

**BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ TEZLİ YÜKSEK LİSANS
PROGRAMI**

**TİROİD ULTRASONOGRAFİSİNDE HİBRİT DERİN ÖĞRENME
YÖNTEMLERİ İLE OTOMATİK NODÜL SEGMENTASYONU**

HAZIRLAYAN

ELİFNUR ERDEMİR AYTEKİN

YÜKSEK LİSANS TEZİ

ANKARA- 2025

**BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ TEZLİ YÜKSEK LİSANS
PROGRAMI**

**TİROİD ULTRASONOGRAFİSİNDE HİBRİT DERİN ÖĞRENME
YÖNTEMLERİ İLE OTOMATİK NODÜL SEGMENTASYONU**

HAZIRLAYAN

ELİFNUR ERDEMİR AYTEKİN

YÜKSEK LİSANS TEZİ

TEZ DANIŞMANI

DR. ÖĞR. ÜYESİ ÇAĞATAY BERKE ERDAŞ

ANKARA - 2025

BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Bilgisayar Mühendisliği Anabilim Dalı Bilgisayar Mühendisliği Tezli Yüksek Lisans Programı çerçevesinde Elifnur ERDEMİR AYTEKİN tarafından hazırlanan bu çalışma, aşağıdaki jüri tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 15 / 08 / 2025

Tez Adı: Tiroid Ultrasonografisinde Hibrit Derin Öğrenme Yöntemleri ile Otomatik Nodül Segmentasyonu

Tez Jüri Üyeleri (Unvanı, Adı - Soyadı, Kurumu)		İmza
Doç. Dr. Emre SÜMER	Başkent Üniversitesi	
Dr. Öğr. Üyesi İbrahim Baran USLU	OSTİM Teknik Üniversitesi	
Dr. Öğr. Üyesi Çağatay Berke ERDAŞ	Başkent Üniversitesi	

ONAY

Prof. Dr. Dilek ÇÖKELİLER SERDAROĞLU

Fen Bilimleri Enstitüsü Müdürü

Tarih : ... / ... /

BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLER ENSTİTÜSÜ
YÜKSEK LİSANS / DOKTORA TEZ ÇALIŞMASI ORJİNALLİK RAPORU

Tarih: 15 / 08 / 2025

Öğrencinin Adı, Soyadı : Elifnur ERDEMİR AYTEKİN

Öğrencinin Numarası : 22220154

Anabilim Dalı : Bilgisayar Mühendisliği Anabilim Dalı

Programı : Bilgisayar Mühendisliği Tezli Yüksek Lisans Programı

Danışmanın Unvanı/Adı, Soyadı : Dr. Öğr. Üyesi Çağatay Berke ERDAŞ

Tez Başlığı : Tiroid Ultrasonografisinde Hibrit Derin Öğrenme Yöntemleri ile Otomatik Nodül Segmentasyonu

Yukarıda başlığı belirtilen Yüksek Lisans tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam 62 sayfalık kısmına ilişkin, 15 / 08 / 2025 tarihinde şahsım/tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı %3'tür. Uygulanan filtrelemeler:

1. Kaynakça hariç
2. Alıntılar hariç
3. Beş (5) kelimeden daha az örtüşme içeren metin kısımları hariç

“Başkent Üniversitesi Enstitüleri Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Usul ve Esaslarını” inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Öğrenci İmzası:.....

ONAY

Tarih: ... / ... / 20...

Öğrenci Danışmanı:

Dr. Öğr. Üyesi Çağatay Berke ERDAŞ

TEŞEKKÜR

Tez çalışmam boyunca bana destek olan ve yol gösteren değerli danışmanım Dr. Öğr. Üyesi Çağatay Berke ERDAŞ'a teşekkür ederim. Bilgisi ve yönlendirmeleri, bu süreci daha verimli kıldı.

Hayatımın her alanında en büyük ilham kaynağım ve örnek aldığım sevgili annem Dr. Arzu ERDEMİR'e sonsuz minnettarım. Onun bilimsel disiplini, azmi ve özverisi benim akademik kariyerimdeki en büyük rehberim oldu. Babam Hakan ERDEMİR'in de sevgisi ve desteği, hayatımda her zaman sağlam bir dayanak oldu. Onların varlığı, beni bugünlere getiren en değerli güçtür.

Sevgili eşim Arş. Gör. Emre Kağan AYTEKİN'e, bana verdiği moral ve motivasyon için teşekkür ederim. Onun desteği ile bu yolculuk kolaylaştı.

Bu tez çalışmam boyunca yanımda olan, desteklerini esirgemeyen herkese içten teşekkürlerimi sunarım.

ÖZET

Elifnur ERDEMİR AYTEKİN

TİROİD ULTRASONOGRAFİSİNDE HİBRİT DERİN ÖĞRENME YÖNTEMLERİ İLE OTOMATİK NODÜL SEGMENTASYONU

Başkent Üniversitesi Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

2025

Tiroid nodülleri, tiroid bezinde normal dokudan farklı olarak gelişen, genellikle iyi huylu olmakla birlikte belirli bir oranda malign potansiyel taşıyan kitlelerdir. Toplumda oldukça yaygın görülen bu oluşumlar, ultrasonografi ile yüksek oranda tespit edilebilmektedir. Ancak nodüllerin doğru şekilde sınırlarının belirlenmesi hem malignite riskinin değerlendirilmesi hem de takip ve tedavi planlamasında kritik öneme sahiptir. Bu tez çalışmasında, tiroid ultrasonografisi görüntülerinden nodüllerin otomatik segmentasyonu amacıyla hibrit bir derin öğrenme modeli geliştirilmiştir. Önerilen yöntem, “EfficientNet” tabanlı yerel özellik çıkarım kapasitesini “Vision Transformer” tabanlı küresel bağlam farkındalığı ile birleştiren çift kodlayıcı mimariye sahiptir. Ek olarak nodül sınırlarının daha hassas belirlenebilmesi için sınır farkındalıklı özel bir kayıp fonksiyonu tasarlanmıştır. Modelin performansı TN3K ve DDTI veri kümeleri üzerinde kapsamlı olarak test edilmiş ve literatürdeki güncel yöntemlerle karşılaştırılmıştır. Elde edilen sonuçlar doğrultusunda geliştirilen hibrit yaklaşım, klinik uygulamalarda tiroid nodüllerinin hızlı, güvenilir ve standart segmentasyonunu sağlayarak tanısal doğruluğu artırma ve hekimlerin iş yükünü azaltma potansiyeline sahiptir.

ANAHTAR KELİMELER: Tiroid nodülleri, Ultrasonografi, Otomatik segmentasyon, Derin öğrenme, Hibrit mimari, EfficientNet, Vision Transformer

ABSTRACT

Elifnur ERDEMİR AYTEKİN

AUTOMATIC NODULE SEGMENTATION IN THYROID

ULTRASONOGRAPHY WITH HYBRID DEEP LEARNING METHODS

Baskent University, Institute of Science

Department of Computer Engineering

2025

Thyroid nodules are masses that develop in the thyroid gland, distinct from normal tissue, and are generally benign but carry a certain degree of malignant potential. These highly prevalent lesions are highly detectable by ultrasonography. However, accurate delineation of nodules is critical for both assessing malignancy risk and planning follow-up and treatment. In this thesis, a hybrid deep learning model was developed for the automatic segmentation of nodules from thyroid ultrasound images. The proposed method utilizes dual-encoder architecture that combines EfficientNet-based local feature extraction capabilities with Vision Transformer-based global context awareness. Additionally, a custom boundary-aware loss function was designed to more precisely define nodule boundaries. The model's performance was extensively tested on the TN3K and DDTI datasets and compared with current methods in the literature. Based on the obtained results, the developed hybrid approach has the potential to increase diagnostic accuracy and reduce physician workload by providing fast, reliable, and standardized segmentation of thyroid nodules in clinical practice.

KEYWORDS: Thyroid nodules, Ultrasonography, Automatic segmentation, Deep learning, Hybrid architecture, EfficientNet, Vision Transformer

İÇİNDEKİLER

TEŞEKKÜR.....	i
ABSTRACT	iii
İÇİNDEKİLER.....	iv
TABLolar LİSTESİ	vii
ŞEKİLLER LİSTESİ	viii
SİMGELER VE KISALTMALAR LİSTESİ	ix
1. GİRİŞ.....	1
1.1. Motivasyon ve Problemin Tanımı.....	1
1.2. Tezin Amacı ve Bilimsel Katkıları.....	1
1.3. Tez Planı.....	2
2. TIBBİ ARKA PLAN: TİROİD VE NODÜLLER	3
2.1. Tiroid Bezi Anatomisi ve Fizyolojisi.....	3
2.2. Tiroid Nodülleri.....	4
2.3. Nodüllerin Benign/Malign Olarak Sınıflandırılması	5
2.4. Epidemiyoloji ve Risk Faktörleri.....	6
2.5. Klinik Semptomlar	6
2.6. Tanı Yöntemleri.....	7
2.7. Tedavi Prosedürleri.....	8
2.8. Ultrasonografinin Temelleri.....	9
3. ALAN BİLGİSİ VE LİTERATÜR TARAMASI	11
3.1. Tıbbi Görüntü Segmentasyonuna Genel Bakış	11
3.2. Hibrit Kodlayıcı-Kod Çözücü Mimarileri	14
3.3. Tiroid Nodül Segmentasyonu Çalışmaları.....	16
4. VERİ KÜMELERİ	22
4.1. TN3K Veri Kümesi.....	22
4.2 DDTI Veri Kümesi	23
4.3. Veri Kümelerinde Ön İşleme	24
5. YÖNTEMLER.....	26
5.1. Kodlayıcı Yaklaşımları	26
5.1.1. Evrişim Tabanlı Yaklaşımlar	26
5.1.2. Transformatör Tabanlı Yaklaşımlar	28

5.2. Önerilen Model: EffiViT3+	29
5.2.1. Kodlayıcı Bileşenleri	29
5.2.2. Çok Özellikli Öznitelik Füzyonu	31
5.2.3. Kod Çözücü Aşaması.....	31
5.2.4. Sonraki İşleme.....	33
5.3. Kayıp Fonksiyonları.....	34
5.3.1. Dice Kaybı	34
5.3.2. Lovász Hinge Kaybı.....	34
5.3.3. Fokal Tversky Kaybı	35
5.3.4. Sınır Kaybı	35
5.3.5. Normalize Edilmiş HD95 Kaybı	35
5.3.6. Bileşik Kayıp	36
6. DENEYSEL SONUÇLAR	37
6.1. Deneysel Kurulum.....	37
6.1.1. Veri Kümelerini Bölme Stratejisi.....	37
6.1.2. Hiperparametreler	38
6.2. Test Metrikleri.....	38
6.2.1. Kesişim Birleşim Oranı	39
6.2.2. Dice Skoru	40
6.2.3. Kesinlik	40
6.2.4. Duyarlılık.....	40
6.2.5. Doğruluk.....	40
6.3. Kodlayıcı Sonuçları.....	41
6.3.1. Evrişim Tabanlı Kodlayıcılara Ait Sonuçlar	41
6.3.2. Transförmator Tabanlı Kodlayıcılara Ait Sonuçlar	42
6.4. Güncel Modellerle Karşılaştırma ve Analiz	44
6.4.1. TN3K Veri Kümesinde Analizler	44
6.4.2. DDTI Veri Kümesinde Analizler.....	47
6.5. Kayıp Fonksiyonu Sonuçları.....	49
6.6. Nitel Sonuçlar	51
7. TARTIŞMA ve SONUÇ	56
7.1. Bulguların Yorumu	56

7.2. Literatür ile Karşılaştırma	58
7.3. Gelecek Çalışmalar	60
7.4. Sonuç	61
KAYNAKLAR.....	63

TABLÖLAR LİSTESİ

	Sayfa
Tablo 3.1. Geçmiş tiroid segmentasyonu çalışmalarına ait bilgiler	19
Tablo 6.1. Model eğitiminde kullanılan hiperparametreler, sembolleri, değerleri ve açıklamaları	39
Tablo 6.2. TN3K veri seti üzerinde beş farklı evrişim tabanlı kodlayıcının performans metrikleri	42
Tablo 6.3. Transformatör tabanlı kodlayıcıların TN3K veri seti üzerindeki test performansları	43
Tablo 6.4. Güncel modeller ile önerilen modelin performans karşılaştırmaları	46
Tablo 6.5. DDTI veri seti üzerinde farklı segmentasyon modellerinin test metrikleri	49
Tablo 6.6. Farklı kayıp fonksiyonlarının metriklerinin karşılaştırılması	51

ŞEKİLLER LİSTESİ

	Sayfa
Şekil 2.1. Tiroid bezinin anatomik görünümü	4
Şekil 2.2. Tiroid nodülünün şematik gösterimi	5
Şekil 2.3. Ultrasonografinin boyun bölgesine uygulanışı	9
Şekil 2.4. Tiroid nodülüne ait örnek ultrasonografi görüntüsü	10
Şekil 4.1. TN3K verisetinden örnek tiroid ultrasonografi görüntüsü ve o görüntüye ait maske	22
Şekil 4.2. DDTI veri kümesinden tiroid ultrasonografi görüntüsü örneği ve bu örneğe ait ikilileştirilmiş maskesi	23
Şekil 5.1. EffiViT3+ modelinin bileşenleri	32
Şekil 5.2. Atröz evrişim, global havuzlama, concat ve 1x1 projeksiyonun beş aşamasını gösteren ASPP bloğunun mimari akış diyagramı	33
Şekil 6.1. TN3K test kümesi ile elde edilen nitel sonuçlar	54
Şekil 6.2. DDTI test kümesi ile elde edilen nitel sonuçlar	55

SİMGELER VE KISALTMALAR LİSTESİ

ACR-TI-RADS	American College of Radiology Thyroid Imaging Reporting and Data System - Amerikan Radyoloji Koleji Tiroid Görüntüleme Raporlama ve Veri Sistemi
Adam	Adaptif Moment Tahmini
ASPP	Atrous Spatial Pyramid Pooling - Atröz Uzaysal Piramit Havuzlama
ATA	American Thyroid Association - Amerikan Tiroid Derneği
BCE	Binary Cross-Entropy - İkili Çapraz Entropi
CBAM	Convolutional Block Attention Module - Konvolüsyonel Blok Dikkat Modülü
CE	Cross-Entropy - Çapraz Entropi
CE-Net	Context Encoder Network - Bağlam Kodlayıcı Ağı
CNN	Convolutional Neural Network - Konvolüsyonel Sinir Ağı
DDTI	Digital Database for Thyroid Images - Dijital Tiroid Görüntü Veritabanı
FN	False Negative - Yanlış Negatif
FP	False Positive - Yanlış Pozitif
FTV	Fokal Tversky Kaybı
GPU	Grafik İşleme Birimi
HD95	Hausdorff Mesafesi (95. persentil)
İİAB	İnce İğne Aspirasyon Biyopsisi
IoU	Intersection over Union - Kesişim Birleşim Oranı
Lovász	Lovász Hinge Kaybı
MAE	Mean Average Error - Ortalama Mutlak Hata
MLP	Multilayer Perceptron - Çok Katmanlı Algılayıcı
MSE	Mean Squared Error - Ortalama Kare Hata
PSPNet	Mean Squared Error - Piramit Sahne Ayrıştırma Ağı
ReLU	Rectified Linear Unit - Düzleştirilmiş Lineer Birim
SE-Net	Squeeze-and-Excitation Network - Sıkıştırma ve Uyarma Ağı
SE	Squeeze-and-Excitation - Sıkıştır ve Uyar
SGD	Stochastic Gradient Descent - Stokastik Gradyan İnişi
STN	Spatial Transformer Network - Mekansal Dönüştürücü Ağı
TN	True Negative - Doğru Negatif
TN3K	Tiroid Nodülü 3000 Veri Kümesi
TP	True Positive - Doğru Pozitif
USG	Ultrasonografi

1. GİRİŞ

1.1. Motivasyon ve Problemin Tanımı

Tiroid bezi, vücudun metabolik dengesini ayarlayan hayati bir endokrin organdır. Tiroid nodülleri, tiroid bezinde normal dokudan bağımsız olarak ortaya çıkan yuvarlak veya oval yapıda olabilir [1]. Tiroid nodülleri nüfusun çoğunu etkileyen ancak genellikle iyi huyludur. Ultrasonografi sayesinde rutin kontrollerde saptanan tiroid nodül sayısı giderek artmaktadır ve popülasyonun %60'ında en az bir nodül ultrasonografi ile tespit edilmiştir [2]. Epidemiyolojik veriler, papiller tiroid karsinom insidansının son otuz yılda yaklaşık 2,5 kat arttığını, buna karşılık mortalite eğrisinin büyük ölçüde sabit kaldığını göstermektedir. Tespit edilen tiroid nodüllerinin sadece %5'inin malign (kötü huylu) potansiyeli olduğu ortaya konmuştur. Buna rağmen tek başına nodül bulgusunun hasta üzerinde yarattığı psikolojik baskı ve hekimleri ek tetkiklere yönelten klinik endişe, tanı sürecini hem karmaşıktırmakta hem de maliyetli hâle getirmektedir. Literatürde “aşırı tanı” olarak nitelenen bu olgu, klinik olarak önemsiz nodüllerin dahi “kansere” etiketiyle takip edilmesine ve çoğu zaman gereksiz ince iğne aspirasyon biyopsisi yapılmasına yol açmaktadır [3].

Tiroid nodüllerinin klinik yönetiminde standartlaştırmayı sağlamak amacıyla ACR-TI-RADS (Amerikan Radyoloji Koleji Tiroid Görüntüleme Raporlama ve Veri Sistemi) ve ATA (Amerikan Tiroid Derneği) gibi risk sınıflandırma sistemleri geliştirilmiştir ve nodülün ultrasonografi görüntüsü üzerinde belirlenen morfolojik özelliklere göre malignite riskini derecelendirmektedir. Ancak bu derecelendirme sistemlerinin etkinliği, nodüllerin görüntüler üzerinde doğru bir biçimde segmentasyonun yapılmasına bağlıdır. Klinik uygulamalarda segmentasyon işlemi çoğunlukla manuel olarak gerçekleştirilmekte olup, bu durum gözlemciler arasında önemli ölçüde tutarsızlıklara yol açmaktadır. Ayrıca manuel segmentasyon zaman alıcıdır ve yüksek klinik iş yükü altında ölçeklenebilirliği sınırlıdır. Dolayısıyla hızlı, objektif ve güvenilir otomatik segmentasyon yöntemlerine duyulan ihtiyaç giderek artmaktadır.

1.2. Tezin Amacı ve Bilimsel Katkıları

Bu tezin temel amacı ultrasonografi görüntülerinde tiroid nodüllerinin otomatik segmentasyonunu sağlamaktır. Bu amaçla, önerilen otomatik segmentasyon yöntemi ile

nodüllerin hızlı ve güvenilir bir şekilde segmente edilmesi sağlanarak doktorların iş yükünün hafifletilmesi amaçlanmıştır.

Bu çalışmada, literatürdeki yöntemlere göre daha iyi performans gösteren yeni bir hibrit kodlayıcı- kod çözücü (encoder- decoder) mimarisi geliştirilmiştir. EfficientNet modelinin yerel detayları bulma yeteneği ile Vision Transformer modelinin küresel bağlamsal dikkat mekanizmasının birleşimi olan bu hibrit kodlayıcı yapısı, yenilikçi hibrit bir yaklaşım sunmaktadır. Bu kodlayıcı ile birlikte kullanılan çok ölçekli UNet3+ kod çözücüsü, özellikle küçük ve düzensiz şekilli nodüllerin segmentasyon doğruluğunu artırmayı hedeflemektedir.

Tez kapsamında geliştirilen yöntem, mevcut en gelişmiş (“state-of-the-art”) modellerle karşılaştırılarak kapsamlı deneysel analizlerle doğrulanmıştır. Çalışmanın bilimsel katkıları şu şekildedir:

- Yeni bir hibrit kodlayıcı- kod çözücü mimarisinin geliştirilmesi,
- Nodül sınırlarının doğruluğunu artırmak üzere yeni bir sınır-farkındalıklı kayıp fonksiyonunun önerilmesidir.

1.3. Tez Planı

Tez toplam yedi bölümden oluşmaktadır. Giriş bölümünde çalışmanın motivasyonu, problemin tanımı, amacı ve bilimsel katkıları ile tez planı özetlenmiştir. İkinci bölümde tiroid bezi, nodüller, tanı yöntemleri ve ultrasonografi gibi klinik temeller aktarılmıştır. Üçüncü bölümde literatür taraması kapsamında tıbbi görüntü segmentasyonu yöntemleri, hibrit kodlayıcı-kod çözücü mimarileri ve tiroid nodülü segmentasyonu çalışmaları incelenmiştir. Dördüncü bölüm, kullanılan veri kümeleri ile görüntü ön işleme adımlarını içermektedir. Beşinci bölümde önerilen hibrit derin öğrenme modelinin genel mimarisi ve temel bileşenleri tanıtılmıştır. Altıncı bölümde kapsamlı deneysel sonuçlar ve istatistiksel değerlendirmeler sunulmuştur. Yedinci bölümde elde edilen bulguların literatürle karşılaştırılarak yorumlanması, çalışmanın kısıtlamaları ve ileriye dönük öneriler tartışılmıştır ve çalışmanın sonuçlarını özetleyerek genel bir değerlendirme sunulmuştur.

2. TIBBİ ARKA PLAN: TİROİD VE NODÜLLER

2.1. Tiroid Bezi Anatomisi ve Fizyolojisi

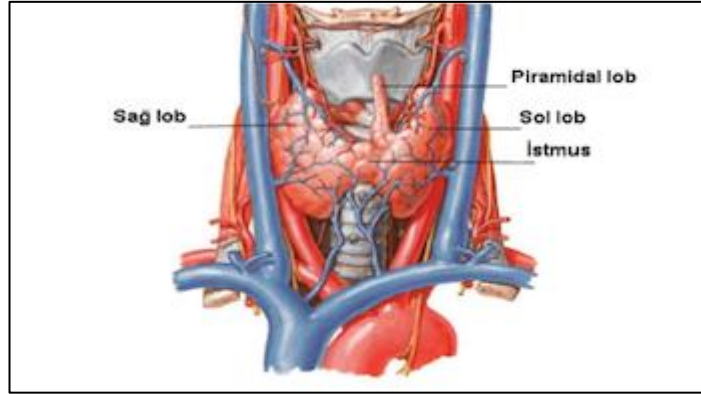
Tiroid bezi, boynun bölgesinin önünde larinks ve trakeaya komşu olarak bulunan endokrin sisteme ait bezdir [4]. Yan yüzlerinde ise vagus siniri ve juguler damarlar gibi önemli yapılar vardır. Bez iç ve dış olmak üzere iki kapsülle çevrilidir. İç kapsül bezle birleşik, dış kapsül ise daha gevşek yapıdadır. Sağ ve sol iki lobdan oluşan bu bez “isthmus” olarak adlandırılan bir köprü ile birbirine bağlıdır [5]. Tiroidin kanlanması superior ve inferior tiroid arterlerinden sağlanır [5, 6]. Yetişkin bireylerde bu bezin ortalama ağırlığı 15-20 gram, iki lobun uzunluğu 4 santimetre ve genişliği 2 santimetredir [6,7]. Embriyolojik gelişimi açısından tiroid bezi, insan embriyosunda gelişen ilk endokrin organdır. Embriyoda, primitif farinksin orta hattından kaynaklanan bir divertikül şeklinde, "tiroid divertikülü" olarak ortaya çıkar ve daha sonra farinks tabanından aşağıya doğru boynun ön kısmına doğru ilerler [4, 6].

Tiroid bezi histolojik olarak "foliküller" adı verilen küçük yapılar halinde bulunmaktadır. Her folikül, tek katlı küboidal epitelyumla çevrili ve içerisinde kolloid adı verilen maddeyi içeren boşluklardan oluşmaktadır [7]. Foliküler hücreler tiroid hormonlarının üretiminden sorumludur ve hormonları ekstraselüler olarak depolamaktadırlar. Foliküllerin arasında dağınık şekilde kalsitonin hormonunu salgılayan parafoliküler veya C hücreleri bulunmaktadır.

Tiroid bezinin fizyolojik olarak temel görevi, metabolizmanın düzenlenmesinde kritik rol oynayan tiroid hormonları olan tiroksin (T4) ve triiyodotironin (T3) üretimidir [4]. T3 ve T4, bezdeki folikül hücrelerinde sentezlenmektedir. Bu hormonların sentezi için iyot gereklidir. Tiroglobulin içerisinde tirozin aminoasitlerinin iyotlanmasıyla hormon sentezi gerçekleşir. T3 ve T4 hormonları metabolizmanın hızını ayarlar ve insan büyümesi üzerinde önemli etkileri vardır [8]. Buna ek olarak bezin parafoliküler hücreleri tarafından salgılanan kalsitonin, kalsiyum homeostazında görev alarak kemik metabolizmasını düzenlemektedir [6]. Buna ek olarak Tiroid hormonları, temel metabolik hız, vücut sıcaklığının düzenlenmesi, fetal dönemde beyin gelişimi ve büyüme gibi önemli fonksiyonlarda rol oynar [5]. Hipotalamus-hipofiz-tiroid aksı üzerinden gerçekleşen negatif geri bildirim mekanizması, tiroid hormonlarının sentezini ve salgılanmasını kontrol eder. Bu düzenleyici sistemde tiroid

stimüle edici hormon (TSH), hipofiz bezi tarafından salgılanarak, tiroid bezinin fonksiyonunu doğrudan düzenler [4, 6].

Sonuç olarak, tiroid bezi anatomik, embriyolojik ve histolojik açıdan karmaşık ve fonksiyonel olarak metabolik süreçler üzerinde önemli etkilere sahip bir endokrin bezdir. Şekil 1’de tiroid bezinin anatomik yapısı gösterilmiştir.



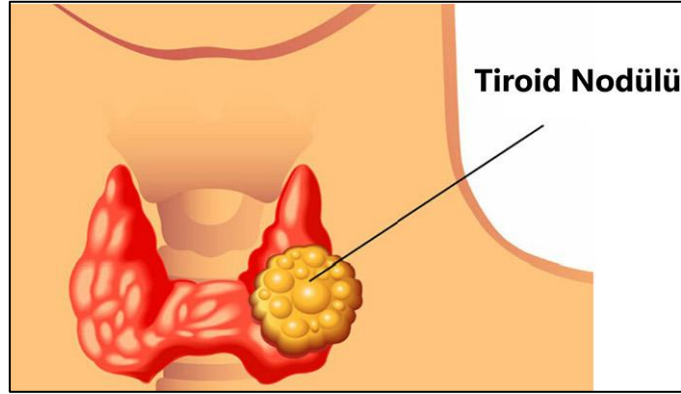
Şekil 2.1. Tiroid bezinin anatomik görünümü [9]

2.2. Tiroid Nodülleri

Tiroid nodülleri, tiroid bezinde normal dokudan farklı olarak gelişen, radyolojik olarak tiroid parankiminden ayırt edilebilen içi sıvı dolu veya katı, oval yapıda olabilen Şekil 2.2’de şematik gösterimi yapılmış olan lezyonlardır [10]. Bu nodüller palpasyon ile hissedilebilecek boyutta olabileceği gibi, ultrasonografi gibi ileri görüntüleme yöntemleriyle tesadüfi olarak da saptanabilmektedir [11]. Tiroid nodüllerinin yaygınlığı toplumlar arasında farklılık göstermekle birlikte, özellikle iyot eksikliğinin yaygın olduğu bölgelerde daha sık görülmektedir [5]. Palpasyonla yapılan klinik değerlendirmelerde saptanan nodül sıklığı %4-7 arasında değişirken, ultrasonografi ile saptanan nodül sıklığı %60’a kadar çıkmaktadır.

Nodüllerin büyük çoğunluğu benign (iyi huylu) özellik taşımaktadır. Malign nodüllerin oranı ise genel nodüller arasında %1-15 arasında değişmektedir [10]. Klinikte tiroid nodüllerine yaklaşımın ana noktasını, maligniteyi dışlamak veya tespit etmek etmektedir [11]. Malign nodüllerin erken dönemde tanısı, hastaların yaşam kalitesini ve prognozunu önemli ölçüde etkilediğinden, doğru tanısal yaklaşım kritik öneme sahiptir [11].

Nodüllerin oluşumu birçok sebebe bağlı olabilmektedir. Basit (diffüz) guatr, multinodüler guatr, benign adenomlar, kronik lenfositik tiroidit (hashimoto), subakut tiroidit ve Graves hastalığı gibi tiroid hastalıkları nodül formasyonu ile birlikte görülebilmektedir [10, 12]. Tiroid nodüllerinin oluşumunda iyot eksikliği temel rol oynarken, iyot fazlalığı da nodül oluşumunu tetikleyebilmektedir.



Şekil 2.2. Tiroid nodülünün şematik gösterimi [13]

2.3. Nodüllerin Benign/Malign Olarak Sınıflandırılması

Tiroid nodülleri, tiroid bezi içinde radyolojik olarak veya fizik muayeneyle çevre dokudan ayırt edilebilen kitlelerdir [14]. Tiroid nodüllerinin klinik açıdan önemi, çoğunluğu benign olmakla birlikte (%85-95), azımsanmayacak oranda (%5-15) malign potansiyel taşımaktadır [15]. Benign nodüller genellikle kolloid nodüller, hiperplastik nodüller, adenomatöz nodüller veya tiroidit gibi inflamatuvar durumlara bağlı oluşurken, malign nodüller papiller, foliküler, medüller ve anaplastik tiroid kanseri gibi maligniteleri içermektedir [16].

Benign nodüller tipik olarak yavaş büyür, çevre dokulara invazyon göstermez ve çoğunlukla tiroid fonksiyonlarını bozmaz [17]. Malign nodüller ise agresif büyüme, çevre doku invazyonu ve lenfatik metastaz gibi özelliklerle kendini gösterir. Ultrasonografide hipoekoik görünüm, mikrokalsifikasyon varlığı, düzensiz sınırlara sahip olması gibi özellikler malignite şüphesini artırır. Bu nedenle nodüllerin görüntüleme özellikleri, sınıflandırma ve takip protokollerinde önemli kriterler olarak kabul edilir [17].

Kısacası, tiroid nodüllerinin benign/malign ayrımı klinik yönetiminde kritik rol oynar, çünkü malign veya şüpheli nodüller cerrahi müdahaleyi gerektirirken benign nodüller genellikle düzenli takip edilmelidir.

2.4. Epidemiyoloji ve Risk Faktörleri

Tiroid nodülleri genel popülasyonda sık görülen endokrin sorunlardan biridir. Epidemiyolojik çalışmalar, bu nodüllerin sıklığının yaşla birlikte arttığını ve kadınlarda erkeklere göre dört kat daha yaygın olduğunu göstermiştir [7]. İyot eksikliği, tiroid nodüllerinin oluşumunda en önemli risk faktörüdür. İyot eksikliği olan bölgelerde nodül prevalansı %25'e kadar yükselirken, iyot yeterli bölgelerde prevalans genellikle %4-7 arasında değişmektedir [14].

Tiroid nodüllerinin oluşum mekanizması multifaktöryeldir. İyot eksikliği dışında, çevresel faktörler (sigara, radyasyon maruziyeti), genetik yatkınlık, yaşlanma ve kadın cinsiyet gibi faktörler de nodül gelişimini artıran önemli risk faktörleri arasında yer almaktadır. Özellikle çocukluk çağında veya erken yetişkin dönemde baş ve boyun bölgesine uygulanan radyasyon, sonraki yaşam döneminde nodül oluşum riskini ve tiroid malignitesi riskini belirgin biçimde arttırmaktadır [18].

Bazı çalışmalar ailesel yatkınlığın tiroid nodülü gelişiminde etkili olduğunu göstermiştir. Aynı aile bireylerinde guatr ve nodül oluşumu görülmesi, genetik faktörlerin nodül oluşumundaki rolüne işaret etmektedir. Bazı ailelerde tiroid stimülan hormon eseptörü, tiroglobulin veya tiroid peroksidaz gibi genlerdeki mutasyonların nodül oluşumunu tetikleyebildiği belirtilmiştir [14].

Son yıllarda görüntüleme yöntemlerinin gelişmesi ve yaygınlaşmasıyla birlikte, asemptomatik nodüllerin daha sık saptanır hale gelmesi nodüllerin genel popülasyondaki prevalansının arttırmıştır.

2.5. Klinik Semptomlar

Tiroid nodüllerinin çoğu başlangıçta asemptomatiktir ve sıklıkla rutin fizik muayenelerde veya ultrasonografi ile tesadüfen saptanmaktadır. Ancak nodüller büyüdükçe boyunda şişlik, kozmetik kaygılar veya bası belirtileri ile ortaya çıkabilmektedir [15]. Büyük nodüller trakea, özofagus ve sinir yapıları üzerine bası yaparak nefes darlığı, ses kısıklığı,

yutma güçlüğü (disfaji), boğazda dolgunluk hissi veya kronik öksürük gibi klinik semptomlara yol açabilmektedir [19].

Hormon aktif nodüller (otonom nodüller) tiroide ait hormonal dengeyi bozarak hipertiroidiye neden olabilir ve bu durum çarpıntı, kilo kaybı, sıcağa tahammülsüzlük ve sinirlilik gibi semptomlarla kendini gösterebilmektedir. Ancak nodüllerin çoğunluğu fonksiyonel değildir, hormonal düzeylerde değişiklik oluşturmazlar ve bu nedenle metabolik semptomlara neden olmazlar [18].

Ağrı veya hassasiyet nadiren görülmekle birlikte, tiroidit gibi inflamatuvar durumlara eşlik eden nodüllerde ortaya çıkabilir. Ayrıca, malignite durumunda lenfadenopati varlığı, nodülün hızlı büyümesi veya çevre dokulara yayılıyor olması önemli klinik bulgulardır ve malignite şüphesini güçlendirmektedir [19].

Semptomların varlığı ve şiddeti nodülün boyutuna, büyüme hızına ve yerleşimine göre değişkenlik gösterebilir. Bu nedenle nodüllerin klinik semptomları tanı ve takip süreçlerinde önemli yol göstericidir.

2.6. Tanı Yöntemleri

Tiroid nodüllerinin doğru şekilde değerlendirilmesi, klinik takip ve tedavi kararları açısından büyük önem taşımaktadır. Günümüzde nodüllerin tanı ve değerlendirilmesinde ilk basamak görüntüleme yöntemi olarak ultrasonografi kullanılmaktadır. Ultrasonografi non-invaziv, ekonomik ve radyasyon içermeyen yapısı sayesinde tiroid nodüllerinin değerlendirilmesinde ideal yöntem olarak kabul edilmektedir [17]. Ultrasonografi nodülün boyutunu, sınırlarını, yapısını (solid, kistik, karma), içeriğini, ekojenitesini ve kalsifikasyon varlığını ayrıntılı bir biçimde tanımlayarak nodüllerin malignite potansiyelini değerlendirmede kritik rol almaktadır [17]. Ultrasonografik bulgularda nodülün hipoekoik olması, düzensiz veya belirsiz sınırlarının bulunması, mikrokalsifikasyon içermesi ve internal vaskülarite artışı gibi özellikler malignite riski yüksek nodüllerin belirlenmesinde önemli göstergelerdir. Bunun yanı sıra, benign nodüller genellikle düzgün sınırlı, izoekoik veya hiperekoik yapıda, makrokalsifikasyonlu veya kistik komponentler içeren özellikler gösterirler [14, 17].

Ultrasonografi eşliğinde gerçekleştirilen ince iğne aspirasyon biyopsisi (İİAB), tiroid nodüllerinin malignite açısından değerlendirilmesinde kullanılmaktadır [15]. İİAB sonucunda elde edilen sitolojik örnekler, Bethesda sınıflaması doğrultusunda benign,

malign, malignite şüpheli veya yetersiz olarak değerlendirilir. Bu yöntemle malign nodüllerin tanısında yüksek duyarlılık (%80-95) ve özgüllük (%70-90) sağlanmakla birlikte, yanlış negatif veya yetersiz örnekleme gibi sınırlılıkları mevcuttur. Bu nedenle İİAB sonuçları hastanın klinik bulguları ve ultrasonografi sonuçları ile birlikte yorumlanmalıdır [15].

Sintigrafi ise nodüllerin çok çalışmasına göre sıcak, soğuk, ılık gibi fonksiyonel durumunu belirlemek amacıyla kullanılan başka bir görüntüleme yöntemidir. Ancak malign-benign ayrımı açısından spesifikliğı ve duyarlılığı düşüktür ve günümüzde rutin olarak ilk değerlendirme yöntemi olarak tercih edilmemektedir. Sintigrafi, esas olarak fonksiyonel nodüllerin belirlenmesinde faydalıdır [17].

2.7. Tedavi Prosedürleri

Tiroid nodüllerinin tedavi yaklaşımı, nodülün benign veya malign oluşuna, büyüklüğüne ve klinik semptomların varlığına göre belirlenmektedir [19]. Benign nodüller, çoğunlukla aktif tedavi gerektirmemektedir. Klinik olarak anlamlı semptomlar veya büyüme eğilimi göstermediğı sürece ultrasonografi eşliğinde periyodik takip edilmesi önerilmektedir [15]. Ancak nodülün büyük olması, kozmetik problemlere yol açması veya solunum yolu basısı gibi klinik problemlere sebep olması durumunda cerrahi müdahale gerekebilir [14].

Malign nodüllerde ise standart yaklaşım cerrahi müdahaledir ve tiroidektomi (tüm ya da parça) tercih edilir. Cerrahi müdahalenin kapsamı nodülün boyutuna, yayılımına ve lenfatik metastaz durumuna göre belirlenmektedir. Özellikle papiller ve foliküler tiroid kanserlerinde tüm tiroidektomi tercih edilerek, nüks riskini azaltmak ve postoperatif radyoaktif iyot tedavisinin etkinliğini artırmak amaçlanmaktadır [18]. Medüller tiroid kanserlerinde ise cerrahiye ek olarak, serum kalsitonin seviyeleriyle yakın takip gerçekleştirilmekte ve hastalığın biyokimyasal kontrolü sağlanmaya çalışılmaktadır [14].

Cerrahi sonrası dönemde malign vakaların tedavisinde sıklıkla radyoaktif iyot tedavisi kullanılmaktadır. Bu tedavi rezidü tiroid dokusunun ve mikroskobik düzeydeki metastatik odakların yok edilmesinde önemli rol oynamaktadır. Ayrıca bazı yüksek riskli durumlarda radyoterapi veya hedefe yönelik moleküler tedaviler gibi ek tedavi seçenekleri düşünülebilir [14, 18].

Son yıllarda benign nodüller için minimal invaziv yöntemler olan perkütan etanol enjeksiyonu, lazer veya radyofrekans ablasyonu cerrahi müdahaleye alternatif olarak

uygulanmaya başlanmıştır. Bu yöntemler, cerrahinin risklerini azaltmakla beraber semptomları hafifletmek, nodül boyutunu küçültmek ve hastanın yaşam kalitesini artırmak açısından avantaj sağlamaktadır [19].

2.8. Ultrasonografinin Temelleri

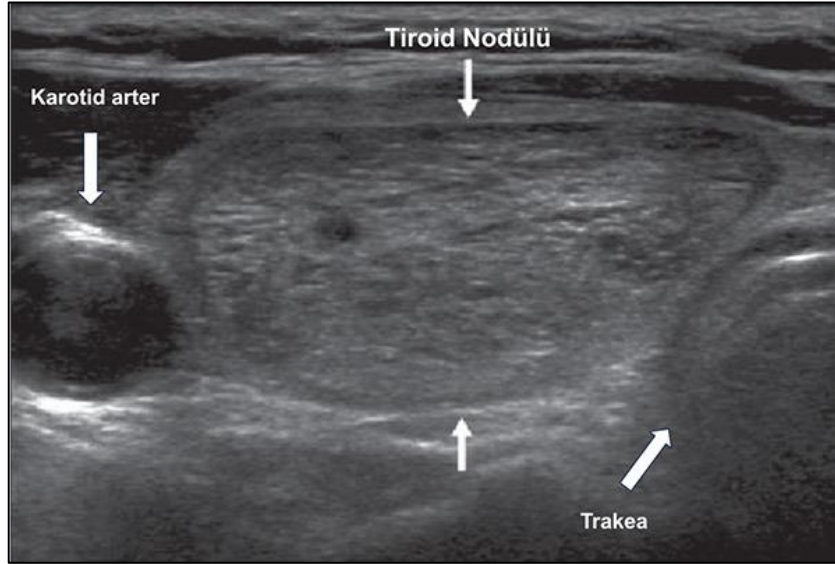
Ultrason, genellikle insan kulağının duyma sınırı olan 20 Hz–20 kHz üzerindeki 1–20 MHz frekanslardaki ses dalgalarını kullanarak vücut dokularının ağrısız görüntülenmesini sağlayan bir tıbbi görüntüleme yöntemidir. Bu frekans aralığı, tiroid bezinin düşük penetrasyon derinliğiyle yüksek mekânsal çözünürlükte değerlendirilmesine imkân tanımaktadır [17].

Ultrasonografi çalışma prensibi, piezoelektrik etki ve “nabız-yankı” (pulse-echo) ilkesine dayanmaktadır. Dönüştürücü (transdüser) içerisindeki piezoelektrik kristallere kısa, güçlü elektrik uyarıları uygulanarak ses dalgaları üretmektedir. Bu dalgalar dokular arasındaki akustik empedans farklılıklarında kısmen yansımakta ve yankı oluşturmaktadır. Dönüştürücü, geri dönen yankı dalgalarını elektrik sinyallerine dönüştürmektedir. Yankının geliş süresine göre mesafe, genliğine göre parlaklığı belirlenmektedir [21]. Daha sonra alınan veriler işlenerek gri tonlamalı kesit görüntüler oluşturulmaktadır. Şekil 2. 3’te boyun ultrasonu çekimi gösterilmiştir.



Şekil 2.3. Ultrasonografinin boyun bölgesine uygulandığı [20]

Tiroid bezinin deęerlendirilmesinde ultrasonografi, nodüllerin boyutu, řekli, sınırları, iç yapısı, ekojenitesi ve kanlanma özellikleri hakkında ayrıntılı bilgi sunmaktadır [14]. Yüksek rezolüsyonlu prob kullanımı sayesinde, nodüllerin çok küçük boyutlarda bile güvenle tanımlanmasını sağlamaktadır [14]. řekil 2.4'te görüntünün merkezinde yer alan ultrasonografi ile çekilmiş tiroid nodülü görüntüsü bulunmaktadır.



řekil 2.4. Tiroid nodülüne ait örnek ultrasonografi görüntüsü [9]

3. ALAN BİLGİSİ VE LİTERATÜR TARAMASI

3.1. Tıbbi Görüntü Segmentasyonuna Genel Bakış

Görüntü segmentasyonu, dijital bir resim üzerindeki pikselleri ortak özelliklere göre gruplandırarak istenilen nesneyi ayırma işlemidir. Bu süreç, renk benzerliği ya da kenar sürekliliği gibi düşük seviyeli ölçütlerle başlayıp nesne konum ve sınırlarının çıkarımına kadar ilerleyen çok katmanlı bir analiz zincirinden oluşmaktadır. Bilgisayarlı görüş uygulamalarında segmentasyon; nesne tanıma, sahne anlaşılması ve robotik navigasyondan artırılmış gerçekliğe kadar geniş bir yelpazede kilit rol oynamaktadır. Sayısal veride mekânsal ilişkileri koruyarak çıkarım yapması, segmentasyonu salt sınıflandırmadan ayıran en temel özelliğidir [22].

Tıbbi görüntü segmentasyonu ise bu genel tanımları klinik bağlamda uygulamaktır. Bilgisayarlı tomografi (BT), manyetik rezonans görüntüleme (MRG), pozitron emisyon tomografisi (PET), ultrasonografi veya dijital röntgen gibi modalitelerden elde edilen ham verilerden organ, doku veya patolojik lezyon sınırlarını otomatik olarak belirlemeyi amaçlamaktadır. Hastaların anatomik varyasyonları, görüntü gürültüsü ve cihaz ayarlarına bağlı kontrast değişkenliği segmentasyonu zorlu kılmaktadır. Buna karşın doğru sınırlar çıkarıldığında tümör hacim ölçümü, radyoterapi hedef sahası planlama, cerrahi navigasyon ve hastalık progresyon takibi gibi hayat kurtarıcı uygulamalar yapılabilmektedir [22].

Tıbbi görüntüler, düşük sinyal-gürültü oranı, metal artefaktları, organ sınırlarında kontrast eksikliği ve hasta bazlı anatomik çeşitlilik nedeniyle klasik segmentasyon algoritmaları için zorlu bir test alanı oluşturmaktadır. El yapımı özniteliklere dayanan kenar operatörleri veya aktif kontur modelleri; kapsamlı parametre ayarı, başlangıç konturu bağımlılığı ve gürültüye hassasiyet nedeniyle klinik pratikte sınırlı kalmıştır. Ayrıca bu yaklaşımlar, üç boyutlu hacimlerde ardışık dilimler arası tutarlı sınır üretmekte zorlanmış, çapraz cihaz genellemesini garanti edememektedir [23]. Bu problemin üstesinden gelmek için 2015'te "Tam Bağlantılı Ağ" (Fully Convolutional Network, FCN) tanıtılmıştır. FCN, geleneksel evrişim ağlarındaki tam bağlantılı katmanları tamamen evrişimli filtrelerle değiştirerek görüntü boyutunu koruyan piksel-piksel olasılık haritaları oluşturmaktadır. Böylece yama (patch) tabanlı yinelenen hesaplamalar ortadan kalkarak ve baştan sona (end-to-end) öğrenme mümkün olmuştur. FCN aynı zamanda atlama (skip) benzeri hiper-bağlantılar aracılığıyla düşük çözünürlüklü semantik öznitelikleri yeniden örnekleyerek ince

ayrıntıları kısmen koruyabilmiştir. Bu sayede ilk otomatik beyin tümörü segmentasyonu çalışmalarında belirgin iyileşme sağlamıştır [23].

FCN'nin ardından, tıbbi görüntü segmentasyonunda çığır açan U-Net tanıtılmıştır. Simetrik kodlayıcı- kod çözücü yapısı, her havuzlama adımında kaybedilen ayrıntıları karşılık gelen çıkarım adımında kopyala-birleştir (skip connection) yoluyla geri kazandırmıştır. U-Net'in bu çift yönlü bilgi akışı, pankreas gibi kontrastı düşük organlarda bile üstün sınır hassasiyeti sunmuştur ve küçük veri setleriyle dahi aşırı öğrenme olmadan eğitilebilmiştir. U-Net bugün, ameliyathanelerde gerçek zamanlı sinir monitörlerinden otomatik mikroskop hücre sayımına kadar yüzlerce makalede referans model konumundadır [23].

U-Net++, U-Net'in atlama bağlantılarını iç içe yoğun (nested dense) bloklarla zenginleştirerek kodlayıcı ve kod çözücü arasındaki semantik uyumsuzluğu azaltmıştır. Ara katmanlara uyguladığı derin gözetim (deep supervision) ile çoklu çözünürlükte geri bildirim sağlayarak öğrenmeyi hızlandırmıştır. Özellikle karaciğer BT'de parankim ve tümörü aynı ağ içinde güvenilir şekilde ayırmasıyla dikkat çekmiştir. Buna koşut geliştirilen Attention U-Net, atlama yolu üzerine kanal-uzaysal dikkat kapıları ekleyerek ağın yalnızca ilgili organ veya lezyon bölgesine odaklanmasını sağlamıştır [23]. Bu iyileştirmelerle yerel ayrıntılar keskinleşse de sahnenin farklı ölçeklerdeki bağlamsal bilgisini derinlemesine yakalamak için algısal alanı genişletecek yeni mimarilere ihtiyaç duyulmuştur. UNet3+ ise tüm çözünürlük seviyelerinden gelen öznitelikleri eş-zamanlı olarak birleştiren tam ölçekli atlama bağlantısı (full-scale skip) stratejisini kazandırmıştır. Bu yaklaşım, yüksek ayrıntının (kenar bilgisi) düşük ayrıntıdaki küresel kontekst ile tek adımda kaynaşmasını sağlayarak segmentasyon doğruluğunu arttırmaktadır. Ayrıca derin gözetim (deep supervision) mekanizması, kod çözücünün her seviyesinde ara çıktılar üretip kayıp fonksiyonuna dâhil ederek ağın çok-ölçekli duyarlılığını kuvvetlendirmektedir [24].

Çok ölçekli bağlam yakalama ihtiyacını karşılayan Piramit Sahne Ayırıştırma Ağı (Pyramid Scene Parsing Network, PSPNet), ResNet ya da DenseNet omurgasından çıkan özellik haritalarını dört farklı havuzlama ölçeğinde toplayarak global ve yerel bilgiyi aynı katmanda birleştirmiştir. DeepLab ailesi ise atröz evrişimlerle algısal alanı büyütürken çözünürlüğü korumayı başarmıştır. DeepLabv3'teki "Atröz Uzaysal Piramit Havuzu" (Atrous Spatial Pyramid Pooling, ASPP) paralel genişletme oranlarıyla çok ölçekli bağlam sunmuştur. DeepLabv3+ hafif bir kod çözücü eklenerek sınır ayrıntısını iyileştirilmiştir [25].

Özellik Piramit Ağı (FPN, Feature Pyramid Network), çok-ölçekli özellik piramidi oluşturarak farklı çözünürlüklerdeki evrişim katmanlarından gelen öznitelikleri yukarıdan aşağıya birleştirmektedir. En üst seviyedeki semantik açıdan zengin fakat düşük çözünürlüklü özellikler, yukarı örneklendirilerek alt seviyelerdeki ayrıntılı haritalarla toplanmaktadır. Bu piramidal yapı hem küçük hem de büyük yapısal hedefleri aynı anda tespit etmeye olanak tanımaktadır [26].

Bellek verimliliği odaklı SegNet, maksimum havuzlamayı havuzdan çıkarma (unpooling) aşamasında kullanarak kod çözücü ağırlıklarını azaltmıştır [27]. CE-Net ise “Inception” türevi bağlamsal kodlayıcı (context encoder) bloklarıyla sınırları bulmada iyi performans elde etmektedir [28]. Mobil donanım için geliştirilmiş MobileNet-V2 V-Net varyantı, derin ayırıcı (depth-wise) evrişimlerle parametre sayısını beşte bire indirmeyi başarmıştır [29].

Örnek (instance) segmentasyonu alanında Mask R-CNN, Faster R-CNN modellerinin bölge öneri mimarisine ilgi alanı hizalama (region of interest Align) ve paralel maske kolu ekleyerek her nesne için piksel düzeyli maske üretmiştir. Meme MRG çalışmalarında 400 ms civarındaki işlem süresiyle klinik iş akışına uyum göstermiştir [30]. “Yol Toplama Ağı” (PANet, Path Aggregation Network), özellik piramidindeki bilgi akışını alt-yukarı yollarla zenginleştirerek küçük lenf nodülü veya mikro kalsifikasyon gibi ayrıntıların tespitini güçlendirdi ve çok organlı BT verilerinde yüksek ortalama hassasiyet kaydetmiştir [23].

Kanal yeniden ölçekleme kavramını getiren SE-Net (Squeeze-and-Excitation) blokları, her kanalı global ortalama havuzlama sonrası elde edilen bir vektörle modüle ederek önemli öznitelikleri vurgulamıştır [31]. Kalın bağırsak polip segmentasyonunda sınıf dengesizliği kaynaklı hataları %20 azaltmasıyla öne çıkmıştır. “Mekansal Transformatör Ağı” (STN, Spatial Transformer Network) ise girdi koordinatlarını optimize edip vertebra numaralandırma ve fetal beyin konumlandırma gibi geometri duyarlı segmentasyonlarda standardizasyon sağlamıştır [23].

Öz dikkat (Self-attention) mekanizmasının ortaya çıkışıyla TransUNet geliştirilmiştir. Evrişimli Sinir Ağı (CNN) kodlayıcı çıktısı, yama (patch) tabanlı Görüntü Transformatörü (ViT, Vision Transformer) katmanlarına aktarılıp yerel-küresel bağlam bir arada işlenmiştir [31]. Buna ek olarak kaydırmalı pencere (shifted window) kullanan Swin-UNet, belleği verimli bir şekilde kullanmayı başarmıştır. Tamamen transformatör tabanlı

nnFormer ise az sayıda CNN bileşeniyle bile yüksek doğruluk elde edilebileceğini kanıtlamıştır [32].

TransFuse, evrişimli ve dönüştürücü yapıları paralel akışta birleştirerek yerel ayrıntı ile küresel bağlamı eşzamanlı işleyen bir hibrit segmentasyon ağıdır. Barındırdığı evrişim katmanlarıyla kenar odaklı ayrıntıları öğrenirken, transformatör kısmı pikseller arası uzun menzilli ilişkileri öğrenmektedir. Ardından öz nitelikler, çapraz dikkat (cross-attention) tabakalarıyla birleştirilmektedirler [33].

SegFormer, Mix-MLP bloklarını ve ölçekler arası hafif başlık (lightweight head) yapısını kullanarak saf transformatör tabanlı segmentasyonu pratik hâle getiren bir mimaridir. Kod çözücü kısmı, farklı çözünürlüklerdeki çok-kafalı dikkat çıktılarının basit birleştirmeyele gerçek zamana yakın hızlarda çalışabilmektedir [34].

3.2. Hibrit Kodlayıcı-Kod Çözücü Mimarileri

Gao ve arkadaşlarının önerdiği UTNet, klasik U-Net iskeletindeki her çözünürlük seviyesinin son evrişiminde öz-dikkat katmanlarıyla değiştirerek CNN ve Transformer bileşenlerini mikro ölçekte kaynaştırmışlardır. Kodlayıcıdaki evrişim blokları doku ve kenar ayrıntılarını daha iyi yakalarken, görelî konum kodlamalı Transformer blokları uzun menzilli bağlamı her ölçekte toplamıştır. Benzer biçimde, kod çözücüdeki öz-dikkat modülleri skip bağlantılardan gelen ince sınırlara küresel tutarlılık kazanmıştır. Kardiyak MR verilerinde Dice skoru ve Kesişim Birleşim Oranı (Intersection over Union, IoU) TransUNet’e göre yükseltirken parametre artışını sınırlı tutması modelin başlıca güçlü yönü olmuştur. Buna karşın, Transformer katmanlarının ek hiper-parametre gereksinimi küçük veri setlerinde hassas ayar zorunluluğu ve ayrıca eğitim süresi saf U-Net’e kıyasla uzamaktadır [35].

He ve arkadaşlarının geliştirdiği H2Former, dört kademeli bir kodlayıcı içinde ResNet-34 evrişim bloğu, Çok Ölçekli Kanal Dikkati (MSCA) ve Swin-benzeri Transformer üçlüsünü tek “hibrit blok”ta bir araya getirmiştir. Evrişim yerel konum bilgisini sağlarken MSCA kanallar arası ilişkiyi 4–32 piksel arası pencerelerde yeniden bileştirmiştir. Pencere tabanlı MHSA ise uzun menzilli bağıntıları yakalamıştır. Model, Kvasir-SEG verisinde TransUNet’ten %2,3 daha yüksek IoU elde ederken parametreleri üçte bire düşürmesiyle dikkat çekicidir. Karmaşık blok tasarımının tekrar kullanılabilirliğini azaltması ve MSCA’nın donanım optimizasyonu gerektirmesi modelin zayıf yönleridir [36].

Perera ve arkadaşlarının MobileUNETR çalışması, mobil cihazlarda gerçek zamanlı deri lezyonu segmentasyonu hedefiyle MobileNetV3 gövdeli CNN ve MobileViT tabanlı Transformer bloklarını aynı hiyerarşik kodlayıcıda birleştirmiştir. CNN gövdesi düşük seviyede derinlik-ayrık evrişim katmanları boyutu küçültürken, MobileViT her aşamada özellik haritasını yamalara bölerek küresel bağıntıları öğrenmiştir. Kod çözücünde transpoze evrişim yerine “piksel-shuffle” kullanılması hem bellek tüketimini azaltmıştır. Yazarlar ISIC ve PH2 veri kümelerinde U-Net++’ı 1,3-2,7 puan Dice skorunda artışla geride bıraktıklarını rapor etmişlerdir. Ancak dar kanal sayısı karmaşık, saçlı veya damarlı lezyon sınırlarında ayrıntı kaybına yol açabilmesi bu modelin temel zayıflığıdır [37].

Lan ve arkadaşlarının BRAU-Net++ mimarisi, BiFormer bloklarında kullanılan İki Seviyeli Yönlendirme Dikkat (Bi-Level Routing Attention) ile hem kodlayıcı hem de kod çözücünde seyrek fakat içerik uyumlu öz-dikkat uygulamasıdır. Geleneksel atlama bağlantıları yerine tasarlanan SCCSA (Kanal-Uzaysal Dikkat) modülü, çok ölçekli özellikleri yeniden ölçekleyerek bilgi kaybını telafi etmiştir. Synapse, ISIC-2018 ve CVC-ClinicDB veri kümelerinde Dice skoru ortalama %1,5 arttırmıştır. Bununla birlikte, GPU’da düzensiz bellek erişimine yol açan seyrek örüntüler bu modelin gerçek zamanlı uygulamalarını zorlaştırabilir [38].

Tang ve arkadaşlarının sunduğu HTC-Net, EfficientNet tabanlı CNN ile Swin Transformer bloklarını üç kollu “Trident Multi-Layer Fusion” (TMF) modülü üzerinden birleştirmiştir. Her kol farklı silme (dilate) oranlı 3×3 evrişim katmanları ve Squeeze-and-Excitation (SE) blokları içerir, çıkışlar United Attention (UA) bloğunda kanal-uzaysal olarak yeniden ağırlıklandırmaktadır. ISIC-2017 ve Kvasir-SEG’de Jaccard’ı %86 bandına taşıyan model, 24 M parametreyle hesap-performans dengesini korumuştur. Üç kollu yapı bellekte fazladan yük getirdiğinden eğitim süresi uzamıştır buna rağmen ablasyon testleri her bileşenin %2-3’lük katkısını ortaya koymuştur [39].

Wang ve arkadaşlarının HDA-ResUNet modeli, U-Net kodlayıcısının dip kısmını Hibrit Silme-Dikkat Evrişim (HDAC) tabakasıyla değiştirmiştir. Atlama bağlantılarına ise Kanal Dikkat (CA) eklemiştir. Üç farklı silme oranı, geniş alandaki bağlamı parametrik artış olmadan toplarken CA gürültülü kanalları bastırmıştır. Karaciğer ve akciğer segmentasyonunda Dice skoru %0,5-1 oranında yükseltip parametreleri U-Net’ten düşük tutması güçlü yön olsa da ardışık yüksek silme oranları ufak yapıları parçalayabildiği için küçük lezyonlarda ayrıntı kaybı yaşatmıştır [40].

Xu ve Wang'ın MCV-UNet çalışması, kodlayıcıda iki atröz evrişim katmanını, darboğazda (bottleneck) ise ViT katmanını kullanarak sinir ultrasonunda küresel bağlamı etkin biçimde modele dahil etmiştir. Atröz katmanlar düşük seviye ayrıntıyı korurken ViT, uzun menzilli bağıntıyı saf CNN'lerden daha iyi öğrenmiştir. 17 temel yöntemle yapılan karşılaştırmada Dice skorunda %0,47'lik iyileşme sağlarken, parametre ve FLOP orta düzeyde kalmıştır. ViT'in spekl gürültüsü yüksek görüntülerde aşırı genelleme yapabilme yeteneği ve sınırlı etiketli veriyle performansı düşmüştür [41].

Zhang ve arkadaşlarının TranSegNet modeli, retina OCT katman segmentasyonu için U-şekilli CNN ile hafif ViT'i aynı kodlayıcı içinde ardışık değil de iç içe kullanmıştır. CNN çoklu ölçekli öznetelik çıkarırken, çok başlı Evrişimli Dikkatli ViT global ilişkileri bulmuştur. Sağlıklı ve DME hastası OCT veri kümelerinde, FCN-SegNet-TransUNet üçlüsünden anlamlı derecede daha yüksek Dice skoru elde edilmesi güçlü yönüdür. Yine de ViT girişiyle aynı boyutta tasarlandığından bellek gereksinimi yüksektir ve 3D hacimler dilim dilim işlenmek zorundadır [42].

3.3. Tiroid Nodül Segmentasyonu Çalışmaları

Tiroid ultrasonografisinde nodül segmentasyonu, klinik tanı süresini kısaltmak ve tanı doğruluğunu artırmak amacıyla son yıllarda derin öğrenme tabanlı mimariler ile yoğun biçimde incelenmiştir. Ma ve arkadaşları, AmSeg modelinde önce gerçek maske verilerinden nodül merkez koordinatlarını belirlemiştir. Ardından her ultrasonografi görüntüsünü bu merkeze göre 224×224 , 518×518 ve 896×896 boyutlarında kırparak üç paralel Swin-UNet bloğuna aktarmıştır. Model, odaklı Dice (Fokal Dice) kaybı ile çok ölçekli L1 karşıt kaybının ağırlıklı toplamını kullanmıştır. Her ölçekteki üç maskeyi piksel bazında toplayarak son çok ölçekli segmentasyonu sağlamıştır. Bu yöntemle TN3K veri kümesinde 614 görüntü için %84,21 Dice skoru ve %72,48 IoU, DDTI veri kümesinde 63 görüntü için %74,81 Dice skoru ve %60,89 IoU elde edilmiştir [43].

Bi ve arkadaşlarının geliştirdiği BPAT-UNet modeli, tiroid ultrasonu görüntülerindeki belirsiz sınırlar ile küçük nodülleri daha doğru ayırmak üzere üç ana bileşen içermektedir. Sınır Noktası Gözetim Modülü kenar piksellerini Canny filtresiyle vurgulamakta ve “şerit” ile “piramit” havuzlamalı öz-dikkat bloklarıyla zenginleştirmektedir. Uyarlanabilir Çok Ölçekli Özellik Birleştirme Modülü kodlayıcının orta katmanlarındaki derinlik farklarını deforme edilebilir evrişim ve deforme edilebilir

dikkat bloklarıyla işlemektedir. Birleştirilmiş Dönüştürücü Modülü ise yerel ve küresel bağlamı pekiştirmektedir. Model, veri kümelerini sekiz-bir-bir oranında eğitim, doğrulama ve sınamaya ayırmış, Yumuşak Dice (Soft Dice) kaybı kullanarak TN3K’de %84,23 Dice skoru ve %71,87 IoU’ya ulaşmıştır [44].

Tao ve arkadaşlarının LCA-Net modeli, beş katmanlı U-şekilli bir iskelete dayanmaktadır. Kodlayıcıda her seviyede 1×1 ve 1×7 çekirdekli evrişimler ince kenar ayrıntılarını yakalamaktadır. En üst seviye, uyarlanabilir çok ölçekli genişletilmiş evrişim katmanları barındırmaktadır. Çözücüde atlama bağlantılarından gelen öznitelikler ardışık ters evrişim ya da yama genişletme işlemlerinden sonra 3×3 evrişimler ve bağlam-dikkat bloklarıyla harmanlanarak hem yerel hem küresel bilgiyi koruyan tahmini maske oluşturmaktadır. TN3K’de %82,08 Dice skoru ve %71,18 IoU elde edilmiştir [45].

Ali ve arkadaşlarının CIL-Net modelinde kodlayıcı bölümü, yoğun bağlantılı yapıyı Üçlü Dikkat bloklarıyla birleştirerek nodül-özel öznitelikleri öne çıkarmaktadır. Öznitelik İyileştirme Bloğu, en derin katmandaki dört farklı genişletilmiş evrişimden gelen çok ölçekli öznitelikleri kanal ve uzamsal dikkatle zenginleştirip çözücüye ek bilgi olarak aktarmaktadır. Çözücüde 3×3 evrişimler ve Artık Bağlam Bloğu kullanılarak yapısal bütünlük ve sınır hassasiyeti korunmaktadır. Yüzde 50 çapraz entropi ile yüzde 50 Dice toplamından oluşan kayıpla TN3K’de %79,68 Dice skoru ve %64,90 IoU, DDTI’de %83,35 Dice skoru ve %73,35 IoU değerleri kaydedilmiştir [46].

Deng ve arkadaşlarının STU3Net modeli, U-şekilli mimariyi “Swin” dönüştürücü ile birleştirerek hem yerel hem küresel öznitelikleri yakalamaktadırlar. Kodlayıcıda klasik evrişimli bloklar ile dönüştürücü blokları paralel çalışmakta ve her seviyedeki çıktılar toplanarak çözücüye aktarılmaktadır. Ayrıca her dönüştürücü bloğundan önce Kanal Öncelikli Evrişimli Dikkat uygulanarak kanal öncelik bilgisi uzamsal dikkatle harmanlanmaktadır. Yüzde 30 ağırlıklı ikili çapraz entropi ve yüzde 70 ağırlıklı genelleştirilmiş Dice kaybıyla TN3K’de %83,68 Dice skoru ve %74,16 IoU elde etmişlerdir [47].

Zhang ve arkadaşlarının DDTransUNet modelinde çift kodlayıcı bulunmaktadır: biri dönüştürücüyle küresel özellikleri, diğeri evrişimli yapıyla yerel ayrıntıları yakalamaktadır. İki daldan gelen öznitelikler çapraz dikkat ile birleştirilmekte ardından Küresel Uzamsal Dikkat ve Küresel Kanal Dikkati modülleri üzerinden çözücüye aktarılmaktadır. TN3K’de %82,31 Dice skoru ve %73,82 IoU elde edilmiştir [48].

Liu ve arkadaşlarının BFG-MSF-Net modelinde ConvNeXt kodlayıcıya dört yeni modül eklenmiştir: Sınır Özniteliği Yönlendirme Modülü kenar bilgisini derin özniteliklerle harmanlamaktadır, Çok Ölçekli Algı Birleştirme Modülü farklı ölçek bilgilerini dikkat temelli biçimde bütünleştirmektedir, Derin Ayırık ASPP Modülü genişletilmiş evrişimlerle geniş bağlam sunmaktadır, İyileştirme Modülü ise CBAM destekli asimetrik evrişimlerle son maskede küçük ayrıntıları belirginleştirmektedir. İkili çapraz entropi ve Dice kaybı ile eğitilen model, TN3K’de %87,26 Dice skoru ve %76,34 IoU ile DDTI veri kümesinde yüzde 85,12 Dice skoru ve yüzde 74,02 IoU değerlerine ulaşmıştır [49].

Wu ve arkadaşlarının DW-Swin modeli, kenarları belirsiz ve gürültülü ultrasonografi görüntülerinde nodül sınırlarını ayırmak için derinlik ayırık öz-dikkat tabanlı kodlayıcı ve çözücü katmanları kullanmaktadır. Çok başlı öz-dikkat ve çok katmanlı algı bloklarına derinlik ayırık evrişim eklenmiştir. Kodlayıcı-kod çözücü atlamalarında Özellik Birleştirme Atlama Bağlantısı modülleri, çok ölçekli evrişim ve SE bloklarıyla kanal etkileşimi ile bağlamsal bilgiyi güçlendirmektedir. Bu mimari TN3K’de %82,26 Dice skoru ve %78,64 IoU, DDTI’de %73 Dice skoru ve %67,93 IoU değerleri sağlamıştır [50].

Yetginler ve Atacak, klasik V-Net kodlayıcısına aynı anda üç ölçekli öznitelik kümesini birleştiren Üçlü Kaynaştırma mekanizmasını eklemiştir. Ek olarak kodlayıcı bloklarının tamamına ve yalnızca son kod çözücü bloğuna SE mekanizması uygulayarak önemli kanalları vurgulamıştır. DDTI’de %84,51 Dice skoru ve %76,27 IoU, TN3K’de %83,88 Dice skoru ve %75,50 IoU ile önceki V-Net türevlerini geride bırakmıştır [51].

Li ve arkadaşlarının geliştirdiği GSE-Nets modeli, klasik U-Net yapısına Küresel Yapı Güçlendirici çözücü ekleyerek düşük seviyeli yapısal bilgileri yüksek seviyeli anlamsal özniteliklerle dikkat ve grafik evrişim blokları aracılığıyla harmanlamıştır. Derinlik ayırık evrişimlerle yüksek çözünürlüklü maskeler üretmişlerdir. DDTI’de %83,52 Dice skoru ve %71,71 IoU ile TN3K’de %81,97 Dice skoru ve %72,24 IoU elde etmişlerdir [52].

Çalışmalara ait yıl, yazarlar, model adları, kullanılan veri kümeleri, tercih edilen kayıp fonksiyonları ile elde edilen Dice skoru ve IoU değerleri, ayrıca temel model özellikleri ayrıntılı biçimde Tablo 3.1’de sunulmuştur.

Tablo 3.1. Geçmiş tiroid segmentasyonu çalışmalarına ait bilgiler

Yıl	Yazarlar	Model	Veri Kümesi	Kayıp Fonksiyonu	Dice Skoru (%)	IoU (%)	Ana Özellikler
2022	Tao, Dang, Shi ve ark.	LCA-Net	TN-SCUI2020 / TN3K	Ağırlıklı CE $(1-p)^2 + p^2$ terimleriyle	90.26 / 82.08	82.65 / 71.18	7×1 & 1×7 Mish omurgası, Bağılamsal Dikkat
2023	Ma, Sun, Liu ve ark.	AMSeg	TN3K / DDTI	Fokal Dice + Çok ölçekli L1 Adverseriyel Kayıp	84.21 / 74.81	72.48 / 60.89	Çok ölçekli ROI kırpma, Adverseriyel birleşim
2023	Bi, Cai, Sun ve ark.	BPAT-Üne	TN3K / DDTI	BCE + Sınır Kaybı + Dice	83.64 / 85.63	71.87 / 74.87	BPSM nokta-dikkat, AMFFM çok ölçekli birleştirme

Tablo 3.1. devam ediyor

2024	Zhang, Wang, Wei ve ark.	DDTransUNet	TN3K / BUS- BRA / CAMUS	BCE + Dice + Adverseriyel Kayıp	82.31 / 88.23 / 90.33	73.82 / 80.75 / 82.51	Çift dallı Swin+CNN kodlayıcı, GSA & GCA + CAF
2024	Liu, Mu, Sun ve ark.	BFG&MSF- Net	TN3K / DDTI	Kombinasyon: %50 BCE + Dice	87.26 / 85.12	76.34 / 74.02	Sınır Yönlendirme Modülü, MSF + DSASPP + İyileştirme Modülü
2024	Ali, Wang, İde	CIL-Net	TN3K / DDTI	BCE + Dice	79.68 / 83.35	64.90 / 73.35	Yoğun bağlantı, Üçlü Dikkat, FIB & RCB

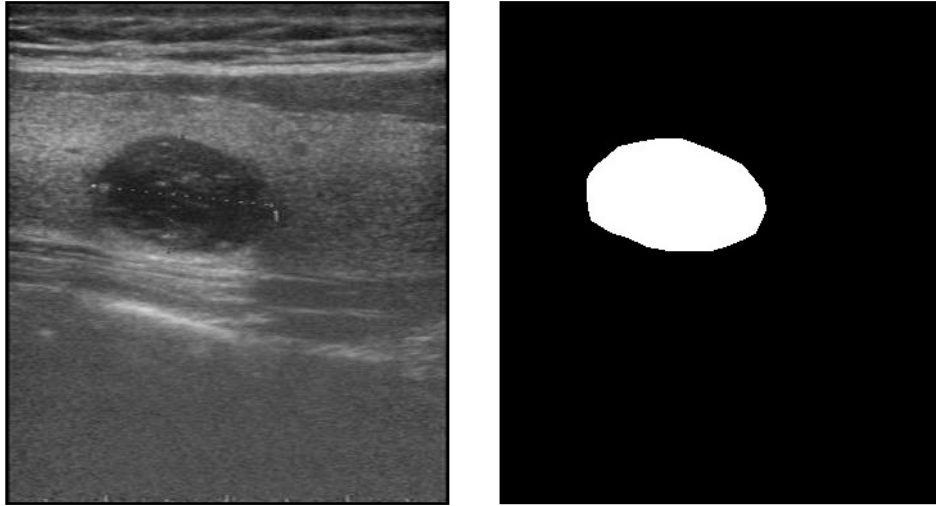
Tablo 3.1. devam ediyor

2024	Deng, Dang, Pan ve ark.	STU3Net	TN3K / BUSI	%30 BCE + %70 Genelleştirilmiş Dice	83.68 / 79.09	74.16 / 66.99	Swin+CNN hibrit, CPCA
2025	Wu, Huang, Yang ve ark.	Improved Swin	TN3K / DDTI	DSA tabanlı Fokal Dice	82.26 / 78.64	73.00 / 67.93	DWConv + MSA/MLP, FFSC
2025	Yetginler & Atacak	Improved V- Net	TN3K / DDTI	Dice + Fokal Tversky	83.88 / 84.51	75.50 / 76.27	SE blokları, Üçlü Birleştirme Kodlayıcı
2025	Li, Fu, Wang ve ark.	GSE-Nets	TUD / DDTI / TN3K	BCE + Dice	89.85 / 83.52 / 81.97	81.60 / 71.71 / 72.24	Yapı Çapraz- Dikkat, Grafik Evrişim Bloğu

4. VERİ KÜMELERİ

4.1. TN3K Veri Kümesi

TN3K (Thyroid Nodule 3000) veri kümesi, Ocak 2016 ile Ağustos 2020 tarihleri arasında GE Logiq E9, Hitachi ARIETTA 850, Mindray RESONA 70B olmak üzere üç farklı ultrasonografi cihazı kullanılarak toplanmış toplam 3.493 adet B-mode gradyan görüntüsünden oluşmaktadır [53]. Bu görüntüler, 2.421 hasta kaydından derlenen ham veri setinden anonimleştirme prosedürleri uygulanarak elde edilmiş ve klinik kullanım amacıyla kayıt altına alınmıştır. Veri toplama sürecinde, aynı hastadan birden fazla görüntü mevcutsa merkeze en yakın nodül içeren kare seçilmiştir. Veri kümesi oluşturulurken bu yaklaşım, segmentasyon doğruluğunu artırmak ve modelin odaklanacağı lezyon bölgelerini standartlaştırmak üzere tasarlanmıştır. Şekil 4.1’de TN3K veri kümesinden örnek bir ultrasonografi görüntüsü ve radyologlar tarafından elde edilmiş maske yer almaktadır.



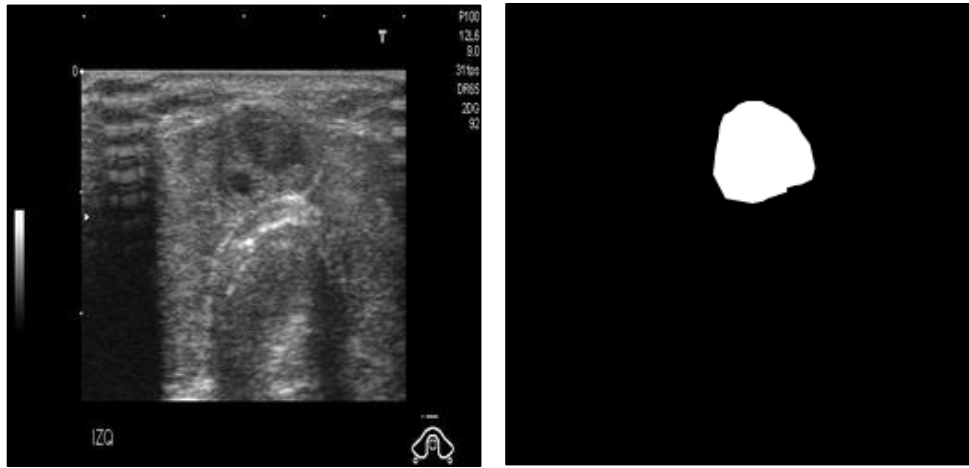
Şekil 4.1. TN3K verisetinden örnek tiroid ultrasonografi görüntüsü ve o görüntüye ait maske

Her bir görüntü için radyologlar tarafından pikseller düzeyinde manuel anotasyon işlemi gerçekleştirilmiş ve maske sınırları iki bağımsız uzman tarafından onaylanmıştır. Sınır belirsizliği ya da düşük kontrastlı alanlardan kaynaklanan anlaşmazlık durumlarında üçüncü bir radyolog görüşü alınarak ortak konsensüs sağlanmıştır. Bu çok adımlı doğrulama protokolü, maskelerin tıbbi gerçekçilik ve tutarlılık açısından yüksek güvenilirlik düzeyi sunmasına katkıda bulunmuştur. Orijinal çözünürlükleri 560×360 veya 512×512 piksel olan

görüntüler, ultrasonografi frekansına bağlı olarak 7–12 MHz arasında değişen lineer probalar ile elde edilmiş olup, tek kanallı (grayscale) DICOM formatında depolanmıştır. Ham görüntüler, ilk aşamada kalite kontrolünden geçirilerek bozuk dosyalar, tekrar kayıtlar ve aşırı artefakt içeren örnekler elenmiştir. Bu adımlar sonucunda 3.235 geçerli görüntü–maske çifti eğitim, validasyon ve test aşamaları için kullanılmaya hazır hale getirilmiştir.

4.2. DDTI Veri Kümesi

DDTI (Digital Database Thyroid Image) veri tabanı, 980 adet tiroid ultrasonografi görüntüsünü ve bu görüntülere ilişkin anotasyon bilgilerini barındırmaktadır. Veri kümesindeki örneklerin 322’si malign, 658’i benign nodül içeriğine sahiptir. Her görüntü, XML formatında sunulan piksel düzeyli maske anotasyonları ile birlikte Bethesda ve TIRADS sınıflandırma skorları olan tamamlayıcı meta-veriler içermektedir. Bu kapsamlı etiketleme hem segmentasyon hem de malignite sınıflandırma deneylerinde çift amaçlı kullanım imkânı tanımaktadır [54]. Şekil 4.2’de DDTI veri kümesinden alınan bir adet tiroid ultrasonografi görüntüsü ve bu görüntüye ait XML dosyasındaki anotasyonların maskeye dönüştürülmüş görüntüsü sunulmaktadır.



Şekil 4.2. DDTI veri kümesinden tiroid ultrasonografi görüntüsü örneği ve bu örneğe ait ikileştirilmiş maskesi

Görsel veriler, 6–15 MHz frekans aralığındaki lineer probalar ile 540×480 piksel çözünürlükte elde edilmiştir. Anotasyon sürecinde üç bağımsız radyolog her görüntü için maske çizimi yapmış, maske tutarlılığı Dice skoru ile ölçülerek ortalama %0,92 oranında yüksek uyum sağlandığı belirlenmiştir. Belirsiz sınır durumunda uzmanlar arası toplantılar

gerçekleştirilerek ortak konsensüs oluşturulmuştur. Malignite risk grubu tayini için FDA onaylı biyopsi sonuçları veri kümesine entegre edilmiş, böylece model eğitimi hem segmentasyon hem de patoloji temelli sınıflandırma odaklı çoklu görev senaryolarına uygun hâle getirilmiştir. Kalite kontrol aşamasında, eksik veya bozuk XML dosyaları, hatalı maske eşleşmeleri ve aşırı artefakt içeren görüntüler çıkarılmıştır. Sonuçta 635 geçerli görüntü–maske çifti elde edilmiştir.

4.3. Veri Kümelerinde Ön İşleme

TN3K ve DDTI kümelerinin birbirinden farklı kaynak, çözünürlük ve anotasyon özelliklerini ortak bir derin öğrenme hattına uyumlu hâle getirmek amacıyla bazı ön işleme adımları uygulanmıştır. Bu adımlar, verilerin belleğe alınmasından başlayarak modelin doğrulama-aşamasındaki girdilerine kadar uzanan ardışık adımlar dizisinden oluşmaktadır. Her adım, klinik görüntü kalitesini korurken modelin genellenebilirliğini azamî düzeye çıkarmayı hedefleyen tutarlı tasarım kararları içermektedir.

İlk aşamada, TN3K kümesinin Kaggle platformundan sağlanan 3.493 adet, 256×256 piksel boyutlu PNG görüntüsü ve karşılık gelen ikili maskeleri, kayıpsız olarak OpenCV aracılığıyla okunmuş ve float32 formatında NumPy dizilerine dönüştürülmüştür. DDTI kümesinde ise ultrasonografi kareleri JPEG biçiminde, maske anotasyonları ise XML dosyalarında çokgen koordinatları şeklinde sunulmuştur. XML etiketleri, standart bir 540×480 piksel tuval üzerine çizilip ikili maske görüntülerine dönüştürülmüş ve benzer biçimde NumPy dizilerine aktarılmıştır. Elde edilen tüm diziler, eğitim, doğrulama ve test alt kümelerine göre bölümlenerek bağımsız “.npy” dosyaları hâlinde saklanmıştır. Böylelikle diskten okuma süresi kısaltılmış ve I/O darboğazı önlenmiştir.

İkinci aşamada, görüntüler aynı boyuta getirilmiştir. TN3K örnekleri özgün hâliyle 256×256 piksel olduğundan ek ölçeklemeye ihtiyaç duymamıştır. Buna karşılık DDTI kareleri, en-boy oranı korunacak biçimde 256 piksele yeniden ölçeklendirilmiş ve oran koruma sırasında oluşan boş kenarlar yansıtmalı dolgu yöntemiyle tamamlanmıştır. Böylece nodülün sınır bölgelere düştüğü vakalarda dahi anatomik bilgi kaybı yaşanmamış. Ayrıca evrişim işlemlerinin kenar etkileri en aza indirilmiştir.

Üçüncü adım piksel düzeyinde standardizasyonu kapsamaktadır. Tüm görüntüler, kendi iç maksimum yoğunluk değerine bölünerek $[0, 1]$ aralığına gerilmiştir. Bu sayede farklı üreticilere ait ultrasonografi cihazlarının parlaklık ve kontrast varyasyonları

dengelenmiştir. Bunu takiben, sayısal kararlılığı artırmak amacıyla z-skor ölçeklendirmesi (ortalama = 0, sapma = 1) uygulanmıştır. Maskeler bu noktada 127 gri-seviye eşiği kullanılarak yeniden ikilileştirilmiş (binarize) ve yalnızca 0 ile 1 değerlerini içerecek şekilde kaydedilmiştir.

Dördüncü adım, eğitim verisinin çeşitlendirilmesine yöneliktir. TN3K eğitim alt kümesi 2 070, DDTI eğitim alt kümesi ise 406 özgün örnekten oluşmaktadır. Aşırı uyumu sınırlamak ve farklı hasta popülasyonlarına aktarım yeteneğini artırmak amacıyla her özgün görüntüye dört rastlantısal dönüşüm olan yatay/dikey çevirme, doksan derecelik rastgele döndürme, elastik deformasyon ve parlaklık–kontrast modifikasyonu uygulanmıştır. Böylelikle TN3K eğitim kümesi 12.420, DDTI eğitim kümesi ise 2.436 örneğe ulaşmıştır.

5. YÖNTEMLER

Bu bölümde, geliştirilen segmentasyon yaklaşımının yapıtaşları sunulmaktadır. Öncelikle, yerel ayrıntı ve küresel bağlam bilgisini en verimli biçimde yakalayabilmek amacıyla değerlendirilen evrişimli ve dönüştürücü tabanlı kodlayıcı yaklaşımları kısaca tanıtılmakta ve seçim sürecine dair genel esaslar açıklanmaktadır. Bunu takiben, çok-ölçekli özelliklerin kayıpsız yeniden yapılandırılmasını hedefleyen kod çözücü tasarımının temel ilkelerine değinilmektedir. Ardından, deneysel kısımda referans alınacak güncel karşılaştırma modelleri yöntemsel bakımdan özetlenerek değerlendirme zemini netleştirilmektedir. Bölümün devamında, EfficientNet-B7 ve ViT-16 tabanlı çift kodlayıcıyı CBAM destekli birleştirme blokları ve UNet3+ tabanlı derin gözetimli kod çözücü ile bütünleştiren önerilen hibrit EffViT3+ ismi konulan mimari betimlenmektedir. Son olarak, alan doğruluğu ile sınır duyarlılığını eşzamanlı optimize etmek üzere tasarlanan kayıp fonksiyonu bileşenleri açıklanarak bölüm tamamlanmaktadır.

5.1. Kodlayıcı Yaklaşımları

5.1.1. Evrişim Tabanlı Yaklaşımlar

ResNet-101, artık bağlantıları (residual connections) sayesinde derin yapılarda rastlanan gradyan sönümlenmesi sorununu hafifleten katman-zengin bir evrişimli sinir ağı omurgasıdır. Her darboğaz (bottleneck) bloğu, 1×1 daraltma, 3×3 derin evrişim ve 1×1 genişletme evrelerinden oluşmaktadır [559]. Blok girdisine eklenen kimlik atlama bağlantısı, özellik akışını ve geriye yayılımı kesintisiz kılarak daha hızlı ve kararlı öğrenme sağlamaktadır. “Batch Normalization” katmanları iç dağılım değişimini dengelemekte, ReLU aktivasyonları ise doğrusal olmayan temsil gücünü artırmaktadır [56]. ImageNet üzerinde ön-eğitilmiş ağırlıkları sayesinde ResNet-101, sınırlı tıbbi görüntü verilerinde dahi güçlü aktarım performansı sunarak yerel dokusal ayrıntıları güvenilir biçimde yakalayabilen sağlam bir evrişim tabanlı kodlayıcı yaklaşımı olarak değerlendirilmiştir.

DenseNet-121, sık bağlantılı (dense connectivity) mimarisiyle her katmanın çıktısını arkasından gelen tüm katmanlara doğrudan ileterek öz nitelik (feature) yeniden kullanımını ve gradyan akışını en üst düzeye çıkaran bir evrişim omurgasıdır. Bu yaklaşım, klasik

artıksal (residual) yapılara kıyasla hem parametre sayısını hem de hesap yükünü önemli ölçüde azaltırken temsil gücünü korumaktadır. Ayrıca erken katmanlarda öğrenilen ince dokusal ayrıntıların derin katmanlarda kaybolmadan kullanılmasına olanak tanımaktadır. Yoğun bağlantı blokları içinde uygulanan büyüyen “growth rate” ilkesi, her katmanda üretilen yeni özellik haritalarının sayısını kontrol ederek bellek tüketimini dengelemektedir [58]. ImageNet üzerinde ön-eğitilmiş ağırlıkları sayesinde DenseNet-121, sınırlı ultrasonografi verisinde dahi ayrıntılı kenar yapılarını yakalayabilen, parametre-verimli bir evrişim tabanlı kodlayıcı yaklaşımı olarak bu çalışmada değerlendirilmiştir.

VGG-16, toplam 13 evrişim ve 3 tam bağlantılı katmandan oluşan derin bir omurgadır. Modelin temel fikri, büyük evrişim çekirdekleri yerine ardışık 3×3 'lük küçük filtreler kullanarak ayrıntılı özellikleri kademeli biçimde çıkarmaktır. Bu sayede hem parametre sayısı tolere edilebilir seviyede kalmakta hem de ağ derinliği arttırılmaktadır. Her iki evrişim katmanından sonra gelen 2×2 maksimum havuzlama birimleri, görüntü boyutunu yarıya indirerek hem bellek tüketimini azaltmaktadır hem de ağın daha geniş alanı görmesini sağlamaktadır [58]. Evrişim bloklarının çıkışında kullanılan ReLU etkinleştirmeleri, öğrenmeyi hızlandırırken gradyan kaybolmasını önlemektedir. Son bölümde yer alan üç tam bağlantılı (fully connected) katman, öğrenilen tüm özellikleri özetleyerek güçlü bir küresel temsil sunmaktadır [59]. Bu sayede VGG-16, sınırlı sayıda tiroid ultrasonografi görüntüsüyle çalışırken bile ince kenar ayrıntılarını güvenilir biçimde yakalayabilen çalışmada kullanılmak üzere yerel özellik çıkarımı için tercih edilmiş olan sağlam bir evrişim tabanlı kodlayıcı yaklaşımıdır.

ResNeXt, evrişimli sinir ağlarında model kapasitesini artırmak için “cardinality” adı verilen ek bir tasarım boyutu kullanan bir kodlayıcı mimarisidir. Her ResNeXt bloğu, giriş özelliğini parçalayarak çok sayıda küçük alt dala ayırmaktadır. Bu dalların her birinde gruplanmış evrişim uygulanmakta ve sonuçlar yeniden toplanmaktadır [60]. Böylece ağ, filtre sayısını veya katman derinliğini büyütmeden farklı görsel örüntüleri paralel olarak öğrenebilmektedir. Blok girişinin doğrudan çıkışa eklendiği artık (residual) bağlantı, gradyan akışını kolaylaştırarak derin modellerde eğitimin kararlı sürmesini sağlamaktadır [61]. ResNeXt hem kenar ayrıntılarını hem de geniş bağlamsal ilişkileri yakalama yeteneği nedeniyle bu tezde evrişim tabanlı kodlayıcı yaklaşımları arasında değerlendirilmiştir.

EfficientNet-B7, ağ derinliği, genişliği ve giriş çözünürlüğünü dengeli biçimde büyüten bileşik ölçekleme (compound scaling) stratejisinin en geniş varyantlarından biridir.

Mimari, MobilNetV2'den uyarlanan MBConv blokları ile çalışmaktadır. Her blokta 1×1 genişletme, derin ayrılabilir (depthwise separable) 3×3 evrişim ve 1×1 daraltma katmanları yer almaktadır. Blok içindeki SE modülleri, kanal düzeyinde öğrenilen dikkat ağırlıkları sayesinde en bilgilendirici öz nitelikleri öne çıkararak ağırlı temsil gücünü arttırmaktadır [62, 63]. ImageNet üzerinde ön-eğitilmiş ağırlıklarıyla, sınırlı tiroid ultrasonografi verilerinde bile hem ince kenar ayrıntılarını hem de geniş bağlamsal bilgiyi yakalayabilen güçlü bir evrişim tabanlı kodlayıcı yaklaşımı olarak bu tezde değerlendirilmiştir. ImageNet üzerinde ön-eğitilmiş ağırlıklarıyla, tiroid ultrasonografi görüntülerinde hem ince kenar ayrıntılarını hem de geniş bağlamsal bilgiyi güvenilir biçimde yakalayabilen güçlü bir evrişim tabanlı kodlayıcı yaklaşım olarak bu tezde tercih edilmiştir.

5.1.2. Transformatör Tabanlı Yaklaşımlar

Swin Transformer (Swin-T), pencere tabanlı kendiliğinden dikkat (window self-attention) mekanizmasıyla ölçeklenebilirliği artırılmış, evrişim benzeri yerel odaklanmayı korurken uzun menzil bağlam bilgisi yakalayabilen bir dönüştürücü mimarisidir. Model, görüntüyü sabit boyutlu pencerelere ayırarak dikkat hesabını her pencerede ayrı yürütmektedir. Ardışık katmanlarda pencereleri kaydırarak (shifted window) bölgesel sınırlar arası bilgi akışını sağlamaktadır [64, 65]. ImageNet üzerinde ön-eğitilmiş Swin-T, tiroid ultrasonografi görüntülerinde hem ayrıntılı kenar yapısını hem geniş bağlamı yakalayabilen, transformatör tabanlı güçlü bir kodlayıcı yaklaşımıdır.

CoAtNet-2, evrişim ve dikkat mekanizmalarını ardışık bloklarda birleştirerek her iki yaklaşımın en kuvvetli yönlerini dengeli biçimde sunan hibrit bir omurgadır. Modelin erken katmanlarında kullanılan derinlik-ayrık evrişimler, yerel dokusal ayrıntıları parametre-verimli şekilde öğrenirken üst katmanlardaki çok-kafalı dikkat blokları, küresel bağlam ilişkilerini modellemeyi mümkün kılmaktadır. Her bloğun sonunda eklenen SE modülleri kanal düzeyinde önem ağırlıkları üretmektedir [66]. Böylece ağ en bilgilendirici özellikleri öne çıkarabilmektedir. Bileşik ölçekleme kuralı sayesinde CoAtNet-2, benzer doğrulukta CNN ve saf dönüştürücü modellere nazaran daha düşük FLOP ile çalışmaktadır. Bu özellikleriyle CoAtNet-2, sınırlı ultrasonografi verisinde hem kenar netliği hem de bağlamsal tutarlılık talep eden segmentasyon görevleri için ideal hibrit kodlayıcı yaklaşımları arasında değerlendirilmiştir.

Vision Transformer (ViT), görüntüyü sabit boyutlu yama (patch) dizilerine bölerek her yamayı dil işleminde kullanılan kelime vektörleri benzeri bir dizi girdi olarak ele alan ilk saf dönüştürücü mimarisidir [67]. Her yama lineer projeksiyonla yüksek boyutlu gömme (embedding) uzayına aktarılmaktadır, ardından eklenen konum kodlamaları yamanın uzamsal bilgisini korumaktadır. Çok-kafalı kendiliğinden dikkat katmanları, tüm yamalar arasında doğrudan bağıntılar kurarak geniş bağlamı tek adımda kapsamaktadır. Çok Katmanlı Algılayıcı (Multilayer Perceptron, MLP) katmanları ise kanal içi özellik etkileşimlerini derinleştirmektedir [68, 69]. Bu tezde ViT, küresel bağlamı eksiksiz yakalama yeteneği sayesinde hibrit sistemlerde evrişim tabanlı kodlayıcılara tamamlayıcı olarak konumlandırılmış güçlü bir dönüştürücü yaklaşımdır.

5.2. Önerilen Model: EffiViT3+

Tiroid nodüllerini ultrasonografi görüntülerinden otomatik olarak segmente etmek amacıyla EffiViT3+ adında hibrit, derin öğrenme tabanlı bir yöntem önerilmiştir. EffiViT3+, yerel öznitelik çıkarımı için EfficientNet-B7'yi, küresel bağlamsal bilgiyi yakalamak için ise Vision Transformer (ViT) mimarisini birleştiren yeni bir mimariye sahiptir. İlk aşamada, her iki ağdan farklı ölçeklerde elde edilen öznitelik haritaları, Evrişimli Blok Dikkat Modülü (Convolutional Block Attention Module, CBAM) ile zenginleştirilmiş özel füzyon blokları aracılığıyla sistematik olarak birleştirilmiştir. Bu hibrit kodlayıcının çıktıları ASPP ile çok-ölçekli ayrıntılarla zenginleştirilmiş ve ardından UNet3+ tabanlı bir kod çözücünden geçirildikten sonra sonraki işleme (post processing) yapıldıktan sonra nihai segmentasyon haritaları oluşturulmuştur. Şekil 5.1'de modele ait bileşenler detaylı olarak gösterilmiştir.

5.2.1. Kodlayıcı Bileşenleri

Modelin kodlayıcı bölümü, yerel ayrıntıları çıkaran EfficientNet-B7 ile öz-dikkat mekanizmaları aracılığıyla yerel bağlam ilişkilerini öğrenen ViT-16'dan oluşan iki paralel bileşenden meydana gelmektedir. Bu iki bileşenden elde edilen çok ölçekli öznitelik haritaları, "FusionBlock" içinde CBAM ile zenginleştirilip birleştirilerek kod çözücüye aktarılmaya hazır hâle getirilmektedir.

EfficientNet-B7, derinlik, genişlik ve çözünürlük boyutlarını dengeli biçimde ölçeklendiren bileşik ölçekleme (compound scaling) prensibine dayalı, yüksek performanslı bir CNN modelidir [70]. MobilNetV2 tabanlı MBConv blokları ve kanal düzeyinde dikkat

sağlayan SE modüllerini kullanarak, daha az parametreyle güçlü temsil gücü sunmaktadır. “Inverted residual” bloklarında önce 1×1 genişletme evrişimi ile kanal sayısı arttırmaktadır. Sonra 3×3 ayırık (depthwise separable) evrişim ile yerel öznitelikler çıkarılmakta ve son olarak bilgi akışını korumak için 1×1 daraltma evrişimi uygulanmaktadır [70]. Her blokta yer alan SE modülleri, küresel ortalama havuzlama ile kanal önemi ağırlıklarını öğrenip en anlamlı öznitelik haritalarını vurgulamaktadır [71]. Tüm evrişim katmanlarından sonra kullanılan “Batch Normalization” ve “Swish” aktivasyonları ise eğitim sırasında gradyan kararlılığını ve öğrenme hızını arttırmaktadır [72, 73]. Önerilen EffiViT3+ modelinde, önceden eğitilmiş EfficientNet-B7 kodlayıcısı sınıflandırma katmanı atlanarak yalnızca beş farklı hiyerarşik düzeyde ara öznitelik haritası çıkarmaktadır. Bu haritalar 1×1 evrişimlerle ViT’den gelen küresel özniteliklerle kanal uyumlu hâle getirilerek CBAM destekli “FusionBlock”larda harmanlanmaktadır ve UNet3+ tabanlı kod çözücü hattına aktarılmaktadır.

Vision Transformer (ViT-16), görüntüyü sabit boyutlu yamalara (patch) bölerek bunlar arasındaki uzun menzilli ilişkileri transformatör kodlayıcı bloklarıyla öğrenen bir mimaridir [74, 75]. Önce görüntü, 16×16 boyutlu 256 adete ayrılmaktadır. Her yama lineer bir katmanla 1024 boyutlu gömme vektörüne dönüştürülmektedir ve hem içerik hem de mekânsal konumu korumak için pozisyon kodlamaları eklenmektedir [74]. Ardından 24 ardışık kodlayıcı bloğu uygulanmaktadır. Her blok LayerNorm, 16 başlı öz-dikkat katmanı, yeniden LayerNorm ve dört katmanlı bir MLP içermektedir. Artık bağlantılar bilgi akışını ve gradyan kararlılığını güçlendirmektedir [75]. Eğitimden sonra 256 token’in ilki sınıflandırma için ayrılmakta geriye kalan 255 yama token’i (B, 1024, 16, 16) boyutlu uzamsal öznitelik haritasına dönüştürülmektedir [74]. EffiViT3+’te önceden eğitilmiş ViT bloğu 256×256 boyutundaki ultrasonografi görüntülerine uygulanmaktadır. Sınıf (CLS) token’i ayrılarak (B, 256, 1024) boyutlu bir dizi elde edilmektedir. Bu dizi (B, 1024, 16, 16) boyutlu bir öznitelik haritasına dönüştürülmektedir. Ardından 1×1 projeksiyon evrişimleriyle EfficientNet-B7’den gelen yerel özniteliklere uyum sağlar duruma getirilmektedir. Son olarak CBAM destekli FusionBlock’lar ile birleştirilip UNet3+ tabanlı kod çözücüye iletilmektedir.

5.2.2. Çok Özellikli Öznitelik Füzyonu

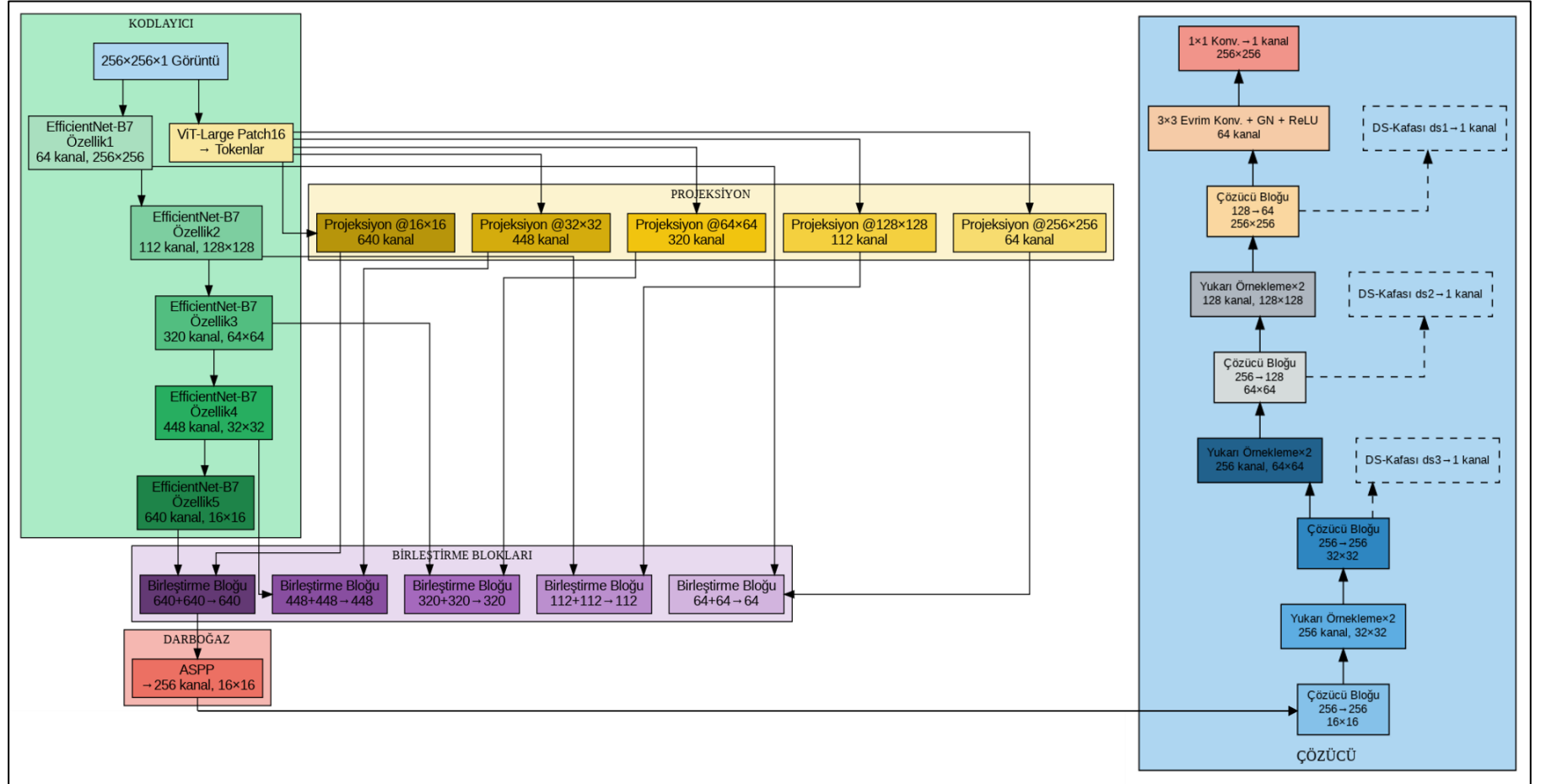
Model, farklı çözünürlüklerdeki öznitelik haritalarını bir araya getirerek hem ince sınır ayrıntılarını hem de geniş bağlamsal bilgiyi aynı anda değerlendirmeyi sağlamaktadır.

Fusion Block, her ölçekten gelen yerel (EfficientNet-B7) ve küresel (ViT) öznitelikleri toplamaktadır. Önce bu haritalar 1×1 evrişimlerle aynı kanal sayısına indirilmekte ve birleştirilmektedir. Ardından CBAM iki aşamada devreye girmektedir. Birinci aşamada Kanal Dikkati ile global ortalama ve maksimum havuzlama yoluyla kanal önem katsayıları öğrenilip sigmoid ile ağırlıklandırılmaktadır. İkinci aşamada Uzamsal Dikkat ile 7×7 evrişimler aracılığıyla en kritik piksel bölgeleri belirlenmektedir. Sonuç olarak hem kanal hem de mekânsal boyutta filtrelenmiş ve bilgi açısından zengin bir öznitelik haritası elde edilmektedir [75].

ASPP, kodlayıcının en derin katmanından gelen düşük çözünürlüklü haritayı dört farklı dilasyon oranıyla (1, 6, 12, 18) işler ve global özet için eklenen ortalama havuzlama dalıyla birlikte beş ayrı çıktı üretmektedir. Bu çıktılar 1×1 evrişim ve normalizasyon katmanlarıyla tek bir haritada birleştirilmektedir. Böylece model, yakın çevre desenlerini ve tüm görüntü düzeyindeki yapısal ilişkileri aynı anda yakalayarak segmentasyon görevinde genelleme kabiliyetini güçlendirmektedir. Şekil 5.2, beş farklı atrous evrişim ve global havuzlama dalında (her biri 256 kanallı) girdi olarak gelen $16 \times 16 \times 640$ özellik haritasını işleyen ve bunları birleştirerek 1×1 evrişim ile 256 kanala indirgenmiş çıktıyı elde eden ASPP bloğunu göstermektedir [75].

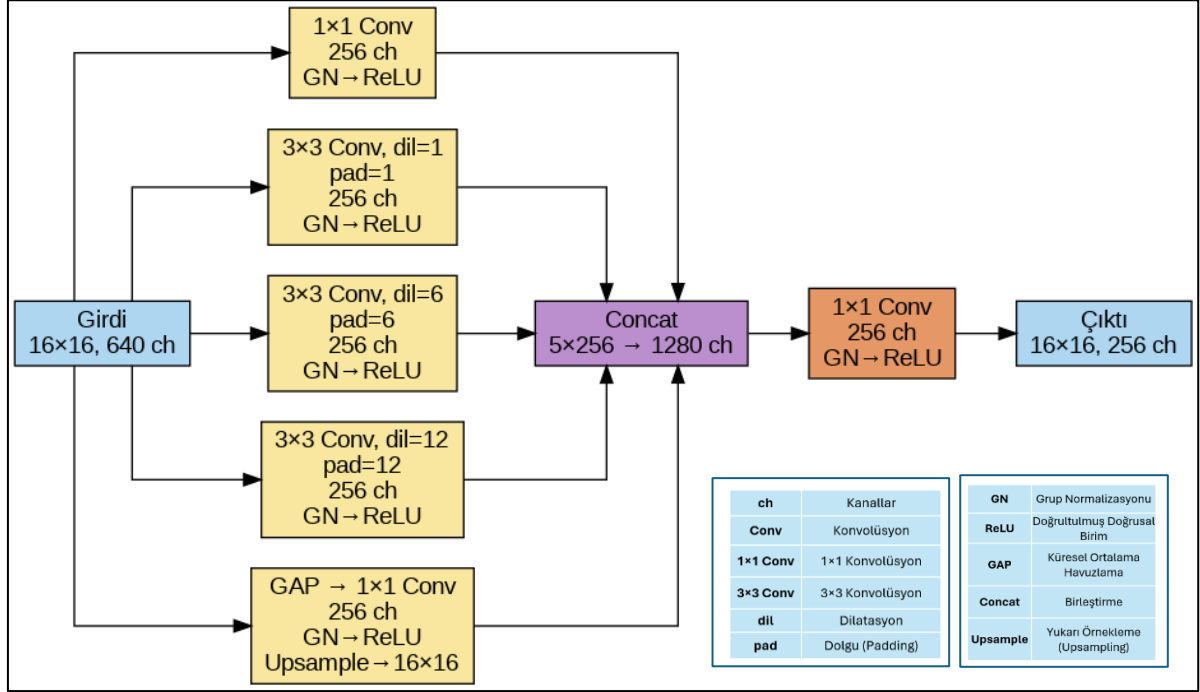
5.2.3. Kod Çözücü Aşaması

Kod çözücü aşamasında, FusionBlock ve ASPP'den gelen çok ölçekli öznitelikler UpSampleBlock ve DecoderBlock yapılarıyla kademeli olarak orijinal görüntü boyutuna yükseltilmektedir ve CBAM destekli dikkat mekanizmalarıyla işlenmektedir. UpSampleBlock, ConvTranspose2d kullanarak mekânsal boyutu iki katına çıkartmakta ardından GroupNorm ve ReLU uygulamaktadır [78, 79]. DecoderBlock ise iki aşamalı 3×3 evrişim, dropout ve CBAM tabanlı dikkat katmanlarından oluşmaktadır. Ek olarak 1×1 artık bağlantısı gradyan akışını korumaktadır. Atlama yapıları, kodlayıcının erken katmanlarındaki ince ayrıntıları doğrudan çözücüye aktararak hem yüksek çözünürlüklü yerel hem de soyut özniteliklerin eşzamanlı kullanılmasını sağlamaktadır.



Şekil 5.1. EffViT3+ modelinin bileşenleri

Son olarak, derin gözetim mekanizması her kod çözücü seviyesinden ara çıktılar üreterek ana kayıp fonksiyonuna dahil edilmesini sağlamaktadır. Böylece model hem sınır duyarlılığını hem de bölge doğruluğunu aynı anda optimize edebilmektedir.



Şekil 5.2. Atröz evrişim, global havuzlama, concat ve 1x1 projeksiyonun beş aşamasını gösteren ASPP bloğunun mimari akış diyagramı

5.2.4. Sonraki İşleme

Sonraki işleme (Post-processing) aşamasında uygulanan morfolojik açma (opening) işlemi, ikili segmentasyon maskesinden küçük parlak artefaktları ve tüy gibi izole gürültü noktalarını uzaklaştırmak amacıyla iki ardışık yapılandırılmış dönüşümden oluşmaktadır. İlk adımda, 3x3 veya 5x5 boyutlu ve tüm elemanları “1” olan kare biçimli bir çekirdek (kernel) kullanılarak erozyon (erosion) uygulanmaktadır. Çekirdek görüntü üzerinde kaydırıldıkça her lokal pencere içindeki en düşük değer yeni merkez piksel olarak seçilmektedir. Böylece çekirdek boyutundan daha küçük yapılar aşındırılarak yok edilmektedir. Erozyon sonucunda inceltilen ana segment bölgeleri, ikinci aşamada aynı çekirdekle gerçekleştirilen genleşme (dilation) işlemiyle kısmen orijinal formuna geri kazandırılmaktadır. Bu aşamada her pencere içindeki en yüksek değer seçilerek erozyonla ortadan kalkan küçük gürültü lekeleri

yok edilirken temel nesne yapısı korunmaktadır. Kernel boyutu ve yineleme sayısı (iterations) parametreleri hem artefakt temizleme düzeyini hem de segment sınırlarının korunma derecesini kontrol edecek şekilde ayarlanarak nihai maskenin daha düzgün ve sürekli olmasını sağlamaktadır.

5.3. Kayıp Fonksiyonları

Model eğitiminde, hem piksel düzeyindeki örtüşme oranını maksimize etmek hem de maskelerin kenar doğruluğunu artırmak amacıyla birden fazla kayıp fonksiyonu eşzamanlı olarak kullanılmıştır.

5.3.1. Dice Kaybı

Gerçek ve tahmin edilen maskeler arasındaki örtüşmeyi ölçen Dice Kaybı, iki maske arasındaki kesişim/birleşim oranını temel almaktadır. Model, tahmin edilen ve gerçek maskelerdeki piksellerin mümkün olduğunca üst üste gelmesini sağlamak için bu fonksiyonla eğitim gerçekleştirilmiştir. Böylece küçük yapısal detaylarda dahi yüksek örtüşme elde edilebilmektedir [80]. Denklem (5.1)'de p_n sigmoid sonrası modele ait o piksel için verilen olasılık; g_n gerçek maskedeki ilgili pikselin 1 veya 0 değerini; ε ise sıfıra bölünmeyi engelleyen küçük bir sabiti ifade etmektedir.

$$\mathcal{L}_{Dc} = 1 - \frac{2 \sum_i p_i g_i + \varepsilon}{\sum_i p_i + \sum_i g_i + \varepsilon} \quad (5.1)$$

5.3.2. Lovász Hinge Kaybı

Lovász Hinge Kaybı, her piksel hatasını Jaccard gradyanları ile ağırlıklandırarak IoU metriğini doğrudan optimize etmektedir [81]. Bu sayede segmentasyonun bölgesel bütünlüğü ve sınır hassasiyeti eşzamanlı olarak iyileştirmektedir. Denklem (5.2)'de x_i model logiti, g_i pikselin gerçek etiketi (0 veya 1), $\Delta J(k)$ ise sıralanmış hatanın Jaccard indeksi gradyanını göstermektedir.

$$\mathcal{L}_{Lvs} = \frac{1}{B} \sum_{b=1}^B \sum_{k=1}^n \max \left(0, 1 - x_{(k)}^{(b)} (2g_{(k)}^{(b)} - 1) \right) \Delta J_{(k)}^{(b)} \quad (5.2)$$

5.3.3. Fokal Tversky Kaybı

Fokal Tversky Kaybı, sınıflar arası dengesizliği dikkate alan Tversky indeksini temel almaktadır ve odak katsayısı (γ) ekleyerek özellikle zor örneklerle ağırlık vermektedir. Böylece ince nodül sınırlarının doğru öğrenilmesini sağlamaktadır [82]. Denklem (5.3)'te

- $\sum_i p_i g_i$ gerçek pozitifleri,
- $\sum_i p_i (1 - g_i)$ yanlış pozitifleri,
- $\sum_i (1 - p_i) g_i$ yanlış negatifleri ifade etmektedir.

α ve β bu hatalara verilen ağırlıklar, γ ise odak katsayısıdır.

$$TI = \frac{\sum_i p_i g_i}{\sum_i p_i g_i + \alpha \sum_i p_i (1 - g_i) + \beta \sum_i (1 - p_i) g_i}, \quad \mathcal{L}_F = (1 - TI)^\gamma \quad (5.3)$$

5.3.4. Sınır Kaybı

Sınır Kaybı, tahmin edilen ve gerçek maskelerin sınır çizgilerindeki farklılıkları vurgulamaktadır. Her iki maskenin sınır bölgelerini çıkararak bu sınırlar arasındaki mesafeyi ölçmekte ve hataları cezalandırmaktadır. Böylece segmentasyon sınırlarında net ayrımlar elde edilmektedir [81]. Denklem (5.4 – 5.6)'da

- P post-sigmoid tahmin maskesi,
- G gerçek maskeyi,
- AvgPool 3×3 ortalama havuzlamayı,
- $\|\cdot\|_1$ ise her pikselde mutlak farklar toplamını ifade etmektedir.

$$E(P) = \text{AvgPool}_{3 \times 3}(P) - P \quad (5.4)$$

$$E(G) = \text{AvgPool}_{3 \times 3}(G) - G \quad (5.5)$$

$$\mathcal{L}_{\text{Budr}} = \|E(P) - E(G)\|_1 \quad (5.6)$$

5.3.5. Normalize Edilmiş HD95 Kaybı

Normalize Edilmiş HD95 Kaybı, maskelerin kenar sapmalarını %95 Hausdorff mesafesi üzerinden değerlendirilmektedir ve görüntünün köşegen uzunluğuna göre

normalize edilmektedir. Amaç, kenar hatalarını en aza indirirken en uç sınır farklılıklarına duyarlı kalmaktır. Denklem (5.7 – 5.8)’de

- $\partial P, \partial G$ sınır noktalarını,
- $d()$ Öklidyen mesafeyi,
- H, W ise görüntü boyutlarını ifade etmektedir.

$$d_{95}(P, G) = \text{percentile}_{95} \left(\left\{ \min_{q \in \partial G} d(p, q) \right\}_{p \in \partial P} \cup \left\{ \min_{p \in \partial P} d(q, p) \right\}_{q \in \partial G} \right) \quad (5.7)$$

$$\mathcal{L}_{H95} = \frac{d_{95}(P, G)}{\sqrt{H^2 + W^2}} \quad (5.8)$$

5.3.6. Bileşik Kayıp

Tüm bu kayıplar, bölgesel örtüşme, sınır hassasiyeti ve sınıflar arası dengesizlik gereksinimlerini dengelemek üzere Denklem (5.9)’daki ağırlıklarla birleştirilmektedir.

$$\mathcal{L}_{\text{total}} = 0.30 \mathcal{L}_{\text{LVS}} + 0.25 \mathcal{L}_{\text{DC}} + 0.25 \mathcal{L}_{\text{F}} + 0.15 \mathcal{L}_{\text{Budr}} + 0.05 \mathcal{L}_{\text{H95}} \quad (5.9)$$

6. DENEYSEL SONUÇLAR

Bu bölümde, önerilen yönteme ait sonuçlar öncelikle ana veri seti olarak kullanılan TN3K üzerinde kapsamlı bir şekilde verilmiştir. TN3K’da beş katlı çapraz doğrulama ile eğitim, doğrulama ve test bölmelerini oluşturarak uygun hiperparametre ayarları hem evrişimli hem de transformaötr tabanlı kodlayıcı mimarilerinin karşılaştırması, eğitim sırasındaki kayıp fonksiyonlarının değişimi ve bileşenlerin katkılarını ortaya koyan ablasyon çalışmaları sunulacaktır. Elde ettiğimiz nicel metrikler ve niteliksel maske görselleştirmeleri de sunulmuştur. Ayrıca, modelin genellenebilirliğini sınamak amacıyla ikinci bir veri seti olan DDT üzerinde de ayrı bir eğitim ve doğrulama süreci gerçekleştirilmiş ve performans ölçümleri verilmiştir. Son olarak, parametre sayısı, bellek kullanımı ve çıkarım süreleri üzerinden hesaplama verimliliğimizi de değerlendirilmiştir.

6.1. DeneySEL Kurulum

Deneyler, Ubuntu 20.04 işletim sistemi ve Python 3.8 tabanlı Google Colab sanal ortamda yürütülmüştür. Donanım olarak NVIDIA A100 GPU (40 GB VRAM), 256 GB RAM ve NVMe SSD depolama kullanılmıştır. PyTorch 1.12 ile timm, albumentations, torch-ema, torchmetrics, OpenCV ve MedPy kütüphaneleri entegre edilmiştir. Eğitim döngüsünün her katında (fold) önce veri yükleyici (DataLoader) ile batch’ler oluşturulmuş, ardından önceden tanımlı kodlayıcı ve kod çözücü bileşenleriyle model kurulumu yapılmış, optimizasyon aşamasında AdamW optimizatörü ve CosineAnnealingLR scheduler kullanılmış ve döngü (epoch) sonunda doğrulama (validation) seti üzerinden IoU metriği hesaplanarak iyileşme durmazsa belirlenen sabır (patience) sayısınca erken durdurma (early stopping) yapılmıştır.

6.1.1. Veri Kümelerini Bölme Stratejisi

Ana deneylerin gerçekleştirildiği TN3K veri seti, hasta düzeyinde gruplanmış yaklaşık 3.493 ultrasonografi görüntüsünden oluşmaktadır ve 5 katlı çapraz doğrulama tekniği uygulanmıştır. Bu görüntülerin rastgele %20’si test seti olarak ayrılmış olup geriye kalan görüntüler ise rastgele beş eşit parçaya bölünmüştür. Beş iterasyon boyunca, her

defasında bir alt küme doğrulama (validation) seti, kalan dört alt küme ise eğitim (training) seti olarak kullanılmıştır.

Önerilen modelin farklı dağılımlara uyum kabiliyetini ölçmek için DDTI kümesinden %20 oranında (94 örnek) bağımsız bir test seti ayrılmıştır. Geri kalan veriyi beş alt kümeye bölünmüştür. Her katta dört alt küme eğitim, bir alt küme validasyon için kullanılmıştır. Beş katlı çapraz doğrulama ile hem TN3K’de hem de farklı bir veri setinde modelimizin tutarlılık ve genelleme gücünü kapsamlı biçimde irdelenmiştir.

6.1.2. Hiperparametreler

Modelin öğrenme sürecini dengelemek ve aşırı uyumu önlemek amacıyla seçilen hiperparametreler, hem literatürde yaygın kullanılan değerler hem de sunucumuzun donanım kısıtlarına göre belirlenmiştir. Tablo 6.1’de seçilen hiperparametreler ve değerleri detaylı olarak gösterilmiştir. Eğitimde maksimum 50 döngü erken durdurma mekanizmasıyla birleştiğinde modelin yeterince öğrenmesini sağlarken gereksiz uzayan eğitimleri keserek aşırı uyumu önlemiştir. GPU bellek kullanımını ve eğitim hızı optimizasyonunu dengelemek için batch boyutu olarak 16 tercih edilmiştir. Öğrenme hızı, AdamW optimizatörünün kararlı ve hızlı yakınsama gösterdiği tipik bir başlangıç düzeyi olup, CosineAnnealingLR scheduler ile dönemsel düşüşler yakalanarak hassas ayar yapılmasına imkân tanımaktadır. Aşırı uyumu sınırlamak için uygulanan “weight decay”, modelin ağırlıklarını hafifçe düzenleyerek genelleme kabiliyetini arttırmıştır. Derin gözetim (deep supervision) katmanlarının ana kayba katkısı için 0,2 değeri seçilerek ara katmanların ürettiği öğrenme sinyalleri ana akışı bozmadan eğitimi hızlandırmıştır. Son olarak, doğrulama IoU’sunda 7 ardışık epoch boyunca iyileşme gözlenmezse eğitimi sonlandıran patience değeri, eğitim sürecinin gereksiz yere uzamasını önlemiştir.

6.2. Test Metrikleri

Test aşamasında modellerin, segmentasyon örtüşmesi IoU ve Dice skoru ile değerlendirilmiştir. Pozitif tespit kalitesini kesinlik ve hassasiyet, tüm piksel sınıflandırmalarının doğruluğunu ise doğruluk metriğiyle değerlendirmiştir. Tüm metriklerin sayısal temelleri bu bölümde formüllerle verilmiştir.

Tablo 6.1. Model eğitiminde kullanılan hiperparametreler, sembolleri, değerleri ve açıklamaları

Hiperparametre	Sembol / Değer	Açıklama
Epoch Sayısı	NUM_EPOCHS = 50	Maksimum eğitim turu sayısı
Batch Boyutu	BATCH_SIZE = 16	GPU bellek ve hız dengesi
Öğrenme Hızı	LEARNING_RATE = 1e-4	AdamW başlangıç öğrenme hızı
Weight Decay	WEIGHT_DECAY = 1e-5	Aşırı öğrenmeyi sınırlamak için regulurizasyon
Deep Supervision Ağırlığı	DEEP_SUPERVISION_WEIGHT = 0.2	Ara katman çıkışlarının kayıp fonksiyonuna katkısı
Erken Durdurma Sabiti	PATIENCE = 7	İyileşme olmazsa durdurma için aranan ardışık döngü sayısı
Optimizatör	AdamW	Parametre güncelleme yöntemi
Scheduler	CosineAnnealingLR	Öğrenme hızını dinamik olarak planlama
Cihaz Seçimi	DEVICE = cuda	GPU/CPU üzerinden model çalıştırma

6.2.1. Kesişim Birleşim Oranı

Kesişim Birleşim Oranı, modelin tahmin ettiği segment ile gerçek segment arasındaki örtüşmenin derecesini ölçmektedir. İki kümenin kesişimindeki piksel sayısının birleşimindeki piksel sayısına bölünmesiyle elde edilmektedir [59]. Segmentasyonun hem genel doğruluğu hem de bölgesel tutarlılığı tek bir değerde özetlenebilmektedir. Denklem (6.1)'de p_i , modelin o pikseli “nodül” olarak sınıflandırma olasılığını (sigmoid çıkışı), g_i ise gerçek maskede o pikselin nodül olup olmadığını (1: nodül, 0: arka plan) göstermektedir.

$$IoU = \frac{|S \cap S'|}{|S \cup S'|} = \frac{\sum_i p_i g_i}{\sum_i p_i + \sum_i g_i - \sum_i p_i g_i} \quad (6.1)$$

6.2.2. Dice Skoru

Dice skoru, IoU'ya benzer şekilde örtüşmeyi değerlendirmektedir ancak özellikle küçük ve düzensiz bölgelerdeki uyumu daha güçlü vurgulamaktadır [59]. Bu sayede ince nodül sınırlarının doğru tespitine dair duyarlılığı arttırmaktadır. Denklem (6.2)'de gözüktüğü üzere, iki katlı kesişim bölgesi toplam bölge büyüklüğüne oranlanmaktadır.

$$\text{Dice} = \frac{2 \sum_i p_i g_i}{\sum_i p_i + \sum_i g_i} \quad (6.2)$$

6.2.3. Kesinlik

Kesinlik (Precision), modelin “nodül” dediği piksellerin gerçekte ne kadarının nodül olduğunu göstermektedir. Yanlış pozitifleri (arka planı nodül olarak etiketleme hatalarını) ne ölçüde azalttığımızı ortaya koymaktadır [83]. Denklem (6.3 – 6.5)'de TP (Gerçek Pozitif), doğru şekilde nodül olarak sınıflanan piksel sayısını; FP (Yanlış Pozitif), gerçekte arka plan olan ancak modelin nodül diye etiketlediği piksel sayısını; TN (Gerçek Negatif), doğru şekilde arka plan olarak sınıflanan piksel sayısını; FN (Yanlış Negatif) ise gerçekte nodül olmasına rağmen modelin arka plan olarak sınıfladığı piksel sayısını temsil etmektedir.

$$\text{Kesinlik} = \frac{TP}{TP+FP} \quad (6.3)$$

6.2.4. Duyarlılık

Duyarlılık (Recall), gerçek nodül piksellerinin ne kadarının başarıyla yakalandığını ölçmektedir. Yanlış negatifleri (nodülleri kaçırma hatalarını) ne kadar azaldığı gösterilmektedir [83]. Formülü Denklem (6.3)'te verilmiştir.

$$\text{Duyarlılık} = \frac{TP}{TP+FN} \quad (6.4)$$

6.2.5. Doğruluk

Doğruluk (Accuracy), modelin hem nodül hem de arka plan piksellerini doğru sınıflandırma oranını vermektedir [84] ve Denklem (6.5)'de gösterilmiştir.

$$\text{Doğruluk} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6.5)$$

6.3. Kodlayıcı Sonuçları

Önerilen U-Net tabanlı segmentasyon mimarisinin omurgasını oluşturan farklı kodlayıcı varyantlarının TN3K üzerinde beş katlı çapraz doğrulama sonuçları bu bölümde verilmiştir. Tüm deneylerde, ağırlık optimizasyonu için DiceLoss kullanılmıştır, erken durdurma (patience = 7) ile Dice skoru, IoU, Hassasiyet, Duyarlılık ve Doğruluk değerleri rapor edilmiştir.

6.3.1. Evrişim Tabanlı Kodlayıcılara Ait Sonuçlar

Tiroid nodülü segmentasyonu amacıyla beş farklı Evrişim tabanlı kodlayıcı mimarisi U-Net omurgasıyla birleştirilerek kapsamlı bir karşılaştırmaya tabi tutulmuştur. Deneylerde ResNet-101, DenseNet-121, VGG16, ResNeXt ve EfficientNet varyantları kullanılmıştır. Her bir model, TN3K veri seti üzerinde beş katlı çapraz doğrulama ile eğitilmiş ve test seti üzerinden değerlendirilmiştir. Tablo 6.2’de görüldüğü üzere elde edilen Dice skoru, IoU, Doğruluk, Kesinlik ve Duyarlılık metrikleri her model için standart sapma değerleriyle raporlanmıştır.

Tablo 6.2’ye bakılarak en yüksek Dice skorunu %85,96 ile EfficientNet tabanlı model elde etmiştir. Aynı modelin IoU’su %75,37, doğruluğu %95,73, kesinliği %87,67 ve duyarlılığı %84,33 olarak ölçülmüştür. Bu değerler, EfficientNet’in hem yanlış pozitifleri hem de yanlış negatifleri düşük tutarak klinik açıdan son derece güvenilir bir segmentasyon sağladığını göstermektedir.

EfficientNet’i takiben gelen DenseNet-121, Dice skorunu %85,53 ve IoU’su %74,72 ile güçlü bir performans sergilemiştir. Doğruluk %95,59, kesinlik %87,09 ve duyarlılık %84,06 değerleri, katmanlar arası sıkı bağlantı mekanizmasının bilgi akışını koruyarak küçük nodül yapılarında bile yüksek başarı sunduğunu kanıtlamıştır. ResNeXt tabanlı kodlayıcı ise Dice skorunu %84,94 ve IoU’su %73,82 ile üçüncü sırada yer almış; doğruluk %95,45, kesinlik %87,10 ve %82,92 duyarlılık göstermektedir.

ResNet-101 + U-Net kombinasyonu, Dice skoru %84,81 ve IoU’su %73,62 değerleriyle makul bir segmentasyon yeteneği sunarken, doğruluk %95,39, kesinlik %86,54 ve duyarlılık %83,15 ile diğer modellere yakın bir stabilite sergilemiştir. Buna karşın,

VGG16 tabanlı mimari en düşük performansı kaydetmiştir. Dice skoru %73,15, IoU %57,75, doğruluk %91,77, kesinlik %74,06 ve duyarlılık %72,34 değerleri, daha sık ve geniş kernel yapısının ince nodül sınırlarını ayırt etmede yetersiz kaldığını ortaya koymuştur. Tüm kodlayıcılar benzer yüksek doğruluk değerleri (%91,77–%95,73) sunsa da esas segmentasyon kalitesini belirleyen Dice skoru ve IoU değerlerindeki yaklaşık 10–20 puanlık farklar, mimari tercihlerinin klinik segmentasyonda kritik öneme sahip olduğunu vurgulamaktadır.

Tablo 6.2. TN3K veri seti üzerinde beş farklı evrişim tabanlı kodlayıcının performans metrikleri

Kodlayıcı	IoU (%)	Dice Skoru (%)	Kesinlik (%)	Duyarlılık (%)	Doğruluk (%)
ResNet-101	73,62 ± 0,35	84,81 ± 0,23	86,54 ± 0,58	83,15 ± 0,70	95,39 ± 0,06
DenseNet-121	74,72 ± 0,40	85,53 ± 0,26	87,09 ± 1,33	84,06 ± 1,19	95,59 ± 0,10
VGG16	57,75 ± 3,59	73,15 ± 2,94	74,06 ± 3,92	72,34 ± 2,78	91,77 ± 0,98
ResNeXt	73,82 ± 0,40	84,94 ± 0,27	87,10 ± 1,37	82,92 ± 1,21	95,45 ± 0,11
EfficientNet	75,37 ± 0,46	85,96 ± 0,30	87,67 ± 0,84	84,33 ± 1,18	95,73 ± 0,06

6.3.2. Transförmator Tabanlı Kodlayıcılara Ait Sonuçlar

Dikkat mekanizmalarının küresel bağlam bilgisini nasıl etkin şekilde yakalayabildiğini göstermek amacıyla Swin Transformer (SwinT), Vision Transformer (ViT) ve CoAtNet-2 kodlayıcılarının tiroid nodülü segmentasyonundaki performansları ele alınmıştır. Farklı kodlayıcıların ince yapı ayırt etme, sınır belirleme ve genel sınıflama başarıları arasındaki farklar sistematik biçimde karşılaştırılmıştır ve Tablo 6.3’te gösterilmiştir.

Öncelikle segmentasyon performans ölçümünün en yaygın ölçütlerinden biri olan Dice skoruna bakıldığında SwinT'in %77,07, ViT'in %83,81 ve CoAtNet-2'nin %82,77 değerleri elde ettiği görülmektedir. ViT, Dice skoru bakımından bu üç model arasında en yüksek değeri göstermiştir. Ardından IoU metriği incelendiğinde SwinT'in %66,30, ViT'in %72,37 ve CoAtNet-2'nin %70,70 gibi oranlarda olduğu görülmektedir. IoU'da da yine ViT en yüksek oranı yakalarken CoAtNet-2 ikinci ve SwinT üçüncü sırada yer almaktadır.

Tablo 6.3. Transformatör tabanlı kodlayıcıların TN3K veri seti üzerindeki test performansları

Kodlayıcı	IoU (%)	Dice Skoru (%)	Kesinlik (%)	Duyarlılık (%)	Doğruluk (%)
SwinT	66,30 ± 1,49	77,07 ± 1,45	85,21 ± 1,01	80,75 ± 2,33	94,83 ± 0,19
ViT	72,37 ± 1,02	83,81 ± 0,92	86,72 ± 0,97	82,23 ± 1,78	95,3 ± 0,26
CoAtNet-2	70,70 ± 3,99	82,77 ± 2,77	87,48 ± 1,27	78,75 ± 5,22	94,96 ± 0,64

Genel sınıflama başarısını tanımlayan Doğruluk metriğinde SwinT %94,83, CoAtNet-2 %94,96 ve ViT %95,30 değerlerine ulaşmıştır. Doğruluk puanları birbirine yakın görünse de, ViT'in en yüksek doğruluk oranı ile hafif bir üstünlük ile sağlamıştır. SwinT ise %94'ün hemen üzerinde kalan değeriyle temel sınıflandırma görevini başarıyla tamamladığı gözlemlenebilir. Kesinlik metriklerine bakıldığında, SwinT %85,21, ViT %86,72 ve CoAtNet-2 %87,48 oranlarında sonuç vermiştir. Kesinlik değerleri, yanlış pozitif sınıflandırmaların oranı üzerine odaklanır; bu bağlamda üç model arasında minimal farklılıklar vardır ancak CoAtNet-2'nin %87,48 ile en yüksek kesinlik oranını elde ettiği bulunmuştur.

Duyarlılık skorlarında ise SwinT %80,75, ViT %82,23 ve CoAtNet-2 %78,75 oranlarındadır. Duyarlılık metrikleri doğru pozitif sınıflandırmaların oranını yansıtmaktadır, bu açıdan ViT %82,23 ile en yüksek duyarlılık skorunu elde ettiği, SwinT ikinci sırada yer aldığı ve CoAtNet-2'nin en düşük duyarlılık oranına sahip olduğu çıkarılabilir.

6.4. Güncel Modellerle Karşılaştırma ve Analiz

6.4.1. TN3K Veri Kümesinde Analizler

TN3K veri seti üzerinde gerçekleştirilen karşılaştırmalı analizde yer alan klasik evrişim tabanlı segmentasyon yöntemleri, güncel transformatör tabanlı yaklaşımlar ve önerilen EffiViT3+ mimarisinin performansları ayrıntılı olarak ele alınmıştır. Tablo 6.4'te özetlenen beş temel metrik sonuçları verilmiştir.

Önerilen EffiViT3+ modeli IoU metriğinde %77,43, Dice skorunda %87,16, Kesinlik oranında %89,10, Duyarlılık oranında %86,18 ve Doğruluk değerinde %96,66 elde ederek tüm rakiplerinden açık ara öne çıkmıştır. Bu mimari, EfficientNet tabanlı çok katmanlı evrişim blokları ile Vision Transformer'ın küresel dikkat mekanizmalarını birleştiren hibrit tasarımı sayesinde, nodül sınırlarını çok yüksek hassasiyetle tanımlayabilmekte ve hem yanlış pozitif hem de yanlış negatif hataları minimuma indirebilmektedir. EffiViT3+'ın %89'u aşan Kesinlik oranı, arka plan piksellerinin neredeyse tamamını doğru şekilde arka plan olarak sınıflandırdığını, %86'nın üzerindeki Duyarlılık oranı ise gerçek nodül piksellerinin büyük çoğunluğunu başarıyla yakaladığını göstermektedir. %96,66'lık Doğruluk skoru da bu üstün segmentasyon yeteneklerinin genel sınıflama başarısıyla tutarlı biçimde birleştiğini teyit etmektedir.

İkinci sırada yer alan geleneksel U-Net yapısının çok katmanlı atlama bağlantılarıyla geliştirilmiş versiyonu olan UNet3+, %74,65'e IoU değerine ulaşması modelin segmentasyon örtüşmesini belirgin ölçüde iyileştirdiğini göstermiştir. Dice skoru %85,23'e çıkartılarak UNet3+'ın özellikle ince sınır yapılarında bile yüksek örtüşme sağladığı ortaya konmuştur. Kesinlik değeri %84,76 ve Duyarlılık değeri %86,84 ile U-Net3+, yanlış pozitif ve yanlış negatif oranlarını dengede tutmuştur. %96,00'lık Doğruluk skoru ise U-Net3+'ın CNN tabanlı modeller içinde en yüksek genel sınıflama başarısını sunduğunu işaret etmektedir.

Transformatör tabanlı modeller arasında üçüncü sırayı alan SegFormer, IoU skoru olarak %70,28 ve Dice skorunda %82,34 ile U-Net ve DeepLabV3+ gibi bazı evrişim tabanlı mimarilerini geride bırakmıştır. SegFormer'in çoklu token grid stratejisi ve hafif ölçeklenebilir dikkat blokları, modelin hem geniş ölçekli bağlam bilgisini hem de ayrıntılı sınır yapısını dengeli biçimde öğrenebilmesine olanak tanımaktadır. %82,10'luk Kesinlik ve %83,94'lük Duyarlılık oranları, SegFormer'in yanlış pozitifleri makul düzeyde sınırlarken

gerçek nodül piksellerinin büyük çoğunluğunu da başarıyla yakaladığını göstermektedir. %95,29'luk Doğruluk değeri ise SegFormer'in genel sınıflama başarısının evrişim tabanlı modellerle yakından rekabet ettiğini ortaya koymaktadır.

Dördüncü model olan FPN, piramit yapısıyla çok ölçekli özellik çıkarımını etkin biçimde gerçekleştirerek IoU değerini %71,55 ve Dice skorunu %83,25 seviyesine taşımıştır. FPN'in Kesinlik oranı %84,80 ve Duyarlılık oranı %82,93 olarak ölçülmüştür. Doğruluk değeri %95,62'ye ulaşmıştır. Bu sonuçlar, FPN'in çoklu ölçeklerdeki semantik bilgiyi bütünleştirme kabiliyetinin, özellikle orta ölçekli nodül yapılarının tanınmasında avantaj sağladığını göstermektedir.

Beşinci sıradaki DeepLabV3+, ASPP ünitesi sayesinde IoU'yu %70,09, Dice skorunu %82,26'ya yükseltmiş; Kesinlik %83,17, Duyarlılık %82,40 ve Doğruluk %95,31 değerleriyle standart CNN mimarileri arasında ortalama üzerinde bir performans sergilemiştir. DeepLabV3+'ın başarısı, farklı dilatasyon oranlarıyla eşzamanlı çok ölçekli bağlam bilgisi toplamasından kaynaklanmakta, ancak UNet3+ ve FPN seviyesine kadar yüksellemediği görülmektedir.

Altıncı sırada bulunan orijinal U-Net, IoU'da %70,83, Dice skorunda %82,63; Kesinlikte %84,30, Duyarlılıkta %82,43 ve Doğrulukta %95,49 değerleriyle temel segmentasyon yeteneklerini ortaya koyar. U-Net, çoklu atlama bağlantılarının olmamasına rağmen güçlü bir başlangıç performansı sunmakta; ancak gelişmiş varyantlarının gerisinde kaldığı açıktır.

Yedinci sırada yer alan U-Net++, ek atlama yollarıyla daha derin bağlantılar kurmaya çalışsa da IoU'da %69,67 ve Dice skorunda %81,83 ile U-Net'in hemen arkasında kalmaktadır. U-Net++'ın Kesinlik %83,27, Duyarlılık %81,77 ve Doğruluk %95,25 değerleri, modelin mimari karmaşıklığının tek başına segmentasyon kalitesini iyileştirmeye yetmediğini, eğitim stratejilerinin ve veri uyarlamasının da kritik rol oynadığını göstermektedir.

Son olarak TransFuse, hibrit bir CNN-transformer yaklaşımı olmasına rağmen IoU'da %64,02 ve Dice skorunda %77,72 ile tüm modeller arasında en düşük örtüşme skorlarını sergilemiştir. Kesinlik %81,56, Duyarlılık %75,59 ve Doğruluk %94,32 değerleri, TransFuse'un hem yanlış pozitif hem de yanlış negatif hataları yüksek oranda ürettiğini; modelin global-context ve lokal-detail entegrasyonunda zorluk yaşadığını ortaya koymaktadır.

Tablo 6.4. Güncel modeller ile önerilen modelin TN3K veri setinde performans karşılaştırmaları

Model		IoU (%)	Dice Skoru (%)	Kesinlik (%)	Duyarlılık (%)	Doğruluk (%)
Evrişim Tabanlı Segmentasyon Modelleri	U-Net	70,83 ± 0,0082	82,63 ± 0,0058	84,30 ± 0,0059	82,43 ± 0,0065	95,49 ± 0,0015
	Unet++	69,67 ± 0,0078	81,83 ± 0,0055	83,27 ± 0,0083	81,77 ± 0,0028	95,25 ± 0,0016
	UNet3+	74,65 ± 0,0273	85,23 ± 0,0200	84,76 ± 0,0470	86,84 ± 0,0139	96,00 ± 0,0071
	FPN	71,55 ± 0,0054	83,25 ± 0,0038	84,80 ± 0,0095	82,93 ± 0,0071	95,62 ± 0,0012
	DeepLabv3+	70,09 ± 0,0043	82,26 ± 0,0030	83,17 ± 0,0039	82,40 ± 0,0046	95,31 ± 0,0008
Transformatör Tabanlı Segmentasyon Modelleri	TransFuse	64,02 ± 0,0054	77,72 ± 0,0044	81,56 ± 0,0068	75,59 ± 0,0065	94,32 ± 0,0012
	SegFormer	70,28 ± 0,0056	82,34 ± 0,0038	82,10 ± 0,0070	83,94 ± 0,0044	95,29 ± 0,0012
Önerilen Model	EffViT3+	77,43 ± 0,0052	87,16 ± 0,0036	89,10 ± 0,0067	86,18 ± 0,0086	96,66 ± 0,0006

Genel sıralamaya baktığımızda, IoU ve Dice skor metrikleri segmentasyon örtüşmesindeki kalitenin birinci derecede belirleyicisi olarak ön plana çıkmaktadır. EffViT3+ ve U-Net3+ bu alanda açık farkla liderliği paylaşmaktadır. Kesinlik ve Duyarlılık oranları, yanlış pozitif ve yanlış negatif hataların dağılımını gösterirken, EffViT3+'ın her iki hatayı da dengeli biçimde minimize ettiği, SegFormer'in ise

transformatör yaklaşımları içinde en yüksek örtüşmeyi sağlarken false negative hataları kontrol etmekte başarılı olduğu anlaşılmaktadır. Doğruluk metrikleri tüm modeller arasında yüksek seyretmekle birlikte, segmentasyon kalitesini farklılaştıran unsur Dice skoru ve IoU değeridir. Bu veriler hem temel CNN modellerinin hem de modern transformatör yaklaşımlarının sınırlarını ve güçlü yanlarını ortaya koyarken, hibrit EffiViT3+ mimarisinin her iki paradigmanın en iyi yönlerini bir araya getirerek segmentasyon performansını zirveye taşıdığını açıkça göstermektedir.

6.4.2. DDTI Veri Kümesinde Analizler

DDTI veri seti, sınırlı sayıda örnek içermesi nedeniyle modelimizin az veri koşullarındaki genellenebilirliğini ve kararlılığını test etmek amacıyla ikinci bir sınav ortamı olarak seçilmiştir. Bu sayede, TN3K’da elde edilen yüksek performansın, veri miktarının belirgin derecede azaldığı bir senaryoda ne ölçüde sürdürülebilir olduğunu değerlendirilmiştir. Bu değerlendirmeye ait test sonuçları Tablo 6.5’te detaylı olarak sunulmuştur.

Tablo 6.5’te yer alan sonuçlara bakılarak önerilen EffiViT3+ modelinin tüm karşılaştırılan mimariler arasında öne çıktığı görülmektedir. EffiViT3+ için test aşamasında hesaplanan IoU değeri %72,38, Dice skoru %83,94, Kesinlik %85,23, Duyarlılık %82,75 ve Doğruluk %97,56 olarak raporlanmıştır.

Evrişim tabanlı modeller grubunda ikinci sırayı alan U-Net3+, %67,95 IoU ve %80,69 Dice skoru sonuçlarıyla EffiViT3+ modelini izlemektedir. U-Net3+ için ölçülen %76,83 Kesinlik ve %86,34 Duyarlılık değerleri, bu modelin de yüksek düzeyde hatasız pozitif tespit ve güçlü gerçek pozitif tespit yeteneği sergilediğini ortaya koymaktadır. Doğruluk bakımından U-Net3+, %96,87 değeriyle ikinci sırada yer almakta, yalnızca EffiViT3+’ın gerisinde kalmaktadır.

Üçüncü sırada yer alan orijinal U-Net için test sonuçları %66,44 IoU, %79,52 Dice skoru, %75,83 Kesinlik ve %96,60 Doğruluk olarak ölçülmüştür. Dördüncü sırayı alan FPN için test sonuçları %64,54 IoU, %78,32 Dice skoru, %80,90 Kesinlik, %76,55 Duyarlılık ve %96,79 Doğruluk şeklindedir. FPN, çok ölçekli özellik çıkarımı ile IoU ve Dice skorlarında U-Net ile U-Net3+’ın arasında bir konumda yer almakta olup doğruluk ve kesinlik değerleriyle evrişim grubu içindeki rekabette orta seviye bir başarı göstermektedir.

U-Net++ için test sonuçları; Ortalama IoU %65,25, Ortalama Dice skoru %78,20, Ortalama Kesinlik %72,99, Ortalama Duyarlılık %87,27 ve Ortalama Doğruluk %96,04 olarak rapor edilmiştir. U-Net++'ın IoU ve Dice skorları, U-Net'in gerisinde kalmakla birlikte Duyarlılık metriğinde %87,27 değeriyle bu listedeki tüm modeller arasında en yüksek recall oranını sunmaktadır. Bu durum, U-Net++'ın gerçek pozitif tespitindeki başarısının öne çıktığını göstermektedir.

Transformatör tabanlı modeller grubunda, SegFormer ikinci sıraya yerleşmiştir. Ölçülen IoU %59,49, Dice skoru %73,42, Kesinlik %69,12, Duyarlılık %79,30 ve Doğruluk %95,66'dır. SegFormer, IoU değeri ve Dice skoruyla evrişim tabanlı grup içindeki U-Net ve U-Net3+ modellerinin hemen ardından gelerek transformatör yaklaşımlarının da segmentasyonda önemli bir örtüşme başarısı elde edebildiğini ortaya koymaktadır. SegFormer'ın Doğruluk değeri de CNN modellerinin ortalamasına yakın düzeyde kalırken Duyarlılık metriğindeki %79,30 değeri doğru tespit kabiliyetinin makul seviyede olduğunu göstermektedir.

TransFuse için test sonuçları dikkate alındığında, IoU değeri %64,02, Dice skoru %77,72, Kesinlik %81,56, Duyarlılık %75,59 ve Doğruluk %94,32 olarak ölçülmüştür. TransFuse, hibrit bir tasarıma sahip olmakla birlikte temel evrişim tabanlı modellere kıyasla IoU ve Dice skorlarında daha geride kalmaktadır. Bununla birlikte Kesinlik değerinin %81,56 düzeyinde olması, modelin yanlış pozitif oranını sınırlama yeteneğine sahip olduğunu göstermektedir.

DeepLabV3+ için test sonuçları ise IoU %40,31, Dice skoru %57,27, Kesinlik %43,74, Duyarlılık %84,71 ve Doğruluk %90,35 şeklinde elde edilmiştir. DeepLabV3+'ın IoU ve Dice skorlarındaki düşük değerler performansı en düşük model olduğunu göstermiştir.

Genel sıralamaya bakıldığında, test IoU ve Dice skorlarının EffViT3+ ve U-Net3+ modellerinde oldukça yüksek olduğu ortaya çıkmıştır. Kesinlik metriğinde EffViT3+'ın %85,23 oranı tüm modeller arasında birinci sıradadır. Duyarlılık metriğinde U-Net++'ın %87,27 değeri en yüksek pozitif tespit oranını sergilemektedir. Doğruluk metriğinde ise EffViT3+ %97,56 ile tablonun en üstünde yer almıştır.

Tablo 6.5. DDTI veri seti üzerinde farklı segmentasyon modellerinin test metrikleri

Model		IoU (%)	Dice Skoru (%)	Kesinlik (%)	Duyarlılık (%)	Doğruluk (%)
Evrişim Tabanlı Segmentasyon Modelleri	U-Net	66,44 ± 6,03	79,52 ± 4,72	96,60 ± 1,10	75,83 ± 8,42	96,60 ± 1,10
	U-Net++	65,25 ± 9,14	78,20 ± 7,78	96,04 ± 2,26	72,99 ± 13,00	87,27 ± 3,49
	U-Net3+	67,95 ± 4,14	80,69 ± 3,07	96,87 ± 0,70	76,83 ± 5,92	86,34 ± 1,32
	FPN	64,54 ± 1,53	78,32 ± 1,13	96,79 ± 0,15	80,90 ± 1,15	76,55 ± 1,77
	DeepLabv3+	40,31 ± 1,72	57,27 ± 1,77	90,35 ± 0,78	43,74 ± 2,14	84,71 ± 1,29
Transformatör Tabanlı Segmentasyon Modelleri	TransFuse	44,10 ± 1,35	61,00 ± 1,30	92,34 ± 0,55	50,43 ± 2,36	78,33 ± 1,90
	SegFormer	59,49 ± 1,88	73,42 ± 0,45	95,66 ± 0,12	69,12 ± 1,22	79,30 ± 1,13
Önerilen Model	EffiViT3+	72,38 ± 0,60	83,94 ± 0,40	97,56 ± 0,05	85,23 ± 0,66	82,75 ± 1,22

6.5. Kayıp Fonksiyonu Sonuçları

Model eğitimi ve değerlendirmelerinde kullanılan kayıp fonksiyonlarının, özellikle tiroid nodülü segmentasyonunda sınır detaylarını ve bölgesel tutarlılığı yeterince optimize edememektedir. Literatürde yaygın olan Dice kaybı, Lovász kaybı veya bunların

kombinasyonlarının hem küçük lezyon kenarlarında hem de geniş doku bölgelerinde ideal sonuçlar vermek bakımından sınırlı kalmaktadır. Bu nedenle yeni, hibrit bir kayıp fonksiyonu geliştirilmiştir. Önerilen kayıp fonksiyonu, klasik segmentasyon kayıplarındaki “kaçırılan ince detay” ve “bölgesel tutarsızlık” sorunlarına, derin gözetim, sınır tabanlı ve HD95 bileşenlerini ekleyerek çözüm sunmayı amaçlamaktadır. Bu bileşenlerin bir arada, tek ve farklı kombinasyonlarda kullanımının etkisi incelenmiş olup validasyon ve test sonuçlarına ait IoU değerleri ve Dice skorları Tablo 6.6’da sunulmuştur.

İlk kombinasyon olarak yalnızca Dice kaybının kullanıldığı senaryoda, validasyonda IoU %74,35 ve Dice skoru %83,10 olarak ölçülmüş, test aşamasında da IoU %74,15 ve Dice skoru olarak %83,10 elde edilmiştir. İkinci olarak Lovasz ve Dice kayıplarının eş ağırlıkla birleştirildiği kombinasyonda, validasyon IoU %75,12, Dice skoru %83,70 düzeylerine ulaşırken test IoU %74,54 ve Dice skoru %83,49 değerlerini görülmüştür. Lovasz bileşeninin eklenmesi, validasyonda IoU’da yaklaşık 0,8 puanlık artış sağlasa da test skorlarında çok küçük bir yükselişi karşılamıştır. Üçüncü kayıp kombinasyonunda, Lovasz, Dice ve sınır kayıp bileşenleriyle validasyonda IoU %78,18, Dice skoru %83,65, testte ise IoU %74,43, Dice skoru %83,45 ölçülmüştür. Sınır kaybının eklenmesi, validasyonda IoU değerini %78’in üzerine taşırken test aşamasında IoU ve Dice skoru önceki kombinasyonlarla benzer düzeyde korunmuştur. Dördüncü kombinasyon Lovasz olarak uygulandığında validasyonda IoU %78,36 ve Dice skoru %87,71’e yükselmiş, test aşamasında ise IoU %74,42 ve Dice skoru %83,44 olarak kaydedilmiştir. HD95 bileşeninin validasyonda Dice skorunda belirgin artışa yol açtığı, test IoU–Dice değerlerinin ise ara bir çizgide kaldığı görülmektedir. Beşinci kayıp bileşimi olan Lovasz, Dice, FTV ve sınır kayıplarının kullanıldığı senaryoda validasyonda IoU %76,91 ve Dice skoru %86,76 elde edilirken test aşamasında ise IoU %76,27 ve Dice skoru %86,41 değerleri elde edilmiştir. Burada FTV ve sınır terimlerinin eklenmesi, test skorlarında yaklaşık 2 puana varan artışlara imkan sağlamıştır. Son ve en kapsamlı kombinasyon olan Lovasz, Dice, FTV, Boundary ve HD95 kayıpları uygulandığında ise validasyonda IoU %78,14, Dice skoru %87,57, testte IoU %77,43 ve Dice skoru %87,16 değerlerine ulaşılmıştır. Bu beş bileşenli yapı, test IoU ve Dice skorunda en yüksek sonuçları vererek hem validasyon hem test aşamalarında en dengeli genelleme performansını sunmuştur.

Tablo 6.6. Farklı kayıp fonksiyonlarının metriklerinin karşılaştırılması

Kayıp Kombinasyonu	Validasyon IoU (%)	Validasyon Dice (%)	Test IoU (%)	Test Dice (%)
Dice Tek (%100)	74,35 ± 0,76	83,10 ± 0,68	74,15 ± 0,37	83,10 ± 0,32
Lovasz + Dice (%50 / %50)	75,12 ± 0,41	83,70 ± 0,50	74,54 ± 0,12	83,49 ± 0,11
Lovasz + Dice + Sınır (%40 / %40 / %20)	78,18 ± 0,92	83,65 ± 0,54	74,43 ± 0,32	83,45 ± 0,26
Lovasz + Dice + HD95 (%40 / %40 / %20)	78,36 ± 0,80	87,71 ± 0,57	74,42 ± 0,33	83,44 ± 0,26
Lovasz + Dice + FTV + Sınır (%30 / %25 / %25 / %20)	76,91 ± 1,49	86,76 ± 1,01	76,27 ± 0,46	86,41 ± 0,31
Lovasz + Dice + FTV + Sınır + HD95 (%25 / %25 / %20 / %20 / %10)	78,14 ± 0,76	87,57 ± 0,53	77,43 ± 0,52	87,16 ± 0,36

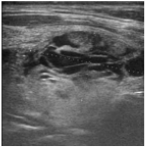
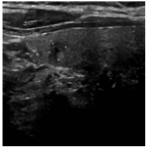
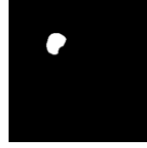
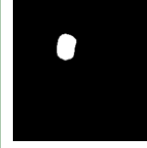
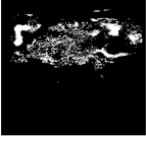
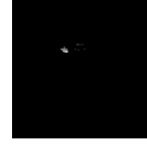
6.6. Nitel Sonuçlar

Çalışmada TN3K ve DDTI olmak üzere iki farklı veri seti kullanılmış olup önerilen model ve diğer modellerin test kümelerinde nitel sonuçları elde edilmiştir. Şekil 7.1’de TN3K test kümesine ait 5 farklı ultrason görüntüsü ve modellerin eğitim veya doğrulama için hiç kullanmadığı ultrason görüntülerinden tiroid nodül segmentasyonu sunulmuştur. Şekil 7.1’de TN3K veri kümesine ait tiroid ultrason görüntüsü, uzmanlar tarafından oluşturulan referans maske ve modellerin tahmin ettiği maskeler yer almaktadır. TN3K veri kümesinden seçilen örnekler üzerinde yapılan görsel karşılaştırmalar, EffViT3+ modelinin farklı boyut ve morfolojiye sahip lezyonlarda yüksek doğruluk ve tutarlılık sağladığını

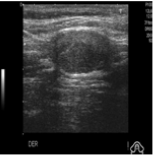
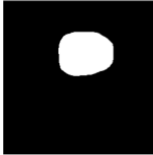
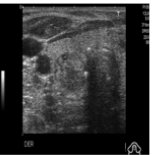
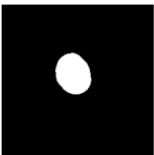
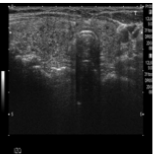
göstermektedir. İlk örnekte, geniş ve yatay uzanan lezyonun segmentasyonu sırasında EffiViT3+ referans maskenin sınırlarını biçimsel olarak korumuş, kenar düzgünlüğü ve iç doluluk açısından hatasız bir çıktı üretmiştir. İkinci örnekte yer alan küçük lezyonda model, boyut ve konum bakımından referansa en yakın tahmini sunarken, diğer modeller ya alanı daraltmış ya da gereksiz genişletmiştir. Üçüncü örnekteki geniş yapılı lezyonda EffiViT3+, yatay ve dikey oranları koruyarak dengeli bir segmentasyon sağlamış, buna karşın FPN ve TransFuse gibi modellerde kenar taşmaları görülmüştür. Dördüncü örnekte, çok küçük hedef yapının doğru konum ve boyutla segmentasyonu EffiViT3+ tarafından başarıyla gerçekleştirilmiş, diğer modellerde ise gürültü, parçalanma veya konum hataları gözlenmiştir. Beşinci örnekte orta boyutlu lezyonun segmentasyonu sırasında EffiViT3+, referans konturuna uyum sağlayarak yapısal bütünlüğü korumuş, TransFuse gibi bazı modellerde pikselleşme ve kenar bozulmaları ortaya çıkmıştır. Bu bulgular, EffiViT3+ modelinin hem küçük hem de büyük lezyonlarda kenar hassasiyeti, iç homojenlik ve yapısal bütünlük bakımından referans maskeye yakın sonuçlar üretilmiştir.

DDTI veri kümesinden seçilen örnekler üzerinde yapılan görsel inceleme Şekil 7.2’de sunulmuştur. EffiViT3+ modelinin düşük kontrast, doku homojenliği ve artefakt yoğunluğu gibi ultrason görüntülerinde sık karşılaşılan zorluklara rağmen yüksek doğrulukta segmentasyon gerçekleştirdiğini göstermektedir. İlk örnekte yer alan geniş ve düzensiz sınırlara sahip lezyonda EffiViT3+, referans maskenin kontur bütünlüğünü eksiksiz şekilde korumuş, kenar pürüzlenmelerinden kaçınarak iç bölgede homojen bir doluluk sağlamıştır. UNet++ ve TransFuse gibi rakip modellerde ise sınır bozulmaları ve kısmi taşmalar gözlenmiştir. İkinci örnekteki küçük lezyonun segmentasyonunda EffiViT3+, boyut ve konum bakımından referansa en yakın tahmini üretmiş; buna karşın UNet3+ ve SegFormer alanı daraltma, UNet ve DeepLabv3 ise gereksiz genişletme eğiliminde olmuştur. Üçüncü örnekte orta boyutlu bir yapı yer almakta olup EffiViT3+, hem yatay hem dikey oranları koruyarak dengeli bir segmentasyon sağlamış, FPN ve DeepLabv3’te ise hedef bölgenin tam kapsanmaması sorunları tespit edilmiştir. Dördüncü örnekte çok küçük ve düşük kontrastlı bir lezyonda EffiViT3+, yalın, gürültüsüz ve doğru konumda bir maske üretmiştir; buna karşın UNet++ ve TransFuse modellerinde parçalanma ve yanlış pozitif alanlar dikkat çekmiştir. Son örnekte ise karmaşık şekilli geniş bir yapıda EffiViT3+, referans konturlarını doğru şekilde takip ederek yapısal bütünlüğü korumuş ve yüksek görsel kalite sağlamıştır. Bu bulgular, EffiViT3+ modelinin hem küçük hem de büyük lezyonlarda biçimsel uyumu,

kenar hassasiyetini ve iç homojenliği koruyabildiğini; ayrıca diğer yaygın modellerde sıkça görülen taşma, eksik kapsama veya pürüzlenme gibi hatalardan kaçındığını ortaya koymaktadır. Sonuç olarak, DDTI veri kümesinde elde edilen görsel değerlendirme, EffiViT3+'ın klinik açıdan güvenilir, dengeli ve istikrarlı bir segmentasyon modeli olduğunu güçlü biçimde desteklemektedir.

Orijinal Görüntü	Referans Maske	UNet Tahmini	Unet++ Tahmini	UNet3+ Tahmini	FPN Tahmini	DeepLabv3 Tahmini	TransFuse Tahmini	SegFormer Tahmini	EffViT3+ Tahmini
									
									
									
									
									

Şekil 6.1. TN3K test kümesi ile elde edilen nitel sonuçlar

Orijinal Görüntü	Referans Maske	UNet Tahmini	Unet++ Tahmini	UNet3+ Tahmini	FPN Tahmini	DeepLabv3 Tahmini	TransFuse Tahmini	SegFormer Tahmini	EffViT3+ Tahmini
									
									
									
									
									

Şekil 6.2. DDTI test kümesi ile elde edilen nitel sonuçlar

7. TARTIŞMA ve SONUÇ

Bu bölümde, deneysel bulguların anlamı ve modelimizin performansındaki kilit noktalar ayrıntılı olarak yorumlanacaktır. Ardından bu sonuçlar güncel literatürdeki benzer yaklaşımlarla karşılaştırılacaktır. Son olarak çalışmanın devamında yapılabilecek geliştirmeler önerilecektir ve sonuç yer almaktadır.

7.1. Bulguların Yorumu

Tiroid nodüllerinin ultrasonografi görüntülerinde etkili bir şekilde segmentasyonu için tasarlanan segmentasyon modelinin geliştirilme aşamasında en önemli unsurlardan biri doğru kodlayıcı mimarisinin seçimidir. Kodlayıcı mimarisi, görüntüdeki temel özellikleri çıkarmak ve daha sonra kod çözücü katmanlarında kullanılmak üzere yüksek düzeyli bilgileri elde etmek amacıyla kullanılmaktadır. Bu nedenle farklı kodlayıcı mimarileri, nodül segmentasyonu başarısında büyük rol oynayabilmektedir. Bu çalışmada kodlayıcı seçiminde hem evrişim tabanlı mimariler hem de transformatör tabanlı mimariler incelenmiştir. Evrişim tabanlı mimariler arasında ResNet-101, DenseNet-121, VGG16, ResNeXt ve EfficientNet yer alırken, transformatör tabanlı kodlayıcılar olarak Swin Transformer (SwinT), Vision Transformer (ViT) ve CoAtNet-2 modelleri kullanılmıştır.

Evrişim tabanlı kodlayıcıların karşılaştırıldığı deneyler sonucunda, EfficientNet en yüksek performansı elde eden mimari olarak öne çıkmıştır. TN3K veri seti üzerinde yapılan testlerde EfficientNet, %85,96 Dice skoru ve %75,37 IoU skoru elde ederek diğer evrişim tabanlı modellerden belirgin biçimde ayrılmıştır. Özellikle küçük nodüllerde sınır ayrıntılarını yakalamadaki başarısıyla klinik uygulamalarda güvenilir bir segmentasyon sağlamıştır. EfficientNet'i takip eden DenseNet-121 ise %85,53 Dice skoru ve %74,72 IoU skorlarıyla güçlü bir alternatif olarak ortaya çıkmaktadır. DenseNet mimarisindeki katmanlar arasındaki sıkı bağlantılar, küçük nodül yapılarında bile tutarlı bir bilgi akışı sağlanmış ve segmentasyon kalitesini arttırmıştır. Diğer yandan, VGG16 modelinin en düşük performansı göstermesi, sıkı ve geniş çekirdekli yapısının küçük ayrıntıları yakalamakta zorlandığını ortaya koymuştur. Bu durum, Dice skorunun %73,15 ve IoU skorunun %57,75 gibi düşük seviyelerde olmasıyla gözlemlenmiştir.

Transformatör tabanlı kodlayıcıların değerlendirilmesinde ViT, SwinT ve CoAtNet-2 kıyaslanmıştır. Bu modeller arasında ViT en yüksek segmentasyon performansını elde etmiştir. ViT modelinin test sonuçlarında Dice skorunun %83,81, IoU'nun ise %72,37 seviyesine ulaşması, transformatör modellerinin global dikkat mekanizmasının etkili olduğunu ve segmentasyonda yüksek doğruluk sağladığını göstermiştir. CoAtNet-2 ise %82,77 Dice skoru ve %70,70 IoU skoru ile ikinci sırada yer almıştır. CoAtNet-2 modelinin hibrit yapısı sayesinde yerel ve global özellikleri dengeleme yeteneği, sınıflandırma ve segmentasyonda iyi bir genel performans göstermesine etkili olmuştur. SwinT modelinin sonuçları ise Dice skoru %77,07 ve IoU %66,30 ile diğer iki Transformer tabanlı modelin gerisinde kalmıştır. Bu durum, Swin Transformer'ın pencereli dikkat mekanizmasının daha küçük ayrıntılarda yeterince etkili olmadığını göstermektedir. Elde edilen bulgular kodlayıcı seçiminde EfficientNet ve ViT birleşiminin üstün performans sağlayabileceğini göstermiştir.

TN3K veri kümesi üzerinde geliştirilen EffiViT3+ modeli, hibrit yaklaşımın konvansiyonel ve transformatör tabanlı modellerden daha üstün sonuçlar sunduğunu açıkça ortaya koymuştur. EffiViT3+'ın %87,16 Dice skoru, %77,43 IoU ve %96,66 doğruluk oranları, EfficientNet ve ViT'in güçlü yönlerinin bir araya gelmesinin nodül sınırlarını doğru şekilde tanımlama ve genel doğruluk açısından belirgin iyileşme sağladığını kanıtlamıştır. Bu model, yüksek hassasiyet ve duyarlılık değerleriyle klinik açıdan kritik olan yanlış tanı oranlarını da minimize etmektedir. EffiViT3+'ı takip eden UNet3+ modeli de yüksek performans göstermiştir. Özellikle derin atlama bağlantıları ile küçük ayrıntıların korunmasında avantaj sağlayan UNet3+'ın IoU (%74,65) ve Dice (%85,23) skorları, klasik evrimsel modellere kıyasla belirgin üstünlüğünü ortaya koymaktadır. Bu durum, atlama bağlantılarının küçük ölçekli segmentasyon problemlerinde kritik öneme sahip olduğunu vurgulamaktadır.

DDTI veri kümesinde yapılan testler, EffiViT3+ modelinin sınırlı veri koşullarında dahi yüksek performansını koruduğunu göstermiştir. EffiViT3+ modelinin %83,94 Dice skoru ve %72,38 IoU değeri, verilerin azalmasına rağmen genelleme kapasitesini koruduğunu göstermektedir. Ayrıca, U-Net++ modelinin duyarlılık değerinde gösterdiği üstünlük (%87,27), modelin gerçek pozitifleri yakalama yeteneğinin yüksek olduğunu ancak yanlış pozitiflerde sorun yaşayabileceğini belirtmektedir.

Kayıp fonksiyonları ile yapılan analizlerde, yalnızca Dice veya Lovász kayıplarının yeterli olmadığı, sınır ve HD95 gibi ek bileşenlerle oluşturulan hibrit kayıp fonksiyonlarının

segmentasyon kalitesini önemli ölçüde artırdığı ortaya çıkmıştır. Lovász, Dice, FTV, sınır ve HD95 kayıp bileşenlerinin birlikte kullanılmasıyla elde edilen %77,43 IoU ve %87,16 Dice skorları, önerilen modelin özellikle ince detayların ve bölgesel tutarlılığın korunmasında büyük avantaj sağladığını göstermiştir.

Görsel incelemelere göre EffViT3+ modelinin her iki veri kümesinde de farklı boyut ve morfolojideki lezyonlarda yüksek doğruluk, kenar hassasiyeti ve yapısal bütünlük sağladığını göstermiştir. Özellikle, küçük, orta ve büyük boyuttaki nodüllerde model, referans maskelere yakın, hatasız segmentasyonlar gerçekleştirmiştir. Ayrıca, düşük kontrast ve artefakt gibi ultrason görüntülerine özgü zorluklara rağmen EffViT3+, yüksek iç homojenlik ve düzgün sınırlar ile üstün performans sergilemiştir. Rakip modellerde gözlemlenen taşma, kenar bozulması, alan daraltma ya da genişletme gibi hataların aksine, EffViT3+ segmentasyonları daha dengeli ve güvenilirdir. Bu sonuçlar, EffViT3+'ın klinik uygulamalarda tiroid nodül segmentasyonu için güçlü ve istikrarlı bir araç olduğunu ortaya koymaktadır.

Genel olarak sonuçlar değerlendirildiğinde, hibrit mimarilerin hem evrışimli hem de transformatör tabanlı tekil modellere kıyasla daha iyi performans gösterdiği net olarak ortaya çıkmıştır. Özellikle EfficientNet ve ViT kombinasyonunun yüksek performansı, hibrit yaklaşımların medikal görüntü segmentasyonu alanında başarılı olabileceğini göstermiştir.

7.2. Literatür ile Karşılaştırma

Tiroid nodülü segmentasyonu üzerine gerçekleştirilen bu çalışmada, önerilen EffViT3+ modelinin performansının değerlendirilmesi için TN3K ve DDTI veri kümeleri üzerinde yakın dönemde geliştirilmiş ve Tablo 2.1'de de gösterilen güncel çalışmalarla karşılaştırma yapılmıştır.

TN3K veri kümesi üzerinde yapılan karşılaştırmalarda literatürde öne çıkan modeller arasında, Tao ve arkadaşları tarafından geliştirilen LCA-Net %82.08 Dice skoru ve %71.18 IoU değerleriyle önemli bir başarı sergilemiştir. Bu modelin başarısında 7×1 ve 1×7 Mish tabanlı kod çözücü ile bağlam dikkat mekanizmasının rolü bulunmaktadır. Ancak, önerilen EffViT3+ modelimiz TN3K üzerinde %87,16 Dice skoru ve %77,43 IoU değeriyle bu modeli belirgin bir farkla geride bırakmıştır. EffViT3+'ın hibrit EfficientNet ve Vision Transformer tabanlı mimarisi, LCA-Net'in dikkat mekanizmasını aşarak global bağlam bilgisini daha etkili şekilde kullanmaktadır.

2023 yılında Ma ve arkadaşlarının önerdiği AMSeg modeli, TN3K veri kümesinde %84.21 Dice ve %72.48 IoU sonuçları elde etmiştir. AMSeg, özellikle çok ölçekli bölge çıkarımı ve adversarial füzyon yöntemini içermekte olup, segmentasyon doğruluğunu artırmaktadır. Ancak EffiViT3+'ın sonuçları AMSeg'e kıyasla Dice skorunda yaklaşık %3, IoU skorunda ise yaklaşık %5 oranında daha yüksek çıkmıştır. Bu durum EffiViT3+ modelinin farklı ölçeklerdeki özellikleri daha güçlü temsil etme kabiliyetini göstermiştir.

BPAT-UNet modeli, TN3K verisetinde %83,64 Dice skoru ve %71,87 IoU değerleri ile AMSeg ile benzer performans sunmuştur. Modelin başarısında BPSM nokta-dikkat ve ATM bottleneck füzyon modüllerinin etkisi öne çıkmaktadır. Buna rağmen önerilen EffiViT3+ modeli BPAT-UNet'in Dice ve IoU skorlarının sırasıyla %3,5 ve %5,5 üzerinde performans göstererek üstünlüğünü sergilemektedir.

Zhang ve arkadaşlarının DDTransUNet modeli TN3K veri kümesinde %82,31 Dice skoru ve %73.82 IoU elde etmiştir. Bu çift dallı Swin+CNN encoder kullanan model, GSA, GCA ve CAF gibi dikkat ve adaptif füzyon modüllerinden yararlanmaktadır. EffiViT3+ bu mimariye kıyasla da Dice skorunda yaklaşık %5 ve IoU skorunda yaklaşık %3,6 daha iyi sonuçlar elde ederek, transformatör ve evrişim temelli karmaşık yapıların daha basit ama etkili hibrit yapı ile geride bırakılabileceğini göstermiştir.

Liu ve arkadaşları tarafından önerilen BFG\&MSF-Net, TN3K veri setinde Dice skorunda %87,26, IoU'da %76,34 oranlarına ulaşarak literatürdeki en yüksek sonuçlardan birini sunmuştur. Özellikle sınır rehberlik modülü ve DSASPP modüllerini kullanan bu model, EffiViT3+'ın Dice skoruna yakın sonuç verirken IoU değerinde yaklaşık %1 daha geride kalmıştır. Ancak EffiViT3+'ın IoU skorundaki %77,43 oranı, BFG\&MSF-Net'in sınır belirlemedeki avantajını bile geride bırakmıştır.

Ali ve arkadaşları tarafından geliştirilen CIL-Net modeli TN3K veri setinde Dice skorunda %79,68, IoU'da %64,90 değerleriyle literatürdeki diğer modellere göre nispeten geride kalmıştır. Bu durum, CIL-Net'in yoğun bağlantı ve üçlü dikkat mekanizmasına rağmen EffiViT3+'ın sunduğu global bağlam bilgisinin üstünlüğünü göstermektedir. EffiViT3+ modelinin Dice skorunda yaklaşık %7,5 ve IoU skorunda %12,5'lik belirgin üstünlük elde edilmiştir.

2025 yılına ait modellerden Wu ve arkadaşları tarafından önerilen Improved Swin, TN3K veri setinde Dice %82,26 ve IoU %73,00 skorları sunarken Yetginler ve Atacak'ın Improved V-Net modeli %83,88 Dice ve %75,50 IoU elde etmiştir. Bu modeller özellikle

DSA tabanlı odaklanmış Dice kaybı ve üçlü füzyon kodlayıcı gibi güçlü modüllere sahip olsalar da geliştirilen EffiViT3+ modeli Dice skoru ve IoU'da belirgin bir farkla üstün çıkmıştır.

Son olarak, Li ve arkadaşlarının geliştirdiği GSE-Nets modeli TN3K üzerinde %81,97 Dice skoru ve %72,24 IoU skorları sunmaktadır. Bu model yapı çapraz-dikkat ve grafik evrişim blokları gibi karmaşık modüllere rağmen EffiViT3+'ın Dice skorunun yaklaşık %5,2, IoU değerinin ise yaklaşık %5,2 altında kalmaktadır.

DDTI veri kümesi açısından yapılan karşılaştırmalarda ise BPAT-UNet (%85,63 Dice skoru, %74,87 IoU) ve BFG\&MSF-Net (%85,12 Dice, %74,02 IoU) modelleri ön plana çıkmıştır. Önerilen EffiViT3+ DDTI üzerinde %83,94 Dice skoru ve %72,38 IoU elde ederek bu modellerle oldukça rekabetçi bir performans göstermiştir. DDTI veri setindeki örnek sayısının TN3K'ya göre daha az olması nedeniyle genelleştirme performansındaki bu hafif düşüş normaldir. Buna rağmen EffiViT3+, CIL-Net (%83,35 Dice skoru, %73,35 IoU), Improved Swin (%78,64 Dice skoru, %67,93 IoU) ve AMSeg (%74,81 Dice skoru, %60,89 IoU) gibi diğer birçok modeli geride bırakmıştır. Bu durum, EffiViT3+'ın farklı veri setlerinde de kararlı performans sağladığını ve genelleme yeteneğinin yüksek olduğunu ortaya koymaktadır.

Genel olarak, önerilen EffiViT3+ modelinin TN3K ve DDTI veri setleri üzerindeki performansı literatürdeki birçok güçlü ve güncel modelden daha yüksek düzeyde olup EfficientNet ve ViT tabanlı hibrit mimari yaklaşımının etkili olduğunu kanıtlamıştır. Bu sonuçlar, hibrit mimarilerin tiroid nodülü segmentasyonunda yerel özellikler ile global bağlam bilgisini dengeli biçimde kullanabilme kabiliyetlerinin önemini ve üstünlüğünü açıkça vurgulamaktadır.

7.3. Gelecek Çalışmalar

Bu çalışma kapsamında önerilen EffiViT3+ modelinin klinik uygulamalara entegrasyonunu hızlandırmak ve modelin etkinliğini daha da artırmak için çeşitli yönlerde ileriye dönük çalışmalar yapılmalıdır. Öncelikle, modelin gerçek zamanlı (real-time) çalışma yeteneğine kavuşması için performans optimizasyonu değerlendirilmelidir. Modelin klinik kullanım sırasında daha hızlı sonuç üretmesini sağlamak amacıyla hafifletilmiş transformatör yapıları ve EfficientNet mimarisinin sadeleştirilmiş versiyonlarının kullanılması gibi yöntemler değerlendirilmelidir. Ek olarak farklı merkezlerden ve çeşitli

görüntüleme cihazlarından elde edilen daha geniş veri setleri toplanmalıdır ve modelin klinik gerçeklik koşullarında test edilerek uzman görüşleri alınmalıdır.

EffiViT3+ modelinin sonuçlarının klinisyenler tarafından daha iyi yorumlanabilmesi için açıklanabilir yapay zeka yöntemlerinin entegrasyonu yapılabilir. Bu bağlamda Grad-CAM, SHAP ve LIME gibi popüler açıklanabilirlik teknikleriyle modelin hangi özelliklere dayanarak karar verdiğinin daha net anlaşılır hale getirilebilir. Böylece radyologların ve uzman klinisyenlerin model çıktılarını daha etkili bir şekilde yorumlamaları sağlanabilir.

Gelecekteki araştırmalarda, tiroid nodülü segmentasyonu ile nodüllerin benign/malign sınıflandırması gibi ek klinik görevleri bir araya getiren çok görevli öğrenme yaklaşımları değerlendirilebilir. Bu sayede tek bir modelin hem segmentasyon doğruluğunu artırması hem de klinik teşhis süreçlerini destekleyerek hasta değerlendirmelerini zenginleştirilmesi mümkün kılınabilir.

7.4. Sonuç

Bu tez çalışmasında, tiroid nodüllerinin ultrasonografi görüntülerinden otomatik segmentasyonu için önerilen EffiViT3+ modelinin geliştirilmesi, kapsamlı değerlendirilmesi ve literatürdeki diğer öncü yöntemlerle karşılaştırılması gerçekleştirilmiştir. Tiroid nodülleri, ultrasonografi görüntülerinde farklı büyüklüklerde, şekillerde ve kontrast özelliklerinde görülebilen, doğru segmentasyonu klinik karar verme süreçlerinde büyük önem taşıyan yapılardır. Bu bağlamda, önerilen segmentasyon modeli, tiroid ultrasonografi görüntülerindeki karmaşık yapıları yüksek doğrulukla yakalamayı ve klinisyenlerin teşhis doğruluğunu artırmayı amaçlamaktadır.

EffiViT3+ modelinin geliştirilmesinde, yerel özellikleri güçlü şekilde yakalayan EfficientNet mimarisi ile global bağlam bilgisini etkin biçimde kullanan Vision Transformer (ViT) yapılarının hibrit birleşimi tercih edilmiştir. EfficientNet'in evrişim katmanları küçük ayrıntılara odaklanarak yerel detayları başarılı bir şekilde yakalarken ViT tabanlı dikkat mekanizmaları görüntü içerisindeki küresel ilişkileri ve bağlam bilgisini daha iyi modelleyerek segmentasyon kalitesini artırmaktadır. Bu hibrit yaklaşım, modelin hem küçük nodüllerde hem de daha geniş yapıların segmentasyonunda yüksek performans göstermesine olanak tanımaktadır.

TN3K ve DDTI olmak üzere iki farklı veri seti üzerinde gerçekleştirilen kapsamlı deneyler, önerilen EffiViT3+ modelinin üstün segmentasyon performansını gösterdiğini ortaya koymaktadır. TN3K veri seti üzerinde yapılan testlerde EffiViT3+ modeli %87,16 Dice ve %77,43 IoU skorlarına ulaşarak literatürdeki birçok gelişmiş modeli geride bırakmıştır. Bu performans, geliştirilen modelin tiroid nodüllerinin sınırlarını yüksek doğrulukla belirleyerek hem yanlış negatif hem de yanlış pozitif oranlarını minimize ettiğini göstermektedir. Özellikle, AMSeg (%84,21 Dice skoru, %72,48 IoU), BPAT-UNet (%83,64 Dice skoru, %71,87 IoU) ve DDTransUNet (%82,31 Dice skoru, %73,82 IoU) gibi yakın dönemde geliştirilmiş modellerle karşılaştırıldığında, EffiViT3+ modelinin belirgin bir farkla daha yüksek sonuçlar sunduğu gözlemlenmiştir. Benzer şekilde, DDTI veri setindeki sınırlı sayıda örnek üzerinde de EffiViT3+ modeli (%83,94 Dice skoru, %72,38 IoU) literatürde yer alan Improved Swin (%78,64 Dice skoru, %67,93 IoU) ve AMSeg (%74,81 Dice skoru, %60,89 IoU) gibi modelleri geride bırakarak stabil bir genelleme yeteneği sergilemiştir.

Çalışmada, farklı mimari yapıların ve kayıp fonksiyonlarının performans üzerindeki etkileri detaylı bir şekilde analiz edilmiştir. Yapılan ablasyon deneyleri, hibrit kayıp fonksiyonlarının (Lovász, Dice, Focal Tversky, Boundary ve HD95 kayıp bileşenlerinin kombinasyonları) kullanılmasının segmentasyon doğruluğunu artırmada katkı sağladığını göstermiştir. Özellikle Lovász, Dice, Focal Tversky, Boundary ve HD95 kayıp bileşenlerinin birlikte kullanıldığı kombinasyonda, test setindeki Dice skoru ve IoU değerinin en yüksek seviyeye (%77,43 IoU, %87,16 Dice) ulaştığı sonucuna varılmıştır. Bu durum, segmentasyon performansının sadece mimari geliştirmelerle değil, aynı zamanda doğru kayıp fonksiyonlarıyla da önemli ölçüde iyileştirilebildiğini göstermektedir.

Sonuç olarak bu tez çalışması, hibrit mimarilerin ve özelleştirilmiş kayıp fonksiyonlarının tiroid nodülü segmentasyonu alanında büyük bir potansiyel taşıdığını göstermiştir. Önerilen EffiViT3+ modeli, yüksek performansı ve genelleştirme yeteneği sayesinde klinik uygulamalarda gelecekte önemli bir rol üstlenebilecek konumdadır.

KAYNAKLAR

- [1] S. Tamhane and H. Gharib, “Thyroid nodule update on diagnosis and management,” *Clinical Diabetes and Endocrinology*, vol. 2, pp. 1–10, 2016.
- [2] B. Çetin, U. Karapınar, K. Duman, Ö. Sağlam ve E. Dursun, “Tiroid Nodüllerine Yaklaşım,” *Arşiv Kaynak Tarama Dergisi*, vol. 22, no. 4, pp. 460–469, 2013.
- [3] R. Wong, S. G. Farrell ve M. Grossmann, “Thyroid nodules: diagnosis and management,” *Medical Journal of Australia*, vol. 209, no. 2, pp. 92–98, 2018.
- [4] Y. Haspolat, R. Tekin, F. Aktar ve M. Küçüköner, *Çocuk ve Ergenlerde Tiroid Hastalıkları*, 1. Baskı, 2019.
- [5] S. Benvenga, G. Tuccari, A. Ieni ve R. Vita, “Thyroid gland: anatomy and physiology,” in *Encyclopedia of Endocrine Diseases*, vol. 4, pp. 382–390, 2018.
- [6] B. A. Policeni, W. R. Smoker ve D. L. Reede, “Anatomy and embryology of the thyroid and parathyroid glands,” *Seminars in Ultrasound, CT and MRI*, vol. 33, no. 2, pp. 104–114, Nisan 2012.
- [7] M. P. Vanderpump, “The epidemiology of thyroid disease,” *British Medical Bulletin*, vol. 99, no. 1, pp. 39–51, 2011.
- [8] R. Prakash, A. Ramachandran, G. B. Savalgi, S. P. Venkata ve V. Mokhasi, “Variations in the anatomy of the thyroid gland: clinical implications of a cadaver study,” *Anatomical Science International*, vol. 87, pp. 45–49, 2012.
- [9] S. Özbaş, “Anatomi,” [Çevrim içi]. Erişim tarihi: Temmuz 2025.
- [10] M. Adaş, G. Adaş, F. Özülker ve O. Yalçın, “Tiroid nodülleri ve klinik önemi,” *Okmeydanı Tıp Dergisi*, vol. 28, pp. 20–25, 2012.

- [11] H. Aksoy ve S. Alıcı, “Tiroid Nodülleri: Genel özellikleri ve tanı metotları,” Van Tıp Dergisi, vol. 4, no. 3, pp. 187–192, 1997.
- [12] T. Şengöz, R. Çubuk, H. Kaya ve E. Arıbal, “Tiroid nodüllerinde ultrason rehberliğinde iğne aspirasyon biyopsisi,” Düzce Medical Journal, vol. 11, no. 2, pp. 26–32, 2009.
- [13] B. Çekiç, “Tiroid nodülü nedir? Belirtileri ve çeşitleri?,” [Çevrim içi]. Erişim tarihi: Temmuz 2025.
- [14] M. Özsan, İ. Üstün ve C. Gökçe, “Tiroid nodülleri,” The Medical Journal of Mustafa Kemal University, vol. 7, no. 27, 2016.
- [15] Ç. İmamoğlu, F. G. İmamoğlu, H. Dizen, A. Argon, Z. H. Adıbelli, B. P. Cengiz ve diğerleri, “Tiroid Nodüllerinde Ultrasonografi Eşliğinde İnce İğne Aspirasyon Sitolojisi: Sitohistolojik Korelasyon,” Muğla Sıtkı Koçman Üniversitesi Tıp Dergisi, vol. 2, no. 3, pp. 7–11, 2015.
- [16] H. M. Paksoy, S. Aydın, E. Ayduran, M. Eken, A. Şanlı ve Ö. Taşdemir, “Tiroid kitlelerinde klinik bulgular ve uyguladığımız tedavi yöntemleri,” The Turkish Journal of Ear Nose and Throat, vol. 18, no. 5, pp. 294–299, 2008.
- [17] K. T. Wong ve A. T. Ahuja, “Ultrasound of thyroid cancer,” Cancer Imaging, vol. 5, no. 1, p. 157, 2005.
- [18] D. A. Şahin ve O. Başpınar, “Tiroid Hastalıkları,” Türkiye Klinikleri Pediatric Sciences – Special Topics, vol. 11, no. 2, pp. 23–28, 2015.
- [19] M. Paksoy, S. Aydın, E. Ayduran, M. Eken, A. Sanlı ve Ö. Taşdemir, “Clinical signs and management strategies in thyroid masses,” Kulak Burun Boğaz İhtisas Dergisi, vol. 18, no. 5, pp. 294–299, 2008.

- [20] “Échographie,” [Çevrim içi]. Erişim tarihi: Temmuz 2025.
- [21] A. Demirdöven, “Ultrason,” *Türkiye Klinikleri Food Sciences – Special Topics*, vol. 7, no. 4, pp. 21–27, 2021.
- [22] R. Wang, T. Lei, R. Cui, B. Zhang, H. Meng ve A. K. Nandi, “Medical image segmentation using deep learning: A survey,” *IET Image Processing*, vol. 16, no. 5, pp. 1243–1267, 2022.
- [23] M. E. Rayed, S. S. Islam, S. I. Niha, J. R. Jim, M. M. Kabir ve M. F. Mridha, “Deep learning for medical image segmentation: State-of-the-art advancements and challenges,” *Informatics in Medicine Unlocked*, vol. 47, p. 101504, 2024.
- [24] V. Badrinarayanan, A. Kendall ve R. Cipolla, “SegNet: A deep convolutional encoder-decoder architecture for image segmentation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, 2017.
- [25] L.C. Chen, G. Papandreou, I. Kokkinos, K. Murphy ve A. L. Yuille, “DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, 2017.
- [26] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan ve S. Belongie, “Feature Pyramid Networks for object detection,” in *IEEE CVPR*, pp. 2117–2125, 2017.
- [27] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao ve J. Zhang, “CE-Net: Context encoder network for 2D medical image segmentation,” *IEEE Trans. Med. Imaging*, vol. 38, no. 10, pp. 2281–2292, 2019.
- [28] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov ve L.-C. Chen, “MobileNetV2: Inverted residuals and linear bottlenecks,” in *IEEE CVPR*, pp. 4510–4520, 2018.

- [29] K. He, G. Gkioxari, P. Dollár ve R. Girshick, “Mask R-CNN,” in IEEE ICCV, pp. 2961–2969, 2017.
- [30] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff ve H. Adam, “Encoder–decoder with atrous separable convolution for semantic image segmentation,” in ECCV, pp. 833–851, 2018.
- [31] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh ve J. Liang, “UNet 3+: A full-scale connected U-Net for medical image segmentation,” in Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pp. 3–11, Springer, 2020.
- [32] Y. Zhang, H. Liu ve Q. Hu, “TransFuse: Fusing transformers and CNNs for medical image segmentation,” arXiv:2102.08005, 2021.
- [33] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez ve P. Luo, “SegFormer: Simple and efficient design for semantic segmentation with transformers,” NeurIPS, vol. 34, pp. 12077–12090, 2021.
- [34] Y. Gao, M. Zhou ve D. N. Metaxas, “UTNet: a hybrid transformer architecture for medical image segmentation,” in MICCAI, pp. 61–71, Springer, Eylül 2021.
- [35] A. He, K. Wang, T. Li, C. Du, S. Xia ve H. Fu, “H2former: An efficient hierarchical hybrid transformer for medical image segmentation,” IEEE Trans. Med. Imaging, vol. 42, no. 9, pp. 2763–2775, 2023.
- [36] H. Tang, Y. Chen, T. Wang, Y. Zhou, L. Zhao, Q. Gao ve diğerleri, “HTC-Net: A hybrid CNN-transformer framework for medical image segmentation,” Biomed. Signal Process. Control, vol. 88, p. 105605, 2024.
- [37] L. Lan, P. Cai, L. Jiang, X. Liu, Y. Li ve Y. Zhang, “BRAU-Net++: U-shaped hybrid

- CNN-transformer network for medical image segmentation,” arXiv:2401.00722, 2024.
- [38] S. Perera, Y. Erzurumlu, D. Gulati ve A. Yılmaz, “MobileUNETR: a lightweight end-to-end hybrid vision transformer for efficient medical image segmentation,” in ECCV, pp. 281–299, Springer, Eylül 2024.
- [39] Z. Wang, Y. Zou ve P. X. Liu, “Hybrid dilation and attention residual U-Net for medical image segmentation,” *Comput. Biol. Med.*, vol. 134, p. 104449, 2021.
- [40] Z. Xu ve Z. Wang, “MCV-UNet: A modified convolution & transformer hybrid encoder-decoder network with multi-scale information fusion for ultrasound image semantic segmentation,” *PeerJ Comput. Sci.*, vol. 10, p. e2146, 2024.
- [41] Y. Zhang, Z. Li, N. Nan ve X. Wang, “TranSegNet: hybrid CNN-vision transformers encoder for retina segmentation of optical coherence tomography,” *Life*, vol. 13, no. 4, p. 976, 2023.
- [42] X. Ma, B. Sun, W. Liu, D. Sui, J. Chen ve Z. Tian, “Amseg: A novel adversarial architecture based multi-scale fusion framework for thyroid nodule segmentation,” *IEEE Access*, vol. 11, pp. 72911–72924, 2023.
- [43] H. Bi, C. Cai, J. Sun, Y. Jiang, G. Lu, H. Shu ve X. Ni, “BPAT-UNet: Boundary preserving assembled transformer UNet for ultrasound thyroid nodule segmentation,” *Comput. Methods Programs Biomed.*, vol. 238, p. 107614, 2023.
- [44] Anonymous, “Local and Context-Attention Adaptive LCA-Net for Thyroid Nodule Segmentation in Ultrasound Images,” 2024.
- [45] H. Ali, M. Wang ve J. Xie, “Cil-net: Densely connected context information learning network for boosting thyroid nodule segmentation using ultrasound images,” *Cognitive Computation*, vol. 16, no. 3, pp. 1176–1197, 2024.

- [46] X. Deng, Z. Dang ve L. Pan, “STU3Net: An Improved U-Net With Swin Transformer Fusion for Thyroid Nodule Segmentation,” *Int. J. Imaging Syst. Technol.*, vol. 34, no. 5, p. e23160, 2024.
- [47] C. Zhang, L. Wang, G. Wei, Z. Kong ve M. Qiu, “A dual-branch and dual attention transformer and CNN hybrid network for ultrasound image segmentation,” *Front. Physiol.*, vol. 15, p. 1432987, 2024.
- [48] J. Liu, J. Mu, H. Sun, C. Dai, Z. Ji ve I. Ganchev, “BFG&MSF-Net: Boundary Feature Guidance and Multi-Scale Fusion Network for Thyroid Nodule Segmentation,” *IEEE Access*, vol. 12, pp. 78701–78713, 2024.
- [49] Y. Wu, L. Huang ve T. Yang, “Thyroid Nodule Ultrasound Image Segmentation Based on Improved Swin Transformer,” *IEEE Access*, 2025.
- [50] B. Yetginler ve İ. Atacak, “An Improved V-Net Model for Thyroid Nodule Segmentation,” *Applied Sciences*, vol. 15, no. 7, p. 3873, 2025.
- [51] X. Li, C. Fu, Q. Wang, W. Zhang, C. Ye ve T. Ma, “GSE-Nets: Global Structure Enhancement decoder for thyroid nodule segmentation,” *Biomed. Signal Process. Control*, vol. 102, p. 107340, 2025.
- [52] H. Gong, J. Chen, G. Chen, H. Li, G. Li ve F. Chen, “Thyroid region prior guided attention for ultrasound segmentation of thyroid nodules,” *Comput. Biol. Med.*, vol. 155, p. 106389, 2023.
- [53] L. Pedraza, C. Vargas, F. Narváez, O. Durán, E. Muñoz ve E. Romero, “An open access thyroid ultrasound image database,” in *10th Int. Symp. Med. Information Processing and Analysis*, pp. 188–193, SPIE, Ocak 2015.
- [54] Q. Zhang, “A novel ResNet101 model based on dense dilated convolution for image

- classification,” SN Applied Sciences, vol. 4, no. 1, p. 9, 2022.
- [55] G. W. Wicaksono ve A. Andreawan, “ResNet101 model performance enhancement in classifying rice diseases with leaf images,” RESTI (Rekayasa Sistem ve Teknologi Informasyon), vol. 7, no. 2, pp. 345–352, 2023.
- [56] A. Bello, S. C. Ng ve M. F. Leung, “Skin cancer classification using fine-tuned transfer learning of DENSENET-121,” Applied Sciences, vol. 14, no. 17, p. 7707, 2024.
- [57] Z. P. Jiang, Y. Y. Liu, Z. E. Shao ve K. W. Huang, “An improved VGG16 model for pneumonia image classification,” Applied Sciences, vol. 11, no. 23, p. 11185, 2021.
- [58] E. Erdemir ve Ç. B. Erdaş, “Development of Hybrid U-Net Models with Pre-trained VGG16 for Thyroid Nodule Segmentation,” in TIPTEKNO 2024, pp. 1–4, IEEE, Ekim 2024.
- [59] A. Bansal, R. Sharma, V. Sharma, A. K. Jain ve V. Kukreja, “Detecting severity levels of cucumber leaf spot disease using ResNeXt deep learning model: a digital image analysis approach,” in INCET 2023, pp. 1–6, IEEE, Mayıs 2023.
- [60] L. Li, Y. Zhu ve Z. Zhu, “Automatic modulation classification using ResNeXt-GRU with deep feature fusion,” IEEE Trans. Instrum. Meas., vol. 72, pp. 1–10, 2023.
- [61] M. Latha, P. S. Kumar, R. R. Chandrika, T. R. Mahesh, V. V. Kumar ve S. Guluwadi, “Revolutionizing breast ultrasound diagnostics with EfficientNet-B7 and Explainable AI,” BMC Medical Imaging, vol. 24, no. 1, p. 230, 2024.
- [62] A. Ghosh, B. Soni ve U. Baruah, “Transfer learning-based deep feature extraction framework using fine-tuned EfficientNet-B7 for multiclass brain tumor classification,” Arabian J. Sci. Eng., vol. 49, no. 9, pp. 12027–12048, 2024.

- [63] S. Ranganathan, “Deep reinforcement learning for sequential recommendations,” arXiv preprint arXiv:1905.12767, 2019.
- [64] Y. Bengio, A. Courville ve P. Vincent, “Representation learning: A review and new perspectives,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, 2013.
- [65] Y. Lecun, Y. Bengio ve G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [66] J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural Networks*, vol. 61, pp. 85–117, 2015.
- [67] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian ve diğerleri, “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [68] F. Shi, J. Cheng, L. Wang, P.-T. Yap ve D. Shen, “LSTM-based BRNNs for MR image sequence prediction,” *IEEE Trans. Med. Imaging*, vol. 37, no. 11, pp. 2490–2501, 2018.
- [69] Ö. Yıldız, M. B. Yıldız, E. Gültepe, A. Yardımcı ve M. Tunç, “Derin öğrenme ile tiroid nodüllerinin sınıflandırılması,” *Bilişim Teknolojileri Dergisi*, vol. 12, no. 2, pp. 123–130, 2019.
- [70] S. Iandola, M. Moskewicz, K. Ashraf, S. Han, W. Dally ve K. Keutzer, “SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size,” arXiv preprint arXiv:1602.07360, 2016.
- [71] A. Krizhevsky, I. Sutskever ve G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.

- [72] K. Simonyan ve A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint arXiv:1409.1556, 2014.
- [73] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov ve diğerleri, “Going deeper with convolutions,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2015, pp. 1–9.
- [74] K. He, X. Zhang, S. Ren ve J. Sun, “Deep residual learning for image recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 770–778.
- [75] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov ve L. Chen, “MobileNetV2: Inverted residuals and linear bottlenecks,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2018, pp. 4510–4520.
- [76] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2017, pp. 1800–1807.
- [77] D. Karimi, H. Dou, S. K. Warfield ve S. Gholipour, “Deep learning with noisy labels: Exploring techniques and remedies in medical image analysis,” *Medical Image Analysis*, vol. 65, p. 101759, 2020.
- [78] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu ve M. Lew, “Deep learning for visual understanding: A review,” *Neurocomputing*, vol. 187, pp. 27–48, 2016.
- [79] C. Shorten ve T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, 2019.
- [80] Y. Wang, L. Yao, H. Zhao, P. Sheng, X. Sun, L. Zhu ve diğerleri, “Machine learning in medical imaging,” *Journal of Biomedical Informatics*, vol. 109, pp. 103–115, 2020.

- [81] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou ve diğerleri, “A guide to deep learning in healthcare,” *Nature Medicine*, vol. 25, no. 1, pp. 24–29, 2019.
- [82] M. Razzak, S. Naz ve A. Zaib, “Deep learning for medical image processing: Overview, challenges and the future,” in *Classification in BioApps*, Cham: Springer, 2018, pp. 323–350.
- [83] A. Ker, L. Wang, J. Rao ve T. Lim, “Deep learning applications in medical image analysis,” *IEEE Access*, vol. 6, pp. 9375–9389, 2018.
- [84] X. Liu, L. Faes, A. Kale, S. Wagner, D. Fu, A. Bruynseels ve diğerleri, “A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis,” *The Lancet Digital Health*, vol. 1, no. 6, pp. e271–e297, 2019.