

# Lab1 手写数字识别

## PART 1

本次实验要求同学们利用 MNIST 数据集作为训练数据，训练出一个能够进行手写数字识别的神经网络，并在单独划分的开发集上验证并优化网络的性能。Lab 提交后，TA 会在单独的测试集上对网络的性能进行验证。

【起始时间】 2015.09.16

【最后期限】 2015.10.08 23:59 (PART1)，2015.10.28 23:59 (PART2)

### 【Lab 提交物】

1. 网络代码（语言不限，但禁止使用现成的工具包）
2. 代码说明文档（需要说明网络的结构，网络参数的调整算法，网络的超参数配置，网络超参数的选择过程等信息）
3. 提交至 FTP Work\_Upload 文件夹下。

### 【系统输入】

系统的输入为文本文件，存储一个  $28 \times 28$  的矩阵，每个数表明该点处的灰度值，范围是 0-255；28 行之后单独一行表明该矩阵对应的数字。本次 lab 会提供 70000 个数字灰度矩阵，训练集和开发集的比例自行划分。

## PART1. BP 神经网络

将归一化后的灰度值矩阵转化成向量直接作为系统的输入，训练一个 BP 神经网络用于手写数字识别任务。

## PART2. 卷积神经网络

训练一个卷积神经网络对图片进行特征提取，将图片转化为特征向量后再用于数字识别任务。完成后，比较卷积神经网络与简单 BP 网络的性能差异。（详见 PART2 文档，PART2 文档会在 PART1 结束后给出，包括了卷积神经网络的建议结构，有能力的同学请不要拘泥于文档）

### 【测试方法】

测试集合由选课同学共同产生，即每个人使用画图工具手写 0 到 9 的数字各 10 遍，上传至 ftp 中 Work\_Upload/Lab1/TestSet 下，文件命名为“学号”。

上传文件夹目录树结构如下（不全），其中每个数字文件夹下为对应的数字图片，编号为 0-9，统一采用 bmp 格式，图像大小为  $28 \times 28$  (建议使用画图中 2px 的线，如下图)，图像为白底黑字（TA 会转成灰度图片）。

12302010026

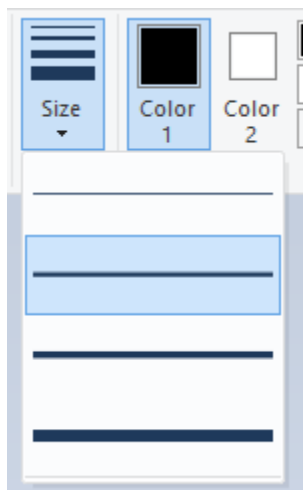
0

1

2

3

(这是目录)



(这是线的粗细)

面试分为两部分，第一部分为算法相关问题提问与回答及实现逻辑检查，第二部分为性能测试。TA 会收取所有人的预测结果并与标准结果比对，得出每个人的预测准确率。

### 【Lab 性能评测方式】

#### 测试数据格式：

test 目录下，共  $n$  个文件，命名为 1.txt 至  $n$ .txt，文件包含一个  $28 * 28$  的灰度矩阵

#### 接口要求：

输入：String 文件目录，int 测试数  $n$

在本地写入一个文件，命名为 “[学号].txt”，文件包含  $n$  行，第  $i$  行，为  $i$ .txt 的对应预测结果。

### 【Lab 评分标准】

Part 1: 60%

Part 2: 40%

For each part:

文档 40% (包含算法原理描述，不同超参数设置对性能影响等)

面试 40% (提问：10% 代码解释：30%)

代码提交 10%

性能 10%

Bonus: 10% (各种 deep learning 的尝试)

### 【备注】

本次 Lab 鼓励同学对网络使用各种优化技巧以提升网络的性能。

准确率占比不会过大，主要是让大家体验实现网络和调整网络参数的过程，请大家不必过分担忧性能问题。

一旦发现抄袭，lab 分数直接为 0，且期末总评-5。

特等奖：lab 中有噪声测试集上精度最高的同学无视其他部分的得分直接满分。

对本次 Lab 有任何疑问，可以联系 TA：

彭浩源 [15212010016@fudan.edu.cn](mailto:15212010016@fudan.edu.cn)

封江涛 [12302010026@fudan.edu.cn](mailto:12302010026@fudan.edu.cn)

### 特殊的写网络技巧：

1. 不要着急实现，避免直接套用公式，仔细推导网络的前向传导和后向反传后进行实现，如果一知半解便进行实现就容易事倍功半；
2. 尽量避免复制粘贴，网络结构很大程度上相似却略有不同，复制粘贴会使得 bug 不容易看出来；
3. 把得到的数据集划分为训练集和开发集，前一部分用于训练调整网络的参数，后一部分用于调整网络超参数；
4. 变量命名要有意义，不然到最后自己都不知道在写什么；
5. 图像初值均值归零，范围控制到(-1, 1)；
6. 网络初值见课上的 PPT
7. 优化技巧提要（鼓励大家进行尝试，不作为评分依据，只会影响精确度）：
  - a) 每次训练对于原始的图片添加很小的随机噪声会增加识别能力；
  - b) Mini-batch；
  - c) 预训练（RBM, AutoEncoder 等）；
  - d) 多个模型进行投票；
  - e) Dropout；
  - f) 不要仅限于以上方法，反思人本身的识别和生物学机理进行优化（也就是通过猜进行优化的意思）；
8. 本次 lab 中 part2 部分难度较大，大家尽早开始动手。

智能系统群内目前仅有 36 位同学，请未加群的同学赶紧加群，群二维码在课程的 material 目录下，TA 会在群内发布通知。