

# 快速识别文字图片

## 一、应用场景

在日常的生活或工作中，我们有时需要识别文字图片，并将图片中的文字提取出来，比如手机的屏幕截图，文档截图等。我们现在使用python和百度云API接口来操作并识别简单的文字图片，将文字提取出来。

## 二、项目代码及结果展示

### 2.1 项目代码

```
1 # encoding:utf-8
2
3 # 权限认证文档代码示例修改，获取access_token
4 import requests
5
6 # client_id 为官网获取的AK， client_secret 为官网获取的SK
7 AK = 【百度云应用的AK】 # 传入自己的AK
8 SK = 【百度云应用的SK】 # 传入自己的SK
9
10 host = 'https://aip.baidubce.com/oauth/2.0/token?grant_type=client_credentials&client_id=%s'
11 response = requests.get(host)
12 if response:
13     print(response.json()) # 返回信息中含有用户特定的信息 (access_token值)
14
15
16 # 文字识别文档代码示例修改，获取图中文字
17 import requests
18 import base64
19
20 request_url = "https://aip.baidubce.com/rest/2.0/ocr/v1/accurate_basic"
21 # 二进制方式打开图片文件
22 f = open(r'【本地文件路径】', 'rb') # 传入自己的文件路径
23 img = base64.b64encode(f.read())
24
25 params = {"image":img}
```

```
26 access_token = response.json()["access_token"] # 获取用户信息access_token
27 request_url = request_url + "?access_token=" + access_token # 加入用户信息
28 headers = {'content-type': 'application/x-www-form-urlencoded'}
29 contents = requests.post(request_url, data=params, headers=headers)
30 if contents:
31     print(contents.json())
32     # 遍历打印内容
33     for content in contents.json()['words_result']:
34         print('识别出来的内容为: \n',content['words'])
```

注：为了方便读者根据官方文档内容进行理解，该代码使用两端拼合的形式展示。

## 2.2 结果展示

示例图片：

人生苦短，  
我用 Python。

识别结果：

人生苦短，  
我用 Python

### 三、项目操作及解析

该代码是调用百度人工智能中文字识别的API，然后实现对图片的识别。由于篇幅问题，很多具体的教程需要自行查看官方文档，此处做一个引导。更多API可查看附录三。

#### 步骤一：注册并接入AI

在调用API时，需要传入access\_token，而获取access\_token，需要使用到API Key和Secret Key。获取这两个参数需要先注册和创建应用，具体操作可查看AI接入指南：<https://ai.baidu.com/ai-doc/REFERENCE/Ck3dwjgn3> <<https://ai.baidu.com/ai-doc/REFERENCE/Ck3dwjgn3>>

#### 步骤二：获取access\_token

创建应用之后，可以获取到API Key和Secret Key，然后参考获取access\_token的文档。

文档链接：<https://ai.baidu.com/ai-doc/REFERENCE/Ck3dwjhhu> <<https://ai.baidu.com/ai-doc/REFERENCE/Ck3dwjhhu>>

文档中的请求代码示例，传入这两个值便可获取到access\_token，具体代码应用见项目代码中的第113行。

#### 步骤三：接入API，识别图片

获取到access\_token之后，可以查看官方文档，根据说明编写代码，实现对图片的识别。

文字识别文档官方链接: <https://cloud.baidu.com/doc/OCR/s/1k3h7y3db>  
<<https://cloud.baidu.com/doc/OCR/s/1k3h7y3db>>

查看文档，文档中重点关注几个板块：接口地址（请求URL）、URL参数、请求参数、请求代码示例、返回参数。

- 接口地址一般是一个url地址，该项目需要传入用户的信息（URL参数）；
- URL参数传入的值一般含有用户信息，可以在网站上获取；
- 请求参数是根据自己需求传入的参数，比如识别不同国家的语言文字，需要传入对应语言的值；
- 返回的参数则是中含有识别结果字符串words，可以根据提示打印出识别的结果文字。

根据文档中的请求代码示例填写自己的文件路径和access\_token，运行即可，具体代码应用见项目代码中的第16~34行。

注意：在实际的使用中，顺序是逆向的，一般是看到文字识别的API文档，发现需要传入access\_token，跳转到获取access\_token的文档，然后看文档需求，发现需要用户的个人信息，然后找到如何获取个人信息，接入网站的文档，获取到个人的关键信息，最终再整合代码，完成整个项目的代码。