

# 포인트넷 기반 포인트 클라우드 객체 분할 알고리즘 개발

Dilkashbek Zukhrudinov, Jiongyi Meng, 최수일\*

전남대학교, 전남대학교, \*전남대학교

xiaomeng199373@gmail.com, zh926720@gmail.com, sichoi@jnu.ac.kr

## Development of point cloud object segmentation algorithm based on PointNet

Dilkashbek Zukhrudinov, Jiongyi Meng, Su-il Choi\*

Chonnam Univ., Chonnam Univ., \*Chonnam Univ.

### 요약

3D 포인트 클라우드에서의 객체를 정확하게 감지하는 것은 자율 주행, 드론, 항공 및 우주 정보화와 같은 많은 응용 분야에서 핵심적인 문제이다. 본 논문에서는 포인트넷 기반의 포인트 클라우드 객체 분할 알고리즘을 제안한다. 제안된 알고리즘은 포인트넷 알고리즘과 포인트시프트(PointSIFT) 모듈을 결합하였으며 S3DIS인 포인트 클라우드 데이터 셋을 사용하여, 최종적으로 물체 분할 결과를 생성한다. 그리고 원래의 포인트넷 알고리즘의 성능과 비교하였다.

### I. 서론

모바일 로봇, 자율 주행, 드론, 항공 및 우주 정보화, 3D공간 대한 연구가 지속적으로 심화되면서 이러한 3D 정보 기술에 대한 수요가 증가함에 따라 점점 더 많은 3D 정보에 대한 연구가 이루어지고 있다. 특히 3D 데이터의 연구에서 3D 목표, 3D 공간에 대한 관심이 가장 중요하다. 3D 정보 분야의 연구는 주로 모델 재구성, 3D 대상 측정 및 인식, 공간 재구성 및 3D 복원에 중점을 둔다. 그 중에 3D 공간의 재구성 기술은 컴퓨터에서 현실적이고 객관적인 장면을 디지털 처리한 다음 컴퓨터에서 실제 재생하여 관련 환경 정보를 보다 포괄적이고 상세하게 이해하는 것이다 [1]. 실제 로봇이 자신의 환경 정보를 보다 명확하게 이해하고 환경 정보와의 정보 교환을 달성하기 위해 3 차원 환경 정보가 필요하다. 이러한 이유로, 3D 공간 정보를 얻는 것에 관한 연구가 본 논문의 연구 과제의 전제이다. 3D 환경 정보 재구성의 문제를 해결하기 위해, 3D 공간 재구성 기술은 항상 컴퓨터 비전에서 많은 관심을 끌었다. 최근에, Qi 등[2]은 포인트 클라우드로부터 직접 포인트 투 피쳐 (point-wise feature)를 학습하는 엔드 투 엔드 (end-to-end) 심층 신경 네트워크 인 포인트넷 (PointNet)을 제안했다. 이 접근법은 3차원 객체 인식, 3D 객체 분류 및 세분화 작업에서 인상적인 결과를 나타냈다. 라이다를 사용하여 얻은 일반적인 포인트 클라우드는 10만개 이상의 포인트를 포함하기 때문에 [2]과 같이 아키텍처기반으로 학하면 높은 계산 및 메모리 요구 사항이 발생한다. 이 논문에서는 3차원 특징 학습 네트워크를 더 많은 점과 3D 탐지 작업으로 확장하는 것이다.

본 논문의 구성은 다음과 같다. II장에서는 기본 포인트넷 아키텍처를 소개한다. 물체 분할하기 위해 새로운 아키텍처를 제시한다. III장에서는 시뮬레이션을 통해 제안한 방법의 효율성을 보인다. 다른 3D 검출 방법과 비교해서 장점과 발전성을 제시한다. IV장은 결론을 맺는다.

### II. 포인트넷 기반 객체 분할 알고리즘

포인트 클라우드에서 3D 포인트의 원래 치수는 다층 퍼셉트론 (MLP: Multilayer Perceptron)에 의해 매우 높은 차원으로 확장된다. 포인트 클라우드의 모든 지역 특징(local feature)를 고려하여 하나의 차원 값 세트를

하나의 지역 특징차원에 있는 포인트넷 지역 특징값 집합으로 만들고, 포인트넷은 대칭 함수를 사용하여 각 집합에 대한 대표 값을 선택하여 포인트넷 집합 전역 특징(global feature)를 출력한다. 이 기술은 포인트 클라우드 표현의 불규칙한 분포 문제를 해결했다. 포인트 클라우드의 전역 특징을 추출한 후에는 이러한 전역 특징을 객체 분류 및 세분화와 같은 다른 목표에 대한 다른 MLP으로 정렬한다.

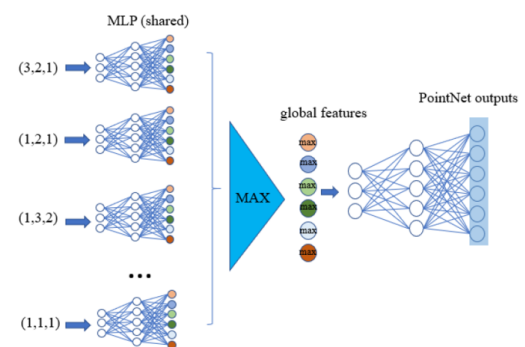


그림 1. 포인트넷 기본 아키텍처

일반적으로 고화질 라이다 포인트 클라우드는 약 100k 개 포인트로 구성된다. 모든 포인트를 직접 처리하면 컴퓨팅 플랫폼의 메모리 및 효율성 부담이 증가뿐만 아니라 공간 전체 매우 높은 포인트 밀도가 검출 오차를 유발할 수 있다. 이를 회피하기 위해 포인트 클라우드에서 고정된 개수의 포인트를 임의로 샘플링 한다. 이 샘플링 전략에는 두 가지 목적이 있다. (1) 계산 양 절감, (2) 샘플링 바이어스 감소 및 딥 러닝 변화 증가.

RPN (Regional Proposal Network)[3] 은 고도로 최적화 된 효과적인 물체 감지 알고리즘이다. 그러나 이 접근법은 일반적인 라이다 포인트 클라우드의 경우가 아닌 텐서 구조 (예 : 이미지, 비디오)로 데이터를 집중적으로 구성해야 한다. 본 논문에서는 라이다에 의해 포착된 포인트 집합의 거리 정보를 이용하여 탐지 물체의 위치를 대략적으로 결정하고 탐지 구역 제안을 제안하고자한다.

제안된 알고리즘의 아키텍처에 대한 입력은 4096 포인트의 3D 좌표이

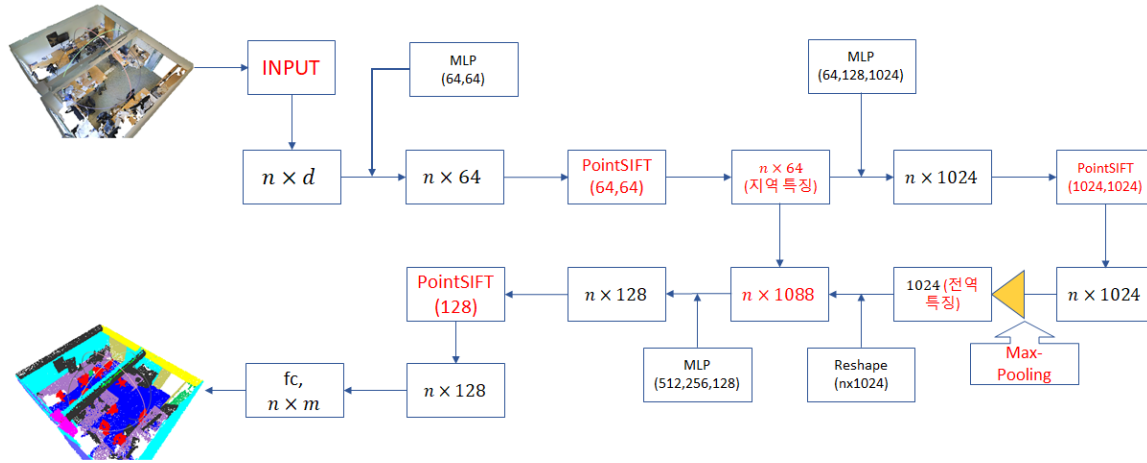


그림 2. 제안된 객체 분할 알고리즘 아키텍처(숫자는 레이어의 크기).

다. 입력 지점은 먼저 MLP를 사용하여 64차원 특징이 있는 벡터로 변환한다. 그런 다음 PointSIFT 모듈을 연결하여 지역 특징을 출력한다. 그 다음, 다층 퍼셉트론(MLP)을 사용하여 벡터를 1024 차원 공간에 매핑하고, PointSIFT 모듈을 연결한 후, 출력 벡터 행렬을 대칭 함수인 최대 풀링에 의해 전역 특징을 출력한다.

그런 다음 첫 번째 PointSIFT 모듈 출력의 지역 특징과 두 번째 PointSIFT 모듈 출력의 전역 특징을 연결하고 다층 퍼셉트론(MLP)을 사용하여 차수를 128 차원으로 줄이고 PointSIFT 모듈과 연결해서 특징을 출력한다. 마지막으로, 완전 연결 계층(fully connection layer)을 사용하여  $m$ 개의 클래스를 추출한다. 모든 다층 퍼셉트론(MLP)과 완전연결 계층에 ReLU 함수와 Batchnorm 함수를 추가된다.

### III. 시물레이션

실험에서는 공개 데이터 세트 S3DIS 데이터를 사용한다. S3DIS는 스탠포드 3D 실내 시맨틱 데이터 세트 (Stanford 3D Indoor Semantics Dataset)의 약자이며, 6개의 공간에 271개의 하우스 스캔 데이터가 포함되어 있으며, 포인트 클라우드의 각 포인트는 분할 카테고리 및 인스턴스 카테고리를 지정된다.

우리는 학습 데이터인 방을  $1m \times 1m \times 1m$  블록으로 나누고, 각 점은 XYZ, RGB 및 공간 정규화 위치 ( $0 \sim 1$ )의 9차원 벡터로 표시된다. 학습하는 동안 각 블록에서 4096개의 포인트를 랜덤으로 선택한다. 우리는 (6)와 같은 프로토콜을 따르고 k-fold 전략을 학습과 테스트 진행한다. 평가 척도에 대해 우리는 합집합에 대한 물체 분류별 교차 평균 (mIoU), 물체 분류별 정확도 평균 (mAcc) 및 전체 점별 정확도 (OA) 등을 사용하여 PointNet 알고리즘과 제안된 알고리즘의 성능을 비교하였다. 표1은 S3DIS 실험결과를 보여준다.

### IV. 결과

본 논문에서는 포인트넷 기반으로 객체 분할 방법을 제시하였다. 본 논문에서는 포인트 클라우드를 직접 조작하고 3차원 정보를 효과적으로 포착할 수 있는 포인트넷 기반 객체 분할 방법을 제안한다. 가장 진보된 몇

가지 방법과 비교할 때, 학습 및 수행 속도를 고려할 때 정확도가 높아진다. 향후 작업에는 공동 라이다 및 이미지 기반 종단 간 3차원 탐지를 확장하여 탐지 및 현지화 정확도를 더욱 향상시키고자 한다.

### ACKNOWLEDGMENT

이 논문은 2020년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(2018R1D1A1B07048868).

### 참고 문헌

- [1] Zhang Cuihong. Indoor 3D scene reconstruction based on Kinect depth camera [D]. Dalian University of Technology, 2013.
- [2] C. R. Qi, H. Su, K. Mo, and L. J. Guibas. 포인트넷: Deep learning on point sets for 3d classification and segmentation. Proc. Computer Vision and Pattern Recognition (CVPR), IEEE, 2017.
- [3] 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 28, pages 91-99. 2015.
- [4] X. Chen, K. Kundu, Z. Zhang, H. Ma, S. Fidler, and R. Urtasun. Monocular 3d object detection for autonomous driving. In IEEE CVPR, 2016.
- [5] X. Chen, K. Kundu, Y. Zhu, A. Berneshawi, H. Ma, S. Fidler, and R. Urtasun. 3d object proposals for accurate object class detection. In NIPS, 2015.
- [6] Yang Bisheng, Liu Yuan, Liang Fuxun, et al. Using Mobile Laser Scanning Data for Features Extraction of High Accuracy Driving Maps[C] // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Prague, Czech Republic: ISPRS, 2016: 433-439

method	mIoU(%)	mACC(%)	OA(%)	ceiling	floor	wall	beam	column	window	door	table	chair	sofa	bookcase	board	clutter
PointNet	47.55	65.50	78.23	88.41	88.15	68.83	39.62	20.22	49.61	51.68	55.04	42.88	7.72	38.61	31.83	35.48
Ours	<b>49.65</b>	<b>66.62</b>	<b>78.81</b>	87.29	<b>91.14</b>	66.49	<b>44.67</b>	<b>26.56</b>	48.17	<b>53.05</b>	53.70	<b>47.33</b>	<b>15.36</b>	<b>39.18</b>	<b>34.62</b>	<b>37.99</b>

표 1. S3DIS 데이터 세트에 대한 분할 실험 결과