CNN

[1. 参考资料 2](#_Toc492160157)

[2. 早期CNN 2](#_Toc492160158)

[2.1. LeNet 2](#_Toc492160159)

[3. 2012 2](#_Toc492160160)

[3.1. AlexNet 2](#_Toc492160161)

[4. 2014 3](#_Toc492160162)

[4.1. VGG 3](#_Toc492160163)

[4.2. GoogleNet 6](#_Toc492160164)

[5. 2015 6](#_Toc492160165)

[5.1. ResNet 6](#_Toc492160166)

1. 参考资料
2. [Deep Learning回顾之LeNet、AlexNet、GoogLeNet、VGG、ResNet](<http://www.cnblogs.com/52machinelearning/p/5821591.html>)
3. [http://cs231n.stanford.edu](http://cs231n.stanford.edu) 来自该课程2017年slides
4. CNN与Visual Recognition

Visual Recognition 相关任务有很多：

1. Action classification 行为分类
2. Image classification 图像分类
3. Image captioning 为图像配文字
4. Object detection 物体识别
5. Object tracking 物体跟踪

CNN：

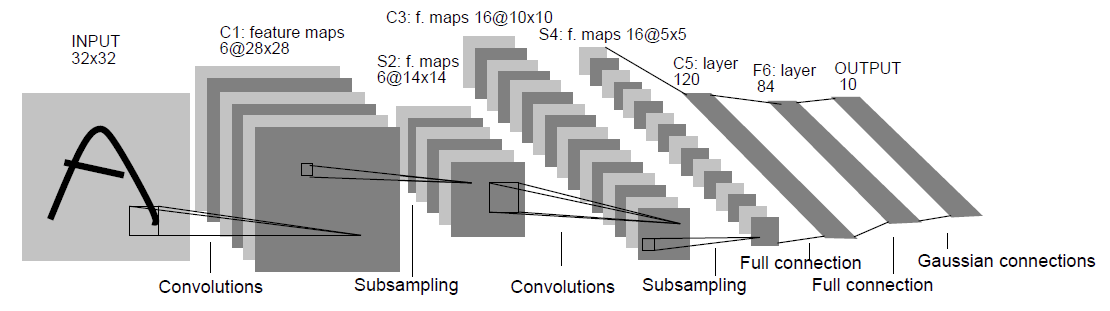
CNN能够抽象局部信息，获取高层特征。

1. 早期CNN
   1. LeNet

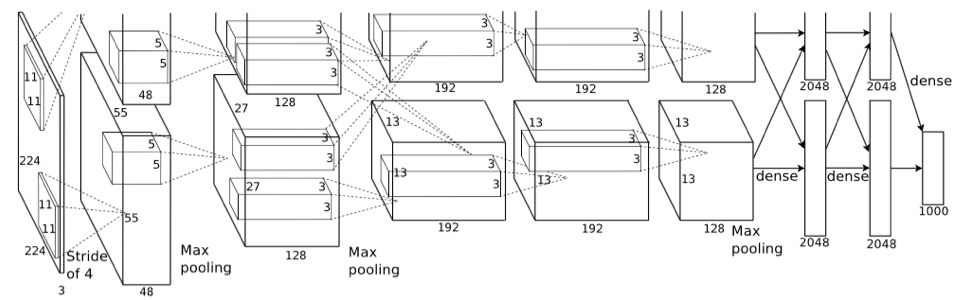
Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324, 1998.

1998年的LeNet5标注着CNN的真正面世，但是这个模型在后来的一段时间并未能火起来，主要原因是费机器（当时没有GPU），而且其他的算法（SVM等）也能达到类似的效果甚至超过。

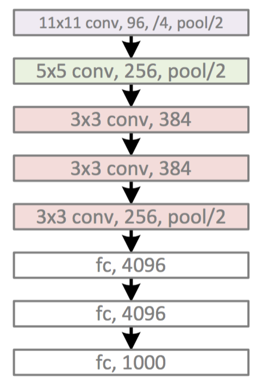
网络结构：



1. 2012
   1. AlexNet



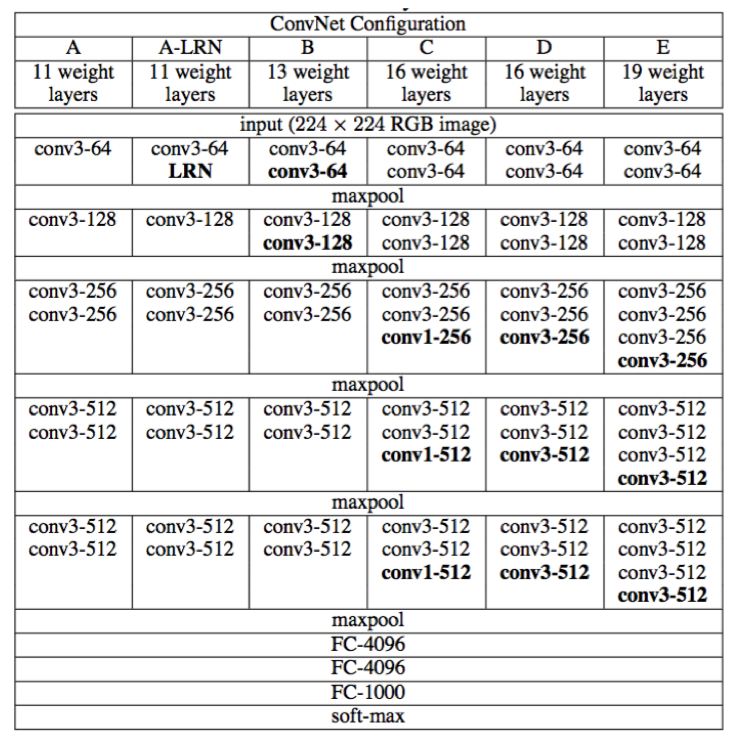
网络结构



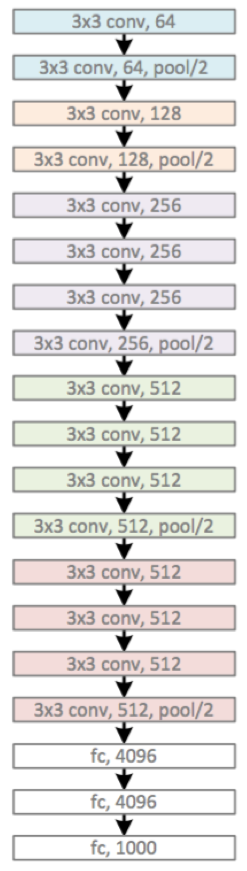
简化版dataflow

重要创新：

1. 数据增强：翻转，裁剪、平移、颜色光照变换等，防止过拟合。
2. Dropout：防止过拟合，后来还出来了一些变种，比如DropConnect等。
3. ReLU激活函数：计算简单、偏导简单、右侧稀疏防止过拟合等优点。
4. Local Response Normalization：简称LRN，实际就是利用临近的数据做归一化。这个策略贡献了1.2%的Top-5错误率。
5. Overlapping Pooling：Overlapping的意思是有重叠，即Pooling的步长比Pooling Kernel的对应边要小。这个策略贡献了0.3%的Top-5错误率。
6. 多GPU并行
7. 2014
   1. VGG



网络结构

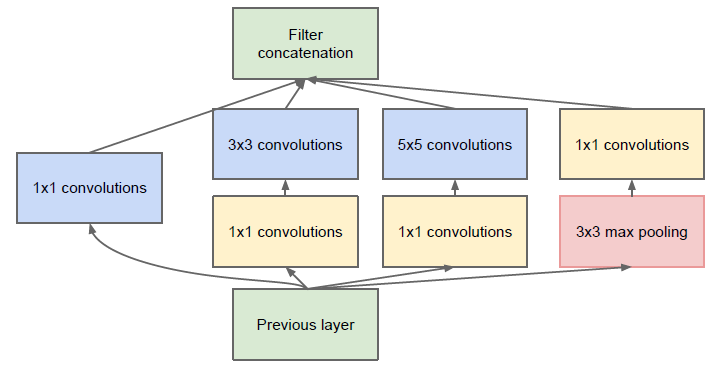


VGG-19简化版dataflow

* 1. GoogleNet

主要的创新在于他的Inception，这是一种网中网（Network In Network）的结构，即原来的结点也是一个网络。Inception一直在不断发展，目前已经V2、V3、V4了。

Inception的结构如下图所示，其中1\*1卷积主要用来降维，用了Inception之后整个网络结构的宽度和深度都可扩大，能够带来2-3倍的性能提升。

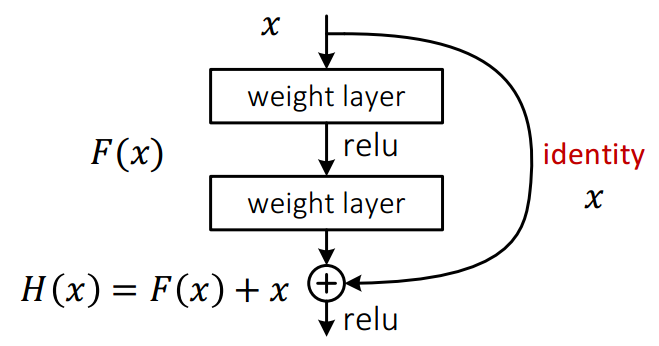


Inception结构

1. 2015
   1. ResNet

没有最深，只有更深（152层）。听说目前层数已突破一千。

主要的创新在残差网络，其实这个网络的提出本质上还是要解决层次比较深的时候无法训练的问题。这种借鉴了Highway Network思想的网络相当于旁边专门开个通道使得输入可以直达输出，而优化的目标由原来的拟合输出H(x)变成输出和输入的差H(x)-x，其中H(X)是某一层原始的期望映射输出，x是输入。



ResNet的block结构