

TOBB EKONOMİ VE TEKNOLOJİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**TRANSFER ÖĞRENME TABANLI EVRİŞİMLİ SİNİR AĞLARINI
KULLANAN OTOMATİK DENTAL PANORAMİK GÖRÜNTÜ
SEGMENTASYONU**

YÜKSEK LİSANS TEZİ

Tülin ÇAYLAK

Elektrik ve Elektronik Mühendisliği Anabilim Dalı

Tez Danışmanı: Doç. Dr. İmam Şamil YETİK

AĞUSTOS 2022

TEZ BİLDİRİMİ

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, alıntı yapılan kaynaklara eksiksiz atıf yapıldığını, referansların tam olarak belirtildiğini ve ayrıca bu tezin TOBB ETÜ Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırlandığını bildiririm.

Tülin Çaylak

İMZA

ÖZET

Yüksek Lisans Tezi

TRANSFER ÖĞRENİMİ TABANLI EVRİŞİMLİ SİNİR AĞLARINI KULLANAN OTOMATİK DENTAL PANORAMİK GÖRÜNTÜ SEGMENTASYONU

Tülin Çaylak

TOBB Ekonomi ve Teknoloji Üniversitesi
Fen Bilimleri Enstitüsü
Elektrik ve Elektronik Mühendisliği Anabilim Dalı

Tez Danışmanı: Doç. Dr. İmam Şamil Yetik

Tarih: Ağustos 2022

Radyolojik incelemelerde kullanılan X-Ray görüntüleri diş hekimlerinin hastanın problemini anlamak ve çözmek için başvurduğu temel verilerden biridir. Bu görüntülerden panoramik X-ray görüntüleri tüm ağız bölgesini göstermesi sebebiyle hastanın probleminin teşhis ve tedavisi açısından daha kapsamlıdır. Ayrıca diş harici birçok bölümü de içerdiğinden en karmaşık yapıya sahip görüntülerdir. Öte yandan kullanılan cihazlardan kaynaklı farklı görüntü niteliklerine sahiptirler. Bunlar göz önünde bulundurulduğunda panoramik X-ray görüntülerinin doğru işlenmesi zorlayıcı bir süreçtir ve başarılı bir şekilde analiz edilmesi büyük önem teşkil etmektedir. Bu süreci otomatik hale getirerek uzmanlara yardımcı olmak için çeşitli klasik görüntü işleme algoritmaları geliştirilmiştir. Yıllar boyunca veri setlerinin boyutlarının artması, kullanılan cihazların gelişmesi gibi etkenlerle birlikte derin öğrenme (DL) teknolojisi daha popüler hale gelmiş ve uygulama alanları genişlemiştir. Dolayısıyla klasik görüntü işleme algoritmalarının yanı sıra X-Ray görüntüleri birçok amaç için derin öğrenme (DL) teknolojisi kullanılarak da analiz

edilmeye başlanmıştır. Fakat dış veri setlerinin boyutları diğer uygulama alanlarında bulunan veri setlerinin boyutlarına kıyasla oldukça azdır. Bu tezin amacı küçük veri seti boyutu sorununa transfer yöntemi ile çözüm bularak panoramik X-ray görüntüleri ile başarılı bir dış segmentasyonu gerçekleştirmektir. Elde bulunan panoramik X-ray görüntülerinin boyutlarının farklılığını gidermek için bütün veri seti yeniden boyutlandırılarak aynı formatta olmaları sağlanmıştır. Seçilen görüntü düzenleme uygulamasıyla modellerin eğitiminde kullanılacak görüntü maskeleri oluşturulmuştur. Transfer öğrenme yöntemi için ihtiyaç duyulan önceden eğitilmiş derin öğrenme (DL) modeli olarak U-Net ve InceptionResNet-V2 modelleri seçilmiş ve iki farklı derin öğrenme mimarisi oluşturulmuştur. InceptionResNet-V2 modeli halka açık görüntü veri setiyle önceden eğitildiği için kütüphane olarak hazır halde bulunmaktadır. U-Net modeli ise halka açık olarak yayınlanan göğüs X-Ray veri seti ile eğitilmiştir. Önceden eğitilmiş modeller ile birleştirilerek panoramik X-ray görüntüleri üzerinden eğitilecek olan temel derin öğrenme (DL) modeli olarak U-Net mimarisi seçilmiştir. Böylece iki farklı derin öğrenme mimarisi oluşturularak eğitim gerçekleştirilmiştir. Modellerin veri ezberlemesine karşı veri artırma yöntemi bütün dış veri setine uygulanmıştır. Ayrıca erken sonlandırma yöntemiyle de yine veri ezberleme probleminin önüne geçilmiştir. Oluşturulan test veri seti kullanılarak modellerin doğrulukları zar katsayısıyla ölçülmüş ve sonuçlar karşılaştırılmıştır.

Anahtar Kelimeler: Medikal görüntü segmentasyonu, U-Net, Transfer öğrenme, Panoramik X-Ray, Derin öğrenme, Inception-ResNet-V2, Zar katsayısı.

ABSTRACT

Master of Science

AUTOMATED DENTAL PANORAMIC IMAGE SEGMENTATION USING TRANSFER LEARNING BASED CONVOLUTIONAL NEURAL NETWORKS

Tülin Çaylak

TOBB University of Economics and Technology
Institute of Natural and Applied Sciences
Department of Electrical and Electronics Engineering

Supervisor: Assoc. Prof. İmam Şamil Yetik

Date: August 2022

X-Ray images used in radiological examinations are one of the basic data that dentists apply to understand and solve the patient's problem. Panoramic X-ray images from these images are more comprehensive in terms of diagnosis and treatment of the patient's problem, since they show the entire mouth region. In addition, since it includes many parts other than teeth, they are the most complex images. On the other hand, they have different image qualities due to the devices used. With these in mind, the correct processing of panoramic X-ray images is a challenging process and successful analysis is of great importance. Various classical image processing algorithms have been developed to assist experts by automating this process. Over the years, deep learning (DL) technology has become more popular and its application areas have expanded with the increase in the size of the data sets and the development of the devices used. Therefore, in addition to classical image processing algorithms, X-Ray images have been started to be analyzed using deep learning (DL) technology for many purposes. However, the dimensions of dental datasets are quite small compared to the sizes of datasets found in other

application areas. The aim of this thesis is to find a solution to the small data set size problem with the transfer method and to perform a successful tooth segmentation with panoramic X-ray images. In order to eliminate the differences in the sizes of the panoramic X-ray images available, the whole data set was resized to ensure that they were in the same format. The image masks to be used in the training of the models were created with the selected image editing application. U-Net and InceptionResNet-V2 models were chosen as the pre-trained deep learning (DL) model needed for the transfer learning method and two different deep learning architectures were created. Since the InceptionResNet-V2 model is pre-trained with the public image dataset, it is available as a library. The U-Net model, on the other hand, was trained with a publicly published chest X-Ray dataset. U-Net architecture was chosen as the basic deep learning (DL) model, which will be trained on panoramic X-ray images by combining with pre-trained models. Thus, training was carried out by creating two different deep learning architectures. Data augmentation method against data memorization of the models was applied to the whole dental data set. In addition, with the early termination method, the data memorization problem is prevented. Using the created test data set, the accuracy of the models was measured with the dice coefficient and the results were compared.

Keywords: Medical image segmentation, U-Net, Transfer learning, Panoramic X-Ray, Deep learning, Inception-ResNet-V2, Dice coefficient.

TEŐEKKÜR

Çalıőmalarım boyunca deęerli yardım ve katkılarıyla beni yönlendiren hocam Doç. Dr. İmam Őamil Yetik'e, kıymetli tecrübelerinden faydalandığım TOBB Ekonomi ve Teknoloji Üniversitesi Elektrik ve Elektronik Mühendislięi Bölümü öğretim üyelerine ve destekleriyle her zaman yanımda olan aileme çok teşekkür ederim.



İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	v
ABSTRACT	vii
TEŞEKKÜR	ix
İÇİNDEKİLER	xi
ŞEKİL LİSTESİ	xiii
ÇİZELGE LİSTESİ	xv
KISALTMALAR	xvii
1. GİRİŞ	1
1.1 Literatür Araştırması	2
1.2 Tezin Amacı	4
1.3 Tezin Akışı	5
2. METODOLOJİ	7
2.1 Derin Öğrenme (DL)	7
2.2 Evrişimli Sinir Ağları (CNN).....	9
2.2.1 U-Net mimarisi	11
2.2.2 Inception-ResNet-v2 mimarisi	13
2.3 Transfer Öğrenme	16
2.4 Önerilen Ağ Mimarisi	18
3. DENEYLER	23
3.1 Veri Seti.....	23
3.1.1 Panoramik X-Ray veri seti.....	23
3.1.2 Göğüs X-Ray veri seti.....	26
3.1.3 ImageNet veri seti	26
3.2 Performans Metrikleri	27
3.3 Model Eğitim Parametreleri	29
3.4 Model Eğitimi	30
3.4.1 Double-UNet modeli eğitimi	32
3.4.2 InceptionResV2-UNet modeli eğitimi	36
4. SONUÇLAR	39
KAYNAKLAR	41

ŞEKİL LİSTESİ

Sayfa

Şekil 2.1 : Basit bir derin sinir ağı yapısı.....	8
Şekil 2.2 : Doğrultulmuş doğrusal birim (ReLU) aktivasyon fonksiyonu.....	10
Şekil 2.3 : Sigmoid aktivasyon fonksiyonu.....	11
Şekil 2.4 : Orijinal U-Net mimarisi.....	13
Şekil 2.5 : Inception-ResNet-v2 mimarisinin şeması.....	14
Şekil 2.6 : (a) 35x35 ızgara modülü şeması, (b) 17x17 ızgara modülü şeması, (c) 8x8 ızgara modülü.....	15
Şekil 2.7 : Kalıntı bağlantılarının ölçeklendirme aşaması.....	15
Şekil 2.8 : (a) Klasik DL öğrenme süreci, (b) Transfer öğrenme süreci.....	17
Şekil 2.9 : Double-UNet mimarisi.....	20
Şekil 2.10 : InceptionResV2-UNet mimarisi.....	21
Şekil 3.1 : (a) Tüm dişlere sahip hastanın X-Ray görüntüsü, (b) Eksik dişlere sahip hastanın X-Ray görüntüsü.....	24
Şekil 3.2 : (a) Model girdi görüntüsü, (b) Maske görüntüsü.....	25
Şekil 3.3 : (a) Girdi görüntüsü, (b) Maske görüntüsü.....	26
Şekil 3.4 : Memeli ve araç alt-ağaçlarından örnek görüntüler.....	27
Şekil 3.5 : Veri artırma yöntemleriyle elde edilen örnek görüntüler.....	31
Şekil 3.6 : (a) Göğüs X-Ray girdi görüntüsü, (b) Gerçek maske görüntüsü ve (c) U-Net modelinin tahmin görüntüsü.....	33
Şekil 3.7 : Double-UNet modelinin öğrenme eğrisi.....	34
Şekil 3.8 : (a) Orijinal girdi görüntüsü, (b) Gerçek maske görüntüsü ve (c) Double-UNet modelinin tahmin maske görüntüsü.....	35
Şekil 3.9 : InceptionResV2-UNet modelinin öğrenme eğrisi.....	37
Şekil 3.10 : (a) Orijinal girdi görüntüsü, (b) Gerçek maske görüntüsü ve (c) InceptionResV2-UNet modelinin tahmin maske görüntüsü.....	38

ÇİZELGE LİSTESİ

Sayfa

Çizelge 3.1 : Ön işlem öncesi ve sonrası örnek diş görüntü boyutları.....	24
Çizelge 3.2 : Model eğitim parametreleri	31
Çizelge 3.3 : Double-UNet modelinin performans sonucu.....	32
Çizelge 3.4 : InceptionResV2-UNet modelinin performans sonucu.....	36



KISALTMALAR

CNN	: Evrişimli Sinir Ağı (Convolutional Neural Network)
DL	: Derin Öğrenme (Deep Learning)
Faster R-CNN	: Daha Hızlı Bölgesel Evrişimsel Sinir Ağları (Faster Region Based Convolutional Neural Networks)
FN	: Yanlış Negatif (False Negative)
FP	: Yanlış Pozitif (False Positive)
Mask R-CNN	: Maskeli Bölgesel Evrişimsel Sinir Ağları (Mask Region Based-Convolutional Neural Networks)
MAE	: Ortalama Mutlak Hatası (Mean Absolute Error)
MSE	: Ortalama Kare Hatası (Mean Square Error)
ML	: Makine Öğrenmesi (Machine Learning)
ReLU	: Doğrultulmuş Doğrusal Birim (Rectified Linear Unit)
RNN	: Tekrarlayan Sinir Ağları (Recurrent Neural Network)
TN	: Gerçek Negatif (True Negative)
TP	: Gerçek Pozitif (True Positive)

1. GİRİŞ

Diş hekimleri için hastanın probleminin doğru teşhis edilmesi ve başarılı bir tedavi yöntemi ile giderilmesi kritik bir süreçtir. Bir teşhisin doğru ve zamanında yapılması hastanın sağlık probleminin çözümü için çok önemlidir. Yanlış teşhis sonrası yanlış tedavi uygulanabileceğinden bu iki aşamanın beraber yönetilmesi gerekmektedir. Teşhis aşamasında diş hekimleri hastanın geçmiş teşhis ve tedavileri, rahatsızlıkları, ilk muayene bulgusu ve radyografiler gibi çeşitli verileri kullanmaktadırlar.

Bu verilerden radyografi görüntüleri uzmanların muayene ile göremeyeceği bulguları görmeye olanak sağladığından sıklıkla kullanılmaktadır. Radyografi görüntüleri ağız bölgesi içinden veya ağız bölgesi dışından alınan görüntüler olmak üzere iki gruba ayrılmaktadır. Ağız bölgesi içinden alınan görüntülerden biri olan bitewing [1], ağız içine yerleştirilen küçük bir film tablet ile elde edilirken alt ve üst dişlerin kuronlarını göstermektedir. Diğer bir ağız bölgesi içinden alınan radyografi görüntüleri ise periapikallerdir. Periapikal görüntüler kurondan diş köküne kadar tüm dişi göstermektedir [1]. Ağız bölgesi dışından alınan radyografiler ise panoramik X-Ray'ler olarak adlandırılmaktadırlar [1]. Panoramik X-Ray görüntüleri tüm ağız bölgesini içerdiğinden diş hekimleri tarafından en başvurulan verilerden biridir. Tüm ağız bölgesinin görülmesine ve incelenmesine imkân vermektedirler. Bu avantajlarının yanı sıra oldukça karmaşık bir yapıya sahip olabilirlerken aynı zamanda kullanılan cihazların farklılıklarından kaynaklı görüntü kalitesinde değişiklikler oluşabilmektedir. Bu sebeple panoramik radyografi görüntülerinin analiz edilmesi zor olabilmektedir. Dolayısıyla bu verilerin doğru analizi ikinci aşama olan tedavi süreci için büyük önem teşkil etmektedir. Bu şartlar göz önünde bulundurularak radyografi görüntülerinin daha kolay, doğru ve otonom analiz edilebilmesi amacıyla birçok yöntem geliştirilmiştir. Bu yöntemler bitewing, periapikal veya panoramik X-Ray görüntüleri kullanılarak tasarlanmıştır. Kullanılan veri çeşidine göre segmentasyon, algılama veya sınıflandırma gibi farklı amaçlar için

kullanılabilecek modeller elde edilmiştir. İlk zamanlarda uzmanlar tarafından klasik görüntü işleme algoritma temelli modeller tasarlanırken gelişen teknolojiyle birlikte sürecin daha otonom hale getirilmesi amacıyla son zamanlarda derin öğrenme (DL) modelleri ile yeni yaklaşımlar geliştirilmektedir.

1.1 Literatür Araştırması

Geçmişten günümüze bitewing, periapikal ve panoramik X-Ray görüntüleri kullanılarak yarı-otomatik veya tam-otomatik olmak üzere diş segmentasyonu için klasik görüntü işleme algoritmaları ve derin öğrenme (DL) modelleri geliştirilmiştir. Klasik görüntü işleme algoritmaları sınır-bazlı, watershed-bazlı, alan-bazlı ve küme-bazlı olmak üzere dört yaklaşım ile uygulanmıştır. Sınır-bazlı yaklaşım da görüntüdeki gri seviyesi devamsızlıklarına bakarak segmentasyon yapan kenar algılama algoritması ve aranan bölgenin dış hattının şeklini çizerek segmentasyon gerçekleştiren aktif dış hat algoritması kullanılan yöntemlerdir. Çalışma [2]'de iki farklı kenar algılama algoritması (Sobel ve Canny) kullanılarak segmentasyon gerçekleştirilmiştir. Elde edilen sonuçlara göre Sobel metodu Canny metoduna göre X-Ray'deki tüm diş bölgesinde yer alan dişleri segmente ederken gürültüyü de filtrelediği belirtilmiştir. Veri seti boyutu belirtilmediği için çalışmanın başarısı yorumlanamamaktadır. Aynı şekilde çalışma [3] ile Canny algoritması kullanılarak yeni bir yaklaşım benimsenmiştir. Önerilen yaklaşım görüntü sınıflandırma, ön işleme, segmentasyon, özellik çıkarımı ve görüntü eşleştirmesi olmak üzere 5 adımdan oluşmaktadır. Tasarlanan algoritmanın başarısı piksel düzeyinde karşılaştırma yapılarak ölçülmektedir. Kenarsız aktif dış hat modeli [4] kullanılarak 340 adet diş üzerinden otomatik diş segmentasyonu gerçekleştirilmiştir. Geliştirilen yöntem diş bölgesinin yoğunluk değerine bağımlı olarak çalıştığından X-Ray kalitesine göre elde edilen başarı değişken olabilirken %58.10 oranında segmentasyon başarısının elde edildiği belirtilmiştir. Çalışma [5]'te verilen tasarım yarı-otomatik ikklendirme, eğime yakınsama ve son-işleme olmak üzere üç aşamadan oluşan aktif dış hat modelinden oluşmaktadır. Farklı tipteki toplam 38 adet X-Ray görüntü üzerinde test edilen bu tasarımın başarısının önceki çalışmalara göre daha iyi olduğu belirtilmiştir. Watershed-bazlı yaklaşımlarda kullanılan watershed

algoritması, bir görüntüyü bitişik bölgelerde segmentlere ayırmak için matematiksel morfolojiyi kullanan bir dönüşümdür [6]. Bu yaklaşımı baz alan çalışma [7]'de iyileştirme, sınır çıkarma ve tohum kombinasyonundan oluşan modifiye edilmiş bir algoritma uygulanmaktadır. Burada 599 adet diş üzerinde test edilen yöntem yaklaşık %86'lık bir başarı elde etmiştir. Yarı-otomatik olarak çalışan bu yöntem manuel ön işlemeye ihtiyaç duymaktadır. Önerilen başka bir çalışma [8] ile her bir dişin ayrı ayrı işlenebilmesi ve aşırı segmentasyonun önüne geçebilmek için ön-işlem aşaması kullanılmaktadır. Kontrastın artırılması, gürültünün azaltılması ve gereksiz bölgelerin filtrelenmesi adımlarından sonra watershed algoritması uygulanmaktadır. Tasarlanan model bitewing X-Ray görüntüleri ile test edilmiştir. Alan-bazlı yaklaşımlar [9, 10] da belirlenen parametre ve koşullara göre görüntüyü ilgili alanlara bölerek çalışmaktadır. Görüntüde yer alan benzerlikler temel alınarak gruplama yöntemi kullanılarak tasarlanan küme-bazlı algoritmalarından biri olan çalışma [11] radyografi segmentasyonu, öznitelik çıkarımı ve öznitelik eşleştirme adımlarını içeren yarı otomatik bir modeldir. Modelin çalışma mantığı gereği görüntüler bulanık olduğunda bu yaklaşımı uygulamak zordur. Küme-bazlı algoritmalarda grup sayısı gibi başlangıç parametrelerinin ayarlanması söz konusu olduğundan genel de bu modeller yarı-otomatik olarak çalışmaktadır.

Teknolojinin gelişmesi ve veri setlerindeki artışlara bağlı olarak klasik görüntü işleme algoritmalarının yanı sıra derin öğrenme (DL) tabanlı modeller de geliştirilmeye başlanmıştır. Bu modellerden bazıları maske-bölge tabanlı evrişimli sinir ağları (Mask R-CNN [12]) veya daha hızlı-bölge tabanlı evrişimli sinir ağları (Faster R-CNN [13]) mimarisi kullanarak segmentasyon veya sınıflandırma gerçekleştirmektedir. Çalışma [14]'de maske-bölge tabanlı evrişimli sinir ağları (Mask R-CNN [12]) kullanılırken çalışma [15]'te daha hızlı-bölge tabanlı evrişimli sinir ağları (Faster R-CNN [13]) ile sonuçlar test edilmiştir. [14] ile önerilen model 193 panoramik X-Ray görüntüsü kullanılarak eğitilmiş ve 1224 görüntü üzerinden %98 doğruluk elde etmiştir. Buna karşın 800 periapikal X-Ray görüntüleri ile eğitildikten sonra 250 test görüntüsüyle başarısı ölçülen [15]'teki modelle %91 oranında başarı elde edilmiştir. İki çalışmada da segmente edilen dişler bir çerçeveye içine alınarak gösterilmektedir. Bazı çalışmalar son zamanlarda tasarlanmış ve

medikal görüntü segmentasyonunda başarılı sonuçlar veren farklı bir evrişimli sinir ağı (CNN) mimarisine sahip olan U-Net [16] modeli kullanılarak geliştirilmiştir. Bu yaklaşımlardan birinde [17], klasik U-Net modeli modifiye edilerek 1201 görüntü ile eğitilmiş %93.4 zar katsayısı değeri elde edilmiştir. Çalışma [18]'de geliştirilen yaklaşım ise 942 görüntü ile eğitilerek 63 görüntü üzerinde test edilmiş ve %94 zar katsayısı değerine ulaşmıştır. Geliştirilen yöntemlerin yanı sıra büyük boyuttaki veri setleriyle önceden eğitilmiş modellerin kullanılmasını mümkün kılan transfer öğrenme yöntemi ile yeni yaklaşımlar geliştirilmiştir. Çalışma [19]'da tek bir evrişimli sinir ağıyla (CNN) oluşturulan model ve transfer öğrenme yöntemiyle oluşturulan model karşılaştırılarak 26 görüntü üzerinden test edilmiş ve sırasıyla %73 ile %88 başarı elde edilmiştir. Aynı yöntem kullanılarak tasarlanan başka bir model [20], belirlenen performans metriğiyle yaklaşık %95 oranında başarıya ulaşmıştır. Çalışma [21]'de göğüs röntgenleriyle eğitilen bir ağ mimarisi farklı model eğitim parametreleriyle mamogram verisini öğrenmek için kullanılmış ve %90.38 oranında doğruluk elde ettiği belirtilmiştir. Bu çalışmaya benzer göğüs kanseri tespiti ve sınıflandırması için tasarlanmış [22, 23] çalışmalar bulunmaktadır. Ayrıca insan deri yanıklarının değerlendirilmesi [24] ve bu yanıklara ait yaraların renkli görüntü segmentasyonu [25] için yine transfer öğrenme yöntemine başvurulmuştur. Bu çalışmalarda modellerin başarısının test edilmesi için genellikle doğruluk, kesinlik gibi metrikler kullanılmaktadır. Örneğin çalışma [25] de iki farklı ağ mimarisi kullanılmakta ve segmentasyon yapılmaktadır. Tasarlanan mimariler sırasıyla %85 ve %95 oranında kesinlik elde etmiştir.

1.2 Tezin Amacı

Bu tez, panoramik X-Ray görüntüleri ile otomatik diş segmentasyonu için tek bir evrişimli sinir ağı (CNN) kullanmak yerine transfer öğrenme yöntemi kullanılarak farklı yapıdaki evrişimli sinir ağlarının birleştirilmesiyle elde edilen ağların kullanımını önermektedir. Bu doğrultuda U-Net [16] ve InceptionResNet-V2 [26] olmak üzere iki farklı evrişimli sinir ağı (CNN) kullanılmaktadır. Transfer öğrenme yönteminin temelinde modellerin aynı alan ile görev veya farklı alan ile görev kombinasyonuna göre 4 farklı algoritma bulunmaktadır. Bu çalışmada aynı alan ile

görev ve farklı alan ile görev kombinasyonu seçilerek 2 farklı transfer öğrenme yöntemi test edilmekte ve sonuçlar karşılaştırılmaktadır. Çalışmanın temelinde U-Net mimarisi kullanılmaktadır. Önceden eğitilmiş ağ mimarileri U-Net modeliyle birleştirilerek iki farklı model tasarlanmıştır. Buna göre birinci tasarlanan model panoramik X-Ray'ler ile aynı alanda yer alan göğüs X-Ray görüntüleri kullanılarak eğitilen ve akciğer bölgesini segmente eden önceden eğitilmiş U-Net modeli ile oluşturulan evrişimli sinir ağı (CNN) mimarisidir. Panoramik X-Ray görüntüleri ile farklı alanda yer alan ImageNet veri seti kullanılarak eğitilen ve farklı objeleri sınıflandıran önceden eğitilmiş InceptionResNet-v2 modeli ile oluşturulan evrişimli sinir ağı (CNN) ikinci tasarlanan modeldir. Panoramik X-Ray görüntüleri veri artırma yöntemi kullanılarak çoğaltılmaktadır. Modellerin eğitimi sırasında kullanılan görüntülerin boyutları ve formatları ön işlem sürecinden sonra eşitlenmektedir. Tasarlanan modellerin eğitim sürecinde erken sonlandırma yöntemiyle veri ezberleme durumu kontrol edilmektedir. Elde edilen sonuçların anlamlandırılması için zar katsayısı performans metriği seçilerek iki model test edilmektedir.

1.3 Tezin Akışı

Tezin geri kalanı şu şekilde organize edilmiştir: Bölüm 2'de, derin öğrenme (DL), transfer öğrenme yöntemi ve tasarımdaki evrişimli sinir ağları (CNN) ile ilgili teorik bilgiler verilmiş ve otomatik diş segmentasyonu için önerilen ağ mimarisinin detayları açıklanmıştır. Bölüm 3'te ise kullanılan veri setleri çeşitleri, modellerin test edilmesi için ayarlanan performans metrikleri, eğitim parametrelerinin neler olduğu ve modelin eğitim aşamaları detaylı olarak sunulmaktadır. Son olarak Bölüm 4'te ise elde edilen sonuçlar, potansiyel araştırma ve gelecek çalışmalara ilişkin tartışmalar ile tez sonuçlandırılmaktadır.

2. METODOLOJİ

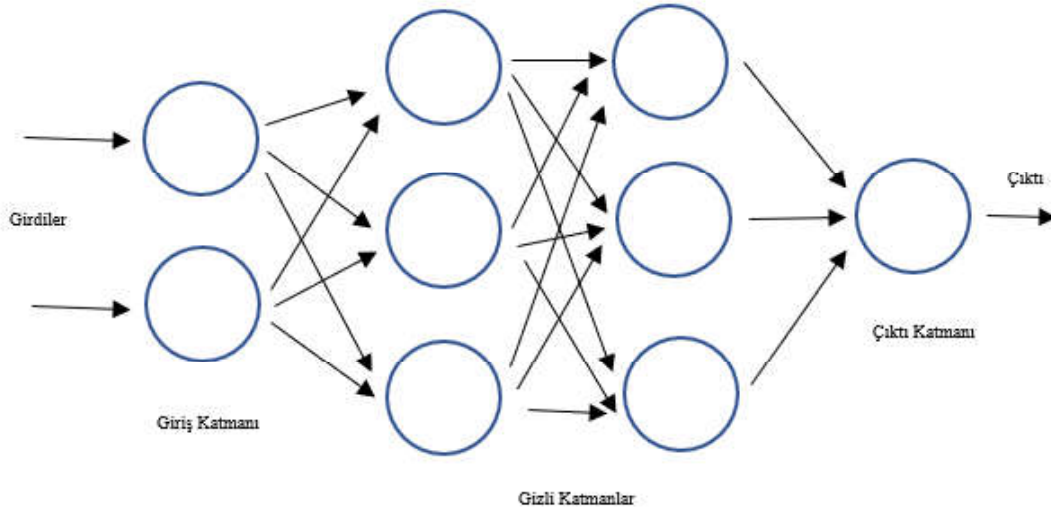
Bu bölümde ilk olarak derin öğrenmenin (DL) ne olduğu, hangi amaçla kullanıldığı ve tercih edildiği uygulama alanları açıklanmaktadır. Derin öğrenme (DL) modellerinden olan ve bu çalışmada kullanılan U-Net [16] ve InceptionResNet-v2 [26] mimarileri tanıtılmaktadır. Bu mimariler ile ilgili teorik bir altyapı oluşturulduktan sonra tezin oluşumunda esas alınan transfer öğrenme yöntemi, neden ve nasıl kullanıldığı açıklanarak tezin genel mantığı belirtilmektedir. Bu temeller ışığında tasarlanan ağ mimarisinin yapısı detaylı bir şekilde verilmektedir.

2.1 Derin Öğrenme (DL)

Derin öğrenme (DL) temelde birden fazla katman içeren ve görüntüdeki objeleri sınıflandırma, cümleleri tercüme edebilme, trafikteki yol işaretlerini algılama benzeri insanların günlük hayatta yapabildikleri görevleri bilgisayarlara öğreten makine öğrenmenin (ML) bir alt kümesidir. İnsan beynine benzeyen yapıya sahip olduğu için derin sinir ağları olarak da adlandırılmaktadırlar. Derin sinir ağları, verilerdeki özellikleri doğru bir şekilde algılamak, sınıflandırmak veya tanımlamak için veri girdileri, ağırlıklar ve sapma etkisinin kombinasyonu ile çalışmaktadır. Derin öğrenme (DL) metin, ses veya görüntüler gibi verileri işleyerek ilgili öznelik çıkarma işlemini otomatikleştirmektedir. Örneğin elimizde birçok farklı markaya ait araba görüntülerinin olduğunu ve her bir markaya ait arabayı sınıflandırmak istediğimizi varsayalım. Derin öğrenme (DL) algoritmaları bu görev için hangi özelliklerin daha önemli olduğuna karar vererek çıkarım yapabilmektedir. Her görüntü sonrası algoritma kendini uyarlayarak yeni bir görüntü üzerinden ilgili markaya ait arabayı yüksek başarı oranıyla sınıflandırabilmektedir.

Derin bir sinir ağının giriş ve çıktı katmanlarına görünür katmanlar denir. Giriş katmanı, derin öğrenme modelinin verileri işlemek üzere aldığı aşamadır ve çıktı katmanı, modelin tahminini yaptığı son aşamadır (Şekil 2.1). Derin sinir ağları,

yapılan tahminleri optimize etmek için her biri bir önceki katman üzerine inşa edilen birbirine bağlı birden çok gizli katmanından oluşmaktadır. Giriş katmanından çıktı katmanına doğru yapılan hesaplamalara ileriye doğru yayılma denir. Yapılan tahminlere göre modelin ilgili parametrelerini güncellemek için çıktı katmanından giriş katmanına doğru yapılan hesaplamalara geriye doğru yayılma denmektedir. Birlikte, ileriye ve geriye doğru yayılma, bir sinir ağının tahminler yapmasına ve elde ettiği tahminlere göre tahmin hatalarını düzeltmesine izin vermektedir. Zamanla, algoritma giderek daha doğru çalışmaktadır.



Şekil 2.1 : Basit bir derin sinir ağı yapısı.

Bununla birlikte, girdi, katman ve çıktı sayısına göre derin öğrenme algoritmalarının karmaşıklığı değişkenlik göstererek farklı işlevler için oluşturulmuş modeller bulunmaktadır. Örneğin, bilgisayarla görü ve görüntü sınıflandırma uygulamalarında kullanılan evrişimli sinir ağları (CNN), bir görüntüdeki özellikleri ve kalıpları algılayarak nesne algılama veya tanıma gibi görevleri yapabilmektedir. Tekrarlayan sinir ağları (RNN) sıralı veya zaman serisi verilerinden yararlandığı için tipik olarak doğal dil ve konuşma tanıma uygulamalarında kullanılmaktadır.

Derin öğrenme (DL) modelleri birçok uygulama alanlarına sahiptir. Örneğin yüz tanıma ile telefonlarda, otomatik sürüş, nesne tanımlama ve sınıflandırma ile arabalarda, ses tanımlama ve tercümesi ile akıllı ev sistemlerinde farklı amaçlar doğrultusunda kullanılmaktadırlar.

2.2 Evrişimli Sinir Ağları (CNN)

Bir evrişimsel sinir ağı (CNN), girdi olarak çeşitli görüntüler alıp bu görüntüdeki nesnelere önem değeri atayarak görüntüdeki objeyi bir diğerinden ayırt edebilen özellikle bilgisayar görüşü alanlarında oldukça yaygın kullanılan bir derin öğrenme (DL) yapısıdır. Şekil 2.1'dekine benzer bir giriş katmanı, bir veya daha fazla gizli katman ve bir çıktı katmanı içeren düğüm katmanlarından oluşmaktadır. Evrişimli sinir ağları (CNN) bir görüntü içindeki kalıpları tanımlamak için doğrusal cebirden, özellikle matris çarpımından yararlanarak görüntü sınıflandırma ve nesne tanıma görevlerini daha etkili gerçekleştirebilmektedir.

Evrişimli sinir ağları temel olarak 3 farklı katman ile oluşturulmaktadır. Bunlardan ilki çıktı katmanında yer alan her bir düğümün bir önceki katmanda yer alan düğüme bağlandığı dolayısıyla tamamen birbirine bağlı bir akış söz konusu olduğu Tamamen Bağlı Katman (FCL) yapısıdır. Bu katman kendinden önceki katmanların elde ettiği özniteliklere göre sınıflandırma işlemini gerçekleştirmektedir. Bir diğer katman yapısı ise Havuzlama Katmanı'dır. Havuzlama katmanı kendinden önceki katman çıkışından elde edilen matris boyutunu azaltarak kendinden sonraki katmana iletmektedir. Boyut azaltma işlemini iki şekilde gerçekleştirmektedir. Maksimum havuzlama ile belirlenen havuzlama boyutu kadar girdi alanındaki maksimum değeri seçilerek o bölge için atama gerçekleştirmektedir. Böylece girdi matrisinin boyutu indirgenmektedir. Ortalama havuzlama ise diğer bir havuzlama katman yöntemi olmakla beraber maksimum değer yerine ortalama olarak ilgili alıcı alana değer atamaktadır. En çok kullanılan katmanlardan biri ise Evrişimli Katman'dır. CNN'de bir evrişim, girdi matrisi ile belirli bir boyuttaki filtrenin (çekirdek) çarpılmasını içeren doğrusal bir işlemdir. Filtreler ise verilen girdilerdeki önemli özelliklerin anlaşılması ve tanımlanmasına yardımcı olan matrislerdir. Kullanılan filtrelerin boyutları genellikle 1x1, 3x3, 5x5 boyutunda olmaktadır. Bu bilgiler ışığında elimizde nxn boyutunda bir girdi olduğunu düşünelim. Seçilen filtre boyutu ise 3x3 olduğunda elde edilecek çıktının boyutu Eşitlik (2-1) ile verilen denklem ile hesaplanmaktadır:

$$boyut_{çikti} = \frac{(I-F+2P)}{s} + 1 \quad (2.1)$$

Girdi matrisinin genellikle kare olduğu varsayıp boyutu I ile belirtilmektedir. F , kullanılan filtrenin boyutunu ifade etmektedir. P ile adlandırılan sıfır dolgu bileşeni filtrenin boyutunun girdi matrisine uymadığı ve girdi matrisinde açıkta kalan değerler olduğunda kullanılmaktadır. S ile adlandırılan adım bileşeni ise filtrenin ne kadar adımla girdi matrisi üzerinde gezdirilmesinin belirlenmesine yardımcı olmaktadır. Bu değerler ışığında çıktı görüntüsünün boyutu elde edilmektedir. Girdi matrisiyle filtrenin nokta çarpımından sonra elde edilen matrise özellik haritası da denmektedir. Evrişim katmanı ile birlikte bir nevi alt örnekleme yapılırken tersine evrişim katmanı ile de tam tersi bir nevi üst örnekleme yapılmaktadır. Her bir evrişim katmanından sonra genellikle aktivasyon fonksiyonu kullanılmaktadır. Aktivasyon fonksiyonu girdi değerlerini belirli bir aralığa eşleyerek çıktı üretilmesini sağlarlar. Aktivasyon fonksiyonları modelin resim, metin veya sesler gibi verilerdeki karmaşık kalıpları öğrenmesini sağlamak için doğrusal olmayan bir yapı sunmaktadır. Farklı matematiksel yöntemleri kullanan aktivasyon fonksiyonlarından biri olan Doğrultulmuş Doğrusal Birim (ReLU) fonksiyonu doğrusal olmayan bir matematiksel ifade kullanmaktadır. ReLU aktivasyon fonksiyonu negatif girdi değerlerini 0'a eşitlerken pozitif değerlerini olduğu gibi geçirmektedir (Şekil 2.2).

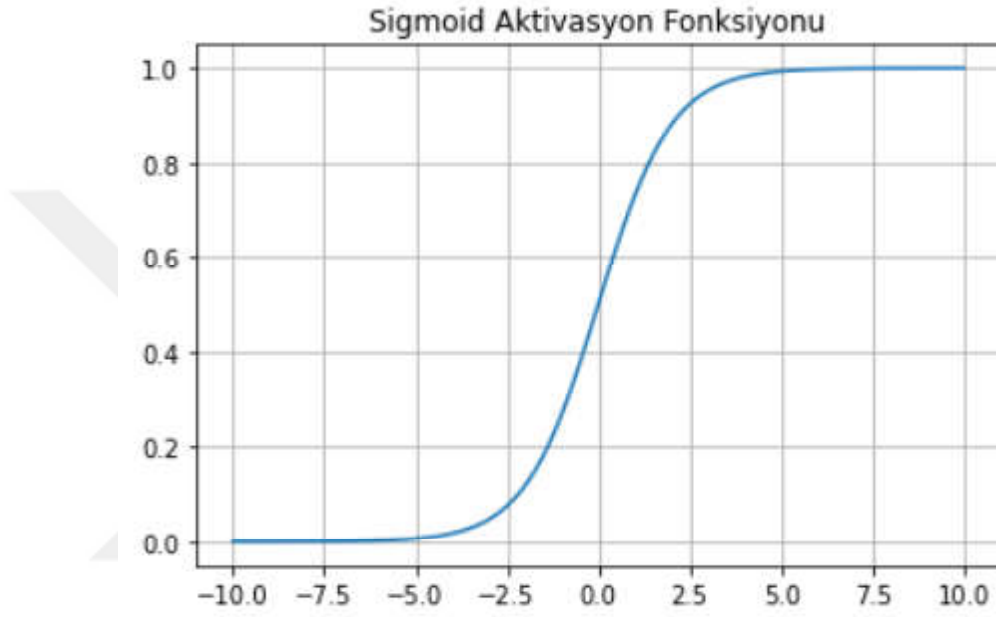
$$f(x) = \max(0, x) \quad (2.2)$$



Şekil 2.2 : Doğrultulmuş doğrusal birim (ReLU) aktivasyon fonksiyonu.

Diğer bir doğrusal olmayan aktivasyon fonksiyonu ise sigmoid fonksiyonudur (Şekil 2.3). Girdi değerleri ne olursa olsun (0,1) aralığına eşleştirmektedir. Tek sınıflı çıktılarda daha çok tercih edilmektedir.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2.3)$$



Şekil 2.3 : Sigmoid aktivasyon fonksiyonu.

Seçilen görev, girdi boyutu, çıktıda yer alan sınıf sayısı gibi parametrelere bağlı olarak bu katmanlar kullanılarak ağ mimarisi tasarlanmaktadır.

2.2.1 U-Net mimarisi

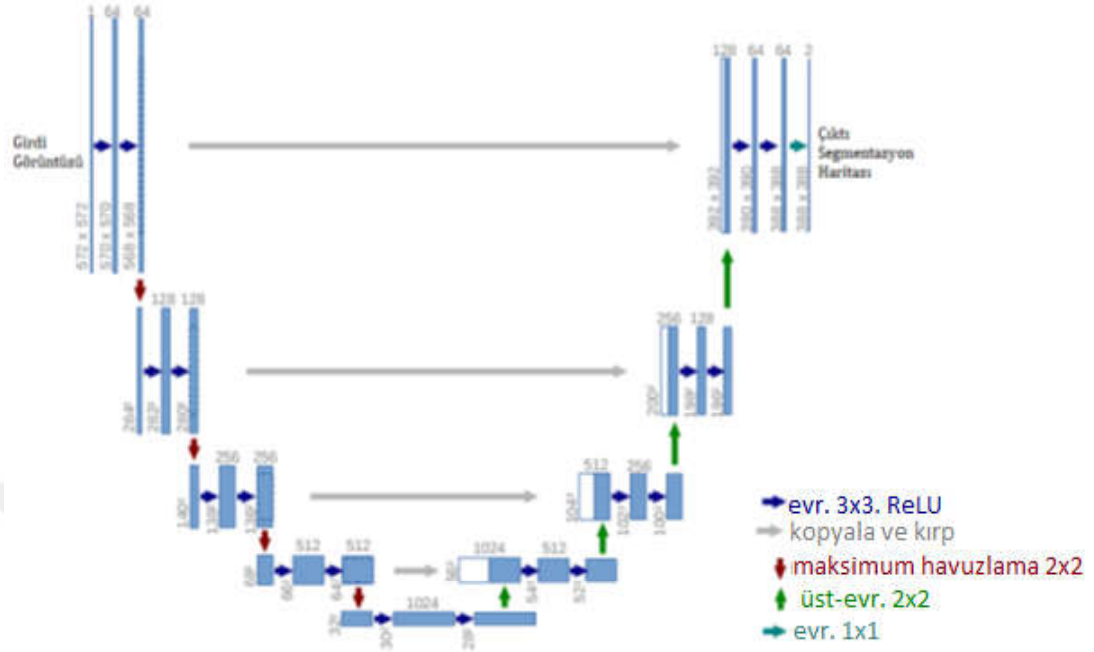
İlk olarak 2015 yılında biyomedikal görüntü segmentasyonu için tasarlanmıştır. Genel olarak içeriği yakalamak için tasarlanan kodlayıcı yolu ve hassas konum tespiti sağlayan simetrik kod çözücü yolu olmak üzere iki temel bölümden oluşmaktadır. Şekil 2.4 ile orijinal U-Net mimarisi [16] gösterilmektedir. Şekilden anlaşılacağı üzere birbirinin simetriği iki bölümden dolayı “U” şekline benzediği için ismini buradan almaktadır.

Kodlayıcı yolu alt örnekleme yaparak ilerlemektedir. İlk olarak evrişim katmanından geçen girdi görüntüsünün boyutu maksimum havuzlama katmanı ile işleme sokularak

düşürülmektedir. Bu işlemlerden sonra düşük çözünürlüğe sahip kodlanmış bir görüntü ile yüksek çözünürlüklü öznitelik haritası elde edilmektedir [16]. Uygulanan bu işlemler aslında görüntüdeki ayırt edici özelliklerin kodlanarak farklı seviyelerdeki öznitelik kümeleri elde edilmesine olanak sağlamaktadır.

Kodlayıcı yolundan sonra yer alan evrişim katmanını takiben kod çözücü yolu üst örnekleme ve burada yer alan katmanların kodlayıcı yolunda yer alan ilgili katmanlarla birleştirilme mantığıyla ilerlemektedir. Havuzlama operatörleri üst örnekleme operatörleriyle yani ters evrişim katmanlarıyla değiştirilmiştir [16]. Böylelikle çıktı görüntüsünün çözünürlüğü arttırılmaktadır. Konumlandırma amacıyla kodlayıcı yolundan elde edilen yüksek çözünürlüklü öznitelik haritaları üst örneklenen çıktılarla birleştirilmektedir [16]. Üst örnekleme kısmı içerik bilgisini daha yüksek çözünürlüklü katmanlara iletmeyi sağlayan yüksek sayıda özellik kanalından oluşmaktadır [16]. Öznitelik haritasını girdi görüntüsünün orijinal boyutuna getirmek için hem ters evrişim katmanı hem de kodlayıcı yolundaki karşılık gelen katmanlarla birleştirme işlemiyle aynı boyutu karşılarken daha yüksek çözünürlüklü piksel alanı elde edilmektedir. Dolayısıyla buradaki temel amaç üst örnekleme yaparken aynı zamanda bir sonraki aşamada yer alan evrişim katmanlarıyla öznitelik kümelerini daha iyi öğrenebilmek için kodlayıcı yolundan gelen öznitelik haritalarını üst örneklenmiş özelliklerle birleştirmektir.

Özetle kodlayıcı yolu, her biri bir doğrultulmuş doğrusal birim (ReLU) ve alt örnekleme için maksimum havuzlama işlemi tarafından takip edilen evrişim katmanından oluşmaktadır. Her alt örnekleme adımında, özellik kanallarının sayısı iki katına çıkarılmaktadır. Kod çözücü yolu, öznitelik haritasının üst örnekleme, özellik kanallarının sayısını yarıya indiren bir ters evrişim katmanı, kodlayıcı yolundan karşılık gelen öznitelik haritasıyla birleştirme ve evrişim katmanından oluşmaktadır. Son katmanda, özellik vektörünü istenen sayıda sınıfa eşlemek için 1×1 evrişim katmanı kullanılmaktadır [16]. Dolayısıyla girdi ve çıktı görüntüleri aynı boyutta olmaktadır. Bu model diğer mimarilere kıyasla sondan sona segmentasyonu öğrenmektedir. Her bir piksel seviyesinde sınıflandırma yaptığından sınırları hassas bir şekilde konumlandırabilir ve ayırt edebilmektedir.

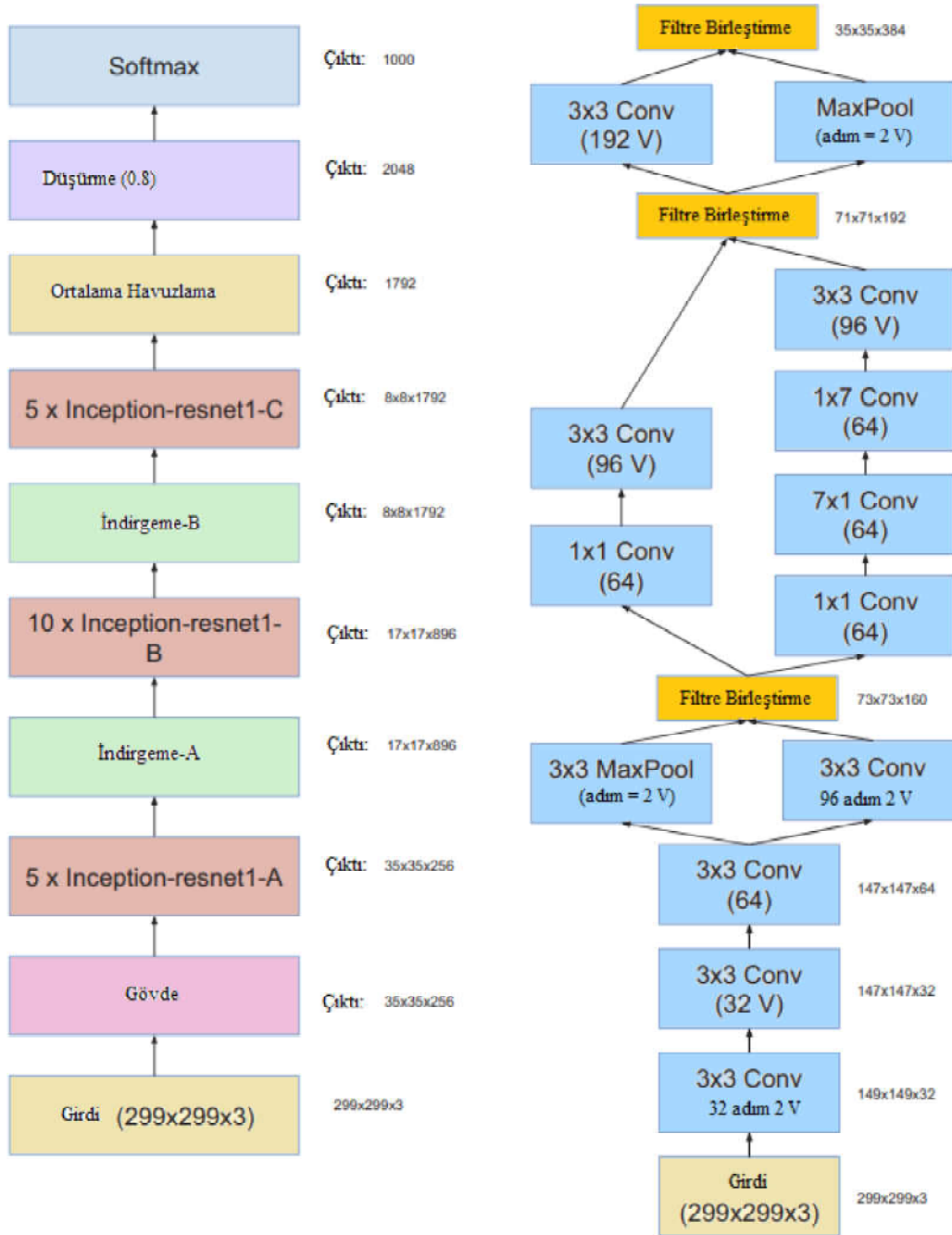


Şekil 2.4 : Orijinal U-Net mimarisi [16].

2.2.2 Inception-ResNet-v2 mimarisi

Görüntüde yer alan ayırt edici bölgeler oldukça farklı değişime sahiptir. İlgili bilgiyi içeren bölgenin konumundaki büyük değişimlerden kaynaklı evrişim katmanı için doğru filtre boyutunu seçmek oldukça zordur. Daha büyük bir filtre daha geniş alanlara yayılmış ilgili bölge için kullanılırken daha küçük filtreler ise daha küçük alanlara yayılmış ilgili bölgeler için kullanılmaktadır. İlk tasarlanan Inception mimarileri genellikle derin yapıdadırlar [26]. Dolayısıyla yeni versiyonlarda çeşitli boyutlardaki filtrelerin birleşimiyle daha geniş bir yapı elde edilmiştir. Inception ve kalıntı bloklarının birleşiminden oluşan yeni Inception versiyonunda (Şekil 2.5) normalde kullanılan filtre birleştirme aşaması kalıntı bağlantısıyla değiştirilmiştir [26].

Yeni versiyonda kalıntı Inception bağlantıları için orijinal Inception mimarisinden daha ucuz Inception blokları kullanılarak daha az hesaplama zamanı elde edilmektedir (Şekil 2.6) [26]. Her Inception bloğunu filtre genişlemesi adı verilen aktivasyonsuz 1x1 boyutundaki evrişim katmanı takip etmektedir [26].

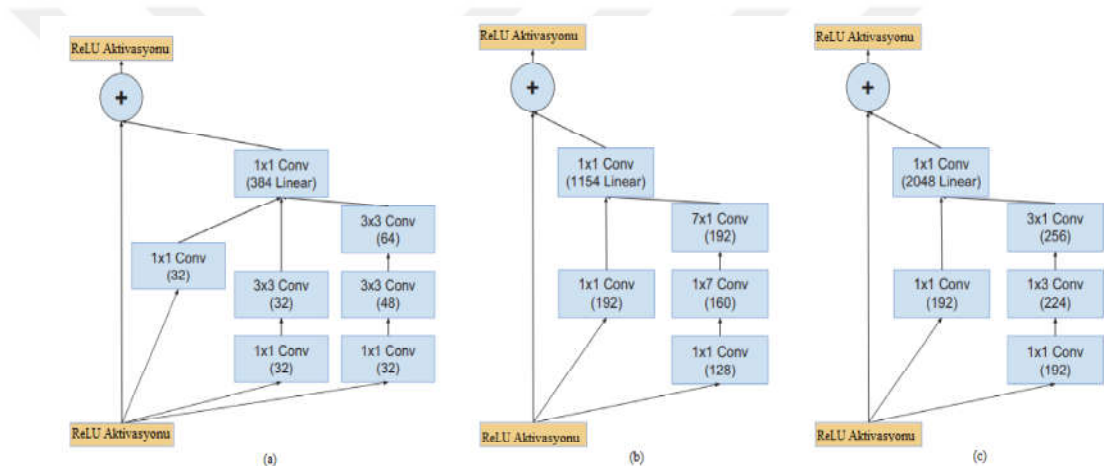


Şekil 2.5 : Inception-ResNet-v2 mimarisinin şeması [26].

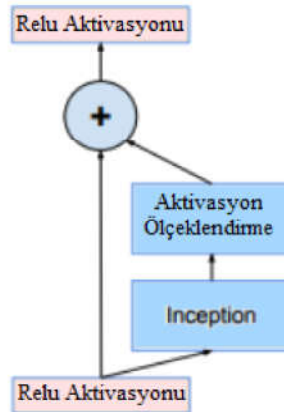
Kalıntı bağlantılarının çalışabilmesi için ilgili katmanın girdi ve çıktı boyutu aynı olmalıdır. İlgili katman çıktısının bir sonraki katmanın girdi derinliğiyle aynı olması adına filtre genişletme ile toplama işleminden önce filtrenin boyutu

ölçeklendirilmektedir. Inception bloklarının içinde yer alan havuzlama katmanı kalıntı bağlantılarıyla değiştirilmiştir.

Çalışma [26] ile kullanılan filtrelerin sayısı 1000'i geçerse kalıntı değişkenlerinin dengesizleştiği ve tüm ağıın eğitimin ilk aşamalarında öldüğü belirtilmiştir. Yani ağıın son ortalama havuzlama katmanından önce sadece sıfır değerini ürettiği anlaşılmıştır. Bu problem öğrenme oranının düşürülmesi ya da fazladan grup-normalizasyon aşaması eklenmesiyle çözülememiştir [26]. Bunun için kalıntı bağlantıları önceki katman aktivasyonuna eklenmeden önce ölçeklendirilmiştir (Şekil 2.7) [26].



Şekil 2.6 : (a) 35x35 ızgara modülü şeması, (b) 17x17 ızgara modülü şeması, (c) 8x8 ızgara modülü [26].



Şekil 2.7 : Kalıntı bağlantılarının ölçeklendirme aşaması [26].

2.3 Transfer Öğrenme

Normal bir derin öğrenme (DL) modelinin eğitim süreci belirli bir görev için ilgili alandan elde edilen verilerle seçilen algoritmanın eğitilmesi ve algoritmanın görmediği bir veri ile test edilmesinden oluşmaktadır. Burada bahsedilen görev nesne sınıflandırma veya yüz tanımlama gibi işlemler olabilirken alan ise modelin eğitimi için kullanılan veri setinin elde edildiği ortamı ifade etmektedir.

Bir modelin görmediği veri setinde de iyi çalışması adına geniş veri setleriyle eğitilmesi gerekmektedir. Böylelikle modelin geliştirilerek gerçek senaryolarla karşılaştığında da daha doğru tahminler yapabilmesi sağlanmaktadır. Bu doğrultuda tasarlanan modeller, geniş çaplı veri setlerinin artmasıyla birlikte neredeyse insan performansına yaklaşan tahminler yaparak başarılı sonuçlar elde etmektedir. Fakat sınırlı ve geliştirilmemiş bir veri setiyle eğitilen model daha önce görmediği bir ortamda iyi çalışmamakta ve nasıl hareket edeceğini bilmemektedir. Bu tür durumlarda başarı yüzdesi düşük olmaktadır. Bu aşamada devreye transfer öğrenme yöntemi girmektedir.

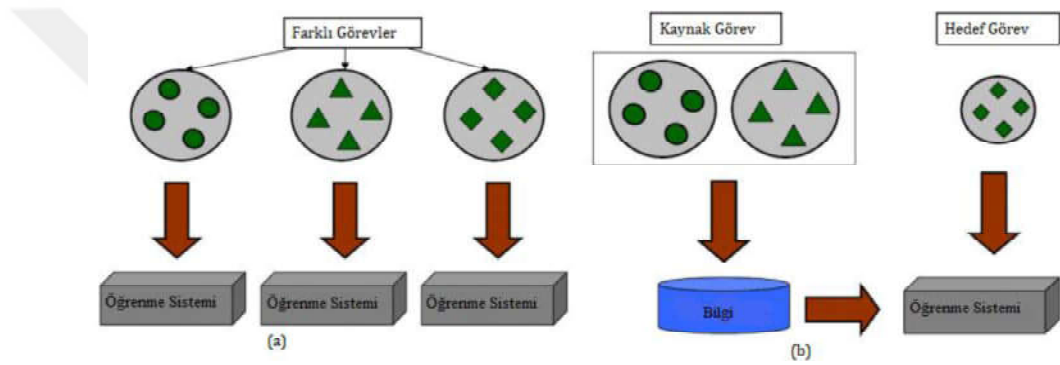
Transfer öğrenme, geniş bir veri setiyle eğitilmiş ve gerçek dünya senaryolarına karşı geliştirilmiş bir modelin öğrenmiş olduğu öznitelik haritalarının ve kazanımlarının yeni bir görev için kullanılmasına olanak sağlayan ve sıfırdan bir modelin eğitilmesi gerekliliğini ortadan kaldıran bir yöntemdir. Tekrar kullanılan kaynak modeline önceden eğitilmiş model denmektedir. Şekil 2.8 (a) ile klasik bir derin öğrenme (DL) süreci ve Şekil 2.8 (b) ile transfer öğrenme süreci gösterilmektedir. Kaynak görev için elde edilen bilgiler başka bir öğrenme sistemine aktarılarak hedef görev için kullanılmaktadır.

Transfer öğrenme yöntemi için iki önemli kriter bulunmaktadır: alan ve görev. Bu kriterler sırasıyla tanımlanırsa,

- Bir *alan*, D , iki bileşenden oluşmaktadır: özellik uzayı χ ve marjinal olasılık dağılımı $P(X)$ ($X = \{x_1, \dots, x_r\} \in \chi$). Örneğin bir modelin görevi belge sınıflandırma ise χ tüm kelime vektör uzayı ve X model eğitiminde

kullanılan belge örneğidir. Genel olarak, iki alan farklı olduğunda farklı özellik uzayına ya da farklı $P(X)$ değerine sahip olmaktadır [22].

- Belirli bir alanda, $D = \{ \chi, P(X) \}$, bir görev, T , iki bileşenden oluşmaktadır: etiket uzayı y ve eğitim verisinden öğrenilen, $x_i \in X$ ile $y_i \in y$ ikilisinden oluşan koşullu olasılık dağılımı $P(Y|X)$. Belge sınıflandırma için y , Doğru ve Yanlış olmak üzere tüm etiket uzayını barındırırken y_i ise etiketlerden birisidir.



Şekil 2.8 : (a) Klasik DL öğrenme süreci, (b) Transfer öğrenme süreci [27].

Tanımlardan yola çıkarak, sırasıyla kaynak bir alan, D_s , kaynak bir görev, T_s , hedef alan, D_t , hedef görev, T_t , olarak ifade edildiğinde transfer öğrenme $D_s \neq D_t, T_s \neq T_t$ kaynak alan ve görevden elde edilen bilgilerle D_t yer alan hedef koşullu olasılık dağılımının, $P(Y_t|X_t)$, öğrenimine olanak sağlamaktadır. Transfer öğrenme temelde dört farklı algorithmadan oluşmaktadır:

1. Kaynak ve hedef alanının özellik uzayı farklıdır: iki farklı dilde yazılmış belgelerden oluşan veri seti [27].
2. Kaynak ve hedef alanının marjinal olasılık dağılımı farklıdır: iki farklı konuda yazılmış belgelerden oluşan veri seti [27].
3. Kaynak ve hedef görevinin etiket uzayı farklıdır: kaynak görevi iki sınıfa sahipse hedef görevi beş sınıfa sahiptir [27].
4. Kaynak ve hedef görevinin koşullu olasılık dağılımı farklıdır [27].

Klasik derin öğrenme (DL) modelleri aynı kaynak alan ile göreve veya hedef alan ile göreve sahiptir. Transfer öğrenme yönteminde ise bu durum daha farklı işlemektedir.

- *Tümevarımsal Transfer Öğrenme* süreci kaynak ve hedef alanına bakmaksızın farklı kaynak ve hedef görevine sahiptir.
- *Transdüktif Transfer Öğrenme* süreci aynı kaynak ve hedef görevlere, farklı kaynak ve hedef alanlarına sahiptir.
- *Denetimsiz Transfer Öğrenme* süreci kaynak alan ve görevinin hedef alan ve görevi ile farklı olduğu durumlarda kullanılmaktadır.

Transfer öğrenme metodunda önceden eğitilmiş modeller iki farklı yöntem ile kullanılabilir. Öznitelik çıkarımı yöntemi ile önceden eğitilmiş modelin öğrenmiş olduğu kazanımlarının aktarılmasıyla yeni veri setlerinin özellikleri öğrenilebilmektedir. Tüm modelin yeniden eğitilmesine gerek olmamakla birlikte sadece belirli bir kısmın yeni modelle birleştirilerek eğitilmesi yeterli olmaktadır. İnce ayar yöntemi ise önceden eğitilmiş modelin donmuş haldeki üst katmanlarından bazılarının çözülerek yeni modelle birlikte eğitilmesidir.

Geçmişten günümüze geliştirilen geniş veri setleriyle eğitilen başarılı modellerin transfer öğrenme yöntemiyle tekrar kullanılması sınırlı sayıdaki verinin daha verimli kullanılmasını, modelin daha genelleştirilebilir olmasını, daha dirençli ve stabil çalışmasını sağlamaktadır.

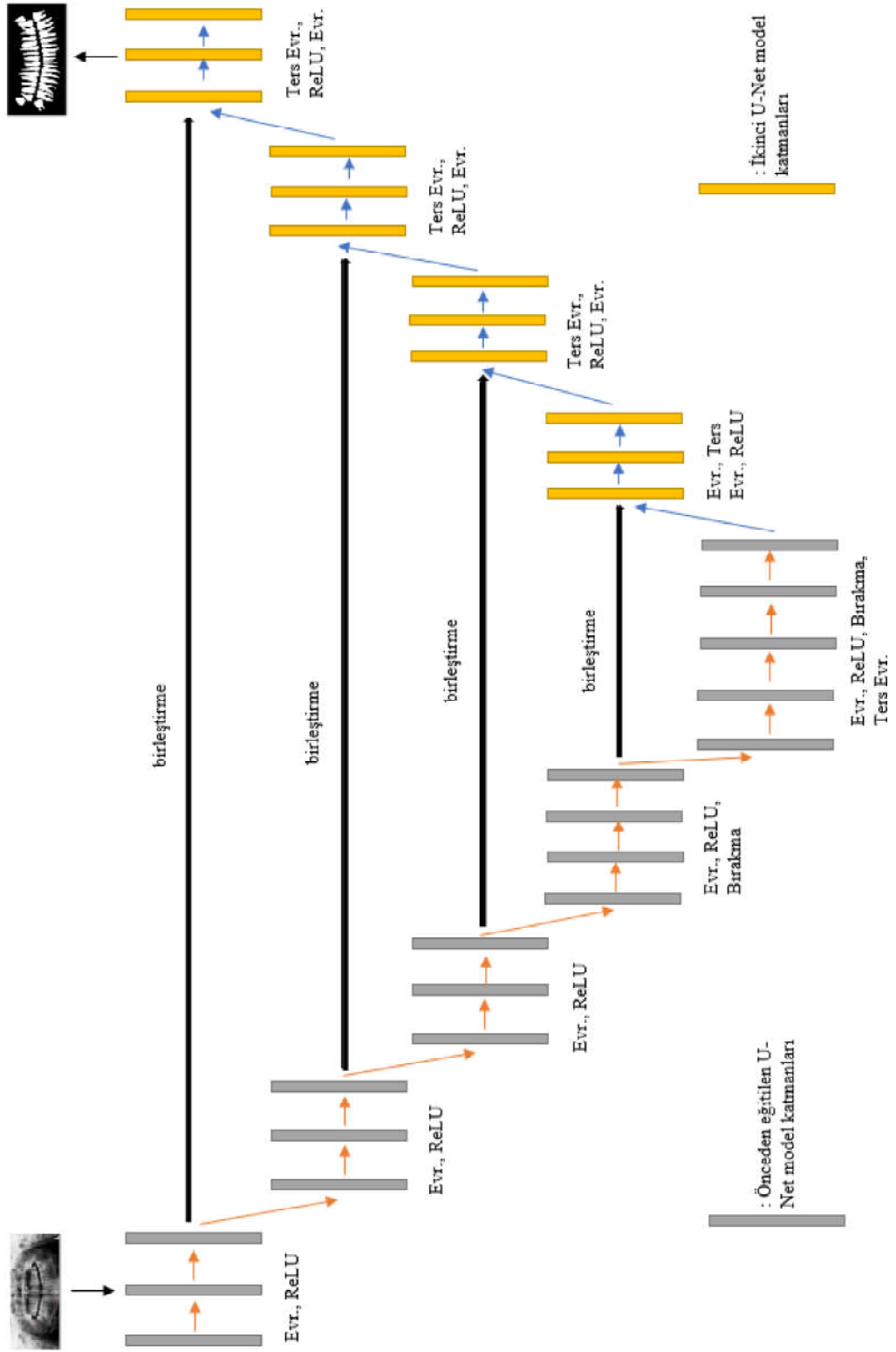
2.4 Önerilen Ağ Mimarisi

Bu çalışmada orijinal U-Net modeli baz alınarak iki farklı ağ mimarisi geliştirilmiştir. U-Net modelinin kod çözücü yolu iki ağ mimarisinde de değiştirilmeden alınmıştır. Kod çözücü yolunda sırasıyla 3x3 filtre boyutuna sahip ters evrişim katmanı, kodlayıcı yolundan ilgili katman ile birleştirme işlemi, 3x3 filtre boyutuna ve ReLU aktivasyon fonksiyonuna sahip evrişim katmanı kullanılmaktadır. Model çıktısı tek sınıf olarak ayarlandığından mimarinin son katmanında 1x1 filtre boyutu ve Sigmoid aktivasyon fonksiyonuna sahip son evrişim katmanı yer almaktadır.

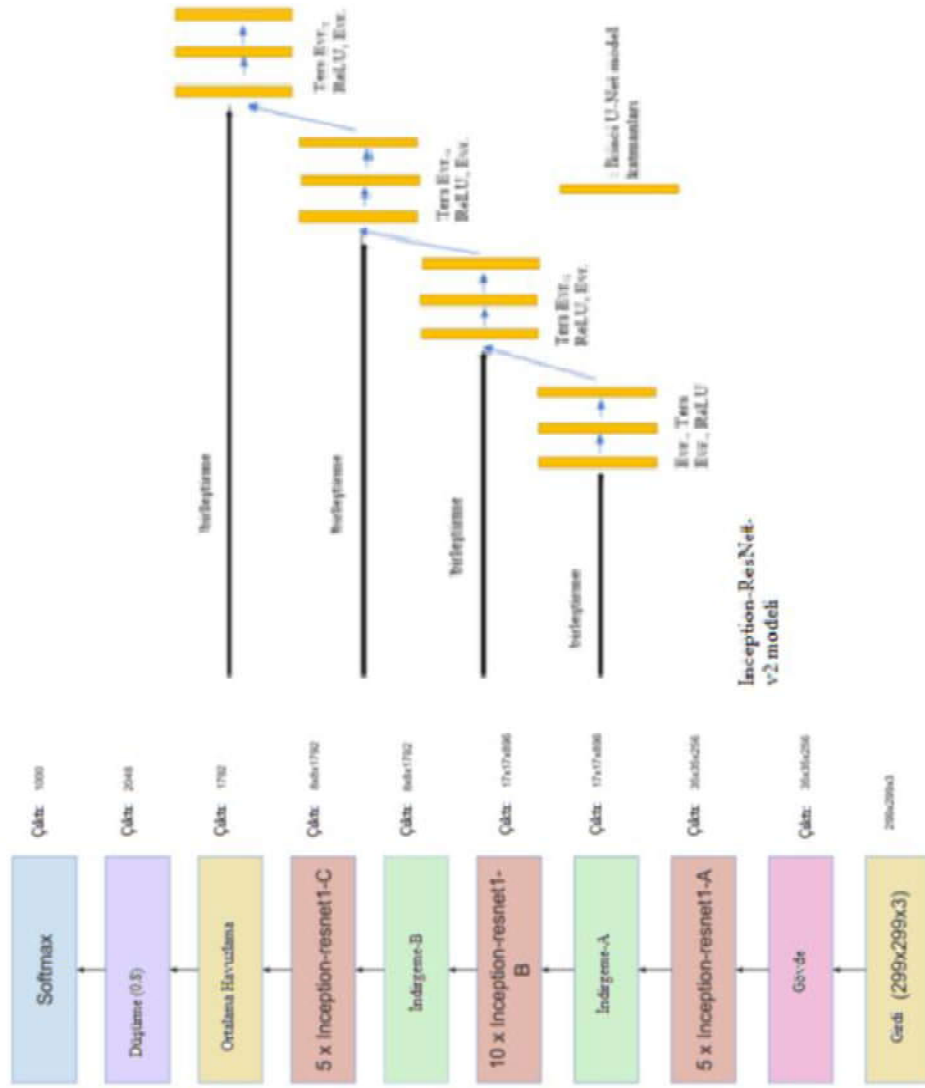
Klasik U-Net modelinin kodlayıcı yolundaki esas katmanların yerini ise önceden eğitilmiş modellerin ilgili katmanları almaktadır. Bu değişimlerle oluşturulan iki farklı mimari sırasıyla,

- **Double-UNet** modeli iki adet U-Net modelinin birleşiminden oluştuğu için bu şekilde isimlendirilmektedir. Transfer öğrenme metodunun uygulanabilmesi adına önceden eğitilmiş model olarak göğüs X-Ray veri setiyle eğitilen bir U-Net modeli kullanılmıştır. Önceden eğitilmiş U-Net modeli ve ikincil U-Net modelinin birleşimi klasik U-Net modeline uygun olarak yapılmaktadır. Şekil 2.9 ile elde edilen ağ mimarisinin şeması verilmektedir. Kodlayıcı yolunda sırasıyla 3x3 filtre boyutuna ve ReLU aktivasyonuna sahip evrişim katmanı ve 2x2 boyutunda maksimum havuzlama katmanı bulunmaktadır. Kodlayıcı ve kod çözücü yolunu birleştiren köprü klasik U-Net modeline uygun olarak 3x3 filtre boyutundaki ReLU aktivasyonuna sahip bir evrişim katmanı ile sağlanmaktadır.
- **InceptionResV2-Unet** modeli Inception-ResNet-v2 ve U-Net modellerinin birleşiminden oluştuğu için bu şekilde isimlendirilmektedir. Transfer öğrenme metoduna uygun olarak ImageNet veri setiyle önceden eğitilmiş InceptionResNet-v2 modelinin her katman sonrası elde edilmesi gereken çıktı boyutuna göre ilgili katmanları seçilerek ikincil U-Net modeliyle birleştirilmiştir. Şekil 2.10 ile elde edilen mimarinin şeması gösterilmektedir. Inception-ResNet-v2 modelinin tasarlandığı girdi boyutu 299x299'dur. Dolayısıyla her katman sonrası elde edilen çıktı boyutları tek sayı olduğu için farklıdır. Bu boyutları 256x256'a ve her katman sonrası U-Net modeline uygun çift sayılı çıktı boyutları alabilmek adına belirtilen oranda ilgili matrisin sonuna 0 ekleme işlemi gerçekleştiren iki boyutta çalışan ZeroPadding2D katmanı kullanılmıştır. Kodlayıcı ve kod çözücü yolunu birleştiren köprü Inception-ResNet-v2 modelinden alınmıştır.

İki model de girdi olarak RGB görüntülerle birlikte etiketlenmiş veri olarak karşılık gelen ikili maske görüntülerini almaktadır.



Şekil 2.9 : Double-UNet mimarisi.



Şekil 2.10 : InceptionResV2-UNet mimarisi.

3. DENEYLER

Bu bölümde ilk olarak tasarlanan modellerin eğitiminde kullanılan farklı türdeki veri setlerinden ve eğitim öncesi görüntüler üzerinde uygulanan ön işlemlerden bahsedilmektedir. Daha sonra modelleri test etmek için seçilen performans metrikleri tanıtılmaktadır. Modellerin eğitiminden önce hangi parametrelerin nasıl ayarlandığı detaylandırılmaktadır. Son olarak yapılan tüm ayarlamalar sonra tasarlanan modellerin eğitim süreci anlatılmaktadır.

3.1 Veri Seti

Bu çalışmada aktif olarak iki olmak üzere toplamda üç farklı veri seti kullanılmaktadır. Tezin amacı diş segmentasyonu olduğu için temel veri seti panoramik X-Ray görüntüleridir. İkinci veri seti göğüs X-Ray görüntüleri ve üçüncü veri seti ise ImageNet [28] veri setidir. İlk iki veri seti akciğer ve diş bölgelerinin segmentasyonunda kullanılırken üçüncü veri seti görüntüde yer alan nesnelere sınıflandırılmasında kullanılmaktadır.

3.1.1 Panoramik X-Ray veri seti

Toplamda 131 adet panoramik X-Ray görüntüsü anonim olarak farklı hastalardan elde edilmiştir. Şekil 3.1 (a)'da tüm dişlere sahip bir hastanın ve Şekil 3.1 (b)'de eksik dişlere sahip bir hastanın X-Ray görüntüsü gösterilmektedir.

Tüm diş görüntüleri model eğitiminden önce iki aşamalı ön işlemden geçirilmiştir. İlk olarak bütün görüntülerin boyutlarının Çizelge 3.1'de belirtildiği gibi birbirinden farklı ve büyük boyutlarda olduğu anlaşılmıştır. Ayrıca kullanılan ağ mimarilerinin daha küçük boyuttaki görüntülerle daha başarılı çalıştığı ve daha az işlemci gücüne ihtiyaç duyduğu gözlenmiştir. Bu sebeple tüm diş görüntülerinin boyutları 256x256 olacak şekilde yeniden ayarlanmıştır. İkinci aşama olarak veri setinin boyutu az olduğu için modellerin eğitiminden önce görüntüler yatay olarak çevrilerek görüntü

sayısı iki katına çıkartılmıştır. İki aşamalı ön işlem süreci yazılım ortamında otomatik olarak gerçekleştirilmiştir.



(a)



(b)

Şekil 3.1 : (a) Tüm dişlere sahip hastanın X-Ray görüntüsü, (b) Eksik dişlere sahip hastanın X-Ray görüntüsü.

Çizelge 3.1 : Ön işlem öncesi ve sonrası örnek diş görüntü boyutları.

Orijinal Boyut	Son Boyut
1158x360	256x256
1269x571	256x256
1545x767	256x256
1622x802	256x256
.....
2751x1221	256x256
3227x1459	256x256

Ön işlemlerden sonra GIMP [29] adı verilen görüntü manipülasyon uygulaması kullanılarak model eğitiminde kullanılacak etiketli veriler bir diğer ifade ile maske görüntüleri oluşturulmuştur. Sırasıyla uygulanan işlemler aşağıda verilmektedir,

- a) Panoramik X-Ray görüntüsü uygulamaya yüklenir.
- b) Görüntü siyah bir katman üstüne bindirilir.
- c) Esas görüntü üzerinden diş bölgesi seçilir.
- d) Seçilen diş bölgesi beyaz renk ile belirtilir.
- e) Esas görüntü ve siyah katman birleştirilerek tek kanallı görüntü olarak kaydedilir.

Bu şekilde 256x256x3 boyutunda görüntü ve 256x256x1 boyutunda maskeler *.png formatında kaydedilerek kullanılmıştır. Şekil 3.2 (a)'da model girdi görüntüsü ve Şekil 3.2(b)'de karşılık gelen maske görüntüsü verilmektedir.



(a)

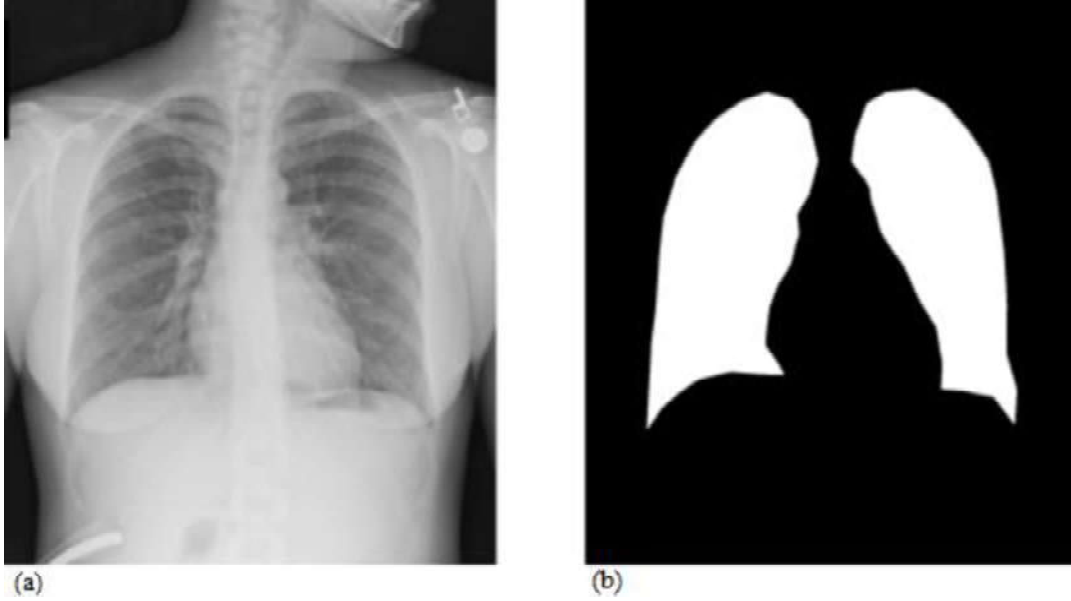


(b)

Şekil 3.2 : (a) Model girdi görüntüsü, (b) Maske görüntüsü.

3.1.2 Göğüs X-Ray veri seti

İkinci veri seti önceden eğitilmiş U-Net modelinin eğitimi için kullanılan göğüs X-Ray görüntüleridir. Göğüs X-Ray görüntüleri ChestX-Ray8 [30] veri tabanından elde edilmiştir. ChestX-Ray8 halka açık olarak yayınlanan 32,717 farklı hastadan alınan toplam 108,948 adet göğüs X-Ray görüntüsüyle oluşturulmuş medikal görüntü veri tabanıdır [30]. İlgili modelin eğitimi için bu görüntüler içerisinde rastgele 2505 görüntü seçilmiştir. Görüntülerin tümü 1024x1024 olarak aynı boyuttadır. İşlemci gücünü azaltmak adına görüntülerin boyutu 256x256 olacak şekilde yeniden ayarlanmıştır. Modelin eğitimi için kullanılacak etiketli veriler için dış maske görüntüleri oluşturulurken izlenen adımlar uygulanmıştır. Tüm işlemler sonrasında 256x256x3 boyutunda görüntüler ve 256x256x1 boyutunda maskeler elde edilmiştir. Şekil 3.3 (a)'da bir göğüs X-Ray görüntüsü ve Şekil 3.3 (b)'de maske görüntüsü gösterilmektedir.



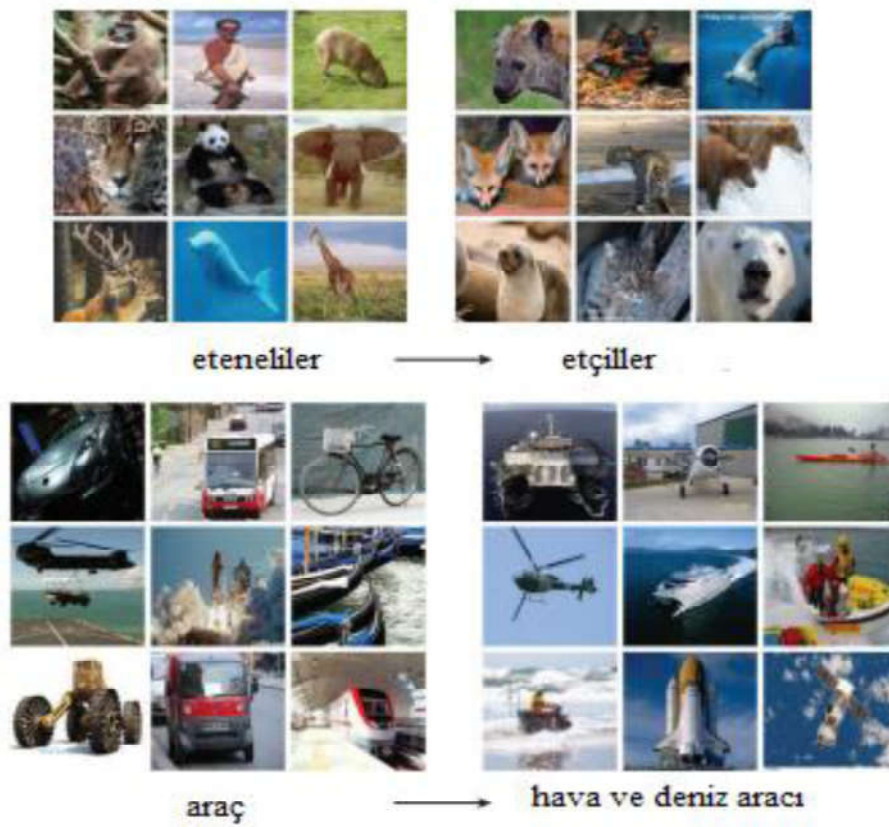
Şekil 3.3 : (a) Girdi görüntüsü, (b) Maske görüntüsü.

3.1.3 ImageNet veri seti

Son veri seti önceden eğitilmiş Inception-ResNet-v2 modelinin eğitiminde kullanılan ImageNet [28] veri tabanıdır. ImageNet, WordNet hiyerarşisine göre organize edilmiştir [28]. WordNet'te birden fazla kelime veya kelime öbeği ile tanımlanan her anlamlı kavram, eş anlamlı küme olarak adlandırılmaktadır [28]. ImageNet'te, her bir

eş anlamlı kümeyi ifade etmek için belirli bir sayıda görüntü kullanılmaktadır. Tamamlanma süreci devam eden ImageNet veri setinin WordNet hiyerarşisindeki kavramların çoğu için on milyonlarca temiz etiketlenmiş ve sıralanmış görüntü sunacağını beklenirken şu an toplamda 14.197.122 resim barındırmaktadır [28]. Şekil 3.4 ile veri setinde bulunan eş anlamlı kümelerden örnek olarak iki tanesi gösterilmektedir.

Halka açık olarak yayınlanan bu veri setiyle birçok model eğitilmiştir. Bunlardan biri olan Inception-ResNet-v2 modeli kullanıma hazır halde kütüphane olarak sunulmaktadır. Dolayısıyla bu çalışmada ImageNet veri seti kullanılarak evrimsel ağ mimarisi eğitimi söz konusu değildir.



Şekil 3.4 : Memeli ve araç alt-ağaçlarından örnek görüntüler [28].

3.2 Performans Metrikleri

Derin öğrenme (DL) modelleri genellikle doğruluk, F1 skoru, kesinlik gibi metrikler ile test edilerek algoritmaların yaptığı tahminlerin gerçek değerlere ne kadar

yakınsadığı ölçülmektedir. Bu çalışmada ise doğruluk parametresinin yanı sıra zar katsayısı metriği seçilerek modellerin performansları test edilmiştir. Ama modellerin başarısının yorumlanması için temel metrik zar katsayısı olarak seçilmiştir.

Zar katsayısı, iki örneğin benzerliklerini ölçmeye yarayan bir yöntemdir. 1945 yılında L.R. Dice [31] ve 1948 yılında T. Sørensen [32] tarafından geliştirilmiştir. Eşitlik (3.1) ile ifade edilen tahmin ve gerçek görüntüsünün toplam piksel sayısına bölünen tahmin ve gerçek görüntüsünün kesişim alanıdır. Bu katsayının değerinin 1.0'a yakın olması, tahmin ve gerçek görüntülerinin benzerliğinin yüksek olduğunu belirtmektedir.

$$zar = \frac{2 * |Maske_{gerçek} \cap Maske_{tahmin}|}{|Maske_{gerçek}| + |Maske_{tahmin}|} \quad (3.1)$$

Denklemden $Maske_{gerçek}$ ile gerçek maske görüntüleri, $Maske_{tahmin}$ ile modelin çıktısı olan tahmin maske görüntülerini ifade etmektedir.

Piksel doğruluğu, model tahmininin gerçek değerine ne kadar yakın olduğunun bir ölçüsüdür. Eşitlik (3.2) ile verilen denklemde, TP=doğru pozitif, TN=doğru negatif, FP=yanlış pozitif, FN=yanlış negatif değerlerini ifade etmektedir. Gerçek bir pozitif, gerçek doğrulukta verilen sınıfa ait olduğu doğru bir şekilde tahmin edilen bir pikseli temsil eder, aksine gerçek bir negatif, verilen sınıfa ait olmadığı doğru olarak tanımlanan bir piksel anlamına gelir.

$$dogruluk = \frac{TP + TN}{TP + FP + TN + FN} \quad (3.2)$$

Kayıp fonksiyonu modelin performansını ölçmek adına kullanılan bir diğer parametredir. Kullanılan modelin gerçek çıktısı ile tahmin çıktısı arasındaki uzaklığı hesaplayan fonksiyondur. Çapraz-entropi, ortalama mutlak hatası (MAE), ortalama kare hatası (MSE) gibi çeşitli kayıp fonksiyonları kullanılmaktadır. Bu çalışmada ise zar katsayısının negatif versiyonu (kayıp=-zar) kayıp fonksiyonu olarak kullanılmıştır.

3.3 Model Eğitim Parametreleri

Klasik bir derin öğrenme (DL) mimarisinin eğitilmeden önce eğitim süreciyle ilgili çeşitli parametreler ayarlanmaktadır. Bu parametreler sırasıyla,

- **Optimize edici:** Bir derin öğrenme (DL) modeli eğilirken temel amaç modelin parametrelerinin en iyi şekilde ayarlanarak kayıp fonksiyon değerinin küçültülmesidir. Optimizasyon, her devirde modelin ağırlık ve öğrenme hızı gibi değerleriyle oynayarak model tahminlerini mümkün olduğunca optimize ve doğru yapılmasını sağlayan algoritmadır. Kayıp fonksiyonu ve model parametrelerini kombine ederek genel doğruluğu artırıp kayıp fonksiyonunun değerini küçülterek çalışmaktadırlar. Bu çalışmada Adam optimizasyon algoritması kullanılmaktadır. Bu optimizasyon algoritması, her ağ ağırlığı için öğrenme oranını ayrı ayrı güncellemektedir. Adam algoritması, öğrenme oranlarını ilk momente (ortalama) göre uyarlamak yerine, eğimlerin merkezlenmemiş varyansını kullanmaktadır. Adam optimizasyon algoritması yaygın olarak kullanılmakla beraber çeşitli avantajlara sahiptir. Algoritmanın uygulanması kolaydır, daha hızlı çalışma süresine ve düşük bellek gereksinimlerine sahipken diğer optimizasyon algoritmalarına göre daha az ayar gerektirmektedir.
- **Devir:** Toplam devir süresi algoritmanın tüm eğitim veri setinde kaç kere çalıştığını ifade etmektedir.
- **Grup boyutu:** Modelin ağırlık, öğrenme oranı gibi değerlerinin güncellenmesi için kullanılacak örnek sayısını belirtmektedir.
- **Öğrenme oranı:** Model ağırlıklarının ne kadarlık bir kısmının güncellenmesi gerektiğini belirten ölçektir.
- **Erken durdurma:** Modelin eğitimi için kullanılan veri setinin model tarafından ezberlenmesini engelleyen yöntemdir.

3.4 Model Eğitimi

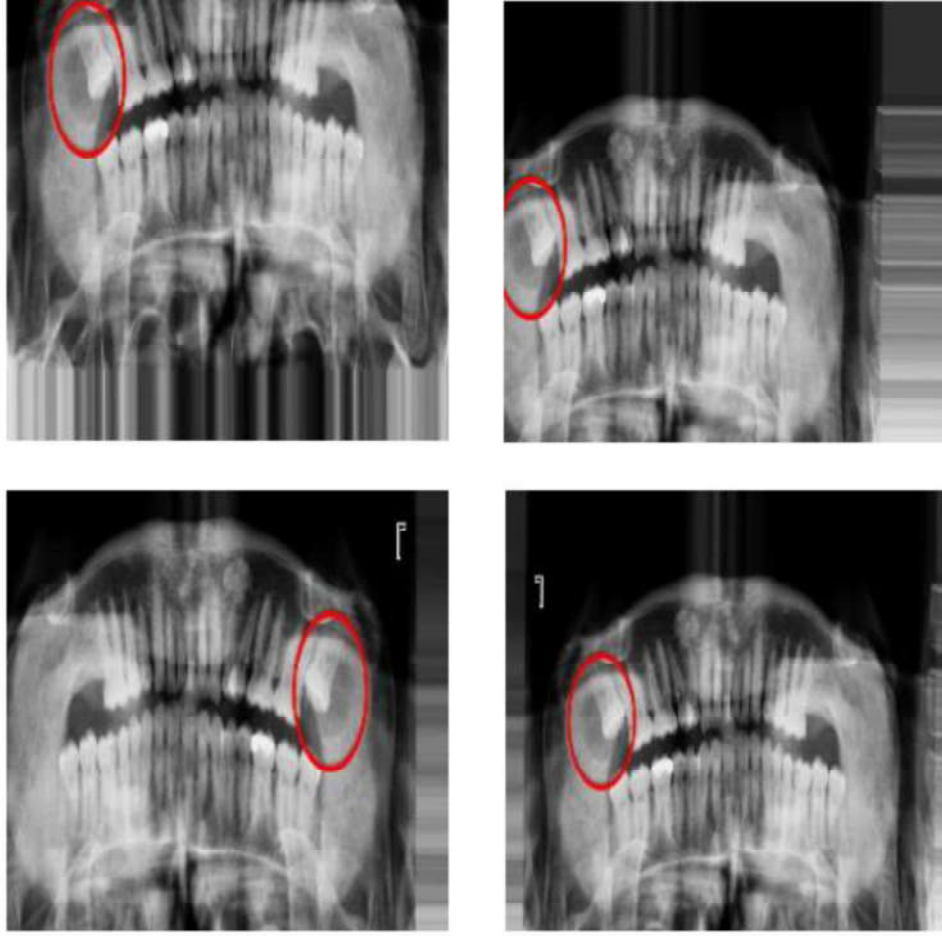
131 panoramik diş röntgeni ve 2505 göğüs X-Ray görüntüsünden manuel olarak modelin eğitiminde kullanılan maske görüntüleri oluşturulmuştur. Dental veri setinin %80'ni eğitim, %10'u doğrulama ve %10'u test için kullanılmıştır. Ancak panoramik diş röntgeni görüntülerinin sayısı az olduğu için bu görüntülere veri artırma yöntemleri uygulanmıştır. Bu yöntemler sırasıyla,

- a) Döndürme aralığı: Görüntüler belirlenen derece ile saat yönünde rastgele döndürülmüştür.
- b) Genişlik-Yükseklik kaydırma aralığı: Görüntülerin tüm pikselleri yatay ve dikey yönlerde hareket ettirilmiştir.
- c) Kesme aralığı: Görüntüler gerdirilmiştir.
- d) Yakınlaştırma aralığı: Görüntülerin yakınlaştırılmış versiyonu, yakınlaştırma aralığı yöntemi kullanılarak oluşturulmuştur.
- e) Yatay döndürme: Görüntüler yatay yönde çevrilmiştir.

Modeller eğitilirken yukarıda verilen yöntemler rastgele uygulanmış ve her devirde farklı görüntüler oluşturulmuştur. Bu yöntemlerden bazılarıyla elde edilen görüntü örnekleri Şekil 3.5 ile verilmektedir.

Görüntülerin çoğaltılması için ilgili yöntemlerin değerleri ayarlandıktan sonra eğitim parametreleri ayarlanmıştır. Bu çalışmada Double-UNet ve InceptionResV2-UNet olmak üzere iki farklı model tasarlanmıştır. Bu modellerin eğitiminde seçilen parametreler Çizelge 3.2 ile verilmektedir.

Modeller Python 3.8.8 yazılım programında NVIDIA GeForce MX150 ve Intel(R) Core (TM) i7-8565U CPU \@ 1.80GHz 1.99 GHz işlemci özelliklerine sahip bilgisayar kullanılarak eğitilmiştir. Erken durdurma yöntemi için doğrulama veri setinden elde edilen kayıp fonksiyonu seçilmiştir ve 5 devir boyunca değeri iyileşmediyse eğitim sonlandırılmıştır. Buna göre Double-Unet adlı model 41 devirde, InceptionResV2-Unet adlı model ise 48 devirde sonlandırılmıştır.



Şekil 3.5 : Veri artırma yöntemleriyle elde edilen örnek görüntüler.

Çizelge 3.2 : Model eğitim parametreleri.

Parametreler	Önceden eğitilmiş U-Net	Double-UNet	InceptionResV2-UNet
Devir	40	100	100
Grup boyutu	32	24	24
Öğrenme oranı	1e-4	1e-4	1e-4
Erken durdurma	-	Doğrulama Kayıp Fonksiyonu	Doğrulama Kayıp Fonksiyonu
Optimizasyon	Adam	Adam	Adam
Aktivasyon Fonksiyonu	Sigmoid	Sigmoid	Sigmoid
Kayıp Fonksiyonu	Kategorisel Çapraz Entropi	1-Zar Katsayısı	1-Zar Katsayısı

3.4.1 Double-UNet modeli eğitimi

Önceden eğitilmiş U-Net modeli için göğüs X-Ray görüntüleri kullanılarak model eğitimi gerçekleştirilmiştir. U-Net modelinde kullanılan tüm evrişim katmanlarının filtre boyutu 3x3 olarak ayarlanmıştır. Son evrişim katmanı hariç tüm katmanlarda ReLU aktivasyon fonksiyonu kullanılmıştır. 2505 görüntünün %10'u modeli test etmek için ayrılmıştır. Şekil 3.6 (a) ile orijinal girdi görüntüsü, Şekil 3.6 (b) ile gerçek maske görüntüsü ve Şekil 3.6 (c) ile tahmin maske görüntüsü gösterilmektedir.

Model eğitimi bittikten sonra kaydedilmiştir. Daha sonra kaydedilen model yüklenerek tüm eğitilebilir katmanları eğitilemez hale getirilmiştir. İkinci U-Net modelinin kod çözücü yolu kaydedilen modelin kodlayıcı yolundaki katmanları ile birleştirilmiştir. İkinci U-Net modelin de kullanılan tüm ters evrişim ve evrişim katmanlarının filtre boyutları 3x3 olarak ayarlanmıştır. Aktivasyon fonksiyonu olarak ReLU seçilmiştir. Modelin son katmanında yer alan aktivasyon fonksiyonu ise sigmoid olarak ayarlanmıştır. Çizelge 3.2 ile verildiği gibi modelin parametreleri ayarlanarak Double-UNet modeli eğitilmiştir. Modelin eğitim süreci boyunca elde edilen devir sayısına karşı performans metriklerinin hem eğitim hem de doğrulama veri setlerindeki başarısı Şekil 3.7 ile verilmektedir.

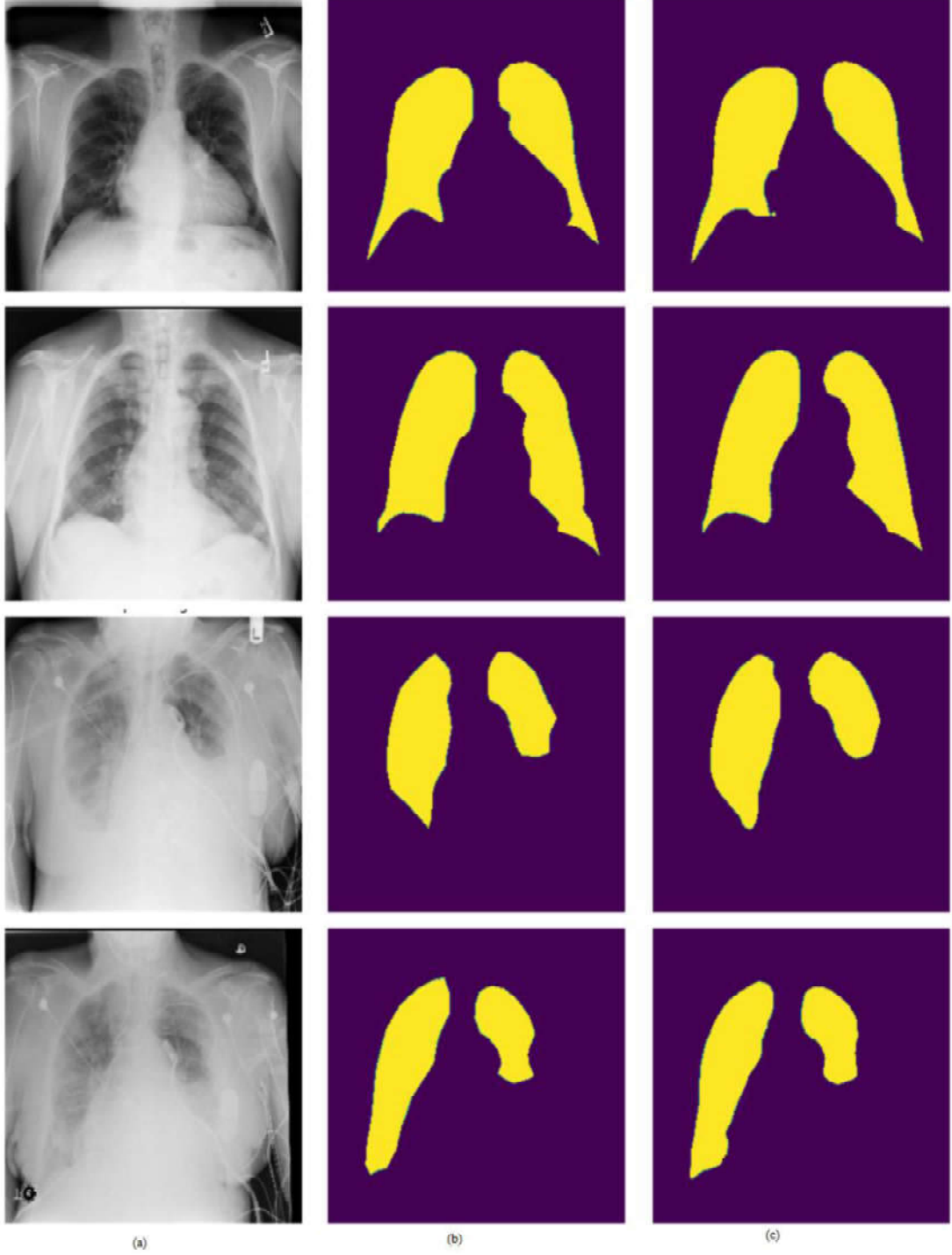
Modelin eğitimi bittikten sonra model hiç görmediği bir veri setiyle test edilmiştir. Elde edilen farklı sonuçlar Şekil 3.8 ile verilmiştir. Modelin başarı metriklerinin sonuçları ise Çizelge 3.3 ile verilmektedir.

Çizelge 3.3 : Double-UNet modelinin performans sonucu.

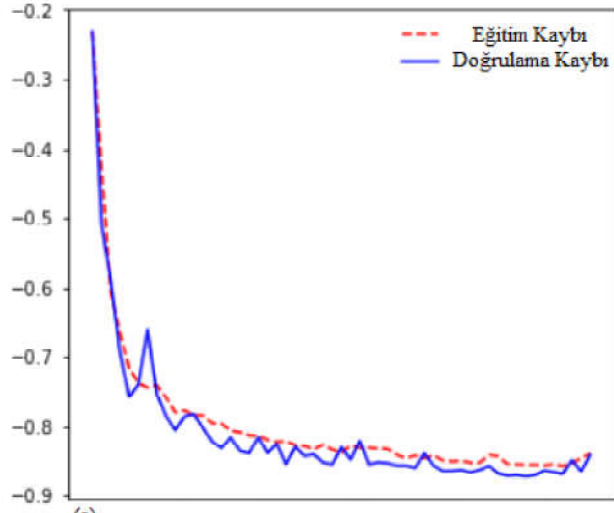
Veri Setleri	Zar Katsayısı	Doğruluk	Kayıp Fonksiyonu
Eğitim	0.86	0.9675	-0.86
Doğrulama	0.8673	0.9728	-0.9728
Test	0.8712	0.9686	-0.8712

Çizelge 3.3'te belirtildiği gibi tasarlanan model %87.12 gibi bir oranla çalışmaktadır. Modelin yüksek oranda bir segmentasyon gerçekleştirdiği görülmektedir. Fakat

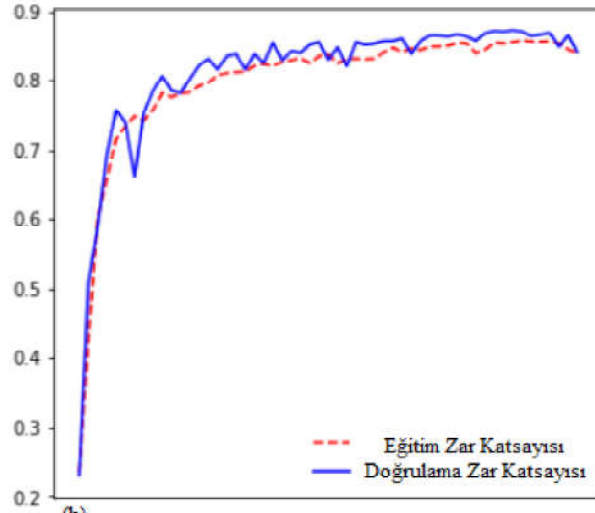
bitişik bulunan dış bölgelerini ayırt etmekte zorlandığı da anlaşılmaktadır. Aynı zamanda ince detayları da kaçırabilmektedir. Sonuç olarak modelin sınırlı sayıdaki veri setiyle iyi bir sonuç çıkardığı söylenebilmektedir.



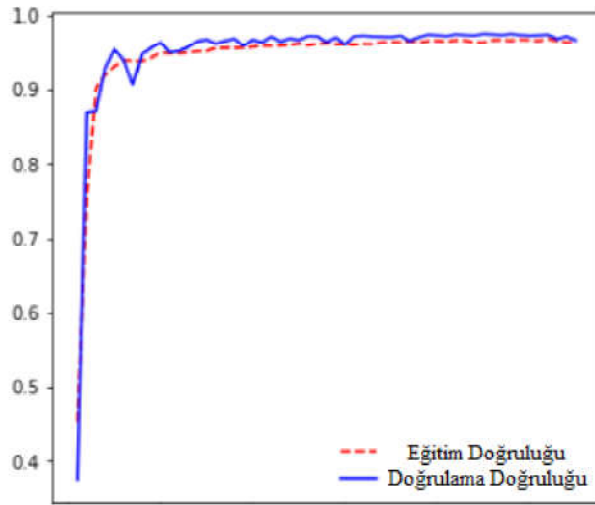
Şekil 3.6 : (a) Göğüs X-Ray girdi görüntüsü, (b) Gerçek maske görüntüsü ve (c) U-Net modelinin tahmin görüntüsü.



(a)

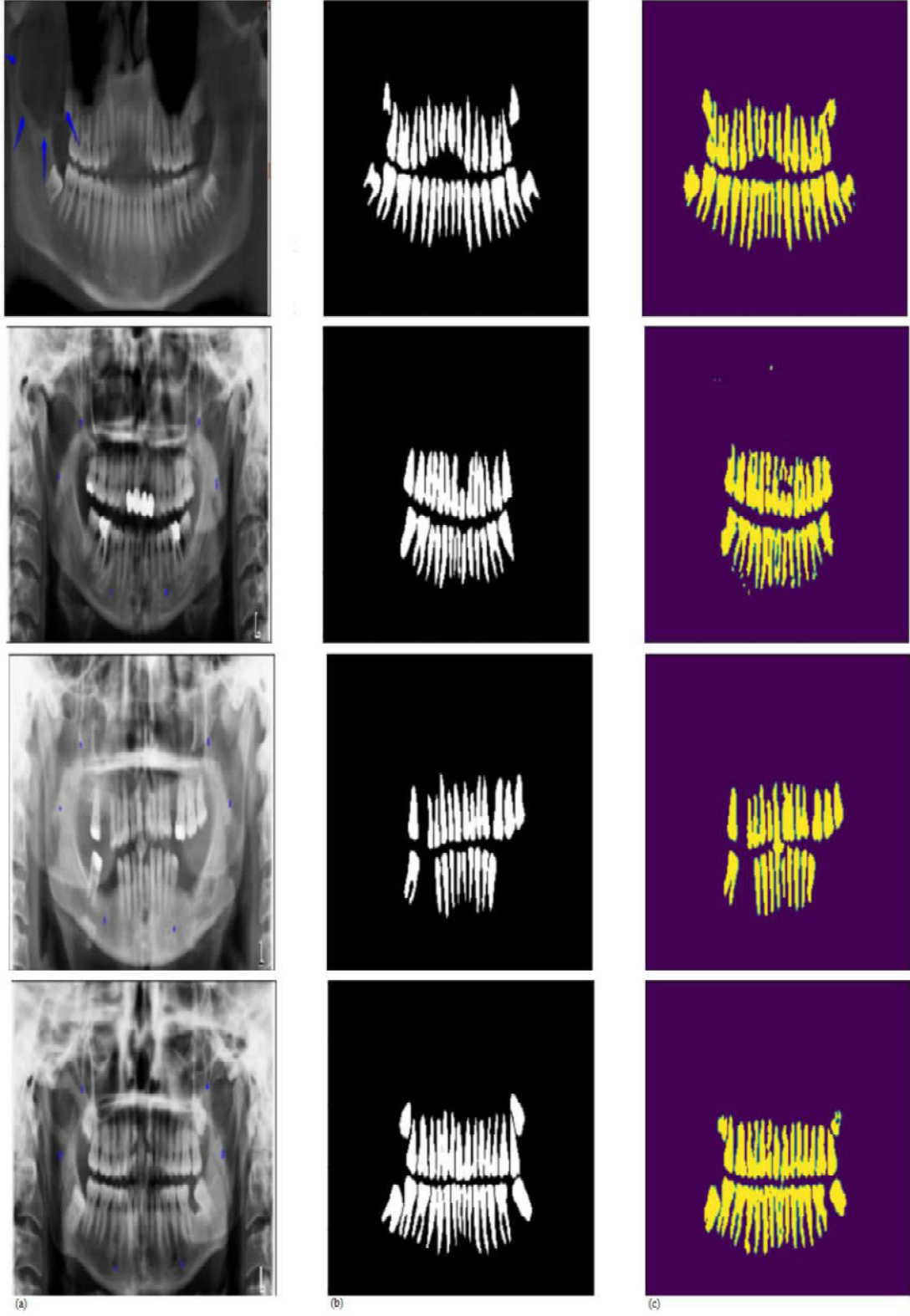


(b)



(c)

Şekil 3.7 : Double-UNet modelinin öğrenme eğrisi.



Şekil 3.8 : (a) Orijinal girdi görüntüsü, (b) Gerçek maske görüntüsü ve (c) Double-UNet modelinin tahmin maske görüntüsü.

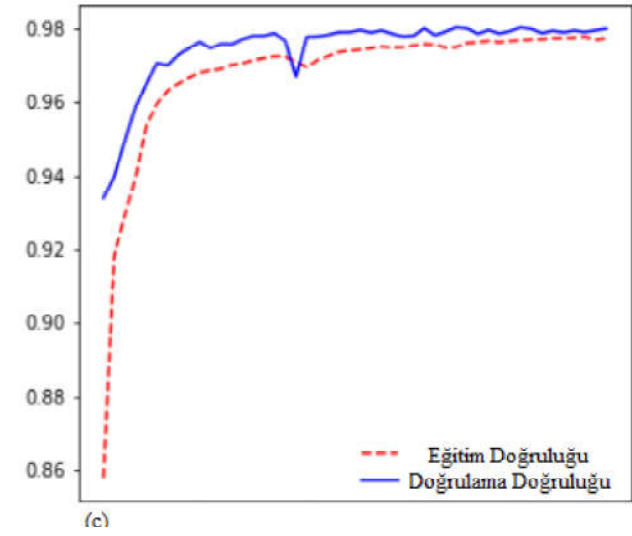
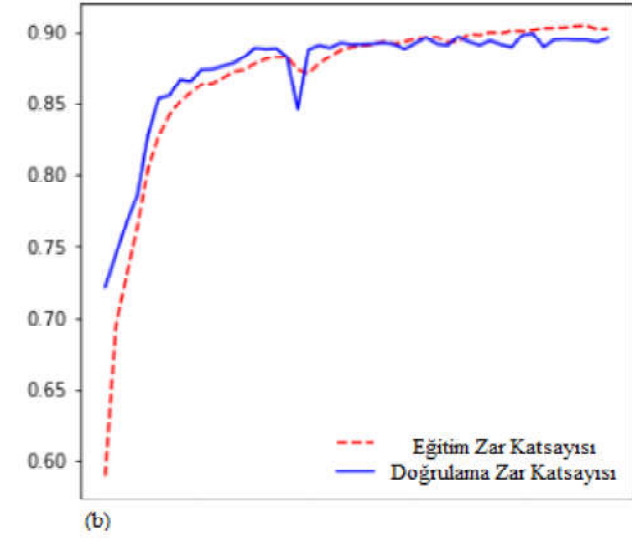
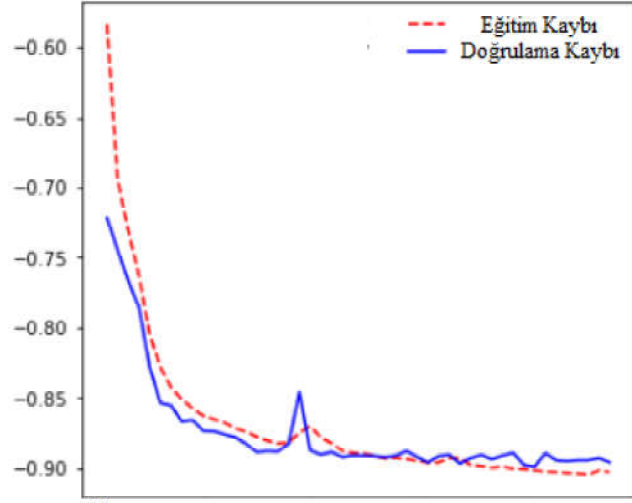
3.4.2 InceptionResV2-UNet Modeli Eğitimi

Önceden eğitilmiş Inception-ResNet-v2 modeli keras kütüphanesi kullanılarak yüklenmiştir. Yüklenen modelin tüm eğitilebilir katmanları eğtilemez hale getirilmiştir. Bu model U-Net modelinin kodlayıcı yoluyla birleştirilmiştir. Kod çözücü yolunda kullanılan ters evrişim ve evrişim katmanları için 3x3 filtre boyutu seçilmiştir. Son katman haricinde aktivasyon fonksiyonu ReLU iken son katmanda sigmoid fonksiyonudur. Modelin eğitimi bittikten sonra devir sayısına karşı elde edilen başarı metriklerinin sonucu Şekil 3.9 ile verilmektedir. Modelin eğitimi bittikten sonra test veri setiyle elde edilen tahminler Şekil 3.10 ile verilmektedir. Modelin başarı metriklerinin sonuçları ise Çizelge 3.4 ile verilmektedir.

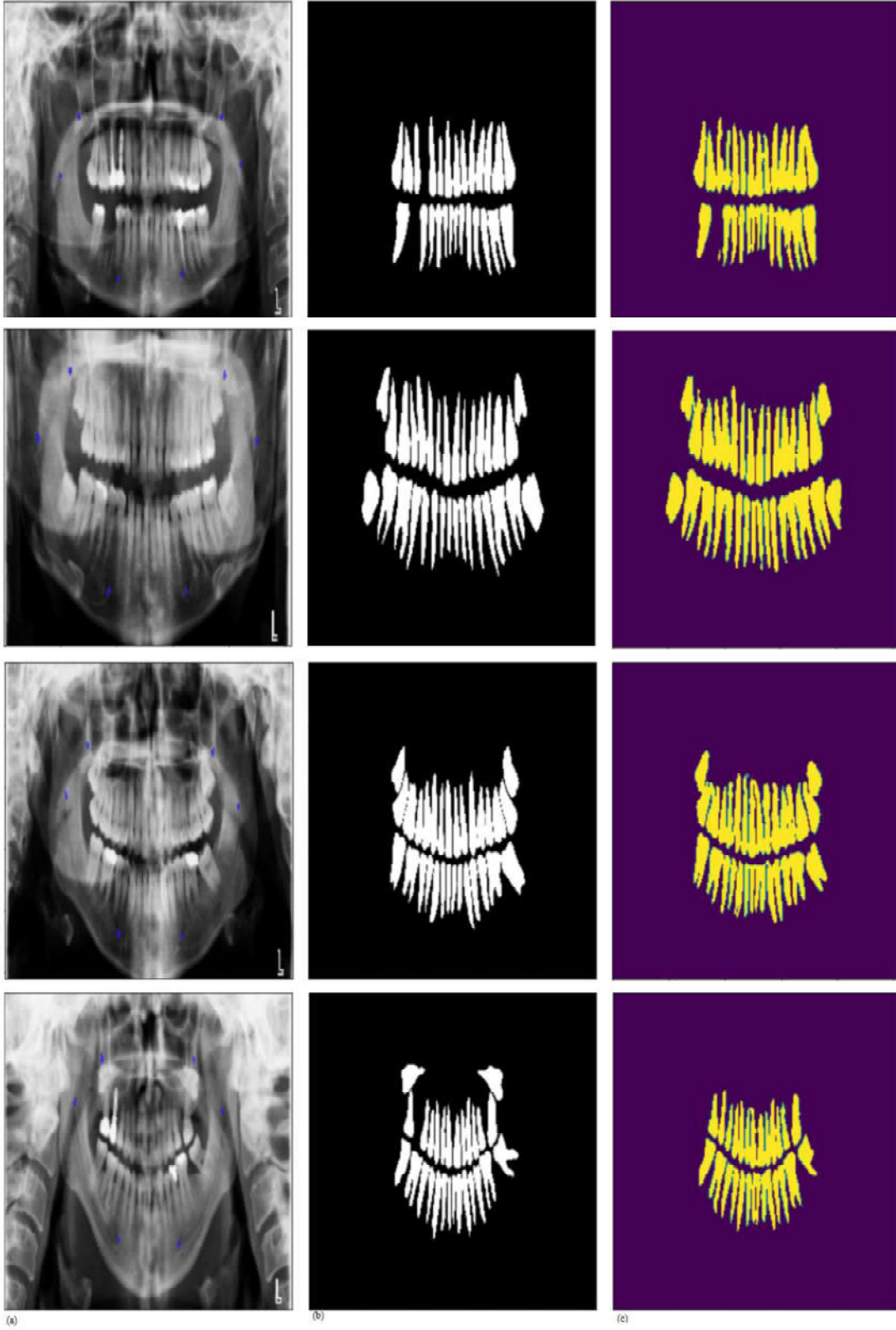
Çizelge 3.4 : InceptionResV2-UNet modelinin performans sonucu.

Veri Setleri	Zar Katsayısı	Doğruluk	Kayıp Fonksiyonu
Eğitim	0.9023	0.9773	-0.9023
Doğrulama	0.8960	0.9798	-0.8960
Test	0.9026	0.9759	-0.9026

Çizelge 3.4 ile görüldüğü üzere tasarlanan model %90.26 gibi bir oran ile görüntüleri segmente etmektedir. Modelin yüksek oranda bir segmentasyon gerçekleştirdiği söylenebilmektedir. Fakat bazı bölgelerde bulunan dış bölgelerini (Şekil 3.10'da 3.satırdaki görüntü) ayırt etmekte zorlandığı da anlaşılmaktadır. Bunun sebebi olarak da görüntüde yer alan dış harici bölgelerin kontrastına benzer seviyede olması gösterilebilir. Sonuç olarak modelin sınırlı sayıdaki veri setiyle iyi bir sonuç çıkardığı görülmektedir.



Şekil 3.9 : InceptionResV2-UNet modelinin öğrenme eğrisi.



Şekil 3.10 : (a) Orijinal girdi görüntüsü, (b) Gerçek maske görüntüsü ve (c) InceptionResV2-UNet modelinin tahmin maske görüntüsü.

4. SONUÇLAR

Bu çalışmanın ana amacı az sayıda bulunan dental veri setiyle derin öğrenme (DL) mimarilerinin veriyi ezberlemeden başarılı bir şekilde eğitilmesine olanak sağlarken tüm segmentasyon sürecinin otomatik gerçekleştirilmesidir. Bu amaç için transfer öğrenme yöntemi kullanılmıştır. Aynı zamanda farklı veya aynı alan/görev de yer alan önceden eğitilmiş iki farklı modelin karşılaştırılması bu çalışmadaki ikincil amaç olarak belirtilebilir.

Tez boyunca göğüs X-Ray, ImageNet ve panoramic X-Ray görüntüleri veri seti olarak kullanılmıştır. Hem görüntülerin boyutları birbirinden farklı olduğu için hem de görüntülerde aynı formatı sağlamak için görüntüler üzerinde ön işlem aşamaları uygulanmıştır. Tasarlanan iki derin öğrenme (DL) mimarisinin de temelini U-Net modeli oluşturmaktadır. Transfer öğrenme yönteminin gerekliliği olan önceden eğitilmiş modeller için göğüs röntgen görüntüleriyle eğitilen farklı bir U-Net modeli ile ImageNet veri setiyle eğitilen Inception-ResNet-V2 modeli seçilmiştir. Diş X-Ray görüntüleriyle eğitilen iki farklı modelin performansları zar katsayısı metriği ile test edilmiştir. İki modelin de eğitiminde kullanılan parametrelerin benzer olmasına dikkat edilmiştir. Böylelikle sonuçların yorumlanması daha kolay ve mantıklı olmaktadır. Metrik sonuçlarına göre InceptionResV2-UNet modelinin Double-UNet modelinden nispeten daha iyi çalıştığı yani daha doğru bir şekilde diş bölgesini segmente ettiği gözlenmiştir. Bu durumun önceden eğitilmiş modellerin eğitiminde kullanılan veri setlerinin boyutunun farklılığı, ağ mimarilerinin daha farklı derinlikte ve genişlikte olması söylenebilir.

Elde edilen sonuçlarla birlikte tasarlanan modeller daha önceki hiçbir çalışmada kullanılmayan mimarilere sahiptir. Çalışmalar [19-25]'de her ne kadar transfer yöntemi kullanılmış olsa da bu çalışmada kullanılan önceden eğitilmiş model çeşitleri ilk olarak kullanılmaktadır.

Yapılan çalışma ile elde edilen sonuçlar ileri de geliştirilecek modeller için oldukça iyi bir temel sağlamaktadır. Tasarlanan modellerin iyileştirilmesi ve geliştirilmesi adına aşağıdaki işlemler gerçekleştirilebilir:

- Eğitim sürecinde daha yoğun veri çoğaltma yöntemi uygulanabilir.
- Double-UNet modelinde kullanılan önceden eğitilmiş U-Net mimarisinin eğitiminde daha büyük boyutlu veri seti kullanılabilir. Böylelikle elde edilecek öznelik haritalarının başarısı artırılabilir.
- Segmentasyon sadece dış üzerinden tek sınıflı olarak gerçekleştirilmiştir. Dolayısıyla çeşitli bölgelerin segmentasyonu için kullanılacak sınıf sayısı artırılabilir. Örneğin çürük, implant, tümör vb. yapılar için de çeşitli parametrelerin değiştirilerek belirlenen yöntem ve modellerle kullanılması söz konusu olabilir.

KAYNAKLAR

- [1] **Pauwels, R.**, (2020). History of Dental Radiography: Evolution of 2d and 3d Imaging Modalities, *Medical Physics International Journal*.
- [2] **Razali, M. R. M., Ahmad, N. S., Hassan, R., Zaki, Z. M. and Ismail W.**, (2015). Sobel and Canny Edges Segmentations for the Dental Age Assessment, *International Conference on Computer Assisted System in Health*, pages 62-66.
- [3] **Trivedi, D. N., Kothari, A. M., Shah, S., and Nikunj, S.**, (2015). Dental Image Matching By Canny Algorithm for Human Identification, *International Journal of Advanced Computer Research*, 4(17):985–990.
- [4] **Shah, S., Abaza, A., Ross, A., Ammar, H.**, (2006). Automatic Tooth Segmentation Using Active Contour Without Edges, *Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, pp. 1-6.
- [5] **Chen, H. and Jain, A. K.**, (2004). Tooth contour extraction for matching dental radiographs, *Proceedings of the 17th International Conference on Pattern Recognition*, pp. 522-525, Vol.3.
- [6] **Beucher, S., Meyer, F.**, (1993). The morphological approach to segmentation: the watershed transformation, In *Mathematical Morphology in Image Processing* (Ed. E. R. Dougherty), pages 433–481.
- [7] **Na, S. D., et al.**, (2014). Individual tooth region segmentation using modified watershed algorithm with morphological characteristic, *Bio-Medical Materials and Engineering*, IOS Press, 3303-3309.
- [8] **Li, H., Sun, G., Sun, H., and Liu, W.**, (2012). Watershed algorithm based on morphology for dental x-ray images segmentation, *Intl. Conference on Signal Processing Proceedings*, volume 2, pages 877–880.
- [9] **Lurie, A., Tosoni, G. M., Tsimikas, J., and Fitz, W.**, (2012). Recursive hierarchic segmentation analysis of bone mineral density changes on digital panoramic images, *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, 113(4):549–558.
- [10] **Modi, C. K. and Desai, N. P.**, (2011). A simple and novel algorithm for automatic selection of ROI for dental radiograph segmentation, *Canadian Conference on Electrical and Computer Engineering*, pages 000504–000507.

- [11] **Chen, H. and Jain, A. K.**, (2004). Matching of dental X-ray images for human identification, *Pattern Recognition*, Volume 37, Issue 7, pp. 1519-1532.
- [12] **He, K., Gkioxari, G., Dollár, P., and Girshick, R.**, (2017). Mask R-CNN, *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2980-2988.
- [13] **Ren, S., He, K., Girshick, R., and Sun, J.**, (2015). Faster R-CNN: Towards real-time object detection with region proposal networks, *Proceedings of the 28th International Conference on Neural Information Processing Systems, Volume 1 (NIPS'15)*, 91-99.
- [14] **Jader, G., Fontineli, J., Ruiz, M., Abdalla, K., Pithon, M. and Oliveira, L.**, (2018). Deep Instance Segmentation of Teeth in Panoramic X-Ray Images, *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pp. 400-407.
- [15] **Chen, H., Zhang, K., Lyu, P. et al.**, (2019). A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films, *Sci Rep* 9, 3840.
- [16] **Ronneberger, O., Fischer, P., Brox, T.**, (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation, *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Springer, LNCS, Vol.9351, 234—241.
- [17] **Koch, T. L., Perslev, M., Igel, C., and Brandt, S. S.**, (2019). Accurate Segmentation of Dental Panoramic Radiographs with U-NETS, *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp. 15-19.
- [18] **Sivagami, S., Chitra, P., Kailash, G. S. R., and Muralidharan, S. R.**, (2020). UNet Architecture Based Dental Panoramic Image Segmentation, *2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)*, pp. 187-191.
- [19] **Prajapati, S. A., Nagaraj, R., and Mitra, S.**, (2017). Classification of dental diseases using CNN and transfer learning, *2017 5th International Symposium on Computational and Business Intelligence (ISCBI)*, pp. 70-74.
- [20] **Mahdi, F. P., Yagi, N., and Kobashi, S.**, (2020). Automatic Teeth Recognition in Dental X-Ray Images Using Transfer Learning Based Faster R-CNN, *2020 IEEE 50th International Symposium on Multiple-Valued Logic (ISMVL)*, pp. 16-21.
- [21] **Pardamean, B., Cenggoro, W. T., Rahutomo, R., Budiarto, A., Karuppiah, K. E.**, (2018). Transfer Learning from Chest X-Ray Pre-trained Convolutional Neural Network for Learning Mammogram Data, *3rd International Conference on Computer Science and Computational Intelligence 2018*.

- [22] **Jouni, H., Issa, M., Harb, A., Jacquemod, G., & Leduc, Y.,** (2016). Neural Network architecture for breast cancer detection and classification, *2016 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET)*.
- [23] **Khan, S., Islam, N., Jan, Z., Din, I. U., & Rodrigues, J. J. C.,** (2019). A novel deep learning based framework for the detection and classification of breast cancer using transfer learning, *Pattern Recognition Letters*, 125, 1-6.
- [24] **Abubakar, A., Ugail, H., & Bukar, A. M.,** (2020). Assessment of human skin burns: A deep transfer learning approach, *Journal of Medical and Biological Engineering*, 40(3), 321-333.
- [25] **Cirillo, M. D., Mirdell, R., Sjöberg, F., & Pham, T. D.,** (2019). Tensor decomposition for colour image segmentation of burn wounds, *Scientific reports*, 9(1), 1-13.
- [26] **Szegedy, C., Ioffe, S., Vanhoucke, V., and Alemi, A. A.,** (2017). Inception-v4, inception-ResNet and the impact of residual connections on learning, *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI'17)*, 4278–4284.
- [27] **Pan, S. J., and Yang, Q.,** (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.
- [28] **Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L.,** (2009). ImageNet: A Large-Scale Hierarchical Image Database, *IEEE Computer Vision and Pattern Recognition (CVPR)*.
- [29] **GNU Image Manipulation Program (GIMP)**, <https://www.gimp.org/>.
- [30] **Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., and Summers, M. R.,** (2017). ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3462-3471.
- [31] **Dice, L. R.,** (1945). Measures of the Amount of Ecologic Association Between Species, *Ecology*, 26 (3): 297–302.
- [32] **Sørensen, T.,** (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons, *Kongelige Danske Videnskabernes Selskab*, 5 (4): 1–34.