運用統計模型與分類器分辨帕金森氏患者產生步態凝凍與步態凝凍前的狀態

指導老師：黃怡婷老師

組員：嚴潔翎、柯怡瑄、黃承平、賴仁凱、許弘叡、葉于廷

1. **摘要**

帕金森氏症是一種神經退化疾病，會緩慢地影響人體的神經系統，最終導致死亡。在疾病晚期常發生一種叫做凝凍步態 (Freezing of Gait, FOG) 的症狀，所謂凝凍步態，是指在正常行走時，雙腳突然間無法依照自主意識活動，開始顫抖、無法跨步、在原地動彈不得。

帕金森氏症的患者隨時都有發生凝凍步態的可能，最常發生在行走時，發作當下，患者會無法移動雙腳，身體不聽使喚，造成容易緊張與恐懼，以至於症狀更不容易緩解，發作時間一旦拉長，便會增加跌倒受傷的風險。發生凝凍步態的主因尚未明朗，目前是依觀察病患步伐啟動困難與步態異常做為此現象發作與否的辨識基準。

為了找到能夠準確判斷的基準以及節省人力的方法，同時，將此方案導入人工智慧中，本研究針對帕金森氏症病患的步態蒐集資料，建置多種不同的步態特徵，參考 [1], [2], [3] 所得的特徵與配適步態之合適樣條函數所得特徵，利用數個分類器進行辨識凝凍步態的現象，最後藉由指標評估各分類器的表現，預期找出適合幫助病患判斷症狀的分類方法。

1. **研究動機與研究目的**

**一、研究動機**

帕金森氏症是繼阿茲海默症後，全球第二常見的神經退化疾病，隨著全球高齡化，患病風險日益增高，60歲以上每增加一歲，盛行率便提高1%；80歲以上則每增加一歲，盛行率上升4%，到了病症的晚期，病患會開始有行走與溝通的困難，甚至開始有容易感到疲倦、憂鬱等心理問題，需要長照的他們日漸為社會帶來龐大的社會成本。隨著病患占世界人口比例的漸增，間接加重了生產者的經濟與勞力負擔，影響的雙方的生活品質。

凝凍步態是帕金森氏症的病徵，此症狀發作時，病患雙腿無法向前邁步，待在原地顫抖，在回復到正常狀態前，病患急需他人攙扶或支撐以保持平衡。此症狀很容易在患者身處人群時發生。例如上下火車時，人群在車門附近迅速地聚集會使病患情緒不穩，間接影響病徵的出現，此刻病患突然無法行動是非常危險的，推擠與催促使病患更加緊張，愈無法緩解病徵，大幅增加受傷的機會。現今對於凝凍步態發作尚未有一個明確的定義，目前是依人工觀察來判斷症狀是否發生，但這並非是長遠之計，長期觀察一個人症狀發生時的特徵並時刻注意他的安危，不僅消耗社會資源，也容易造成患者的心理負擔。此研究的目的是期望能讓帕金森氏症的患者能夠盡可能擁有自主生活的能力。

隨著資料收集技術與機器學習理論的成熟，已經可以將凝凍步態問題導入人工智慧中，透過攜帶式裝置收集病人的步態資料，並加以訓練人工智慧，能獲取比主觀判別更優秀的辨別標準且更準確地偵測步態的異常，甚至可在凝凍步態發生前發現到步態的異常。未來將裝置導入實際生活獲取大量步態資料後即能完整的訓練模型，得到更精確的判斷，並在發生病症時，發出提示聲提醒病患與周遭的人，除了病患能夠知道本身正在發作凝凍步態，他人也能較易觀察異樣在旁提供協助，使病患免於跌倒受傷的風險，擁有更完善的生活品質，同時減輕醫療照護負擔，舒緩社會壓力。

**二、研究目的**

本研究藉由多個回顧文獻來找出凝凍步態的可能特徵變數，再者，多數資料是利用感測器收集後，利用無線傳輸傳送到電腦。收集訊號有可能會受到干擾，而造成部分資料異常，因此本研究會採用樣條函數配適收集訊號的軌跡，利用多個點決定特徵，該做法第一個好處是可以避免異常訊號的影響，第二個好處是樣條函數的係數可視為步態特徵納入至模型中。運用文獻與樣條函數估計出的特徵變數，使用數個分類器來分類該視窗是否為凝凍步態，試圖找出較少特徵與計算時間較短的分類方法。以下為本次的研究目的：

1. 找出對凝凍步態的辨識較有貢獻的特徵。
2. 配適步態合適的樣條函數。
3. 運用交叉驗證後的各統計評估指標區找出較適當的縮減維度方法。
4. 將特徵依照部位分開，用交叉驗證後的各種統計評估指標判斷是否有單一部 位特徵足以讓分類器的分類預測能力效果良好。
5. 運用交叉驗證後的各種指標檢視分類器的辨別預測能力。
6. 找出在分類預測凝凍步態用時較短的分類方法。
7. **文獻回顧與探討**

近年來，已有多篇論文探討過帕金森氏症患者穿戴步態感測器所獲取的資料，比較不同特徵擷取和相異分類器的的辨識成效。

1. **研究源起**

凝凍步態會使患者短暫失去行走的能力，50 % 的帕金森氏症患者們常會有凝凍步態的症狀 [3]。然而，此症狀的起因是相當難確定的，其對藥物、環境或個人感知等相當敏感，例如凝凍步態較常在家發生而非實驗室或醫生的診間，而發作與否的判別是由病患描述或直接觀察來確認。

藥物對帕金森氏症的治療是相當無效的，服藥之後，藥效會在二到六小時間漸減。隨著病情的惡化，藥效時間會縮短且需更頻繁用藥。對於還沒發作凝凍步態的病患，往往傾向拒絕藥物治療，導致惡化程度無法顯著控制，以至於病患往後會更需要照護或復健等非藥物治療[4]。

在家中，病患會穿戴有節奏的聽覺刺激(Rhythmic auditory stimulation, RAS)儀器，以患者正常步態速度的110 % 發出常規的滴答聲，目的是提高患者的步態速度並減少步態變異程度 [5]，但實際訪問患者的結果是他們並未感受到凝凍步態的症狀有顯著改善 [6]。物理治療師指出，藉由聲音、影像、觸覺能使患者從凝凍步態中更快回到正常的行走方式，因此想要從步態資料找到凝凍步態的特徵，利用機器學習 (Machine Learning) 中的分類器 (Classifier) 找出步態異常的跡象後，機器發出聲響、震動或顯示訊號來提示病患，加速症狀的舒解。

**二、實驗過程與資料獲取 [3]**

實驗室中我們記錄病人的步態資料，分別要求好發凝凍步態的受試者在各個部位穿戴3D加速度感測儀（綁在脛骨(腳踝上方處)、大腿(膝蓋上方處)、後背下方）後，沿著研究路徑進行十分鐘實驗，其中動作包括隨意走動、轉彎和停止，並記錄 x、y和z方向的加速度。過程中，凝凍步態的辨別是依照身旁物理治療師的觀察。

資料中，十位受試者平均年齡 66 - 67 歲，Hoehn-Yahr 分級表 [7]平均分數 2.6，嚴重視力或聽力不足、失智、有神經或骨科疾病徵兆的人都並不會參與試驗。患者在用藥週期的 OFF 階段的早晨（在他們最後一次服用抗帕金森氏症藥物後超過 12 小時） 接受了測試。在 ON 期間頻繁發生凝凍步態的兩位病患沒有因實驗而停止藥物治療。

帕金森氏症常用Hoehn-Yahr 分級表 [7] 來分級，針對相對的身體能力的喪失分成五個階段：

第一階段：只有身體的一側有症狀

第二階段：身體兩側都有症狀但不影響平衡

第三階段：平衡受損但還能控制身體

第四階段：失去許多功能，但可在不須協助下行走或站立

第五階段：需要輪椅或臥床

實驗的研究路徑(圖1)：

(a) 在走廊直線來回走動，包括中途的180度轉彎。

(b) 受試者會根據實驗人員的指示，在隨意走動間做出包括一連串的停止與360 度轉彎等動作，其中至少會有六次轉彎，並且保證一定有三種不同方向。

(c) 模擬日常行走 (activities of daily living (ADL) part)，進出房門、去實驗室廚房拿食物吃，盛水回房間等。

|  |
| --- |
|  |
| 圖 1 步態軌跡示意圖 (參考[3]) |

**三、視窗切割與定義**

收集完的資料是頻率為 64 次/秒的資料，由於凝凍步態會持續數秒，故會將單點的資料整併成固定秒數的視窗(window)以利觀察較為完整的現象，而Bachlin [3] 等人將資料的視窗定義成 4 秒一個視窗，且視窗間的重疊長度為 3.5 秒，即視窗一次滑動 0.5 秒。

加速度的資料可視為多個波段相加的樣貌，利用傅立葉轉換(Fourier transform)對加速進行分解，可觀察加速度為哪幾個頻率的波段所組成的資訊，而發現在正常走路與站立時，加速度 3 – 8 Hz 波段的資料並不明顯，Moore等人 [8] 介紹了一個凝凍指標(Freezing Index, FI)，以客觀地識別凝凍步態。這個凝凍指標定義為凝凍帶(freezing band) 3 – 8 Hz中的功率除以運動帶(locomotor band) 0.5 – 3 Hz 中的功率。當凝凍指數值高於某個凝凍閾值(threshold)，視窗便被辨識為凝凍步態事件。

在 Pardoel [1] 的研究中，視窗分為四種狀態：無凝凍步態(No-FOG)、凝凍步態前(Pre-FOG)、凝凍步態過渡期(Pre-FOG Transition)、凝凍步態(FOG)，在廣義的分類中，後三種都屬凝凍步態。以圖2說明:凝凍步態事件的前兩秒稱做前凝凍步態片段(Pre-FOG segment)，發生期間稱做凝凍步態片段(FOG segment)，其他期間稱之為無凝凍視窗片段(No-FOG segment)。完整視窗落前凝凍步態片段中在被定義為前凝凍步態(Pre-FOG)，像圖 2 中的 W7 - W11，若視窗的開頭落在前凝凍步態片段，結束落在凝凍步態片段中，則視窗被分類為凝凍步態過渡期(Pre-FOG transition)，像圖 2 中的 W12 - W16，視窗完整落在凝凍步態片段中則被分類為凝凍步態(FOG)，像圖 2 中的 W17，其餘則被分類為無凝凍步態(No-FOG)。

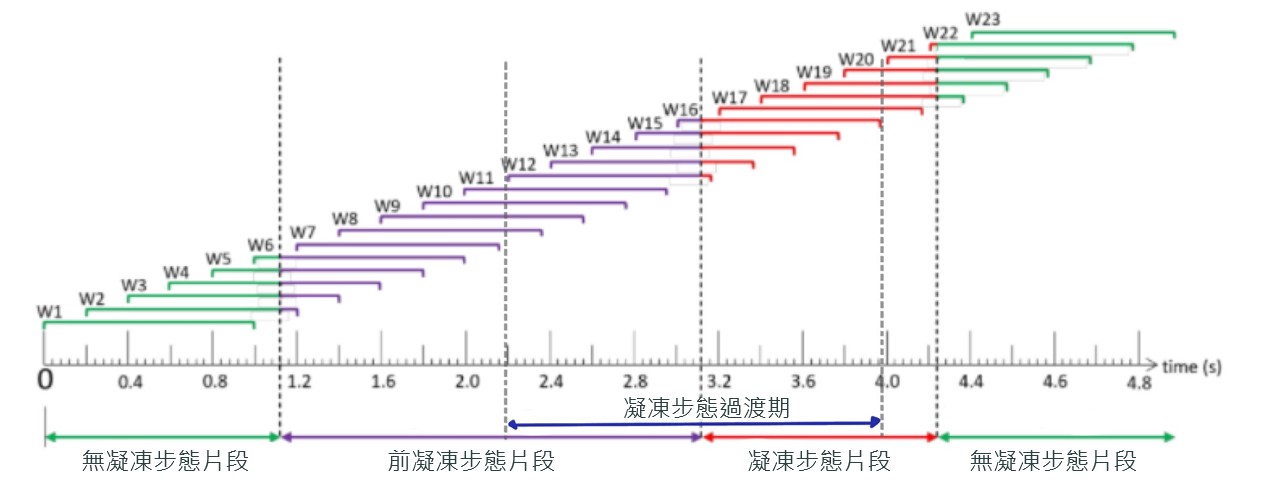


圖2

**四、偵測凝凍步態**

將視窗資料的每個特徵輸入至系統中，系統需要在發作後的兩秒內偵測出凝凍步態，這樣才會被認定是一次成功偵測的事件。每個病患輪流進行辨識後，得到平均特異性(Specificity)為 81.6 % 平均敏感度(Sensitivity)為 73.1 %，最好的預測發生在病患 4 和病患 10，最差的預測發生在病患 8 和病患 1，利用不同部位模擬的結果，在大腿部位的垂直方向加速度表現最好[3]。

Pardoel 等人 [1] 將視窗切割成長度為 1 秒，每次滑動 0.2 秒，且定義凝凍視窗為整個視窗位於凝凍狀態中，並利用原本資料擷取出一百多個特徵，包含視窗的最大值、最小值、中位數、平均值、標準差、傅立葉轉換後的最大頻率、傅立葉轉換後介於 3 – 8 Hz 的能量大小、傅立葉轉換後介於0.5 – 3 Hz的能量大小、傅立葉轉換後介於3 – 8 Hz的能量大小與 0.5 – 3 Hz的能量大小的比率和傅立葉轉換後的總能量總和的特徵，接者使用relief-F和最小冗餘最大特徵相關（mRMR）進行特徵篩選，最後輸入至RUSboost 集成分類器訓練，執行留一交叉驗證法，產出多個資料的測試結果，計算所有病患的平均特異性 (Specificity) 為 91.6 %、平均敏感度 (Sensitivity) 為61.9% [1]。

1. **研究方法及步驟**
2. **研究數據**

資料取自於網路資料庫，其由美國加利福尼亞州爾灣的公立大學─爾灣加州大學(University of California ,Irvine)所建構的UCI資料庫系統中心，此系統中心目前包含了 622 多個不同資料分析方法及各個領域的資料集以供研究使用。本研究選擇資料庫中的Daphnet Freezing of Gait Data Set進行統計分析。資料集內共有 10 位帕金森氏症病患在走路測試中經歷凝凍步態 (FOG) 時的三部位 (腳踝、軀幹、大腿) 三軸 (水平向前、垂直、水平橫移) 加速度資料 [3], [9], [10]。

根據 [1], [2], [3] 與樣條函數所產生特徵，視窗的特徵值可大致分為三種。第一種，根據原始加速度值得到的敘述性統計，如表1。第二種，原始資料經過傅立葉轉換後得到的數值經過各種運算而得，如表2，因為三個部位各有三軸，所以以下特徵值除了平均加速度能量 (AAE) 外都是衍伸出9個特徵值，而平均加速度能量(AAE) 則為 3 個特徵值，故總共有 129 個根據文獻所得特徵值。第三種，用視窗內64個加速度資料點配適樣條函數，估計出合適的軌跡，取出B-樣條 (B-spline) 基底的座標為特徵和自然三次樣條的每段的方程式係數作為特徵。以下為特徵定義：

表1 原始加速度值所得特徵

|  |  |
| --- | --- |
| 特徵名稱 | 描述 |
| 最大值 | 原始加速度數值在各個視窗所得敘述性統計 |
| 最小值 |
| 平均值 |
| 標準差 |
| 中位數 |
| 主導方向的特徵值(eigenvalues) (EVA) | 加速度數據沿 x、y 和 z 軸的共變異數矩陣的特徵值 |

表2 經過傅立葉轉換後所得特徵

|  |  |
| --- | --- |
| 特徵名稱 | 描述 |
| 最大值 | 原始加速度值經過傅立葉轉換後所得能量在各個視窗的敘述性統計 |
| 最小值 |
| 平均值 |
| 平均加速度能量 | 三軸能量的平均(AAE) |
| 最大頻率 | 最大能量值所對應的頻率 |
| 凝凍帶 | 3 – 8 Hz間的能量加總(Freezing band) |
| 運動帶 | 0.5 – 3 Hz間的能量加總(Locomotor band) |
| 凝凍指數 | 3 – 8 Hz除以0.5 – 3 Hz間的能量加總(Freezing Index) |
| 總能量 | 視窗內能量的加總(Totalpower) |

表3 樣條函數估計視窗特徵

|  |  |
| --- | --- |
| 特徵名稱 | 描述 |
| 一階三次自然樣條方程式係數 | 三次自然樣條估計原始加速度數值在各個視窗所得各個係數 |
| 二階三次自然樣條方程式係數 |
| 三階三次自然樣條方程式係數 |
| 四階三次自然樣條方程式係數 |
| B-樣條基底座標元素 | B-樣條估計原始加速度數值在各個視窗所得各個基底座標元素 |

1. **維度縮減方法**

依據上述的研究數據及多個特徵值進行維度縮減，使用的維度縮減方法包含法包含：

1. 最小冗餘最大特徵法(mRMR)
2. 互資訊（mutual information, 簡稱：MI）
3. Relief-F
4. **統計分析方法**

利用維度縮減後所保留的特徵值進行統計分析，統計分析方法包含：

1. 單純貝氏分類器（Naive Bayes classifier）
2. K近鄰分類演算法（k Nearest Neighbor）
3. 支援向量機(Support Vector Machine, 簡稱：SVM)
4. RUSboost
5. 隨機森林(Random Forest)
6. **評估指標**

上述所有統計分析方法皆運用除了第四位及第十位病患外的資料做 7:3 等比例的交叉驗證來檢視模型預測能力的好壞，其值越高代表預測能力越佳。此外，各分析結果將呈現如表1形式：真負類(TN)代表為該視窗實際為非凝凍步態且預測該視窗亦為非凝凍步態、真正類(TP)代表為該視窗實際為凝凍步態且預測該視窗亦為凝凍步態、偽負類(FN)代表該視窗實際為凝凍步態且預測該視窗為非凝凍步態、偽正類(FP)該視窗實際為非凝凍步態且預測該視窗為凝凍步態，並依此建構各項統計評估指標評估統計方法的分類表現之優劣。各項評估指標陳列於表2。

表1 預測結果表示

|  |  |  |  |
| --- | --- | --- | --- |
|  | | 預測結果 | |
| 正類 | 負類 |
| 真實狀態 | 正類 | TP | FN |
| 負類 | FP | TN |

表2 統計評估指標

|  |  |  |
| --- | --- | --- |
| 評估指標 | 說明 | 定義 |
| 召回率(Recall) | 真實狀態為陽性其預測結果亦陽性的比率 |  |
| 特異度(Specificity) | 真實狀態為陰性其預測結果亦為陰性的比率 |  |
| 準確率(Accuracy) | 測量整體分類準確率 |  |
| 精確率(Precision) | 預測結果為陽性其真實狀態亦為陽性的比率 |  |
| F1-分數(F1-score) | 考慮精確率及召回率下，測量準確程度 |  |

1. **研究步驟**

本論文研究進行步驟如下：

1. 根據文獻找出數據的特徵。
2. 配適步態的樣條函數得到特徵。
3. 運用維度縮減找出對分類較有貢獻的特徵。
4. 運用交叉驗證後的統計評估指標判斷各分類器的分類與預測能力。
5. 找出對分類與預測是否有凝凍步態的最佳縮減維度法、最有代表性的特徵與 最有效率的分類器。
6. **先導試行分析結果和預期結果**
7. **先導試行分析結果**

將資料切割成 1 秒視窗，每次滑動 0.25 秒後產生下一個視窗。視窗分類為四種的比例與次數統計呈現在表 3。

表3 四種的各視窗次數與比例

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 病患代號 | 無凝凍步態 | 凝凍步態 | 凝凍步態過渡期 | 前凝凍步態 |
| 01 | 12248 (96.1%) | 332 (2.6%) | 80 (0.6%) | 88 (0.7%) |
| 02 | 9323 (92.02%) | 632 (6.2%) | 92 (0.9%) | 84 (0.8%) |
| 03 | 14492 (91.6%) | 956 (6%) | 198 (1.3%) | 170 (1.1%) |
| 04 | 12230 (100%) |  |  |  |
| 05 | 11013 (83.7%) | 1640 (12.5%) | 256 (1.9%) | 248 (1.9%) |
| 06 | 13171 (95.9%) | 488 (3.6%) | 39 (0.3%) | 41 (0.3%) |
| 07 | 10205 (96.2%) | 228 (2.1%) | 87 (0.8%) | 89 (0.8%) |
| 08 | 7672 (89.9%) | 749 (8.8%) | 54 (0.6%) | 58 (0.7%) |
| 09 | 9591 (89.1%) | 975 (9.1%) | 100 (0.9%) | 100 (0.9%) |
| 10 | 12078 (100%) |  |  |  |
| 全部病患 | 112023 (94%) | 6000 (5%) | 906 (1%) | 1. (1%) |

1. **根據蒐集資料之預期結果**

(1) 隨機森林在分類有凝凍步態時有較好的能力。

(2) 腳踝得到之數據可以相較其他部位得到之數據有更好的預測效果。

(3) 最小冗餘特徵選擇維度縮減法所挑出的特徵值有較好的分類效果。

(4) 切割視窗長度為1秒，重疊0.75秒時的分類效果最好。

(5) 傅立葉轉換原始資料後得到的變數提供較好辨別凝凍步態的特徵。

1. **遇到的困難與解決方法**

分類器有別於以往的統計方法，較依賴大量的演算法來達成成效，此新穎的方法需要熟知演算法的流程，方可順暢的使用，需要短時間內填補機器學習領域相關知識。此外，在分析實務數據時會面對的問題包括：龐大的資料處理程序(切割視窗、取出上百個特徵、視窗的定義)、不平衡數據使得分類結果過度配適或凝凍時間短無法辨別其凝凍特徵等。

解決的方式包含團隊縝密的分工與討論，每人研讀不同方法後，分享給其他組員，達到有效率的學習。遇到障礙時，可諮詢具有專業經驗的指導老師或查詢更多相關文獻，從中效仿的解決方法與思考模式，查看其做法是否可以成為解套自身問題的關鍵。

1. **參考文獻**

[1] Pardoel, S., Shalin, G., Nantel, J., Lemaire, E. D., Kofman, J., Early Detection of Freezing of Gait during Walking Using Inertial Measurement Unit and Plantar Pressure Distribution Data, Sensors, Vol. 21, no. 6, 2021.

[2] Mazilu, S., Calatroni, A., Gazit, E. and Roggen, D. Feature Learning for Detection and Prediction of Freezing of Gait in Parkinson’s Disease, Machine Learning and Data Mining in Pattern Recognition Conference, 2013. DOI:[10.1007/978-3-642-39712-7\_11](http://dx.doi.org/10.1007/978-3-642-39712-7_11)

[3] Bachlin, M., Plotnik, M., Roggen, D., Maidan, I., Hausdorff, J. M., Giladi, N. and Troster, G., Wearable assistant for Parkinson’s disease patients with the freezing of gait symptom, IEEE Trans Inf Technol Biomed, Vol. 14, no. 2, 436-46, 2010.

[4] N. Giladi, M. P. McDermott, S. Fahn, S. Przedborski, J. Jankovic,M. Stern, and C. Tanner, “Freezing of gait in PD: Prospective assessment in the DATATOP cohort,” Neurology, vol. 56, no. 12, pp. 1712–1721, 2001.

[5]J. M. Hausdorff, J. Lowenthal, T. Herman, L. Gruendlinger, C. Peretz, and N. Giladi, “Rhythmic auditory stimulation modulates gait variability in parkinson’s disease,” European Journal of Neuroscience, vol. 26, no. 8, pp. 2369–2375, October 2007.

[6] A. Nieuwboer, K. Baker, A.-M. Willems, D. Jones, J. Spildooren,I. Lim, G. Kwakkel, E. Van Wegen, and L. Rochester, “The Short-Term Effects of Different Cueing Modalities on Turn Speed in People With Parkinson’s Disease,” Neurorehabil Neural Repair, 2009, [Epub ahead of print].

[7] M. M. Hoehn and M. D. Yahr, Parkinsonism: onset, progression and mortality, Neurology, vol. 17, no. 5, pp. 427–42, May 1967.

[8] Moore, T. S., MacDougall, H. G. and Ondo, W. G. Ambulatory monitoring of freezing of gait in Parkinson’s disease, J. Neurosci Methods, Vol. 167, no. 2, 340-8, 2008.

[9] UC Irvine Machine Learning Repository (https://archive.ics.uci.edu/ml/index.php)

[10] UC Irvine Machine Learning Repository.(2013). Daphnet Freezing of Gait Data Set [Data file]. https://archive.ics.uci.edu/ml/datasets/Daphnet+Freezing+of+Gait