

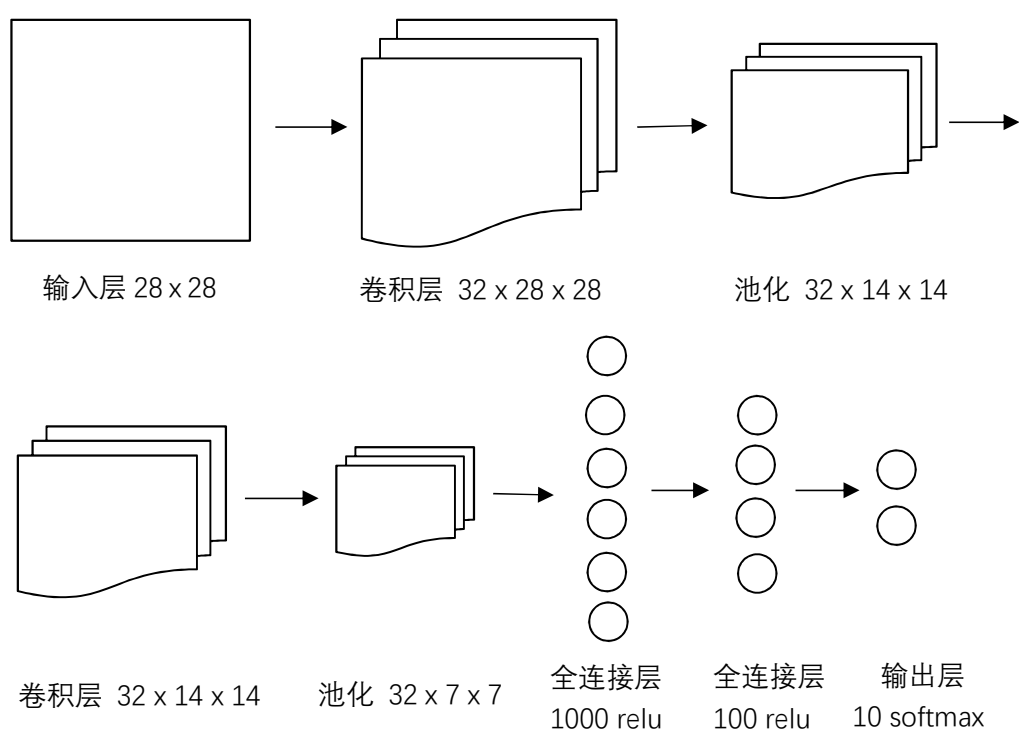
手写数字识别实验报告

计 64 侯林洋 2016011336

Mail:houly16@mails.tsinghua.edu.cn

一、实验模型设计

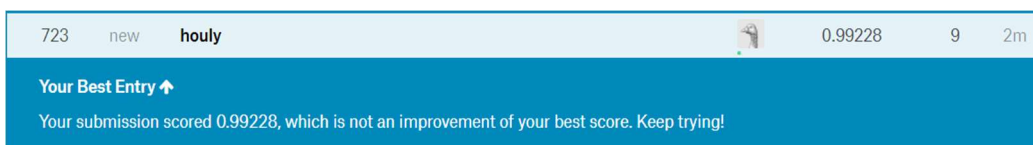
本次实验我借助 keras 深度学习框架实现 CNN 用于手写数字识别任务，设计了两层卷积+池化+两层全连接层的模型，模型结构图如下：



输入层大小为 28×28 ，第一个卷积层有 32 个卷积核，局部感受野大小为 5×5 ，激活函数为 relu，池化采取 2×2 最大值混合，第二个卷积、池化层相同。两个全连接层分别拥有 1000、100 个 relu 神经元，输出层有 10 个 softmax 神经元。

二、实验结果

本次实验我共提交 9 次结果，最佳精度为 0.99228。



三、 参数选择

在实验中，我主要针对学习效率，参数初始化，过拟合三个方面进行了优化。

1) 学习效率

学习速率	评测结果	单次迭代时间
0.0002	0.97528	20s
0.0005	0.99228	15s
0.0007	0.99000	13s
0.0010	0.97714	9s

2) 参数初始化

在实验中采用零均值初始化，高斯分布正交化两种初始化方法，在迭代多次后结果没有明显差别。

3) 过拟合

采用两种方式抗过拟合：

1. 划分验证集，验证集大小为 2000，每训练一轮后进行一次验证，当识别率不再增加时停止训练，采用此方法最终训练次数为 80 轮左右，在测试集上测试结果为 0.99000。
2. 采用弃权函数 dropout，固定迭代次数，经过 300 轮训练后测试结果为 0.99228。

最终程序采用 keras 框架提供的 adam 优化器，lr=0.0005，交叉

熵代价函数。迭代次数 300，批量大小 50。

四、 附件

代码环境为 TensorFlow + Keras + Python2.7 + Ubuntu16.04。