

第 8 讲 视觉应用场景认知

计算机视觉（Computer Vision）是一门研究如何使机器“看”的科学，它是使用计算机及相关设备对生物视觉的一种模拟。其主要任务就是通过对采集的图片或视频进行处理以获得相应场景的三维信息，就像人类和许多其他类生物每天所做的那样。形象地说，就是给计算机安装上眼睛（照相机）和大脑（算法），让计算机能够感知环境。进入 21 世纪，在各个行业都涌现出成熟的应用场景，包括计算机视觉技术，如面部识别、无人驾驶、智能安防、智慧医疗、智慧教育、智能制造等。

计算机理解图片：分类 classification、检测 detection、分割 segmentation。



是什么？



在哪里？



哪些像素组成？

8.1 图像分类

- 旨在判断该图像所属类别。
- 给定一张输入图像，解决“是什么？”的问题。
- 即是将图像结构化为一类信息，用事先确定好的标签（tag）、类别(category)或实例 ID 来描述图片。

8.2 目标检测



目标检测(object detection)，主要是解决“是什么？在哪里？”的问题。

- （1）目标定位任务：就是通常是以包围盒的(bounding box)形式进行标注。
- （2）目标检测任务：即定位出这个目标的位置并且知道目标物是什么，通常也会细分该任务可得到两个子任务，即目标检测，与目标识别。

8.3 图像分割



图像分割是对图像的像素级描述，它赋予每个像素类别（实例）意义，主要解决“每一个像素属于哪个目标物或场景”的问题，包括语义分割（semantic segmentation）和实例分割（instance segmentation）。



（1）语义分割是对前背景分离的拓展，要求分离开具有不同语义的图像部分，它是目标检测更进阶的任务，目标检测只需要框出每个目标的包围盒，语义分割需要进一步判断图像中哪些像素属于哪个目标。



（2）实例分割区分属于相同类别的不同实例。例如，当图像中有多只猫时，语义分割会将两只猫整体的所有像素预测为“猫”这个类别。与此不同的是，实例分割需要区分出哪些像素属于第一只猫、哪些像素属于第二只猫。

8.4 视频分析

视频分析技术就是使用计算机进行图像视觉分析的技术，是通过将场景中背景和目标分离进而分析并追踪在视频内出现的目标，从而识别目标显示结果的一项技术。



8.5 项目任务

8.5.1 图像分类应用——植物识别

实验摘要：根据拍摄照片，识别图片中植物的名称，可配合其它识图能力对识别的结果进一步细化，提高用户体验，广泛应用于拍照识图类 APP 中。

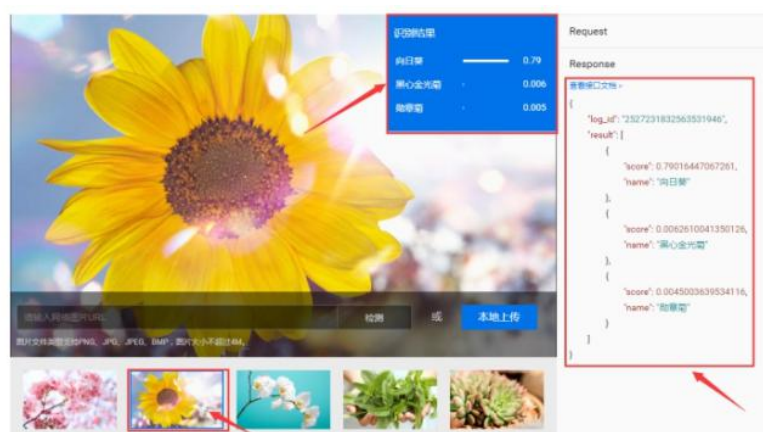
1. 访问植物识别应用

访问 <https://ai.baidu.com/tech/imagerecognition/plant>，然后点击功能演示：



2. 上传图片验证效果

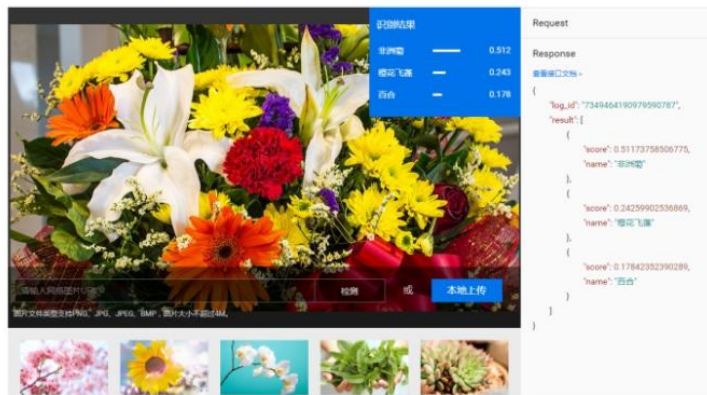
选择任意示例图片，即可见到相应的识别结果。如单击第二张图片，即可见到图片右上方显示其可能结果，其是向日葵的置信度为 0.79，是黑心金光菊的置信度为 0.006，是勋章菊的置信度为 0.005，对比可以见到植物识别功能可以正常识别，同时在最右侧可以见到接口返回的 Response 的 json 数据。



还可以使用本地图片进行验证，单击“本地上传”，上传本地图片，查看识别效果，如上传一张玫瑰的图片，即显示其为玫瑰的置信度为 0.797：

3. 尝试多目标测试

在实际应用中，也会遇到一张图片中有多种植物的情况，所以还需要测试一下多目标的情况识别是否准确，选择一张含有多种植物的图片进行上传，可以看到是非洲菊的置信度为 0.512，是橙花飞蓬的置信度为 0.243，是百合的置信度为 0.178。



4. 植物识别 API 接口调用（查看技术文档）

(1) 创建应用，获取 API Key 和 Secret Key：分别为 89Cb65osjaHdTjLRFASmfygR
UA3dNTmCa8xxbXXmAOf7EYTwcroqyKx，并领取免费资源。

The screenshot shows the 'Application List' (应用列表) page in the Baidu Cloud console. It contains a table with columns: 应用名称 (Application Name), AppID, API Key, Secret Key, 创建时间 (Creation Time), and 操作 (Actions). Two applications are listed. The first application, '视觉应用图像识别2022', is highlighted with a red border. The second application is 'TUXIANG'.

应用名称	AppID	API Key	Secret Key	创建时间	操作
视觉应用图像识别2022	26137160	89Cb65osjaHdTjLRFASmfygR	UA3dNTmCa8xxbXXmAOf7EYTwcroqyKx	2022-04-30 16:18:39	报表 管理 删除
TUXIANG	25403568	R31ow3GH4KTR3xVbCWTCdK Bz	***** 显示	2021-12-23 19:53:29	报表 管理 删除

(2) 以 pthon 为例，获取 access_token

浏览器：

https://aip.baidubce.com/oauth/2.0/token?grant_type=client_credentials&client_id=89Cb65osjaHdTjLRFASmfygR&client_secret=UA3dNTmCa8xxbXXmAOf7EYTwcroqyKx

或者运行以下程序：

```
# encoding:utf-8
import requests

# client_id 为官网获取的 AK， client_secret 为官网获取的 SK
host = 'https://aip.baidubce.com/oauth/2.0/token?grant_type=client_credentials&client_id=89Cb65osjaHdTjLRFASmfygR&client_secret=UA3dNTmCa8xxbXXmAOf7EYTwcroqyKx'

response = requests.get(host)

if response:
```



```
print(response.json())

{'refresh_token': '25.a6cd4f671d03626201940c823cd2429e.315360000.1966667875.282335-26137160', 'expires_in': 2592000, 'session_key': '9mzdCPOCuRKtkJyQspq/D0w31hLC8qIcDJ2/BWjqWUPQEhukGNLTctPJgpeZJtAgaW2TmFxQn2HHEv7EE100uXjXWvVGmw==', 'access_token': '24.37097b4c1cffd13878f6b31de7b6cbb5.2592000.1653899875.282335-26137160', 'scope': 'public vis-classify_dishes vis-classify_car brain_all_scope vis-classify_animal vis-classify_plant brain_object_detect brain_realtime_logo brain_dish_detect brain_car_detect brain_animal_classification brain_plant_classification brain_immediat brain_advanced_general_classification brain_quote'}
'24.37097b4c1cffd13878f6b31de7b6cbb5.2592000.1653899875.282335-26137160'
```

(3) 以 pthon 为例，调用接口程序

```
# encoding:utf-8
import requests
import base64
'''
植物识别
'''

request_url = "https://aip.baidubce.com/rest/2.0/image-classify/v1/plant"
# 二进制方式打开图片文件
f = open(r'D:\03 教学\21-22-01,02 计算机视觉应用开发\2021 讲义\flowers\1223.jpg', 'rb')
img = base64.b64encode(f.read())

params = {"image":img}
access_token = '24.37097b4c1cffd13878f6b31de7b6cbb5.2592000.1653899875.282335-26137160'
request_url = request_url + "?access_token=" + access_token
headers = {'content-type': 'application/x-www-form-urlencoded'}
response = requests.post(request_url, data=params, headers=headers)
if response:
    print(response.json())

{'result': [{'score': 0.80854875, 'name': '荷花'}, {'score': 0.5669627, 'name': '莲'}, {'score': 0.16101351, 'name': '睡莲'}], 'log_id': 1520322559046039450}
```

8.5.2 目标检测应用——图像主体检测

自行实验

8.5.3 图像分割应用——天空分割

自行实验

8.5.4 视频技术应用——视频内容分析