[Web逆向] 逆向世界某读切块图片链接与Python还原切割图片

发表于 2024-1-13 10:20 | 只看该作者 精华 楼主 电梯直达

本帖最后由 T4DNA 于 2024-1-13 15:28 编辑

免责声明

1、本贴仅作为技术讨论，本人不会利用以下技术盈利、或从事任何侵害该网站的行为。
2、如觉得此贴不妥，请联系本人将第一时间删除。

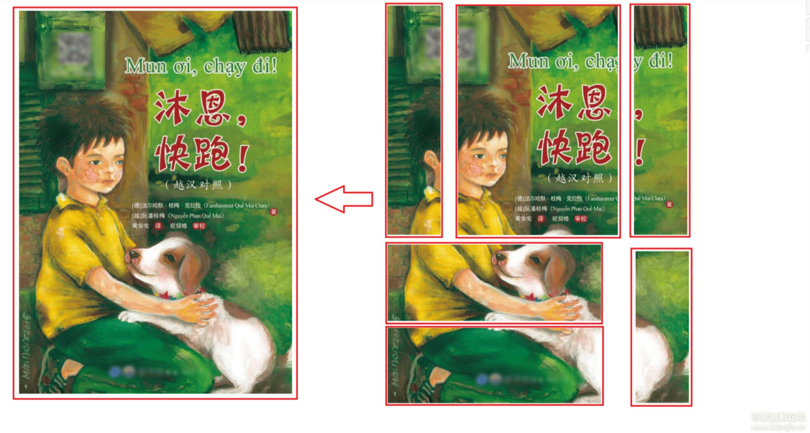
背景

网站链接："aHR0cHM6Ly9hcHAucmVhZG9vci5jbi9hcHAvZHQvcGQvMTU2NTgzNzEzOT9zPTE="

背景：

该网站是某某电子书平台，展示的电子书效果是分割图片在前端页面还原，也是一个老生常谈的反爬措施了。

与龙X期刊的整齐5*5方块，或是XX学堂5*1乱序不同，该网站采用了新颖的（对我来说）不规则方块（乱序）。

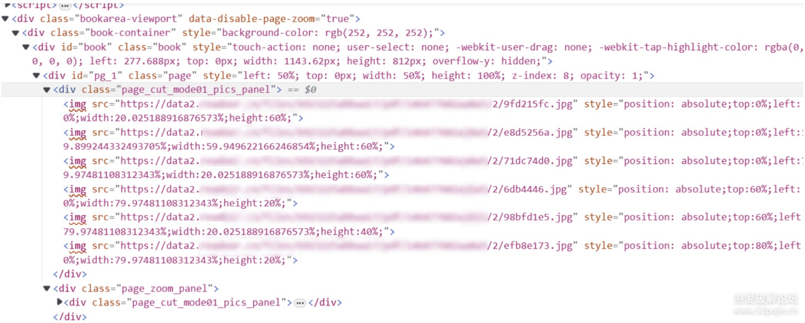

Mun ơi, chạy đi!
沐恩，快跑！
（越汉对照）
吾爱破解论坛
www.52pojie.cn

逆向过程

吸取了https://www.52pojie.cn/thread-1871407-1-1.html的教训，我们先盘逻辑，再谈解密。

请求逻辑

刷新页面，查看图片，他是以Img标签载入HTML的，通过CSS排列位置，所以首要任务是拿到Img的src是如何生成与其对应的位置信息。



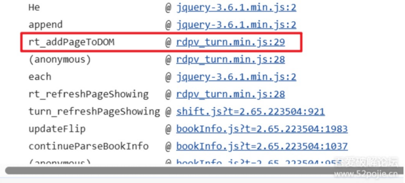
```
▶<script>…</script>
▼<div class="bookarea-viewport" data-disable-page-zoom="true">
  ▼<div class="book-container" style="background-color: rgb(252, 252, 252);">
    ▼<div id="book" class="book" style="touch-action: none; user-select: none; -webkit-user-drag: none; -webkit-tap-highlight-color: rgba(0,
      0, 0, 0); left: 277.688px; top: 0px; width: 1143.62px; height: 812px; overflow-y: hidden;">
      ▼<div id="pg_1" class="page" style="left: 50%; top: 0px; width: 50%; height: 100%; z-index: 8; opacity: 1;">
        ▼<div class="page_cut_mode01_pics_panel"> == $0
            <img src="https://data2.[redacted]2/9fd215fc.jpg" style="position: absolute;top:0%;left:
            0%;width:20.025188916876573%;height:60%;">
            <img src="https://data2.[redacted]/2/e8d5256a.jpg" style="position: absolute;top:0%;left:
            9.899244332493705%;width:59.949622166246854%;height:60%;">
            <img src="https://data2.[redacted]/2/71dc74d0.jpg" style="position: absolute;top:0%;left:
            9.97481108312343%;width:20.025188916876573%;height:60%;">
            <img src="https://data2.[redacted]/2/6db4446.jpg" style="position: absolute;top:60%;left:
            0%;width:79.97481108312343%;height:20%;">
            <img src="https://data2.[redacted]/2/98bfd1e5.jpg" style="position: absolute;top:60%;left
            79.97481108312343%;width:20.025188916876573%;height:40%;">
            <img src="https://data2.[redacted]/2/efb8e173.jpg" style="position: absolute;top:80%;left
            0%;width:79.97481108312343%;height:20%;">
          </div>
        ▼<div class="page_zoom_panel">
          ▶<div class="page_cut_mode01_pics_panel">…</div>
          </div>
```

吾爱破解论坛
www.52pojie.cn

查看网络，第一个引人注意的是这个addPageToDom字面意思，是加载页面的函数。进入，下断点，在网页上进入下一页。

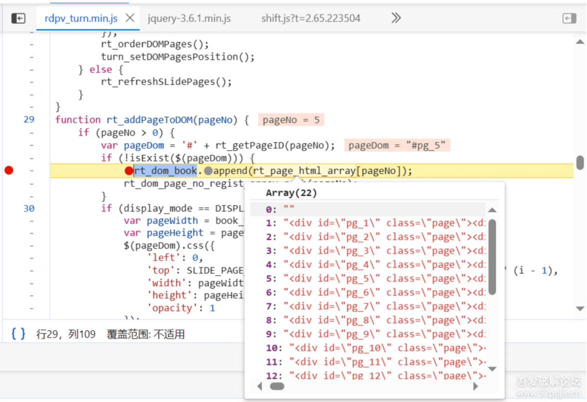
标头
预览
响应
发起程序
计时
Cookie

```
xe                    @ jquery-3.6.1.min.js:2
```



```
He                          @ jquery-3.6.1.min.js:2
append                      @ jquery-3.6.1.min.js:2
rt_addPageToDOM             @ rdpv_turn.min.js:29
(anonymous)                 @ rdpv_turn.min.js:28
each                        @ jquery-3.6.1.min.js:2
rt_refreshPageShowing       @ rdpv_turn.min.js:28
turn_refreshPageShowing     @ shift.js?t=2.65.223504:921
updateFlip                  @ bookInfo.js?t=2.65.223504:1983
continueParseBookInfo       @ bookInfo.js?t=2.65.223504:1037
```

吾爱破解论坛 www.52pojie.cn

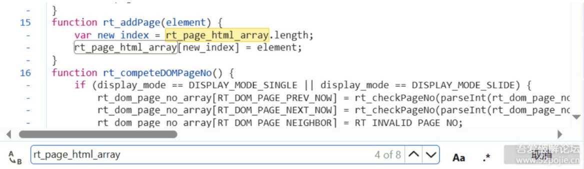


```
        });
        rt_orderDOMPages();
        turn_setDOMPagesPosition();
    } else {
        rt_refreshSLidePages();
    }
}
function rt_addPageToDOM(pageNo) {  pageNo = 5
    if (pageNo > 0) {
        var pageDom = '#' + rt_getPageID(pageNo);  pageDom = "#pg_5"
        if (!isExist($(pageDom))) {
            rt_dom_book.append(rt_page_html_array[pageNo]);
            rt_dom_page_no_regist
        }
        if (display_mode == DISPL
            var pageWidth = book_
            var pageHeight = page
            $(pageDom).css({
                'left': 0,
                'top': SLIDE_PAGE                                  (i - 1),
                'width': pageWidt
                'height': pageHei
                'opacity': 1
```

```
Array(22)
0: ""
1: "<div id=\"pg_1\" class=\"page\"><d
2: "<div id=\"pg_2\" class=\"page\"><d
3: "<div id=\"pg_3\" class=\"page\"><d
4: "<div id=\"pg_4\" class=\"page\"><d
5: "<div id=\"pg_5\" class=\"page\"><d
6: "<div id=\"pg_6\" class=\"page\"><d
7: "<div id=\"pg_7\" class=\"page\"><d
8: "<div id=\"pg_8\" class=\"page\"><d
9: "<div id=\"pg_9\" class=\"page\"><d
10: "<div id=\"pg_10\" class=\"page\">
11: "<div id=\"pg_11\" class=\"page\">
12: "<div id=\"pg_12\" class=\"page\">
```

吾爱破解论坛 www.52pojie.cn

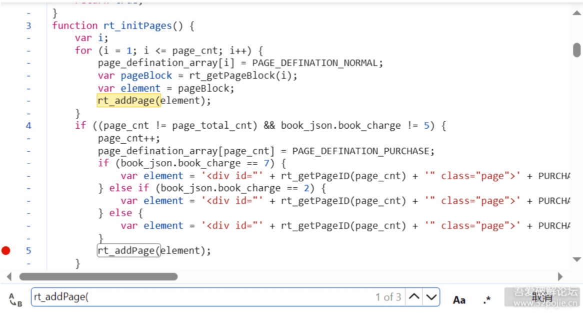
可以看到rt_page_html_array已经是写好准备创建的html对象，包含了所有图片链接，那么目标就是他如何生成，看到这儿我们可以发现这位非常老实，函数名和变量名都没有混淆，所以直接搜索即可。



```
}
function rt_addPage(element) {
    var new_index = rt_page_html_array.length;
    rt_page_html_array[new_index] = element;
}
function rt_competeDOMPageNo() {
    if (display_mode == DISPLAY_MODE_SINGLE || display_mode == DISPLAY_MODE_SLIDE) {
        rt_dom_page_no_array[RT_DOM_PAGE_PREV_NOW] = rt_checkPageNo(parseInt(rt_dom_page_nc
        rt_dom_page_no_array[RT_DOM_PAGE_NEXT_NOW] = rt_checkPageNo(parseInt(rt_dom_page_nc
        rt_dom_page_no_array[RT_DOM_PAGE_NEIGHBOR] = RT_INVALID_PAGE_NO;
```

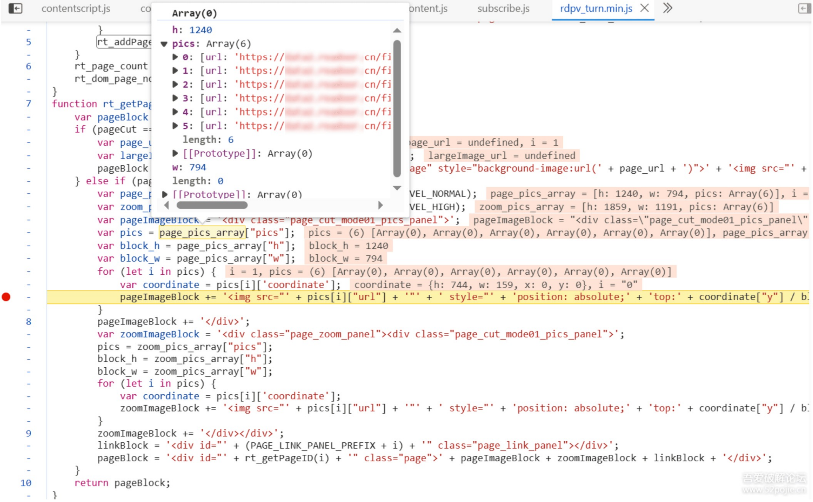
吾爱破解论坛 www.52pojie.cn

继续搜索rt_addPage、rt_getPageBlock直到打上断点后看到未经修饰的链接，说明你找到了



```
}
function rt_initPages() {
    var i;
    for (i = 1; i <= page_cnt; i++) {
        page_defination_array[i] = PAGE_DEFINATION_NORMAL;
        var pageBlock = rt_getPageBlock(i);
        var element = pageBlock;
        rt_addPage(element);
    }
    if ((page_cnt != page_total_cnt) && book_json.book_charge != 5) {
        page_cnt++;
        page_defination_array[page_cnt] = PAGE_DEFINATION_PURCHASE;
        if (book_json.book_charge == 7) {
            var element = '<div id="' + rt_getPageID(page_cnt) + '" class="page">' + PURCHA
        } else if (book_json.book_charge == 2) {
            var element = '<div id="' + rt_getPageID(page_cnt) + '" class="page">' + PURCHA
        } else {
            var element = '<div id="' + rt_getPageID(page_cnt) + '" class="page">' + PURCHA
        }
        rt_addPage(element);
    }
```

吾爱破解论坛 www.52pojie.cn

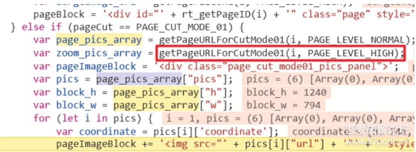


```
Array(0)
  h: 1240
▼ pics: Array(6)
  ▶ 0: [url: 'https://[redacted]cn/fi
  ▶ 1: [url: 'https://[redacted]cn/fi
  ▶ 2: [url: 'https://[redacted]cn/fi
  ▶ 3: [url: 'https://[redacted]cn/fi
  ▶ 4: [url: 'https://[redacted]cn/fi
  ▶ 5: [url: 'https://[redacted]cn/fi
    length: 6
  ▶ [[Prototype]]: Array(0)
  w: 794
  length: 0
▶ [[Prototype]]: Array(0)
```

```
            }
            rt_addPage
        }
        rt_page_count
        rt_dom_page_nc
    }
    function rt_getPag
        var pageBlock
        if (pageCut ==
            var page_u                                                 page_url = undefined, i = 1
            var largeI                                                 ;  largeImage_url = undefined
            pageBlock                                                  age" style="background-image:url(' + page_url + ')">' + '<img src="' +
        } else if (pag
            var page_p                                                 VEL_NORMAL);  page_pics_array = [h: 1240, w: 794, pics: Array(6)], i =
            var zoom_p                                                 VEL_HIGH);  zoom_pics_array = [h: 1859, w: 1191, pics: Array(6)]
            var pageImageBlock = '<div class="page_cut_mode01_pics_panel">';  pageImageBlock = "<div class=\"page_cut_mode01_pics_panel\"
            var pics = page_pics_array["pics"];  pics = (6) [Array(0), Array(0), Array(0), Array(0), Array(0), Array(0)], page_pics_array
            var block_h = page_pics_array["h"];  block_h = 1240
            var block_w = page_pics_array["w"];  block_w = 794
            for (let i in pics) {  i = 1, pics = (6) [Array(0), Array(0), Array(0), Array(0), Array(0), Array(0)]
                var coordinate = pics[i]['coordinate'];  coordinate = {h: 744, w: 159, x: 0, y: 0}, i = "0"
                pageImageBlock += '<img src="' + pics[i]["url"] + '"' + ' style="' + 'position: absolute;' + 'top:' + coordinate["y"] / bl
            }
            pageImageBlock += '</div>';
            var zoomImageBlock = '<div class="page_zoom_panel"><div class="page_cut_mode01_pics_panel">';
            pics = zoom_pics_array["pics"];
            block_h = zoom_pics_array["h"];
            block_w = zoom_pics_array["w"];
            for (let i in pics) {
                var coordinate = pics[i]['coordinate'];
                zoomImageBlock += '<img src="' + pics[i]["url"] + '"' + ' style="' + 'position: absolute;' + 'top:' + coordinate["y"] / bl
            }
            zoomImageBlock += '</div></div>';
            linkBlock = '<div id="' + (PAGE_LINK_PANEL_PREFIX + i) + '" class="page_link_panel"></div>';
            pageBlock = '<div id="' + rt_getPageID(i) + '" class="page">' + pageImageBlock + zoomImageBlock + linkBlock + '</div>';
        }
        return pageBlock;
    }
```

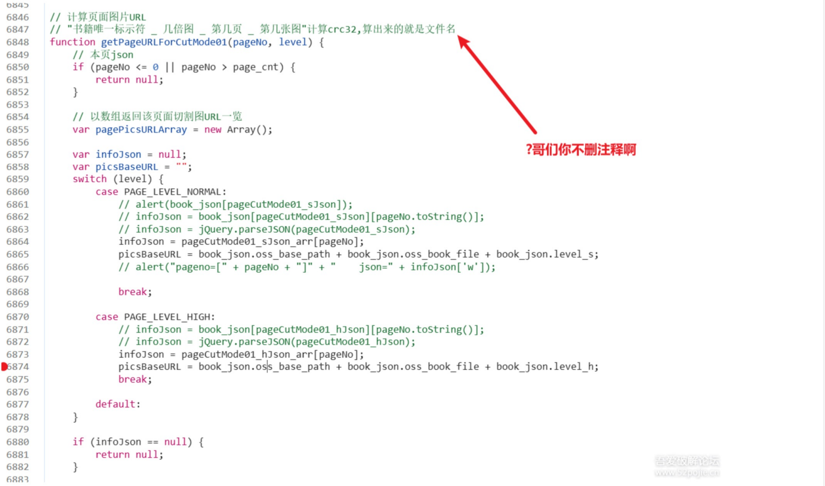
吾爱破解论坛
www.52pojie.cn

名字太直白，显然第二个PAGE_LEVEL_HIGH是更高质量的分割图片获取，它们都调用了一个getPageURLForCutModel01，这也太直白了，就是获取1类分割图片的页面链接（这个帖子要变成英语翻译了）。



```
        pageBlock = '<div id="' + rt_getPageID(i) + '" class="page" style='
    } else if (pageCut == PAGE_CUT_MODE_01) {
        var page_pics_array = getPageURLForCutMode01(i, PAGE_LEVEL_NORMAL);
        var zoom_pics_array = getPageURLForCutMode01(i, PAGE_LEVEL_HIGH);
        var pageImageBlock = '<div class="page_cut_mode01_pics_panel">';
        var pics = page_pics_array["pics"];  pics = (6) [Array(0), Array(0
        var block_h = page_pics_array["h"];  block_h = 1240
        var block_w = page_pics_array["w"];  block_w = 794
        for (let i in pics) {  i = 1, pics = (6) [Array(0), Array(0), Array
            var coordinate = pics[i]['coordinate'];  coordinate = {h: 744,
            pageImageBlock += '<img src="' + pics[i]["url"] + '"' + ' style
```

吾爱破解论坛
www.52pojie.cn

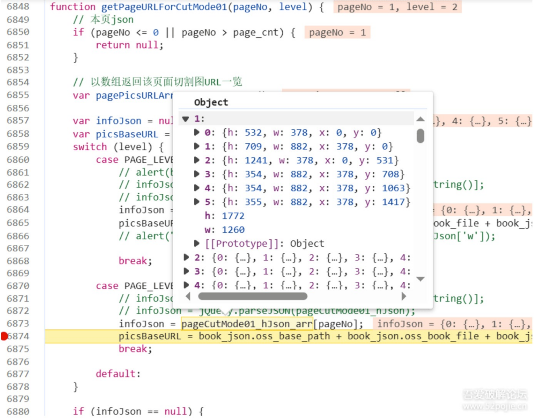


```
// 计算页面图片URL
// "书籍唯一标示符 _ 几倍图 _ 第几页 _ 第几张图"计算crc32,算出来的就是文件名
function getPageURLForCutMode01(pageNo, level) {
    // 本页json
    if (pageNo <= 0 || pageNo > page_cnt) {
        return null;
    }

    // 以数组返回该页面切割图URL一览
    var pagePicsURLArray = new Array();

    var infoJson = null;
    var picsBaseURL = "";
    switch (level) {
        case PAGE_LEVEL_NORMAL:
            // alert(book_json[pageCutMode01_sJson]);
            // infoJson = book_json[pageCutMode01_sJson][pageNo.toString()];
            // infoJson = jQuery.parseJSON(pageCutMode01_sJson);
            infoJson = pageCutMode01_sJson_arr[pageNo];
            picsBaseURL = book_json.oss_base_path + book_json.oss_book_file + book_json.level_s;
            // alert("pageno=[" + pageNo + "]" + "    json=" + infoJson['w']);

            break;

        case PAGE_LEVEL_HIGH:
            // infoJson = book_json[pageCutMode01_hJson][pageNo.toString()];
            // infoJson = jQuery.parseJSON(pageCutMode01_hJson);
            infoJson = pageCutMode01_hJson_arr[pageNo];
            picsBaseURL = book_json.oss_base_path + book_json.oss_book_file + book_json.level_h;
            break;

        default:
    }

    if (infoJson == null) {
        return null;
    }
```

?哥们你不删注释啊

吾爱破解论坛
www.52pojie.cn

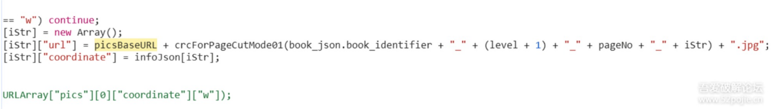
图片位置则是在pageCutMode01_hJson_arr中存储，而图片链接是crc32直接算出来的。



```
function getPageURLForCutMode01(pageNo, level) {  pageNo = 1, level = 2
    // 本页json
    if (pageNo <= 0 || pageNo > page_cnt) {  pageNo = 1
        return null;
    }

    // 以数组返回该页面切割图URL一览
    var pagePicsURLArr
    
    var infoJson = nul                              }, 4: {…}, 5: {…}
    var picsBaseURL =
    switch (level) {
        case PAGE_LEVE
            // alert(t
            // infoJsc                              tring()];
            // infoJsc
            infoJson =                              {0: {…}, 1: {…},
            picsBaseUF                              ook_file + book_j
            // alert('                              Json['w']);

            break;

        case PAGE_LEVE
            // infoJsc                              tring()];
            // infoJson = jQue y.parseJSON(pageCutMode01_hJson);
            infoJson = pageCutMode01_hJson_arr[pageNo];  infoJson = {0: {…}, 1: {…},
            picsBaseURL = book_json.oss_base_path + book_json.oss_book_file + book_j
            break;

        default:
    }

    if (infoJson == null) {
```

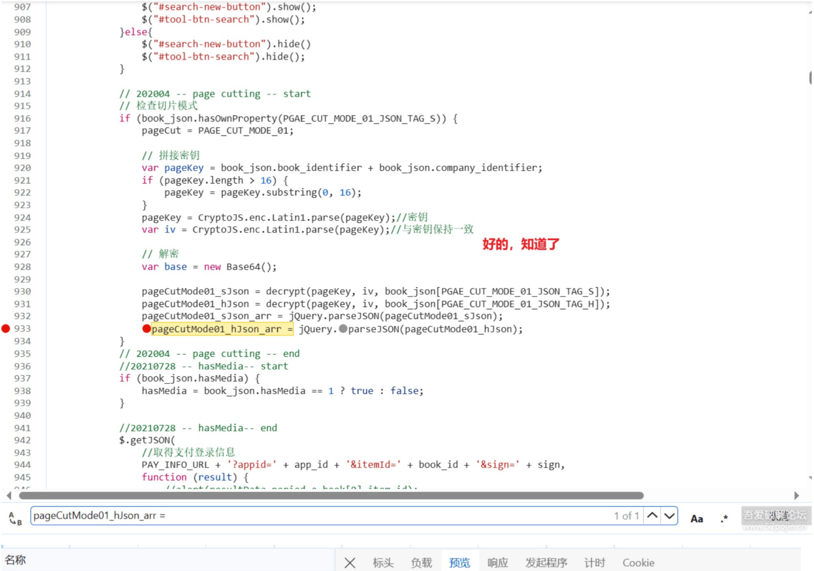
```
Object
▼ 1:
  ▶ 0: {h: 532, w: 378, x: 0, y: 0}
  ▶ 1: {h: 709, w: 882, x: 378, y: 0}
  ▶ 2: {h: 1241, w: 378, x: 0, y: 531}
  ▶ 3: {h: 354, w: 882, x: 378, y: 708}
  ▶ 4: {h: 354, w: 882, x: 378, y: 1063}
  ▶ 5: {h: 355, w: 882, x: 378, y: 1417}
    h: 1772
    w: 1260
  ▶ [[Prototype]]: Object
▶ 2: {0: {…}, 1: {…}, 2: {…}, 3: {…}, 4:
▶ 3: {0: {…}, 1: {…}, 2: {…}, 3: {…}, 4:
▶ 4: {0: {…}, 1: {…}, 2: {…}, 3: {…}, 4:
```

吾爱破解论坛
www.52pojie.cn



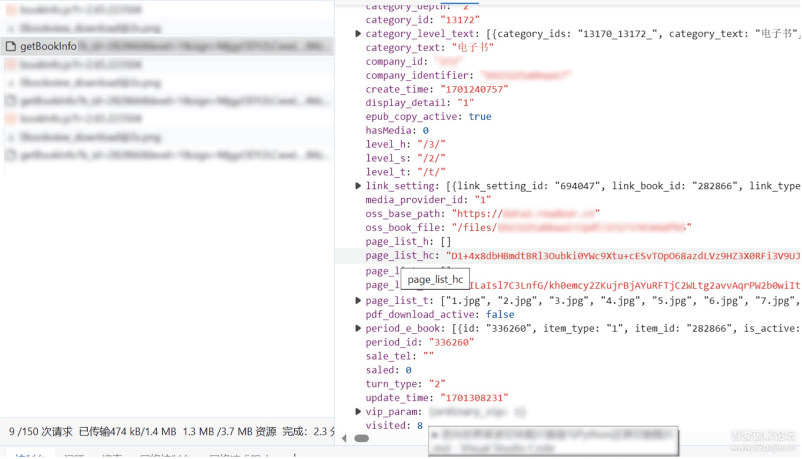
```
== "w") continue;
[iStr] = new Array();
[iStr]["url"] = picsBaseURL + crcForPageCutMode01(book_json.book_identifier + "_" + (level + 1) + "_" + pageNo + "_" + iStr) + ".jpg";
[iStr]["coordinate"] = infoJson[iStr];


URLArray["pics"][0]["coordinate"]["w"]);
```

最后就是一个pageCutMode01_hJson_arr了，是一个AES解密，bookjson获取到的数据就是加密后的。



```
                $("#search-new-button").show();
                $("#tool-btn-search").show();
            }else{
                $("#search-new-button").hide()
                $("#tool-btn-search").hide();
            }

            // 202004 -- page cutting -- start
            // 检查切片模式
            if (book_json.hasOwnProperty(PGAE_CUT_MODE_01_JSON_TAG_S)) {
                pageCut = PAGE_CUT_MODE_01;

                // 拼接密钥
                var pageKey = book_json.book_identifier + book_json.company_identifier;
                if (pageKey.length > 16) {
                    pageKey = pageKey.substring(0, 16);
                }
                pageKey = CryptoJS.enc.Latin1.parse(pageKey);//密钥
                var iv = CryptoJS.enc.Latin1.parse(pageKey);//与密钥保持一致

                // 解密
                var base = new Base64();

                pageCutMode01_sJson = decrypt(pageKey, iv, book_json[PGAE_CUT_MODE_01_JSON_TAG_S]);
                pageCutMode01_hJson = decrypt(pageKey, iv, book_json[PGAE_CUT_MODE_01_JSON_TAG_H]);
                pageCutMode01_sJson_arr = jQuery.parseJSON(pageCutMode01_sJson);
                pageCutMode01_hJson_arr = jQuery.parseJSON(pageCutMode01_hJson);
            }
            // 202004 -- page cutting -- end
            //20210728 -- hasMedia-- start
            if (book_json.hasMedia) {
                hasMedia = book_json.hasMedia == 1 ? true : false;
            }

            //20210728 -- hasMedia-- end
            $.getJSON(
                //取得支付登录信息
                PAY_INFO_URL + '?appid=' + app_id + '&itemId=' + book_id + '&sign=' + sign,
                function (result) {
```

好的，知道了



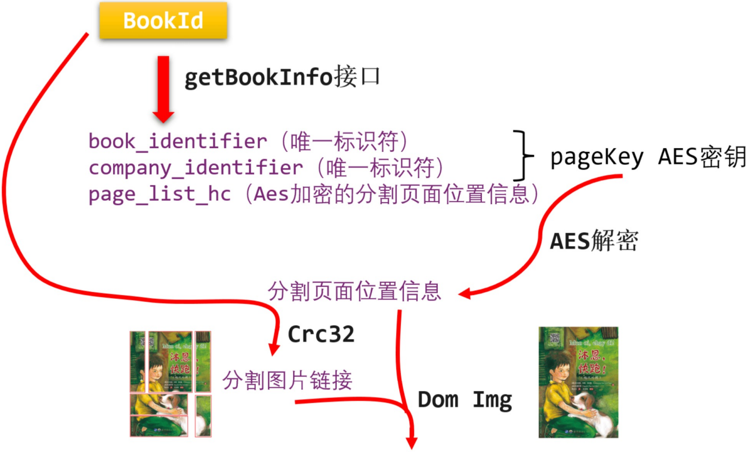
```
getBookInfo [redacted]
```

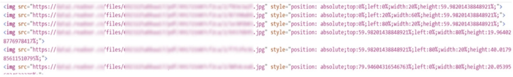
```
category_depth: "2"
category_id: "13172"
▶ category_level_text: [{category_ids: "13170_13172_", category_text: "电子书",
category_text: "电子书"
company_id: [redacted]
company_identifier: [redacted]
create_time: "1701240757"
display_detail: "1"
epub_copy_active: true
hasMedia: 0
level_h: "/3/"
level_s: "/2/"
level_t: "/t/"
▶ link_setting: [{link_setting_id: "694047", link_book_id: "282866", link_type
media_provider_id: "1"
oss_base_path: "https://[redacted]"
oss_book_file: "/files/[redacted]"
page_list_h: []
page_list_hc: "D1+4x8dbHBmdtBRl3Oubki0YWc9Xtu+cESvTOpO68azdLVz9HZ3X0RFi3V9UJ
page_l
page_l                      ILaIsl7C3LnfG/kh0emcy2ZKujrBjAYuRFTjC2WLtg2avvAqrPW2b0wiIt
▶ page_list_t: ["1.jpg", "2.jpg", "3.jpg", "4.jpg", "5.jpg", "6.jpg", "7.jpg",
pdf_download_active: false
▶ period_e_book: [{id: "336260", item_type: "1", item_id: "282866", is_active:
period_id: "336260"
sale_tel: ""
saled: 0
turn_type: "2"
update_time: "1701308231"
▶ vip_param: [redacted]
visited: 8
```

9 /150 次请求 已传输474 kB/1.4 MB 1.3 MB /3.7 MB 资源 完成: 2.3 分

吾爱破解论坛
www.52pojie.cn

流程


BookId
getBookInfo接口
book_identifier（唯一标识符）
company_identifier（唯一标识符）
page_list_hc（Aes加密的分割页面位置信息）
pageKey AES密钥
AES解密
分割页面位置信息
Crc32
分割图片链接
Dom Img



```
<img src="https://[redacted]/files/[redacted].jpg" style="position: absolute;top:0%;left:0%;width:20%;height:59.98201438848921%;">
<img src="https://[redacted]/files/[redacted].jpg" style="position: absolute;top:0%;left:20%;width:60%;height:59.98201438848921%;">
<img src="https://[redacted]/files/[redacted].jpg" style="position: absolute;top:0%;left:80%;width:20%;height:59.98201438848921%;">
<img src="https://[redacted]/files/[redacted].jpg" style="position: absolute;top:59.98201438848921%;left:0%;width:80%;height:19.964028776978417%;">
<img src="https://[redacted]/files/[redacted].jpg" style="position: absolute;top:59.98201438848921%;left:80%;width:20%;height:40.017985611510795%;">
<img src="https://[redacted]/files/[redacted].jpg" style="position: absolute;top:79.94604316546763%;left:0%;width:80%;height:20.05395
```

吾爱破解论坛
www.52pojie.cn

Python还原

AES解析BookInfo获取分割图片位置信息

复制代码 隐藏代码

```
var pageKey = book_json.book_identifier + book_json.company_identifier;
if (pageKey.length > 16) {
    pageKey = pageKey.substring(0, 16);
}
pageKey = CryptoJS.enc.Latin1.parse(pageKey);//密钥
var iv = CryptoJS.enc.Latin1.parse(pageKey);//与密钥保持一致
```

注意要把book_identifier和company_identifier拼接的Key削减为标准的16位，否则不是标准的AES key长度。

复制代码 隐藏代码

```
AESKey = response['book_identifier'] + response['company_identifier']
AESKey = AESKey[0:16] if len(AESKey) > 16 else AESKey
key_bytes = AESKey.encode('latin-1')
```

此处AES加密模式是CBC，ZeroPadding，Python一样写

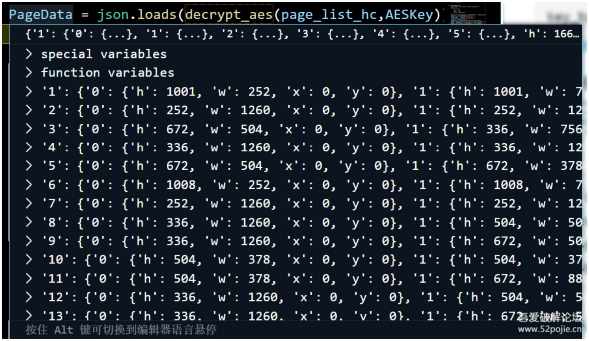
```
function decrypt(k, iv, d) {
    var decrypted = CryptoJS.AES.decrypt(d, k, { iv: k, mode: CryptoJS.mode.CBC, padding: CryptoJS.pad.ZeroPadding });
    decrypted = decrypted.toString(CryptoJS.enc.Utf8);
    return decrypted;
}
```

吾爱破解论坛
www.52pojie.cn

复制代码 隐藏代码

```
from Crypto.Cipher import AES
cipher = AES.new(key_bytes, AES.MODE_CBC, iv=key_bytes)
decrypted = cipher.decrypt(base64.b64decode(ciphertext))
decrypted = ''.join([chr(i) for i in decrypted if i != 0 ])
```

然后我们就成功获得了坐标Json decrypted，当然要json.loads()一下，使其成为json对象。



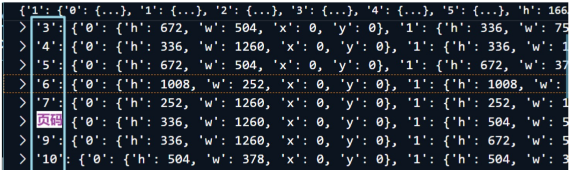
```
PageData = json.loads(decrypt_aes(page_list_hc,AESKey))
 {'1': {'0': {...}, '1': {...}, '2': {...}, '3': {...}, '4': {...}, '5': {...}, 'h': 166…
 > special variables
 > function variables
 > '1': {'0': {'h': 1001, 'w': 252, 'x': 0, 'y': 0}, '1': {'h': 1001, 'w': 7
 > '2': {'0': {'h': 252, 'w': 1260, 'x': 0, 'y': 0}, '1': {'h': 252, 'w': 12
 > '3': {'0': {'h': 672, 'w': 504, 'x': 0, 'y': 0}, '1': {'h': 336, 'w': 756
 > '4': {'0': {'h': 336, 'w': 1260, 'x': 0, 'y': 0}, '1': {'h': 336, 'w': 12
 > '5': {'0': {'h': 672, 'w': 504, 'x': 0, 'y': 0}, '1': {'h': 672, 'w': 378
 > '6': {'0': {'h': 1008, 'w': 252, 'x': 0, 'y': 0}, '1': {'h': 1008, 'w': 7
 > '7': {'0': {'h': 252, 'w': 1260, 'x': 0, 'y': 0}, '1': {'h': 252, 'w': 12
 > '8': {'0': {'h': 336, 'w': 1260, 'x': 0, 'y': 0}, '1': {'h': 504, 'w': 50
 > '9': {'0': {'h': 336, 'w': 1260, 'x': 0, 'y': 0}, '1': {'h': 672, 'w': 50
 > '10': {'0': {'h': 504, 'w': 378, 'x': 0, 'y': 0}, '1': {'h': 504, 'w': 37
 > '11': {'0': {'h': 504, 'w': 378, 'x': 0, 'y': 0}, '1': {'h': 672, 'w': 88
 > '12': {'0': {'h': 336, 'w': 1260, 'x': 0, 'y': 0}, '1': {'h': 504, 'w': 5
 > '13': {'0': {'h': 336, 'w': 1260, 'x': 0, 'y': 0}, '1': {'h': 67
 按住 Alt 键可切换到编辑器语言悬停
```

非标准Crc获取分割图片链接

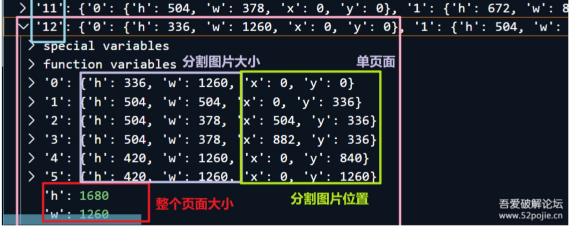
复制代码 隐藏代码

```
for (var iStr in infoJson) {
    if (iStr == "h" || iStr == "w") continue;
    pagePicsURLArray["pics"][iStr] = new Array();
    pagePicsURLArray["pics"][iStr]["url"] = picsBaseURL + crcForPageCutMode01(book_json.book_identifier + "_" + (level + 1) + "_" + pageNo + "_" + iStr) + ".jpg";
    pagePicsURLArray["pics"][iStr]["coordinate"] = infoJson[iStr];
}
```

level+1可以写死，最高等级,iStr是infoJson的每个Key，其实也就是分割的0、1、2、3、4、5几个数字。

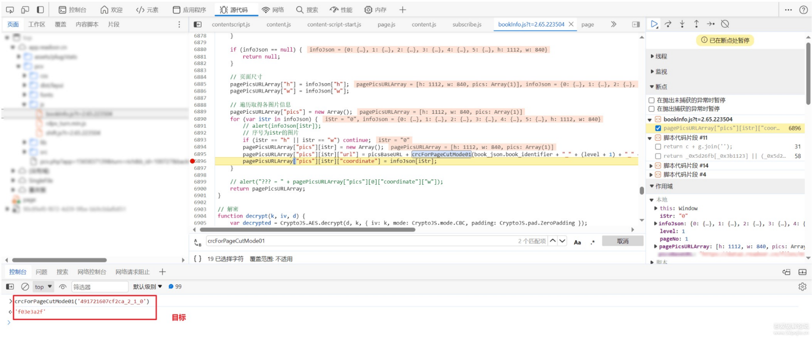


```
 {'1': {'0': {...}, '1': {...}, '2': {...}, '3': {...}, '4': {...}, '5': {...}, 'h': 166
 > '3': {'0': {'h': 672, 'w': 504, 'x': 0, 'y': 0}, '1': {'h': 336, 'w': 75
 > '4': {'0': {'h': 336, 'w': 1260, 'x': 0, 'y': 0}, '1': {'h': 336, 'w': 1
 > '5': {'0': {'h': 672, 'w': 504, 'x': 0, 'y': 0}, '1': {'h': 672, 'w': 37
 > '6': {'0': {'h': 1008, 'w': 252, 'x': 0, 'y': 0}, '1': {'h': 1008, 'w':
 > '7': {'0': {'h': 252, 'w': 1260, 'x': 0, 'y': 0}, '1': {'h': 252, 'w': 1
 > 页码 {'0': {'h': 336, 'w': 1260, 'x': 0, 'y': 0}, '1': {'h': 504, 'w': 5
 > '9': {'0': {'h': 336, 'w': 1260, 'x': 0, 'y': 0}, '1': {'h': 672, 'w': 5
 > '10': {'0': {'h': 504, 'w': 378, 'x': 0, 'y': 0}, '1': {'h': 504, 'w': 3
```



```
> '11': {'0': {'h': 504, 'w': 378, 'x': 0, 'y': 0}, '1': {'h': 672, 'w': 8
v '12': {'0': {'h': 336, 'w': 1260, 'x': 0, 'y': 0}, '1': {'h': 504, 'w':
  > special variables
  > function variables
  > '0': {'h': 336, 'w': 1260, 'x': 0, 'y': 0}
  > '1': {'h': 504, 'w': 504, 'x': 0, 'y': 336}
  > '2': {'h': 504, 'w': 378, 'x': 504, 'y': 336}
  > '3': {'h': 504, 'w': 378, 'x': 882, 'y': 336}
  > '4': {'h': 420, 'w': 1260, 'x': 0, 'y': 840}
  > '5': {'h': 420, 'w': 1260, 'x': 0, 'y': 1260}
    'h': 1680
    'w': 1260
```

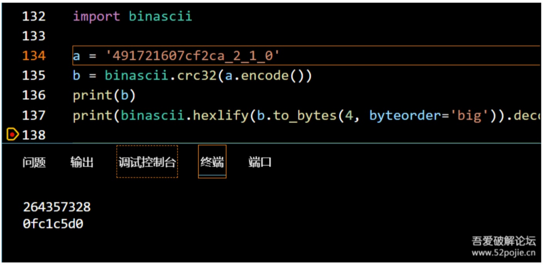
首先断下一个Js的crcForPageCutMode01函数，看看在浏览器的表现如何。



```
        }

        if (infoJson == null) {  infoJson = {0: {…}, 1: {…}, 2: {…}, 3: {…}, 4: {…}, 5: {…}, h: 1112, w: 840}
            return null;
        }

        // 页面尺寸
        pagePicsURLArray["h"] = infoJson["h"];  pagePicsURLArray = [h: 1112, w: 840, pics: Array(1)], infoJson = {0: {…}, 1: {…}, 2: {…},
        pagePicsURLArray["w"] = infoJson["w"];

        // 遍历取得各图片信息
        pagePicsURLArray["pics"] = new Array();  pagePicsURLArray = [h: 1112, w: 840, pics: Array(1)]
        for (var iStr in infoJson) {  iStr = "0", infoJson = {0: {…}, 1: {…}, 2: {…}, 3: {…}, 4: {…}, 5: {…}, h: 1112, w: 840}
            // alert(infoJson[iStr]);
            // 序号为iStr的图片
            if (iStr == "h" || iStr == "w") continue;  iStr = "0"
            pagePicsURLArray["pics"][iStr] = new Array();  pagePicsURLArray = [h: 1112, w: 840, pics: Array(1)]
            pagePicsURLArray["pics"][iStr]["url"] = picsBaseURL + crcForPageCutMode01(book_json.book_identifier + "_" + (level + 1) + "_"
            pagePicsURLArray["pics"][iStr]["coordinate"] = infoJson[iStr];
        }

        // alert("??? = " + pagePicsURLArray["pics"][0]["coordinate"]["w"]);
        return pagePicsURLArray;
    }

    // 解密
    function decrypt(k, iv, d) {
        var decrypted = CryptoJS.AES.decrypt(d, k, { iv: k, mode: CryptoJS.mode.CBC, padding: CryptoJS.pad.ZeroPadding });
```

```
> crcForPageCutMode01('491721607cf2ca_2_1_0')
< 'f03e3a2f'
```

Crc32是常见的hash算法，同样的输入不会出现两个结果，这个f03e3a2f就是我们的标准值。首先用Python内置的binascii库调用一下标准的Crc32。



```python
import binascii

a = '491721607cf2ca_2_1_0'
b = binascii.crc32(a.encode())
print(b)
print(binascii.hexlify(b.to_bytes(4, byteorder='big')).dec
```

```
264357328
0fc1c5d0
```

与浏览器不同，这说明可能使用了非标准的处理，我们进去crcForPageCutMode01与标准的Crc32对比看看。

复制代码 隐藏代码

```
function crcForPageCutMode01(Instr) {
    if (typeof (window.Crc32Table) == "undefined") {
        window.Crc32Table = new Array(256);
        var i, j;
        var Crc;
        for (i = 0; i < 256; i++) {
            Crc = i;
            for (j = 0; j < 8; j++) {
                if (Crc & 1)
                    Crc = ((Crc >> 1) & 0x7FFFFFFF) ^ 0xEDB88320;
                else
                    Crc = ((Crc >> 1) & 0x7FFFFFFF);
            }
            Crc32Table[i] = Crc;
        }
    }
    if (typeof Instr != "string") Instr = "" + Instr;
    Crc = 0xFFFFFFFF;
    for (i = 0; i < Instr.length; i++)
        Crc = ((Crc >> 8) & 0x00FFFFFF) ^ Crc32Table[(Crc & 0xFF) ^ Instr.charCodeAt(i)];
```

```
    Crc ^= 0xFFFFFFFF;
    Crc = (Crc ^ (-1)) >>> 0
    return (Crc).toString(16);
}
```

与标准的Crc32对比，其实就是多了一句Crc = (Crc ^ (-1)) >>> 0。

/>>>是Javascript的无符号右移运算符，在这儿>>> 0 将前面的数强制转为无符号32位整数，但是Python中没有直接的无符号右移。

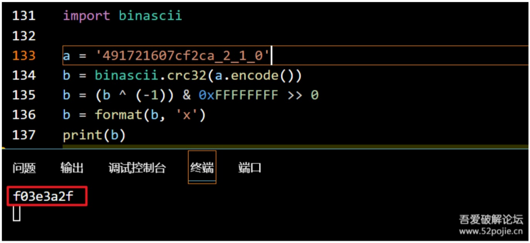
我们可以使用& 0xffffffff 确保值在无符号32位整数范围内。

所以我们在标准的Crc32后，先进行这一步，再转成16进制字符串即可。

复制代码 隐藏代码

```
import binascii

a = '491721607cf2ca_2_1_0'
b = binascii.crc32(a.encode())
b = (b ^ (-1)) & 0xFFFFFFFF >> 0
b = format(b, 'x')
print(b)
```

获得目标值f03e3a2f，把它封装到Python的crcForPageCutMode01函数就可了。



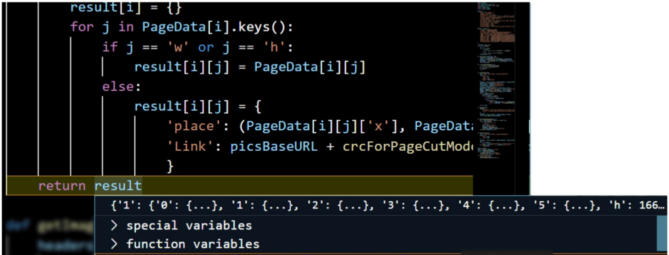
```
import binascii

a = '491721607cf2ca_2_1_0'
b = binascii.crc32(a.encode())
b = (b ^ (-1)) & 0xFFFFFFFF >> 0
b = format(b, 'x')
print(b)
```

```
f03e3a2f
```

同时PageData中每一个整图片总大小w、h，分块的位置x，y后续拼接图片是需要的，那么我们都放入result中，最终获得了如图result。

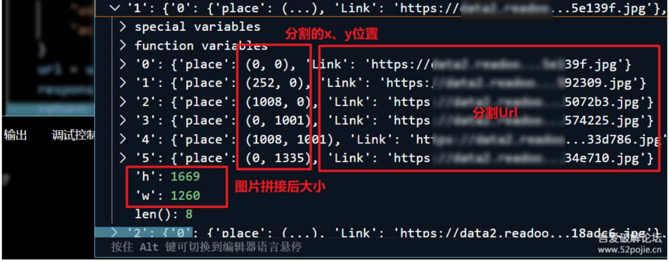
复制代码 隐藏代码

```
for i in PageData.keys():
    result[i] = {}
    for j in PageData[i].keys():
        if j == 'w' or j == 'h':
            result[i][j] = PageData[i][j]
        else:
            result[i][j] = {
                'place': (PageData[i][j]['x'], PageData[i][j]['y']),
                'Link': picsBaseURL + crcForPageCutMode01(response['book_identifier'] + "_3_" + i + "_" + j) + ".jpg"
            }
```



```
        result[i] = {}
        for j in PageData[i].keys():
            if j == 'w' or j == 'h':
                result[i][j] = PageData[i][j]
            else:
                result[i][j] = {
                    'place': (PageData[i][j]['x'], PageDat
                    'Link': picsBaseURL + crcForPageCutMod
                }
    return result
```

```
{'1': {'0': {...}, '1': {...}, '2': {...}, '3': {...}, '4': {...}, '5': {...}, 'h': 166…
> special variables
> function variables
```



```
∨ '1': {'0': {'place': (...), 'Link': 'https://[redacted]...5e139f.jpg'},
  > special variables
  > function variables
  > '0': {'place': (0, 0), 'Link': 'https://[redacted]39f.jpg'}
  > '1': {'place': (252, 0), 'Link': 'https:[redacted]92309.jpg'}
  > '2': {'place': (1008, 0), 'Link': 'https[redacted]5072b3.jpg'}
  > '3': {'place': (0, 1001), 'Link': 'https[redacted]574225.jpg'}
  > '4': {'place': (1008, 1001), 'Link': 'ht[redacted]...33d786.jpg'
  > '5': {'place': (0, 1335), 'Link': 'https[redacted]34e710.jpg'}
    'h': 1669
    'w': 1260
    len(): 8
> '2': {'0': {'place': (...), 'Link': 'https://data2.readoo...18ade6.jpg'},
按住 Alt 键可切换到编辑器语言悬停
```

吾爱破解论坛
www.52pojie.cn

Python分割图片的还原

终于我们获得了每个分割图片的链接、和图片在整个图片的位置，接下来我们用PIL实现分离图片的合并。

requests请求图片链接获取图片的二进制流，接下来要创建Image对象。因为Image.open()接受的是文件名，所以我们可以先把response.content存入字节流管道，再创建Image对象。

复制代码 隐藏代码

```
import requests
from PIL import Image
from io import BytesIO

response = requests.get(url, headers=headers).content
imageBlock = Image.open(BytesIO(response))
```

然后创建一个新的完整图片大小的空Image，也就是result中的h、w。

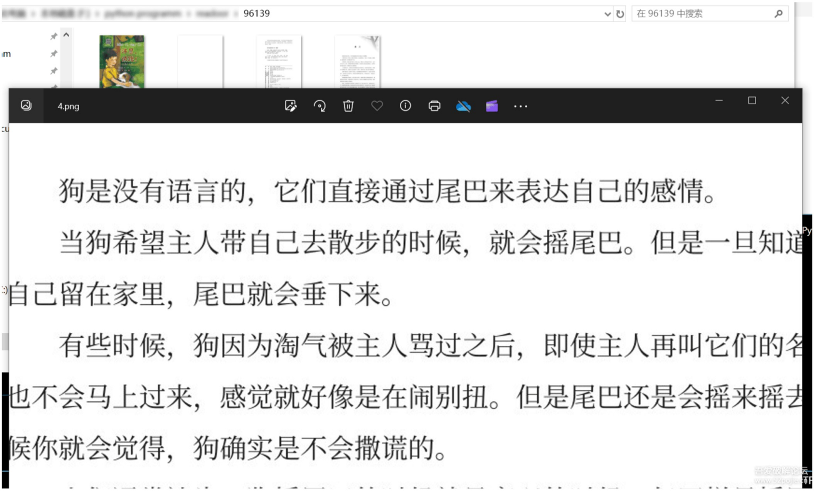
复制代码 隐藏代码

```
pic = Image.new("RGB", (data['w'], data['h']))
```

对每一个分块，用.paste方法加入pic指定位置中（也就是result中的place），最后保存到png中，最终获得完整、清晰的页面图片。

复制代码 隐藏代码

```
for j in data.keys():
    if j == 'w' or j == 'h':
        continue
    Link = data[j]['Link']
    imageData = getImage(Link)
    pic.paste(imageData, data[j]['place'])
pic.save(f'{bookid}/{i}.png')
```

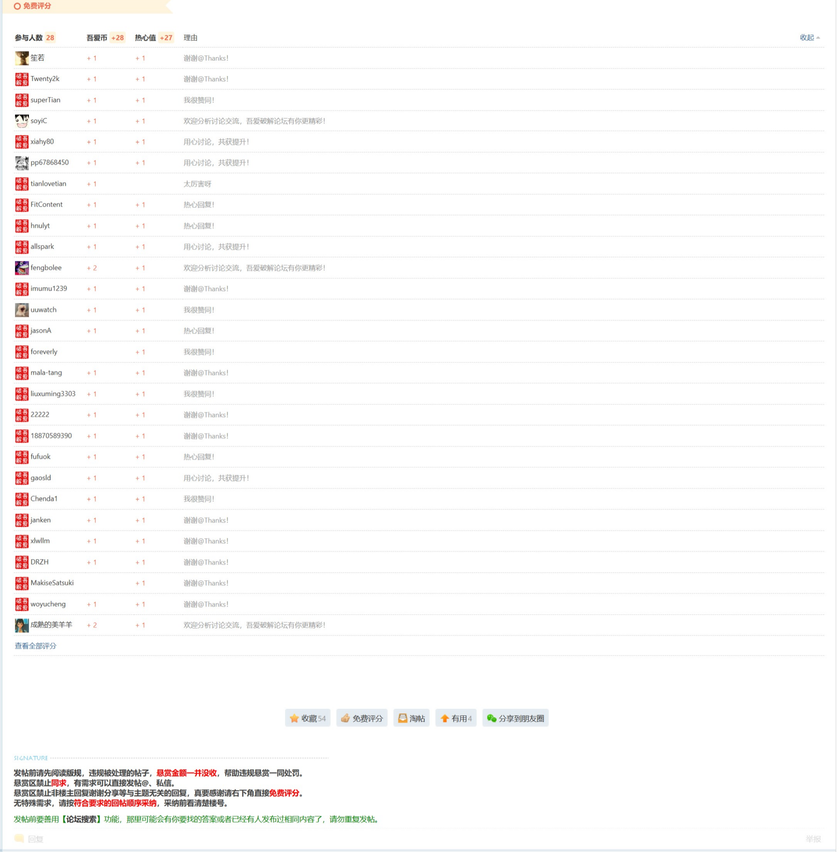


狗是没有语言的，它们直接通过尾巴来表达自己的感情。

当狗希望主人带自己去散步的时候，就会摇尾巴。但是一旦知道
自己留在家里，尾巴就会垂下来。

有些时候，狗因为淘气被主人骂过之后，即使主人再叫它们的名
也不会马上过来，感觉就好像是在闹别扭。但是尾巴还是会摇来摇
候你就会觉得，狗确实是不会撒谎的。

吾爱破解论坛
www.52pojie.cn



参与人数 28	吾爱币 +28	热心值 +27	理由
笙若	+ 1	+ 1	谢谢@Thanks!
Twenty2k	+ 1	+ 1	谢谢@Thanks!
superTian	+ 1	+ 1	我很赞同!
soyiC	+ 1	+ 1	欢迎分析讨论交流，吾爱破解论坛有你更精彩!
xiahy80	+ 1	+ 1	用心讨论，共获提升!
pp67868450	+ 1	+ 1	用心讨论，共获提升!
tianlovetian	+ 1		太厉害呀
FitContent	+ 1	+ 1	热心回复!
hnulyt	+ 1	+ 1	热心回复!
allspark	+ 1	+ 1	用心讨论，共获提升!
fengbolee	+ 2	+ 1	欢迎分析讨论交流，吾爱破解论坛有你更精彩!
imumu1239	+ 1	+ 1	谢谢@Thanks!
uuwatch	+ 1	+ 1	我很赞同!
jasonA	+ 1	+ 1	热心回复!
foreverly		+ 1	我很赞同!
mala-tang	+ 1	+ 1	谢谢@Thanks!
liuxuming3303	+ 1	+ 1	我很赞同!
22222	+ 1	+ 1	谢谢@Thanks!
18870589390	+ 1	+ 1	谢谢@Thanks!
fufuok	+ 1	+ 1	热心回复!
gaosld	+ 1	+ 1	用心讨论，共获提升!
Chenda1	+ 1	+ 1	我很赞同!
janken	+ 1	+ 1	谢谢@Thanks!
xlwllm	+ 1	+ 1	谢谢@Thanks!
DRZH	+ 1	+ 1	谢谢@Thanks!
MakiseSatsuki		+ 1	谢谢@Thanks!
woyucheng	+ 1	+ 1	谢谢@Thanks!
成熟的美羊羊	+ 2	+ 1	欢迎分析讨论交流，吾爱破解论坛有你更精彩!

发帖前请先阅读版规，违规被处理的帖子，悬赏金额一并没收，帮助违规悬赏一同处罚。
悬赏区禁止同求，有需求可以直接发帖@、私信。
悬赏区禁止非楼主回复谢谢分享等与主题无关的回复，真要感谢请右下角直接免费评分。
无特殊需求，请按符合要求的回帖顺序采纳，采纳前看清楚楼号。

发帖前要善用**【论坛搜索】**功能，那里可能会有你要找的答案或者已经有人发布过相同内容了，请勿重复发帖。