

**计算机视觉作业报告**

|  |  |
| --- | --- |
| 作业名称： | 利用CNN进行手写数字识别与物体分类 |
| 姓 名： | 刘柯 |
| 学 号： | 12021232 |
| 电子邮箱： | 12021232@zju.edu.cn |
| 联系电话： | 15858367198 |
| 导 师： | 卜佳俊 |

2021年 1 月 15 日

**利用CNN进行手写数字识别与物体分类 实验报告**

1. 作业已实现的功能简述及运行简要说明

在本次实验中，主要实现了两个任务。第一个任务是利用MNIST数据集自创构建一个新的数据集，然后利用LeNet-5在这个新数据集上进行训练得到测试集的识别率。第二个任务是是利用CIFAR-10数据库，增加了一个新的物体类别及其图像，然后利用LeNet-5和ResNet18对这个新数据库进行分类功能训练。

前一个任务的代码对应./mnist\_test文件夹，运行python main.py可以看到新数据集的生成，以及LeNet-5在新数据集上的训练结果。后一个任务对应./cifar\_test文件夹，运行python main.py –test –net=resnet18可以看到训练好的Resnet18在测试集上的效果。

1. 作业的开发与运行环境

开发环境：Mac PyCharm 2020.3.2(professional edition)

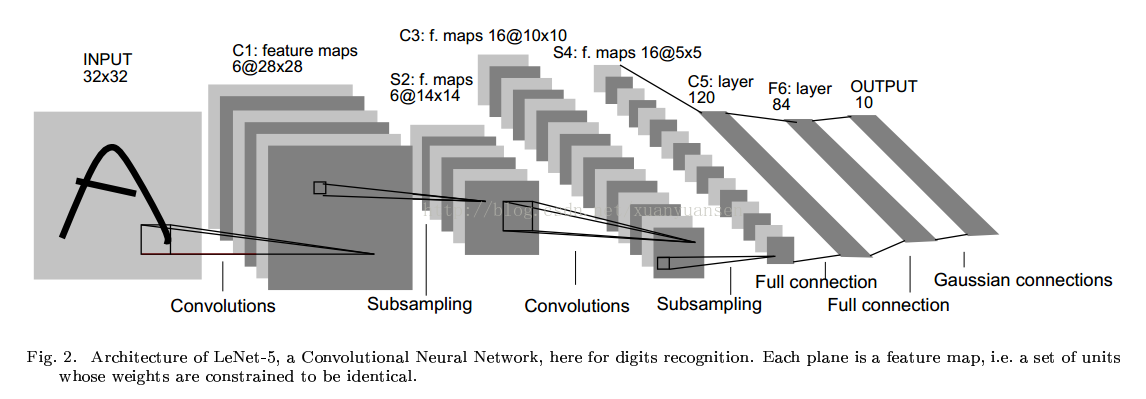
运行环境:Linux python3,torch==1.6.0+cu101, torchsummary==1.5.1, torchvision==0.7.0+cu101, matplotlib==3.2.2, numpy==1.18.5, Pillow==8.1.0

训练GPU：RTX2080

1. 系统或算法的基本思路、原理、流程或步骤等
   1. MNIST任务：

数据集更改方式：把labels为4的图片按主对角线翻转作为背景放在label为8的图片中。

使用的LeNet-5图示如下：



LeNet-5中包含了3个卷积层和2个全连接层，最后得到每一幅图在10个数字labels上的概率，取最大的那个labels作为最后的预测。

* 1. Cifar10任务：

数据集更改方式：从Cifar100中取得所有的flowers和tree的类的图片，定义他们的label为10，加入到cifar10数据集之后。

这个任务用了Lenet-5网络和ResNet18网络，前者的示意图在上面已经贴出，后者的示意图如下：

图片包含 游戏机

描述已自动生成

1. 具体如何实现，例如关键（伪）代码、主要用到函数与算法等

ResNet拟合思路：在普通平原网络的基础上增加了一个shortcut， 构成一个residual block。此时拟合目标就变为F(x)，F(x)就是残差。如果深层网络的后面那些层是恒等映射，那么模型就退化为一个浅层网络。那现在要解决的就是学习恒等映射函数了。 但是直接让一些层去拟合一个潜在的恒等映射函数H(x) = x，比较困难，这可能就是深层网络难以训练的原因。但是，如果把网络设计为H(x) = F(x) + x,如下图。我们可以转换为学习一个残差函数F(x) = H(x) - x. 只要F(x)=0，就构成了一个恒等映射H(x) = x. 而且，拟合残差肯定更加容易。

1. 实验结果与分析
   1. MNIST 任务

新数据集示例:

* 1. Cifar10任务

1. 结论与心得体会
2. 参考文献