

侦测走神司机开题报告

施梦杰 2018.01.29

项目背景

每年都有数以万计的人死于交通事故，其中人为原因占绝大部分。开车走神，疲劳驾驶这些不经意的行为都有可能造成严重的后果。这些情况可能会导致车辆的突然加速、减速，行驶状态不及时避让。这些行为对行车安全埋下了重大的隐患。

中国社科院与某保险机构的联合调研显示，道路交通风险最大的是分心驾驶，分心驾驶主要表现在疲劳驾驶和开车使用手机等方面，是道路安全的头号潜在杀手。

这个项目的目标是，通过车载的摄像头对司机的行为进行自动监测，通过这样的方式去改善分心驾驶的隐患。

我们可以将这个问题转换成图片分类的问题，在该领域已经获得很多成果。比如将数据集 [\(CIFAR-10\)](#) 图片十分类的问题，优秀的深度学习模型的正确率可以达到 95%（如 [Fractional Max-Pooling](#)）。我们可以参照优秀的深度学习模型来建立分类司机状态照片的模型来完成这个项目。

问题描述

通过车载的摄像头，按一定的频率获取照片，对照片上司机的状态进行分类。如果是安全驾驶状态，则维持正常状态。如果监测到司机有走神的行为，则发出警报，提醒司机。

我们侦测的走神行为有这么几种：驾驶过程中打字、打电话，调节收音机，喝饮料、拿后座东西，整理头发和化妆以及同其他乘客说话。

输入数据

输入数据来自于 Kaggle ([数据下载地址](#))

其中提供了三个文件：

- driver_imgs_list.csv

这个文件中，包含了训练数据中，每个图片文件名，司机的编号以及这张图中司机的状态。

- imgs.zip
 - train (10 个 class，代表不同的司机状态，已经分类成 c0-c9)
 - Test (测试数据)
- sample_submission.csv (kaggle 提交测试数据的样本)

接下来，我将抽样一些图片数据，了解一下数据。

c0: 安全驾驶



c1: 右手打字



c2: 右手打电话



c3: 左手打字



c4: 左手打电话



c5: 调收音机



c6: 喝饮料

c7: 拿后面的东西



c8: 整理头发和化妆

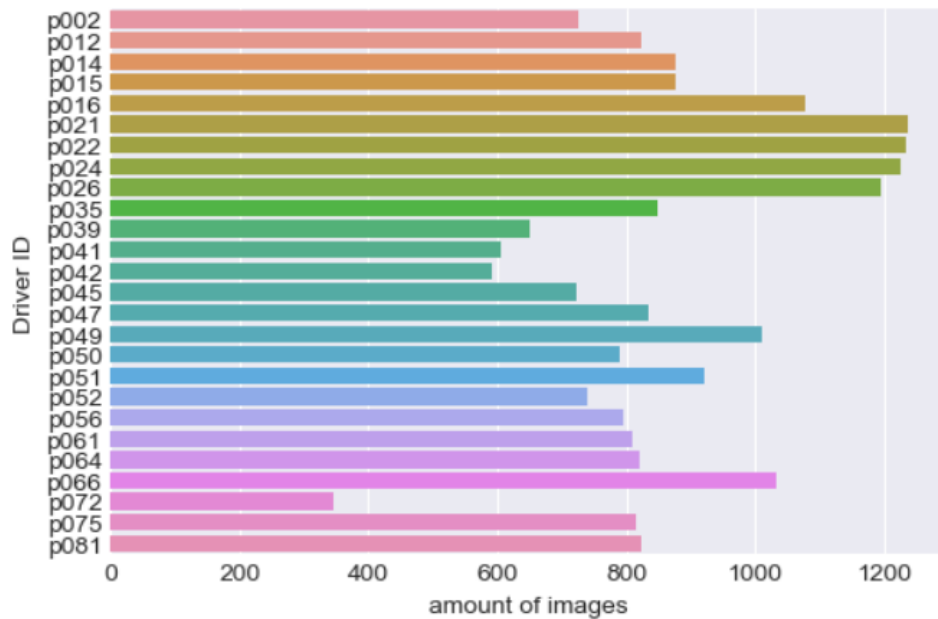


c9: 和其他乘客说话



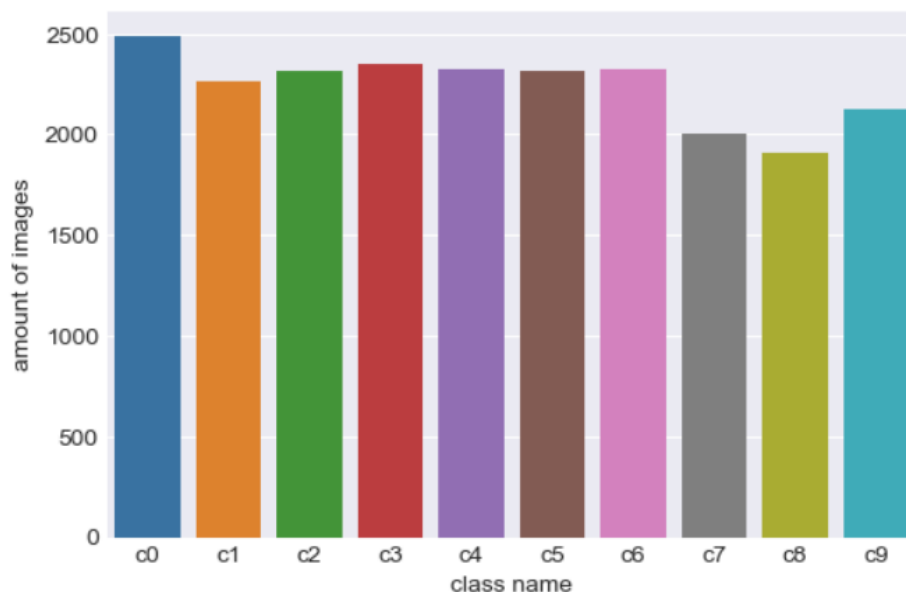
我们可以看到，这些图片是通过副驾驶上方位置的摄像头拍摄的，司机是有重复的，接下来我们来看一下数据的分布。

通过分析 driver_list_imgs 文件，可以知道总共有 26 位司机，22424 张图片，下面是每位司机的照片数量分布。



为了在训练的过程中，为了避免过拟合，我将 list 中最后两个司机的照片作为验证集，其他图片作为训练集。接下来我们看一下每个类别的分布。

```
Mean: 2242.4
Standard Deviation: 175.39
Text(0.5,0,'class name')
```



从数据和图表来看，数据分布是比较均匀的，均值为 2242.4，方差为 175.39。

解决办法

我们要做的是个分类器模型，输入一张图片，然后将图片分为 C0-C9 中的一类。

首先，对数据进行预处理，将图片格式转换成训练数据。然后将训练数据分割成两部分，训练集和数据集。我在观察数据集的时候，发现训练数据是少量司机的各种状态的图片，如果随机分割训练集和验证集，可能会出现过拟合的现象。所以我按照司机的 ID 分成训练集和验证集。第三步，建立模型，用迁移学习的方式，导入 imagenet 上训练好的权重，开发少量的层数，或者增加少量的层数进行训练，观察预测精度。在此基础上，优化模型。最后，挑选表现最好的模型对测试数据进行预测，生成 submission 文件，上传至 kaggle 上进行评分。

基准模型

我的基准模型选择为用 imageNet 数据预训练的 VGG16 模型，去掉全连接层，增加自己定义的全连接层，最后用 softmax 进行 10 分类。将这个模型作为基准模型，进行优化，进一步提升性能。

我的目标是 $\text{logloss} < 0.25$ ，处于 kaggle 前 10% 的水平。

评估指标

我选用的评估指标为 accuracy 和 log loss。

Accuracy 指标可以看出模型分类的准确度，分类正确和分类错误的比重一样的。

此外，kaggle 评估的方式是 muti-class logarithmic loss，公式如下：

$$\text{logloss} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^M y_{ij} \log(p_{ij}),$$

这个评估方式，对不同信心的预测惩罚不一样，比如更大信心的错误判断会比稍微小一点的信心的错误判断的惩罚更大，而更大信心的正确判断的惩罚会比更小信心的正确判断更小。

设计大纲

1. 划分训练集，验证集

在所有训练集中，选取两位司机的照片最为验证集，其他作为训练集。验证集在原训练集占比大致为 10%。

2. 预处理数据

在这个项目中，我将尝试多个模型，VGG16，RESTNet50，InceptionV3，用 imageNet 预训练过的权重。数据预处理的过程中，先缩放至模型默认的大小，vgg16 模型是缩放至（224，224）然后减去均值。

3. 训练多个模型并评估比较

训练选定的模型，进行对比。训练的方式是用迁移学习，在原先模型的基础上添加层数，将原先模型所有层的权重冻结，不进行训练。只训练后来添加的层。

4. 优化模型并评估

在观察训练集的时候，训练数据比较单一，可能出现过拟合的现象。我会尝试一下策略：

- Dropout
- 降低模型复杂度，增加正则化项。
- 将多个模型的预测值取均值，作为最终的预测值。

5. 挑选最佳模型对测试数据进行预测，上传至 kaggle 进行评分