

ALIBABA CLOUD

阿里云

智能双录质检
产品简介

文档版本：20220527

 阿里云

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
 危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 危险 重置操作将丢失用户配置数据。
 警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告 重启操作将导致业务中断，恢复业务时间约十分钟。
 注意	用于警示信息、补充说明等，是用户必须了解的内容。	 注意 权重设置为0，该服务器不会再接受新请求。
 说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击设置> 网络> 设置网络类型。
粗体	表示按键、菜单、页面名称等UI元素。	在结果确认页面，单击确定。
Courier字体	命令或代码。	执行 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid</code> <i>Instance_ID</i>
[] 或者 [a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ } 或者 {a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

目录

1.什么是智能双录质检	05
2.计费说明	07
3.功能概览	08

1.什么是智能双录质检

“双录”是指在金融行业对产品完整的销售过程进行的录音和录像。近些年随着国家对金融行业监管力度的加强，各大金融机构积极响应国家监管制度，建设了内部的双录平台，并已经与业务系统紧密集成。

产品介绍

智能双录质检（Intelligent Dual-Recording System, 简称 IDRS）是蚂蚁集团移动开发平台（mPaaS）团队与阿里巴巴达摩院基于实时音视频以及 AI 识别服务，共同研制的一款智能化的音视频内容录制、检测及审核产品。该产品通过 AI 识别算法，对视频中的各种关键动作、证件文稿、以及销售过程中的关键角色进行实时识别和追踪，支持对视频全过程实时检测和录制。基于 mPaaS 移动端技术，优化端上检测服务，该产品可以在终端视频录制过程中，实时检测，实时提示，实时纠正，提升检测过程的友好度和一次性通过率。

除了智能化的检测能力，该产品在云端还提供了配套的双录流程编排、双录流水查询以及双录视频托管等端到端的服务，帮助用户智能化地建立、使用、管理和分析整个音视频双录业务。

产品优势

基于达摩院 AI 能力

阿里巴巴达摩院团队人工智能（Artificial Intelligence，简称 AI）实验室在语音识别、图像识别、视觉理解、语言理解等方面开展大量研究，并沉淀出 AI 相关的大量技术成果。智能双录质检产品基于达摩院 AI 技术，能够精准识别语音、图像等，并处于行业领先地位。

集成蚂蚁实时音视频通话能力

智能双录质检中的实时音视频通话能力源自蚂蚁集团 mPaaS 团队研发的音视频通话服务，该服务基于支付宝产品，并经多年技术沉淀而来。该服务具有高安全性、弱网环境下视频质量自适应、以及高可靠、低时延的优势。

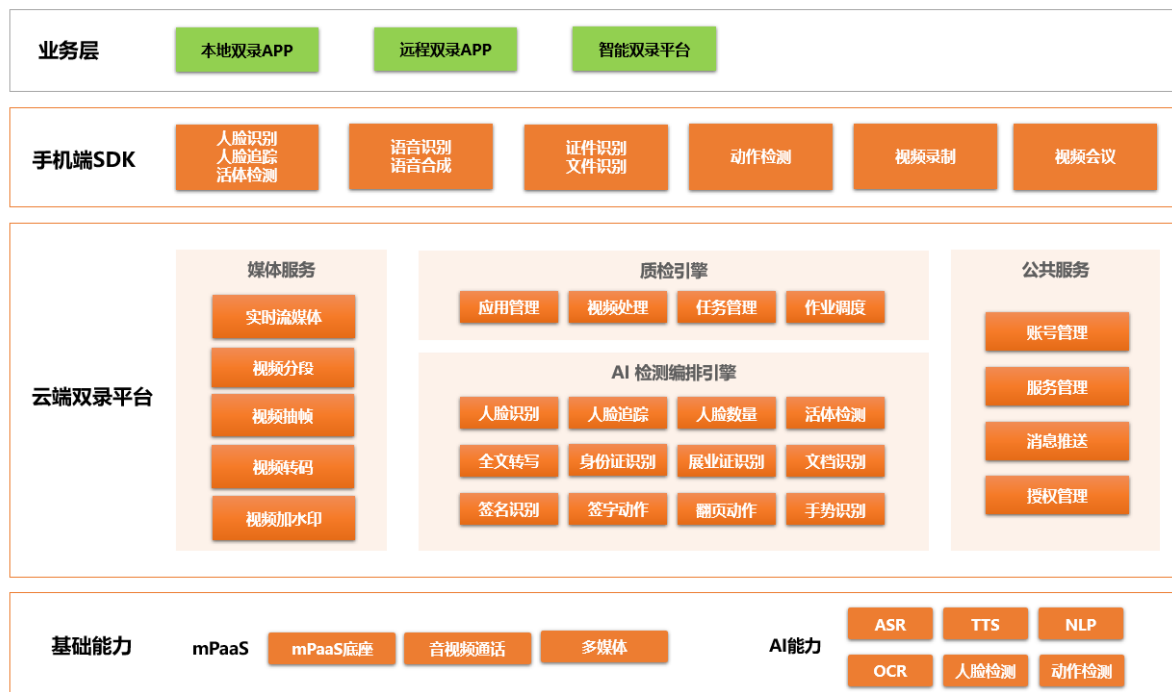
金融合规，数据保密

智能双录质检部署在阿里云的金融云中，具有稳定可靠、金融合规以及数据保密的优势。相比普通公有云，金融云是专门服务于银行、证券、保险、基金等金融机构的行业云，采用独立的机房集群提供满足一行两会监管要求的云产品，并为金融客户提供更加专业的服务，如等保四级。金融云符合人民银行云计算金融应用规范，IDC A 级机房资质认证，金融监管风险评估调查，金融行业监管报告提交，金融行业 IT 合规服务等。

产品架构

智能双录质检将易于使用的 AI 检测能力封装为 SDK 提供给企业或软件开发商，并由企业或软件开发商根据业务需求完成业务双录应用的开发。

智能双录质检的产品架构如下图所示。



智能双录检测架构说明如下：

- **基础能力**：包括存储和 AI 模型引擎。主要提供视频文件的存储和读取，以及核心 AI 模型的运行环境。
- **云端双录平台**：包括媒体服务、质检引擎、公共管理服务、AI 检测编排引擎等。对外提供针对双录质检场景流程的编辑，发布等功能。
- **手机端 SDK**：包括手机 SDK 和 OpenAPI，提供端上的检测能力和调用能力。
- **业务层**：在智能双录质检产品的基础上，企业或软件开发商可以根据具体业务需求，开发双录质检的业务流程及应用，并实现与保险公司内部系统和外部系统的集成。

使用场景

在金融行业的产品销售过程中，通常分为面对面音视频业务、远程音视频以及事后审计业务。

本地双录

针对线下面对面销售的业务的场景，本产品将人脸检测、语音识别、动作检测等所有检测项涉及的能力，封装在统一的 SDK 中，客户端可随时调用相应检测能力对本地销售过程进行实时质检，并支持将录音录像，检测过程保存及上传到云端。同时产品提供了双录的流程配置服务，可按业务流程需求自定义检测环节的顺序与内容，服务将按照已配置流程自动调用相应检测项。

远程双录

针对异地/线上销售的业务场景，本产品提供了音视频通话服务（Real-Time Communication，简称 RTC）及检测能力。支持销售方与购买方通过移动客户端接通视频通话，在视频通话中实时调用端上的各项检测能力进行检测，并将销售过程录音录像，保存及上传到云端。

云端质检

针对销售事后审核的场景，本产品提供了云端视频检测服务，将事先录制好的双录视频，批量上传到云端，并调用智能双录质检的云检测服务，对所上传的视频进行全面的检测。云端质检支持对双录视频的检测流程进行自定义，从而提高对双录音视频事后检测的效率。

2. 计费说明

智能双录质检提供三种模式的检测服务，分别是本地双录质检、远程双录质检和云端双录质检。这三种检测服务均采用后付费方式进行计费，本文分别对三种服务的计费规则进行说明。

本地双录质检

当您使用智能双录质检提供的 SDK（Android 或 iOS），在手机端对本地视频录制并进行实时质检时，本地双录质检的收费规则如下。

服务	价格	说明
本地检测服务	0.08 元/分钟	按照每台设备调用质检服务 SDK 的时长进行计费。

远程双录质检

当您使用智能双录质检提供的 SDK（Android 或 iOS），在手机端对远程多端视频录制进行实时质检时，远程双录质检的收费规则如下。

服务	价格	说明
远程双录（720P）	0.08 元/分钟	包含实时通话和端上质检能力，并按照每路视频流的时长进行计费。
云端单流录制服务（720P）	0.06 元/分钟	将各端实时通话的视频流分别录制为单独的视频文件（720P），并按照每路视频流的录制时长分别进行计费。
云端混流录制服务（720P）	0.12 元/分钟	将各端实时通话的视频流合成录制为一个视频文件（720P），并按照录制视频的时长进行计费。

云端双录质检

当您使用智能双录质检提供的云端检测服务，将已录制完成的视频上传到云端进行质检时，云端双录质检的计费规则如下。

服务	价格	说明
云端检测服务	0.12 元/分钟	按照云端质检的检测视频时长进行计费。

 注意

视频需要按照打点规范进行录制，具体技术说明请通过 [提交工单](#) 的方式获取。

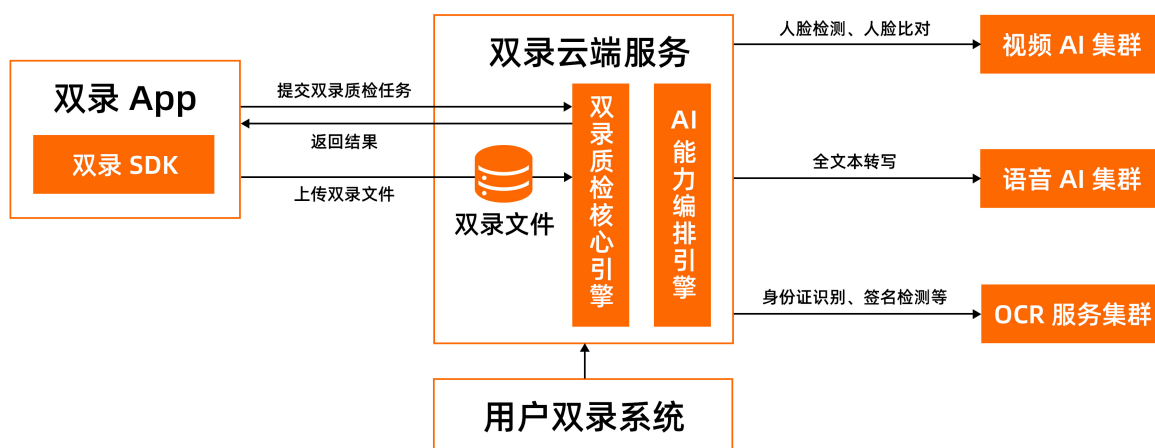
3.功能概览

针对保险、银行等金融领域的双录（录音、录像）场景合规性要求，智能双录质检利用 AI 技术实现对销售过程中各类关键动作、语音、证件的检测，以及关键角色的面部追踪，实现对销售过程的智能化监控。其中，智能双录质检提供的终端检测能力，可以让绝大部分销售合规检测在现场完成，实时发现问题、实时提示、实时纠正，极大地提升了销售的友好度和一次性通过率。

智能双录质检将 AI 检测能力封装成 SDK，并将 SDK 提供给企业或开发者，并由企业或开发者根据业务具体需求开发终端 App，并使之实现双录质检业务。

业务流程

终端 App 中集成了双录 SDK 后，与双录云端服务的业务调用流程如下：



本地双录质检

针对线下面对面展业的双录场景，本地双录质检功能为该类型的展业双录活动提供现场交互式检测。

本地双录 SDK 功能

本地双录 SDK 在录制视频流时，根据预设的检测规则，调用相应的检测模型，并根据检测模型返回的结果，统一对视频进行处理；并将处理后的视频及检测结果返回给双录 App。

本地双录质检特点

本地双录质检功能具有如下特点：

- 多角色同框面部检测和跟踪：对现场多角色进行实时面部识别、检测和跟踪。
- 视频检测防欺诈：识别视频中的欺诈行为。
- 图片信息识别：识别身份证号码、姓名、有效期等信息。
- 用户回答检测：离线识别用户的回答内容。
- 离线 AI 能力：支持在离线无网状态下对本地双录进行交互式检测，扩展展业场景。

本地双录 SDK 检测能力

本地双录 SDK 提供如下检测能力：

- 人脸检测和信息采集

检测和采集视频中的所有人脸及人脸信息。

- 支持 4 张以上的人脸信息采集。
- 从输入视频流到返回视频，SDK 的处理时间不超过 100ms，且检测耗时控制在 50ms 以内。
- 支持动态设置检测频率，比如每秒钟 5 次。
- 人脸匹配和追踪

检测视频中的人脸信息与预设的人脸信息是否匹配，并持续追踪该人脸。

 - 支持检测和追踪四组以上的预置人脸（人脸图像+角色名称）信息。
 - 支持设置人脸匹配相识度。
 - 支持在视频中把匹配到的人脸进行标记。
 - 提示人脸上线消息。当第一次在视频中检测到预设的人脸图像信息时，返回人脸上线消息。
 - 提示人脸离线消息。如果在指定的时间间隔内检测不到已标记的人脸，返回人脸离线消息。
 - 支持动态设置检测频率，比如每秒钟 5 次。SDK 根据检测频率调用模型 SDK 进行检测并将结果返回给 App
 - 从输入视频流到返回视频，SDK 的处理时间不超过 100ms，且检测耗时控制在 50ms 以内。
- 特定词语识别

检测语音流中的特定激活词，并返回该激活词，检测耗时控制在 50ms 以内。

 - 预置识别 6 个激活词。
 - 支持扩展、修改和删除激活词。
- 活体检测

检测视频中的人脸是不是活体，以防用照片或视频等进行仿冒，检测耗时控制在 50ms 以内。

检测完成后，返回检测结果和结果置信度。
- 翻屏动作识别

检测视频中是否存在在手机屏幕上滑动翻屏的动作，检测耗时控制在 50ms 以内。
- 签名动作识别

检测视频中是否存在在手机或 iPad 中签字的动作，检测耗时控制在 50ms 以内。

远程双录质检

针对异地线上展业的双录场景，远程双录质检功能为该类型的展业双录活动提供实时交互式检测，帮助用户以远程视频的方式进行展业活动。

远程双录质检特点

远程双录质检功能具有如下特点：

- 多角色分离：支持如保险代理人、投保人等不同角色的人员进行远程展业。
- 有网即可展业：只要有网络即可远程展业。
- 针对角色的实时交互检测：按照展业角色对会议各端分别设置检测项目。
- 检测结果多端同步：各端检测结果实时同步，协同交互控制和流程驱动。
- 云端视频录制：云端录制各端视频和合流视频，无需本地录制和上传。

- 远程屏幕共享：支持更复杂的保险代理业务操作和协同。

远程双录 SDK 能力

远程双录是将音视频通话能力与双录质检服务相结合，当对话的双方身处异地时，可以音视频通话过程进行实时的通讯，同时进行音频视频的录制，并且结合了实时质检的能力。其主要能力单元分为音视频通话、录制服务与本地智能检测能力。

- 音视频通话

远程双录 SDK 中的音视频通话能力由蚂蚁集团移动开发平台（mPaaS）团队提供，可实现一对一通话及多人会议。通话过程中支持屏幕录制、屏幕共享、截图等功能，同时支持即时文字消息和文件传输。关于音视频通话的更多信息请参见 [音视频通话产品简介](#)。

- 录制服务

远程双录 SDK 检测能力中提供了音视频录制服务，可对远程音视频通话全程录音录像，包括音视频通话、截图、语音播报、交互式白板、屏幕共享、交易操作过程、文字聊天内容、签字画面等。对远程双录中的音视频通话进行录制和存储，是为了后续对业务办理流程进行回溯。

录制中支持设置主、备双录模式，满足高可用的要求，从而保障双录的稳定性。

- 检测能力

远程双录 SDK 中的检测能力与本地双录中的检测能力一致，详细信息请参见 [本地双录 SDK 检测能力](#)。

云端双录质检

针对销售事后审核的场景，智能双录质检在云端中提供了云端视频检测服务，即将录制好的视频，批量上传到云端，由智能双录质检在云端批量完成视频检测。云端双录质检还支持以 API 的形式，对外提供服务。

云端双录质检特点

云端双录质检功能具有如下特点：

- 多角色同框面部检测和跟踪：对现场多角色进行实时面部识别、检测和跟踪。
- 视频检测防欺诈：识别视频中的欺诈行为。
- 离线 AI 能力：支持对离线无网状态下录制的视频进行检测。

云端双录检测能力

- 人脸检测和信息采集

检测和采集视频中的所有人脸及人脸信息。

- 支持 4 张以上的人脸信息采集。
- 从输入视频流到返回视频，SDK 的处理时间不超过 100ms，且检测耗时控制在 50ms 以内。
- 支持动态设置检测频率，比如每秒钟 5 次。

- 人脸匹配和追踪

检测视频中的人脸信息与预设的人脸信息是否匹配，并持续追踪该人脸。

- 支持检测和追踪四组以上的预置人脸（人脸图像+角色名称）信息。
- 支持设定检测频率。
- 支持返回人脸上线和人脸离线的消息。
- 支持通过检测项动态设置判断人脸离线的时间间隔。

- 支持通过检测项动态设置人脸追踪的检测频率，比如每秒 2 次。
- 活体检测

检测视频中的人脸是不是活体，以防用照片或视频等进行仿冒，检测耗时控制在 50ms 以内。

检测完成后，返回检测结果和结果置信度。