

# 설진 영상의 혀 영역 자동 분할을 위한 U-Net 기반 딥러닝 알고리즘

김민솔·서진우·남동현<sup>1\*</sup>

상지대학교 한의과대학 한방신경정신과, 1: 상지대학교 한의과대학 진단·생기능의학교실

## A Deep Learning Algorithm Based on U-Net for Automatic Tongue Region Segmentation from Images Acquired through a Computerized Tongue Image Acquisition System

Minsol Kim, Jinwoo Suh, Donghyun Nam<sup>1\*</sup>

Department of Korean Neuropsychiatry, College of Korean Medicine, Sangji University,  
1: Department of Biofunctional Medicine and Diagnosis, College of Korean Medicine, Sangji University

Research on improving the accuracy of tongue region segmentation in the deep learning process of Korean medicine's tongue diagnosis is actively ongoing. This study aims to propose a segmentation model based on the U-Net using Convolutional Neural Networks (cNN). Binary mask images representing the tongue region were created, and the dataset was randomly divided into an 8:2 ratio for the train and test set. To prevent overfitting, the changes in validation loss were monitored at each epoch. Finally, Dice similarity coefficient (DSC) and Jaccard index (JI) were computed to measure the accuracy of the segmentation model in identifying the target regions. The modified U-Net network showed a significant level of accuracy in tongue region segmentation and minimal overfitting despite a small number of training data sets. The accuracy in the train set reached 0.997 and in the test set the accuracy was 0.993. The Dice similarity coefficient (DSC) score on the test dataset was  $0.981 \pm 0.017$  and the Jaccard index (JI) score was  $0.964 \pm 0.031$ . The proposed model based on U-Net for tongue region extraction is anticipated to be effective for practical applications such as computerized tongue analysis systems.

keywords : Tongue image segmentation, U-Net, Tongue diagnosis, Tongue region segmentation, Deep learning, Convolutional neural network

### 서론

한의학에서 설진은 망진(望診), 문진(聞診), 문진(問診), 절진(切診)의 사진(四診) 중 망진에 해당하는 진단법으로, 육안으로 혀의 색, 모양, 설태(舌苔) 및 설열(舌裂) 등을 관찰하여 장부의 건강 상태를 알아내고 한의학적 변증에 판단 근거를 제공하는 중요한 진단 방법 중 하나이다.

설진은 외부에서 용이하게 관찰할 수 있는 점막 부위인 혀를 관찰하는 정교하고 섬세한 진단 방법으로, 정확한 설진을 위해서는 혀와 설태의 색상, 설태의 두께, 혀의 갈라짐 정도, 건조 및 습윤 정도 등의 다양한 정보를 종합적으로 관찰해야 한다. 그러나 현대 한의학에서 설진은 의사가 육안으로 혀를 관찰한다는 한계로 인해 시간과 조명에 따라 변화하는 혀의 색과 설태 등을 정량화하기 어

렵다. 또한, 의사의 주관적인 판단이 개입되어 객관적인 데이터를 얻기 쉽지 않다는 문제가 있다. 이를 극복하기 위해서 국내외 연구자들이 디지털 혀 영상장치가 내장된 설진기를 개발하고 있으며, 이 과정에서 촬영된 혀 이미지에서 혀 영역을 정확히 분리시키는 작업은 가장 중요한 초기 과정 중 하나로 간주된다. 이와 같이 영상의 특정 구조 또는 관심 영역을 구분하여 분석 가능하게 만드는 것을 영상 분할(segmentation)이라고 지칭하며, 이는 의료 영상 내의 특정 구조나 병변을 식별하기 위해 필수적인 요소이다.

전통적인 이미지 처리 기법을 활용한 혀 분할한 방법으로는, 밝기와 명암대비, 그리고 경계선 추출 등을 기반으로 수동적으로 이미지 제거하거나 영역 확장 기법(Region Growing)을 사용하는 방식이 주로 사용되어 왔다. 히스토그램, 휘도 기반의 질감 정보, 색조-채도(HIS) 색상 모델, 임계값 계산 등을 이용하여 효율적인

\* Corresponding author

Donghyun Nam, Department of Biofunctional Medicine & Diagnostics, College of Korean Medicine, Sangji University, 83 Sangjidae-gil Wonju-si Gangwon-do, Republic of Korean, #26339

E-mail : omdnam@naver.com · Tel : +82-33-738-7504

Received : 2024/08/20 · Revised : 2024/12/10 · Accepted : 2024/12/24

© The Society of Pathology in Korean Medicine, The Physiological Society of Korean Medicine

pISSN 1738-7698 eISSN 2288-2529 http://dx.doi.org/10.15188/kjopp.2024.12.38.6.259

Available online at https://kmpath.jams.or.kr

혀 영역 분할을 한 연구들이 그 예시이다.<sup>1-5)</sup> 이후 디지털 템플릿 등을 이용해 학습 기반 접근법을 시도하거나,<sup>6)</sup> 소벨 마스크(Sobel Mask), Snake 알고리즘<sup>7,8)</sup>, Grab cut<sup>9)</sup> 등을 이용한 연구들이 등장했다. 그러나 전통적인 이미지 처리 기법은 사용자인 한의사가 알고리즘의 도움을 받아 혀 이미지에서 직접 혀 영역을 추출해야 하는 번거로움이 있다.

최근 급속도로 발전한 딥러닝 기술은 대량의 데이터를 학습하여 영상 내 복잡한 패턴을 자동으로 추출하여 자동화와 높은 정확도를 달성한다는 장점으로 많은 분야에 큰 기여를 하고 있다. 의료 영상 분석에서도 전통적인 이미지 처리 기법의 수작업 문제를 해결하며 시간과 비용을 감소시키면서도 높은 정확도로 인하여 딥러닝을 활용한 다양한 연구들이 진행 중에 있다. 특히, 딥러닝의 한 유형인 합성곱 신경망(Convolutional neural network, CNN)은 합성곱 계층과 필터를 이용하여 이미지의 특징을 추출하는데 탁월하다. 이를 기반으로 한 이미지 분석 모델에는 LeNet, Alexnet, VGGNet, ResNet, SegNet, U-Net 등이 있으며, 혀 영상 이미지 분석에서도 앞서 말한 딥러닝 모델들을 활용하여 혀 영상을 분할하는 시도들이 보고되고 있다.<sup>10-13)</sup>

합성곱 신경망을 기반으로 한 모델인 U-net<sup>14)</sup>은 Segmentation을 위해 특별히 설계된 모델로서, 세포, 혈관 등과 같은 작은 이미지의 정확한 분할에 탁월하여 현재 널리 사용되고 있다.<sup>15)</sup> U자의 대칭적인 구조를 통해 인코딩 경로에서 추출된 이미지를 디코딩 경로에서 재조합함으로써, 적은 양의 고해상도 데이터로도 높은 정확도를 낼 수 있다는 장점이 있다. 이는 데이터의 희귀성과 높은 정확도가 요구되는 의료 영상 분석에 적합하기 때문에, 현재까지 U-Net은 안저영상<sup>16)</sup>, 폐CT<sup>17,18)</sup>, HIFU 초음파 영상<sup>19)</sup>, 치아 X-ray<sup>20)</sup>, 척추CT<sup>21,22)</sup> 등의 다양한 의료 영상 분할에 이용되고 있다. 혀 영상 이미지 분할에서도, U-net을 응용하여 SpurNet<sup>23)</sup>, IAUNet<sup>24)</sup>, Chen의 TSNet<sup>25)</sup>, Ahn의 VGG16기반의 U-Net모델<sup>26)</sup> 등이 개발되었다. 또한 Li의 병렬 합성곱 계층을 사용한 U-Net연구<sup>27)</sup> 등도 있었다. U-Net은 의료 영상 분석 연구에서 뛰어난 성과를 보이며, 혀 영역 분할에서 높은 정확도를 보여주었다.

그러나 이와 같은 기존 혀 영상 분할 모델들은 탐색적 단계에서 이루어진 시도들로서, 장비에 의존적인 인공지능 모델들의 특성상 설진기와 같은 의료 기기에 직접 적용할 수 없으므로, 실제 임상에서 설진기 영상 분석에 활용하기 어렵다는 단점이 존재한다. 즉 U-Net을 이용하여 혀 영역 분할을 한 선행 연구들은 일반 카메라와 모바일 기기, 공개 데이터 등 다양한 방법으로 수집된 혀 이미지로 데이터셋을 구축하여 높은 정확도를 얻었지만, 이는 장비 의존적인 특성을 고려했을 때 설진기에 그대로 적용할 수 없다는 한계가 있다.

이에 본 연구자는 합성곱 신경망(convolutional neural network, cNN) 모델인 U-Net을 기반으로 하여 설진기를 통해 촬영된 혀 영상으로부터 혀 영역을 추출하는 알고리즘을 개발하고, 해당 알고리즘의 정확도를 평가하였다. 더 나아가 이를 설진기에 적용하여 임상 환경에 적용할 수 있는 가능성을 검토해보고자 한다.

## 연구대상 및 방법

### 1. 연구윤리

이 연구는 2023년 3월 31일 상지대학교 부속한방병원 생명윤리위원회의 승인을 받아(SJIRB-Human-23-001) 동 병원에서 수행되었으며, 임상시험 관리기준과 헬싱키 선언의 원칙이 준수되었다. 모든 연구대상 참여자는 이 연구에 대한 정보를 받았으며, 등록 전 예 서면 동의서를 제공하였다.

### 2. 연구대상

본 연구는 2023년 4월 1일부터 2023년 12월 31일까지 모집된 만 20세 이상의 건강한 성인 121명을 대상으로 수행되었다. 건강한 성인의 선별하는 기준은 최근 1년 이내에 실시한 건강검진 상 고용노동부고시(제2023-8호) 근로자 건강진단 실시기준에서 정하는 질병이 의심되지 않는 경우에 해당하며, 건강검진 이후 새롭게 발생한 지속되는 자각적 또는 객관적 이상 소견이 없는 자로 한정하였다. 다만 50대 이상에서 관리되고 있는 고혈압은 연구대상에 포함하였다. 이 외에도 치아교정기를 거치한 경우와 턱관절 장애가 있어 입을 크게 벌리기 어려운 경우에는 배제되었다.

### 3. 혀 영상의 촬영

연구참여자는 스크리닝과 혀 촬영을 위해 두 차례 방문하였다. 스크리닝 방문시에는 참여 동의서 작성, 인구학적 조사, 병력 확인, 선택 및 제외기준에 대한 평가 후 등록 여부를 결정하였다. 촬영 방문은 오전에 이루어졌으며, 혀 촬영 전에 혀에 대한 자극이나 착색을 피하기 위해, 물을 제외한 음식물의 섭취와 혀에 대한 직접적인 양치질은 금지되었다.

혀 촬영은 사전에 표준 색상차트(ColorChecker Passport, X-rite, USA)를 이용하여 사전에 색상보정이 이루어진 설진기(CTS-2000, 대승의료기기, 대한민국)를 이용하여 이루어졌으며, 혀 옆에 거치 시킨 18% 회색 색편을 함께 촬영하여 영상 획득과정에서 발생할 수 있는 화이트 발란스와 색상 편위를 모니터링하고, 편위가 발생한 경우 추가적으로 교정하였다.

촬영된 혀 영상의 해상도는 4024 X 3036이었고, 이미지 저장 과정에서 데이터 손실을 최소화하기 위해 이미지는 24 비트, PNG(portable network graphics) 포맷으로 저장되었다. 혀 이미지를 저장한 후, 연구참여자에게 이상반응의 발생 유무를 확인하였다.

### 4. 혀 영역 마스크 이미지 제작

각 이미지별로 혀 영역에 대한 ground truth로서, 혀 영역을 나타내는 이진화된 마스크 이미지를 제작하였다. 설진기 전문가는 촬영된 혀 영상 위에 혀 가장자리를 따라 펜 타블렛(CTL-6100WL, Wacom, Korea)을 이용하여 혀 윤곽선을 그린 후, 윤곽선 내부를 흰 색으로 채워 마스크 이미지를 제작하였다. 마스크 이미지를 구성하는 화소의 색상값이 0과 255값만 사용되도록 OpenCV threshold 함수를 이용하여 마스크 이미지는 최종적으로 이진화되었다. 따라서 단일 채널 마스크 이미지에서 혀 영역의 모든 화소 색상값은 255, 배경 영역은 0이었다.

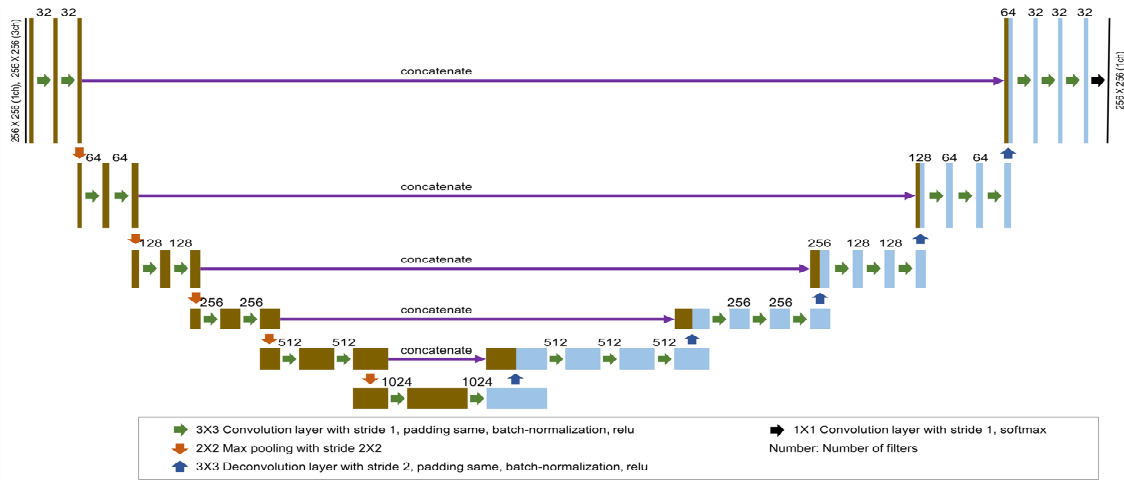


Fig. 1. The structural schematic diagram of the convolutional neural network used in this study.

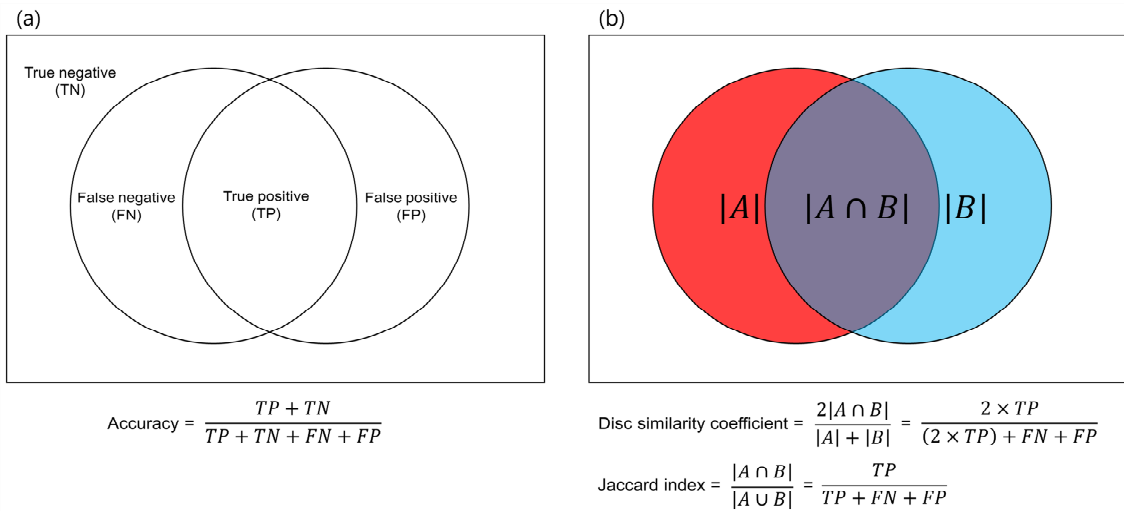


Fig. 2. The concepts of (a) accuracy and (b) Dice similarity coefficient and Jaccard index.

5. 인공지능 학습방법

데이터의 누락 등 데이터 관리상에 오류가 발생하지 않도록 전체 학습과정을 데이터 관리 및 분할, 학습, 성능평가 단계로 나누고, 각 단계별로 일괄 처리로 이루어지도록 code를 작성하였으며, 파일의 결손여부를 확인할 수 있도록 중간에 각 이미지 파일의 경로와 숫자를 확인하였다.

Segmentation 학습을 위해 121장의 허 이미지와 마스크 이미지는 쌍을 이루어 8:2의 비로 훈련용 데이터 셋(96쌍)과 테스트용 데이터 셋(25쌍)으로 무작위 분할되었다. 각각 256 X 256 해상도 크기로 resize된 후, 허 이미지와 마스크 이미지는 효과적인 학습을 위해서는 충분한 훈련용 데이터 셋의 확보가 필요하므로, 훈련용 데이터 셋은 좌우 뒤집기 (tensorflow.image.random\_flip\_left\_right())와 명암 변화 (tensorflow.image.random\_brightness())를 주어 데이터 셋을 증강하였다. 명암 변화시 최대 델타값은 0.1이었으며, 증강된 데이터 셋에 대한 파이프라인을 구성하여 훈련용 데이터 셋은 구성되었다. 데이터 셋은 순서를 섞어서(shuffle) 무작위 순서로 학습이 이루어지도록 하였으며, 이때 buffer size는 1,000

이었다.

학습시 과적합 현상을 줄이기 위해 매 epoch마다 validation loss의 변화를 모니터링하여, 15회 연속 validation loss가 줄어들지 않으면 callback이 작동하여 학습이 중지되도록 하였다. 또 각 epoch마다 training 및 validation에 대한 accuracy와 loss를 모니터링하여, 학습이 끝난 후 학습이 원활하게 이루어졌는지, 과적합이 발생하지는 않았는지, 정확도 향상은 충분히 포착되었는지 여부를 확인하였다.

학습을 위한 base model로는 의료용 이미지 분할에 널리 사용되고 있는 U-Net이 선택되었으며, 학습 성능을 보다 높이기 위해 upsampling 단계에서 convolution layer를 3층으로 구성하였다. 본 연구에 사용된 U-Net 모델의 상세한 구조는 Fig. 1에 도시하였다. 학습을 위한 batch size는 32, 최대 epoch는 500, steps per epoch은 20으로 설정하였으며, 모델 compile시 loss 계산은 sparse categorical crossentropy, optimizer는 Adam이 적용되었다. 허 영역은 하나의 폐곡선 도형이므로, 별도의 mini patch를 구성하지는 않았다.

학습을 통해 얻어진 weight matrix는 별도의 화일로 저장되었으며, 저장된 weight matrix 파일을 다시 불러와서 학습된 분할 모델이 테스트용 혀 이미지에 대하여 혀 영역을 예측하도록 하였다.

#### 6. Segmentation 정확도 평가방법

모델의 정확도는 테스트용 마스크 이미지와 모델이 예측한 혀

영역 사이에 일치하는 정도를 통해 평가하였다. 모델 정확도의 정량적인 평가를 위해 의료용 인공지능 분할 모델 정확도 계산에 널리 사용되는 Dice similarity coefficient (DSC)와 Jaccard index (JI)를 계산하여 제시하였다. 일반적인 정확도 평가지표인 accuracy와 함께 Dice similarity coefficient, Jaccard index를 구하는 식은 Fig. 2에 도시하였다.

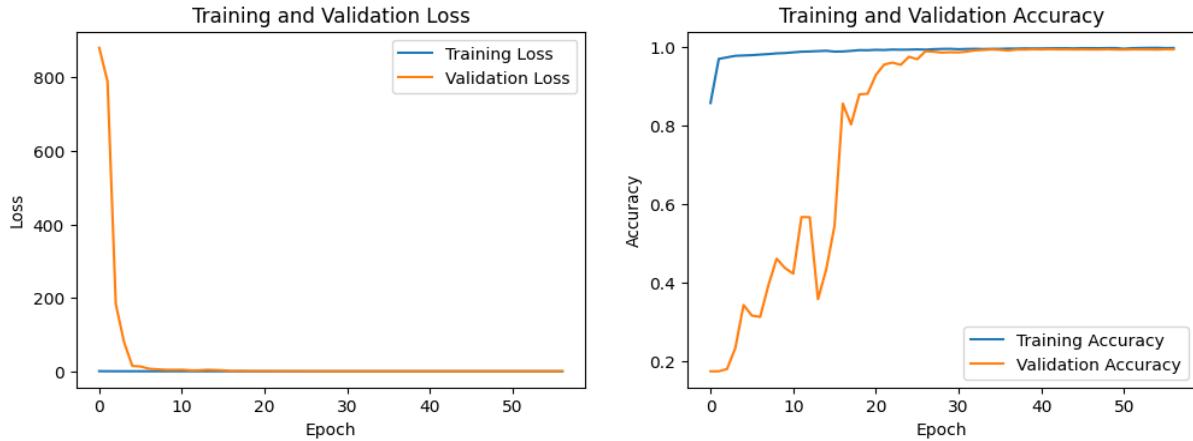


Fig. 3. Changes of loss and accuracy in training and validation according to epoch.

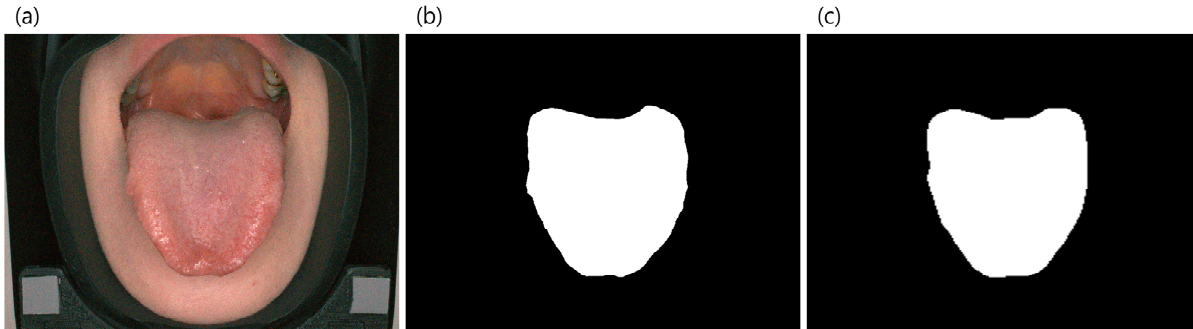


Fig. 4. Tongue area predicted by the developed tongue segmentation model. (a) an input tongue image, (b) a mask image for the input tongue image (as ground truth), and (c) the tongue area image predicted by the developed tongue segmentation model (for easy comparison with the mask image, the image has been resized and presented so that its width-to-height ratio is the same with the mask image). It is confirmed that the mask image (b) and the predicted image (c) are very similar.

## 결 과

### 1. 연구참여자 특성

연구참여자의 인구학적 특성은 Table 1에 제시하였다.

Table 1. Characteristics of the study participants (N=121).

Parameters	Value
Gender (Male : Female)	33 : 88
Age (year)	43.26 ± 14.51
Height (cm)	163.67 ± 8.41
Weight (kg)	62.61 ± 11.61
Systolic blood pressure (mmHg)	120.73 ± 13.15
Diastolic blood pressure (mmHg)	82.43 ± 11.01
Pulse rate (bpm)	72.91 ± 11.06
Body temperature (°C)	36.64 ± 0.26

Data presented as mean ± standard deviation

### 2. 설진기 안전성

모든 연구참여자에서 혀 촬영 후 이상반응은 관찰되지 않았다.

### 3. 학습 결과

57번째 epoch에서 callback이 작동하여 학습은 중지되었다. Callback이 작동한 시점에서 training loss는 0.008, training accuracy는 0.997이었으며, validation loss는 0.026, validation accuracy는 0.993였다. 57번의 epoch가 진행되는 과정에서 정확도의 변화는 Fig. 3에 제시하였다.

Fig. 3을 통해 30번째 epoch 근처까지 validation accuracy가 지속적으로 증가하는 경향이 뚜렷하게 나타남을 확인되었으며, 이후에도 accuracy가 감소하지 않고 포화된 상태에서 조금씩 증가하는 경향이 나타났음을 확인할 수 있었다.

### 4. 모델의 정확도

테스트용 데이터 셋에 포함된 마스크 이미지와 모델이 예측한

혀 영역 이미지 사이에는 높은 일치도가 확인되었다. Dice similarity coefficient는  $0.981 \pm 0.017$ 였고, Jaccard index는  $0.964 \pm 0.031$ 이었다. 테스트용 데이터 셋에서 혀 이미지를 분할한 예시를 Fig. 4에서 확인할 수 있다.

## 고찰

본 연구는 설진기에 적용된 혀 이미지 분석에서 혀 영역 추출의 자동화를 실현하고, 높은 정확도를 달성하기 위해 설계되었다. 설진기는 디지털 혀 영상 촬영장치를 사용하여 설진의 객관화와 과학화를 가능하게 했지만, 임상가가 직접 혀 영역을 추출해야 하는 번거로움이 있었다. 딥러닝을 이용하여 SegNet, FCN, VGGNet 등을 이용한 다양한 최신 혀 영상 분석 연구들이 있었지만 복잡한 고해상도 이미지를 처리하는 데 적합하지 않거나 정확도에서 한계를 드러냈다. U-Net은 이미지 처리가 뛰어나 의료 영상 분석 등에 많이 활용되는 CNN 모델로, U-Net을 활용하여 높은 정확도의 결과를 낸 혀 영역 분할 연구들이 있다. 그러나 기존 연구는 설진기와 같은 의료 기기나 임상 환경에 직접 적용하기에는 한계가 있다. 이러한 한계를 극복하기 위해 본 연구는 cNN 딥러닝 모델을 기반으로 한 U-Net을 선택하여 데이터를 학습시켰다.

연구 결과, training accuracy가 0.997, validation data가 0.993로, 높은 수준의 분할 성능을 보이며 안정적인 학습이 이루어졌음을 확인할 수 있었다. Dice similarity coefficient (DSC) 평가에서는  $0.981 \pm 0.017$ , Jaccard index (JI)에서는  $0.964 \pm 0.031$ 를 달성하였다. 적은 학습 데이터에도 불구하고 DSC 평가지표에서는 FCN, SegNet, DeepTongue, TongueNet 등의 기존 모델에 비해 상대적으로 높거나 열등하지 않은 정확도를 나타냄을 확인할 수 있었다<sup>20</sup>. DSC는 두 집단의 유사성을 측정하는 통계적 지표로, 컴퓨터 비전이나 의료 영상 분석에서 객체의 영역을 비교하는데 유용하다. 이들 지표를 통해 본 연구 모델이 높은 정확도로 설진기에 실용화될 수 있음을 알 수 있다. 또 본 연구에서는 윤곽 추출을 통한 내부 고립 영역(소위 islands나 lakes) 제거와 같이 분할 정확도를 높이기 위한 후처리 과정을 적용하지 않고도 높은 분할 정확도를 나타냄으로써 수정된 U-Net 모델이 설진기 혀 영역 추출에 유용함을 확인할 수 있었다.

본 연구는 다음과 같은 몇 가지 중요한 시사점을 제공한다.

첫째, 의료 영상 데이터의 경우 높은 정확도 및 신뢰도가 요구되지만 다른 분야에 비해 학습에 사용될 충분한 데이터를 확보하기 어렵다는 특징이 있다. 이로 인해 학습에 앞서 데이터 증강이 필수적이다. 본 연구에서는 훈련용 데이터 셋을 좌우 뒤집기와 명암 변화를 주어 데이터를 증강했다. 설진기의 디지털 혀 영상 촬영장치를 통해 조명 조건과 혀의 위치, 촬영 각도 등 표준화된 촬영조건에서 얻어진 혀 영상 데이터를 사용하였기에 적은 수의 학습용 데이터를 사용했음에도 학습은 효과적으로 이루어졌으며 뚜렷한 과적합 현상없이 높은 정확도를 얻을 수 있었다. 또한, 제안된 모델은 설진기로 찍은 영상 데이터로 학습하였기에 설진기에 바로 응용할 수 있는 가능성을 제공한다는 장점이 있다.

둘째, U-Net의 인코더 및 디코더 초기 계층은 각각 4개의 블

록으로 구성되어 있다. U-Net은 U자의 대칭 구조로 인코딩 과정에서 추출한 이미지의 특징을 디코딩 과정에서 재조합하여 정확도를 높인다. 본 연구에 사용된 U-Net 모델의 경우, up-sampling 과정에서 convolution layer를 하나 더 추가하여 혀 영역 검출의 효과를 제고하였다.

셋째, 본 연구에 사용된 혀 mask 이미지가 설진 전문가에 의해 고해상도로 정교하게 제작되었다. 이로 인해 무작위 학습 이후 추가적인 후처리 공정이 없었으며, 이는 연구 모델에 사용될 데이터 셋의 수를 증가시킬 경우, 더 높은 정확도를 보여줄 가능성을 시사한다.

본 연구에서 혀 영역 분할 알고리즘이 상당한 정확도를 보였음에도 불구하고 몇몇 제한점이 있다. 먼저, 학습에 사용된 혀 영상 데이터 셋의 크기가 121장으로 충분하지 않다. 설진에서 사용되는 임상지표들이 혀의 색, 모양, 형태의 상태 등과 같이 매우 정확한 분석을 요구한다는 점과 혀 영역의 분할 성능이 모든 분석의 첫 단계가 된다는 점을 고려했을 때, 혀 영역 분할 성능은 개선은 최종 분석결과의 정확도 개선과 직접적으로 연결된다. 따라서 혀 영역 분할 알고리즘의 정확도를 보다 개선시키기 위한 노력이 지속적으로 요구되며, 이를 위해서는 더 많은 학습용 데이터 셋의 확보가 필요하다.

둘째, 다양한 환자와 환경에서 설진기를 이용하여 획득된 데이터 확보가 필요하다. 본 연구에서는 단일 기관에서 건강한 성인의 혀 영상만 이용되었다. 본 연구를 통해 얻어진 혀 영역 분할 알고리즘이 혀에서 관찰되는 다양한 변이를 포함하는 혀 영상에서도 얼마나 효과적으로 작동하는지 여부는 불명확하다. 인공지능 알고리즘은 학습된 데이터에 기반하여 작동하기 때문에, 특이한 혀 영상이 입력되었을 경우에도 효과적으로 작동하는 강건한 모델을 만들기 위해서는 보다 다양한 상태의 설진기로 촬영된 혀 영상을 학습해야 한다. 적은 학습용 영상으로도 상당한 정확도로 혀 영역을 분할했다는 것은 알고리즘 모델링에 사용된 혀 영상들 사이에 큰 변이가 없었음과 혀 영상과 마스크 영상이 잘 제어된 조건(같은 설진기, 같은 검사자, 표준화된 운영지침, 같은 기준으로 제작된 마스크 영상 등)에서 얻어졌음을 시사하기도 한다.

더 강건하고 높은 정확도의 분할 모델로 개발하기 위해서는 학습된 데이터 셋의 수가 더 많이 필요하다. 따라서 본 연구팀은 모델의 정확도를 더 개선하기 위해 추가 데이터 셋을 확보하고 학습시킬 예정이다. 다양하고 많은 데이터들로 연구 모델이 보다 향상된 정확도와 일반화 기능을 정교하게 갖추게 되면 임상 의료용으로 활용될 가능성이 높아질 것이다. 또한, 향후 기존의 혀 영역 분할 모델들과 본 연구 모델의 성능을 정량적으로 비교하는 연구를 시행할 계획이다.

종합적으로 볼 때, 본 모델은 수정된 U-Net을 활용해 혀 영역을 효과적으로 segmentation 할 수 있음을 확인했다. 또한, 본 모델은 설진기로 촬영된 혀 영상 데이터에 특화되어 있어 설진기의 의료 영상 분석 목적으로 사용하기에 적합할 것으로 보인다. 연구 모델을 설진기에 도입함으로써 임상 한의사가 직접 혀 영역을 검출해야 하는 불편함을 크게 개선할 수 있을 것으로 기대된다. 또한, 높은 정확도를 가진 본 모델이 향후 다양한 연구에 이용된다면 한

의학 설진의 객관화 및 과학화에 크게 기여할 수 있을 것으로 생각된다.

## 결 론

U-Net을 베이스 모델로 하여 설진기로 촬영한 혀 이미지의 혀 영역을 추출하는 알고리즘을 개발하였다. 개발된 모델의 정확도는, Dice similarity coefficient 기준으로 측정하였을 때  $0.981 \pm 0.017$ 인 것으로 나타나 실용화가 가능한 수준의 높은 정확도를 달성하였음을 확인하였다.

## 감사의 글

이 연구는 한국연구재단 바이오·의료기술개발사업(NRF-2022M3A9E4017033) 지원에 의하여 수행되었습니다.

## References

- Kim KH, Do JH, Ryu H, Kim YS. Tongue diagnosis method for extraction of effective region and classification of tongue coating. In: Proceedings of the 1st Workshops on Image Processing Theory, Tools and Applications. Sousse, Tunisia: IEEE, 2008. 1-7.
- Wu WJ, Ma LZ, Xiao XZ. Method of tongue image segmentation based on luminance and roughness information. *Journal of System Simulation*, 2006;18(S1):374-6, 379.
- Zhang ZS, Liu Y. Tongue extraction algorithm based on dynamic thresholds and correction model. *Computer and Modernization*, 2014;(11):49-52.
- Li DX, Wei YK. Tongue image segmentation method based on adaptive thresholds. *Computer Technology and Development*, 2011;21(9):63-5.
- Zhi L, Yan JQ, Zhou T, Tang QL. Tongue shape detection based on b-spline. In: Proceedings of the 2006 International Conference on Machine Learning and Cybernetics. Dalian, China: IEEE, 2006. 3829-32.
- Li X, Li J, Wang D. Automatic tongue image segmentation based on histogram projection and matting. In: Proceedings of the 2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2014:76-81.
- Wang YQ, Wei BG, Cai YH, Shen LS. A knowledge-based arithmetic for automatic tongue segmentation. *Acta Electronica Sinica*, 2004, 32(3):489-91.
- Jang MS, Lee WB. Improved snakes algorithm for tongue image segmentation in Oriental tongue diagnosis. *The Journal of the Institute of Internet, Broadcasting and Communication*. 2016;16(4):125-31.
- Liu B, Hu G, Zhang X, Cai Y. Application of an Improved Grab Cut method in tongue image segmentation. *Intelligent Computing Methodologies*. In: Huang DS, Gromiha M, Han K, Hussain A (eds). ICIC 2018. Lecture Notes in Computer Science, Springer. 10956.
- Jeong JW. Demonstration of a deep learning model for extracting tongue image regions using SegNet convolutional neural networks. *Sangji Korean Medical Journal*. 2021;9(1):11-23.
- Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017;39(12):2481-95.
- Rhee S. Tongue area segmentation and tongue cracks detection using CNN. *J Korean Inst Intell Syst*. 2020;30:258-64.
- Hu J, Yan Z, Jiang J. Classification of fissured tongue images using deep neural networks. *Technology and Health Care*. 2022;30:271-83.
- Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A (eds). Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture Notes in Computer Science. Springer. 9351.
- Zhang Z, Zheng J, Zhao R, Liu S, Liu Z, Wang J. Cracked tongue extraction model based on improved U-Net Method. *IEEE Access*. 2023;11:126352-64.
- Ding J, Zhang Z, Tang J, Guo F. A multichannel deep neural network for retina vessel segmentation via a fusion mechanism. *Front Bioeng Biotechnol*. 2021;9:697915.
- Zhang N, Lin J, Hui B, Qiao B, Yang W, Shang R, Wang X, Lei J. Lung nodule segmentation and recognition algorithm based on multiposition U-Net. *Comput Math Methods Med*. 2022; 2022:5112867.
- Wang H, Shi Q, Zhang B, Xiao L. Multilevel attention Unet segmentation algorithm for lung cancer based on CT images. *Computers, Materials & Continua*. 2024;78(2):1569-89.
- Ni B, Liu Z, Cai X, Nappi M, Wan S. Segmentation of ultrasound image sequences by combing a novel deep siamese network with a deformable contour model. *Neural Computing and Applications*. 2023;35(20):14535-49.
- Hou S, Zhou T, Liu Y, Dang P, Lu H, Shi H. Teeth

- U-Net: A segmentation model of dental panoramic X-ray images for context semantics and contrast enhancement. *Computers in Biology and Medicine*. 2023;152:106296.
21. Rehman F, Ali Shah SI, Riaz MN, Gilani SO, R F. A Region-based deep level set formulation for vertebral bone segmentation of osteoporotic fractures. *J Digit Imaging*. 2020;33(1):191-203.
  22. Li B, Wu S, Zhang S, Liu X, Li G. Fast segmentation of vertebrae CT image based on the SNIC algorithm. *Tomography*. 2022;8(1):59-76.
  23. Zhang H, Jiang R, Yang T, Gao J, Wang Y, Zhang J. Study on TCM tongue image segmentation model based on convolutional neural network fused with superpixel. *Evid Based Complement Alternat Med*. 2022;2022:3943920.
  24. Liu W, Chen J, Liu B, Hu W, Wu X, Zhou H. Tongue image segmentation and tongue color classification based on deep learning. *Digital Chinese Medicine*. 2022;5(3):253-63.
  25. Chen H. Research on tongue segmentation based on convolutional neural network. Doctoral dissertation. Department of Computer Engineering Graduate School of Wonkwang University. 2021.
  26. Ahn IK, Bae KH, Lee SW. Tongue image segmentation using CNN and various image augmentation techniques. *Journal of Biomedical Engineering Research*. 2021;42(5):201-10.
  27. Li YJ, Jung ST. Tongue segmentation using the receptive field diversification of U-net. *Journal of the Korea Society of Computer and Information*. 2021;26(9):37-47.