



周围的人都比你厉害，你才会慢慢变强

公告

昵称： 山上有风景
园龄： 1年11个月
粉丝： 170
关注： 19
+加关注

< 2019年12月 >

日	一	二	三	四	五	六
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	1	2	3	4
5	6	7	8	9	10	11

搜索

找找看

常用链接

- 我的随笔
- 我的评论
- 我的参与
- 最新评论
- 我的标签

我的标签

- STL(18)
- SDN(15)
- ThinkPHP3.2(1)

积分与排名

积分 - 211835
排名 - 1998

随笔分类

- C/C++(74)
- Html(2)
- Java(33)

OpenCV---数字验证码识别

目录

- 一：依赖环境安装
- 二：安装tesseract-ocr
 - (一) 介绍
 - (二) 下载地址
 - (三) 下载traineddata训练数据
- 三：代码实现

0

0

关注 | 顶部 | 评论

Javascript(19)
OpenCV(29)
PHP(2)
Python(155)
STL泛型编程(18)
单片机笔记 (复习用) (3)
计算机网络(32)
其他知识(14)
设计模式(27)
数据结构(57)
数据库(8)
算法习题(43)
算法训练营
随笔所想(4)
图形界面编程
正则表达式(2)
转载推文(4)

随笔档案

2019年11月(5)
2019年10月(32)
2019年9月(21)
2019年7月(10)
2019年5月(8)
2019年4月(25)
2019年3月(8)
2019年2月(1)
2019年1月(12)
2018年12月(19)
2018年9月(5)
2018年8月(95)
2018年7月(78)
2018年6月(26)
2018年5月(17)
2018年4月(22)
2018年3月(111)

最新评论

1. Re:python---websocket的使用

网上的方法都不行，换成gbk会报如下的错：

IndexError: index out of range

--缘分天空0320

2. Re:python---websocket的使用

@ 缘分天空0320太久没用，忘了。应该是字符编码问题吧。这类问题网上应该可以很容易找到方法解决。你看看Python的默认编码和代码是不是一致。一般就是gbk和utf8之间出错...

--山上有风景

案例实战-数字验证码识别

- ▶ [OpenCV+Tesseract-OCR](#)
- ▶ [OpenCV预处理](#)
- ▶ [Tesseract-OCR验证码识别](#)

知识点

- ▶ 预处理 - 去除干扰线与点
- ▶ 不同的结构元素中选择
- ▶ Image与numpy array相互转化
- ▶ 识别与输出

推文：Python验证码识别 安装Pillow、tesseract-ocr与pytesseract模块的安装以及错误解决

一：依赖环境安装

```
pip install Pillow
pip3 install pytesseract
```

二：安装tesseract-ocr

(一) 介绍

其中pytesseract会直接调用tesseract模块，我们需要进行安装

不然可会报错

```
pytesseract.TesseractNotFoundError: tesseract is not installed or it's not in your path
```

(二) 下载地址

github地址：<https://github.com/tesseract-ocr/tesseract>

0

0

关注 | 顶部 | 评论

3. Re:python---websocket的使用

您好，用了您的代码，报如下错误，麻烦问下如何解决呢？

Traceback (most recent call last): File "server3.py", line 101, in <module...

--缘分天空0320

4. Re:数据结构（六）查找---线性索引查找

请问最后倒序排序的那个代码怎么实现的？可以发一下吗？

--Viki-

5. Re:SDN实验---Ryu的源码分析

@ 山上有风景谢谢！ ...

--iRoy_33

阅读排行榜

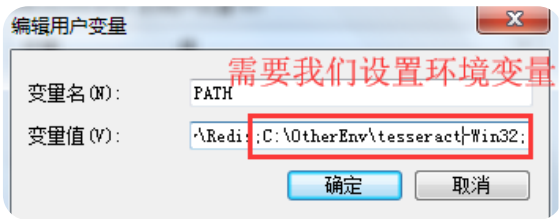
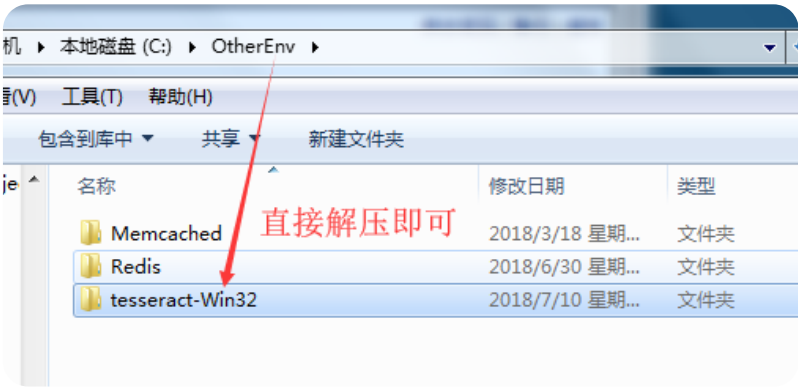
- 1. python---websocket的使用 (17253)
- 2. OpenCV---图像二值化(12513)
- 3. OpenCV---模板匹配matchTemplate(11534)
- 4. OpenCV---直线检测(8208)
- 5. python---基础知识回顾 (九) 图形用户界面-----WxPython(7986)

评论排行榜

- 1. python---基础知识回顾 (九) 图形用户界面-----Tkinter(4)
- 2. python---websocket的使用 (3)
- 3. SDN实验---Ryu的源码分析 (3)
- 4. 数据结构（三）串---KMP模式匹配算法之获取next数组(2)
- 5. 数据结构（四）树---树的存储结构(1)

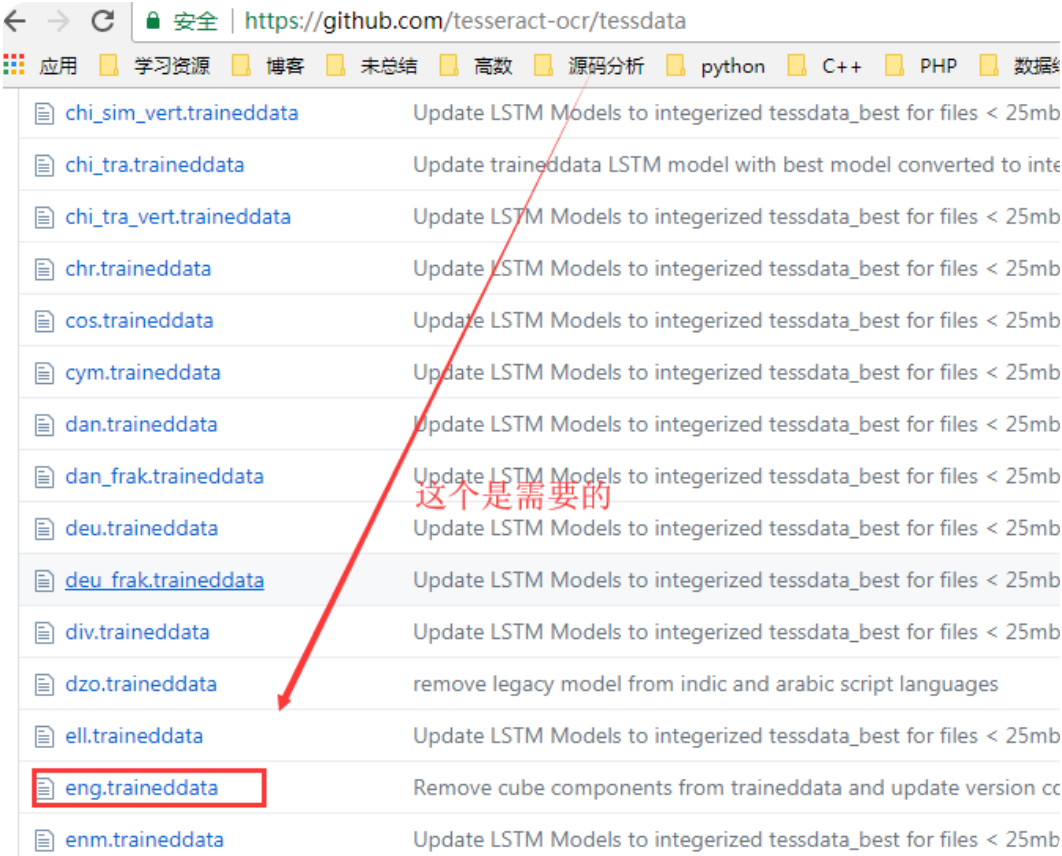
推荐排行榜

- 1. 数据结构（七）排序---堆排序(11)
- 2. python---aiohttp的使用(6)
- 3. python---websocket的使用 (4)
- 4. python---基础知识回顾 (九) 图形用户界面-----WxPython(3)
- 5. Python图像处理库PIL中图像格式转换(3)



(三) 下载traineddata训练数据

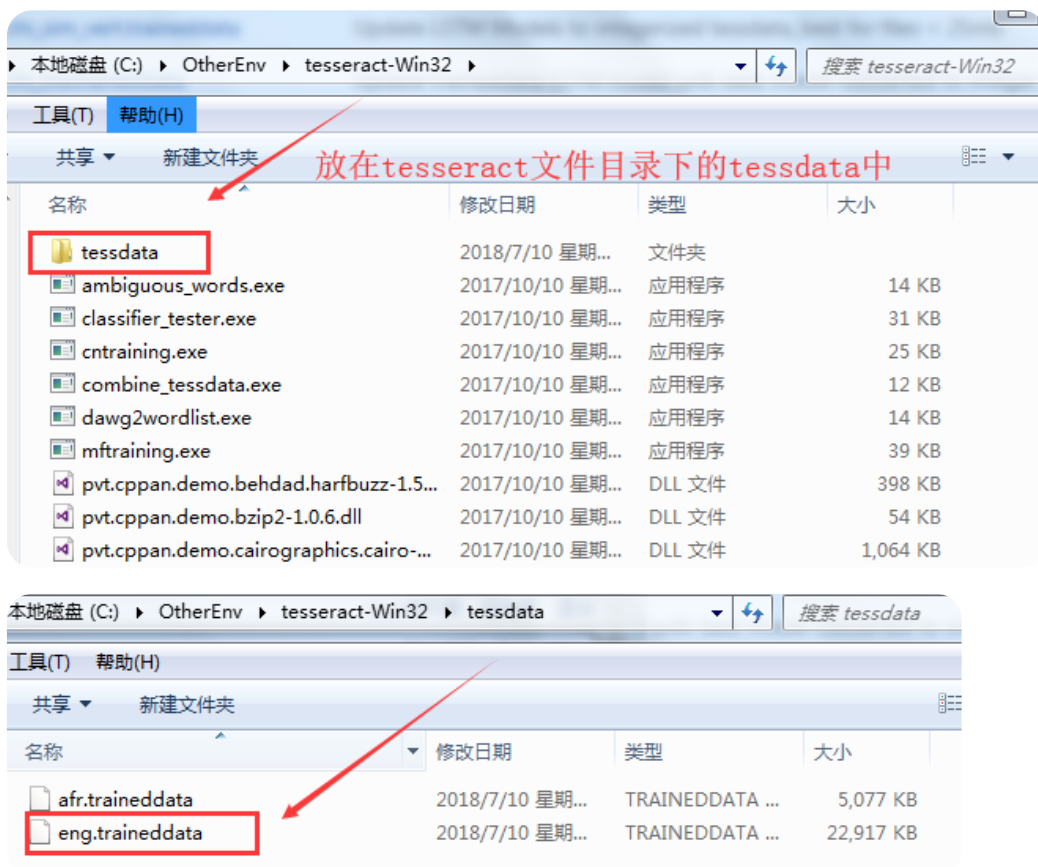
github地址: <https://github.com/tesseract-ocr/tessdata>



0

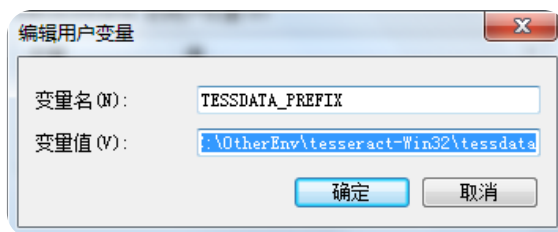
0

关注 | 顶部 | 评论



注意：我们还是要将其设置环境变量

```
pytesseract.TesseractError: (1, 'Error opening data file \\OtherEnv\\tesseract-Win32\\tessdata\\eng.traineddata Please make sure the TESSDATA_PREFIX environment variable is set to the parent directory of your "tessdata" directory. Failed loading language \'eng\' Tesseract couldn\'t load any languages! Could not initialize tesseract.')
```



现在我们重新启用cmd命令行，可以在cmd命令行调用python文件，获取到验证数据

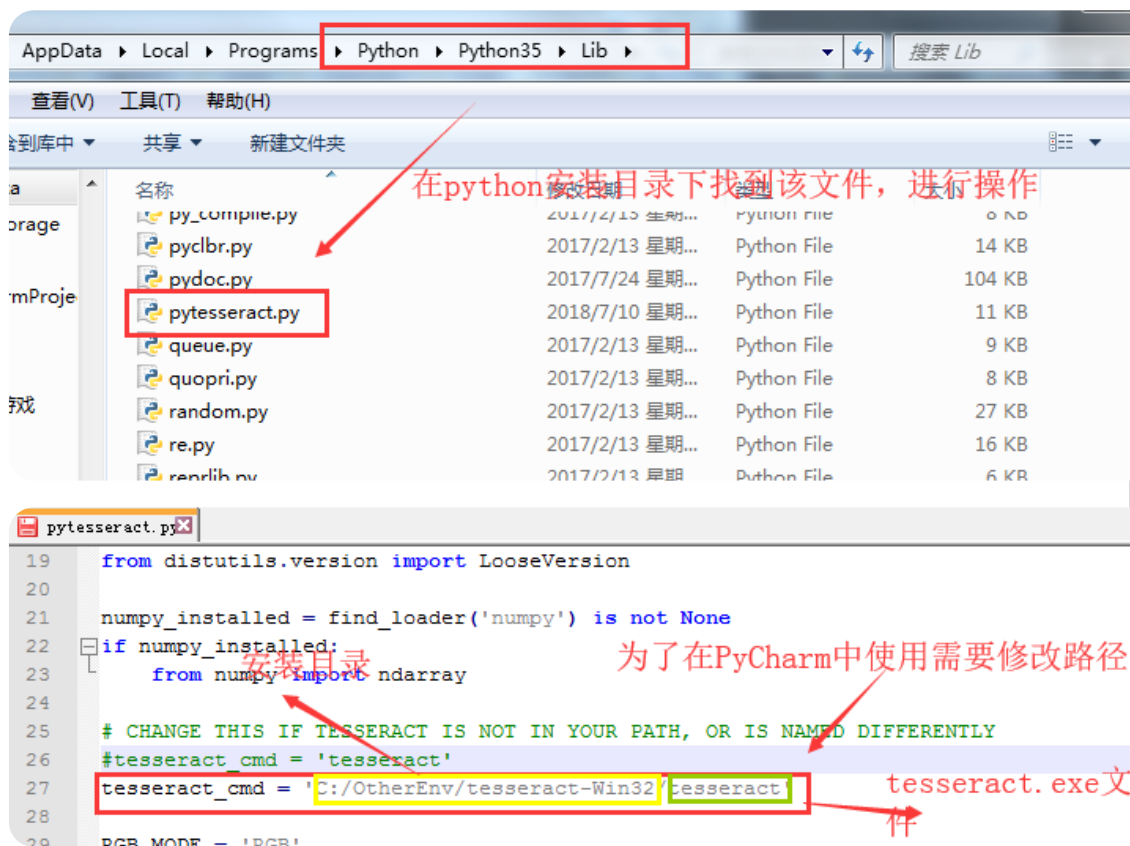
```
D:\MyPython\day25\pv>python op21.py
result:PF8124
```

但是我们在PyCharm中使用时还是需要修改python文件

0

0

关注 | 顶部 | 评论



还有在我们的程序文件中加入环境变量

```
os.environ['TESSDATA_PREFIX'] = "C:/OtherEnv/tesseract-Win32/tessdata"
```

三：代码实现

```
import cv2 as cv
import numpy as np
from PIL import Image
import os
import pyesseract as tess

os.environ['TESSDATA_PREFIX'] = "C:/OtherEnv/tesseract-Win32/tessdata"

def recognize_text(image):
    gray = cv.cvtColor(image, cv.COLOR_BGR2GRAY)
    ret, binary = cv.threshold(gray, 0, 255, cv.THRESH_BINARY_INV | cv.THRESH_OTSU)
    kernel = cv.getStructuringElement(cv.MORPH_RECT, (1, 2))
    mid1 = cv.morphologyEx(binary, cv.MORPH_OPEN, kernel)
    kernel = cv.getStructuringElement(cv.MORPH_RECT, (2, 1))
    open_out = cv.morphologyEx(mid1, cv.MORPH_OPEN, kernel)
    cv.imshow("bin1", open_out)

    cv.bitwise_not(open_out, open_out)  # 变白色背景
    textImage = Image.fromarray(open_out)

    text = tess.image_to_string(textImage)
    print("result:%s"%text)

src = cv.imread("./y4.png")  # 读取图片
cv.namedWindow("input image", cv.WINDOW_AUTOSIZE)  # 创建GUI窗口, 形式为窗口名, 大小, 标志
```

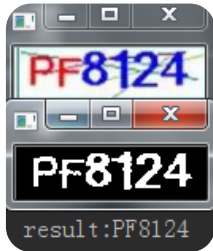
0 0

关注 | 顶部 | 评论

```
cv.imshow("input image",src)    #通过名字将图像和窗口联系

recognize_text(src)

cv.waitKey(0)    #等待用户操作，里面等待参数是毫秒，我们填写0，代表是永远，等待用户操作
cv.destroyAllWindows()    #销毁所有窗口
```



作者：山上有风景

欢迎任何形式的转载，但请务必注明出处。

限于本人水平，如果文章和代码有表述不当之处，还请不吝赐教。

分类： OpenCV

好文要顶

关注我

收藏该文



山上有风景

关注 - 19

粉丝 - 170

+加关注

« 上一篇： OpenCV---人脸检测

» 下一篇： 数据分析与展示---anaconda的使用

posted @ 2018-07-10 10:51 山上有风景 阅读(1411) 评论(0) 编辑 收藏

[刷新评论](#) [刷新页面](#) [返回顶部](#)

注册用户登录后才能发表评论，请 [登录](#) 或 [注册](#)， [访问](#) 网站首页。

Copyright © 2019 山上有风景

Powered by .NET Core 3.1.0 on Linux

0

0

[关注](#) | [顶部](#) | [评论](#)