國立雲林科技大學資訊工程系

碩士論文

Department of Computer Science and Information Engineering National Yunlin University of Science & Technology

Master Thesis

步態分析與跌倒預防的關鍵因素研究

Study on Key Factors of Gait Analysis and Fall Prevention

洪郁晴

Yu-Ching Hung

指導教授：王文楓 博士

Advisor: Wen-Fong Wang, Ph.D.

中華民國 113 年 6 月  
June 2024

摘要

跌倒對老年人來說是一個重大問題，因為它可能會導致嚴重的身體傷害，甚至是長期的健康問題。隨著年齡的增長，老年人會經歷肌肉力下降、平衡能力衰弱和反應時間變慢等變化，這些變化往往會影響步態，從而增加跌倒的風險。預防跌倒的策略包括進行定期的運動和平衡訓練，如瑜伽和太極，這些活動有助於增強肌肉力量和改善平衡，且可以顯著降低老年人跌倒的風險，以提高他們的生活質量。

共有13名非跌倒者(年齡70±5歲，男性4人)和16名易跌倒者(年齡70±5歲，男性4人)參與了該研究，他們沒有明顯的肌肉骨骼或神經系統疾病。此外，還納入了22名健康的成年人(年齡20±2歲，男性6人)作為正常對照。記錄了性別、身體質量指數、之前的跌倒事件和10公尺步行測試的結果。利用來自每次10公尺步行測試中部分五個步態循環的加速度來計算步態穩定性的單側交叉相關係數和步態對稱性的雙側交叉相關係數。採用支援向量機和K-近鄰演算法來對老年非跌倒者、老年易跌倒者和健康的成年人進行分類。

關鍵字：步態、三軸加速度計、老年、互相關、KNN、SVM、RF

ABSTRACT

Falls are a significant issue for the elderly, as they can lead to serious physical injuries and even long-term health problems. With aging, older adults experience a decline in muscle strength, weakened balance, and slower reaction times, which often affect their gait and increase the risk of falling. Strategies to prevent falls include regular exercise and balance training, such as yoga and tai chi, which help strengthen muscle power and improve balance, significantly reducing the risk of falls in the elderly and enhancing their quality of life.

A total of 13 non-fallers (aged 70±5 years, 4 males) and 16 fallers (aged 70±5 years, 4 males), without significant musculoskeletal or neurological diseases, participated in the study. Additionally, 22 healthy adults (aged 20±2 years, 6 males) were included as a normal control group. Gender, body mass index, previous fall events, and results of a 10-meter walk test were recorded. The unilateral cross-correlation coefficient for gait stability and the bilateral cross-correlation coefficient for gait symmetry were calculated using acceleration from five gait cycles in each 10-meter walk test. Support Vector Machines and K-Nearest Neighbor algorithms were used to classify elderly non-fallers, elderly fallers, and healthy adults.

Keywords：Gait, three-axis accelerometer, elderly, cross-correlation、KNN、SVM、RF

目錄

[摘要 i](#_Toc174458253)

[ABSTRACT ii](#_Toc174458254)

[目錄 iii](#_Toc174458255)

[表目錄 v](#_Toc174458256)

[圖目錄 viii](#_Toc174458257)

[第一章 緒論 1](#_Toc174458258)

[1.1 研究背景 1](#_Toc174458259)

[1.2 相關研究 1](#_Toc174458260)

[1.3 研究動機 2](#_Toc174458261)

[1.4 研究目的 2](#_Toc174458262)

[1.5 論文架構 3](#_Toc174458263)

[第二章 材料和方法 4](#_Toc174458264)

[2.1 步態動作及感測訊號 4](#_Toc174458265)

[2.2 研究儀器 5](#_Toc174458266)

[2.3 實驗方法 6](#_Toc174458267)

[2.4 自相關/互相關運算 7](#_Toc174458268)

[2.5 左/右雙腳步態相關性資料模型建立 8](#_Toc174458269)

[2.6 步態訊號特徵參數之設計 9](#_Toc174458270)

[第三章 研究方法 11](#_Toc174458271)

[3.1 訊號前處理 11](#_Toc174458272)

[3.2 步態週期切割 12](#_Toc174458273)

[3.3 步態特徵提取 13](#_Toc174458274)

[3.4 窮舉法之特徵分析 16](#_Toc174458275)

[3.5 步態種類之分類 18](#_Toc174458276)

[3.5.1 KNN 18](#_Toc174458277)

[3.5.2 SVM 19](#_Toc174458278)

[3.5.3 RF(Random forest) 20](#_Toc174458279)

[第四章 結果 22](#_Toc174458280)

[4.1 Mann-Whitney U檢定：特徵之顯著性分析 23](#_Toc174458281)

[4.1.1 Mann-Whitney U檢定：KNN 24](#_Toc174458282)

[4.1.2 Mann-Whitney U檢定：SVM 27](#_Toc174458283)

[4.1.3 Mann-Whitney U檢定：RF 31](#_Toc174458284)

[4.2 窮舉法分析 35](#_Toc174458285)

[4.2.1 KNN分析結果 37](#_Toc174458286)

[4.2.2 SVM分析結果 41](#_Toc174458287)

[4.2.3 RF(Random Forest)分析結果 44](#_Toc174458288)

[第五章 討論 49](#_Toc174458289)

[參考文獻 50](#_Toc174458290)

表目錄

[表一 F、NF之自相關/互相關序列參數Mann-Whitney U檢定結果 14](#_Toc174457964)

[表二 F、HA之自相關/互相關序列參數Mann-Whitney U檢定結果 14](#_Toc174457965)

[表三 NF、HA之自相關/互相關序列參數Mann-Whitney U檢定結果 15](#_Toc174457966)

[表四 F、NF之傳統步態參數Mann-Whitney U檢定結果 16](#_Toc174457967)

[表五 F、HA之傳統步態參數Mann-Whitney U檢定結果 16](#_Toc174457968)

[表六 NF、HA之傳統步態參數Mann-Whitney U檢定結果 16](#_Toc174457969)

[表七 電腦配備 17](#_Toc174457970)

[表八 混淆矩陣示意圖 22](#_Toc174457971)

[表九 F&NF於KNN之顯著性參數混淆矩陣 24](#_Toc174457972)

[表十 F&NF顯著性參數之KNN分類結果 25](#_Toc174457973)

[表十一 F&HA於KNN之顯著性參數混淆矩陣 25](#_Toc174457974)

[表十二 F&HA顯著性參數之KNN分類結果 25](#_Toc174457975)

[表十三 NF&HA於KNN之顯著性參數混淆矩陣 26](#_Toc174457976)

[表十四 NF&HA顯著性參數之KNN分類結果 26](#_Toc174457977)

[表十五 F&NF&HA於KNN之顯著性參數混淆矩陣 27](#_Toc174457978)

[表十六 F&NF&HA顯著性參數之KNN分類結果 27](#_Toc174457979)

[表十七 F&NF於SVM之顯著性參數混淆矩陣 28](#_Toc174457980)

[表十八 F&NF顯著性參數之SVM分類結果 28](#_Toc174457981)

[表十九 F&HA於SVM之顯著性參數混淆矩陣 29](#_Toc174457982)

[表二十 F&HA顯著性參數之SVM分類結果 29](#_Toc174457983)

[表二十一 NF&HA於SVM之顯著性參數混淆矩陣 30](#_Toc174457984)

[表二十二 NF&HA顯著性參數之SVM分類結果 30](#_Toc174457985)

[表二十三 F&NF&HA於SVM之顯著性參數混淆矩陣 31](#_Toc174457986)

[表二十四 F&NF&HA顯著性參數之SVM分類結果 31](#_Toc174457987)

[表二十五 F&NF於RF之顯著性參數混淆矩陣 32](#_Toc174457988)

[表二十六 F&NF顯著性參數之RF分類結果 32](#_Toc174457989)

[表二十七 F&HA於RF之顯著性參數混淆矩陣 33](#_Toc174457990)

[表二十八 F&HA顯著性參數之RF分類結果 33](#_Toc174457991)

[表二十九 NF&HA於RF之顯著性參數混淆矩陣 34](#_Toc174457992)

[表三十 NF&HA顯著性參數之RF分類結果 34](#_Toc174457993)

[表三十一 F&NF&HA於RF之顯著性參數混淆矩陣 35](#_Toc174457994)

[表三十二 F&NF&HA顯著性參數之RF分類結果 35](#_Toc174457995)

[表三十三 F&NF於KNN之最佳特徵組合混淆矩陣 39](#_Toc174457996)

[表三十四 F&NF之KNN最佳特徵組合分類結果 39](#_Toc174457997)

[表三十五 F&HA於KNN之最佳特徵組合混淆矩陣 39](#_Toc174457998)

[表三十六 F&HA之KNN最佳特徵組合分類結果 40](#_Toc174457999)

[表三十七 NF&HA於KNN之最佳特徵組合混淆矩陣 40](#_Toc174458000)

[表三十八 NF&HA之KNN最佳特徵組合分類結果 40](#_Toc174458001)

[表三十九 F&NF&HA於KNN之最佳特徵組合混淆矩陣 41](#_Toc174458002)

[表四十 F&NF&HA之KNN最佳特徵組合分類結果 41](#_Toc174458003)

[表四十一 最佳特徵組合之SVM分類結果 42](#_Toc174458004)

[表四十二 F&NF於SVM之最佳特徵組合混淆矩陣 42](#_Toc174458005)

[表四十三 F&NF於SVM之RBF最佳特徵組合分類結果 42](#_Toc174458006)

[表四十四 F&HA於SVM之最佳特徵組合混淆矩陣 43](#_Toc174458007)

[表四十五 F&HA於SVM之RBF最佳特徵組合分類結果 43](#_Toc174458008)

[表四十六 NF&HA於SVM之最佳特徵組合混淆矩陣 43](#_Toc174458009)

[表四十七 NF&HA於SVM之RBF最佳特徵組合分類結果 44](#_Toc174458010)

[表四十八 F&NF&HA於SVM之最佳特徵組合混淆矩陣 44](#_Toc174458011)

[表四十九 F&NF&HA於SVM之RBF最佳特徵組合分類結果 44](#_Toc174458012)

[表五十 F&NF於RF之最佳特徵組合混淆矩陣 45](#_Toc174458013)

[表五十一 F&NF於RF之最佳特徵組合分類結果 46](#_Toc174458014)

[表五十二 F&HA於RF之最佳特徵組合混淆矩陣 46](#_Toc174458015)

[表五十三 F&HA於RF之最佳特徵組合分類結果 46](#_Toc174458016)

[表五十四 NF&HA於RF之最佳特徵組合混淆矩陣 47](#_Toc174458017)

[表五十五 NF&HA於RF之最佳特徵組合分類結果 47](#_Toc174458018)

[表五十六 F&NF&HA於RF之最佳特徵組合混淆矩陣 47](#_Toc174458019)

[表五十七 F&NF&HA於RF之最佳特徵組合分類結果 48](#_Toc174458020)

圖目錄

[圖一 步態週期示意圖 4](#_Toc174458021)

[圖二 穿戴式慣性步態感測器 5](#_Toc174458022)

[圖三 (a)感測器穿戴位置示意圖、(b)從感測器接收到的加速度訊號 6](#_Toc174458023)

[圖四 10公尺步行實驗之工作流程圖 7](#_Toc174458024)

[圖五 交叉互相關運算所產生之相關性序列示意圖 8](#_Toc174458025)

[圖六 (a)老年易跌倒者、(b)老年非跌倒者、(c)健康成年人 9](#_Toc174458026)

[圖七 加速度X軸訊號於SG濾波前/後 11](#_Toc174458027)

[圖八 右腳RSS訊號於Butterworth濾波前/後 12](#_Toc174458028)

[圖九 經過濾波和切割後的步態訊號數據(橘色:右腳;藍:左腳) 12](#_Toc174458029)

[圖十 F&NF不同k值之效能圖 37](#_Toc174458030)

[圖十一 F&HA不同k值之效能圖 37](#_Toc174458031)

[圖十二 NF&HA不同k值之效能圖 38](#_Toc174458032)

[圖十三 F&NF&HA不同k值之效能圖 38](#_Toc174458033)

1. 緒論
   1. 研究背景

跌倒是對老年人健康和身體功能造成嚴重影響的主要因素之一，統計數據顯示每年約有30%的老年人經歷跌倒[1]，平均每 100,000 名老年人中就約有13,840人(範圍從每 100,000 人的 7,594 人到 19,796 人)因跌倒相關傷害尋求醫療幫助[2]。這些跌倒事件中，步態變化所誘發在身體移動中動作姿態的平衡問題可能扮演了重要的角色[3]。隨著年齡增長，尤其是在50歲之後，人體的肌肉質量每年會減少大約1-2%[4]，這可能對步態穩定性造成影響。儘管如此，有研究指出，在步行速度[5]或步態變異性[6]方面並未發現明顯與年齡相關的差異。因此，對老年人步態變化的研究仍存在爭議，找出可以明確分析和識別老年人步態變化的具體參數，對於有效識別具高跌倒風險的老年人至關重要。

* 1. 相關研究

一些研究人員採用機器學習演算法，透過使用慣性測量單元(IMUs)所獲取的步行數據來識別高跌倒風險的步態，其準確率介於79%–94%之間[7, 8]。儘管研究人群、獲取方法、數據處理和分類方面存在相當大的變化，但對於如何使用監督式和非監督式機器學習演算法來研究老年人和高跌倒風險老年人的步態變化，尚未達成共識[9]。近年來，互相關運算(Cross-correlation)已被用來研究帕金森氏症[10]中雙側腿部步態訊號耦合的程度[11]，以及區分社區居住的老年人中的跌倒者和非跌倒者[8]。在以上的研究中，互相關運算之應用顯現了顯著的研究潛能。

在之前的研究中，由k-最近鄰居(k-Nearest Neighbor: kNN)辨識演算法揭示出近似週期性的IMU訊號時序，包含了關於步態穩定性和對稱性的必要細節，並在跌倒風險評估中表現良好[8]。先前的研究顯示，機器學習可以對 IMU 的運動訊號進行分類，區分老年易跌倒者和非跌倒者，準確率為81%-86% [12, 13]。其中有一個模擬研究使用了微多普勒雷達訊號和主成分分析來評估健康成年人、老年非跌倒者和老年易跌倒者之間的步態特性，包括平均值、標準差和最大步行速度[14]。

* 1. 研究動機

為了更準確地識別潛在跌倒風險較高的老年人，本研究之重點是透過步態的三軸加速度訊號，尋找能夠凸顯出非正常步態變化的參數，以利於識別出個別高跌倒風險之老年人。本研究工作涉及利用互相關運算針對IMUs訊號結合各種類機器學習技術，來進行步態數據分析。探討這種方法之目的，在於提升對老年人跌倒風險的評估精準度。本研究的創新之處，在於探索綜合使用監督式機器學習演算法、自相關/互相關運算、高辨識度步態訊號特徵之篩選策略研究、以及統計分析，來建立、分析和評估健康成年人、老年非跌倒者和老年易跌倒者的雙腿步態加速度訊號之顯著異常特徵，進而更利於使用IMU測量老年人之步態，用於高跌倒風險者的識別。

* 1. 研究目的

本研究旨在透過監督式機器學習技術，對老年人在步態的微小變化上，進行細緻的辨識和分類，特別是在識別老年非跌倒者和易跌倒者的類別方面。研究將重點放在使用雙側下肢的穿戴式感測器來收集步態數據，並利用監督式學習演算法來區分細微的步態變化。此外，研究還計劃比較其他機器學習方法在分類健康成年人、老年非跌倒者和易跌倒老年人的步態模式方面的效能，以此來驗證和評估不同技術在老年人步態分析領域的應用效果，這包括對跌倒風險的評估和老化相關步態變化的辨識。

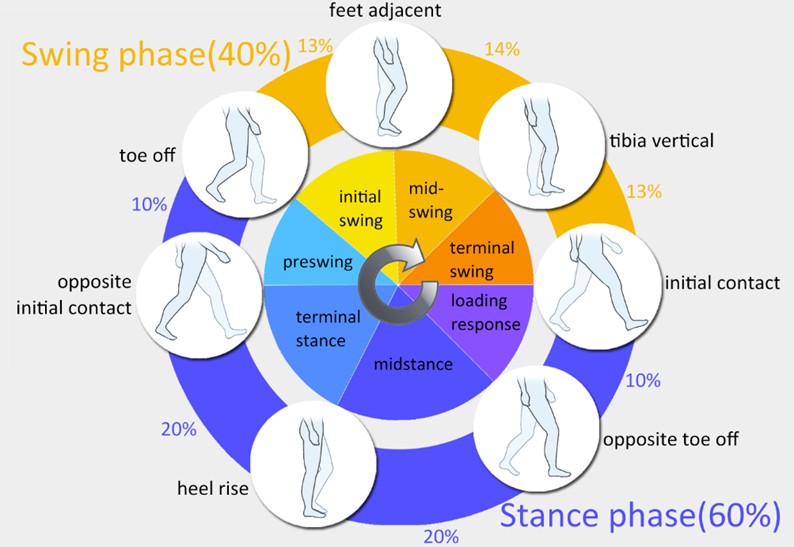
這項研究將有助於深入理解老年步態變化的機制，並為開發更有效的跌倒預防和治療策略提供科學依據。以期能為基於精準醫學的理念，建立智慧辨識技術之應用基礎。

* 1. 論文架構

論文共有五個主要章節，首先，第一章為研究背景，闡述了進行此研究的動機和主要目的。接著，第二章會詳細介紹實驗所使用的材料和儀器與分析軟體，並描述從這些儀器中收集到的訊號特徵。此外，本章還包含了對實驗參與者的選擇標準以及實驗的具體流程和執行方法。第三章則專注於訊號處理方法之介紹，包括濾波器的使用以及特徵計算和提取技術的應用。第四章展示了機器學習技術在對參與者實驗數據進行訓練和分類後得到的結果。最後，第五章總結了整個研究的主要發現，並對未來的研究方向提出了展望。

1. 材料和方法
2. 步態動作及感測訊號

圖一展示了步態週期(gait cycle)的兩個主要階段：擺動階段(Swing Phase)和站立階段(Stance Phase)，以及各自包含的細分階段和其所佔的時間比例。步態週期是指從週期內的某一個步態動作特徵開始，經過一系列動作，直到同一步態動作特徵再次出現的一整個過程，例如：如圖一所示，從一個腳跟接觸地面開始，經過一系列動作，直到同一腳跟再次接觸地面為止的整個過程。



圖一 步態週期示意圖

站立階段(Stance Phase)約佔整個步態週期的60%，它包括以下幾個步態間隔(Gait Event Interval: GEI)：負重反應(Loading Response)、中期站立(Midstance)、末期站立(Terminal Stance)、預擺動(Preswing)。擺盪階段(Swing Phase)佔整個步態週期的40%，它包括以下幾個階段：擺動初期(Initial Swing)、擺動中期(Midswing)和擺動末期(Terminal Swing)。此外，步態週期可以細分為七個步態事件:

1.初始接觸或腳跟著地，當腳跟接觸地面時；

2.對側腳趾離地；

3.腳跟離地，即腳跟開始離開地面；

4.對側腳跟著地；

5.當腳離開地板時，腳趾離地；

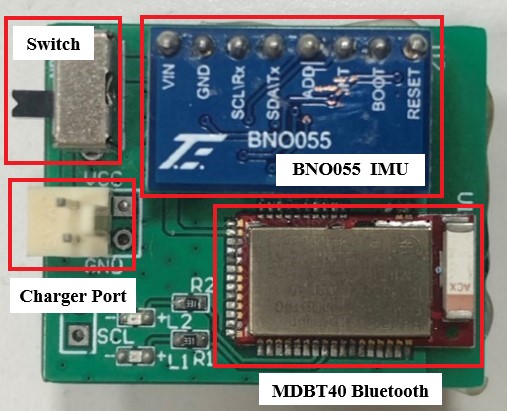
6.腳相鄰，這是擺動肢體經過站立肢體的時間；

7.脛骨垂直，當擺動肢體的脛骨與垂直軸相對應時。

使用穿戴式慣性感測器在收到訊號時，加速規會在不同方向上測量加速度。同時，又因環境因素的影響，使得不同個人在行走動作上會產生不同的加速度測量誤差。為了更全面地去除和降低3度空間內在各維度方向上，測量加速度訊號時所受到的測量誤差之影響，我們需要使用將三軸加速度訊號的平方和取平方根(公式一)之值，作為去除和降低方向性誤差對後續研究結果所產生顯著影響的手段，其公式如下:

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | 公式一 |

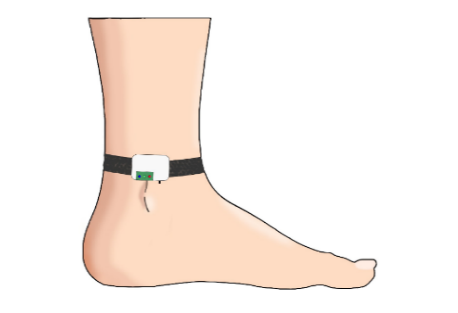
1. 研究儀器



圖二 穿戴式慣性步態感測器

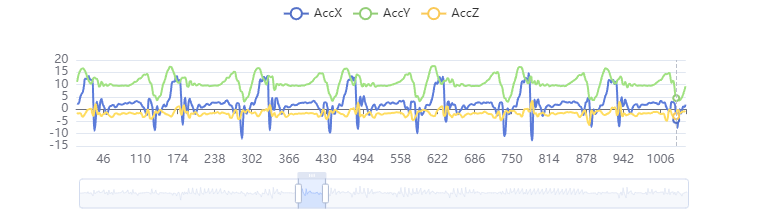
本論文使用的步態測量儀器是我們實驗室自行研發之穿戴式慣性感測器，如圖二所示。其核心部件為穿戴式慣性感測器，這款感測器包括BNO055 IMU模組，具備加速度計、陀螺儀和磁力計，使其能準確追蹤物體在空間中的動作和方向變化。此外，儀器還整合了MDBT40藍芽模組以實現無線連接，並通過電力供應單元和鋰電池提供所需的能源。這些元件的綜合運作，使得感測器能夠高效地進行步態測量。

在步態訊號感測中，我們使用魔鬼氈綁帶將感測器穩固地綁在受試者的雙腳腳踝上方約2.5 cm處(圖三(a))。待慣性感測器固定後，開啟感測器的電源開關，同時啟動平板電腦藍牙功能，以確保平板電腦能夠向步態感測模組發出控制指令，並準確接收來自感測模組回傳的IMU感測訊號。



(a)

(b)



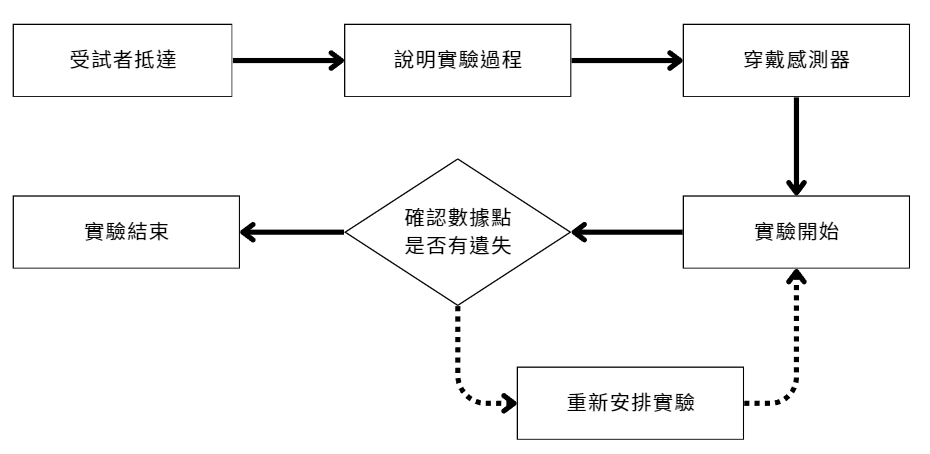
圖三 (a)感測器穿戴位置示意圖、(b)從感測器接收到的加速度訊號

1. 實驗方法

本研究使用於 2018 年 6 月 1 日至 2019 年 11 月 30 日在台灣南部的兩個社區服務中心招募的 27 名年齡在 65 歲或以上的當地居民(從 31 名定期到這些服務中心進行健康檢查的社區居民中選出，檢查包括身高、體重和血壓測量)，他們在其社區居住均已超過 1 年，這些老年成年人在日常生活活動中表現出能自主獨立行動，並能夠不倚賴外力行走超過 10 公尺[15]。

我們還納入了 22 名健康的成年人(年齡範圍 18–22 歲)，他們在過去 1 年內沒有跌倒的歷史，以便比較和老年成年人之間的步態特徵差異。這些成年人是志願者，並接受與老年成年人相同的評估。本研究已得到機構審查委員會(Institutional Review Board: IRB)的批准(批准編號 B-ER-110-152)。

首先，向實驗參與者說明實驗流程(圖四)，接著，引導實驗參與者根據測試人員的指示，站立於步行實驗的起始點位置，於聽到測試人員發出的語音提示後，開始步行並向前直行10公尺。當遇到三角錐後繞行迴轉，並直行返回起始點。至起始點後，再次迴轉後繼續直行，直至返回起始點。最後，依照所收集的步態數據紀錄上的時間序列標記作檢查，確認所擷取和接收的步態數據的完整性。若步態數據紀錄有5%以上在訊號擷取或通訊中所造成的訊號遺失，則需要重新進行實驗，反之則結束實驗。



圖四 10公尺步行實驗之工作流程圖

1. 自相關/互相關運算

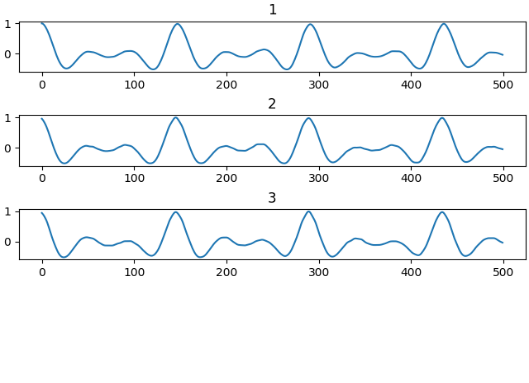
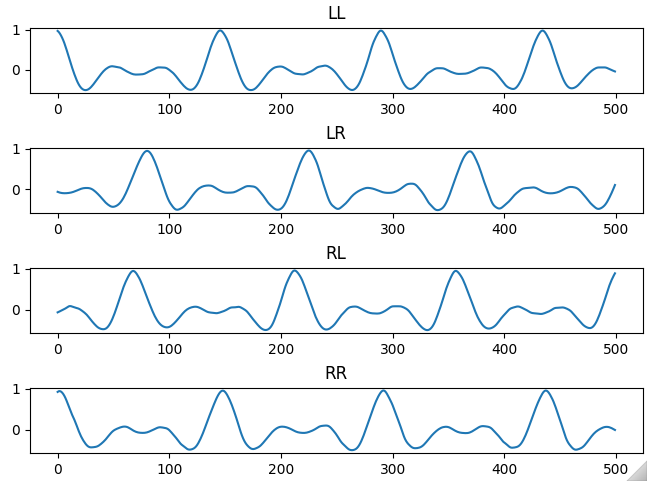
自相關/互相關運算因其能夠準確衡量兩組訊號之間的相關程度，在多個研究領域中被廣泛應用，例如：影像處理和電腦視覺、訊號分析、聲音和語音識別等。通過對原始步態訊號的自相關/互相關運算，計算處理後的相關係數值(例如：公式二的)會落在 -1 至 +1 的範圍內，這種方法不僅可以顯示兩組原始步態訊號序列之間的相關程度，同時還可以顯示出兩組步態訊號序列所對應的步態動作之間的相對時間位移，的詳細公式如公式二所示[11]。

|  |  |  |
| --- | --- | --- |
|  |  | 公式二 |

其中，N為原始訊號序列的長度，表訊號序列 p 在時間點 *i* 的值，是訊號序列 p 的平均值， 指 q 部分訊號序列(,…,)的部分平均值， 值的範圍在 1 到 -1 之間，若 的值接近 1，這意味著在延遲 t 的情況下，訊號序列 p 和 q 為強烈正相關；如果值接近 -1，則代表為強烈負相關；如果值接近 0，則表示沒有顯著的相關性。

1. 左/右雙腳步態相關性資料模型建立

在10公尺步態行走實驗中，我們選取行走過程中的中間三個步態週期的訊號，作為步態分析之依據。為了顯現出左腳和右腳在一個完整的行走步態週期中的差異，我們對整體三個步態週期的數據做歸一化互相關係數(Normalized Cross-correlation Coefficient: NCC)之運算，我們會將三步中的每一步之RSS訊號資料，輪流與整段三步的RSS訊號做互相關性計算，隨後再將三組數據做平均獲得一組最終的NCC序列(圖五)。



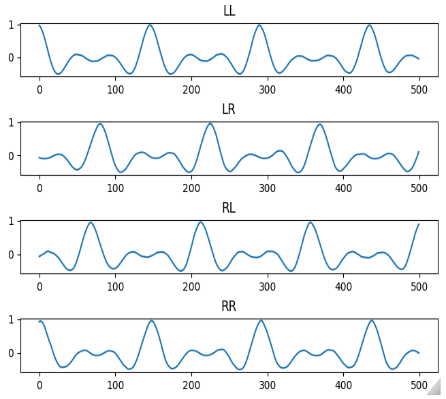
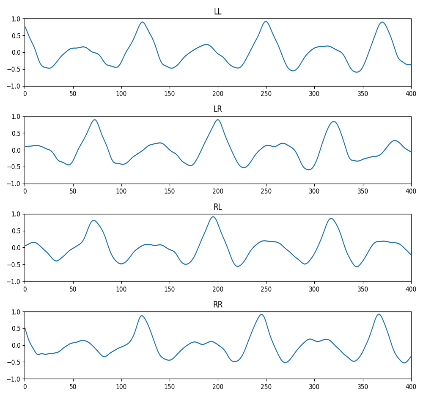
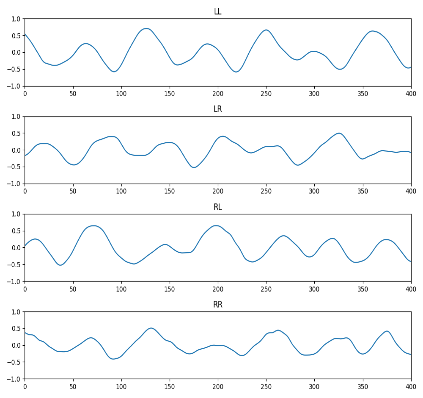
圖五 交叉互相關運算所產生之相關性序列示意圖

LL、LR、RR 和 RL 是四種不同的互相關運算所產生之相關性序列，每種序列是由三個不同長度的單一步態週期訊號序列p 用於計算 所產生之相關性序列。因此，圖六中的每一個 LL、LR、RR 和 RL 序列都是來自每種類型的三個不同長度 p 序列的平均序列。

(a)

(b)

(c)



圖六 (a)老年易跌倒者、(b)老年非跌倒者、(c)健康成年人

1. 步態訊號特徵參數之設計

為了研究和分析易跌倒老年人、不易跌倒老年人和健康成年人等三類實驗群組的行走步態之間的相關性和差異性，我們提出將有辨識分類潛力的步態訊號分成兩類步態特徵。首先，定義這三實驗群組的基本步態訊號特徵，成為第一類的步態特徵。接著，建立經由自相關/互相關運算等步態比較之相關性序列的統計參數，成為第二類的步態特徵。由圖六可看出3種類的受試者在左/右腳的自相關性(LL、RR)，以及左對右(LR)/右對左(RL)互相關性上，有明顯的差異。透過這二類的步態特徵，我們將比較在不同實驗群組之間，行走時左/右腳步態對比上的相關性變化。

為了分析第一類步態特徵，我們使用自相關/互相關數列運算的統計特性，我們選擇了幾個特定的訊號序列集合，其中包含了LL、RR、LR 和 RL 等，來進行深入的分析。為了量化這些數列的特徵，我們分別計算了各個集合的平均值(Mean: μ)、標準差(Standard Deviation: σ)和變化程度(Degree of Variation: Var)。前兩個參數幫助我們理解步態資料集的中心趨勢和數據核心之散佈程度，而變化程度則反映了步態資料集之極限範圍的分散程度。透過這些步態訊號之統計特徵的計算，可使我們更能夠全面地理解這三類實驗群組之步態訊號的特性和差異，從而對其進行更有效的分析和解讀。

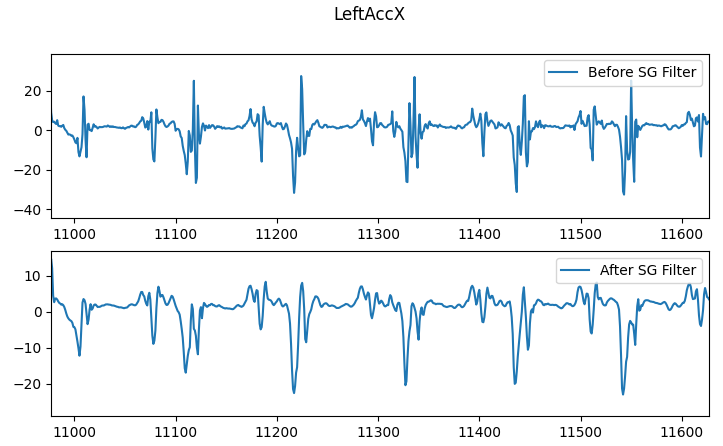
在第二類步態特徵中，被選用的參數包括步頻(Cadence)、單一步態週期的長度(Single-Stride Gait Cycle Interval: GCI)、以及連續3個步態週期的總長度(3-Stride Gait Cycle Period: GCP)。Cadence、GCI、GCP的計算公式，分別顯示於公式三、公式四和公式五中。

|  |  |  |
| --- | --- | --- |
|  |  | 公式三 |
|  |  | 公式四 |
|  |  | 公式五 |

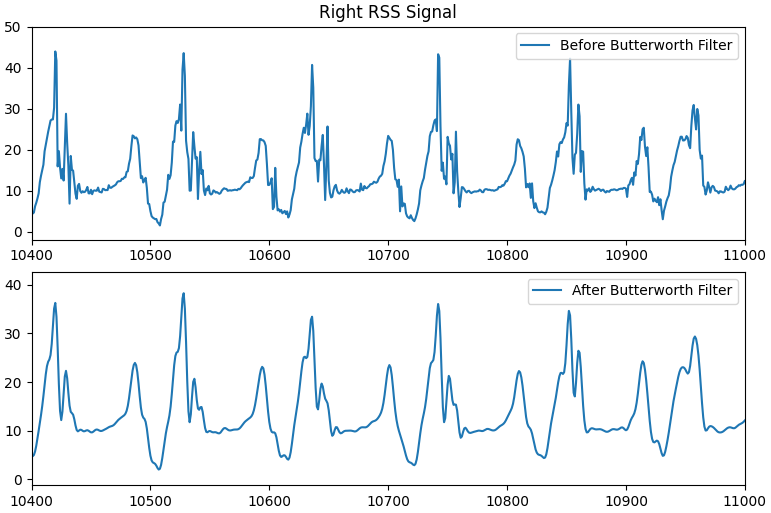
其中，*Num\_Gait*為步態週期的個數，*Gait\_Time*為完成*Num\_Gait*個步態週期的總花費時間，而GCI(*i*)公式中的為x軸之訊號過零點前的第i個波谷(最低點)的時間點。在分析步態時，步頻提供了步行速度的相關訊息。而單一步態週期長度(GCI(*i*))和連續三個步態週期長度(GCP(*i*))，則可提供關於步態動作在瞬時穩定性和長時穩定性等兩個動態特徵上的具體比較。這些數據不僅反映了行走的節奏，還能揭示潛在的步態不規則性。

1. 研究方法
2. 訊號前處理

為了精確提取原始步態訊號中的辨識特徵，通常必須先對原始訊號進行訊號前處理作業，以去除雜訊的干擾。而雜訊的可能來源，包括來自慣性感測器內部電路操作所產生的電器雜訊、穿戴方式不良所造成的量測雜訊、行走動作與地面交互作用所產生的晃動雜訊、或是外在地面環境因素所產生的步態雜訊等。本研究選用 Savitzky-Golay (SG)濾波器[16]和巴特沃斯濾波器(Butterworth filter)[17]，來進行慣性感測器三軸的加速度訊號及RSS加速度訊號等的訊號前處理工作。針對以上所述之濾波器，我們將採用以下的參數設置：SG濾波器的窗口大小設為 9，多項式階數設為 3，以提升處理後訊號的平滑度和保留其訊號細節(圖七)。另外，巴特沃斯濾波器的截止頻率設為15Hz，濾波器階數設為4，可以有效地去除高頻噪聲而不會過度影響訊號的有用成分(圖八)。這樣的設置能夠在保證訊號質量的同時，有效地對X/Y/Z軸的加速度訊號及其RSS加速度值等，進行訊號前處理的工作。



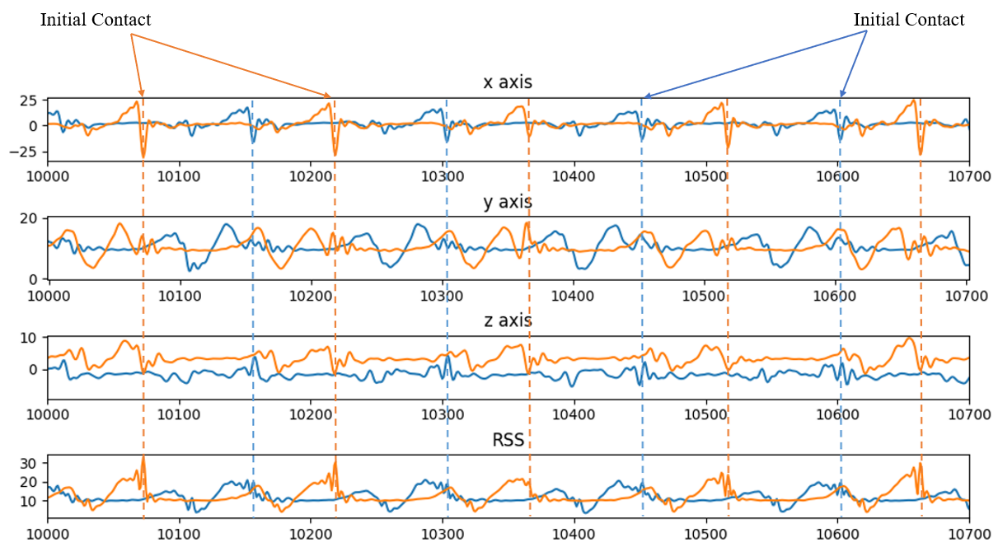
圖七 加速度X軸訊號於SG濾波前/後



圖八 右腳RSS訊號於Butterworth濾波前/後

1. 步態週期切割

我們透過穿戴式慣性感測器來分別量測左腳踝(LA)/右腳踝(RA)位置，在三維空間中的加速度大小的變化[8]。在圖九中，我們特別關注IMU加速度訊號在x軸上，在過零點前的最低值(波谷)來辨識腳跟觸地事件(Initial Contact)，這在步態週期中是至關重要的一個動作[8]。這些數據點對於揭示步態動作的動態變化具有重大意義，因為腳跟觸地為單一步態週期訊號分割的關鍵。



圖九 經過濾波和切割後的步態訊號數據(橘色:右腳;藍:左腳)

1. 步態特徵提取

為了評估各個步態特徵參數在智慧辨識應用上的效能差異，我們將使用無母數統計分析方法中的 Mann-Whitney U檢定，針對本研究的三類實驗群組的步態加速度訊號，進行 Mann-Whitney U檢定，以獲取代表該步態特徵參數在三類實驗群組的智慧辨識應用中，是否可成為提升辨識的顯著性指標。Mann-Whitney U檢定是一個強大的無母數統計工具，用於比較兩個獨立樣本之間的分布差異。Mann-Whitney U檢定不需要假設數據符合常態分布，因此在處理非常態分布數據或具有異常值的數據時尤為有用。

在進行Mann-Whitney U檢定時，我們會先設定兩個假設：一是虛無假設，它認為兩個樣本組的分布沒有顯著差異；二是對立假設，指出兩個樣本組的分布存在顯著差異。透過樣本資料的分析，我們可以判斷虛無假設是否成立。如果計算出的p值小於0.05，我們有足夠的證據拒絕虛無假設，接受對立假設，這表示兩個樣本組在統計上顯著不同。反之，如果p值大於0.05，則沒有足夠證據拒絕虛無假設，我們將認為兩個樣本組間沒有顯著的統計差異。

為了在本研究中比較三類實驗群組，我們將使用逐對比較的方法進行多次Mann-Whitney U檢定。我們會進行以下比較：

1. 易跌倒老年人(Faller: F) vs. 老年非跌倒者(Non-faller: NF)
2. 易跌倒老年人(Faller: F) vs. 健康成年人(Healthy Adult: HA)
3. 健康成年人(Healthy Adult: HA) vs. 老年非跌倒者(Non-faller: NF)

Mann-Whitney U檢定的結果將為我們提供每對比較的檢驗結果，從而可以確定在不同組別之間是否存在著顯著性的統計差異。表一至表六中顯示了步態特徵參數的各項統計值，以及進行Mann-Whitney U檢定後所得的p值檢定結果。

表一 F、NF之自相關/互相關序列參數Mann-Whitney U檢定結果

|  |  |  |  |
| --- | --- | --- | --- |
| **Variables** | **F** | **NF** | ***p*-Value** |
| RR-Mean | 0.0007±0.0083 0.00031(-0.00612-0.00547) | -0.0068±0.0031  -0.00736(-0.00861--0.00394) | 0.843 |
| RR-Var | 0.0229±0.0263 0.01902(0.00325-0.047) | 0.1184±0.043  0.11638(0.07637-0.15046) | \* |
| RR-Std | 0.1487±0.0625 0.14554(0.08829-0.2031) | 0.1178±0.0353 0.10084(0.09814-0.13063) | 0.911 |
| LL-Mean | -0.001±0.0066 0.00086(-0.00515-0.00409) | -0.0078±0.0067 -0.00870(-0.01327--0.00136) | \* |
| LL-Var | 0.0412±0.0331 0.0458(0.01386-0.06606) | 0.1231±0.034 0.12067(0.09907-0.12913) | 0.87 |
| LL-Std | 0.5534±1.3444 0.16598(0.0802-0.19697) | 0.088±0.1557 0.10653(0.09142-0.16087) | \* |
| RL-Mean | -0.0001±0.0092 -0.00155(-0.00674-0.00506) | -0.0058±0.0059 -0.00450(-0.00810--0.00190) | \* |
| RL-Var | 0.0305±0.0292 0.03896(0.00502-0.05133) | 0.1126±0.0343 0.11363(0.08485-0.12867) | \* |
| RL-Std | 0.1426±0.069 0.15375(0.07451-0.1835) | 0.1214±0.0481 0.11308(0.09302-0.12419) | \* |
| LR-Mean | -0.0003±0.0069 0.00169(-0.00296-0.0047) | -0.008±0.0073 -0.00921(-0.01409--0.00470) | \* |
| LR-Var | 0.0255±0.0232  0.02744 (0.0089-0.03689) | 0.1185±0.0407 0.11301(0.08994-0.15682) | 0.575 |
| LR-Std | 0.1491±0.0611 0.14816(0.08817-0.21039) | 0.1191±0.0311 0.11747(0.10138-0.12875) | \* |

*Mann-Whitney U-Test* ;\*表示p值小於0.05

表二 F、HA之自相關/互相關序列參數Mann-Whitney U檢定結果

|  |  |  |  |
| --- | --- | --- | --- |
| **Variables** | **F** | **HA** | ***p*-value** |
| RR-Mean | 0.0007±0.0083 0.00031(-0.00612-0.00547) | 0.0085±0.0087 0.00973(0.00208-0.01264) | 0.003\* |
| RR-Var | 0.0229±0.0263 0.01902(0.00325-0.047) | 0.1722±0.0178 0.17554(0.16300-0.18029) | \* |
| RR-Std | 0.1487±0.0625 0.14554(0.08829-0.2031) | 0.4145±0.0217 0.41898(0.40373-0.42461) | \* |
| LL-Mean | -0.001±0.0066 0.00086(-0.00515-0.00409) | 0.0053±0.0083 0.00415(0.00149-0.01117) | 0.014 |
| LL-Var | 0.0412±0.0331 0.0458(0.01386-0.06606) | 0.167±0.023 0.16945(0.15622-0.18242) | \* |
| LL-Std | 0.5534±1.3444 0.16598(0.0802-0.19697) | 0.4076±0.0296 0.41164(0.39524-0.42710) | \* |
| RL-Mean | -0.0001±0.0092 -0.00155(-0.00674-0.00506) | 0.0064±0.0092 0.00614(-0.00023-0.01151) | 0.022 |
| RL-Var | 0.0305±0.0292 0.03896(0.00502-0.05133) | 0.1616±0.0213 0.16271(0.15144-0.18038) | \* |
| RL-Std | 0.1426±0.069 0.15375(0.07451-0.1835) | 0.4011±0.0274 0.40337(0.38913-0.42471) | \* |
| LR-Mean | -0.0003±0.0069 0.00169(-0.00296-0.0047) | 0.0066±0.0083 0.00878(0.00449-0.01078) | 0.002\* |
| LR-Var | 0.0255±0.0232  0.02744 (0.0089-0.03689) | 0.1726±0.0161 0.17932(0.16024-0.18314) | \* |
| LR-Std | 0.1491±0.0611 0.14816(0.08817-0.21039) | 0.415±0.0197 0.42346(0.40030-0.42795) | \* |

*Mann-Whitney U-Test* ;\*表示p值小於0.05

表三 NF、HA之自相關/互相關序列參數Mann-Whitney U檢定結果

|  |  |  |  |
| --- | --- | --- | --- |
| **Variables** | **NF** | **HA** | ***p*-value** |
| RR-Mean | -0.0068±0.0031  -0.00736(-0.00861--0.00394) | 0.0085±0.0087 0.00973(0.00208-0.01264) | \* |
| RR-Var | 0.1184±0.043  0.11638(0.07637-0.15046) | 0.1722±0.0178 0.17554(0.16300-0.18029) | \* |
| RR-Std | 0.1178±0.0353 0.10084(0.09814-0.13063) | 0.4145±0.0217 0.41898(0.40373-0.42461) | 0.029\* |
| LL-Mean | -0.0078±0.0067 -0.00870(-0.01327--0.00136) | 0.0053±0.0083 0.00415(0.00149-0.01117) | 0.008\* |
| LL-Var | 0.1231±0.034 0.12067(0.09907-0.12913) | 0.167±0.023 0.16945(0.15622-0.18242) | 0.123 |
| LL-Std | 0.088±0.1557 0.10653(0.09142-0.16087) | 0.4076±0.0296 0.41164(0.39524-0.42710) | 0.008\* |
| RL-Mean | -0.0058±0.0059 -0.00450(-0.00810--0.00190) | 0.0064±0.0092 0.00614(-0.00023-0.01151) | 0.154 |
| RL-Var | 0.1126±0.0343 0.11363(0.08485-0.12867) | 0.1616±0.0213 0.16271(0.15144-0.18038) | \* |
| RL-Std | 0.1214±0.0481 0.11308(0.09302-0.12419) | 0.4011±0.0274 0.40337(0.38913-0.42471) | \* |
| LR-Mean | -0.008±0.0073 -0.00921(-0.01409--0.00470) | 0.0066±0.0083 0.00878(0.00449-0.01078) | \* |
| LR-Var | 0.1185±0.0407 0.11301(0.08994-0.15682) | 0.1726±0.0161 0.17932(0.16024-0.18314) | \* |
| LR-Std | 0.1191±0.0311 0.11747(0.10138-0.12875) | 0.415±0.0197 0.42346(0.40030-0.42795) | 0.03\* |

*Mann-Whitney U-Test* ;\*表示p值小於0.05

表四 F、NF之傳統步態參數Mann-Whitney U檢定結果

|  |  |  |  |
| --- | --- | --- | --- |
| **Variables** | **F** | **NF** | **p-value** |
| Cadence | 33.53±15.43 27.49(22.59-38.32) | 46.66±9.24 49.63(36.8-55.53) | 0.024\* |
| GCI | 88.17±35.31 74.15(60-114.81) | 111.23±24.93 103(98-121) | 0.048\* |
| GCP | 286.93±129.67 221.5(180-364.75) | 333.61±33.86 320(317-373) | 0.119 |

*Mann-Whitney U-Test* ;\*表示p值小於0.05

表五 F、HA之傳統步態參數Mann-Whitney U檢定結果

|  |  |  |  |
| --- | --- | --- | --- |
| **Variables** | **F** | **HA** | **p-value** |
| Cadence | 33.54±15.44  27.49(22.59-38.32) | 52.87±5.8271 54.18(51.1-55.75) | 0.001\* |
| GCI | 88.18±35.32  74.15(60-114.81) | 108.72±6.67 107.5(105-111) | 0.027\* |
| GCP | 286.94±129.68  221.5(180-364.75) | 326.31±21.18 321(314-338) | 0.081 |

*Mann-Whitney U-Test* ;\*表示p值小於0.05

表六 NF、HA之傳統步態參數Mann-Whitney U檢定結果

|  |  |  |  |
| --- | --- | --- | --- |
| **Variables** | **NF** | **HA** | **p-value** |
| Cadence | 46.66±9.24 49.63(36.8-55.53) | 52.87±5.8271 54.18(51.1-55.75) | 0.167 |
| GCI | 111.23±24.93 103(98-121) | 108.72±6.67 107.5(105-111) | 0.329 |
| GCP | 333.61±33.86 320(317-373) | 326.31±21.18 321(314-338) | 0.527 |

*Mann-Whitney U-Test* ;\*表示p值小於0.05

1. 窮舉法之特徵分析

窮舉法，是通過嘗試所有可能的解決方案，來找到最佳參數組合的方法。使用窮舉法的原因在於其簡單直觀，不需要複雜的數學推導或特別的算法設計，只需列舉出所有可能的解，並檢查每個解是否符合要求。窮舉法的全面性，保證了能夠找到問題的所有可能解，這在需要全面搜索的問題中尤為重要。它的基本思想是遍歷所有的可能性，以確保每個可能的解決方案儘可能的都被納入評估，特別是在問題規模較小或計算資源充足的情況下。然而，窮舉法的一個主要缺點是其計算量可能會非常龐大，特別是當問題規模較大時。例如，當涉及到大量變量或組合時，窮舉法可能需要非常長的時間來計算所有可能的解決方案。因此，在實際應用中，需要根據具體情況評估是否採用窮舉法。

在本研究中，我們透過窮舉法將全部15個步態特徵做排列組合，逐一分析每種組合的辨識效果，從中找出辨識率最優的特徵組合。具體過程包括以下幾個步驟：

1. 特徵提取：從步態數據中提取15個潛在的特徵，這些特徵包括步態平均值、標準差、變異性、步頻、單一步態週期長度、3個步態週期長度等。
2. 特徵組合生成：生成所有可能的特徵組合，確保每個組合包含至少兩個或更多特徵。
3. 效能評估：對每個特徵組合進行逐一分析，使用機器學習模型來評估其在三類實驗群組中的辨識效果。這包括計算每個特徵組合在不同組別(健康成年人、老年非跌倒者、易跌倒老年人)之間的分類準確率。
4. 最佳特徵組合選擇：根據分析結果，選擇辨識率最優的特徵組合，作為本研究的最終特徵集。

由於這種方法需要對大量的特徵組合進行計算和分析，我們使用的電腦配備如下：

表七 電腦配備

|  |  |
| --- | --- |
| **名稱** | **型號** |
| Operating System | Windows 10 22H2 |
| CPU | Intel Core i5-10400 |
| GPU | NVIDIA GeForce RTX2080 |
| RAM | 40GB |
| Storage | 931G |
| Programming Language | Python 3.7.9 |

1. 步態種類之分類

在本研究中，我們將採用三種不同的機器學習模型，包括K-Nearest Neighbors (KNN)、Support Vector Machine (SVM) 和 Random Forest (RF)，來進行跌倒步態的分類。透過比較這三種模型所產生的效能，我們期望能夠找出最適合於跌倒步態分類的模型。

1. KNN

K-Nearest Neighbors (KNN) 是一種基本且廣泛應用的監督式機器學習演算法，適用於分類和回歸問題。KNN 演算法的核心思想是通過分析一個樣本的鄰近樣本來進行預測。在特徵空間中，對於一個未知類別的樣本點，我們可以通過測量其與已知類別樣本點之間的距離，來確定其最近的 K 個鄰居。這些最近鄰居的類別標籤可被用來對未知樣本進行類別預測。具體來說，對於分類問題，預測結果是由最近鄰居中最頻繁出現的類別決定的；對於回歸問題，預測結果是由最近鄰居的目標值的平均值或加權平均值決定的。

KNN演算法中最常用的距離計算公式是歐幾里德距離(Euclidean distance)，用於計算兩點之間的直線距離。 對於兩個點 和 ，它們在n維空間中的歐幾里得距離公式為：

|  |  |  |
| --- | --- | --- |
|  |  | 公式六 |

這種基於局部信息進行預測的方法，使得 KNN 演算法能夠適應不同的數據分佈，並且對於非線性問題也有良好的表現。

1. SVM

支援向量機(Support Vector Machine, SVM)是一種在監督學習和模式識別領域中非常流行的機器學習演算法。這種演算法的一大優點是能夠從有限的訓練數據中提取重要的統計特性，同時擅長處理高維度數據集和非線性問題。作為一個線性分類器，SVM 的主要目的是在特徵空間中找到一個最佳的超平面，這個超平面能夠最大化不同類別數據點之間的間隔，從而確保分類的準確性。這使得 SVM 在多個領域，包括文字分類、圖像識別以及生物醫學分類等，透過訓練過程，SVM 能夠識別出這樣的超平面，其方程式可以表示為：

|  |  |  |
| --- | --- | --- |
|  |  | 公式七 |

其中為法向量，用來表示超平面的方向，為位移項，用來表示超平面與原點之間的距離，確定了超平面以後，SVM通過在超平面的兩側各自擴展出一條邊界線來建立分類決策界線。這些邊界線由以下方程式定義：

|  |  |  |
| --- | --- | --- |
|  |  | 公式八 |

這些邊界線與超平面共同定義了一個區域，稱爲“間隔”，它幫助我們區分不同的數據類別。在訓練過程中，SVM 的目標是找到一个最大化这個間隔的超平面， 從而提高分類的準確性，具體的分類規則如下列所示：

|  |  |  |
| --- | --- | --- |
|  |  | 公式九 |

此外，SVM的效能受到C值和Gamma參數的顯著影響。C值控制了模型對錯誤分類的容忍程度，較高的C值使得模型對分類錯誤容忍度更低，從而能夠獲得更嚴格的分類界限，但也可能導致過度擬合；較低的C值則增加了對分類錯誤的容忍度，使得模型更加重視並保持較大的分類間隔，但這可能會導致一些分類錯誤的存在。

Gamma參數則決定了每個資料點在形成決策界面時的影響範圍，較低的gamma值表示資料點的影響範圍較廣，導致決策邊界更加平滑；較高的gamma值則表示資料點的影響範圍較小，使得決策邊界變得複雜。

支援向量機(SVM)已經成為近年在各種實驗應用中頗受歡迎的一種演算法，不僅因為它在處理分類問題上的高效表現，也因為它在回歸分析和異常值檢測等多方面的可靠性。其多樣性和出色的效能使 SVM 顯示出強大的應用潛力。

1. RF(Random forest)

隨機森林(Random Forest, RF)是一種集成學習演算法，廣泛應用於數據科學領域中的分類和迴歸問題，該演算法由多棵決策樹組成，通過結合這些決策樹的預測結果來提升模型的準確性。在隨機森林演算法中，決策樹選擇的方式為CART(Classification and Regression Tree)樹，訓練過程包括兩個主要的隨機性引入方式：自助聚集(Bootstrap aggregating，或稱bagging)和特徵隨機選擇。首先，它通過自助聚集的方式從原始訓練數據集中重複抽樣選取多個子樣本來訓練多個決策樹;其次，在每個決策樹的分裂過程中，隨機森林不是考慮所有特徵，而是隨機選擇一個特徵子集來尋找最佳分裂特徵，這種隨機性的引入不僅增加了模型的多樣性，也防止模型過度擬合。

對於分類問題，CART樹使用基尼不純度(Gini Impurity)作為衡量分割的標準，基尼不純度是利用基尼系數來衡量一個節點中類別的混亂程度，基尼不純度越小，數據的純度越高，公式如下:

|  |  |  |
| --- | --- | --- |
|  |  | 公式十 |

|  |  |  |
| --- | --- | --- |
|  |  | 公式十一 |

其中是指在集合x中，屬於第i類的樣本比例，n為類別總數。而對於迴歸問題，則是採用二元分割的方式，並計算分割後各節點的加權基尼不純度，並求其和，以找到最佳分割點：

其中和為分割後的左右兩個節點，N為分割前的總樣本數。隨機森林因高效性、易於實現、對於大量數據和高維度數據的適應性強等優點，在機器學習領域得到了廣泛的應用，它既可以處理類別數據，也可以處理連續數據，且對於缺失數據有較好的容忍性，但隨機森林也有一定的局限性，例如當樹太多時，模型的訓練和預測速度會變慢，且模型解釋性較差，特別是在需要理解模型決策過程的應用場景中。

1. 結果

本節展示將分成兩部分，第一部分，對51名實驗參與者(F: 16、NF: 13、HA: 22)的步態特徵資料進行Mann-Whitney U檢定(使用JASP分析)，將有顯著性的特徵挑選出來使用於智慧辨識。第二部分，將採用由窮舉法分析所獲得的最佳特徵組合，使用於智慧辨識。接著，我們將這些顯著性特徵或是最佳特徵組合，應用於三種監督式機器學習演算法：K近鄰(KNN)、支持向量機(SVM)和隨機森林(RF)，以進行智慧分類和預測。

為了增加模型預測結果，我們選擇使用 K-fold 交叉驗證， K-fold 交叉驗證是一種在機器學習和統計建模中常用的模型評估方法，用於評估模型的泛化能力， 這種方法主要是將原始資料分成 K 個大小相等的子集，在這 K 個子集中，每個子集輪流作為驗證集來評估模型，而剩下的 K-1 個子集則作為訓練集來訓練模型， 整個過程會重複K次，每次都將選擇不同的子集作為驗證集，而其他子集作為訓練集。最後，將這 K 次評估的結果平均，作為模型表現的估計。在本研究中，我們選用了 5 作為 K-fold 交叉驗證的 K 值，這是進行 K-fold 交叉驗證時較為普遍的選擇。

後續將詳細介紹分析結果，其中包含混淆矩陣(表八)和分類器效能評估之標準，例如：精確度(Precision)、準確率(Accuracy)、靈敏度(Sensitivity)、特異度(Specificity)和F1-score(公式十二至公式十六)。

表八 混淆矩陣示意圖

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | 預測值 | | |
| F | NF | HA |
| 實際值 | F | TP | FN | FN |
| NF | FP | TP | FN |
| HA | FP | FP | TP |

F：易跌倒老年人；NF：不易跌倒老年人；HA：健康成年人

|  |  |  |
| --- | --- | --- |
|  |  | 公式十二 |
|  |  | 公式十三 |
|  |  | 公式十四 |
|  |  | 公式十五 |
|  |  | 公式十六 |

1. Mann-Whitney U檢定：特徵之顯著性分析

我們將利用Mann-Whitney U檢定篩選出的顯著性特徵(p值小於0.05)，進行K近鄰(KNN)、支持向量機(SVM)和隨機森林(RF)分類預測。每種演算法皆針對以下四組分類進行分析：

1. 易跌倒老年人(Faller: F) vs. 老年非跌倒者(Non-faller: NF)
2. 易跌倒老年人(Faller: F) vs. 健康成年人(Healthy Adult: HA)
3. 老年非跌倒者(Non-faller: NF) vs. 健康成年人(Healthy Adult: HA)
4. 易跌倒老年人(Faller: F) vs. 老年非跌倒者(Non-faller: NF)   
   vs. 健康成年人(Healthy Adult: HA)

第一組使用的顯著性特徵為:

Cadence、GCI、RR-Var、LL-Mean、LL-Std、RL-Mean、RL-Std、LR-Mean、LR-Std、RL-Var

第二組使用的顯著性特徵為:

RR-Mean、RR-Var、RR-Std、LL-Mean、LL-Var、LL-Std、RL-Mean、RL-Std、RL-Var、LR-Var、LR-Mean、LR-Std、Cadence、GCI

第三組使用的顯著性特徵為:

RR-Mean、RR-Var、RR-Std、LL-Mean、LL-Std、RL-Var、RL-Std、LR-Var、LR-Mean、LR-Std

第四組使用的顯著性特徵為:

RR-Var、LL-Mean、LL-Std、RL-Var、RL-Std、LR-Mean、LR-Std

1. Mann-Whitney U檢定：KNN

表九展示了使用K近鄰(KNN)分類演算法針對F(易跌倒老年人)和NF(老年非跌倒者)兩類群體，採用顯著性參數進行分類後的混淆矩陣。混淆矩陣中的數據顯示，True lable為F的樣本中，有10個被正確分類為F，6個被錯誤分類為NF；而True lable為NF的樣本中，全部13個樣本被正確分類為NF。這表明KNN分類器在此情境下對NF的辨識效果較好，而對F的辨識存在一定誤差。

表九 F&NF於KNN之顯著性參數混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | F | 10 | 6 |
| NF | 0 | 13 |
|  |  | | F | NF |
|  |  | | Predicted label | |

表十總結了使用顯著性參數進行KNN分類後的各項效能指標。結果顯示，儘管KNN分類器在各項指標上均達到一定水準，但仍有提升空間，尤其是在對F(易跌倒老年人)的分類中，分類器的精確度和召回率雖然較高，但混淆矩陣顯示的錯誤分類情況表明，對於跌倒風險的精確預測尚需進一步優化。

表十 F&NF顯著性參數之KNN分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.793 | 0.842 | 0.812 | 0.842 | 0.79 |

表十一顯示了F和HA兩類別在KNN分類器下的混淆矩陣，從混淆矩陣中可以看出，對於易跌倒老年人類別(F)，KNN分類器正確分類了11個樣本，但錯誤分類了5個樣本為HA。對於健康成年人類別(HA)，KNN分類器的分類效果較好，22個樣本均被正確分類，沒有任何錯誤分類。

表十一 F&HA於KNN之顯著性參數混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | F | 11 | 5 |
| HA | 0 | 22 |
|  |  | | F | HA |
|  |  | | Predicted label | |

藉由表十二這些效能指標值可以看出，使用統計分析方法進行篩選和排序後，分類器的辨識效能高達85%以上。這表明，應用自相關和互相關運算於步態加速度訊號，並通過統計分析方法篩選和排序特徵參數，能夠非常有效地識別具有高度跌倒風險的老年人群體。

表十二 F&HA顯著性參數之KNN分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.868 | 0.907 | 0.843 | 0.907 | 0.856 |

表十三展示了NF和HA兩類別在KNN分類器下的混淆矩陣。從混淆矩陣中可以看出，對於老年人不易跌倒類別(NF)，KNN分類器正確分類了12個樣本，錯誤分類了1個樣本為健康成年人(HA)，對於健康成年人類別(HA)，KNN分類器的分類效果較好，正確分類了21個樣本，錯誤分類了1個樣本為老年不易跌倒類別(NF)。

表十三 NF&HA於KNN之顯著性參數混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | NF | 12 | 1 |
| HA | 1 | 21 |
|  |  | | NF | HA |
|  |  | | Predicted label | |

從表十四可以看出，KNN分類器在這組顯著性參數的應用中，我們發現模型在各項效能評估指標上均表現理想，準確率達到94.2%，這表明我們篩選出的顯著性參數集合，能夠準確地識別出具有高跌倒風險的老年人群體。

表十四 NF&HA顯著性參數之KNN分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.942 | 0.954 | 0.954 | 0.923 | 0.954 |

表十五為使用KNN進行分類後的混淆矩陣，在使用KNN進行訓練時，顯著性參數的準確率可達90%。根據表十六中顯示的KNN模型實驗結果，我們發現模型在各項效能評估指標上均表現卓越，準確率達到90.2%，這表明我們篩選出的顯著性參數集合，能夠準確地識別出具有高跌倒風險的老年人群體。

表十五 F&NF&HA於KNN之顯著性參數混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | F | 13 | 1 | 2 |
| NF | 2 | 11 | 0 |
| HA | 0 | 0 | 22 |
|  |  | F | NF | HA |
|  |  | Predicted label | | |

表十六 F&NF&HA顯著性參數之KNN分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.902 | 0.90 | 0.886 | 0.949 | 0.892 |

1. Mann-Whitney U檢定：SVM

由表十七、

表十八可以看出，SVM分類器在使用顯著性參數進行分類時，達到了較高的分類效能。具體來說，分類器對於兩類別的辨識能力均較強，準確率達到82.6%。此外，精確度和召回率均接近或超過85%，這表明分類器在區分老年易跌倒和不易跌倒老年人群體時，具有較高的可靠性和穩定性。這些結果表明，通過使用顯著性參數，SVM分類器能夠有效地識別具有跌倒風險的老年人群體，為相關的跌倒預防措施提供了有力的支持。

表十七 F&NF於SVM之顯著性參數混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | F | 12 | 4 |
| NF | 1 | 12 |
|  |  | | F | NF |
|  |  | | Predicted label | |

表十八 F&NF顯著性參數之SVM分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.826 | 0.881 | 0.858 | 0.858 | 0.822 |

表十九展示了使用支持向量機(SVM)進行分類後的混淆矩陣。從混淆矩陣中可以觀察到，對於易跌倒(F)類別，SVM分類器正確地識別了13個樣本，但錯誤地將3個樣本分類為健康成年人(HA)，對於健康成年人類別(HA)，SVM分類器正確地識別了21個樣本，但錯誤地將1個樣本分類為易跌倒(F)。

表十九 F&HA於SVM之顯著性參數混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | F | 13 | 3 |
| HA | 1 | 21 |
|  |  | | F | HA |
|  |  | | Predicted label | |

從表二十可以看出，SVM分類器在使用顯著性參數進行分類時，達到了非常高的分類效能。具體來說，分類器對於兩類別的辨識能力均強，準確率達到90%。此外，精確度和召回率均接近或超過90%，這表明我們篩選的關鍵參數集合能夠準確地識別出具有高跌倒風險的老年人群。

表二十 F&HA顯著性參數之SVM分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.90 | 0.91 | 0.888 | 0.888 | 0.892 |

表二十一顯示了使用支持向量機(SVM)進行分類後的混淆矩陣。從混淆矩陣可以看到，對於老年不易跌倒類別(NF)，SVM分類器正確地識別了11個樣本，但錯誤地將2個樣本分類為健康成年人(HA)，對於健康成年人類別(HA)，SVM分類器正確地識別了20個樣本，但錯誤地將2個樣本分類為老年不易跌倒(NF)。

表二十一 NF&HA於SVM之顯著性參數混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | NF | 11 | 2 |
| HA | 1 | 21 |
|  |  | | NF | HA |
|  |  | | Predicted label | |

從表二十二發現，SVM分類器在使用顯著性參數進行分類時，達到了不錯的分類效能。準確率達到91.4%。此外，精確度、召回率和F1分數均超過90%，顯示出分類器在區分老年不易跌倒和健康成年人群體時具有很高的準確性。然而，特異性略低，僅為84.6%，這表明分類器在某些情況下仍有錯誤分類的風險

表二十二 NF&HA顯著性參數之SVM分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.914 | 0.913 | 0.954 | 0.846 | 0.932 |

根據表二十三中顯示的SVM模型實驗結果中，在使用SVM進行分類預測時，顯著性參數的準確率有達80.4%。在分類NF時，有許多的分類錯誤。藉由表二十四可以發現，在各項效能評估指標上表現出一定的分類程度，準確率達到80.4%，這表明我們篩選出的顯著性參數集合，能夠一定地識別出具有高跌倒風險的老年人群體，但在識別老年不易跌倒群體的效果仍有進步的空間。

表二十三 F&NF&HA於SVM之顯著性參數混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | F | 14 | 1 | 1 |
| NF | 8 | 5 | 0 |
| HA | 0 | 0 | 22 |
|  |  | F | NF | HA |
|  |  | Predicted label | | |

表二十四 F&NF&HA顯著性參數之SVM分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.804 | 0.809 | 0.753 | 0.904 | 0.747 |

1. Mann-Whitney U檢定：RF

表二十五展示了使用隨機森林(RF)進行分類後的混淆矩陣，從混淆矩陣可以觀察到，對於易跌倒(F)類別，RF分類器正確地識別了13個樣本，但錯誤地將3個樣本分類為不易跌倒(NF)，對於不易跌倒(NF)類別，RF分類器正確地識別了11個樣本，但錯誤地將2個樣本分類為易跌倒(F)。

表二十五 F&NF於RF之顯著性參數混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | F | 13 | 3 |
| NF | 2 | 11 |
|  |  | | F | NF |
|  |  | | Predicted label | |

表二十六表明隨機森林分類器在F和NF分類任務中具有較高的分類性能。雖然精確率略低於其他指標，但整體的準確率、召回率、特異度和F1分數仍然表現出色，顯示出該模型在識別跌倒風險時的有效性。

表二十六 F&NF顯著性參數之RF分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.846 | 0.706 | 0.816 | 0.833 | 0.811 |

表二十七展示了使用隨機森林(RF)進行分類後的混淆矩陣。從混淆矩陣可以觀察到，對於老年易跌倒類別(F)，RF分類器正確地識別了15個樣本，但錯誤地將1個樣本分類為健康成年人(HA)，對於健康成年人(HA)類別，RF分類器正確地識別了22個樣本，沒有誤分類為老年易跌倒(F)的樣本。

表二十七 F&HA於RF之顯著性參數混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | F | 15 | 1 |
| HA | 0 | 22 |
|  |  | | F | HA |
|  |  | | Predicted label | |

由表二十八可以發現，RF分類器在使用顯著性參數進行分類時，達到了非常高的分類效能。準確率達到97.3%。此外，精確度、召回率和特異性均超過95%，這表明分類器在區分老年易跌倒和健康成年人群體時具有極高的可靠性和穩定性。這些結果顯示，通過使用顯著性參數，RF分類器能夠非常有效地識別具有跌倒風險的老年人群體

表二十八 F&HA顯著性參數之RF分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.973 | 0.978 | 0.968 | 0.978 | 0.973 |

表二十九為使用隨機森林(RF)進行分類後的混淆矩陣。從混淆矩陣可以觀察到，對於老年不易跌倒類別(NF)，RF分類器正確地識別了12個樣本，但錯誤地將4個樣本分類為健康成年人(HA)，對於健康成年人(HA)類別，RF分類器正確地識別了21個樣本，但錯誤地將1個樣本分類為老年不易跌倒(NF)。

表二十九 NF&HA於RF之顯著性參數混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | NF | 12 | 1 |
| HA | 0 | 22 |
|  |  | | NF | HA |
|  |  | | Predicted label | |

從表三十可以看出，RF分類器在使用顯著性參數進行分類時，展示了優異的分類效能。具體而言，分類器在兩類別的辨識能力上表現出色，準確率達到97.1%。此外，精確度、召回率和F1分數均超過95%，顯示出分類器在區分老年不易跌倒和健康成年人群體時具有很高的準確性。然而，特異性略低，為92.3%，這表明分類器在某些情況下仍有錯誤分類的風險。

表三十 NF&HA顯著性參數之RF分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.971 | 0.956 | 0.95 | 0.923 | 0.977 |

表三十一 F&NF&HA於RF之顯著性參數混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | F | 15 | 1 | 0 |
| NF | 1 | 12 | 0 |
| HA | 0 | 0 | 22 |
|  |  | F | NF | HA |
|  |  | Predicted label | | |

表三十一的混淆矩陣展示了顯著性參數訓練後的結果，在使用隨機森林（RF）模型進行訓練時，顯著性參數使得模型的準確率達到了96%。從表三十二可以看到，採用顯著性參數的RF模型在各項效能評估指標上都表現出色，尤其是準確率高達96%。這表明，我們篩選出的顯著性參數集能夠精準識別出具有高跌倒風險的老年族群。

表三十二 F&NF&HA顯著性參數之RF分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.96 | 0.953 | 0.953 | 0.981 | 0.953 |

1. 窮舉法分析

每種演算法都有屬於他們各自的最佳特徵組合，我們利用窮舉法分別找出KNN、SVM、RF的最佳特徵組合，透過最佳特徵組合的結果與使用Mann-Whitney U篩選出的組合的結果進行比較，通過這種比較，我們能夠評估那些最佳特徵組合參數對於提升機器學習模型在智慧辨識應用上的預測效能會有多大的影響。其組別如下:

1. 易跌倒老年人(Faller: F) vs. 老年非跌倒者(Non-faller: NF)

KNN : RR-Mean、RR-Std、RR-Var、LL-Mean、LL-Std、LL-Var、RL-Mean、RL-Std、RL-Var、LR-Mean、LR-Std、LR-Var

SVM : RR-Mean、RR-Std、RR-Var、LL-Mean、RL-Mean、RL-Std、RL-Var、LR-Mean、LR-Std、LR-Var

RF : RR-Var、LL-Mean、LL-Std、RL-Mean、RL-Std、RL-Var、LR-Mean、LR-Std、Cadence、GCI、GCP

1. 易跌倒老年人(Faller: F) vs. 健康成年人(Healthy Adult: HA)

KNN: RR-Mean、RR-Std、RR-Var、LL-Mean、LL-Std、RL-Mean、RL-Var、LR-Mean

SVM : RR-Mean、RR-Std、RR-Var、LL-Mean、LL-Var、LL-Std、RL-Mean、RL-Var、RL-Std、LR-Mean、LR-Std

RF : RR-Mean、RR-Var、LL-Mean、LL-Std、RL-Mean、RL-Var、RL-Std、LR-Mean、LR-Std、Cadence、GCI

1. 老年非跌倒者(Non-faller: NF) vs. 健康成年人(Healthy Adult: HA)

KNN : RR-Mean、RR-Var、RR-Std、LL-Mean、LL-Std

SVM : RR-Mean、RR-Var、LL-Mean、LL-Var、LL-Std、RL-Mean

RF : RR-Mean、RR-Var、RR-Std、LL-Mean、LL-Var、LL-Std、RL-Std

1. 易跌倒老年人(Faller: F) vs. 老年非跌倒者(Non-faller: NF)   
   vs. 健康成年人(Healthy Adult: HA)

KNN : RR-Mean、RR-Std、LL-Mean、LL-Var 、RL-Var

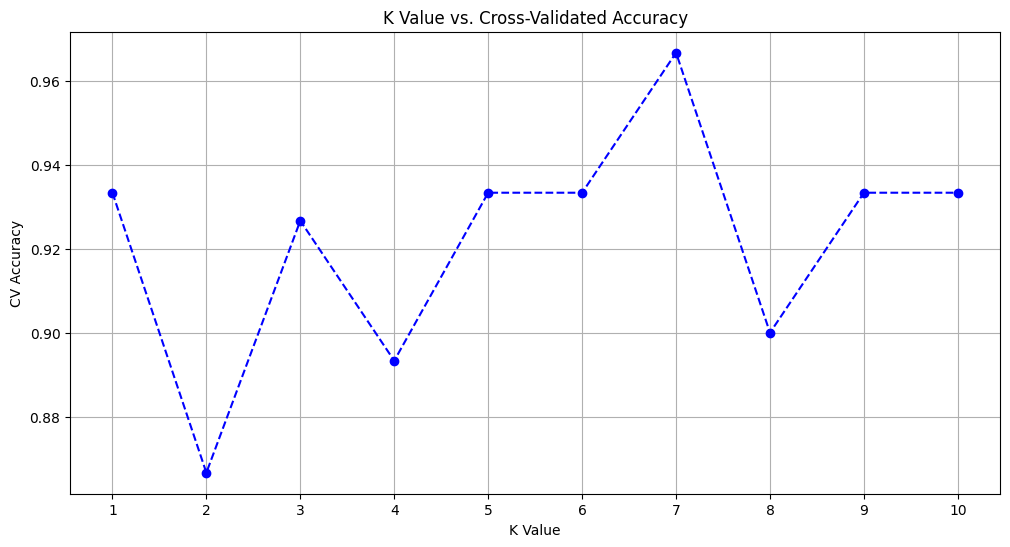
SVM : RR-Mean、RR-Var、RR-Std、LL-Var、RL-Var

RF : RR-Mean、RR-Var、RR-Std、RL-Var、RL-Std

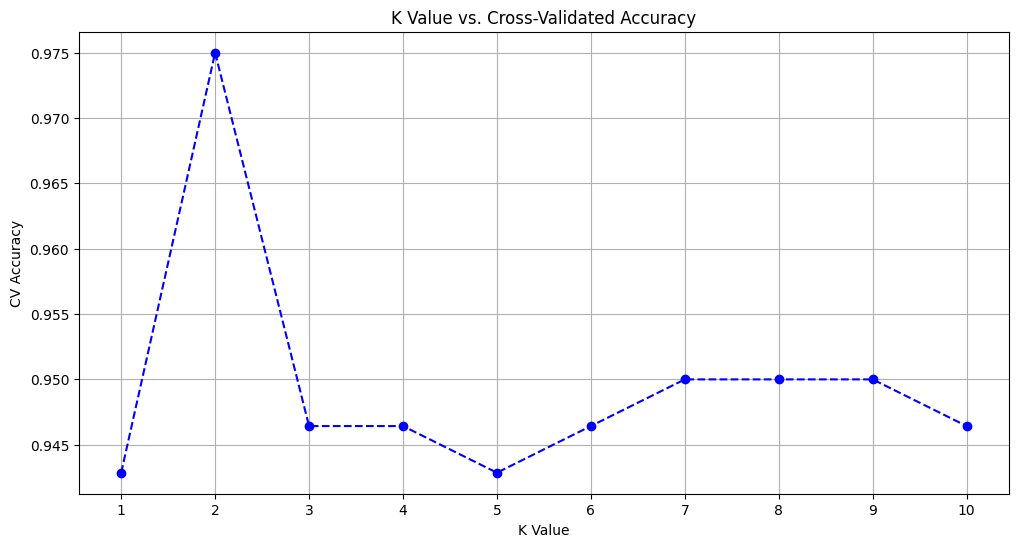
而在分類辨識中，Gait Label值為0則代表該筆數據為易跌倒老年人；Gait Label值為1則代表該筆數據為不易跌倒老年人；Gait Label值為2代表該筆數據為健康成年人。

1. KNN分析結果

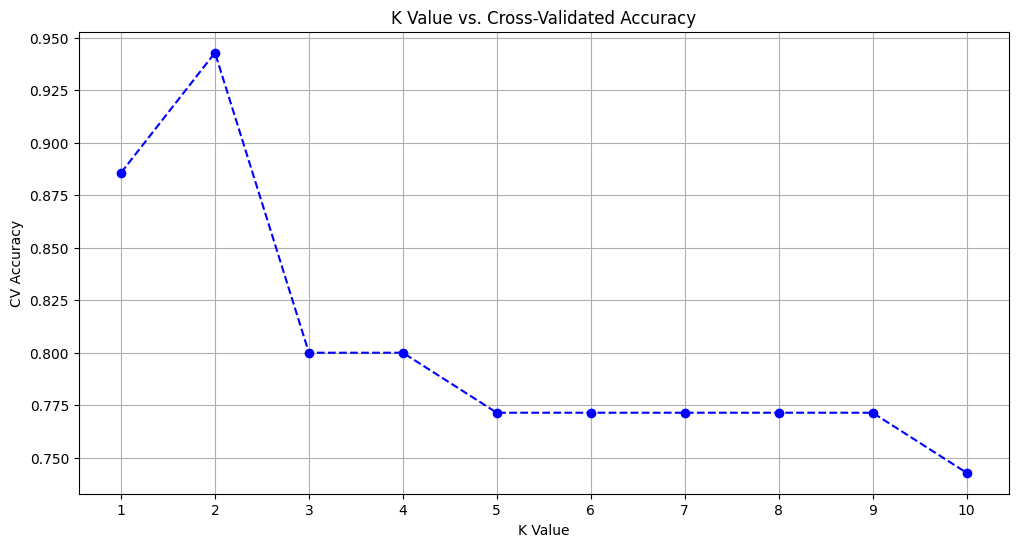
在本研究中，鑑於k值在KNN演算法中扮演著至關重要的角色，並且能夠顯著影響模型訓練的表現，我們對實驗資料集進行了一系列不同K值的測試和比較。橫軸為不同的k值，縱軸表示使用該k值分析的準確率，觀察圖十至圖十三可以發現在分類以上4個組別時，當 k 值分別等於7、2、9，我們獲得了最高的準確率。



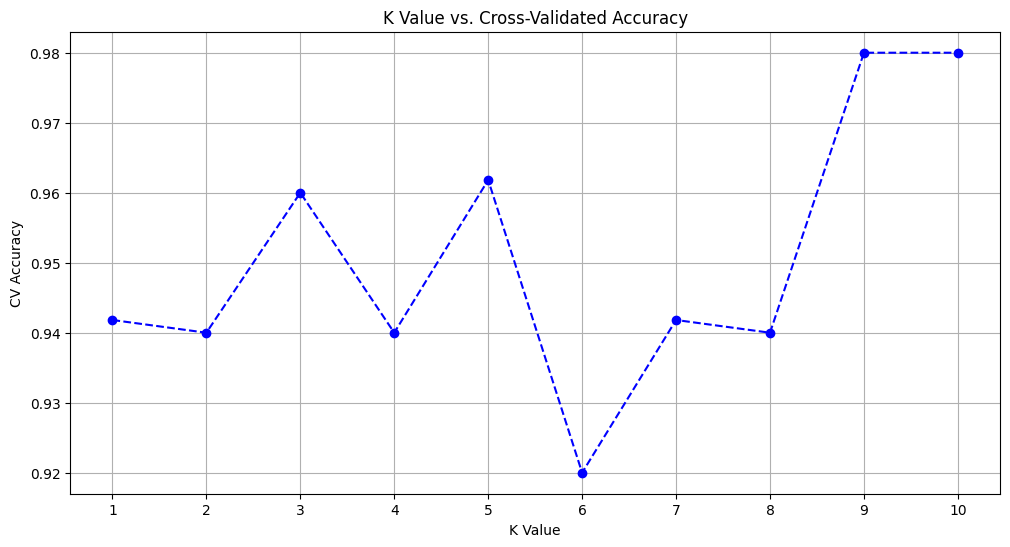
圖十 F&NF不同k值之效能圖



圖十一 F&HA不同k值之效能圖



圖十二 NF&HA不同k值之效能圖



圖十三 F&NF&HA不同k值之效能圖

因此，本研究旨在透過比較不同K值下的模型表現，找出最優化K值，以確保模型達到最佳的表現。這個過程不僅揭示了K值選擇對模型精確度的影響，也為如何調整KNN演算法中的關鍵參數提供了實驗依據。

如表三十三至表四十所示，皆展示了最佳特徵組合使用KNN訓練後的結果，得知在使用KNN訓練顯著性參數後其準確率可達或趨近於95%。

表三十三 F&NF於KNN之最佳特徵組合混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | F | 15 | 1 |
| NF | 0 | 13 |
|  |  | | F | NF |
|  |  | | Predicted label | |

表三十四 F&NF之KNN最佳特徵組合分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.965 | 0.964 | 0.968 | 0.964 | 0.965 |

表三十五 F&HA於KNN之最佳特徵組合混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | F | 15 | 1 |
| HA | 0 | 22 |
|  |  | | F | HA |
|  |  | | Predicted label | |

表三十六 F&HA之KNN最佳特徵組合分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.973 | 0.978 | 0.968 | 0.978 | 0.972 |

表三十七 NF&HA於KNN之最佳特徵組合混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | NF | 12 | 1 |
| HA | 1 | 22 |
|  |  | | NF | HA |
|  |  | | Predicted label | |

表三十八 NF&HA之KNN最佳特徵組合分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.942 | 0.954 | 0.954 | 0.923 | 0.954 |

表三十九 F&NF&HA於KNN之最佳特徵組合混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | F | 16 | 0 | 0 |
| NF | 1 | 12 | 0 |
| HA | 0 | 0 | 22 |
|  |  | F | NF | HA |
|  |  | Predicted label | | |

表四十 F&NF&HA之KNN最佳特徵組合分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.98 | 0.98 | 0.974 | 0.98 | 0.976 |

表三十四、表三十六、表三十八、表四十展示了最佳特徵組合於KNN辨識實驗後的各種效能評估指標，我們可以發現所有的效能指標值皆高過95%，這顯示出了我們所篩選出的顯著性參數集合可以很精準地篩選出有高度跌倒風險的老人群體。由此可見，使用自相關/互相關運算於步態加速度訊號之上，並透過統計分析方法或是窮舉法進行特徵參數之篩選和排序，對於有高度潛在跌倒風險的老人群體，皆可以進行非常有效的辨識。

1. SVM分析結果

SVM是一種流行的機器學習方法，它透過使用核函數來實現非線性映射，將原始資料映射到更高維度的空間。這種映射讓原本在低維空間中線性不可分的數據，在高維空間中變得線性可分，從而允許SVM有效地分類數據。此外，這種映射策略有效避免在處理高維資料時可能遇到的維度災難(Curse of dimensionality)問題。

表四十一展示了在使用不同核函數時，SVM在不同組別的最佳特徵組合上的分類結果。具體而言，使用RBF核函數的效果最佳，準確率達到0.973。多項式核函數次之，準確率為0.881，而線性核函數的準確率最低，為0.829。

表四十一 最佳特徵組合之SVM分類結果

|  |  |  |
| --- | --- | --- |
| RBF | Polynomial | Linear |
| 0.973 | 0.881 | 0.829 |

確定SVM的參數與核函數後，表四十二、表四十四、表四十六、表四十八的混淆矩陣展示了最佳特徵組合使用SVM訓練後的結果，得知在使用SVM訓練顯著性參數後其準確率可達96%。

表四十二 F&NF於SVM之最佳特徵組合混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | F | 16 | 0 |
| NF | 1 | 12 |
|  |  | | F | NF |
|  |  | | Predicted label | |

表四十三 F&NF於SVM之RBF最佳特徵組合分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.966 | 0.975 | 0.966 | 0.966 | 0.965 |

表四十四 F&HA於SVM之最佳特徵組合混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | F | 15 | 1 |
| HA | 0 | 22 |
|  |  | | F | HA |
|  |  | | Predicted label | |

表四十五 F&HA於SVM之RBF最佳特徵組合分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.973 | 0.978 | 0.968 | 0.978 | 0.973 |

表四十六 NF&HA於SVM之最佳特徵組合混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | NF | 12 | 1 |
| HA | 0 | 22 |
|  |  | | NF | HA |
|  |  | | Predicted label | |

表四十七 NF&HA於SVM之RBF最佳特徵組合分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.971 | 0.956 | 0.95 | 0.923 | 0.977 |

表四十八 F&NF&HA於SVM之最佳特徵組合混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | F | 16 | 0 | 0 |
| NF | 2 | 11 | 0 |
| HA | 0 | 0 | 22 |
|  |  | F | NF | HA |
|  |  | Predicted label | | |

表四十九 F&NF&HA於SVM之RBF最佳特徵組合分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.96 | 0.963 | 0.948 | 0.981 | 0.952 |

表四十三、表四十五、表四十七、表四十九呈現了在SVM辨識實驗中採用關鍵參數後，得到的各種效能評估指標，我們可以觀察到，所有效能指標值均超過95%，這表明我們篩選的關鍵參數集合能夠精確地識別出具有高跌倒風險的老年人群。

1. RF(Random Forest)分析結果

隨機森林(Random Forests, RF)是一種集成式的方法，通過組合多個決策樹來提高模型的預測準確性和穩定性，它既可以用於分類問題也可以用於回歸問題。隨機森林的核心思想是利用多棵決策樹的預測結果進行投票或平均，從而得到最終的預測結果。這種方法有效地降低了單棵決策樹可能出現的過擬合問題，並提高了模型對新數據的泛化能力。

隨機森林中還有兩個極其重要的參數，分別是n\_estimators 和 max\_feature， n\_estimators 參數代表隨機森林中樹木的數量，當增加樹的數量時，模型的穩定性和性能通常會提高，因為多棵樹可以更好地捕捉數據中的各種模式，並通過平均或投票的方式減少過度擬合的風險。然而，增加樹的數量也會導致模型的訓練時間更長，並增加模型的計算成本。

而另一個參數 max\_features 指的是決策樹在每次分裂時，隨機選擇的特徵數量的最大值，這個參數對決策樹的分裂方式有極其重要的影響。如果參數值較小，那麼樹間的差異性會增加，這有助於提高模型的多樣性，從而增強模型對於新資料的泛化能力。但如果這個值設置得過小，又可能會導致每棵樹的性能不是很好，因為它們無法利用數據的所有特徵。反之，如果 max\_features 設置得過大，則樹之間的差異性會降低，這可能會導致模型過於關注訓練數據，從而增加過度擬合的風險。

表五十 F&NF於RF之最佳特徵組合混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | F | 16 | 0 |
| NF | 1 | 12 |
|  |  | | F | NF |
|  |  | | Predicted label | |

表五十一 F&NF於RF之最佳特徵組合分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.966 | 0.975 | 0.966 | 0.966 | 0.965 |

表五十二 F&HA於RF之最佳特徵組合混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | F | 15 | 1 |
| HA | 0 | 22 |
|  |  | | F | HA |
|  |  | | Predicted label | |

表五十三 F&HA於RF之最佳特徵組合分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.973 | 0.978 | 0.968 | 0.978 | 0.973 |

表五十四 NF&HA於RF之最佳特徵組合混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | | NF | 12 | 1 |
| HA | 0 | 22 |
|  |  | | NF | HA |
|  |  | | Predicted label | |

表五十五 NF&HA於RF之最佳特徵組合分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.971 | 0.956 | 0.95 | 0.923 | 0.977 |

表五十六 F&NF&HA於RF之最佳特徵組合混淆矩陣

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| True label | F | 15 | 0 | 0 |
| NF | 1 | 12 | 0 |
| HA | 0 | 0 | 22 |
|  |  | F | NF | HA |
|  |  | Predicted label | | |

表五十七 F&NF&HA於RF之最佳特徵組合分類結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | Specificity | F1-score |
| 0.98 | 0.98 | 0.974 | 0.98 | 0.976 |

表五十、表五十二、表五十四、表五十六的混淆矩陣展示了最佳特徵組合訓練後的結果，在使用RF訓練時，最佳特徵組合其準確率可達96%。表五十一、表五十三、表五十五、表五十七所示的隨機森林(RF)模型使用最佳特徵組合的實驗結果中，我們可以觀察到模型在各項效能評估指標上均展示出卓越的表現，準確率更是高達96%，這表明我們篩選的顯著性參數集合能夠精確地識別出具有高跌倒風險的老年人群。

1. 討論

本研究通過對健康成年人與老年人群進行的步態分析，旨在更好地理解和預測跌倒風險。通過結合慣性測量單元(IMU)和先進的機器學習技術，我們對老年人的步態進行了深入分析，以識別出高跌倒風險的個體。本研究的主要發現支持了我們的假設，即利用互相關和機器學習方法可以有效區分老年非跌倒者、老年易跌倒者和健康成年人。

雖然步態分析在跌倒預防方面已被廣泛研究，但通過結合互相關和機器學習這兩種技術，我們能夠更深入地瞭解步態數據中的模式，從而更精確地預測跌倒風險，互相關技術可以揭示不同步態參數之間的時間依賴性，而機器學習模型則可以從這些關係中學習，以識別跌倒的潛在預兆。

本研究不僅加深了對老年人步態特徵的理解，也提高了對其行走模式微妙差異的識別能力。通過詳細分析，揭示了導致跌倒的關鍵因素，這對醫療專業人員制定個性化的跌倒預防策略及開發跌倒檢測和預警系統至關重要，有望推進老年人跌倒預防，顯著提升他們的生活品質，增強自信和獨立性，並降低跌倒帶來的健康風險和醫療成本，為個人與社會帶來廣泛利益。

儘管研究結果看似不錯，但也存在一些限制。首先，樣本大小相對較小，可能影響結果的一般化。其次，雖然我們使用的機器學習模型表現出色，但這些模型的解釋性較差，未來研究可以探索更透明的模型以提高結果的可解釋性。

未來研究建議：未來的研究可以擴大樣本規模，包括不同背景和條件的實驗參與者，以驗證我們的發現。此外，進一步探索機器學習模型在預測其他健康結果方面的應用，將是一個有價值的研究方向。

參考文獻

[1] Peel, N.M., *Epidemiology of falls in older age.* Canadian Journal on Aging/La Revue canadienne du vieillissement, 2011. **30**(1): p. 7-19.

[2] Haagsma, J.A., et al., *Falls in older aged adults in 22 European countries: incidence, mortality and burden of disease from 1990 to 2017.* Injury prevention, 2020. **26**(Suppl 2): p. i67-i74.

[3] Azizah Mbourou, G., Y. Lajoie, and N. Teasdale, *Step length variability at gait initiation in elderly fallers and non-fallers, and young adults.* Gerontology, 2003. **49**(1): p. 21-26.

[4] Muscaritoli, M., et al., *Muscle atrophy in aging and chronic diseases: is it sarcopenia or cachexia?* Internal and emergency medicine, 2013. **8**: p. 553-560.

[5] Kobsar, D., et al., *Evaluation of age-related differences in the stride-to-stride fluctuations, regularity and symmetry of gait using a waist-mounted tri-axial accelerometer.* Gait & posture, 2014. **39**(1): p. 553-557.

[6] Grabiner, P.C., S.T. Biswas, and M.D. Grabiner, *Age-related changes in spatial and temporal gait variables.* Archives of physical medicine and rehabilitation, 2001. **82**(1): p. 31-35.

[7] Hua, A., et al., *Accelerometer-based predictive models of fall risk in older women: a pilot study.* NPJ digital medicine, 2018. **1**(1): p. 25.

[8] Lien, W.-C., et al., *Intelligent fall-risk assessment based on gait stability and symmetry among older adults using tri-axial accelerometry.* Frontiers in bioengineering and biotechnology, 2022. **10**: p. 887269.

[9] Howcroft, J., J. Kofman, and E.D. Lemaire, *Review of fall risk assessment in geriatric populations using inertial sensors.* Journal of neuroengineering and rehabilitation, 2013. **10**: p. 1-12.

[10] Crispim-Junior, C.F., et al., *Alzheimer's patient activity assessment using different sensors.* Gerontechnology, 2012. **11**: p. 266-267.

[11] Park, K., H. Dankowicz, and E.T. Hsiao-Wecksler, *Characterization of spatiotemporally complex gait patterns using cross-correlation signatures.* Gait & Posture, 2012. **36**(1): p. 120-126.

[12] Roshdibenam, V., et al., *Machine learning prediction of fall risk in older adults using timed up and go test kinematics.* Sensors, 2021. **21**(10): p. 3481.

[13] Martinez, M. and P.L. De Leon, *Falls risk classification of older adults using deep neural networks and transfer learning.* IEEE journal of biomedical and health informatics, 2019. **24**(1): p. 144-150.

[14] Saho, K., et al., *Gait classification of young adults, elderly non-fallers, and elderly fallers using micro-Doppler radar signals: Simulation study.* IEEE sensors journal, 2017. **17**(8): p. 2320-2321.

[15] Punt, M., et al., *Characteristics of daily life gait in fall and non fall-prone stroke survivors and controls.* Journal of neuroengineering and rehabilitation, 2016. **13**: p. 1-7.

[16] Schafer, R.W., *What is a Savitzky-Golay filter?[lecture notes].* IEEE Signal processing magazine, 2011. **28**(4): p. 111-117.

[17] Mello, R.G., L.F. Oliveira, and J. Nadal, *Digital Butterworth filter for subtracting noise from low magnitude surface electromyogram.* Computer methods and programs in biomedicine, 2007. **87**(1): p. 28-35.