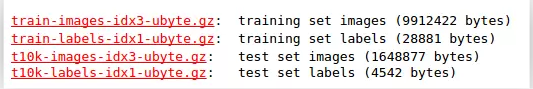
**利用LeNet-5实现手写数字识别**

**一,Minist数据集**

MNIST 数据集来自美国国家标准与技术研究所, National Institute of Standards and Technology (NIST). 训练集 (training set) 由来自 250 个不同人手写的数字构成, 其中 50% 是高中学生, 50% 来自人口普查局 (the Census Bureau) 的工作人员. 测试集(test set) 也是同样比例的手写数字数据.

该数据集由四部分组成 ，也就是一个训练图片集，一个训练标签集，一个测试图片集，一个测试标签集（60,000 个训练样本和 10,000 个测试样本.）



**二，卷积神经网络LeNet-5**

卷积神经网络： 是一种常见的深度学习架构，受生物自然视觉认知机制(动物视觉皮层细胞负责检测光学信号)启发而来，是一种特殊的多层前馈神经网络。它的人工神经元可以响应一部分覆盖范围内的周围单元，对于大型图像处理有出色表现。

一般神经网络VS卷积神经网络：

相同点：卷积神经网络也使用一种反向传播算法(BP)来进行训练

不同点：网络结构不同。卷积神经网络的网络连接具有局部连接、参数共享的特点。（局部连接：是相对于普通神经网络的全连接而言的，是指这一层的某个节点只与上一层的部分节点相连。

参数共享：是指一层中多个节点的连接共享相同的一组参数。）

（一）卷积神经网络的主要组成

卷积层（Convolutional layer），卷积运算的目的是提取输入的不同特征，第一层卷积层可能只能提取一些低级的特征如边 缘、线条和角等层级，更多层的网络能从低级特征中迭代提取更复杂的特征。

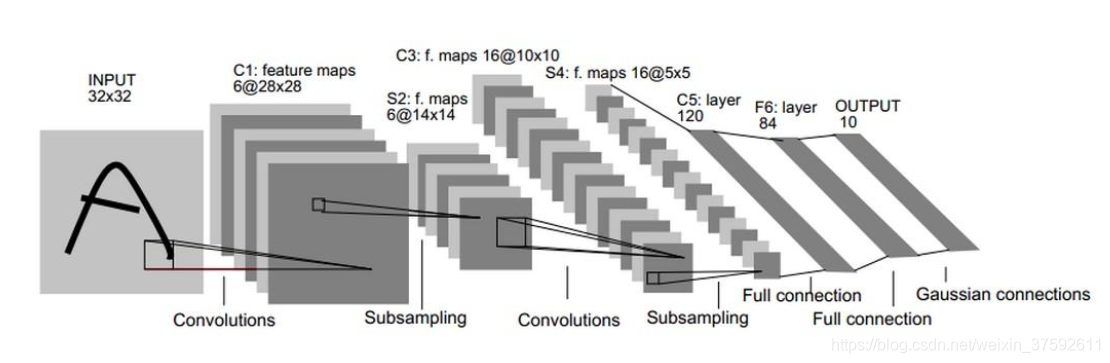
池化层（Pooling），它实际上一种形式的向下采样。有多种不同形式的非线性池化函数，而其中最大池化（Max pooling）和平均采样是最为常见的。pooling层的作用：Pooling层相当于把一张分辨率较高的图片转化为分辨率较低的图片； pooling层可进一步缩小最后全连接层中节点的个数，从而达到减少整个神经网络中参数的目的

全连接层（Full connection）, 与普通神经网络一样的连接方式，一般都在最后几层

（二）Lenet-5卷积神经网络模型

LeNet-5：是Yann LeCun在1998年设计的用于手写数字识别的卷积神经网络，当年美国大多数银行就是用它来识别支票上面的手写数字的，它是早期卷积神经网络中最有代表性的实验系统之一。

LenNet-5共有7层（不包括输入层），每层都包含不同数量的训练参数，主要有2个卷积层、2个下抽样层（池化层）、3个全连接层3种连接方式。如下图所示。



LeNet-5各层结构：

输入： 首先 是 数据输入 INPUT 输入图像的尺寸归一化为32\*32

C1层（卷积层）

第一层卷积层的输入为原始的图像，原始图像的尺寸为32×32×1。卷积层的过滤器尺寸为5×5，深度为6，不使用全0补充，步长为1。由于没有使用全0补充，所以这一层的输出的尺寸为32-5+1=28，深度为6。这一个卷积层总共有5×5×1×6+6=156个参数，其中6为偏置项参数个数，卷积层的参数个数只和过滤器的尺寸，深度以及当前层节点矩阵的深度有关。因为下一层节点矩阵有28×28×6=4704个节点，每个节点和5×5=25个当前层节点相连，所以本层卷积层总共有4704×(25+1)=122304个连接。但是我们只需要学习156个参数，主要是通过权值共享实现的。卷积相对于全连接是稀疏的。优势：1、参数更少 2、计算量降低。

2，S2层（池化层）

这一层的输入为第一层的输出，是一个28×28×6的节点矩阵。本层采用的过滤器大小为2×2，步长为2，所以本层的输出矩阵大小为14×14×6。

3，C3层（第二次卷积）

C3层是一个卷积层，卷积和和C1相同，不同的是C3的每个节点与S2中的多个图相连。 本层的输入矩阵大小为14×14×6，采用的过滤器大小为5×5，深度为16，不使用全0补充，步长为1。这一层的输出的尺寸为14-5+1=10，深度为16，即输出矩阵大小为10×10×16。本层参数有5×5×6×16+16=2416个，连接有10×10×16×(5×5+1)=41600个。

4，S4层（第二个池化层）

S4是pooling层，窗口大小仍然是2\*2，共计16个feature map，C3层的16个10x10的图分别进行以2x2为单位的池化得到16个5x5的特征图。这一层有2x16共32个训练参数，5x5x5x16=2000个连接。连接的方式与S2层类似。

5，C5（全连接层）

本层的输入矩阵大小为5×5×16，在LeNet-5模型的论文中将这一层称为卷积层，但是因为过滤器的大小就是5×5，所以和全连接层没有区别，这里直接看成全连接层。本层输入为5×5×16矩阵，将其拉直为一个长度为5×5×16的向量，即将一个三维矩阵拉直到一维空间以向量的形式表示，这样才可以进入全连接层进行训练。本层的输出节点个数为120，所以总共有5×5×16×120+120=48120个参数

6，F6(全连接层)

F6层是全连接层。本层的输入节点个数为120个，输出节点个数为84个，对应于一个7x12的比特图，-1表示白色，1表示黑色，这样每个符号的比特图的黑白色就对应于一个编码。该层的训练参数和连接数是(120 + 1)x84=10164。

7，Output层（全连接层）

Output层也是全连接层，共有10个节点，分别代表数字0到9，且如果节点i的值为0，则网络识别的结果是数字i。采用的是径向基函数（RBF）的网络连接方式。假设x是上一层的输入，y是RBF的输出，则RBF输出的计算方式是：



上式w\_ij 的值由i的比特图编码确定，i从0到9，j取值从0到7\*12-1。RBF输出的值越接近于0，则越接近于i，即越接近于i的ASCII编码图，表示当前网络输入的识别结果是字符i。该层有84x10=840个参数和连接。

（三）Lenet训练算法步骤

Tensorflow训练Lenet-5模型的卷积网络包括两个过程，为前向传播过程和训练过程。前向传播过程与神经网络的前向传播过程一样通过训练数据的输入经过网络得出实际输出，训练过程就是不断调整网络的输出与实际输出的之间网络的权值和偏值从而使两者之间的相差达到最小。

LeNet-5的训练算法

训练算法与传统的BP算法差不多。主要包括4步，这4步被分为两个阶段：

第一阶段，向前传播阶段：

a）从样本集中取一个样本(X,Yp)，将X输入网络；

b）计算相应的实际输出Op。

在此阶段，信息从输入层经过逐级的变换，传送到输出 层。这个过程也是网络在完成训练后正常运行时执行的过程。在此过程中，网络执行的是计算（实际上就是输入与每层的权值矩阵相点乘，得到最后的输出结果）：

Op=Fn（…（F2（F1（XpW（1））W（2））…）W（n））

第二阶段，向后传播阶段

a）算实际输出Op与相应的理想输出Yp的差；

b）按极小化误差的方法反向传播调整权矩阵。

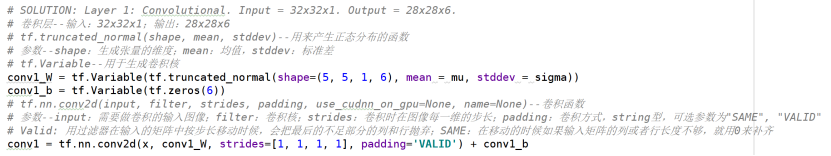
**三，LeNet5详细实现**

设置参数，mu为均值，sigma为标准差

mu=0

sigma=0.1

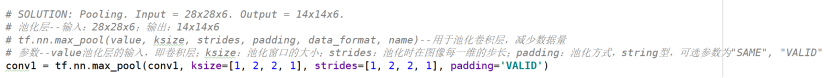
卷积层：对原图像使用5\*5\*6的卷积核进行卷积处理，抛弃用过滤器在输入的矩阵中按步长移动时最后的不足部分



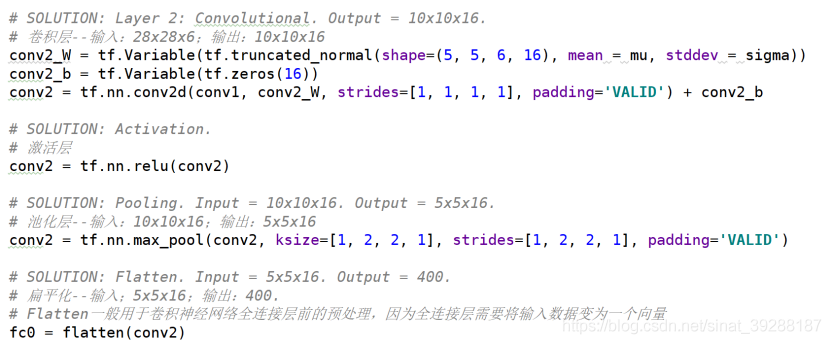
激活：进入非线性特征，将所有负激活变为0



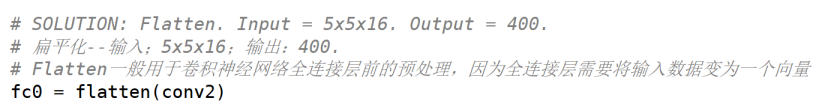
池化层：对卷积生成的图像进行池化处理，池化矩阵尺寸为2\*2



同上再次进行卷积、激活、池化



扁平化：因为全连接层需要将输入数据变为一个向量，所以对数据进行扁平化处理



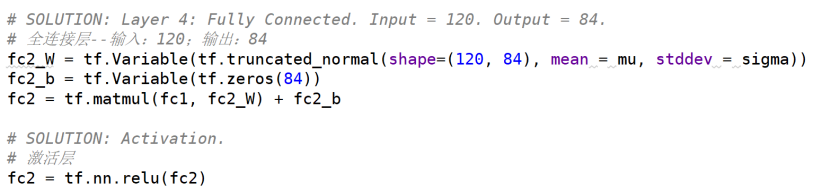
全连接层：类似卷积层处理

第一次全连接：

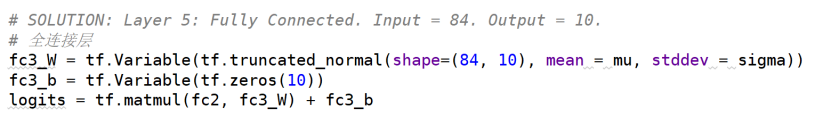
一張含有 文字 的圖片

自動產生的描述

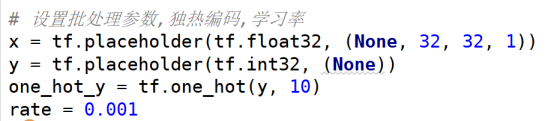
第二次全连接：



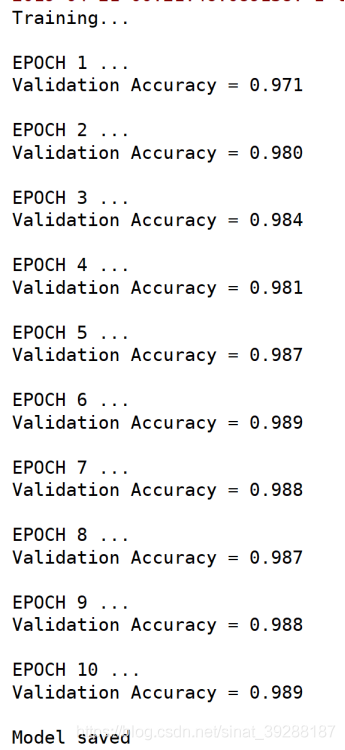
第三次全连接：



设置批处理参数,独热编码,学习率等相关参数



**四，实验结果**





**五，总结**

神经网络权值和偏值的调整与梯度的下降方向和速度有关，用梯度下降算法乘以学习率可以计算下一个定点的位置。当LEARNING\_RATE\_BASE=0.8时，由于初始学习率过大，损失函数的输出值较大说明输出值与实际值差别较大，测试结果准确率较低。当LEARNING\_RATE\_BASE=0.01时，损失函数的输出值较小，说明输出值与实际差别较小，测试结果较为准确。由实验可知，当学习率较大时，神经网络的学习速度较快但是容易导致损失值爆炸，振荡无法拟合，误差大；当学习率较小时，神经网络的学习速度较慢，损失值较小与实际的输出较为接近，但是学习率较小容易导致过渡拟合，无法快速地找到好的下降方向，随着迭代次数的增大损失值基本不变。

本实验中发现该神经网络对minist数据集本身的测试图像识别的准确率较高，如上图6为minist本身自带的图片；可能由于该卷积神经网络的结构设置的不够合理，导致过拟合现象。同时，对比图1和图2，图4和图5发现，在输入图像的背景为黑色时识别的效果较好；原因可能时用于该背景与测试图像的背景相同，同时可能用于实验测试的图像要转化为位图导致图像质量的损失而导致图像的别率较低。对比图3和图4，数字3被识别为7，从图像上看出数字3写的有点类似与7，所以在识别过程中被误认为7。