

T.C.
BANDIRMA ONYEDİ EYLÜL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

KOLONOSKOPİK GÖRÜNTÜLERDE POLİP SEGMENTASYONU

Yaren AKGÖL

Lisansüstü Eğitim Enstitüsü

Yazılım Mühendisliği Programı

HAZİRAN 2025

T.C.
BANDIRMA ONYEDİ EYLÜL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

KOLONOSKOPİK GÖRÜNTÜLERDE POLİP SEGMENTASYONU

Yaren AKGÖL

DANIŞMAN
Dr. Öğr. Üyesi Buket TOPTAŞ

Lisansüstü Eğitim Enstitüsü

Yazılım Mühendisliği Programı

HAZİRAN 2025

ONAY

Yaren AKGÖL tarafından hazırlanan “**Kolonoskopik Görüntülerde Polip Segmentasyonu**” adlı tez çalışması 30/06/2025 tarihinde yapılan sınavla aşağıdaki danışman tarafından Bandırma Onyedİ Eylül Üniversitesi Fen Bilimleri Enstitüsü Yazılım Mühendisliği Anabilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

Dr. Öğr. Üyesi Buket TOPTAŞ

(Danışman)

Doç. Dr. İlyas ÖZER

(Üye)

Prof. Dr. Davut Hanbay

(Üye)

İmza

Bu tezin kabulü, Lisansüstü Eğitim Enstitüsü Yönetim Kurulu’nun/...../.....gün ve/..... sayılı kararı ile onaylanmıştır.

Prof. Dr. Serhat DUMAN

Enstitü Müdürü

BANDIRMA ONYEDİ EYLÜL ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS TEZİ
ETİK BEYANI

Bandırma Onyedi Eylül Üniversitesi Fen Bilimleri Enstitüsü tez yazım kılavuzuna göre hazırlamış olduğum “**Kolonoskopik Görüntülerde Polip Segmentasyonu**” adlı tezimin özgün bir çalışma olduğunu, tez hazırlanırken tüm aşamalarda bilimsel etik ilkelerine uygun davrandığımı, tez kapsamında sunulan tüm verileri bilimsel etik ilkelerine uygun elde ettiğimi, tezde faydalandığım tüm eserlere atıf yaptığımı ve kaynaklar kısmında bu eserleri gösterdiğimi beyan ederim.

11/06/2025

YAREN AKGÖL
İmza

ÖNSÖZ

Bu tez çalışmasının odak noktası, kolorektal kanserin erken teşhisinde önemli bir risk faktörü olan poliplerin tespiti ve segmentasyonudur. Dünya çapında en yaygın ve ölümcül kanser türlerinden biri olan kolorektal kanser, poliplerin zamanında ve doğru şekilde tespit edilmesi yoluyla önlenabilir bir evrede yakalanabilmektedir. Ancak, kolonoskopi gibi geleneksel yöntemler, polip tespitinde hatalara ve zaman kaybına neden olabilmektedir.

Bu doğrultuda, çalışmada görüntü işleme tekniklerinden yararlanılarak kolonoskopik görüntüler üzerinde otomatik ve yüksek doğrulukla gerçekleştirilecek polip tespiti ve segmentasyonu hedeflenmiştir.

Bu tez çalışması, yüksek lisans eğitimim süresince edindiğim akademik birikimin bir ürünü olarak sunulmuştur. Araştırma sürecinde karşılaştığım zorluklar, akademik ve kişisel gelişimime önemli katkılar sağlamıştır.

Yüksek lisans sürecimde sabırla destek veren ve bilgi ve tecrübesiyle beni yönlendiren değerli danışmanım Dr. Öğr. Üyesi Buket TOPTAŞ'a şükranlarımı sunarım. Bu günlere gelmemde büyük emekleri olan dedem Fikret AKBAŞ, ailem ve akademik hayatımda bana ilham veren hocalarıma teşekkür ederim.

Bu çalışmanın, ilgili alanda yapılacak olan araştırmalara katkı sağlamasını temenni ederim.

ÖZET

YÜKSEK LİSANS TEZİ

Kolonoskopik Görüntülerde Polip Segmentasyonu

Yaren AKGÖL

Bandırma Onyedİ Eylöl Üniversitesi

Lisansüstü Eğitim Enstitüsü

Yazılım Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Buket TOPTAŞ

Haziran 2025, 65 sayfa

Polipler, normal olmayan doku büyümeleridir. Bu anormal doku büyümeleri, başlangıçta iyi huylu doku büyümeleri olarak bilinse de, bazı polip türleri zamanla kanserli hücrelere dönüşebilmektedir. Kolorektal kanser, poliplerin kötü huylu büyümelere dönüşmesiyle ortaya çıkan bir kanser türüdür. Poliplerin erken tanısı ve uygun tedavisi, Kolorektal kanserden korunmak için önemlidir.

Kolorektal polip tanısında klinik uzman değerlendirmesi, zaman, maliyet ve insan hatası problemlerine açıktır. Bilgisayar destekli tanı sistemleri bu eksikleri gidererek otomatik tanı gerçekleştirir ve böylece bu sorunları minimize eder. Bu tez çalışmasında, Kolorektal poliplerin otomatik segmentasyonunu gerçekleştiren literatür çalışmalarının detaylı bir incelemesi yapılmıştır. Ardından, segmentasyonda sağlam ve hassas sonuçlar veren DeepLabv3+ mimarisinin farklı omurga ağlarındaki segmentasyon görevine odaklanılmıştır. Bu çalışmada, omurga ağı olarak DenseNet, ResNet50, SqueezeNet ve VGG16 mimarisi kullanılmıştır. Bu dört evrimsel sinir ağı mimarisini omurga ağı olarak alan DeepLabv3+ mimarisi Kvasir-SEG ve CVC-ClinicDB veri setleri üzerinde test edilmiştir. Kolorektal polip segmentasyon görevinde elde edilen sonuçlar yedi farklı performans metriği üzerinde değerlendirilmiştir. Bu metrikler; ortalama Dice, ortalama Kesişim/Birleşim oranı, doğruluk, geri çağırma, özgüllük, keskinlik ve ortalama en

küçük karesel hata metrikleridir. Kvasir-SEG veri seti üzerinde omurga ağı olarak DenseNet mimarisinde en iyi değerlendirme sonuçları elde edilmiştir. Bu sonuçlar sırası ile şöyledir: 0,858 oranında ortalama Dice, 0,850 oranında ortalama Kesişim/Birleşim oranı, 0,948 oranında doğruluk, 0,824 oranında geri çağırma, 0,896 oranında keskinlik ve 0,045 oranında ortalama en küçük karesel hata metriğidir. Özgüllük metriğinde ise en iyi sonuç 0,984 oranında VGG16 mimarisi ile elde edilmiştir. Tez çalışması kapsamında gerçekleştirilen diğer bir çalışma ise U-Net mimarisinin farklı kayıp fonksiyonları üzerinde ki segmentasyon görevi üzerinedir. Bu çalışmada ise, İkili Çapraz Entropi, Tversky, Dice ve Bileşik kayıpları içeren dört farklı model eğitilmiştir. Bu modellerin performansları CVC-ColonDB ve Kvasir-SEG veri setleri üzerinde kapsamlı bir şekilde ölçülmüştür. Deneysel veriler, her iki veri setinde en yüksek segmentasyon doğruluğunun İkili Çapraz Entropi kayıp fonksiyonunun ortaya koyduğunu doğrulamıştır. CVC-ColonDB veri seti için en yüksek doğruluk performansı 0.9761 oranında iken, Kvasir-SEG için bu oran 0.9313 olarak ölçülmüştür. Elde edilen bulgular, kayıp fonksiyonu optimizasyonunun doğruluk ve kararlılık üzerinde belirleyici rol oynadığını göstermektedir.

Anahtar kelimeler: Evrişimsel Sinir Ağları, Kayıp Fonksiyonları, Kolon Kanseri, Kolorektal Kanseri, Polip Segmentasyonu

ABSTRACT

M.Sc. THESIS

Segmentation of Polyps in Colonoscopic Images

Yaren AKGÖL

Bandırma Onyedi Eylül University

Graduate Education Institute

Software Engineering

Supervisor: Asst. Prof. Dr. Buket TOPTAŞ

June 2025, 65 pages

Polyps are abnormal tissue outgrowths that, although initially benign, can evolve into malignant lesions over time. Colorectal cancer arises upon the malignant transformation of such polyps; thus, early detection and appropriate treatment are essential for prevention.

Clinical assessment of colorectal polyps by expert endoscopists is susceptible to delays, high costs and human error. Computer-aided diagnostic systems address these limitations by automating the detection process, thereby reducing both time expenditure and false-negative or false-positive rates. A comprehensive literature review was conducted to examine current methods for automatic colorectal polyp segmentation. Subsequently, the DeepLabv3+ architecture—renowned for its robust and precise segmentation capabilities—was evaluated using four convolutional backbones: DenseNet, ResNet50, SqueezeNet and VGG16. Each DeepLabv3+ variant was tested on the Kvasir-SEG and CVC-ClinicDB datasets, employing seven performance metrics: mean Dice coefficient, mean Intersection over Union, accuracy, recall, specificity, precision and mean squared error. On the Kvasir-SEG dataset, the best overall performance was achieved using the DenseNet backbone. Specifically, the model obtained a mean Dice score of 0.858, a mean IoU of 0.850, accuracy of 0.948, recall of 0.824, precision of 0.896, and a mean squared error of 0.045. In terms of specificity, the highest result (0.984) was observed with the

VGG16 backbone. Another component of this thesis involved evaluating segmentation performance using the U-Net architecture with different loss functions. Four models were trained using Binary Cross-Entropy, Tversky, Dice, and Composite loss functions. The performance of these models was thoroughly evaluated on the CVC-ColonDB and Kvasir-SEG datasets. Experimental results demonstrated that the Binary Cross-Entropy loss function consistently yielded the highest segmentation accuracy across both datasets. The highest accuracy was 0.9761 on the CVC-ColonDB dataset and 0.9313 on the Kvasir-SEG dataset. These findings highlight the critical role of loss function optimization in achieving high segmentation accuracy and model stability.

Keywords: Convolutional Neural Networks, Colon Cancer, Colorectal Cancer, Loss Functions, Polyp Segmentation



İÇİNDEKİLER

	Sayfa
ÖNSÖZ	5
ÖZET	I
ABSTRACT	III
İÇİNDEKİLER.....	V
ŞEKİL LİSTESİ	VII
TABLO LİSTESİ.....	IX
SİMGE VE KISALTMA LİSTESİ	X
1. GİRİŞ.....	1
1.1 İLGİLİ ÇALIŞMALAR	6
1.1.1 Evrişimsel Sinir Ağı Kullanılarak Tasarlanmış Yöntemler	6
1.1.2 Dikkat Modülleri Kullanılarak Tasarlanmış Yöntemler	10
1.1.3 Transformatörler Kullanılarak Tasarlanmış Yöntemler	12
1.2 TEZİN AMACI	15
1.3 TEZİN GEREKÇESİ	15
1.4 TEZİN ORGANİZASYONU	16
2. MATERYAL VE YÖNTEM	17
2.1 EVRİŞİMSEL SİNİR AĞLARI	17
2.2 KAYIP FONKSİYONLARI	22
2.3 U-Net MİMARİSİ	25
2.4 DenseNet MİMARİSİ	27
2.5 ResNet50 MİMARİSİ	29
2.6 SqueezeNet MİMARİSİ	31
2.7 VGG16 MİMARİSİ	33
2.8 DeepLabv3+ MİMARİSİ	35

3. KULLANILAN VERİ SETLERİ VE PERFORMANS ÖLÇÜTLERİ	37
3.1 VERİ SETLERİ	37
3.2 PERFORMANS METRİKLERİ.....	39
4. UYGULAMALAR.....	42
4.1 UYGULAMA 1: FARKLI OMURGA AĞLARI KULLANILARAK DeepLabv3+ MİMARİSİ İLE POLİP SEGMENTASYONU	42
4.1.1 Uygulama Ortamı ve Eğitim Parametreleri.....	42
4.1.2 Önerilen Yöntem	43
4.1.3 Sonuçlar.....	44
4.2 UYGULAMA 2: U-Net TABANLI POLİP SEGMENTASYONUNDA KAYIP FONKSİYONLARININ KARŞILAŞTIRMALI ANALİZİ	47
4.2.1 Uygulama Ortamları ve Eğitim Parametreleri.....	48
4.2.2 Önerilen Yöntem	48
4.2.3 Sonuçlar	49
5. SONUÇ VE ÖNERİLER	52
6. KAYNAKLAR	54

ŞEKİL LİSTESİ

<u>Sekil</u>	<u>Sayfa</u>
Şekil 1-1: Çeşitli polip tanılama yöntemleri	2
Şekil 1-2: Polip segmentasyon görevini zorlaştıran nedenler: (a) düşük kaliteli görüntüler, (b) poliplerin şekil ve boyut farklılıkları, (c) arka planda yabancı maddelerin bulunması. Görseller CVC-ClinicDB veri setinden alınmıştır [20].	4
Şekil 2-1: Evrişim katmanı, havuzlama katmanı ve tam bağlı katmanları içeren geleneksel CNN mimarisi	17
Şekil 2-2: 4x4 boyutundaki bir girdi görüntüsü üzerinde 3x3 bir maske ile gerçekleştirilen evrişim işlemi	20
Şekil 2-3: 4x4 boyutundaki bir girdi görüntüsü üzerinde 3x3 bir maske ile gerçekleştirilen evrişim işlemi	20
Şekil 2-4: Tam bağlı katman	22
Şekil 2-5: U-Net mimarisinin genel yapısı	25
Şekil 2-6: DenseNet mimarisi.....	28
Şekil 2-7: Artık blok	29
Şekil 2-8: ResNet50 mimarisi.....	30
Şekil 2-9: Sıkıştırma-genişletme modülü	32
Şekil 2-10: SqueezeNet mimarisinin genel yapısı	33
Şekil 2-11: VGG16 mimarisi.....	34
Şekil 2-12: DeepLabv3+ mimarisinin genel yapısı	36
Şekil 3-1: CVC-ClinicDB veri setine ait görseller: (a) orijinal polip görüntüsü, (b) orijinal görüntülere karşılık gelen gerçek ikili görüntüleri temsil etmektedir.	38
Şekil 3-2: Kvasir-SEG veri setine ait görseller: (a) orijinal polip görüntüsü, (b) orijinal görüntülere karşılık gelen gerçek ikili görüntüleri temsil etmektedir.	38
Şekil 3-3: CVC-ColonDB veri setine ait görseller: (a) orijinal polip görüntüsü, (b) orijinal görüntülere karşılık gelen gerçek ikili görüntüleri temsil etmektedir.	39
Şekil 4-1: Kvasir-SEG veri seti üzerinde farklı modellerin segmentasyon performans metrikleri.....	45
Şekil 4-2: CVC-ClinicDB veri seti üzerinde farklı modellerin segmentasyon performans metrikleri.....	46
Şekil 4-3: Kvasir-SEG [118] veri seti üzerindeki segmentasyon sonuçları	47

Şekil 4-4: CVC-ClinicDB [20] veri seti üzerindeki segmentasyon sonuçları	47
Şekil 4-5: CVC-ColonDB veri seti için modeller arası <i>mDice</i> karşılaştırılması.....	51
Şekil 4-6: Kvasir-SEG veri seti için modeller arası <i>mDice</i> karşılaştırılması.....	51



TABLO LİSTESİ

<u>Tablo</u>	<u>Sayfa</u>
Tablo 1-1: CNN mimarisi kullanılarak tasarlanmış yöntemlerin özeti.....	9
Tablo 1-2: Dikkat modülleri kullanılarak tasarlanmış yöntemlerin özeti	11
Tablo 1-3: Transformatörler kullanılarak tasarlanmış yöntemlerin özeti	14
Tablo 3-1: Literatürde kullanılan polip segmentasyon veri setlerine ait genel bilgiler	37
Tablo 4-1: Eğitim parametrelerinin değerleri	42
Tablo 4-2: Kullanılan modeller ve omurga ağları	43
Tablo 4-3: Kvasir-SEG ve CVC-ClinicDB veri setleri üzerinde elde edilen segmentasyon performans sonuçları	45
Tablo 4-4: Eğitim parametrelerinin değerleri	48
Tablo 4-5: CVC-ColonDB [119] veri seti üzerinde kayıp fonksiyonları ile elde edilen performans metrikleri.....	50
Tablo 4-6: Kvasir-SEG [118] veri seti üzerinde kayıp fonksiyonları ile elde edilen performans metrikleri.....	50

SİMGE VE KISALTMA LİSTESİ

Simgeler

Açıklama

$\approx$	: Yaklaşık
σ	: Aktivasyon Fonksiyonunu
e	: Euler Sabiti
$*$	: Evrişim İşlemi
$\oplus$	: Kanal Ekseninde Birleştirme İşlemi

Kısaltmalar

Açıklama

Adam	: Uyarlanabilir Moment Tahmini (Adaptive Moment Estimation)
Acc	: Doğruluk (Accuracy)
ASPP	: Aralıklı Uzamsal Piramit Havuzlama Modülleri (Atrous Spatial Pyramid Pooling)
BCE	: İkili Çapraz Entropi (Binary Cross Entropy)
BDT	: Bilgisayarlı Destekli Tanı
CNN	: Evrişimsel Sinir Ağları (Convolutional Neural Network)
DenseNet	: Yoğun Bağlantılı Evrişimsel Ağlar (Densely Connected Convolutional Networks)
DRE	: Dijital Rektal Muayene (Digital Rectal Examination)
FCN	: Tam Konvolüsyonel Ağ (Fully Convolutional Network)
FN	: Yanlış Negatif (False Negative)
FP	: Yanlış Pozitif (False Positive)
MAE	: Ortalama Mutlak Hata (Mean Absolute Error)
mDice	: Ortalama Dice Katsayısı (Mean Dice Coefficient)
mIoU	: Ortalama Kesişim Bölme Oranı (Mean Intersection over Union)
MRG	: Manyetik Rezonans Görüntüleme (Magnetic Resonance Imaging)
Pr	: Kesinlik (Precision)
ReLU	: Doğrusal Düzeltme Birimi (Rectified Linear Unit)
ResNet	: Artık Ağ (Residual Network)
Sp	: Özellik / Doğru Negatif Oranı (Specificity)
TN	: Doğru Negatif (True Negative)

TP	: Doğru Pozitif (False Positive)
ViT	: Görsel Dönüştürücü (Vision Transformer)
VGG	: Görsel Geometri Grubu (Visual Geometry Group)
WHO	: Dünya Sağlık Örgütü (World Health Organization)
YOLO	: Sadece Bir Kez Bak (You Only Look Once)

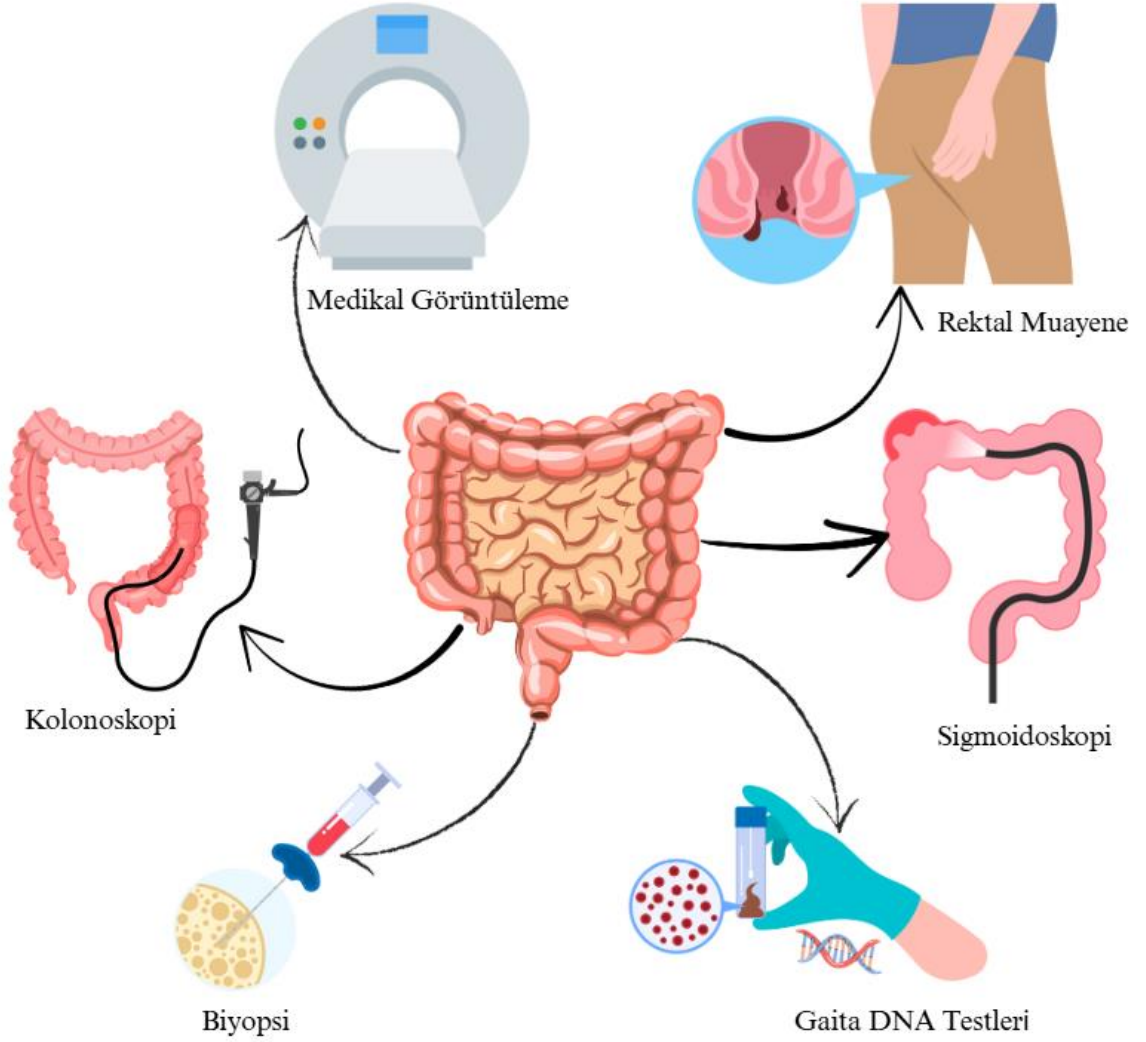


1. GİRİŞ

Polipler, vücudun içinde oluşan küçük ve anormal doku büyümeleridir. Bu anormal doku büyümeleri, burun [1], mide [2], bağırsak [3] ve rahim [4] gibi organların iç yüzeyinde görülebilirler. Polipler, başlangıçta iyi huylu lezyonlar olarak tanımlansa da bazı polip türleri zamanla kansere dönüşme riski taşımaktadır. Dünya Sağlık Örgütü'nün (World Health Organization-WHO) 2022 verilerine göre, kalın bağırsak ve rektumda gelişen kötü huylu tümörleri ifade eden kolorektal kanser, dünya çapında en yaygın üçüncü kanser türü olarak sınıflandırılmıştır ve yıllık yaklaşık 1,9 milyondan fazla yeni vaka bildirilmiştir [5]. Ayrıca, 2023 verilerine göre, 2040 yılına kadar kolorektal kanserin küresel yükünün yılda 3,2 milyon yeni vaka (%63 artış) ve yılda 1,6 milyon ölüm (%73 artış) şeklinde artması öngörülmektedir [6]. Sağlık Bakanlığı Halk Sağlığı Genel Müdürlüğü'nün kanser istatistik verilerine göre Türkiye'de kolorektal kanserin görülme sıklığı hem erkeklerde hem de kadınlarda üçüncü sırada yer almaktadır [7]. Bu istatistikler, kolorektal kanserin poliplerin malign dönüşümüyle yakından ilişkili olduğunu ve erken tanı ile müdahalenin hastalığın seyri üzerinde belirleyici bir rol oynayabileceğini ortaya koymaktadır.

Kolorektal poliplerin tanısında çeşitli yöntemler uygulanmaktadır. Bu yöntemler, fiziksel muayene, medikal görüntüleme, kolonun iç kısmının incelenmesi, histopatolojik inceleme ve genetik testlerden oluşmaktadır. Fiziksel muayene, hastanın anamnezi bilgisinin detaylı bir şekilde incelenmesine dayanır. Burada, hekim hastanın dışkıda kanama, ani bağırsak alışkanlığı değişikliği gibi semptomlarını değerlendirerek polip varlığına işaret edebilecek bulguları tespit etmeyi amaçlar. Ayrıca, Dijital Rektal Muayene (Digital Rectal Examination-DRE) yöntemi ile eldiven ile parmak kullanılarak rektum palpasyonu yapılmakta; bu yöntem, özellikle rektuma yakın ve büyük boyutlu poliplerin varlığının saptanmasında kullanılmaktadır. Ancak, DRE yöntemi küçük veya kolorektal sistemin diğer bölgelerinde yer alan lezyonların tespitinde yetersiz kalmaktadır. Medikal görüntüleme ise, ultrason ya da Manyetik Rezonans Görüntüleme (Magnetic Resonance Imaging-MRG) teknikleri kullanılarak poliplerin varlığını değerlendirmeyi amaçlar. Ancak, bu yöntemler, uzamsal çözünürlük, kontrast farklılıkları ve görüntüleme artefaktları gibi sınırlamalar nedeniyle, özellikle küçük boyutlu veya yassı poliplerin saptanmasında yetersiz kalabilmektedir. Daha kapsamlı ve ayrıntılı bir değerlendirme için genellikle kolonoskopi yöntemi tercih edilmektedir. Poliplerin tanısında kullanılan çeşitli yöntemlerin gösterimi Şekil 1-1'de görsel olarak sunulmuştur. Bu görselde gösterilen kolonoskopi işlemi, kolorektal sistemin tamamının doğrudan görselleştirilmesine olanak tanır [8] ; bu sayede poliplerin boyut, konum ve

morfolojik özellikleri detaylı bir şekilde incelenebilir. Kolonoskopinin daha kısa bir versiyonu olan sigmoidoskopi [8] yöntemi ise yalnızca kalın bağırsağın alt kısmını incelemekte olup daha sınırlı bir görüntüleme imkânı sunmaktadır.



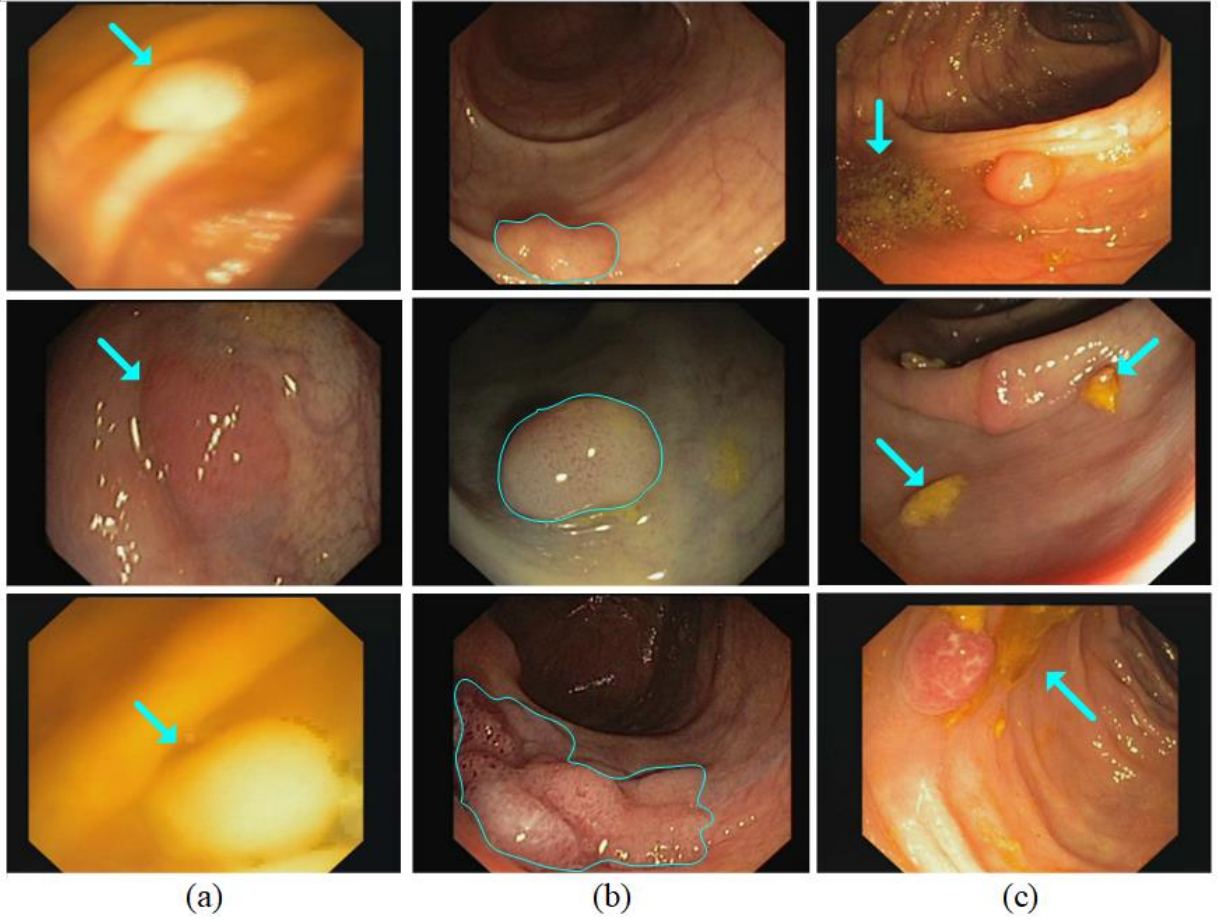
Şekil 1-1: Çeşitli polip tanılama yöntemleri

Kolonoskopi, kolorektal polip tespitinde güvenilir tanı yöntemlerinden biri olarak kabul edilmektedir [9]. Ancak, bu yöntem klinik uygulamalarda, karmaşık ve zaman alıcı yapısı nedeniyle hekimin deneyimine ve teknik yeterliliğine önemli ölçüde bağlıdır [10]. Söz konusu bağımlılık, yanlış veya eksik tespitlere yol açarak erken tanı sürecini geciktirebilmektedir. Yapılan araştırmalar, manuel inceleme sürecinde poliplerin yaklaşık %30'unun gözden kaçırılabilirliğini öne sürmektedir [11]. Bu nedenle, erken tanının doğruluğunu ve verimliliğini artırmak için

Bilgisayarlı Destekli Tanı (BDT) sistemleri kritik bir araştırma ve uygulama alanı haline getirmiştir [12].

Polip tespitinin hassasiyetini ve tutarlılığını artırmak için, BDT sistemlerine olan ilgi son yıllarda artış göstermiştir [13]. BDT sistemleri, kolonoskopi sırasında elde edilen endoskopik görüntüleri otomatik veya yarı otomatik biçimde analiz ederek hekime erken ve doğru tanı konusunda destek sağlamayı amaçlayan sistemlerdir. Bu sistemler; gelişmiş görüntü işleme, makine öğrenmesi ve derin öğrenme tabanlı yaklaşımlar sayesinde, poliplerin renk, doku ve geometrik özelliklerini inceleyerek tespit, sınıflandırma ve segmentasyon görevlerini yerine getirir.

BDT sistemlerinde, poliplerin tespiti; geleneksel görüntü ön işleme teknikleri, kenar tespiti ve bölge çıkarım yöntemleri kullanılarak veya Sadece Bir Kez Bak (You Only Look Once-YOLO) gibi son teknolojik yöntemler kullanılarak potansiyel lezyon alanlarının belirlenmesi işlemi kapsar [14] [15]. Poliplerin sınıflandırılması ise, polip görüntülerinden çıkarılan özniteliklerin analiz edilme sürecini kapsar. Son dönemde, derin ağ mimarilerinin sınıflandırma görevindeki başarısını kanıtlamasıyla, poliplerin sınıflandırılmasında Evrimsel Sinir Ağlarının kullanımı literatürde popüler bir yaklaşım olarak öne çıkmıştır [16] [17]. Polip segmentasyonu, polip araştırmalarında diğer görevler arasında daha az çalışılmış olanıdır [18]. Segmentasyon, dijital görüntüdeki piksellerin benzer özelliklere göre gruplandırılarak, anlamlı bölgelerin veya nesnelerin ayrıştırılma sürecidir. Polip segmentasyonu ise kolonoskopi görüntüleri üzerinde poliplerin morfolojik özelliklerinin, konumlarının ve sınırlarının belirlenerek sağlıklı dokudan ayrılmasını sağlayan bir analiz sürecidir. Polip segmentasyonu, poliplerin sadece konum ve sınır bilgilerini ortaya koyması nedeniyle, semantik düzeyde bilgi sunan tespit ve sınıflandırma görevlerine göre daha sınırlı bir görev tanımı taşımaktadır. Bu sınırlılığın temel nedenlerinden biri, poliplerin arka planla benzer renklere sahip olması ve polip ile arka plan arasındaki sınırların net bir şekilde ayırt edilememesidir. Ayrıca, poliplerin şekil, boyut, renk ve diğer yapısal özelliklerindeki farklılıkları, segmentasyon görevini daha karmaşık hale getirmektedir [19]. Polip segmentasyon görevini zorlaştıran nedenlerden bazıları Şekil 1-2’de görsel olarak sunulmuştur. Bu görselde, bahsedilen zorluklar mavi ok ile gösterilmiştir.



Şekil 1-2: Polip segmentasyon görevini zorlaştıran nedenler: (a) düşük kaliteli görüntüler, (b) poliplerin şekil ve boyut farklılıkları, (c) arka planda yabancı maddelerin bulunması. Görseller CVC-ClinicDB veri setinden alınmıştır [20].

Polip segmentasyon görevinde, geleneksel yöntemler genellikle manuel işlem gerektirirken, BDT sistemleri daha hızlı ve doğru analizler sunmaktadır. Son yıllarda, donanım teknolojilerindeki gelişmelere bağlı olarak derin öğrenme tabanlı yöntemler yaygınlaşmış ve geleneksel BDT yöntemlerine kıyasla daha güvenilir sonuçlar elde edilmiştir. Özellikle, Tam Konvolüsyonel Ağ (Fully Convolutional Network-FCN) Mimarisi [21], U-Net mimarisi [22] ve Transformer mimarisi [23] çeşitli alanlardaki segmentasyon görevlerinde güvenilir biçimde yüksek doğruluk oranları sunmuştur.

FCN mimarisinin [21] 2014 yılında tanıtılması ile birlikte segmentasyon görevinde önemli ilerlemeler sağlanmıştır. Bu mimari, geleneksel tam bağlı katmanlar yerine yalnızca evrimsel katmanları kullanmıştır. Geleneksel tam bağlı katmanlar sabit boyutlu girdiler gerektirirken, evrimsel katmanlar yerel filtreler aracılığıyla uzaysal boyutları koruyarak çalışır. FCN mimarisi [21], yalnızca evrimsel katmanlar kullandığı için herhangi bir boyuttaki görüntüyü işleyebilir,

böylece esnek bir şekilde farklı boyutlardaki görüntülerle piksel bazında segmentasyon yapabilir. Bu mimari, hareketli hedef tanımda [24], tıbbi alanda [25] [26], tarımda [27] [28] ve farklı birçok alandaki segmentasyon görevinde stratejik bir temel oluşturmuştur. FCN mimarisinin [21], düşük çözünürlüklü çıktılar üretmesi küçük nesneleri ayırt etmekte zorlanması ve detay kaybına neden olması segmentasyon görevindeki performansını sınırlamaktadır [29]. Bu sınırlamalar, FCN mimarisini [21] temel alan ve yapısal olarak U şeklinde tasarlanan U-Net mimarisi [22] ile büyük ölçüde aşılmıştır.

U-Net mimarisi [22], görüntü segmentasyonu alanında yaygın olarak kullanılan ve FCN mimarisine [21] dayalı geliştirilmiş bir mimari [30]. Bu mimari, bağlam bilgisi ve kesin konumlandırma elde etmek için kodlayıcı, kod çözücü ve atlama bağlantılarından oluşmaktadır. Kodlayıcı, girdi görüntüsünden çok ölçekli özellik haritalarının çıkarılmasına odaklanırken, kod çözücü ise bu özellik haritalarının yeniden yapılandırılmasını sağlar. Kodlayıcı ve kod çözücü arasında bulunan atlama bağlantıları ise, aynı seviyedeki katmanları bağlayarak kodlayıcıdan gelen yüksek çözünürlüklü bilgiyi kod çözücüye aktarır. U-Net mimarisinin [22] verimli tasarımı, esnek yapısı ve uygulama kolaylığı çeşitli alanlardaki segmentasyon sorunlarında mükemmel bir performans sergilemiş [31] ve bu durum U-Net mimarisini [22] görüntü segmentasyonunda hızlı yaygın olarak benimsenen bir mimari haline getirmiştir [32]. Tıbbi alan [33] [34] [35], tarım [36] [37] ve uzaktan algılama [38] [39] gibi birçok farklı alanda çeşitli uygulamalara sahiptir.

U-Net mimarisi [22], güçlü bir mimari yapıya sahip olmasına rağmen kodlayıcı ve kod çözücü katmanları arasındaki özellik birleştirme aşamasında anlamsal bilgi kayıpları yaşamaktadır [40]. Yaşanan bilgi kaybı özellikle küçük ölçekli nesnelerin ve ince detayların segmentasyonunda performans düşüşlerine yol açmaktadır [41]. Bu sorunu gidermek için son yıllarda dikkat mekanizmaları geliştirilmiştir. Dikkat, insan algısında önemli bir rol oynamaktadır [42]. İnsan dikkat mekanizmasının, derin öğrenme mimarilere uygulanmasına dikkat mekanizmaları adı verilmektedir [43]. Bu mekanizmalar, U-Net mimarisinin [22] atlama bağlantılarına eklenerek önemli bilgileri vurgulamakta ve alakasız bilgilerin etkilerini azaltmaktadır [44]. Bu durum, segmentasyon doğruluğunu artırırken aynı zamanda anlamsal bütünlüğü korumaktadır.

Dikkat mekanizmaları, başlangıçta doğal dil işleme alanında popülerleşmiş ve oldukça başarılı sonuçlar elde etmiştir [45]. Bununla birlikte, son yıllarda dikkat mekanizmaları, görsel veriler üzerinde de etkili bir şekilde uygulanarak başarılı sonuçlar elde etmiştir [46]. Özellikle U-Net tabanlı dikkat mekanizmaları, örneğin Attention U-Net [47], Residual Attention U-Net [48], Multi-Scale Attention U-Net [49] gibi varyasyonlar, segmentasyon doğruluğunu önemli ölçüde iyileştirmiştir. Bu mekanizmalar, U-Net'in segmentasyon görevlerinde karşılaştığı çok ölçekli

bağlam bilgisini yeterince kullanamama, uzun menzilli bağımlılıkları yakalamada zorlanma ve sınır detaylarını net biçimde belirleyememe gibi sınırlamaları ortadan kaldırarak daha yüksek bir performans elde edilmesini sağlamıştır. Dikkat mekanizmalarının başarıları, özellikle görüntü sınıflandırma [50], yüz tanıma [51], görüntü üretimi [52] gibi alanlarda önemli bir rol oynarken, bu mekanizmaların daha verimli bir şekilde kullanılabilmesi için Transformer mimarisi [53] geliştirilmiştir.

Transformer mimarisi [53], 2017 yılında önerilmiş olup, tamamen dikkat mekanizmasına dayalı bir yapı sunmaktadır [53]. Başlangıçta makine çevirisi gibi doğal dil işleme alanında tasarlanan Transformer mimarisi [53], son yıllarda görsel veriler üzerinde de kullanılmaya başlanmıştır [54]. Günümüzde derin öğrenme mimarilerinde yüksek performans ve esneklik sağlayan Transformer mimarisi [53], özellikle doğal dil işleme, görüntü işleme ve diğer yapay zekâ uygulamalarında kritik bir rol oynamaktadır [55]. Ayrıca, Transformer mimarisi [53] tabanlı modellerin görüntü sınıflandırma [56], nesne tespiti [57] ve segmentasyon [58] gibi alanlarda doğruluk artışları sağlamış ve bu görevlerde yaygın olarak tercih edilmeye başlanmıştır.

1.1 İLGİLİ ÇALIŞMALAR

Bu tez çalışmasında, polipleri bölütlenme görevinin literatürdeki yeri üç başlık altında incelenmiştir. Bu üç başlık sırasıyla; Evrişimsel Sinir Ağı Kullanılarak Tasarlanmış Yöntemler, Dikkat Modülleri Kullanılarak Tasarlanmış Yöntemler ve U-Net Mimarisi Tabanlı Yöntemlerdir. Bu yöntemleri kullanan literatür çalışmaları, tezin bu bölümünde alt başlıklar halinde sunulmuştur.

1.1.1 Evrişimsel Sinir Ağı Kullanılarak Tasarlanmış Yöntemler

Evrişimsel Sinir Ağları (Convolutional Neural Network-CNN) mimarisi ilk kez 1998 yılında doküman tanıma alanında tanıtılmış bir derin ağ mimarisidir [59]. 2015 yılında “ImageNet” [60] ismi verilen bir veri setinin duyurulmasıyla, CNN mimarileri büyük ölçekli görüntü sınıflandırma görevlerinde kullanılmaya başlanmıştır [61]. Bu mimarilerin görüntü analizlerindeki başarılı sonuçları, birçok araştırmacının dikkatini çekmesine yol açmıştır [59].

Huang ve arkadaşları [62], kodlayıcı ve kod çözücü katmanları ile tasarlanmış bir derin öğrenme ağı önermişlerdir. Kod çözücü, polip segmentasyon görevini daha hızlı ve daha doğru gerçekleştirilmesini sağlamaktadır. Bu çalışmada, çıkarım süresini azaltmak ve işlem verimliliğini artırmak amacıyla HarDNet68 mimarisi, omurga ağ olarak kullanılmıştır. Geliştirilen ağ, beş farklı

kolonoskopi veri seti üzerinde test edilmiştir. Ayrıca, çalışmanın kaynak kodları açık erişimli olarak yayınlanmıştır.

Mahmud ve arkadaşları [19], kolonoskopik görüntülerin daha doğru ve verimli bir şekilde analiz edilmesini sağlamak amacıyla PolypSegNet adında bir model önermişlerdir. Bu çalışmada, geleneksel segmentasyon ağlarının sahip olduğu sınırlılıkları aşmak amacıyla kodlayıcı ve kod çözücünden yararlanılmıştır. Önerilen modelde, görüntüler modele verilmeden önce ön işlemden geçirilmiştir. Ön işlem aşamasında, tüm görüntüler ağ girişine uygun olacak şekilde yeniden boyutlandırılmış ve genlik normalizasyonu yöntemi kullanılarak piksel değerleri ölçeklendirilmiştir.

Wang ve arkadaşları [63], polip segmentasyonu görevi için ISCNet adında bir model önermişlerdir. Önerilen model, dört aşamadan oluşmaktadır. İlk aşama, ResNet34 omurgasından oluşmaktadır. Bu omurga, beş farklı derinlik düzeyinde özellik haritaları çıkararak kolonoskopi görüntülerinde detayların etkili biçimde yakalanmasını sağlamaktadır. İkinci aşama, poliplerin kolonoskopi görüntüleri üzerinde tespit edildiği aşamadır. Üçüncü aşama ise, çeşitli çevresel bağlamlara odaklanma modüllerinin kullanıldığı aşamadır. Bu modül sayesinde poliplerin sınırları daha net ve kesin bir şekilde tespit edilmekte, modelin performansı artırılmaktadır. Son aşama ise kod çözücü aşamasıdır.

Zhang ve arkadaşları [64], karmaşık kolonoskopi görüntülerini etkili bir şekilde işlemek için bağlam seçim ağı modelini önermişlerdir. Model kodlayıcı ve kod çözücünden oluşmaktadır. Ayrıca önerilen model, yerel bağlam dikkati modülü, küresel bağlam modülü ve adaptif seçim modülünden oluşmaktadır. Yerel bağlam dikkati modülü, zor bölgelerde dikkat artırarak yerel özelliklerin kodlayıcıdan kod çözücüye aktarılmasını sağlamaktadır. Küresel bağlam modülü, tüm görüntünün genel bilgisini toplarken; adaptif seçim modülü bu bilgidен özellikleri seçip birleştirmektedir.

Su ve arkadaşları [65], polip segmentasyon görevinde performans arttırmak amacıyla özellik-zenginleştirilmiş piramit ağı adında bir model önermişlerdir. Model, üç modülden oluşmaktadır: çapraz gömme modülü, tahmine dayalı kalibrasyon modülü ve hiyerarşik özellik birleştirme modülü. Rastgele ölçeklendirme, döndürme, genişletme, aşındırma ile görüntü kontrastı, keskinliği ve parlaklığının rastgele ayarlanmasını gibi çeşitli veri artırma yöntemleri kullanılmıştır. Modelde ayrıca halka açık veri setleri kullanımı tercih edilmiştir.

Gökhan ve arkadaşları [66], polip segmentasyon görevinde veri dengesizliği sorununu çözmek ve daha duyarlı ve genelleştirilebilir sonuçlar elde etmek için yeni bir kayıp fonksiyonu önermişlerdir. Önerilen kayıp fonksiyonunun performansını değerlendirmek amacıyla iki farklı

yöntem uygulanmıştır. Birinci yöntemde, 18 katmanlı bir ResNet omurgası kullanılarak U-Net mimarisi içinde kod çözücü görevi üstlenmiştir. İkinci yöntemde ise, 34 katmanlı bir ResNet ağı kodlayıcı olarak ve U-Net yapısı kod çözücü olarak kullanılmıştır.

Xia ve arkadaşları [67], polip segmentasyonu görevinde daha iyi performans elde etmek amacıyla ağırlık dengesizliği ve örnek karmaşıklığını optimize eden bir model önermişlerdir. Önerilen model çok ölçekli özellik geliştirme ve birleştirme ağı olarak adlandırılmıştır. Ayrıca önerilen modelde ağırlık optimizasyonu, özellik geliştirme ve dikkat mekanizmaları gibi bileşenler kullanılarak polip segmentasyon görevinin iyileştirilmesi sağlanmıştır.

Zhou ve arkadaşları [68], polip segmentasyon görevi için çapraz düzey özellik birleşim ağı önermişlerdir. Önerilen model üç ana bileşenden oluşmaktadır. Önerilen model, düşük seviyeli özellikleri birleştirerek ve sınır farkındalığını belirleyerek segmentasyon sonuçlarını iki çözücünün çıktılarıyla elde etmektedir. Bu durum polip segmentasyon görevinde daha hassas sonuçlar elde edilmesini sağlamaktadır. Ayrıca çalışmada kod açık kaynaklı kod olarak paylaşılmıştır.

Hu ve arkadaşları [69], polip segmentasyon görevi sürecinde karşılaşılan zorlukları aşmak amacıyla Görsel Dönüştürücü (Vision Transformer-ViT) tabanlı yeni bir model olan DAN-PD'yi önermişlerdir. Önerilen modelde, daha hassas özellikler elde etmek amacıyla paralel kodlayıcılar kullanılmıştır. Ayrıca, daha iyi bir genelleme sağlanması amacıyla model U-Net mimarisi ile tasarlanmıştır. Bu modelde, polip segmentasyonunda yenilikçi bir çözüm sunmak amacıyla denetimsiz alan uyarlaması yöntemleri ile karşıt öğrenme stratejileri bir arada kullanılmıştır.

Su ve arkadaşları [70], polip segmentasyonu görevinde kullanılmak üzere yeni tasarlanmış bir sinir ağı modeli önermişlerdir. Önerilen model, ana semantik dal ve yardımcı sinir dal olmak üzere iki bölümden oluşmaktadır. Ana semantik dal, çok seviyeli özellik haritaları üreterek segmentasyon sonuçlarını tahmin ederken, yardımcı sinir dalı sınır bilgisini işleyerek segmentasyon haritalarının daha net ve keskin olmasını sağlamaktadır. Ayrıca önerilen modelde, doğruluğu ve keskinliği arttırmak için çeşitli modüller de kullanılmıştır. Modelin birçok model ile karşılaştırılması yapılmıştır.

Liu ve arkadaşları [71], kolonoskopik görüntülerin segmentasyonunu iyileştirmek amacıyla kaba-yumuşak segmentasyon çerçevesi içeren bir model önermişlerdir. Önerilen model, temelde üç bileşenden oluşmaktadır. Modelde ilk olarak, polip veri seti Kenar-Anlamsal Normalizasyon Modülü tabanlı, Ablasyon-Tutarlı Segmentasyon Ağı kullanılarak eğitilmekte ve tahminler gerçekleştirilmektedir. Tahmin edilen haritalar, otomatik sınıflandırma mekanizması

yardımıyla basit ve karmaşık örnekler şeklinde ayrılmakta; ardından bu haritalar iyileştirilerek doğruluğu artırılmaktadır.

Wickström ve arkadaşları [72], polip segmentasyon görevi için üç farklı model kullanmışlardır. Bu modellerden biri olan tam evrişimli ağ, görüntülerdeki önemli piksellerin tespiti amacıyla kullanılmıştır. Segmentasyon görevinde ise U-Net ve SegNet mimarileri tercih edilmiştir. Kullanılan SegNet mimarisi simetrik bir kodlayıcı ve kod çözücü yapısına sahiptir.

Zhu ve arkadaşları [73], ise kolon kanserinde poliplerin tespiti ve segmentasyonu için bir model önermişlerdir. Önerilen model 5 katmandan oluşmakta, her katmanda kodlayıcı ve kod çözücüler yer almaktadır. Ayrıca, detaylı ve genel bilgilerin doğru şekilde yakalanabilmesi için Global-Yerel Bağlam Modülü, Uyarlanabilir Seçim Modülü önermişlerdir. Birden fazla veri türünü işleyebilmek için ise Çoklu Modallik Çapraz Dikkat Modülü kullanılmıştır. Önerilen modelin kodu ise açık kaynaklı kod olarak yayınlanmıştır.

Tablo 1-1’de, polip segmentasyonu görevi için CNN mimarilerini kullanan araştırmaların özeti sunulmuştur. Bu tablo, önerilen yöntemler ve bu yöntemlerin kullandığı veri setlerini ele almaktadır. Ayrıca, yöntemin kullanılan her veri setindeki Ortalama Kesişim Bölme Oranı (Mean Intersection over Union- *mIoU*) değerleri de tabloda sunulmuştur.

Tablo 1-1: CNN mimarisi kullanılarak tasarlanmış yöntemlerin özeti

Çalışma (Yıl)	Önerilen Yöntem	Kullanılan Veri Setleri (<i>mIoU</i>)	
[62] (2021)	HarDNet-MSEG	Kvasir- SEG(0.857) CVC- ClinicDB(0.882) CVC-T(0.821)	ETIS-Larib(0.613) CVC-ColonDB(0.660)
[19](2021)	PolypSegNet	<i>Kullanılan veri setlerinde mIoU değeri hesaplanmamıştır.</i>	
[63](2023)	ISCNet	Kvasir-SEG(0.893) CVC-ClinicDB(0.92)	CVC-EndoScene STIL(0.778)
[64](2020)	ACSNet	Kvasir-SEG(0.914) EndoScene(0.8829)	
[65](2022)	FAPN	CVC-ClinicDB(0.892) CVC-T(0.847)	Kvasir(0.865) CVC-ColonDB(0.706) Etis(0.715)
[66](2022)	Derin Sinir Ağı İçin Kullanılacak Yeni Bir	Kvasir-SEG(0.7318) CVC-ClinicDB(0.6166)	ETIS-Larib(0.625)

	Dengesizlik Farkındalıklı Kayıp Fonksiyonu		CVC-ColonDB(0.7003)
[67](2023)	MFEFNet	Kvasir-SEG(0.874) CVC-ClinicDB(0.916) ETIS-Larib(0.686)	CVC-ColonDB(0.682) CVC-300(0.828)
[68](2023)	CFA-Net	CVC-ClinicDB(0.883) Kvasir(0.861)	CVC-ColonDB(0.665) Etis(0.655) CVC-300(0.827)
[69](2022)	DAN-PD	<i>Kullanılan veri setlerinde mIoU değeri hesaplanmamıştır.</i>	
[70](2023)	Yeni bir sinir ağı	CVC-300(0.837) CVC-612(0.893)	
[71](2023)	Aşamalı segmentasyon çerçevesi	Kvasir-SEG(0.8651) CVC-ClinicDB(0.7188)	Etis-Larib(0.5483) CVC-ColonDB(0.6424) CVC-300(0.7898)
[72](2020)	FCN	EndoScene(0.767)	
[73](2023)	CRCNet	<i>Kullanılan veri setlerinde mIoU değeri hesaplanmamıştır.</i>	

1.1.2 Dikkat Modülleri Kullanılarak Tasarlanmış Yöntemler

Dikkat modülleri, CNN mimarilerinin gözden kaçırdığı ya da önemsiz olarak değerlendirdiği nesnelerin segmentasyonunu iyileştirmek; bu nesnelere daha fazla ağırlık atayarak öne çıkarmak amacıyla segmentasyon görevlerinde CNN'lere entegre edilmektedir. Polip segmentasyonu alanında, bu modülleri kullanan bazı literatür yaklaşımları şunlardır:

Alam ve arkadaşları [74], U-Net mimarisine benzer bir model önermişlerdir. Önerdikleri modelde, iki yol üzerinde boşluklu ve boşluksuz (normal) evrişim filtreleri kullanmışlardır. Bu model temelde kodlayıcı ve kod çözücü yapısına dayanmaktadır. Modelin kodu açık kaynaklı kod olarak yayınlanmıştır.

Song ve arkadaşları [75], polip segmentasyon görevini kolaylaştırmak için Res2Net mimarisini temel ağ bir model önermişlerdir. Önerilen modelde görüntüdeki genel desenleri ve yüksek seviyeli özellikleri çıkarmada önemli rol oynayan omurga ağı kullanılmıştır. Ayrıca dikkat mekanizmaları ile polip tespitiindeki performans arttırılmıştır.

Yang ve arkadaşları [76], çeşitli modüllerden oluşan bir model önermişlerdir. Bu modüller sayesinde fazla bağlam özelliği elde etmişler, bağlamsal bilgiyi arttırmışlar ve polip lezyonlarının

sınırlarını daha etkili bir şekilde elde etmişlerdir. Ayrıca modelin kodu açık kaynaklı kod olarak paylaşılmıştır.

Lin ve arkadaşları [77], polip segmentasyonu için bir model geliştirmişlerdir. Modelde omurga kullanarak görüntülerden farklı seviyelerde özellikler çıkarmışlardır. Ayrıca yeni bir modül önererek segmentasyon sürecinde sınırlamaları aşmayı hedeflemişlerdir. Modelde kod çözücü kullanarak özellik çıkarımı yapmışlardır.

Xia ve arkadaşları [78], polip segmentasyonunda Parlaklık Önceliğiyle Yönlendirilen Dikkat Ağı modelini önermişlerdir. Önerdikleri model temelde beş bölümden oluşmaktadır. Model kodlayıcı ile başlamaktadır. Bu model ile poliplerin parlaklık özelliklerini kullanarak daha doğru ve başarılı bir segmentasyon yapmayı amaçlamaktadırlar.

Nguyen ve arkadaşları [79], polip segmentasyonunda başarıyı arttırmak için kademeli bağlam ve dengeleyici dikkat mekanizmasını önermişlerdir. Model iki modülden oluşmaktadır. Doğru bir polip segmente işlemi için, bağlamsal bilgiler kademeli olarak işlenerek doğruluk artırılmaktadır. Dengeleyici dikkat mekanizması ise, modelin farklı özellikler arasındaki ağırlıkları dengeli bir şekilde öğrenmesini mümkün kılmaktadır. Önerilen model, poliplerin kanser teşhisi ve tedavisinde ne kadar önemli olduğunu göstermektedir. Yapılan çalışmanın kodu açık kaynaklı kod olarak paylaşılmıştır.

Yang ve arkadaşları [80], poliplerin tespiti için boyut küçültme işlemi olmadan çalışan yeni bir karıştırma verimli kanal dikkat ağı önermişlerdir. Bu model, Mask R-CNN'nin mimarisine dayanarak polipleri daha iyi bulmak ve segmentasyonunu geliştirmek için tasarlanmıştır.

Tablo 1-2'de, polip segmentasyonu görevi için dikkat modüllerini CNN mimarilerine entegre eden araştırmaların özeti sunulmuştur. Bu tablo, önerilen yöntemler ve bu yöntemlerin kullandığı veri setlerini ele almaktadır. Ayrıca, yöntemin kullanılan her veri setindeki *mIoU* değerleri de tabloda sunulmuştur.

Tablo 1-2: Dikkat modülleri kullanılarak tasarlanmış yöntemlerin özeti

Çalışma (Yıl)	Önerilen Yöntem	Kullanılan Veri Setleri (<i>mIoU</i>)	
[74](2023)	Dikkat Mekanizmalı Derin Sinir Ağı	<i>Kullanılan veri setlerinde mIoU değeri hesaplanmamıştır.</i>	
[75](2022)	Çok Ölçekli Dikkat Çözümlemesi için Paralel Bir Ağ	CVC-ClinicDB (0.888) Kvasir (0.865) Etis (0.679)	CVC-ColonDB (0.690)

[76](2023)	Çapraz Ölçekli Özellik Füzyonu ile Hibrit Dikkat Mekanizmasını Birleştiren Derin Sinir Ağı	Kvasir-SEG (0.907) CVC-ClinicDB (0.941)	Endotect (0.884)
[77](2022)	Bit Dilimleme Bağlamsal Dikkat Ağı	<i>Kullanılan veri setlerinde mIoU değeri hesaplanmamıştır.</i>	
[78](2023)	Ön-Parlaklık Yönlendirmeli Dikkat Ağı	Kvasir -SEG (0.850) CVC – ClinicDB (0.907) Etis (0.819)	CVC-ColonDB (0.962)
[79](2021)	Kademeli Bağlam Ve Dengeleyici Dikkat Mekanizması	<i>Kullanılan veri setlerinde mIoU değeri hesaplanmamıştır.</i>	
[80](2022)	Karıştırmalı Verimli Kanal Dikkat Ağı	<i>Kullanılan veri setlerinde mIoU değeri hesaplanmamıştır.</i>	

1.1.3 Transformatörler Kullanılarak Tasarlanmış Yöntemler

Transformer tabanlı segmentasyon modelleri, görüntü yamasını gömme, konumsal kodlama ve öz-dikkat mekanizmalarıyla küresel ve yerel bağlamı aynı anda modelleyerek çok katmanlı yukarı örnekleme ve atlama-bağlantılarıyla hassas ve esnek segmentasyon sunar. Polip segmentasyonu görevinde, bu modelleri kullanana bazı araştırmalar şöyledir:

Nachmani ve arkadaşları [81], polip segmentasyonunda daha doğru ve daha hızlı ağırlar oluşturmak için ResBlock içeren ResPVT adlı yeni bir model önermişlerdir. Önerilen model, kodlayıcı bileşeni olarak Piramidal Görsel Dönüştürücü tabanlı bir yapı, ayrıca derin özellik öğrenimini desteklemek amacıyla Artık Blok ve üst düzey özelliklerin bütünleştirilmesini sağlayan bir Birleşim Modülü içermektedir. Omurga olarak transformatörleri kullanılmaktadır. Yeni yöntemlerle karşılaştırıldığında, model daha iyi performans göstermiştir.

Lu ve arkadaşları [46], kolonoskopi görüntülerinde polipleri tespit etmek için transformatör tabanlı bir model önerilmişlerdir. Önerilen model, dört modülden oluşmaktadır. Daha kapsamlı bağlamsal bilgi elde etmek amacıyla, Alıcı Alanı Genişleten Çok Ölçekli Küresel Algı Modülü entegre edilmiştir. Sınır belirsizliğini gidermek amacıyla Sınır Rehberli Özellik Geliştirme Modülü, yanlış tespit bölgelerinin düzeltilmesi için Tamamlayıcı Füzyon Modülleri ve mekânsal detay bilgilerini desteklemek üzere Detay İyileştirme Modülü kullanılmıştır. Kullanılan bu modüllerin birlikte çalışması ile yüksek doğrulukta bir segmentasyon ağı önerilmiştir.

Li ve arkadaşları [82], polip segmentasyonundaki tutarsızlık sorununu çözmek amacıyla transformatör ve evrişimsel özellikleri birleştirerek bir model önermişlerdir. Polipleri daha iyi

tespit etmek için özellik çıkarma modülüne sahip Dönüştürücü Tabanlı Kodlayıcı kullanılmaktadır. Aynı zamanda, Bağlamsal Özellik Çıkarma Modülünün ve Çoklu Bilgi Özellik Toplama Modülünün performansı değerlendirilmiştir.

Wu ve arkadaşları [83], tıbbi görüntü segmentasyonu için geliştirilmiş bir model önermişlerdir. Önerilen modelde, piramit kodlayıcı olarak Swin Transformer kullanılmıştır. Ayrıca, çok ölçekli kanal dikkat mekanizması ve basit fakat etkili bir mekansal dikkat mekanizması geliştirilmiştir. Segmentasyonun kenar ve yapı detaylarını aşamalı olarak güçlendirmek için dikkat mekanizması tasarlanmış ve ters dikkat mekanizması ile birleştirilmiştir. Bu model ile daha doğru ve yüksek performanslı segmentasyon sonuçları elde edilmektedir.

Liu ve arkadaşları [84], Yeni bir polip segmentasyon modeli önermişlerdir. Modelde CNN ve transformer kullanılarak, yerel ve küresel bilgiler paralel olarak işlenmiştir. CNN ve Transformer'ı birleştirmenin avantajını verimli bir şekilde kullanabilmek amacıyla Özellik Süper Kod Çözücü tasarlanmıştır. Ayrıca modelde, Polarize Öz Dikkat Modülü kullanılmıştır. Bu modül, gereksiz bilgileri azaltmakta, polip bölgesinde yerel ve küresel ilişkileri güçlendirmektedir.

Hu ve arkadaşları [85], PPNet adında bir model önermişlerdir. Önerilen modelde, polip segmente işleminin performansını arttırmak için Piramit Havuzlama kullanılmıştır. Bu model, Transformer tekniklerini birleştiren bir modeldir. Bellek Tutma Piramit Havuzlama Modülü ve Piramit Özellik Füzyon Modülü, katmanlar arasında bağlamsal özellikleri yakalamak üzere tasarlanmıştır. Ayrıca model, dikkat mekanizmasıyla desteklenerek kod çözücü aşamasında özellik çıkarımını optimize etmektedir.

Wang ve arkadaşları [86], polip segmentasyonunu geliştirmek için CNN ve Transformer mimarilerini birleştiren bir model önermişlerdir. CNN ve Transformer mimarilerinin bir arada kullanılması, poliplerin şekil özelliklerini etkili şekilde iletmesine olanak sağlamıştır. Ayrıca modelde çeşitli yönlerdeki detayları bütün şekilde değerlendirmek amacıyla, Hiyerarşik Piramit Füzyon modülü kullanılmaktadır.

Duc ve arkadaşları [87], polip segmentasyonu için Transformer ve CNN mimarisini birleştiren ColonFormer adında bir model önermişlerdir. Kodlayıcı, kod çözücü ve geliştirme modülleri, tasarlanan modelin temel bileşenleridir. Kodlayıcının omurgası olarak Mix Transformer kullanılmıştır. Kod çözücü için ise Piramit Havuzlama Modülü kullanılmıştır.

Khan ve arkadaşları [88], polip segmentasyonu için bir model geliştirmişlerdir. Önerilen model, polipleri daha doğru bir şekilde tanımlamak için farklı boyut ve çözünürlüklerdeki özellikleri kullanmaktır. Aynı zamanda önerilen model, yoğun bağlantılar ile özellikleri

birleştirmektedir ve eğitilebilir parametreleri optimize ederek daha verimli çalışma imkânı sunmaktadır.

Liu ve arkadaşları polip segmentasyonu için Çok Ölçekli Özellik Sınır Grafi Akıl Yürütme Ağı adında bir model önerilmişlerdir. Önerilen model temelde üç ana bileşenden oluşmaktadır. İlk bileşende farklı ölçeklerden gelen bilgiler birleştirmektedir. İkinci bileşen, bu bilgileri polipleri doğru şekilde segment etmek için birleştirmektedir ve ölçekler arasındaki ilişkileri yakalamaktadır. Son olarak üçüncü bileşen, sınırları daha hassas hale getirmektedir ve segmentasyonu geliştirerek performansı arttırmaktadır [89].

Dong ve arkadaşları tarafından önerilen model polip segmentasyonunu geliştirmek amacıyla geliştirilmiştir. Önerilen model, daha güçlü ve etkili bir transformer kodlayıcı kullanarak poliplerin doğru bir şekilde segmente edilmesini sağlamaktadır. Ayrıca modelde kodlayıcı ve kod çözücü arasındaki iletişim üç ana bileşen ile gerçekleştirilmektedir [90].

Tablo 1-3'te, polip segmentasyonu görevi için Transformer tabanlı mimarileri kullanan araştırmaların özeti sunulmuştur. Bu tablo, önerilen yöntemler ve bu yöntemlerin kullandığı veri setlerini ele almaktadır. Ayrıca, yöntemin kullanılan her veri setindeki *mIoU* değerleri de tabloda sunulmuştur.

Tablo 1-3: Transformatörler kullanılarak tasarlanmış yöntemlerin özeti

Çalışma (Yıl)	Önerilen Yöntem	Kullanılan Veri Setleri (<i>mIoU</i>)	
[81] (2023)	Piramidal Görsel Dönüştürücü ve Artık Bloklarla Polip Segmentasyonu	Kvasir-SEG(0.918) CVC-ClinicDB(0.833) Endotect(0.830)	CVC-ColonDB(0.679) Etis(0.671)
[46](2023)	Sınır Rehberli Çok Ölçekli Algısal Polip Segmentasyon Ağı	Kvasir(0.866) CVC-ColonDB(0.736)	Etis(0.721) CVC-300(0.834)
[82](2022)	Çoklu Bilgi Toplama Ağı	Kvasir-SEG(0.876) CVC-ClinicDB(0.942) CVC-ColonDB(0.739)	Etis(0.725) CVC-300(0.835)
[83](2022)	Çok Ölçekli Mekansal Ters Dikkat Ağı	CVC-ClinicDB(0.879) Kvasir(0.880)	CVC-ColonDB(0.712) Etis (0.698)
[84](2022)	Çift Dalı Çok Ölçekli Özellik Birleştirme Ağı	<i>Kullanılan veri setlerinde mIoU değeri hesaplanmamıştır.</i>	
[85](2023)	Piramit Havuzlama Tabanlı Ağ	EndoScene(0.839) Kvasir-SEG(0.874) Etis (0.716)	CVC-ClinicDB(0.878) CVC-ColonDB(0.726)

[86](2023)	Hiyerarşik Bilgi Yönlendirmeli Birleştirme Ağı	Kvasir – SEG(0.871) CVC–ClinicDB(0.895)	CVC-ColonDB(0.736) CVC-300(0.840)
[87](2022)	ColonFormer	CVC–ClinicDB(0.883) CVC-T(0.826) CVC–ColonDB(0.730)	ETIS–LARİB(0.711) Kvasir(0.877)
[88](2023)	Basit ve Sağlam Derinlik Bazlı Ardışık Ağ	CVC–ClinicDB(0.874)	Kvasir(0.903)
[89](2023)	Çok Ölçekli Özellik Sınır Grafiği Çıkarım Ağı	<i>Kullanılan veri setlerinde mIoU değeri hesaplanmamıştır.</i>	
[90](2021)	Piramit Görsel Dönüştürücüler ile Polip Segmentasyonu	EndoScene(0.833) Kvasir-SEG(0.864) Etis (0.706)	CVC–ClinicDB(0.889) CVC–ColonDB(0.727)

1.2 TEZİN AMACI

Hollanda'nın Utrecht kentinde gerçekleştirilen bir çalışmada, uygulanan kolonoskopi işlemleri sonucunda poliplerin %25'inin ve adenomların %26'sının ilk incelemede tespit edilemediği ortaya konmuştur. Elde edilen bulgular, kolonoskopi sırasında her dört polipten ve her dört adenomdan birinin gözden kaçma olasılığının bulunduğunu göstermektedir [91].

Tezin amacı, kolorektal kanserini erken teşhis etmek için geliştirilen algoritmaların etkinliğini değerlendirmeyi hedeflemektedir. Ayrıca, kalın bağırsak endoskopik işlemlerinin kolon ve rektum kanseri gelişme riskini %50 oranında azalttığı ve bu koruyucu etkinin 6 yıl süresince devam ettiği göz önünde bulundurularak [92], bu prosedürlerin erken tespit ve müdahale açısından önemini vurgulamak tezin hedeflerindendir.

1.3 TEZİN GEREKÇESİ

Kolorektal kanser, dünya çapında en yaygın üçüncü kanser türü olarak belirlenmiştir [5]. Bu kanser türünde erken tanı ve tedavi, hastalığın kontrol altına alınmasında kritik bir rol oynamaktadır. Kolonoskopi ile yapılan polip tespiti, erken tanı açısından hayati öneme sahiptir. Ancak kolonoskopi görüntülerinin uzman hekimler tarafından değerlendirilmesi ve tanı koyulması, zaman alıcı ve yorucu olabilmektedir. Bu nedenle, derin öğrenme tabanlı bilgisayarlı tanı sistemlerine olan ihtiyaç artmaktadır. Literatürde çeşitli yöntemler önerilmiştir. Ancak poliplerin şekil, boyut, konum gibi çeşitlilikleri segmentasyon görevinde performansı sınırlamaktadır. Bu tez kapsamında, alandaki mevcut eksiklikler dikkate alınarak ilgili literatüre bilimsel katkı sağlanması amaçlanmıştır.

1.4 TEZİN ORGANİZASYONU

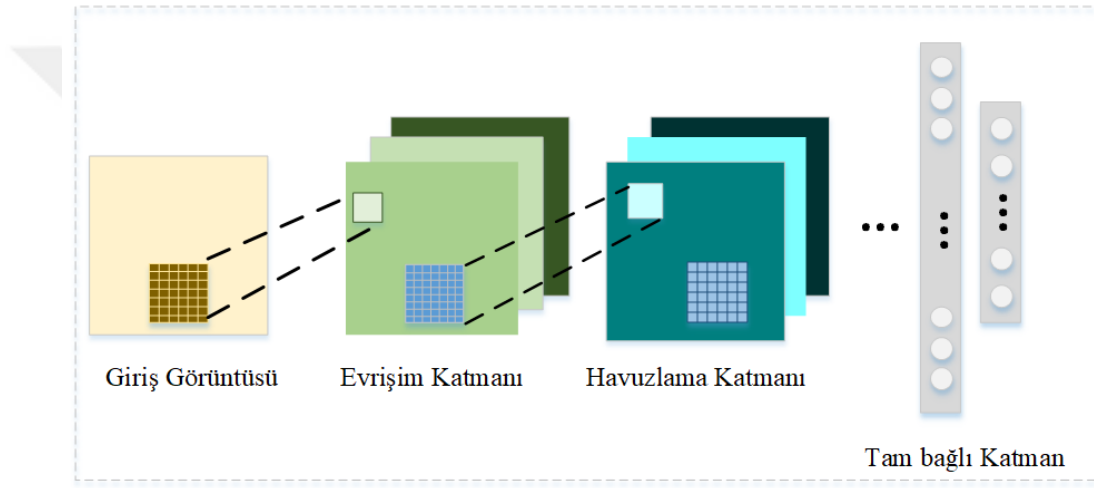
Bu tez, beş ana bölümden oluşmaktadır. Birinci bölümde, çalışmanın amacı, kapsamı ve önemi açıklanmış; ayrıca ilgili literatürde yer alan çalışmalar alt başlıklar altında incelenmiştir. İkinci bölümde, araştırmada kullanılan yöntemler ve teknik ayrıntılar detaylı bir şekilde sunulmuştur. Üçüncü bölümde, çalışmada kullanılan veri setleri tanıtılmış ve performans ölçütleri açıklanmıştır. Dördüncü bölümde ise geliştirilen yöntemlerin uygulama adımları, deneysel sonuçları ve bu sonuçlara ilişkin analizler yer almaktadır. Son olarak beşinci bölümde, sonuçlar tartışılmış ve gelecekte yapılabilecek çalışmalara yönelik öneriler sunulmuştur.



2. MATERYAL VE YÖNTEM

2.1 EVRİŞİMSEL SİNİR AĞLARI

CNN mimarileri, verideki mekânsal ilişkileri koruyan derin öğrenme yapılarıdır. Bu yapılar, görüntülerden anlamlı özellikleri çıkararak sınıflandırma [93], nesne tanıma [94] ile segmentasyon [96] [97] [98] görevlerinde başarılı sonuçlar sunan mimarilerdir. Bir CNN mimarisi genel olarak, evrişim katmanı, havuzlama katmanı ve tam bağlantılı katman olarak bilenen üç katmandan oluşmaktadır. Şekil 2-1’de geleneksel bir CNN mimarisinin genel yapısı gösterilmiştir.



Şekil 2-1: Evrişim katmanı, havuzlama katmanı ve tam bağlı katmanları içeren geleneksel CNN mimarisi

Evrışim Katmanı: Evrişim katmanı, bir girdi görüntüsü üzerinde, maskenin evrişim işlemine tabi tutulmasıdır. Bu katman sayesinde, girdi görüntüsündeki kenar, köşe gibi yerel öznitelikler otomatik olarak tespit edilir. Sonraki katmanlarda, bu özellikler üzerinde daha yüksek seviyeli temsiller oluşturularak sınıflandırma ve tanıma performansı artırılır. Evrişim katmanının da, adım büyüklüğü, maske boyutu, giriş görüntü boyutunun korunup korunmayacağı önemli unsurlardır. Bir evrişim katmanı, girdi matrisi ve çekirdek matrisi kullanılarak tanımlandığında, bu katman aracılığıyla elde edilen çıktı özellik haritası Denklem (2-1)’de sunulan matematiksel ifade ile tanımlanmaktadır.

$$Y_{(i,j)} = \sum_{u=0}^{k_h-1} \sum_{v=0}^{k_w-1} X(is + u - p, js + v - p) K(u, v) \quad (2-1)$$

Burada, u parametresi çekirdeğin yüksekliği boyunca satır indeksini, v parametresi çekirdeğin genişliği boyunca sütun indeksini, k_h parametresi çekirdeğin yüksekliğini, k_w parametresi çekirdeğin genişliğini, is parametresi dikey yönde yapılan adım büyüklüğünü, js parametresi yatay yönde yapılan adım büyüklüğünü ifade etmektedir. K parametresi ise evrişim işleminde kullanılan çekirdeğin (filtrenin) kendisini, yani ağırlık matrisini temsil etmektedir. Y parametresi çıkış özelliği haritasının elemanını, p parametresi kullanılan sıfır dolgusun her bir kenarda toplamda kaç satır/sütun eklendiğini, s parametresi ise kaydırma adımını ifade etmektedir.

Ağın öğrenme kapasitesi ve çıktı boyutu, evrişim katmanlarında kullanılan çeşitli evrişim türlerinden doğrudan etkilenmektedir. Bu durumda, tercih edilen evrişim türlerinden biri Geçerli (Valid) Evrişimdir. Bu işlem, hiç dolgu eklenmeden yapılan bir evrişim türüdür. Geçerli Evrişim, yalnızca giriş görüntüsünün sınırları içerisinde kalan bölgeler üzerinde işlem gerçekleştirir. Bu işlem, matematiksel olarak Denklem (2-2) ile ifade edilmektedir.

$$Y_{(i,j)} = \sum_{u=0}^{k_h-1} \sum_{v=0}^{k_w-1} X(i+u-p, j+v) K(u, v), \quad 0 \leq i \leq H-k_i, 0 \leq j \leq W-k_w \quad (2-2)$$

Burada, H parametresi giriş görüntüsünün yüksekliğini, W parametresi genişliğini temsil etmekte ve k_i parametresi ise çekirdek matrisinin yüksekliğini ifade etmektedir. Burada, sıfır dolgusu $p=0$, adım değeri s olarak gösterilmektedir. Bu koşullar altında çıktı boyutlarının hesaplanmasında kullanılan matematiksel ifade Denklem (2-3)'te sunulmaktadır.

$$H_{çıkıtlı} = \left\lceil \frac{H-k_h}{s} \right\rceil + 1, W_{çıkıtlı} = \left\lceil \frac{W-k_w}{s} \right\rceil + 1 \quad (2-3)$$

Bazı çalışmalarda, evrişim işlemi sonrasında elde edilen çıktı boyutunun, giriş verisinin boyutlarıyla birebir korunması hedeflenmektedir. Bu gereksinim doğrultusunda, Aynı Boyutlu (Same) Evrişim olarak adlandırılan bir diğer evrişim türü tercih edilmektedir. Adım büyüklüğü, evrişim çekirdeğinin giriş verisi üzerinde her işlemde kaç birimlik kayma yapacağını belirleyen temel bir parametredir. Bu değer, çekirdeğin giriş üzerinde hangi aralıklarla konumlanacağını tanımlar ve çıktı boyutunun belirlenmesinde doğrudan rol oynamaktadır. Çekirdeğin yüksekliği k_h ve çekirdeğin genişliği k_w değerlerinin tek sayı olması durumunda, ilgili matematiksel ifade Denklem (2-4)'te tanımlanmaktadır.

$$p_h = \frac{k_h-1}{2}, p_w = \frac{k_w-1}{2} \quad (2-4)$$

Bu denklemde, eğer çıktı boyutları 1 olacak şekilde ayarlanmışsa, ilgili durum Denklem (2-5) ile ifade edilmektedir.

$$H_{çıkıtı} = H, W_{çıkıtı} = W \quad (2-5)$$

Aynı Boyutlu Evrişimin genel matematiksel ifadesi aşağıda Denklem (2-6) ile gösterilmektedir.

$$Y(i,j) = \sum_{u=0}^{k_h-1} \sum_{v=0}^{k_w-1} X(i+u-p_h, j+v-p_w)K(u,v), \quad 0 \leq i < H, 0 \leq j < W \quad (2-6)$$

Burada, çekirdek K , giriş görüntüsü X , (i,j) konumundaki çıkış piksel değeri $Y(i,j)$, dolgu değerleri p_h ve p_w , çıkış pikselleri i,j ve çekirdek koordinatları u,v olmak üzere ifade edilmiştir.

Örneğin, boyutları 4×4 olan bir giriş matrisi ile 3×3 boyutunda bir filtre ve kaydırma adımı 1 olarak belirlendiğinde gerçekleştirilen evrişim hesaplama işlemi şöyledir:

$$\text{Adım 1: } H_{out} = 4 - 3 + 1 = 2, W_{out} = 4 - 3 + 1 = 2$$

$$\text{Adım 2: } Y_{valid}(i,j) = \sum_{u=0}^2 \sum_{v=0}^2 X(i+u, j+v)K(u,v), \quad i,j \in \{0,1\}$$

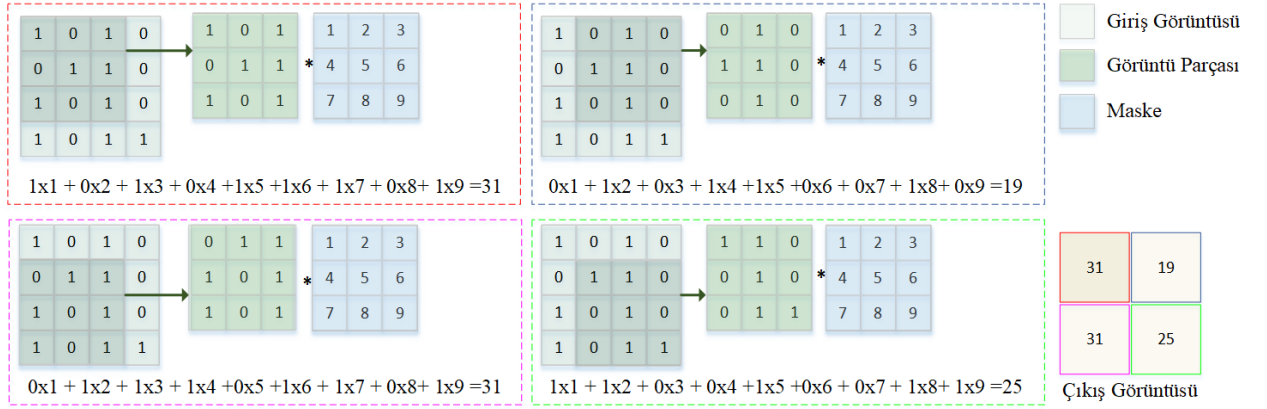
$$\text{Adım 3: } p_h = p_w = \frac{3-1}{2} = 1, H_{out} = 4, W_{out} = 4$$

$$\text{Adım 4: } Y_{same}(i,j) = \sum_{u=0}^2 \sum_{v=0}^2 X(i+u-1, j+v-1)K(u,v), \quad i,j \in \{0,1,2,3\}$$

İlk adımda, dolgusuz evrişim (geçerli dolgu) kullanılarak çıktı boyutları hesaplanmıştır. İkinci adımda, dolgusuz evrişim durumunda çıkış değerleri hesaplanmıştır. Üçüncü adımda, giriş ve çıkış boyutlarının aynı kalmasını sağlayan dolgu miktarı ve çıktı boyutları belirlenmiştir. Dördüncü adımda ise aynı dolgu uygulanarak evrişim işlemi gerçekleştirilmiştir.

Şekil 2-2’te, 4×4 boyutlarındaki bir giriş görüntüsü üzerinde kayan pencere mantığıyla uygulanan 3×3 boyutundaki bir maske ile gerçekleştirilen evrişim işlemi gösterilmektedir. Bu işlemin her adımında maske, giriş matrisinin geçerli 3×3 alt-bölgesiyle noktasal çarpılıp toplanarak tek bir skaler değer (örneğin sol üst alt-bölge için 31) elde etmekte ve bu değer çıktı matrisinin ilgili hücresine (2×2 boyutlu olacak şekilde) yerleştirilmektedir. Böylece, maske kayan pencere mantığıyla dörder konumda (iki yatay iki dikey) kaydırma yapılarak girişteki mekânsal

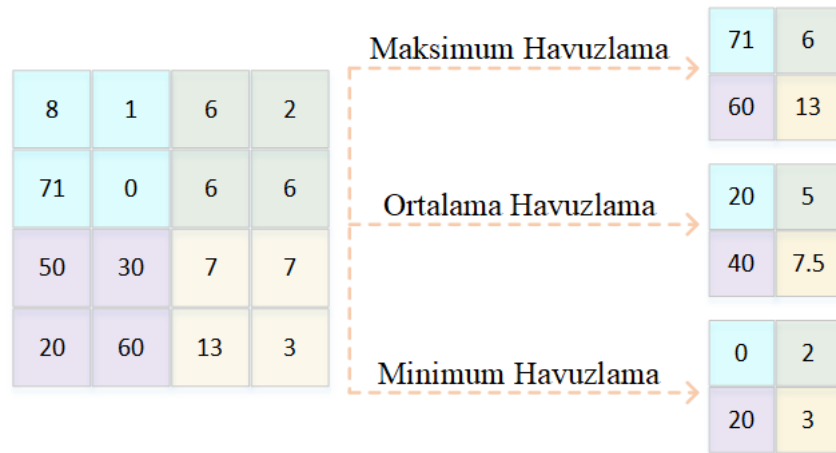
ilişkileri koruyan dört skaler değer (31, 19, 31, 25) üretmek suretiyle 2×2 boyutlu bir özellik haritası oluşturulur.



Şekil 2-2: 4x4 boyutundaki bir girdi görüntüsü üzerinde 3x3 bir maske ile gerçekleştirilen evrişim işlemi

Havuzlama katmanı: Evrişim katmanını takiben gelen Havuzlama katmanı, Evrişim katmanından elde edilen özniteliklerin boyutunu küçültür, sonraki katmanlarda işlenecek veri miktarını azaltmayı, aşırı uyumu engellemeyi ve eğitim süresini kısaltmayı amaçlamaktadır [98]. Bu iki katman, girdi verisinden önemli özellikleri çıkarma ve bu özellikleri soyut bir şekilde gösterme sürecini içermektedir.

Havuzlama katmanlarının; maksimum havuzlama (maksimum değer alma), ortalama havuzlama (ortalama alma) ve minimum havuzlama (en küçük değeri alma) gibi çeşitleri vardır. Şekil 2-3'de 4x4 boyutundaki bir matris üzerinde üç farklı havuzlama (pooling) işlemi gösterilmektedir.

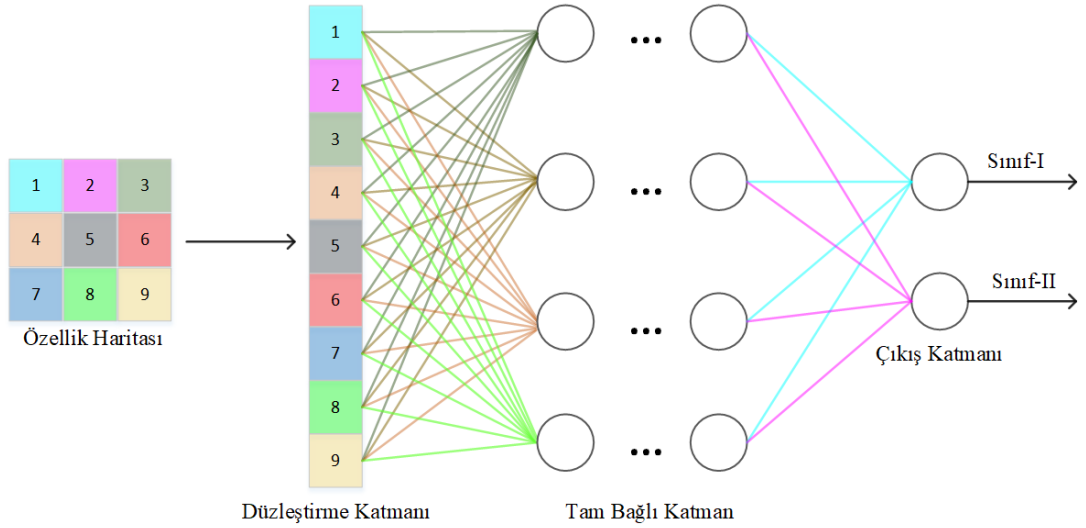


Şekil 2-3: 4x4 boyutundaki bir girdi görüntüsü üzerinde 3x3 bir maske ile gerçekleştirilen evrişim işlemi

Şekil 2-3’de, 4×4 boyutlu bir giriş matrisi üzerinde eş boyutlu (2×2) dört farklı alt bölgenin tanımlanması ve her bir bölgeden “maksimum”, “ortalama” ile “minimum” değerlerin seçilmesi yoluyla üç ayrı havuzlama işleminin sonuçlarının nasıl elde edildiği gösterilmektedir. Öncelikle, sol üst 2×2 bölgedeki değerler (8, 1; 71, 0) için maksimum havuzlama çıktısı 71, ortalama havuzlama çıktısı $(8 + 1 + 71 + 0)/4 = 20$ ve minimum havuzlama çıktısı 0’dır. Sağ üst 2×2 bölgedeki değerler (6, 2; 6, 6) için havuzlama sonuçları sırasıyla 6, $(6 + 2 + 6 + 6)/4 = 5$ ve 2’dir. Sol alt 2×2 bölgedeki değerler (50, 30; 20, 60) için sırasıyla 60, $(50 + 30 + 20 + 60)/4 = 40$ ve 20; son olarak sağ alt 2×2 bölgedeki değerler (7, 7; 13, 3) için sırasıyla 7, $(7 + 7 + 13 + 3)/4 = 7,5$ ve 3 olarak elde edilmiştir. Bu işlemler sonucunda, her bir havuzlama türü kendi skaler değerlerini bir araya getirerek 2×2 boyutlu ayrı bir çıktı matrisi oluşturmaktadır. Maksimum havuzlama yöntemi, baskın ve belirgin özelliklerin vurgulanmasında etkili olurken, ortalama havuzlama yöntemi, verinin genel yapısını dengeli ve tutarlı bir biçimde temsil etmektedir. Minimum havuzlama yöntemi ise, düşük değerli bölgelerin ön plana çıkarılmasına olanak tanımaktadır.

Tam Bağlı Katman: CNN mimarilerinde kullanılan diğer bir katman ise tam bağlı katmandır. Tam bağlı katmanın genel yapısı Şekil 2-4’de sunulmuştur. Tam bağlı katman, evrişim katmanı ve havuzlama katmanından elde edilen özelliklerin sınıflandırma, tahmin etme, karar süreçlerinde kullanılmak üzere yapılandırılmasını ve değerlendirilmesini gerçekleştirir. Bu katmanda, gelen özellik haritaları tek boyutlu vektöre dönüştürülür. Böylece, farklı boyutlardaki özellikler sabit bir boyuta getirilirken, tam bağlı katmanlarda işlem yapılması kolaylaşmaktadır.

Şekil 2-4’de, bir CNN tarafından elde edilen 3×3 boyutlu bir özellik haritasının düzleştirme katmanına geçirilmesiyle üretilen tek boyutlu vektörün tam bağlı katmana aktarılması ve buradan da sınıflandırıcı çıkış katmanına ulaşan ileri beslemeli süreç gösterilmektedir. İlk olarak, özellik haritasındaki her bir piksel değeri, evrişim ve aktivasyon işlemlerini takiben tek boyutlu bir dizi halinde düzenlenir. Bu işlem düzleştirme katmanı olarak adlandırılır. Düzleştirilmiş vektör, tam bağlı katmanda her bir nöronun tüm giriş ağırlıklarıyla bağlantı kurması sayesinde lineer dönüşümler ve isteğe bağlı aktivasyon fonksiyonları aracılığıyla daha anlamlı ve karmaşık öznitelikler üretir. Son aşamada ise bu katmandan elde edilen çıkış değerleri, ikili sınıfı (“Sınıf-I” ve “Sınıf-II”) tanımlayan softmax veya benzeri bir aktivasyon fonksiyonuna sahip çıkış katmanına beslenerek olasılık dağılımı biçiminde sınıflandırma sonuçları elde edilir.



Şekil 2-4: Tam bağlı katman

2.2 KAYIP FONKSİYONLARI

Kayıp fonksiyonu, modelin tahmin ettiği sonuçlar ile gerçek gözlemler arasındaki sapmayı hesaplar [99]. Böylece, modelin optimizasyon sürecinde en düşük sapma seviyesine ulaşmaya odaklanmasını sağlar. Burada, modelin optimizasyon süreci, öğrenme algoritmasının amacı olan kayıp fonksiyonunu en aza indirecek şekilde model parametrelerini güncelleme döngüsünü ifade eder. İyi tasarlanmış bir kayıp fonksiyonu, modelin yalnızca doğruluk odaklı optimizasyonundan öte, eğitim sürecini verilerdeki anlamlı örüntüleri öğrenmeye yönlendirdiği için derin öğrenmede kritik bir rol oynar [100].

Derin öğrenme modelleri genellikle ön tanımlı kayıp fonksiyonlarını kullanırken, güncel araştırmalar [101] [102] [103] sınıf dengesizliği, aykırı değerler ve göreve özgü performans kriterlerini ele almak üzere farklı kayıp fonksiyonlarının yenilikçi kombinasyonlarını tanıtmaktadır. Bu bağlamda, ideal bir kayıp fonksiyonu modelin hem hataları doğru ve duyarlı bir şekilde ölçmesine hem de göreve özgü performans hedeflerini dengeleyerek optimize etmesine izin veren, sınıf dengesizliği ve aykırı değerlere karşı dayanıklı esnek bir kriter olarak söylenebilir.

Poliplerin değişken şekil ve konumları ile segmentasyon zorlukları göz önüne alındığında, ideal bir kayıp fonksiyonu net sınır belirleme ve sağlam özellik çıkarımını ön plana çıkararak modelin duyarlılığını artırır. Bu sayede tanısal doğruluk yükselir ve tedavi planlaması daha etkili hâle gelebilir. Tez kapsamında, dört farklı kayıp fonksiyonunun polip bölütleme görevindeki performansı incelenmiş ve bu fonksiyonların matematiksel ifadesi bu bölümde verilmiştir.

İkili Çapraz Entropi Kayıp Fonksiyonu: İkili Çapraz Entropi (Binary Cross Entropy - BCE) [104] kayıp fonksiyonu, birden fazla etiketi aynı anda tahmin edebilme yeteneğine sahip, bir görüntü üzerinde belirli özelliklerin tespitinde yaygın olarak kullanılan bir kayıp fonksiyonudur [105]. Bu fonksiyon, her bir örneğin pozitif veya negatif olma olasılığını tahmin ederken, gerçek etiketler ile modelin tahminleri arasındaki farkı ölçmektedir. BCE kayıp fonksiyonunun [104] matematiksel ifadesi Denklem (2-7)'de verildiği gibidir.

$$L_{BCE} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (2-7)$$

Burada, N parametresi gözlem sayısını, y_i parametresi ise her i 'inci örnek için gerçek sınıf etiketini belirler ve burada $y_i \in \{0,1\}$ 'dir. p_i parametresi ise modelin her i 'inci sınıf için tahmin ettiği olasılığı belirler ve burada $p_i \in (0,1)$ 'dir. Örneğin, $N=3$, $y = [1, 0, 1]$, ve $p = [0.9, 0.2, 0.7]$ olsun. Bu durumda, toplam kayıp terimi ve ortalama BCE kaybı şöyle hesaplanır;

$$y_1 \log(p_1) + (1 - y_1) \log(1 - p_1) = 1 \cdot \log(0.9) + 0 \cdot \log(0.1) \approx -0.1054$$

$$y_2 \log(p_2) + (1 - y_2) \log(1 - p_2) = 0 \cdot \log(0.2) + 1 \cdot \log(0.8) \approx -0.2231$$

$$y_3 \log(p_3) + (1 - y_3) \log(1 - p_3) = 1 \cdot \log(0.7) + 0 \cdot \log(0.3) \approx -0.3567$$

$$\sum_{i=1}^3 [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \approx -0.6852$$

$$L_{BCE} = -\frac{1}{3} \times (-0.6852) \approx 0.2284$$

Tversky Kayıp Fonksiyonu: Tversky Kayıp Fonksiyonu [106], temelini 1977 yılında geliştirilen Tversky benzerlik ölçütünden almaktadır [107]. Bu kayıp fonksiyonu, dengesiz veri bölütleme görevini gerçekleştirmek için geliştirilmiş bir fonksiyondur. Tversky Kayıp Fonksiyonu [106], yanlış pozitif ve yanlış negatif hataları ayrı ağırlıklarla ele alabilme yetisi sayesinde, modelin özellikle kritik sınıflandırma hatalarını minimize etmeye odaklanan daha hassas bir öğrenme süreci geliştirmesine imkân tanır. Bu özellik, veri kümesinde az temsil edilen veya küçük boyutlu nesnelerin segmentasyonunda kayda değer performans artışı sağlar. Tversky Kayıp Fonksiyonunun matematiksel ifadesi Denklem (2-8)'de ölçüldüğü gibidir.

$$L_{Tversky(S,G)} = \frac{|S \cap G|}{|S \cap G| + \lambda_1 |S - G| + \lambda_2 |G - S|} \quad (2-8)$$

Burada, S parametresi tahmin edilen segmentasyon maskesini, G parametresi ise gerçek segmentasyon maskesini temsil etmektedir. λ_1 ve λ_2 parametreleri, sırasıyla yanlış pozitif ve yanlış negatif sınıflandırmaların kayıp fonksiyonundaki ağırlıklı katkılarını dengelemek amacıyla tanımlanan katsayılardır. $S \cap G$ tahmin edilen ve gerçek segmentasyon maskeleri arasındaki ortak bölgeyi $S - G$ ve $G - S$ ifadeleri ise bu maskeler arasındaki örtüşmeyen alanları temsil etmektedir.

Örneğin, $|S \cap G| = 4$, $|S \setminus G| = 1$ ve $|G \setminus S| = 2$ olarak ölçülen değerler ile $\lambda_1 = 0.5$, $\lambda_2 = 1.0$ parametre değerlerine sahip olsun. Bu durumda, Tversky Kayıp Fonksiyonu şöyle hesaplanır;

$$L_{Tversky} = \frac{4}{4 + 0.5 \times 1 + 1.0 \times 2} \approx 0.6154$$

Dice Kayıp Fonksiyonu: Dice Kayıp Fonksiyonu [108]'nın temelini oluşturan Dice Katsayısı, ilk olarak 1945 yılında tanıtılmıştır [109]. Özellikle segmentasyon görevleri ve tıbbi görüntüleme alanlarında yüksek performans sergileyen Dice Kayıp Fonksiyonu [108], model tarafından tahmin edilen segmentasyon maskesi ile gerçek segmentasyon maskesi arasındaki benzerliği ölçmektedir. Bu kayıp fonksiyonu, matematiksel olarak Denklem (2-9)'da tanımlanmaktadır.

$$L_{Dice} = 1 - \frac{2|S \cap G|}{|S| + |G|} \quad (2-9)$$

Burada S ve G parametreleri, sırasıyla tahmin edilen ve gerçek segmentasyon maskelerini ifade etmektedir. Dice Kayıp Fonksiyonu [108] değeri sıfıra yaklaştıkça modelin tahmini ile gerçek segmentasyon maskesi arasındaki benzerlik artmaktadır.

Örneğin, tahmin edilen segmentasyon maskesi S ile gerçek segmentasyon maskesi G arasında örtüşen doğru pozitif piksel sayısı $|S \cap G| = 8$, tahmin edilen toplam pozitif piksel sayısı $|S| = 12$ ve gerçek pozitif piksel sayısı $|G| = 7$ parametre değerlerine sahip olsun. Bu durumda Dice Kayıp Fonksiyonu [108], şöyle hesaplanır;

$$L_{Dice} = 1 - \frac{2 \times 8}{12 + 7} \approx 0.158$$

Bileşik Kayıp Fonksiyonları: Bileşik Kayıp Fonksiyonu (Compound), özellikle tıbbi görüntü segmentasyonu, nesne tespiti gibi farklı birçok alanda kullanılmaktadır. Bu kayıp fonksiyonu, modelin tahminleri ile gerçek etiketler arasındaki uyumsuzluğu ölçmektedir. Veri setlerinde sınıf dengesizliği ve küçük nesnelerin tespiti gibi zorlukların bulunduğu durumlarda, standart kayıp

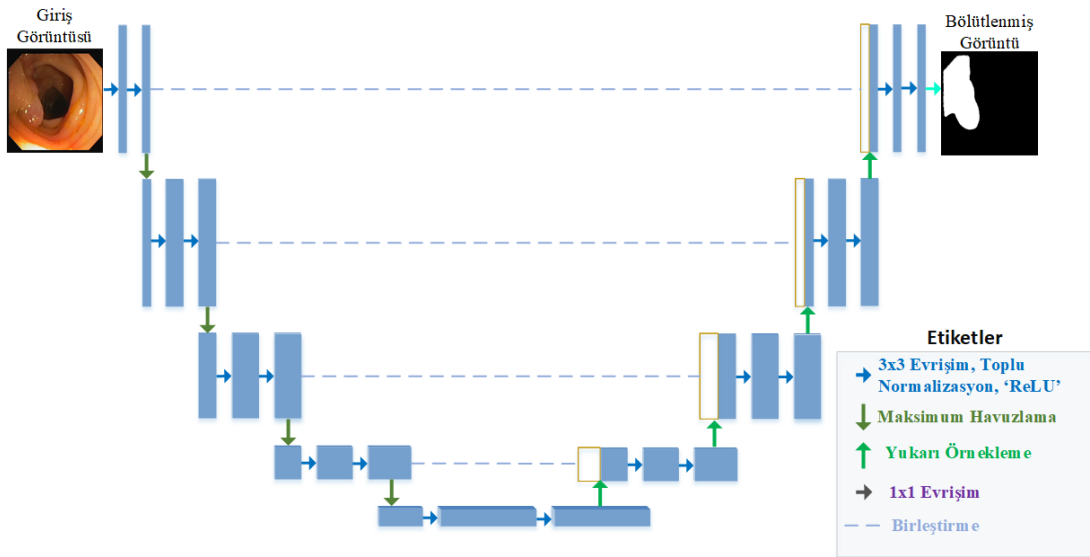
fonksiyonlarının yetersiz kalması sebebiyle, Bileşik Kayıp Fonksiyonları tercih edilmektedir. Bileşik Kayıp Fonksiyonunun matematiksel ifadesi Denklem (2-10)'da verildiği gibidir.

$$L_{Compound} = \alpha L_{BCE} + \beta L_{Dice} \quad (2-10)$$

Burada, L_{BCE} İkili Çapraz Entropi Kayıp Fonksiyonu, L_{Dice} Dice Kayıp Fonksiyonunu ifade etmektedir. α ve β parametreleri ise bu iki bileşene atanan ağırlıkları kontrol etmektedir.

2.3 U-Net MİMARİSİ

U-Net mimarisi [22], tam evrişimli bir sinir ağı olarak Ronneberger ve arkadaşları tarafından 2015 yılında geliştirilmiş bir mimaridir. Bu mimari, piksel düzeyinde hassas sınıflandırma gerektiren görevlerde yüksek doğruluk sağlamaktadır [22]. U-Net mimarisi [22], adını aldığı "U" harf biçimindeki simetrik yapısı ile kodlayıcı ve kod çözücü kısımlarından oluşur. U-Net mimarisinin [22] genel yapısı Şekil 2-5'de gösterilmiştir.



Şekil 2-5: U-Net mimarisinin genel yapısı

X , ağı besleyen giriş görüntüsü olmak üzere, $X \in R^{H \times W \times C}$ 'dir. Burada H ve W parametreleri sırası ile yüksekliği ve genişliği temsil etmektedir. C parametresi ise kanal sayısını ifade etmektedir. Her bir katman, Denklem (2-11)'de verildiği gibi modellenmektedir.

$$f_0 = \sigma (W^{(0)} + X + b^{(0)}) \quad (2-11)$$

Burada, f_0 parametresi ilk katmandan çıkan özellik haritasını; $W^{(0)}$ parametresi filtre ağırlıklarını; $b^{(0)}$ parametresi filtreye eklenen bias terimini; X parametresi giriş verisini ve σ parametresi aktivasyon fonksiyonunu göstermektedir.

U-Net mimarisinde [22], aktivasyon fonksiyonu olarak *Sigmoid* veya *Softmax* aktivasyon fonksiyonları kullanılmaktadır. Bu aktivasyon fonksiyonlarının matematiksel ifadeleri Denklem (2-12) ve Denklem (2-13)'de sunulmaktadır.

$$Sigmoid = \frac{e^x}{1+e^x} \quad (2-12)$$

$$Softmax = a_x + \frac{e^{z_i}}{\sum_{i=0}^m e^{z_i}} \quad (2-13)$$

Burada sunulan Denklem (2-12)'de, x parametresi, genellikle bir nöronun ağırlıklı toplamını temsil ederken; e parametresi, *Euler Sabiti* olup yaklaşık olarak 2,718 değerine sahiptir. *Sigmoid* aktivasyon fonksiyonu, özellikle ikili sınıflandırma ile ilgili sorunlarda bir pikselin belirli bir sınıfa ait olma olasılığını tahmin etmek için kullanılmaktadır.

Softmax aktivasyon fonksiyonunun matematik ifadesi olan Denklem (2-13)'de ise, z_i parametresi, i . sınıfa ait ham skor değerini; m parametresi, toplam sınıf sayısını; e^{z_i} parametresi, i . sınıf için üstel dönüşümü ve paydadaki toplam ise, tüm sınıfların üstel skorlarının toplamını ifade etmektedir. Bu matematiksel ifade ile her sınıf için 0'dan 1'e kadar değişen bir olasılık değeri elde edilmektedir.

U-Net mimarisinde [22], giriş katmanını takiben her evrişim katmanı, giriş verisinden anlamlı özellikler çıkarmak amacıyla ardışık evrişim ve aktivasyon işlemleri gerçekleştirmektedir. Denklem (2-14)'teki genel ifade ile modellenmektedir.

$$f^{(l)} = \sigma (W^{(l)} * f^{(l-1)} + b^{(l)}) \quad (2-14)$$

Burada, $f^{(l-1)}$ parametresi bir önceki katmandan elde edilen özellik haritasını; $W^{(l)}$ parametresi l katmanına ait evrişim filtrelerin ağırlıklarını; $b^{(l)}$ parametresi aynı katmana ait bias değerini; $*$ parametresi evrişim işlemini; σ parametresi aktivasyon fonksiyonunu temsil etmektedir.

Kodlayıcı: Simetrik yapısı ile tanınan U-Net mimarisinin [22] kodlayıcı katmanı, giriş görüntüsündeki temel renk, doku ve şekil gibi düşük çözünürlüklü özellik haritalarını anlamlı bir forma dönüştürmektedir. Kodlayıcı katmanının matematiksel ifadesi Denklem (2-15)'te sunulmaktadır.

$$f^{(l)} = \text{MaxPool}(\sigma(W^{(l)} * f^{(l-1)} + b^{(l)})) \quad (2-15)$$

Burada, $f^{(l-1)}$ parametresi bir önceki katmandan elde edilen özellik haritasını; $W^{(l)}$ parametresi l katmanına ait evrişim filtrelerin ağırlıklarını; $b^{(l)}$ parametresi aynı katmana ait bias değerini; $*$ parametresi evrişim işlemini; σ parametresi aktivasyon fonksiyonunu; MaxPool parametresi ise maksimum havuzlama işlemini temsil etmektedir.

Kod Çözücü: Bir diğer katman olan kod çözücü, kodlayıcı tarafından dönüştürülen formları özgün çözünürlüğe yeniden çıkararak detaylı segmentasyon maskesi üretmektedir. Kod çözücü katmanının, genel matematiksel ifadesi Denklem (2-16)'da ifade edilmektedir.

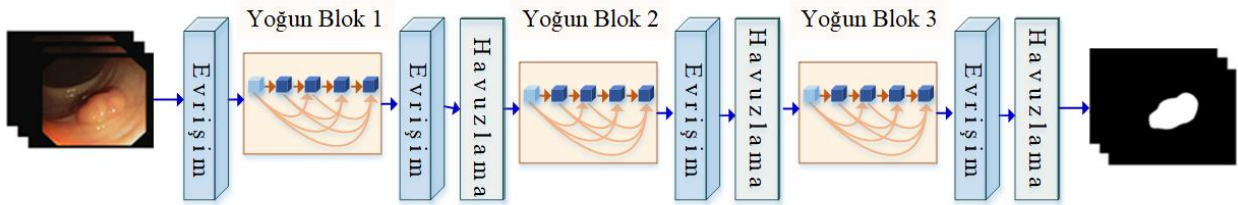
$$f^{(l)} = \sigma(W^{(l)} * [Up(f^{(l+1)}) \oplus f_{enc}^{(l)}] + b^{(l)}) \quad (2-16)$$

$Up(f^{(l+1)})$ parametresi bir kod çözücünün daha derin katmanından gelen düşük çözünürlüklü özellik haritasını; $f_{enc}^{(l)}$ parametresi kodlayıcının l . katmanından gelen özellik haritasını; $\oplus$ parametresi kanal ekseninde birleştirme işlemini; $W^{(l)}$ parametresi l katmanına ait evrişim filtrelerin ağırlıklarını; σ parametresi aktivasyon fonksiyonunu; $b^{(l)}$ parametresi bias terimini; $f^{(l)}$ parametresi kod çözücünün l katmanından çıkan özellik haritasını temsil etmektedir.

2.4 DenseNet MİMARİSİ

Yoğun Bağlantılı Evrişimsel Ağlar (Densely Connected Convolutional Networks- DenseNet), 2017 yılında Huang vd. tarafından tanıtılmıştır [110]. DenseNet mimarisi [110], katmanları birbirine bağlı olan CNN yapısına sahiptir. Ayrıca, ağ parametrelerinin sayısını azaltarak özelliklerin yeniden kullanımını mümkün kılmaktadır.

DenseNet mimarisi [110], yan yana dizilmiş yoğun blok ve aralarındaki geçiş bloklarından oluşmaktadır. Bir yoğun blokta, çıktılar her katmandan doğrudan diğer katmanlara aktarılmaktadır [110]. Bu yapı, özelliklerin yeniden kullanılmasına, modelin parametreleri daha verimli hale getirmesine olanak tanımaktadır. Daha az parametre ile derin ağlarda yüksek performans elde edilen DenseNet mimarisi [110] yapısı Şekil 2-6’da sunulmuştur.



Şekil 2-6: DenseNet mimarisi

Şekilde 2-6’da sunulan DenseNet mimarisinin [110], giriş görüntüsü CVC-ClinincDB [20] veri setinden alınan bir kolonoskopik görüntüdür. Görüntü, üç adet yoğun bloktan geçirilerek işlenmektedir. Her yoğun blok, dört evrilim katmanından oluşmaktadır. Bu katmanlar, kendisinden önceki tüm katmanların çıktılarını girdi olarak almaktadır. Yoğun bloklar arasında yer alan evrilim katmanları ve havuzlama katmanları, görüntüdeki önemli bilgileri koruyarak boyut indirgeme işlemini gerçekleştirmektedir. Şekil 2-6’daki yapı, bu işlem sırasını ve katmanlar arasındaki bilgi akışı gösterilmektedir.

DenseNet mimarisinde [110], her katman kendisinden önceki tüm katmanların çıktısını girdi olarak almaktadır. Bu durumun matematiksel ifadesi Denklem (2-17)’de gösterilmektedir.

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (2-17)$$

Burada, x_l parametresi l katmanının çıktısını; $H_l(.)$ parametresi bu katmanda uygulanan işlemler dizisini temsil etmektedir.

Her Yoğun Blok içerisinde, katmanlar arasında yoğun bağlantılar kurularak her bir katmanın yeni özellik haritaları üretilmektedir. Çıkış kanal sayısı, giriş kanal sayısına, blok içerisindeki toplam katman sayısı ile büyüme oranının çarpımı eklenerek hesaplanır. Yoğun bloğun çıkış kanal sayısı, Denklem (2-18)’de verilen matematiksel ifade ile hesaplanmaktadır.

$$C_{çıkış} = C_0 + k \cdot L \quad (2-18)$$

Burada, $C_{çıkış}$ parametresi yoğun bloktan çıkan toplam kanal sayısını; C_0 parametresi blok girişindeki kanal sayısını; k her katmanın ürettiği yeni kanal sayısını; L ise blok içerisindeki katman sayısını ifade etmektedir.

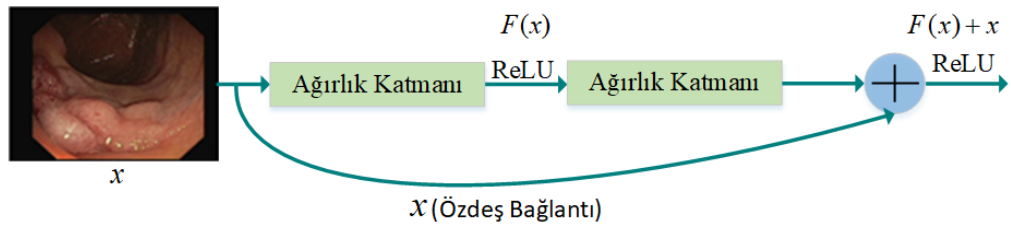
Her yoğun blok arasında boyut indirgeme işlemi gerçekleştirilmektedir. Bu indirgeme işlemi, hesaplama maliyetini azaltmakta, modelin parametre sayısını kontrol altında tutmak için uygulanmaktadır. Matematiksel ifadesi Denklem (2-19)'da sunulmaktadır.

$$x' = AvgPool (Conv_{1 \times 1} (BN(x))) \quad (2-19)$$

Burada, x parametresi önceki yoğun bloktan gelen çıkışı; $BN(x)$ parametresi toplu normalizasyon işlemi; $Conv_{1 \times 1}$ parametresi 1x1 boyutunda evrişim işlemi; $AvgPool$ parametresi ortalama havuzlama işlemi; x' parametresi ise boyutu küçültülmüş yeni özellik haritasını temsil etmektedir.

2.5 ResNet50 MİMARİSİ

Artık Ağ (Residual Network - ResNet), kaybolan gradyan sorununu çözmek için geliştirilmiş bir derin öğrenme mimarisidir. Bu mimarinin, katman sayıları değişiklik göstermektedir [111]. Katman sayısına göre öne çıkan versiyonlarından biri olan ResNet50 mimarisi [111], 50 katmandan oluşmakta ve bu katmanlar artık bloklar halinde gruplanmaktadır. Gruplanmış artık bloklar, giriş verisinin birden fazla katman aracılığıyla işlenmesini sağlayan özel yapılar olup, bu yapıların genel görünümü Şekil 2-7'de sunulmaktadır.



Şekil 2-7: Artık blok

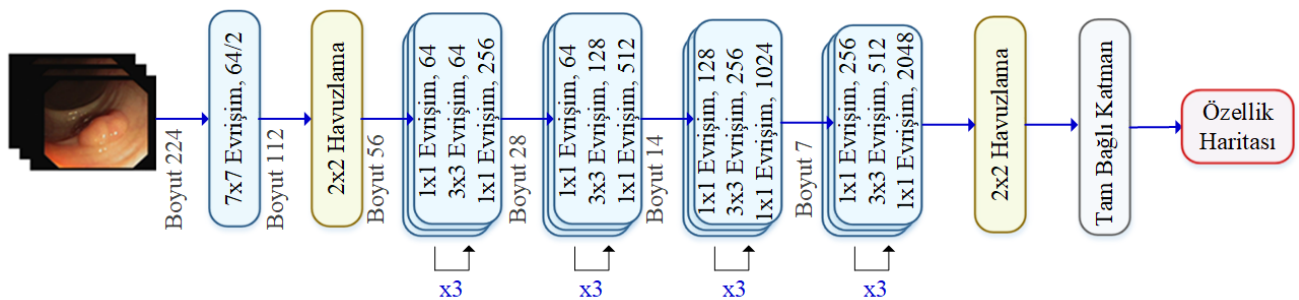
Artık blok yapısı, öğrenme sürecini kolaylaştırmakta ve derin ağlarda karşılaşılan kaybolan gradyan problemini azaltmaktadır. Kaybolan gradyan problemi, eğitim aşamasında karşılaşılan bir problemdir. Derin öğrenme modellerinde, model ilk olarak giriş verilerini kullanarak tahminler üretilmektedir. Ardından, modelin tahmini ile gerçek değer arasındaki hata hesaplanmakta; hesaplanan hata geri yayılım yöntemiyle ağırlık önceli katmanlarına doğru iletilmektedir. Bu süreç, her katmanın çıktısına etkisi belirler ve ağırlıkları bu etkiye göre günceller. Hatanın katmanlara iletilmesi sırasında gradyan adı verilen sayılar kullanılır. Gradyanlar, bazı durumlarda çok küçük olup neredeyse sıfıra yaklaşabilmektedir. Bu duruma, kaybolan gradyan problemi denilmektedir.

Bu probleme karşı önerilen yapılardan biri olan artık bloklar, giriş verisi sırasıyla iki adet ağırlık katmanından geçirilmekte ve her katmandan sonra ReLU aktivasyon fonksiyonu uygulanmaktadır. ResNet50 mimarisinin temel yapı taşı olan artık blokların, matematiksel ifadesi Denklem (2-20)'de verildiği üzere tanımlanmaktadır.

$$y = F(x, \{W_i\}) + x_2 \quad (2-20)$$

Burada, x parametresi giriş verisini; $F(x, \{W_i\})$ parametresi ağırlık parametresini; x_2 parametresi giriş verisinin kısa yol bağlantısı üzerinden çıkışa aktarılan halini; y parametresi ise artık blok çıkışını temsil etmektedir.

ResNet50 mimarisinde, evrişim işlemleri sırasıyla 1×1 , 3×3 ve tekrar 1×1 filtreler ile gerçekleştirilir. İlk 1×1 evrişim, giriş verisinin boyutunu azaltarak modelin hesaplama yükünü hafifletmektedir. Bunu takip eden 3×3 evrişim, daha düşük boyuttaki veriyle çalıştığı için daha verimli sonuçlar üretmektedir. Son olarak uygulanan ikinci 1×1 evrişim, çıktının özgün boyutuna geri dönmesini sağlamaktadır. Bu üç aşamalı yapı sayesinde hem bilgi kaybı minimize edilmekte hem de daha derin ağların eğitimi daha sabit hâle gelmektedir. Bu mimarinin yapısı Şekil 2-8'de gösterilmiştir.



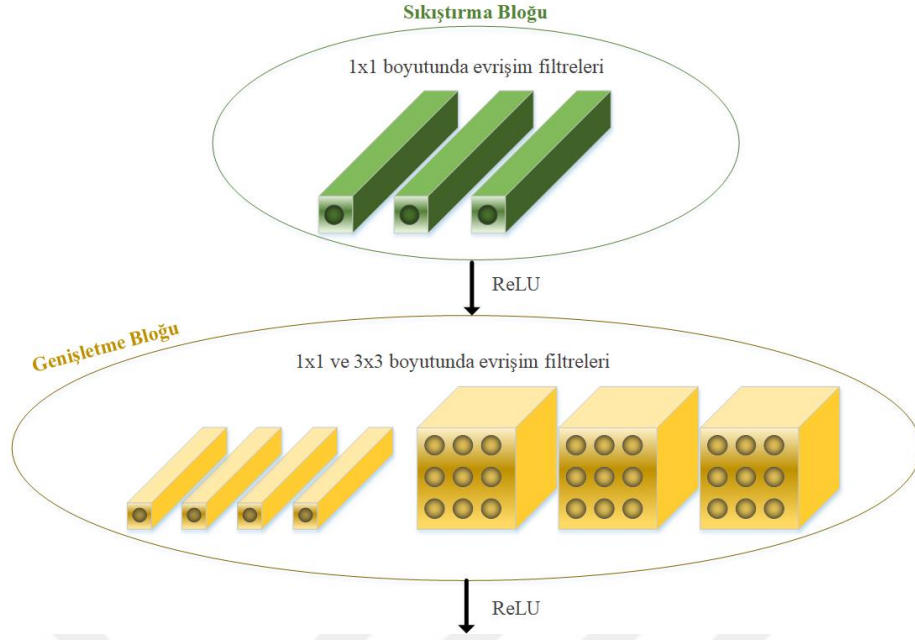
Şekil 2-8: ResNet50 mimarisi

Şekil 2-8’de sunulan mimari, giriş görüntüsü olarak $224 \times 224 \times 3$ piksel boyutunda, bir polip görüntüsü almaktadır. İlk katmanda, 7×7 çekirdek boyutuna sahip ve 64 filtreli bir evrişim işlemi uygulanmaktadır. Bu işlem ile 2 birimlik kaydırma adımı gerçekleştirilerek giriş görüntüsü daha düşük boyutlu bir özellik haritasına dönüştürülmektedir. Bu katmanın amacı, düşük seviyeli özelliklerin çıkarılmasıdır. Evrişim işleminin ardından, 2×2 boyutunda bir maksimum havuzlama işlemi uygulanmakta ve özellik haritasının boyutu yarıya indirilmektedir. Bu sayede modelin parametre sayısı azalmaktadır. Bu işlemleri takiben, artık bloklar bulunmaktadır. Artık blok yapıları, öğrenme sırasında giriş bilgisinin doğrudan ileriye taşınmasına olanak tanımaktadır. Bu durum, katmanlar arası bilgi kaybını önlenmektedir.

Her artık blok, üç aşamadan oluşmaktadır. İlk aşamada, 1×1 boyutlu evrişim uygulanarak giriş verisinin kanal sayısı azaltılmaktadır. Bu sayede, hesaplama maliyeti düşürülmektedir. Ardından, özellik çıkarımının yapıldığı 3×3 boyutlu evrişim uygulanmaktadır. Son olarak, tekrar 1×1 boyutlu evrişim ile kanal sayısı artırılarak verinin boyutu eski haline getirilmektedir. Artık blokların ardından, 2×2 boyutunda ortalama havuzlama işlemi gerçekleştirilir. Bu katman, tüm özellik haritasını özetleyerek veriyi sınıflandırma için uygun hale getirir. Son aşamada, elde edilen özellik vektörü tam bağlı katmana aktarılmakta ve sınıflandırma işlemi tamamlanmaktadır.

2.6 SqueezeNet MİMARİSİ

SqueezeNet mimarisi, alt örnekleme katmanları ve tam bağlı katmanlarla entegre edilmiş bir evrişimli sinir ağıdır [112]. Bu mimari, sıkıştırma-genişletme modülleri (Fire modul) olarak adlandırılan bileşenlerden oluşan bir yapıya sahiptir. Bu modüllerin yapısal gösterimi Şekil 2-9’da sunulmuştur.



Şekil 2-9: Sıkıştırma-genişletme modülü

Şekil 2-9’da sunulan sıkıştırma-genişletme modülü mimarisi, sıkıştırma ve genişletme olmak üzere iki temel aşamadan oluşmaktadır. Sıkıştırma aşamasında, 1x1 evrişim uygulanarak kanal sayısı azaltılmaktadır. Bu aşamanın matematiksel ifadesi Denklem (2-21)’de sunulmuştur.

$$x_s = \sigma (W_s * x + b_s) \quad (2-21)$$

Burada, x_s parametresi sıkıştırılmış çıkışı; x parametresi giriş verisini; W_s parametresi 1x1 evrişim filtresini; b_s parametresi bias terimini; σ parametresi aktivasyon fonksiyonunu temsil etmektedir.

Bir diğer aşama olan genişletme aşamasında, sıkıştırılmış veri hem 1×1 hem de 3×3 evrişim işlemlerine tabi tutulur. İlgili matematiksel ifadeler, sırasıyla Denklem (2-22) ve Denklem (2-23)’te sunulmaktadır.

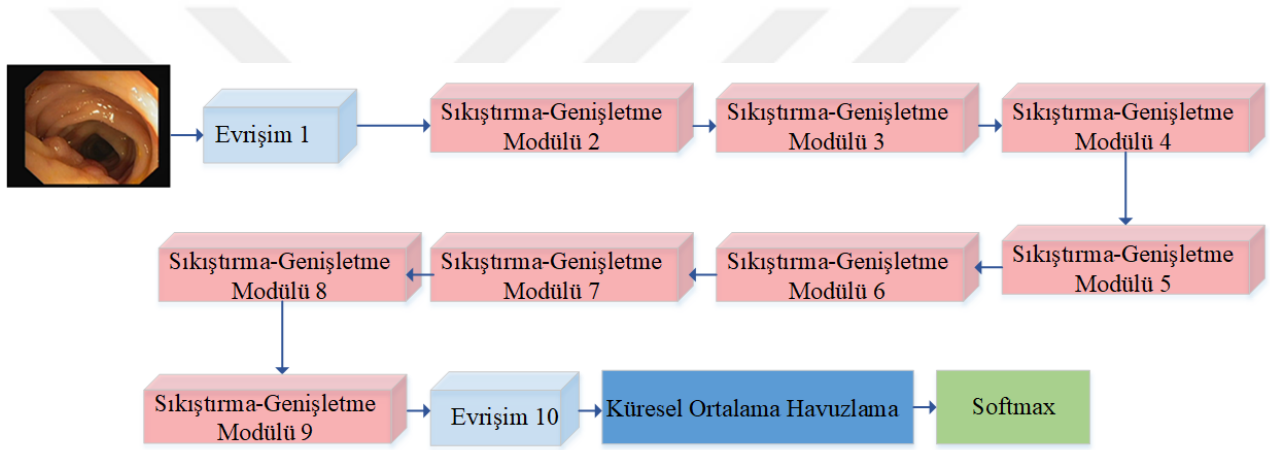
$$x_{e1} = \sigma (W_{e1} * x_s + b_{e1}) \quad (2-22)$$

$$x_{e3} = \sigma (W_{e3} * x_s + b_{e3}) \quad (2-23)$$

Burada Denklem (2-22)’de, W_{e1} parametresi 1×1 evrişim filtresine ait ağırlıkları; x_s parametresi sıkıştırılmış çıktıyı; b_{e1} parametresi 1×1 evrişim filtresine ait bias terimini; x_{e1}

parametresi 1×1 evriřim sonucu ıktısını; σ parametresi aktivasyon fonksiyonunu temsil etmektedir. Denklem (2-23)'te ise; W_{e3} parametresi 3×3 evriřim filtresine ait ağırlıkları; x_s parametresi sıkıştırılmış ıktıyı; b_{e3} parametresi 3×3 evriřim filtresine ait bias terimini; x_{e3} parametresi 3×3 evriřim sonucu ıktısını temsil etmektedir.

SqueezeNet mimarisi [112], genellikle bařlangıta bir evriřim katmanı, bunu takiben sekiz adet sıkıştırma-geřiřletme modl ve son blmde bir evriřim katmanı ile global ortalama havuzlama katmanından oluřan bir yapıya sahiptir. Bu mimari, sınıflandırma grevlerinde yksek doęruluk oranını korurken, model parametre sayısını ve boyutunu nemli lde azaltmayı hedeflemektedir. Geleneksel tam baęlantılı katmanlar yerine kresel ortalama havuzlama kullanılarak, zellikle parametre sayısında nemli lde tasarruf saęlanmaktadır. SqueezeNet mimarisinin genel yapısı [112] řekil 2-10'da gsterilmektedir.

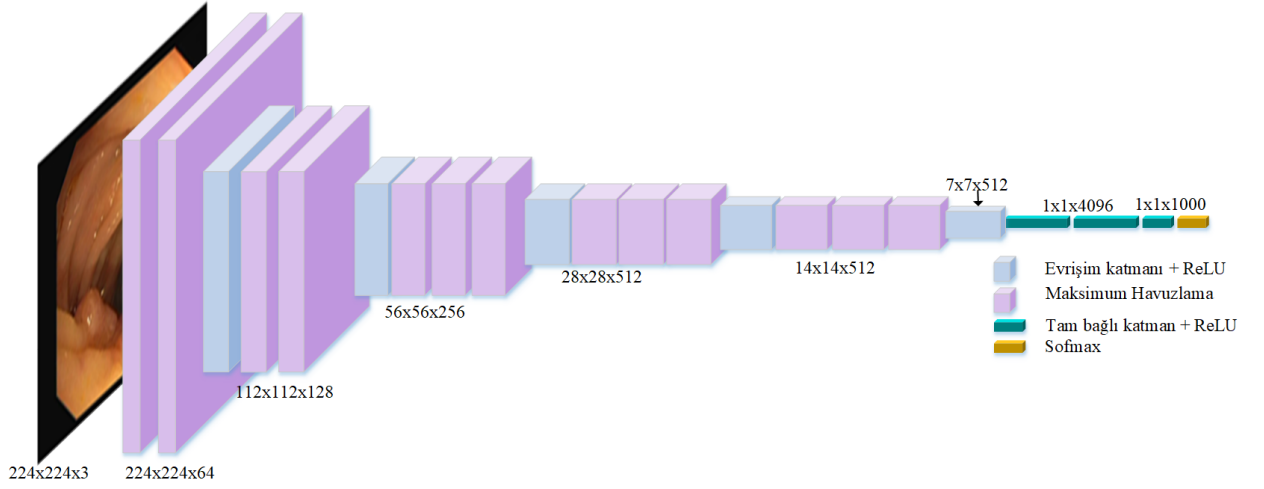


řekil 2-10: SqueezeNet mimarisinin genel yapısı

řekil 2-10'da sunulan mimaride, ilk olarak Evriřim 1 katmanı ile grntden temel zellikler ıkartılmaktadır. Ardından, 8 adet Sıkıştırma-geřiřletme modl yer almaktadır. Bu modller, grnty gereksiz bilgilerden arındırarak nemli bilgileri ıkartmaktadır. Modller boyunca iřlenen grnt, Evriřim 10 katmanına ulařtıęında, grntdeki nemli ve anlamlı bilgileri ne ıkartılmaktadır. Bylece, modelin doęru sınıflandırma yapabilmesi iin son bilgi hazırlıęını tamamlanmıř olur. Ardından gelen Kresel Ortalama Havuzlama katmanı, veriyi zetleyerek, sınıflandırma iin uygun boyuta getirir. Son olarak, Softmax katmanı ile model, grntnn hangi sınıfa ait olduęunu belirlemektedir.

2.7 VGG16 MİMARİSİ

VGG16, 2014 yılında Oxford Üniversitesi'nden Görsel Geometri Grubu (Visual Geometry Group-VGG) tarafından geliştirilen ve önemli başarı elde eden bir evrişimsel sinir ağı mimarisidir [113]. VGG16 mimarisi [113], 13 evrişim, 3 tam bağlı katman, *ReLU* aktivasyon fonksiyonu, maksimum havuzlama katmanlarından oluşmaktadır. VGG16 mimarisinin genel yapısı [113], Şekil 2-11'de sunulmuştur.



Şekil 2-11: VGG16 mimarisi

Şekil 2-11'de polip görüntüsü VGG16 mimarisine giriş olarak verilmiştir. Bu görüntü, öncelikle bir dizi evrişim ve havuzlama katmanlarından geçirilmektedir. VGG16 mimarisi [113], daha fazla derinlik ve daha az parametreyle daha iyi sonuçlar için 3×3 boyutunda küçük evrişim filtreleri kullanmaktadır. Her bir evrişim bloğu, artan filtre sayısı ile daha karmaşık ve ayrıntılı özelliklerin çıkarımını mümkün kılmaktadır [114]. Matematiksel olarak, bir evrişim katmanının çıkışı Denklem (2-24)'te gösterilmiştir.

$$Y_{i,j,k} = b_k + \sum_{u=1}^3 \sum_{v=1}^3 \sum_{c=1}^C W_{u,v,c,k} \cdot X_{i+u-1,j+v-1,c} \quad (2-24)$$

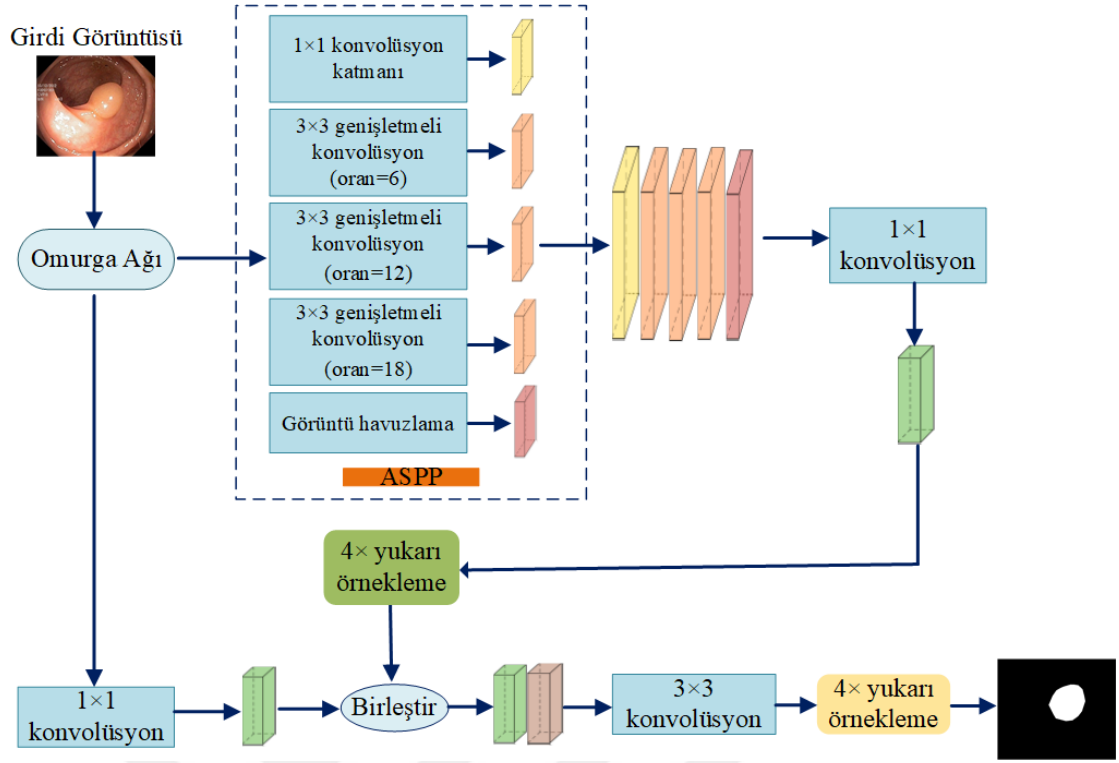
Burada, $Y_{i,j,k}$ parametresi evrişim katmanının k . filtreye ait, i, j konumundaki çıktı değerini; b_k parametresi k . filtreye ait bias terimini; $\sum_{u=1}^3 \sum_{v=1}^3$ işlemleri 3×3 boyutundaki evrişim filtresi üzerinde, yatay (u) ve dikey (v) olmak üzere tüm konumlardaki toplama işlemini; $\sum_{c=1}^C$ giriş verisindeki tüm kanal özelliklerinin C boyunca toplamını; $W_{u,v,c,k}$ 3×3 boyutundaki evrişim filtresinin k . çıkış kanalı için, u, v konumundaki ve c . giriş kanalıyla ilişkili ağırlık

parametresini; $X_{i+u-1,j+v-1,c}$ ise giriş özellik haritasının c . kanalındaki $i + u - 1, j + v - 1$ uzaysal konumundaki değerini temsil etmektedir.

2.8 DeepLabv3+ MİMARİSİ

DeepLabv3+ mimarisi segmentasyon görevi için tasarlanmış bir mimaridir [115]. Bu mimari, DeepLabv3 [116] mimarisini temel alarak oluşturulmuştur. İki mimari arasındaki en belirgin fark ise, özellikle nesne sınırlarında segmentasyon doğruluğunu artırmak amacıyla DeepLabv3+ mimarisine bir çözücü modülünün eklenmiş olması ve Xception [115] modeline dayanan bir omurga ağı içermesidir. Bu mimaride, 1×1 genişletilmiş evrişim, Atrous oranları sırasıyla 6, 12 ve 18 olan üç adet 3×3 atrous evrişim katmanı, küresel ortalama havuzlama katmanı içeren bir Aralıklı Uzamsal Piramit Havuzlama (Atrous Spatial Pyramid Pooling-ASPP) modülü kullanılmaktadır. ASPP modülü, çok ölçekli bağlamsal bilgilerin etkin bir biçimde yakalanmasını sağlamaktadır.

ASPP modülünden elde edilen çok ölçekli bağlamsal özellikler, çözücü modüle aktarılmadan önce dört kat alt örnekleme işlemine uygulanmaktadır. Kod çözücü ise, omurga ağından elde edilen düşük seviyeli konumsal özellikler ile ASPP modülünden gelen yüksek seviyeli bağlamsal temsilleri birleştirerek detaylı segmentasyon sağlamaktadır. DeepLabv3+ mimarisinin genel yapısı [115], Şekil 2-12’de sunulmaktadır.



Şekil 2-12: DeepLabv3+ mimarisinin genel yapısı

3. KULLANILAN VERİ SETLERİ VE PERFORMANS ÖLÇÜTLERİ

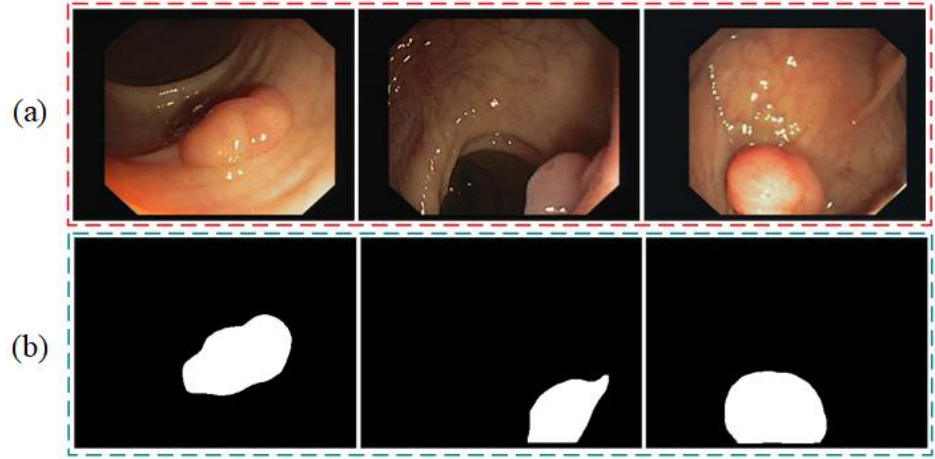
3.1 VERİ SETLERİ

Hastaların, sağkalım oranlarını iyileştirmek için kanser öncesi poliplerin erken teşhisi ve tedavisi oldukça önemlidir [117]. Bu amaçla geliştirilen polip segmentasyon algoritmalarının performansını değerlendirmek için çeşitli halka açık endoskopik görüntü veri setleri kullanılmaktadır. Bu tez çalışmasında kullanılan polip segmentasyon veri setlerine ait genel bilgiler Tablo 3-1’de sunulmaktadır.

Tablo 3-1: Literatürde kullanılan polip segmentasyon veri setlerine ait genel bilgiler

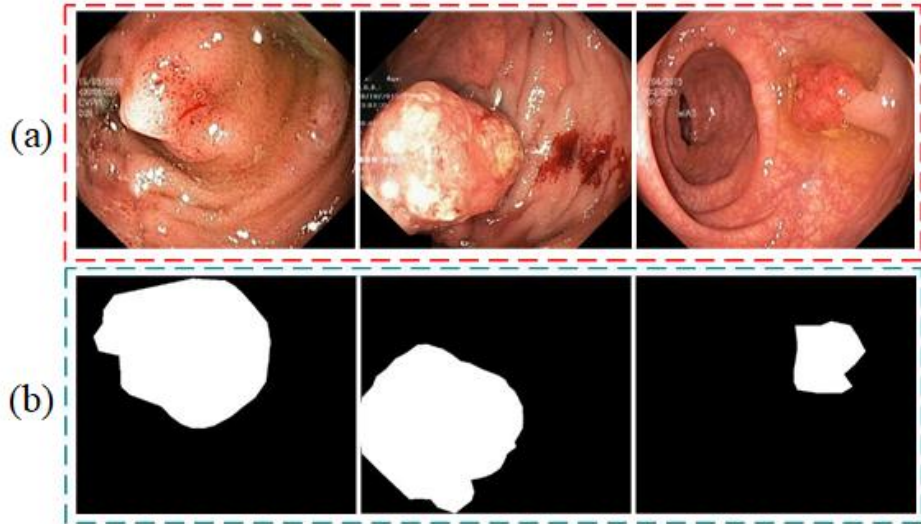
Veri Seti (Yıl)	Referans	Erişim Bağlantısı	Format	Çözünürlük
CVC-ClinicDB (2015)	[20]	https://polyp.grand-challenge.org/CVCClinicDB/	Görüntü (tiff)	384 × 384
Kvasir-SEG (2020)	[118]	https://datasets.simula.no/kvasir-seg/	Görüntü	332 × 487 – 1920 × 1072
CVC-ColonDB (2011)	[119]	http://mv.cvc.uab.es/projects/colon-qa/cvccolondb	Görüntü	574 × 500

CVC-ClinicDB Veri Seti: 2015 yılında yayınlanan CVC-ClinicDB veri seti, İspanya'nın Barselona şehrindeki Hospital Clinic ile iş birliği içinde derlenmiştir [20]. Bu veri seti, 23 farklı kolonoskopi video kesitinden oluşmakta ve her birinden 31 farklı polip görüntüsü elde edilmektedir [20]. Düşük kaliteye sahip görüntüler veri setine dahil edilmemiştir. CVC-ClinicDB [20] veri seti, 612 görüntüden oluşmakta olup, her bir görüntü 576x768 piksel çözünürlüğündedir [20]. Şekil 3-1’de, CVC-ClinicDB [20] veri setinden alınan üç farklı polip görüntüleri ile bu görüntülerin gerçek zemin görüntüleri (maske) sunulmuştur.



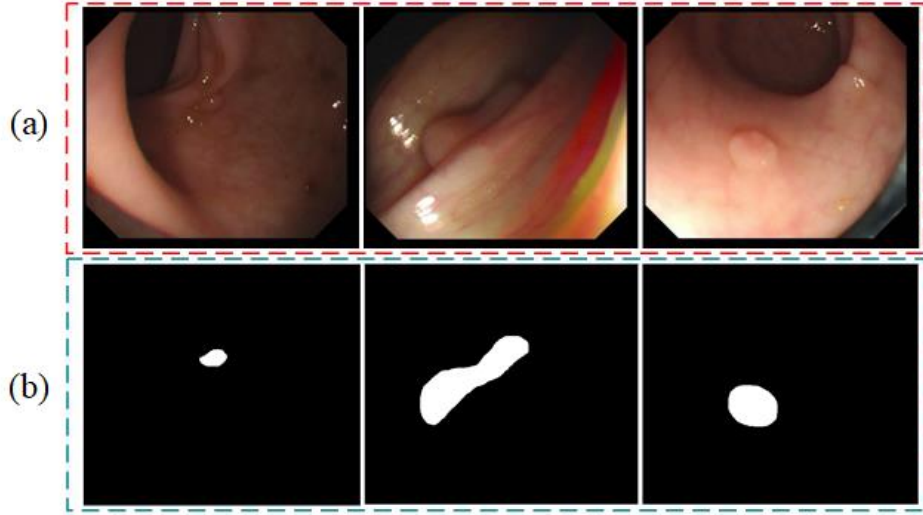
Şekil 3-1: CVC-ClinicDB veri setine ait görseller: (a) orijinal polip görüntüsü, (b) orijinal görüntülere karşılık gelen gerçek ikili görüntüleri temsil etmektedir.

Kvasir-SEG Veri Seti: Kvasir-SEG veri seti, 2020 yılında Norveç'teki Oslo Üniversitesi Hastanesi'nde uzman hekimlerin kolorektal polip segmentasyonu amacıyla oluşturduğu açık erişimli bir kaynaktır [118]. Bu veri setindeki 1000 polip görüntüsü ve bu görüntülere karşılık gelen maskeler, 332x487 piksel ile 1920x1072 piksel arasında değişmektedir. Şekil 3-2'de, Kvasir-SEG [118] veri setinden alınan üç farklı polip görüntüleri ile bu görüntülerin gerçek zemin görüntüleri (maske) sunulmuştur.



Şekil 3-2: Kvasir-SEG veri setine ait görseller: (a) orijinal polip görüntüsü, (b) orijinal görüntülere karşılık gelen gerçek ikili görüntüleri temsil etmektedir.

CVC-ColonDB Veri Seti: CVC-ColonDB veri seti, Mayo Kliniği iş birliğiyle yürütülen projede 15 kısa kolonoskopi videosundan elde edilen karelerin derlenmesiyle oluşturulmuş bir veri setidir [120]. Bu veri seti 380 polip ve maske görüntülerinde oluşmaktadır. Tüm görüntüler 574x500 piksel çözünürlüğüne sahiptirler. Şekil 3-3'de, CVC-ColonDB [119] veri setinden alınan üç farklı polip görüntüleri ile bu görüntülerin gerçek zemin görüntüleri (maske) sunulmuştur.



Şekil 3-3: CVC-ColonDB veri setine ait görseller: (a) orijinal polip görüntüsü, (b) orijinal görüntüye karşılık gelen gerçek ikili görüntüleri temsil etmektedir.

3.2 PERFORMANS METRİKLERİ

Tıbbi görüntü segmentasyonu, normal olmayan doku büyümelerinin doğru ve erken bir şekilde tespit edilmesinde kritik bir rol oynamaktadır. Segmentasyon işleminde, model başarısını ölçmek için kullanılan çeşitli performans metrikleri bulunmaktadır. Bu tez çalışmasında, uygulamaların performans değerlendirmeleri yedi farklı performans metriği ile yapılmıştır. Bu kapsamda değerlendirilen ilk metrik, Ortalama Dice Katsayısı (Mean Dice Coefficient-mDice) metriğidir. *Dice* Katsayısı, tahmin edilen bölge ile gerçek bölgenin ne kadar örtüştüğünü ifade eden bir metriktir. *mDice* ise, birden fazla sınıf veya görüntü için hesaplanan *Dice* Katsayısının ortalamasını ifade eder. Bu performans metriğine ait matematiksel ifade Denklem (3-1)'te sunulmuştur.

$$mDice = \frac{1}{1+k} \sum_{i=0}^k \frac{2TP}{2TP+FP+FN} \quad (3-1)$$

Ortalama Kesişim Bölme Oranı (Mean Intersection over Union-*mIoU*) performans metriği *mIoU* değerlerinin ortalaması ile hesaplanmaktadır. Bu metrik, tahmin edilen bölge ile gerçek bölgenin kesişim alanını ifade etmekte olup değeri 0 ile 1 arasında değişmektedir. Segmentasyon görevinde *mIoU* değerinin 0'a yaklaşması performansın düşük olduğunu bir göstergesi iken, 1'e yaklaşması segmentasyon görevinin başarılı olduğunu gösterir. Bu metriğin matematiksel ifadesi Denklem (3-2)'de sunulmuştur

$$mIoU = \frac{1}{1+k} \sum_{i=0}^k \frac{TP}{TP+FP+FN} \quad (3-2)$$

Doğruluk (Accuracy-Acc) performans metriği ise, modelin doğru sınıflandırdığı örneklerin, toplam örnek sayısına oranını ifade eder. Bu metrik, dengesiz veri setlerinde yanıltıcı bilgi verebilir. Bu nedenle, medikal segmentasyon görevi gibi sınıf dengesizliğinin çok olduğu görevlerde tek başına kullanılması doğru değildir. Geri çağırma (Recall-R) metriği, gerçek pozitiflerin model tarafından doğru şekilde tespit edilme oranını ifade etmektedir. Bu metriğin, yüksek sonuç vermesi pozitif olan durumları doğru şekilde tespit ettiğini ifade etmektedir. Uygulamalarda geri çağırma performans metriğinin, yüksek sonuç vermesi modelin performansının iyi olduğu göstermektedir. Metriğin düşük sonuçlar vermesi ise, pozitif örneklerin gözden kaçırıldığını ifade etmektedir. Düşük olduğu durumlarda, sınıflarda dengesizliği söz konusu olabilmektedir. Doğruluk performans metriği Denklem (3-3)'de, geri çağırma performans metriği ise Denklem (3-4)'de matematiksel olarak ifade edilmiştir.

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \quad (3-3)$$

$$R = \frac{TP}{TP+FN} \quad (3-4)$$

Özgüllük (Specificity -Sp) performans metriği, doğru negatiflerin model tarafından doğru şekilde tanımlanma oranını sunmaktadır. Eğer modelin özgüllük değeri yüksekse, bu modelin negatif sınıfları doğru bir şekilde tahmin ettiğini; düşükse bu modelin negatif sınıfları ayırt etmede zorlandığını göstermektedir. Özgüllük performans metriğinin düşük sonuçlara sahip olması, yanlış

pozitif sonuçlara yol açmaktadır. Sp performans metriğinin matematiksel ifadesi Denklem (3-5)'de sunulmuştur.

$$Sp = \frac{TN}{TN+FP} \quad (3-5)$$

Kesinlik (Precision- Pr) metriği ise, modelin pozitif olarak doğru tahmin ettiği örneklerin, doğru pozitif tahmin ve yanlış pozitif tahminlerinin toplamına oranını ifade etmektedir. Kesinlik değerinin yüksek bir değer sahip olması pozitif örneklerin doğru tahmin edildiğini ifade etmektedir. Düşük bir değere sahip olması ise, pozitif değerleri negatif değer olarak tahmin ettiğini göstermektedir. Denklem (3-6)'da matematiksel olarak ifade edilmiştir.

$$Pr = \frac{Tp}{TP+FP} \quad (3-6)$$

Ortalama Mutlak Hata (Mean Absolute Error- MAE) performans metriği, model tarafından tahmin edilen değerler ile gerçek değerler arasındaki mutlak farkların ortalamasını ifade etmektedir. MAE performans metriğinin yüksek değere sahip olması modelin gerçek değerlerden uzak olduğunu ifade ederken; düşük değere sahip olması gerçek değerlere yakın tahminler yaptığını ifade etmektedir. Matematiksel ifadesi Denklem (3-7)'de sunulmuştur.

$$MAE = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H |S(i,j) - G(i,j)| \quad (3-7)$$

Burada, W parametresi, görüntünün genişliğini; H parametresi, görüntünün yüksekliğini temsil etmektedir. $S(i,j)$ parametresi, modelin (i,j) pikseline ait sınıf tahminini, $G(i,j)$ ise aynı piksele karşılık gelen gerçek sınıf değerini göstermektedir.

4. UYGULAMALAR

4.1 UYGULAMA 1: FARKLI OMURGA AĞLARI KULLANILARAK DeepLabv3+ MİMARİSİ İLE POLİP SEGMENTASYONU

Bu tez kapsamında, DeepLabv3+ mimarisi ile dört farklı omurga kullanılmıştır. Çeşitli değerlendirme metrikleri ile performans karşılaştırılması gerçekleştirilmiştir. Ayrıca farklı veri setleri ile değerlendirmeler gerçekleştirilmiştir.

4.1.1 Uygulama Ortamı ve Eğitim Parametreleri

Bu tez kapsamında, kolorektal polip segmentasyonu üzerine yapılan deneyler Python 3.0 ve TensorFlow kütüphanesi kullanılarak gerçekleştirilmiş ve eğitim süreçleri NVIDIA RTX A4000 GPU'lu bir sistem üzerinde yürütülmüştür.

Tezin bu uygulamasında, tasarlanan dört model için, aynı eğitim parametreleri kullanılmıştır. BU eğitim parametreleri Tablo 4-1'de sunulmuştur.

Tablo 4-1: Eğitim parametrelerinin değerleri

Hiperparametre	Değer
Öğrenme Oranı	0,0001
Yığın boyutu	8
Aktivasyon Fonksiyonu	ReLU
Optimizasyon Algoritması	Adam

Modelin her eğitim adımında ağırlıklarının ne kadar değişeceği öğrenme oranı tarafından belirlenmektedir. Öğrenme oranı 0,001 olarak ayarlanmıştır. Bu değer, modelin her adımda ağırlıklarının küçük ve kontrollü güncellemesini sağlamıştır. Böylece model, yavaş ve istikrarlı bir şekilde öğrenme gerçekleştirmiştir. Yığın boyutu, modelin ağırlıklarını güncellemeden önce aynı anda işlediği eğitim örneği sayısıdır; bu tez kapsamında 8 olarak ayarlanmıştır. Bu değer büyük bir değer olarak ayarlanması, daha fazla bellek kullanmasına yol açmakta; ayrıca modelin genelleme yeteneğini olumsuz etkilemektedir. Yığın boyutunun 8 olarak ayarlanması, eğitim süresini uzatmadan, güncellemeleri sık yaparak daha hızlı bir öğrenme imkanı sunmuştur. Uyarlanabilir Moment Tahmini (Adaptive Moment Estimation-Adam) optimizasyon algoritması,

her parametre için öğrenme oranını değiştirerek eğitimi daha hızlı ve kararlı bir şekilde gerçekleştiren bir optimizasyon algoritmasıdır. Bu algoritma, geçmiş gradyanların ortalaması ve karesini kullanarak ağırlık güncellemeleri gerçekleştirir. Tez kapsamında, modelin etkin ve hızlı öğrenmesi için Adam optimizasyon algoritması tercih edilmiştir. Son olarak, Doğrusal Düzeltme Birimi (Rectified Linear Unit-ReLU) aktivasyon fonksiyonu tüm katmanlar arasında geçişlerde kullanılmıştır. ReLU aktivasyon fonksiyonunun, modelin öğrenme sürecini hızlandırması ve gradyan kaybı problemini büyük ölçüde azaltması nedeniyle tercih edilmiştir.

4.1.2 Önerilen Yöntem

Bu tez kapsamında önerilen mimari, polip segmentasyonu için dört farklı CNN tabanlı omurga ağı kullanılarak tasarlanmış, DeepLabv3+ tabanlı bir mimaridir. Önerilen mimaride, kamuya açık olarak sunulan Kvasir-Seg [118] veri seti ve CVC-ClinincDB [20] veri seti kullanılmıştır.

Önerilen DeepLabv3+ mimarisi, ASPP modülü ve omurga ağından oluşmaktadır. Bu çalışmada, omurga ağı olarak DenseNet mimarisi [110], ResNet50 mimarisi, SqueezeNet mimarisi [112] ve VGG16 mimarisi [113] DeepLabv3+ mimarisine entegre edilmiştir. Deneysel çalışmaları daha iyi ifade edebilmek için, farklı omurga ağı kullanılarak çalıştırılan DeepLabv3+ mimarisi deneysel çalışma sırasına göre numaralandırılmış ve bu işlem Tablo 4-2'de verilmiştir.

Tablo 4-2: Kullanılan modeller ve omurga ağları

Model Adı	Omurga Ağı
Model 1	DenseNet Mimarisi
Model 2	ResNet50 Mimarisi
Model 3	SqueezeNet Mimarisi
Model 4	VGG16 Mimarisi

DenseNet mimarisi [110], omurga ağı olarak tercih edilmesinin temel nedenleri, katmanlar arası yoğun bağlantıları sayesinde düşük ve yüksek seviyeli özelliklerin etkin paylaşımıyla zengin temsil öğrenimi sunmasıdır. Gradyan akışının iyileşmesiyle daha derin ağlarda bile kararlı eğitim sağlaması; parametre verimliliğiyle sınırlı tıbbi veri setlerinde aşırı öğrenmeyi (overfitting) azaltması; ayrıca, DeepLabv3+'ın ASPP modülleriyle elde edilen çok ölçekli bağlam bilgisinin,

DenseNet'in bol bağlantılı yapısıyla kenar detayları ve global morfoloji bilgisini eşzamanlı işleyerek polip segmentasyonunun doğruluk ve tutarlılığını artırmasıdır.

ResNet50 mimarisi, derinliği sayesinde karmaşık verilerden zengin ve ayrıntılı özellikler çıkarabilmektedir. En önemli özelliği ise artık bloklar sayesinde, gradyan kaybı sorununu önlemesidir. Bu nedenle, öğrenme daha düzenli ve etkili bir şekilde gerçekleşir. ResNet50'nin güçlü özellik çıkarımı yeteneği, DeepLabv3+ ile birlikte kullanıldığında segmentasyon performansını önemli ölçüde artırmakta ve modelin genelleme kabiliyetini geliştirmektedir. ResNet50 mimarisinin omurga ağı olarak tercih edilmesi, bu özelliklerin tümünün bir sonucudur.

Bir diğer mimari olan SqueezeNet mimarisi [112], az parametre sayısına sahip bir mimaridir. Bu durum, modelin eğitim sürecini ve işlem süresini önemli ölçüde kısaltmasına izin vermektedir. DeepLabv3+ mimarisi, SqueezeNet'in bu etkili yapısı sayesinde karmaşık segmentasyon görevlerini daha hızlı bir şekilde gerçekleştirebilmektedir. Bu bağlamda, SqueezeNet mimarisinin [112] omurga ağı olarak kullanılması, modelin hafifliği ve hesaplama verimliliği nedeniyle tercih edilmiştir. Ayrıca, segmentasyon doğruluğundan ödün vermeden yüksek performans elde edilmesini sağlamıştır.

Son olarak, önerilen mimaride omurga ağı olarak VGG16 mimarisi [113], giriş görüntüsünden ayrıntılı ve anlamlı özellikler çıkartmaktadır. Daha modern omurga ağlarına kıyasla derinliği sınırlıdır, ancak daha az karmaşık yapısı sayesinde eğitim sürecinde daha hızlı sonuçlar üretmektedir. Bu doğrultuda, VGG16 omurgasının DeepLabv3+ mimarisi ile birlikte kullanılması, doğruluk odaklı segmentasyon uygulamalarında önemli avantajlar sunmaktadır. VGG16 mimarisinin [113] sunduğu ayrıntılı özellik çıkarımı, DeepLabv3+'ın çok ölçekli bağlamsal bilgi yakalama yeteneğiyle birleştiğinde, modelin genel segmentasyon performansında belirgin bir artış sağlanmaktadır.

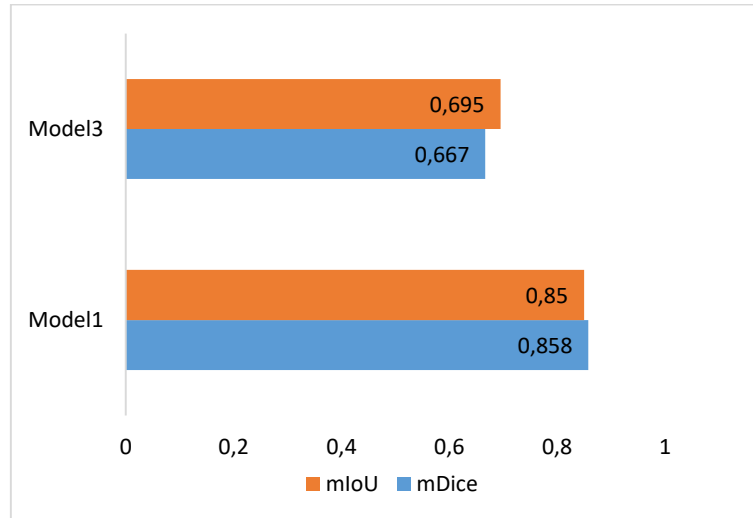
4.1.3 Sonuçlar

Bu tez kapsamında kullanılan veri setleri, eğitim, doğrulama ve test kümeleri olarak ayrılmıştır. Her veri kümesi, önce rastgele karıştırılmış daha sonra %20 test, %80 eğitim verisi olarak ayrılmıştır. Daha sonra, eğitim verisinin %20'si ise doğrulama için kullanılmıştır. Uygulamanın polip segmentasyon görevinde elde ettiği performans sonuçları farklı metrikler ile değerlendirilmiştir. Elde edilen sonuçlar Tablo 4-3'de sunulmuştur. Bu tabloda yeşil renk ile yazılan değerler ilgili sütundaki performans metriğinde elde edilen en yüksek sonucu temsil ederken, mavi renk ile yazılan değerler en düşük sonucu temsil etmektedir.

Tablo 4-3: Kvasir-SEG ve CVC-ClinicDB veri setleri üzerinde elde edilen segmentasyon performans sonuçları

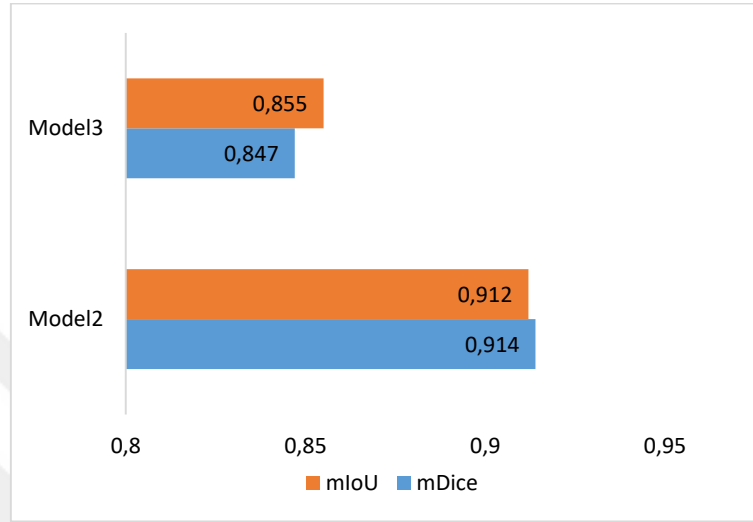
Model	mDice	mIoU	Acc	R	Sp	Pr	MAE
Kvasir-SEG [118] Veri Seti							
Model 1	0,858	0,850	0,948	0,824	0,981	0,896	0,045
Model2	0,825	0,821	0,939	0,776	0,979	0,881	0,054
Model3	0,667	0,695	0,893	0,634	0,949	0,703	0,102
Model4	0,818	0,818	0,943	0,754	0,984	0,895	0,500
CVC-ClinicDB [20] Veri Seti							
Model1	0,907	0,906	0,960	0,860	0,996	0,959	0,016
Model2	0,914	0,912	0,961	0,921	0,972	0,928	0,015
Model3	0,847	0,855	0,972	0,813	0,990	0,884	0,026
Model4	0,871	0,874	0,974	0,824	0,993	0,923	0,023

Kvasir-SEG [118] veri seti ile elde edilen en yüksek *mDice* ve *mIoU* performans metrik değerleri DenseNet mimarisinin [110] omurga ağı olarak kullanıldığı modelde (Model1) en yüksek segmentasyon performansını gösterirken; bu metriklerin en düşük segmentasyon performans değeri SqueezeNet mimarisinin [112] omurga ağı olarak kullanıldığı modelde (Model3) elde edilmiştir. İlgili karşılaştırmalar Şekil 4-1’de sunulmuştur.



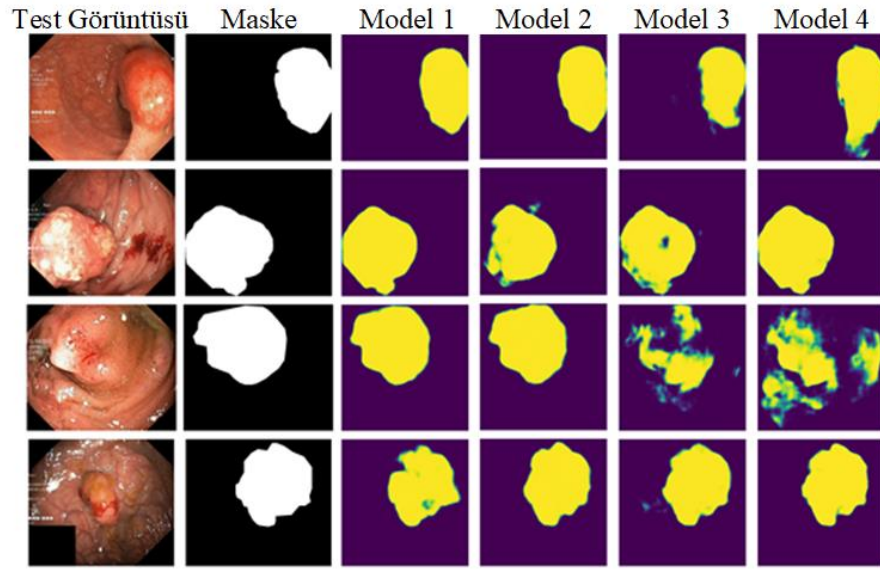
Şekil 4-1: Kvasir-SEG veri seti üzerinde farklı modellerin segmentasyon performans metrikleri

CVC-ClinicDB [20] veri seti ile elde edilen en yüksek *mDice* ve *mIoU* performans metrik değerleri ResNet50 mimarisinin omurga ağı olarak kullanıldığı modelde (Model2) en yüksek segmentasyon performansını gösterirken; bu metriklerin en düşük segmentasyon performans değeri SqueezeNet mimarisi [112] omurga ağı olarak kullanıldığı modelde (Model3) elde edilmiştir. Karşılaştırılmalı analizi Şekil 4-2’de sunulmuştur.

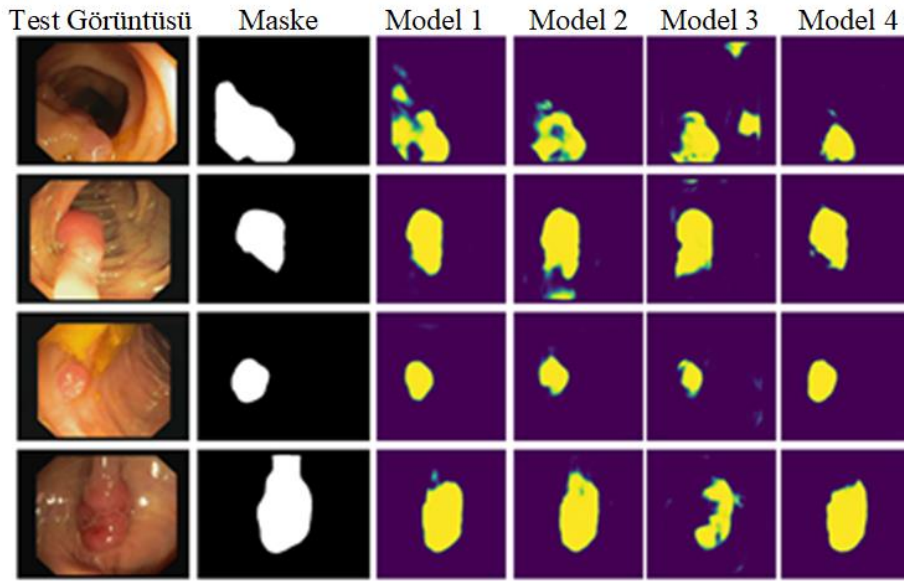


Şekil 4-2: CVC-ClinicDB veri seti üzerinde farklı modellerin segmentasyon performans metrikleri

Şekil 4-3 ve Şekil 4-4, sırasıyla Kvasir-SEG [118] veri seti ve CVC-ClinicDB [20] veri setlerine ait deneylerden elde edilen dört farklı veri seti görüntüsüne ait segmentasyon sonuçlarını göstermektedir. Görsellerde ilk sütun orijinal kolonoskopik görüntüleri, ikinci sütun ilk sütunda verilen görüntülerin gerçek zemin görüntülerini (maske görüntüleri) temsil etmektedir. Sonraki dört sütun ise sırası ile Model 1, Model 2, Model 3 ve Model 4 ‘ün ilk sütundaki görüntüden elde ettikleri segmentasyon sonuçlarını temsil etmektedir. Şekil 4-3 ve Şekil 4-4’de bahsedilen görseller [114]’da bahsedilen çalışmadan alınmış deneysel çalışmalardır.



Şekil 4-3: Kvasir-SEG [118] veri seti üzerindeki segmentasyon sonuçları



Şekil 4-4: CVC-ClinicDB [20] veri seti üzerindeki segmentasyon sonuçları

4.2 UYGULAMA 2: U-Net TABANLI POLİP SEGMENTASYONUNDA KAYIP FONKSİYONLARININ KARŞILAŞTIRMALI ANALİZİ

Bu tez kapsamında, U-Net mimarisinin farklı dört kayıp fonksiyonu üzerinde etkisi değerlendirilmiştir. Değerlendirme sürecinde, çeşitli performans metrikleri kullanılmış ve farklı veri setleri ile mimariler kapsamlı bir şekilde analiz edilmiştir.

4.2.1 Uygulama Ortamları ve Eğitim Parametreleri

Bu tez kapsamında, olrektal polip segmentasyonu ve eğitim süreçleri için Python 3.7 ile TensorFlow framework kullanılarak ve NVIDIA RTX A4000 GPU donanımına sahip bir sistem üzerinde yürütülmüştür.

Tezin bu uygulamasında, tasarlanan her model için aynı eğitim parametreleri kullanılmış; bu eğitim parametreleri Tablo 4-4’de sunulmuştur.

Tablo 4-4: Eğitim parametrelerinin değerleri

Hiperparametre	Değer
Öğrenme Oranı	0,0001
Yığın boyutu	8
Aktivasyon Fonksiyonu	ReLU
Optimizasyon Algoritması	Adam
Toplam Eğitim Döngüsü (Epoch)	100

Adam optimizasyon algoritması kullanılmıştır. *Adam* optimizasyon algoritması, her parametre için öğrenme hızını otomatik ayarlayan ve modelin daha hızlı ve dengeli öğrenmesini sağlayan bir algoritmadır. Bu tez kapsamın, öğrenme hızı 0,0001; her iterasyonda modele beslenen örnek sayısını 16 olarak ayarlanmıştır. Örnek sayısının 16 olarak ayarlanması, gradyan hesaplamalarının yeterince doğru ve güvenilir olmasına olanak tanırken; aynı zamanda modelin daha hızlı ve kararlı şekilde öğrenmesine olanak tanımıştır. Eğer örnek sayısı az tutulursa, öğrenmede dalgalanmalar söz konusu olurken, örnek sayısı fazla tutulursa modelde genelleme problemleri ortaya çıkmaktadır. Toplam eğitim döngüsü 100 olarak ayarlanmıştır. Bu sayede, model eğitim verisi üzerinde yeterince tekrarlar sağlamış; modelin performansı optimize edilmiştir. Ayrıca, girdi görüntüleri, 256x256 piksel olarak yeniden boyutlandırılmıştır. Veri setleri rastgele karıştırılmış, %80 eğitim verisi, %20 test verisi için kullanılmıştır. Eğitim verilerinin %20’si, doğrulamada kullanılmıştır.

4.2.2 Önerilen Yöntem

Bu tez kapsamında, polip segmentasyonu amacıyla, temeli U-Net mimarisine dayanan ve dört farklı kayıp fonksiyonu ile geliştirilmiş bir mimari tasarlanmıştır. Bu tez kapsamında önerilen çalışmada, iki halka açık veri seti, CVC-ColonDB [119] veri seti ve Kvasir-SEG [118] veri seti kullanılmıştır.

Sınıflandırma ve segmentasyon performansını doğru ölçmek için, segmentasyon görevlerinde kayıp fonksiyonları kullanılmaktadır. Kayıp fonksiyonları, rastgele bir olayın veya ilgili değişkenlerin değerlerini negatif olmayan reel sayılara eşleyen bir fonksiyondur.

Polip segmentasyonundaki U-Net mimarisinin [22] performansı, veri setleri üzerinde İkili Çapraz Entropi [104], Tversky Kayıp Fonksiyonu [106], Dice Kayıp Fonksiyonu [108] ve Bileşik Kayıp Fonksiyonu kullanılarak değerlendirilmiştir.

İlk kayıp fonksiyon olan İkili Çapraz Entropi [104] kayıp fonksiyonunun, matematiksel olarak basit olması, eğitim sürecinde optimizasyonun iyi bir şekilde ilerlemesine olanak tanımaktadır. Ayrıca, modelin çıktısını olasılık olarak değerlendirerek gerçek değerler ile tahmin edilen değerler arasındaki farklılıkları etkili bir şekilde ölçmektedir.

Tversky Kayıp Fonksiyonu [106], özellikle dengesiz veri setlerinde model performansını iyileştirmede kritik bir rol oynamaktadır. Segmentasyon görevinde, küçük ve nadir yapıları doğru şekilde tespit etmesi bu kayıp fonksiyonunun avantajlarından biridir. Medikal görüntü analizi gibi kritik uygulamalarda, dengesiz veri setlerinde sağladığı yüksek performans nedeniyle bu tez kapsamında U-Net mimarisi ile entegre edilerek uygulanmıştır.

Dice Kayıp Fonksiyonu [108], küçük nesnelerin doğru segmentasyonda başarılı bir performans göstermektedir. Ayrıca, tahmin edilen ve gerçek segmentasyon bölgeleri arasındaki örtüşmeyi doğrudan optimize etmesi ve hem yanlış pozitif hem de yanlış negatif hataları dengeli bir şekilde ele alması, bu kayıp fonksiyonunun önemli avantajları arasında yer almaktadır.

Bileşik Kayıp Fonksiyonu ise, Dice Kayıp Fonksiyonu ve İkili Çapraz Entropi birleşiminden oluşmaktadır. Bu yaklaşım, her iki kayıp fonksiyonun avantajlarını bir araya getirme imkanı sunmaktadır. Bileşik Kayıp Fonksiyonu, farklı kayıp fonksiyonları ile harmanlayarak genel performansı artırmaktadır.

4.2.3 Sonuçlar

Bu tez kapsamında gerçekleştirilen karşılaştırmalı çalışma, kayıp fonksiyonlarının entegrasyonunun modelin genelleme yeteneğini artırmada ve aşırı öğrenme problemlerini azaltmada kritik bir rol oynadığını ortaya koymaktadır. U-Net mimarisi [22] kullanılarak, farklı

kayıp fonksiyonları ile yapılan segmentasyon performansı karşılaştırması, CVC-ColonDB [119] veri seti için Tablo 4-5’de ve Kvasir-SEG [118] veri seti için Tablo 4-6’da sunulmuştur. Bu tablolarda, ilgili sütundaki performans metriğinde en yüksek değeri gösteren sonuçlar yeşil renkle, en düşük değeri gösteren sonuçlar ise mavi renkle belirtilmiştir.

Tablo 4-5: CVC-ColonDB [119] veri seti üzerinde kayıp fonksiyonları ile elde edilen performans metrikleri.

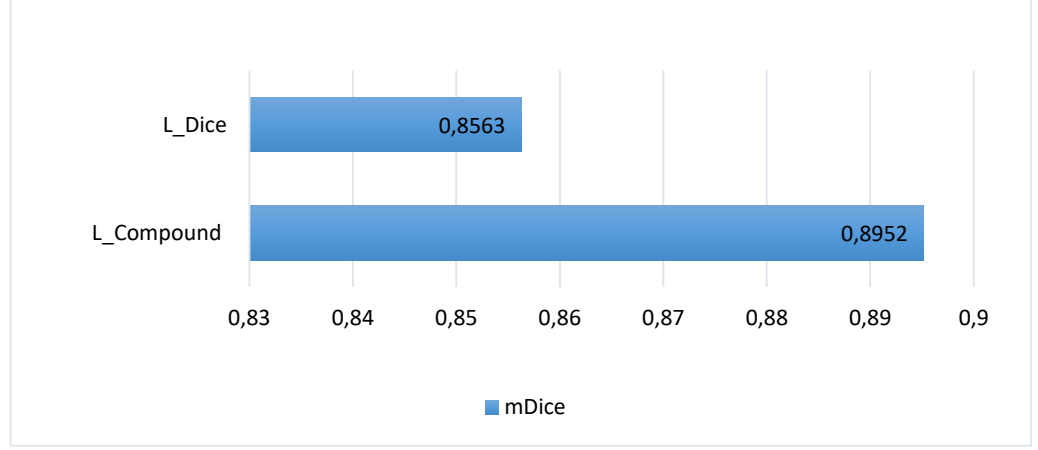
Metrik	Acc	R	Sp	Pr	mDice	mIoU	MAE
L_{BCE}	0.9761	0.8524	0.9916	0.9056	0.8782	0.8809	0.0240
L_{Dice}	0.9724	0.7702	0.9957	0.9447	0.8486	0.8563	0.0250
$L_{Tversky}$	0.9747	0.8339	0.9919	0.9099	0.8703	0.8734	0.0247
$L_{Compound}$	0.9731	0.8730	0.9913	0.9194	0.8956	0.8952	0.0227

Tablo 4-6: Kvasir-SEG [118] veri seti üzerinde kayıp fonksiyonları ile elde edilen performans metrikleri.

Metrik	Acc	R	Sp	Pr	mDice	mIoU	MAE
L_{BCE}	0.9313	0.6926	0.9807	0.8598	0.7672	0.7770	0.0639
L_{Dice}	0.9216	0.6508	0.9773	0.8339	0.7310	0.7487	0.0716
$L_{Tversky}$	0.9287	0.7674	0.9629	0.7677	0.7676	0.7787	0.0644
$L_{Compound}$	0.9219	0.6231	0.9831	0.8668	0.7250	0.7452	0.0716

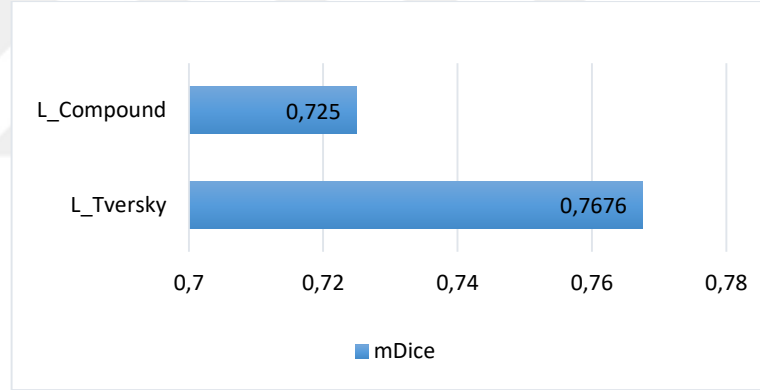
Sonuçlar değerlendirildiğinde, $L_{Tversky}$ Kayıp Fonksiyonunda, özellikle yüksek R ve $mIoU$ değerleri ile dikkat çekmiştir. Bu durum, segmentasyon kalitesinde kayda değer iyileşmeler sağladığını göstermektedir. $L_{Compound}$ Kayıp Fonksiyonu ise düşük MAE değeri sağlayarak, tahminlerin doğruluğu ve tutarlılığı açısından önemli bir avantaj sunmaktadır.

Elde edilen sonuçlara göre, CVC-ColonDB [119] veri setinde $L_{Compound}$ Kayıp Fonksiyonu ile en yüksek $mDice$ değeri elde edilirken; en düşük $mDice$ değeri L_{Dice} Kayıp Fonksiyonu ile elde edilmiştir. Karşılaştırılmalı analizi Şekil 4-5’de gösterilmektedir.



Şekil 4-5: CVC-ColonDB veri seti için modeller arası *mDice* karşılaştırılması

Kvasir-SEG [118] veri setinde $L_{Tversky}$ Kayıp Fonksiyonu ile elde edilen en yüksek *mDice* değeri 0,7676; en düşük *mDice* değeri ile $L_{Compound}$ Kayıp Fonksiyonunda elde edilmiştir. Şekil 4-6’da karşılaştırılması sunulmaktadır.



Şekil 4-6: Kvasir-SEG veri seti için modeller arası *mDice* karşılaştırılması

Ayrıca $L_{Tversky}$ Kayıp Fonksiyonu başlangıçta $\lambda_1 = 0.7$ değeriyle kullanılırken, performans üzerindeki etkileri gözlemlemek amacıyla λ_1 değeri 0.3’e düşürülmüştür. Bu düşüş sonucunda *mDice* ve *mIoU* metriklerinde kısmi bir azalma kaydedilmiştir. Bu durum, model üzerinde daha yüksek bir λ değeri seçmenin daha uygun olabileceğini göstermektedir.

5. SONUÇ VE ÖNERİLER

Polipler, normal olmayan doku büyümeleri olup, zamanla kötü huylu yapıya dönüşme potansiyeline sahiptirler. Bu durum, kolorektal kansere yol açmaktadır. Yüksek ölüm oranları nedeniyle kolorektal kanser ciddi bir sorun teşkil etmektedir. Poliplerin erken tespit edilmesi, kolorektal kanser kaynaklı ölüm oranlarının azaltılması açısından büyük önem arz etmektedir. Bu tez kapsamında, kolonoskopi görüntüleri üzerinde poliplerin segmentasyonu için geliştirilen yöntemler, erken teşhis ve tedavi süreçlerinin iyileştirilmesine katkı sağlamayı hedeflemektedir.

Bu tez çalışmasında, kolorektal poliplerin segmentasyonu için derin öğrenme tabanlı iki farklı uygulama geliştirilmiştir. İlk uygulamada, DeepLabv3+ mimarisi [115], DenseNet mimarisi [110], ResNet50 mimarisi, SqueezeNet mimarisi [112] ve VGG16 mimarisi [113] gibi farklı omurgalar ile güçlendirilmiş ve Kvasir-SEG [118] veri seti ve CVC-ClinicDB [20] veri setleri üzerinde test edilmiştir. Kvasir-SEG [118] veri seti ile elde edilen *mDice* (0,858) ve *mIoU* (0,850) değerleri DenseNet mimarisi [110] ile en yüksek performansı göstermiştir. Bu değerler, DenseNet mimarisinin [110] özellik çıkarımındaki başarısını ortaya koymaktadır. En düşük performans ise, SqueezeNet mimarisi [112] ile *mDice* (0,667) ve *mIoU* (0,695) değerleriyle kaydedilmiştir. Bu durum ise, SqueezeNet mimarisinin [112] daha az parametreye sahip yapısı ile poliplerin segmente işleminde sınırlamalar getirdiğini göstermektedir. CVC-ClinicDB [20] veri setinde en yüksek *mDice* (0,914) ve *mIoU* (0,912) değerleri ResNet50 mimarisi ile elde edilirken, en düşük performans *mDice* (0,847) ve *mIoU* (0,855) değerleriyle SqueezeNet mimarisi [112] tarafından gözlemlenmiştir. Bu sonuçlar doğrultusunda, ResNet50 mimarisinin artık bloklar sayesinde derin katmanlarda üstün performans sergilediği; buna karşın SqueezeNet mimarisinin [112] daha az karmaşık yapısının, daha zorlu veri setlerinde performansını olumsuz etkilediği ortaya konmuştur.

İkinci uygulamada ise, U-Net mimarisi [22] kolorektal kanseri görüntü segmentasyonunda kayıp fonksiyonlarının etkilerini incelemektedir. Kayıp fonksiyonları, CVC-ColonDB [119] veri seti ve Kvasir-SEG [118] veri setleri üzerinde değerlendirilmiştir. Değerlendirme sonuçlarına göre CVC-ColonDB [119] veri setinde, en yüksek performans *R* (0,8730), *mIoU* (0,8952) ve *mDice* (0,8956) değerleriyle Bileşik Kayıp fonksiyonu tarafından elde edilmiştir. Kvasir-SEG [118] veri setinde ise en yüksek değerler *Acc* (0,9313), *mIoU* (0, 7770) ve *mDice* (0,7672) değerleriyle İkili Çapraz Entropi Kayıp Fonksiyonu [104] ile kaydedilmiştir. Bu sonuçlar farklı kayıp fonksiyonlarının değişen performanslarını orta koymaktadır. Bileşik Kayıp fonksiyonunun yüksek bir performans göstermesi, farklı kayıp fonksiyonlarının avantajlarını bir araya getirerek segmentasyon başarısını artırdığını göstermektedir. Kvasir-SEG [118] veri setinde elde edilen

yüksek performans sonuçları, İkili Çapraz Entropi Kayıp Fonksiyonunun [104] belirli veri setlerinde iyi bir performans sunduğunu ortaya koymaktadır. Bu sonuçlar, polip segmentasyonunda kayıp fonksiyonları ve mimari seçimlerin performans üzerinde önemli bir etkiye sahip olduğunu göstermektedir.

Gelecekteki çalışmalarda, literatürde yer alan mimarilerin kısıtlamaları ele alınarak iyileştirilmesi önem taşımaktadır. Veri setlerinin çeşitlendirilmesi ve özellikle mimarilerinin genelleme yeteneğini artırmak için yenilikler önerilmektedir. DeepLabv3+ mimarisinde omurga ağı olarak çeşitli mimariler tercih edilebilir; bu durum hesaplama maliyeti ve performans açısından çeşitli avantajlar sunabilir. Gelecekteki çalışmalar için daha gelişmiş kayıp fonksiyonları ve veri artırma tekniklerinin kullanılması, segmentasyon doğruluğunu artırmak için kullanılabilir.



6. KAYNAKLAR

- [1] L. Kapusuz, Z., Özkırış, M., Sakallıoğlu, Ö., & Saydam, “NAZAL POLİPLER,” *Bozok Tıp Derg.*, vol. 2, pp. 35–41, 2012.
- [2] F. ÇELİK *et al.*, “Endoskopik incelemede saptanan mide poliplerinin yerleşimi, histopatolojik tipleri ve tedavi yönetimi,” *Akad. Gastroenteroloji Derg.*, vol. 20, no. 3, pp. 129–135, Jan. 2022, doi: 10.17941/agd.1055273.
- [3] Y. ŞAHİNTÜRK and A. H. ÇEKİN, “Kolon polipleri: Lokalizasyon, histoloji, boyut - 5 yıllık kolonoskopik değerlendirme,” *Endosk. Gastrointest.*, vol. 26, no. 2, pp. 57–60, Sep. 2018, doi: 10.17940/endoskopi.461413.
- [4] M. D. C. Vieira, A. Vitagliano, M. C. Rossette, L. C. de A. Neto, A. Gallo, and A. D. S. Sardo, “Endometrial Polyps: Update Overview on Etiology, Diagnosis, Natural History and Treatment,” *Clin. Exp. Obstet. Gynecol.*, vol. 49, no. 10, Sep. 2022, doi: 10.31083/j.ceog4910232.
- [5] H. Sung *et al.*, “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” *CA. Cancer J. Clin.*, vol. 71, no. 3, pp. 209–249, May 2021, doi: 10.3322/caac.21660.
- [6] “No Title.” [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/colorectal-cancer>
- [7] “No Title.” [Online]. Available: <https://hsgm.saglik.gov.tr/tr/kanser-istatistikleri>
- [8] N. Forbes, “Impact of Screening-related Sigmoidoscopy on the Incidence and Mortality of Colorectal Cancer,” *Gastroenterology*, vol. 165, no. 3, p. 802, Sep. 2023, doi: 10.1053/j.gastro.2023.03.210.
- [9] I. A. Issa and M. Nouredine, “Colorectal cancer screening: An updated review of the available options,” *World J. Gastroenterol.*, vol. 23, no. 28, p. 5086, 2017, doi: 10.3748/wjg.v23.i28.5086.
- [10] X. Yang, Q. Wei, C. Zhang, K. Zhou, L. Kong, and W. Jiang, “Colon Polyp Detection and Segmentation Based on Improved MRCNN,” *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021, doi: 10.1109/TIM.2020.3038011.
- [11] P. Wang *et al.*, “Lower Adenoma Miss Rate of Computer-Aided Detection-Assisted Colonoscopy vs Routine White-Light Colonoscopy in a Prospective Tandem Study,” *Gastroenterology*, vol. 159, no. 4, pp. 1252–1261.e5, Oct. 2020, doi: 10.1053/j.gastro.2020.06.023.

- [12] K. Lin, J. Hong, Y. Huang, Y. Yu, and R.-G. Tsai, "PD-YOLO: Colon Polyp Detection Model Based on Enhanced Small-Target Feature Extraction," *Comput. Mater. Contin.*, vol. 82, no. 1, pp. 913–928, 2025, doi: 10.32604/cmc.2024.058467.
- [13] X. Yang *et al.*, "YOLO-OB: An improved anchor-free real-time multiscale colon polyp detector in colonoscopy," *Biomed. Signal Process. Control*, vol. 103, p. 107326, May 2025, doi: 10.1016/j.bspc.2024.107326.
- [14] M. D. Manshadi, M. Mousavi, A. Goltzarian, M. Soltani, and A. Mosavi, "Colorectal Polyp Localization: From Image Restoration to Real-time Detection with Deep Learning," in *2023 IEEE 17th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, IEEE, May 2023, pp. 739–744. doi: 10.1109/SACI58269.2023.10158661.
- [15] C. Xu *et al.*, "CE-YOLO: A channel-efficient YOLO algorithm for colorectal polyp detection," *Biomed. Signal Process. Control*, vol. 104, p. 107529, Jun. 2025, doi: 10.1016/j.bspc.2025.107529.
- [16] B. Ay, C. Turker, E. Emre, K. Ay, and G. Aydin, "Automated classification of nasal polyps in endoscopy video-frames using handcrafted and CNN features," *Comput. Biol. Med.*, vol. 147, p. 105725, Aug. 2022, doi: 10.1016/j.compbiomed.2022.105725.
- [17] S. Pratik, P. Sharma, D. R. Nayak, and B. K. Balabantaray, "WMCF-Net: Wavelet pooling-based multiscale contextual fusion network for polyp classification," *Biomed. Signal Process. Control*, vol. 107, p. 107727, Sep. 2025, doi: 10.1016/j.bspc.2025.107727.
- [18] M. M. Hasan and M. M. Rahman, "Privacy-preserving polyp segmentation using federated learning with differential privacy," *Smart Heal.*, vol. 36, p. 100551, Jun. 2025, doi: 10.1016/j.smhl.2025.100551.
- [19] T. Mahmud, B. Paul, and S. A. Fattah, "PolypSegNet: A modified encoder-decoder architecture for automated polyp segmentation from colonoscopy images," *Comput. Biol. Med.*, vol. 128, p. 104119, Jan. 2021, doi: 10.1016/j.compbiomed.2020.104119.
- [20] J. Bernal, F. J. Sánchez, G. Fernández-Esparrach, D. Gil, C. Rodríguez, and F. Vilariño, "WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians," *Comput. Med. Imaging Graph.*, vol. 43, pp. 99–111, Jul. 2015, doi: 10.1016/j.compmedimag.2015.02.007.
- [21] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," Nov. 2014, [Online]. Available: <http://arxiv.org/abs/1411.4038>
- [22] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical

- Image Segmentation,” 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.
- [23] E. Sanderson and B. J. Matuszewski, “FCN-Transformer Feature Fusion for Polyp Segmentation,” 2022, pp. 892–907. doi: 10.1007/978-3-031-12053-4_65.
- [24] S. Zhang, P. Rao, X. Chen, and Y. Li, “An FCN-based transfer-learning method for spatial infrared moving-target recognition,” *Infrared Phys. Technol.*, vol. 137, p. 105145, Mar. 2024, doi: 10.1016/j.infrared.2024.105145.
- [25] S. Barın and G. E. Güraksın, “An automatic skin lesion segmentation system with hybrid FCN-ResAlexNet,” *Eng. Sci. Technol. an Int. J.*, vol. 34, p. 101174, Oct. 2022, doi: 10.1016/j.jestch.2022.101174.
- [26] Ü. Budak, Z. Cömert, Z. N. Rashid, A. Şengür, and M. Çıbuk, “Computer-aided diagnosis system combining FCN and Bi-LSTM model for efficient breast cancer detection from histopathological images,” *Appl. Soft Comput.*, vol. 85, p. 105765, Dec. 2019, doi: 10.1016/j.asoc.2019.105765.
- [27] S. Kamal *et al.*, “FCN Network-Based Weed and Crop Segmentation for IoT-Aided Agriculture Applications,” *Wirel. Commun. Mob. Comput.*, vol. 2022, pp. 1–10, May 2022, doi: 10.1155/2022/2770706.
- [28] A. Taravat, M. P. Wagner, R. Bonifacio, and D. Petit, “Advanced Fully Convolutional Networks for Agricultural Field Boundary Detection,” *Remote Sens.*, vol. 13, no. 4, p. 722, Feb. 2021, doi: 10.3390/rs13040722.
- [29] B. Shuai, T. Liu, and G. Wang, “Improving Fully Convolution Network for Semantic Segmentation,” 2016, [Online]. Available: <http://arxiv.org/abs/1611.08986>
- [30] O. Ozturk, B. Sarıtürk, and D. Z. Seker, “Comparison of Fully Convolutional Networks (FCN) and U-Net for Road Segmentation from High Resolution Imageries,” *Int. J. Environ. Geoinformatics*, vol. 7, no. 3, pp. 272–279, Dec. 2020, doi: 10.30897/ijegeo.737993.
- [31] D. Liu, Z. Wang, and A. Liang, “MiM-UNet: An efficient building image segmentation network integrating state space models,” *Alexandria Eng. J.*, vol. 120, pp. 648–656, May 2025, doi: 10.1016/j.aej.2025.02.035.
- [32] C. Zhang, Y. Zhang, and X. Xu, “Dilated inception U-Net with attention for crop pest image segmentation in real-field environment,” *Smart Agric. Technol.*, vol. 11, p. 100917, Aug. 2025, doi: 10.1016/j.atech.2025.100917.
- [33] L. S V V, S. E. Darmanayagam, and S. R. R. Cyril, “Semantic Segmentation of Lumbar Vertebrae Using Meijering U-Net (MU-Net) on Spine Magnetic Resonance Images,”

- Comput. Model. Eng. Sci.*, vol. 142, no. 1, pp. 733–757, 2025, doi: 10.32604/cmesci.2024.056424.
- [34] H. Gao, J. Gu, S. Liu, S. Xu, and Q. Zhao, “MSFA-Net: Multi-scale feature aggregation and attention-enhanced U-Net for microscopic hyperspectral pathology images segmentation,” *Opt. Laser Technol.*, vol. 187, p. 112652, Sep. 2025, doi: 10.1016/j.optlastec.2025.112652.
- [35] A. Garbaz, Y. Oukdach, S. Charfi, M. El Ansari, L. Koutti, and M. Salihoun, “MLFA-UNet: A multi-level feature assembly UNet for medical image segmentation,” *Methods*, vol. 232, pp. 52–64, Dec. 2024, doi: 10.1016/j.ymeth.2024.10.010.
- [36] Q. Liu and J. Zhao, “MA-Res U-Net: Design of Soybean Navigation System with Improved U-Net Model,” *Phyton (B. Aires)*, vol. 93, no. 10, pp. 2663–2681, 2024, doi: 10.32604/phyton.2024.056054.
- [37] Z. Guo *et al.*, “InstaCropNet: An efficient Unet-Based architecture for precise crop row detection in agricultural applications,” *Artif. Intell. Agric.*, vol. 12, pp. 85–96, Jun. 2024, doi: 10.1016/j.aiia.2024.05.002.
- [38] D. ARIKAN and F. YILDIZ, “Göktürk-1 Uydu Görüntülerinden U-Net Modeli Kullanılarak Binaların Çıkarılması,” *Turkish J. Remote Sens.*, Jun. 2023, doi: 10.51489/tuzal.1300939.
- [39] P. Ulmas and I. Liiv, “Segmentation of Satellite Imagery using U-Net Models for Land Cover Classification,” 2020, [Online]. Available: <http://arxiv.org/abs/2003.02899>
- [40] T. Hussain, H. Shouno, M. A. Mohammed, H. A. Marhoon, and T. Alam, “DCSSGA-UNet: Biomedical image segmentation with DenseNet channel spatial and Semantic Guidance Attention,” *Knowledge-Based Syst.*, vol. 314, p. 113233, Apr. 2025, doi: 10.1016/j.knosys.2025.113233.
- [41] W. Dai *et al.*, “Exploiting Scale-Variant Attention for Segmenting Small Medical Objects,” 2024, [Online]. Available: <http://arxiv.org/abs/2407.07720>
- [42] M. Corbetta and G. L. Shulman, “Control of goal-directed and stimulus-driven attention in the brain,” *Nat. Rev. Neurosci.*, vol. 3, no. 3, pp. 201–215, Mar. 2002, doi: 10.1038/nrn755.
- [43] Y. Akgöl, B. Toptaş, and M. Toptaş, “Polyp segmentation with colonoscopic images: a study,” *Neural Comput. Appl.*, Mar. 2025, doi: 10.1007/s00521-025-11144-2.
- [44] D. Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., ... & Rueckert, “Attention u-net: Learning where to look for the pancreas”.

- [45] M. T. Luong, H. Pham, and C. D. Manning, “Effective approaches to attention-based neural machine translation,” *Conf. Proc. - EMNLP 2015 Conf. Empir. Methods Nat. Lang. Process.*, pp. 1412–1421, 2015, doi: 10.18653/v1/d15-1166.
- [46] L. Lu, S. Chen, H. Tang, X. Zhang, and X. Hu, “A multi-scale perceptual polyp segmentation network based on boundary guidance,” *Image Vis. Comput.*, vol. 138, p. 104811, Oct. 2023, doi: 10.1016/j.imavis.2023.104811.
- [47] T. M. Angona and M. R. H. Mondal, “An attention based residual U-Net with swin transformer for brain MRI segmentation,” *Array*, vol. 25, p. 100376, Mar. 2025, doi: 10.1016/j.array.2025.100376.
- [48] Z. Wang, Y. Zou, and P. X. Liu, “Hybrid dilation and attention residual U-Net for medical image segmentation,” *Comput. Biol. Med.*, vol. 134, p. 104449, Jul. 2021, doi: 10.1016/j.compbimed.2021.104449.
- [49] X. Shu, X. Li, X. Zhang, C. Shao, X. Yan, and S. Huang, “MRAU-net: Multi-scale residual attention U-shaped network for medical image segmentation,” *Comput. Electr. Eng.*, vol. 118, p. 109479, Sep. 2024, doi: 10.1016/j.compeleceng.2024.109479.
- [50] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, “CBAM: Convolutional block attention module,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11211 LNCS, pp. 3–19, 2018, doi: 10.1007/978-3-030-01234-2_1.
- [51] G. Yang, J. Ren, P., Zhang, D., Chen, D., Wen, F., Li, H., & Hua, “Neural aggregation network for video face recognition,” *Proc. IEEE Conf. Comput. Vis. pattern Recognit.*, pp. 4362–4371, 2017.
- [52] A. Zhang, H., Goodfellow, I., Metaxas, D., & Odena, “Self-attention generative adversarial networks,” *Int. Conf. Mach. Learn.*, pp. 7354–7363, 2019.
- [53] A. Vaswani *et al.*, “Attention is all you need,” *Adv. Neural Inf. Process. Syst.*, vol. 2017-Decem, pp. 5999–6009, 2017.
- [54] A. Dosovitskiy *et al.*, “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” *Int. Conf. Learn. Represent.*, Oct. 2021, [Online]. Available: <http://arxiv.org/abs/2010.11929>
- [55] Y. Tay, M. Dehghani, D. Bahri, and D. Metzler, “Efficient Transformers: A Survey,” *ACM Comput. Surv.*, vol. 55, no. 6, pp. 1–28, Jun. 2022, doi: 10.1145/3530811.
- [56] Z. Liu *et al.*, “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,” *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 10012–10022, 2021, doi: 10.1109/ICCV48922.2021.00986.

- [57] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, “End-to-End Object Detection with Transformers,” 2020, pp. 213–229. doi: 10.1007/978-3-030-58452-8_13.
- [58] A. Hatamizadeh *et al.*, “UNETR: Transformers for 3D Medical Image Segmentation,” *Proc. - 2022 IEEE/CVF Winter Conf. Appl. Comput. Vision, WACV 2022*, pp. 574–584, 2022.
- [59] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998, doi: 10.1109/5.726791.
- [60] O. Russakovsky *et al.*, “ImageNet Large Scale Visual Recognition Challenge,” *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y.
- [61] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Adv. Neural Inf. Process. Syst.*, p. 8, 2012.
- [62] C.-H. Huang, H.-Y. Wu, and Y.-L. Lin, “HarDNet-MSEG: A Simple Encoder-Decoder Polyp Segmentation Neural Network that Achieves over 0.9 Mean Dice and 86 FPS,” 2021, [Online]. Available: <http://arxiv.org/abs/2101.07172>
- [63] C. Wang, R. Xu, S. Xu, W. Meng, and X. Zhang, “Automatic polyp segmentation via image-level and surrounding-level context fusion deep neural network,” *Eng. Appl. Artif. Intell.*, vol. 123, p. 106168, Aug. 2023, doi: 10.1016/j.engappai.2023.106168.
- [64] R. Zhang, G. Li, Z. Li, S. Cui, D. Qian, and Y. Yu, “Adaptive Context Selection for Polyp Segmentation,” 2020, pp. 253–262. doi: 10.1007/978-3-030-59725-2_25.
- [65] Y. Su, J. Cheng, M. Yi, and H. Liu, “FAPN: Feature Augmented Pyramid Network for polyp segmentation,” *Biomed. Signal Process. Control*, vol. 78, p. 103903, Sep. 2022, doi: 10.1016/j.bspc.2022.103903.
- [66] O. Gökkan and M. Kuntalp, “A new imbalance-aware loss function to be used in a deep neural network for colorectal polyp segmentation,” *Comput. Biol. Med.*, vol. 151, p. 106205, Dec. 2022, doi: 10.1016/j.compbimed.2022.106205.
- [67] Y. Xia, H. Yun, and Y. Liu, “MFEFNet: Multi-scale feature enhancement and Fusion Network for polyp segmentation,” *Comput. Biol. Med.*, vol. 157, p. 106735, May 2023, doi: 10.1016/j.compbimed.2023.106735.
- [68] T. Zhou *et al.*, “Cross-level Feature Aggregation Network for Polyp Segmentation,” *Pattern Recognit.*, vol. 140, p. 109555, Aug. 2023, doi: 10.1016/j.patcog.2023.109555.
- [69] J. Hu, Y. Xu, and Z. Tang, “DAN-PD: Domain adaptive network with parallel decoder for

- polyp segmentation,” *Comput. Med. Imaging Graph.*, vol. 101, p. 102124, Oct. 2022, doi: 10.1016/j.compmedimag.2022.102124.
- [70] Y. Su, J. Cheng, C. Zhong, C. Jiang, J. Ye, and J. He, “Accurate polyp segmentation through enhancing feature fusion and boosting boundary performance,” *Neurocomputing*, vol. 545, p. 126233, Aug. 2023, doi: 10.1016/j.neucom.2023.126233.
- [71] G. Liu, Y. Jiang, D. Liu, B. Chang, L. Ru, and M. Li, “A coarse-to-fine segmentation frame for polyp segmentation via deep and classification features,” *Expert Syst. Appl.*, vol. 214, p. 118975, Mar. 2023, doi: 10.1016/j.eswa.2022.118975.
- [72] K. Wickstrøm, M. Kampffmeyer, and R. Jenssen, “Uncertainty and interpretability in convolutional neural networks for semantic segmentation of colorectal polyps,” *Med. Image Anal.*, vol. 60, p. 101619, Feb. 2020, doi: 10.1016/j.media.2019.101619.
- [73] J. Zhu, M. Ge, Z. Chang, and W. Dong, “CRCNet: Global-local context and multi-modality cross attention for polyp segmentation,” *Biomed. Signal Process. Control*, vol. 83, p. 104593, May 2023, doi: 10.1016/j.bspc.2023.104593.
- [74] M. J. Alam and S. A. Fattah, “SR-AttNet: An Interpretable Stretch–Relax Attention based Deep Neural Network for Polyp Segmentation in Colonoscopy Images,” *Comput. Biol. Med.*, vol. 160, p. 106945, Jun. 2023, doi: 10.1016/j.compbiomed.2023.106945.
- [75] P. Song, J. Li, and H. Fan, “Attention based multi-scale parallel network for polyp segmentation,” *Comput. Biol. Med.*, vol. 146, p. 105476, Jul. 2022, doi: 10.1016/j.compbiomed.2022.105476.
- [76] L. Yang, C. Zhai, Y. Liu, and H. Yu, “CFHA-Net: A polyp segmentation method with cross-scale fusion strategy and hybrid attention,” *Comput. Biol. Med.*, vol. 164, p. 107301, Sep. 2023, doi: 10.1016/j.compbiomed.2023.107301.
- [77] Y. Lin, J. Wu, G. Xiao, J. Guo, G. Chen, and J. Ma, “BSCA-Net: Bit Slicing Context Attention network for polyp segmentation,” *Pattern Recognit.*, vol. 132, p. 108917, Dec. 2022, doi: 10.1016/j.patcog.2022.108917.
- [78] H. Xia, Y. Qin, Y. Tan, and S. Song, “BA-Net: Brightness prior guided attention network for colonic polyp segmentation,” *Biocybern. Biomed. Eng.*, vol. 43, no. 3, pp. 603–615, Jul. 2023, doi: 10.1016/j.bbe.2023.08.001.
- [79] T.-C. Nguyen, T.-P. Nguyen, G.-H. Diep, A.-H. Tran-Dinh, T. V. Nguyen, and M.-T. Tran, “CCBANet: Cascading Context and Balancing Attention for Polyp Segmentation,” 2021, pp. 633–643. doi: 10.1007/978-3-030-87193-2_60.
- [80] K. Yang *et al.*, “Automatic polyp detection and segmentation using shuffle efficient

- channel attention network,” *Alexandria Eng. J.*, vol. 61, no. 1, pp. 917–926, Jan. 2022, doi: 10.1016/j.aej.2021.04.072.
- [81] R. Nachmani, I. Nidal, D. Robinson, M. Yassin, and D. Abookasis, “Segmentation of polyps based on pyramid vision transformers and residual block for real-time endoscopy imaging,” *J. Pathol. Inform.*, vol. 14, p. 100197, 2023, doi: 10.1016/j.jpi.2023.100197.
 - [82] W. Li, Y. Zhao, F. Li, and L. Wang, “MIA-Net: Multi-information aggregation network combining transformers and convolutional feature learning for polyp segmentation,” *Knowledge-Based Syst.*, vol. 247, p. 108824, Jul. 2022, doi: 10.1016/j.knosys.2022.108824.
 - [83] C. Wu, C. Long, S. Li, J. Yang, F. Jiang, and R. Zhou, “MSRAformer: Multiscale spatial reverse attention network for polyp segmentation,” *Comput. Biol. Med.*, vol. 151, p. 106274, Dec. 2022, doi: 10.1016/j.compbiomed.2022.106274.
 - [84] F. Liu, Z. Hua, J. Li, and L. Fan, “DBMF: Dual Branch Multiscale Feature Fusion Network for polyp segmentation,” *Comput. Biol. Med.*, vol. 151, p. 106304, Dec. 2022, doi: 10.1016/j.compbiomed.2022.106304.
 - [85] K. Hu, W. Chen, Y. Sun, X. Hu, Q. Zhou, and Z. Zheng, “PPNet: Pyramid pooling based network for polyp segmentation,” *Comput. Biol. Med.*, vol. 160, p. 107028, Jun. 2023, doi: 10.1016/j.compbiomed.2023.107028.
 - [86] J. Wang, S. Tian, L. Yu, Z. Zhou, F. Wang, and Y. Wang, “HIGF-Net: Hierarchical information-guided fusion network for polyp segmentation based on transformer and convolution feature learning,” *Comput. Biol. Med.*, vol. 161, p. 107038, Jul. 2023, doi: 10.1016/j.compbiomed.2023.107038.
 - [87] N. T. Duc, N. T. Oanh, N. T. Thuy, T. M. Triet, and V. S. Dinh, “ColonFormer: An Efficient Transformer Based Method for Colon Polyp Segmentation,” *IEEE Access*, vol. 10, pp. 80575–80586, 2022, doi: 10.1109/ACCESS.2022.3195241.
 - [88] T. M. Khan, M. Arsalan, I. Razzak, and E. Meijering, “Simple and robust depth-wise cascaded network for polyp segmentation,” *Eng. Appl. Artif. Intell.*, vol. 121, p. 106023, May 2023, doi: 10.1016/j.engappai.2023.106023.
 - [89] F. Liu, Z. Hua, J. Li, and L. Fan, “MFBGR: Multi-scale feature boundary graph reasoning network for polyp segmentation,” *Eng. Appl. Artif. Intell.*, vol. 123, p. 106213, Aug. 2023, doi: 10.1016/j.engappai.2023.106213.
 - [90] B. Dong, W. Wang, D.-P. Fan, J. Li, H. Fu, and L. Shao, “Polyp-PVT: Polyp Segmentation with Pyramid Vision Transformers,” *CAAI Artif. Intell. Res.*, p. 9150015,

Dec. 2023, doi: 10.26599/AIR.2023.9150015.

- [91] P. D. Leufkens, A. M., Van Oijen, M. G. H., Vleggaar, F. P., & Siersema, “Factors influencing the miss rate of polyps in a back-to-back colonoscopy study,” *Endoscopy*, vol. 470–475, 2012.
- [92] A. D. Muller and A. Sonnenberg, “Prevention of Colorectal Cancer by Flexible Endoscopy and Polypectomy: A Case-Control Study of 32 702 Veterans,” *Ann. Intern. Med.*, vol. 123, no. 12, pp. 904–910, Dec. 1995, doi: 10.7326/0003-4819-123-12-199512150-00002.
- [93] V. S. Dhaka *et al.*, “A Survey of Deep Convolutional Neural Networks Applied for Prediction of Plant Leaf Diseases,” *Sensors*, vol. 21, no. 14, p. 4749, Jul. 2021, doi: 10.3390/s21144749.
- [94] S. Alghyaline, “Optimised CNN Architectures for Handwritten Arabic Character Recognition,” *Comput. Mater. Contin.*, vol. 79, no. 3, pp. 4905–4924, 2024, doi: 10.32604/cmc.2024.052016.
- [95] L. Schneider, A. Krasowski, V. Pitchika, L. Bombeck, F. Schwendicke, and M. Büttner, “Assessment of CNNs, transformers, and hybrid architectures in dental image segmentation,” *J. Dent.*, vol. 156, p. 105668, May 2025, doi: 10.1016/j.jdent.2025.105668.
- [96] H. Jiang, H. Ma, W. Qian, M. Gao, and Y. Li, “An Automatic Detection System of Lung Nodule Based on Multigroup Patch-Based Deep Learning Network,” *IEEE J. Biomed. Heal. Informatics*, vol. 22, no. 4, pp. 1227–1237, Jul. 2018, doi: 10.1109/JBHI.2017.2725903.
- [97] J. Bertels, D. Robben, R. Lemmens, and D. Vandermeulen, “Convolutional neural networks for medical image segmentation,” Nov. 2022, [Online]. Available: <http://arxiv.org/abs/2211.09562>
- [98] K. Zhao, J. Xiao, C. Li, Z. Xu, and M. Yue, “Fault diagnosis of rolling bearing using CNN and PCA fractal based feature extraction,” *Measurement*, vol. 223, p. 113754, Dec. 2023, doi: 10.1016/j.measurement.2023.113754.
- [99] J. Ma *et al.*, “Loss odyssey in medical image segmentation,” *Med. Image Anal.*, vol. 71, p. 102035, Jul. 2021, doi: 10.1016/j.media.2021.102035.
- [100] R. Dakhli and W. Barhoumi, “Improving skin lesion classification through saliency-guided loss functions,” *Comput. Biol. Med.*, vol. 192, p. 110299, Jun. 2025, doi: 10.1016/j.compbiomed.2025.110299.

- [101] H.-C. Park, S. Poudel, R. Ghimire, and S.-W. Lee, "Polyp segmentation with consistency training and continuous update of pseudo-label," *Sci. Rep.*, vol. 12, no. 1, p. 14626, Aug. 2022, doi: 10.1038/s41598-022-17843-3.
- [102] A. Qayoom, J. Xie, and H. Ali, "Polyp segmentation in medical imaging: challenges, approaches and future directions," *Artif. Intell. Rev.*, vol. 58, no. 6, p. 169, Mar. 2025, doi: 10.1007/s10462-025-11173-2.
- [103] J. Long, J. Lin, J. Liu, and L. Liu, "Toward Robust Segmentation of Polyp via Box-supervised and Feature-Embedded," *Arab. J. Sci. Eng.*, Nov. 2024, doi: 10.1007/s13369-024-09762-4.
- [104] Usha Ruby Dr.A, "Binary cross entropy with deep learning technique for Image classification," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 4, pp. 5393–5397, Aug. 2020, doi: 10.30534/ijatcse/2020/175942020.
- [105] V. O. Yazici, J. van de Weijer, and A. Ramisa, "Color Naming for Multi-color Fashion Items," 2018, pp. 64–73. doi: 10.1007/978-3-319-77700-9_7.
- [106] Q. Guo, C. Wang, D. Xiao, and Q. Huang, "A novel multi-label pest image classifier using the modified Swin Transformer and soft binary cross entropy loss," *Eng. Appl. Artif. Intell.*, vol. 126, p. 107060, Nov. 2023, doi: 10.1016/j.engappai.2023.107060.
- [107] A. Tversky, "Features of similarity.," *Psychol. Rev.*, vol. 84, no. 4, pp. 327–352, Jul. 1977, doi: 10.1037/0033-295X.84.4.327.
- [108] N. Ottakath *et al.*, "MSEUnet: Refined Intima-media segmentation of the carotid artery based on a multi-scale approach using patch-wise dice loss," *Biomed. Signal Process. Control*, vol. 100, p. 107077, Feb. 2025, doi: 10.1016/j.bspc.2024.107077.
- [109] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, no. 3, pp. 297–302, Jul. 1945, doi: 10.2307/1932409.
- [110] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 2261–2269, Aug. 2016, doi: 10.1109/CVPR.2017.243.
- [111] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.
- [112] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," Feb. 2016, [Online]. Available: <http://arxiv.org/abs/1602.07360>

- [113] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” Sep. 2014, [Online]. Available: <http://arxiv.org/abs/1409.1556>
- [114] Y. Akgöl and B. Toptaş, “Polyp Segmentation with Deep Learning: Utilizing DeeplabV3+ Architecture and Various CNN Backbones,” *DÜMF Mühendislik Derg.*, Oct. 2024, doi: 10.24012/dumf.1517112.
- [115] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-decoder with atrous separable convolution for semantic image segmentation,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11211 LNCS, pp. 833–851, 2018, doi: 10.1007/978-3-030-01234-2_49.
- [116] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, “Rethinking Atrous Convolution for Semantic Image Segmentation,” 2017, [Online]. Available: <http://arxiv.org/abs/1706.05587>
- [117] K. Wang, Y. Lu, J. Liu, and X. Zhang, “GAML: Geometry-Aware Mutual Learning for polyp segmentation,” *Biomed. Signal Process. Control*, vol. 108, p. 107965, Oct. 2025, doi: 10.1016/j.bspc.2025.107965.
- [118] D. Jha *et al.*, “Kvasir-SEG: A Segmented Polyp Dataset,” 2020, pp. 451–462. doi: 10.1007/978-3-030-37734-2_37.
- [119] D. Vázquez *et al.*, “A Benchmark for Endoluminal Scene Segmentation of Colonoscopy Images,” *J. Healthc. Eng.*, vol. 2017, pp. 1–9, 2017, doi: 10.1155/2017/4037190.
- [120] N. Tajbakhsh, S. R. Gurudu, and J. Liang, “Automated Polyp Detection in Colonoscopy Videos Using Shape and Context Information,” *IEEE Trans. Med. Imaging*, vol. 35, no. 2, pp. 630–644, Feb. 2016, doi: 10.1109/TMI.2015.2487997.