

侦测分神司机

(机器学习纳米学位毕业项目开题报告)

方无迪 2017 年 10 月 15 日

1 项目背景

本项目名为侦测分神司机，来源于数据分析竞赛网站 Kaggle 在 2016 年 8 月的比赛项目[1]。我们知道，司机的分神行为（比如打电话、发短信）对行车安全和交通效率有很大影响。该项目的愿景是通过车内摄像机来自动检测司机驾驶行为，来据此更合理地为顾客车险投保，有助于改善所述现状。



插图 1：分神司机样照

比赛发起方为 State Farm，美国最大的车险公司。State Farm 公司在 2011 年就开始运用 Usage-Based Insurance (UBI) 车险商业模式，与车联网厂商 Hughes Telematics 合作开展基于驾驶行为数据的保费模式[2]。可以看出，该项目的发起也体现了 State Farm 正在对未来进行探索性布局。



插图 2：UBI 保险模式

该项目是从汽车保险这一大规模消费市场的实际需求出发，着眼于未来车联网大数据环境下的驾驶风险评估系统，利用机器视觉和机器学习技术，搭建驾驶行为自动检测模型，其商业价值和应用前景值得期待。

2 问题描述

比赛主办方提供给我们已经分为 10 个类别的照片，示例如下图所示，作为机器学习模型训练所用。并要求我们的模型能够对测试照片进行类别预测，以判断司机当前是处于哪种状态。

表格 1 10 类别及示例照片

 C0 安全驾驶	 C1 右手发短信	 C2 右手打电话
 C3 左手发短信	 C4 左手打电话	 C5 调收音机
 C6 喝水	 C7 伸手到后面	
 C8 整理头发和化妆	 C9 和乘客说话	

该问题实质上属于有监督机器学习(supervised learning)的分类方向(classification)，并且是计算机视觉(computer vision)范畴。模型的目标是利用大量的分类图片进行训练，来学到不同驾驶行为的特征，然后可以正确识别未见过的驾驶照片。

3 输入数据

输入数据可以从该 Kaggle 比赛的 Data 页面下载。主要包括 22424 张训练用照片和 79726 张测试用照片，以及训练照片的列表（包含所属类别和司机 ID）。

其中，照片尺寸为 640×480，并经过 metadata 元数据去除处理（以保证问题是纯粹的机器视觉）。测试集还包含了一些扩充数据（不参与分数计算），以抵御手动标注测试集。

此外，训练集和测试集是按照不同司机分割的。这提示我们，在后续训练时也应应对训练集按照不同司机划分出验证集，来判断模型是否具备应用于未知司机的泛化能力。

4 解决办法

针对这个具体问题，给出解决方案如下。

- (1) 从训练集的 26 个司机中，划出数个司机作为验证集
- (2) 设定图像数据增强的多种参数，产生训练数据
- (3) 利用迁移学习，选取预训练模型，设定某些层作后续 fine-tune
- (4) 设定好超参数，开始多个 epoch 的模型训练，直到验证集指标达到要求
- (5) 重复(1)~(4)的步骤多次，并将测试集结果进行集成，提交至 kaggle
- (6) 对结果进行可视化

5 基准模型

为了和解决方案做对比，拟采用经过 ImgeNet 预训练的 Vgg16 模型[3]，并全新训练末端全连接层，作为基准模型。

VGG16 模型简洁和直观，在考虑使用卷积神经网络解决图像问题时，通常可以先使用 VGG16 作为基准模型，来对数据集进行初步验证，然后再考虑对模型进行优化或使用层数更多的模型，并用于后续的客观对比。

6 评估指标

为量化基准模型和解决方案，采用 Kaggle 比赛 Private Leaderboard 的得分作为评估指标。该得分是，利用服务器上测试集的真实标签，对 69%比例的测试集计算得到的多类对数损失。其中，多类对数损失公式如下。

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

式中， N 为样本数量， M 为类别数量， y_{ij} 为表征样本 i 是否实际属于类别 j 的0-1指示函数， p_{ij} 为样本 i 在类别 j 上的预测概率。

这个评估标准对于问题本身、数据集以及解决方案来说都是合适：①问题本身是属于多分类问题；②69%比例测试集多达55000，足够用来验证模型泛化能力；③解决方案中还可以将该指标应用于划分出的验证集，便于超参数选取等。

7 设计大纲

为实现解决方案和获取结果，制定实施流程如下。

- (1) 数据探索。对类别分布、类间特征差异进行探索性分析。
- (2) 基准模型。搭建基准模型，并初步评估。
- (3) fine-tune 模型。选取不同验证集、不同超参数、不同的迁移模型，进行训练和测试。
- (4) 输出结果。将 fine-tune 模型的测试集结果进行集成，并提交 kaggle。
- (5) 可视化及讨论。使用 CAM 可视化，分析模型的类激活图，并做后续改进讨论。

参考文献

1. [State Farm Distracted Driver Detection \(Kaggle\)](#)
2. [UBI 车险海外案例简析：State Farm（车云网）](#)
3. [K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale imagerecognition. In ICLR, 2015.](#)