

毕业项目开题报告——DeepTesla

一、项目背景

自动驾驶是汽车产业与人工智能、物联网、高性能计算等新一代信息技术深度融合的产物，是当前全球汽车与交通出行领域智能化和网联化发展的主要方向。自动驾驶涉及的软硬件有传感器、高精度地图、V2X（车辆与周围的移动通信控制系统实现交互的技术）以及 AI 算法，其中 AI 算法处于相对最重要的地位。目前在算法方面已经有了大量研究成果并已开始成熟。本项目旨在实现自动驾驶 AI 算法中的行驶中方向盘角度预测。本项目采用了端到端的实现方式，其中涉及到深度学习、计算机视觉相关领域的内容。

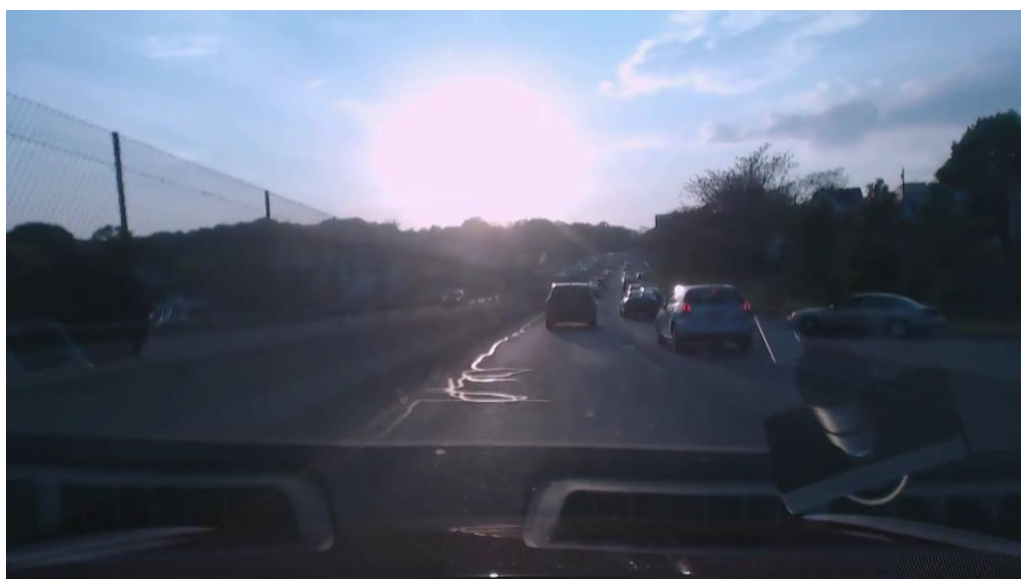
二、问题描述

本项目中，需要利用 MIT 6.S094 这门公开课中的 Tesla 数据集训练深度学习模型，根据车辆的前置相机所拍摄的路况图像，实现对车辆转向角度的预测。

三、输入数据

数据集总共分为两部分：

1) tesla 在两种不同驾驶模式（human driving 和 autopilot）下的前置相机录制的视频（其中一个 csv 文件表明各段视频所属类别），视频为 mkv 格式。共有十段视频，第一段 50 秒，第二段 120 秒，其余均为 90 秒。第十段视频将作为测试集，其余作为训练集。视频分辨率为 1280x820，其中下部为车盖部分训练时需要进行裁剪。图片样例如下：



2) 视频对应的行驶过程中的控制信号，使用 CSV 格式，具体内容以及样例如下：

ts_micro	frame_index	wheel
1464305394391807	0	-0.5
1464305394425141	1	-0.5
1464305394458474	2	-0.5

其中，ts_micro 是时间戳，frame_index 是帧编号，wheel 是转向角度（以水平方向为基准，+为顺时针，-为逆时针）

四、解决办法

将视频转换为图片集合并进行预处理作为特征，车辆转向角度作为标签来实现端到端的深度学习。使用深度卷积神经网络进行的模型训练，模型采用若干卷积层池化层，若干全连接层，最后输出一个方向盘角度的值，具体层数以及超参数的选取将在实现过程中进行不断调节。

五、基准模型

使用此前用来训练 CIFAR-10 中使用的较浅的简单卷积神经网络作为比较。此网络具体为 64-128-256 三层卷积池化层，其中 stride 为 2x2，padding 为 SAME；512-256-128 三层全连接层；各层均使用 ReLU 作为激活函数；dropout 采用 0.5。

六、评估指标

以 y_{-} 作为角度预测值， y 作为角度准确值， n 为样本数量，使用 MSE（Mean Squared Error）作为损失函数，对模型准确度进行衡量。公式如下：

$$loss = \frac{1}{n} \sum_{i=0}^n (y_{-} - y)^2$$

七、设计大纲

1、数据处理：

- 1) 对数据进行相关理解。
- 2) 通过 opencv 将视频转换为图片并进行一定的预处理。
- 3) 将图像数据以及对应的方向盘角度数据整合到一起进行存储，其中保留部分作为测试集。
- 4) 将得到的数据分为训练集和验证集放入模型进行训练。

2、模型训练：

- 1) 使用 keras 完成几个模型的搭建。
- 2) 将小部分的数据放入模型观测模型搭建是否准确, 以及模型效果是否显著。
- 3) 选取一个较为可靠的模型进行整个数据集的训练。
- 4) 使用测试集进行预测, 如不够好回到第一步继续尝试。

3、输出结果：

使用已实现的 `run.py` 来调用模型并对得到的方向盘角度进行可视化显示输出一个新的视频。