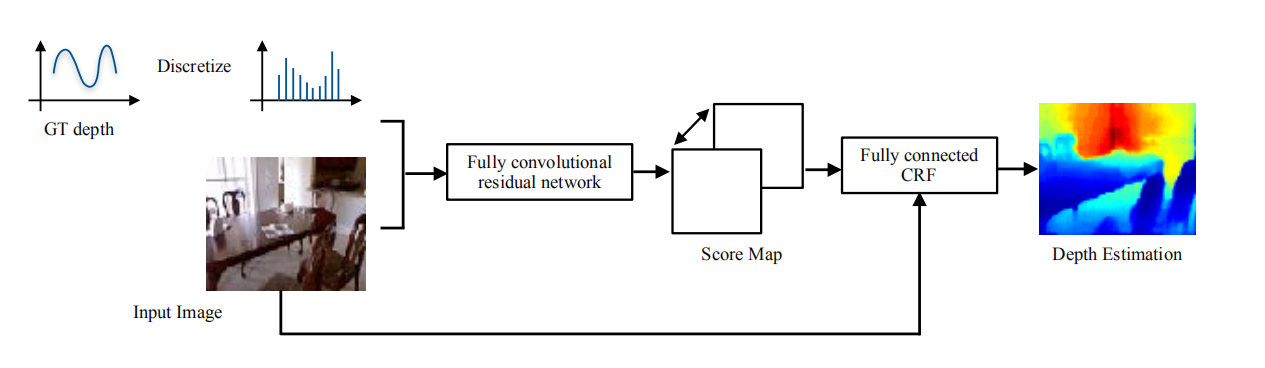
论文阅读笔记：

**Estimating Depth from Monocular Images as Classification Using Deep Fully Convolutional Residual Network**

**内容**：提出了一个很不一样的想法，传统的深度估计都是把它当做是回归任务来完成，而在这篇文章中，作者提出要深度当做是像素分类问题。首先将连续的地面实况深度离散化为几个箱子，并根据箱子的深度范围标记箱子。然后通过训练完全卷积深度残差网络将深度估计问题解决为分类。与估计单个点的精确深度相比，更容易估计其深度范围。更重要的是，可以以概率分布的形式轻松获得深度预测的置信度。可以应用信息增益损失来利用在训练期间接近地面实况的预测，以及用于后处理的完全连接的条件随机场（CRF）以进一步改善性能。

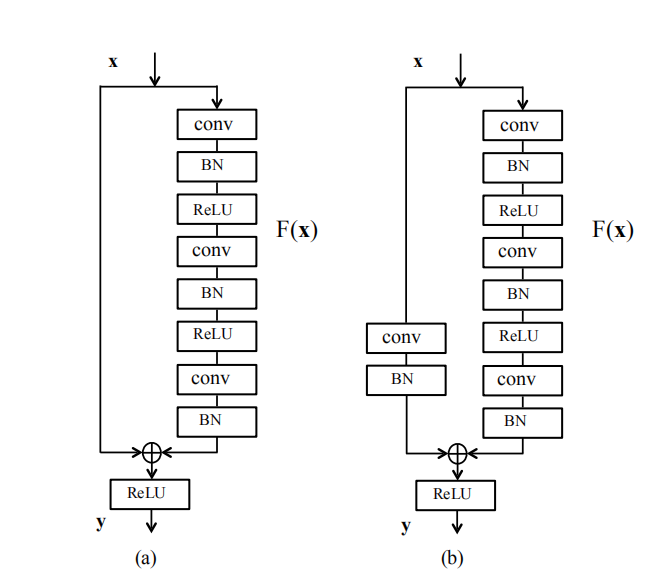


**理解**：在这篇论文中，它可以自然地预测了输出空间上概率分布形式的置信度。 不同的点具有可能的深度值的不同分布。典型的回归模型仅输出可能的深度值的平均值而没有方差（即，缺少预测的置信度）。后处理应用完全连接的条件随机场（CRF）利用完全连接的CRF，可以通过与其连接的其他像素来改善低置信度的像素深度估计。

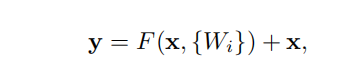
**细节方面的理解：**

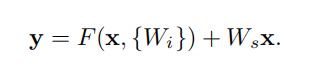
1.网络结构

应用resnet网络（跳过一个或多个层的连接），用卷积层取代了传统CNN架构中的完全连接层。 通过这样做，它使得完全卷积网络能够输入任意大小的图像并输出下采样预测图，再使用简单的双线性插值上采样具有原图的大小。

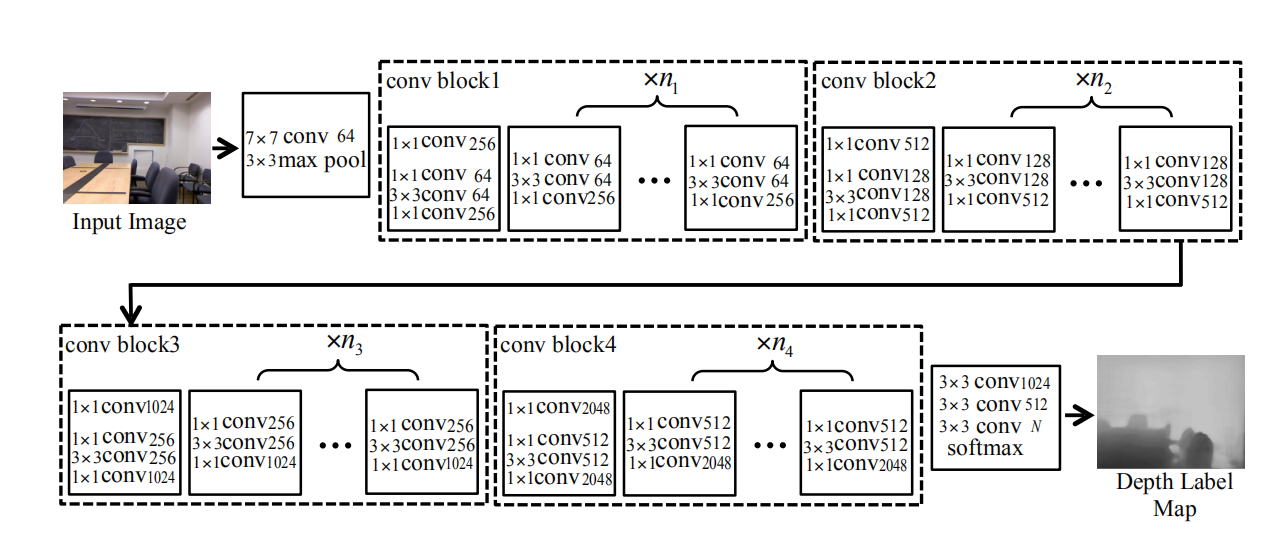


两个残差块，a（身份映射）和b（线性投影）两种单元映射关系





网络组成的整体构架如下：

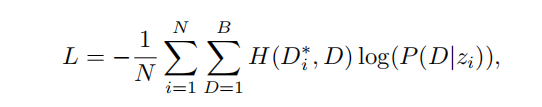


输入图像被馈入卷积层，最大池化层后跟4个卷积块

每个卷积块以具有线性投影的构建块开始，随后是具有身份映射的不同数量的构建块。最后输出图通过双线性插值恢复原图大小。

2.损失函数

使用像素多项逻辑损失函数。与其他的分类损失函数不同，其它的分类与标签值不同就被认为是错误的，并且在更新网络参数方面没有任何贡献。但是深度估计不同接近地面实况深度标签的预测也可以帮助更新网络参数。 这是通过损失函数中的“信息增益”矩阵实现的。

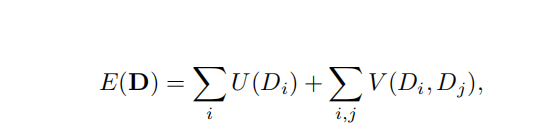


关于信息增益矩阵

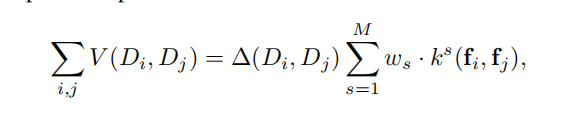
捕获

以概率分布的形式获得每个预测的置信度。 在通过完全连接的CRF进行后处理期间也可以应用这种置信度。

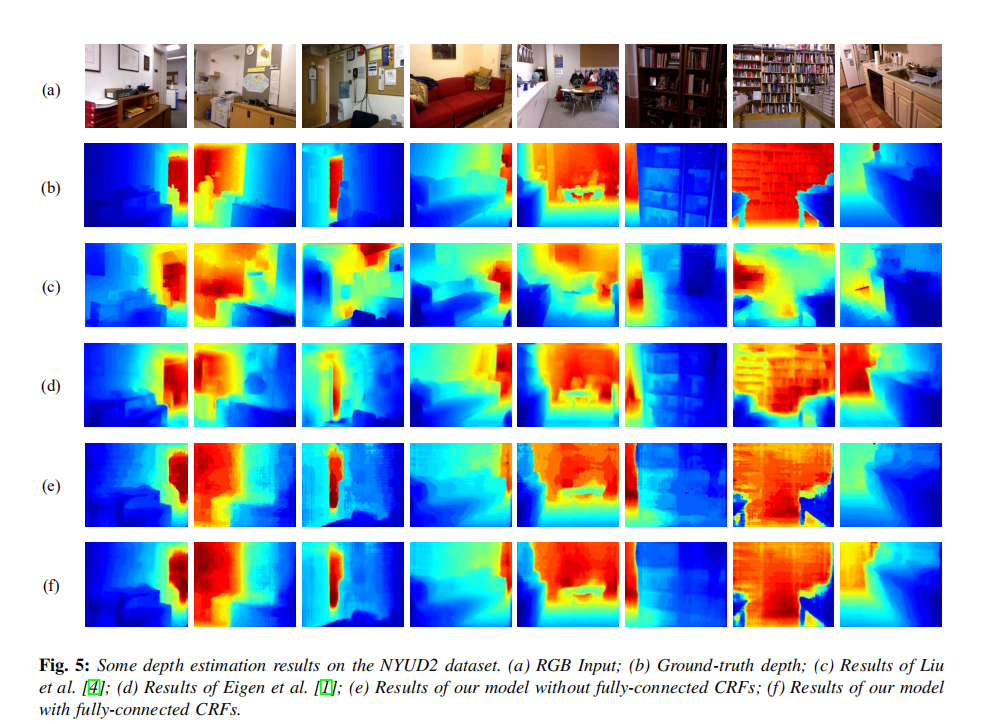
1. **完全连接的CRF**

****

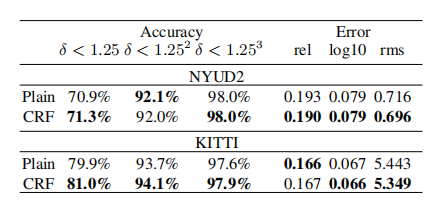
捕获



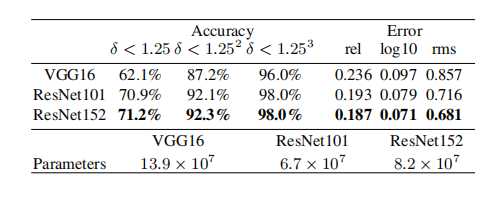
**实验与结果**：



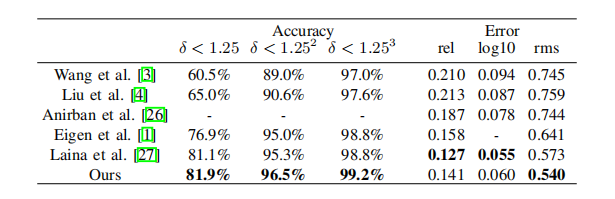
作者通过和其他情况的对比以及他自己不用全连接的CRF的情况可以看出，他的深度图效果比其他都要好更加贴近深度图的情况，但是与此同时它本身深度图在物体的边缘部分并没有表现地很好。



对比了应用CRF和没应用CRF的情况可以看出CRF对它有所帮助。



对比了不同的网络架构可以看出RESnet的准确率明显比VGG较高



根据数据的对比很明显可以看出该论文的方法在准确率上都高于其他的方法

**问题**：

文章在提出的resnet的网络很是复杂理解起来有点难。