# 计算机视觉实践-练习2

目录

[计算机视觉实践-练习2实验报告 1](#_Toc7415)

[一、 实验目的 1](#_Toc10363)

[二、 实验原理 1](#_Toc29385)

[三、 实验步骤 3](#_Toc19007)

[四、 数据集 4](#_Toc27115)

[五、 程序代码 4](#_Toc32044)

[六、 实验结果 8](#_Toc32587)

[七、 实验分析与总结 9](#_Toc19770)

### 一、实验目的

1、熟悉卷积神经网络的基本结构，明白LeNet-5由什么组成，各层如何连接。

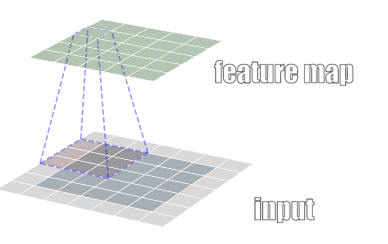
2、实现LeNet-5在MINIST数据集上的训练和测试。

### 二、实验原理

**卷积层：**

卷积层是神经网络中的一种基础计算层，本质是在数据上滑动一个固定大小的小窗口，并将核内元素与对应数据位置的元素相乘，然后对所有乘积求和得到一个标量值，将这个值作为输出数据的对应位置的值。通过对不同位置的卷积操作，能够在输入特征图中提取出不同的特征，如边缘、纹理、形状等。

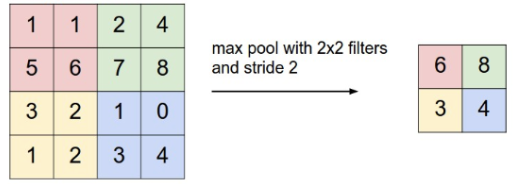
卷积层原理图：

****

**池化层**

池化层是可以用于减小特征图的尺寸，减少后续神经网络的计算量，并提高模型的效率和泛化性能。池化层的操作类似于卷积层，也是在输入的特征图上滑动一个小窗口，但池化层的窗口内不是进行加权求和操作，而是从窗口内选择一个最大或者平均值作为输出

池化层原理图：

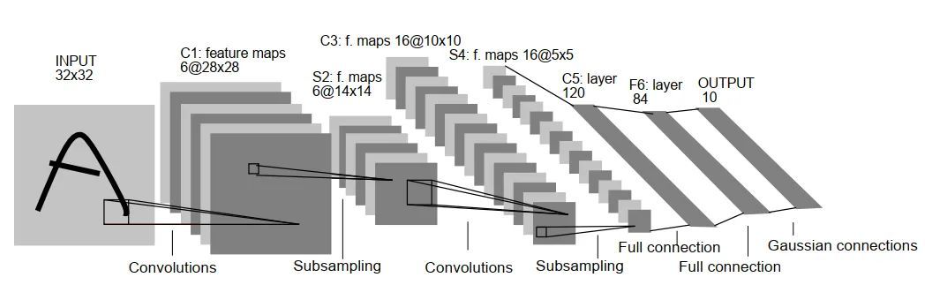


**Sigmoid激活函数**

Sigmoid是一种常用的激活函数，原理是通过一个非线性变换，将输入值映射到0~1之间的取值范围，从而实现对神经元输出的正规化。Sigmoid函数的公式如下：



**LeNet-5网络结构**

****

整个 LeNet-5 网络总共包括7层，分别是：C1、S2、C3、S4、C5、F6、OUTPUT

**输入层**

32x32 像素的图像，通道数为1

**C1层**

C1 层是卷积层，使用 6 个 5×5 大小的卷积核，通过卷积，得到 6 个 28×28 大小的特征图。采用Sigmoid作为激活函数。

**S2层**

S2 层是池化层，使用 6 个 2×2 大小的卷积核进行池化，得到 6 个 14×14 大小的特征图。整个过程相当于降采样层+激活层。先是降采样，然后激活函数 sigmoid 非线性输出。先对 C1 层 2x2 的视野求和，然后进入激活函数。

**C3 层**

C3 层是卷积层，使用 16 个 5×5xn 大小的卷积核，得到 16 个 10×10 大小的特征图，采用Sigmoid作为激活函数。

**S4 层**

S4池化层，使用 16 个 2×2 大小的卷积核进行池化，得到 16 个 5×5 大小的特征图。

**C5 层**

C5 层是卷积层，使用 120 个 5×5x16 大小的卷积核，得到 120 个 1×1 大小的特征图：5-5+1=1。

**F6层**

F6 是全连接层，共有 84 个神经元，每个神经元都与 C5 层的 120 个特征图相连。计算输入向量和权重向量之间的点积，再加上一个偏置，结果通过 sigmoid 函数输出。

**OUTPUT层**

Output 层是全连接层，采用了 RBF 函数计算输入向量和参数向量之间的欧式距离，RBF输出的计算方式是：



### 三、实验步骤

1、下载数据集

2、编写代码，搭建 LeNet-5 神经网络结构

3、MNIST 数据集进行训练

4、保存模型

5、手写图片的测试

### 四、数据集

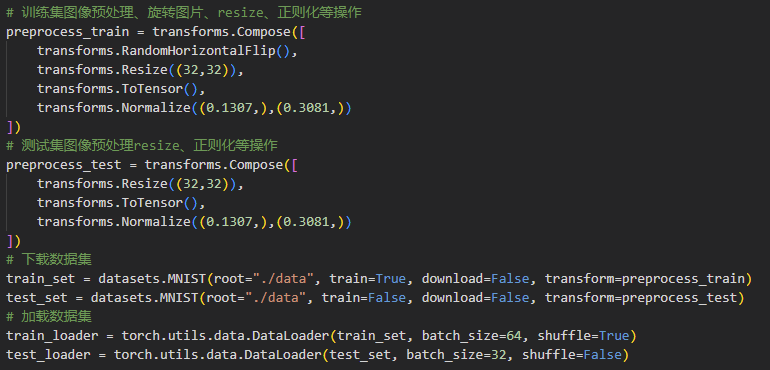
MNIST数据集是一个手写数字识别数据集。该数据集包含了6万张28x28像素的灰度图像和1万张测试图像，涵盖了0~9十个数字的手写图像，每个数字有约6000张图片。

数据集展示

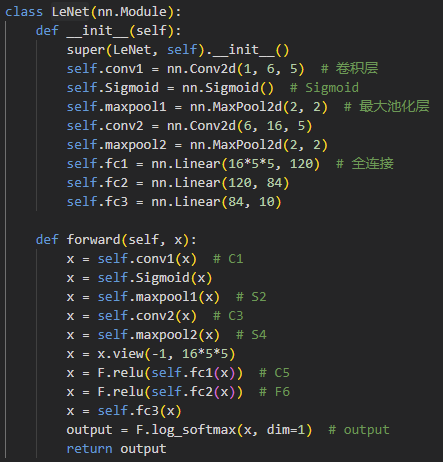


### 五、程序代码

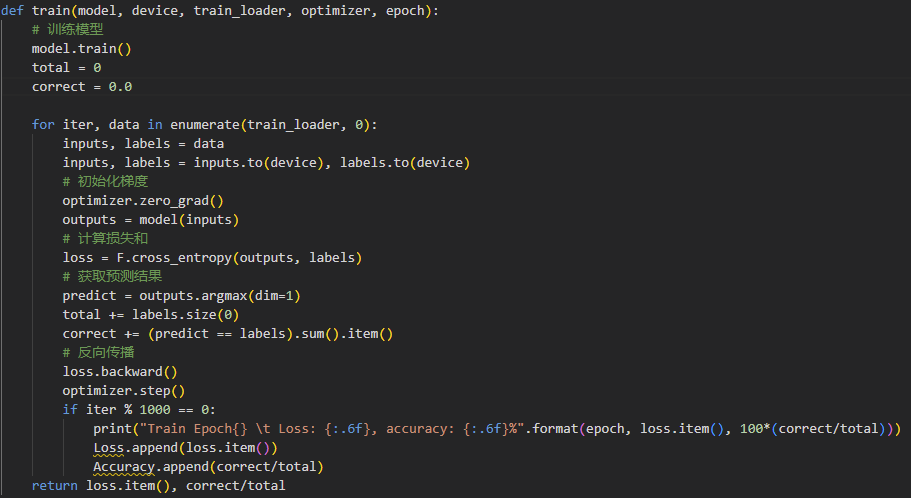
1、加载数据集以及数据集预处理：



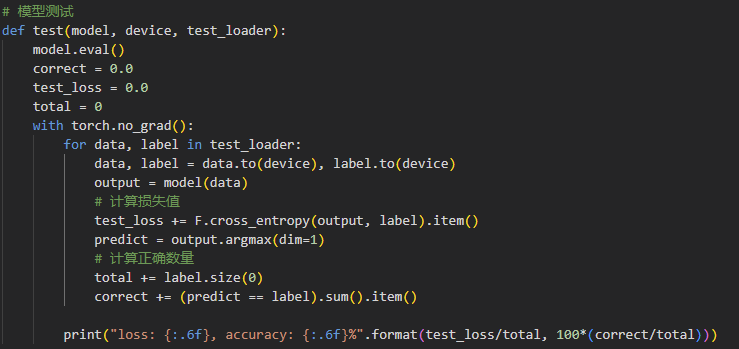
2、搭建LeNe网络结构。



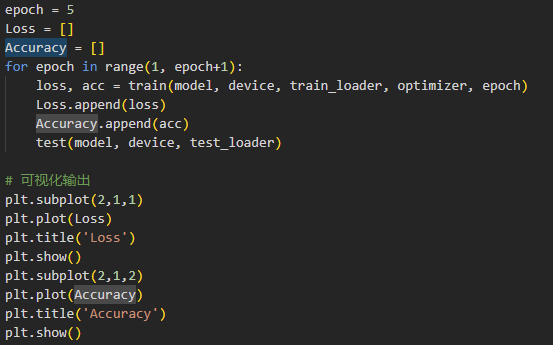
3、定义训练过程



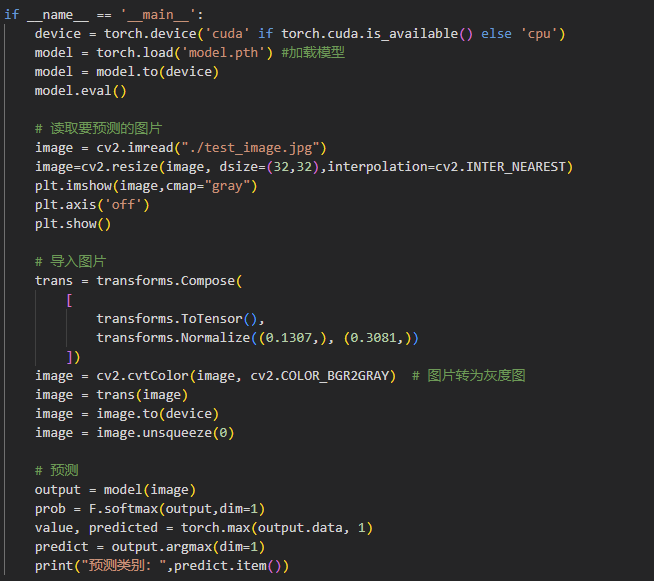
4、定义测试过程



5、输出loss以及accuracy曲线

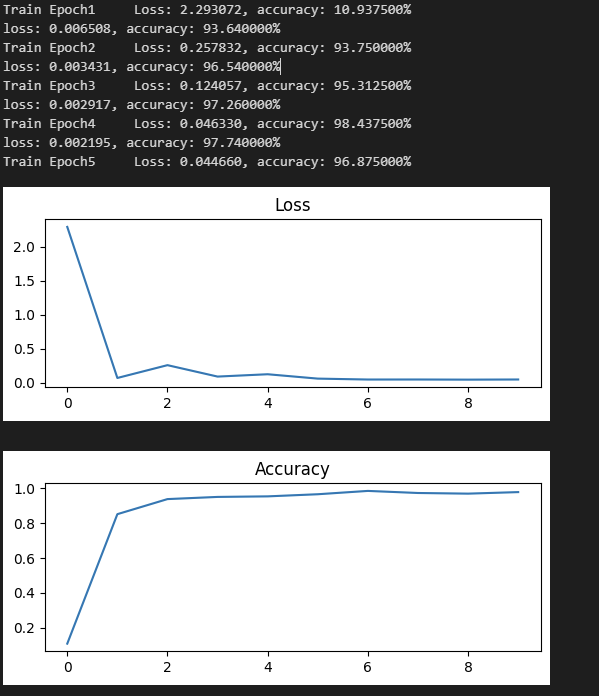


6、最终预测结果

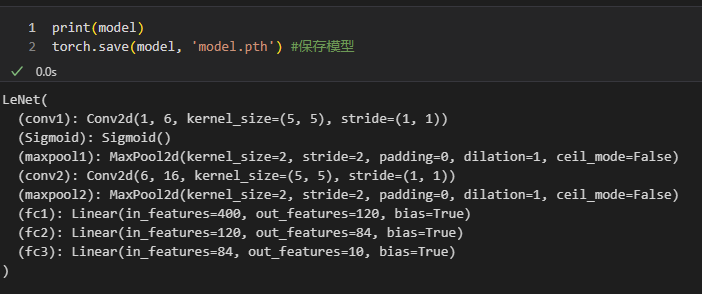


### 六、实验结果

训练迭代过程的loss值和accuracy的变化曲线



模型输出



预测结果



### 七、实验分析与总结

LeNet-5是一种经典的卷积神经网络结构，结构比较简单，但高效。通过多次卷积和下采样，对特征更好的处理，提取出了有效的特征，降低计算量，最终使用这些特征进行分类，可以完成区分简单的手写数字任务。

它的结构包含了卷积层、池化层、全连接层等模块。在本次实验中编写代码，熟悉了整个网络结构框架，对于每一层的功能和原理都有了进一步了解。