

# 딥러닝 기반의 Semantic Segmentation을 위한 Residual U-Net에 관한 연구

신석용<sup>1</sup>, 이상훈<sup>2\*</sup>, 한현호<sup>3</sup>

<sup>1</sup>광운대학교 플라즈마바이오디스플레이학과 석사과정, <sup>2</sup>광운대학교 인제니움학부 교수,  
<sup>3</sup>울산대학교 교양대학 교수

## A Study on Residual U-Net for Semantic Segmentation based on Deep Learning

Seokyong Shin<sup>1</sup>, SangHun Lee<sup>2\*</sup>, HyunHo Han<sup>3</sup>

<sup>1</sup>Master Student, Department of Plasma Bio Display, Kwangwoon University

<sup>2</sup>Professor, Ingenium College of Liberal Arts, Kwangwoon University

<sup>3</sup>Professor, College of General Education, University of Ulsan

요약 본 논문에서는 U-Net 기반의 semantic segmentation 방법에서 정확도를 향상시키기 위해 residual learning을 활용한 인코더-디코더 구조의 모델을 제안하였다. U-Net은 딥러닝 기반의 semantic segmentation 방법이며 자율주행 자동차, 의료 영상 분석과 같은 응용 분야에서 주로 사용된다. 기존 U-Net은 인코더의 얇은 구조로 인해 특징 압축 과정에서 손실이 발생한다. 특징 손실은 객체의 클래스 분류에 필요한 context 정보를 부족을 초래하고 segmentation 정확도를 감소시키는 문제가 있다. 이를 개선하기 위해 제안하는 방법은 기존 U-Net에 특징 손실과 기울기 소실 문제를 방지하는데 효과적인 residual learning을 활용한 인코더를 통해 context 정보를 효율적으로 추출하였다. 또한, 인코더에서 down-sampling 연산을 줄여 특징맵에 포함된 공간 정보의 손실을 개선하였다. 제안하는 방법은 Cityscapes 데이터셋 실험에서 기존 U-Net 방법에 비해 segmentation 결과가 약 12% 향상되었다.

주제어 : 딥러닝, 인코더-디코더, 영상처리, Residual learning, Semantic segmentation

Abstract In this paper, we proposed an encoder-decoder model utilizing residual learning to improve the accuracy of the U-Net-based semantic segmentation method. U-Net is a deep learning-based semantic segmentation method and is mainly used in applications such as autonomous vehicles and medical image analysis. The conventional U-Net occurs loss in feature compression process due to the shallow structure of the encoder. The loss of features causes a lack of context information necessary for classifying objects and has a problem of reducing segmentation accuracy. To improve this, The proposed method efficiently extracted context information through an encoder using residual learning, which is effective in preventing feature loss and gradient vanishing problems in the conventional U-Net. Furthermore, we reduced down-sampling operations in the encoder to reduce the loss of spatial information included in the feature maps. The proposed method showed an improved segmentation result of about 12% compared to the conventional U-Net in the Cityscapes dataset experiment.

Key Words : Deep learning, Encoder-Decoder, Image processing, Residual learning, Semantic segmentation

\*Corresponding Author : SangHun Lee(leesh58@kw.ac.kr)

Received April 21, 2021

Accepted June 20, 2021

Revised May 4, 2021

Published June 28, 2021

## 1. 서론

최근 컴퓨터 비전에서 segmentation 및 객체 검출은 자율주행 자동차, 의료 영상 분석과 같은 분야에서 주로 사용되며, 많은 응용 프로그램에서 중요한 역할을 한다.[1] Segmentation에는 semantic segmentation, instance segmentation 등 다양한 종류가 있다. 그 중, 본 논문에서 다루는 semantic segmentation은 이미지에 존재하는 객체의 범주 (category)와 그 객체를 구성하는 픽셀들을 분류하는 기법이다. 그러나 semantic segmentation은 같은 범주의 객체들을 서로 구분하지 않으며, 이를 추가로 구분하는 기법은 panoptic segmentation [2]이라 한다.

딥러닝 알고리즘이 도입된 이후로 semantic segmentation은 빠르게 발전되었다. 딥러닝 기반의 semantic segmentation은 Convolutional Neural Network (CNN)을 이용한 방법들이 많이 연구되고 있다. 대표적인 방법에는 Fully Convolutional Networks (FCN) [3], U-Net [4], SegNet [5] 등이 있으며, 인코더-디코더 (Encoder-Decoder) [6] 구조로 설계되어 있다. 인코더 (Encoder)는 입력 이미지에서 특징을 추출하고 압축하여 context 정보를 생성하며, 디코더 (Decoder)는 context 정보가 포함된 특징맵을 확장하여 segmentation map을 출력한다. Context 정보는 객체의 범주를 구분하는데 필수적인 정보이며, segmentation map은 이미지에서 각 픽셀이 해당하는 범주를 표시한 것이다.

FCN은 딥러닝 기반 semantic segmentation 방법 중 가장 처음 제안되었으며, classification을 위한 네트워크를 segmentation에 사용하는 전략을 제시하였다. 이때, FCN은 segmentation map을 생성하는 과정에서 윤곽선이나 세밀한 정보가 손실되어 객체의 외형이 부정확하게 분할되는 문제가 있다. 따라서, 이 문제를 해결하기 위해 U-Net은 concatenation 연산 [7]으로 인코더의 중간 특징 정보를 디코더에서 활용하여 정보 손실을 최소화하였다. 그 결과 FCN과 비교하여 segmentation 정확도가 개선되었다. 그러나 U-Net은 인코더에서 추출하는 context 정보가 부족하여 객체의 분할이 잘못되거나 실패하는 문제가 있다. 이러한 문제는 응용 분야인 자율주행 자동차에서 객체의 위치 및 동작 인식을 방해하며, 의료 영상 분석에서 병변의 위치와 크기에 대한 오차를 높인다.

본 논문은 기존 연구의 문제를 해결하고 segmentation

정확도를 높이기 위한 연구이다. 제안하는 방법은 기존 U-Net에 residual learning을 활용한 인코더를 통해 특징 손실을 줄이고 context 정보를 효율적으로 추출하였다. 또한, 인코더에서 down-sampling 연산을 줄여 특징맵에 포함된 공간 정보의 손실을 개선하였다. 이 연구는 응용 분야에서 긍정적인 효과를 보일 것으로 기대한다.

## 2. 관련 연구

### 2.1 U-Net

U-Net [4]은 segmentation을 목적으로 설계된 U자 형태의 네트워크이며 인코더-디코더 구조이다. U-Net은 FCN의 문제점을 해결하기 위해 concatenation 연산으로 각 단계의 특징 정보를 그대로 활용하고 후속 convolution 연산으로 특징 정보를 정제하였다. Fig. 1은 U-Net의 구조이다. Fig. 1에서 왼쪽 부분은 인코더이고, 오른쪽 부분은 디코더이다. 인코더는 각 단계마다 convolution 2개와 pooling 1개, 디코더는 transposed convolution 1개와 convolution 2개로 구성되어 있다. 그리고 concatenation 연산은 인코더에서 각 단계의 특징맵을 디코더에서 대응되는 단계의 transposed convolution을 수행한 특징맵과 연결하여 하나의 특징맵으로 구성한다. 이때, U-Net은 인코더에서 convolution이 각 단계마다 2개씩, 총 10개만 존재하므로 다른 네트워크 (ResNet, MobileNet [8] 등)과 비교하여 convolution이 현저히 적다. 이는 인코더의 context 정보 추출 능력이 떨어져 segmentation이 부정확해지는 문제를 초래한다.

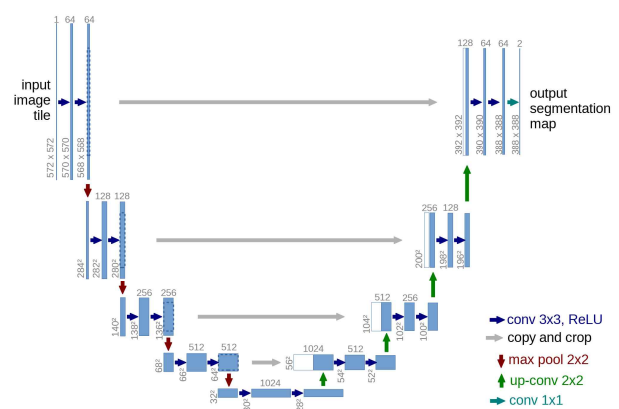


Fig. 1. U-Net [4]

## 2.2 ResNet

ResNet [9]은 residual learning을 통해 기울기 소실 문제 (gradient vanishing)를 해결하도록 설계한 네트워크이다. 기울기 소실 문제는 역전파 과정에서 입력 계층으로 갈수록 기울기가 점진적으로 0에 수렴하는 현상이며, 신경망의 학습을 방해한다. 특히, 신경망에 계층 수가 많을수록 쉽게 발생한다. 이 문제를 해결하기 위해 ResNet은 residual learning을 제안하였다. Residual learning은 신경망의 각 계층에서 출력에 입력을 더해 기울기가 적어도 1 이상이 되도록 하여 기울기 소실을 방지한다.

## 3. 제안하는 방법

제안하는 방법은 context 정보를 추출하기 위한 인코더와 정밀한 segmentation map 생성을 위한 디코더로 구성되었다. Fig. 2는 제안하는 방법의 구조이다. 인코더는 2개의 3x3 convolution과 residual learning이 포함된 16개의 Encoding block으로 구성되었고, 디코더는 3개의 Decoding block으로 구성되었다. 이때, 디코더에서 출력한 특징맵의 채널 수는 데이터셋의 범주 개수 (20개)를 초과한다. 따라서 1x1 convolution을 사용하여 특징맵의 채널 수를 맞춘 뒤, 최종 segmentation map을 생성하였다.

### 3.1 Residual learning을 이용한 인코더

Fig. 2에 표시한 것처럼 제안하는 방법은 인코더-디코더 구조이다. 제안하는 방법은 인코더에서 residual

learning이 적용된 ResNet을 백본 네트워크로 활용하였다. ResNet은 ResNet-18, ResNet-50, ResNet-101 등 여러 종류가 있다. ResNet-n에서 n은 네트워크의 계층 수이며 n이 커질수록 연산량이 증가하지만, 성능이 향상된다. 본 논문에서는 segmentation 성능과 속도의 trade-off를 고려하여 ResNet-34를 선택하였다.

Fig. 2에서 제안하는 방법의 인코더는 특징맵의 해상도에 따라 4개의 블록 (i.e. Encoding block1, Encoding block2, Encoding block3, Encoding block4)으로 구분하였다. 블록들은 각각 입력 이미지의 1/1, 1/2, 1/4, 1/8 해상도를 가지며, 각 블록 위의 숫자 (i.e. x3, x4, x6)는 해당 블록의 반복 횟수이다. 각 Encoding block의 residual learning 구조는 모두 같으며, 수식 1과 Fig. 3은 residual learning이 적용된 Encoding block을 나타낸다.

$$H(x) = F(x) + x \tag{1}$$

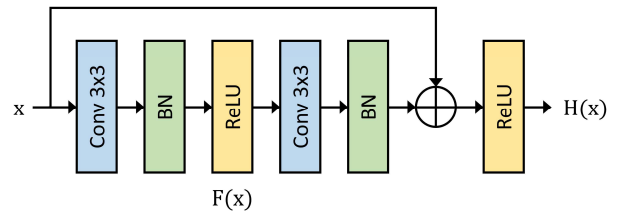


Fig. 3. Encoding block with Residual learning

제안하는 방법은 인코더에서 convolution 수가 총 34개로 Fig. 1의 기존 U-Net보다 약 3배 많도록 설계하였다. 이 구조는 인코더의 context 정보 추출 능력을 개선하여 segmentation 정확도를 높인다. 그러나 인코더

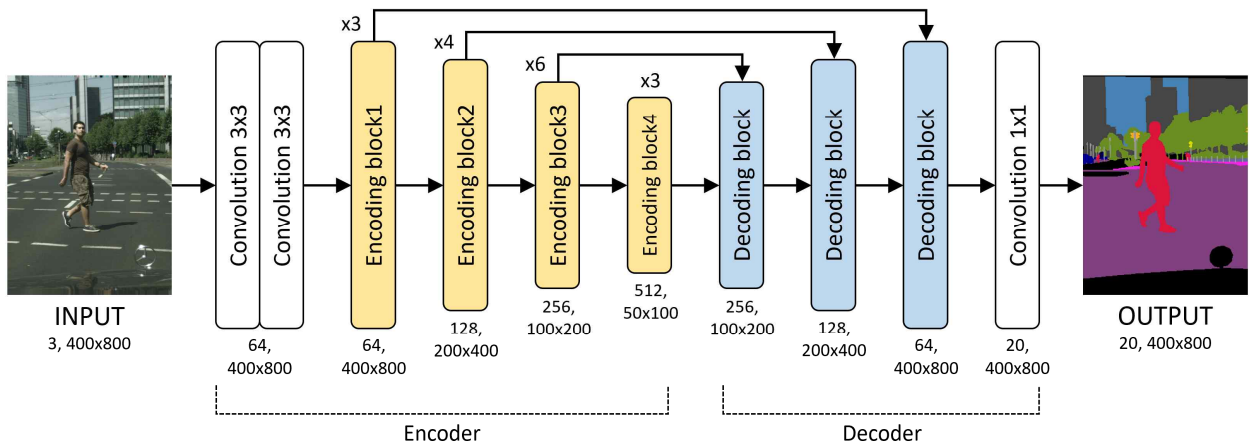


Fig. 2. Proposed method architecture

가 깊어짐에 따라 특징 손실과 기울기 소실 문제가 발생하기 쉽다. 이러한 문제를 해결하기 위해 제안하는 방법에 residual learning을 적용하였다. Residual learning은 연속적인 convolution의 특징 추출 및 압축 과정에서 발생하는 특징 손실과 네트워크가 깊어질수록 발생하기 쉬운 기울기 소실 문제를 최소화한다.

일반적인 ResNet은 Encoding block을 진행하기 전에 stride 2를 가진 7x7 convolution과 max-pooling을 수행하여 특징맵의 해상도를 축소한다. 또한, 여러 semantic segmentation 방법들도 ResNet을 백본으로 사용하면서 이러한 down-sampling 과정을 그대로 수행한다. 그러나 convolution과 pooling으로 특징맵의 해상도를 축소할 경우, 특징맵에 포함된 공간 정보의 손실을 초래할 수 있다. 따라서 제안하는 방법은 7x7 convolution과 max-pooling을 2개의 3x3 convolution으로 교체하여 특징맵의 해상도를 축소시키지 않고 특징 추출 전 처리 과정만 수행하였다. 추가로 언급하지 않으면 convolution 계층 뒤에 Batch normalization (BN), Rectified Linear Unit (ReLU) 계층이 포함된다.

### 3.2 정밀한 segmentation을 위한 디코더

디코더는 인코더에서 출력한 특징맵을 입력 이미지 크기로 확장 및 복원하는 과정이다. 이때 단순히 보간법 또는 transposed convolution으로 특징맵을 확장하면 윤곽선과 같은 세밀한 부분을 복원할 수 없어 segmentation 정확도가 크게 떨어진다. 따라서 제안하는 방법은 concatenation으로 인코더에서 추출된 중간 특징맵들을 활용하여 이 문제를 개선하였다. 인코더의 중간 특징맵들은 인코더의 최종 출력 특징맵보다 상세 정보가 풍부하다.

Fig. 2에서 Decoding block의 구조는 다음과 같다. 먼저 stride 2를 가진 2x2 transposed convolution으로 특징맵을 2배로 확장하였다. 그런 다음, 확장된 특징맵과 이와 같은 해상도를 가진 인코더의 중간 특징맵을 concatenate한다. concatenation은 두 특징맵을 채널 차원 기준으로 이어 붙이는 과정이다. Fig. 4와 수식 2는 concatenation을 나타내며, Fig. 5와 수식 3은 Decoding block을 나타낸다.

수식 2 및 3에서  $T(x)$ 는 transposed convolution으로 확장된 특징맵이고,  $E(y)$ 는 인코더에 있는 Encoding block이 생성한 중간 특징맵이다. 그리고  $c$ ,  $h$ ,  $w$ 는 각각 채널 수, 높이, 너비를 나타낸다.

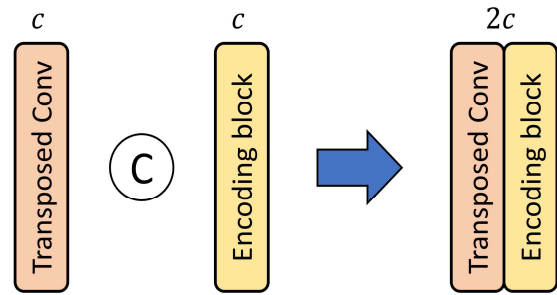


Fig. 4. Concatenation diagram

$$\text{Concat}(2c, h, w) = [T(x)_{c,h,w}; E(y)_{c,h,w}] \quad (2)$$

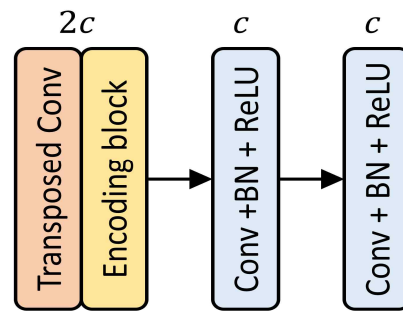


Fig. 5. The details of Decoding block

$$DB = f_{conv}^{3 \times 3}(f_{conv}^{3 \times 3}(\text{Concat}(T(x); E(y)))) \quad (3)$$

마지막으로 결합된 특징맵에 3x3 convolution, Batch normalization, ReLU를 2번 반복 수행하여 특징맵을 재구성하였다. 특히, 제안하는 방법은 Fig. 1의 기존 U-Net과 다르게 디코더에서 batch normalization을 사용하여 학습 과정을 안정화하였다. Batch normalization은 학습 시간을 단축하며 최저점 (global minimum)에 수렴하지 못해 학습이 더 이상 진행되지 않는 문제를 최소화한다. 결합된 특징맵의 채널 수는 수식 3에 따라 2배로 증가하며, 재구성한 특징맵의 채널 수는 Fig. 5에 따라 다시 2배 감소하여 결합 전의 채널 수를 유지한다.

### 3.3 손실 함수

제안하는 방법은 교차 엔트로피 (Cross entropy) 손실 함수를 사용하였다. 이는 semantic segmentation과 같이 범주의 개수가 여러 개인 다중 분류 (multinomial classification)에서 주로 사용한다. 교차 엔트로피 손실 함수는 제안하는 방법이 생성한 segmentation map과

groundtruth를 비교하여 오차를 출력한다. 이러한 오차를 줄이기 위해 신경망의 가중치를 조절하는 과정이 학습이다. 수식 4는 교차 엔트로피 손실 함수이다.

$$CE = - \sum_k t_k \log_e(y_k) \quad (4)$$

여기서  $t_k$ ,  $y_k$ ,  $k$ 는 각각 groundtruth, 제안하는 방법의 출력, 학습 데이터의  $k$ 번째 요소이다. 교차 엔트로피 손실 함수는 자연로그를 포함하므로  $t_k$ 가 1일 때,  $y_k$ 가 0에 가까울수록 오차는 기하급수적으로 커지게 된다. 따라서 평균 제곱 오차 (MSE: Mean Squared Error)와 같이 오답에 대한 영향을 크게 받는 특성이 있다.

## 4. 실험 및 결과

### 4.1 실험 환경

본 논문에서는 semantic segmentation에서 널리 사용되는 Cityscapes 데이터셋 [10]으로 실험하였다. 이 데이터셋은 50개의 서로 다른 도시에서 수집한 5000장의 도로 장면 이미지 (street scenes image)가 포함되어 있다. 그리고 학습을 위한 2975장의 이미지, 검증을 위한 500장의 이미지, 테스트를 위한 1525장의 이미지로 구성되어 있다. 실험에서는 한정된 VRAM 용량으로 인해 이미지 크기를 800x400, 배치 크기 (batch size)를 8로 설정하였다. 그런 다음, 에폭 (epoch)은 100, 초기 학습률 (learning rate)은 0.001, 가중치 감쇠 (weight decay)는 0.00001로 설정하고 Adam optimizer (Adaptive Moment Estimation optimizer)를 사용하였다. 또한 학습을 진행하면서 검증 손실 (validation loss)이 5 epoch 동안 줄지 않으면 학습률을 1/2만큼 축소하여 최저점에 더욱 근접할 수 있도록 하였다. 그리고 1/2의 확률로 이미지를 수평으로 뒤집는 데이터 증강 (data augmentation) 기법을 적용하였다. 실험에서 사용한 하드웨어와 소프트웨어는 다음과 같다: Intel i7 9700 CPU, NVIDIA RTX 3090 GPU, 32GB RAM, Ubuntu 20.10 x64, Python 3.8, PyTorch 1.8.1.

실험 결과는 Mean Intersection over Union (MIoU)을 사용하여 비교 및 분석하였다. MIoU는 semantic segmentation에서 정확도 측정을 위한 평가

지표이며 수식 5과 같다. 수식 5에서 TP, FP, FN,  $k$ 는 각각 true positive, false positive, false negative, class number를 나타낸다.

$$MIoU = \frac{1}{k} \sum_{i=0}^k \frac{TP}{TP + FP + FN} \quad (5)$$

### 4.2 실험 결과 및 고찰

본 논문은 여러 기존 방법들과 제안하는 방법을 정확도 (MIoU)와 파라미터 개수 (Params)의 측면에서 비교하였다. Table 1은 Cityscapes 데이터셋에서 제안하는 방법과 다른 방법들의 정확도와 파라미터 개수를 나타낸다.

Table 1. The results on Cityscapes dataset.

Method	Input size	Para.(M)	MIoU(%)
FCN-8s [3]	1024x2048	35.3	65.3
U-Net [4]	400x800	31.0	55.8
SegNet [5]	256x512	29.5	57.0
ENet [11]	360x640	<b>0.4</b>	57.0
SQ [12]	1024x2048	-	59.8
CRF-RNN [13]	512x1024	-	62.5
DeepLab [14]	512x1024	262.1	64.8
DPN [15]	512x1024	-	66.8
<b>Proposed</b>	400x800	24.3	<b>68.2</b>

제안하는 방법의 MIoU는 약 68.2%로 다른 방법들보다 높은 정확도를 보였다. 특히, 제안하는 방법은 U-Net보다 MIoU가 약 12% 향상되었다. 제안하는 방법의 파라미터 개수는 FCN, U-Net, SegNet, DeepLab과 비교하여 적은 수를 가졌다.



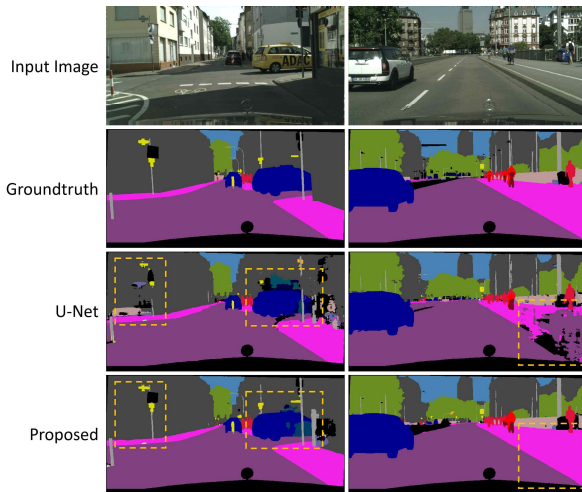


Fig. 6. Visual comparisons of complex scene on the Cityscapes dataset.

Fig. 6은 Cityscapes 데이터셋을 사용하여 복잡한 경관에서 제안하는 방법과 U-Net의 segmentation 결과를 비교하였다. 그림에서 주황색 상자는 기존 방법보다 좋은 분할 결과를 보인 부분이다. U-Net은 클래스 분류에 중요한 context 정보가 부족하여 복잡한 경관에서 전체적으로 segmentation이 불확실한 결과를 보였다. 특히 주황색 상자 영역의 기둥, 자동차, 보도 등의 객체에서 검출에 실패한 픽셀이 많다. 그러나 제안하는 방법은 residual learning을 활용하여 context 정보를 효율적으로 추출하였기 때문에 U-Net과 비교하여 더 정밀한 segmentation 결과를 보였다. 주황색 상자 영역에서 검출에 실패한 픽셀을 최소화하였다.

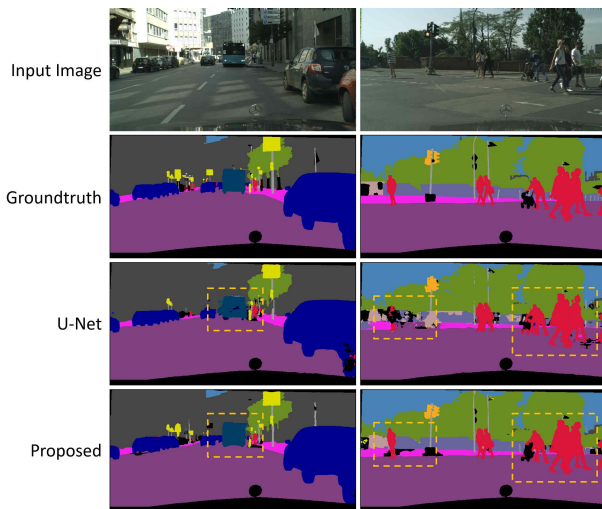


Fig. 7. Visual comparisons of multiple objects on the Cityscapes dataset.

Fig. 7은 여러 객체에서 제안하는 방법과 U-Net의 segmentation 결과를 비교하였다. 그림에서 주황색 상자는 기존 방법보다 좋은 분할 결과를 보인 부분이다. U-Net은 버스, 자동차, 사람 등 다양한 객체에서 segmentation 정확도가 떨어지는 결과를 보였다. 특히 객체의 일부가 다른 클래스로 분류되는 문제가 있다. 그러나 제안하는 방법은 다양한 객체들을 온전히 분할하여 높은 정확도를 보였다. 또한, 그림에서 버스, 사람들, 주변 객체 등을 누락되는 픽셀이 거의 없이 분할하였다.

## 5. 결론

본 논문에서는 U-Net 기반의 semantic segmentation 방법에서 정확도를 향상시키기 위해 residual learning을 활용한 방법을 제안하였다. 기존 U-Net은 인코더의 얇은 구조로 인해 특징 압축 과정에서 손실이 발생하며, 특히 클래스 분류에 필수적인 context 정보가 부족한 문제가 있다. 이 문제를 해결하기 위해 본 논문은 기존 U-Net에 residual learning을 활용한 인코더를 제안하여 특징 손실을 줄이고 context 정보를 효율적으로 추출하였다. 또한 인코더에서 down-sampling 연산을 줄여 특징맵에 포함된 공간 정보의 손실을 개선하였다. 실험 결과는 제안하는 방법이 residual learning과 특징맵의 공간 정보 손실 억제를 통해 segmentation 정확도가 개선되었다. 이를 통해 제안하는 방법은 정확도가 중요한 자율주행 자동차, 의료 영상 분석 등의 응용 분야에서 활용될 수 있다. 향후 연구는 특징 강조를 위한 Attention mechanism을 사용하여 특징 추출 및 복원을 강화할 필요가 있다고 사료된다.

## REFERENCES

- [1] Shin, S., Han, H., & Lee, S. H. (2021). Improved YOLOv3 with duplex FPN for object detection based on deep learning. *The International Journal of Electrical Engineering & Education*, 002072092098352. <https://doi.org/10.1177/0020720920983524>
- [2] Kirillov, A., He, K., Girshick, R., Rother, C., & Dollar, P. (2019). Panoptic Segmentation. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019–June, 9396–9405. <https://doi.org/10.1109/CVPR.2019.00963>
- [3] Shelhamer, E., Long, J., & Darrell, T. (2017). Fully

- Convolutional Networks for Semantic Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(4), 640-651.  
<https://doi.org/10.1109/TPAMI.2016.2572683>
- [4] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (Vol. 9351, Issue Cvd, pp. 234-241).  
[https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- [5] Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), 2481-2495.  
<https://doi.org/10.1109/TPAMI.2016.2644615>
- [6] Sovetkin, E., Achterberg, E. J., Weber, T., & Pieters, B. E. (2021). Encoder-Decoder Semantic Segmentation Models for Electroluminescence Images of Thin-Film Photovoltaic Modules. IEEE Journal of Photovoltaics, 11(2), 444-452.  
<https://doi.org/10.1109/JPHOTOV.2020.3041240>
- [7] Estrada, S., Conjeti, S., Ahmad, M., Navab, N., & Reuter, M. (2018). Competition vs. Concatenation in Skip Connections of Fully Convolutional Networks (pp. 214-222).  
[https://doi.org/10.1007/978-3-030-00919-9\\_25](https://doi.org/10.1007/978-3-030-00919-9_25)
- [8] Howard, A., Sandler, M., Chen, B., Wang, W., Chen, L.-C., Tan, M., Chu, G., Vasudevan, V., Zhu, Y., Pang, R., Adam, H., & Le, Q. (2019). Searching for MobileNetV3. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 1314-1324.  
<https://doi.org/10.1109/ICCV.2019.00140>
- [9] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016-Decem, 770-778.  
<https://doi.org/10.1109/CVPR.2016.90>
- [10] Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., & Schiele, B. (2016). The Cityscapes Dataset for Semantic Urban Scene Understanding. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016-Decem, 3213-3223.  
<https://doi.org/10.1109/CVPR.2016.350>
- [11] Paszke, A., Chaurasia, A., Kim, S., & Culurciello, E. (2016). ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation. 1-10.  
<http://arxiv.org/abs/1606.02147>
- [12] Treml, M., Arjona-medina, J., Unterthiner, T., Durgesh, R., Friedmann, F., Schuberth, P., Mayr, A., Heusel, M., Hofmarcher, M., Widrich, M., Nessler, B., & Hochreiter, S. (2016). Speeding up Semantic Segmentation for Autonomous Driving. NIPS 2016 Workshop MLITS, Nips, 1-7.  
<https://openreview.net/pdf?id=S1uHiFyyg%0Ahttps://openreview.net/forum?id=S1uHiFyyg>
- [13] Zheng, S., Jayasumana, S., Romera-Paredes, B., Vineet, V., Su, Z., Du, D., Huang, C., & Torr, P. H. S. (2015). Conditional Random Fields as Recurrent Neural Networks. 2015 IEEE International Conference on Computer Vision (ICCV), 2015 Inter, 1529-1537.  
<https://doi.org/10.1109/ICCV.2015.179>
- [14] Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 834-848.  
<https://doi.org/10.1109/TPAMI.2017.2699184>
- [15] Liu, Z., Li, X., Luo, P., Loy, C.-C., & Tang, X. (2015). Semantic Image Segmentation via Deep Parsing Network. 2015 IEEE International Conference on Computer Vision (ICCV), 2015 Inter, 1377-1385.  
<https://doi.org/10.1109/ICCV.2015.162>

신 석 용(Seokyong Shin)

[학사학위]



- 2019년 8월 : 광운대학교 컴퓨터공학과(공학사)
- 2020년 3월 ~ 현재 : 광운대학교 대학원 플라즈마바이오 디스플레이학과(석사 과정)
- 관심분야 : Segmentation, 딥러닝, 영상처리

· E-Mail : kistssy@gmail.com

이 상 훈(SangHun Lee)

[중신학위]



- 1983년 2월 : 광운대학교 응용전자공학과(공학사)
- 1987년 2월 : 광운대학교 전자공학과(공학석사)
- 1992년 2월 : 광운대학교 전자공학과(공학박사)
- 1990년 ~ 현재 : 광운대학교 정교수

- 2012년 2월 ~ 2013년 2월 : 광운대학교 정보통신처장
- 2013년 2월 ~ 2015년 2월 : 광운대학교 학생복지처장
- 2015년 2월 ~ 2017년 12월 : 한국융합학회 편집위원장
- 2018년 1월 ~ 2020년 1월 : 한국융합학회 회장
- 관심분야 : 영상인식, 3D 영상처리, 기계 학습, 영상처리와 생체 인식, IOT 영상처리, 딥러닝
- E-Mail : leesh58@kw.ac.kr

한 현 호(HyunHo Han)

[중신학위]

- 2009년 2월 : 광운대학교 컴퓨터공학과(공학사)
- 2011년 2월 : 광운대학교 정보콘텐츠 대학원 유비쿼터



스컴퓨팅학과 (공학석사)

· 2019년 8월 : 광운대학교 정보디스플레이학과(박사)

· 2020년 3월 ~ 현재 : 울산대학교 교양대학 교수

· 관심분야 : 무선 네트워크, 영상인식, 3D 영상처리, 기계 학습, 딥러닝

· E-Mail : hhhan@ulsan.ac.kr