

# 深度學習應用於步態偵測

陳逸星<sup>1</sup>、蘇木春<sup>2</sup>、余昌翰<sup>3</sup>、陳奕蘋<sup>4</sup>、賴映如<sup>5</sup>

<sup>1</sup>國立中央大學、<sup>2</sup>國立中央大學、<sup>3</sup>國立中央大學、<sup>4</sup>國立中央大學、<sup>5</sup>國立中央大學

<sup>1</sup>yihsinchen@g.ncu.edu.tw、<sup>2</sup>muchun.su@g.ncu.edu.tw、<sup>3</sup>ke511081177@gmail.com、

<sup>4</sup>apple870728@gmail.com、<sup>5</sup>lai860423@gmail.com

## 論文摘要

在本研究中提出了一種利用雙向長短記憶神經網路的模型，藉由九軸感測器所接收到雙腳之各項慣性數據，來分類出人類走路步態的各個階段。在現代醫學與照護服務中，步態是一個能幫助醫師或者專業人士來判斷一個人的身體狀態。因此除了直接用人眼觀測外，一個能藉由儀器快速辨識並且客觀分類人類步態的系統，可以加速整個診療或者復健的過程，並且省略了大量的人力。本研究運用深度學習的運算模式，經過訓練去符合最後的模型，最終完成初步的分類。本篇計畫希望藉由提出一個輕便的系統，來幫助診療人士能在初步診療時能快速地得出一個人走路週期的7個步態。

**關鍵詞：**深度學習、智慧型機器人、智慧照護

## 1. 前言

步態分析在近年來已經被廣泛應用在各類的研究上作為重要的生物特徵，而人類的步態則可以看出一個人的身體狀態。因此有效的研究步態並且做出分類與篩選，可以幫助人們去應用在醫療與老人照護的服務中。

目前收集步態訊號的技術，主要分為三類：視覺影像設備、地面壓力傳感器與穿戴式感測器。而在這三者當中，使用具加速度、陀螺儀、磁力計之九軸穿戴式感測器具有便於攜帶、低耗能且最低成本之優點，但不同於另外兩類訊號收集技術，穿戴式感測裝置所分割之步態階段較為粗糙。

本研究計畫目的是希望藉由感測器所蒐集之九軸訊號特徵進行定義，透過定義能夠標示出精準的步態事件以此切割出精細的步態階段，希望能藉由深度學習的模型，快速地標示出各個步態階段，以此補足上述穿戴式感測裝置資料較為粗糙之缺點。

## 2. 文獻探討

步態分析在近年來被應用在各項研究上作為相當重要的生物特徵，人類的步態會反映出一個人的身體狀態。而步態週期(Gait Cycle)是由七個步態

階段(圖 1) (Periods)所組成[1][2]。

由於步態時間週期差異與步態階段之特徵可以直接觀察到人體肌肉、神經或是骨骼在行走時的狀態，因此步態分析常被應用在醫療照顧中[3]，像是辨識帕金森氏症的病患之步態，應用在物理治療[4][5]當中，分析中風患者的步態並予以復健治療[6]等。

步態訊號收集，主要分為三類：地面壓力傳感器(Floor Sensors)、視覺擷取設備(Machine Vision)與穿戴式感測器(Wearable Sensors)。

**視覺影像**的優點是可以直接觀察到整個人體的步態與生物特徵，相較於所有的方法，電腦視覺影像將會是最直接且直觀的判斷，運用圖像或是影片來當作實驗訊號[7]，然而此類影像分析相對會耗費較大的運算資源。

使用**地面壓力傳感器**收集資料時，會將感測器、壓力板[8]或是壓力感測鞋墊[9]安裝於地面或腳上，量測並取得各樣實驗數據，因感測器需安裝在一固定的場地，因此受測者活動空間跟活動的種類有一定的限制。

而**穿戴式感測器**收集訊號時，會將**慣性感測元件(Inertial Measurement Units)**配戴置身體各個部位[9][10]，例如腰、膝蓋或是腳踝等。而穿戴式裝置的優點就是便於攜帶，且不需在特定地點架設裝置，使用起來較為自由，然而穿戴式感測裝置卻相較於另外兩者去分割之步態階段較為粗糙。

使用穿戴式裝置要切割出精準的步態階段相對困難，[15]嘗試使用 IMU 參數的閾值(Threshold-based)來幫助分類各個步態階段，然而運用此類方法需要耗費大量的時間去移除感測器的雜訊跟誤差，而當感測器數量增加時會使閾值的擷取以及雜訊的去除變得複雜。而使用機械學習的方法像是 KNN[12]、決策樹(Decision trees)[13]、貝式分類(Bayesian network classifier)[14]等，藉由電腦處理大量的實驗數據，並且運用監督式學習能夠快速地處理所有的資料，並且快速的切割步態。

## 3. 研究方法

為了能夠精準切割出每一個人的步態階段，本研究需要建造一個良好的模型，能夠適用在不同的人身上，並且能夠有效便捷的輸出步態階段，以此來節省大量人力的觀測，也能避免單由人眼觀測紀錄所產生的誤差。

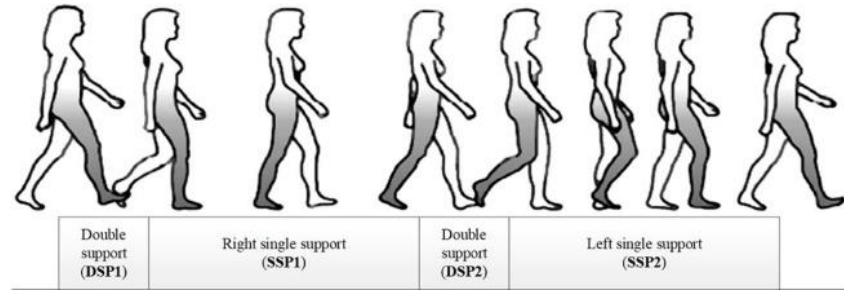


圖 1 [2]、一般人的七個步態階段

### 3.1 資料前處理

使用九軸感測器收集的資料總共有 16 個維度，分別為三軸加速度、三軸角速度、三軸地磁計、歐拉角以及四元數，而綁在雙腳上因此共有 16\*2，共 32 個維度，並且研究人員將根據攝影機拍出來的畫面根據步態的七個階段(圖 1)標示在資料上用以訓練。

而在感測器資料上，根據每筆資料去做平滑化(Filtering)，以作為訓練的輸入。

為了能符合 LSTM 模型之輸入需求，本研究採取 30 值為一次的 Time Step，因此將把所有人的資料從原本 32 維的資料，經過再構(Reshape)變成每一個 Time Step 皆取 30 值共 30\*32 的資料型態進行訓練。

### 3.2 神經網路之輸入

本研究使用雙向長短記憶網路(Bidirectional LSTM)[16][17]，因使用的資料有上下時序關係，且本研究的任務是要辨識出在時間序列裡相對的步態階段，且 LSTM 可以透過訓練學習到哪些資訊該被記憶以及遺忘，因此研究採用 BiLSTM 模型。希望能夠透過當前的資訊並且根據前後資訊觀察後，來確認當下是在步態的哪一階段。

研究透過輸入前 30 格(Frame)感測器所收集的所有 IMU 資訊作為 Features，其資料大小為 30\*32，並且輸出未來 1 格(Frame)的預測階段作為 Feature 丟回模型做位移往下一格繼續訓練。

### 3.3 神經網路模型

研究採用的模型如表(一)使用雙層 Bidirectional LSTM，神經元數量為 64 個，並且中間為一層 Dropout 層，dropout rate 為 0.5。最後使用兩層全連接層(Fully Connected Layer)，第一層輸出神經元為 32 個，激活函數(Activation Function)使用 relu 函數，第二層對應模型最後輸出 7 個階段，研究最終為分類問題，所以使用 softmax 函數。

### Bidirectional LSTM

Input(訓練資料總長,30,32)
Bidirectional LSTM-64
Dropout(0.5)
Bidirectional LSTM-64
Dropout(0.5)
Dense-32, Activation='relu'
Dropout(0.5)
Dense-7, Activation='softmax'
Output(預測階段 1~7)

表(一)神經網路模型設定

### 3.4 損失函數

以下(式 1)為我們訓練模型時的目標函數：

$$L = - \sum_{i=1}^n \sum_{t=1}^c (y_{i,t} * \log(\hat{y}_{i,t})) \quad (式 1)$$

其中 n 為總樣本數，c 為類別種類，y 為步態分類的 Ground Truth Label， $\hat{y}$  則是預測之步態階段。模型會透過不斷的遞迴來減少損失函數。

### 3.5 訊號同步設計

在使用攝影機擷取人體行走畫面，並且加上九軸感測器所收集的資訊時，必須確認感測器所收集到的資訊是與攝影機的每一 frame 是完全同步的。實驗時必須確認攝影機儲存資訊的輸入將與無線感測器所傳回來的資料在同一個 frame 時資料是同步的，因此研究必須設計一個機制來幫助實驗人員不斷確認當下的感測器訊號是同步於攝影機的畫面。

由於 ZED 攝影機開放了 API(Application Programming Interface)使得使用者可以自行運用

程式技術來去做攝影機與感測器的同步整合,本研究根據無線傳感器傳回來的訊號來執行迴圈,當每一次有訊號傳進來時,程式會將當前的各項感測器數據紀錄,並且同時記錄當前那一偵(Frame)的畫面,為此我們可以確認在電腦儲存端兩項數據匯市同步儲存的。

而因為九軸感測器是經由無線傳播到電腦外接 USB 的接收器再將電子訊號從不同接口轉成各項變量,因此確認數據在傳輸上的延遲時間是非常重要的。為此實驗時會請受試者在行走一段時間後,主動去敲擊左右兩腳的感測器,並且對應攝影機上的畫面,來確認當前的資料傳輸沒有延遲。由於敲擊瞬間將產生一個瞬間的加速度,研究人員只需觀測感測器上是否在攝影機的同步時間內有當下動作時同時作動的加速度資訊,實驗人員就可以確認各個實驗器材彼此的訊號是否有同步,藉由這樣各項的訊號收集事件來多重驗證我們各項器材同步機制的可靠性。因此在同步的精度上,誤差將不會超過一偵的時間,即誤差將控制在 1/50 秒的時間,每步時間誤差小於 2%。

#### 4. 實驗

本研究想要藉由雙向長短記憶網路(Bidirectional LSTM)[16][17]作為模型來訓練由九軸感測器得到的資訊,看提出的模型是否能夠直接的根據輸入的資料,去做精準的切割。

在[10]中有提到使用了 7 個 IMU 分別放在腹部中央、兩側大腿前、膝蓋下小腿正前方以及兩腳腳背、並且使用腳上四個點去偵測是否當下腳有著地,其使用了複雜的 DCNN(Deep Convolutional Neural Network)模型來幫助預測,而[15]裡則與本研究一樣使用兩個感測器並且裝在雙腳腳踝上,論文裡使用不同濾波器,並且針對波形去進行特徵解析來完成步態週期分割。

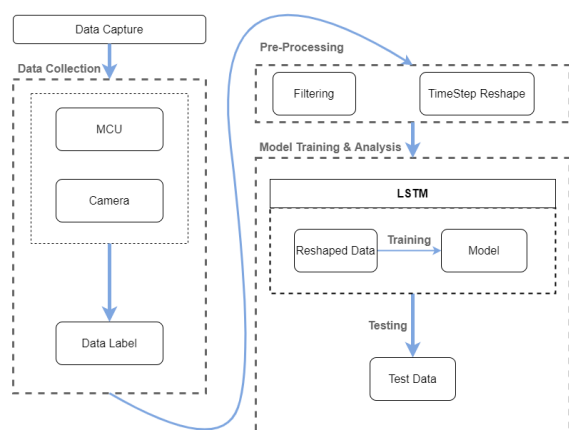


圖 2, 實驗流程圖

本研究的受測者共有 10 位, 6 位男性與 4 位女性, 年齡介於 22~30 歲之間。實驗時, 受測者將兩個感測器綁於雙腳外腳踝上, 並且請他們用最自然的方式前行。



圖 3, 實驗之訊號採集過程與設備安裝說明

#### 4.1 實驗設計

本實驗流程如圖 2, 將由受測者雙腳腳踝上綁上受試者穿戴式九軸感測器(如圖 3), 並且以受試者平時最自由的方式於一給定範圍內來回直行。平均每位受試者來回一趟將走十六步, 大約有完整的八個步態週期。實驗時本研究不只記錄兩個感測器的數據, 同時本研究會用高速攝影機來作為影像的紀錄以及實驗數據標籤之依據, 而兩者資料紀錄的頻率皆為 50hz, 亦即每秒蒐集 50 筆資料。

#### 4.2 實驗設備介紹

本研究使用的穿戴式九軸感測模組來收集人類的步態訊號。而感測器的模組是超核電子的 HI221 無限姿態感測節點, 其可輸出三軸的角速度、加速度以及三軸磁力計的原始數據, 並且模組會藉由九項數據混合計算出四元數以及歐拉角。設備加速度的感測範圍是  $\pm 8G$ , 而陀螺儀的感測範圍是  $\pm 2000^\circ/s$ 。而設備採樣的頻率固定在 50Hz。

資料測量時將九軸感測模組固定在人體雙腳的外腳踝上方(圖 3)。理想的狀態下, 當雙腳站直立正, 感測器模組的底板將平行於地面, 並且整個感測器垂直於地面上, 並且感測器 X 軸將收到來自地心引力方向大約 -1G 的加速度。

實驗時, 將使用 Stereolabs ZED 3D 攝影機, 並且設定他的 fps(frame per second)同樣為 50fps, 為了要同步於感測器輸入的速度。本研究將運用攝影機之影像幫助實驗去為九軸感測器收集到的訓練集(圖 4)做人工標籤步態的七個階段, 並且使用這些資料來為此模型進行訓練。



圖 4, ZED 所蒐集畫面

預測值 實際值	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6	Phase 7
Phase 1	181	2	0	0	0	0	15
Phase 2	19	348	3	0	0	0	0
Phase 3	0	15	161	12	0	0	0
Phase 4	0	0	5	239	21	0	0
Phase 5	0	0	0	14	238	8	0
Phase 6	0	0	0	0	9	156	16
Phase 7	7	0	0	0	0	4	173

表 2、A 組之 Confusion Matrix

Phase	True Positive(TP)	True Negative(TN)	False Positive(FP)	False Negative(FN)	Recall	Precision	Accuracy	F1_Score
1	181	1422	26	17	0.91	0.87	0.97	0.893827
2	348	1259	17	22	0.94	0.95	0.98	0.946939
3	161	1450	8	27	0.86	0.95	0.98	0.901961
4	239	1355	26	26	0.90	0.90	0.97	0.901887
5	238	1356	30	22	0.91	0.89	0.97	0.901515
6	156	1453	12	25	0.86	0.93	0.98	0.893983
7	173	1431	31	11	0.94	0.85	0.98	0.891753

表 3、A 組各階段之統計數據

### 4.3 實驗數據

在 10 位受測者中，本研究收集了共 10122 筆資料，每一筆資料皆代表當下一偵(Frame)時雙腳感測器收集到的 16\*2 筆資料，其中 AccX、AccY、AccZ 代表三軸加速度且測量範圍為 $\pm 8G$ ，GyrX、GyrY、GyrZ 為三軸之陀螺儀腳速度，測量範圍為 $\pm 2000^\circ/s$ 。MagX、MagY、MagZ 為三軸方向地磁感測器，測量範圍為 $\pm 800mG$ ，Roll、Pitch、Yaw 為三軸航向角，範圍為 $\pm 180^\circ$ 。Qw、Qx、Qy、Qz 為節點的四元數。每個感測器 16 筆資料，左右腳兩個感測器每偵共 32 筆。

而模型輸出目標 Phase 為我們人工標示共 1~7 七種階段[2]，分別為：

- 1.Loading Response(首次觸地)：**腳跟首次接觸地面，落實腳在支撐相的位置和動作，降低下肢前向運動的速度。
- 2.Early Stance(承重反應)：**重心從腳跟轉向全足的過程，骨盆運動在此階段逐漸穩定。
- 3.Mid Stance：**支撐腳全部著地，另一隻腳處於擺動相，是唯一單腳支撐全部重力的階段，控制脛骨前向慣性運動，為下肢向前推進做準備。

**4.Terminal Stance：**將身體重心轉移至對側下肢，又稱為擺盪前期。

**5.Initial Swing：**腳底離開地面，屈髖帶動屈膝，加速肢體向前擺動。

**6.Mid Swing：**雙腳對立階段，腳仍然是離開地面的。

**7.Terminal Swing：**下肢前向運動減速，腳跟準備著地的姿勢。

我們將同樣的數據分成兩組，其中 A 組(表 2、表 3、表 4)為全部資料取 75%作為資料訓練集、25%為資料測試集。B 組(表 5、表 6、表 7)為取其中 5 男 3 女做為資料的訓練集，另外 1 男 1 女做為資料的測試集，目的是希望能歸納出一個通用的模型能從不同母體當中，延伸精準推測出個別個體的步態階段。

A 組顯示了運用本研究之模型後，能有效的在有限的資料當中去精準的切出各個步態階段，其中模型給出的準確率為 90.9%。然而在觀察 Confusion Matrix(表 2)後，可以發現錯誤的預測大多是發生在該階段的前一階段，或者後一階段，與原本之階段相距甚近，如此一來，而在觀察整體每

階段預測時間平均(表 4)時，與實際值最多也只相差總體的 1%。

Phase	Real%	Predict%	Gap%
1	12%	13%	0.5%
2	22%	22%	0.3%
3	11%	10%	1.1%
4	16%	16%	0.0%
5	16%	16%	0.4%
6	11%	10%	0.7%
7	11%	12%	1.2%

表 4、A 組所預測各階段時間與實際時間占百分比差異

B 組之目的是希望該實驗不只是在實驗訓練模型時，只針對訓練資料的那幾個人有良好的預測，而是能夠把模型推廣到所有不在訓練資料裡的人群，能夠去預測各種不同人的步態資料。可以看出本研究模型在抽掉一男一女後，所訓練之模型，也能夠運用在這兩名測試集上，而模型給出的精準率為 88.9%。與 A 組同樣地可以發現在 Confusion Matrix 上(表 5)，錯誤的預測皆是發生在該階段的前一階段，或者後一階段。在觀察預測相對全體平均時間(表 7)時，在 Phase 3 與 Phase 4 有較差的表現，可以看到模型較易混淆此兩個階段。

預測值 實際值	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6	Phase 7
Phase 1	146	5	0	0	0	0	23
Phase 2	10	312	0	0	0	0	0
Phase 3	0	15	115	0	0	0	0
Phase 4	0	0	37	162	13	0	0
Phase 5	0	0	0	3	210	8	0
Phase 6	0	0	0	0	14	126	6
Phase 7	0	0	0	0	0	12	119

表 5、B 組之 Confusion Matrix

Phase	True Positive(TP)	True Negative(TN)	False Positive(FP)	False Negative(FN)	Recall	Precision	Accuracy	F1_Score
1	146	1152	15	25	0.85	0.91	0.97	0.879518072
2	312	996	20	10	0.97	0.94	0.98	0.95412844
3	115	1171	37	15	0.88	0.76	0.96	0.815602837
4	162	1123	3	50	0.76	0.99	0.96	0.859416446
5	210	1090	27	11	0.95	0.89	0.97	0.917030568
6	126	1172	20	20	0.86	0.86	0.97	0.863013699
7	119	1176	26	17	0.88	0.82	0.97	0.846975089

表 6、B 組各階段之統計數據

Phase	Real%	Predict%	Gap%
1	13%	12%	0.7%
2	24%	25%	0.7%
3	10%	11%	1.6%
4	16%	12%	3.5%
5	17%	17%	1.1%
6	11%	11%	0.0%
7	10%	11%	0.7%

表 7、B 組所預測各階段時間與實際時間占百分比差異



## 5. 結論與討論

利用雙向長短記憶神經網路(Bidirectional LSTM)，經過本研究的訓練後，能藉由僅僅兩個簡單的 IMU 來快速切割出一個人的步態階段。且不同於影像擷取設備或是地面壓力傳感器，透過穿戴式的 IMU 來收集雙腳的步態能較上述兩者更不受空間以及活動種類來限制，同時也能更易普及在各樣的臨床測量當中。

而在本次的實驗當中，因受測人員皆為較年輕且健康正常的男女，因此所收集到的步態特徵因步長、步寬、步幅等，收集的訊號可能會較年長者或行動不便者特徵相對明顯。然而在實驗 B 可以知道模型確實能夠有效辨識不在訓練資料集的人之步態，因此未來更進一步希望能提出適合各年齡層有效標準化各項訊號的方法，並且增加訓練集的樣本，藉此能夠使得模型能夠更加廣泛的應用在所有的人上。

## 誌謝

本研究經費來自於科技部 110-2634-F-008-005-、109-2221-E-008-059-MY3、交通大學人工智慧普適研究中心。

## 參考文獻

- [1] Nordin, M., & Frankel, V. H. (2012). Basic biomechanics of the musculoskeletal system. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins.
- [2] DeLisa, J. A., & United States. (1998). Gait analysis in the science of rehabilitation. Washington, D.C: Dept. of Veterans Affairs, Veterans Health Administration, Rehabilitation Research and Development Service, Scientific and Technical Publications Section.
- [3] Magee, D. J., & Manske, R. C. (2021). Orthopedic physical assessment.(7th ed.) St. Louis: Elsevier Saunders.
- [4] Herman, T., Giladi, N., Gruendlinger, L., & Hausdorff, J. M. (2007). Six weeks of intensive treadmill training improves gait and quality of life in patients with Parkinson's disease: a pilot study. Archives of physical medicine and rehabilitation, 88(9), 1154–1158.
- [5] Ebersbach, G., Edler, D., Kaufhold, O., & Wissel, J. (2008). Whole body vibration versus conventional physiotherapy to improve balance and gait in Parkinson's disease. Archives of physical medicine and rehabilitation, 89(3), 399–403. <https://doi.org/10.1016/j.apmr.2007.05.015>
- [6] Bohannon R. W. (1987). Gait performance of hemiparetic stroke patients: selected variables. Archives of physical medicine and rehabilitation, 68(11), 777–781. <https://doi.org/10.1016/j.apmr.2007.09.031>
- [7] S. D. Choudhury and T. Tjahjadi, “Gait recognition based on shape and motion analysis of silhouette contours,” Computer Vision and Image Understanding, vol. 117, no. 12, pp. 1770–1785, 2013.
- [8] Razak, A. H., Zayegh, A., Begg, R. K., & Wahab, Y. (2012). Foot plantar pressure measurement system: a review. Sensors (Basel, Switzerland), 12(7), 9884–9912. <https://doi.org/10.3390/s120709884>
- [9] Muro-de-la-Herran, A., Garcia-Zapirain, B., & Mendez-Zorrilla, A. (2014). Gait analysis methods: an overview of wearable and non-wearable systems, highlighting clinical applications. Sensors (Basel, Switzerland), 14(2), 3362–3394. <https://doi.org/10.3390/s140203362>
- [10] Su, B., Smith, C., & Gutierrez Farewik, E. (2020). Gait Phase Recognition Using Deep Convolutional Neural Network with Inertial Measurement Units. Biosensors, 10(9), 109. <https://doi.org/10.3390/bios10090109>
- [11] Djuric, M. (2008). Automatic recognition of gait phases from accelerations of leg segments. 2008 9th Symposium on Neural

Network Applications in Electrical Engineering, 121-124.

- [12] Derlatka, M., & Bogdan, M. (2015). Ensemble kNN classifiers for human gait recognition based on ground reaction forces. 2015 8th International Conference on Human System Interaction (HSI), 88-93.
- [13] Derlatka M., Ihnatouski M. (2010) Decision Tree Approach to Rules Extraction for Human Gait Analysis. In: Rutkowski L., Scherer R., Tadeusiewicz R., Zadeh L.A., Zurada J.M. (eds) Artificial Intelligence and Soft Computing. ICAISC 2010. Lecture Notes in Computer Science, vol 6113. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-13208-7\\_74](https://doi.org/10.1007/978-3-642-13208-7_74)
- [14] Martinez-Hernandez, U., & Dehghani-Sani, A. A. (2018). Adaptive Bayesian inference system for recognition of walking activities and prediction of gait events using wearable sensors. Neural networks : the official journal of the International Neural Network Society, 102, 107–119. <https://doi.org/10.1016/j.neunet.2018.02.017>
- [15] 鄭宇倫(2016)。基於穿戴式三軸加速度感測器的步態週期分割與特徵分析。國立雲林科技大學資訊工程系碩士論文，未出版。雲林縣。
- [16] Schuster, M. & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. IEEE Transactions on Signal Processing, 45, 2673--2681. doi: 10.1109/78.650093
- [17] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>