

書報討論書面報告 碩資工一甲 11363149 吳冠穎

日期：11/12

講者：大同大學 許超雲 講座教授

題目：生成式人機互動技術研究

心得報告：

今天的書報討論課程邀請了，大同大學的許超雲講座教授，和我們分享生成式人機互動的技術研究，讓我們在兩節課的時間內，學習到一些新的技術，首先跟我們分享一段詩詞表示致意，虎尾春冰心不驚，意思是踩著老虎的尾巴，走在春天即將解凍的冰上，處境相當地危險，但因為技術成熟所以也不會緊張。接著和我們分享監督式學習以及非監督學習，在視覺追蹤的應用，多數文獻是採用監督式學習的方式進行，而研究採用的是非監督學習方法的生成式模型：變分自編碼器(VAE)。

因為 VAE 能夠很好地解析並重建臉部特徵點，也和我們分享，傳統自動編碼器(AE)，和 VAE 的差別在於，AE 較為依賴標籤數據，但沒有經過標籤的判斷，會比較沒辦法辨別，而 VAE 不同，VAE 較能夠在可能的兩個可能性中，找到一個中間的特徵，接著使用 Mediapipe 套件的影像辨識技術，捕捉人臉的特徵，也提出了可能會影像判斷的因素(e.g. 環境光、面部遮擋、面部轉動角度過大、臉部重疊)，需要將有問題的資料作處理，評估指標使用的是 R^2 Score、MAE、MSE、RMSE，數據處理上，去除了中心點，也修改了 Leaky Relu 函數，並且添加了適量的雜訊至訓練集。

為了使研究能更好的應用到各種情境，進行了延伸實驗，增加了不同的距離測試，還有增加更多的使用者進行測試，結果發現，在 50cm 的條件下，各評估指標的評分皆有所上升，結論，本研究的注視點偏差角度，相對於 Zdarsky 等人提出的方法、Lemley 等人提出的方法偏差的角度都還要好，也提到了未來可能的應用：影像標籤輔助、車輛輔助駕駛。

本週書報討論的演講和我們分享研究，真的受益良多，提到了遇到什麼問題，還有問題的解決，使用的方法，加入了什麼優化方式，以及和其他學者的研究做比較，這些都是未來撰寫論文會用到的，學習到了不少。

關鍵字：VAE、AE、Mediapipe、 R^2 Score、MAE、MSE、RMSE、Leaky Relu

參考文獻：

[1] 當天演講的投影片

[2] 自己做的筆記