

# 机器学习纳米学位

## 毕业项目开题报告

张鸣一

2017年7月15日

## 用深度神经网络侦测走神司机

(Using Deep Neural Network to Detect Distracted Driver)

### 项目背景

驾驶员未能完全将精力集中于道路上,致使他们的判断出现偏差,谓之"分心驾驶(distracted driving)".导致分心驾驶的原因很多,其中包括:驾驶员用手机通话,传短信,阅读,使用全球定位系统,观看视频或电影,吃东西,吸烟,补妆,跟乘客聊天,身心疲劳等.

驾驶期间,驾驶员必须在任何时候都要全神贯注.否则会带来严重后果.研究表明,驾驶过程中使用移动电话的驾驶员发生交通事故的风险大约是不使用者的四倍.在所有交通事故中,近三成涉及分心驾驶.

世界卫生组织WHO的资料[1]显示,随着移动电话用户的增长和新型车载通讯系统的迅速采用,分心驾驶对安全造成的严重威胁正逐年增加.无人驾驶虽然取得了长足进步,但完全替代人类还尚需时日.如何有效地侦测司机走神并加以提醒就变得尤为重要.通过车载监视器监测驾驶员状态,通过深度卷积神经网络的分析,可以一定程度上达到这个目的.

### 问题描述

项目的目标是根据车载监视器显示的画面自动判断驾驶员是否走神,若驾驶员走神,则给予提醒.通过对镜头中二维图像的分析,我们可对驾驶员的状态进行分类.为了达成项目目标,我们将建立一个深度神经网络模型,输入监测器中的图像,分析图像中的特征,将图像内容分入相应的类中,并显示出分类结果.分类的好坏将以准确率来衡量.

### 数据集与输入

数据集来自Kaggle的"State Farm Distracted Driver Detection"竞赛[2].数据集分为三个文件.

- imgs.zip 所有训练集和测试集图片的压缩文件:
  - train文件夹 (训练集图片,按驾驶状态分类如下)
    - c0文件夹 安全驾驶
    - c1文件夹 右手手机打字
    - c2文件夹 右手打电话

- c3文件夹 左手手机打字
  - c4文件夹 左手打电话
  - c5文件夹 调收音机
  - c6文件夹 喝水
  - c7文件夹 拿后面的东西
  - c8文件夹 整理头发和化妆
  - c9文件夹 和乘客交谈
- test文件夹 (测试集图片)
- driver\_imgs\_list.csv.zip 所有训练集图像文件名(img)列表, 包含对应的驾驶员(subject)和行为分类(classname)
- sample\_submission.csv.zip 比赛提交样本

## 解决方法

此项目为一个分类任务. 最终目的是将预测集的图片分类. 为达到这个目的, 我们首先得预处理数据, 将训练集的图片分成训练和验证两部分. 为了节省计算资源, 我们将先采用迁移学习, 用经过ImageNet预训练模型做全连接层的训练. 并在验证集上看预测精度. 在此基础上改进模型并训练. 最后选出最高精度的模型对测试集的图片进行分类, 上传Kaggle进行测试.

## 基准模型

我将使用ImageNet上预训练的模型来作为基准模型. 我选择使用Keras的ResNet50[3], 并自训练全连接层的模型来作为基准模型.

## 评估指标

评估指标采用准确率和Logloss来衡量. 准确率可以粗略评估模型的好坏. 而Logloss可以更加精确地评估模型对于分类的确定程度. Kaggle的Leading Board采用的方式为Logloss.

Logloss的表达式为

$$\text{Logloss} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^M y_{ij} \ln p_{ij}$$

其中  $N$  是测试集的图片数,  $M$  是图像类别数(在此项目中  $M = 10$ ),  $\ln$  是自然对数, 当第  $i$  张图属于第  $j$  类,  $y_{ij}$  是1, 反之则为0.  $p_{ij}$  是预测第  $i$  张图属于第  $j$  类的概率, 取值在0到1之间. 从中我们可以看出, 当模型预测得越准确, 置信概率越高, 那么Logloss就越小. 因此Logloss的值越小表明模型越好.

根据我对Kaggle上面讨论的浏览, 我决定将我的目标定在准确率大于90%, Logloss小于0.5. 希望可以达成.

## 项目设计

1. 图像预处理. 这里可能要对图像进行加强, 用kNN(k-Nearest-Neighbor)来加强图像.

2. 用基准模型进行评估. 可分别对预处理过的图片和未处理的图片进行模型训练, 观察过拟合情况.
3. 改进模型. 也可能会用上其他的预训练模型, 如Vgg-16等, 观察模型在验证集上的表现.
4. 模型可视化以及微调. 可将模型结果融合, 得到最终模型.
5. 将模型预测预测集的结果上传Kaggle, 查看模型Leading Board上的表现.

## 参考文献

- [1] Mobile phone use: a growing problem of driver distraction  
[http://www.who.int/violence\\_injury\\_prevention/publications/road\\_traffic/distracted\\_driving/en/](http://www.who.int/violence_injury_prevention/publications/road_traffic/distracted_driving/en/)
- [2] State Farm Distracted Driver Detection  
<https://www.kaggle.com/c/state-farm-distracted-driver-detection>
- [3] Deep Residual Learning for Image Recognition  
<https://arxiv.org/abs/1512.03385>