# 网络中的网络[[1]](#footnote-1)

赵淼译

(北京交通大学电子与信息工程学院，电子科学与技术，15650761201)

这里使用局部全连接层代替卷积层获得局部感受野的信息。这种mlpconv更具有抽象能力。经过多层mlpconv的堆叠获得深度网络，在最后一层mlpconv后面进行空间平均池化，然后将得到的向量输入softmax层获得最后的预测。



Mlpconv：在每个局部感受野加一个全连接层获得这一层的输出。然后将这个全连接层进行滑动，MLP在各个局部感受野是共享的。

在原先由于全连接层的存在，很难解释类别信息如何传递回原先的卷积层，因为全连接层会起到黑盒子的作用，相比之下，全局池化更有意义和可解释性，将特征图的各层解释为高级的语义信息的区别所在。

## 1 卷积

卷积核表现为一种线性滤波运算，如果潜在概念线性可分，线性卷积足够使用，但是如果当抽象概念需要更强大的抽象能力的时候，需要更多线性卷积滤波器组合来表示更复杂的概念。

## 2 网络中的网络

### MLP 卷积层

Mlpconv卷积层如下所示

|  |  |
| --- | --- |
|  | 1 |

从交叉通道池化的角度来看，等式2等同于在一个正常的卷积层上加一个级联的交叉通道参数池。每个池层是表现为对于输入特征图的线性重组。跨通道池化的特征图在下一层再次进行跨通道的池化。这允许复杂的跨通道信息和进行科学系的交互。

可以将其理解为1乘1卷积的层。实现NIN就是在原来的卷积后再加上1\*1的卷积即可，如下图，有三个NIN层，那么第一个NIN的实现应该是conv1[3\*3],(kernal) conv2[1\*1],conv3[1\*1]这种。CNN里的卷积大都是多通道的feature map和多通道的卷积核之间的操作（输入的多通道的feature map和一组卷积核做卷积求和得到一个输出的feature map），如果使用1x1的卷积核，这个操作实现的就是多个feature map的线性组合，可以实现feature map在通道个数上的变化。接在普通的卷积层的后面，配合激活函数，就可以实现network in network的结构了

1. 实现跨通道的交互和信息整合

2. 进行卷积核通道数的降维和升维

因为NIN中的MLP层可以用两层1×1卷积核来代替，比如当前这一层是54×54×96的图像层，然后过一个1×1×96的卷积核，还是一个54×54×96的卷积层，然后再过一个1×1×96的卷积核，还是一个54×54×96的卷积层。

但是这样但看最开始那个96个特征层的图像同一个位置不同层之间的像素点，相当于过了一个96×96×96的MLP网络

1×1卷积核确实可以实现全连接层

具体的操作是，输入是224x224x3 的图像，假设经过变换之后最后一层是[7x7x512]的，那么传统的方法应该将其展平成为一个7x7x512长度的一层，然后做全连接层，假设全连接层为4096×1000层的（假设有1000个分类结果）。

那么用1×1卷积核怎么做呢，因为1×1卷积核相当于在不同channel之间做线性变换，所以：

先选择7×7的卷积核，输出层特征层数为4096层，这样得到一个[1×1×4096]层的

然后再选择用1×1卷积核，输出层数为1000层，这样得到一个[1×1×1000]层的

这里先用一个普通卷积作用于一个局部区域形成一个区部感受野，再使用1\*1卷积，将已经经过加权后的局部区域输出值进行通道间的组合。这样会形成类似全连接的效果。

### 全局平均池化

下面是一个简单的三层mlpconv结构，后面接着全局池化。全局池化对整个通道信息进行聚合，但是缺乏空间信息的保留。



实验证明mlpconv搭配全局池化效果较好，conv搭配全局池化效果不如conv搭配droout，理解为全局平均池化对线性卷积的要求较高。

1. 赵淼 电信研1706班 电子科学与技术 17120036 15650761201 [↑](#footnote-ref-1)