

정제 모듈을 포함한 컨볼루션 뉴럴 네트워크 모델을 이용한 라이다 영상의 분할

LiDAR Image Segmentation using Convolutional Neural Network Model with Refinement Modules

박병재¹, 서범수¹, 이세진[†]

Byungjae Park¹, Beom-Su Seo¹, Sejin Lee[†]

Abstract: This paper proposes a convolutional neural network model for distinguishing areas occupied by obstacles from a LiDAR image converted from a 3D point cloud. The channels of a LiDAR image used as input consist of the distances to 3D points, the reflectivities of 3D points, and the heights of 3D points from the ground. The proposed model uses a LiDAR image as an input and outputs a result of a segmented LiDAR image. The proposed model adopts refinement modules with skip connections to segment a LiDAR image. The refinement modules with skip connections in the proposed model make it possible to construct a complex structure with a small number of parameters than a convolutional neural network model with a linear structure. Using the proposed model, it is possible to distinguish areas in a LiDAR image occupied by obstacles such as vehicles, pedestrians, and bicyclists. The proposed model can be applied to recognize surrounding obstacles and to search for safe paths.

Keywords: LiDAR image, Segmentation, Convolutional neural network

1. 서 론

이동 로봇이 주어진 임무를 수행하기 위해서는 우선 환경에 대한 다양한 정보를 획득할 수 있어야 한다. 환경에 대한 정보는 이동 로봇의 현재 위치 정보나 이동 로봇 주변에 장애물 정보가 포함된다. 이 중에서 이동 로봇 주변의 장애물 정보는 이동 로봇이 안전하게 이동하기 위한 경로를 계획할 때 필요하다. 이동 로봇 주변의 장애물의 위치만이 아닌 이동 로봇의 장애물의 종류까지 이동 로봇이 알 수 있다면 이동 로봇이 경로를 계획할 때 효율적이고 안전한 경로를 계획하는 것이 가능하다^{1,2}. 이동 로봇이 주변의 장애물 정보를 획득하기 위해서는 스테레오 카메라(stereo camera)나 라이다(lidar)와 같은

센서를 주로 이용한다. 이 중에서 라이다는 스테레오 카메라에 비해서 넓은 탐지 범위를 가지고 있으며 조명의 영향을 거의 받지 않아 실내 및 실외 환경 모두에서 사용할 수 있다는 장점을 가지고 있다. 그러나 라이다에서 얻은 3D 포인트 클라우드(point cloud)는 스테레오 카메라에서 얻을 수 있는 RGB 영상이나 깊이 영상보다 데이터의 밀도가 낮아서 이동 로봇 주변의 장애물 정보를 획득하기 어렵다는 이슈(issue)가 존재한다.

라이다에서 얻은 3D 포인트 클라우드로부터 이동 로봇 주변의 장애물 정보를 획득하기 위한 다양한 방법이 제안되어 왔다. 제안되었던 방법들은 주로 라이다에서 얻은 3D 포인트 클라우드 중에서 바닥이나 벽과 같은 배경에 속하는 3D 포인트를 제거한 다음 나머지 3D 포인트를 분류한다. 3D 포인트를 분류하기 위해서는 RANSAC (random sample consensus)을 응용한 군집화(clustering)를 사용하거나³⁻⁵ 3D 포인트로부터 미리 정의한 피쳐(hand crafted feature)를 추출하여 SVM (support vector machine)을 이용한다⁶.

최근 들어 컨볼루션 뉴럴 네트워크(convolutional neural network)가 다양한 응용에서 기존의 군집화 알고리즘이나 미리 정의한 피쳐 기반의 분류 알고리즘보다 성능이 뛰어난 것

Received : Jan. 15, 2018; Revised : Feb. 13, 2018; Accepted : Feb. 19, 2018

※ This work was supported by Institute for Information & communications technology Promotion (IITP) grant funded by the Korea government (MSIT) (No.2017-0-00067, Development of ICT Core Technologies for Safe Unmanned Vehicles).

1. ETRI, Daejeonl, South Korea ({bjp,bsseo}@etri.re.kr)

† Corresponding author: Division of Mechanical and Automotive Engineering, Kongju National University, Cheonan, South Korea (sejiny3@kongju.ac.kr)

이 보여짐에 따라 컨볼루션 뉴럴 네트워크를 이용하여 라이다에서 얻은 3D 포인트 클라우드로부터 이동 로봇 주변의 장애물 정보를 획득하는 방법이 제안되었다^[7-11]. 이러한 방법은 주로 영상에서 물체를 포함하는 테두리 상자(bounding box) 정보 찾는 것과 유사한 방법으로 이동 로봇 주변의 장애물을 포함하는 테두리 상자 정보를 찾는다.

이동 로봇이 경로를 계획하기 위해서는 이동 로봇 장애물의 테두리 상자 정보뿐 아니라 장애물이 존재하는 영역과 그 장애물의 종류에 대한 정보가 필요하다. 이를 위하여 본 논문에서는 라이다로부터 얻은 3D 포인트 클라우드를 분할(segmentation)할 수 있는 컨볼루션 뉴럴 네트워크 모델(model)을 제안한다. 제안하는 모델은 영상 분할을 위해 사용되는 컨볼루션 뉴럴 네트워크를 응용하여^[9-11] 라이다로부터 얻은 3D 포인트 클라우드를 분할한다. 기존의 영상 분할을 위한 컨볼루션 뉴럴 네트워크는 RGB 영상을 입력 받아서 분할 결과를 출력한다. 이와 같이 본 논문에서 제안하는 모델은 먼저 라이다로부터 입력 받은 3D 포인트 클라우드를 라이다 영상으로 변환하여 이를 입력 받아 입력 받은 라이다 영상과 동일한 높이와 너비를 가지는 라이다 영상 분할 결과를 출력한다.

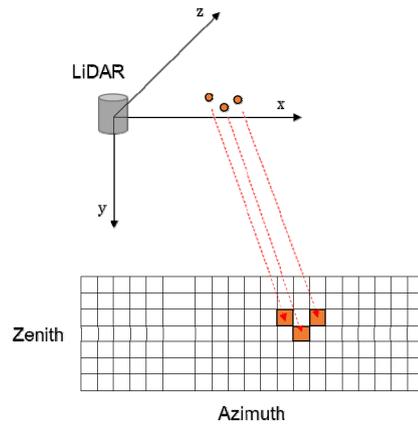
본 논문의 구성은 다음과 같다. 2장에서는 라이다에서 입력 받은 3차원 포인트 클라우드를 제안하는 모델의 입력으로 사용하기 위한 라이다 영상으로 변환하는 과정에 대해 설명한다. 3장과 4장에서는 제안하는 컨볼루션 뉴럴 네트워크 모델의 구조 및 학습 방법에 대해 설명한다. 5장에서는 제안하는 모델을 사용한 라이다 영상의 분할 실험 결과를 보여준 다음 6장에서는 결론을 맺는다.

2. 라이다 영상 변환

제안하는 컨볼루션 뉴럴 네트워크 모델을 이용하여 3차원 포인트 클라우드를 분할하기 위해서는 3차원 포인트 클라우드를 제안하는 모델에 입력하기 위한 라이다 영상 형태로 변환해야 한다.

라이다 영상은 일반적인 RGB 영상과 유사하게 H×W×C 형태를 가진 3차원 텐서(tensor)로 H와 W는 각각 수직 방향과 수평 방향 격자의 숫자를 의미하고 C는 채널(channel) 숫자를 의미한다. 라이다 영상에서 수직 방향 격자와 수평 방향 격자는 일정한 간격으로 나누어진 구면 좌표계(spherical coordinate)에서 양각(zenith)과 방위각(azimuth)을 의미한다. 라이다 영상에서 채널은 3차원 포인트까지의 거리나 3차원 포인트의 반사율 정보 등을 나타낸다.

3차원 포인트 클라우드를 라이다 영상으로 변환하기 위해



[Fig. 1] LiDAR image conversion from 3D point cloud

서는 [Fig. 1]과 같이 하나의 3차원 포인트를 다음 식을 이용하여 구면 좌표계로 변환한 다음 라이다 영상의 수직 및 수평 격자의 위치를 계산하여 해당하는 격자의 채널 정보를 갱신한다.

$$\theta = \sin^{-1} \frac{x}{\sqrt{x^2 + z^2}}, \tag{1}$$

$$\varphi = \sin^{-1} \frac{y}{\sqrt{x^2 + y^2 + z^2}}.$$

만일 여러 개의 서로 다른 3차원 포인트가 라이다 영상에서 같은 격자로 변환될 경우 3차원 포인트까지의 거리가 가장 짧은 라이다 포인트가 해당 격자의 채널 정보를 결정한다.

제안하는 모델의 경우 전방 90도 범위에 포함되는 3차원 라이다 포인트 클라우드만을 라이다 영상을 변환하는데 사용하였다. 라이다 영상의 수직 방향 격자 숫자는 Velodyne 라이다의 수직 주사선 개수와 동일한 64개로 되어 있으며 수평 방향 격자의 숫자는 512개로 되어 있다. 라이다 영상의 채널 정보는 라이다 포인트까지의 거리와 라이다 포인트의 반사율 외에 추가적으로 지면으로부터 라이다 포인트까지의 높이를 사용하였다.

3. 컨볼루션 뉴럴 네트워크 구조

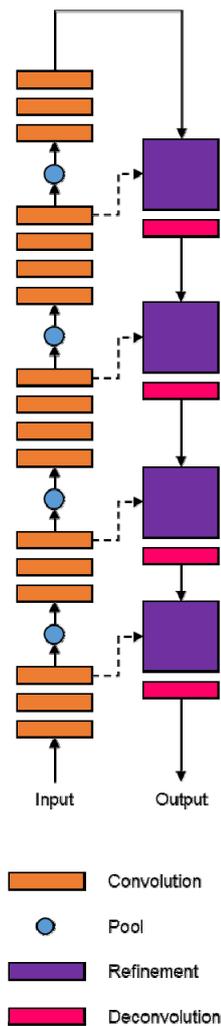
제안하는 컨볼루션 뉴럴 네트워크 모델의 구조는 [Fig. 2]에 나타나 있다. 제안하는 모델은 RGB 영상을 분할하기 위한 Facebook의 SharpMask 모델과^[12] 유사하게 건너 뛰는 연결(skip connection)을 이용하는 여러 개의 정제 모듈(refinement module)을 가지고 있다.

입력된 라이다 영상은 제안된 컨볼루션 뉴럴 네트워크에서 먼저 세 개에서 네 개의 컨볼루션 레이어(convolutional

layer)가 연속으로 적용된 다섯 개의 컨볼루션 모듈과 네 개의 풀링(pooling) 레이어를 거쳐 텐서 형태로 추출된다. 추출된 텐서는 세 개의 정제 모듈과 디컨볼루션(deconvolutional) 레이어를 통과시켜 라이다 영상 분할 결과로 출력된다.

컨볼루션 모듈 중 처음 세 개의 컨볼루션 모듈을 구성하는 컨볼루션 레이어는 3×3의 필터(filter)로 이루어져 있으며 마지막 컨볼루션 모듈을 구성하는 컨볼루션 레이어는 1×1의 필터로 이루어져 있다.

제안하는 모델에서 특징적인 부분 중 하나인 정제 모듈은 이전 단계에서의 출력뿐 아니라 컨볼루션 모듈의 출력을 함께 사용하는 특징을 가지고 있다. 정제 모듈의 내부 구조는 [Fig. 3]에 나타나 있다. 정제 모듈은 [Fig. 3]과 같이 이전 단계에서의 출력과 컨볼루션 모듈의 출력을 각각 서로 다른 컨볼루션 레이어를 통과시킨 다음 그 결과를 채널 방향으로

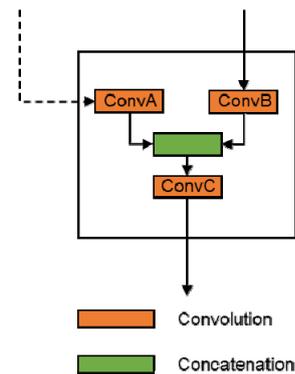


[Fig. 2] Proposed convolutional neural network model for LiDAR image segmentation

쌓은 다음 다시 컨볼루션 레이어를 통과시키는 구조를 가지고 있다. 정제 모듈의 내부에서 사용되는 세 개의 컨볼루션 레이어는 모두 3×3 필터로 이루어져 있다. [Table 1]에서는 첫 번째 정제 모듈을 구성하는 세 개의 컨볼루션 레이어의 입력 및 출력 채널 숫자를 보여준다. [Table 1]에서 나타난 바와 같이 정제 모듈에서 이전 단계에서의 출력을 통과시키는 컨볼루션 레이어와 컨볼루션 모듈의 출력을 통과시키는 컨볼루션 레이어의 채널 숫자가 같다.

제안하는 모델 내부에 있는 모든 컨볼루션 또는 디컨볼루션 레이어 뒤에는 배치 정규화(batch normalization) 레이어와 ReLU(rectifier linear unit) 활성화 함수가 결합되어 있다. 그리고 앞에서부터 네 번째 컨볼루션 모듈을 통과할 때마다 출력되는 텐서 채널의 숫자는 64, 128, 256, 512로 증가하며 뒤에서부터 네 번째부터 두 번째까지의 디컨볼루션 레이어를 통과할 때마다 텐서 채널을 숫자는 256, 128, 64로 감소한다. 풀링 레이어는 모두 2×2 필드에서 최대 값을 남기는 최대값 풀링(max pooling)을 사용하였다.

제안하는 모델에서 라이다 영상을 입력 받아 라이다 영상 분할 결과를 출력 하는데 이 때 출력하는 라이다 영상 분할 결과의 너비와 높이는 입력 받은 라이다 영상과 동일하다. 제안하는 모델의 출력은 H×W×K의 형태를 가진 3차원 텐서로 K는 분할하고자 하는 물체의 클래스 숫자를 의미한다. 제안하는 모델이 입력 받은 라이다 영상의 높이 및 너비가 동일한 라이다 영상 분할 결과를 출력하기 위해서는 제안하는 모델 내부에 풀링 레이어와 동일한 숫자의 디컨볼루션 레이어를 사



[Fig. 3] Internal structure of refinement module

[Table 1] Input and output channels of the 1st refinement module

	Input channels	Output channels
ConvA	512	128
ConvB	512	128
ConvC	256	256

용한다. 이 때 디컨볼루션 레이어는 트랜스(transposed) 컨볼루션 레이어를 이용하는 방법과 선형 보간을 이용한 업스케일(upscale) 후 컨볼루션 레이어를 이용하여 채널 숫자를 조정하는 두 가지 방법으로 구현할 수 있다. 제안하는 모델에서는 두 가지 방법을 혼용하여 첫 번째 정제 모듈에서는 첫 번째 방법을 사용하고 두 번째와 세 번째 정제 모듈에서는 두 번째 방법을 사용하였다.

4. 컨볼루션 뉴럴 네트워크 학습

4.1 데이터셋(Data set)

제안하는 컨볼루션 뉴럴 네트워크 모델을 학습하고 성능을 검증하기 위한 데이터 셋으로는 KITTI 장애물 데이터셋을 사용하였다^[3]. KITTI 장애물 데이터셋은 7,481장의 라이다 데이터 및 라벨로 이루어져 있다. KITTI 장애물 데이터셋을 컨볼루션 뉴럴 네트워크의 학습, 검증(validation) 그리고 테스트(test)를 위하여 64:16:20의 비율로 나누었다. 라이다 데이터는 앞의 2장에서 설명한 방법을 통하여 라이다 영상으로 변환하였고, 분할하고자 하는 물체의 클래스(class)는 자동차, 사람, 자전거/오토바이 세 가지로 나누었다.

4.2 데이터 증강

제안하는 모델을 학습시킬 때 과적합(overfitting)의 발생을 방지하여 강인성을 향상시키기 위해서 데이터 증강(augmentation)을 수행하였다. 데이터 증강 방식으로는 수평 방향 반전(flip)과 수평 이동(translation) 두 가지를 사용하였다.

4.3 손실 함수

제안하는 모델의 파라미터(parameter)가 라이다 영상을 분할할 수 있도록 올바르게 학습되고 있는지 확인하기 위하여 크로스 엔트로피 손실(cross entropy loss)을 손실 함수로써 사용하였다:

$$L(x, c) = -x[c] + \log \sum_j \exp(x[j]), \quad (2)$$

위 식에서 x 는 라이다 영상의 픽셀을 c 는 클래스 라벨(class label)을 의미하며 $x[c]$ 는 라이다 영상 픽셀 x 가 클래스 라벨 c 를 가지고 있을 확률이다.

제안하는 모델은 $H \times W \times K$ 형태를 가진 3차원 텐서를 출력하는데 계산 편의성을 위하여 텐서의 형태를 $(H \times W) \times K$ 형태

를 가진 2차원 텐서로 변형하여 크로스 엔트로피 손실을 계산하였다.

4.4 클래스 가중치

KITTI 장애물 데이터셋에 포함된 3차원 포인트 클라우드에 포함되어있는 3차원 포인트 중 대부분은 분할하고자 하는 클래스에 속하는 물체로부터 반사된 것이 아닌 도로나 하늘과 같은 배경으로부터 반사된 것이다. 따라서 KITTI 장애물 데이터셋으로 라이다 영상을 만들어 제안하는 모델을 학습시킬 경우 컨볼루션 뉴럴 네트워크가 분할하고자 하는 클래스를 잘 구분하기 보다는 배경을 더 잘 구분하도록 학습될 수 있다. 이를 방지하기 위하여 본 논문에서는 KITTI 장애물 데이터셋에 포함된 라이다 데이터의 클래스별 분포를 반영한 가중치를 컨볼루션 뉴럴 네트워크를 손실 함수에 반영되도록 하였다. 라이다 데이터의 클래스별 분포를 고려한 가중치를 구하기 위해서는 우선 학습을 위한 라이다 데이터를 라이다 영상으로 변환한 다음 변환된 라이다 영상에서 배경을 포함한 각각의 클래스의 출현 빈도를 구하여 다음과 같이 계산한다^[10].

$$w[c] = \frac{f_M}{f[c]}, \quad (3)$$

위 식에서 f_M 은 배경을 포함한 각각의 클래스의 출현 빈도의 중간 값을 나타내고 $f[c]$ 은 각각의 클래스의 출현 빈도를 나타낸다. KITTI 장애물 데이터셋에서 등장 빈도가 높은 배경의 경우 가중치가 낮게 계산되고 배경에 비하여 등장 빈도가 낮은 자동차, 사람, 자전거/오토바이 클래스의 경우 가중치가 높게 계산되어 컨볼루션 뉴럴 네트워크를 학습할 때 계산되는 손실 함수에 반영하면 컨볼루션 뉴럴 네트워크가 배경보다 자동차, 사람, 자전거/오토바이 클래스를 더 잘 구분할 수 있도록 학습될 수 있다. 클래스 가중치를 고려한 크로스 엔트로피 손실은 아래와 같이 나타낼 수 있다.

$$L(x, c) = w[c] \cdot \left(-x[c] + \log \sum_j \exp(x[j]) \right). \quad (4)$$

4.5 학습 환경 및 방법

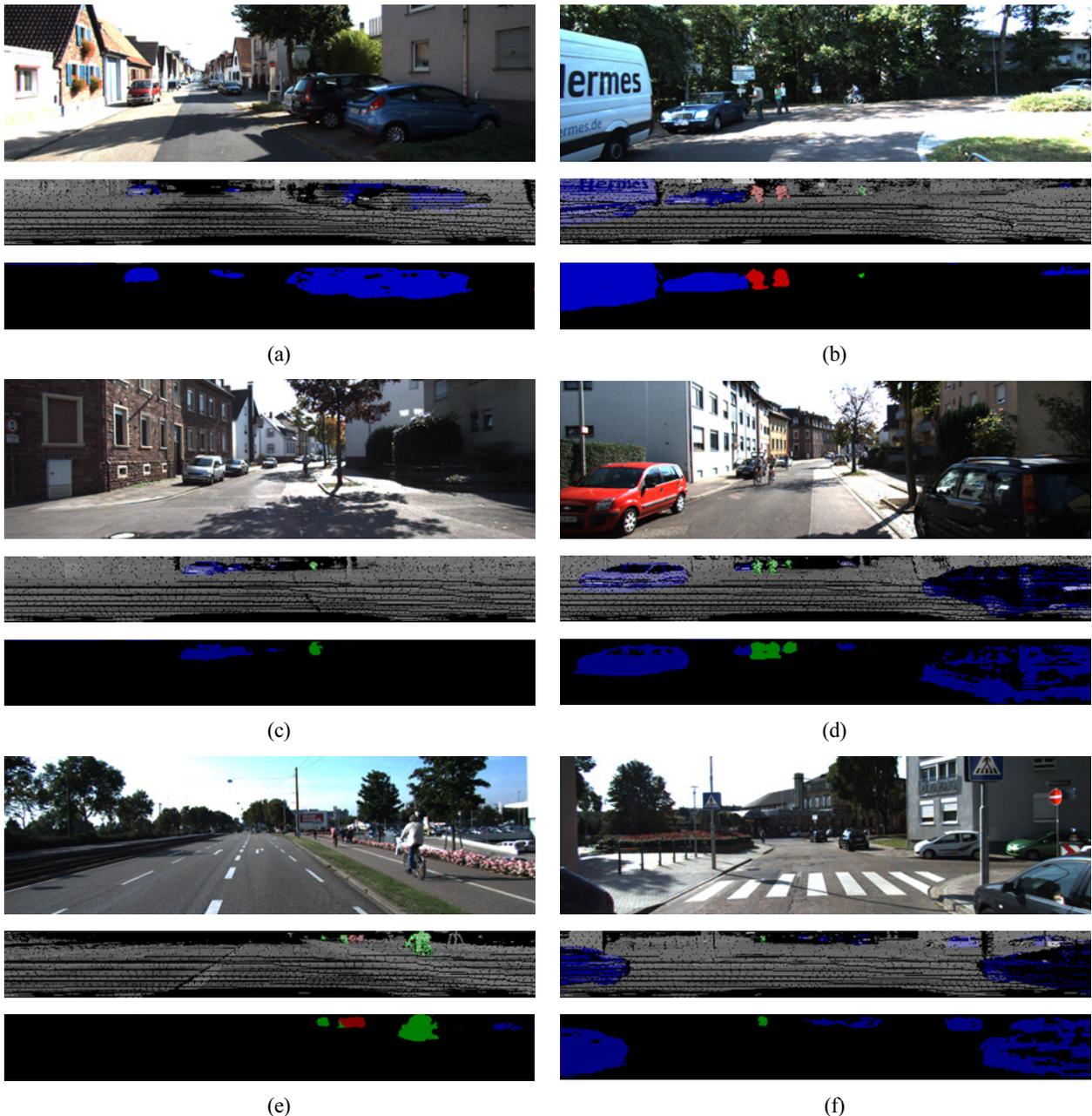
제안하는 모델을 KITTI 장애물 데이터셋으로 학습시키기 위하여 1대의 Nvidia GTX 1080Ti를 이용하였다. 제안하는 모델의 파라미터를 학습시키기 위한 최적화 알고리즘으로는 Adam(adaptive moment estimator)를 사용하였고 학습률(learning rate)은 0.001부터 시작하여 학습이 진행될 수록 지수

함수의 역수 형태로 감소하도록 하였다. KITTI 장애물 데이터셋의 배치(batch) 크기를 16으로 나누어 500 에폭(epoch) 반복하여 제안하는 컨볼루션 뉴럴 네트워크의 파라미터를 학습하였다.

5. 실험 결과

본 논문에서 제안하는 컨볼루션 뉴럴 네트워크 모델을

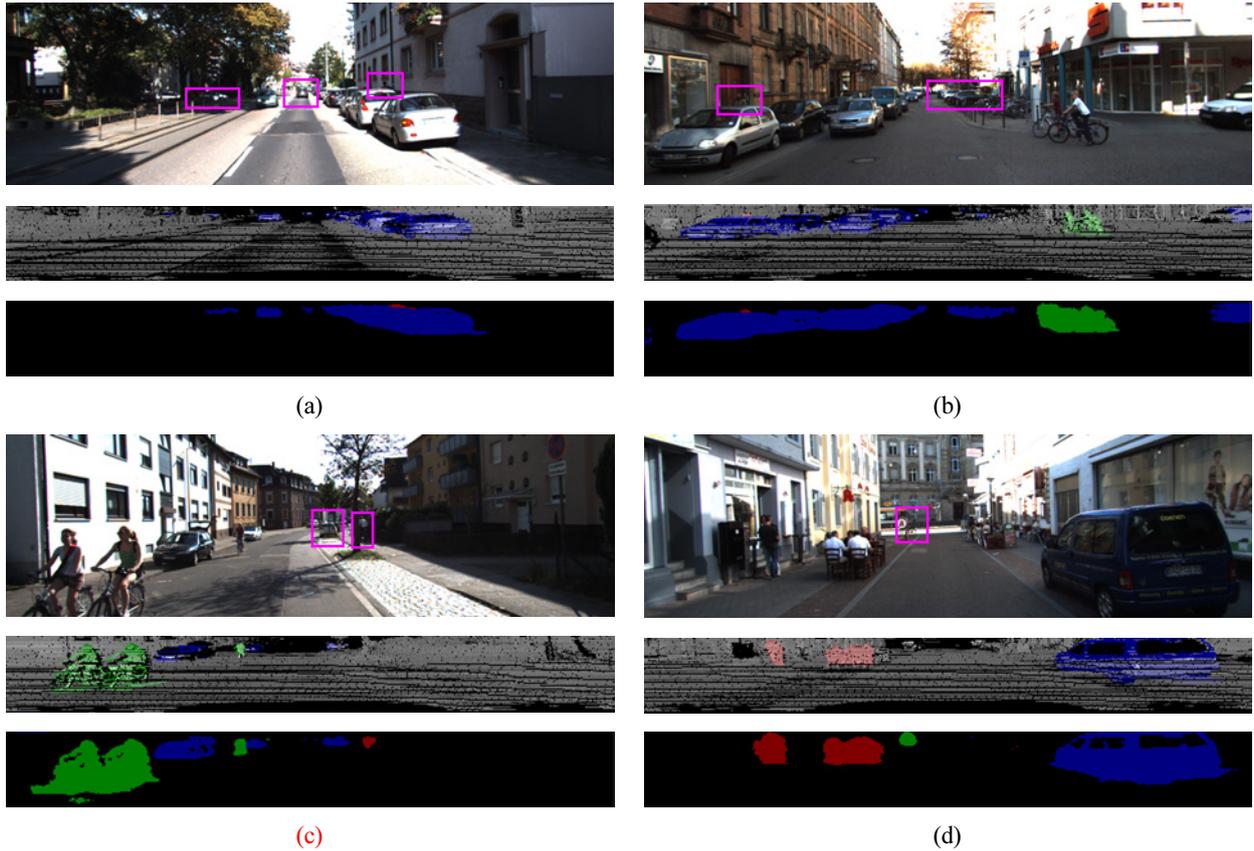
사용하여 KITTI 장애물 데이터셋을 학습시켜 라이다 영상을 분할한 결과가 [Fig. 4]에 나타나 있다. [Fig. 4]에서 각각의 그림은 위에서 아래 순서대로 RGB 영상, 라이다 영상 및 물체의 실제 클래스 라벨 그리고 라이다 영상의 분할 결과를 나타낸다. 이 때 라이다 영상은 라이다 영상의 채널 정보 중 반사율을 보여주고 있다. 클래스 라벨의 색 중에서 파란색은 자동차를, 빨간색은 사람을 그리고 초록색은 자전거/오토바이 클래스를 나타낸다. [Fig. 4]에서 나타난 바와 같이 제안하는 모델을 이용하면



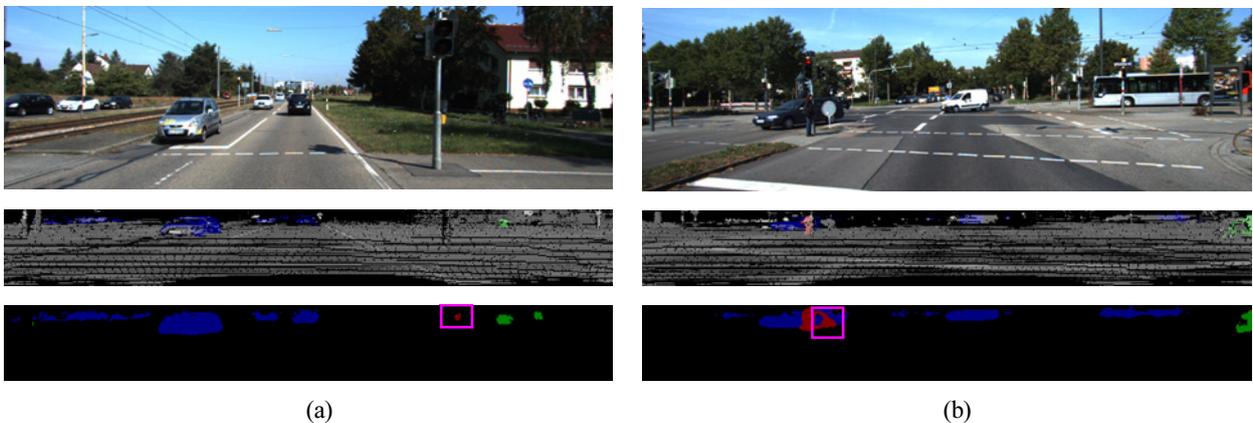
[Fig. 4] LiDAR image segmentation results using the proposed convolutional neural network model. Upper figure: RGB image for reference, middle figure: LiDAR image (intensity channel) with ground truth labels, lower figure: LiDAR image segmentation result. Blue, green and red colors represents vehicles, pedestrians and bicyclists, respectively

라이다 영상을 분할하는 것이 가능하다. 특히 [Fig. 4 (b)], [Fig. 4 (e)] 그리고 [Fig. 4 (f)]에서 나타난 바와 같이 먼 거리에 있는 크기가 작은 물체도 제안하는 모델이 라이다 영상을 분할할 수 있다는 것을 확인할 수 있다. 또한 [Fig. 4 (d)]와 [Fig. 4 (e)]에서 나타난 바와 같이 서로 다른 클래스에 속하는 물체가 라이다 영상에서 겹쳐 있는 경우에도 제안하는 모델이 이들을

구분해서 라이다 영상을 분할해 내는 것이 가능하다는 것을 확인할 수 있다. [Fig. 5]는 제안하는 모델이 라이다 영상으로부터 먼 거리에 있거나 다른 물체에 가려져서 일부만 보이는 경우에도 라이다 영상을 올바르게 분할하는 것을 보여준다. [Fig. 5 (a)]에서 RGB 영상의 왼쪽 및 중간 마젠타(magenta) 색 박스 내에 있는 차량이 먼 거리에 있지만 제안하는 모델이 차



[Fig. 5] The proposed convolutional neural network model can distinguish distant or occluded objects. Magenta boxes represent distant objects or occluded objects



[Fig. 6] LiDAR image segmentation results using the proposed convolutional neural network model contain incorrectly labeled areas. Magenta boxes represent incorrectly labeled areas

[Table 2] Label correctness of pixel

Class	Label correctness of pixel
Car	0.977
Pedestrian	0.928
Bicycle/Bike	0.906

량이 있는 것을 라이다 영상으로부터 분할하는데 성공하였다. [Fig. 5 (a)] 에서 RGB 영상의 오른쪽 마젠타 색 박스 내에 있는 사람은 앞에 있는 차량에 가려 일부만 보이지만 제안하는 모델이 이를 분할하는데 성공한 것을 보여준다. [Fig. 5 (b)], [Fig. 5 (c)] 그리고 [Fig. 5 (d)]에서 나타난 바와 같이 제안하는 모델이 먼 거리에 있는 물체나 다른 물체에 가려져 일부만 보이는 물체를 분할해 내는 것이 가능하다는 것을 확인할 수 있다. [Fig. 6]은 제안하는 모델을 이용하여 라이다 영상을 분할할 때 잘못된 부분이 포함된 결과를 보여준다. [Fig. 6]에서 영상 분류 결과의 마젠타 색 박스 내에 있는 부분이 라이다 영상의 분할이 잘못된 부분을 나타낸다.

제안하는 모델이 라이다 영상을 얼마나 잘 분할하고 있는지 정량적으로 평가하기 위하여 픽셀 레벨의 정확도와 IoU(Intersection over union) 두 가지를 사용하였다. 픽셀 레벨의 정확도는 [Table 2]에 나타나 있다.

IoU는 분류한 물체의 형태와 실제 클래스 라벨의 형태가 얼마나 일치하는지를 다음 식을 사용하여 계산한다^[9].

$$IoU = \frac{N_{TP}}{N_{TP} + N_{FP} + N_{FN}} \quad (5)$$

위 식에서 N_{TP} 는 라이다 영상의 참 양성(true positive) 픽셀(pixel)의 숫자를, N_{FP} 는 라이다 영상의 거짓 양성(false positive) 픽셀의 숫자를, 그리고 N_{FN} 는 라이다 영상의 거짓 음성(false negative) 픽셀의 숫자를 의미한다. 제안하는 컨볼루션 뉴럴 네트워크로 라이다 영상을 분할한 IoU는 0.707이다.

제안하는 모델을 온라인(online)으로 라이다 영상 분류를 위해 사용할 경우 3차원 포인트 클라우드를 라이다 영상으로 변환하는 것과 같은 전처리 작업에 걸리는 시간을 포함하여 64×512 크기의 라이다 영상을 1대의 Nvidia GTX 1080Ti 상에서 약 20 ms 이내에 처리하는 것이 가능하다.

6. 결 론

본 논문에서는 라이다 영상을 입력 받아 장애물에 해당하는 부분을 분할하기 위한 컨볼루션 뉴럴 네트워크 모델을 제안하였다. 이를 위해서는 라이다 센서로 수집한 3차원 포인

트 클라우드를 라이다 영상으로 변환한 다음 변환한 라이다 영상을 제안하는 모델에 입력한다. 제안하는 모델은 입력 받은 라이다 영상과 동일한 높이 및 너비를 가진 라이다 영상 분할 결과를 출력한다.

제안하는 모델을 사용하면 라이다 센서를 사용하여 자동차, 사람 그리고 자전거/오토바이에 해당하는 영역을 분할하여 구분하는 것이 가능하다. 제안하는 모델의 라이다 영상 분할 성능은 KITTI 물체 인식 데이터셋을 사용하여 검증하였다.

Reference

- [1] J. J. Park, C. Johnson, and B. Kuipers, "Robot Navigation with MPEPC in Dynamic and Uncertain Environments: From Theory to Practice," *IEEE/RSJ International Conference on Intelligent Robots and Systems Workshop on Progress, Challenges and Future Perspectives in Navigation and Manipulation Assistance for Robotic Wheelchairs*, Algarve, Portugal, 2012.
- [2] S. Lee and D. Kim, "Spherical Signature Description of 3D Point Cloud and Environmental Feature Learning based on Deep Belief Nets for Urban Structure Classification," *Journal of Korea Robotics Society*, Vol. 11, No. 3, pp. 115-126, September, 2016.
- [3] B. Douillard, J. Underwood, N. Kuntz, V. Vlaskine, A. Quadros, P. Morton, and A. Frenkel, "On the Segmentation of 3D LIDAR Point Clouds," *2011 IEEE International Conference on Robotics and Automation*, Shanghai, China, pp. 2798-2805, 2011.
- [4] D. Z. Wang, I. Posner, and P. Newman, "What Could Move? Finding Cars, Pedestrians and Bicyclists in 3D Laser Data," *2012 IEEE International Conference of Robotics and Automation*, Saint Paul, MN, USA, pp. 4038-4044, 2012.
- [5] D. Zermas, I. Izzat, and N. Papanikolopoulos, "Fast Segmentation of 3D Point Clouds: A Paradigm on Lidar Data for Autonomous Vehicle Applications," *2017 IEEE International Conference of Robotics and Automation*, Singapore, pp. 5067-5073, 2017.
- [6] M. Himmelsbach, A. Muller, T. L. Lüttel, and H.-J. Wünsche, "LIDAR-based 3D Object Perception," *International Workshop on Cognition for Technical Systems*, Munich, Germany, 2008.
- [7] M. Engelcke, D. Rao, D. Z. Wang, C. H. Tong, and I. Posner, "Vote3Deep: Fast Object Detection in 3D Point Clouds Using Efficient Convolutional Neural Networks," *2017 IEEE International Conference on Robotics and Automation*, Singapore, pp. 1355-1361, 2017.
- [8] B. Li, T. Zhang, and T. Xia, "Vehicle Detection from 3D Lidar Using Fully Convolutional Network," *Robotics: Science and Systems*, Michigan, USA, 2016.
- [9] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *2015 IEEE Conference on Computer Vision and Pattern Recognition*, Boston, USA, pp. 3431-3440, 2015.
- [10] P. O. Pinheiro, R. Collobert, and P. Dollár, "Learning to Segment

Object Candidates,” *Advances in Neural Information Processing Systems* 28, Montreal, Canada, pp. 1990-1998, 2015.

[11] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481-2495, Dec., 2017.

[12] P. O. Pinheiro, T. -Y. Lin, R. Collobert, and P. Dollár, “SharpMask: Learning to Refine Object Segments”, *European Conference on Computer Vision*, Amsterdam, Netherlands, pp. 75-91, 2016.

[13] *Object Detection Evaluation 2012*, [Online], http://www.cvlibs.net/datasets/kitti/eval_object.php, Accessed: Jan. 30, 2018.



박 병 재

2005 포항공과대학교 기계공학과 (학사)
2007 포항공과대학교 기계공학과 (석사)
2013 포항공과대학교 기계공학과 (박사)
2016 ~ 현재 한국전자통신연구원 선임연구원

관심분야: 로봇 인식, 기계 학습, 센서 융합, 지도 작성



서 범 수

1996 서울시립대 전산통계학과 (학사)
1998 서울시립대 전산통계학과 (석사)
2014 충남대학교 컴퓨터공학과 (박사 수료)
1998 ~ 현재 한국전자통신연구원 책임연구원

관심분야: 필드 로봇, 환경 인지, 안전 평가



이 세 진

2003 한양대학교 기계공학과 (학사)
2005 포항공과대학교 기계공학과 (석사)
2009 포항공과대학교 기계공학과 (박사)
2017 ~ 현재 공주대학교 기계자동차공학부 부교수

관심분야: 이동 로봇, 지도 작성 및 위치 추정, 기계 학습, 센서 융합