### SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation

参考<https://blog.csdn.net/Chunfengyanyulove/article/details/79418814>

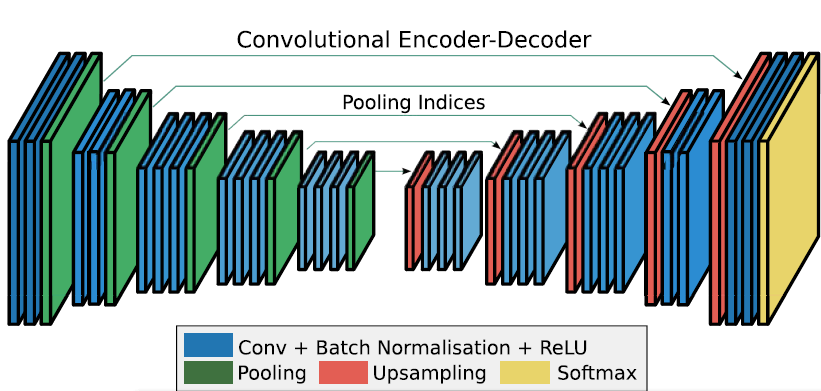
Our architecture, SegNet, is designed to be an efficient architecture for pixel-wise semantic segmentation.

The encoder network in SegNet is topologically identical to the convolutional layers in VGG16

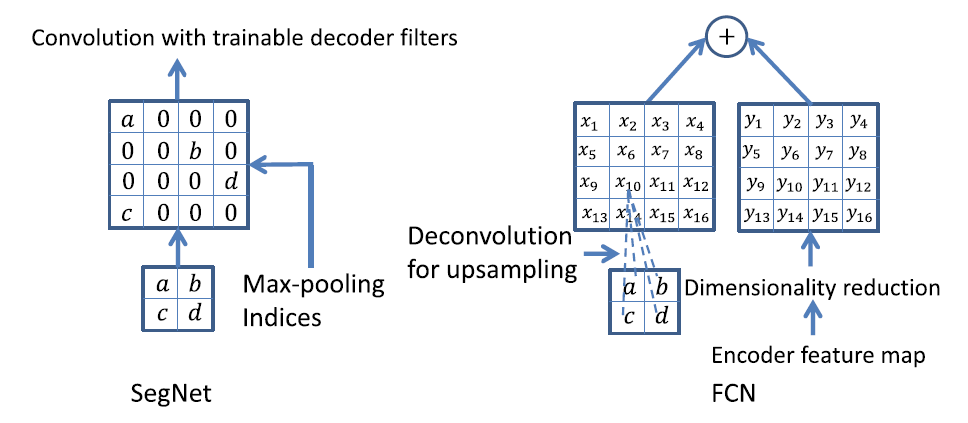
The key component of SegNet is the decoder network which consists of a hierarchy of decoders one corresponding to each encoder.

One of the main contributions of this paper is our analysis of the SegNet decoding technique and the widely used Fully Convolutional Network (FCN)

SegNet包含两部分，分别为encoder网络以及decoder网络，encoder网络采用VGG结构，由VGG16的前13个卷积层构成，去掉全连接层，一方面可以得到较高分辨率的特征图，另一方面可以使网络的参数数量有较大的减少（134M减少到14.7M）。与之对应decoder也具有13层，decoder的最后连接多类别的softmax分类器进行像素的分类。



由于在segment的过程中，边界轮廓信息是至关重要的，然而在进行下采样的过程中，许多信息会丢失，所以需要在下采样之前对轮廓信息进行保存，本文采用一种有效的方法，即max-pooling indices。



 在SegNet中的Pooling与其他Pooling多了一个index功能（该文章亮点之一），也就是每次Pooling，都会保存通过max选出的权值在2x2 filter中的相对位置

 Upsamping就是Pooling的逆过程（index在Upsampling过程中发挥作用），Upsamping使得图片变大2倍。我们清楚的知道Pooling之后，每个filter会丢失了3个权重，这些权重是无法复原的，但是在Upsamping层中可以得到在Pooling中相对Pooling filter的位置。所以Upsampling中先对输入的特征图放大两倍，然后把输入特征图的数据根据Pooling indices放入，下图所示，Unpooling对应上述的Upsampling，switch variables对应Pooling indices。

