

Hackathon比赛项目需求文档

一、前期准备情况

1、可应用技术

Mask Detecting Technique: 口罩检测技术

Speech Recognition Technique: 语音识别技术

2、已做工作

二、项目应用场景及相关需求

1、基于语音识别技术的在线教学实时字幕功能

①应用背景

②项目Demo需求

基本功能

额外功能

2、基于语音识别技术和口罩检测技术的防疫门禁功能

①应用背景

②项目Demo需求

基本功能

额外功能

Hackathon比赛项目需求文档

一、前期准备情况

1、可应用技术

Mask Detecting Technique: 口罩检测技术

通过机器学习，对模型进行训练，可初步进行人是否佩戴口罩的识别工作。

Speech Recognition Technique: 语音识别技术

通过调用接口，可实现对输入语音的识别。

2、已做工作

①创建口罩检测模型，并进行了小规模训练，目前识别能力尚弱；

②调用百度语音识别API，进行语音测试后发现识别能力尚佳；

③搭建、学习Electron平台，用于生成GUI.

二、项目应用场景及相关需求

1、基于语音识别技术的在线教学实时字幕功能

①应用背景

2020年新冠疫情的肆虐，让网络会议、在线课堂的利用率达到了空前之高。在疫情期间，几乎所有上班族、学生群体都是通过各种视频会议或在线课堂软件来组织会议或进行在线教学的。然而，使用这些软件数月以来，我们也发现了一些问题：对于一些内容较多、需要不时记录的会议或课程，在发言人语速较快或带有方音时，就会给参会者的理解和记录带来一定的困难；此外，如果身处不便于播放声音的环境或是嘈杂环境，无法听清发言者的语音，同样会使直播效果大打折扣。

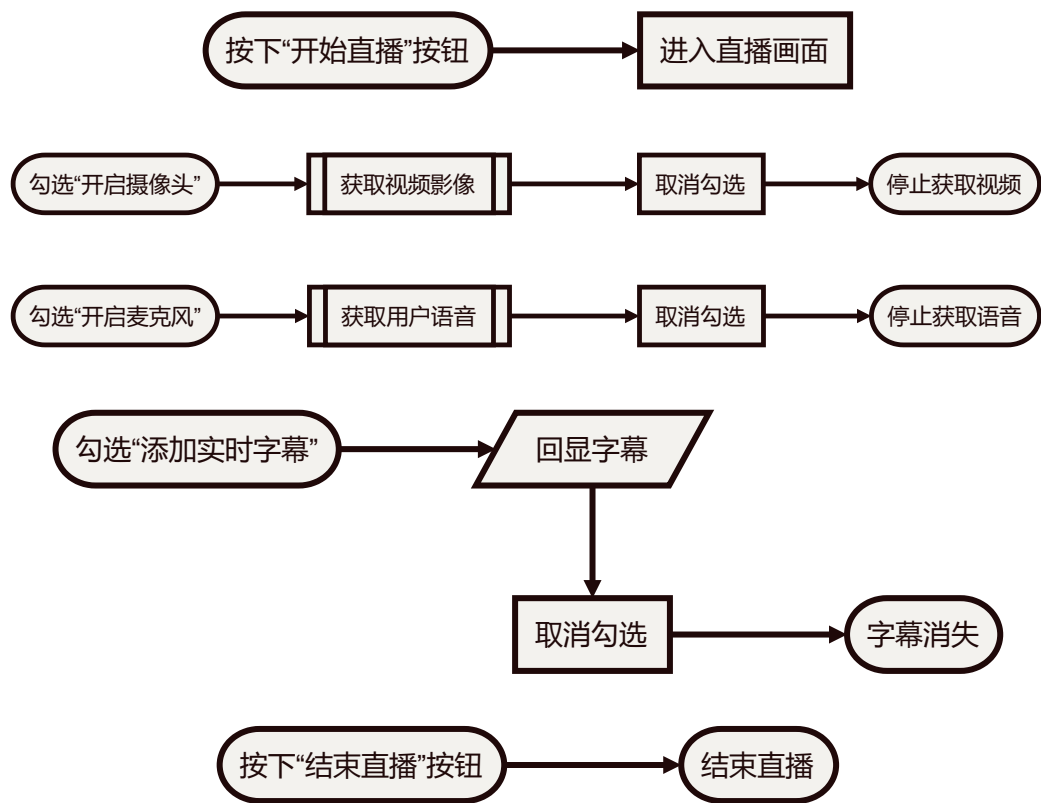
因此，产生了所需的一个功能：为会议/网络课堂直播添加实时字幕。目前相关的语音识别技术API已经比较成熟，尤其是中国有关公司如百度、讯飞等提供的语音识别技术，不仅识别率高，而且支持多种方言、多种语言，然而在网络会议、教学的场景中，鲜有实时字幕的功能。尽管目前市场上已有相关的公司开发了此功能，如Google的Live Caption插件已经支持在多个平台上使用，但是由于地区限制和语言限制，我们还无法在日常生活中正常使用它。

中国目前仍有多地疫情形势紧张，世界上亦有许多国家疫情形势严峻，需要使用在线视频会议软件进行开会或授课，考虑到以上问题，开发一项为直播添加字幕的功能具有一定的实际意义。

②项目Demo需求

基本功能

- a.可以获取用户的实时视频影像（摄像头）；
- b.可以获取用户的语音（麦克风）；
- c.在主界面，用户选择“开始直播”按钮后，进入直播界面；
- d.画面下方应有可勾选的功能“添加实时字幕”，可勾选的“开启摄像头”和“开启麦克风”功能，默认均为不勾选；
- e.勾选“开启摄像头”，开始获取用户影像，显示到画面上，取消勾选，则停止获取用户影像；勾选“开启麦克风”，开始获取用户语音，取消勾选，则停止获取用户语音；若选择“添加实时字幕”功能，软件开始将获取的语音输入到“语音识别”模型中，将模型识别出的结果实时回显到视频画面下方区域。回显字幕时，有两种可行方式：
 - I. 视频画面下方固定方块区域（可用深色区域）内用于放置多句字幕，当方块区域空间占满后即清空，以便后续字幕的继续写入，当勾选“添加实时字幕”功能后，画面中出现该方块区域；
 - II. 视频画面下方特定区域内只用于放置一句字幕，语音识别技术每识别好一句即回显到画面上，当下一句话识别完成后，前一句字幕清空，并回显当前识别结果；
- f. 画面下方应有“结束直播”按钮，用户按下后，退出直播画面，软件结束获取用户摄像头、麦克风。



额外功能

- 识别多种方言：勾选“添加实时字幕”功能后，附带选择按钮，可选择普通话、粤语、东北话、四川话、河北话、天津话、西安话.....等多种方言识别；
- 识别其他语种：可选择英语、日语、汉语等多语种识别；
- 会议记录：将语音识别出的每句话分行记录到本地文档中。

2、基于语音识别技术和口罩检测技术的防疫门禁功能

①应用背景

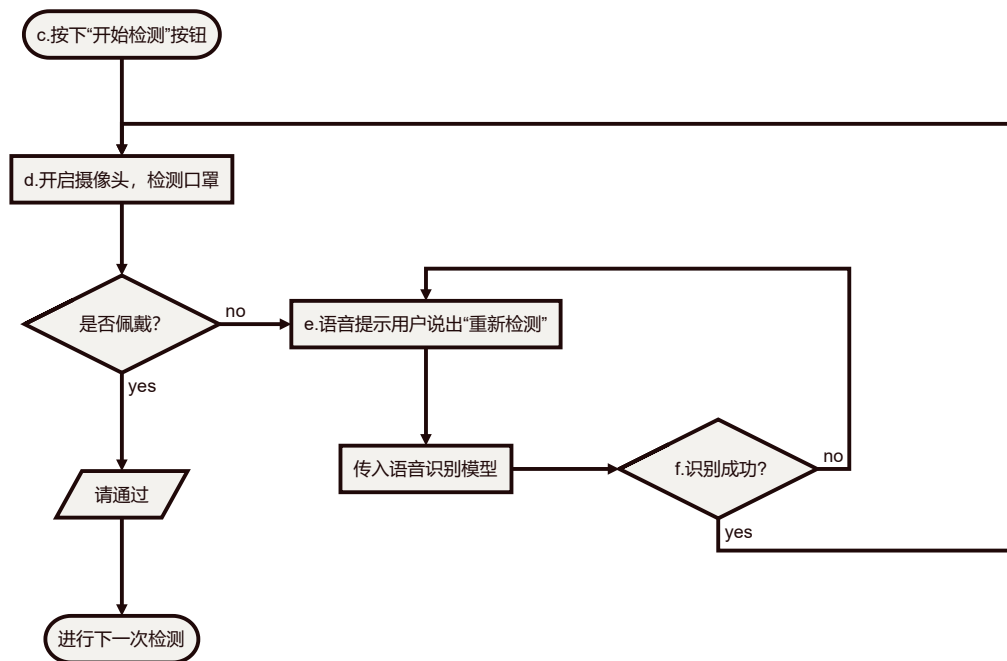
2020年新冠疫情的肆虐，让从前并无佩戴口罩习惯的我们，都自觉佩戴上了医用口罩。在疫情期间，我们进入任何公共场景，都需要佩戴口罩，既是为了自己的健康，也为国家联防联控贡献一份力量。然而，疫情持续近一年来，我们也发现了戴口罩带来的一些不便：对于眼镜佩戴者来说，在冬天戴口罩会让镜片蒙雾，对日常生活产生一定影响；在炎热的夏天，佩戴口罩让本就闷热的夏天更加难过；而医用外科口罩普遍偏紧，耳朵、面部又常常留下红印.....但无论多么艰难，在疫情防控尚未结束之时，我们都不应该摘下口罩。所以，在后疫情时代，加强检测公共场所进出人员是否佩戴口罩的工作尤为重要。那么除了工作人员站岗外，这项工作是否可以用软件来完成呢？

因此，产生了所需的一个功能：利用口罩检测技术并辅以语音识别技术，实现内置防疫门禁系统的软件。软件不仅能够检测是否佩戴口罩，还可以实现对进出人员的记录。这对于解放人力和病例溯源具有一定的实际意义。

②项目**Demo**需求

基本功能

- a.可以获取实时视频影像（摄像头）；
- b.可以获取实时语音（麦克风）；
- c.在主界面，选择“开始检测”按钮后（默认开启前置摄像头），进入视频画面界面；
- d.检测到画面中人脸后，传入“口罩检测”模型，传出检测结果。
 - 若检测结果为“未佩戴”，在收到下一次用户传入的“重新检测”语音之前，停止向“口罩检测”模型输入人脸数据；
 - 若检测结果为“已佩戴”，在视频画面中显示“请通过”，并向门禁系统发送“通过”信号，人员可通过；
- e.若检测结果为“未佩戴”，输出语音“请佩戴口罩，佩戴完成后，请说出‘重新检测’。”
- f.接收用户传入的“重新检测”语音信息，并传入“语音识别”模型，若语音识别出“重新”、“检测”相关信息，即发送信号给“口罩检测”模型，告知可以开始检测人脸口罩（回到d步骤），否则再次提示用户；
- g.重复d~f步骤，直至该受检测人员通过门禁；



额外功能

检测口罩是否佩戴正确（佩戴歪、佩戴了不合规的口罩等）。