

후두 내시경 영상에서의 성문 분할 및 성대 점막 형태의 정량적 평가

이선민[†], 오 석^{**}, 김영재^{***}, 우주현^{****}, 김광기^{*****}

Segmentation of the Glottis and Quantitative Measurement of the Vocal Cord Mucosal Morphology in the Laryngoscopic Image

Seon Min Lee[†], Seok Oh^{**}, Young Jae Kim^{***}, Joo Hyun Woo^{****}, Kwang Gi Kim^{*****}

ABSTRACT

The purpose of this study is to compare and analyze Deep Learning (DL) and Digital Image Processing (DIP) techniques using the results of the glottis segmentation of the two methods followed by the quantification of the asymmetric degree of the vocal cord mucosa. The data consists of 40 normal and abnormal images. The DL model is based on Deeplab V3 architecture, and the Canny edge detector algorithm and morphological operations are used for the DIP technique. According to the segmentation results, the average accuracy of the DL model and the DIP was 97.5% and 94.7% respectively. The quantification results showed high correlation coefficients for both the DL experiment ($r=0.8512$, $p<0.0001$) and the DIP experiment ($r=0.7784$, $p<0.0001$). In the conclusion, the DL model showed relatively higher segmentation accuracy than the DIP. In this paper, we propose the clinical applicability of this technique applying the segmentation and asymmetric quantification algorithm to the glottal area in the laryngoscopic images.

Key words: Laryngoscopy, Vocal Cord, Segmentation, Quantitative measurement, Deep Learning, Digital Image Processing

1. 서 론

후두내시경 검사(Laryngoscopy)는 음색, 크기, 음

질의 이상으로 특징되는 음성 장애가 의심될 경우 후두와 성대(vocal cord)를 관찰하기 위해 시행한다 [1]. 성대 점막의 형태적 변화가 심할 경우 판독이

※ Corresponding Author : Kwang Gi Kim, Address: 38-13, Docjeom-ro 3beon-gil, Namdong-gu, Incheon, 21565, Republic of Korea, TEL : +82-32-458-2770, FAX : +82-32-460-2361, E-mail : kimkg@gachon.ac.kr

※ Corresponding Author : Joo Hyun Woo, Address: 21, Namdong-daero 774beon-gil, Namdong-gu, Incheon, 21565, Republic of Korea, TEL : +82-32-458-2567, FAX : +82-32-467-9044, E-mail : woojh@gilhospital.com

Receipt date : Mar. 2, 2022, Revision date : Apr. 14, 2022
Approval date : Apr. 26, 2022

[†] Dept. of Nursing, College of Nursing, Gachon University (E-mail : smlee7182@gachon.ac.kr)

^{**} Dept. of Biomedical Engineering, Gil Medical Center, College of Medicine, Gachon University (E-mail : seogo1216@gmail.com)

^{***} Dept. of Biomedical Engineering, Gil Medical Center, College of Medicine, Gachon University (E-mail : youngjae@gachon.ac.kr)

^{****} Dept. of Otorhinolaryngology-Head and Neck Surgery, Gil Medical Center, College of Medicine, Gachon University

^{*****} Dept. of Health Sciences & Technology, Gachon Advanced Institute for Health Sciences & Technology (GAIHST), Gachon University

※ This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2021-2017-0-01630) supervised by the IITP(Institute for Information & communications Technology Promotion), and by the GRRC program of the Gyeonggi province (No. GRRC-Gachon2020(B01), respectively. In addition, the research was supported by the G-ABC FRD2019-11-02(3).

수월하지만, 사람의 성대는 1초에 100~300회 정도로 빠르게 진동하기 때문에, 일반적인 내시경 장비만을 사용하여 모든 성대를 정확하게 관찰하는 것은 쉽지 않다[2]. 이에 내시경 영상을 모니터로 출력하여 성대의 움직임을 시각화한 후 좌우 성대의 운동 거리, 성문 폐쇄 정도, 점막 파동 등을 관찰하는 후두 스트로보스코피(Laryngeal Stroboscopy), 초고속 비디오 후두내시경 검사(High-speed videolaryngoscopy, HSV) 등이 도입되어 왔다[3-5]. 그러나 두 방법 모두 성대 점막 간의 형태에 주관적 시지각 평가를 적용하므로 정량적 평가로서는 부족하다는 한계가 있다[6].

이러한 한계점을 해결하기 위해 초고속 후두내시경 영상을 정량적으로 분석하는 연구가 2010년 미국의 루이지애나 대학에서 진행되었다[7]. Phonovibrogram을 사용하여 해부학적 구조를 기반으로 영상에서 양측 성대를 각각 자동 분할(segmentation)하고, 성대 점막의 윤곽선을 확보하여 중심부에서 좌우 성대 점막의 윤곽선까지의 거리를 측정하였다. 해당 연구는 컴퓨터 보조 진단 시스템(Computer-aided diagnosis system, CADs)으로 발전 가능한 자동 분할 및 정량적 분석법의 효용성을 제시하였다는 점에서 의의가 있으나, 정상 소견의 데이터만 사용하였다는 제한점이 존재한다. 이를 보완하여 다양한 소견의 후두 성대를 대상으로 성대 점막의 형태적 비대칭을 정량화하는 자동 시스템이 더욱 활발히 연구될 필요가 있다.

내시경 영상에서 성대를 정량적으로 평가하기 위해서는 정확한 후두 성대 영역의 분할이 우선되어야 한다[8]. 현재까지 진행된 영상 분할 연구들에 따르면 성대 영역을 분할하기 위한 방법에는 디지털 영상처리(Digital image processing, DIP)와 딥러닝(Deep learning, DL) 등이 있다. 디지털 영상처리 방법을 사용한 연구의 사례로, 2018년 김진필 등은 OpenCV 라이브러리를 사용하여 24비트 컬러의 홍채 영상에서 라플라시안(Laplacian), 소벨(Sobel), 캐니 엣지 검출(Canny edge detector) 알고리즘 등을 사용하여 홍채 얼굴을 분할하였고, 캐니 엣지 검출 알고리즘이 타 알고리즘에 비해 정확함을 확인하였다[9]. 2020년 조용현은 RGB 색상의 피부 영상에서 색상 특징과 질감 특징을 조합한 3가지 방식으로 병변을 분할하고 그 성능을 비교하였다[10]. 그 결과, 평균 정확도는 3가지 방식 모두 87% 이상이었으며 그 중에서도

두 가지의 특징을 조합한 방식은 89.34%로 우수한 성능을 나타냈다.

딥러닝을 사용한 연구 사례로는 2018년 김범상 등이 축소된 U-Net 네트워크를 사용하여 컬러 망막 영상에서 망막 혈관을 분할한 바 있다[11]. 제안한 방법의 정확도는 95.23%로, 최대 94.70%의 정확도를 나타낸 기존 방법들에 비해 높은 성능을 나타냈다. 또한 2021년 권관영 등이 U-Net 모델을 이용하여 피부 발적의 경계를 분할하는 연구를 진행하였다[12]. 제안한 방법의 평균 유사도는 94.93%, 평균 정밀도와 민감도는 각각 95.19%, 95.24%로 우수한 결과를 보였다.

다양한 의료 영상에서 분할 알고리즘들이 활용되고 있는 반면, 후두 성대는 분할을 시도한 연구 사례가 미비하다. 따라서 후두 성대의 자동 분할을 시도하고, 분할한 성대의 정량적 평가 가능성을 확인한 필요가 있다.

본 논문에서는 후두 성대 영역의 분할을 위해 디지털 영상처리 및 딥러닝 기반의 분할 방법들을 제시하고, 그 결과를 비교해보고자 한다. 또한 각 방법을 통해 분할된 후두 성대 영역을 대상으로 좌우 성대 점막의 형태적 비대칭 정도를 측정하고, 그 결과를 비교 분석함으로써 분할한 영역의 정량화를 제안하는 두 방법의 활용 가능성을 평가해보고자 한다.

2. 연구 방법

2.1 데이터 수집

본 연구는 가천대 길병원에서 2015년 01월 01일부터 2020년 12월 31일까지 후두내시경을 시행한 환자 40명을 대상으로 하였다. 640×480 픽셀의 해상도와 24비트 컬러의 후두내시경 원본 이미지를 사용하였고, 원본 데이터는 판독이 완료된 정상 후두 10건, 양성 및 악성 후두 질환을 포함한 비정상 후두 30건으로 구성하였다. 모든 데이터는 IRB(GDIRB2021-165) 승인 하에 수집된 후두내시경 데이터를 선별하여 사용하였다.

2.2 개발환경

실험 환경은 NVIDIA GeForce RTX 2080 TI (NVIDIA, Santa Clara, Calif) 그래픽 처리 장치와 3.60 GHz Intel Xeon E5-1650(Intel, Santa Clara,

Calif) CPU 및 128GB RAM으로 구성되었다. 데이터 전처리는 ImageJ(NIH, Bethesda, MD, USA) 프로그램을 사용하였고, 딥러닝 학습은 Ubuntu 18.04.05 운영체제에서 Python 3.7.0(Python Software Foundation, Wilmington, DE, USA)을 통해 수행되었으며 Pytorch 1.9.0 프레임워크가 사용되었다. 디지털 영상처리에 기반한 알고리즘에서는 OpenCV 3.4.1 라이브러리가 사용되었다. 분할 결과는 Scikit-learn 0.24.1 프레임워크를 사용하여 검증하였고, 정량적 측정 결과에 대한 통계 분석은 MedCalc version 19.6.1(MedCalc Software, Ostend, Belgium) 프로그램으로 수행하였다.

2.3 이미지 전처리

본 실험에서 사용된 후두내시경의 원본 이미지는 성대 주위 조직 및 불필요한 영역을 포함한다. 또한 관찰하고자 하는 대상이 성대 점막의 윤곽이므로, 이미지를 성대 점막의 사이 공간인 성문(glottis) 영역만 관찰할 수 있는 사각형 형태로 크롭(Crop)하였다(Fig. 1의 (a)). 후두내시경 영상의 노이즈(Noise)를 감소시키기 위해, 현재 픽셀(Pixel) 값과 주변 이웃 픽셀 값들의 가중 평균(Weighted average)을 이용해서 현재 픽셀의 값을 대체하는 가우시안 필터(Gaussian filter)가 적용되었다(Fig. 1의 (b)). 이후 딥러닝을 사용한 분할에서는 별도로 영상의 대비 및 선명도를 높이기 위해 대비 제한 적응 히스토그램 평활화(Contrast-limited adaptive histogram equalization, CLAHE)를 시행하였고[13](Fig. 1의 (c)), 이미지를 256×256 크기로 변환(Resize)하였다. 최종적으로 픽셀 값의 범위는 0-1로 스케일링(Scaling)되었다.

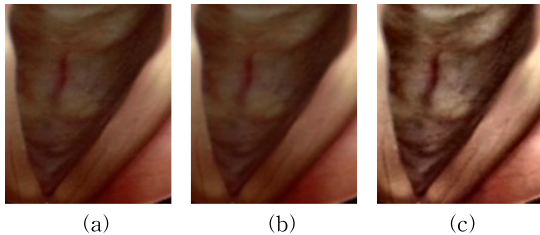


Fig. 1. Pre-processing process in cropped image. (a) Cropped image, (b) Gaussian filtered image, and (c) CLAHE image.

2.4 딥러닝 기반의 성문 분할

본 실험 중 딥러닝을 사용한 방법에서는 ResNet50을 기본 구조로 갖는 Deeplab V3 아키텍처를 사용하여 성문 분할을 수행하였다. Deeplab V3은 필터가 수용하는 픽셀에 간격을 둔 Atrous convolution을 사용한다[14]. 이로 인해, 파라미터 수와 계산량을 동일하게 유지하면서, 다양한 크기의 영역에서 특징 추출이 가능하다는 장점을 갖는 딥러닝 분할 알고리즘이다[15]. Deeplab V3의 백본은 ResNet50[16]으로 구성되었으며, Atrous convolution을 병렬 구조로 적용하는 인코더(Encoder), 이를 다시 합쳐주는 ASPP(Atrous Spatial Pyramid Pooling), Up-sampling을 하는 디코더(Decoder)로 구성되었다. Fig. 2는 Deeplab V3의 기본 구조를 보여준다.

딥러닝 모델의 학습은 배치 사이즈(batch size) 4, 학습 에폭(Epochs) 300 및 학습률(Learning rate) 0.0001을 설정하여 시행되었으며, 최적화 기법으로는 Adam[17]을 사용하였다. 적은 데이터 수를 극복하고 과적합(Overfitting)을 방지하기 위해 실시간 데이터 증강(Online data augmentation)을 시행하였고, 데이터셋을 5개의 부분집합으로 겹치지 않게 분할한 후, 모델 학습 시 순차적으로 각 부분집합이 테스트 데이터로 사용되었으며, 랜덤하게 이동 변환 $[-0.125, 125]$, 크기 변환 $[0.5, 1.5]$, 회전 변환 $[-22.5^\circ, 22.5^\circ]$ 이 적용되었다. 10번 반복하는 동안 손실 함수(Loss function) 값이 감소하지 않으면 학습율을 0.5배 감소시켰으며, 20번 반복하는 동안 손실 함수 값이 감소하지 않으면 학습을 조기 종료하였다. 이후 후처리 과정으로써, 분할 결과에 5×5 중간값 필터

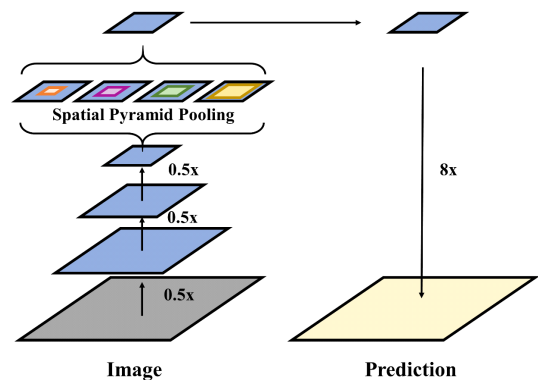


Fig. 2. Architecture of Deeplab V3 model used for the glottis segmentation.

(Median filter)를 적용하였다.

2.5 디지털 영상처리 기반의 성문 분할

전처리된 이미지에서 성문 영역을 분할하기 위해, 노이즈를 제거하고 관심 영역의 경계선을 검출하는 캐니 엣지 검출 알고리즘[18]을 사용하였다(Fig. 3의 (b)). 그 결과, 일부 경계선의 끊김 현상이 나타나 모폴로지(Morphology) 연산 중 팽창(dilation) 연산을 적용하여 경계선 검출을 완성하였고, 지나치게 두꺼워진 경계선은 침식(erosion) 연산을 적용하여 두께를 조절하였다(Fig. 3의 (c)). 이때 모폴로지는 컴퓨터 비전 분야에서 노이즈를 제거하거나 영역간 경계를 허무는 형태학적 연산의 수행을 의미한다[19].

성문 영역의 완전한 분할을 위해서는 경계선의 내부 영역이 필요하므로, 경계선을 검출하는 알고리즘을 응용하여 내부 영역 전체를 탐지하고 해당 영역의 픽셀 값은 모두 255로 변경하였다. 이후에도 나타나는 갈라짐 현상은 모폴로지 연산의 닫힘(closing) 연산으로 해결하였다(Fig. 3의 (d)).

2.6 성대 점막의 형태적 비대칭 분석

본 연구에서는 성대 점막의 형태적 비대칭 정도를 분석하기 위하여 분할한 성문의 좌우 너비를 측정하였고, 이때 좌우 너비는 성문의 중심축과 직교하는 선분의 좌우 길이와 같다. 중심축은 성문의 무게중심 점과 맨 끝 변곡점의 좌표 값을 찾아 두 지점을 이어서 구하였다(Fig. 4의 (b)). 중심축의 기울기는 후두 내시경 영상의 촬영 각도에 따라 달라질 수 있으므로, 이미지의 중앙을 기준으로 반 시계 방향 회전하는 알고리즘을 적용하여 중심축의 기울기가 수직이 되도록 조정하였다(Fig. 4의 (c)). 이후, 축에 직교하면서 성문 영역 내에 존재하는 선분을 좌우로 각각 9개씩 일정한 간격으로 설정하였다(Fig. 4의 (d)).

설정된 선분 9쌍의 길이, 즉 성문의 좌우 너비를 측정하고, 9쌍 전부 좌우 너비 간의 차이를 확인하였다. 9개의 값 중 특정 값이 평균과의 편차가 클수록 표준편차 값은 커지고, 성대 점막에서 해당 영역의 형태적 비대칭은 심해진다. 한편 전반적으로 편차가 고를수록 표준편차의 값은 작아지고, 성대 점막은 형태적으로 대칭에 가까워진다. 본 논문에서는 정상군

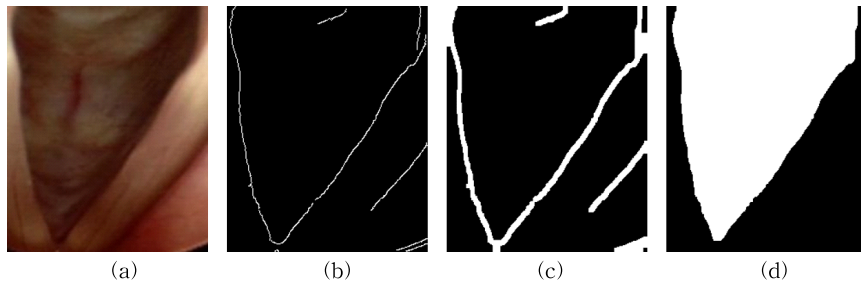


Fig. 3. The process of segmentation of the glottis using DIP model, (a) raw image, (b) Image with contours detected, (c) Morphological operated image, and (d) Segmented image.

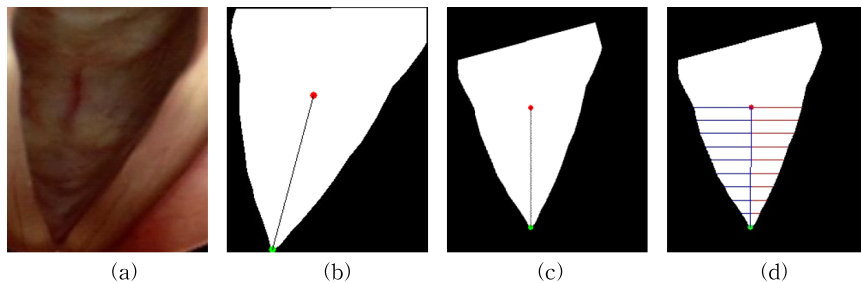


Fig. 4. The process of measuring widths of the glottis based on Ground truth, (a) raw image, (b) Setting central axis of the glottis, (c) Rotation the central axis to perpendicular, and (d) Setting 9 pairs of orthogonal line to the central axis.

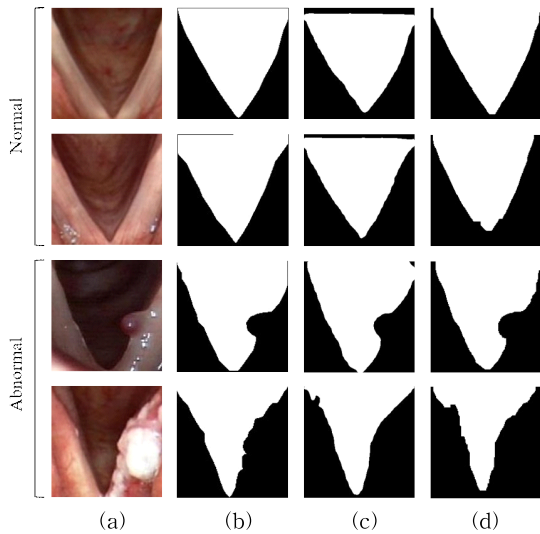


Fig. 5. Result of segmentation of the glottis in the laryngoscope image, (a) Cropped image, (b) Ground truth, (c) Result of DL segmentation, and (d) Result of DIP segmentation.

과 비정상군의 편차 분포를 비교하여 분포도 간의 차이가 있는지 확인하기 위해 각 데이터의 평균값과 표준편차를 구하였다. 딥러닝 기법의 분할 결과와 디지털 영상처리 기법의 분할 결과 모두 앞서 제시한 방식을 동일하게 적용하여 분석하였다.

3. 결 과

딥러닝 방법과 디지털 영상처리 방법의 분할 결과는 Fig. 5와 같다. 딥러닝 기법으로 분할한 결과는 5-fold 교차 검증(5-fold cross validation)을 통해 모든 데이터가 최소 한번씩 평가되도록 하였다[20,21]. 이후 딥러닝과 디지털 영상처리 실험 각각의 분할 결과는 정확도(Accuracy), 민감도(Sensitivity), 특이도(Specificity), 다이스 계수(Dice similarity coefficient, DSC)의 4가지 척도에 기반하여 Ground Truth와 비교하였다. 오차 행렬(confusion matrix) 지표의 구성요소인 TP(True Positive), FP(False Positive), TN(True Negative), FN(False Negative)

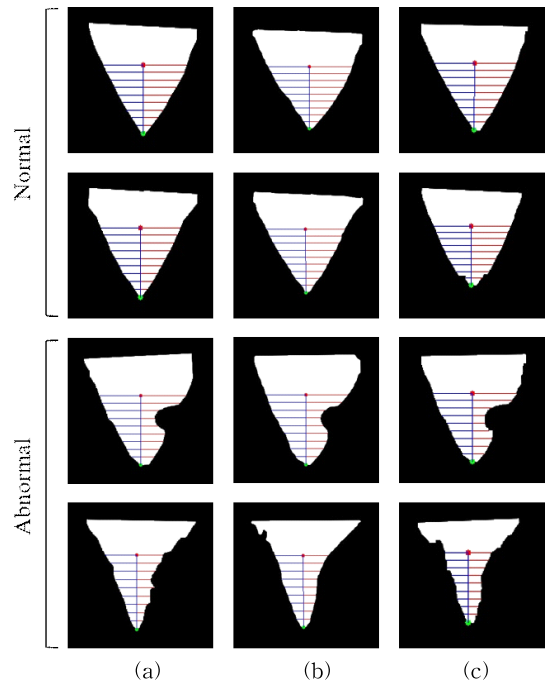


Fig. 6. Result of width measurement of segmented glottis, (a) Result of measurement from Ground truth, (b) Result of measurement from DL segmentation, and (c) Result of measurement from DIP segmentation.

을 구하고 식 (1)~(4)에 따라 4가지 척도를 계산하여 최종적인 결과는 계산한 값의 백분율로 나타내었다. 그 결과는 Table 1에 제시하였다.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Specificity} = \frac{TN}{FP + TN} \quad (3)$$

$$\text{DSC} = 2 \times \frac{TP}{(TP + FP) + (TP + FN)} \quad (4)$$

딥러닝 실험의 평균 정확도는 97.50%, 평균 민감도는 96.96%, 평균 특이도는 97.98%, 평균 다이스 계

Table 1. Result of segmentation of the glottis using DL and DIP model.

	Accuracy (%)	Sensitivity (%)	Specificity (%)	DSC (%)
Deep Learning	97.50	96.96	97.98	96.95
Digital Image Processing	94.70	91.02	98.59	94.72

수는 96.95%, 디지털 영상처리 실험의 평균 정확도는 94.70%, 평균 민감도는 91.02%, 평균 특이도는 98.59%, 평균 다이스 계수는 94.72%로 나타났다.

두 실험의 분할 결과에서 측정된 성문 좌우 너비는 Fig. 6과 같다. 정량화한 성대 접막의 형태적 비대칭 정도는 상관 분석(Correlation Analysis) 및 Bland-Altman Plot을 사용하여 각각 Ground truth의 값과 비교 분석하였고, 데이터의 정상군과 비정상군의 측정 결과를 비교하기 위해 두 실험 모두 T-test를 수행하였다. 그 결과 정상군과 비정상군의 비대칭 정도는 유의한 차이가 있었고, 딥러닝 실험($r=0.8512$, $p<0.0001$), 디지털 영상처리 실험($r=0.7784$, $p<0.0001$) 모두 양의 상관관계를 보였으며(Fig. 7), 오차의 평균에서 ± 1.96 표준편차 내에 위치하는 값은 총 40건의 데이터 중 딥러닝 실험 39건, 디지털 영상처리 실험 38건이었다(Fig. 8).

4. 고 찰

본 연구에서는 딥러닝과 디지털 영상처리의 두 방법을 사용하여 후두내시경 영상에서 성문 영역을 분할하고, 각각의 분할 영역에서 좌우 너비 및 성대 접막의 비대칭 정도를 측정하여 비교하였다. 비대칭 정도는 9쌍의 좌우 선분 간 차이를 구하고, 그 차이 값들의 표준편차로 나타내었다.

디지털 영상처리에 기반한 분할 방식으로, 김진필 등은 경계선을 검출하는 라플라시안, 소벨, 캐니 엣지 검출 알고리즘 등을 제시한 바 있다[9]. 그 중에서 가장 정확도가 높은 캐니 엣지 검출 알고리즘이 본

실험에 사용되었다. 본 연구에서 비교 분석한 딥러닝과 디지털 영상처리 기법은 추후 조용현 등[10]이 제시한 분할 방법과도 비교 가능하다.

후두 성대 영역에 대한 평균 분할 정확도는 딥러닝 기법 및 디지털 영상처리 기법 각각 97.5%, 94.7%로, 딥러닝의 분할 성능이 상대적으로 더 우수하였다. 또한 비대칭 정도의 정량화 결과를 Ground truth와 비교했을 때 두 실험 모두 높은 상관관계를 보였다. 그럼에도 불구하고, Bland Altman-plot 분석에서 딥러닝 기법 및 디지털 영상처리 기법 모두 일부 데이터가 95% 신뢰 구간에 포함되지 않는 문제를 보였다. 해당 데이터는 딥러닝 기법 및 디지털 영상처리 기법 각각 정확도 97.96%, 정확도 94.90%로 모두 높은 분할 정확도를 보였으나, 일부 영역에서 미세한 분할 오류가 발생하였고, Ground truth의 비대칭 측정 과정에서 이미지의 회전에 오류가 발생하여 두 실험 결과와 큰 차이를 보인 것으로 확인된다. 이는 추가 연구 시 다양한 데이터를 대규모로 투입하여 측정 알고리즘의 정확도를 향상시킨다면 해결 가능할 것으로 판단된다. 본 연구에 사용된 데이터 수는 알고리즘을 학습시키고 성능을 평가하기에 충분하지 않다는 제한점이 있고, 데이터의 규모는 연구 결과의 신뢰성에도 중요한 역할을 하므로, 추후 이를 보완한 후속 연구가 필요하다.

이 외에도, 분할 이미지에서 측정된 성문의 너비가 이미지의 픽셀 수를 의미한다는 제한점이 있다. 임상 현장에서 활용할 경우 측정값의 오차 범위를 고려해야 한다는 문제가 대두될 수 있다. 본 논문의 최종 목적은 대상의 형태적 정보를 정량적으로 제시

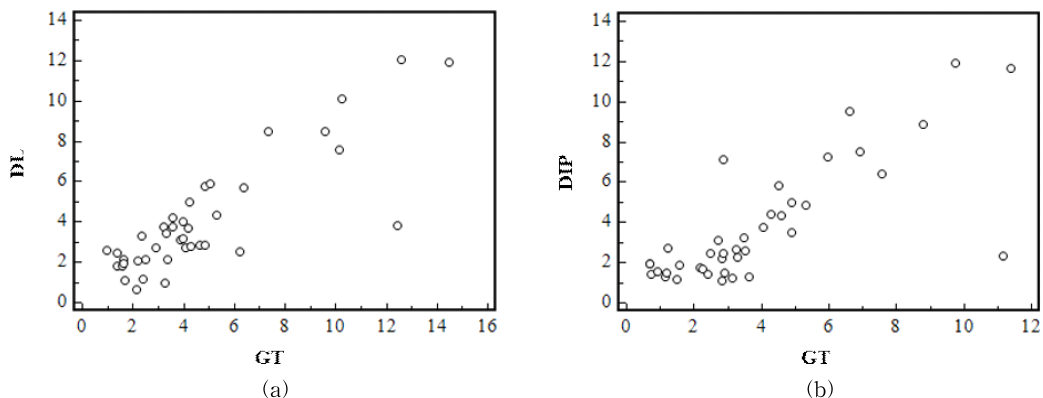


Fig. 7. Scatter plots of values of quantifying asymmetry comparing DL results and DIP results. (a) Comparison of DL results and Ground truth and (b) Comparison of DIP results and Ground truth.

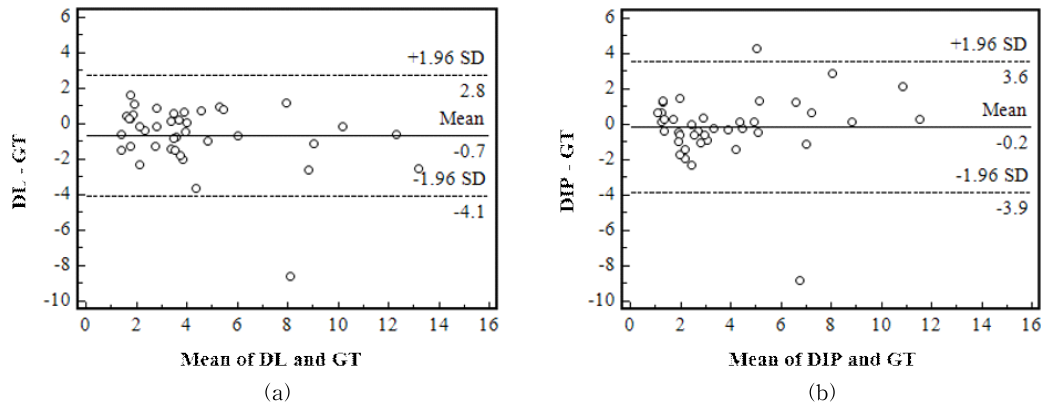


Fig. 8. Bland–Altman plots of values of quantifying asymmetry comparing DL results and DIP results. (a) Comparison of DL results and Ground truth and (b) Comparison of DIP results and Ground truth.

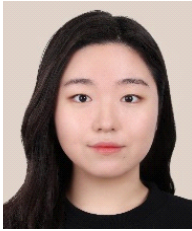
하는 데에 있으므로 정량화 결과의 후처리를 별도로 수행하지 않았으나, 임상 적합성을 높이기 위해서는 정량화한 형태적 정보를 데이터 스케일링(Data Scaling) 처리한 후속 연구가 필요하다.

본 연구에서 제시한 방법은 향후 모든 후두내시경 검사에 영상 분할 및 정량적 평가 방식으로 응용이 가능하다. 딥러닝의 우수한 분할 가능성을 확인하였으므로 추후 학습 데이터를 양적, 질적인 측면에서 보완한다면 더욱 정교한 이미지 분할 뿐만 아니라 후두내시경 영상에서의 즉각적인 분할 및 정량화가 가능하다. 이는 후두 스트로보스코피, 초고속 후두내시경 영상 등 기존 성대 진동 검사의 한계를 극복하여 객관적이고 효율적인 평가 방법으로써 상용될 것으로 예상된다. 본 연구의 결과는 컴퓨터 보조 진단 시스템의 연구개발 및 다양한 후두 질환의 진단 과정에 도움이 될 것이다.

REFERENCE

- [1] S.H. Lee, “Diagnostic Role of Stroboscopy,” *Journal of the Korean Society of Laryngology, Phoniatrics and Logopedics*, Vol. 21, No. 1, pp. 13–16, 2010.
- [2] T.K. Kwon and M.S. Kim, “Current Standpoints on the Pathophysiology of Benign Vocal Fold Lesions,” *Journal of the Korean Society of Laryngology, Phoniatrics and Logopedics*, Vol. 26, No. 2, pp. 91–93, 2015.
- [3] S.G. Wang and H.J. Park, “Laryngeal Stroboscopy,” *Journal of Clinical Otolaryngology Head and Neck Surgery*, Vol. 26, No. 1, pp. 3–9, 2015.
- [4] W.J. Cha, S.G. Wang, J.Y. Jang, G.H. Kim, and Y.W. Lee, “Post-Processing of High-Speed Video-Laryngoscopic Images to Two-Dimensional Scanning Digital Kymographic Images,” *Journal of the Korean Society of Laryngology, Phoniatrics and Logopedics*, Vol. 28, No. 2, pp. 89–95, 2017.
- [5] J.C. Lee and I.H. Bae, “Recent Advances in Examination of Vocal Fold Vibration,” *Journal of the Korean Society of Laryngology, Phoniatrics and Logopedics*, Vol. 32, No. 1, pp. 1–8, 2021.
- [6] S.H. Choi and C.H. Choi, “Qualitative Analysis of Vocal Fold Vibration in Normal Adults with a High-Speed Video System,” *Communication Sciences and Disorders*, Vol. 17, No. 4, pp. 591–602, 2012.
- [7] M. Kunduk, M. Doellinger, A.J. McWhorter, and J. Lohscheller, “Assessment of the Variability of Vocal Fold Dynamics within and between Recordings with High Speed Imaging and by Phonovibrogram,” *The Laryngoscope*, Vol. 120, No. 5, pp. 981–987, 2010.
- [8] Y.J. Kim and K.G. Kim, “Development of an Optimized Deep Learning Model for Medical

- Imaging,” *Journal of the Korean Society of Radiology*, Vol. 81, No. 6, pp. 1274–1289, 2020.
- [9] J.P. Kim and Y.B. Cho, “Image Processing of Lesion in Iris Image Using OpenCV,” *Journal of Digital Contents Society*, Vol. 19, No. 11, pp. 2035–2040, 2018.
- [10] Y.H. Cho, “An Efficient Lesion Segmentation by Using ROI Extraction and Combination of Features in Digital Skin Images,” *Journal of Korean Institute of Intelligent Systems*, Vol. 30, No. 6, pp. 472–481, 2020.
- [11] B.S. Kim and I.H. Lee, “Retinal Blood Vessel Segmentation using Deep Learning,” *The Journal of Korean Institute of Information Technology*, Vol. 17, No. 5, pp. 77–82, 2019.
- [12] G.Y. Kwon, J.H. Kim, Y.J. Kim, S.M. Lee, and K.G. Kim, “Detecting Boundary of Erythema Using Deep Learning,” *Journal of Korea Multimedia Society*, Vol. 24, No. 11, pp. 1492–1499, 2021.
- [13] K. Zuiderveld, “Contrast Limited Adaptive Histogram Equalization,” *Graphics Gems*, pp. 474–485, 1994.
- [14] Y. Wang, S. Sun, J. Yu, and D. Yu, “Skin Lesion Segmentation using Atrous Convolution via DeepLab V3,” *arXiv Preprint*, arXiv:1807.08891, 2018.
- [15] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, “Rethinking Atrous Convolution for Semantic Image Segmentation,” *arXiv Preprint*, arXiv:1706.05587, 2017.
- [16] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.
- [17] D.P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” *arXiv Preprint*, arXiv:1412.6980, 2014.
- [18] J. Canny, “A Computational Approach to Edge Detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, No. 6, pp. 679–698, 1986.
- [19] R.M. Haralick, S.R. Sternberg, and X. Zhuang, “Image Analysis using Mathematical Morphology,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, No. 4, pp. 532–550, 1987.
- [20] R. Kohavi, “A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection,” *International Joint Conferences on Artificial Intelligence Organization*, Vol. 14, No. 2, pp. 1137–1145, 1995.
- [21] S. Subramaniam, K. Jayanthi, C. Rajasekaran, and R. Kuchelar, “Deep Learning Architectures for Medical Image Segmentation,” *IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 579–584, 2020.



이 선 민

2020년 3월~현재 가천대학교 간
호학과 재학
관심분야: 의료영상, 의료정보, 빅
데이터, 인공지능



우 주 현

2006년 이비인후과 전문의
2008년 가천대 이비인후과학교실
전임강사
2010년 가천대 이비인후과학 박사
2014년 위스콘신메디슨 연수
2020년~현재 가천대 이비인후과
학교실 교수, 대한후두음
성언어의학회 편집이사

관심분야: 후두질환, 음성질환, 두경부암



오 석

2014년 3월~2020년 2월 울산대
학교 공과대학 전기공학부
(학사)
2020년 3월~2021년 9월 울산대
학교 의과대학 의과학과
(석사 과정 중퇴)

2021년 6월~현재 가천대 길병원 의료기기 R&D센터 연
구원

관심분야: 생체신호, 의료영상, 의료전자기기



김 광 기

2006년 2월 서울대학교 공과대학
박사
2007년 7월 국립암센터 선임/책임
연구원
2017년 2월 가천대학교 의공학과
교수-현재

2017년 2월 가천대길병원 의료기기 R&D센터 센터장, 의
용생체공학회/대한의료정보학회/대한의료로봇
학회/대한혁신의료기술학회 이사로 활동중임

관심분야: 의료기기, 의료로봇, 의료정보, 인공지능



김 영 재

2013년 2월 광운대학교 플라즈마
바이오디스플레이학과
(석사)
2018년 2월 광운대학교 플라즈마
바이오디스플레이학과
(박사)

2008년 1월~2017년 4월 국립암센터 의공학연구과
2019년 9월~현재 가천대학교 의용생체공학과 조교수
관심분야: 의료영상, 인공지능, 빅데이터, 컴퓨터보조진단