**计算机视觉与应用实践报告二**

**姓名：齐浩泽 学号：823104010001**

**1.实验介绍**

计算机视觉与应用实践第二次作业是实现LeNet-5在MNIST数据集上的训练和测试。本次作业是使用华为自研的深度学习框架昇思（**MindSpore**）完成的。

**2.实验原理**

a). LeNet-5网络出自论文Gradient-Based Learning Applied to Document Recognition，原本是一种用于手写体字符识别的非常高效的卷积神经网络，包含了深度学习的基本模块：卷积层，池化层，全连接层。

其网络结构包含如下：

INPUT（输入层） 32∗32的图片。

C1（卷积层）选取6个5∗5卷积核(不包含偏置),得到6个特征图，每个特征图的的一个边为32−5+1=28，也就是神经元的个数由1024减小到了28∗28=784。

S2（池化层）池化层是一个下采样层，输出14∗14∗6的特征图。

C3（卷积层）选取16个卷积核大小为5∗5，得到特征图大小为10∗10∗16。

S4（池化层）窗口大小为2∗2，输出5∗5∗16的特征图。

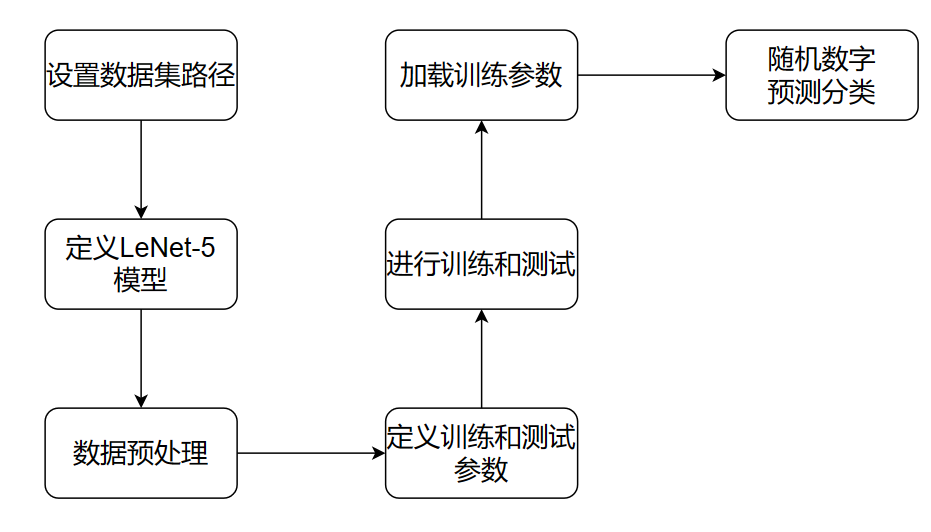
F5（全连接层）120个神经元。

F6（全连接层）84个神经元。

OUTPUT（全连接层）10个神经元，10分类问题。

b). MNIST数据集是由10类28∗28的灰度图片（手写数字0~9）组成，训练数据集包含60000张图片，测试数据集包含10000张图片。

**3.实验过程**

****

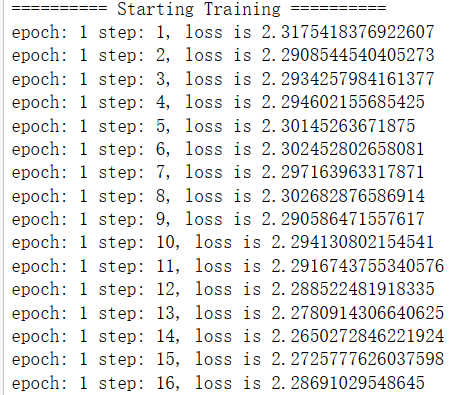
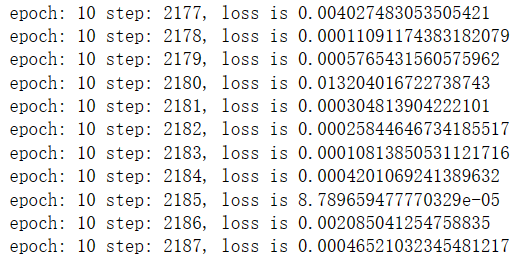
代码部分如下所示：





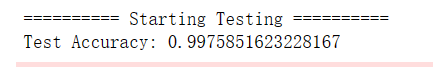


程序运行截图如下：

a).训练过程：

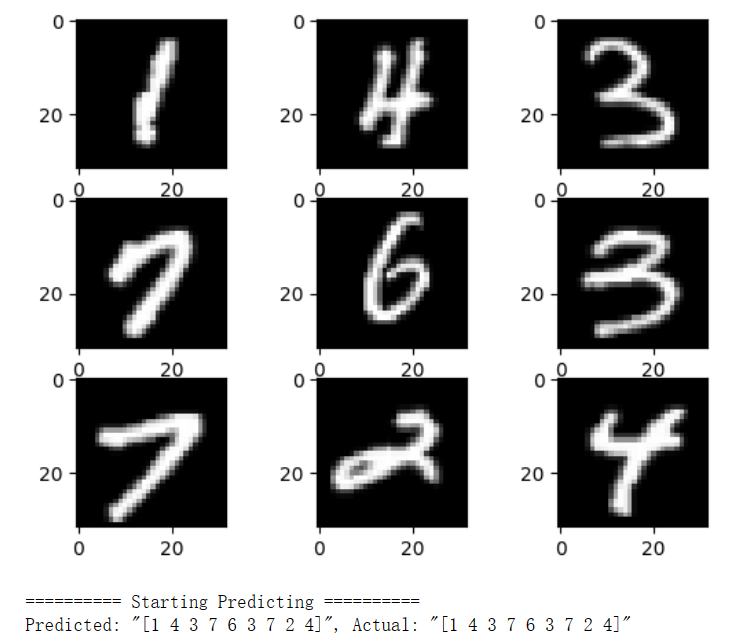
共10个epoch，单次处理32张图像，一个epoch共1875次。可以看到最终训练得到的损失函数值为0.00046。

b).测试过程：



最终训练好的模型对于测试数据集的准确率达到99.76%。

c). 预测过程：



对于随机抽取的9张手写数字图片，该模型全部预测准确。

**4.实验分析**

本次实践作业体现了LeNet-5网络对手写数字识别的高准确率，可能具有以下优势：

1. **专门设计**：LeNet-5是一种专门为手写字符识别任务设计的卷积神经网络（CNN）结构。它由Yann LeCun等人于1998年提出，是深度学习领域的开创性工作之一。
2. **利用图像结构信息**：卷积神经网络能够很好地利用图像的结构信息。LeNet-5通过其卷积层捕捉局部特征，并通过池化层（subsampling layers）降低特征的维度，同时保留重要的空间信息。
3. **参数较少**：LeNet-5的卷积层参数较少，这是因为卷积层具有局部连接和权重共享的特性。这意味着网络可以学习到可以在图像中不同位置通用的特征，而不需要为每个位置单独学习特征。
4. **多层结构**：LeNet-5包含多个卷积层和全连接层，这种结构有助于网络学习复杂的特征表示。随着网络深度的增加，模型能够捕捉到更加抽象和高级的特征。
5. **数据集适配**：LeNet-5被广泛应用于MNIST手写数字识别数据集，该数据集包含了大量的训练样本，这有助于网络学习到丰富的特征和模式。
6. **计算效率高**：由于其较小的模型尺寸和参数数量，LeNet-5在当年的计算资源下也能高效运行，这使得它在实际应用中非常实用。