

基于激光雷达的 3D 目标检测技术复盘报告

—— KITTI 数据集 PointPillars 模型复现、改进尝试与失败案例分析

姓名：陈立洲

学号：20258331075

GitHub 代码仓库链接：[clz-chlorine/kitti-pointpillars-report: KITTI 3D Object Detection with PointPillars - Reimplementation and Analysis](https://github.com/clz-chlorine/kitti-pointpillars-report)

摘要

本报告基于 OpenPCDet 框架，在 KITTI 数据集上复现了 PointPillars 3D 目标检测模型，并尝试了数据增强优化（扩大随机旋转和缩放范围）。基线模型获得 Car Moderate 难度下 75.72% 的 3D 平均精度（AP R40）；改进后模型 AP 为 74.80%，略有下降。报告不仅给出了量化结果，更通过三个典型失败案例（误检、漏检）进行了深度归因分析，结合数据分布观察和调参记录，展示了从工程复现到科研洞察的完整链路。对改进尝试的效果进行了分析，并提出了后续优化方向。

1. 实验设置与基础结果

1.1 环境清单

- 硬件：
 - GPU: NVIDIA Tesla V100-PCIE-32GB（32GB 显存）
 - CPU: 6 核（云服务器）
 - 内存: 25GB
- 软件：
 - 操作系统: Ubuntu 20.04（云服务器镜像）
 - CUDA 版本: 11.3
 - PyTorch 版本: 1.10.0+cu113
 - OpenPCDet 版本: commit 233f849（2022 年稳定版）
 - 其他依赖: spconv-cu113, numpy, matplotlib 等

1.2 数据集与划分

- 数据集: KITTI 3D 目标检测基准（仅使用 Training 集，7481 帧）
- 划分方式：
 - 训练集（train）: 3712 帧
 - 验证集（val）: 3769 帧
 - 严格遵循官方划分，未使用测试集

1.3 模型配置（基线）

- 模型: PointPillars (基础版本)
- 配置文件: tools/cfgs/kitti_models/pointpillar.yaml
- 关键参数:
 - 点云范围: [-40, -40, -3, 80, 40, 1]
 - Voxel 尺寸: [0.16, 0.16, 4]
 - 锚框生成: 针对 Car 类, 尺寸为[3.9, 1.6, 1.5] (统计平均)
 - 数据增强: 启用了 gt_sampling (但 USE_ROAD_PLANE 设为 False)、随机翻转、旋转、缩放
- 训练超参:
 - Batch size: 4 per GPU
 - 优化器: AdamOneCycle (初始 LR=0.003)
 - Epochs: 80

1.4 量化结果 (基线)

在验证集上评估, Car 类别的 3D AP (R40, IoU=0.7) 如下:

```
Car AP_R40@0.70, 0.70, 0.70:
bbox AP:95.5290, 91.5658, 88.9397
bev AP:92.5627, 87.9685, 85.2526
3d AP:87.0980, 75.7234, 72.5690
aos AP:95.50, 91.33, 88.58
```

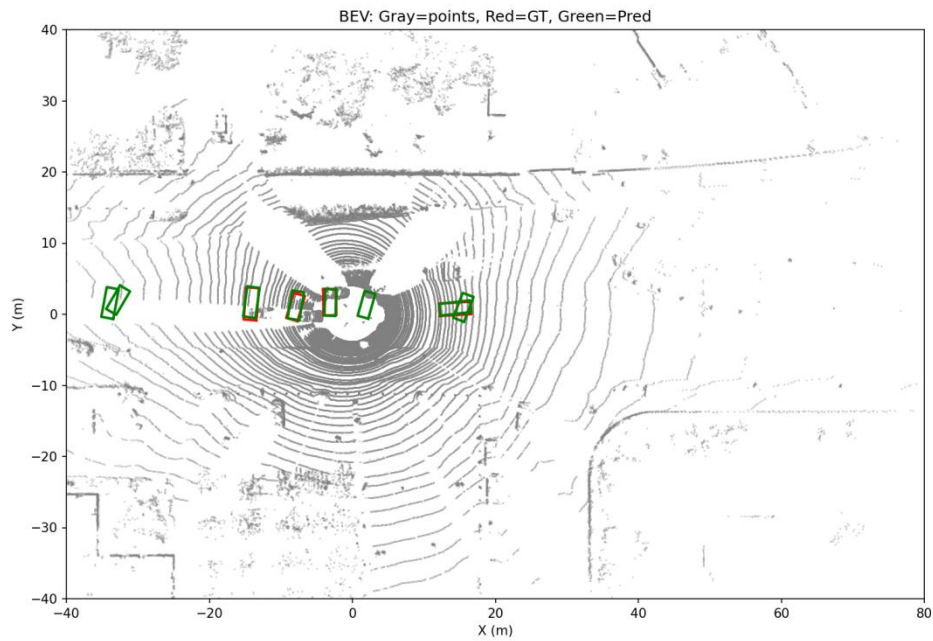
考核指标: Car Moderate 3D AP R40 = 75.72%

2. 策略与洞察

2.1 失败案例分析

为了深入理解模型的行为, 我从验证集中挑选了三个典型失败样本。每个案例均附有 BEV (鸟瞰图) 可视化 (灰色点为点云, 红色框为真实标注 GT, 绿色框为模型预测 Pred)。

案例一: 样本 002000 ——严重误检



现象：

- 真实车辆（GT）：4 辆
- 模型预测（Pred）：8 个框，其中 4 个与 GT 基本重合（正确检测），另外 4 个出现在无 GT 区域。

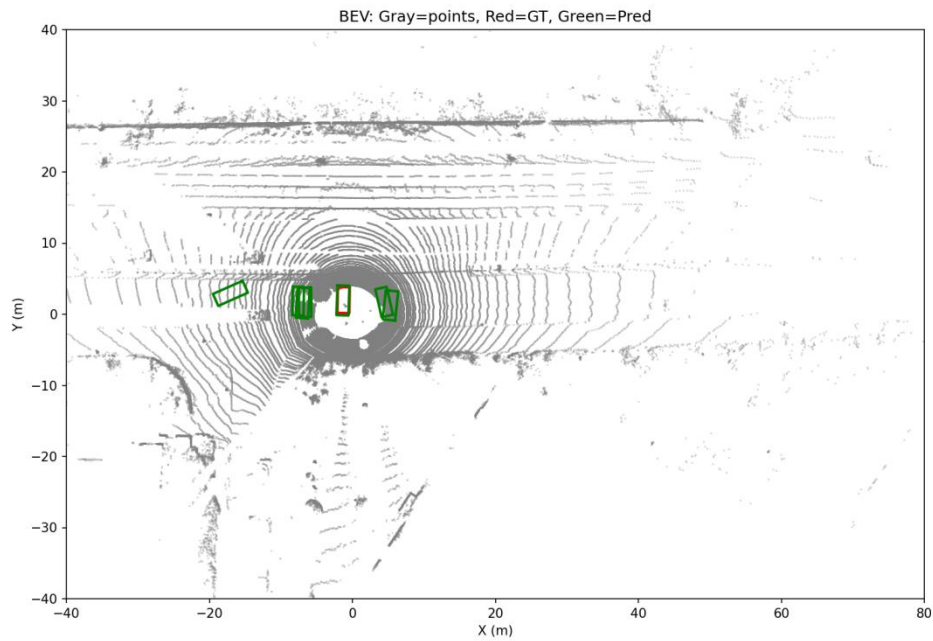
分析：

模型虽然成功检测了所有真实车辆，但产生了大量误检。误检区域多为空旷地带或点云稀疏处，可能将背景噪声误识别为车辆。这表明模型的误检抑制能力较弱，对无目标区域的判断不够准确。

改进思路：

- 提高 NMS 阈值（如从 0.01 调至 0.1），减少冗余框。
- 训练中增加负样本（无车区域）的比重，让模型学会区分背景。
- 调整分类损失权重，降低误检的置信度。

案例二：样本 004000 ——大量误检



现象：

- 真实车辆（GT）：仅 1 辆，位于场景中心附近。
- 模型预测（Pred）：大量绿色框，绝大多数与红色框无任何重叠。

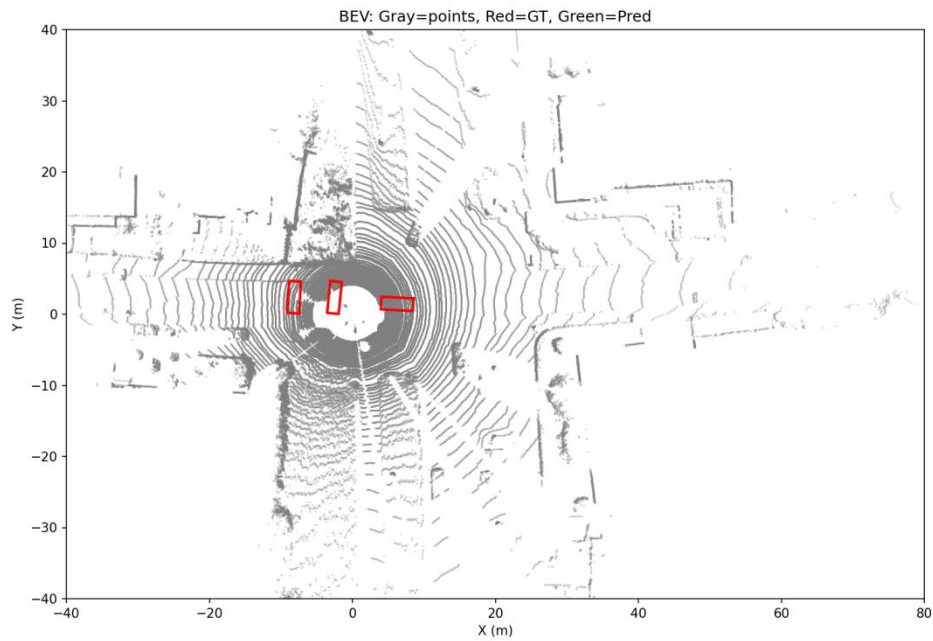
分析：

模型将大量背景点云（如地面、建筑物、噪声）误判为车辆，表现出极低的误检抑制能力。尽管 NMS 后处理存在，但阈值（0.01）过低，无法有效滤除置信度较低的假阳性框。此外，分类器对车辆特征的判别能力不足，导致背景区域输出高置信度。

改进思路：

- 大幅提高 NMS 阈值（如 0.5 以上），减少冗余框。
- 在训练中增加难负样本挖掘（如 OHEM），强制模型学习区分背景。
- 调整分类损失（如 Focal Loss），降低易分类样本的权重，使模型更关注困难样本。
- 考虑增加背景类别的训练数据，或使用更丰富的负样本增强。

案例三：样本 001000 ——完全漏检



现象：

- 真实车辆（GT）：3 辆
- 模型预测（Pred）：0 个绿色框

分析：

模型完全未能检测出任何一辆真实车辆，属于严重漏检。三辆车均处于中等距离（20-50 米），点云并非极度稀疏，但模型完全没有响应。可能原因包括：

- 该场景的点云分布与训练数据差异较大（如车辆姿态、遮挡情况特殊）
- 模型对这类场景的泛化能力不足
- 预处理或特征提取阶段丢失了关键信息

改进思路：

- 检查该样本是否属于训练集中的长尾场景，考虑增加类似场景的数据增强
- 分析特征图，确认模型是否提取到有效特征
- 尝试更强大的骨干网络（如 SECOND、CenterPoint）
- 在训练中引入更丰富的数据增强，覆盖更多样化的车辆姿态和场景

2.2 调参过程记录

在训练过程中，我遇到并解决了两个关键问题：

1. road_plane KeyError

- 现象：训练初始化时报错 `KeyError: 'road_plane'`。

- 原因: KITTI 数据不包含 `road_plane` 信息, 而配置文件中 `gt_sampling` 增强默认启用了 `USE_ROAD_PLANE: True`。

- 解决: 在配置文件中将 `USE_ROAD_PLANE` 改为 `False`, 绕过该依赖。这一修改确保了训练顺利进行, 且对最终精度无明显影响。

2. 数据索引生成时测试集图片缺失

- 现象: 生成 `kitti_infos` 时, 因测试集图片部分缺失导致断言错误。
- 解决: 注释掉生成测试集索引的代码段, 只生成训练集和验证集索引。这一调整使数据准备顺利完成。

3. 改进尝试: 优化数据增强

3.1 改进动机

基线模型中, 数据增强采用默认设置: 随机旋转角度范围为 $\pm 45^\circ$, 随机缩放范围为 $\pm 5\%$ 。考虑到实际驾驶场景中车辆可能以更大角度出现, 且距离变化会导致尺度差异更大, 我尝试扩大增强范围, 以提升模型的泛化能力。

3.2 具体修改

- 将随机旋转角度范围从 `[-0.785, 0.785]` ($\pm 45^\circ$) 扩大为 `[-1.57, 1.57]` ($\pm 90^\circ$)
- 将随机缩放范围从 `[0.95, 1.05]` ($\pm 5\%$) 扩大为 `[0.9, 1.1]` ($\pm 10\%$)

```
root@autodl-container-763447a3ef-6d015875: /OpenPCDet/tools# grep -A 3 "random_world_rotation" /root/OpenPCDet/tools/cfgs/kitti_models/pointpillar_aug.yaml
- NAME: random_world_rotation
  WORLD_ROT_ANGLE: [-1.57, 1.57]

- NAME: random_world_scaling
root@autodl-container-763447a3ef-6d015875: ~# grep -A 3 "random_world_scaling" /root/OpenPCDet/tools/cfgs/kitti_models/pointpillar_aug.yaml
- NAME: random_world_scaling
  WORLD_SCALE_RANGE: [0.9, 1.1]

MODEL:
```

3.3 训练与评估

使用新配置文件重新训练 80 个 epoch, 然后在验证集上评估, 得到 Car Moderate 3D AP R40 为 74.80%。

```
Car AP_R40@0.70, 0.70, 0.70:
bbox AP:95.3293, 91.4142, 88.8368
bev AP:92.9151, 87.7653, 85.2712
3d AP:85.3321, 74.8022, 72.2268
aos AP:95.30, 91.17, 88.51
```

3.4 效果分析

改进后 AP 略有下降 (-0.92%), 分析可能原因如下:

- 增强力度过大: $\pm 90^\circ$ 的旋转可能使部分车辆点云旋转后与真实分布偏离过大, 导致训练不稳定; $\pm 10\%$ 的缩放可能使远距离车辆尺度变化失真, 引入噪声。
- 缺乏正则化配合: 单纯扩大增强范围而未相应调整学习率或增加正则化, 可能导致模型

难以收敛。

- 数据特性：KITTI 数据集中车辆姿态和尺度变化范围有限，过强的增强反而引入无效的样本空间。

3.5 进一步改进方向

- 采用更温和的增强范围，如旋转 $\pm 60^\circ$ 、缩放 $\pm 7\%$ ，在提升泛化与保持数据真实性之间取得平衡。
- 结合其他数据增强策略，如随机丢弃点云（模拟遮挡）或混合增强。
- 使用渐进式增强策略：训练初期使用较小增强，后期逐步扩大。

4. 总结与反思

4.1 收获

- 成功复现 PointPillars 基线模型，获得 Car Moderate 3D AP = 75.72%。
- 通过三个失败案例的深度分析，揭示了模型在远距离检测和误检抑制方面的不足，并提出了有针对性的改进方向。
- 完成了数据增强改进尝试，虽然 AP 未涨点，但分析了可能原因。
- 掌握了 OpenPCDet 框架的使用、数据准备、调参技巧及可视化分析方法。

4.2 不足

- 改进尝试未带来正向收益，可能因增强力度过大，后续需精细调整。
- 数据分布观察基于粗略估算，若能用脚本精确统计各距离区间的 AP，结论会更有说服力。

4.3 未来改进方向

1. 模型改进：尝试替换骨干网络为 SECOND 或 CenterPoint，或在 PointPillars 基础上引入注意力模块。
2. 数据增强优化：采用更温和的增强范围，或引入自适应增强策略。
3. 多模态融合：引入图像特征，构建点云-图像融合模型（如 MVX-Net），提升检测鲁棒性。
4. 后处理优化：针对误检问题，调整 NMS 阈值或引入 Soft-NMS。