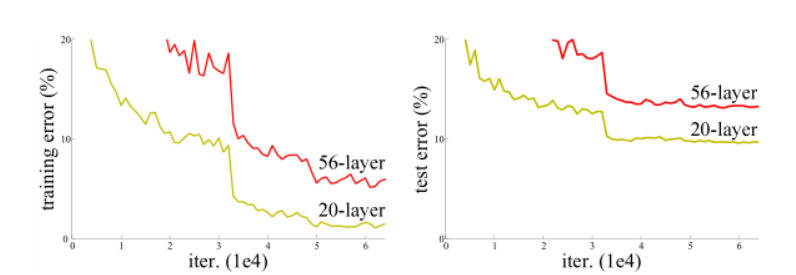
# ResNet原理及其tensorflow 2.0实现

**ResNet介绍：**

ResNet是微软研究院华裔学者何凯明等人提出的一种深层卷积神经网络，ResNet 全名为resdual network，残差网络。

ResNet在2015 ILSVRC和COCO比赛上以巨大的优势取得了5项第一。ResNet的提出有效解决了网络层数增加导致的train error和 test error增加的问题，使得神经网络的层数大大提高。

神经网络层数的增加，堆叠更深层的网络结构时，退化问题却出现了：随着网络深度的增加，网络的准确度会慢慢饱和，然后迅速下降，并且这种下降不是由于过拟合导致的，对于一个合适的深度神经网络而言，增加层数使得train error增加。



**Figure 1. Training error (left) and test error (right) on CIFAR-10 with 20-layer and 56-layer “plain” networks.**

ResNet通过一个深度残差学习框架来解决退化问题，如figure 2所示，其主要思想是增加short connections。对于一个层堆积模块（building block），当输入为x时，其特征映射为,我们将特征映射改为残差：,原始的特征映射为.残差映射比起原始映射更加容易学习，因为当恒等映射（identity mapping）为最优映射时，只需将残差映射训练为0即可，而不必训练堆积层映射为恒等映射；残差映射的最差也是恒等映射，即映射的性能不会因为堆积更多层而下降。

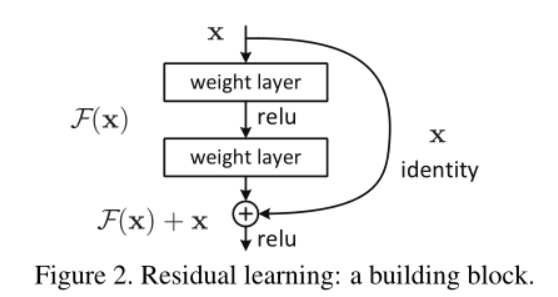


Figure2的结构可以定义为 , 表示将要训练的残差映射，x和y分别代表输入和输出向量。在figure2中可以更加具体定义为,表示激活函数reLU，和x的维度必须相等，才能进行运算，所以在实际情况中，由于卷积层，可能使得的维度减小，需要减小identity x 的维度，可以通过1x1卷积核实现。

**ResNet结构：**

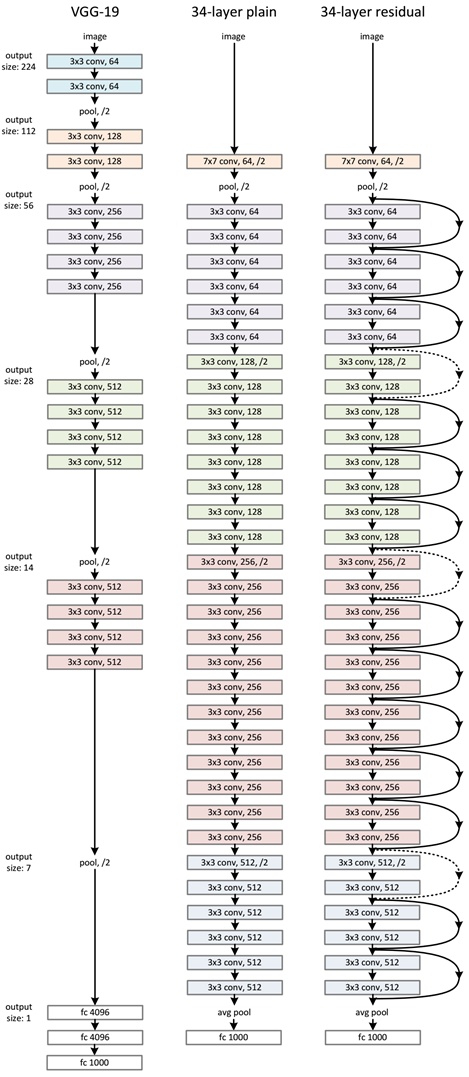
Plain network是不加short connection的网络结构，其baselines很大程度上参考了VGG19的网络结构，其卷积层的设计遵循两个准则：

1: 输出feature map的数目与卷积核数目相同；

2：如果feature map的尺寸减半，卷积核的数目加倍，确保层的时间复杂度不变。

除此之外，卷积层使用stride = 2进行下采样，并在最后使用global average pooling layer和full connection全连接层。

Figure 3右是一个34层ResNet网络，每两层卷积层间增加short connection。虚线的short connection表示和x的维度不一致，有两种解决方案：（1）采用zero-padding增加维度，为了保持参数，卷积层中进行了strides =2的下采样，（2）建立projection shortcut，采用1x1的卷积核来匹配维度

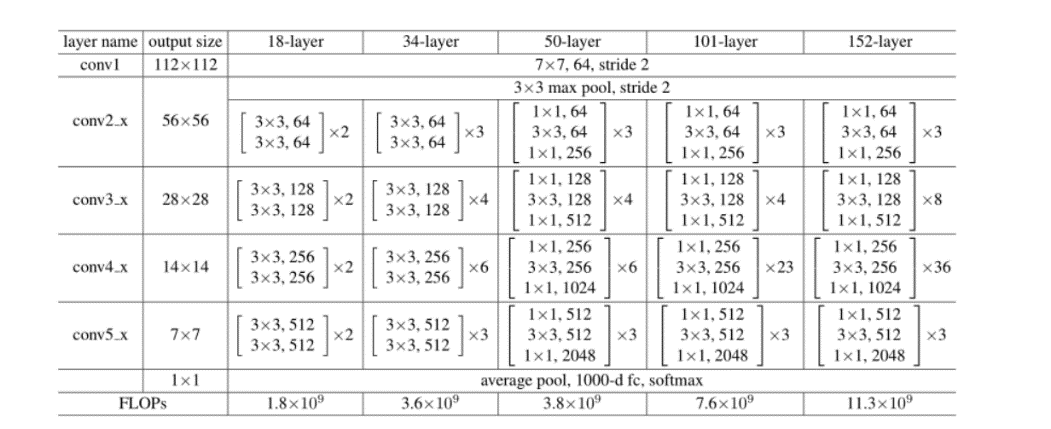
****

**Figure 3. Example network architectures for ImageNet. Left: the VGG-19 model (19.6 billion FLOPs)as a reference. Middle: a plain network with 34 parameter layers (3.6 billion FLOPs).Right: a residual network with 34 parameter layers(3.6 billion FLOPs)**

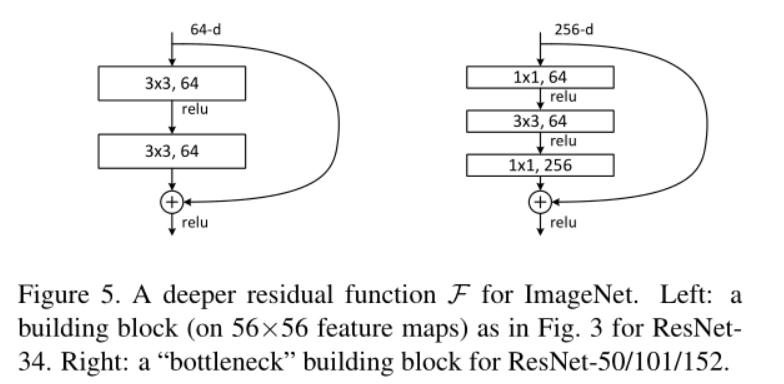
值得注意的是，尽管ResNet-34的层数相对于VGG-19增加了，但是ResNet-34卷积核数目更少，复杂度更小。

ResNet网络结构有18-layer，34-layer，50-layer，101-layer，152-layer几种，其网络结构如table1所示。18层和34层的building block是相同的，为2层卷积结构，卷积核的大小为33；而50层，101层和152层ResNet网络的building block结构是相同的，为3层卷积结构，卷积核大小分别为11，3和11, 如figure 4所示。经过数个building block后，其输出再经过一个平均池化层（average pooling）和一个对应类别维度的全连接层。

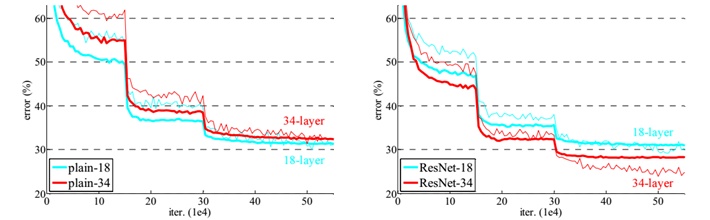
对比plain-18,34网络和ResNet-18,34的训练和validation效果，可以看出残差网络有效解决了增加层数导致退化问题。



**Table 1. Architectures for ImageNet.**



**Figure 4. A deeper residual function F for ImageNet. Left: a building block (on 56×56 feature maps) as in Fig. 3 for ResNet-34. Right: a “bottleneck” building block for ResNet-50/101/152.**



**Figure 5. Training on ImageNet. Thin curves denote training error, and bold curves denote validation error of the center crops. Left: plain networks of 18 and 34 layers. Right: ResNets of 18 and 34 layers.**

**ResNet-18，34 Tensorflow 2.0实现：**