文字识别 产品简介





【版权声明】

©2013-2025 腾讯云版权所有

本文档(含所有文字、数据、图片等内容)完整的著作权归腾讯云计算(北京)有限责任公司单独所有,未经腾讯云 事先明确书面许可,任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成 对腾讯云著作权的侵犯,腾讯云将依法采取措施追究法律责任。

【商标声明】



腾讯云

及其它腾讯云服务相关的商标均为腾讯云计算(北京)有限责任公司及其关联公司所有。本文档涉及的第三方主体的 商标,依法由权利人所有。未经腾讯云及有关权利人书面许可,任何主体不得以任何方式对前述商标进行使用、复 制、修改、传播、抄录等行为,否则将构成对腾讯云及有关权利人商标权的侵犯,腾讯云将依法采取措施追究法律责 任。

【服务声明】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况,部分产品、服务的内容可能不时有所调整。 您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定,除非双方另有约定,否则, 腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【联系我们】

我们致力于为您提供个性化的售前购买咨询服务,及相应的技术售后服务,任何问题请联系 4009100100或 95716。



文档目录

产品简介

产品概述

通用文字识别

卡证文字识别

票据单据识别

特定场景识别

文档智能

智能扫码

文本图像增强

基本概念



产品简介产品概述

最近更新时间: 2025-05-07 16:56:32

腾讯云文字识别(Optical Character Recognition,OCR)基于行业前沿的深度学习技术,将图片上的文字内容智能识别成为可编辑的文本。支持通用文字、卡证文字、票据单据、特定场景等多场景下的印刷体、手写体文字识别,同时支持票据核验功能,支持提供定制化服务,可以有效地代替人工录入信息。

通用文字识别

卡证文字识别

票据单据识别

特定场景识别

文档智能

文本图像增强

智能扫码

OCR 文字识别接入指引

观看视频

功能体验

文字识别 OCR 功能演示 Demo 单击 免费体验 。

使用要求

用户需要具备编程开发能力,熟悉 Java/Python/PHP/Node.js/C++ 等编程语言。

文字识别 OCR 服务需要用户进行二次开发对接,然后通过调用 API 接口,实现将图片或者扫描件中的文字识别成可编辑的文本。文字识别提取之后返回的结果是不能够直接编辑使用的,需要用户通过编码将识别结果对接到用户的业务系统,或者是保存为 TXT、Excel 等文件格式。

用户可以通过 API 3.0 Explorer 进行在线调用、签名验证、SDK 代码生成和快速检索接口等能力。



通用文字识别

最近更新时间: 2024-08-23 14:52:45

产品介绍

通用文字识别(General Optical Character Recognition,General OCR)基于行业前沿的深度学习技术,提供通用印刷体识别、通用印刷体识别(高精度版)、通用手写体识别、英文识别、表格识别等多种服务,支持将图片上的文字内容,智能识别为可编辑的文本,可应用于随手拍扫描、纸质文档电子化、电商广告审核等多种场景,大幅提升信息处理效率。

观看视频

产品功能

通用印刷体识别

支持多场景、不同版面下整图文字的识别。支持自动识别语言类型,同时支持自选语言种类(推荐),除中英文外, 支持日语、韩语、西班牙语、法语、德语、葡萄牙语、越南语、马来语、俄语、意大利语、荷兰语、瑞典语、芬兰 语、丹麦语、挪威语、匈牙利语、泰语等多种语言。

通用印刷体识别(高精度版)

支持图像整体文字的检测和识别,返回文字框位置与文字内容。相比通用印刷体识别接口,准确率和召回率更高,覆盖场景更广泛,应用场景包括:印刷文字识别、网络图片文字识别、广告图文字识别、街景店招文字识别、菜单文字识别、视频标题文字识别、头像文字识别等。

通用手写体识别

支持多场景、不同版面下整图文字中的手写中文、英文、字母、数字、常见字符的识别。针对手写字体无规则、字迹 潦草、模糊等特点进行了识别能力的增强。可应用于银行、保险、金融等手写单据录入、教育行业笔记文档化等场 景。

英文识别

支持图像英文文字的检测和识别,返回文字框位置与文字内容。支持多场景、不同版面下的英文、字母、数字和常见 字符的识别,同时覆盖英文印刷体和英文手写体识别。可应用于英文笔记作业文档化、英文试卷答题卡识别等场景。

表格识别(V2)

支持中英文图片/ PDF 内常规表格、无线表格、多表格的检测和识别,支持日文有线表格识别,返回每个单元格的文字内容,支持旋转的表格图片识别,支持将识别结果保存为 Excel 格式。

表格识别(V3)

版权所有:腾讯云计算(北京)有限责任公司 第5 共18页



支持中英文图片/PDF内常规表格、无线表格、多表格的检测和识别,返回每个单元格的文字内容,支持旋转的表格图片识别,且支持将识别结果保存为 Excel 格式。识别效果比表格识别更好,覆盖场景更加广泛,对表格难例场景,如无线表格、嵌套表格(有线表格中包含无线表格)的识别效果均优于表格识别V2,同时能避免部分印章干扰、表格线断裂等影响因素。适用于对接口准召率要求更高的客户。

广告文字识别

支持广告商品图片内文字的检测和识别,返回文本框位置与文字内容。支持中英文、横排和竖排文字识别,支持90度、180度、270度翻转以及倾斜场景文字识别。文字识别的召回率和准确率能达到96%以上。



卡证文字识别

最近更新时间: 2024-12-31 09:40:02

产品介绍

卡证文字识别(Card Optical Character Recognition,Card OCR)基于行业前沿的深度学习技术,提供身份证识别、银行卡识别、驾驶证、行驶证、营业执照识别、护照识别、港澳台通行证识别、港澳台来往内地通行证识别、港澳台居住证识别、中国香港身份证、名片识别、机动车登记证书识别等多种服务,支持将图片上的文字内容,智能识别为结构化的文本,可应用于用户注册、银行开户、交通出行、政务办事等多种场景,大幅提升信息处理效率。

产品功能

身份证识别

支持中国大陆居民二代身份证正反面所有字段的识别,包括姓名、性别、民族、出生日期、住址、公民身份证号、签发机关、有效期限;具备身份证照片、人像照片的裁剪功能和翻拍、PS、复印件告警功能,以及边框和框内遮挡告警、临时身份证告警和身份证有效期不合法告警等扩展功能。可应用于各种身份证信息有效性校验场景,例如用户注册、银行开户等。

身份证识别(安全加密版)

身份证识别(安全加密版)接口实现了数据加密传输,能够有效防止个人身份证隐私信息不被窃取泄露。本接口支持中国大陆居民二代身份证正反面所有字段的识别,包括姓名、性别、民族、出生日期、住址、公民身份证号、签发机关、有效期限,识别准确度达到99%以上。产品还支持多种扩展能力,满足不同场景的需求。如身份证照片、人像照片的裁剪功能,同时具备9种告警功能。

银行卡识别

支持对中国大陆主流银行卡正反面关键字段的检测与识别,包括卡号、卡类型、卡名字、银行信息、有效期。支持竖排异形卡识别、多角度旋转图片识别。可应用于各种银行卡信息有效性校验场景,如金融行业身份认证、第三方支付 绑卡等场景。

驾驶证识别

支持驾驶证主页和副页所有字段的自动定位与识别,主页识别字段包括证号、姓名、性别、国籍、住址、出生日期、初次领证日期、准驾车型、有效期限、发证单位;副页识别字段包括证号、姓名、档案编号、记录。驾驶证识别重点字段的识别准确度达到99%以上,同时具备复印件、翻拍和 PS 告警功能。同时支持识别交管12123 App 发放的电子驾驶证正页。电子驾驶证正页识别字段包括证号、姓名、性别、国籍、出生日期、初次领证日期、准驾车型、有效期开始时间、有效期截止时间、档案编号、状态、累积记分。

行驶证识别



支持对行驶证主页和副页所有字段的自动定位与识别,包含车牌号码、车辆类型、所有人、住址、使用性质、品牌型 号、车辆识别代码、发动机号、注册日期、发证日期、发证单位等。

营业执照识别

支持快速精准识别营业执照上的字段,包括统一社会信用代码(三合一之前为注册号)、公司名称、主体类型、法定 代表人、注册资本、组成形式、成立日期、营业期限和经营范围等字段。可应用于企业在线审核、企业线上办公等场 景,为企业服务提供便利。

护照识别(中国大陆地区护照)

支持中国大陆地区护照个人资料页多个字段的检测与识别。已支持字段包括英文姓名、中文姓名、国家码、护照号、 出生地、出生日期、国籍英文、性别英文、有效期、签发地点英文、签发日期、持证人签名、护照机读码(MRZ 码)等。

护照识别(港澳台地区及境外护照)

支持中国港澳台地区以及其他国家、地区的护照。识别字段包括护照 ID、姓名、出生日期、性别、有效期、发行国、国籍、国家地区代码等;具备护照人像照片的裁剪功能和翻拍、复印件告警功能。

港澳台通行证识别

支持对卡式港澳台通行证的识别,包括签发地点、签发机关、有效期限、性别、出生日期、英文姓名、姓名、证件号 等字段。

港澳台来往内地通行证识别

支持港澳台来往内地通行证的识别,智能识别并结构化港澳台居民来往内地通行证正面全部字段,包含中文姓名、英文姓名、性别、出生日期、签发机关、有效期限、证件号、签发地点、签发次数、证件类别。

港澳台居住证识别

支持港澳台居住证正反面全字段内容检测识别功能,包括姓名、性别、出生日期、地址、身份证号、签发机关、有效 期限、签发次数、通行证号码关键字段识别。可以应用于港澳台居住证信息有效性校验场景,例如银行开户、用户注 册等场景。

中国香港身份证识别

支持中国香港身份证人像面中关键字段的识别,包括中文姓名、英文姓名、姓名电码、出生日期、性别、证件符号、 首次签发日期、最近领用日期、身份证号、是否是永久性居民身份证;具备防伪识别、人像照片裁剪等扩展功能。

名片识别

支持对名片各字段的自动定位与识别,包含姓名、电话、手机号、邮箱、公司、部门、职位、网址、地址、QQ、微信、MSN等。可应用于社交平台商务应用、CRM软件等场景。

机动车登记证书识别

版权所有:腾讯云计算(北京)有限责任公司 第8 共18页



支持中国内地(大陆)机动车登记证书主要字段的结构化识别,包括机动车所有人、身份证明名称、号码、车辆型 号、车辆识别代号、发动机号、制造厂名称等。

智能卡证分类

支持身份证、护照、名片、银行卡、行驶证、驾驶证、港澳台通行证、户口本、港澳台来往内地通行证、港澳台居住证、不动产证、营业执照的智能分类。



票据单据识别

最近更新时间: 2024-08-23 10:47:00

产品介绍

票据单据识别(Invoice Optical Character Recognition,Invoice OCR)基于行业前沿的深度学习技术, 提供增值税发票识别、火车票识别、出租车票识别、机票行程单识别、完税证明识别等多种服务,支持将图片上的文 字内容,智能识别为结构化的文本,可应用于企业票据报销、金融票据识别、快递单据录入等多种场景,大幅提升信 息处理效率。

通用票据识别(高级版)

针对可报销发票支持 PDF多页、一页中单张、多张、类型票据混合识别,已支持票种包括:增值税发票(专票、普票、卷票、区块链发票、通行费发票)、全电发票(专票、普票)、非税发票(通用票据、统一缴纳书)、定额发票、通用机打发票、购车发票(机动车销售发票、二手车发票)、火车票、出租车发票、机票行程单、汽车票、轮船票、过路过桥费发票共14大类标准报销发票,同时支持非上述类型的其他发票的智能识别。

增值税发票核验(新版)

支持增值税发票的准确性核验,您可以通过输入增值税发票的关键字段提供所需的验证信息,接口返回真实的票面相 关信息,包括发票代码、发票号码、开票日期、金额、消费类型、购方名称、购方税号、销方名称、销方税号等多个 常用字段。支持多种发票类型核验,包括增值税专用发票、增值税普通发票(含电子普通发票、卷式发票、通行费发 票)、全电发票、机动车销售统一发票、货物运输业增值税专用发票、二手车销售统一发票、通用机打电子发票(广 东和浙江)。

增值税发票识别

支持增值税发票(专票、普票)、全电发票(专票、普票)的字段结构化识别,包括发票代码、发票号码、打印发票代码、打印发票号码、开票日期、合计金额、校验码、税率、合计税额、价税合计、购买方识别号、复核、销售方识别号、开票人、密码区1、密码区2、密码区3、密码区4、发票名称、购买方名称、销售方名称、服务名称、备注、规格型号、数量、单价、金额、税额、收款人等字段。

OFD 发票识别

支持OFD格式的增值税发票(专票、普票)、全电发票(专票、普票)的结构化识别,返回发票代码、发票号码、 开票日期、验证码、机器编号、密码区,购买方和销售方信息,包括名称、纳税人识别号、地址电话、开户行及账 号,以及价税合计、开票人、收款人、复核人、税额、不含税金额等字段信息。

银行回单识别

本接口支持银行回单全字段的识别,包括付款开户行、收款开户行、付款账号、收款账号、回单类型、回单编号、币种、流水号、凭证号码、交易机构、交易金额、手续费、日期等字段信息。

版权所有:腾讯云计算(北京)有限责任公司 第10 共18页



完税证明识别

支持对企业通用完税证明、个人纳税证明结构化识别,支持字段包括:税号、纳税人识别号、纳税人名称、金额合计 大写、金额合计小写、填发日期、税务机关、填票人等关键字段的识别。

医疗发票识别

医疗发票识别目前支持全国统一门诊发票、全国统一住院发票、以及部分地方的门诊和住院发票的识别。

运单识别

支持市面上主流版式电子运单的识别,字段包括收件人和寄件人的姓名、电话、地址以及运单号等字段。

网约车行程单识别

支持网约车行程单关键字段的识别,包括行程起止日期、上车时间、起点、终点、里程、金额等字段。

保险单据识别

支持常见版式的保险理赔单据的字段识别,包含病案首页、费用清单、结算单、医疗发票的姓名、性别、出生日期、 出院诊断、疾病编码、入院病情等。

产品优势

能力全面

支持可报销票据单据的全种类识别,如增值税发票、定额发票、购车发票、火车票、出租车发票、机票行程单、电子运单等,覆盖字段多、识别信息全面。

准确率高

各字段精度均处于业界领先水平,以增值税发票为例,发票代码、价税合计、合计金额、购买方识别号、开票日期等识别准确率达到95%以上。

支持核验

支持多种类型的发票核验,包括增值税(专票、普票)、全电发票(专票、普票)、机动车销售统一发票、二手车销售统一发票。

应用场景

企业费控报销管理

适用于企业的费控报销管理系统,通过对票面信息的快速获取及核验,实现财务发票的电子化存档,提高费控报销的 审核效率。

财税代理记账

版权所有:腾讯云计算(北京)有限责任公司 第11 共18页



适用于代理记账企业,通过对票据信息的获取及核验,实现财务票据电子化管理,减少人工操作及票据丢失成本,提 高代账管理效率。

第13 共18页



特定场景识别

最近更新时间: 2024-08-23 10:47:18

产品介绍

特定场景识别(Specific Scenario Optical Character Recognition,Specific Scenario OCR)基于行业前沿的深度学习技术,提供车牌识别、车辆 VIN 码识别、印章识别等特定场景识别服务,支持将图片上的文字内容,智能识别为结构化的文本,可应用于汽车保险理赔、车辆租赁、企业方案管理等多种行业场景,大幅提升信息处理效率。

车牌识别

支持对中国大陆机动车车牌的自动定位和识别,返回地域编号和车牌号码与车牌颜色信息。

车辆vin码识别

支持图片内车辆识别代号(VIN)的检测和识别。

印章识别

支持各类印章,包括发票章,财务章等,适用于公文,票据等场景。

产品优势

能力全面

支持自动定位和识别车牌号码、17位的车辆 VIN 码信息、行驶证主页和副页及驾驶证主页正面的所有字段信息。

算法领先

基于业内领先的深度学习算法、车牌识别、车辆 VIN 码的准确率达98%以上,算法精度高。

鲁棒性强

适应多场景、不同背景,模糊、可容忍透视畸变、光照不均等实际应用中存在的复杂场景。

应用场景

车主信息服务

在汽车保险理赔、二手车交易、车辆租借和年审等场景,帮助用户快速录入车辆相关信息,提高业务人员的办公效率 和服务准确性。

企业档案管理

识别公司印章,方便档案管理和查找。



文档智能

最近更新时间: 2025-05-08 15:17:12

产品介绍

文档智能(Document AI)深度融合 OCR 与多模态大模型,实现高精度识别、智能解析与结构化信息抽取。覆盖货运单证、跨境物流、快递面单、教育作业、保险理赔及国际结算等场景,助力企业自动化升级,提升运营效率与数据准确性。

产品功能

自定义键值

支持自定义创建个性化键值(key),用户通过传入自定义 key,模型自动建立图片中文字的键值对应关系,实现 对任意版式图片的结构化识别。

智能匹配

支持对已上传图片的智能配准,实现对不同版式图片与已发布模板的自动匹配,减少人工分类成本,快速实现图片的 结构化识别。

自定义字段类型

支持自定义创建字段类型,支持针对不同识别区内容类型进行专项优化,如小写金额、日期、纯数字等,可根据需求选择合适的字段类型以提升识别准确率,也可通过穷举可能的输出值范围自定义字段类型,对识别结果进行智能纠正和规范。

产品优势

高精度

支持任意版式的卡证票据识别,各字段精度均处于业界领先水平,制式卡证票据识别精度97%,复杂场景高达 95%,行业领先。

泛化性

基于百亿参数的专属场景的多模态大模型底座,模型泛化性强,全面的文档问答覆盖,强大的算力支撑。

易用性

用户仅需简单的几步配置,即可定制个性化需求模板,快速提取结构化数据,实现高效录入。

模板自定义

版权所有:腾讯云计算(北京)有限责任公司 第14 共18页



支持用户上传样本自定义模板,定义字段规则,可以通过配置只输出自定义字段或者默认字段+自定义字段。

应用场景

交通场景

适用于各种版式不固定、版面复杂、中英文混排的场景,针对进出口物流、货代、进出口贸易的单据,完成高精度的 抽取识别:

- 大宗货运场景, 涉及单据: 收发货磅单、道路运输证、从业资格证。
- 跨境货代场景,涉及单据运单、海外发票、装箱单、提单、进出口报关单、托书、货物保险单、货运到达通知。
- 快递&快运场景,涉及单据: 快递运单、出货单、出库单、签回单。
- 网络运输场景,涉及单据:从业资格证、网约车运输证、网约车驾驶证、网约车行程单识别。

教育场景

针对教育应用场景,我们提供以下专项内容抽取能力:

- 拍照搜题场景; 如: 试卷切题(仅出题目边框)、试题识别。
- 试卷录入场景; 如: 公式识别、试卷切题、试题识别。
- 作业批改场景; 如: 试卷切题、试题识别、中英文手写作文识别。

金融保险

针对金融保险下的保险理赔场景,针对医疗单据、诊断证明、检验报告、保单等单据,完成了高精度的抽取识别;

- 医疗单据类;如:门诊发票(医保报销审核)、住院发票(纸质版)、费用清单、费用结算单。
- 诊断证明类;如:诊断证明、门诊病历、出院记录、住院病案首页。
- 检验报告类; 如: 检验报告、病理报告。
- 保单类;如: 人寿保险单。

其他

适用于各种个人企业证件的识别场景,针对下列典型版式统一、未定制的材料,可完成高精度的抽取识别:

- 政务类个人证件,如:临时身份证、出生证、结婚证、军官证、警察证、残疾人证、导游证等。
- 教育类个人证件,如:学生证、四六级证、教师资格证、毕业证书、学位证、学历认证、办学许可证等。
- 交通类个人证件,如: 道路运输证、从业资格证、网约车运输证、网约车驾驶证、临时车牌、车辆合格证等。
- 企业资质类证照,如: 医疗许可证、食品经营许可证、药品生产许可证、烟草专卖许可证、运输许可证等。
- 其他单据材料,如:车险保单、快递运单、银行回单、商业汇票、仪器仪表盘、芝麻分等。



智能扫码

最近更新时间: 2024-08-27 10:14:22

产品介绍

智能扫码(Intelligent Scan, IS)包含了二维码和条形码识别 API、智能扫码 SDK,为您提供全套扫码相关的服务。其中,智能扫码 SDK 针对一图多码、大图小码等不同场景做了针对性优化。

产品功能

二维码和条形码识别

提供识别二维码和条形码的 API 接口,能够准确识别二维码、条形码、PDF417码、DataMatrix 码。

智能扫码 SDK

提供方便用户集成的扫码 SDK,支持大图小码、一图多码、 模糊码的识别,并且支持自动补光和自动调焦。

产品优势

多码适配

智能扫码提供以下四种码制的识别能力:二维码、条形码、PDF417码、DataMatrix码。

接入全面

智能扫码提供多种接入方式:目前已提供 API 接入、iOS SDK 接入、Android SDK 接入三种接入方式,后续还 会推出小程序 SDK 接入和 Linux 平台 SDK 接入的方式,能够满足客户在不同场景的扫码需求。

效果突出

智能扫码识别准确率高、鲁棒性强,SDK 支持大图小码、一图多码、模糊识别,能够满足多样的实际场景需求。

性能优异

智能扫码基于腾讯海量的服务经验,提供高性能的服务模式,其 FPS 高于50,安装容量包小,同时保证了高内聚、低耦合、接入快速、方便开发。

配套完善

智能扫码 SDK 除了基本扫码功能外,还支持多种扩展服务,如暗光开灯,自动调焦等。

第17 共18页



文本图像增强

最近更新时间: 2024-08-23 10:47:51

产品介绍

文本图像增强(Text Image Enhancement)基于行业领先的计算机视觉技术,面向文件类图片场景提供图像处理服务,提供切边增强、弯曲矫正、阴影去除、摩尔纹去除、图像提亮等多项功能特性,旨在帮助客户更便捷、更快速地获取清晰度更高、阅读性更强的文档类图片,更好地服务客户后续业务流程。

文本图像增强

面向文档类图片提供的图像增强处理能力,包括切边增强、图像矫正、阴影去除、摩尔纹去除等;可以有效优化文档 类的图片质量,提升文字的清晰度。

产品优势

高精度

本产品基于深度学习技术和图像处理技术,可实现高精度的图像处理效果。

完备性

本产品提供切边增强、弯曲矫正、阴影去除、摩尔纹去除、图像提亮等多项功能特性。

应用场景

纸质文件电子化

通过移动端拍照上传文档,用于后续流程审批,或对文件备份存档,减少精准拍摄的工作量。

OCR 图像预处理

对于存在特殊场景的业务,如果图片存在相似的低质量问题,可以先对图像进行增强处理,再请求 OCR 服务,可提升 OCR 识别效果。



基本概念

最近更新时间: 2024-08-23 14:54:13

OCR

OCR(Optical Character Recognition),是光学字符识别的缩写,通过扫描等光学输入方式将各种票据、报刊、书籍、文稿及其它印刷品的文字转化为图像信息,再利用文字识别技术将图像信息转化为可以使用的计算机输入技术。

API

API(Application Programming Interface,应用程序编程接口)是一些预先定义的函数,应用将自身的服务能力封装成 API,并通过 API 开放给用户调用。API 包括基本信息、前后端的请求路径和参数以及请求相关协议。

JSON

JSON(JavaScript Object Notation,JS 对象简谱)是一种轻量级的数据交换格式,任何 JavaScript 语言 支持的类型都可以通过 JSON 来表示,例如字符串、数字、对象、数组等。

QPS

QPS 指每秒请求的并发数,1QPS 即每秒同时支持请求接口1次,50QPS 即每秒同时支持请求接口50次。

SecretKey

SecretId 和 SecretKey 合称为云 API 密钥,是用户访问腾讯云 API 进行身份验证时需要用到的安全凭证。 SecretKey 是用于加密签名字符串和服务器端验证签名字符串的密钥。一个 APPID 可以创建多个云 API 密钥。

SecretId

SecretId 和 SecretKey 合称为云 API 密钥,是用户访问腾讯云 API 进行身份验证时需要用到的安全凭证。 SecretId 用于标识 API 调用者身份。一个 APPID 可以创建多个云 API 密钥。

SDK

SDK(Software Development Kit)是软件工程师为特定的软件包、软件框架、硬件平台、操作系统等建立应用软件时的开发工具的集合。

版权所有:腾讯云计算(北京)有限责任公司 第18 共18页