

Proposal

2018 年 5 月 5 日

1 监测走神司机 - 开题报告

王晨龙

2018 年 5 月 4 日

1.1 项目背景

随着人们生活水平以及交通道路的发展，现在道路上行驶着越来越多的车辆。我们会经常遇到这种情况，在交通路口等待红灯时，交通灯已经变成了绿色但是前面的车并没有走，只有当后面的司机鸣笛提醒时才开始行驶。还有一种情况就是前车突然减速，并且不再按照直线行驶出现了左右摇摆的情况。

当发生这些情况时通常会影响到交通效率，并且有可能造成交通事故，给路上的行人带来了安全隐患。之所以出现以上情况大部分是因为司机在驾驶过程中走神导致的，比如开车时发短信、打电话甚至是浏览社交网站。根据美国 CDC 汽车安全部门的统计，大约有五分之一的交通事故是因为司机走神导致的，每年因此而导致的伤亡人数分别达到了 42.5 万人和 3000 人。因此我们希望通过一些技术手段来监测司机的驾驶行为，对不良的驾驶行为能够及时提醒，减少因此发生的交通事故。

1.2 问题描述

Kaggle 为我们提供了一些数据，这些数据是司机驾驶过程中的 2D 图片，我们希望通过这些图片训练一个可以检测司机驾驶行为的模型，能够正确的判断不同驾驶行为的特征，从而实现对司机驾驶行为的监测及提醒。很显然该问题是一个监督学习的分类问题。

1.3 数据集和输入

数据集来自[Kaggle](#)。

- **imgs.zip** - 所有图片（训练、测试）文件压缩包，训练集已经将图片按照司机行为进行分类 c0 ~ c9

- **sample_submission.csv** - 提交到 Kaggle 时正确的格式模板
- **driver_imgs_list.csv** - 训练图片列表，包括司机 ID 和类别 ID

1.4 解决方案

由于数据及是大规模的图片，属于深度学习中计算机视觉范畴的问题，本文将选择卷积神经网络作为模型。

首先，针对训练集数据进行预处理，了解数据的基本特性。

第二步，选取合适的卷积神经网络基准模型。

第三步，将图片的像素信息作为模型输入，通过模型提取图片的特征，并基于训练集数据不断优化模型参数。

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。

1.5 基准模型

我将使用 NASNet 作为基准模型，基于这个模型进行改进，期望进入 Kaggle LeaderBoard 上排名的前 10%，模型 logloss 需要小于 0.2563。

1.6 评价指标

通过 logloss 进行评估，这也是 Kaggle 的评估方式

$$\text{logloss} = -\frac{1}{N} \sum_{i=0}^N \sum_{j=0}^M y_{ij} \log(p_{ij})$$

- N 是测试集中图片的数量
- M 是图片分类标签的个数
- log 是自然对数
- y_{ij} 当样本 i 标签为 j 时值为 1，否则为 0
- p_{ij} 为样本 i 归属于标签 j 的预测概率

1.7 设计大纲

1. 图像数据加载分析

对数据集中的数据进行加载，并进行可视化分析，分析数据集特点。

2. 图像数据预处理

利用 opencv 对图像数据进行处理，去除无关的数据；对图像进行去中心化处理；完成训练集和验证集的划分。

3. 基准模型创建及评估

利用 **Keras** 建立基准模型，然后对基准模型进行训练和测试，利用测试集数据对测试结果进行评估，将预测结果提交到 **Kaggle** 得到评估结果。

4. 模型建立及评估

基于 **Keras** 提供的基准模型建立初始模型，对初始模型进行训练评估。

5. 模式可视化和调优，得到最终模型

使用迁移学习、模型微调和模型融合等技术对初始模型进行修改，多次将预测结果提交到 **Kaggle** 得到评估结果，得到最终模型。

6. 用 Kaggle 测试集生成 submission.csv，提交到 Kaggle 查看排名情况

根据 **Kaggle** 要求，输出数据集。

1.8 参考文献

- [State Farm Distracted Driver Detection](#)
- [Learning Transferable Architectures for Scalable Image Recognition](#)
- [入门 | 从 VGG 到 NASNet，疑问概览图像分类网络](#)