# 第十周报告

校园一卡通图片识别程序设计（5）

字符识别（分类）

一、实验方法与期望

在上一周尝试使用Tesseract-OCR模块进行字符识别获得了不错的识别效果。本周尝试使用卷积神经网络的方法进行。考虑到数据集不易获得，本周主要研究二分类，即对“姓”和“名”字符进行分类识别。

本周的报告主要包括：字符区域分割成字符块、数据集的产生、搭建CNN网络对分割得到的字符块进行识别。

二、字符区域的分割

上周我们是直接将字符区域输入到Tesseract-OCR模块中进行识别，如果想使用CNN的方法，我们首先需要做的工作是把图1中的字符区域进行切割，获得字符块，即待识别的类型。

使用投影分析法，得到水平投影和竖直投影，如图2、图3，获得每个字符区域的轮廓如图4，切割出来字符，如图5。

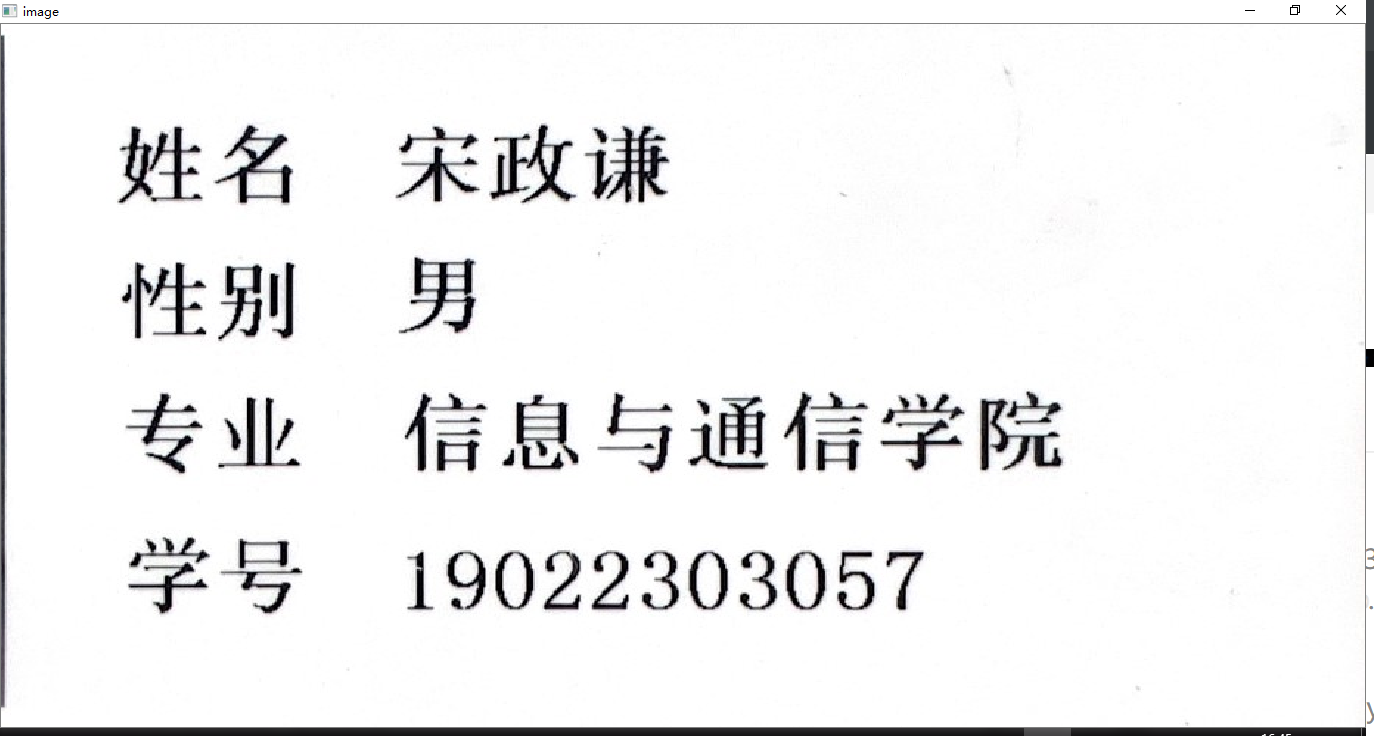


图1 字符区域

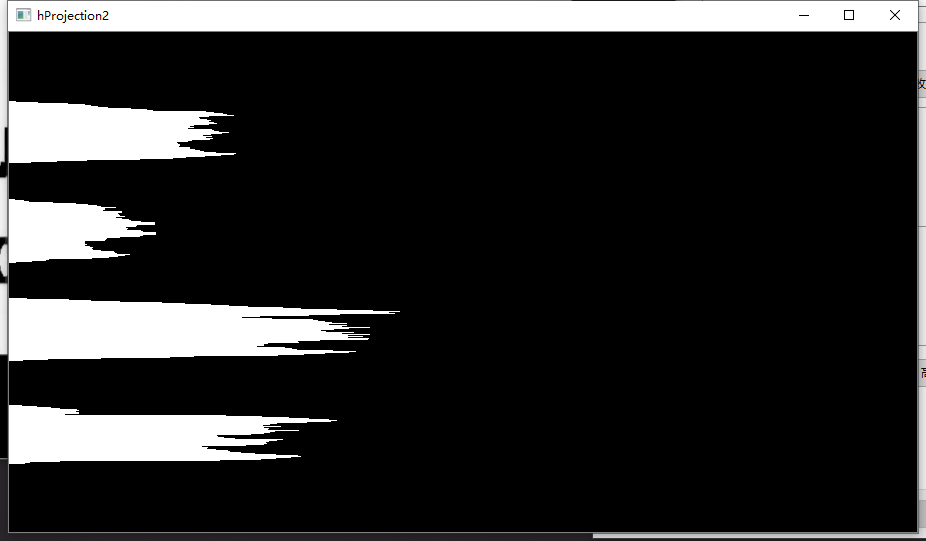


图2 字符水平投影



图3 字符竖直投影

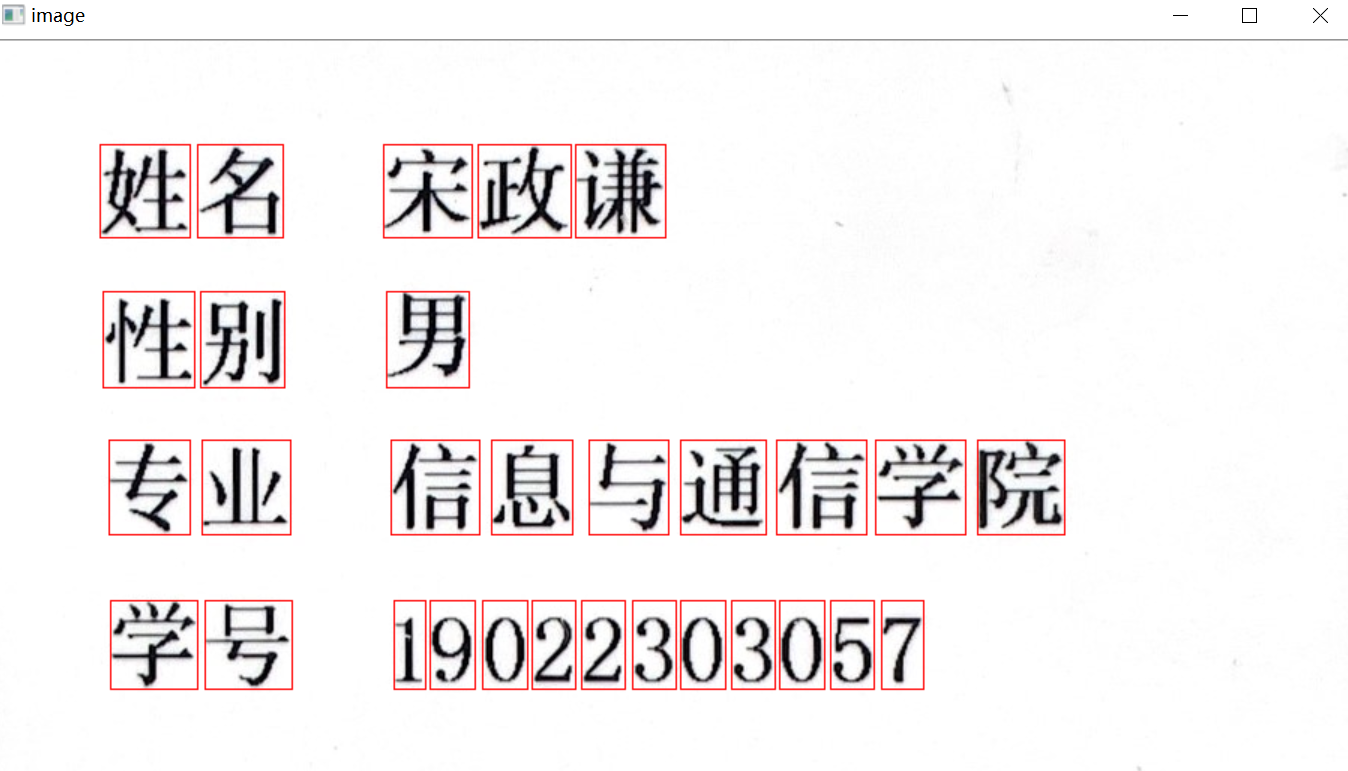


图4 获得字符轮廓





图5 切割得到的字符块

三、数据集的产生

使用程序产生与待识别字符字体相似的汉字字符块。本次主要是产生“姓”和“名”字符块。产生的图片被分为训练集（每个类100张图片）和测试集（每个类50张图片）。其中“姓”被标注为0，“名”被标注为1。

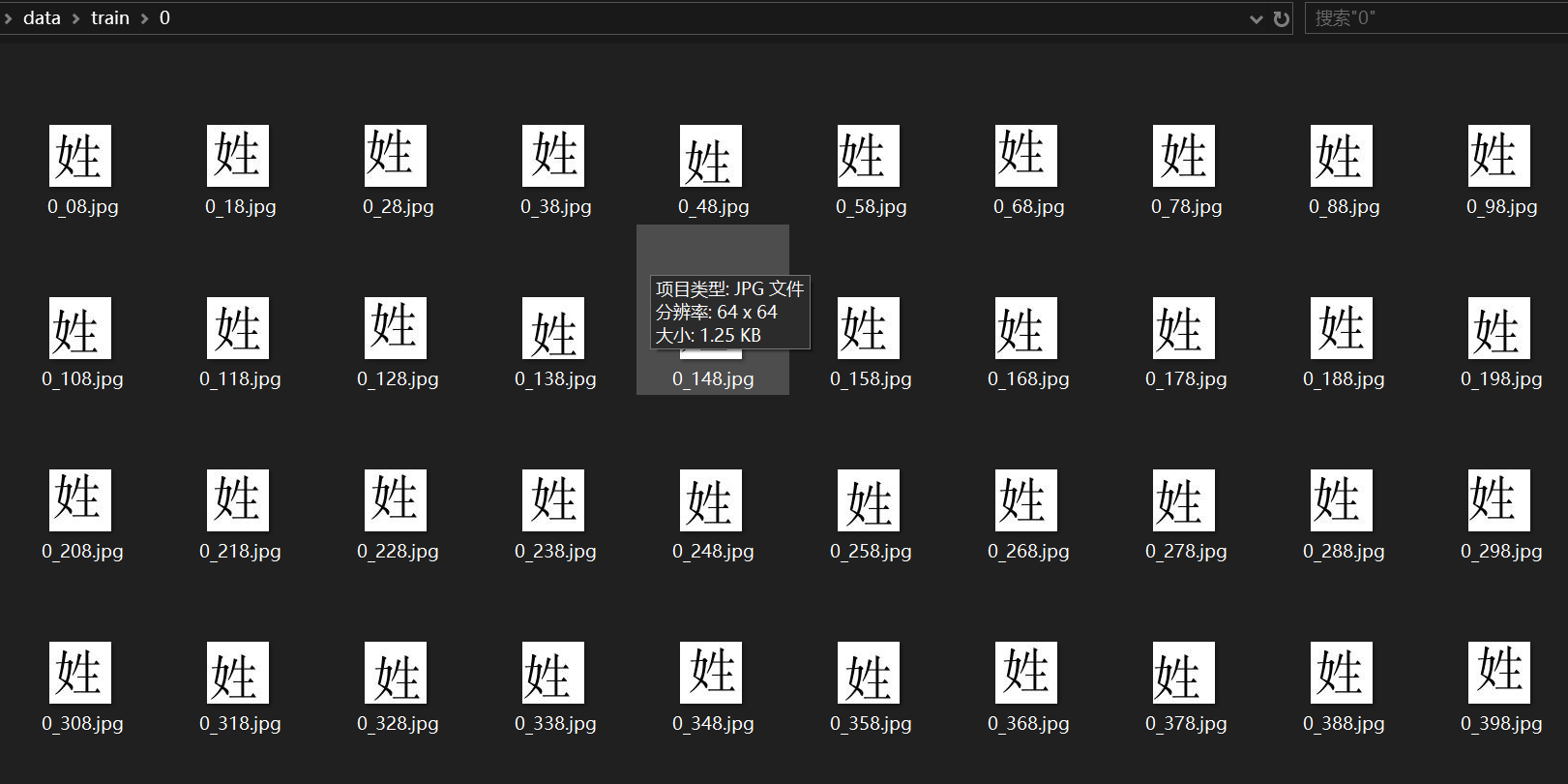


图 6 部分数据集展示

四、搭建网络进行图片分类

至此，我们有了数据集和待识别的图片。

使用Keras框架可以快速的搭建卷积神经网络。借助Keras中的ImageDataGenerator类中的flow\_from\_directory()函数将训练图片读取。输入搭建好的卷积神经网络中进行训练。训练结束之后将图5中待识别的“姓”和“名”输入模型进行预测。输出预测结果。

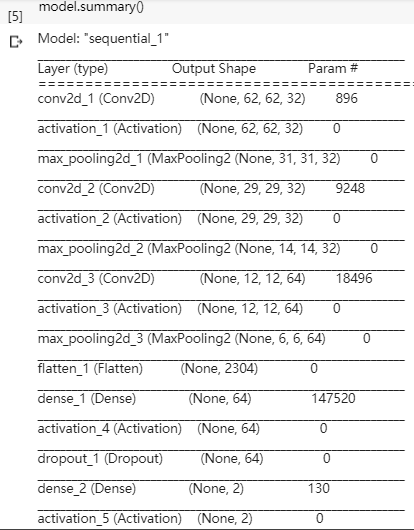


图7 网络模型

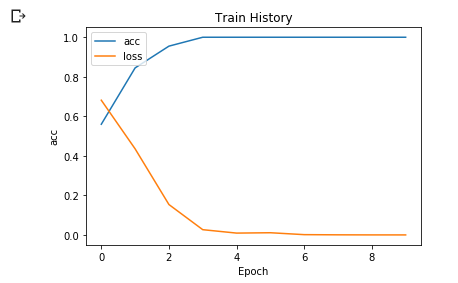


图 8 训练过程示意图

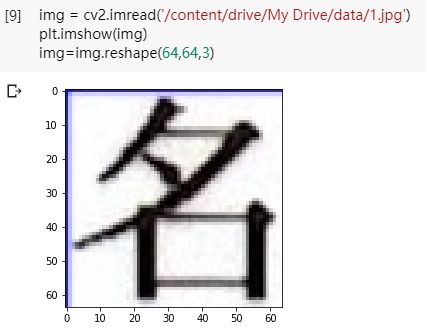


图 9 读取待识别图片

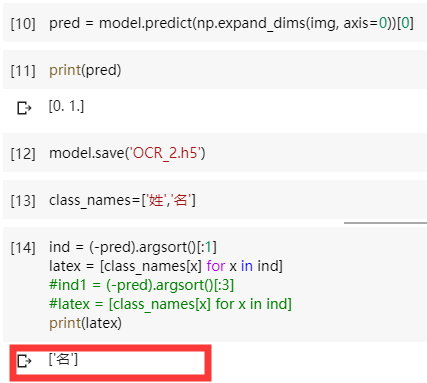


图10 输出预测结果

五、分析

这次实验对于主要目的是对于“姓”、“名”图片进行二分类，训练集是每个类100张图片，卷积神经网络使用3层的卷积层和池化层相连接的经典结构，之后进行flatten和全连接，并加入了dropout层，网络结构完整。可能由于数据量较小和类别较少的原因，在训练5个epoch时就达到了较好的效果（如图8）。

在之后应当对网络结构进行调整，选择最适合实验需求的网络结构。调整网络参数，使网络更加快速的收敛。

报告人：宋政谦

时间：2019/11/10