

國立臺北大學資訊工程學系專題報告

利用人工智慧判斷高跟鞋高度合適範圍之手機應用程式評估系統

Artificial Intelligence Based Mobile Application for Identifying Suitable Range of Heel Height

專題組員：蘇曼文、黃怡庭、吳思妤、康璣尹

專題編號：PRJ-NTPUCSIE-109-001

執行期間：2020 年 09 月 至 2021 年 06 月

1. 摘要

對現代女性而言，高跟鞋為一項重要的時尚單品，不但能拉長腿部線條，也能改善身材比例，然而穿著過高的高跟鞋將對健康帶來不可逆的傷害，但是大多數女性卻不清楚自己適穿的高跟鞋跟高的極限。基於這個需求，本研究開發了一個具有人工智慧評估的手機應用程式(application；APP)，可以很簡單方便地替女性評估可穿高跟鞋高度的極限值。我們共招募了 100 位女性受測者，由醫生評估每位受測者合適的跟高極限並提供作為模型的標準答案，另外透過 APP 收集受測者由自然站立踮高至到極限的後腳跟連續影像，將此連續影像餵入人工智慧模型進行訓練及預測受測者的高跟鞋跟高極限。最後，使用者僅需使用此手機 APP 自行錄製後腳跟踮腳影片，APP 即可快速評估每個人最適合的高跟鞋鞋跟高度，平均絕對誤差為 1.29 公分。

2. 簡介

對女性而言，穿著高跟鞋能拉長腿部線條及改善身材比例，即使會帶來行動不便、磨腳、肌肉痠痛等負面影響，高跟鞋仍為現代女性愛不釋手的時尚單品。然而不同人之間存在個體差異，足部構造的不

同使得每個人能承受的高跟鞋高度因人而異。已有相關研究指出穿著不同高度的高跟鞋或長期穿著高跟鞋對於人體足部所造成的影響。研究顯示所有高跟鞋造成的傷害當中，有 72%與腳踝及腳部有關[1]，如拇趾外翻、肌肉骨骼疼痛以及骨關節炎[2]，更有研究指出，當穿著過高的高跟鞋，輕則下背痛，拇趾外翻、變形，小腿痠攣；重則造成踝關節韌帶拉傷，罹患膝關節炎，甚至引發椎間盤突出 [3] - [4]，對健康帶來不可逆的傷害。然而女性卻難以透過簡單的方式知道自己合適的高跟鞋跟高，傳統上必須至醫院由專業醫師進行觸診，透過觀察後腳跟在不同高度的形變，並經過換算才能得到適穿的高跟鞋跟高，這對一般人來說是非常不便的。

然而過往極少有文獻探討女性在挑選高跟鞋時最適合的高度極限，目前僅有的研究為國立臺北大學資訊工程學系的嵌入式系統與應用實驗室的研究團隊與臺北榮總復健醫學部的聯合團隊提出的基於人工智慧 (artificial intelligence；AI) 的高跟鞋高度評估平台 [5]。此平台能夠快速且精準地評估使用者合適的高跟鞋高度，評估誤差 1.21 公分。然而此系統的缺點在於此系統必須要搭配一特殊平台才能使用，但此平台過於龐大，會有空間及成本的限制。除了在操作上不便外，也

不容易在市場上廣泛推廣。因此我們期望改良此系統的不便，開發出一套能由使用者單獨操作，不需仰賴額外平台的評估系統。

因此，我們開發出一款基於 AI 的高跟鞋跟高極限評估手機應用程式(application；APP)，能夠客觀評估使用者最合適的高跟鞋高度，讓使用者能隨時隨地透過本研究所開發之 APP 來了解自身最適合的高跟鞋跟高，是本研究欲達成之目的。

3. 研究方法

此系統的整體系統架構如 Fig. 1.所示。本系統的軟體構成包括 APP 使用者操作介面、影像處理模組、自動框選模組、跟高極限預測模組、高跟鞋極限高度推算模組。除了 APP 使用者操作介面在手機上運行外，其餘的四個模組皆在電腦端上運行。操作時，使用者開啟智慧型手機上的高跟鞋跟高極限評估 APP 來錄製自己憑自身腿部肌力踮起腳尖的影片並上傳至電腦端，電腦端的程式會將此踮腳尖影片轉換為每秒 30 張的連續影像，餵入第一個 AI 模型(自動框選模組)，第一個 AI 模型將自動圈選每張影像中的左右腿以及比例尺紙盒模型。圈選後的影像則會餵入第二個 AI 模型(跟高極限預測模組)判斷出每張照片為可穿或不可穿，最後所有的影像的判斷結果會透過高跟鞋極限高度推算模組進行高度轉換，透過找出可穿與不可穿之間的臨界影像即可推算出該使用者可穿的高跟鞋極限高度。

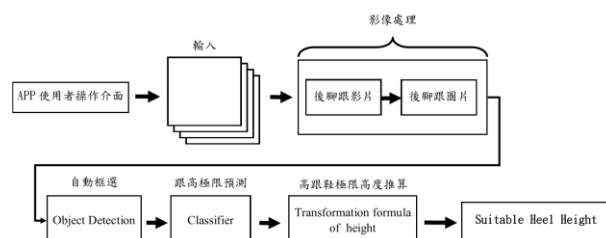


Fig. 1. 整體系統架構圖

3.1. 資料收集

本研究的資料集收集流程分為三個階段：受測者篩選及足部資料蒐集、醫師診斷並提供正確答案及後腳跟影片錄製。

3.1.1. 受測者篩選及足部資料蒐集

本研究共收測了 100 位女性受測者，受測者平均年齡為 21.2 歲，受測對象納入條件如下：足歲年齡為 20 至 35 歲之女性、弓率介於 11 至 14%、後腳跟歪斜角度 ≤ 4 度、大拇指外翻 ≤ 20 度、小拇指內翻 ≤ 20 度、不具不明原因性脊髓側彎、足底不具繭及過去無嚴重下肢受傷病史。再針對符合條件者進一步蒐集其腳長、腳寬、腳幅等足部數據。納入足歲年齡為 20 至 35 歲之女性的原因為在此年齡區間的人大多骨骼已發育健全且尚未進入骨骼退化階段。

3.1.2. 醫師診斷並提供正確答案

所有受測者個別由醫師進行人工診斷來獲得可穿的高跟鞋極限高度的標準答案。受測時，醫師操作由林伯星教授團隊所開發之升降平台[5]來觀察受測者後腳跟的形變並診斷其跟高極限，此跟高極限結果將作為後續跟高極限預測模型的正確答案。此平台後側的數位游標尺會隨千斤頂抬升顯示其目前抬升的高度，方便醫師了解目前後腳跟位於何種高度，如

Fig. 2.所示。評估時，受測者須保持自然站立於系統平台上，雙手垂放於大腿兩側，腳底橫弓需對齊在平台上升的轉折處。此平台可由醫師自行調整平台內部的電動千斤頂抬升高度來模擬不同的高跟鞋跟高，方便醫師評估並提供該受測者的跟骨發生內翻的高度。



Fig. 2. 醫師診斷平台操作示意圖



Fig. 3. 錄製踮腳尖影片示意圖

3.1.3. 後腳跟影片錄製

在後腳跟影片錄製方面，我們固定使用一台搭載 Android 10 的智慧型手機 (OPPO R17)，其後置相機鏡頭為 1600 萬畫素，光圈為 F1.7，手機相機鏡頭與受測對象後腳跟相距 73 公分，並且鏡頭對齊雙腳連線中點，以維持拍攝的視角，避免影響錄下的資料。除此之外，在右腳外側距離雙腳連線中點 15.5 公分處擺上用來作為比例尺的各邊長均為 5 公分之正方形紙盒，放置此正方形紙盒之目的在於後續

用來作為推算高度的比例尺，正方形紙盒的被拍攝面需與兩後腳跟連線切齊。受測者依指示將雙手垂放於大腿兩側，並維持穩定速率漸漸踮起腳尖直到最高處，組員同時開始錄影記錄腳跟跟骨與阿基里斯腱的彎曲角度變化。Fig. 3. 為實際拍攝踮腳尖影片的示意圖。

3.2. 深度學習模型

本研究採用兩個深度學習模型來分別實現自動框選及跟圖片辨識。電腦端用來開發的程式語言為 Python 3.8，使用的深度學習框架為 Tensorflow 2.4，用來進行深度學習的顯示卡為 GeForce RTX 3080 (ASUSTeK Computer Inc., Taipei, Taiwan)。

3.2.1. 自動圈選模組

本研究使用 EfficientDet 物件偵測模型 [6] 來實現自動圈選，Fig. 4. 為 EfficientDet 的架構圖。EfficientDet 為 EfficientNet 的延伸，使用了新的特徵融合方法加權雙向特徵金字塔網路 (Bi-directional Feature Pyramid Network ; BiFPN)，並以 ImageNet 預訓練的 EfficientNet-B0 作為基底。BiFPN 從基底取得特徵後，能夠自上而下及自下而上進行雙向特徵融合，並將融合的特徵輸入 Class/Box 預測網路，同時預測出圈選區域以及類別。

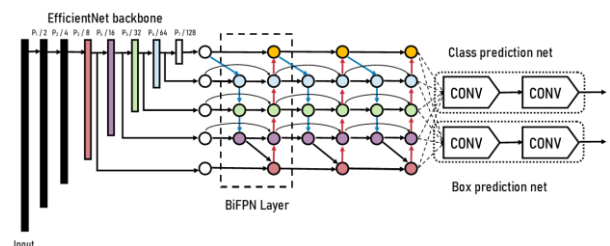


Fig. 4. EfficientDet 架構圖

本研究透過手機 APP 錄製好的每段影片都會先上傳至電腦端，電腦端則開始進行每秒 30 張的自動擷圖產生影像，每張影像內包含受測者的左右後腳跟、比例尺紙盒及白色布景，白色布景的目的在於避免模型在學習過程中受到環境其他雜物影響導致框選的準確度下降。所有影像均會先人工標記出左右腿以及比例尺紙盒，來作為自動框選模型的訓練資料，經標記處理的照片如 Fig. 5. 中綠色框所示。經人工框選出的左右腿及比例尺紙盒將餵入 EfficientDet 進行訓練，將測試集餵入訓練好的自動圈選模型即可自動圈選出左右腿及比例尺紙盒。

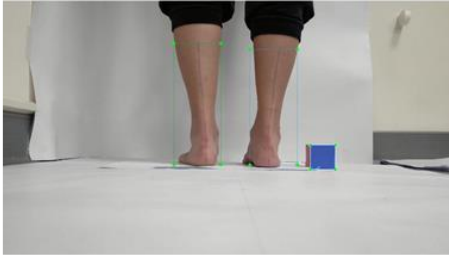


Fig. 5. 人工標記之左右腿及比例尺紙盒示意圖

除此之外，此模型內部含有高度換算程式，可以利用紅色比例盒和被圈選的足部兩者之比例推算出該後腳跟離地的實際高度。根據將影片轉圖片的前處理及 EfficientDet 模型，我們可以得到 k 組含左腳、右腳及紅色盒子的 bounding 座標，bounding 座標包含了框選區的左上角及右下角座標。

先將 k 個紅色盒子 bounding 座標的左上角 Y 座標及右下角 Y 座標相減，我們可以得到 k 個紅色盒子的高度，單位為

pixel，再將 k 個盒子的高度差除以 k 可以得到紅色盒子的平均高度，單位為 pixel。

接著，將第 n 張圖右腳 bounding 座標的右下角 Y 座標和第一張圖右腳 bounding 座標的右下角 Y 座標相減，我們可以得到第 n 張圖相對於第一張圖腳踎高了多少高度，單位為 pixel。之後將此數值除上紅色盒子的平均高度，我們可以得到第 n 張圖相對於第一張圖腳踎高了多少高度與紅色盒子平均高度的比值，將此比值乘上已知的紅色盒子真實高度 5 公分，便可得到第 n 張圖相對於第一張圖腳踎高了多少公分。左腳同理。實際計算的各公式如(1)–(5)所示。

- 盒子在照片內的高度 (pixel) :

Average box height =

$$\frac{\sum_{t=1}^k |YRD_B(t) - YLU_B(t)|}{k}$$

(1)

- 右腳在照片內的高度 (pixel) :

$nth \text{ Right leg height}_{pixel} =$

$$|YRD_R(n) - YRD_R(1)| \quad (2)$$

- 右腳在現實中的高度 (cm) :

$$\begin{aligned} nth \text{ Right leg height}_{cm} \\ = \frac{nth \text{ Right leg height}_{pixel}}{\text{Average box height}} \\ \times 5 \end{aligned} \quad (3)$$

- 左腳在照片內的高度 (pixel) :

$nth \text{ Left leg height}_{pixel} =$

$$|YRD_L(n) - YRD_L(1)| \quad (4)$$

- 左腳在現實中的高度 (cm)：

$$\begin{aligned} & \text{nth Left leg height}_{cm} \\ &= \frac{\text{nth Left leg height}_{pixel}}{\text{Average box height}} \\ & \times 5 \quad (5) \end{aligned}$$

k 為一部影片切割後照片的張數； N 為 k 張照片中的第 n 張，範圍為 1 到 k ；(XLUL, YLUL) 代表左腳左上角的 XY 座標；(XRDL, YRDL) 代表左腳右下角的 XY 座標；(XLUR, YLUR) 代表右腳左上角的 XY 座標；(XRDR, YRDR) 代表右腳右下角的 XY 座標；(XLUB, YLUB) 代表紅色盒子左上角的 XY 座標；(XRDB, YRDB) 代表紅色盒子右下角的 XY 座標。圈選出的影像會被標記為左腳影像或是右腳影像，並餵入相對應的跟高極限預測模組進行預測。

3.2.2. 跟高極限預測模組及高跟鞋極限高度推算模組

本研究透過 Keras 2.4.0 [7] 實作卷積神經網路[8](convolutional neural network; CNN) 來實現跟高極限預測模組，並配合醫生給予的跟高極限作為正確答案進行分類，來判斷此影像為可穿或不可穿。跟高極限預測模組之 CNN 架構圖如 Fig. 6. 所示。此 CNN 架構 kernel size 為 3×3 ，pool size 為 2×2 ，dropout 比例為 0.25。首先將圈選出來的左右腳後腳跟影像分別餵入所屬的模型，左腳則餵入左腳預測模型，右腳則餵入右腳預測模型，經過三層的卷積 (convolution)、dropout、最大池化 (maxpooling)，再將特徵展平後輸入三個全連接層進行訓練並取得權重，最後模型會將此影像分類為 1 或 0。若為可穿高度則輸出為 1，否則輸出為 0，如 Fig. 7. 所

示的每個點。這些點在經過高跟鞋極限高度推算模組可以得到一斷層，如 Fig. 7 中的 0/1 斷層圖。之後取出分類結果 1 與 0 之間的臨界影像，透過內建於 EfficientDet 模型的高度推算公式即可推算出適穿的極限跟高並將極限跟高回傳並呈現在手機 APP 上。

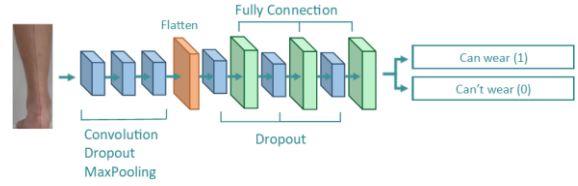


Fig. 6. CNN 架構圖

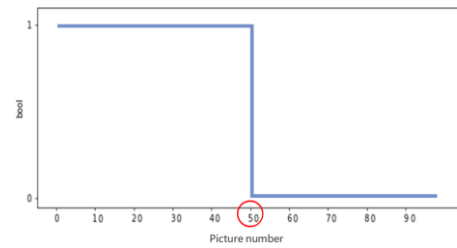


Fig. 7. 0/1 斷層圖及極限跟高預測

3.3. APP 使用者操作介面

此跟高極限評估 APP 於 Android Studio 上開發，此為 Google 發布的 Android 平台開發程式之整合式開發環境，並使用 Java 作為開發語言及搭配 API 16: Android 4.1 (Jelly Bean)。APP 介面圖如 Fig. 8 所示，包含使用說明、開啟相機、上傳按鈕。

此 APP 使用流程如下，首先點選使用說明按鈕將透過 YouTube 播放預錄好的操作示範影片供使用者方便理解操作方式。接著點選開啟相機後便會開啟內建相機，讓使用者能進行錄影。拍攝完的影片會呈現在介面中間，選取欲上傳的影片後再點

選右上方的上傳按鈕，即可將錄製好的影片上傳到電腦端。電腦端程式即可進行極限跟高的預測，預測結果會回傳並顯示於使用者的 APP 介面上。Fig. 9 為已回傳結果的 APP 使用者操作介面。

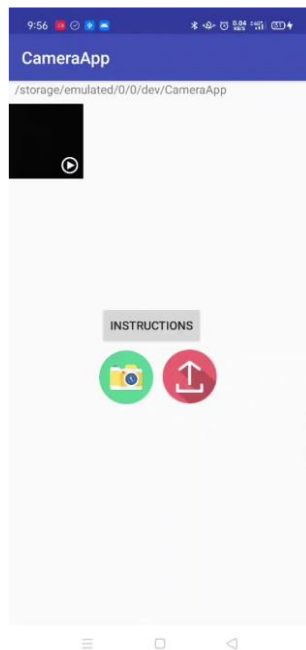


Fig. 8. APP 使用者操作介面

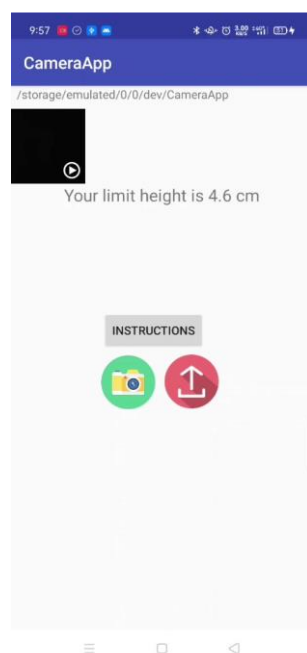


Fig. 9.已回傳結果的 APP 使用者操作介面。

4. 主要成果與評估

此研究開發了一套能夠透過一手機 APP 讓使用者方便且快速進行自動高跟鞋極限跟高預測的系統。

物件偵測方面, EfficientDet 模型以 80 人為訓練測資, 共輸入 7815 張照片, 訓練完畢再使用另外 20 人的資料對左腳、右腳及紙盒各自使用 1953 張照片做預測, 並且額外輸入不在訓練測資的 4 人共 385 張照片進行交比(intersection over union; IOU), 這 4 人框選左腳部分平均 IOU 為 0.96, 框選右腳部分的 IOU 平均 IOU 為 0.97, 框選紙盒部分的平均 IOU 為 0.97, 整體而言各模型之平均 IOU 高達 0.97。

而跟高極限預測模組中的腳部圖片辨識方面, 我們以 100 人為訓練測資, 並進行 10-fold 交叉驗證, 每批訓練的照片為 8791 張, 預測的張數為 977 張。訓練完畢的 CNN 模型可以得到平均 accuracy (0.96)、loss (0.09)。另外, 我們也餵入了不被 CNN 模型 training 與 testing 的額外 4 人進行預測。4 個人的 accuracy 分別為 98.23%, 95.65%, 100.00%, 98.90%, 平均 accuracy 為 98.20%。

將跟高極限預測模組與高跟鞋極限高度推算模組的兩種模型串接後, 以沒被訓練過的 4 人測資分別對左右腳進行預測, 左腳平均平均絕對誤差(mean absolute error; MAE)為 2.22 公分, 右腳平均 MAE 為 0.37 公分, 整體的平均 MAE 為 1.29 公分, 只比半吋多 0.02 公分, 小於市售高跟

鞋跟高度的常見級距。故此系統在鞋跟高度的預測上有著良好的實用性。

未來我們希望能針對模型架構做更多的處理與改良，以期提高系統的準確度，並讓整體執行時間縮短，讓 APP 能更即時回傳結果。

5. 結語與展望

此研究開發出一個簡易且方便的 APP，能夠讓使用者不需額外的平台即可自動評估出自己的高跟鞋跟高極限。此外，本研究的深度學習模型的整體準確率為 97%，MAE 為 1.29 公分。未來我們期望對模型架構進行改良，或是使用不同的模型架構，來提升系統的跟高預測的準確率以及加快 APP 跟高回傳的速度，讓使用者的體驗更加即時。

6. 銘謝

非常感謝指導教授提供大量資源與豐富的研究經驗分享，以及實驗室學長姐的協助，讓我們在進行專題的路上更加順利，也謝謝醫生與助理不辭辛勞，專程來三峽為我們的受測者們診斷出 AI 模型的正确答案。最後也感謝組員們的合作才能使我們的專題能如期完成。

7. 參考文獻

- [1] P. Prokop and J. Švancarova, "Wearing high heels as female mating strategy," *Pers. and Individ. Differ.*, vol. 152, art. no. 109558, pp. 1–5 2020.
- [2] J. X. Moore, B. Lambert, G. Jenkins, and G. M. Jr, "Epidemiology of High-Heel Shoe

Injuries in U.S.," *J. Foot & Ankle Surg.*, vol. 54, pp. 615–619, 2015.

[3] M. S Barnish and J. Barnish, "High-heeled shoes and musculoskeletal injuries: a narrative systematic review," *BMJ open*, vol. 6, no. 1, pp. 1–8, Jan. 2016.

[4] Y.-Y. Zhuang, "Study on Design of Parallel Reconfigurable Convolutional Neural Network Processor," M.S. thesis, Dept. Phy. Edu., Natl. Univ. of Tainan, Tainan, Taiwan, 2019.

[5] S.-H. Lee, B.-S. Lin, H.-C. Lee, X.-W. Huang, Y.-C. Chi, B.-S. Lin, and K. Abe, "Artificial Intelligence Based Assessment System for Evaluating Suitable Range of Heel Height," *IEEE Access*, vol. 9, pp. 38374–38385, March 2021.

[6] M. Tan, R. Pang, and Q. V. Le, EfficientDet: Scalable and Efficient Object Detection. [Online]. Available: <https://arxiv.org/abs/1911.09070>, Accessed on: July 27, 2020.

[7] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[8] Keras, Keras: The Python deep learning library, 2018. [Online]. Available: <https://keras.io>, Accessed on: August 18, 2020.