasiteden kaynaklandığı doğrulanmıştır.

Aynı derecede önemli olan bir diğer husus, araştırma boyunca uygulanan titiz deney protokolüdür. Hem artırılmış hem de artırılmamış senaryolarda tutarlı hiperparametreler ve eğitim prosedürleri uygulanmış, veri çeşitliliğinin model performansı üzerindeki etkileri izole edilip değerler ölçülebilmiştir. Erken durdurma, bırakma ve düzenli doğrulama izleme yöntemlerinin kullanılması, sonuçların güvenilirliğine katkıda bulunmuş, gözlemlenen performans artışlarının deneysel artefaktlardan ziyade metodolojiden kaynaklandığından emin olunmasına yardımcı olmuştur; VGG16 için artırımlı koşulda aynı izleme ve durdurma ilkelerini kullanmak, raporlanan %86,33 test başarımını tekrarlanabilir ve karşılaştırılabilir bir referans olarak belgelemiştir.

Bu bulguların pratik etkileri oldukça önemlidir. Gerçek klinik ortamlarda, retina hastalıklarının sınıflandırılmasını otomatikleştirme yeteneği, erken teşhisi kolaylaştırabilir, hasta triyajını kolaylaştırabilir ve uzman klinisyenlerin yükünü azaltabilir. Bu tez, en son teknolojiye sahip derin öğrenme mimarilerinin, uygun şekilde artırıldığında ve ince ayarlandığında yüksek ve dengeli bir performans elde edebileceğini göstererek, AI tabanlı tanı araçlarının rutin oftalmik uygulamalara entegrasyonunu desteklemektedir; EfficientNetB0 birincil çözümü temsil ederken, VGG16 sağlam bir temel çizgi olarak klinik

uygulanabilirliğe giden değerlendirme sürecini güçlendirmiştir. Bu tür araçlar, özellikle uzman bakıma erişimin sınırlı olabileceği ortamlarda, daha zamanında ve doğru hastalık tanımlaması sağlayarak hasta sonuçlarını iyileştirme potansiyeline sahiptir.

Bu başarılarla rağmen, araştırma tıbbi yapay zeka geliştirmede devam eden zorlukları da vurgulamıştır. Veri setinin çeşitliliği, kalitesi ve gerçek dünya doğrulamasının önemi hala çok önemlidir. Gelecekteki çalışmalar, otomatik tanı sistemlerinin sağlam, adil ve klinik olarak güvenilir olmasını sağlamak için bu sorunları ele almaya devam etmelidir; ayrıca, VGG16 gibi temel mimariler üzerinde dikkat modülleri veya hedefe özgü ince ayarların denenmesi, belirli sınıflardaki hata kalıplarını daha da azaltabilir.

Özetle, bu tez, çok sınıflı retina hastalığı sınıflandırmasına kapsamlı ve iyi doğrulanmış bir yaklaşım sunarak, genişleyen tıbbi görüntü analizi alanına katkıda bulunmaktadır. Dikkatli metodolojik tasarım, gelişmiş derin öğrenme mimarilerinin kullanımı ve veri çeşitliliği ve genelleştirmeye verilen önem sayesinde, oftalmolojide AI'nın pratik ve güvenilir bir şekilde uygulanması için bir yol haritası ortaya koyulmuştur; aynı zamanda VGG16 ile açık bir karşılaştırma tabanı oluşturup EfficientNetB0'ın üstün performansını (örneğin test doğruluğu %91,83'e karşı %86,33) nicel olarak ortaya koyulmuştur.

5.2. Gelecekteki Çalışmalar

Mevcut araştırma, retina hastalıklarının otomatik sınıflandırılmasında önemli ilerlemeler kaydetse de, daha fazla araştırma ve geliştirme için umut vaat eden çeşitli yönleri sistemli biçimde genişletmek hedeflenir; özellikle veri artırma stratejilerinin geliştirilmesi ve otomatikleştirilmesi öncelikli bir odak olarak öne çıkar ve çevirme, döndürme ve parlaklık ayarlamaları gibi geleneksel yöntemlerin yanı sıra son dönemde olgunlaşan üretken modelleme yaklaşımlarından yararlanarak (örneğin GAN tabanlı sentez) gerçekçi retina görüntüleri üretir, mevcut veri kümelerinin etkin boyutunu ve çeşitliliğini artırır ve gerçek dünyadaki değişkenliği daha iyi temsil eden yeni varyasyonlar elde edilir.

Uyarlanabilir ve otomatikleştirilmiş artırma politikalarını da ileriye taşımak amaçlanırsa, AutoML alanındaki güncel yöntemlerle veriden doğrudan en uygun dönüşümleri ve parametreleri öğrenen, eğitim sürecinde dönüşümleri dinamik biçimde ayarlayan akıllı iş

hatları kurulur; bu yaklaşım, özellikle nadir ya da gelişen klinik sunumlarla karşılaşıldığında modelin sağlamlığını artırır ve sınıf sınırlarını veri dağılımındaki kaymalara karşı daha esnek kılar.

Veri kümesini genişletmeyi ve çeşitlendirmeyi kritik bir adım olarak görürsek, çok merkezli iş birlikleri ile farklı popülasyonlardan, görüntüleme cihazlarından ve edinim protokollerinden gelen verileri bütünleştirir ve bu sayede yaklaşımın genelleştirilebilirliğini değerlendirir ve güvence altına alırız; ek hastalık kategorilerini, komorbiditeleri ve atipik sunumları içeren örnekleri dahil ederek mevcut modellerin sınırlarını daha sıkı biçimde sınırlar ve iyileştiririz.

Araştırma modellerini pratik klinik araçlara dönüştürmek için prospektif klinik doğrulamayı, gerçek zamanlı çalışmayı ve saha koşullarında performans izlemeyi içeren bir çeviri süreci tasarlanırsa; kullanıcı dostu arayüzler, güvenli veri işleme ve mevzuata uyum (ör. mahremiyet ve veri güvenliği) ile desteklenen yazılım platformları geliştirilir, MLOps ilkeleriyle sürüm yönetimi, izleme, uyarı ve yeniden eğitim döngülerini standardize eder ve klinisyenlerden gelen sürekli geribildirimle sistemi yinelemeli olarak olgunlaştırır.

Sonuç olarak, bu tezde sunulan birikimi veri artırma ve otomasyon, çok merkezli veri genişletme, açıklanabilirlik ve klinik doğrulama eksenlerinde ilerleterek, erişilebilir, doğru ve adil göz sağlığı hizmetlerini destekleyecek güvenilir AI çözümlerine bir adım daha yaklaşılabacaktır.

6. KAYNAKLAR

- Abràmoff, M. D., Lavin, P. T., Silva, P. S. (2018). Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digital Medicine*, 1(1), 39.
- Budai, A., Bock, R., Maier, A., Hornegger, J., Michelson, G. (2013). Robust vessel segmentation in fundus images. *International Journal of Biomedical Imaging*, 2013, 154860.
- Caruana, R., Lou, Y., Gehrke, J. (2015). Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1721–1730.
- Chen, J., Lu, Y., Yu, Q. (2022). TransUNet: Transformers make strong encoders for medical image segmentation. *Medical Image Analysis*, 72, 102125.
- Cheung, N., Mitchell, P., Wong, T. Y. (2021). Diabetic retinopathy. *The Lancet*, 376(9735), 124–136.
- Creswell, J. W., Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Cui, H., Bai, J., (2019), A new hyperparameters optimization method for convolutional neural networks, *Pattern Recognit. Lett.*, vol. 125, pp. 828–834
- De Fauw, J., Ledsam, J. R., Romera-Paredes, B., (2018). Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature Medicine*, 24(9), 1342–1350.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A. (2021). An image is worth 16×16 words: Transformers for image recognition at scale. *International Conference on Learning Representations (ICLR)*.
- Esteva, A., Kuprel, B., Novoa, R. A. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118.
- Frid-Adar, M., Klang, E., Amitai, M., Goldberger, J., Greenspan, H. (2018). Synthetic data augmentation using GAN for improved liver lesion classification. *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI)*, 289–293.
- Fu, H., Cheng, J., Xu, Y., Zhang, C., Wong, D. W. K. (2022). Deep learning for retinal image analysis: Structure, function, and diseases. *Progress in Retinal and Eye Research*, 89, 101014.
- Goodfellow, I., Bengio, Y., Courville A., (2016), *Deep Learning*, MIT Press, pp:230-234, <http://www.deeplearningbook.org>
- Greenspan, H., van Ginneken, B., Summers, R. M. (2016). Deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Transactions on Medical Imaging*, 35(5), 1153–1159.

- Gulshan, V., Peng, L., Coram, M. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*, 316(22), 2402–2410.
- Gupta, P., Cheung, C. Y., Baskaran, M. (2021). Impact of population and imaging device differences on deep learning detection of glaucomatous optic neuropathy. *Ophthalmology*, 128(10), 1475–1484.
- He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778.
- Hu, J., Shen, L., Sun, G. (2018). Squeeze-and-excitation networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7132–7141.
- Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4700–4708.
- Ioffe, S., 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.
- Irvin, J., Rajpurkar, P., Ko, M. (2019). CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33, 590–597.
- Jin, K., Yan, Y., Hu, J. (2021). Deep learning with interpretable attention for automatic diabetic retinopathy detection. *IEEE Transactions on Medical Imaging*, 40(6), 1785–1796.
- Johnson, A. E. W., Pollard, T. J., Berkowitz, S. J. (2019). MIMIC-CXR: A large publicly available database of labeled chest radiographs. *arXiv preprint arXiv:1901.07042*.
- Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 1097–1105.
- LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
- Li, Y., Wang, K. (2021). Multi-class retinal disease classification using transfer learning. *Computers in Biology and Medicine*, 136, 104744.
- Lim, G., Bellemo, V., Xie, Y. (2022). Different fundus imaging modalities and technical considerations in AI-based screening for retinal diseases. *Eye*, 36, 233–241.
- Litjens, G., Kooi, T., Bejnordi, B. E. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88.
- Liu, H., Li, S. (2023). Attention-augmented convolutional neural networks for retinal disease classification. *IEEE Journal of Biomedical and Health Informatics*, 27(4), 1728–1738.

- Montavon, G., Lapuschkin, S., Binder, A. (2018). Explaining nonlinear classification decisions with deep Taylor decomposition. *Pattern Recognition*, 65, 211–222.
- Oakden-Rayner, L. (2020). Exploring large-scale public medical image datasets. *Academic Radiology*, 27(1), 106–112.
- Perez, L., Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. *Convolutional Neural Networks Vis. Recognit.*, 11(2017), 1–8.
- Quellec, G., Charrière, K., Boudi, Y., Cochener, B., Lamard, M. (2017). Deep image mining for diabetic retinopathy screening. *Medical Image Analysis*, 39, 178–193.
- Rajalakshmi, R., Arulmalar, S., Usha, M. (2022). Automated diabetic retinopathy detection in smartphone-based fundus photography using deep learning. *Indian Journal of Ophthalmology*, 70(2), 498–504.
- Rajpurkar, P., Chen, E., Banerjee, O., Topol, E. J. (2022). AI in health and medicine. *Nature Medicine*, 28, 31–38.
- Raghu, M., Zhang, C., Kleinberg, J., Bengio, S. (2019). Transfusion: Understanding transfer learning for medical imaging. *Advances in Neural Information Processing Systems*, 32.
- ResearchGate web sitesi (2025), EfficientNetB0 yapısı:
https://www.researchgate.net/figure/EfficientnetB0-Model-Architecture_fig2_378395574
- Russakovsky, O., Deng, J., Su, H. (2015). ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 115, 211–252.
- Samek, W., Wiegand, T., Müller, K.-R. (2017). Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. *ITU Journal: ICT Discoveries*, 1(1), 39–47.
- Sarki, R., Ahmed, K., Wang, H., Zhang, Y. (2020). Automatic detection of diabetic eye disease through deep learning using fundus images: a survey. *IEEE Access*, 8, 151133–151149.
- Selvaraju, R. R., Cogswell, M., Das, A. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 618–626.
- Shamsi, A., Asghari, M. H., Afshar, P. (2021). EfficientNet-based deep transfer learning model for diagnosis of COVID-19 in computed tomography images. *Frontiers in Artificial Intelligence*, 4, 645299.
- Shin, H. C., Roth, H. R., Gao, 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.
- Shorten, C., Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6, 60.

- Simonyan, K., Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations (ICLR)*.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., (2014), Dropout: A Simple Way to Prevent Neural Networks from Overfitting, *Journal of Machine Learning Research*, 15(56):1929–1958
- Szegedy, C., Liu, W., Jia, Y. (2015). Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1–9.
- Tan, M., Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 6105–6114.
- Tajbakhsh, N., Shin, J.Y., Gurudu, S.R., Hurst, R.T., Kendall, C.B., Gotway, M.B., Jianming, L. (2016), Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? *IEEE Trans Med Imaging*. 2016 May;35(5):1299-1312.
- Ting, D. S. W., Cheung, C. Y., Lim, G. (2019). Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA*, 318(22), 2211–2223.
- Tjoa, E., Guan, C. (2020). A survey on explainable artificial intelligence (XAI): Towards medical XAI. *IEEE Transactions on Neural Networks and Learning Systems*, 32(11), 4793–4813.
- Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56.
- Vaswani, A., Shazeer, N., Parmar, N. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.
- Wang, J., Chen, X. (2022). DenseNet-based transfer learning for automated multi-class retinal disease classification. *IEEE Access*, 10, 35412–35423.
- Woo, S., Park, J., Lee, J. Y., Kweon, I. S. (2018). CBAM: Convolutional block attention module. *Proceedings of the European Conference on Computer Vision (ECCV)*.
- Wong, T. Y., Cheung, C. M. G., Larsen, M., Sharma, S., Simo, R. (2023). Diabetic retinopathy. *Nature Reviews Disease Primers*, 9(1), 1–21.
- Yim, J., Chopra, R., Spitz, T. (2020). Predicting conversion to wet age-related macular degeneration using deep learning. *Nature Medicine*, 26(6), 892–899.
- Yun, S., Han, D., Oh, S. J. (2019). CutMix: Regularization strategy to train strong classifiers with localizable features. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 6023–6032.
- Zhang, H., Cisse, M., Dauphin, Y. N., Lopez-Paz, D. (2018). mixup: Beyond empirical risk minimization. *International Conference on Learning Representations (ICLR)*.
- Zhou, L., Wang, S. (2022). Automated multi-disease detection in fundus images using deep learning. *Medical Image Analysis*, 78, 102403.

Zhu, J. Y., Park, T., Isola, P., Efros, A. A. (2022). Unpaired image-to-image translation using cycle-consistent adversarial networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10), 6144–6159.



ÖZGEÇMİŞ

