Going Deeper with Convolutions

本文主要关注针对计算机视觉的高效深度神经网络结构，通过改进神经网络的结构达到不增加计算资源需求的前提下提高网络的深度，从而达到提高效果的目的。

1. **Main Contribution**

1.提高了网络内部计算资源的利用率，在保持计算量预算不变的前提下增加了网络的深度和广度。

1. 结构体系的设计决策是基于Hebbian准则和多尺度处理机制
2. 竞赛中的一个22层深度神经网络

**2. Ralated Work**

1. 本文提出的网络结构为Inception，得名于论文参考文献12（network in network）。

1. CNN最近的趋势是通过增加层数和层大小,并使用dropout来解决过度拟合的问题
2. 论文参考文献15使用不同尺度的Gabor过滤器来处理多尺度问题，同本文的Inception Model类似。

4. 本文借鉴参考论文12，使用了很多1×1的卷积核。卷积核在本文中的作用主要在于降维，以此来去除计算瓶颈

**3. Motivation and High level considerations**

**3.1 直接增加CNN大小的缺点：**

1. 更倾向于overfitting

2. 大量的计算资源

**3.2 如何解决：**

1. 基本方法根本上将完全连接转换成稀疏连接的架构；

2. 论文的参考文献2表明，考虑到统计相关性，一个稀疏网络结构可以重新构建出最优结构。并产生了Hebbian principle——neurons that fire together, wire together；

3. 从更底层考虑，现在的硬件在非一致稀疏数据结构上的计算非常不高效，尤其在这些数据上使用已经为密集矩阵优化过的库时。原来自论文参考文献9以来，都会使用随机稀疏的网络结构来打破对称性，提高学习率。但论文参考文献11中又重新使用全连接的结构，以图利用密集计算的高效性；

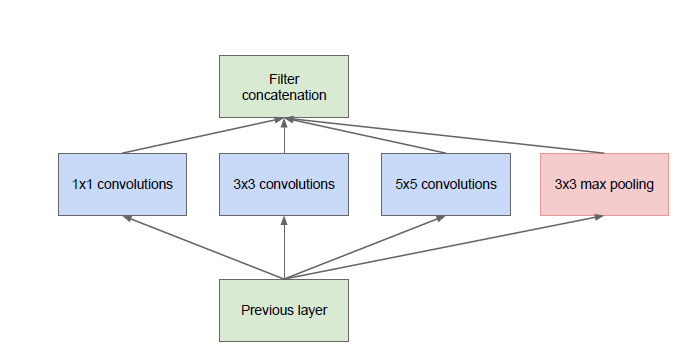
4. 所以，现在的问题是有没有一种方法，既能保有网络结构的稀疏性，又能利用密集矩阵的高计算性能。论文提出了一种Inception Module，可以达到此等效果

**4. Architectural Detail**

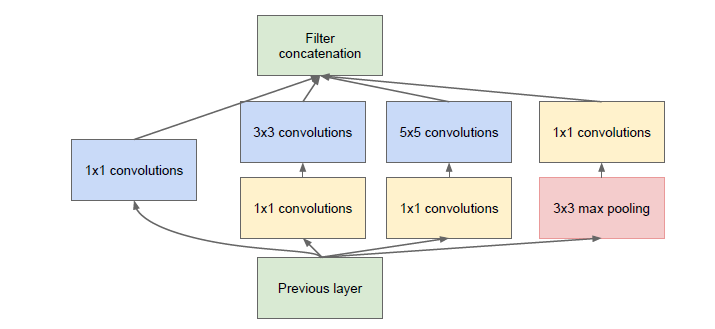
The main idea of the Inception architecture is based on finding out how an optimal local sparse structure in a convolutional vision network can be approximated and covered by readily available dense components。

Inception架构的设计理念是找到完美的局部稀疏结构可以被现成的稠密组件近似或完全覆盖。

如何发现最优结构呢？ 可以这样考虑，较低的层次对应着图像的某个区域，使用1×1的卷积核仍然对应这个区域，使用3×3的卷积核，可以得到更大的区域对应，以此类推，因而设计如文中的图2(a)。



为了降维，使用1×1的核进行降维，设计如图2(b)。降维能够起效主要得益于embedding技术的发展，即使较低的维度仍然可以包含很多信息。在Filter concatenation层将1×1/3×3/5×5的卷积结果连接起来。如此设计的好处在于防止了层数增多带来的计算资源的爆炸性需求。从而使网络的宽度和深度均可扩大。使用了Inception层的结构可以有2-3×的加速。



**5．Googlenet**

