

图片转文档（华为）概述：

3 种转换方式：

单一图片转为文档，图片是一个下载链接，用 HTTP GET 方式，见：[单图转换 GET](#)

单一图片转为文档，图片 POST 到服务器，用 HTTP POST 方式，见：[单图转换 POST](#)

多个图片转为文档，图片是多个下载链接，用 HTTP POST 方式，见：[多图转换 POST](#)

基本用法：

由于转换需要时间，文件越大页数越多，转换越久，故系统采用**异步**的方式获得转换结果。

调用转换接口后会获得 token，随后有 2 种方式查询转换结果：

1. 用 HTTP GET 方式，轮询“查询 query 接口”获得结果。详细见：[查询 QUERY](#)
2. 设置 callbackurl，当转换结束后，系统会回调该 URL 直接推送转换结果。详细见：[回调 URL](#)

API 调用需要签名，详细见[华为官方签名](#)：

1. <https://support.huaweicloud.com/usermanual-apig/apig-ug-0011.html>

单图转换 GET

将一张图片下载链接 url 转换为其他格式，type 是目标文档的类型，比如要转为 docx，type 就是 docx

请求参数：

参数	类型及范围	备注	是否必须发送
url	string	文件 url，必须 http(s)，ftp 开头，需要 URL Encoding	是
type	string	小写，需转换为的文件类型，例如 docx	是
language	int	OCR 识别语言选项，默认 2 简体中文： 1: 英语 2: 简体中文 3: 繁体中文 4: 法语 5: 德语 6: 意大利语 7: 俄语 8: 日文 9: 韩文 10: 西班牙语 11: 葡萄牙语 12: 丹麦语 13: 荷兰语 14: 芬兰语 15: 挪威语	否

参数	类型及范围	备注	是否必须发送
		16: 瑞典语 17: 土耳其语	
clean	int	是否清除图像背景只显示文字，默认 0 否，1 是	否
outfilename	string	生成的文件的文件名，默认随机	否
callbackurl	string	回调 URL，转换结束后，会回调该 URL，需要 URL Encoding，详细见 回调 URL	否

请求示例:

```
http://img2doc.apistore.huaweicloud.com/v1/convert?url=https%3a%2f%2fxxx%2fxxx.jpg&type=docx&ocr=0
```

将所在 url 地址的 jpg 文件转为 docx，type 就是需转换为的文件类型，这个例子里就是 docx

输出文件类型 type 可取值: doc, docx, pptx, xlsx, rtf, txt, ofd。

必须签名才能调用成功，签名见华为签名规则:

<https://support.huaweicloud.com/usermanual-apig/apig-ug-0011.html>

返回数据结构:

名称	含义	类型及范围	是否必须返回	备注
code		number	是	10000: 请求成功
msg		string	是	
result		Dictionary	否	成功后返回

result:

名称	含义	类型及范围	是否必须返回	备注
token		string	是	用于 query 接口

返回示例(成功状态):

```
{
  "code":10000,
  "msg": "",
  "result":{"token":"xxx"}
}
```

返回示例(失败状态):

```
{
  "code":40001,
  "msg":"ParmNotRight"
}
```

单图转换 POST

直接将单个图片 POST 到服务器，大小限制 12M

请求参数：

参数	类型及范围	备注	是否必须发送
file	file	要转换的文档，Content-Type 使用 multipart/form-data，最大 12M	是
type	string	小写，需转换为的文件类型，例如 docx	是
language	int	OCR 识别语言选项，默认 2 简体中文，值同 GET 方法	否
clean	int	是否清除图像背景只显示文字，默认 0 否，1 是	否
outfilename	string	生成的文件的文件名，默认随机	否
callbackurl	string	回调 URL，转换结束后，会回调该 URL，详见 回调 URL	否

请求示例：

http://img2doc.apistore.huaweicloud.com/v1/convert_post

Header 中的 Content-Type 必须是 multipart/form-data

输出文件类型 **type** 可取值：**doc, docx, pptx, xlsx, rtf, txt, ofd**。

必须签名才能调用成功，签名见华为签名规则：

<https://support.huaweicloud.com/usermanual-apig/apig-ug-0011.html>

返回数据结构：

名称	含义	类型及范围	是否必须返回	备注
code		number	是	10000: 请求成功
msg		string	是	
result		Dictionary	否	成功后返回

result:

名称	含义	类型及范围	是否必须返回	备注
token		string	是	用于 query 接口

返回示例(成功状态):

```
{
  "code":10000,
  "msg": "",
  "result":{"token":"xxx"}
}
```

返回示例(失败状态):

```
{
  "code":40001,
  "msg":"ParmNotRight"
}
```

多图转换 POST

将多张图片转换为一个文档，url 传入多张图片的地址，传入非图片格式无效

Header 中的 Content-Type 传入 application/json

POST Body 传入 JSON，格式如下：

请求参数 BODY:

参数	类型及范围	备注	是否必须发送
url	string 数组	要转换的图片 URL 数组，支持 http(s)，ftp 开头	是
type	string	小写，需转换为的文件类型，例如 docx	是
language	int	OCR 识别语言选项，默认 2 简体中文，值同 GET 方法	否

参数	类型及范围	备注	是否必须发送
clean	int	是否清除图像背景只显示文字，默认 0 否，1 是	否
outfilename	string	生成的文件的文件名，默认随机	否
callbackurl	string	回调 URL，转换结束后，会回调该 URL，详细见回调 URL	否

请求示例:

`http://img2doc.apistore.huaweicloud.com/v1/convert`

传入的 Body 为 JSON 格式，如下：

```
{ "url": [ "http://xxx/xxx1.png", "http://xxx/xxx2.png" ], "type": "docx", "clean": 0 }
```

输出文件类型 **type** 可取值：**doc**, **docx**, **pptx**, **xlsx**, **rtf**, **txt**, **ofd**。

必须签名才能调用成功，签名见华为签名规则：

<https://support.huaweicloud.com/usermanual-apig/apig-ug-0011.html>

支持几乎所有图片格式

返回数据结构:

名称	含义	类型及范围	是否必须返回	备注
code		number	是	10000:请求成功
msg		string	是	
result		Dictionary	否	成功后返回

result:

名称	含义	类型及范围	是否必须返回	备注
token		string	是	用于 query 接口

返回示例(成功状态):

```
{
  "code":10000,
  "msg":"",
  "result":{"token":"xxx"}
}
```

返回示例(失败状态):

```
{
  "code":40001,
  "msg":"ParmNotRight"
}
```

查询 QUERY

请求参数:

参数	类型及范围	备注	是否必须发送
token	string		是

请求示例:

```
https://all2pdf.apistore.huaweicloud.com/v1/query?token=7b799e09e0838919d3ae63d0566683a2
```

无需签名, 无调用次数限制

由于转换需要时间, 文件越大页数越多, 转换越久, 故需要**轮询**查询接口来获得结果。查询频率可以是 1s 一次, 也可以更长一些。查询后先看 status, 如果是 Done 或 Failed, 则转换结束, 停止轮询。如果是 Doing 或 Pending, 则继续轮询。

返回数据结构:

名称	含义	类型及范围	是否必须返回	备注
code		number	是	10000: 请求成功
msg		string	是	
token		string	是	请求的 token
result		Dictionary	否	成功后返回

result:

名称	含义	类型及范围	是否必须返回	备注
status	状态	string	是	Pending: 还未开始 Doing: 正在转换 Done: 转换成功 Failed: 转换失败
progress	进度	number(0.00	否 (status 为	比如 0.88 表示 88%

名称	含义	类型及范围	是否必须返回	备注
		- 1.00)	Doing 时返回)	
fileurl	文件地址	string	否 (status 为 Done 时返回)	转换出来的文件地址, http 和 https 都支持
reason	失败原因	string	否 (status 为 Failed 时可能返回)	转换失败的原因

返回示例(成功状态):

```
{
  "code":10000,
  "msg": "",
  "token": "xxx",
  "result":
  {
    "progress":0.02,
    "status": "Doing"
  }
}
```

```
{
  "code":10000,
  "msg": "",
  "token": "xxx",
  "result":
  {
    "status": "Done",

    "fileurl": "https://file.duhitech.com/o/7b799e09e0838919d3ae63d0566683a2/cc9c7f1e-03f8-4742-bc37-aab9da191c26.docx"
  }
}
```

返回示例(失败状态):

```
{
  "code":40000,
  "msg": "No such token"
}
```

注意:

- 上传文件大小 **不能超过 1000M**。
- 转换完成后, 在 **1 小时** 内下载文件。

回调 URL:

用途: 客户可以自行部署服务器, 系统转换结束后会调用客户提供的回调 URL, 直接发送转换结果, 从而无需再轮询 Query。

当设置了回调 URL, 转换结束后 (无论成功失败), 系统都会尝试调用该 URL, 具体如下:

以 POST 方式调用该 URL, Header 头中 Content-Type: application/json

Body 为 JSON 格式, 内容和 Query 的结果相同, 例如:

```
{"code":10000,"msg":"","token":"xxx","result":{"status":"Done","fileurl":"https://file.duhuitech.com/o/7b799e09e0838919d3ae63d0566683a2/cc9c7f1e-03f8-4742-bc37-aab9da191c26.docx"}}
```

服务端收到该 POST 后需在 10 秒内返回 HTTP STATUS CODE 200, 视为调用成功, 否则系统认为回调失败, 会再次尝试。规则如下:

系统共计最多会调用 3 次回调 URL, 如果第一次失败, 则等待 3 秒后尝试第二次, 如果第二次失败, 则等待 5 秒后尝试第三次, 如果第三次失败, 则不再尝试。

回调 URL 超时时间 10 秒。

错误码表:

JSON 里返回的 code	错误信息
40000	通用错误
40001	参数错误
40002	参数不符合规范