# LeNet-1998 Gradient-Based Learning Applied to Document Recognition 学习笔记

本文主要提出了LeNet-5并在字母识别中取得良好的效果。LetNet-5成功进行多层训练的卷积神经网络CNN。

## 基础知识学习：

卷积：图像与卷积核【或称滤波器】进行像素的加权求和。

卷积的优点：采用**局部连接、权值共享**降低了参数量【需要训练的权值数量只跟卷积核大小、卷积核数量有关，使训练复杂度大大下降】，同时减轻了过拟合。

1.局部连接：每一个像素点在空间上和周围的像素点实际上是有紧密联系的，但是和太远的像素点就不一定有什么关联了。我们把关系紧密的区域称为感受野，因此一个神经元只需将感受野区域的像素点作为输入，而不需要将所有像素点都作为输入，即局部连接。

2.权值共享：一个卷积核就是一个特征提取器，每个卷积核与输入图像进行一次卷积得到的新的图像称为一个Feature Map【卷积核个数越多，提取的特征越多，理论上来说精度也会更高】，一个新的Feature Map的所有像素点都是使用同一个卷积核得到的，即权值共享。

池化-Max pooling：

把下图的4X4输入看作某些特征的集合，数字大意味着特征提取器提取到的特征最明显。最大池化就是保留每个象限提取到的最明显的特征。大量的实验表明最大池化有非常好的效果，它降低了输出参数量、提高了计算速度，同时提高了模型的泛化能力。

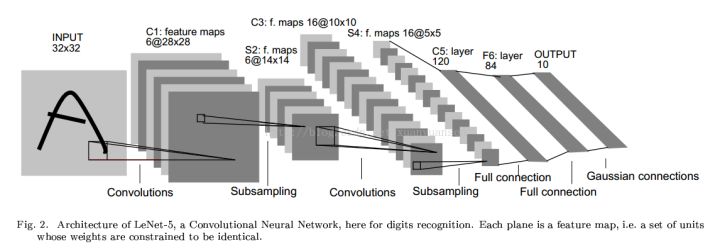
关于CNN具体配置——卷积层，Relu层，池化层具体流程学习参考了：

<https://zhuanlan.zhihu.com/p/27908027?utm_source=qq&utm_medium=social&utm_oi=1014976979178848256>

<https://www.jianshu.com/p/2a0f3a4a9d1d>

**Andrew Ng 机器学习第四五周课程**

在学习完基础知识之后，再回头看论文，主要是letnet-5这部分



**具体过程：（此处学习参考**<https://zhuanlan.zhihu.com/p/34311419>**）**

**输入：**32\*32=1024的手写字体图片，相当于1024个神经元。

**C1层：**作者选择了6个特征卷积核，然后卷积核大小选择5\*5，这样我们可以得到6个特征图，然后每个特征图的大小为32-5+1=28，也就是神经元的个数为6\*28\*28=784。

**S2层：**下采样层，也就是使用最大池化进行下采样，池化的滤波器大小f选择(2,2)，步长stride为2。这样我们可以得到输出大小为14\*14，且有6个这样的图片。

**C3层：**卷积层，这一层我们选择卷积核的大小依旧为5\*5，据此我们可以得到新的图片大小为14-5+1=10，此处采用16个卷积核，所以最终输出16个10\*10的图片。

**S4层：**下采样层，对C3的16张10\*10的图片进行最大池化，池化的滤波器大小f选择(2,2)，步长stride为2。因此最后S4层为16张大小为5\*5的图片。至此我们的神经元个数已经减少为：16\*5\*5=400。

**C5层**：将S4层的输出平铺为一个400的一维向量。然后用这400个神经元构建下一层，C5层有120个神经元。S4层的400个神经元与C5层的每一个神经元相连【C5层有120个神经元】，这就是全连接层，可看作一个标准的神经网络层。

**F6层：**对C5层的120个神经元再添加一个全连接层【F6层含有84个神经元】。

**Outut：**最后将F6层的84个神经元填充到一个SoftMax函数，得到输出长度为10的张量，张量中为1的位置代表所属类别。（例如[0,0,0,1,0,0,0,0,0,0]的张量，1在index=3的位置，故该张量代表的图片属于第三类）

具体计算过程学习Youtube上视频：

<https://www.youtube.com/watch?v=LxfUGhug-iQ&list=PLkt2uSq6rBVctENoVBg1TpCC7OQi31AlC&index=7>

## 实验：另附