中華民國第 64 屆中小學科學展覽會

作品說明書

高級中等學校組　電腦與資訊學科

團隊合作獎

052511

基於影像識別與感測裝置之坐姿偵測輔助系統

學校名稱： 國立新竹女子高級中學

作者：	指導老師：
高二 紀欣妤 高二 陳宥妤 高二 蕭伃肴	古佳怡

關鍵詞： 影像識別、感測裝置、智慧醫療

摘要

本研究旨在開發一套系統，判斷使用者是否久坐與使用電腦的姿勢是否正確，並即時提供回饋。此坐姿輔助系統僅需使用電腦內建視訊鏡頭以 MediaPipe BlazePose 影像辨識模型抓取上半身節點資訊，並透過雙腳穿戴 micro:bit 感測裝置取得腳部三維加速度值。研究者蒐集各年齡層不同姿勢下的數據，討論其差異並計算 F-score 值以訂立閥值，判斷使用電腦的姿勢。同時，研究者運用 PyQt 套件建立使用者介面，使得系統能夠進行個人化設定，並即時地提供回饋。相較現有研究，本研究所提出之系統更輕量化、方便使用且成本更低。

壹、研究動機

一、研究動機

根據國民健康署 2017 年的調查[1]，臺人每日平均久坐高達 5.9 小時。許多民眾因久坐或以不良姿勢使用電腦，導致諸多生理問題，如肩頸及背部疼痛。世界衛生組織指出[2]，全球約有 60%到 85%的人口有久坐的生活型態，此型態被列為十大死亡及患病主因。

研究者希望能從源頭解決此些問題，研發一套輔助系統，透過影像辨識判斷上半身姿勢，與簡便穿戴式裝置偵測腳部資訊。此系統可在使用者姿勢錯誤或久坐時即時提醒，警示使用者應調整坐姿和適時走動，並提供錯誤姿勢統計報表。

二、文獻回顧

以下對現有姿勢輔助相關文獻方法進行探討與分析：

產品或論文	內容	待改良
俄羅斯工程師所開發的姿勢輔助系統[3]	透過電腦鏡頭拍攝使用者正確姿勢，錯誤時模糊瀏覽畫面	系統啟動時使用者需擺出正確姿勢、僅偵測上半身、僅模糊指定畫面
SitApp 姿勢輔助軟體[4]	透過電腦鏡頭判斷使用者姿勢，並在螢幕中即時提醒	需先拍攝正確姿勢、僅偵測上半身
以智慧椅墊進行坐姿分析之研究(2019 年，臺灣師範大學)[5]	透過 6 個壓力感測器分類坐姿	感測坐墊難以普及，無法即時提供回饋

坐姿檯燈偵測(已商品化，7990 元)[6]	透過超音波及紅外線感測頭部位置，一定時間內未改善則透過語音提醒使用者	僅偵測頭部、細部未分析、改善強制力不足、售價過高
Posture Pal APP 偵測坐姿[7]	透過 AirPods 的動作傳感器，偵測使用者頭部動作	僅偵測頭部、硬體設備門檻較高

表 1 姿勢輔助相關文獻探討(研究者自繪)

研究者發現現有研究與產品，存在價格過高、功能受限或使用不便等問題，因此將研發一套更輕量化、方便使用且成本更低的系統。

貳、研究目的

一、結合影像辨識與穿戴式裝置，偵測使用者當下姿勢。

二、辨識使用者使否使用電腦、是否呈現坐姿、及姿勢是否正確。

三、蒐集資料集，設定不同姿勢之閥值，並探討其準確率。

四、給予使用者即時的回饋，督促使用者改善坐姿或警示使用者離坐走動。

五、設計使用者介面，使其方便使用、貼合使用者需求並提供數據統計。

六、增加此系統的實用性，使其方便攜帶及降低成本。

參、研究設備及器材

中央處理器：CPU 11th Gen Intel(R) Core(TM) i5-1135G7 @ 2.40GHz 2.42 GHz

記憶體：8.00 GB

硬碟：475GB

程式語言：Python 3.11

函示庫套件：PyQt5 / MediaPipe / opencv

感測裝置硬體：micro:bit V1.3B

肆、 研究過程或方法

一、研究架構

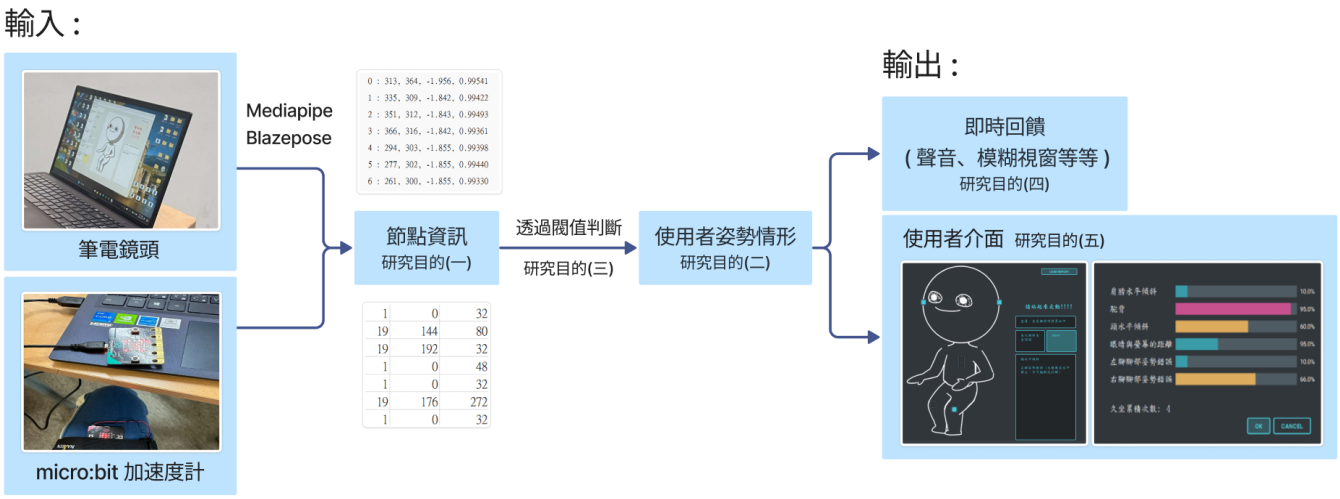


圖 1 研究流程與架構圖(研究者自繪)

二、裝置示意圖



圖 2 系統裝置示意圖(研究者自繪)

伍、研究結果

本研究欲透過「影像辨識」與「穿戴式感測裝置」偵測姿勢，以下為技術方法描述：

一、影像辨識模型 MediaPipe BlazePose

本研究實作目前常用影像辨識框架，並比較其效果及分析優缺：

	特點	**優缺點**
Yolo	廣泛運用，能快速辨別是否為物體及物體類別	與本研究針對骨架節點分析比較，較無關聯
Vibe	根據骨架二維資訊推估三維人體模型	實作發現此模型計算速度較慢，不夠即時，且易發生版本衝突

MediaPipe	由 Google 釋出影像辨識框架，根據骨架二維資訊推估三維人體模型	符合本實驗目的，輕量化，速度足夠即時且具泛用性

表 2 Yolo、Vibe、MediaPipe 比較表格(研究者自繪)

考量效能與準確性，研究者選擇 MediaPipe 為此研究使用的影像辨識框架，其有許多運用及功能，其中 BlazePose 為具身體姿勢追蹤功能的模型，能偵測身體關鍵節點。BlazePose 運作步驟分為輸入圖片、偵測人體、放置節點與輸出，以下針對此四個步驟進行更詳細說明。

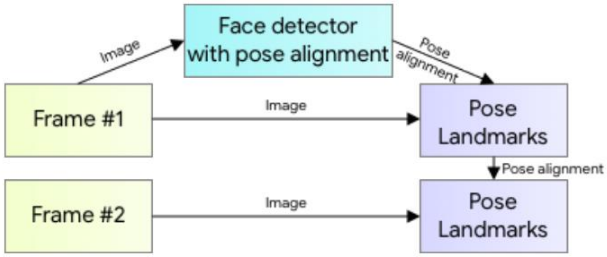


圖 3 BlazePose 運作流程(擷取自參考資料[9])

(一) 輸入圖片

即時輸入圖片以進行接下來的步驟。

(二) 偵測人體

使用 NMS 演算法(Non-Maximum Suppression)偵測人體臉部，因其為人體中不易改變且最具特色的部位。NMS 演算法將從多個物件中保留最符合臉部的物件，如圖 4 所示。

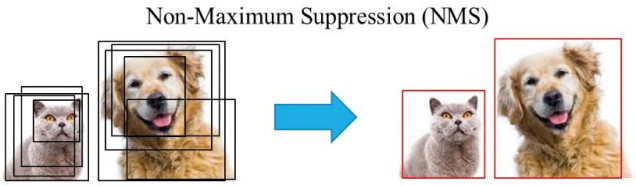


圖 4 NMS 示意圖(擷取自參考資料[8])

圖中每一個框，皆為一個 Bounding Box(簡稱 BBox)，每個 BBox 皆被賦予五個值，框的中心位置(x, y)、長寬(h, w)及信心程度(confidence score)。NMS 演算法將計算各 BBox IoU 值(Intersection Over Union)，以下為 IoU 計算公式：

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

實際篩選流程如圖 5 所示：

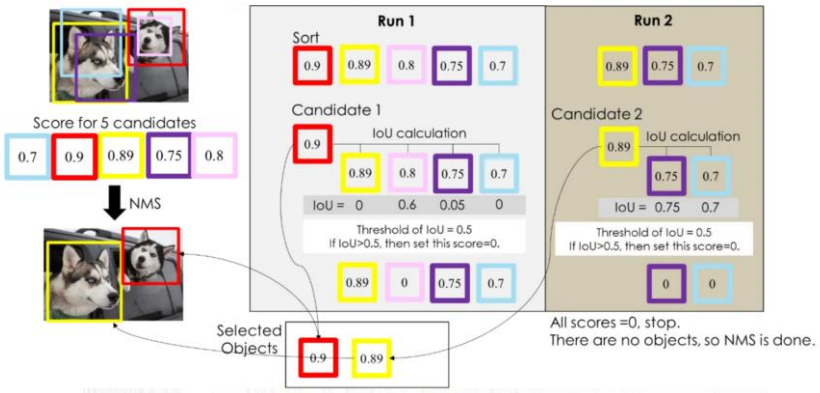


圖 5 NMS 演算法篩選流程圖(擷取自參考資料[8])

每次執行皆選定一信心程度最高的 BBox，並將其與其他 BBox 進行 IoU 計算，當其值大於設定閥值，表示該 BBox 與選定的 BBox 高度重疊，則判斷此非獨立物件，便將此 BBox 信心程度調整為零以排除。此演算法會一直執行直到除了選定的物件，其他信心程度皆為零為止。

藉由此演算法找到一個最具代表性的臉部區域後，程式將根據臉部資訊，推估人體中心（臀部中點）及其與肩膀中點的連線，並定義此為中心垂線，以及圍繞人體的圓圈大小，如圖 6 所示。人體體型及節點相對位置大致相同，因此模型以上述推估之關鍵資訊有效縮小檢索範圍，以加速計算其餘骨架節點。之後步驟將聚焦於此感興趣區域(ROI，Region of Interest)，直到區域中人體變動或消失才會重新偵測，以增加模型效率。

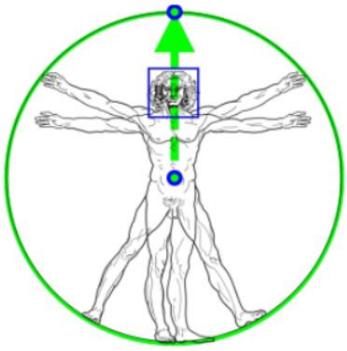


圖 6 完美人形示意圖(擷取自參考資料[9])

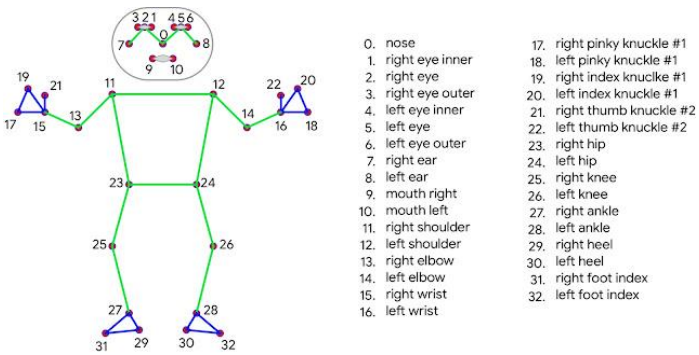


圖 7 骨架節點示意圖(擷取自參考資料[9])

(三) 放置節點

模型放置 33 個節點，位置如圖 7 所示。節點選擇參考其他模型，如 BlazeFace、BlazePalm 和 Coco 資料集，保留其中必要節點，透過最少節點數取得關鍵資訊，以增加放置節點效率。

BlazePose 運用熱圖(heatmap)和偏移量分析圖(offset map)以及回歸分析，放置 33 個節點。heatmap 預測節點位置分為兩個步驟，第一個步驟透過交叉熵函數(sigmoid)為每一個節點產生一個初階熱圖，此圖中顏色深淺表示該位置為此特定節點的概率；第二個步驟透過

confidence map 對之前每一張初階熱圖進行回歸，取得更加準確的節點位置，最終產生回歸熱圖，標示每個像素為該節點的概率。透過 heatmap 找到節點的初步位置後，將以 offset map 預測偏移量，使節點預測更加準確，實際準確率如圖 8 所示。

實際流程圖如圖 9 所示，分成左中右三個區塊進行訓練，左中兩區塊估算 heatmap 及 offset，右側為 regression。輸入圖片後，BlazePose 將圖片降維，在過程中為了在高維和低維特徵中取得平衡，會經由圖中 skip connection 跳過某些步驟。在左中兩區塊訓練後，得到一個輸出，該輸出會被右側的塔捨去，接著進行 regression 訓練。虛線為單向箭頭，故右側 regression 得到的結果不會回傳給左側的 heatmap 重新進行訓練，這樣不僅改善了 heatmap 的預測結果，也可提高整體的精準度。

	Mean 2D Euclidean error, normalized by torso size	*PCK@0.2*
Heatmaps	16.2	83.6
Regression without Heatmaps loss	15.9	79.9
Regression with heatmap regularization	**14.4**	**84.1**

圖 8 使用三種不同方式的錯誤度與精確度比較 (擷取自參考資料[9])

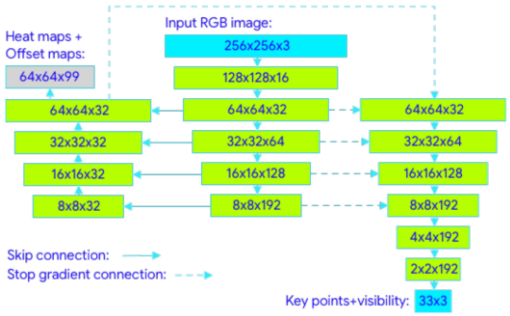


圖 9 放置節點流程圖(擷取自參考資料[10])

以下介紹放置節點的訓練資料：

初始訓練資料集使用整個人都清晰可見或是臀部及肩膀中間可被清楚標示的圖片 。爾後，進行遮蔽模擬(occlusion augmentation)，遮擋照片的一部分，訓練模型判斷人體被遮蔽時的節點位置。此模型訓練總共使用 85000 張照片 ，其中 60000 張照片中有一人或多人展示簡單的動作，另外 25000 張照片中展示較複雜的動作。訓練使用的照片節點皆有被人工標註。

(四) 輸出

輸出各骨架節點 x, y, z 以及 visibility 數值，如圖 10 所示。x,y,z 為節點三維資訊，皆以人體中心為原點。visibility 為節點可視程度，數值範圍為 0 至 1。

	time	0_x	0_y	0_z	0_vis	1_x	1_y	1_z	1_vis
2	2024-01-27-14_11_38	244	187	-0.732	0.99989	253	158	-0.672	0.9997
3	2024-01-27-14_11_38	244	189	-1.122	0.99988	253	158	-1.058	0.99968
4	2024-01-27-14_11_38	245	191	-1.125	0.99987	254	159	-1.06	0.99965
5	2024-01-27-14_11_38	245	191	-1.098	0.99986	254	159	-1.033	0.99962

圖 10 骨架節點數據(研究者自繪)

其中 z 值是透過 x, y 值與弱透視投影(weak perspective projection)推估而來，靠近鏡頭呈負值，遠離鏡頭呈正值。弱透視投影法為透視投影(perspective projection)與正交投影(orthographic projection)的混和體。透視投影是由一點發射光線所產生的投影，會有縮放及扭曲的現象；正

交投影則是假設焦距無限大，每一條投影線接互相平行，如圖 11 所示。弱透視投影法保留透視投影之縮放特性，及正交投影平行光投影的特性，使物體具有深淺差異且不會變形，如圖 12 所示。此時即可用 x 值及 y 值回推 z 值，公式如下，其中 f 表示焦距，x 及 y 表示物體實際座標，x’ 及 y’ 表示投影後的座標。

$$x' = \frac{xf}{z}, y' = \frac{yf}{z}$$

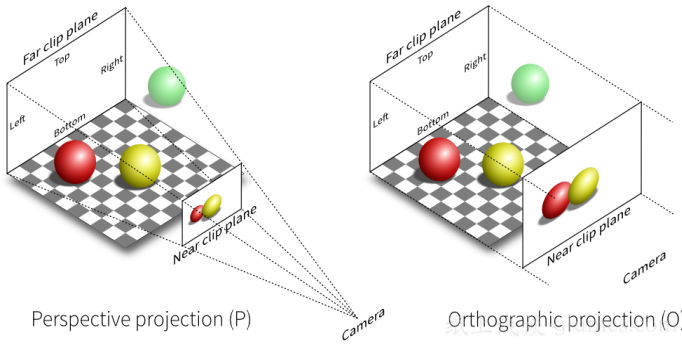


圖 11 投影法示意圖(擷取自參考資料[11])

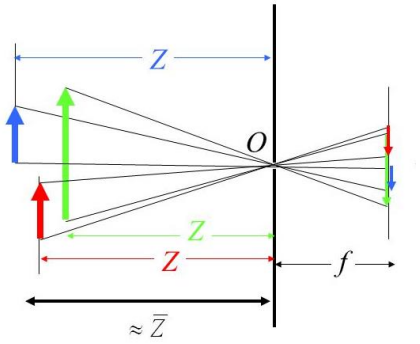


圖 12 弱透視投影法示意圖(擷取自參考資料[14])

二、穿戴式感測裝置

根據 2022 年台灣網路報告統計資料指出[15]，全台約有 20.33%的人具有使用智慧手環、手錶方面的相關經驗。另根據 Statista 統計報告顯示，2022 全球穿戴式感測裝置售出金額高達 11 億美金，且各統計皆預測此市場會持續擴大[16]。鑒於具有配戴穿戴式感測裝置習慣的使用者逐漸增加，本研究使用輕便的穿戴式裝置，在不增加使用者負擔下，取得腳部節點三維資訊。本研究使用 micro:bit，以簡便的穿戴裝置取得雙腳三維加速度。

micro:bit 是由英國廣播公司研發的微型電腦開發板，此板子小巧、操作簡易且成本不高，每片市價約 500-600 元，內部配置加速度感測器、光感測器以及溫度感測器。micro:bit 開發板中加速度感測元件可得知目標物三維加速度值。該模組原理為當重力讓測量元件位移時，將使電路中兩極間電容產生變化。透過計算電容差，即可推出相對加速度大小。加速度感測器可測量在-2g 至+2g 範圍內的加速度值，其加速度值被記錄在-2048 至+2048 的範圍內。

如圖 13 所示，傳播端會分別被綁在雙腳大腿上，接收端則與電腦連接。待傳播端利用加速度感測器收集到數據後，利用廣播模組將數據以無線電波形式回傳給接收端。在同個頻道下的接收端可透過開發板左上角的天線接收電波，而板上的處理器可將其轉換為原始訊息。接收端收到訊息後，利用 pySerial 進行序列通訊。pySerial 是一個 python 函式庫，其運用序列通訊，使微控制器和電腦間互相交換數據，如圖 14 所示。



圖 13 接收端與電腦連接示意圖(研究者自製)

```
Written to CSV: ['2024-03-04 15:44:34', 544, 542, 112]
Received: (528,542,80)

Written to CSV: ['2024-03-04 15:44:34', 528, 542, 80]
Received: None

Received: (544,542,96)

Written to CSV: ['2024-03-04 15:44:34', 544, 542, 96]
Received: None

Received: (528,542,80)

Written to CSV: ['2024-03-04 15:44:34', 528, 542, 80]
Received: (544,542,96)
```

圖 14 傳輸數據的螢幕截圖(研究者自製)

三、判別是否在使用電腦

透過影像辨識技術判斷使用者是否在使用電腦，再進行錯誤姿勢與久坐之判斷。若非使用中則不進行姿勢判斷，研究者定義以下幾點為非使用中之情形：

(一) 離坐：透過 BlazePose 影像辨識模型取得臉部的 11 個節點(節點 0 ~ 10)，若有五個節點以上 visibility 小於一定值，即系統未成功辨識出人臉。另外，依據光的直線性及鏡頭的成像原理，當眼睛離螢幕越近，所偵測之雙眼間距越大，雙眼 x 差值越大；反之，當眼睛離螢幕越遠，x 差值越小。當五個節點以上 visibility 小於一定值或雙眼 x 差小於一定值，研究者定義此情況為使用者離座。

(二) 頭部非面對螢幕：當臉部的 11 個節點之中有五個節點以上 visibility 大於一定值，但雙眼 z 值差大於一定值，表示臉部有一半以上為可視狀態，且視線與螢幕的夾角大於 45 度，研究者定義此為使用者頭部非面對螢幕。

若臉部距螢幕在一定距離內且頭部面對螢幕，即符合使用電腦之情形，將繼續進行錯誤姿勢判斷。反之，則被判定為非使用電腦之情形，會在使用者介面以文字提醒。

四、定義久坐與判別是否呈現坐姿

本研究探討文獻並針對久坐進行定義：

世界衛生組織[18]表示不論任何年齡層，皆應避免久坐。歐洲聯盟表示維持坐姿的時間越少越好，且建議每 20-30 分鐘起身一次。台灣衛生福利部國民健康署[19]建議，使用 3C 產品 30 分鐘，休息 10 分鐘。因此，本研究定義 30 分鐘即為久坐。

當使用者未被判斷為離座，即臉部五個節點以上 visibility 大於一定值且雙眼 x 差值若大於一定值，則判斷使用者為坐姿。當維持坐姿時間超過 30 分鐘，將視為久坐。

五、判別姿勢正確與否

本研究依據生物學相關文獻資料[17]定義正確姿勢，以程式判斷以下幾點：

(一) 頭部水平傾斜：透過 BlazePose 影像辨識模型取得雙眼節點 y 值(節點 3、節點 6)，當兩者差大於一定值，表示使用者雙眼未呈水平，研究者定義此為頭部水平傾斜。本研究雖偵測

許多頭部節點，但此規則與以下規則皆以眼睛作為頭部衡量標準，因眼睛較不易被遮蔽且辨識準確度較高，較不易造成本研究之系統誤判。

(二) 肩膀水平傾斜：透過 BlazePose 影像辨識模型取得雙肩的節點 y 值(節點 11、節點 12)，當兩者 y 值差大於一定值，表示使用者的雙肩未呈水平，研究者定義此為肩膀水平傾斜。

(三) 駝背：透過 BlazePose 影像辨識模型取得雙眼及雙肩的 z 值，當左側肩膀及左側眼睛 z 值差大於一定值，或右側肩膀及右側眼睛 z 值差大於一定值，表示其前後距離差異大，研究者定義此為駝背。

(四) 身體側斜：透過 BlazePose 影像辨識模型取得雙眼及雙肩的 x 值(節點 3、節點 6、節點 11、節點 12)，並計算其中點。當雙肩中點與雙眼中點 x 值差大於一定值，表示其身體向側傾斜程度越大，研究者定義此為身體側斜。

(五) 臉部距離螢幕過近：當雙眼 x 差值大於一定值，則可推斷臉部距離螢幕在一定距離內；當身體相對於臀部越靠近螢幕，雙眼 z 值越小。當雙眼 x 差值大於一定值且雙眼 z 值平均小於一定值，研究者定義此為臉部距離螢幕過近。

(六) 左腳/右腳姿勢錯誤：透過 micro:bit 開發版，取得雙腳三維加速度值。使用者需先將雙腳垂直於地面，保持正確坐姿後，手動按下重置鈕，程式將紀錄此時的加速度值並定義此為初始正確數據。當偵測時 x 加速度值或 z 加速度值與初始正確數據差值大於一定值，表示腳部位置改變，可能為翹腳或抖腳等，研究者定義此為腳部姿勢錯誤。

六、資料集收集與評估指標

研究者邀請若干名年齡層介在 11 到 76 歲的受測者分別呈現上半身正確姿勢、頭部水平傾斜、肩膀水平傾斜、駝背、身體側斜、臉部距離螢幕過近、臉過於遠離螢幕、頭部非面對螢幕、下半身右腳靜止、左腳靜止、右腳上下抖動、左腳上下抖動、右腳左右抖動、左腳左右抖動，共 14 組動作，且每組動作皆維持至少 70 秒。因應硬體效能不同，每秒可收集 10 到 20 組不等數據，研究者對同一秒數據計算平均，定義其為一筆數據。由於剛開始及快結束測試時，受測者可能未維持指定的姿勢而造成干擾，因此將前 5 秒與後 5 秒數據剔除。並且剔除資料集中同一受測者數據的異常值 (Outlier)，以減少受測者未維持特定姿勢產生的誤差。根據 z 分數法（z-score），定義異常值為超出平均值正負三個標準差之數據。剔除極值後， 共收集 107687 筆數據，研究者將其以 8: 2 之比例分為訓練集與測試集。

本研究針對訓練集，以精確度(Precision)、召回率(Recall)計算調和平均(F1-score)定義判斷不同姿勢之閥值，其公式如下：

$$Precision = \frac{TP}{TP + FP}, Recall = \frac{TP}{TP + FN}, F1-score = 2\frac{precision \times recall}{precision + recall}$$

	實際為正樣本	實際為負樣本
預測為正樣本	$TP = True\ Positive$	$FP = False\ Positive$
預測為負樣本	$FN = False\ Negative$	$TN = True\ Negative$

表 3 混淆矩陣 Confusion Matrix(研究者自繪)

本研究定義正樣本為正確姿勢，其他姿勢皆為負樣本。F1-score 之值落在 0 ~ 1 值間，當 F1-score 越接近 1 時，代表閥值越準確，反之則越不準確。研究者在比對不同閥值的 F1-score 後，皆取 F1-score 最高的閥值做為最終結果。

其中，本系統之錯誤姿勢判斷的嚴格程度可由使用者自行選擇，以增加其使用意願。調整後的閥值則會透過改變 β 值，調整 Precision、Recall 的權重，其公式如下：

$$F-score = \frac{(1+\beta^2)\,Precision \times Recall}{\beta^2\,Precision + Recall}$$

本系統將以 F2-score 最高分數對應值作為較不嚴格的閾值。 F2-score 即 β=2 時的 F-score，此時 Precision 被賦予權重是 Recall 的兩倍，強調將正確姿勢判斷為正確的能力。

本研究針對測試集，以準確率(Accuracy)判斷此系統之準確率，其公式如下：

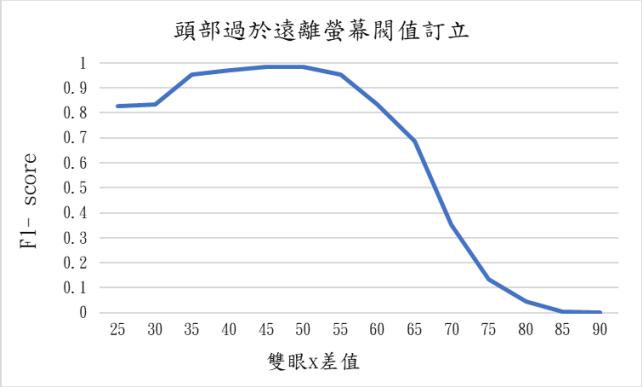
$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

七、姿勢閥值訂立

以下將討論不同姿勢下以 F- score 定義的閥值：

(一) 離坐與否判定

判定是否離坐，即判斷臉部是否可視或是否過於遠離螢幕，分別以臉部節點 visibility 值與雙眼 x 差值進行判斷。研究者分析所有有效數據，計算每秒臉部 11 個節點 visibility 平均值，發現只要使用者在鏡頭前，不論姿勢是否正確或是否遠離座位，約有 98.8%之數據 visibility 值大於 0.97。研究者計算正確姿勢與臉過於遠離螢幕時的雙眼 x 差值，發現當 x 差值訂為 47 時，其 F1-score 最接近 1，能最有效的區分此二種情形，如圖 15 所示。研究者定義當臉部一半即 5 個節點以上 visibility 小於 0.97，或當雙眼 x 差值小於 47，系統即判定其為離坐。

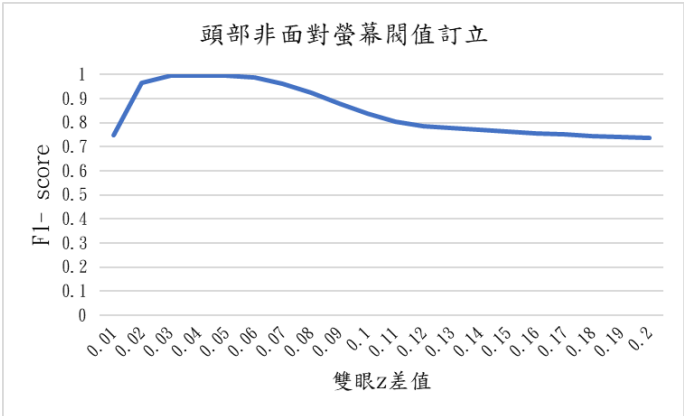


閥值	47
TP	8156
FP	107
FN	63
TN	5425
Precision	0.98705071
Recall	0.99233483
F1-score	0.98968572

圖 15 臉部距離螢幕過遠與正確姿勢雙眼 x 差值分析圖表(研究者自繪)

(二) 頭部是否面對螢幕判定

系統以雙眼 z 差值大小，判斷頭部是否面對螢幕。研究者計算正確姿勢與臉部非面對螢幕時的雙眼 z 差值，發現當 z 差值訂為 0.04 時，其 F1-score 最接近 1，能最有效的區分此二種情形，如圖 16 所示。研究者定義當雙眼 z 差值大於 0.04，系統即判定其為頭部非面對螢幕。



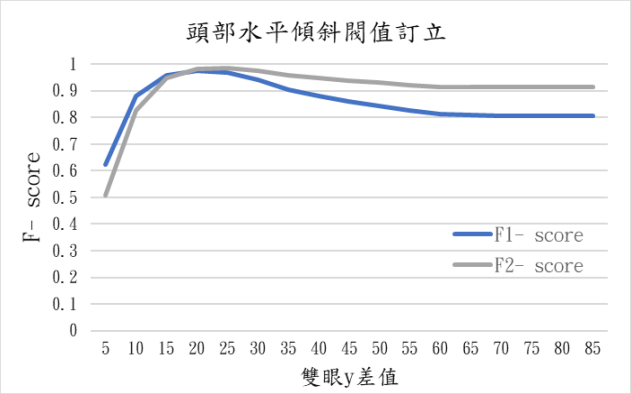
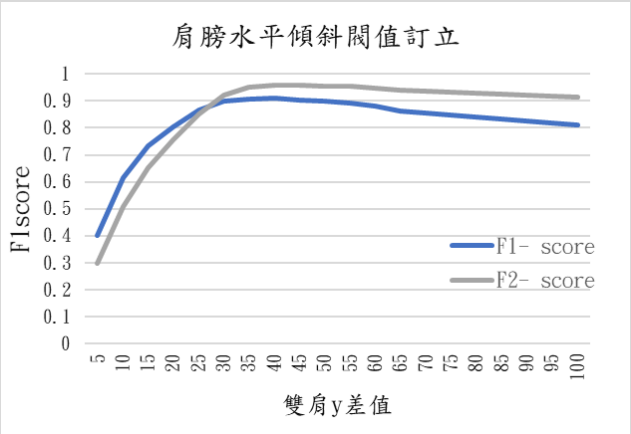
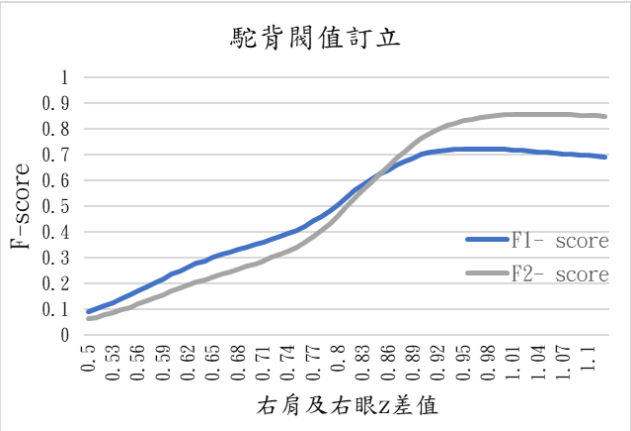
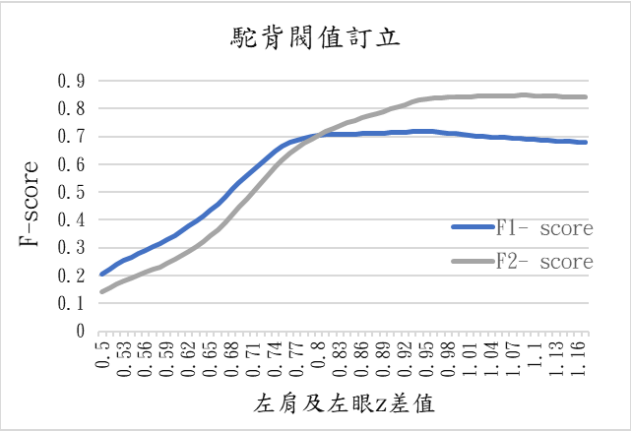
閥值	0.04
TP	10091
FP	2
FN	67
TN	7258
Precision	0.999802
Recall	0.993404
F1- score	0.996593

圖 16 臉部非面對螢幕與正確姿勢雙眼 z 差值分析圖表(研究者自繪)

(三) 正確姿勢判定

1. 上半身姿勢

研究者計算正確姿勢與錯誤姿勢在不同閥值下的 F1-score 與 F2- score，並取 F-score 最接近 1 的閥值作為判定姿勢錯誤與否的標準，分析如下表所示：

錯誤姿勢種類	判斷依據在不同閥值下的 F-score	閥值對應評估指標之值
頭部水平傾斜		種類 / F1- score / F2- score 閥值 22 23 TP 10357 10371 FP 471 525 FN 14 0 TN 4481 4427 Precision 0.95650166 0.9518172 Recall 0.99865008 1 對應最高分數 0.97712156 0.9899771
肩膀水平傾斜		種類 / F1- score / F2- score 閥值 40 42 TP 10500 10540 FP 2063 2124 FN 42 2 TN 3762 3701 Precision 0.83578763 0.8322805 Recall 0.99601594 0.9998103 對應的分數 0.90889418 0.9611176
駝背		種類 / F1- score / F2- score 閥值 0.97 1.04 TP 9508 9930 FP 6807 8029 FN 494 72 TN 3026 1804 Precision 0.58277659 0.5529261 Recall 0.95060988 0.9928014 對應最高分數 0.72257476 0.8565218
		種類 / F1- score / F2- score 閥值 0.94 1.08 TP 9526 10246 FP 6670 9068 FN 754 34 TN 3275 877 Precision 0.58816992 0.530496 Recall 0.9266537 0.9966926 對應最高分數 0.71959511 0.8477016

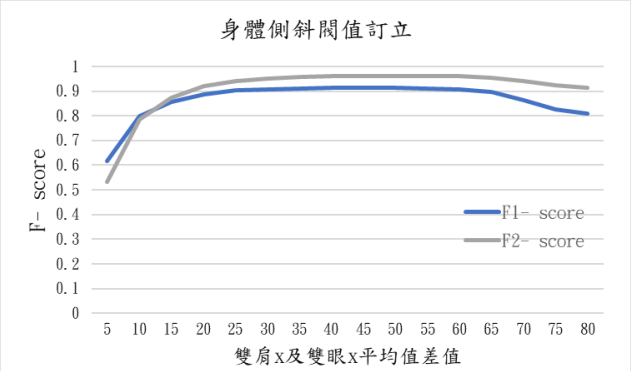
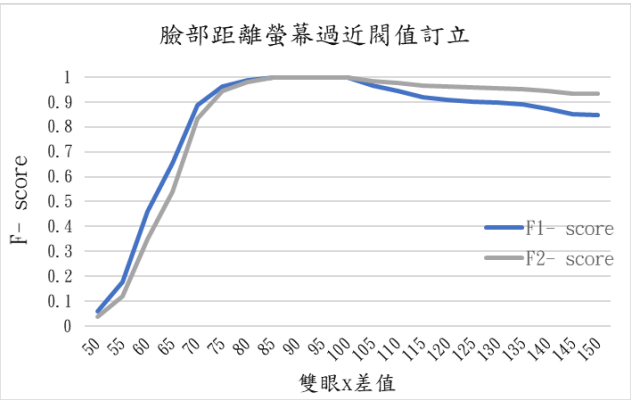
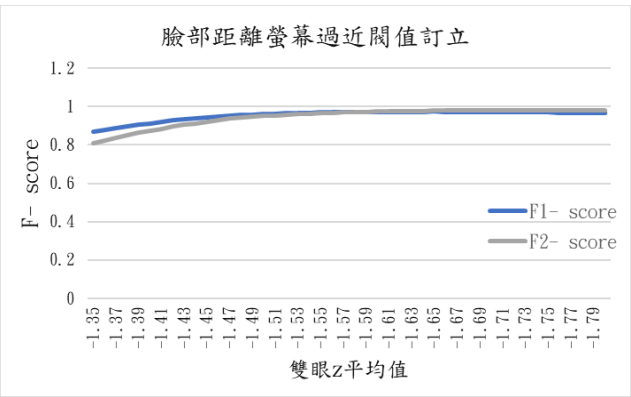
<table>
<tr>
<td>身體側斜</td>
<td>
</td>
<td>
<table>
<tr><th>種類</th><th>F1- score</th><th>F2- score</th></tr>
<tr><td>閥值</td><td>45</td><td>50</td></tr>
<tr><td>TP</td><td>10568</td><td>10581</td></tr>
<tr><td>FP</td><td>1943</td><td>1967</td></tr>
<tr><td>FN</td><td>40</td><td>27</td></tr>
<tr><td>TN</td><td>4987</td><td>4963</td></tr>
<tr><td>Precision</td><td>0.84469667</td><td>0.843242</td></tr>
<tr><td>Recall</td><td>0.99622926</td><td>0.9974548</td></tr>
<tr><td>對應最高分數</td><td>0.91422639</td><td>0.962259</td></tr>
</table>
</td>
</tr>
<tr>
<td rowspan="2">臉部距離螢幕過近</td>
<td>
</td>
<td>
<table>
<tr><th>種類</th><th>F1- score</th><th>F2- score</th></tr>
<tr><td>閥值</td><td>90</td><td>91</td></tr>
<tr><td>TP</td><td>8219</td><td>8219</td></tr>
<tr><td>FP</td><td>4</td><td>4</td></tr>
<tr><td>FN</td><td>0</td><td>0</td></tr>
<tr><td>TN</td><td>6193</td><td>6193</td></tr>
<tr><td>Precision</td><td>0.99951356</td><td>0.9995136</td></tr>
<tr><td>Recall</td><td>1</td><td>1</td></tr>
<tr><td>對應最高分數</td><td>0.99975672</td><td>0.9999027</td></tr>
</table>
</td>
</tr>
<tr>
<td>
</td>
<td>
<table>
<tr><th>種類</th><th>F1- score</th><th>F2- score</th></tr>
<tr><td>閥值</td><td>-1.65</td><td>-1.74</td></tr>
<tr><td>TP</td><td>10427</td><td>10493</td></tr>
<tr><td>FP</td><td>376</td><td>523</td></tr>
<tr><td>FN</td><td>181</td><td>115</td></tr>
<tr><td>TN</td><td>7498</td><td>7351</td></tr>
<tr><td>Precision</td><td>0.96519485</td><td>0.9525236</td></tr>
<tr><td>Recall</td><td>0.98293741</td><td>0.9891591</td></tr>
<tr><td>對應最高分數</td><td>0.97398533</td><td>0.9816083</td></tr>
</table>
</td>
</tr>
</table>

表 4 錯誤姿勢與正確姿勢閥值分析圖表(研究者自繪)

研究者定義在正常模式中，當雙眼 y 差值大於 22，系統即判定其為頭部水平傾斜；當雙肩 y 差值大於 40 時，判定為肩膀水平傾斜；當左眼與肩 z 差值大於 0.97，或右眼與肩 z 差值大於 0.94 時，判定為駝背；當雙肩中點與雙眼中點 x 值差大於 45，判定為身體側斜；當雙眼之 x 差值大於 90，且雙眼 z 平均值小於-1.65 時，判定其為臉部距離螢幕過近。

研究者定義在較不嚴格的模式中，當雙眼 y 差值大於 23，系統即判定其為頭部水平傾斜；當雙肩 y 差值大於 42 時，判定為肩膀水平傾斜；當左眼與肩 z 差值大於 1.04，或右眼與肩 z 差值大於 1.08 時，判定為駝背；當雙肩中點與雙眼中點 x 值差大於 50，判定為身體側斜；當雙眼之 x 差值大於 91，且雙眼 z 平均值小於-1.74 時，判定其為臉部距離螢幕過近。

2.下半身姿勢

判定是否左右腳姿勢錯誤，即分別判斷左右腳是否左右抖動、上下抖動或蹺腳，分別 x 加速度量值變化量與 z 加速度量值變化量進行判斷。以下圖表為左右腳靜止與不同方向抖腳的分析：

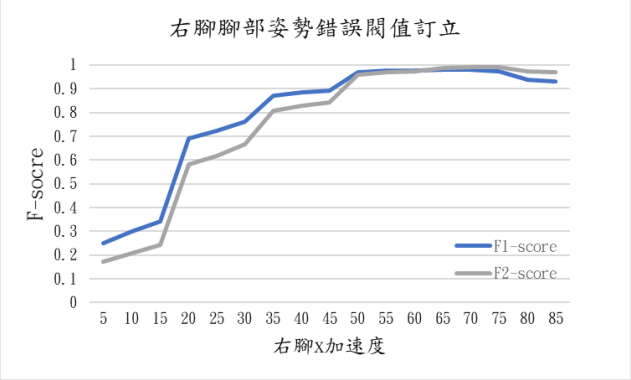
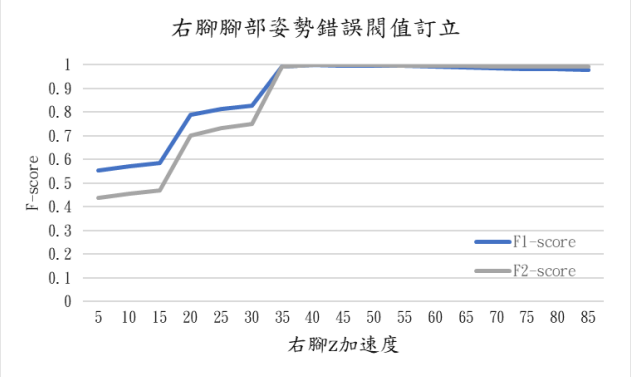
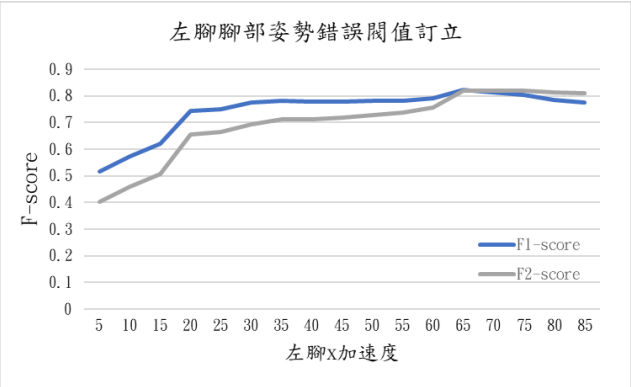
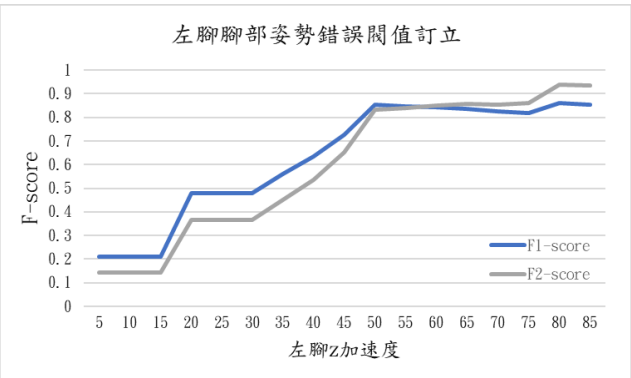
<table>
<tr><th>錯誤姿勢種類</th><th>判斷依據在不同閥值下的 F-score</th><th>閥值對應的值</th></tr>
<tr><td rowspan="2">右腳姿勢錯誤</td><td>
</td><td>
<table>
<tr><th>種類</th><th>F1- score</th><th>F2- score</th></tr>
<tr><td>閥值</td><td>50</td><td>70</td></tr>
<tr><td>TP</td><td>4161</td><td>4220</td></tr>
<tr><td>FP</td><td>42</td><td>165</td></tr>
<tr><td>FN</td><td>63</td><td>4</td></tr>
<tr><td>TN</td><td>2623</td><td>2500</td></tr>
<tr><td>Precision</td><td>0.9900071</td><td>0.9623717</td></tr>
<tr><td>Recall</td><td>0.9850852</td><td>0.999053</td></tr>
<tr><td>對應最高分數</td><td>0.98754005</td><td>0.9914948</td></tr>
</table>
</td></tr>
<tr><td>
</td><td>
<table>
<tr><th>種類</th><th>F1- score</th><th>F2- score</th></tr>
<tr><td>閥值</td><td>50</td><td>50</td></tr>
<tr><td>TP</td><td>4220</td><td>4220</td></tr>
<tr><td>FP</td><td>23</td><td>23</td></tr>
<tr><td>FN</td><td>4</td><td>4</td></tr>
<tr><td>TN</td><td>2544</td><td>2544</td></tr>
<tr><td>Precision</td><td>0.99457931</td><td>0.9945793</td></tr>
<tr><td>Recall</td><td>0.99905303</td><td>0.999053</td></tr>
<tr><td>對應最高分數</td><td>0.996811</td><td>0.9981551</td></tr>
</table>
</td></tr>
<tr><td rowspan="2">左腳姿勢錯誤</td><td>
</td><td>
<table>
<tr><th>種類</th><th>F1- score</th><th>F2- score</th></tr>
<tr><td>閥值</td><td>64</td><td>65</td></tr>
<tr><td>TP</td><td>6559</td><td>6565</td></tr>
<tr><td>FP</td><td>1096</td><td>1110</td></tr>
<tr><td>FN</td><td>39</td><td>33</td></tr>
<tr><td>TN</td><td>5497</td><td>5483</td></tr>
<tr><td>Precision</td><td>0.8568256</td><td>0.8553746</td></tr>
<tr><td>Recall</td><td>0.99408912</td><td>0.9949985</td></tr>
<tr><td>對應最高分數</td><td>0.82190737</td><td>0.8926079</td></tr>
</table>
</td></tr>
<tr><td>
</td><td>
<table>
<tr><th>種類</th><th>F1- score</th><th>F2- score</th></tr>
<tr><td>閥值</td><td>80</td><td>81</td></tr>
<tr><td>TP</td><td>6582</td><td>6596</td></tr>
<tr><td>FP</td><td>2089</td><td>2110</td></tr>
<tr><td>FN</td><td>16</td><td>2</td></tr>
<tr><td>TN</td><td>2544</td><td>2523</td></tr>
<tr><td>Precision</td><td>0.759082</td><td>0.7576384</td></tr>
<tr><td>Recall</td><td>0.99757502</td><td>0.9996969</td></tr>
<tr><td>對應最高分數</td><td>0.86213897</td><td>0.9396547</td></tr>
</table>
</td></tr>
</table>

表 5 錯誤姿勢與正確姿勢閥值分析圖表(研究者自繪)

研究者觀察翹腳數據，發現其 x 或 z 加速度變化明顯高於系統定義閥值。因此研究者定義在正常模式中，當 x 加速度量變化量大於 50 或 z 加速度量變化量大於 50，系統即判定其為右腳姿勢錯誤；當 x 加速度量變化量大於 64 或 z 加速度量變化量大於 80，系統即判定其為左腳姿勢錯誤。

在較不嚴格的模式中，當 x 加速度量變化量大於 70 或 z 加速度量變化量大於 50，系統即判定其為右腳姿勢錯誤；當 x 加速度量變化量大於 65 或 z 加速度量變化量大於 81，系統即判定其為左腳姿勢錯誤。

八、使用者互動與回饋介面

本研究設計一個操作簡易的使用者介面，讓其選擇偏好的回饋方式，以貼合使用者基本需求，增加其使用意願。並在姿勢錯誤或久坐時，給予即時回饋，如在介面中顯示與執行提醒機制，以便使用者即時調整姿勢。最終提供錯誤姿勢分析，以供使用者改善自身習慣。

本研究採用 PyQt 製作使用者介面：

PyQt 是一個使用 Python 製作程式介面的套件，此套件配備 QtDesigner，以視覺化的方式讓開發者設計界面，其操作介面如圖 17 所示。可將 QtDesigner 製作出的 ui 檔，轉成 py 檔，並以 python 語言撰寫功能，使按下按鈕時執行功能或顯示不同文字。

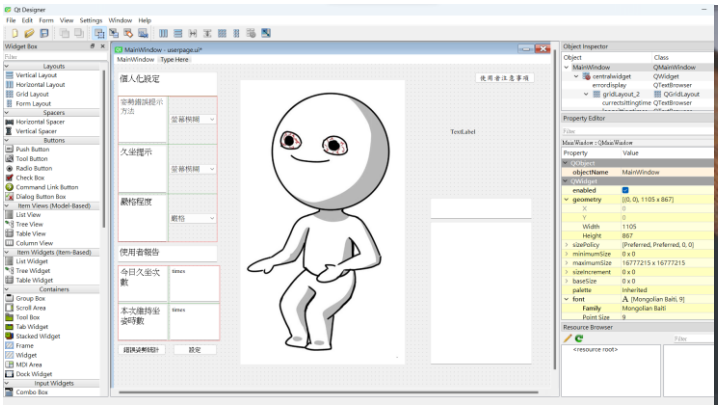


圖 17 Qt Designer 操作介面示意圖(研究者自製)



圖 18 系統主介面(研究者自製)



圖 19 姿勢判斷介面(研究者自製)

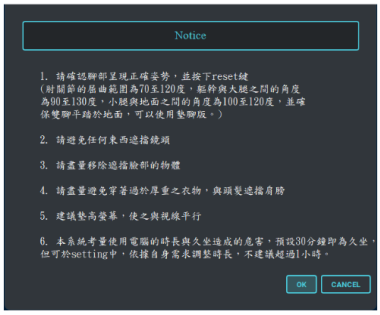


圖20 注意事項介面 (研究者自製)



圖 21 錯誤姿勢統計圖表介面(研究者自製)



圖 22 久坐提醒 -- 模糊螢幕(研究者自製)

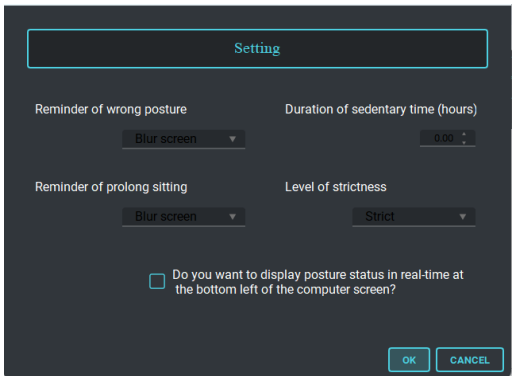


圖 23 設定介面(研究者自製)

圖 18 為本系統的主介面，按下 start detecting 即可開始姿勢偵測，其中偵測時的介面如圖 19 所示，圖中，可見使用者目前呈現坐姿的時間已達久坐標準，系統於介面顯示粗體文字並將對應時間的文字框以藍底呈現，此時會執行使用者設定的提醒機制，模糊螢幕畫面的效果如 22 圖所示。notice 按鈕連結使用者提醒事項，如圖 20 所示。setting 則連結使用者設定介面，如圖 23 所示，使用者可自行設定久坐或姿勢錯誤的提醒機制；設定判斷姿勢正確與否的的嚴格程度，預設為嚴格；調整判斷為久坐的時長，系統預設為 30 分鐘，以貼合使用者需求，增加其使用意願。user report 按鈕連結錯誤姿勢分析介面，分析使用者各姿勢的錯誤率，以供使用者改善自身習慣，如圖 21 所示。

陸、討論

一、姿勢判斷準確率分析

研究者針對測試集，以 accuracy 討論各姿勢判斷的準確率。

(一) 離座：當臉部 5 個節點以上 visibility 小於 0.97，或當雙眼 x 差值小於 47 判斷為離坐的準確率約為 90.57%。

(二) 頭部非面對螢幕：當雙眼 z 差值大於 0.04 判斷為離座的準確率約為 94.13%，顯示該閥值在判定頭部是否面對螢幕的效果極佳。

(三) 頭部水平傾斜：當雙眼 y 差值大於 22 判斷為頭部水平傾斜的準確率約為 96.22%，顯示該閥值在區分是否頭部水平傾斜的效果極佳。

(四) 肩膀水平傾斜：當雙肩 y 差值大於 40 判斷為肩膀水平傾斜的準確率約為 83.91%。

(五) 駝背：當左眼與肩 z 差值大於 0.97，或右眼與肩 z 差值大於 0.94 時判斷為駝背的準確率為 63.01%。研究者推測，駝背的動作對於眼睛和肩膀的深度差異本身不大，因此透過 Blazepose 取得的 z 差值在正確或前傾時的差異亦不大，因此容易造成誤判。研究者認為單透過電腦內建鏡頭，難以準確判斷是否駝背，可考慮另架鏡頭，或是透過紀錄使用者體型交叉對比，以更精確的判斷是否駝背。

(六) 身體側斜：當雙肩中點與雙眼中點 x 值差大於 45，判斷為身體側斜 45 的準確率為 70.02%。研究者推測，不同受測者身體側斜幅度差異較大，有些使用者雖然軀幹稍微向側傾斜，但頭部傾斜幅度不大，因此造成準確度不如預期。此閥值對於傾斜幅度較大的判斷效果較佳，研究者認為另外尋找特徵，以兩個條件共同判斷即可改善此問題。

(七) 臉部距離螢幕過近：當雙眼之 x 差值大於 90，且雙眼 z 平均值小於-1.65 時，判斷為臉部距離螢幕過近的準確率為 96.50%。顯示該閥值在區分臉部是否距離螢幕過近的效果極佳。

(八) 右腳姿勢錯誤：當 x 加速度量變化量大於 50 或 z 加速度量變化量大於 50，判斷為右腳姿勢錯誤的準確率為 87.33%，顯示該閥值在判定右腳姿勢是否錯誤的效果佳。

(九) 左腳姿勢錯誤：當 x 加速度量變化量大於 64 或 z 加速度量變化量大於 80，判斷為左腳姿勢錯誤的準確率為 86.99%。

二、影像辨識特殊情形討論(上半身)

(一) 使用者頭部未面對鏡頭，系統誤判為離坐

當使用者以側臉面對鏡頭時，因非正臉將導致節點擠壓偏移而使雙眼 x 差值極小，如圖 25 所示。所得數據與離鏡頭過遠時相似，將使本研究之系統誤判側臉為離坐，導致坐姿維持時間累計中斷而錯誤統計久坐時間。為解決此問題，研究者進一步發現側坐與正坐雙眼的 z 值差有顯著的差異，如圖 24 所示。本系統將在判定是否離坐前，判斷是否有頭部未面對鏡

頭之情形，定義此情形為臉部五個節點以上 visibility 大於一定值，且雙眼 z 差值大於一定值。

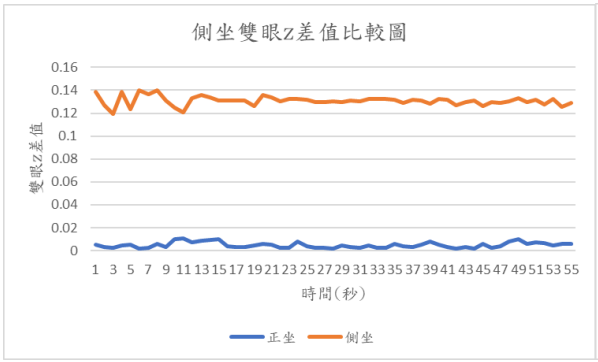


圖 24 側坐雙眼 z 差值比較圖(研究者自繪)

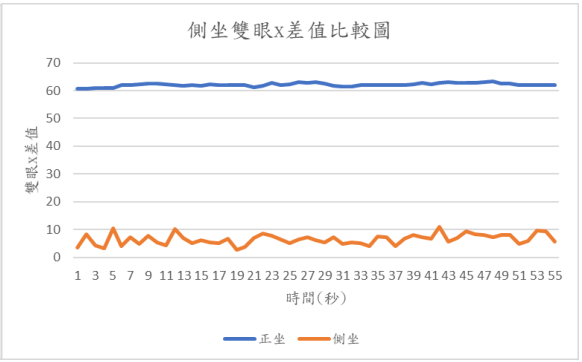


圖 25 側坐雙眼 x 差值比較圖(研究者自繪)

(二) 雙肩的節點遮擋，系統誤判為肩膀水平傾斜

當厚重衣物或是頭髮遮擋住肩膀，將使雙肩 y 值差產生些微誤差，可能使系統判定為肩膀水平傾斜，如圖 26 所示。在圖 27 中，請相同受試者維持同個正確姿勢，分別展示衣著厚重外套與衣著輕便衣物各 70 秒，並將兩者 y 值差進行比較。研究者發現穿著衣物，遮擋肩膀時 y 值差明顯比較大。

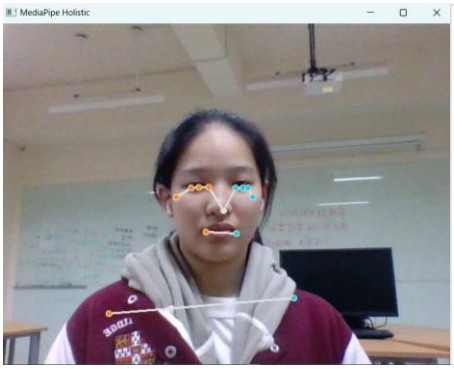


圖 26 衣服導致肩膀節點誤差示意圖 (研究者自製)

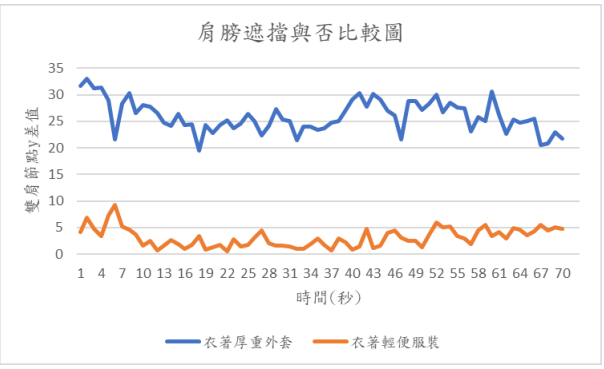


圖 27 肩膀節點遮擋與否 y 值差比較圖 (研究者自繪)

模型在訓練時因有進行遮蔽模擬(occlusions augmentation)，因此遮蔽肩膀左右側的節點依然會進行預測，不過無法像未被遮蔽的節點一樣準確而有誤差。

研究者在蒐集資料時有加入穿著較厚重衣物的數據來訂立閥值，經測試後發現，衣著帽踢、薄外套的誤差不會造成系統誤判，但衣著羽絨外套或有較厚的帽子時，將導致系統誤判為肩膀水平傾斜。為避免發生此情形，研究者於系統中，加入使用者注意事項介面以提醒使用者，避免衣著過於厚重之衣物。並於即時顯示錯誤姿勢的文字框中加入描述，若持續顯示錯誤且認為肩膀未傾斜亦可能是衣物遮擋造成。

(三) 使用者離坐期間，模型誤抓取環境物件

當使用者離坐時，模型有時會在環境背景尋找物件並放上節點，如圖 30 所示。原先系統僅採用 visibility 判斷是否離座，然而此時節點 visibility 可能會大於 0.99，導致系統誤判為使用者呈現坐姿。

研究者推測原因如下：

當使用者離坐，區域中臉部消失，BlazePose 以 NMS 演算法重新偵測人臉。此時會保留 BBox 信心值最高的物件，然而此時無明顯的臉部區域，故保留框有其他物體的 BBox。接著，追蹤此區域並放置節點。在圖 28 中，請受測者在第 20 秒時離開座位，並於第 115 秒重回位子，並記錄其臉部 11 個節點 visibility 值平均的變化。從圖中可觀察到，離開後臉部平均可視程度降低，此時模型重新偵測可能的臉部區域；離坐期間，可視程度明顯上升，表示找到了可能的臉部區域，但其值不穩定；回坐後，模型偵測到受測者的臉部，而可視程度快速提高，最後趨於穩定。

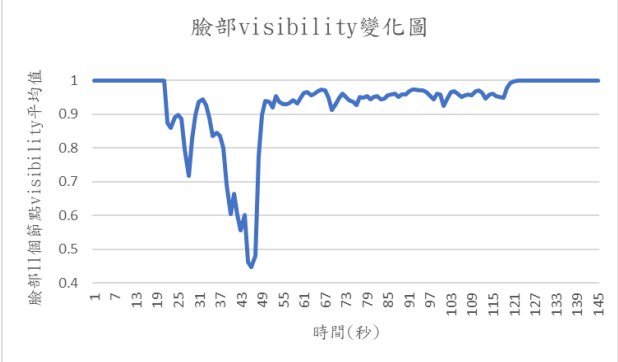


圖 28 臉部 visibility 變化圖(研究者自繪)

研究者蒐集五個不同環境離坐後 60 秒之臉部節點平均 visibility 變化值，如圖 34 所示。研究者發現，其中共有 30 筆數據 visibility 值大於閥值，將導致系統誤判為使用者呈現坐姿，久坐統計錯誤，並接續判斷側臉及坐姿正確與否。由圖 29 至圖 33 可知，不同環境背景，BlazePose 偵測到的背景物件之 visibility 有顯著的差異。研究者發現當背景中有圓形或較不具稜角的輪廓時，模型較容易偵測到此物體且 visibility 值較高。

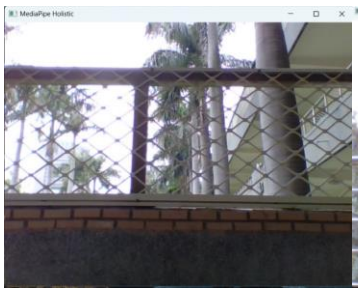


圖 29 環境一

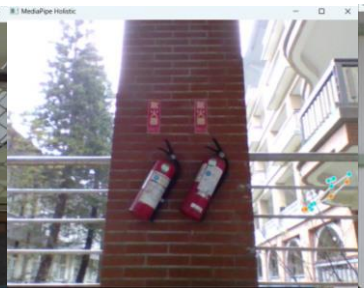


圖 30 環境二

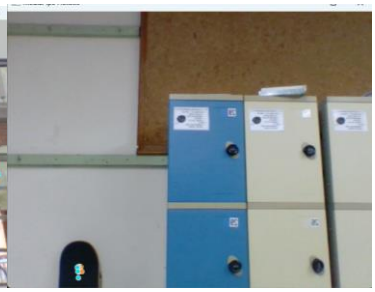


圖 31 環境三

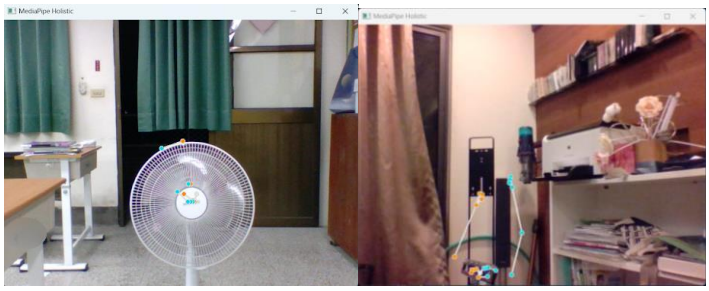


圖 32 環境四　　圖 33 環境五

(圖 29、圖 30、圖 31、圖 32、圖 33 研究者自製)

研究者亦發現背景抓取到的物件因其不具五官，故臉部 11 個節點較集中，導致雙眼 x 差值較小，如圖 35 所示。各種不同環境中，抓取到的 x 差值皆小於 20，而使用者呈現正確姿勢時的差值平均為 65.36，兩者有明顯落差。研究者加入 x 差值進行判斷即可排除此誤判情形，故重新定義離坐為臉部 5 個以上的節點 visibility 平均小於一定值或雙眼 x 差值小於一定值。

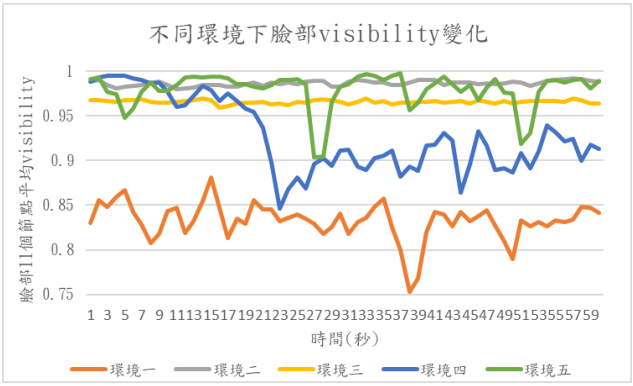


圖 34 不同環境下臉部 visibility 變化(研究者自繪)

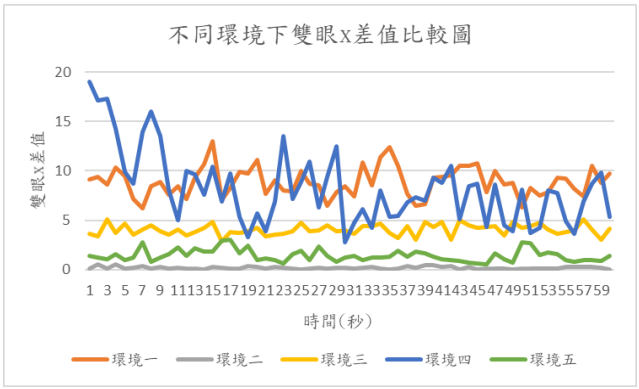


圖 35 不同環境下雙眼 x 差值比較圖(研究者自繪)

(四) 臉部被遮擋

研究者依據被遮擋的時機分為兩種情形討論，第一種為程式一開始臉部即被遮擋，第二種為程式已抓到臉部，但使用者於中途遮擋。

1. 程式一開始臉部即被遮擋

此情形常造成節點完全不準確，如圖 36 所示，將造成多數姿勢判斷錯誤，其中被判定為離坐最為顯著，因其偵測到的物體多雙眼 x 差值過小或 visibility 小於 0.99。研究者推測，此時 BlazePose 模型在第一步驟，偵測臉部即發生錯誤。因遮擋住臉部重要特徵，模型在進行 NMS 演算法時，框住臉部的 BBox 信心值很低，而被刪除，最後保留信心值最高的物件。

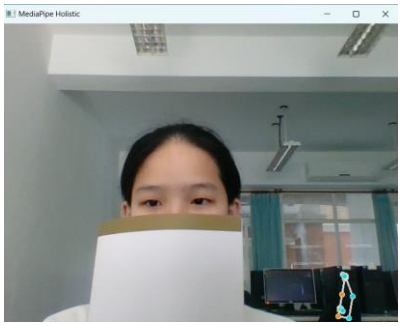


圖 36 程式一開始臉部即被遮擋誤判示意圖(研究者自製)

2. 程式已抓到臉部，使用者於中途遮擋

此情形將使節點偏移。研究者將依遮擋程度分成兩個情況討論，輕微遮擋及嚴重遮擋。輕微遮擋為臉部小部分被遮擋或遮擋物貼合臉型，如配戴口罩，如圖 37 所示。第二種為嚴重遮擋，此情形中臉部有一半以上的節點被遮蔽，且此遮擋物不貼合臉型，如圖 38 所示。

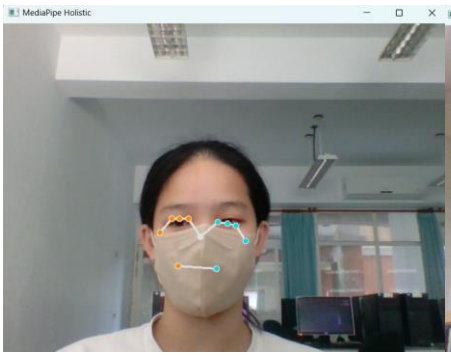


圖 37 臉部輕微遮擋示意圖(研究者自製)

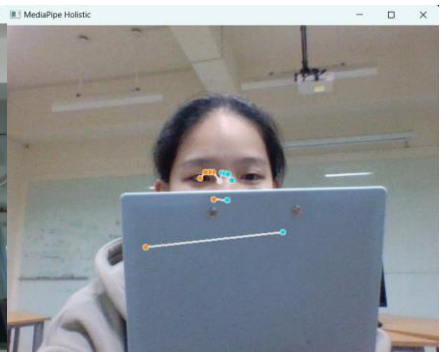


圖 38 臉部嚴重遮擋示意圖(研究者自製)

中途遮擋臉部將導致許多姿勢判斷錯誤，其中系統誤判為離坐的情形特別顯著。在圖 39 中，請相同受測者維持同一個正確姿勢，依序分別展示臉部無遮擋、配戴口罩、及嚴重遮擋的情形，各維持 60 秒，並比較三者雙眼 x 差值，發現遮擋越嚴重，x 差值越小，越有機會造成系統誤判為離坐。

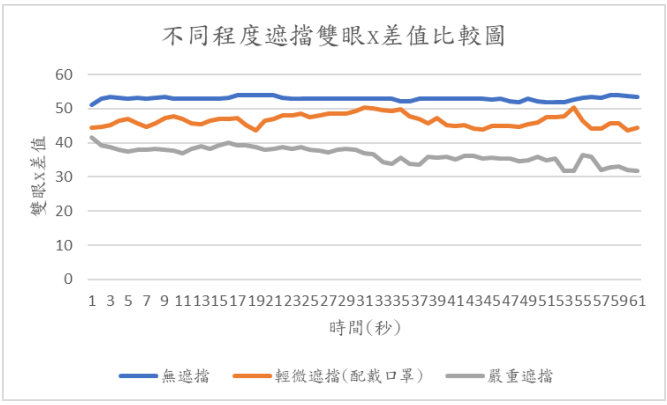


圖 39 不同程度遮擋雙眼 x 差值比較圖(研究者自繪)

研究者推測節點產生偏移的原因如下：

當 BlazePose 模型開啟並抓取到臉部之後會持續追蹤此區域。當臉部被遮擋時，模型會依據先前抓取到的區域，與模型訓練時進行的遮蔽模擬(occlusions augmentation)推測節點的位置，但因資訊不足，導致節點位置不準確。

研究者發現，輕微遮擋所造成的節點偏移在多數時間，不會影響系統判斷。研究者在蒐集資料時，特別加入配戴口罩的數據來訂立閥值。

臉部於系統一開始就被遮擋以及中途嚴重遮擋，皆可能導致系統誤判為離坐。為避免發生此情形，研究者於使用者注意事項介面提醒使用者，盡量避免遮擋臉部。並於即時顯示錯誤姿勢的文字框中加入描述，系統判定為離坐，請避免遮擋臉部，使臉部為可視狀態。

(五) 鏡頭前出現第二個人

1. 兩人臉部位置相近，臉部節點偏移

當經過者臉部比受測者的更靠近鏡頭，且兩個距離位置相近時，受測者的臉部節點偏移，如圖 40 所示。此時臉部節點 visibility 值多大於 0.99，且雙眼 x 差值大於 49，系統判斷使用者呈現坐姿，進而判斷其坐姿正確性，導致多數姿勢判斷錯誤，研究者推測原因如下：

BlazePose 模型在第一次尋找完人臉後，會持續追蹤此臉部區域。當區域中臉部消失或模型判斷臉部有變動，準確率降低時，模型才再以 NMS 演算法計算人臉。當經過者臉部比受測者的更靠近鏡頭，且兩個臉部距離相近時，原先偵測的臉部區域準確率降低，將觸發模型再次以 NMS 演算法計算人臉位置。此時，同時框著兩個臉部的 BBox 信心值，較框著單一臉部的 BBox 高。進行 IoU 運算後，將保留框著兩個臉部的 BBox，並將臉部節點放置於此 BBox 中，故造成節點位置不準確。

但若經過之人的臉部比使用者更靠近螢幕，且兩者距離差異大時，則不會影響原先追蹤的臉部區域，故不會重新偵測臉部位置，模型不會誤判，如圖 41 所示。



圖 40 臉部靠近節點誤抓示意圖　圖 41 臉部相對距離遠節點不誤抓示意圖

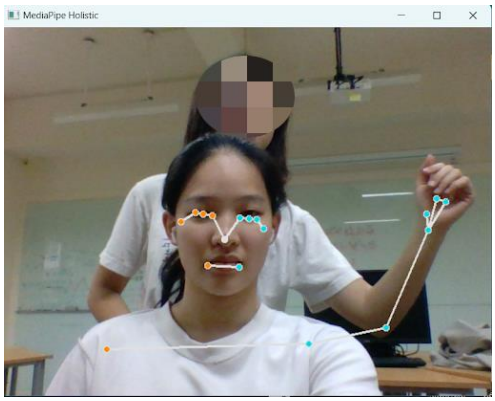
(研究者自製)　　(研究者自製)

為解決此問題，研究者另加入 FaceMesh 影像辨識模型找出 visibility 最高的臉部，並將影像切割成其人體大小，以排除其餘人臉之干擾。

2. 抓取經過之人的手部節點

當後面有人經過，且手出現在使用者合理的手部位置時，程式有機會抓取經過之人的手部節點，如圖 42 所示。研究者推測使用者手部節點 visibility 值過低，當找到另一個 visibility 值較高的節點，則將其判定其為使用者之手部節點。在圖 43 中，請受測者坐於電腦前方，在第 61 秒時，後方加入第二個人，將手舉至受測者合理之手部位置，並維持此姿勢至第 100 秒。第 61 秒時，系統抓到經過的人之手肘，visibility 逐漸上升。

當經過之人的手部與使用者合理的手部位置差異過大，則不會抓取錯誤節點，如圖 68 所示。研究者推測 BlazePose 模型依據人臉與中心垂線推測節點大致的相對位置，當經過者之手部節點位置與使用者合理的位置差異過大，即不會判斷其為使用者的節點。



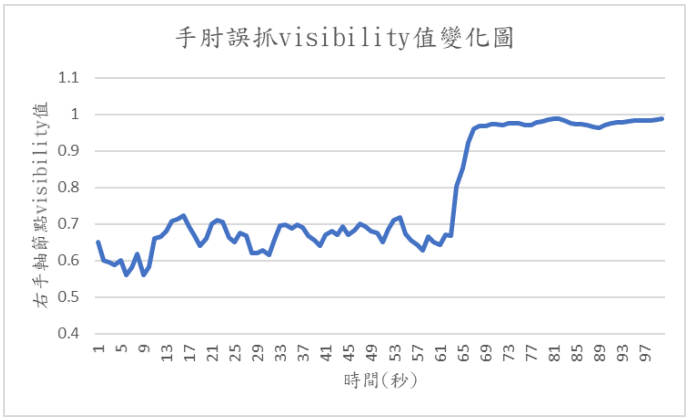


圖 42 抓取錯誤的手肘節點示意圖　圖 43 抓取錯誤的手肘節點時 visibility 變化圖

(研究者自製)　　(研究者自繪)



圖 44 手肘遠離系統不誤抓示意圖(研究者自製)

研究者在判定坐姿是否正確時，並不會使用手部節點資訊，因此此類誤判不會對本研究的結果造成影響。

(六) 太靠近螢幕時可能被判定為離座

當臉部過度靠近螢幕，鏡頭無法抓取到臉部節點，此時 visibility 小於 0.99，與離座的表現相同。離螢幕過近時即會被判定為離座，此時維持坐姿時間累計會被中斷，造成久坐統計不正確。研究者將於即時顯示錯誤姿勢的文字框中描述，目前判定為離座，請將臉部放置於鏡頭可拍攝範圍，以提醒使用者。

(七) 肩膀位於鏡頭抓取到的交界

使用筆電內建鏡頭時，為配合使用角度，肩膀剛好位於鏡頭抓取到的邊界下方，有時模型會將雙肩節點上移，放置於鏡頭抓取的邊界，造成系統的肩膀水平傾斜誤判。研究者發現，此情形的重現性不高，對於系統整體未有顯著影響。但若是使用可調節角度的外接鏡頭，使鏡頭可以拍攝到完整的肩膀，即可完全避免此情形的發生。

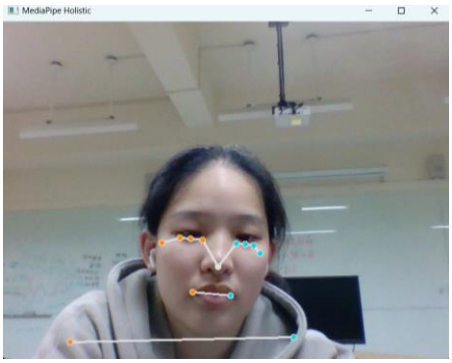


圖 45 肩膀位於鏡頭交界傾斜誤判示意圖(研究者自製)

(八) 離座後，重回位置瞬間節點不正確

離座後，重回位置的瞬間，節點位置常會不正確或是無法抓取到節點，導致久坐及姿勢判斷錯誤。而經過一段時間後，節點位置會變得更準確。研究者推測其原因如下：

根據 BlazePose 的底層技術，當人臉消失時，模型會重新開始偵測臉部，之後節點則會根據前一次偵測找到的人臉放置，以使模型更即時。當人臉消失時，此模型將嘗試抓取臉

部，而這些臉部節點的 visibility 值不高。此時 BlazePose 將判定此為人臉消失，不斷透過 NMS 演算法尋找臉部區域，因而減低整體運算效率。當使用者回坐時，要隔一段時間才能抓到臉部區域再進行節點放置，過程中節點會變得較準確，visibility 值較高。

重回座位後，所需放置正確節點的時間不大相同，取決於許多因素，如離坐時間長短及電腦效能等。研究者發現，此情形的重現性不高，對於系統整體未有顯著影響。

三、感測裝置資料傳輸(下半身)

(一) micro:bit 資料傳輸

當 micro:bit 使用無線廣播模組傳輸數據時，有時會有多傳或少傳數據之現象。無線電波為電磁波的一種，所以在傳播的途中會發生反射、折射、繞射或被吸收等現象。研究者因此推測這可能是因為 micro:bit 在進行無線傳輸時受到天氣、地形、傳播距離，亦或者是受到其他訊號干擾所造成之錯誤，將於系統內對不完全之下半身感測資料予以過濾和排除。

	A	B	C	D	E
1	16	7	0	1	
2	16	7	16		
3	16	7	16		
4	16	7	16		
5	16	7	16		

圖 46 錯誤傳輸示意圖(研究者自製)

柒、結論

一、總結

(一) 姿勢訓練資料集

目前多數公開資料集僅標記骨架節點，而未標記整體姿勢正確與否。因此雖可用於骨架辨識訓練用途，但無法用於後續判別姿勢正確與否使用。因此本研究邀請若干名 11 歲到 76 歲不同年齡層男性與女性受測者，於鏡頭前呈現正確與錯誤不同坐姿，最終共收集 107687 筆有效數據。

(二) 不同姿勢閥值訂立

上半身透過電腦內建鏡頭抓取骨架節點資訊，下半身透過雙腳穿戴 micro:bit 感測裝置取得腳部三維加速度值，取得上下半身資訊，並透過以下閥值判斷使用者姿勢。另外以 F2 - score 定義較不嚴格的錯誤姿勢判斷閥值。

姿勢類別	規則判斷與閥值	此閥值之 F1-score	此閥值之準確率
離座	雙眼 x 差<47 或 一半以上	0.989686	90.57%

	visibility<0.97		
頭部非面對螢幕	雙眼 z 差>0.04	0.996593	94.13%
頭部水平傾斜	雙眼 y 差>22	0.977122	96.22%
肩膀水平傾斜	雙肩 y 差>40	0.908894	81.37%
駝背	右眼與右肩 z 差>0.97 或 左眼與左肩 z 差>0.94	0.722575 / 0.719595	63.01%
身體側斜	雙肩中點與雙眼 中點 x 差>45	0.914226	70.02%
臉部距離螢幕過近	雙眼 x 差>90 且 雙眼 z 平均< -1.65	0.991615 / 0.973653	96.50%
右腳姿勢錯誤	右腳 x 加速度差>50 且 右腳 z 加速度差>50	0.987540 / 0.996811	87.33%
左腳姿勢錯誤	左腳 x 加速度差>64 且 左腳 z 加速度差>80	0.920368 / 0.862139	86.99%

表 6 各姿勢閥值訂立表格(研究者自繪)

(三) 互動介面系統實作

本研究開發一套互動介面系統，讓使用者選擇其偏好互動方式，並於姿勢錯誤時或久坐時即時回饋，亦提供錯誤姿勢分析，供使用者調整自身習慣。

(四) 誤判情形討論

本研究發現當使用者頭部非面對鏡頭，系統易誤判為離坐，研究者以雙眼 z 差值判別此情形，並於系統中增加頭部非面對鏡頭的判斷以修正此問題。而當使用者離坐期間，模型誤抓取環境物件，因此系統加入雙眼 x 差值判斷以排除此類問題。當節點被遮蔽，如穿著厚重衣物或配戴口罩時，系統易產生誤判，研究者加入遮擋後數據訓練並在出現誤判時提醒使用者，以修正此問題。

(五) 本研究相較於現有研究

相較於現有姿勢偵測之產品，本研究不需讓使用者先擺出正確姿勢，即可偵測使用者坐姿正確與否。且本研究系統僅需使用內建鏡頭與三塊 MicroBit 開發板即可同時偵測上半身與下半身關鍵資訊，裝置輕量化方便使用者攜帶使用。另本研究商品成本相較現有市價為 7990 元之產品，僅需 1500 元花費，有效減少使用者負擔。

二、未來展望

(一) 姿勢訓練資料集

本研究後續欲邀請更多年齡層、不同身高與體重之若干名男性與女性受測者，於鏡頭前以不同坐姿拍攝，收集更大量與多元之標記後姿勢訓練資料，並規劃將所收集姿勢資料集公開以供後續研究者使用。

(二) 規則個人化

本研究目前透過醫學文件，歸納出適當閥值，以判斷姿勢。後續欲增加應用程式介面功能讓使用者輸入身高、性別等資訊，並透過前述的資料訓練集，以大數據分析歸納出更符合各體型的人之規則。也希望讓使用者能自行微調閥值，讓規則更人性化，避免姿勢設定太嚴格導致使用者放棄使用。本研究的久坐時間為依據文獻提出的建議歸納而成。後續欲在應用程式介面整理久坐相關資訊，以供使用者參考並依自身情形設定多少時間即為久坐。

(三) 增加細部規則

本研究為求裝置輕便，有部分正確姿勢規則無法判斷，如腳跟要與地面垂直以及手腕使用鍵盤的角度。後續欲將 micro:bit 配戴在更多節點之上，或在側邊另外架設鏡頭，收集目前無法取得的資訊，以進行更多姿勢判斷與提醒。

(四) 功能個人化

本研究後續欲增加更符合每個人需求的功能，如讓使用者自行上傳音檔，並以此音檔作為提示音效以增加其使用意願。亦可繪製出更詳細的事後報告，讓使用者更了解自身習慣。

捌、參考文獻資料

[1]衛生福利部國民健康屬. (2021, January 29). 民國 2017 年國民健康訪問調查.https://www.hpa.gov.tw/Pages/ashx/File.ashx?FilePath=~/File/Attach/13636/File_15832.pdf

[2]WHO. (2002, April 4). Physical Inactivity a Leading Cause of Disease and Disability, Warns WHO. World Health Organization. https://reurl.cc/13NzzG

[3]Chernyavskaya, O. (2020, June 27). Fix Posture. Fix Posture.https://reurl.cc/eLKZQK

[4]Huang, E. (2021, July 21). SitApp 改善駝背坐姿軟體，遊戲化提醒電腦前久坐的工作健康.https://reurl.cc/OGKeQy

[5]張雅婷. (2019, September). 以智慧椅墊進行坐姿分析之研究. 台灣師範大學. https://reurl.cc/E4dzqn

[6]BenQ.MindDuo 2 Plus 產品介紹頁面.https://reurl.cc/yY9lg8

[7]陳錦洪(2022 年 4 月 15 日)。AirPods 可改善駝背／坐姿 不再頸痛！Posture Pal 免費 App 設定教學。https://reurl.cc/VNbqLQ

[8]Huang, T. (2018, September 28). 機器/深度學習: 物件偵測 Non-Maximum Suppression (NMS). Medium. https://reurl.cc/krKvkG

[9]Bazarevsky, V., & Grishchenko, I. (2020, June 17). Arxiv. BlazePose: On-Device Real-Time Body Pose Tracking. https://reurl.cc/80xO04

[10]Bazarevsky, V., & Grishchenko, I. (2020, August 13). On-Device, Real-Time Body Pose Tracking with MediaPipe BlazePose. Google Research. https://reurl.cc/VzQpDn

[11]From Perspective Picture to Orthographic Picture [Closed]. (2016, April 12). Stackoverflow. https://reurl.cc/gGpkM4

[12]Bulat, A., & Tzimiropoulos, G. (2016, September 6). Human Pose Estimation via Convolutional Part Heatmap Regression. Arxiv. https://reurl.cc/13NpOD

[13]Huang, Y. (2021, May 12). 運用 ThreeJS 貼圖材質特性，將電影平面海報視覺 3D 化. TPIU. https://reurl.cc/v0R8Ey

[14]Tsagkatos, T. (2019, January 24). Difference between Two Perspective Projection Representations. Game Development. https://reurl.cc/bDKxxd

[15]2022 年台灣網路報告. (2022). 財團法人台灣網路資訊中心. https://reurl.cc/bDKyYv

[16]Laricchia, federica. (2023, May 15). Number of Connected Wearable Devices Worldwide from 2019 to 2022. Statista. https://reurl.cc/2zRkDr

[17]長庚醫療財團法人. (n.d.). 電腦族姿勢之保養.https://reurl.cc/80xOv4

[18]WHO. (2022, October 5). Physical Activity. World Health Organization. https://reurl.cc/13NpvX

[19]衛生福利部國民健康署建議 3C 產品加註警語行政指導原則. (n.d.). 衛生福利部國民健康署

附錄、程式流程圖

以下是主程式和姿勢辨識程式之流程圖：

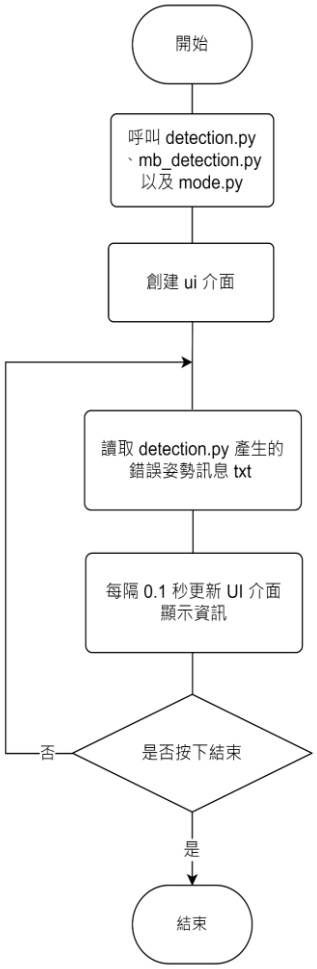


圖 47 main.py 執行流程圖(研究者自繪)

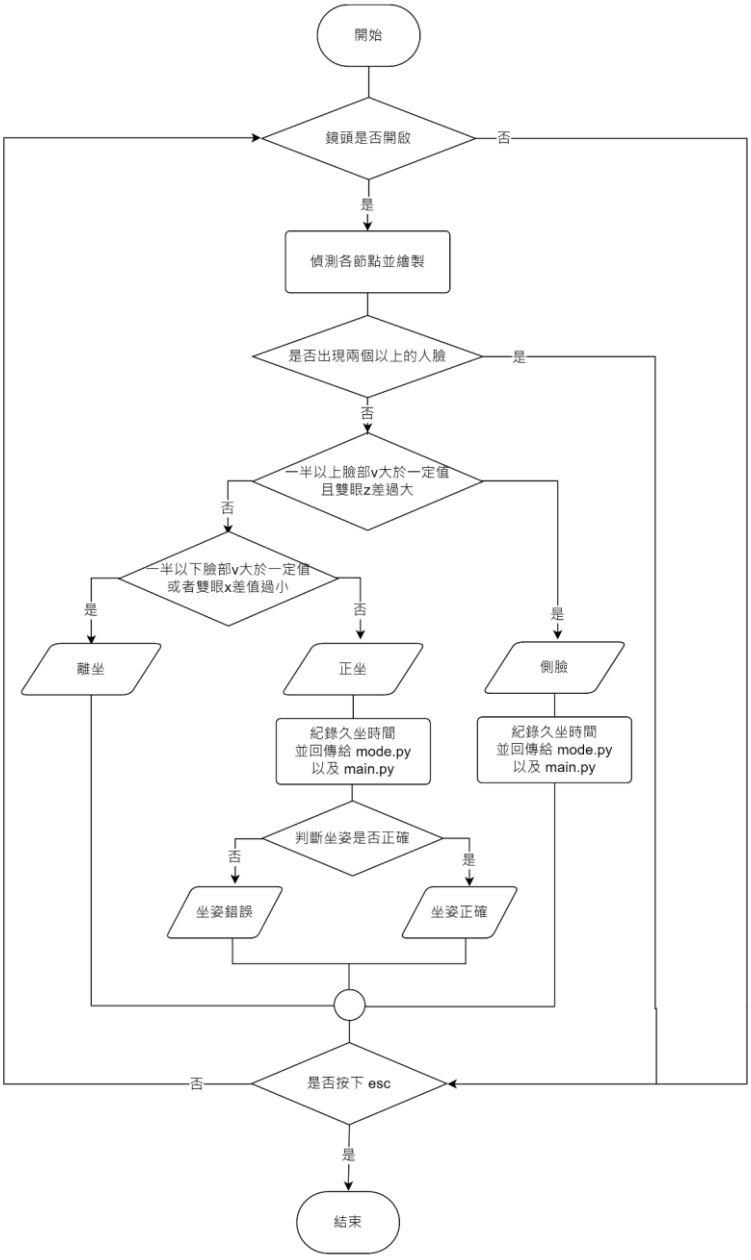


圖 48 detection.py 執行流程圖(研究者自繪)

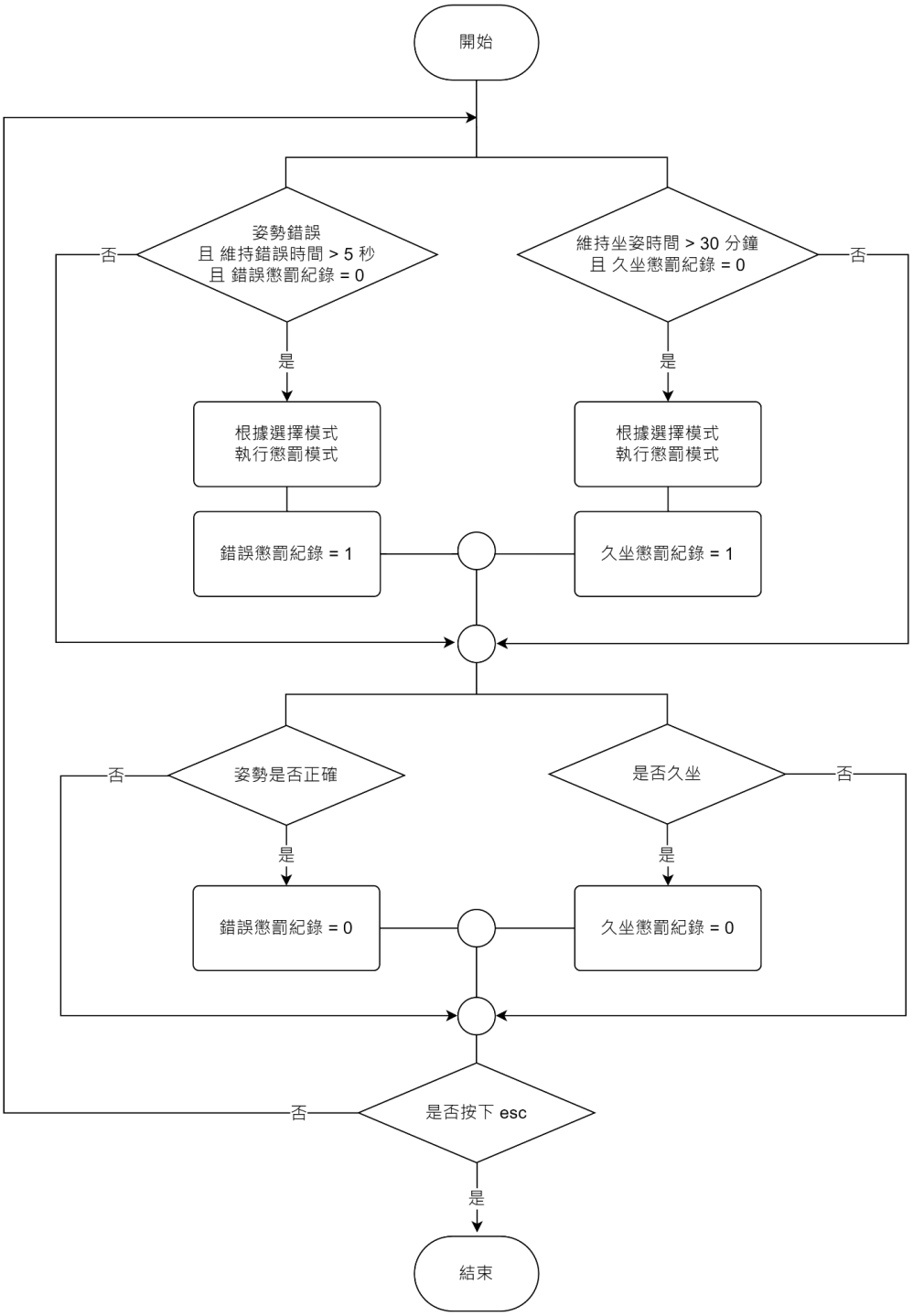


圖 49 mode.py 提醒機制判斷執行流程圖(研究者自繪)

【評語】052511

本作品主題為坐姿偵測系統，判斷使用者是否久坐及使用電腦的姿勢是否正確。研究方法採用電腦內建視訊鏡頭以 MediaPipe BlazePose 影像辨識模型抓取上半身節點資訊，並透過雙腳穿戴 micro:bit 感測裝置取得腳部三維加速度值。但在姿勢上僅利用簡單的 heuristic 進行判斷，過於簡單，缺乏創新性，且結果容易受到環境影響。

作品簡報

基於影像識別與感測裝置之坐姿偵測輔助系統





摘要

本研究旨在開發一套系統，判斷使用者是否久坐與使用電腦的姿勢是否正確，並即時提供回饋。此坐姿輔助系統僅需使用電腦內建視訊鏡頭以MediaPipe BlazePose影像辨識模型抓取上半身節點資訊，並透過雙腳穿戴micro:bit感測裝置取得腳部三維加速度值。研究者蒐集各年齡層不同姿勢下的數據，討論其差異並計算F-score值以訂立閥值，判斷使用電腦的姿勢。同時，研究者運用PyQt套件建立使用者介面，使得系統能夠進行個人化設定，並即時地提供回饋。相較現有研究，本研究所提出之系統更輕量化、方便使用且成本更低。

研究目的

一、結合影像辨識與穿戴式裝置，偵測使用者當下姿勢。

二、辨識使用者是否使用電腦、是否呈現坐姿及姿勢是否正確。

三、蒐集資料集，設定不同姿勢之閥值，並探討其準確率。

四、給予使用者即時的回饋，督促使用者改善坐姿或警示使用者離坐走動。

五、設計使用者介面，使其方便使用、貼合使用者需求並提供數據統計。

六、增加此系統的實用性，使其方便攜帶及降低成本。

研究動機

根據國民健康署的調查，臺人每日平均久坐高達5.9小時。許多民眾因久坐或以不良姿勢使用電腦，導致諸多生理問題，如肩頸及背部疼痛。世界衛生組織指出全球約有60%到85%的人口有久坐的生活型態，此型態被列為十大死亡及患病主因。研究者希望從源頭解決此些問題，研發一套輔助系統。探討現有研究與產品，發現其存在價格過高、功能受限或使用不便等問題，因此將研發一套更輕量化、方便使用且成本更低的系統。

研究流程及方法



圖片來源：研究者自繪

研究結果

影像識別：MediaPipe BlazePose

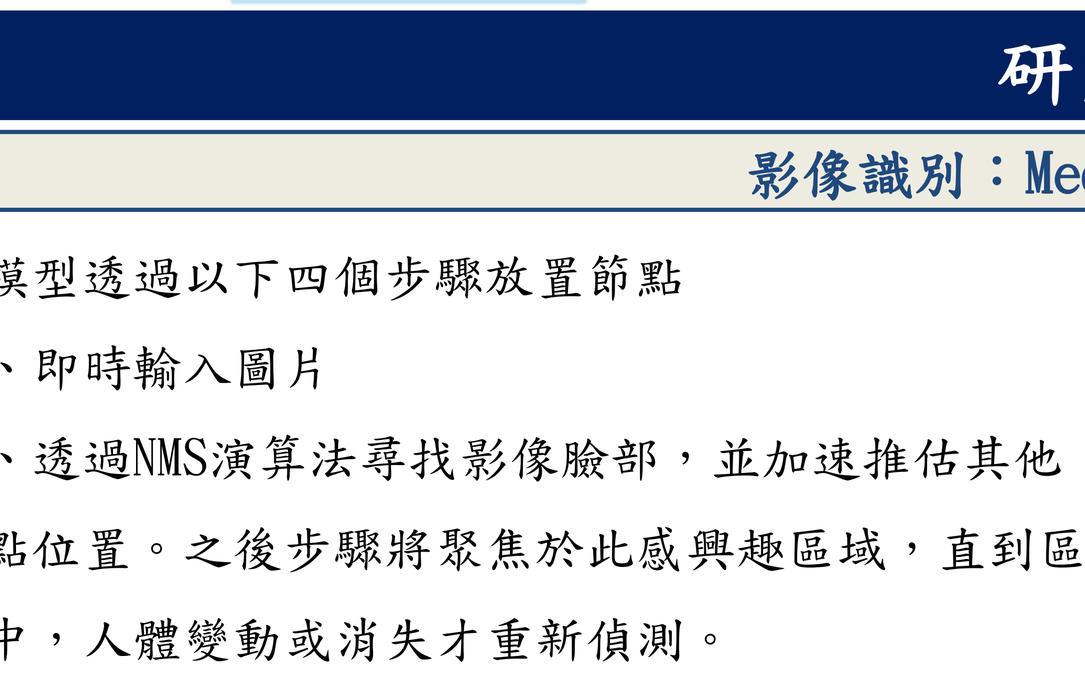
此模型透過以下四個步驟放置節點

一、即時輸入圖片

二、透過NMS演算法尋找影像臉部，並加速推估其他節點位置。之後步驟將聚焦於此感興趣區域，直到區域中，人體變動或消失才重新偵測。

圖片來源：擷取自參考資料[1]

圖片來源：擷取自參考資料[1]



圖片來源：擷取自參考資料[2]

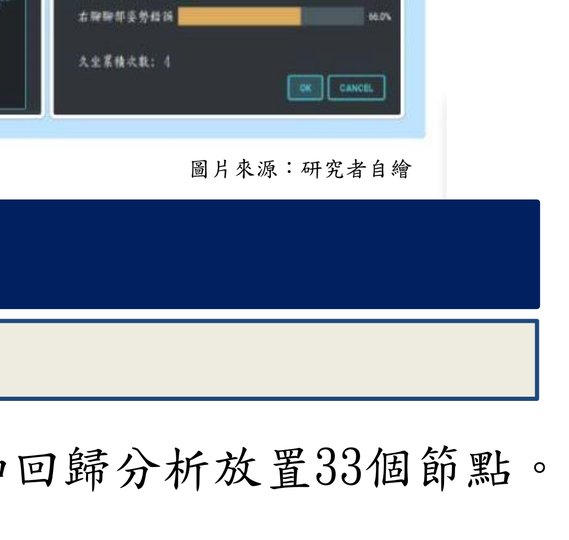
三、透過熱圖、偏移量分析圖和回歸分析放置33個節點。

0. nose
1. right eye inner
2. right eye
3. right eye outer
4. left eye inner
5. left eye
6. left eye outer
7. right ear
8. left ear
9. mouth right
10. mouth left
11. right shoulder
12. left shoulder
13. right elbow
14. left elbow
15. right wrist
16. left wrist
17. right pinky knuckle #1
18. left pinky knuckle #1
19. right index knuclke #1
20. left index knuckle #1
21. right thumb knuckle #2
22. left thumb knuckle #2
23. right hip
24. left hip
25. right knee
26. left knee
27. right ankle
28. left ankle
29. right heel
30. left heel
31. right foot index
32. left foot index

圖片來源：擷取自參考資料[1]



圖片來源：擷取自參考資料[1]



圖片來源：擷取自參考資料[3]

四、輸出各骨架節點x, y, z以及visibility數值。x, y, z皆以身體臀部為原點，其中z值是透過x, y值與弱透視投影推估。visibility表示此節點可視程度。

穿戴式感測裝置：micro:bit

使用micro:bit取得雙腳三維加速度。此裝置是每片約500-600元的微型電腦開發板，小巧且操作簡易，內部配置的加速度感測器可得知目標物三維加速度值。

判別是否在使用電腦

透過影像辨識技術以程式判斷：

一、離坐：臉部半數節點visibility小於一定值或雙眼x差小於一定值。

二、頭部非面對螢幕：雙眼 z 值差大於一定值。

定義久坐與判別是否呈現坐姿

世界衛生組織表示不論任何年齡層，皆應避免久坐。本研究參考各醫療機構資料，定義30分鐘即為久坐。當未被判斷為離座，則判斷使用者呈現坐姿。

判別姿勢正確與否

依據生物學文獻資料定義正確姿勢，以程式判斷：

一、頭部水平傾斜：雙眼y值差大於一定值

二、肩膀水平傾斜：雙肩y值差大於一定值

三、駝背：左右側眼睛與肩膀z值差大於一定值

四、身體側斜：雙肩中點與雙眼中點x值差大於一定值

五、臉距螢幕過近：雙眼x值差大於一定值且雙眼平均z值小於一定值

六、左/右腳姿勢錯誤：x或z加速度值與初始值差值大於一定值

資料集收集與評估指標

研究者邀請若干名年齡層介在11到76歲的受測者。去除前後段干擾值與異常值後，共收集107687筆數據，研究者將其以8: 2之比例分為訓練集與測試集。針對訓練集以F1-score定義判斷不同姿勢之閥值，並以F2-score定義較不嚴格模式的閥值，其公式如下：

	實際為正樣本	實際為負樣本
預測為正樣本	$TP = True\ Positive$	$FP = False\ Positive$
預測為負樣本	$FN = False\ Negative$	$TN = True\ Negative$

$$Precision = \frac{TP}{TP+FP},\ Recall = \frac{TP}{TP+FN}$$

$$F-score = \frac{(1+\beta^2)\ Precision \times Recall}{\beta^2\ Precision + Recall}$$

本研究針對測試集，以準確率(Accuracy)判斷此系統之精確程度，其公式如下：

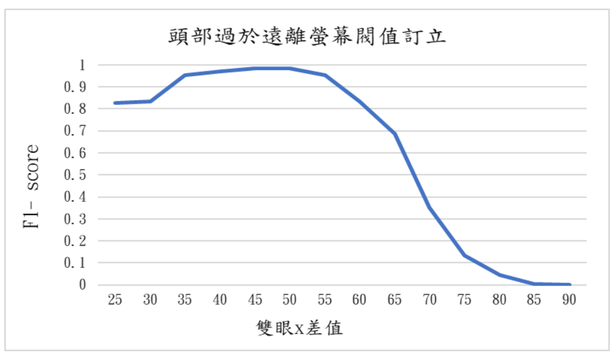
$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

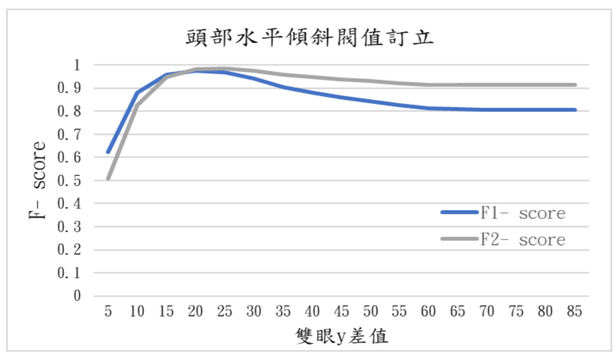
上述公式皆由研究者自製

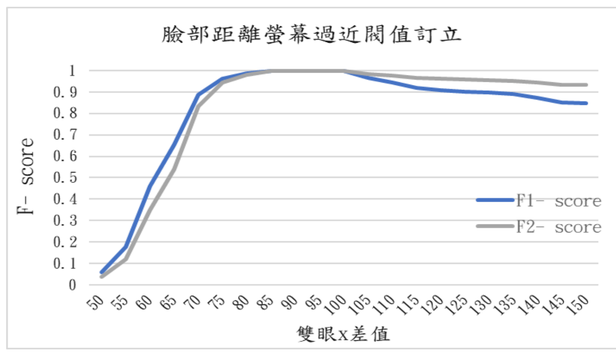
姿勢閥值訂立

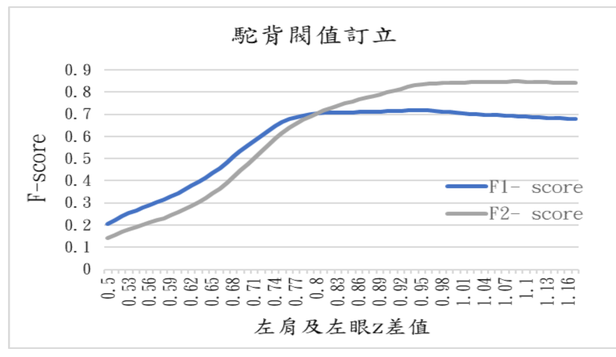
透過以上規則，針對訓練集判斷不同閥值下的F-score，並取最高分數對應的閥值做為系統的判斷依據。

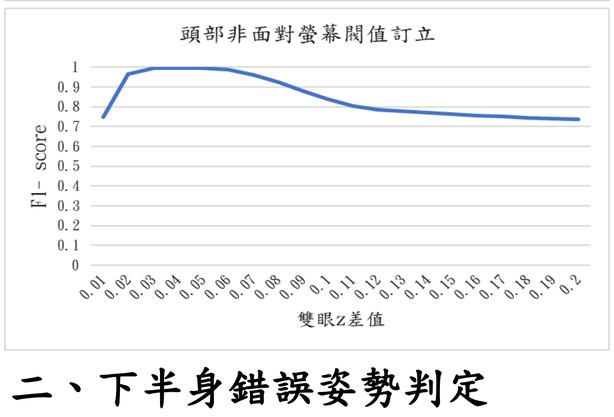
一、判斷是否維持坐姿

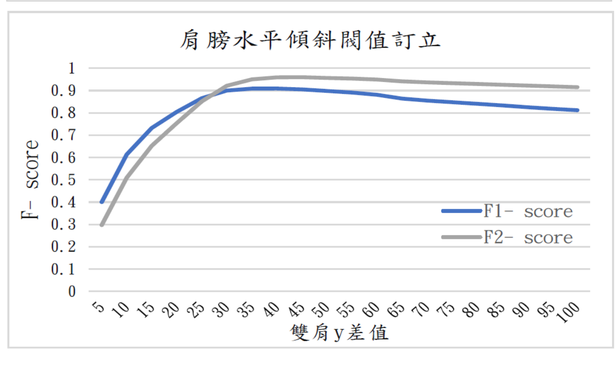


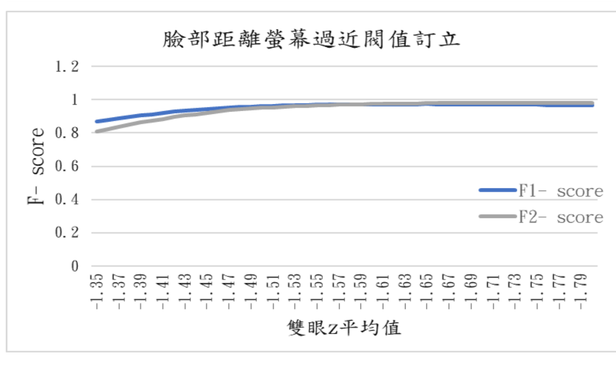


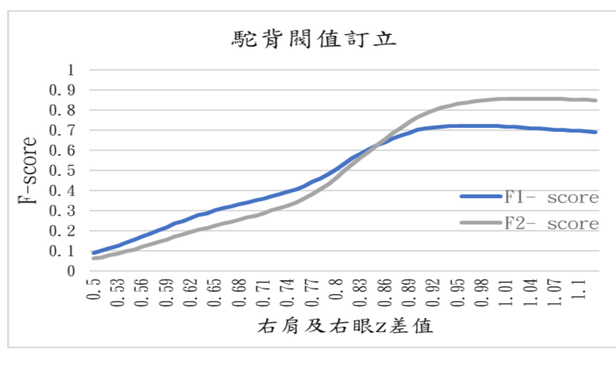




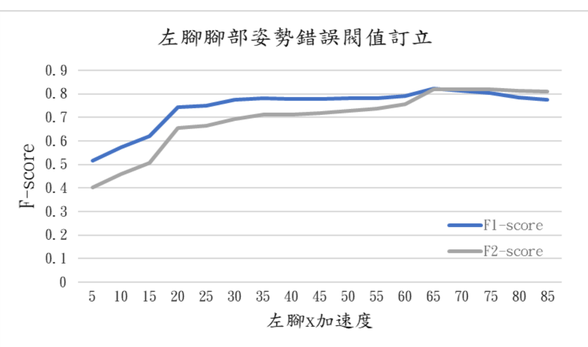


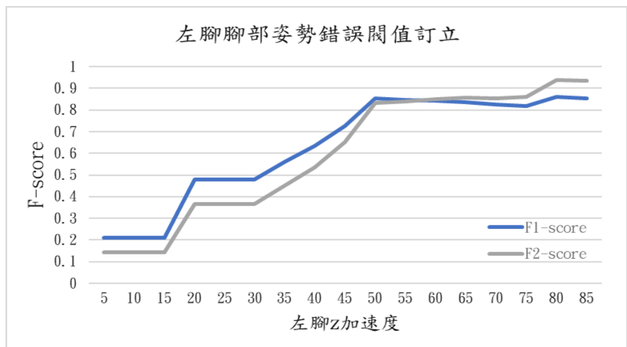


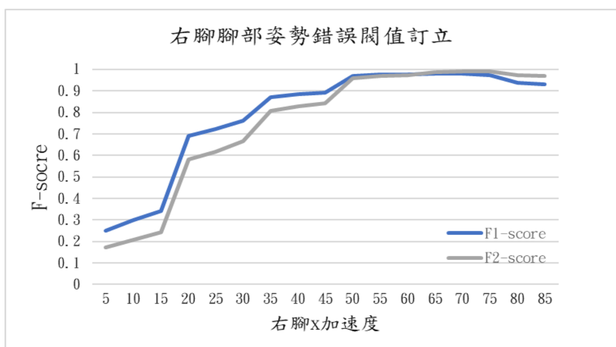


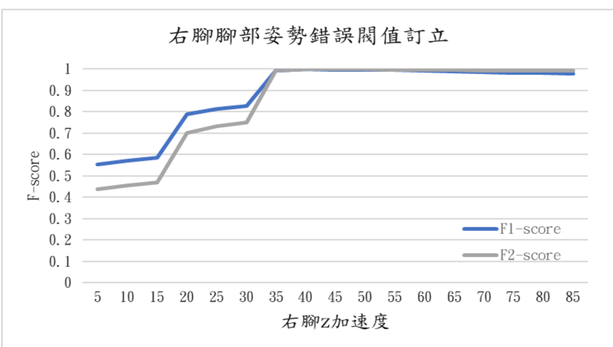


二、上半身錯誤姿勢判定





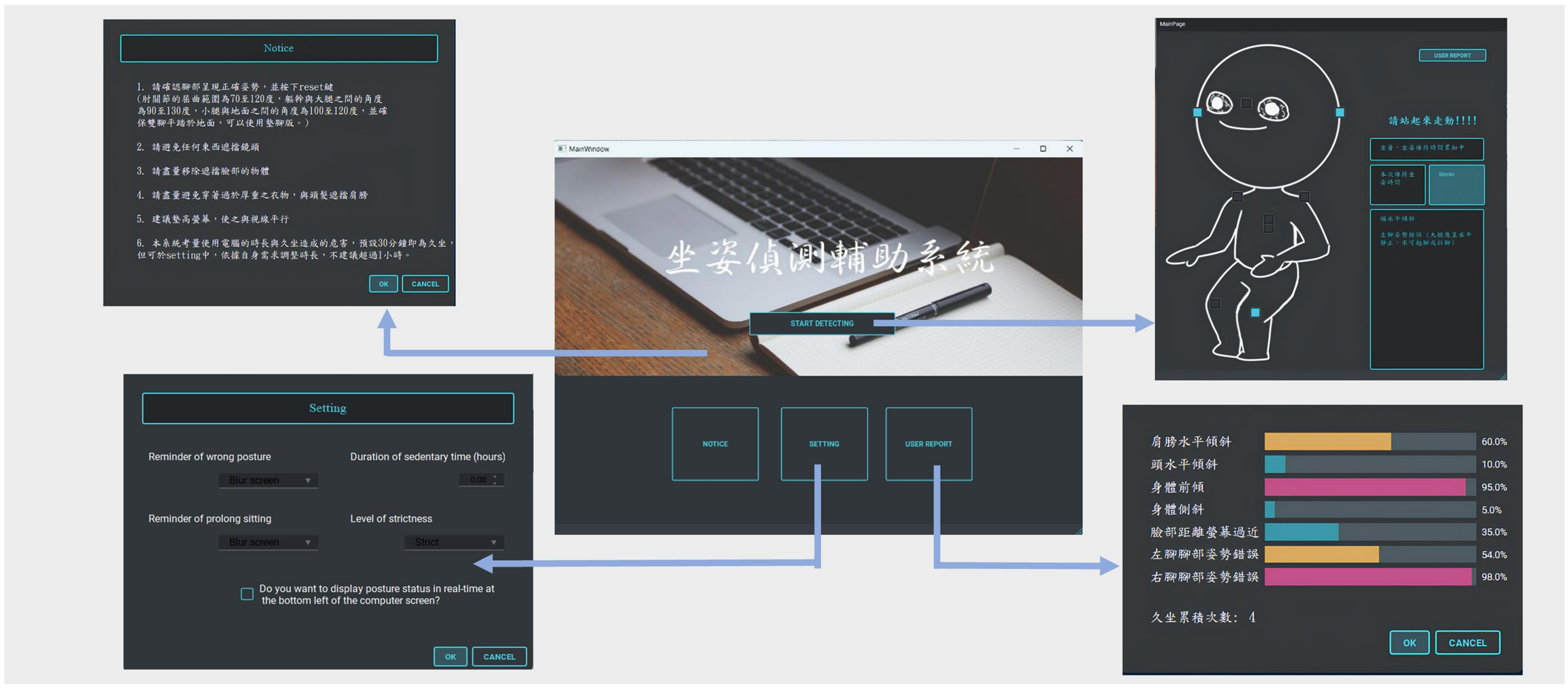




上述圖片皆由研究者自繪

使用者互動與回饋介面

透過PyQt製作操作簡易使用者介面，以提供即時回饋，並使此系統更符合使用者的需求，增加其使用意願。



圖片來源：研究者自繪

討論

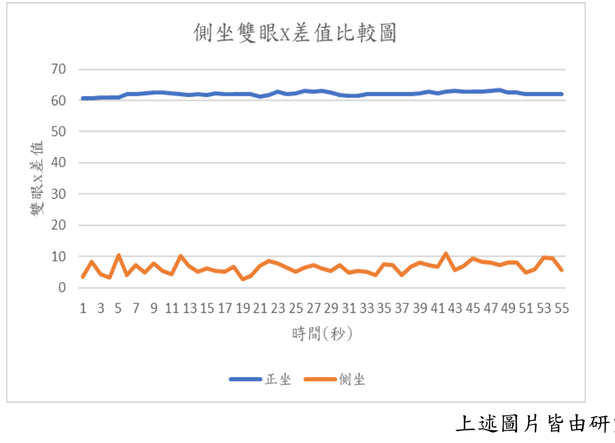
姿勢判斷準確率分析

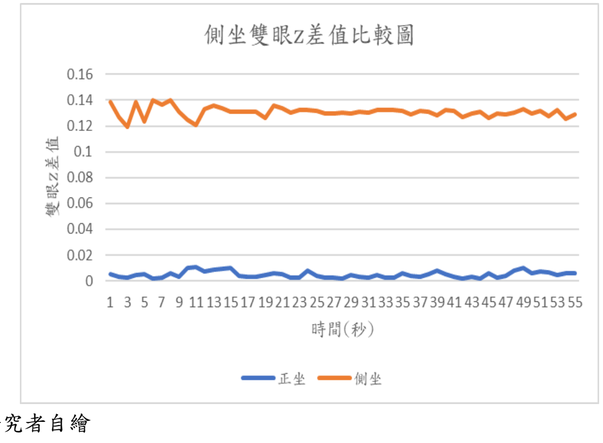
針對測試集，以accuracy討論各姿勢判斷的準確率。此系統於判斷離座、頭部非面對螢幕、頭部水平傾斜、臉部距離螢幕過近效果極佳。而判斷駝背與身體側斜時效果不如預期，研究者推測，駝背對於眼睛和肩膀的深度差異小，易造成駝背誤判；不同受測者傾斜程度差異較大，導致傾斜幅度較小的姿勢被誤判。

特殊情形討論

一、使用者頭部未面對鏡頭，誤判與螢幕的距離

以側臉面對鏡頭時，雙眼x差值極小。本系統將在判定離坐前，先以雙眼z差值判斷頭部是否面對鏡頭。

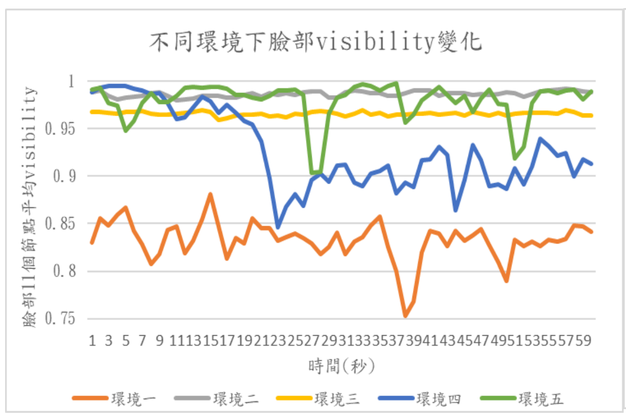


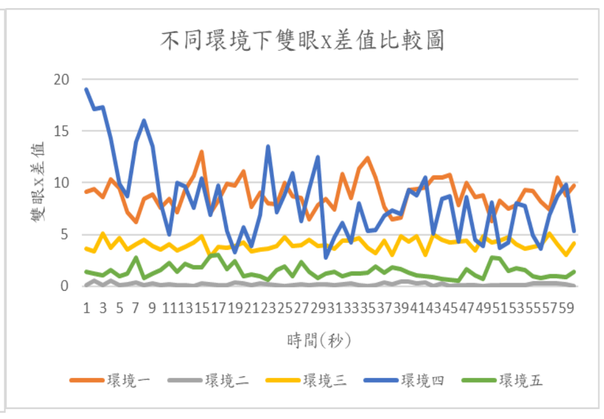


上述圖片皆由研究者自繪

二、使用者離坐期間，模型誤抓取環境物件

背景中有較不具稜角的輪廓時，模型較容易偵測到此物體且visibility值較高。背景抓取到的物件不具五官，節點較集中，加入雙眼x差值判斷是否離坐。





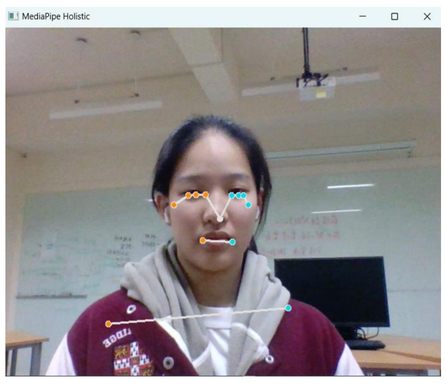


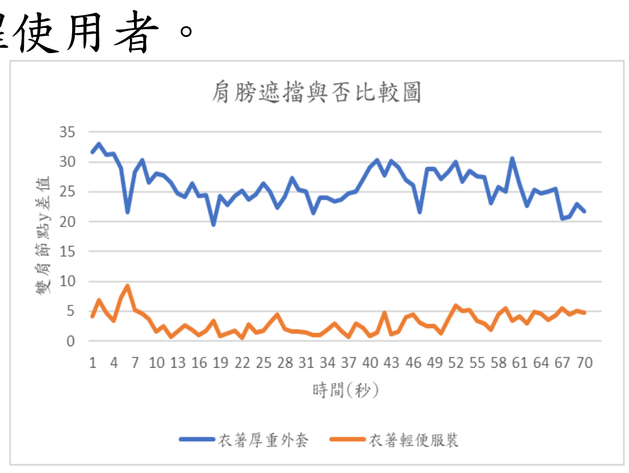
環境一　環境二　環境三　環境四　環境五

上述圖片皆由研究者自繪、自製

三、雙肩的節點遮擋，誤判肩膀水平傾斜

穿著厚重衣物或頭髮遮擋肩膀，依據模型訓練時的遮蔽模擬放置節點。雙肩節點產生偏移，加入特殊數據訂立閥值，並以介面提醒使用者。

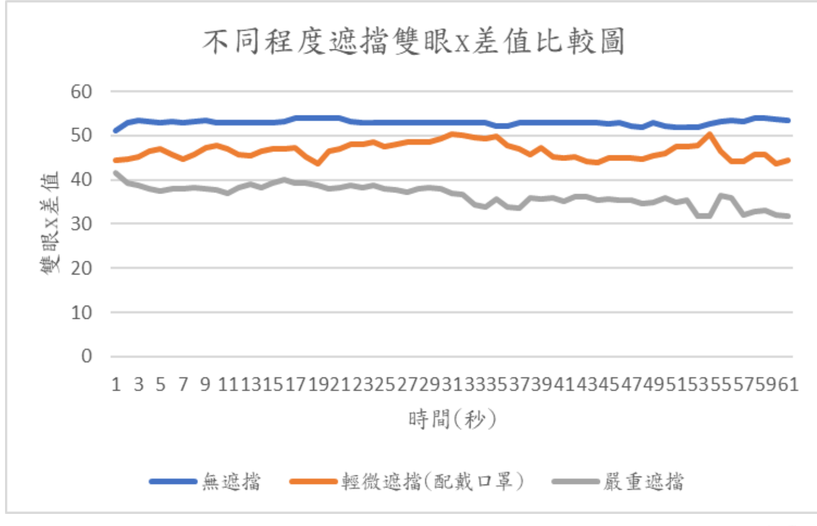




上述圖片皆由研究者自繪、自製

四、臉部被遮擋，誤判為離坐

臉部一開始即被遮擋與臉部於中途被遮擋，雙眼X差值小。加入特殊數據訂立閥值，並以介面提醒。



上述圖片皆由研究者自繪、自製

五、有兩個距離相近的臉部，誤抓臉部節點

原先追蹤的臉部區域產生變動，臉部節點偏移。加入FaceMesh模型排除其餘人臉之干擾。



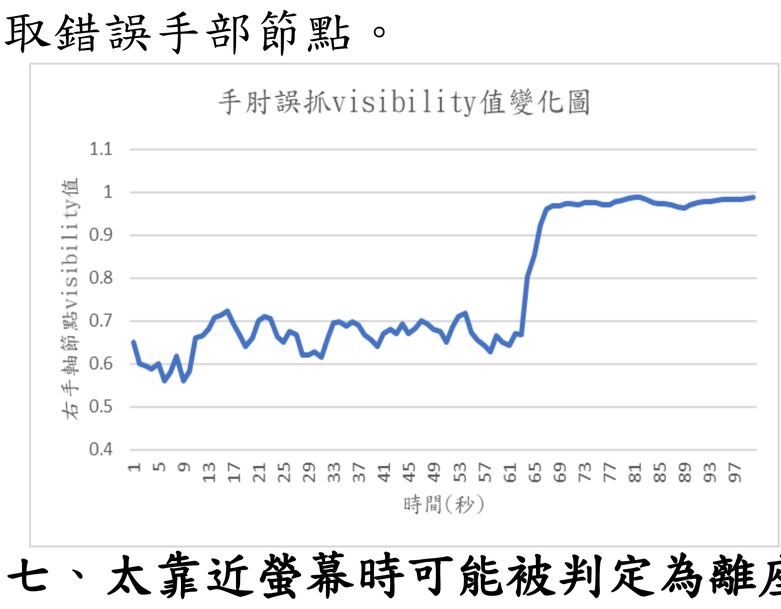
圖片來源：研究者自製



圖片來源：研究者自製

六、鏡頭前出現第二個人，誤抓手部節點

第二個人手出現在使用者合理的手部位置時，模型抓取錯誤手部節點。







上述圖片皆由研究者自繪、自製

七、太靠近螢幕時可能被判定為離座

八、肩膀位於鏡頭抓取到的邊界

九、離座後，重回位置瞬間節點不正確

十、micro:bit 資料傳輸數量錯誤

訊號干擾導致資訊數量錯誤，系統將對不完全之下半身感測資料予以排除。

結論

總結

一、姿勢訓練資料集：本研究蒐集11歲到76歲不同年齡受測者不同坐姿數據，共107687筆有效資料。並將其分為訓練集與測試集，分別計算F-score與準確率。

二、姿勢閥值訂立：

姿勢類別	規則判斷與閥值	此閥值之F1-score	此閥值之準確率
離座	雙眼x差<47 或 一半以上 visibility<0.97	0.989686	90.57%
頭部非面對螢幕	雙眼z差>0.04	0.996593	94.13%
頭部水平傾斜	雙眼y差>22	0.977122	96.22%
肩膀水平傾斜	雙肩y差>40	0.908894	81.37%
駝背	右眼與右肩z差>0.97 或 左眼與左肩z差>0.94	0.722575 / 0.719595	63.01%
身體側斜	雙肩中點與 雙眼中點x差>45	0.914226	70.02%
臉部距離 螢幕過近	雙眼x差>90且 雙眼z平均< -1.65	0.991615 / 0.973653	96.50%
右腳姿勢錯誤	右腳x加速度差>50且 右腳z加速度差>50	0.987540 / 0.996811	87.33%
左腳姿勢錯誤	左腳x加速度差>64且 左腳z加速度差>80	0.920368 / 0.862139	86.99%

圖片來源：研究者自繪

三、誤判情形討論：非面對鏡頭易誤判離坐，增以雙眼z差值判別。離坐易誤抓環境物件，增以雙眼x差值判別。節點被遮蔽如穿著厚重衣物或配戴口罩易誤判，增以遮擋後數據並於可能誤判時提醒。

四、互動介面實作：開發互動介面以選擇互動方式，並於姿勢錯誤或久坐時回饋，亦提供錯誤姿勢分析。

五、本研究相較現有研究：不需讓使用者擺出正確姿勢，以內建鏡頭與少量開發板即可偵測全身資訊，輕量且相較現有產品，僅需不到1/5花費，有效減少使用者負擔。

六、本研究對社會的助益：透過此操作簡易且成本較低的系統，協助長時間使用電腦的族群改善其錯誤姿勢，並避免久坐，從源頭減少此些行為導致的諸多生理問題。

未來展望

一、增加並公開姿勢訓練資料集

二、分析數據使判別更個人化，且增加客製功能

三、增加感測裝置與鏡頭，以取得更細部資訊

參考文獻

[1]Bazarevsky, V., & Grishchenko, I. (2020, June 17). Arxiv. BlazePose: On-Device Real-Time Body Pose Tracking. https://reurl.cc/80xO04

[2]Huang, T. (2018, September 28). 機器/深度學習: 物件偵測 Non-Maximum Suppression (NMS). Medium. https://reurl.cc/krKvkG

[3]Tsagkatos, T. (2019, January 24). Difference between Two Perspective Projection Representations. Game Development. https://reurl.cc/bDKxxd

[4]WHO. (2002, April 4). Physical Inactivity a Leading Cause of Disease and Disability, Warns WHO. World Health Organization. https://reurl.cc/13NzzG

[5]Chernyavskaya, O. (2020, June 27). Fix Posture. Fix Posture. https://reurl.cc/eLKZQK

[6]Huang, E. (2021, July 21). SitApp 改善駝背坐姿軟體，遊戲化提醒電腦前久坐的工作健康.https://reurl.cc/OGKeQv

[7]陳錦洪(2022年4月15日)。AirPods可改善駝背／坐姿 不再頸痛！Posture Pal免費App設定教學。https://reurl.cc/VNbqLQ

[8]WHO. (2022, October 5). Physical Activity. World Health Organization. https://reurl.cc/13NpvX