

딥러닝 기반 거리 영상의 Semantic Segmentation을 위한 Atrous Residual U-Net

신석용¹, 이상훈^{2*}, 한현호³

¹광운대학교 플라즈마바이오디스플레이학과 석사과정, ²광운대학교 인제니움학부 교수, ³울산대학교 교양대학 교수

Atrous Residual U-Net for Semantic Segmentation in Street Scenes based on Deep Learning

SeokYong Shin¹, SangHun Lee^{2*}, HyunHo Han³

¹Master Student, Department of Plasma Bio Display, Kwangwoon University

²Professor, Ingenium College of Liberal Arts, Kwangwoon University

³Professor, College of General Education, University of Ulsan

요약 본 논문에서는 U-Net 기반의 semantic segmentation 방법에서 정확도를 개선하기 위한 Atrous Residual U-Net (AR-UNet)을 제안하였다. U-Net은 의료 영상 분석, 자율주행 자동차, 원격 감지 영상 등의 분야에서 주로 사용된다. 기존 U-Net은 인코더 부분에서 컨볼루션 계층 수가 적어 추출되는 특징이 부족하다. 추출된 특징은 객체의 범주를 분류하는 데 필수적이며, 부족할 경우 분할 정확도를 저하시키는 문제를 초래한다. 따라서 이 문제를 개선하기 위해 인코더에 residual learning과 ASPP를 활용한 AR-UNet을 제안하였다. Residual learning은 특징 추출 능력을 개선하고, 연속적인 컨볼루션으로 발생하는 특징 손실과 기울기 소실 문제 방지에 효과적이다. 또한 ASPP는 특징맵의 해상도를 줄이지 않고 추가적인 특징 추출이 가능하다. 실험은 Cityscapes 데이터셋으로 AR-UNet의 효과를 검증하였다. 실험 결과는 AR-UNet이 기존 U-Net과 비교하여 향상된 분할 결과를 보였다. 이를 통해 AR-UNet은 정확도가 중요한 여러 응용 분야의 발전에 기여할 수 있다.

주제어 : ASPP, 딥러닝, 인코더-디코더, 잔차 학습, Semantic segmentation, U-Net

Abstract In this paper, we proposed an Atrous Residual U-Net (AR-UNet) to improve the segmentation accuracy of semantic segmentation method based on U-Net. The U-Net is mainly used in fields such as medical image analysis, autonomous vehicles, and remote sensing images. The conventional U-Net lacks extracted features due to the small number of convolution layers in the encoder part. The extracted features are essential for classifying object categories, and if they are insufficient, it causes a problem of lowering the segmentation accuracy. Therefore, to improve this problem, we proposed the AR-UNet using residual learning and ASPP in the encoder. Residual learning improves feature extraction ability and is effective in preventing feature loss and vanishing gradient problems caused by continuous convolutions. In addition, ASPP enables additional feature extraction without reducing the resolution of the feature map. Experiments verified the effectiveness of the AR-UNet with Cityscapes dataset. The experimental results showed that the AR-UNet showed improved segmentation results compared to the conventional U-Net. In this way, AR-UNet can contribute to the advancement of many applications where accuracy is important.

Key Words : ASPP, Deep learning, Encoder-Decoder, Residual learning, Semantic segmentation, U-Net

*This article is extended and excerpted from the conference paper presented at the 7th ICSMB 2021. (Manuscript ID: 0055)

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

Received August 4, 2021

Revised September 17, 2021

Accepted October 20, 2021

Published October 28, 2021

1. 서론

최근, segmentation 및 객체 검출은 의료 영상 분석, 원격 감지 영상, 자율주행 자동차 등의 분야에서 중요하게 사용되며, 많은 응용 프로그램에서 핵심적인 역할을 한다[1,2]. segmentation에는 semantic segmentation, panoptic segmentation[3] 등 다양한 종류가 있다. 이 중, semantic segmentation은 이미지에 존재하는 객체들의 범주와 그 객체를 구성하는 픽셀을 분류하는 문제이다[1].

딥러닝 기술이 컴퓨터 비전에 도입되면서 semantic segmentation도 크게 발전하였다. 대부분의 최근 semantic segmentation 방법들은 convolutional neural networks (CNN)을 기반으로 연구되고 있다 [4]. 이러한 방법들은 대표적으로 fully convolutional networks (FCN)[5], U-Net[6], SegNet[7] 등이 있으며, 대부분 인코더-디코더 (Encoder-Decoder)[8] 구조로 설계되어 있다. 인코더는 이미지에서 객체의 범주를 분류하는데 필수적인 특징을 추출 및 압축하며, 디코더는 추출한 특징을 확장하여 segmentation map을 생성한다. segmentation map은 이미지에서 각 픽셀마다 예측한 범주를 표시한 것이다.

FCN은 classification 문제에서 사용하는 모델을 semantic segmentation에 적용하는 전략을 제시하였다. FCN은 segmentation 분야에 딥러닝을 접목할 수 있는 기반을 마련하였으며, 대다수의 연구가 FCN의 전략을 사용한다. 그러나 FCN은 디코더에 딥러닝이 적용되지 않아 객체가 멍둑하게 분할되는 문제가 있다. 이는 자율주행 자동차에서 도로 환경 인식을 낮추고 의료 영상에서 병변 예측이 잘못될 수 있다. 따라서 이 문제를 해결하려는 방법이 다수 연구되었다. U-Net은 디코더에서 concatenation[9] 연산과 컨볼루션으로 인코더의 중간 특징맵을 활용하였다. 그 결과 FCN과 비교하여 분할 정확도가 향상되었다. 그러나 U-Net은 인코더에서 컨볼루션 계층 수가 적어 추출되는 특징이 부족하여 분할이 부정확하거나 실패하는 문제가 있다. 따라서 본 논문에서는 분할 정확도를 개선하기 위해 Atrous Residual U-Net (AR-UNet)을 제안하였다. 실험은 Cityscapes 데이터셋에서 AR-UNet을 다른 기존 방법들과 비교하였다. 절제 연구는 AR-UNet에서 ASPP에 대해 분석하였다.

2. 관련 연구

2.1 U-Net

U-Net은 인코더-디코더 구조이며, U자 형태를 가진다. U-Net은 FCN의 단점을 해결하기 위해 lateral connection과 concatenation 연산을 사용하였다. Lateral connection은 인코더의 중간 특징맵을 디코더로 가져오는 과정이며, concatenation은 중간 특징맵과 디코더의 특징맵을 하나로 이어 붙이는 과정이다. Fig 1은 U-Net의 구조이며, 왼쪽 부분은 인코더이고 오른쪽은 디코더이다. 인코더는 컨볼루션 2개와 풀링 1개가 반복되며, 디코더는 concatenation 1개, transposed convolution 1개, 컨볼루션 2개로 구성된다. 이때, U-Net은 인코더에서 컨볼루션이 총 10개만 존재하며, 다른 네트워크 (Xception[10], ResNet[11] 등)과 비교하면 컨볼루션이 크게 적다. 이는 인코더에서 추출하는 특징이 부족하여 분할 정확도를 저하시킨다.

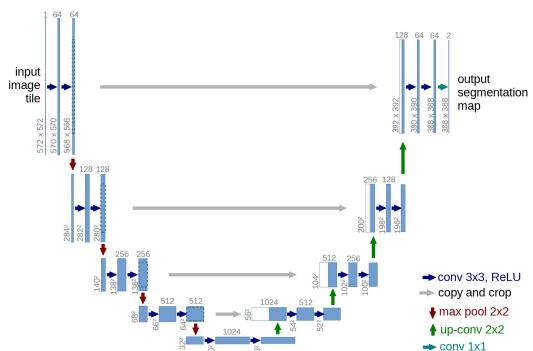


Fig. 1. U-Net architecture

2.2 Atrous Convolution

Atrous convolution은 DeepLab[12,13]에서 제안되었으며, 수용 영역을 확장하기 위한 방법이다. 수용 영역은 컨볼루션 커널이 한 번의 연산으로 처리할 수 있는 영역이다. 커널 크기를 확장하면 커널이 처리할 수 있는 영역, 즉 커널이 볼 수 있는 영역(field of view)가 확장되어 특징을 추출하기가 유리하다. 그러나 이는 매개변수(parameters)의 개수와 연산량이 증가하는 문제가 있다. 따라서 매개변수의 개수와 연산량을 유지하면서 수용 영역을 확장할 수 있는 atrous convolution이 연구되었다. atrous convolution은 rate에 따라 확장된 커널 사이 공간을 0으로 처리하여

효율적으로 넓은 수용 영역을 가질 수 있다. Fig. 2는 atrous convolution의 도식이다. Fig. 2 (a)는 rate가 1인 기존 컨볼루션이며, (b)는 rate가 2인 atrous convolution이다. Fig. 2 (b)에서 파란색과 주황색 커널만 실제 컨볼루션 연산에 사용하고, 나머지 커널은 0으로 처리한다. 따라서 연산량은 Fig. 2 (a)와 같지만 수용 영역은 5×5 로 확장된다. 식 1은 atrous convolution을 나타낸다. 식 1에서 x, y, w, i, r 은 각각 입력, 출력, 커널, 커널의 위치, rate이다.

$$y[i] = \sum_k x[i + r \cdot k]w[k] \quad (1)$$

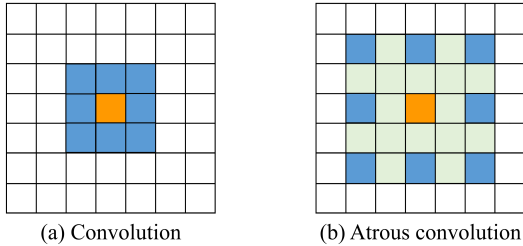


Fig. 2. Atrous convolution diagram

3. 제안하는 방법

AR-UNet은 특징 추출을 위한 인코더와 정확한 segmentation map의 생성을 위한 디코더로 구성되었다. Fig. 3은 AR-UNet의 구조이다. 인코더는 3×3 컨볼루션 2개, encoding block 16개, ASPP 1개로 구성된다. 디코더는 concatenation 연산 1개와 컨볼루션 2개로 구성된다. 마지막으로, 1×1 컨볼루션

을 처리하여 최종 segmentation map을 출력한다[5].

3.1 Residual Learning

Fig. 3과 같이 AR-UNet은 인코더-디코더 구조이다. AR-UNet은 인코더에서 백본 네트워크로 residual learning이 적용된 ResNet을 채택하였다. ResNet은 ResNet-18, ResNet-34, ResNet-50 등 여러 종류가 있다. 본 연구에서는 분할 속도와 성능의 trade-off에 따라 ResNet-34를 채택하였다.

Fig. 3에서 백본 네트워크는 특징맵의 크기에 따라 4개의 블록으로 구분하였다. 블록들은 각각 입력 이미지에 대비 1/1, 1/2, 1/4, 1/8 크기의 특징맵을 가진다. 각 encoding block의 구조는 모두 같으며, 식 2와 Fig. 4는 residual learning이 적용된 encoding block의 구조이다.

$$H(x) = F(x) + x \quad (2)$$

AR-UNet은 인코더의 컨볼루션 수가 총 34개로 기존 U-Net보다 약 3배 많다. 이 구조는 인코더의 특징 추출 능력을 향상시켜 분할 성능을 개선한다. 그러나 인코더의 계층 수가 증가하면서 특징 손실과 기울기 소실 문제가 발생한다. 따라서 AR-UNet은 residual learning을 적용하였다. Residual learning은 연속적인 컨볼루션의 특징 추출 과정에서 발생하는 특징 손실을 방지한다. 또한, 네트워크의 계층 수가 증가할수록 발생할 확률이 높은 기울기 소실 문제를 방지한다.

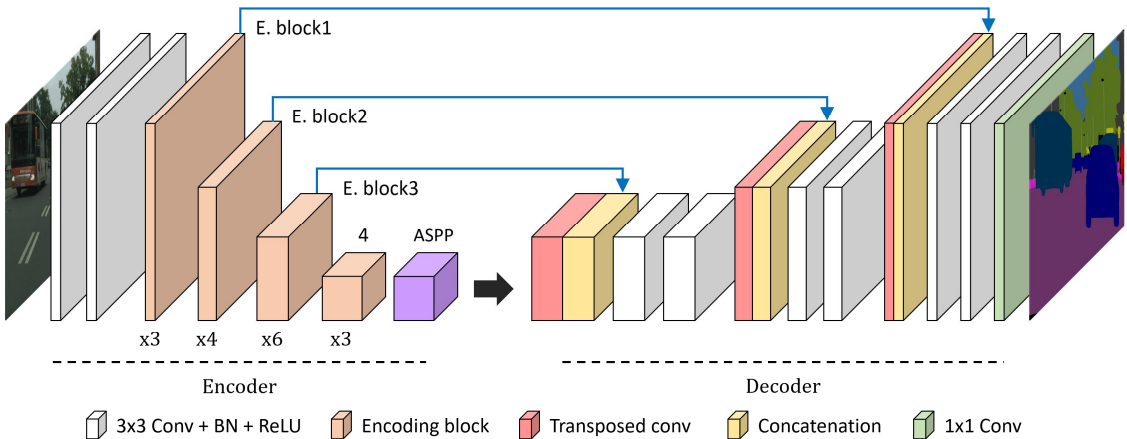


Fig. 3. Proposed method architecture

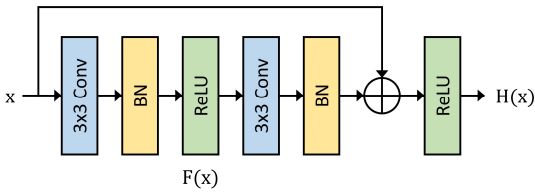


Fig. 4. The details of encoding block including residual learning

일반적인 백본 네트워크는 encoding block을 처리하기 전에 stride 2를 가진 컨볼루션 또는 풀링을 수행하여 특징맵의 해상도를 축소시킨다. 또한, 대다수의 semantic segmentation 방법들도 이러한 과정을 그대로 수행한다. 그러나 이는 특징맵에 포함된 공간 정보가 손실되는 문제를 초래한다. 따라서 AR-UNet은 ResNet의 7×7 컨볼루션과 max-pooling을 2개의 3×3 컨볼루션으로 대체하여 특징맵의 해상도를 축소하지 않고 특징 추출 전처리만 수행하였다. 추가로 기술하지 않으면 컨볼루션 계층 뒤에 배치 정규화(Batch normalization, BN)와 ReLU 활성화 함수가 포함된다.

3.2 Atrous Spatial Pyramid Pooling

일반적인 CNN은 수용 영역을 확장하고 연산량을 줄이기 위해 연속적인 컨볼루션 과정에서 max-pooling 또는 stride를 사용하여 특징맵의 공간 해상도를 축소한다. 그러나 이러한 과정은 깊은 계층에서 공간 정보를 잃게 만들어 semantic segmentation과 같은 픽셀 단위 예측에서 정확도를 저하시키는 문제가 있다. 따라서 AR-UNet은 특징맵의 해상도를 축소시키지 않고 수용 영역을 확장할 수 있도록 인코더의 마지막에 atrous spatial pyramid pooling (ASPP)를 활용하였다. ASPP는 DeepLab에서 처음 제안되었으며,

rate가 서로 다른 여러 atrous convolution을 병렬로 사용하였다. 이는 CNN이 수용 영역의 확장과 동시에 다양한 크기의 수용 영역을 가질 수 있도록 하였다. Fig 5는 ASPP의 구조이다.

Fig. 5에서 기호 C, H, W, r은 각각 채널 수, 높이, 너비, rate를 나타낸다. ASPP는 총 5개의 분기로 병렬 처리되며 각각 1×1 컨볼루션, 3×3 컨볼루션 rate (3, 6, 9), global average pooling (GAP) + 1×1 컨볼루션 + 업샘플링이다. 그런 다음, 각 분기의 특징맵을 concatenation하고 1×1 컨볼루션으로 결합된 특징맵을 재구성하였다. Fig. 5에 표시한 것처럼 ASPP의 입력 및 출력 특징맵의 해상도와 채널 수는 서로 동일하다. 식 3은 ASPP를 나타낸다.

$$ASPP = W_1^1(W_1^1(x) \cdot W_3^3(x) \cdot W_3^6(x) \cdot W_3^9(x) \cdot \theta W_1^1 \sigma(x)) \quad (3)$$

여기서 $x, W_{kernel}^{rate}, \theta, \sigma$ 는 각각 입력, 컨볼루션, GAP, 업샘플링을 나타낸다. DeepLab에서 제시한 기존 ASPP는 rate가 (6, 12, 18)로 설정되었다. 그러나 이러한 높은 rate 값은 컨볼루션 연산을 수행하지 않는 픽셀을 증가시켜 체커보드 아티팩트를 유발할 가능성이 높아진다. 체커보드 아티팩트는 인접한 픽셀이 처리되지 않아 특징맵에 격자 무늬가 나타나는 현상이다. 따라서 AR-UNet은 Fig 5와 같이 rate를 (3, 6, 9)로 설정하여 이러한 문제를 최소화하였다. 또한 DeepLab은 ASPP의 출력 특징맵의 채널 수를 절반으로 줄였다. 그러나 AR-UNet은 이전 encoding block에서 추출한 특징들을 최대한 보존하기 위해 해당 특징맵의 채널 수를 입력 특징맵의 채널 수와 일치시켰다.

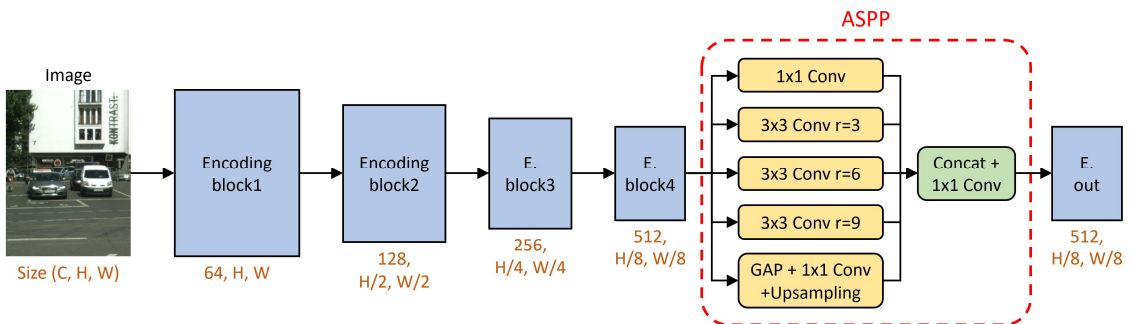


Fig. 5. Atrous spatial pyramid pooling (ASPP)

3.3 정확한 분할을 위한 디코더

AR-UNet의 디코더는 기존 U-Net과 유사한 구조로 설계하였다. Fig 6은 디코더의 구조이다. 첫 번째로, stride 2를 가진 2×2 transposed convolution으로 특징맵을 2배 확장하였다. 그리고 확장된 특징맵과 해상도가 같은 인코더의 중간 특징맵을 concatenation하였다. 두 번째로, 결합된 특징맵에 3×3 컨볼루션, 배치 정규화, ReLU를 2번 수행하여 특징맵을 재구성하였다. 특히, AR-UNet은 기존 U-Net과 달리 디코더에서 배치 정규화를 사용하여 학습 효율성을 높였다. 식 4와 식 5는 각각 concatenation과 디코더를 나타낸다.

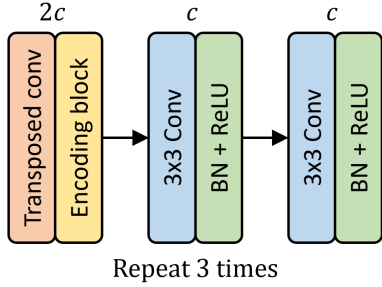


Fig. 6. Decoder architecture

$$\text{Concat}(2c, h, w) = [U(x)_{c,h,w}; E(y)_{c,h,w}] \quad (4)$$

$$\text{Decode} = f_{\text{conv}}^{3 \times 3}(f_{\text{conv}}^{3 \times 3}(\text{Concat}(T(x); E(y)))) \quad (5)$$

여기서 $E(y)$ 는 encoding block이 생성한 중간 특징맵이며, $U(x)$ 는 transposed convolution으로 확장된 특징맵이다. 그리고 h, w, c 는 각각 높이, 너비, 채널 수를 나타낸다.

3.4 손실 함수

AR-UNet은 교차 엔트로피(Cross-entropy) 손실 함수를 사용하였다. 이는 정답 이미지 (groundtruth)와 segmentation map을 비교하여 오차를 출력한다. 학습은 손실 함수가 출력한 오차를 감소시키기 위해 가중치를 조절하는 과정이다. 식 6은 교차 엔트로피 손실 함수를 나타낸다. 식 6에서 k, t_k, y_k 는 각각 학습 데이터의 k 번째 요소, groundtruth, AR-UNet의 출력이다.

$$CE = \sum_k t_k \log_e(y_k) \quad (6)$$

4. 실험 및 결과

4.1 실험 환경

AR-UNet은 다음과 같은 환경에서 실험하였다. 하이퍼파라미터는 배치 크기(batch size)를 4, 에폭(epoch)을 100, 학습률(learning rate)을 0.001로 설정하였다. 또한 AdamW[14] 최적화기와 poly learning rate scheduler [11]를 사용하였다. 데이터 증강 (data augmentation) 기법은 random horizontal flip, random crop, color jitter를 적용하였다. 학습 이미지 크기는 512×1024 이다. 하드웨어와 소프트웨어는 다음과 같다: AMD Ryzen 3700x, NVIDIA RTX 3090 2장, PyTorch 1.9

실험 결과는 mean intersection over union (MIoU)을 사용하여 평가 및 비교하였다. MIoU는 정확도를 위한 평가 지표이며 식 7과 같다. 식 7에서 k, TP, FN, FP 는 각각 클래스 번호, true positive, false negative, false positive를 나타낸다.

$$MIoU = \frac{1}{k} \sum_{i=0}^k \frac{TP}{TP + FP + FN} \quad (7)$$

4.2 Cityscapes

Cityscapes[15] 데이터셋은 semantic segmentation 연구에 널리 사용된다. 이 데이터셋은 50개의 도시에서 수집한 5000장의 거리 영상이 포함되어 있다. 그리고 2975장의 학습 이미지, 500장의 검증 이미지, 1525장의 테스트 이미지로 분할되어 있다.

Table 1. Results on Cityscapes dataset.

Method	Params (M)	MIoU (%)
FCN-8s [5]	35.3	65.3
U-Net [6]	31.0	57.3
SegNet [7]	29.5	70.4
DeepLabv2 [12]	262.1	63.1
ICNet [16]	26.5	70.6
DeepLabv3 [13]	58.0	72.0
AR-UNet	33.3	72.1

본 논문은 AR-UNet과 여러 기존 방법들을 매개변수 개수 (Params)와 정확도 (MIoU) 관점에서 비교하였다. Table 1은 Cityscapes 데이터셋에서 AR-UNet

과 기존 방법의 정확도와 매개변수 개수를 나타낸다. AR-UNet의 MIoU는 약 72.1%로 다른 기존 방법들보다 높은 성능을 보였다. U-Net과 비교하여 MIoU가 약 14.8%, DeepLabv3와 비교하여 약 0.1% 향상되었다. AR-UNet의 매개변수 개수는 FCN, U-Net, SegNet과 비교하여 큰 차이를 보이지 않았으며, DeepLabv2, DeepLabv3와 비교하여 적은 수를 가졌다. 이는 비슷한 연산량으로 더 높은 정확도를 달성함을 증명하였다.

Fig. 7은 Cityscapes 데이터셋을 사용하여 여러 객체에서 AR-UNet과 기존 U-Net의 분할 결과를 비교하였다. 그림에서 주황색 상자는 기존 U-Net보다 정확한 분할 결과를 보인 부분이다. 기존 U-Net은 객체의 범주를 분류하는데 필요한 특징이 부족하여 버스, 사람 등 다양한 객체에서 분할 정확도가 낮은 결과를 보였다. 특히, 객체의 일부가 분할되지 않는 문제가 있다. 그러나 AR-UNet은 residual learning과 ASPP를 활용하여 추가 특징을 추출했으므로 U-Net과 비교하여 더 정확한 분할 결과를 보였다. 또한 주황색 상자 영역에서 여러 객체의 누락되는 픽셀을 최소화하여 분할하였다.

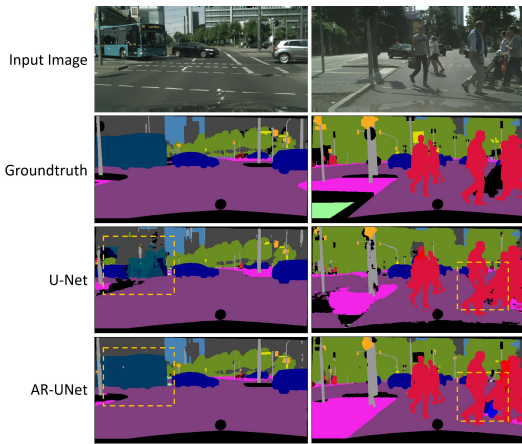


Fig. 7. Qualitative comparisons of multiple object on the Cityscapes dataset.

4.3 절제 연구

Table 2는 AR-UNet에 사용된 ASPP의 정확도 기여와 rate에 따른 정확도의 변화를 분석하였다. “BS”는 baseline으로 AR-UNet에서 ASPP만 제거한 모델이다. “BS + ASPP”는 baseline에 ASPP가 추가된 모델이며 atrous convolution에 설정된 rate를 순차적으로 표기하였다. 표의 결과에서 ASPP가 분할 정확도를

높이는데 큰 기여를 한다는 것을 알 수 있다. 또한 rate를 (3, 6, 9)로 설정한 때가 가장 높은 정확도를 보였다. Rate는 낮을수록 수용 영역은 좁지만, 실제 연산을 수행하는 픽셀의 수가 많아 체커보드 아티팩트가 거의 없다. 반대로 rate가 높을수록 수용 영역은 넓지만 체커보드 아티팩트가 발생하기 쉽다. 따라서 최적의 rate 값을 찾는 것이 중요하며, 실제 연구를 통해 이를 증명하였다.

Table 2. Ablation study for ASPP analysis on Cityscapes dataset.

Method	Params	MIoU
BS	24.3 M	68.2 %
BS+ASPP (r=2, 4, 6)	33.3 M	71.5 %
BS+ASPP (r=3, 6, 9)	33.3 M	72.8 %
BS+ASPP (r=4, 8, 12)	33.3 M	71.8 %

5. 결론

본 논문은 U-Net 기반 semantic segmentation 방법에서 정확성을 개선하기 위해 residual learning과 ASPP를 활용한 Atrous Residual U-Net (AR-UNet)을 제안하였다. 기존 U-Net은 인코더에서 추출되는 특징이 부족하여 분할 정확도를 저하시키는 문제를 초래한다. 이 문제는 응용 프로그램의 신뢰도를 낮추게 된다. 따라서, AR-UNet은 기존 U-Net에 residual learning과 ASPP를 활용하여 특징 추출 능력을 개선하였다. 실험 결과는 AR-UNet이 기존 방법과 비교하여 분할 정확도가 향상되었음을 보였다. 이를 통해 AR-UNet은 정확도가 중요한 원격 감지 영상, 자율주행 자동차 등의 응용 분야에서 활용될 수 있다. 향후 연구는 모바일 장치에서 실행할 수 있는 모델을 설계할 필요가 있다고 사료된다.

REFERENCES

- [1] S. Y. Shin, S. H. Lee & H. H. Han (2021). A Study on Residual U-Net for Semantic Segmentation based on Deep Learning. *Journal of Digital Convergence*, 19(6), 251-258. DOI : 10.14400/JDC.2021.19.6.251
- [2] S. Shin, H. Han & S. H. Lee. (2021). Improved YOLOv3 with duplex FPN for object detection

- based on deep learning. *The International Journal of Electrical Engineering & Education*, 002072092098352.
DOI : 10.1177/0020720920983524
- [3] A. Kirillov, K. He, R. Girshick, C. Rother & P. Dollar. (2019). Panoptic Segmentation. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019-June, 9396-9405.
DOI : 10.1109/CVPR.2019.00963
- [4] S. Y. Shin, S. H. Lee & J. S. Kim (2021, January) Modified Encoder-Decoder model of U-Net for Semantic Segmentation based on Deep Learning. *The 7th International Conference on Small & Medium Business*. (pp.379-380). Jeju : SMB.
- [5] E. Shelhamer, J. Long & T. Darrell. (2017). Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4), 640-651.
DOI : 10.1109/TPAMI.2016.2572683
- [6] O. Ronneberger, P. Fischer & T. Brox. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer.
- [7] V. Badrinarayanan, A. Kendall & R. Cipolla. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12), 2481-2495.
DOI : 10.1109/TPAMI.2016.2644615
- [8] E. Sovetkin, E. J. Achterberg, T. Weber & B. E. Pieters. (2021). Encoder-Decoder Semantic Segmentation Models for Electroluminescence Images of Thin-Film Photovoltaic Modules. *IEEE Journal of Photovoltaics*, 11(2), 444-452.
DOI : 10.1109/JPHOTOV.2020.3041240
- [9] S. Estrada, S. Conjeti, M. Ahmad, N. Navab & M. Reuter. (2018). Competition vs. Concatenation in Skip Connections of Fully Convolutional Networks, *In International Workshop on Machine Learning in Medical Imaging* (pp. 214-222).
DOI : 10.1007/978-3-030-00919-9_25
- [10] F. Chollet. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017-Janua, (pp. 1800-1807).
DOI : 10.1109/CVPR.2017.195
- [11] K. He, X. Zhang, S. Ren & J. Sun. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [12] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy & A. L. Yuille. (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.
DOI : 10.1109/TPAMI.2017.2699184
- [13] L. Chen, G. Papandreou, F. Schroff & H. Adam. (2017). Rethinking Atrous Convolution for Semantic Image Segmentation. *arXiv preprint arXiv:1706.05587*
- [14] I. Loshchilov & F. Hutter (2019). Decoupled Weight Decay Regularization. *7th International Conference on Learning Representations, ICLR 2019*. . *arXiv preprint arXiv:1711.05101*.
- [15] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth & B. Schiele. (2016). The Cityscapes Dataset for Semantic Urban Scene Understanding. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016-Decem, (pp. 3213-3223).
DOI : 10.1109/CVPR.2016.350
- [16] H. Zhao, X. Qi, X. Shen, J. Shi & J. Jia. (2018). ICNet for Real-Time Semantic Segmentation on High-Resolution Images. *In European Conference on Computer Vision* (pp. 418-434).
DOI : 10.1007/978-3-030-01219-9_25

신 석 용 (SeokYong Shin)

[학생회원]



- 2019년 8월 : 광운대학교 컴퓨터 공학과(공학사)
- 2020년 3월 ~ 현재 : 광운대학교 대학원 플라즈마바이오 디스플레이학과(석사 과정)
- 관심분야 : Segmentation, 딥러닝, 영상처리
- E-Mail : kistssy@gmail.com

이 상 훈 (SangHun Lee)

[종신회원]



- 1983년 2월 : 광운대학교 응용전자 공학과(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)

[종신회원]



- 2009년 2월 : 광운대학교 컴퓨터 공학과(공학사)
- 2011년 2월 : 광운대학교 정보콘텐츠 대학원 유비쿼터스컴퓨팅학과(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