



昆仑银行集中采购项目需求说明书

(2021 年网贷业务系统建设项目-生物特征识别平台子
项目)

昆仑银行股份有限公司

2021 年

谢彬 吴宇斌

1) 项目背景

网贷平台上线近两年来，业务量累计 1200 万笔，累放金额约 340 亿元，有力地支持了全行互联网贷款业务的快速发展。随着 2020 年 7 月 12 号《商业银行互联网贷款暂行管理办法》发布（下称“监管新规”），对我行业务发展提出了新的适应性改造要求。此外网贷系统运营期间问题频发，性能堪忧，为快速解决业务问题，更好的助力网贷业务向更高的目标迈进，需对现有系统进行优化。

2) 项目需求

- ✧ 项目名称：2021 年网贷业务系统建设项目-生物特征识别平台子项目
- ✧ 原供应厂商：重庆中科云从科技有限公司
- ✧ 项目需求：升级目前的引擎版本，在已有生物特征识别平台中新增增值税发票 OCR 模块，提供增值税发票 OCR 的不限机器，不限次数的永久授权。

2.1. 引擎平台升级内容如下：

原则基于现有平台新增增值税发票接口，作为统一接口输出渠道调用增值税发票 OCR 识别。需提供增值税发票 OCR 的永久授权且不限制机器，不限时间，不限授权数量但仅限昆仑银行使用。

增值税发票 OCR 引擎对接引擎平台。

引擎平台开通新的增值税发票 OCR 渠道，为渠道应用服务。

同时引擎平台保存渠道业务调用 OCR 流水，提供用户查看。

2.2. 增值税发票 OCR 要求内容如下：

➤ 简介

增值税票据 OCR 识别引擎支持识别各种增值税专票、普票、电子发票图片上面的文字信息并支持返回各字段的位置信息，支持多角度任意方向的增值税发票检测识别。

产品中使用的接口有增值税票据 OCR 识别和增值税票据 OCR 识别 Echo，均使用 Http 协议实现。调用时，将 FaceGo 增值税票据 OCR 识别引擎所在的 IP 地址、端口配置填入 URL 中的 ip、port 字段部分，进行调用。

➤ **识别要求**

1. 图片大小要求：小于 3M
2. 图片格式要求：jpg、bmp、png

功能说明：

1. 支持识别字段：发票类型、发票代码、发票号码、开票日期、校验码、购买方名称、购买方纳税人识别号、购买方地址（电话）、购买方开户行及账号、合计金额、合计税额、价税合计（大写）、小写、销售方名称、销售方纳税人识别号、销售方地址（电话）、销售方开户行及账号、税额、金额、货物或应税劳务（服务名称）、开票人、复核、收款人、单位、规格型号、数量、单价、税率、机器编号、密码区
2. 支持多角度输入识别，包括旋转和透视的图片输入
3. 对非增值税发票图片可拒识，即识别为空。

2.2.1. 请求格式规范

2.2.1.1. HTTP 接口协议

协议适用于和应用平台应用进行数据交换对接，采用 TCP/IP 通信方式。

应用平台为服务端，开放指定 TCP 端口进行监听，默认为 20007；其它应用平台作为客户端，根据其配置的端口，主动向平台服务器发送请求连接。

2.2.1.1.1. 安全认证方式

用户在每次发送请求时，需要带上公用字段 app_id,app_secret，每次请求时验证 id 与 secret 是否正确，其中，超级用户的 app_id，app_secret 分别默认为 system，12345

2.2.1.1.2. 通用请求头

下表是其它第三方接入协议通用请求头(HTTP Headers):

Header Name	DescrIPtion
Content-Type	消息体的类型，如:application/json
Host	服务器的域名或IP地址，如192. 168. 1. 10
Port	服务使用的监听端口，如23001
Connection	是否启用超时时间，如果启用超时时间可以设置TCP长连接有效时长（建议大于30秒）

表 0-1

HTTP POST 格式定义

```
POST /RequestName HTTP/1.1
Host:192. 168. 1. 10
Accept-language: zh-cn
Date: Sat, 31 Dec 2015 23:59:59 GMT
Content-Type:application/json;charset=UTF-8
Content-Length: length
X-Authorization: Bearer Token 字符串
Connection: Keep-Alive / close
```

发送的文本信息

2.2.1.1.3. 通用响应头

增值税票据 OCR 识别引擎返回的通用头部:

Header Name	DescrIPtion
Content-Length	消息体的长度，不包括请求头部
Content-Type	消息体的类型，如:application/json;charset=UTF-8
Date	请求端的当前本地时间，例如:12 Dec 2014 12:00:00 GMT
Server	响应该请求的服务器的域名或IP地址

表 0-2

HTTP 响应格式

HTTP/1.1 200 OK

Content-Type: application/json; charset=utf-8

Date: Sat, 31 Dec 2015 23:59:59 GMT

Content-Length: length

响应的文本信息（采用JSON格式）

2.2.1.1.4. 常见响应状态码

Status-Code	DescrIPtion
200 OK	请求成功。
302 Found	请求的资源临时具有不同URI（重定向）。
400 Bad Request	请求的语法不对，报文格式错误。
401 Unauthorized	无权访问该资源，请求需要用户授权。
403 Forbidden	服务器拒绝该请求，一般为账号异常。
404 Not Found:	服务器没有找到与请求URI相符的资源。
500 Internal Server Error	服务器发生了不可预期的错误。
503 Server Unavailable	服务器请求暂时不可用。

表 0-3

2.2.1.2. 响应编码规范

响应数据包的格式默认为JSON格式，输出内容使用UTF-8编码。

2.2.1.3. 公共响应报文格式

每次请求响应的报文格式必然包括如下部分。

名称	类型	描述	名称	类型
result	int	非0表示错误码	result	int
seq	string	会话序号	seq	string
Info	string	返回的错误信息	Info	string

表 0-4

2.2.2. 接口定义

2.2.1.1. 增值税票据 OCR 识别引擎 - 增值税票据 OCR 识别

对增值税票据图片进行识别

● POST
http://ip:port/ocr/note

● 使用示例
http://ip:port/ocr/note
{
 "app_id":"system",
 "app_secret":"12345",
 "img": "",
 "seq": "f428b43e-5bc1-4bea-a93a-ed177ca4552e"
}

● 参数

表 3-1

字段	类型	描述
app_id	string	应用访问账号
app_secret	string	应用访问密码
seq 可选	string	会话序号(非必填)
img	string	图片数据(base64 编码), 原始图片建议小于 3M, 格式限定为 jpg bmp png 三种格式

● 成功返回参数

表 3-2

字段	类型	描述
result	number	0 表示成功
seq	string	会话序号
noteOcrRet	object	票据 OCR 识别结果

字段	类型	描述
NoteType	object[]	票据类型（即票据抬头部分，譬如 xxx 省增值税专用发票）
data	string	票据类型识别结果
pos	object	票据类型位置信息 (p1:左上坐标; p2:右上坐标; p3:右下坐标; p4:左下坐标;)
NoteCode	object[]	票据代码
NoteNo	object[]	票据号
NoteDate	object[]	开票日期
BuyerName	object[]	购货单位名称
BuyerTaxID	object[]	购货单位纳税人识别号
BuyerAddrTel	object[]	购货单位地址电话
BuyerBank	object[]	购货单位开户行及账号
SellerName	object[]	销货单位名称
SellerTaxID	object[]	销货单位纳税人识别号
SellerAddrTel	object[]	销货单位地址电话
SellerBank	object[]	销货单位开户行及账号
CipherCode	object[]	密码区
TotalAmount	object[]	金额总计
TotalTax	object[]	税额总计
PriceTaxTotalKanji	object[]	价税合计大写
PriceTaxTotalDigit	object[]	价税合计小写
Payee	object[]	收款人
Checker	object[]	复核
Drawer	object[]	开票人
CheckCode	object[]	普票校验码
MachineID	object[]	机器编码
ItemName	object[]	货物或应税劳务服务名称
Specification	object[]	规格型号
Unit	object[]	单位
Quantity	object[]	数量
UnitPrice	object[]	单价
ItemAmount	object[]	金额

字段	类型	描述
Tax	object[]	税率
ItemTax	object[]	税额

● 成功返回示例

```
{
  "debugInfo": {
    "elapsed": 2577,
    "path": "/ocr/note"
  },
  "info": "success",
  "noteOcrRet": {
    "BuyerAddrTel": [
      {
        "data": "公主岭市八里桥村 0434659261",
        "pos": {
          "p1": {
            "x": 1236,
            "y": 641
          },
          "p2": {
            "x": 1952,
            "y": 676
          },
          "p3": {
            "x": 1951,
            "y": 733
          },
          "p4": {
            "x": 1241,
            "y": 704
          }
        }
      }
    ],
    .....
    "TotalTax": [
      {
        "data": "¥1452.9",
        "pos": {
          "p1": {
            "x": 3519,
            "y": 1438
          },
          "p2": {
            "x": 3848,
            "y": 1442
          }
        }
      }
    ]
  }
}
```



```

        "p3": {
            "x": 3847,
            "y": 1499
        },
        "p4": {
            "x": 3518,
            "y": 1503
        }
    }
}
]
},
"result": 0,
"seq": "1550815625549"
}

```

- 失败返回参数

表 3-3

名称	类型	描述
result	number	非 0 表示错误码
seq	string	会话序号

- 失败返回示例

```

{
    "debugInfo": {
        "elapse": 299,
        "path": "/ocr/note"
    },
    "info": "票据 OCR 识别错误",
    "result": 20001051,
    "seq": "598f2a91-7fde-465c-95fa-425d35dfff0e"
}

```