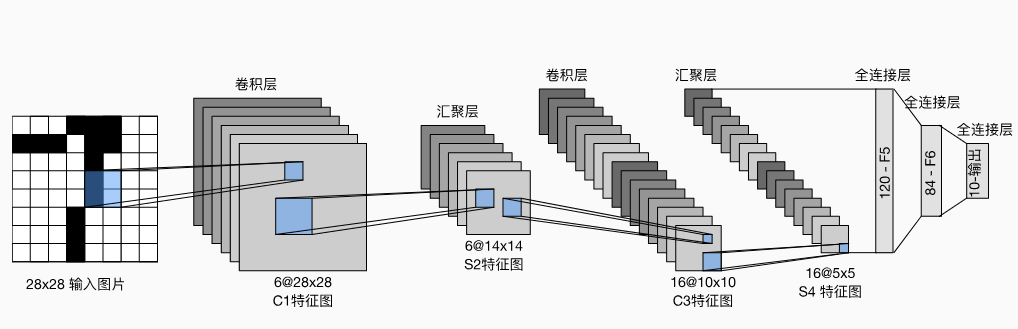
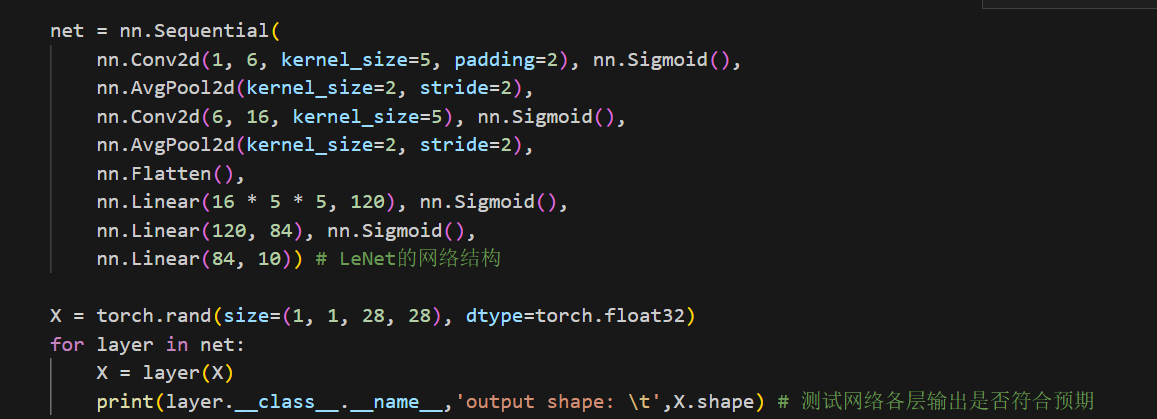
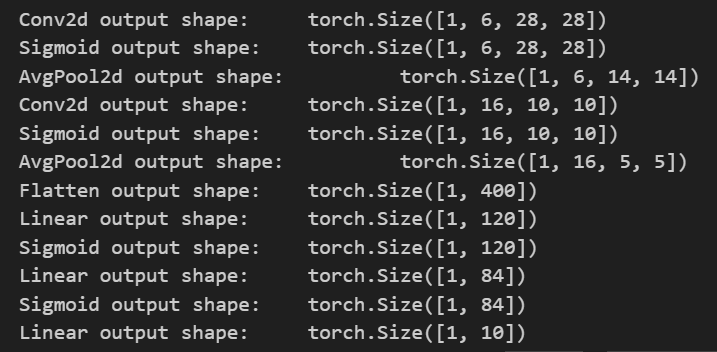
1. LeNet网络架构说明和代码实现

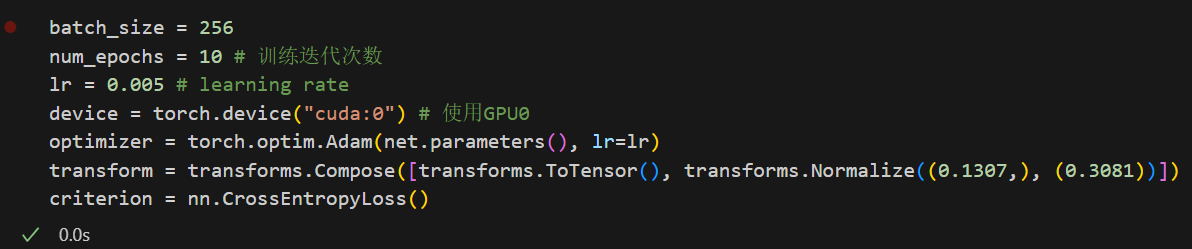


每个卷积块中的基本单元是一个卷积层、一个sigmoid激活函数和平均汇聚层。每个卷积层使用卷积核和一个sigmoid激活函数。这些层将输入映射到多个二维特征输出，通常同时增加通道的数量。第一卷积层有6个输出通道，而第二个卷积层有16个输出通道。每个池操作（步幅2）通过空间下采样将维数减少4倍。卷积的输出形状由批量大小、通道数、高度、宽度决定。为了将卷积块的输出传递给稠密块，必须在小批量中展平每个样本。换言之，将这个四维输入转换成全连接层所期望的二维输入。这里的二维表示的第一个维度索引小批量中的样本，第二个维度给出每个样本的平面向量表示。LeNet的稠密块有三个全连接层，分别有120、84和10个输出。因为在执行分类任务，所以输出层的10维对应于最后输出结果的数量。

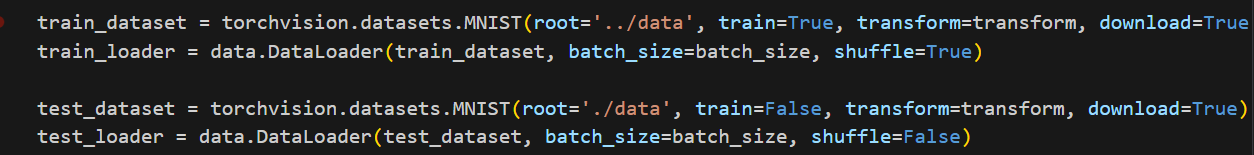




1. 预备工作和数据集读取

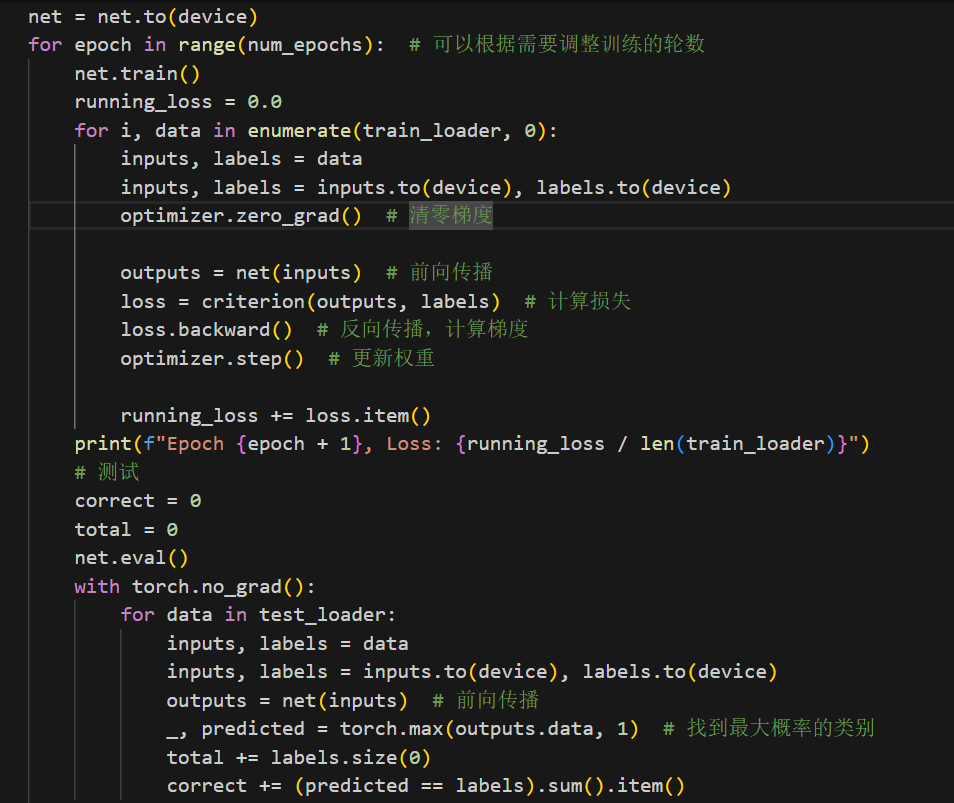


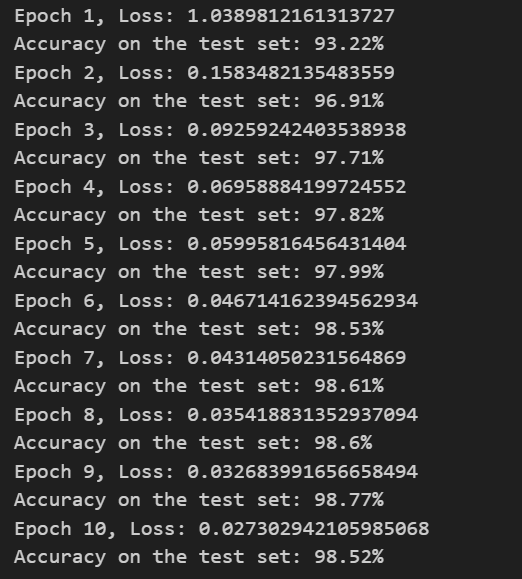
为了使用GPU进行训练，创建device类对象，在这里表示使用编号为0的GPU。



如果本地没有MNIST数据集存在，就从网络上下载。

1. 训练和测试结果





可以看到随着网络的不断训练，在训练集上的损失不断下降，同时网络在测试集上的精度稳步上升，说明网络的结果和对应的超参数设置合理，能够拟合到对应任务上。