2023 졸업프로젝트 진행 보고서

주제: 3D Object Detection

2023. 04. 05. ~ 2023. 05. 23.

신상윤, 윤성우

진행 사항

* KITTI dataset 해석
* OpenPCDet train.py 코드 분석

KITTI dataset 해석

* **calib**
  + P0, P1, P2, P3 🡨 12개 element == 4x3 matrix

숫자는 카메라 번호를 의미 (총 4개의 카메라)

* + R0\_rect 🡨 9개 element == 3x3 matrix
  + Tr\_velo\_to\_cam 🡨 12개 element == 3x4 matrix

참고자료: <https://darkpgmr.tistory.com/190>



(사진 출처: <https://github.com/bostondiditeam/kitti/blob/master/resources/devkit_object/readme.txt>)

* **label**

element 14개 (type 포함 15개)

(score label이 있다면 15개, type 포함 16개)

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

자동 생성된 설명

(사진 출처: <https://github.com/bostondiditeam/kitti/blob/master/resources/devkit_object/readme.txt>)

dimensions label의 width가 전방축, height가 상향축, length가 좌우축

location label 🡪 카메라에선 x가 오른쪽, y가 위쪽, z가 앞쪽

차량, 육상 차량, 바퀴, 텍스트이(가) 표시된 사진

자동 생성된 설명

(사진 출처: <https://www.cvlibs.net/datasets/kitti/setup.php>)

score label은 없을 수도 있음

* **velodyne**

x, y, z, r 값이 float type으로 4byte씩 연속으로 찍혀있음

참고자료: <https://darkpgmr.tistory.com/190>

train.py 분석

**line 129: build\_network() 분석**

OpenPCDet 구조

텍스트, 스크린샷, 폰트, 라인이(가) 표시된 사진

자동 생성된 설명

OpenPCDet 아키텍처를 이용하여 모델에 부여할 수 있는 옵션

(cfg file 🡨 /tools/cfgs/kitti\_models/\*.yaml)

* VFE – VFETemplate, MeanVFE, PillarVFE, ImageVFE, DynMeanVFE, DynPillarVFE, DynamicPillarVFESimple2D
* BACKBONE\_3D – VoxelBackBoneBx, UNetV2, PointNet2Backbone, PointNet2MSG, VoxelResBackBone8x, VoxelBackBone8xFocal, VoxelResBackBone8xVoxelNeXt, VoxelResBackBone8xVoxelNeXt2D, PillarBackBone8x, PillarRes18BackBone8x
* MAP\_TO\_BEV – HeightCompression, PointPillarScatter, Conv2DCollapse
* PFE – VoxelSetAbstraction
* BACKBONE\_2D – BaseBEVBackbone, BaseBEVBackboneV1
* DENSE\_HEAD – AnchorHeadTemplate, AnchorHeadSingle, PointIntraPartOffsetHead, PointHeadSimple, PointHeadBox, AnchorHeadMulti, CenterHead, VoxelNeXtHead
* POINT\_HEAD – AnchorHeadTemplate, AnchorHeadSingle, PointIntraPartOffsetHead, PointHeadSimple, PointHeadBox, AnchorHeadMulti, CenterHead, VoxelNeXtHead
* ROI\_HEAD – RoIHeadTemplate, PartA2FCHead, PVRCNNHead, SECONDHead, PointRCNNHead, VoxelRCNNHead, MPPNetHead, MPPNetHeadE2E

VoxelNet 모델 구조

텍스트, 스크린샷, 라인, 폰트이(가) 표시된 사진

자동 생성된 설명

총 5단계 (초록색): VFE (MeanVFE) 🡪 BACKBONE\_3D (VoxelBackBone8x) 🡪 MAP\_TO\_BEV (HeightCompression) 🡪 BACKBONE\_2D (BaseBEVBackbone) 🡪 DENSE\_HEAD (AnchorHeadSingle)

질문 /home/student\_01/models/second/pcdet/models/detectors/second\_net.py line: 10

for loop의 batch\_dict는 training 모드에선 무의미한가? 사용된다면 어떻게 사용되는지?

질문 /home/student\_01/models/second/pcdet/models/detectors/second\_net.py line: 27

dense\_head는 어디에서 찾을 수 있는지?

get\_loss()에서 loss 구하려면 y와 y\_hat이 있어야 하는데 어디서 그 값들을 가져오는지?

**line 175: train\_model() 분석**

단순히 “loss 🡪 backward 🡪 update parameter” + “로그 기록” 기능으로 해석함

향후 계획

VoxelNet 모델에서 사용된 각 모듈 (MeanVFE, VoxelBackBone8x, HeightCompression, BaseBEVBackbone, AnchorHeadSingle)에 대한 소스코드를 논문과 비교하며 의미 해석하기

기존 VoxelNet (SECONDNet) 모델을 약간 변경하여 아이디어를 반영한 커스텀 모델을 만들어보고 성능 평가하기