* **机器学习毕业项目** [**侦测走神司机**](https://github.com/nd009/capstone/tree/master/distracted_driver_detection) **开题**

王闻宇  
2017年5月31日

**项目背景**

我们经常遇到这样的场景：一盏灯变成绿色，你面前的车不走。 另外，在没有任何意外发生的情况下，前面的车辆突然减速，或者转弯变道。等等这些现象，给道路安全带来了很大的影响。

那么造成这样现象的原因是什么，主要有因为司机疲劳驾驶，或者走神去做其他事情，想象身边的例子，开车时候犯困，开始时候打电话，发短信，喝水，拿后面东西，整理化妆的都有。这对道路安全和行车效率形成了极大的影响。



据中国安全部门介绍，五分之一的车祸是由司机分心引起的。令人遗憾的是，这样一来，每年有42.5万人受伤，3000人因分心驾驶而死亡。

我们希望通过车内摄像机来自动检测驾驶员走神的行为，来改善这一现象，并更好地保证客户的安全。

**问题描述**

我们要做的事情，就是根据车内摄像机的画面自动检测驾驶员走神的行为。如果是安全驾驶则一切正常，如果有走神行为，给予警报提醒。

驾驶员可能存在的走神的行为，大概有如下几种，左右手用手机打字，左右手用手持方式打电话， 调收音机（玩车机），喝饮料，拿后面的东西，整理头发和化妆，和其他乘客说话。

侦测的准确率accuracy就是衡量解决这个问题好坏的重要指标

**输入数据集**

输入数据集来自Kaggle 下载地址如下：

<https://www.kaggle.com/c/state-farm-distracted-driver-detection/data>

下载下来解压后有3个文件

- driver\_imgs\_list.csv.zip (92.89K)

- imgs.zip (4G) 所有的图片数据, 解压后

- train (训练集数据)

- c0 ~ c9 分别表示不同状态的训练集

- test (测试集数据，用于提交Kaggle比赛的测试集)

- sample\_submission.csv.zip (206.25K) Kaggle比赛需要提交的样本

下面是10种状态下每个状态的示例图片：图片大小 640x480

|  |  |
| --- | --- |
| c0 安全驾驶  C:\Users\elnin\AppData\Local\Microsoft\Windows\INetCache\Content.Word\img_414.jpg | c1 右手手机打字  C:\Users\elnin\AppData\Local\Microsoft\Windows\INetCache\Content.Word\img_4217.jpg |
| c2 右手打电话  C:\Users\elnin\AppData\Local\Microsoft\Windows\INetCache\Content.Word\img_4221.jpg | c3 左手手机打字  C:\Users\elnin\AppData\Local\Microsoft\Windows\INetCache\Content.Word\img_204.jpg |
| c4 左手手机打字  C:\Users\elnin\AppData\Local\Microsoft\Windows\INetCache\Content.Word\img_3187.jpg | c5 调收音机（玩车机）  C:\Users\elnin\AppData\Local\Microsoft\Windows\INetCache\Content.Word\img_4810.jpg |
| c6 喝水  C:\Users\elnin\AppData\Local\Microsoft\Windows\INetCache\Content.Word\img_3979.jpg | c7 拿后面东西  C:\Users\elnin\AppData\Local\Microsoft\Windows\INetCache\Content.Word\img_3691.jpg |
| c8整理头发和化妆  C:\Users\elnin\AppData\Local\Microsoft\Windows\INetCache\Content.Word\img_12358.jpg | c9 和其他乘客说话  C:\Users\elnin\AppData\Local\Microsoft\Windows\INetCache\Content.Word\img_5341.jpg |

其中 driver\_imgs\_list.csv.zip 的是对分类标号和人分类编号的csv文件。这个csv表格有三列

* subject：人的ID，不同的人，这个值不同
* classname：状态，c0 ~ c9
* img：图片名称

**解决方法**

这是一个分类器分为，预测的时候是将图片进行归类 C0~C9

首先，第一步，对图片进行预处理，保留有用的部分

第二步，将train数据，划分为训练和验证集，

第三步，建模调参，首先采用迁移学习(transfer learning) 的方式，对imagenet上的已经训练好的模型拿过来，只对以已经预测过的数据做全连接层的训练。在验证集上看精度accuracy. 然后尝试改造模型和自己建模，在验证集上看精度accuracy。

第四步，选择最高精度accuracy的模型和参数，生成Kaggle测试集提交文件，手动拿出20个数据开人眼判断是否正确

**基准模型**

我用ImageNet上已经成熟的模型来做基准模型来和我的计算结果做对比。

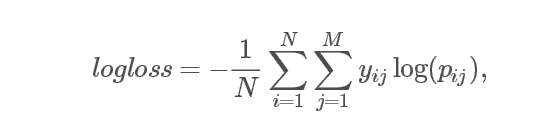
我选择选择 ResNet50 的去掉原有全连接层之后，自己训练全连接层来做为基准模型。 我基于这个基准模型再做改进。

我的目标是 accuracy > 0.9 并且 logloss < 0.6

**评估指标**

这是典型的分类问题，评估指标采用

1. 精度accuracy来评估结果好坏。
2. Logloss的评估方式，这也是kaggle比赛的评估方式



对比这两种方案。Accuracy对于判断正确和错误的比重是一样的，也就是对了就多一个，错了就少一个，最终看正确的百分比

而logloss的评估方式对判断是不是是有明显的方法，如果正确了，Pij=1 => log(Pij)=0, 而Pij=0.999 => log(Pij)=-0.001。最后增加的log差不多。但如果判断错误，如 Pij=0 => log(Pij) = -无穷。Pih=0.001 => log(Pij)=-6.9 也就是判断错误一个，对等分影响会非常大。

我认为，在accuray>0.95的情况下，看logloss更有意义

**设计大纲**

1. 预处理
2. 基准模型评估
3. 改造模型或者自建模型评估
4. 模式可视化和调优，得到最终模型
5. 用kaggle测试集数据生成pred.csv, 然后提到到kaggle看全球排名情况