



<搜象 OCR 增值税卷票结构化识别> 接口开发参考

搜象（厦门）信息技术有限公司

2022 年 02 月 21 日

目 录

1	第一章. 引擎接口开发	3
1.1	接口概述	3
1.2	服务请求报文	3
1.2.1	请求报文示例.....	3
1.2.2	请求数据项说明.....	3
1.3	服务返回报文	3
1.3.1	返回报文示例.....	3
1.3.2	返回数据项说明.....	6
1.4	失败返回码定义	7
1.4.1	失败返回示例.....	7
1.4.2	失败返回码定义.....	7
1.5	附-识别字段.....	7
1.5.1	增值税卷票.....	7

1 第一章. 引擎接口开发

1.1 接口概述

请求地址：请参考商品页 API 接口调用地址

请求方式：POST

请求格式：Content-Type: application/json

返回类型：JSON

1.2 服务请求报文

1.2.1 请求报文示例

```
{
  "fileType": 1,
  "filePath": "http://192.168.0.2:8080/test.jpg"
}
```

1.2.2 请求数据项说明

数据项	数据类型	描述
fileType	int	识别文件类型，1-图片，2-PDF，不传则自动判断
filePath	string	识别文件路径（filePath 和 fileBase64 二者择一传参即可）
fileBase64	string	识别文件 Base64 编码（需带 Base64 头，如：图片【data:image/jpeg;base64,】, PDF【data:application/pdf;base64,】）
pageIndex	int	识别 PDF 第几页，不传则默认首页，首页从 0 开始，该参数对 PDF 有效

1.3 服务返回报文

1.3.1 返回报文示例

```
{
  "code": 0,
```

```

"data": {
  "angle": 0.0,
  "anglePic": 0,
  "columnResultList": [
    {
      "columnKey": "invoice_type",
      "columnName": "发票类型",
      "columnValue": "普通发票"
    },
    {
      "columnKey": "couplet_type",
      "columnName": "抵扣/发票联",
      "columnValue": "发票联"
    },
    {
      "columnKey": "invoice_code",
      "columnName": "发票代码",
      "columnValue": "033000100101"
    },
    {
      "columnKey": "invoice_number",
      "columnName": "发票号码",
      "columnValue": "14547954"
    },
    {
      "columnKey": "machine_code",
      "columnName": "机器编号",
      "columnValue": "661535951900"
    },
    {
      "columnKey": "seller_name",
      "columnName": "销售方名称",
      "columnValue": "杭州萧山宁东万丰加油站有限公司"
    },
    {
      "columnKey": "buyer_identifier",
      "columnName": "购买方纳税人识别号",
      "columnValue": "913301097441052544"
    },
    {
      "columnKey": "invoice_date",
      "columnName": "开票日期",
      "columnValue": "2019-07-14"
    }
  ]
}

```

```
    },  
    {  
      "columnKey": "cashier",  
      "columnName": "收款员",  
      "columnValue": "高高"  
    },  
    {  
      "columnKey": "buyer_name",  
      "columnName": "购买方名称",  
      "columnValue": "浙江"  
    },  
    {  
      "columnKey": "seller_identifier",  
      "columnName": "销售方纳税人识别号",  
      "columnValue": "913302087572467338"  
    },  
    {  
      "columnKey": "project_col",  
      "columnName": "项目列",  
      "columnValue": "*汽油*924 车用汽油 (VIA) "  
    },  
    {  
      "columnKey": "unit_price_col",  
      "columnName": "单价列",  
      "columnValue": "6.75"  
    },  
    {  
      "columnKey": "number_col",  
      "columnName": "数量列",  
      "columnValue": "29.63"  
    },  
    {  
      "columnKey": "amount_col",  
      "columnName": "金额列",  
      "columnValue": "200.00"  
    },  
    {  
      "columnKey": "total_amount_low",  
      "columnName": "合计金额(小写/元)",  
      "columnValue": "200.00"  
    },  
    {  
      "columnKey": "total_amount_cap",
```

```

        "columnName": "合计金额(大写)",
        "columnValue": "贰佰圆"
    },
    {
        "columnKey": "check_code",
        "columnName": "校验码",
        "columnValue": "83931023061128542450"
    }
],
"templateCatelId": "825454941819633664",
"templateCateName": "增值税发票",
"templateId": "855496572866134016",
"templateName": "增值税发票(卷票)"
},
"message": "success"
}

```

1.3.2 返回数据项说明

数据项	数据类型	描述
code	int	返回编码，0-代表成功，其余失败
data	object	返回具体数据项
+angle	float	文字顺时针偏转方向
+anglePic	int	图片四分类方向，0-正向, 90-右向, 180-反向, 270-左向
+templateCateId	string	模板分类 ID
+templateCateName	string	模板分类名称
+templateId	string	模板 ID
+templateName	string	模板名称
+columnResultList	objectArray	文字识别数据项列表
++columnName	string	字段中文名称
++columnKey	string	字段英文名称
++columnValue	string	字段值（\n 字符表示字段值中的换行或较大间隔）
++columnBox	intArray	字段坐标->[x1,y1,x2,y2,x3,y3,x4,y4]（非固定项，坐标方式提取的字段有返回，规则方式提取的字段无返回）
message	string	返回信息，包括成功失败信息

1.4 失败返回码定义

1.4.1 失败返回示例

```
{
  "code": 1011,
  "data": {},
  "message": "无识别文件"
}
```

1.4.2 失败返回码定义

返回编码	返回信息	描述
1011	无识别文件	filePath 和 fileBase64 同时为空
1012	不支持的文件类型	支持图片和 PDF，图片包括但不限于 jpg、jpeg、png、bmp、gif、tif、tiff、jfif、heic、jpe、dib 等格式
1014	文件 URL 错误	文件 URL 有误，读取不到资源
1015	文件 BASE64 串错误	文件 BASE64 串有误，解析文件异常
1099	服务器端异常	返回具体异常信息

1.5 附-识别字段

1.5.1 增值税卷票

字段中文名称	字段英文名称	字段中文名称	字段英文名称
发票类型	invoice_type	购买方名称	buyer_name
抵扣/发票联	couplet_type	销售方纳税人识别号	seller_identifier
发票代码	invoice_code	项目列	project_col
发票号码	invoice_number	单价列	unit_price_col
机打号码	print_number	数量列	number_col
机器编号	machine_code	金额列	amount_col
销售方名称	seller_name	合计金额(小写/元)	total_amount_low
购买方纳税人识别号	buyer_identifier	合计金额(大写)	total_amount_cap
开票日期	invoice_date	校验码	check_code
收款员	cashier		