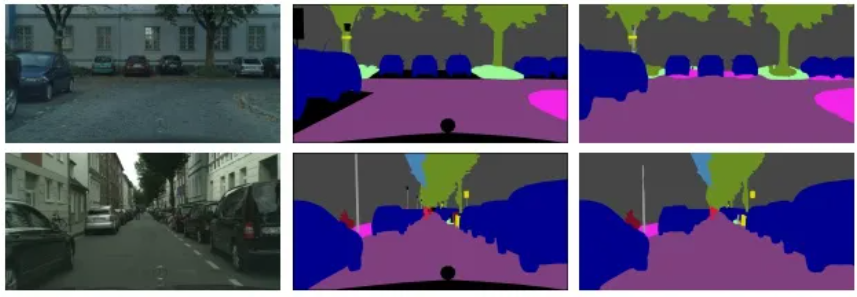
***什么是语义分割？***

按照文献【1】中的介绍，语义分割目的是给图像中每个像素分配唯一的标签（或者类别）。例子如下：



左边图像为真实场景拍摄的图像，右边为逐像素类别标注的结果。行人、车辆、道路和数目等分别采用不同的颜色进行区分，但是并不区分同一类别的不同个体。

***什么是RefineNet？***

在RefineNet【1】发表之时，部分语义分割方法由于采用卷积或者池化层，造成了图像分辨率的降低。

为此，Lin等人提出了RefineNet，一种多路径强化网络。**RefineNet的显式的利用了下采样过程的所有信息，使用远程残差连接来实现高分辨率的预测**。此时，浅层的完善特征可以直接的用于强化高级的语义特征。

***贡献是什么？***

提出了多路径网络，利用多级别的抽象用于高分辨率语义分割；

通过使用带残差连接的同态映射构建所有组件，梯度能够在短距离和长距离传播，从而实现端到端的训练；

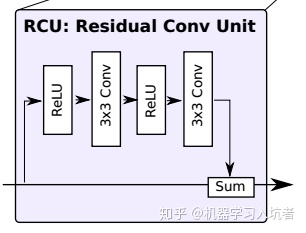
提出了链式残差池化模块，从较大的图像区域俘获背景上下文。使用多个窗口尺寸获得有效的池化特征，并使用残差连接和学习到的权重融合到一起。

***整体架构如何？***

RefineNet共包含下面三个核心模块：

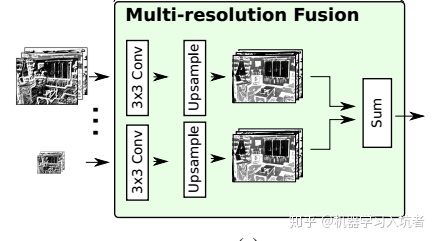
* 残差卷积模块RCU（Residual Conv Unit）；
* 多分辨率融合模块（Multi-Resolution fuse）；
* 链式残差池化模块（Chained Residual Pooling）；

***（1）残差卷积模块***



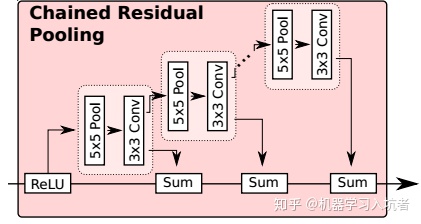
模块中包含激活（ReLU）和卷积（Conv3x3）操作，然后使用加法将前后的特征图融合，这在设计上ResNet思想相同。

***（2）多分辨率融合模块***



将前面多种分辨率的特征图输入到融合模块内后，首先采用卷积层获得尺寸不变的特征图；然后使用上采样操作将所有特征图扩展为尺寸相同的新特征图；最后，使用Sum操作融合所有的特征图。

***（3）链式残差池化模块***



此模块的目的是从大的背景区域中俘获上下文信息，多个池化窗口能获得有效的特征，并使用学习到的权重进行融合。

***实验结果如何？***

RefineNet采用了交并比（IoU）、像素准确率和平均准确率在多个数据集上（如【2】【3】等）进行了分割相关的实验。

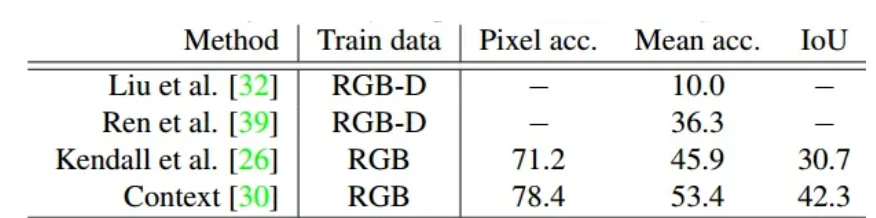
***（1）对象分析任务***



标注的位置为头、躯干、上下臂和上下腿，从上图的预测结果中可以看出RefineNet效果非常好，非常准确的预测出了各个部位。

***（2）语义分割***

RefineNet提供了多个标准数据集的结果，定量指标值均非常出色，比如在SUN-RGBD数据集的结果如下：



***（3）消融实验***

RefineNet是一种灵活的级联结构，下表展示了几种连接方式下的表现：

