

# 자가학습과 지식증류 방법을 활용한 LiDAR 3차원 물체 탐지에서의 준지도 도메인 적응

## Semi-Supervised Domain Adaptation on LiDAR 3D Object Detection with Self-Training and Knowledge Distillation

우정완<sup>1</sup>·김재열<sup>2</sup>·임성훈<sup>†</sup>

Jungwan Woo<sup>1</sup>, Jaeyeul Kim<sup>2</sup>, Sunghoon Im<sup>†</sup>

**Abstract:** With the release of numerous open driving datasets, the demand for domain adaptation in perception tasks has increased, particularly when transferring knowledge from rich datasets to novel domains. However, it is difficult to solve the change 1) in the sensor domain caused by heterogeneous LiDAR sensors and 2) in the environmental domain caused by different environmental factors. We overcome domain differences in the semi-supervised setting with 3-stage model parameter training. First, we pre-train the model with the source dataset with object scaling based on statistics of the object size. Then we fine-tune the partially frozen model weights with copy-and-paste augmentation. The 3D points in the box labels are copied from one scene and pasted to the other scenes. Finally, we use the knowledge distillation method to update the student network with a moving average from the teacher network along with a self-training method with pseudo labels. Test-Time Augmentation with varying  $z$  values is employed to predict the final results. Our method achieved 3rd place in ECCV 2022 workshop on the 3D Perception for Autonomous Driving challenge.

**Keywords:** Deep Learning, LiDAR, Computer Vision, 3D Object Detection

### 1. 서 론

인공지능의 발전과 함께 자율주행 기술은 활발히 연구되고 있으며 이에 따라 딥러닝 방법 기반의 자율주행 인지(perception) 모델을 학습하기 위한 데이터셋 구축 사업 또한 다양하게 이루어지고 있다. 주행 환경에서 외부 요인에 강인하게 높은 신뢰도로 객체들을 인지하기 위해 Camera, LiDAR, RADAR

등의 다양한 센서들의 조합을 사용하여 학습 데이터를 수집하여 활용하고 있고, 이렇게 수집한 데이터와 레이블을 학술적 목적으로 공개하기도 한다<sup>1-4)</sup>.

하지만 공개된 여러 개의 데이터셋들 간에는 센서 위치 및 각도, 센서의 종류, 다른 센서와의 조합 등의 이유로 동일한 장면을 촬영한다 해도 센서 측면에서의 차이가 존재한다. 또한 각 데이터셋들이 촬영된 국가와 도로 상황, 날씨 등이 모두 다르기 때문에 외부 환경적으로도 차이가 존재한다. 특히 LiDAR 센서의 경우, 하나의 센서에서 레이저(Laser)를 쏘아, 돌아오는 빛의 이동시간(TOF, Time Of Flight)을 측정하는 방식이기 때문에 센서의 위치에 따라 얻어지는 데이터에 큰 차이가 존재한다. 예를 들면, 차량의 지붕보다 낮은 높이에 LiDAR 센서를 부착하는 경우, 이 센서는 다른 비슷한 차량들의 지붕에 맺히는 3차원 점들을 하나도 얻을 수 없다.

현 시점에서 실용적으로 사용되는 LiDAR 센서 중에는 기계식(mechanical) 방법과 MEMS (Microelectromechanical) 거

Received : May. 2. 2023; Revised : Jun. 24. 2023; Accepted : Jul. 5. 2023

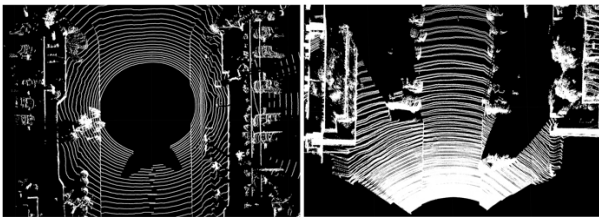
※ This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2023-00210908)

1. Integrated M.S. / Ph.D. Program Student, Department of Electrical Engineering and Computer Science DGIST, Daegu, Korea (friendship1@dgist.ac.kr)

2. Ph.D. Student, Department of Electrical Engineering and Computer Science DGIST, Daegu, Korea (jykim94@dgist.ac.kr)

† Associate Professor, Corresponding author: Department of Electrical Engineering and Computer Science DGIST, Daegu, Korea (sunghoonim@dgist.ac.kr)

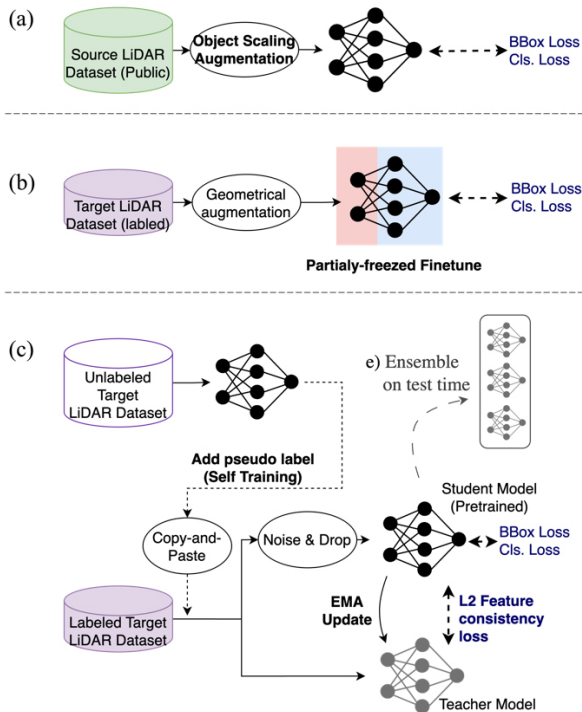
울 기반 고정형(solid-state) 방법이 존재하는데, 정해진 개수의 수직으로 정렬된 광선을 모터를 이용해 360도 회전하며 쏘아 보내 데이터를 얻는 기계식과는 달리, 고정형 LiDAR 센서는 하나의 레이저 광선만을 빠르게 회전하는 작은 거울에 반사시켜 데이터를 얻기 때문에 패턴이 외부 진동 등에 따라 광선의 발사 각도가 규칙적이지 않으며 [Fig. 1]에서 확인할 수 있듯이 이 두 센서 간에는 맺히는 패턴이 다르다<sup>7,8)</sup>. 이와 같이 센서와 환경 차이에 따라 데이터셋 사이에 도메인 차이가 존재하기 때문에 하나의 도메인에서 학습한 모델을 가지고 다른 처음 보는 도메인에 적용하려고 했을 때, 도메인 차이에 따른 인지 모델의 성능 하락 문제가 필연적으로 발생한다. 또한 환



[Fig. 1] 3D point cloud distribution difference according to LiDAR sensor types. Left: 360 degrees rotating mechanical lidar sensor used in the Waymo dataset. Right: Solid state lidar sensor used by Innoviz dataset

경적인 측면에서도 국가마다 다른 차량의 크기에 의해 3차원 객체 탐지에서 도메인 적응이 어렵다는 점은 이미 잘 알려져 있다<sup>9)</sup>.

본 논문에서는 이와 같은 센서 및 환경 특성 변화에 따른 도메인 차이를 극복하여 하나의 데이터셋에서 학습된 모델을 적은 학습 데이터를 가진 다른 도메인으로 적용시키는 준지도 도메인 적응(semi-supervised domain adaptation) 방법을 연구한다. 이를 위해 Waymo와 Innoviz 데이터셋을 사용하여 도메인 차이에 따른 성능 하락을 보완할 수 있는 방법을 연구한다. 공중에 공개된 데이터셋 중 하나인 Waymo<sup>11)</sup> 데이터셋은 360도 기계식 LiDAR 센서를 차량 상부에 탑재하여 미국에서 촬영하였으며, Innoviz 데이터셋은 Innoviz에서 개발한 고정형 LiDAR 센서를 차량 전방 하단에 두고 이스라엘에서 전방만을 촬영한다. 1150개의 경로에서 찍은 23만개의 LiDAR 센서 프레임에서 annotate된 1200만개의 3D 박스 레이블을 가지고 있는 Waymo 데이터셋과는 달리, Innoviz 데이터셋에서는 오직 103개의 프레임만이 주어진다. 우리는 Waymo의 풍부한 학습 데이터를 활용하여 Innoviz 데이터셋에서도 높은 3차원 객체 탐지 정확도를 얻는 것을 목표로 하는 준지도 도메인 적응 문제 해결을 위해 지식 증류(knowledge distillation)과 자가학습(self-training) 방법을 사용한다.



[Fig. 2] A schematic diagram of the domain adaptation method we designed. (a) Pre-training is performed in the source domain based on statistical information, and (b) fine-tuning is performed in the target domain. And finally, (c) further adjustments involving knowledge distillation and self-training

## 2. 방법 소개

본 연구에서는 준지도 도메인 적응을 수행하는 과정을 세 단계로 나눈다. [Fig. 2]에서 볼 수 있듯이, (a) 먼저 원천(source) 데이터셋에서 미리 학습을 진행한 후, (b) 목표(target) 도메인에서 1차적인 학습을 진행한다. (c) 마지막으로 학습된 모델을 사용하여 추가적인 자가 학습을 진행하여 모델의 정확도를 향상시킨다.

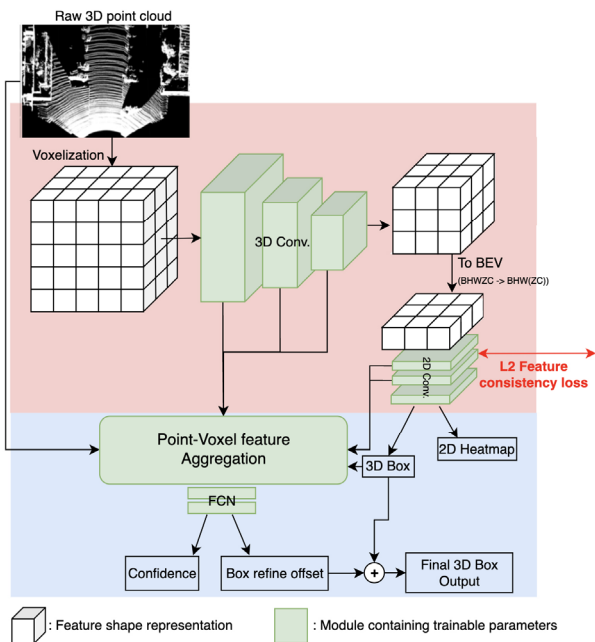
### 2.1 통계 정보를 활용한 원천 도메인에서의 사전 훈련

우리는 Waymo에서 먼저 학습된 CenterPoint<sup>10)</sup> 기반의 PV-RCNN++<sup>[11,12]</sup> 모델을 기본모델(baseline)으로 사용한다. Waymo 데이터셋에서 학습할 때, 우리는 객체 클래스별로 목표(target)와 원천(source) 데이터셋의 물체 크기 평균 크기가 일치하도록, 클래스별로 스케일링 변환을 적용하여 원천 도메인에서 객체의 박스 안 점군의 위치를 객체 중심을 기준으로 뺀 채로 목표 도메인에 맞추었다. 또한, 3차원 점들의 좌표를 옮겨 센서의 위치를 목표 데이터셋과 동일하게 보정하였으며, 센서별로 상이한 LiDAR 센서의 반사 강도(intensity) 값을 무시하였다.

Waymo 3차원 객체 탐지 데이터셋에는 트럭 레이블이 별도로 존재하지 않고 일반차량 클래스에 포함되어 있는데, Innoviz 데이터셋에서 객체 탐지 레이블에는 트럭 클래스가 따로 존재한다. 이러한 차이를 해결하기 위해, Waymo 3차원 객체 탐지 데이터셋에서 레이블을 추출하여 Waymo 3차원 객체 탐지 데이터셋의 차량 클래스를 일반 차량과 트럭 클래스로 세분화하여 학습한다. 그리고 원천 데이터셋은 360도에서 얻어지는 반면 목표 데이터셋 전방만 존재하기 때문에 원천 데이터셋의 전방과 후방을 반으로 잘라 각기 따로 학습에 사용하여 데이터셋을 두 배로 이용할 수 있는 증강(augmentation) 방법을 사용한다.

## 2.2 목표 도메인에서의 미세조정 학습

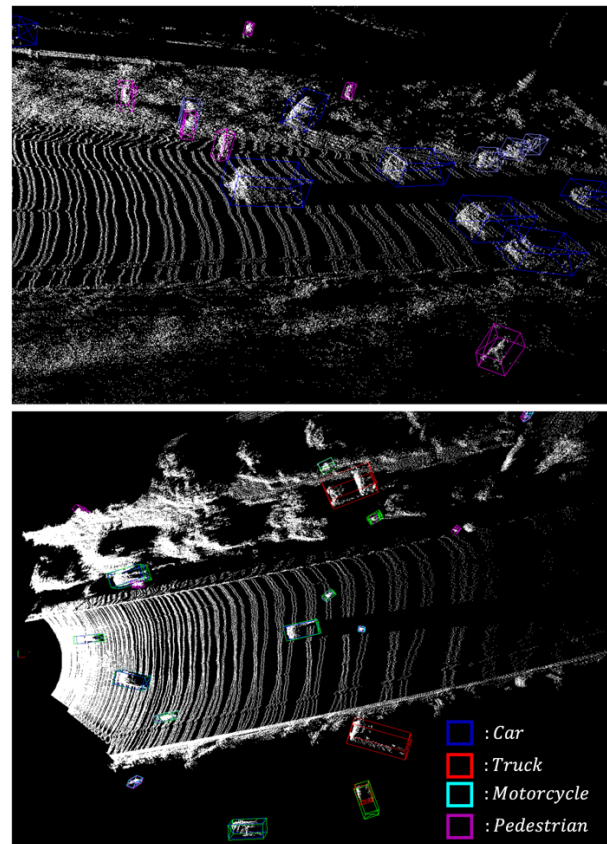
본 논문에서 기본 모델로 사용한 PV-RCNN+<sup>[11]</sup> 모델은 2-stage 기반 3차원 객체 탐지 방법으로, RPN (Region Proposal Network) 에서 제시된 영역에 대해 추가적인 추론을 거쳐 최종 정제된 결과를 예측한다. 우리는 기존에 학습된 모델에서 출력에 가까운 RPN 네트워크 부분까지만을 학습하였으며, 그 외의 네트워크 부분에서는 batch normalization<sup>[13]</sup>의  $\alpha$ ,  $\beta$  값을 선택적으로 학습하였다. 이를 통해, 기 학습된 네트워크 중에서 후보 영역을 제한하는 RPN 부분만을 새로운 도메인에서 추가



[Fig. 3] Detailed description of our model architecture. In the case of partially frozen finetune, only the red background area is trained, and only the batch norm parameter is updated for the blue color area. Red arrows indicate L2 feature consistency loss. The loss is given by the L2 distance from the 2D convolution layer of the completely identical teacher network structure

학습을 진행하여, 변화하는 입력단의 네트워크만 효율적이고 안정적인 학습을 목표로하였다. 네트워크의 어떤 부분을 부분적으로 학습하는지는 [Fig. 3]에서 그림으로 확인할 수 있다.

목표 데이터셋은 소량의 레이블만을 가지고 있기 때문에, 목표 데이터셋에서 학습이 길어질수록 과적합을 쉽게 초래한다. 더 일반화된 학습을 위해 우리는 다양한 데이터 증강 방법을 적용하였다. 소량의 레이블을 기반으로 전체 장면에서 객체 부분의 점군 데이터만을 떼어낸 후, 이 객체들을 다른 레이블 되지 않은 장면에 붙여넣기 하는 방식의 복사-붙여넣기 증강<sup>[14]</sup>을 진행하였다. 단순히 같은 위치로 복사 붙여넣기 하는 기존 방법과는 다르게, 물체의 거리는 일정 범위 내에서 그대로 유지하지만 위치할 수 있는 각도를 다양하게 주는 것으로 더 다양한 증강이 가능하도록 하였다. 증강 기법을 사용한 학습 데이터 예시는 [Fig. 4]에서 확인할 수 있다. LiDAR 센서의 직진성으로 인해 생겨나는 음영 부분이 붙여 넣어진 물체에 대해서는 생기지 않는 것을 확인할 수 있다. 또한, 다양한 각도로 회전시켜 위치시키기 때문에 도로를 벗어난 위치에 붙여 넣어진 물체도 존재한다. 이러한 현실과 다른 데이터 증강 기법의 미비점은 본 연구에서는 무시하였지만, 미래에 해결할



[Fig. 4] Training data sample with Copy-and-Paste augmentation method. There is no LiDAR shadow on pasted sample. Green box is prediction

과제로 남겨둔다.

또한 장면 전체, 또는 레이블 된 객체의 점군 데이터에 임의의 회전과 이동 변환을 주는 증강 방법 또한 사용했다. 추가로 [0.95, 1.05] 범위의 랜덤 스케일링과 전방 축 기준 좌우 반전을 사용한다.

### 2.3 Knowledge Distillation 및 Self-Training

이렇게 소량의 목표 학습 데이터에서 추가 미세 조정 학습을 완료한 모델을 남아있는 다른 레이블 되지 않은 목표 데이터셋에서 추론하여 pseudo-label 박스를 얻고, 얻어진 박스 영역 안의 3차원 점들을 다시 복사하여 붙여 넣는 방식으로 추가 학습을 진행하는 자가 학습(self-training) 방식을 적용하였다.

적은 데이터에서도 높은 일반화 성능을 얻기 위해, 자기지도 학습(self-supervised learning)에서 주로 사용되는 BYOL<sup>[15]</sup>을 참고하여, 우리는 두 개의 동일한 구조의 학생과 교사 모델을 두어 지식 증류(knowledge distillation) 방법을 적용한다. 학생 모델에는 더 강한 수준의 가우시안 노이즈와 point random drop 방법을 적용한 입력을 사용하고, 교사 모델에는 상대적으로 약한 수준의 왜곡을 적용한 입력을 넣어 지식 증류 학습을 진행한다.

$$loss_{feat} = \frac{1}{HWC} \sum_{h,w} \sum_c \| F_{h,w,c}^s - F_{h,w,c}^t \|_2 \quad (1)$$

식 (1)을 따라, 모델 backbone에서 학생과 교사 모델에서 H,W,C의 크기를 가지는 2D backbone의 마지막 특징치  $F^s$ 와  $F^t$ 의 모든 원소가 같아야 한다는 L2 손실 함수를 주어 학생 모델을 학습한다. 교사 모델은 역전파 과정 없이 학생 모델로부터 지수 이동 평균(Exponential Moving Average) 방식을 통해서만 업데이트 된다. 구체적으로는 [Fig. 3]에서 볼 수 있듯이 2D Backbone 네트워크의 레이어 하나에 손실 함수를 적용한다. 이와 같이 강한 수준의 임의의 왜곡을 knowledge distillation 방법을 통해 학습하게 하는 것으로, 타겟 도메인의 적은 양의 데이터셋에서도 과적합 상황을 회피하고 높은 일반화 성능을 얻도록 설계하였다.

마지막으로, 두 도메인 간 도로 표면 높이의 차이를 대처하기 위해 입력 3차원 점군 데이터의 Z축 방향으로 [-15, -10, -5, +5, +10] 센티미터로 5가지 오프셋을 주어 얻은 다양한 결과를 얻고, 이 결과들을 하나로 합치고 예측된 점수 기반으로 NMS (Non-Maximum Suppression)을 진행하는 Test-Time-Augmentation 방법을 적용하여 최종 결과를 얻었다.

## 3. 결과 및 토론

### 3.1 평가 방법

우리는 결과 3차원 객체 탐지 능력을 평가하기 위해 XYIoU metric을 사용한다. 기존 3차원 객체 탐지 연구들은 일반적으로 AP (Average Precision)을 평가 metric으로 사용하지만, 본 연구는 ECCV 워크샵 대회에 제출되면서 기존의 AP 외에도 대회에서 사용된 새로운 metric을 통해 평가를 진행한다. 3차원 공간 상에서 연산되는 AP는 결과 박스의 점수 내지는 불확실성 정보를 통해 예측된 결과의 우선순위를 지정해야만 측정 가능하다는 점과, 연산에 비교적 많은 컴퓨팅 자원을 필요로 한다는 단점이 있어 워크샵 대회 환경에서는 AP를 사용하는 대신, BEV (Bird's-Eye-View)에서 IoU (Intersection over Union)를 기반으로 측정하는 평가 방식을 사용한다. 구체적인 평가 수식은 다음과 같다.

$$Pred\ IoU = \frac{1}{|P|} \sum_{pred \in P} \max_{gt \in G} (IoU(gt, pred)) \quad (2)$$

$$GT\ IoU = \frac{1}{|G|} \sum_{gt \in G} \max_{pred \in P} (IoU(gt, pred)) \quad (3)$$

$$XYIoU = \frac{GT\ IoU + Pred\ IoU}{2} \quad (4)$$

식 (2)과 같이 예측한 모든 박스들의 집합인 P 안의 모든 예측된 박스 pred에 대해 가장 좋은 IoU를 가지는 정답 박스 gt를 정답 박스의 집합인 G에서 찾고, 그 IoU 값을 평균내는 방식으로 계산된다. 식 (3)에서는 같은 과정을 정답 박스를 기준으로 진행하고, 식 (4)에서는 그 두 값을 평균내는 것으로 XYIoU를 계산하여 최종 결과를 평가한다.

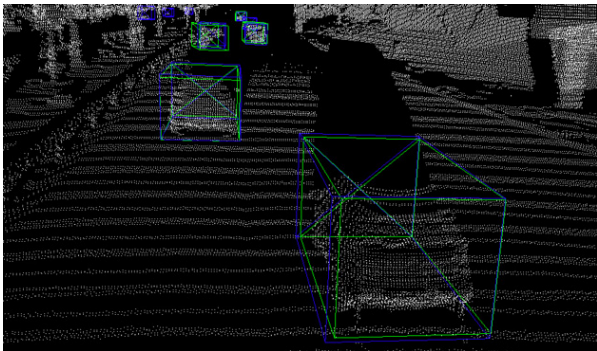
AP (Average Precision) 평가는 Waymo<sup>[1]</sup>에서 사용하는 방법을 따라, 정답 3차원 박스와 예측한 3차원 박스 사이의 IoU가 일정 수준 이상인 경우 정답으로 인정하는 방법을 기반으로 PR (precision recall) 곡선을 그려 측정한다. 차량의 경우 IoU (Intersection over Union) 값이 0.7 이상인 경우 참으로, 차량 외 class의 경우 IoU가 0.5 이상인 경우 참으로 취급한다.

### 3.2 실험 결과

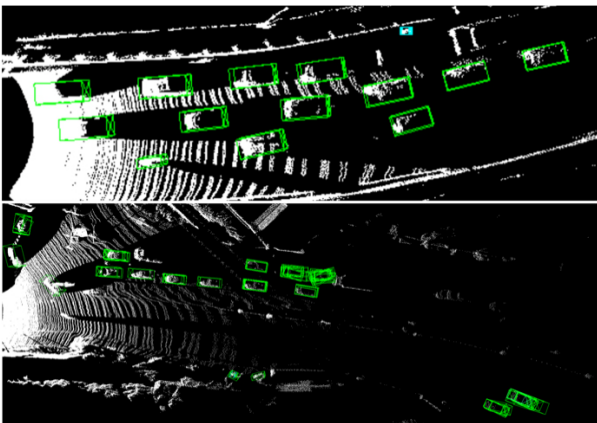
[Fig. 5]와 [Fig. 6]에서 확인할 수 있듯이, 학습 단계에서 본 적 없는 목표 도메인의 데이터셋에서도 3차원 객체 탐지 알고리즘이 작동하는 것을 정성적으로 확인할 수 있다. [Fig. 5]는

학습 레이블이 있는 103개의 프레임들을 학습과 평가 데이터셋으로 4:1의 비율로 나누어 평가 데이터셋에서 평가한 결과이며, [Fig. 6]는 정답 레이블이 없는 테스트 데이터셋에서 평가한 결과이다.

또한 우리는 제안한 방법이 각기 최종 성능에 어떤 영향을 끼치는지 평가하기 위해 랜덤하게 4:1의 비율로 분할한 학습-평가 데이터셋을 통해 정량적으로 검증하였다. 위 [Table 1]은 (a) 아무런 방법 없이 실험할 때, (b) 학생-교사 지식 증류 방법을 사용할 때, (c) 추가로 복사-붙여넣기 데이터 증강 방법을



[Fig. 5] 3D object detection result sample from Innoviz dataset evaluation set. Blue: ground-truth. Green: prediction



[Fig. 6] 3D object detection prediction sample from Innoviz dataset test set. Results are before NMS post-processing

사용했을 때 각각의 검증 결과이다. 우리가 제안하는 방법은 Innoviz에서 주관한 ECCV 2022 workshop 대회에서 제공하는 별도의 공개되지 않은 테스트 세트에서 Z축에 5 cm 간격으로 5가지 offset을 주는 Test-Time-Augmentation 방법까지 사용하여 평균 XYIoU 0.6088의 성능으로 3위를 달성하였다 (d).

정답 레이블이 존재하는 데이터 샘플 자체가 103개로 부족하며, 검증을 위해 이를 분할하였기 때문에, 82개의 샘플에서 학습이 진행되고 21개의 샘플에서 검증이 진행된다. 이때 103개의 샘플 중에서 보행자 또는 사이클리스트가 존재하는 장면은 27개뿐 이므로 [Table 1] 에서 복사-붙여넣기 방법을 사용하여 소량의 class에 대해 충분한 augmentation이 진행되지 전까지는 두 소량의 class에 대해 정답을 맞추는 경우가 나오지 않는 것을 확인할 수 있다.

#### 4. 결론

우리는 본 논문에서 자율 주행 환경에서 사용되는 데이터 도메인의 차이가 문제가 된다는 인식을 가지고 이를 도메인 적응 과정에서 해결할 수 있는 방법을 제안했다. 소수의 타겟 도메인 레이블이 주어지는 3차원 객체 탐지의 준지도 도메인 적응 문제 해결을 위해 딥러닝 기반 방법의 학습에서 도메인 간 물체 크기의 통계치를 활용하였고, 모델의 일부만 효율적으로 추가 학습하는 방식을 사용하기도 하였다. 또한 적은 레이블만을 가지는 목표 도메인에서의 효과적인 적응을 위해 복사-붙여넣기 방식의 자가 학습 방법을 적용하였을 뿐만 아니라 자기지도학습의 방법을 빌려와 지식 증류 방법을 사용하여 일반화 성능 또한 높일 수 있도록 하였다. 하지만 여전히 고속도로 및 도심 주행과 같은 외부 환경적 도메인 차이를 직접적으로 다루는 방법은 적용되지 않아, 향후 추가적인 성능 향상이 기대된다. 또한 지식 증류 방법에서 단순히 두 특징치를 가깝게 하는 방법뿐만 아니라 contrastive loss와 같이 멀어지기도 할 수 있는 손실 함수를 활용하는 방식으로 추후 연구가 가능하다.

[Table 1] Ablation study on Innoviz evaluation dataset. (a): PV-RCNN++ baseline without any other methods. (b): with Teacher-Student Knowledge Distillation. (c): additional copy-and-paste augmentation applied. (d): additional Test-Time-Augmentation applied. (a), (b) and (c) is conducted on evaluation set and (d) is the result on test set. AP cannot be measured because test set data is provided only through the competition server, and evaluation on the server is conducted only by XYIoU

Methods	GT IoU	Pred IoU	XYIoU	AP (Vehicle)	AP (Pedestrian)	AP (Cyclist)	mAP
(a)	0.1927	0.7757	0.4842	0.1713	0.0000	0.0653	0.0789
(b)	0.2374	0.7404	0.4889	0.2627	0.0000	0.0000	0.0876
(c)	0.2910	0.7601	0.5255	0.3148	0.1667	0.2667	0.2494
(d) (test set)	0.5275	0.6901	0.6088	-	-	-	-

## References

- [1] P. Sun, H. Kretzschmar, X. Dotiwalla, A. Chouard, V. Patnaik, P. Tsui, J. Guo, Y. Zhou, Y. Chai, B. Caine, V. Vasudevan, W. Han, J. Ngiam, H. Zhao, A. Timofeev, S. Ettinger, M. Krivokon, A. Gao, A. Joshi, Y. Zhang, J. Shlens, Z. Chen, and D. Anguelov, "Scalability in Perception for Autonomous Driving: Waymo Open Dataset," *Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, USA, 2020, DOI: 10.1109/CVPR42600.2020.00252.
- [2] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *The International Journal of Robotics Research*, vol. 32, no. 11, Aug., 2013, DOI: 10.1177/0278364913491297.
- [3] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, and O. Beijbom, "nuScenes: A Multimodal Dataset for Autonomous Driving," *Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, USA, 2020, DOI: 10.1109/CVPR42600.2020.01164.
- [4] J. Houston, G. Zuidhof, L. Bergamini, Y. Ye, L. Chen, A. Jain, S. Omari, V. Iglovikov, and P. Ondruska, "One Thousand and One Hours: Self-driving Motion Prediction Dataset," *Conference on Robot Learning (CoRL)*, Cambridge, USA, pp. 409-418, 2020, [Online], <https://proceedings.mlr.press/v155/houston21a.html>
- [5] M.-F. Chang, J. Lambert, P. Sangkloy, J. Singh, S. Bak, A. Hartnett, D. Wang, P. Carr, S. Lucey, D. Ramanan, and J. Hays, "Argoverse: 3D Tracking and Forecasting With Rich Maps," *Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, USA, 2019, DOI: 10.1109/CVPR.2019.00895.
- [6] P. Xiao, Z. Shao, S. Hao, Z. Zhang, X. Chai, J. Jiao, Z. Li, J. Wu, K. Sun, K. Jiang, Y. Wang, and D. Yang, "PandaSet: Advanced Sensor Suite Dataset for Autonomous Driving," *International Conference on Intelligent Transportation*, Indianapolis, USA, 2021, DOI: 10.1109/ITSC48978.2021.9565009.
- [7] R. Roriz, J. Cabral, and T. Gomes, "Automotive LiDAR Technology: A Survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6282-6297, Jul., 2022, DOI: 10.1109/TITS.2021.3086804.
- [8] L. Jie, Z. Jin, J. Wang, L. Zhang, and X. Tan, "A SLAM System with Direct Velocity Estimation for Mechanical and Solid-State LiDARs," *Remote Sensing*, vol. 14, no. 7, Apr., 2022, DOI: 10.3390/rs14071741.
- [9] Y. Wang, X. Chen, Y. You, L. E. Li, B. Hariharan, M. Campbell, K. Q. Weinberger, and W.-L. Chao, "Train in Germany, Test in the USA: Making 3D Object Detectors Generalize," *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, USA, 2020, DOI: 10.1109/CVPR42600.2020.01173.
- [10] T. Yin, X. Zhou, and P. Krähenbühl, "Center-Based 3D Object Detection and Tracking," *Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, USA, 2021, DOI: 10.1109/CVPR46437.2021.01161.
- [11] S. Shi, L. Jiang, J. Deng, Z. Wang, C. Guo, J. Shi, X. Wang, and H. Li, "PV-RCNN++: Point-Voxel Feature Set Abstraction with Local Vector Representation for 3D Object Detection," *International Journal of Computer Vision*, vol. 131, pp. 531-551, Nov., 2022, DOI: 10.1007/s11263-022-01710-9.
- [12] 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," *Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, USA, 2020, DOI: 10.1109/CVPR42600.2020.01054.
- [13] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *International Conference on Machine Learning*, Lille, France, pp. 448-456, 2015, DOI: 10.5555/3045118.3045167.
- [14] Y. Yan, Y. Mao, and B. Li, "SECOND: Sparsely Embedded Convolutional Detection," *Sensors*, vol. 18, no. 10, Oct., 2018, DOI: 10.3390/s18103337.
- [15] J.-B. Grill, F. Strub, F. Altché, C. Tallec, P. Richemond, E. Buchatskaya, C. Doersch, B. A. Pires, Z. Guo, M. G. Azar, B. Piot, K. Kavukcuoglu, R. Munos, and M. Valko, "Bootstrap Your Own Latent - A New Approach to Self-Supervised Learning," *Advances in Neural Information Processing Systems*, Red Hook, USA, pp. 21271-21284, 2020, DOI: 10.5555/3495724.3497510.



### 우정완

2019 대구경북과학기술원 기초학부(학사)  
2019~현재 대구경북과학기술원 전기전자  
컴퓨터공학과(석박통합과정)

관심분야: 3D Object Detection, LiDAR Perception



### 김재열

2018 서울과학기술대학교 기계자동차공학과  
(학사)  
2021 서울과학기술대학교 자동차공학과(석사)  
2021~현재 대구경북과학기술원 전기전자  
컴퓨터공학과(박사과정)

관심분야: LiDAR, Domain Adaptation / Generalization



### 임성훈

2014 서강대학교 전기전자공학과(학사)  
2016 한국과학기술원 전기전자공학과(석사)  
2019 한국과학기술원 전기전자공학과(박사)  
2019~현재 대구경북과학기술원 전기전자  
컴퓨터공학과 조교수/부교수

관심분야: Computer Vision, Domain Adaptation