

---

## 利用 Python 实现简单图片下载功能

我们有时候会需要在网上查找并下载图片，当数量比较少的时候，点击右键保存，很轻松就可以实现图片的下载，但是有些图片进行了特殊设置，点击右键没有显示保存选项，或者需要下载很多图片，这样的情况，写一段 Python 爬虫代码就可以轻松解决！

### 一、页面抓取

```
#coding=utf-8

import urllib

def getHtml(url):

    page = urllib.urlopen(url)

    html = page.read()

    return html

html = getHtml("https://tieba.baidu.com/p/5582243679")

print html
```

页面数据抓取过程定义了 `getHtml()` 函数，其作用是给 `getHtml()` 传递一个网址，最终进行整个页面的下载。

### 二、页面数据筛选

```
import re

import urllib

def getHtml(url):

    page = urllib.urlopen(url)
```

---

```
    html = page.read()

    return html

def getImg(html):

    reg = r' src="(.\+?.jpg)" pic_ext'

    imgre = re.compile(reg)

    imglist = re.findall(imgre, html)

    return imglist

html = getHtml("https://tieba.baidu.com/p/5582243679")

print getImg(html)
```

页面数据筛选中，定义了一个新的函数 `getImg()`，该函数的功能是筛选出 .jpg 格式的图片地址。

### 三、图片下载

```
#coding=utf-8

import urllib

import re

def getHtml(url):

    page = urllib.urlopen(url)

    html = page.read()

    return html

def getImg(html):

    reg = r' src="(.\+?.jpg)" pic_ext'

    imgre = re.compile(reg)
```

---

```
imglist = re.findall(imgre, html)

x = 0

for imgurl in imglist:

    urllib.urlretrieve(imgurl, '%s.jpg' % x)

    x+=1

html = getHtml("https://tieba.baidu.com/p/5582243679")

print getImg(html)
```

通过 for 循环获得所有符合条件的图片网址, 并采用 `urllib.urlretrieve()` 方法, 将远程数据下载到本地, 并重新命名!