

# 점군 위치 정보 인코딩 활용을 통한 딥러닝 기반 포인트 클라우드 샘플링 성능 개선 연구

이 후 경<sup>1)</sup> · 전 재 승<sup>1)</sup> · 홍 석 진<sup>1)</sup> · 유 진 우<sup>\*2)</sup>

국민대학교 자동차공학전문대학원<sup>1)</sup> · 국민대학교 자동차IT융합학과<sup>2)</sup>

## L-Net: Deep Learning-based Point Cloud Sampling Network through Positional Encoding

Hookyung Lee<sup>1)</sup> · Jae Seung Jeon<sup>1)</sup> · Seokjin Hong<sup>1)</sup> · Jinwoo Yoo<sup>\*2)</sup>

<sup>1)</sup>Graduate School of Automotive Engineering, Kookmin University, Seoul 02707, Korea  
<sup>2)</sup>Department of Automobile and IT Convergence, Kookmin University, Seoul 02707, Korea  
(Received 25 April 2023 / Revised 22 May 2023 / Accepted 24 May 2023)

**Abstract** : With the increased interest in point cloud processing, point cloud sampling techniques have also been attracting attention. As more models can calculate point clouds directly, working time has become an essential factor. The down-sampling process can solve this problem. Existing sampling algorithms have performed constant sampling methods regardless of the characteristics of task-model learning. However, this weakness has limitations in improving the performance of the task model in learning, and such task-agnostic methods perform too low when the sampling rate is high. Therefore, this paper proposes a novel down-sampling model network based on deep learning tasks. The proposed network utilizes fully connected layers to extract meaningful features from input sequences, and adds positional encoding instead of a conventional convolution concept. By introducing positional encoding into down sampling, the proposed network learns about the relationship between point clouds, and generates a task-oriented sampling methodology. Furthermore, the network incorporates skip connections to preserve important information during the down-sampling process. The proposed model outperforms several state-of-the-art models in terms of classification accuracy. With the fast inference time of the proposed network, it can be used in various applications, and our approach provides a promising solution for down-sampling tasks in different point cloud applications.

**Key words** : Deep learning(딥러닝), Positional encoding(위치 인코딩), Point cloud(점군), Down sampling(다운 샘플링), Classification(분류)

### 1. 서론

RGB-D 카메라, LiDAR 등 3D 센싱을 이용한 포인트 클라우드 생성 기술이 빠르게 발전하고 있으며, 컴퓨팅 속도가 증가하고 있으며, 3D 포인트 클라우드 분야에 대한 관심도 높아지고 있어 주목받고 있다.<sup>6-8)</sup> 또한, 딥러닝을 활용하여 기술을 결합한 논문도 나오고 있다.<sup>39-41)</sup> 이로 인해 다양한 분야에서 포인트 클라우드의 중요성이 커지고 있다. 포인트 클라우드는 실제 객체와 환경을 상

세하고 정확하게 표현하여 3D 데이터의 정밀한 측정, 분석 및 조작을 가능하게 한다. 로봇공학, 자율주행차, 가상현실, 건축, 문화유산 보존 등 다양한 분야에 응용되고 있다. 포인트 클라우드를 생성하는 기술이 지속적으로 발전됨에 따라 광범위한 데이터 셋에 포인트 클라우드 데이터 구조에 대한 의존도가 더욱 높아질 것으로 예상할 수 있다.

3D 포인트 클라우드 데이터는 일반적으로 이미지나

\*Corresponding author, E-mail: [jwyooy@kookmin.ac.kr](mailto:jwyooy@kookmin.ac.kr)

<sup>†</sup>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.

자연어 처리 데이터와 다른 불규칙하고 입체적인 특성을 가지고 있기 때문에 이를 처리하기 위한 새로운 방법이 필요했다. 이것은 점군이 밀도가 다양한 불규칙한 공간에 있기 때문이다.

연구 초기에는 3D 포인트 클라우드 데이터를 2D 이미지로 변환하여 처리하는 방법이 제안되었다. 이 방법은 3D 포인트 클라우드 데이터의 포인트를 이미지의 픽셀로 변환하여 이미지로 처리한다. 이 방법은 이미지 처리 분야에서 성공적으로 적용되었지만, 3D 포인트 클라우드 데이터의 복잡성과 다양성을 충분히 반영하지 못한다는 문제점이 있다. 따라서 3D 상태인 포인트 클라우드 딥러닝을 처리하기 위해서는 몇 가지 다른 측정 기준이 필요했다. 처음에는 투영 기반<sup>1,2)</sup> 및 Volumetric 기반 방법<sup>3-5)</sup>이 제안되어 각 포인트를 그리드로 변환하여 3D 포인트 클라우드를 처리하고 그리드의 기존 2D 이미지와 동일한 방식으로 컨볼루션 레이어를 사용하여 특징 추출을 수행한다. 이러한 방법들은 3D 공간의 불규칙한 점들을 격자 형태로 변환했기 때문에 격자 셀의 점들의 수가 고르지 못하여 특정 셀에서 정보의 손실이나 계산의 낭비를 초래하였다. 따라서, 그리드 변환의 문제를 극복하기 위해 Direct point 기반 전략이 등장했다. 일부 방법은 다중 공유 MLP(Multi-Layer Perceptron)를 사용하여 각 포인트를 독립적으로 모델링,<sup>9,11)</sup> 컨볼루션 커널의 유형에 따라 3D 컨볼루션 방법 등이 등장하였다.<sup>12-16)</sup>

포인트 클라우드는 분류,<sup>9,10)</sup> Instance segmentation,<sup>20,21)</sup> Semantic segmentation<sup>17,18)</sup> 및 Registration<sup>19)</sup>과 같은 다양한 분야에 적용되고 있다. 이 방법은 포인트 클라우드 데이터를 입력으로 사용하고 마지막 단계에서 로컬 함수를 통해 피쳐 정보를 모은다. 이러한 방법은 정확한 위치 정보를 유지하지만, 계산 비용은 포인트 수에 따라 선형적으로 증가하며 고용량의 고밀도 3D 포인트 클라우드 데이터를 처리하는 것은 여전히 어렵다. 따라서, 3D 포인트 클라우드의 데이터양을 줄이고 처리 효율을 향상시키기 위해 데이터를 샘플링하는 방법이 제안된다. 기존에는 작업에 구애받지 않는 Random 샘플링, Furthest Point 샘플링, Grid voxel 샘플링과 같은 휴리스틱 기반 샘플링 방법이 사용되었다. 그러나 이러한 방법은 다운스트림 작업에서 정보를 손실하거나 의미 없는 데이터를 선택할 수 있으므로 성능이 저하될 수 있다. 또한, 뒤의 작업을 전혀 고려하지 않아 성능 발전에 한계가 있다. 최근 작업 지향 샘플링 네트워크<sup>22-24)</sup>가 제안되어서 다운스트림 작업의 성능을 최적화하는 샘플링을 생성할 수 있다. S-Net 및 Samplenet은 딥 러닝을 사용한 샘플링 전략으로 특정 작업에 대해 우수한 성능을 보였다. 또한 APSNET은 포인트와의 관계에 더 많은 주의를 기울이기

위해 주의 기반 방법을 사용했다. 그러나 이러한 모델은 점 구름 사이의 전체 정보를 완전히 고려하지는 않는다. 그리고 순차적으로 입력받기 때문에 추론 속도가 느려지는 한계가 있다. 따라서 본 논문에서는 작업 지향 샘플링을 위해 포인트 클라우드 입력을 받아 선형 변환과 위치 인코딩 기능을 추가하여, 빠르고 강건한 모델을 제안한다. 기존 논문들에서는 컨볼루션 레이어를 사용하는 데 이는 위치정보는 보존해 주지만, 계산 과정이 오래걸릴 수 있다. 이에 반해 제안하는 L-Net은 계산 속도가 상대적으로 더 빠른 선형 레이어를 거친 후, 위치 인코딩을 더해주는 형식으로 모델을 구상했다. 제안된 모델은 포인트 클라우드 분류 분야에서 최고 수준의 성능 향상을 달성했다. 요약하자면, 주요 기여는 세 가지이다.

1. 우리는 작업 지향적인 목표로 Point cloud 특성과 맞지 않는 컨볼루션 레이어를 사용하지 않고 점들의 위치를 더욱 고려한 포인트 클라우드 샘플링 네트워크인 L-Net을 제안한다.
2. 제안하는 접근 방식은 포인트 클라우드의 분류 및 재구성과 관련된 작업에서 향상된 성능을 보여주며, 정확도와 속도 측면에서 기존의 Non-task oriented 방법과 Task oriented 방법 모두를 능가한다.
3. 이 접근 방식은 포인트 클라우드에서 일반적으로 발생하는 장거리 종속성 문제를 효과적으로 해결한다. 따라서 기본 기하학적 구조를 효과적으로 포착하여 희박한 샘플링 포인트 클라우드에 주목할 만한 영향을 미친다.

## 2. 관련 연구

### 2.1 포인트 클라우드 딥러닝

포인트 클라우드에 대한 딥러닝은 객체 감지, 세분화 및 분류와 같은 포인트 클라우드와 관련된 다양한 작업에 적용되었다. 예를 들어 Qi, Charles 등이 있다. Qi 등<sup>10)</sup> Charles 등<sup>9)</sup>은 포인트 클라우드 분류 및 세분화를 위해 제안되었으며 여러 벤치마크 데이터 세트에서 최고 수준의 성능을 달성했다. 또한, 포인트 클라우드에서 객체 감지를 위한 딥러닝 모델이 제안되었다.<sup>26,27)</sup> 포인트 클라우드 처리를 위해 제안된 다른 모델들이 있는데, 각각 로컬 기능과 적응형 컨볼루션 연산을 활용한다.<sup>16,28)</sup> 또한 일부 연구에서는 포인트 클라우드 생성 및 재구성 작업에 생성 모델<sup>29,30)</sup>을 사용하였다. Zhao 등<sup>31)</sup>은 이전 연구에서 사용된 글로벌 주의 체계와 달리 로컬 자기주의 메커니즘을 사용한다.<sup>34-37)</sup> 또한 벡터 주의 방법이 스칼라 주의 방법을 능가한다는 것을 입증하고 포인트 클라우드에서 위치 정보를 적절하게 처리하기 위한 위치 인코딩 방법을 도입한다. 본 논문에서는 적용하는 분야는

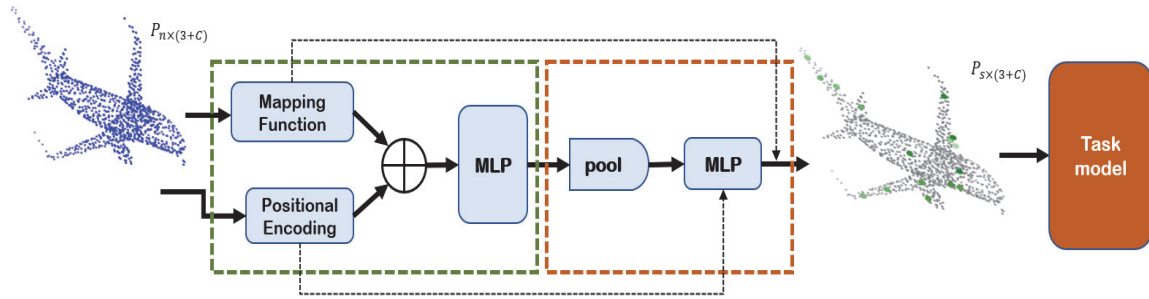


Fig. 1 Model architecture of the proposed L-Net

다르지만, Point transformer에서 영감을 받아 유사한 방법을 도출하도록 설계하였다. 또한, Positional encoding을 추가하여 위치 정보를 잃어버리지 않도록 하여 많은 단계를 거치지 않고 디코더를 구성했다.

### 2.2 점 구름 샘플링

랜덤 샘플링, 균일 샘플링, 원거리 포인트 샘플링(FPS) 및 그리드 샘플링과 같은 포인트 샘플링 작업에 구애받지 않는 알고리즘이 과거에 널리 사용되었다. 그 중에서도, FPS는 많은 최근 연구에서 여전히 인기 있는 선택이다.<sup>31-33)</sup> FPS는 빠르고 비교적 균일한 아웃풋을 내준다는 점에서 널리 사용되었지만 샘플링된 포인트가 사용될 다운스트림 작업을 완전히 고려하지 않아 잠재적인 성능 저하를 초래할 수 있다. 따라서, 최근에는 딥러닝을 활용한 다운샘플링 방법이 제안되었다.<sup>22-24)</sup> Dovrat 등<sup>22)</sup> 제안 샘플링의 효율성과 정확성은 학습 가능한 포인트 클라우드 샘플링 방법을 통해 향상되었다. Lang 등<sup>23)</sup>은 포인트 클라우드 샘플링을 위한 새로운 매커니즘을 소개한다 컨볼루션 레이어를 활용한 딥러닝 네트워크에 점 구름을 넣고 학습시키는 방법을 사용하였다. Yang 등<sup>24)</sup>는 주의 매커니즘을 활용하여 중요도 가중치를 할당함으로써 포인트 간의 관계를 향상시켜 보다 효과적인 샘플링 프로세스를 허용한다. 반면에, 제안하는 L-Net은 전체적인 것과 세부적인 포인트의 관계를 전부 보면서 장기적인 의존성을 완화하는 작업 지향적인 샘플링 방법론이다.

### 2.3 최근점 이웃 선택 알고리즘

최근 연구에서 가장 가까운 이웃(NN) 방법은 정보 융합에 널리 사용되었다. 그러나 신경망의 맥락에서, 주요 단점은 선택 규칙이 차별화될 수 없다는 것이다. 이를 해결하기 위해 Goldberger 등은 일련의 인접 후보에 대한 범주형 분포를 정의하여 NN 규칙의 확률적 완화를 제안했다. 제안하는 연구에서, 우리는 가장 가까운 이웃 선택

알고리즘을 점 구름 사이의 관계를 파악하기 위해 적용한다. 이렇게 하면 점 구름 사이의 관계를 더 잘 파악할 수 있다. 또한, 글로벌 영역의 정보를 고려하기 위해 Skip connection을 추가한다.

### 2.4 위치 인코딩

딥러닝 모델 영역에서 위치 인코딩은 일반적으로 입력 데이터의 위치 정보를 인코딩하는 데 사용되었다. 점 구름과 관련하여, 이전 연구에서는 점의 위치 정보를 통합하기 위해 Cartesian 좌표, 구면 좌표 또는 극좌표와 같은 기본 인코딩 기술을 사용했다. 그러나 이러한 방법은 정보 손실 및 표현력 부족 측면에서 한계가 있다. 이 문제를 해결하기 위해 일부 연구에서는 학습된 위치 인코딩 기술을 사용하여 더 많은 정보를 모델에 표현하려고 노력하였다. 이러한 기술은 일반적으로 위치 정보를 나타내는 연속 함수를 학습하는 것을 포함하며, 이는 점 구름에서 복잡한 공간 관계와 패턴을 더 잘 포착할 수 있다. 우리는 기존 MLP에서 사용했던 Transformer 아키텍처의 위치인코딩 방법을 사용하였다.

## 3. 제안하는 샘플링 네트워크 방법론(L-Net)

3장에서는 본 연구에서 제안한 알고리즘에 대하여 설명하고 전체적인 아키텍처의 구성을 설명한다.

L-Net은 크게 3가지 단계로 나뉜다. PointNet 모델로 피쳐를 추출하는 단계, 피쳐를 수집하는 단계 마지막으로 피쳐를 뽑는 단계로 나누어져 있다. 이번 섹션에서는 어떤 방식으로 피쳐를 뽑는지 각 단계 별로 상세하게 설명한다.

첫 번째 Feature extraction 단계에서는(Fig. 1의 초록색 점선 박스) 초기 포인트 클라우드 인풋에 Mapping function을 통해 Positional Encoding(P.E)과 Element-wise summation을 적용한다. 이때, Convolution layer를 사용하지 않는데 그 이유는 기존 Convolution layer로 레이어를 쌓는 방식은 인접한 지역의 정보를 공유하고, 이러한 지

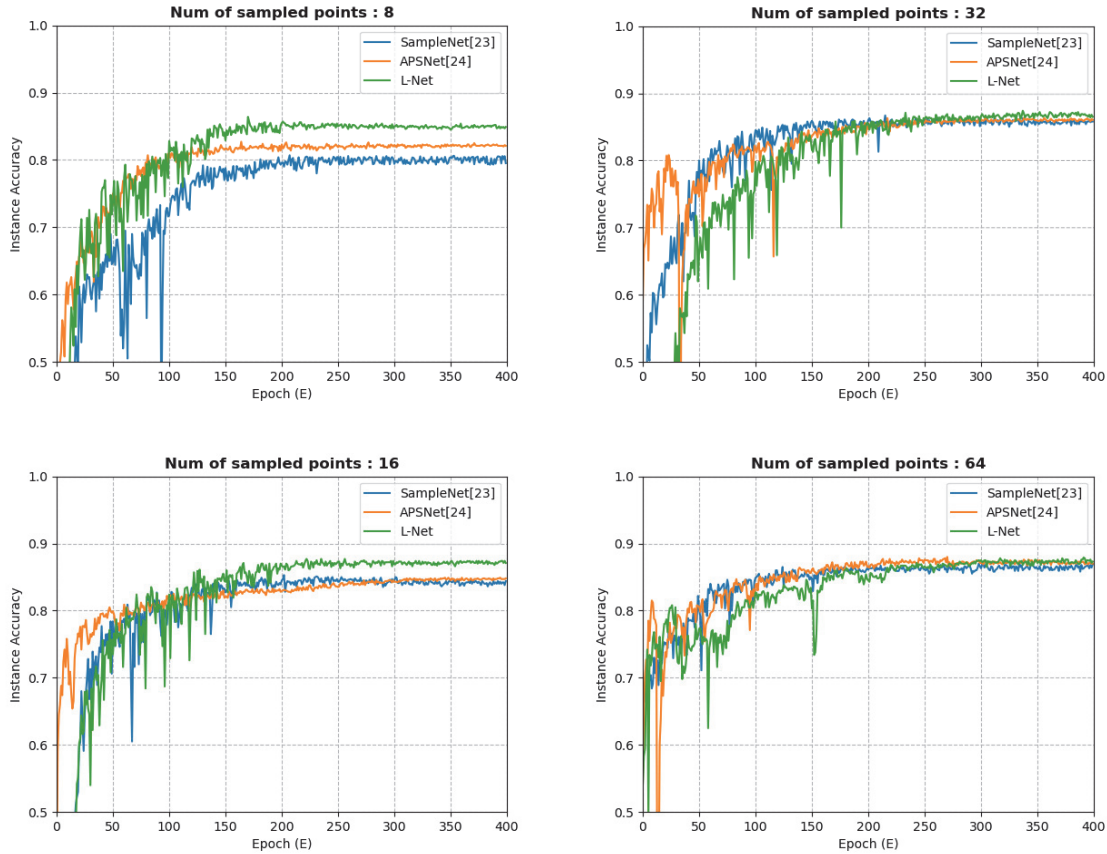


Fig. 2 Instance accuracy: This is the result of training three sampling models(SampleNet, APSNet, L-Net) on the ModelNet40 dataset after uniformly sampling 1024 examples. The sampling ratios used were 128, 64, 32, and 16. For example, if 8 examples were sampled from 1024, the sampling ratio would be 128. L-Net achieved much better performance in sparse areas. The classification accuracies are shown in Table 1

역 정보의 패턴을 학습하여 이미지나 Natural language processing 같은 분야에서의 특징 추출에 적합하지만, Point cloud는 데이터가 불규칙하게 분포하고 있어 적합한 형태의 학습 방법은 아니다. 따라서, 본 논문에서는 선형 레이어와 Positional encoding을 추가하였다. 그렇게 추출한 특징들을 종합하여서 특징을 수집하기 위해 Max-pooling 작업을 수행했다. 또한, 훈련 중에 매칭 단계, 즉 가장 가까운 이웃 선택에 미분 가능한 완화를 도입한다. 이것은 학습에서 미분 가능하지 않은 부분들을 없애주며 최종 출력에서 Weight를 최적화할 때 중요한 역할을 수행한다. 마지막으로, 최종 출력을 생성하기 위해 완전 연결 네트워크(Fully connected layer)를 포함한 다층 퍼셉트론 레이어(Multi-layer perceptron)가 사용되었다. 또한, 이 부분에서 위치 정보에 대한 표현력이 약해질 수 있는 문제가 있지만, 선형 다층 퍼셉트론 레이어를 사용함으로써 발생하는 위치 정보의 손실을 완화하기 위해 위치 인코딩이 추가적으로 삽입되었다. 이러한

위치 인코딩은 위치 정보를 보존하는 데 도움이 된다. Point cloud의 순서를 규정하기 어려운 데이터 점을 해결할 수 있다. 그뿐만 아니라, 위치 인코딩을 통해 입력 데이터에서 추가되는 정보는 기존의 특성 정보를 강화시키는데 큰 효과를 발휘할 수 있다. Point cloud의 특성 정보는 한정적이므로 모델의 정확도 향상에 크게 기여할 수 있다. 마지막으로, 과적합을 방지하기 위해 정규화 기법으로 드롭아웃을 사용했다. 이렇게 드롭아웃을 사용하면 기존에 비해 피팅이 되지 않는 범용적인 결과값을 얻을 수 있다. 결과에는 작업에 대한 점들의 수가 표시된다. 이 결과를 통해 멀티태스킹 학습을 수행한다. 멀티태스킹 학습 접근 방식을 채택하여 단일 모델 내에서 모든 작업을 효율적으로 처리할 수 있다. 샘플링 모델과 기존 Task model의 동시 학습을 통해 파라미터를 공유하는 효과를 가져오는데, 이는 관심 있는 영역에 대한 데이터 분포를 더욱 효과적으로 활용하고 기존보다 더욱 많은 작업 간 상관관계를 고려하여 모델의 성능을 더욱 향상시



Table 1 Classification accuracies of five sampling methods on ModelNet40. All experimented in the same environment

Sampling ratio	128	64	32	16
RS	8.7	24.87	54.53	79.26
FPS	24.31	55.12	76.92	87.53
Samplenet	80.71	85.32	86.38	87.10
APSnet	82.72	84.89	86.66	<b>88.00</b>
L-Net	<b>86.06</b>	<b>87.44</b>	<b>87.49</b>	87.88

킬 수 있다. 또한 Task model을 Frozen 시키고 해당 모델만 학습을 시킨 후, 다른 Task에도 적용하는 실험을 해볼 수 있다. 모델 네트워크 구조에 대한 자세한 내용은 Fig. 1에 나와 있다.

#### 4. 실험 결과 및 토론

4절에서는, 제안하는 L-Net의 성능이 기존의 샘플링 방법보다 우수하다는 것을 보여줄 것이다. 제안하는 L-Net은 분류 문제에 대해서 실험을 수행했고, 포인트 클라우드 샘플링 비율이 높은 영역에서 성능이 특히 우수함을 입증했다. 또한, Positional encoding을 여러 가지 위치에 더해보면서 다양한 버전의 L-Net에 대해 실험을 여러 차례 진행하여 최적의 값을 찾았다. 제안하는 L-Net은 PyTorch에서 구현하였으며, Learning rate와 Weight decay가 각각 128, 0.9 및 0.0001로 설정된 배치 크기인 SGD 옵티마이저를 사용하였다. 게다가, 우리는 모든 실험에서 400회를 반복하여 훈련을 실시했다. 분류 문제에서 ModelNet 40 데이터 세트<sup>38)</sup>를 사용했다. 초기의 데이터 셋에서 균일하게 샘플링된 1024개 지점에 대한 실험을 수행하고 공식 웹사이트에 제공된 열차 테스트 분할 세트를 사용하여 훈련 및 평가했다.

또한 이전에 모든 논문에서 저조한 성능을 보였던 샘플링 비율이 16, 32, 64, 128인 희소 포인트에 실험을 집중했다. 한 가지의 중요한 전제는 원래의 작업 모델의 성능을 초과할 수 없다고 가정한다. 예를 들어, 분류의 경우, PointNet이 기본 모델로 정의되기 때문에, 샘플링되지 않은 PointNet 즉, 1024개 중 1024개를 뽑는 샘플링 비율이 1인 것에서 성능이 90.4라면, PointNet의 성능은 샘플링된 모델을 포함한 모델이 90.4를 초과할 수 없다는 전제가 있다. Fig. 3을 보면, ModelNet40의 8, 16, 32개를 샘플링한 포인트에 대한 결과인데, 회색이 원래의 Ground truth 값이며, 녹색이 샘플링 된 결과값이다. 왼쪽의 결과는 L-Net의 결과이고 오른쪽은 APSNet의 결과이다. 전체적인 모양에서, 제안하는 L-Net의 결과 값

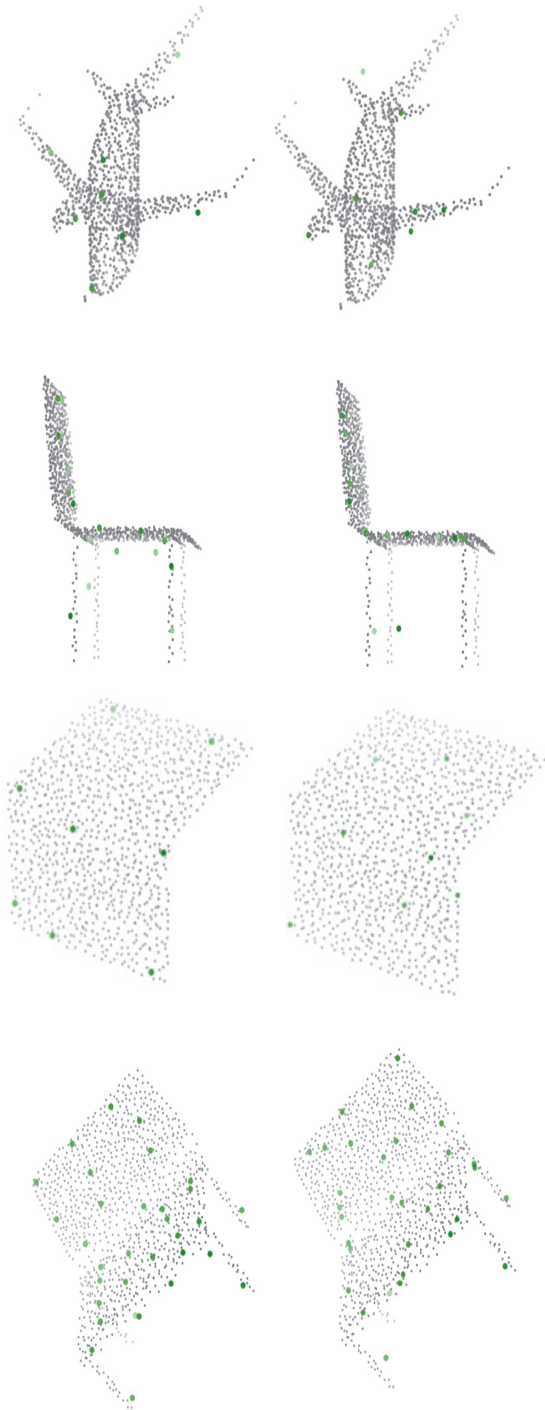


Fig. 3 Visualized Sampled Points

은 APSNet의 결과 값보다 더 균일하게 표현된다. 전체적으로 제안된 방법이 물체의 모서리를 더 잘 파악하는 경향이 있으며, 의자와 같은 물체의 경우 다리를 더 잘 파악하는 경향이 있다. 제안하는 L-Net은 샘플링 비율

Table 2 Class-specific accuracies of two sampling methods on ModelNet40

Sampling ratio		Laptop	Chair	Nightstand	Airplane
128	SampleNet	81.32	91.38	38.18	97.51
	APSNNet	83.33	95.56	42.18	<b>98.52</b>
	L-Net	<b>92.68</b>	<b>95.61</b>	<b>68.88</b>	96.6
64	SampleNet	85.23	94.38	59.38	98.11
	APSNNet	88.88	95.52	64.67	<b>99.49</b>
	L-Net	<b>97.56</b>	<b>97.53</b>	<b>72.72</b>	97.55
32	SampleNet	89.29	86.38	67.21	98.89
	APSNNet	90.90	94.68	67.39	<b>99.50</b>
	L-Net	<b>90.91</b>	<b>97.49</b>	<b>74.03</b>	98.20

이 높을 때, 예를 들어 1024개 중 8개를 뽑는 샘플 비율이 128인 경우, 다른 비교군들과의 성능 차이가 확연하게 나타나는 것을 확인할 수 있다. 자세한 클래스별 결과를 4가지 종에 대해 정확도를 분석한 결과는 Table 2에 나와 있다.

### 5. 결론

본 논문에서는 포인트 클라우드 샘플링을 위한 트랜스포머 기반 샘플링 네트워크인 L-Net을 소개한다. S-Net과 SampleNet은 MLP 기반 방법을 사용하여 샘플링 프로세스를 생성하고 APSNet은 주의 기반 모델을 사용하는 반면, 제안하는 L-Net은 점 구름의 이산적인 특성을 제대로 반영하지 못하는 컨볼루션 레이어를 사용하지 않는 대신 선형레이어를 사용함으로써 속도를 높이고 점군 사이의 관계를 잘 이해할 수 있는 모델을 구축함으로써 정확도 또한 높였다. Positional Encoding과 잘 설계된 모델을 통해 각 지점 간의 관계를 보다 잘 이해할 수 있는 장점이 있다. 또한, 작업 모델과 동시에 학습을 하는 기법을 추가하여 학습의 관계를 보다 잘 이해할 수 있으므로 작업 목적에 잘 맞는 가변적인 샘플링 기법이다. 제안하는 L-Net은 특히 희소 데이터에서 우수한 성능을 보여준다. 또한 학습된 샘플링 데이터를 작업 네트워크와 별도로 사용할 수 있기 때문에 많은 시나리오에서 유용성을 입증하고 있다.

### 후 기

본 연구는 산업통상자원부 산업핵심기술개발사업 “혼합현실 기반 자율주행 부품 시스템 평가 기술개발(20014476)” 과제의 지원으로 수행되었음.

### References

- 1) F. J. Lawin, M. Danelljan, P. Tosteberg, G. Bhat, F. S. Khan and M. Felsberg, “Deep Projective 3D Semantic Segmentation,” Computer Analysis of Images and Patterns: 17th International Conference (CAIP), Part I 17, pp.95-107, 2017.
- 2) A. Boulch, B. Le Saux and N. Audebert, “Unstructured Point Cloud Semantic Labeling using Deep Segmentation Networks,” Proceedings of the Workshop on 3D Object Retrieval, pp.17-24, 2017.
- 3) D. Maturana and S. Sebastian, “Voxnet: A 3D Convolutional Neural Network for Real-time Object Recognition,” 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.922-928, 2015.
- 4) G. Riegler, A. Osman Ulusoy and A. Geiger, “Octnet: Learning Deep 3D Representations at High Resolutions,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3577-3586, 2017.
- 5) P. -S. Wang, Y. Liu, Y. -X. Guo, C. -Y. Sun and X. Tong, “O-CNN: Octree-based Convolutional Neural Networks for 3D Shape Analysis,” ACM Transactions on Graphics (TOG), Vol.36, No.4, pp.1-11, 2017.
- 6) Z. Liang, Y. Guo, Y. Feng, W. Chen, L. Qiao, L. Zhou, J. Zhang and H. Liu, “Stereo Matching Using Multi-level Cost Volume and Multi-scale Feature Constancy,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.43, No.1, pp.300-315, 2019.
- 7) Y. Guo, F. Sohel, M. Bennamoun, M. Lu and J. Wan, “Rotational Projection Statistics for 3D Local Surface Description and Object Recognition,” International Journal of Computer Vision, Vol.105,

- pp.63-86, 2013.
- 8) Y. Guo, M. Bennamoun, F. Sohel, M. Lu and J. Wan, "3D Object Recognition in Cluttered Scenes with Local Surface Features: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.36, No.11, pp.2270-2287, 2014.
  - 9) R. Q. Charles, H. Su, M. Kaichun and L. J. Guibas, "Pointnet: Deep Learning on Point Sets for 3D Classification and Segmentation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.652-660, 2017.
  - 10) C. R. Qi, L. Yi, H. Su and L. J. Guibas, "Pointnet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space," *Advances in Neural Information Processing Systems* 30, pp.5105-5114, 2017.
  - 11) M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Póczos, R. R. Salakhutdinov and A. J. Smola, "Deep Sets," *Advances in Neural Information Processing Systems* 30, 2017.
  - 12) W. Song, Z. Liu, Y. Tian and S. Fong, "Pointwise CNN for 3D Object Classification on Point Cloud," *Journal of Information Processing Systems*, Vol.17, No.4, pp.787-800, 2021.
  - 13) N. Thomas, T. Smidt, S. Kearnes, L. Yang, L. Li, K. Kohlhoff and P. Riley, "Tensor Field Networks: Rotation-and Translation-equivariant Neural Networks for 3D Point Clouds," *arXiv:1802.08219*, 2018.
  - 14) F. Groh, P. Wieschollek and H. P. Lensch, "Flex-convolution: Million-scale Point-cloud Learning Beyond Grid-worlds," *Asian Conference on Computer Vision*, pp.105-122, 2018.
  - 15) W. Wu, Z. Qi and F. Li, "Pointconv: Deep Convolutional Networks on 3D Point Clouds," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.9621-9630, 2019.
  - 16) H. Thomas, C. R. Qi, J. -E. Deschaud, B. Marcotegui, F. Goulette and L. Guibas, "KPConv: Flexible and Deformable Convolution for Point Clouds," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp.6411-6420, 2019.
  - 17) J. Li, B. M. Chen and G. H. Lee, "So-net: Self-organizing Network for Point Cloud Analysis," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.9397-6406, 2018.
  - 18) Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein and J. M. Solomon, "Dynamic Graph Cnn for Learning on Point Clouds," *ACM Transactions on Graphics (TOG)*, Vol.38, No.5, pp.1-12, 2019.
  - 19) Y. Aoki, H. Goforth, R. A. Srivatsan and S. Lucey, "PointNetLK: Robust & Efficient Point Cloud Registration using Pointnet," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.7163-7172, 2019.
  - 20) W. Wang, R. Yu, Q. Huang and U. Neumann, "SGPN: Similarity Group Proposal Network for 3D Point Cloud Instance Segmentation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.2569-2578, 2018.
  - 21) X. Wang, S. Liu, X. Shen, C. Shen and J. Jia, "Associatively Segmenting Instances and Semantics in Point Clouds," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.4096-4105, 2019.
  - 22) O. Dovrat, I. Lang and S. Avidan, "Learning to Sample," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019.
  - 23) I. Lang, A. Manor and S. Avidan, "SampleNet: Differentiable Point Cloud Sampling," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.7578-7588, 2020.
  - 24) Y. Yang, X. Yang and S. Ji, "APSNet: Attention Based Point Cloud Sampling," *arXiv:2210.05638*, 2022.
  - 25) A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin, "Attention is All You Need," *Advances in Neural Information Processing Systems* 30, pp.6000-6010, 2017.
  - 26) S. Shi, X. Wang and H. Li, "PointRCNN: 3D Object Proposal Generation and Detection from Point Cloud," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.770-779, 2019.
  - 27) S. Shi, C. Guo, L. Jiang, Z. Wang, J. Shi, X. Wang and H. Li, "PV-RCNN: Point-Voxel Feature Set Abstraction for 3D Object Detection," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.10529-10538, 2020.
  - 28) M. Jiang, Y. Wu, Z. Zhao and C. Lu, "PointSIFT: A SIFT-like Network Module for 3D Point Cloud Semantic Segmentation," *ArXiv:1807.00652*, 2018.
  - 29) G. Yang, X. Huang, Z. Hao, M. -Y. Liu, S. Belongie and B. Hariharan, "PointFlow: 3D Point Cloud Generation

- with Continuous Normalizing Flows,” Proceedings of the IEEE/CVF International Conference on Computer Vision, pp.4541-4550, 2019.
- 30) M. Vakalopoulou, G. Chassagnon, N. Bus, R. Marini, E. I. Zacharaki, M. -P. Revel and N. Paragios, “AtlasNet: Multi-atlas Non-linear Deep Networks for Medical Image Segmentation,” Medical Image Computing and Computer Assisted Intervention-MICCAI 2018: 21st International Conference, Proceedings, Part IV 11, pp.658-666, 2018.
- 31) H. Zhao, L. Jiang, J. Jia, P. Torr and V. Koltun, “Point Transformer,” Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021.
- 32) F. Wu, A. Fan, A. Baeovski, Y. N. Dauphin and M. Auli, “Pay Less Attention with Lightweight and Dynamic Convolutions,” ArXiv:1901.10430, 2019.
- 33) X. Lai, J. Liu, L. Jiang, L. Wang, H. Zhao, S. Liu, X. Qi and J. Jia, “Stratified Transformer for 3D Point Cloud Segmentation,” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.8500-8509, 2022.
- 34) H. Hu, Z. Zhang, Z. Xie and S. Lin, “Local Relation Networks for Image Recognition,” Proceedings of the IEEE/CVF International Conference on Computer Vision, pp.3464-3473, 2019.
- 35) Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin and B. Guo, “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,” Proceedings of the IEEE/CVF International Conference on Computer Vision, pp.10012-10022, 2021.
- 36) H. Zhao, J. Jia and V. Koltun, “Exploring Self-attention for Image Recognition,” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.10076-10085, 2020.
- 37) S. Hochreiter and J. Schmidhuber, “Long Short-term Memory,” Neural Computation, Vol.9, No.8, pp.1735-1780, 1997.
- 38) Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang and J. Xiao, “3D Shapenets: A Deep Representation for Volumetric Shapes,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.1912-1920, 2015.
- 39) H. Lee, J. S. Jeon and J. Yoo, “A Study on Improving the Accuracy of 3D Point Cloud Data Classification by Deep-learning Based Point Sampling,” KSAE Spring Conference Proceedings, p.1050, 2022.
- 40) B. Yu and S. Kee, “3D LiDAR Depth Completion Based on Shift-convolution Using Adjacent Pixel Information,” Transactions of KSAE, Vol.29, No.9, pp.803-809, 2021.
- 41) M. Sun, D. Paek and S. -H. Kong, “A Study on Deep Learning Based Lidar Object Detection Neural Networks for Autonomous Driving,” Transactions of KSAE, Vol.30, No.8, pp.635-647, 2022.