

## 文档转图片（阿里）概述：

在线试用：

<https://try.dhconvert.com/>

小程序搜索：

度慧文档转换

### 3 种转换方式：

单一文档转为图片，文档是一个下载链接，用 HTTP GET 方式，见：[文档转换 GET](#)

单一文档转为图片，文档 POST 到服务器，用 HTTP POST 方式，见：[文档转换 POST](#)

多个图片转为图片，文档是多个下载链接，用 HTTP POST 方式，见：[多张图片转换 POST](#)

### 查询转换结果：

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。有 2 种方式：

1. 调用转换接口后获得 token，再用 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](#)

阿里云支持从 OSS 内网直接下载文件，节约流量，见：[阿里云独有部分](#)

## 文档转换 GET

将文档下载地址 url 转换为图片，type 是源文档的 type，比如要把 docx 转为长图，type 就是 docx，outtype 就是 2

### 请求参数：

参数	类型及范围	备注	是否必须发送
url	string	文件 url，支持 http(s)，ftp 开头，需要 URL Encoding	是
outtype	int	输出类型：1：每页一个图；2：长图；3：动图	是
type	string	要转换的原始文件扩展名，如果不传，则取 url 中的扩展名， <b>注意如果获取失败会导致转换异常或出错</b>	否
imagesize	int	图片大小。	否

参数	类型及范围	备注	是否必须发送
		<p>如果 outtype 是 1（每页一个图），该值是每个图最大的宽或高。</p> <p>如果 outtype 是 2（长图），该值是长图的宽。</p> <p>如果 outtype 是 3（动图），该值是动图最大的宽或高。</p> <p>默认是 0，自动。长图，动图宽最大 2000，其他图宽高最大 4096。</p>	
imageduration	int	<p>动图多少秒切换下一张。如果 outtype 是 3（动图）才有效，单位是秒，比如传入 10 就是每隔 10s 切换一张图</p>	否
imageformat	string	<p>如果输出是每页一个图或长图，输出的图片格式。默认空：jpg，可填入 png</p>	否
ismobileurl	int	<p>如果 type 是“url”才有效，抓取的网页是否移动端显示。默认 0：桌面端网页，1：移动端网页</p>	否
urlmargin	string	<p>如果 type 是 url，生成的页面边距。输入的是 URLEncoding 后的字符串，按照顺序：left top right bottom。单位可以是：px, in, cm, mm。例如左上右下分别是 1px, 2px, 3px, 4px，传的字符串就是：1px%202px%203px%204px。默认值：左右 0，上下各 1cm</p>	否
excelislandscape	int	<p>如果是 Excel 文件，是否横屏，默认 0 否（竖屏），1 是</p>	否
exceliscenter	int	<p>如果是 Excel 文件，是否横竖居中，默认 0 否，1 是</p>	否

参数	类型及范围	备注	是否必须发送
excelmargin	int	如果是 Excel 文件，四边的边距，默认 10px，单位是像素	否
excelsheetindex	int	如果是 Excel 文件，指定转换的 Sheet 索引，默认 0：全部，第一个 sheet 就是 1，以此类推	否
excelnotshowgridlines	int	如果是 Excel 文件，不显示网格线，默认 0 显示，1 不显示	否
exceluseprintarea	int	如果是 Excel 文件，是否使用打印区域，默认 0 不使用，1 使用，2 不使用且只显示有内容的区域	否
excelpagesize	int	如果是 Excel 文件，设定页面大小，默认 A4。0:A4，1:A3，2:A4，3:A5，4:B4，5:B5，6:Letter，7:Legal，8:Tabloid，9:Ledger	否
wordshowmarkup	int	如果是 Word 文件，是否显示审阅标记，默认 0 不显示，1 显示	否
powerpointoutputtype	int	如果是 PPT 文件，设定导出样式，默认幻灯片。0:幻灯片，0 以上是讲义模式：1:每页一个幻灯片，2:每页二个幻灯片，3:每页三个幻灯片，4:每页四个幻灯片，5:每页六个幻灯片，6:每页九个幻灯片	否
powerpointhandoutorder	int	如果是 PPT 文件，当设置为讲义时，设定顺序，默认 0 水平，1 垂直	否
powerpointhandoutorientation	int	如果是 PPT 文件，当设置为讲义时，输出文件的方向，默认 0 不改变，1 横向，2 纵向	否

参数	类型及范围	备注	是否必须发送
watermark	string	添加水印，字符个数最大 10 个，仅对非图片文件有效，需要 URL Encoding	否
watermarkfontsize	int	水印的字体大小，默认 24pt	否
watermarkfontcolor	string	水印的颜色，输入颜色码，默认黑色#000000，必须 7 位，需要 URL Encoding	否
watermarkfontalpha	int	水印的透明度，取值范围 1-100，越小越透明，默认 20	否
watermarkstyle	int	水印的样式，默认 0：文档中央一个水印；1：文档铺满水印	否
password	string	源文件的密码，支持有密码的 PDF，Word，PPT，Excel 文件类型，默认空无密码	否
pageindexes	string	如果源文件是 PDF 文件，指定转换的 PDF 页数，默认空全部页，例如：1, 3, 5-7 就是 1, 3, 5, 6, 7 共 5 页	否
callbackurl	string	回调 URL，转换结束后，会回调该 URL，需要 URL Encoding，详细见 <a href="#">回调 URL</a>	否

### 请求示例：

```
https://all2img.market.alicloudapi.com/v1/convert?url=https%3a%2f%2fxxx%2fxxx.docx&outtype=2
```

将所在 url 地址的 docx 文件转为长图，type 就是源文件的 type，这个例子里就是 docx，outtype 就是 2（长图）

签名见阿里签名规则。

支持多种文件格式，具体如下（**type** 可传入如下格式）：

PDF 文档：pdf

微软 Office 文档: doc, docx, ppt, pptx, xls, xlsx, pot, pps, ppsx

WPS 文档: wps, wpt, dps, dpt, et, ett

苹果 iWork 文档: pages, key, numbers

开放版式文档: ofd

电子刊物: caj

电子书: epub, chm, mobi, azw, azw3, fb2, cbr, cbz, djvu

Markdown 格式: md

SVG 格式: svg

CAD 文档: dwg, dxf, dwt, dws

Sketch 文档: sketch

网页文件: html, htm, mht, eml

图片文件: 几乎所有格式例如 png, jpg, jpeg, gif, tif, tiff, bmp, **可以统一传 img, 代表一切图片**

文本文件: txt, rtf, java, js, c, cpp, jsp, css, xml, properties, log, 其他任意文本文件都可以传 txt

网址网页: url (例如: http://www.baidu.com, 不支持文件下载链接)

### 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"
}
```

## 文档转换 POST

直接将单个文档 POST 到服务器，大小限制 8M

### 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 8M	是
outtype	int	输出类型：1：每页一个图；2：长图；3：动图	是
type	string	要转换的原始文件扩展名，如果不传，则取 file 中的文件扩展名， <b>注意如果获取失败会导致转换异常或出错</b>	否
imagesize	int	图片大小。 如果 outtype 是 1（每页一个图），该值是每个图最大的宽或高。 如果 outtype 是 2（长图），该值是长图的宽。 如果 outtype 是 3（动图），该值是动图最大的宽或高。 默认是 0，自动。长图，动图宽	否

参数	类型及范围	备注	是否必须发送
		最大 2000，其他图宽高最大 4096。	
imageduration	int	动图多少秒切换下一张。如果 outtype 是 3（动图）才有效，单位是秒，比如传入 10 就是每隔 10s 切换一张图	否
imageformat	string	如果输出是每页一个图或长图，输出的图片格式。默认空：jpg，可填入 png	否
excelislandscape	int	如果是 Excel 文件，是否横屏，默认 0 否（竖屏），1 是	否
exceliscenter	int	如果是 Excel 文件，是否横竖居中，默认 0 否，1 是	否
excelmargin	int	如果是 Excel 文件，四边的边距，默认 10px，单位是像素	否
excelsheetindex	int	如果是 Excel 文件，指定转换的 Sheet 索引，默认 0：全部，第一个 sheet 就是 1，以此类推	否
excelnotshowgridlines	int	如果是 Excel 文件，不显示网格线，默认 0 显示，1 不显示	否
exceluseprintarea	int	如果是 Excel 文件，是否使用打印区域，默认 0 不使用，1 使用，2 不使用且只显示有内容的区域	否
excelpagesize	int	如果是 Excel 文件，设定页面大小，默认 A4。0:A4，1:A3，2:A4，3:A5，4:B4，5:B5，6:Letter，7:Legal，8:Tabloid，9:Ledger	否
wordshowmarkup	int	如果是 Word 文件，是否显示审阅标记，默认 0 不显示，1 显示	否

参数	类型及范围	备注	是否必须发送
powerpointoutputtype	int	如果是 PPT 文件，设定导出样式，默认幻灯片。0:幻灯片，0 以上是讲义模式：1:每页一个幻灯片，2:每页二个幻灯片，3:每页三个幻灯片，4:每页四个幻灯片，5:每页六个幻灯片，6:每页九个幻灯片	否
powerpointhandoutorder	int	如果是 PPT 文件，当设置为讲义时，设定顺序，默认 0 水平，1 垂直	否
powerpointhandoutorientation	int	如果是 PPT 文件，当设置为讲义时，输出文件的方向，默认 0 不改变，1 横向，2 纵向	否
watermark	string	添加水印，字符个数最大 10 个，仅对非图片文件有效，需要 URL Encoding	否
watermarkfontsize	int	水印的字体大小，默认 24pt	否
watermarkfontcolor	string	水印的颜色，输入颜色码，默认黑色#000000，必须 7 位，需要 URL Encoding	否
watermarkfontalpha	int	水印的透明度，取值范围 1-100，越小越透明，默认 20	否
watermarkstyle	int	水印的样式，默认 0: 文档中央一个水印；1: 文档铺满水印	否
password	string	源文件的密码，支持有密码的 PDF, Word, PPT, Excel 文件类型，默认空无密码	否
pageindexes	string	如果源文件是 PDF 文件，指定转换的 PDF 页数，默认空全部页，例如：1, 3, 5-7 就是 1, 3, 5, 6, 7 共 5 页	否

参数	类型及范围	备注	是否必须发送
callbackurl	string	回调 URL，转换结束后，会回调该 URL，详见 <a href="#">回调 URL</a>	否

### 请求示例：

`https://all2img.market.alicloudapi.com/v1/convert_post`  
Header 中的 Content-Type 必须是 multipart/form-data

签名见阿里签名规则。

支持多种文件格式，具体如下（**type** 可传入如下格式）：

PDF 文档：pdf

微软 Office 文档：doc, docx, ppt, pptx, xls, xlsx, pot, pps, ppsx

WPS 文档：wps, wpt, dps, dpt, et, ett

苹果 iWork 文档：pages, key, numbers

开放版式文档：ofd

电子刊物：caj

电子书：epub, chm, mobi, azw, azw3, fb2, cbr, cbz, djvu

Markdown 格式：md

SVG 格式：svg

CAD 文档：dwg, dxf, dwt, dws

Sketch 文档：sketch

网页文件：html, htm, mht, eml

图片文件：几乎所有格式例如 png, jpg, jpeg, gif, tif, tiff, bmp, **可以统一传 img，代表一切图片**

文本文件：txt, rtf, java, js, c, cpp, jsp, css, xml, properties, log, 其他任意文本文件都可以传 txt

### 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"
}
```

## 多张图片转换 POST

将多张图片转换为长图/动图, 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

参数	类型及范围	备注	是否必须发送
type	string	固定传 img	是
outtype	int	输出类型：1：每页一个图；2：长图；3：动图	是
imagesize	int	图片大小。 如果 outtype 是 1（每页一个图），该值是每个图最大的宽或高。 如果 outtype 是 2（长图），该值是长图的宽。 如果 outtype 是 3（动图），该值是动图最大的宽或高。 默认是 0，自动。长图，动图宽最大 2000，其他图宽高最大 4096。	否
imageduration	int	动图多少秒切换下一张。如果 outtype 是 3（动图）才有效，单位是秒，比如传入 10 就是每隔 10s 切换一张图	否
imageformat	string	如果输出是长图，输出的图片格式。默认空：jpg，可填入 png	否
watermark	string	添加水印，字符个数最大 10 个，仅对非图片文件有效，需要 URL Encoding	否
watermarkfontsize	int	水印的字体大小，默认 24pt	否
watermarkfontcolor	string	水印的颜色，输入颜色码，默认黑色 #000000，必须 7 位，需要 URL Encoding	否
watermarkfontalpha	int	水印的透明度，取值范围 1-100，越小越透明，默认 20	否
watermarkstyle	int	水印的样式，默认 0：文档中央一个水印；1：文档铺满水印	否
callbackurl	string	回调 URL，转换结束后，会回调该 URL，详细见 <a href="#">回调 URL</a>	否

### 请求示例：

<https://all2img.market.alicloudapi.com/v1/convert>

传入的 Body 为 JSON 格式，如下：

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

签名见阿里签名规则。

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://api.duhuitech.com/q?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 - 1.00)	否 (status 为 Doing 时返回)	比如 0.88 表示 88%
imageurls	图片文件地址	string 数组	否 (status 为 Done 时返回)	转换出来的图片文件地址，http 和 https 都支持
count	文件总数	integer	否 (status 为 Done 时返回)	图片文件总数
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",
    "imageurls":[
      "https://file.duhuitech.com/o/xxx/1.jpg",
      "https://file.duhuitech.com/o/xxx/2.jpg"
    ],
    "count":2
  }
}
```

#### 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 **2 小时内** 下载文件。 (\*\*2025. 1. 1 之后变更为 1 小时)

## 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","imageurls":["https://file.duhuitech.com/o/xxx/1.jpg","https://file.duhuitech.com/o/xxx/2.jpg"],"count":2}}
```

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。

## 阿里云独有部分:

支持从阿里云 OSS 内网直接下载文件, 目前支持的是上海地区的阿里云 OSS 内网:

oss-cn-shanghai-internal.aliyuncs.com

文档转换 GET 或多张图片转换 POST 里的 url 地址包含上述域名则自动支持

## 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