# 机器视觉实验报告（二）

目录

[机器视觉实验报告（六） 1](#_Toc8990)

[一． 实验目的 1](#_Toc8991)

[二． 实验原理 1](#_Toc8992)

[三． 实验步骤 4](#_Toc8993)

[四． 数据集 4](#_Toc8994)

[五． 程序代码 4](#_Toc8995)

[六． 实验结果 11](#_Toc8996)

[七． 实验分析与总结 13](#_Toc8997)

## 一．实验目的

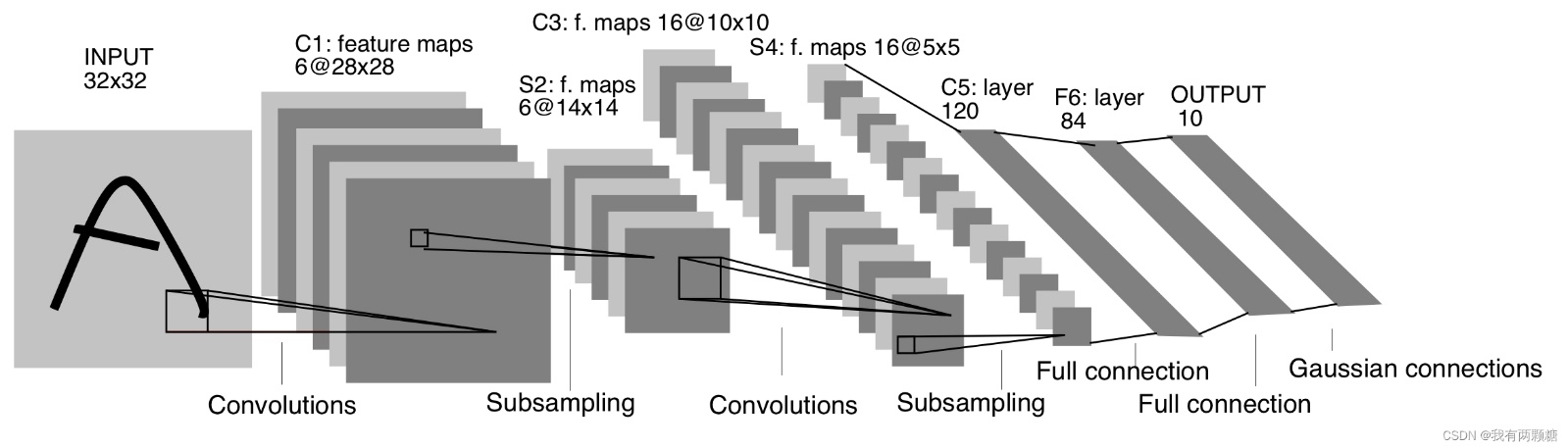
* 熟悉卷积神经网络的基本结构，包括卷积层，池化层，激活函数及最后的全连接层等。
* 学习经典手写数字识别网络LeNet及其他经典特征提取网络。
* 在Mnist数据集上完成图像分类任务，使训练集或测试集的准确率达到98%以上

## 二．实验原理

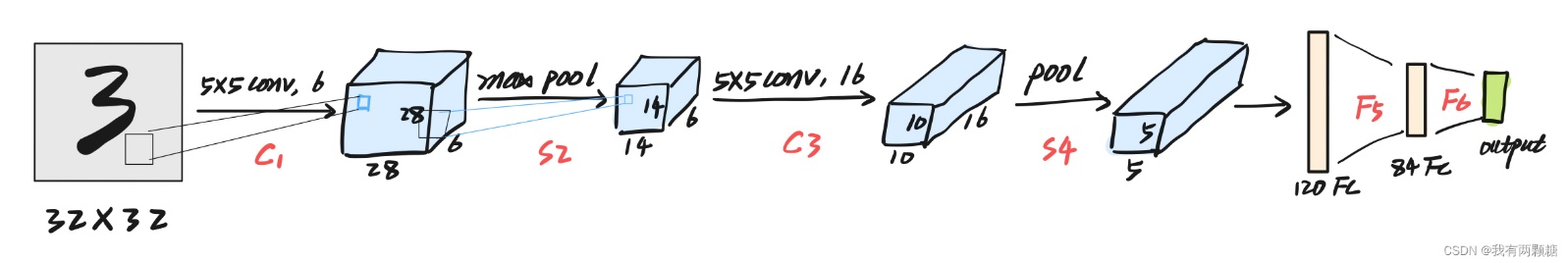
主要包括LeNet5网络。

**2.1** LeNet5

LeNet5 由 LeCun 等人在论文 Gradient-based learning applied to document recognition 中提出，其模型架构图如下：



改进的 LeNet5 模型如下：



特点：先用卷积层来学习图片的空间信息，再用全连接层来转换到类别空间

三 实验步骤

1.首先构建模型框架

2.LetNet5 模型比较简单，接着检查一下模型

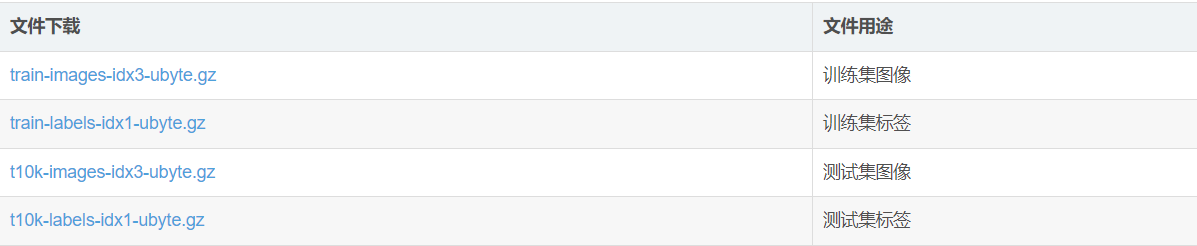
3.使用 datasets.MNIST() 下载数据到本地，并使用 DataLoader 加载数据：

4.为了使用 GPU，首先设置模型和数据的 device:

5. 接着常规操作，先定义损失函数，优化方法选择 Adam，设置train mode，再将数据放到 device，预测模型输出，计算损失，误差反向传播，参数更新 … 最后打印一下当前的 epoch 和 loss 值

6. 使用训练好的模型测试数据，加载模型，读取图片，转换为灰度图，改变大小，设置背景黑色，最后送到模型训练

## 四．数据集

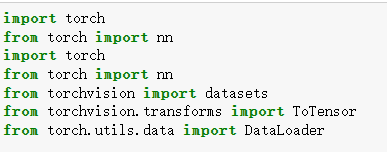


在上述文件中，训练集一共包含了 60,000 张图像和标签，而测试集一共包含了 10,000 张图像和标签。测试集中前5000个来自最初NIST项目的训练集.，后5000个来自最初NIST项目的测试集。前5000个比后5000个要规整，这是因为前5000个数据来自于美国人口普查局的员工，而后5000个来自于大学生。

该数据集自1998年起，被广泛地应用于机器学习和深度学习领域，用来测试算法的效果，例如线性分类器（Linear Classifiers）、K-近邻算法（K-Nearest Neighbors）、支持向量机（SVMs）、神经网络（Neural Nets）、卷积神经网络（Convolutional nets）等等。

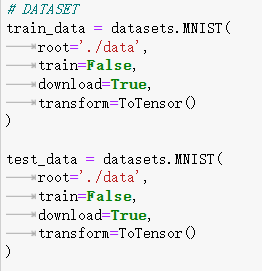
## 五．程序代码

 首先导入需要的模块：



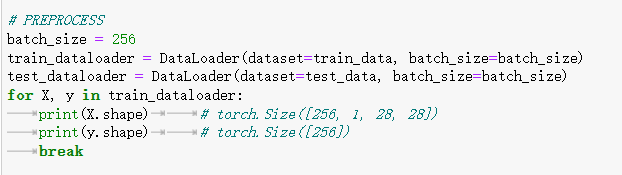
构建模型框架，为了使用 Sequential，代码中自定义了一个 Reshape 模块，将图片 resize 一下

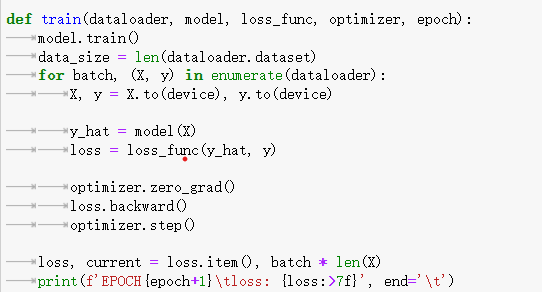
LetNet5 模型比较简单，接着检查一下模型

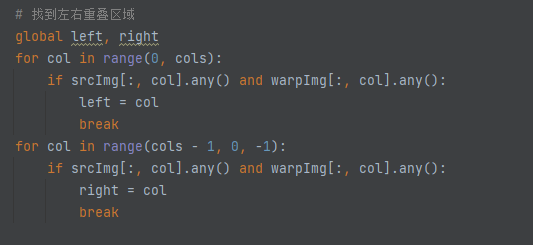
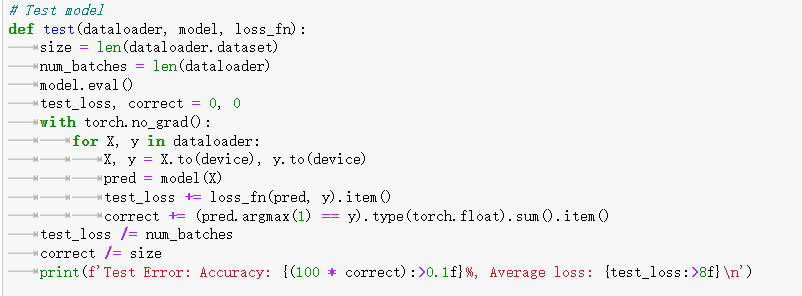


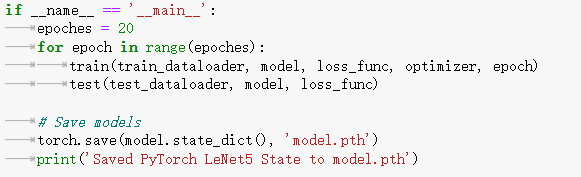
使用 datasets.MNIST() 下载数据到本地，并使用 DataLoader 加载数据：

可以得到每个 batch 大小为 256，batch 中图像为单通道图像，大小为 28 × 28 28 \times 2828×28， 但论文里面输入数据是 32 × 32 32 \times 3232×32， 这也是我们在第一个 CONV 中设置 padding=2 的原因



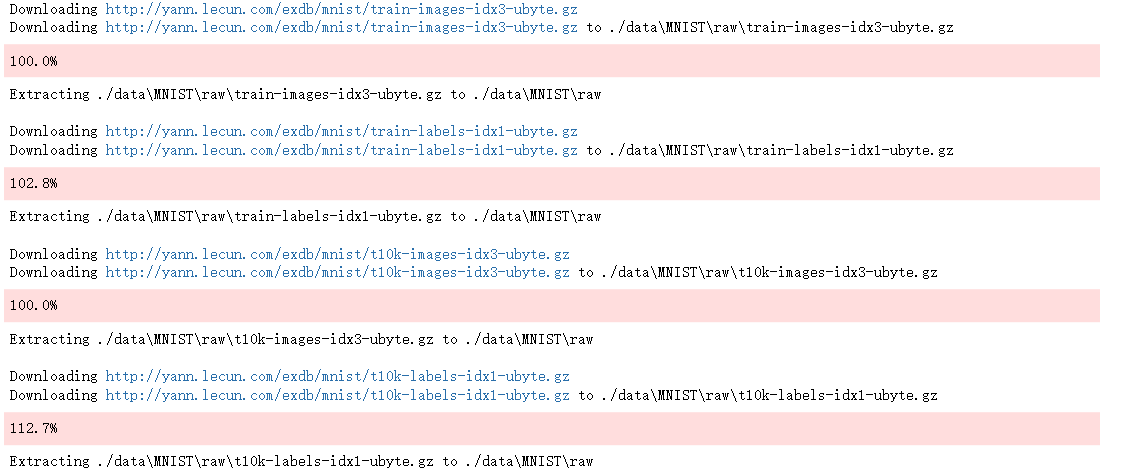


训练代码

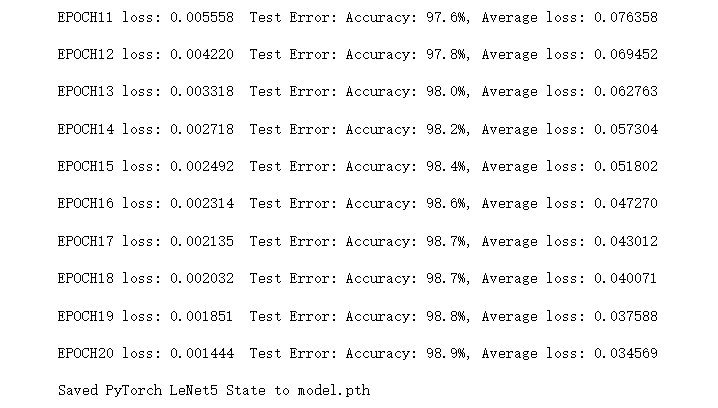


设置轮数二十，开始模型迭代，并且保存训练好的模型

## 六．实验结果



数据集下载



训练结果

## 七．实验分析与总结

本次实验，我们通过LeNet5实现了在Mnist数据集上的训练。



谢谢

老师

(

●

’

◡

’

●

)