



**PANORAMİK RADYOGRAFİ GÖRÜNTÜLERİ  
ÜZERİNDE DİŞ ÇÜRÜĞÜNÜN DERİN ÖĞRENME  
TABANLI YÖNTEMLER İLE ANALİZİ VE TESPİTİ**

**2023  
YÜKSEK LİSANS TEZİ  
BİLGİSAYAR MÜHENDİSLİĞİ**

**Dilara ÖZDEMİR**

**Tez Danışmanı  
Doç. Dr. Caner ÖZCAN**

**PANORAMİK RADYOGRAFİ GÖRÜNTÜLERİ ÜZERİNDE DİŞ  
ÇÜRÜĞÜNÜN DERİN ÖĞRENME TABANLI YÖNTEMLER İLE ANALİZİ  
VE TESPİTİ**

**Dilara ÖZDEMİR**

**Tez Danışmanı  
Doç. Dr. Caner ÖZCAN**

**T.C.  
Karabük Üniversitesi  
Lisansüstü Eğitim Enstitüsü  
Bilgisayar Mühendisliği Anabilim Dalında  
Yüksek Lisans Tezi  
Olarak Hazırlanmıştır**

**KARABÜK  
Ağustos 2023**

Dilara ÖZDEMİR tarafından hazırlanan “PANORAMİK RADYOGRAFİ GÖRÜNTÜLERİ ÜZERİNDE DİŞ ÇÜRÜĞÜNÜN DERİN ÖĞRENME TABANLI YÖNTEMLER İLE ANALİZİ VE TESPİTİ” başlıklı bu tezin Yüksek Lisans Tezi olarak uygun olduğunu onaylarım.

Doç. Dr. Caner ÖZCAN

.....

Tez Danışmanı, Bilgisayar Mühendisliği Anabilim Dalı

Bu çalışma, jürimiz tarafından Oy Birliği ile Bilgisayar Mühendisliği Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir. 28/08/2023

Ünvanı, Adı SOYADI (Kurumu)

İmzası

Başkan : Dr. Öğr. Üyesi Ferhat ATASOY (KBÜ)

.....

Üye : Doç. Dr. Caner ÖZCAN (KBÜ)

.....

Üye : Dr. Öğr. Üyesi Yusuf Yargı BAYDİLLİ (HÜ)

ONLINE

KBÜ Lisansüstü Eğitim Enstitüsü Yönetim Kurulu, bu tez ile, Yüksek Lisans derecesini onamıştır.

Doç. Dr. Zeynep ÖZCAN

.....

Lisansüstü Eğitim Enstitüsü Müdürü

*“Bu tezdeki tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu; ayrıca bu kuralların ve ilkelerin gerektirdiği şekilde, bu çalışmadan kaynaklanmayan bütün atıfları yaptığımı beyan ederim.”*

Dilara ÖZDEMİR

## **ÖZET**

**Yüksek Lisans Tezi**

### **PANORAMİK RADYOGRAFİ GÖRÜNTÜLERİ ÜZERİNDE DİŞ ÇÜRÜĞÜNÜN DERİN ÖĞRENME TABANLI YÖNTEMLER İLE ANALİZİ VE TESPİTİ**

**Dilara ÖZDEMİR**

**Karabük Üniversitesi**

**Lisansüstü Eğitim Enstitüsü**

**Bilgisayar Mühendisliği Anabilim Dalı**

**Tez Danışmanı:**

**Doç. Dr. Caner ÖZCAN**

**Ağustos 2023, 56 sayfa**

Bu çalışmada, uzmanı desteklemek için çevresel faktörlerden etkilenmeyen derin öğrenme yöntemlerine dayalı diş çürüğü lezyonlarının tespitini panoramik görüntüler üzerinde tamamen otomatikleştirebilen bir model önerilmektedir. Çürük, bireylerin en yaygın karşılaştığı ve en çok diş sağlığı hizmetlerine başvurduğu bir diş lezyonudur. Diş çürüğü lezyonları yemek artıklarından ya da ağız içinde aktivite gösteren bakterilerin diş yapısında bozulma oluşturmamasından kaynaklanmaktadır. Bu lezyonlar klinik ortamlarda uzmanların gerçekleştirdiği görsel muayeneler sonucunda tespit edilmektedir. Görsel muayeneler X-ışını kullanarak görüntüleme yapabilen cihazların ürettiği radyografik görüntüler ile gerçekleştirilmektedir. Diş çürüğü lezyonlarının tespiti ve teşhisi uzmana klinik ortamda bir yük teşkil etmektedir. Bu durumdan dolayı kliniksel zaman yönetimi ve tedavi süreci zorlaşabilmektedir. Diş çürüğü lezyonlarının derin öğrenme yöntemleri ile sınıflandırılması, bölütlenmesi ve tespiti bu tür

zorlukların aşılabilmesi için önemli bir araştırma konusudur. Bu çalışmada, uzmanlar tarafından oluşturulmuş çürük etiketlerine sahip panoramik veri seti Mask R-CNN mimarisine girdi olarak verilmiş ve performans ölçümü gerçekleştirilmiştir. Daha sonrasında kırılan her görüntüden çürük etiketine sahip olanlar çürük sınıfına ve çürük etiketine sahip olmayanlar ise çürük olmayan sınıfına dahil edilmiştir. Kategorilenmiş bu ana veri seti farklı sınıf oranları dikkate alınarak alt veri setlerine ayrılmıştır. Toplamda 2709 kırılmış görüntüden oluşan alt veri seti üzerinde öz-dikkat mekanizmasına sahip ConvNeXt modeli ile %80,34 doğruluk, %86,54 kesinlik, %91,47 duyarlılık ve %88,94 f1-skoru performansına ulaşılmıştır. Aynı zamanda ConvNeXt modeli ile k-katman çapraz doğrulama yöntemi ile modelin dış çürüğü görüntülerini ne kadar iyi sınıflandırabildiği ölçülmüştür. Alt ve ana veri setleri üzerinde derin öğrenme sınıflandırıcılarından GoogLeNet, ShuffleNet, ResNet-50, MobileNet, DenseNet-169, DenseNet-121, VGG-19 mimarileri ile eğitimler gerçekleştirilmiştir. Tez çalışması kapsamında yapılandırılmış sınıflandırıcısıyla önerilen yöntem ile de eğitim gerçekleştirilmiştir. Çürük ve çürük olmayan sınıflara göre kategorize edilen ana veri seti üzerinde önerilen yöntem, diğer sınıflandırma modellerine ve alt veri setlerine göre %93,17 doğruluk, %89,43 kesinlik, %85,84 duyarlılık ve %87,49 f1-skoru metriği ile daha yüksek performans göstermiştir. Bu sonuçlar ışığında sınıflandırma, bölütleme ve tespit işlemleri başarılı bir şekilde gerçekleştirilmiştir.

**Anahtar Sözcükler :** Panoramik radyografi, Derin öğrenme, Dış çürüğü lezyonları, Görüntü sınıflandırma, Nesne tespiti, Bölütleme, Örnek bölütleme.

**Bilim Kodu :** 92431

## **ABSTRACT**

**M. Sc. Thesis**

### **ANALYSIS AND DETECTION OF DENTAL CARIES ON PANORAMIC RADIOGRAPHY IMAGES WITH DEEP LEARNING-BASED METHODS**

**Dilara ÖZDEMİR**

**Karabük University  
Institute of Graduate Programs  
Department of Computer Engineering**

**Thesis Advisor:**

**Assoc. Prof. Dr. Caner ÖZCAN**

**August 2023, 56 pages**

In this study, a system that can fully automate the detection of caries lesions on panoramic images based on deep learning methods unaffected by environmental factors is provided to support the specialists. Caries is a dental lesion commonly encountered by a person and most frequently applied to dental health services. Tooth caries lesions are the deterioration of tooth structure caused by food residues or bacteria formed active in the mouth. These lesions are detected during visual examinations performed by specialists in clinical settings. Visual assessments are performed with radiographic images produced by devices capable of imaging using X-rays. Detection and diagnosis of dental caries lesions burden the specialist in the clinical setting. Because of this situation, clinical time management and treatment processes can be complicated. Classification, segmentation, and detection of dental caries lesions with deep learning methods is an important research topic to overcome such difficulties. A dataset of panoramic dental images with caries annotations created

by specialists was given as input to Mask R-CNN architecture, and performance was measured. Afterwards, each cropped image with caries annotations were included in the caries class, and each cropped image without annotations was included in the non-caries category. This categorized main data set was divided into sub-datasets considering different class ratios. 80.34% accuracy, 86.54% precision, 91.47% sensitivity and 88.94% f1-score performance were achieved with the ConvNeXt model with self-attention mechanism on the sub-dataset consisting of 2709 cropped images in total. At the same time, the k-fold cross validation method with the ConvNeXt model measured how well the model could classify dental caries images. GoogLeNet, ShuffleNet, ResNet-50, MobileNet, DenseNet-169, DenseNet-121, and VGG-19 deep learning architectures have been trained with main and sub datasets. The training was also carried out with the method proposed with the structured classifier within the scope of the thesis study. The proposed method on the main dataset, categorized according to caries and non-carious classes, showed higher performance with 93.17% accuracy, 89.43% precision, 85.84% sensitivity and 87.49% f1-score metric compared to other classification models and sub-datasets. According to these results, classification, segmentation, and detection processes are carried out successfully.

**Key Word** : Panoramic radiography, Deep learning, Dental caries lesions, Image classification, Object detection, Segmentation, Instance segmentation.

**Science Code** : 92431



## TEŞEKKÜR

Tez çalışmamızın araştırma, planlama ve yürütülme süreçlerinde ilgisini ve desteğini hiçbir zaman esirgemeyen, alanda sahip olduğu bilgi ve tecrübelerinden tez sürecim boyunca faydalandığım, bilgilendirme ve yol göstermesiyle çalışmamı bilimsel olarak şekillendiren sayın danışmanım Doç. Dr. Caner ÖZCAN'a sonsuz teşekkürlerimi sunarım.

Çalışma süreci boyunca yol gösterici ve bilgilendirici yaklaşımıyla bana destek olan Öğr. Gör. Buse Yaren KAZANGİRLER'e, bana ümit verip her zaman yanımda olan ve desteğini benden hiç esirgemeyen Arş. Gör. Elif MEŞECİ'ye ve beni her zaman çalışmaya teşvik ederek destekleyen Süheda AKDAĞ'a teşekkür ederim. Veri etiketleme platformu DentiAssist'i geliştirerek çalışmada büyük rol oynayan sayın hocam Öğr. Gör. Dr. Ahmet KARAOĞLU'na teşekkürlerimi sunarım. Çalışmanın ana temelini oluşturan veri kümesinin elde edilmesinde büyük role sahip olan sayın Dr. Öğr. Üyesi Adem PEKİNCE'ye ve sayın Doç. Dr. Yasin YAŞA'ya sonsuz teşekkür ederim.

Bu tez çalışmamızı "KBUBAP-23-YL-066" proje numarası ile desteklemeye layık gören Karabük Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi'ne ve "2200272" numarası ile desteklemeye layık gören Türkiye Bilimsel ve Teknolojik Araştırma Kurumu TÜBİTAK'a teşekkür ederim.

Hayatım boyunca beni destekleyen ve manevi yardımlarını asla esirgemeyen aileme sonsuz teşekkürlerimi sunarım.

## İÇİNDEKİLER

### Sayfa

KABUL.....	ii
ÖZET .....	iv
ABSTRACT.....	vi
TEŞEKKÜR.....	viii
İÇİNDEKİLER .....	ix
ŞEKİLLER DİZİNİ.....	xi
ÇİZELGELER DİZİNİ .....	xii
SİMGELER VE KISALTMALAR DİZİNİ .....	xiii
BÖLÜM 1 .....	1
GİRİŞ .....	1
BÖLÜM 2 .....	5
DIŞ ÇÜRÜĞÜ .....	5
2.1. DIŞ GÖRÜNTÜLERİ VE GÖRÜNTÜ TİPLERİ .....	5
2.2. DIŞ ÇÜRÜĞÜNÜN KLİNİK ORTAMDA TEŞHİSİ .....	6
2.3. DIŞ ÇÜRÜĞÜ VERİ KÜMESİ.....	7
2.3.1. Verilerin Etiketlenmesi .....	8
2.3.2. Panoramik Görüntülerin Kırpılması .....	11
BÖLÜM 3 .....	17
GÖRÜNTÜ BÖLÜTLEME VE NESNE TESPİTİ .....	17
3.1. BÖLGE TABANLI EVRİŞİMLİ SİNİR AĞLARI .....	17
3.2. HIZLI BÖLGE TABANLI EVRİŞİMLİ SİNİR AĞLARI.....	18
3.3. DAHA HIZLI BÖLGE TABANLI EVRİŞİMLİ SİNİR AĞLARI .....	19
3.4. MASKE BÖLGE TABANLI EVRİŞİMLİ SİNİR AĞLARI .....	20

	<b><u>Sayfa</u></b>
BÖLÜM 4 .....	22
GÖRÜNTÜ SINIFLANDIRMA.....	22
4.1. EVRİŞİMLİ SİNİR AĞLARI .....	22
4.2. ÖN EĞİTİMLİ EVRİŞİMLİ SİNİR AĞLARI.....	23
4.2.1. GoogLeNet .....	23
4.2.2. ShuffleNet.....	25
4.2.3. ResNet.....	26
4.2.4. MobileNet .....	27
4.2.5. VGG.....	29
4.2.6. DenseNet.....	30
4.3. ÖZ-DİKKAT TABANLI DERİN ÖĞRENME YÖNTEMLERİ .....	31
4.4. ÖNERİLEN YÖNTEM.....	32
BÖLÜM 5 .....	35
DENEYSEL ÇALIŞMALAR.....	35
5.1. PERFORMANS METRİKLERİ .....	35
5.2. NESNE TESPİT MODELLERİNİN SONUÇLARI .....	37
5.3. ÖZ-DİKKAT TABANLI DERİN ÖĞRENME YÖNTEMLERİNİN SONUÇLARI .....	39
5.4. ÖN EĞİTİMLİ EVRİŞİMLİ SİNİR AĞLARI VE ÖNERİLEN YÖNTEM SONUÇLARI .....	41
BÖLÜM 6 .....	47
SONUÇLAR VE ÖNERİLER .....	47
KAYNAKLAR .....	49
ÖZGEÇMİŞ .....	56

## ŞEKİLLER DİZİNİ

### Sayfa

Şekil 2.1. Veri kümesinde bulunan radyografi örnekleri. ....	7
Şekil 2.2. Panoramik görüntüler ve çürük etiketleri. ....	8
Şekil 2.3. DentiAssist görüntü etiketleme platformunun arayüzü. ....	9
Şekil 2.4. Çekim hataları işaretleme paneli. ....	9
Şekil 2.5. Görüntü etiketleme paneli. ....	10
Şekil 2.6. Etiket türlerinin ve etiketlerin görüntülediği panel. ....	11
Şekil 2.7. Diş çürüğü lezyonlarının açıklamaları. ....	15
Şekil 2.8. Panoramik görüntülerin kırılması ve kategorilenmesi. ....	16
Şekil 3.1. R-CNN mimarisinin nesne algılama sistemi. ....	18
Şekil 3.2. Fast R-CNN mimarisi çalışma prensibi. ....	18
Şekil 3.3. Faster R-CNN mimarisinin bileşenleri. ....	19
Şekil 3.4. Mask R-CNN mimarisinin akış şeması. ....	21
Şekil 4.1. GoogLeNet mimarisinin katmanları. ....	24
Şekil 4.2. ShuffleNet mimarisinin katmanları ve kapsamlı görünümü. ....	25
Şekil 4.3. ShuffleNet bileşenleri. a) DWConv darboğaz ünitesi; b) GConv ve kanal karıştırmalı bir ShuffleNet birimi; c) iki adımlı ünite. ....	26
Şekil 4.4. Artık blok gösterimi. ....	27
Şekil 4.5. MobileNet bloğu. ....	28
Şekil 4.6. VGG ağ mimarisi. ....	30
Şekil 4.7. DenseNet katmanlar arası aktarım. ....	30
Şekil 4.8. ConvNeXt katmanları. ....	32
Şekil 4.9. Önerilen yöntemin sınıflandırıcısı. ....	34
Şekil 5.1. IOU gösterimi. ....	36

## ÇİZELGELER DİZİNİ

### Sayfa

Çizelge 5.1. Panoramik görüntülerdeki diş çürüklerinin örnek bölütleme sonuçları.	38
Çizelge 5.2. Kırpılmış ve uzman etiketli diş görüntülerindeki çürük lezyonlarının örnek bölütleme sonuçları. ....	39
Çizelge 5.3. Öz-dikkat tabanlı yöntem sonuçları. ....	40
Çizelge 5.4. K-katman çapraz doğrulama sonuçları. ....	40
Çizelge 5.5. %30 çürük, %70 çürük olmayan sınıf oranlarından oluşan veri seti üzerinde sınıflandırma ağlarının performansı. ....	42
Çizelge 5.6. Sınıflandırma ağlarının sınıf oranları dengeli veri seti üzerindeki performansı. ....	43
Çizelge 5.7. %70 çürük, %30 çürük olmayan sınıf oranlarından oluşan veri seti üzerinde sınıflandırma ağlarının performansı. ....	44
Çizelge 5.8. %85 çürük, %15 çürük olmayan sınıf oranlarından oluşan veri seti üzerinde sınıflandırma ağlarının performansı. ....	45
Çizelge 5.9. Sınıflandırma ağlarının kırpılmış diş görüntülerinden oluşan ana veri seti üzerinde performansı. ....	46

## SİMGELER VE KISALTMALAR DİZİNİ

### SİMGELER

- $p$  : başlangıç noktası  
 $q$  : sorgu noktası  
 $r$  : bitiş noktası  
 $n$  : matris boyutu  
 $y_{min}$  : minimum y noktası  
 $x_{min}$  : minimum x noktası  
 $y_{max}$  : maksimum y noktası  
 $x_{mx}$  : maksimum x noktası  
 $j$  : matris indisi  
 $k$  : matris indisi  
 $w$  : ağırlık matrisi  
 $x_{in}$  : giriş vektörü  
 $f$  : aktivasyon fonksiyonu  
 $w_0$  : sapma değeri  
 $r^{(l)}$  : bernoulli rasgele değişkeni  
 $\tilde{y}$  : eleman bazında çarpım sonuçları  
 $l$  : katman  
 $z$  : bir sonraki katman girdileri  
 $y_{out}$  : ağ çıktı vektörü

## KISALTMALAR

BERT	: Bidirectional Encoder Representations from Transformers (Dönüştürücülerden Çift Yönlü Enkoder Gösterimleri)
CNN	: Convolutional Neural Networks (Evrışimli Sinir Ağları)
DWConv	: Depthwise Convolution (Derinlemesine Evrişim)
Faster R-CNN	: Faster Regions With CNN Features (Daha Hızlı Bölge Tabanlı Evrişimli Sinir Ağı)
Fast R-CNN	: Fast Regions With CNN Features (Hızlı Bölge Tabanlı Evrişimli Sinir Ağı)
FDI	: Fédération Dentaire Internationale (Uluslararası Dişhekimliği Federasyonu)
FN	: False Negative (Yanlış Negatif)
FP	: False Positive (Yanlış Pozitif)
FPN	: Feature Pyramid Network (Öznitelik Piramit Ağı)
FReLU	: Flexible Rectified Linear Unit (Esnek Doğrultulmuş Lineer Birim)
GConv	: Grouped Convolution (Grup Evrişimi)
GeLU	: Gaussian Error Linear Unit (Gauss Hatası Doğrusal Birimi)
GPT	: Jeneratif Önceden Eğitilmiş Dönüştürücü (Generative Pre-Trained Transformer)
IOU	: Intersection Over Union (Birleşim Üzerinden Kesişim)
Mask R-CNN	: Mask Regions With CNN Features (Maske Bölge Tabanlı Evrişimli Sinir Ağı)
ReLU	: Rectified Linear Unit (Doğrultulmuş Lineer Birim)
ROI	: Region Of Interest (İlgi Bölgesi)
RPN	: Region Proposal Network (Bölge Teklif Ağı)
R-CNN	: Bölge Tabanlı Evrişimli Sinir Ağı (Regions With CNN Features)
SVM	: Support Vector Machine (Destek Vektör Makinesi)
TN	: True Negative (Doğru Negatif)
TP	: True Positive (Doğru Pozitif)
ViT	: Vision Transformer (Görüntü Dönüştürücü)

## BÖLÜM 1

### GİRİŞ

Ağız içinde bulunan çeşitli besin atıkları ve bakterilerin oluşturduğu asitlerden kaynaklı olarak dişin mineral dokusu bozulabilmektedir. Bu bozulma sonucunda, dişin çiğneme yüzeylerinden köklere doğru olan her bir bölgesinde çürük lezyonları oluşabilir. Diş çürüğü, herkesin hayatında en az bir kere karşılaştığı yaygın bir dental problemdir [1]. Diş çürüğü yaygın bir dental lezyon olmasına karşın erken teşhisi sayesinde tedavi edilerek önlenmektedir [2]. Diş üzerinde oluşan çoğu lezyon gibi diş çürükleri de uzmanlar tarafından görsel muayene ile teşhis edilmektedir. Fakat dişin derin kısımlarında oluşan çürük lezyonları, görsel muayene sırasında tespit edilememektedir [3]. Gözle görülemeyen lezyonların incelenmesi ve analizi için nesnelere anatomik yapılarıyla birlikte sunabilen X-ışını radyografisi kullanılmaktadır [4]. Radyografiler, farklı ışın tekniklerini kullanabildiği ve bu sayede nesnelere derinlemesine incelenmesini sağladığı için dental muayenelerde sıklıkla kullanılmaktadır [5]. Dişlerin farklı açılardan ve farklı görünümünü içeren bitewing, panoramik, periapikal ve sefalometrik gibi pek çok radyografi türü bulunmaktadır [6]. Diş çürüğü lezyonlarının radyografik görüntüler üzerinde tespiti gerçekleştirilirken çevresel koşullar tespit sürecini olumsuz etkileyebilmektedir. Ayrıca çoğu durumda diş çürükleri gözden kaçabilmektedir. Çoğu alanda da olduğu gibi diş çürüklerinin sınıflandırılması ve tespiti için de derin öğrenme modelleri kullanılmaktadır. Derin öğrenme modelleri, görüntülerdeki nesne tespit ve görüntü sınıflandırma problemlerine çözüm sunabilmektedir.

Çalışmanın akademik temeller ışığında sunulabilmesi için öncelikle benzer literatür çalışmaları incelenmiştir. Diş çürüğü lezyonlarının radyografiler üzerinde tespiti, bölütlenmesi ve sınıflandırılması konusunda yıllar içinde birçok çalışmanın yapıldığı görülmektedir. Mertens vd. [7], bitewing radyografilerinde çürük tespiti için tamamen evrişimli bir ağ kullanarak %94,00 oranında doğruluğa ulaşmıştır. Cantu vd. [8],



bitewing radyografiler üzerinde çürük lezyonlarının tespiti ve bölütlenmesi için U-Net [9] ve EficientNet-B5 [10] mimarilerini kullanmıştır. Bu mimari içerisinde, EfficientNet-B5 ile U-Net entegrasyonunu gerçekleştirilmiştir. Ayrıca son katman olan tam bağlı katman (fully connected) kaldırılmıştır. Bu entegrasyonunun eğitimi için panoramik görüntülerde daha önce eğitilmiş model ağırlıkları kullanılmıştır. Bayrakdar vd. [11] ve Lee vd. [12], diş çürüğü lezyonlarını bölütleyebilmek için U-Net mimarisini bitewing görüntüler üzerinde başarılı bir şekilde eğitmişlerdir. Bayrakdar vd. U-Net kullanarak %81,00 duyarlılık, %86,00 kesinlik ve %84,00 oranında f1-skoruna ulaşmıştır. Lee vd. ise %63,29 kesinlik, %65,02 %64,14 oranında f1-skoruna ulaşmıştır. Ayrıca, Bayrakdar vd. sırası ile %84,00, %84,00 ve %84,00 oranında duyarlılık, kesinlik ve f1-skoru performansı ile diş çürüğü lezyonlarını tespit etmek için VGG-16 [13] ağını da başarılı bir şekilde eğitmiştir. Bu çalışma içerisinde test kümesi ile klinik ortamdaki veriler arasındaki seçim ön yargısından kaynaklı performans düşüşünü önlemek için çoklu çapraz doğrulama evrişimi kullanılmıştır. Khan vd. [14] ve Lian vd. [15] periapikal ve panoramik görüntüler üzerinde çürük lezyonlarının oluşum ve ilerleme aşamalarını belirleyerek diş çürüklerini, alveoler kemik durgunluğunu ve interradiküler radyolusensi U-Net ve DenseNet [16] mimarileri ile bölütlendirip sınıflandırmışlardır. Zhu vd. [17], temelinde PraNet [18], U-Net ve Res-U-Net [19] mimarilerini kullanarak panoramik görüntülerindeki çürükleri tanıyabilen CariesNet'i sunmuştur. Farklı mimarileri temel alarak oluşturulan CariesNet, çürük tanıma görevini %93,64 zar katsayısıyla (dice coefficient) ve %93,61 doğrulukla gerçekleştirmiştir.

Öz-dikkat tabanlı yöntemler son zamanlarda derin öğrenme araştırmalarına konu olmaktadır. Jiang vd. [20] RDFNet mimarisini önerdi ve dijital görüntülerdeki çürük lezyonlarını hafif, orta ve derin düzeyinde ele alarak sınıflandırabilen bir mimari geliştirmişlerdir. RDFNet, %62,30 doğruluk, %56,90 duyarlılık ve %57,90 oranında skoruna ulaşarak Maske Bölge Tabanlı Evrişimli Sinir Ağı (Mask Regions with CNN features Mask R-CNN) [21], Daha Hızlı Bölge Tabanlı Evrişimli Sinir Ağı (Faster Regions with CNN features, Faster R-CNN) [22] ve YOLO [23] mimarisinin sürümlerinden daha iyi bir performans göstermiştir. Ayrıca RDF-Net mimarisine bir dönüştürücü (transformer) mekanizması ve Esnek Doğrultulmuş Lineer Birim (Flexible Rectified Linear Unit, FReLU) [24] dahil edilerek %1,6 oranında bir

iyileşme sağlanmıştır. Che ve Hu [25], öz-dikkat mekanizmasının ve dönüştürücülerin birleşiminden bir çürük lezyon tespit modeli önermiştir. Bu çalışmada özneliklerin daha iyi temsil edilmesi için yeni bir yaklaşıma sahip dikkat mekanizması sunulmaktadır. Sunulan model %63,81 oranında bir ortalama kesinlik değerine ulaşarak radyografik görüntülerde etiketlenmemiş çürük lezyonlarının tespitini bile mümkün kılmaktadır. Zhou vd. [26], panoramik radyografilerinde süt dişlerini tanıyabilecek Swin Dönüştürücüsü (Swin Transformer) [27] tabanlı çalışan bir derin öğrenme sistemi önerdi. Bu çalışma ile birlikte dönüştürücü modellerinin diş çürüğü lezyonlarını tanıma görevinde başarılı olabileceği gösterilmiştir. Bu model ile %85,57 doğruluk, %88,32 kesinlik, %83,17 duyarlılık ve %85,67 f1-skoru performansına ulaşılmıştır. Hossain vd. [28], bir akıllı telefon yardımı ile elde edilen diş görüntüleri üzerinde bulunan diş çürüğü lezyonlarını sınıflandırmak için bir Görüntü Dönüştürücü (Vision Transformer, ViT) [29] kullanmıştır. Görüntüleri çürük yok, erken safhadaki çürük ve ileri seviyedeki çürük olarak sınıflandıran bu sistem sırası ile bu sınıflarda %95,00, %91,00 ve %100,00 oranlarında duyarlılığa ulaşmıştır.

Bu tez çalışmasının amacı panoramik görüntülerde çürük lezyonlarını algılayabilen ve klinik ortamda uygulanabilen derin öğrenmeye dayalı tam otomatik bir sistem geliştirmektir. Tez çalışmasının temelini oluşturulan veri seti, diş bölgeleri ve çürük bölgeleri olarak iki sınıfta etiketlenmiştir. Bu çalışmada kullanılan veri seti farklı oranlarda, farklı sınıf dağılımlarında, etiketli ve etiketsiz olmak üzere bölünerek sonuçlar elde edilmiştir. Bölünen veri setlerine göre sınıflandırma ve nesne tespit görevleri gerçekleştirilmiştir ve çalışma oluşturulan alt veri setlerine göre üç bölüme ayrılmıştır. 1) Uzman etiketlerine sahip panoramik görüntüler Mask R-CNN mimarisine girdi olarak verilerek, örnek bölütleme görevi gerçekleştirilmiştir. 2) Uzman etiketlerine sahip geniş panoramik görüntülerdeki sadece çürük içeren dişler bireysel olarak kırılarak çürük etiketleri ile birlikte Mask R-CNN mimarisi eğitilmiştir. 3) Panoramik görüntülerdeki her bir diş kendi süper kategorisinin koordinatları ele alınarak sınırlayıcı kutularından kırıldı. Kırılan her bir görüntü uzman etiketleri bilgisine göre çürük ve çürük değil sınıflarına ayrıldı. Elde edilen ham veri farklı sınıf oranlarına bölünerek sınıflandırıcılara girdi olarak sunuldu. Literatürde sunulan son teknoloji derin öğrenme sınıflandırıcıları temel alınarak diş çürüğü

lezyonlarını algılayabilen bir sınıflandırıcı önerildi. Aynı zamanda, bu kırılmış veriler öz-dikkat mekanizması içeren bir derin öğrenme modeli ile de eğitildi.

Tez çalışması, beş ana bölüm altında sunulmaktadır. İlk bölümde panoramik radyografi görüntülerindeki çürükleri tanıyabilen derin öğrenme modelleri ile ilgili çalışmalardan oluşan literatür aşaması sunulmaktadır. İkinci bölümde diş çürüğü problemi, diş çürüklerinin incelenmesinde kullanılan görüntü tipleri, diş çürüklerinin klinik ortamda teşhisi anlatılmaktadır. Aynı zamanda bu bölümde tez çalışmasının temelini oluşturan veri kümesi, veri kümesinin elde edilmesi ve bu verilerin uzmanlar tarafında anlamlandırılması ile ilgili bilgiler verilmektedir. Üçüncü bölümde görüntü bölütleme ve nesne tespit mimarileri sunulmakta ve bölge tabanlı yaklaşımlar anlatılmaktadır. Dördüncü bölümde, görüntü sınıflandırma yöntemleri anlatılmaktadır. Evrişimli Sinir Ağları (Convolutional Neural Networks, CNN) [30], ön eğitilmiş CNN'ler ve öz-dikkat tabanlı yöntemler bu bölümün konusunu oluşturmaktadır. Bunların yanı sıra bu bölümde, derin öğrenme mimarilerinden esinlenerek tasarlanan yeni bir sınıflandırıcı sunulmaktadır. Bölüm beşte ise deneysel çalışmalar sunulmaktadır. Nesne tespit modellerinin, CNN'lerin ve öz-dikkat tabanlı derin öğrenme yöntemlerinin diş çürüğü lezyonunu tanıma sonuçları bu bölümde verilmektedir.

## BÖLÜM 2

### DIŞ ÇÜRÜĞÜ

Son yıllarda derin öğrenme modelleri, diş çürüğü teşhisi aşamalarında önemli bir rol oynamaktadır [31]. Diş çürüğü, çoğunlukla yemek artıklarının ağızda kalması sonucuyla ağız içinde çoğalan çeşitli bakteriler sonucunda diş üzerinde oluşan bir lezyondur. Diş çürüğü lezyonları oldukça yaygındır ve yapılan son çalışmalara göre 2,3 milyar insan bu lezyonlarla karşılaşabilmektedir [32]. Diş çürüklerinin erken teşhisi ve tedavisi diş kaybı gibi durumların oluşmaması için oldukça önemlidir. Derin öğrenme yaklaşımları çevresel koşullardan etkilenmediği için çürük teşhisinde doğruluk ve hassasiyeti artırabilmektedir. Öte yandan, teşhise kattığı hız sayesinde klinik ortamda uzmanı destekleyebilmektedir. Derin öğrenme yöntemleri, büyük miktarda diş çürüğü verisini analiz edip işleyebilmektedir ve tanıyabilmektedir. Bu yöntemler, klinik ortamda dişlerin görsel muayenesi için kullanılan radyografik görüntüleri girdi olarak almaktadır. Çok katmanlı yapılara sahiptir ve farklı yaklaşımlar ile görüntüdeki nesnelere derinlemesine inceleyebilmektedir. Bir derin öğrenme modeli, radyografik görüntüleri alır ve düşük etkili öznelikleri tanıyarak yüksek özelliklere sahip olan çürük bölgelerini tanıyabilir. Çürük teşhisi ve tedavi planlamasının derin öğrenme yöntemleri ile desteklenebilmesi sayesinde uzmanlara daha fazla hız ve kaynak sağlanarak hasta bakım olanakları iyileştirilebilmektedir.

#### 2.1. DIŞ GÖRÜNTÜLERİ VE GÖRÜNTÜ TIPLERİ

Diş çürüğü teşhisinde radyografik görüntülerden faydalanılmaktadır. Radyografi, röntgen ışınları aracılığıyla nesnelere iç anatomisini incelemeye yarayan görüntülerdir. Radyografiler, X-ışınlarını kullanarak ağız içerisinde bulunan bütün yapıların görünümünü sağlayabilmektedir. Bu tarz görüntüleri elde edebilmek için radyografik cihazlar kullanılmaktadır. Bu cihazlar düşük dozda X-ışını üretebilen tüpler içermektedir. Bu tüpler tarafından üretilen radyasyon film dedektörü sayesinde

görüntüye dönüştürülür. Radyografide dijital görüntü oluşturabilmek için ise bu film dedektörleri yerine dijital dedektörler kullanılmaktadır [33,34]. Radyografi çekim sürecinde, görüntüde bozuklukların görülmemesi için hastanın belirli pozisyona alması oldukça önemlidir. Bu yüzden bu cihazların kurulduğu ortamlarda hastayı kısıtlayabilen çeşitli aparatlar ve destekler bulunmaktadır. Radyografilerin çekim aşamasında, düşük düzeyde iyonlaştırıcı kullanılmaktadır [35]. Bu sayede çekim yapan kişi ve hastaların radyasyon maruziyeti azaltılmaktadır.

Radyografik görüntülerin birçok kullanım alanı bulunmaktadır. Bu kullanım alanlarına sağlık başta olmak üzere endüstri, güvenlik, sanat ve tarım gibi alanlar örnek verilebilir [36]. Radyografinin, iskelet sistemi lezyonları, iç organlarda oluşabilecek durumlar, ağız içi problemleri gibi yapıları sunabilmesinden dolayı tıp biliminin pek çok branşında kullanımı bulunmaktadır. Pek çok radyografi görüntüsü farklı amaçlar için kullanılmaktadır. Diş hekimliğinde radyografik görüntüleme süreci ağız içi ve ağız dışı olmak üzere ikiye ayrılmaktadır. Ağız içi ve ağız dışı görüntülemeler sonucu elde edilen görüntüler bitewing, periapikal, dijital, panoramik, tomografik, sefalometrik ve oklüzal radyografi olarak ele alınmaktadır.

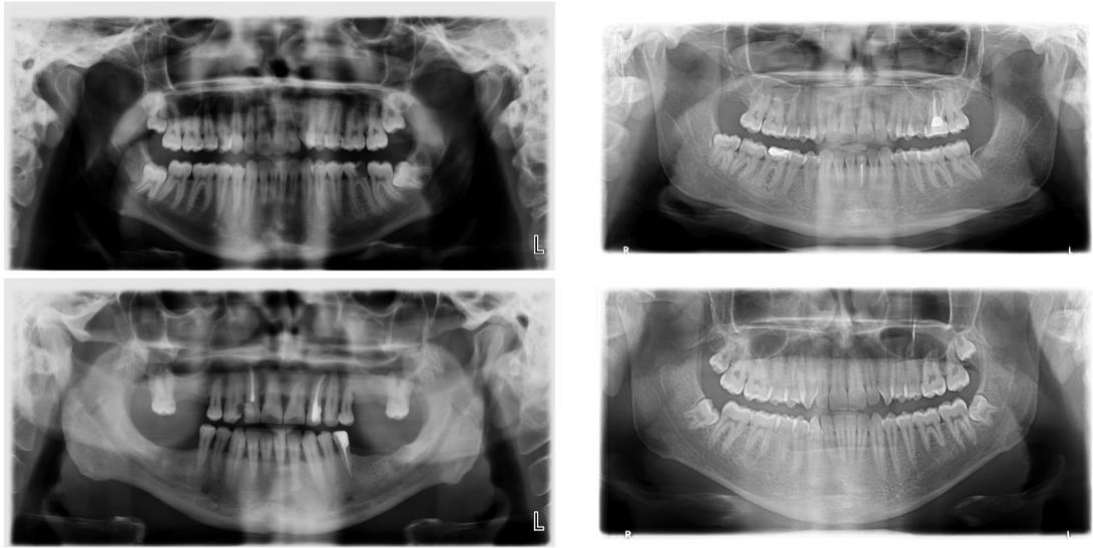
## **2.2. DIŞ ÇÜRÜĞÜNÜN KLİNİK ORTAMDA TEŞHİSİ**

Radyografik görüntüler, diş çürüğü lezyonlarının teşhisinde yaygın bir kullanıma sahiptir. Diş çürüğü tespitinde çoğunlukla uzmanlar bitewing radyografileri tercih etmektedir. Bitewing radyografiler dişlerin üzerindeki bölgeleri daha yakından görüntüleyerek daha kaliteli görüntüler sunabilmektedir. Fakat bitewing görüntüler diş çürüğünün görsel analizinde sıklıkla tercih edilse de hasta sağlığı ele alındığında panoramik diş görüntülerine kıyasla daha dezavantajlıdır. Bitewing görüntülerin çekim sürecinde farklı koşullardan dolayı tekrar çekim gerekebilmektedir. Bu durumda hasta daha çok radyasyona maruz kalmaktadır ve sağlık açısından oldukça sakıncalıdır. Özellikle, özel hasta grubundaki kişilerin çekim süreçleri pozisyon korumasının sağlanamadığından oldukça zor olduğu için tekrar çekim gerçekleştirme oranı oldukça yüksektir. Ayrıca, bitewing görüntüler dişlerin aralarında bulunan çürüklerin analizinde etkili olmasına rağmen dişlerin kök bölgelerinde, çiğneme yüzeylerinde ve farklı açılardaki görünümde kalan çürükleri algılamada dezavantajlı olabilmektedir.

Bitewing diř görüntüleri, diřlerin yatay olarak düzlemini ele alır ve tek bir kesit sağlayabilir bu yüzden de farklı diř yüzeylerini ve diř bölgelerinin boyutlarını sunamaz. Bu gibi durumlarda daha fazla perspektif ve detay sunabilen panoramik radyografi görüntüleri tercih edilmektedir. Panoramik radyografiler, ağızın geniş bir görünümünü sunar. Bu görünüm içerisinde çene kemiđi, sinüs boşluđu dikeyde sınırları oluştururken yatayda ise bütün çene kemiđini kapsar. Panoramik görüntü, daha çok detay ve sađlık açısından avantajı beraberinde getirmesine rağmen, düşük kalitedeki görüntüler olduđu için derin öğrenme modellerinin uygulanabilmesi oldukça zordur.

### 2.3. DIř ÇÜRÜĐÜ VERİ KÜMESİ

Çalıřmada kullanılan veriler, Karabük Ağız ve Diř Sađlığı Eđitim ve Arařtırma Hastanesi'nden gerekli izinler alınarak elde edilmiřtir. Çalıřmada kullanılan bütün görüntüler hasta bilgilerinden arındırılarak anonimleřtirilmiřtir. řekil 2.1'de diř çüğürü veri kümesine ait örnek görüntüler bulunmaktadır.



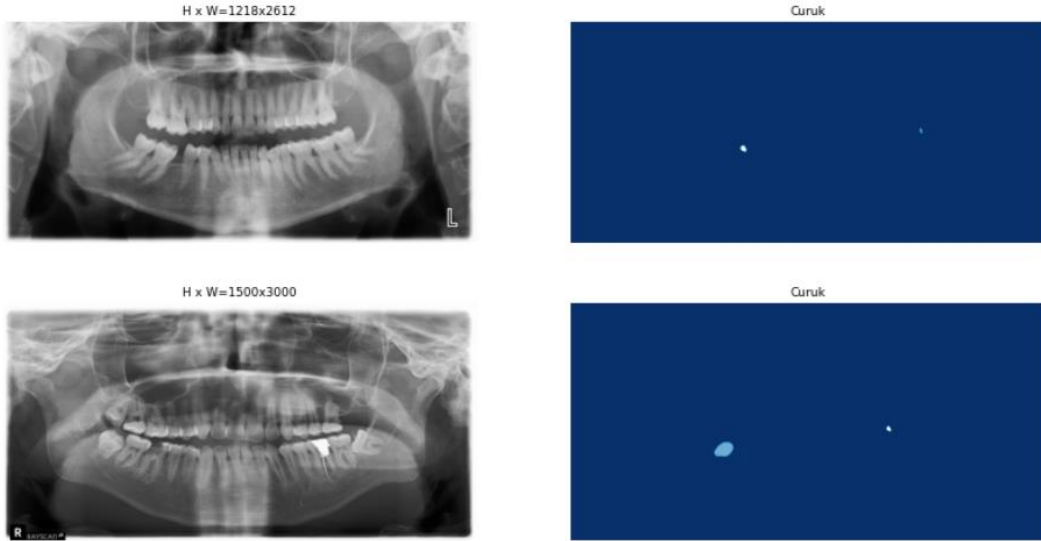
řekil 2.1. Veri kümesinde bulunan radyografi örnekleri.

Veri kümesindeki görüntüler I-Max Touch Owandy Radyoloji cihazı tarafından yakalanmıřtır. Bu radyoloji cihazının 60-80 kV oranında tüp voltajı bulunmaktadır. Tüp akımı ise 6-12 mA deđerindedir. Veri kümesi üzerinde veri artırımı, veri kırpma

ve yeniden boyutlandırma gibi çeşitli ön işleme adımları uygulanmıştır. Bu sayede veriler derin öğrenme yöntemlerinin eğitimi için uygun durumlara getirilmiştir.

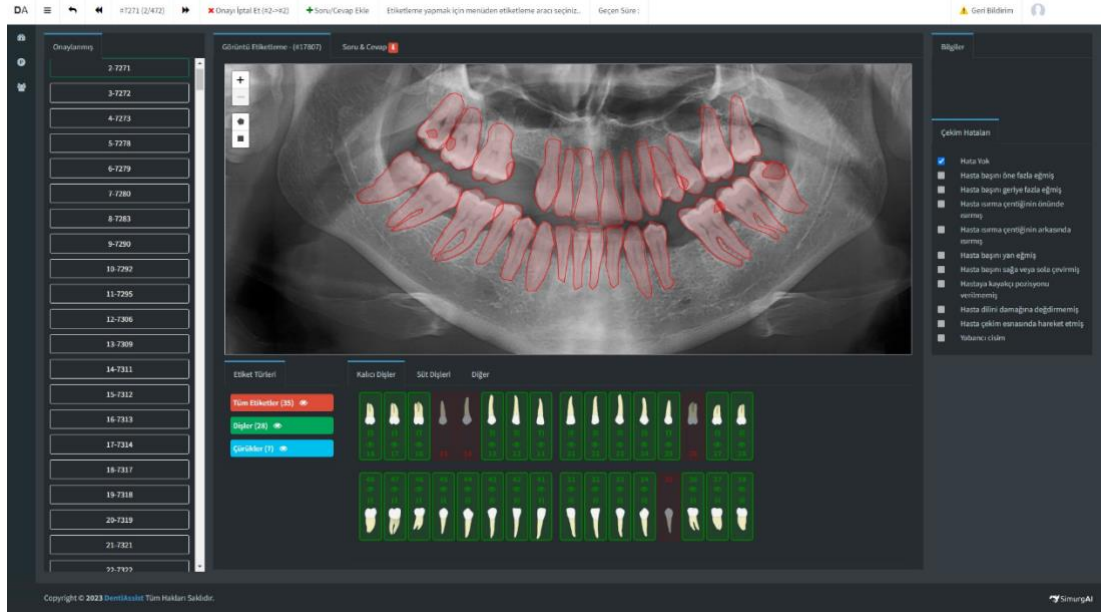
### 2.3.1. Verilerin Etiketlenmesi

Elde edilen verilerin etiketlenmesi tespit, sınıflandırma ve bölütleme çalışmaları için oldukça önemlidir. Verilerin etiketlenmesi radyografi uzmanları tarafından el ile gerçekleştirilmiştir. Diş bölgeleri ve bu diş bölgeleri içerisinde bulunan çürük alanları etiketlenmiştir. Bu sayede hem diş nesnelerin süper kategorileri hem de her bir çürük bölgesi için poligon açıklamaları oluşturulmuştur. Uzman etiketleriyle oluşturulan veri setine ait örnek görüntüler ve çürük maskeleri Şekil 2.2’de gösterilmektedir.



Şekil 2.2. Panoramik görüntüler ve çürük etiketleri.

Uzmanlar veri etiketleme süreçlerini DentiAssist [37] yazılımını kullanarak gerçekleştirmişlerdir. DentiAssist platformu, pek çok diş görüntüsü üzerinde etiketleme ve bu etiketlere sınıf atama özelliği sunmaktadır. DentiAssist platformunun etiketleme arayüzü Şekil 2.3’te gösterilmektedir.



Şekil 2.3. DentiAssist görüntü etiketleme platformunun arayüzü.

Veri etiketleme sürecinde görüntülerdeki çekim hataları uzmanlar tarafından belirtilebilmektedir. Etiketleme platformunun sağ panelinde Şekil 2.4'te gösterildiği gibi hastanın pozisyonuna göre hatalar belirtilebilmektedir. Aynı zamanda hasta çekim sürecine izin verilmeyen yabancı cisimler ile girmesi ya da hastanın pozisyonu bozacak hareketler gerçekleştirilmesi gibi hatalar da uzmanlar tarafından bu platform üzerinde belirtilebilmektedir.



Şekil 2.4. Çekim hataları işaretleme paneli.



Veri etiketlemesi gerçekleştirilirken etiketlenecek bölgeye sınıf isimleri verilmektedir. Uzmanlar diş nesnelerini belirlemek için diş hekimliği alanında evrensel olarak kullanılan Uluslararası Dişhekimliği Federasyonu (Fédération Dentaire Internationale, FDI) notasyonunu kullanmışlardır. Bu sayede her bir diş çürüğü alını kapsayan diş bölgeleri belirlenebilmektedir. Kırpılarak oluşturulacak panoramik görüntüler için FDI notasyonu ile etiketlenmiş bölgelerin koordinatları referans olarak ele alınmaktadır. Şekil 2.5'te gösterildiği gibi platformun orta alanında bulunan görüntü etiketleme panelinden etiketler, koordinatlar seçilerek oluşturulabilmektedir. Poligon etiketlemeleri bu panel üzerinde gerçekleştirilebilmektedir.



Şekil 2.5. Görüntü etiketleme paneli.

Bütün diş etiketleri platformun alt panelinde görüntülenebilmektedir. Şekil 2.6'da etiketlerin sıralı bir yapıda görüntülenebildiği panel verilmektedir. Aynı zamanda bu panel üzerinde bulunan diş gösterimi sayesinde, ağızdaki eksik dişler de incelenebilmektedir. Her bir etiketin görünürlüğü bu panel üzerinden değiştirilebilmektedir. Burada yapılan görünürlük değişiklikleri etiketleme paneli üzerindeki etiket görünürlüğünü de etkilemektedir.



Şekil 2.6. Etiket türlerinin ve etiketlerin görüntülediği panel.

### 2.3.2. Panoramik Görüntülerin Kırılması

Geniş alanları içeren panoramik görüntüler otomatik bir sistem ile kırılmıştır. Kırpma işlemi gerçekleştirilirken çürükler bir dişin içinde olup olmama durumuna göre kontrol edilir. Çürük bölgesini kaplayan poligon nesnesinin diş bölgesi içinde olma durumunu kontrol eden algoritmanın adımları ilk olarak  $q$  olarak verilen bir noktanın  $p$  ve  $r$  noktaları arasında olup olmadığını kontrol eden bir fonksiyon ile başlar. Bu fonksiyona gönderilen nokta  $x$  ve  $y$  koordinatları ile birlikte 2 boyutlu bir matris olarak gönderilen bir nokta  $p$  ve  $r$  noktalarının  $x$  ve  $y$  koordinatlarına göre kontrol edilir. Eğer nokta  $p$  ve  $r$  noktalarının arasında ise fonksiyon doğru, eğer arasında değilse fonksiyon yanlış değerini döndürür. Bir noktanın verilen diğer noktaların arasında mı olduğunu kontrol edene bu fonksiyonun algoritma adımları aşağıdaki gibidir:

fonksiyon arasındaMı( $p, q, r$ ):

eğer ( $q[0]$  küçük eşitse  $\max(p[0], r[0])$  ve

$q[0]$  büyük eşitse  $\min(p[0], r[0])$ ) ve

( $q[1]$  küçük eşitse  $\max(p[1], r[1])$  ve

$q[1]$  büyük eşitse  $\min(p[1], r[1])$ ):

döndür Doğru

döndür Yanlış

Bir noktanın diğer noktalar arasında mı olduğu kontrol edildikten sonra, bu noktaların birbirleri ile durumlarına bakılmaktadır. Oryantasyon fonksiyonu sayesinde kendisine parametre olarak gönderilen üç noktanın birbirlerine göre konumları incelenebilmektedir. Öncelikle bu üç noktanın sırası ile bir çizgi üzerinde mi olduğu kontrol edilir. Her bir nokta birbirleri ile aynı doğru üzerindeyse yani doğrusal bir konuma sahip ise fonksiyon 0 değerini göndermektedir. Eğer fonksiyona parametre olarak gönderilen noktalar bir çizgi üzerinde değilse, bu noktaların saat yönünün hangi tarafında olduğu kontrol edilir. Bu sayede noktaların birbirlerine göre yönü hesaplanabilmektedir. Bu ifadelerin farkı 1 ise bu üç nokta aynı saat yönüne doğru bir yapı içerisindedir. Eğer değer 2 ise noktaların yönü saat yönünün tersine doğru ilerlemektedir. Bu kontrolleri gerçekleştiren oryantasyon fonksiyonunun adımları aşağıdaki gibidir:

```

fonksiyon oryantasyon(p, q, r):
    deger = ((q[1] - p[1]) * (r [0] - q[0])) -
            ((q[0] - q[0]) * (r[1] - q [1]))
    eğer deđer == 0 ise:
        döndür 0
    eğer deđer > 0 ise:
        döndür 1
    aksi halde:
        döndür 2

```

Kesişim fonksiyonu,  $(p1, q1)$  ve  $(p2, q2)$  noktalarından oluşturulan çizgilerin birbiri ile kesişip kesişmediğinin değerini döndürmektedir. Her bir noktanın birbiri ile oluşturduğu alanlar hesaplanmaktadır. İki nokta arasındaki en kısa bağlantıyı tanımlayan doğru parçaları olan çizgi bölütlerinin birbirlerine göre konumları ve yönleri belirlenmektedir. Eğer ikili nokta kombinasyonları birbirinden farklı ise  $(p1, q1)$  çizgi bölütü,  $(p2, q2)$  çizgi bölütleri birbirlerini kesmektedir. Bu durum özel durum olarak belirtilmektedir. Eğer her bir nokta birbiri ile doğrusal bir yönde ise ve parametre olarak gönderilen nokta diğer iki noktanın arasındaysa o zaman fonksiyon kesiştiğine dair bir çıktı üretir. Bu şartları sağlayan durumların hepsi özel durum olarak

belirtilir. Genel ve özel durumların dışında kalan durumlarda ise kesişim sağlanmaz ve fonksiyon yanlış değerini üretir. Bu şekilde çizgi bölütlerinin kesişimini kontrol eden fonksiyonun algoritma adımları aşağıdaki gibidir:

```
fonksiyon kesişim( $p1, q1, p2, q2$ ):  
    o1 = oryantasyon( $p1, q1, p2$ )  
    o2 = oryantasyon( $p1, q1, q2$ )  
    o3 = oryantasyon( $p2, q2, p1$ )  
    o4 = oryantasyon( $p2, q2, q1$ )  
  
    # Genel durum  
    eğer (o1 farklı o2) ve (o3 farklı o4):  
        döndür Doğru  
  
    # Özel Durumlar  
    eğer (o1 == 0) ve arasındaMı( $p1, p2, q1$ ):  
        döndür Doğru  
    eğer (o2 == 0) ve arasındaMı( $p1, q2, q1$ ):  
        döndür Doğru  
    eğer (o3 == 0) ve arasındaMı( $p2, p1, q2$ ):  
        döndür Doğru  
    eğer (o4 == 0) ve arasındaMı( $p2, q1, q2$ ):  
        döndür Doğru
```

$p$  noktasının, noktalar listesinden oluşturulan bir poligonun içinde olup olmadığı kontrol edilmektedir. Noktalar olarak adlandırılan matrisin boyutu  $n$  değerine kaydedilmektedir. Bu sayede poligonun kaç köşesi olduğu tutulmaktadır. Eğer noktaların uzunluğu 3'ten daha küçük ise fonksiyon yanlış değerini döndürür. Çünkü bir poligon en az üç köşelidir ve bu durum fonksiyonun hatasız çalışması için temel durum olarak kontrol edilmektedir. Daha sonrasında araştırdığımız  $p$  noktası arasında çizgi bölütleri oluşturmak için sonsuz uzunluktaki bir nokta oluşturulmaktadır. Çizgi bölütleri arasında kaç tane kesişme olduğunu tutabilmek için bir sayaç

kullanılmaktadır. Mevcut noktaların bir sonraki nokta koordinatlarıyla oluşturulan çizgi bölütlerinin  $p$  noktası ile kesişim kontrolü gerçekleştirilmektedir. Bir sonraki noktayı tutabilmek için fonksiyonda sonraki olarak adlandırılan bir değer kullanılır. Eğer kesişim var ise ve bütün noktalar bir doğru üzerinde ise  $p$  noktasının bir sonraki ile mevcut nokta arasında mı olduğu kontrol edilir. Eğer bu kontrol sonucu doğru ise bu blok doğru değerini döndürür ve kesişim sayacı artırılır. Eğer döngü sayacı 0 ise çizgi bölütleri döngüsü tamamlanmıştır ve döngü durdurulur. Son olarak sayaç değerinin tek mi çift mi olduğu kontrol edilir ve eğer çift ise doğru değeri döndürülmektedir. Çift değilse de  $p$  noktasının poligonun dışında kaldığını belirten yanlış değeri döndürülür. Kendisine parametre olarak gönderilen bir noktanın poligon içerisinde olup olmadığını kontrol eden fonksiyonun adımları aşağıdaki gibidir:

fonksiyon poligonİçinde(noktalar,  $p$ ):

$n = \text{uzunluk}(\text{noktalar})$

eğer  $n < 3$  ise:

döndür Yanlış

sonsuz = (1000,  $p[1]$ )

sayaç = 0

$z = 0$

döngü:

sonraki =  $(z + 1) \% n$

eğer kesişim(points[i], noktalar[sonraki],  $p$ , sonsuz):

eğer oryantasyon(noktalar[i],  $p$ , noktalar[sonraki]) = 0 ise:

eğer arasındaMı(noktalar[i],  $p$ , noktalar[sonraki]):

döndür Doğru

sayaç += 1

$z = \text{sonraki}$

eğer  $z = 0$  ise:

dur

döndür(sayaç % 2 == 1)

Her bir çürük etiketinin koordinatları ve konumları, dışları kapsayan poligonların içinde mi yoksa dışında mı durumları kontrol edilmektedir. Her bir etiket minimum x

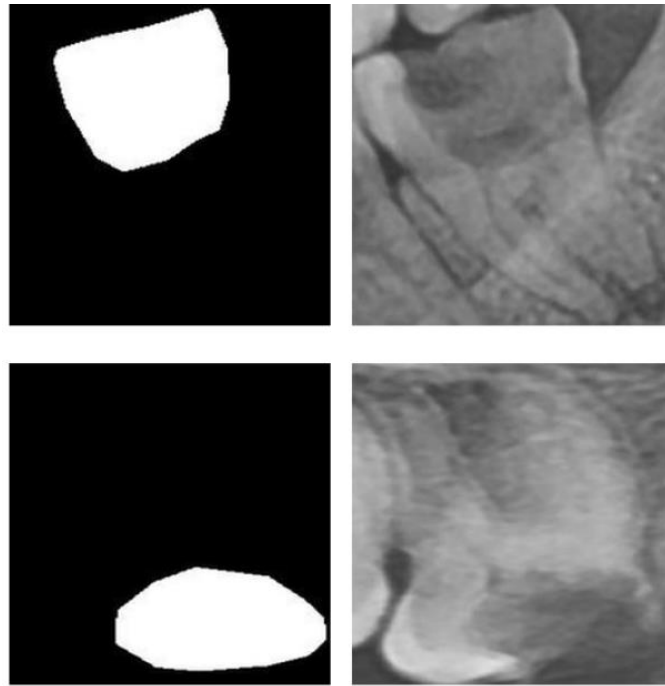
noktasını belirten  $x_{min}$ , minimum y noktasını belirten  $y_{min}$ , maximum x noktasını belirten  $x_{max}$  ve maksimum y noktasını belirten  $y_{max}$  değerleriyle tutulmaktadır. Dış etiketlerinin boyutları Eşitlik 2.1'deki gibi hesaplanmaktadır.

$$\text{genişlik } (W) = x_{max} - x_{min}; \text{yükseklik } (H) = y_{max} \quad (2.1)$$

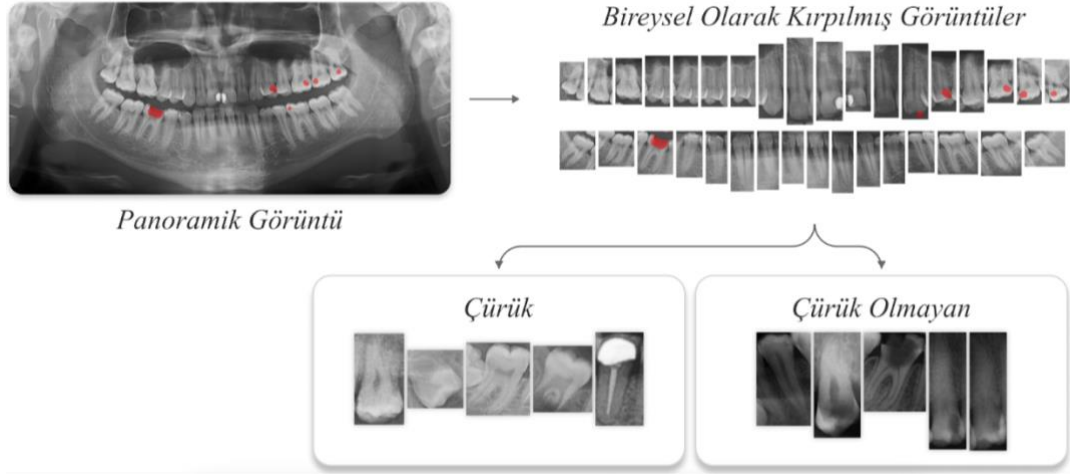
Çürük etiketleri oluşturulurken koordinat değerlerinde kayma gerçekleşmektedir. Bu kayma problemini çözebilmek ve dış etiketlerinin boyutuna sahip olan yeni görüntülerde dış çürüğü etiketlerini ölçeklendirebilmek için Eşitlik 2.2'deki bağıntı kullanılmaktadır.

$$\begin{aligned} \text{çürükEtiketi}(\tilde{x}_i, \tilde{y}_i) &= (\text{çürükEtiketi}(x_i) - \text{dışEtiketi}(x_i), \\ &\text{çürükEtiketi}(y_i) - \text{dışEtiketi}(y_i)) \end{aligned} \quad (2.2)$$

Bu işlem sayesinde dış görüntülerinin boyutlarından oluşan yeni görüntü üzerinde çürük etiketleri ölçeklendirilebilmektedir. Kırpma süreci sonucunda oluşturulan veri setindeki çürük etiketli örnek dış görüntüleri ve ikili maskeleri Şekil 2.7'de verilmektedir.



Şekil 2.7. Dış çürüğü lezyonlarının açıklamaları.



Şekil 2.8. Panoramik görüntülerin kırılması ve kategorilenmesi.

Sınıflandırma ağları çalışmasına dahil edilmek üzere önceki sürece ek olarak çürük ve çürük olmayan sınıflar kontrol edilerek ayrı kategorilere ayrıldı. Bu sayede toplamda 14498 görüntüden oluşan bir veri seti elde edildi. Kırma süreci ve görüntülerin sınıflara ayrılma süreci Şekil 2.8’de görselleştirilmiştir.

## BÖLÜM 3

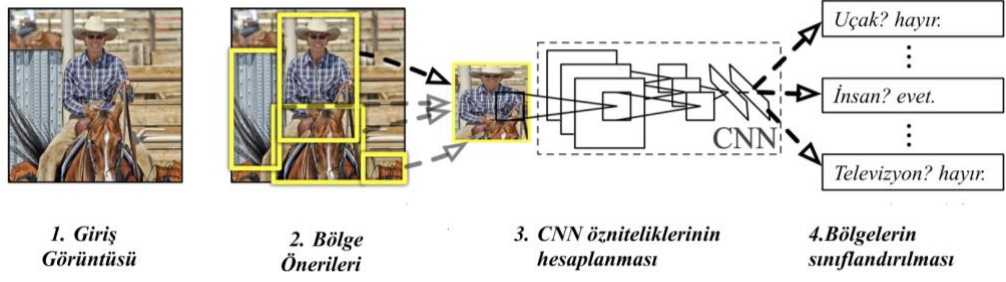
### GÖRÜNTÜ BÖLÜTLEME VE NESNE TESPİTİ

Nesne tespiti ve görüntü bölütleme, görüntüler üzerindeki nesneleri algılama ve işlemede önemli görevlerdir. Görüntü bölütleme, girişlerdeki farklı bölgeleri ayırarak, her bir bölgeyi temsil eden öznitelikleri tanıma işlevidir. Bu tanıma işlevi sayesinde bölge gruplandırmaları gerçekleştirilebilmektedir. Görüntü bölütleme, özel sınırlayıcı kutularla çalışabilen nesne tespit yaklaşımları için temel bir adım olarak literatürde sunulmaktadır [38]. Nesne tespitinin temel hedefi görüntülerdeki nesnelere koordinatlarını ve bu nesnelere sınıf etiketlerinin üretilmesidir [39].

#### 3.1. BÖLGE TABANLI EVRİŞİMLİ SINIR AĞLARI

Bölge Tabanlı Evrişimli Sinir Ağı (Regions with CNN features, R-CNN) [40], nesne tespit işlemini yürütebilmek için üç modüle sahip bir yapı kullanır. Bu modüllerden ilki, görüntüler üzerinde bölge önerileri gerçekleştirir. Bu modül bölge öneri modülü olarak adlandırılmaktadır. Bu öneriler R-CNN mimarisinin ilk modülü tarafından sınıflardan tamamen bağımsız olarak gerçekleştirilmektedir. Aynı zamanda, oluşturulan öneriler ile aday bölgeler tespit edilir. Bir diğer modül ise bu aday bölgelerden öznitelik çıkaran bir evrişim katmanıdır. Bu katman tamamen bir CNN yapısıyla çalışmaktadır. Son modül ise Destek Vektör Makinesi (Support Vector Machine, SVM) olarak sunulmaktadır. Bu modül tamamen doğrusaldır. Şekil 3.1’de R-CNN mimarisinin çalışma prensibi gösterilmektedir. R-CNN mimarisi görüntüde bulunan nesnelere önerilerle aday olarak belirler ve her bölge önerisi için öznitelik çıkarır. Ancak, R-CNN öznitelik çıkarma sürecini seçici arama (selective search) [41] algoritmasıyla yönettiği için bu süreç oldukça yavaştır.

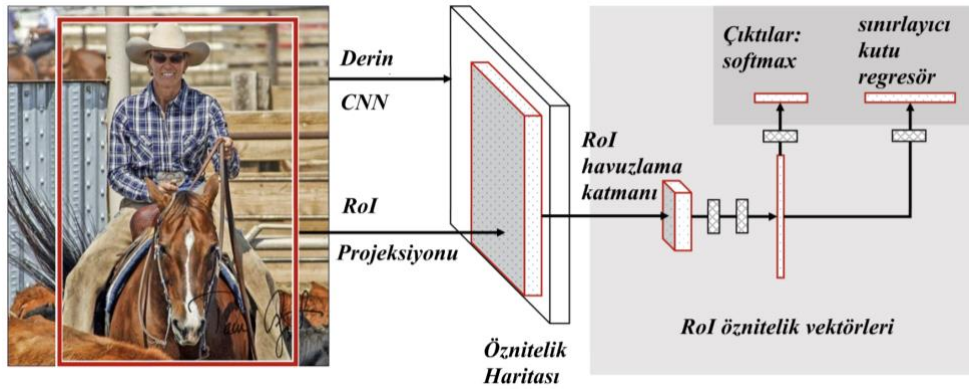




Şekil 3.1. R-CNN mimarisinin nesne algılama sistemi [40].

### 3.2. HIZLI BÖLGE TABANLI EVRİŞİMLİ SİNİR AĞLARI

Şekil 3.2’de gösterilen Hızlı Bölge Tabanlı Evrişimli Sinir Ağı R-CNN (Fast Regions with CNN features, Fast R-CNN) [42] mimarisi R-CNN nesne tespit modellerinin temel yaklaşımlarını kullanır. Fast R-CNN mimarisi, R-CNN modeline göre nesne tespit işlevinde daha hızlı bir performans göstermektedir [43]. R-CNN’ye kıyasla Fast R-CNN her aday bölge ile CNN’yi beslemek zorunda değildir. Bu işlev öznetelik haritaları ile sağlanır. Öznetelik haritaları sayesinde Fast R-CNN görüntü üzerinde her bir bölge önerisi için CNN katmanını tekrar tekrar kullanmaz. Bu sayede maliyet azaltılır ve hız artırılır. Aynı zamanda, Fast R-CNN, R-CNN mimarisine göre daha hızlı bir test performansı sunar.



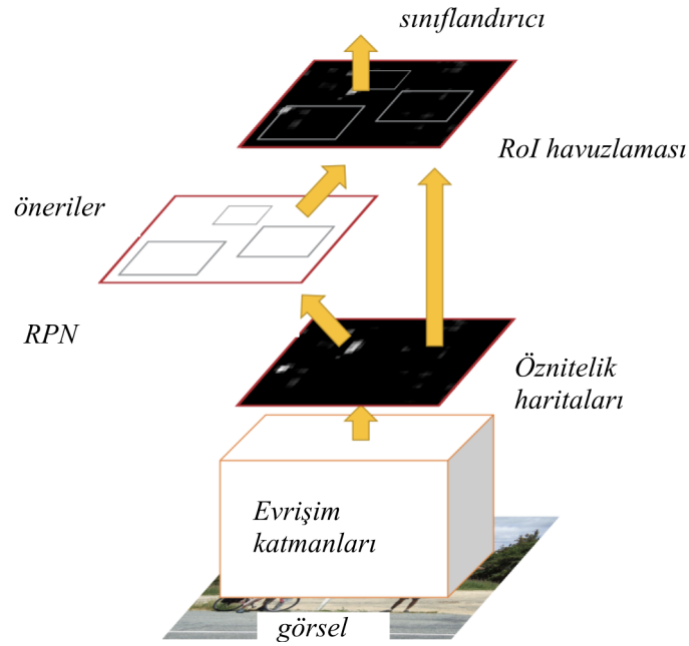
Şekil 3.2. Fast R-CNN mimarisi çalışma prensibi [42].

Fast R-CNN mimarisinde bölge önerilerinin sınırlayıcı kutuları, İlgi Bölge’leri (Region Of Interest, ROI) projeksiyonu referans alınarak öznetelik haritalarının boyutuna indirgenir. Bu indirgeme ROI havuzlama işlemi için kullanılmaktadır.

Havuzlama katmanı ile elde edilen öznitelik vektörleri sınırlayıcı kutu regresörleri kullanılarak düzenlenir.

### 3.3. DAHA HIZLI BÖLGE TABANLI EVRİŞİMLİ SİNİR AĞLARI

Faster R-CNN [22], seçici arama algoritmasının eğitim süresini yavaşlatmasına çözüm olarak sunulmuştur. Faster R-CNN mimarisinde girdiler sınırlayıcı kutular, etiketler ve olasılık değerleri olarak ele alınır. Girdi görüntüleri, önceden eğitilmiş bir CNN yapısından geçerek evrişimli öznitelik haritaları oluşturulur. Bu mimaride girdi görüntüleriyle beslenen CNN, otomatik öznitelik çıkarım işlemini yerine getirir. Bu yaklaşım literatürde öğrenme aktarımı (transfer learning) olarak adlandırılır. Öğrenme aktarımı büyük veri kümeleri üzerinde daha önce gerçekleştirilen eğitim sonucunda kaydedilen ağırlıkların farklı bir hedef için çalıştırılan sınıflandırıcı üzerinde kullanılmasını sağlar.



Şekil 3.3. Faster R-CNN mimarisinin bileşenleri [22].

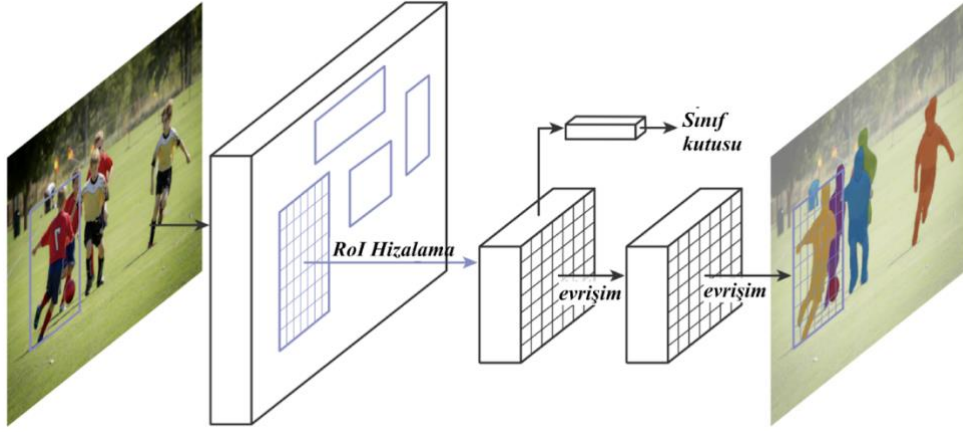
Şekil 3.3'te Faster R-CNN mimarisinin çalışma modeli gösterilmektedir. Faster R-CNN mimarisinde CNN eğitiminden elde edilen öznitelikler Bölge Teklif Ağı'na (Region Proposal Network, RPN) gönderilir. Bu sayede evrişimli katmanlar ile öznitelikler paylaşılarak maliyet önemli derecede azaltılmaktadır [44]. RPN, içerisinde

nesne bulunabilecek bölgeleri önerir ve bu bölgelere sınırlayıcı kutular atar. RPN yaklaşımında farklı boyutlarda sınırlayıcı kutular üretilir ve bu eğitim sürecini zorlaştırır. Bu sorunu çözebilmek için çapalar kullanılmaktadır. Çapalar sayesinde ilgili nesne içeren bölgeler ve bu bölgelerin nesnelere oturabilme düzeyi ele alınabilmektedir. RPN tarafından üretilen bölgeler ve bu bölgelerin sınırlayıcı kutuları üzerinde ROI havuzlama işlevi uygulanır. Son olarak ise ROI havuzlama sayesinde ROI dışındaki içerik arka plan olarak ağda ele alınır ve sınırlayıcı kutuların koordinatları üretilir.

### **3.4. MASKE BÖLGE TABANLI EVRİŞİMLİ SINIR AĞLARI**

Mask R-CNN [21] ağı, kendisinden önce gelen Faster R-CNN mimarisinin içeriğindeki bütün modülleri kapsayacak şekilde geliştirilmiştir. Mask R-CNN mimarisi örnek bölütleme işlemini verimli bir şekilde gerçekleştirebilmektedir [45]. Mask R-CNN mimarisi omurga ağı olarak ResNet kullanır ve bu şekilde otomatik öznitelik çıkarımını gerçekleştirir. Bu çıkarılan öznitelikler bir Öznitelik Piramit Ağı'na (Feature Pyramid Network, FPN) [46] gönderilir. FPN piramit şeklindeki yapısıyla, alttan yukarıya olacak şekilde dar ölçekli ve geniş ölçekli öznitelik haritalarını oluşturur. FPN sayesinde farklı ölçeklerde özniteliklere sahip oldukça kapsamlı bir öznitelik haritası oluşturulabilmektedir. Bu da FPN ile kullanılan Mask R-CNN performansına olumlu katkılar sağlamaktadır. Ayrıca, Mask R-CNN mimarisi ROI havuzlaması gerçekleştirilirken yapılan maske hizalama hatalarını azaltmak için bir ROI Hizalama (ROI Align) katmanı kullanır.

Şekil 3.4'te yapısı gösterilen Mask R-CNN mimarisinde sınıf ve sınırlayıcı kutu koordinatlarına ek bir şekilde her bir ROI bölgesi için maske çıktısı üretilir. Bu maske çıktıları ikili maskeler olarak üretilmektedir. Mask R-CNN mimarisinde poligon maskeleri üretilirken, örnekleri oluşturulan ROI bölgeleri için ikili maskeleri kodlayabilen maske dalı kullanılmaktadır. Bu maske çıktılarında piksel düzeyinde bir sigmoid [47] fonksiyonu kullanılmaktadır.



Şekil 3.4. Mask R-CNN mimarisinin akış şeması [21].

Maske dalına paralel olarak çalışan sınıflandırma dalı, hedef sınıf etiketlerinin üretilmesi için kullanılmaktadır. Sınıflandırma dalında ise piksel düzeyinde bir softmax ve tam bağlı katman kullanılmaktadır. Tam bağlı katman içerisinde çok terimli çapraz entropi kayıp fonksiyonu kullanılmaktadır. Bu süreçler ile maske ve sınıf çıktıları birbirinden ayırt edilebilmektedir.

## BÖLÜM 4

### GÖRÜNTÜ SINIFLANDIRMA

Görüntü sınıflandırma, görüntünün bütününe ele alarak işlev gösterebilen görüntü işleme alanında kullanılan önemli bir yaklaşımdır. Bu yaklaşım, görüntüdeki öznitelikler ile görüntüleri inceleyerek görüntüleri belirgin kategorilere ayırmayı hedef almaktadır. Görüntü sınıflandırma tıbbi teşhis, karakter tanıma ve uzaktan algılama problemleri gibi pek çok uygulama alanına sahiptir [48].

#### 4.1. EVRİŞİMLİ SİNİR AĞLARI

Derin öğrenme, insanların düşünme yeteneklerinden ve beyin fonksiyonlarının matematik modellenmesiyle ortaya çıkan yapay sinir ağlarının daha derin katmanlar yaklaşımı ile büyük ölçekli verilerden öğrenebildiği bir makine öğrenmesi dalıdır. Derin öğrenme modelleri, içerisinde pek çok katmanla birlikte derinliği artırarak öğrenme kapasitesini genişletmektedir. Bu modeller, veri içerisindeki karmaşık öznitelikleri diğer yapay zeka yöntemlerine göre daha etkin bir şekilde öğrenebilmektedir. Derin öğrenme modelleri öğrenme türüne, katman yapısına ve karmaşıklığına göre birçok alana bölünmektedir. Literatürde sunulan pek çok derin öğrenme yöntemi olsa da CNN'lerin görüntüler üzerinde etkin bir performans gösterdiği ortaya konulmuştur.

CNN'lerin uygulama alanı genelde görüntü sınıflandırma, görüntü üzerinde tespit ve bölütleme uygulamalarına konu olmaktadır. CNN'ler insan beyninin görüş prensibi modellenerek tasarlanmıştır. Bu ağlar, görüntü üzerindeki öznitelikleri otomatik bir şekilde çıkararak yüksek bir performans gösterebilmektedir. CNN'ler üç ana katman ile görüntü işleme süreçlerine katılmaktadır. Bu katmanlardan ilki evrişim katmanıdır. Bu katman ile görüntü matrisleri filtrelenerek görüntülerden öznitelik haritaları (feature maps) [49] oluşturulur. Evrişim katmanı ile öğrenilebilir parametreler bir

öznitelik haritası üzerinde tutularak diğer katmanlara iletilir. Bir diğer katman ise havuzlama (pooling) [50] katmanıdır. Evrişim katmanında oluşturulan öznitelik haritaları havuzlama katmanında boyut indirgeme işlemlerinden geçer. Hedef öznitelikler havuzlama katmanı ile elenir ve yüksek boyutluluk probleminin önüne geçilir. Havuzlama katmanında görüntüler maksimum ve ortalama havuzlama olmak üzere iki farklı işlemde geçirilebilmektedir. CNN'lerin son katmanı tam bağlı [51] katmanıdır. Tam bağlı katman üzerinde, öznitelik haritaları tek boyutlu bir vektöre dönüştürülmektedir. Bu katman içerisinde geleneksel sinir ağlarının yaklaşımı kullanılarak, her girdi bir sonraki katmanın nöronu ile ilişkilendirilerek tam bağlı yapı oluşturulur ve bu sayede öğrenme gerçekleştirilir.

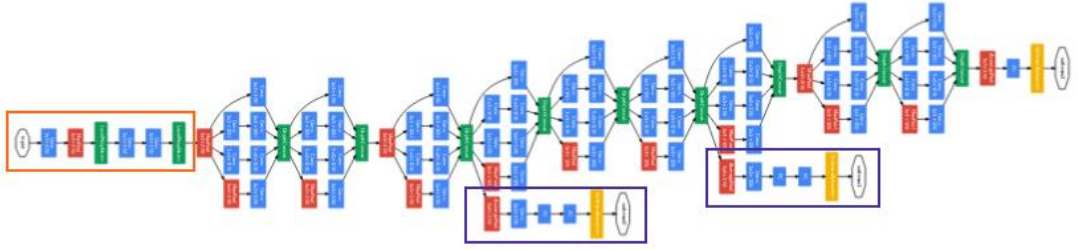
## **4.2. ÖN EĞİTİMLİ EVRİŞİMLİ SİNİR AĞLARI**

Ön eğitilmiş CNN'ler, büyük veri kümeleri üzerinde önceden eğitilerek, farklı öğrenim görevlerinde daha hızlı ve verimli bir eğitim süreci sağlayabilmektedir. Bu özelliği sayesinde birçok derin öğrenme uygulamasında oldukça tercih edilmektedir. Eğitim sürecine kattığı bu avantajlar sayesinde sınıflandırma ve nesne tespiti gibi pek çok görevde omurga ağı ya da temel ağ olarak kullanılmaktadır. Büyük veri kümelerini başlangıç noktası olarak ele alan ön eğitilmiş CNN'ler, sağladığı verimlilik sayesinde de zor özniteliklere sahip sağlık görüntülerinin işlenmesinde de oldukça tercih edilmektedir. Öte yandan, öğrenme aktarımı mantığına dayanan ön eğitilmiş CNN'ler, daha az veri ile de yüksek performans gösterebilmektedir. Tüm bu nedenler ele alındığında öne eğitilmiş CNN'ler, literatürde birçok çalışmada araştırma konusu olarak karşımıza çıkmaktadır.

### **4.2.1. GoogLeNet**

GoogLeNet [52] mimarisi, derin katmanlara dayanan bir ağ olarak literatürde sunulmaktadır. GoogLeNet mimarisi geliştirilirken "Inception V1" olarak adlandırılan özel bir yaklaşım ortaya konulmuştur. Inception yaklaşımında geliştirilen modüller, çeşitli boyutlara sahip filtreler ile çalışabilen evrişim işlevlerine ve katmanların çıkışlarına boyut indirgemesi sağlayabilen havuzlama katmanlarına sahiptir. GoogLeNet mimarisi, sunduğu bu olanaklar ile eşzamanlı olarak çeşitli ölçeklerdeki

öznitelikleri katmanlarında işleyebilmektedir. Bu Inception yaklaşımının sunduğu avantajlar sayesinde GoogLeNet mimarisi, oldukça işlevsel bir öznitelik temsili sağlayabilmektedir. Öte yandan, bu modüller katmanlardan diğer katmanlara gönderilen özellik haritalarını ele alabilmektedir ve özellik haritalarından aktarılan bilgilerin birleşim noktalarını öğrenebilmektedir.



Şekil 4.1. GoogLeNet mimarisinin katmanları [52].

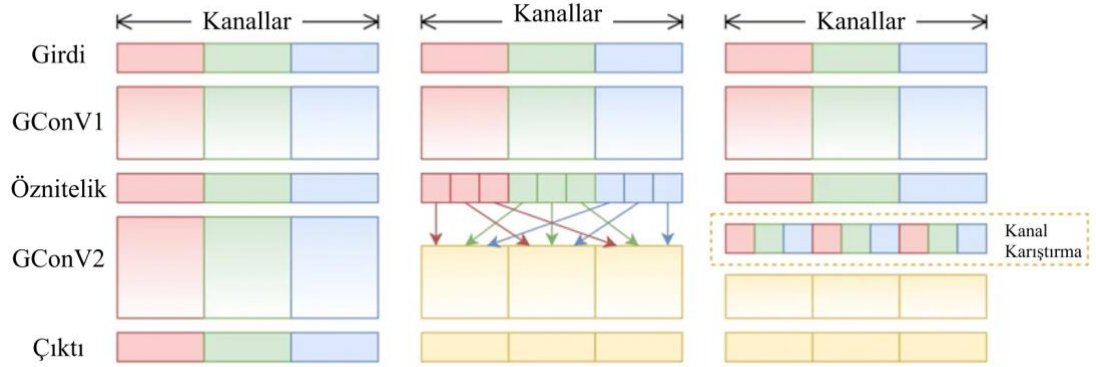
GoogLeNet mimarisi, Şekil 4.1’de gösterildiği gibi boyut indirgeme sağlayan 128 filtreli 1x1 evrişimler kullanmaktadır. Bu mimari, çıkışta 1024 birime sahip Doğrultulmuş Lineer Birim (Rectified Linear Unit, ReLU) [53] ile çalışan tam bağlı bir katmanı kullanmaktadır. Gereksiz özniteliklerin %70’inin çıktıya verilmediği bir bırakma (dropout) [54] katmanını da bünyesinde barındırmaktadır. GoogLeNet mimarisi, sınıflandırıcı olarak ise model çıktılarını olasılık değerleri olarak sunabilen softmax matematik fonksiyonuna dayanan bir katman kullanır. Bu sayede hem iki sınıfta hem de çok sınıfta çıktı üretilebilmektedir.

GoogLeNet mimarisi, çok fazla katman bulundurduğu için oluşan derinliği azaltmak için 1x1 boyutlarında evrişimler kullanmaktadır. Aynı zamanda, bu evrişim işlemleri hesaplama yükünü ve karmaşıklığını da azaltmaktadır. Bu evrişim işlemleri sayesinde, çeşitli öznitelik kanalları da birbirine bağlanabilmektedir. Bağlanma işlemleri ile birlikte de öznitelikler arası bilgi aktarımı ve etkileşim sağlanmaktadır. GoogLeNet mimarisi, paralel ve anlık olarak çalışabilen pek çok Inception modülü ile geliştirilmiştir. Paralel çalışabilen modüller ile ağ oldukça geniş bir ölçekte öznitelik öğrenimi gerçekleştirebilmektedir. Ayrıca, farklı ölçekteki öznitelikler paralel işlevler

sağlayabilen modüller sayesinde daha iyi algılanabilmektedir. GoogLeNet mimarisi, çıkışa yaklaşan katmanlarında küresel ortalama işlevi gösterebilen havuzlama katmanı kullanmaktadır. Bu havuzlama katmanı, öznitelik haritasının boyutunu değiştirerek çıkışı üretir ve tam bağlı katmanlara sabit bir giriş sağlar. Ayrıca GoogLeNet oldukça derin ve geniş bir mimari olduğu için küresel ortalama işlevini sağlayan havuzlama katmanı, parametre açısından verimlilik sağlar.

#### 4.2.2. ShuffleNet

ShuffleNet [55] mimarisinin ana bileşeni, özniteliklerin kanal aralarında aktarımını sağlamak için kullanılan bir kanal karıştırma işlevi olarak karşımıza çıkmaktadır. Bu mimari robot, telefon ve drone gibi cihazların üstündeki donanımlar hedeflenerek tasarlanmıştır. ShuffleNet geliştirilirken derin öğrenme uygulamalarının bu tarz cihazlarda hızlı ve verimli bir şekilde aktivite gösterebilmesi hedeflenmiştir. Donanım odaklı bu mimari ile sınırlı hesaplama maliyeti ile en iyi performansı elde etmek hedeflenmektedir.

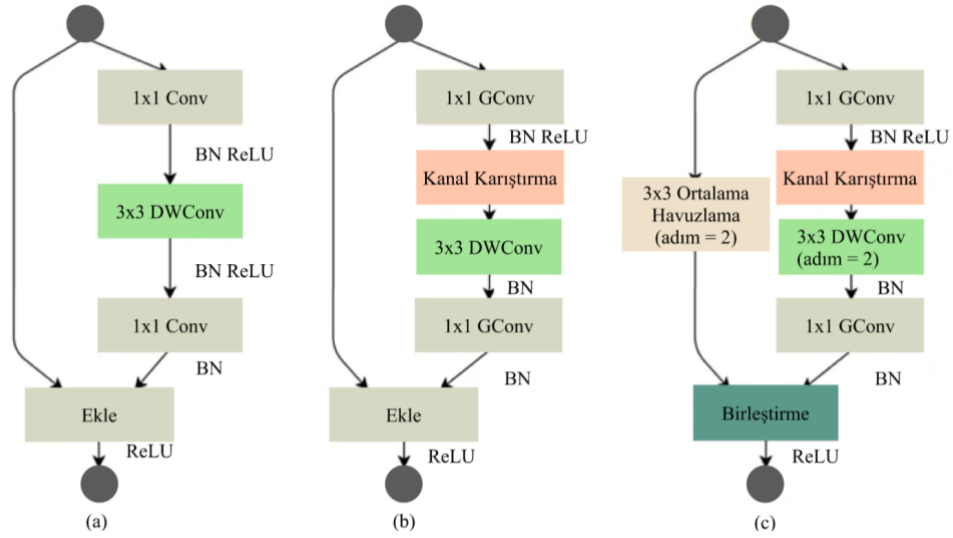


Şekil 4.2. ShuffleNet mimarisinin katmanları ve kapsamlı görünümü [55].

ShuffleNet ile aynı araştırma alanlarına konu olabilen Xception [56] ve ResNeXt [57] mimarileri küçük ağlarda verimlilik açısından yetersiz kaldığı için ShuffleNet mimarisinde noktasal Grup Evrişimi (Grouped Convolution, GConv) önerilmektedir. Bu sayede hesaplama yükü ve karmaşıklığı daha da azaltılmıştır. GConv, öznitelikleri matrislerini işlerken her bir grup içerisinde yer alan kanalları karıştırmak için bir kanal karıştırıcı (channel shuffle) işlemini Şekil 4.2’de gösterildiği gibi sunar. Bu işlem ile, özniteliklerden elde edilen bilgiler kanallar arasındaki geçişlere katılarak daha yüksek



performanslı bir eğitim işlemi sağlanır. Aynı zamanda, modelin hesaplamaya harcadığı güç azaltılmış olur. Kanal karıştırma işleminden önce her grup içinde bulunan kanalların çıktıları birbirinden ayrı olarak elde edilir. Elde edilen çıktılar grup ikili birleştirme (grouped binary concatenation) işlevi ile kanalların sırası bozulmadan grup çıktıları birbirine yan yana olacak şekilde birbirine eklenir. Son olarak ise kanal karıştırma işlevi ile bilgi aktarımlarında rastgelelik ve denge sağlanır.



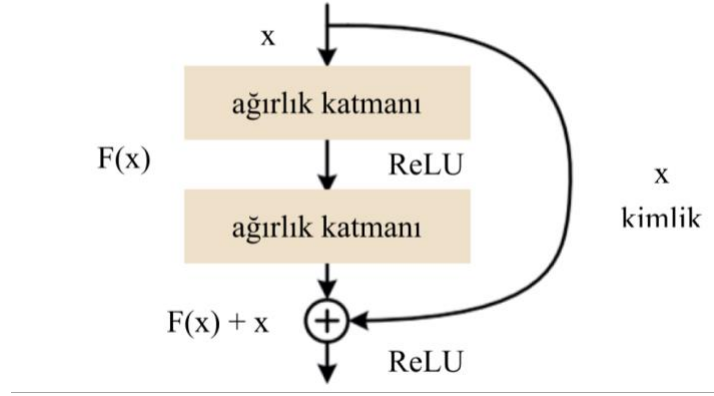
Şekil 4.3. ShuffleNet bileşenleri. a) DWConv [56,58] darboğaz ünitesi [59]; b) GConv ve kanal karıştırmalı bir ShuffleNet birimi; c) iki adımlı ünite [55].

Şekil 4.3'te ShuffleNet birimleri üç aşamalı olarak gösterilmektedir. Şekil 4.3(a)'da 3x3 Derinlemesine Evrişim (Depthwise Convolution, DWConv) kullanan bir darboğaz birimi gösterilmektedir. Şekil 4.3(b)'de ise GConv ve kanal karıştırma işlevine sahip bir ShuffleNet elemanı gösterilmektedir. Burada grup evrişim işlevinin hedefi, kanalın boyutunu kısayol yolu ile eşlemektir. Son olarak Şekil 4.3(c)'de iki adımdan (stride) oluşan bir ShuffleNet birimi ifade edilmektedir. Kanal karıştırma ve GConv kullanan bu ünite ile ShuffleNet ağından geçen bütün öznetelikler verimli bir şekilde öğrenilebilir.

#### 4.2.3. ResNet

ResNet [59] mimarisi, geri yayılım sırasında ağırlık güncellemesi yapılırken ortaya çıkan gradyan patlamalarına (vanishing gradient) çözüm olarak sunulan bir derin sinir

ağdır. ResNet ağı temelinde artık bloklardan (residual blocks) oluşturulmaktadır. Şekil 4.4'te gösterilen artık bloklar giriş verilerini alarak çıktıyı hesaplar ve çıktılar hesaplanırken giriş özneliklerinin üzerine yeni katman bilgileri ekler. Bu bloklar kısayol bağlantılarını (shortcut connection) kullanarak giriş verilerini çıktıya doğrudan gönderebilmektedir. Kısayol bağlantıları öznelikler arasında  $x$  kimlik (identity) parametresini kullanarak bilgi eşleşmesi yaparak yığınlanmış katman çıktılarına oluşturulan haritaları ekler. Bu bağlantılar sayesinde artık bloklarda öğrenme kolaylaştırılır ve öğrenme performansı yükseltilir. ResNet ağları derinliklerine göre farklı versiyonlarıyla literatürde sunulmaktadır. ResNet'in modellerinde daha fazla katman eklenerek karmaşıklık ve performans artırılmaktadır.



Şekil 4.4. Artık blok gösterimi.

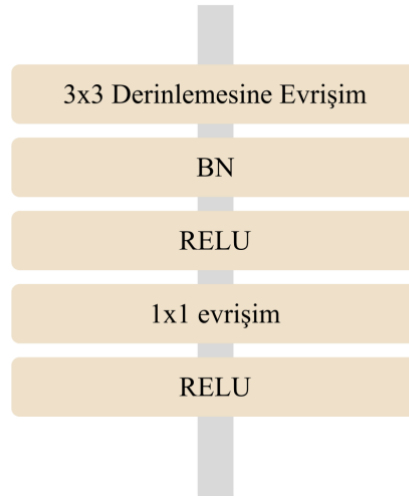
ResNet ağı, pek çok avantaj sağlayabilmektedir. Farklı versiyonları ve derin ağ yapısı ile literatürde sunulan diğer yöntemlere göre daha yüksek bir performans gösterebilmektedir. Gradyan kaybı problemine farklı çözümler sunarak daha kolay eğitim aşaması sağlayabilmektedir. Artık ağlar içerisinde kurulan yapı sayesinde optimizasyon süresine daha kolay adapte olabilme yeteneğine sahiptir.

#### 4.2.4. MobileNet

MobileNet [58] derin öğrenme ağı, kaynak yetersizliğini aşabilmek için inşa edilmiş küçük ve düşük gecikme olanağı sağlayan ağlardır. Bu ağlar tıpkı GoogLeNet gibi, mobil cihazlar başta olmak üzere donanımsal cihazların üzerinde performans sağlama amacıyla sunulmuştur. Küçük bir model olmasına rağmen daha derin ölçekli modellere

benzer olarak görüntü bölütlemesi, sınıflandırması ve görüntü üzerinde nesne tespiti gibi birçok hedefte konfigüre edilebilmektedir. MobileNet mimarileri, donanımlar üzerindeki performansları ele alınarak daha küçük olarak tasarlanmasına rağmen boyut, gecikme ve doğruluk arasında dengeyi kurarak yüksek performans göstermektedir.

MobileNet, geleneksel modellerden farklı olarak iki ana strateji ile ilerlemektedir. Bu stratejiler sayesinde daha hafif ve verimli bir eğitim süreci sunmaktadır. Bu stratejilerden ilki derinlemesine ayrılabilir evrişim (depthwise separable convolution) olarak karşımıza çıkmaktadır. Bu işlev öznelikler üzerinde iki aşamalı bir evrişim işlemi uygulamaktadır. Derinlemesine ayrılabilir evrişimin parametre sayısını ve hesaplama karmaşıklığını azaltması sayesinde MobileNet mimarisinin hafif yapısı sunulmaktadır. Derinlemesine ayrılabilir evrişim her bir girdiyi bireysel olarak işler ve tek kanallı bir çıktı haritası üretir. Şekil 4.5'te gösterildiği gibi bu evrişim işlevlerini barındıran MobileNet katmanı gösterilmektedir. Katmanlarda bu evrişim işlemi bir Toplu Normalleştirme (Batch Normalization) [60], daha sonrasında ReLU ve 1x1 boyutlarında evrişimler takip eder.



Şekil 4.5. MobileNet bloğu.

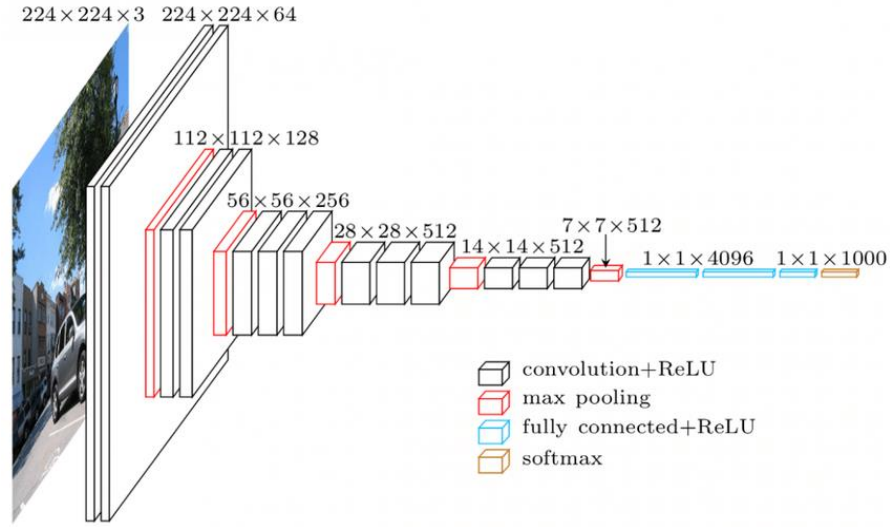
MobileNet modelinin bir diğer ana stratejisi olarak, ağın genişliği etkileyen kanal sayısını kontrol edebilmek için genişlik çarpanı (width multiplier) hiperparametresi sunulmaktadır. Genişlik çarpanı ile kanal sayısının artırılmasının yanı sıra ağ boyutu

da dikkate değer ölçüde küçültülmektedir. Bu iki ana strateji sayesinde MobileNet kaynak bakımından sınırlı olan cihazlar üzerine konfigüre edildiğinde oldukça hafif ve verimli bir eğitim süreci sunabilmektedir.

#### 4.2.5. VGG

VGG [13] mimarisi, VGG-16 ve VGG-19 olmak üzere iki versiyonla literatürde sunulmaktadır. Bu mimariler içerisinde bulundurduğu evrişim katmanları sayısına göre isimlendirilmektedir. Şekil 4.6'da VGG mimarisinin katmanları ve çalışma prensibi gösterilmektedir. VGG, kendinden önce gelen ağlardan farklı olarak evrişim katmanlarında 3x3 boyutlarında filtreler kullanmaktadır. Bu sayede en küçük boyutlu yakalama filtreleri elde edilmiştir ve sonuçlar üzerinde iyileştirmeler sağlanmıştır. Aynı zamanda VGG mimarisinde, ek evrişim filtreleri bulunmaktadır. Bu evrişim filtreleriyle birlikte katman çıktılarına ReLU aktivasyon fonksiyonu uygulanmadan önce girdiler üzerinde bir doğrusal değişim işlevi sağlanır. Bu evrişim işlemleri gerçekleştirilirken uzamsal çözünürlüğün korunabilmesi adına adım sayısı bir piksel olacak şekilde ayarlanmaktadır. VGG ağında bulunan bütün gizli katmanların hepsi ReLU aktivasyon fonksiyonunu kullanmaktadır. ReLU aktivasyon fonksiyonu dikkate alınabilecek bir şekilde çalışma zamanını düşürmektedir. VGG mimarisi, üç adet tam bağlı katman içermektedir ve bu katmanların ilki ve ikincisi toplamda 4096 kanal içerirken, sonuncusu ise 100 kanal içermektedir.

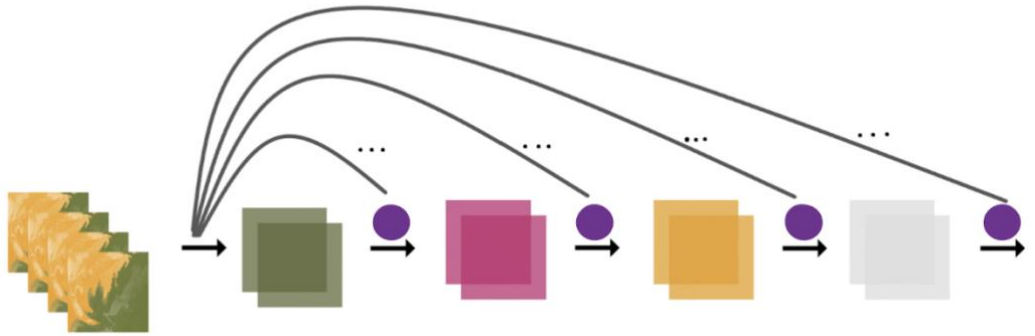
VGG-16 mimarisi 13 evrişim katmanına sahiptir. Bu katmanlar arasındaki öznitelik haritalarına havuzlama işlevleri uygulanmaktadır. Ayrıca VGG-16 mimarisinde üç adet tam bağlı katman bulunmaktadır. Bu yapı ile mimari toplamda 16 katmandan oluşmaktadır. VGG-19 mimarisi ise 16 evrişim katmanı ve VGG ağının ana yaklaşımını olan üç adet tam bağlı katman içermektedir. Daha çok evrişim işlevi sunan katman kullanarak VGG-19 mimarisi ile daha derin bir ağ elde edilmektedir. Her iki mimaride de parametre sayısını azaltmak için 3x3 boyutlu evrişim filtreleri kullanılmaktadır.



Şekil 4.6. VGG ağ mimarisi [13].

#### 4.2.6. DenseNet

ResNet mimarisinde, katmanlardan bir sonrakine aktarım gerçekleştirilirken gradyan yayılımını sağlamak için kimlik işlevleri önerilmektedir. DenseNet [16] mimarisinde ise bu kimlik işlevleri ile önceki katmanlardan sonraki katmanların hepsine ek öznetelik haritaları gönderilmektedir. Bu sayede bütün katmanlar arasında bilgi aktarımı sağlanır. Önceki katmanlardan alınan bilgiler güncel katmanda öznetelik haritalarında toplanarak katman işlevlerinden geçer. Bu aktarım işlevleri yoğun (dense) bağlantılar olarak literatürde sunulmaktadır ve bu bağlantılar sayesinde daha başarılı sonuçlar elde edilebilmektedir.

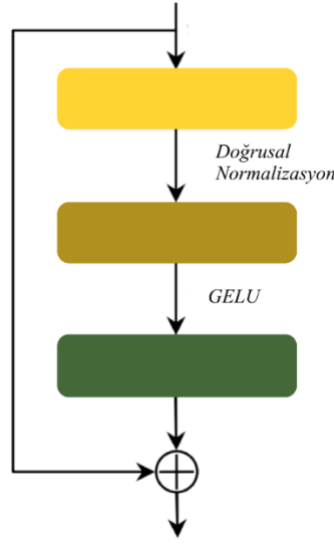


Şekil 4.7. DenseNet katmanlar arası aktarım.

Şekil 4.7’de DenseNet mimarisindeki katmanlar arası bilgi aktarımı gösterilmektedir. DenseNet mimarisinde yoğun bağlantılar ile yapılan aktarımlardan kaynaklı olarak kanal sayısı diğer mimarilere kıyasla daha azdır. Bu mimari içerisinde ağıın büyüme oranını kontrol eden bir  $k$  hiperparametresi kullanılmaktadır. Bu hiperparamete ile birlikte katmanlar arasında ne kadar bilgi aktarılacağı belirlenmektedir. Kontrol edilen bilgi miktarı küresel olarak ağıın bütün katmanlarına gönderilebilmektedir.

### 4.3. ÖZ-DİKKAT TABANLI DERİN ÖĞRENME YÖNTEMLERİ

Öz-dikkat tabanlı derin öğrenme yöntemleri, görüntü ve dil işleme, zamansal tahminleme modelleri gibi pek çok uygulama alanına sahip yaklaşımlardır. Bu yaklaşımlarda girdiler birbirinden farklı bölgeler arasındaki ilişkileri hesaplamaya odaklanmaktadır. Öz-dikkat tabanlı yöntemler, hedef çıktılar için belirgin olan öznitelikleri vurgulamaya amaçlamaktadır. Bu yöntemler ana bileşen olarak, öz-dikkat matrislerini ele almaktadır. Bu matrisler ile ilişki hesaplamaları gerçekleştirilir ve ilişki hesaplamaları öznitelik haritalarındaki ağırlıklarla ilişkilendirilir. Bu sayede her bir özniteliğin diğer özniteliklerle etkileşimi ele alınarak bilgilerin daha iyi öğrenilmesi sağlanır. Literatürde pek çok öz-dikkat tabanlı yöntem bulunmaktadır [61–65]. Dönüştürücü Modeli [66], Dönüştürücülerden Çift Yönlü Enkoder Gösterimleri (Bidirectional Encoder Representations from Transformers, BERT) [67] ve Jeneratif Önceden Eğitilmiş Dönüştürücü (Generative Pre-trained Transformer) [68] gibi modeller genelde çeviri, modelleme, duygu analizi gibi dil işleme gibi uygulama alanlarına sahiptir. Öz-dikkat tabanlı modellerin görüntü işleme alanında da birçok uygulama alanı bulunmaktadır. Görüntü alanında uyarlanmış ilk versiyon olarak ViT [29] modeli karşımıza çıkmaktadır. Bu modelde her bir girdi birimi diğer girdi birimleri ile etkileşim halindedir. Bu model, girdi görüntülerini öz-dikkat mekanizmaları ile piksel düzeyinde işleyebilmektedir. Bir diğer görüntüleri işleyebilen öz-dikkat tabanlı yöntem ise ConVNeXt [69] modelidir. Bu model, ViT ve literatürde sunulan Swin Dönüştürücüsü [27] modelini temel alarak geliştirilmiştir.



Şekil 4.8. ConvNeXt katmanları.

ConvNeXt modeli, referans alınan yöntemlerle yaklaşık olarak aynı parametre sayısına ve bellek kullanım oranına sahiptir fakat, bu model ile birlikte belirli modüller basitleştirilmiştir. Şekil 4.8’de katman prensibi gösterilen ConvNeXt modeli öncelikle bir evrişim katmanı ile girdiyi işler. Daha sonrasında bu evrişim işlemleriyle oluşturduğu özniteliklere bir doğrusal normalizasyon işlemi uygular. Doğrusal normalizasyon ile evrişim katmanı çıktıları belli bir aralığa getirilir. ConvNeXt bloğunun son işlevi olarak ise Gauss Hatası Doğrusal Birimi (Gaussian Error Linear Unit, GeLU) [70] aktivasyonudur. Bu aktivasyon işlemi çıktıları sıkıştırır ve dönüşüm işlemini tamamlar.

#### 4.4. ÖNERİLEN YÖNTEM

DenseNet mimarisinin çıkışını sağlayan sınıflandırıcıya konfigürasyon işlemi uygulanarak bir derin öğrenme yöntemi önerilmiştir. DenseNet modeli içerisinde küresel ortalama havuzlama işlevi kullanarak alt örnekleme uygulanan öznitelikler tam bağlı bir katmana gönderilmiştir. Çıktılar için ideal ağırlık ve bias değerleri öğrenildikten sonra ReLU aktivasyon ve bırakma işlevi ile rastgelelik sağlanarak aşırı öğrenme probleminin önüne geçilmiştir. Aynı zamanda çıktılar sıkıştırılarak geleneksel yöntemlere kıyasla iyileşme sağlanmıştır. Model çıktıları, yeni girdiler tam bağlı bir katmanda yeniden işlenerek elde edilmektedir. Bu uygulamalar sonucunda klasik

DenseNet mimarisinin sınıflandırıcısına göre daha derin bir sınıflandırıcı elde edilmiştir.

$$y_{jk}(x_{in}) = f\left(\sum_{c=1}^n w_{jk}x_{inc} + w_{j0}\right) \quad (4.1)$$

Eşitlik 4.1'deki gibi gösterilen matematiksel ifade'de  $j$  ve  $k$  matris indislerini temsil etmektedir. Burada  $w$  ağırlık matrisi ile  $x_{in}$  olarak gösterilen girişler arasında iç çarpım işlemi yapılır ve  $f$  ile gösterilen aktivasyon fonksiyonuna gönderilir.  $w$  ağırlık matrisini ve  $w_0$  ise sapma (bias) değerini temsil etmektedir. Bu sayede  $y_{jk}$  ile gösterilen tam bağlı katmanın çıktısı oluşturulur.

$$\tilde{y}^{(l)} = r^{(l)} * y_{jk}^{(l)} \quad (4.2)$$

Tam bağlı katmanın çıktısı iki nokta dağılımına ilişkin bağımsız Bernoulli rastgele değişkenlerini gösteren  $r^{(l)}$  vektörüyle eleman bazında çarpıma katılır. Çarpım sonuçları Eşitlik 4.2'deki gibi  $\tilde{y}$  ile belirtilmektedir. Burada  $l$  matematiksel ifadesi ilgili katmanı belirtmektedir.

$$z^{(l+1)} = w^{(l+1)}\tilde{y}^{(l)} + w_0^{(l+1)} \quad (4.3)$$

Eşitlik 4.3'te bir sonraki katman girdilerinin vektörünü temsil eden değer  $z$  ile gösterilmektedir.  $l$  değeri bir sonraki katmanın ağırlıklarıyla çarpılarak sapma değerleriyle toplanır:

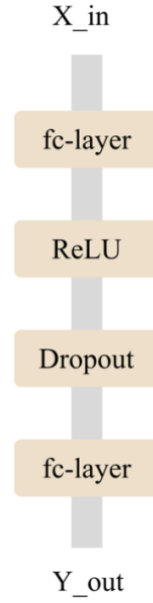
$$y^{(l+1)} = ReLU(z^{(l+1)}) \quad (4.4)$$

Eşitlik 4.4'te katman çıktıları vektörü ReLU aktivasyon fonksiyonuna gönderilerek katman çıktıları üretilmektedir.



$$y_{out}(y^{(l+1)}) = f\left(\sum_{c=1}^n w_{jk}y^{(l+1)}_c + w_{j0}\right) \quad (4.5)$$

Dağıtım vektörü kullanılarak inceltmiş çıktılarla birlikte aşağı örnekleme elde edilmiş olur. Son olarak aşağı örneklenmiş çıktılar tekrar bir tam bağlı katman kullanılarak Eşitlik 4.5'te gösterildiği üzere  $y_{out}$  gerçek çıktıları üretilir.



Şekil 4.9. Önerilen yöntemin sınıflandırıcısı.

Şekil 4.9'da üzerinde yapılandırma uygulanan sınıflandırıcı gösterilmektedir. Şekildeki yapıya göre girişler bir tam bağlı katmandan geçirilerek bir aktivasyon işlevine gönderilir. Aktivasyonu gerçekleştirilmiş girdiler daha sonrasında bir tam bağlı katmandan geçerek çıktı üretilir.

## BÖLÜM 5

### DENEYSEL ÇALIŞMALAR

Bu bölümde, derin öğrenme yöntemlerinin uygulanması gerçekleştirilirken Python programlama dili kullanılmıştır. Python programlama dili pek çok modül ve kütüphane dağıtımıyla birlikte derin öğrenme yöntemlerinin uygulanmasında esneklik sağlamaktadır. Pytorch [71], TensorFlow [72] ve Keras [73] gibi dağıtımlar bu derin öğrenme yöntemleri uygulanırken başlıca kullanılan kütüphanelerdir. Derin öğrenme modellerinin eğitim ve oluşturulmasında çeşitli avantajlar sağlayabilen Keras kütüphanesi TensorFlow kütüphanesi ile entegre bir şekilde çalışabilmektedir. Aynı zamanda keras çekirdeği ile birlikte Pytorch kütüphanesine çeşitli derin öğrenme katmanları entegre edilebilmektedir. Tez çalışması gerçekleştirilirken bu kütüphanelerin sürümleriyle uyumlu bir bilgisayar kullanılmıştır. Bilgisayar, NVIDIA Quadro RTX 5000 grafik kartına ve i9 10980XE işlemciye sahiptir.

#### 5.1. PERFORMANS METRİKLERİ

Çürük nesnelere örnek bölütleme görevinin performansı ölçülürken Birleşim Üzerinden Kesişim (Intersection Over Union, IOU) metriği kullanılmaktadır. IOU değeri uzman etiketlerinden oluşan dış çürüğü poligonları ile Mask R-CNN mimarisinin ürettiği poligonlar arasındaki değerler ile hesaplanmaktadır. Jaccard katsayısı olarak da bilinen IOU metriği, gerçek etiketin sınırlayıcı kutusu ile tahmin edilen kutunun kesişim değerini bulur ve bu değeri kutuların birleşim değerine böler [74]. Şekil 5.1’de bölme işlemi gösterilmektedir. Bu sayede Doğru Pozitif (True Positive, TP), Doğru Negatif (True Negative, TN), Yanlış Pozitif (False Positive, FP) ve Yanlış Negatif (False Negative, FN) değerleri üretilmektedir. TP değeri, gerçek etiketlemedeki sınırlayıcı kutu ile tahmin edilen sınırlayıcı kutunun aynı olduğu durumu göstermektedir. TN değeri, arka plan olan bir bölgenin arka plan olarak tahmin edilmesi durumudur. FN değeri, gerçek etiketlemede var olan nesnenin

tahminde bulunamadığını gösterme durumudur. FP değeri ise, etiketlenmemiş bir bölgenin nesne olarak tahmin edildiği durumları saymaktadır.



Şekil 5.1. IOU gösterimi.

Doğruluk metriği, sınıflandırma modelinin ne kadar doğru tahmin gerçekleştirdiğini ölçmek için hesaplanmaktadır. Doğruluk metriği ile modelin doğru tahminlerinin, toplam tahminlere oranı ölçülmektedir. Eşitlik 5.1’de bağıntı ile doğruluk metriği hesaplanmaktadır.

$$\text{Doğruluk} = \frac{(TP + TN)}{(TP + FP + TN + FN)} \quad (5.1)$$

Kesinlik metriği, modelin ne kadar doğru bir tahmin yaptığını ölçmek için hesaplanmaktadır. Doğru tahminlerin tüm pozitif durumlara olan oranı Eşitlik 5.2’deki bağıntı ile hesaplanmaktadır.

$$\text{Kesinlik} = \frac{TP}{TP + FP} \quad (5.2)$$

Duyarlılık metriği, modelin pozitif sınıfının ne kadar iyi şekilde tanımlandığını yansıtmaktadır. Doğru tahmin edilen nesne sayısının tüm tahminlere oranı, hesaplamadaki gerçek nesnelere için elde edilmektedir. Duyarlılık metriğinin hesaplama ifadesi Eşitlik 5.3’te gösterilmektedir.

$$Duyarluluk = \frac{TP}{TP + FN} \quad (5.3)$$

F1-skoru metriği, kesinlik ve duyarlılık değerlerinin ne kadar dengeli olduğunu belirlemektedir. Bu iki metrik ile hesaplanan f1-skoru, modelin TP değerinin sayabileceği özellikleri yakalama yeteneğini ve FP değerlerini kontrol etme yeteneğini değerlendirmektedir. F1-skoru, Eşitlik 5.4'teki bağıntı ile hesaplanmaktadır.

$$F1 - skoru = 2 * \frac{Duyarluluk * Kesinlik}{Duyarluluk + Kesinlik} \quad (5.4)$$

## 5.2. NESNE TESPİT MODELLERİNİN SONUÇLARI

Mask R-CNN nesne tespit modelinin eğitimi için ortam kurulumu Anaconda dağıtımının [75] conda paketleri ile sağlanmıştır. Eğitim arayüzleri ve model yapılandırılması için kodlar jupyter notebook üzerinde çalıştırılmıştır. Jupyter notebook etkileşimli bir ortam sağlayan ve Python kütüphanelerinin çalışmasına olanak sağlayan bir programdır. Daha sonrasında TensorFlow kütüphanesi tanımlanarak derin öğrenme modellerinin çalışması ve eğitilmesi için ortam hazırlıkları tamamlanmıştır. TensorFlow, açık kaynak koduna sahip bir kütüphanedir. Derin öğrenme modellerini yapılandırmak, eğitmek ve uygulayabilmek için kullanılmaktadır. Bu kütüphane içinde bulundurduğu çoğu modül ile birlikte karmaşık matematiksel işlemleri gerçekleştirerek hesaplama açısından verimlilik sağlamaktadır. TensorFlow kütüphanesi, matematiksel ifade ve işlemleri tensör olarak adlandırılan çok boyutlu dizilerde tutmaktadır. Bu sayede girdi matrisleri modellere kolay bir şekilde iletilerek esneklik sağlanmaktadır. Aynı zamanda tensörler, yüksek güç gerektiren işlemleri ve karmaşık yapıları daha verimli şekilde inşa etmeyi mümkün kılmaktadır.

Mask R-CNN mimarisi oldukça büyük bir veri kümesi olan Microsoft COCO [76] verisi üzerinde eğitilmiş ağırlıkları kendine transfer edebilir ve bu sayede farklı nesne tanıma hedeflerinde verimli bir eğitim gerçekleştirebilir. Eğitim süreci hızlandırılmak için bir grafik karta sahip olan bir bilgisayar üzerinde gerçekleştirilmiştir. Mask R-CNN mimarisi omurga ağı olarak ResNet-101 ağını kullanmaktadır. ResNet-101 ağı

ile birlikte girdi tensörlerinden öznitelik haritaları oluşturulmaktadır. Bu öznitelikler, diş görüntülerinin farklı özelliklerini saklayabilmektedir. Girdiler Mask R-CNN mimarisinin bütün katmanlarından geçtikten sonra diş görüntülerini tanıyan model üretilmiş olmaktadır. Bu model, bütün diş görüntülerindeki çürük nesnelere sınırlayıcı kutularını, maske kutularını, güven puanlarını ve sınıf ismini üretebilmektedir.

Nesne tespit modelinin eğitildiği veri seti %80 oranında eğitim ve %20 oranında test olmak üzere ayrılmıştır. Uzman etiketli diş çürüğü lezyonlarının açıklamalarını içeren veri seti Mask R-CNN mimarisine girdi olarak verilmiştir. Bu veri seti içerisinde toplamda 2152 panoramik görüntü bulunmaktadır. 32 dişin görünümünü içeren panoramik görüntülerde ise toplamda 2303 adet diş çürüğü nesnesi bulunmaktadır. Her dişte bir çürük olabileceği gibi, bir dişte birden fazla çürük olabildiği durumlar da bulunmaktadır. Ağzın geniş bir görünümü sunan panoramik görüntülerdeki diş çürüğü lezyonunun örnek bölütlemesi için Mask R-CNN mimarisinin performans ölçümü kesinlik, duyarlılık ve f1-skoru metrikleri ile gerçekleştirilmektedir. Bu metrik sonuçları için IOU değerleri hesaplanmaktadır. Bu sayede uzman etiketleri ile modelin ürettiği etiketlerin kesişimlerinin skorları üretilerek Mask R-CNN mimarisinin sonuçları Çizelge 5.1’de gösterildiği gibi elde edilmiştir. Sonuçlara bakıldığında Mask R-CNN mimarisi diş çürüklerinin örnek bölütleme görevinde yetersiz kalmıştır. Çürük lezyonlarını kapsayan etiketler oldukça küçüktür. Ağzda bulunan her dişin yapısı ve kapladığı oran farklıdır. Bu nedenle diş çürüğü lezyonlarını diş yapılarının üzerinde olduğundan yapıları ve boyutları farklılık göstermektedir. Diş çürüğü lezyonları geniş panoramik görüntülerde çok küçük alanları kaplamaktadır. Bu yüzden Mask R-CNN mimarisinin geniş panoramik görüntülerde diş çürüğü algılaması oldukça zorlu bir görevdir.

Çizelge 5.1. Panoramik görüntülerdeki diş çürüklerinin örnek bölütleme sonuçları.

Model	Kesinlik (%)	Duyarlılık (%)	F1-skoru (%)
Mask R-CNN	45,44	49,27	47,28

Diş çürüğü lezyonlarına daha iyi odaklanabilmek için her bir panoramik görüntü diş süper kategorileri ele alınarak kırılmıştır. Kırpma işlemi, ağzın geniş görünümünü sunan panoramik görüntünün her pikselinde hesaplama yapmak yerine, her bir dişin üzerindeki diş çürüğü özelliklerini algılama imkanı sunmuştur. Görüntülerin kırpma işlemi dişin içerisinde çürük etiketi olmasına göre gerçekleştirilmiştir. Bu sayede toplamda 2303 adet diş görüntüsü çürük etiketleriyle birlikte Mask R-CNN ağına girdi olarak gönderilmiştir. Çizelge 5.2’de kırılmış görüntülerdeki çürük nesnelere örnek bölütlemesini gerçekleştirebilen Mask R-CNN mimarisinin performans sonuçları verilmektedir. Performans sonuçları için gerçekleştirilen kırpma süreci için önceden eğitilmiş diş bölgelerini tanıyabilen bir model [77] kullanılmıştır.

Çizelge 5.2. Kırılmış ve uzman etiketli diş görüntülerindeki çürük lezyonlarının örnek bölütleme sonuçları.

Model	Kesinlik (%)	Duyarlılık (%)	F1-skoru (%)
Mask R-CNN	57,20	69,15	62,61

Mask R-CNN mimarisi, geniş görüntülerdeki küçük alanları kapsayan etiketleri tanıma görevinde daha büyük alanlar içeren nesnelere tanıma görevine kıyasla yüksek bir performans gösteremez [20]. Bu sonuçlar ışığında Mask R-CNN mimarisinin aynı etiket boyutlarını kapsayan küçük görüntülerde, büyük görüntülere kıyasla daha iyi bir performans gösterdiği sonucuna varılmaktadır.

### 5.3. ÖZ-DİKKAT TABANLI DERİN ÖĞRENME YÖNTEMLERİNİN SONUÇLARI

Öz-dikkat tabanlı yöntemlerden ConvNeXt ağına toplamda 2709 görüntüden oluşan diş çürüğü veri seti girdi olarak verilmiştir. Bu veri seti içerisinde 2303 çürük ve 406 çürük olmayan diş görüntüsü bulunmaktadır. Eğitim ve test seti %80’e %20 oranında ayrılmıştır. Bu sayede 2168 görüntüden oluşan eğitim seti ve 541 görüntüden oluşan test seti elde edilmiştir.

Çizelge 5.3. Öz-dikkat tabanlı yöntem sonuçları.

Model	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1-skoru (%)
ConvNeXt	80,34	86,54	91,47	88,94

Çizelge 5.3'te ConvNeXt ağının çürük lezyonlarını ne kadar doğru sınıflandırıldığına dair sonuçlar verilmektedir. ConvNeXt ağının performansı doğruluk kesinlik ve duyarlılık metrikleri ile ölçülmektedir. Sonuçlar, kırılmış panoramik radyografilerde küçük piksel oranlarına sahip çürük bölgelerinin öz-dikkat tabanlı yöntemlerle iyi bir şekilde sınıflandırılabilceğini göstermektedir.

Çizelge 5.4. K-katman çapraz doğrulama sonuçları.

Katman	Eğitim Seti Doğruluğu (%)	Test Seti Doğruluğu (%)
1.	84,91	85,42
2.	85,28	83,95
3.	84,91	85,42
4.	84,82	85,79
5.	84,64	86,51
<b>Ortalama</b>	<b>84,91</b>	<b>85,42</b>

Her bir katmanda elde edilen sonuçların ortalaması alınarak model performansı genel olarak değerlendirilmiştir. K-katman çapraz doğrulama modelin güvenilirliğinin ölçülmesi için kullanılan bir yöntemdir. Çizelge 5.9'da sonuçlara bakılacak olursa farklı katmanlarda alınan sonuçların birbirine yakın değerlere sahip olması ConvNeXt modelinin çürük görüntülerini sınıflandırma işleminde güvenilirliğini göstermektedir. Aynı zamanda çapraz doğrulama yöntemi kullanılarak elde edilen sonuçların normal eğitim sürecinden daha yüksek olması, ConvNeXt modelinin daha dengeli bir

performans gösterdiğini kanıtlamaktadır. Aynı zamanda model, veri setinin farklı katmanlarında yakın sonuçlar göstererek hataları daha iyi tespit edebildiğini ve genelleme yapabildiğini göstermektedir.

ConvNeXt modelinin performansının ve modelin genelleme yeteneğinin ölçülebilmesi için k-katman çapraz doğrulama (k-fold cross validation) çalışması yapılmıştır. K-katman çapraz doğrulama ile modelin veri setinin farklı katmanlarında nasıl çalıştığı araştırılmıştır. Çapraz doğrulama yapılırken veri seti rastgele alt kümelere bölünmüştür. Bölünme işlemi gerçekleştirilirken katman değeri 5 olarak ele alınmıştır. Bu sayede oluşturulan alt veri kümeleri test olarak ele alınır ve 5 katman değeri sayesinde veri setinin %20'lik kısmını oluşturmaktadır. Eğitim verisi ise %80 oranında ayrılmıştır. Her bir katman üzerinde iterasyon gerçekleştirilirken eğitim ve test oranı korunmaktadır. Çizelge 5.4'te ConvNeXt modelinin çapraz doğrulama ile her bir katman üzerindeki eğitim ve test performansı verilmektedir. Aynı zamanda ortalama doğruluk değerleri de gösterilmektedir.

#### **5.4. ÖN EĞİTİMLİ EVRİŞİMLİ SİNİR AĞLARI VE ÖNERİLEN YÖNTEM SONUÇLARI**

Panoramik görüntülerde her bir diş üzerinde çürük lezyonu bulunmamaktadır. Diğer bir ifadeyle, bir hastanın her bir dişinde çürük bulunmayabilir. Bu yüzden panoramik görüntüler kırılarak elde edilen veri setinde, çürük içeren görüntülerin oranı çürük içermeyen görüntülere göre oldukça düşüktür. Bu oran klinik ortamda karşılaşılabilecek çürük oranı ile uyum içerisindedir. Fakat diş çürüğü sınıfına ait görüntülerin azınlık olmasından kaynaklı oluşan sınıf dengesizliği bulunmaktadır. Bu sınıf dengesizliğinin derin öğrenme sonuçları üzerinde nasıl bir performans gösterdiğini algılayabilmek için ana veri seti farklı sınıf oranlarında bölünmüştür. Farklı oranlarda bölme yaparak, azınlık olan diş çürüğü sınıfının yüksek oranlara sahip olduğunda derin öğrenme modellerinin nasıl bir performans gösterdiği araştırılmıştır. Sınıf dengesizliğini engellemek için sınıf oranları eşitlenmiştir ve derin öğrenme modellerinin performansı ölçülmüştür. Aynı zamanda, diş çürüğü sınıfı azınlık olduğu için farklı oranlarda veri setleri oluşturularak modelin genelleme performansı da araştırılmıştır. Farklı sınıf oranları ile yapılan veri bölme işlemi, dengeli bir model



sağlamak, çürük sınıflandırma görevinin klinik ortamda kullanılabilirliğinin incelenmesi ve modelin çürük sınıfındaki performansını algılamak ve artırmak için önemli bir çalışmadır.

Sınıflandırma ağlarına girdi olarak gönderilmek için panoramik görüntülerdeki her bir diş görüntüsü kırıldı. Kırılan diş görüntüleri çürük açıklamalarına göre çürük içerenler çürük sınıfına dahil edildi. Çürük içermeyen diş görüntüleri ise çürük olmayan kategorisine dahil edildi. Toplamda 2303 çürük içeren diş görüntüsü, 12195 çürük içermeyen diş görüntüsü ile bir veri seti oluşturuldu. Oluşturulan veri seti %80 eğitim ve %20 test veri seti olarak ayrıldı.

Ön eğitilmiş CNN'lerin ve önerilen yöntemin eğitim aşaması gerçekleştirilirken tek bir grafik işlemci kullanılmıştır. Test sonuçlarının alınmasında gerçekleştirilen kırma süreci için FDI diş numaralarına göre dişleri algılayabilen önceden eğitilmiş bir derin öğrenme modeli [77] kullanılmıştır.

Çizelge 5.5. %30 çürük, %70 çürük olmayan sınıf oranlarından oluşan veri seti üzerinde sınıflandırma ağlarının performansı.

Model	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1-skoru (%)
GoogLeNet	81,80	78,33	80,24	79,11
ShuffleNet	78,19	74,10	74,43	74,26
ResNet-50	85,13	82,19	82,88	82,52
MobileNet	80,72	77,23	79,64	78,13
DenseNet-169	84,14	81,13	81,24	81,18
DenseNet-121	85,22	84,94	81,64	82,89
VGG-19	85,80	86,47	79,20	81,63
<b>Önerilen Yöntem</b>	<b>85,94</b>	<b>84,70</b>	<b>80,82</b>	<b>82,36</b>

Azınlık olan dış çürüğü sınıfının oranı artırılarak modeller üzerinde nasıl bir performans değişikliği olduğunun araştırılması için ana veri seti bölünmüştür. Kırılmış görüntülerden oluşan ana veri seti çürük oranı %30 olarak ele alınarak, %70'i ise çürük olmayan sınıfı olacak şekilde ayrılmıştır ve alt bir veri seti oluşturulmuştur. Bu alt veri seti %80 eğitim ve %20 test olarak ayrılarak sınıflandırma ağlarına girdi olarak verilmiştir. Sınıflandırma ağlarının sonuçları Çizelge 5.5'te gösterilmektedir.

Çürük ve çürük olmayan görüntü sayısı eşitlenerek sınıflandırıcılar ile performans ölçümü gerçekleştirilmiştir. Burada her iki sınıfın oranı %50 olarak ele alınmaktadır. Çizelge 5.6'da önerilen yöntem ve geleneksel sınıflandırma yöntemlerinin sınıfların dengeli olduğu veri seti üzerindeki performans sonuçları verilmektedir. Dengeli veri setindeki veriler %80 eğitim ve %20 test seti olarak ayrılmıştır.

Çizelge 5.6. Sınıflandırma ağlarının sınıf oranları dengeli veri seti üzerindeki performansı.

Model	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1-skoru (%)
GoogLeNet	84,05	84,08	84,10	84,05
ShuffleNet	79,05	79,42	79,31	79,04
ResNet-50	83,91	84,04	83,82	83,86
MobileNet	83,24	83,35	83,40	83,2
DenseNet-169	84,86	84,85	84,88	84,85
DenseNet-121	85,54	85,63	85,45	85,50
VGG-19	79,18	80,38	79,65	79,12
<b>Önerilen Yöntem</b>	<b>87,29</b>	<b>87,33</b>	<b>87,35</b>	<b>87,29</b>

Çürük sınıfı toplamın %70'ini oluşturacak ve çürük olmayan sınıf toplam veri setinin %30'unu oluşturacak şekilde yeni bir alt veri seti oluşturulmuştur. Daha az olan dış çürüğü sınıfına daha fazla oran vererek oluşturulan bu alt veri seti ile modellerin bu sınıfı ne kadar iyi öğrendiği araştırılabilmektedir. Bu alt veri seti üzerinde VGG-19 mimarisi diğer yöntemleri geride bırakarak daha yüksek bir performans göstermektedir. Çizelge 5.7'de bu veri seti üzerinde sınıflandırma ağlarının performans sonuçları gösterilmektedir. Veri seti, eğitim %80 ve test %20 oranında olacak şekilde ayrılmıştır.

Çizelge 5.7. %70 çürük, %30 çürük olmayan sınıf oranlarından oluşan veri seti üzerinde sınıflandırma ağlarının performansı.

Model	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1-skoru (%)
GoogLeNet	80,72	76,91	77,19	77,04
ShuffleNet	77,83	73,73	70,59	71,72
ResNet-50	80,54	76,96	79,68	77,92
MobileNet	80,18	77,84	76,11	76,83
DenseNet-169	81,44	79,27	77,75	78,40
DenseNet-121	80,72	77,90	79,03	78,39
<b>VGG-19</b>	<b>86,84</b>	<b>84,86</b>	<b>85,17</b>	<b>85,01</b>
Önerilen Yöntem	82,34	78,89	81,66	79,91

Son olarak veri seti %85'i çürük ve %15'i çürük olmayan sınıfını oluşturacak şekilde bölünmüştür. Çizelge 5.8'de sınıflandırma ağlarının ve önerilen yöntemin bu veri seti üzerindeki performans sonuçları gösterilmektedir.

Çizelge 5.8. %85 çürük, %15 çürük olmayan sınıf oranlarından oluşan veri seti üzerinde sınıflandırma ağlarının performansı.

Model	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1-skoru (%)
GoogLeNet	90,22	80,86	74,73	77,28
ShuffleNet	89,85	82,50	69,48	73,65
ResNet-50	90,41	84,88	75,88	79,34
MobileNet	90,95	81,11	80,76	80,93
DenseNet-169	90,77	81,49	83,38	82,38
DenseNet-121	90,66	85,25	76,66	80,02
VGG-19	89,68	85,56	71,63	76,05
<b>Önerilen Yöntem</b>	<b>91,64</b>	<b>82,63</b>	<b>79,49</b>	<b>80,94</b>

Çizelge 5.9’da sınıflandırma ağlarının 14290 görüntüden oluşan veri seti üzerinde doğruluk, kesinlik, duyarlılık ve f1-skoru performans ölçütleri gösterilmektedir. Derin öğrenme modelleri kırılmış görüntülerden oluşan ana veri setinde, belli sınıf oranlarına sahip alt veri setlerine kıyasla daha yüksek bir performans göstermiştir. Bu durum, dış çürüğü sınıfının azınlık olduğu durumlarda modelin bu sınıfı öğrenmek için daha dikkatli olduğunu göstermektedir. Ayrıca en yüksek performansı gösteren model, dış çürüklerine özgü yapılandırılmış sınıflandırıcısıyla, bu sınıfa daha fazla odaklanabilmektedir. Çürük sınıfına özel yapılandırılmış sınıflandırıcısı ile önerilen yöntem, gerçek hayata uygun sınıf dağılımına sahip veri setinde daha yüksek performans göstererek klinik ortamda uygulanabilirliğini kanıtlamaktadır.

Çizelge 5.9. Sınıflandırma ağlarının kırpılmış dış görüntülerinden oluşan ana veri seti üzerinde performansı.

Model	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1-Skoru (%)
GoogLeNet	91,48	84,04	83,73	83,88
ShuffleNet	89,24	79,67	80,09	79,88
ResNet-50	91,17	85,83	78,57	81,55
MobileNet	91,44	84,38	82,65	83,48
DenseNet-169	90,93	83,93	83,32	83,62
DenseNet-121	91,72	84,44	84,32	84,08
VGG-19	91,27	86,66	78,40	81,71
<b>Önerilen Yöntem</b>	<b>93,17</b>	<b>89,43</b>	<b>85,84</b>	<b>87,49</b>

## BÖLÜM 6

### SONUÇLAR VE ÖNERİLER

Bu çalışmada, panoramik görüntülerdeki diş çürüğü verileri üzerinde sınıflandırma, tespit ve örnek bölütleme görevleri literatürde sunulan güncel yöntemler ile gerçekleştirilmiştir. Ayrıca diş çürüğü sınıflandırma görevi için sınıflandırıcısı yapılandırılmış bir yöntem sunulmuştur. Örnek bölütleme ve tespit işlemleri için uzman etiketli görüntüler kullanılmaktadır. Örnek bölütleme görevi Mask R-CNN mimarisi ile gerçekleştirilmiştir. Mask R-CNN mimarisi, kırılmış görüntülerdeki diş çürüklerini, geniş panoramik görüntülerdeki çürüklere göre daha iyi algılamaktadır. Bu durum, büyük görüntülerde küçük piksel oranlarını kapsayan çürük bölgelerinin daha zor algılanabildiğini kanıtlamaktadır. Aynı zamanda sonuçlar, Mask R-CNN mimarisinin, kırma gibi görevlerle görüntü boyutları çürük etiketine yaklaştıkça daha iyi bir performans gösterdiğini kanıtlamaktadır.

Sınıflandırma görevi için GoogLeNet, ShuffleNet, ResNet-50, MobileNet, DenseNet-169, DenseNet-121, VGG-19 ve öz-dikkat tabanlı yöntemlerden ConvNeXt mimarisi kullanılmıştır. Çalışma kapsamında sınıflandırma problemine bir sınıflandırıcı önerilmektedir. Sınıflandırma mimarilerinin veri dağılımında performanslarının ölçülebilmesi için bir veri seti araştırması gerçekleştirilmiştir. Veri seti araştırması gerçekleştirilirken, uzman etiketleri dikkate alınarak diş bölgelerinden kırılan görüntüler ile ana bir veri seti oluşturulmuştur. Daha sonrasında ana veri seti farklı sınıf oranlarında bölünmüştür. Çürük ve çürük değil sınıfları ile kategorilendirilen ana ve her bir alt veri setinde sınıflandırma ağlarının ve önerilen yöntemin performansı ölçülmüştür. Bu veri setleri içerisinde ana veri seti üzerinde önerilen yöntem %93,17 doğruluk, %89,43 kesinlik, %85,84 duyarlılık ve %87,49 f1-skoru metriğine ulaşarak en yüksek performansı göstermiştir. Bu durum, bir panoramik görüntüde çürük lezyonu içeren dişlerin bütün dişlere oranları dikkate alınarak oluşturulan doğal dağılımlı veri seti üzerinde yüksek performans gösteren sınıflandırıcıların

kliniŒel ortamda uygulanabilirliđinin bir göstergesidir. Aynı zamanda öz-dikkat tabanlı yöntemlerin de ürük teŒhisi için uyarlanabileceđi görölmektedir.

## KAYNAKLAR

1. Majanga, V. and Viriri, S., “A survey of dental caries segmentation and detection techniques”, *The Scientific World Journal*, 1–19 (2022).
2. Cheng, L., Zhang, L., Yue, L., Ling, J., Fan, M., Yang, D., Huang, Z., Niu, Y., Liu, J., Zhao, J., Li, Y., Guo, B., Chen, Z., and Zhou, X., “Expert consensus on dental caries management”, *International Journal Of Oral Science*, 14 (1): 1–8 (2022).
3. Fejerskov, O., and von der Fehr, F. R.. "Caries epidemiology with special emphasis on diagnostic standards.: In dental caries-the disease and its clinical management." *Caries Epidemiology With Special Emphasis On Diagnostic Standards*, Blackwell, Munksgaard Copenhagen (2003).
4. Banumathi, A., Vijayakumari, B., Geetha, A., Shanmugavadivu, N., and Raju, S., “Performance analysis of various techniques applied in human identification using dental x-rays”, *Journal Of Medical Systems*, 31 (3): 210–218 (2007).
5. Thillaikkarasi, R. and Saravanan, S., “An enhancement of deep learning algorithm for brain tumor segmentation using kernel based CNN with M-SVM”, *Journal Of Medical Systems*, 43 (4): 84 (2019).
6. Yaren Tekin, B., Ozcan, C., Pekince, A., and Yasa, Y., “An enhanced tooth segmentation and numbering according to FDI notation in bitewing radiographs”, *Computers In Biology And Medicine*, 146: 105547 (2022).
7. Mertens, S., Krois, J., Cantu, A. G., Arsiwala, L. T., and Schwendicke, F., “Artificial intelligence for caries detection: Randomized trial”, *Journal Of Dentistry*, 115: 103849 (2021).
8. Cantu, A. G., Gehrung, S., Krois, J., Chaurasia, A., Rossi, J. G., Gaudin, R., Elhennawy, K., and Schwendicke, F., “Detecting caries lesions of different radiographic extension on bitewings using deep learning”, *Journal Of Dentistry*, 100: 103425 (2020).
9. Ronneberger, O., Fischer, P., and Brox, T., “U-Net: Convolutional networks for biomedical image segmentation”, *18. Springer International Publishing*, Germany (2015).
10. Tan, M. and Le, Q., “EfficientNet: Rethinking model scaling for convolutional neural networks”, *International Conference On Machine Learning*, (2019).
11. Bayrakdar, I. S., Orhan, K., Akarsu, S., Çelik, Ö., Atasoy, S., Pekince, A., Yasa, Y., Bilgir, E., Sağlam, H., Aslan, A. F., and Odabaş, A., “Deep-learning approach for caries detection and segmentation on dental bitewing radiographs”, *Oral Radiology*, 38 (4): 468–479 (2022)



12. Lee, S., Oh, S., Jo, J., Kang, S., Shin, Y., and Park, J., “Deep learning for early dental caries detection in bitewing radiographs”, *Scientific Reports*, 11 (1): 16807 (2021).
13. Simonyan, K. and Zisserman, A., “Very deep convolutional networks for large-scale image recognition”, *arXiv preprint arXiv:1409.1556*, (2014).
14. Khan, H. A., Haider, M. A., Ansari, H. A., Ishaq, H., Kiyani, A., Sohail, K., Muhammad, M., and Khurram, S. A., "Automated feature detection in dental periapical radiographs by using deep learning", *Oral Surgery, Oral Medicine, Oral Pathology And Oral Radiology*, 131 (6): 711–720 (2021).
15. Lian, L., Zhu, T., Zhu, F., and Zhu, H., "Deep learning for caries detection and classification", *Diagnostics*, 11 (9): 1672 (2021).
16. Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q., "Densely connected convolutional networks", *2017 IEEE Conference On Computer Vision And Pattern Recognition (CVPR)*, Honolulu (2017).
17. Zhu, H., Cao, Z., Lian, L., Ye, G., Gao, H., and Wu, J., "CariesNet: A deep learning approach for segmentation of multi-stage caries lesion from oral panoramic X-ray image", *Neural Computing And Applications*, 35 (22): 16051–16059 (2023).
18. Fan, D.-P., Ji, G.-P., Zhou, T., Chen, G., Fu, H., Shen, J., and Shao, L., “PraNet: Parallel reverse attention network for polyp segmentation”, *International Conference On Medical Image Computing And Computer-Assisted Intervention* (2020).
19. Diakogiannis, F. I., Waldner, F., Caccetta, P., and Wu, C., "ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data", *ISPRS Journal Of Photogrammetry And Remote Sensing*, 162: 94–114 (2020).
20. Jiang, H., Zhang, P., Che, C., and Jin, B., "RDFNet: A Fast caries detection method incorporating transformer mechanism", *Computational And Mathematical Methods In Medicine* (2021).
21. He, K., Gkioxari, G., Dollar, P., and Girshick, R., "Mask R-CNN", *Proceedings Of The IEEE International Conference On Computer Vision*, pp. 2961-2969 (2017).
22. Ren, S., He, K., Girshick, R., and Sun, J., “Faster R-CNN: Towards real-time object detection with region proposal networks”, *Advances In Neural Information Processing Systems*, 28 (2015).
23. Redmon, J., Divvala, S., Girshick, R., and Farhadi, A., "You only look once: Unified, real-time object detection", *Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition*, pp. 779-788 (2016).
24. Qiu, S., Xu, X., and Cai, B., "FReLU: Flexible rectified linear units for improving convolutional neural networks", *24th International Conference On Pattern Recognition (ICPR)* (2018).

25. Sun, C., and Chen, H., "An attention-based transformer model for dental caries detection", *In International Conference On Electronic Information Engineering, Big Data, And Computer Technology (EIBDCT 2022)*, Vol. 12256. SPIE (2023).
26. Zhou, X., Yu, G., Yin, Q., Yang, J., Sun, J., Lv, S., and Shi, Q., "Tooth type enhanced transformer for children caries diagnosis on dental panoramic radiographs", *Diagnostics*, 13 (4): 689 (2023).
27. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., and Guo, B., "Swin Transformer: Hierarchical vision transformer using shifted windows", *Proceedings Of The IEEE/CVF International Conference On Computer Vision*, (2021).
28. Hossain, M., Rahman, Md. M., Syeed, M., Ummae, H., Uddin, M., and Mumu, S., "CaViT: Early stage dental caries detection from smartphone-image using vision transformer", (2023).
29. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., and Houlsby, N., "An image is worth 16x16 words: Transformers for image recognition at scale.", *arXiv preprint arXiv:2010.11929* (2020).
30. O'Shea, K. and Nash, R., "An introduction to convolutional neural networks." *arXiv preprint arXiv:1511.08458* (2015).
31. Mohammad-Rahimi, H., Motamedian, S. R., Rohban, M. H., Krois, J., Uribe, S. E., Mahmoudinia, E., Rokhshad, R., Nadimi, M., and Schwendicke, F., "Deep learning for caries detection: A systematic review", *Journal Of Dentistry*, 122: 104115 (2022).
32. GBD 2017 Oral Disorders Collaborators, Bernabe, E., Marcenes, W., Hernandez, C. R., Bailey, J., Abreu, L. G., ... and Kassebaum, N. J., "Global, regional, and national levels and trends in burden of oral conditions from 1990 to 2017: a systematic analysis for the global burden of disease 2017 study." *Journal Of Dental Research*, 99.4 362-373, (2020).
33. Parks, E. T. and Williamson, G. F., "Digital radiography: An overview", *The Journal Of Contemporary Dental Practice*, 3 (4): 23–39 (2002).
34. Muhamedagic, B. and Muhamedagic, L., "Digital radiography versus conventional radiography in dentistry", *Acta Informatica Medica*, 17.2 85 (2009).
35. Utami, A. P., Anggraeni, A., and Faesol, A., "Radiographer characteristics relationship with application patient safety radiation in sleman district", *Jurnal Ilmu Dan Teknologi Kesehatan STIKES Widya Husada*, 13 (2): (2022).
36. Yüce, K. and Barlas, B., "Tersanelerdeki radyografi uygulamaları ve risk değerlendirmesi", *Gemi Ve Deniz Teknolojisi*, (2020): 69–85 (2021).

37. İnternet: DentiAssist Veri Etiketleme Platformu, “DentiAssist uygulama ekranı”, <http://app.dentiassist.com> (2021).
38. Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., and Pietikäinen, M., "Deep learning for generic object detection: A survey", *International Journal Of Computer Vision*, 128 (2): 261–318 (2020).
39. Zaidi, S. S. A., Ansari, M. S., Aslam, A., Kanwal, N., Asghar, M., and Lee, B., "A survey of modern deep learning based object detection models", *Digital Signal Processing*, 126: 103514 (2022).
40. Girshick, R., Donahue, J., Darrell, T., and Malik, J., “Rich feature hierarchies for accurate object detection and semantic segmentation”, *Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition* (2014).
41. Uijlings, J. R. R., van de Sande, K. E. A., Gevers, T., and Smeulders, A. W. M., "Selective search for object recognition", *International Journal Of Computer Vision*, 104 (2): 154–171 (2013).
42. Girshick, R., "Fast R-CNN", *Proceedings Of The IEEE International Conference On Computer Vision*, (2015).
43. Hmidani, O. and Ismaili Alaoui, E. M., "A comprehensive survey of the R-CNN family for object detection", *2022 5th International Conference On Advanced Communication Technologies And Networking (Commnet)*, (2022).
44. Fan, Q., Brown, L., and Smith, J., "A closer look at Faster R-CNN for vehicle detection", *2016 IEEE Intelligent Vehicles Symposium (IV)*, (2016).
45. Hafiz, A. M. and Bhat, G. M., "A survey on instance segmentation: State of the art", *International Journal Of Multimedia Information Retrieval*, 9 (3): 171–189 (2020).
46. Lin, T.-Y., Dollar, P., Girshick, R., He, K., Hariharan, B., and Belongie, S., "Feature pyramid networks for object detection", *2017 IEEE Conference On Computer Vision And Pattern Recognition (CVPR)*, Honolulu (2017).
47. Nwankpa, C., Ijomah, W., Gachagan, A., and Marshall, S., “Activation functions: Comparison of trends in practice and research for deep learning.”, *arXiv preprint arXiv:1811.03378* (2018).
48. Nath, S. S., Mishra, G., Kar, J., Chakraborty, S., and Dey, N., "A survey of image classification methods and techniques", *2014 International Conference On Control, Instrumentation, Communication And Computational Technologies (ICCICCT)* (2014).
49. Thaller, H., Linsbauer, L., and Egyed, A., “Feature maps: A comprehensible software representation for design pattern detection”, *2019 IEEE 26th International Conference On Software Analysis, Evolution And Reengineering (SANER)* (2019).

50. Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P., "Gradient-based learning applied to document recognition", *Proceedings Of The IEEE*, 86 (11): 2278–2324 (1998).
51. Basha, S. H. S., Dubey, S. R., Pulabaigari, V., and Mukherjee, S., "Impact of fully connected layers on performance of convolutional neural networks for image classification", *Neurocomputing*, 378: 112–119 (2020).
52. Szegedy, C., Wei Liu, Yangqing Jia, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A., "Going deeper with convolutions", *2015 IEEE Conference On Computer Vision And Pattern Recognition (CVPR)*, Boston, MA, USA (2015).
53. Agarap, A. F., "Deep learning using rectified linear units (relu).", *arXiv preprint arXiv:1803.08375* (2018).
54. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R., "Dropout: A simple way to prevent neural networks from overfitting", *The Journal Of Machine Learning Research*, 15.1 1929-1958 (2014).
55. Zhang, X., Zhou, X., Lin, M., and Sun, J., "ShuffleNet: An extremely efficient convolutional neural network for mobile devices", *2018 IEEE/CVF Conference On Computer Vision And Pattern Recognition (CVPR)*, Salt Lake City, UT (2018).
56. Chollet, F., "Xception: Deep learning with depthwise separable convolutions", *2017 IEEE Conference On Computer Vision And Pattern Recognition (CVPR)*, Honolulu (2017).
57. Xie, S., Girshick, R., Dollar, P., Tu, Z., and He, K., "Aggregated residual transformations for deep neural networks", *2017 IEEE Conference On Computer Vision And Pattern Recognition (CVPR)*, Honolulu (2017).
58. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H., "Mobilenets: Efficient convolutional neural networks for mobile vision applications.", *arXiv preprint arXiv:1704.04861* (2017).
59. He, K., Zhang, X., Ren, S., and Sun, J., "Deep residual learning for image recognition", *2016 IEEE Conference On Computer Vision And Pattern Recognition (CVPR)*, Las Vegas, NV, USA, (2016).
60. Ioffe, S. and Szegedy, C., "Batch normalization: Accelerating deep network training by reducing internal covariate shift." *International Conference On Machine Learning* (2015).
61. Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., and Sutskever, I., "Learning transferable visual models from natural language supervision.", *International Conference On Machine Learning*, PMLR, pp. 8748-8763 (2021)

62. Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., and Soricut, R., “Albert: A lite bert for self-supervised learning of language representations.”, *arXiv preprint arXiv:1909.11942* (2019).
63. Jia, C., Yang, Y., Xia, Y., Chen, Y.-T., Parekh, Z., Pham, H., Le, Q. V., Sung, Y., Li, Z., and Duerig, T., “Scaling up visual and vision-language representation learning with noisy text supervision.”, *International Conference On Machine Learning*, PMLR, pp. 4904-4916 (2021).
64. Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., and Zettlemoyer, L., “BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension.”, *arXiv preprint arXiv:1910.13461* (2019).
65. Kolesnikov, A., Beyer, L., Zhai, X., Puigcerver, J., Yung, J., Gelly, S., and Houlsby, N., “Big transfer (BiT): General visual representation learning.”, *Computer Vision–ECCV 2020: 16th European Conference, August 23–28, Proceedings, Part V 16. Springer International Publishing*, Glasgow, UK (2020).
66. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I., “Attention is all you need”, *Advances In Neural Information Processing Systems*, 30 (2017).
67. Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K., “BERT: Pre-training of deep bidirectional transformers for language understanding”, *arXiv preprint arXiv:1810.04805*, (2018).
68. Radford, A., Narasimhan, K., Salimans, T., and Sutskever, I., “Improving language understanding by generative pre-training”, (2018).
69. Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., and Xie, S., “A convnet for the 2020s.” *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11976-11986 (2022).
70. Hendrycks, D. and Gimpel, K., “Gaussian error linear units (gelus).”, *arXiv preprint arXiv:1606.08415* (2016).
71. Internet: “PyTorch”, <https://pytorch.org/> (2023).
72. Internet: “TensorFlow”, <https://www.tensorflow.org/> (2023).
73. Internet: “Keras”, <https://keras.io/> (2023).
74. Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., and Savarese, S., “Generalized intersection over union: A metric and a loss for bounding box regression.”, *In Proceedings Of The IEEE/CVF Conference On Computer Vision And Pattern Recognition* (2019).
75. Internet: “Anaconda”, <https://www.anaconda.com/> (2023).

76. Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C. L., "Microsoft COCO: Common objects in context", *Computer Vision–ECCV 2014: 13th European Conference, September 6-12, 2014, Proceedings, Part V 13*. Springer International Publishing, Zurich, Switzerland (2014).
77. Karaoglu, A., Ozcan, C., Pekince, A., and Yasa, Y., "Numbering teeth in panoramic images: A novel method based on deep learning and heuristic algorithm", *Engineering Science And Technology, An International Journal*, 37: 101316 (2023).

## ÖZGEÇMİŞ

Dilara ÖZDEMİR ilk, orta ve lise öğrenimini Antalya’da tamamladı. 2016 yılında Karabük Üniversitesi Bilgisayar Mühendisliği Bölümü’nde lisans öğrenimine başladı. Kasım 2021 ve Ocak 2022 tarihleri arasında Karabük Üniversitesi Teknoloji Geliştirme Bölgesi’nde bulunan SimurgAI yazılım şirketinde Yapay Zeka Geliştiricisi olarak görev yaptı. Şubat 2021 ve Temmuz 2021 tarihleri arasında TÜBİTAK STAR kapsamında stajyer araştırmacı olarak görev aldı. 2021 yılında lisans eğitimini tamamlayarak onur derecesi ile mezun oldu. Mezun olduktan sonra Zonguldak Bülent Ecevit Üniversitesi Bilgisayar Mühendisliği Bölümü’nde Şubat 2022 ve Ocak 2023 tarihleri arasında Araştırma Görevlisi olarak çalıştı. 2023 Şubat ayı itibarıyla Amerika’da yer alan Mitsubishi Electric Innovation Center (MELIC) şirketinde UI/UX Designer olarak görev yapmaktadır.