

拟公示算法机制机理内容

算法名称	喜马拉雅语音驱动数字人算法
算法基本原理	<p>喜马拉雅语音驱动数字人算法基于深度学习 GAN 和 Transformer 结构。</p> <p>算法的学习输入为真人播讲视频，通过深度学习训练，使网络学习到真人形象的生成能力，并建立语音和数字人头部表情、口唇和身体运动的对应关系。</p> <p>生成数字人视频时，算法的输入为一段语音，该段语音通过经过学习的神经网络，输出数字人的播讲视频。</p>
算法运行机制	<p>算法服务运行分为两阶段：</p> <p>数字人形象学习：用户提交需定制形象的正面播讲视频，在提交过程中需验证视频中为用户本人；将该视频输入到算法，经过算法模型的训练，得到用户的数字人神经网络模型，该模型对应了用户提交形象的数字人。</p> <p>数字人语音驱动：语音特征提取模块接受语音数据，并提取得到语音特征，再将语音特征输入到经过训练的数字人神经网络模型，该模型将语音特征对应到数字人的口唇、表情、动作上，最终输出数字人播讲视频。</p>
算法应用场景	主播数字人短视频创作工具（珠峰有话说）

算法目的意图	辅助创作者生产虚拟数字人短视频
算法公示情况 (选填)	