

利用Transformer判斷人臉健康狀況系統

賴麗玉/Lily



ABSTRACT

人臉是人類獨特的身份標誌，提供年齡、情緒和健康等資訊。人臉與健康息息相關，通過觀察臉部，如眼睛、皮膚，可以初步推斷是否有健康問題，對於病症的及時治療和改善相當重要。本研究使用 Dlib 函式庫進行人臉偵測，以及使用 Vision Transformer 模型訓練，此模型是基於多頭自注意力機制 (Multi-Head Self-Attention) 的圖像式深度學習模型，利用自注意力機制提升模型訓練速度，及其可擴展性，被廣泛應用於電腦視覺領域，因此本研究使用 Vision Transformer 模型識別人臉的健康與帶有病徵的狀況。訓練模型使用的公開資料集來自 Kaggle、Roboflow，以及使用 Google、Microsoft Bing 等搜尋引擎蒐集圖像。此研究使用5種病徵及無病徵，共6種類別，作為圖像資料集的類別來訓練 Vision Transformer 模型，得到的 Accuracy 為 80.24%，Loss 值為0.9988。

INTRODUCTION

人臉的特徵變化與健康狀況具有密切關係，去醫院看診時，醫生可以透過觀察人臉的皮膚、五官等特徵，初步推斷是否有健康問題，判斷有何種疾病，或人體異常現象反應在面部的變化，再深入進行健康檢查或治療，有助於及早發現和解決健康問題。本研究希望使用人臉無病徵及病徵狀態的圖像資料集，以 Vision Transformer (ViT) [1] 模型為訓練模型，完成人臉健康狀況辨識系統，並希望透過此系統，輔助檢測人臉是否有病徵，提升辨識的效率。

BACKGROUND

人臉病徵辨識相關文獻

研讀了4篇關於面部識別的論文，並彙整各論文的研究範圍、訓練結果及優勢，為本專題提供了研究方向。

Jiaqi Qiang et al. (2022) [2] 主要是對面部識別疾病診斷應用的綜述，統整不同深度學習模型，及疾病造成的面部特徵，包含內分泌及代謝疾病、遺傳和染色體異常、神經肌肉疾病及其他類型的疾病。在遺傳和染色體異常上，使用FDNA (Facial dysmorphism novel analysis)，辨識狄蘭氏症候群，有最高的準確率94%。在神經肌肉疾病上，使用OpenFace 2.0，辨識帕金森氏症，有最高的準確率95.6%。

Zhaohui Su et al. (2021) [3] 是對基於深度學習的面部影像分析在醫學研究中的特徵和效果。其方法是找尋 PubMed、PsycINFO、CINAHL、IEEEExplore 和Scopus 等資料庫中相關研究，進行系統性回顧(Systematic review) 及統合分析(Meta-analysis)。使用CNN (卷積神經網路) 辨識臉部皮膚病，具有較高的準確率為95.24%。

Li Zhang and Bob Zhang [4] 使用Multi-Disease CNN (MD-CNN) 辨識糖尿病、肺部疾病及健康面部，進行多疾病分類。以非侵入式設備提取面部，找出瞳孔位置，區分額頭、鼻子及左、右臉頰四個區域，使MD-CNN 可有效辨識面部特徵。此研究將資料集，分為不同比例進行訓練，比較準確度、迭代次數，並與SVM、k-NN、multi-class AdaBoost、ScSPM等模型進行比較，其中MD-CNN 分成 7:3時，有最高的準確率為 75%。

Omneya Attallah [5] 提出一個能檢測單一和多種疾病的面部診斷系統FaceDisNet，此系統整合來自不同結構的卷積神經網絡構建而成。使用Bo Jin 的《Disease-Specific Faces》數據集，包含β-地中海貧血、甲狀腺功能亢進症、唐氏綜合症和麻風病等疾病，以及健康面部。FaceDisNet在二元分類和多類分類類別經過集成分類和特徵選擇步驟後，分別達到了98.57%和98%的最高準確率。

METHODOLOGY

Dlib 面部區塊提取

模型訓練前，先將圖像使用 Dlib [6] 進行人臉切割，Dlib 提供圖像處理的函式庫，可用於人臉偵測。需使用 shape_predictor_68_face_landmarks.dat 模型辨識人臉，該模型可以檢測及標示人臉圖像中的 68 個面部特徵點。使用 Dlib 檢測出資料夾中所有圖像的人臉，取得臉部特徵，進行圖片切割，儲存為新的圖像檔，去除了大部分非人臉的特徵後，可提升臉部辨識的準確率。



圖1 面部區塊提取

Vision Transformer

本文使用的 Transformer 模型為 Vision Transformer 模型，是 Google 在2020年提出的圖像辨識深度學習模型，採用多頭自注意力機制，使模型能在整張圖像中學習全局性的特徵，而不只是局部特徵，具有可擴展性、長距離依賴性等特點。

系統架構

系統主要由人臉健康狀況辨識及介紹介面構成，以網頁介面為系統整體之主要架構。

圖2 為網頁的運作流程，使用者只會看到自己上傳的圖片及檢測結果。使用者登入帳號後，選擇進入檢測頁面，可以選擇進行人臉健康檢測，上傳圖像，圖像會傳進資料庫裡，選擇檢測按鈕，資料庫裡的圖像會進入 Dlib 進行面部提取，再進入 ViT 模型進行檢測，最後回傳檢測結果至頁面。

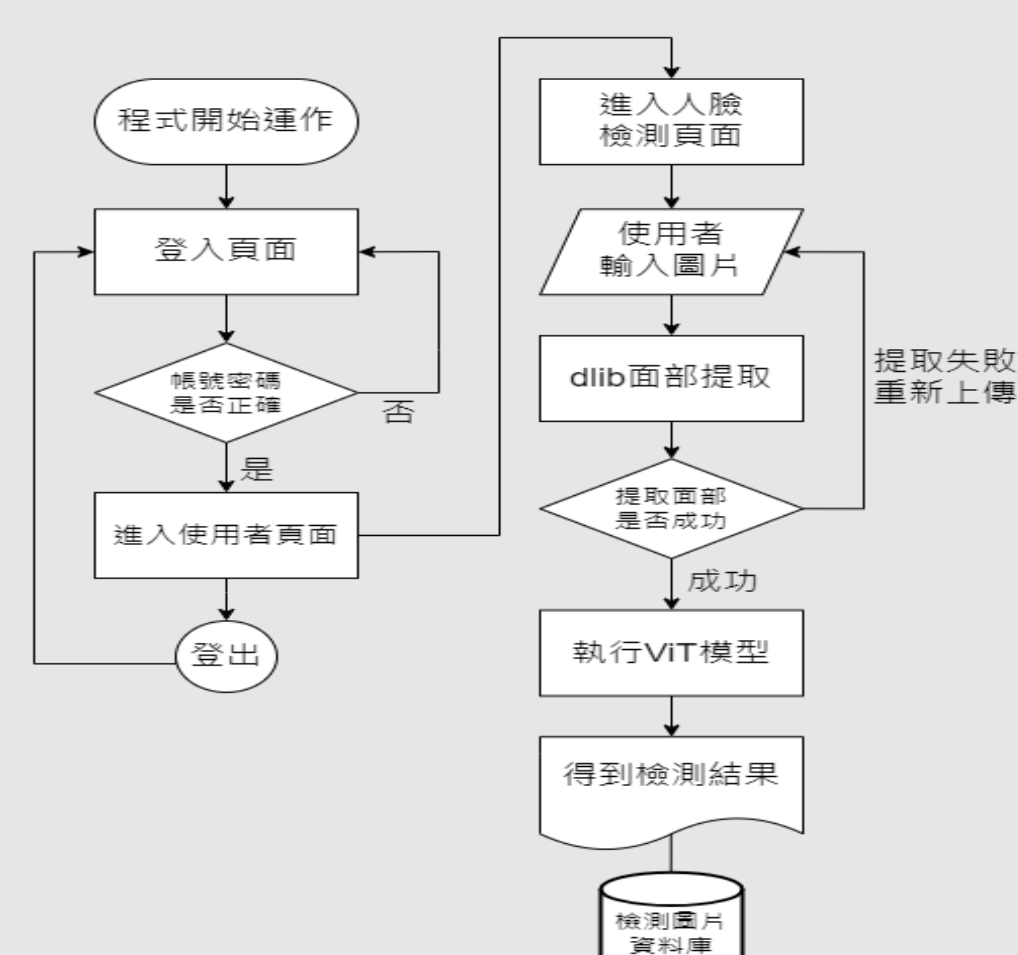


圖2 網頁流程圖

系統介面實作之工具

Flask 是使用 Python 編寫的輕量級Web應用框架，使用簡單的核心，擴充增加其他功能。SQLite 是量級的嵌入式關聯式資料庫管理系統，支援 SQL 語法，可整合至應用程式中，減少資料庫訪問延遲。本系統需快速返還辨識結果，所以選擇使用 SQLite 搭配 Flask 框架，來降低系統延遲的問題，使 ViT 模型能在網站上運作。

RESULT AND DISCUSSION

資料集收集

本研究使用 Kaggle 及 Roboflow 的公開資料集，以及 Google、Microsoft Bing 等搜尋引擎蒐集圖像。資料集分為六種類別，包括無病徵、痤瘡、黃疸、水腫、顏面神經麻痺及結膜炎。圖片共 892 張，為提高訓練資料集的品質，在訓練中使用數據增強來增加資料多樣性，包括改變飽和度、曝光度。

Vision Transformer的訓練結果

ViT 模型的訓練結果整體準確度為80.24%，Loss 值為0.9988。

表1 各類別資料集張數

類別	Train(張)	Valid(張)	合計
無病徵	180	44	224
痤瘡	146	37	183
黃疸	90	22	112
水腫	99	24	123
顏面神經麻痺	132	33	165
結膜炎	68	17	85

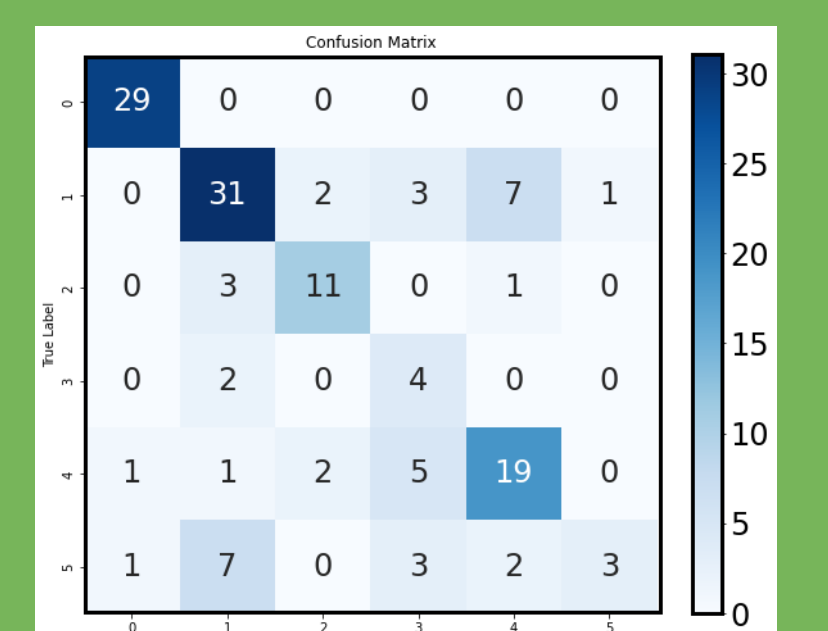
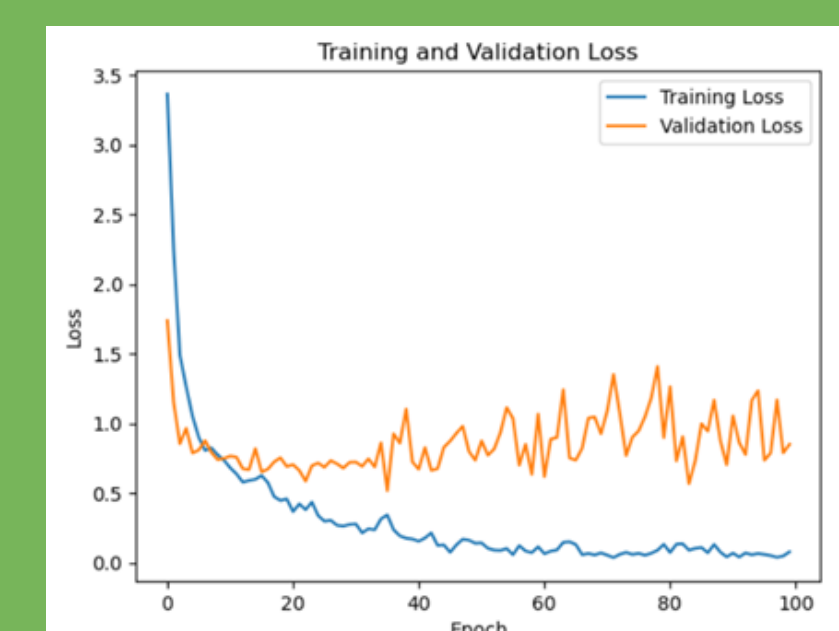


圖3(a) loss 函數圖 (b) 混淆矩陣圖

系統介面

系統介紹介面作為起始畫面，介紹頁面可選擇模型或病徵分類介紹。模型介紹頁面介紹系統架構，病徵分類頁面介紹病徵類別。系統介面的面部健康檢測，點擊新增圖片選擇欲檢測的人臉圖片，按下檢測按鈕，頁面會顯示經過 Dlib 人臉檢測與 ViT 模型處理後的圖片及文字呈現檢測結果。



圖4 病徵分類頁面

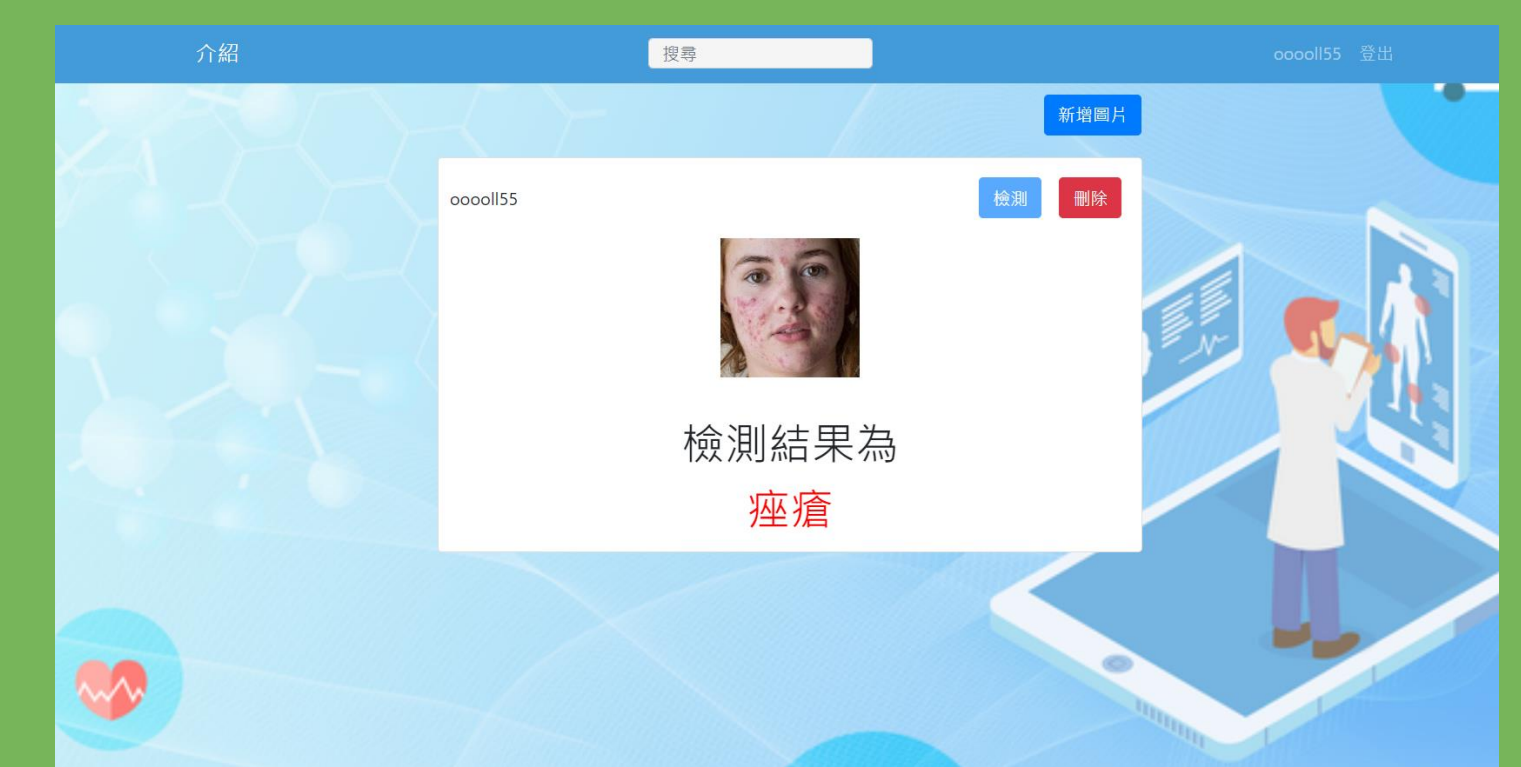


圖5 檢測結果畫面

CONCLUSIONS

鑒於分辨人臉的健康狀況，ViT 模型利用多頭自注意力機制提升模型訓練效率。然而，資料集的大小及圖像也會影響模型準確性，由於帶有病徵的完整人臉圖像較少，且部分特徵具相似性，使辨識的難度增加，造成辨識的準確性下降。因此，在資料集及訓練上還有許多改進空間，未來期望加入更多不同疾病特徵的人臉圖像、具有醫學診斷的資料集，並改進模型架構，以提升該研究的準確度及可靠性。透過這些方向的深入研究，將提升 ViT 模型在人臉健康狀況辨識方面的性能和應用價值。

REFERENCES

- [1] Khalid Salama, "Image classification with Vision Transformer," Keras, https://keras.io/examples/vision/image_classification_with_vision_transformer/, Jan. 2021(accessed May, 2023).
- [2] Jiaqi Qiang, Danning Wu, Hanze Du, Huijuan Zhu, Shi Chen, Hui Pan, "Review on Facial-Recognition-Based Applications in Disease Diagnosis," MDPI, <https://www.mdpi.com/2306-5354/9/7/273>, Jul.2022 (accessed Oct. 2023).
- [3] Zhaohui Su, Bin Liang, Feng Shi, J Gelfond, Sabina Šegalo, Jing Wang, Peng Jia, Xiaoning Hao, "Deep learning-based facial image analysis in medical research: a systematic review protocol," BMJ j. BMJ Open. vol. 11, no. 11:e047549, pp. 1-6, Nov. 2021. [Online]. Available: <https://bmjopen.bmj.com/content/11/11/e047549>
- [4] Li Zhang, Bob Zhang, "Non-Invasive Multi-Disease Classification via Facial Image Analysis Using a Convolutional Neural Network," in: Proceedings of the International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR), Chengdu, China, 2018, [Online]. Available: <https://ieeexplore.ieee.org/document/8521262/citations?tabFilter=papers#citations>
- [5] Omneya Attallah, "A deep learning-based diagnostic tool for identifying various diseases via facial images," Sage J. Digital Health Vol.8: 1–22, Sept. 2022. Available: <https://journals.sagepub.com/doi/pdf/10.1177/20552076221124432>
- [6] coneypo, "Python 3 利用 Dlib 实现人脸检测和剪切," cnblogs, <https://www.cnblogs.com/adaminxie/p/8339863.html>, Jun. 2021(accessed May, 2023)