

會議記錄

專案名稱	尚未確認		
系統名稱	尚未確認		
文件編號	250101-會議記錄模板		
會議主題	小組第六次會議	編輯者	金殿倫
會議日期	2025 年 08 月 18 日	編輯日期	2025/08/18
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

【人臉辨識-林湘凌(Cecilia)負責】

這禮拜有針對程式碼分成大至三個模組作運用，Mediapipe 有測試針對點的多少去測試，但未根據點的多寡去影響到追蹤表情的速度跟明確程度，但因為只是針對點的判斷表情所以沒有很實際。

保羅(Paul)建議：

希望現階段針對表情偵測可以去判斷的套件去往下延伸，像是 deepface 學習模組是可以參考的例子。

【stt/tts-蕭伊庭(Sherry)&周易暉(Ethan)負責】

stt 部分有嘗試把程式碼從 base 改到 medium 之後在抓字的辨識度時候有比較精準，另位有尋找網路上很多英文的資料但使用後發現有不合或是不太能使用的狀況出現。

tts 在聲音的部分有找到 chatgpt 的付費模型，不包含分析的話成本是 5 塊左右

保羅(Paul)建議：

Stt 可以試試看先轉成簡體字看看，網路上有專門 python 套件可以使用，資料的部分如果都是英文可以先翻譯參考看看，如果看不懂在看原文，但還是建議找中文的資料或是影片，因為因為現在越來越多人拍影片的時代應該也不難找。Tts 的話就還需要花時間找一個明確的模組去使用。

【情緒辨識-羅靖傑(Morris)負責】

有嘗試讓模型去自動學習，並把情緒特徵的圖片上傳到各個情緒的資料庫讓模型能在即時偵測的時候變辨識到情緒，但不知道是不是因為沒有加辨識點的關係讓模組在偵測的時候容易受到光線的影響判斷，也有可能是因為資料庫的需求照片比較多。

保羅(Paul)建議：

情緒的圖片可以去用 paper with code 這個網站去增加資料庫中的情緒，但網站今年停掉了，最近有一個 github 網站比較熱門，網路上有類似的表情資料庫的。機器學習還是建議有千張以上才能對情緒判斷更精確。擴增學習是機器模型常用的技術，運用水平翻轉，改顏色，放大縮小讓機器訓練可以增強，另外，這部分可以晚點處理，現階段能先做基本的辨識就可以了，但如果能針對微表情做偵測就成功了。這部分能與林湘凌(Cecilia)和在一起做表情辨識的部分，因為特徵點的問題已經有解決方案了，。但能夠針對各個基本表情的程度去分析會取決這個專案的成功度。

【avatar-金殿倫(Darren)負責】

影片用 minimax 生成聲音再配合 heygen 生成個人影片，完成度雖高但會根據字數影響生成時間，有些不確定性。

保羅(Paul)建議：

完成度不錯但生成影片的時間還不太即時，要是能夠接近馬上根據學生講得馬上去講出即時的問題或是學生沒有呈現清楚的事。使用研究 heygenAPI 看能否適合我們的需求。

貳、 小組規劃

【系統藍圖-金殿倫(Darren)負責】

保羅(Paul)建議：

做的資料比較像是就是使用系統用了哪些技術，是這樣沒有錯，但老師希望的是一頁一頁像用畫的去清晰流程的功能，向第一頁案了按鈕會跑到下一頁類似這樣的 blueprint，一頁一頁的呈現在之後檢查溝通的時候才能一看就知道功能開發的如何。