

数字媒体技术作业 4

罗剑杰 数字媒体 15331229

作业内容

手写数字识别, 识别出A4纸上的数字

实验过程

实验环境：

- Ubuntu16.04 LTS
- tensorflow 1.3.0

直接运行 `python main.py` 即可

1. 使用作业1的代码，校正A4纸。
2. 修改作业2的代码，把所有的数字识别并且分割出来，并且把图片转换为和mnist风格相似的图片，存在 `process` 文件夹下。
3. 使用作业3训练出来的模型，用来识别作业4提供的数字

其中，识别的主要过程是：

1. 数据预处理
2. 读数据
3. 预测

预处理的代码嵌入到分割数字的代码中，具体改动的细节如下：

1. 先对A4纸的4周做裁剪，然后统计它的水平方向的直方图，分出3行数字
2. 然后对每一行数字进行识别框出，通过每个框的距离图片左边的距离来对每个数字的图片排好序
3. 将每个数字的图片处理成mnist格式的图片
 1. 图片反转，变成黑底白字
 2. 用最近邻把长边放缩至28，再补短边至28，形成一个28*28的图片

读取数据的时候，将每张图片读入，然后将它reshape成1x784的向量，然后17个数字所对应的17个向量一个一个添加到 `test_data` 中。使用作业3训练好的模型的参数构建出一个model，将 `test_data` 传入进行预测，即可得到结果。

实验结果

实验结果截图如下：

```
2018-01-20 19:47:19.1
eForce 820M, pci bus
[1 2 7 2 6 7]
[7 7 6 5 2 8 1 7]
[6 6 6]
(tensorflow) longj@l
```

可以看到，其中3，9，4的识别出现了问题，没有办法正常输出。和同学们交流后发现都存在这种情况。思考原因：**TA**给出的输入图片中的手写的**3，4，9**和**mnist**中的出入挺大，导致没有办法正常识别出来，其中做过实验来证实这个假设：通过将TA写的数字分割放入训练集中，出来的模型可以将该input的所有数字正确识别，但是用这个模型去检测tensorflow中mnist的测试集，准确率从99%降到92%，说明出现了过拟合。所以的话3，4，9识别不出属于正常的情况。