

3 차원 라이다 스캐너 기반 저밀도 포인트 클라우드의 의미론적 분할 방법

김예림, 김민석, 이유철*
동국대학교 정보통신공학과

{yr2819, batandy, yulee}@dongguk.edu

Sparse Point Cloud Semantic Segmentation Based on 3D LiDAR Scanner

Ye Rim Kim, Min Seok Kim, Yu-Cheol Lee*

Department of Information and Communication Engineering, Dongguk University

요 약

본 논문은 3 차원 라이다 스캐너로부터 얻은 저밀도 포인트 클라우드를 이용한 공간의 의미론적 분할 방법에 관한 것이다. 가격적인 측면에서 고밀도 포인트 클라우드에서 동작되는 의미론적 분할 기법들에 대한 저해상도의 라이다 스캐너의 활용 가능성이 점차 증대되고 있다. 따라서 기존의 고밀도 포인트 클라우드 기반의 분할 방법을 실제 저밀도 포인트 클라우드 환경에서 동작할 수 있도록 데이터 세트를 구성하고, 실제 성능 측면에서의 검증이 필요하다. 실제 실험을 통해 실시간성이 뛰어난 RangeNet++ 을 활용하여 고밀도 환경을 대상으로 제작된 의미론적 분할 방법이 저밀도 포인트 클라우드에서도 동작 가능함을 입증하였다.

I. 서 론

포인트 클라우드의 의미론적 분할은 다양한 분야에서 환경 이해를 위한 핵심 요소이며, 딥러닝 기술의 발전으로 성능이 크게 향상되었다. 이러한 기술의 발전을 바탕으로 RangeNet++ [1]와 같이 실시간성과 정확성이 높은 의미론적 분할 방법이 개발되고 있다. 다만, RangeNet++ 를 비롯한 대다수의 기존 연구들은 주로 64 채널과 같은 고밀도 라이다 스캐너를 기반으로 실험 및 평가가 이루어진다. 그러나 실제 응용 분야에서는 센서의 비용, 전력 소모, 데이터 처리 부담 등의 이유로 16 채널과 같은 저채널 라이다가 더 널리 사용되고 있다. 이는 상대적으로 낮은 밀도의 포인트 클라우드를 생성하므로, 기존 모델의 성능이 저하될 수 있다는 우려가 있다.

본 논문에서는 RangeNet++ 의 저밀도 포인트 클라우드에 대한 의미론적 분할 과정을 서술한다. 또한, 실험을 통해 저밀도 포인트 클라우드에서 RangeNet++ 의 성능을 검증하고, 다양한 밀도의 포인트 클라우드에 대한 RangeNet++ 의 적용 가능성을 논의하며 본 연구의 기대효과와 향후 전망을 제시하고자 한다.

II. 저밀도 포인트 클라우드 의미론적 분할

RangeNet++ 는 회전하는 3D 라이다 스캐너에서 얻은 포인트 클라우드의 의미론적 분할을 목표로 제안된 딥러닝 모델이다. RangeNet++ 모델의 전체적인 과정은 그림 1 과 같다. 이 모델은 포인트 클라우드 데이터를 범위 이미지로 변환한 후, 의미론적 분할 네트워크에 입력한다. 이는 3D 공간 정보를 효과적으로 인코딩하면서도 2D CNN 을 적용할 수 있게 하여 계산 효율성을 크게 향상시키는 효과가 있다. 의미론적 분할 네트워크는 범위 이미지를 처리하여 의미론적으로 분할된 범위 이미지를 출력한다. 그러나 의미론적 범위 이미지를 다시 포인트 클라우드로 투영하는 과정에서 그림자 현상과 정보 손실 문제가 발생할 수 있다. RangeNet++ 는 이러한 문제를 해결하기 위해 KNN 기반의 효율적인 후처리 기법을 제안하였다.

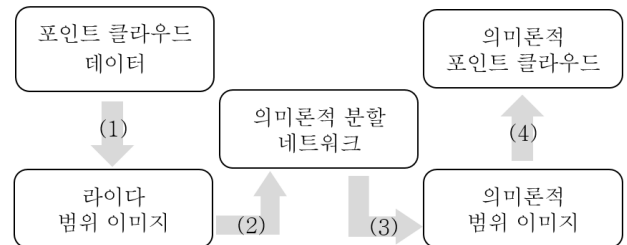


그림 1 RangeNet++ 개념도

RangeNet++ 는 SemanticKITTI[2] 데이터셋의 평가에서 의미론적 분할 성능에서 우수한 성능을 달성하였으며 임베디드 GPU 환경에서도 실시간으로 동작 가능한 속도를 보여 자율주행차량을 비롯한 실제 응용 분야에서의 활용 가능성을 입증하였다. 그러나 RangeNet++ 의 우수한 성능은 64 채널과 같은 고밀도 라이다 스캐너를 사용하여 얻은 포인트 클라우드를 기반으로 평가되었기에, 상대적으로 낮은 밀도의 포인트 클라우드를 생성하는 저채널 라이다 스캐너에 대한 성능 검증이 필요하다. 따라서, 본 연구에서는 16 채널 라이다 스캐너로 수집한 저밀도 포인트 클라우드에 대해 RangeNet++ 를 적용하고 그 성능을 평가하고자 한다.

본 연구에서는 실내 환경에서 수집된 3D 포인트 클라우드 데이터를 사용하여 커스텀 데이터 셋을 구축하였다. 데이터 수집에는 Velodyne VLP-16 라이다 스캐너를 사용하였으며, 다양한 실내 공간(연구실, 복도, 강의실, 계단 등)에서 총 2,600 개의 스캔을 취득하였다. 각 스캔은 라이다가 360 도 회전하며 얻어진 포인트 클라우드 데이터로, 약 20,000~30,000 개의 3D 점으로 구성된다. 수집된 포인트 클라우드 데이터는 SemanticKITTI 와 동일한 포맷(.bin)으로 저장되었으며, 각 점은(x, y, z)좌표와 반사도 값으로 구성된다. 수집된 포인트 클라우드 데이터에 대한 의미론적 레이블링은 수작업으로 진행되었다. 총 5 개의 클래스를 정의하였으며, 이는 벽(wall), 천장(ceiling), 바닥(floor), 계단(stair), 사람(person)으로 이루어진다.

표 1 하드웨어 및 소프트웨어 사양

Development environment		Specification
HW	CPU	Intel®Xeon®Bronze3204
	RAM	64GB
	GPU	NVIDIA Quadro RTX 8000
	OS	UBUNTU 22.04
SW	GPU development tool	CUDA 11.3
	Deep learning framework	PyTorch 10.1
	Programming language	Python 3.7

표 2 검증 데이터 세트에 대한 RangeNet53 의 IoU [%]

입력 크기	천장	바닥	계단	벽	사람	mIoU	정확도
16x1024px	96.1	83.4	79.1	97.2	67.0	84.5	92.3

효과적인 학습을 진행하기 위하여 전체 2,600 개의 스캔 중 70% (1,820 개)는 학습 세트, 10% (260 개)는 검증 세트, 20% (520 개)는 테스트 세트로 할당하였다. 각 세트는 다양한 실내 환경과 레이플 분포를 고르게 포함하도록 구성하였다. 또한, 본 연구에서는 16 채널 라이다 스캐너 기반 데이터를 사용하므로 모델의 입력 값을 VLP-16 라이다 스캐너의 수직 채널 값인 16 으로 설정하였다.

III. 학습 및 실험

본 연구에서는 RangeNet++ 의 저밀도 포인트 클라우드에 대한 의미론적 분할 성능 평가를 위해 16 채널 라이다로 수집한 데이터셋을 활용하여 학습 및 평가를 수행하였다. RangeNet++ 학습 및 평가에 사용된 하드웨어와 소프트웨어는 표 1 와 같다.

RangeNet++ 모델은 학습 과정에서 weighted cross-entropy 손실함수를 사용한다. 이 손실함수는 클래스 불균형 문제를 다루기 위해 각 클래스의 빈도에 기반한 가중치를 적용하며 다음과 같다.

$$L = -\sum_{i=1}^C w_i y_i \log(\hat{y}_i), \text{ where } w_i = \frac{1}{\log(f_i + \epsilon)}, \quad (1)$$

여기서 C 는 전체 클래스의 개수이다. 그리고 w_i 는 클래스 i 의 역빈도로 계산된 가중치이고, y_i 는 클래스 i 의 실제 레이블, \hat{y}_i 는 예측된 확률값을 의미한다. f_i 는 i 번째 클래스에 해당하는 포인트의 개수이다. ϵ 은 분모가 0 이 되는 것을 막기 위해 사용되는 조정 상수이다. 백본 네트워크는 Darknet53[3] 아키텍처를 사용하고, 학습률은 0.001, 감쇠율은 0.99, 학습의 배치 사이즈는 8 로 설정하여 100 epoch 동안 수행하였다.

모델의 성능은 모든 클래스에 일반적으로 적용되는 mIoU(mean intersection-over-union)지표를 사용하여 다음 수식과 같이 평가하였다.

$$mIoU = \frac{1}{C} \sum_{i=1}^C \frac{TP_i}{TP_i + FP_i + FN_i}, \quad (2)$$

여기서 TP_i , FP_i 및 FN_i 는 클래스 i 에 대한 참 양성(true positive), 거짓 양성(false positive) 및 거짓 음성(false negative) 예측의 수에 해당한다.

표 2 는 전체 클래스에 대한 IoU, mIoU 그리고 정확도를 보여준다. mIoU 84.5%, 정확도 92.3%의 성능을 보였다. 벽, 바닥 등 주요 클래스에서는 90% 이상의 높은 IoU 를 달성했지만, 작은 물체인 사람 클래스에서는 67.0%로 낮은 IoU 를 보였다.

그림 2 는 테스트 데이터 셋의 추론 결과로 얻어진 각 클래스에 대한 혼동 행렬(Confusion Matrix)이다. 혼동 행렬로부터 계산된 평균 정밀도, 재현률 그리고 F1 sco-

True label	Predicted label					
	unlabeled	ceiling	floor	stair	wall	person
unlabeled	0.00	0.00	0.00	0.00	0.00	0.00
ceiling	0.02	0.96	0.00	0.00	0.02	0.00
floor	0.00	0.00	0.83	0.00	0.16	0.00
stair	0.00	0.00	0.00	0.79	0.21	0.00
wall	0.02	0.00	0.00	0.01	0.97	0.00
person	0.00	0.00	0.01	0.02	0.31	0.67

그림 2 테스트 데이터셋에 대한 혼동행렬

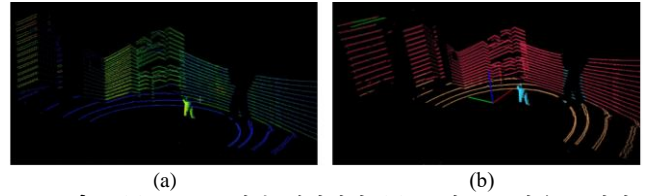


그림 3 (a) LiDAR 원본 데이터와 (b) 포인트 클라우드 의미론적 분할 후

re 는 각각 76.1%, 71.0%, 73.3%이다.

그림 3(b)는 RangeNet++ 를 이용하여 테스트 셋에 포함된 포인트 클라우드인 그림 3(a)를 분할한 결과이다. 그림 3(b)를 통해 벽, 천장, 바닥 등으로 클래스에 맞게 분할된 것을 알 수 있다.

IV. 결론 및 차후 계획

본 연구에서는 16 채널 라이다 데이터를 활용하여 RangeNet++ 로 의미론적 분할을 수행하였으며, 실험을 통해 기존의 고밀도 포인트 클라우드 기반 분할 방법이 저밀도 포인트 클라우드에 대해서도 우수한 의미론적 분할 성능을 발휘할 수 있음을 검증하였다.

향후 연구에서는 데이터 증강 기법과 손실 함수 설계를 통해 작은 물체 분할 성능을 개선하고, 경량화된 네트워크와 효율적 후처리 기법을 적용하여 다양한 밀도의 포인트 클라우드에 대한 강건하고 효율적인 의미론적 분할을 목표로 한다.

ACKNOWLEDGMENT

본 연구는 산업통상자원부 재원으로 한국산업기술기획평가원 (KEIT)의 지원을 받아 수행한 연구 과제임 (No. 20024008, AI 및 로봇을 활용한 조선/항공/건설 도장 공정 자동화 플랫폼 기술 개발)

참고 문헌

- [1] A. Milioto et al., "RangeNet++: Fast and Accurate LiDAR Semantic Segmentation," Proc. of IEEE/RSJ IROS, pp.4586-4593, 2019.
- [2] J. Behley et al., "SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences," Proc. of ICCV, pp.9297-9307, 2019.
- [3] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv:1804.02767, 2018.