

## 會議記錄

專案名稱	尚未確認		
系統名稱	尚未確認		
文件編號	250811-小組第六次會議		
會議主題	小組第六次會議	編輯者	蕭伊庭
會議日期	2025 年 08 月 11 日	編輯日期	2025/08/11
會議地點	Teams 線上會議		
出席人員	保羅(Paul)、林湘凌(Cecilia)、蕭伊庭(Sherry)、周易暉(Ethan)、羅靖傑(Morris)、金殿倫(Darren)		

### ■ 本次會議討論事項：

#### 壹、上週分工探討結果

【stt/tts-蕭伊庭(Sherry)&周易暉(Ethan)負責】

有完成簡易 stt 和 tts 的初步嘗試，分別使用 whisper 和 TTS 套件：stt 字音正確但字型準確率不高；tts 則是語言和聲線彈性不多，可能要再找找支援輸出中文的套件。

保羅(Paul)建議：

程式碼都可以再優化成模組的方式，更好的是寫成物件，最後或許可以做成我們自己的超物件更實用。有達到預期的效果可以再繼續深入研究，也可以多嘗試各種參數，試著用關鍵字去 google 查，網路上應該很常見相關的教學。

【avatar-金殿倫(Darren)負責】

目前找到的都需要付錢，另外有類似功能的中國軟體是離線的，但使用在專題可能比較不好，[HeyGen 網站\(https://www.heygen.com/\)](https://www.heygen.com/)沒有付費有限制生成三個影片。

保羅(Paul)建議：多測試看看，可以先試做一個影片。專題之後至少上輔大新聞會是很好的成果。

【人臉辨識-林湘凌(Cecilia)& 羅靖傑(Morris)負責】

初步嘗試辨識表情，判斷是什麼情緒，臉鏡頭正前方可以辨識的心情有七種：喜、怒、哀、樂、厭惡、害怕和一般，發現厭惡比較難辨識出來，嘗試改參數

但作用不大，可能需要自訂參數。Mediapipe 可以偵測表情和舉手的動作，透過嘴巴開合的距離去判斷是開心還是驚訝；若偵測到手掌高於肩膀就判斷為舉手。

保羅(Paul)建議：

程式碼可以優化成模組化，也再看看有沒有更靈活的其他套件。cv2 會是專題很重要的部分，它是很多套件的 root 套件，主要做圖形處理。可以試試把特徵點都存成 csv 檔之類的，想知道存一個點的資料大小要多少，臉部表情串流面試要多少資料量。

## **貳、未來一週規劃**

### **【小組規劃】**

至少一人規劃系統藍圖，讓大家知道大概怎麼設計，有概念哪個功能做到什麼程度和確定適不適用我們的目標等。藍圖包含使用的流程和有什麼功能（例如哪個按鈕連到什麼功能），用文字或草圖設計，不過粗略就可以了。

### **【stt/tts-蕭伊庭(Sherry)&周易暉(Ethan)負責】**

模組化程式碼。stt 要有中英文嘗試，辨識率要提高；tts 要能自訂更多參數，調語速或角色等，可以參考 Audify 閱讀電子書。

### **【avatar-金殿倫(Darren)負責】**

廣泛搜索，先使用 HeyGen 免費額度試做一個影片看效果。

### **【人臉辨識-林湘凌(Cecilia)&羅靖傑(Morris)負責】**

模組化程式碼。試著把特徵點都存起來，看存一個點的資料大小以及臉部表情串流面試要多少資料量。