

解决方案实践

# 硅基智能多模态仿真数字人解决方案实践

文档版本 1.0  
发布日期 2023-12-07



版权所有 © 华为技术有限公司 2024。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

## 商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

## 注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

# 安全声明

## 漏洞处理流程

华为公司对产品漏洞管理的规定以“漏洞处理流程”为准，该流程的详细内容请参见如下网址：

<https://www.huawei.com/cn/psirt/vul-response-process>

如企业客户须获取漏洞信息，请参见如下网址：

<https://securitybulletin.huawei.com/enterprise/cn/security-advisory>

# 目录

1 方案概述.....	1
2 资源和成本规划.....	4
3 操作流程.....	7
4 实施步骤.....	9
4.1 硅基数字人形象克隆.....	9
4.2 硅基数字人视频制作.....	11
4.3 硅基数字人直播搭建.....	21
4.4 硅基数字人交互实施.....	41
5 修订记录.....	50

# 1 方案概述

## 应用场景

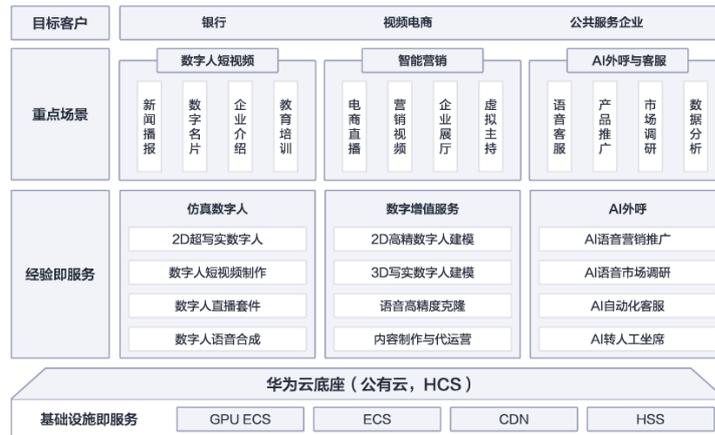
目前数字人已经广泛应用在众多行业的各类业务场景中，典型的行业客户及应用场景包括如下：

- 银行：数字经理、理财视频、在线服务、直播营销、迎宾大屏、展示大屏、智能机柜、学习服务等。
- 保险：计划书播报、保全服务、直播营销、客服咨询、代理人形象打造、内部服务等。
- 基金：产品介绍、基金经理和销售员克隆、直播营销、在线服务、内部服务等。
- 党政：政策宣传、在线服务、党史学习、迎宾指导大屏等。
- 广电：主持人克隆、新闻播报、活动发布、天气预报、内部服务等。
- 教培：校园介绍、名师克隆、课件制作、展馆名人复刻、招生宣传、学习服务平台、健康校园打造等。
- 文旅：景点介绍视频、直播宣传、线上咨询服务、线下服务大屏等。
- 电商直播：店主克隆、产品介绍视频、直播带货、客服服务等。
- 展示：历史人物克隆、重量嘉宾克隆、展示活动宣传、产品介绍视频、直播播报等。

## 方案架构

方案主要由华为云计算底座+硅基智能多模态仿真数字人产品形成面向银行、保险、教培、党政、文旅、电商直播、基金理财等领域的AI服务解决方案，硅基智能以自研DUIX操作系统为核心，通过打造开放开源，美观智能，7\*24不间断服务的硅基劳动力，为客户提供近乎无限供给的数字员工，大幅度降低组织成本，快速提升组织效率，让客户在数字化转型中拥有以算力为核心的创新竞争力。

图 1-1 架构图



## 方案优势

- 市场侧竞争力

较好的品牌知名度：华为云在虚拟数字人领域经过持续的投入，已经积累了较好的品牌知名度，同时作为合作伙伴的硅基智能在行业里也有较好的品牌知名度。

很好的客户需求匹配度：硅基智能多模态仿真数字人解决方案包含的内容完整同时可以模块化单独部署，可以很好匹配客户各类应用场景的需要，包括部分客户分阶段规划分阶段部署的需要考虑。

良好的服务支撑能力：华为云联合服务商合作伙伴硅基智能，具有全面的服务支撑能力，包括售前的咨询、方案的制定以、售后的部署、售后的运维运营支持等内容。

- 技术侧竞争力

交互能力先进：支持高还原度仿真数字人形象克隆、动作拟合、声音克隆与合成(TTSA)、支持语音与音乐合成(STA, AI歌唱)等多项仿真数字人核心技术；支持DUIX平台人机对话交互，以及2D卡通形象阿凡达模式数字人等特色功能。

产品能力强：完善的数字人视频制作、数字人直播、数字人AI交互系统，并支持数字人和真人同屏交互，支持手机APP、H5、直播、大屏、电视、投影、全息等展示方式；支持全面SAAS服务，根据客户需要，按时长和并发数计费，弹性低门槛。

实施效率高：快速定制数字人形象和进行语音克隆，最少仅需5分钟真人视频，甚至1张真人照片即可训练，语音克隆仅需2-3小时录音采样；最快一周即可完成形象和语音训练，投入商用。

- 云服务竞争力

灵活弹性的高性能云服务器：部署HPC头节点和工作节点，提供H1、H2计算增强型、M2内存优化型的ECS计算实例，该计算实例类型是H系列和M系列针对HPC业务的增强型实例，承载该实例的物理服务器提供最为强劲的计算能力。H2类型云服务器会挂载本地NVMe PCIe SSD，最大提供3.2T容量、300K IOPS、3000MBps带宽的存储能力。

优良性能的裸金属服务器：按照专属物理服务器方式提供优良的计算性能，采用2路18核服务器，无虚拟化损耗，具备最佳的性能优势，用户可通过管理控制台实现裸金属服务器的自动化发放，满足HPC业务按需灵活弹性的需求。

Lustre高性能并行文件系统：使用经过优化的专用ECS、EVS规格部署Lustre文件系统，单OSS可提供最大4.5GBps吞吐量，最大提供2Tbps吞吐量，单文件支持最大32PB满足HPC的文件系统性能要求。

分布式可扩展的块存储服务：基于分布式架构提供的可弹性扩展的块存储服务，每个弹性云服务器可按需挂载云硬盘，为HPC计算节点和存储节点提供存储空间，单盘最大支持32TB容量、30K IOPS、1Tbps吞吐量。

优良性能的VPC网络：为HPC用户在公有云上构建安全隔离的虚拟网络环境，HPC计算网络通过RDMA EDR Infiniband 网络互连，节点间最低延时90ns，单节点最大带宽100Gbps，为用户提供高带宽的网络环境。

海量的对象存储服务：基于对象的海量存储服务，为HPC应用提供存储空间，存储HPC算例数据结果数据，用户可通过对象存储服务灵活的上传、下载HPC计算数据。

# 2 资源和成本规划

本节介绍解决方案实践中资源规划情况，包含以下内容：

- 硅基数字人视频

图 2-1 部署架构图 1

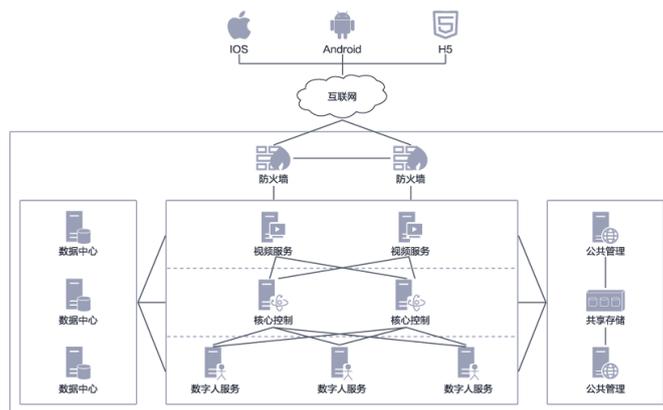


表 2-1 资源需求 1

服务	类型	数量	配置	部署系统
基础服务	公共管理服务器	2	CPU: 16C MEM: 64G DISK: 500G	配置管理 资源管理 用户中心 监控告警 运营分析
	视频服务服务器	2	CPU: 32C MEM: 128G DISK: 1T	视频合成 视频管理

服务	类型	数量	配置	部署系统
	核心控制服务器	2	CPU: 32C MEM: 128G DISK: 1T	控制调度 任务管理
数字人渲染服务	数字人服务器	X	2U服务器 CPU: 64C MEM: 128G DISK: 2T GPU: 4*T4	数字人驱动 数字人渲染
数据服务	数据服务器	3	CPU: 32C MEM: 128G DISK: 3T	关系型数据库 全文搜索引擎 消息中心
	实时交互, 1080p, 需要的外网带宽	6Mb	600kb/路	/

数字人渲染服务模块，一个渲染节点2块T4 GPU支持2路并发，可根据并发量进行水平扩展。

基础服务当前配置可支持数千路并发；

上述配置为高可靠和负载均衡配置；如果不考虑高可靠，可以对基础服务部分进行简配。

- 硅基数字人直播及交互

图 2-2 部署架构图 2

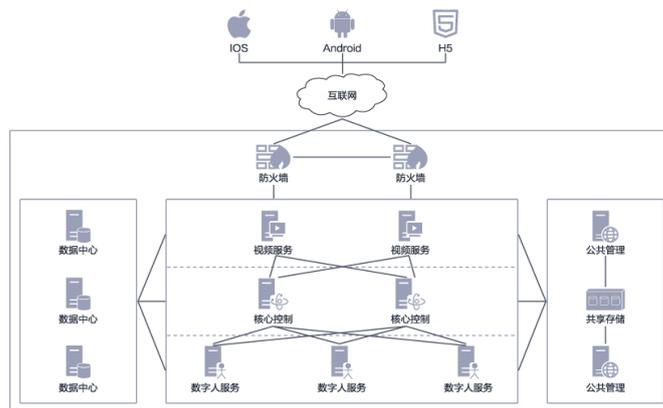


表 2-2 资源需求 2

服务	类型	数量	配置	部署系统
基础服务	公共管理服务器	2	CPU: 16C MEM: 64G DISK: 500G	配置管理 资源管理 用户中心 监报告警 运营分析
	视频服务服务器	2	CPU: 32C MEM: 128G DISK: 1T	视频合成 视频管理
	核心控制服务器	2	CPU: 32C MEM: 128G DISK: 1T	控制调度 任务管理
duix	数字人服务器	X	2U服务器 CPU: 64C MEM: 256G DISK: 2T GPU: 2*V100	数字人驱动 数字人渲染
直播	数字人服务器	X	2U服务器 CPU: 32C MEM: 256G DISK: 2T GPU: 2*V100	数字人驱动 数字人渲染
数据服务	数据服务器	3	CPU: 32C MEM: 128G DISK: 3T	关系型数据库 全文搜索引擎 消息中心
	实时交互, 1080p, 需要的外 网带宽	XMbps	600kb/路	/

数字人渲染服务模块，一个渲染节点1块V100 GPU支持2路并发，可根据并发量进行水平扩展。

基础服务当前配置可支持数千路并发；

上述配置为高可靠和负载均衡配置；如果不考虑高可靠，可以对基础服务部分进行简配。

# 3 操作流程

硅基智能多模态仿真数字人解决方案专业服务已经上架为联运商品，操作流程如下：

图 3-1 流程图



各流程活动的具体工作和要求如下表格所示：

表 3-1 工作和要求

序号	工作内容	具体描述	责任人	备注/模版
1	需求方案	华为云及硅基协助客户沟通确认具体的需求，包括所需产品、部署方式、并发量、场景定制等内容，明确方案及具体下单路径及下单内容。	客户+硅基+华为云	/
2	完成下单	基于需求的澄清及方案的确定，客户完成在华为云上的下单。	客户	/
3	实施计划	基于需求澄清及下单内容，华为云及硅基同客户对交付方案进一步沟通明确，包括部署方式、并发量、场景、硬件资源、交付计划等关键内容。	硅基+华为云+客户	/

序号	工作内容	具体描述	责任人	备注/模版
4	交付执行	基于交付方案的确定，硅基与华为云完成交付执行。	硅基+华为云	/
5	验收确认	基于交付执行的结果，客户与硅基及华为云完成交付验收，确认收入。	客户+硅基+华为云	/

# 4 实施步骤

- 4.1 硅基数字人形象克隆
- 4.2 硅基数字人视频制作
- 4.3 硅基数字人直播搭建
- 4.4 硅基数字人交互实施

## 4.1 硅基数字人形象克隆

图 4-1 硅基数字人形象克隆



### 实施步骤介绍

1. 真人模特选择：根据业务的具体需要，客户选择好数字人真人模特，可以是客户的营销、服务或者相关的主管领导的形象。
2. 绿幕视频录制：在数字人真人模特选好之后，根据业务需要及录制指导，录制2到5分钟的绿幕视频，并把录好的视频给到硅基。
3. 数字人模型训练：基于上面录好的绿幕视频，硅基基于训练平台加以专业的服务完成对模型训练，并发给客户进行确认。
4. 数字人模型绑定：上面训练好的数字人模型可以用于视频制作、直播或交互，根据客户的需要硅基把形象绑到相应的应用。

## 具体过程介绍如下

### 步骤1 基础必读

定制数字人需要拍摄一段3-5分钟的真人出镜口播视频，经过AI训练后可以得到一比一还原的数字人形象。

具体出镜姿势、口播内容，可以根据最终的使用场景来确定，如后期用来制作财经知识视频，即可选择正装出镜，口播财经内容，这样整体情景比较一致。

定制好的数字人口型可以根据文字、录音内容自动匹配。

### 步骤2 拍摄前(根据自己需要决定拍绿幕还是实景)

#### ● 绿幕布景

绿幕布景：绿幕拍摄后期可以自定义更换其它背景，拍摄时绿幕需要铺满整个画面，尽量不要有褶皱、阴影，否则会影响抠图效果，请务必得到满意布景后再开始拍摄。

发型佩饰：避免佩戴会晃动的耳坠，耳坠晃动会影响口型学习，可佩戴固定耳钉。避免碎发太多，容易导致绿幕扣不干净，如果是实景拍摄则不受影响。

衣服颜色：使用绿幕背景拍摄视频时，请不要穿搭绿幕相近色(如黄色、绿色)的衣服，以免影响抠图效果。

#### ● 实景布景

实景布景：实景拍摄后期无法更换其它背景，拍摄时无特殊要求，请务必得到满意布景后再开始拍摄。

灯光：均匀打光的光源，推荐使用"三点式打光"，如果想要下巴区域无阴影，推荐使用反光板、泡沫板(俗称"米菠萝")来补光。

#### ● 相机拍摄参数设置(推荐)

- a. 视频分辨率尽量4K(3840\*2160)，最低1080P(1920\*1080)
- b. 视频尺寸为9:16或16:9，请根据自己的使用场景选择横屏或竖屏拍摄
- c. 视频帧率：30帧或60帧
- d. 镜头焦距：推荐50mm镜头(等效全画幅相机焦距)左右，想要脸瘦一点用85mm
- e. 镜头光圈：  
如拍摄绿幕，则光圈不宜太大  
如拍摄实景且想要虚化背景，建议根据实际情况调整光圈，保证人物主体清晰且背景虚化
- f. 如使用手机拍摄，请选择最高画质(1080P以上)进行拍摄，苹果手机请选择“兼容性最佳”格式拍摄

拍摄时(视频时长控制在5分钟以内)

#### ● 拍摄步骤

- a. 先保持静默10秒左右，闭口不要漏牙齿；
- b. 正常口播3-4分钟，可以有自然的手势动作，但动作不要靠近脖子、嘴巴、面部区域；具体出镜姿势、口播内容，可以根据最终的使用场景来确定，如后期用来制作财经知识视频，即可选择正装出镜，口播财经内容，这样整体情景比较一致。
- c. 实景拍摄到此即可，最终导出一段完整的视频即可；

- d. 绿幕拍摄需要在口播之后，人物需要走出画面并将桌椅等前景道具搬离画面，仅保留绿幕继续拍摄10秒左右空镜头。
- 注意事项
  - a. 眼神：注视镜头且录制连贯，有提词设备为佳
  - b. 收音：保持环境安静，无背景音及他人说话，有收音设备为佳说话时需吐字清晰，不能纯念数字1234
  - c. 手势动作：不要出现在面部、嘴巴、脖子周边附近

### 步骤3 拍摄后(美颜处理)

- 视频拍摄完成后直接导出一整条视频，不要剪辑，否则训练出来的数字人模型会跳帧!
- 素材建议导出4k 25/30p，最低1080P
- 由于相机直出素材动态范围更大，细节保留更多，建议完成调色美颜后，再进行数字人训练
- 定制数字人仅口型区域由AI生成，其他的头部动作(比如眨眼)、身体动作(比如手势)、美颜效果等都是保留录制时的样子，即根据所拍视频就可以预览数字人效果。

----结束

## 4.2 硅基数字人视频制作

图 4-2 硅基数字人视频制作



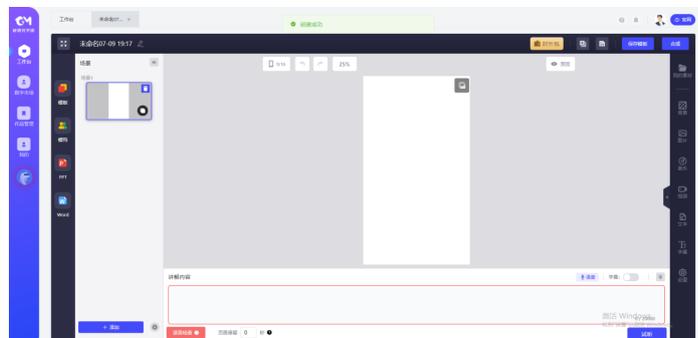
### 实施步骤介绍

1. 主题确定：根据业务的具体需要，客户选择好对应的业务主题，比如说银行理财产品介绍还是学校老师教学课件的制作。
2. 模特选择：基于上述主题的确定，选择对应的数字人模特，模特选择可以是库里已有的，也可以是前面已经帮客户定制好的数字人模特。
3. 背景选择：基于上述主题的确定，选择对应的视频背景模板，如果没有合适的，客户也可自行制作并上传使用。
4. 文案编辑：基于上述主题的确定，要整体准备好相应的文案内容，也就是需要数字人在视频中要讲的话。如果需要，客户也可以把录好的声音上传到数字人系统来代替文案内容的文字编辑。
5. 其它配置：基于上述主题的确定，其它配置包括了其它声音模型、图片素材、视频素材、背景音乐、是否需要字幕等配置。
6. 一键合成：以上所有选择及配置完成之后，即可进行一键合成，合成之后可以把视频导出使用或者提供链接的方式使用。

## 具体过程介绍如下

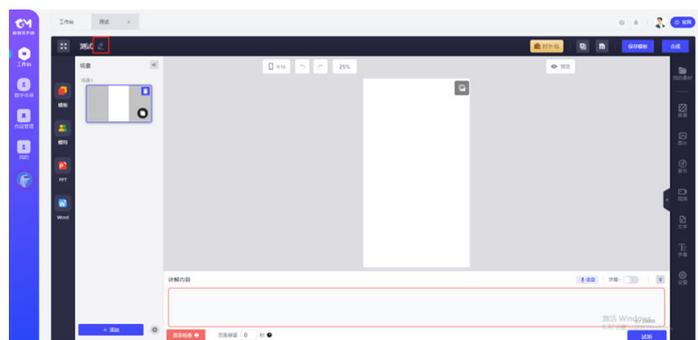
步骤1 首先进入视频编辑页面，如下图：

图 4-3 视频编辑页面



步骤2 单击页面左上方“”按钮，可修改作品名称，如下图：

图 4-4 修改作品名称



步骤3 单击“”按钮，可查看当前可用合成时长余额，选择充值可以进行视频合成时长充值，如下图：

图 4-5 充值 1

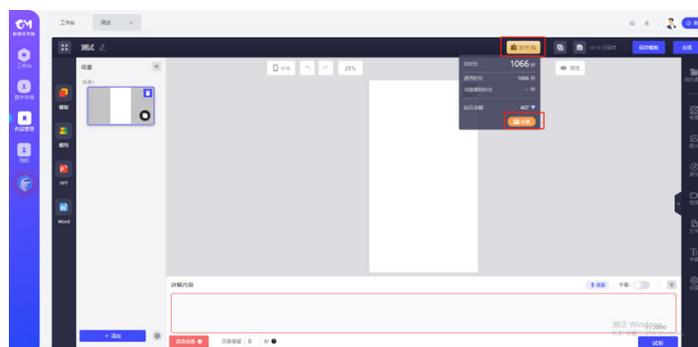
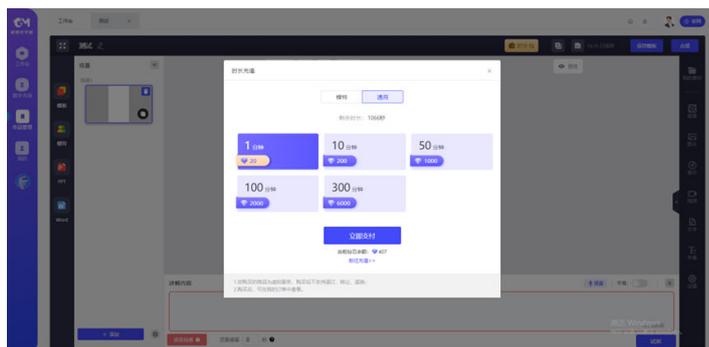


图 4-6 充值 2



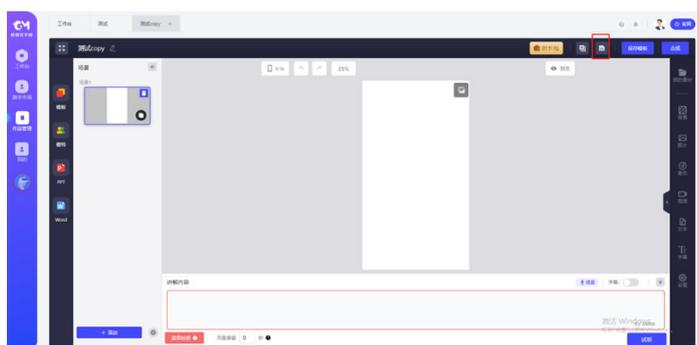
步骤4 单击“”按钮，可以复制当前编辑视频至新的操作界面进行编辑，如下图：

图 4-7 编辑



步骤5 单击“”按钮，可以保存当前设置，如下图

图 4-8 保存当前设置



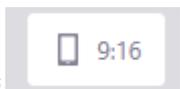
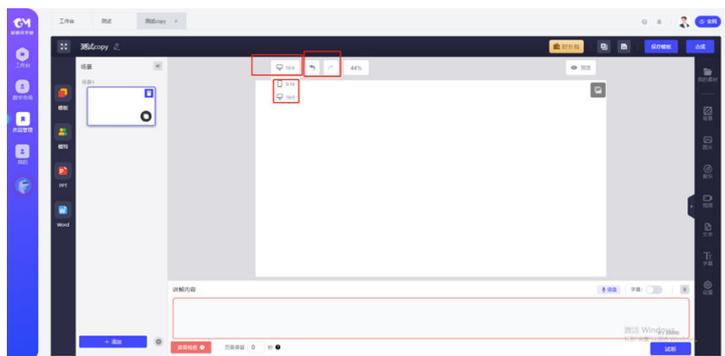
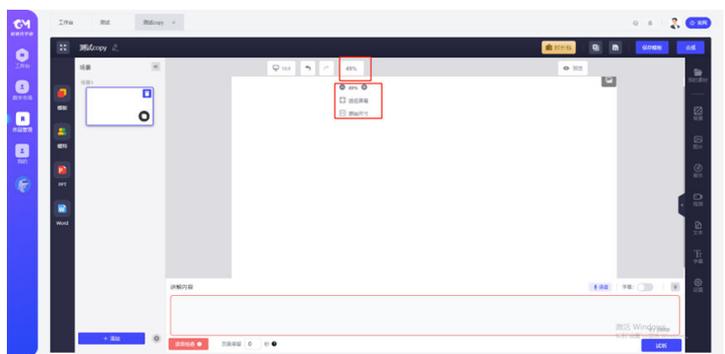
步骤6 单击“ 9:16”可切换视频横屏/竖屏展示，单击按钮，可撤销操作，如下图：

图 4-9 可撤销操作



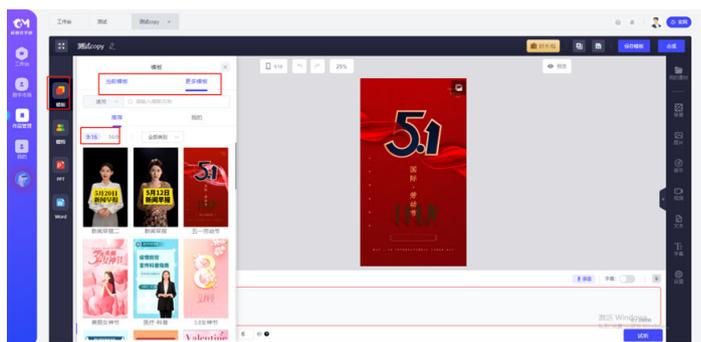
步骤7 单击“”可以扩大或者缩小页面比例，如下图：

图 4-10 扩大或缩小比例



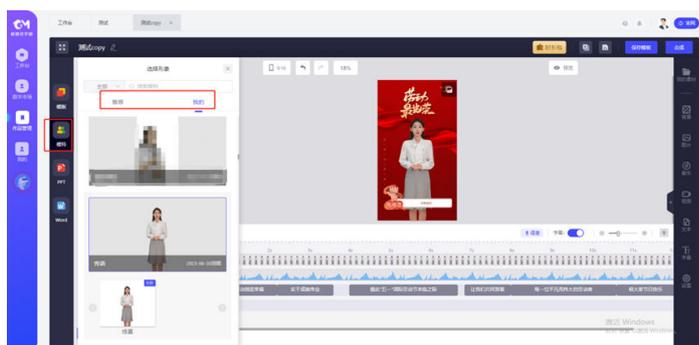
步骤8 单击“模板”，可选择切换不同的视频模板，如下图：

图 4-11 切换不同的视频模板



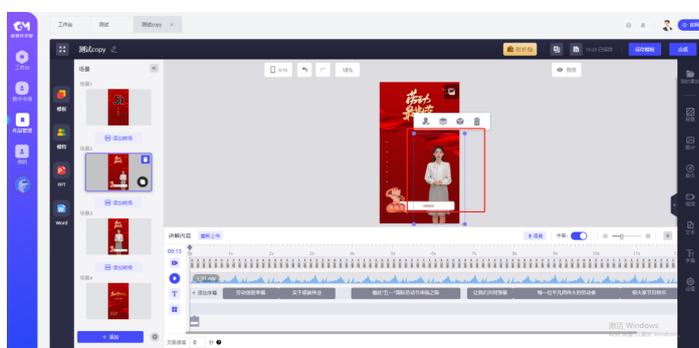
步骤9 单击“模特”可选择不同的模特，如下图：

图 4-12 可选择不同的模特



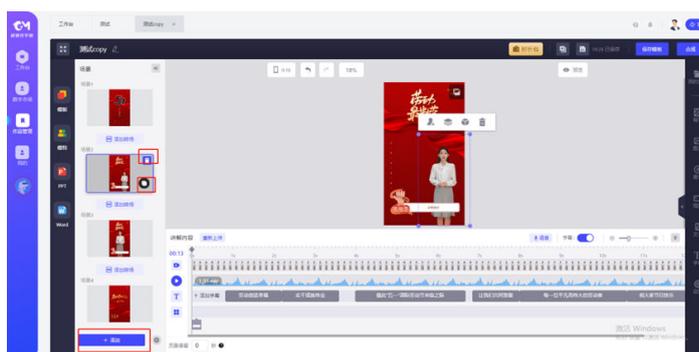
步骤10 选择好模特后，单击模特图层，可进行调整大小，移动位置等操作，如下图：

图 4-13 调整



步骤11 单击“添加”，可新增空白场景，选中场景单击场景上“”按钮，可复制当前场景，单击“”按钮可删除当前场景。如下图：

图 4-14 添加



步骤12 也可上传PPT或Word文档，作为视频场景或画中画，一键合成视频，PPT/Word即转化成视频，快捷又高效。如下面两图（PPT和WORD）：

图 4-15 上传 PPT 或 Word 文档 1

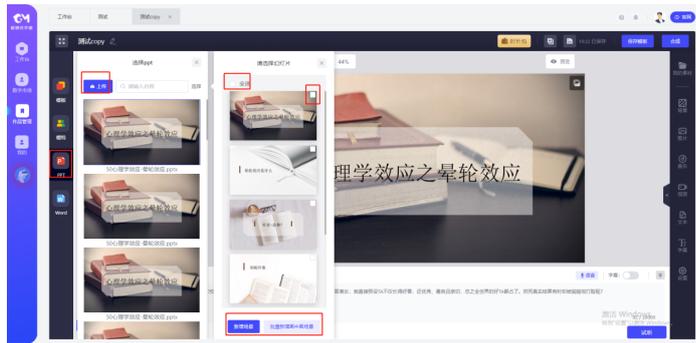


图 4-16 上传 PPT 或 Word 文档 2



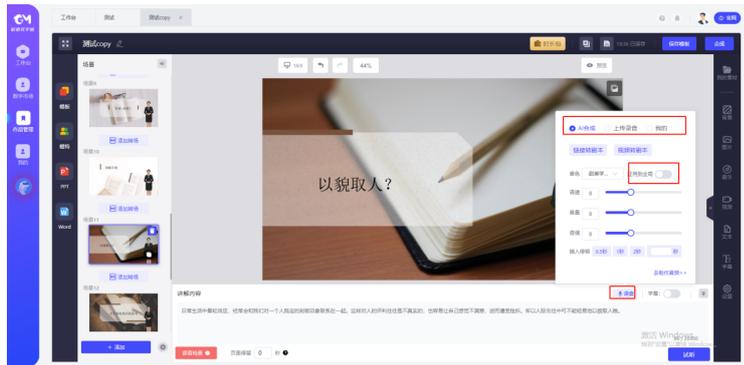
步骤13 PPT、Word场景如下图所示，自动识别PPT、Word中讲解内容

图 4-17 上传 PPT 或 Word 文档 2



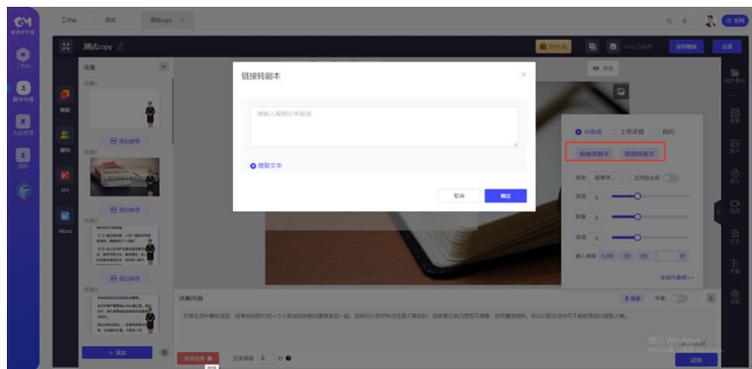
步骤14 场景设置完成后，接下来进行讲解内容设置。可在“讲解内容”栏输入需要讲解的内容，设置语音。可选择AI合成/上传录音两种方式，确定好音色，语速、音量及语调默认，也可根据需要手动调整；“应用到全局”可将调整好的音色应用到该视频下所有场景中。“我的”可选用用户在声音市场制作的音频文件。如下图：

图 4-18 内容设置



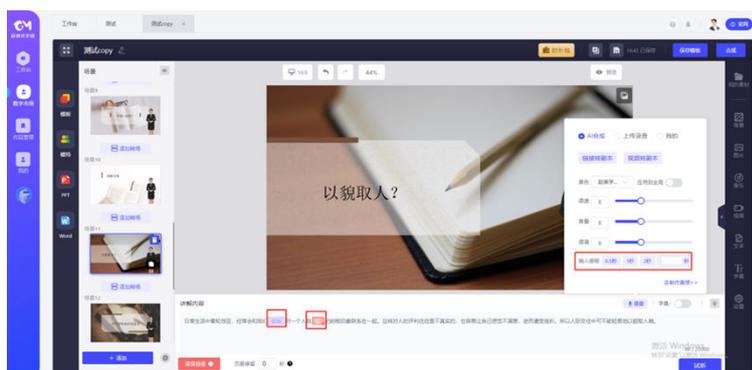
**步骤15** 也可单击“链接转剧本”，支持输入快手、小红书、美拍、微视、微博等平台链接，提取文本或录音；或者单击“视频转剧本”，支持导入本地视频提取文

图 4-19 提取文



**步骤16** 试听该语音，将鼠标置于需要插入停顿的文字后，单击插入停顿时长，可插入停顿，鼠标右击可设置该文本读音，如下图：

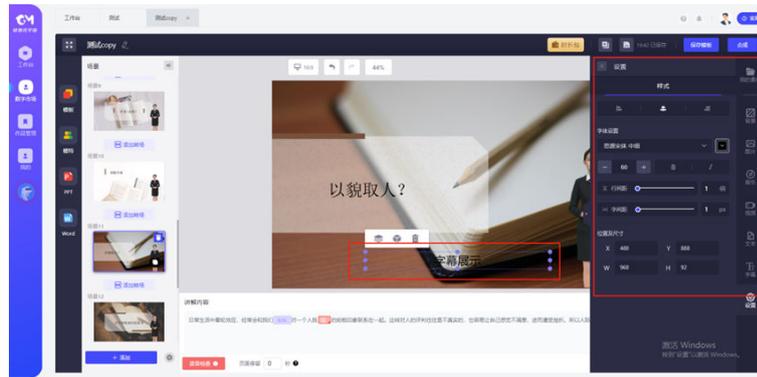
图 4-20 文本读音



**步骤17** 接下来进行字幕设置，点开“字幕”按钮，可插入视频字幕，双击可设置字幕字体、颜色、间距等属性。如下图：

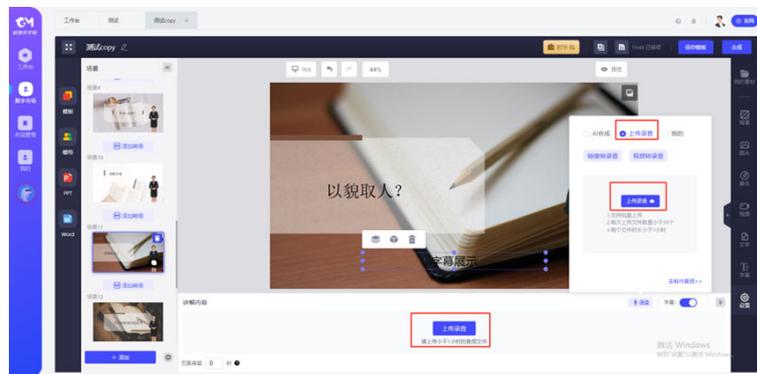


图 4-21 字幕设置



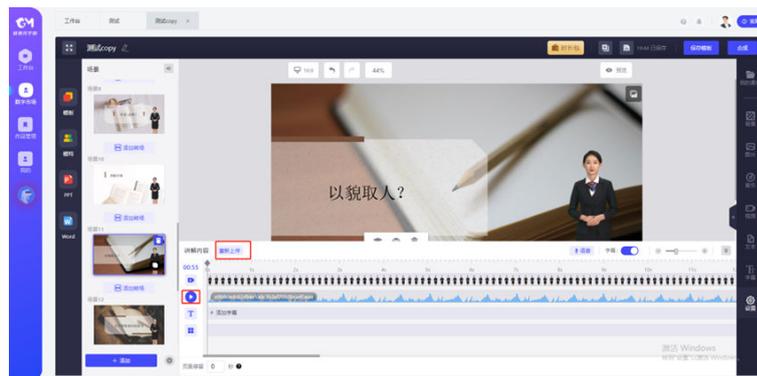
步骤18 上传录音：选择语音--上传录音，可从本地选择录音进行上传，如下图：

图 4-22 上传录音



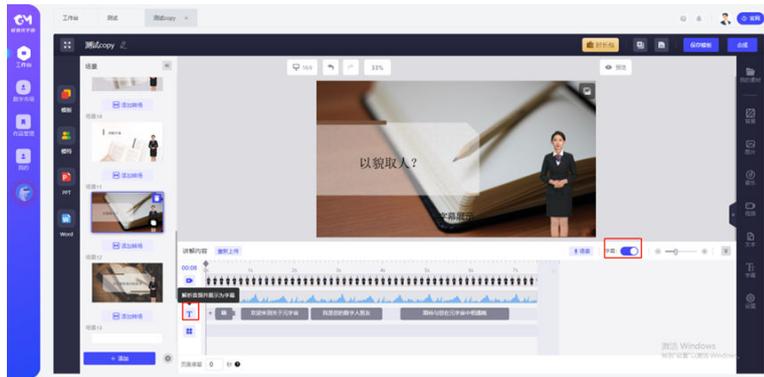
步骤19 上传成功后，单击左下方“”按钮，可试听该语音，单击“重新上传”，可重新上传语音，如下图：

图 4-23 试听语音



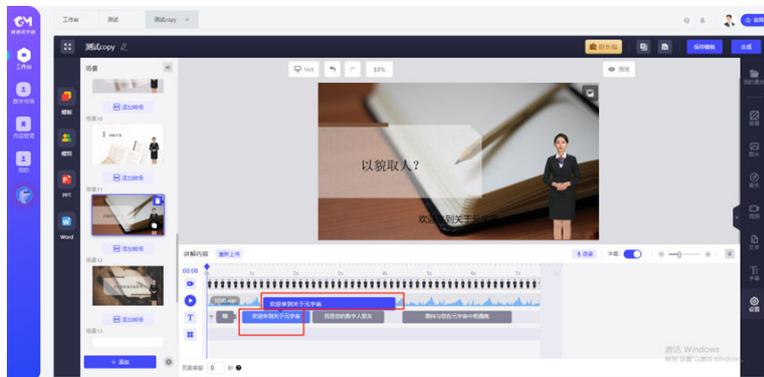
步骤20 单击第三轨道“”，会自动解析添加字幕，如下图：

图 4-24 解析添加字幕



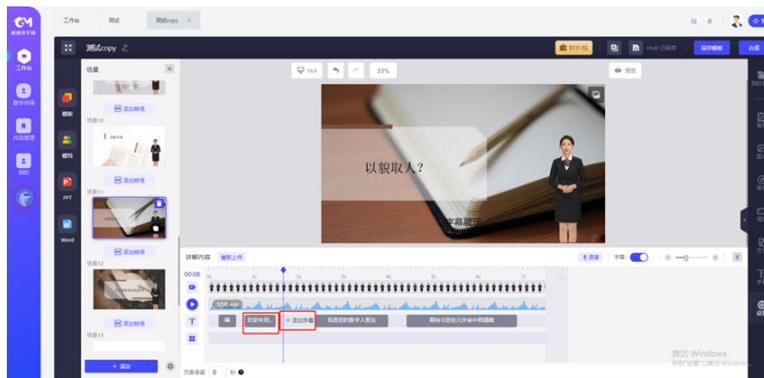
步骤21 解析成功后，选中下方字幕进度条，双击可修改字幕，如下图：

图 4-25 修改字幕



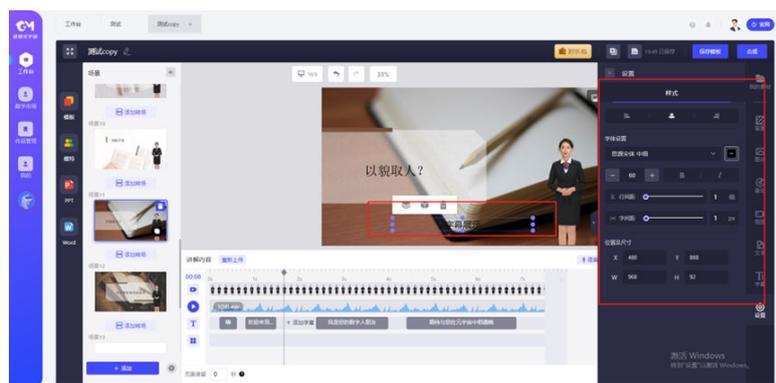
步骤22 拖动字幕条，可修改字幕长度，添加新字幕，如下图：

图 4-26 添加新字幕



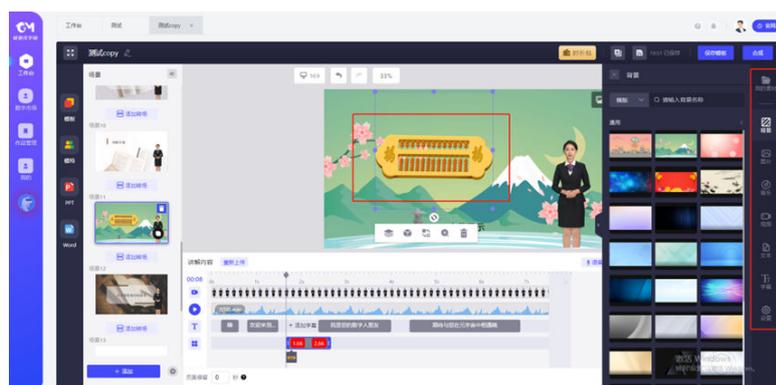
步骤23 双击页面字幕框，或者单击右侧“字幕”可打开设置字幕属性，可设置字体、颜色、粗心、间距等属性，如下图：

图 4-27 字幕属性



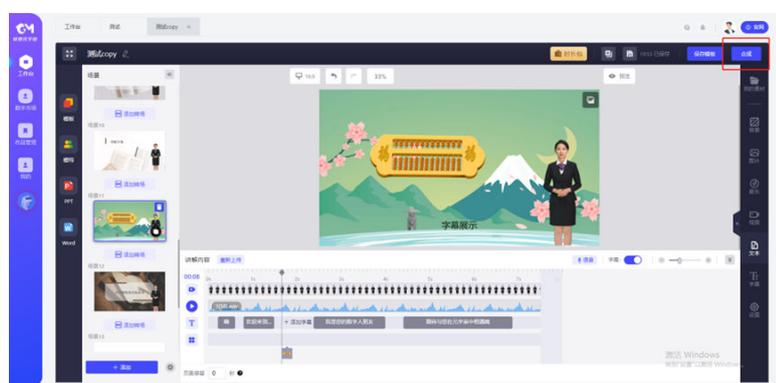
**步骤24** 页面右侧素材库中，“我的素材”可从本地上传图片、背景、视频、音乐等至我的素材中，插入到视频；也可使用系统提供的背景、图片、音乐、视频等素材一键替换视频背景，插入图片、音乐、视频等；“文本”中可输入文本插入视频中；选择添加字幕样式，设置字幕间距；设置该场景停留时长。在视频展示区选中该图层，可灵活拖动图层位置、大小，设置图层位置，查看及删除图层等操作。如下图：

图 4-28 素材库



**步骤25** 所有场景设置完成后，单击“合成”按钮，即可进行视频合成。如下图：

图 4-29 视频合成



----结束

## 4.3 硅基数字人直播搭建

图 4-30 图示 1



### 实施步骤介绍

1. 直播主题确定：根据业务的具体需要，客户确定好直播主题，比如说带货直播还是招生直播等。
2. 模特选择：基于上述直播主题确定，选择对应的数字人主播，主播选择可以是库里已经有的，也可以是前面已经帮客户定制好的数字人模特。
3. 背景选择：基于上述直播主题确定，选择对应的直播背景模板，如果没有合适的，客户也可自行制作并上传使用。配置也可以包括图片素材、视频素材、背景音乐等方面的配置。
4. 产品话术：基于上述直播主题确定，要整体准备好相应的话术内容，也就是需要数字人在直播中要讲的话。如果需要，客户也可以把录好的声音上传到数字人直播系统来代替话术内容的文字编辑。
5. 规则配置：基于上述直播主题确定，包括对数字人基于抓取弹幕信息的回复话术进行配置，包括直播单双人模式等相关规则的配置等。第四步，一键合成：以上所有选择及配置完成之后，即可进行一键合成，合成之后可以把视频导出使用或者提供链接的方式使用。
6. 一键推流：以上所有选择及配置完成之后，即可进行一键推流，也就是把抖音、视频号或其它直播终端的地址配置好完成把实时的直播视频流推过去。

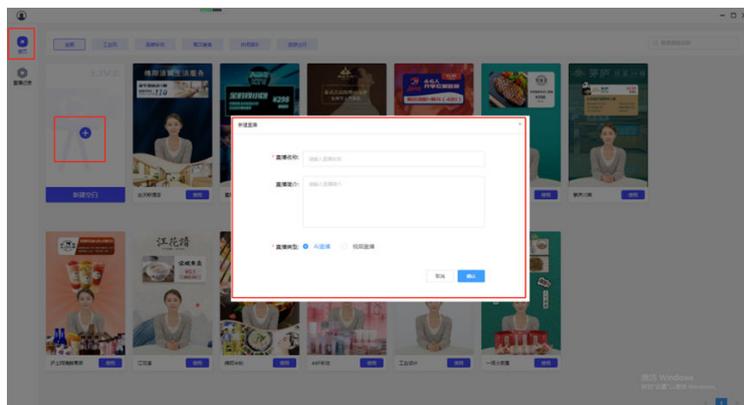
### 具体过程介绍如下

#### 步骤1 新建空白直播

平台有2个入口可以新建空白直播：

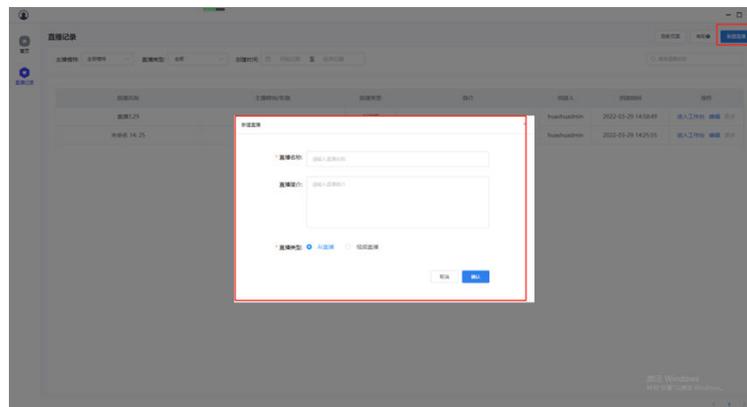
单击“首页 > 新建空白”，打开新建直播窗口；如下图：

图 4-31 图示 2



单击“直播记录 > 新建直播”，开始创建直播，如下图：

图 4-32 图示 3



输入直播名称、直播简介（选填），选择好直播类型后，单击确认按钮，如图2-3-3，即成功创建直播，如下图：

图 4-33 图示 4

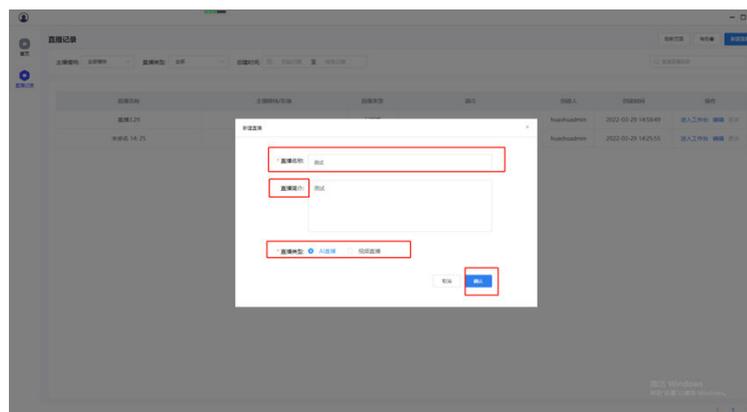
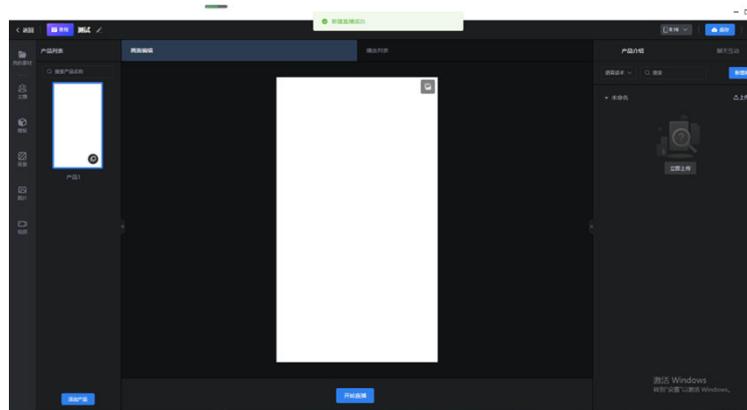
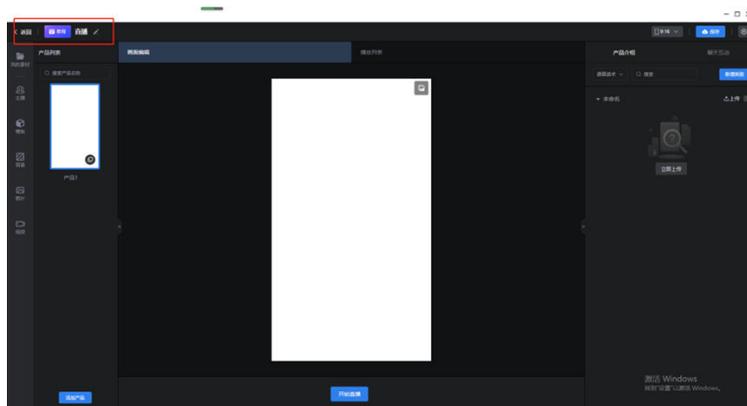


图 4-34 图示 5



在工作台页面，单击左上角“返回”按钮，可退出工作台页面，单击“教程”可以查看指导教程，单击“”按钮可修改直播名称，如下图：

图 4-35 图示 6



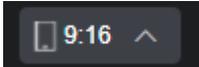
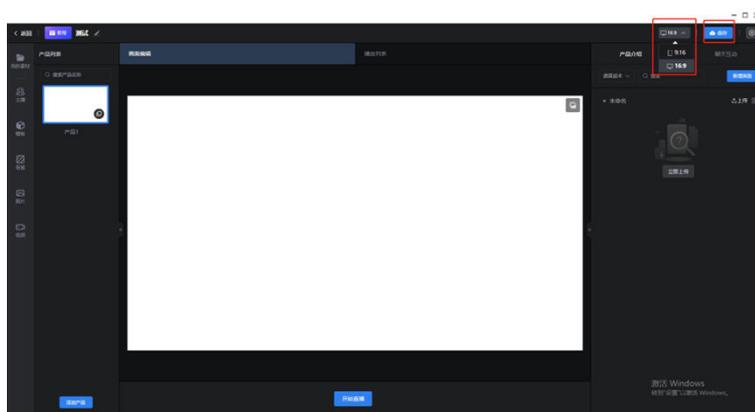
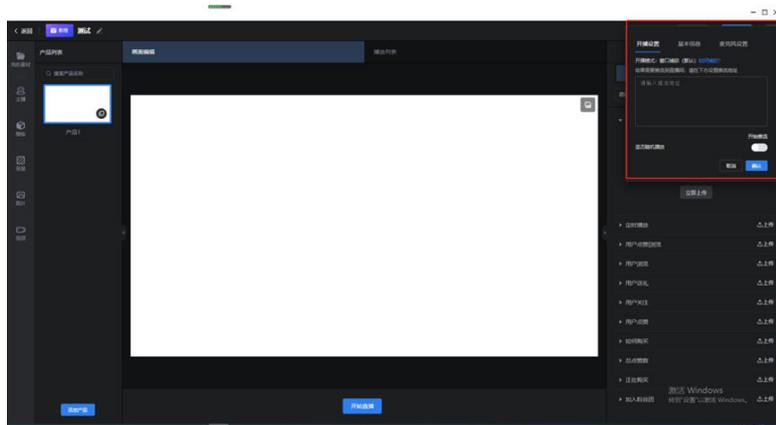
工作台右上方单击“”按钮可以调整直播间尺寸竖屏（9:16）/横屏（16:9），单击“”可以保存当前设置，如下图：

图 4-36 图示 7



单击“”进行开播设置，选择直播模式窗口捕获（默认）/推流（输入推流地址），直播基本信息设置（直播名称、直播简介），麦克风设置，如下图：

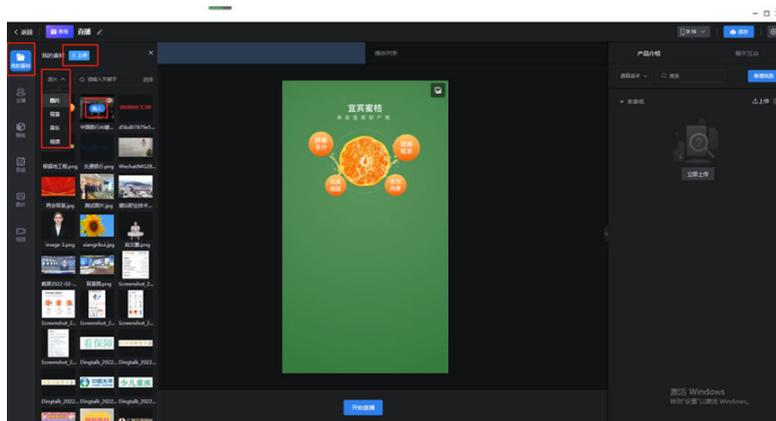
图 4-37 图示 8



### 步骤2 画面编辑

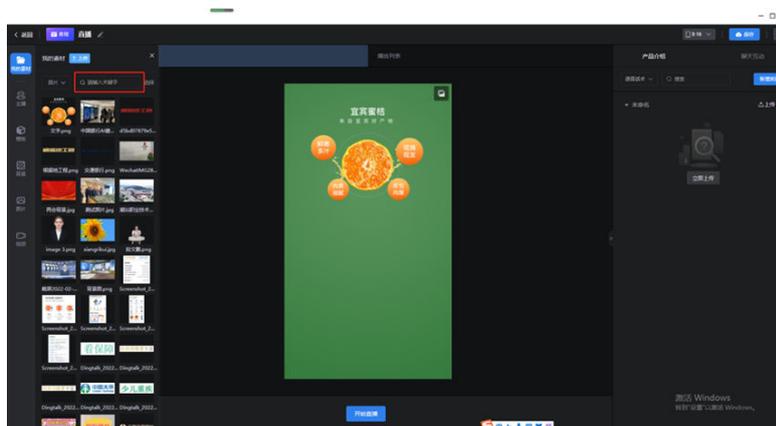
接下来进行直播间设置。单击“页面编辑”，单击页面左侧“我的素材”按钮，可从本地选择图片、背景、音乐及视频上传，并将素材插入到直播间。如下图：

图 4-38 图示 9



单击搜索按钮，可输入素材名称在素材库搜索相关素材，如下图：

图 4-39 图示 10



单击“选择”按钮，可以全选或部分选择已上传的素材进行删除操作，单击“取消”可取消选择，如下面两图：

图 4-40 图示 11

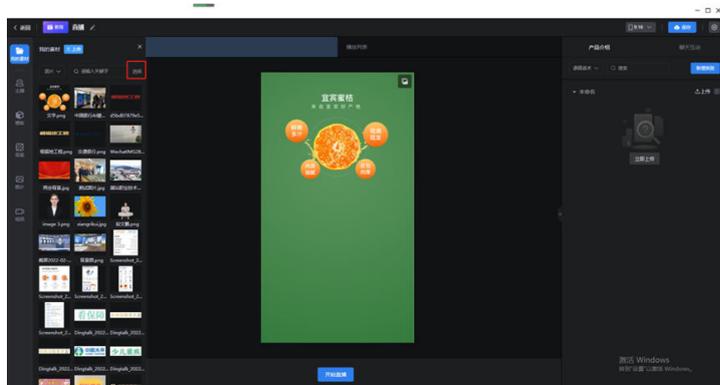
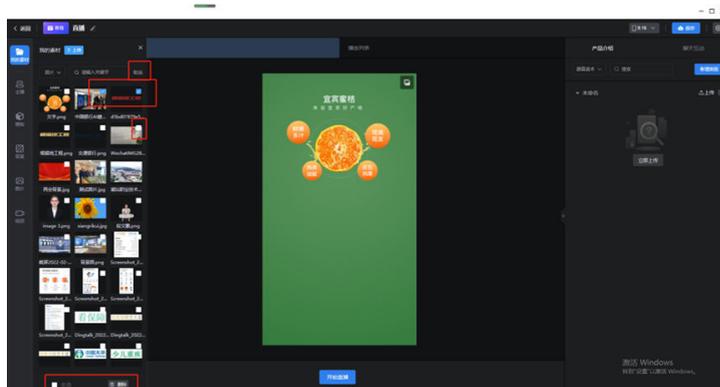
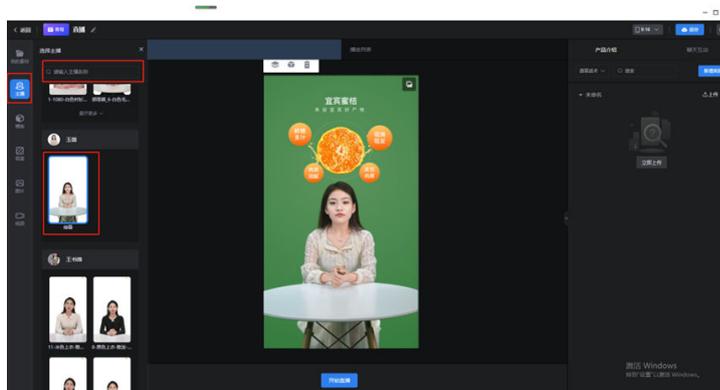


图 4-41 图示 12



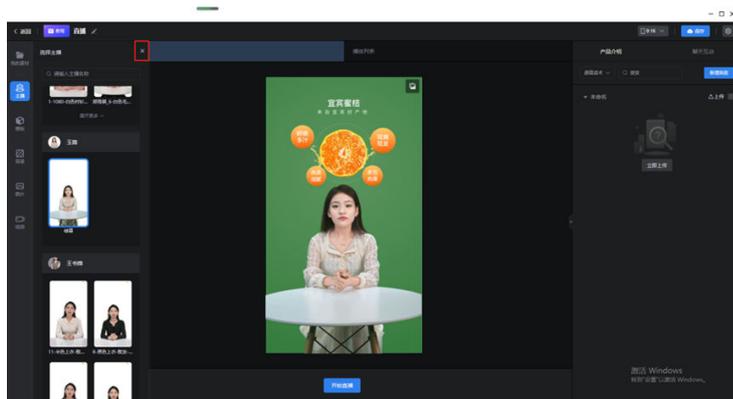
单击“主播”，选择视频主播及场景，也可通过搜索主播名称搜索指定主播，如下图：

图 4-42 图示 13



单击左上方“”可关闭选择框，如下图：

图 4-43 图示 14



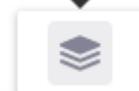
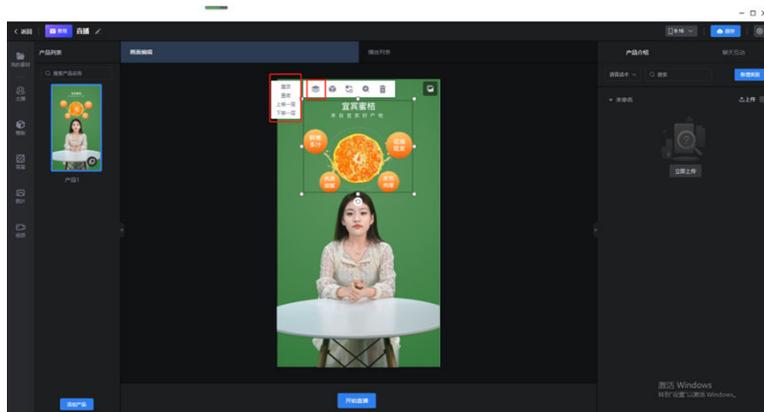
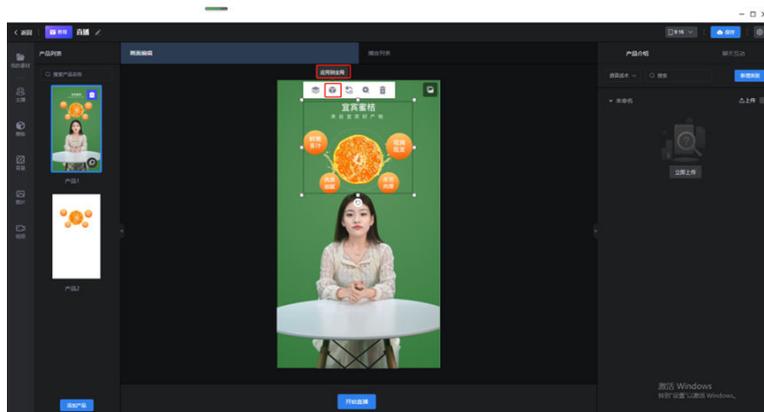
单击画布预览区，选中图层，单击“”按钮，可设置图层位置，如下图：

图 4-44 图示 15



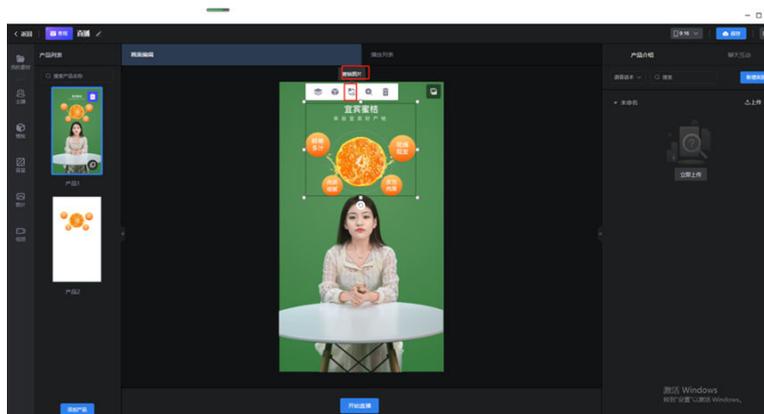
单击“”按钮，可以将当前选中内容应用到全部场景中，如下图：

图 4-45 图示 16



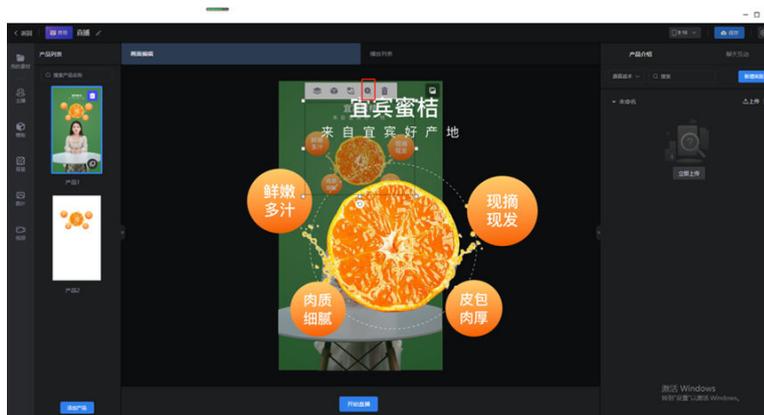
单击“”按钮，可以从本地选择图片进行替换，所替换图片大小默认与被替换图片大小一致，如下图：

图 4-46 图示 17



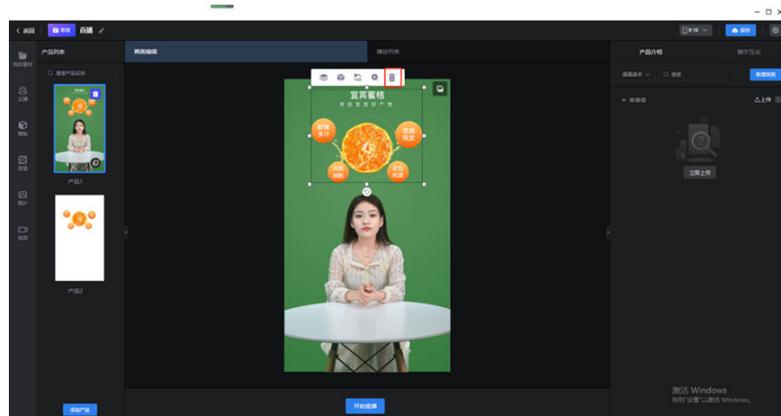
单击“”可以放大查看当前图片，如下图：

图 4-47 图示 18



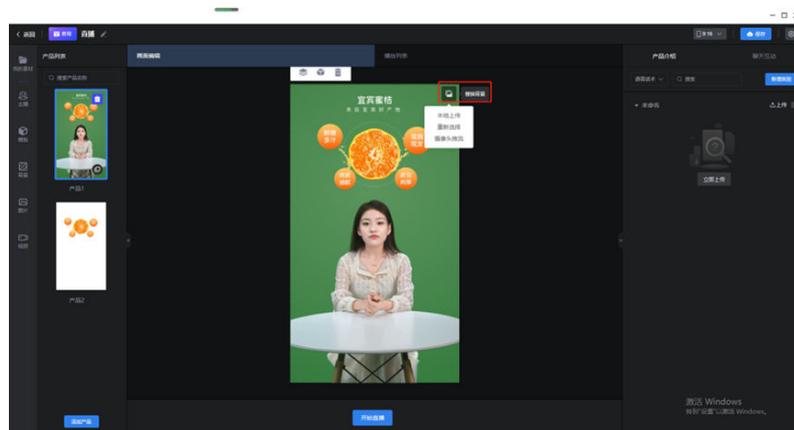
单击“”按钮，可删除当前选择图片，如下图：

图 4-48 图示 19



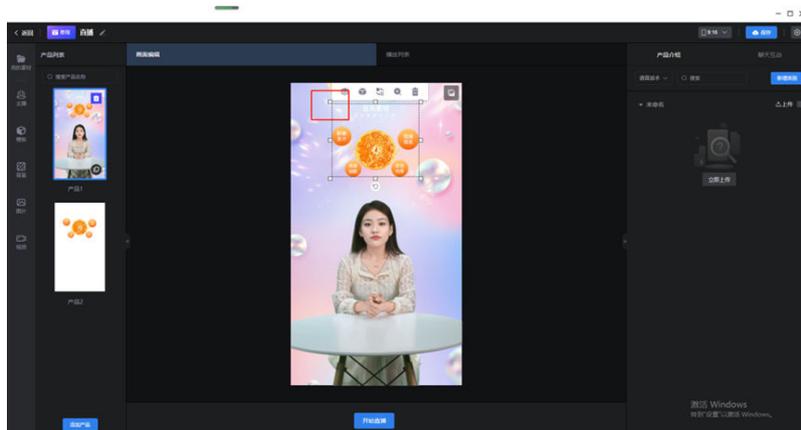
单击“”按钮，可以替换直播间背景，可以选择从本地上传以及重新选择系统提供的背景以及摄像头推流背景三种方式，如下图：

图 4-49 图示 20



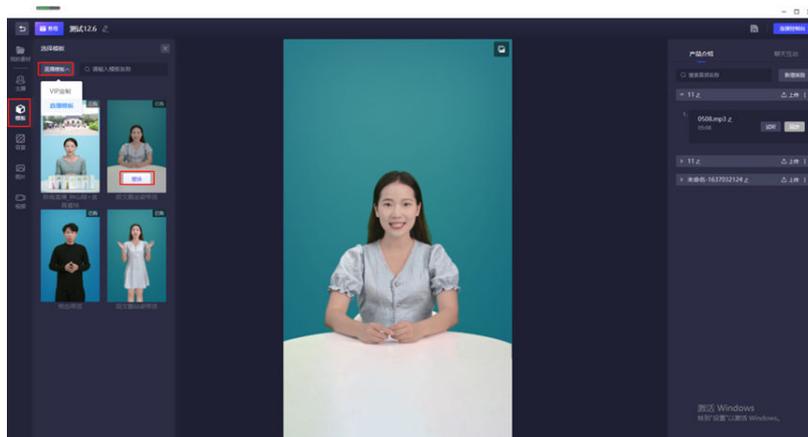
选中图层，可以通过鼠标单击节点，进行放大、缩小以及移动位置等操作，如下图：

图 4-50 图示 21



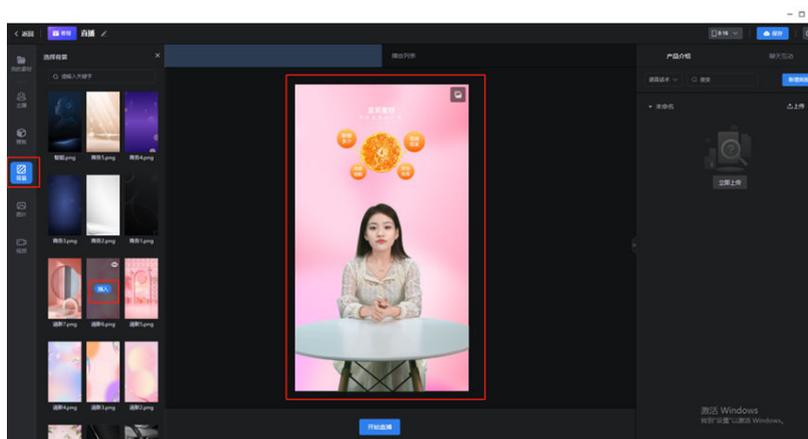
单击“模板”可选择直播模板作为直播场景，模板包含“VIP定制模板”以及“直播模板”，类别包含所有直播模特，单击模板“替换”按钮，可使用模板设置直播间，如下图：

图 4-51 图示 22



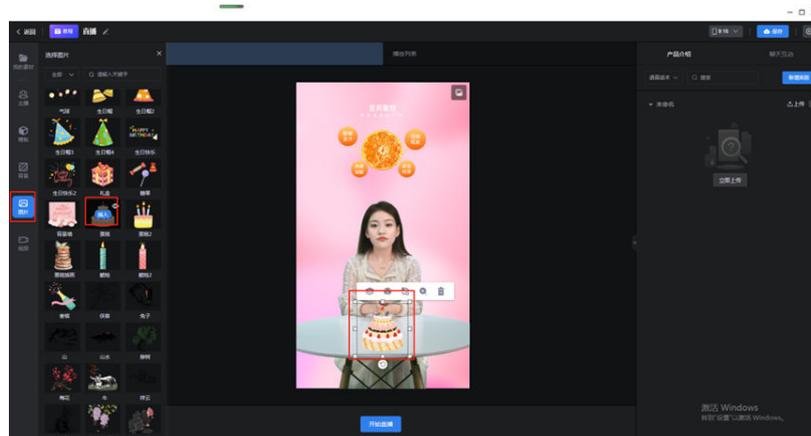
单击“背景”，可选择系统提供的背景图片进行插入、替换，如下图：

图 4-52 图示 23



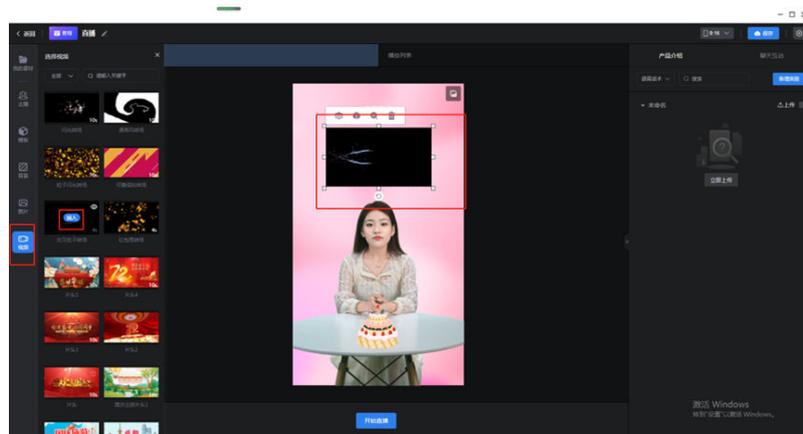
单击“图片”按钮，可选择系统提供的图片素材进行插入，并可在画布中选中该图片，进行调整图层位置、选择应用到全局，替换图片、查看图片以及删除、移动位置及调整大小等操作。如下图：

图 4-53 图示 24



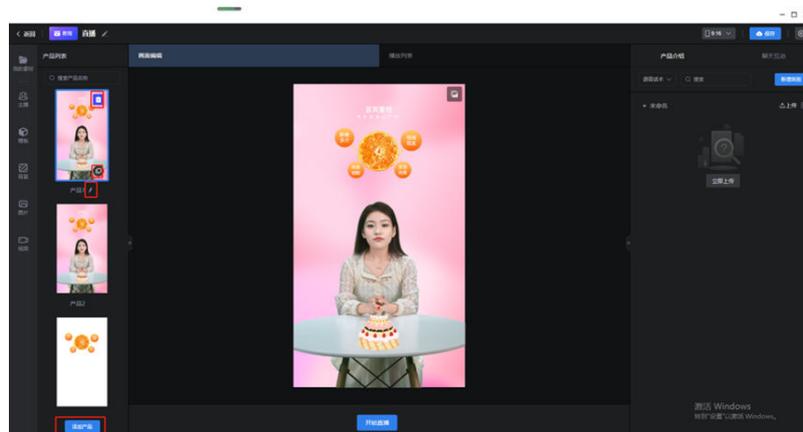
单击“视频”按钮，可选择视频插入直播间，并可在画布中选中该视频，进行调整图层位置、选择应用到全局、查看及删除视频、移动位置及调整大小等操作。如下图

图 4-54 图示 25



选中场景，单击场景上“”按钮，可删除该直播场景，单击“”按钮可复制该直播场景，或者单击下方“添加产品”按钮，可添加多个直播场景进行设置，并在直播时切换场景。单击场景名称下方按钮，可修改场景名称。如下图：

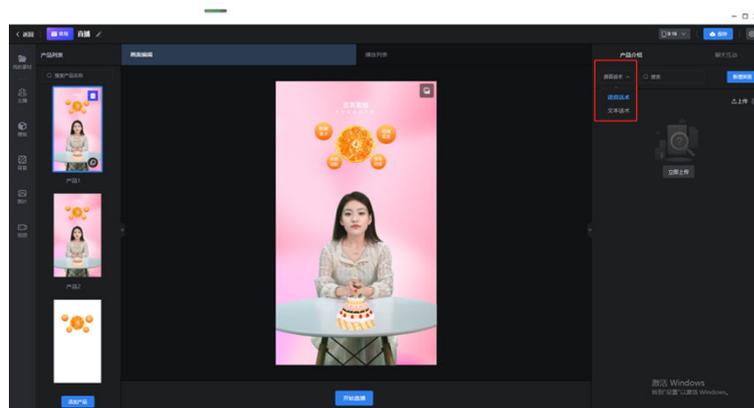
图 4-55 图示 26



### 步骤3 产品介绍

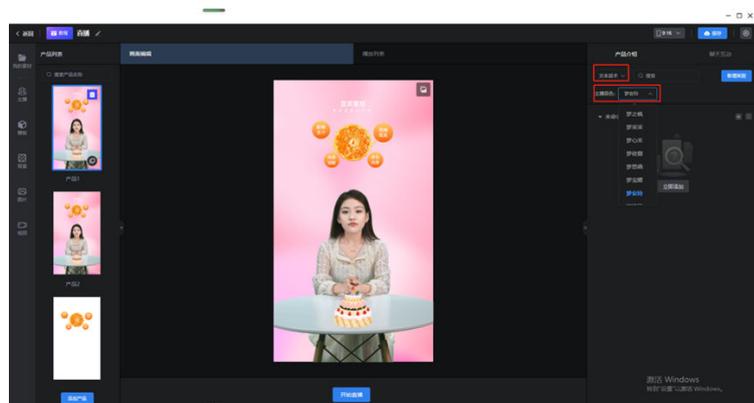
页面右侧可根据需要选择上传产品介绍音频，插入到播放列表中即可实现数字人自主直播。支持上传语音及文本合成两种方式。如下图：

图 4-56 图示 27



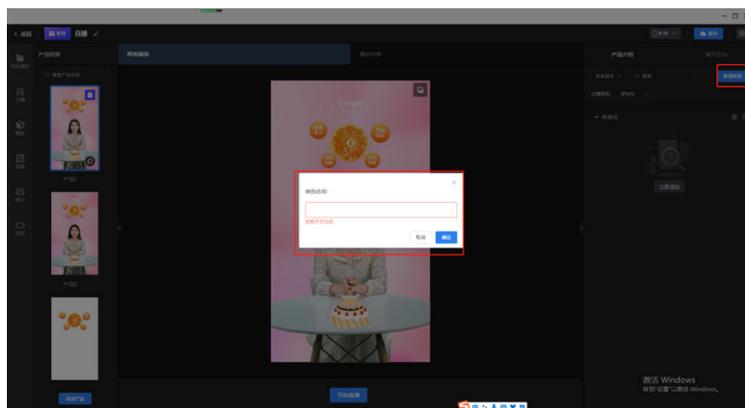
文本合成：单击主播音色可以选择合适的主播声音，如下图：

图 4-57 图示 28



单击“新增类别”可以添加不同类别的产品介绍语音，如下图：

图 4-58 图示 29



单击“”或者“立即添加”按钮，如下图，可以输入介绍的文本，如下图：

图 4-59 图示 30

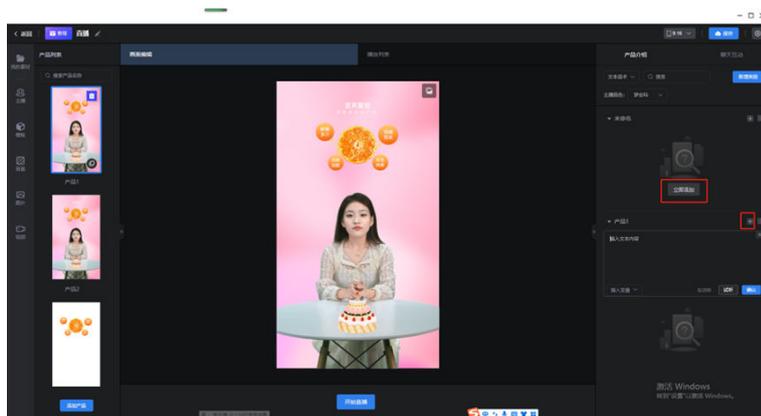
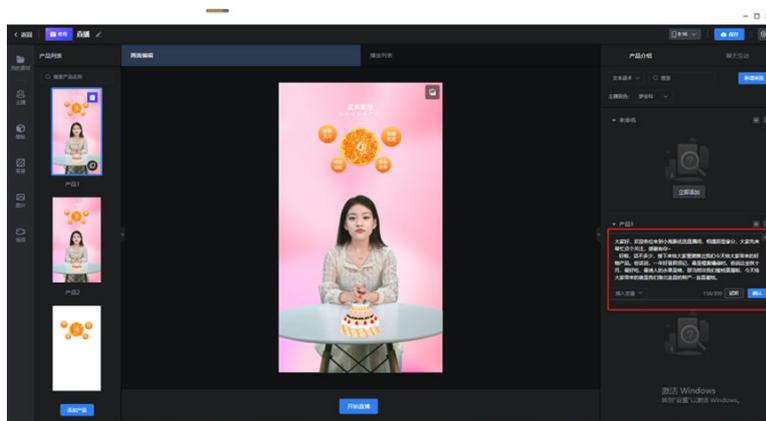


图 4-60 图示 31



单击“试听”可以试听该音频，单击“确认”即可生成音频文件，如下面两图：

图 4-61 图示 32

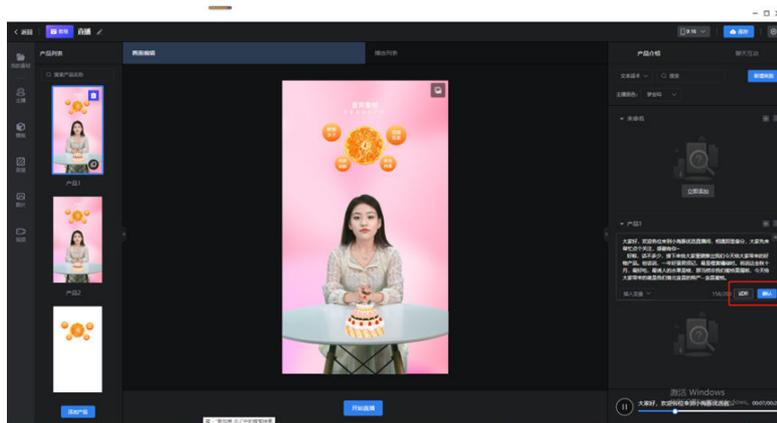
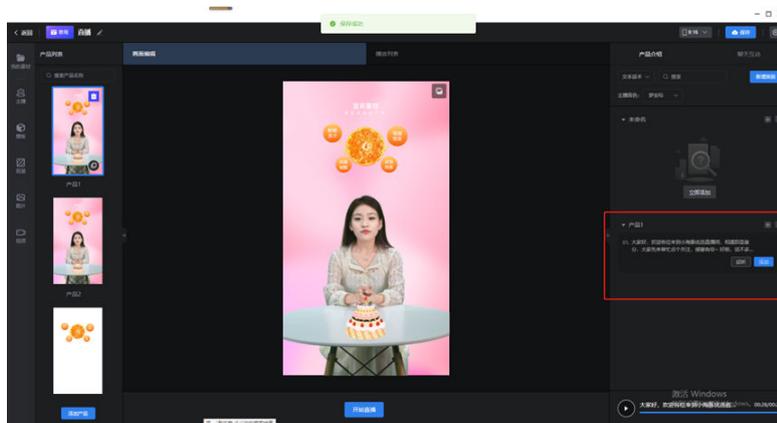
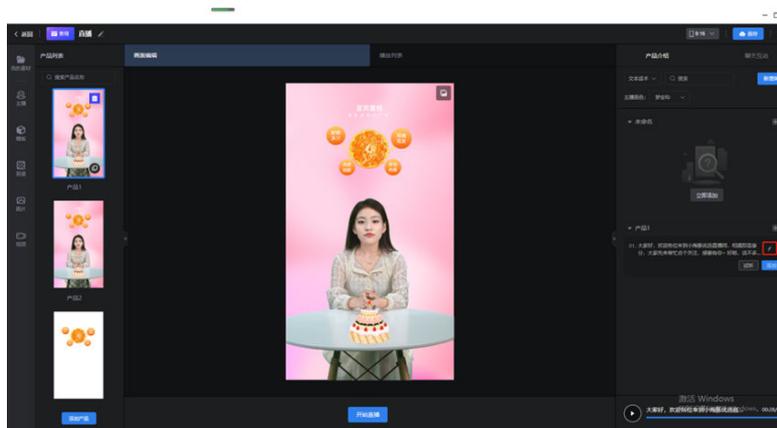


图 4-62 图示 33



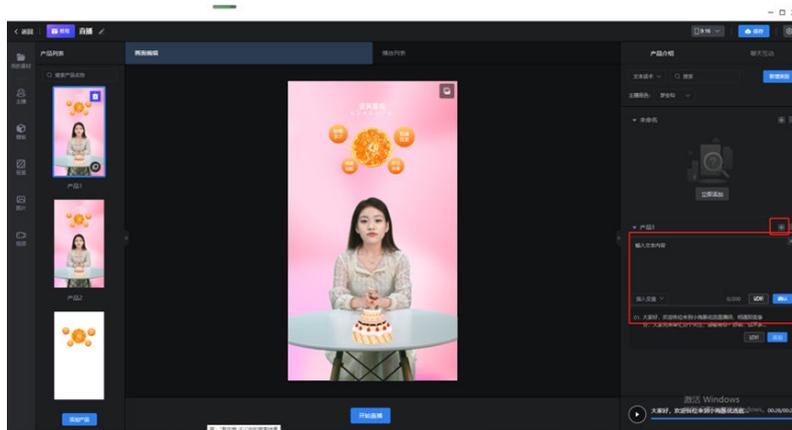
单击语音右侧“”按钮，可以进入文本编辑页面修改文本，修改完成后单击确认按钮即可根据修改后的文本生成语音，如下图：

图 4-63 图示 34



再次单击“”可在该类别下新增多个不同的本文内容，如下图：

图 4-64 图示 35



单击“确认后”可生成音频，如下图，通过上下拖动可移动顺序，如下图：

图 4-65 图示 36

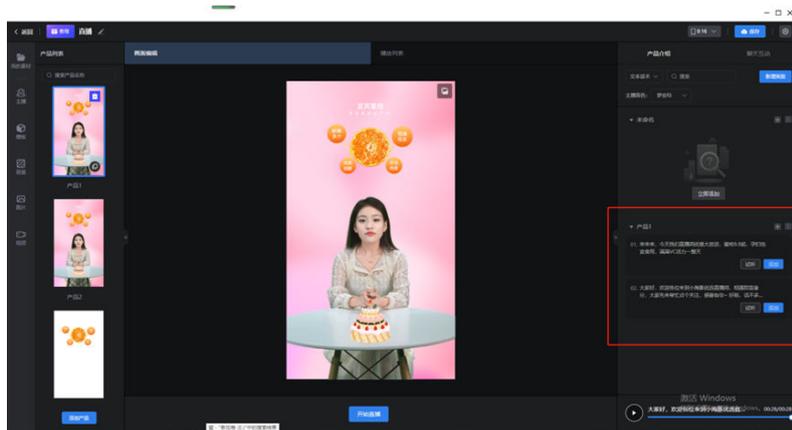
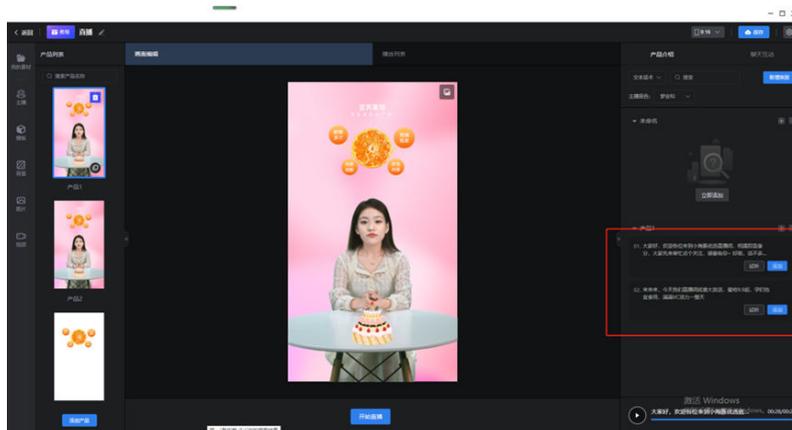
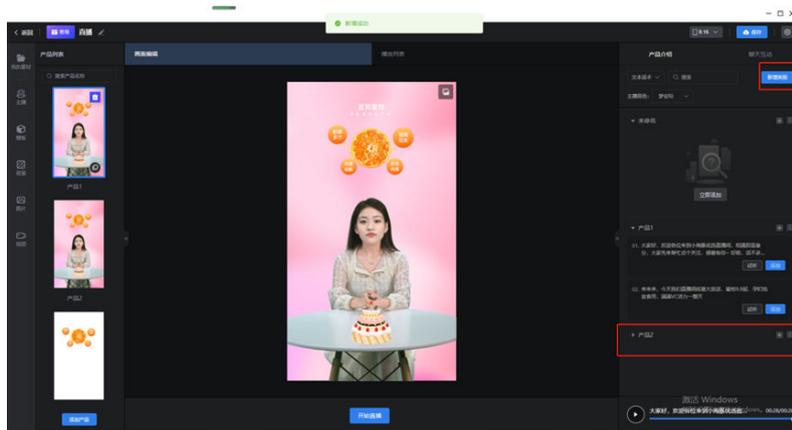


图 4-66 图示 37



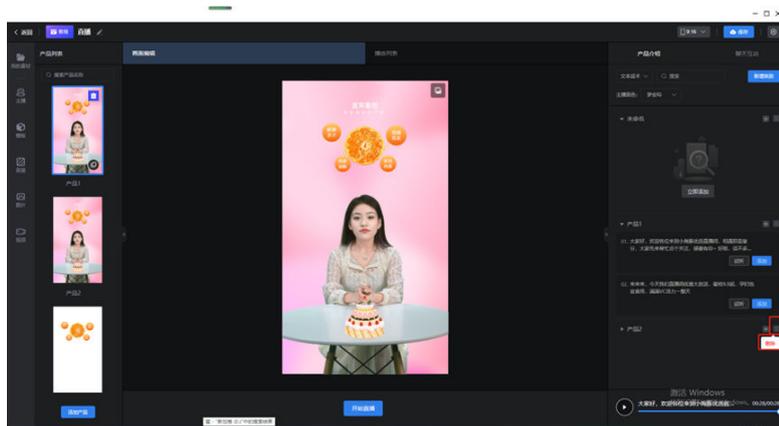
单击“新增类别”可新增不同的类别，如下图：

图 4-67 图示 38



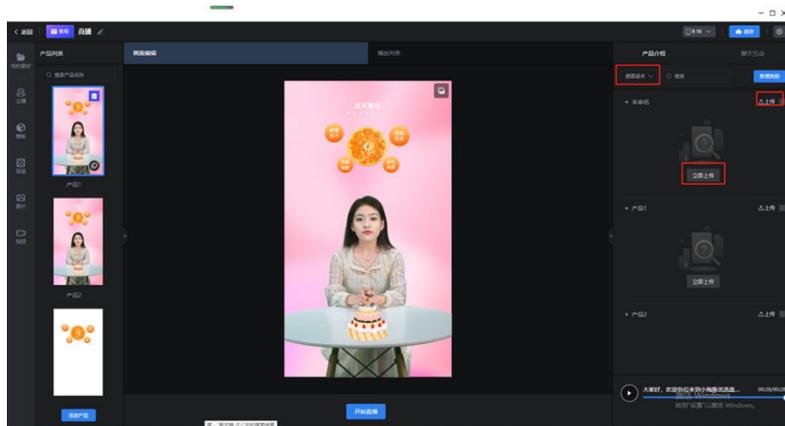
单击类别右侧“”按钮，可删除该类别，如下图：

图 4-68 图示 39



语音话术：选择类型“语音话术”，可从本地选择录音上传，如下图：

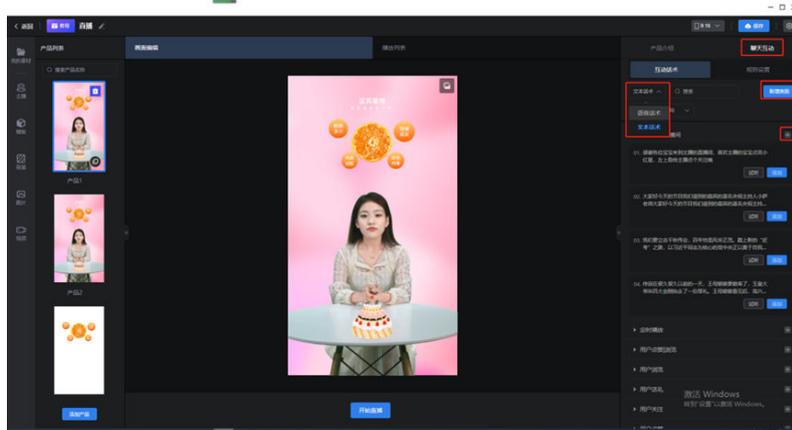
图 4-69 图示 40



#### 步骤4 聊天互动

参考“产品介绍”话术，聊天互动话术同样可以选择“文本合成”以及“上传语音”两种方式，可以设置不同的聊天互动话术，如下图：

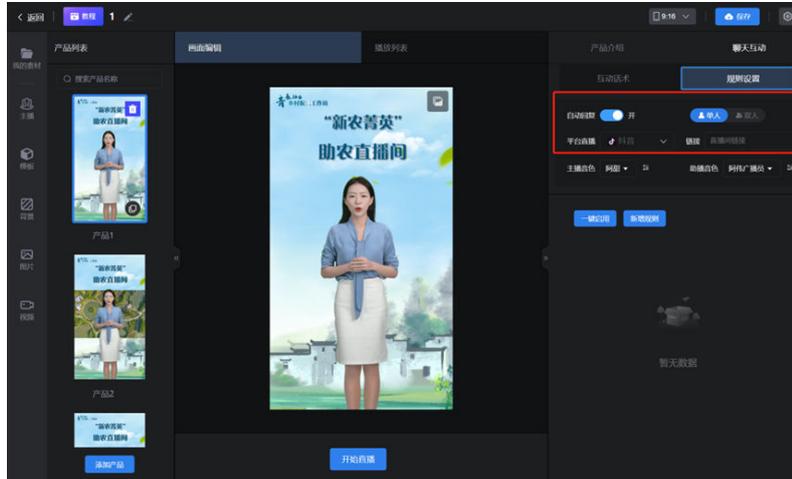
图 4-70 图示 41



### 步骤5 自动回复

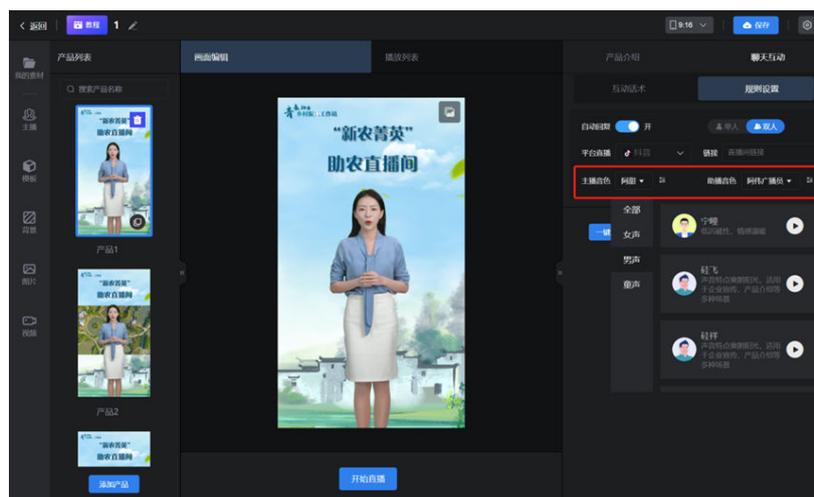
规则设置可以设置自动回复的规则，点开“自动回复”按钮，选择“单人直播”/“双人直播”，单人直播默认有主播进行回复，双人直播会营造一个小助理的角色，小助理无需露脸，会根据设置的自动回复规则定时引导用户点赞+关注等。选择直播平台并输入直播间地址，如下图：

图 4-71 图示 42



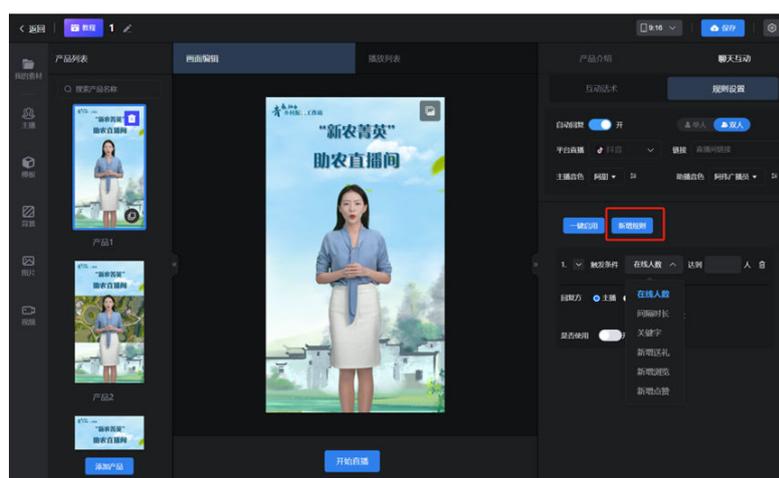
单击选择主播或助播的音色，单击播放按钮可以试听该音色，如下图：

图 4-72 图示 43



单击新增规则可以设置自动回复的规则，包括在线人数、间隔时长、关键字、新增送礼、新增浏览、新增点赞等规则，如下图：

图 4-73 图示 44



在线人数：可以设置直播间人数达到多少人时主播/助播回复的话术；

间隔时长：可以设置每隔多长时间主播/助播可以在直播间里说一些欢迎或者邀请点赞+关注的话术；

关键字：可以通过设置关键字，当直播间里有用户问问题时触发到该关键字，例如价格、多少钱等，主播/助播回复的话术；

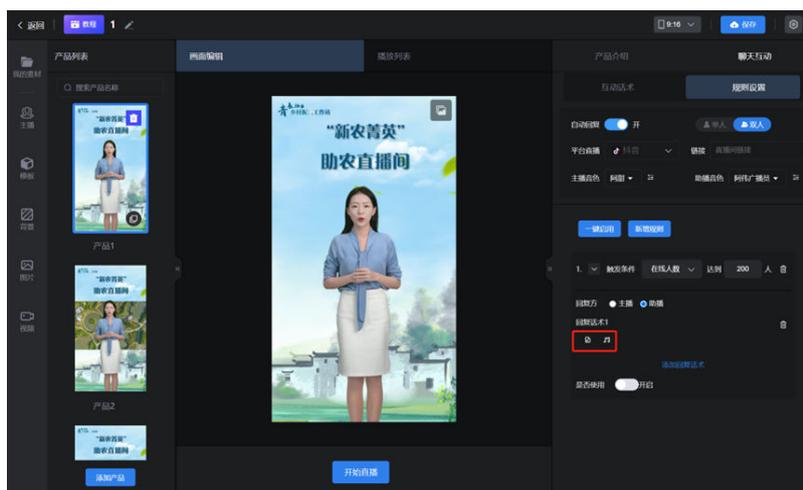
新增送礼：可以设置当直播间里有人送礼时，主播/助播回复感谢的话术；

新增浏览：可以设置当直播间里新增几个用户时，主播/助播回复的话术；

新增点赞：可以设置当直播间里有多少个人点赞时，主播/助播回复的话术

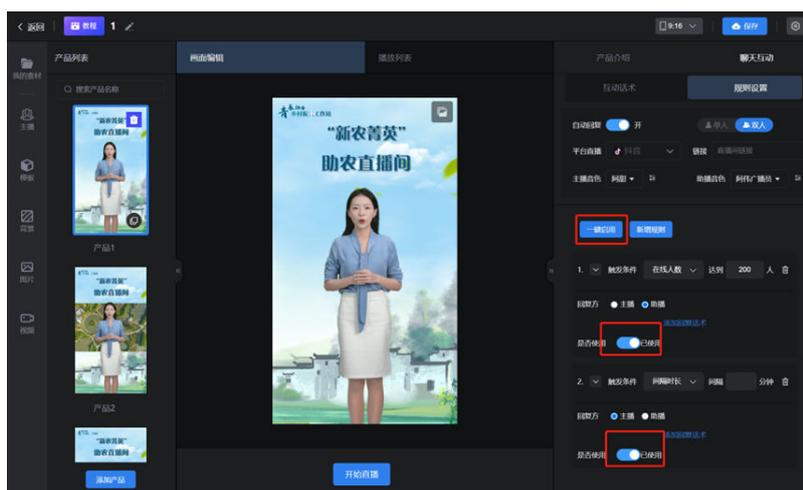
选择回复方主播或者是助播，单击“添加回复话术”，单击“”按钮，可以上传文本，单击“”可以从本地上传回复的音频，如下图：

图 4-74 图示 45



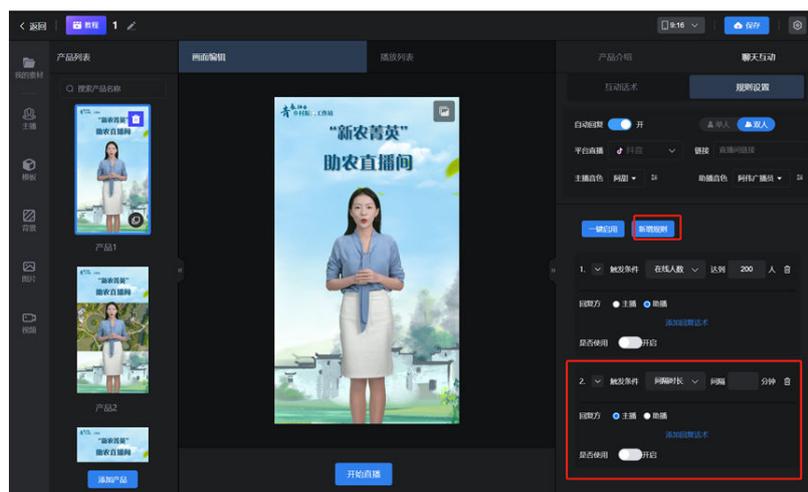
每个规则可以单独设置是否启用，也可以单击上方“一键启用”启用所有规则，如下图：

图 4-75 图示 46



再次单击新增规则，可以设置多个不同的规则，同一个规则也可以添加多个自动回复话术。如下图：

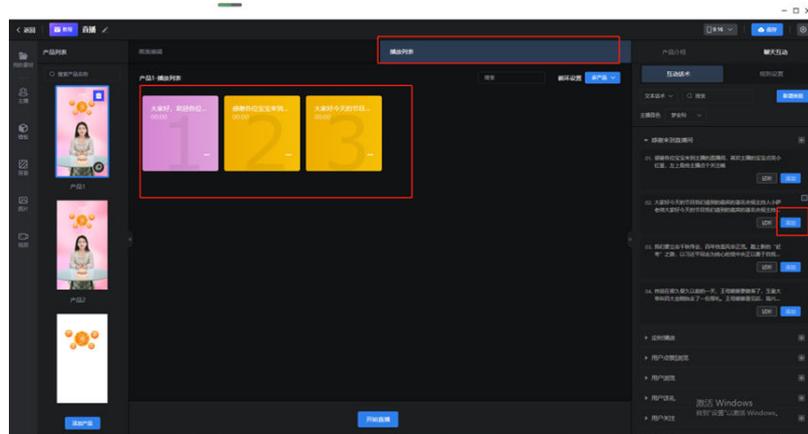
图 4-76 图示 47



### 步骤6 播放列表

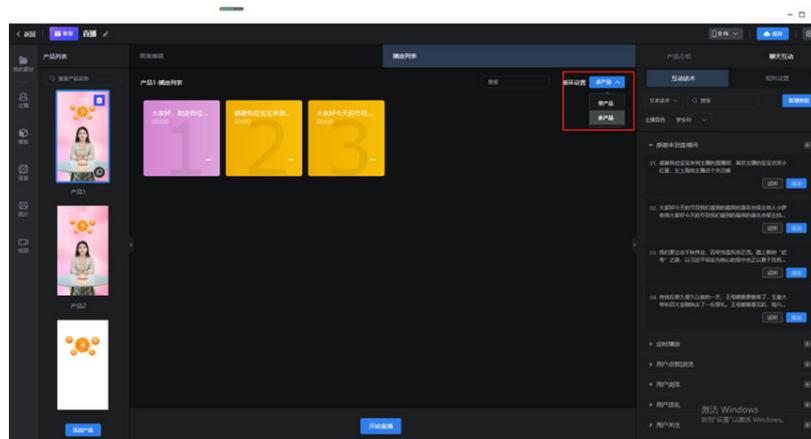
拖动或单击“添加”按钮，可以将左侧设置好的音频添加至“播放列表中”，如下图：

图 4-77 图示 48



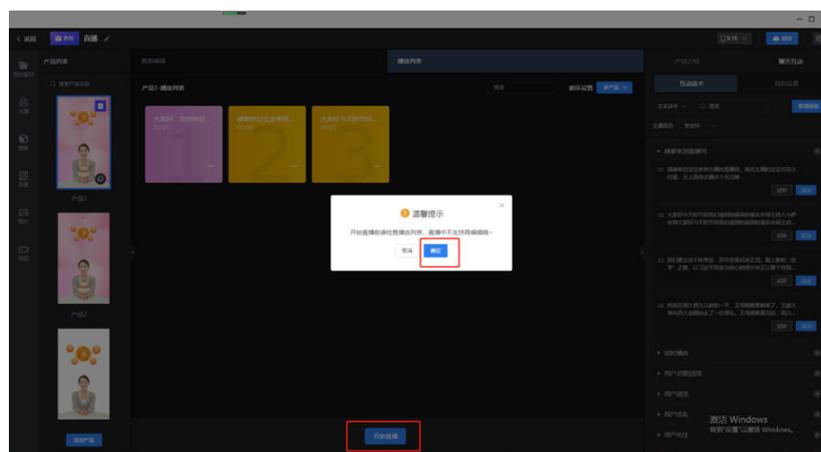
设置播放规则，包括单场景、多场景的循环播放，如下图：

图 4-78 图示 49



设置完成后单击“开始直播”即可进入到直播间，如下图：

图 4-79 图示 50



### 步骤7 使用直播模板创建直播

在首页，单击模板“使用”按钮，如下图，进入工作台，如下图。

图 4-80 图示 51

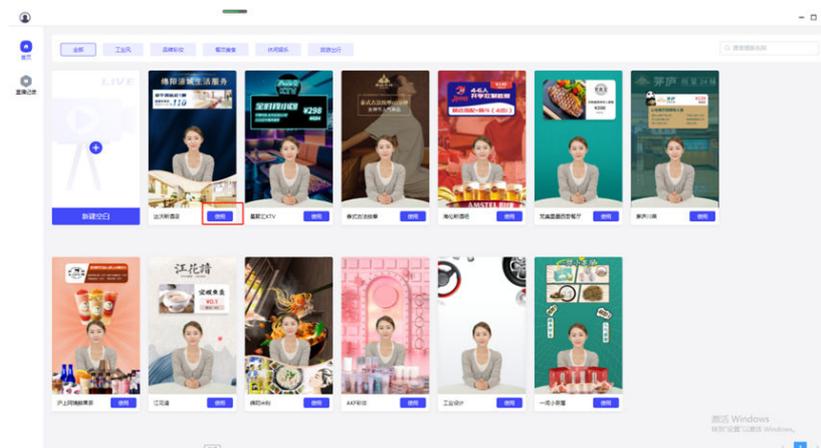
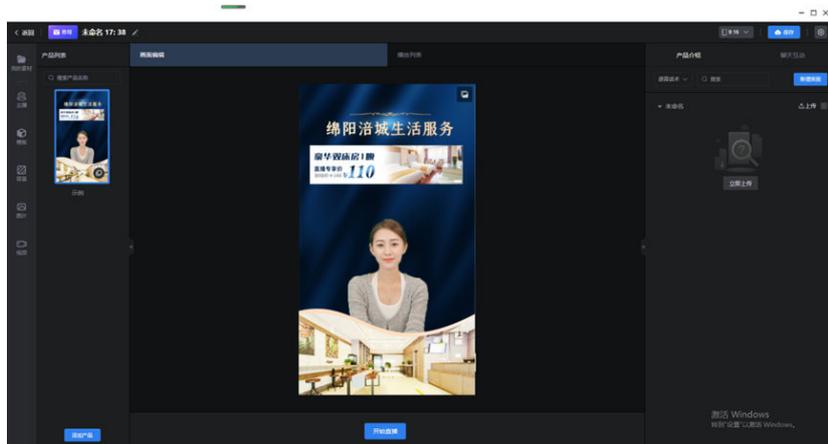
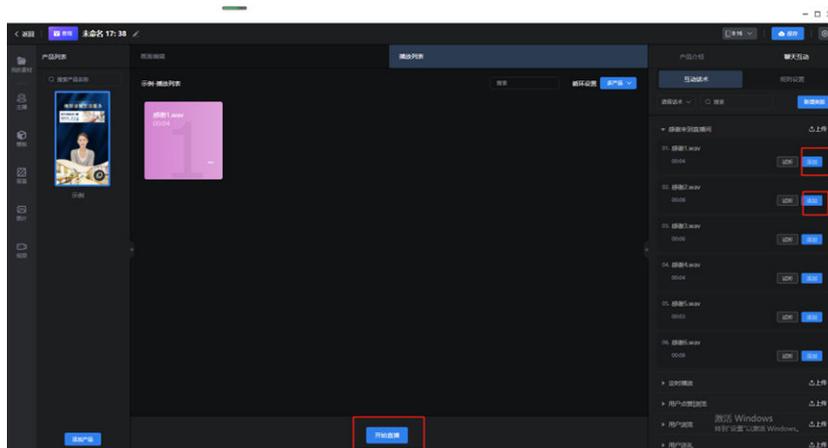


图 4-81 图示 52



模板的主播可以免费使用，主播、背景、素材等都是配置好的，如需修改素材图片可以参考插入图片的操作替换模板中的图片、背景等素材；如不需要修改，将互动话术添加到播放列表，单击“开始直播”即可开启直播，如下图：

图 4-82 图示 53



----结束

## 4.4 硅基数字人交互实施

图 4-83 样例 1



### 实施步骤介绍

1. 需求确定：根据业务的需要，客户确定好具体的业务交互场景。
2. 场景编辑：根据所确定业务场景，编辑好业务场景，特别是场景的背景及其它相关要素。

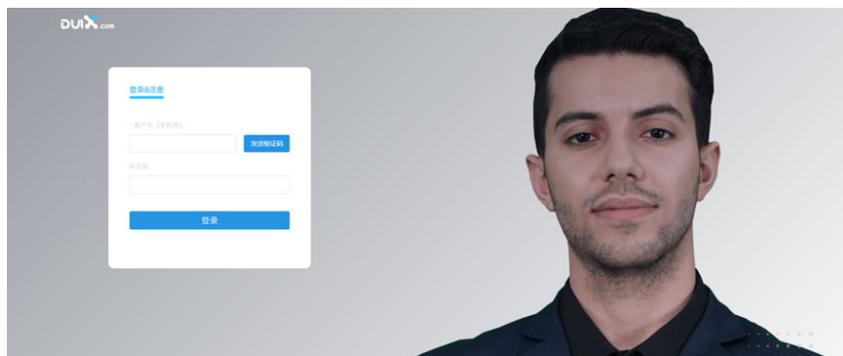
3. 模特选择：基于上述业务交互场景的确定，选择对应的数字人模特，模特选择可以是库里已经有的，也可以是前面已经帮客户定制好的数字人模特。
4. 交互配置：基于上述业务交互场景的确定，对知识库进行配置，比如问题和答案的对应关系配置，如果需要，还可以对接第三方的知识库系统。
5. 系统集成：在以上业务场景相关的配置完成之后，也需要完成系统集成，即把数字人交互系统集成到客户的系统中，比如说客户已有的手机APP系统。
6. 终端使用：上述工作完成之后，终端用户即可下载安装手机APP或者搜索对应的小程序进行使用了。

## 具体过程介绍如下

### 步骤1 登录

手机号验证码登录；（未注册手机号登录并注册）

图 4-84 样例 2



### 步骤2 首页

单击【会话演示】新窗口打开本地服务系统预设的会话demo，进行对话；

单击【创建会话】按钮，可跳转到创建会话界面；

单击【技术文档】按钮跳转到【SDK开发者平台】；

单击【交流论坛】按钮跳转到【交流论坛平台】；

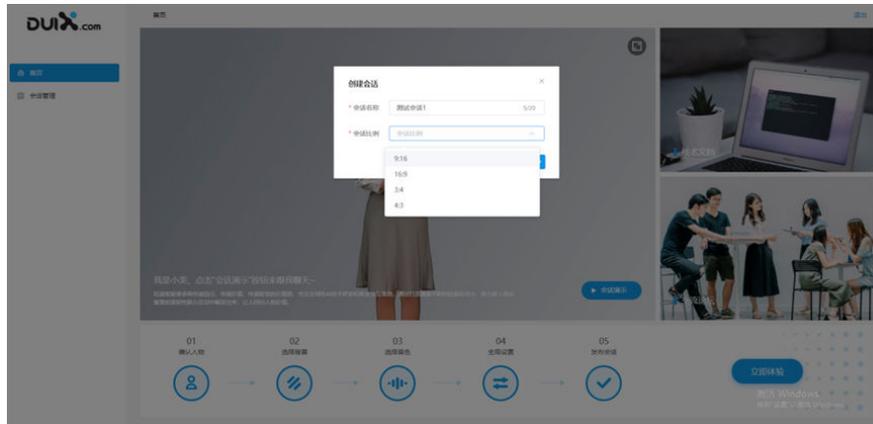
图 4-85 样例 3



### 步骤3 快速创建-确认人物

单击【立即体验】，显示【创建会话】弹出，输入会话名称及会话比例；

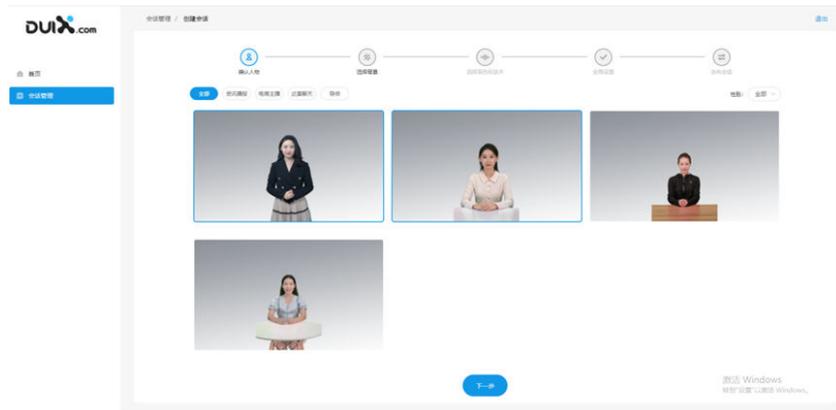
图 4-86 样例 4



### 步骤4 快速创建-确认人物

1. 单击选中数字人；
2. 单击[下一步]按钮，跳转到【选择背景】页面；

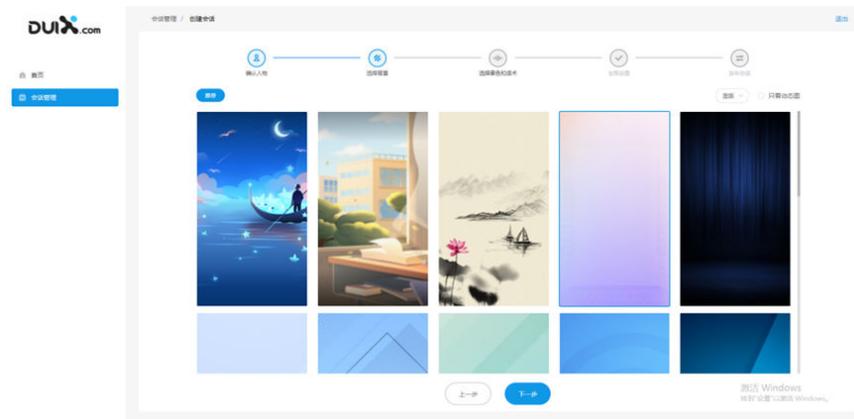
图 4-87 样例 5



### 步骤5 快速创建-选择背景

1. 支持横竖屏筛选，单击[只看动态图]，只显示动态图；
2. 单击选中背景图片；
3. 单击[上一步]跳转到【选择数字人】页面；
4. 单击[下一步]按钮，跳转到【选择音色和话术】页面；

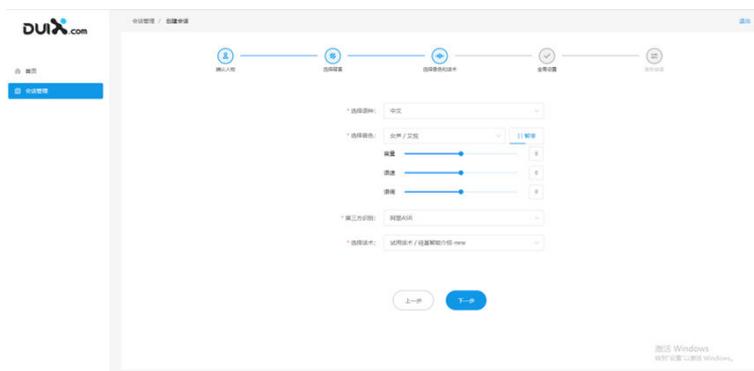
图 4-88 样例 6



### 步骤6 快速创建-选择音色和话术

1. 选择语种：  
默认中文，单击下拉菜单选择中文/英文；
2. 选择音色：  
默认一个音色，单击级联下拉菜单选择音色：男声/女声/童声一对应的音色；  
选中后显示音量和语速、语调设置；  
单击[试听]按钮，播放10s默认话术；
3. 第三方识别引擎：  
默认chrome识别引擎，单击下拉菜单选择；
4. 选择话术：  
默认试用话术/硅基介绍，单击级联下拉菜单选择话术，对应选择话术；
5. 操作按钮：  
单击[上一步]跳转到【选择音色和话术】页面；  
单击[下一步]按钮，跳转到【全局设置】页面；

图 4-89 样例 7

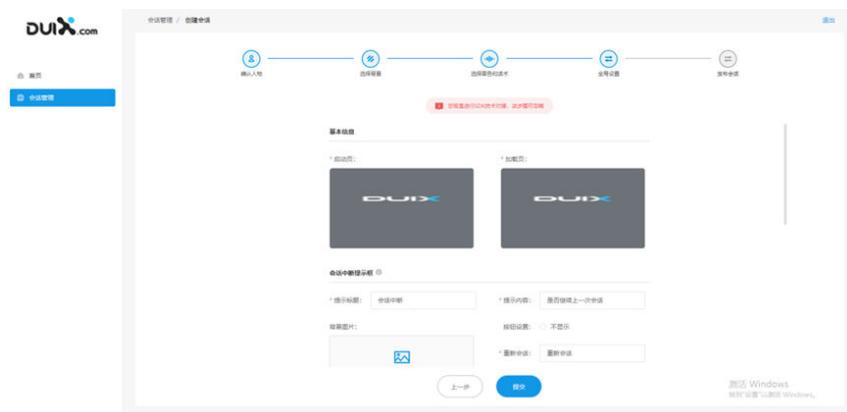


### 步骤7 快速创建-全局设置

1. 基本信息：  
启动页：必填，默认系统预设启动图，单击可选择上传本地启动图，上传成功后可查看大图或删除图片重新上传；

- 加载页：必填，默认系统预设启动图，单击可选择上传本地加载图，上传成功后可查看大图或删除图片重新上传；
2. 会话中断提示框：  
提示标题：必填，单击输入修改文案；  
提示内容：必填，单击输入修改文案；  
背景图片：非必填，单击本地上传图片，上传成功后可查看大图或删除图片重新上传；  
按钮设置：单击“不显示”，前端不显示按钮，反之不显示按钮；  
- 重新会话，单击输入框可输入修改文案（按钮事件重新加载会话）  
- 继续会话，单击输入框可输入修改文案（按钮事件继续上一次会话节点）
3. 会话结束提示框：  
提示标题：必填，单击输入修改文案；  
提示内容：必填，单击输入修改文案；  
背景图片：非必填，单击本地上传图片，上传成功后可查看大图或删除图片重新上传；  
按钮设置：单击“不显示”，前端不显示按钮，反之显示按钮；  
- 结束会话，单击输入框可输入修改文案（按钮事件结束会话并退出程序）  
- 重新会话，单击输入框可输入修改文案（按钮事件重新加载会话）
4. 报错提示语：  
会话拥挤提示：默认系统提示文案“我正在忙，待会儿上线”，单击输入修改文案；  
网络异常提示：默认系统提示文案“网络不太好，检查一下再来哦~”，单击输入修改文案；  
其他异常报错提示：默认系统提示文案“抱歉，打个盹儿，请稍后再试~”，单击输入修改文案；
5. 操作按钮：  
单击[上一步]跳转到【全局设置】页面；  
单击[下一步]按钮，跳转到【创建成功】页面；

图 4-90 样例 8

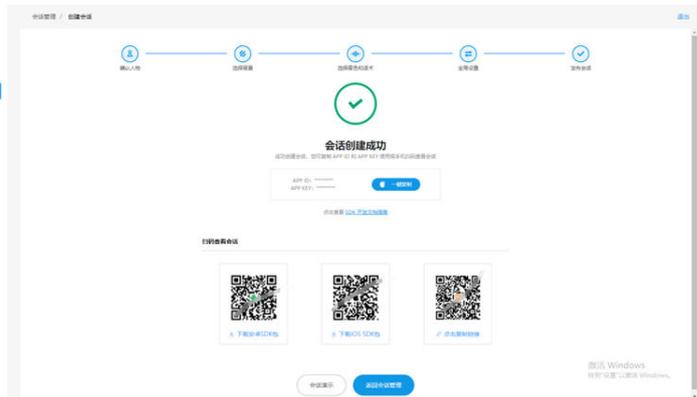


### 步骤8 快速创建-创建成功

1. 单击【一键复制】，复制会话APP ID\APP KEY，可进行SDK对接；

2. 单击SDK开发者文档指南新窗口打开【SDK文档平台】；
3. 手机扫码查看会话：  
生成对应的二维码，用户手机扫码后可安装对应安装包查看初始化数据的APP；  
单击下载安装包：单击跳转到对应的SDK开发文档；
4. 操作按钮：  
单击[会话演示]按钮，新窗口打开H5页面，显示已配置的会话数据；  
单击[返回会话管理]返回【会话列表】页面；

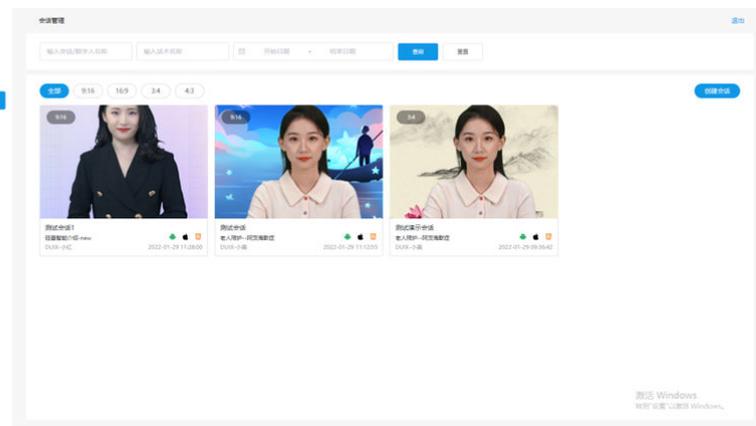
图 4-91 样例 9



**步骤9 会话管理**

1. 查询条件  
会话/数字人名称：关键字模糊查询；  
话术名称：关键字模糊查询  
更新时间：YYYY-MM-DD~YYYY-MM-DD
2. 列表操作：  
复制：单击[复制]按钮，复制APPid、APPkey；  
编辑：单击[编辑]按钮，跳转到【编辑会话】页面；  
删除：单击[删除]按钮，删除会话数据；  
会话演示：单击[会话演示]按钮，新窗口打开【会话演示】页面；
3. 创建会话：  
单击[创建会话]按钮，弹出【创建会话】页面；

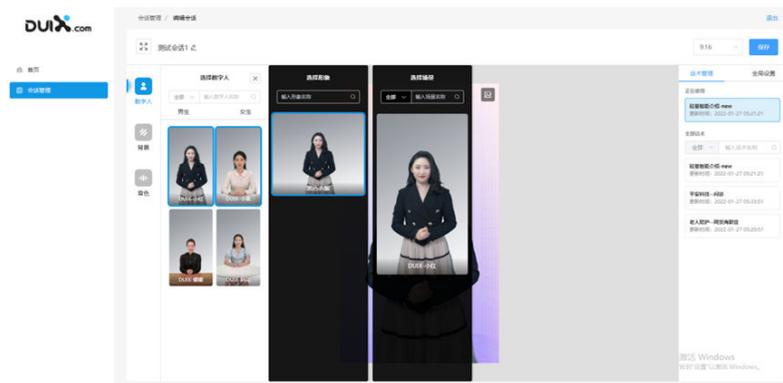
图 4-92 样例 10



### 步骤10 编辑会话—选择数字人

1. 头部区域：  
会话名称：单击可编辑输入新的会话名称；  
比例：单击可以选择切换比例，需二次确认操作；  
保存：单击“保存”按钮，保存修改数据，返回【会话列表】页面；
2. 左侧选择数字人  
选择模特-形象-场景，三级级联选择数字人，单击【插入】按钮，把数字人插入到画布中；

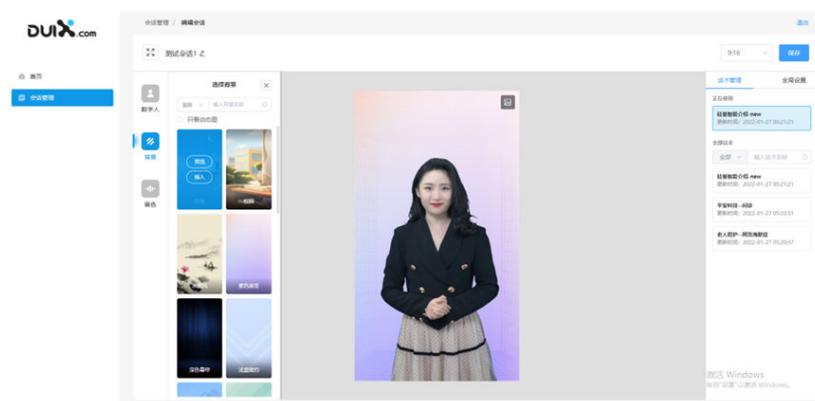
图 4-93 样例 11



### 步骤11 编辑会话—选择背景

1. 查询条件：  
背景名称：支持关键字模糊查询；  
只看动态图：单击[只看动态图]，只显示动态图；
2. 左侧选择背景图片：  
获取焦点显示[插入]按钮，单击按钮插入背景图片到画布；  
获取焦点显示[预览]按钮，单击按钮弹出显示大图并可以关闭；

图 4-94 样例 12

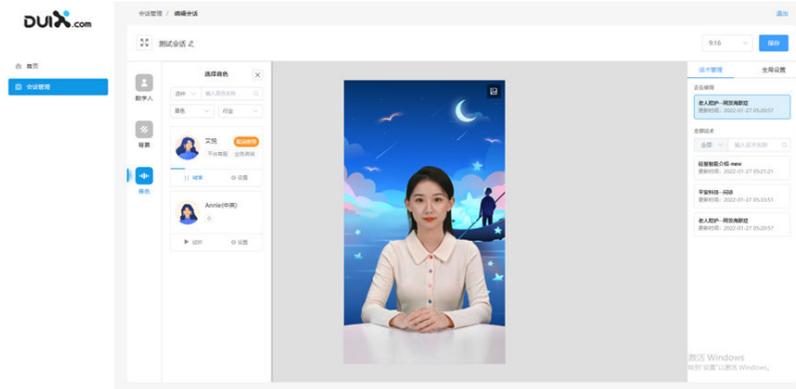


### 步骤12 编辑会话—选择音色

1. 查询条件：  
音色类型：默认全部；

- 音色名称：支持关键字模糊查询；
- 语种：全部/中文/英文；
- 声音：全部/男生/女生/童声；
- 2. 左侧选择音色
- 单击[试听]按钮，播放10s声音；
- 单击[设置]按钮，可调整音量、语速、语调；

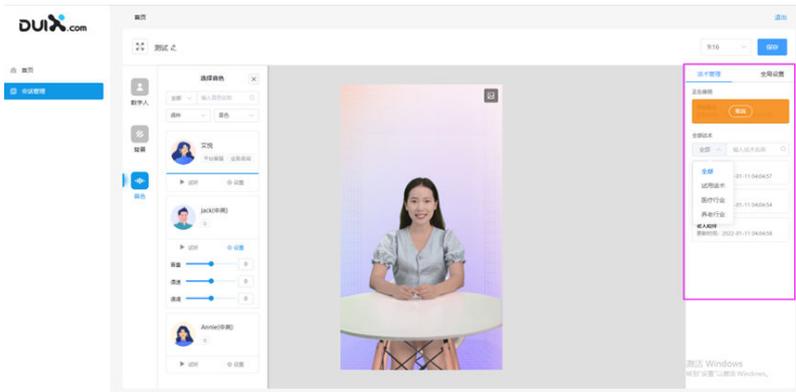
图 4-95 样例 13



**步骤13 编辑会话—选择话术**

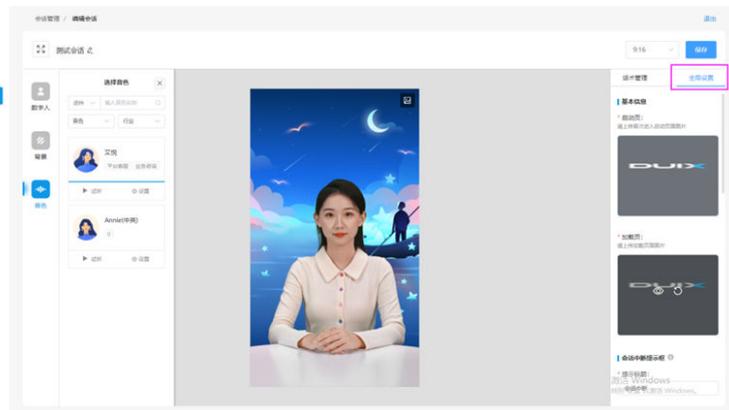
1. 正在使用：
  - 显示当前正在使用的话术，单击[取消]按钮，弹出操作确认提示框“取消当前话术后线上会同步更新，确认取消吗？”，单击“确认”按钮，提示“话术取消成功”；
2. 查询条件：
  - 话术行业：全部/医疗/养老，默认显示全部；
3. 操作按钮：
  - 取消：获取焦点时，显示[取消]按钮，弹出操作确认提示框“确认取消使用中话术吗？”，单击“确认”按钮取消使用中话术；
  - 使用：获取焦点时，显示[使用]按钮，弹出操作确认提示框“当前已有使用中的话术，确认使用该话术替换吗？”，单击“确认”按钮替换话术；

图 4-96 样例 14



**步骤14 编辑会话—全局设置**  
可修改已编辑的全局设置

图 4-97 样例 15



----结束

# 5 修订记录

表 5-1 修订记录

发布日期	修订记录
2023-05-19	第一次正式发布