

# 3D点云目标检测算法汇总\_qq\_29462849的博客-CSDN博客

作者：Tom Hardy

Date：2020-2-12

来源：[汇总|3D点云目标检测算法](#)

## 前言

前面总结了几种基于激光雷达点云数据的3D目标检测算法，还有一些算法不再单独列出，这里做个简单总结来分享下！

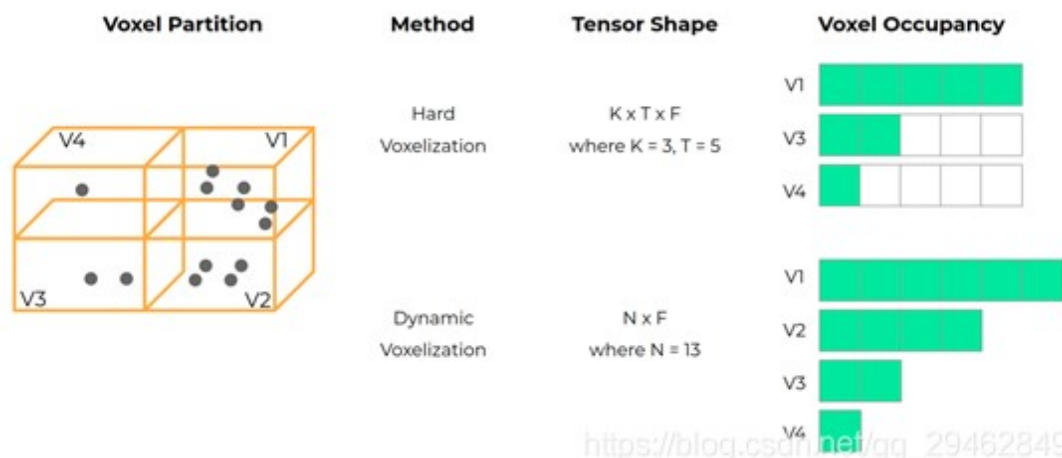
## 基于激光雷达点云的3D目标检测算法

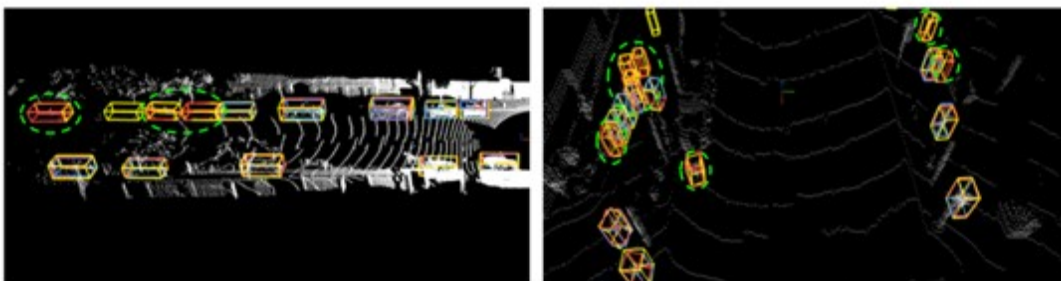
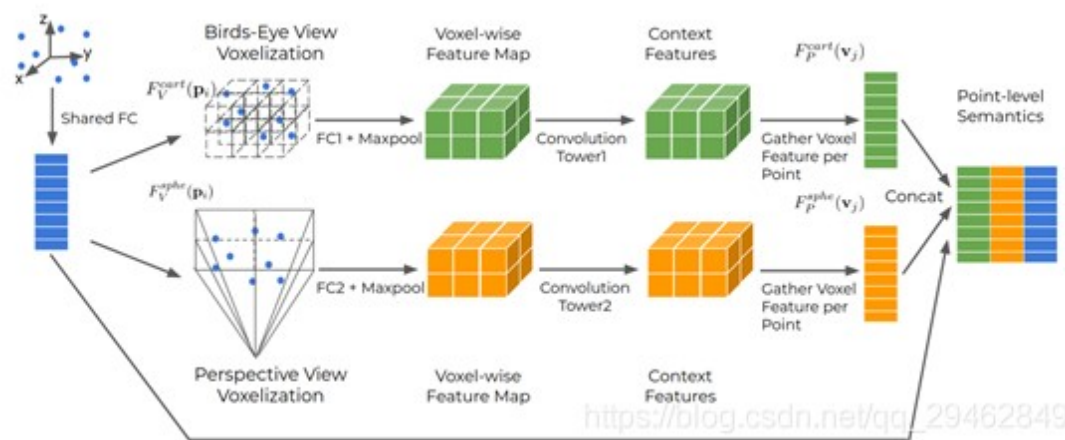
### 1、End-to-End Multi-View Fusion for 3D Object Detection in Lidar Point Clouds (Waymo和Google联合提出)

主要提出了一种新的端到端多视图融合（MVF）算法，该算法能有效地学习利用透视图和点云信息。具体地说，论文介绍了动态体素化，它与现有的体素化方法相比有四个优点：

- 1、消除了预先分配具有固定大小的张量需要；
- 2、克服了由于随机点/体素丢失引起的信息损失；
- 3、产生确定的体素嵌入和更稳定的检测结果；
- 4、建立点和体素之间的双向关系，这为跨视点特征融合奠定了基础；

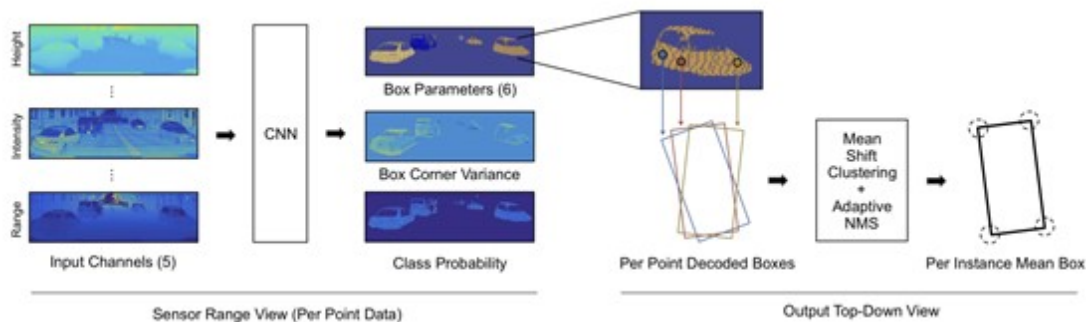
通过采用动态体素化，提出的特征融合体系结构可以使每个点学习融合来自不同视图的信息。MVF对点进行操作，可以自然地激光雷达点云扩展到其他方法。在最新发布的Waymo开放数据集和KITTI数据集上广泛评估了MVF模型，并证明它比可比较的单视图点柱baseline显著提高了检测精度。

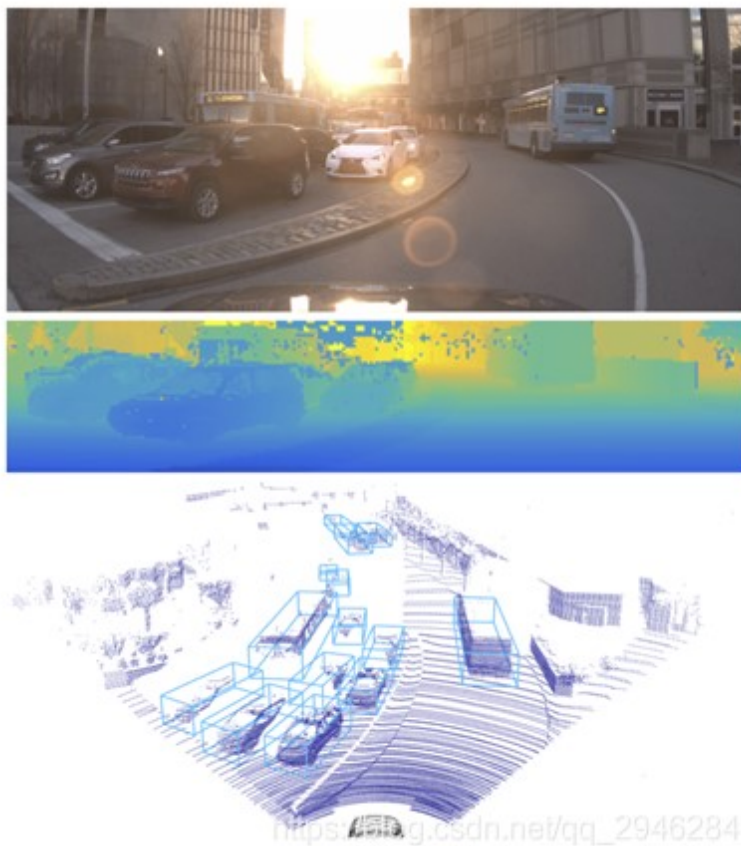




## 2、LaserNet: An Efficient Probabilistic 3D Object Detector for Autonomous Driving (Uber提出, CVPR2019)

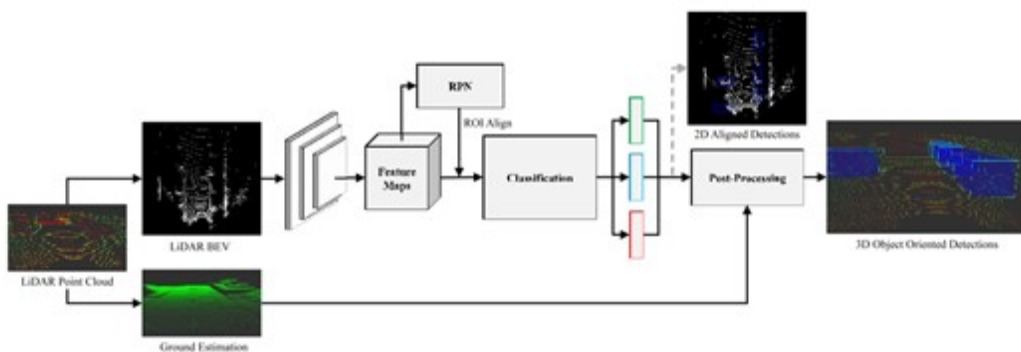
本文提出了一种基于激光雷达数据的自动驾驶三维目标检测算法LaserNet。这项工作提出了一种有效的方法来学习一个端到端的概率三维目标检测器。当有足够的训练数据时，通过使用一个小而密集的范围图像，而不是一个大而稀疏的鸟瞰图像，可以在显著降低运行时间的情况下获得最新的检测性能。该方法不仅为每个检测产生一个类概率，而且在检测边界盒上产生一个概率分布。本文提出的方法是第一个通过模拟包围盒角点的分布来捕捉检测的不确定性。通过估计检测的精度，该方法可以使全自动驾驶系统中的下游部件在具有不同不确定性的物体周围表现出不同的行为。





### 3、BirdNet: a 3D Object Detection Framework from LiDAR information

本文针对3D检测任务，提出了一种面向激光雷达数据的新的网络框架BirdNet。首先，将激光雷达数据投影到一种新的用于鸟瞰投影的cell编码中。然后，通过一个最初设计用于图像处理的卷积神经网络来估计目标在平面上的位置和方向。最后，在后处理阶段计算面向3D的检测任务。



本文描述了一种优化的单级深层卷积神经网络LMNet，它只使用点云数据来检测城市环境中的目标。此功能使该方法能够在一天中的任何时间和照明条件下工作。提出的网络结构采用扩展卷积，随着深度的增加，感知场逐渐增大，计算时间减少约30%。网络输入包括无组织点云数据的五种透视表示，网络为每个点输出对象贴图和边界框偏移值。实验表明，使用反射、范围和三个轴上的每个轴上的位置有助于改善输出边界框的位置和方向。在KITTI数据集评估服务器的帮助下进行了定量评估，获得了最快的处理速度，使其适合实时应用。本文在一辆装有Velodyne HDL-64的实车上实现并测试了网络。基于桌面GPU实现了高达50 FPS的执行速度，在单个Intel Core i5 CPU上实现了高达10 FPS的执行速度。



和实时性。论文在两个数据集上验证了PIXOR: KITTI-BEV目标检测数据集和large-scale 3D车辆检测基准。两个数据集表明, 所提出的检测器在平均精度 (AP) 方面明显优于其他最先进的方法, 而且速度上仍高于28fps。

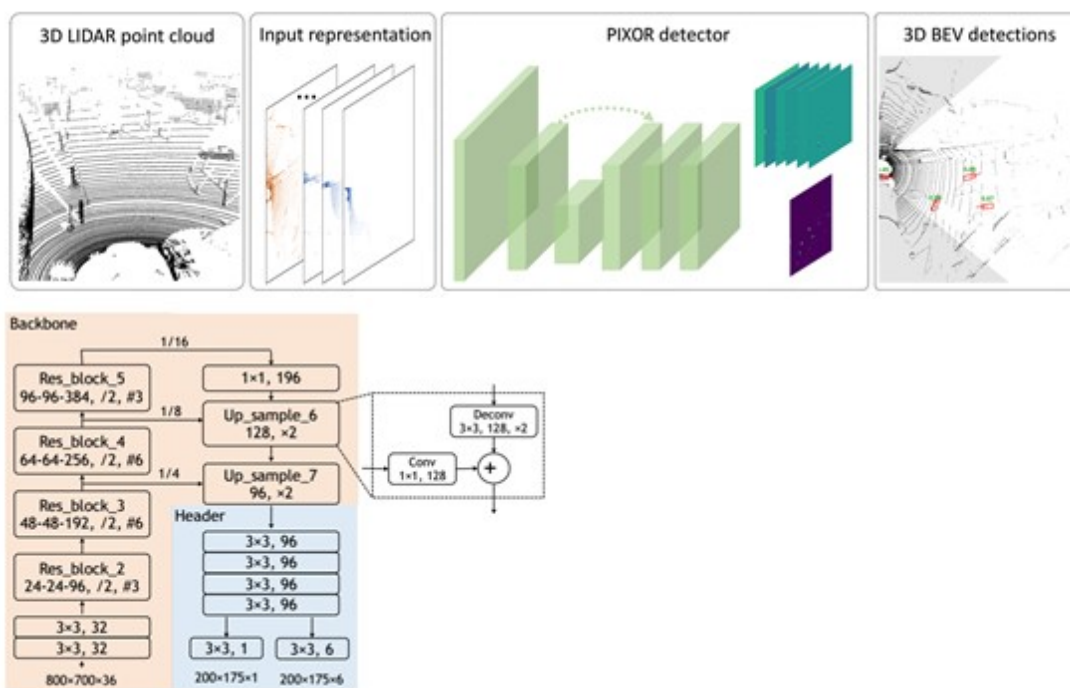
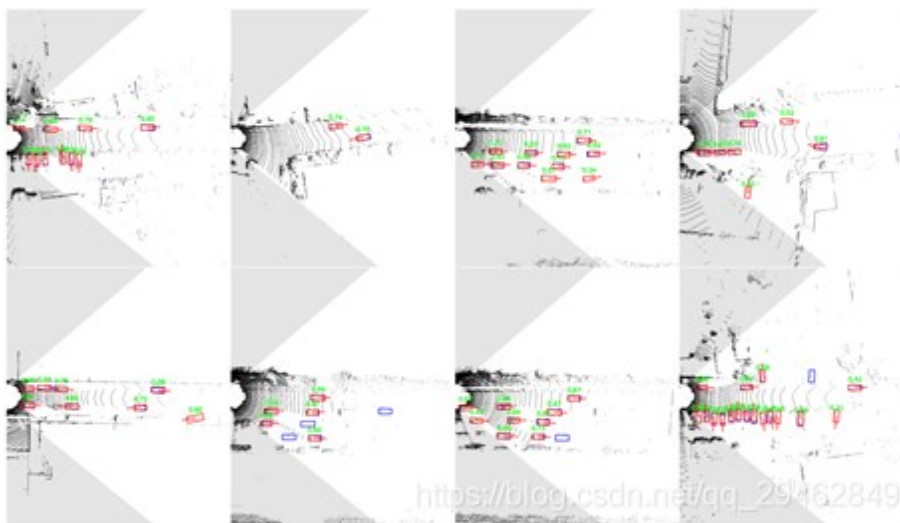


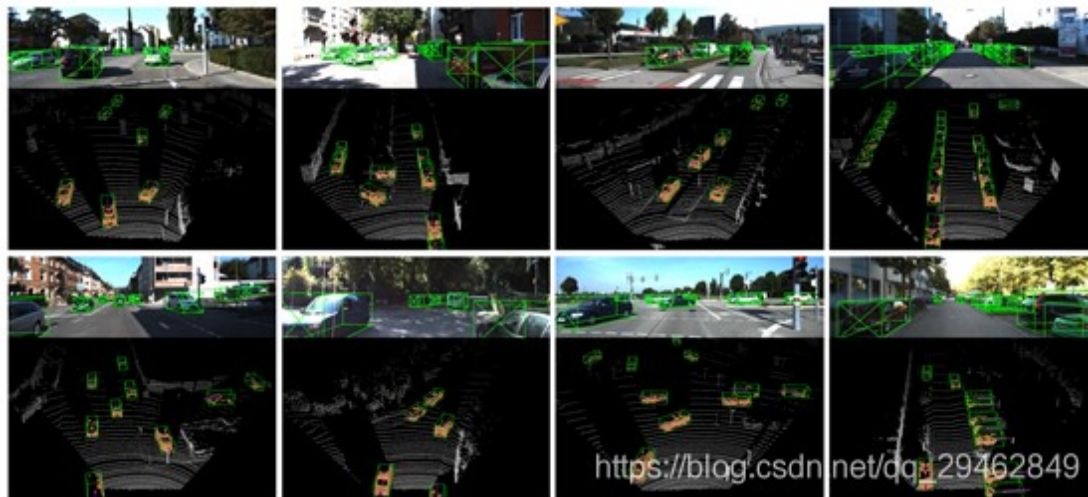
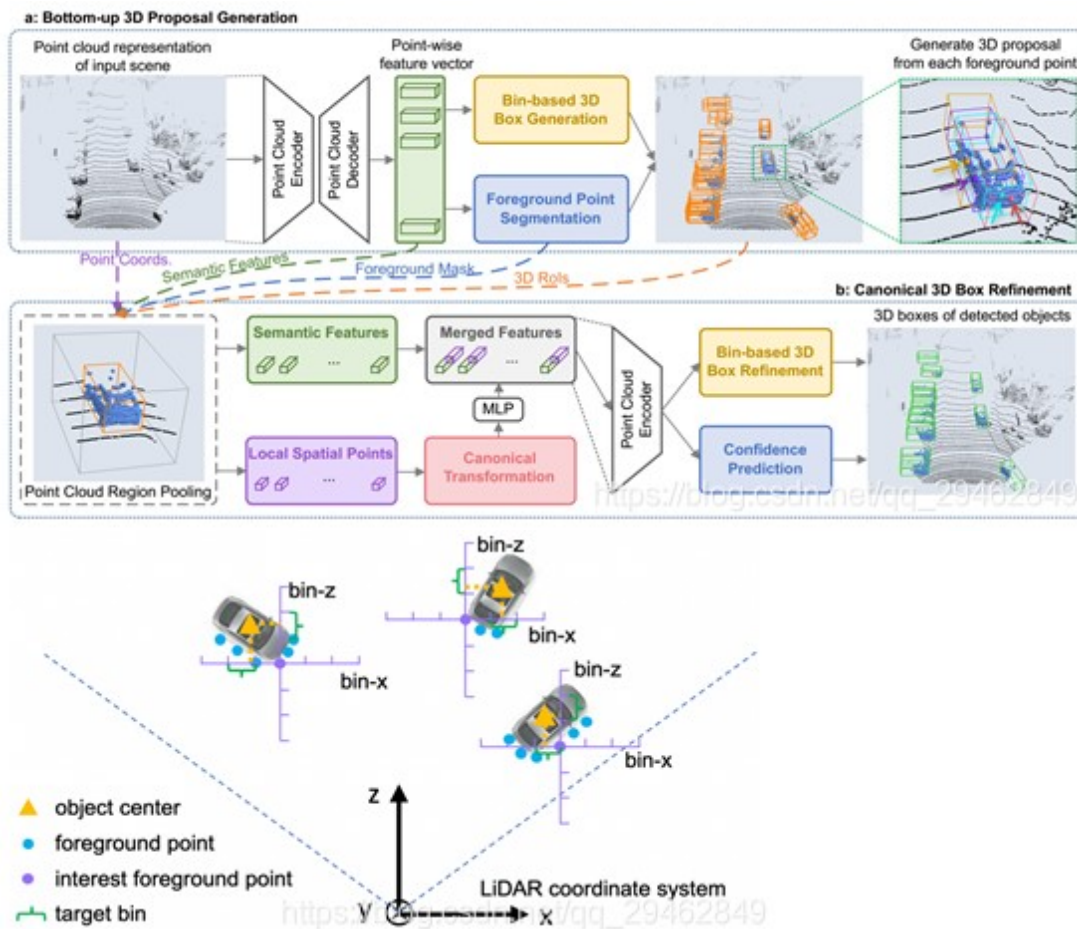
Figure 2. The network architecture of PIXOR.



## 6、PointRCNN: 3D Object Proposal Generation and Detection from Point Cloud (香港大学提出, CVPR2019)

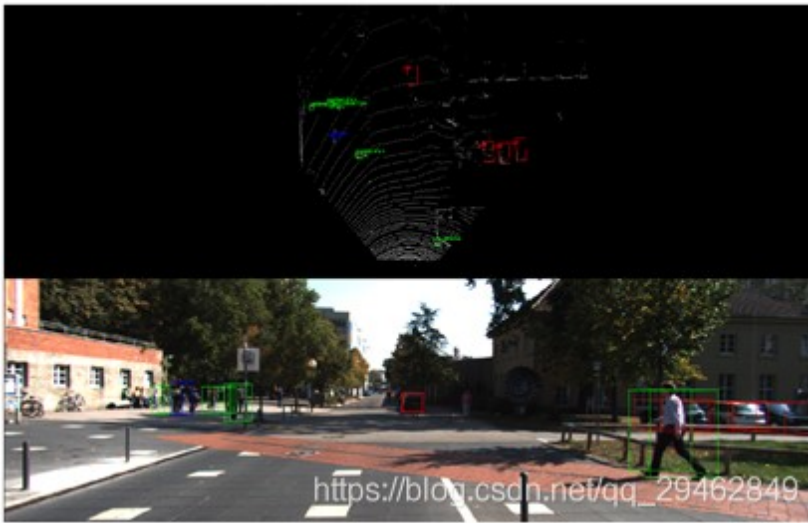
本文提出了一种基于点云的三维目标检测方法。整个框架由两个阶段组成: 第一阶段用于自下而上的3D region proposal, 第二阶段用于在标准坐标系中细化proposal以获得最终的检测结果。第一阶段子网络没有像以前的方法那样从RGB图像或投影点云到鸟瞰图或体素中生成建议, 而是通过将整个场景的点云分割成前景点和背景, 以自下而上的方式直接从点云生成少量高质量的3D建议。第二阶段子网络将每个方案的集合点转换为规范坐标, 以学习更好的局部空间特征, 并结合第一阶段学习到的每个点的全局语义特征, 进行精确的box细化和置信度预测。在KITTI数据集的三维检测基准上进行的大量实验表明, 本文提出的架构仅使用点云作为输入, 其性能优于最新

的方法，并且具有显著的边缘。



## 7、YOLO3D: End-to-end real-time 3D Oriented Object Bounding Box Detection from LiDAR Point Cloud

基于YOLOv2进行改进的文章，主要将3D点云在二维透视图像空间进行映射，并在在二维透视图像空间的一次回归元结构成功的基础上，对其进行扩展，从LiDAR点云生成面向三维对象的bounding box。本文的主要贡献是将YOLO v2的损失函数扩展为包含偏航角、笛卡尔坐标系中的3D box中心和box高度的直接回归问题。文章创新度不大，但是效果和速度值得借鉴：Titan X GPU上实现了实时性能（40 fps）。



$$\begin{aligned}
 L = & \lambda_{coord} \sum_{i=0}^{s^2} \sum_{j=0}^B L_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (z_i - \hat{z}_i)^2] \\
 & + \lambda_{coord} \sum_{i=0}^{s^2} \sum_{j=0}^B L_{ij}^{obj} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{l_i} - \sqrt{\hat{l}_i})^2 \\
 & + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] \\
 & + \lambda_{yaw} \sum_{i=0}^{s^2} \sum_{j=0}^B L_{ij}^{obj} (\phi_i - \hat{\phi}_i)^2 \\
 & + \lambda_{conf} \sum_{i=0}^{s^2} \sum_{j=0}^B L_{ij}^{obj} (C_i - \hat{C}_i)^2 \\
 & + \lambda_{conf} \sum_{i=0}^{s^2} \sum_{j=0}^B L_{ij}^{noobj} (C_i - \hat{C}_i)^2 \\
 & + \lambda_{classes} \sum_{i=0}^{s^2} \sum_{j=0}^B L_{ij}^{obj} \sum_{c \in classes} (p_i(c) - \hat{p}_i(c))^2
 \end{aligned}$$

## 8、FVNet: 3D Front-View Proposal Generation for Real-Time Object Detection from Point Clouds (上海交大&腾讯优图)

与二维点云相比，原始点云和稀疏点云的三维目标检测得到的处理要少得多。本文提出了一种新的基于点云的三维前视图生成和目标检测框架FVNet。它包括两个阶段：前视图proposal的生成和三维bounding box参数的估计。本文没有从相机图像或鸟瞰图中生成proposal，而是先将点云投影到柱面上，生成保留丰富信息的前视特征图。然后引入一个region proposal网络，从生成的地图中预测三维区域proposal，并进一步从整个点云中提取出感兴趣的对象。最后，论文提出了另一个网络，从提取出的目标点中提取点特征，并在标准坐标系下回归最终的三维bounding box参数。FVNet以每个点云样本12毫秒的速度实现实时性能。在三维检测基准KITTI上进行的大量实验表明，所提出的结构在精度上和速度上优于以相机图像或点云为输入的最新技术。



