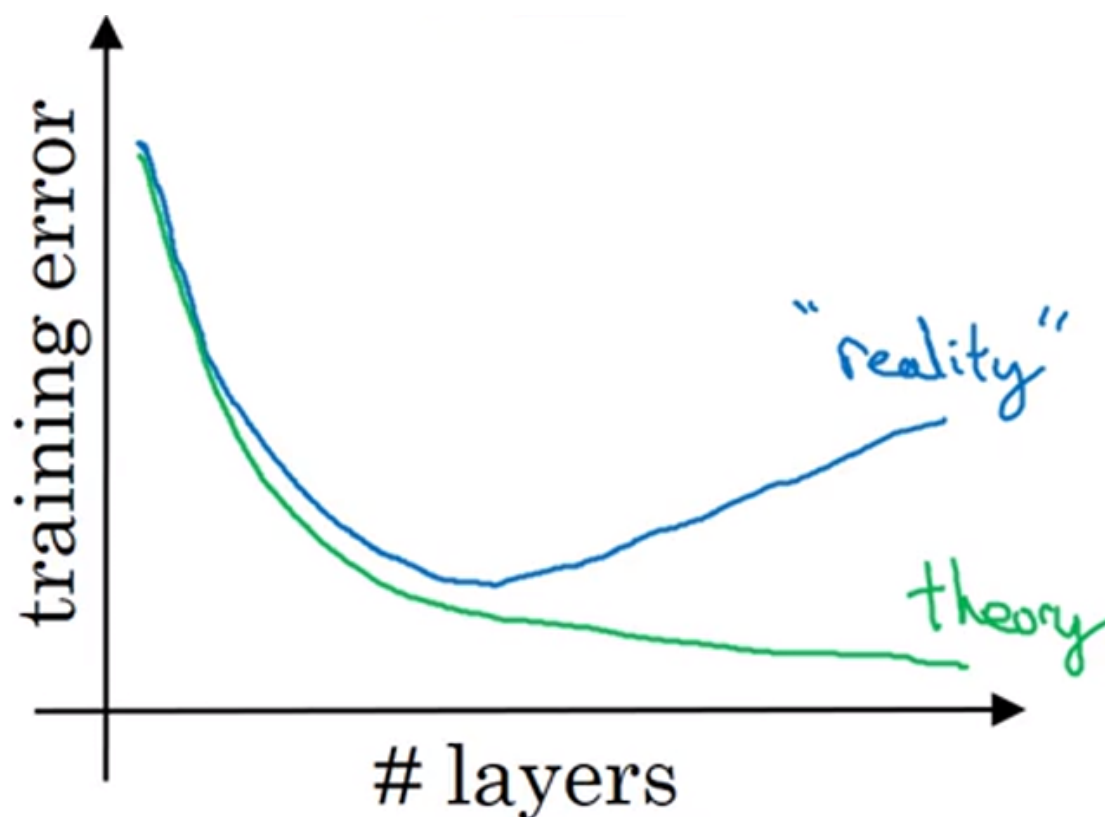
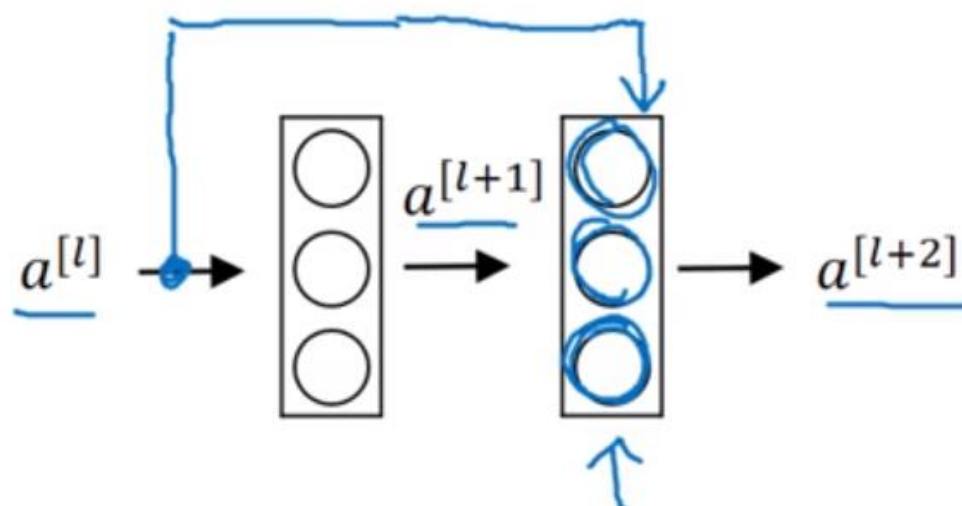


1、为什么引入残差网络

在一般的认识里，随着神经网络层数的加深，训练效果会越来越好，即 loss 递减，accuracy 递增，然而并不然，下图是训练误差随层数的理论和实际值的变化，可以看



到，当层数增加到一定数量时，实际误差反而增大，这种现象我们称之为梯度消失或梯度爆炸。而残差网络 ResNet 可以有效解决这样的问题，随着层数的增加，误差值会越来越小。



2、ResNet

一幅图可以很好的说明残差网络工作原理，在 $a^{[l+2]}$ 生成过程中，也就是 $z^{[2]}$ 线性生成之后，非线性激活之前，将 $a^{[l]}$ 加在里面去，学术上称之为捷径（或跳远连接）即跳过若干层。