

人工智能课程总结
&
AI-Car 实验报告

张倩

2016202127

一、课程总结

通过对本学期课程的学习，我加深了对人工智能这一热点技术的认识。例如图像识别，对于传统人力来说，我们既可以在一瞬间获取一幅图片的细节，也可以很容易感知到图片的种类与意境，从图片中获取大量信息。而传统技术或许可以通过不同的图片格式对图片进行呈现，但对于如何获取图片中更多的语义信息，“读取”图片内容则是十分被动的。对于计算机来说，通过巧妙的机器学习及神经网络学习等技术使机器有了“学习”的能力，并且使机器通过学习具备一定人类的能力，此之谓“人工智能”。

人工智能的研究可以分为几个技术问题。其分支领域主要集中在解决具体问题，其中之一是，如何使用各种不同的工具完成特定的应用程序。AI 的核心问题包括推理、知识、规划、学习、交流、感知、移动和操作物体的能力等。

在本学期内，我们对各问题的基本方法与原理等进行了学习，也学会了通过机器学习与神经网络等方法解决实际问题。伴随着整个学期进程的，是我们通过小车的拼装与测试，对小车功能的设计与实施，对深度学习技术有了更进一步的了解，通过自己的动手实施提高了自己的动手能力，也通过团队的完美配合在收获友谊的同时也通过互相帮助对课程内容有了更深一步的理解。在本学期的人工智能课程学习过程中，我认为从一堆零件开始，到小车可以具有语音识别、图像识别等功能本身就是一件十分有意思的事，虽然课程已经将要结束了，但在这段时间里收获到的知识与技能也将在未来的学习和工作中帮助到我们。我也将在后续的学习中继续对机器学习与深度学习等知识进行更进一步的学习与实验，加强自己在人工智能方面的知识储备与技能，提高自己的专业素养。

二、人工智能小车 AI-Car 设计报告

1、材料准备与小车拼装

在第一阶段，我们完成了对小车的拼装与测试。

我们使用了以下配件对小车实验环境进行搭建。

Arduino UNO 开发板	面包板
电机驱动扩展板 L293D	杜邦线若干
驱动马达	Arduino IDE
BT06 蓝牙串口模块	蓝牙串口 APP
超声波传感器 HC-SR04	DroidCam 软件

虽然充电宝供电稳定，但是电压较小，因此我们选择使用 18650 电池并配合电池相应充电器对小车进行供电与续航。

之后，我们对小车的使用蓝牙控制、超声波避障与使用红外避障等功能进行了实现。

对于小车的具体安装与基本避障，可以参考以下博客：

2、小车电脑图片传输

由于需要用手机摄像头对小车行进道路上的情况进行拍摄,而如果通过手机摄像头对画面进行拍摄再将图片传输到电脑的速度较慢,无法满足我们对小车前进的实时要求。因此,我们选择利用 DroidCam 软件,将小车摄像头拍摄到画面投射在电脑显示屏上,获取小车前进过程中的画面。

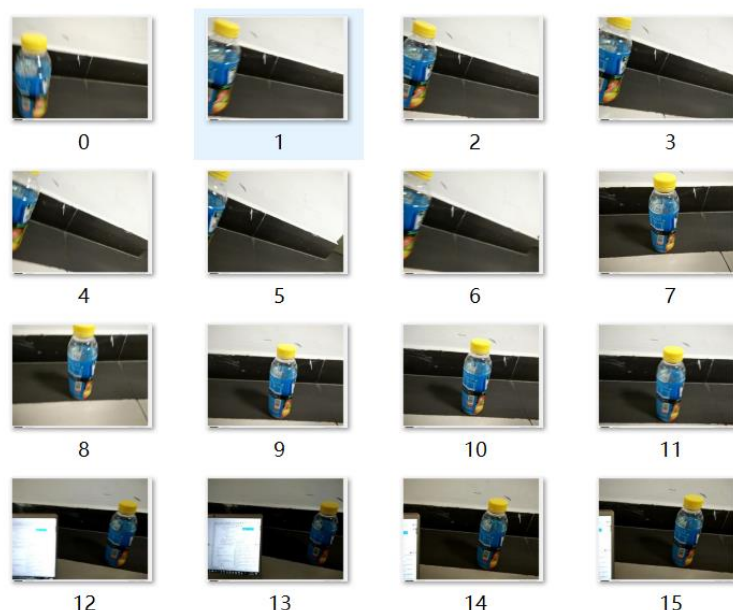


图 1 python 程序捕捉到的 DroidCam 画面图像

由于 DroidCam 的画面无明显导出接口,因此我们选择使用 python 程序对 DroidCamPC 端软件投射出来的图像进行截取,按照一定的时间间隔获取图像,将截取到的图像作为参数代入到后续的图像识别与分类模型中,根据模型解读图像信息,生成相应的小车行进指令,通过蓝牙串口发送指令,控制小车行进,实现小车前进道路上的实时识别。

截图代码如下(由于手机在小车上竖直放置,因此需要使用旋转函数对拍摄到的图像进行 90 度旋转后进行测试):

```
def screen():  
  
    im = ImageGrab.grab((20,120,620,560)) #截图  
  
    img = im.rotate(-90)  
  
    img.save('D:\Image\image4/'+str(c)+'.jpg','JPEG') #图片存储
```

3、图像识别——障碍物识别

在该部分,我们实现了识别图像中障碍物类别及障碍物在图片中的位置以调整小车的行

进方向及行进速度。

我们使用 OpenCV 的 dnn 模块，利用预先训练过的模型，实现这部分的障碍物识别。该模型主要利用 SSD 技术与 MobileNets 网络结构进行训练，MobileNets 网络结构使用深度可分离的卷积技术，减少了网路参数数量，从而有效减小了模型的大小，并且通过将其与 SSD（Single Shot Detector）技术结合起来。得到一个执行速度足够快、同时识别准确率足够好的模型。

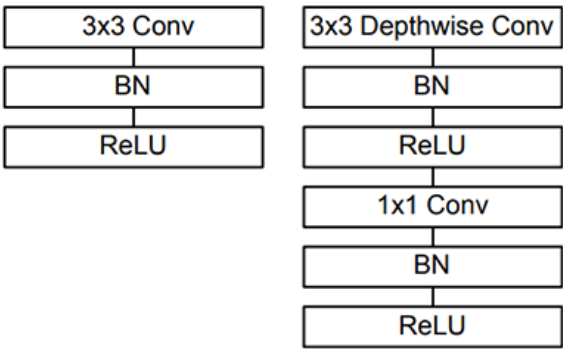


图 2 MobileNet 网络结构

该模型输入是一张张图片，输出是每张图片上，模型识别出来的一系列对象，及其 bounding box。

使用网络图片进行识别，示例如下：

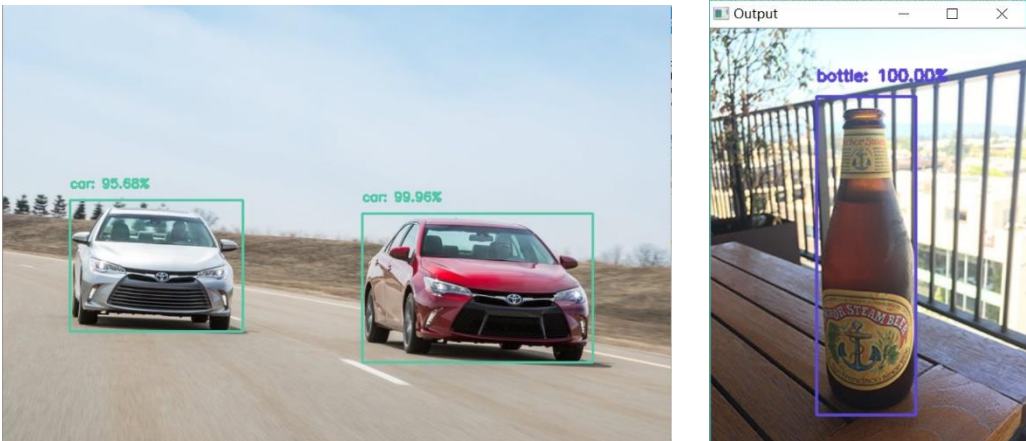


图 3 使用网络图片测试模型运行结果

使用我们在前进道路中有手机摄像头捕捉到的图像，获取模型测试结果，发现小车对障碍物的识别效果较好，程序可以很好的识别到图片中存在的物体，在实际测试中，我们选择 bottle 作为小车行进的障碍物，对具有 bottle 这一障碍物的图片进行特殊处理。

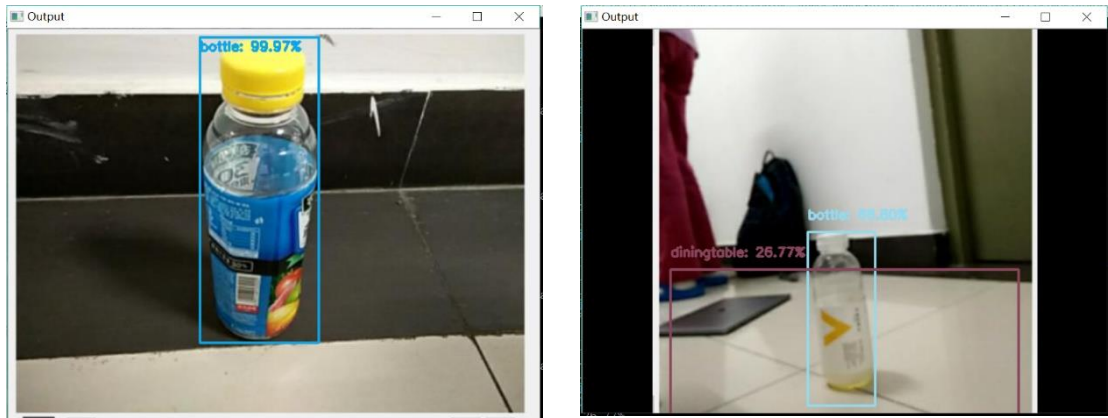


图 4 小车障碍物图测试效果

在部分图片中，程序会对内容较为复杂的图片产生识别错误，因此我们对于与实验无关的物体种类进行过滤处理，降低判断小车是否准备识别的阈值，减小实验误差。

在识别到障碍物的种类之后，我们根据模型返回的物体在图像当中的位置(startX, startY), (endX, endY)确定障碍物在小车视野中的大概位置，根据该位置确定小车应该采取何种方式绕过障碍物继续前进，或者当小车识别到前方有 person 出现时降低行驶速度，采取较慢的速度继续前行。

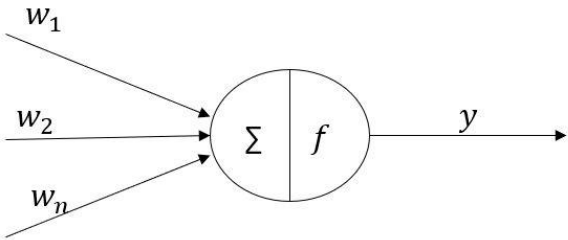
4、CNN 卷积神经网络——图片指示识别

卷积神经网络 (Convolution Neural Network, CNN) 是一种特殊类型的前向神经网络，普遍适用于图像识别、语音识别等领域，是为识别二维物体而特殊设计的多层网络结构，对平移、缩放、倾斜等变形具有高度不变性。

以下对 CNN 的背景知识及基本原理进行介绍：

(1) 神经元

神经元 (感知机 Perceptron) 是人工神经网络的基本处理单元，一般为多输入单输出，对输入信号进行加权求和，通过函数转化继而输出，若用 x_i 表示神经元的输入， y 表示神经元的输出， $f: \mathbb{R} \rightarrow \mathbb{R}$ 表示激活函数，则输出可用 $y = f(b + \sum_{i=1}^n x_i w_i)$ 计算得到。



1 一个神经元

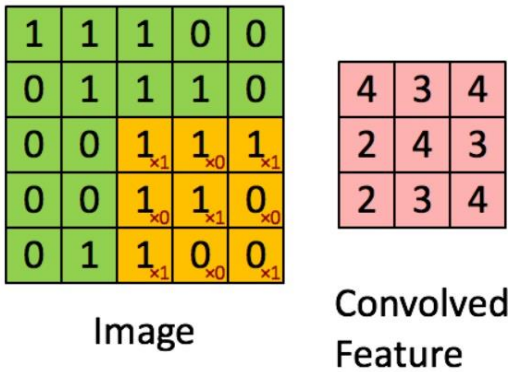
通过激活函数，神经网络具有对非线性关系进行建模的能力，我们可以通过对每个输入权重的不断调整，对神经元进行训练。虽然一个神经元通过训练可以完成简单的线性分类任务，但要处理更为复杂的模型则要对神经元进行进一步的组合利用才可完成。

(2) CNN 简介

CNN 由于一个映射面上的神经元共享权值，因而减少了网络自由参数的个数，降低了网络参数选择的复杂度。由于这个原因，可以直接将图像作为网络的输入，避免了传统识别算法中复杂的特征提取过程，并且提高了识别模型对图像变形的高度不变性。

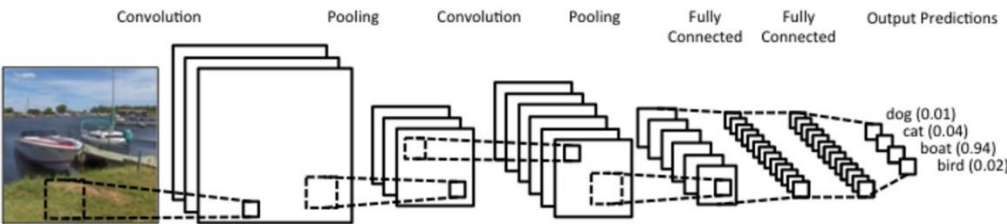
CNN 的基本结构由输入层，卷积层（convolutional layer）、池化层（pooling layer）、全连接层及输出层构成。卷积层和池化层一般取多个，采样卷积层和池化层交替设置，即一个卷积层连接一个池化层。卷积层中输出特征面的每个神经元与其输入进行局部连接，并通过对应的连接权值与局部输入进行加权求和再加上偏置值，得到该神经元输入值，该过程等同于卷积过程。

卷积层对输入使用一个或多个过滤器进行转换，应用在图像上，得到特征图，CNN 使用多个卷积核，形成对输入的多个特征图。卷积过程包括用一个可训练的滤波器 f_x 去卷积一个输入的图像，然后加一个偏置 b_x ，得到卷积层 C_x 。



池化层对输入图像进行子采样，子采样的方法一般采用最大池化、平均池化和随机池化等。子采样（池化）过程包括，将每 $2 \times 2 = 4$ 个像素构成的区域求和变成一个像素，然后通过 W_{x+1} 加权，再增加偏置 b_{x+1} ，然后通过一个激活函数，产生一个缩小 4 倍的特征映射图 S_{x+1} 。池化可以避免过度学习问题。

CNN 的最后一个卷积层或者池化层，通过全连接方式连接到输出层，输出层输出分类标签或者预测值。



(3) 方向识别

我们利用上述所述的 CNN 模型，使用控制小车拍摄的数据集训练出一个模型，模型共有 7 层，前四层为卷积层，均包括卷积、激活、池化等效果，后三层为全连接层。前四层原理和方法基本相同，差别在于通道数量及其他参数的变化，需根据每层的不同需要修改函数参数。可以调用 cnn 模型，提前对模型使用时可能需要的参数进行设置，并且定义模型训练过程中需要的具体函数，设置模型训练过程中的整体学习率及预测精度。

模型训练好之后，使用实时拍摄到的图片检测小车接下来要运行的方向，再将检查结果沿蓝牙串口发送到小车，控制小车行进。

三、 实验总结

- 1、在小车装配阶段，由于各设备硬件性能不同会对小车实际行驶状况产生较大阻碍，因提前考虑可能发生的情况，想清楚应对措施进行解决。
- 2、环境配置的细节问题。在第三阶段，由于每个人电脑系统、python 版本等的不同，需要安装的 tensorflow 及其他库的版本也有所不同，因此在未统一的环境下进行实验，代码大概率无法共享，在小组合作时，最好统一实验环境。
- 3、过拟合的问题。由于在进行 CNN 的实验时，训练数据集为我们通过手机摄像头实地拍摄，因此实验结果不具有普适性，且由于第二次扩充后的数据库中单类图片数量较少，因此在迭代达到一定次数后，将出现过拟合的问题，考虑扩充数据集之后，过拟合问题消失。但由于扩充数据集会增大计算负担，因此需根据具体情况对模型的各项参数进行调整，以提高训练精度与预测精度，并避免产生过拟合的问题。