2021.3.30

运行VisualizeLip.py的input和output

终端下

“ python VisualizeLip.py -i in.mp4 -o out.mp4

问题1：距离过远时（大概人脸占屏幕比例小于十分之一..?），画框不准确（框了半张脸）

设想：识别装置距离用户半米到一米，不要过远，否则提示用户靠近

问题2： 处理效率不高。目前来看，大概20s视频，完全导出需要30-40s；5min视频，导出了近15-20min

设想：输入和处理的同步化！！！（知识盲区

限制用户的输入时间，比如10-15s..??如果未再规定时间内完成应答，则refuse

问题3：导出视频的码率较低，且fps必须固定为25-->如何根据输入视频的fps决定输出的fps??（暂未解决

目前解决：利用终端下ffmpeg对输入的视频帧率进行修改

“ ffmpeg -i in.mp4 -r target\_fps(25) out.mp4

下一步学习方向：skvideo和ffmpeg的应用（网上的教程有点少。。。

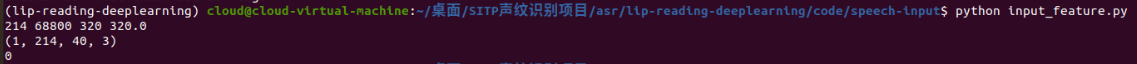
2021.3.31

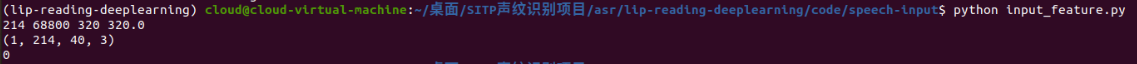
运行input-speech模块

测试文件需为单个.wav 放置在code/speech-input/Audio/subject/ 目录下，命名为sound.wav

目前测试的处理音频长度在10s以内

处理速度ok





下一步：音频中提取的特征参量的含义与利用