

효율적인 비정형 도로영역 인식을 위한 Semantic segmentation 기반 심층 신경망 구조

박세진¹ · 한정훈¹ · 문영식^{2*}

Efficient Deep Neural Network Architecture based on Semantic Segmentation for Paved Road Detection

Sejin Park¹ · Jeong Hoon Han¹ · Young Shik Moon^{2*}

¹Ph.D. Student, Department of Computer Science and Engineering, Hanyang University, Ansan, 15588 Korea

^{2*}Professor, Department of Computer Science and Engineering, Hanyang University, Ansan, 15588 Korea

요 약

컴퓨터 비전 시스템의 발달로 보안, 생체인식, 의료영상, 자율주행 등의 분야에 많은 발전이 있었다. 자율주행 분야에서는 특히 딥러닝을 이용한 객체인식, 탐지 기법이 주로 사용되는데, 자동차가 갈 수 있는 영역을 판단하기 위한 도로영역 인식이 특히 중요한 문제이다. 도로 영역은 일반적인 객체탐지에서 활용되는 사각영역인식과는 달리 비정형적인 형태를 띠므로, ROI 기반의 객체인식 구조는 적용할 수 없다. 본 논문에서는 Semantic segmentation 기법을 사용한 비정형적인 도로영역 인식에 맞는 심층 신경망 구조를 제안한다. 또한 도로영역에 특화된 네트워크 구조인 Multi-scale semantic segmentation 기법을 사용하여 성능이 개선됨을 입증하였다.

ABSTRACT

With the development of computer vision systems, many advances have been made in the fields of surveillance, biometrics, medical imaging, and autonomous driving. In the field of autonomous driving, in particular, the object detection technique using deep learning are widely used, and the paved road detection is a particularly crucial problem. Unlike the ROI detection algorithm used in general object detection, the structure of paved road in the image is heterogeneous, so the ROI-based object recognition architecture is not available. In this paper, we propose a deep neural network architecture for atypical paved road detection using Semantic segmentation network. In addition, we introduce the multi-scale semantic segmentation network, which is a network architecture specialized to the paved road detection. We demonstrate that the performance is significantly improved by the proposed method.

키워드 : 컴퓨터 비전, 딥러닝, 의미적 분할, 자율 주행, 도로영역 인식

Keywords : Computer vision, Deep learning, Semantic segmentation, Autonomous driving, Road detection

Received 13 July 2020, Revised 16 July 2020, Accepted 16 July 2020

* Corresponding Author Young Shik Moon (E-mail: ysmoon@hanyang.ac.kr, Tel: +82-31-400-4049)

Professor, Department of Computer Science and Engineering, Hanyang University, Ansan, 15588 Korea

Open Access <http://doi.org/10.6109/jkiice.2020.24.11.1437>

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서론

최근 자율주행에 대한 산업계와 학계의 관심이 매우 높다. 구글, 애플, 엔비디아와 같은 IT 기업 뿐 아니라, 제너럴 모터스, 아우디, 테슬라와 같은 자동차 회사들도 앞다투어 자율주행 자동차를 만들기 위한 연구를 지속적으로 수행하고 있다. 자율주행의 발달에 큰 영향을 끼친 기술 중 대표적인 것으로 컴퓨터 비전과 딥러닝을 들 수 있다.

컴퓨터 비전 시스템의 발달은 2012년 이후 딥러닝의 급격한 도약과 함께 비약적으로 이루어져 왔다[1], [2],[3],[4]. 생체인식, 의료영상, 자율주행 등의 분야에서 특히 인간의 인식 수준에 이르는 성과들을 보였는데, 자율주행 분야에서는 특히 딥러닝을 이용한 객체인식 및 탐지 기법이 주로 사용되었다. 자동차가 갈 수 있는 영역을 판단하기 위해, 자동차 전방카메라를 이용해 촬영된 영상에서 도로영역을 인식하는 것이 특히 중요한 문제이다[5], [6], [7]. 정밀하고 정확한 도로영역인식을 통해 자율주행을 위한 주행 방향 추론이 가능하고, 긴급 상황에서 급제동 등의 대처를 할 수 있으므로, 향상된 도로영역 인식 개발은 시급한 문제이다.

도로영역은 비정형적인 형태를 띠므로, 일반적인 객체 탐지에서 활용되는 Faster RCNN [4], You Only Look Once (YOLO) [8]와 같은 ROI 기반의 객체인식 구조는 적용할 수 없다. 비정형적인 구조의 객체를 인식하기 위한 딥러닝 네트워크의 대표적인 종류로 Fully Convolutional Network (FCN)를 들 수 있다[9]. FCN의 특징은 입력과 출력이 같은 크기의 영상이라는 것이며, 통상적으로 입력은 일반적인 흑백영상 혹은 RGB 영상이 되고, 출력은 입력영상을 의미적으로 분할한 객체의 Class map이 된다. 이러한 기법을 Semantic segmentation 이라 칭하는데, Faster RCNN, YOLO 등이 가지는 사각영역 객체 인식에 제한 되는 문제를 극복한 대표적인 딥러닝 방법이다.

한편 도로영상에 존재하는 객체들은 그 크기가 다양하고, 서로 다른 종류의 것들이 동시에 중첩해서 나타난다는 특징이 있다[10],[11]. 또한 도로영역은 자동차의 전방으로 가까이 있는 영역들과 멀리 있는 영역들을 동시에 하나의 대상으로 인식해야 하는 특수한 문제를 가진다[12]. 전통적인 객체인식 알고리즘은 대상 객체의 크기에 따라 서로 다른 알고리즘을 적용하거나, 다양한 크기의 영상을 동시에 다루는 기법을 사용한다.

본 논문에서는 Semantic segmentation 기법을 사용한 비정형적이고, 다양한 크기의 영역을 동시에 가지는 도로영역 인식에 맞는 심층 신경망 구조를 제안한다. 또한 도로영역에 특화된 네트워크 구조인 Multi-scale semantic segmentation 모델을 사용하여 성능이 개선됨을 입증하였다.

II. 관련 연구

2.1. Semantic segmentation network

일반적인 Convolutional Neural Network (CNN) 은 Down-pooling을 사용하여 회전에 강건하면서, 특징공간의 크기를 줄이는 효과를 내는 구조를 가지고 있다. 이러한 구조는 원본 영상의 해상도를 줄어뜨리게 하고, 객체가 존재하는 영역의 위치정보를 잃게 만드는 단점이 있는데, 물체가 있는 위치를 찾아야 하는 Semantic segmentation 기법에서 이러한 점은 치명적인 단점이 된다. 이를 극복하기 위해 SegNet [13]과 FCN [9]이 제안되었다.

대표적인 Semantic segmentation network 인 SegNet 이나 FCN 은 Un-pooling이나 Up-sampling을 사용하여 Down-pooling 된 특징들의 위치를 다시 복원한다. 특히 FCN의 Up-sampling은 한 번의 양방향 선형보간연산을 사용하여 해상도가 줄어든 영상을 원본영상으로 확대하는 방법을 사용한다. 그러나 결과의 정밀도가 떨어져 이를 보완하기 위해 여러 번의 Up-sampling 을 하면서 객체가 존재하는 위치를 학습하도록 파라미터를 가지는 전치 합성곱 (Transposed convolution) 연산을 사용한 Deconvolutional network이 사용된다. 그러한 네트워크 형태에서는 작은 객체부터 큰 객체까지 다양한 크기의 객체 위치를 추적하기 위해 Up-sampling 이 여러 번 이루어지는 구조를 가지지만, 점진적으로 커지는 특징 때문에 매우 큰 객체와 작은 객체를 동시에 인식하기 어렵

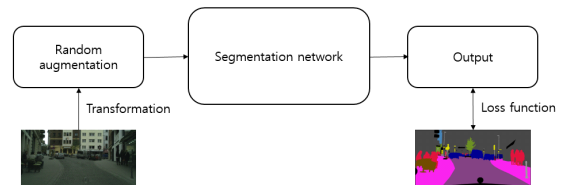


Fig. 1 Schematic diagram of algorithm

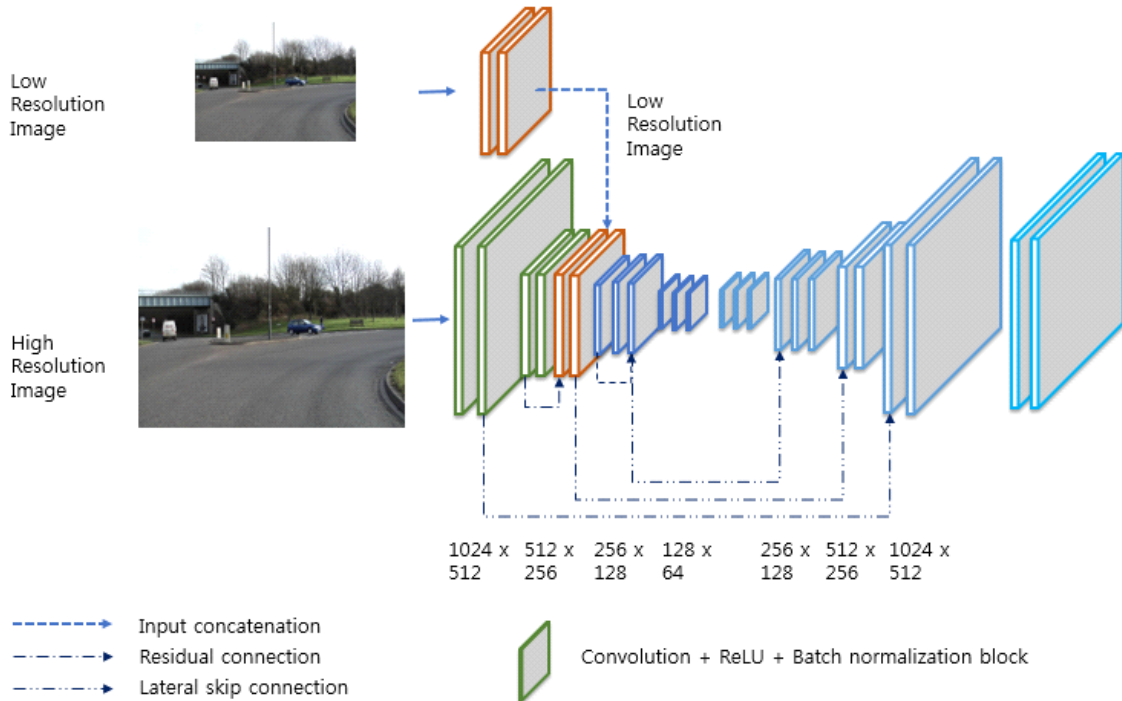


Fig. 2 Architecture of semantic segmentation network for paved road detection

다는 단점이 있다. 한편 [14]에서는 같은 연산량으로 더 큰 특징을 잡아낼 수 있도록, 일반적인 convolution 사이에 공간을 집어넣는 Atrous convolution을 제안하였다. 하지만, Atrous convolution을 사용하는 구조 역시 큰 객체와 작은 객체를 동시에 인식하기는 어렵다는 단점을 가진다.

2.2. Object detection network

자동차, 보행자와 같은 상대적으로 객체의 크기가 작은 대상들은 인식의 효율성, 위치 정확성의 개선 등을 이유로 사각 영역기반 객체인식기법들이 주로 사용된다. 그러한 딥러닝 모델에는 대표적으로 Faster RCNN, YOLO 와 같은 것들이 있다.

Faster RCNN은 사각영역 내의 객체 인식을 위해 CNN 의 마지막 Convolution layer에 후보영역을 인식하는 Region proposal network (RPN) 와 인식 대상 객체를 다시 재분류하는 Region classification network (RCN) 을 연결하는 방식으로 구성된다. 이 방식에서는 하나의 대상객체의 주변부에서 여러 개의 사각인식영역이 출력

될 수 있는 문제를 극복하고자, Non-maximum suppression layer 를 사용하였다. RPN 과 RCN 으로 나누어진 2단계 학습기법을 사용하여 거짓양성 (False positive) 를 억제할 수 있다는 것이 장점이다.

한편 YOLO 모델은 예측하고자 하는 영상을 서로 중첩되지 않는 사각영역으로 나누고, 각 영역마다 하나의 객체를 예측한다. 이후 미리 설정된 Bounding box 를 통해 객체의 정확한 위치와 크기를 파악한다. YOLO 모델은 전체 영상에서 한 번의 평가를 통해서 직접적으로 bounding box들을 예측하고, class의 확률을 계산하므로, 다른 실시간 탐지모델보다 높은 정확도를 보이고, Faster RCNN 과 같은 2단계 학습기법에 거의 근접한 성능을 보이면서도 155 fps의 빠른 수행 속도를 가진다는 것이 장점이다.

이러한 Object detection network 들은 2개 이상의 중첩된 객체를 분류해 낼 수 있다는 장점이 있는 반면, 사각영역이 아니면서 매우 큰 범위를 가지는 도로영역과 같은 객체 인식에서는 적절치 않다는 단점이 있다.

2.3. Instance segmentation network

Semantic segmentation network 의 단점은 2개 이상 중첩된 객체를 분리하여 인식하지 못한다는 점이다. 이것을 개선하기 위해 Object detection network 과 Semantic segmentation network를 결합함으로써 서로 중첩되어 있는 객체를 각기 다른 객체로 구별하는 모델이 제안되었는데, 이것을 Instance segmentation network 이라고 부른다. 대표적인 기법으로 Mask-RCNN [15] 이 있는데, 이것은 Faster R-CNN 에 존재하는 Bounding box 인식을 위한 네트워크 경로에 병렬로 object mask prediction 경로를 추가한 것이다. 또한 Faster R-CNN에 있는 ROI Pooling 계층 대신에 ROI Align 계층을 사용한다. Instance segmentation network 은 최종 결과물이 segmentation mask 이긴 하지만, 사각영역을 우선 인식한 후 인식된 영역 내에서 다시 한번 mask 를 찾는 방식이므로, 도로영역과 같이 사각 영역으로 정의될 수 없는 비정형 객체 인식 문제에서는 적합하지 않다.

앞에서 설명된 Semantic segmentation network, Object detection network, Instance segmentation network중에서 도로영역 인식에 가장 적합한 방법은 인접한 도로를 여러 개의 개별 instance 로 나눌 필요가 없고, 그 형태가 비정형적인 도로영역의 특성을 가장 잘 반영하는 Semantic segmentation network 이다. 기존의 도로영역 인식 신경망은 U-net 이나 DeepLab 등을 이용한 단일 해상도 영상을 입력으로 하는 Semantic segmentation 구조가 주로 쓰였다.

III. Semantic Segmentation for Paved Road Detection

제안된 방법은 Semantic segmentation network을 사용하여, 자동차 전방 카메라로부터 수집된 영상을 입력으로 받아 도로영역을 출력하는 구조를 가지고 있다. 제안된 방법의 알고리즘의 전체적인 구조는 그림 1 과 같으며, 입력영상, 데이터 증강, Semantic segmentation network, 출력, 정답영상으로 이루어져 있다.

도로영역은 고해상도의 영상의 근거리 영역부터 저해상도 영상의 원거리 영역으로 점진적으로 변해가는데, 이러한 해상도의 변화는 한 해상도 입력만으로 Semantic segmentation network을 구성했을 때 성능 저

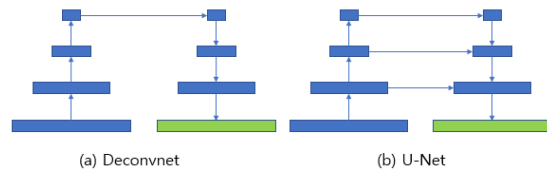


Fig. 3 Illustration of Deconvnet and U-Net architecture

하를 가져 온다는 단점이 있다. 하나의 네트워크 구조를 두 개의 서로 다른 해상도를 가지는 두 개의 네트워크 구조로 나누고, 다시 병합하는 Multi-scale Semantic segmentation network 가 이것을 극복하기 위한 대안이 될 수 있다. 제안된 Semantic segmentation network의 상세 구조는 그림 2 에 나타내었다. 기본적으로 제안된 네트워크 구조는 ResNet[16]과 U-Net[17]의 구조를 따른다. 즉, Skip connection 과 Lateral skip connection 구조를 채용하여 매우 깊은 구조의 네트워크 학습이 효과적으로 이루어지도록 하였다.

3.1. 주행가능 영역 인식을 위한 Semantic segmentation network

본 논문에서는 Deconvolutional network (Deconvnet)의 형태를 기본 Segmentation network 구조로 사용하였다. Deconvnet 은 일반적인 CNN 이 pooling 계층을 통해 해상도를 절반씩 줄여나가기 때문에 마지막 convolution 계층의 출력이 원본 영상보다 작아진다는 문제를 극복하기 위해 Deconvolution을 사용한다. Deconvolution 은 convolution 의 반대과정으로 입력영상의 어떤 부분으로부터 특정 feature 가 출력되었는지 역으로 연산하는 과정이라고 볼 수 있다. CNN 의 pooling을 통해 작아진 크기의 영상을 pooling 계층의 횟수만큼 Deconvolution 을 반복하면, (이때 Deconvolution stride 는 pooling 의 것과 같아야함. 통상적으로는 2를 사용) 최종출력이 원본영상의 크기와 같아지도록 만들 수 있다. 특정 객체가 존재하는 pixel에 mask를 적용한 class map을 정답레이블로 하고, 최종 출력을 담당하는 convolution 계층에 Softmax 연산을 적용하며, Cross entropy loss 등을 loss function으로 사용한다. Semantic segmentation의 성능 개선을 위해 기존에 제안된 Focal loss 와 Dice loss 등을 사용하여 추가적인 실험을 진행하였다. Focal loss는 상대적으로 영역구분이 쉬운 픽셀들에 대한 loss 값을 저감시키는 함수 수식 1을 사용하여, 네트워크가 어려운 샘플

플에 더 집중하도록 만들었다. 수식 1에서 p 는 네트워크의 출력, \hat{p} 는 정답 레이블, α 는 balancing factor를 의미한다.

$$FL(p, \hat{p}) = -(\alpha(1-\hat{p})^\gamma p \log(\hat{p}) + (1-\alpha)\hat{p}^\gamma (1-p) \log(1-\hat{p})) \quad (1)$$

Dice loss 는 데이터 불균형 특징이 존재하는 semantic segmentation network 구조에 많이 사용되는 손실함수이며 수식 2 와 같이 정의된다. 수식 2에서 p 는 네트워크의 출력, \hat{p} 는 정답 레이블을 의미한다.

$$DL(p, \hat{p}) = 1 - \frac{2 \sum_i^N p_i \hat{p}_i}{\sum_i^N p_i^2 + \sum_i^N \hat{p}_i^2} \quad (2)$$

3.2. Multi-scale Semantic segmentation network

도로영역에 거리에 따라 다르게 존재하는 저해상도와 고해상도 영상을 동시에 활용하기 위하여, 제안된 Semantic segmentation network 은 두 가지 서로 다른 해상도의 영상을 동시에 입력으로 한다. 서로 다른 크기의 입력영상을 합성하기 위해 더 큰 해상도의 영상에 convolution 계층과 pooling 계층을 순차적으로 적용하여 작은 크기의 영상으로 만들면서, 큰 해상도 영상의 특징을 추출하는 역할을 수행하도록 하였다. Pooling 계층의 출력을 작은 해상도 영상의 입력과 채널별로 연결(concatenation) 함으로써 두 해상도의 영상을 합성할 수 있고, 이후 다시 convolution layer 를 연속적으로 배치 함으로써, 두개의 입력에 단일 CNN 을 적용하는 효과를 내도록 하였다. 이러한 방식을 통해, 멀리 있는 객체는 크기가 작아 세밀한 영상 정보가 부족하므로 원본해상도를 그대로 사용하여 인식해야 하지만, 가까이 있는 객체는 상대적으로 정보가 많아 해상도를 줄여서 작은 크기의 영상으로 만들어도 정보손실이 덜하다는 사전 지식을 하나의 단일 네트워크에 반영할 수 있다.

일반적인 Semantic segmentation network은 그림 3의 (a)와 같은 Deconvolution 구조를 가진다. [16]는 영상의 전반적인 컨텍스트 정보를 얻기 위해 하위층 네트워크와 상위층 네트워크가 대칭형태로 구성되어 있고, 상위층 네트워크에서 좀 더 정확한 위치 정확도를 가지도록 하기 위해 하위계층 네트워크의 데이터를 복사해서 불

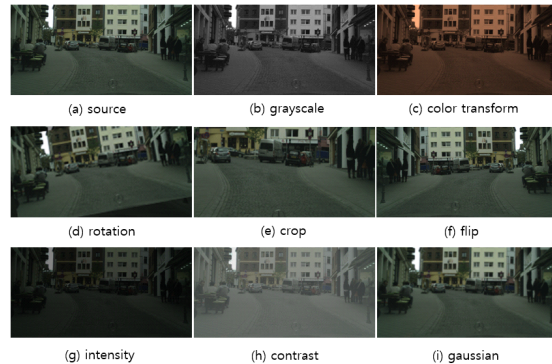


Fig. 4 Illustration of data augmentation

여 넣는 구조(그림 3의 (b))를 가지고 있다. 본 논문에서 제안된 방법에서는 [16]에서 제안된 바와 같이, 심층신경망의 층이 깊어질 때 학습의 성능이 떨어지는 문제를 극복하기 위하여 Skip connection 구조를 3, 7 번째 계층에 각각 적용하였다.

3.3. 데이터 증강 기법

일반화 성능 개선을 위해서 입력영상을 다양한 전처리 기법을 사용하여 데이터 증강(Augmentation)을 수행하였다. 사용된 전처리 기법은 자르기, 크기조절, 회전, 색상 변형, 밝기 변형, 대비 변형, 좌우 대칭 등이다. 사용된 증강기법으로 변환된 영상의 예시를 그림 4에 나타내었다.

3.4. 학습 기법 및 구현

심층신경망의 파라미터는 He initializer [18]로 초기화 하였다. 최적화 알고리즘으로 Adam optimizer [19]를 사용하였고, 초기 학습률은 0.001을 사용하였다. 모든 계층에는 Batch normalization[20]이 사용되었고, 학습단계에서 Dropout 은 최상단 5개 계층에 대해서 0.8의 확률로 사용되었다. 실험에 사용된 코드는 Tensorflow ver 2.0 로 구현되었고, Ubuntu 18.0 환경에서 실행되었으며, 학습시간은 Pascal Titan X를 이용하여 약 4시간이 소요되었다. 영상 당 추론시간은 0.5 초가 소요되었다.

IV. 실험

제안된 방법의 성능을 검증하기 위하여 Cityscape

dataset [21]을 사용하였다. Cityscape 데이터셋은 차량 전방에 설치된 카메라로 촬영된 2048×1024 해상도를 가지는 RGB 3채널로 이루어진 영상과 33개의 클래스로 분류된 Segmentation mask를 제공한다. 33개의 클래스 중 도로영역에 해당하는 데이터만을 정답 데이터로 사용하였다. 학습에 사용된 데이터 수는 2978개 이며, 검증에 사용된 데이터의 수는 503개 이다. 성능측정을 위한 지표로는 IoU (Intersection over Union)을 사용하였다. IoU 는 정답 영역과 예측 영역의 겹치는 비율로 계산되는 성능척도이며, 높은 값을 가질수록 높은 성능을 나타낸다.

제안된 방식에 의한 성능 개선을 표 1에 보였다. 비교 실험을 위하여 가장 기본이 되는 Deconvolutional neural network 와 최신의 Semantic segmentation 모델인 U-Net과 DeepLab V3를 사용하여 제안된 방법과 비교하였다. 표 1은 비교된 3개의 방식과 제안된 모델의 성능 통계인데, 제안된 방식이 $\text{IoU } 0.934 \pm 0.031$ 로 기존 방법보다 우수한 성능을 보였다.

Table. 1 Performance comparison with conventional semantic segmentation network

Model	IoU
DeconvNet	0.852 ± 0.025
U-Net	0.908 ± 0.034
Deeplab V3	0.915 ± 0.027
Ours	0.934 ± 0.031

제안된 방법에 사용된 기법인 Multi-scale semantic segmentation 방법의 효과성을 검증하기 위하여, Multi-scale 입력이 있는 방법과 없는 방법의 성능 개선을 표 2에 비교하였다. 또한, 추가적으로 Loss function 의 차이에 대한 성능 개선도 실험으로 보였다. 표 2에 실험에 사용된 여러 가지 기법들의 적용에 따른 성능 변화를 보였다. 실험에서 Multi-scale 입력을 가지며 Dice loss를 사용하였을 때 가장 높은 $\text{IoU } 0.934 \pm 0.031$ 의 성능을 보였다. 그림 5는 본 논문에서 제안한 방법으로 인식된 결과이다. 그림의 왼쪽은 정답 레이블을, 오른쪽은 인식된 결과 마스크를 나타낸다. Cityscape 데이터의 특성상 주로 포장된 도로에 대해서만 데이터가 수집되어 있으므로, 본 논문에서 제시된 결과도 포장도로에서 검증된 결과임을 명시한다.

Table. 2 Performance improvement by including proposed features

Input scale	Loss function	IoU
Single scale	Weighted CE	0.852 ± 0.025
Single scale	Focal loss	0.861 ± 0.034
Single scale	Dice loss	0.884 ± 0.016
Multi scale	Weighted CE	0.866 ± 0.032
Multi scale	Focal loss	0.908 ± 0.041
Multi scale	Dice loss	0.934 ± 0.031



Fig. 5 Inference result of proposed method (left) groundtruth (right) method output

V. 결론

본 논문에서는 심층신경망을 사용하여 자동차 전방 카메라에서 수집된 영상에서 도로영역을 효율적으로 인식하는 방법을 제안하였다. 자동차 전방 카메라 영상의 특수성을 이용하여 심층신경망의 입력이 Multi-scale 영상 형태를 가지는 구조를 제안하였고, 이를 학습하기 위한 손실함수와 네트워크 구조를 제안하였다. 제안된 방법은 Cityscape dataset을 이용하여 성능을 측정하고, 다른 방법들과의 비교를 통해 비교우위를 입증하였다. 실험결과, 3가지 기존 심층신경망 구조에 비해 우월한 성능을 냈을 것으로, 논문에서 제시한 손실함수와 Multi-scale 영상 입력에 의한 성능 변화 효과성도 실험을 통해 입증하였다. 이는 제안된 도로영역 인식을 위한 효율적인 신경망 구조가 자율주행 등 도로영상을 이용한 도로영역인식 문제 해결에 도움이 될 수 있음을 시사한다.

References

- [1] Z. Cai, Q. Fan, R. S. Feris, and N. Vasconcelos, "A unified multi-scale deep convolutional neural network for fast object detection," in *European Conference on Computer Vision*, pp. 354-370, 2016.
- [2] T. Bengio, A. Courville, and P. Vincent "Representation learning: A review and new perspectives," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1798-1828, 2013.
- [3] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pp. 315-323, 2011.
- [4] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, pp. 91-99, 2015.
- [5] J. Zhang and H.-H. Nagal, "Texture-based segmentation of road images," *Proceedings of the Intelligent Vehicles Symposium*, 1994.
- [6] S. Beucher and M. Bilodeau, "Road segmentation and obstacle detection by a fast watershed transformation," *Proceedings of the Intelligent Vehicles '94 Symposium*, 1994.
- [7] X. Yu, S. Beucher, and M. Bilodeau, "Road tracking, lane segmentation and obstacle recognition by mathematical morphology," *Proceedings of the Intelligent Vehicles '92 Symposium*, 1992.
- [8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779-788, 2016.
- [9] J. Long, F. Shelhamer, and T. Darrell "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3431-3440, 2015.
- [10] G. L. Oliveira, W. Burgard, and T. Brox, "Efficient deep models for monocular road segmentation," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016.
- [11] G. Mátyus, S. Wang, S. Fidler, and R. Urtasun, "Hd maps: Fine-grained road segmentation by parsing ground and aerial images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3611-3619, 2016.
- [12] D. Levi, N. Garnett, E. Fetaya, and I. Herzlyia, "StixelNet: A Deep Convolutional Network for Obstacle Detection and Road Segmentation," in *British Machine Vision Conference*, pp. 109.1-109.12, 2015.
- [13] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 12, pp. 2481-2495, 2017.
- [14] L. C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," *CoRR*, abs/1706.05587, 2017.
- [15] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, pp. 2961-2969, 2017.
- [16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778, 2016.
- [17] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234-241, 2015.
- [18] K. He, X. Zhang, X. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, pp. 1026-1034, 2015.
- [19] D. P. Kingma and J. Ba "Adam: A Method for Stochastic Optimization," *International Conference on Learning Representations (ICLR)*, pp. 13-24, 2015.
- [20] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," in *International Conference on Machine Learning*, pp. 448-456, 2015.
- [21] M. Cordts, M., Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3213-3223, 2016.



박세진(Sejin Park)

2007년 8월 아주대학교 컴퓨터공학 석사
2012년 9월 ~ 현재 한양대학교 컴퓨터공학과 박사과정
2015년 ~ 2016년 LG 전자 책임연구원

※관심분야 : 컴퓨터 비전, 패턴인식, 딥러닝



한정훈(Jeong Hoon Han)

2016년 2월 한림대학교 컴퓨터공학 학사
2016년 3월 ~ 현재 한양대학교 컴퓨터공학과 석박사통합과정

※관심분야 : 컴퓨터비전, 패턴인식, 딥러닝



문영식(Young Shik Moon)

1980년 2월 서울대학교 전자공학과 공학사
1982년 2월 한국과학기술원 전기 및 전자공학과 공학석사
1990년 Ph. D. in Department of Electrical and computer Engineering University of California at Irvine
1982년 ~ 1985년 한국 전자통신연구소 연구원
1989년 ~ 1990년 Inno Vision Medical 선임연구원
1990년 ~ 1992년 생산기술연구소 선임연구원
1992년 ~ 현재 한양대학교 컴퓨터공학과 교수
※관심분야 : 영상처리, 컴퓨터비전, 패턴인식