

# 結合 Mediapipe Pose 和支持向量機的坐姿辨識系統

陳楷傑<sup>1</sup>, 蕭如淵<sup>2</sup>

<sup>1</sup> 國立彰化師範大學資訊工程系

E-mail: kai20010209@gmail.com

<sup>2</sup> 國立彰化師範大學資訊工程系

E-mail: hsiao jy@cc.ncue.edu.tw

## 摘要

本研究開發一種具實用性的坐姿辨識系統，以改善不良坐姿對健康可能造成的問題。我們採用了開源計算機視覺庫二版 (OpenCV2) 來擷取外部鏡頭影像，並利用谷歌開發的 Mediapipe Pose 套件來提取上半身的特徵點位置，例如眼睛、鼻頭，再將些數據投入支援向量機 (SVM) 進行模型訓練，使系統能夠辨識三種不同坐姿，實驗結果準確率高達 99.90%。

當系統運作時，如果偵測到不良坐姿持續一段時間，系統將發出警報，提醒使用者調整姿勢以維護健康。這項研究結合了計算機視覺、機器學習的技術，旨在提供一個實用的坐姿辨識系統，我們期望該系統在實際應用中能夠為使用者提供有效的坐姿監控，進而改善長時間坐姿對身體的潛在危害。

關鍵詞: 坐姿辨識, Mediapipe Pose, 支持向量機

## Abstract

This study develops a practical seated posture recognition system to improve potential health issues caused by poor sitting posture. We utilize the open-source computer vision library (OpenCV2) to capture external camera images and leverage the Mediapipe Pose package developed by Google to extract key points of the upper body, such as eyes and nose. These data are then fed into a Support Vector Machine (SVM) for model training, enabling the system to recognize three different sitting postures with a high accuracy rate of 99.90%.

During system operation, if a poor sitting posture is detected for a prolonged period, the system will issue an alert to remind the user to adjust their posture to maintain their health.

This research combines computer vision and machine

learning techniques to provide a practical posture recognition system. We anticipate that this system will effectively monitor sitting posture in practical applications, thereby mitigating potential health risks associated with prolonged sitting.

**Keywords:** posture recognition, Mediapipe Pose, Support Vector Machine

## 1. 引言

現代生活中，長時間使用電腦和手機成為人們不可避免的一部分。然而，不正確的坐姿在長時間使用這些電子設備時可能導致健康問題。儘管有關坐姿的健康風險已經引起廣泛關注，但很少有系統能夠實時監測和提醒使用者，以維護良好的坐姿。本研究旨在開發一個基於計算機視覺和機器學習的坐姿辨識系統，該系統能夠透過外接鏡頭實時檢測上半身的坐姿，提供即時的警告，以幫助使用者維持良好的坐姿習慣。

我們使用 OpenCV2 和 Mediapipe Pose 進行鏡頭影像的擷取和特徵提取，並將這些特徵投入 SVM 進行模型訓練。透過 Python 的 Pygame 套件，我們開發了一個直觀的使用者介面，向使用者發出及時的警報。本系統具有應用於辦公室、學習環境和家庭等場景的潛在價值，有望改善人們在數位時代長時間坐姿不當的問題。

在接下來的章節中，我們將詳細描述所開發的系統、實驗結果、討論優勢和限制，以及未來的改進方向。透過本研究，我們期望提高對坐姿健康的警覺性，同時為未來相關研究提供一個有價值的參考。

## 2. 文獻探討

關於坐姿辨識的研究大致可分為三種方向，基於壓力感測器的系統[2][4][5][6]、基於加速度計的系統[1]與基於視覺辨識的系統[3][8]，基於壓力感測器的系統將感測器設置在椅子上來辨識坐姿，此種系統的缺點是需要設置多個壓力感測器，且壓力感測器損壞後不易更換，基於加速度計的系統將可穿戴慣性量測單元(IMU)連接到使用者身上以感知運動，並將數據透過傳輸線送回電腦來判斷坐姿，此種系統的缺點是要讓使用者連接感測器，在實際應用上較為不便，基於視覺辨識的系統則通過深度相機取得坐姿影像進行辨識，我們的研究是基於視覺辨識的系統，使用一般相機來取得影像，與深度相機相比在實際應用上更為方便，利用谷歌開發的 Mediapipe Pose 套件來取得影像中上半身的特徵點數據，並使用 SVM 訓練坐姿辨識的模型，以完成能即時回應且高準確率、高便利性的坐姿辨識系統。

### 3. 我們的系統

#### 3.1. 系統概述

本研究開發了一個坐姿辨識的系統，經由外接鏡頭取得人體上半身的正面影像，系統會透過影像即時判斷出坐姿，並在不良坐姿持續一段時間後發出警訊。

本研究使用特徵點(例如眼睛、鼻頭、肩膀)間的距離而不是整個圖像作為訓練資料，此方法主要有三個優勢

- 一. 維度較低：相對於整個圖像，特徵距離通常具有更低的維度。這可以減少模型的複雜性，提高訓練和預測的效率。
- 二. 降低雜訊：特徵距離捕捉了坐姿的特定特徵，有助於降低噪聲和不相關信息的影響，這使得模型更專注於與坐姿相關的重要特徵。
- 三. 穩健性提高：特徵距離的正規化有助於提高模型對不同尺寸和比例的影像的穩健性，從而增強了模型的泛化性能。

#### 3.2. 硬體設備

本研究所使用的硬體設備包括一台搭載第 11 代

Intel(R) Core(TM) i7-1165G7 處理器的筆記型電腦。記憶體配置為 16GB RAM，用於系統的模型訓練。使用的鏡頭型號為 USB2.0 HD UVC WebCam。

#### 3.3. 數據集蒐集

本研究根據參考文獻[7]，將坐姿分類為直立(圖1)、俯臥(圖2)與扶下巴或前傾(圖3)三種。測試者首先被要求坐在椅子上面對筆記型電腦的鏡頭且眼睛距離螢幕70公分處，之後做出直立、俯臥與扶下巴或前傾三種坐姿，本研究將三種姿勢分別蒐集300多張影像共1006張影像，如表1所示。將蒐集的影像利用Mediapipe捕捉上半身的特徵節點。這些節點包括肩膀、臉部等關鍵點。再透過歐基里德距離公式(1)計算這些節點之間的相關距離以建立數據集，為了正規化數據，每組相關距離都除以鼻子的長度，即雙眼中間至鼻頭的距離。每張影像皆會透過公式轉為25個特徵距離。

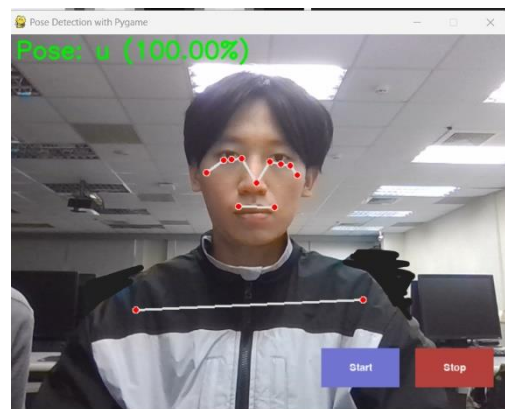


圖 1 直立

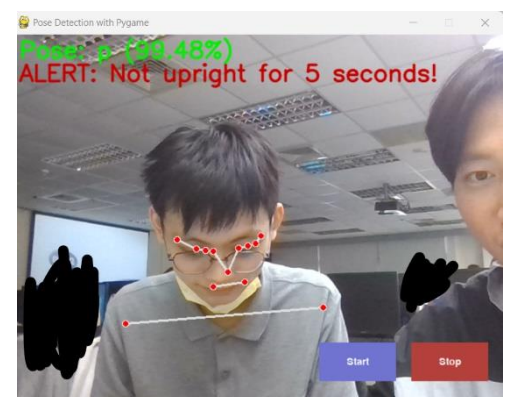


圖 2 俯臥

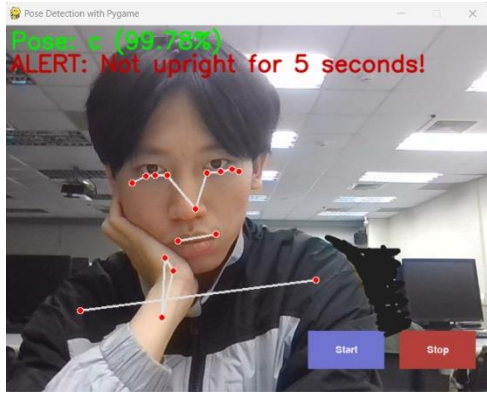


圖 3 扶下巴或前傾

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2} \quad (1)$$

表 1 坐姿種類

坐姿種類	數量
直立	331
俯臥	371
扶下巴或前傾	303

### 3.4. 影像正規化

為了確保模型對不同尺寸的影像具有穩健性，將數據集中的相關距離進行正規化。這是透過將每個相關距離除以影像中鼻子的長度實現的，確保特徵的一致性，首先先算出雙眼中央座標，假設左眼的座標為  $(x_1, y_1)$ ，右眼的座標為  $(x_2, y_2)$ ，則雙眼中間座標  $(x_e, y_e)$  可以用以下公式(2)表示

$$x_e = (x_1 + x_2) \div 2 \quad (2)$$

$$y_e = (y_1 + y_2) \div 2$$

假設鼻頭的座標為  $(x_n, y_n)$ ，雙眼中央的座標為  $(x_e, y_e)$ ，則鼻子長度  $(d_n)$  可用歐基里德距離公式(3)表示

$$d_n = \sqrt{(x_n - x_e)^2 + (y_n - y_e)^2} \quad (3)$$

### 3.5. SVM 模型訓練

本研究收集了三種姿勢的特徵點距離，每種姿勢各約有 300 張影像。這些數據經過正規化後，被用於訓練模型。經過實驗，模型的核函數使用線

性核函數(linear kernel)，這種核函數有計算速度快的優點，此外，為了能進行三種以上類別的分類，我們將決策函數結構(decision\_function\_shape)設為一對其他 (One-vs-Rest)，在使用十倍交叉驗證後，模型準確率為99.90%。

### 3.6. 系統呈現

本研究統利用了OpenCV2來取得外部鏡頭的即時影像，並提取影像中上半身的特徵節點計算其相關距離，再利用預先訓練好的SVM模型來判斷當前的坐姿類型。隨後，利用Mediapipe Pose和Pygame將坐姿判斷結果呈現在畫面上。展示方式包括標註特徵點、顯示坐姿類別標籤，並在持續一段時間出現不良坐姿時發出警報。

### 3.7. 系統運作流程圖

圖4為本系統運作的流程圖，當系統開始執行時會透過鏡頭取得即時影像，待使用者按下開始(start)按鈕後即持續判斷坐姿，按下停止(stop)按鈕後則停止判斷。系統運行時會透過筆電鏡頭或是外接鏡頭取得坐姿正面影像，若坐姿非直立坐姿超過5秒中，則會以文字及聲音警訊通知使用者。

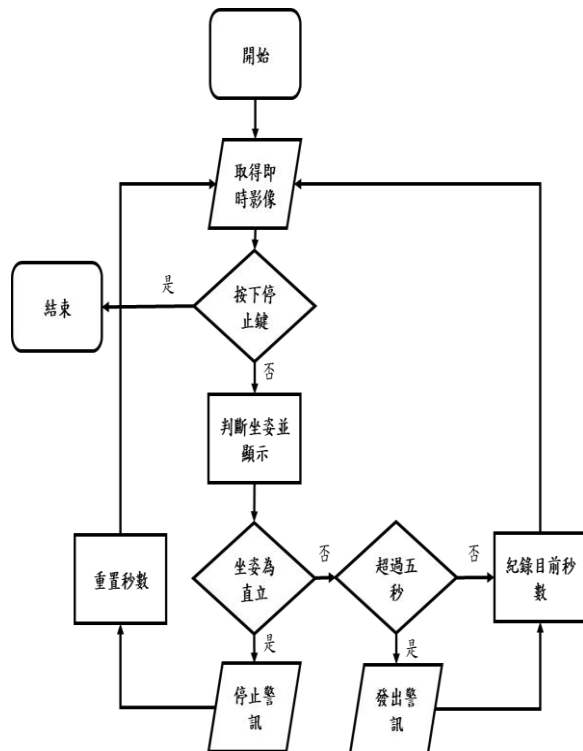


圖 4 系統運作流程圖



## 4. 結果

本研究之SVM模型以python3.9.7與scikit-learn1.3.2來開發，我們使用十倍交叉驗證來評估該模型的準確性，數據集被隨機分為十個相等的子集，並且做十輪訓練，每輪訓練將其中九個子集作為訓練集，剩下一個子集當作驗證集，且十輪的驗證集皆不同，最後將十輪訓練的準確率取算術平均數作為模型準確率，經過實驗結果模型準確率為99.90%。

表2比較了本系統與其他系統的優勢。首先，本系統只需要一個普通的相機鏡頭作為感測器，這使得實際應用的門檻較低，即使設備損壞也容易更換。在模型方面，我們的輸入是特徵點距離而非整張影像，這不僅減少了運算量，還提高了準確率。此外，本究使用了谷歌推出的 Mediapipe Pose 來獲取特徵點位置。Mediapipe Pose 具備高幀率 and 低延遲的特性，使系統能夠實現實時回應。在所比較的系統中，我們的系統也具有最高的準確率，充分展示了其優越性。

## 5. 結論

本研究成功開發了一個基於計算機視覺和機器學習的坐姿辨識系統，利用 OpenCV2 和 Mediapipe Pose 進行影像擷取和特徵提取，並通過 SVM 進行模型訓練。實驗結果顯示，模型在辨識三種不同坐姿具有高達 99.90% 的準確率。

本系統能夠即時監控使用者的坐姿，並在不良坐姿持續一段時間後發出警報，提醒使用者調整姿勢，以維護健康。這一功能對於長時間使用電腦和手機的人群尤其重要，能夠有效預防因不良坐姿引起的各種健康問題。

透過本研究，我們展示了計算機視覺和機器學習技術在健康監控領域的潛力。本系統具有應用於辦公室、學習環境和家庭等場景的廣泛潛力，能夠幫助使用者養成良好的坐姿習慣，提升生活質量。

未來的研究可以進一步優化系統，例如引入更多的坐姿類型進行辨識，操作介面也可進一步改

善，以提供更直觀和友好的使用體驗。我們期望這項研究能夠促進關於坐姿辨識技術的發展，並實現更廣泛的應用。

## 6. 參考文獻

- [1] S. Ma, W.-H. Cho, C.-H. Quan, and S. Lee, "A sitting posture recognition system based on 3 axis accelerometer," in *2016 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, Oct. 2016, pp. 1–3. doi: [10.1109/CIBCB.2016.7758131](https://doi.org/10.1109/CIBCB.2016.7758131).
- [2] Q. Hu, X. Tang, and W. Tang, "A Smart Chair Sitting Posture Recognition System Using Flex Sensors and FPGA Implemented Artificial Neural Network," *IEEE Sensors Journal*, vol. 20, no. 14, pp. 8007–8016, Jul. 2020, doi: [10.1109/JSEN.2020.2980207](https://doi.org/10.1109/JSEN.2020.2980207).
- [3] A. Kulikajevs, R. Maskeliunas, and R. Damaševičius, "Detection of sitting posture using hierarchical image composition and deep learning," *PeerJ Comput. Sci.*, vol. 7, p. e442, Mar. 2021, doi: [10.7717/peerj-cs.442](https://doi.org/10.7717/peerj-cs.442).
- [4] H. Jeong and W. Park, "Developing and Evaluating a Mixed Sensor Smart Chair System for Real-Time Posture Classification: Combining Pressure and Distance Sensors," *IEEE J. Biomed. Health Inform.*, vol. 25, no. 5, pp. 1805–1813, May 2021, doi: [10.1109/JBHI.2020.3030096](https://doi.org/10.1109/JBHI.2020.3030096).
- [5] J. Wang, B. Hafidh, H. Dong, and A. E. Saddik, "Sitting Posture Recognition Using a Spiking Neural Network," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 1779–1786, Jan. 2021, doi: [10.1109/JSEN.2020.3016611](https://doi.org/10.1109/JSEN.2020.3016611).
- [6] Z. Fan, X. Hu, W.-M. Chen, D.-W. Zhang, and X. Ma, "A deep learning based 2-dimensional hip pressure signals analysis method for sitting posture recognition," *Biomedical Signal Processing and Control*, vol. 73, p. 103432, Mar. 2022, doi: [10.1016/j.bspc.2021.103432](https://doi.org/10.1016/j.bspc.2021.103432).
- [7] S. Jiao, Y. Xiao, X. Wu, Y. Liang, Y. Liang, and Y. Zhou, "LMSPNet: Improved Lightweight Network for Multi-Person Sitting Posture Recognition," in *2023 IEEE 3rd International Conference on Computer Communication and Artificial Intelligence (CCAI)*, May 2023, pp. 289–295. doi: [10.1109/CCAI57533.2023.10201258](https://doi.org/10.1109/CCAI57533.2023.10201258).
- [8] B.-S. Lin, K.-J. Liu, W.-H. Tseng, A. M. Ahmed, H.-C. Wang, and B.-S. Lin, "A Deep Learning-Based Chair System That Detects Sitting Posture," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 1, pp. 482–490, Jan. 2024, doi: [10.1109/JBHI.2023.3337125](https://doi.org/10.1109/JBHI.2023.3337125).

表 2 不同系統特色與比較

	Ma et al.	Hu et al.	Kulikaievas et al.	Jeong et al.	Wang et al.	Fan et al.	Lin et al.	我們的系統
感測器	IMU	彎曲感測器	深度相機	壓力和距離感測器	壓力感測電阻器	壓力感測器	深度相機	普通相機
感測形式	加速度計	壓力感測器	視覺辨識	壓力感測器	壓力感測器	壓力感測器	視覺辨識	視覺辨識
感測器數量	1	6	1	12	171	2288	2	1
姿勢數量	5	7	6	8	15	5	42	3
實時回應	否	否	否	否	是	否	是	是
硬體平台	個人電腦 Arduino	Arduino FPGA	個人電腦	個人電腦	個人電腦	個人電腦	樹莓派	個人電腦
模型	SVM	ANN	DRHN	KNN	Liquid state machines(LSMs)	SVM and CNN	EfficientNet Lite2	SVM
準確率	95.33%	97.80%	91.47%	92%	99.90%	99.82%	99.71%	99.90%