

KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

PREMATÜRE RETİNOPATİ (ROP) TEŞHİSİNDE DERİN ÖĞRENME
TABANLI RETİNA GÖRÜNTÜ SEGMENTASYONU VE SINIFLANDIRMASI

YÜKSEK LİSANS TEZİ

Ayhan MURAT

KASIM 2025
TRABZON



KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

**PREMATÜRE RETİNOPATİ (ROP) TEŞHİSİNDE DERİN ÖĞRENME TABANLI
RETİNA GÖRÜNTÜ SEGMENTASYONU VE SINIFLANDIRMASI**

Ayhan MURAT

ORCID:

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünce
"BİLGİSAYAR YÜKSEK MÜHENDİSİ"
Unvanı Verilmesi İçin Kabul Edilen Tezdir.

Tezin Enstitüye Verildiği Tarih : 24 / 10 / 2025

Tezin Savunma Tarihi : 19 / 11 / 2025

Tez Danışmanı : Doç. Dr. Vasif NABİYEY

ORCID :

Trabzon 2025

**KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**Bilgisayar Mühendisliği Anabilim Dalında
Ayhan MURAT Tarafından Hazırlanan**

**PREMATÜRE RETİNOPATİ (ROP) TEŞHİSİNDE DERİN ÖĞRENME
TABANLI RETİNA GÖRÜNTÜ SEGMENTASYONU VE SINIFLANDIRMASI**

**başlıklı bu çalışma, Enstitü Yönetim Kurulunun 30 / 10 / 2025 gün ve 2126 sayılı
kararıyla oluşturulan jüri tarafından yapılan sınavda
YÜKSEK LİSANS TEZİ
olarak kabul edilmiştir.**

Jüri Üyeleri

Başkan : Doç. Dr. Vasif NABİYEV

Üye : Doç. Dr. Selen AYAS

Üye : Doç. Dr. Ali Kürşat ERÜMİT

Prof. Dr. Asim KADIOĞLU

Enstitü Müdürü

ÖNSÖZ

Prematüre Retinopatisi (ROP), prematüre bebeklerde görülen, erken tanı gerektiren ve körlüğe yol açabilen ciddi bir göz hastalığıdır. Hastalığın teşhisi, retina damar yapılarının morfolojik özelliklerine, vasküler dilatasyona ve avasküler bölgelerin tespitine dayandığından son derece karmaşık bir süreçtir. Deneyimli oftalmologlar arasında bile evreleme tutarsızlıkları görülebilmektedir. Kalıcı görme kaybını önlemek için erken ve doğru teşhis kritik önem taşımaktadır. Bu tez çalışmasında, derin öğrenme yöntemleriyle retina görüntülerinin otomatik segmentasyonu ve sınıflandırılmasını gerçekleştiren bir ROP teşhis sistemi geliştirilmiştir.

Çalışmamın başından sonuna kadar akademik bilgi ve tecrübesiyle bana eşlik eden, her zorlu aşamada destek olan danışmanım Doç. Dr. Vasif NABİYEV'e derin teşekkürlerimi sunarım. Prematüre Retinopatisi hakkında değerli bilgilerini benimle paylaşan Doç. Dr. Murat GÜNEY'a ve katkılarıyla beni destekleyen diğer tüm hocalarıma teşekkürü bir borç bilirim. Ayrıca bu süreçte anlayışlı yaklaşımları ve destekleriyle yanımda olan çalışma arkadaşlarıma teşekkür ederim.

Bu zorlu yolculukta sevgileri bana güç veren, her anımda yanımda hissettiğim biricik eşime ve çocuklarıma; hayatımın her döneminde olduğu gibi bu süreçte de manevi ve maddi desteklerini hiç eksik etmeyen, kıymetli annem Türkan MURAT'a, babam Şevket MURAT'a sevgilerimi sunarım.

Her zaman benimle olduğunu bildiğim, varlığını hep yanımda hissettiğim canım ablam Aysun MURAT'a en derin şükran ve minnettarlığımı ifade ederim.

Ayhan MURAT

Trabzon 2025

TEZ ETİK BEYANNAMESİ

Yüksek Lisans Tezi olarak sunduğum “Prematüre Retinopati (ROP) Teşhisinde Derin Öğrenme Tabanlı Retina Görüntü Segmentasyonu ve Sınıflandırması” başlıklı bu çalışmayı baştan sona kadar danışmanım Doç. Dr. Vasif NABİYEY’in sorumluluğunda tamamladığımı, verileri/örnekleri kendim topladığımı, deneyleri/analizleri ilgili laboratuvarlarda yaptığımı/yaptırdığımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim. 19/11/2025

Ayhan MURAT

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ	III
TEZ ETİK BEYANNAMESİ.....	IV
İÇİNDEKİLER	V
ÖZET	VII
SUMMARY	VIII
ŞEKİLLER DİZİNİ	IX
TABLolar DİZİNİ.....	XI
SEMBOLLER VE KISALTMALAR DİZİNİ	XII
1. GENEL BİLGİLER	1
1.1. Giriş.....	1
1.2. Tezin Amacı ve Kapsamı	2
1.3. Prematüre Retinopatisi (ROP).....	5
1.3.1. ROP'un Tanımı ve Klinik Önemi	5
1.3.2. ROP Evreleri ve Tanı Kriterleri	5
1.3.3. ROP Tanısında Karşılaşılan Zorluklar	6
1.3.4. Görüntüleme Teknikleri (RetCam vb.)	6
1.4. Retina Görüntülerinde Segmentasyonun Önemi	7
1.4.1. Retina Damar Segmentasyonu: Tanım ve Zorluklar.....	7
1.4.2. Damar Yapısının ROP Tanısındaki Rolü.....	8
1.4.3. Otomatik Segmentasyonun Klinik Katkısı.....	8
1.5. Tıbbi Görüntü Segmentasyonunda Derin Öğrenme Yöntemleri.....	8
1.5.1. Evrişimsel Sinir Ağları (ESA)	8
1.5.2. U-Net: Biyomedikal Görüntü Segmentasyonu İçin Temel Mimari.....	16
1.5.3. Attention U-Net ve Dikkat Mekanizmaları.....	20
1.6. Model Performans Ölçütleri.....	24
1.6.1. Doğruluk ,Duyarlılık,Kesinlik	24
1.6.2. F1 Skoru (F1 Score)	25
1.6.3. Intersection over Union (IoU)	25
1.6.4. ROC Eğrisi ve AUC	26
1.6.5. Macro ve Micro Ortalama Metrikler	26

1.6.6.	Karışıklık Matrisi (Confusion Matrix)	27
1.7.	Literatür Taraması	28
2.	YAPILAN ÇALIŞMALAR	35
2.1.	Retinal Damar Segmentasyonu Aşaması	37
2.1.1.	VesselSegNet: Önerilen Segmentasyon Ağı	37
2.1.2.	VesSegNet Eğitimi için Kullanılan Veri Seti: DRIVE	42
2.1.3.	Hibrit Kayıp Fonksiyonu Bileşenleri	44
2.2.	Segmentasyon Destekli ROP Sınıflandırma Yaklaşımı	50
2.2.1.	FusionROPNet: Önerilen Sınıflandırma Ağı	50
2.2.2.	FusionROPNet Eğitimi İçin Kullanılan Veri Seti: FARFUM-ROP	51
2.2.3.	ROP Veri Setleri: Mevcut Kaynaklar, Zorluklar ve Kısıtlamalar	52
2.2.4.	Öznitelik Çıkarıcı Omurga: EfficientNet-B0 Mimarisi	53
2.2.5.	EfficientNet-B0 Tercih Nedenleri	54
2.2.6.	Transfer Öğrenme	55
2.2.7.	Özellik Füzyonu ve Sınıflandırma Katmanı	55
3.	BULGULAR VE TARTIŞMA	57
4.	SONUÇLAR	65
5.	ÖNERİLER	67
6.	KAYNAKLAR	68
7.	EKLER	74
ÖZGEÇMİŞ		

Yüksek Lisans Tezi

ÖZET

PREMATÜRE RETİNOPATİ (ROP) TEŞHİSİNDE DERİN ÖĞRENME TABANLI RETİNA GÖRÜNTÜ SEGMENTASYONU VE SINIFLANDIRMASI

Ayhan MURAT

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı
Danışman: Doç. Dr. Vasif NABİYEV
2025, 79 Sayfa

Bu tez çalışması, prematüre retinopatisi (ROP) tanısında kullanılmak üzere iki aşamalı, derin öğrenme temelli bir yaklaşım önermektedir. Çalışmanın ilk aşamasında, retinal damar yapılarının hassas segmentasyonu için Attention U-Net mimarisini temel alan VesSegNet geliştirilmiş ve dikkat mekanizmalarıyla güçlendirilmiştir. Model, DRIVE veri seti üzerinde eğitilmiş; farklı damar yapılarının etkili şekilde öğrenilmesi için hibrit bir kayıp fonksiyonu kullanılmıştır. Elde edilen segmentasyon çıktıları, özellikle ince damar yapılarının ayrıntılı olarak ortaya konmasını sağlamış ve daha önce segmente edilmemiş FARFUM-ROP veri setinde başarılı damar haritaları üreterek ikinci aşamada kullanılmak üzere transfer edilmiştir. İkinci aşamada, segmente edilen damar maskeleri ve orijinal ROP fundus görüntüleri birlikte değerlendirilerek ROP plus hastalığı sınıflandırılması (no-plus, pre-plus, plus) hedeflenmiştir. Bu amaçla, EfficientNet tabanlı çift kollu (dual-branch) bir füzyon mimarisi tasarlanmıştır. Deneysel sonuçlar, VesSegNet'in DRIVE test setinde 0.807 F1 skoru, FusionROPNet'in ise ROP plus sınıflandırmasında 0.8137 makro F1 skoru ile yüksek performans gösterdiğini ortaya koymaktadır. ROP tanısında damar morfolojisinin klinik önemi göz önünde bulundurularak, segmentasyon bilgisi sınıflandırma sürecine entegre edilmiştir.

Anahtar Kelimeler: Prematüre Retinopatisi, Retinal Damar Segmentasyonu, Attention U-Net, Hibrit Kayıp Fonksiyonu, Çift Kollu Füzyon Mimarisi

Master's Thesis

SUMMARY

DEEP LEARNING-BASED RETINAL IMAGE SEGMENTATION AND CLASSIFICATION FOR DIAGNOSIS OF RETINOPATHY OF PREMATURITY (ROP)

Ayhan MURAT

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Department of Computer Engineering
Supervisor: Assoc. Prof. Vasif NABIYEV
2025, 79 Pages

This thesis proposes a two-stage, deep learning-based approach for the diagnosis of retinopathy of prematurity (ROP). In the first stage, VesSegNet, based on the Attention U-Net architecture, was developed and enhanced with attention mechanisms for precise segmentation of retinal vessel structures. The model was trained on the DRIVE dataset, utilizing a hybrid loss function to effectively learn different vessel structures. The obtained segmentation outputs enabled detailed representation of fine vessel structures and successfully generated vessel maps on the previously unsegmented FARFUM-ROP dataset, which were then transferred for use in the second stage. In the second stage, the segmented vessel masks and original ROP fundus images were evaluated together to classify ROP plus disease (no-plus, pre-plus, plus). For this purpose, an EfficientNet-based dual-branch fusion architecture was designed. Experimental results reveal that VesSegNet achieves an F1 score of 0.807 on the DRIVE test set, whereas FusionROPNet attains a macro F1 score of 0.8137 for ROP plus disease classification, demonstrating robust performance. Given the clinical importance of vascular morphology in ROP diagnosis, segmentation information was incorporated into the classification pipeline.

Keywords: Retinopathy of Prematurity, Retinal Vessel Segmentation, Attention U-Net, Hybrid Loss Function, Dual-Branch Fusion Architecture

ŞEKİLLER DİZİNİ

	<u>Sayfa No</u>
Şekil 1. Önerilen iki aşamalı mimari	4
Şekil 2. Evrişim katmanında özellik çıkarımı.....	10
Şekil 3. 2×2 Max Pooling işlemi ile 7×7 boyutundaki özellik haritasının 3×3 boyutuna indirgenmesi.....	11
Şekil 4. Evrişimsel sinir ağlarında çekirdek, adım ve dolgu parametrelerinin görselleştirilmesi.....	13
Şekil 5. Aktivasyon fonksiyonları: (a) ReLU (Rectified Linear Unit) fonksiyonu, (b) Leaky ReLU fonksiyonu.....	14
Şekil 6. Dropout regularization tekniğinin görselleştirilmesi.....	15
Şekil 7. U-Net mimarisi	16
Şekil 8. Attention U-Net mimarisi	21
Şekil 9. VesSegNet mimarisinde kullanılan kanal dikkat mekanizmasının (channel attention) yapısı.....	22
Şekil 10. VesSegNet mimarisinde kullanılan uzamsal dikkat mekanizmasının (spatial attention) yapısı.....	23
Şekil 11. EfficientNet model ölçeklendirme stratejileri. (a) temel ağ mimarisi, (b) genişlik ölçeklendirme, (c) derinlik ölçeklendirme, (d) çözünürlük ölçeklendirme, (e) bileşik ölçeklendirme (compound scaling)	29
Şekil 12. 1.Aşamada önerilen VesSegNet mimarisi iç yapısı ve eğitim bileşenler	35
Şekil 13. 2.Aşamada önerilen FusionROPNet mimarisi iç yapısı ve eğitim bileşenleri.....	35
Şekil 14. U-Net mimarilerinin test veri kümesi F1 skoru karşılaştırması	39
Şekil 15. U-Net mimarilerinin test veri kümesi IoU karşılaştırması	40
Şekil 16. DRIVE veri seti örneği: (a) fundus görüntüsü, (b) damar segmentasyon maskesi	43
Şekil 17. Retina görüntüsünün ön işleme aşamaları: (1) orijinal RGB fundus, (2) gri tonlamaya dönüştürülmüş görüntü, (3) CLAHE ile kontrast iyileştirmesi, (4) gamma düzeltme ($\gamma=1.2$) uygulanmış görüntü-kontrast optimizasyonu, (5) orijinal maske	43
Şekil 18. DRIVE test veri setinden seçilen örnek bir fundus görüntüsünün VesSegNet mimarisi tarafından segmentasyon sonuçlarının görsel analizi.....	47
Şekil 19. VesSegNet modelinin FAPFUM-ROP veri setindeki segmentasyon sonuçları: (a,c) orijinal fundus görüntüleri, (b,d) tahmin edilen damar maskeleri.....	48

Şekil 20.	Farklı ROP evrelerindeki ve sağlıklı kontrol grubundaki fundus görüntülerinin VesSegNet segmentasyon sonuçları	49
Şekil 21.	FARFUM-RoP veri setinden alınan retina görüntüleri	51
Şekil 22.	VesSegNet modelinin eğitim ve doğrulama kayıp grafiği	57
Şekil 23.	VesSegNet modelinin eğitim sürecinde doğruluk değerlerinin değişimi	58
Şekil 24.	Modelin doğrulama kümesinde F1 skoru ve IoU değerlerinin eğitim süreci boyunca epochlara göre değişimi	58
Şekil 25.	Modelin doğrulama kümesinde kesinlik (precision) ve duyarlılık (recall) metriklerinin eğitim süreci boyunca epochlara göre değişimi	59
Şekil 26.	Test veri setinde elde edilen karmaşıklık matrisi	60
Şekil 27.	FusionROPNet'in doğrulama kümesinde sınıf bazlı F1 skorları (no-plus, pre-plus, plus) ve makro F1 değerinin karşılaştırılması	60
Şekil 28.	FusionROPNet için normalize edilmiş karışıklık matrisi	61
Şekil 29.	FusionROPNet'in modelinin test verisi üzerindeki performans metrikleri	62

TABLÖLAR DİZİNİ

	<u>Sayfa No</u>
Tablo 1. U-Net varyantlarının segmentasyon başarısının karışıklık matrisi analizi.....	40
Tablo 2. U-Net varyantlarının model eğitiminde kullanılan hiperparametreler	41
Tablo 3. Kayıp fonksiyonlarının karışıklık matrisi analizi	46
Tablo 4. EfficientNet-B0 mimarisinin katman yapısı ve parametreleri.....	54
Tablo 5. Tez kapsamında önerilen FusionROPNet modelinin literatürdeki güncel ROP tespit sistemleriyle performans karşılaştırması.....	63
Tablo 6. Önerilen iki aşamalı model ile doğrudan RGB tabanlı tek aşamalı modelin ROP sınıflandırma performansları.....	64

SEMBOLLER VE KISALTMALAR DİZİNİ

α, β	: Hybrid Loss içindeki BCE ve Dice kayıp ağırlıkları
n	: Örneklem büyüklüğü (Görüntü sayısı)
L	: Toplam Kayıp Fonksiyonu (Hybrid Loss)
γ	: Focal Loss Fonksiyonundaki odaklanma parametresi
ADAMW	: Adam Optimize Edicisinin Ağırlık Azaltma (Weight Decay)
BCE	: İkili Çapraz Entropi (Binary Cross-Entropy)
CLAHE	: Kontrast Limitli Adaptif Histogram Eşitleme
ESA	: Evrişimli Sinir Ağları
ICROP	: Uluslararası ROP Sınıflandırma Sistemi
IoU	: Birleşim Üzerinden Kesişim Oranı (Jaccard Benzerliği)
ReLU	: Doğrultulmuş Doğrusal Birim
ROP	: Prematüre Retinopatisi

1. GENEL BİLGİLER

1.1. Giriş

Retinadaki kan damarları, canlı insan bedeninde doğrudan incelenebilen tek terminal damar yapısı olarak benzersiz bir konuma sahiptir (Lucas vd., 2013). Bu özelliği sayesinde retina görüntüleri, sistemik hipertansiyon, diyabetik retinopati ve özellikle prematüre retinopatisi (ROP) gibi ciddi oftalmik hastalıkların erken teşhisi ve takibi için kritik bir biyobelirteç işlevi görmektedir. Damar morfolojisindeki düzensizlikler, düşük kontrast ve gürültü gibi etkenlerle birleşerek damar yapılarının belirginliğini azaltmaktadır. Ayrıca damar bölgelerinin arka plana göre oldukça küçük bir alan kaplaması, sınıf dengesizliği problemine neden olmaktadır. Söz konusu kısıtlamalar aynı zamanda geliştirilen otomatik sistemlerin genelleme kapasitesini zorlamaktadır. Özellikle medikal görüntü verileri üzerine geliştirilen derin öğrenme modellerinde en büyük sorun, eğitildiği veri setinde yüksek doğruluk sağlarken, daha önce görmediği ve farklı dağılıma sahip veri setlerinde genelleme yapamamasıdır. Bu zorlukların üstesinden gelmek ve genelleme kapasitesini artırmak amacıyla, bu tez çalışması yenilikçi mimari ve yöntem üzerine temellendirilmiştir.

VesSegNet; Damar segmentasyonu için önerilen bu mimari, özellikle retina damar segmentasyonu problemine genelleme kapasitesi yüksek olacak şekilde tasarlanmıştır. Modelin mimarisi, farklı veri dağılımlarına uyum sağlayabilme yeteneği sayesinde, kazanılan bilginin modelin daha önce karşılaşmadığı veri setlerine etkin bir şekilde aktarılmasını sağlamaktadır. Eğitim sürecinde kullanılan hibrit kayıp fonksiyonu (HybridLoss), Binary Cross-Entropy (BCE), Dice Loss ve Focal Loss bileşenlerini beraber kullanarak modelin sadece piksel doğruluğuna değil, aynı zamanda zorlu ve az temsil edilen damar piksellerine de odaklanmasını sağlamıştır. Bu yaklaşım, genelleme kapasitesinin artırılmasında kritik bir rol oynamış ve literatürdeki genelleme sorununa yönelik bir yaklaşım önermektedir.

FusionROPNet; ROP sınıflandırması için önerilen bu mimari, sınıflandırma başarısını artırmak için Dual Fusion yaklaşımı kullanmaktadır. Bu yöntemde, modele yalnızca orijinal fundus görüntüsü değil, aynı zamanda segmentasyon görevi için eğitilen ön model olan VesSegNet modelinden elde edilen segmentasyon haritalarında ikinci bir girdi olarak

FusionROPNet modeline dahil edilmiştir. Bu girdi akışı sayesinde, FusionROPNet modeli, ham görüntüdeki bağlamsal bilgileri (dokusal ve global yapı) korurken, VesSegNet'in çıktı olarak verdiği damar yapısı gibi kritik ayrıştırma bilgileri üzerinde de çalışabilmektedir. Bu strateji, özellikle zorlu görüntülerde, modelin hem ince ayrıntıları yakalama hem de yanlış tahmin edilen bölgeleri düzeltme başarısını maksimize etmiş, böylece ince ayrıntıları yakalama yeteneğini artırmıştır. Bu yaklaşım sayesinde geliştirilen mimari, ROP gibi zorlu veri setleri üzerinde yüksek doğruluk oranları elde etmiştir. Önerilen iki aşamalı mimari, otomatik teşhis uygulamalarında uzmanlara önemli klinik destek sağlama potansiyeline sahiptir.

1.2. Tezin Amacı ve Kapsamı

Bu tez çalışması, Prematüre Retinopatisi (ROP) teşhisinde karşılaşılan klinik kısıtları gidermek için yüksek doğruluklu ve özgün bir derin öğrenme mimarisi önermeyi amaçlamaktadır. Bu kapsamda, ROP için yeterli büyüklükte etiketli veri setinin bulunmaması ve damar morfolojisinin ROP tanısındaki kritik önemi göz önünde bulundurularak, damar yapısını açık ve sistematik biçimde modele entegre eden iki aşamalı ve uçtan uca bir derin öğrenme çerçevesi önerilmiştir. Çalışmanın ilk aşamasında VesSegNet mimarisi, fundus veri setlerinden türetilen damar yapıları üzerinde eğitilmiştir. Takip eden aşamada, FusionROPNet mimarisi entegre edilerek damar bilgileri ve orijinal görüntüler füzyona tabi tutulmuş; bu sayede sınırlı sayıdaki veri seti zenginleştirilerek etkin bir ROP sınıflandırması gerçekleştirilmiştir. Bu iki aşamalı strateji sayesinde, sınırlı ROP verisi ile çalışılırken, damar yapısının klinik önemi modele açıkça yansıtılmış ve genelleme kapasitesi performans metriklerinde görüldüğü gibi artırılmıştır.

Önerilen iki aşamalı derin öğrenme çerçevesinin, tek aşamalı sınıflandırma yaklaşımına göre performans üstünlüğü deneysel olarak doğrulanmıştır. Test setinde, tek aşamalı mimari 0.7843 F1-macro skoru elde ederken, VesSegNet segmentasyon çıktılarını kullanılan iki aşamalı strateji 0.8137 (%3.75 artış) F1-macro skoruna ulaşmış, Cohen's Kappa skorunda 0.6614'ten 0.7360'a (~%11.3 artış) anlamlı bir performans artışı sağlamıştır. Bu iyileşme, birinci aşamada retinal damar yapısının hassas bir şekilde segmente edilmesi sayesinde, ikinci aşamadaki sınıflandırma modelinin klinik açıdan en önemli özelliklere (damar kıvrımlılığı, genişlik, çatal noktaları gibi) odaklanabilmesinden kaynaklanmaktadır. Özellikle ROP gibi sınırlı etiketli veriye sahip tıbbi görüntüleme

problemlerinde, bu iki aşamalı yaklaşım, modelin ilgisiz arka plan bilgisinden arındırılmış, damar morfolojisi üzerinden öğrenmesini sağlayarak genelleme kapasitesini artırmıştır.

Mevcut ROP teşhis süreci, tecrübeli uzmanlara olan yüksek bağımlılığı, teşhisin uzun sürmesi ve özellikle uzman eksikliğinin yaşandığı coğrafi bölgelerde veya merkezlerde tanı kalitesinin düşme riski gibi önemli kısıtlamalara sahiptir. Bu kapsamda tezin temel hedefleri şunlardır:

- Klinik düzeyde yardımcı sistem geliştirilmesi: Uzman oftalmologlara ROP plus hastalığı sınıflandırılması ve takibinde hızlı, objektif ve güvenilir segmentasyon destekli sınıflandırma sonuçları sunarak karar verme süreçlerini desteklemek.
- Genelleme yeteneği yüksek mimari tasarımı: Derin öğrenme modellerinin en büyük handikapı olan, eğitildiği veri setinde iyi sonuç verip, modelin daha önce görmediği veri setlerinde oluşan genelleme sorununu aşmak. Bu amaçla, VesSegNet ve FusionROPNet gibi mimariler ile hibrit kayıp fonksiyonları kullanılarak sistemin klinik ortamlarda kullanılabilirliği artırılmıştır. Binary Cross-Entropy (BCE), Dice Loss ve Focal Loss bileşenlerini birleştiren hibrit kayıp fonksiyonu (Hybrid Loss) ile sınıf dengesizliğine etkin bir çözüm sunarak, segmentasyon problemlerinde yüksek Dice skoru, IoU (Intersection over Union), hassasiyet ve F1 değerleri elde etmek suretiyle literatüre katkıda bulunmak.

Tezin birinci bölümünde öncelikle ROP hastalığının tanımı, klinik önemi, evreleri, plus hastalığı sınıflandırılması ve teşhis sürecinde karşılaşılan zorluklar detaylıca incelenmektedir.

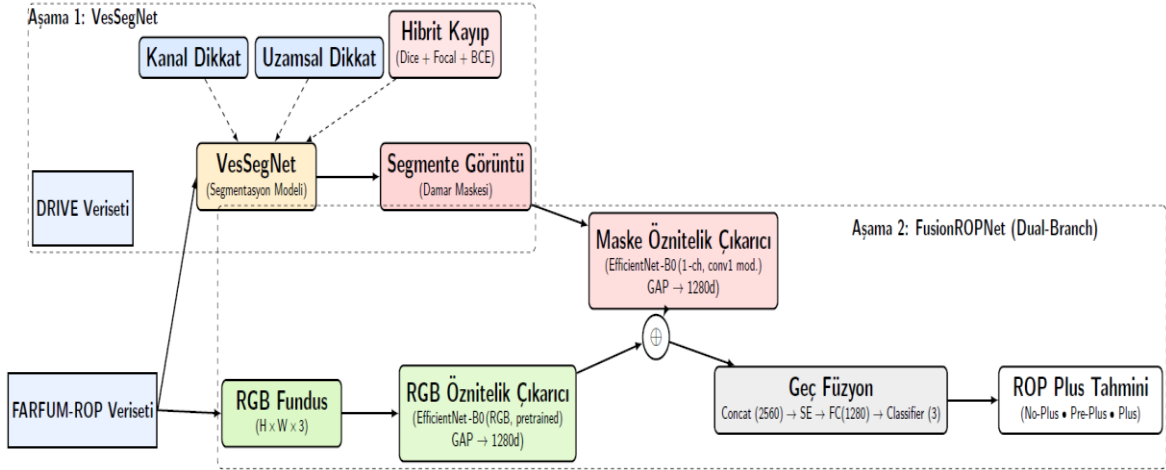
Daha sonra, retina damar segmentasyonunun klinik katkısı ve bu alandaki zorluklar ele alınarak, çözümde kullanılan derin öğrenme yaklaşımları ile CNN, U-Net, Attention U-Net gibi segmentasyon mimarileri tanıtılmaktadır. Bölüm, geliştirilen modellerin performansını değerlendirmek için kullanılacak doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall), F1-skoru, ROC eğrisi, AUC ve karışıklık matrisi gibi temel performans ölçütlerinin tanımlanması ve ilgili literatür taraması ile tamamlanmaktadır.

Tezin ikinci bölümü çalışmanın metodolojik yeniliklerini ve uygulama aşamalarını içermektedir. Bu bölümde, VesSegNet segmentasyon ağının mimarisi, BCE, Dice Loss ve Focal Loss bileşenlerini içeren Hybrid Loss fonksiyonu ile eğitim süreci açıklanmıştır. VesSegNet tarafından üretilen segmentasyon çıktılarının FARFUM-ROP veri setine transfer süreci ve bu transferin nitel analizi ele alınmıştır. Devamında, ROP plus hastalığı sınıflandırılması için çift kollu (dual-branch) yapıya sahip, FusionROPNet mimarisi

tanıtılmış; VesSegNet çıktılarını girdi olarak kullanan füzyon yaklaşımı ve EfficientNet-B0 omurgasının kullanımı açıklanmıştır.

Tezin üçüncü bölümünde, geliştirilen iki aşamalı sistemin deneysel sonuçları kapsamlı bir şekilde sunulmaktadır. VesSegNet için DRIVE veri seti üzerinde elde edilen segmentasyon performansı, accuracy, Dice katsayısı, IoU ve F1-skoru gibi nicel metriklerle değerlendirilmiş; eğitim ve doğrulama kayıplarının değişimi, öğrenme eğrileri ve farklı veri dağılımlarına genelleme kapasitesi detaylı olarak analiz edilmiştir. Özellikle ince damar yapılarının segmentasyonu ve Hybrid Loss fonksiyonu'nun sınıf dengesizliği sorununa etkisi görsel örneklerle desteklenmiştir. FARFUM-ROP veri setine transfer edilen segmentasyon çıktılarının kalitesi, nitel ve nicel yöntemlerle incelenmiştir. FusionROPNet mimarisinin ROP sınıflandırma performansı ise karışıklık matrisi analizi, Macro F1-skoru, hassasiyet ve duyarlılık metrikleri ile değerlendirilmiş; çift kollu füzyon yaklaşımının ve segmentasyon maskesinin sınıflandırma kararına olan katkısı incelenmiştir.

Bu kapsamlı deneysel bulgular, dördüncü bölümde sonuçlar olarak özetlenmiş, literatüre yapılan yöntemsel katkılar ve klinik uygulamalara potansiyel etkisine değinilmiştir. Son olarak beşinci bölümde, gelecek çalışmalar için öneriler ve olası araştırma yönleri sunulmuştur.



Şekil 1. Önerilen iki aşamalı mimari

Şekil 1, önerilen iki aşamalı mimariyi göstermektedir. İlk aşamada VesSegNet modeli, DRIVE veri seti ile eğitilerek retinal damar segmentasyonu gerçekleştirmektedir. İkinci aşamada ise FusionROPNet mimarisi, FARFUM-ROP veri setine ait görüntülerden elde

edilen damar maskeleri ve orijinal RGB fundus görüntülerini birleştirerek ROP plus hastalık sınıflandırması yapmaktadır.

1.3. Prematüre Retinopatisi (ROP)

1.3.1. ROP'un Tanımı ve Klinik Önemi

Prematüre retinopatisi (Retinopathy of Prematurity, ROP), prematüre doğan bebeklerde retinal damarlanmanın tamamlanamaması veya anormal seyretmesi sonucunda ortaya çıkan, anormal damar oluşumu ve fibro-vasküler doku çoğalmasıyla seyreden bir hastalıktır (Gilbert, 2008). Hastalık özellikle düşük doğum ağırlıklı ve erken doğan bebeklerde görülür (Blencowe vd., 2013). ROP, tedavi edilmediği takdirde retina dekolmanı ve kalıcı görme kaybı ile sonuçlanabilir. Hastalığın erken evrelerde tespit edilmesi ve uygun tedavi yaklaşımlarının (lazer ablasyon tedavisi veya vasküler endotelyal büyüme faktörü inhibitörleri gibi) uygulanması, klinik seyrin olumlu yönde değiştirilmesine olanak tanımaktadır. Hastalığın önlenabilir olması ve teşhisinin yüksek düzeyde uzmanlık gerektirmesi, ROP'u bilgisayar destekli tanı sistemleri için önemli bir araştırma konusu haline getirmektedir (Fierson vd., 2018).

1.3.2. ROP Evreleri ve Tanı Kriterleri

ROP, Uluslararası ROP Sınıflandırma Sistemi (ICROP)'na göre beş evrede değerlendirilir (International Committee for the Classification of Retinopathy of Prematurity, 2005).

- Evre 1: Retina ile avasküler alan arasında demarkasyon hattı
- Evre 2: Bu hattın kabarık sırt halini alması
- Evre 3: Fibrovasküler proliferasyonların oluşması
- Evre 4: Parsiyel retina dekolmanı
- Evre 5: Total retina dekolmanı

Ayrıca, retinal damar değişiklikleri plus hastalığı varlığına göre sınıflandırılmaktadır. Plus hastalığı, retinanın arka segmentinde en az iki kadranda arteriyel genişleme ve venöz tortuözitenin (kıvrımlanma) artması ile karakterize edilir. Bu bulgular tam gelişmemişse

durum pre-plus olarak tanımlanır. Plus hastalığının varlığı, tedavi kararında kritik bir faktördür ve hastalığın seyrini gösterir. Klinik pratikte, plus hastalığının objektif değerlendirilmesi, gözlemciler arası değişkenlik nedeniyle zorluklar içermektedir ve bu nedenle bilgisayar destekli tanı sistemlerine olan ihtiyaç artmaktadır.

1.3.3. ROP Tanısında Karşılaşılan Zorluklar

Prematüre bebeklerde ROP kontrolü için belirli aralıklarla uzman tarafından muayene edilme gerekliliği, oftalmologların iş yükünü önemli ölçüde artırmaktadır. Aynı zamanda ROP alanında uzman hekim sayısının yetersizliği, bu sürecin etkin yönetimini güçleştirmektedir.

Ayrıca, hastalığın evre geçişlerinde ve belirsiz bulgular taşıyan vakalarda, klinisyenler arasında görüş ayrılıkları ortaya çıkmakta ve bu durum teşhis tutarsızlığına yol açabilmektedir. Retinadaki ince damar değişikliklerinin doğru şekilde ayırt edilebilmesi ise yalnızca ileri düzey uzmanlık değil, aynı zamanda uzun süreli klinik deneyim de gerektirmektedir. Bu durum, ROP tanısının standardizasyonunu güçleştiren önemli bir engel oluşturmaktadır. Özellikle büyük merkezler dışında kalan hastanelerde gerekli donanım ve uzman sayısının yetersizliği, düzenli taramaların aksamasına veya gecikmesine yol açabilmektedir. ROP alanında deneyimli retina uzmanlarının her merkezde bulunmaması, tanının doğruluğunu ve tedaviye erişimi sınırlayan ek bir sorun olarak karşımıza çıkmaktadır.

1.3.4. Görüntüleme Teknikleri (RetCam vb.)

ROP tanısı konulması sırasında yaygın olarak kullanılan sistemlerden biri geniş açılı dijital fundus kamerası olan RetCam'dir. Bu sistem, yüksek çözünürlüklü görüntü sağlayarak hem tanı hem de tedavi sürecinde karşılaştırmalı değerlendirme imkânı sunar. Son yıllarda taşınabilir kamera sistemleri ve yapay zeka destekli analiz yöntemleri de tanısal süreçlerde kullanılmaya başlanmıştır.

Son yıllarda taşınabilir fundus kameralar, özellikle donanım ve uzman eksikliği bulunan sağlık merkezleri için önemli bir alternatif olmaktadır. Bu cihazlar, mobilite avantajı sayesinde prematüre bebeklerin yoğun bakım ünitelerinde veya küçük hastanelerde de ROP taramalarının gerçekleştirilmesine olanak tanır. Taşınabilir sistemlerin kullanımı, bebeklerin

uzun süreli transportasyonundan kaynaklanan riskleri ortadan kaldırarak hasta güvenliğini artırmaktadır. Ayrıca, bu cihazların maliyet etkinliği ve kolay kullanım özellikleri, özellikle gelişmekte olan ülkelerde ROP tarama programlarının yaygınlaştırılmasına katkı sağlamaktadır. Günümüzde, akıllı telefon tabanlı fundus görüntüleme sistemleri ve bulut tabanlı analiz platformları ile kullanılan taşınabilir sistemler, tele-oftalmoloji uygulamalarının gelişmesine de zemin hazırlamaktadır.

1.4. Retina Görüntülerinde Segmentasyonun Önemi

1.4.1. Retina Damar Segmentasyonu: Tanım ve Zorluklar

Retina damar segmentasyonu, fundus görüntülerinde kan damarlarının piksel düzeyinde tespit edilerek damar olmayan bölgelerden ayırt edilmesi sürecidir (Fraz vd., 2012). Bu süreç, damar sınırlarının belirlenmesi, ince kılcal damarların tespiti ve dallanma noktalarının doğru şekilde çıkarılması gibi adımları içerir. Retina damar segmentasyonu bilgisayarlı görme için oldukça zorlu bir problemdir. Bunun sebepleri:

- Kalın damarlar daha kolay bir şekilde ayırt edilebilirken, ince damarların düşük kontrast ve gürültü sebebi ile ayırt edilmesi zordur.
- Fundus görüntülerde aydınlatma homojen olmayabilir. Bazı bölgelerde damar arka plandan çok az ayrışır.
- Görüntüleme sırasında oluşan yansıma, bulanıklık, kamera kaynaklı bozulmalar, hasta hareketleri gibi sebeplerden fundus görüntülerinde bozulma ve farklılıklar oluşabilmektedir.
- Damarların kompleks geometrisinin olması, özellikle damar kavşakları ve çap değişimlerinin tespiti zordur.
- En önemli problemlerden biri genelleme problemidir. Bir model eğitildiği veri setinde başarılı olurken farklı veri setlerinde kullanılan cihaz, çözünürlük veya hasta popülasyonundan dolayı başarılı olamayabilir.

Sonuç olarak retina damar segmentasyonu klinik öneme sahip olmasına rağmen, yukarıda belirtilen zorluklardan dolayı bilgisayarlı görüntü işleme ve derin öğrenme yöntemleri için geliştirmeye açık bir araştırma alanıdır.

1.4.2. Damar Yapısının ROP Tanısındaki Rolü

ROP tanısında damar yapısının analizi, plus hastalığının belirlenmesinde kritik bir adımdır. Özellikle gözdeki atardamarların genişlemesi ve toplardamarların aşırı kıvrımlı hale gelmesi (arteriyel genişleme ve venöz kıvrımlanma), damar kalınlığındaki değişimler hastalığın evresini belirlemede doğrudan etkili olmaktadır. Damar morfolojisindeki değişimler, plus hastalığının tespitinde belirleyici bir parametre olarak kullanılır.

Retinal damar ağının geometrik özellikleri, hastalığın ilerleyişi hakkında önemli bilgiler sunmaktadır. Damar çapı, kıvrımlılık derecesi, dallanma açıları ve vasküler yoğunluk gibi nicel parametreler, klinik değerlendirmede objektif kriterler olarak kullanılabilir. Ancak bu parametrelerin manuel olarak ölçülmesi zaman alıcı ve gözlemciler arası değişkenliğe açık bir süreçtir. Bu nedenle, damar yapısının otomatik segmentasyonu ve analizi, ROP tanısında standardizasyon sağlamak ve erken teşhis doğruluğunu artırmak açısından büyük önem taşımaktadır.

1.4.3. Otomatik Segmentasyonun Klinik Katkısı

Yapay zeka tabanlı segmentasyon sistemleri, retinal hastalıklarında tespitinde birçok önemli katkı sağlamaktadır. Bu sistemler retina damar yapısının yüksek hacimli veri setlerinde hızlı, güvenilir ve tekrarlanabilir şekilde analiz edilmesine olanak tanır. Klinik açıdan, Otomatik damar segmentasyonu uzman değerlendirmesine yardımcı olarak tanı süresini önemli ölçüde kısaltır ve uzmanlar tarafından konulan tanı farklılıklarını azaltır (Abramoff et al, 2010).

1.5. Tıbbi Görüntü Segmentasyonunda Derin Öğrenme Yöntemleri

1.5.1. Evrişimsel Sinir Ağları (CNN)

Evrişimli Sinir Ağları, özellikle görsel veriler üzerinde güçlü bir performans sağlayan özelleşmiş bir derin öğrenme mimarisidir. Geneksel yapay sinir ağlarından ayrılan yönü verinin uzamsal yapısını koruyarak işlem yapabilme yeteneğidir. Bu mimari özellik, bilgisayarlı görü problemlerinde özellikle görüntü sınıflandırma, nesne algılama ve görsel

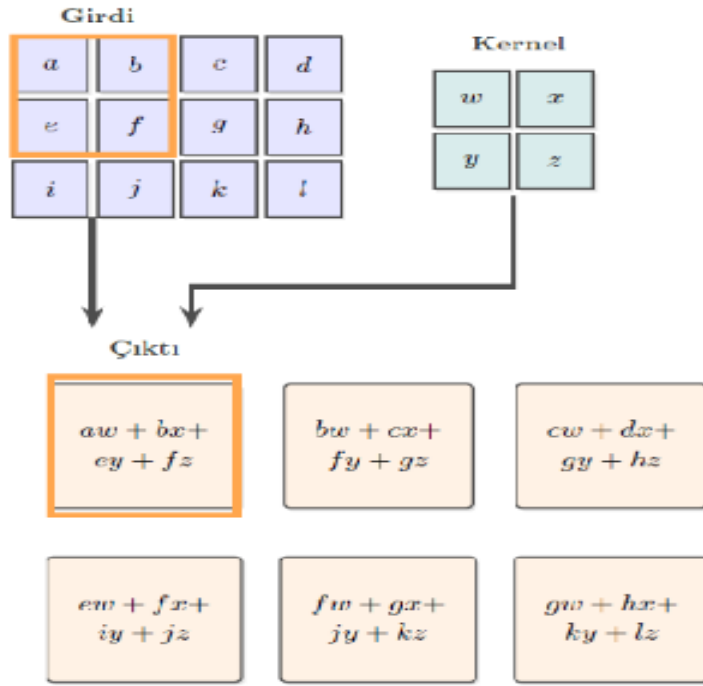
desen çıkarımı gibi görevlerde kayda değer başarılar elde edilmesini sağlamaktadır (LeCun vd., 2015).

Evrişimli sinir ağlarının geleneksel sinir ağlarından en temel farkı ve avantajları şu şekilde sıralanabilir:

- Uzamsal bilginin korunması: Evrişimli sinir ağları, görüntü üzerindeki uzamsal dağılımı ve uzamsal bilgiyi anlayarak her pikseli diğer piksellerle ilişkisi içinde değerlendirir. Bu özellik, evrişimli sinir ağlarının en önemli avantajlarından birini oluşturmaktadır.
- Yerel özelliklerin tespiti: Küçük boyutlu filtreler kullanması sayesinde, evrişimli sinir ağları görüntü içerisindeki ince ayrıntıları etkin bir şekilde tespit edebilir. Yakalanan bu kritik bilgiler, ilerleyen katmanlarda birleştirilerek önemli özelliklerin kaybolması önlenir ve hiyerarşik bir özellik çıkarımı gerçekleştirilir.
- Hesaplama verimliliği: Geleneksel sinir ağlarına kıyasla evrişimli sinir ağlarının en belirgin avantajı hesaplama maliyetinin düşük olmasıdır. Evrişim işlemleri ve havuzlama (pooling) katmanları aracılığıyla görüntünün boyutu küçültülürken, önemli öznitelikler korunur. Bu yaklaşım, parametre sayısını ve hesaplama yükünü önemli ölçüde azaltarak modelin verimliliğini artırır.

Evrişimsel sinir ağlarının mimarisi, belirli işlevlere sahip katmanların sistematik bir şekilde ardışık olarak yerleştirilmesiyle oluşturulur. Bu katmanlar üç temel kategoride incelenebilir:

Evrişim katmanı, ESA mimarisinin temel yapı taşıdır. Bu katmanda, girdi görüntüsünün tamamına erişmek yerine, yerel alanlardaki özellik çıkarımı gerçekleştirilir. Bu amaçla filtre veya çekirdek (kernel) olarak adlandırılan küçük boyutlu matrisler kullanılır. Bir filtre, girdi matrisinin üzerinde belirli bir adım büyüklüğü (stride) ile kaydırılarak evrişim işlemini gerçekleştirir. Bu işlem sonucunda, ilgili bölgenin özelliklerini temsil eden bir özellik haritası (feature map) elde edilir. Her bir filtre, görüntüdeki farklı bir özelliği (örneğin kenar, köşe, doku veya renk gradyanı) tespit etmek üzere eğitilir. Matematiksel olarak, evrişim işlemi girdi matrisi ile filtre matrisi arasındaki çapraz korelasyon operasyonu şeklinde ifade edilebilir (Goodfellow vd., 2016).



Şekil 2. Evrişim katmanında özellik çıkarımı

Şekil 2'de gösterilen evrişim işleminde, 2×2 boyutlu kernel matrisi (w, x, y, z elemanlarından oluşan), 4×4 boyutlu girdi matrisi üzerinde sistematik olarak kaydırılmaktadır. Her kaydırma adımı (stride), kernel elemanları ile örtüşen girdi bölgesindeki elemanlar arasında eleman-bazında çarpım işlemi uygulanır ve bu çarpımların toplamı çıktı matrisinde bir değer olarak kaydedilir. Örneğin, turuncu ile vurgulanan ilk pozisyonda çıktı değeri ($aw + bx + ey + fz$) formülü ile hesaplanır. Bu işlem, kernelin tüm girdi üzerinde taranmasıyla tekrarlanır ve sonuçta 3×2 boyutlu bir özellik haritası elde edilir. Evrişim operasyonu, ESA mimarilerinde yerel uzamsal özelliklerin öğrenilmesinde temel rol oynar.

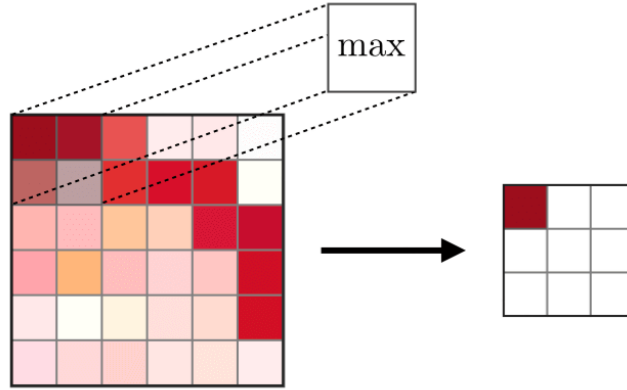
Birden fazla filtrenin kullanılması, görüntüden çok sayıda özelliğin eş zamanlı olarak çıkarılmasına olanak tanır. İlk katmanlarda genellikle düşük seviyeli özellikler (kenarlar, çizgiler) öğrenilirken, daha derin katmanlarda yüksek seviyeli ve soyut özellikler (şekiller, dokular, nesneler) öğrenilir.

Havuzlama katmanı, özellik haritalarının boyutsal indirgenmesini sağlayan ve hesaplama karmaşıklığını azaltan kritik bir bileşendir. Bu katman, özellik haritaları üzerinde aşağı örnekleme (downsampling) işlemi uygulayarak, uzamsal boyutları küçültürken en önemli bilgilerin korunmasını hedefler.

Havuzlama işlemi, genellikle örtüşmeyen veya kısmi örtüşen küçük pencereler kullanılarak gerçekleştirilir. En yaygın kullanılan havuzlama yöntemleri şunlardır:

- Maksimum Havuzlama (Max Pooling): Belirlenen pencere içindeki en büyük değeri seçer. Bu yöntem, en belirgin özelliklerin korunmasını sağlar ve görüntüdeki küçük değişimlere karşı dayanıklılık (invariance) kazandırır.
- Ortalama Havuzlama (Average Pooling): Pencere içindeki tüm değerlerin ortalamasını alır. Bu yaklaşım, daha yumuşak bir özellik geçişi sağlar ve genel doku bilgisini korur.

Havuzlama katmanının başlıca avantajları şunlardır: Özellik haritalarının boyutunu küçülterek hesaplama maliyetini düşürür. Aşırı öğrenmeyi (overfitting) azaltır. Modelin konumsal değişimlere karşı daha dayanıklı olmasını sağlar. Ağın daha derin katmanlarda daha geniş alanlardaki özellikleri algılamasına olanak tanır.



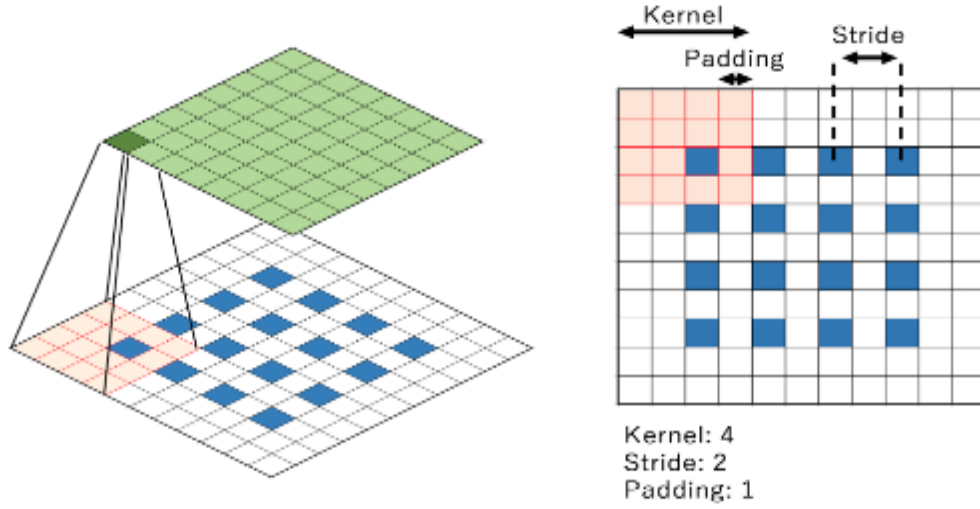
Şekil 3. 2×2 Max Pooling işlemi ile 7×7 boyutundaki özellik haritasının 3×3 boyutuna indirgenmesi

Şekil 3, maksimum havuzlama (max pooling) işleminin çalışma prensibini görselleştirmektedir. 7×7 boyutundaki girdi özellik haritası üzerinde, 2×2 boyutlu bir pencere (filtre) stride=2 (adım=2) değeri ile kaydırılmaktadır. Her kaydırma adımında, pencere içindeki dört piksel değerinden en büyüğü seçilerek çıktı matrisine aktarılır. Koyu kırmızı ile gösterilen ilk bölgede, pencere içindeki maksimum değer belirlenerek çıktı matrisinin sol üst köşesine yerleştirilir. Bu işlem, tüm özellik haritası taranana kadar tekrar eder ve sonuçta 3×3 boyutlu indirgenmiş bir özellik haritası elde edilir. Bu mekanizma, uzamsal boyutların %50'den fazla azaltılmasını sağlarken, her bölgedeki en güçlü özelliklerin korunmasına olanak tanır.

Tam bağlantılı katmanlar, genellikle ESA mimarisinin son bölümünde yer alır ve sınıflandırma görevini gerçekleştirir. Evrişim ve havuzlama katmanları tarafından çıkarılan yüksek seviyeli özellikleri alarak, bunları nihai çıktıya (sınıf etiketlerine) dönüştürür. Bu katmanda, önceki katmandan gelen tüm nöronlar, mevcut katmandaki her bir nörona bağlanır. Dolayısıyla, öğrenilen özelliklerin tamamı dikkate alınarak karar verme süreci gerçekleştirilir. Evrişim katmanlarından gelen çok boyutlu özellik haritaları, önce düzleştirme (flatten) işlemi ile tek boyutlu bir özellik vektörüne dönüştürülür. Geleneksel yapay sinir ağlarındaki gizli katmanlarla aynı yapıya sahip olan tam bağlantılı katmanlar, özellik vektörünü alarak sınıf olasılıklarını hesaplar. Çok sınıflı sınıflandırma problemlerinde, son tam bağlantılı katmanın çıkışında genellikle softmax aktivasyon fonksiyonu kullanılır. Bu fonksiyon, her bir sınıfa ait olasılık değerlerini hesaplayarak, en yüksek olasılığa sahip sınıfın tahmin olarak seçilmesini sağlar.

Padding (Dolgu): Girdi görüntüsünün kenarlarına sıfır veya başka değerler eklenerek, evrişim işlemi sonrası özellik haritasının boyutunun kontrol edilmesini sağlar. Valid padding (padding=0) kullanıldığında her katmanda boyut küçülür ve kenar bilgileri daha az işlenir. Same padding ise kenarlara sıfır eklenerek orijinal boyutun korunmasını sağlar, bu da derin ağlarda boyut yönetimini kolaylaştırır ve kenar piksellerdeki bilgi kaybını önler.

Stride (Adım): Filtrenin girdi üzerinde kaydırılırken attığı adım sayısını belirtir. Stride=1 her piksel üzerinden geçerken, daha büyük stride değerleri ($\text{stride} \geq 2$) çıktı boyutunu önemli ölçüde küçültür ve hesaplama maliyetini azaltır. Büyük stride, bazen havuzlama katmanının alternatifi olarak kullanılır.



Şekil 4. Evrişimsel sinir ağlarında çekirdek, adım ve dolgu parametrelerinin görselleştirilmesi

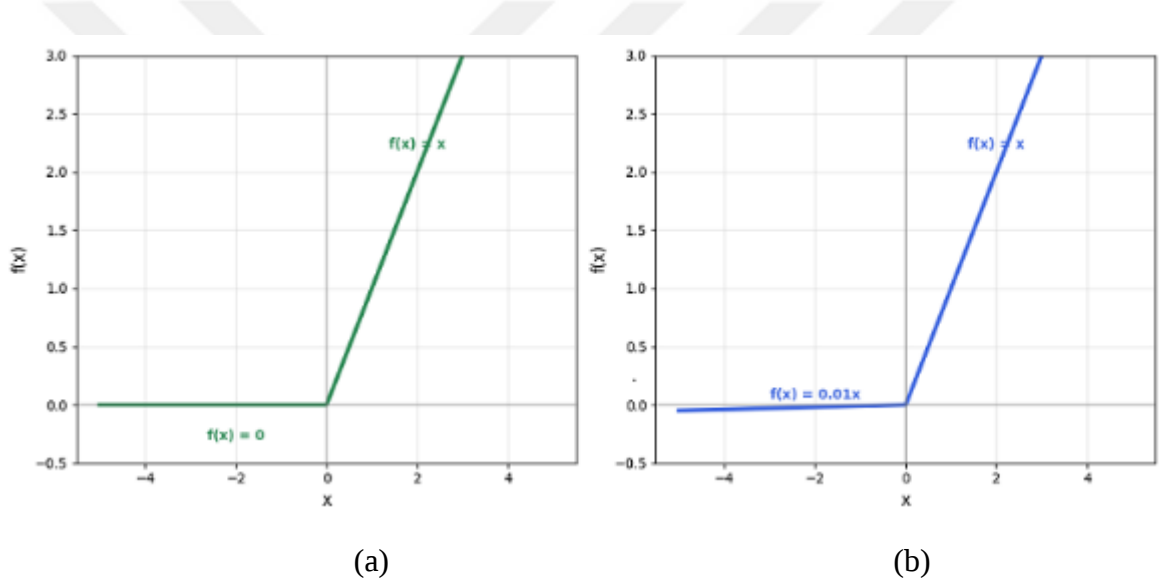
Şekil 4, evrişimsel sinir ağlarının temel yapıtaşlarından olan evrişim (convolution) işleminin parametrelerini görsel olarak sunmaktadır. Sol tarafta, çekirdek (kernel) matrisinin giriş verileri üzerinde sistematik bir şekilde kaydırılarak özellik çıkarımı gerçekleştirdiği üç boyutlu gösterim yer almaktadır. Şeklin sağ bölümünde, 4×4 boyutundaki bir çekirdeğin 2 birimlik adım (stride) parametresi ile hareket ettiği iki boyutlu olarak detaylı gösterilmektedir. Çekirdek tarafından işlenen bölgeler, evrişim işleminin uygulandığı konumları ifade etmektedir. Giriş matrisinin çevresine eklenen genişletilmiş alan ise sıfır dolgu (zero-padding) tekniğini temsil etmekte ve bu sayede çıktı boyutlarının kontrolü sağlanmaktadır. Şekilde belirtildiği üzere, kernel boyutu 4, stride değeri 2 ve padding değeri 1 olarak yapılandırılmıştır.

Dolgu işlemi, özellikle kenar piksellerinin kaybolmasını önlemek ve çıktı boyutlarını kontrol edebilmek amacıyla uygulanmaktadır. Adım parametresinin 2 olarak seçilmesi, örnekleme frekansını azaltarak çıktı öznitelik haritasının uzamsal çözünürlüğünü düşürmektedir. Bu parametrelerin optimal konfigürasyonu, ağın öğrenme kapasitesi, hesaplama karmaşıklığı ve bellek gereksinimleri arasındaki dengenin sağlanmasında kritik rol oynamaktadır. Bu parametrelerin optimize edilmesi, ESA mimarilerinin derinlik, verimlilik ve performans dengesi açısından kritiktir. Modern mimariler (ResNet, VGG, Inception vb.), padding ve stride değerlerini stratejik olarak kullanarak görüntü işleme görevlerinde yüksek başarı elde etmektedir.

Aktivasyon fonksiyonları, ESA mimarisinde doğrusal olmayan (non-linear) özellikler öğrenmeyi mümkün kılan kritik bileşenlerdir. Bu fonksiyonlar olmadan, ağ ne kadar derin olursa olsun yalnızca doğrusal dönüşümler öğrenebilir ve karmaşık örüntüleri modelleyemez. ESA'larda en yaygın kullanılan aktivasyon fonksiyonları:

ReLU (Rectified Linear Unit): En popüler aktivasyon fonksiyonu olup, $f(x) = \max(0, x)$ şeklinde tanımlanır. Negatif değerleri sıfırlayarak basit ama etkili bir doğrusal olmama özelliği sağlar. Hesaplama açısından verimlidir ve gradyan kaybı (vanishing gradient) problemini azaltır.

Leaky ReLU: ReLU'nun geliştirilmiş versiyonu olup, negatif değerler için küçük bir eğim (genellikle 0.01) kullanır. Bu sayede "ölü nöron" problemini azaltır.



Şekil 5. Aktivasyon fonksiyonları: (a) ReLU (Rectified Linear Unit) fonksiyonu, (b) Leaky ReLU fonksiyonu

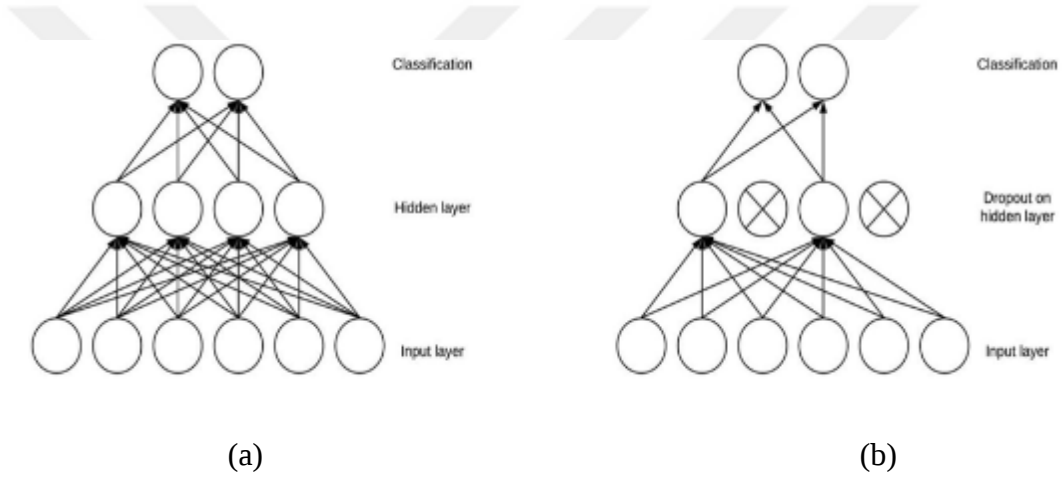
Şekil 5, her iki fonksiyonun da pozitif bölgede aynı doğrusal davranışı sergilediğini, ancak negatif bölgede farklılaştığını göstermektedir. ReLU (a) negatif değerleri tamamen sıfırlarken, Leaky ReLU (b) negatif bölgede küçük bir eğim (0.01) koruyarak gradyan akışını sürdürür.

Sigmoid: Çıkış değerlerini 0 ile 1 arasında sınırlayan bir fonksiyondur. İkili sınıflandırma problemlerinin çıkış katmanında sıklıkla kullanılır.

Tanh: Sigmoid'in ölçeklendirilmiş versiyonu olup, çıkış aralığı -1 ile 1 arasındadır. Sıfır merkezli olması nedeniyle bazı durumlarda sigmoid fonksiyonu yerine tercih edilir.

ESA'larda aşırı öğrenmeyi önlemek ve modelin genelleme yeteneğini artırmak için çeşitli düzenleme (regularization) teknikleri kullanılır:

Dropout: Eğitim sırasında rastgele seçilen nöronların geçici olarak devre dışı bırakılması işlemidir. Bu yöntem, ağırların belirli nöronlara aşırı bağımlı olmasını önler ve daha ayırt edici özellikler öğrenmesini teşvik eder (Srivastava vd., 2014). Şekil 6(a) standart tam bağlantılı sinir ağı mimarisini göstermektedir; tüm nöronlar ve bağlantılar aktif durumdadır. Şekil 6(b)'de ise eğitim sırasında dropout uygulandığında, rastgele seçilen nöronlar (X ile işaretli) ve bunlara ait bağlantılar geçici olarak devre dışı bırakılmaktadır. Bu işlem, her eğitim iterasyonunda farklı bir alt-ağ yapısının öğrenmesini sağlayarak modelin genelleme yeteneğini artırır.



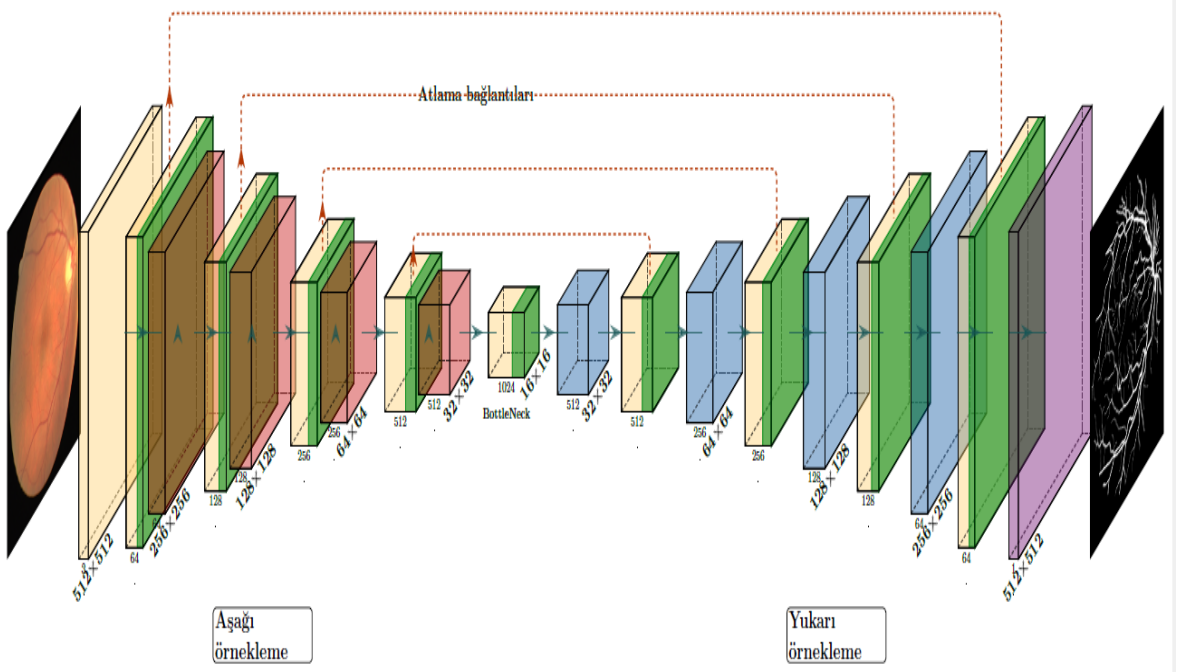
Şekil 6. Dropout regularizasyon tekniğinin görselleştirilmesi

Batch Normalization: Her katmandaki aktivasyonların normalize edilmesini sağlayarak, eğitim sürecini hızlandırır ve daha yüksek öğrenme oranlarının kullanılmasına olanak tanır. Ayrıca, düzenleyici bir etki göstererek aşırı öğrenmeyi azaltır (Ioffe & Szegedy, 2015). **L1 ve L2 Düzenleme:** Kayıp fonksiyonuna ağırlık değerlerinin cezalandırılmasını ekleyerek modelin karmaşıklığını kontrol eder. L1 düzenleme seyrek ağırlık matrisleri oluştururken, L2 düzenleme (weight decay) ağırlık değerlerini küçük tutar ve yaygın olarak tercih edilir. **Data Augmentation (Veri Çoğaltma):** Eğitim verisi üzerinde döndürme, yansıtma, kırpmaya, renk değiştirme gibi dönüşümler uygulanarak yapay olarak veri setini genişletir. Bu yöntem, modelin farklı varyasyonlara karşı dayanıklılığını artırır ve en etkili düzenleme tekniklerinden biridir. **Early Stopping:** Doğrulama hatasının artmaya başladığı noktada eğitimi durdurarak aşırı öğrenmeyi önler. Model, doğrulama performansının en iyi olduğu iterasyondaki ağırlıklarla kaydedilir.

1.5.2. U-Net: Biyomedikal Görüntü Segmentasyonu İçin Temel Mimari

U-Net, özellikle tıbbi görüntü bölütleme (segmentation) görevlerinde yüksek başarı gösteren ve 2015 yılında Ronneberger, Fischer ve Brox tarafından geliştirilen özel bir evrimsel sinir ağı mimarisidir. Adını karakteristik U-şeklindeki mimari yapısından alan bu ağ, sınırlı eğitim verisi olan problemlerde de etkili sonuçlar üretebilme yeteneği sayesinde biyomedikal görüntü analizi alanında devrim niteliğinde bir gelişme olarak kabul edilmektedir (Ronneberger vd., 2015).

U-Net mimarisi, piksel düzeyinde hassas tahminler gerektiren görüntü bölütleme problemleri için özel olarak tasarlanmıştır. Geleneksel sınıflandırma görevlerinden farklı olarak, bölütleme görevlerinde her piksel için ayrı bir sınıf tahmini yapılması gerekmektedir. U-Net, bu zorlu görevi yerine getirmek için kodlayıcı (encoder) ve kod çözücü (decoder) iki temel bileşeni içeren simetrik bir yapı kullanır.



Şekil 7. U-Net mimarisi

Şekil 7, U-Net'in simetrik encoder-decoder mimarisini göstermektedir. Soldaki daralan yol (encoder), ardışık 3x3 evrişim ve 2x2 max pooling katmanları ile görüntü boyutunu kademeli olarak azaltırken özellik kanal sayısını artırır (64→128→256→512→1024). Ortadaki bottleneck bölümü, en derin özellik temsilini

oluşturur. Sağdaki genişleyen yol (decoder), 2×2 transposed convolution ile yukarı örnekleme yapar ve özellik kanal sayısını azaltır. Kesikli oklar ile gösterilen atlama bağlantıları (skip connections), encoder'daki yüksek çözünürlüklü özellikleri decoder'ın karşılık gelen seviyelerine birleştirir (concatenation). Bu mekanizma, ince yapıların ve kenar detaylarının korunmasını sağlar. Son aşamada 1×1 evrişim katmanı ile piksel-düzeyinde sınıf tahmini üretilir.

Daralma yolu, mimarinin sol tarafını oluşturur ve klasik bir evrişimsel sinir ağı yapısını takip eder. Bu bölümün temel işlevleri şunlardır:

- **Özellik Çıkarımı:** Ardışık evrişim katmanları aracılığıyla görüntüden hiyerarşik özellikler çıkarılır. İlk katmanlarda düşük seviyeli özellikler (kenarlar, dokular), derin katmanlarda ise yüksek seviyeli semantik özellikler öğrenilir.
- **Boyutsal İndirgeme:** Her evrişim bloğundan sonra 2×2 maksimum havuzlama işlemi uygulanır. Bu işlem, uzamsal boyutları yarıya indirirken özellik kanallarının sayısını iki katına çıkarır (örneğin, $64 \rightarrow 128 \rightarrow 256 \rightarrow 512$).
- **Bağlam Bilgisinin Yakalanması:** Havuzlama işlemleri sayesinde, ağı alıcı alanı (receptive field) genişler ve daha geniş bağlamsal bilgi elde edilir.
- **Tipik bir daralma yolunda,** her seviyede iki adet 3×3 evrişim katmanı (padding olmadan), ardından ReLU aktivasyon fonksiyonu ve 2×2 maksimum havuzlama katmanı yer alır.

Bu yapı, giderek artan soyutlama seviyelerinde özellik temsillerinin öğrenilmesini sağlar. Genişleme yolu, mimarinin sağ tarafını oluşturur ve daralma yolunda kaybedilen uzamsal bilgiyi geri kazandırır. Bu bölümün temel özellikleri:

- **Yukarı Örnekleme (Upsampling):** Her adımda, özellik haritaları 2×2 ters evrişim (transposed convolution veya deconvolution) işlemi kullanılarak boyutsal olarak iki katına çıkarılır. Aynı zamanda özellik kanalları yarıya indirilir.
- **Atlama Bağlantıları (Skip Connections):** U-Net'in en yenilikçi ve ayırt edici özelliği, daralma yolundaki karşılık gelen katmanlardan gelen yüksek çözünürlüklü özellik haritalarının, genişleme yolundaki yukarı örneklenmiş özellik haritalarıyla birleştirilmesidir. Bu birleştirme işlemi, özellik haritalarının kanal boyutunda birleştirilmesi (concatenation) şeklinde gerçekleştirilir.
- **Hassas Lokalizasyon:** Atlama bağlantıları sayesinde, derin katmanlarda öğrenilen semantik bilgi ile sık katmanlardaki hassas konum bilgisi bir araya getirilir. Bu mekanizma, piksel düzeyinde kesin bölütleme sonuçlarının elde edilmesini sağlar.

Genişleme yolunda her seviyede, birleştirme işleminden sonra iki adet 3×3 evrişim katmanı ve ReLU aktivasyonu uygulanır. Son katmanda, 1×1 evrişim kullanılarak özellik vektörü istenilen sınıf sayısına eşlenir.

U-şeklindeki mimarinin alt noktasında yer alan köprü katmanı, en derin ve en soyut özelliklerin öğrenildiği bölümdür. Bu katmanda, en küçük uzamsal boyut ve en fazla özellik kanalı bulunur. Köprü katmanı, kodlayıcı ve kod çözücü arasında bir geçiş görevi görür ve görüntünün en yoğun semantik temsilini içerir.

U-Net'in başarısının temel nedeni, atlama bağlantılarının kullanılmasıdır. Bu bağlantılar birkaç kritik avantaj sağlar:

- Gradyan Akışının İyileştirilmesi: Geri yayılım sırasında, gradyanların ağı derin katmanlarına daha kolay ulaşmasını sağlayarak eğitim sürecini iyileştirir ve gradyan kaybı problemini azaltır.
- Çok Ölçekli Bilgi Entegrasyonu: Farklı soyutlama seviyelerindeki bilgilerin kombinasyonu hem yerel detayların hem de global bağlamın yakalanmasını sağlar.
- Hassas Sınır Belirleme: Özellikle nesne sınırlarının kesin olarak belirlenmesi gereken uygulamalarda, yüksek çözünürlüklü özellik haritalarının korunması kritik öneme sahiptir.

U-Net, sınırlı eğitim verisi ile çalışabilmek için veri artırma (data augmentation) tekniklerinden yararlanır. Özellikle tıbbi görüntüleme alanında, etiketlenmiş veri elde etmenin maliyetli ve zaman alıcı olması nedeniyle bu özellik büyük önem taşır:

- Elastik Deformasyon: Görüntülere rastgele elastik bozulmalar uygulanarak, doğal doku varyasyonları simüle edilir.
- Döndürme ve Yansıtma: Görüntüler rastgele açılarda döndürülür ve yansıtılır.
- Ölçeklendirme ve Kırpma: Farklı büyütme seviyelerini simüle etmek için kullanılır.
- Parlaklık ve Kontrast Değişimi: Farklı görüntüleme koşullarına karşı dayanıklılık sağlar.

U-Net, özellikle biyomedikal görüntü segmentasyonunda karşılaşılan iki kritik zorluğu ele almak için özel bir ağırlıklı kayıp fonksiyonu kullanır: sınıf dengesizliği ve bitişik nesnelerin birbirinden ayrılması. Standart segmentasyon görevlerinde, arka plan pikselleri genellikle nesne piksellerinden çok daha fazladır ve bu durum modelin arka plana aşırı odaklanmasına neden olabilir. Ayrıca, birbirine çok yakın veya değen hücrelerin sınırlarının doğru tespit edilmesi, biyomedikal analizlerde hayati öneme sahiptir. U-Net'in kayıp

fonksiyonu, her piksel için özel olarak hesaplanan bir ağırlık haritası (weight map) kullanır. Bu ağırlık haritası iki temel bileşenden oluşur:

- Sınıf dengesi ağırlığı (w_c): Her sınıfın frekansına göre hesaplanan ve az temsil edilen sınıflara (örneğin hücre sınırları) daha yüksek ağırlık veren bir katsayıdır. Bu sayede, model nadir görülen ama kritik öneme sahip piksellere daha fazla dikkat eder.
- Nesne ayrımı ağırlığı (w_0): Bitişik nesneler arasındaki sınır bölgelerine ekstra ağırlık verir. Her piksel için, en yakın iki nesne merkezine olan uzaklıklar (d_1 ve d_2) hesaplanır. Bu uzaklıklar kullanılarak, nesnelerin birleştiği bölgelerde ($d_1 + d_2$ küçük olduğunda) ağırlık değeri artırılır. Bu mekanizma, modelin bitişik hücreleri ayırt etmesini teşvik eder. Özel kayıp fonksiyonu sayesinde U-Net hem sınıf dengesizliği problemini çözer hem de birbirine çok yakın nesnelerin sınırlarını başarıyla tespit edebilir. Özellikle hücre segmentasyonunda, değen hücrelerin ayrı bireyler olarak tanınması kritik öneme sahiptir ve bu ağırlıklı kayıp fonksiyonu bu zorluğu başarıyla ele alır.

U-Net mimarisi, başlangıçta tıbbi görüntü analizi için geliştirilmiş olsa da zamanla birçok farklı alanda başarıyla kullanılmıştır:

- Tıbbi görüntü bölütleme: Hücre bölütleme, organ segmentasyonu, tümör tespiti, beyin MR görüntülerinin analizi, retina damar segmentasyonu ve Prematüre Retinopatisi (ROP) teşhisi gibi oftalmolojik uygulamalarda kullanılmaktadır.
- Uydu görüntü analizi: Uzaktan algılama uygulamalarında arazi kullanım sınıflandırması, bina ve yapı tespiti, yol ağı çıkarımı, orman alanlarının haritalanması ve şehir planlaması çalışmalarında kullanılmaktadır.
- Endüstriyel uygulamalar: Üretim süreçlerinde otomatik kalite kontrolü, yüzey kusurlarının tespiti, malzeme yapısı analizi ve yarı iletken üretiminde hata tespiti gibi uygulamalarda kullanılmaktadır.
- Otonom araçlar: Sürücüsüz araçlarda yol ve şerit segmentasyonu, trafik işaretlerinin ve diğer araçların tespiti, sürüş senaryolarının anlaşılması ve robotik uygulamalarda çevre algılama görevlerinde yer almaktadır.

U-Net'in başarısı, birçok varyasyon ve iyileştirmenin geliştirilmesine ilham vermiştir:

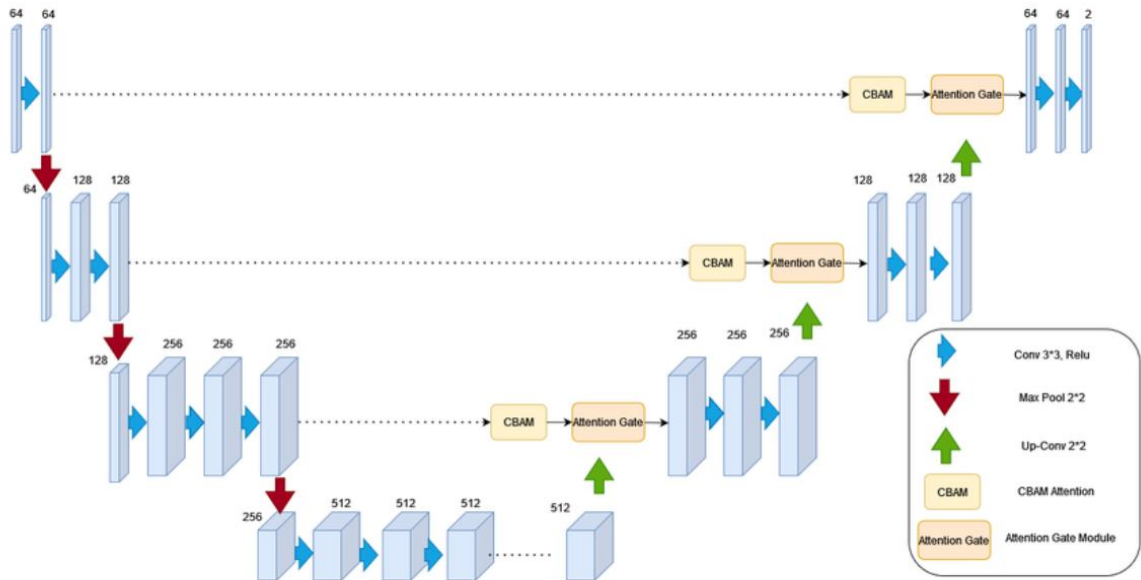
- 3D U-Net: Orijinal U-Net'in üç boyutlu görüntüler için uyarlanmış versiyonudur (Çiçek vd., 2016). 2D evrişim katmanları yerine 3D evrişim kullanarak, hacimsel tıbbi görüntülerde (CT, MRI) daha iyi bağlamsal bilgi çıkarımı sağlar.
- Residual U-Net: U-Net mimarisine rezidüel bağlantılar (residual connections) eklenerek daha derin ağların eğitilmesini kolaylaştırır (Alom vd., 2018). ResNet'ten ilham alan bu yaklaşım, gradyan akışını iyileştirerek daha karmaşık özellik öğrenimini mümkün kılar ve performansı artırır.
- Attention U-Net: Encoder'dan decoder'a aktarılan özelliklere otomatik dikkat ağırlıkları uygulayarak, modelin hedef bölgelere seçici odaklanmasını sağlar. Attention gate'ler, alakasız arka plan özelliklerini baskılamak, önemli anatomik yapılara ait özellikleri vurgular. Özellikle küçük, ince veya düşük kontrastlı hedef yapıların segmentasyonunda önemli performans artışı gösterir. Bu tez çalışmasında geliştirilen VesSegNet modeli, Attention U-Net mimarisine dayanmaktadır.
- U-Net++: Encoder ve decoder arasında yoğun (nested) atlama bağlantıları ve derin gözetim stratejisi kullanır. Her encoder seviyesi, sadece karşılık gelen decoder seviyesine değil, ara seviyelere de bağlanır. Bu yoğun bağlantı yapısı, farklı semantik seviyelerdeki özelliklerin daha iyi füzyonunu sağlar ve değişken boyutlardaki nesnelerin segmentasyonunda üstün performans gösterir.

U-Net mimarisi, sınırlı veri ile yüksek hassasiyetli piksel seviyesi tahminler gerektiren problemlerde, özellikle de tıbbi görüntü analizi alanında, altın standart haline gelmiştir. Basit ama etkili tasarımı, güçlü performansı ve uyarlanabilirliği sayesinde, görüntü bölütleme alanında en çok tercih edilen mimarilerden biri olmaya devam etmektedir.

1.5.3. Attention U-Net ve Dikkat Mekanizmaları

Attention-UNet, özellikle tıbbi görüntü segmentasyonu gibi görevlerde sıklıkla kullanılan popüler bir evrişimsel sinir ağı (CNN) mimarisi olan U-Net'in dikkat mekanizmaları ile iyileştirilmiş bir versiyonudur. Geleneksel U-Net'teki atlama bağlantıları, kodlayıcıdan gelen tüm özellik haritalarını (feature maps) kod çözücüye doğrudan geçirir. Ancak bu hem ilgili hem de ilgisiz arka plan gürültüsü/özelliklerinin geçmesine neden olabilir. Attention U-Net, bu atlama bağlantılarına dikkat geçitleri (attention gates- AG'ler) ekler.

- Dikkat geçidi (attention gate): Kodlayıcıdan gelen özellikler ile kod çözücüdeki daha derin katmanlardan gelen özellikler (bağlam bilgisi) kullanılarak bir dikkat haritası oluşturulur.
- Bu dikkat haritası, kodlayıcı özelliklerini filtreler. Modelin segmentasyon hedefi için en alakalı olan uzamsal alanlara (örneğin bir tümörün sınırlarına veya bir organın kenarlarına) odaklanmasını sağlar.
- Sonuç olarak, Attention U-Net, ilgisiz arka plan bölgelerindeki aktivasyonu otomatik olarak azaltır ve hedef yapının şekline ve boyutuna daha iyi uyum sağlar.



Şekil 8. Attention U-Net mimarisi

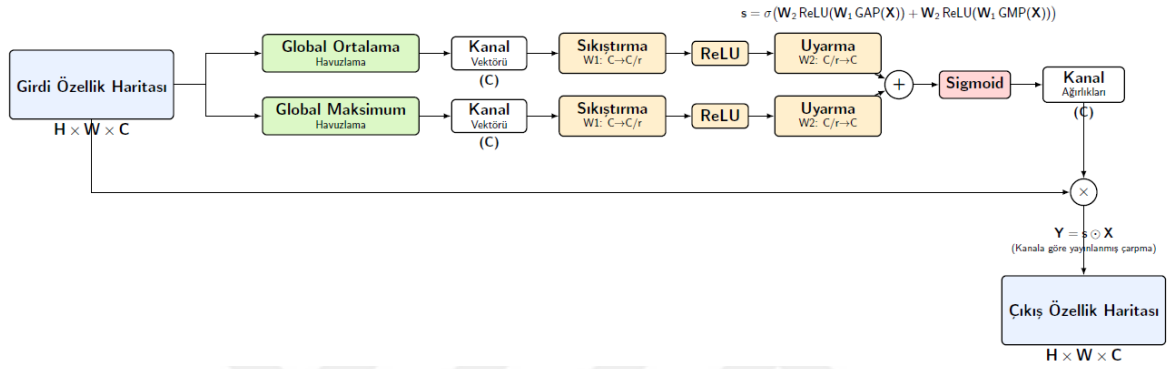
Dikkat mekanizması, derin öğrenme modellerinin bir girdi içerisindeki hangi kısımların (örneğin bir görüntüdeki piksellerin, bir cümledeki kelimelerin) belirli bir görev için daha önemli olduğunu otomatik olarak öğrenmesini sağlayan bir tekniktir. Tıpkı insan beyninin bir işe odaklanırken ilgili bilgileri öne çıkarması gibi, yapay zeka modellerinin de kritik özelliklere daha fazla "dikkat" etmesine olanak tanır.

Bu mekanizma sayesinde model, tüm girdi verisine eşit ağırlık vermek yerine, çıktı üretimi için en alakalı olan kısımlara ağırlıklar atar.

Bağlamı güçlendiren temel dikkat mekanizması türleri:

Kendi Üzerine Dikkat (Self-Attention): Girdinin kendi içindeki farklı kısımlarının birbirleriyle olan ilişkisini hesaplar. Özellikle Transformer gibi doğal dil işleme (NLP) modellerinde yaygın olarak kullanılır. Bir cümledeki her kelimenin diğer kelimelerle olan bağlamını anlamasına yardımcı olur.

Kanallar Arası Dikkat (Channel Attention): Farklı özellik kanallarının (feature channels) göreve ne kadar katkıda bulunduğunu belirleyerek daha önemli kanalları vurgular.



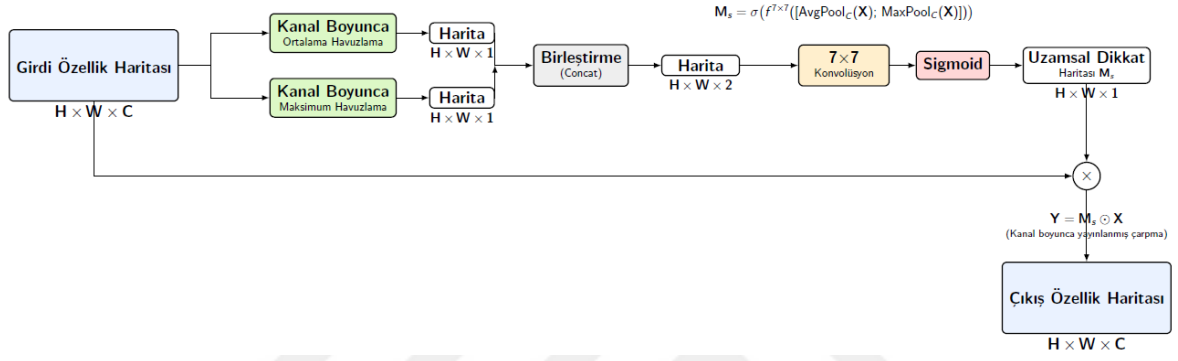
Şekil 9. VesSegNet mimarisinde kullanılan kanal dikkat mekanizmasının (channel attention) yapısı

Kanal dikkat mekanizması (Şekil 9), derin öğrenme modellerinin özellik haritalarındaki farklı kanalların önem derecelerini adaptif olarak ayarlamasını sağlayan bir yöntemdir. Bu mekanizma, ağın hangi özellik kanallarının belirli bir görev için daha bilgilendirici olduğunu öğrenmesine olanak tanır. İlk aşamada, girdi özellik haritasına ($H \times W \times C$ boyutunda) global ortalama havuzlama (Global Average Pooling- GAP) işlemi uygulanarak uzamsal bilgiler sıkıştırılır ve her kanal için tek bir temsili değer elde edilir. Bu işlem, her kanalın genel bağlamsal bilgisini yakalayarak ($1 \times 1 \times C$) boyutunda bir kanal tanımlayıcı vektör üretir. Ardından, iki tam bağlantılı katmandan oluşan bir darboğaz (bottleneck) yapısı kullanılır. İlk katman, hesaplama verimliliğini artırmak amacıyla kanal boyutunu bir azaltma oranı (r) ile küçültür ($C \rightarrow C/r$) ve ReLU aktivasyon fonksiyonu ile doğrusal olmayan bir dönüşüm uygular. İkinci katman ise boyutu tekrar orijinal kanal sayısına genişletir ($C/r \rightarrow C$) ve sigmoid aktivasyon fonksiyonu kullanarak 0 ile 1 arasında normalize edilmiş dikkat ağırlıkları üretir.

Elde edilen bu ağırlıklar, her kanalın önem derecesini temsil eder ve orijinal özellik haritası ile kanal bazında çarpılarak önemli kanalların vurgulanması, önemsiz olanların ise baskılanması sağlanır. Bu yaklaşım, modelin temsil gücünü artırarak daha ayırt edici

özellikler öğrenmesine ve sonuç olarak segmentasyon performansının iyileşmesine katkı sağlar. Kanal dikkat mekanizması, özellikle U-Net gibi encoder-decoder mimarilerinde atlama bağlantılarında (skip connections) kullanıldığında, farklı ölçeklerdeki özelliklerin daha etkili bir şekilde birleştirilmesini mümkün kılar.

Uzamsal Dikkat (Spatial Attention): Görüntü içindeki farklı konumların (piksellerin) göreve ne kadar katkıda bulunduğunu belirleyerek daha kritik uzamsal alanları vurgular.



Şekil 10. VesSegNet mimarisinde kullanılan uzamsal dikkat mekanizmasının (spatial attention) yapısı

Uzamsal dikkat mekanizması (Şekil 10), özellik haritalarının farklı uzamsal konumlarının önem derecelerini belirleyerek modelin ilgili bölgelere odaklanmasını sağlayan bir yöntemdir. Kanal dikkat mekanizmasından farklı olarak, bu yaklaşım "nerede" önemli bilgilerin bulunduğu odaklanır. İlk aşamada, girdi özellik haritası ($H \times W \times C$) üzerinde kanal boyunca iki farklı havuzlama işlemi uygulanır: ortalama havuzlama (average pooling) ve maksimum havuzlama (maximum pooling).

Bu işlemler, her bir uzamsal konumdaki tüm kanallar boyunca istatistiksel bilgileri toplar ve iki ayrı $H \times W \times 1$ boyutunda özellik haritası üretir. Ortalama havuzlama, her konumdaki ortalama aktivasyonu yakalarken, maksimum havuzlama en belirgin özellikleri vurgular. Bu iki özellik haritası birleştirme (concatenation) işlemi ile $H \times W \times 2$ boyutuna getirilir ve böylece her uzamsal konum için zenginleştirilmiş bir temsil elde edilir. Ardından, 7×7 boyutunda bir konvolüsyon katmanı uygulanarak uzamsal bağlam bilgisi yakalanır ve $H \times W \times 1$ boyutunda tek kanallı bir özellik haritası üretilir. Bu konvolüsyon işlemi, komşu pikseller arasındaki ilişkileri öğrenerek her uzamsal konumun önem derecesini belirler. Son olarak, sigmoid aktivasyon fonksiyonu kullanılarak 0 ile 1 arasında normalize edilmiş

uzamsal dikkat ağırlıkları elde edilir. Bu ağırlık haritası, orijinal girdi özellik haritası ile eleman bazında çarpılarak (element-wise multiplication) önemli uzamsal bölgelerin vurgulanması ve önemsiz bölgelerin baskılanması sağlanır. Uzamsal dikkat mekanizması, özellikle tıbbi görüntü segmentasyonunda hastalıklı dokuların veya ilgi alanlarının bulunduğu bölgelere odaklanmayı kolaylaştırarak segmentasyon doğruluğunu artırır. Kanal dikkat mekanizması ile kullanıldığında, hem "ne" hem de "nerede" sorularına cevap vererek modelin temsil kapasitesini önemli ölçüde güçlendirir.

1.6. Model Performans Ölçütleri

1.6.1. Doğruluk, Duyarlılık, Kesinlik

Doğruluk (Accuracy), segmentasyon modelinin doğru sınıflandırdığı piksel sayısının toplam piksel sayısına oranını ifade eder ve genel performansın en temel göstergesidir. Ancak, özellikle sınıf dengesizliğinin bulunduğu segmentasyon problemlerinde doğruluk metriği yanıltıcı olabilir (Garcia-Garcia vd., 2017). Matematiksel olarak doğruluk şu şekilde hesaplanır:

$$\text{Doğruluk} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Burada TP (True Positive), TN (True Negative), FP (False Positive) ve FN (False Negative) değerleri sırasıyla doğru pozitif, doğru negatif, yanlış pozitif ve yanlış negatif tahminleri temsil etmektedir.

Sınıflandırma problemlerinde modelin gerçek pozitif örnekleri tespit etme yeteneğini ölçen temel bir performans metriğidir (Sokolova & Lapalme, 2009). Görüntü segmentasyonu bağlamında, hedef sınıfa ait piksellerin ne oranda doğru şekilde sınıflandırıldığını gösterir ve piksel düzeyinde hesaplanır.

$$\text{Duyarlılık} = \frac{TP}{TP + FN} \quad (2)$$

Kesinlik (Precision), modelin pozitif olarak sınıflandırdığı piksellerin gerçekten pozitif sınıfa ait olma oranını ifade eder ve yanlış pozitiflerin kontrolünde önemlidir (Minaee vd., 2020):

$$\text{Kesinlik} = \frac{TP}{TP + FP} \quad (3)$$

1.6.2. F1 Skoru (F1 Score)

F1 Skoru, kesinlik ve duyarlılık metriklerinin harmonik ortalamasıdır ve her iki metriği dengeli bir şekilde dikkate alarak model performansının bütüncül bir değerlendirmesini sunar. Özellikle dengesiz veri setlerinde model performansını değerlendirmek için güçlü bir ölçüttür:

$$F1 = 2 \times \frac{\text{Kesinlik} \times \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (4)$$

F1 Skoru, 0 ile 1 arasında değer alır ve 1'e yaklaştıkça model performansının daha iyi olduğunu gösterir. Segmentasyon uygulamalarında, özellikle medikal görüntü analizinde yaygın olarak kullanılmaktadır.

1.6.3. Intersection over Union (IoU)

Intersection over Union (IoU), Jaccard indeksi olarak da bilinir. Özellikle nesne tespiti ve segmentasyon görevlerinde en yaygın kullanılan örtüşme tabanlı değerlendirme ölçütlerinden biridir (Everingham vd., 2010). IoU, tahmin edilen segmentasyon maskesi ile gerçek segmentasyon maskesi arasındaki kesişim alanının, birleşim alanına oranını hesaplar:

$$\text{IoU} = \frac{\text{Kesişim Alanı}}{\text{Birleşim Alanı}} = \frac{TP}{TP + FP + FN} \quad (5)$$

IoU metriği, 0 ile 1 arasında değer alır; 0 hiçbir örtüşme olmadığını, 1 ise mükemmel örtüşmeyi ifade eder. Literatürde başarılı bir segmentasyon için sıklıkla 0.5 eşik değeri kabul

edilir (Lin vd., 2014). Pascal VOC ve COCO gibi popüler segmentasyon veri setlerinin değerlendirme protokollerinde temel metrik olarak kullanılmaktadır (Everingham vd., 2010). Mean IoU (mIoU), tüm sınıflar için hesaplanan IoU değerlerinin ortalamasıdır ve çok sınıflı segmentasyon problemlerinde genel performansı özetlemek için kullanılır:

$$mIoU = \frac{1}{N} \sum_{i=1}^N IoU_i \quad (6)$$

1.6.4. ROC Eğrisi ve AUC

Receiver Operating Characteristic (ROC) eğrisi, farklı eşik değerlerinde sınıflandırıcının performansını görselleştiren bir grafikdir ve True Positive Rate (TPR) ile False Positive Rate (FPR) arasındaki ilişkiyi gösterir (Fawcett, 2006).

$$TPR \text{ (Duyarlılık)} = \frac{TP}{TP + FN} \quad (7)$$

$$FPR = \frac{FP}{FP + TN} \quad (8)$$

Area Under the Curve (AUC), ROC eğrisinin altında kalan alanı ifade eder ve 0 ile 1 arasında değer alır. AUC = 0.5 rastgele tahmin yapan bir modeli, AUC = 1.0 ise mükemmel ayırım yapabilen bir modeli temsil eder. Segmentasyon görevlerinde ROC eğrisi ve AUC, özellikle ikili segmentasyon problemlerinde ve olasılık haritalarının kalitesini değerlendirmek için kullanılır. Ancak, çok sınıflı segmentasyon problemlerinde, her sınıf için ayrı ayrı ROC eğrileri oluşturularak tek sınıfa karşı tüm sınıflar (one-vs-all) yaklaşımı benimsenebilir.

1.6.5. Macro ve Micro Ortalama Metrikler

Çok sınıflı segmentasyon problemlerinde, sınıflar arası performans farklılıklarını analiz etmek için macro ve micro ortalama metrikler kullanılır. Macro Ortalama, her sınıf için ayrı ayrı hesaplanan metriklerin aritmetik ortalamasıdır. Bu yaklaşım, her sınıfa eşit

ağırlık vererek sınıf dengesizliğinden bağımsız bir değerlendirme sağlar (Sokolova & Lapalme, 2009):

$$\text{Macro-F1} = \frac{1}{N} \sum_{i=1}^N F1_i \quad (9)$$

Micro Ortalama ise, tüm sınıfların TP, FP ve FN değerlerini toplamadan sonra metrikleri hesaplar, dolayısıyla daha fazla örneğe sahip sınıflara daha fazla ağırlık verir (M.-L. Zhang & Zhou, 2014):

$$\text{Micro-F1} = \frac{2 \times \sum_{i=1}^N TP_i}{2 \times \sum_{i=1}^N TP_i + \sum_{i=1}^N FP_i + \sum_{i=1}^N FN_i} \quad (10)$$

1.6.6. Karışıklık Matrisi (Confusion Matrix)

Karışıklık matrisi (confusion matrix), sınıflandırma modelinin performansını görselleştirmek ve detaylı analiz yapmak için kullanılan temel bir araçtır. N sınıflı bir segmentasyon problemünde, NxN boyutunda bir matris oluşturulur ve her hücre, belirli bir sınıfın başka bir sınıf olarak tahmin edilme sayısını gösterir.

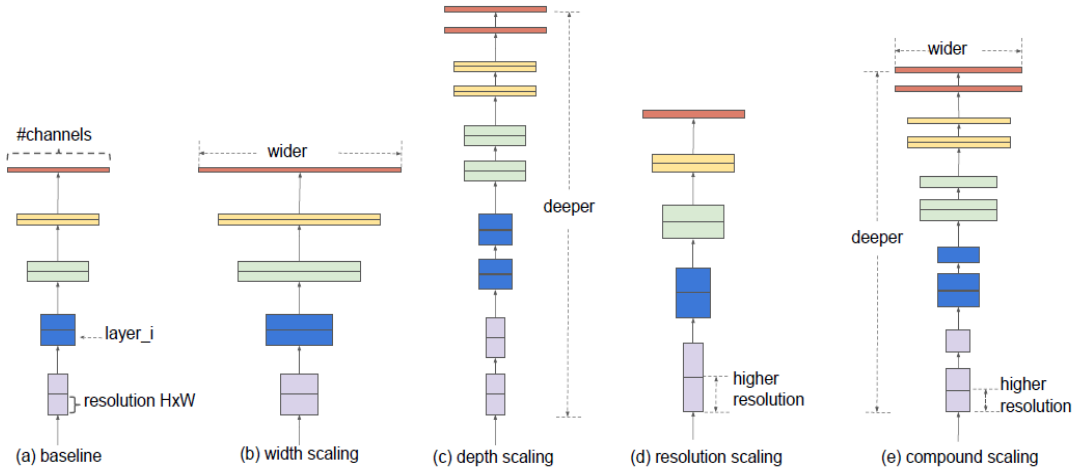
Segmentasyon görevlerinde karışıklık matrisi, piksel düzeyinde hesaplanır ve her piksel için tahmin edilen sınıf ile gerçek sınıf karşılaştırılır. Matrisin köşegen elemanları doğru sınıflandırmaları, köşegen dışı elemanlar ise sınıflandırma hatalarını temsil eder. Karışıklık matrisi, hangi sınıfların birbiriyle karıştırıldığını ve modelin hangi sınıflarda daha zayıf performans gösterdiğini anlamak için kritik bilgiler sunar (Powers, 2020).

1.7. Literatür Taraması

Modern ESA Mimarileri: Derin öğrenme alanında görüntü sınıflandırma görevlerinde kullanılan evrişimsel sinir ağları, son yıllarda mimari inovasyonlar sayesinde dikkat çekici performans artışları göstermiştir. ResNet (Residual Network) mimarisi, He ve arkadaşları tarafından 2015 yılında önerilen artık bağlantılar (residual connections) sayesinde performans kaybı problemini çözerek çok derin ağların eğitilmesini mümkün kılmıştır (He vd., 2015). ResNet'in temel fikri, katmanların doğrudan eşlemeleri öğrenmesi yerine artık fonksiyonları öğrenmesine dayanmaktadır; bu yaklaşım, gradyan akışını iyileştirerek 152 katmana kadar olan ağların başarıyla eğitilmesini sağlamıştır. ImageNet veri setinde %3.57'lik hata oranıyla o dönem için state-of-the-art performans elde eden ResNet, tıbbi görüntüleme uygulamalarında da yaygın olarak tercih edilmektedir.

MobileNet mimarisi, Google tarafından mobil ve gömülü sistemler için özel olarak tasarlanmış hafif bir ESA modelidir ve derinlemesine ayrılabilir evrişimler (depthwise separable convolutions) kullanarak hesaplama maliyetini önemli ölçüde azaltmaktadır (Howard vd., 2017). Standart evrişim işlemlerini derinlemesine evrişim ve noktasal evrişim olmak üzere iki aşamaya ayıran bu yaklaşım, parametre sayısını ve hesaplama yükünü yaklaşık 8-9 kat azaltarak mobil cihazlarda gerçek zamanlı çalışmayı mümkün kılmaktadır. MobileNet'in V2 ve V3 versiyonları, ters darboğaz (inverted residual) blokları ve sıkıştırma-genişletme (squeeze-and-excitation) mekanizmalarıyla performansı daha da geliştirmiştir.

EfficientNet, Tan ve Le tarafından 2019 yılında önerilen ve model ölçeklendirmesi için yeni bir paradigma sunan bir mimari ailesidir (Tan & Le, 2019). Geleneksel yaklaşımların aksine, EfficientNet derinlik, genişlik ve çözünürlük boyutlarını bileşik bir katsayı kullanarak dengeli bir şekilde ölçeklendiren bileşik ölçeklendirme yöntemini kullanmaktadır. Neural Architecture Search (NAS) ile optimize edilen EfficientNet-B0 temel modeli, ImageNet üzerinde %77.1 top-1 doğruluk oranıyla önceki modellere göre 8.4 kat daha az parametre ve 6.1 kat daha az hesaplama maliyeti sunmaktadır. EfficientNet-B0'dan B7'ye kadar farklı ölçeklerdeki modeller, hesaplama kaynakları ve doğruluk arasında optimal denge sağlayarak tıbbi görüntü analizi gibi kaynak kısıtlı uygulamalar için ideal bir seçenek haline gelmiştir. Özellikle retinal görüntü analizi çalışmalarında, EfficientNet'in transfer learning yeteneği ve yüksek parametre verimliliği sayesinde sınırlı veri setleriyle bile başarılı sonuçlar elde edilebilmektedir.



Şekil 11. EfficientNet model ölçeklendirme stratejileri. (a) temel ağ mimarisi, (b) genişlik ölçeklendirme, (c) derinlik ölçeklendirme, (d) çözünürlük ölçeklendirme, (e) bileşik ölçeklendirme (compound scaling)

EfficientNet, geleneksel tek boyutlu ölçeklendirme yöntemlerinin aksine, bileşik ölçeklendirme (compound scaling) yaklaşımı kullanır (Şekil 11). Bu yöntemde, ağın derinliği, genişliği ve girdi çözünürlüğü sabit bir katsayı kullanılarak dengeli bir şekilde birlikte ölçeklendirilir. Bu yaklaşım, tek boyutlu ölçeklendirme yöntemlerine göre daha iyi doğruluk ve verimlilik sağlar (Tan & Le, 2019).

Transfer Learning ve Fine-Tuning Stratejileri: Transfer learning, derin öğrenme alanında özellikle sınırlı veri setlerine sahip uygulamalarda yaygın olarak kullanılan güçlü bir tekniktir ve büyük ölçekli veri setleri üzerinde önceden eğitilmiş modellerin bilgisini yeni bir görev için yeniden kullanma prensibine dayanmaktadır (Pan & Yang, 2010). Bu yaklaşım, özellikle tıbbi görüntüleme alanında kritik öneme sahiptir; çünkü etiketlenmiş tıbbi görüntülerin toplanması hem maliyetli hem de uzmanlık gerektiren bir süreçtir. ImageNet gibi milyonlarca görüntü içeren büyük veri setleri üzerinde eğitilmiş modeller, görüntülerdeki kenarlar, dokular, şekiller gibi düşük seviye özelliklerden, yüksek seviye semantik özelliklere kadar geniş bir özellik yelpazesi öğrenmiş olduklarından, bu öğrenilmiş temsiller farklı görsel tanıma görevlerine başarıyla transfer edilebilmektedir.

Transfer learning'in tıbbi görüntü analizindeki başarısı, özellikle Tajbakhsh ve arkadaşlarının (2016) çalışmasıyla kanıtlanmıştır; bu çalışmada ImageNet üzerinde önceden eğitilmiş ESA'ların farklı tıbbi görüntüleme modalitelerinde (göğüs röntgeni, kolonoskopi, ultrason) sıfırdan eğitilen modellere göre önemli ölçüde daha iyi performans gösterdiği ortaya konmuştur (Tajbakhsh vd., 2016). Benzer şekilde, Litjens ve arkadaşları (2017) tıbbi

görüntü analizinde derin öğrenme üzerine yaptıkları kapsamlı derlemede, transfer learning'in özellikle küçük veri setlerinde overfitting problemini azaltarak model genelleme yeteneğini artırdığını vurgulamışlardır (Litjens vd., 2017).

Fine-tuning stratejileri, transfer learning sürecinin en kritik bileşenlerinden birini oluşturmaktadır ve önceden eğitilmiş modelin hangi katmanlarının dondurulacağı, hangi katmanların yeniden eğitileceği ve öğrenme hızının nasıl ayarlanacağı gibi kararları içermektedir. Yosinski ve arkadaşları (2014), derin sinir ağlarının farklı katmanlarının transfer edilebilirliğini inceleyen öncü çalışmalarında, ilk katmanların daha genel özellikler (kenarlar, renkler) öğrendiğini ve bu özelliklerin farklı görevlere kolayca transfer edilebildiğini, son katmanlara doğru gidildikçe ise özelliklerin daha görev-spesifik hale geldiğini göstermişlerdir (Yosinski vd., 2014). Bu bulgular ışığında, yaygın bir strateji ilk katmanları dondurup (freeze) son katmanları hedef veri seti üzerinde fine-tune etmektir; ancak optimal strateji veri seti büyüklüğüne ve kaynak ile hedef domain arasındaki benzerliğe bağlı olarak değişmektedir.

Tıbbi görüntülemede fine-tuning için farklı stratejiler mevcuttur. Raghu ve arkadaşları (2019), tıbbi görüntüleme uygulamalarında transfer learning'in etkinliğini sistematik olarak incelemişler ve küçük tıbbi veri setlerinde bile ImageNet transfer learning'in faydalı olduğunu, ancak tıbbi görüntülerin ImageNet'ten farklı görsel özelliklere sahip olması nedeniyle agresif fine-tuning'in gerekli olabileceğini ortaya koymuşlardır (Raghu vd., 2019). Tipik bir fine-tuning protokolü, tüm ağı düşük bir öğrenme hızıyla (örneğin, 0.0001) eğitmeyi, ardından son katmanları daha yüksek bir öğrenme hızıyla fine-tune etmeyi içermektedir. Ayrıca, katmanları kademeli olarak çözerek eğitmek (progressive unfreezing stratejisi) bazı çalışmalarda daha iyi sonuçlar vermiştir. Retinopati tespiti gibi oftalmolojik görüntü analizi uygulamalarında, Gulshan ve arkadaşları (2016) diyabetik retinopati tespiti için Inception-v3 modelini fine-tune ederek oftalmolog seviyesinde performans elde etmişlerdir (Gulshan vd., 2016).

Güncel yaklaşımlar arasında, alan-adaptif transfer öğrenme (domain-adaptive transfer learning) ve çoklu-görev öğrenme (multi-task learning) gibi gelişmiş stratejiler de bulunmaktadır. Farklı görüntüleme modaliteleri arasında transfer learning için adversarial domain adaptation yöntemleri etkili sonuçlar vermektedir (Ganin & Lempitsky, 2015). Ayrıca, kendi-kendine denetimli öğrenme (self-supervised learning) yaklaşımları, etiketlenmemiş görüntülerin bolluğundan yararlanarak modellerin güçlü görsel temsiller öğrenmesini ve daha iyi başlangıç ağırlıklarıyla transfer edilmesini sağlamaktadır (Chen vd.,

2020). Bu stratejilerin seçimi, mevcut veri miktarına, hesaplama kaynaklarına ve hedef görevin karmaşıklığına bağlı olarak değişmekte olup, en iyi sonuçlar genellikle deneysel çalışmalar ve hiperparametre optimizasyonu ile elde edilmektedir.

Fusion Yaklaşımları: Tıbbi görüntü analizinde fusion (birleştirme) yaklaşımları, farklı bilgi kaynaklarını veya görüntü modalitelerini birleştirerek daha zengin özellik temsilleri elde etmeyi amaçlayan güçlü stratejilerdir. Özellikle retinal görüntü analizinde, ham fundus görüntülerinin segmente edilmiş damar yapılarıyla birleştirilmesi hem global görüntü özelliklerinin hem de lokal anatomik yapıların eşzamanlı olarak değerlendirilmesini sağlayarak sınıflandırma performansını önemli ölçüde artırmaktadır (Gu vd., 2019). Bu yaklaşım, özellikle retinopati tespiti gibi damar morfolojisinin kritik bir rol oynadığı uygulamalarda, modelin hem renk bilgisini, eksüda ve kanama gibi lezyonları içeren orijinal görüntüden, hem de damar genişliği, kıvrımlılık (tortuozite) ve dal yapısı gibi vasküler özellikleri segmentasyon maskesinden öğrenmesine olanak tanımaktadır.

Multi-stream ve dual-stream ESA mimarileri, fusion yaklaşımlarının en yaygın uygulama biçimlerini oluşturmaktadır. Bu mimarilerde, orijinal fundus görüntüler ve segmentasyon maskeleri ayrı evrimsel akışlar (stream) aracılığıyla işlenerek bağımsız özellik haritaları elde edilmekte, ardından bu özellikler farklı seviyelerde birleştirilerek final sınıflandırma yapılmaktadır. Wang ve arkadaşları (2021), diyabetik retinopati tespiti için geliştirdikleri dual-pathway modelinde, orijinal fundus görüntülerini bir ESA kolunda işlerken, damar segmentasyon maskelerini paralel bir kolda işlemiş ve bu iki akıştan elde edilen özellikleri concatenation yöntemiyle birleştirerek tek akışlı modellere göre %4.2 daha yüksek doğruluk elde etmişlerdir (Wang vd., 2021).

Fusion stratejilerinin uygulanmasında farklı birleştirme seviyeleri ve yöntemleri mevcuttur. Erken birleştirme (early fusion) yaklaşımında, ham görüntü ve segmentasyon maskesi giriş seviyesinde kanal olarak birleştirilmekte ve tek bir ESA ile işlenmektedir; bu yöntem, düşük seviye özellikler arasındaki etkileşimlerin erken öğrenilmesini sağlamaktadır (Huang vd., 2021). Orta seviye birleştirme (intermediate fusion) stratejisinde, her iki akışın orta katmanlarından elde edilen özellik haritaları birleştirilerek hem düşük hem de orta seviye özelliklerin entegrasyonu sağlanmaktadır. Geç birleştirme (late fusion) yaklaşımı ise, her akışın tam olarak işlenmesinden sonra yüksek seviye özelliklerin veya karar skorlarının birleştirilmesini içermektedir. Attention mekanizmalarının entegrasyonu, füzyon sürecini daha da iyileştirmekte ve modelin hangi bilgi kaynağına ne kadar ağırlık vereceğini dinamik olarak öğrenmesini sağlamaktadır.

Segmentasyon maskelerinin füzyon yaklaşımlarında kullanılması için farklı stratejiler geliştirilmiştir. U-Net gibi state-of-the-art segmentasyon modelleri kullanılarak elde edilen yüksek kaliteli damar maskeleri doğrudan kanal olarak eklenebileceği gibi, bu maskelerden türetilen morfolojik özellikler (damar kalınlığı haritaları, bifurkasyon noktaları, tortuoosite değerleri) de ek kanallar olarak modele dahil edilebilmektedir. Bazı yaklaşımlarda, segmentasyon maskesi binary (ikili) olarak değil, olasılık haritası (probability map) olarak kullanılmakta ve bu sayede modelin belirsiz bölgelerdeki bilgiyi de değerlendirmesi sağlanmaktadır. Ayrıca, multi-task learning paradigması altında, segmentasyon ve sınıflandırma görevleri eşzamanlı olarak öğrenilerek her iki görevin birbirini desteklemesi sağlanabilmektedir; bu yaklaşımda segmentasyon kolunun ürettiği attention haritaları, sınıflandırma kolunun odaklanması gereken bölgeleri belirlemektedir.

Retinopathy of Prematurity (ROP) gibi spesifik uygulamalarda, füzyon yaklaşımlarının etkinliği daha da belirginleşmektedir. Brown ve arkadaşları (2018), i-ROP çalışmasında plus hastalığı tespiti için damar özelliklerini ekstrakte ederek klasik görüntü özellikleriyle birleştirmişler ve bu hibrit yaklaşımın uzman oftalmologlarla karşılaştırılabilir performans gösterdiğini rapor etmişlerdir (Brown vd., 2018). Güncel çalışmalar, end-to-end öğrenilebilir füzyon mimarilerinin geliştirilmesine odaklanmakta ve otomatik segmentasyon ile sınıflandırmanın tek bir modelde entegre edildiği birleşik öğrenme yaklaşımları öne çıkmaktadır. Bu tür modeller, segmentasyon hatalarının sınıflandırma performansı üzerindeki etkisini minimize etmekte ve her iki görevin birbirini iyileştirmesini sağlamaktadır.

Literatürde retinal damar segmentasyonu üzerine yapılan çalışmalar incelendiğinde, geleneksel yöntemlerden derin öğrenme tabanlı yaklaşımlara kadar geniş bir yelpaze görülmektedir. Retina damarlarının doğru segmentasyonu, diyabetik retinopati, glokom, hipertansif retinopati ve prematüre retinopatisi (ROP) gibi oftalmolojik hastalıkların erken teşhis ve tedavi planlamasında kritik öneme sahiptir. Geleneksel segmentasyon yöntemleri, genellikle eşleşen filtreler (Chaudhuri vd., 1989), morfolojik işlemler (Zana & Klein, 2001) ve vasküler izleme algoritmalarına (Liu & Sun, 1993) dayanmaktadır. Ancak bu yöntemler, düşük kontrastlı damar yapıları ve arka plan gürültüsü gibi zorluklar karşısında yetersiz kalabilmektedir. Son yıllarda, derin öğrenmeye dayalı yöntemler bu sınırlamaları aşmak adına önemli bir ivme kazanmıştır. U-Net mimarisi, encoder-decoder yapısı ve atlamalı bağlantıları sayesinde, tıbbi görüntü segmentasyonunda en yaygın kullanılan derin öğrenme modellerinden biri haline gelmiştir (Ronneberger vd., 2015).

Attention U-Net atlama bağlantılarında dikkat mekanizmaları kullanarak özellikle ince damar yapılarının segmentasyonunda başarı sağlamıştır. UNet++ (Zhou vd., 2018), derin denetim ve iç içe geçmiş skip bağlantılarla segmentasyon başarımını artırmıştır. R2U-Net (Alom vd., 2018) ise residual ve recurrent blokları birleştirerek sınırlı veriyle daha güçlü temsil kabiliyeti elde etmeyi hedeflemiştir. MobileNetV2'nin bütünleştirildiği M2U-Net (Laibacher vd., 2019), UNet'in hesaplama verimliliğini artırarak gerçek dünya uygulamaları için uygun hale getirilmiştir.

LadderNet (Zhuang, 2019), U-Net mimarisini çoklu yollarla genişleterek daha fazla bağlamsal bilgi akışı sağlamış, bu da damar yapılarının daha hassas şekilde tespitine olanak tanımıştır. Guo ve arkadaşları, U-Net mimarisine uzamsal dikkat modülü ekleyerek SA-UNet modelini önermiştir (Guo vd., 2020). Bu model, özellikle düşük kontrastlı ve ince damarların bulunduğu bölgelerde segmentasyon performansını artırmaktadır. Khan tarafından önerilen T-Net ise düşük kaynaklı cihazlarda çalışabilecek hafif bir U-Net varyantıdır (Khan vd., 2022).

Segmentasyon doğruluğunu artırmak için üretici çekişmeli ağlar (GAN) da kullanılmaktadır. Özellikle RV-GAN (Kamran vd., 2021), hem ince damarların sürekliliğini hem de yanlış pozitif oranını dengeleyerek yüksek doğrulukta sonuçlar elde etmiştir. Transformer tabanlı modeller ve GAN ile kullanılan dikkat mekanizmaları ve çok ölçekli özellik rafine blokları segmentasyon kalitesini önemli ölçüde artırmıştır. VGA-Net (Jalali vd., 2024), grafik evrimsel ağları dikkat mekanizmalarıyla entegre ederek ince damarların segmentasyonunda üstün başarı sağlamıştır. GCC-UNet (Wei vd., 2024), kapsül evrimlerini kullanarak hem yerel hem de küresel damar yapılarının etkili öğrenimini sağlamıştır. SFNet (Yang vd., 2025), uzamsal ve frekans domain ağlarını birleştirerek özellikle geniş alanlı OCTA görüntülerinde etkili performans göstermiştir. UNet ve türevleri gerek klasik yapıları gerekse GAN ve transformer destekli gelişmiş versiyonları ile retina damar segmentasyonu alanında halen en etkin çözümler arasında yer almaktadır.

Prematüre retinopatisinin (ROP) otomatik teşhisi ve sınıflandırılması için literatürde farklı yaklaşımlar önerilmiştir. (Ataer-Cansizoglu vd., 2015), 2015 (i-ROP): Imaging and Informatics in ROP (i-ROP) sistemini geliştirilmiş, 77 görüntü üzerinde makine öğrenimi (SVM sınıflandırıcı) kullanarak %95 doğruluk elde edilmiş ve görüntü özelliklerinin uzman tanısıyla korelasyonu incelenmiştir. (Worrall vd., 2016) ilk tam otomatik ROP tespiti yaklaşımlarından biri olarak pretrained GoogLeNet ince-ayarı ve Bayesçi belirsizlik tahmini ile ROP görüntülerinde otomatik tespit sunmuştur. Brown ve ark. (2018), i-ROP DL

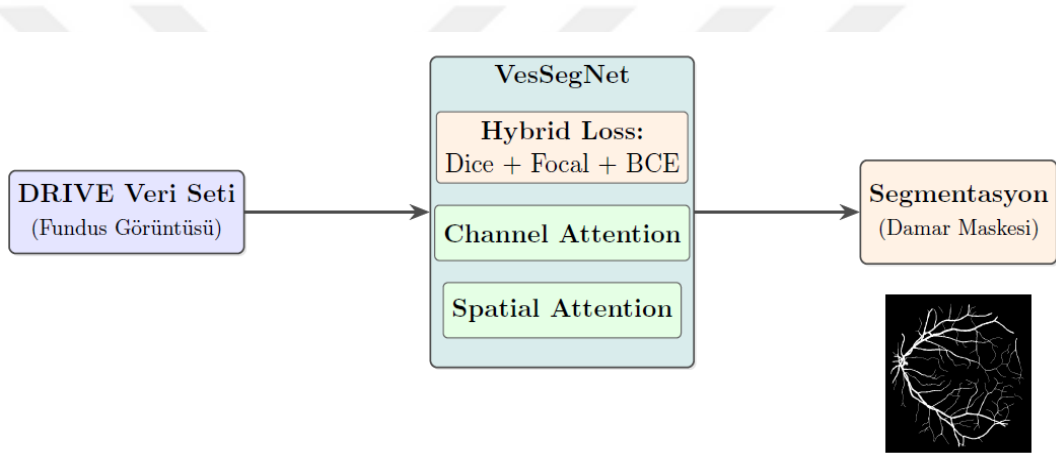
çalışmasında plus disease için büyük, çok-merkezli etiketlenmiş veriyle eğitilmiş derin CNN kullanarak algoritmanın uzman hekimlerle karşılaştırılabilir performans gösterdiğini raporlamıştır. (Wang vd., 2018) DeepROP adlı iki aşamalı (identification ve grading) derin öğrenme tabanlı otomatik ROP tarama sistemi geliştirmiştir (Mulay vd., 2019). Mask R-CNN tabanlı yaklaşım ile demarcation line / ridge tespiti yaparak özellikle ROP evre 2 tespitini destekleyecek yerelizasyon odaklı bir yöntem sunmuştur.

(Wang vd., 2021) multidimensional deep learning platform ile ROP lezyonlarını çok-boyutlu olarak tanımlayan ve sınıflandıran geniş ölçekli bir DL platformunun tarama için uygun olabileceğini göstermiştir. (R. Zhang vd., 2022) özellikle AP-ROP (Aggressive Posterior ROP) tespiti için özelleştirilmiş otomatik tanı yaklaşımları (iki aşamalı görev bölünmesi) önermiştir. (Kıran Yenice vd., 2024) RegNetY002 Modeli (2024), 634 fundus görüntüsü üzerinde Tip 1, Tip 2 ve A-ROP tespiti için geliştirilmiştir. Model, Tip 2 ROP tespitinde %98 doğruluk ve %98 özgüllük; Stage 2 vs Stage 3 ROP ayrımında %90 doğruluk ve %95 özgüllük elde etmiştir. eriferik retina görüntülerinde Canny edge detection algoritması ile ROP sınırları tespit edilmiş, sonrasında YOLO v7 detector algoritması kullanılarak ilgi bölgesi (ROI) kırmızı dikdörtgenlerle belirlenmiştir. (Rahim vd., 2024) i-ROP DL System, iki aşamalı ESA kullanarak ROP şiddet skorunu hesaplayan ilk sistem olmuştur. 6.000 görüntü ile eğitilen sistem, normal, pre-plus ve plus ROP evrelerini sınıflandırmada %91 doğruluk, plus hastalığı tespitinde %93 duyarlılık ve %94 özgüllük elde etmiştir.

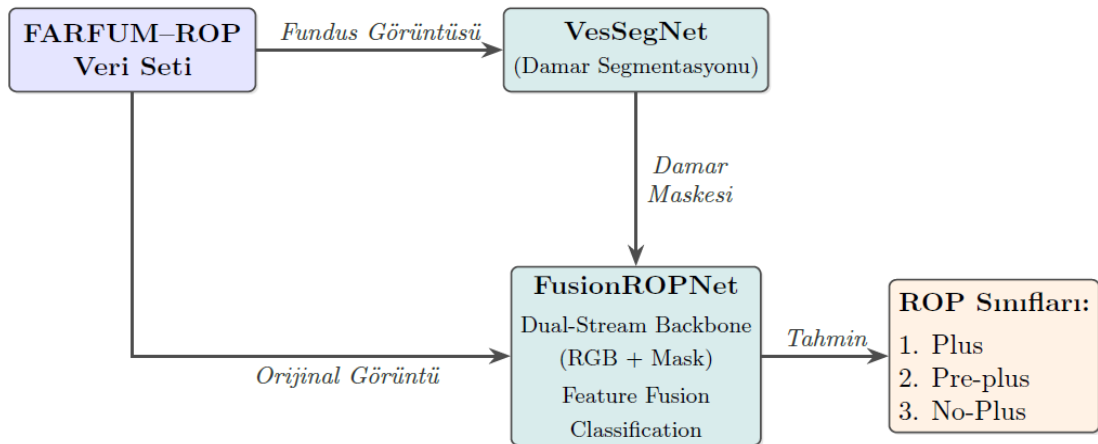
Bu çalışmalar, derin öğrenme tabanlı yaklaşımların otomatik ROP tanısında önemli bir potansiyele sahip olduğunu göstermektedir; farklı hastalık evreleri ve alt tipleri için %90 ile %98 arasında doğruluk oranları elde edilmiştir. Geleneksel görüntü işleme tekniklerinden sofistike çok aşamalı derin öğrenme mimarilerine geçiş, alanın daha doğru ve klinik olarak uygulanabilir tarama sistemlerine doğru evrimini vurgulamaktadır. Bu umut verici sonuçlara rağmen, metodolojilerin standardizasyonu, veri seti çeşitliliğinin genişletilmesi ve farklı klinik ortamlarda genellenebilirliğin sağlanması konularında zorluklar devam etmektedir ve bu konular gelecek araştırmalarda daha fazla incelemeyi gerektirmektedir.

2. YAPILAN ÇALIŞMALAR

Bu tez çalışmasının amacı prematüre retinopatisi (ROP) tanısında damar morfolojisini açık ve sistematik biçimde modele entegre eden, iki aşamalı ve uçtan uca bir derin öğrenme çerçevesi geliştirmektir. Önerilen yaklaşımın birinci aşamasında, retinal damar segmentasyonu için özel olarak tasarlanmış VesSegNet mimarisi geliştirilmiştir. VesSegNet, literatürde yaygın olarak kullanılan ve referans damar segmentasyon veri setlerinden biri olan DRIVE (Digital Retinal Images for Vessel Extraction) veri seti üzerinde eğitilmiştir. VesSegNet, rop fundus görüntülerinden güvenilir damar maskeleri üretmiştir (Şekil 12).



Şekil 12. 1.Aşamada önerilen VesSegNet mimarisi iç yapısı ve eğitim bileşenleri



Şekil 13. 2.Aşamada önerilen FusionROPNet mimarisi iç yapısı ve eğitim bileşenleri

İkinci aşamada ise (Şekil 13), ROP veri setine ait fundus görüntüleri VesSegNet modeli aracılığıyla işlenerek her bir görüntü için damar segmentasyon maskeleri elde edilmiştir. Ardından, orijinal fundus görüntüsü ile VesSegNet tarafından üretilen damar maskeleri, her bir kaynaktan elde edilen derin özniteliklerin yan yana birleştirilmesi prensibine dayanan iki kollu bir füzyon ağı kullanılarak birleştirilmiştir.

Bu füzyon işlemi, hem ham görüntü verilerinin içerdiği genel özellikler hem de damar yapılarının morfolojik özellikleri gibi spesifik bilgilerin aynı anda modele sunulmasını sağlamaktadır. Füzyon sonrası elde edilen çok kanallı görüntü temsili, ROP plus hastalığı sınıflandırılması için tasarlanan derin öğrenme modeline giriş olarak verilmiştir.

Geleneksel yaklaşımlarda, derin öğrenme modelleri bu özellikleri dolaylı olarak öğrenmeye çalışmakta, ancak öğrenme süreci kara kutu özelliği taşımakta ve hangi özelliklerin karar mekanizmasında etkili olduğu belirsiz kalmaktadır. Önerilen iki aşamalı yaklaşımda ise, damar morfolojisi açık biçimde modellendiği için, sınıflandırma sürecinin yorumlanabilirliği artmakta ve klinik karar destek sistemleri açısından güvenilirlik yükselmektedir. Ayrıca, damar segmentasyonunun ayrı bir modül olarak tasarlanması, sistemin modüler yapısını güçlendirmekte ve her iki aşamanın bağımsız olarak optimize edilebilmesine olanak tanımaktadır.

Önerilen iki aşamalı derin öğrenme çerçevesinin birinci aşaması, damar segmentasyonu görevini üstlenen VesSegNet olarak adlandırılmıştır. İkinci aşama ise, damar maskeleri ile orijinal görüntülerin füzyonunu gerçekleştiren ve ROP evre sınıflandırmasını yapan FusionROPNet olarak tanımlanmıştır. Bu isimlendirme, tez boyunca kullanılacak olup, metodolojinin anlaşılabilirliğini ve takip edilebilirliğini kolaylaştırmayı amaçlamaktadır. VesSegNet ve FusionROPNet mimarilerinin entegrasyonu, tezin orijinal katkısını oluşturmakta ve ROP tanısında damar morfolojisinin sistematik kullanımına yönelik yenilikçi bir çerçeve sunmaktadır.

Geliştirilen iki aşamalı derin öğrenme yaklaşımı, ROP tanısında doğruluğun artırılması, yorumlanabilirliğin iyileştirilmesi ve klinik kullanıma uygun otomatik karar destek sistemlerinin geliştirilmesi hedeflerine odaklanmaktadır.

2.1 Retinal Damar Segmentasyonu Aşaması

2.1.1. VesSegNet: Önerilen Segmentasyon Ağı

Önerilen iki aşamalı derin öğrenme çerçevesinin birinci aşamasında, retinal damar segmentasyonu görevini üstlenen VesSegNet mimarisi geliştirilmiştir. VesSegNet, tıbbi görüntü segmentasyonu alanında yaygın olarak kullanılan ve literatürde referans mimarilerden biri olan Attention U-Net mimarisini temel almaktadır.

Klasik U-Net yapısından farklılaşan VesSegNet modeli; Attention Block (AG) yapıları, ağı her seviyesine entegre edilen kanal ve uzamsal dikkat mekanizmaları ve hibrit kayıp fonksiyonu olmak üzere üç temel özelliikle zenginleştirilmiştir. Bu özelleştirilmiş mimari, retinal damar yapılarının ince morfolojik detaylarını çok daha yüksek bir hassasiyetle yakalayabilmek amacıyla tasarlanmıştır.

Bu çalışma kapsamında, retina damar segmentasyonu görevinde en yüksek performansı sağlayan mimariyi tespit etmek amacıyla; U-Net, Attention U-Net, U-Net++ ve R2U-Net olmak üzere dört farklı U-Net türevi kodlanmış ve DRIVE veri seti üzerinde eğitilmiştir. En iyi sonucu vereni belirlemek adına tüm mimariler, standartlaştırılmış bir karşılaştırma zemininde, aynı hiperparametreler ve eğitim koşulları altında PyTorch ortamında analiz edilmiştir.

VesSegNet Mimari Yapısı ve Dikkat Mekanizmaları: VesSegNet mimarisi, giriş katmanından çıkış katmanına kadar toplam dört seviyeli hiyerarşik bir yapıya sahiptir ve sırasıyla 64, 128, 256 ve 512 filtre sayıları kullanılmıştır. Kodlayıcı (encoder) bloklarında, her konvolüsyon işleminin ardından Batch Normalization ve ReLU aktivasyon fonksiyonu uygulanmıştır. Mimari, iki farklı dikkat mekanizması ile zenginleştirilmiştir: (i) Kanal dikkat mekanizması (Channel Attention), her bir kodlayıcı ve kod çözücü katmanında AdaptiveAvgPool2d ve AdaptiveMaxPool2d işlemleri kullanılarak hesaplanmış, indirgeme oranı 8 olarak belirlenmiş ve sigmoid aktivasyonu ile $[0,1]$ aralığına normalize edilmiştir. (ii) Uzamsal dikkat mekanizması (Spatial Attention), kanal boyutunda ortalama ve maksimum değerlerin hesaplanması ile elde edilen 2 kanallı öznitelik haritası üzerine 5×5 boyutunda konvolüsyon filtresi uygulanarak önemli uzamsal bölgelerin belirlenmesini sağlamıştır. Geçit dikkat bloğu (Gating Attention Block), kod çözücü katmanlarında skip connection'lar üzerinde uygulanmış, kodlayıcıdan gelen öznitelikler (x) ile kod çözücünden gelen öznitelikler (g) 1×1 konvolüsyon ile dönüştürülmüş, LeakyReLU

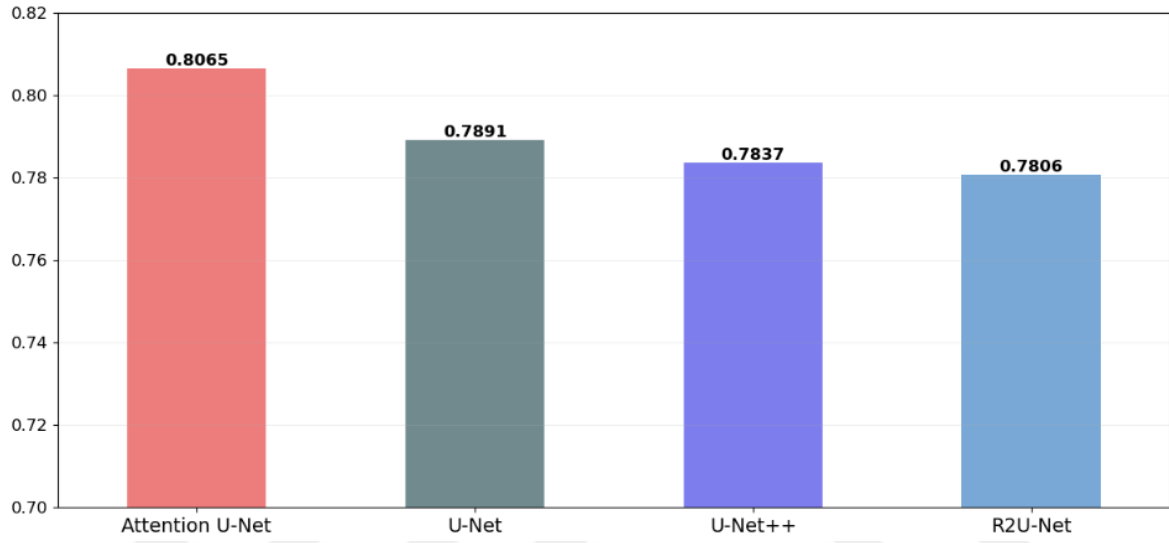
(negative_slope=0.2) aktivasyonu sonrası sigmoid fonksiyonu ile dikkat ağırlıkları hesaplanmış ve kodlayıcı öznitelikleri bu ağırlıklar ile çarpılarak filtrelenmiştir. Her dikkat bloğunun çıktısı, ilgili öznitelik haritası ile eleman bazında çarpılarak (element-wise multiplication) adaptif öznitelik zenginleştirme işlemi gerçekleştirilmiştir. Kod çözücü bloklarında, 2×2 boyutunda stride=2 parametrelili ConvTranspose2d katmanları kullanılarak öznitelik haritalarının boyutu kademeli olarak artırılmış, her yukarı örnekleme işleminin ardından dikkat bloğu ile filtrelenmiş kodlayıcı öznitelikleri concatenate edilerek DoubleConv modülünden geçirilmiş ve son olarak kanal ve uzamsal dikkat mekanizmaları ile çıktı öznitelikleri zenginleştirilmiştir.

Veri Ön İşleme ve Augmentasyon Pipeline: Retinal görüntülerin eğitime hazırlanması için standart bir ön işleme boru hattı uygulanmıştır. RGB görüntüler gri tonlamaya dönüştürülmüş, CLAHE algoritması (clipLimit=2.0, tileGridSize=8×8) ile kontrast iyileştirmesi yapılmış ve gamma=1.2 ile kontrast düzeltilmesi uygulanmıştır. Son aşamada görüntüler mean=0.5 ve std=0.5 ile [-1,1] aralığına normalize edilmiştir.

Eğitim setinin sınırlı boyutu nedeniyle, modelin genelleme kapasitesini artırmak amacıyla geometrik ve fotometrik dönüşümlerden oluşan kapsamlı bir augmentasyon stratejisi benimsenmiştir. Eğitim modunda 128×128 boyutunda RandomCrop, %50 olasılıkla yatay ve dikey çevirme, $\pm 30^\circ$ rotasyon ve $\pm 20\%$ parlaklık/kontrast değişimi (ColorJitter) uygulanmıştır. Test modunda ise sadece 512×512 boyutuna Resize işlemi gerçekleştirilmiştir.

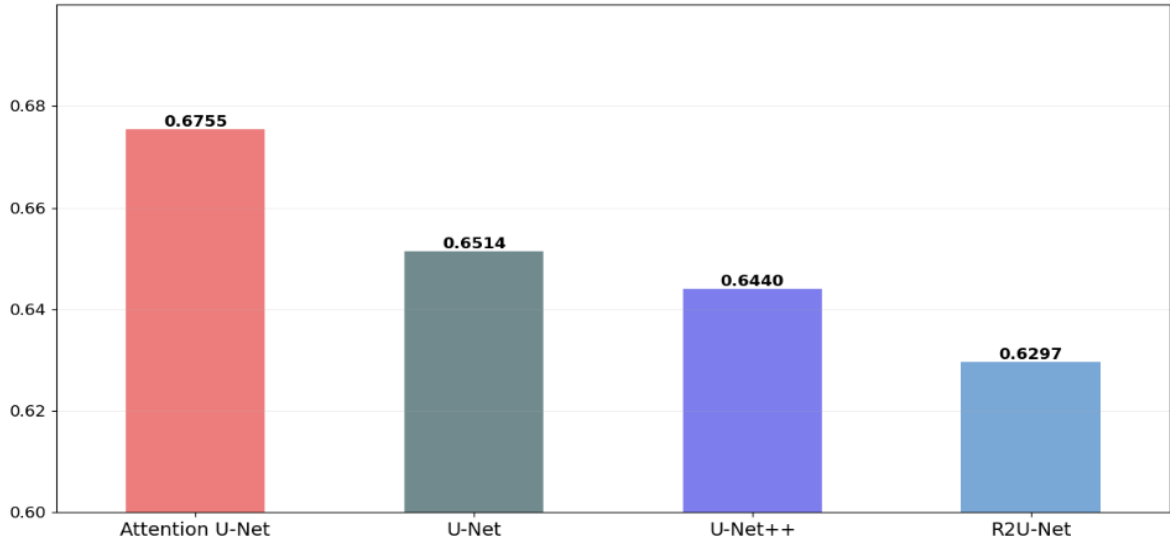
Hibrit Kayıp Fonksiyonu ve Optimizasyon: VesSegNet eğitiminde sınıf dengesizliğini ele almak ve segmentasyon kalitesini artırmak için üç bileşenli hibrit kayıp fonksiyonu kullanılmıştır: BCEWithLogitsLoss (pos_weight=5.0 ile damar piksellerine öncelik), Dice Loss (smooth=1e-6 ile örtüşme optimizasyonu) ve Focal Loss (gamma=2.0 ile zor örneklerle odaklanma). Kayıp ağırlıkları başlangıçta alpha=0.3, beta=0.4 olarak ayarlanmış ancak nn.Parameter ile öğrenilebilir hale getirilmiş ve softmax normalizasyonu ile eğitim sırasında adaptif olarak optimize edilmiştir. Gradient patlamasını önlemek için max_grad_norm=1.0 ile gradient clipping uygulanmıştır. Optimizasyon için AdamW optimizer (lr=0.001, weight_decay=1e-4, betas=(0.9, 0.999)) tercih edilmiş, OneCycleLR scheduler ile super-convergence stratejisi benimsenmiştir (max_lr=0.01, pct_start=0.3, div_factor=25, final_div_factor=1e4, anneal_strategy='cos'). Learning rate her batch sonrası scheduler.step() ile dinamik olarak güncellenmiştir.

Model performansları, altı temel metrik üzerinden değerlendirilmiştir: Dice katsayısı (Dice coefficient), birleşim üzerindeki kesişim (intersection over union- IoU), kesinlik (precision), duyarlılık (recall), F1 skoru (F1-score) ve doğruluk (accuracy). Tüm modeller aynı veri seti, ön işleme adımları, hiperparametreler ve değerlendirme kriterleri ile test edilmiştir. Böylece modellerin yalnızca mimari yapılarından kaynaklanan performans farkları karşılaştırılabilir şekilde analiz edilmiştir.



Şekil 14. U-Net mimarilerinin test veri kümesi F1 skoru karşılaştırması

Şekil 14'te, test veri kümesi üzerinde farklı U-Net varyantlarının F1 skoru performansları karşılaştırılmaktadır. Attention U-Net, 0.8065 F1 skoru ile en yüksek performansı sergilerken, standart U-Net (0.7891), U-Net++ (0.7837) ve R2U-Net (0.7806) daha düşük skorlar elde etmiştir. Dikkat mekanizmalarının entegrasyonu, modelin damar yapılarına odaklanma yeteneğini artırarak segmentasyon başarısını iyileştirmiştir.



Şekil 15. U-Net mimarilerinin test veri kümesi IoU karşılaştırması

Şekil 15'te U-Net varyantlarının IoU performansları karşılaştırılmaktadır. Attention U-Net, 0.6755 IoU skoru ile en başarılı mimari olurken, standart U-Net (0.6514), U-Net++ (0.6440) ve R2U-Net (0.6297) daha düşük skorlar elde etmiştir. IoU metriği, tahmin edilen ve gerçek segmentasyon maskeleri arasındaki örtüşme oranını ölçtüğü için, dikkat mekanizmalarının piksel düzeyinde doğruluğu artırdığını göstermektedir. Sonuçlar, F1 skoru bulgularıyla tutarlı olup Attention U-Net'in üstünlüğünü doğrulamaktadır.

Tablo 1. U-Net varyantlarının segmentasyon başarısının karışıklık matrisi analizi

Model	Gerçek Pozitif (TP) Vessel	Gerçek Negatif (TN) Background	Yanlış Pozitif (FP) Background	Yanlış Negatif (FN) Vessel
Attention U-Net	355k (355.000)	4.72m* (4.720.000)	69.3k (69.300)	103k (103.000)
R2U-Net	387k* (387.000)	4.63m (4.630.000)	151k (151.000)	70.8k* (70.800)
U-Net ++	354k (354.000)	4.69m (4.690.000)	92.8k (92.800)	103k (103.000)
U-Net	352k (352.000)	4.70m (4.700.000)	82.7k* (82.700)	106k (106.000)

Tablo 1, farklı U-Net mimarilerinin karışıklık matrisi analiz sonuçlarını göstermektedir. R2U-Net, 387.000 doğru pozitif (TP) ile en yüksek damar tespit oranına sahipken, 151.000 yanlış pozitif (FP) değeri ile yüksek hata içermektedir. Attention U-Net, 355.000 TP ile dengeli bir performans sergilerken, 69.300 FP değeri ile en düşük yanlış alarm oranına sahiptir. Doğru negatif (TN) açısından, Attention U-Net'in 4.720.000 piksel ile arka plan tespitinde en başarılı model olduğu görülmektedir. Yanlış negatif (FN) değerleri incelendiğinde, R2U-Net 70.800 ile en az damar kaybına sahipken, U-Net 106.000 ile en yüksek kaçırma oranını göstermektedir. Attention U-Net hem yüksek doğruluk hem de düşük hata oranları ile en dengeli segmentasyon performansını sunmaktadır.

Tablo 2. U-Net varyantlarının model eğitiminde kullanılan hiperparametreler

Hyperparameter	Value	Hyperparameter	Value
Adam_betas	[0.9,0.999]	Optimizer	AdamW
Batch_size	4	Weight_decay	0,0001
Epochs	100	Loss_alpha	0,4
Validation_split	0,2	Loss_beta	0,4
Loss_function	Hybrid Loss	Loss_gamma	0,5
Lr_scheduler	OneCycleLR	Learning_rate	0,001

Tablo 2, model eğitiminde kullanılan hiperparametreleri göstermektedir. Eğitim, AdamW optimizasyonu ve OneCycleLR zamanlayıcısı ile 0.001 başlangıç öğrenme oranında gerçekleştirilmiştir. 100 epoch boyunca 4 batch boyutu kullanılmış, veri setinin %20'si doğrulama için ayrılmıştır. Aşırı öğrenmeyi önlemek amacıyla 0.0001 ağırlık azalması uygulanmıştır. Kayıp fonksiyonu olarak Hybrid Loss tercih edilmiş ve BCE, Dice Loss ile Focal Loss bileşenleri sırasıyla 0.4, 0.4 ve 0.2 ağırlıklarıyla dengelenmiştir.

Yapılan çalışmalar sonucunda, U-Net varyantları içinde Attention U-Net, 0.8085 F1 skoru ile en yüksek performansı göstermiştir. Model, tüm değerlendirme metriklerindeki dengeli başarısı ve düşük doğrulama kaybı sayesinde en güvenilir ve önerilebilir model olarak öne çıkmıştır. Bununla birlikte, kullanılan veri setinin sınırlı sayıda görüntü içermesi, modeller arasındaki performans farklarının yeterince belirginleşmesini engellemiş olabilir.

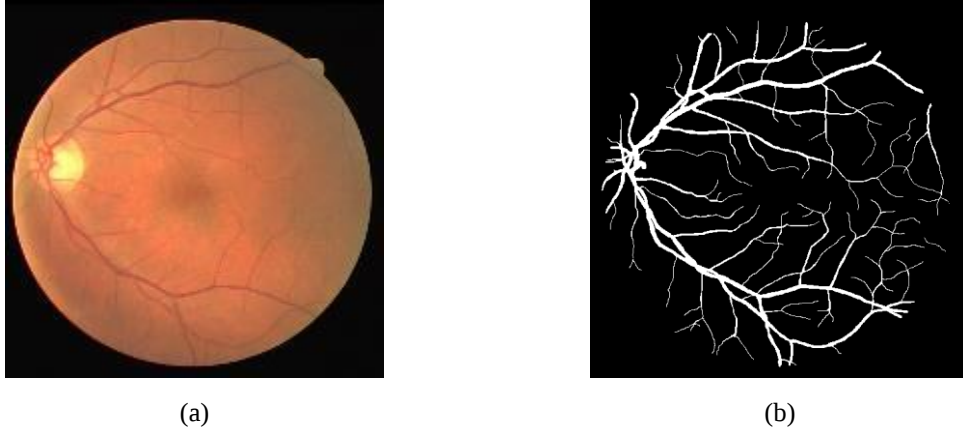
Bu nedenle, daha büyük ve çeşitlendirilmiş veri kümeleriyle yürütülecek ileri çalışmaların, model performansları arasında daha net ayrımlar ortaya koyması beklenmektedir.

Attention U-Net; Klasik U-Net'teki doğrudan birleştirme yerine, AttentionBlock kullanarak kodlayıcıdan gelen bilgiyi, kod çözücüdeki bağlama uygun şekilde filtrelemektedir. Bu, alakasız arka plan bilgisini engellemektedir. Model, sadece atlama bağlantılarında değil, ağın her seviyesinde (hem encoder hem decoder çıkışlarında) kanal ve uzamsal dikkat uygulayarak damar segmentasyonunun gerektirdiği ince morfolojik detaylara ve kritik kanallara sürekli olarak odaklanmasını sağlamaktadır.

VesSegNet mimarisi, çözümleyici (encoder) ve çözücü (decoder) olmak üzere iki temel bileşenden oluşmaktadır. Encoder bloğu, giriş görüntüsünden hiyerarşik özellik haritaları çıkarmaktan sorumludur ve görüntünün alt seviye (kenar, doku) ile üst seviye (şekil, yapısal bağlam) özelliklerini öğrenir. Decoder bloğu ise, encoder tarafından üretilen soyut özellik temsillerini orijinal görüntü çözünürlüğüne geri dönüştürerek piksel seviyesinde segmentasyon haritası oluşturur. VesSegNet'in ayırt edici özelliği, bu iki katman içerisinde entegre edilen dikkat mekanizmaları ve kullanılan dinamik ağırlıklı hibrit kayıp fonksiyonudur. Bu mekanizmalar, modelin öğrenme kapasitesini artırarak damar yapılarına odaklanmasını ve ilgisiz arka plan bilgilerini baskılamasını sağlamaktadır.

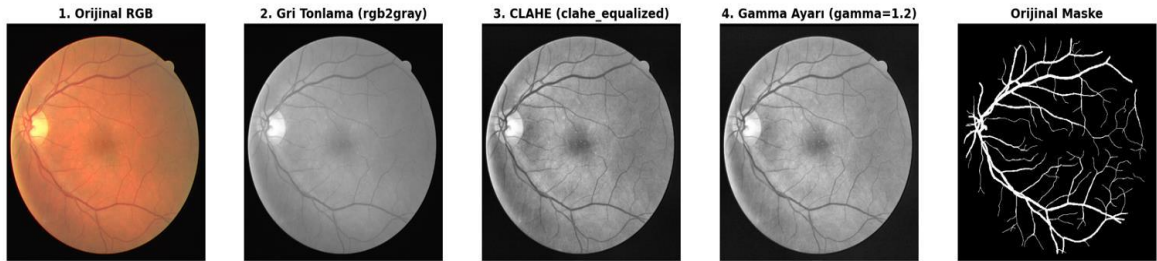
2.1.2. VesSegNet Eğitimi İçin Kullanılan Veri Seti: DRIVE

VesSegNet modelinin eğitimi için DRIVE (Digital Retinal Images for Vessel Extraction) veri seti kullanılmıştır. Hollanda'daki bir diyabetik retinopati tarama programından elde edilmiştir. Görüntüler, 45 derecelik görüş alanına (FOV) sahip Canon CR5 non-midriyatik 3CCD kamera kullanılarak elde edilmiştir. Bu veri seti, 565x584 piksel boyutunda 40 adet fundus görüntüsünden oluşmakta olup, 20 eğitim ve 20 test görüntüsü olarak ayrılmıştır. Veri miktarı az olmasına rağmen DRIVE veri seti, retina damar segmentasyonu problemlerinde yaygın olarak kullanılan ve referans kabul edilen bir kıyaslama veri setidir.(Staal vd., 2004) Veri seti; sağlıklı ve hafif diyabetik retinopati olgularını içermektedir.



Şekil 16. DRIVE veri seti örneği: (a) fundus görüntüsü, (b) damar segmentasyon maskesi

Şekil 16, DRIVE veri setinden örnek bir fundus görüntüsünü (a) ve karşılık gelen damar segmentasyon maskesini (b) göstermektedir. Görüntüler, bu çalışma kapsamında geliştirilen RetinaDataset sınıfı aracılığıyla işlenmiştir. İlk olarak, Attention U-Net modelinin giriş gereksinimlerini karşılamak ve hesaplama verimliliğini artırmak amacıyla görüntüler 512x512 piksel boyutuna yeniden boyutlandırılmıştır.



Şekil 17. Retina görüntüsünün ön işleme aşamaları: (1) orijinal RGB fundus, (2) gri tonlamaya dönüştürülmüş görüntü, (3) CLAHE ile kontrast iyileştirmesi, (4) gamma düzeltme ($\gamma=1.2$) uygulanmış görüntü- kontrast optimizasyonu, (5) orijinal maske

Şekil 17, damar yapılarının daha belirgin hale gelmesi için görüntüler gri tonlamaya çevrilmiştir. Düşük kontrastlı bölgelerdeki damar detaylarını iyileştirmek ve yerel kontrast farklılıklarını dengelemek için CLAHE (Contrast Limited Adaptive Histogram Equalization) uygulanmıştır. Görüntülerin kontrastını optimize etmek ve damar yapılarının belirginliğini artırmak amacıyla gamma düzeltme ($\gamma = 1.2$) işlemi gerçekleştirilmiştir. Son olarak, model eğitiminin kararlılığını sağlamak ve gradyan güncellemelerini dengelemek için görüntülerin piksel değerleri $[-1, 1]$ aralığına normalize edilmiştir. Son sütunda yer alan

orijinal maske, segmentasyon hedefini göstermekte ve ön işleme adımlarının damar yapısını koruyarak kontrast iyileştirmesi sağladığı doğrulanmaktadır. Bu ardışık işlem hattı, modelin eğitim aşamasında daha zengin özellik öğrenmesini ve veri setlerinde daha tutarlı segmentasyon performansı sergilemesini sağlamaktadır.

2.1.3. Hibrit Kayıp Fonksiyonu Bileşenleri

Bu tez çalışmasında segmentasyon için kullanılan hibrit kayıp fonksiyonu, sınıf dengesizliği ve ince damar sınırlarının duyarlılıkla yakalanması hedefleri doğrultusunda birden çok kayıp fonksiyonun birleştirilmesiyle tasarlanmıştır. Temel bileşenler Dice Loss, Focal Loss ve Binary Cross Entropy Loss'tur; ayrıca bu bileşenleri ağırlıklı toplayan birleştirilmiş bir yapı (Hybrid Loss) mevcuttur. Eğitim stabilitesini artırmak ve patlayan gradyan (exploding gradient) problemini önlemek amacıyla, gradyanların belirli bir eşik değerini (threshold) aşmaması için gradient clipping uygulanmıştır.

Dice benzerlik katsayısının (Dice similarity coefficient- DSC) bir türevi olan bu kayıp fonksiyonu, özellikle sınıf dengesizliğinin (örneğin, fundus görüntüsünde arka plan piksellerinin damar piksellerine oranla çok daha fazla olması) yaygın olduğu segmentasyon görevlerinde kullanılır.

$$\text{Dice}(p, y) = \frac{2 \sum_i (p_i \cdot y_i) + \varepsilon}{\sum_i p_i + \sum_i y_i + \varepsilon} \quad (11)$$

$$L_{\text{Dice}} = 1 - E[\text{Dice}] \quad (12)$$

- p_i : piksel için model tarafından tahmin edilen olasılık değeri (0-1 arası)
- y_i : piksel için gerçek etiket değeri (0 veya 1)
- $\sum_i$: tüm pikseller üzerinden toplam
- ε : sıfıra bölünmeyi önlemek için küçük bir sabit (genellikle),
- $\text{Dice}(p, y)$: tahmin ile gerçek etiket arasındaki örtüşme ölçütü (0-1 arası)

Dice kayıp fonksiyonu iki küme arasındaki örtüşmeyi ölçer. Formül, temel olarak tahmin edilen (pred) ve gerçek etiket (target) maskeleri arasındaki kesişimin iki katının, bu

iki maskenin ayrı ayrı toplamalarına oranlanmasıyla elde edilir. Amaç: Modelin, özellikle nadir görülen ve kritik olan ön plan (ilgi alanı) piksellerini doğru tahmin etmeye odaklanmasını sağlamaktır.

Kolay örneklerin etkisini azaltıp zor örneklerle odaklanır; dengesiz sınıflarda pozitif piksel hatalarını daha görünür kılar. Uygulamada önce sigmoid ile olasılığa geçilir ve ikili çapraz entropi üzerine odak (focusing) terimi uygulanır:

$$L_{\{BCE\}} = -[y \log(p) + (1 - y) \log(1 - p)] \quad (13)$$

$$L_{\{Focal\}} = (1 - \exp(-L_{\{BCE\}}))^{\gamma} \times L_{\{BCE\}} \quad (14)$$

Binary Cross Entropy (BCE) kayıp fonksiyonu, ikili sınıflandırma problemlerinde en yaygın kullanılan temel kayıp fonksiyonlarından biridir. Segmentasyon bağlamında her piksel, bir sınıfa (örneğin “damar” veya “arka plan”) ait olma olasılığıyla değerlendirilir. BCE, her piksel için tahmin edilen olasılık ile gerçek etiket arasındaki farkı ölçer ve modelin bu farkı minimize etmesini sağlar.

Matematiksel olarak BCE kaybı şu şekilde tanımlanır:

$$BCE(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (15)$$

- y_i : piksel i için gerçek etiket değeri (0 veya 1)
- $\hat{y}_i$: model tarafından tahmin edilen olasılık değeri
- N : toplam piksel sayısıdır

BCE, modelin yanlış tahminlerde bulunduğu pikselleri cezalandırarak tahmin olasılıklarını doğru sınıf yönünde güncellemesini sağlar. Ancak, sınıf dengesizliğinin yüksek olduğu veri setlerinde (örneğin, damar piksellerinin toplam piksellere oranının çok düşük olduğu fundus görüntülerinde) BCE, baskın sınıfı öğrenmeye eğilimlidir.

Hybrid Loss fonksiyonu üç bileşeni doğrulama performansına göre ayarlanabilir sabit ağırlıklarla toplar:

$$Hybrid = W_{Dice} * L_{Dice} + W_{Focal} * L_{Focal} + W_{BCE} * L_{BCE} \quad (16)$$

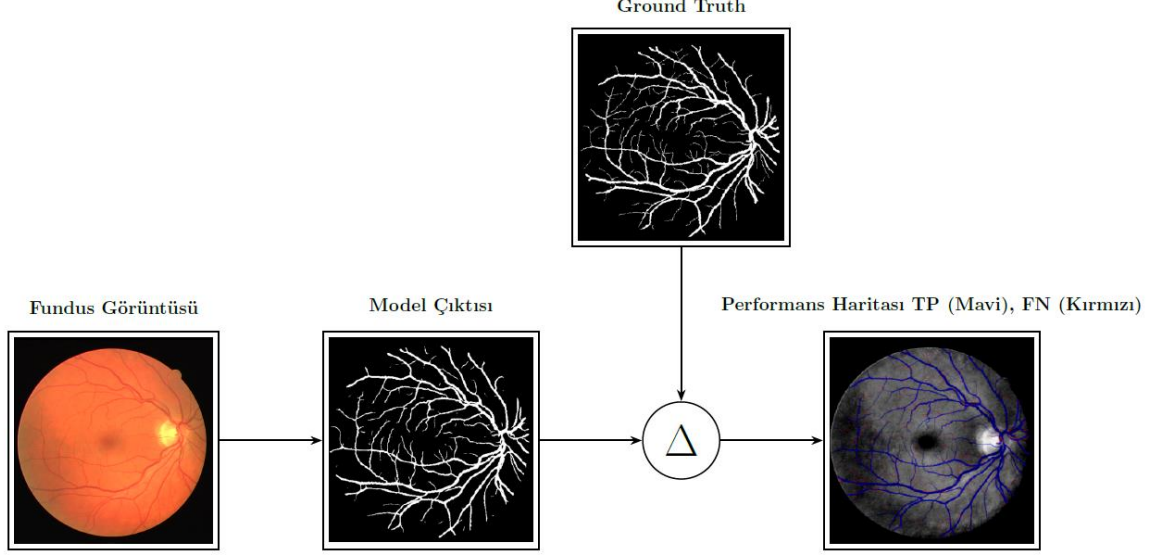
Ağırlıklar ($w_{\cdot}$) hiperparametrelerdir. BCE (nn.BCEWithLogitsLoss) standart ikili sınıflandırma performansını sağlar. Logit girişini kabul ettiği için sayısal kararlılığı artırır. Dice Loss sınıf dengesizliğini ele alır ve örtüşmeyi maksimize etmeye odaklanır. Focal Loss kolay örneklerin etkisini azaltarak, zorlu piksellere odaklanmayı sağlar. Bu hibrit kayıp fonksiyonu, özellikle küçük, seyrek ve kritik nesnelerin segmentasyonunda (örneğin damarlar, lezyonlar) performansı artırmak için güçlü bir araçtır.

Tablo 3. Kayıp fonksiyonlarının karışıklık matrisi analizi

Model	Gerçek Pozitif (TP) Vessel	Gerçek Negatif (TN) Background	Yanlış Pozitif (FP) Background	Yanlış Negatif (FN) Vessel
Hybrid Loss	355k (355.000)	4.72m* (4.720.000)	69.3k (69.300)	103k (103.000)
Dice Loss	387k* (387.000)	4.63m (4.630.000)	151k (151.000)	70.8k* (70.800)
Focal Loss	354k (354.000)	4.69m (4.690.000)	92.8k (92.800)	103k (103.000)
BCE	352k (352.000)	4.70m (4.700.000)	82.7k* (82.700)	106k (106.000)

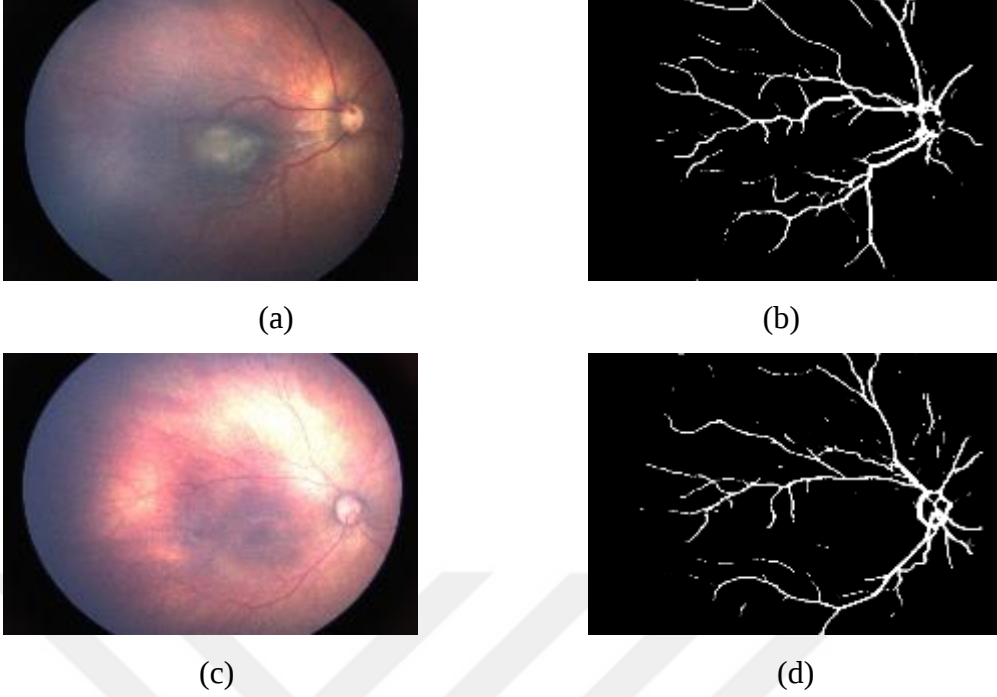
Tablo 3, farklı kayıp fonksiyonlarının karışıklık matrisi sonuçlarını karşılaştırmaktadır. Dice Loss, 387.000 doğru pozitif (TP) ile en yüksek damar tespitini sağlarken, 151.000 yanlış pozitif (FP) ile yüksek yanlış alarm oranına sahiptir. BCE, 106.000 yanlış negatif (FN) ile en fazla damar kaybını göstermektedir. Önerilen Hybrid Loss, 355.000 TP ve 69.300 FP değerleriyle en dengeli performansı sergilemektedir. 4.720.000 doğru negatif (TN) değeri, Hybrid Loss'un arka plan tespitinde de en başarılı model olduğunu göstermektedir. Sonuçlar, hibrit yaklaşımın farklı kayıp bileşenlerini dengeleyerek

hem duyarlılık hem de hassasiyet açısından üstün performans sağladığını ortaya koymaktadır.



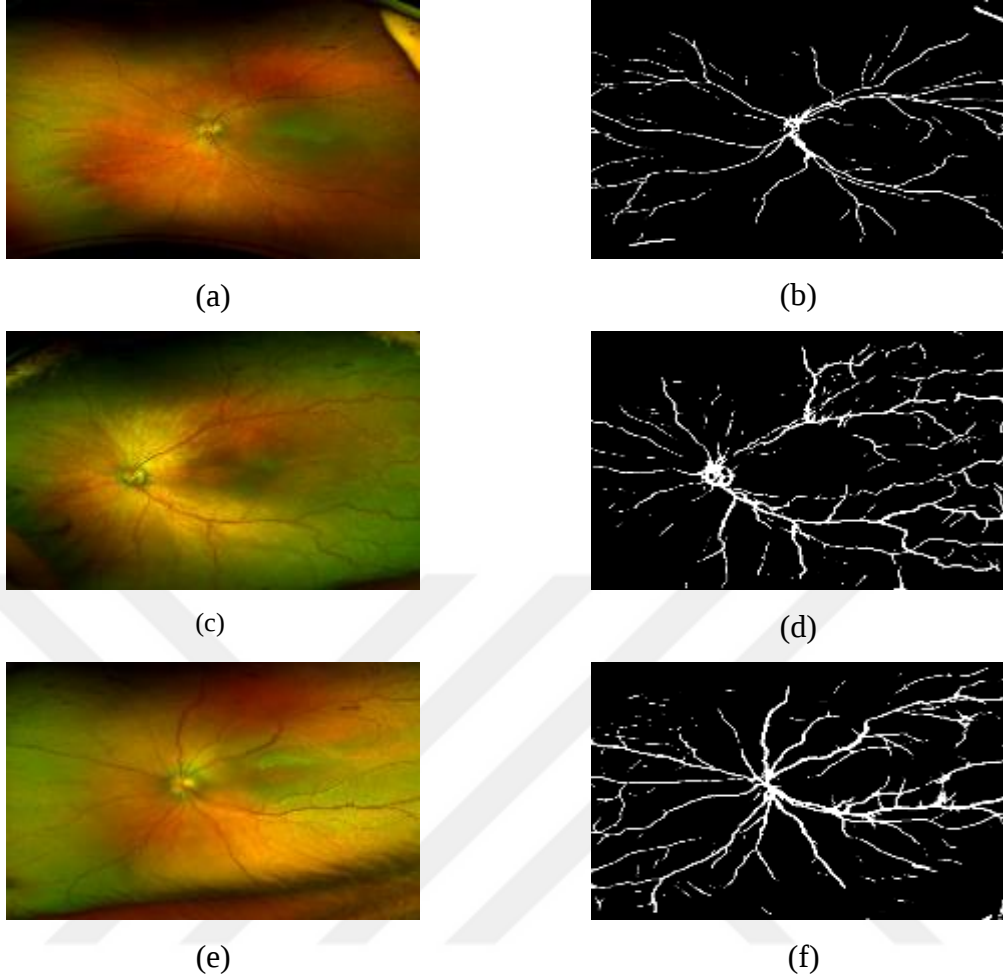
Şekil 18. DRIVE test veri setinden seçilen örnek bir fundus görüntüsünün VesSegNet mimarisi tarafından segmentasyon sonuçlarının görsel analizi

Şekil 18’de soldan sağa doğru sırasıyla; DRIVE test veri setinden alınan orijinal fundus görüntüsü, VesSegNet modeli tarafından üretilen tahmin maskesi, uzman oftalmolog tarafından manuel olarak işaretlenmiş gerçek damar maskesi (Ground Truth) ve son olarak performans haritası gösterilmektedir. Performans haritasında, mavi renkle gösterilen bölgeler modelin doğru tespit ettiği damar piksellerini (True Positive – TP) temsil ederken, kırmızı renkle işaretlenen bölgeler modelin tespit edemediği veya atladığı gerçek damar yapılarını (False Negative – FN) göstermektedir. Bu görsel karşılaştırma, önerilen modelin ince damar yapılarını yakalama başarısını ve hata yapma eğiliminde olduğu bölgeleri (genellikle düşük kontrastlı ve ince çaplı periferik damarlar) açıkça ortaya koymaktadır. Model çıktısının ground truth ile piksel piksel karşılaştırılması ile elde edilen bu performans haritası, segmentasyon kalitesinin piksel düzeyinde değerlendirilmesine olanak tanımaktadır.



Şekil 19. VesSegNet modelinin FAPFUM-ROP veri setindeki segmentasyon sonuçları: (a,c) orijinal fundus görüntüleri, (b,d) tahmin edilen damar maskeleri

Geliştirilen VesSegNet modelinin genelleme kapasitesini değerlendirmek amacıyla, eğitim sürecinde hiç görmediği FAPFUM-ROP veri seti üzerinde test edilmiştir. Şekil 19'da sunulan görsel sonuçlar, modelin yeni ve farklı bir veri setindeki performansını ortaya koymaktadır. Sol sütunda yer alan orijinal fundus görüntüleri, prematüre retinopatisi (ROP) vakalarına ait olup, retinal damar yapılarında belirgin morfolojik değişiklikler içermektedir. Sağ sütunda gösterilen segmentasyon haritaları, VesSegNet modelinin bu damar yapılarını başarılı bir şekilde tespit edebildiğini göstermektedir. Her iki test örneğinde de ana damar gövdeleri, dallanma noktaları ve ince kapiller yapılar yüksek doğrulukla segmente edilmiştir. Özellikle optik disk çevresindeki yoğun damar ağının ve periferik bölgelerdeki ince damar yapılarının net bir şekilde tanımlanması, modelin farklı veri setleri üzerinde genelleme yeteneğinin güçlü olduğunu kanıtlamaktadır. Bu sonuçlar, VesSegNet'in farklı görüntüleme protokolleri ve hasta populasyonlarından elde edilen veriler üzerinde de güvenilir performans sergileyebileceğini göstermektedir.



Şekil 20. Farklı ROP evrelerindeki ve sağlıklı kontrol grubundaki fundus görüntülerinin VesSegNet segmentasyon sonuçları

Şekil 20’de, normal (a-b), Pre Plus ROP (c-d) ve Plus ROP (e-f) evrelerindeki fundus görüntülerinin VesSegNet ile segmentasyon sonuçları sunulmaktadır. ROP hastalarında (c, e) görülen belirgin damar genişlemesi, kıvrımlılık ve düzensiz morfoloji, normal damar yapısından (a) açıkça farklılaşmaktadır. Klinik ortamda çekilen bu görüntülerde yoğun arka plan gürültüsü, düzensiz ışıklandırma ve düşük kontrast gibi kalite sorunları bulunmasına rağmen, VesSegNet hem normal hem de her iki ROP evresinde ana ve periferik damar yapılarını başarıyla tespit etmiştir (b, d, f). Model, DRIVE gibi standart veri setlerinde görülmeyen bu karmaşık damar yapılarına ve görüntü gürültüsüne karşı güçlü bir genelleme yeteneği sergilemektedir. Özellikle fundus görüntülerinde gözlemlenen yoğun gürültü, model performansını olumsuz etkileyebilecek önemli bir faktör olmasına rağmen, VesSegNet mimarisi bu zorlu koşullarda dahi tutarlı segmentasyon sonuçları üretmeyi başarmıştır.

2.2. Segmentasyon Destekli ROP Sınıflandırma Yaklaşımı

2.2.1. FusionROPNet: Önerilen Sınıflandırma Ağı

Sınıflandırma performansını artırmak amacıyla literatürde füzyon (fusion) ve çok dallı (multi-branch) yaklaşımlar sıklıkla kullanılmaktadır. Aynı hastadan elde edilen farklı öznetelik gruplarının birleştirilmesi, daha zengin ve temsil gücü yüksek özelliklerin öğrenilmesini sağlamaktadır. Bu doğrultuda, çalışmanın ikinci aşamasında iki dallı (dual-branch) bir mimari tasarlanmıştır: Görüntü dalı (Image Branch) orijinal RGB fundus görüntüsünden genel özelliklerini (renk, doku, aydınlanma, genel yapı) çıkarır. Maske dalı (Mask Branch) VesSegNet tarafından üretilen tek kanallı damar maskesinden vasküler morfoloji ve damar ağı geometrisine ilişkin özellikleri öğrenir.

Her iki dalda çıkarılan özellikler geç füzyon (late fusion) yöntemi ile birleştirilmekte ve tam bağlantılı katmanlar üzerinden ROP plus hastalığı tahmini gerçekleştirilmektedir. Bu yaklaşım, yalnızca fundus görüntüsüne dayalı modellerin gözden kaçırabileceği ince damar yapısındaki morfolojik değişimlerin ve vasküler anormalliklerin karar sürecine açık bir şekilde dahil edilmesini mümkün kılmaktadır. Dual-Branch mimari yaklaşımı, bazı temel avantajlar sunmaktadır:

- Tamamlayıcı bilgi kaynakları: Görüntü dalı makro düzeyde bilgi sunarken, maske dalı mikro düzeyde vasküler detaylara odaklanır. Bu iki bilgi kaynağının kombinasyonu, tek başına kullanıldıklarında elde edilebilecek performansın ötesinde sonuçlar üretir.
- Özelleşmiş öğrenme: Her dal, kendi girdi tipine özgü en uygun özellikleri öğrenmek için bağımsız olarak optimize edilebilir.
- Klinik ilişkilendirme: ROP plus hastalığı sınıflandırılması damar kıvrımlılığı, damar genişlemesi ve vasküler anormallikler kritik göstergelerdir. Maske dalı, bu özelliklerin doğrudan öğrenilmesine olanak tanır.
- Yorumlanabilirlik: Her dalın katkısı ayrı ayrı analiz edilebildiği için, modelin karar verme sürecinin anlaşılması ve klinik yorumlama kolaylaşır.

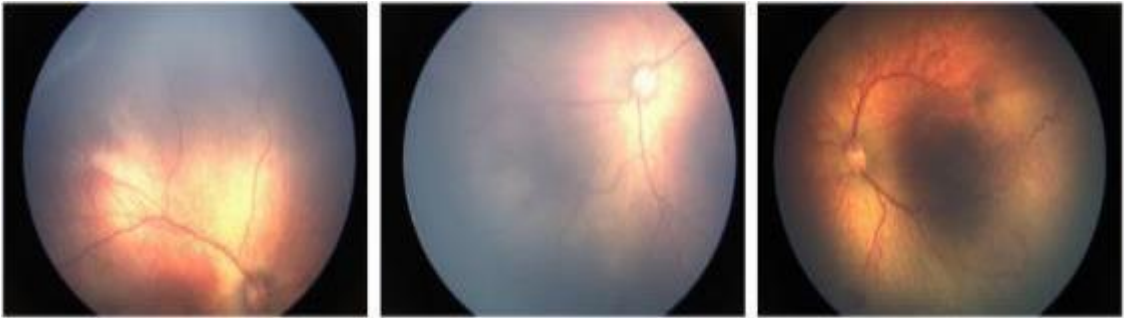
Veri Ön İşleme ve Augmentasyon: FusionRopNet sınıflandırma modeli, çift dallı mimari yapısı gereği her örnek için iki farklı girdi almaktadır: RGB fundus görüntüsü ve birinci aşamada VesSegNet tarafından üretilen damar segmentasyon maskesidir. Etiket

bilgileri excel dosyasından okunarak, hasta klasörü ve görüntü adı eşleştirmesiyle ilgili fundus görüntüsü (.jpg) ve segmentasyon maskesi (.png) çiftleri yüklenmiştir. Sınıf etiketleri No-Plus (1), Pre-Plus (2) ve Plus (3) olarak tanımlanmış, model girişi için 0, 1, 2 indekslerine dönüştürülmüştür. Tüm görüntüler 512×512 piksel boyutuna yeniden ölçeklendirilmiştir.

Eğitim setinin genelleme kapasitesini artırmak için kapsamlı bir veri augmentasyonu stratejisi uygulanmıştır. Geometrik dönüşümler olarak ± 20 derece rastgele rotasyon (RandomRotation), %50 olasılıkla yatay çevirme (RandomHorizontalFlip), %20 olasılıkla dikey çevirme (RandomVerticalFlip) ve %90-110 ölçekleme faktörü aralığında rastgele afin dönüşüm (RandomAffine) kullanılmıştır. Fotometrik augmentasyon olarak ColorJitter ile brightness ve contrast parametreleri ± 20 oranında rastgele değiştirilmiştir. Ayrıca görüntü kalitesi varyasyonlarını simüle etmek için 3×3 kernel boyutunda ve $\sigma=0.1-1.0$ aralığında Gaussian bulanıklaştırma (GaussianBlur) eklenmiştir.

2.2.2. FusionROPNet Eğitimi İçin Kullanılan Veri Seti: FARFUM-ROP

FARFUM-ROP (Farabi ve Ferdowsi Üniversitesi Mashhad ROP veri seti), 68 hastadan toplanan 1533 ROP fundus görüntüsü içermektedir ve beş deneyimli çocuk göz doktoru tarafından normal, pre-plus ve plus olarak bağımsız şekilde etiketlenmiştir. Veri seti 2016 Nisan ile 2019 Mayıs tarihleri arasında İran'daki prematüre bebeklerden toplandı ve doğum ağırlığı 2000 gramdan az ve gebelik haftası 34 haftadan az olan bebekleri içermektedir. FARFUM-ROP, ROP için kamuya açık erişilebilir olan ilk önemli görüntü bankalarından biri olarak sunulmaktadır (Akbari vd., 2024).



Şekil 21. FARFUM-RoP veri setinden alınan retina görüntüleri

Veri setinin yapısı oldukça detaylı ve organize edilmiştir; 68 sıkıştırılmış .rar dosyası, her biri hasta koduna karşılık gelecek şekilde adlandırılmış ve 1200 x 1600 piksel boyutlarında .jpg formatında görüntüler içermektedir. Her hasta için 1 ile 41 arasında değişen sayıda görüntü bulunmakta olup, hasta başına ortalama 23 görüntü mevcuttur. Şekil 21'de görüldüğü üzere, veri seti farklı ROP plus hastalık seviyelerini (no-plus, pre-plus, plus) yansıtan çeşitli damar morfolojilerine sahip fundus görüntüleri içermekte ve bu çeşitlilik modelin farklı hastalık düzeylerini ayırt etme yeteneğini test etmek için önemli bir imkan sunmaktadır.

Veri setiyle birlikte iki ayrı excel dosyası sunulmaktadır. İlk dosya olan datasetdetails.xlsx, 68 satır ve 4 sütundan oluşmakta ve hastaların kimlik kodları (Patient01'den Patient68'e kadar), doğum ağırlıkları (gram cinsinden), gebelik haftası (hafta cinsinden) ve cinsiyetleri gibi demografik bilgileri içermektedir. İkinci dosya olan datasetlabels.xlsx ise 1533 satır ve sekiz sütundan oluşmakta olup, hasta kodları, görüntü dosya isimleri ve etiketleri vermektedir.

2.2.3. ROP Veri Setleri: Mevcut Kaynaklar, Zorluklar ve Kısıtlamalar

FARFUM-ROP veri setine ek olarak mevcut veri setleri: i-ROP veri seti, ABD'den yayınlanmış olup 5,561 görüntü içermektedir, ancak profesyonel topluluk için erişilebilir değildir (Brown vd., 2018).

Çek Cumhuriyeti merkezli veri seti, Ostrava Üniversite Hastanesi'nden toplanan 188 yenidoğana ait 6,004 retinal görüntüden oluşmakta ve Figshare, Kaggle ve GitHub platformlarında CC0 lisansı altında erişime açıktır (Timkovič vd., 2024).

Çin merkezli veri seti, 483 prematüre bebekten toplanan 1,099 fundus görüntüsünden oluşmakta olup, gerçek dünya klinik ortamlarında uygulanabilir yapay zeka sistemlerinin geliştirilmesi için tasarlanmıştır (Zhao vd., 2024).

ROP veri setlerinin kullanımı ve geliştirilmesi sürecinde birçok zorluk ve kısıtlama mevcuttur. Yetişkin diyabetik retinopatisi için geniş kapsamlı ve ücretsiz veri setleri bulunmasına rağmen, bebeklerin retina görüntüleri veya doğrudan ROP veri setleri için durum oldukça farklıdır ve bu alanda ciddi bir veri eksikliği söz konusudur. ROP ile ilgili mevcut görüntü bankalarının çoğu özel (private) statüsündedir ve detaylar yalnızca ilgili yayınlarından elde edilebilmektedir, bu da araştırmacılar arası iş birliğini ve sonuçların karşılaştırılabilirliğini önemli ölçüde sınırlamaktadır. Metodolojik açıdan bakıldığında, veri

setlerinin önemli sınırlamaları arasında ROP Evre V (total retina dekolmanı) gibi ileri evre vakalarının sınırlı sayıda olması, görüntülerin tek bir klinik merkezden toplanması nedeniyle veri çeşitliliğinin az olması ve sistemlerin normal ile çok erken evre ROP vakaları arasında ayırım yapmakta zorlanması yer almaktadır. Özellikle pre-plus ve plus-form tanımının klinik literatürde belirsiz olması ve pre-plus'lu hastaların plus-form kategorisine dahil edilebilmesi, tüm ROP veri setlerinin ortak sınırlamasını oluşturmaktadır; ayrıca pre-plus ve plus-form görülme sıklığının nadir olması da etiketleme sürecini zorlaştırmaktadır.

Klinik uygulamadaki en büyük zorluklardan biri, mevcut yapay zeka sistemlerinin çoğunun tek merkezli veri setleri kullanılarak deneysel araştırma amaçlı tasarlanmış olması ve bu sistemlerin gerçek dünya klinik ortamlarında uygulanabilirlik açısından yetersiz kalmasıdır. Gelecekteki çalışmalar için çoklu klinik merkezlerden ve daha büyük hasta popülasyonlarından alınan veri setlerine ihtiyaç duyulmakta olup, bu tür kapsamlı veri setleri yapay zeka sistemlerinin pratik ve güvenilir klinik araçlar haline gelmesi için kritik önem taşımaktadır. Ek olarak, bebek hastalardan görüntü toplama sürecinde karşılaşılan teknik zorluklar- yetersiz hasta iş birliği, kısa dikkat süreleri, hareket artefaktları, odaklama problemleri ve aydınlatma tutarsızlıkları- yüksek kaliteli veri setlerinin oluşturulmasını güçleştirmekte ve kapsamlı kalite kontrol süreçlerini zorunlu kılmaktadır.

2.2.4. Öznitelik Çıkarıcı Omurga: EfficientNet-B0 Mimarisi

Bu çalışmada sınıflandırma omurgası olarak EfficientNet mimarisi tercih edilmiştir. EfficientNet, bileşik ölçekleme ilkesiyle ağ derinliği, genişliği ve giriş çözünürlüğünü dengeli biçimde büyütürken parametre/verim oranını iyileştirir; Mobile Inverted Bottleneck (MBConv) blokları ve entegre Squeeze and Excitation (SE) mekanizmasıyla temsil gücünü artırır (Tan & Le, 2019). Çalışmamızda iki kollu bir füzyon tasarımı benimsenmiştir: Fundus görüntüler için ImageNet üzerinde ön eğitilmiş EfficientNet B0 omurgası kullanılmış, tek kanallı maske girişi için ise ilk evre konvolüsyonu 1 kanalı destekleyecek şekilde uyarlanmış bir EfficientNet varyantı oluşturulmuştur. Her iki koldan çıkarılan özellikler global ortalama havuzlama ile sıkıştırıldıktan sonra birleştirilmekte, kanal dikkatini modellemek üzere SE bloğu uygulanmakta ve ardından tam bağlantılı bir füzyon katmanı ile sınıflandırıcıya iletilmektedir. Bu transfer öğrenme mimarisi, ön eğitilmiş ağırlıkların genelleme kabiliyetinden yararlanırken, görev özel ince ayar sayesinde problem'e özgü ayrımların daha etkin öğrenilmesini sağlamaktadır. EfficientNet, bileşik ölçekleme ile ağ

derinliği-genişliği-giriş çözünürlüğünü dengeli büyütüp parametre/verim oranını iyileştiren bir mimaridir. Temel yapıtaşı, ters çevrilmiş artık (inverted residual) yapıya sahip MBConv bloklarıdır; her blokta derinlikçe ayrılabilir evrişim ve kanal dikkatini modelleyen Squeeze-and-Excitation (SE) bulunur (Tan ve Le, 2019).

Tablo 4. EfficientNet-B0 mimarisinin katman yapısı ve parametreleri

Katman	Parametreler ve Çıkış Boyutu
İlk Katman (Stem)	3×3 evrişim (stride 2, 32 kanal), batch normalization ve Swish/SiLU aktivasyonu → 112×112
Blok Grubu 1	MBConv1, 3×3, 16 kanal, tekrar=1, stride=1 → 112×112
Blok Grubu 2	MBConv6, 3×3, 24 kanal, tekrar=2, ilk blok stride=2 → 56×56
Blok Grubu 3	MBConv6, 5×5, 40 kanal, tekrar=2, ilk blok stride=2 → 28×28
Blok Grubu 4	MBConv6, 3×3, 80 kanal, tekrar=3, ilk blok stride=2 → 14×14
Blok Grubu 5	MBConv6, 5×5, 112 kanal, tekrar=3, stride=1 → 14×14
Blok Grubu 6	MBConv6, 5×5, 192 kanal, tekrar=4, ilk blok stride=2 → 7×7
Blok Grubu 7	MBConv6, 3×3, 320 kanal, tekrar=1, stride=1 → 7×7
Son Katmanlar (Head)	1×1 konv (1280 kanal), global ortalama havuzlama, dropout ve tam bağlantılı sınıflandırıcı

2.2.5. EfficientNet-B0 Tercih Nedenleri

Seçimi destekleyen başlıca gerekçeler: Parametre ve hesaplama verimliliği; EfficientNet-B0 yaklaşık 5.3M parametre ve 224×224 girişte ~0.39 GFLOPs ile, benzer doğruluk düzeylerini daha düşük hesaplama maliyetiyle sunar; bu da kısıtlı GPU belleği ve zaman bütçelerinde avantaj sağlar. Ters artık yapıları MBConv blokları ve Squeeze-and-Excitation dikkat mekanizması, doku/kenar gibi ince ayrımları güçlü ve verimli biçimde yakalar; medikal görüntüleme için uygundur.

Transfer öğrenme erişimi; ImageNet üzerinde ön-eğitilmiş ağırlıkların yaygın ve güvenilir olması, RGB kolunun hızlı yakınsama ve daha iyi başlangıç temsilleriyle ince ayarlanmasını sağlar. Füzyon mimarisiyle uyum; Mevcut mimaride RGB kolu için ön-eğitilmiş EfficientNet-B0, maske kolu içinse ilk konvolüsyon katmanı 1 kanallı destekleyecek şekilde uyarlanmış EfficientNet kullanılmıştır. Her iki koldan gelen global ortalama havuzlanmış özellikler birleştirilip SE bloğu ile kanal dikkati güçlendirilerek sınıflandırmaya iletilir. Sınırlı veriyle genelleme, Daha az parametre ve ön-eğitilmiş başlangıç, küçük/orta ölçekli medikal veri kümelerinde aşırı öğrenme riskini azaltır; istikrarlı genelleme sunar (Tan ve Le, 2019). Bu nedenlerle EfficientNet-B0, bu çalışmanın iki kollu (RGB + 1-kanal maske) füzyon tasarımında doğruluk-maliyet dengesini optimize eden uygun bir omurga olarak belirlenmiştir.

2.2.6. Transfer Öğrenme

Transfer öğrenme, büyük ölçekli genel amaçlı veri kümeleri (ör. ImageNet; yaklaşık 1.4 milyon görüntü, 1000 sınıf) üzerinde ön-eğitilmiş ağırlıkların yeniden kullanılması ilkesine dayanır. Mimarimizde iki kollu füzyon tasarımının RGB kolu ImageNet üzerinde ön-eğitilmiş EfficientNet-B0 ağırlıklarıyla başlatılmış ve görev için ince ayar (fine-tuning) yapılmıştır; tek kanallı maske kolu ise 1-kanal girişe uyarlama gereği nedeniyle sıfırdan eğitilmiştir. Ön-eğitilmiş sınıflandırıcı başı çıkarılarak (özellik çıkarımına odaklı yapı) her iki koldan gelen küresel ortalama havuzlanmış özellikler birleştirilmiş, kanal dikkatini güçlendirmek üzere SE bloğu uygulanmış ve görev-özel füzyon katmanı üzerinden sınıflandırma gerçekleştirilmiştir. Bu hibrit kurulum, genel görsel temsillerden yararlanırken maske kolunun görev-özel ipuçlarını öğrenmesine olanak tanır.

2.2.7. Özellik Füzyonu ve Sınıflandırma Katmanı

Her iki EfficientNet omurgasından çıkarılan özellik haritaları, küresel ortalama havuzlama sonrası vektörleştirilir ve geç füzyon (late fusion) yaklaşımıyla birleştirilir. EfficientNet-B0 için her bir dalın özellik boyutu 1280 olup, birleştirme sonrası 2560 boyutlu bir vektör elde edilir. Füzyon vektörü üzerinde kanal dikkatini modellemek amacıyla SE bloğu uygulanır; akabinde tek katmanlı bir füzyon başı ile sınıflandırıcıya iletilir. Eğitimde

çapraz-entropi kaybı kullanılır; olasılık raporlaması gerektiğinde çıkış lojitlerine softmax uygulanır.

- Özellik boyutu (B0, her dal): $1280 \rightarrow$ Birleşik vektör: 2560
- Dikkat: SE bloğu (kanal düzeyinde sıkıştır-uyar)
- Füzyon başı: $2560 \rightarrow 1280$, ReLU aktivasyonu, Dropout $p=0.3$
- Sınıflandırıcı: $1280 \rightarrow 3$ (ROP evre sayısı).

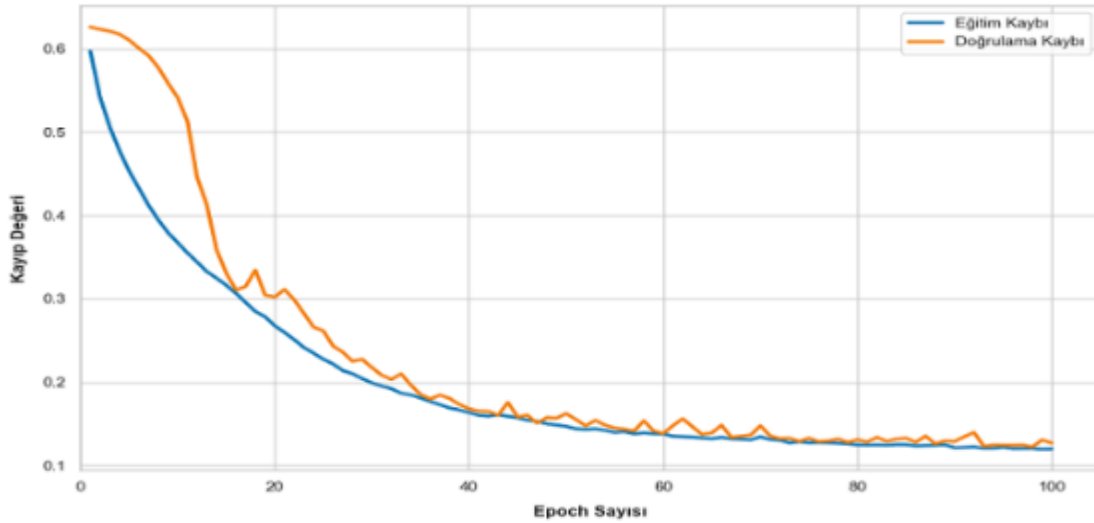
Bu tasarım, RGB dalının global görsel bağlam ve doku örüntülerini, maske dalının ise damar morfolojisine ilişkin yapısal ipuçlarını birlikte kullanarak daha sağlam ve klinik olarak yorumlanabilir tahminler üretmeyi amaçlar.



3. BULGULAR VE TARTIŞMA

Önerilen VesSegNet mimarisinin retina damar segmentasyonu görevindeki performansı, elde edilen nicel metrikler ve nitel gözlemler ışığında detaylı olarak analiz edilmektedir. OneCycleLR scheduler ile learning rate'in dinamik olarak ayarlanması, eğitim sürecinin daha hızlı yakınsamasını sağlamıştır. Hybrid Loss fonksiyonu, BCE, Dice Loss ve Focal Loss'un ağırlıklı kombinasyonu olarak tasarlanmıştır. Bu hibrit yaklaşım, hem bölgesel örtüşmeyi (Dice) hem de piksel bazlı sınıflandırmayı (Focal) optimize ederek dengeli bir öğrenme sağlamıştır. Gamma değerinin 2 olarak seçilmesi, modelin zor örnekler üzerinde daha fazla odaklanmasını sağlamıştır.

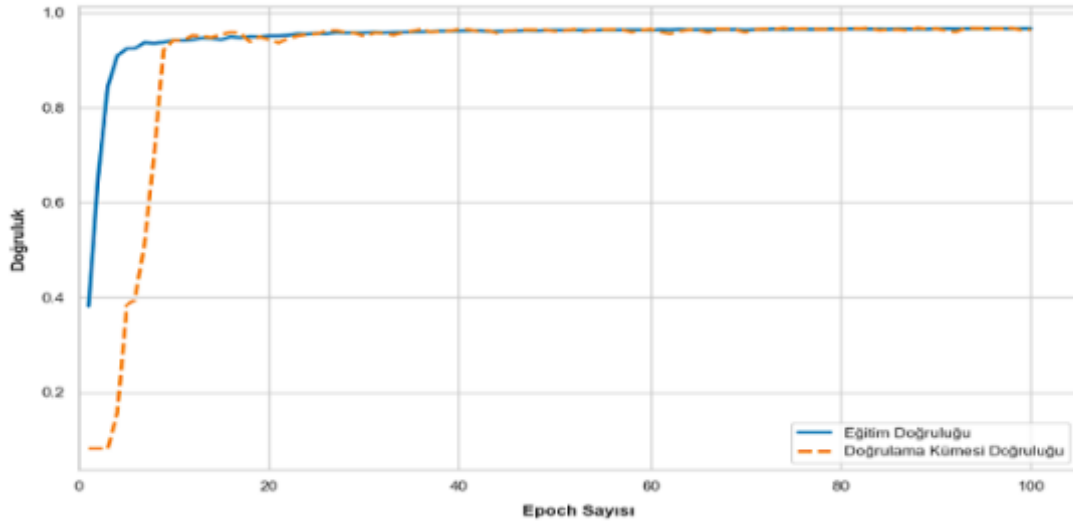
VesSegNet modeli, 100 epoch boyunca eğitilmiştir. Modelin eğitim sürecinde uygulanan Hibrit Kayıp Fonksiyonu (HybridLoss), AdamW optimize edici ve OneCycleLR öğrenme oranı çizelgeleyicisi ile başarılı bir yakınsama göstermiştir.



Şekil 22. VesSegNet modelinin eğitim ve doğrulama kayıp grafiği

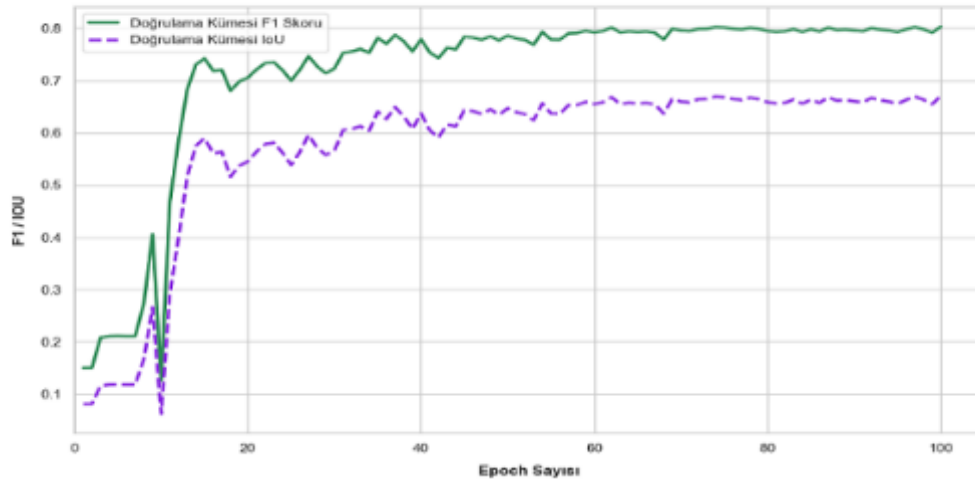
Şekil 22'de, eğitim ve doğrulama kayıp değerlerinin epoch bazındaki değişimi sunulmuştur. Model eğitiminin ilk 20 epoch'unda kayıp değerlerinde hızlı bir düşüş gözlemlenmiş, sonrasında ise stabil bir seyir izlemiştir. Eğitim süreci sonunda train loss değeri 0.12, validation loss değeri ise 0.13 olarak ölçülmüştür. Bu iki değer arasındaki minimal fark (0.01), modelin aşırı öğrenme (overfitting) yapmadığını göstermektedir.

Model, eğitim verisi üzerinde öğrendiği özellikleri doğrulama verisine başarıyla genelleylebilmiştir.



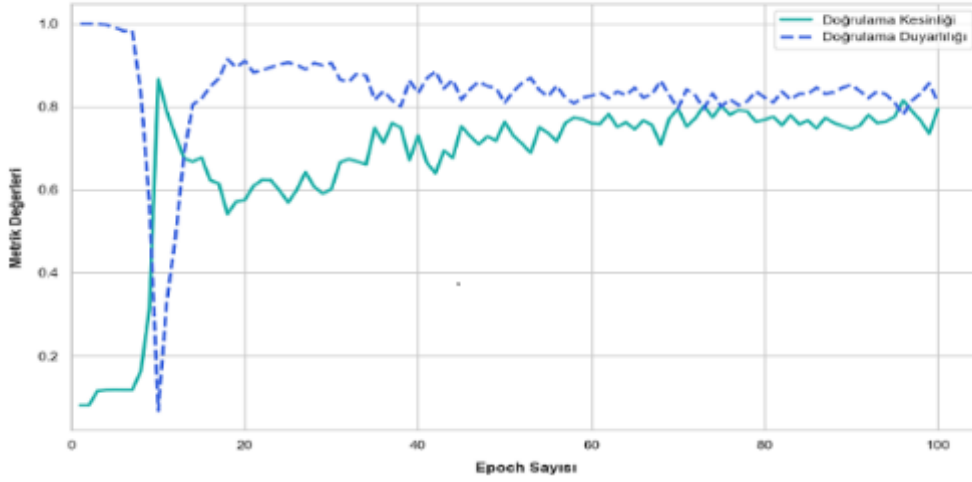
Şekil 23. VesSegNet modelinin eğitim sürecinde doğruluk değerlerinin değişimi

Şekil 23'te doğrulama accuracy değerinin eğitim adımlarına göre değişimi gösterilmektedir. Model başlangıçta ~0.10 accuracy değerinden başlayarak, yaklaşık 100 adım içerisinde 0.97 seviyesine ulaşmış ve bu seviyeyi korumuştur. Bu hızlı yakınsama, modelin veri setindeki damar yapılarını etkin bir şekilde öğrendiğini göstermektedir.



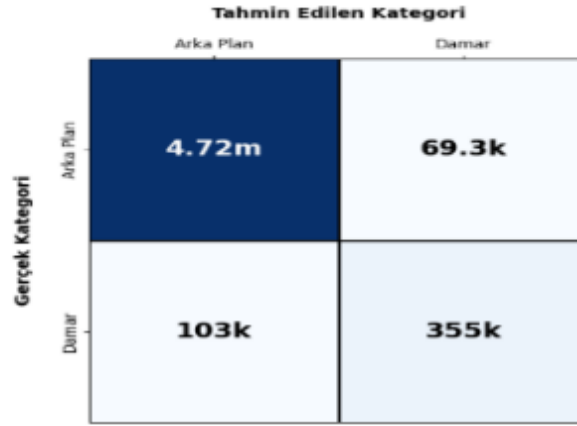
Şekil 24. Modelin doğrulama kümesinde F1 skoru ve IoU değerlerinin eğitim süreci boyunca epochlara göre değişimi

Retina damar segmentasyonu gibi dengesiz sınıf dağılımlarında F1 skoru ve IoU kritik öneme sahiptir. Şekil 24’te test veri setinde elde edilen F1 skoru (0.807), dice katsayısı (0.806) ve IoU (0.675), modelin maske örtüşmesinde güçlü bir performans sergilediğini ve literatürdeki rekabetçi sonuçlarla uyumlu olduğunu göstermektedir. Ek olarak doğruluk (0.9672) ve kayıp (0.1206) değerleri, genel tahmin kalitesinin yüksek ve hata oranının düşük olduğuna işaret eder. Hassasiyet (0.8214) ile duyarlılık (0.7936) değerlerinin birbirine yakın olması, modelin sağlıklı bir denge kurduğunu göstermektedir: yüksek hassasiyet, damar olmayan bölgelerin yanlış pozitif olarak işaretlenmesi riskini azaltarak tanısal güvenilirliği artırırken; yüksek duyarlılık, ince damarların gözden kaçırılma (yanlış negatif) riskini düşürür. Yüksek duyarlılık, ince damar yapılarının veya patolojik damarların gözden kaçırılması (Yanlış Negatif) riskini düşürür. Modelin her iki metrikte de yüksek değerler elde etmesi hem damar tespiti hem de arka plan ayrımı konusunda başarılı olduğunu gösterir.



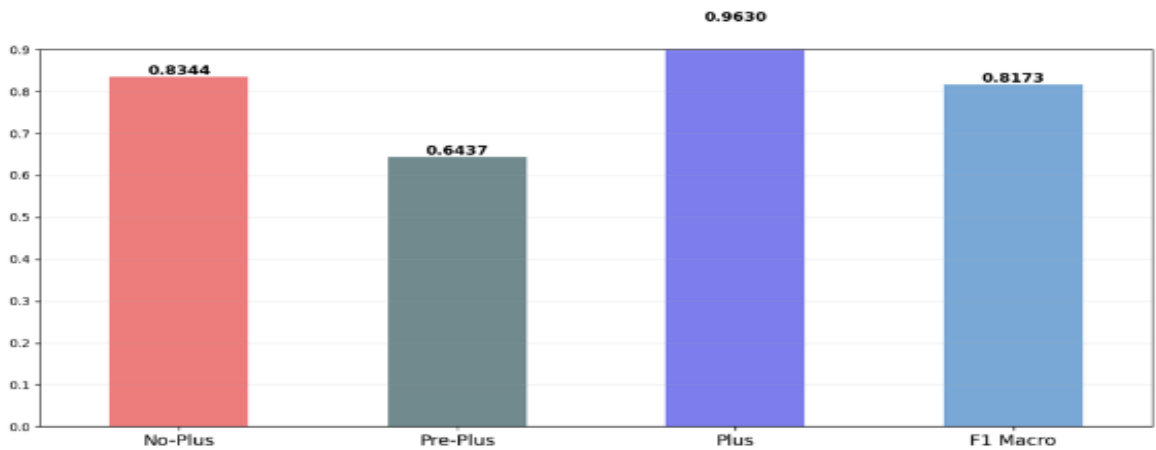
Şekil 25. Modelin doğrulama kümesinde kesinlik (precision) ve duyarlılık (recall) metriklerinin eğitim süreci boyunca epochlara göre değişimi

Şekil 25’te doğrulama metriklerinin epoch boyunca evrimi görülüyor: İlk ~10 epoch’ta sınıf dengesizliğinin etkisiyle duyarlılık çok yüksek seyrederken (model “her yeri damar” demeye eğilimli), kesinlik düşüktür. 15. epoch civarında arka plan daha iyi öğrenildiği için kesinlik hızla toparlanıyor; 20–40 arası her iki metrik 0.78–0.86 bandında dengeye oturuyor. 70. epoch’tan itibaren fark daralıyor ve son bölümde kesinlik ≈ 0.80 , duyarlılık ≈ 0.83 –0.85 seviyelerinde istikrar kazanıyor. Bu çıktılar, test sonuçlarındaki $F1 \approx 0.807$ ve $Dice = 0.806$ ile tutarlı olup aşırı öğrenme belirtisi göstermemektedir; gözlenen küçük dalgalanmalar veri/augmentasyon kaynaklı normal varyasyondur.



Şekil 26. Test veri setinde elde edilen karmaşıklık matrisi

VesSegNet modelinin performansı, 20 adet 512×512 boyutundaki test görüntüsü üzerinde değerlendirilmiştir. Toplamda 5,238,100 piksel üzerinde yapılan sınıflandırma sonucunda model %96.70 genel doğruluk (accuracy) oranına ulaşmıştır. Şekil 26'daki karmaşıklık matrisi sonuçlarına göre, 4.72 milyon piksel doğru şekilde arka plan olarak (doğru negatif), 355 bin piksel ise doğru şekilde damar yapısı olarak (doğru pozitif) tespit edilmiştir. Model, arka plan piksellerini ayırt etmede oldukça başarılı olup %98.31 özgüllük (specificity) oranı elde etmiştir. Damar segmentasyonu için hesaplanan kesinlik (precision) değeri %81.91, duyarlılık (recall) değeri ise %79.78'dir. Bu sonuçlar, modelin özellikle ince damar yapılarının tespitinde bazı zorluklar yaşasa da genel olarak yüksek performans sergilediğini göstermektedir.



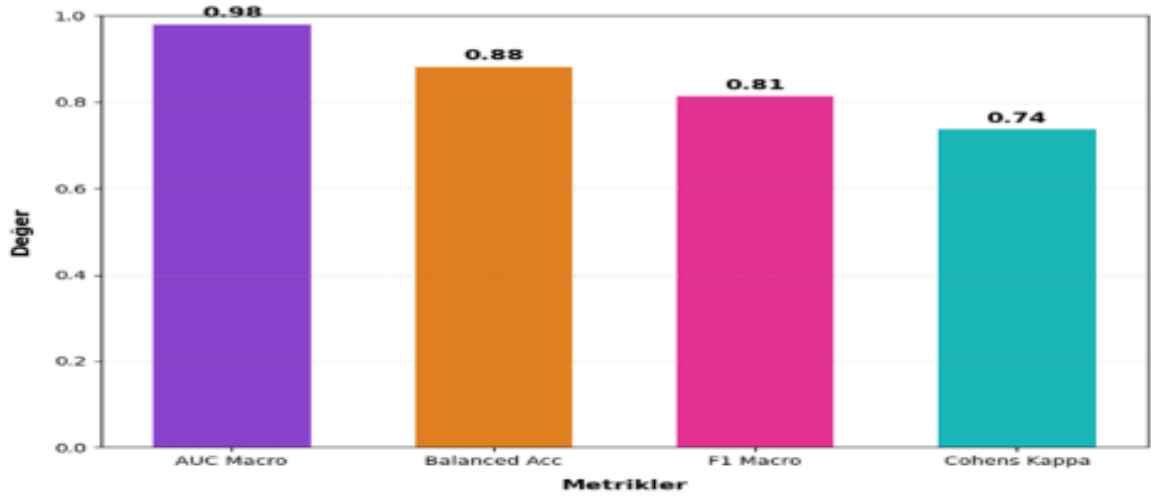
Şekil 27. FusionROPNet'in doğrulama kümesinde sınıf bazlı F1 skorları (no-plus, pre-plus, plus) ve makro F1 değerinin karşılaştırılması

FusionROPNet, doğrulama kümesinde sınıflar arası dengeli bir performans göstererek genel makro F1 değerinde 0.8173'e ulaşmıştır (Şekil 27). Sınıf bazında bakıldığında, klinik açıdan en kritik sınıf olan plus için performansın en yüksek olduğu görülmekte (plus: 0.9630), no-plus sınıfı için sonuçlar tatmin edici seviyededir (no-plus: 0.8344). En zorlayıcı sınıf pre-plus (0.6437) olup, bu düşüş sınıfın ara form/ara şiddet profili ve sınıflar arası görsel örtüşme nedeniyle beklenen bir zorluktur. Buna rağmen yüksek plus duyarlılığı, ileri evre vakaların atlanma riskini azaltırken; no-plus başarımı yanlış alarm oranını sınırlı tutmaktadır. Genel olarak, elde edilen makro F1 ve sınıf-bazlı skorlar, FusionROPNet'in ROP evrelemesinde klinik açıdan anlamlı ve rekabetçi bir doğruluk düzeyi sunduğunu göstermektedir.

		Tahmin Edilen Kategori		
		No-Plus	Pre-Plus	Plus
Gerçek Kategori	No-Plus	0.72 (68)	0.28 (27)	0.00 (0)
	Pre-Plus	0.00 (0)	1.00 (28)	0.00 (0)
	Plus	0.00 (0)	0.07 (4)	0.93 (52)

Şekil 28. FusionROPNet için normalize edilmiş karışıklık matrisi

Şekil 28'deki normalize karışıklık matrisi, modelin sınıflar arası ayırım gücünü net biçimde ortaya koymaktadır: No-plus örneklerinin %72'si doğru sınıflanırken %28'i pre-plus'a kaymıştır; pre-plus sınıfı %100 doğrulukla tanınmıştır ancak diğer sınıflardan bu sınıfa hatalı atamalar olduğu için özgüllüğü görece düşebilir (no plus→pre plus %28, plus→pre plus %7); plus örneklerinin %93'ü doğru sınıflanmış ve yalnızca %7'si pre plus'a karışmıştır; no-plus ile plus arasında doğrudan karışma olmaması, uç sınıfların iyi ayrıldığını gösterirken, belirsizliğin esasen pre-plus ile komşu sınıflar arasında yoğunlaştığını ve karar eşiği/özellik ayrıştırmasının bu bölgede daha kritik olduğunu düşündürmektedir.



Şekil 29. FusionROPNet'in modelinin test verisi üzerindeki performans metrikleri

Şekil 29'da FusionROPNet modelinin test verisi üzerindeki genel başarımını bütüncül biçimde ortaya koymaktadır. Makro AUC'nin 0.98 düzeyinde olması, modelin tüm sınıflar için yüksek ayırt ediciliğe sahip olduğunu göstermektedir. Sınıf dengesizliğinden daha az etkilenen dengeli doğruluk (balanced accuracy) 0.88 ile adil bir doğruluk seviyesine işaret ederken, Cohen'in kappa katsayısının 0.7460 civarında olması tesadüfi uyumdan arındırılmış güçlü bir sınıflandırıcı-gerçeklik anlaşmasını göstermektedir. Eşik bağımlı bir ölçüt olan makro F1'in 0.8137 düzeyinde olması ise duyarlılık ve kesinlik arasında dengeli bir performans elde edildiğini göstermektedir. Birlikte değerlendirildiğinde bu metrikler, modelin klinik olarak anlamlı eşiklerde de rekabetçi ve dengeli bir performans sergilediğini göstermektedir.

Tablo 5. Tez kapsamında önerilen FusionROPNet modelinin literatürdeki güncel ROP tespit sistemleriyle performans karşılaştırması

Çalışma / Sistem (Yıl)	ROP Sistemi (2015) Ataer-Cansizoglu vd.	i-ROP DL (2018) Brown ve ark.	DeepROP (2018) Wang vd.	RegNetY00 2 Modeli (2024) Kıran Yenice vd.	McROP DPFRr-CLAHE (2024) Rahim vd.	FusionRO PNet (2025) Murat vd.
Veri Seti (Boyut)	77 geniş açılı görüntü	~6.000 etiketli görüntü	20.795 görüntü, 3722 vaka, 1273 hasta	189 fundus görüntüsü (A-ROP, Tip 1, Tip 2)	Calgary ROP 1778 RetCam görüntüsü	FARFUM-ROP (1533 görüntü)
Görev / Sınıflandırma Seviyesi	Plus Hastalığı (Normal / Pre-plus / Plus)	Plus Hastalığı Tespiti (Normal - Plus)	Id-Net — ROP Tespiti (Normal vs ROP)	Tip 2 ROP Tespiti	Evre 0–3 Sınıflandırması	Plus Hastalığı (Normal / Pre-plus / Plus)
Accuracy	0.95	0.91	0.9555	0.98	0.9224	N/A
Sensitivity	Plus: 0.93	0.93	0.8491	N/A	0.7241	N/A
Specificity	N/A	0.94	0.9690	0.98	0.9621	0.9402
Precision	N/A	N/A	N/A	N/A	0.7925	0.8249
F1 Skoru	N/A	N/A	N/A	N/A	0.7567	0.8136 Plus:0.9629
Cohen's Kappa	N/A	N/A	N/A	N/A	0.7107	0.74

Tablo 5'te önerilen FusionROPNet modelinin performansı, literatürdeki güncel ROP tespit yöntemleriyle karşılaştırılmıştır. Model, 0.8136 F1 skoru ve Plus sınıfı için 0.9629 F1

skoru ile dengeli bir sınıflandırma performansı sergilemiştir. Özellikle 0.9402 specificity ve 0.8249 precision değerleri, sistemin yanlış pozitif oranını minimize ederek klinik güvenilirliğini artırdığını göstermektedir. 0.74 Cohen's Kappa skoru, modelin rastgele tahminlerin üzerinde tutarlı bir sınıflandırma yeteneğine sahip olduğunu ortaya koymaktadır. Bu sonuçlar, FusionROPNet'in erken ROP teşhisi için klinik ortamlarda kullanılabilecek potansiyele sahip olduğunu desteklemektedir.

Tablo 6. Önerilen iki aşamalı model ile doğrudan RGB tabanlı tek aşamalı modelin ROP sınıflandırma performansları

Performans Metrikleri	Tek Aşamalı Yaklaşım (Doğrudan RGB Sınıflandırma)	Önerilen İki Aşamalı Yaklaşım (VesSegNet + FusionROPNet)
F1-Score (Macro)	0.7843	0.8137
Cohen's Kappa	0.6614	0.7360
Matthews Korelasyon Katsayısı	0.6725	0.7682
Precision (Macro)	0.8166	0.8249
Recall (Macro)	0.7831	0.8689
Specificity	0.8860	0.9402
AUC-ROC	0.9686	0.9895

Tablo 6, önerilen iki aşamalı derin öğrenme çerçevesinin, geleneksel tek aşamalı sınıflandırma yaklaşımına göre performans üstünlüğünü göstermektedir. İki aşamalı yaklaşımda, VesSegNet ile retinal damar yapıları segmente edilmiş ve bu segmentasyon maskeleri FusionROPNet sınıflandırma modelinin girdi özellikleri olarak kullanılmıştır. Tek aşamalı yaklaşımda ise fundus görüntüler doğrudan sınıflandırma ağına verilmiştir. Test veri seti üzerinde yapılan değerlendirmeler, damar segmentasyonu tabanlı iki aşamalı stratejinin tüm performans metriklerinde daha yüksek değerler elde ettiğini ortaya koymuştur. Bu sonuçlar, retinal damar morfolojisinin ROP tanısı için kritik öneme sahip olduğunu ve segmentasyon aşamasının sınıflandırma performansını anlamlı ölçüde artırdığını deneysel olarak doğrulamaktadır.

4. SONUÇLAR

Bu çalışmada retina damar segmentasyonu için önerilen VesSegNet mimarisi, Attention Block'lar ve Channel/Spatial Attention birimlerine ek olarak sınıf dengesizliğini gözetken Hybrid Loss (BCE+Dice+Focal) ile eğitilerek test kümesinde yüksek örtüşme başarımı sergilemiştir: F1 skoru=0.807, dice=0.806, IoU=0.675, doğruluk=0.9672, duyarlılık=0.7936, kesinlik=0.8214 ve kayıp=0.1206. Dengesiz sınıf yapısında doğruluk tek başına yanıltıcı olabildiğinden, F1/Dice/IoU'nun bu düzeyi modelin gerçekten damar piksellerine odaklandığını ve arka planı etkili biçimde bastırdığını göstermektedir. Nitel gözlemler, özellikle ince damarların sürekliliğinde ve sınır keskinliğinde belirgin bir iyileşmeye işaret ederken, kalan hataların çoğunlukla optik disk çevresi ve yüksek parlaklık/kontrast bölgelerinde kümелendiğini ortaya koymaktadır. Kesinlik ile duyarlılığın birbirine yakın ve yüksek olması, yanlış pozitifleri sınırlarken ince damarların gözden kaçırılma riskini de düşürdüğünü doğrulamaktadır.

ROP sınıflandırmasında FusionROPNet, doğrulama/test değerlendirmelerinde makro F1=0.8173 ile klinik açıdan anlamlı ve dengeli bir performans sunmuştur. Sınıf bazında en yüksek skor plus (F1=0.9630) iken, no-plus (F1=0.8344) tatmin edici düzeydedir; pre plus (F1=0.6437) sınıfı, sınıf belirsizliği nedeniyle beklendiği gibi daha zorlu kalmıştır. Karışıklık matrisi belirsizliğin ağırlıkla pre-plus etrafında yoğunlaştığını göstermektedir. Klinik kullanım açısından, karar eşiğinin duyarlılığı önceliklendirecek biçimde kalibre edilmesi (özellikle plus vakalarını kaçırmamak için) ve gerektiğinde kesinliği artıracak ikincil bir inceleme/uzman onayı akışı ile desteklenmesi, hata profiline uygun bir strateji olarak öne çıkmaktadır. Genel olarak, segmentasyon ve sınıflandırma modüllerinden oluşan önerilen çerçeve hem damar maskesi üretiminde hem de ROP evrelemede rekabetçi ve klinik olarak faydalı bir doğruluk düzeyi sunmaktadır. Bulgular, mimarinin dikkat mekanizmaları sayesinde açıklanabilirlik ve odaklanma açısından güçlü sinyaller verdiğini göstermektedir; bununla birlikte çok merkezli ve daha geniş veri kümeleriyle dış doğrulama, pre-plus sınıfında sınıf dengesizliğini azaltmaya yönelik yöntemler (ör. odak kaybı/örnekleme) ve karar eşiği/olasılık kalibrasyonu ile performansın daha da güçlendirilmesi beklenmektedir.

İki aşamalı yaklaşımın üstünlüğü, tek aşamalı RGB tabanlı referans model ile kıyaslandığında net bir şekilde ortaya çıkmaktadır: VesSegNet segmentasyonu kullanan

sistem, F1-macro skorunu 0.7843'ten 0.8137'ye (%3.75 artış) taşırken; modelin karar tutarlılığını gösteren Cohen's Kappa skorunda 0.6614'ten 0.7360'a (~%11.3 artış) ve dengesiz veri setlerinde modelin kalitesini en net yansıtan Matthews Korelasyon Katsayısı (MCC) değerinde 0.6725'ten 0.7682'ye (%14.2 artış) çarpıcı bir iyileşme sağlamıştır. Kappa ve MCC metriklerindeki bu belirgin yükseliş, segmentasyon desteğinin modelin tahminlerini şans faktöründen arındırdığını ve klinik açıdan daha güvenilir kararlar üretmesini sağladığını kanıtlamaktadır. Bu bulgular, damar segmentasyonunun ROP sınıflandırması için kritik klinik özelliklerin etkin öğrenilmesini sağladığını ve sınırlı veri koşullarında genelleme kapasitesini artırdığını deneysel olarak kanıtlamaktadır.



5. ÖNERİLER

Elde edilen sonuçlar klinik açıdan anlamlı olsa da sistemin gerçek dünya koşullarında yaygın kullanıma hazır hale gelmesi için ek çalışmalara ihtiyaç vardır: Öncelikle farklı hastaneler ve cihazlardan toplanmış daha büyük, çok merkezli veri setleriyle dış doğrulama yapılarak genellenebilirlik test edilmeli; etnik köken, gebelik süresi ve görüntü kalitesi gibi alt-gruplar üzerinde performans ayrı ayrı raporlanmalıdır. Pre-plus sınıfındaki görece düşük performansı iyileştirmek için bu sınıfa ait örneklerin artırılması, veri dengeleme stratejilerinin kullanılması ve sınıf sınırındaki belirsizliği azaltmak amacıyla uzmanlar arası uyumu artıran etiketleme protokollerinin (ör. uzlaşmalı etiket, hakem değerlendirmesi) devreye alınması önerilir.

Klinik kullanımda modelin yalnızca tahmin değil, aynı zamanda belirsizlik de raporlaması; düşük güven skorlu vakaların otomatik olarak ikinci bir uzman incelemesine yönlendirilmesi güvenilirliği artıracaktır; buna eşlik eden attention/görselleştirme haritalarının klinisyenlerle birlikte yorumlanması açıklanabilirliği güçlendirir. Dağıtıma yönelik olarak, daha hafif mimariler ve basit topluluk yaklaşımlarıyla hız ve tutarlılık dengesi iyileştirilebilir; karar eşiği uygulamanın önceliklerine göre kalibre edilmeli ve sahada kullanım sırasında performans izleme, veri/dağılım kayması tespiti ve geri bildirimle sürekli öğrenme döngüsü sürdürülmelidir.

Dağıtıma yönelik olarak, daha hafif mimariler ve basit topluluk yaklaşımlarıyla hız ve tutarlılık dengesi iyileştirilebilir; karar eşiği uygulamanın önceliklerine göre kalibre edilmeli ve sahada kullanım sırasında performans izleme, veri/dağılım kayması tespiti ve geri bildirimle sürekli öğrenme döngüsü sürdürülmelidir. Uzun vadede, teleoftalmoloji platformlarına entegrasyon ve taşınabilir fundus cihazlarıyla uyumluluk, ROP taramasının yaygınlaştırılmasına katkı sağlayacaktır.

6. KAYNAKLAR

- Abràmoff, M. D., Garvin, M. K., & Sonka, M. (2010). Retinal imaging and image analysis. *IEEE Reviews in Biomedical Engineering*, 3, 169–208. <https://doi.org/10.1109/RBME.2010.2084567>
- Akbari, M., Pourreza, H.-R., Khalili Pour, E., Dastjani Farahani, A., Bazvand, F., Ebrahimiadib, N., Imani Fooladi, M., & Ramazani K., F. (2024). FARFUM-RoP, A dataset for computer-aided detection of Retinopathy of Prematurity. *Scientific Data*, 11(1), 1176. <https://doi.org/10.1038/s41597-024-03897-7>
- Alom, M. Z., Hasan, M., Yakopcic, C., Taha, T. M., & Asari, V. K. (2018). *Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation* (No. arXiv:1802.06955). arXiv. <https://doi.org/10.48550/arXiv.1802.06955>
- Ataer-Cansizoglu, E., Bolon-Canedo, V., Campbell, J. P., Bozkurt, A., Erdogmus, D., Kalpathy-Cramer, J., Patel, S., Jonas, K., Chan, R. V. P., Ostmo, S., & Chiang, M. F. (2015). Computer-Based Image Analysis for Plus Disease Diagnosis in Retinopathy of Prematurity: Performance of the “i-ROP” System and Image Features Associated With Expert Diagnosis. *Translational Vision Science & Technology*, 4(6), 5. <https://doi.org/10.1167/tvst.4.6.5>
- Blencowe, H., Lawn, J. E., Vazquez, T., Fielder, A., & Gilbert, C. (2013). Preterm-associated visual impairment and estimates of retinopathy of prematurity at regional and global levels for 2010. *Pediatric Research*, 74 Suppl 1(Suppl 1), 35–49. <https://doi.org/10.1038/pr.2013.205>
- Brown, J. M., Campbell, J. P., Beers, A., Chang, K., Ostmo, S., Chan, R. V. P., Dy, J., Erdogmus, D., Ioannidis, S., Kalpathy-Cramer, J., Chiang, M. F., & Imaging and Informatics in Retinopathy of Prematurity (i-ROP) Research Consortium. (2018). Automated Diagnosis of Plus Disease in Retinopathy of Prematurity Using Deep Convolutional Neural Networks. *JAMA Ophthalmology*, 136(7), 803–810. <https://doi.org/10.1001/jamaophthalmol.2018.1934>
- Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M., & Goldbaum, M. (1989). Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE Transactions on Medical Imaging*, 8(3), 263–269. <https://doi.org/10.1109/42.34715>
- Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). *A Simple Framework for Contrastive Learning of Visual Representations* (No. arXiv:2002.05709). arXiv. <https://doi.org/10.48550/arXiv.2002.05709>
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). *3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation* (No. arXiv:1606.06650). arXiv. <https://doi.org/10.48550/arXiv.1606.06650>

- Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J., & Zisserman, A. (2010). The Pascal Visual Object Classes (VOC) Challenge. *International Journal of Computer Vision*, 88(2), 303–338. <https://doi.org/10.1007/s11263-009-0275-4>
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- Fierson, W. M. (2018), American Academy of Pediatrics Section on Ophthalmology, American Academy of Ophthalmology, American Association for Pediatric Ophthalmology and Strabismus, & American Association of Certified Orthoptists. Screening Examination of Premature Infants for Retinopathy of Prematurity. *Pediatrics*, 142(6), e20183061. <https://doi.org/10.1542/peds.2018-3061>
- Fraz, M. M., Remagnino, P., Hoppe, A., Uyyanonvara, B., Rudnicka, A. R., Owen, C. G., & Barman, S. A. (2012). Blood vessel segmentation methodologies in retinal images—A survey. *Computer Methods and Programs in Biomedicine*, 108(1), 407–433. <https://doi.org/10.1016/j.cmpb.2012.03.009>
- Ganin, Y., & Lempitsky, V. (2015). *Unsupervised Domain Adaptation by Backpropagation* (No. arXiv:1409.7495). arXiv. <https://doi.org/10.48550/arXiv.1409.7495>
- Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J. (2017). *A Review on Deep Learning Techniques Applied to Semantic Segmentation* (No. arXiv:1704.06857). arXiv. <https://doi.org/10.48550/arXiv.1704.06857>
- Gilbert, C. (2008). Retinopathy of prematurity: A global perspective of the epidemics, population of babies at risk and implications for control. *Early Human Development*, 84(2), 77–82. <https://doi.org/10.1016/j.earlhumdev.2007.11.009>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Gu, Z., Cheng, J., Fu, H., Zhou, K., Hao, H., Zhao, Y., Zhang, T., Gao, S., & Liu, J. (2019). CE-Net: Context Encoder Network for 2D Medical Image Segmentation. *IEEE Transactions on Medical Imaging*, 38(10), 2281–2292. <https://doi.org/10.1109/TMI.2019.2903562>
- Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J. L., & Webster, D. R. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*, 316(22), 2402–2410. <https://doi.org/10.1001/jama.2016.17216>
- Guo, C., Szemenyei, M., Yi, Y., Wang, W., Chen, B., & Fan, C. (2020). *SA-UNet: Spatial Attention U-Net for Retinal Vessel Segmentation* (No. arXiv:2004.03696). arXiv. <https://doi.org/10.48550/arXiv.2004.03696>
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition* (No. arXiv:1512.03385). arXiv. <https://doi.org/10.48550/arXiv.1512.03385>
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H. ... (2017). *MobileNets: Efficient Convolutional Neural Networks for*

- Mobile Vision Applications* (No. arXiv:1704.04861). arXiv. <https://doi.org/10.48550/arXiv.1704.04861>
- Huang, Y.-P., Basanta, H., Yu-Chuan Kang, E., Chen, K.-J., Hwang, Y.-S., Lai, C.-C., Campbell, J. P., Chiang, M. F., Paul Chan, R. V., Kusaka, S., Fukushima, Y., & Wu, W.-C. (2021). Automated detection of early-stage ROP using a deep convolutional neural network. *The British Journal of Ophthalmology*, 105(8), 1099–1103. <https://doi.org/10.1136/bjophthalmol-2020-316526>
- International Committee for the Classification of Retinopathy of Prematurity. (2005). The International Classification of Retinopathy of Prematurity revisited. *Archives of Ophthalmology (Chicago, Ill.: 1960)*, 123(7), 991–999. <https://doi.org/10.1001/archophth.123.7.991>
- Ioffe, S., & Szegedy, C. (2015). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift* (No. arXiv:1502.03167). arXiv. <https://doi.org/10.48550/arXiv.1502.03167>
- Jalali, Y., Fateh, M., & Rezvani, M. (2024). VGA-Net: Vessel graph based attentional U-Net for retinal vessel segmentation. *IET Image Process.*, 18, 2191–2213. <https://doi.org/10.1049/ipr2.13102>
- Kamran, S. A., Hossain, K. F., Tavakkoli, A., Zuckerbrod, S. L., Sanders, K. M., & Baker, S. A. (2021). RV-GAN: Segmenting Retinal Vascular Structure in Fundus Photographs using a Novel Multi-scale Generative Adversarial Network (Vol. 12908, pp. 34–44). https://doi.org/10.1007/978-3-030-87237-3_4
- Khan, T. M., Robles-Kelly, A., & Naqvi, S. S. (2022). T-Net: A Resource-Constrained Tiny Convolutional Neural Network for Medical Image Segmentation. 644–653. https://openaccess.thecvf.com/content/WACV2022/html/Khan_T-Net_A_Resource-Constrained_Tiny_Convolutional_Neural_Network_for_Medical_Image_WACV_2022_paper.html
- Kıran Yenice, E., Kara, C., & Erdaş, Ç. B. (2024). Automated detection of type 1 ROP, type 2 ROP and A-ROP based on deep learning. *Eye*, 38(13), 2644–2648. <https://doi.org/10.1038/s41433-024-03184-0>
- Laibacher, T., Weyde, T., & Jalali, S. (2019). M2U-Net: Effective and Efficient Retinal Vessel Segmentation for Resource-Constrained Environments (No. arXiv:1811.07738). arXiv. <https://doi.org/10.48550/arXiv.1811.07738>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014* (pp. 740–755). Springer International Publishing. https://doi.org/10.1007/978-3-319-10602-1_48

- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- Liu, I., & Sun, Y. (1993). Recursive tracking of vascular networks in angiograms based on the detection-deletion scheme. *IEEE Transactions on Medical Imaging*, 12(2), 334–341. <https://doi.org/10.1109/42.232264>
- Long, J., Shelhamer, E., & Darrell, T. (2015). *Fully Convolutional Networks for Semantic Segmentation* (No. arXiv:1411.4038). arXiv. <https://doi.org/10.48550/arXiv.1411.4038>
- Lucas, W. J., Groover, A., Lichtenberger, R., Furuta, K., Yadav, S.-R., Helariutta, Y., He, X.-Q., Fukuda, H., Kang, J., Brady, S. M., Patrick, J. W., Sperry, J., Yoshida, A., López-Millán, A.-F., Grusak, M. A., & Kachroo, P. (2013). The plant vascular system: Evolution, development and functions. *Journal of Integrative Plant Biology*, 55(4), 294–388. <https://doi.org/10.1111/jipb.12041>
- Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2020). *Image Segmentation Using Deep Learning: A Survey* (No. arXiv:2001.05566). arXiv. <https://doi.org/10.48550/arXiv.2001.05566>
- Mulay, S., Ram, K., Sivaprakasam, M., & Vinekar, A. (2019). Early Detection of Retinopathy of Prematurity stage using Deep Learning approach. *Medical Imaging 2019: Computer-Aided Diagnosis*, 107. <https://doi.org/10.1117/12.2512719>
- Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359. <https://doi.org/10.1109/TKDE.2009.191>
- Powers, D. M. W. (2020). *Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation* (No. arXiv:2010.16061). arXiv. <https://doi.org/10.48550/arXiv.2010.16061>
- Raghu, M., Zhang, C., Kleinberg, J., & Bengio, S. (2019). *Transfusion: Understanding Transfer Learning for Medical Imaging* (No. arXiv:1902.07208). arXiv. <https://doi.org/10.48550/arXiv.1902.07208>
- Rahim, S., Sabri, K., Ells, A., Wassyn, A., Lawford, M., Chu, L., & He, W. (2024). *Novel Fundus Image Preprocessing for Retcam Images to Improve Deep Learning Classification of Retinopathy of Prematurity* (No. arXiv:2302.02524). arXiv. <https://doi.org/10.48550/arXiv.2302.02524>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation* (No. arXiv:1505.04597). arXiv. <https://doi.org/10.48550/arXiv.1505.04597>
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>

- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, 15(1), 1929–1958.
- Staal, J., Abramoff, M. D., Niemeijer, M., Viergever, M. A., & van Ginneken, B. (2004). Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4), 501–509. <https://doi.org/10.1109/TMI.2004.825627>
- Tan, M., & Le, Q. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *Proceedings of the 36th International Conference on Machine Learning*, 6105–6114. <https://proceedings.mlr.press/v97/tan19a.html>
- Timkovič, J., Nowaková, J., Kubiček, J., Hasal, M., Varyšová, A., Kolarčík, L., Maršolková, K., Augustynek, M., & Snášel, V. (2024). Retinal Image Dataset of Infants and Retinopathy of Prematurity. *Scientific Data*, 11, 814. <https://doi.org/10.1038/s41597-024-03409-7>
- Wang, J., Ji, J., Zhang, M., Lin, J.-W., Zhang, G., Gong, W., Cen, L.-P., Lu, Y., Huang, X., Huang, D., Li, T., Ng, T. K., & Pang, C. P. (2021). Automated Explainable Multidimensional Deep Learning Platform of Retinal Images for Retinopathy of Prematurity Screening. *JAMA Network Open*, 4(5), e218758. <https://doi.org/10.1001/jamanetworkopen.2021.8758>
- Wang, J., Ju, R., Chen, Y., Zhang, L., Hu, J., Wu, Y., Dong, W., Zhong, J., & Yi, Z. (2018). Automated retinopathy of prematurity screening using deep neural networks. *EBioMedicine*, 35, 361–368. <https://doi.org/10.1016/j.ebiom.2018.08.033>
- Wei, X., Lin, X., Liu, H., Zhao, S., & Li, Y. (2024). *Retinal Vessel Segmentation with Deep Graph and Capsule Reasoning* (No. arXiv:2409.11508). arXiv. <https://doi.org/10.48550/arXiv.2409.11508>
- Worrall, D. E., Wilson, C. M., & Brostow, G. J. (2016). Automated Retinopathy of Prematurity Case Detection with Convolutional Neural Networks. In G. Carneiro, D. Mateus, L. Peter, A. Bradley, J. M. R. S. Tavares, V. Belagiannis, J. P. Papa, J. C. Nascimento, M. Loog, Z. Lu, J. S. Cardoso, & J. Cornebise (Eds.), *Deep Learning and Data Labeling for Medical Applications* (pp. 68–76). Springer International Publishing. https://doi.org/10.1007/978-3-319-46976-8_8
- Yang, X., Liu, M., Xu, Y., Yang, X., Si, Z., Li, Z., & Zhao, Z. (2025). *SFNet: A Spatial-Frequency Domain Deep Learning Network for Efficient Alzheimer's Disease Diagnosis*. arXiv. <https://doi.org/10.48550/arXiv.2507.16267>
- Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). *How transferable are features in deep neural networks?* (No. arXiv:1411.1792). arXiv. <https://doi.org/10.48550/arXiv.1411.1792>
- Zana, F., & Klein, J.-C. (2001). Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. *IEEE Transactions on Image Processing*, 10(7), 1010–1019. <https://doi.org/10.1109/83.931095>

- Zhang, M.-L., & Zhou, Z.-H. (2014). A Review on Multi-Label Learning Algorithms. *IEEE Transactions on Knowledge and Data Engineering*, 26(8), 1819–1837. <https://doi.org/10.1109/TKDE.2013.39>
- Zhang, R., Zhao, J., Xie, H., Wang, T., Chen, G., Zhang, G., & Lei, B. (2022). Automatic diagnosis for aggressive posterior retinopathy of prematurity via deep attentive convolutional neural network. *Expert Systems with Applications*, 187, 115843. <https://doi.org/10.1016/j.eswa.2021.115843>
- Zhao, X., Chen, S., Zhang, S., Liu, Y., Hu, Y., Yuan, D., Xie, L., Luo, X., Zheng, M., Tian, R., Chen, Y., Tan, T., Yu, Z., Sun, Y., Wu, Z., & Zhang, G. (2024). A fundus image dataset for intelligent retinopathy of prematurity system. *Scientific Data*, 11(1), 543. <https://doi.org/10.1038/s41597-024-03362-5>
- Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, S...*, 11045, 3–11. https://doi.org/10.1007/978-3-030-00889-5_1
- Zhuang, J. (2019). *LadderNet: Multi-path networks based on U-Net for medical image segmentation* (No. arXiv:1810.07810). arXiv. <https://doi.org/10.48550/arXiv.1810.07810>

7. EKLER

Ek 1. CBAM (Convolutional Block Attention Module) tabanlı kanal ve uzamsal dikkat mekanizmaları

Bu modül, CBAM (Convolutional Block Attention Module) yapısının iki temel bileşeni olan kanal dikkat ve uzamsal dikkat mekanizmalarını uygular. Kanal dikkat mekanizması, özellik haritalarına uygulanan ortalama ve maksimum havuzlama işlemlerinin çıktısını paylaşımlı çok katmanlı algılayıcı (MLP) üzerinden geçirerek hangi özellik kanallarının daha önemli olduğunu, yani modelin “neye” odaklanması gerektiğini belirler. Ardından uzamsal dikkat mekanizması, kanal boyutunda sıkıştırılmış bu özellikleri birleştirip bir konvolüsyon katmanından geçirerek görüntü üzerindeki hangi bölgelerin veya piksellerin daha kritik olduğunu, yani “nereye” odaklanması gerektiğini öğrenir. Bu sayede model, gürültüyü bastırarak ayırt edici detaylara daha fazla ağırlık verebilir.

```
# Kanal Dikkat Mekanizması
# Her kanalın önemini öğrenerek kanal bazında ağırlıklandırma yapar.
class ChannelAttention(nn.Module):
    def __init__(self, in_channels, reduction_ratio=8):
        super(ChannelAttention, self).__init__()
        self.avg_pool = nn.AdaptiveAvgPool2d(1)

        # Global Max Pooling - her kanalın maksimum değerini alır
        self.max_pool = nn.AdaptiveMaxPool2d(1)

        # Paylaşımlı MLP (çok katmanlı algılayıcı) - kanal önemini
        self.fc = nn.Sequential(
            nn.Linear(in_channels, in_channels // reduction_ratio, bias=False),
            nn.ReLU(inplace=True),
            nn.Linear(in_channels // reduction_ratio, in_channels, bias=False)
        )
        self.sigmoid = nn.Sigmoid()

    def forward(self, x):
        b, c, _, _ = x.size()
        avg_out = self.fc(self.avg_pool(x).view(b, c))
        max_out = self.fc(self.max_pool(x).view(b, c))
```

```

        out = avg_out + max_out
        return self.sigmoid(out).view(b, c, 1, 1)

# Uzamsal Dikkat Mekanizması
# Görüntünün hangi bölgelerinin önemli olduğunu öğrenir
class SpatialAttention(nn.Module):
    def __init__(self, kernel_size=5):
        super(SpatialAttention, self).__init__()

        # 2 kanallı girişi (avg + max) tek kanallı dikkat haritasına dönüştürür
        self.conv=nn.Conv2d(2,1,kernel_size=kernel_size,padding=kernel_size //
2)

        self.sigmoid = nn.Sigmoid()

    def forward(self, x):
        # Kanal boyutu boyunca ortalama al -> (batch, 1, H, W)
        avg_out = torch.mean(x, dim=1, keepdim=True)

        # Kanal boyutu boyunca maksimum al -> (batch, 1, H, W)
        max_out, _ = torch.max(x, dim=1, keepdim=True)

        # İki haritayı birleştir -> (batch, 2, H, W)
        x = torch.cat([avg_out, max_out], dim=1)

        return self.sigmoid(self.conv(x))

```

Ek 2. Dengesiz sınıf problemlerinde BCE, Dice ve Focal kayıplarını öğrenilebilir ağırlıklarla birleştiren hibrit kayıp fonksiyonu

Bu modül, damar segmentasyonu için Binary Cross-Entropy (piksel bazlı sınıflandırma), Dice Loss (bölge örtüşme ölçümü) ve Focal Loss (zor örneklerle odaklanma) kayıp fonksiyonlarını öğrenilebilir ağırlıklarla birleştiren hibrit bir kayıp fonksiyonu ile, dengesiz sınıf dağılımına sahip medikal görüntülerde daha doğru segmentasyon sonuçları elde edilmesini sağlar.

```
class HybridLoss(nn.Module):
    def __init__(self, alpha=0.3, beta=0.4, gamma=2.0, pos_weight=5.0):
        super().__init__()
        self.bce=nn.BCEWithLogitsLoss(pos_weight=torch.tensor([pos_weight]))
        self.dice = DiceLoss()
        self.focal = FocalLoss(gamma=gamma)
        # Ağırlıkların toplamı 1 olacak
        self.weights = nn.Parameter(torch.tensor([alpha, beta, 1-alpha-beta]))
        self.softmax = nn.Softmax(dim=0)

    def forward(self, pred, target):
        weights = self.softmax(self.weights)
        return (weights[0] * self.bce(pred, target) +
                weights[1] * self.dice(pred, target) +
                weights[2] * self.focal(pred, target))

# Dice Loss (Dice Katsayısı tabanlı kayıp fonksiyonu)
class DiceLoss(nn.Module):
    def __init__(self, smooth=1e-6):
        super(DiceLoss, self).__init__()
        self.smooth = smooth

    def forward(self, pred_logits, target):
        pred = torch.sigmoid(pred_logits)
        pred = pred.contiguous()
        target = target.contiguous()

        intersection = (pred * target).sum(dim=2).sum(dim=2)
        dice = (2. * intersection + self.smooth) / (
            pred.sum(dim=2).sum(dim=2)+target.sum(dim=2).sum(dim=2)+ self.smooth)
        return 1 - dice.mean()
```

```

class FocalLoss(nn.Module):
    def __init__(self, gamma=2):
        """
        Args:
            gamma: Odaklama parametresi
            gamma=0: standart cross-entropy'ye eşit
        """
        super(FocalLoss, self).__init__()
        self.gamma = gamma

    def forward(self, pred, target):
        """
        Args:
            pred: Model çıktısı (logitler)
            target: Gerçek etiketler (0 veya 1)
        """
        # Logitleri olasılığa dönüştür
        pred = torch.sigmoid(pred)
        # Standart binary cross-entropy hesapla
        bce = F.binary_cross_entropy(pred, target)
        # pt = doğru sınıfın olasılığı
        pt = torch.exp(-bce)
        # Focal ağırlık:  $(1-pt)^{\gamma}$  - zor örnekleri vurgular
        focal_loss = (1 - pt) ** self.gamma * bce
        return focal_loss.mean()

```

Ek 3: Orijinal fundus görüntüsü ve segmentasyon maskesini ayrı EfficientNet kollarıyla işleyip SE-Block ile birleştiren çift-dallı füzyon tabanlı ROP sınıflandırma modeli

Bu modül, ROP (Prematüre Retinopatisi) hastalığını sınıflandırmak amacıyla orijinal fundus görüntüsünü ve segmentasyon maskesini iki ayrı kolda (EfficientNet varyantları ile) işleyip birleştiren çift dallı bir füzyon mimarisini tanımlamaktadır. Bu yapının içinde yer alan SE-Block (Squeeze-and-Excitation) modülü, görüntü ve maske kollarından gelen öznitelikler birleştirildikten sonra devreye girerek kanal dikkati uygular; yani ağın, sınıflandırma kararı verirken hangi koldan gelen bilginin veya hangi öznitelik kanalının daha önemli olduğuna karar vermesini sağlar ve ağırlıkları buna göre dinamik olarak yeniden düzenler.

```

"""
Args:
model_name: EfficientNet varyantı
num_classes: Çıkış sınıf sayısı ( Normal, Pre-Plus, Plus )
"""

class ROPFusionEfficientNet(nn.Module):
    def __init__(self, model_name='efficientnet-b0', num_classes=3):
        super().__init__()
        self.backbone_img = EfficientNet.from_pretrained(model_name)
        feat_dim = self.backbone_img._fc.in_features
        self.backbone_img._fc = nn.Identity()
        self.backbone_mask = EfficientNet1ch(model_name, pretrained=False)
        self.se = SEBlock(feat_dim * 2)
        self.fusion_fc = nn.Sequential(
            nn.Linear(feat_dim * 2, feat_dim),
            nn.ReLU(inplace=True),
            nn.Dropout(0.3)
        )
        self.classifier = nn.Linear(feat_dim, num_classes)

    def forward(self, img, mask):
        feat_img = self.backbone_img.extract_features(img)
        feat_img = F.adaptive_avg_pool2d(feat_img, 1).squeeze(-1).squeeze(-1)
        feat_mask = self.backbone_mask(mask)
        fused = torch.cat([feat_img, feat_mask], dim=1)

```

```
        fused = self.se(fused)
        x = self.fusion_fc(fused)
        return self.classifier(x)

class SEBlock(nn.Module):
    def __init__(self, channels, reduction=16):
        super(SEBlock, self).__init__()
        self.fc1 = nn.Linear(channels, channels // reduction)
        self.relu = nn.ReLU(inplace=True)
        self.fc2 = nn.Linear(channels // reduction, channels)
        self.sigmoid = nn.Sigmoid()
    def forward(self, x):
        y = self.fc1(x)
        y = self.relu(y)
        y = self.fc2(y)
        y = self.sigmoid(y)
        return x * y.expand_as(x)
```

ÖZGEÇMİŞ

Lise öğrenimini Selma Yiğitalp Lisesi'nde görmüştür. Yüksek öğrenimini Karadeniz Teknik Üniversitesi Bilgisayar Mühendisliği Bölümü'nde tamamlayarak 2003 yılında mezun olmuştur. Profesyonel kariyerine mezuniyetinin ardından Ankara ODTÜ Teknokent'te Likom Yazılım A.Ş.'de çalışarak başlamıştır. 6 yıl süren bu dönemde önemli projelerde aktif rol alarak sektörel deneyim ve teknik uzmanlık kazanmıştır. 2009 yılından bu yana Karadeniz Teknik Üniversitesi Bilgi İşlem Dairesi'nde görev yapmaktadır. 2024 yılında Karadeniz Teknik Üniversitesi tarafından verilen "Yılın Teknolojik Çözümler Üreten Personeli" ödülünü almıştır. Yabancı dil olarak İngilizce bilmektedir.

Murat, A., & Nabiyeu, V., (2025) Deep Learning-based Retinal Vessel Segmentation: Attention U-net. Scientific Journal "Bulletin of K. Zhubanov Aktobe Regional University" <https://doi.org/10.70239/arsu.2025.t81.n3.04>

Murat, A., ve Nabiyeu, V., Retina Damar Segmentasyonunda U-Net ve Geliştirilmiş Mimari Modeller, Proceeding Book of 5th International Conference on Frontiers in Academic Research ICFAR 2025, Konya, Eylül 2025 ,612-620.