

Ali Erol

Kalibrasyondan Bağımsız Geometrik Normalizasyon ile Glottis Alan Ölçümü

Calibration-Independent Glottal Area Measurement via Geometric Normalization

Özet

Endoskopik görüntüler üzerinden glottis alanının güvenilir ve karşılaştırılabilir biçimde ölçülmesi, klinik değerlendirme açısından önemli bir gereksinimdir. Ancak muayene sırasında dinamik olarak değişen kamera–doku mesafesi ve piksel–milimetre kalibrasyon verilerinin eksikliği, standart piksel tabanlı ölçümlerin güvenilirliğini sınırlandırmaktadır. Bu çalışma, söz konusu ölçek bağımlılığı problemini aşmak amacıyla geometrik olarak normalize edilmiş, kalibrasyondan bağımsız bir ölçüm yaklaşımı sunmaktadır. Çalışma kapsamında, BAGLS veri kümesi kullanılarak nnU-Net v2 mimarisi üzerinde yüksek doğruluklu bir segmentasyon modeli eğitilmiştir. Elde edilen maskelerdeki glottis alanı, Temel Bileşenler Analizi (PCA) ile belirlenen baskın anatomik eksen uzunluğunun karesine göre normalize edilmiştir. Bu sayede, çekim mesafesinden bağımsız ve farklı görüntüler arasında doğrudan karşılaştırılabilir bir alan metriği elde edilmiştir. DeneySEL sonuçlar, geliştirilen modelin yüksek segmentasyon başarımı sağladığını ve önerilen normalizasyon yaklaşımının farklı çekim ölçekleri altında kararlı ve tutarlı nicel ölçümler sunduğunu göstermektedir.

Anahtar Kelimeler: Glottis segmentasyonu, nnU-Net, kalibrasyondan bağımsız ölçüm, geometrik normalizasyon, PCA

Abstract

Reliable and comparable measurement of the glottal area from endoscopic images is an important requirement for objective clinical assessment. However, during endoscopic examinations, the camera–tissue distance dynamically changes, and the absence of pixel-to-millimeter calibration data limits the reliability of standard pixel-based measurements. This variability makes direct quantitative comparison across different images challenging. To address this scale dependency problem, this study proposes a geometrically normalized, calibration-independent measurement approach. Within the scope of the study, a high-accuracy glottis segmentation model was trained using the nnU-Net v2 architecture on the BAGLS dataset. The glottal area extracted from the resulting segmentation masks was normalized with respect to the square of the dominant anatomical axis length, determined using Principal Component Analysis (PCA). Through this normalization strategy, a scale-independent area metric was obtained, enabling consistent comparison across varying imaging conditions. Experimental results indicate that the proposed model achieves robust segmentation performance and that the normalization

approach provides stable and comparable quantitative measurements under different capture scales.

Keywords: Glottis segmentation, nnU-Net, calibration-independent measurement, geometric normalization, PCA

1. Giriş

Ses teli fonksiyonlarının daha objektif ve tekrarlanabilir bir şekilde analiz edilebilmesi için nicel ölçümlere ihtiyaç duyulmaktadır [1,3]. Ancak endoskopik görüntülemelerde karşılaşılan en temel zorluk, kamera ile doku arasındaki mesafenin sabit olmamasıdır [2]. Endoskopik muayene sırasında kamera hastanın boğazına yaklaştıkça glottis alanı görüntüde daha büyük, uzaklaştıkça ise daha küçük görünmektedir. Bu durum, piksel biriminden fiziksel birime dönüşümü sağlayan bir kalibrasyon bilgisi olmadığında, mutlak alan ölçümlerini güvenilir hale getirmektedir [2,6].

Bu çalışmada, söz konusu kısıtlamaları aşmak amacıyla, derin öğrenme tabanlı segmentasyon ve şekil analizine dayalı hibrit bir yaklaşım önerilmektedir. Çalışmanın temel katkıları şunlardır:

- Glottis bölgesinin zorlu aydınlatma ve bulanıklık koşullarında bile yüksek doğrulukla ayrıştırılması için son teknoloji nnU-Net v2 mimarisi kullanılarak sağlam bir segmentasyon modeli geliştirilmiştir [4].
- Kamera mesafesindeki değişimlerden etkilenmeyen, ölçekten bağımsız (scale-invariant) ve boyutsuz yeni bir alan metriği (A/L^2) tanımlanmıştır [3,6].
- Glottisin karakteristik uzunluğu (L), Temel Bileşenler Analizi (PCA) yöntemiyle otomatik olarak hesaplanmış ve alanın bu uzunluğa göre normalize edilmesi sağlanmıştır [7].

2.5 İlgili Çalışmalar

Endoskopik görüntüleme sırasında kamera–doku mesafesinin değişken olması ve piksel–fiziksel birim dönüşümüne yönelik kalibrasyon bilgilerinin çoğu zaman mevcut olmaması, mutlak ölçümlere dayalı yaklaşımların güvenilirliğini sınırlayan temel bir problem olarak literatürde geniş biçimde ele alınmıştır [2,3]. Bu durum, özellikle glottis açıklığı gibi geometrik ölçümlerin farklı çekim koşulları altında karşılaştırılmasını zorlaştırmakta ve klinik yorumlamayı doğrudan etkilemektedir [6].

Literatürde glottis açıklığının analizi için farklı yöntemler önerilmiştir. Erken dönem çalışmalarda, manuel veya yarı otomatik yaklaşımlar ile eşikleme, kenar tespiti ve morfolojik işlemler kullanılarak glottis alanının çıkarılmaya çalışıldığı görülmektedir [3]. Bu yöntemler,

basit yapıları ve düşük hesaplama maliyetleri nedeniyle tercih edilmekle birlikte, gürültüye, aydınlatma değişimlerine ve kullanıcı bağımlılığına karşı sınırlı dayanıklılık göstermiştir [1].

Evrişimsel sinir ağlarının (CNN) yaygınlaşmasıyla birlikte, glottis segmentasyonu için derin öğrenme tabanlı yöntemler ön plana çıkmıştır. Özellikle U-Net mimarisi ve türevleri, endoskopik ve yüksek hızlı laringoskopi görüntülerinde glottis bölgesinin otomatik olarak segmentasyonu için yaygın biçimde kullanılmış; klasik görüntü işleme yaklaşımlarına kıyasla daha yüksek doğruluk ve tutarlılık sağladığı raporlanmıştır [4,5,9]. Güncel çalışmaların önemli bir kısmı, segmentasyon performansını artırmaya odaklanmakta ve başarımı ağırlıklı olarak Dice, IoU, Precision ve Recall gibi piksel tabanlı metrikler üzerinden değerlendirmektedir [8].

Bununla birlikte, literatürdeki mevcut çalışmaların büyük çoğunluğunda, segmentasyon çıktılarından elde edilen ölçümler doğrudan piksel cinsinden hesaplanmakta ve bu ölçümler görüntü ölçeğine ve kamera konumuna bağımlı kalmaktadır [3,6]. Bu durum, farklı kamera mesafelerinde veya farklı sistemlerle elde edilen görüntüler arasında nicel karşılaştırmaların yapılmasını güçleştirmekte ve ölçümlerin klinik anlamlılığını sınırlamaktadır. Kalibrasyon temelli yaklaşımlar bu sorunu kısmen ele alabilse de, pratik klinik ortamlarda gerekli kalibrasyon bilgilerinin çoğu zaman mevcut olmaması, bu yöntemlerin yaygın kullanımını zorlaştırmaktadır [2].

Bu çalışmada, literatürde tanımlanan söz konusu sınırlamalardan yola çıkılarak, kalibrasyon bilgisine ihtiyaç duymadan glottis açıklığını değerlendirebilen segmentasyon tabanlı bir yaklaşım önerilmektedir. nnU-Net v2 tabanlı bir derin öğrenme modeli kullanılarak glottis bölgesi yüksek doğrulukla segmente edilmekte; elde edilen ikili maskeler üzerinden alan bilgisi hesaplanmakta ve glottis şeklinin baskın anatomik eksenini Temel Bileşenler Analizi (PCA) ile belirlenmektedir [4,7]. Alan bilgisinin bu ana eksene göre normalize edilmesiyle, ölçekten bağımsız ve boyutsuz bir alan metriği (A/L^2) elde edilmekte ve böylece kamera mesafesinden kaynaklanan ölçek değişimlerine karşı daha kararlı bir nicel değerlendirme amaçlanmaktadır.

3. Yöntem

Önerilen yöntem dört ana aşamadan oluşmaktadır:

1. Glottis segmentasyonu
2. Piksel alanı hesaplanması
3. PCA tabanlı uzunluk (ana eksen) tahmini
4. Normalize glottis alanı metriğinin elde edilmesi

Aşağıda her bir aşama matematiksel olarak açıklanmaktadır.

3.1 Glottis Segmentasyonu

Segmentasyon problemi, görüntüdeki her pikselin glottis bölgesine ait olup olmadığının belirlenmesi şeklinde tanımlanır. Bu amaçla, glottis yapısını arka plandan ayıran ikili bir segmentasyon maskesi üretilir:

$$\hat{M} \in \{0,1\}^{H \times W}$$

Burada $\hat{M}(x, y) = 1$ glottis bölgesini, $\hat{M}(x, y) = 0$ ise arka planı temsil eder. Bu tür piksel düzeyinde glottis segmentasyonu, endoskopik ve yüksek hızlı laringoskopi görüntülerinde daha önce de başarıyla uygulanmış ve literatürde yaygın kabul görmüştür [9]. İkili maske, derin öğrenme modeli tarafından üretilen olasılık haritası SSS üzerinden sabit bir eşik değeri τ kullanılarak aşağıdaki şekilde elde edilir:

$$\hat{M}(x, y) = \begin{cases} 1, & S(x, y) \geq \tau \\ 0, & otherwise \end{cases}$$

3.2 Glottis Alanının Piksel Biriminde Hesaplanması

Segmentasyon maskesi kullanılarak glottis alanı, piksel sayımı yöntemiyle hesaplanmıştır. Alan A, maske içerisindeki ön plan piksellerinin toplamı olarak tanımlanır:

$$A = \sum_{x=1}^W \sum_{y=1}^H M(x, y)$$

Bu büyüklük, glottis alanını piksel kare (px²) cinsinden temsil etmektedir. Piksel sayımına dayalı alan hesaplama yaklaşımı, glottis segmentasyonu literatüründe yaygın olarak kullanılan temel bir yöntemdir [3,6]. Glottis alanına dayalı nicel değerlendirmelerin, klinik glottis analizlerinde anlamlı bir ölçüt sunduğu daha önceki çalışmalarda da vurgulanmıştır [10]. Bununla birlikte, bu tür alan ölçümleri kamera mesafesi ve görüntü ölçeğindeki değişimlere doğrudan duyarlıdır.

3.3 PCA Tabanlı Glottis Uzunluğunun Hesaplanması

Kamera mesafesinden daha az etkilenen bir referans uzunluk elde etmek amacıyla, glottis maskesinin geometrik yapısı analiz edilmiştir. Bu tür geometrik şekil analizlerinde, nesnenin baskın yönelimini belirlemek amacıyla Temel Bileşenler Analizi (PCA) sıklıkla tercih edilen bir yöntemdir [7].

Öncelikle, glottis maskesine ait tüm ön plan piksel koordinatları elde edilir:

$$P = \{ (x_i, y_i) \mid M(x_i, y_i) = 1 \}, \quad i = 1, \dots, N$$

Bu noktalar kullanılarak aşağıdaki adımlar uygulanır.

3.3.1 Nokta Merkezleme

Koordinatların ortalaması hesaplanır:

$$\mu = \left(\frac{1}{N} \right) \sum_{i=1}^N \begin{bmatrix} x_i \\ y_i \end{bmatrix}$$

Merkezlenmiş noktalar:

$$\tilde{p}_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix} - \mu$$

3.3.2 Kovaryans Matrisi

Merkezlenmiş noktalardan kovaryans matrisi oluşturulur:

$$\Sigma = \frac{1}{N-1} \sum_{i=1}^N \tilde{p}_i \tilde{p}_i^T$$

Bu matris, glottis maskesinin uzaysal dağılımını temsil eder.

3.3.3 Özdeğer–Özvektör Ayrışımı

Kovaryans matrisi üzerinde özdeğer ayrışımı yapılır:

$$\Sigma v_j = \lambda_j v_j$$

Burada λ_j özdeğerleri, v_j ise karşılık gelen özvektörleri ifade etmektedir. En büyük özdeğere karşılık gelen özvektör. PCA'nın temel varsayımlarına uygun olarak glottisin ana geometrik eksenini kabul edilmiştir [7].

$$V_{\{\max\}}$$

glottisin ana geometrik eksenini olarak kabul edilmiştir.

3.3.4 Uzunluk Hesabı

Tüm merkezlenmiş noktalar bu eksene projekte edilir:

$$\alpha_i = \tilde{p}_i^T V_{\{\max\}}$$

Glottis uzunluğu L, bu projeksiyonların maksimum ve minimum değerleri arasındaki fark olarak tanımlanır:

$$L = \max(\alpha_i) - \min(\alpha_i)$$

Bu uzunluk değeri piksel birimindedir ve glottisin ana eksenini boyunca uzamsal yayılımını temsil eder.

3.4 Normalize Glottis Alanı Metriği

Piksel alanı A tek başına kamera mesafesine duyarlı olduğundan, alan ölçümünün normalize edilmesi amaçlanmıştır. Bu çalışmada glottisin karakteristik uzunluğu olarak L kullanılarak aşağıdaki normalize alan metriği tanımlanmıştır:

$$A_{norm} = \frac{A}{L^2}$$

Bu ifade boyutsuzdur ve kamera mesafesi deęişimlerine karşı daha kararlı bir ölçüm sunar. Göreceli ve karşılaştırmalı glottis ölçümlerinin klinik yorumlamada daha anlamlı olduęu, daha önceki çalışmalarda da vurgulanmıştır [10]. Alan bilgisinin karakteristik bir uzunluk ile normalize edilmesi, mutlak ölçekten bağımsız nicel deęerlendirmeler elde etmeyi amaçlayan şekil tabanlı analiz yaklaşımlarıyla uyumludur [3,6].

3.5 Deęerlendirme Yaklaşımı

Elde edilen segmentasyon maskeleri ve geometrik ölçümler, hem bölge tabanlı hem de sınır tabanlı metrikler kullanılarak deęerlendirilmiştir. Segmentasyon doğruluęu klasik metriklerle, sınır uyumu ise Hausdorff Mesafesi (HD95) ve Ortalama Yüzey Mesafesi (ASD) ile analiz edilmiştir. Bu metrikler, medikal görüntü segmentasyonunda sınır doğruluęunu deęerlendirmek için yaygın olarak kullanılmaktadır [8].

3.6 Yöntemin Genel Akışı

Önerilen yöntemin genel iş akışı aşağıdaki gibidir:

$$I \rightarrow \text{nnU-Net} \rightarrow M \rightarrow A \rightarrow \text{PCA} \rightarrow L \rightarrow \frac{A}{L^2}$$

Bu akış sayesinde, kalibrasyon bilgisi olmayan endoskopik görüntülerden glottis açıklığına ilişkin ölçekten bağımsız bir ölçüm elde edilmektedir.

4. Deneyisel Sonuçlar

Bu çalışmada kullanılan veriler, Kaggle platformunda yayımlanan ve glottis segmentasyonu için referans niteliğinde olan Benchmark for Automatic Glottis Segmentation (BAGLS) veri kümesinden elde edilmiştir [11].

Bu çalışmada, **nnU-Net** modeli, **BAGLS** veri kümesinden rastgele seçilen **6000** adet tek kanallı (**grayscale**) endoskopik görüntü kullanılarak eğitilmiştir. Tüm görüntüler, modelin girdi gereksinimlerine uygun olacak şekilde tek renk kanalı (grayscale) formatında kullanılmıştır. nnU-Net mimarisinin bu tür medikal görüntülerde yüksek genellenebilirlik sağladığı literatürde daha önce gösterilmiştir [4].

Eğitim sürecinin ardından model performansı, aynı veri kümesinin test klasöründen seçilen 50 adet bağımsız test görüntüsü üzerinde deęerlendirilmiştir. Test aşamasında, modelden elde edilen segmentasyon çıktıları ile yer gerçeęi (ground truth) maskeleri karşılaştırılarak **Dice**, **IoU**, **Precision** ve **Recall** metrikleri hesaplanmıştır. Bu metrikler, medikal görüntü segmentasyonunda performans deęerlendirmesi için yaygın olarak kullanılan standart ölçütlerdir [4,8].

Her bir metrik için **ortalama (Mean)**, **standart sapma (Std)** ve **medyan (Median)** deęerleri raporlanmış olup, sonuçlar **Tablo 1**'de sunulmaktadır.

4.1 Performans Metrikleri

Test kümesi üzerindeki genel performans sonuçları **Tablo 1’de**, sadece glottisin açık olduğu(pozitif) vakaların sonuçları ise **Tablo 2’de** verilmiştir.

Tablo 1: Test Kümesi Performansı (50 Test Örneği)

Stat	Dice	IoU	Precision	Recall	AbsErr_Norm	RelErr_Norm
Mean	0.746	0.678	0.744	0.763	0.021	1.65×10^5
Std	0.337	0.320	0.347	0.348	0.057	1.17×10^6
Median	0.877	0.781	0.912	0.901	0.006	5.89×10^{-2}

Açıklama:Yüksek precision ve recall değerleri, yöntemin glottis bölgesini güvenilir biçimde segmentleyebildiğini göstermektedirStandart sapmanın görece yüksek olması ise özellikle glottisin kısmen veya tamamen kapalı olduğu örneklerde segmentasyon performansının düşmesinden kaynaklanmaktadır. Benzer davranışlar, glottis segmentasyonu literatüründe de raporlanmıştır [3,6].

4.1.1 Pozitif Vakalar Üzerinde Ayrıntılı Performans Analizi

- **A_gt > 0** olan **42 vaka**
- **A_gt = 0** olan **8 negatif vaka**

Segmentasyon performansını glottisin gerçekten mevcut olduğu durumlar için daha doğru şekilde karakterize etmek amacıyla, metrikler **yalnızca pozitif vakalar (n=42)** üzerinde ayrıca raporlanmıştır. **Tablo 2’de** sunulan bu sonuçlar, modelin glottis açıklığı bulunan örneklerde daha kararlı (daha düşük varyanslı) bir performans sergilediğini göstermektedir. Pozitif vakalara odaklanarak yapılan bu tür analizlerin, piksel tabanlı ölçümlerin sınırlılıklarını azaltmada etkili olduğu literatürde vurgulanmaktadır [3,6].

Tablo 2: Sadece Pozitif Vakalar (Positive Only) İçin Performans Özeti

Stat	Dice	IoU	Precision	Recall	AbsErr	RelErr
Mean	0.888	0.807	0.885	0.909	0.023	1.637
Std	0.082	0.125	0.123	0.098	0.061	9.781
Median	0.903	0.823	0.934	0.948	0.007	0.077

Görüldüğü üzere, yalnızca pozitif vakalar dikkate alındığında performans metriklerinin standart sapmaları daha anlamlı bir aralığa oturmakta ve modelin glottis açık vakalardaki segmentasyon başarısı daha tutarlı biçimde raporlanabilmektedir.

4.2 Confusion Matrix

Tablo 3: Confusion Matrix (Glottis Varlığı Tespiti)

	Pred Empty	Pred Non-empty
GT Empty	7 (TN)	1 (FP)
GT Non-empty	0 (FN)	42 (TP)

Confusion matrix analizi, modelin glottis varlığını tespit etmede yüksek güvenilirliğe sahip olduğunu göstermektedir. Glottisin bulunmadığı 8 vakanın 7'si doğru şekilde boş olarak sınıflandırılmış, yalnızca bir vakada yanlış pozitif tahmin üretilmiştir. Yanlış negatif durumun gözlemlenmemesi, modelin glottis varlığını kaçırma eğiliminin düşük olduğunu ortaya koymaktadır. Bu tür var/yok analizleri, glottis segmentasyonu çalışmalarında klinik açıdan önemli bir değerlendirme adımı olarak ele alınmaktadır [1,9].

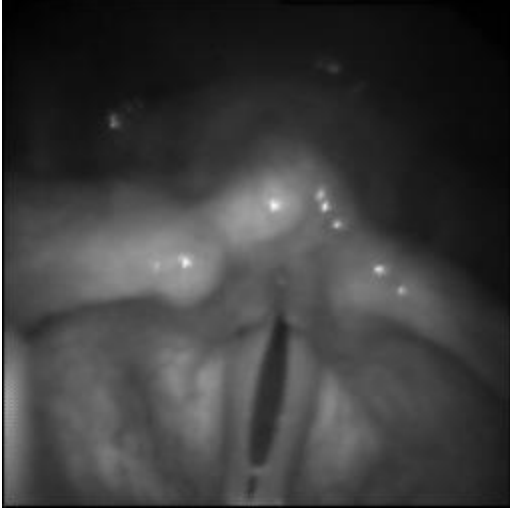
4.3 Rastgele Seçilen Test Örnekleri Üzerinde Ayrıntılı Vaka Analizi

Modelin farklı glottis geometrilerine sahip örnekler üzerindeki davranışını daha ayrıntılı incelemek amacıyla, seçili 10 test vakasına ait nicel sonuçlar **Tablo 3'da** sunulmuştur. Vaka bazlı nicel ve nitel analizlerin birlikte raporlanması, glottis segmentasyonu literatüründe yaygın bir değerlendirme yaklaşımıdır [1,5].

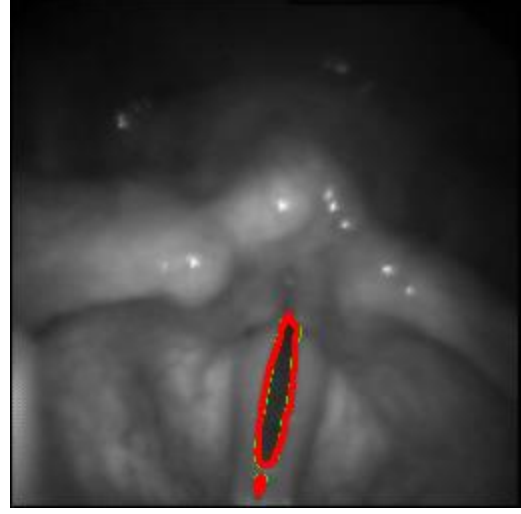
Tablo 3: Rastgele Seçilen 10 Test Vakasının Sonuçları

Case	Dice	IoU	Prec.	Recall	A_GT	A_Pred	L_GT	L_Pred	N_GT	N_Pred	AbsErr	RelErr
Case 0	0.906	0.827	0.950	0.865	266	242	108.7	108.2	0.023	0.021	0.002	8.20e-2
Case 1	0.970	0.941	0.968	0.972	892	896	89.4	91.4	0.112	0.107	0.004	3.83e-2
Case 2	0.656	0.488	0.488	1.000	84	172	21.1	27.0	0.189	0.235	0.046	2.45e-1
Case 3	0.914	0.841	0.849	0.989	187	218	40.0	42.2	0.117	0.122	0.006	4.86e-2
Case 4	0.977	0.956	0.974	0.981	3572	3599	201.3	200.6	0.088	0.089	0.001	1.44e-2
Case 5	0.000	0.000	0.000	0.000	0	183	0.0	47.1	0.000	0.083	0.083	8.26e6
Case 6	0.922	0.855	0.986	0.866	656	576	93.9	93.0	0.074	0.067	0.008	1.03e-1
Case 7	0.894	0.809	0.920	0.870	937	886	94.0	95.3	0.106	0.097	0.009	8.05e-2
Case 8	0.900	0.819	0.834	0.978	369	433	163.0	168.0	0.014	0.015	0.001	1.05e-1
Case 9	0.976	0.954	0.962	0.991	5449	5618	129.4	128.9	0.325	0.338	0.013	3.89e-2

Şekil 1'de modelin örnek bir segmentasyon çıktısı sunulmuştur.



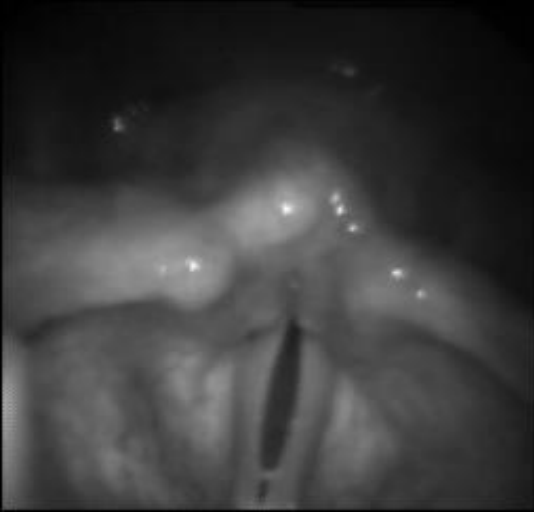
(a)Orjinal Görüntü



(b)Segmentasyon Maskesi

Şekil 1: nnU-Net modeli ile test kümesinden rastgele seçilen bir örnek üzerinde elde edilen segmentasyon sonucu.

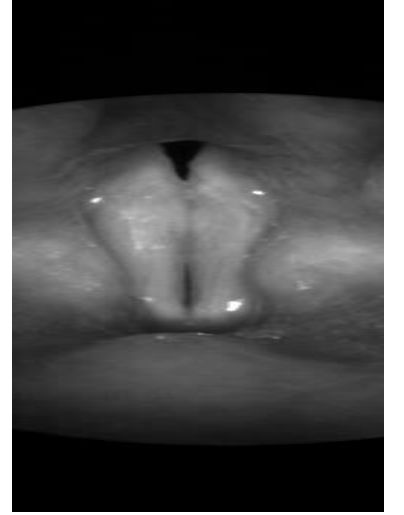
Not: Case 1’de segmentasyon metriklerinin yüksek ve tutarlı değerler alması, modelin glottis bölgesini doğru biçimde ayırt edebildiğini göstermektedir. Buna bağlı olarak, segmentasyon maskesinden türetilen normalize alan ölçümleri de anlamlı ve güvenilir kabul edilebilir.



(a) Norm Pred:0.10



(b) Norm Pred:0.06



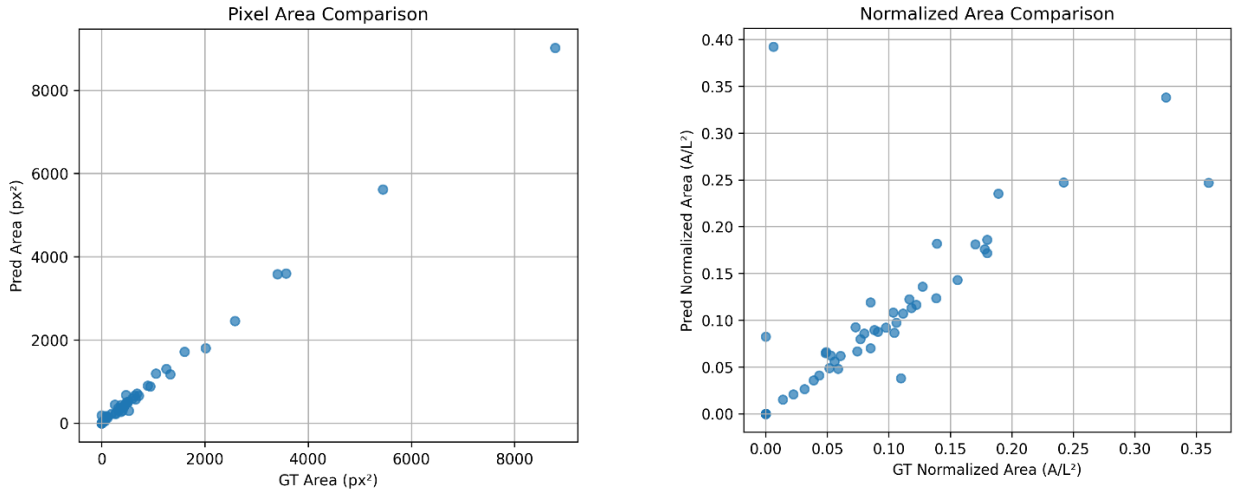
(c) Norm Pred:0.01

Şekil 2: Farklı glottis açıklık seviyelerine sahip ham (etiketsiz) test görüntüleri ve modelin hesapladığı normalize alan (A/L^2) değerleri. **(a)** Geniş açık glottis (0.10), **(b)** Orta seviye açıklık (0.06), **(c)** Neredeyse kapalı glottis (0.01).

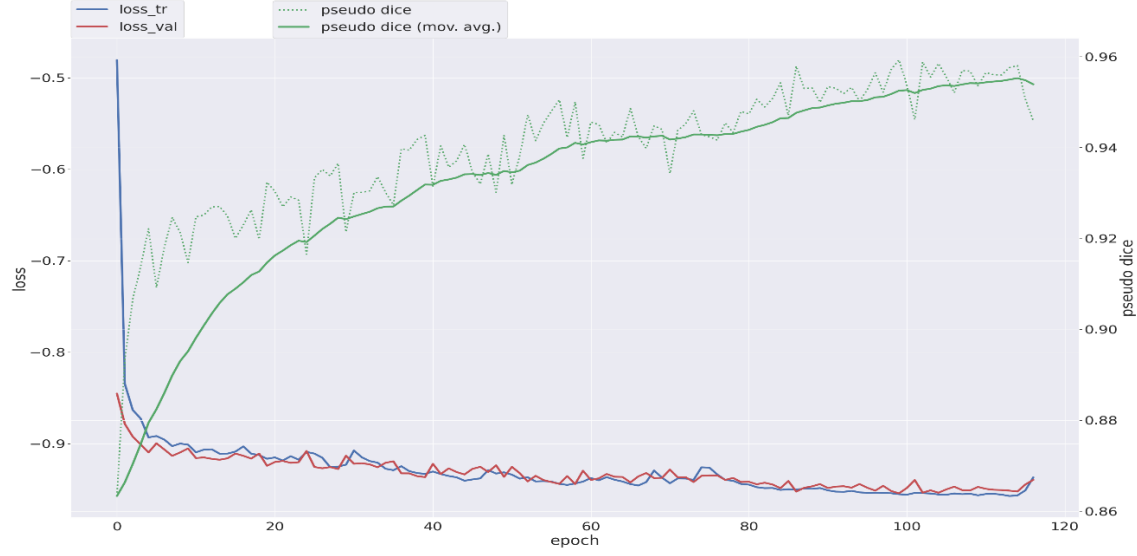
Görsel karşılaştırma incelendiğinde, soldan sağa doğru ilerledikçe glottis açıklığının kademeli olarak daraldığı açık biçimde gözlemlenmektedir. Bu morfolojik değişim, glottis alanının ve buna bağlı olarak normalize glottis alanı değerlerinin azalmasının beklendiğini göstermektedir. Nitekim tablo sonuçları da bu görsel gözlemi doğrulamakta; glottis boşluğunun daraldığı örneklerde normalize alan değerlerinin sistematik biçimde azaldığı görülmektedir. Bu uyum, önerilen yöntemin hem görsel hem de nicel olarak tutarlı sonuçlar ürettiğini ortaya koymaktadır.

Bu görüntü, modelin gerçek test koşulları altındaki segmentasyon performansını görsel olarak temsil etmektedir.

4.4 Eğitim Grafikleri ve Alan Analizi



Şekil 3: Yer gerçeği (GT) ve model tahmini (Pred) arasındaki korelasyon analizi. **(Sol)** Piksel biriminden ham alan karşılaştırması (px^2), **(Sağ)** Önerilen normalize alan metriğinin (A/L^2) karşılaştırması.



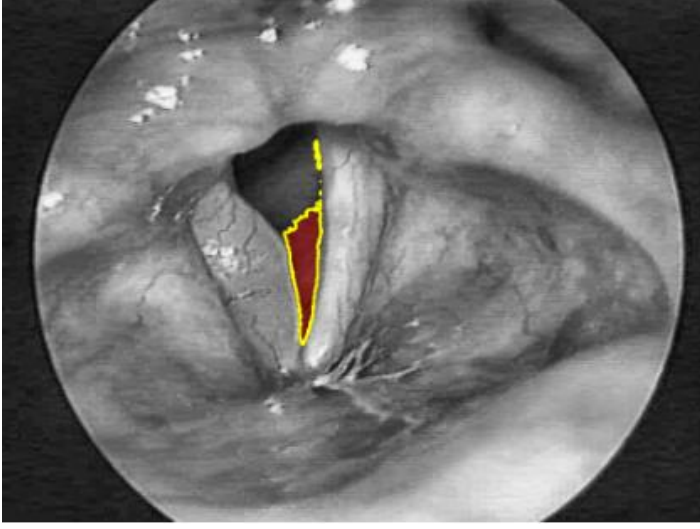
Şekil 4: nnU-Net modelinin eğitim süreci boyunca kayıp (loss) ve başarıım (Dice) grafiği. **Mavi:** Eğitim kaybı ($loss_{tr}$), **Kırmızı:** Doğrulama kaybı ($loss_{val}$), **Yeşil:** Ortalama Dice skoru.

Eğitim ve doğrulama kayıpları arasında belirgin bir ayrışma gözlenmemesi, modelin overfitting yapmadığını göstermektedir. Artan ve kararlı pseudo Dice eğrisi, öğrenmenin stabil ve genellenebilir olduğunu doğrulamaktadır. Bu tür eğitim davranışları, nnU-Net mimarisine dayalı çalışmalarda beklenen bir durumdur [4].

5. Tartışma

Kapalı glottis durumlarında modelin istisnai olarak küçük yanlış pozitif segmentasyonlar üretebildiği ve bunun normalize alan değerlerini yapay olarak artırabildiği gözlemlenmiştir. Ayrıca BAGLS veri kümesi dışında kalan bazı klinik görüntülerde, eğitim verisi ile test verisi arasındaki dağılım farkına (domain shift) bağlı olarak segmentasyon performansında düşüşler meydana gelmiştir. Bu bulgular, gerçek dünya uygulamalarında ek kontrol mekanizmalarının ve daha geniş veri çeşitliliğinin önemine işaret etmektedir.

Gelecek çalışmalarda, aynı hastaya ait farklı kamera mesafelerinde çekilmiş görüntüler kullanılarak normalize alan metriğinin ölçek değişimlerine karşı davranışının doğrudan test edilmesi planlanmaktadır. Buna ek olarak, dış klinik verilerle modelin iyileştirilmesi ve kapalı glottis durumlarını ayırt edebilecek ek bir “glottis var/yok” karar mekanizmasının geliştirilmesi, yöntemin klinik uygulanabilirliğini artırmaya yönelik önemli adımlar olacaktır.



a)Klinik Görüntü

Alan (A) : 3102 px²
Uzunluk (L) : 209.45739158287515 px
A / L² : 0.07070504642038451

BAGLS veri kümesinden farklı olarak, gerçek klinik örnekte glottis açıklığı belirgin şekilde daha büyük olup görüntüleme koşulları (ölçek, zoom ve kontrast) farklılık göstermektedir. Bu dağılım farkı (domain shift), modelin eğitim sürecinde karşılaşmadığı bir görsel istatistikle karşılaşmasına neden olmuş ve segmentasyon hatalarının artmasına yol açmıştır.

6. Sonuç

Normalize alan metriği açısından yapılan değerlendirmeler, önerilen ölçütün glottis açıklığındaki göreceli değişimleri tutarlı ve kararlı biçimde yansıttığını göstermektedir. Glottis açıklığının arttığı vakalarda normalize alan değerlerinin sistematik olarak yükselmesi, açıklığın azaldığı veya kapandığı durumlarda ise düşmesi, yöntemin anatomik ve geometrik açıdan anlamlı bir davranış sergilediğini ortaya koymaktadır. Bu bulgular, normalize alan metriğinin karşılaştırmalı glottis analizi için güvenilir bir nicel gösterge olarak kullanılabileceğini desteklemektedir.

Sonuç olarak, geliştirilen nnU-Net tabanlı segmentasyon modeli ve PCA destekli normalize alan ölçütü, pozitif glottis vakalarında yüksek segmentasyon başarımı ve göreceli açıklık ölçümünde tutarlı sonuçlar üretmiştir. Elde edilen bulgular, kalibrasyon bilgisinin bulunmadığı endoskopik görüntülerde glottis açıklığının nicel ve karşılaştırmalı analizi için önerilen yaklaşımın uygulanabilir ve güvenilir bir temel sunduğunu göstermektedir.

7)Kaynaklar

[1]

Deliyski, D. D., Hillman, R. E.,
Quantitative laryngeal imaging: Clinical applications and challenges,
Journal of Voice, Cilt 24, Sayı 5, ss. 620–631, 2010.

[2]

Deliyski, D. D.,
Endoscopic imaging challenges in quantitative voice assessment,
Journal of Voice, Cilt 20, Sayı 3, ss. 450–465, 2006.

[3]

Lohscheller, J., Eysholdt, U.,
Phonovibrography: Quantification of vocal fold vibration using endoscopic imaging,
Journal of Voice, Cilt 22, Sayı 4, ss. 422–433, 2008.

[4]

Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., Maier-Hein, K. H.,
nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation,
Nature Methods, Cilt 18, ss. 203–211, 2021.

[5]

Ding, Y., Li, Y., Wang, Q.,
Automatic glottis segmentation for laryngeal endoscopic images based on U-Net,
Biomedical Signal Processing and Control, Cilt 68, 2021.

[6]

Lohscheller, J., Eysholdt, U.,
Limitations of pixel-based glottal area measures in endoscopic imaging,
Journal of Voice, Cilt 21, Sayı 4, ss. 457–464, 2007.

[7]

Jolliffe, I. T.,
Principal Component Analysis,
Springer Series in Statistics, 2. Baskı, 2002.

[8]

Taha, A. A., Hanbury, A.,
Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool,
BMC Medical Imaging, 2015.

[9]

Fehling, M. K., et al.,
Automatic detection and segmentation of the glottis in high-speed videoendoscopy,
PLOS ONE, 2020.

[10]

Lohscheller, J., et al.,
Clinical evaluation of vocal fold vibration using endoscopic image analysis,
Journal of Voice, 2014.

[11]

Gómez, P., *et al.*,

Benchmark for Automatic Glottis Segmentation (BAGLS) Dataset,

Kaggle, 2020.

Erişim adresi: <https://www.kaggle.com/datasets/gomezp/benchmark-for-automatic-glottis-segmentation>