

腾讯云智能数智人 定制服务



腾讯云

【 版权声明 】

©2013–2026 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 】



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

定制服务

2D 形象、声音定制 API

形象、声音定制 API 概述

形象定制

接口调用逻辑示意图

获取上传临时令牌

上传素材到腾讯云 COS

定制接口

进度查询接口

效果确认接口

人脸检测任务创建

人脸检测任务状态查询

声音复刻（基础版）

接口调用逻辑示意图

音质检测任务创建接口

音质检测任务状态查询接口

获取上传临时令牌

上传素材到腾讯云 COS

定制接口

进度查询接口

效果确认接口

声音复刻（极速版）

接口调用逻辑示意图

获取声音训练文本

获取上传临时令牌

上传素材到腾讯云 COS

提交音质检测任务

查询音质检测任务进度

定制接口

进度查询接口

结果码说明

接口对接常见问题

2D 端渲染形象定制及下载流程

定制服务

2D 形象、声音定制 API

形象、声音定制 API 概述

最近更新时间：2024-11-13 15:25:52

本文档主要描述 腾讯云智能2D数字人 aPaas 平台的开放接口协议。

一、接口调用环境

正式环境：`https://gw.tv.s.qq.com`

二、接口调用方式

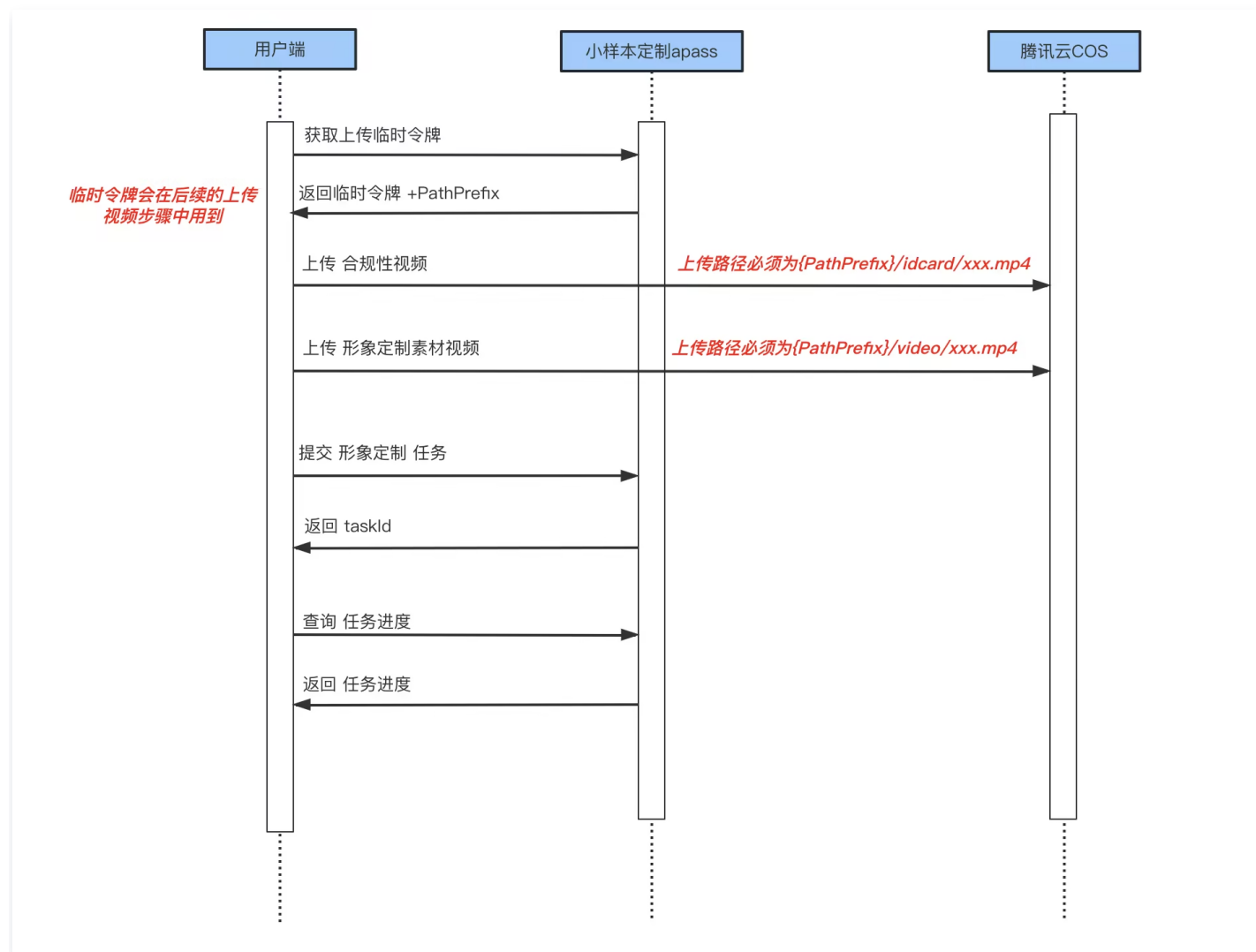
数智人 aPaas 接口调用方式

⚠ 注意：

如果通过腾讯云计费直接采购。appkey 和token请到 [数智人平台资产管理页面](#) 查询。

形象定制 接口调用逻辑示意图

最近更新时间：2025-03-14 16:06:02



1. 初始说明

- 用户端需先获取临时令牌，该令牌将在后续上传视频时使用。

2. 获取临时令牌

- 用户端发起“获取上传临时令牌”请求。
- 服务端返回包含临时令牌 + PathPrefix的响应，用于后续文件上传的鉴权和路径定位。

3. 视频上传阶段

- 用户端依次上传两类视频：
 - 合规性视频（用于内容审核），上传路径必须为 `{PathPrefix}/idcard/xxx.mp4`。
 - 形象定制素材视频（用于生成定制形象），上传路径必须为 `{PathPrefix}/video/xxx.mp4`。

4. 任务提交

- 客户端提交“形象定制”任务。
- 服务端返回 `taskId`（任务唯一标识），供后续查询进度使用。

5. 任务进度查询

- 客户端通过 `taskId` 定期查询任务进度。
- 服务端返回当前任务进度状态（如处理中、已完成等）。

获取上传临时令牌

最近更新时间：2024-11-18 17:24:52

在调用 [定制接口](#) 前，需要将素材文件上传到指定的云端存储。上传时，首先通过该接口获取上传临时令牌和指定的上传路径，再通过 [上传素材到腾讯云 COS](#) 接口上传对应的文件到云端存储。

调用协议

HTTPS + JSON

POST /v2/ivh/assetmanager/customservice/getuploadcredentials

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

请求参数

无

返回参数

参数	类型	必须	说明
Credentials	object	是	临时密钥信息
Credentials.Token	string	是	临时密钥令牌
Credentials.TmpSecretId	string	是	临时证书密钥 ID
Credentials.TmpSecretKey	object	是	临时证书密钥 Key
StartTime	int	是	临时证书有效起始时间，返回 Unix 时间戳，精确到秒
ExpiredTime	int	是	临时证书有效截止时间，返回 Unix 时间戳，精确到秒
PathPrefix	string	是	上传到 COS 上的路径前缀

返回示例

```
{  
  "Header": {
```

```
"RequestID": "c18ae4e49d89e408231c5b3ceea23c",
"SessionID": "gzfa9105ca17310669340273673",
"Code": 0,
"Message": "ok"
},
"Payload": {
  "Credentials": {
    "TmpSecretId": "AKID*****ib",
    "TmpSecretKey": "DVki*****k0=",
    "Token":
"8oz1l*****49esQlG0940LX2tacXiCv4YS2qdBsiQ"
  },
  "ExpiredTime": 1731074134,
  "PathPrefix": "https://cos-test.ap-
guangzhou.myqcloud.com/customer-pipeline/123/4506c29b-2904-45e2-b7d7-
3052fe2493c2/",
  "StartTime": 1731066934
}
}
```



```
"Payload": {
  "Credentials": {
    "TmpSecretId": "AKIDM6p_7lbof_ow5_DTYz...Qzb1MlvtFfjxqbg1FrN9",
    "TmpSecretKey": "9pkrcGy6duNWhj4/...J87PGk=",
    "Token": "fGSqNKc3Q90WcLEz5jn2pdHo3cGEuPKa23738bd8cb693d875f4e2ead9f6ff446w9B4mmLxcks10FXXJAW-jtudiJnfdPht0RvkEC3PshPrv2nrXPLjwTst8A-mGxVxkpM-YIE_RpeuQG2nIKF9izYreejnrzx6oCuCL61AXy39v6CdCHXoWkdWkZyc1LWio2QVMrshS0uKr26H770MmDvjH9Evk19I9V8K4KR7mJaE3GDisk5tJpAqFtNd5tSRGSPsIErdHuf2jUef9WvnWuV1yysSpJy7qr0KqQh7GwzWkv4j2jtr1B6fbNzYfTxPHvijLwDC50Dg613pgiZqwtu7GtoUxdFOXWQny1aHiiz1NnRR25ELot_o7tAsdL6Ae7uj-31bkyQbU88e5JHSRvS4p0QVe58LUNw5WhpTeJlXh-LRsPmc2z29MmXtSvWkgx2Y-KiQBYHsF81paEMuHK4eU60303xdjd5_h6N4C0u6QFet3PUJY7Kgt1t7L4K0670QJERrikEgBYppyZYqvVHaD4YZO-TABrhymYPUVKdbmzW4Idim8tZUI6Fc93FGyYRAFfql3zux4IX-NrieAGMRifks1q1XfRfHfFo2iQeaLyUubddGoyVvK6y5eIj1kou5mM0drPBpPzYoZJWalHT12mbm-hDKHvcgBe-bMe7qs7ze6NiyqwrECDCrVmdTA8j0EVKPGuE94xAnMyytYHnhmQvC8o77hBuQvgeCPiX-R_xAUD-6ydv5Gi_8mLcd07h35Nvk0TrXyNuMTmqjTEbo1sVC74jP68pQMzNy--UU1Z9C_tI6Y938xwf5xPB70MzZMTP7IEnMPqsDZgGwWg1R-U-JbMzq-HkDBU...B0d6Qf3XN2PCL1q1gkJuS8jZBLQ4K4c5XVvfbEpSw4TcLn500yqp6ks7cEXBZeAhAowRj_APBDNTI72-XnlydgBbcjY6G3xXuXDRbBGUs8QrCi_3mMYNMyNq4fLbAnR1sdAnjFKV7UNQ0FMsEFWpKyLD4cs4bEbhbatMt-lu37oVXD4EUcnNyR6Yz1jNws27n8gAL0I9FDEb-aHICXaYyG5cEgELghf16gJsomjwZa0FpXtI4RMPBP_6D0f6RGE0IF_0r5Awh_9_QsGwIXkUW18aeVw7C0fexQ8Cf0vWu-zk0q2F4x9CpQwUdsN18FIYC3I3sYTViS5gIh2P_-L0dkzwjplz5-yRglbdjH5ba1jXm1GqhipeohMsYxQSHDfgkAm1WFwqpp4vKt-0S-1kvY94BqTRkm8M2_0voRNZIDx8n3DZJiI657HDL33d4cGncSv0Q_R49SRiNK471Cne-NDoqRmCBvEXY-mgDS7Fht2F5vKxueYvX0iRnmP4_d75qDpYn1_PNYCoNj4ViYSICPXJUNeofRTJc2miwSENL7vH0H0kM02hNikB46LhLPZPtK51sd4g1JozesLRcbD0sAZyJd_5e1pE7SaulMdm7FMkQq3th0CwacrIy062DncwEQBBM-00uwr7CpBdhakLAXn-RAO-C12ogMw-Wx1BQ57ixN7yjE28ukSzG8QjC14_QUQsG2Y8FKQHVY5hZp1AGhjpBCN_7XzH-WAPGfzGXJW1DT-zkHJYRPjMbB19zx2WcIfCyChtbz7XReQCN6BmWmXvtqApjKAdr9ZQctTK0aNdvm53zz1jxdoHttAkJL9YCPp0iN_2YjHI5t4e7PMc_Itsxg2RgDy83EL8Iz0lqs57WtRErFQspFgtAWr3dmjgTYtHehPIP_K42k5WA2t6g3031z_3_zaGMTzngv4u0PTvYz-QpFqAQMDQPaRL132QTKarriImA8Ff8d10DhjuvBDPzIwzA6085tXTJMHBk1DZwDpw3KLJUiqAgfWcNWh0JmrdmJo_AvYz9yUaqNqH0GAIEUtY1zsh3emcD6vb1V2kZCc3jGjzrTsg"
  }
},
"ExpiredTime": 1736241459,
"PathPrefix": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipeline/352/59c6f6cd-8f92-45f9-8352-64ca8fb92eaa/",
"StartTime": 1736234259
```

- bucket名称: 解析出来的值为 virtualhuman-cos-prod-1251316161。
- region名称: 解析出来的值为 accelerate。
- key(bucket 下的存储路径): 解析出来的值为 customer-pipeline/352/59c6f6cd-8f92-45f9-8352-64ca8fb92eaa/。

注意:

解析出来的值为 key 的前半部分，还需要根据上传者素材类型来拼接后半部分，如上传的是 pdf 授权书素材，则需要拼接上 idcard/xxx.pdf，如下图所示，具体后半部分值的定义请参考 [定制接口](#) 中的相关参数说明。

腾讯云智能数智人	参数名称	数据类型	是否必填	说明	备注
<ul style="list-style-type: none"> 购买指南 数智人形象介绍 数智人平台操作指引 服务端 API 接入 <ul style="list-style-type: none"> 数智人 aPaas 接口调用方式 2D真人形象、声音定制 API <ul style="list-style-type: none"> 概述 形象定制 <ul style="list-style-type: none"> 接口调用逻辑示意图 获取上传临时令牌 上传素材到腾讯云 COS 定制接口 进度查询接口 效果确认接口 人脸检测任务创建 人脸检测任务状态查询 声音复刻 (基础版) <ul style="list-style-type: none"> 接口调用逻辑示意图 音质检测任务创建接口 音质检测任务状态查询接口 获取上传临时令牌 上传素材到腾讯云 COS 定制接口 进度查询接口 效果确认接口 声音复刻 (极速版) 	IdentityWrittenCosUrl	string	否	pdf 格式授权书 URL，要求： 1. url 地址为通过 上传者素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名 <code>customer-pipeline/(数字)/(uuid)/idcard/b.pdf</code> 2. 该格式主要适用于“ 书面版授权书 ”，以完整清晰的扫描件形式提交 备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与 IdentityCosUrl 二选一或二者都填	本项目目录： 调用协议 请求参数 返回参数 请求示例 返回示例
	MaterialCosUrl	string	否	形象定制素材 url 地址要求： 1. url 地址为通过 上传者素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 video 路径，例如 <code>customer-pipeline/(数字)/(uuid)/video/c.mp4</code> 2. 视频大小：不超过5GB，4K不超过10G 3. 视频时长：专属口型版本定制时长为2-10分钟；通用口型版本定制时长为1-10分钟；高精版本定制时长为2-10分钟 4. 视频分辨率：1080P、4K (3840*2160)，如定制高精版，则必须为4K 5. 视频长宽比：16:9 (9:16) 6. 视频帧率：不低于25fps，不超过60fps 7. 视频格式：mp4, mov 备注：形象定制时 video 路径必传 声音定制素材 url 地址要求： 1. url 地址为通过 上传者素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如 <code>customer-pipeline/(数字)/(uuid)/audio/c.zip</code> 2. 压缩包格式：.zip 格式的压缩包，单个 zip 包用于定制1个音色；压缩时不要新建文件夹，直接选中所有 wav 压缩 3. 对于单个 zip 包里的音频注意事项： ①音频数量：每个 zip 包里可以包含1条或多条 wav 格式音频，总量不超过10条 ②音频大小：每个 zip 包里的音频大小总和不超过1G ③音频格式：每条音频需为 wav 格式，其他格式的音频请先转为 wav 格式后再再压缩成 zip 包 ④音频采样率：采样率大于等于24k即可，建议24k、36k	

定制接口

最近更新时间：2025-12-19 14:24:31

使用该接口提交定制的请求，通过 [进度查询接口](#) 查询定制的各个阶段以及各阶段的相关信息。

调用协议

HTTPS + JSON

POST /v2/ivh/assetmanager/customservice/make

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

请求参数

参数	类型	必须	说明
AnchorName	string	是	<p>主播名称：</p> <ol style="list-style-type: none">1. 该名称主要用于标识该定制形象/声音，可以根据实际情况自定义2. 命名参考：如果该主播定制的数量只有1个，可以直接以主播名称命名，例如“张三”；也可加上衣服名称来更好的标识，例如“张三蓝西服”等3. 不超过50个字符，不少于2个字符，只允许有汉字、字母、数字、下划线、横线4. 不能重名
MakeType	string	是	<p>定制类别：</p> <p>IMAGE：2D小样本-标准版（专属口型）形象定制</p> <p>IMAGE_GENERAL：2D小样本-标准版（通用口型）形象定制</p> <p>IMAGE_4K：2D小样本-高精版（专属口型）形象定制</p> <p>IMAGE_PHOTO：2D小样本（照片数智人）形象定制</p> <div style="border: 1px solid #00a88f; padding: 5px; margin: 10px 0;"><p>注意：选择 IMAGE_PHOTO 类型时，强烈建议预先执行 人脸检测任务创建 以提高成功率</p></div> <p>VOICE：声音复刻（基础版）</p> <p>ZERO_SHOT_VOICE：声音复刻（极速版）</p>
IdentityCosUrl	string	否	<p>视频格式授权书 URL，要求：</p> <ol style="list-style-type: none">1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名/customer-pipeline/{数字}/{uuid}/idcard/a.mp42. 该格式主要适用于“口述版授权书”，也可将“书面版授权书”拍成完整清晰的视频形式提交

			<p>备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与IdentityWrittenCosUrl 二选一或二者都填</p> <p>3. 大小要求：不超过5GB</p>
IdentityWrittenCosUrl	string	否	<p>pdf 格式授权书 URL，要求：</p> <ol style="list-style-type: none"> 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名/customer-pipeline/{数字}/{uuid}/idcard/b.pdf 2. 该格式主要适用于“书面版授权书”，以完整清晰的扫描件形式提交 <p>备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与IdentityCosUrl 二选一或二者都填</p> <p>3. 大小要求：不超过10MB</p>
MaterialCosUrl	string	否	<p>形象定制素材 url 地址要求：</p> <ol style="list-style-type: none"> 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源url 地址后增加 video 路径，例如/customer-pipeline/{数字}/{uuid}/video/c.mp4 2. 视频大小，不超过5GB，4K不超过10GB 3. 视频时长：专属口型版本定制时长为2-10分钟；通用口型版本定制时长为1-10分钟；高精版本定制时长为2-10分钟 4. 视频分辨率：1080P、4K（3840*2160），如定制高精版，则必须为4K 5. 视频长宽比：16:9（9:16） 6. 视频帧率：不低于25fps，不超过60fps 7. 视频格式：mp4，mov <p>备注：形象定制时 video 路径必传</p> <p>声音定制素材 url 地址要求：</p> <ol style="list-style-type: none"> 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如/customer-pipeline/{数字}/{uuid}/audio/c.zip 2. 支持音频文件或压缩包。压缩包格式：.zip 格式的压缩包，单个 zip 包用于定制1个音色 3. 音频注意事项： <ol style="list-style-type: none"> ①音频数量：每个 zip 包里可以包含1条或多条音频，建议总量不超过10条 ②音频大小：每个 zip 包里的音频大小总和不超过1GB ③音频格式：支持 wav、mp3、m4a、aac、wma 和 asf 格式 ④音频采样率：采样率大于等于24k即可，建议24k、36k ⑤音频命名：名称不要有空格或者特殊字符

			<p>备注：除 ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型必填</p> <p>照片数智人定制素材 url 地址要求：</p> <ol style="list-style-type: none"> url 地址为通过上传素材到腾讯云 COS上传到指定路径的资源 url 地址后增加 photo 路径，例如/customer-pipeline/{数字}/{uuid}/photo/example.png 图片名称：不少于2个字符，只允许有汉字、字母、数字、下划、横线；图片格式：支持jpg、jpeg、png、webp；图片大小：不超过16MB；图片宽高比：支持 1:1 9:16 16:9 4:3 人物清晰正面照，脸部居中，表情自然，嘴部闭合 <p>备注：照片定制时 photo 路径必传</p>
SexType	string	是	<p>性别：</p> <p>MALE：男</p> <p>FEMALE：女</p>
Notes	string	否	定制备注信息，100字以内
TextDriver	string	否	用于生成驱动 demo 的文本内容，字符长度在4-1000之间（包括 ssm1 标签,单个中文算一个字符）
VoiceDriver CosFile	string	否	<p>生成驱动 demo 的音频文件路径，要求：</p> <ol style="list-style-type: none"> url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如:/customer-pipeline/{数字}/{uuid}/audio/example.wav 音频文件大小不超过10MB，格式支持wav、mp3、wma、m4a、aac
Audiold	string	否	ZERO_SHOT_VOICE 定制类型必填：填写 查询音质检测任务进度 通过后返回的 Audiold
TimbreInfo	Timbre Template	否	形象定制时若需要指定音色，或尝试用视频里的声音快速复刻，则填充该字段。
IsHaveBackground	bool	否	<p>形象定制类型，训练出来的形象是否带原始背景。默认“否”，即不带原始背景，在应用过程中可以按需更换背景</p> <p>true：保留拍摄背景</p> <p>false：抠除拍摄背景（支持更换背景，仅支持绿幕视频和白墙视频）</p>

RemoveBackground	int	否	当 IsHaveBackground=false 时，进一步填充该参数。默认值是 1。 1: 抠除拍摄背景-绿幕视频 2: 抠除拍摄背景-白幕视频
AlphaCosUrl	string	否	自定义 Alpha 通道视频地址，视频上传路径参照 形象定制素材 路径
ZeroshortLang	string	否	ZERO_SHOT_VOICE 定制类型的语言类型，chi-中文普通话，eng-英文，can-粤语，默认chi
PhotoVersion	int	否	默认0 0: 快速版。在10分钟内即可完成数智人定制，只支持头部轻微晃动，面部特征稳定。建议执行 人脸检测任务 (CheckerVersion=1) 1: 优享版。在1个小时内可完成数智人定制，头部和身体都支持轻微晃动。建议执行 人脸检测任务 (CheckerVersion=2)。

TimbreTemplate 信息参数

参数	类型	必须	说明
TimbreCode	string	否	制作文本驱动 demo 时指定的音色 code
Speed	string	否	制作文本驱动 demo 时指定音色的语速,默认1.0。语速（范围[0.5-1.5]，值为0.5时播报语速最慢，值为1.5时播报语速最快）
IsWithTimbre	bool	否	尝试用视频里的声音快速复刻，额外生成 demo（由于视频里的音质质量不能保证，可能会存在快速复刻失败的情况）。该 demo 不影响其他 demo 的正常合成。

返回参数

参数	类型	必须	说明
TaskId	string	是	制作的任务 ID，携带 TaskId 访问 进度查询接口 ，可获得制作进度和制作结果

请求示例

```
{
  "Header": {},
  "Payload": {
```

```
"AnchorName": "马老师",
"MakeType": "IMAGE_GENERAL",
"IdentityCosUrl": "https://virtualhuman-cos-prod-
1251316161.cos.accelerate.myqcloud.com/customer-pipeline/100/f30e4d52-
00ba-4e36-8a26-26c348bf786e/idcard/identity.mp4",
"MaterialCosUrl": "https://virtualhuman-cos-prod-
1251316161.cos.accelerate.myqcloud.com/customer-pipeline/100/78675441-
08bf-43f2-93db-f3936068bce7/video/test.mp4",
"SexType": "FEMALE",
"RemoveBackground": 2
}
}
```

返回示例

```
{
  "Header": {
    "RequestID": "c18aeee4e49d89e408231c5b3ceea23c",
    "SessionID": "gzfa9105ca17310669340273673",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "TaskId": "1212_29934_w"
  }
}
```

进度查询接口

最近更新时间：2025-03-07 12:14:53

通过TaskId来查询任务的制作进度和结果。

调用协议

HTTPS + JSON

POST /v2/ivh/assetmanager/customservice/getprogress

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

请求参数

参数	类型	必须	说明
TaskId	string	是	定制接口 返回的 TaskId

返回参数

参数	类型	必须	说明
Status	string	是	任务制作状态： MAKING：进行中 SUCCESS：制作成功 FAIL：制作失败
FailMessage	string	否	当 Status: FAIL 时，返回制作失败原因，便于排查问题
Stage	string	否	当前定制阶段（当前字段是为了兼容历史定制任务，新定制可忽略，使用 StageV2）
StageId	string	否	定制阶段 Id（当前字段是为了兼容历史定制任务，新定制可忽略，使用 StageV2）
StageV2	string	否	（自2024.1.1之后发起的进度查询请求，以当前字段为准） 当 Status: MAKING 时，返回任务进行中的具体阶段： SUBMIT：素材提交 AUDIT：人工审核 PROCESS：预处理（仅形象定制有） TRAIN：智能训练 CONFIRM：效果确认（照片定制无此阶段） END：完成

StatusV2	string	否	当 StageV2 处于某个阶段时，该阶段的具体状态： WAIT_PROCESS: 待处理 PROCESSING: 处理中 SUCCESS: 成功 FAIL: 失败
StageInfo	string	否	处于某个阶段的额外信息，JSON 格式的字符串，不同阶段的 JSON 格式类型见下述表格，目前只有效果确认阶段（StageV2 为 CONFIRM）有值。
AssetList	Array of [AssetInfo]	否	返回定制的形象或音色信息，目前支持返回的定制类型有： IMAGE: 2D小样本-标准版（专属口型）形象定制 IMAGE_GENERAL: 2D小样本-标准版（通用口型）形象定制 IMAGE_4K: 2D小样本-高精版（专属口型）形象定制 IMAGE_PHOTO: 2D小样本（照片数智人）形象定制 ZERO_SHOT_VOICE: 声音复刻（极速版） VOICE: 声音复刻（基础版）

不同阶段 StageInfo 额外信息见下表：

StageV2	类型	描述	示例
CONFIRM: 效果确认阶段	string of [CustomerConfirmContent]	当阶段处于效果确认阶段（StageV2: CONFIRM），且状态处于待处理中时（StatusV2: WAIT_PROCESS），返回待确认效果 Demo 视频，客户可以根据返回的 Demo 视频通过调用 效果确认接口 对 Demo 视频效果进行认同/驳回操作	<pre>{\TrainResult\":{\url1\", \url2\"}, \RejectReason\": \"\", \TrainRecord\": [{\StartTime\": \"2023-11-01 23:00:17\", \Record\": {\url1\", \url2\"}}]}</pre>

CustomerConfirmContent

参数	类型	必须	说明
TrainResult	Array of [string]	是	供客户进行效果确认的 Demo 视频 url 地址
RejectReason	string	是	数智人侧不接受重训的理由
TrainRecord	Array of [TrainRecord]	否	历史训练的 Demo 视频清单

TrainRecord

参数	类型	必须	说明
----	----	----	----

StartTime	string	是	历史训练时间
Record	Array of [string]	是	历史训练的 Demo 视频 url 地址

AssetInfo

参数	类型	必须	说明
VirtualmanTypeCode	string	否	数智人类型 Code，参考 附录3-数智人类型 ，形象定制返回
VirtualmanKey	string	否	数智人 virtualmankey，形象定制返回
Resolution	string	否	数智人分辨率，形象定制返回
OriginZoom	string	否	缩放系数，保留三位小数，形象定制返回
VirtualmanResourceId	int32	否	数智人资源 ID，形象定制返回
TimbreKey	string	否	音色key，声音复刻返回
TimbreSample	string	否	音色demo链接，声音复刻返回
MakeType	string	是	定制类型： IMAGE_PHOTO: 2D小样本（照片数智人）形象定制 ZERO_SHOT_VOICE: 声音复刻（极速版）

请求示例

```
{
  "Header": {},
  "Payload": {
    "TaskId": "1212_29934_w"
  }
}
```

```
}  
}
```

返回示例

```
{  
  "Header": {  
    "RequestID": "gz1a74c25f17030396653416949",  
    "SessionID": "gz1a74c25f17030396653416950",  
    "DialogID": "",  
    "Code": 0,  
    "Message": "ok"  
  },  
  "Payload": {  
    "StageInfo": "{\\"TrainResult\\":  
[\\"url1\\",\\"url2\\"],\\"RejectReason\\":\\"\\",\\"TrainRecord\\":  
[{\\"StartTime\\":\\"2023-11-01 23:00:17\\",\\"Record\\":  
[\\"url1\\",\\"url2\\"]}]}",  
    "Status": "MAKING",  
    "StatusV2": "WAIT_PROCESS",  
    "FailMessage": "",  
    "StageV2": "CONFIRM"  
  }  
}
```

效果确认接口

最近更新时间：2024-11-13 15:22:12

当通过进度查询接口查到处于待确认效果阶段时（Status:MAKING，StageV2:CONFIRM，StatusV2:WAIT_PROCESS），对 demo 视频进行效果确认。

调用协议

HTTPS + JSON

POST /v2/ivh/assetmanager/customservice/customerconfirm

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

请求参数

参数	类型	必须	说明
TaskId	string	是	定制接口 返回的 TaskId
Operate	string	是	操作： AGREE：认同 REJECT：驳回
Reason	string	否	驳回原因，当 Operate:REJECT 时必填，最多300字

返回参数

请求示例

```
{
  "Header": {},
  "Payload": {
    "TaskId": "1212_29934_w",
    "Operate": "AGREE"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "gz1a74c25f17030396653416949",
  }
}
```

```
"SessionID": "gz1a74c25f17030396653416950",  
"DialogID": "",  
"Code": 0,  
"Message": "ok"  
},  
"Payload": {}  
}
```

人脸检测任务创建

最近更新时间：2024-12-25 16:47:53

传入 RequestId, 图片的 url 和检测版本, 创建人脸检测任务, 返回 TaskId。使用 TaskId 可以进行进度查询。通过预先的人脸检测, 可以提高照片数智人的定制成功率。

调用协议

HTTPS + JSON

POST /v2/ivh/facedetectserver/facedetectservice/facedetect

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

请求参数

参数	类型	必须	说明
RequestId	string	是	本次请求的唯一标识ID 示例值: 569ae497-f266-4565-93bd-93aaf822524e
ImageUrl	string	是	要进行检测的图片的 url
CheckerVersion	int	否	检测版本, 支持项: [1,2]。不传默认为1。 值为1时: <ul style="list-style-type: none">检查图片的比例, 目前支持[9:16, 16:9, 1:1, 4:3]检查结果只包含人脸框信息和角度 值为2时: <ul style="list-style-type: none">不检查图片比例检查结果包含遮挡比例, 建议裁剪的区域坐标等信息。根据建议裁剪的区域对原图片进行裁剪, 然后再进行数智人制作效果会更好。

返回参数

参数	类型	必须	说明
TaskId	string	是	携带TaskId访问 人脸检测任务状态查询 接口, 可获得检测进度和检测结果

请求示例

```
{
  "Header": {},
  "Payload": {
    "RequestId": "569ae497-f266-4565-93bd-93aaf822524e",
    "ImageUrl": "https://virtualhuman-cos-prod-1251316161.cos.ap-gangzhou.myqcloud.com/customer-pipeline/704/e60fe7c8-ec69-44cf-bc5e-c4c754614e82/photo/48922.png",
    "CheckerVersion": 2
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "gz1a74c25f17030396653416949",
    "SessionID": "gz1a74c25f17030396653416950",
    "DialogID": "",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "TaskId": "569ae497-f266-4565-93bd-93aaf822524e"
  }
}
```

人脸检测任务状态查询

最近更新时间：2024-11-13 15:25:52

通过 TaskId 来查询人脸检测任务的进度和结果。

调用协议

HTTPS + JSON

POST /v2/ivh/facedetectserver/facedetectservice/getfacedetectstate

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

请求参数

参数	类型	必须	说明
TaskId	string	是	人脸检测任务创建 接口返回的TaskId

返回参数

参数	类型	必须	说明
TaskId	string	是	任务ID
TaskState	string	是	任务状态： <ul style="list-style-type: none">• WAITING: 任务等待中• PROCESSING: 任务执行中• SUCCESS: 任务执行成功• FAILED: 任务执行失败
ErrorCode	int	否	任务执行失败时有值，错误码，可参考附录
ErrorMsg	string	否	任务执行失败时有值，错误信息
IsPassDetect	bool	否	任务执行成功时有值，是否通过人脸检测
Faces	Array of FaceInfo	否	任务执行成功时有值，检测到的人脸框

FaceInfo

参数	类型	必须	说明
Box	Box Info	否	检测到的人脸框
Rotation	float	否	人脸的角度
Score	float	否	人脸置信度 *
AreaRatio	float	否	人脸区域占比 *
FaceRadio	float	否	脸部遮挡比例 0~1, 越大遮挡越大 *
FeatureRadio	float	否	五官遮挡比例 0~1, 越大遮挡越大 *
Yaw	float	否	yaw角度, -90~90 表明侧脸程度, 越大姿态越大 *
Pitch	float	否	pitch角度, -90~90 表明俯仰程度, 越大姿态越大 *
MaxCrop	Box Info	否	建议的最大裁剪区域 *
MinCrop	Box Info	否	建议的最小裁剪区域 *

说明:

表格中带 * 的参数只有创建时 CheckerVersion=2 的任务才会有检测结果。

BoxInfo

参数	类型	必须	说明
Left	int	否	左上角的x坐标点
Top	int	否	左上角的y坐标点
Width	int	否	宽度

Height	int	否	高度
--------	-----	---	----

请求示例

```
{
  "Header": {},
  "Payload": {
    "TaskId": "569ae497-f266-4565-93bd-93aaf822524e"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "gz1a74c25f17030396653416949",
    "SessionID": "gz1a74c25f17030396653416950",
    "DialogID": "",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "ErrorCode": 0,
    "IsPassDetect": true,
    "Faces": [
      {
        "Yaw": 11.440315246582,
        "Pitch": -3.8738098144531,
        "MaxCrop": {
          "Top": 0,
          "Height": 512,
          "Width": 512,
          "Left": 0
        },
        "MinCrop": {
          "Top": 0,
          "Height": 512,
          "Width": 512,
          "Left": 0
        }
      }
    ]
  }
}
```

```
    },  
    "Box": {  
      "Top": 103,  
      "Height": 406,  
      "Width": 320,  
      "Left": 85  
    },  
    "Rotation": -7,  
    "Score": 1,  
    "AreaRatio": 0.49560546875,  
    "FaceRadio": 0,  
    "FeatureRadio": 0  
  }  
],  
"TaskState": "SUCCESS",  
"ErrorMsg": "",  
"TaskId": "569ae497-f266-4565-93bd-93aaf822524e"  
}  
}
```

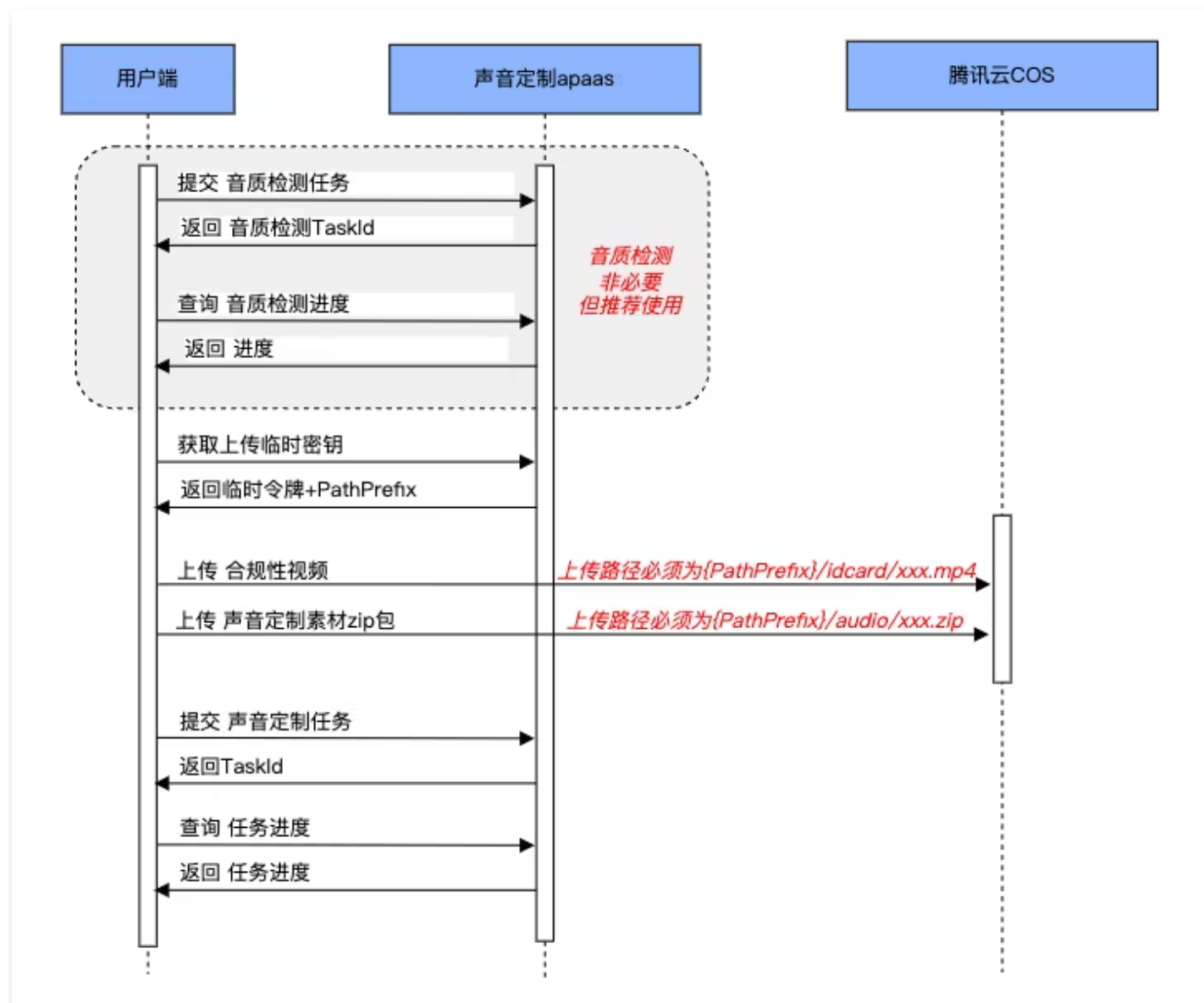
附录一：错误码字典

代码	含义	处理方式
100001	请求参数错误	修正请求入参
100002	参数取值错误	修正请求入参
100003	图片格式不支持	选择新的图片
100004	图片尺寸不支持	调整图片尺寸重新检测
100005	图片比例不支持	调整图片比例重新检测
200001	内部错误	联系管理员
200002	文件下载失败	联系管理员
200003	任务不存在	联系管理员
300001	图片中没有检测到人脸	更换新的图片
300002	图片中检测到多个人脸	更换新的图片

300003	人脸框不满足要求，可能过大或者过小	更换新的图片
--------	-------------------	--------

声音复刻（基础版） 接口调用逻辑示意图

最近更新时间：2025-03-14 16:06:02



1. 音质检测阶段 (可选)

- 用户提交音质检测任务，系统返回一个音质检测 TaskId。
- 用户可选择性查询任务进度，系统返回当前进度。

2. 文件上传准备

- 用户需先获取上传临时密钥，系统返回临时令牌 + PathPrefix。

3. 文件上传要求

- 合规性视频：必须上传至路径 `(PathPrefix)/idcard/xxx.mp4` 。
- 声音定制素材压缩包：必须上传至路径 `(PathPrefix)/audio/xxx.zip` 。

4. 声音定制任务提交

- 用户提交声音定制任务，系统返回声音定制 **TaskId**。
- 用户可查询任务进度，系统返回当前状态（如处理中、完成等）。

音质检测任务创建接口

最近更新时间：2025-12-01 17:18:11

可以通过该接口提前检测声音质量，提升定制效率。

支持的音质指标：

- 信噪比指数：大于等于25合格
- 混响指数：大于等于30合格
- 截幅指数：小于等于10合格
- 有效时长：大于等于5分钟合格
- 文本准确率：（当传入“音频参考文本”时该指标有效）
 - 中文普通话：大于等于85%合格
 - 英文：大于等于50%合格
 - 粤语：大于等于70%合格

调用协议

HTTPS + JSON

POST /v2/ivh/algotaskserver/algotaskservice/audioevaluation

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

请求参数

参数	类型	必须	说明
AudioUrl	String	是	待检测音频的 URL，音频格式目前支持 wav、mp3、m4a、aac、wma 和 asf 格式，建议文件大小不超过 20MB
ReferenceText	String	否	音频参考文本，可选。不填表示不需要“文本准确率”指标
Language	String	否	语言类型。chi-中文普通话，eng-英文，can-粤语。默认chi

返回参数

参数	类型	必须	说明
TaskId	String	是	已创建音质检测任务的 ID

请求示例

```
{
  "Header": {},
  "Payload": {
    "AudioUrl": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipeline/100/63232275-a349-4f27-8e33-aa4ac791094c/voice_record/hello.wav",
    "ReferenceText": "蓝蓝的天空银河里，有只小白船，船上有棵桂花树，白兔在游玩。渡过那条银河水，走向云彩国，走过那个云彩国，飘向明天",
    "Language": "chi"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "1e2c786806e6a59d8c2c70bdbd94f70e",
    "SessionID": "gza802cc9317328691207007711",
    "DialogID": "",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "TaskId": "837a583863b745e59391456943ffc22f"
  }
}
```

音质检测任务状态查询接口

最近更新时间：2024-11-29 11:26:32

通过 TaskId 来查询音质检测任务的状态和结果。

调用协议

HTTPS + JSON

POST /v2/ivh/algotaskserver/algotaskservice/getaudioevaluationstate

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

请求参数

参数	类型	必须	说明
TaskId	string	是	音质检测任务创建接口 返回的 TaskId

返回参数

参数	类型	必须	说明
TaskState	string	是	任务状态： CREATED：已创建 QUEUING：排队中 PROCESSING：处理中 SUCCESS：任务完成 FAIL：任务失败
Result	object	是	音质检测任务结果
Result.Snr	object	是	音频的信噪比指标
Result.Snr.Value	float	是	音频的信噪比指标测量值
Result.Snr.QualifiedRange	Array of [float]	是	音频的信噪比指标合格范围 [lower_limit,upper_limit]
Result.Snr.Description	string	是	音频的信噪比指标说明
Result.C50	object	是	音频的混响指标
Result.C50.Value	float	是	音频的混响指标测量值

Result.C50.QualifiedRange	Array of [float]	是	音频的混响指标合格范围 [lower_limit,upper_limit]
Result.C50.Description	string	是	音频的混响指标说明
Result.Clipping	object	是	音频的截幅指标
Result.Clipping.Value	float	是	音频的截幅指标测量值
Result.Clipping.QualifiedRange	Array of [float]	是	音频的截幅指标合格范围 [lower_limit,upper_limit]
Result.Clipping.Description	string	是	音频的截幅指标说明
Result.Duration	object	是	音频的时长指标
Result.Duration.Value	float	是	音频的时长
Result.Duration.QualifiedRange	Array of [float]	是	音频的时长合格范围 [lower_limit,upper_limit]
Result.Duration.Description	string	是	音频的时长指标说明
Result.Wer	object	是	音频的文本准确率指标
Result.Wer.Value	float	是	文本准确率，仅“参考文本”不为空时有效
Result.Wer.QualifiedRange	Array of [float]	是	文本准确率合格范围 [lower_limit,upper_limit]
Result.Wer.Description	string	是	文本准确率说明
Result.AlignDetails	array of [TextAlign]	是	文本对齐结果
Result.Sentences	array of [Sentences]	是	短句截取时间
ErrorMsg	string	是	任务失败提示信息

TextAlign

参数	类型	必须	说明
Detail	Array of [string]	是	文本每个字符对齐结果，如["好","好"]/"好", "<eps>"], 其中"<eps>"为占位符

Sentences

参数	类型	必须	说明
StartMs	int	是	句子开始毫秒数
EndMs	int	是	句子结束毫秒数
Sentence	string	是	句子内容

请求示例

```
{
  "Header": {},
  "Payload": {
    "TaskId": "63232275-a349-4f27-8e33-aa4ac791094c"
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "Message": "",
    "RequestID": "28499a-cf6d-4f7b-b7fa-4a5a54662b38"
  },
  "Payload": {
    "Status": "SUCCESS",
    "Result": {
      "Snr": {
        "Value": 72.67048,

```

```
"QualifiedRange": [
  30,
  2147483600
],
>Description": "音频的信噪比, 越大越好"
},
"C50": {
  "Value": 57.082405,
  "QualifiedRange": [
    30,
    2147483600
  ],
>Description": "混响指标, 越大越好"
},
"Clipping": {
  "QualifiedRange": [
    -2147483600,
    0
  ],
>Description": "截幅指数: 小于等于10合格",
  "Value": 0
},
"Duration": {
  "Value": 24.21,
  "QualifiedRange": [
    600,
    2147483600
  ],
>Description": "有效时长: 单位秒, 大于等于10分钟合格"
},
"Wer": {
  "Value": 0.91358024,
  "QualifiedRange": [
    0.7,
    1
  ],
>Description": "文本准确率: 大于等于70%合格"
},
"Sentences": [
  {
```

```
    "StartMs": 340,
    "EndMs": 4510,
    "Sentence": "激光祛斑是利用激光的光热解原理，"
  },
  {
    "StartMs": 4510,
    "EndMs": 6040,
    "Sentence": "激光照射皮肤后，"
  },
  {
    "StartMs": 6040,
    "EndMs": 8640,
    "Sentence": "会被皮肤病变部位的色素吸收，"
  },
  {
    "StartMs": 8900,
    "EndMs": 12660,
    "Sentence": "而激光产生的高能量就会将这些色素击碎，"
  },
  {
    "StartMs": 12890,
    "EndMs": 17155,
    "Sentence": "色素碎屑会随人体的各项循环排除体外，"
  },
  {
    "StartMs": 17155,
    "EndMs": 19050,
    "Sentence": "从而达到祛斑的效果。"
  }
],
"AlignDetails": [
  {
    "Detail": [
      "<eps>",
      "大"
    ]
  },
  {
    "Detail": [
      "家",

```

```
        "家"  
    ]  
  },  
  {  
    "Detail": [  
      "<eps>",  
      "好"  
    ]  
  }  
]  
},  
"ErrorMsg": ""  
}  
}
```

获取上传临时令牌

最近更新时间：2024-11-18 17:24:52

在调用 [定制接口](#) 前，需要将素材文件上传到指定的云端存储。上传时，首先通过该接口获取上传临时令牌和指定的上传路径，再通过 [上传素材到腾讯云 COS](#) 接口上传对应的文件到云端存储。

调用协议

HTTPS + JSON

POST /v2/ivh/assetmanager/customservice/getuploadcredentials

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

请求参数

无

返回参数

参数	类型	必须	说明
Credentials	object	是	临时密钥信息
Credentials.Token	string	是	临时密钥令牌
Credentials.TmpSecretId	string	是	临时证书密钥 ID
Credentials.TmpSecretKey	object	是	临时证书密钥 Key
StartTime	int	是	临时证书有效起始时间，返回 Unix 时间戳，精确到秒
ExpiredTime	int	是	临时证书有效截止时间，返回 Unix 时间戳，精确到秒
PathPrefix	string	是	上传到 COS 上的路径前缀

返回示例

```
{  
  "Header": {
```

```
"RequestID": "c18ae4e49d89e408231c5b3ceea23c",
"SessionID": "gzfa9105ca17310669340273673",
"Code": 0,
"Message": "ok"
},
"Payload": {
  "Credentials": {
    "TmpSecretId": "AKID*****ib",
    "TmpSecretKey": "DVki*****k0=",
    "Token":
"8oz1l*****49esQlG0940LX2tacXiCv4YS2qdBsiQ"
  },
  "ExpiredTime": 1731074134,
  "PathPrefix": "https://cos-test.ap-
guangzhou.myqcloud.com/customer-pipeline/123/4506c29b-2904-45e2-b7d7-
3052fe2493c2/",
  "StartTime": 1731066934
}
}
```



```

"Payload": {
  "Credentials": {
    "TmpSecretId": "AKIDM6p_7lbof_ow5_DTYz...Qzb1MlvtFfjxqbg1FrN9",
    "TmpSecretKey": "9pkrcGy6duNWhj4/...J87PGk=",
    "Token": "fGSqNKc3Q90WcLEz5jn2pdHo3cGEuPKa23738bd8cb693d875f4e2ead9f6ff446w9B4mmLxcks10FXXJAW-jtudiJnfdPht0RvkEC3PshPrv2nrXPLjwTst8A-mGxVxkpM-YIE_RpeuQG2nIKF9izYreejnrzx6oCuCL61AXy39v6CdCHXoWkdWzyc1LWio2QVMrshS0uKr26H770MmVdjH9Evk19I9V8K4KR7mJaE3GDisk5tJpAqFtNd5tSRGSIeRdHuf2jUef9WvnWuV1yysSpJy7qr0KqQh7GwzWkv4j2jtr1B6fbNzYfTxPHvijLwDC50Dg613pgiZqwtu7GtoUxdFOXWQny1aHiiz1NnRR25ELot_o7tAsdL6Ae7uj-31bkyQbU88e5JHSRvS4p0QVe58LUNw5WhpTeJlXh-LRsPmc2z29MmXtSvWkgx2Y-KiQBYHsF81paEMuHK4eU60303xdjd5_h6N4C0u6QFet3PUJY7Kgt1t7L4K0670QJERrikEgBYppyZYqvVHaD4YZO-TABrhymYPUVkdmbzW4Idim8tZUI6Fc93FGyYRAFfql3zux4IX-NrieAGMRifks1q1XfRfHfFo2iQeaLyUubddGoyVvK6y5eIj1kou5mM0drPBpPzYoZJWalHT12mbm-hDKHvcgBe-bMe7qs7ze6NiyqwrECDCrVmdTA8j0EVKPGuE94xAnMyytYHnhmQvC8o77hBuQvgeCPiX-R_xAUD-6ydv5Gi_8mLcd07h35Nvk0TrXyNuMTmqjTEbo1sVC74jP68qPMZny--UU1Z9C_tI6Y938xwf5xPB70MzZMTP7IEnMPqsDZgGwWg1R-U-JbMzq-HkDBU...B0d6Qf3XN2PCL1q1gkJuS8jZBLQ4K4c5XVvfbEpSw4TcLn500yqp6ks7cEXBZeAhAowRj_APBDNTI72-XnlydgBbcjY6G3xXuXDRbBGUs8QrCi_3mMYNMyNq4flBanR1sdAnjFKV7UNQ0FMsEFWpkyLD4cs4bEbhbatMt-lu37oVXD4EUcnNyR6Yz1jNws27n8gAL0I9FDEb-aHICXaYyG5cEgELghf16gJsomjwZa0FpXtI4RMPBP_6D0f6RGE0IF_0r5Awh_9_QsGwIXkUW18aeVw7C0fexQ8Cf0vwU-zk0q2F4x9CpQwUdsN18FIYC3I3sYTViS5gIh2P_-L0dkzwjplz5-yRglbdjh5ba1jXm1GqhipeohMsYxQSHDfgkAm1WFwqpp4vkt-0S-1kvY94BqTRkm8M2_0vornZIDx8n3DZJiI657HDL33d4cGncSv0Q_R49SRiNK471Cne-NDoqRmCBvEXY-mgDS7Fht2F5vKxueYvX0iRnmP4_d75qDpYn1_PNYCoNj4ViYSICPXJUNeofRTJc2miwSENL7vHOH0kM02hNikB46LhLPZtK51sd4g1JozesLRcbD0sAZyJd_5e1pE7SaulMd7FMkQq3th0CwacrIy062DncwEQBBM-00uwr7CpBdhakLAXn-RAO-C12ogMw-Wx1BQ57ixN7yjE28ukSzG8QjC14_QUQsG2Y8FKQHVY5hZp1AGhjpBCN_7XzH-WAPGfzrGXJW1DT-zkHJYRPMbB19zx2WCifCyChtbz7XReQCN6BmmXvtqApjKAdr9ZQctTK0aNdvm53zz1jxdoHttAkJL9YCPp0iN_2YjHI5t4e7PMc_Itsxg2RgDy83EL8Iz0lqs57WtRErFQspFgtAWr3dmjgTYtHehPIP_K42k5WA2t6g3031z_3_zaGMTzngv4u0PTvYz-QpfqAQMDQPaRL132QTKariImA8Ff8d10DhjuvBDPzIwzA6085tXTJMhJbK1DZwDpw3KLJUiqAgfWcNhoHJmrdmJo_AvYz9yUaqNqH0GAIEUtYlZsh3emcD6vb1V2kCci3GjzrTsg"
  }
},
"ExpiredTime": 1736241459,
"PathPrefix": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipeline/352/59c6f6cd-8f92-45f9-8352-64ca8fb92eaa/",
"StartTime": 1736234259

```

- bucket名称: 解析出来的值为 virtualhuman-cos-prod-1251316161。
- region名称: 解析出来的值为 accelerate。
- key(bucket 下的存储路径): 解析出来的值为 customer-pipeline/352/59c6f6cd-8f92-45f9-8352-64ca8fb92eaa/。

注意:

解析出来的值为 key 的前半部分，还需要根据上传者类型来拼接后半部分，如上传的是 pdf 授权书素材，则需要拼接上 idcard/xxx.pdf，如下图所示，具体后半部分值的定义请参考定制接口中的相关参数说明。

腾讯云智能数智人	IdentityWrittenCosUrl	string	否	pdf 格式授权书 URL，要求： 1. url 地址为通过上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名/customer-pipeline/(数字)/(uuid)/idcard/b.pdf 2. 该格式主要适用于“书面版授权书”，以完整清晰的扫描件形式提交 备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与 IdentityCosUrl 二选一或二者都填	本项目目录： 调用协议 请求参数 返回参数 请求示例 返回示例
形象定制	MaterialCosUrl	string	否	形象定制素材 url 地址要求： 1. url 地址为通过上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 video 路径，例如/customer-pipeline/(数字)/(uuid)/video/c.mp4 2. 视频大小，不超过5GB，4K不超过10G 3. 视频时长：专属口型版本定制时长为2-10分钟；通用口型版本定制时长为1-10分钟；高精版本定制时长为2-10分钟 4. 视频分辨率：1080P、4K (3840*2160)，如定制高精版，则必须为4K 5. 视频长宽比：16:9 (9:16) 6. 视频帧率：不低于25fps，不超过60fps 7. 视频格式：mp4, mov 备注：形象定制时 video 路径必传 声音定制素材 url 地址要求： 1. url 地址为通过上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如/customer-pipeline/(数字)/(uuid)/audio/c.zip 2. 压缩包格式：.zip 格式的压缩包，单个 zip 包用于定制1个音色；压缩时不要新建文件夹，直接选中所有 wav 压缩 3. 对于单个 zip 包里的音频注意事项： ①音频数量：每个 zip 包里可以包含1条或多条 wav 格式音频，总量不超过10条 ②音频大小：每个 zip 包里的音频大小总和不超过1G ③音频格式：每条音频需为 wav 格式，其他格式的音频请先转为 wav 格式后再再压缩成 zip 包 ④音频采样率：采样率大于等于24k即可，建议24k、36k	

定制接口

最近更新时间：2025-12-19 14:24:31

使用该接 提交定制的请求，通过 [进度查询接口](#) 查询定制的各个阶段以及各阶段的相关信息。

调用协议

HTTPS + JSON

POST /v2/ivh/assetmanager/customservice/make

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

请求参数

参数	类型	必须	说明
AnchorName	string	是	<p>主播名称：</p> <ol style="list-style-type: none">1. 该名称主要用于标识该定制形象/声音，可以根据实际情况自定义2. 命名参考：如果该主播定制的数量只有1个，可以直接以主播名称命名，例如“张三”；也可加上衣服名称来更好的标识，例如“张三蓝西服”等3. 不超过50个字符，不少于2个字符，只允许有汉字、字母、数字、下划线、横线4. 不能重名
MakeType	string	是	<p>定制类别：</p> <p>IMAGE：2D小样本-标准版（专属口型）形象定制</p> <p>IMAGE_GENERAL：2D小样本-标准版（通用口型）形象定制</p> <p>IMAGE_4K：2D小样本-高精版（专属口型）形象定制</p> <p>IMAGE_PHOTO：2D小样本（照片数智人）形象定制</p> <div style="border: 1px solid #00a88f; padding: 5px; margin: 10px 0;"><p>⚠ 注意：选择 IMAGE_PHOTO 类型时，强烈建议预先执行 人脸检测任务创建 以提高成功率</p></div> <p>VOICE：声音复刻（基础版）</p> <p>ZERO_SHOT_VOICE：声音复刻（极速版）</p>
IdentityCosUrl	string	否	<p>视频格式授权书 URL，要求：</p> <ol style="list-style-type: none">1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名/customer-pipeline/{数字}/{uuid}/idcard/a.mp42. 该格式主要适用于“口述版授权书”，也可将“书面版授权书”拍成完整清晰的视频形式提交

			<p>备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与IdentityWrittenCosUrl 二选一或二者都填</p>
IdentityWrittenCosUrl	string	否	<p>pdf 格式授权书 URL，要求：</p> <ol style="list-style-type: none"> 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名/customer-pipeline/{数字}/{uuid}/idcard/b.pdf 2. 该格式主要适用于“书面版授权书”，以完整清晰的扫描件形式提交 <p>备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与IdentityCosUrl 二选一或二者都填</p>
MaterialCosUrl	string	否	<p>形象定制素材 url 地址要求：</p> <ol style="list-style-type: none"> 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源url 地址后增加 video 路径，例如/customer-pipeline/{数字}/{uuid}/video/c.mp4 2. 视频大小，不超过5GB，4K不超过10G 3. 视频时长：专属口型版本定制时长为2-10分钟；通用口型版本定制时长为1-10分钟；高精版本定制时长为2-10分钟 4. 视频分辨率：1080P、4K（3840*2160），如定制高精版，则必须为4K 5. 视频长宽比：16:9（9:16） 6. 视频帧率：不低于25fps，不超过60fps 7. 视频格式：mp4，mov <p>声音定制素材 url 地址要求：</p> <ol style="list-style-type: none"> 1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如/customer-pipeline/{数字}/{uuid}/audio/c.zip 2. 支持音频文件或压缩包。压缩包格式：.zip 格式的压缩包，单个 zip 包用于定制1个音色 3. 音频注意事项： <ol style="list-style-type: none"> ①音频数量：每个 zip 包里可以包含1条或多条音频，建议总量不超过10条 ②音频大小：每个 zip 包里的音频大小总和不超过1G ③音频格式：支持 wav、mp3、m4a、aac、wma 和 asf 格式 ④音频采样率：采样率大于等于24k即可，建议24k、36k ⑤音频命名：名称不要有空格或者特殊字符 <p>备注：除 ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型必填</p> <p>照片数智人定制素材 url 地址要求：</p> <ol style="list-style-type: none"> 1. url地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url

			<p>地址后增加 photo 路径，例如/customer-pipeline/{数字}/{uuid}/photo/example.png</p> <p>2. 图片名称：不少于2个字符，只允许有汉字、字母、数字、下划线、横线；图片格式：支持jpg、jpeg、png、webp；图片大小：不超过16M；图片宽高比：支持 1:1 9:16 16:9 4:3</p> <p>3. 人物清晰正面照，脸部居中，表情自然，嘴部闭合</p>
SexType	string	是	<p>性别：</p> <p>MALE：男</p> <p>FEMALE：女</p>
Notes	string	否	定制备注信息，100字以内
TextDriver	string	否	用于生成驱动 demo 的文本内容，字符长度在4-1000之间（包括ssml 标签,单个中文算一个字符）
VoiceDriverCosFile	string	否	<p>生成驱动 demo 的音频文件路径，要求：</p> <p>1. url 地址为通过上传素材到腾讯云 COS上传到指定路径的资源 url 地址后增加 audio 路径，例如:/customer-pipeline/{数字}/{uuid}/audio/example.wav</p> <p>2. 音频文件大小不超过10M，格式支持wav、mp3、wma、m4a、aac</p>
Audioid	string	否	ZERO_SHOT_VOICE 定制类型必填:填写 查询音质检测任务进度 通过后返回的 Audioid
TimbreInfo	TimbreTemplate	否	形象定制时若需要指定音色，或尝试用视频里的声音快速复刻，则填充该字段。
IsHaveBackground	bool	否	<p>形象定制类型，训练出来的形象是否带原始背景。默认“否”，即不带原始背景，在应用过程中可以按需更换背景</p> <p>true：保留拍摄背景</p> <p>false：抠除拍摄背景（支持更换背景，仅支持绿幕视频和白墙视频）</p>
RemoveBackground	int	否	<p>当 IsHaveBackground=false 时，进一步填充该参数。默认值是 1。</p> <p>1：抠除拍摄背景-绿幕视频</p> <p>2：抠除拍摄背景-白幕视频</p>
AlphaCosUrl	string	否	自定义 Alpha 通道视频地址，视频上传路径参照 形象定制素材 路径

	ng		
ZeroshortLang	string	否	ZERO_SHOT_VOICE 定制类型的语言类型，chi-中文普通话，eng-英文，can-粤语，默认chi
PhotoVersion	int	否	默认0 0: 快速版。在10分钟内即可完成数智人定制，只支持头部轻微晃动，面部特征稳定。建议执行 人脸检测任务 (CheckerVersion=1) 1: 优享版。在1个小时内可完成数智人定制，头部和身体都支持轻微晃动。建议执行 人脸检测任务 (CheckerVersion=2)。

TimbreTemplate 信息参数

参数	类型	必须	说明
TimbreCode	string	否	制作文本驱动 demo 时指定的音色 code
Speed	string	否	制作文本驱动 demo 时指定音色的语速，默认1.0。语速（范围[0.5-1.5]，值为0.5时播报语速最慢，值为1.5时播报语速最快）
IsWithTimbre	bool	否	尝试用视频里的声音快速复刻，额外生成 demo（由于视频里的音质质量不能保证，可能会存在快速复刻失败的情况）。该 demo 不影响其他 demo 的正常合成。

返回参数

参数	类型	必须	说明
TaskId	string	是	制作的任务 ID，携带 TaskId 访问 进度查询接口 ，可获得制作进度和制作结果

请求示例

```
{
  "Header": {},
  "Payload": {
    "AnchorName": "张三",
    "MakeType": "VOICE",
  }
}
```

```
"IdentityCosUrl": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipeline/100/f30e4d52-00ba-4e36-8a26-26c348bf786e/idcard/identity.mp4",
"SexType": "FEMALE",
"MaterialCosUrl": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipeline/110/77ef6290-0991-4e45-9451-0ef8587985a3/audio/张三.zip"
}
}
```

返回示例

```
{
  "Header": {
    "RequestID": "c18aeee4e49d89e408231c5b3ceea23c",
    "SessionID": "gzfa9105ca17310669340273673",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "TaskId": "1212_29934_w"
  }
}
```

进度查询接口

最近更新时间：2025-03-07 12:14:53

通过TaskId来查询任务的制作进度和结果。

调用协议

HTTPS + JSON

POST /v2/ivh/assetmanager/customservice/getprogress

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

请求参数

参数	类型	必须	说明
TaskId	string	是	定制接口 返回的 TaskId

返回参数

参数	类型	必须	说明
Status	string	是	任务制作状态： MAKING：进行中 SUCCESS：制作成功 FAIL：制作失败
FailMessage	string	否	当 Status: FAIL 时，返回制作失败原因，便于排查问题
Stage	string	否	当前定制阶段（当前字段是为了兼容历史定制任务，新定制可忽略，使用 StageV2）
StageId	string	否	定制阶段 Id（当前字段是为了兼容历史定制任务，新定制可忽略，使用 StageV2）
StageV2	string	否	（自2024.1.1之后发起的进度查询请求，以当前字段为准） 当 Status: MAKING 时，返回任务进行中的具体阶段： SUBMIT：素材提交 AUDIT：人工审核 PROCESS：预处理（仅形象定制有） TRAIN：智能训练 CONFIRM：效果确认（照片定制无此阶段） END：完成

StatusV2	string	否	当 StageV2 处于某个阶段时，该阶段的具体状态： WAIT_PROCESS: 待处理 PROCESSING: 处理中 SUCCESS: 成功 FAIL: 失败
StageInfo	string	否	处于某个阶段的额外信息，JSON 格式的字符串，不同阶段的 JSON 格式类型见下述表格，目前只有效果确认阶段（StageV2 为 CONFIRM）有值。
AssetList	Array of [AssetInfo]	否	返回定制的形象或音色信息，目前支持返回的定制类型有： IMAGE: 2D小样本-标准版（专属口型）形象定制 IMAGE_GENERAL: 2D小样本-标准版（通用口型）形象定制 IMAGE_4K: 2D小样本-高精版（专属口型）形象定制 IMAGE_PHOTO: 2D小样本（照片数智人）形象定制 ZERO_SHOT_VOICE: 声音复刻（极速版） VOICE: 声音复刻（基础版）

不同阶段 StageInfo 额外信息见下表：

StageV2	类型	描述	示例
CONFIRM: 效果确认阶段	string of [CustomerConfirmContent]	当阶段处于效果确认阶段（StageV2: CONFIRM），且状态处于待处理中时（StatusV2: WAIT_PROCESS），返回待确认效果 Demo 视频，客户可以根据返回的 Demo 视频通过调用 效果确认接口 对 Demo 视频效果进行认同/驳回操作	<pre>{\TrainResult\":{\url1\", \url2\"}, \RejectReason\": \"\", \TrainRecord\": [{\StartTime\": \"2023-11-01 23:00:17\", \Record\": {\url1\", \url2\"}}]}</pre>

CustomerConfirmContent

参数	类型	必须	说明
TrainResult	Array of [string]	是	供客户进行效果确认的 Demo 视频 url 地址
RejectReason	string	是	数智人侧不接受重训的理由
TrainRecord	Array of [TrainRecord]	否	历史训练的 Demo 视频清单

TrainRecord

参数	类型	必须	说明
----	----	----	----

StartTime	string	是	历史训练时间
Record	Array of [string]	是	历史训练的 Demo 视频 url 地址

AssetInfo

参数	类型	必须	说明
VirtualmanTypeCode	string	否	数智人类型 Code，参考 附录3-数智人类型 ，形象定制返回
VirtualmanKey	string	否	数智人 virtualmankey，形象定制返回
Resolution	string	否	数智人分辨率，形象定制返回
OriginZoom	string	否	缩放系数，保留三位小数，形象定制返回
VirtualmanResourceId	int32	否	数智人资源 ID，形象定制返回
TimbreKey	string	否	音色key，声音复刻返回
TimbreSample	string	否	音色demo链接，声音复刻返回
MakeType	string	是	定制类型： IMAGE_PHOTO: 2D小样本（照片数智人）形象定制 ZERO_SHOT_VOICE: 声音复刻（极速版）

请求示例

```
{
  "Header": {},
  "Payload": {
    "TaskId": "1212_29934_w"
  }
}
```

```
}  
}
```

返回示例

```
{  
  "Header": {  
    "RequestID": "gz1a74c25f17030396653416949",  
    "SessionID": "gz1a74c25f17030396653416950",  
    "DialogID": "",  
    "Code": 0,  
    "Message": "ok"  
  },  
  "Payload": {  
    "StageInfo": "{\\"TrainResult\\":  
[\\"url1\\",\\"url2\\"],\\"RejectReason\\":\\"\\",\\"TrainRecord\\":  
[{\\"StartTime\\":\\"2023-11-01 23:00:17\\",\\"Record\\":  
[\\"url1\\",\\"url2\\"]}]}",  
    "Status": "MAKING",  
    "StatusV2": "WAIT_PROCESS",  
    "FailMessage": "",  
    "StageV2": "CONFIRM"  
  }  
}
```

效果确认接口

最近更新时间：2024-11-13 15:22:12

当通过进度查询接口查到处于待确认效果阶段时（Status:MAKING，StageV2:CONFIRM，StatusV2:WAIT_PROCESS），对 demo 视频进行效果确认。

调用协议

HTTPS + JSON

POST /v2/ivh/assetmanager/customservice/customerconfirm

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

请求参数

参数	类型	必须	说明
TaskId	string	是	定制接口 返回的 TaskId
Operate	string	是	操作： AGREE：认同 REJECT：驳回
Reason	string	否	驳回原因，当 Operate:REJECT 时必填，最多300字

返回参数

请求示例

```
{
  "Header": {},
  "Payload": {
    "TaskId": "1212_29934_w",
    "Operate": "AGREE"
  }
}
```

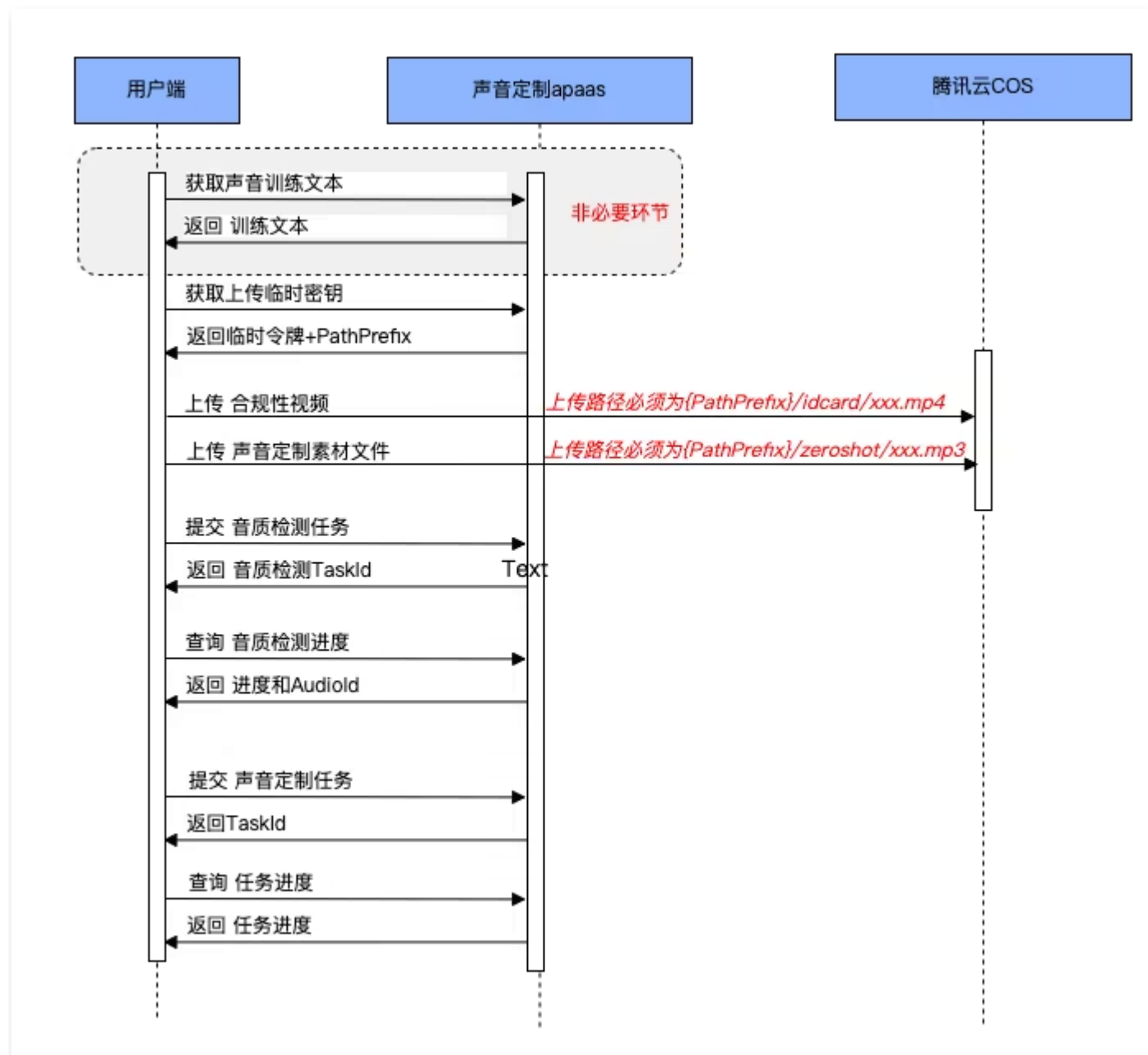
返回示例

```
{
  "Header": {
    "RequestID": "gz1a74c25f17030396653416949",
  }
}
```

```
"SessionID": "gz1a74c25f17030396653416950",  
"DialogID": "",  
"Code": 0,  
"Message": "ok"  
},  
"Payload": {}  
}
```

声音复刻（极速版） 接口调用逻辑示意图

最近更新时间：2025-03-14 16:06:02



1. 初始化阶段（可选）

- 用户可获取声音训练文本，系统返回训练文本。

2. 文件上传准备

- 用户获取上传临时密钥，系统返回临时令牌 + PathPrefix。

3. 文件上传规则

- **合规性视频**: 必须上传至路径 `{PathPrefix}/idcard/xxx.mp4` 。
- **声音定制素材文件**: 必须上传至路径 `{PathPrefix}/zeroshot/xxx.mp3` 。

4. 音质检测任务

- 用户提交音质检测任务，系统返回**音质检测 TaskId**。
- 用户可查询检测进度，系统返回**进度**和 **Audioid**。

5. 声音定制任务

- 用户提交声音定制任务，系统返回 **TaskId**。
- 用户可查询任务进度，系统返回**任务进度**。

获取声音训练文本

最近更新时间：2024-11-29 17:15:42

接口说明

调用该接口获取需要朗读的文档内容，然后使用该文本内容诵读成音频文件当作声音制作素材。

调用协议

HTTPS + JSON

POST /v2/ivh/assetmanager/zeroshotservice/gettext

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

请求参数

参数	类型	必须	说明
Language	String	否	获取对应语言的文本：chi-中文普通话，eng-英文，can-粤语

返回参数

参数	类型	必须	说明
Texts	Array of [Text]	是	可选择的文本列表

Text

参数	类型	必须	说明
TextId	string	是	文本ID
Text	string	是	文本

请求示例

```
{
  "Header": {},
  "Payload": {
    "Language": "chi"
  }
}
```

```
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "DialogID": "",
    "Message": "",
    "SessionID": "gz92439a3317328659145933517",
    "RequestID": "9e2bf1f4dafe4250409132f78defcaea"
  },
  "Payload": {
    "Texts": [
      {
        "TextId": "ec28499a-cf6d-4f7b-b7fa-4a5a54662b38",
        "Text": "本次降价的300多款产品 其中肉类零食占比较多一些"
      }
    ]
  }
}
```

获取上传临时令牌

最近更新时间：2024-11-18 17:24:52

在调用 [定制接口](#) 前，需要将素材文件上传到指定的云端存储。上传时，首先通过该接口获取上传临时令牌和指定的上传路径，再通过 [上传素材到腾讯云 COS](#) 接口上传对应的文件到云端存储。

调用协议

HTTPS + JSON

POST /v2/ivh/assetmanager/customservice/getuploadcredentials

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

请求参数

无

返回参数

参数	类型	必须	说明
Credentials	object	是	临时密钥信息
Credentials.Token	string	是	临时密钥令牌
Credentials.TmpSecretId	string	是	临时证书密钥 ID
Credentials.TmpSecretKey	object	是	临时证书密钥 Key
StartTime	int	是	临时证书有效起始时间，返回 Unix 时间戳，精确到秒
ExpiredTime	int	是	临时证书有效截止时间，返回 Unix 时间戳，精确到秒
PathPrefix	string	是	上传到 COS 上的路径前缀

返回示例

```
{  
  "Header": {
```

```
"RequestID": "c18ae4e49d89e408231c5b3ceea23c",
"SessionID": "gzfa9105ca17310669340273673",
"Code": 0,
"Message": "ok"
},
"Payload": {
  "Credentials": {
    "TmpSecretId": "AKID*****ib",
    "TmpSecretKey": "DVki*****k0=",
    "Token":
"8oz1l*****49esQlG0940LX2tacXiCv4YS2qdBsiQ"
  },
  "ExpiredTime": 1731074134,
  "PathPrefix": "https://cos-test.ap-
guangzhou.myqcloud.com/customer-pipeline/123/4506c29b-2904-45e2-b7d7-
3052fe2493c2/",
  "StartTime": 1731066934
}
}
```



```

"Payload": {
  "Credentials": {
    "TmpSecretId": "AKIDM6p_7lbof_ow5_DTYz...Qzb1M1vtFfjxqbg1FrN9",
    "TmpSecretKey": "9pkrcGy6duNWhj4/...J87PGk=",
    "Token": "fGSqNKc3Q90WcLEz5jn2pdHo3cGEuPKa23738db8c693d875f4e2ead9f6ff446w9B4mmLxcks10FXXJAW-jtudiJnfdPht0RvkEC3PshPrv2nrXPLjwTst8A-mGxVxkpM-YIE_RpeuQG2nIKF9izYreejnrzx6oCuCL61AXy39v6CdCHXoWkdWkZyc1LWio2QVMrshS0uKr26H770MmVdjH9Evk19I9V8K4KR7mJaE3GDisk5tJpAqFtd5tSRGpSiERdGuf2jUef9WvnWuV1yysSpJy7qr0KqQhj7GwzWkv4j2jtr1B6fbNzYfTxPHvjlWDC50Dg613pgiZqwtu7GtoUxdFOXWQny1aHiiz1NnRR25ELot_o7tAsdL6Ae7uj-31bkyQbU88e5JHSRvS4p0QVe58LUNw5WhpTeJlXh-LRsPmc2z29MmXtSvWkgx2Y-KiQBYHsF81paEMuHK4eU60303xdjd5_h6N4C0u6QFet3PUJY7Kgt1t7L4K0670QJERrikEgBYppyZYqvVHaD4YZO-TABrhymYPUVkdmbzW4Idim8tZUI6Fc93FGyYRAFfql3zux4IX-NrieAGMRifks1q1XfRfHf02iQeaLyUubddGoyVvK6y5EiJkou5mM0drPBpPzYoZJWalHT12mbm-hDKHvcgBe-bMe7qs7ze6NiyqwrECDCrVmdTA8j0EVKPGuE94xAnMyytYHnhmQvC8o77hBuQvqgcCPiX-R_xAUD-6ydv5Gi_8mLcd07h35Nvk0TrXyNuMTmqjTEbo1sVC74jP68pQMzNy--UU1Z9C_tI6Y938xwf5xPB70MzZMTP7IEnMPqsDZgGwWg1R-U-JbMzq-HkDBU...B0d6Qf3XN2PCL1q1gkJuS8jZBLQ4K4c5XVvfbEpSw4TcLn500yqp6ks7cEXBZeAhAoWRJ_APBDNTI72-Xn1ydgBbcjY6G3xXuXDRbBGUs8QrCi_3mMYNMyNq4fLbAnR1sdAnjFKV7UNQ0FMsEFWpKyLD4cs4bEbhbtMt-lu37oVXD4EUcnNyR6Yz1jNws27n8gAL0I9FDEb-aHICXaYyG5cEgELghf16gJsomjwZa0FpXtI4RMPBP_6D0f6RGE0IF_0r5Awh_9_QsGwIXkUW18aeVw7C0fexQ8Cf0vwU-zk0q2F4x9CpQudsN18FIYC3I3sYTViS5gIh2P_-L0dkzwjplz5-yRglbdjH5ba1jXm1GqhipeohMsYxQSHDfgkAm1WFwqpp4vkt-0S-1kvY94BqTRkm8M2_0vornZIDx8n3DZJiI657HDL33d4cGncSv0Q_R49SRiNK471Cne-NDoqRmCBvEXY-mgDS7Fht2F5vKxueYvX0iRnmP4_d75qDpYn1_PNYCoNj4ViYSICPXJUNeofRTJc2miwSENL7vH0H0kM02hNikB46LhLPZPtK51sd4g1JozesLRcbD0sAZyJd_5e1pE7SaulMdm7FMkQq3th0CwacrIy062DncwEQBBM-00uwr7CpBdhakLAXn-RAO-C12ogMw-Wx1BQ57ixN7yjE28ukSzG8QjC14_QUQsG2Y8FKQHVY5hZp1AGhjpBCN_7XzH-WAPGfzrGXJW1DT-zkHJYRPMbB19zx2WcIfCyChtbz7XReQCN6BmWmXvtqApjKAdr9ZQctTk0aNdvm53zz1jxdoHttAkJL9YCPp0iN_2YjHI5t4e7PMc_Itsxg2RgDy83EL8Iz0lqs57WtRErFQspFgtAWr3dmjgTYtHehPIP_K42k5WA2t6g3031z_3_zaGMTzngv4u0PTvYz-QpfqAQMDQPaRL132QTKarriImA8Ff8d10DhjuvBDPzIwzA6085txTJMhJbK1DZwDpw3KLJUiqAgfWcNhoHJmrdmJo_AvYz9yUaqHqH0AIEUtYlZsh3emcD6vb1V2kZkCci3GjzrTsg"
  }
},
"ExpiredTime": 1736241459,
"PathPrefix": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipeline/352/59c6f6cd-8f92-45f9-8352-64ca8fb92eaa/",
"StartTime": 1736234259

```

- bucket名称: 解析出来的值为 virtualhuman-cos-prod-1251316161。
- region名称: 解析出来的值为 accelerate。
- key(bucket 下的存储路径): 解析出来的值为 customer-pipeline/352/59c6f6cd-8f92-45f9-8352-64ca8fb92eaa/。

注意:

解析出来的值为 key 的前半部分，还需要根据上传者类型来拼接后半部分，如上传的是 pdf 授权书素材，则需要拼接上 idcard/xxx.pdf，如下图所示，具体后半部分值的定义请参考定制接口中的相关参数说明。

腾讯云智能数智人	参数名称	数据类型	是否必填	说明	备注
定制接口	IdentityWrittenCosUrl	string	否	pdf 格式授权书 URL，要求： 1. url 地址为通过上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名/customer-pipeline/(数字)/(uuid)/idcard/b.pdf 2. 该格式主要适用于“书面版授权书”，以完整清晰的扫描件形式提交 备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与 IdentityCosUrl 二选一或二者都填	本项目目录： 调用协议 请求参数 返回参数 请求示例 返回示例
定制接口	MaterialCosUrl	string	否	形象定制素材 url 地址要求： 1. url 地址为通过上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 video 路径，例如/customer-pipeline/(数字)/(uuid)/video/c.mp4 2. 视频大小，不超过5GB，4K不超过10G 3. 视频时长：专属口型版本定制时长为2-10分钟；通用口型版本定制时长为1-10分钟；高精版本定制时长为2-10分钟 4. 视频分辨率：1080P、4K (3840*2160)，如定制高精版，则必须为4K 5. 视频长宽比：16:9 (9:16) 6. 视频帧率：不低于25fps，不超过60fps 7. 视频格式：mp4, mov 备注：形象定制时 video 路径必传 声音定制素材 url 地址要求： 1. url 地址为通过上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如/customer-pipeline/(数字)/(uuid)/audio/c.zip 2. 压缩包格式：.zip 格式的压缩包，单个 zip 包用于定制1个音色；压缩时不要新建文件夹，直接选中所有 wav 压缩 3. 对于单个 zip 包里的音频注意事项： ①音频数量：每个 zip 包里可以包含1条或多条 wav 格式音频，总量不超过10条 ②音频大小：每个 zip 包里的音频大小总和不超过1G ③音频格式：每条音频需为 wav 格式，其他格式的音频请先转为 wav 格式后再再压缩成 zip 包 ④音频采样率：采样率大于等于24k即可，建议24k、36k	

提交音质检测任务

最近更新时间：2025-12-19 14:24:31

朗读 [获取声音训练文本](#) 内容，录制好音频文件上传到指定 COS 路径后，需要进行音频质量检测，检测通过的音频会返回 Audioid，使用该 Audioid 调用定制接口进行一句话声音定制。

支持的音质指标：

- 信噪比指数：大于等于20合格
- 混响指数：大于等于20合格
- 截幅指数：小于等于10合格
- 有效时长：大于等于5秒合格
- 文本准确率：
 - 中文普通话：大于等于70%合格
 - 英文：大于等于50%合格
 - 粤语：大于等于70%合格

调用协议

HTTPS + JSON

POST /v2/ivh/assetmanager/zeroshotservice/detectaudioquality

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

请求参数

参数	类型	必须	说明
TextId	String	否	获取声音训练文本 接口返回的 TextId
ReferenceText	String	否	音频参考文本，ReferenceText 和 TextId 为两种不同方式获取声音训练文本，可以都不填或二选一，能获得更加准确音质检测结果
AudioUrl	String	是	语音文件 URL 地址，要求： <ol style="list-style-type: none">1. URL 地址为通过4.1上传到指定路径的资源 url 地址后增加 zeroshot 路径，例如 /customer-pipeline/{数字}/{uuid}/zeroshot/xx.x.mp32. 建议上传音频时长为10-20S，单个文件不超过20MB3. 音频格式支持 wav、mp3、m4a、wma、asf 和 aac，采样率需大于等于16K；对于压缩格式，建议码率大于等于128kbps4. 音频命名为2-50个字符，只允许有汉字、字母、数字、下划线、横线

Language	String	否	语言类型。chi-中文普通话，eng-英文，can-粤语。默认chi
----------	--------	---	------------------------------------

返回参数

参数	类型	必须	说明
TaskId	String	是	制作的任务 ID，携带 TaskId 访问 查询音质检测任务进度 ，可获得检测进度和检测结果

请求示例

```
{
  "Header": {},
  "Payload": {
    "AudioUrl": "https://virtualhuman-cos-prod-1251316161.cos.accelerate.myqcloud.com/customer-pipeline/1322/c9d435f3-9dfb-42c8-9b39-c0117d3cd35c/zeroshot/90450181730994661.mp3"
  }
}
```

返回示例

```
{
  "Header": {
    "Code": 0,
    "DialogID": "",
    "Message": "ok",
    "RequestID": "c9d435f3-9dfb-42c8-9b39-c0117d3cd35c"
  },
  "Payload": {
    "TaskId": "6da6f935-de23-4337-b753-64cca3627362"
  }
}
```

查询音质检测任务进度

最近更新时间：2024-11-13 15:22:12

通过 TaskId 来查询音质检测的进度和结果。

调用协议

HTTPS + JSON

POST /v2/ivh/assetmanager/zeroshotservice/getdetectprogress

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

请求参数

参数	类型	必须	说明
TaskId	string	是	提交音质检测任务 返回的 TaskId

返回参数

参数	类型	必须	说明
Audioid	string	是	音频ID
DetectionCode	int	是	检测code: 0: 表示当前语音通过 -1: 表示检测失败, 需要重试 -2: 表示语音检测不通过, 提示用户再重新录制一下 (通常漏读, 错读, 或多读) -3: 表示语音中噪声较大, 不通过
DetectionMsg	string	否	检测提示信息
TaskState	string	否	检测任务状态: PROCESSING: 处理中 SUCCESS: 任务处理完成 FAIL: 任务处理失败

请求示例

```
{
  "Header": {},
  "Payload": {
    "TaskId": "c9d435f3-9dfb-42c8-9b39-c0117d3cd35c"
  }
}
```

返回示例

```
{
  "Header": {
    "RequestID": "gz1a74c25f17030396653416949",
    "SessionID": "gz1a74c25f17030396653416950",
    "DialogID": "",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "TaskState": "SUCCESS",
    "AudioId": "17xsjw8-9dfb-42c8-9b39-sjdue972hds",
    "DetectionCode": 0,
    "DetectionMsg": "success"
  }
}
```

定制接口

最近更新时间：2025-12-19 14:24:31

使用该接口提交定制的请求，通过 [进度查询接口](#) 查询定制的各个阶段以及各阶段的相关信息。

调用协议

HTTPS + JSON

POST /v2/ivh/assetmanager/customservice/make

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

请求参数

参数	类型	必须	说明
AnchorName	string	是	<p>主播名称：</p> <ol style="list-style-type: none">1. 该名称主要用于标识该定制形象/声音，可以根据实际情况自定义2. 命名参考：如果该主播定制的数量只有1个，可以直接以主播名称命名，例如“张三”；也可加上衣服名称来更好的标识，例如“张三蓝西服”等3. 不超过50个字符，不少于2个字符，只允许有汉字、字母、数字、下划线、横线4. 不能重名
MakeType	string	是	<p>定制类别：</p> <p>IMAGE：2D小样本-标准版（专属口型）形象定制</p> <p>IMAGE_GENERAL：2D小样本-标准版（通用口型）形象定制</p> <p>IMAGE_4K：2D小样本-高精版（专属口型）形象定制</p> <p>IMAGE_PHOTO：2D小样本（照片数智人）形象定制</p> <div style="border: 1px solid #00aaff; padding: 5px; margin: 10px 0;"><p>注意：选择 IMAGE_PHOTO 类型时，强烈建议预先执行 人脸检测任务创建 以提高成功率</p></div> <p>VOICE：声音复刻（基础版）</p> <p>ZERO_SHOT_VOICE：声音复刻（极速版）</p>
IdentityCosUrl	string	否	<p>视频格式授权书 URL，要求：</p> <ol style="list-style-type: none">1. url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名/customer-pipeline/{数字}/{uuid}/idcard/a.mp42. 该格式主要适用于“口述版授权书”，也可将“书面版授权书”拍成完整清晰的视频形式提交

			<p>备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与IdentityWrittenCosUrl 二选一或二者都填</p>
IdentityWrittenCosUrl	string	否	<p>pdf 格式授权书 URL，要求：</p> <ol style="list-style-type: none"> url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 idcard 路径，例如：域名/customer-pipeline/{数字}/{uuid}/idcard/b.pdf 该格式主要适用于“书面版授权书”，以完整清晰的扫描件形式提交 <p>备注：除 IMAGE_PHOTO、ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型与IdentityCosUrl 二选一或二者都填</p>
MaterialCosUrl	string	否	<p>形象定制素材 url 地址要求：</p> <ol style="list-style-type: none"> url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源url 地址后增加 video 路径，例如/customer-pipeline/{数字}/{uuid}/video/c.mp4 视频大小，不超过5GB，4K不超过10GB 视频时长：专属口型版本定制时长为2-10分钟；通用口型版本定制时长为1-10分钟；高精版本定制时长为2-10分钟 视频分辨率：1080P、4K（3840*2160），如定制高精版，则必须为4K 视频长宽比：16:9（9:16） 视频帧率：不低于25fps，不超过60fps 视频格式：mp4，mov <p>声音定制素材 url 地址要求：</p> <ol style="list-style-type: none"> url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如/customer-pipeline/{数字}/{uuid}/audio/c.zip 支持音频文件或压缩包。压缩包格式：.zip 格式的压缩包，单个 zip 包用于定制1个音色 音频注意事项： <ol style="list-style-type: none"> ①音频数量：每个 zip 包里可以包含1条或多条音频，建议总量不超过10条 ②音频大小：每个 zip 包里的音频大小总和不超过1G ③音频格式：支持 wav、mp3、m4a、aac、wma 和 asf 格式 ④音频采样率：采样率大于等于24k即可，建议24k、36k ⑤音频命名：名称不要有空格或者特殊字符 <p>备注：除 ZERO_SHOT_VOICE 定制类型无需填写外，其他定制类型必填</p> <p>照片数智人定制素材 url 地址要求：</p>

			<ol style="list-style-type: none"> url地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 photo 路径，例如/customer-pipeline/{数字}/{uuid}/photo/example.png 图片名称：不少于2个字符，只允许有汉字、字母、数字、下划、横线；图片格式：支持jpg、jpeg、png、webp；图片大小：不超过16M；图片宽高比：支持 1:1 9:16 16:9 4:3 人物清晰正面照，脸部居中，表情自然，嘴部闭合
SexType	string	是	性别： MALE：男 FEMALE：女
Notes	string	否	定制备注信息，100字以内
TextDriver	string	否	用于生成驱动 demo 的文本内容，字符长度在4-1000之间（包括 ssm1 标签,单个中文算一个字符）
VoiceDriver CosFile	string	否	生成驱动 demo 的音频文件路径，要求： <ol style="list-style-type: none"> url 地址为通过 上传素材到腾讯云 COS 上传到指定路径的资源 url 地址后增加 audio 路径，例如:/customer-pipeline/{数字}/{uuid}/audio/example.wav 音频文件大小不超过10M，格式支持wav、mp3、wma、m4a、aac
Audiold	string	否	ZERO_SHOT_VOICE 定制类型必填：填写 查询音质检测任务进度 通过后返回的 Audiold
TimbreInfo	Timbre Template	否	形象定制时若需要指定音色，或尝试用视频里的声音快速复刻，则填充该字段。
IsHaveBackground	bool	否	形象定制类型，训练出来的形象是否带原始背景。默认“否”，即不带原始背景，在应用过程中可以按需更换背景 true：保留拍摄背景 false：抠除拍摄背景（支持更换背景，仅支持绿幕视频和白墙视频）
RemoveBackground	int	否	当 IsHaveBackground=false 时，进一步填充该参数。默认值是 1。 1：抠除拍摄背景-绿幕视频 2：抠除拍摄背景-白幕视频

AlphaCosUrl	string	否	自定义 Alpha 通道视频地址，视频上传路径参照 形象定制素材 路径
ZeroshortLang	string	否	ZERO_SHOT_VOICE 定制类型的语言类型，chi-中文普通话，eng-英文，can-粤语，默认chi
PhotoVersion	int	否	默认0 0: 快速版。在10分钟内即可完成数智人定制，只支持头部轻微晃动，面部特征稳定。建议执行 人脸检测任务 (CheckerVersion=1) 1: 优享版。在1个小时内可完成数智人定制，头部和身体都支持轻微晃动。建议执行 人脸检测任务 (CheckerVersion=2)。

TimbreTemplate 信息参数

参数	类型	必须	说明
TimbreCode	string	否	制作文本驱动 demo 时指定的音色 code
Speed	string	否	制作文本驱动 demo 时指定音色的语速，默认1.0。语速（范围[0.5-1.5]，值为0.5时播报语速最慢，值为1.5时播报语速最快）
IsWithTimbre	bool	否	尝试用视频里的声音快速复刻，额外生成 demo（由于视频里的音质质量不能保证，可能会存在快速复刻失败的情况）。该 demo 不影响其他 demo 的正常合成。

返回参数

参数	类型	必须	说明
TaskId	string	是	制作的任务 ID，携带 TaskId 访问 进度查询接口 ，可获得制作进度和制作结果

请求示例

```

{
  "Header": {},
  "Payload": {
    "AnchorName": "马老师",
    "MakeType": "ZERO_SHOT_VOICE",
  }
}

```

```
"AudioId": "audiof1d85a5c766e47c690ff61cbf43d5364",
"IdentityCosUrl": "https://virtualhuman-cos-prod-
1251316161.cos.accelerate.myqcloud.com/customer-pipeline/100/f30e4d52-
00ba-4e36-8a26-26c348bf786e/idcard/identity.mp4",
"MaterialCosUrl": "https://virtualhuman-cos-prod-
1251316161.cos.accelerate.myqcloud.com/customer-pipeline/100/78675441-
08bf-43f2-93db-f3936068bce7/video/test.mp4",
"SexType": "FEMALE"
}
}
```

返回示例

```
{
  "Header": {
    "RequestID": "c18aeee4e49d89e408231c5b3ceea23c",
    "SessionID": "gzfa9105ca17310669340273673",
    "Code": 0,
    "Message": "ok"
  },
  "Payload": {
    "TaskId": "1212_29934_w"
  }
}
```

进度查询接口

最近更新时间：2025-03-07 12:14:53

通过TaskId来查询任务的制作进度和结果。

调用协议

HTTPS + JSON

POST /v2/ivh/assetmanager/customservice/getprogress

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

请求参数

参数	类型	必须	说明
TaskId	string	是	定制接口 返回的 TaskId

返回参数

参数	类型	必须	说明
Status	string	是	任务制作状态： MAKING：进行中 SUCCESS：制作成功 FAIL：制作失败
FailMessage	string	否	当 Status: FAIL 时，返回制作失败原因，便于排查问题
Stage	string	否	当前定制阶段（当前字段是为了兼容历史定制任务，新定制可忽略，使用 StageV2）
StageId	string	否	定制阶段 Id（当前字段是为了兼容历史定制任务，新定制可忽略，使用 StageV2）
StageV2	string	否	（自2024.1.1之后发起的进度查询请求，以当前字段为准） 当 Status: MAKING 时，返回任务进行中的具体阶段： SUBMIT：素材提交 AUDIT：人工审核 PROCESS：预处理（仅形象定制有） TRAIN：智能训练 CONFIRM：效果确认（照片定制无此阶段） END：完成

StatusV2	string	否	当 StageV2 处于某个阶段时，该阶段的具体状态： WAIT_PROCESS: 待处理 PROCESSING: 处理中 SUCCESS: 成功 FAIL: 失败
StageInfo	string	否	处于某个阶段的额外信息，JSON 格式的字符串，不同阶段的 JSON 格式类型见下述表格，目前只有效果确认阶段（StageV2 为 CONFIRM）有值。
AssetList	Array of [AssetInfo]	否	返回定制的形象或音色信息，目前支持返回的定制类型有： IMAGE: 2D小样本-标准版（专属口型）形象定制 IMAGE_GENERAL: 2D小样本-标准版（通用口型）形象定制 IMAGE_4K: 2D小样本-高精版（专属口型）形象定制 IMAGE_PHOTO: 2D小样本（照片数智人）形象定制 ZERO_SHOT_VOICE: 声音复刻（极速版） VOICE: 声音复刻（基础版）

不同阶段 StageInfo 额外信息见下表：

StageV2	类型	描述	示例
CONFIRM: 效果确认阶段	string of [CustomerConfirmContent]	当阶段处于效果确认阶段（StageV2: CONFIRM），且状态处于待处理中时（StatusV2: WAIT_PROCESS），返回待确认效果 Demo 视频，客户可以根据返回的 Demo 视频通过调用 效果确认接口 对 Demo 视频效果进行认同/驳回操作	<pre>{\TrainResult\":{\url1\", \url2\"}, \RejectReason\": \"\", \TrainRecord\": [{\StartTime\": \"2023-11-01 23:00:17\", \Record\": {\url1\", \url2\"}}]}</pre>

CustomerConfirmContent

参数	类型	必须	说明
TrainResult	Array of [string]	是	供客户进行效果确认的 Demo 视频 url 地址
RejectReason	string	是	数智人侧不接受重训的理由
TrainRecord	Array of [TrainRecord]	否	历史训练的 Demo 视频清单

TrainRecord

参数	类型	必须	说明
----	----	----	----

StartTime	string	是	历史训练时间
Record	Array of [string]	是	历史训练的 Demo 视频 url 地址

AssetInfo

参数	类型	必须	说明
VirtualmanTypeCode	string	否	数智人类型 Code，参考 附录3-数智人类型 ，形象定制返回
VirtualmanKey	string	否	数智人 virtualmankey，形象定制返回
Resolution	string	否	数智人分辨率，形象定制返回
OriginZoom	string	否	缩放系数，保留三位小数，形象定制返回
VirtualmanResourceId	int32	否	数智人资源 ID，形象定制返回
TimbreKey	string	否	音色key，声音复刻返回
TimbreSample	string	否	音色demo链接，声音复刻返回
MakeType	string	是	定制类型： IMAGE_PHOTO: 2D小样本（照片数智人）形象定制 ZERO_SHOT_VOICE: 声音复刻（极速版）

请求示例

```
{
  "Header": {},
  "Payload": {
    "TaskId": "1212_29934_w"
  }
}
```

```
}  
}
```

返回示例

```
{  
  "Header": {  
    "RequestID": "gz1a74c25f17030396653416949",  
    "SessionID": "gz1a74c25f17030396653416950",  
    "DialogID": "",  
    "Code": 0,  
    "Message": "ok"  
  },  
  "Payload": {  
    "StageInfo": "{\\"TrainResult\\":  
[\\"url1\\",\\"url2\\"],\\"RejectReason\\":\\"\\",\\"TrainRecord\\":  
[{\\"StartTime\\":\\"2023-11-01 23:00:17\\",\\"Record\\":  
[\\"url1\\",\\"url2\\"]}]}",  
    "Status": "MAKING",  
    "StatusV2": "WAIT_PROCESS",  
    "FailMessage": "",  
    "StageV2": "CONFIRM"  
  }  
}
```

结果码说明


最近更新时间：2024-08-23 10:39:03

代码	含义	措施
0	成功	无
100000	系统内部错误	重试或联系管理员定位错误
100001/100003	请求参数错误（包括参数格式、类型等错误，缺少参数错误，必传参数没填）	对照 API 协议核查参数格式、类型
100002	参数取值错误	对照 API 协议核查参数列表
100005	未授权操作	检查对请求的会话是否有操作权限
110006	记录不存在	检查 taskid 是否正确
100008	超过配额限制	分配的配额已经用完

接口对接常见问题

最近更新时间：2024-08-23 17:05:43

序号	问题描述	解决方式	备注
1	调用接口时出现鉴权不通过	<ol style="list-style-type: none"> 1. 确认签名时参数按照字典序排列了 2. 确认时间戳参数为当前时间 3. 参考右侧的示例参数，按照自己的签名步骤计算一下，如果计算结果与示例一致则表示签名计算步骤无误。 	<p>参数：</p> <pre>appkey = e38267c0e86411ebb02aed82acb0ed99 accesstoken = f68f2d10ae9e4604b76fb05cf46bccec timestamp=1646636485</pre> <p>签名结果：</p> <pre>BfWuaC9kmaicCggXc693uK%2BsZQ8qe88O4HVQNTdwZuo%3D</pre>
2	调用接口出现"miss Content Length"错误	<ol style="list-style-type: none"> 1. 请确认请求的 body 是否加了 header 参数和 payload 参数，不加的话会报错。如右图所示。 2. 参数的类型要严格按照接口文档中填充。 3. 检查 url 是否拼写正确 	
3	UnauthorizedOperation: 未授权操作: 未给该 appkey 创建项目	未给客户开通小样本形象定制接口的使用权限，请联系腾讯相关人员开通	
4	视频素材是否需要上传到客户的腾讯云 COS 上	不需要。通过临时令牌上传会上传到数智人内部的 COS 上，不需要客户额外开通腾讯云 COS	

<p>5</p>	<p>调用 cos 的 sdk 时的 Bucket、Region、Key 如何获取</p>	<p>可以从 prefix 中解析出来，如右图所示</p>	
----------	---	-------------------------------	--

2D 端渲染形象定制及下载流程

最近更新时间：2026-03-23 18:22:42

前言

2D 端渲染形象是指借助用户本地设备（如 PC、手机、智能硬件等），通过设备自身的 CPU 完成渲染计算，并在本地实时显示的一类形象。数智人的端渲染形象训练一次可以产生多个分辨率的形象，其中 720P 较为特殊，除了可以作为端渲染形象使用外，还可以通过云渲染的方式使用，为了区分之前依赖 GPU 进行使用的云渲染形象，我们称此类 720P 形象为轻量版云渲染形象。（可以理解为它有两个名字：2D 端渲染形象或 2D 轻量版云渲染形象）

本文将从头介绍 2D 端渲染形象的下单、定制、使用等流程。

下单流程

下单流程主要为购买形象配额，操作步骤如下：

1. 登录 [数智人购买页](#)。
2. 选择并购买形象配额。您可以选择购买 2D 小样本通用口型、2D 小样本专属口型、2D 小样本照片三种类型的“形象定制”配额，如下图所示。

⚠ 注意：

此三类形象定制配额可以同时用于云渲染形象和端渲染形象定制流程。由于训练流程互斥，如同时需要云渲染和端渲染形象，则需要分两次提交训练。后面流程会详述。

云智能数智人 | [返回产品详情](#)

- 2D形象 3D形象 声音复刻 更多服务

购买须知

- 计费概述** 形象购买不可单独使用，需要加购配合服务资产一同使用；具体可参见腾讯云智能数智人的[计费概述](#)
- 使用说明** 购买完成后资源为生效状态，可前往[数智人平台](#)或者通过调用API使用该产品。该产品仅限在中国境内使用。
- 退费说明** 购买前请详细了解[退费说明](#) 腾讯云有权根据退费说明拒绝相应退款要求

选择配置

形象类型

<p>2D小样本通用口型 通过一段真人视频素材进行训练数智人 ✓ 适用于对数智人唇齿特征无要求、无良好拍摄条件的客户。</p>	<p>2D小样本专属口型 通过一段真人视频素材进行训练数智人 ✓ 适用于对数智人形象复刻有要求、有较好拍摄条件的客户。</p>	<p>2D小样本高精 通过一段4K的真人视频素材进行训练数智人 ✓ 适用于大型会议、面对面对话、产品发布会、大屏场景。</p>
<p>2D小样本照片 通过一张照片即可训练数智人 ✓ 适用于泛互、娱乐场景。</p>	<p>2D精品 通过专业影棚中录制动作素材 ✓ 适用于金融、传媒类对数智人形象、动作有要求的客户。</p>	<p>2D小样本免训练(视频素材) 无需训练即可快速生成一个数智人视频 ✓ 适用于上传素材为视频的场景</p>
<p>2D小样本免训练(照片素材) 无需训练即可快速生成一个数智人视频 ✓ 适用于上传素材为照片的场景</p>	<p>2D端渲染 通过本地渲染呈现数字人效果，多场景通用 ✓ 适用于通用口型、专属口型、高精、照片等形象类型</p>	

[形象类型说明](#)

采购类型

- 形象租赁
 形象定制
 定制形象续期

选择数量

- 0 +

注意：1、形象定制配额自下单之日起有1年使用有效期，过期后配额自动失效；2、已定制完成的形象（2D小样本照片除外）自效果确认验收之日起有1年有效期，过期后可通过“定制形象续期”进行续费使用。3、2D小样本照片形象可永久使用。

定制流程

定制流程主要为消耗配额训练形象，操作步骤如下：

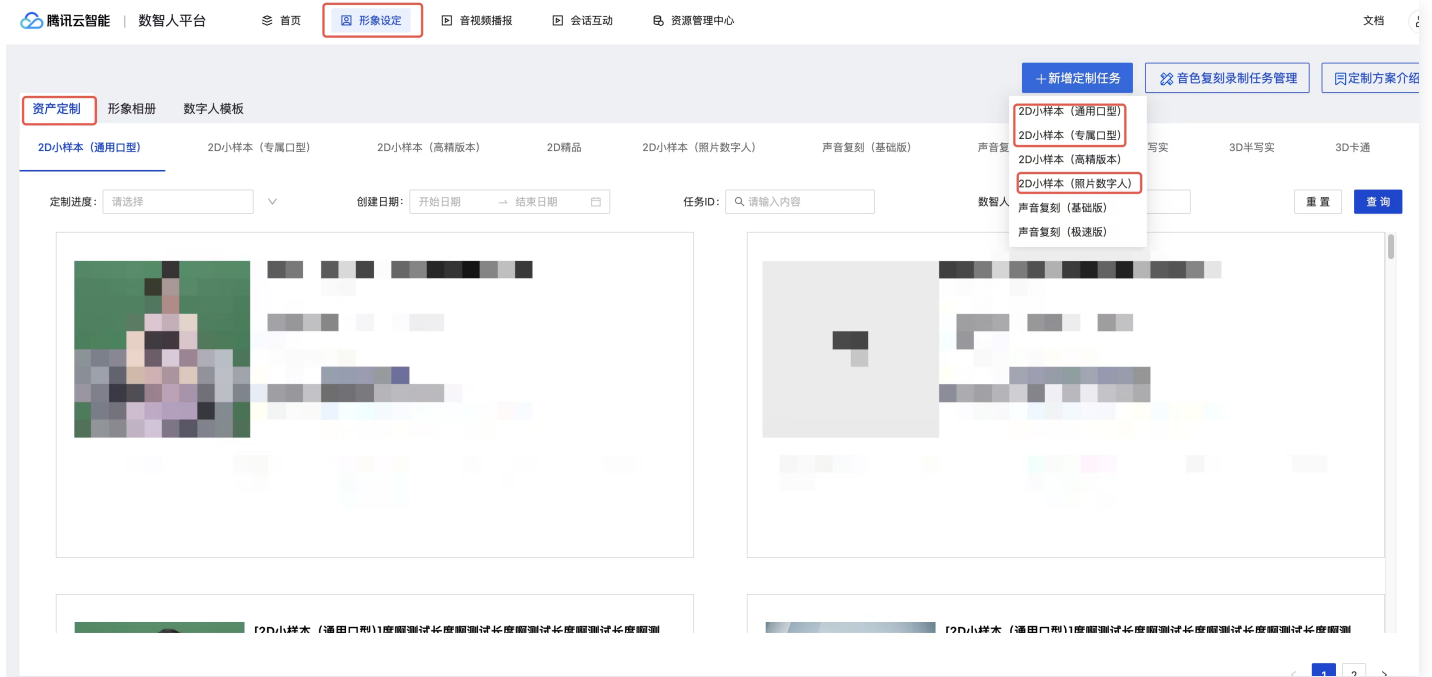
1. 登录 [数智人平台](#)。
2. 在**形象设定 > 资产定制**中，根据配额的使用情况进行端渲染形象的定制。

场景1：配额没有使用过，直接开始训练端渲染形象。

场景2：配额已经用于训练过云渲染形象，需要补训端渲染形象。

场景1：配额未使用过，直接开始训练端渲染形象。如果您所拥有的配额已经用于训练过云渲染形象，则跳过该步骤介绍，直接参考 **场景2**。

通过图示位置，选择对应已经拥有的定制配额，开启新增定制任务，如下图。



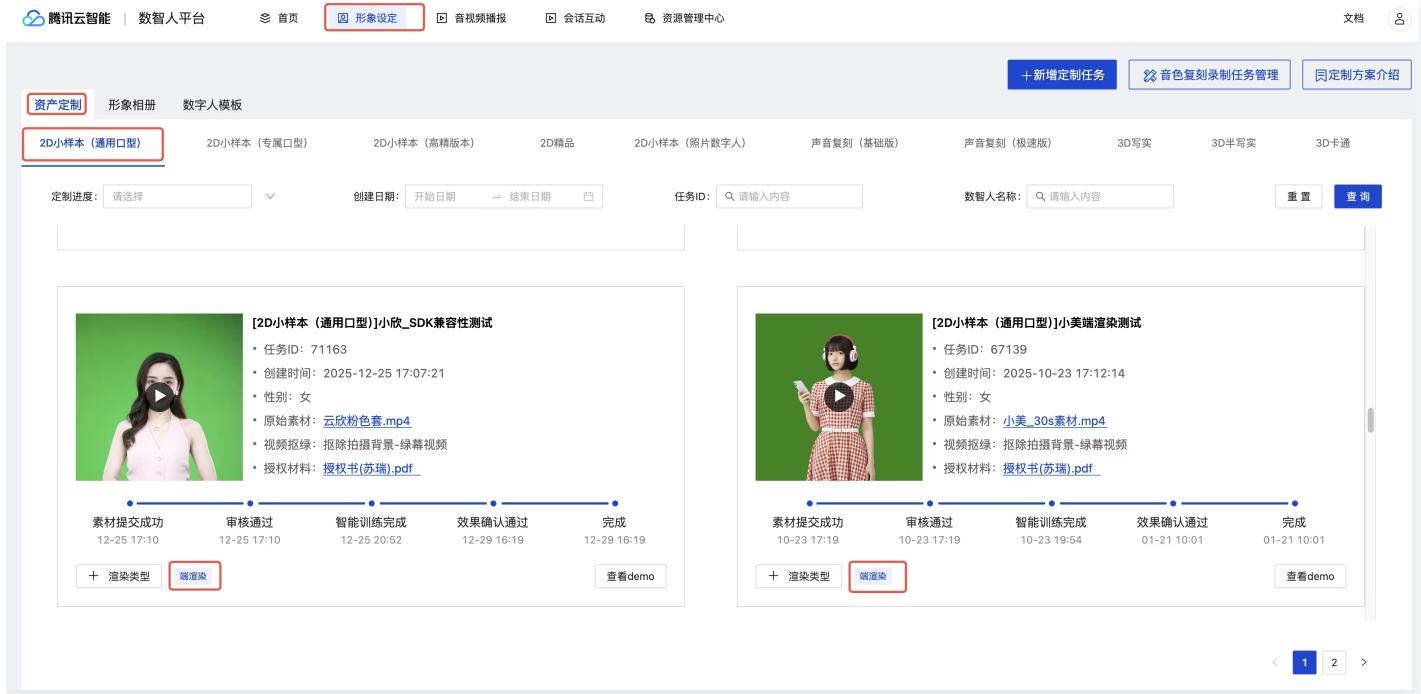
进入定制流程后，默认为“云渲染”，需要手动切换至“端渲染”。后续流程跟云渲染流程一致，上传素材，确认提交即可。



流程提交后，可以在资产定制中心查看定制进度，端渲染的形象，一般都标注有“端渲染”字样提示，如下图所示。训练确认完成后，即可在平台下载使用。

说明：

如果要补训云渲染形象，点击旁边的“渲染类型”选择“云渲染”即可触发补训流程。

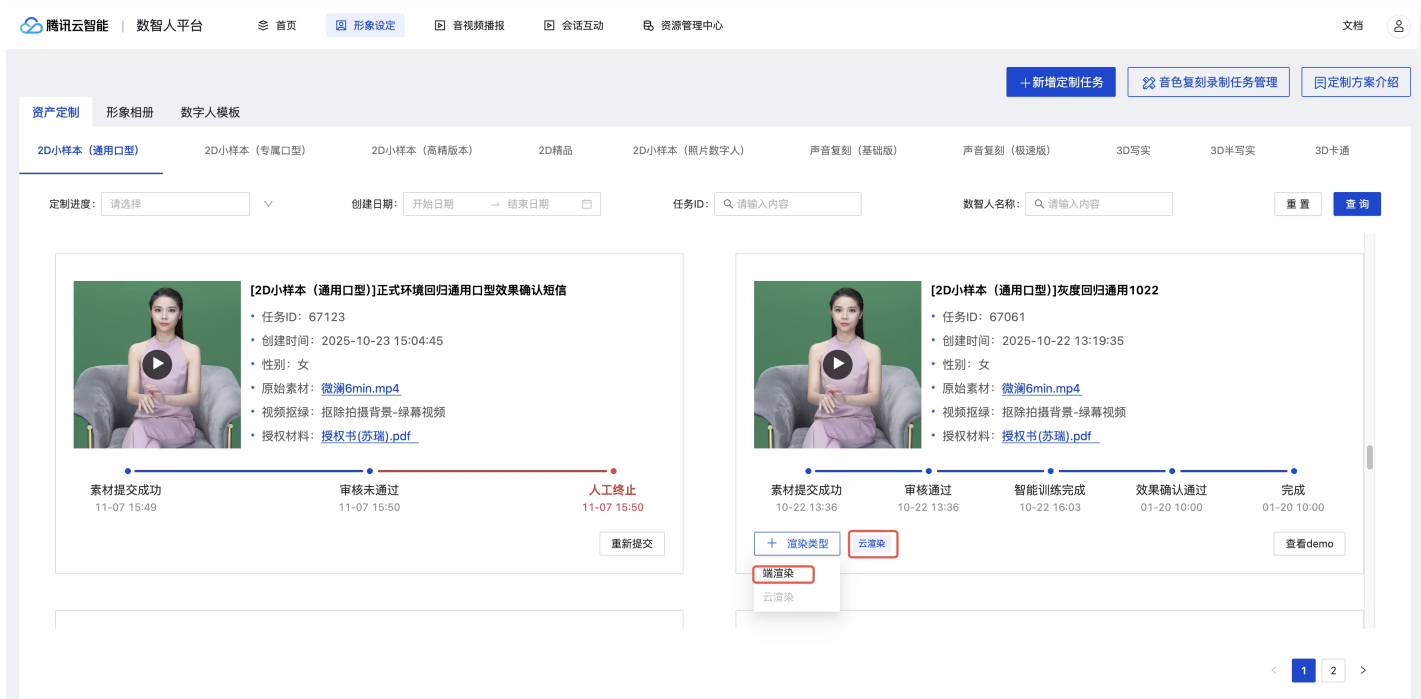


场景2：配额已经用于训练过云渲染形象，需要补训端渲染形象。

注意：

补训场景需要特别注意，补训的配额只有半年有效期，即从云渲染形象训练成功后半年内，可二次发起端渲染形象训练，超期不可补训，相关训练入口会自动关闭。反之亦然，即先训练端渲染形象，后补训云渲染形象场景也符合上述要求。

找到已经训练好的云渲染形象任务单，如下图。单击渲染类型 > 端渲染即可快速发起补训端渲染形象。

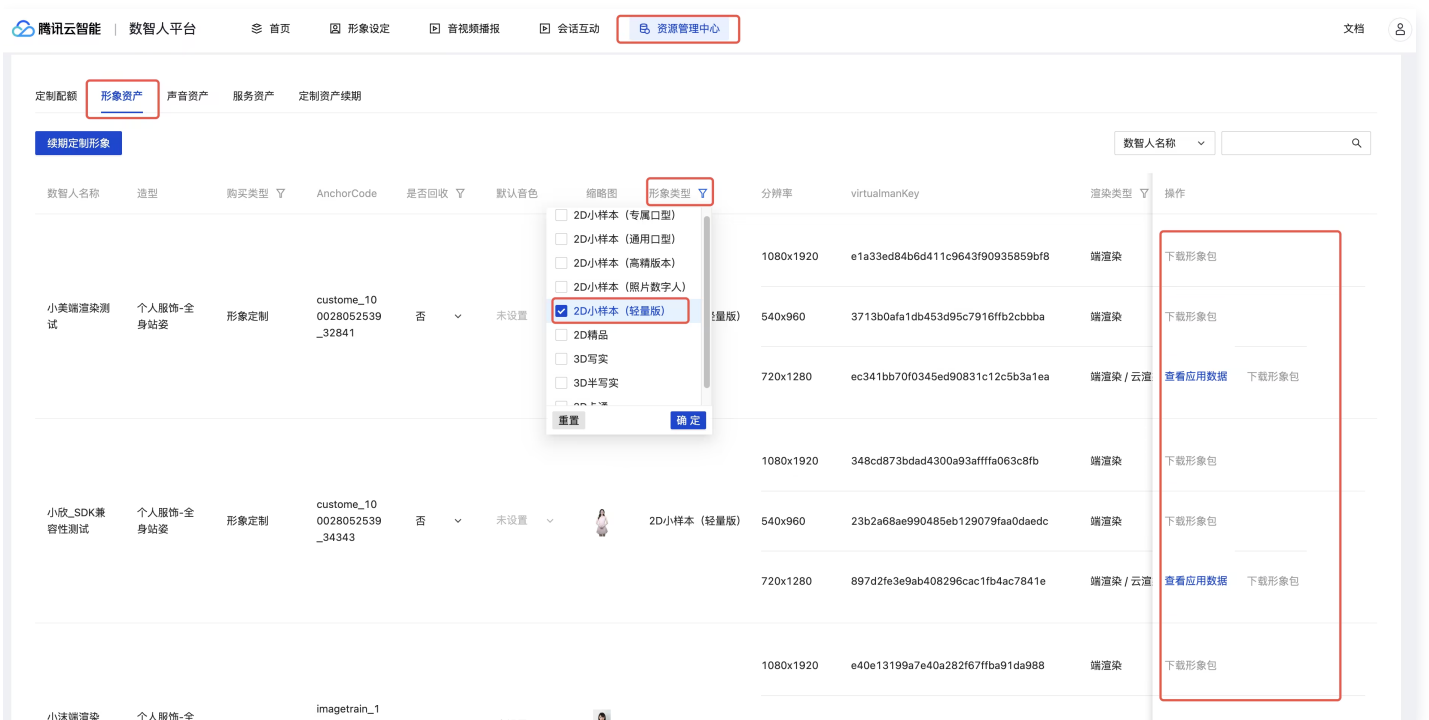


下载使用流程

1. 登录 [数智人平台](#)。
2. 在[资源管理中心](#) > [形象资产](#)中，通过形象类型过滤筛选出“2D小样本（轻量版）”，自动过滤出所有端渲染形象。
3. 在右侧的操作列，有专门的形象包下载地址。

⚠ 注意：

下载形象包默认只有7天的下载开放时间，7天后下载入口会自动关闭。在形象训练完成后，请尽快完成下载。



相关 SDK、Demo 包请联系商务同学获取，待相关流程完善后，我们会尽快提供自助下载入口。

SDK 的使用教程详见：[2D 端渲染场景概述](#)

2D 端渲染场景所需购买的清单

SKU 名称	是否必选	数量
定制配额（小样本专属、小样本通用、小样本照片任选其一即可）	必选	1
端渲染授权 license（订阅制，可以包年或包月）	必选	1

轻量版云端服务小时包（端云一体方案专用）	可选	1
----------------------	----	---

定制配额：

云智能数智人 [返回产品详情](#)

- 2D形象
- 3D形象
- 声音复刻
- 更多服务

购买须知

- 计费概述** 形象购买不可单独使用，需要加购配合服务资产一同使用；具体可参见腾讯云智能数智人的**计费概述**
- 使用说明** 购买完成后资源为生效状态，可前往**数智人平台**或者通过调用**API**使用该产品。该产品仅限在中国境内使用。
- 退费说明** 购买前请详细了解**退费说明** 腾讯云有权根据退费说明拒绝相应退款要求

选择配置

形象类型

<p>2D小样本通用口型</p> <p>通过一段真人视频素材进行训练数智人</p> <p>✓ 适用于对数智人唇齿特征无要求、无良好拍摄条件的客户。</p>	<p>2D小样本专属口型</p> <p>通过一段真人视频素材进行训练数智人</p> <p>✓ 适用于对数智人形象复刻有要求、有较好拍摄条件的客户。</p>	<p>2D小样本高精</p> <p>通过一段4K的真人视频素材进行训练数智人</p> <p>✓ 适用于大型会议、面对面对话、产品发布会、大屏场景。</p>
<p>2D小样本照片</p> <p>通过一张照片即可训练数智人</p> <p>✓ 适用于泛互、娱乐场景。</p>	<p>2D精品</p> <p>通过在专业影棚中录制动作素材</p> <p>✓ 适用于金融、传媒类对数智人形象、动作有要求的客户。</p>	<p>2D小样本免训练（视频素材）</p> <p>无需训练即可快速生成一个数智人视频</p> <p>✓ 适用于上传素材为视频的场景</p>
<p>2D小样本免训练（照片素材）</p> <p>无需训练即可快速生成一个数智人视频</p> <p>✓ 适用于上传素材为照片的场景</p>	<p>2D端渲染</p> <p>通过本地渲染呈现数字人效果，多场景通用</p> <p>✓ 适用于通用口型、专属口型、高精、照片等形象类型</p>	

[形象类型说明](#)

采购类型

- 形象租赁
- 形象定制**
- 定制形象续期

选择数量

- 0 +

定制配额

注意：1、形象定制配额自下单之日起有1年使用有效期，过期后配额自动失效；2、已定制完成的形象（2D小样本照片除外）自效果确认验收之日起有1年有效期，过期后可通过“定制形象续期”进行续费使用。3、2D小样本照片形象可永久使用。

资源包（服务资产）

会话互动并发数（云渲染场景）

会话互动场景必选，支持对应路数的会话交互，月包和年包购买下单即生效，日包（1天）和周包（7天）需要通过主账号手动激活，激活指引详见：[个人资产管理](#)

- 0 + 路

请选择

端渲染授权 license 和轻量版云端服务小时包：

云智能数智人 [返回产品详情](#)

2D形象

3D形象

声音复刻

更多服务

购买须知

- 计费概述** 形象购买不可单独使用，需要加购配合服务资产一同使用；具体可参见腾讯云智能数智人的[计费概述](#)
- 使用说明** 购买完成后资源为生状态，可前往[数智人平台](#)或者通过调用API使用该产品。该产品仅限在中国境内使用。
- 退费说明** 购买前请详细了解[退费说明](#) 腾讯云有权根据退费说明拒绝相应退款要求

选择配置

形象类型

<p>2D小样本通用口型 通过一段真人视频素材进行训练数智人 ✓ 适用于对数智人唇齿特征无要求、无良好拍摄条件的客户。</p>	<p>2D小样本专属口型 通过一段真人视频素材进行训练数智人 ✓ 适用于对数智人形象复刻有要求、有较好拍摄条件的客户。</p>	<p>2D小样本高精 通过一段4K的真人视频素材进行训练数智人 ✓ 适用于大型会议、面对面对话、产品发布会、大屏场景。</p>
<p>2D小样本照片 通过一张照片即可训练数智人 ✓ 适用于泛互、娱乐场景。</p>	<p>2D精品 通过在专业影棚中录制动作素材 ✓ 适用于金融、传媒类对数智人形象、动作有要求的客户。</p>	<p>2D小样本免训练 (视频素材) 无需训练即可快速生成一个数智人视频 ✓ 适用于上传素材为视频的场景</p>
<p>2D小样本免训练 (照片素材) 无需训练即可快速生成一个数智人视频 ✓ 适用于上传素材为照片的场景</p>	<p>2D端渲染 通过本地渲染呈现数字人效果，多场景通用 ✓ 适用于通用口型、专属口型、高精、照片等形象类型</p>	

[形象类型说明](#)

轻量版云端服务小时包

资源包 (服务资产)

<p>通用音频合成100小时包 可用于交互会话场景 (端渲染)、单独生成音频等场景，购买下单即生效。 默认支持 20 路并发。有效期自购买之日起一年内有效</p>	<p>- 0 + 个</p>
<p>轻量版云端服务10小时包 本服务主要用于端云一体场景。在使用端渲染SDK时，当终端设备的本地渲染性能不足时，可无缝切换至云端渲染，确保体验流畅。有效期自购买之日起一年内有效</p>	<p>- 0 + 个</p>
<p>通用音频合成并发资源 需搭配“通用音频合成 100 小时包”使用，购买该资源可以扩充通用音频合成的并发路数。购买下单即生效</p>	<p>- 0 + 路 请选择</p>
<p>2D端渲染SDK授权使用年包-按应用 通过派发秘钥 (AppID+SecretKey) 实现授权时长的管理，将SDK连同秘钥集成到终端应用中使用。不限终端数量，自秘钥派发之日起生效。该权益不包含TTS服务，如有涉及需购买“通用音频合成 100 小时包”</p>	<p>- 0 + 个 请选择</p>
<p>2D端渲染SDK授权使用年包-按设备 针对单台设备授权方式进行售卖，有效期自设备激活之日起算。激活动作由设备端自主触发，最晚不超过6个月。若超期未激活，会在6个月后的第1天默认自动激活</p>	<p>- 0 + 个 请选择</p>

端渲染授权 license 二选一

注意：购买下单前，请详细阅读[价格指南](#) (涉及优惠折扣)

服务合约

我已阅读并同意 [产品服务协议](#) 和 [服务等级协议](#)