

SKU-Net: Improved U-Net using Selective Kernel Convolution for Retinal Vessel Segmentation

Dong-Hwan Hwang*, Gwi-Seong Moon*, Yoon Kim*

*Student, Dept. of Computer Science and Engineering, Kangwon National University, Chuncheon, Korea

*Student, Dept. of Computer Science and Engineering, Kangwon National University, Chuncheon, Korea

*Professor, Dept. of Computer Science and Engineering, Kangwon National University, Chuncheon, Korea

[Abstract]

In this paper, we propose a deep learning-based retinal vessel segmentation model for handling multi-scale information of fundus images. we integrate the selective kernel convolution into U-Net-based convolutional neural network. The proposed model extracts and segment features information with various shapes and sizes of retinal blood vessels, which is important information for diagnosing eye-related diseases from fundus images. The proposed model consists of standard convolutions and selective kernel convolutions. While the standard convolutional layer extracts information through the same size kernel size, The selective kernel convolution extracts information from branches with various kernel sizes and combines them by adaptively adjusting them through split-attention. To evaluate the performance of the proposed model, we used the DRIVE and CHASE DB1 datasets and the proposed model showed F1 score of 82.91% and 81.71% on both datasets respectively, confirming that the proposed model is effective in segmenting retinal blood vessels.

▶ **Key words:** Deep Learning, Retinal Vessel Segmentation, Convolutional Neural Network, Selective Kernel Convolution, U-Net

[요 약]

본 논문에서는 안저영상의 다중 스케일 정보를 다루기 위한 딥러닝 기반의 망막 혈관 분할 모델을 제안한다. 제안 모델은 이미지 분할 딥러닝 모델인 U-Net과 선택적 커널 합성곱을 통합한 합성곱 신경망으로 안저영상에서 눈과 관련된 질병을 진단하는데 중요한 정보가 되는 망막 혈관의 다양한 모양과 크기를 갖는 특징 정보를 추출하고 분할한다. 제안 모델은 일반적인 합성곱과 선택적 커널 합성곱으로 구성된다. 일반적인 합성곱 층은 같은 크기 커널 크기를 통해 정보를 추출하는 반면, 선택적 커널 합성곱은 다양한 커널 크기를 갖는 브랜치들에서 정보를 추출하고 이를 분할 주의집중을 통해 적응적으로 조정하여 결합한다. 제안 모델의 성능 평가를 위해 안저영상 데이터인 DRIVE와 CHASE DB1 데이터셋을 사용하였으며 제안 모델은 두 데이터셋에 대하여 F1 점수 기준 82.91%, 81.71%의 성능을 보여 망막 혈관 분할에 효과적임을 확인하였다.

▶ **주제어:** 딥러닝, 망막 혈관 분할, 합성곱 신경망, 선택적 커널 합성곱, U-Net

-
- First Author: Dong-Hwan Hwang, Co-author: Gwi-Seong Moon, Corresponding Author: Yoon Kim
 - *Dong-Hwan Hwang (donghwan1994@kangwon.ac.kr), Dept. of Computer Science and Engineering, Kangwon National University
 - *Gwi-Seong Moon (msg0812@geo-vision.co.kr), Dept. of Computer Science and Engineering, Kangwon National University
 - *Yoon Kim (yooni@kangwon.ac.kr), Dept. of Computer Science and Engineering, Kangwon National University
 - Received: 2021. 03. 25, Revised: 2021. 04. 14, Accepted: 2021. 04. 14.

I. Introduction

망막 혈관의 구조는 고혈압, 동맥 경화증, 당뇨병의 미세혈관 합병증인 당뇨병 망막증과 같은 망막 질환을 발견하는데 중요한 정보가 된다. 안구 검사를 위해서는 신경과 전문의, 심장 전문의, 안과 전문의 및 기타 망막 혈관 질환 전문가의 참여와 작업이 필요하지만, 끊임없이 증가하는 망막 이미지를 처리하는 데 한계가 있다. 과거 많은 연구는 이러한 한계를 극복하기 위해 망막 혈관 구조를 검출하는 자동화 기술을 제안하였다.[1,2,3]. 그러나 망막 혈관이 미세하고 안저영상의 배경과 큰 차이가 없어 망막 혈관 검출 자동화 기술은 많은 어려움이 있다.

최근 발전한 딥러닝은 컴퓨터비전 분야에서 뛰어난 성능을 보였다. 특히 이미지넷 챌린지 데이터셋에서 우승한 AlexNet[4]이 등장하면서 많은 CNN 모델들이 제안되었다. 이에 따라 망막 혈관 검출 또한 딥러닝을 이용하는 방법이 제안되었다[5,6,7]. 그리고 이미지 분할 모델인 U-Net[9] 기반의 방법들은 각각 안저영상 40장과 28장으로 구성된 DRIVE, CHASE DB1과 같은 적은 수의 데이터에도 뛰어난 망막 혈관 분할 성능을 보였다. 하지만 U-Net 기반의 방법들은 대부분 3×3 합성곱 층을 사용한다. 3×3 합성곱 층으로 이루어진 인코더, 디코더는 같은 수용 영역 크기를 가지고 있어 고정적인 스케일의 공간 정보를 다룬다. 본 논문에서 다중 합성곱 브랜치를 사용하여 동적인 수용 영역을 통해 정보를 추출하고 추출된 다중 스케일의 공간 정보들을 각각 다른 가중치를 적용하여 결합하는 U-Net 기반 모델을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서 망막 혈관 검출 및 분할에 관한 이전 연구를 소개한다. 3장에서 제안모델에서 사용한 선택적 커널 방법을 소개하고 제안모델의 자세한 내용을 설명한다. 4장에서는 DRIVE와 CHASE DB1 데이터셋을 이용한 실험을 통하여 제안 모델의 우수성을 입증하고 마지막 5장에서 결론을 맺는다.

II. Related works

1. Previous Works

망막 혈관 분할을 목적으로 하는 초기 방법에는 픽셀의 밝기나 지식기반으로 설계된 특징들을 이용한 방법들이 있다. Akram 등[2]은 웨이블릿 변환을 기반으로 이미지에서 혈관의 특징 향상 기술을 제안하였다. Onkaew 등[3]은 이미지의 그래디언트 방향을 사용하여 망막 혈관을 분할

하는 방법을 제안하였다. 그리고 지도 학습 기반의 방법에서는 픽셀의 특징들을 추출하고 레이블 된 데이터를 이용하여 특징을 분류한다. Marin 등[1]은 픽셀 분류를 위해 신경망을 사용하고 픽셀 표현을 위해 그레이 레벨 및 모멘트 기반 특징으로 구성된 7D 벡터를 계산하였다. Aslani 등[8]은 여러 특징 추출기를 통해 계산된 하이브리드 특징 벡터로 픽셀을 특성화하고 랜덤 포레스트 분류기를 이용하여 이진 분류하는 분할 방법을 제안하였다.

2. U-Net based Retinal Vessel Segmentation

최근 딥러닝이 발달하면서 이미지 분할 분야에서 완전 합성곱 신경망(fully convolutional network)[8] 기반의 모델들이 등장하고 높은 성능을 내고 있다. Ronneberger 등이 제안한 U-Net[9]은 U자 모양의 FCN 모델로 다양한 이미지 인식 분야에 사용되는 모델이다[10,11]. U-Net은 공간 차원을 줄이는 인코더와 공간 차원을 늘리는 디코더로 구성된다. 또한 U-Net의 디코더에서는 같은 공간 차원을 갖는 낮은 수준(low-level)의 특징 맵과 높은 수준(high-level)의 특징 맵을 결합한다. 최근 망막 혈관 분할을 비롯한 의료 영상 분야에서 U-Net 기반의 다양한 모델들이 제안되었다. Residual U-Net[12]은 깊은 모델의 학습과 특징 맵의 정보 전파를 위해 사전 활성화 잔차(pre-activation residual)구조로 이루어진 모델이다. R2U-Net[13]은 반복 잔차(recurrent residual)구조를 활용하여 U-Net과 동일한 파라미터에서 향상된 성능을 내는 모델이다. 그리고 DenseBlockU-Net[14]은 U-Net의 합성곱 블록을 밀집 블록(dense block)으로 변형하여 학습 과정에서 발생하는 기울기 소실 및 폭발 문제를 완화하고 계층 간의 특징 전파를 강화한 모델이다. LadderNet[15]은 체인구조의 다중 U-Net으로 이루어진 모델로 다중 인코더-디코더 쌍을 이용하여 성능을 높였다. IterNet[16]은 U-Net과 여러 mini-U-Net의 반복적인 구조로 분할 결과를 반복적으로 분할하여 분할 결과에 가려진 혈관의 미세한 부분을 찾아낸다.

이러한 많은 U-Net 기반의 모델들은 많은 발전을 이루었지만, 대부분의 계층이 3×3 커널 크기의 합성곱 층으로 이루어져 있다. 이는 모델의 수용 영역의 크기를 제한하고 안저영상의 공간 정보를 고정적으로 다루어 크기와 모양이 다양한 망막 혈관을 검출하는 데 한계가 있다. DU-Net[17]은 이러한 한계를 해결하기 위해 변형 가능한 합성곱(deformable convolution)과 변형 가능한 관심 영역 풀링(deformable ROI pooling)을 이용하여 안저영상에서 실제 객체의 크기에 따라 수용 영역을 조정하였다.

본 논문은 DUNet[17]에서 더 나아가 다른 커널 크기의 브랜치들을 통해 안저영상의 다중 스케일 공간 정보를 추출하고 다중 스케일 공간 정보를 분할 주의집중 모듈을 통해 적응적으로 조정하여 결합하는 방법을 사용한다.

III. The Proposed Scheme

이번 장에서는 제안 모델의 핵심인 Selective kernel networks(SK-Net)[18]에서 제안한 선택적 커널 합성곱(selective kernel convolution)을 소개하고 제안모델에서 사용한 선택적 커널 합성곱 블록을 설명한다. 또한 U-Net[9]과 선택적 커널 합성곱 블록을 통합한 제안 모델 Selective kernel U-Net(SKU-Net)의 구조를 설명한다.

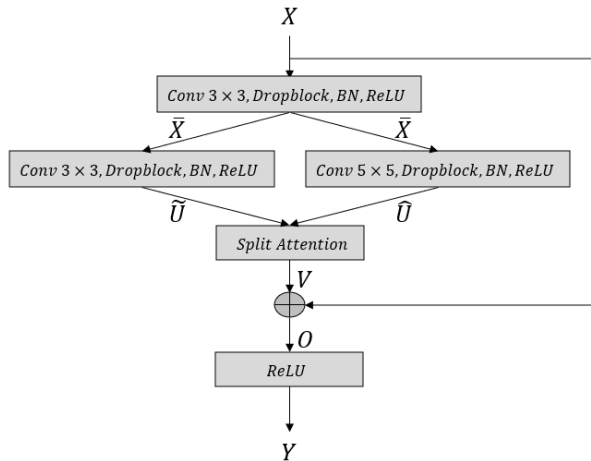


Fig. 1. Selective Kernel Convolution Block

1. Selective kernel convolution

선택적 커널 합성곱은 Selective kernel networks[18]에서 제안한 방법으로 동적으로 $2i + 1$ 의 커널 크기를 갖는 다중 브랜치와 분할 주의집중(split-attention)으로 이루어진 모듈이다. Fig 1은 제안 모델에서 사용된 선택적 커널 합성곱 블록 구조이다. 제안 모델의 선택적 커널 합성곱은 먼저 3×3 커널 크기의 합성곱 층이 있고, [18]과 유사하게 두 개의 브랜치와 분할 주의집중(split-attention) 모듈로 구성된다. 두 개의 브랜치는 3×3 커널 크기의 합성곱 층과 5×5 커널 크기의 합성곱 층으로 이루어진다. 또한 선택적 커널 합성곱 블록은 ResNet[19]의 잔차(residual) 구조를 따른다. 선택적 커널 합성곱 블록에는 합성곱 층 다음으로 드롭블록(Dropblock)[20], 배치 정규화(Batch normalization)[21], ReLU 활성화 함수를 추가한다. 드롭블록(Dropblock)[20]은 무작위로 활성

화 단위를 선택하고 그 주변 영역의 활성화 단위들을 같이 생각하여 모델의 과적합을 완화하는 효과가 있다.

입력 특징 맵 $X \in \mathbb{R}^{H \times W \times C}$ 가 주어졌을 때, 3×3 커널 크기의 합성곱 층에서 특징 맵 $\bar{X} \in \mathbb{R}^{H \times W \times C}$ 를 추출한다. 다음으로 3×3 커널 크기의 합성곱 층과 5×5 커널 크기의 합성곱 층에서 각각 특징 맵 $\tilde{U} \in \mathbb{R}^{H \times W \times C}$ 와 $\hat{U} \in \mathbb{R}^{H \times W \times C}$ 을 추출한다. 수식은 다음과 같다.

$$\begin{aligned} \bar{X} &= \bar{F}(X), \\ \tilde{U} &= \tilde{F}(\bar{X}), \\ \hat{U} &= \hat{F}(\bar{X}), \end{aligned} \quad (1)$$

식 (1)에서 $\bar{F}(\cdot)$ 와 $\tilde{F}(\cdot)$ 는 3×3 커널 크기의 합성곱 연산이고 $\hat{F}(\cdot)$ 는 5×5 커널 크기의 합성곱 연산이다. 다중 브랜치에서 추출한 특징 맵은 분할 주의집중 모듈을 통해 브랜치 별로 가중치를 구한다. 제안 모델에서 사용한 분할 주의집중 모듈은 [18]의 방법과 유사한 방법으로 다른 스케일의 정보를 적응적으로 조정하는 역할을 한다.

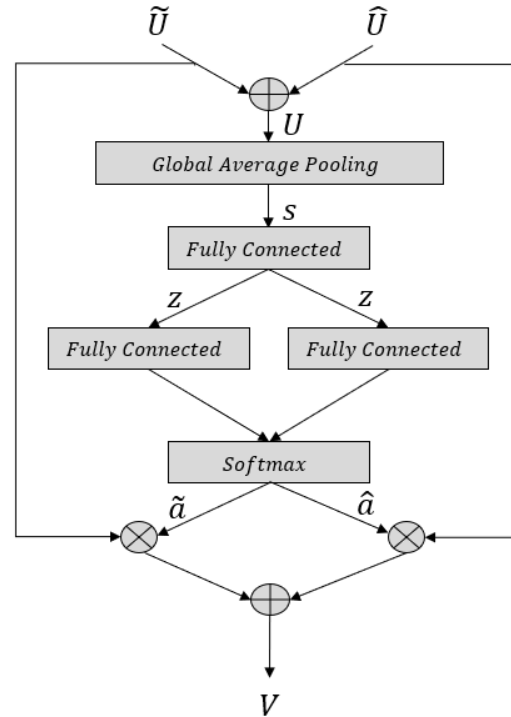


Fig. 2. Split-Attention

Fig 2는 제안 모델에서 사용한 분할 주의집중 모듈이다. 분할 주의집중 연산을 위해 먼저 다중 경로에서 생성된 특징 맵들을 원소별 합을 통하여 특징 맵 $U \in \mathbb{R}^{H \times W \times C}$ 를 구한다. 수식은 다음과 같다.

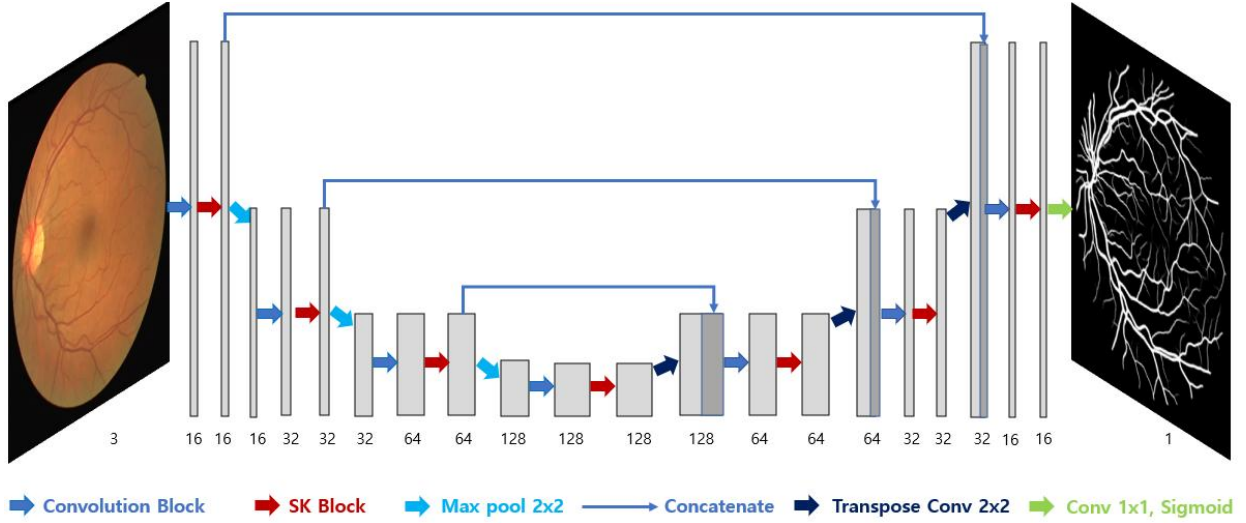


Fig. 3. Network Architecture

$$U = \tilde{U} + \hat{U}. \quad (2)$$

다음으로 대표 특징 맵 U 를 전역 평균 풀링(global average pooling) 연산을 통해 채널별로 공간 차원 정보를 압축하여 특징 벡터 $s \in \mathbb{R}^C$ 를 구한다.

$$s_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W U_c(i, j), \quad (3)$$

식 (3)에서 U_c 는 U 의 c 번째 채널이다. s_c 는 c 번째 채널에서 계산된 하나의 벡터로 특징 맵의 공간 차원 정보를 압축한 정보이다. 채널별 대표 정보 벡터 s 는 더 작은 채널 차원인 벡터 $z \in \mathbb{R}^{d \times 1}$ 로 변형한다.

$$z = Ws, \quad (4)$$

식 (4)에서 $W \in \mathbb{R}^{d \times C}$ 는 완전 연결 층(fully connected layer)의 가중치이다. d 는 압축을 위한 매개변수로 다음 식을 통해 구한다. 본 논문에서는 $r = 8$, $L = 32$ 를 사용하였다.

$$d = \max(C/r, L). \quad (5)$$

압축 벡터 z 를 다시 원래의 차원으로 변형한 후 소프트 맥스 연산을 하여 커널 크기별로 다른 가중치를 부여한다. c 번째 요소에 대한 계산식은 다음과 같다.

$$\begin{aligned} \tilde{a}_c &= \frac{\exp(\tilde{A}_c z)}{\exp(\tilde{A}_c z) + \exp(\hat{A}_c z)}, \\ \hat{a}_c &= \frac{\exp(\hat{A}_c z)}{\exp(\tilde{A}_c z) + \exp(\hat{A}_c z)}, \end{aligned} \quad (6)$$

식 (6)에서 $\tilde{A}, \hat{A} \in \mathbb{R}^{C \times d}$ 는 특징 벡터 z 에 선형 결합하는 완전 연결 층(fully connected layer)의 가중치이다.

\tilde{a}, \hat{a} 는 소프트 맥스를 통해 계산된 벡터로 브랜치에 따라 다른 가중치를 갖는다. \tilde{a}_c 와 \hat{a}_c 의 합은 1이며 각각 \tilde{U}_c 와 \hat{U}_c 에 곱합 후 원소별로 합하여 특징 맵 $V \in \mathbb{R}^{H \times W \times C}$ 를 구한다. c 번째 요소에 대한 계산식은 다음과 같다.

$$V_c = \tilde{a}_c \tilde{U}_c + \hat{a}_c \hat{U}_c, \quad (7)$$

식 (7)에서 $V_c \in \mathbb{R}^{H \times W}$ 는 c 번째 요소에 대한 연산 결과이다. 특징 맵 V 는 선택적 커널 합성곱 블록의 입력 X 와 더한 후 ReLU 활성화 함수를 거쳐 최종 특징 맵을 얻는다. 수식은 다음과 같다.

$$\begin{aligned} O &= V + X, \\ Y &= \sigma(O), \end{aligned} \quad (8)$$

식 (8)에서 σ 는 ReLU 활성화 함수이다. 최종적으로 구한 $Y \in \mathbb{R}^{H \times W \times C}$ 는 공간 차원을 조절한 후 다음 인코더나 디코더의 입력이 된다.

2. Network architecture

Fig 3은 제안 모델의 구조이다. 제안 모델은 다양한 수용 영역을 위해 선택적 커널 합성곱 블록(SK Block)이 추가된 인코더와 디코더로 구성되어 있다. Fig 4의 왼쪽은 U-Net, 오른쪽은 제안모델의 합성곱 블록이다. U-Net은 3×3 합성곱 층과 ReLU 활성화 함수로 이루어진 블록을 사용한다. 제안 모델의 합성곱 블록은 U-Net에서 사용하는 합성곱 블록에서 드롭블록(Dropblock)[20]과 배치 정규화(BN)[21]가 추가된다.

제안 모델의 인코더와 디코더는 드롭블록이 추가된 합성곱 블록과 선택적 커널 합성곱 블록으로 구성된다. 인코더에서 공간 차원을 줄이기 위한 방법으로 2×2 최대 풀

링(max pooling) 연산을 사용한다. 제안 모델의 디코더는 인코더와 동일한 구조이다. 디코더에서 공간 차원을 증가하기 위한 방법으로 2×2 전치 합성곱(transpose convolution)을 사용한다. 전치 합성곱은 공간 차원을 증가시키고 특징 맵의 채널을 반으로 줄이는 역할을 한다. 줄어든 특징 맵은 인코더에서 같은 해상도인 특징 맵과 접합되어 다음 디코더의 입력이 된다. 제안 모델의 마지막 특징 맵은 1×1 합성곱 층과 시그모이드 활성화 함수를 거쳐 출력된다. 학습에서 사용되는 손실함수로 이진 교차 엔트로피(binary cross-entropy)를 사용한다.

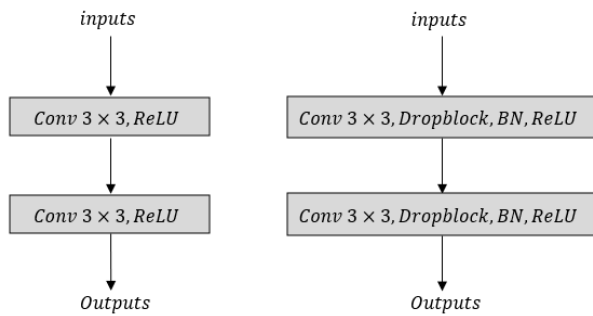


Fig. 4. Convolution block for U-Net(left), Convolution block for SKU-Net(right)

IV. Experimental Results

1. Dataset

제안 모델의 성능 평가를 위해 DRIVE와 CHASE DB1 데이터셋을 사용하였다. 두 데이터셋은 전문가에 의해 망막 혈관을 전경, 나머지를 배경으로 이진 분할되었다. DRIVE 데이터셋은 훈련을 위한 20장과 평가를 위한 20장을 포함한 40장의 안저영상을 포함한다. CHASE DB1 데이터셋은 28장의 안저영상으로 이루어져있다. CHASE DB1은 훈련 데이터셋과 시험 데이터셋을 구분하지 않으므로 다른 방법들[5,13,15,16,17,22,23]과의 비교를 위해 첫 20장은 훈련 데이터셋으로 나머지 8장은 시험 데이터셋으로 나누었다. DRIVE와 CHASE DB1의 해상도는 각각 565×584 , 999×960 이고 제안모델의 입력에 맞추기 위해 양선형 보간법을 사용하여 각각 592×592 , 1008×1008 로 크기 조정하였다. 또한 적은 수의 데이터 문제를 해결하기 위해 RGB 이미지에서 21개, 회색조 이미지로 변환 후 3개의 데이터 증강 기법을 사용하여 회색조 변환을 포함한 총 25개의 데이터 증강 기법을 DRIVE와 CHASE DB1 훈련 데이터셋에 적용하였다. 데이터 증강 기법에 대한 자세한 내용은 Table 1과 같다.

Table 1. Data Augmentations

Channel	DRIVE	CHASE DB1
RGB	Center Crop, Random Crop, Random Rotate, Transpose, Elastic Transform, Grid Distortion, Optical Distortion, Vertical Flip, Horizontal Flip, Random Brightness Contrast, Random Gamma, Hue Saturation Value, RGB Shift, Gaussian Noise, Random Brightness, Random Contrast, Motion Blur, Median Blur, Gaussian Blur, Channel Shuffle, Cutout[24]	
Gray Scale	Vertical Flip, Horizontal Flip, Center Crop	

2. Evaluation metrics

제안 모델의 정량 평가를 위해 안저영상의 모든 픽셀에 대한 실제 값과 분할 결과를 true positive(TP), false positive(FP), false negative(FN), true negative(TN)로 산출하고 정확도(Accuracy)와 F1 점수를 계산하였다. 정확도(Accuracy)와 F1 점수(F1-score)의 계산식은 다음과 같다.

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP}, \quad (9)$$

$$Precision = \frac{TP}{TP + FP}, \quad (10)$$

$$Recall = \frac{TP}{TP + FN}, \quad (11)$$

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}, \quad (12)$$

모델의 성능 평가하기 위해 추가로 ROC 곡선(AUC)을 계산하였다. AUC 값이 1이면 완벽한 분할을 의미한다.

3. Experiments details

제안 모델 SKU-Net의 성능 평가를 위해 U-Net과 성능을 비교하였다. 또한 드롭블록의 유효성을 검증하기 위해 드롭블록이 추가된 U-Net을 추가하여 총 세 개의 모델을 실험하였다. 세 개의 모델들은 스크래치에서 DRIVE와 CHASE DB1 데이터셋 모두 50 epoch 동안 학습하였다. 학습을 위한 파라미터로 학습률은 1×10^{-3} , 배치 크기는 DRIVE와 CHASE DB1 각각 8과 4로 적용하였다. 손실함수에 대한 최적화 알고리즘으로는 아담 최적화 방식(Adam optimizer)을 사용하고 학습률을 줄이는 스케줄러로 ReduceLROnPlateau를 사용하였다. 본 논문에서 사용한 드롭블록의 파라미터로 블록 크기 7×7 을 적용하고 keep rate는 DRIVE에서 0.85, CHASE DB1에서 0.80으

로 지정하였다. 또한 훈련 데이터셋의 10%는 검증 데이터셋으로 사용하였다. 제안 모델의 프레임워크는 Pytorch로 구현하였으며, 자세한 실험환경은 Table 2와 같다.

Table 2. Experiment Environment

OS	Ubuntu Server 18.04
CPU	Xeon Silver 4114
GPU	RTX3090
RAM	256GB
Deep learning framework	Pytorch 1.7.1

4. Results

Table 3과 4는 각각 DRIVE, CHASE DB1 데이터셋에 대한 실험 모델들의 망막 혈관 검출 성능 결과이다. Table 3과 Table 4에서 DRIVE, CHASE DB1 데이터셋 모두 U-Net+Dropblock이 U-Net보다 F1 점수(F1), 정확도(Accuracy), AUC에서 향상된 성능을 보인다. 이는 정규화를 위한 드롭블록이 망막 혈관 분할의 성능개선에 효과적임을 보여준다. 제안 모델인 SKU-Net은 U-Net+Dropblock과 비교했을 때 두 데이터에 대한 F1 점수, 정확도, AUC 지표에서 모두 향상된 성능을 보였다.

Table 3. DRIVE Dataset Result

Methods	F1	Accuracy	AUC
U-Net	0.8120	0.9682	0.9813
U-Net+Dropblock	0.8174	0.9687	0.9851
SKU-Net	0.8291	0.9699	0.9874

Table 4. CHASE DB1 Dataset Result

Methods	F1	Accuracy	AUC
U-Net	0.7927	0.9736	0.9870
U-Net+Dropblock	0.7930	0.9739	0.9902
SKU-Net	0.8171	0.9768	0.9909

Fig 5와 6은 DRIVE와 CHASE DB1 두 데이터셋의 몇 가지 이미지, 제안 방법의 분할 결과, 정답 레이블이다.

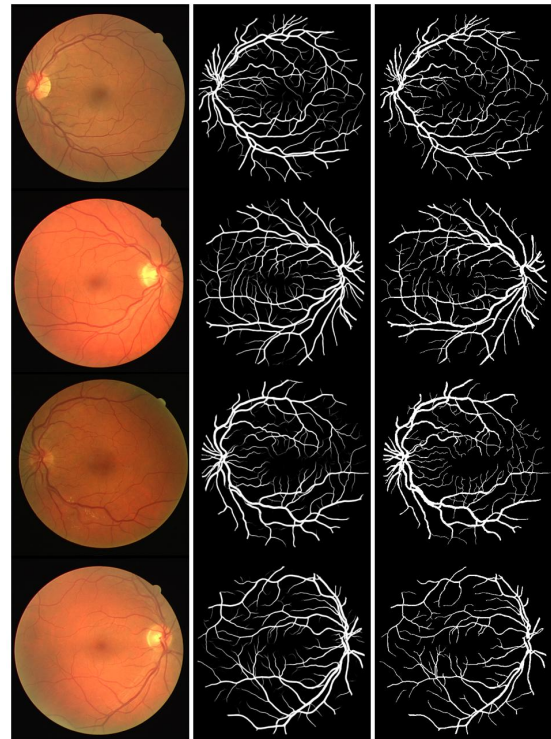


Fig. 5. DRIVE example original image(left), prediction(middle), ground truth(right)

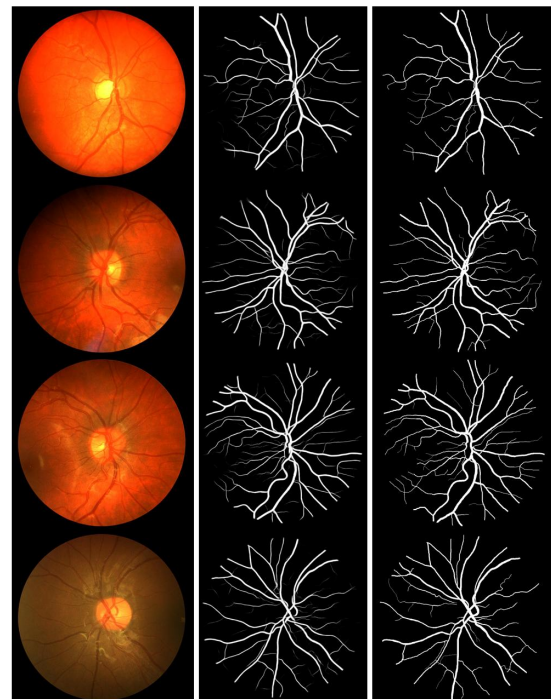


Fig. 6. CHASE DB1 example original image(left), prediction(middle), ground truth(right)

5. Comparisons with other methods

Table 5와 6은 각각 DRIVE와 CHASE DB1에 대한 제안 모델과 다른 방법들의 F1 점수, 정확도, AUC 지표에 대한 성능 비교를 보여준다. Table 5에서 제안모델을 다른 방법들과 비교하였을 때 F1 점수(F1), 정확도(Accuracy), AUC 기준 82.91%, 96.99%, 98.74%로 가장 높은 성능을 보인다. 그리고 Table 6에서 제안모델은 F1 점수, 정확도, AUC 기준 81.71%, 97.68%, 99.09%로 가장 높은 성능을 보인다.

Table 5. DRIVE Dataset Comparisons

Methods	F1	Accuracy	AUC
Orlando et. al. [5]	0.7857	0.9454	0.9506
R2UNet[13]	0.8171	0.9556	0.9784
M2U-Net[22]	0.8091	0.9630	0.9714
LadderNet[15]	0.8202	0.9561	0.9793
DUNet[17]	0.8203	0.9697	0.9856
IterNet[16]	0.8218	0.9574	0.9813
RSAN[23]	0.8203	0.9691	0.9855
SKU-Net(Ours)	0.8291	0.9699	0.9874

Table 6. CHASE DB1 Dataset Comparisons

Methods	F1	Accuracy	AUC
Orlando et. al. [5]	0.7332	0.9458	0.9524
R2UNet[13]	0.7928	0.9634	0.9815
M2U-Net[22]	0.8006	0.9703	0.9666
LadderNet[15]	0.8031	0.9656	0.9839
DUNet[17]	0.8001	0.9644	0.9834
IterNet[16]	0.8073	0.9655	0.9851
RSAN[23]	0.8111	0.9751	0.9894
SKU-Net(Ours)	0.8171	0.9768	0.9909

V. Conclusions

본 논문에서는 안저영상에서 망막 혈관 분할을 위한 SKU-Net을 제안한다. SKU-Net은 선택적 합성곱을 사용하여 안저영상에서 다중 스케일의 공간 정보를 추출하고 이를 적응적으로 결합하여 수용 영역을 조정한다. 실험을 통해 선택적 커널 합성곱 블록의 유효성을 확인하였고, 실험 결과 제안 모델은 DRIVE와 CHASE DB1 데이터에서 F1 점수 기준 각각 82.91%, 81.71%의 높은 성능을 보였다. 이는 같은 실험환경에서 DRIVE와 CHASE DB1 데이터에 대하여 U-Net보다 F1 점수 기준 1.71%p, 2.44%p의 성능 향상을 보임으로써 제안 모델 SKU-Net이 망막 혈관 분할에 효과적임을 증명하였다.

ACKNOWLEDGEMENT

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2021-2018-0-01433) supervised by the IITP(Institute for Information & communications Technology Promotion)

REFERENCES

- [1] D. Marín, A. Aquino, M. E. Gegundez-Arias and J. M. Bravo, "A New Supervised Method for Blood Vessel Segmentation in Retinal Images by Using Gray-Level and Moment Invariants-Based Features," in IEEE Transactions on Medical Imaging, Vol. 30, No. 1, pp. 146-158, Jan. 2011. DOI: <https://doi.org/10.1109/TMI.2010.2064333>
- [2] M. U. Akram, A. Atzaz, S. F. Aneque and S. A. Khan, "Blood vessel enhancement and segmentation using wavelet transform," 2009 International Conference on Digital Image Processing. IEEE, pp. 34-38, Bangkok, Thailand, Aug. 2009. DOI: <https://doi.org/10.1109/ICDIP.2009.70>
- [3] D. Onkaew, R. Turior, B. Uyyanonvara and T. Kondo, "Automatic Extraction of Retinal Vessels Based on Gradient Orientation Analysis," IEEE Eighth International Joint Conference on Computer Science and Software Engineering(JCSSE), pp. 102-107, Nakhonpathom, Thailand, Jun. 2011. DOI: <https://doi.org/10.1109/JCSSE.2011.5930102>
- [4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," Advances in neural information processing systems, Vol. 25, pp. 1097-1105, 2012. DOI: <https://doi.org/10.1145/3065386>
- [5] J. I. Orlando, E. Prokofyeva, and M. B. Blaschko, "A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images," IEEE transactions on Biomedical Engineering, Vol. 64, No. 1, pp. 16-27, Jan. 2017. DOI: <https://doi.org/10.1109/TBME.2016.2535311>
- [6] D. Maji, A. Santara, S. Ghosh, D. Sheet and P. Mitra, "Deep neural network and random forest hybrid architecture for learning to detect retinal vessels in fundus images," 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 3029-3032, Milan, Italy, Nov. 2015. DOI: <https://doi.org/10.1109/EMBC.2015.7319030>
- [7] S. Aslani and H. Samel, "A new supervised retinal vessel segmentation method based on robust hybrid features," Biomedical Signal Processing & Control, Vol. 30, pp. 1-12, Sep.

2016. DOI: <https://doi.org/10.1016/j.bspc.2016.05.006>
- [8] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *Proceedings of the IEEE conference on computer vision and pattern recognition(CVPR)*, pp. 3431-3440, Boston, MA, USA, Mar. 2015. DOI: <https://doi.org/10.1109/cvpr.2015.7298965>
- [9] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *International Conference on Medical image computing and computer-assisted intervention(MICCAI)*, Vol. 234-241, pp. 234-241, Springer, Cham, 2015. DOI: https://doi.org/10.1007/978-3-319-24574-4_28
- [10] J. Son, G. Moon and Y. Kim, "Automatic Detection System of Underground Pipe Using 3D GPR Exploration Data and Deep Convolutional Neural Networks," *Journal of the Korea Society of Computer and Information*, Vol. 26, No. 2, pp. 47-55, Feb. 2021. DOI: <https://doi.org/10.9708/jksoci.2021.26.02.027>
- [11] S. Kim, J. Sohn and D. Kim, "A method for concrete crack detection using U-Net based image inpainting technique," *Journal of the Korea Society of Computer and Information*, Vol. 25, No. 10, pp. 35-42, Oct. 2020. DOI: <https://doi.org/10.9708/jksoci.2020.25.10.035>
- [12] Z. Zhang, Q. Liu and Y. Wang, "Road extraction by deep residual u-net," *IEEE Geoscience and Remote Sensing Letters*, Vol. 15, No. 5, pp. 749-753, Mar. 2018. DOI: <https://doi.org/10.1109/LGRS.2018.2802944>
- [13] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha and V. K. Asari, "Recurrent residual U-Net for medical image segmentation," *Journal of Medical Imaging*, Vol. 6, No. 1, pp. 1-16, Mar. 2019. DOI: <https://doi.org/10.1117/1.jmi.6.1.014006>
- [14] X. Li, H. Chen, X. Qi, Q. Dou, C. Fu, and P. Heng, "H-DenseUNet: Hybrid densely connected UNet for liver and tumor segmentation from CT volumes," *IEEE Transactions on Medical Imaging*, Vol. 37, No. 12, pp. 2663-2674, 2018. DOI: <https://doi.org/10.1109/TMI.2018.2845918>
- [15] J. Zhuang, "Laddernet: Multi-path networks based on u-net for medical image segmentation," *arXiv preprint arXiv:1810.07810*, 2018.
- [16] L. Li, M. Verma, Y. Nakashima, H. Nagahara, R. Kawasaki, "Iternet: Retinal image segmentation utilizing structural redundancy in vessel networks," *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 3656-3665, CO, USA, Mar. 2020. DOI: <https://doi.org/10.1109/wacv45572.2020.9093621>
- [17] Q. Jin, Z. Meng, T. D. Pham, Q. Chen, L. Wei, and R. Su, "DUNet: A deformable network for retinal vessel segmentation," *Knowledge-Based Systems*, Vol. 178, pp. 149-162, 2019. DOI: <https://doi.org/10.1016/j.knosys.2019.04.025>
- [18] G. Ghiasi, T. Y. Lin and Q. V. Le, "Selective Kernel Networks," *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 510-519, Long Beach, CA, USA, June. 2019. DOI: <https://doi.org/10.1109/cvpr.2019.00060>
- [19] K. He, X. Zhang, S. Ren and J. Jun, "Deep residual learning for image recognition," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, Las Vegas, NV, USA, Jun. 2016. DOI: <https://doi.org/10.1109/cvpr.2016.90>
- [20] X. Li, W. Wang, H. Xiaolin and J. Yang, "Dropblock: A regularization method for convolutional networks," In *Neural Information Processing Systems*, 2018.
- [21] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.
- [22] T. Laibacher, T. Weyde, and S. Jalali, "M2U-Net: Effective and efficient retinal vessel segmentation for resource-constrained environments," *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 115-124, Long Beach, CA, USA, June. 2019. DOI: <https://doi.org/10.1109/cvprw.2019.00020>
- [23] C. Guo, M. Szemenyei, Y. Yi, W. Zhou and H. Bian, "Residual Spatial Attention Network for Retinal Vessel Segmentation," In *International Conference on Neural Information Processing, ICONIP 2020*, pp. 509-519, Springer, Cham, Nov. 2020. DOI: https://doi.org/10.1007/978-3-030-63830-6_43
- [24] T. DeVries and G. W. Taylor, "Improved Regularization of Convolutional Neural Networks with Cutout," *arXiv preprint arXiv:1708.04552*, 2017.

Authors



Dong-Hwan Hwang received the B.S. degrees in Computer Science and Engineering from Kangwon National University, Korea, in 2019. He is currently a M.S. Student in the department of Computer Science and

Engineering at Kangwon National University. His research interests are in the areas of machine learning and computer vision.



Gwi-Seong Moon received the B.S. degrees in Mechanical Engineering from Sungkyunkwan University, Korea, in 2015, and he received the M.S. degrees in Computer Science and Engineering from

Kangwon National University, Korea, in 2018. He is currently a Ph.D. Student in the department of Computer Science and Engineering at Kangwon National University. His research interests are in the areas of machine learning and computer vision.



Yoon Kim received a B.S. degree in 1993, an M.S. degree in 1995, and a Ph.D. degree in 2003, in electronic engineering with the Department of Electronic Engineering from Korea University. In 2004, he joined the

Department of Computer Science and Engineering, Kangwon National University, where he is currently a professor. From 1995 to 1999, he was with the LG-Philips LCD Co., where he was involved in research and development on digital image equipment. His research interests are in the areas of machine learning, multimedia communications, and computer vision.