## Ctalk\_coding demonstration

**简介**

本项目利用手机自带的前置摄像头获得用户练习公众演讲（public speaking）的音视频数据，从中提取整合信息以评估用户每次演讲的效果。我们搭建的APP提供对用户演讲的点评与打分，并提供标注出扣分点的音视频回放，以满足用户对自身公众演讲练习-评估-改良的需求。

**代码实现**

**音频处理\_人为定义评估标准：**

**语速分析-**想用语音识别进行切词-（暂未实现）

**特征识别-**自相关，断续程度，总体音量与环境音量对比；某些需要切词的功能（0.005537253694581281）：发音准确程度，咬字效果，单语速单纯重复，语速控制不当等（暂未实现）。

**特征标注-**利用app的实时弹幕功能，对用户的公众演讲做出in-real-time feedback；同时将相关弹幕整合到记录视频中，供用户回放分析用。

**特征打分-**首先将声音离散成01序列，分析时间窗口中语音信号的特征，并给予其量化打分。

**无关因素-降噪：**利用环境噪音及谱减法，获得降噪效果；同时环境音量大小与音频本体的声音强度比值，作为一个重要的量度标准。

**Confident or not-**利用监督学习，利用带tag样本对Score（*pattern*）进行学习，从而获得较为合理与个性化的评判标准。

**视频处理**

**特征识别-**利用已有的面部特征识别数据库与识别算法，识别半身照视频中的脸、眼睛、嘴等面部特征，并得到其大小，尺寸等信息。

**特征标注-**对识别出来的特征进行标注；并将相关标注整合到记录视频中，供用户回放分析用。

**特征打分-**面部表情的变化/用户在摄像机视野中的移动程度作为模型的input，output为用户该次公众演讲的自信程度；利用监督学习，利用带tag样本对Score（*pattern*）进行学习，从而获得较为合理与个性化的评判标准。

**无关因素-用户与摄像头的相对位置与角度的校正：**引入起点校正机制，并利用用户自定义的体型信息，做到自定义化的运动校正。

**整合**

整合音频处理结果与视频处理结果，利用监督学习得到最好的Score（*pattern*），以期提升用户的公众演讲能力。

