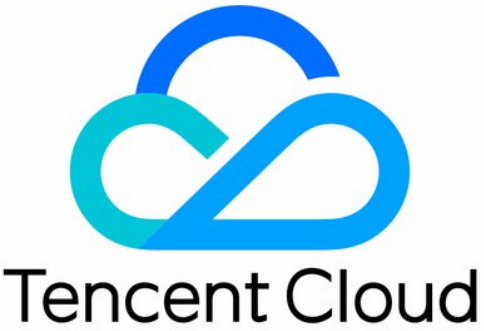
Optical Character Recognition
Product Introduction



Copyright Notice

Trademark Notice

Tencent Cloud

Service Notice

This document provides an overview of the as-is details of Tencent Cloud's products and services in their entirety or part. The descriptions of certain products and services may be subject to adjustments from time to time.

The commercial contract concluded by you and Tencent Cloud will provide the specific types of Tencent Cloud products and services you purchase and the service standards. Unless otherwise agreed upon by both parties, Tencent Cloud does not make any explicit or implied commitments or warranties regarding the content of this document.

Contact Us

We are committed to providing personalized pre-sales consultation and technical after-sale support. Don't hesitate to contact us at 4009100100 or 95716 for any inquiries or concerns.

Contents

Product Introduction
Overview

Last updated: 2025-02-06 16:03:20

Tencent Cloud Optical Character Recognition (OCR) is based on industry-leading deep learning technology, which intelligently recognizes text in images and converts it into editable text. It supports the recognition of printed and handwritten text in multiple scenarios, including general text, card text, invoice documents, and specific scenarios. It also supports invoice verification and provides custom services, effectively replacing manual entry.

General OCR
Card OCR
Invoice OCR
Specific Scenario Recognition
SmartStructure Optical Character Recognition
Text Image Enhancement
Intelligent Scan

OCR Text Recognition Access Guide

Watch video

Feature Experience

OCR feature demo click Free trial.

Use Requirements

Users need to have programming skills and be familiar with programming languages such as Java/Python/PHP/Node.js/C++.
The OCR service requires users to perform secondary development integration and then call the API to recognize text from images or scanned copies into editable text. The results returned after text recognition cannot be directly edited and used. Users need to encode the recognition results into their business system or save them as TXT, Excel, etc.
Users can use API 3.0 Explorer for online calling, signature verification, SDK code generation, and quick retrieval of APIs.

General Character Recognition

Last updated: 2025-02-06 16:03:29

Overview

General Optical Character Recognition (General OCR) is based on industry-leading deep learning technology, providing various services such as general printed text recognition, high-accuracy general printed text recognition, general handwritten text recognition, English recognition, and table recognition. It supports intelligently recognizing text on images and converting it into editable text. It can be used in photo scanning, paper document digitization, e-commerce ad moderation, and many other scenarios to greatly improve information processing efficiency.
Watch video

Product Features

General Print Recognition

Supports text recognition in entire images under multi-scenario and different layouts. It can automatically detect the language type and also supports selecting the language type manually (recommended). In addition to Chinese and English, it supports Japanese, Korean, Spanish, French, German, Portuguese, Vietnamese, Malay, Russian, Italian, Dutch, Swedish, Finnish, Danish, Norwegian, Hungarian, and Thai.

General Print Recognition (High-Precision)

Supports the detection and recognition of text in entire images, returning the text box positions and text content. Compared to the general printed text recognition API, it offers higher accuracy and recall rate, covering a wider range of scenarios. Application scenarios include printed text recognition, online image text recognition, ad image text recognition, street view store sign text recognition, menu text recognition, video title text recognition, and avatar text recognition.

General Handwriting Recognition

Supports the recognition of handwritten Chinese, English, letters, numbers, and common characters in entire images under multi-scenario and different layouts. The recognition capability has been enhanced to handle irregular, messy, and blurry handwriting. It can be applied to scenarios such as handwritten document entry in banks, insurance, finance, and documentation of notes in the education sector.

English Recognition

Supports the detection and recognition of English text in images, returning the text box positions and text content. It supports the recognition of English, letters, numbers, and common characters in multi-scenario and different layouts, covering both printed and handwritten English text. It can be applied to scenarios such as documentation of English notes, recognition of English exam answer sheets, and other similar applications.

Table Recognition (V2)

Supports the detection and recognition of regular tables, wireless tables, and multi-tables in Chinese and English images/PDFs, and supports the recognition of wired tables in Japanese. It returns the text content of each cell, supports the recognition of rotated table images, and allows saving the recognition results in Excel format.

Table Recognition (V3)

Supports the detection and recognition of regular tables, wireless tables, and multi-tables in Chinese and English images/PDFs, returning the text content of each cell, supporting the recognition of rotated table images, and allowing saving the recognition results in Excel format. The recognition effect is better than table recognition, covering a wider range of scenarios. It performs better in difficult table scenarios such as wireless tables and nested tables (wireless tables within wired tables), and can avoid interference from stamps and broken table lines. It is suitable for customers with higher accuracy and recall rate requirements for the API.

Advertisement Text Recognition

Supports the detection and recognition of text in advertisement product images, returning the text box positions and text content. It supports the recognition of Chinese and English, horizontal and vertical text, and text in 90-degree, 180-degree, 270-degree flip, and tilted scenarios. The recall rate and accuracy of text recognition can reach over 96%.

Card Recognition

Last updated: 2025-02-06 16:03:37

Overview

Card Optical Character Recognition (Card OCR), based on industry-leading deep learning technology, provides various services including identity card recognition, bank card recognition, driving license, driving license, business license recognition, passport recognition, Hong Kong (China), Macao (China), and Taiwan (China) travel permit recognition, Hong Kong (China), Macao (China), and Taiwan (China) mainland travel permit recognition, Hong Kong (China), Macao (China), and Taiwan (China) residence permit recognition, Hong Kong (China) identity card, business card recognition, and motor vehicle registration certificate recognition. It supports intelligently recognizing text on images and converting it into structured text, applicable in scenarios such as user registration, bank account opening, transportation, and government services, greatly improving information processing efficiency.

Product Features

ID Card Recognition

Supports recognition of all fields on the front and back of a second-generation Chinese mainland ID card, including name, gender, ethnicity, date of birth, address, identification number, issuing authority, and validity period. It can also crop ID card photos and face photos, and provides alarms for rephotographed, doctored, and photocopied images, border and frame occlusions, temporary ID cards, and invalid validity periods. It can be applied in various ID information validity check scenarios, such as user registration and bank account opening.

ID Card Recognition (Encryption Version)

The Identity Card Recognition (Encrypted Version) API implements encrypted data transfer, effectively preventing the theft and leakage of personal ID information. This API supports recognition of all fields on the front and back of a second-generation Chinese mainland ID card, including name, gender, ethnicity, date of birth, address, ID number, issuing authority, and validity period, with an accuracy rate of over 99%. The product also supports various scaling capabilities to meet different scenarios, such as cropping ID card photos and face photos, and provides nine types of alarms.

Bank Card Recognition

Supports detection and recognition of key fields on the front and back of mainstream Chinese mainland bank cards, including card number, card type, cardholder name, bank information, and valid period. It supports recognition of vertically oriented and irregularly shaped cards, as

well as images rotated at various angles. It can be applied in various bank card information validity check scenarios, such as identity verification in the financial industry and third-party payment card binding.

Driver'S License Recognition

Supports automatic positioning and recognition of all fields on the home page and sub-page of a driving license. Home page fields include license number, name, gender, nationality, address, date of birth, date of first issue, vehicle type, validity period, and issuing authority. Sub-page fields include license number, name, file number, and records. The recognition accuracy of key fields on the driving license exceeds 99%, and it provides alarms for copies, rephotographed, and doctored images. It also supports recognition of the electronic driving license issued by the Traffic Management 12123 App. Fields on the electronic driving license include license number, name, gender, nationality, date of birth, date of first issue, vehicle type, validity start date, validity end date, file number, status, and accumulated points.

Vehicle License Recognition

Supports automatic positioning and recognition of all fields on the home page and sub-page of a driving license, including license plate number, vehicle type, owner, address, nature of use, brand model, vehicle identification code, engine number, registration date, issue date, and issuing authority.

Business License Recognition

Supports fast and accurate recognition of fields on a business license, including unified social credit code (registration number before the three-in-one reform), company name, entity type, legal representative, registered capital, organization form, date of establishment, business term, and business scope. It can be applied in scenarios such as online business verification and online office work, providing convenience for business services.

Passport Recognition (Chinese Mainland Passport)

Supports detection and recognition of multiple fields on the personal information page of passports from the Chinese mainland. Supported fields include English name, Chinese name, Country Code, passport number, birthplace, date of birth, nationality in English, gender in English, validity period, issuing place in English, issuing date, holder's signature, and passport machine-readable zone (MRZ) code.

Passport Recognition (Hong Kong (China), Macao (China), Taiwan (China), and Foreign Passports)

Supports passports from Hong Kong (China), Macao (China), Taiwan (China), and other countries and regions. Recognized fields include passport ID, name, date of birth, gender,

validity period, issuing country, nationality, and country/region code. It also provides cropping of passport portrait photos and alarms for rephotographed and photocopied images.

Hong Kong (China), Macao (China), and Taiwan (China) Permit Recognition

Supports recognition of card-type Mainland Travel Permits for Hong Kong (China), Macao (China), and Taiwan (China), including issuing place, issuing authority, validity period, gender, date of birth, English name, name, and ID number.

Hong Kong, Macao, and Taiwan Travel Permit Recognition

Supports recognition of Mainland Travel Permits for Hong Kong (China), Macao (China), and Taiwan (China), with intelligent recognition and structuring of all fields on the front side, including Chinese name, English name, gender, date of birth, issuing authority, validity period, ID number, issuing place, number of issuances, and document category.

Hong Kong (China), Macao (China), and Taiwan (China) Residence Permit Recognition

Supports detection and recognition of all fields on the front and back of residence permits for Hong Kong (China), Macao (China), and Taiwan (China), including name, gender, date of birth, address, identity card number, issuing authority, validity period, number of issuances, and passport number. It can be applied in scenarios for validity checks of residence permit information, such as bank account opening and user registration.

Hong Kong (China) ID Card Recognition

Supports recognition of key fields on the portrait side of the Hong Kong (China) identity card, including Chinese name, English name, name telex code, date of birth, gender, document symbol, date of first issue, date of latest issue, identity card number, and whether it is a permanent identity card. It also features anti-counterfeiting recognition, face photo cropping, and other extended functions.

Business Card Recognition

Supports automatic positioning and recognition of various fields on business cards, including name, phone, mobile number, mailbox, company, department, position, website address, address, QQ, WeChat, and MSN. It can be applied in social platform business applications, CRM software, and other scenarios.

Motor Vehicle Registration Certificate Recognition

Supports structured recognition of key fields on Mainland China motor vehicle registration certificates, including motor vehicle owner, identity proof name, number, model, vehicle identification number, engine number, and manufacturer name.

Smart Card Categorization

Supports intelligent classification of identity cards, passports, business cards, bank cards, driving licenses, driving permits, Mainland Travel Permits for Hong Kong (China), Macao (China), and Taiwan (China), household registers, Mainland Travel Permits for Hong Kong (China), Macao (China), and Taiwan (China), residence permits for Hong Kong (China), Macao (China), and Taiwan (China), real estate certificates, and business licenses.

Invoice Optical Character Recognition

Last updated: 2025-02-06 16:03:48

Introduction

Invoice Optical Character Recognition (Invoice OCR) is based on industry-leading deep learning technology, providing services such as VAT invoice recognition, train ticket recognition, taxi receipt recognition, flight itinerary recognition, and tax payment certificate recognition. It supports intelligent identification of text content in images into structured text, applicable to various scenarios like corporate invoice reimbursement, financial invoice recognition, and express bill entry, significantly improving information processing efficiency.

General Invoice Recognition (Advanced Edition)

For reimbursable invoices, it supports PDF multiple pages, single page with single or multiple invoices, and mixed-type invoice recognition. Supported invoice types include: VAT invoices (special, general, roll, blockchain, toll), fully electronic invoices (special, general), non-tax invoices (general receipts, unified payment slips), fixed amount invoices, general machine-printed invoices, vehicle purchase invoices (motor vehicle sales invoices, second-hand car invoices), train tickets, taxi receipts, flight itineraries, bus tickets, ship tickets, and toll invoices, covering 14 major types of standard reimbursable invoices. It also supports intelligent identification of other invoice types not listed above.

Value-Added Tax Invoice Verification (New Edition)

Supports the accuracy verification of VAT invoices. By entering key fields of the VAT invoice, you can provide the required verification information. The API returns the actual invoice-related information, including invoice code, invoice number, invoice date, currency, consumption type, buyer's name, buyer's taxpayer identification number, seller's name, seller's taxpayer identification number, and other common fields. It supports verification of various invoice types, including special VAT invoices, general VAT invoices (including electronic general invoices, roll invoices, toll invoices), fully electronic invoices, unified motor vehicle sales invoices, special VAT invoices for the transportation industry, unified second-hand car sales invoices, and general machine-printed electronic invoices (Guangdong and Zhejiang).

VAT Invoice Recognition

Supports the structured recognition of fields for VAT invoices (special, general) and fully electronic invoices (special, general), including invoice code, invoice number, printed invoice

code, printed invoice number, invoice date, total currency, check code, tax rate, total tax amount, total amount including tax, buyer's identification number, verification, seller's identification number, issuer, password area 1, password area 2, password area 3, password area 4, invoice name, buyer's name, seller's name, service name, remark, specification, quantity, unit price, currency, tax amount, payee, and other fields.

OFD Invoice Recognition

Supports the structured recognition of VAT invoices (special, general) and fully electronic invoices (special, general) in OFD format, returning invoice code, invoice number, invoice date, verification code, machine number, password area, buyer and seller information including name, taxpayer identification number, address and phone number, bank of account and account, as well as total amount including tax, issuer, payee, verifier, tax amount, and amount excluding tax.

Bank Receipt Recognition

This API supports the recognition of all fields of bank receipts, including payer's bank, recipient's bank, payer's account number, recipient account number, receipt type, receipt number, currency type, serial number, voucher number, transaction institution, transaction amount, transaction fee, date, and other field information.

Tax Payment Certificate Recognition

Supports the structured recognition of general tax payment certificates for businesses and personal tax payment certificates, including key fields such as tax number, taxpayer identification number, taxpayer name, total amount in words, total amount in figures, issuance date, tax authority, and issuer.

Medical Invoice Recognition

Medical invoice recognition currently supports the recognition of national unified outpatient invoices, national unified inpatient invoices, and some local outpatient and inpatient invoices.

Waybill Recognition

Supports the recognition of mainstream electronic waybills on the market, including fields such as recipient and sender's name, phone, address, and waybill number.

Online Ride-Hailing Itinerary Recognition

Supports the recognition of key fields in online ride-hailing itineraries, including trip start and end dates, boarding time, starting point, end, mileage, and currency.

Insurance Document Recognition

Supports the recognition of fields in common insurance claim documents, including the first page of the medical record, expense list, settlement sheet, medical invoice's name, sex, date of birth, discharge diagnosis, disease code, admission condition, etc.

Advantages

Comprehensive Capability

Supports the recognition of all types of reimbursable documents, such as VAT invoices, quota invoices, car purchase invoices, train tickets, taxi invoices, flight itineraries, electronic waybills, etc., covering many fields and providing comprehensive information.

High Accuracy

The accuracy of each field is industry-leading. For example, the identification accuracy of VAT invoices for fields such as invoice code, total amount including tax, total amount, buyer's identification number, and invoice date exceeds 95%.

Support Verification

Supports the verification of various types of invoices, including VAT (special, general), fully electronic invoices (special, general), unified invoices for motor vehicle sales, and unified invoices for second-hand car sales.

Use Cases

Corporate Expense Control and Reimbursement Management

Suitable for corporate expense control and reimbursement management systems, enabling the rapid acquisition and verification of invoice information, achieving electronic archiving of financial invoices, and improving review efficiency for expense control and reimbursement.

Finance, Taxation, and Agency Accounting

Suitable for agency accounting firms, enabling the acquisition and verification of invoice information, achieving electronic management of financial invoices, reducing manual operation and the cost of lost invoices, and improving agency accounting management efficiency.

Specific Scenario Recognition

Last updated: 2025-02-06 16:03:57

Introduction

Specific Scenario Optical Character Recognition (Specific Scenario OCR) based on industry-leading deep learning technology provides specific scenario recognition services such as license plate recognition, vehicle VIN code recognition, and seal recognition. It supports intelligently recognizing text on images into structured text, applicable in various industry scenarios like car insurance claims, vehicle rental, and enterprise solution management, greatly improving information processing efficiency.

License Plate Recognition

Supports automatically positioning and recognizing motor vehicle license plates in the Chinese mainland, returning region codes, license plate numbers, and license plate color information.

Vehicle VIN Code Recognition

Supports detection and recognition of vehicle identification numbers (VIN) in images.

Seal Recognition

Supports various seals, including invoice seals and financial seals, suitable for documents and invoices.

Advantages

Comprehensive Capability

Supports automatically positioning and recognizing license plate numbers, 17-digit vehicle VIN code information, all field information on the main and secondary pages of the driving license, and the front page of the driving license.

Leading Algorithms

Based on industry-leading deep learning algorithms, license plate recognition and vehicle VIN code accuracy exceed 98%, with high algorithm precision.

Strong Robustness

Adapts to multi-scenario and different backgrounds, tolerates perspective distortion, uneven lighting, and other complex scenarios in practical applications.

Use Cases

Car Owner Information Service

In car insurance claims, second-hand car trade, vehicle rental, and annual inspection scenarios, helps users quickly input vehicle information, improving office efficiency and service accuracy.

Enterprise Profile Management

Recognizes company seals, facilitating archive management and search.

SmartStructure Optical Character Recognition

Last updated：2025-02-06 16:04:08

Introduction

Smart OCR (SmartStructure Optical Character Recognition) is based on industry-leading deep learning and image detection technologies, capable of structured text extraction and recognition for any fixed format cards, certificates, and invoices. The product intelligently establishes key-value relationships through pre-learning and allows customers to set key-value pairs with custom templates, enhancing the efficiency of information data extraction and entry. It can be used in government services, special invoice verification, industry form customization, and other scenarios.

Intelligent Structured Recognition V2

Supports recognition and structuring of information in general formats, including various non-standard format business licenses, personal certificates, invoice forms, contracts, and other structured scenarios' field information. It also supports synchronous return of full text OCR results to assist in business information extraction. No configuration is required, making it flexible and efficient. Suitable for various structured information entry scenarios.

Advantages

Comprehensive Capability

Supports recognition of any fixed format cards and certificates, with field accuracy at industry-leading levels, achieving an identification accuracy of over 90%.

High Generalization

Supports structured data extraction for various common fixed formats, such as police certificates, teacher qualification certificates, and road transport certificates, suitable for multi-industry scenarios.

Use Cases

Common Document Recognition

Suitable for various personal and business certificate recognition scenarios, capable of high-precision structured recognition for the following typical unified, non-customized materials:

1. Government affairs personal documents, such as: temporary identity card, birth certificate, marriage certificate, military officer ID card, police certificate, disability certificate, tour guide certificate, etc.
2. Educational personal documents, such as: student ID, CET-4 and CET-6 certificates, teacher qualification certificate, diploma, degree certificate, Ministry of Education academic certification, school operating license, etc.
3. Transportation personal documents, such as: road transport certificate, online car-hailing transport certificate, online car-hailing driver's license, temporary license plate, vehicle conformity certificate, etc.
4. Enterprise qualification licenses, such as: medical license, food business license, pharmaceutical production license, tobacco monopoly license, transport license, etc.

Ticket and Document Recognition

Suitable for various invoice document recognition scenarios, capable of high-precision structured recognition for the following typical unified, non-customized materials:

1. Financial bills and documents, such as: bank paper receipt, bank electronic receipt, large remittance voucher, bookkeeping agency license, customer account notification, etc.
2. Financial reimbursement documents, such as: product receipts, payment screenshot vouchers, Didi itinerary, hotel receipts, customs payment slips, product order pages, etc.
3. Transportation bills and documents, such as: customs declaration, waybill, English invoice, English receipt, shipping and unloading list, purchase order, delivery order, etc.

Industry Material Recognition

Suitable for recognition scenarios of various special materials in sub-industry fields, capable of high-precision structured recognition of short texts and tables for the following typical unified, non-customized materials:

1. Industrial industry materials, such as: product tags, equipment nameplates, price tags, etc.
2. Service industry materials, such as: contracts, bids, vehicle insurance documents, lottery, etc.
3. Healthcare materials, such as: drug instructions, discharge records, hospital test reports, etc.

Intelligent Scan

Last updated: 2025-02-06 16:04:16

Overview

Intelligent Scan (IS) includes QR code and barcode recognition APIs, Intelligent Scan SDK, providing a full set of scan code services. The Intelligent Scan SDK is optimized for scenarios like scan multiple codes in one image and small code in large image.

Product Features

QR Code and Barcode Recognition

Provides APIs for recognizing QR codes and barcodes, capable of accurately identifying QR codes, barcodes, PDF417 codes, and DataMatrix codes.

Intelligent Scan SDK

Provides a scan code SDK for easy user integration, supporting the recognition of small code in large image, scan multiple codes in one image, and fuzzy code, as well as automatic lighting and autofocus.

Advantages

Multi-Code Adaptability

Intelligent Scan offers recognition capabilities for four types of codes: QR codes, barcodes, PDF417 codes, and DataMatrix codes.

Comprehensive Integration

Intelligent Scan provides multiple access methods: currently offering API access, iOS SDK access, and Android SDK access. In the future, Mini Program SDK access and Linux platform SDK access will also be available to meet customers' scan code needs in different scenarios.

Outstanding Effect

Intelligent Scan has high accuracy and strong robustness. The SDK supports small code in large image, scan multiple codes in one image, and fuzzy recognition, meeting diverse practical scenario demands.

High Performance

Intelligent Scan, based on Tencent's extensive service experience, offers a high-performance service model with FPS over 50 and a small installation package, ensuring high-cohesion, low-coupling, quick integration, and easy development.

Well-Equipped

In addition to basic scan code features, the Intelligent Scan SDK also supports various extended services, such as low-light lighting and autofocus.

Text Image Enhancement

Last updated：2025-02-06 16:04:29

Introduction

Text Image Enhancement, based on industry-leading computer vision technology, provides Image Process services for document images. It offers features such as edge enhancement, distortion correction, shadow removal, moiré removal, and image brightening. It aims to help customers obtain clearer and more readable document images more conveniently and quickly, better serving subsequent business processes.

Text Image Enhancement

The image enhancement capabilities for document images include edge enhancement, image correction, shadow removal, and moiré removal. These can effectively optimize the image quality of documents and improve text clarity.

Advantages

High Precision

This product, based on deep learning and Image Process technology, can achieve high-precision image processing effects.

Completeness

This product offers features such as edge enhancement, distortion correction, shadow removal, moiré removal, and image brightening.

Use Cases

Paper File Electronic

By taking photos and uploading documents via the mobile terminal, it can be used for subsequent Process approval or file backup and archive, reducing the workload of precise shooting.

OCR Image Preprocessing

For businesses with special scenarios, if images have similar low-quality issues, you can perform Enhanced Processing on the image before requesting OCR service to improve OCR recognition effect.

Basic Concepts

Last updated: 2025-02-06 16:04:39

OCR

Optical character recognition (OCR) is a technology that converts the text in various documents, newspapers, books, manuscripts, and other printed materials into image information through optical input methods such as scanning and then converts such information into computer-readable input through text recognition technology.

API

API (Application Programming Interface) is a set of predefined functions. The application encapsulates its own service capabilities into API and makes them accessible to users. API includes basic information, front-end and back-end request paths and parameters, and related protocols.

JSON

JSON (JavaScript Object Notation) is a lightweight data exchange format. Any type supported by the JavaScript language can be represented by JSON, such as strings, numbers, objects, arrays, etc.

QPS

QPS (Queries per Second) measures the requests processed concurrently per second. 1 QPS means that the API processed 1 request per second; 50 QPS means that the API processed 50 requests per second.

SecretKey

SecretId and SecretKey collectively refer to the TencentCloud API key, which is the security credential required for authentication when you access TencentCloud API. SecretKey is used to encrypt signature strings and verify them on the server. You can create multiple TencentCloud API keys for one APPID.

SecretId

SecretId and SecretKey collectively refer to the TencentCloud API key, which is the security credential required for authentication when you access TencentCloud API. SecretId is used to identify the API caller. You can create multiple TencentCloud API keys for one APPID.

SDK

A software development kit (SDK) is a collection of development tools used by software engineers to create application software for specific software packages, software frameworks, hardware platforms, and operating systems.