**VGG阅读总结**

15331416 赵寒旭

VGGNet是牛津大学计算机视觉组(Visual Geometry Group)和Google Deepmind公司研究员一起研发的深度卷积 神经网络。VGGNet在AlexNet的基础上探索了卷积神经网络的深度与性能之间的关系，通过反复堆叠3\*3的小型 卷积核和2\*2的最大池化层，VGGNet构筑的16~19层卷积神经网络模型取得了很好的识别性能，同时VGGNet的拓展性很强，迁移到其他图片数据上泛化能力很好，而且VGGNet结构简洁，现在依然被用来提取图像特征。

/\*

注：部分理解内容来源于网上的blog，

这里是写的很好的一篇参考：（详细描述了Alex和VGG）

<https://blog.csdn.net/u011974639/article/details/76146822#vggnet>

\*/

**VGG的特点**

**1. 针对网络架构**

1）全部使用的卷积核和的池化核，通过不断加深网络结构来提升性能。

2）使用多个小卷积核串联组成卷积层，和先前的大卷积核相比，拥有同样的感受野，却有着更少的参数，更强的非线性变换，因此有着更强的特征提取能力。

3）使用卷积层，意义在于线性变换，输入通道和输出通道数不变，没有发生降维。

4）LRN层作用不大（并不能提升性能，且会增加计算量），仅在一个网络中使用。

**2. 针对过拟合现象**

1）数据增强

使用Multi-scale的方法，将原始图像缩放到不同尺寸，再随机裁取固定大小的图片，这样可以增加很多数据量，防止模型过拟合。

2）在预测时采用多裁剪的多尺度配合（两者配合使用，效果比单使用好），可以很好的提升模型的性能。

**3. 针对训练速度**

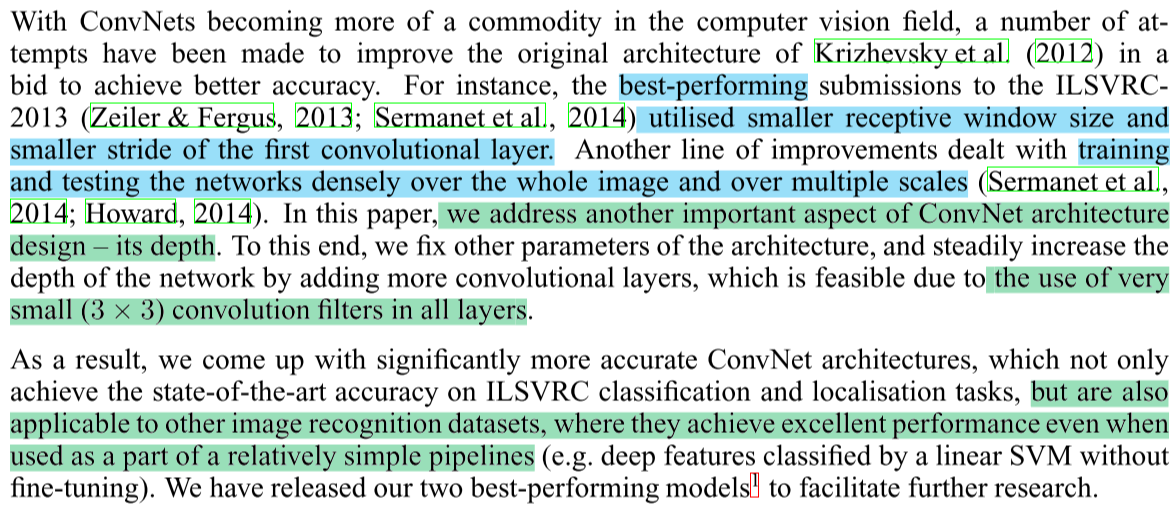
可以先训练底深度的A网络，再复用A网络的权重初始化后面的几个复杂模型，这样训练收敛的速度更快

**Abstract**

研究了卷积网络深度对大型图像识别准确性的影响，主要评估小的卷积核（）同等架构下随着网络深度的增加卷积网络的性能变化，随着网络深度达到16-19层，网络的性能也有着显著的提升。

主要的贡献：展示出网络的深度是算法优良性能的关键部分(使用了非常小的滤波器（）。

**1.** **Introduction**



为在AlexNet原始架构上实现更高的准确率，可以在第一个卷积层上使用更小的感受野和更短的步长，另一种路线是使用多尺度的密集训练和测试网络（后文讨论了Multi-crop和Multi-scale的使用对模型性能的影响）。

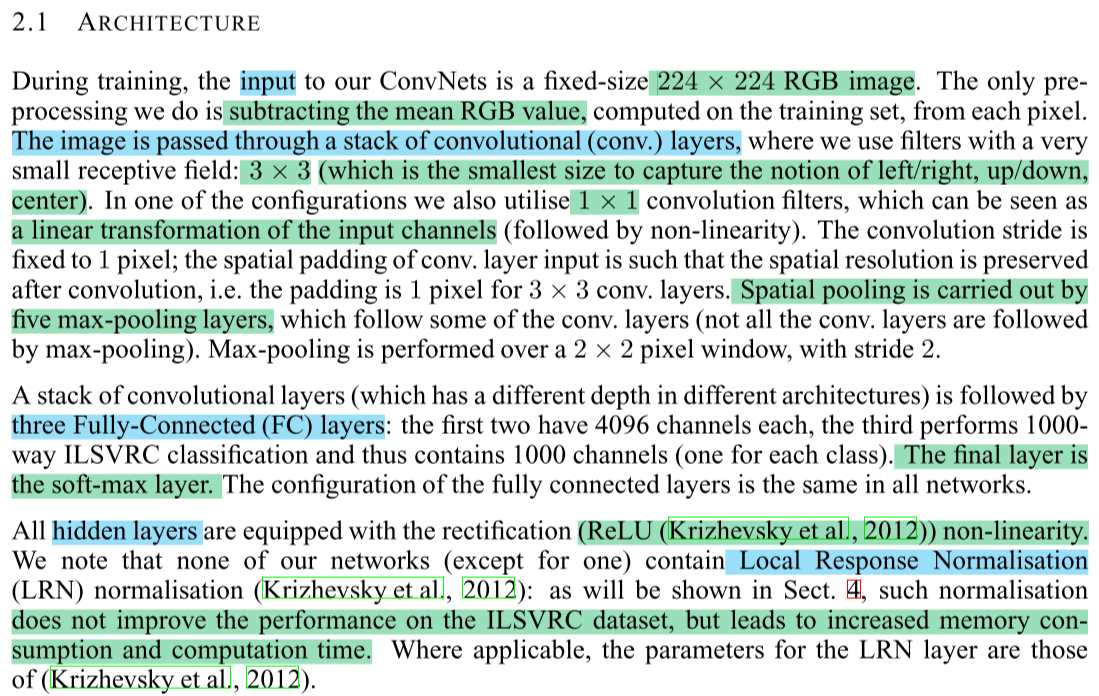
本文主要研究卷积神经网络架构设计的另一个重要影响因素：深度。

结果：实现了在ILSVRC上更精确的卷积神经网络，在其他数据集上也有很好的表现，迁移性很强。

**2. ConvNet** **Configurations（卷积网络配置）**

为在公平的环境下测量网络深度增加带来的提升，我们所有的卷积网络层结构使用相同的原则来设计。

**2.1 Architecture**



1）input：a fixed-size 224 × 224 RGB image

pre-processing：从每个像素减去RGB值的均值

2）The image is passed through a stack of convolutional layers:

使用非常小的的滤波器（是足够捕获一个区域信息的最小的卷积核）

其中一个配置中使用的卷积滤波器（对输入通道的线性变换），这样的卷积不改变输入通道的维度，且可以提高模型的学习能力。

convolution stride：1 pixel

spatial padding：对卷积层输入的扩展要保证卷积操作前后分辨率一致（eg. 的卷积层要扩展1 pixel）

Spatial pooling：由5个最大池化层实现（跟在某些卷积层后面）

Max-pooling：在的窗口进行，步长为2

3）3 Full-Connect layers

4096,4096,1000

最后一层是softmax层。

5）Hidden layers

整流函数：ReLU

6）Local Response Normalisation

除了一个网络之外，我们所有的网络都不含局部响应归一化，它不能提升在ILSVRC数据集上的性能，但会导致内存消耗和计算时间的增加。

7）LRN layer parameters

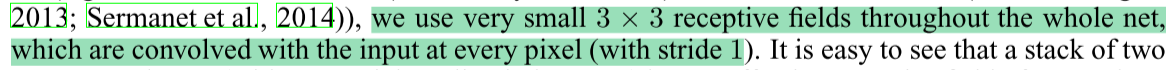
Where applicable, the parameters for the LRN layer are those of (Krizhevsky et al., 2012).

**2.2 Configurations**

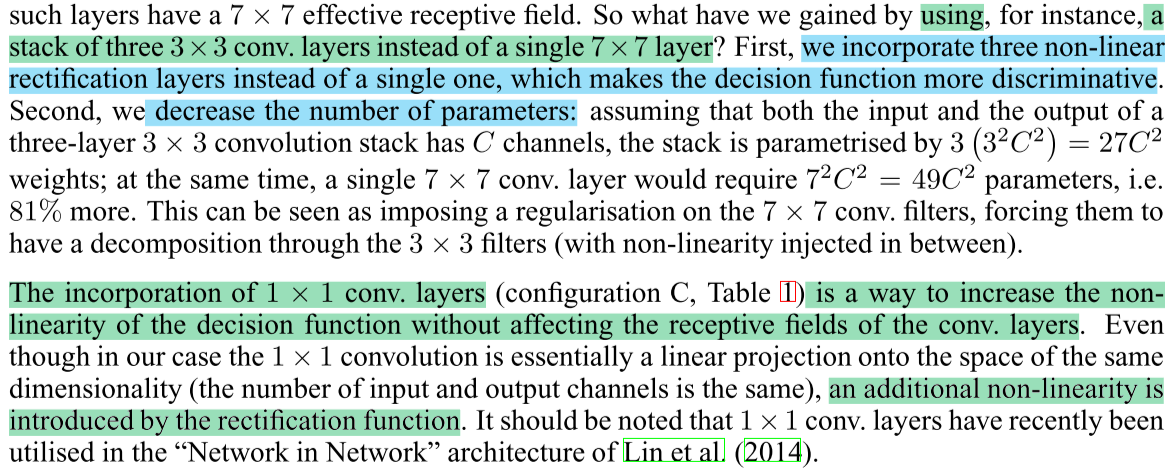
卷积核尺寸很小，卷积核的数目初始为64，每过一次max-pooling层，卷积核数目翻倍，最终增加到512个。

尽管网络深度很大，网络中的权重数目并不会比有更大卷积层和感受野的更浅的网络多。

**2.3 Discussion**



------

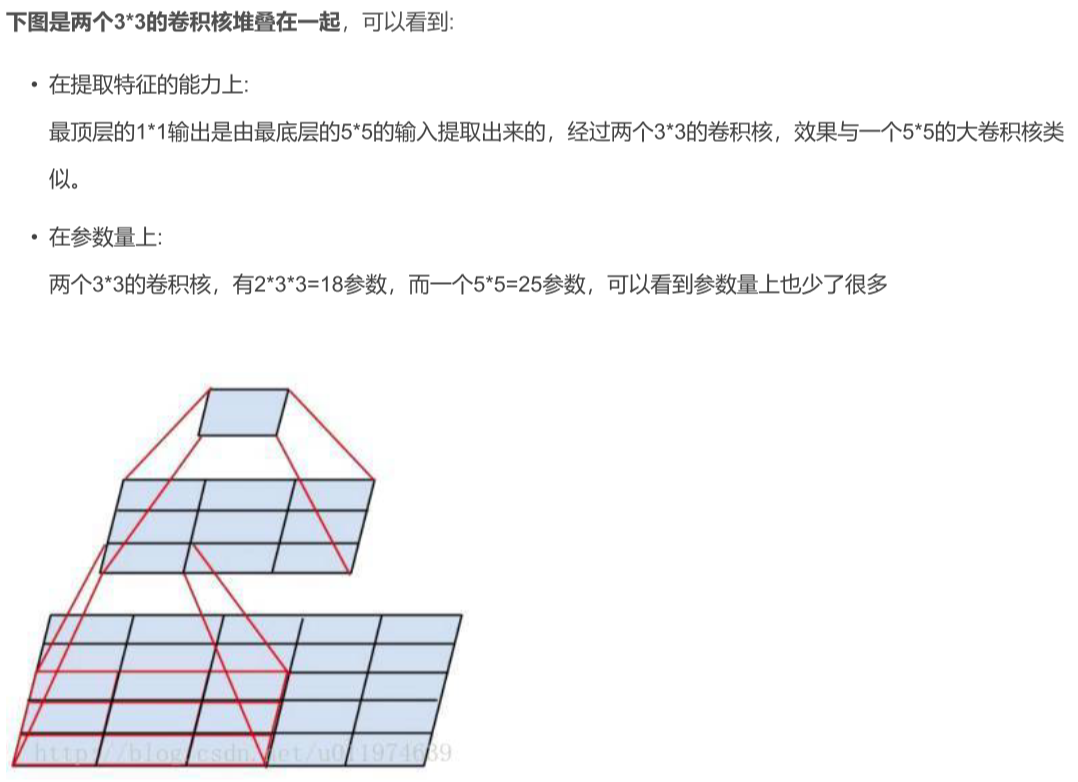


使用多个小卷积核堆叠，感受野与大卷积核相同。

相比于单层，使用三个非线性整流层会使决策函数更有判别力。（多个小卷积核的堆叠会比单个大卷积核有更多激活函数的转换（ReLU），学习能力更强）

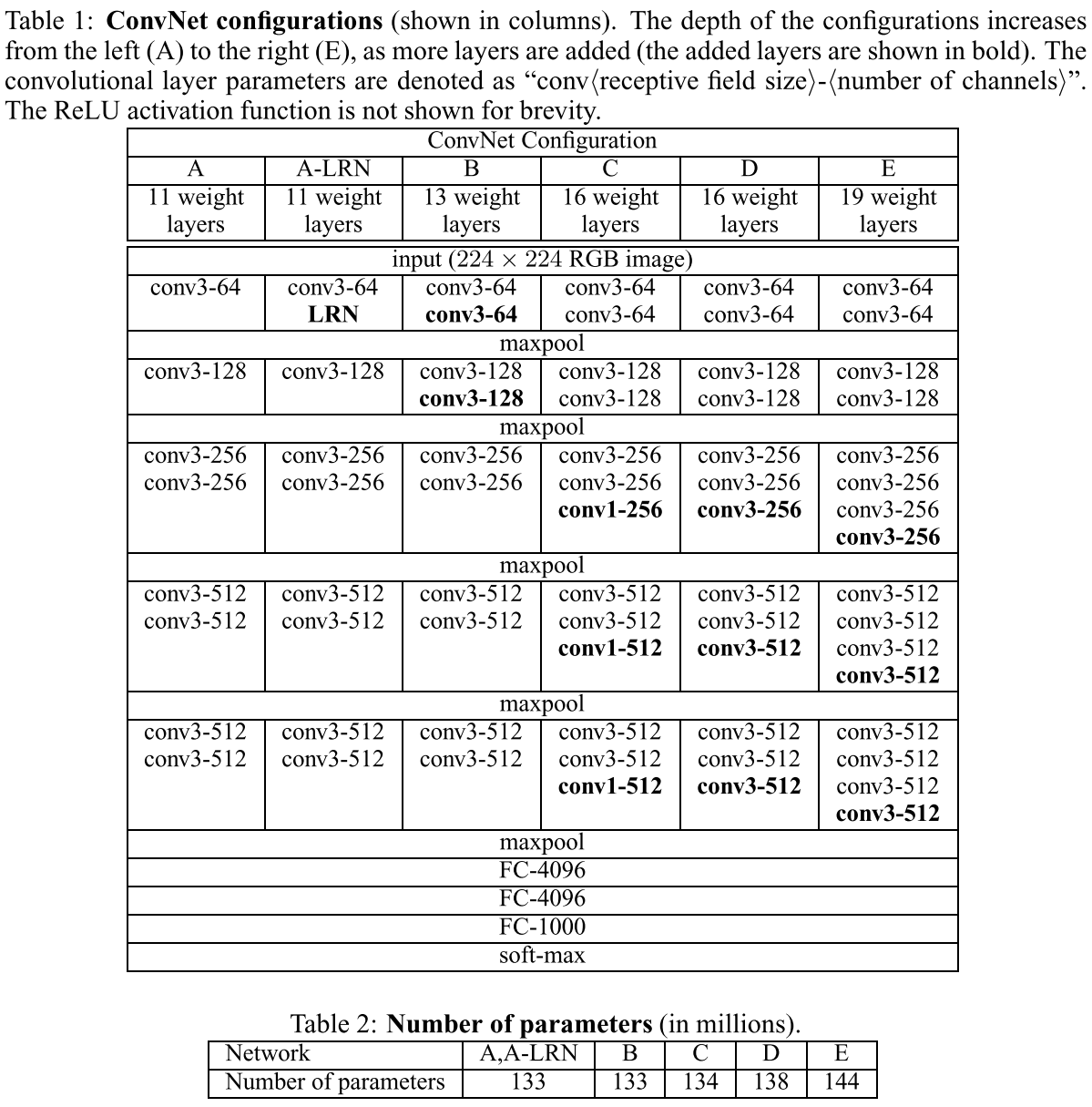
同时，多个小卷积核会有更少的参数。

使用的卷积核，可以在不影响卷积层感受野的情况下增强决策函数的非线性，即使在本文模型使用卷积核是同维度上的线性变换。（卷积后会经过ReLU处理，而ReLU在输入大于0是线性的）

左图是两个3\*3的卷积核堆叠在一起，可以看到:

在提取特征的能力上: 最顶层的1\*1输出是由最底层的5\*5的输入提取出来的，经过两个3\*3的卷积核，效果与一个5\*5的大卷积核类似。

在参数量上:两个3\*3的卷积核，有2\*3\*3=18参数，而一个5\*5=25参数，可以看到参数量上也少了很多



**3. Classification Framework（分类框架）**

**3.1 Training**

**3.2 Testing**

**3.3 Implementation Details**

4. Classification Experiments（分类实验）

4.1 Single Scale Evaluation

4.2 Multi-scale Evaluation

4.3 Multi-crop Evaluation

4.4 ConvNet Fusion

4.5 Comparison with the State of the Art

5. Conclusion（总结）

A Localisation

A.1 Localisation ConvNet

A.2 Localisation Experiments

B Generalisation of Very Deep Features