

應用YOLO姿態估計與ST-GCN進行時空動作辨識以輔助高齡照護與復健

組員: 徐皓揚、鄧顏可、哈又馨、蔡子謙 指導教授: 王士豪 教授

01 研究動機與目的

隨著全球人口老化，因肢體限制產生的復健與跌倒偵測需求劇增。傳統穿戴式偵測器因儀器昂貴、配戴不適，難以普及於居家環境，現有AI 框架多採用多階段模型，導致擷取影像關鍵點時出現延遲，且高精度的 CNN 並聯特徵融合模型在臨床小樣本資料集下易面臨失效。

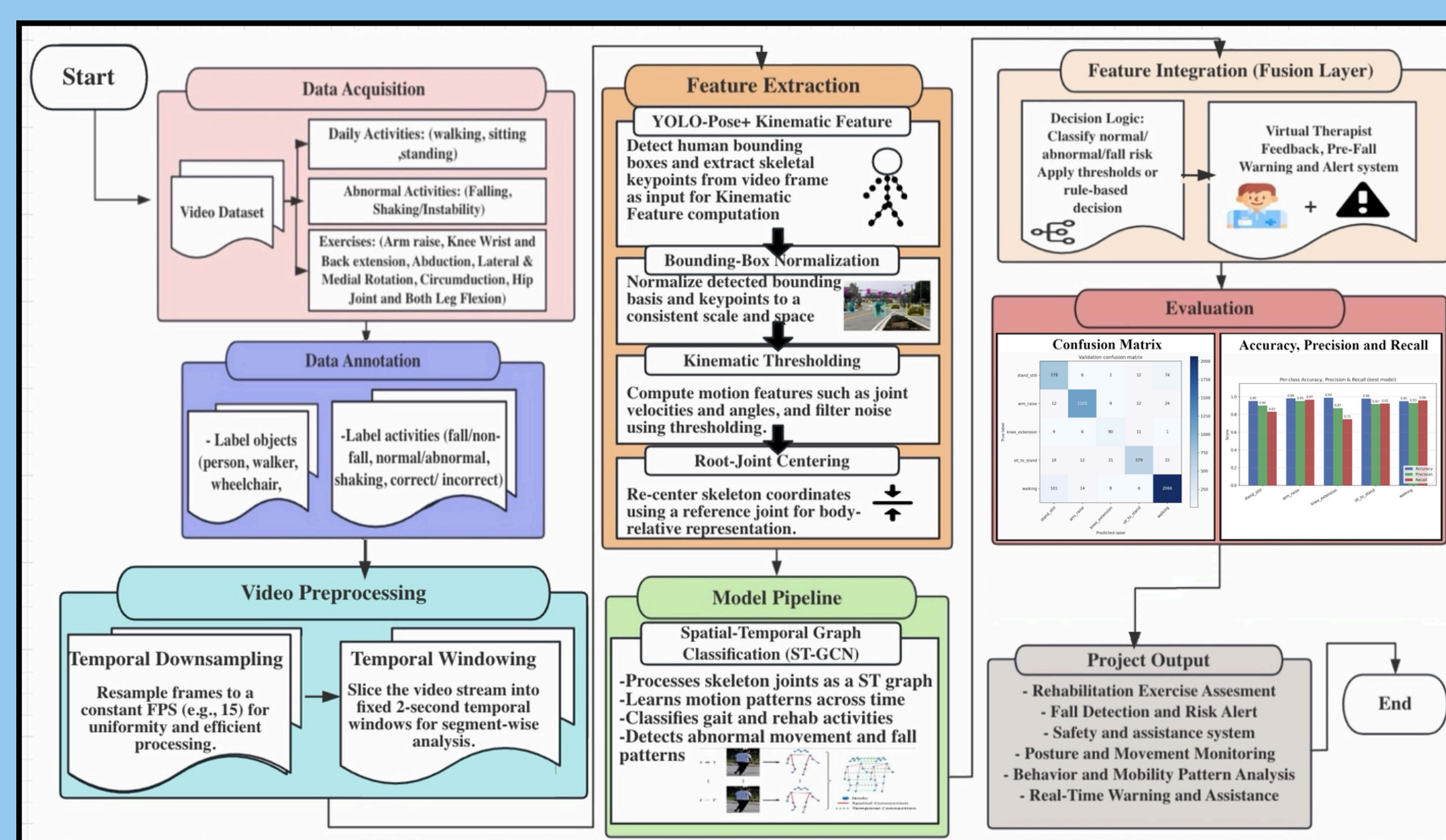


圖一、復健示意圖

此研究導入 YOLOv11-pose 實現單階段提取人體目標與骨架關鍵點，大幅提升非侵入式步態分析的推理速度，並利用 ST-GCN（時空圖卷積網路）確保在小樣本數下能準確推斷動作，並應用在可部署於居家設備的軟體，以優化居家復健動作追蹤與長者跌倒即時偵測。

02 系統架構

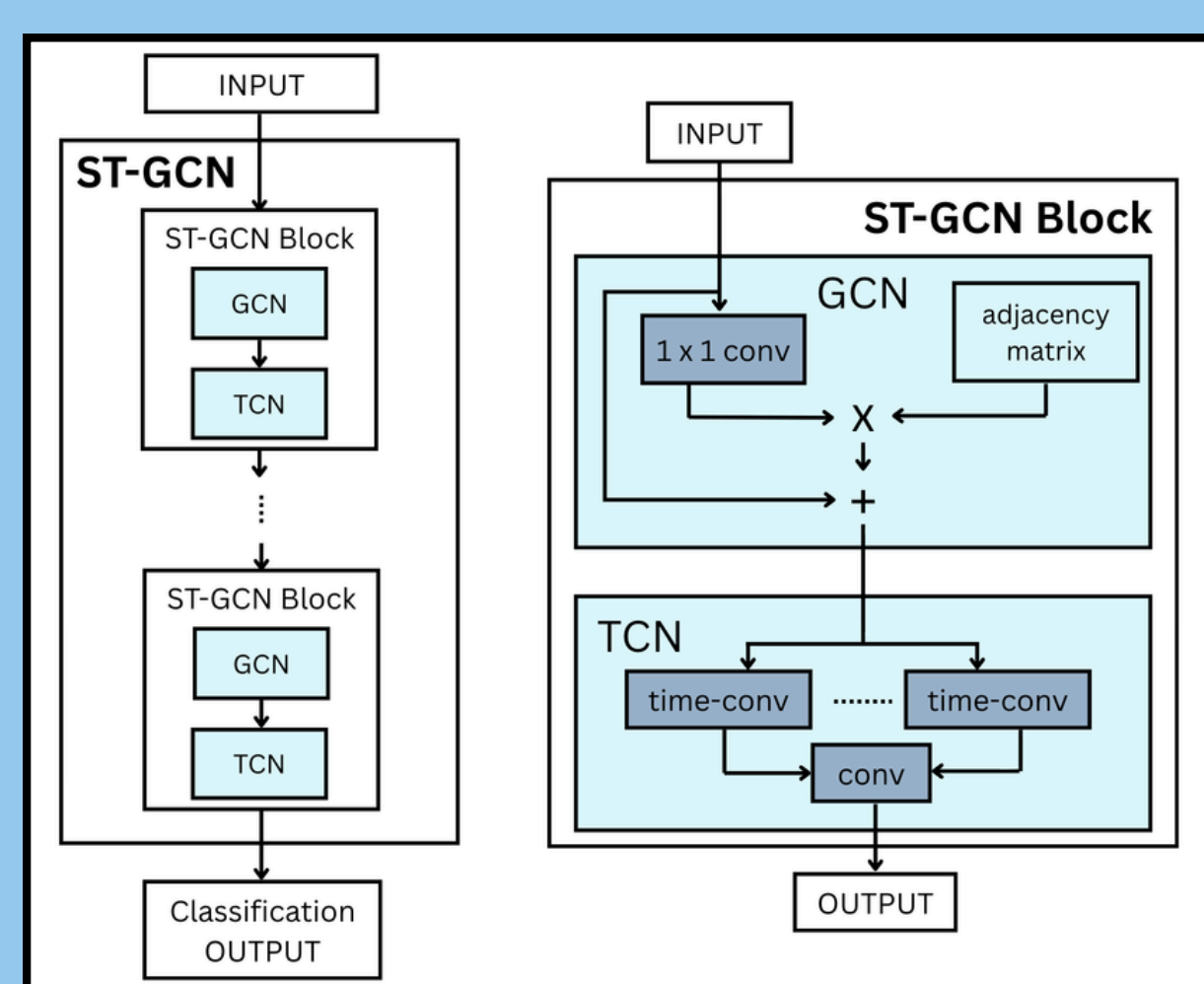
WLU、MobiPhysio、HAR 等數個資料集內的肢體運動影片被用來訓練影像辨識模型。這些影片是以動作項目以及示範正確與否標註。資料預處理環節中，時間方面對影片降低幀數 (downsample) 並分段 (window) 來確保影片長度的一致性，空間方面則透過邊界框裁切雜訊，再以 YOLOv11-Pose 擷取重要關節點。



圖二、系統架構圖

03 實作技術與環境

程式在 Windows 11 以 Python 開發。訓練模型時採用時空圖卷積網路 (ST-GCN) 分析每個動作所對應的關節點移動方式。



圖三、ST-GCN 示意圖

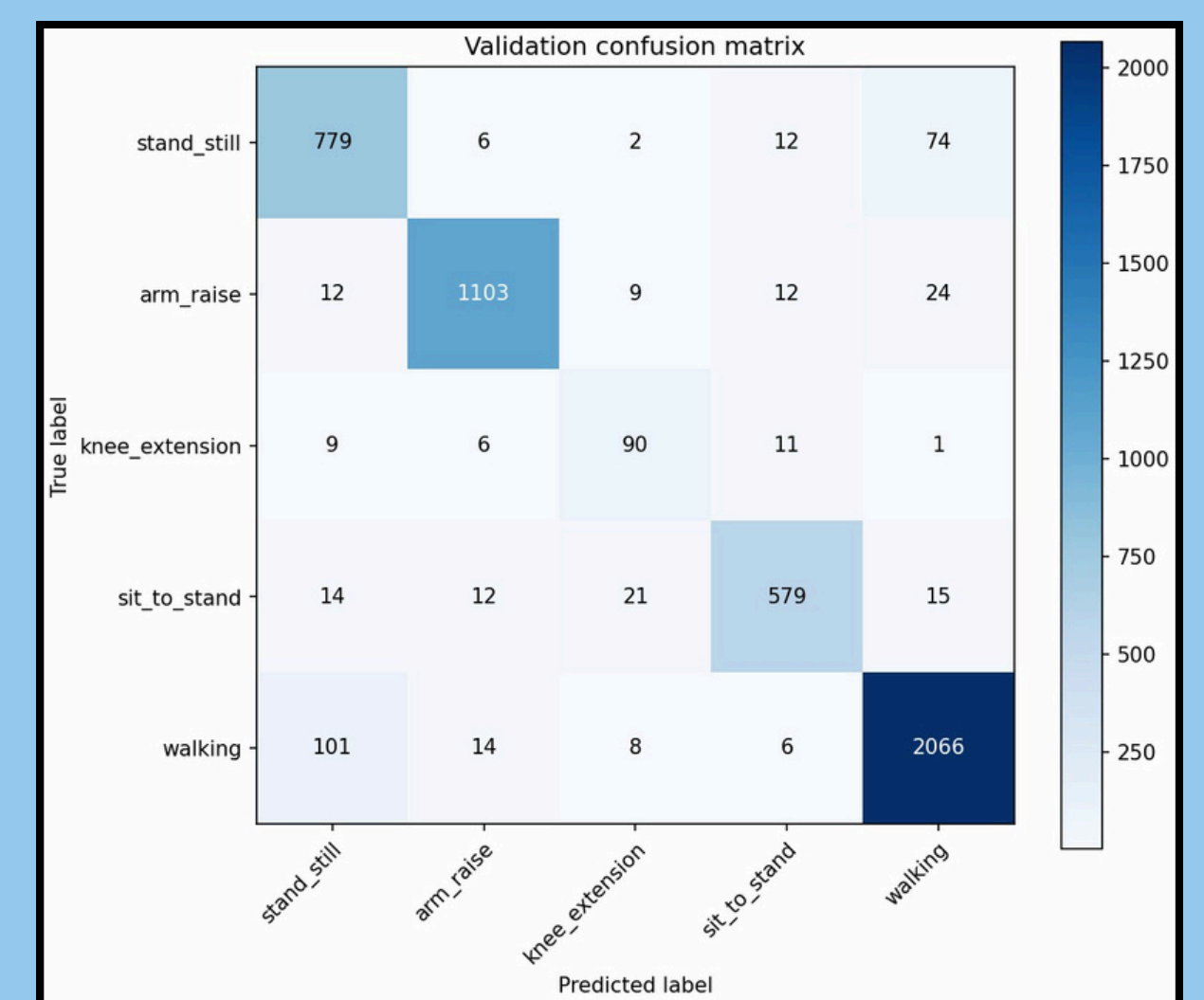
04 成果展示

經過 20 週期的訓練後，模型偵測動作的能力如下：

模型整體準確率
92.7%

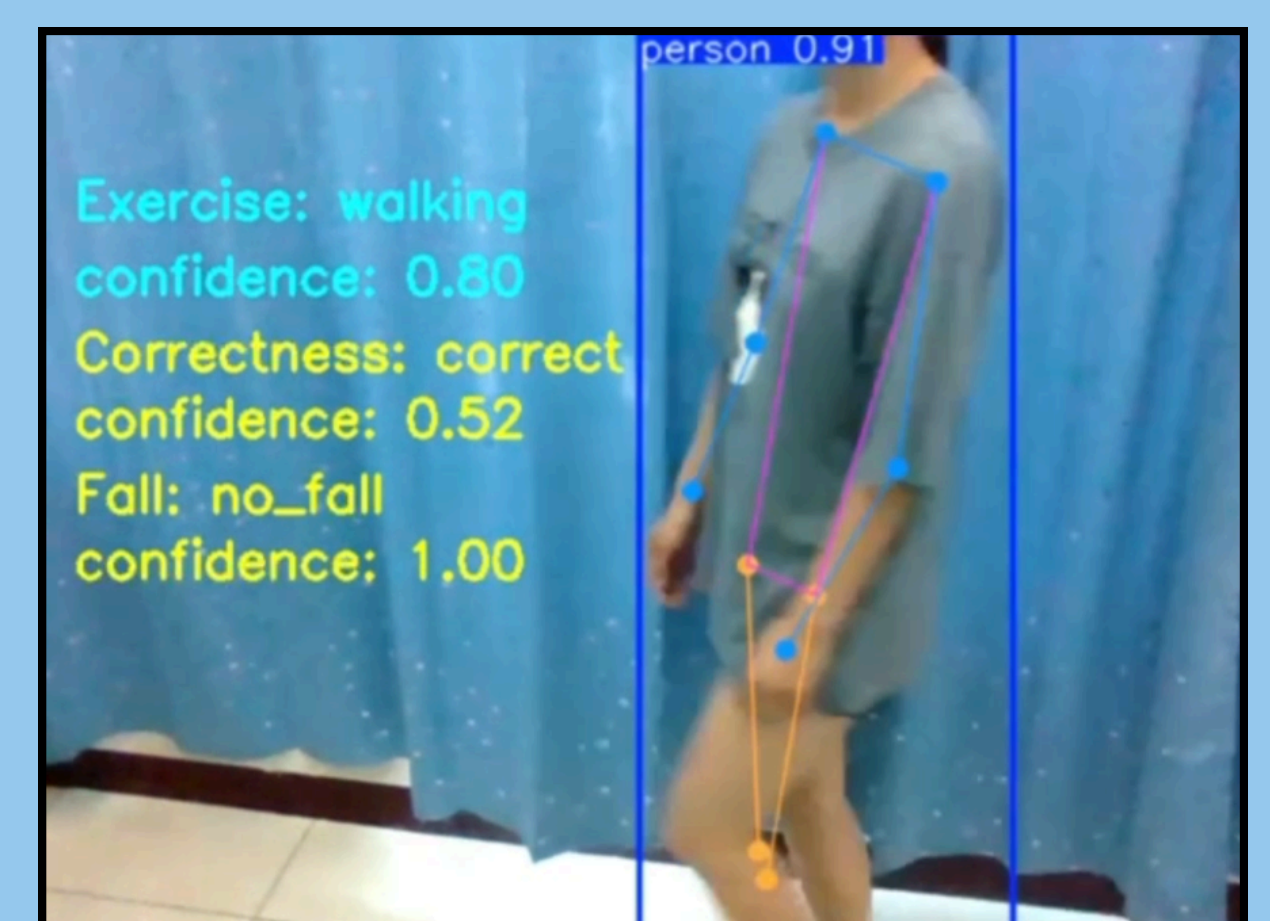
分類最低精確度
87.2%

分類最低喚回度
74.6%



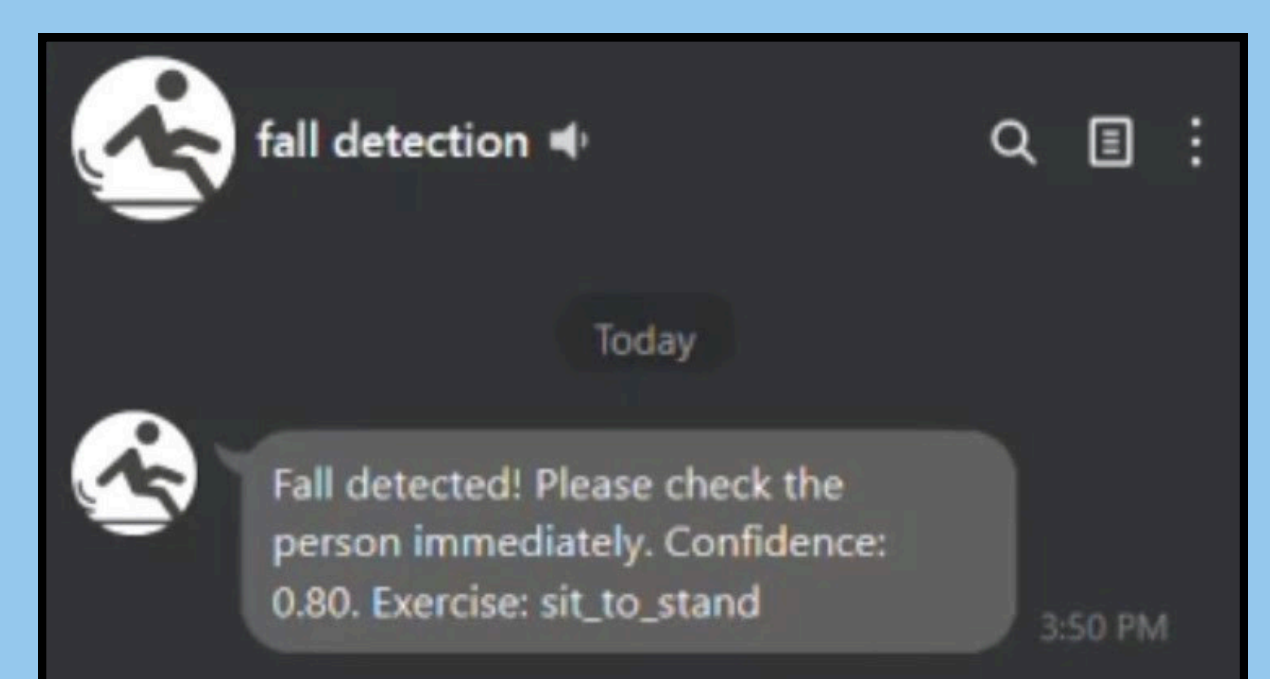
圖四、訓練模型混淆矩陣

程式啟動後，會開啟使用者的電腦鏡頭。隨著活動的影像輸入至系統裡，系統會在對其進行前述的預處理後讓模型進行分類，並在輸出影像中顯示其判定受測者執行的動作、動作是否正確、及其信心指數。



圖五、程式使用示意圖

若是系統偵測到使用者者跌倒動作，不僅會在螢幕上警示，還會透過 LINE 發送警告訊息。此功能可望降低使用者受傷至就醫之間的時間延遲。



圖六、警示訊息示例

05 結論與未來展望

透過 YOLOv11-pose 和 STGCN，鏡頭能以極低延遲的捕捉速度，透過解析拍攝者的連續動作關鍵點判斷動作以及進行摔倒預警，且達到不錯的準確率。

未來可以利用準確度更高的模型，在維持捕捉速度的同時提高動作預測的準確度，程式也可以透過顏色標記使用者錯誤的姿勢，逐步指導修正動作。

06 示範影片

掃描右側 QR code 可以查看成果示範影片

