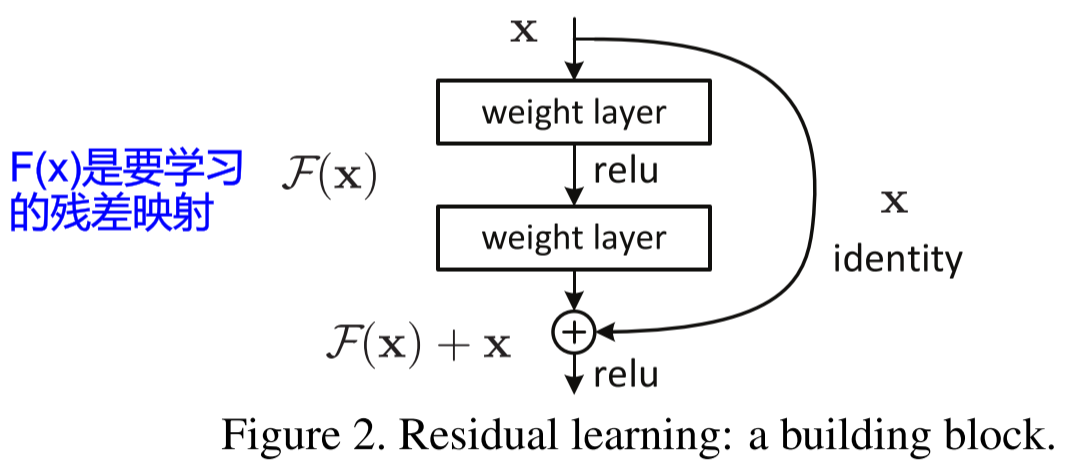
ResNet Notes

1. Motivation

* 随着网络的加深，出现了训练准确率下降的现象，我们可以确定这不是由于过拟合造成的(过拟合的情况训练集应该准确率很高)；所以作者针对这个问题提出了一种全新的网络，叫深度残差网络，它允许网络尽可能的加深，其中引入了全新的结构，如图所示：



* identity mapping

恒等映射，指的就是图中“弯弯的曲线”（x）

* residual mapping

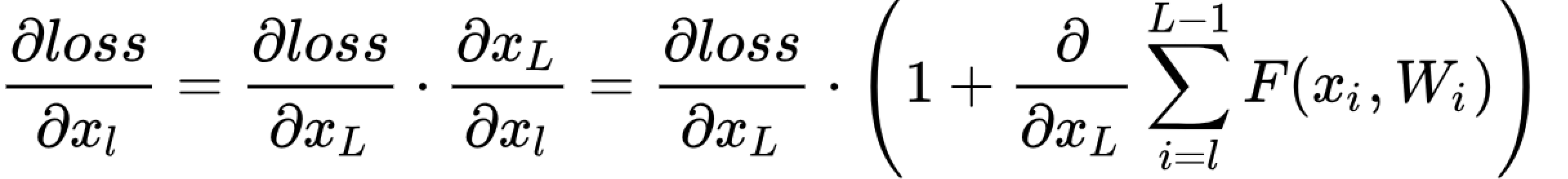
指的就是除了“弯弯的曲线”那部分（F(x) ），residual mapping指的是“**差**”

1. 为什么ResNet可以解决神经网络的退化问题

* 实验证明
* 理论证明

首先对于一个堆积层结构（几层堆积而成），当输入为 x 时其学习到的特征记为 H(x) ，现在我们希望其可以学习到残差 F(x)=H(x)-x ，这样其实原始的学习特征是 F(x)+x。

为什么残差学习相对更容易，从直观上看残差学习需要学习的内容少，因为残差一般会比较小，学习难度小。



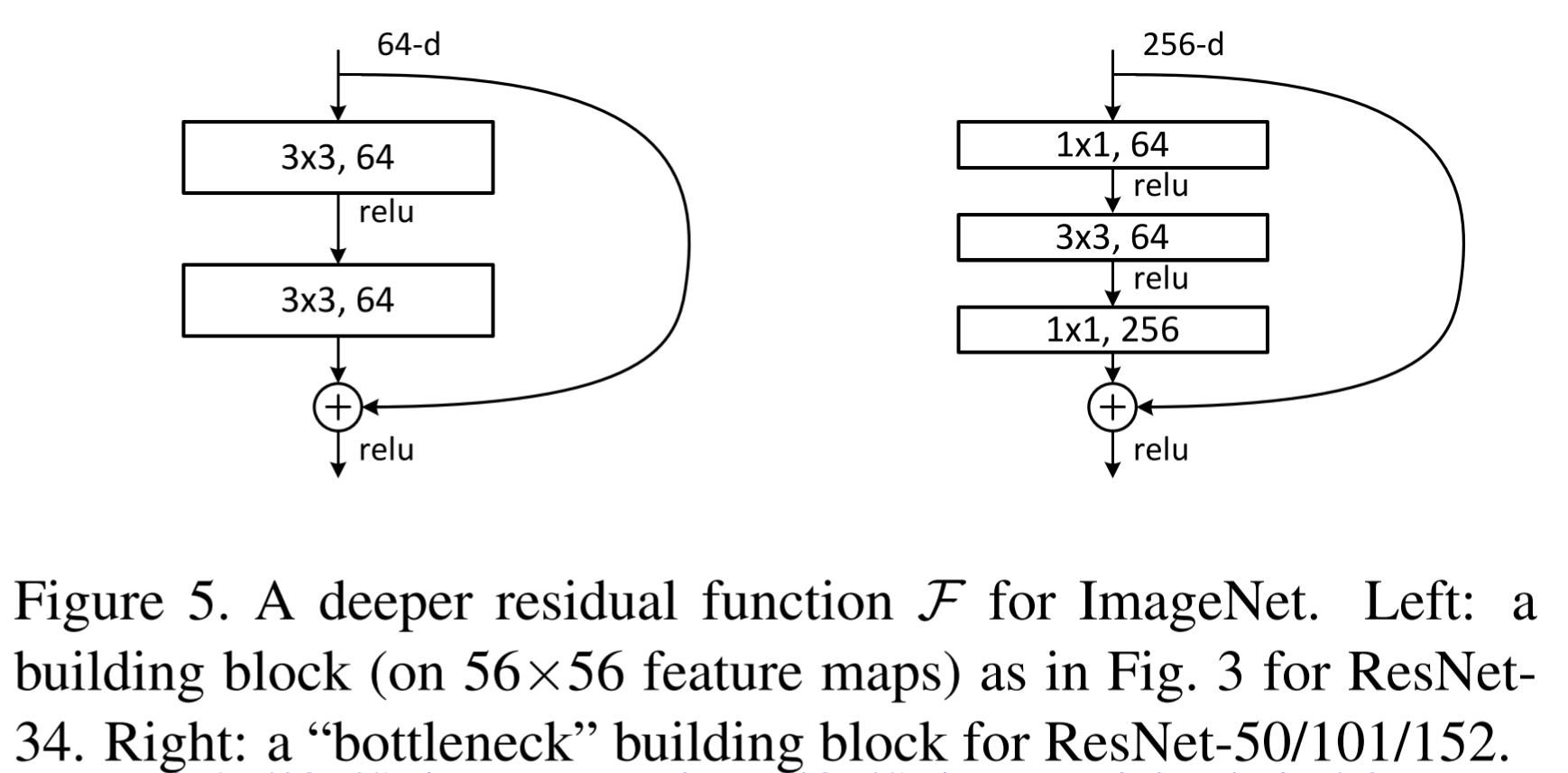
小括号中的1表明短路机制可以无损地传播梯度，而另外一项残差梯度则需要经过带有weights的层，梯度不是直接传递过来的。残差梯度不会那么巧全为-1，而且就算其比较小，有1的存在也不会导致梯度消失。所以残差学习会更容易。要注意上面的推导并不是严格的证明。

1. ResNet网络结构

* Details

ResNet网络是参考了VGG19网络，在其基础上进行了修改，并通过短路机制加入了残差单元，变化主要体现在ResNet直接使用stride=2的卷积做下采样，并且用global average pool层替换了全连接层。ResNet的一个重要设计原则是：当feature map大小降低一半时，feature map的数量增加一倍，这保持了网络层的复杂度。

* 两种残差结构



这两种结构分别针对ResNet34（左图）和ResNet50/101/152（右图），一般称整个结构为一个“building block”。其中右图又称为“bottleneck design”，目的一目了然，就是为了降低参数，第一个1x1的卷积把256维channel降到64维，然后在最后通过1x1卷积恢复。

* F(x) + x，两个map相加的问题

对于短路连接，当输入和输出维度一致时，可以直接将输入加到输出上。当维度不一致时（channel维度增加一倍，width、height维度减小一倍），这就不能直接相加。有两种策略：（1）采用zero-padding增加维度，此时一般要先做一个降采样，可以采用stride=2的avg pooling，这样不会增加参数；（2）采用新的映射（projection shortcut），一般采用1x1的卷积，这样会增加参数，也会增加计算量。

