**机器学习框架 第三次竞赛 说明文档 蒋晓伟2017011850**

通过分析题目要求，本次竞赛的任务是能够实现对验证码进行识别。计划采用的算法是：使用深度学习的卷积神经网络（CNN）算法实现所给数据集模糊验证码图像的识别。

初步想尝试去过验证码全部都是数字的情况下，使用降噪处理+图像分割+数字识别+字符拼接组合的方法得到五位验证码，采用老师所给资料中的minist数据集在kaggle平台上提供免费的gpu进行训练，使用从其他网站上下载的验证码作为训练集训练，准确率达到了80%以上，我认为可以添加进行字母的识别了。考虑到字母分为大写字母和小写字母，通过搜集相关资料发现这个想法可能不是很好，于是尝试换方法。



结合《数字图像处理》这门课以及实训项目，我发现使用CNN卷积神经网络和全连接效果应该是最好的。其中，数字+大写字母+小写字母=62个字符。所以，向量长度范围是5\*（10+26+26）=310。文本转向量：通过某种规则(char2pos），计算字符数值，然后根据该字符在4个字符中的位置，计算向量索引idx = i \* CHAR\_SET\_LEN + char2pos(c)

向量转文本：跟文本转向量操作相反（vec2text）。通过构想，前向传播组成为3个卷积层（3\*3\*1\*32，3\*3\*32\*64，3\*3\*64\*64），3个池化层，4个dropout防过拟合层，2个全连接（(8\*20\*64,1024）(1024\*MAX\_CAPTCHA\*CHAR\_SET\_LEN])）4个Relu激活函数。反向传播组成：计算损失（sigmoid交叉熵），计算梯度，目标预测，计算准确率，参数更新。卷积过程中，通过报错信息知道要求相乘的两个矩阵宽、高必须相同，全连接中，要第一个矩阵的列和第二个矩阵的行必须相同。通过tensorflow中的参数保存功能获取保存对象，调用保存方法。其中，函数gen\_captcha\_text\_and\_image（）用于读取图片的名称获取真实值和将图片信息转化为二维数组。函数text2vec（）和vectext（）用于文本和向量之间的互转，函数get\_next\_batch（）用于每次进行下一次迭代，函数crack\_captcha\_cnn（）用于构建定义3层卷积神经网络，卷积 + Relu激活函数， 池化，dropout 防止过拟合，返回输出结果。

函数train\_crack\_captcha\_cnn（）用于对训练集的训练并保存数据模型。函数crack\_captcha（）用于调用之前惨老的模型用测试集训练。

通过pandas.DataFrame()方法把预测结果输出到csv文件中。整理格式加上列号，提交。