



汉王云识别 SDK

(Server 版)

API-V1.0

表格

(v1.0.160115)

汉王科技-研发中心-互联网

2016/1/15

目录

- 1. 汉王开发者中心简介.....1
- 2. 开发环境及说明.....1
 - 2.1. 要求.....1
 - 2.2. 请求安全认证.....1
- 3. OCR 接口.....2
 - 3.1. 表格识别.....2
 - 3.1.1. 接口描述.....2
 - 3.1.2. 请求函数.....2
 - 3.1.3. 请求消息体：.....2
 - 3.1.4. 响应消息体（增强版 json）：.....2
 - 3.1.5. 返回值示例(增强版 json)：.....5
 - 3.1.6. 响应码列表：.....12
- 附录.....12

1. 汉王开发者中心简介

汉王立足于智能人机交互技术，提供成熟的“多元、智能、交互”技术与产品。汉王云授权业务是将汉王的多种人机交互模式识别技术，借助庞大的云计算能力，以云服务的方式提供给广大企业与个人，让模式识别触手可及。我们针对每种识别 API 都提供了免费额度，也同时提供了多种授权方式。每名注册用户首先要创建一个应用，然后再根据此应用申请多个授权，供您使用。

2. 开发环境及说明

2.1. 要求

支持所有开发语言，支持 http/https 协议。我们提供 Rest Webservice 用户，所有请求请使用 Post 请求。

2.2. 请求安全认证

请采用标准 https 方式进行所有请求的发送。用户首先需要登录汉王开发者中心进行用户注册，为了给真实的开发者提供服务，您申请的开发者账号将被我们进行审核。在您创建应用后，进行授权服务申请，获取您要开发的平台 Key,即可开始您的云识别开发。详情，请参考网站帮助。

网站地址：<http://developer.hanvon.com/>

3. OCR 接口

3.1. 表格识别

3.1.1. 接口描述

根据用户传递来的图片数据进行表格识别。推荐图片格式为 jpg，图片为正立图片，图片大小为 200K 以内，图片要尽可能清晰，字体不要过小，以获得更好的识别效果。

为了您的访问更加安全，我们提供了白名单认证功能，请到汉王开发者中心的 Key 管理栏目，在您的 Key 列表中找到对应的 Server Key，来指定您的白名单(参照网站帮助文档)。

3.1.2. 请求函数

```
public static String recgTable(String key, String path, String type);
```

3.1.3. 请求消息体：

字段	出现次数	可为空	数据类型	描述
key	1	否	String	请到汉王开发者中心的 Key 管理栏目，在您的 Key 列表中找到对应的 Server Key
path	1	否	String	文件全路径
type	1	否	String	返回值类型。 1:纯文本 2: json 3:增强版 json

3.1.4. 响应消息体（增强版 json）：

字段	父字段	数据类型	描述
code	response	数值	返回码

result	response	字符串	返回码所对应的信息
block_count	response	数值	文本块的总数
table_count	response	数值	表格的总数
block_list	response	数组	文本块的集合
table_list	response	数组	表格的集合
block_list 结构如下			
order	block_list	字符串	在全篇文本的块顺序号
line_list	block_list	数组	一个文本块里行的集合
line_num	line_list	数值	一个文本块的行号
text	line_list	字符串	识别出来的文字内容，第一候选
confidence	line_list	字符串	每个文字的置信度,从 A 到 E，置信度一次降低

pos	line_list	字符串	每个文字的位置，文字间用 分隔，位置之间用-分隔
table_list 结构如下			
order	table_list	数值	一个表格在全篇文本的块顺序号
tr_count	table_list	数值	一个表格行数
td_count	table_list	数值	一个表格列数
table	table_list	数组	表示一个表格
tr	table	数组	一个表格里的一行
td	tr	对象	某一行里的一列
colspan	td	数值	单元格可横跨的列数
rowspan	td	数值	单元格可纵跨的行数
line_list	td	数组	一个单元格里行的集合

line_num	line_list	数值	一个文本块的行号
text	line_list	字符串	识别出来的文字内容，第一候选
confidence	line_list	字符串	每个文字的置信度,从 A 到 E，置信度一次降低
pos	line_list	字符串	每个文字的位置，文字间用 分隔，位置之间用-分隔

3.1.5. 返回值示例(增强版 json) :

Success :

```
{
  "code": "0",
  "result": "ESC_SUCCESS",
  "block_count": "2",
  "table_count": "1",
  "block_list":
  [
    {
      "order": "1",
      "line_list":
      [
        {
          "line_num": "0",
          "text": "表头 1 表头 2",
          "confidence": "AAAAAA",
```

```

        "pos": "160-42-192-74|197-42-229-75|237-47-246-71|466-42-498-74|503-
42-535-75|540-47-554-71"
    }
]
},
{
    "order": "3",
    "line_list":
    [
        {
            "line_num": "0",
            "text": "表尾",
            "confidence": "AA",
            "pos": "323-354-355-385|360-355-392-385"
        }
    ]
}
],
"table_list":
[
    {
        "order": "2",
        "tr_count": "3",
        "td_count": "3",
        "table":
        [
            {
                "tr":
                [
                    {
                        "td":
                        {

```



```

        "colspan": "1",
        "rowspan": "1",
        "line_list":
        [
        ]
    },
    {
        "td":
        {
            "colspan": "1",
            "rowspan": "1",
            "line_list":
            [
                {
                    "line_num": "0",
                    "text": "第一列",
                    "confidence": "AAA",
                    "pos": "304-145-335-177|341-157-373-160|378-
145-409-177"
                }
            ]
        }
    },
    {
        "td":
        {
            "colspan": "1",
            "rowspan": "1",
            "line_list":
            [
                {

```

```

        "line_num": "0",
        "text": "笔一列",
        "confidence": "DAA",
        "pos": "507-145-537-171|547-151-572-152|581-145-611-171"
    }
}
],
{
    "tr":
    [
        {
            "td":
            {
                "colspan": "1",
                "rowspan": "1",
                "line_list":
                [
                    {
                        "line_num": "0",
                        "text": "第一行",
                        "confidence": "AAA",
                        "pos": "102-198-133-230|139-210-171-213|176-198-208-230"
                    }
                ]
            }
        }
    ]
},
{

```

```
"td":
{
  "colspan": "1",
  "rowspan": "1",
  "line_list":
  [
    {
      "line_num": "0",
      "text": "11",
      "confidence": "AA",
      "pos": "344-203-353-227|363-203-372-227"
    }
  ]
},
{
  "td":
  {
    "colspan": "1",
    "rowspan": "1",
    "line_list":
    [
      {
        "line_num": "0",
        "text": "12",
        "confidence": "AA",
        "pos": "546-203-555-227|562-203-576-227"
      }
    ]
  }
}
```

```

    },
    {
      "tr":
      [
        {
          "td":
          {
            "colspan": "1",
            "rowspan": "1",
            "line_list":
            [
              {
                "line_num": "0",
                "text": "第二行",
                "confidence": "AAA",
                "pos": "102-251-133-283|139-257-171-278|176-
251-208-283"
              }
            ]
          }
        }
      ],
    },
    {
      "td":
      {
        "colspan": "1",
        "rowspan": "1",
        "line_list":
        [
          {
            "line_num": "0",
            "text": "21",
            "confidence": "AA",

```

```

        "pos": "341-256-355-280|363-256-372-280"
    }
}
],
},
{
    "td":
    {
        "colspan": "1",
        "rowspan": "1",
        "line_list":
        [
            {
                "line_num": "0",
                "text": "22",
                "confidence": "AA",
                "pos": "543-256-557-280|562-256-576-280"
            }
        ]
    }
}
]
}
]
}
]
}
}

Failure:
{
    "code": "437",

```

```
"result": ": input is null"
}
```

3.1.6. 响应码列表：

代码	描述
0	success
430	key illegal 没有有效的 key
431	whitelist illegal 请求不在白名单范围
434	serviceCode illegal 您没有申请相应的服务，或服务的次数已到临界值，或服务已到期
437	input is null 您的请求参数为空
438	input json format invalid 您的请求参数 json 格式非法
439	image data is null 您的请求参数中图片数据为空
8001	recognize service exception 识别服务错误
8002	recognize core exception 您的图片样本不够清晰

附录

1. 文本尽可能充满整个画面，不要出现半个字的图片，背景不要复杂。
2. 拍摄时，请保持光线充足，不要有阴影和反光

3. 请对焦后进行拍摄，尽量避免抖动。
4. 为提高识别率，文本分辨率最少 300dpi 以上。