

# 人工智能课程项目报告三

小组成员 王辉 2015201943

梁晓周 2015201921

吴红薇 2015201928

## 1.背景知识

TensorFlow Object Detection API 是一个构建在 TensorFlow 之上的开源框架，可以轻松构建，培训和部署对象检测模型。在 17 年 11 月 17 日发布的最新版本中，它提供了 15 种在 COCO 数据集，Kitti 数据集和 Open Images 数据集上预先训练的检测模型集，可供直接调用。COCO 训练模型中 speed 项指的是每 600\*600 图像的运行时间；COCO mAP 可以理解为准确度的一种度量值，该数值越高代表准确度越好。在我们的实验中，用到的是速度最快的 ssd\_mobilenet 模型。

Model name	Speed (ms)	COCO mAP@10	Outputs
ssd_mobilenet_v1_coco	30	21	Bboxes
ssd_inception_v2_coco	42	24	Bboxes
fasterrcnn_inception_v2_coco	58	28	Bboxes
fasterrcnn_resnet50_coco	89	30	Bboxes
fasterrcnn_resnet50_ineroposals_coco	64		Bboxes
rfcn_resnet50_coco	92	30	Bboxes
fasterrcnn_resnet50_coco	106	32	Bboxes
fasterrcnn_resnet50_ineroposals_coco	82		Bboxes
fasterrcnn_inception_resnet_v2_atrous_coco	620	37	Bboxes
fasterrcnn_inception_resnet_v2_atrous_ineroposals_coco	241		Bboxes

图 1 模型列表

随着深度学习的发展，卷积神经网络变得越来越普遍。当前发展的总体趋势是，通过更深和更复杂的网络来得到更高的精度，但是这种网络往往在模型大小和运行速度上没多大优势。一些嵌入式平台上的应用比如机器人和自动驾驶，它们的硬件资源有限，就十分需要一种轻量级、低延迟、同时精度尚可接受的网络模型。

MobileNet 基于一个流线型的架构，使用深度可分离的卷积来构建轻量级的深层神经网络。其引入两个简单的全局超参数，在延迟度和准确度之间有效地进行平衡。MobileNets 模型使用深度可分解的卷积，它可以将标准卷积分解成一个深度卷积和一个点卷积也就是  $1 \times 1$  卷积核，深度卷积将每个卷积核应用到每一个通道，而  $1 \times 1$  卷积用来组合通道卷积的输出。这种分解可以有效减少计算量，降低模型大小，计算量减少了近 90%，但精度只下降了 1%。



图 2 MobileNet 模型结构

MobileNet 致力于打造一个轻量化的深度神经网络，以便应用于移动终端（如手机或者 tesla 电动汽车）进行多对象识别。目前 MobileNet 是服务器上训练，未来甚至会考虑直接在终端训练。比较理想的模式是一个混合终端和云端的方式，一部分的计算在本地进行，其他

的对延时要求较低的任务就交给云端处理。



图3 MobileNet 模型应用

## 2.实验过程

本阶段目标：利用小车装载的手机摄像头传输的视频图像，依托深度神经网络模型进行物体图像识别，同时通过小车与电脑的蓝牙连接，实现信息交互和操作控制，最终实现“小车寻物”的功能。

### 2.1 图片物体识别

我们利用 Tensorflow 提供的 Object Detection API，选用其中的 `ssd_mobilenet_v1_coco` 模型训练出可以识别人、车、椅子等 90 种物体的深度神经网络模型。对于图片对象，该模型的物体识别函数能够识别图片中的物体，并重新渲染图片，显示出物体的名称以及预测可能性指数。



图4 图片物体识别结果

### 2.2 视频物体识别

接下来要将模型应用到视频的物体识别中，我们利用 MoviePy 提供的 `VideoFileClip` 接口读入视频片段，并以物体识别函数为参数调用 `fl_image()` 函数，这样就可以用每一个重新渲染后的帧替换原帧，得到一个动态识别物体的视频片段。



图5 视频物体识别结果

### 2.3 实时视频物体识别

为了实现实时视频的物体识别，我们先在手机端下载了可以通过 wifi 传输实时图像的软件“IP Webcam”，并得到视频流的 URL 地址。然后我们利用 OpenCV 提供的 `VideoCapture` 接口读入该 URL 的视频流内容，再对其进行视频物体识别。

实时视频物体识别最大的问题是延迟太严重，由于 wifi 传输速度以及深度神经网络模型识别速度的限制，往往会产生长达十余秒的渲染图像延迟。



图 6 实时视频物体识别结果

## 2.4 “小车寻物”功能加载

我们利用 PyBluez 提供的 BluetoothSocket 接口进行电脑与小车蓝牙之间的连接，从而实现电脑对小车的遥感控制。电脑端输入目标物体名称，如 chair，那么装载着手机摄像头的小车就会开始按照随机方向行走，电脑端根据进行实时视频物体识别，若识别到的物体名称与目标物体名称匹配，则通过蓝牙发送命令让小车停止运动并长鸣喇叭，表示已找到目标物体。

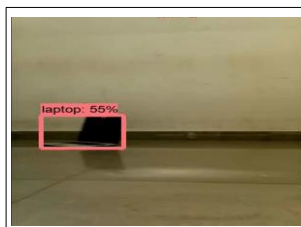


图 7 “小车寻物”识别结果

## 3 实验总结

这次实验时间跨度长，难度逐级上升，给我们人工智能课程小组的成员的实验时间安排、任务分配以及实验实施等方面都带来了一定的挑战。本次人工智能小车的实验基于课本知识，又高于课本知识，不但需要我们不断补充最新的、与我们的实验最相关的人工智能知识及学习相关工具使用方法，还对我们的动手能力提出了更高的要求，这不同于其他课程仅局限于编程实践的要求，更加有助于我们的全面发展。

我们的实验过程中存在着许许多多的问题，很遗憾的是，有些问题我们最终也未能解决，为了呈现一个比较好的效果，我们采取了一些折中方案，最后也取得了不错的成效，这说明了我们要充分发挥主观能动性去克服苦难。

记录人 梁晓周

记录时间 2017/12/24