摘要：

卷积网络技术最近在基于视觉的检测任务中取得了巨大的成功。 本文介绍了将全卷积网络技术移植到三维距离扫描数据检测任务上的研究进展。 具体而言，这个情景是从Vodyodyne 64E激光雷达的数据中获得的。 我们建议将数据呈现在二维点图上，并使用一个单一的2D端到端完全卷积网络来同时预测对象的一致性和绑定。 通过仔细设计边界框编码，即使使用二维卷积网络，也能够预测完整的三维边界框。 KITTI数据集上的实验显示了所提出方法的最新性能。

多年来机器人研究的发展，三维激光雷达已被广泛应用于各种机器人平台。典型的3D激光雷达数据通过在距离扫描中组织的3D点云呈现环境信息。在机器人任务中利用范围扫描数据进行了大量研究，包括定位，映射，对象检测和场景解析[16]。在物体检测任务中，范围扫描在定位检测到的物体时比照相机图像具有特定的优势。由于范围扫描本质上包含三维点云的空间坐标，所以更容易获得检测到的物体的姿态和形状。在包括感知和控制模块的机器人系统上，例如准确定位障碍物车辆的三维坐标对于后续的计划和控制阶段是至关重要的。在本文中，我们设计了一个完全卷积网络（FCN）来检测和定位物体作为距离扫描数据的3D盒子。 FCN在基于计算机视觉的检测任务中取得了显着的成绩。本文将FCN移植到3D距离扫描的检测任务上。我们严格的使用Velodyne 64E激光雷达作为自动驾驶系统的3D车辆检测场景。该方法可以概括为其他类似激光雷达设备上的其他物体检测任务。

介绍：

数据准备我们考虑由Velodyne 64E激光雷达捕获的点云。 像其他距离扫描数据一样，Velodyne扫描的点可以粗略投影，并使用以下投影函数离散成2D点图。

其中p =（x，y，z）>表示3D点，（r，c）表示其投影的2D地图位置。 θ和φ表示观察点时的方位角和仰角。 Δθ和Δφ分别是连续射束发射器之间的平均水平和垂直角分辨率。 投影点图像类似于圆柱形图像。 我们使用2通道数据（d，z）填充2D点图中（r，c）的元素，其中d = px2 + y2。 请注意，x和y在z周围以d为旋转不变性。 图1a显示了2D点图的d通道的一个例子。 很少有一些点可能被投影到同一个2D位置，在这种情况下，靠近观察者的点被保留下来。 （d，z）=（0,0），填充没有3D点的2D位置的元素。

网络体系结构CNN体系结构的主干部分类似于Huang等人。 [11]，龙等人。[18]。 如图2所示，CNN特征图在前3个卷积层中连续下采样，并在去卷积层中连续上采样。 然后主干在第四层分裂为对象分类分支和三维边界框回归分支。 我们描述其实施细节如下：

在所提出的方法的不同阶段生成数据可视化。 （a）输入点图，d通道可视化。 （b）oa p处的对象分支的输出置信度图。 红色代表更高的信念。 （c）对应于预测为正的所有点的边界框候选，即（b）中的高置信度点。 （d）非最大抑制后的剩余边界框。 红点是车辆上的重点，供参考。

建议的FCN结构同时预测车辆的物体和边界框。 conv1 / deconv5a，conv1 / deconv5b和conv2 / deconv4的输出特征映射首先连接，然后分别移植到它们的连续层。

（a）（3）的说明。 对于每个车辆点p，我们定义一个以p为中心的特定坐标系。 坐标系的x轴（rx）与从Velodyne原点到p（虚线）的射线一起。 （b）观察车辆时的旋转不变性的示例。 车辆A和车辆B具有相同的外观。 有关详细信息，请参阅第III-C部分的（3）。