# 7eam Project Report 垃圾识别与分类

小组名称: invalid syntax

小组成员: 顾吴量 诸圣怡 杨启航

指导老师:鲍杨

# 目录

- 1.简介
- 2.问题缘起
  - 2.1 垃圾分类政策
  - 2.2 垃圾分类现状
- 3.垃圾分类程序介绍
  - 3.1 关于 TensorFlow
  - 3.2 程序运行主要步骤
  - 3.3 具体展示
    - 3.3.1 我们的数据库
    - 3.3.2 训练新的分类器
    - 3.3.3 具体程序分步解析
- 4.项目缺陷与展望
- 5.成果与结论
  - 5.1 成果展示
  - 5.2 结论

# 垃圾识别与分类

7eam: invalid syntax

# 1.简介:

我们使用 Python+tensorflow 实现图片重新分类和识别的功能,并针对目前垃圾分类过程中用户难以区分垃圾种类的现状,将此照片识别功能用于垃圾分类应用中。我们从人脸识别功能出发,参考 Github 上 tf-classifier 的程序脚本以及训练过程,自行建立不同种垃圾的图片数据库,生成预训练模型,使用 Tensorflow 解析每一张图片的详细特征点,重新训练图片分类器,生成图片分类数据。最后运行程序完成识别新提供的垃圾图片并分类的功能。用户将需要进行识别的垃圾以图片形式输入,通过运行程序,用户可以对垃圾进行有效的识别并正确分类。

#### 2.问题缘起:

#### 2.1 垃圾分类政策

今年 1 月 31 日,上海市十五届人大二次会议表决通过了《上海市生活垃圾管理条例》,7 月 1 日起正式开始实施。

也就是说,还有不到一个月,在上海扔垃圾不分类要变违法事了!按照条例规定,个人混合投放垃圾,今后最高可罚 200 元;单位混装混运,最高则可罚 5 万元。

同时,上海党政机关内部办公场所不得使用一次性杯具,旅馆不得主动提供一次性日用品,餐馆、外卖不得主动提供一次性餐具。

但是以全球卓越城市为目标的上海、垃圾分类的道路可谓坎坷曲折。

a.试点阶段,1995-1998

阶段	年份	实施工作	分类方式
	1995	曹杨五村第七居委会的一个居住区启动垃圾分类试点	有机垃圾、无机垃圾、 有害垃圾; 废电池、 玻璃专项分类
试点阶段	1998	开展废电池、废玻璃专项分类回收	

b.推广阶段,1999-2006

	阶段	年份	实施工作	分类方式
		1999	垃圾分类工作纳入上海市环保三年行动计划,出台 《上海市区生活垃圾分类收藏、处置实施方案》等文 件	有机垃圾、无机垃圾、有害垃圾; 废电池、玻璃专项分类
		2000	首批100个小区启动垃圾分类试点,上海成为我国8个 垃圾分类试点城市之一	2000-2003年
Ħ	IA -> PA CII	2002	重点推进焚烧区垃圾分类工作	"有机垃圾、无机垃圾"调整为"干垃圾、湿垃圾"
	推广阶段	2006	全市有条件的居住区垃圾分类覆盖率超过60%	2003-2006年 焚烧区域:不可燃垃圾、有害垃圾、可燃垃圾 圾、可燃垃圾 其他区域:可堆肥垃圾、有害垃圾、其他垃圾

# c. 调整阶段, 2007-2013

阶段	年份	实施工作	分类方式
	2007	逐步推行垃圾四分类、五分类新方式	2007年-2010年
	2009	世博园区周边区域垃圾分类覆盖率达100%	居住区:有害垃圾、玻璃、可回收物、其他垃圾 办公场所:有害垃圾、可回收物、其他垃 圾
调整阶段	2010	全市有条件的居住区垃圾分类覆盖率超过70%	公共场所:可回收物、其他垃圾 其他:装修垃圾、大件垃圾、餐厨垃圾、 一次性塑料饭盒等实施专项收运、专项处 置。
	2011	"百万家庭低碳行,垃圾分类我先行", 1080个试点小区	2010-2013 大分流:装修垃圾、单位餐厨垃圾、大件 绿化枯枝落叶等 小分类:有害垃圾、玻璃、废旧衣物、湿 垃圾、其他干垃圾等

# d. 实施阶段, 2014-至今

阶段	年份	实施工作	分类方式	
	2014	《上海市促进生活垃圾分类减量办法》		
实施阶段	2017	《上海市单位生活垃圾强制分类实施方案》 可回收 圾、湿	可回收物、有害垃 圾、湿垃圾、干垃	
	2018	《关于建立完善本市生活垃圾全程分类体系的实施方案》	圾	

第一阶段:分类标准为"有机垃圾"、"无机垃圾"和"有毒有害垃圾"。

第二阶段:分类标准为"干垃圾"、"湿垃圾"和"有害垃圾"

第三阶段:分类标准为"干垃圾"、"湿垃圾"和"有害垃圾"

第四阶段:分类标准分出居住区和企事业单位两大类,前者按照"有害垃圾、玻璃、可回收

物、其他垃圾"四类分,后者按照"可回收物、其他垃圾"二类分

第五阶段:分类标准为"四分类":可回收物、有害垃圾、湿垃圾、干垃圾

#### 2.2 垃圾分类现状

#### a. 分类政策的不延续性造成居民困惑

垃圾分类的种类和名称的几经修改让居民产生分类行为上的困惑,也对垃圾分类的成功失去信心。

### b. 以信息提供为主、自上而下的,宣传方式传播效果不佳

社区开展垃圾分类,一般的宣传策略为宣传册发放、海报张贴或电子屏通知的方式。这种以信息提供为主,但往往不能取得很好的效果。

#### c. 垃圾分类设施不完善导致垃圾分类无法实施

垃圾桶数量不足、垃圾桶混用,使得居民源头分类好的垃圾无法分类投放,大大降低居民 参与的积极性。

#### d. 责任分工不明确导致垃圾的混装混运

很多居民认为垃圾分类是保洁员的责任。清运公司不负责任将干湿混装收运的行为也打击了很多认真做分类居民的积极性。

由此可见,虽然垃圾分类势在必行,但是在居民实际操作情况来看仍然存在比较大的问题,我们的程序设计方案针对日常生活中,居民并不明确各种垃圾的分类情况,提出以TensorFlow为基础的垃圾图片识别功能,并在用户输入图片之后告知用户该物品所属哪一类垃圾

#### 3. 垃圾分类程序介绍

#### 3.1 关于 TensorFlow

TensorFlow 是 Google 的开源深度学习库,你可以使用这个框架以及 Python 编程语言,构建大量基于机器学习的应用程序。深度学习让我们能够以极高的准确性构建复杂的应用程序。图像、视频、文本、音频等领域的问题,都可以通过深度学习解决。

我们使用 TensorFlow 进行图像分类/物体识别。通过识别需要分类的垃圾品种,我们进一步帮助用户更好的明确垃圾如何分类,并更加有效的实行垃圾分类。

Google 的 TensorFlow 图像识别系统是目前最准确的图像分类软件。所谓图像识别,就是对图片中的内容进行识别,然而这并非对任意图片都能识别。只有被训练过的对象,系统才能识别。

首先就需要创建一个 TensorFlow 图像分类器,之后再添加一些新的分类器。

# 3.2程序运行主要步骤

- a. 下载: 下载需要预训练的模型、图片和训练新分类器的脚本。
- b. 设置图像文件夹: 设置文件夹结构以便数据流图可以简单地拾取分类。
- c. 运行脚本形成新分类器: 输入脚本运行命令,进行模型预训练。脚本会分析每一张图片中的元素并生成只有 0 和 1 的二进制数据。之后程序会根据每一个分类文件夹中图片的特征点分析形成这个分类的图片的特征。输入的命令为: python retrain\_new.py --model\_dir E:\tfClassifier-master\image\_classification\inception --image\_dir

E:\tfClassifier-master\image classification\dataset

--output graph

E:\tfClassifier-master\image\_classification\output\_dir\output\_graph.pb

--output\_labels

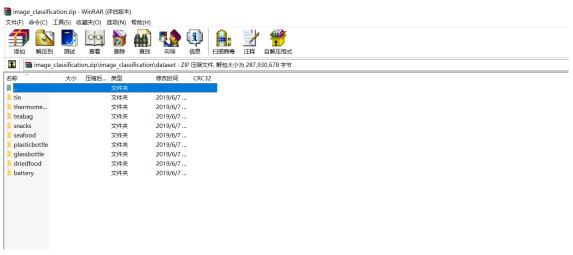
E:\tfClassifier-master\image\_classification\output\_dir\output\_labels.txt

--how\_many\_training\_steps 500 --learning\_rate 0.3 --testing\_percentage

--validation\_percentage 10

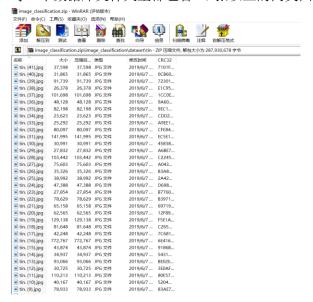
#### 3.3 具体展示

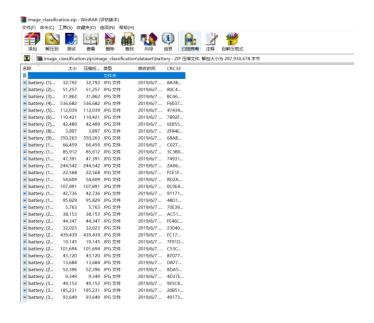
# 3.3.1 我们的数据库



10

每一个数据库文件夹里都包含 20 张以上的同类图片以训练机器识别,如:





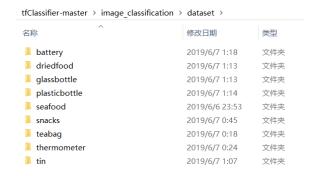
# 3.3.2 训练新的分类器

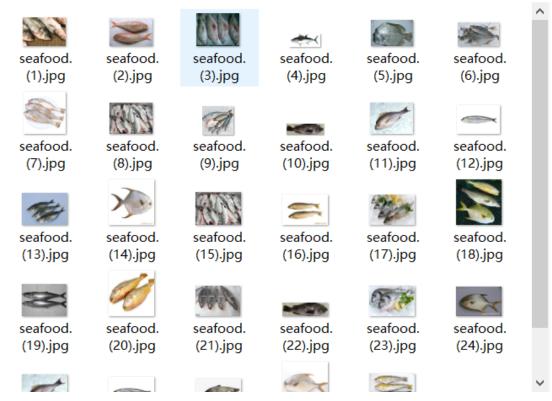
关于 tensorflow 重新训练图片识别分类:

如下所示

调用 retrain\_new.py

我们建立了一个数据库文件夹,文件夹下面的子文件夹里就是不同类别的图片





然后在系统的 cmd 命令窗口里 进入到 retrain new.py 所在的文件夹

# E:\tfClassifier-master\image\_classification>\_

我们参考了 Github 中的如下命令:

python retrain\_new.py --model\_dir --image\_dir--output\_graph --output\_labels --how\_many\_training\_steps 500 --learning\_rate 0.3 --testing\_percentage 10 --validation\_percentage 10

这些参数所对应的含义如下

model\_dir 参数: 指定了 model 的存放位置,就是我们的 inception 文件夹

image dir 参数: 指定了数据集的位置,就是我们的 dataset 文件夹

output\_graph 参数: 产生的 output\_graph.pb 文件的存放路径(后面会用到)

output\_labels 参数:产生的 output\_labels.txt 的存放路径(后面会用道)

how\_many\_training\_steps 参数:训练步数,和学习速率配合调整

learning\_rate 参数: 学习速率,和训练步数配合调整(常用的有 0.001, 0.01, 0.1, 0.3, 1, 3, 我们使用了 0.3)

testing\_percentage 参数:测试集比例

validation\_percentage 参数:验证集比例

我们根据自己文件实际存放位置修改的命令如下:

python retrain\_new.py --model\_dir E:\tfClassifier-master\image\_classification\inception

 $--image\_dir \\ E: \verb|\tfC| assifier-master \verb|\image\_c| assification \verb|\dataset| \\ --output\_graph$ 

E:\tfClassifier-master\image\_classification\output\_dir\output\_graph.pb

--output labels

E:\tfClassifier-master\image\_classification\output\_dir\output\_labels.txt

--how\_many\_training\_steps 500 --learning\_rate 0.3 --testing\_percentage 10 --validation\_percentage 10

然后, 那几个分类的识别就训练好了

之后在 output\_dir 文件夹里会生成 output\_graph.pb 和 output\_labels.txt。

tfClassifier-master > image_classification > output_dir			
修改日期	类型	大	
2018/1/5 4:52	文件		
2019/6/7 1:34	PB 文件		
2019/6/7 1:34	文本文档		
	修改日期 2018/1/5 4:52 2019/6/7 1:34	修改日期 类型 2018/1/5 4:52 文件 2019/6/7 1:34 PB 文件	

output\_graph.pb 对应不同图片分类的特征点,output\_labels.txt 则是不同分类的类别名称标签。

这两个文件在后续图片识别程序中会用到。

driedfood glassbottle plasticbottle seafood snacks teabag thermometer tin

# 3.3.3具体图片识别程序分步解析

我们运用的是 retrain\_model\_classifier1.py 在安装包中调用程序

```
import sys
import tensorflow as tf
```

录入带检测的图片的文件名,等待对比识别

这里使用的 try-except 过程可以在用户输入错误图片名称,或者图片位置放置错误时提醒用户。

#### 程序识别了新输入的图片,并与之前训练的分类器中特征点进行对比

#### 输出分类结果:

- a.如果机器识别对比出的结果显示被检测图片的特征点与数据库中已有图片的特征点相似度都小于百分之五十,则打印输出"this garbage is not among our sorts"
- b.如果机器识别对比出的结果显示被检测图片的特征点与塑料瓶数据库文件夹(plasticbottle)/玻璃瓶数据库文件夹(glassbottle)/易拉罐数据库文件夹(tin)高相似的话,打印输出"this is recyclable rubbish",指示图片中的垃圾属于可回收物。

如果机器识别对比出的结果显示被检测图片的特征点与废电池数据库文件夹(battery)/体温计(含汞)数据库文件夹(thermometer)高相似的话,打印输出"this is harmful rubbish",指示图片中的垃圾属于有害垃圾。

如果机器识别对比出的结果显示被检测图片的特征点与废茶包数据库文件夹(teabag)/干垃圾数据库文件夹(driedfood)/水产数据库文件夹(seafood)/零食数据库文件夹(snacks)高相似的话,此处直接归为 else 一类,打印输出"this is wet garbage",指示图片中的垃圾属于湿垃圾。

```
if standard<=0.5:
    print('this garbage is not among our sorts')
elif (sort=="plasticbottle") or (sort=="glassbottle") or (sort=="tin"):
    print('this is recyclable rubbish')
elif (sort=="battery") or (sort=="thermometer"):
    print('this is harmful rubbish')
else:
    print('this is wet garbage')
break
except:
    print('this picture does not exist')
    continue</pre>
```

#### 4. 项目缺陷与展望

此程序现阶段必须要把待检测的图片和图片库 (已做的 database) 放在一个文件夹里才能执行。

我们的展望是在实际应用时能用手机现场拍照即可检测,要实现这个更为便捷的目标, 我们预见需要涉及微信小程序以及用到 JavaScript 写网页的相关工具。

我们未来也可以再进一步扩大数据库,以此推广到对更多垃圾种类的识别,如盆栽植物、蔬果、废荧光灯管等,形成更完善、强大的识别。

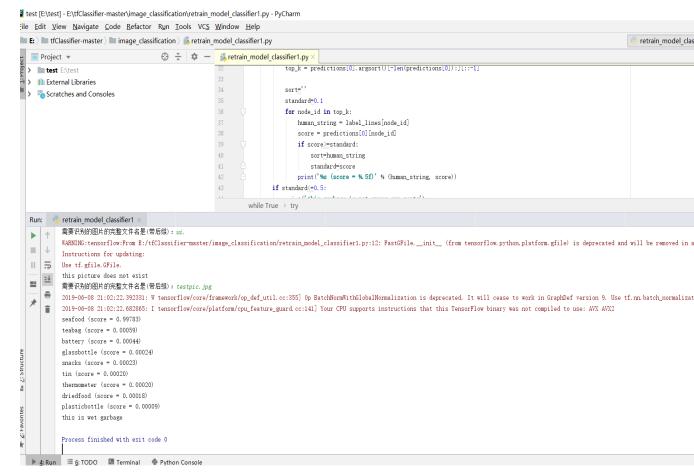
#### 5.成果与结论

# 5.1 成果展示

当我们放入一张如下的待检测图片时:



在 PyCharm 中的运行结果为:



需要识别的图片与 seafood 的特征点相似度最高,程序显示"this is wet garbage",指示图片中的垃圾属于湿垃圾。

#### 5.2 结论

这一次的小组作业让我们积极地思考与发现可应用 python 编程来解决的实际生活热点问题,将理论付诸于实践,进一步感受学科魅力。

在设计垃圾识别与分类这一程序的过程中,我们在已有的 python 编程基础上,走出舒适区,敢于挑战: 自学 Github 上开源文件,研究 Tensorflow 识别图片的功能……期间经历了许多困难和瓶颈,我们通过向老师寻求指导建议、网上查找相关资料与解决方法、小组探究和个人实践,最终完成了"垃圾识别与分类"这一程序。

过程中,小组同学的齐心协力献策献计,全力以赴地完成负责的部分并共同完善,和面对未知的识别领域的攻坚克难给我们带来的编程自信,以及应用编程解决实际问题的社会担当感,是我们最大的收获。