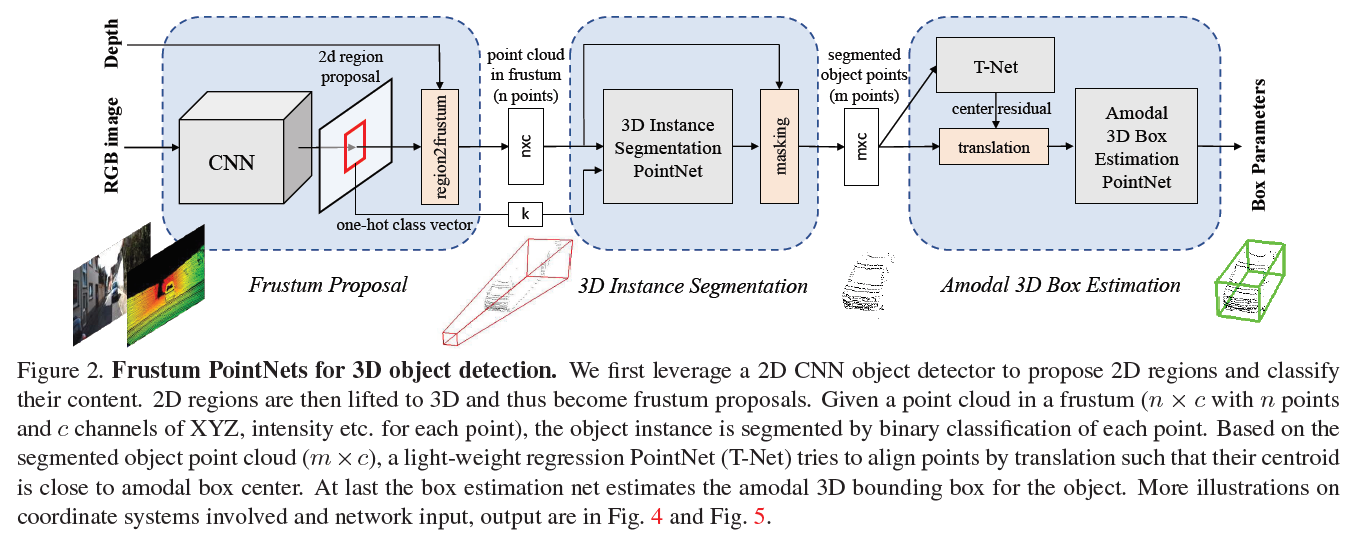
Frustum pointnets for 3d object detection from rgb-d data

3D object detection from RGB-D Data

* Front view image based methods
* Bird’s eye view based methods
* 3D based methods

问题：目前单纯基于Lidar数据的3D目标检测算法通常对小目标检测效果不佳

F-PointNet 结合基于图像的2D检测来定位目标，再用其对应的点云数据视锥进行bbox回归以实现3D目标检测



Frustum proposal 视锥生成

首先通过2D目标检测器来定位图片中的目标以及判断它们的类别。对每一个检测到的目标，通过标定好的传感器的内参和它们之间的转换矩阵得到其对应的点云数据中的各点，即点云视锥。作者使用的2D目标检测模型是基于VGG网络的FPN作为特征提取器，并用Fast R-CNN来预测最终的2D bbox

3D instance segmentation 3D实例分割

对每个得到的点云视锥，通过旋转得到以中心视角为坐标轴的点云数据。对转换后的点云数据用PointNet（或PointNet++）进行实例分割。实例分割是一个二分类问题，用于判断每个点属于某个目标或者不属于。

3D box estimation 3D边界框回归

将上一步实例分割的结果作为mask得到属于某个实例的所有点云，计算其质心作为新的坐标系原点。通过一个T-Net进行回归得到目标质心和当前坐标原点的残差。将点云平移到计算得到的目标质心后，通过PointNet（或PointNet++）对3D bbox的中心、尺寸和朝向进行回归得到最终的输出。此步骤采用的回归方式和Faster R-CNN中类似，不直接回归，而是回归到不同尺寸和朝向的锚点（anchors）。