大连理工大学计算机学院

程序设计训练实验报告

|  |  |
| --- | --- |
| 实验名称 ： |  |
| 学生姓名 ： |  |
| 学生学号 ： |  |
| 联系方式 ： |  |
|  |  |
|  |  |

基于各种学习的物体识别工具箱

# 实验简介

随着机器学习和计算机视觉的快速发展，目前各种模式识别算法的都已经很成熟了，从机器学习的各种经典算法到深度学习的AlexNet、GooGleNet。借这次大作业的机会，我们组一致决定一起学习相关的机器学习、模式识别算法。

（我编不下去了，靠大家了）

我们组的分工如下：

实验成果如下：



Figure 主界面



Figure 车牌识别

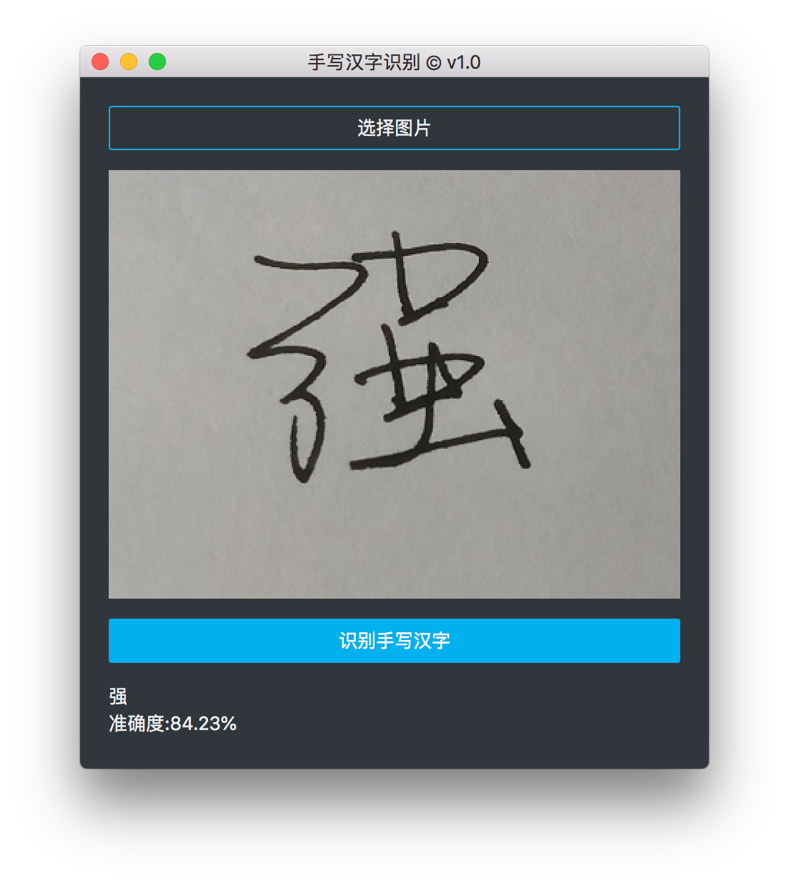


Figure 手写汉字识别



Figure 印刷文字识别

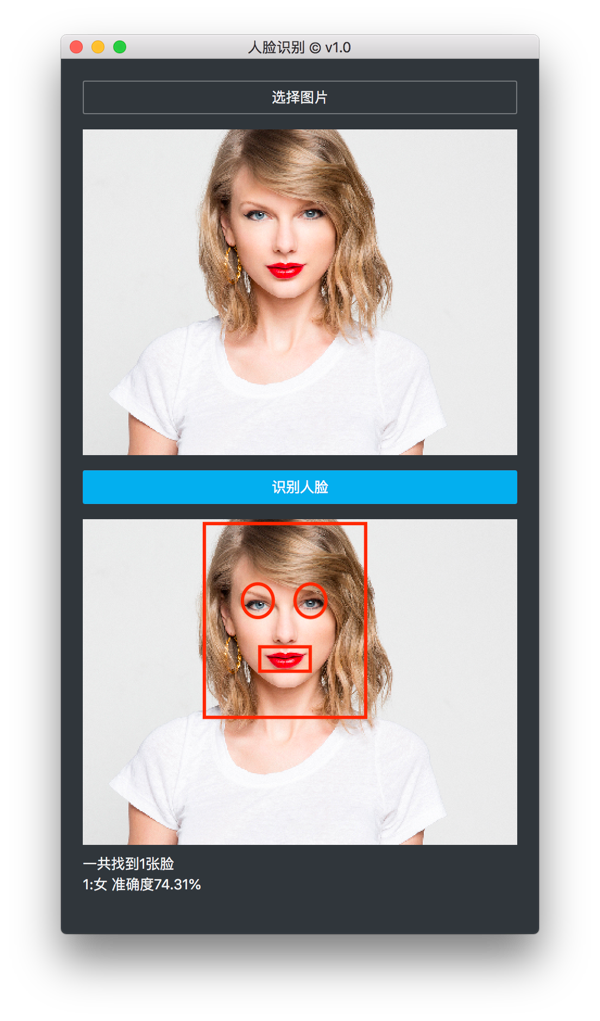


Figure 人脸识别

# 实验目的

1. 了解GoogLeNet的原理，掌握其基本使用方法

# 实验相关原理与技术

1. 车牌识别
2. 光学字体识别/印刷字体识别
3. 手写汉字识别

GoogLeNet/Inception是2014年Christian Szegedy提出的一种全新的深度学习结构，在这之前的AlexNet、VGG等结构都是通过增大网络的深度（层数）来获得更好的训练效果，但层数的增加会带来很多负作用，比如过拟合、梯度消失、梯度爆炸等。GoogLeNet的提出则从另一种角度来提升训练结果：更高效的利用计算资源，在相同的计算量下能提取到更多的特征，从而提升训练结果。

GoogLeNet的inception，结构如下图所示。整个GoogLeNet是由多个下述模块串联而成。主要作用在于使用1x1的卷积来进行升、降维和在多个尺寸上同时进行卷积再聚合。

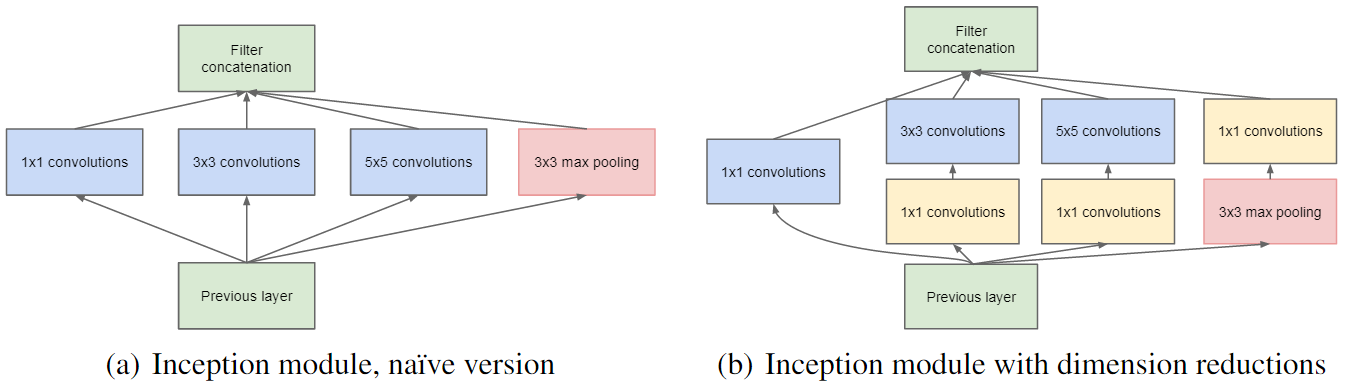


Figure GoogLeNet/Inception module

在NIN（Network inNetwork）的结构中是利用MLP代替传统的线性卷积核，从而提高网络的表达能力。文中同时利用了跨通道pooling的角度解释，认为文中提出的MLP其实等价于在传统卷积核后面接cccp层，从而实现多个feature map的线性组合，实现跨通道的信息整合。而cccp层是等价于1×1卷积的，因此细看NIN的caffe实现，就是在每个传统卷积层后面接了两个cccp层（其实就是接了两个1×1的卷积层）。用上图的模块举个例子，假设输入的feature map是28×28×192，模块中1×1卷积通道为64，3×3卷积通道为128,5×5卷积通道为32，那么对3×3和5×5卷积层前分别加入了通道数为96和16的1×1卷积层，这样卷积核参数就变成了1×1×192×64+(1×1×192×96+3×3×96×128)+(1×1×192×16+5×5×16×32)，参数大约减少到原来的三分之一。同时在并行pooling层后面加入1×1卷积层后也可以降低输出的feature map数量，结构中在pooling后面加了通道为32的1×1卷积，使得输出的feature map数降到了256。GoogLeNet利用1×1的卷积降维后，得到了更为紧凑的网络结构，虽然总共有22层，但是参数数量却只是8层的AlexNet的十二分之一。由此可见，GoogLeNet的优点十分突出。

所以在此次实验中，我使用GoogLeNet来训练手写汉字分类器。

1. 人脸识别

# 实验环境

# 实验方案与过程

1. 车牌识别
2. 光学字体识别/印刷字体识别
3. 手写汉字识别

最开始我们需要下载图片，生成相关文件然后用caffe自带的GoogLeNet进行训练。项目结构如下图所示。

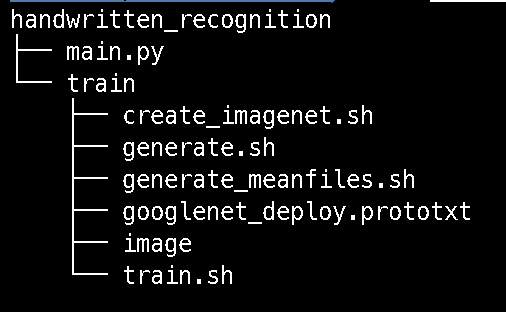


Figure 项目结构

训练的具体步骤如下：

1. 下载中科院NLPR[1]提供的数据集到`train/image`
2. 运行`generate.sh`生成训练集和验证集，会得到`train.txt`和`val.txt`，包含了测试集和验证集的图片路径
3. 运行`generate\_meanfiles.sh`来生成均值文件，会得到`meantiles`，包好生成的均值文件
4. 运行`train.sh`来进行训练，会到得`models`，包含训练好的模型

训练时使用的数据集一共有3755个汉字。GooGleNet的原理在“实验相关原理与技术部分”已经详细说明，此处不再赘述。我在训练过程中记录了不同迭代次数时，对于不同k的Top K可能的结果的准确度，情况如下。

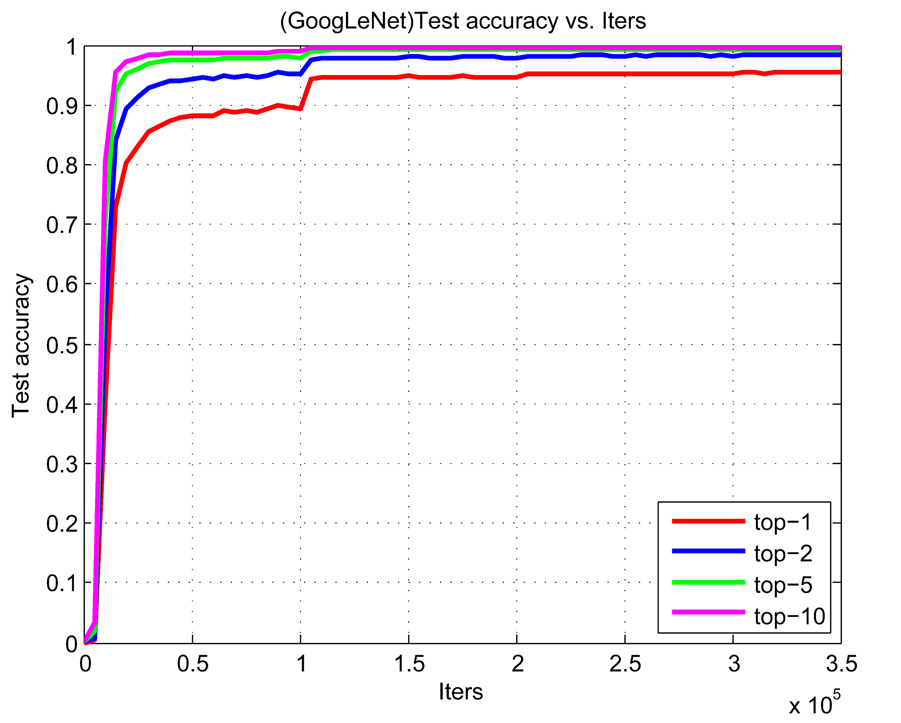
****

Figure 准确度-迭代次数

在最后的识别过程中，需要先将图片进行二值化和裁剪，然后将处理后的图片传入训练出的模型，获取top1的预测结果。运行s结果展示见“实验简介” 部分。

1. 人脸识别

# 实验总结

# 参考文献

[1] <http://www.nlpr.ia.ac.cn/databases/handwriting/Offline_database.html> CASIA Online and Offline Chinese Handwriting Databases

[2] <https://ieeexplore.ieee.org/document/7333881/all-figures> High performance offline handwritten Chinese character recognition using GoogLeNet and directional feature maps

[3]