JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 27 No. 9, pp. 1-12, September 2022
https://doi.org/10.9708/jksci.2022.27.09.001

Corneal Ulcer Region Detection With Semantic Segmentation Using Deep Learning

Jinhyuk Im*, Daewon Kim**

*Graduate Student, Graduate School of Computers, Dankook University, Yong-In, Korea
**Professor, Department of Computer Engineering, Dankook University, Yong-In, Korea

[Abstract]

Traditional methods of measuring corneal ulcers were difficult to present objective basis for diagnosis because of the subjective judgment of the medical staff through photographs taken with special equipment. In this paper, we propose a method to detect the ulcer area on a pixel basis in corneal ulcer images using a semantic segmentation model. In order to solve this problem, we performed the experiment to detect the ulcer area based on the DeepLab model which has the highest performance in semantic segmentation model. For the experiment, the training and test data were selected and the backbone network of DeepLab model which set as Xception and ResNet, respectively were evaluated and compared the performances. We used Dice similarity coefficient and IoU value as an indicator to evaluate the performances. Experimental results show that when '*crop & resized*' images are added to the dataset, it segment the ulcer area with an average accuracy about 93% of Dice similarity coefficient on the DeepLab model with ResNet101 as the backbone network. This study shows that the semantic segmentation model used for object detection also has an ability to make significant results when classifying objects with irregular shapes such as corneal ulcers. Ultimately, we will perform the extension of datasets and experiment with adaptive learning methods through future studies so that they can be implemented in real medical diagnosis environment.

▸ **Key words**: Semantic Segmentation, Corneal Ulcer, DeepLab, Xception, ResNet, Dice Similarity Coefficient (DSC), IoU

[요 약]

안과 환자의 질병을 판단하기 위해서는 특수 촬영 장비를 통해 찍은 안구영상을 이용한 안과의사의 주관적 판단의 개입이 전통적으로 활용되고 있다. 본 연구에서는 안과 의료진이 질병을 판단할 때 보조적 도움이 될 수 있도록 객관적 진단결과를 제시해주는 각막궤양 의미론적 분할방법에 대하여 제안하였다. 이를 위해 DeepLab 모델을 활용하였고 그 중 Backbone network으로 Xception과 ResNet 네트워크를 이용하였다. 실험결과를 나타내기 위한 평가지표로 다이스 유사계수와 IoU 값을 이용하였고 ResNet101 네트워크를 사용하였을 때 '*crop & resized*' 이미지에 대해 최대 평균 정확도 93%의 다이스 유사계수 값을 보였다. 본 연구는 객체 검출을 위한 의미론적 분할모델 또한 안구의 각막궤양 부분과 같은 불규칙하고 특이한 모양을 추출하고 분류하는데 뛰어난 결과를 도출할 수 있는 성능을 보유하고 있음을 보여주었다. 향후 학습용 Dataset을 양적으로 보강하여 실험결과의 정확도를 제고할 수 있도록 하고 실제 의료진단 환경에서 구현되어 사용되어 질 수 있도록 할 계획이다.

▸ **주제어**: 의미론적 분할, 각막궤양, 딥랩, Xception, ResNet, Dice Similarity Coefficient (DSC)

• First Author: Jinhyuk Im, Corresponding Author: Daewon Kim
*Jinhyuk Im (jhim0730@naver.com), Graduate School of Computers, Dankook University
**Daewon Kim (drdwkim@dku.edu), Department of Computer Engineering, Dankook University
• Received: 2022. 07. 04, Revised: 2022. 08. 01, Accepted: 2022. 08. 03.

http://www.ksci.re.kr pISSN:1598-849X | eISSN:2383-9945

I. Introduction

인공지능과 Deep Learning 기술의 빠른 발전으로 현대 사회에서 ICT 분야는 뗄레야 뗄 수 없는 일부분이 되었다. 그 중에서도 인공지능을 이용한 영상분할 기술은 일반산업계, 영화계, 보안분야 등에서의 얼굴인식, 의료계 등 다양한 분야에서 활용되는 비중이 매우 높은 영상처리 방법이며 현재 여러 다양한 분야에서 이 기술의 효율을 높이기 위한 연구가 활발히 이루어지고 있다[1-4]. 영상분할이란 임의로 주어진 디지털 영상을 특정 목적에 따라 여러 픽셀의 부분집합으로 나누는 것을 말한다. 이때, 사용분야의 목적에 따라 동일한 영상이라도 분할 결과가 다를 수 있으며 일반적으로 분할된 픽셀 집합을 의미 있는 영역으로 간주하여 그 다음 과정에 적절하도록 처리한다. 따라서 영상분할은 주로 Image Processing의 전처리 과정에서 이루어진다. 전통적 방법의 영상분할 과정은 Binarization, Region Growing, Edge Detection 등의 기법과 같이 대부분 픽셀 값들의 수학적 계산에 따른 결과로 분류하여 부분집합으로 나누는 방법을 통해 이루어졌다. 하지만 최근에는 인공지능 기술의 한 분야인 딥러닝 기법이 발전하여 매우 복잡한 영상 내에서도 Machine Learning을 통하여 기계 스스로 이미지의 특징들을 학습하여 사용자가 찾고자 의도한 영역을 인간의 눈과 비슷하게 구분하여 찾아내는 정도에 이르렀다. 이러한 기술들은 Satellite Image[5], Pedestrian Recognition[6], Robot Industry[7], Autonomous Driving[8] 등 다양한 분야에서 응용되고 있다. 특히 의료분야에서의 영상분할은 환자의 질병을 의료영상을 통하여 분석하고 진단하기 위한 중요한 전처리 과정으로 인식되고 있다. 의료영상 분할 작업은 주로 CT, MRI, 특수촬영 영상 등을 이용하여 인체장기에 발생한 다양한 종양 및 궤양 등을 검출하거나 특정 질병의 진료를 위한 체내객체의 부피측정 및 진단을 수행하기 위해 이루어진다. 여러 의료분야 중 특히 안과분야는 인간의 감각 중 매우 중요한 시각을 다루는 영역이기 때문에 의료진들도 불확실한 진단과 치료를 진행하지 않고 실수를 최소화 할 수 있도록 의료영상을 이용한 초기 진단이 상당히 중요하다. 안과에서 다루는 인간의 질병 중 각막궤양은 안구의 각막 표피에 발생하는 질병으로서 감염경로에 따라 세균성궤양과 진균성 궤양으로 나눌 수 있다. 각막궤양의 진단은 안구를 직접 촬영하는 정밀촬영 장비를 통해 이루어지는데 안구 상의 궤양을 촬영한 후 감염부위를 의료진이 직접 육안으로 보고 판단한 후 기타 추가 검사를 통해 세균성 요인에 의한 것인지 아니면 진균성 요인에 의한 것인지 등을 판별하게 된다. 이후 치료과정에서 궤양의 크기와 성질을 의료진이 직접 판별하여 치료의 효과와 향후 치료방향을 결정한다. 이때 각막궤양에 대한 의료적 판단을 함에 있어 안과 의료진의 주관적 판단력이 개입할 수 있는 여지가 있다. 그리고 각막궤양의 형태와 성질, 치료법은 의료진 개개인마다 판단하는 내용과 결과가 각기 다를 수 있다. 본 연구를 통하여 안과 의료진이 각막궤양에 대한 의료적 판단과 처치를 하고자 할 때 주관적 판단의 결과를 보조할 수 있는, ICT 기술을 활용한 객관적 판단의 의료적 근거를 제시하고자 한다. 본 논문에서는 의료영상 분할을 통해 안구상에서 각막의 궤양영역을 자동으로 검출하기 위해 딥러닝 모델 중 Semantic Segmentation 방법을 통해 학습한 모델을 사용하였다. 이를 토대로 각막의 궤양영역을 이미지의 픽셀단위로 검출하였고 Ground Truth와 비교하여 검출 정확도를 판단하였다. 또한 사용된 딥러닝 모델의 Backbone Network에 따른 결과를 비교하여 각막궤양 검출에 적합한 네트워크를 차별화하고 향후 연구에 대한 방향을 가늠하고자 하였다.

II. Related Works

Semantic Segmentation이란 영상 내에 검출하고자 하는 객체가 있을 때 이를 단순히 Bounding Box 등을 이용하여 위치정보만을 나타내는 것이 아니라 픽셀단위로 분류함에 있어 의미있는 영역을 검출해내는 것을 말한다. 전통적 영상분할 방법 중 영역성장법, Thresholding 등 픽셀값 기반 분할방법과의 가장 큰 차이점은 단순히 이미지 분할만 하는 것이 아니고 어떤 성질을 지닌 객체인지 그 의미를 구분하여 검출한다는 점이다. 의미론적 분할방법은 인공 신경망을 적용하여 큰 발전을 이루었는데 Fully Convolutional Network (FCN) 구조가 발표되면서 깊은 신경망을 이용한 의미론적 분할이 가능하다는 것을 보여주었다. J. Long[9][10]은 일반적으로 분류를 위해 사용되는 Convolutional Neural Network (CNN)의 레이어 구조에 마지막 단인 FCN을 1x1크기의 합성곱 층으로 간주하여 특징 맵 상에서의 위치정보를 유지하고 업샘플링 개념을 통해 특징 맵 크기를 입력영상과 맞추어줌으로써 영상에서 객체를 구분할 수 있는 특징을 얻을 수 있다는 것을 보여주었다. 이후 FCN 구조의 큰 틀은 유지하되 보다 성능을 높이기 위해 다양한 방법이 시도되었는데 H. Noh[11][12]는 FCN의 업샘플링 과정에서 손실되는 상세정보를 담은 픽셀을 보완하는 DeConvNet을 발표하였다.

F. Yu[13][14]는 FCN에서 손실되는 상세정보를 보완하는 방법으로 Dilated Convolution을 제시하였다. 이 연구에서는 기본적으로 FCN과 구조는 같으나 Pooling layer를 5개에서 3개로 줄여 특징맵의 크기가 현저히 줄어드는 것을 방지하여 상세정보를 비교적 잘 보존할 수 있도록 하였다. 그로인해 FCN보다 연산량이 늘어났는데 Receptive Field에서 일정한 간격의 Point만 이용하고 나머지는 0으로 채우는 방법인 Dilated Convolution을 통해 다시 연산량을 줄임으로써 속도는 비슷하지만 검출성능이 훨씬 정밀한 결과를 도출 하였다. 이처럼 향상된 성능은 의료분야에도 적용될 수 있을 만큼 많은 연구가 이루어졌다. H. Fu[15][16]는 FCN과 Fully-connected Conditional Random Field (CRF)를 이용하여 망막 스캔영상에서의 혈관검출을 성공하였으며 2016년 DRIVE and STARE 데이터셋을 이용한 혈관 분할 및 검출 성능 competition에서 최고성능을 보였다. CRF는 인접한 데이터들의 특징을 이용하여 해당 데이터의 특징을 유추하는 기계학습 기법으로서 개별적인 혈관들의 가능성 맵과 멀리 떨어져있는 픽셀간의 상호작용을 하나로 묶는 방법을 이용하였다. H. Lee[17]는 신체의 형태학적 분석을 위해 CT 영상에서 골격근 단면을 자동으로 분할하는 픽셀단위의 깊은 분할 알고리즘을 발표하였다. 분석을 가속화하기 위해 후처리된 영상을 사용하였으며 다이스 유사계수로 평가하였을 때 정답과 약 3.68%의 오차율을 보이면서 자동분할에 성공하였다. Y. Yuan[18]은 피부 근접 촬영 영상에서 피부병변을 검출하기 위해 19개의 Layer를 가진 깊은 FCN과 자카드 계수를 평가지표로 사용하였다. 또한 최소한의 전처리와 후처리과정만 필요로 하여 다양한 의료 영상 분할 업무에 적용이 가능하며 의료 영상분할이 CT 이미지나 MRI 뿐만이 아닌 RGB 포맷의 컬러영상에도 적용이 가능함을 보여주었다. G. Wang[19]은 Bounding Box와 Fine Tuning에 기반한 분할 기법을 이용해 CT 영상 내에서 특정 장기 영역을 분할하였다. 이때 후처리를 위한 방법으로 ScribbleSup[20] 기법을 이용하여 학습 후 평가가 이루어진 영상의 분할 성능을 높일 수 있다고 발표하였다. 각막궤양은 안구의 각막 상에 발생하는 질병으로 보통 CT 영상이나 MRI 보다는 안구촬영에 특화된 특수 장비를 통해 근접촬영된 영상을 활용한다. N. A. T. Otoum[21][22]은 각막궤양을 검출하기 위해 Fluorescein으로 각막을 염색한 후 푸른빛을 비추어 궤양이 좀 더 드러나게 하는 방법으로 촬영된 영상을 활용하였다. 파란색 계열의 배경과 초록색 계열의 염색된 궤양영역을 쉽게 분할하기 위해 영상의 색 공간을 RGB에서 HSV 색 공간으로 변환하여 검출하였다. 그러나 염색된 영상만을 이용해야 하고 검출의 범위가 원형으로 지정된 후 사용자가 직접 그 범위를 조정해주어야 하는 한계가 있어 완전 자동화에는 못 미쳤다고 볼 수 있다. T. F. Chen[23]과 T. P. Patel[24]은 각막궤양 검출을 위해 랜덤 포레스트 조직 분류기와 Active Contouring without Edges 방식을 활용하여 궤양 영역을 분할하였다. 궤양을 포함하고 있는 원 영상을 랜덤 포레스트 조직 분류기를 통해 각각 궤양과 배경을 뜻하는 백색과 흑색으로 맵핑하여 픽셀단위의 가능성 맵을 구성하였다. 그 후 3명의 다른 사용자를 통해 에지 없는 능동 윤곽선 검출을 위한 시작점을 얻어 검출결과를 도출 하였다. Q. Sun[25]은 의미론적 분할 방법을 구현하기 위해 입력영상을 각 픽셀을 중심으로 19x19 크기의 Patch 단위로 분리하여 CNN 모델에 학습시킴으로써 자동 절차를 통해 궤양 영역을 분리하였다. 이 연구에서는 fluorescein으로 염색된 영상을 사용하였으며 다이스 유사계수를 평가지표로 사용하였다. 또한 각 영상의 픽셀 한 개마다 Patch를 학습데이터로 활용함으로써 상대적으로 적은 양의 학습 데이터로도 성공적인 궤양 영역 검출이 가능함을 보여주었고 이러한 시스템이 깊은 신경망을 통해 절차적 자동화가 가능함을 보여주었다.

III. Corneal Ulcer Region Detection

본 연구에서는 각막궤양 영역을 검출하기 위해 특수촬영 장비를 통한 근접촬영 영상을 사용하였다. 이러한 영상들의 집합을 이용하여 Dataset을 구성하고 Ground Truth를 미리 제작하여 깊은 신경망에 학습되도록 한 후 실험결과를 비교평가 하였다. 각막궤양 영역 검출 과정을 자동화하기 위해 깊은 신경망을 이용한 의미론적 분할 방법을 사용하였으며 모델 구조 변경을 위한 필터크기 및 풀링 계층 다양화 또는 객체와 배경 간 비율 조정을 위한 저주파 필터링, 허프변환 등을 통해 각막궤양 영역 검출 성능이 가장 높은 모델을 평가하고 분석하였다. Fig. 1에 각막궤양을 포함하고 있는 안구이미지를 이용한 의미론적 분할 과정이 나타나 있다. Fig. 1은 입력이미지가 임의의 인코더에 의해 다운샘플링 되고 Convolution Layer를 거쳐 디코더에 의해 업샘플링 되는 과정을 보이고 있다.

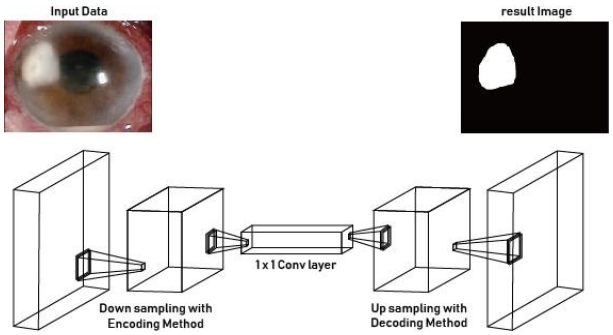


Fig. 1. Experimental Process of Semantic Segmentation

한편, 각막궤양 영역을 픽셀단위로 검출하여 분할하려는 의미론적 분할 분야에서 주로 사용되는 DeepLab[26][27] 모델은 발전을 거듭하면서 높은 정밀도와 성능향상을 보이고 있다. 본 연구에서는 궤양검출을 위해 DeepLab 모델인 v3 모델을 기반으로 하여 연구를 진행하였다. DeepLab 모델의 구조는 기본적으로 FCN 모델 구조와 마찬가지로 Convolution layer와 Pooling Layer를 통해 학습한 후 생성된 특징 맵의 크기를 키워가며 입력영상의 크기로 확대시키는 구조로 되어있다. 이 과정을 Encoding과 Decoding이라 부르는데 FCN 모델은 부호화 단계에서 일반적으로 알려진 CNN 모델의 합성계층을 사용하여 구현된 반면, DeepLab은 Fig. 2와 같이 Atrous Convolution 기법을 통해 구현되었다.

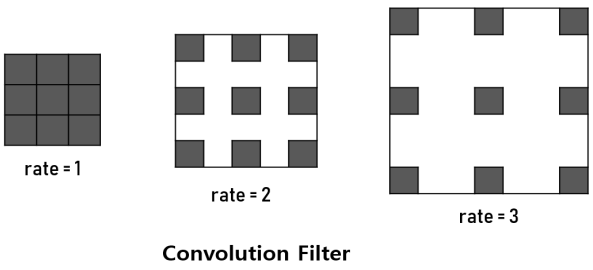


Fig. 2. Form of Atrous Convolution Filters

Atrous Convolution은 기존 합성계층과 다르게 컨볼루션 필터 셀 사이의 간격을 의미하는 파라미터인 rate를 가지며 이 rate 값이 커질수록 컨볼루션 필터의 범위는 넓어진다. 이는 곧 입력영상에서의 픽셀 당 수용필드의 범위가 넓어짐을 의미한다. 의미론적 분할 방법에서는 수용필드가 넓을수록 세부정보를 유지하는 능력이 다를 수 있으므로 모델의 성능이 향상될 가능성이 존재한다. 넓은 수용필드는 많은 연산량을 초래하지만 Atrous Convolution 기법은 컨볼루션 필터의 셀 사이간격의 hole에 0을 채움으로써 일반적인 컨볼루션과 동일한 연산비용으로 더 넓은 수용필드를 고려하는 방법으로 이 연산량 이슈를 해결하였다. DeepLab v1의 성능이 강화된 v2에서는 Atrous Convolution을 응용한 Atrous Spatial Pyramid Pooling (ASPP)를 활용 하였다. ASPP란 특징맵에 Atrous Convolution 필터의 rate이 서로 다른 여러 개의 필터들을 병렬로 묶은 합성계층을 적용하여 다양한 크기로 특징이 추출될 수 있도록 하고 결과적으로 더 정확한 분할이 될 수 있도록 한 것이다.

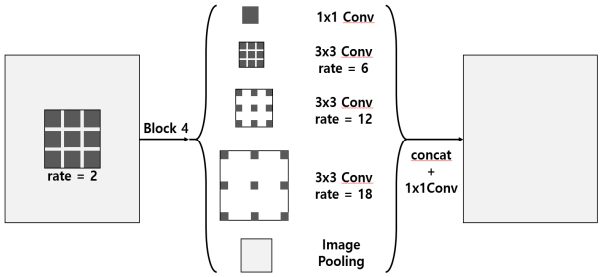
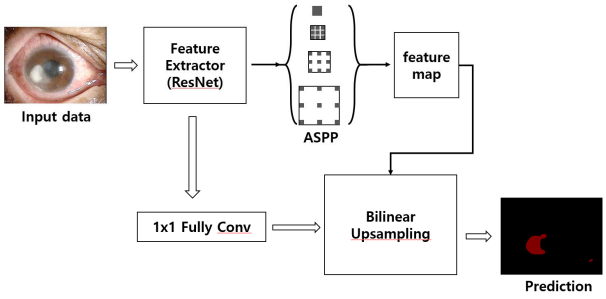
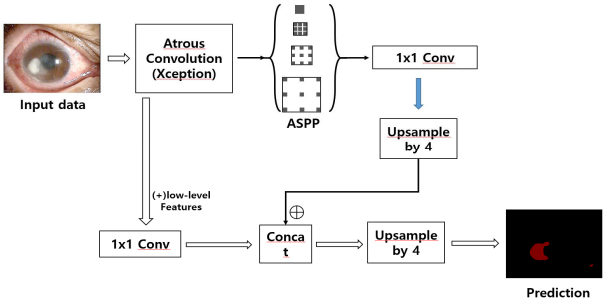


Fig. 3. Form of Atrous Spatial Pyramid Pooling

Fig. 3에 Atrous Spatial Pyramid Pooling 계층의 형태가 나타나 있다. 이후 DeepLab v3와 DeepLab v3+를 통해 성능 향상을 이루었는데 v3에서는 부호화 단계의 백본 네트워크로 ResNet을 사용하여 보다 더 깊은 학습을 통한 특징맵의 추출 성능을 향상시켰다.



(a)



(b)

Fig. 4. Network Architectures of DeepLab: (a) v3, (b) v3+

v3+에서는 Xception 네트워크와 깊이와 방향의 분리가 가능한 Depthwise Separable Convolution을 이용하여 늘어난 파라미터를 통한 성능 향상과 효율적인 연산이 가능하도록 하였다. 또한 복호화 단계에서는 기존의 FCN에서부터 활용되던 Bilinear Upsampling 대신 U-Net에 기반한 복호기를 사용하여 효율을 높였다. 본 연구에서는 v3와 v3+ 두 종류의 백본 네트워크를 활용하여 각막궤양 검출을 진행하였으며 각 네트워크의 궤양영역 분할성능을 측정하고 분석하였다. Fig. 4에 연구에 활용된 각 모델의 네트워크 구조가 나타나 있다. DeepLab 모델에 입력된 자료로는 각막궤양 영역을 포함하고 있는 전체 안구 이미지와 궤양영역 주변만이 나타나도록 사전 신호처리 후 잘라내고 크기를 재조정 한 이미지가 각각 사용되었다. 원 영상에서 각막궤양 영역의 경계가 모호함으로 인해 검출이 잘 되지 않는 현상을 방지하기 위해 회색조 변환, 히스토그램 평활화, Median 필터링, Gamma correction 등의 전처리 과정을 거치고 검출 정확도를 높이고자 하였다. 이렇게 원본 안구영상을 Preprocessing 단계를 거쳐 전처리하고 나아가 DeepLab 알고리즘으로 각막궤양 영역을 검출하는 전체 프로세스 구조가 Fig.5에 나타나 있다. 컬러값을 갖고 있는 원 영상은 RGB, HSV등 다양한 채널의 형태로 변환할 수 있는데 Fig. 5에 나타난 과정 중 사전 영상처리의 첫 번째 단계인 회색조 변환은 수 식 (1)과 같이 컬러픽셀의 R, G, B 값에 특정 상수를 곱한 결과를 각 픽셀에 적용함으로써 화소의 값이 0부터 255사이의 값으로 맵핑되도록 하는 과정이다.

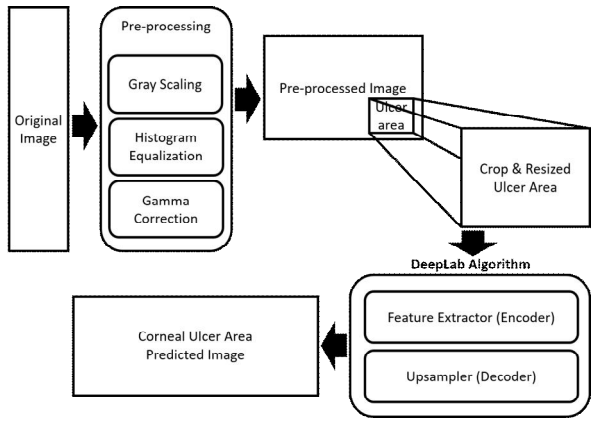


Fig. 5. Blockdiagram of Corneal Ulcer Region Detection

그 후 히스토그램 평활화 과정을 거치는데 회색조로 변환된 영상은 각 픽셀이 단일값을 가지므로 하나의 히스토그램으로 표현할 수 있다.

$$GRAY=(R\times 0.2125)+(G\times 0.7154)+(B\times 0.0721) \quad (1)$$

회색조 영상의 경우 픽셀값이 명암을 나타내는데 이 명암의 분포가 한쪽으로 치우쳐있거나 분포가 균일하지 않은 경우 경계가 명확히 인식되지 않을 수 있는데 이러한 문제를 히스토그램 평활화를 통해 해결할 수 있다. 이 과정을 통하여 영상에 Mapping 하는 방법은 수 식 (2)와 같은 사상함수를 적용함으로써 이루어지는데 여기서 $cdf(x)$는 누적 히스토그램에서 픽셀 값 x의 누적값을 의미한다.

$$p(x')=\frac{cdf(x)-cdf_{\min}}{(M\times N)-cdf_{\min}}\times(L-1) \quad (2)$$

L은 정규화 된 누적 히스토그램 범위의 크기를 의미하며 회색조 영상에서는 256이 된다. M과 N은 입력 이미지의 가로, 세로 크기를 뜻하며 결과적으로 이 과정을 통하여 이미지의 명암대비가 높아진다. 이어서 히스토그램 평활화를 진행한 후 영상의 잡음제거를 위해 Median Filter를 적용하였다. 3×3 크기의 중간값 필터를 영상에 적용하여 필터 내 픽셀 값들을 오름차순 또는 내림차순으로 정렬한 뒤 중간값으로 필터 내 모든 값을 변환하는 방법으로 잡음이 제거된 결과를 제공하였다. 이 필터가 적용된 후 안구 이미지는 각막궤양 영역의 윤곽선 부분에 잡음이 제거되어 원활한 분할에 도움이 된다. 덧붙여, 잡음이 제거된 영상의 배경과 궤양영역의 차이를 더욱 뚜렷하게 나타내기 위해 감마보정 필터를 통한 보정작업을 진행하였다. 일반적으로 사람의 눈은 명도가 낮은 부분에서는 영역의 구분이 쉬우나 명도가 높은 경우에는 구분력이 떨어진다. 이러한 특성은 디지털 영상이 표현되는 출력매체의 감마값에 따라 다르게 나타난다. 따라서 밝은 궤양 영역의 표현을 단순화시키고 이후 이루어지는 영상분할 단계에서 각막궤양 영역을 더 정확하게 검출하고자 하였다.

$$v'_{(i,j)}=\left(\frac{v_{(i,j)}}{n(l)}\right)^{\gamma}\times n(l) \quad (3)$$

수 식 (3)에 감마보정 과정이 나타나 있다. 수 식 (3)에서 $v'_{(i,j)}$은 영상에서 (i,j)에 위치한 픽셀의 보정된 밝기를 의미하며 $n(l)$은 픽셀이 가지는 색상의 Level 개수를 의미한다. 이와 같은 처리에 의해 영상의 밝은 부분이 단순화되어 표현된다. 각 모델 별로 분석을 진행한 후 검출

성능이 떨어지는 영상에 대하여 지금까지 제시된 사전 신호처리 후 배경과 객체의 픽셀 비율을 조정하기 위해 Fig. 6과 같이 영상의 궤양영역을 중심으로 Cropping과 Resizing을 진행한 후 학습 및 테스트 작업을 진행하였다. 원본 이미지인 640×480 크기의 안구영상은 각막과 공막, 홍채영역, 눈꺼풀, 속눈썹 등 궤양영역 검출에 불필요한 부분까지 모두 포함하고 있어, 보다 정확한 궤양영역 확보와 검출성능 향상을 위해 포토샵 작업을 통해 각막상의 궤양영역이 집중되어 보이도록 Cropping 한 후 일관성 있는 연산작업을 위해 320×240 크기로 Resizing 하였다.

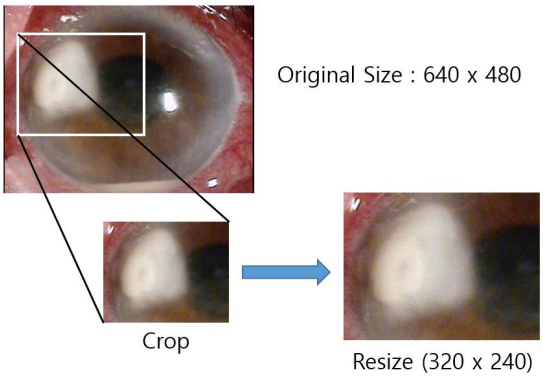


Fig. 6. Crop & Resize process of Corneal Ulcer Images

Deeplab v3에 사용된 ResNet은 Residual Neural Network의 약자로서 CNN 구조를 기반으로 한 분류 작업에 특화된 신경망 모델이다. ResNet의 기본 개념은 네트워크의 망 깊이가 깊어질수록 학습 효율이 더 좋아진다는 아이디어에서 출발하였으나 적용분야에 따라 일정 깊이를 넘어서면 오히려 학습 효율이 떨어질 수도 있다. 부적절한 네트워크의 깊이에 따른 비효율을 뜻하는 Degradation 문제를 해결하기 위해 ResNet에서는 Skip Connection 개념을 적용하였다. 이는 입력값 x에 대하여 일정 계층을 거친 F(x) 값이 다음 계층으로 진행될 때, 입력값을 출력값에 더한 H(x)를 다음계층에 전달함으로써 학습효율이 떨어지는 것을 방지하고자 한 것이다. Skip Connection 개념이 적용된 하나의 블록을 Residual Block이라 명명하였으며 수 식으로 나타내면 식 (4)와 같다.

$$H(x)=f(F(x)+x) \tag{4}$$

DeepLab v3+에 적용된 Xception 모델은 기존의 Inception 네트워크를 기반으로 하여 파생 변형된 네트워크로서 Inception 네트워크에 사용되던 병렬적 합성계층에 깊이방향의 분리 가능한 합성 방법이 더해진 네트워크이다. 깊이방향의 분리 가능한 합성 방법은 합성 계층에서 연산을 수행할 때 필터가 채널을 함께 연산하는 것이 아니라 채널 별로 수행한 합성의 결과를 더함으로써 전체 연산량을 줄이는 방법이다. Fig. 7에 이 과정이 R, G, B 각 채널에 대하여 적용되는 절차가 나타나 있다.

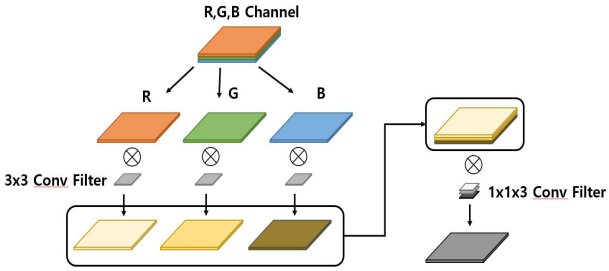


Fig. 7. Form of Depthwise Separable Convolution

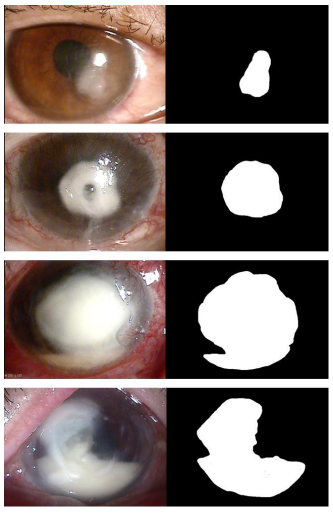
합성단계에서 줄어든 연산량으로 인해 학습 성능에 영향을 끼치는 다른 파라미터를 늘릴 수 있는 여유가 생겼고 이는 학습 성능의 증가로 이어질 수 있다. 본 연구에서는 안구촬영 이미지에서 각막궤양 영역을 자동으로 검출하고자 Semantic Segmentation 모델을 제안한 것이며 학습효율을 높이기 위해 앞서 설명한 방법 등을 이용하여 전처리 단계를 거쳐 Fine-Tuning 한 영상을 DeepLab의 입력 이미지로 활용하였다. 계속해서 사전 영상처리가 이루어진 이미지를 딥러닝 알고리즘에 입력하고 깊이가 다른 각각 두 개의 ResNet과 Xception 네트워크를 백본 네트워크로 활용하여 분할모델에 적용시킨 후 그 결과를 비교하고 분석하였다.

IV. Experiment and Results

환자의 안구 내 각막궤양 영역을 특수촬영한 장비를 통해 획득한 영상으로부터 자동으로 검출하기 위한 모델들의 성능결과를 비교하고 분석하기 위해 실험을 수행하였다.

Fig. 8. Slit Lamp BQ-900

각막궤양을 포함한 안구영상은 단국대학교 의과대학병원 안과학교실에서 제공 받았고 다양한 연령대의 환자 안구 촬영영상 95장 중 73장과 22장을 다양한 형태로 Augmentation 작업을 진행하여 각각 657장을 Training 단계에 사용하고 198장을 Testing 단계에서 사용하였다. 모든 영상은 Fig. 8에 보이는 특수촬영 장비인 Haag-Streit 사의 BQ 900 LED Slit Lamp를 이용하여 촬영되었다. 실험 시 학습 효율을 위해 원본영상을 VGA (640×480) 크기로 재조정하여 일치시켜 주었다. 또한 학습과정 중 검증단계에서와 Test 단계에서의 결과 성능을 평가하고 평가지표 값을 계산하여 정리하기 위해 정답영상이라 할 수 있는 Ground Truth를 제작하였다. Fig. 9에 실험에 사용된 샘플 영상과 그에 해당하는 Ground Truth가 나타나있다. 실험은 Windows 10 Pro OS, Intel(R) Core(TM) i7-10700K(3.8GHz) CPU, 32GB RAM, NVIDIA GeForce GTX 3070 8GB GPU의 개발환경에서 진행되었다. DeepLab 모델의 구동을 위해 영상을 처리함에 있어 CPU와 함께 GPU의 자원을 활용하는 Tensorflow-GPU 2.0 버전과 CUDA v10.2 프로세서를 통해 빠르고 효율적인 학습이 진행될 수 있도록 하였다. 각 모델들의 각막궤양 영역 검출 성능을 확인하고 비교평가하기 위해 Xception65, Xception71, ResNet50, ResNet101 네트워크를 사용하여 학습을 진행하였다.



(a) (b)

Fig. 9. (a) Original Image (b) Ground Truth

표 1에 사용된 데이터의 형태와 사용된 인코더, 디코더의 목록이 나타나 있다. 실험에 활용된 딥러닝 네트워크의 학습 후 Test 결과를 평가하기 위해 평가지표로 다이스 유사계수와 Intersection of Union (IoU) 값을 활용하였다.

Table 1. Explanation of Experiment

No.	Data	Encoder	Decoder
1	Original image	Xception 65	U-Net
2	Original image	Xception 71	U-Net
3	Original image	ResNet 50	Bilinear
4	Original image	ResNet 101	Bilinear
5	Crop & Resized	Xception 65	U-Net
6	Crop & Resized	ResNet 101	Bilinear

Dice Similarity Coefficient (DSC)는 서로 다른 두 집합의 유사도를 계산하기 위한 방법이며 계산식이 수 식 (5)에 나타나 있다.

$$DSC = \frac{2\left|n_t \cap n_s\right|}{n_t + n_s} \tag{5}$$

수 식 (5)에서 n_t는 Ground Truth의 객체에 대한 픽셀 수를 의미하고 n_s는 테스트 후 결과영상에서 궤양영역으로 검출된 영역의 픽셀수를 의미한다. 두 집합의 교집합이 되는 픽셀 개수에 2를 곱한 후 이를 정답영상과 결과영상의 픽셀수를 더한 값으로 나눠줌으로써 두 집합간의 유사도를 판단할 수 있으며 이 방법으로 각 네트워크의 테스트 결과를 평가하고 비교분석 하였다. 또 다른 평가지표인 IoU 값 역시 영상분할 분야에서 분할결과의 정확도를 구하기 위한 방법으로 수 식 (6)에 계산법이 나타나 있다.

$$IoU = \frac{n_t \cap n_p}{n_t \cup n_p} \tag{6}$$

수 식 (6)에서 n_t는 정답영상의 객체에 대한 픽셀 수를 의미하고 n_p는 검출결과의 객체에 대한 픽셀수를 의미한다. 두 픽셀집합의 교집합을 합집합으로 나눔으로써 정답영상에 대한 검출결과의 정확도를 판단할 수 있다. Xception 인코더의 경우 네트워크 명 뒷부분의 숫자가 의미하는 것은 깊이방향의 분리 가능한 합성방법을 이용하는 계층의 수를 뜻하며 망의 깊이가 다른 것이라 할 수 있다. ResNet 네트워크 역시 네트워크 명 뒷부분의 숫자가 의미하는 것은 Residual Block 계층의 깊이를 뜻한다. 각 실험은 20000번의 반복학습을 수행하였으며 ResNet의 경우 학습률은 0.0001로 설정하였고 Xception의 경우 Poly Learning Rate 기법을 이용하여 학습이 진행되면서 점차 학습률을 줄여나가는 방법을 사용하였다.

$$PLR = (1 - \frac{i}{(\max - i)})^p \quad (7)$$

수 식 (7)에 Poly Learning Rate (PLR) 계산법이 나타나 있다. 여기서 i는 iteration 횟수를 의미하며 p는 0.9를 적용하였다. Xception과 ResNet 각 네트워크는 ImageNet으로 사전학습 된 모델을 기반으로 Training을 진행하였다. 표 2와 3은 각 실험을 통해 측정된 다이스 유사계수와 IoU 값을 나타낸 것이다.

Table 2. Results of Accuracy using Dice Similarity Coefficients for Corneal Ulcer Detection (1=100%)

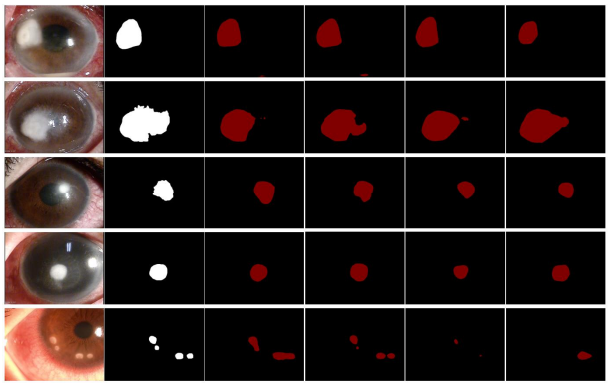
File No.	Xception 65	Xception 71	ResNet50	ResNet 101
0_00006	0.8206	0.7298	0.9269	0.9215
0_00007	0.7825	0.5896	0.9156	0.9765
0_00008	0.7373	0.9762	0.8102	0.7674
0_00009	0.9156	0.9351	0.5045	0.8651
0_00012	0.9288	0.7152	0.8906	0.8544
0_00013	0.7941	0.8559	0.7835	0.8578
0_00022	0.7793	0.7721	0.8765	0.9057
0_00025	0.8482	0.8921	0.8759	0.8963
0_00030	0.7402	0.8621	0.9163	0.9413
0_00034	0.8039	0.8882	0.9436	0.9548
0_00052	0.8156	0.8561	0.8719	0.8951
1_00003	0.8563	0.8452	0.9364	0.9561
1_00007	0.8525	0.7604	0.7563	0.8456
1_00013	0.8265	0.6089	0.7397	0.8113
1_00016	0.8764	0.8542	0.7183	0.8549
1_00019	0.8956	0.8573	0.8497	0.8037
1_00020	0.8845	0.9874	0.8621	0.8642
1_00022	0.8542	0.8695	0.9845	0.9785
1_00026	0.8419	0.7671	0.7428	0.9762
1_00031	0.7215	0.7654	0.9761	0.7462
1_00032	0.8652	0.8654	0.9875	0.9768
1_00034	0.8452	0.8152	0.8654	0.8653
Average	0.8311	0.8212	0.8515	0.8870

실험은 총 198장의 테스트용 영상에 대하여 진행되었는데 본 논문에서는 22개의 실험결과를 무작위로 추출하여 표2 ~ 표5에 제시하였다. 다이스 유사계수를 평가지표로 이용한 표2의 실험결과 Xception65 네트워크가 Xception71 네트워크 보다는 평균으로 약 1%의 우세한 결과를 보였다. 더불어 ResNet50 보다는 ResNet101 네트워크가 평균으로 3.6%정도 더 우세한 검출결과를 보였다. 표3에는 IoU 값을 평가지표로 한 결과가 나타나 있는데 이 결과 역시 Xception65 네트워크가 Xception71 네트워크 보다는 평균으로 약 1%의 우세한 결과를 보였다.

Table 3. Results of Accuracy using IoU value for Corneal Ulcer Detection (1=100%)

File No.	Xception 65	Xception 71	ResNet50	ResNet 101
0_00006	0.6957	0.5746	0.8637	0.8545
0_00007	0.8523	0.8623	0.9356	0.9526
0_00008	0.9762	0.6695	0.8526	0.8526
0_00009	0.8623	0.9236	0.8652	0.9452
0_00012	0.8671	0.8623	0.8028	0.8526
0_00013	0.6585	0.7481	0.8441	0.9565
0_00022	0.6383	0.6155	0.7802	0.8277
0_00025	0.7364	0.8052	0.7792	0.8121
0_00030	0.9562	0.9432	0.8455	0.8892
0_00034	0.8265	0.7989	0.8932	0.9136
0_00052	0.9523	0.8615	0.8729	0.8101
1_00003	0.8562	0.9564	0.9256	0.9462
1_00007	0.8456	0.8795	0.6081	0.8623
1_00013	0.8523	0.9868	0.9856	0.9152
1_00016	0.8623	0.8982	0.8652	0.8625
1_00019	0.8623	0.7503	0.7387	0.9526
1_00020	0.7929	0.6488	0.9452	0.9152
1_00022	0.9523	0.9768	0.9736	0.8625
1_00026	0.8463	0.8689	0.8256	0.8152
1_00031	0.8751	0.9126	0.9523	0.9456
1_00032	0.9562	0.8659	0.9765	0.9425
1_00034	0.8623	0.9752	0.8625	0.9462
Average	0.8448	0.8356	0.8633	0.8923

또한 ResNet50 보다는 ResNet101 네트워크가 평균으로 2.9%정도 더 우세한 검출결과를 보였다 종합 실험결과 Xception 네트워크보다 ResNet 네트워크의 분할정확도가 평균대비 상대적으로 약 3~4% 정도 높은 결과를 보였다. Fig. 10은 각 네트워크의 분류 결과 영상을 정답영상과 비교한 결과 중 일부를 나타내고 있다.



(a) (b) (c) (d) (e) (f)

Fig. 10. Results of Corneal Ulcer Detection: (a) Source Image, (b) Ground Truth, (c) Results from ResNet50, (d) Results from ResNet101, (e) Results from Xception65, (f) Results from Xception71

Fig. 10에서 (a)열은 입력 안구이미지이고 (b)열은 테스트 결과를 정확히 비교하기 위한 Ground Truth 영상이

다. (c)열은 ResNet50 네트워크를 이용한 검출결과이고 각각 (d), (e), (f)열은 ResNet101, Xception65, Xception71 네트워크를 이용하여 검출한 결과영상이다. 이 그림은 표2와 표3에 보이고 있는 다이스 유사계수와 IoU 값을 이용한 평가지표 값들에 해당하는 결과들을 이미지로 시각화하여 보였을 때 얻어지는 영상이다. 각막궤양을 포함한 원본 이미지를 이용하여 궤양영역 검출을 진행한 것에 이어 안구 이미지에서 궤양영역만을 미리 잘라내어 크기조절을 한 후 타겟 영역을 확대시킨 영상을 이용하여 데이터로 학습한 후 테스트를 진행하였다. 이 테스트는 원본 이미지를 그대로 이용했을 때 가장 높은 성능을 보였던 Xception65와 ResNet101 네트워크를 이용하여 진행하였다. 그 결과 두 네트워크 모두 원본 이미지를 그대로 학습했던 기존의 결과보다 향상된 결과를 보였으며 특히 ResNet101 네트워크의 경우 평균 93%의 다이스 계수 정확도를 보였다. Fig. 11에 Cropping과 Resizing을 각각 진행한 후 Xception65와 ResNet101에 적용하여 각막궤양 검출을 실험한 결과가 나타나있다. Fig. 11의 (a)행은 입력에 사용된 이미지를 보이고 있고 (b)행은 Ground Truth 영상이다. (c)행과 (d)행은 각각 Xception65와 ResNet101 네트워크를 이용하여 도출된 결과를 보이고 있다.

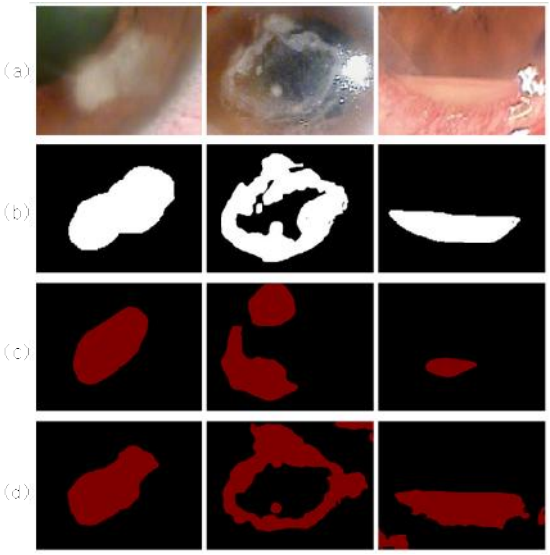


Fig. 11. Results of Corneal Ulcer Detection using Cropped and Resized Images: (a) Source Image, (b) Ground Truth, (c) Results from Xception65, (d) Results from ResNet101

두 번째와 세 번째 열에 보이는 이미지에 대한 (c)행의 Xception65를 이용한 검출결과보다는 (d)행의 ResNet101 네트워크를 이용한 검출결과가 더 낫다는 것을 육안으로 확인할 수 있다. 이를 수치로 표현하여 표4와 표5에 각각 나타내었다.

Table 4. Results of Accuracy using Dice Similarity Coefficient for Corneal Ulcer Detection from using Cropped and Resized Images (1=100%)

File No.	Xception65	ResNet101
0_00006	0.9059	0.9314
0_00007	0.9258	0.8915
0_00008	0.8256	0.9509
0_00009	0.8979	0.9425
0_00012	0.8591	0.9192
0_00013	0.8956	0.8491
0_00022	0.8507	0.9365
0_00025	0.8879	0.9461
0_00030	0.9365	0.9442
0_00034	0.8451	0.9391
0_00052	0.9396	0.9149
1_00003	0.9523	0.9258
1_00007	0.8256	0.9256
1_00013	0.8219	0.9625
1_00016	0.9526	0.9856
1_00019	0.8925	0.8954
1_00020	0.8925	0.8902
1_00022	0.8562	0.9156
1_00026	0.9236	0.9526
1_00031	0.8652	0.9485
1_00032	0.9258	0.9357
1_00034	0.9156	0.9812
Average	0.8906	0.9310

Table 5. Results of Accuracy using IoU Value for Corneal Ulcer Detection from using Cropped and Resized Images (1=100%)

File No.	Xception65	ResNet101
0_00006	0.8280	0.9256
0_00007	0.9556	0.9236
0_00008	0.9632	0.9064
0_00009	0.8147	0.9156
0_00012	0.8526	0.8504
0_00013	0.9356	0.8925
0_00022	0.7402	0.9256
0_00025	0.8456	0.8975
0_00030	0.7652	0.8943
0_00034	0.8956	0.8852
0_00052	0.9885	0.8431
1_00003	0.9458	0.9563
1_00007	0.8415	0.9145
1_00013	0.8756	0.8523
1_00016	0.9256	0.9258
1_00019	0.8759	0.9158
1_00020	0.9258	0.9852
1_00022	0.8563	0.9365
1_00026	0.9256	0.8826
1_00031	0.8452	0.9256
1_00032	0.9256	0.9258
1_00034	0.9258	0.9456
Average	0.8842	0.9102

두 표에는 잘라내서 크기를 재조정한 후 궤양영역을 확대하여 학습하고 테스트한 결과가 각각 다이스 유사계수와 IoU 값으로 나타나 있다. 표 4에 나타난 다이스 유사계수 평가지표 결과를 보면 Xception65의 경우 이전의 83%에서 89%로 6% 가량 증가한 것을 볼 수 있고 ResNet101 네트워크의 경우 88%에서 93%로 약 5% 증가한 것을 볼 수 있다. 이어서 표 5에 보이고 있는 IoU 값의 경우 Xception65에서 이전의 84%에서 88%로 4% 가량 증가한 것을 볼 수 있고 ResNet101 네트워크의 경우 89%에서 91%로 약 2% 검출 정확도가 증가한 것을 볼 수 있다. 실험에서 사용한 각각의 네트워크는 Backbone Network으로서 딥러닝 학습 시 Layer의 개수 또는 Convolution Filter의 개수 등 망의 깊이를 달리하여 실험한 후 각막궤양 영역 검출결과를 비교하고 분석하여 검출 효율을 높이는데 그 목적을 두었다. 따라서 본 연구에서 활용한 Xception65, Xception71, ResNet50, 그리고 ResNet101은 네트워크의 깊이에 따라 각기 다른 실험결과를 도출한 것이며 본 논문을 통해 최고의 성능을 보인 Backbone Network을 정량적으로 평가하고 비교하였다. 실험결과의 분석을 통해 의미론적 분할 모델인 DeepLab을 이용하여 안구 내 각막궤양 영상에서 궤양영역을 효과적으로 검출할 수 있고 오검출 또는 미검출 된 결과를 최소화하기 위하여 Crop & Resize와 같은 전처리과정을 거친 후 학습할 경우 그 성능이 향상되는 것을 확인하였다. 이어서 본 연구를 통해 제안한 방법의 실험결과와 기존의 연구논문 결과와의 비교 평가를 진행하였다. 표6에 본 논문의 연구결과와 Patel[24], 그리고 Sun[25]의 실험결과를 비교 분석한 내용이 나타나있다.

Table 6. Methods and Results Comparison for Corneal Ulcer Detection and Measurements

Compared Methods & Results		DC	IoU
Image-Based Analysis [24]	Epithelial Defects	0.83~0.86	N/A
	Stromal Infiltrate	0.78~0.83	N/A
Patch-Based Deep CNN [25]		0.86	N/A
Proposed Method		0.9108	0.89

이미지 기반의 처리를 통해 각막궤양의 크기를 측정하고 검출한 [24]의 경우 Epithelial Defects (상피손상)과 Stromal Infiltrate (사이질 각막염)에 대해서 각각 0.83 ~ 0.86과 0.78 ~ 0.83의 Dice Coefficients를 이용한 Accuracy를 보였고, [25]의 경우에는 패치 기반의 Deep CNN을 사용하여 평균 0.86의 Dice Coefficients 정확도를 보였다. 본 연구를 통해 제안한 방법으로는 평균 0.9108의 Dice Coefficients와 0.8972의 IoU 값을 보임으로써 각막궤양 영역 검출에 더 양호한 성능을 보였다. [24]와 [25]의 경우 각각 영상처리와 딥러닝 기법을 활용하여 각막궤양 영역을 검출하였는데 본 논문에서는 사전에 영상처리를 진행하고 후처리로 딥랩 기법을 활용하여 두 방법의 장점을 적절히 혼용한 것이 두 방법과의 차별점이라 할 수 있다. 향후 이러한 복합적인 방법으로 지속적인 각막궤양 질병 영역 검출 연구를 진행할 예정이다.

V. Conclusions

본 연구에서는 각막궤양이 포함된 안구 영상에서의 궤양영역을 픽셀단위로 검출하기 위해 의미론적 분할 방법을 사용하였고 DeepLab 모델을 활용하였다. 그리고 백본 네트워크 등 다양한 모델들의 내부 구조 변경을 통해 실험결과가 의미하는 성능을 분석하였다. DeepLab 모델은 물체검출 및 영역검출 등에서 수준급의 성능을 보이고 있는데 이를 이용해 자율주행, 위성영상 분석 등의 활용이 가능하다. 본 논문에서는 DeepLab 모델을 기반으로 백본 네트워크인 Xception과 ResNet을 각각 적용하였고 네트워크의 깊이를 다양하게 하여 안구 이미지 상 각막궤양 영역을 검출하는 방법을 제안하였다. 안과 환자들로부터 획득한 임상 데이터의 직접 활용을 위해 단국대학교 의과대학 안과학교실에서 제공한 환자들의 데이터를 사용하였고 Training과 Testing을 위해 Augmentation을 진행한 후 데이터별로 Ground Truth 이미지를 확보하였다. 상기 자료를 바탕으로 Xception과 ResNet 네트워크의 깊이를 달리하였고 Learning Rate 등을 조절하여 학습데이터로 각각 20000번의 반복학습을 진행하였다. 테스팅 후 실험 결과를 다이스 유사계수와 IoU 값을 이용한 지표로 평가하였다. 다이스 유사계수 평가지표로는 Xception65 네트워크에서 평균 83%, Xception71 네트워크에서는 평균 82%, ResNet50 네트워크에서는 평균 85%, 그리고 ResNet101 네트워크에서는 평균 88%의 정확도를 보였다. IoU 값을 이용한 평가지표로는 Xception65 네트워크에서 평균 84%, Xception71 네트워크에서 평균 83%, ResNet50 네트워크에서 평균 86%, 그리고 ResNet 101 네트워크에서는 평균 89%의 정확도를 보였다. 이후 두 종류의 네트워크 중 상대적으로 더 높은 정확도를 보이는 Xception65와 ResNet101 네트워크를 이용하여 원본이미지를 각막궤양 영역을 중심으로 잘라낸 후 크기를 재조정한 상태의 영상으로 만들어 데이터학습을 시켰을 때 검출 정확도가 최대 6.5% 정도 향상되는 것을 확인하였다. 다

이스 계수를 이용하여 정확도를 측정한 경우 평균 6.5%의 향상된 결과를 보였고 IoU 값을 이용하여 측정한 경우 평균 4% 정도의 향상된 결과를 보였다. 특히 ResNet101 네트워크의 경우 상대적으로 Xception65 네트워크보다 우월한 결과를 도출하였으며 DeepLab 모델의 의료분야 영상에의 확장사용 가능성을 높였다. 연구에 사용된 데이터셋의 크기가 현저히 작을 경우 테스팅 시 과적합 오류와 같은 이슈가 발생할 수 있으므로 향후 더 많은 안구질환 환자들로부터의 데이터셋 수집을 통한 실험을 계획 중이며 많은 양의 데이터로부터 도출된 일반적인 실험결과를 제시하고 해당내용 분석을 진행할 예정이다. 또한 의미론적 분할 모델의 적응적 학습 등의 기법을 이용하여 학습기능을 강화함으로써 의료분야에서 높은 신뢰성으로 활용이 가능하도록 연구를 진행할 계획이다.

REFERENCES

[1] Y. Tsai, W. C. Hung, S. Schulter and K. Sohn, "Learning to adapt structured output space for semantic segmentation," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7472-7481, Feb. 2018.

[2] Z. Wang, and S. Sarcar, "Outline Objects using Deep Reinforcement Learning," arXiv preprint arXiv : 1804.04603, Apr. 2018.

[3] S. Zheng, S. Jayasumana, B. Romera-Paredes, V. Vineet, Z. Su, D. Du, C. Huang and P. H. S. Torr, "Conditional random fields as recurrent neural networks," Proceedings of the IEEE international conference on computer vision, pp. 1529-1537, 2015.

[4] Y. Chu, J. Fei and S. Hou, "Adaptive Global Sliding-Mode Control for Dynamic Systems Using Double Hidden Layer Recurrent Neural Network Structure," IEEE Trans. on Neural Networks and Learning Systems, Vol. 31, No. 4, pp. 1297-1309, Apr. 2020. DOI: 10.1109/TNNLS.2019.2919676.

[5] V. Iglovikov, S. Mushinskiy and V. Osin, "Satellite imagery feature detection using deep convolutional neural network: A kaggle competition," arXiv preprint arXiv:1706.06169, Jun 2017.

[6] X. Du, M. El-Khamy, J. Lee and L. Davis, "Fused DNN: A deep neural network fusion approach to fast and robust pedestrian detection," 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 953-961, March 2017.

[7] L. Tai, G. Paolo and M. Liu, "Virtual-to-real deep reinforcement learning: Continuous control of mobile robots for mapless navigation," IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 31-36, Sep. 2017.

[8] M. Teichmann, M. Weber, M. Zöllner, R. Cipolla and R. Urtasun, "Multinet: Real-time joint semantic reasoning for autonomous driving," 2018 IEEE Intelligent Vehicles Symposium (IV), pp. 1013-1020, Changshu, Suzhou, China, Jun. 2018.

[9] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3431-3440, 2015.

[10] D. Wang, D. Zhang, G. Yang, B. Xu, Y. Luo and X. Yang, "SSRNet: In-Field Counting Wheat Ears Using Multi-Stage Convolutional Neural Network," IEEE Trans. on Geoscience and Remote Sensing, Vol. 60, pp. 1-11, 2022. Art No. 4403311, DOI: 10.1109/TGRS.2021.3093041.

[11] H. Noh, S. Hong, and B. Han, "Learning Deconvolution Network for Semantic Segmentation," The IEEE Conference on Computer Vision (ICCV), pp. 1520-1528, 2015.

[12] C. Peng, K. Zhang, Y. Ma and J. Ma, "Cross Fusion Net: A Fast Semantic Segmentation Network for Small-Scale Semantic Information Capturing in Aerial Scenes," IEEE Trans. on Geoscience and Remote Sensing, Vol. 60, pp. 1-13, 2022. Art no. 5601313, DOI: 10.1109/TGRS.2021.305 3062.

[13] F. Yu, and V. Koltun, "Multi-Scale Context Aggregation By Dilated Convolutions," IEEE Trans. on Parallel and Distributed Systems, Vol. 16, No. 3, pp. 219-232, Mar. 2005.

[14] Y. -J. Ma, H. -H. Shuai and W. -H. Cheng, "Spatiotemporal Dilated Convolution With Uncertain Matching for Video-Based Crowd Estimation," IEEE Trans. on Multimedia, Vol. 24, pp. 261-273, 2022. DOI: 10.1109/TMM.2021.30500 59.

[15] H. Fu, Y. Xu, D. W. K. Wong and J. Liu, "Retinal Vessel Segmentation via Deep Learning Network And Fully-connected Conditional Random Fields," 2016 IEEE 13th International Symposium on Biomedical Imaging(ISBI), pp. 698-701, Apr. 2016.

[16] C. Chen, J. H. Chuah, R. Ali and Y. Wang, "Retinal Vessel Segmentation Using Deep Learning: A Review," IEEE Access, Vol. 9, pp. 111985-112004, 2021. DOI: 10.1109/AC CESS.2021. 3102176.

[17] H. Lee, F. M. Troschel, S. Tajmir, G. Fuchs, J. Mario, F. J. Fintelmann and S. Do, "Pixel-level deep segmentation: artificial intelligence quantifies muscle on computed tomography for body morphometric analysis," Journal of digital imaging, Vol. 30, No. 4, pp. 487-498. Jun. 2017.

[18] Y. Yuan, M. Chao and Y. Lo, "Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks With Jaccard Distance," IEEE Trans. on Medical Imaging, Vol. 36, No. 9, pp. 1876-1886, Sep. 2017.

[19] G. Wang, W. Li, M. A. Zuluaga, R. Pratt, P. A. Patel, M. Aertsen, T. Doel, A. L. David, J. Deprest, S.Ourselin and T. Vercauteren, "Interactive Medical Image Segmentation Using Deep Learning With Image-Specific Fine Tuning," IEEE Trans. on Medical

Imaging, Vol. 37, No. 7, pp. 1562-1573, Jul. 2018.

[20] D. Lin, J. Dai, J. Jia, K. He and J. Sun, "ScribbleSup: Scribble-Supervised Convolutional Networks for Semantic Segmentation," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3159-3167, Apr. 2016.

[21] N. A. T. Otoum, "Medical Iamge Processing : applications in ophthalmology and total hip replacement," Loughborough University, pp. 52-74, 2012.

[22] Y. Ma, J. Liu, Y. Liu, H. Fu, Y. Hu, J. Cheng, H. Qi, Y. Wu, J. Zhang, and Y. Zhao, "Structure and Illumination Constrained GAN for Medical Image Enhancement," IEEE Trans. on Medical Imaging, Vol. 40, No. 12, pp. 3955-3967, Dec. 2021. DOI: 10.1109/TMI.2021.3101937.

[23] T. F. Chen, and L. A. Vese, "Active Contours Without Edges," IEEE Trans. on Image Processing, Vol. 10, No. 2, pp. 266-277, Feb. 2001.

[24] T. P. Patel, N. V. Prajna, S. Farsiu, N. G. Valikodath, L. M. Niziol, L. Dudeja, K. H. Kim and M. A. Woodward, "Novel Image Based Analysis for Reduction of Clinician-Dependent Varability in Measurement of the Corneal Ulcer Size," Clinical Science “Cornea”, pp. 331-339, Mar. 2018.

[25] Q. Sun, L. Deng, J. Liu, H. Huang, J. Yuan and X. Tang, "Patch-Based Deep Convolutional Neural Network for Corneal Ulcer Area Segmentation," Fetal, Infant and Ophthalmic Medical Image Analysis, Springer, Cham, pp. 101-108, 2017.

[26] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy and A. L. Yuille, "Semantic Image Segmentation With Deep Convolutional Nets and Fully Connected CRFs," arXiv preprint arXiv:1412.7062, 2014.

[27] G. Lenczner, A. Chan-Hon-Tong, B. Le Saux, N. Luminari and G. Le Besnerais, "DIAL: Deep Interactive and Active Learning for Semantic Segmentation in Remote Sensing," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 15, pp. 3376-3389, 2022. DOI: 10.1109/JSTARS.2022.3166551.

Authors

Jinhyuk Im received the B.S. (2017) from the Graduate school of Dankook University, Yong-In, Kyunggi-Do, Republic of Korea. He worked as a graduate student researcher at Next Generation Terminals in Multimedia & SW Lab. in Dankook University. His research interests include image processing, deep learning, and neural networks.

Daewon Kim received the Ph. D. (2002) in Electrical and Computer Engineering from Iowa State University, Ames, Iowa, USA. He is currently a professor in Department of Computer Engineering at Dankook University, Republic of Korea. His research interests include image and signal processing, deep learning, mobile applications and nondestructive evaluation.