# 计算机视觉实践-练习2实验报告

122106222837 张潇

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)

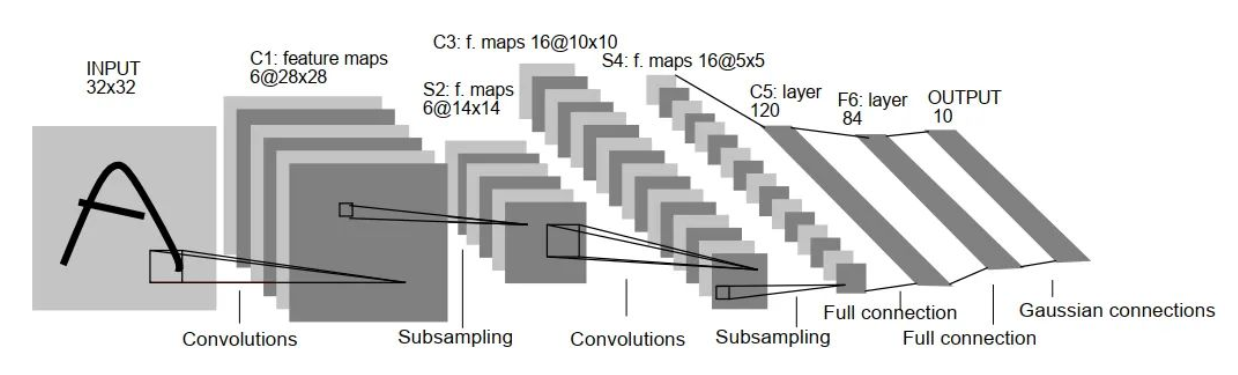
[七、 实验分析与总结 9](#_Toc19770)

# 实验目的

* 熟悉卷积神经网络的基本结构，包括卷积层，池化层，激活函数及最后的全连接层等。
* 实现LeNet-5在MINIST数据集上的训练和测试。

# 实验原理

LetNet-5是一个较简单的卷积[神经网络](https://so.csdn.net/so/search?q=%E7%A5%9E%E7%BB%8F%E7%BD%91%E7%BB%9C&spm=1001.2101.3001.7020" \t "_blank)。输入二维灰度图像，先经过两次卷积层到池化层，再经过全连接层，最后为输出层。整个 LeNet-5 网络总共包括7层（不含输入层），分别是：C1、S2、C3、S4、C5、F6、OUTPUT。网络结构图如下图所示：



**Input**

输入层（INPUT）是 32x32 像素的灰度图像。

**C1**

C1 层是卷积层，使用 6 个 5×5 大小的卷积核，padding=0，stride=1进行卷积，得到 6 个 28×28 大小的特征图。

**S2**

S2 层是下采样层，使用 6 个 2×2 大小的卷积核进行池化，padding=0，stride=2，得到 6 个 14×14 大小的特征图。S2 层其实相当于下采样层+激活层。先是下采样，然后激活函数 sigmoid 非线性输出。

**C3**

C3 层是卷积层，使用 16 个 10×10 大小的卷积核，padding=0，stride=1 进行卷积，得到 16 个 10×10 大小的特征图。

**S4**

S4 是下采样层，使用 16 个 2×2 大小的卷积核进行池化，padding=0，stride=2，得到 16 个 5×5 大小的特征图。

**C5**

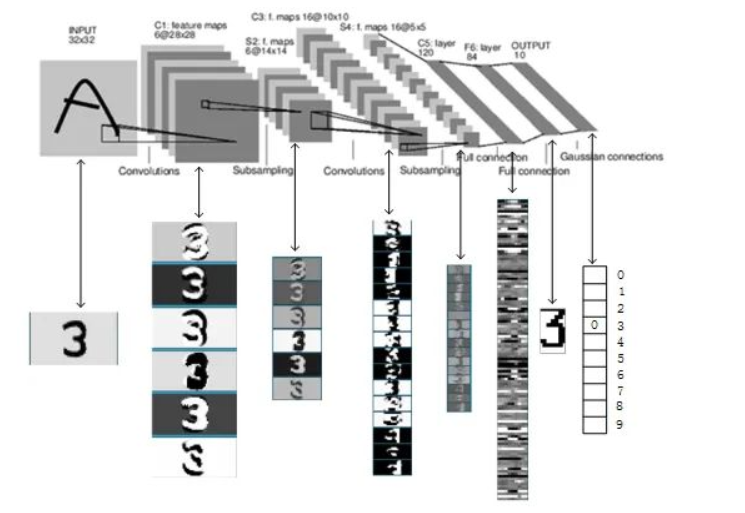
C5 层是卷积层，使用 120 个 5×5×16 大小的卷积核，padding=0，stride=1进行卷积，得到 120 个 1×1 大小的特征图。

**F6**

F6 是全连接层，共有 84 个神经元，与 C5 层进行全连接，即每个神经元都与 C5 层的 120 个特征图相连。使用 sigmoid 激活函数。

**OUTPUT**

最后的 Output 层也是全连接层，产生了10个输出，分别对应了10个类别。整体过程如下图所示：



# 实验步骤

1. 下载并加载数据，并做出一定的预先处理
2. 搭建 LeNet-5 神经网络结构，并定义前向传播的过程
3. **将定义好的网络结构搭载到 GPU/CPU，并定义优化器**
4. **定义训练过程**
5. **定义测试过程**
6. 使用 MNIST 数据集进行训练
7. **保存模型**
8. **手写图片的测试**

# 数据集

[MNIST](https://so.csdn.net/so/search?q=MNIST&spm=1001.2101.3001.7020" \t "_blank)是一个非常有名的手写体数字识别数据集，训练样本：共60000个，其中55000个用于训练，另外5000个用于验证；测试样本：共10000个。MNIST数据集每张图片是单通道的，大小为28x28。部分展示如下：



# 程序代码

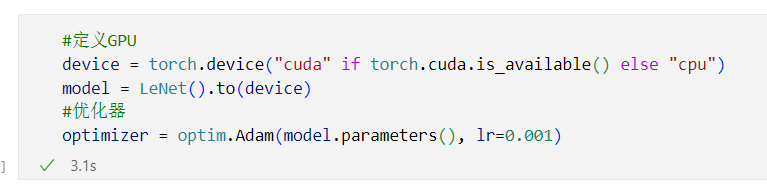
* 首先导入包并加载数据集，如下：



* 搭建LeNet-5网络结构，并定义前向传播的过程。



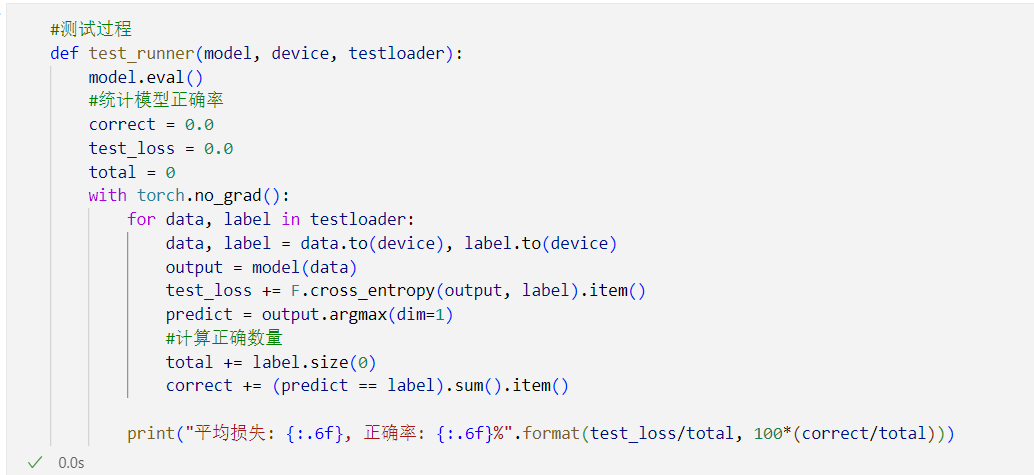
* 定义好的网络结构搭载到GPU/CPU，并定义优化器



* 定义训练过程



* 定义测试过程



* 运行调用



* 预测类别

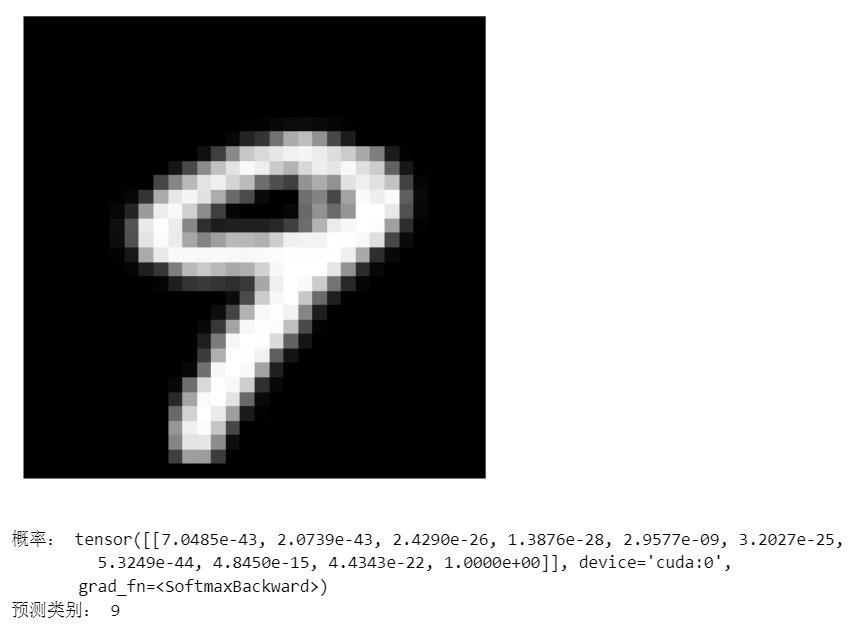


# 实验结果

* 训练结果如下：



* 预测结果如下：



# 实验分析与总结

LeNet-5 是一种用于手写体字符识别的非常高效的卷积神经网络。通过多次卷积和下采样，可以提取出图像中较为重要的信息，空间维度降低的同时，提取出了有效的特征，使用这些特征进行分类，可以区分简单的手写数字。