

國立台北大學資訊工程學系專題報告

基於畫鐘法之失智症評估系統

Dementia Assessment System based on CDT

專題組員:王鈺豪、蘇芊慧

專題編號:PRJ-NTPUCSIE-107-005

執行期間:107 年 09 月 至 108 年 05 月

1. 摘要

傳統評估是否罹患失智症的方式，國際通常採用”短期智能測驗(Mini-Mental State Examination；以下稱 MMSE)”及”畫鐘法測驗(Clock-Drawing Test；以下稱 CDT)”此兩種方法。此兩種方法原先皆需要經由具有失智症專長的醫生或專業人士進行評估。而 CDT 相較於 MMSE 具有以下幾個特點：快速且不失準度(>85%)的篩檢、易於上手，不需要有高度教育及文字能力需求、具良好耐受性，不因測驗時間過長而影響受測結果以及低成本，只需一張紙與筆，無須測驗費用。

本計畫”基於畫鐘法之失智症評估系統”旨在追求開發一套能夠簡易、快速、具高準度、適用於醫學領域與大眾使用的軟體，將原本需要經由專業人士判斷的難題，利用影像辨識、RCNN 範圍型卷積網路深度學習、與系統化評分方式之相關論文等技術與文獻，將原本檢測方式門檻降低，創造醫療商機，使更多潛在失智症患者受惠。本計畫針對受測者畫出之類比時鐘之影像進行分析。透過相機、掃描器等影像擷取工具將時鐘影像傳

送至電腦。首先將輸入之原始影像進行中值濾波與輪廓描繪，對影像進行複雜光線與雜訊第一次前處理，後使用邊緣增強與高斯模糊並二值化，針對第一次前處理影像進行增強，突顯時鐘線條。接著使用 YOLO(You Only Look Once)卷積神經網路將時鐘上數字進行分類與定位，並使用霍夫變換對時鐘上指針進行分析。最後利用醫学期刊之系統化評分方式對受測者之影像進行評分。

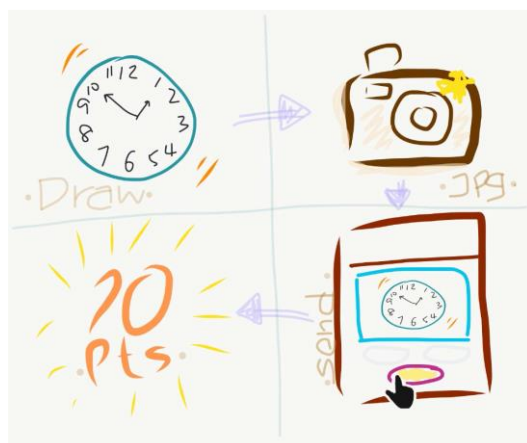
2. 簡介

現今許多國家邁入高齡化社會，隨著醫療進步，人類平均年齡也隨之越來越長。依衛生福利部(民國 100 年)委託台灣失智症協會進行之失智症流行病學調查結果，以及內政部 107 年 12 月底人口統計資料估算：台灣 65 歲以上老人約 340 萬人(全人口的 14.56%)，其中輕微認知障礙(MCI)有約 62 萬人，佔 18.23%；失智症有約 27 萬人，佔 7.86% (其中極輕度失智症有約 11 萬人)[1]。

老人相關醫療成為一個課題，而最常聽見的就是失智症，又被大眾稱為老人癡呆症。但要檢測失智症往往

都需要做一些測驗且經判斷，可能須掛號等候許久，浪費時間。若能將此測驗轉換為一套自動化軟體，並廣泛應用於醫院檢測或商用檢測，便能造福潛在患者，且提升醫療商機。

為實現此目的，所遇到的問題約略分為三種：光線與雜訊、影像辨識與判斷方式。在進行影像辨識前，必須將影像進行前處理，避免使用者拍攝時因光線、角度、雜訊等問題影響影像辨識準確度。之後經由 YOLO 卷積神經網路與霍夫變換將影像上物件進行辨識與分類。最後利用系統化判斷方式進行評分與觀察時間準確度。操作示意圖如下圖一。

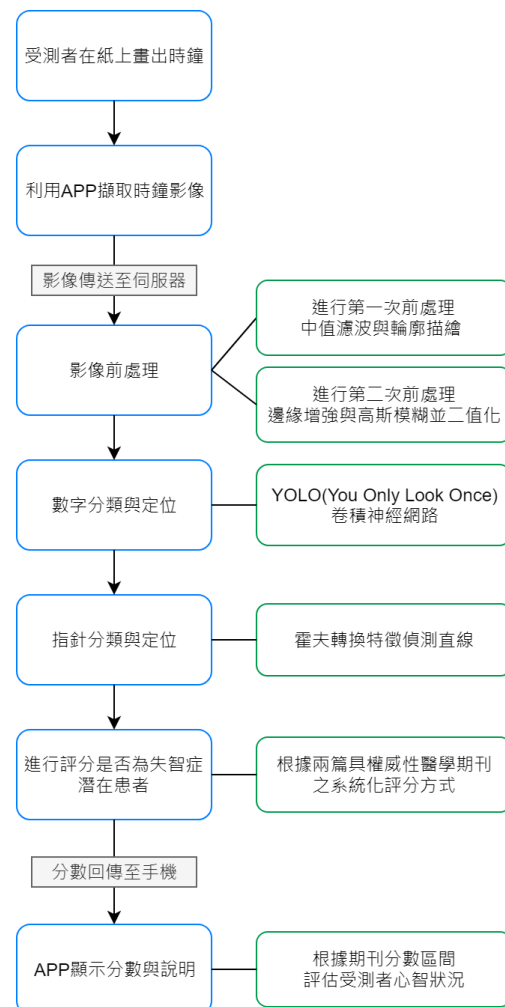


圖一、操作示意圖

3. 專題進行方式

系統之架構如下：第一步驟，透過中值濾波(Median Filter)與輪廓描繪(Contour depiction)，對影像進行第一次前處理。第二步驟，透過邊緣增強(Edge Enhance)與高斯模糊(Gaussian blur)並二值化，針對第一次前處理影像進行增強。第三步驟，使用 YOLO(You Only Look Once)卷積神經網路將時鐘上數字進行分類與定位。第四步驟，使用霍夫變換對時鐘

上指針進行分析。第五步驟，利用鐘面數字位置評分法與指針指向評分法之系統化評分方式對受測者之影像進行評分。系統架構圖如下圖二所示，詳細研究內容與步驟說明如後。



圖二、系統架構圖

3.1 影像的輸入

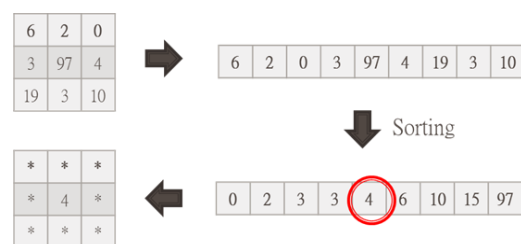
時鐘的影像經由拍照或將照片檔案輸入電腦後(如下頁圖五)，利用程式進行影像處理。

3.2 影像第一次與第二次前處理

本節之目的為將輸入影像進行兩次前處理，將雜訊及複雜光線干擾濾除，以利後續 YOLO 卷積神經網路能夠更精準辨識其內容與定位。

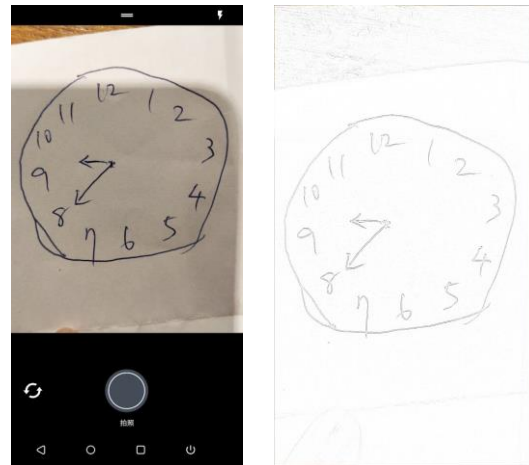
3.2.1 第一次前處理

將輸入影像執行中值濾波。在進行輪廓描繪前，通常需要先把圖像中的雜訊濾除。中值濾波是一種非線性的數字濾波技術，常用於去除圖像中的雜訊與噪點，且能保持圖像中的邊緣，以避免其他濾波器會出現模糊的狀況。以 3×3 的遮罩矩陣為例，透過空間濾波(Spatial Convolution)，將矩陣內數值由小到大排列，並取得排列後的中間值重新回填至原本的矩陣內，其示意圖如下圖四[2]。



圖四、中值濾波示意圖[2]

接著對執行完中值濾波的圖像使用輪廓偵測，輪廓偵測為在圖像中找出變化大的兩相鄰像素，並判斷其是否為輪廓並輸出。執行完中值濾波與輪廓偵測之後的預期效果如下圖六：



圖五(左)、擷取影像並輸入

圖六(右)、預期效果

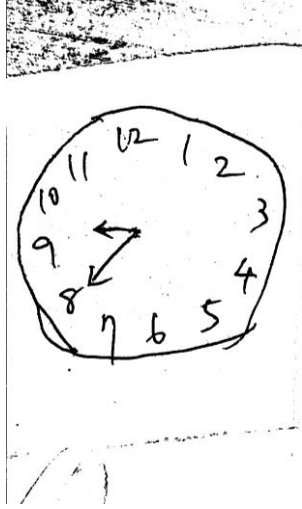
3.2.2 第二次前處理

將執行完第一次前處理之影像執行邊緣強化，使其更能加強輪廓描繪後的資訊。接著進行高斯模糊。高斯模糊是一種在圖像處理上廣泛使用的處理效果，在此研究中，高斯模糊被使用來減少邊緣強化後殘留噪點的影響。從數學角度中，高斯模糊的過程是將圖像與常態分佈做卷積，由於常態分佈又稱高斯分布，故此模糊技術就被稱為高斯模糊。在二維空間定義如下，其中 σ 是常態分布的標準偏差：[3]。

$$G(u, v) = \frac{1}{2\pi\sigma^2} e^{-(u^2+v^2)/(2\sigma^2)} \quad (1)$$

在二維空間中，此公式生成的曲面等高線是從中心開始呈常態分布的同心圓。分布不為零的像素組成的矩陣與原始圖像做卷積。每個像素值都是周圍相鄰像素值的加權平均。隨著標準偏差提高，模糊效果會逐漸增強。原始像素的值有最大的高斯分布值，所以擁有最大的權重比值，相鄰

像素隨著距離原始像素越來越遠，其權重也越來越小。這樣進行模糊處理比其它的模糊濾波器，能夠保留更多邊緣效果，故在此使用此模糊濾波器[3]。最後經第二次前處理的預期效果如下圖七所示。

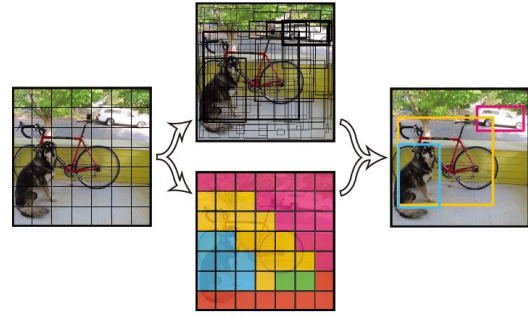


圖七、經第二次前處理後之預期效果

3.3 YOLO 卷積神經網路

YOLO(You Only Look Once)卷積神經網路架構為物件偵測的類神經網路演算法，與現今較著名的物件偵測網路架構，例如：RCNN、Fast RCNN、Faster RCNN、SSD 不同。使用 end-to-end 架構，將原本傳統物件偵測的三個流程：區域抽取、特徵取樣、影像分類合併成一個，故僅使用了一個 CNN 實現其架構，可以避免傳統物件偵測需要分開訓練的缺點，加快運算速度[4]。其判斷流程為：將影像切成 $S \times S$ 大小的格子(grid)，如果格子中有物件則該格子會負責去偵測該物件。每個格子會預測 B 個 bounding boxes 與 confidence scores。Confidence scores 代表對應 bounding box 含有物件的信心程度以及該 bounding box 中物件的精準度，

其示意圖如下圖八所示[4]：



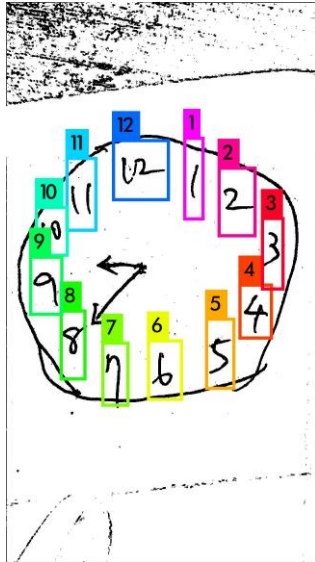
圖八、YOLO 預測示意圖[4]。

YOLO 在訓練時，主要是對於每個 grid cell 的去做 loss function 的最佳化，但只有在其 grid cell 內含有物件時才會計算其分類錯誤的 loss。其 loss function 如下所示[4]：

$$\begin{aligned} \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{\text{obj}} (x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \\ + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{\text{obj}} (\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2 \\ + \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 \\ + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2 \\ + \sum_{i=0}^{S^2} \mathbb{1}_i^{\text{obj}} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2 \end{aligned} \quad (2)$$

YOLO 第三版與第一版差別主要是增加了殘差網路與特徵金字塔網路[5]。利用殘差網路技術能使類神經網路有更敏銳的感知能力，網路層數能夠提升更高一個級數，且避免掉梯度爆炸或梯度消失的問題，訓練速度加快[5]。特徵金字塔網路則可以改善原本 YOLO 對於小物件偵測較不擅長的問題[5]。且引入 Faster RCNN 中的 anchor box，稱為 RPN(Region Proposal Network)，預先猜測數個 anchor box，而不再像 YOLO 第一版直接映射 bounding box 的座標[5]。

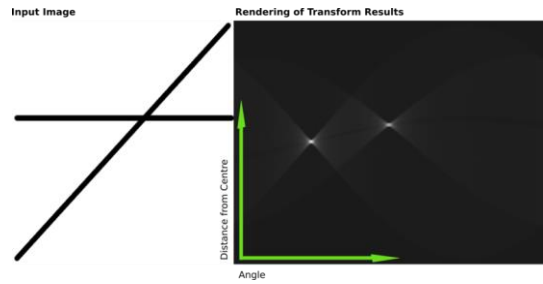
本研究辨識的數字為小物件，且須達到快速辨識，故預計使用 YOLO 第三版訓練，其預期辨識數字效果如下圖九所示：



圖九、預期辨識效果

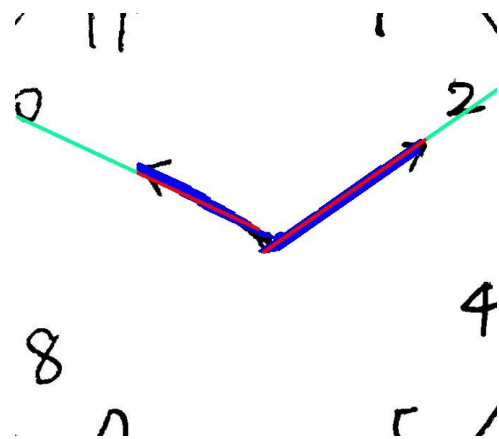
3.4 霍夫變換特徵偵測直線

霍夫變換是一種特徵檢測方式，它可以用來辨識出物件中的特徵，例如：線條、圓形、橢圓形等…，物體的形狀則是由累加空間(accumulator space)裡的局部最大值(local maximum)來決定。實作偵測直線的霍夫變換演算法流程為：建立一個二維的累加器(accumulator)，計算圖中有那些像素可以用 $r = x \cdot \cos \theta + y \cdot \sin \theta$ 表示。最後將 accumulator 中重複率高於 threshold 的直線回傳[6]。範例如下圖十所示：



圖十、直線霍夫變換。左圖為輸入圖片，共有兩條直線；右圖為 accumulator，accumulator 內兩光點表示有輸入的圖中有兩條直線[6]。

本研究將使用霍夫變換求出時鐘內兩指針，並計算出指針與分針以及兩者分別指向的方向，以利後續評分使用，其預計效果如下圖十一所示：



圖十一、霍夫變換求指針效果

3.5 基於兩篇醫學期刊之系統化評分方式

本節的目的將根據兩篇登載於美國老年醫學協會期刊上之研究，規劃出針對時鐘上之數字與時鐘上之指針的兩種系統化評分方式，最後給予分數以判斷受測者是否有罹患失智症的風險，其兩種評分方法如下頁：

3.5.1 時鐘數字之系統化評分方式

以 Watson YI、Arfken CL 與 Birge SJ 三位在第12屆美國老齡研究聯合會年度科學會議上發表，並被刊登於1993年11月美國老年醫學協會期刊上之研究[7]。其評分方法為：第一步驟，將鐘面上12與中心點連線畫出第一條直線穿越整個鐘面，再用第二條線垂直於第一條直線且穿越中心點，將整個鐘面劃分成四個象限，如下圖十二所示。第二步驟，歸納每一個數字所屬的象限，若數字位於兩條線上則順時針移動進入所屬象限，以數字12、1、2所屬象限為第一象限，順時針依此類推，如下圖十三所示[7]。



圖十二、第一步驟 圖十三、第二步驟

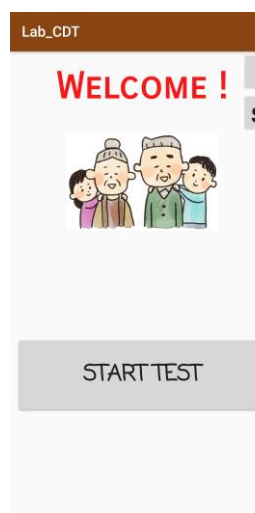
第三步驟，比對標準時鐘與受測者之時鐘的12個數字所屬的象限。前三個象限內的任一個象限內如果有出現標準時鐘該象限內沒出現的數字則扣1分，第四象限如果有出現標準時鐘該象限內沒出現的數字則扣4分，最多扣7分。正常人的扣分分數會介於0~3分，若高於3分的話有87%的敏感性與82%的特異性被判斷為失智症高風險患者[7]。

3.5.2 時鐘指針之系統化評分方式

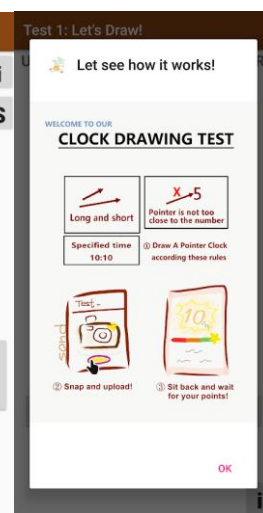
以 Sunderland T、Hill JL 與 Mellow AM 等人在美國老年醫學協會期刊發表之研究，利用畫鐘法測試及評估失智症患者的視覺空間能力[8]。其評分方法為利用期刊內的Table，判斷受測者畫出之時鐘有無符合Table之內容，其評分方式因含有數字評分成分，需與3.5.1之評分系統連接。滿分為10分，最低1分，罹患失智症的患者平均表現分數為4.9+/-2.7分，而正常人平均表現分數為8.7+/-1.1分[8]。

4. 主要成果與評估

本系統介面為APP，如圖十四。〈START TEST〉為測試進入按鈕。在點擊〈START TEST〉之後，會先顯示測試的教學，如圖十五。

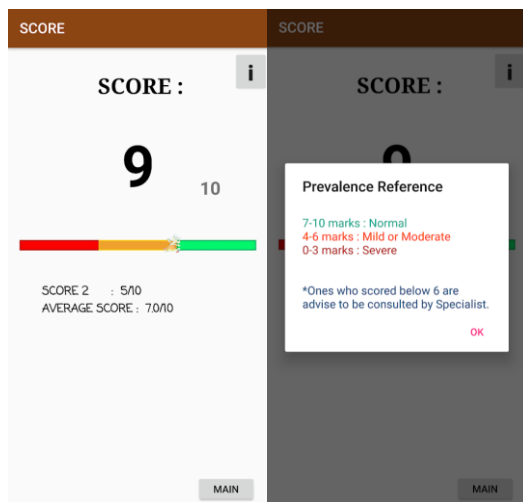


圖十四



圖十五

接著拍照並傳送至伺服器，經過伺服器計算之後會回傳分數，如圖十六，並且顯示分數評估，如圖十七，使用者可透過此評估了解自身狀況。

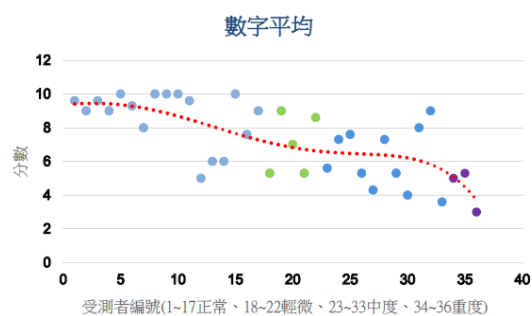


圖十六

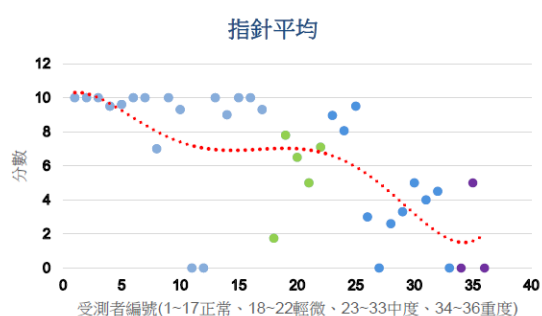
圖十七

本計畫經教授與花蓮當地老人社區之秘書合作，於 108/4/21 在兩處老人照護中心進行共三場測試活動。經統計後，共有 36 個有效樣本與 5 個無效樣本(無法動筆、離席等...)。其中 36 個有效樣本分別有正常、輕度、中度、重度的失智症患者，受測者平均年齡為 72.4 歲。每位受測者我們蒐集年齡、心智狀況與三張時鐘樣本。

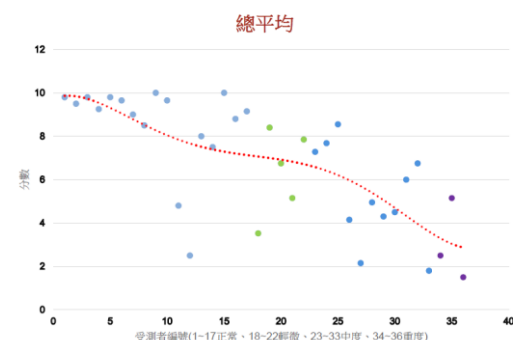
經過本系統評估所有有效樣本，統計數據之後，分別得到三張趨勢圖，如圖二十一、圖二十二與圖二十三。可以看到分數與心智狀況呈現正相關，心智狀況越差分數就越低。



圖二十一、數字平均趨勢圖



圖二十二、指針平均趨勢圖



圖二十三、總平均趨勢圖

5. 結語與展望

本計畫預期讓有失智症潛在風險者能夠透過此系統，快速且高準確度的檢查，並回報相關評分分數(指針分數、數字分數、時間正確性等...)，判斷是否有失智症的風險。

因失智症好發於中老年人，醫療院所的老人相關科目或療養院等等老人照護相關場所也能透過此系統去快速篩檢老人失智症的程度，進一步進行追蹤或更深入的檢查及治療。透過此系統，能夠創造醫療進步、醫療產業商機與對社會的相關福祉。

6. 銘謝

感謝教授與實驗室的學長姐在過程中的協助，提供資源與建議，使我們的專題的流程能夠如此順利。

感謝花蓮生命線的顏育德秘書長與當地的老人養護中心志工的協助，讓我們有機會實際測試這套系統的準確度，也體會到社會服務的價值。

最後感謝專題的夥伴，在開發過程中給予許多幫助與鼓勵，因為這份鍥而不捨的精神，才能讓我們完成這次的專題。

7. 參考文獻

- [1] 老年人口突破 14% 內政部：臺灣正式邁入高齡社會. 中華民國內政部全球資訊網. 107 年 4 月 10 號
- [2] Lung-Yu, Tsai; Image Processing – 中值濾波(Median Filter)
- [3] Wikipedia: Gaussian blur
- [4] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).
- [5] Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental

improvement. *arXiv preprint arXiv:1804.02767*.

- [6] Wikipedia: Hough transform
- [7] Watson, Y. I., Arfken, C. L., & Birge, S. J. (1993). Clock completion: an objective screening test for dementia. *Journal of the American Geriatrics Society*, 41(11), 1235-1240.
- [8] Sunderland, T., Hill, J. L., Mellow, A. M., Lawlor, B. A., Gundersheimer, J., Newhouse, P. A., & Grafman, J. H. (1989). Clock drawing in Alzheimer's disease: a novel measure of dementia severity. *Journal of the American Geriatrics Society*, 37(8), 725-729.