주행가능영역 판별을 위한 LiDAR 기반 의미론적 분할 연구

김나형* · 최 선**

Abstract

딥러닝 기술의 등장과 고차원 센서 데이터의 활용은 정확도와 속도가 핵심인 자율주행 분야를 크게 변화시켰습니다. 딥러닝은 심층 신경망을 활용한 컴퓨터 비전 기술을 활용하여 차선, 장애물 등을 감지하고 주행가능영역 판별을 가능케 했습니다. 이와 더불어 LiDAR 센서로 생성된 포인트 클라우드(Point cloud)는 카메라로부터 얻은 2D 이미지를 보완하는 풍부한 3D 공간 정보를 제공함으로써 인공지능에게 더 많은 정보를 담은 데이터를 제공하였습니다. 포인트 클라우드를 사용하는 방식은 다양하지만 본문에서는 주행가능영역 판별에서 주로 사용되는 포인트 클라우드를 그대로 사용하는 방식 및 2D 이미지 형태로 변환하여 사용하는 방식을 소개하고자 합니다. 또한 포인트 클라우드를 활용한 딥러닝 기반 자율주행 주행가능영역 판별에 대한 주요 방법과 예시를 정리하며, 해당 기술이 나아가야 할 방향을 고찰하고자 합니다.

I. 서 론

과거에 비해 자율주행 분야에서의 기술적 혁신은 빠르게 진행되고 있습니다. 특히 딥러닝과 고차원 센서 데이터의 융합은 자율주행 차량의 정확성과 효율성을 크게 향상시켰습니다. 자율주행 분야에서는 정확한 환경 인식과 신속한 결정이 매우 중요한데, 이러한 요구사항을 충족시키기 위한 딥러닝 알고리즘과 다양한 센서 데이터를 결합한 접근법이 사용되고 있습니다.

자율주행은 인지, 판단, 제어의 총 3가지 단계로 이루어지는데 그 중 딥러닝의 발전과 함

* 가천대학교 AI/SW학부 지능형 차량 인지 연구실, skgud99@gachon.ac.kr
** 가천대학교 AI/SW학부 지능형 차량 인지 연구실, sun4194@gachon.ac.kr

께 인공지능이 '인지'를 수행하는 데에 많은 주목을 하고 있습니다. 물론 이전에 전통적인 컴퓨터 비전 기술을 사용하여 차선인식이나 장애물 감지와 같은 작업을 수행했지만, 딥러닝은 심층 신경망을 활용하여 해당 기술을 더욱 정교하고 정확하게 수행할 수 있게 되었습니다. 딥러닝을 활용한 컴퓨터 비전 기술은 앞서 언급한 차선, 장애물 감지 등에 있어서 자동으로 feature(특징)를 학습하기 때문에 기존의 수동으로 설계하던 feature들의 한계를 넘어서 더 유연성 있는 판별을 가능하게 하였습니다.

자율주행 분야에서 컴퓨터 비전의 발전을 살펴보면, 초기에는 단순한 물체 검출 및 추적과 같은 작업에 중점을 두었습니다. 그러나 최근에는 semantic segmentation(의미론적 분할), instance segmentation(객체 분할)과 같은 복잡한 문제를 해결하는 방향으로 발전하고 있습니다. 특히, 주행 가능 영역 판단은 자율주행 차량이 안전하게 주행하기 위한 핵심 요소입니다.

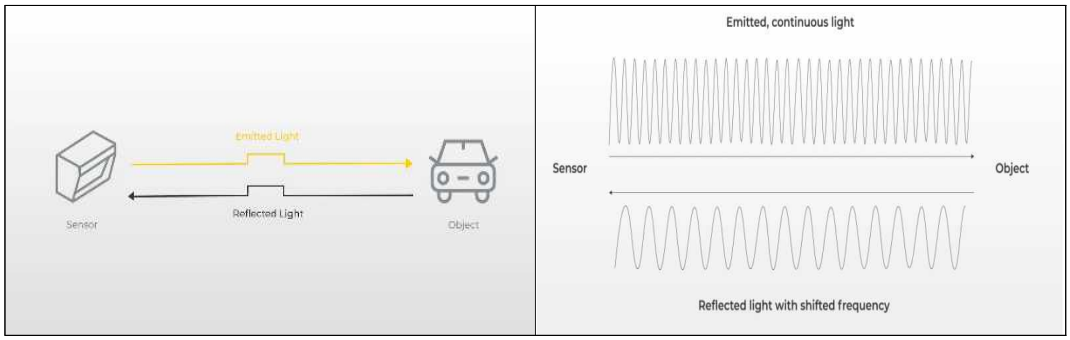
자율주행에서 가장 중요한 것은 차량을 안전하게 주행할 수 있게 하는 것입니다. 이를 위해서는 주행가능 영역이 명확하게 판단되어야 합니다. 인간의 눈과 뇌를 대신하여 물체를 정확하게 인지하고, 이를 통해 해당 영역이 안전하게 주행할 수 있는 영역인지 아닌지를 명확하게 구분합니다.이때 입력으로 들어오는 정보가 이미지가 될 수도, 포인트 클라우드가 될 수도 있으며 어떤 처리방식을 사용할지 알맞은 모델을 찾는 과정이 필요합니다.

기존 자율주행에서 카메라만을 사용할 때는 조명 변화, 날씨의 영향, 거리 정보의 부재 등으로 인한 인식의 한계가 있었습니다. 그러나 빛을 사용하는 LiDAR(Light Detection and Ranging)의 등장은 이러한 문제점들을 크게 줄이면서, 특히 3D 환경에서의 정밀한 거리 정보 제공을 통해 자율주행차량의 환경 인식 능력을 혁신적으로 향상시켰습니다.

LiDAR센서는 딥러닝의 학습 수단을 이미지가 아닌 포인트 클라우드로 대체할 수 있게 하였습니다. LiDAR는 빛을 사용하여 원격 거리 측정을 수행하는 센서로 레이저를 발사하고, 그 레이저의 반사 시간을 측정함으로써 대상까지의 거리를 계산합니다. LiDAR는 크게 두가지 원리로 나누어 볼 수 있는데 먼저 [그림 1]의 (a)와 같이 ToF(Time-of-Flight) 방식을 이용하는 경우는 라이다 센서가 발사한 광선이 대상 객체에 도달하고 다시 센서로 돌아오는 데 걸리는 시간을 측정하여 거리를 계산하는 원리에 기반합니다. 다음으로 [그림 1]의 (b)와 같이 FMCW(Frequency-Modulated Continuous-Wave) 방식을 이용하는 경우는 일정한 주파수로 조절된 레이더 빔을 사용하여 대상 객체와의 거리를 측정합니다. 광선은 주파수가 변화하는 파동으로 발사되고, 이 파동이 객체에서 반사되어 돌아오면 다시 감지됩니다. 이러한 주파수 변화의 차이를 분석함으로써 거리를 계산하며 ToF는 비교적 저렴하

고 짧은 거리에서 높은 정확성을 제공하는 반면 FMCW는 더 긴 거리에서도 좋은 정확도를 보입니다. 두 방식 모두 수집된 데이터를 이용하여 3D좌표로 변환하고, 각각의 3D좌표는 하나의 포인트로 표현되어 해당 포인트들의 모임이 포인트 클라우드를 형성합니다.

그림 1 라이다 센서 종류 [31]



(a) (b)

이러한 포인트 클라우드를 주행가능 영역 판별에 활용되는 semantic segmentation[28]에 적용하는 방식은 크게 두 가지로 나뉩니다. 첫째, 포인트 클라우드를 그대로 활용하여 환경 정보를 추출하고 분석하는 방식이고 둘째, 포인트 클라우드를 2D 이미지 형태로 투영하여 활용하는 방식입니다. 이러한 방식들은 포인트 클라우드로 인공지능에게 더 많은 정보를 제공하고 환경 정보를 보다 효율적으로 처리할 수 있게끔 하며 자율주행의 주행가능 영역을 판별하는데 자주 사용됩니다. [그림 2]는 포인트 클라우드를 그대로 활용하여 주행가능영역을 판별하는 예시입니다. 주어진 포인트 클라우드에 도로, 나무, 차량과 같은 객체들이 색상으로 구분되어 있고 해당 이미지에서는 분홍색 영역이 주행가능영역이라는 것을 확인할 수 있습니다.

그림 2 2dpass[2]를 이용한 주행가능영역 semantic segmentation 판별

Ⅱ. 주행가능영역 인식을 위한 이미지 기반 연구

자율주행 기술은 주행 환경을 이해하고 조작하는 능력이 요구되는 시스템으로, 이를 위해 다양한 센서와 컴퓨팅 파워를 필요로 합니다. 영상처리 기술은 이러한 센서 데이터를 활용하여 주변 환경을 인식하고 해석하는 핵심 기술 중 한가지입니다.

과거에는 주로 전통적인 컴퓨터 비전 기술이 자율주행의 영상처리에 활용되었습니다. 이미지를 활용하여 텍스처를 분석하거나, edge를 검출하는 등의 방식으로 주행 가능영역을 판별해왔지만 이러한 기술은 복잡한 도로 상황에서의 정확한 객체 탐지 및 추적에 한계가 있었으며, 다양한 주행 환경에 대응하기 어려웠습니다. 또한, 환경 변화에 적응하는데 한계가 있어 유연성이 부족한 문제가 있었습니다.

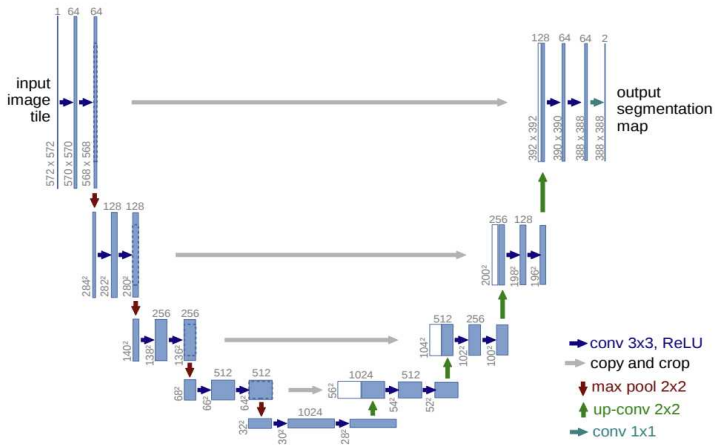
딥러닝은 기존의 영상처리 방법과는 달리 데이터 기반 학습을 통해 특징을 스스로 추출하고 판단할 수 있는 능력을 가집니다. 이로써 자율주행 영상처리 기술은 더욱 정확하고 유연한 판단을 할 수 있게 되었으며 Convolutional Neural Networks(CNN)[29] 및 Recurrent Neural Networks(RNN)[30]과 같은 딥러닝 아키텍처의 활용으로 객체 탐지, 차선 인식, 교통 표지판 인식 등 다양한 주행 요소의 처리가 가능해졌습니다.

딥러닝을 통한 자율주행 영상처리 기술의 발전은 정확도와 안정성에서 보다 훌륭한 성능 향상을 이뤘습니다. 실시간으로 도로 환경을 분석하고 판단함으로써 사고 예방 및 운전 편의성 증대에 크게 기여하고 있습니다. 또한, 딥러닝의 학습 능력을 활용하여 실제 도로 환경에서 발생하는 다양한 변수 및 노이즈 처리에 대해서도 발전을 이뤄내고 있습니다.

그 중에서도 Semantic Segmentation으로 불리는 의미론적 분할은 자율주행의 주행가능 영역을 판별하는 핵심기술이라 볼 수 있습니다. Semantic Segmentation은 이미지 속 각 픽셀에 레이블을 할당하는 컴퓨터 비전 영역의 작업을 말합니다. 보통 해당 픽셀이 어떤 객체에 속하는지, 어떤 클래스에 속하는지를 판별하기 위해 사용됩니다. 자율주행에서는 도로, 차량, 보행자, 건물, 나무 등의 영역을 구분하여 해당 클래스들로 이미지를 분할하는데 사용됩니다. 특히나 인식되는 객체의 종류가 피해야 할 장애물인지 아닌지를 구분해야 하는 자율 주행에서는 정확하고 빠른 객체 인식 및 판단이 중요하기에 의미론적 분할 작업이 중요한 역할을 합니다.

Semantic Segmentation은 CNN에서의 Fully Connected Layer를 Fully Convolutional Layer로 대체함으로써 시작되었습니다[1]. 대표적으로 인코더-디코더 구조를 가지며 픽셀 수준의 예측을 위해 디코더에서 저해상도의 feature map을 업샘플링하는 SegNet[7]이 있습니다.

그림 3 U-Net 의 구조



또한, [그림 3]과 같이 디코더가 대칭적인 U자 형태를 보이는 U-Net[8]은 의료 영상 분석에 주로 사용되었습니다. 그리고 기존 방식보다 더 넓은 영역의 특징을 포착하도록 하는 atrous convolution과 픽셀 간의 공간적인 관계와 색상 정보를 고려하여 분할 경계를 개선하는 DeepLab[9] 등이 있습니다. [그림 4]는 모양과 텍스처를 함께 처리하는 방식의 Gated-SCNN[10] 모델을 이용하여 입력 이미지에 대해 semantic segmentation 을 진행한 그림입니다. 왼쪽과 같은 원본 이미지에서 등장하는 물체들을 클래스 단위로 예측하여 오른쪽 색

칠된 이미지로 표현하였습니다.

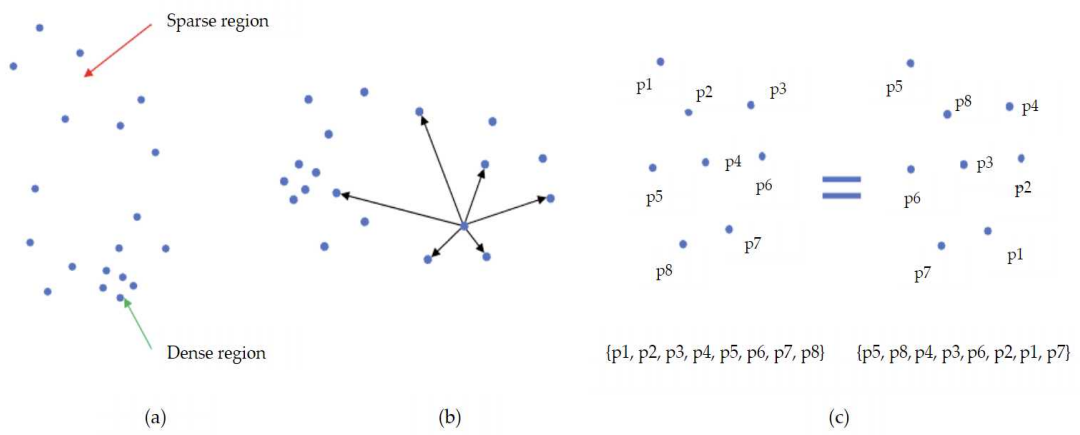
그림 4 GSCNN을 이용한 semantic segmentation

Ⅲ. LiDAR 데이터인 포인트 클라우드의 특성

[그림 3]과 같이 기존의 딥러닝에서 주행가능영역 판별은 이미지를 통한 학습이 보편적이었습니다. 그러나 포인트 클라우드의 등장으로 깊이 및 공간정보가 반영된 주행가능영역 판별이 가능하게 됐습니다. LiDAR는 Point cloud 형태의 데이터로 좌표정보와 반사강도를 나타내는 intensity와 같은 정보들을 포함합니다. 이는 지형의 고도나 형상을 정밀하게 측정하거나 도시 환경에서의 건물이나 도로 등 3D 모델링에 활용될 수 있으며, 자율 주행에서는 주변 환경을 인식하고 장애물을 탐지하는데 활용됩니다.

그러나 포인트 클라우드는 몇 가지 한계점을 가집니다[12]. 먼저 [그림 5]의 (a)는 포인트 클라우드의 분포가 균일하지 않다는 특징을 보여줍니다(Irregular/비정형성). 이는 가까이 있는 물체에서는 포인트 클라우드가 dense하게 계측되며, 멀리 있는 물체에서는 sparse 하게 계측된다는 것입니다. 다음으로 [그림 5]의 (b)는 각 포인트 클라우드가 독립적이며, 서로 인접한 포인트 클라우드 간의 거리가 고정되지 않는다는 특징을 보여줍니다(Unstructured/비구조성). 마지막으로 (c)는 포인트 클라우드의 계측된 순서가 달라져도 동일한 공간을 표현한다는 특징을 보여줍니다(Unordered /비정렬성).

그림 5 포인트 클라우드의 한계 [32]



위와 같은 제약들을 해결하기 위해 포인트 클라우드를 사용한 다양한 방식들이 고안되어 왔습니다. 그 중 크게 포인트 클라우드의 raw data를 그대로 사용하는 방식과, range image로 변환하여 사용하는 방식으로 나눌 수 있습니다. Range image란, 포인트 클라우드의 x, y, z 데이터를 u, v 좌표계의 이미지 형태로 투영한 것입니다. 포인트 클라우드를 그대로 사용하는 방식은 많은 정보를 그대로 사용할 수 있어 보다 정확한 공간 정보를 사용할 수 있다는 장점이 있습니다. 반대로 range image를 사용하는 방식은 이미지 처리 기법을 활용할 수 있어 효율적이라는 장점이 있습니다.

일반적으로 이미지는 CNN과 같은 형태의 합성곱 신경망을 주로 활용하였습니다. 하지만 포인트 클라우드의 비정형성으로 인해 행렬형태의 kernel을 직접 적용하기 어렵습니다. 이를 해결하기 위해 PointNet[13]은 각 포인트를 독립적으로 처리한 후, max pooling 연산을 사용하여 전반적인 구조를 학습하는 방식을 제안하였습니다. 이러한 방식을 통해 포인트 클라우드의 비정렬성을 보장하면서도 전체적인 형태를 인식할 수 있게 만들었습니다. [그림 5]는 PointNet을 이용해 semantic segmentation을 진행했을 때의 정성적 평가를 위한 그림입니다. [그림 6]의 첫번째 행은 포인트 클라우드에 색상 정보를 입힌 모습이고, 두번째 행은 각 포인트 클라우드에 semantic segmentation이 진행된 모습입니다. 그러나 PointNet은 지역적인 구조를 충분히 반영하지 못하는 한계가 있었고, 이를 해결하기 위해 PointNet++[14]가 등장했습니다. PointNet++는 계층적인 방식을 도입하여 포인트 클라우드의 다양한 스케일에서 특징을 추출하여 성능을 개선하였습니다.

그림 6 PointNet을 이용한 semantic segmentation

이 외에도 포인트 클라우드의 지역적인 구조를 더욱 효과적으로 학습하기 위한 다양한 방법들이 제안되었습니다. DGCNN[15]은 동적 그래프 구조를 사용하여 포인트 간의 관계를 학습하였고, PointCNN[16]은 포인트를 적절한 순서로 재배열하여 전통적인 CNN과 유사한 방식으로 처리하였습니다. 또한, 이미지 처리에서의 성공적인 기법들을 포인트 클라우드에 적용하는 시도도 있었습니다. PointSIFT[17]는 이미지에서의 특징점 추출 방식인 SIFT를 포인트 클라우드에 적용하였고, VoteNet[18]은 포인트 클라우드에서 3D 객체의 위치를 예측하기 위해 투표 메커니즘을 도입하였습니다. 이러한 방식은 포인트 클라우드의 복잡한 구조 속에서도 객체의 중심을 정확하게 찾아내는 데 효과적이었습니다.

Ⅳ. 의미론적 분할을 위한 LiDAR 데이터셋

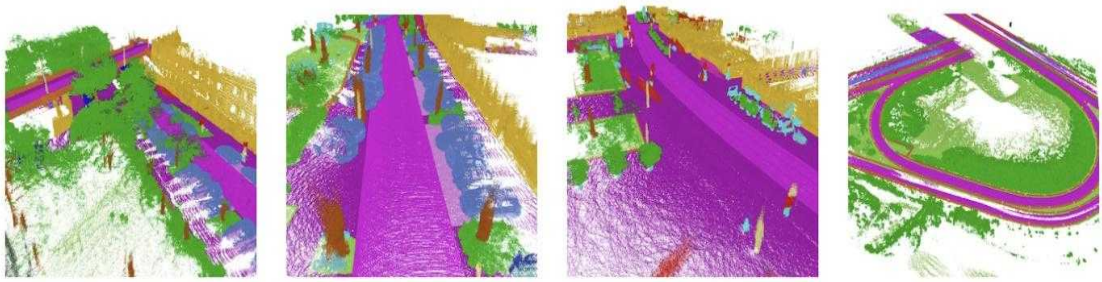
Semantic Segmentation을 수행하기 위해서는 딥러닝 모델을 학습시킬 대량의 데이터가 요구됩니다. 그 중 보편적으로 해당 분야에 활용되는 데이터셋들과 그 활용 예시들에 대해서 소개하고자 합니다. 현재 자율주행에서 주로 사용하는 포인트 클라우드 데이터셋으로는 KITTI[19], SemanticKITTI[20], Semantic3D[21], RELLIS-3D[22], Waymo Open Dataset[23] 등이 존재합니다.

KITTI는 자율주행 연구에 있어 가장 많이 활용되는 데이터셋입니다. Velodyne HDL-64

센서와 카메라를 기반으로 포인트 클라우드 뿐 아니라 projection matrices, calibration files, raw data 등의 세부 정보도 제공하기에 다양하게 활용될 수 있습니다.

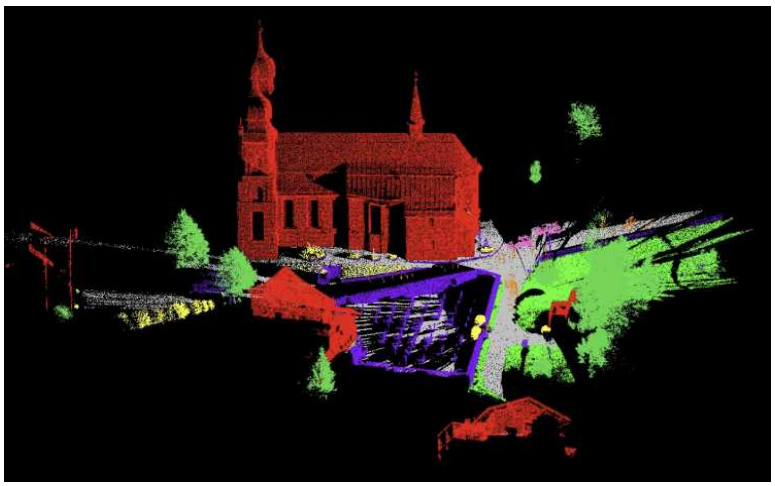
SemanticKITTI는 [그림 7]과 같이 KITTI 데이터셋에 annotation이 더해져 semantic segmentation 연구에 활용 가능한 데이터 셋입니다. 약 43,552 개의 스캔으로 구성되어 있으며, 도로, 보행자, 차량과 같은 28개의 클래스로 이루어져 있습니다. 광범위한 자율주행 및 컴퓨터 비전 연구에 사용되는 KITTI 와는 다르게, SemanticKITTI는 주로 3D semantic segmentation 및 object detection 연구에 초점을 맞추고 있습니다.

그림 7 SemanticKITTI 데이터셋 예시



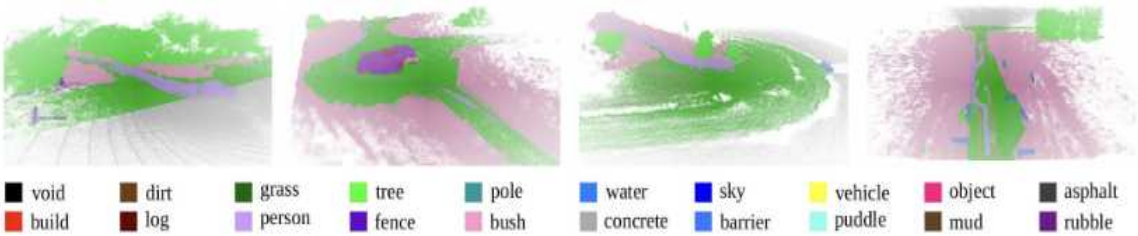
[그림 8]과 같이 Semantic3D 데이터셋은 약 4억 개 이상의 포인트로 구성된 대규모의 3D 포인트 클라우드 데이터 셋입니다. 건물, 나무 등 총 8개의 클래스를 갖으며 같이 주로 도시 환경에서의 데이터를 포함하고 있습니다.

그림 8 Semantic3D 데이터셋 예시



[그림 9]와 같은 RELLIS-3D는 그림9에서 볼 수 있듯 자율주행에서 주로 사용되는 다른 데이터셋들과는 다르게 농경지와 공원과 같은 비도심 환경에서 수집된 데이터 셋입니다. 약 2억 개의 포인트로 구성되어 있으며 6,235 개의 이미지와 13,556 개의 LiDAR 스캔을 포함합니다. 하늘, 잔디, 수풀, 진흙과 같은 20개의 클래스로 구성되어 있으며, 각 클래스별 분포가 고르지 못하다는 한계점이 있습니다.

그림 9 RELLIS-3D 데이터셋 예시



[그림 10]의 Waymo Open 데이터셋은 Waymo에서 자율주행 자동차에서 수집한 센서에서 얻은 포인트 클라우드 데이터입니다. 다른 데이터셋과는 다르게 1개의 중거리 LiDAR, 4개의 단거리 LiDAR를 활용하여 수집한 데이터셋을 제공합니다. 다양한 날씨, 번개 같은 조건이 적용되어 있다는 것이 특징입니다.

그림 10 Waymo Open 데이터셋 예시

Ⅴ. 주행가능 영역에서의 LiDAR 인지 기술

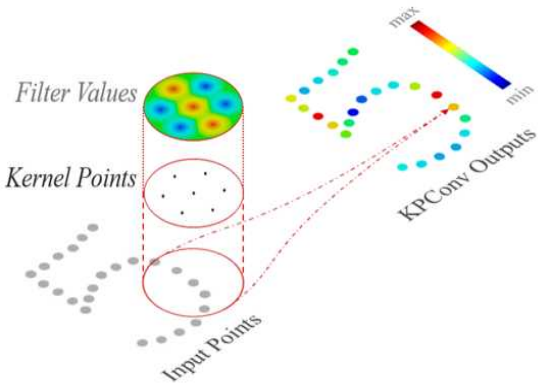
위와 같은 데이터셋을 활용하여 주행가능 영역 판별을 위한 semantic segmentation에서의 포인트 클라우드를 처리하는 방법에 기반하여 설명 드리겠습니다.

1. Point-wise method

먼저 SemanticKITTI 데이터셋을 이용한 KPConv[24]은 포인트 클라우드를 처리함에 있어서 grid나 mlp 방식을 사용하지 않고 변환 없이 그대로 convolution을 적용한 모델입니다. 기존의 이미지 기반 convolution와 비슷하지만 kernel pixels 대신에 구 형태의 kernel points 집합을 정의하여 특정 weight가 부여될 영역을 정의하는 방식을 사용합니다. Rigid KPConv와 Deformable KPConv로 나눌 수 있으며, Deformable KPConv는 각 convolution 위치에 따라 다른 shifts를 생성하게 되는데, 이를 통해 input 포인트 클라우드의 서로 다른 영역에 대해서 kernel 모양도 그에 맞게 형성될 수 있게 됩니다.

포인트 클라우드를 담는 kernel을 최적화시키기 위해 KPConv는 K-Nearest Neighbors 방식이 아닌 ERF(Effective Receptive Field)방식을 사용합니다. ERF 방식을 통해 입력으로 들어오는 포인트 클라우드의 순서에 관계없이 동일한 결과를 낼 수 있도록 하며, 지역적인 특징을 더욱 잘 잡아낼 수 있습니다. KNN에 비해 균일하지 않은 데이터에 강하기에 다양한 밀도를 가지는 원본 포인트 클라우드를 그대로 사용하면서도 효과적인 결과를 낼 수 있습니다.

그림 11 KPConv에서의 구형 kernel



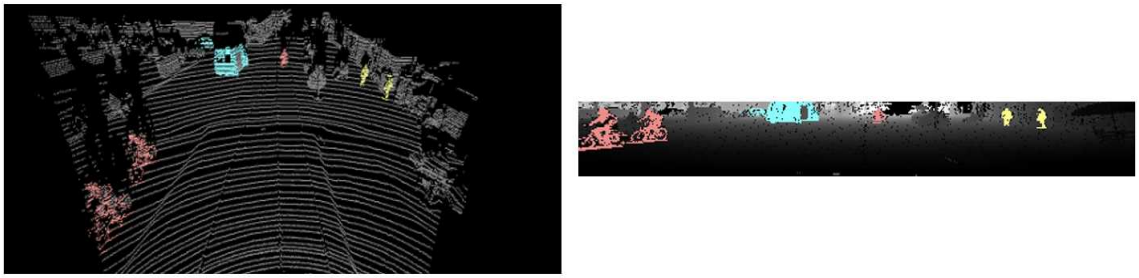
[그림 11]은 KPConv에서 사용하는 구 모양의 kernel 개념을 시각적으로 표현한 그림입니다. 특정 구 kernel을 기준으로 weight가 부여되는 것을 볼 수 있으며, 이를 통해 효과적으로 포인트 클라우드로 이루어진 데이터셋을 표현하는 것을 확인할 수 있습니다.

Rigid KPConv는 object classification이나, 작은 segmentation 데이터셋 같은 심플한 작업에서 우수한 성능을 보이며, Deformable KPConv는 더 다양하고 많은 물체에 대해 적용되어야 하는 큰 segmentation 데이터셋과 같은 보다 복잡한 작업에서 뿐 아니라 적은 수의 kernel points만이 존재할 때도 우수한 성능을 보였습니다. 이렇듯 KPConv는 포인트 클라우드를 활용할 수 있는 다양한 접근 방식에 도전하였으며 그 결과, 포인트 클라우드를 이용한 semantic segmentation에서 우수한 성능을 보였습니다.

2. Range image method

다음으로 LU-Net[25]은 2D range image의 각 픽셀 별 이웃한 픽셀을 활용하여 3D feature를 추출해내고 추출된 결과를 기반으로 U-net[26]에 적용하여 보다 효과적인 range image 의미론적 분할을 수행합니다. LiDAR 스캔 방식은 일정한 패턴을 가지기 때문에 각 점의 특정 각도와 거리를 이용하여 2D 이미지 상의 위치를 결정할 수 있습니다. [그림 12]의 (a)에 해당하는 포인트 클라우드가 해당 과정을 거쳤을 때 (b)와 같은 형태로 변할 수 있는 것입니다. 해당 방식 외에도 range image를 제작하는 방식은 더 있지만 LU-Net에서는 위와 같은 방법을 활용합니다.

그림 12 LU-Net에서의 3D 포인트 클라우드 – range image 변환

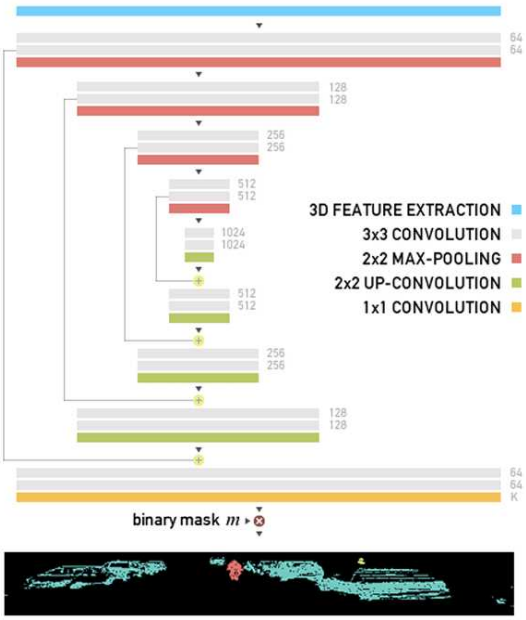


(a) (b)

그러나 이는 기존의 U-net의 input으로 사용하기 적절하지 않기에 x, y, z좌표에 대해 한 픽셀과 그 주변의 8개 픽셀들에 대해 reflectance값과 병합하여 한 번 더 MLP 레이어

를 거치게 됩니다. 이를 통해 x, y, z 좌표의 특성을 놓치지 않고 U-net의 input으로 활용할 수 있습니다. [그림 13]는 해당 방법론을 전체적으로 표현하고 그 결과를 보입니다.

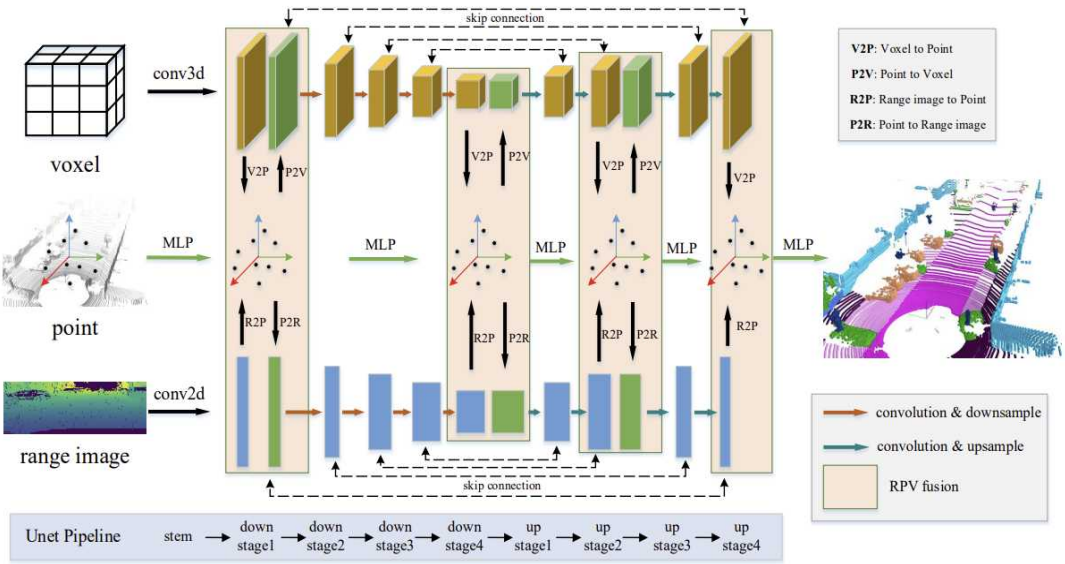
그림 13 LU-Net 전체 구조 및 결과



3. Multi-modal

각 방식들이 가지고 있는 장점들은 최대한 활용하고, 단점들은 보완하기 위해 하나의 입력만 사용하지 않고 여러 개의 입력 형태를 갖는 모델도 있습니다. RPVNet[27]은 명확한 위치 정보를 갖지만 sparse한 특징을 갖는 포인트 클라우드, 규칙적이지만 sparse하고 해상도에 따라 계산량도 크게 증가하는 voxel, 그리고 규칙적이고 밀도가 높지만 투영 과정에서 물리적인 왜곡이 일어나는 range image를 모두 입력 정보로 사용합니다. 이를 위해 세 가지 특징을 병합할 수 있는 게이트 퓨전 모듈을 제안하는 방법이 있습니다. [그림 14]와 같이 각각의 입력들은 유사한 Unet 형태의 구조를 공유하여 다중 상호작용을 갖는 3개의 브랜치 네트워크를 갖습니다.

그림 14 RPVNet 전체 구조



Point cloud는 이와 같이 특정한 절차를 거쳐 기존의 이미지를 활용한 딥러닝 또한 그대로 적용할 수 있으며, 이는 대용량의 포인트 클라우드 데이터를 보다 더 빠르게 처리할 수 있는 방법 중 한가지로 주목받고 있습니다.

물론 해당 방식들 외에도 많은 방식들이 연구되고 있으며 이러한 딥러닝 모델들의 발전으로 주행 환경에서의 보다 빠르고 정교한 semantic segmentation 을 가능하게 하고 있습니다. 이로 인해 자율주행 분야에서의 주행 가능 영역 판별에도 큰 변화가 생기고 있음을 느낄 수 있습니다.

이러한 딥러닝 모델들의 발전으로 주행 환경에서의 보다 빠르고 정교한 semantic segmentation을 가능하게 하고 있습니다. 이로 인해 자율주행 분야에서의 주행 가능 영역 판별에도 큰 변화가 생기고 있음을 느낄 수 있습니다.

Ⅵ. 결 론

자율주행 분야에서의 딥러닝 기술과 고차원 센서 데이터의 활용은 주행 가능 영역의 판별에 큰 변화를 가져왔습니다. 그 중에서도 현재 많은 주목을 받고 있는 Point cloud에 대하여 Raw Point cloud, 2D range image와 같은 형태의 데이터 활용 방식을 중심으로 다

양한 주행가능영역 판별 기술들이 연구되고 있습니다. 이러한 기술들은 차선, 장애물 등을 정확하게 감지하고 주행 가능 영역을 판별하는 데 큰 도움을 제공하며, 자율주행 차량의 안전성과 성능 향상에 기여하고 있습니다. 특히, 포인트 클라우드와 같은 3D 정보는 카메라로부터 얻은 2D 이미지만으로는 파악하기 어려운 공간 정보를 제공하여 인공지능의 인지 능력을 향상시킵니다. 따라서, 포인트 클라우드를 활용한 딥러닝 기반의 자율주행 기술은 앞으로도 지속적인 연구와 발전이 필요하며, 이를 통해 자율주행 분야에서의 더욱 정확하고 안전한 주행이 가능해질 것으로 기대됩니다.

참고문헌

[1] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2015, pp. 3431-3440.

[2] X. Yan, J. Gao, C. Zheng, C. Zheng, R. Zhang, S. Cui, and Z. Li, "2DPASS: 2D priors assisted semantic segmentation on LiDAR point clouds," in Proc. Eur. Conf. Comput. Vis. Tel Aviv, Israel: Springer, Oct. 2022, pp. 677-695.

[3] Y. Alon, A. Ferencz, and A. Shashua, "Off-road path following using region classification and geometric projection constraints," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., Jun. 2006, pp. 689-696.

[4] A. A. Assidiq, O. O. Khalifa, M. R. Islam, and S. Khan, "Real time lane detection for autonomous vehicles," in Proc. IEEE International Conference on Computer and Communication Engineering, 2008, pp. 82-88.

[5] Lecun, Yann & Bottou, Leon & Bengio, Y. & Haffner, Patrick. (1998). "Gradient-Based Learning Applied to Document Recognition," in Proc. IEEE. 86. 2278 - 2324.

[6] David E. Rumelhart; James L. McClelland, "Learning Internal Representations by Error Propagation," in Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations, MIT Press, 1987, pp.318-362.

[7] Badrinarayanan, Vijay, Alex Kendall, and Roberto Cipolla. "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," in Proc. IEEE transactions on pattern analysis and machine intelligence 39.12, 2017.s

[8] Olaf Ronneberger, Philipp Fischer, Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation", In International Conference on Medical image computing and computer-assisted intervention, pages 234-241. Springer, 2015.

[9] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs", in Proc. IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 4, pp. 834-848, Apr. 2018.

[10] T. Takikawa, D. Acuna, V. Jampani, S. Fidler, "Gated-SCNN: Gated shape CNNs for semantic segmentation", in Proc. IEEE Int. Conf. Comput. Vis., pp. 5229-5238, 2019.

[11] Wang, R. "3D Building Modeling using Images and LiDAR: A Review", Int. J. Image Data Fusion 2013, 4, 273-292

[12] Kharroubi, A.; Poux, F.; Ballouch, Z.; Hajji, R.; Billen, R. "Three Dimensional Change Detection Using Point Clouds: A Review", Geomatics 2022, 2, 457-485.

[13] C. R. Qi, H. Su, K. Mo, L. J. Guibas, "PointNet: Deep learning on point sets for 3D classification and segmentation", in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 77-85, 2017.

[14] C. R. Qi, L. Yi, H. Su, L. J. Guibas, "PointNet++: Deep hierarchical feature learning on point sets in a metric space", in Proc. 31st Int. Conf. Neural Inf. Process. Syst., pp. 5105-5114, 2017.

[15] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, J. M. Solomon, "Dynamic graph CNN for learning on point clouds", ACM Trans. Graph., vol. 38, 2019.

[16] Y. Li, R. Bu, M. Sun, W. Wu, X. Di, B. Chen, "PointCNN: Convolution on X-transformed points", in Proc. 32nd Int. Conf. Neural Inf. Process. Syst., pp. 820-830, 2018.

[17] M. Jiang, Y. Wu, C. Lu, "PointSIFT: A sift-like network module for 3D point cloud semantic segmentation", 2018.

[18] C. R. Qi, O. Litany, K. He and L. J. Guibas, "Deep hough voting for 3D object detection in point clouds", in Proc. IEEE/CVF Int. Conf. Comput. Vis., pp. 9276-9285, 2019.

[19] Geiger, A., Lenz, P., Stiller, C. and Urtasun, R., "Vision meets robotics: The kitti dataset", The International Journal of Robotics Research 32.11, 2013 : 1231-1237.

[20] Behley, J., Garbade, M., Milioto, A., Quenzel, J., Behnke, S., Stachniss, C. and Gall, J., "Semantickitti: A dataset for semantic scene understanding of lidar sequences", in Proc. IEEE/CVF international conference on computer vision, 2019.

[21] Hackel T, Savinov N, Ladicky L, Wegner JD, Schindler K, Pollefeys M., "Semantic3d.net: A new large-scale point cloud classification benchmark", arXiv, 2017.

[22] Jiang, P., Osteen, P., Wigness, M. and Saripalli, S., "Rellis-3d dataset: Data, benchmarks and analysis." 2021 IEEE international conference on robotics and automation (ICRA). IEEE, 2021.

[23] Sun, P., Kretzschmar, H., Dotiwalla, X., Chouard, A., Patnaik, V., Tsui, P., Guo, J., Zhou, Y., Chai, Y., Caine, B. and Vasudevan, V., "Scalability in perception for autonomous driving: Waymo open dataset." In Proc. IEEE/CVF conference on computer vision and pattern recognition, 2020.

[24] Thomas, H., Qi, C.R., Deschaud, J.E., Marcotegui, B., Goulette, F. and Guibas, L.J., "Kpconv: Flexible and deformable convolution for point clouds." In Proc. IEEE/CVF international conference on computer vision, 2019.

[25] Biasutti, P., Lepetit, V., Aujol, J.F., Bré dif, M. and Bugeau, A., "Lu-net: An efficient network for 3d lidar point cloud semantic segmentation based on end-to-end-learned 3d features and u-net", in Proc. IEEE/CVF International Conference on Computer Vision Workshops, 2019.

[26] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, in Proc. Part III 18. Springer International Publishing, 2015.

[27] Xu, J., Zhang, R., Dou, J., Zhu, Y., Sun, J., Pu, S. "Rpvnet: a deep and efficient range-point-voxel fusion network for lidar point cloud segmentation", arXiv

[28] Wang P, Chen P, Yuan Y, Liu D, Huang Z, Hou X, Cottrell G., "Understanding Convolution for Semantic Segmentation", in Proc. IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, 2018.

[29] Girshick, Ross. "Fast r-cnn", in Proc. IEEE international conference on computer vision, 2015.

[30] Sherstinsky, Alex. "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network." Physica D: Nonlinear Phenomena 404, 2020.

[31] https://www.blickfeld.com/blog/time-of-flight-vs-fmcw/

[32] https://www.researchgate.net/figure/Challenges-related-to-3D-point-cloud-processing-a-Irregular-with-dense-and-sparse_fig1_364466128