

DİŞ SEGMENTASYONUNDA SEGFORMER YÖNTEMİNİN MODEL PARAMETRELERİ ÜZERİNDEKİ ETKİSİNİN ARAŞTIRILMASI

Egemen BELGE¹, Seçkin KARASU^{1*}

¹Zonguldak Bülent Ecevit University, Faculty of Engineering, Department of Electrical and Electronics Engineering, 67100, Zonguldak, Türkiye

Özet: Son yıllarda bilgisayar destekli tedavilerde yapay zekâ temelli uygulamalar diş hekimleri için hastalık teşhisinde kullanımı giderek artmaktadır. Diş hastalığının belirlenmesi sürecinde doğru bir şekilde bölütleme işlemi oldukça önemlidir. Diş bölütlenmesinin manuel olarak yapılması diş hekimleri tarafından yapılan teşhis için geçen süreyi ve işlem yükünü önemli ölçüde arttırmaktadır. Bu aşamada, makine öğrenmesi ve yapay zekâ yöntemleriyle otomatik olarak diş bölgesinin bölütlenmesi araştırmacıların ilgi duyduğu güncel bir konu olmaktadır. Çalışmada 12 diş stajyeri tarafından 15318 poligonlu 598 hastadan alınan X-Ray diş görüntüleri kullanılmaktadır. Kullanılan veri seti eğitim, doğrulama ve test olarak %70, %15, %15 olarak bölünmüştür. Bu veri seti otomatik olarak diş bölütlemeyi amaçlayan derin öğrenme ağının eğitim sürecinde kullanılmaktadır. SegFormer eğitim bloğu hiper parametrelerinin değişimine bağlı oluşturulan mimarilerinin performansları incelenmektedir. Burada MiT B0-B5 mimarilerine göre oluşturulan modellerin Dice benzerlik katsayılarına göre test verisi için performansları sırasıyla %92,61, %92,82, %93,25, %93,13, %93,17 ve %93,09 olarak elde edilmektedir. Elde edilen test sonuçlarına göre geliştirilen yapay zekâ tabanlı SegFormer ağı diş bölütlemeyi yüksek doğrulukla gerçekleştirmektedir. Geliştirilen derin öğrenme ağı özellikle diş hastalıklarının teşhisinde girdi olarak verimli bir şekilde kullanılabilir. Yüksek Dice benzerlik katsayıları, çalışmada sunulan SegFormer ağının diş bölgesini doğru bir şekilde tespit edebildiğini ifade etmektedir.

Anahtar kelimeler: Yapay zekâ, Diş bölütlenmesi, SegFormer, X-Ray görüntüleme, Dice benzerlik katsayısı


Investigation of the Effect on Model Parameters in SegFormer Method for Tooth Segmentation


Abstract: In recent years, AI-based applications in computer-aided treatments have been increasingly used by dentists for disease diagnosis. Accurate segmentation is crucial in the process of identifying dental diseases. Manual segmentation of teeth significantly increases the time and workload for dentists when making diagnoses. At this stage, the automatic segmentation of the dental region using machine learning and artificial intelligence methods has become a topic of great interest for researchers. In the study, X-Ray tooth ratios taken from 598 patients with 15318 polygons by 12 dental interns are used. The dataset is split into 70% for training, 15% for validation, and 15% for testing. This dataset is used in the training process of a deep learning network for automatic tooth segmentation. The performances of architectures generated based on changes in the hyperparameters of the SegFormer training block are examined. Here, according to the Dice similarity coefficients of the models according to the Mix B0-B5 architectures, the performances for the test data are obtained as 92.61%, 92.82%, 93.25%, 93.13%, 93.17% and 93.09%, respectively. According to the test results obtained, the developed artificial intelligence-based SegFormer network performs tooth segmentation with high accuracy. The developed deep learning network can be efficiently used especially in the diagnosis of dental diseases. High Dice similarity coefficients indicate that the SegFormer network presented in this study can accurately detect the tooth region.

Keywords: Artificial intelligence, Tooth segmentation, SegFormer, X-Ray imaging, Dice similarity coefficient

*Sorumlu yazar (Corresponding author): Zonguldak Bülent Ecevit University, Faculty of Engineering, Department of Electrical and Electronics Engineering, 67100, Zonguldak, Türkiye

E mail: seckin.karasu@beun.edu.tr (S. KARASU)

Egemen BELGE  <https://orcid.org/0000-0001-5852-1085>

Seçkin KARASU  <https://orcid.org/0000-0001-5277-5252>

Gönderi: 18 Ekim 2024

Kabul: 28 Kasım 2024

Yayınlanma: 15 Ocak 2025

Received: October 18, 2024

Accepted: November 28, 2024

Published: January 15, 2025

Cite as: Belge E, Karasu S. 2025. Investigation of the effect on model parameters in SegFormer method for tooth segmentation. BSEng Sci, 8(1): 132-141.

1. Giriş

Yapay zekanın ortaya çıkışıyla birlikte son on yıl içerisinde önemli teknolojik gelişmeler meydana gelmektedir. Her sektörde olduğu gibi yapay zekanın diş bilimi alanında da kullanımı artmaktadır. Yapay zekâ, diş ve tıp biliminde hastaların tıbbi verilerinin toplanmasında, makine öğrenmesi kullanılarak diş çürüklerinin tahmininde ve hastaların verilerini işleyerek hastalık tanısının belirlenmesi gibi farklı uygulama alanlarında tercih edilmektedir (Lee vd., 2022). Günümüzde yapay zekâ yöntemleri kullanılarak diş

segmentasyonu gerçekleştirilmektedir. Diş segmentasyonu bilgisayar destekli işlemlerde ve klinik tanılamada önemli bir rol oynamaktadır. Diş bölgesinin ve diğer dokuların arasındaki özelliklerin ayrımı segmentasyon tekniğiyle mümkün olmaktadır. Uygun tanının belirlenmesi ve gerekli cerrahi işlemlerin anında yapılabilmesi aşamalarında dişin doğru bir şekilde segmentasyonu gereklidir. Ortodonti alanında diş hekimleri, hastaların diş hareketi ve diş kökü emilimi hakkında bilgilere ihtiyaç duymaktadır. Elde edilen bu bilgiler kullanılarak ortodonti tedavi aşamaları önemli



ölçüde kısaltılmaktadır. Dişlerin panoramik X-ışını görüntülerinden segmentasyon işlemi gerçekleştirilmektedir. Bunun yanı sıra, diş segmentasyonunun doğru yapılması, hastaların diş yaşının tahmin edilmesi, kişinin adli olarak biyometrik özelliklerinin tanımlanması ve gizli diş yapılarının ortaya çıkarılmasında kullanılabilir (Zhao vd., 2020). Tüm bu sebepler birlikte değerlendirildiğinde diş segmentasyonu hem araştırma hem de klinik uygulamalarda kritik rol oynamaktadır.

Bununla birlikte diş segmentasyonu diş hekimleri tarafından manuel ya da yarı otomatik etkileşimli yöntemlerle gerçekleştirilmektedir (Kang vd., 2015; Wang vd., 2016; Li ve Wang, 2016). Bu durum diş hekimlerinin bilgisine doğrudan bağlı olup, oldukça zaman alıcı bir işlemdir. Düşük kaliteli diş görüntülerinden dolayı diş hekimleri diş bölgesini belirleme işleminde zorlanmaktadır. Görüntü kalitesinin düşük olması, diş bölgesinin belirleme işlemini olumsuz etkilemektedir. Tüm sebeplerden dolayı diş segmentasyonu için daha iyi bir görüntüleme ve analiz yöntemine ihtiyaç duyulmaktadır.

Son yıllarda diş segmentasyonunun zorluklarına karşı bilim insanları yenilikçi çözümler üretmektedir. Geliştirilen segmentasyon metotları geleneksel ve derin öğrenme tabanlı olarak iki ana alanda açıklanmaktadır (Zhao vd., 2020). Geleneksel teknikler çoğu zaman görüntünün gradyanları ve bölge tabanlı benzerlik ölçümü gibi ön bilgilere ve görüntü özelliklerine dayanmaktadır (Gao vd., 2013; Indraswari vd., 2015; Gao vd., 2018). Bazı panoramik X-ışını görüntülerinde dişlerin segmentasyonu işleminde lokal adaptif eşik yöntemi kullanılmaktadır (Indraswari vd., 2015). Bu yöntem görüntünün istatistiksel özelliklerini çıkarmaktadır. Başka bir çalışmada ise Otsu eşiği ve Mahalanobis mesafesinin birlikte değerlendirildiği şablon eşleştirme yöntemi kullanılmaktadır (Poonsri vd., 2016). Eşikleme stratejisine dayanan bu modelin diş köklerini tanımlaması oldukça zordur. Piksel yoğunluk seviyesindeki süresizliğe göre görüntüyü bölgelere ayırarak segmentasyon işlemi gerçekleştirilmektedir (Modi ve Desai, 2011). Matematiksel morfolojiyi kullanan Watershed dönüşümü gri tonda verilen görüntünün segmentasyonunda kullanılmaktadır (Li vd., 2012). Diş görüntüsünün segmentasyonunda morfolojik metot kullanılmaktadır. Çalışmada metodun yüksek doğruluklu ve düşük zaman karmaşıklığı vurgulanmaktadır (Said vd., 2006). Aktif kontur metodu ve diş segmentasyonu için şekil analizi birleştirilmektedir (Lira vd., 2009). Şekil modelleri dişin algılanmasında kullanılmaktadır (Lira vd., 2009). Geleneksel görüntü segmentasyonu teknikleri, önceden belirlenmiş bilgiler ve görüntü özelliklerinin eşleştirilmesine doğrudan bağlıdır (Zhao vd., 2020). Özel uzmanlık bilgisi gerektiren geleneksel segmentasyon tekniğinin uygulama alanını önemli ölçüde kısıtlamaktadır. Derin öğrenme tabanlı diş segmentasyonu uygulamaları son yıllarda önemli ölçüde artmaktadır. Son yıllarda küresel ve yerel özelliklerin kodlanmasının özellik gösterimi ve görsel özellik

öğrenimini güçlendirdiği vurgulanmaktadır (Wu vd., 2019; Zhu ve Yang, 2020). Görüntü içerisindeki arka planı çıkarma ve kendi kendine öğrenme alanında çalışmalar yer almaktadır (Yong vd., 2017; Meng vd., 2017). Panoramik X ışını görüntüsünden diş segmentasyonu probleminin çözümü için birkaç yaklaşım geliştirilmektedir (Yang vd., 2018; Chen vd., 2019; Wirtz vd., 2018). Yang vd. (2018) 196 vakanın yer aldığı veri setinden diş görüntü analizi sunulmaktadır. Çalışmada, diş hastalıklarının teşhisi ve sınıflandırılması için Convolutional Neural Networks (CNN) kullanılmaktadır. Chen vd. (2019) diş alanının segmentasyonu için Mask R-CNN yapısında (He vd., 2017) derin öğrenme ağı geliştirmektedir. Geliştirdikleri bu yöntem ile örnek segmentasyon yöntemlerinin diş segmentasyonunda kullanılabileceğini gösterilmektedir. Yağılan bir çalışmada yeni bir ölçüm yaklaşımının önerildiği ve diş segmentasyonu için önerilen modelin performansının değerlendirilebilmesi için kompakt bir fonksiyon önerildiği belirtilmiştir (Kato ve Hotta, 2024). Dice kaybı yeniden ele alınarak kosinüs benzerliği yerine t-vMF benzerliğiyle daha kompakt bir kayıp fonksiyonu olan t-vMF Dice kaybı önerilmiştir. Ayrıca, sınıf zorluğuna göre adaptif eğitim sağlayan ve parametre η 'yı otomatik belirleyen "Adaptive t-vMF Dice loss" algoritması geliştirilmiştir. CVC-ClinicDB, Kvasir-SEG, Automated Cardiac Diagnosis Challenge ve Synapse veri setleri üzerinde yapılan deneylerde, önerilen yöntemle Dice skorunun orijinal Dice kaybına göre iyileştirildiği gösterilmiştir. Alveolar yarıklı hastalarda diş şekil varyasyonu, interdental benzerlik ve bitişik dişlerin örtüşmesi gibi zorlukları aşmak için yeni bir "diş tanımlayıcı" (tooth descriptor) göstergesi önerilmiştir (Gong vd., 2024). İki aşamalı bir ağ kullanarak önce diş tanımlayıcılarını ve merkez noktalarını tahmin eder, ardından bu bilgilerle diş küpleri oluşturmakta ve segmentasyon ağı bu küplerden tam diş yapısını çıkartmaktadır. Yapılan deneylerde, önerilen yöntem %94,4 Dice doğruluğu ile mevcut diş segmentasyon yöntemlerinden daha üstün performans gösterilmiştir. Panoramik radyografilerde dişlerin otomatik segmentasyonu için derin öğrenmeye dayalı çok ölçekli mekansal havuzlama (ms-SP) tekniği test edilmiştir (Nagaraju ve Sudha, 2024). Dişler, bir oral radyolog tarafından etiketlenmiş ve aşırı öğrenmeyi önlemek için veri artırma stratejisi kullanılmıştır. Sonuçlar, ortalama IoU %87, F1 skoru %98,9 ve doğruluk %98,5 olarak elde edilmiştir. Görsel analiz, elde edilen sonuçların gerçek verilere benzer olduğunu göstermiştir. Önerilen teknik panoramik diş fotoğraflarında segmentasyonu otomatikleştirmenin yanı sıra adli kimlik tespiti ve tanı otomasyonu gibi alanlarda da faydalı olabileceği bildirilmiştir (Nagaraju ve Sudha, 2024). Panoramik radyografelerde süt, karma ve kalıcı dişlerin otomatik olarak sınıflara ayrılması ve numaralandırılması için derin öğrenme tabanlı bir yöntem geliştirilerek model doğrulaması yapılmıştır (Xu vd., 2023). Çalışmada kullanılan derin öğrenme ağı, U-Net tabanlı ilgilenilen bölgenin çıkarılması ve Hibrit Görev Kaskatı tabanlı bir

diş segmentasyonu ve numaralandırma modelinden oluşmaktadır. Derin öğrenme ağı, 4232 diş görüntü üzerinde eğitilmiş olup, 605 diş görüntüsü doğrulamada ve 1209 görüntü ise test aşamasında kullanılmıştır. Panoramik radyografi bölütlemesini gerçekleştirmek için atlama bağlantılarının yer aldığı transformatör tabanlı Ushaped kodlayıcı-kod çözücü mimarisi olan SWinUnet tanıtılmaktadır (Sheng vd., 2023). SWin-Unet'in diş segmentasyon performansını analiz edebilmek için, çalışmada PLAGH-BH veri seti tanıtılmıştır. Bu veri seti üzerinde çalışmada önerilen derin öğrenme ağının model doğruluğu %88,52 olarak elde edilmektedir. Önerilen diş bölütleme yönteminin doğruluğu U-Net, Link-Net ve FPN ağlarıyla kıyaslanmaktadır. Performans sonuçlarına göre maksimum doğruluğun önerilen SWinUnet ağında elde edildiği vurgulanmaktadır. Diş tanıma modülü tasarlamak için R-CNN ağı kullanılmakta ve VGG-16 konvolüsyon mimarisine dayanan diş numaralandırma sistemi çalışmada geliştirilmektedir (Alam vd., 2023). Çalışmada, panoramik diş görüntüleri içeren veri seti eğitim ve doğrulama olarak bölümlere ayrılarak önerilen derin öğrenme ağı eğitilmiştir. Bu çalışmada önerilen ağın model doğruluğu %89,8 olarak elde edilmiştir. Üç boyutlu olarak diş bölütlenme ve sınıflandırma amacıyla evrimsel sinir ağlarına dayanan yeni bir yaklaşım önerilmiştir (Tian vd., 2019). Burada geliştirilen evrimsel sinir ağının diş bölütleme doğruluğu %89,81 olarak elde edilmiştir.

Bu çalışmada diş segmentasyonu için derin öğrenme tabanlı SegFormer yöntemi sunulmaktadır. SegFormer yönteminin içerisindeki hiyerarşik transformatör kodlayıcısının model büyüklüğü değiştirilerek diş segmentasyonunun başarımı incelenmektedir. Burada, 12 diş stajyeri tarafından toplanan 15318 poligonlu 598 X-ray diş görüntüleri üzerinde derin öğrenme ağı eğitilmiştir. Çalışmada kullanılan veri seti %70 eğitim, %15 doğrulama ve %15 test olarak ayrılmaktadır. SegFormer segmentasyon tekniğinin içerisinde yer alan MiT-BO ile MiT-B5 aralığındaki kodlayıcı parametrelerinin model büyüklükleri değiştirilerek derin öğrenme ağının performansı incelenmektedir.

Bu çalışmanın geri kalan kısmı şu şekilde organize edilmektedir. Çalışmada kullanılan veri seti, performans değerlendirme kriteri ve SegFormer segmentasyon modeli Bölüm 2'de sunulmaktadır. Çalışmada kullanılan SegFormer derin öğrenme ağının performansı Bölüm 3'de açıklanmaktadır. Derin öğrenme ağının performans sonuçları Bölüm 4'de tartışılmaktadır.

2. Materyal ve Yöntem

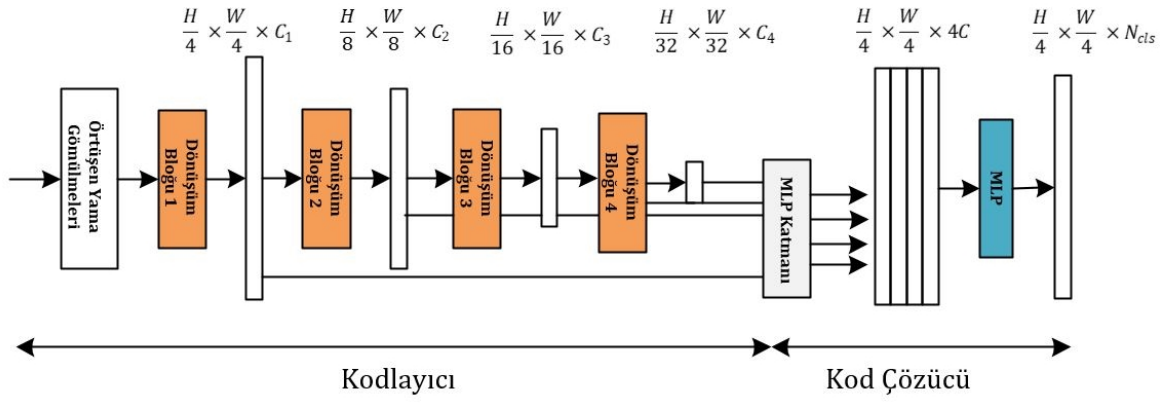
2.1. Veri Seti

Segmentasyon işlemi, Kongo Demokratik Cumhuriyeti'ndeki 12 Humans in Loop stajyeri tarafından, Lopez ve arkadaşları tarafından yayınlanan Panoramik radyografi veri tabanını kullanarak eğitimlerinin bir parçası olarak manuel olarak gerçekleştirilmektedir (Human In The Loop, 2023). Çalışmada kullanılan veri seti, her dişin farklı bir sınıfa göre bölütlendiği toplam 15318 poligona sahip 598

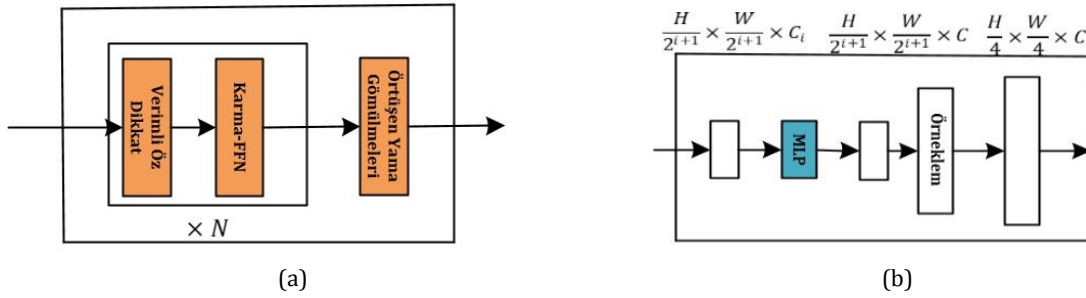
görüntüden oluşmaktadır. Veri setinin içerisinde 32 farklı diş sınıfı yer almaktadır. Dişlerin numaralandırılması üst sol dişlerden başlar ve saat yönünde alt dişlere doğru devam etmektedir. Her dişin pozisyonuna göre benzersiz bir numarası bulunmaktadır.

2.2 SegFormer Segmentasyon Modeli

SegFormer görüntü segmentasyonu amacıyla tasarlanmış bir derin öğrenme modelidir. Anlamsal segmentasyon işleminde yüksek performans elde etmek için hem dönüşüm bloklarının (Şekil1) hem de evrimsel sinir ağlarının (CNN) güçlü yanlarını birleştirmektedir (Xie vd., 2021). SegFormer derin öğrenme modeli iki farklı modülün bir arada kullanıldığı bir yapıdan oluşmaktadır. Burada kullanılan dönüşüm kodlayıcı blokları yüksek çözünürlüklü özelliklerin çıkarılmasında kullanılmaktadır. Bunun yanı sıra kod çözücü kısmında kullanılan derin öğrenme ağları, anlamsal segmentasyon maskesini üretmek için çok seviyeli özellikleri birleştirmektedir. SegFormer modelinin genel şeması Şekil 2'de sunulmaktadır. Burada kodlayıcı ve kod çözücü katmanlarının sıralı yapısı yer almaktadır. Kodlayıcı bloğu 4 farklı birimden oluşmaktadır. Kod çözücü bloğunda ise, iki farklı derin öğrenme ağı bulunmaktadır. SegFormer modelinde verilen bir $H \times W \times 3$ boyutundaki görüntü ilk olarak 4×4 boyutundaki parçalara bölünmektedir. Burada H görüntünün yüksekliği, W görüntünün genişliğini ifade etmektedir. İkinci olarak orijinal görüntünün $\{\frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}\}$ çok seviyeli çözünürlüklerini çıkarmak için hiyerarşik dönüşüm bloğu kullanılmaktadır. $\frac{H}{4} \times \frac{W}{4} \times N_{cls}$ çözünürlüğün de segmentasyon maskesini tahmin etmek için MPL kod çözücüsü kullanılmaktadır. Burada N_{cls} kategori sayısını ifade etmektedir. C_i ise özellik haritasını elde etmek için kullanılan vektördür. Hiyerarşik dönüşüm kodlayıcılarının sadece boyutları değiştirilerek MiT-BO ile MiT-B5 arasındaki yapılar segmentasyon işlemlerinde kullanılmaktadır. Tablo 1'de bu model yapılarının parametre büyüklükleri gösterilmektedir. Şekil 2(a)' da görüldüğü gibi SegFormer dönüşüm bloğu; verimli öz dikkat, karma ileri beslemeli sinir ağı (FFN) ve örtüşen yama birleştirme olmak üzere üç farklı yapıdan oluşmaktadır. Verimli öz dikkat bloğunda, giriş özelliklerindeki dikkat puanları hesaplanarak, modelin pikseller arasındaki ilişkileri yakalaması sağlanmaktadır. Karma FFN bloğu, doğrusal olmayan bir aktivasyon fonksiyonuna sahip iki doğrusal katmandan oluşmaktadır. Örtüşen yama birleştirmeleri, görüntü üzerindeki üst üste binen bölgelerin özelliklerini birleştirmek için kullanılan bir katmandır. Bu katman, modelin karmaşık bölgeleri anlama yeteneğini geliştirmektedir. Şekil 2(b)' de MLP katmanının iç yapısı gösterilmektedir. Burada kanal boyutunu birleştirebilmek amacıyla kodlayıcıdan gelen çok seviyeli özellikler MLP katmanından geçmektedir. Daha sonrasında bu özellikler $\frac{1}{4}$ kadar yukarı örneklenmekte ve birleştirilmektedir.



Şekil 1. SegFormer modelinin blok şema gösterimi.



Şekil 2. Dönüşüm (a) ve MLP katmanının (b) iç yapısı.

Çalışmada SegFormer modelinin içerisinde yer alan AdamW optimizasyonu ile eğitilmiş olan hiper parametreler kullanılmaktadır. Bu hiper parametreler, DE20K ve Cityscapes üzerinde 160K ve COCO-Stuff üzerinde ise 80K iterasyon boyunca eğitilerek belirlenmektedir (Xie vd., 2021).

Tablo 1. SegFormer modelinin Enkoder ve Kod çözücü parametreleri

Enkoder Model Boyutu	Enkoder Parametresi	Dekoder Parametresi
MiT-B0	3,4	0,4
MiT-B1	13,1	0,6
MiT-B2	24,2	3,3
MiT-B3	44	3,3
MiT-B4	60,8	3,3
MiT-B5	81,4	3,3

Bu çalışmada SegFormer modeli içerisinde yer alan MiT-B0 ile MiT-B5 yapılarının dış bölütleme başarımına olan etkisi incelenmektedir. Aynı zamanda model büyüklüğündeki değişimine bağlı performans kıyaslaması da çalışmada yer almaktadır.

2.3 Dice Değerlendirme Kriteri

Dice, görüntü segmentasyonunun model performansını değerlendirmek amacıyla geliştirilen bir katsayıdır. Bu katsayı eşitlik 1'de verildiği şekilde tanımlanmaktadır (Yan, 2018).

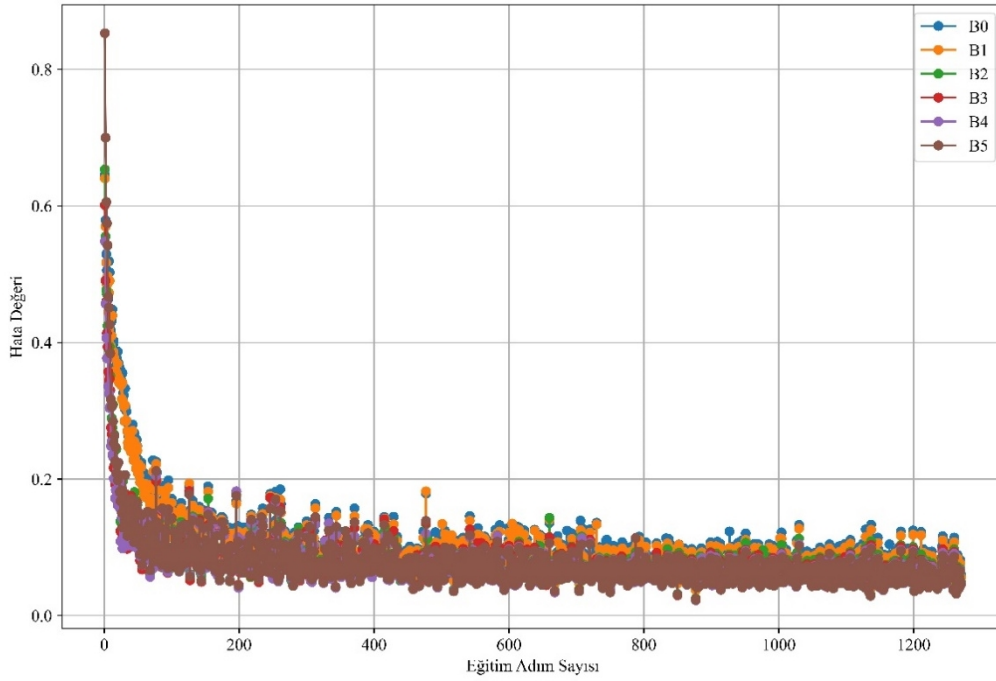
$$Dice = \frac{2|A \cap B|}{|A| + |B|} \quad (1)$$

Burada A ve B sırasıyla tahmin edilen dış segmentasyonu bölgesini, gerçek dış bölgesini ifade etmektedir. Dice

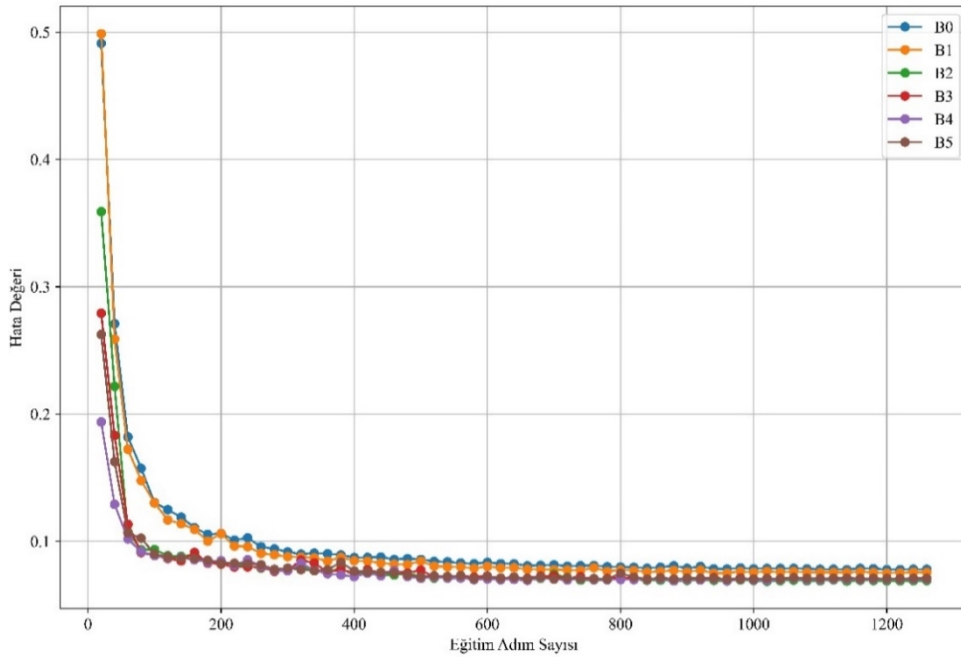
katsayısının maksimum değeri 1'e eşittir. Bu durum segmentasyon işleminin doğruluk oranının en yüksek olduğu belirtmektedir. Segmentasyon başarımının en düşük olduğu durumda ise Dice değeri 0 değerini almaktadır. Çalışmada kullanılan Dice benzerlik katsayısı SegFormer ağının eğitim aşamasında dış bölütleme başarımını ölçen bir değerlendirme kriteridir. Modelin eğitimi sırasında her bir iterasyonda Dice benzerlik katsayısı hesaplanarak dış segmentasyon başarımı elde edilmekte ve bu değerlendirme kriterine göre SegFormer ağı her iterasyonda güncellenmektedir.

3. Bulgular

Şekil 3'de SegFormer modelinin B0-B5 versiyonları için eğitim adım sayısına karşılık hata değerinin değişimi gösterilmektedir. Yatay eksen, eğitim adım sayısını, dikey eksen ise hata değerini temsil etmektedir. Grafikte her bir model (B0-B5) farklı renkte gösterilmiştir. Grafikte ilk aşamada hata değerinin hızla düştüğü, ardından belirli bir adım sayısından sonra stabil hale geldiği gözlemlenmiştir. Eğitim aşamasının ilk adımlarında özellikle 100-200 eğitim adımından sonra hata değerlerinde yaklaşık %80 üzerinde bir azalma yaşanmaktadır. Bu noktadan sonra hatalar kısmen düşük seviyelerde, sabit bir şekilde ilerlemektedir. Farklı modeller arasında (B0'dan B5'e) belirgin bir performans farkı görünmemekle birlikte, bazı modellerin daha hızlı yakınsadığı, ancak genel hata seviyelerinin birbirine yakın olduğu söylenebilir. Bu durum, SegFormer modelinin tüm versiyonlarının benzer eğitim performansı gösterdiğini ve eğitim sürecinde hataların hızla azaldığını, ancak belirli bir noktadan sonra daha az iyileşme olduğunu ortaya koymaktadır.



Şekil 3. SegFormer B0-B5 için eğitim adım sayısına karşılık hata değerinin değişimi.



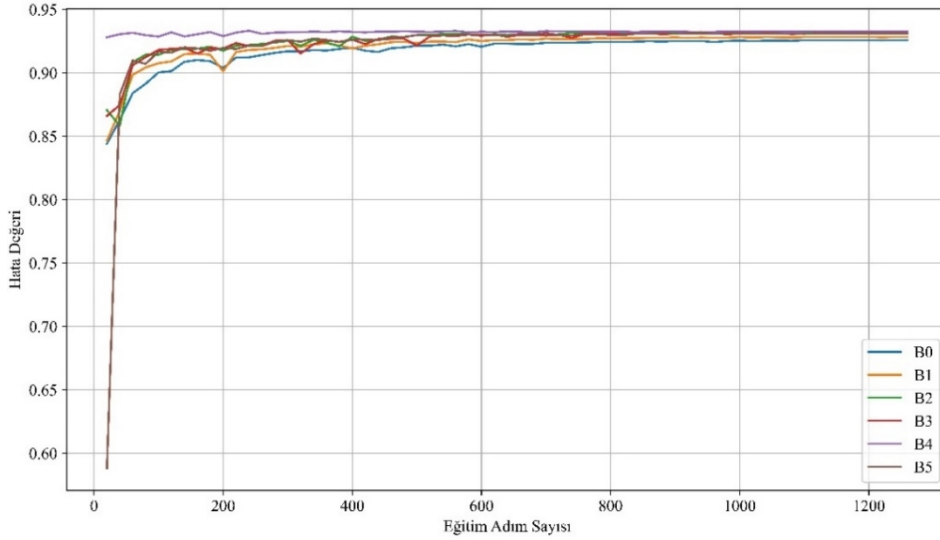
Şekil 4. SegFormer B0-B5 için doğrulama adım sayısına karşılık hata değerinin değişimi.

Şekil 4'te, Grafikte, B0, B1, B2, B3, B4 ve B5 olmak üzere altı farklı model mimarisinin eğitim adımları boyunca doğrulama hatası (validation loss) değerlerindeki değişim gözlemlenmektedir. Grafikte, x eksenini eğitim adımlarını, y eksenini ise hata değerini (loss) temsil etmektedir. Grafik incelendiğinde, tüm modellerin başlangıç aşamasında özellikle 0-100 adımları arasında hızlı bir hata azalması yaşandığı görülmektedir. Bu, modelin veriden temel özellikleri öğrenmeye başladığı ve eğitim sürecine hızla adapte olduğu anlamına geldiği belirtilebilir. Bu aşamada hata değeri yaklaşık olarak 0,5 seviyesinden 0,1 seviyesine kadar hızlı bir düşüş göstermektedir. Eğitim süreci ilerledikçe, özellikle 100.

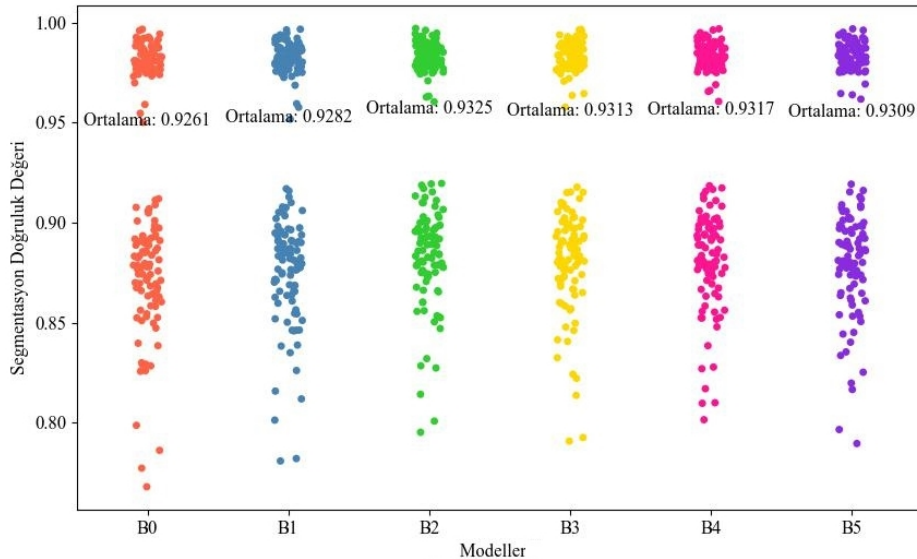
adımdan sonra hata azalmasında bir yavaşlama meydana gelmektedir. Tüm modeller bu noktada daha detaylı örüntüleri öğrenmeye başlamakta ve bu da daha yavaş bir hata azalması ile sonuçlanmaktadır. Yaklaşık 400. adıma kadar, tüm modellerin hata değeri 0,05 ile 0,1 aralığına düşmekte ve bu noktada modellerin öğrenme sürecinde bir dengeye ulaşmaya başladıkları gözlemlenmektedir. 400. adımdan itibaren, modellerin hata değerlerinde önemli bir değişiklik olmadığı, hatanın belirli bir seviyeye oturduğu ve artık kayda değer bir öğrenmenin gerçekleşmediği görülmektedir. Bu durum, modellerin veriden öğrenebileceği özellikleri büyük ölçüde öğrendiğini ve daha fazla eğitimle yalnızca

minimal iyileşmeler elde edilebileceğini göstermektedir. B0 modeli, başlangıç aşamasında (ilk 100 adım) diğer modellere göre daha hızlı bir düşüş sergilemekte ve erken aşamada daha düşük hata değerlerine Şekil 3'teki grafik incelendiğinde ulaşmaktadır. Bu, B0 modelinin başlangıçta daha hızlı öğrenme gösterdiğini düşündürmektedir. Ancak eğitim sürecinin ilerleyen aşamalarında (100 adımdan sonra) tüm modellerin performanslarının birbirine yaklaştığı görülmektedir. B1, B2, B3, B4 ve B5 modelleri arasında hata değerleri açısından çok az fark vardır ve genellikle benzer bir hata seviyesinde stabil hale gelmektedirler. Son aşamada B5 modeli, diğer modellere kıyasla biraz daha düşük hata değeriyle (yaklaşık 0,05) en iyi performansı sergilediği söylenebilir. Ancak bu fark oldukça azdır. Genel olarak tüm modeller başlangıçta hızla öğrenim hatalarını azaltmakta ve yaklaşık 400. adımdan sonra stabilize olmaktadır. Farklı mimariler (B0-B5) arasında erken aşamalarda bazı farklılıklar görülse de özellikle ilerleyen adımlarda bu farklar oldukça azalmaktadır. B5 modeli, diğerlerine göre biraz daha iyi sonuç verse de

tüm modellerin genel performanslarının birbirine oldukça yakın olduğu söylenebilir. Şekil 5'te, çeşitli modellerin eğitim süreçlerindeki segmentasyon başarımları (y-ekseni) ve eğitim adım sayıları (x-ekseni) karşılaştırılmaktadır. Eğitim adımları arttıkça tüm modellerin segmentasyon başarımları önemli ölçüde artmakta ve belirli bir noktadan sonra sabitlenmektedir. İlk 100 adımda, başarımlar hızla artarak yaklaşık %90 seviyesine ulaşmakta, ardından artış daha yavaş bir tempoda devam etmektedir. B0, B1, B2, B3 ve B5 modelleri benzer bir başarı eğrisi izlerken, B4 modeli diğerlerinden ayrılmakta ve daha dalgalı bir performans sergilemektedir. Bu segmentasyon modelinin diğerlerine göre ayrışmasının sebebi veri çeşitliliğinin eksikliğinden kaynaklandığı düşünülmektedir. Bütün modeller, yaklaşık 300 adım sonrasında segmentasyon başarımlarını büyük ölçüde sabitlemekte ve %90 ile %95 bandında stabil bir başarı oranına ulaşmaktadır. Özetle, çoğu model belirli bir eğitim sürecinden sonra yüksek segmentasyon başarımına ulaşırken, B4 modelinin performansı diğerlerine kıyasla daha düşük kalmaktadır.



Şekil 5. SegFormer B0-B5 için eğitim verisine karşılık segmentasyon doğruluk değeri değişimi.



Şekil 6. SegFormer B0-B5 için test verisine karşılık segmentasyon doğruluk değeri değişimi.

Şekil 6'da modellerin test verisi için genel performanslarının karşılaştırılması verilmiştir. B0 modelinin ortalama doğruluk değeri %92.61 olarak belirlenmiştir. Bu, diğer modellere kıyasla en düşük doğruluk oranını temsil etmektedir. B0 modelinin doğruluk dağılımında, doğruluk değerlerinin genel olarak daha düşük bir aralıkta seyrettiği ve birkaç düşük performanslı noktanın olduğu gözlemlenmektedir. B1 modeli, ortalama doğruluk açısından B0'dan daha iyi bir performans sergileyerek %92,82 ortalama doğruluk değeri elde etmiştir. B1 modelindeki dağılım, daha dar bir aralık göstermekte ve bu durum daha tutarlı bir performans işaret etmektedir. B2 modeli ise ortalama %93,25 doğruluk değeriyle tüm modeller arasında en yüksek ortalama doğruluğa sahiptir. Doğruluk dağılımı da oldukça yoğun ve sıkı bir şekilde yüksek değerlerde toplanmıştır; bu da modelin yüksek bir performansa sahip olduğunu göstermektedir. B3, B4 ve B5 modellerinin ortalama doğruluk değerleri sırasıyla %93,13, %93,17 ve %93,09'dur. Bu üç modelin doğruluk sonuçları birbirine oldukça yakındır ve aralarındaki farklar çok azdır. Grafik üzerinde, B2 ile B5 modelleri arasındaki performans farkı oldukça azdır ve bu modeller genel olarak yüksek doğruluk oranlarına sahiptir.

B0 ve B1 modellerinin doğruluk dağılımları, diğer modellere göre daha geniş bir aralığa yayılmıştır. Özellikle B0 modeli, doğruluk açısından bazı uç (outlier) noktalara sahiptir; yani bazı denemelerde çok düşük doğruluk değerleri elde edilmiştir. Bu durum, B0 modelinin bazı veri örneklerinde zayıf performans gösterdiğini işaret etmektedir. Diğer taraftan, B2, B3, B4 ve B5 modellerinin doğruluk dağılımı daha yoğun ve belirli bir aralığa sıkışmış durumdadır. Bu ise bu modellerin daha tutarlı bir performans gösterdiğini ve doğruluk değerlerinin daha homojen dağıldığını göstermektedir. Özellikle B2 modeli hem yüksek ortalama doğruluk değeri hem de dar bir dağılım ile diğer modellerden daha başarılı görünmektedir.

Farklı modeller arasındaki doğruluk değerlerinde büyük sapmalar olmamakla birlikte, küçük farklar gözlemlenmektedir. B2 modeli, en yüksek ortalama doğrulukla diğer modellerden biraz daha önde görünmektedir. Ancak B3, B4 ve B5 modellerinin ortalamaları da B2 modeline çok yakındır. Bu modeller, neredeyse aynı doğruluk seviyelerinde performans sergilemektedir. B0 modeli ise doğruluk açısından en düşük performansı göstermekte olup, dağılımın geniş olması, bu modelin daha değişken bir performansa sahip olduğunu düşündürmektedir.

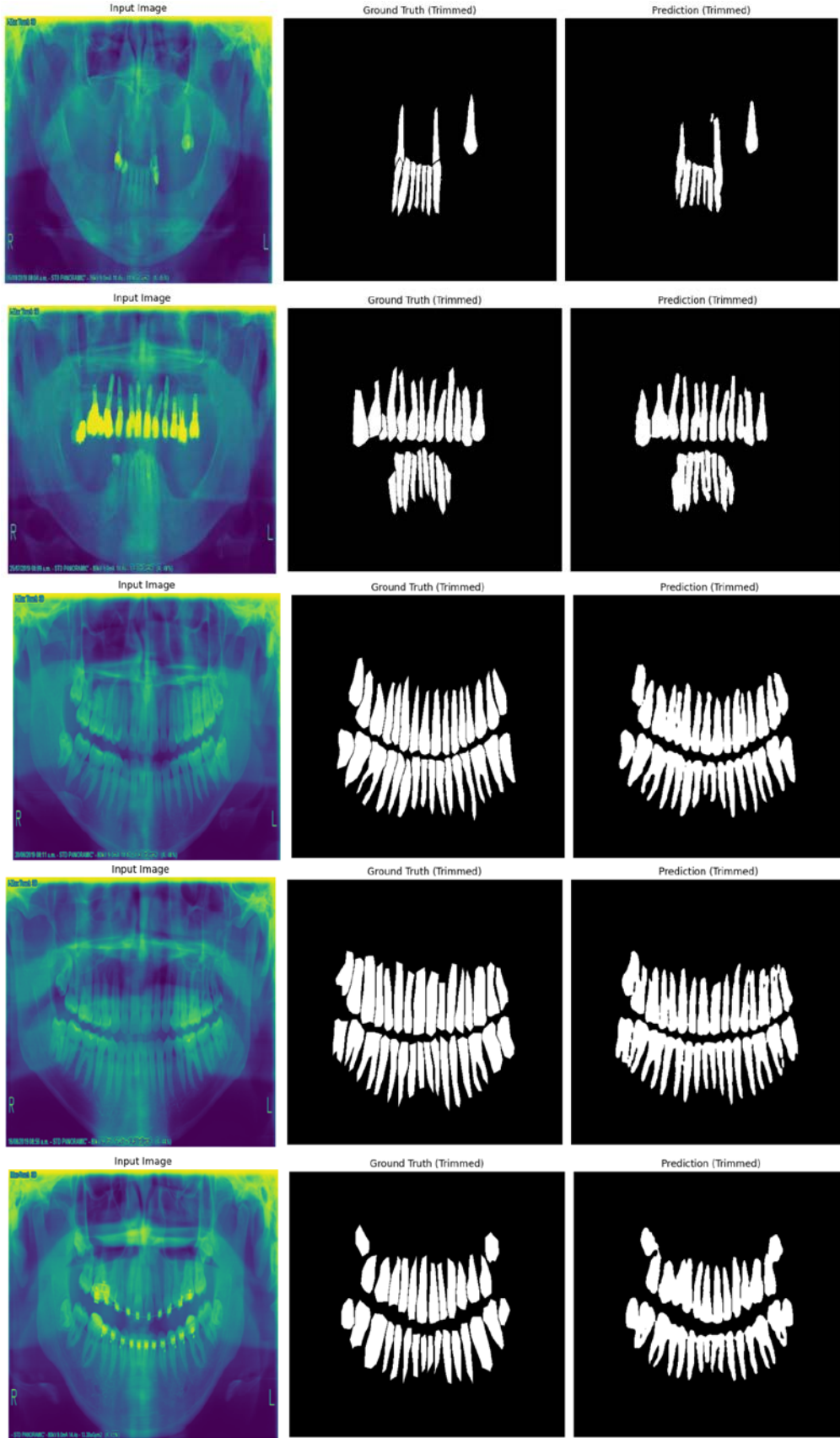
Şekil 7'de, beş farklı durum için X-ray giriş görüntüsü, referans (ground truth) görüntü ve model segmentasyon görüntüsü içeren üç farklı görüntü verilmiştir. Burada X-ray giriş görüntüsü modelin segmentasyon işlemi için temel aldığı ham veriyi temsil etmektedir. İkinci görüntü, ground truth yani referans görüntü olup, X-ray verisinde beklenen doğru segmentasyonu veya bölütleme göstermektedir. Bu, modelin ne kadar doğru sonuçlar ürettiğini değerlendirmek için kullanılmaktadır. Son olarak, model segmentasyon görüntüsü, modelin giriş

verisi üzerinden gerçekleştirdiği tahmini sonuçları göstermektedir. Bu üç görüntü yan yana incelendiğinde, modelin segmentasyon doğruluğu referans görüntü ile karşılaştırılarak başarı seviyesi değerlendirilebilir. Bu tür bir görsel analiz, modelin hatalarını ve iyileştirme alanlarını tespit etmek açısından önemli olmaktadır.

4. Tartışma ve Sonuç

Çalışmada SegFormer modelinin farklı versiyonlarının (B0-B5) eğitim sürecinde benzer performans sergilediği, başlangıçta hızlı bir öğrenme ve hata azaltma sağladıkları, ancak belirli bir adım sayısından sonra bu iyileşmenin yavaşladığı ve hataların stabilize olduğu görülmektedir. B2 modeli, doğruluk açısından diğer modellerden hafif bir üstünlük gösterdiği bulunmuştur. Modellerin segmentasyon başarımları ve doğruluk sonuçları, eğitim adımlarının artmasıyla büyük ölçüde sabitlenmiş ve %90 ile %95 arasında yüksek bir performansa ulaşmıştır. Ancak, B0 modeli diğer modellere kıyasla daha düşük doğruluk ve daha geniş bir dağılım sergileyerek daha değişken bir performans göstermiştir. Özellikle B2 modeli hem doğruluk hem de tutarlılık açısından en başarılı model olarak öne çıkmıştır. Sonuçlar, SegFormer modelinin çeşitli versiyonlarının dış segmentasyonu gibi karmaşık görevlerde oldukça etkili olduğunu göstermektedir. Başlangıç aşamasında hızlı öğrenme yetenekleri olsa da modellerin eğitim sürecinin ileri aşamalarında küçük iyileştirmeler dışında daha fazla gelişim sağlamadığı gözlemlenmiştir. Test veri setinde farklı modeller için segmentasyon performansları arasında bazı farklar olsa da genelde ortalamaların birbirine oldukça yakın olduğu görülmüştür. Ancak Şekil 5'te, her modelde ağırlıklı olarak %95 ve üzeri başarı sergileyen bir küme ile %80-92 aralığında bir başka küme olmak üzere iki farklı dağılım gözlemlenmektedir. Bu durum, test verisinde hangi model kullanılırsa kullanılsın belirli bir grubun sürekli daha iyi veya az da olsa daha düşük performans sergilediğini göstermektedir. Test veri setinde farklı çözünürlükte görüntülerin bulunması ya da belirli dış tiplerini segmente ederken modelin doğruluğunu olumsuz etkileyen yapısal farklılıklar olabileceğini düşündürmektedir. Ayrıca dışların şekli, büyüklüğü ve yapısı gibi faktörler ışığında, bazı dış tiplerinde modelin yüksek doğruluk gösterdiği, diğerlerinde ise ince detayların kaybolarak modelin performansının düştüğü düşünülmektedir. Sonuç olarak, bu bulgular, SegFormer modelinin dental segmentasyon gibi tıbbi görüntüleme alanlarında etkili bir şekilde kullanılabileceğini göstermektedir. Özellikle B2 versiyonu, en iyi genel performansı sunmaktadır. Ancak modellerin test verisinde iki farklı küme oluşturması, veri setindeki heterojenlikten kaynaklanabilecek olası sorunlara da işaret etmektedir.

Literatürdeki çalışmalarla kıyaslandığında, R-CNN ağı ve VGG-16 konvolüsyon mimarisine dayanan dış segmentasyonu modelinin model doğruluk oranı %89.8 olarak elde edilmiştir (Alam vd., 2023).



Şekil 7. X-ray giriş görüntüsü, referans (ground truth) görüntü ve model segmentasyon görüntüsü.

SWin-UNET'in diş segmentasyon performansını analiz edebilmek için, çalışmada PLAGH-BH veri seti için geliştirilen derin öğrenme ağının model doğruluğu %88.52 olarak elde edilmiştir (Sheng vd., 2023). Bir

başka çalışmada ise U-Net tabanlı ilgilenilen bölgenin çıkarılması ve Hibrit Görev Kaskatı tabanlı bir diş segmentasyonu ve numaralandırma modeline dayanan diş segmentasyonu için model başarımı %92,00 olarak

elde edilmiştir (Xu vd., 2023). Bu çalışmada önerdiğimiz diş segmentasyonu için elde edilen modelin başarımı %93,25 olarak elde edilmiştir.

Gelecek çalışmalarda, SegFormer modelinin performansını daha da iyileştirmek amacıyla farklı veri setleri üzerinde detaylı incelemeler yapılacaktır. Özellikle farklı çözünürlüklerdeki dental görüntüler ve çeşitli diş tiplerine ait veriler kullanılarak modelin genelleme yeteneği test edilecektir. Diş sınıfı veya çürük gibi daha özel verilerin test ve doğrulama aşamalarında daha belirgin bir şekilde ele alınması, modelin doğruluğunu ve pratik uygulanabilirliğini artırabilir. Bu yöndeki katkılar, özellikle klinik uygulamalarda segmentasyonun daha ayrıntılı bir şekilde değerlendirilmesine olanak tanıyacaktır. Gelecekteki çalışmalarda bu tür detaylı analizlere yer verilmesi planlanmaktadır. Ayrıca, veri setindeki çeşitliliğin artırılması, modelin segmentasyon doğruluğunu nasıl etkilediğini değerlendirmek açısından önemli bir adım olacaktır. Hibrit yaklaşımlar ile modelin daha geniş bir yelpazede test edilmesi ve genelleme kabiliyetinin artırılması için çeşitli stratejiler (örneğin, veri artırma teknikleri, transfer öğrenme) planlanmaktadır.

Katkı Oranı Beyanı

Yazarların katkı yüzdesi aşağıda verilmiştir. Tüm yazarlar makaleyi incelemiş ve onaylamıştır.

	E.B.	S.K.
K	50	50
T	50	50
Y	50	50
VTI	50	50
VAY	50	50
KT	50	50
YZ	50	50
KI	50	50
GR	50	50
PY	50	50
FA	50	50

K= kavram, T= tasarım, Y= yönetim, VTI= veri toplama ve/veya işleme, VAY= veri analizi ve/veya yorumlama, KT= kaynak tarama, YZ= Yazım, KI= kritik inceleme, GR= gönderim ve revizyon, PY= proje yönetimi, FA= fon alımı.

Çatışma Beyanı

Yazarlar bu çalışmada hiçbir çıkar ilişkisi olmadığını beyan etmektedir.

Etik Onay Beyanı

Bu çalışmada hayvanlar ve insanlar üzerinde herhangi bir çalışma yapılmadığı için etik kurul onayı alınmamıştır.

Kaynaklar

Alam MK, Haque T, Akhte, F, Albagieh HN, Nabhan AB, Alsenani MA, Islam S. 2023. Teeth segmentation by optical radiographic images using VGG-16 deep learning convolution architecture with R-CNN network approach for biomedical sensing applications. *Optic Quant Electron*, 55(9): 808.

Chen H, Zhang K, Lyu P, Li H, Zhang L, Wu J, Lee CH. 2019. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. *Sci Rep*, 9(1): 3840.

Gao C, Meng D, Yang Y, Wang Y, Zhou X, Hauptmann AG. 2013. Infrared patch-image model for small target detection in a single image. *IEEE Transact Image Proces*, 22(12): 4996-5009.

Gong Y, Zhang J, Cheng J, Yuan W, He L. 2024. Automatic tooth segmentation for patients with alveolar clefts guided by tooth descriptors. *Biomed Signal Proces Cont*, 90: 105821.

He K, Gkioxari G, Dollár P, Girshick R. 2017. Mask R-CNN. In *Proc IEEE Int Conf Comput Vision*, 2017: 2961-2969.

Humans In The Loop. 2023. Teeth Segmentation on dental X-ray images [Data set]. Kaggle.

Indraswari R, Arifin AZ, Navastara DA, Jawas N. 2015. Teeth segmentation on dental panoramic radiographs using decimation-free directional filter bank thresholding and multistage adaptive thresholding. *International Conference on Information & Communication Technology and Systems (ICTS)*, September 16, Surabaya, Indonesia, pp: 49-54.

Kang HC, Choi C, Shin J, Lee J, Shin YG. 2015. Fast and accurate semiautomatic segmentation of individual teeth from dental CT images. *Computat Math Meth Med*, 2015(1): 810796.

Kato S, Hotta K. 2024. Adaptive t-vMF dice loss: An effective expansion of dice loss for medical image segmentation. *Comput Biol Medic*, 168: 107695.

Lee SJ, Chung D, Asano A, Sasaki D, Maeno M, Ishida Y, Kobayashi T, Kuwajima Y, Silva JDD, Nagai S. 2022. Diagnosis of tooth prognosis using artificial intelligence. *Diagnostics*, 12(6): 1422.

Li H, Sun G, Sun H, Liu W. 2012. Watershed algorithm based on morphology for dental x-ray images segmentation. *International Conference on Signal Processing*, August 12-15, Hong Kong, China, pp: 877-880.

Li Z, Wang H. 2016. Interactive tooth separation from dental model using segmentation field. *PLoS One*, 11(8): e0161159.

Lira PHM, Giraldo GA, Neves LA. 2009. An automatic morphometrics data extraction method in dental x-ray image. *International Conference on Biodental Engineering*, June 26-27, Porto, Portugal, pp: 77-82.

Meng D, Zhao Q, Jiang L. 2017. A theoretical understanding of self-paced learning. *Info Sci*, 414: 319-328.

Modi CK, Desai NP. 2011. A simple and novel algorithm for automatic selection of ROI for dental radiograph segmentation. *Canadian Conference on Electrical and Computer Engineering*, May 08-11, Niagara Falls, Canada, pp: 000504-000507. <https://doi.org/10.1109/CCECE.2011.6030501>.

Nagaraju P, Sudha SV. 2024. Design of a novel panoptic segmentation using multi-scale pooling model for tooth segmentation. *Soft Comput*, 28(5): 4185-4196.

Poonsri A, Aimjirakul N, Charoenpong T, Sukjamsri C. 2016. Teeth segmentation from dental x-ray image by template matching. *9th Biomedical Engineering International Conference*, December 7-9, Laung Prabang, Laos, pp: 1-4.

Said EH, Nassar DEM, Fahmy G, Ammar HH. 2006. Teeth segmentation in digitized dental X-ray films using mathematical morphology. *IEEE Transact Info Forens Secur*, 1(2): 178-189.

Sheng C, Wang L, Huang Z, Wang T, Guo Y, Hou W, Yan X. 2023. Transformer-based deep learning network for tooth segmentation on panoramic radiographs. *J Syst Sci Compl*, 36(1): 257-272.

Tian S, Dai N, Zhang B, Yuan F, Yu Q, Cheng X. 2019. Automatic classification and segmentation of teeth on 3D dental model

- using hierarchical deep learning networks. *IEEE Access*, 7: 84817-84828.
- Wang CW, Huang CT, Lee JH, Li CH, Chang SW, Siao MJ, Lai TM, Ibragimov B, Vrtovec T, Ronneberger O, Fischer P, Cootes TF, Lindner C. 2016. A benchmark for comparison of dental radiography analysis algorithms. *Medic Image Analy*, 31: 63-76.
- Wirtz A, Mirashi SG, Wesarg S. 2018. Automatic teeth segmentation in panoramic X-ray images using a coupled shape model in combination with a neural network. In *Medical Image Computing and Computer Assisted Intervention-MICCAI 2018: 21st International Conference*, September 16-20, Granada, Spain, Part IV(11): 712-719.
- Wu A, Zhu L, Han Y, Yang Y. 2019. Connective cognition network for directional visual commonsense reasoning. *Adv Neural Info Proces Syst*, 32.
- Xie E, Wang W, Yu Z, Anandkumar A, Alvarez JM, Luo P. 2021. SegFormer: Simple and efficient design for semantic segmentation with transformers. *Adv Neural Info Proces Syst*, 34: 12077-12090.
- Xu M, Wu Y, Xu Z, Ding P, Bai H, Deng, X. 2023. Robust automated teeth identification from dental radiographs using deep learning. *J Dentistry*, 136: 104607.
- Yan M, Guo J, Tian W, Yi Z. 2018. Symmetric convolutional neural network for mandible segmentation. *Know Based Syst*, 159: 63-71.
- Yang J, Xie Y, Liu L, Xia B, Cao Z, Guo C. 2018. Automated dental image analysis by deep learning on small dataset. *42nd Annual Computer Software and Applications Conference (COMPSAC)*, July 23-27, Tokyo, Japan, pp: 492-497.
- Yong H, Meng D, Zuo W, Zhang L. 2017. Robust online matrix factorization for dynamic background subtraction. *IEEE Transact Pattern Analy Machine Intel*, 40(7): 1726-1740.
- Zhao Y, Li P, Gao C, Liu Y, Chen Q, Yang F, Meng D. 2020. TSASNet: Tooth segmentation on dental panoramic X-ray images by Two-Stage Attention Segmentation Network. *Know Based Syst*, 206: 106338
- Zhu L, Yang Y. 2020. Actbert: Learning global-local video-text representations. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 13-19, Seattle, US, pp: 8746-8755.