Real-Time Road Segmentation Using LiDAR Data Processing on an FPGA

Abstract:

作者提出一个CNN模型利用Lidar数据做语义分割。更重要的是作者采用FPGA，能在16.9ms内处理一次Lidar扫描所得的数据。在KITTI数据集上取得了较高的准确度。

Introduction：

本文主要任务是道路语义分割，这是可通行区域检测的基础。通过FPGA运行Lidar based road segmentation算法来获得实时低功耗的嵌入式系统。本文将道路分割问题构建为使用深度神经网络的球形图像中的语义分割任务，来代替通常在传统神经网络中实现的编码器 - 解码器结构。

Algorithms design:

本文中，LiDAR数据是input，利用深度神经网络输出top-view predictions来做road segmentation。作者还提供了相机视图上的结果以获得更好的可视化。 该算法具有以下三个步骤：预处理，神经网络处理和后处理。

1. Pre-processing:

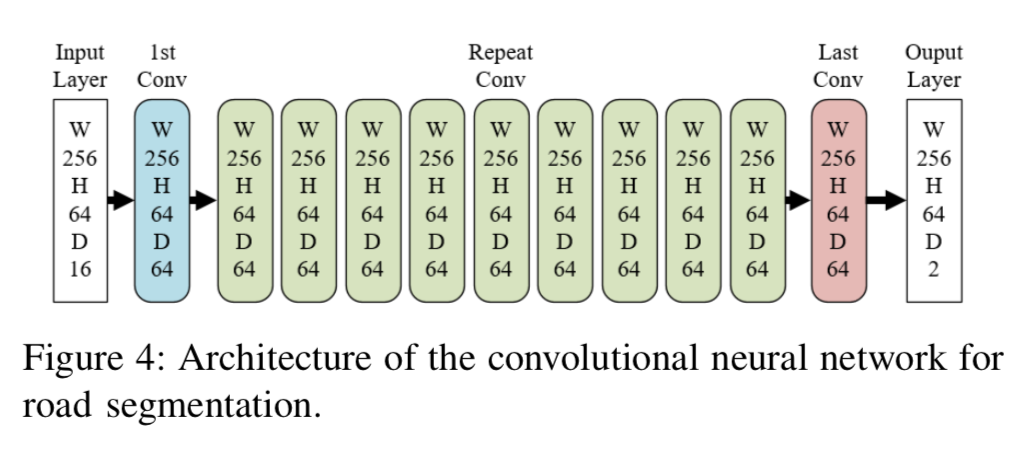
在预处理期间，输入数据点被排列并投影到具有M×N张量和C通道的3-D斑点中，使得张量可以流过神经网络中的层以产生输出。

自动驾驶任务有四种类型的视图：图像视图（也称为摄像机视图），俯视图（也称为鸟瞰图），圆柱视图和球面视图。 图像视图和顶视图通常是选择，因为在这两个视图中，LiDAR数据可以与相机数据融合，并且这些视图对于人眼来说是自然的。 但是，这些视图中的LiDAR点数很稀疏。

圆柱视图和球面视图与LiDAR传感器匹配，数据点可以覆盖映射图上高达91％的像素。因此，作者选择球面视图作为投影方案。 基于LiDAR分辨率选择极角θ和方位角φ的分辨率。在水平方向上，LiDAR点按0.4°分组，使LiDAR的设计分辨率加倍，使没有LiDAR测量的单元数减少。虽然选择球面视图用于数据投影，但我们仍然可以从其他视图添加额外的特征通道，以提高训练的神经网络的准确性。 这里我们选择16个通道，前7个通道来自LiDAR点，其中单元格中的海拔高度最低，接下来的7个通道来自具有最高海拔高度的LiDAR点。 7个通道是笛卡尔坐标（x，y，z）中测量点的位置，球面坐标（θ，φ，r）中测量点的位置，以及测量点（H）的反射强度。 其他2个通道是2D地图（i，j）上的单元格的位置。

1. Neural network processing:

作者提出一个基于FPGA的新的网络结构：



The convolutional layer is built with 64 filters and each filter has a 5×5 kernel with stride size of 1 and padding size of 2.

激活函数是RLU；

6th和10thblock后加了drop-out层

1. posting-processing:

后处理中，将从深度网络中获得的结果重新投回到targeted views中，例如camera view或者top view中。后处理的挑战在于投影后神经网络输出中的点在目标视图上非均匀分布。在本文的后处理步骤中，首先确定可驱动区域的轮廓，然后将轮廓内的区域标记为目标视图上的分割结果。

为了确定可通行区域的轮廓，选择每个角度θ中对应于神经网络输出的每列的最远点，并将其投影到目标视图上。然后，沿目标视图上所有角度的最远点绘制折线。 如果我们在底部添加一条直线，折线图将变为多边形。 然后将多边形视为可通行区域的轮廓并填充语义像素标签。