# DeepLearning

## 知识点

### 基本知识

* 为什么要加深网络？
* 层的加深，减少了网络的参数。（1 个5 x 5 == 2 个 3 x 3，但是25 > 18）
* 通过叠加层，将ReLU等激活函数夹在卷积层的中间，进一步提高了网络的表现力。这是因为向网络添加了基于激活函数的“非线性” 表现力，通过非线性函数的叠加，可以表现更加复杂的东西。
* 分层学习，由简单到复杂，学习更加高效。
* 在深度模型中，每层的输出特征图的尺寸大都随着网络深度而变化，主要是高和宽越来越小，输出特征图的深度随着网络层数的深度而增加，高和宽的减小有助于减小计算量，而特征图（一个卷积核的输出）深度的增加则使每层输出中可用特征数量的增多。
* 梯度爆炸/消失

随着层数的增多，在网络中反向传播的梯度会随着连乘变得不稳定，变得特别大或者特别小。

解决办法：

* 使用BatchNorm

<https://blog.csdn.net/ygfrancois/article/details/90382459>

* 激活函数换为ReLu
* 使用Xaiver初始化
* 网络恶化

随着深度的增加，网络的性能会越来越差，直接体现为在训练集上的准确率会下降，不是由过拟合造成的，过拟合的话在训练集上的表现应该越来越好。

应该是：越深的网络反向梯度越难传导。

解决办法：

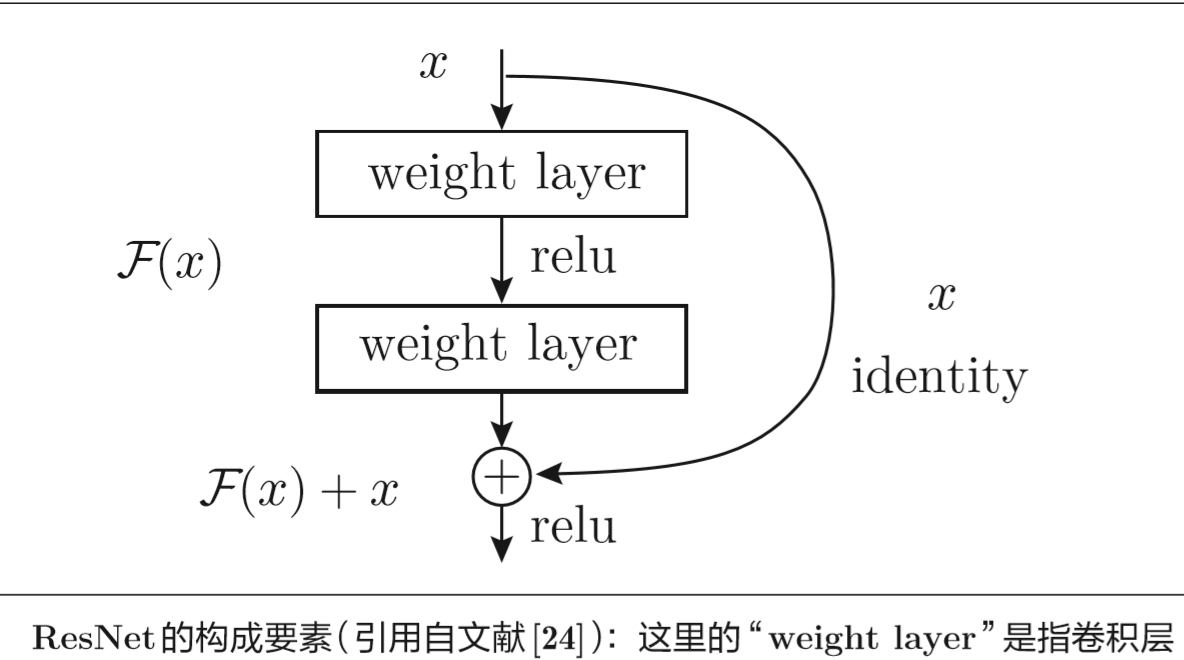
残差网络 resNet

## VGG

* GlobalAveragePooling层
* 可以实现任意图像大小的输入
* <https://www.jianshu.com/p/510072fc9c62>

## resNet

* 结构



* 作用

通过快捷结构，反向传播时信号可以无衰减地传递。

不对来自上游 的梯度进行任何处理，将其原封不动地传向下游。因此，基于快捷 结构，不用担心梯度会变小（或变大），能够向前一层传递“有意义 的梯度”。