**计算机视觉实践-练习2\_2**

**一、引言**

Lenet是卷积神经网络的开山之作，由Yann LeCun于上世纪90年代提出，他首次使用了卷积层，池化层这两个全新的神经网络组件，在手写字符识别任务上取得了瞩目的准确率。

Lenet网络有Lenet-1至Lenet-5一系列版本，其中Lenet-5效果最佳。Lenet共7层，由3个卷积层，2个池化层，2个全连接层组成。

图示

描述已自动生成

图1 Lenet-5网络结构

**二、实验任务**

实现Lenet-5在MNIST数据集上的训练和测试，并进行分析。

**三、网络结构**

文本

描述已自动生成

图2 Lenet-5网络结构

本次实验使用的Lenet-5网络结构如图2所示，输入为28×28大小的灰度图。网络一共包括2个卷积层、2个池化层和3个全连接层，其中由于第三层卷积层起到的作用是把特征图展平，每个特征图的大小为1×1，因此使用全连接层代替原始Lenet-5中的第三个卷积层。

通过卷积层，可使得信号特征增强，降低噪声；故其主要作用是降低图像噪声，提取图像重要特征。此外，卷积层能够保证图像的平移不变性，保留图像的空间信息，卷积后的图像位置关系没有改变。

池化层一般参加训练，它的主要作用是减少数据，在降低数据维度的同时保留特征图中重要的特征信息，同时也避免了网络参数太多而造成的过拟合问题。

全连接层一般接在卷积神经网络的最后，用于提取卷积和池化之后的特征向量，并基于提取的特征向量进行图像分类。

原始的Lenet-5网络使用全局平均池化和tanh激活函数，我们现在更倾向于使用最大池化和ReLU激活函数，能够加速收敛并且取得更好的正确率。

**四、实验**

训练：本次实验使用MNIST数据集进行训练，共60000张训练图片和10000张测试图片。在训练之前先使用repeat函数将输入图像转化成空间维度为3×28×28的张量。Batch-size设置为36，训练10个epoch，使用交叉熵损失函数，使用Adam优化器，学习率设置为0.001。

文本

描述已自动生成

图3 训练过程

测试：使用MNIST测试集进行测试，使用训练阶段保存的模型，测试正确率0.989。



图4 测试正确率

**五、分析**

Lenet-5作为卷积神经网络中的开创性工作，提取了三大思想：1.局部感知2.权值共享3.下采样。

因为图像特征分布在图像的像素上，利用卷积操作可以在多个位置提取相类似的特征，于是有了局部感知。另外由于当年并没有计算能力强悍的GPU来辅助训练，因此通过下采样可以有效加快训练和提取更高维的特征，能够节省参数和计算，这与当年的技术相比是一个关键的优势。另外原论文中提到，全卷积不应该放在第一层，图像特征有着高度的空间相关性，因此权值共享可以有效利用图像上的空间相关性。