



# Derin öğrenme ile panoramik radyografi görüntülerinden diş segmentasyonu: UNet, FPN ve PSPNet karşılaştırması

## Tooth segmentation from panoramic radiography images with deep learning: comparison of UNet, FPN and PSPNet

Salih Taha Alperen Özçelik<sup>1,\*</sup>, Hüseyin Üzen<sup>2</sup>, Abdülkadir Şengür<sup>3</sup>, Adalet Çelebi<sup>4</sup>

<sup>1</sup> Bingöl Üniversitesi, Elektrik-Elektronik Mühendisliği Bölümü, Mühendislik Fakültesi 12000, Bingöl, Türkiye

<sup>2</sup> Bingöl Üniversitesi, Bilgisayar Mühendisliği Bölümü, Mühendislik Fakültesi 12000, Bingöl, Türkiye

<sup>3</sup> Fırat Üniversitesi, Elektrik-Elektronik Mühendisliği Bölümü, Teknoloji Fakültesi 23000, Elazığ, Türkiye

<sup>4</sup> Mersin Üniversitesi, Ağız, Diş ve Çene Cerrahisi Bölümü, Diş Hekimliği Fakültesi 33000, Mersin, Türkiye

### Öz

Panoramik radyografi, diş hastalıklarının erken teşhis ve tedavisinde yaygın olarak kullanılan bir araçtır. Ancak bu tekniklerin başarısı, diş hekiminin deneyimine ve doğru yorum yapma yeteneğine bağlıdır. Bu çalışmada, insan hatasını azaltmak ve hızlı çözümler sunmak amacıyla panoramik radyografi görüntülerinden derin öğrenme ile diş segmentasyonu yapılmıştır. 18-65 yaş arasındaki hastalardan alınan 313 panoramik radyografi görüntüsü, deneyimli bir ağız ve çene cerrahisi tarafından piksel seviyesinde etiketlenmiştir. UNet, FPN ve PSPNet mimarileri ile segmentasyon yapılmış, VGG16 omurga ağı ile UNet modeli %93,74 F1 skoru, %88,22 KüB ve %98,25 doğruluk ile en iyi sonucu vermiştir. Ayrıca EfficientNet, ResNet50, InceptionV3, DenseNet121 ve MobileNet omurga ağları ile denemeler yapılmış, en yüksek performans EfficientNet omurga ağı ile %93,91 F1 skoru, %88,52 KüB ve %98,26 doğruluk olarak elde edilmiştir.

**Anahtar kelimeler:** Derin öğrenme, Diş segmentasyonu, UNet

### 1 Giriş

Ağız ve diş sağlığı sorunları, sadece günlük yaşam kalitesini ciddi şekilde etkilemekle kalmaz, aynı zamanda bazı durumlarda ciddi sağlık komplikasyonlarına ve hatta ölümlere yol açabilir [1, 2] Dünya sağlık örgütünün verilerine göre dünya çapında yaklaşık 3,5 milyar insan ağız ve diş sağlığı sorunlarıyla karşı karşıyadır [3]. Diş çürükleri, periodontal hastalıklar, ağız kanserleri ve diş kaybı bu sorunlardan başlıcalarıdır. Bu sorunlara erken tanı koyulması ve doğru bir tedavi planlanması hastalar için büyük önem arz etmektedir. Panoramik radyografi gibi ileri düzey görüntüleme teknikleri bu aşamada uzman diş hekimlerine yardımcı olarak kullanılan en yaygın araçlardır. Fakat bu süreçler, diş hekimi deneyimi ve doğru yorum yapma yeteneğiyle doğrudan ilişkilidir. İşte bu iki özelliğin yetersiz olması durumu yanlış tanı ve gereksiz tedavi gibi birçok sorunu da yanında getirmektedir [4]. Derin öğrenme yöntemleri tıbbi görüntülerin işlenmesi konusunda son yıllarda oldukça etkili bir araç haline gelmiştir [5].

### Abstract

Panoramik radiography is a widely used tool for the early diagnosis and treatment of dental diseases. However, the success of these techniques depends on the dentist's experience and ability to interpret correctly. In this study, dental segmentation was performed using deep learning on panoramic radiography images to reduce human error and provide faster solutions. A total of 313 panoramic radiographs from patients aged 18-65 were pixel-wise labeled by an experienced oral and maxillofacial surgeon. Segmentation was performed using UNet, FPN, and PSPNet architectures, and the UNet model with a VGG16 backbone achieved the best result with a 93.74% F1 score, 88.22% IoU, and 98.25% accuracy. Additionally, tests were conducted with EfficientNet, ResNet50, InceptionV3, DenseNet121, and MobileNet backbones, with EfficientNet showing the highest performance with a 93.91% F1 score, 88.52% IoU, and 98.26% accuracy.

**Keywords:** Deep learning, Tooth segmentation, UNet

Panoramik radyografi görüntülerinin kullanılarak dişlerin derin öğrenme ile otomatik segmente edilmesi diş hekimlerine erken teşhis ve tedavi planlamada yardımcı olabilecektir. Derin öğrenme sayesinde daha az hatayla ve daha yüksek hızda diş hastalıklarını tespit edebilecek ve daha fazla hastayla ilgilenebilmesini sağlayacaktır. Diş taramada derin öğrenme tabanlı segmentasyon tekniklerinin gelişmesi son yıllarda önemli bir ivme kazanmıştır. Dişlerin tespit edilmesi, segmentasyonu ve numaralandırılması için Mask-RCNN Evrişimsel Sinir Ağını (ESA) kullanan Oktay vd. [6], tarafından geliştirilen yöntem, her diş türüne benzersiz bir sınıf adı vererek çoklu sınıflandırma gerçekleştirmektedir. 200 görüntü üzerinde eğitilen ve 278 panoramik diş röntgeni üzerinde test edilen yöntemin, diş tespitinde ortalama başarı oranı %98, segmentasyon için F1 skoru ise %93 olarak belirlenmiştir. Diş görüntülerinin segmentasyonu için derin evrişimli sinir ağları ve seviye set yöntemini kullanan Yang vd. [7] Dental pulpu segmente ederek dişlerin konumunu belirlemeye odaklanmışlar ve bu amaçla U-Net modelini eğitmişlerdir. Model, 10 örnek içeren bir CBCT veri setinde

\* Sorumlu yazar / Corresponding author, e-posta / e-mail: sozcelik@bingol.edu.tr (S. T. A. Özçelik)

Geliş / Received: 07.06.2024 Kabul / Accepted: 23.08.2024 Yayınlanma / Published: 15.10.2024

doi: 10.28948/ngumuh.1497540

test edilmiştir. Önerdikleri model, ortalama Dice katsayısı 0.9791 ve ortalama Jaccard katsayısı 0.9595 olacak şekilde yüksek doğrulukta diş segmentasyonu sağlamıştır. Diş tanımlaması için Faster R-CNN ve diş numaralandırması için VGG-16 evrişim mimarisi kullanan Alam vd. [8], FDI iki haneli numaralandırma sistemine göre etiketleme yapmıştır. Kullanılan veri seti, 1500 panoramik diş röntgeni içerir ve bu görüntüler, The Southwest State University of Bahia (UESB)'deki teşhis görüntüleme merkezinde toplanmıştır. Diş tanımlama için %87.85 duyarlılık, %89.80 kesinlik ve %88.81 F-ölçüsü; diş numaralandırma için ise %84.03 duyarlılık ve %86.50 kesinlik elde edilmiştir. Diş segmentasyonu için Mask R-ESA ağ mimarisi kullanan Rubiu vd. [9] Tufts Dental veri kümesinden alınan görüntülere ISO 3950 standardına göre etiketleme gerçekleştirmiştir. Veri seti, 1000 panoramik dental radyograf içeren Tufts Dental Database'den elde edilmiştir. Ön işleme olarak, görüntü parlaklığı, kontrastı, kırılması ve boyutunun değiştirilmesi gibi veri artırma teknikleri uygulanmıştır. Elde edilen değerlendirme metrikleri arasında test setinde doğru sınıflandırılmış dişlerin yüzdesi %98,4 ve Dice skoru 0.87 olarak belirlenmiştir. Diş tanımlamak için U-Net tabanlı bir İlgili Alanı (İA) çıkarma modeli ve Hybrid Task Cascade tabanlı diş segmentasyonu ve numaralandırma modeli kullanan Xu vd. [10], 6046 panoramik radyograf içeren bir veri kümesi toplamış ve bu verilerin etiketlenmesi manuel olarak gerçekleştirilmiştir. Segmentasyon için %98.67 doğruluk ve %98.56 geri çağırma oranı elde eden araştırmacılar, numaralandırma için %97.83 doğruluk ve %97,73 geri çağırma oranı elde etmiştir. Ön işleme sürecinin daha da iyileştirilmesi gerekliliği ve diş numaralandırma hatalarını çözme ihtiyacı da çalışmanın eksiklikleri olarak sayılabilir. Panoramik diş röntgenlerinde otomatik diş numaralandırma için bir ESA geliştiren Privado vd. [11], 8000 anonimleştirilmiş panoramik diş görüntüsü kullandı. Bu veri seti, diş hekimliği alanında üç yıldan fazla deneyime sahip iki uzman tarafından etiketlendi. Nesne tespiti ve sınıflandırma olmak üzere iki ana katmandan oluşan bir ESA mimarisi kullanıldı. Nesne tespiti için Matterport Mask-R ESA ve sınıflandırma için ResNet101 kullanıldı. Araştırmacılar, modelin eğitimi için transfer öğrenme yöntemini de uyguladılar. Geliştirdikleri model, diş tespitinde %99.24 ve diş numaralandırmada %93.83 doğruluk oranı elde etti. Çalışmanın eksiklikleri arasında, bazı durumlarda diş eksikliğinin yanlış tespit edilmesi ve protez taçlarının tek diş olarak algılanması gösterilebilir. 1005 panoramik X-ray görüntüsünden oluşan bir veri seti kullanan Yüksel vd. [12], tedavi tespiti ve numaralandırma için toplam 1005 görüntü etiketlediler. Diş bölgelerinin etiketlenmesi için ise 500 görüntü etiketlediler. YOLO ve diğer nesne tespit modelleri gibi çeşitli derin öğrenme mimarileri kullanılmıştır. Ön işleme olarak, görüntülerin dönüştürülmesi ve veri artırma teknikleri uygulayan araştırmacılar, modelin performansını, %89.4 Ortalama Hassasiyet skoru (numaralandırma için) ve %59.0 (tedavi tespiti için) ile değerlendirmiştir. Çalışmanın eksiklikleri arasında, veri setinin küçük olması ve nadir tedavi türlerinin yetersiz temsili bulunmaktadır. Lin ve Chang [13], diş segmentasyonu ve numaralandırması için U-Net ve FCN-8s

mimarilerini kullandılar. Segmentasyon için poligonal etiketleme, numaralandırma için sınırlayıcı kutu yöntemleri tercih edildi. Veri seti, Tayvan'daki bir diş kliniğinden alınan panoramik diş röntgenlerinden oluşuyordu. Ön işleme olarak histogram eşitleme ve normalizasyon uygulandı. Değerlendirme metrikleri arasında %97.54 Kesişim Üzerinden Birleşim (KüB), %95.62 diş numaralandırma doğruluğu, %98.33 diş durumu tanımlama doğruluğu elde ettiler. Tian vd. [14], yüksek çözünürlüklü 3D diş modelleri içeren bir veri seti kullandılar. Segmentasyon için, derin ESA kullanan hiyerarşik bir diş segmentasyon mimarisi uygulandı. Bu mimari, iki seviyeli hiyerarşik diş sınıflandırma ağı ve üç seviyeli hiyerarşik diş segmentasyon modelini içeriyor. Segmentasyon sınırlarını iyileştirmek için koşullu rastgele alan modeli de kullanıldı. Ağ, Level-1'de %95.96 ve Level-2'de %88.06 sınıflandırma başarısı sağladı. Diş segmentasyonu başarısı ise %89.81 olarak kaydedildi. Tek bir 3D intraoral taramadan ne kadar temsilci bilginin öğrenilebileceğini inceleyen Jana vd. [15], on farklı yöntem kullanılarak bu durum nicel olarak değerlendirilmiştir. Tek bir 3B intraoral tarama ile eğitim yapıldığında, Dice skoru 0.86'ya kadar yükselebilirken, tam eğitim seti ile elde edilen Dice skoru 0.94 olarak bulunmuştur. Sonuç olarak, uygun koşullar altında, örneğin veri artırma ile, segmentasyon yöntemlerinin tek bir 3B diş nokta bulutu taramasından önemli miktarda bilgi öğrenebileceği sonucuna varılmıştır. KIBT taramalarında çene ve dişlerin doğru bir şekilde segmentasyonu için Wang vd. [16] tarafından bir yöntem geliştirildi. Otuz KIBT taraması kullanılarak bir karışık ölçekli yoğun (KÖ-Y) evrişimli sinir ağı eğitildi. Ağın performansı, Dice benzerlik katsayısı ve yüzey sapması gibi ölçümlerle değerlendirildi. Sonuçlar, KÖ-Y ağıyla segmente edilen KIBT taramalarının altın standart segmentasyonlarla büyük bir örtüşme gösterdiğini (Dice benzerlik katsayısı: çene için  $0.934 \pm 0.019$ , dişler için  $0.945 \pm 0.021$ ) ve yüzey sapmasının da küçük olduğunu (çene için  $0.390 \pm 0.093$  mm, dişler için  $0.204 \pm 0.061$  mm) gösterdi. Ayrıca, KÖ-Y ağının bir KIBT taramasını segmente etmek için yaklaşık 25 saniye sürdüğü, manuel segmentasyonun ise yaklaşık 5 saat aldığı bulundu. Bu çalışma, çene ve dişlerin otomatik segmentasyonunda KÖ-Y ağının başarılı olduğunu ve bu yöntemin KIBT taramalarında çoklu sınıf segmentasyonunun gerçekleştirilmesini sağladığını gösterdi.

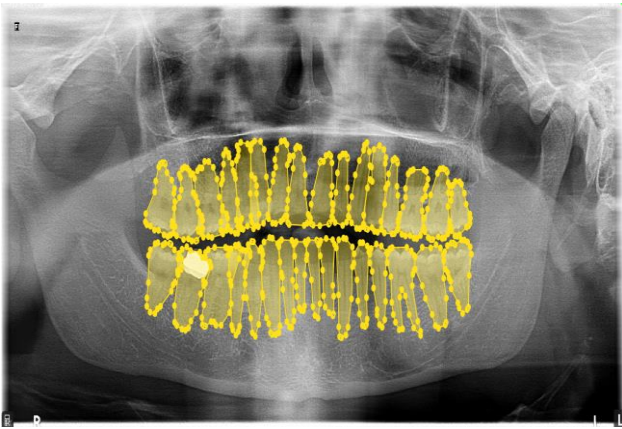
Bu çalışmada 313 panoramik radyografi görüntüsünden oluşan bir veri kümesi oluşturulmuştur. Bu görüntülerdeki dişler uzman diş hekimi tarafından etiketlenmiştir. Elde edilen veri kümesi UNet, FPN ve PSPNet mimarileri kullanılarak eğitilmiş ve segmente edilmiştir. Kullanılan verilerin %80'i eğitime %20'si teste ayrılmıştır. Bu üç mimariyi kıyaslama amaçlı ilk segmentasyon VGG16 omurga ağı ile gerçekleştirilmiştir. Elde edilen segmentasyon sonucunda %93.74 F1 skor ve %88,22 KüB ve %98.25 doğruluk ile en yüksek segmentasyon başarısını UNet göstermiştir. Ayrıca en yüksek başarıyı veren UNet ağ mimarisi omurga ağları değiştirilerek tekrar eğitilmiştir. Omurga ağı olarak Efficientnet [17], ResNet50 [18], InceptionV3 [19], DenseNet121 [20] ve MobileNetv2 [21] mimarileri kullanılmıştır. UNet üzerinden tüm omurga ağları performans açısından kıyaslanmıştır. En yüksek başarıyı

%93.91 F1 skoru ve %88.52 KüB ve %98.26 doğruluk ile EfficientNet omurga ağı sağlamıştır.

## 2 Materyal ve metod

### 2.1 Veri kümesi

Bu çalışmada kullanılan veri kümesi, çeşitli yaş gruplarından ve farklı diş problemleri olan hastaların özel diş kliniklerinden elde edilen 500 panoramik radyografi görüntüsünden oluşmuştur. Görüntüler, 18-65 yaş arasındaki hastaları kapsamaktadır ve veri seti, kist, çene kırığı gibi spesifik durumlar hariç tutularak oluşturulmuştur. Veri kümesi, 18-65 yaş arasındaki hastaları içermekte olup kist, çene kırığı gibi durumlar ve çözünürlüğü düşük görüntüler veri kümesine dahil edilmemiştir. Bu şartlara uyularak gerçekleştirilen eliminasyon süreci sonucunda 313 panoramik radyografi görüntüsünden oluşan bir veri kümesi oluşturulmuştur. Veri kümesi hastaların isim ve cinsiyet gibi kişisel bilgilerini içermeyen anonim olarak toplanmıştır. Bu retrospektif araştırma 2023/13-37 numaralı etik kurul kararıyla Fırat Üniversitesi Hastanesi Etik kurulu tarafından onaylanmıştır. Panoramik radyografi görüntüleri (2700-3000) x 1316 piksel çözünürlükte ve 8 bit derinlikte jpeg formatında kaydedilmiştir. Elde edilen görüntüler 10 yıl üzerinde deneyime sahip ağız ve çene cerrahı tarafından her bir diş piksel seviyesinde çokgensel etiketleme (PSÇE) yöntemiyle etiketlenmiştir. Etiketlenmiş bir diş radyografisi görüntüsü Şekil 1’de verilmiştir.



Şekil 1. Etiketlenmiş bir panoramik diş radyografisi

Literatürdeki çalışmaların birçoğunda kullanılan sınırlayıcı kutu (SK) yöntemi ile diş etiketlemesi gerçekleştirilmiştir. Bu yöntem dişin tamamını bir dikdörtgen kutuya çerçevesine alarak etiketlenmesine verilen isimdir. PSÇE ile SK karşılaştırıldığında PSÇE, gerçek diş sınırlarını daha keskin bir şekilde temsil etmesi, arka plan bilgisini minimize etmesi sayesinde dişlerin daha doğru bir şekilde arka plandan ayrışmasını sağlayacaktır.

### 2.2 Önerilen model

Elde edilen veri kümesi FPN [22], PSPNet [23] ve UNet [24] ile eğitilmiştir. Modelde verilerin %80’i eğitim için %20’si test için kullanılmıştır. Giriş görüntüleri 256x512 boyutuna uygulanarak veri artırma tekniği uygulanmıştır.

### 2.2.1 FPN

256x512 boyutlarındaki panoramik diş röntgen görüntüsü ağ mimarisine giriş olarak verilmiştir. Ağ bu gelen giriş görüntüsünden farklı çözünürlüklerde özellik haritaları çıkarır. VGG16’nın [25] belli katmanlarından elde edilen öznelik haritaları  $C_2, C_3, C_4, C_5$  Denklem (1)’deki gibi tanımlanır:

$$\begin{aligned} C_2 &= Conv2_2(I) \\ C_3 &= Conv3_3(C_2) \\ C_4 &= Conv4_3(C_3) \\ C_5 &= Conv2_2(C_4) \end{aligned} \quad (1)$$

Burada  $I$  giriş görüntüsü  $ConvX_y$  ise VGG16’nın  $x$ ’inci katmanın  $y$ ’inci evrişim bloğunu temsil eder.

Üst seviyeden alt seviyeye doğru bir yol kullanarak daha yüksek seviyedeki özneliklerle düşük seviyedeki öznelik haritalarıyla birleştirir. Yukarı örnekleme işlemi bilinear interpolasyon yöntemi ile gerçekleştirilir. Bu işlem Denklem (2)’de verilmiştir.

$$P_l = Upsample(P_{l+1}) \quad (2)$$

Burada  $P_{l+1}$ ,  $l + 1$  seviyesindeki öznelik haritasını  $P_l$  ise yukarı örneklenmiş  $l$  seviyesindeki özellik haritasını temsil eder. Bu yukarı aşağı yolunda daha yüksek seviyeli öznelik haritaları yukarı örneklenir. Aynı boyutlardaki düşük seviyeli öznelik haritalarıyla birleştirilir. Yan bağlantılar kullanılarak temel ağın farklı seviyelerinden gelen öznelikler yukarı örneklenen haritalarla birleştirilerek her seviyede daha detaylı öznelikler elde edilir. Bu birleşim Denklem (3)’te verilmiştir.

$$P_l = Conv(C_l) + Upsample(P_{l+1}) \quad (3)$$

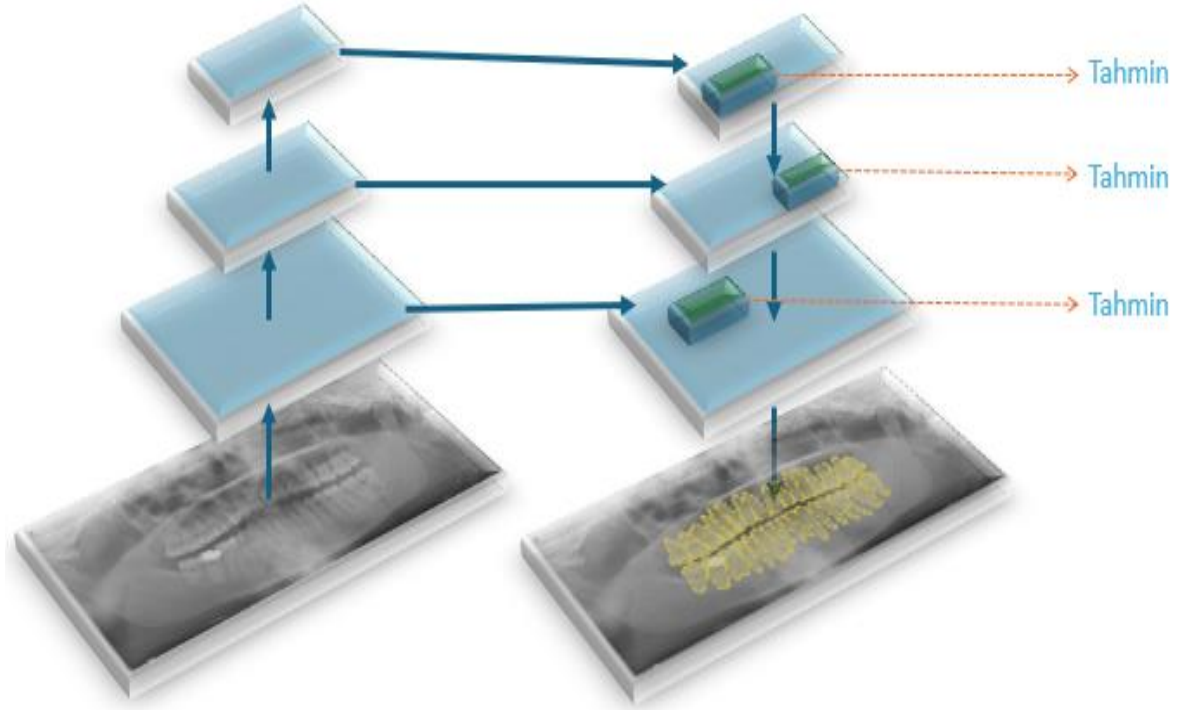
Burada  $C_l$  temel ağın  $l$  seviyesindeki özellik haritasını conv ise 1x1 evrişim işlemini temsil eder. Bu birleşim her seviye için tekrarlanarak gerçekleştirilir.

$$\begin{aligned} P_5 &= Conv(C_5) \\ P_4 &= Conv(C_4) + Upsample(P_5) \\ P_3 &= Conv(C_3) + Upsample(P_4) \\ P_2 &= Conv(C_2) + Upsample(P_3) \end{aligned} \quad (4)$$

Birleşim işlemlerinden sonra, her seviyede oluşturulan özellik haritaları daha fazla işlenir ve son özellik haritaları (FPN’nin çıktıları) elde edilir:

$$\begin{aligned} P_2' &= Smooth(P_2) \\ P_3' &= Smooth(P_3) \\ P_4' &= Smooth(P_4) \\ P_5' &= Smooth(P_5) \end{aligned} \quad (5)$$

Burada  $Smooth$  işlemi 3x3 evrişimlerle özellik haritalarının daha yumuşatılmasını ifade eder. FPN, bu birleşimden sonra her seviyedeki öznelik haritalarını işleyerek segmentasyonda kullanılmak üzere çıktı öznelik haritaları oluşturur.

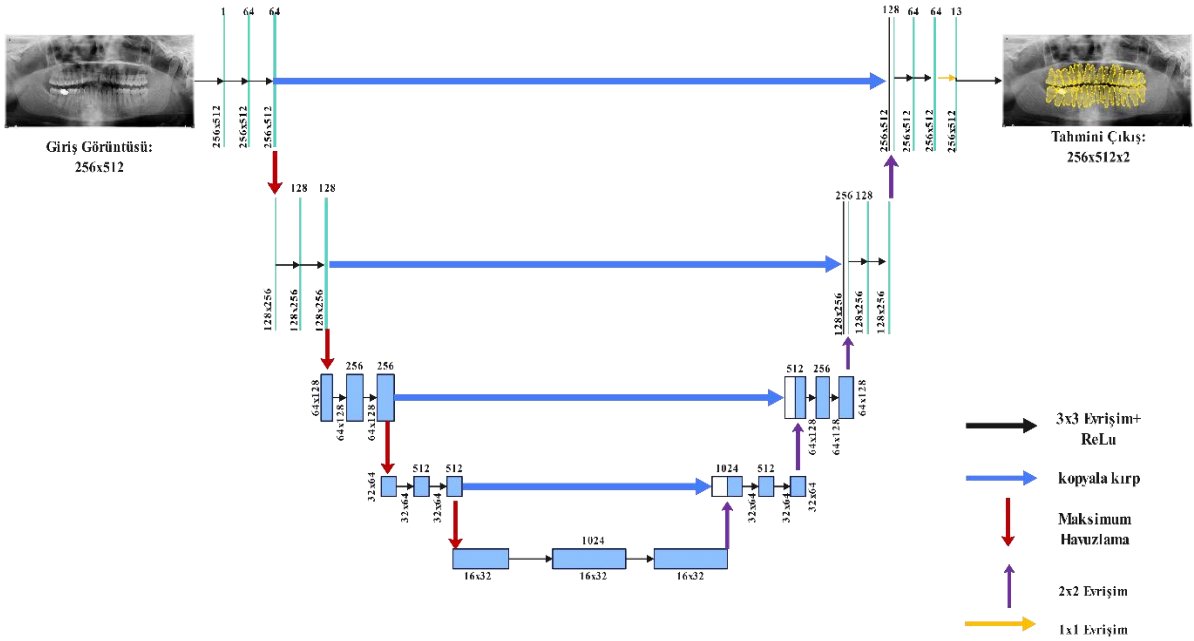


Şekil 2. Kullanılan FPN modeli

Şekil 2’de kullanılan FPN mimarisinin görseli verilmiştir. Mavi olan bloklar FPN modelinin farklı katmanlarını temsil etmektedir. Bu katmanlar girdiden başlayarak yukarı doğru çıkar ve her seviyede öznetelik haritaları oluşturur. Her katmanda belirli bir seviyesinde öznetelik çıkarılır. Lacivert oklar ise bir katmandan başka bir katmana öznetelik taşınmasını ifade eder. FPN’de, her katmanın çıkışı, bir sonraki üst katmanın girişi olarak kullanılır ve bu şekilde özelliklerin daha yüksek seviyelere taşınması sağlanır.

### 2.2.2 Unet

Unet mimarisi kodlayıcı ve kod çözücü yapısına sahip bir ESA’dır. Bu yapı kodlayıcı kod çözücü yapısı sayesinde diş segmentasyonu probleminde yüksek çözünürlüklü çıktılar üretebilir. Adımı U şeklindeki yapısından almıştır. Bu yapı Şekil 3’te verilmiştir.



Şekil 3. Kullanılan Unet modeli



**Kodlayıcı:** Kodlayıcı kısım girişte verilen 256x512 panoramik radyografi görüntüsünden öznetelik çıkarımı yapar. Bu Öznetelikler daha düşük çözünürlüklü özneteliklere dönüştürülür. Kodlayıcı kısım Evrişim ve ReLu katmanları (siyah ok) ve maksimum havuzlama (kırmızı ok) katmanlarından oluşur. Her bir katmanda, 3x3 konvolüsyon işlemi uygulanır ve ardından ReLu aktivasyon fonksiyonu kullanılır. Konvolüsyon işlemlerinden sonra, 2x2 maksimum havuzlama işlemi uygulanarak görüntü boyutları yarıya indirilir. Bu, ağına daha büyük ölçekli özellikleri öğrenmesini sağlar. Encoder bölümünde her bir aşamada kanal sayısı artırılır. Böylece model daha fazla ve daha karmaşık özellikler öğrenebilecektir.

**Bottleneck Katmanı:** Kodlayıcı ve kod çözücü bölümleri arasında bulunan bu katman, en düşük boyuttaki öznetelik haritalarını içerir ve en fazla kanal sayısına sahiptir (örneğin, 1024 kanal). Bu katman, ağına derin özellikleri öğrenmesi için kullanılır.

**Kod Çözücü:** Kodlayıcıdan gelen öznetelik haritalarını kullanarak boyutları tekrar artırır ve orijinal giriş boyutuna geri döndürür. 2x2 transpoze konvolüsyon işlemi (mor oklar) ile boyutlar iki katına çıkarılır. Bu işlemler, öznetelik haritalarının yeniden yapılandırılmasına yardımcı olur. Kodlayıcı bölümünden gelen karşılık gelen öznetelik haritaları ile birleştirme işlemi yapılır. Bu, yüksek çözünürlüklü bilgilerin korunmasını sağlar ve segmentasyonun doğruluğunu artırır. Yine, her bir yukarı örnekleme adımından sonra 3x3 konvolüsyon ve ReLU aktivasyon fonksiyonu uygulanır. Bu, yeniden yapılandırılan öznetelik haritalarının işlenmesini sağlar. Son katmanda, 1x1 konvolüsyon işlemi uygulanarak kanal sayısı hedef sınıf sayısına eşitlenir. Bu, her pikselin hangi sınıfa ait olduğunu belirlemek için kullanılır.

### 2.2.3 PSPNet

Farklı ölçeklerdeki bağlamsal özellikleri öğrenerek segmentasyon gerçekleştiren bir derin öğrenme modelidir. Giriş katmanında 256x512 boyutundaki panoramik diş radyografi görüntüsünü alarak bunu özellik çıkarma katmanına aktarılır. Evrişimsel katmanlara ve aktivasyon fonksiyonlarına tabi tutularak diş radyografi görüntüsünden derin öznetelikler elde edilir. Bu öznetelikler piramitsel havuzlama modülünü aktarılır. Bu modül farklı ölçeklerdeki bağlamsal bilgilerin yakalanmasını sağlar.

Şekil 4'te kullanılan PSPNet bloğu verilmiştir. Modelin girişinde yer alan mavi blok, ilk evrişim katmanı olarak görev yapar ve giriş görüntüsünden temel özellikleri çıkarır. Bu özellikler daha sonra yeşil blok tarafından işlenir; bu ara katman mavi bloktan gelen özellikleri daha yüksek seviyede

analiz eder. Ardından, pembe blok daha derin evrişim katmanı olarak işlev görür ve yeşil bloktan gelen bilgileri daha detaylı bir şekilde işler. Turuncu blok ise en üst evrişim katmanı olup, en soyut ve yüksek seviyedeki özellikleri çıkarır. Elde edilen bu özellikler, beyaz blok tarafından birleştirilir; bu blok havuzlama katmanı olarak işlev görür. Havuzlama işlemi sonrası, özellik haritaları renkli bloklar aracılığıyla üst örnekleme ile orijinal çözünürlüğe geri dönüştürülür.

### 2.3 Değerlendirme metrikleri

Önerilen modelin sınıflandırma performansı Kesişim Üzerinden Birleşim(KüB), F1 skoru, kesinlik, doğruluk ve hatırlama gibi çeşitli değerlendirme metrikleri kullanılarak değerlendirildi.

#### 2.3.1 Kesişim üzerinden birleşim (KüB)

KüB, tahmin edilen ve gerçek segmentasyon bölgelerinin ne derece iyi örtüşüğünü ölçen bir metriktir. KüB kesişim bölgesinin birleşim bölgesine oranı olarak hesaplanır ve şu şekilde ifade edilir:

$$KüB = \frac{|Gerçek Pozitif \cap Tahmin Pozitif|}{|Gerçek Pozitif \cup Tahmin Pozitif|} \quad (6)$$

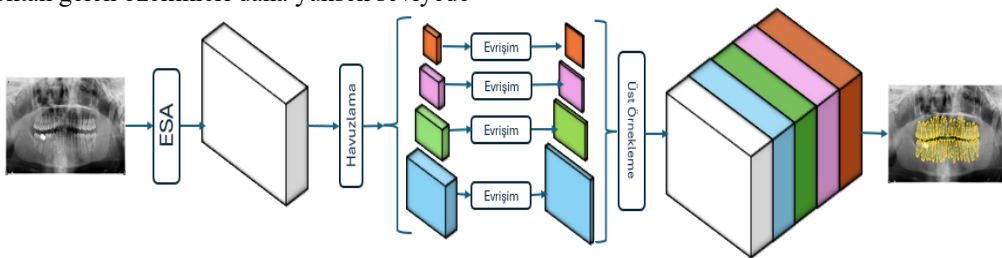
Burada kesişim bölgesi tahmin edilen ve gerçek pozitif bölgelerin ortak kısmını temsil eder. Birleşim bölgesi ise tahmin edilen ve gerçek pozitif bölgelerin toplamını temsil etmektedir. Sonuç olarak KüB değeri 1'e ne kadar yakınsa segmentasyon o kadar iyidir.

#### 2.3.2 F1 skoru

F1 skoru, modelin performansını değerlendirirken hem kesinlik hem de duyarlılık değerlerini göz önünde bulunduran bir metriktir. Bu skor, özellikle veri setindeki sınıfların dengesiz olduğu durumlarda büyük önem taşır. F1 skoru, kesinlik ve duyarlılığın harmonik ortalaması olarak hesaplanır ve şu şekilde formüle edilir:

$$F1 Skoru = 2 \times \frac{Kesinlik \times Duyarlılık}{Kesinlik + Duyarlılık} \quad (7)$$

Harmonik ortalama kullanımı hem kesinlik hem de duyarlılık değerlerinin dengelenmesine yardımcı olur ve her iki metriğin de düşük olduğu durumlarda daha düşük bir F1 skoru elde edilir. Bu sayede, modelin genel performansı hakkında daha dengeli bir bilgi edinilmiş olur.



Şekil 4. Kullanılan PSPNet modeli

### 2.3.3 Kesinlik

Kesinlik, modelin pozitif olarak tahmin ettiği örneklerin gerçekte ne kadarının doğru olduğunu belirtir. Bu metrik, yanlış pozitiflerin (modelin pozitif olarak sınıflandırdığı fakat aslında negatif olan örnekler) oranını belirlemek için kullanılır. Kesinlik şu formül ile hesaplanır:

$$Kesinlik = \frac{Gerçek Pozitifler}{Gerçek Pozitifler + Yanlış Pozitifler} \quad (8)$$

Burada yanlış pozitif modelin pozitif olarak sınıflandırdığı fakat aslında negatif olan örnekleri temsil eder. Gerçek pozitif ise, modelin doğru bir şekilde pozitif olarak sınıflandırdığı örneklerdir.

### 2.3.4 Duyarlılık

Duyarlılık, modelin gerçek pozitif örnekleri doğru bir şekilde tanımlama yeteneğini ifade eder. Bu metrik, yanlış negatiflerin etkisini ölçmek için kullanılır. Gerçek pozitif, modelin doğru bir şekilde pozitif olarak sınıflandırdığı örneklerdir. Yanlış negatif ise modelin negatif olarak tahmin ettiği ancak gerçekte pozitif olan örneklerdir. Duyarlılık şu formül ile hesaplanır:

$$Duyarlılık = \frac{Gerçek Pozitifler}{Gerçek Pozitifler + Yanlış Negatifler} \quad (9)$$

## 3 Bulgular ve tartışma

Deneysel çalışmalar RTX 2080Ti i5 12400f işlemci ve 32 GB RAM'e sahip bir sistemle yürütülmüştür. Önerilen modelin etkinliği F1 skor, KüB, hatırlama, duyarlılık ve doğruluk metrikleriyle ölçülmüştür. Etiketlenmiş 313 panoramik radyografi görüntüsü 256x512 boyutuna getirilmiştir. Veri artırma uygulanmıştır. Verilerin %80'i eğitime %20'si teste ayrılmıştır. Veri kümesi FPN, PSPNet ve Unet kullanarak eğitilmiş ve test edilmiştir. Default olarak VGG16 omurga ağı ile gerçekleştirilen eğitim için elde edilen metrikler Tablo 1'de verilmiştir. 16 batch boyutu ile gerçekleştirilen eğitimde, başlangıç öğrenme oranı 0,0001 olarak belirlenmiştir. Eğitim süreci boyunca Adam optimizer kullanılmıştır ve eğitim 100 epoch boyunca sürdürülmüştür. Maksimum yığın boyutu, eğitim sırasında bellek yönetimi ve hesaplama verimliliği göz önünde bulundurularak optimize edilmiştir. Öğrenme oranı sabit tutulmuş ve herhangi bir öğrenme oranı azaltıcı (learning rate reducer) kullanılmamıştır. Model performansı her epoch sonunda doğrulama verileri ile değerlendirilmiş ve doğrulama doğruluğu ile eğitim kaybı izlenmiştir.

Tablo 1'de görüldüğü üzere, en iyi performansı UNet VGG16 ağı sağlamıştır. UNet VGG16 ile elde edilen F1 skoru 93.74, KüB 88.22, kesinlik 94.16, duyarlılık 93.32 ve doğruluk 98,25'tir. FPN VGG16 ağı da yüksek performans göstermiştir; F1 skoru %92,16, KüB %85.47, kesinlik %90,72, duyarlılık %93,65 ve doğruluk %97.76 olarak ölçülmüştür. PSPNet VGG16 ağı ise diğer iki modele göre

daha düşük performans sergilemiştir; F1 skoru %89.60, KüB %81.16, kesinlik %87.18, duyarlılık %92.16 ve doğruluk %97.00 olarak elde edilmiştir. Sonuçlar, UNet VGG16 ağının diğer modellere kıyasla daha üstün bir performans sergilediğini göstermektedir. UNet modelinin diğer modellere kıyasla daha yüksek performans sergilemesi çalışma devamında diğer omurga ağları ile kıyaslanması için UNet üzerinden devam edilmesine karar kılınmıştır.

UNet ağ modeli için farklı omurga ağları denenerek omurga ağlarının dış segmentasyonu problemi için performansları kıyaslanmıştır. Omurga ağı olarak efficientnet, ResNet50, InceptionV3, DenseNet121 ve MobileNetv2 mimarileri kullanılmıştır. Elde edilen sonuçlar Tablo 2'deki gibidir

UNet EfficientNetb0, 93.91 gibi yüksek bir F1 skoru ve 95,34 duyarlılık değeri ile genel olarak çok güçlü bir performans sergilemiştir. Bu model, doğru pozitif ve doğru negatifleri oldukça iyi tanımlayarak duyarlılığı yüksek tutmuştur. Ayrıca, 88.52 Küçük Boyut (KüB) değeri ile diğer modellere göre biraz daha büyük boyutlu olsa da, performansı ile bu farkı dengelemektedir. UNet ResNet50, 93.85 F1 skoru ve 98.27 doğruluk değeri ile en yüksek doğruluk oranını elde etmiştir. Bu model, 93.73 kesinlik değeri ile doğru pozitif tahminlerinde yüksek isabet oranına sahiptir. ResNet50, genel doğruluk açısından en iyi performansı sergileyerek güvenilir bir model olduğunu göstermektedir.

UNet InceptionV3, 93.81 F1 skoru ve 88.35 KüB değeri ile öne çıkarken, 94.82 kesinlik değeri ile en yüksek kesinlik oranına sahiptir. Bu model, doğru pozitif tahminlerinde üstün başarı gösterirken, duyarlılık değeri 92.83 ile diğer modellere göre biraz daha düşüktür. Bu da modelin yanlış negatif tahminlere biraz daha açık olduğunu göstermektedir. UNet MobileNetv2, 93.79 F1 skoru ve 88.31 KüB değeri ile dengeli bir performans sergilemektedir. 93.35 kesinlik ve 94.24 duyarlılık değerleri ile hem doğru pozitif hem de doğru negatif tahminlerde iyi bir denge sağlamaktadır. MobileNetv2, hem kesinlik hem de duyarlılık açısından güvenilir bir modeldir. UNet VGG16, 93.74 F1 skoru ve 88.22 KüB değeri ile dikkat çekmektedir. Bu model, 94.16 kesinlik değeri ile doğru pozitif tahminlerinde yüksek isabet oranına sahip olup, 93.32 duyarlılık değeri ile doğru negatif tahminlerde de başarılıdır. VGG16, genel olarak dengeli ve yüksek performans sunarak diğer modeller arasında öne çıkmaktadır. UNet DenseNet121, 93.65 F1 skoru ve 88.06 KüB değeri ile en yüksek duyarlılık (95.36) değerine sahiptir. Bu model, doğru negatif tahminlerde oldukça başarılı olup, duyarlılığı yüksek tutarak yanlış negatif oranını minimize etmektedir. Ancak, 92.00 kesinlik değeri ile doğru pozitif tahminlerde diğer modellere nazaran biraz daha düşük performans sergilemektedir.

Genel değerlendirmede, EfficientNetb0 ve ResNet50, genel doğruluk ve F1 skoru açısından öne çıkarken, InceptionV3 ve DenseNet121 belirli ölçütlerde üstün performans göstermektedir. MobileNetv2 ve VGG16 ise dengeli performans sunarak hem kesinlik hem de duyarlılık açısından iyi sonuçlar elde etmişlerdir.

**Tablo 1.** UNet FPN ve PSPNet için elde edilen sonuçlar

Ağ Mimarisi	F1 Skor (%)	KüB (%)	Kesinlik (%)	Duyarlılık (%)	Doğruluk (%)
Unet VGG16	93.74	88.22	94.16	93.32	98.25
FPN VGG16	92.16	85.47	90.72	93.65	97.76
PSPNet VGG16	89.60	81.16	87.18	92.16	97.00

**Tablo 2.** Farklı omurga ağları için elde edilen sonuçlar

Kullanılan Omurga Ağı	F1 Skor (%)	KüB (%)	Kesinlik (%)	Duyarlılık (%)
Unet Efficientnetb0	93.91	88.52	92.52	95.34
Unet Resnet50	93.85	88.41	93.73	93.96
Unet İnceptionv3	93.81	88.35	94.82	92.83
Unet Mobilenetv2	93.79	88.31	93.35	94.24
Unet Vgg16	93.74	88.22	94.16	93.32
Unet Densenet121	93.65	88.06	92.00	95.36

#### 4 Sonuçlar

Bu çalışma, panoramik radyografi görüntüleri üzerinde diş segmentasyonu için farklı segmentasyon modellerinin etkinliğini değerlendirmek amacıyla gerçekleştirilmiştir. Özellikle UNet, FPN ve PSPNet gibi derin öğrenme tabanlı mimarilerin performansları karşılaştırılmış ve her bir modelin doğruluk, hassasiyet ve F1 skoru gibi metrikler açısından analizi yapılmıştır. Kullanılan UNet, FPN ve PSPNet mimarileri arasında, özellikle UNet modeli VGG16 omurga ağı ile en yüksek segmentasyon doğruluğunu elde etmiştir. Bu sonuç, literatürdeki diğer çalışmalarda da sıkça vurgulanan UNet'in diş segmentasyonu için uygun bir model olduğunu doğrulamaktadır [26]. UNet modeli, EfficientNet, ResNet50, InceptionV3, DenseNet121 ve MobileNet gibi farklı omurga ağlarıyla yeniden eğitilmiş ve performansları karşılaştırılmıştır. Bu omurga ağlarının her biri, segmentasyon doğruluğu ve diğer değerlendirme metrikleri açısından farklı performans sergilemiştir. Özellikle EfficientNetb0 ve ResNet50, genel doğruluk ve F1 skoru açısından öne çıkarken, InceptionV3 ve DenseNet121 bazı metriklerde iyi bir performans göstermiştir. MobileNetv2 ve VGG16 ise dengeli bir performans sunarak hem kesinlik hem de duyarlılık açısından oldukça başarılı sonuçlar elde etmiştir. Elde edilen bu sonuçlar, derin öğrenme tabanlı diş segmentasyonu modellerinin, el ile gerçekleştirilen segmentasyona kıyasla büyük bir zaman tasarrufu sağladığını ve yüksek doğrulukla çalıştığını göstermektedir. Ayrıca, kullanılan veri artırma teknikleri ve farklı omurga ağları ile yapılan denemeler, model performansını optimize etme konusunda önemli katkılar sağlamıştır. Bu çalışmada, özellikle veri kümesinin etiketleme sürecindeki dikkat ve titizlik, elde edilen yüksek performanslı sonuçların temelini oluşturmuştur. Ancak, çalışmanın bazı sınırlamaları da bulunmaktadır. Öncelikle, kullanılan veri kümesi belirli bir bölgeden ve belirli demografik özelliklere sahip hastalardan elde edilmiştir. Bu nedenle, model başarısını artırmak için daha geniş ve çeşitli veri kümelerinde denenebilir. Ayrıca, segmentasyon doğruluğunun artırılması için doğrudan diş segmentasyonu problemine özgü bir mimari tasarlanabilir.

Gelecek çalışmalar, farklı diş hastalıklarının ve anomalilerin otomatik tespiti ve sınıflandırılması için daha gelişmiş derin öğrenme modellerinin geliştirilmesine odaklanmalıdır. Ayrıca, modellerin etkinliği ve güvenilirliğini test etmek için bir diş hekimi ve derin

öğrenme modellerinin segmentasyon konusunda kıyaslanması gibi daha geniş kapsamlı deneyler de klinik ortamında yapılabilir. Sonuç olarak, derin öğrenme tabanlı segmentasyon teknikleri, diş hekimliği alanında erken teşhis ve tedavi planlamada önemli bir araç olarak gelecekteki çalışmalarda kullanılabilir ve bu alanda yapılan çalışmaların devamı, diş sağlığı hizmetlerinin kalitesini artırmaya yönelik önemli katkılar sağlayacaktır.

#### Çıkar çatışması

Yazarlar çıkar çatışması olmadığını beyan etmektedir.

#### Benzerlik oranı (iThenticate): %8

#### Kaynaklar

- [1] S.A. Gill, Quinonez R.B., M. Deutchman, C.E. Conklin, D. Rizzolo, D. Rabago, H. Silk. Integrating Oral Health into Health Professions School Curricula. Medical Education Online, 27 (1). 2022. <https://doi.org/10.1080/10872981.2022.2090308>.
- [2] Oral health: A window to your overall health - Mayo Clinic. (n.d.). Retrieved May 26, 2024, from <https://www.mayoclinic.org/healthy-lifestyle/adult-health/in-depth/dental/art-20047475>
- [3] WHO. Oral health. Retrieved December 25, 2023, from <https://www.who.int/news-room/fact-sheets/detail/oral-health>
- [4] R. Izzetti, , M. Nisi, G. Aringhieri, L. Crocetti, F. Graziani and C. Nardi. Basic knowledge and new advances in panoramic radiography imaging techniques: a narrative review on what dentists and radiologists should know. Applied Sciences, 11 (17), 7858. 2021. <https://doi.org/10.3390/APP11177858>
- [5] H. Zhang, and Y. Qie. Applying deep learning to medical imaging: a review. Applied Sciences 13 (18), 10521. 2023. <https://doi.org/10.3390/APP131810521>
- [6] A.B. Oktay, and A. Gurses. Detection, segmentation, and numbering of teeth in dental panoramic images with mask regions with convolutional neural network features. State of the Art in Neural Networks and Their Applications: 1, 73–90. 2021. <https://doi.org/10.1016/B978-0-12-819740-0.00004-8>.
- [7] Y. Yang, R. Xie, W. Jia, Z. Chen, Y. Yang, L. Xie and B.X. Jiang. Accurate and automatic tooth image segmentation model with deep convolutional neural

- networks and level set method. *Neurocomputing*, 419, 108–125. 2021. <https://doi.org/10.1016/J.NEUCO.M.2020.07.110>
- [8] M.K. Alam, T. Haque, F. Akhter, H.N. Albagieh, A. Bin Nabhan, M.A. Alsenani, S. Islam. Teeth segmentation by optical radiographic images using VGG-16 deep learning convolution architecture with R-CNN network approach for biomedical sensing applications. *Optical and Quantum Electronics*, 55(9), 1–19. 2023. <https://doi.org/10.1007/S11082-023-05096-X/TABLES/4>.
- [9] G. Rubiu, M. Bologna, M. Cellina, M. Cè, D. Sala, R. Pagani, ... M. Ali. Teeth segmentation in panoramic dental x-ray using mask regional convolutional neural network. *Applied Sciences* 2023, Vol. 13, Page 7947, 13 (13), 7947. 2023. <https://doi.org/10.3390/APP13137947>.
- [10] M. Xu, Y. Wu, Z. Xu, P. Ding, H. Bai and X. Deng. Robust automated teeth identification from dental radiographs using deep learning. *Journal of Dentistry*, 136, 104607. 2023. <https://doi.org/10.1016/J.JDEN.T.2023.104607>.
- [11] Prados-Privado, M., J. García Villalón, A. Blázquez Torres, C.H. Martínez-Martínez and C. Ivorra. A convolutional neural network for automatic tooth numbering in panoramic images. *BioMed Research International*, 2021. <https://doi.org/10.1155/2021/3625386>.
- [12] A.E. Yüksel, S. Gültekin, E. Simsar, Ş.D. Özdemir, M. Gündoğar, S.B. Tokgöz and İ.E. Hamamcı. Dental enumeration and multiple treatment detection on panoramic X-rays using deep learning. *Scientific Reports*, 11(1), 1–10. 2021. <https://doi.org/10.1038/s41598-021-90386-1>
- [13] S.Y. Lin, and H.Y. Chang. Tooth numbering and condition recognition on dental panoramic radiograph images using CNNs. *IEEE Access*, 9, 166008–166026. 2021. <https://doi.org/10.1109/ACCE.SS.2021.3136026>.
- [14] S. Tian, N. Dai, B. Zhang, F. Yuan, Q. Yu and X. Cheng. Automatic classification and segmentation of teeth on 3D dental model using hierarchical deep learning networks. *IEEE Access*, 7, 84817–84828. 2019. <https://doi.org/10.1109/ACCESS.2019.2924262>.
- [15] A. Jana, H.M. Subhash and D. Metaxas. Automatic tooth segmentation from 3D dental model using deep learning: a quantitative analysis of what can be learnt from a single 3D dental model, 12567, 42–51. 2023. <https://doi.org/10.1117/12.2669716>.
- [16] H. Wang, J. Minnema, K.J. Batenburg, T. Forouzanfar, F.J. Hu and G. Wu. Multiclass CBCT image segmentation for orthodontics with deep learning. *Journal of Dental Research*, 100 (9), 943–949. 2021. <https://doi.org/10.1177/00220345211005338>.
- [17] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna. Rethinking the inception architecture for Computer Vision. 2016.
- [18] K. He, X. Zhang, S. Ren and J. Sun. Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 770–778. 2015. <https://doi.org/10.1109/CVPR.2016.90>.
- [19] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna. Rethinking the inception architecture for computer vision. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2818–2826. 2016. <https://doi.org/10.1109/CVPR.2016.308>.
- [20] G. Huang, Z. Liu, L. Van Der Maaten and K.Q. Weinberger. Densely connected convolutional networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2261–2269*. 2017. <https://doi.org/10.1109/CVPR.2017.243>
- [21] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L.C. Chen. MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 4510–4520. 2018. <https://doi.org/10.1109/CVPR.2018.00474>
- [22] T.Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan and S. Belongie. Feature pyramid networks for object detection. 2016. Retrieved from <https://arxiv.org/abs/1612.03144v2>
- [23] H. Zhao, J. Shi, X. Qi, X. Wang and J. Jia. Pyramid scene parsing network. *proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 6230–6239. 2017. <https://doi.org/10.1109/CVPR.2017.660>.
- [24] O. Ronneberger, P. Fischer and T. Brox. U-net: Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9351, 234–241. 2015. [https://doi.org/10.1007/978-3-319-24574-4\\_28/COVER](https://doi.org/10.1007/978-3-319-24574-4_28/COVER).
- [25] K. Simonyan, and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. Retrieved from, 2014. <https://arxiv.org/abs/1409.1556v6>
- [26] K. Becker, É. Da, S. Rocha and P.T. Endo. A comparative study of deep learning models for dental segmentation in panoramic radiograph. *Applied Sciences* 12 (6), 3103. 2022. <https://doi.org/10.3390/APP12063103>

