

浙江大学实验报告

专业： 计算机科学与技术
姓名： 吴同
学号： 3170104848
日期： 2019 年 12 月 19 日

课程名称： 计算机视觉 指导老师： 宋明黎 电子邮件： wutongcs@zju.edu.cn
实验名称： 利用 CNN 进行手写数字识别 实验类型： 综合型 同组同学： 无

1 实验内容

本实验使用 Tensorflow 框架搭建 LeNet 网络结构，以 MNIST 数据集作训练，进行手写数字的识别。

2 实验环境

- macOS Catalina 10.15
- clang version 11.0.0
- OpenCV version 4.1.0
- CMake version 3.13.4

3 实验原理

LeNet-5 诞生于 1994 年，是一种用于手写体字符识别的卷积神经网络。整个网络由 7 层（不包含输入层）构成，如下图所示：

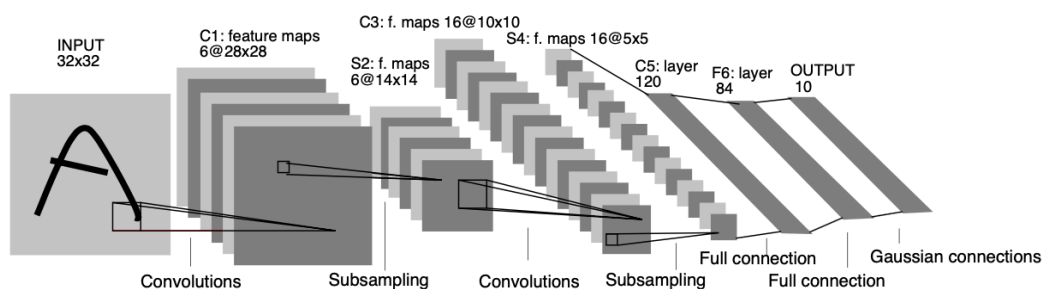


图 1: LeNet-5 的 7 层结构

- 输入层的原始图像大小为 $32 * 32$ 。
- C1 层为卷积层，使用 6 个卷积核，每个核大小为 $5 * 5$ ，得到 6 个特征图，每个特征图的大小为 $28 * 28$ 。
- S2 层为池化层，池化单元大小为 $2 * 2$ ，6 个特征图经池化后大小变为 $14 * 14$ 。
- C3 层为卷积层，使用 16 个卷积核，每个核大小为 $5 * 5$ ，得到 16 个特征图，每个特征图的大小为 $10 * 10$ 。

- S4 层为池化层，池化单元大小为 2×2 ，16 个特征图经池化后大小变为 5×5 。
- C5 层为卷积层，使用 120 个卷积核，每个核大小为 5×5 ，得到 120 个特征图，每个特征图的大小为 1×1 。
- F6 层为全连接层，使用 84 个单元，得到 84 个特征图，每个特征图的大小为 1×1 ，与 C5 层全连接。
- 第 7 层为输出层，也是全连接层，共有 10 个节点，分别代表数字 0 到 9，如果第 i 个节点的值 0，则表示网络识别的结果是数字 i 。

4 实验过程

4.1 模型搭建

在搭建卷积层时，调用 `tf.Variable()` 函数分配 `weight` 和 `bias` 变量，并调用 `tf.nn.conv2d` 搭建卷积层，激活函数使用 `relu`。在搭建池化层时，调用 `tf.nn.max_pool()` 函数。在全连接层中，用矩阵的乘法进行实现。

4.2 模型训练

TensorFlow 中自带从 MNIST 数据集读入数据的函数。读入数据后，将每张 28×28 图片扩展为 32×32 ，即上下左右的边界各加入两行/列空白。调用 `tf.placeholder()` 定义网络的输入输出，定义 softmax 交叉熵，定义 Adam 优化算法。

接下来，进行模型的训练。在每次迭代前，对数据进行打乱，以防过拟合。每次迭代后，计算验证集上的准确率。

4.3 模型测试

训练完成后，得到模型文件。使用测试集的数据进行测试。

5 实验结果

在本实验中，`epochs` 设为 20，`batchSize` 设为 128，使用含有 10000 个样例的测试集，准确率为 99.1%，实验结果达到了较高的准确率。

在每次迭代的过程中，验证集上的准确率变化情况如下：

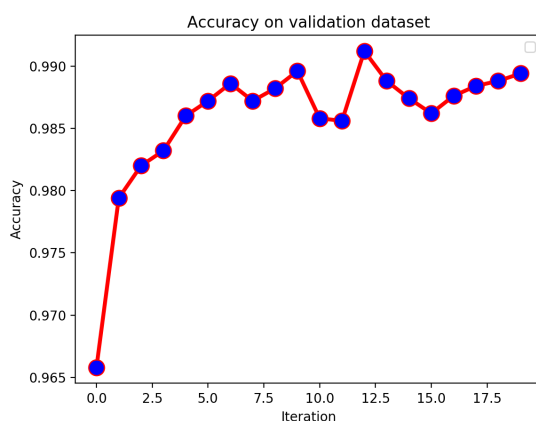


图 2: 验证集上的准确率变化